NAIST-IS-DD1161014

Doctoral Dissertation

Exploring Human Factors in Augmented Reality-Mediated Communication Systems

Igor de Souza Almeida

March 06, 2014

Department of Information Science
Graduate School of Information Science
Nara Institute of Science and Technology

A Doctoral Dissertation
submitted to Graduate School of Information Science,
Nara Institute of Science and Technology
in partial fulfillment of the requirements for the degree of
Doctor of ENGINEERING

Igor de Souza Almeida

Thesis Committee:

Professor Hirokazu Kato	(Supervisor)
Professor Satoshi Nakamura	(Co-supervisor)
Associate Professor Goshiro Yamamoto	(Co-supervisor)
Associate Professor Kiyoshi Kiyokawa	(Osaka University)

Exploring Human Factors in Augmented Reality-Mediated Communication Systems*

Igor de Souza Almeida

Abstract

Augmented Reality (AR) has recently grown out of being a new way to interact with virtual contents to become a way to enhance communication. This new found niche of AR, referred in this work as Augmented Reality Mediated Communication (ARMC), can be defined as any form of active communication between two or more persons that benefits from the assistance of virtual imagery displayed in their real world view. Two particularities of ARMC serve as motivation for this thesis: human factors are often overlooked in the conception of ARMC systems, and the fact that there are considerably fewer works focusing on co-located ARMC than on the remote case.

In this thesis, the applicability and effectiveness of AR technology for improving communication were assessed by exploring human factors. Two prototypes systems were developed for this purpose, the first is the remote ARMC called HANDY which proposes an AR video conferencing in which a user is able to virtually " reach out " to another ' s real world by using a two cameras setup. This system evaluated the effect of ARMC on the human factor Social Presence. A comparison between the use of the traditional video chat and HANDY was evaluated under gesture and communication oriented tasks. The second system is a new ARMC approach to co-located meeting support for small audiences. The system, named Meetsu, consists of virtual icons and text annotations (containing meeting participants' comments) displayed on a live video feed of the meeting room. It was targeted as a method to promote willingness to communicate (WTC) among meeting participants, arguably the first work to attempt it.

*Doctoral Dissertation, Department of Information Science, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-DD1161014, March 06, 2014.

The experiments with Meetsu measured the levels of WTC in two distinct groups for a period of time, before and after using the system, and compared the use of AR and Non-AR views as display method.

Keywords:

augmented reality, communication, human factors, handy, meetsu

拡張現実感を用いたコミュニケーションシステムにおけるヒューマンファクターに関する研究*

イーゴ　デ　ソウザ　アウメイダ

内容梗概

近年，拡張現実感 (Augmented Reality, 以下 AR) は，仮想物体との新しいインタラクション手法という枠組みを超え，コミュニケーション向上の手段としても発展を遂げてきた．新しく発見された AR のこの応用分野は（以下，本研究内ではこれを Augmented Reality Mediated Communication (ARMC) とする），実世界上に重畳表示される仮想物体の支援による恩恵を受ける二人もしくはそれ以上の人々の間の活発なコミュニケーションの種々の形態として定義される．ARMC に関する以下の二つ点が本研究の動機として挙げられる．一点目はヒューマンファクターが ARMC システムの概念において，しばしば見落とされがちであるということである．二点目は使用者が同じ場所に存在する ARMC に着目した研究は，遠隔地にいる使用者の ARMC のケースと比較し，驚くほど少ないという事実である．

本稿において，コミュニケーション向上のための AR 技術の適用可能性と有効性について述べる．この目的のため，二つのプロトタイプシステムを構築した．一つ目は遠隔 ARMC システム HANDY であり，二つのカメラを使用することで，その使用者が，遠隔地に仮想的に「手を差し伸べる」ことを可能とする AR ビデオ会議を提案するものである．このシステムではヒューマンファクターであるソーシャルプレゼンスにおける ARMC の効果を評価した．従来のビデオチャットと HANDY の比較はジェスチャーやコミュニケーション指向のタスクで評価した．二つ目のシステムは新しい ARMC のアプローチであり，同じ場所にいる少人数の参加者での会議のサポートを対象としている．このシステム Meetsu は会議室の実映像上に重畳表示される仮想的なアイコンやテキストの注釈（会議参加者の

*奈良先端科学技術大学院大学 情報科学研究科 専攻 博士論文, NAIST-IS-DD1161014, 2014 年 03 月 06 日.

コメントを含む）により構成される．これは会議参加者間のコミュニケーションの意欲性 (willingness to communicate，WTC) の促進手法であることを意図しており，それを試みた最初の研究である．Meetsu を用いた実験では二つのグループに一定期間システムを使用させ，その前後における WTC の水準を測定し，さらに表示方法として AR 使用時と未使用時の比較を行った．

キーワード

拡張現実感, コミュニケーション, ヒューマンファクター, handy, meetsu

to those who taught me that laughter is the best medicine:
my parents, Roberto and Nina, and
Roberto Gómez Bolaños

Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

1.1. Background and Motivation

The premise of Augmented Reality (AR) is to generate a real world view boosted by virtual imagery. This proposition initially led to the development of applications focused on guiding users through step by step maintenance [2][3] and assembly operations [4]. A few years later, TV broadcasters would embrace AR during the broadcasting of sports and news programs [5][6]. Today, AR is transitioning into people's daily lives allowing them to play with virtual pets in their living room [7] and try on clothes without having to go to a store [8].

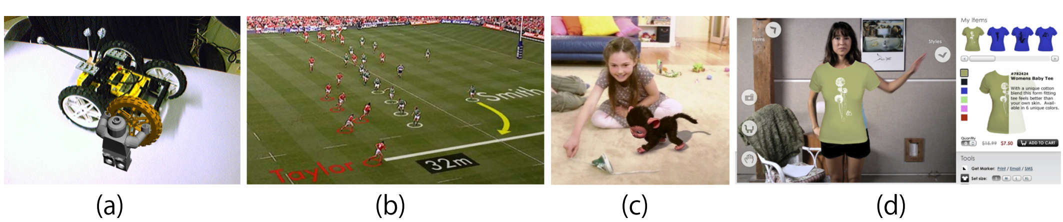

Figure 1.1. (a) Assembly task assisted by a virtual agent [4] (b) Sports broadcasting using virtual lines to provide [6] (c) Sony's EyePet [7] (d) Zugara's Webcam Social Shopper [8].

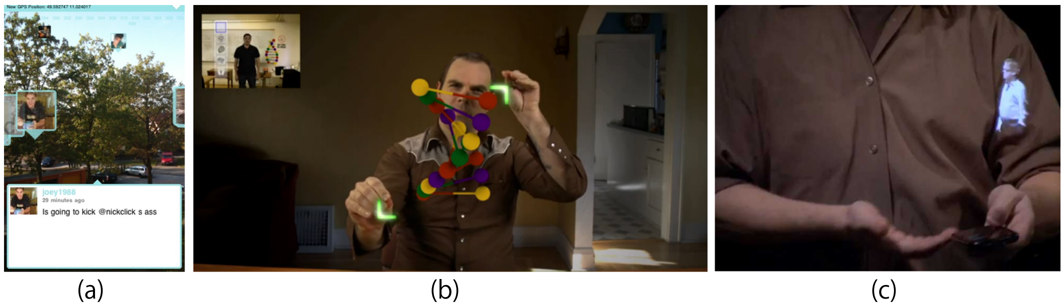

Figure 1.2. (a) Mobile app TwittARound showing location-based tweets in the real world [10] (b) Zugara's ZugStar [12] (c) Microsoft Research's Holoflector [11].

AR has recently grown out of being a new way to interact with the virtual to become a way to enhance communication with one another. The idea of augmenting a chat experience by either bringing remote users to the same space virtually [9] or displaying real-time annotations in the real world (such as *Tweets* [10]) is becoming a prospective trend for the future of communication. If the Microsoft's Holoflector [11] is envisioning the potential future of video-based AR communication using holographics and rendering graphics, Zugara's ZugStar [12] is currently achieving it through a web-based AR interactive video conferencing system which allows users to be assisted by virtual contents while communicating.

In this work, this newfound niche of AR is referred to as *AR-mediated Communication* (ARMC). ARMC can be defined as any form of active communication between two or more persons that benefits from the assistance of virtual contents being displayed in their real world. This definition is not limited by the nature of the contents, accepting any dimension (2D, 3D and 4D) and format (image, text and volumetric models), nor by the nature of the communication, including any space (remote and co-located) and time (synchronous and asynchronous).

ARMC has two particularities (challenges) which came to be the motivation of this work:

1. Human communication factors are often overlooked in the conception of ARMC systems.

2. There are few works focusing on co-located ARMC.

The first particularity indicates that AR-mediated communication requires

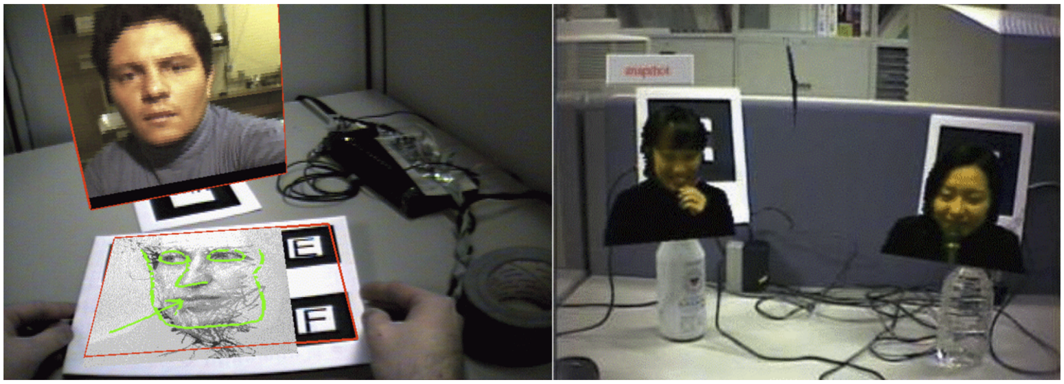

Figure 1.3. Live images of remote users are rendered on fiducial markers for a teleconference session [9].

more than just improved tracking and interaction techniques in order to achieve effective communication. The interdisciplinary study of how the technology handles and affects human factors such as *Social Presence* and *Willingness to Communicate (WTC)* must play a key role on directing the development of future AR solutions. [13] supports this idea by noting that as one focuses on achieving Azuma's proposed "Interactive in real-time" component, the study of human factors may be considered a critical component altogether. Nevertheless, the study of human factors has faced a paucity of investigations compared to the enabling technologies.

The second particularity refers to the continuous focus of ARMC works on the challenge of connecting two people remotely located. A quick Google Trends search illustrates such fact as a general trend of interest. In Figure 1.4, a chart of interest over time on the keywords *face-to-face communication* and *remote communication* shows the unequal amount of attention paid between them in recent years. Early ARMC works addressed primarily remote communication, such as the AR conferencing presented by [9] in which remote users, wearing a Head-mount display (HMD), can see their real-time virtual image rendered on a fiducial marker (see Figure 1.3). Another approach to this issue can be seen in [14], in which the capability to use fiducial marker to render virtual objects is added to the usual desktop-based videoconferencing system. With the ARMC works tackling mainly the seeming bigger challenge of remote scenarios, the co-located scenarios received little attention and works on these were up to a slow start.

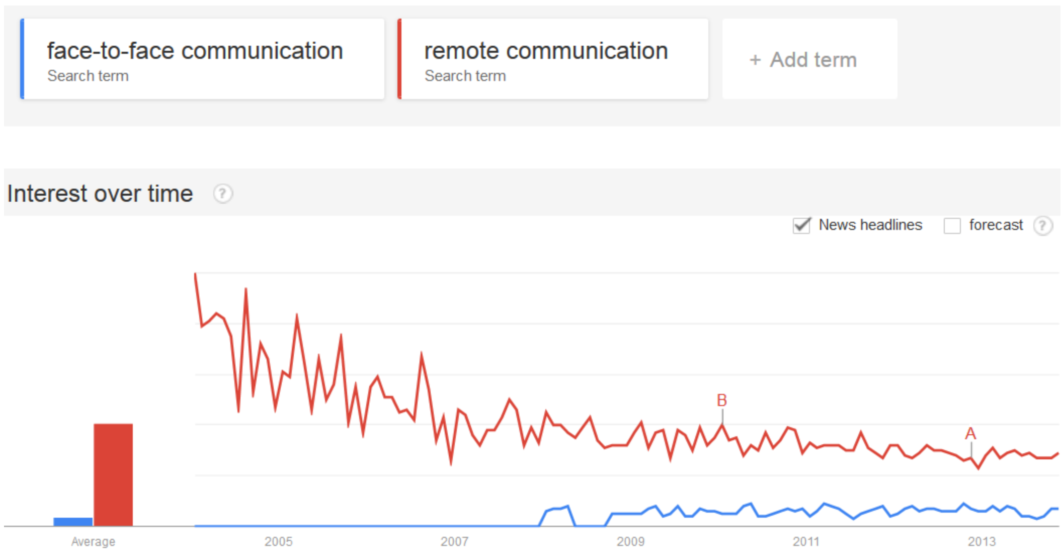

Figure 1.4. Google Trends' *Interest over time* chart indicating a higher interest rate on remote communication as a general topic than face-to-face communication over the past years *(as of December 5, 2013)*.

One reason for the lack of works on co-located ARMC is perhaps the belief that the nature of the communication in this case is ideal. In other words, the challenges faced in the remote case to provide an experience as close as possible to co-located communication is not considered as a challenge in an already co-located environment. Thus, the following question arised: *How can co-located communication benefit from AR?* Pioneer works on co-located ARMC such as [15] demonstrated that "virtual annotations enhanced understandability of the discussed topic" while the direct manipulation of virtual contents "improved insight in complex problems". This process of maximizing the understanding between communicating parties is the alternative definition of *Effective Communication* as discussed by [16][17].

In this thesis, the applicability and effectiveness of AR technology for improving communication were assessed by exploring human factors. Both scenarios of communication, remote and co-located, were investigated through the development of prototype systems and further analysis of related human factors. The remote ARMC system, called HANDY, proposes AR video chat in which a user is able to virtually coexist with a remotely located user. By using a two cameras setup, one for the user's face and another for the user's hand, the resulting image displays the hand image of one user into the other's face image. The proposed system aims at providing an increased sense of presence, thus the system evaluated the effect of ARMC on the human factor Social Presence. A series of gesture and communication oriented tasks were performed using the traditional video chat and HANDY in order to establish a significant difference between the approaches.

The co-located system is a new ARMC approach to meeting support for small audiences is presented. The AR feature of the system was developed on the web aiming at high accessibility since it could be accessed from any browser. A webcam provides the live video feed of the real world (the meeting room) on which text annotations, containing meeting participants' comments, and virtual icons

Figure 1.5. Two separate screens are used during the meeting: one contains the presentation slides (left), another displays Meetsu's AR view (right).

are rendered. The system uses a manual registration process to associate the contents with the user's location. Two screens are presented to the audience, one containing the speaker's presentation slides and another containing the AR view (Figure 1.5). The target scenario is Presentation Meeting. This meeting style consists of a speaker in front of a small audience presenting a topic of interest aided by presentation slides, followed by a Q&A session follows as the presentation finishes (details on types of meeting in section **??**). A prototype system called Meetsu was developed to provide means to evaluate our approach. Furthermore, a new human factor measure for ARMC called Willingness to Communicate was chosen to indicate the extent to which our approach affects human communication.

1.2. Research problem

Although remote and co-located communication have been successfully achieved in technical terms by a variety of AR approaches, only a handful of them have provided further investigation on their effect on human factors. The application of human factors study to technology design (cognitive ergonomics) aims at enhancing human well-being and system performance altogether. In ARMC, it is critical to allow the users to experience a comfortable and engaging blend of natural communication skills and computer interfacing. This discussion gives rise to two important questions:

- *What are the human factors which can be affected by ARMC?*.
- *What is a suitable ARMC design to affect human factors?*.
- *How can the effect of ARMC on human factors be measured?*.

1.3. Research Goal and Adopted Approach

The primary goal of this thesis is:

- *To develop ARMC systems which can support remote and co-located communication while also positively affecting human factors.*

The approach taken to achieve the research goal can be summarized as follows:

1. **Literature review**.
   Previous works and current trends in ARMC were reviewed in order to identify ways to employ AR in remote and co-located scenarios. Similarly, the review on Human Factors allowed a broader view on the complexity of the subject, which culminated in the selection of two measures (social presence and WTC) to be further studied.

2. **Remote ARMC: HANDY**.
   A prototype system named HANDY (details in Chapter 3) was developed targeting a new approach to AR video chat. A system focused on hand interaction during a video chat experience was designed to increase sense of presence. This system evaluated the effect of the ARMC on the human factor Social Presence.

3. **Co-located ARMC: Meetsu**.
   The co-located meeting support system Meetsu (details in Chapter 4) was initially designed to attend the needs of our lab's weekly research meeting. The requirements elicited from our meeting participants provided early directions for the design and development of the system. The continuous development of the system took place along with the experience during our meetings for over one year in order to increase the maturity of the system. In addition, the feedback received from the participants provided support to our assumptions which encouraged further investigation with impartial subjects. A user study was conducted to investigate if our approach was able to increase the levels of the human factor Willingness to Communicate of two distinct meeting groups.

1.4. Original Contributions

The following is a summary of the original contributions of this research:

1. **A new approach to ARMC in video chat**.

   The new approach is a combination of an original system design, achieved by means of an established technique (background subtraction). The addition of hand interaction into the traditional video chat setting, by using an extra camera and the merging of face and hand video images, allows participants to video chat while being able to partially interact with the contents of the remote video image. A formal study evaluated the effect of ”reaching out” to a remote environment on social presence.

2. **A new approach to ARMC in co-located meeting scenario**.

   Different from previous approaches, participants are not restrained by any wearable devices as no HMD or 3D glasses are used in this approach. Meetsu uses a separate screen, which contains a shared live video feed of the meeting room, to display the augmented image. Thus, it is guaranteed that all participants visualize the same contents at the same time. Moreover, this approach allows participants to maintain visual contact with all other participants. It also includes a strategic control to minimize the effect of occlusion among annotations.

3. **A formal user study addressing the WTC measure for an ARMC system**.

   Even though the use of ARMC has been demonstrated as capable of improving communication between two or more people, it is still unknown whether ARMC is capable of affecting a person's motivation to initiate communication. The user study regarding the WTC measure attempts to answer this questioning.

1.5. Structure of the Thesis

This thesis consists of seven chapters. Chapter 1 is an introductory chapter which includes Background, Motivation, Goals and Contributions of this research. Fundamental references, which open the discussion on ARMC and human communication factors, were also included.

Chapter 2 provides an overview on electronic meeting support and a description of the existing scenarios. Also included is the literature review on the AR approaches to co-located meeting support.

Chapter 3 introduces the remote ARMC system HANDY. A detailed description of the system is presented along with the results of the user study regarding Social Presence.

Chapter 4 discusses the details of the co-located AR meeting support system, Meetsu, including the system's design and development along with a breakdown of the main features. The results of two user studies (the long-term and short-term) are also described including details on the participants, methodology, procedures, statistical results, and discussion.

Chapter 5 summarizes this thesis, with final conclusions and possible future works. Subsequently, the appendices detail the material used during user studies and the relevant pieces of code that are part of the developed systems.

CHAPTER 2

Background and Related Work

2.1. Remote Communication and Handy-related works

Video mediated communication systems such as video chat systems continue to be widely spread owning to accessibility and low cost. It is one of the most popular communication channels being used between remotely located people. Traditional video-chat systems rely heavily on the two-way video/audio feedback in order to allow people to communicate. In this manner, however, natural communication cues such as gaze direction, proximity behaviour and pointing in space are not fully supported [18].

Works such as [19] have pointed out the unfulfilled expectations of this type of systems highlighting important factors, such as social presence (details can be found in section 2.3.1), that have not been fully realized by the available technologies. Picturing a situation where a person wishes to give a gift to another or do a handshake, it is easy to understand why the available technologies do not match the real world interaction.

Researchers have explored new concepts that can offer a more complete experience. The idea of Shared Space is one of them. Shared space draws similarities with Collaborative Virtual Environments (CVE) in which participants and information share a common display space [20][21]. Even though it has been proven that using a dedicated shared space environment improves the results of certain tasks [22][23], it does not assess the effect of having an additional separated environment aside from the users' real world. The HyperMirror [24] proposed a shared scene is created by capturing the front view of users in different rooms and merging the images of all users in one video image (Figure 2.1). The resulting image displays all users as if they were side-by-side in front of a mirror; users have to keep facing forward at all times making it difficult to interact with another user that is virtually by their side.

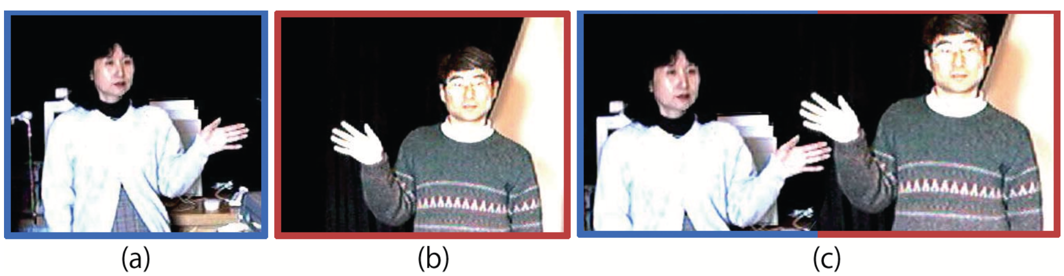

Figure 2.1. Morikawa's Hypermirror: (a) Video image of a woman captured in room A; (b) Video image of a man captured in room B; (c) Resulting video image showing both people as if they were in the same room [25].

We aim at improving upon the traditional video-chat setup, which typically consists of two windows being presented to both ends: one to display the user's own video image and another one to show the remote person. By adding one extra camera focusing on a shared space with monochromatic background, a user can place his/her hands inside this space and have them combined with the other user's face image. The simulation of coexistence stimulates different interaction patterns as reported by [26]. We advocate that this has the potential to enhance social presence.

Systems such as the GestureCam and SharedView [27] have used the superimposition of video images to allow for remote communication and virtual coexistence. In this approach, a person at a remote location wears a HMD with a shared camera. The output from the camera goes to a screen at a remote location, where a person interacts with it by pointing at the objects on the scene. This interaction is captured by another camera, and the image is transmitted back to the HMD. Now, the user wearing the HMD can see the combined image of the real object and the remote user's pointing gestures. Although this approach has been successfully applied to task-space collaboration, it does not support communication-oriented tasks which require high degree of presence. In our proposed approach, it is possible to quickly switch the focus from task-space (both sides can see the objects) to communication (both sides can see each other). Moreover, the use of HMD remains a constraint to natural communication cues.

Figure 2.2. Screenshot of the augmented reality chat application used for social presence and co-presence study in [28].

Social presence has been defined as "the salience of the partner in a mediated communication and the consequent salience of their interpersonal interactions" [29] (see section 2.3.1). However, [30] also distinctively points out that social presence has been shown to relate more to the user's perception of a medium's ability to provide the salience of another as opposed to measuring the actual perceived salience of another person. In this work, social presence is targeted as the measure to perceive how much the presence of one user's video image into another user's video image can enhance sense of "being part of" each other's environment.

A few works have attempted to measure levels of social presence in video-mediated communication such as [31]. In this work, it was hypothesized that task performance by a remote co-worker was related to the degree of social presence conveyed by the used media. Performance was evaluated in two conditions, using a two-way video setup and using an application sharing system (users can view and interact with the same software simultaneously). Findings from self-reported metrics suggested that visual feedback is not necessary to create a sense of presence.

AR has also been measured regarding its effect on social presence. [28] compared an AR chat with a traditional web conferencing system (CiTrix's *GoToMeeting*[32]) in search of significant differences in the measurement of social presence. Figure 2.2 shows a screenshot of their developed AR application. Object matching and object identification tasks were used to encourage deliberation. In the AR case, the objects were virtual and rendered on tracked fiducial markers; the traditional system was used along with *Google SketchUp*[33]. Results from self-reported surveys featuring Semantic Differential Scales (SDS) indicated no significant difference between the system.

2.2. Co-located Meeting Support and Meetsu-related works

The establishment of Computer Supported Collaborative Work (CSCW) field ignited a general interest in supporting group work. According to [34], in its most general form, CSCW "examines the possibilities and effects of technological support for humans involved in collaborative group communication and work processes". CSCW is widely recognized as interdisciplinary effort, involving knowledge of network and distributed systems, psychological and cognitive science, and human factors (HCI)[35]. The technical systems resulted from CSCW research and development are named *Groupware*.

Groupware is defined by [36] as "computer-based systems that support groups of people engaged in a common goal and that provide an interface to a shared environment". It explicitly provides awareness of the co-workers and their activities and does not separate the users from each other as it is common in distributed systems [35]. As a product of CSCW, groupware is also accomplished through different approaches: technology and computer hardware and software, and group work ans social interaction. The updated Johansen's Matrix in Figure 2.3 shows the CSCW quadrants in which groupwares can be classified into.

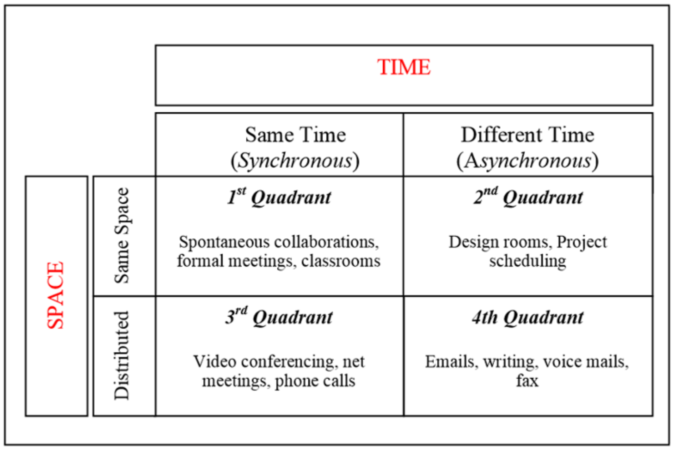

Figure 2.3. Rama's update on Johansen's Matrix describing the CSCW quadrants [37].

The 1st quadrant covers groupwares that focus on conventional co-located environments where the participants share the same space and time such as co-located meetings and classrooms. The 2nd quadrant consists of information sharing at the same place albeit at different times. In this case, the information is formed over time thus the system does not need to provide for immediate sharing. The 3rd quadrant describes cases in which the participants collaborate remotely but are connected at the same time such as video conferencing. The 4th and final quadrant covers technologies that allow for individuals to interact remotely on their demand. Communication through email for example is subject to each user's time.

A further classification of groupware has been discussed by [38] based on the main interaction modes (Figure 2.4). Since the proposed classification attempts to place groupwares according to their supported features, it allows for systems with the same purpose to have different placements as their focus evolves or branches out. For example, conferencing systems becoming increasingly cooperative as seen in [39][40] will move towards the cooperation support corner.

Among the different types of groupware, *Electronic Meeting Support* (EMS) systems represent the most fundamental approach to group support. Defined by

[41] as systems that "use information technology to support the group work that occurs in meetings", they aim at supporting one of the most basic needs of human society: to meet. The following section details the effects of EMS systems and the scenarios in which they are most commonly applied to.

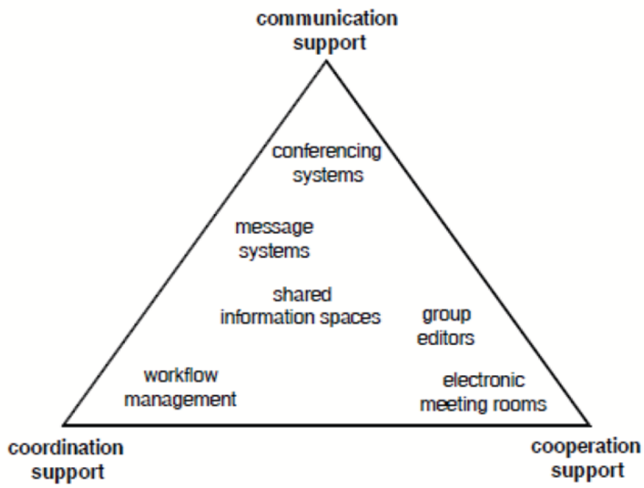

Figure 2.4. Groupwares categorized in between three different interaction modes: communication, coordination and cooperation [38].

2.2.1 Electronic Meeting Support

Organizations are influenced by its time. The changes in society or economy create new challenges for companies to deal with. As a result, new technologies emerge and new trends arise to support processes such as a group meeting. EMS systems can be described as one of these trends.

As organizations become more geographically spread and dynamic, gathering people in the same place becomes hard and normally more costly. [1] reports that companies have turned to technology as an alternative to co-located meetings in order to lower travel and meeting costs. According to the survey, the benefits of using EMS include time and cost saving, flexibility in location and timing, and increase in productivity. Further discussion on the characteristics and challenges of remote and co-located meeting scenario can be found in section **??**.

Different from groupware, there are three distinctive dimensions that can be integrated into the taxonomy of EMS environments: group size, participant location and the timing of the meeting [41] (Figure 2.5). While Time Dispersion maintains the two levels seeing in groupware, Group Proximity (equivalent to the Space in groupware) has three levels related to geographical distribution of the group: multiple individual sites correspond to all participants being physical separated working in individual offices; one group site denotes that all participants are physically in the same place; and multiple group sites consist of participants meeting in separate locations in subgroups. A third dimension is described called Group Size. Although a small group has been stated as having 10 members or less and a large group as having more than 10 members, it is a concept that may depend on the organizational culture.

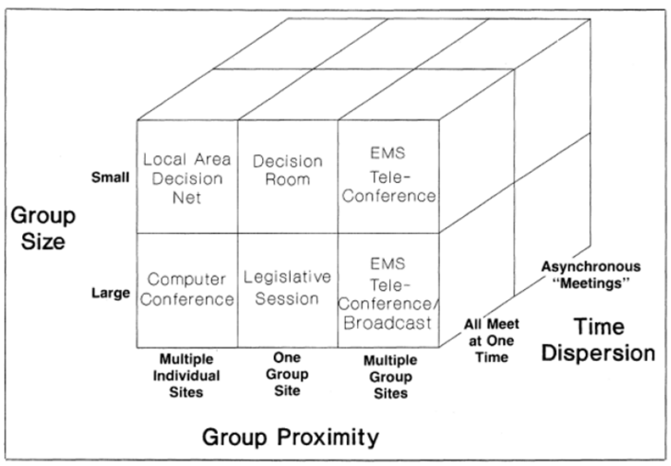

Figure 2.5. Taxonomy of EMS environments [41].

[42] states that information processing in organizations revolves around two factors: uncertainty reduction (eliminating the lack of information) and equivocality resolution (reducing ambiguity of information). It has been suggested that communication media differ in their relative abilities to reduce uncertainty and resolve equivocality. A medium strong in one could be relatively weak on the other. The differences in the relative strengths of the media provided by EMS systems could cause differences in the perceptions and performance of groups using them.

Figure 2.6 shows a variety of media with different degrees of richness (from left to right - less rich media to rich media) and personal value (from left to right - impersonal media to personal media). The uncertainty reduction and equivocality resolution levels are confronted as inverse scales. This means that when one is at its highest level, the other is at its lowest. E-mail, for example, has the highest level of uncertainty reduction ability which implies that enough information is shared at once. However it has the lowest level of equivocality resolution for being more restricted in communication cues making it difficult to resolve ambiguity.

According to [29], media that facilitates the immediate exchange of a wide range of communication cues is regarded as rich media, while those that allow the exchange of limited cues over a long period are considered lean media. Regardless of its richness, it is noted that each medium can provide a particular kind of information which could be equally valuable. It might be a better choice to make use of videoconferencing for discussion and debates, for example, or utilize emails for routine questions which demand more objective information.

In the following subsections, a more detailed description of remote and co-located meeting scenarios is presented including characteristics and challenges faced by technology. A review of previous co-located AR meeting support sys-

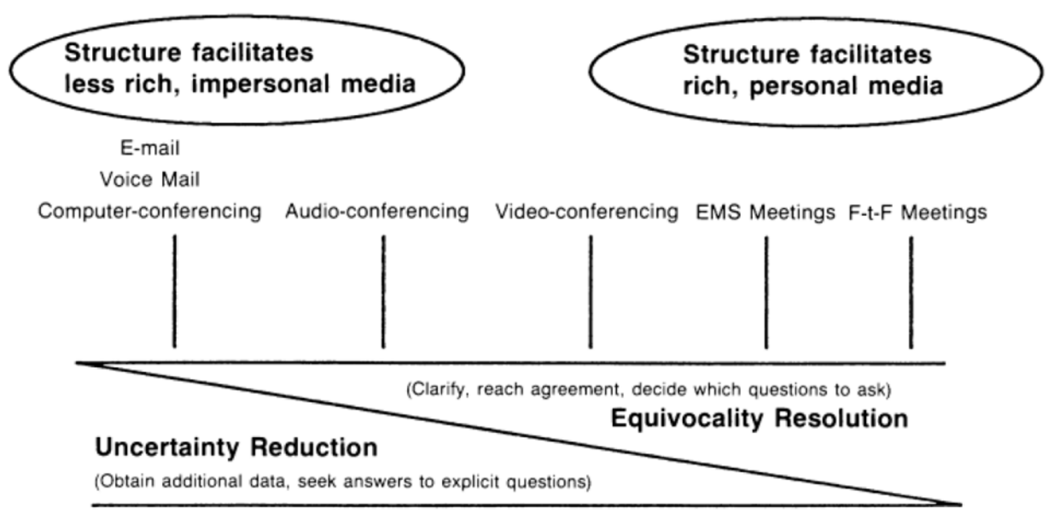

Figure 2.6. Continuum of Communication Media [42].

tems is also presented. The pros and cons of each technology are discussed and a parallel with our approach is established. Furthermore, a brief discussion regarding types of meeting further explains this work's target meeting.

2.2.1.1 Remote and Co-located Scenarios

Remote and co-located scenarios have different at times opposite sets of characteristics and challenges which have been the topic of several previous works [43][44][45][46][47][48][49]. In this section, the most relevant differences to this work are presented. Table 2.1 summarizes the main points further discussed.

A recurring point when differences between remote and co-located communication are discussed is the non-verbal cues. When two or more people are physically present, the natural communication that happens between them is aided by a number of implicit cues such as body language, hand gestures, glance, and facial expression, among others [50]. These cues help the verbal message to be conveyed in all its nuances. Even with the great strides made on the technological side, these cues are only partially supported in a remote communication scenario and with inferior degree of efficiency. The gaze, for example, that occurs in a videoconferencing environment is not as accurate, nor expressive, as the provided in a face-to-face case situation.

The non-verbal point is a recurring point due to its broad effect to other characteristics that differentiate both scenarios, including the human factor named Social Presence (see section 2.3.1). Shortly defined as the degree to which a person can perceive the presence of another in a mediated environment, the level of social presence depends directly on the medium's supported non-verbal cues. [51] reports gaze direction as a predictor of conversational attention and establishes it as an important cue for social attention. Therefore, remote communication shows lower degrees of social presence than co-located.

Multiple interactions can be achieved by both scenarios; however, the remote scenario relies strongly on the available communication channels and, in general, is less efficient. If only audio is available, it may become a challenge to support multiple interactions since all participants are usually sharing the same channel. A similar challenge is faced by videoconferencing systems if multiple users speak at the same time. Text-based systems support multiple interactions, but face the

Table 2.1. Characteristics and challenges of remote and co-located scenarios.

Remote scenario	**Co-located scenario**
Non-verbal cues are partially supported	Non-verbal cues are fully supported
Low accuracy on message delivery	High accuracy on message delivery
Social presence degree is low	Social presence degree is high
Multiple interactions can be achieved	Multiple interactions can be achieved
Less awareness of the other's environment	Full awareness of the other's envir.
Physical objects cannot be shared	Physical objects can be shared
Unplanned interactions are not supported	Unplanned interactions are supported
Less off topic discussion	More off topic discussion
Participant's id is evident	Participant's id is not always evident
More inclusive	Less inclusive

challenge of sorting and presenting all the information in a smart way.

The awareness of the other person's environment is another factor that contributes for the social interaction between the parts. Being able to "read" the environment allows for suitable choices in social interaction. An example is the choice of volume and tone of the voice depending on where the other person is speaking from and who is around that person. While the remote scenario only has a window of information in most cases, in the co-located scenario both people are sharing the same environment therefore both have full awareness of the environment. Other example is spatial awareness.

The inability of the remote scenarios to share physical objects between communicators is another characteristic that differentiates it from co-located scenarios. This limitation hampers, for example, task-space collaboration, in which [52] argues participants would often more usefully share a task space (and its physical elements), whereas remote systems, such as videoconferencing systems, usually provide individual spaces.

Remote communication is generally planned in advance which prevents unplanned interactions and informal conversations to happen naturally. These rather common interactions impact the socialization processes necessary for effective teaming, according to [46]. On the other hand, the necessary scheduling involved in remote communication has been found to increase the group's commitment to the main discussion while decreasing off topic interactions. This is an aspect of co-located communication that is very challenging to address.

The necessary identification of participants communicating through a remote system (except in cases where anonymity is allowed) gives the opportunity for communication among them to happen without hesitation. It is known that just by being in the place does not mean that people will communicate with each other [43], sometimes due to simple information not being available, such as a person's name, some people may feel more hesitant to approach others.

Another beneficial characteristic of the remote scenario, over the co-located, is scalability. Remote systems are more inclusive since a greater number of people is able to join owning to the fact that a person can be virtually anywhere and that there are no physical space restraints.

Table 2.2. Reasons for preference for technology-based systems [1].

Reason to why people prefer EMS	
Saves time	More flexibility in location and timing
Saves money	Increases productivity
Allows me to multitask	Ability to archive sessions for after viewing
Less peer pressure	Easier to follow data-heavy presentations

Table 2.3. Reasons for preference for co-located meetings [1].

Reasons to why people prefer co-located meetings
Build stronger, more meaningful business relationships
Ability to read body language and facial expressions
More social interaction, ability to bond with co-workers/clients
Allow for more complex strategic thinking
Better environment for tough, timely decision-making
Less opportunity for unnecessary distraction
Lead to higher quality decision making
Easier to focus
Meeting often result in disruption and delays

Some of the aspects described in this section can also be found in a recent survey which asked respondents (business executives) to indicate the reasons why they would prefer remote or co-located scenarios [1]. Tables 2.2 and 2.3 summarize the collected answer. They provide additional characteristics to be discussed and another perspective on the expectations that each scenario raises. Currently, the value of co-located meetings still outweighs that of the present "virtual alternatives" as investigated in [53].

2.2.1.2 Co-located AR Meeting Support Systems (CAMSS)

Despite all the benefits of EMS systems collected from the respondents of the survey in [1], 84% of them chose co-located meetings as their preferred type of meeting over technology-based meetings. Even though it is not a surprising reality on its own, it is enough evidence to highlight the importance of increasing the number of works on co-located EMS systems. In fact, the referred survey uses the term *technology-based meetings* meaning exclusively remote EMS systems. Another indication that technology-based meetings are hardly held in a co-located environment. Remote cases can be exemplified by [54] (Figure 2.7), in which telepresence and pos-meeting information analysis is provided.

Even though the co-located scenario involves a different set of problems, they tend to focus on similar issues as the remote case such as artifact sharing and broadcasting of information. Works such as [55] and [56] have few notes on the effectiveness of the communication or evaluation of human factors, two elements that are more evident and relevant when people share the same space.

In [57], it is argued that in the majority of AR systems, synchronous collaboration is supported in a co-located arrangement. However the majority of the works target task-oriented collaboration, thus task performance and success becomes the main measure to be evaluated instead of the effectiveness of the communication. Few works have attempted to attend the needs of co-located meetings using AR, for instance.

AR-based co-located meeting support has been mostly experimental solutions with little real world integration. Nevertheless, they provide some insights on the challenges faced by this new type of AR system. One example is the work in [58] which describes an approach to mediate collaboration in three dimensional AR named EMMIE (Figure 2.8). The first notorious feature is the amount of hardware and devices (including see-through head-worn displays, laptops and stylus-based display, wall-sized projection display and ceiling-mounted trackers) used for a three persons experiment. This hybrid system allowed collaboration with 3D objects in the real world while also receiving additional information on the screen. While impractical, it was a starting point for the use of AR in co-located meeting contexts.

[59] used augmented videoconferencing to enhance classroom learning (a variation of co-located meeting). In this work, physical objects are tracked using computer vision techniques which allows virtual data to be rendered on them, following the global coordinate system established by a fiducial marker. These objects are used as props by the teacher during a class. The augmented image is projected on a large screen to all students. Figure 2.9 shows the resulting augmented image presented to the students.

[60] reports an experiment called *chat-augmented conference*. This is perhaps one of the first examples of a current trend to use technology, mostly social networking services (SNS), in international conferences. In [60], side-by-side screens

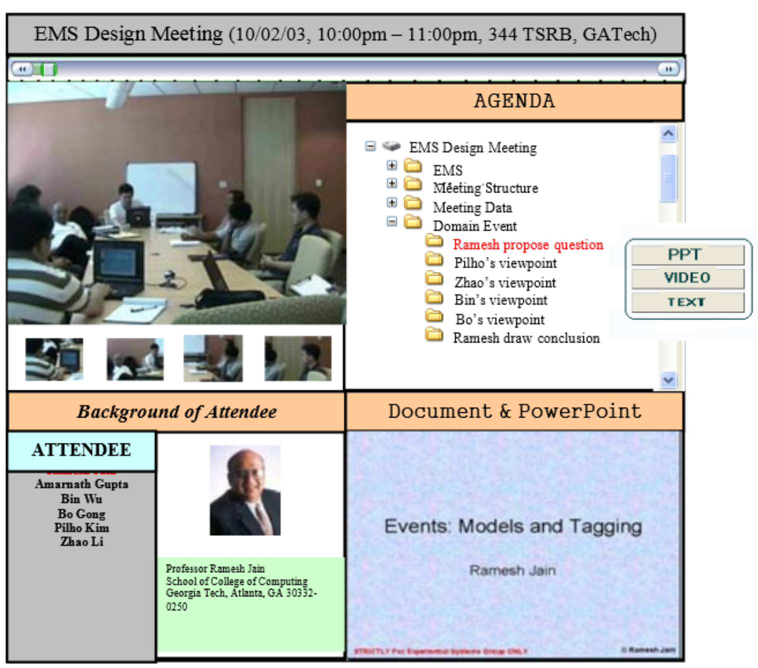

Figure 2.7. User interface in the remote EMS system by [54] which provides a multi-view live video image and information sharing.

(one showing presentation slides, another displaying a chat with input from the audience) were used to help participants engage in discussion at a conference (Figure 2.10). The chat screen was mainly used to show the conversation flow of the participants in a text-based format, which was displayed at all times to the whole conference room during the presentation and the following Q&A (Question & Answer) session. The results showed a high message generation ratio. Furthermore,

Figure 2.8. A meeting scenario using EMMIE [58].

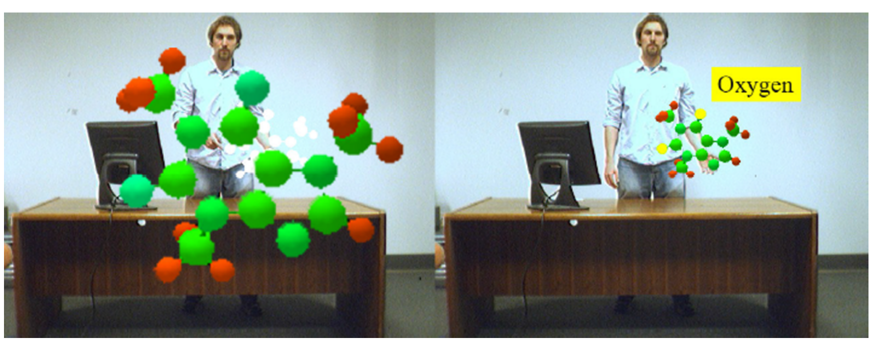

Figure 2.9. Screenshots of the resulting augmented video image projected to the students in the classroom [59].

the audience showed a strong interest in using such a system. Shortcomings of this work included the difficulty in identifying the chat speakers due to the limited number of avatars made available. Also, the attention given to the chat-screen distracted from real world discussion.

While the augmentation in [60] expanded the frontiers of the conference discussion, it gave us the idea of using AR technology in the place of text-based live chat alone. It is not hard to imagine that AR annotations onto a user's real world might help participants identify others and engage more in discussion. [61] defines AR annotations as virtual information that describes in some way, and is registered to, an existing object, and considers them to be a powerful way to give users more information about the world around them. This motivated us to explore the further use of AR annotations as a communication tool in co-located meetings.

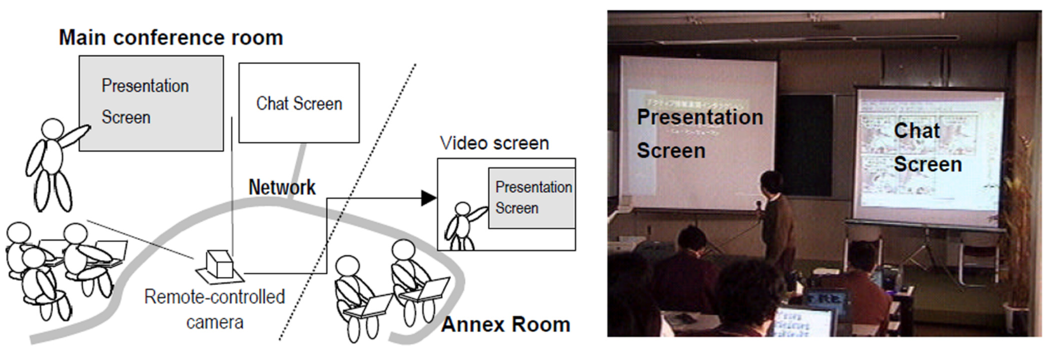

Figure 2.10. Chat-augmented conference system setup [60].

Most recently, [62] introduced a system for supporting educational presentations which enables the speaker to visualize the feedback from an audience regarding their level of understanding (Figure 2.11). Each audience member inputs feedback using mobile devices. The feedback is displayed as augmented icons to the speaker, who is equipped with a head-mounted display (HMD). A fiducial marker is used as a location reference for the positioning of the virtual icons above a user's head. Our work, on the other hand, transmits the audience's feedback to the audience themselves as well as to the speaker, much attention given to the effectiveness of the communication among participants.

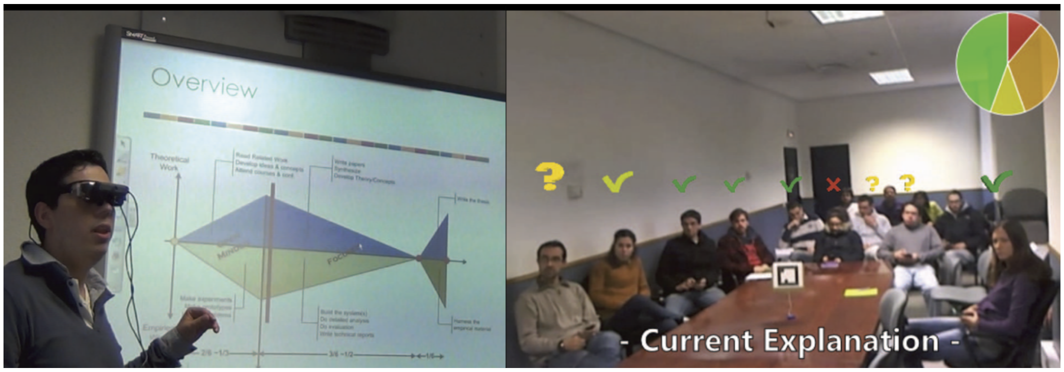

Figure 2.11. Augmented Presentation Feedback system (APFs) [62].

2.2.1.3 Types of Meeting

A meeting can be classified according to a number of criteria such as person-to-person interaction, purpose and arrangement. Thus, a meeting normally can be described by combining these criteria as presented in the classification below:

- **Criterion a: Person-to-person**
  Examples include informal meeting, and interviews. The number of people involved in the meeting process is crucial to the scheduling and management of the meeting. Knowing how many people a person has to speak too, or how many people will be sharing the attention of the speaker directly influences a person's approach to the meeting in terms of communication and behaviour. Based on a person-to-person criteria [63], meetings can have four basic configurations:

  – **One-to-One**: a single person interacting with another person.
    Examples include informal meeting, and interviews.

  – **One-to-Many**: a single person interacting with multiple persons as lecturer or speaker.
    Examples include presentation, announcement, and board meeting.

  – **Many-to-One**: multiple people interacting with a single person.
    Examples include interview by multiple interviewers, and hiring of a person by a panel of people.

  - **Many-to-Many**: multiple people interacting with another group of people.
    Examples include team presentation.

  In this work, the One-to-Many meeting criteria is the main focus since the target scenario is a presentation meeting where one speaker presents to a audience of many participants. However, in terms of communication, this work also considers the Many-to-One criteria, since multiple participants are able to "speak" to the speaker at once.

- **Criterion b: Purpose**
  Most of the literature classifies meeting by purpose within an organizational environment such as a business corporation. Therefore, the basic assumed denomination for this type of meeting is *business meeting*. A business meeting is "a type of gathering or encounter where focused interaction occurs when people agree to sustain for a time a focus of attention in a conversation or task" [53]. In this definition, the word business has the connotation of "purposeful" which can be applied to many different types of organizational environments, including universities and research labs.

  There is no strict standardization thus the concept of purpose is presented in different degrees of specification. According to [64][65], the seven common types of workplace meetings are:

  - **Regular staff meetings**: often simply a status report from a supervisor to a selected group of staff (the most common and frequent type).
  - **Project team or group meetings**: discuss the life of the team and team strategies for special projects.
  - **Cross-departmental meetings**: work out agreements regarding work handoffs between departments.
  - **Problem-solving meetings**: discuss issues that have been causing difficulty to many within the organization.
  - **Information-sharing meetings**: share updates on the current work and key informations to all members of the organization.

  - **Impromptu meetings**: focus on a single topic; share important information with many people at once or gather feedback from the participants to help informing a pending decision.
  - **Combination meetings**: combination of the types above.

  Even though the literature does not explicitly addresses research labs, the types of meeting in a lab context are very similar since these are organizational environments as well. Taking this view into account, this work tackles information-sharing meetings, which are named research meetings in a lab context. In research meetings, a student reports his/hers research progress to all lab members in a weekly basis (following a predetermined agenda of speakers).

- **Arrangement**
  Examples include interview by multiple interviewers, and hiring of a person by a panel of people. The structure and length of a meeting is addressed by the criteria of arrangement. It is important to define the appropriate roles and procedures to achieve the purpose of a meeting. The following describes some types of meeting as discussed by [66]:

  - **Standing Meeting**: a scheduled appointment at regular intervals such as weekly.
  - **Presentation**: a highly structured meeting where one or more people speak and a moderator leads the proceedings. Attendees may have an opportunity to ask questions.
  - **Conference**: a highly structured, moderated meeting, like a presentation, where various participants contribute following a fixed agenda.
  - **Emergency Meeting**: a meeting called with very little notice, but mandatory attendance.
  - **Seminar**: a structured meeting led by people with expertise in the subject matter.

It is usual to find combinations of these arrangements in order to satisfy the needs of the organization. A weekly standing meeting arranged as a presentation, for example, configures a research meeting. In this work, research meetings are the target, in which a student plays the role of speaker, another is the moderator and the remaining lab members contribute in a follow-up Q&A session.

This meeting style is much preferred when, for example, ideas have to be pitched to a board of directors or a research work has to be presented to a knowledgeable audience. Even though co-located meetings support all the natural communication cues, they tend to empower participants with a more dynamic personality or those who are higher up in the hierarchy. A co-located meeting can be an exciting experience for some and a terrifying experience to others, depending on their communication skills and their position in the organization. As a result, turn-taking in a Q&A session, for instance, might be restricted to specific participants.

Previous co-located AR meeting support systems such as [62] unintentionally contribute to this scenario by providing limited access to meeting contents, such as comments from the participants. Furthermore, the use of gadgets such as HMD's prevents users from keeping eye contact. While [67] demonstrated that this problem does not prevent task-focused collaboration to be performed successfully, the same can not be applied to collaboration scenarios that rely more heavily on social interactions and interpersonal communication skills.

Aside from the social aspect of the meeting, this troubling scenario can also be caused by a person's low predisposition toward approaching communication. This factor can be stated differently as low levels of WTC, which belongs to one of the main groups of measures in communication research (approach-avoidance) [68]. Therefore, one way to approach the problems in co-located scenario is improving the overall levels of WTC. However, a study of WTC for ARMC systems is yet to be conducted.

2.3. Human Factors

Human factors is a key term that has been at times synonymous of ergonomics. It is a research discipline "concerned with the understanding of the interactions among human and other elements that may optimize human well-being and overall system performance" [69]. The factors that influence the design, development, or evaluation of a system will vary according to the purpose and the target audience as well.

Cognition is another keyword in human factors research. It is defined as the mental process of knowing, including aspects such as awareness, perception, reasoning, and judgement. [13] suggests cognitive tasks as more indicative of how users react to AR systems. However, it is noted that cognitive measures tend to be harder to interpret.

Human factors are generally evaluated according to their importance to the context of the work. They are usually issues that can be classified into broad categories such as sensation and perception, information processing, situation awareness, information visualization, among others [70]. In AR research, [71] presents cognitive issues in visual AR (recognition, resolution of an object, navigation, among others), and audio AR (sensory information received in a passive manner), along with mobile applications and solutions involving head-worn displays.

[72] has discussed human factors in AR-based manuals, such as the effect of wearing HMDs (an ergonomic issue) and the effect of superimposing information on the real view (an information visualization issue). [73] proposed an extended use of visual, auditory, and tactile tasks as a more efficient method to identify AR human factors issues.

The following subsections detail two human factors, social presence and willingness to communicate, discussed and evaluated in the next chapters.

2.3.1 Social Presence

Social Presence has been defined as "the salience of the partner in a mediated communication and the consequent salience of their interpersonal interactions" [29]. However, [30] also distinctively points out that social presence has been shown to relate more to the user's perception of a medium's ability to provide salience of another as opposed to measuring the actual perceived salience of another person. It is a complex concept that share similarities with the other distinct concepts such as Co-Presence and Presence as differentiated by [74].

In this work, social presence is targeted as the measure to assess how much the presence of one user's video image into another user's video image can enhance the sense of "being part of" each other's environment. The social presence measure is a factor particular to remote scenarios, since co-located people share the same (ideal) degree of social presence. Similar work can be found in [31], which targets the measure of social presence in video and application sharing. Other examples include [75], which analysed social presence in online distance education according to five factors: social respect, social sharing, open mind, social identity, and intimacy.

Even though the body of work targeting social presence is considerably large, few works involve AR systems. One example can be seen in [28], which compares degrees of social presence and co-presence between a traditional videoconferencing system and an AR conferencing application, however results showed no significant differences between them. [76] compared the sense of presence during social interaction in AR and VR environments. Results indicated higher degree of spatial presence in AR systems, while engagement and naturalness factors showed no significant differences.

Perhaps the most challenging aspect of evaluating social presence, similarly to other human factors, is the evaluation method. [77] points out that researchers often define and conceptualize social presence differently, sometimes as social interaction or immediacy, others as connectedness or emotion. Similarly, the measures to perceive social presence suffer from the same problem.

A common approach is the use of self-reported metrics in which the elaboration of the questions or statements require special attention. [29] proposed early on the use of semantic differential scales (SDS), in which opposite terms (such as impersonal/personal, cold/warm, among others) figure at the endpoints of each scale (the number of points and terms can vary). This is still arguably the most common evaluation method for social presence according to [78], and it has been used in works such as [18][28]. In the first stage of this thesis, SDS were used to evaluate social presence along with additional supporting measures.

[77] developed an alternative social presence scale that measures perceived degree of social presence in computer-supported collaborative learning environments based in previous telepresence research. Though it is claimed a step in the right direction, no further studies have been carried out by other works.

2.3.2 Willingness to Communicate (WTC)

According to [68], there are three basic requirements to communicate efficiently: *being motivated to communicate*, *being knowledgeable about the situation*, and *being skilled at sending and receiving messages*. Although they are all considered to be the foundation of competent communication, the motivation to communicate is considered the first priority.

By definition, motivation can be perceived as “something that motivates” or “the state of being motivated”. In this work, the latter is the main focus. Furthermore, motivation can be divided into two types based on the factors that induce a person to the state of being motivated: intrinsic (“to do something because it is inherently interesting or enjoyable”) and extrinsic (“to do something because it leads to a separable outcome”)[79]. In general, intrinsic motivation is a complex matter and highly dependable on each person’s feeling therefore it becomes very hard to affect and measure objectively. On the other hand, we argue that extrinsic motivation must be taken into account in the conception of a meeting support system using AR.

In a meeting context, extrinsic motivation to communicate includes explicitly external factors, such as receiving a grade or praise, achieving approval from others, or aiming at self-affirmation, as well as less explicit factors, such as the pressure of being one of the few participants not actively participating in the meeting or the approachable atmosphere in which a person feels motivated by seeing others in the same position participating.

Motivation can be assessed from a negative perspective, when the person’s anxiety about communication is the target. However, it has been more successfully assessed from a positive perspective, when the person’s willingness to communicate (WTC) is the main focus. WTC is known in communication research as one of the most developed measures and it can be briefly defined as a measure of the predisposition toward either approaching or avoiding communication [80].

WTC is well-known to second-language learning research as being a fundamental goal in order to increase the likelihood of learners to use the target language naturally [81]. Figure 2.12 conceptualizes the variables that influence WTC and demonstrates that WTC represents the last instance before the actual communication behaviour.

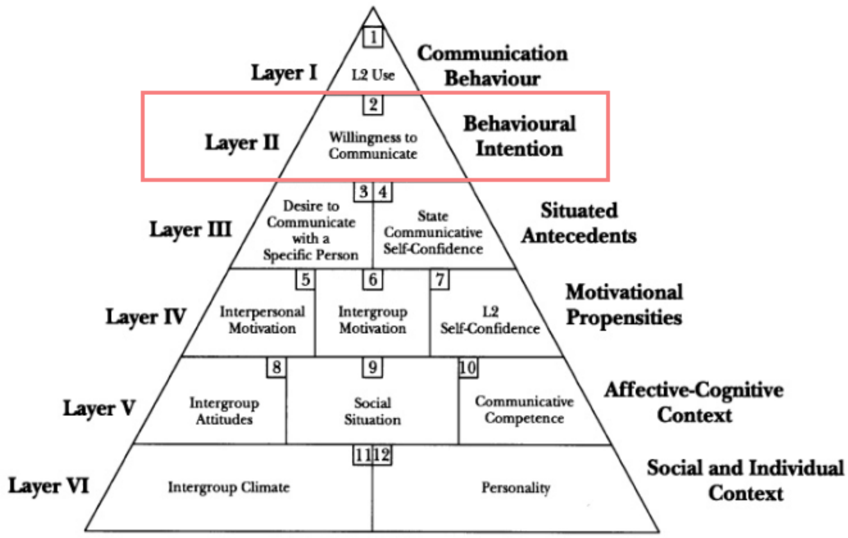

Figure 2.12. Heuristic Model of Variables Influencing Willingness to Communicate [81].

The most common approach to evaluate WTC is the use of a scale in a survey format. A person must indicate the percentage of times he/she would choose to communicate in 20 situations. Of the 20 items on the instrument, eight are used merely to distract from the meaningful items. The scores range from 0 to 100 percent, 0 representing *Never* and 100 representing *Always*.

The average of the scores will account for the overall WTC score. In addition, receiver-type and context-type sub-scores can also be calculated. Receiver-type sub-scores are divided into three categories: stranger, acquaintance and friend. The score for each category is calculated by adding the scores of their three corresponding items of the questionnaire. Similarly, the context-type sub-scores are divided into four contexts: group discussion, meetings, interpersonal speaking and public speaking.

A classification based on ranges of scores is also applied as norms for WTC scores. There are three groups into which a person can be classified: Low WTC (<52), Average WTC (>=52 and <=82) and High WTC (>82) groups. The low WTC group is considered the critical group since it consists of the people with the lowest WTC scores and therefore those who need more attention or incentive.

The sub-scores classification is different for each type of receiver and context.

In this research, special attention is paid to the context-type sub-score for the meeting context. The lowest (critical) group is also of particular interest since it is the one which might reflect more clearly the effectiveness of the system. The score range to classify the context-type sub-score for the meeting context was altered in order to incorporate the people on the verge of Low WTC scores in the molds of [14]. The following ranges were determined: Low WTC (<70), Average WTC (>=70 and <80) and High WTC (>=80).

2.4. Summary

This chapter has summarized the concepts and works related to the proposed remote and co-located ARMC systems. Moreover, the human factors featured in this work, Social Presence and WTC, were defined and exemplified in accordance with the existent literature.

CHAPTER 3

Handy System: AR-based Video Chat

This chapter details the development of HANDY (an augmented video chat system focused on remote ARMC), along with details of the system and experiments.

3.1. System Development

The main idea of HANDY is to minimize the feeling of being geographically separated by allowing one user to reach out to another. We achieved the coexistence of two users through the use of video image segmentation. Figure 3.1 shows the system setup and the resulting image generated by the augmentation process. Two cameras are part of the environment: one placed towards the face (face camera), in the same position used in conventional video-chat system; another placed beside the user (hand camera), capturing the area where the hands can be inserted. This area consists of a monochrome static background (referred to as shared space) and is represents the metaphorical crossing to each other's real environments.

The shared space requires a training phase using the first n initial frames containing the monochrome background. The training step is necessary to allow the segmentation of foreground and background pixels from an incoming frame

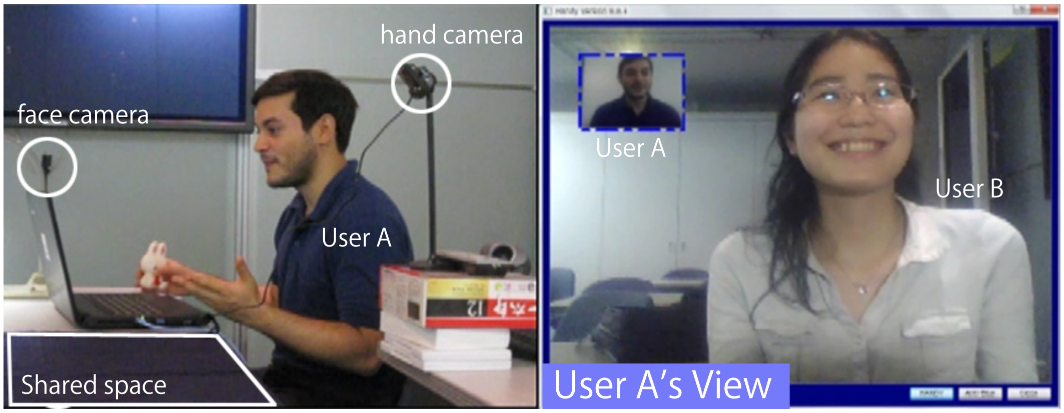

Figure 3.1. HANDY system setup.

according to a threshold value t. Images captured by the face camera along with images captured by the hand camera are combined alternately between both ends. In other words, a hand image of user A is merged with a face image of user B and vice-versa. The combined image can be seen by both sides. The OpenCV library [82] was used for the capturing and merging of images and the Qt library [83] was used for network communication architecture.

Azuma [84] has defined AR in terms of required system characteristics as follows:

- Combination of reality and virtual objects or information
- Interactive in real time
- Registered in 3D

However, [85] discusses AR as a technology that enhances the user's perception of and interaction with the real world. It comes to redefine AR as a less strict concept as "a real-time direct or indirect view of a physical real-world environment that has been enhanced/augmented by adding virtual computer-generated information to it". This definition imposes no limitation regarding the technical means to accomplish AR.

In our system we opted for the overlay of 2D images onto live video streams. The relatively simple use of background segmentation technique allowed for higher affordability and smoother integration in what is known as a traditional video chat system. Moreover, by using a simple setup, the system could be easily reproduced. Nevertheless, the use of a high quality 3D imagery and registration could potentially improve interaction capabilities.

3.1.1 User Interaction

In this section, we describe a practical scenario for the use of HANDY in comparison to the traditional video-chat system. The view shown in Figure 3.2 corresponds to User A's side.

Scenario: giving an object

This scenario corresponds to a very simple action in which User A would like to give an object to User B. In HANDY (Figure 5), user A would initially show the object to user B as it would happen in the traditional system. However, as soon as user A places the object in the shared space, it would appear on user B's environment. The effect of seeing an object that was on user A's video image being then part of user B's video image is one of the motivational factors that could impact social presence [12].

Figure 3.2. (a) Giving a gift in a traditional video chat setting (b) Giving a gift using HANDY.

3.2. Experiment

We designed a pilot user study using HANDY against the traditional video-chat system to evaluate the degree of Social Presence as well as Ease of Use, Enjoyment, Ease of Communication, Intuitiveness, Ease of understanding and Closeness. The study ' s scenario consisted mainly of remote users performing interactive tasks under equal roles, being the exception the final task where subjects played different roles (teacher/student).

3.2.1 Participants and Tasks

We had a total of 10 participants in the study comprising 8 Japanese students and 2 foreign students, 23 to 35 years old (average 25.5). They were all male graduate students from the Interactive Media Design Lab and have experienced Augmented Reality before. All participants had met before the experiment which helped creating a relaxed atmosphere. Furthermore, since they were same-sex pairs, any tensions that may arise from inter-gender interactions were disregarded. The participants performed in pairs over two days with each session taking about one hour.

For each pair, each participant was taken to a separate room and seated in front of a large display. Initially, they were asked to fill out a demographic survey which was followed by a brief walkthrough of the study by an experiment assistant. Participants were instructed on how to use HANDY. A questionnaire was given after each condition for every task. Another questionnaire was given after both conditions had been performed for them to evaluate comparatively both conditions under that given task. A final survey was given at the end of the session for the participants to evaluate comparatively both conditions considering the experience as a whole.

Three tasks were performed by each pair. The first task was a Rock-paper-scissors game. Participants were asked to play six games for each condition. Even though the wins were counted, a winner was not declared at the end. The second task was a Puzzle Matching (Figure 3.3a) where participants were given an equal number of puzzle pieces (six). Each piece had a unique match (unique connecting edges), once assembled they all completed the same image, a rabbit.

Their task was to find the match for each puzzle. One subject was given six pieces containing the rabbit's head and the other subject was given six pieces containing the rabbit's body. They used the same pieces for both conditions.

The third one was an Origami Training task (Figure 3.3b). One subject was given the role of a Teacher while the other played the role of a Student. The Teacher was given a print out containing the instructions on how to fold an origami. Two different origamis with distinct degrees of difficulty were chosen: a piano (easy level) and a balloon (medium level). For each pair, the order of the selected origami for each condition was changed. For this task, the cameras were targeting the hands on the table from a top view instead of the face. In addition, the only video image being displayed was the one with the merging.

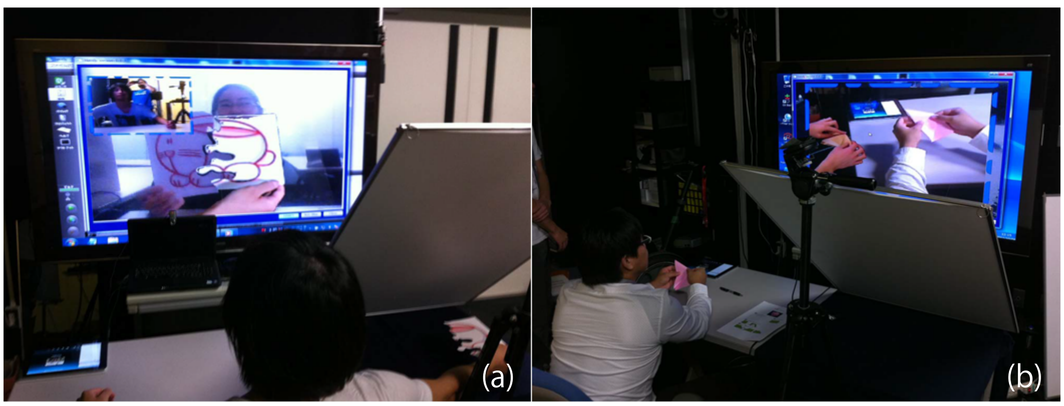

Figure 3.3. Pictures taken during the experiments: (a) Puzzle matching (b) Origami training.

Subjects were asked to perform under two conditions:

- **HANDY Off**: this condition simulated the traditional video-chat setup where real-time audio and video feed are available to the subjects only.
- **HANDY On**: this condition assumed the same setup as HANDY Off, except for the added augmented image previously described in section 3.1.

During condition HANDY On, subjects were asked to use HANDY alternately to have both of them experiencing the system equally. For Task 3, however, only the user who took the Student role used HANDY actively, placing the hand in front of the second webcam.

3.2.2 Methodology

We used a within-subjects design with a single independent variable (HANDY On & HANDY Off). The order of the tasks was the same throughout the experiments. Observation notes were taken while watching participants' performance in each task to record different interaction and behaviour patterns.

7-point Likert scales, 1 = strongly disagree, 7 = strongly agree, were used in the exploratory questionnaire items to evaluate the overall performance of the system according to six measures: ease of use, enjoyment, ease of communication, intuitiveness, ease of understanding and closeness. SDS were used to measure social presence with focus on the medium as defined by [29]. In this experiment, we used 7-point scales with a total of nine items. The items and their respective translations used in the Japanese surveys were as follows:

1. Impersonal (人間味がない) - Personal (人間味がある)
2. Cold (冷たい) - Warm (温かい)
3. Ugly (醜い) - Beautiful (美しい)
4. Small (小さい) - Large (大きい)
5. Insensitive (繊細でない) - Sensitive (繊細な)
6. Colourless (華やかでない) - Colourful (華やかな)
7. Unsociable (社交的でない) - Sociable (社交的な)
8. Closed (閉鎖的な) - Open (開放的な)
9. Passive (受動的な) - Active (能動的な)

Our hypothesis targeting Social Presence can be described as follows:

- H_0: HANDY On and HANDY Off generate a similar degree of measured social presence.
- H_a: There is a significant difference in the degree of social presence generated by HANDY On and HANDY Off.

3.2.3 Results

The first set of results describe the system's performance and overall impression by the users according to the six measures assessed through the Liker scales. A two factor (task, Handy On/Off) ANOVA test was performed on the results.

We found significant differences between conditions for Enjoyment ($F(2,54)=4.019$, $P<0.05$), Ease of Communication ($F(2,54)=4.019$, $P<0.05$) and Intuitiveness ($F(2,54)=4.019$, $P<0.05$). The average scores for all measures are shown in Figure 3.4 showing the comparison between mean scores of HANDY Off and HANDY On conditions.

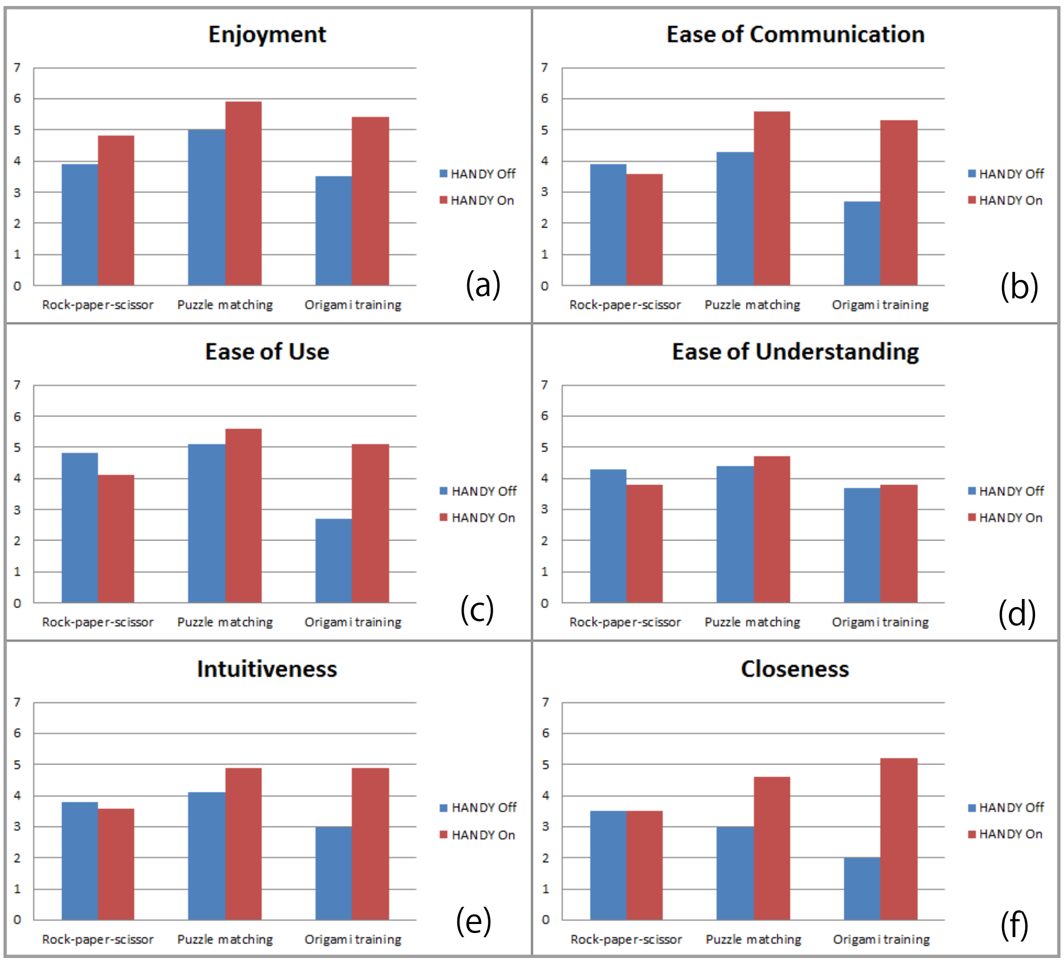

Figure 3.4. Six charts reporting the Likert Scale average scores for each of the six items during each task.

For the puzzle matching and origami training tasks, users thought the HANDY On condition was more fun as well as it made the communication easier to be performed and more natural. These are aspects that could signal a possible higher degree of Social Presence under the given condition. For the rock-paper-scissor task however even though users thought the HANDY On condition was more fun, it was not more natural or easier to communicate with than HANDY Off condition.

The user's feedback reflected the results obtained from the questionnaires: most of the participants agree that when using HANDY system (HANDY On condition), the experience becomes more fun and it improves communication abilities. Other comments however referred to non-ideal settings such as the position of the second camera. Ideally, this second camera should be placed behind the user's back and in a height close to the user's viewpoint. However, due to limitations when using the background subtraction (training stage), the camera was placed beside the user. Eye contact and some video latency were also reported as technical points that could affect the communication if the task relies on any of those factors. For example, the rock-paper-scissor task became slightly troublesome since the users had to adapt the timing of their hands due to latency.

The second set of results addresses the perceived degree of social presence in both conditions as reported through the SDS. A two factor (task, Handy On/Off) ANOVA test was performed on the results. Despite the small number of samples, significant differences were found between conditions HANDY On (Mean=4.5481, StDev=1.2707, $p<0.05$) and HANDY Off (Mean=3.5925, StDev=1.1423, $p<0.05$) across all nine factors. Figure 3.5 shows the means of social presence for each condition over the three tasks. The average of social presence was consistently rated higher during condition HANDY On.

Post-hoc comparisons showed that social presence was significantly higher in HANDY On condition, making the current results enough to at least indicate HANDY On's tendency towards having a higher degree of social presence by statistically rejecting the null hypothesis.

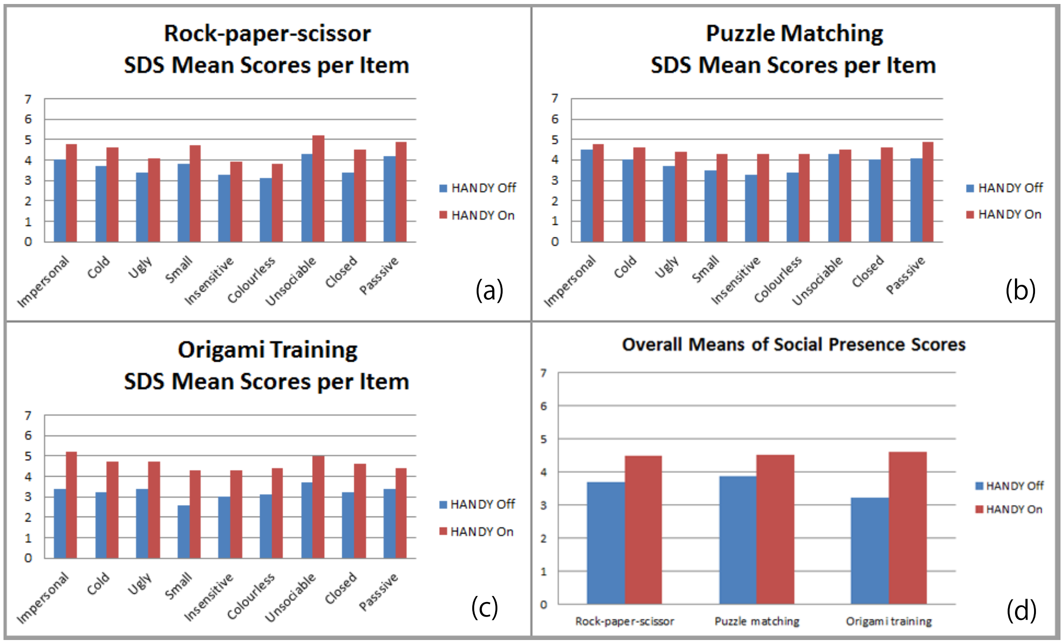

Figure 3.5. Four charts report the SDS scores: (a)(b)(c) present the SDS mean scores per item during each task (d) summarizes the overall SDS mean scores (sum of all SDS items) during each task.

3.2.4 Discussion

Subjects had mixed feelings towards Task 1 (Rock-paper-scissors) and Task 2 (Puzzle Matching) tasks in contrast with the generally positive feedback given for the experience in Task 3 (Origami training). This may be due to the nature of the interaction involved in the given tasks.

Task 1 had a predetermined set of gestures which represent the options in the game. The hands were held in the air while showing the shape of the hand. During the experiment, subjects were often trying to reposition their hands due to hands occlusion. Different from the face-to-face game, there is only one angle of view; therefore users need to rearrange their spaces to play the game successfully. Trying to find the right position in the 2D displayed image while moving the hand in the 3D real world was not an easy task.

Task 2 generated more active and open gestures. Different strategies arisen in order to perform the match more efficiently such as dropping the piece inside the shared space to make it static and describing the contour of the piece to the other user since trying to actively match the pieces brought up difficulties.

Task 3 included a sheet with instructions on how to fold the origami but it was up to the subject to decide how to pass that information without showing them explicitly. In preparation for the task, the users naturally adjusted their hand position. In contrast to Task 1 and 2, it was like preparing your space on a table, similar to a face-to-face situation. In this case, occlusion was not frequent. The complete focus on the hands for the entire time of Task 3 might have helped in the participants' higher involvement in the task, but it also provided less communication cues since no face image was being displayed.

The hand eye coordination in HANDY may not be considered optimal for tasks 1 and 2 due to their large dependence on the use of 3D space (hands are moving above the table). This perception however changed for task 3 where the participants interacted mostly on the table (2D). Perhaps, the gesture-oriented tasks required a more precise coordination between hand and eye than the communication-oriented task. Further analysis could be done by having participants performing a communication-oriented task which requires the use of 3D space.

The higher degree of social presence reported in all tasks indicates that the users perceived their partners more efficiently however it does not guarantee one feels he/she is "being perceived" simultaneously. In this case, a co-presence study would be needed in order to clarify this possibility.

The simultaneous use of the shared space by both users has not being investigated. This case is likely to cause higher degree of distraction and loss of information since it requires a user to share focus between windows (both would show distinct augmented images).

3.2.4.1 On Social Presence

Further investigation is to target additional factors that could influence the social presence degree of HANDY such as gestures. In particular, the different types of gestures the user is successfully making use of for communication and/or tasks purposes as well as to what degree the system enables those gestures to be correctly expressed.

The results achieved with HANDY indicate that the investigation of ARMC as potential technologies for enhancing human factors is relevant and current. As technologies become more social, it becomes clear the importance of this type of interdisciplinary investigation.

3.3. Summary

This chapter has summarized an early experimental research which resulted on the development of a remote ARMC system, named HANDY, an AR approach to video-chat communication which served as for a study on social presence.

The results of the study can be summarized as:

- A significant difference was found between HANDY and the traditional video chat in degrees of enjoyment, ease of communication and intuitiveness. Average scores between both systems and additional reported feedback indicate higher preference for HANDY.
- A significant difference was found between HANDY and the traditional video chat in degrees of social presence. Average scores between both systems and additional reported feedback indicate higher degree of social presence was perceived using HANDY.

CHAPTER 4

Meetsu: AR-based Co-located Electronic Meeting Support

This chapter describes our proposed approach to co-located ARMC in a meeting context. The system, Meetsu, employs AR annotations overlaid on live video stream as means to enhance communication. The augmented video image is displayed on a large screen which remains visible to all participants at all times during the meeting. The background, motivation and related work presented in Chapter 1 and 2 provide the foundation for the system design and development discussed in this chapter. Moreover, two experiments are described: a long-term (involving our lab's meeting group) and a short-term (involving two outside groups).

4.1. Approach

The question raised earlier in this paper regarded the best way to improve communication in co-located meetings. Communication research showed us that motivation is one of the key components of efficient communication. Therefore, our system's approach proposes a method to target motivation by applying AR.

The extrinsic motivation factors described in section 5.1 contain important elements in order to be achieved. A student interested in receiving a grade or praise needs ways to have himself and his contribution (question) noticed by the professors; a student that feels pressured to participate when seeing others actively participating can easily understand this setting when showed who and how much others are participating; a person that builds confidence to participate by seeing his peers participating benefits from clearly understanding who and to what extent they are participating. Thus, perceiving who are participating and how much they are participating are key elements to perceiving motivation factors in a meeting scenario. As a result, it is reasonable to assume that the easier it is for the participants to perceive these elements, the higher the chance they can get motivated.

We advocate that AR makes the task of identifying who asked a question and how much contribution has been put out more intuitive to be perceived. By showing a comment or icon above the real world image of a participant, the key elements of motivation become instantly recognizable.

4.2. System Design

In the first development phase of Meetsu, professors and students were interviewed in order to identify possible problems in a presentation meeting scenario that could be overcome using AR. Most of the respondents considered the role of the audience members. The Professors wanted a system that would raise the students' interest in participating with questions and comments, while the students mentioned the long meeting hours (3 to 4 hours on average) and their reluctance to ask questions (some due to shyness; others concerned about interrupting the flow of the presentation).

The answers also addressed the roles of chairman and speaker. It was considered tiresome for the chairman to manage the turn-taking during the Q&A session which involved constant inquiry for questions or comments. Moreover, increasing feedback from the audience was considered desirable for the speakers.

While the professors' comments were taken as the main criteria for the development of the system, the students' remarks raised an interesting discussion about the Q&A phenomenon in a typical meeting scenario. Every participant has different levels of WTC in a meeting, with some people asking more questions than others. It is believed that if the WTC levels of the participants were improved, this would increase the likelihood of them asking questions and making comments, which would result in a more participative meeting.

In light of the information gathered, we have designed a web-based meeting support system to motivate them to communicate in a meeting. In our proposed system, Meetsu, the audience is allowed to submit questions at any time in the meeting. These questions are displayed to all participants on a large screen (Meetsu screen) which acts as a virtual mirror in which all participants can be seen. This system configuration was deemed more suitable for use in the meeting scenario than other approaches to AR, such as the use of HMDs. Requiring each participant to wear a HMD would have been unaffordable, and problematic for noticing new questions from participants who are out of sight. Other discarded approaches included the sole use of smartphones, which was deemed to affect meeting manners since it would require targeting of the device towards other participants.

The use of a top-front view of the audience along with comments placed above the participants' heads was thought to recall the speech balloon convention often present in comic books. This type of representation, which allows text and icon annotations to be understood as a person's speech or thought, can not be achieved by a top view of audience.

We supposed that by displaying the questions and comments from the participants through the system, we could encourage the audience to communicate more actively since there would be no need to wait until the Q&A session to ask a question. This procedure could also mitigate a reluctance to formulate questions and, through writing instead of speaking, help shy persons participate. We also developed a set of interactive icons which could be used to express pre-defined ideas such as "I have a question". This feature is designed for participants who may want to ask a question without having it completely formulated. This would help the chairman during the turn-taking process by seeing in advance who has a question.

Each of these features can help with the management of a meeting and the creation of a more dynamic meeting environment that might shorten the overall meeting time. All the features are described in more detail in the following section.

4.3. System Development

4.3.1 Overview of the System

Meetsu is a web-based AR system developed as a combination of PHP and JQuery libraries. PostgreSQL is used as a database management system to record actions performed by the audience through the Web UI such as submitted questions and clicked icons. The combination of this information with the live image captured by the webcam creates the web-based AR view (see section 4.3.3) which is displayed on the Meetsu screen. Figure 4.1 shows the overview of the system.

A PointGrey Flea2 camera is positioned above the presentation screen, overlooking the audience, capturing a live image of the room. All participants face a wide white canvas onto which two computer screens are projected. One displays the presenter's slides and the other contains the live camera feed of the room on top of which Meetsu's AR feature is displayed.

Participants access the system through a web browser using Notebooks and iPads. They have to create and access their personal accounts in Meetsu before they can virtually attend a session. When attending a session, users are required to perform a manual registration step by locating themselves in the live video feed using a mouse click or finger tap, depending on the device.

In this system, all displayed annotations are two-dimensional images, and the registration required for the positioning of the annotations onto the real-world view is done manually. The resulting image is a real-world view enhanced by virtual information, which falls within the concept of AR proposed by [85].

4.3.2 Web-based User Interface

The system contains an online version of the presentation slide, which is uploaded by the speaker before the meeting. The digital file is made available within the web interface, from which users can follow the presentation (Figure 4.2a). An optional automatic mode allows the online slides to be changed in sync with the slides being presented (Figure 4.2b). If the user is browsing through the slides,

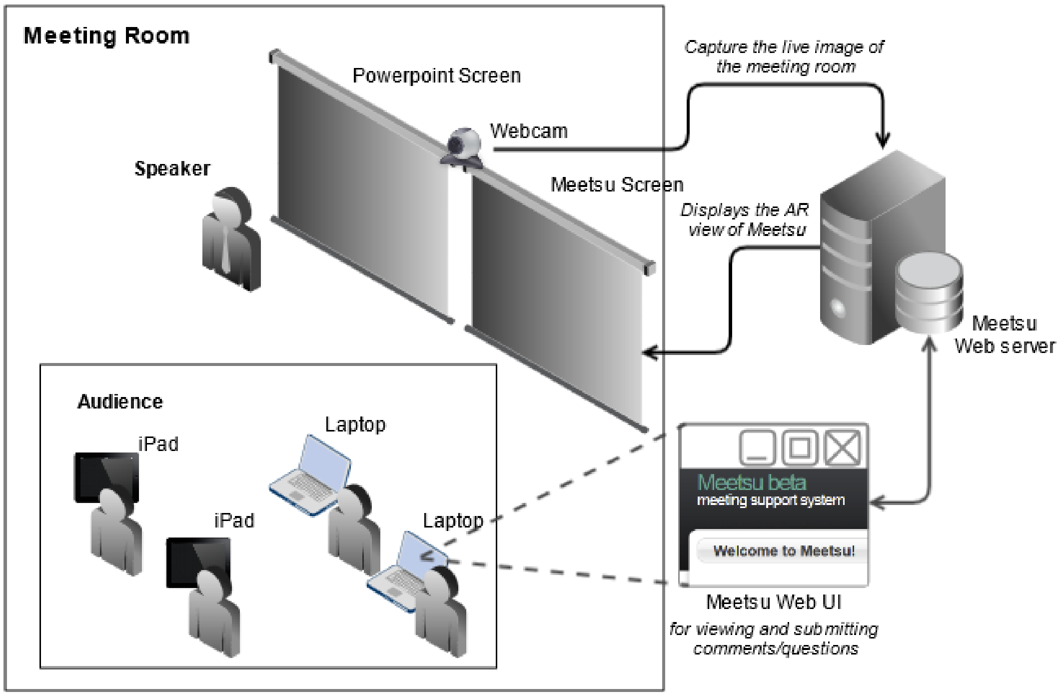

Figure 4.1. Overview of Meetsu system.

this mode allows the user to change to the slide currently being presented. This eliminates the need for printed copies and makes it easier for the audience to navigate through the slides.

A remote presentation control was also implemented. This allows the participants to address their question to a specific slide easily, without having the speaker to search for the requested slide (Figure 4.2c). Users can submit questions and comments through a web submission form by typing in the text field and clicking the submit button (Figure 4.2d). Once submitted, a question is automatically included in the question list, which is updated in real-time and made available to all attendees (Figure 4.2e). At the same time, an AR annotation is displayed on the video at the user's location (acquired from the manual registration step). In regards to the question list interface, users have two buttons available for each comment.

The first button is the Good! button (Figure 4.2f). It is similar in function to the Like button in Facebook, being used by participants to show their agreement

Figure 4.2. Available features in the web-based UI of Meetsu.

with a particularly interesting remark. The second button is the Highlight button (Figure 4.2g). This button visually highlights the AR annotation (by changing the background color) of the corresponding comment. Also displayed in the same AR fashion as the submitted questions and comments, the interactive icons are used to express a limited range of expressions such as “I have a question”, “I'm confused”, “I agree”, “I disagree” (Figure 4.2h, from top to bottom). The icons were designed to allow introverted participants to interact readily. Lastly, participants were encouraged to use a bug report feature (Figure 4.2i) to record problems or technical difficulties while using the system.

4.3.3 AR View

The AR view displays virtual information such as text annotations and icons onto the users in the real meeting room environment (Figure 4.3). This view provides an easy understanding of the users' identity and location in the real world alongside their respective entry. Therefore, in a meeting, it becomes easy to understand who asked a question and where in the room this person is located. In addition, the meeting participants and the speaker can grasp the overall impression of the audience while being able to track which topics are being discussed. The speaker also benefits from the AR view by being able to locate the focus of questions in the meeting room which allows for a change of gaze and attention to a specific region in the audience. In addition, the speaker is able to adapt his explanation depending on the question owner. If the speaker know the question owner is a knowledgeable person, the speech can be more technical, otherwise it can be simplified for easy understanding.

The priority of the AR view is to display the latest question of each participant in the foreground, while the other questions stay in the background. Currently, there is no strategy as to handle the timeline of questions in the AR view, however the list of questions in web UI provide the timeline. Also, the highlight button managed to get around the occlusion problem by allowing any question to be brought to the foreground.

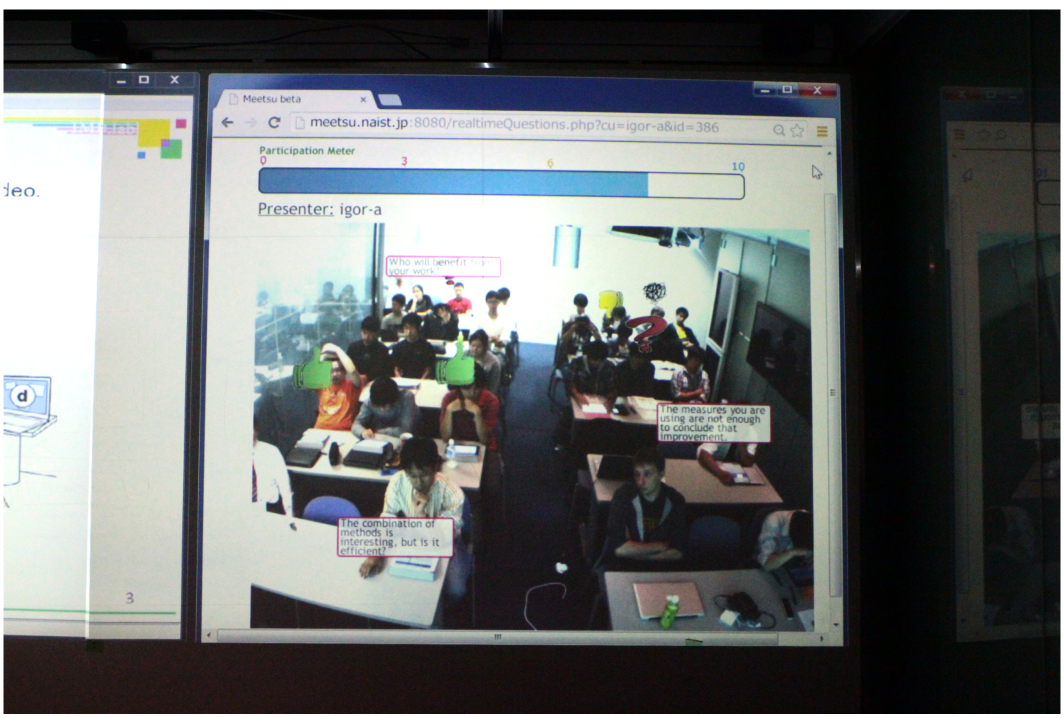

Figure 4.3. Meetsu's AR view in full.

The participation meter, which can be seen in Figure 4.3 as a blue bar above the video image, is a visual reference of how many questions/comments have been submitted and how many times the icons have been used. Questions account for one point in the scale while icons account for half a point. The distinction is based on the idea that text annotations are a more pronounced form of communication than icons. As the entries come in the system, the meter progresses accordingly. A change of color can be seen when the number surpasses rating milestones as shown in Figure 4.4. From 0 to 3, the meter is red; from 3.5 to 6.5, the meter turns yellow; and from 7 to 10, the meter changes to blue. The limit of 10 on the meter was discussed with the professors as a good number to indicate an excellent participation rate.

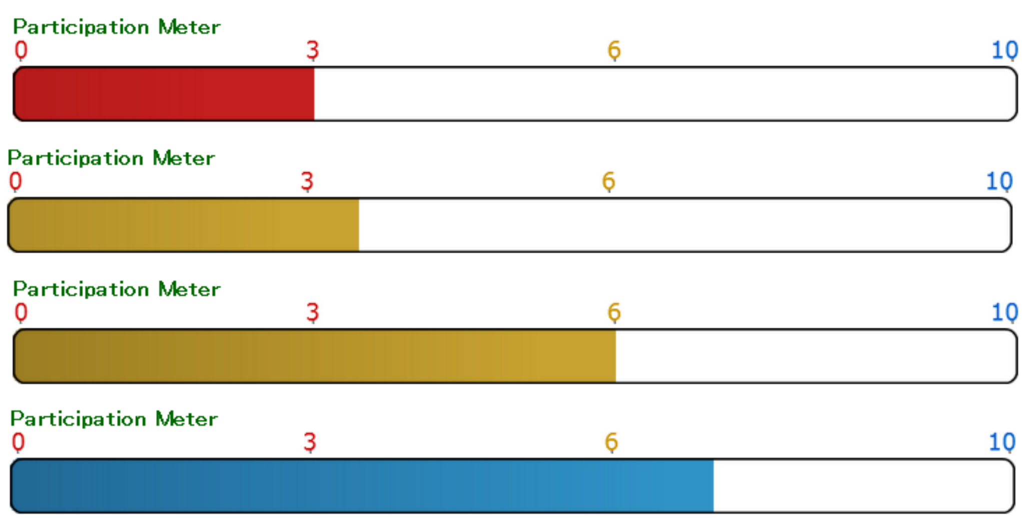

Figure 4.4. Participation meter changes color when surpassing rating milestones.

The real world view allows the participants to observe themselves and others in the environment, therefore we suppose that the AR view in the meeting support system provides a more personal and engaging experience to the participants.

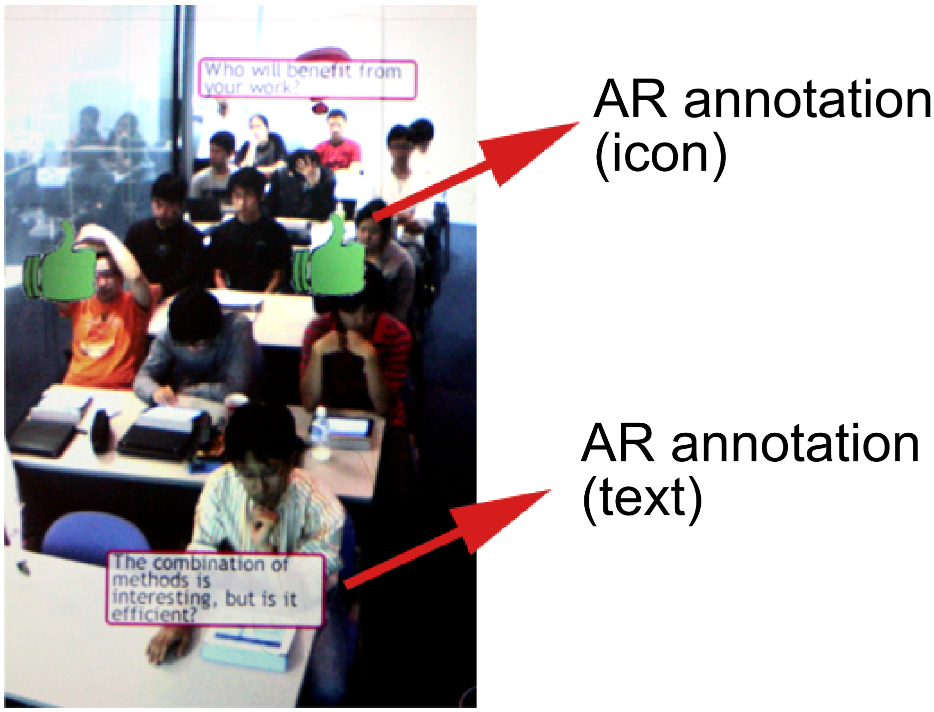

Figure 4.5. Close-up of the AR View showing in details two types of AR annotations, icon and text, rendered on the live video feed of the meeting room.

4.4. Experiments

4.4.1 Long-term Study

A participant observational study (moderate participation) was conducted within the first year from the integration of Meetsu in our lab's weekly research meeting. Although it is not reliable as a form of system evaluation, this type of study provides qualitative descriptions that can be used to formulate measurements as well as generalizations and hypotheses [86]. Works such as [87][88][89] have previously made use of this type of study in the process of acquiring thorough understanding of the context and finding behavioural trends.

Our aims were to explore how meeting participants responded to the system during the meeting, how they interacted with the system, how the interaction affected their communication patterns and meeting experience, and how the system changed the meeting dynamic. Furthermore, since some participants did not use the system actively (they did not use the web user interface to submit questions), it was important to analyze how they were affected by the passive use of the system (they were exposed to the AR view regardless).

4.4.1.1 Study Design and Method

An average of 23 participants per meeting participated in the study. They were a mix of professors, graduate and research students with an average age of 27. The group was predominantly of male averaging 95% of the group. All participants had previous exposure to AR technology. The meetings were held in our usual meeting room arranged in a classroom style (Figure 4.6).

Figure 4.6. Meeting room where all the experiments took place.

All meetings were the information meeting type arranged as presentation (Section 2.2.1.3 for details) where the speakers either presented their research progress or rehearsed for a conference presentation. The schedule of speakers had been determined in advance (an average of two speakers per meeting session) and participation was not mandatory.

Prior to the start of the meeting session, the student under the role of meeting facilitator encouraged participants to use Meetsu by submitting questions or comments and/or using the icons. However, no further measures were taken to enforce the use of the system. An average of 75% of the participants used the system throughout the study.

A total of 24 meeting sessions were observed by an evaluator who also had the role of attendee. Each meeting session averaged 2.3 hours, discarding the extra time spent with other meeting activities such as progress report by all participants and general announcements.

4.4.1.2 General Findings and Observations

The **meeting role** defines what is expected of a person in a meeting context. It is a combination of two instances described here as *conditional role* and *situational role* (see Table 4.1). Conditional roles correspond to hierarchical status such as student, associate professor and professor; situational roles are based on the assumed duties during the meeting session, such as speaker, attendee and facilitator.

During the study, five out of the nine possible combinations were observed: *Student-Attendee*, *Student-Speaker*, *Student-Facilitator*, *Professor-Attendee*, and *Associate Professor-Attendee*. As expected, **student-attendee** was the most common role since the majority of participants were students. In a research meeting, as an educational context, a student is expected to take an active role by making comments/questions. Students generally took a passive approach, demonstrating low willingness to participate, either due to lack of interest in the topic being presented or shyness.

Table 4.1. Meeting roles in research meeting.

Conditional Roles	Situational Roles
Student	Speaker
Associate Professor	Attendee
Professor	Facilitator

Students expressed concern over how much exposure they were under. Different from the student-speaker, which is inevitably the center of attention, the student-attendee is not usually prepared to handle “overexposure” (when all eyes are on one person, such as question making). Since the AR view used by Meetsu displays all participants in the room in real-time, some participants perceived it as a factor that increased social pressure, as discussed by [90], and it was *"unnecessary to display it constantly"*. Others perceived it as a video mediation of face-to-face communication that lowered the level of actual exposure. From this perspective, the fact that participants were not looking directly at an actual person's location, but to a video image of one instead, generated a less stressful situation.

Another recurrent role was **student-speaker**. This role is responsible for making a presentation during the meeting. In general, participants under this meeting role combination showed clear signs of apprehension, mostly related to the feeling of insecurity about the quality of his/her presentation. Some students can be quoted saying: *"Do I have enough progress for this presentation?"* and *"I worry that my presentation is not understandable"*. Others displayed a more personality-related aversion to speaking in public (Quote: *"I just don't like to present at all"*).

Speakers paid close attention to the annotations displayed on Meetsu's AR view in order to read the questions/comments and see the icons submitted by the audience. In this situation, speakers were trying to perceive the overall impression of the audience while seeking for approval of their presentation. The most common scenario had annotations expressing questions regarding unclear points in different passages of the presentation. Initially, receiving negative feedback was thought to cause embarrassment to the speakers, however the annotations also provided a way for them to preview the questions and identify the questioners.

Different from the previous roles, the duties of a **student-facilitator** concern the meeting as a whole, including the perception of the environment and the building up of the meeting atmosphere. It is a role in charge of conducting the meeting flow and handling its events, especially the Q&A session. It is also the person assuring the order in which participants willing to contribute can do so. Initially, a different student would take on this role every meeting session, however the constant need to instruct the facilitator on how to manage the Q&A session using Meetsu created the need to assign this role to a fixed student who was familiar with Meetsu. This measure was found to be very effective for the definite integration of the system in the meeting process. At times, the facilitator managed the Q&A session using the question list as a reference of order, but still requiring the questioner to orally address the question.

The roles of professors and associate professors were only combined with the situational role of attendee (**professor-attendee** and **assistant professor-attendee**). It was noticed that the situational role has great influence in the turn-taking procedures during the Q&A session. Students respectfully waited for the professors to express themselves first even if all attendees are given virtually equal chances to participate. A consequence of this cultural fact is that a Professor's speech tended to be very long. This procedure changed over time to allow students to make questions first. It was noted that the Professor's speech length shorten over time. Before the change, professor tended to have long speeches due to the uncertainty of whether students would make questions or not. Using Meetsu, professors become easily aware of other participants' intention to make questions, which could have affected the later length of the speech.

The **general meeting flow** consisted of three main processes: pre-meeting procedures, presentation and Q&A session. The pre-meeting procedures included the preparation of the speaker and the audience with the system, such as logging in Meetsu and enter the current meeting session (for attendees), creating a meeting session in Meetsu and loading the AddIn needed for Meetsu in the PowerPoint file. This process is followed immediately by presentation and Q&A session, which leads to a conditional step to determine whether the Q&A continues or not. Following the end of the Q&A for one presentation, another conditional step determines whether a next presentation restarts the flow or the meeting session ends (Figure 4.7).

The available **interaction models** in the meeting supported by Meetsu are expanded in comparison to the standard (unsupported) meeting. In the latter, a presenter speaks to many participants at once during the presentation and the participants have a chance to speak one-to-one with the presenter during the Q&A session. Using Meetsu, the models from the standard meeting are preserved while also supporting the many-to-one interaction since many participants have the chance to "speak" at the same time (Figure 4.8). The speaker is able to perceive this interaction through the augmented view and/or the list of questions in the web UI. Interestingly, [91] says that effective meetings must be interactive

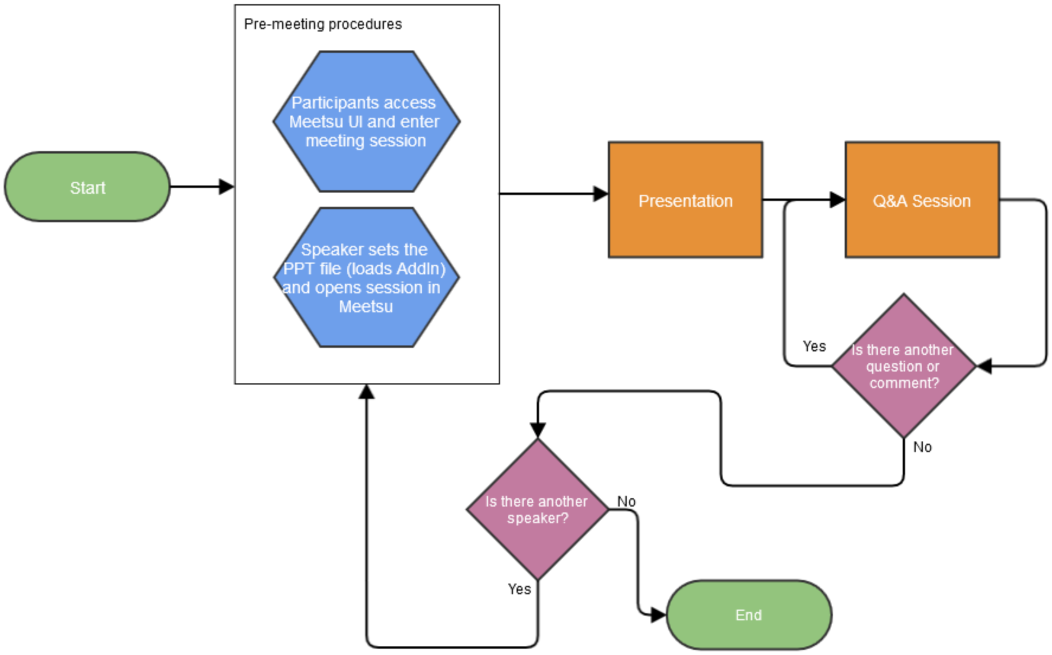

Figure 4.7. A diagram of the average research meeting flow as observed.

while maintaining balance participation among attendees. These models have a close relation to the person-to-person meeting criteria (see section 2.2.1.3) and help identify potential communication enhancements.

4.4.2 Short-term Study

4.4.2.1 Participants and Context

The data were collected for this study from graduate students, researchers and professors of two different lab groups. The average ages for Group 1 and Group 2 were 26.8 and 24.6, respectively. Each group consisted of over 20 participants; however the data of the participants who did not attend all meeting sessions were eventually discarded. The final sample sizes for Group 1 and Group 2 were 16 (five females and 11 males) and 20 (one female and 19 males), respectively. Group 1 was composed of one staff researcher and 15 students. Group 2 was comprised of three Professors and 17 students.

According to the demographic survey, 44% of the participants in Group 1 and

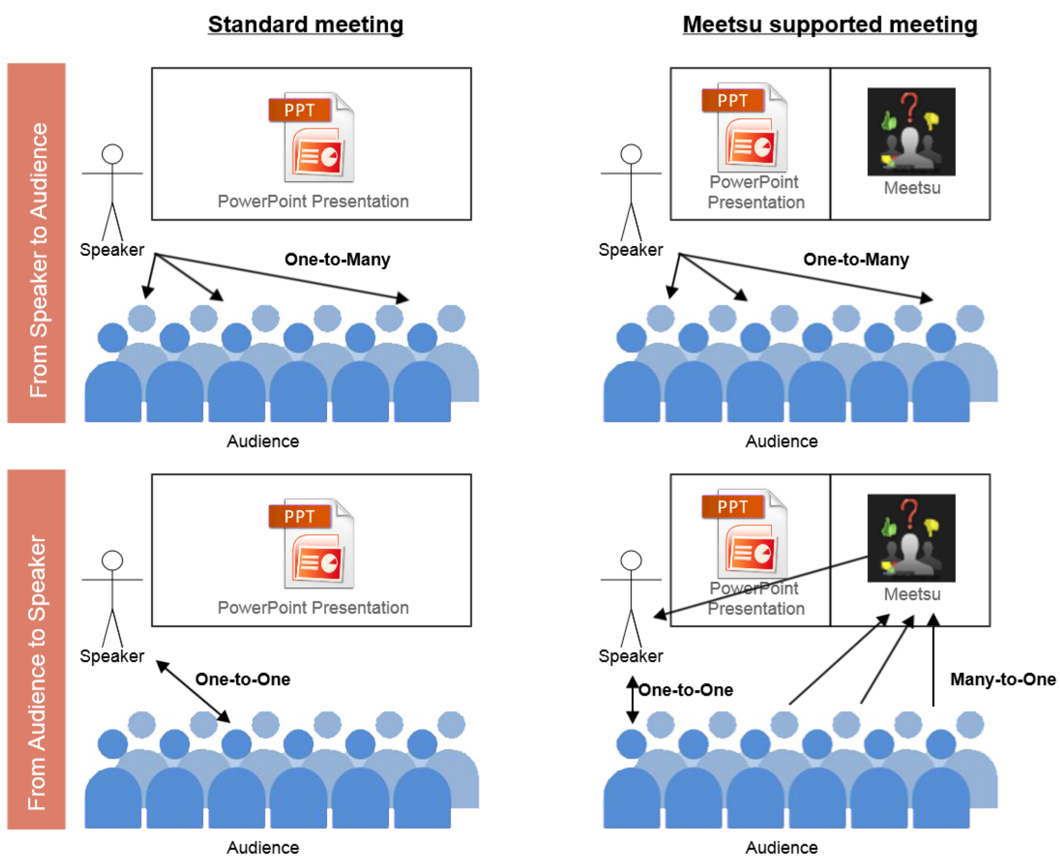

Figure 4.8. Comparison of interaction models between a standard (unsupported) meeting and a meeting supported by Meetsu.

100% in Group 2 had prior experience with an AR system. 6% of the participants in Group 1 and 20% in Group 2 had previous experience with a meeting support system.

The meetings of both groups were hosted in the same meeting room where Meetsu had been previously set up. Both wired and wireless network connections were made available. The groups' meeting schedules were maintained as had been decided on in advance according to the groups' agenda. Therefore, it was predetermined who the speaker would be and how many presentations would be held each day. The meetings were conducted by the participants themselves following their group's usual meeting style regarding turn-taking strategy and time limit.

4.4.2.2 Measures

The WTC scale was used in order to obtain a measure of the likelihood of each group initiating communication. This likelihood was measured twice for each group, on the first and last days of the experiments. Given the focus of this study on the meeting scenario, only the average WTC scores in the meeting context were considered for further analysis. Moreover, each group was subdivided into High WTC, Average WTC group and Low WTC, according to scores on the first WTC survey, following the scale in appendix C.1. The critical group (Low WTC group) became the main target since it was assumed to be the one to benefit most from the experience.

Apart from the WTC scale, the 7-point Likert-scale was used in an additional questionnaire made available online and was submitted by the participants after each meeting session, 1 representing Disagree and 7 representing Agree. The following statements were assessed: (Q1) I was distracted by the system, (Q2) The system made me anxious to participate, (Q3) The system made me willing to participate, and (Q4) The system made me engage in real-world conversation with other participants. Subjective comments were also collected.

The Likert-scale was intended as a measure for the participant to indicate how each system contributed to their meeting experience. A final comparison questionnaire asked the participants their preference regarding the AR and Non-AR systems.

The Likert-scale measures were used to indicate how each version of the system contributed to the participant's meeting experience. In addition, a final comparison questionnaire, asking their preference regarding each version was given. These measures aimed at indicating which system was more effective, and, therefore, had the biggest impact on any significant change of WTC scores before and after using the system.

4.4.2.3 Methodology

A formal user study was designed to investigate the following hypothesis: - The use of an AR meeting support system can significantly improve the levels of WTC in co-located presentation meeting contexts.

This study used a within-subjects design where all participants experienced the system under the same circumstances of environment (same meeting place), period of experiments (once every week for three weeks) and conditions (AR and Non-AR). First, a paired two-tailed t-test was used to analyze the change in the WTC mean scores before and after using Meetsu. The critical significance level used was 0.05. These results were used to note the impact of Meetsu on communication.

Following this analysis, the Wilcoxon signed-rank test and the frequencies of the Likert-scale questionnaires were used to evaluate the performance of the AR and the Non-AR versions of the system. Both versions were compared and evaluated according to measures of distraction, anxiety, willingness and engaging in communication.

4.4.2.4 Procedure

Both groups gathered separately for weekly presentation meetings for three weeks. The two versions of the system were used alternately by each group, in a different order to minimize carry over effects. Group 1 used the AR version of the system in the first two weeks and the Non-AR version in the third week. Group 2 used the AR version of the system in the first and third weeks and the Non-AR version in the second week. Both groups used the AR version on the first day of the experiments in order to instruct the participants on how to use the AR features of the system. A non-participative observation process took place during all the meetings.

The following table describes the methodological and time triangulation used for this experiment.

Prior to the experiment days (Day 2, Day 3 and Day 4), the researcher joined each group's weekly meeting for a day to observe their meeting style prior to the use of Meetsu. The difference in the meeting flow between the two groups was mainly the timing to ask a question. Group 1 would ask questions mostly at the end of the presentation, while Group 2 would interrupt the presentation to ask a question. Neither of the groups had a chairman.

Table 4.2. Table of methodological and time triangulation.

	Observation Report	WTC Survey	AR Survey	Non-AR Survey
Day 1	X			
Day 2		X		
Day 3			X*	
Day 4		X		X*

* The order of the AR and Non-AR surveys were alternated between the groups.

4.4.2.5 Conditions: AR and Non-AR

Meetsu was originally developed as a web-based AR system. However, the need to compare our system with previous approaches led us to the development of a Non-AR version. The Non-AR version differs from the AR version in two ways: the way to display participants' input and features available. While the AR version of Meetsu uses the second presentation screen to show the live video feed of the meeting room where all the AR annotations are displayed, the Non-AR version displays all the input from the audience as plain text much like the question list available in the web UI (Figure 4.9). All the features directly related to the AR view, such as the interactive icons and the highlight button, are disabled in the Non-AR version.

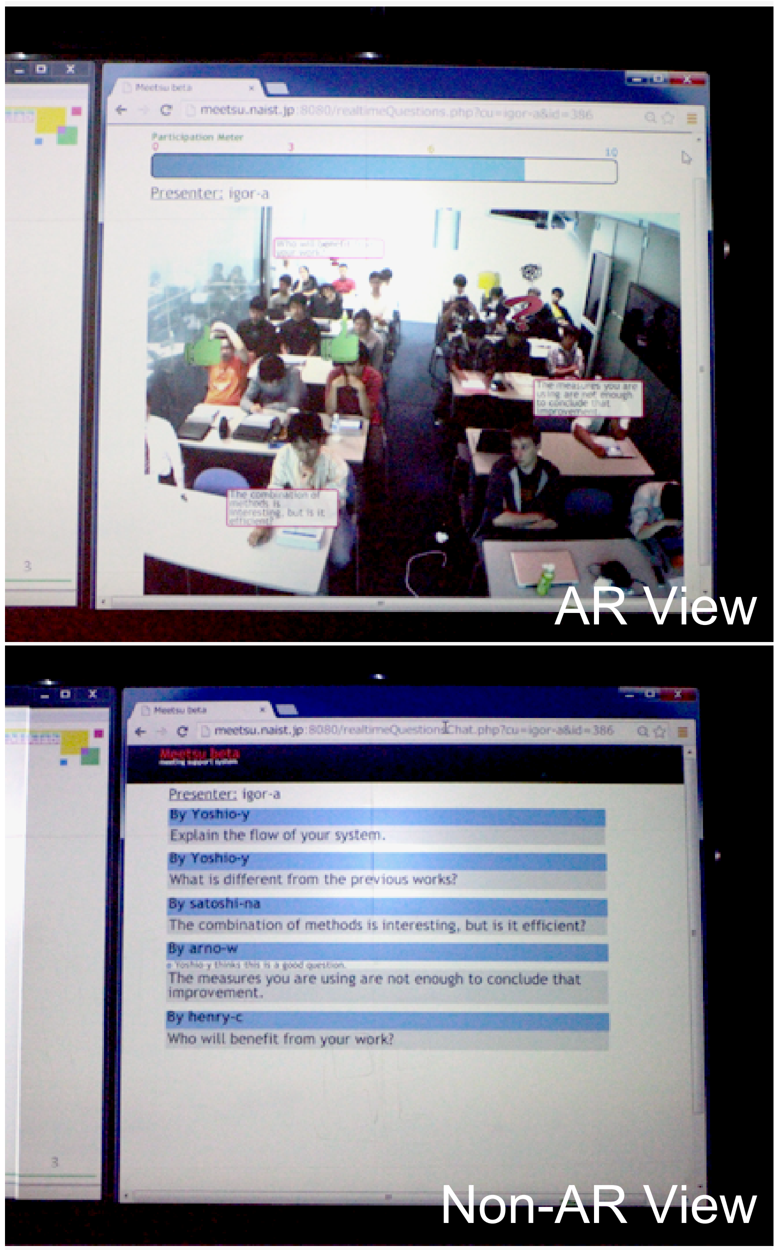

Figure 4.9. AR and the Non-AR views of Meetsu.

4.4.2.6 Results

Group 1 had seven speakers making presentations in the three days of experiments, while Group 2 had eight presentations altogether. Both groups averaged around 40 minutes per presentation with an average time per day of 93 minutes for Group 1 and 118 minutes for Group 2. The overall meeting time throughout the three weeks for Group 1 was 281 minutes and 354 minutes for Group 2. The overall meeting time can also be interpreted as the amount of time each group was exposed to the system over the three weeks of experiments (Table 4.3).

Table 4.3. Total number of presentations per group, average time of presentations and overall meeting time.

	Group 1	Group 2
Total number of presentations	7	8
Average time per presentation (minutes)	40.14	44.25
Average time per meeting day (minutes)	93	118
Overall meeting time (minutes)	281	354

The research question seeks to investigate if the use of Meetsu affects the levels of WTC in a meeting scenario. The descriptive statistics present the meeting sub-scores before and after the use of Meetsu. The results (see Table 4.4) indicate an increase in the scores of both groups. Group 1 increased around 10 points (from 56.69 to 66.27) and Group 2 increased around 3 points (from 55.65 to 58.55).

Table 4.4. Average of Meeting sub-scores per group.

	N	Mean Before	Mean After	StDev Before	StDev After
Group 1	16	56.69	66.27	22.854	22.507
Group 2	20	55.65	58.55	21.034	23.543

In order to determine the significance of this improvement, a paired t-test was employed to analyze the data (see Table 4.5). The results were t(15)=-2.692, two-tail p=0.017, suggesting that a statistically significant result for a confidence level of 95% was achieved for Group 1. At t(19)=-0.804, two-tail p=0.431, no significant difference was found for Group 2.

Both groups had a high concentration of subjects in the Low WTC group, with 56% of Group 1 (9 out of 16) and 75% of Group 2 (15 out 20). Tables 4.6 and 4.7 summarize how the system affected the Low WTC groups in Group 1

Table 4.5. Paired t-Test for Meeting sub-scores per group.

	t	*df*	Sig. (2-tailed)
Group 1	-2.692	15	0.017
Group 2	-0.804	19	0.431

and Group 2. Similar to the overall meeting sub-score, the meeting sub-score of Low WTC groups increased for both groups. Group 1 increased around 14 points (from 40.59 to 54.11) and Group 2 increased around 3 points (from 47.76 to 50.09). With a result of t(8)=-2.698, two-tail p=0.027 for Group 1, a statistically significant difference between before and after using Meetsu can be observed in the Low WTC group. No significant difference was observed for Group 2 at t(14)=-0.525, two-tail p=0.608.

Table 4.6. Average of Meeting sub-scores from Low WTC group.

	N	Mean Before	Mean After	StDev Before	StDev After
Low WTC (Group 1)	9	40.59	54.11	16.481	19.018
Low WTC (Group 2)	15	47.76	50.09	17.620	19.812

Table 4.7. Paired t-Test for Meeting sub-scores of Low WTC.

	t	*df*	Sig. (2-tailed)
Low WTC (Group 1)	-2.698	8	0.027
Low WTC (Group 2)	-0.525	14	0.608

The results for the four measures assessed using Likert-scale are shown in Table 4.8. A Wilcoxon signed-rank test was used to determine whether there is a difference in the scores for the AR and the Non-AR versions. The results indicate that there is significant difference in how Group 1 perceived distraction (Q1), z=-2.289, p=0.022. In other words, the participants from Group 1 were more distracted using the Non-AR version (Median=4.000) than the AR version (Median=3.000). The remaining measures had equal median scores for both versions therefore no significant difference was found.

The distraction (Q1) measure for Group 2 indicated that the participants were more distracted using the AR version (Median=2.500) than the Non-AR (Median=2.000). The measure for anxiety (Q2) indicated that the participants felt that the system made them more anxious to participate when using the Non-AR (Median=2.000) version than the AR (Median=1.500). No significant difference was found in any of the measures from Group 2.

Table 4.8. Wilcoxon Signed-Rank test results showing the median score of each version of the system within each group in each assessed measure (Q1=distraction, Q2=anxiety, Q3=willingness and Q4=engage in real world conversation).

	Measures		N	Median	Z	Sig. (2-tailed)
Group 1	Q1	Non-AR	16	4.000	-2.289	0.022
		AR	16	3.000		
	Q2	Non-AR	16	2.000	-0.104	0.917
		AR	16	2.000		
	Q3	Non-AR	16	4.000	-0.586	0.558
		AR	16	4.000		
	Q4	Non-AR	16	4.000	-0.333	0.739
		AR	16	4.000		
Group 2	Q1	Non-AR	20	2.000	-0.106	0.916
		AR	20	2.500		
	Q2	Non-AR	20	2.000	-0.749	0.454
		AR	20	1.500		
	Q3	Non-AR	20	4.000	-1.338	0.181
		AR	20	4.000		
	Q4	Non-AR	20	4.000	-0.447	0.655
		AR	20	4.000		

Table 4.9 shows the frequency table of answers for the forced-choice question

comparison between AR and Non-AR versions. The participants were asked to choose the preferred version for each of the measures from the Likert-scale. An estimated 69% of the participants in Group 1 and 85% of Group 2 considered the AR version more distracting. The system was considered to make one willing to participate by an estimated 62.5% of the participants in Group 1 and 45% of Group 2. Similarly, about 62.5% of Group 1 and 75% of Group 2 considered the system to have made them engage in real-world conversation.

Q2 was the only measure which showed a disagreement between the groups. While an estimated 62.5% of Group 1 considered that the system made them more anxious to participate, participants with the same opinion represented 45% of Group 2.

Table 4.9. Frequency table of the comparison results showing which system was chosen in each assessed measured.

	Group 1 (N=16) Frequency (Percent)		Group 2 (N=20) Frequency (Percent)	
	AR	Non-AR	AR	Non-AR
Q1	11 (69.0%)	5 (31.0%)	17 (85.0%)	3 (15.0%)
Q2	10 (62.5%)	6 (37.5%)	9 (45.0%)	11 (55.0%)
Q3	11 (69.0%)	5 (31.0%)	16 (80%)	4 (20.0%)
Q4	10 (62.5%)	6 (37.5%)	15 (75.0%)	5 (25.0%)

Apart from asking the participants their opinions on how much they think each system stimulated their participation, a rundown of their actual participation during the meetings was extracted from the system's database. These additional results present the number of questions and comments produced by each group over the three days of experiments. Table 4.10 details the numbers per WTC group and per system version. Both groups participated more under the AR version of the system with Group 1 and Group 2 producing 14 and 36 questions/comments, respectively. For Group 1, a breakdown of these numbers showed that the inputs from the Low WTC Group accounted for the highest participation among the three WTC groups with 43% of all the inputs when using the AR version. The same group accounted for 20% of all the inputs when using the Non-AR version, the lowest participation rate.

Table 4.10. Number and percentage of questions/comments collected by each type of system per WTC group.

	Group 1		Group 2	
	AR	Non-AR	AR	Non-AR
High WTC	4 (28.5%)	5 (50%)	0	0
Avg WTC	4 (28.5%)	3 (30%)	0	0
Low WTC	6 (43%)	2 (20%)	36 (100%)	10 (100%)
Total	14	10	36	10

4.4.2.7 Discussion

Regarding **communication** , both groups showed an increase in WTC meeting sub-scores, overall and for the Low WTC group. Given the higher participation rate and overall preference, the AR version can be credited as the agent of this change. With Group 1 accounting for a statistically significant result in both cases, this provides the first measure of WTC for an AR system and should stimulate research on further communication scenarios in which AR can be applied.

While both WTC scores improved for Group 2, the same statistical significance could not be reproduced in their results. Nevertheless, the observation that took place during the meetings provided a good indicator of the participants' positive response towards the system. In a comparison based on the observation report produced prior to the use of the system, it was noted that the feedback came from beyond the usual circle of active audience members. The more proactive participants used speech to clearly express their ideas, while the less prominent participants used the icons. In the end, both contributed to the meeting which shows that the interest in communicating was raised and at times concretized.

On the analysis of **AR versus Non-AR** , the comparison revealed interesting results, especially for the Distraction and Anxiety measures. Both groups perceived the AR version as being more distracting. The first interpretation of this result was negative, since distraction was understood as something that interferes in a meeting. However, this did not prevent the participants from using the AR version more often than its counterpart. An additional survey was conducted in order to further investigate the perceived distraction. Both groups were asked to answer a two-question questionnaire using 7-point Likert scales, this time focused on the AR view instead of the whole system. The first question asked to what degree the subject felt the AR view was distracting in order to verify the consistency with the previous results. The second question asked to what degree the subject ’ s concentration was hindered by the AR view.

Table 4.11. Comparison of distraction average from AR view between the previous data and the new data.

	Non-AR	AR (prev.)	AR (new)
Group 1	4.0	3.0	2.3
Group 2	2.0	2.5	3.0

As shown in Table 4.11, the results of the first question showed that the average of distraction associated with the AR view remained consistently lower than the Non-AR view for Group 1, and higher than the Non-AR view for Group 2. Both groups scored low averages in the second question (Group 1 scored 2.4 and Group 2 scored 2.3), which demonstrated that the AR view did not prevent the participants from concentrating in the meeting. Following the questionnaire, a short interview took place to directly ask the participants to detail their perceived distraction.

Overall, the participants who considered the AR view as more distracting reported that it was more intuitive, since they could easily identify the owner of the comment, allowing them to quickly return to the real world discussion. On the other hand, the participants who considered the Non-AR view as more distracting reasoned that the text-based nature of view compelled them to read the entries, which made them spend more time than desired at it, either reading the comments or identifying the owner. Furthermore, the AR view was considered the livelier due to the animations (icons and text popping) and the faces of the people in the room, which prompted participants to pay more attention to it, while the Non-AR was considered unattractive.

An additional comparison of perceived distraction for two groups based on how long they have used the system was also conducted. The short-term group includes the results gathered from Group 1 and Group 2 over three weeks of

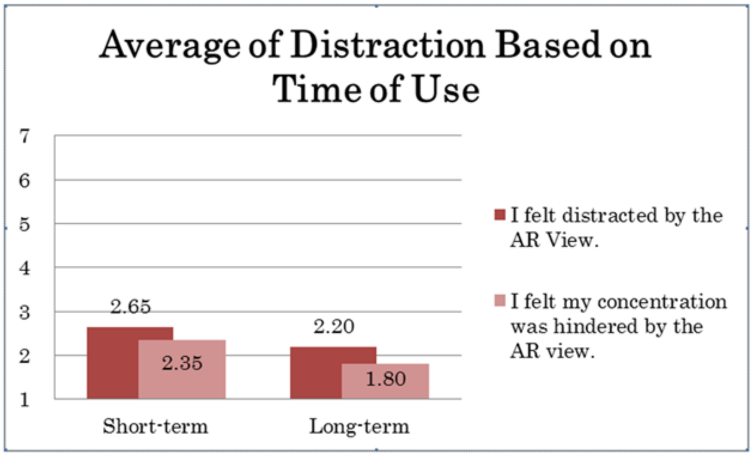

Figure 4.10. Distraction based on time of use for two groups: short-term (three weeks) and long-term (over six months).

experiments. The long-term group shows results of a survey conducted in our lab with the participants who have been using the system for over a year (13 people at the time of this survey). As shown in Figure 4.10, the perceived distraction is highly affected by how familiar the participants are to the technology, as it can be noted by the decrease in both measures for the long-term group.

Based on the results from the survey, we can conclude that the distraction associated with the AR view carried a positive meaning for the participants, unlike the Non-AR. Nevertheless, the perceived distraction is likely to decrease as the participants become more used to the technology.

The Anxiety measure was the only measure in which the groups did not agree. Group 1 felt the AR version caused more anxiety than the Non-AR version while Group 2 felt otherwise. Group 1 is particularly interesting, since while they considered the AR version to cause more anxiety, they also considered this version to make them more willing to participate. It may be that the AR version has have a quality that raises the levels of anxiety, but this does not seem to be detrimental, given that the Group 1 had the best results out of the two groups.

Following the analysis of the results, we concluded that the AR version performed better in three out of the four assessed measures (Q1, Q3 and Q4), which confirmed the effectiveness of the AR version over the non-AR. Moreover, it is possible to affirm that the AR version had a stronger impact on the accounted increase of the WTC scores.

Some of the shortcomings of this work include some technical issues (online slides were flickering for some users) and unclear procedures for the speaker regarding the displayed questions. Since the questions were displayed at any time, the speakers got confused as to whether they should answer immediately or not. To quote one of the speakers, "I don't know when to answer the questions". Moreover, the displayed questions were not numbered or ordered, the highlight button being the only option available to identify the current topic. This can be considered a downside of the current AR implementation. The Non-AR view did not undergo the same problem, since the questions were shown as a list arranged by order of submission.

4.5. Summary

This chapter has described Meetsu system as a novel approach to co-located meetings using AR technology. The use of text and icon annotations overlaid on the live video feed of the meeting was used as a communication enhancer and a meeting management assisting tool. Furthermore, a WTC study was conducted alongside a comparison study of AR and Non-AR views of text annotations.

The results of the study can be summarized as:

- Group 1 and Group 2 showed an increase in their average WTC meeting scores after using Meetsu, for their entire group (from 56.69 to 66.27 in Group 1, and from 55.65 to 58.55 in Group 2) and for their low WTC group (from 40.59 to 54.11 in Group 1, and 47.76 to 50.09 in Group 2). A significant difference was found in Group 1. No significant difference was found in Group 2.

- A comparison study between AR and Non-AR views of Meetsu using Likert scales showed lower distraction rate in Group 1 when using the AR view, stressed by a significant difference. No significant difference was found for the remaining measures.

- A comparison study between AR and Non-AR views of Meetsu using forced-choice questions showed that the AR view was the more distracting but also the more engaging in both groups. The AR view was considered the one to cause more anxiety in Group 1 while Group 2 considered the Non-AR view in the same matter.

- Higher participate rate was demonstrated while using the AR view in both groups. In similar fashion, the Low WTC group was identified as the group of people who participated the most using Meetsu.

CHAPTER 5

Conclusions and Future Work

5.1. Review of the Thesis

This thesis has proposed two new approaches to ARMC in remote and co-located scenarios. A remote ARMC system called HANDY, a different take on the AR video chat system, was developed and further investigated as a social presence enhancing system. A web-based AR meeting support system, Meetsu, was developed, which aims at increasing willingness to communicate in co-located presentation meeting. A first measure of WTC for AR technology was investigated and the results showed that AR can increase the overall WTC meeting sub-scores with great effect on the critical WTC group (Low WTC Group). The summary of these contributions follows:

- A prototype system named HANDY (details in Chapter 3) was developed targeting AR video conferencing. This system evaluated the effect of the ARMC on Social Presence. Experiment results with HANDY showed significant higher degree of social presence when the AR feature was used (HANDY On condition). Furthermore, significant results were found for enjoyment, ease of communication and intuitiveness measures during HANDY

On condition.

- A novel approach to AR meeting support was proposed with the development of Meetsu. The use of a shared live video feed of the meeting room along with the display of AR annotations created an effective AR co-located EMS solution, which increased the levels of willingness to communicate of the studied groups. The study demonstrated that co-located ARMC has the potential to promote communication.

5.2. Design findings

This work studied, developed and evaluated ARMC systems in remote and co-located scenarios. The design of the proposed approaches was largely affected by the parallel study of human factors. Here are some of the findings which can provide further insight to future systems:

- **Shared augmented view**
  The nature of augmented reality technology combined with the communication process is often associated with an enjoyable and more engaging experience. Therefore, it is appropriate to establish, in the design of the ARMC, a way to share any augmentation created by the system. In doing so, all participants can take advantage of the beneficial characteristics that come along with a shared augmented view. Further discussion can take place to argue whether a centralized shared AR view has more advantages over a distributed shared view. The findings of this work are enough to indicate that in a co-located meeting situation, for example, a centralized approach converges all the discussion to one big shared screen, instead of having the AR view being displayed in individual devices. By centralizing a AR view, it is possible to mitigate dispersion while increasing awareness to all.

- **The speech balloon convention**
  The real-time face image along with annotations on top of it conveys a stronger message of communication due to the similarity with the speech balloon convention. Also, the combination of face image and words conveys

a stronger message than other arguably similar methods such as combining avatar and words, or just simply using a user's location and words. The face identification of a person is also part of the communication process related to cognition, and therefore deserves to be taken into account as a potential ARMC factor.

- **The use of mirror images**
  When the user is able to see himself during the ARMC process, the communication experience presents a new dimension of awareness. While the perception of another may stimulate communication and provide the sense of presence, the perception of oneself may allow for a stronger sense of being, which can also motivate one's involvement in the communication process. Both of these are pertinent factors to communication.

- **2D representation of AR**
  The development of ARMC systems does not have to be solely associated with the traditional AR characteristics (3D registration and tracking). Emulating the visual outcome of AR (the resulting combination of real image and virtual imagery) by using 2D virtual contents onto video images can be also regarded as a valid approach to evaluating the applicability of AR technology. It is important however to have a deep understanding of the problem it in order to determine whether the 2D representation is the best way to go. The 2D approach can not emulate all the potential benefits of a traditional AR approach, therefore the purpose of the system has to be well defined.

- **AR conception with human factors**
  Any system focused on providing efficient human communication cannot avoid integrating human factors from the conception process until the evaluation methods. The new trend of using AR as a communication tool requires a new trend in how the AR solutions are conceived. It is not enough to just allow two people to meet virtually, it is necessary to fully understand how this setting is affecting the way they feel, communicate and understand each other. Although ideally ARMC system would be created using the understanding of the factor to create a better solution, having

a AR solution being measured in terms of how it affects human factors is also a valid approach to integrating these factors in the process of creating a more cognitive and ergonomic solution.

5.3. Future Work

The further improvement of the developed systems and the further study of human factors are some of the suggested future work:

- **General**
  1. Both system could benefit from the incorporation of standard AR technology. The use of human tracking along with 3D registration could greatly improve the interaction methods.
- **HANDY & Social Presence**
  1. Implementing a one camera solution would be ideal from an interaction point of view, since the user would only need to move his hand in the front area, instead of the side. In addition, it would make the solution even more affordable and closer to what is the standard setup in video chat communication.
  2. Despite the considerable good results achieved with the experiments, a larger sample size could drastically increase any consistent differences between the tested conditions.
  3. The use of the alternative social presence measures could be used to assure the reported results, along with correlated measures such as co-presence.
- **Meetsu & WTC**
  1. Investigating the use of the system in different contexts (such as classroom lectures), different audience sizes (e.g., groups of five, 20 and 50) or different configurations (involving people who do not know each other, for example) may impact the communication measures.

2. Even though the use of WTC provided a first attempt to measure the effect of AR on communication, other important factors such as anxiety can be formally assessed by an appropriate measure, such as the Personal Report on Communication Anxiety (PRCA-24).
3. Extended work can also include remote uses of Meetsu in order to evaluate the communication measures when the participants do not share the cues of being face-to-face.

Publication List

Journal Paper

1. Igor de Souza Almeida, Goshiro Yamamoto, Takafumi Taketomi, Jun Miyazaki and Hirokazu Kato, “Increasing Willingness to Communicate in Co-located Presentation Meetings using Augmented Reality”. VRSJ Journal. [to be published in March/2014] (related to *Chapter 4*).

International Conferences

1. Igor de Souza Almeida, Marina Oikawa, Jordi Carres Polo, Jun Miyazaki, Mark Billinghurst, Hirokazu Kato, “AR-based video-mediated communication”, Proc. of *The XIV Symposium on Virtual and Augmented Reality (SVR2012)*, pp. 125-130, Niteroi, Brazil, May 2012 (related to *Chapter 3*).

2. Igor de Souza Almeida, Jun Miyazaki, Goshiro Yamamoto, Makoto Fujisawa, Toshiyuki Amano, Hirokazu Kato, “Meetsu: Augmented Reality Co-located Meeting Support System”, Proc. of *The 5th Korea-Japan Workshop in Mixed Reality (KJMR2012)*, Extended Abstract, Seoul, South Korea, April 2012 (related to *Chapter 4*).

3. Igor de Souza Almeida, Marina Atsumi Oikawa, Jordi Carres Polo, Jun Miyazaki, Mark Billinghurst, Hirokazu Kato, "AR-based Social Presence Enhancement in Video-chat Communication", Proc. of *IEEE Symposium on 3D User Interfaces 2012 (3DUI2012)*, pp. 135-136, California, USA, March 2012 (related to *Chapter 3*).

4. Igor de Souza Almeida, Jun Miyazaki, Goshiro Yamamoto, Makoto Fujisawa, Toshiyuki Amano and Hirokazu Kato, " AR based Co-located Meeting Support System", Proc. of *The 21st International Conference on Artifical Reality and Telexistence (ICAT2011)*, pp. 146, Osaka, Japan, November 2011 (related to *Chapter 4*).

Award

1. ICAT2011 Best Poster Award, November 2011.

Acknowledgments

First and foremost, my deepest gratitude to my supervisor Prof. Hirokazu Kato for continuously believing in me through the ups and downs of the past six years. He made it possible for me to realize my dream of coming to Japan and pursuing a post graduation. When Prof. Kato said 'yes' to that small town boy in Brazil, he made him believe that he was ready to take on the world. Working with him allowed me to learn, even if just a little, about how to think as a leader, how to conduct myself as a researcher, how to laugh when all seems out of place and how to keep pushing myself towards a goal. I would like to thank him for sharing his vision and knowledge about work, and sometimes ice skating.

I also would like to thank my thesis committee Professor Satoshi Nakamura and Associate Professor Kiyoshi Kiyokawa, from Osaka University, for the valuable advices and for believing in me when I needed the most. Special thanks to my co-supervisor Assistant Professor Goshiro Yamamoto, who helped me tremendously in the final decisive moments of my doctoral course. I would like to thank him for all the help with my journal publication, thesis review, and meetings after meetings. I would not have made without his guidance and dedication. お疲れ様でした。

I would also like to extend my appreciation to the Interactive Media Design lab members. I am particularly grateful to Assistant Professor Takafumi Taketomi, former Associate Professor Jun Miyazaki, Yuichiro Fujimoto, Atsushi Keyaki and Angie Chen for your unconditional help and support.

I also want to acknowledge the importance of my Japanese family, Henry Chu, Nico Prananta, and Marie Engelene Obien. They are my rock, the ones that always reach out to me even when I was not deserving. I heartily thank them for being a little bit crazy, sometimes a little bit musical, but at the end they are a little bit of me. In similar fashion, I would like to thank Marina Oikawa for her everlasting support and care. The days in Japan became that much funnier, brighter, and easier by having her around. I never thanked her enough for paving the way for me, and for showing by example that dedication can do wonders.

Special thanks to everyone in Japan who have helped me at some point during these years of hard work. Especially, Erlyn Manguilimotan, Marcos Villagra, Jordi Polo, Mai Takeuchi, and Tatsuya Hiramatsu.

I would like to thank my friends in Brazil for keeping me in their prayers and wishing me all the best in any circumstance. My success is their success: Inara Cavalcante, Luciana Kellen da Matta, Daniel Creao, Aline Figueiredo, Daniel Moreira, Francisco Werther Junior, the Wanzeller family and the Emmi family.

I am beyond grateful to my family for always being there for me and making me feel like I'm good enough. They are the reason I am able to be where I am, do what I do and be who I am. Thank you!

Appendix

A. HANDY Questionnaires (English)

A.1 Social Presence

The following questionnaire was used to measure Social Presence from English-speaking subjects:

Social Presence Survey

Please rate the condition HANDY OFF.

Impersonal	1 2 3 4 5 6 7	Personal
Cold	1 2 3 4 5 6 7	Warm
Ugly	1 2 3 4 5 6 7	Beautiful
Small	1 2 3 4 5 6 7	Large
Insensitive	1 2 3 4 5 6 7	Sensitive
Colourless	1 2 3 4 5 6 7	Colourful
Unsociable	1 2 3 4 5 6 7	Sociable
Closed	1 2 3 4 5 6 7	Open
Passive	1 2 3 4 5 6 7	Active

A.2 Post-Task Survey

The following questionnaire was used to evaluate additional measures from English-speaking subjects after every task:

.PLEASE READ THE STATEMENTS BELOW AND INDICATE YOUR PREFERENCE

1.	How easy was to use each system?	**Which system was easier to use?**	
		☐ HANDY OFF	☐ HANDY ON
2.	How fun was to use each system?	**Which system was more fun to use?**	
		☐ HANDY OFF	☐ HANDY ON
3.	How easy was to communicate with the other person?	**In which system was easier to communicate with the other person?**	
		☐ HANDY OFF	☐ HANDY ON
4.	How natural was the communication?	**In which system the communication was more natural?**	
		☐ HANDY OFF	☐ HANDY ON
5.	How easy was to understand the other person?	**In which system was easier to understand the other person?**	
		☐ HANDY OFF	☐ HANDY ON
6.	How close was to a face-to-face communication?	**Which system was closer to a face-to-face communication?**	
		☐ HANDY OFF	☐ HANDY ON

In which of the two conditions (HANDY OFF and HANDY ON) did you feel like the other person was most like being in the same room as you?

If you had to do a similar task again, which system would you prefer?

Did using the augmented system feel more natural?

A.3 Post-Session Survey

The following questionnaire was used to evaluate additional measures from English-speaking subjects after every session:

Indicate your preferred answer by marking an "X" in the appropriate box of the seven-point scale. Please consider the entire scale when making your responses, as the intermediate levels may apply.

PLEASE READ THE STATEMENTS BELOW AND INDICATE YOUR DEGREE OF AGREEMENT WITH EACH STATEMENT							
1. How easy was to use each system?	**HANDY OFF**						
	☐ easy	☐	☐	☐ neutral	☐	☐	☐ not easy
	HANDY ON						
	☐ easy	☐	☐	☐ neutral	☐	☐	☐ not easy
	Which system was easier to use?						
			☐ HANDY OFF		☐ HANDY ON		
2. How fun was to use each system?	**HANDY OFF**						
	☐ fun	☐	☐	☐ neutral	☐	☐	☐ not fun
	HANDY ON						
	☐ fun	☐	☐	☐ neutral	☐	☐	☐ not fun
	Which system was more fun to use?						
			☐ HANDY OFF		☐ HANDY ON		
3. How easy was to communicate with the other person?	**HANDY OFF**						
	☐ easy	☐	☐	☐ neutral	☐	☐	☐ not easy
	HANDY ON						
	☐ easy	☐	☐	☐ neutral	☐	☐	☐ not easy
	In which system was easier to communicate with the other person?						
			☐ HANDY		☐ HANDY		

			OFF		ON		
4. How natural was the communication?	**HANDY OFF**						
	☐	☐	☐	☐	☐	☐	☐
	natural			neutral			not natural
	HANDY ON						
	☐	☐	☐	☐	☐	☐	☐
	natural			neutral			not natural
	In which system the communication was more natural?						
			☐ HANDY OFF		☐ HANDY ON		
5. How easy was to understand the other person?	**HANDY OFF**						
	☐	☐	☐	☐	☐	☐	☐
	easy			neutral			not easy
	HANDY ON						
	☐	☐	☐	☐	☐	☐	☐
	easy			neutral			not easy
	In which system was easier to understand the other person?						
			☐ HANDY OFF		☐ HANDY ON		
6. How close was to a face-to-face communication?	**HANDY OFF**						
	☐	☐	☐	☐	☐	☐	☐
	close			neutral			not close
	HANDY ON						
	☐	☐	☐	☐	☐	☐	☐
	close			neutral			not close
	Which system was closer to a face-to-face communication?						
			☐ HANDY OFF		☐ HANDY ON		

1. How different was the experience during condition HANDY ON compared with HANDY OFF?

2. Comments

3. What did you like best about the experience during condition HANDY ON?

4. What could be improved in the experience?

B. HANDY Questionnaires (Japanese)

B.1 Social Presence

The following questionnaire was used to measure Social Presence from Japanese-speaking subjects:

Social Presence Survey　ソーシャルプレゼンス調査

（ソーシャルプレゼンス：人がメディアを介したコミュニケーションにおいて，相手を実在する人と認識する度合い）

HANDY 未使用時の状態の評価をしてください.

人間味がない	1 2 3 4 5 6 7	人間味がある
冷たい	1 2 3 4 5 6 7	温かい
醜い	1 2 3 4 5 6 7	美しい
小さい	1 2 3 4 5 6 7	大きい
繊細でない	1 2 3 4 5 6 7	繊細な
華やかでない	1 2 3 4 5 6 7	華やかな
社交的でない	1 2 3 4 5 6 7	社交的な
閉鎖的な	1 2 3 4 5 6 7	開放的な
受動的な	1 2 3 4 5 6 7	能動的な

B.2 Post-Task Survey

The following questionnaire was used to evaluate additional measures from English-speaking subjects after every task:

QUESTIONNAIRE – Task

日付: ________________________ タスク: ____________________

以下の設問を読み、各設問に対し７段階のいずれかで評価してください。

1.	それぞれのシステムの使い易さはどうでしたか。	どちらのシステムが使いやすかったですか ☐ HANDY OFF ☐ HANDY ON	
2.	システムを使ってみて楽しかったですか	使ってみてより楽しかったのはどちらですか ☐ HANDY OFF ☐ HANDY ON	
3.	相手とのコミュニケーションの取りやすさはどうでしたか	どちらのシステムが相手とコミュニケーションし易いですか ☐ HANDY OFF ☐ HANDY ON	
4.	コミュニケーションは自然でしたか	コミュニケーションがより自然だったのはどちらですか ☐ HANDY OFF ☐ HANDY ON	
5.	相手の感情は理解し易かったですか	相手の感情がわかりやすいのはどちらですか ☐ HANDY OFF ☐ HANDY ON	
6.	対面のコミュニケーションにどの程度近かったですか	より対面のコミュニケーションに近かったのはどちらですか ☐ HANDY OFF ☐ HANDY ON	

2 つの状態(HANDY 使用時と未使用時)のうち，パートナーがあなたと同じ部屋にいるかのような感覚に近かったのはどちらですか．

もし同じタスクをもう一度しなければならない時，どちらの状態で行いたいですか．

拡張現実感を用いることが自然なコミュニケーションを促せていると感じましたか．

B.3 Post-Session Survey

The following questionnaire was used to evaluate additional measures from English-speaking subjects after every session:

あなたの答えに合わせて適切なボックスにチェックしてください。
賛成・反対の度合いに合わせて中間のチェックボックスも利用してください。

以下の設問を読み、各設問に対し７段階のいずれかで評価してください。							
1. それぞれのシステムの使い易さはどうでしたか。	**HANDY OFF**						
	☐	☐	☐	☐	☐	☐	☐
	簡単			どちらでもない			簡単ではない
	HANDY ON						
	☐	☐	☐	☐	☐	☐	☐
	簡単			どちらでもない			簡単ではない
	どちらのシステムが使いやすかったですか						
			☐		☐		
			HANDY OFF		HANDY ON		
2. システムを使ってみて楽しかったですか	**System A – HANDY OFF**						
	☐	☐	☐	☐	☐	☐	☐
	楽しかった			どちらでもない			楽しくなかった
	System B – HANDY ON						
	☐	☐	☐	☐	☐	☐	☐
	楽しかった			どちらでもない			楽しくなかった
	使ってみてより楽しかったのはどちらですか						
			☐		☐		
			HANDY OFF		HANDY ON		
3. 相手とのコミュニケーションの取りやすさはどうでしたか	**System A – HANDY OFF**						
	☐	☐	☐	☐	☐	☐	☐
	簡単			どちらでもない			簡単ではない

	System B – HANDY ON						
	☐	☐	☐	☐	☐	☐	☐
	簡単			どちらでもない			簡単ではない
	どちらのシステムが相手とコミュニケーションし易いですか						
			☐		☐		
			HANDY OFF		HANDY ON		
4. コミュニケーションは自然でしたか	**System A – HANDY OFF**						
	☐	☐	☐	☐	☐	☐	☐
	自然			どちらでもない			自然ではない
	System B – HANDY ON						
	☐	☐	☐	☐	☐	☐	☐
	自然			どちらでもない			自然ではない
	コミュニケーションがより自然だったのはどちらですか						
			☐		☐		
			HANDY OFF		HANDY ON		
5. 相手の感情は理解し易かったですか	**System A – HANDY OFF**						
	☐	☐	☐	☐	☐	☐	☐
	理解し易い			どちらでもない			理解しやすくはない
	System B – HANDY ON						
	☐	☐	☐	☐	☐	☐	☐
	理解し易い			どちらでもない			理解しやすくはない
	相手の感情がわかりやすいのはどちらですか						
			☐		☐		
			HANDY OFF		HANDY ON		

6. 対面のコミュニケーションにどの程度近かったですか	**System A – HANDY OFF**						
	☐	☐	☐	☐	☐	☐	☐
	近い			どちらでもない			近くはない
	System B – HANDY ON						
	☐	☐	☐	☐	☐	☐	☐
	近い			どちらでもない			近くはない
	より対面のコミュニケーションに近かったのはどちらですか						
			☐		☐		
			HANDY OFF		HANDY ON		

1. 提案システムを使った場合、使わなかった場合でどのような違いが有りましたか。

2. 自由にコメントをお願いします。

3. 提案システムを使ってみて最も良かった点を教えてください。

4. このシステムにあったら良いと思う機能や改善点があれば教えてください。

C. Meetsu Questionnaires (English)

C.1 WTC Questionnaire

The following questionnaire was used to measure Willingness to Communicate from English-speaking subjects:

interactive Media Design Lab
Graduate School Of Information Science

#1 Questionnaire

Name: ______________________________________ **Date:** _____________________

Directions: Below are 20 situations in which a person might choose to communicate or not to communicate. Presume you have completely free choice. Indicate the percentage of times you would choose to communicate in each type of situation. Indicate in the space at the left of the item what percent of the time you would choose to communicate.
(0 = Never to 100 = Always)

_____1. Talk with a service station attendant.
_____2. Talk with a physician.
_____3. Present a talk to a group of strangers.
_____4. Talk with an acquaintance while standing in line.
_____5. Talk with a salesperson in a store.
_____6. Talk in a large meeting of friends.
_____7. Talk with a police officer.
_____8. Talk in a small group of strangers.
_____9. Talk with a friend while standing in line.
_____10. Talk with a waiter/waitress in a restaurant.
_____11. Talk in a large meeting of acquaintances.
_____12. Talk with a stranger while standing in line.
_____13. Talk with a secretary.
_____14. Present a talk to a group of friends.
_____15. Talk in a small group of acquaintances.
_____16. Talk with a garbage collector.
_____17. Talk in a large meeting of strangers.
_____18. Talk with a spouse (or girl/boyfriend).
_____19. Talk in a small group of friends.
_____20. Present a talk to a group of acquaintances.

C.2 Additional Questionnaire - Audience Type

The following questionnaire was used to measure distraction, anxiety, willingness to communicate and engaging in real world conversation from English-speaking meeting attendees:

interactive
Media Design Lab
Graduate School Of Information Science

#2 Questionnaire (Audience Type)

Name: Age: Sex: *M* *F*

Username:

Have you experienced AR technology before? *Yes* *No*

Have you used a meeting support system before? *Yes* *No*

How often did you use Meetsu today?

Never *Once* *2-3 times* *More than 3 times*

Closed-ended questions

I was distracted by the system.

Disagree 1 2 3 4 5 6 7 Agree

The system made me anxious to participate.

Disagree 1 2 3 4 5 6 7 Agree

The system made me willing to participate.

Disagree 1 2 3 4 5 6 7 Agree

The system made me engage in real-world conversation with other participants.

Disagree 1 2 3 4 5 6 7 Agree

Open-ended questions

What was good about this system?

Comments.

C.3 Additional Questionnaire - Speaker Type

The following questionnaire was used to measure distraction, anxiety, willingness to communicate and engaging in real world conversation from English-speaking meeting speakers:

interactive Media Design Lab
Graduate School Of Information Science

#2 Questionnaire (Speaker Type)

Demographics

Name: Age: Sex: *M* *F*

Username:

Have you experienced AR technology before? *Yes* *No*

Have you used a meeting support system before? *Yes* *No*

iHow often did you use Meetsu today?

Never *Once* *2-3 times* *More than 3 times*

Closed-ended questions

I was distracted by the system.

Disagree *1* *2* *3* *4* *5* *6* *7* *Agree*

It was easy to acquire the feedback from the audience.

Disagree *1* *2* *3* *4* *5* *6* *7* *Agree*

Open-ended questions

What was good about this system?

Comments.

C.4 Additional Questionnaire - Comparison

The following questionnaire was used to measure preference between AR and Non-AR versions of Meetsu through forced-choice questions from English-speaking subjects:

interactive
Media Design Lab
Graduate School Of Information Science

#3 Questionnaire (Comparison)

<u>Demographics</u>

Username:

Which system was more distracting?

AR ☐ *Non-AR* ☐

Which system made you more anxious?

AR ☐ *Non-AR* ☐

Which system made you more willing to participate?

AR ☐ *Non-AR* ☐

Which system made you engage more in real-world conversation?

AR ☐ *Non-AR* ☐

D. Meetsu Questionnaires (Japanese)

D.1 WTC Questionnaire

The following questionnaire was used to measure Willingness to Communicate from Japanese-speaking subjects:

interactive Media Design Lab
Graduate School Of Information Science

第1回アンケート

氏名::________________________　　日付:________________________

以下は人がコミュニケーションを行うか行わないかを決める20のシチュエーションです．あなたは完全に自由な選択権をもっていることを仮定してください．各シチュエーションにおいて，あなたがその選択を選ぶであろうパーセンテージを示してください．あなたがコミュニケーションするパーセンテージを左側の空欄に記入してください．

(0 = 決してしない 〜　100 = 常にする)

_____1. ガソリンスタンドの店員と話す．
_____2. 医師と話す．
_____3. 見知らぬ人ばかりのグループで発表する．
_____4. 列に並んでいるとき,知り合いと会話する．
_____5. 店で販売員と会話する．
_____6. 友人ばかりの大きなグループの中で話す．
_____7. 警察官と話す．
_____8. 知り合いでない人々の小さなグループの中で話す．
_____9. 列に並んでいるとき,友人と会話する．
_____10. レストランでウェイターやウェイトレスと会話する．
_____11. 知り合いばかりの大きなグループの中で話す ．
_____12. 列に並んでいるとき,知り合いでない人と話す．
_____13. 秘書さんと話す．
_____14. 友人ばかりのグループで発表する．
_____15. 知り合いばかりの小さなグループの中で話す．
_____16. ごみ収集業者の人と話す．
_____17. 見知らぬ人ばかりの大きなグループの中で話す．
_____18. 妻(もしくは夫)と話す(いなければ彼氏,彼女)．
_____19. 友人ばかりの小さなグループの中で話す．
_____20. 知り合いの人ばかりのグループで発表する．

D.2 Additional Questionnaire - Audience Type

The following questionnaire was used to evaluate additional measures from Japanese-speaking meeting attendees in both AR and Non-AR conditions:

interactive
Media Design Lab
Graduate School Of Information Science

Meetsu アンケート(聴衆タイプ)

基本情報調査

氏名:　　　　年齢:　　　　性別: 男性　女性

ユーザネーム:

以前に AR 技術を体験したことがありますか　はい　いいえ

以前にミーティング支援システムを使用したことがありますか　はい　いいえ

本日どのくらいの頻度で Meetsu を使用しましたか?

1 回も使用していない　1 回　2-3 回　3 回以上

選択形式の質問

システムに気を散らされた.

反対　*1*　*2*　*3*　*4*　*5*　*6*　*7*　同意

システムによってミーティング参加に不安を感じた.

反対　*1*　*2*　*3*　*4*　*5*　*6*　*7*　同意

システムのおかげでミーティングに進んで参加できた.

反対　*1*　*2*　*3*　*4*　*5*　*6*　*7*　同意

システムのおかげで他の参加者との実環境での会話に集中できた.

反対　*1*　*2*　*3*　*4*　*5*　*6*　*7*　同意

自由記述形式の質問

システムの良いところはどこでしたか.

他になにかコメントがあればお書きください.、

D.3 Additional Questionnaire - Speaker Type

The following questionnaire was used to measure distraction, anxiety, willingness to communicate and engaging in real world conversation from Japanese-speaking meeting speakers in both AR and Non-AR conditions:

interactive
Media Design Lab
Graduate School Of Information Science

Meetsu アンケート (話者タイプ)

基本情報調査
氏名:
年齢:
性別: *男性*　*女性*
ユーザネーム:
以前に AR 技術を体験したことがありますか　*はい*　*いいえ*
以前にミーティング支援システムを使用したことがありますか　*はい*　*いいえ*
本日どのくらいの頻度で Meetsu を使用しましたか?
1 回も使用していない　*1 回*　*2-3 回*　*3 回以上*

選択形式の質問
システムに気を散らされた.
反対　*1*　*2*　*3*　*4*　*5*　*6*　*7*　同意
聴衆からフィードバックを得るのが容易だった.
反対　*1*　*2*　*3*　*4*　*5*　*6*　*7*　同意

自由記述形式の質問
システムの良いところはどこでしたか.

他になにかコメントがあればお書きください.、

D.4 Additional Questionnaire - Comparison

The following questionnaire was used to measure preference between AR and Non-AR versions of Meetsu through forced-choice questions from Japanese-speaking subjects:

interactive Media Design Lab
Graduate School Of Information Science

Meetsu アンケート（比較）

<u>基本情報調査</u>
ユーザネーム：

<u>選択形式の質問</u>
どちらのシステムを用いた際に、より気が散らされましたか。

AR ☐ *Non-AR* ☐

どちらのシステムを用いた際に、より不安な気持ちになりましたか。

AR ☐ *Non-AR* ☐

どちらのシステムが、よりミーティングに参加したいという気持ちにさせてくれましたか。

AR ☐ *Non-AR* ☐

どちらのシステムが、実世界での会話を促進させましたか。

AR ☐ *Non-AR* ☐

References

[1] J. Koyen. Business meetings - the case for face-to-face. http://www.forbes.com/forbesinsights/Business_Meetings_FaceToFace/ (Last accessed: December 7, 2013).

[2] H. Alvarez, I. Aguinaga, and D. Borro. Providing guidance for maintenance operations using automatic markerless augmented reality system. In *Proceedings of the 10th IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR11)*, pages 181–190, 2011.

[3] S. J. Henderson and S. Feiner. Evaluating the benefits of augmented reality for task localization in maintenance of an armored personnel carrier turret. In *Proceedings of The 8th IEEE International Symposium on Mixed and Augmented Reality (ISMAR09)*, pages 135–144, 2009.

[4] I. Barakonyi, T. Psik, and D. Schmalstieg. Agents that talk and hit back: Animated agents in augmented reality. In *Proceedings of The 3rd IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR04)*, pages 141–150, 2004.

[5] J. Chandaria, G. A. Thomas, and D. Stricker. The matris project: real-time markerless camera tracking for augmented reality and broadcast applications. *Journal of Real-Time Image Processing*, 2(2-3):69–79, 2007.

[6] G. A. Thomas. Mixed reality techniques for tv and their application for on-set and pre-visualization in film production. In *Proceedings of The International Workshop on Mixed Reality Technology for Filmmaking*, October 2006.

[7] Sony. Eyepet. http://www.eyepet.com/ (Last accessed: November 28, 2013).

[8] Zugara. Webcam social shopper. http://webcamsocialshopper.com/ (Last accessed: November 29, 2013).

[9] M. Billinghurst and H. Kato. Real world teleconferencing. In *Proceedings of Extended Abstracts on Human Factors in Computing Systems (CHI99)*, CHI EA '99, pages 194–195, New York, NY, USA, 1999. ACM.

[10] Michael Zoellner. Twittaround. http://i.document.m05.de/twittaround/ (Last accessed: November 30, 2013).

[11] Microsoft Research. Holoflector. http://research.microsoft.com/apps/video/default.aspx?id=159487 (Last accessed: November 29, 2013).

[12] Zugara. Zugstar. http://corporate.zugara.com/augmented-reality/video-conferencing (Last accessed: November 29, 2013).

[13] M. A. Livingston. *Human Factors in Augmented Reality Environments*, chapter Issues in Human Factors Evaluations of Augmented Reality Systems, pages 3–9. Springer, 2013.

[14] I. Barakonyi, T. Fahmy, and D. Schmalsteig. Remote collaboration using augmented reality videoconferencing. In *Proceedings of Graphics Interface*, pages 89–96, 2004.

[15] D. Schmalsteig, A. Fuhrmann, Z. Szalavari, and M. Gervautz. Studierstube - an environment for collaboration in augmented reality. Technical report, Vienna University of Technology, January 1997.

[16] W. Gudykunst. *Theorizing About Intercultural Communication*, volume 1, chapter An Anxiety/Uncertainty Management Theory of Effective Communication, pages 281–322. Sage, 2004.

[17] M. P. McIntosh and R. A. Luecke. Interpersonal communication skills in the workplace. http://www.flexstudy.com/catalog/schpdf.cfm?coursenum=97003 (Last accessed: December 7, 2013).

[18] J. Hauber. Understanding remote collaboration in video collaborative virtual environments. Master's thesis, University of Canterbury, Christchurch, New Zealand, 2008.

[19] J. Galegher, R. E. Kraut, and C. Egido. *Intellectual Teamwork: Social and Technological Foundations of Cooperative Work*, chapter Teleconferencing as a Technology to Support Cooperative Work: Its Possibilities and Limitations, pages 351–372. Lawrence ErlBaum Associates, 1990.

[20] S. D. Benford, C. C. Brown, G. T. Reynard, and C. M. Greenhalgh. Shared spaces: Transportation, artificiality, and spatiality. In *Proceedings of The ACM Conference on Computer-Supported Cooperative Work (CSCW1996)*, pages 77–86, 1996.

[21] M. Billinghurst, I. Poupyrev, H. Kato, and R. May. Mixing realities in shared space: An augmented reality interface for collaborative computing. In *Proceedings of The IEEE International Conference on Multimedia and Expo (ICME2000)*, volume 3, pages 1641–1644, 2000.

[22] J. S. Olson, G. M. Olson, and D. K. Meader. What mix of video and audio is useful for small groups doing remote realtime design work? In *Proceedings of The ACM Special Interest Group on Computer-Human Interaction (SIGCHI1995)*, pages 362–368, 1995.

[23] S. Whittaker, E. Geelhoed, and E. Robinson. Shared workspaces: how do they work and when are they useful? *International Journal of Man-Machine Studies*, 39(5):813–842, November 1993.

[24] O. Morikawa and T. Maesako. Hypermirror: toward pleasant-to-use video mediated communication system. In *Proceedings of The ACM Conference on Computer-Supported Cooperative Work (CSCW1998)*, pages 149–158, 1998.

[25] O. Morikawa. Hypermirror: Toward pleasant-to-use video mediated communication system. http://staff.aist.go.jp/morikawa.osamu/cscw/cscw98.htm (Last accessed: December 2, 2013).

[26] C. N. Gunawardena. Social presence theory and implications for interaction and collaborative learning in computer conferences. *International Journal of Educational Telecommunications*, 1(2):147–166, 1995.

[27] H. Kuzuoka, T. Kosuge, and M. Tanaka. Gesturecam: Platform for augmented reality based collaboration. In *Proceedings of The 4th International Conference on Artificial Reality and Tele-existence (ICAT1994)*, pages 165–180, 1994.

[28] T. B. Faas. An examination of social presence in video conferencing vs. an augmented reality conferencing application. Master's thesis, Purdue University, 2010.

[29] J. Short, E. Williams, and B. Christie. *The Social Psychology of Telecommunications*. John Wiley & Sons Ltd, 1976.

[30] E. Sallnas. Effects of communication mode on social presence, virtual presence and performance in collaborative virtual environments. *Presence: Teleoperators and Virtual Environments*, 14(4):434–449, August 2005.

[31] E. Bradner and G. Mark. Social presence with video and application sharing. In *Proceedings of The International ACM SIGGROUP Conference on Supporting Group Work (GROUP2001)*, pages 154–161, 2001.

[32] CiTrix. Gotomeeting. http://www.gotomeeting.com/online/ (Last accessed: November 28, 2013).

[33] Trimble Navigation. Sketchup. http://www.sketchup.com/ (Last accessed: December 3, 2013).

[34] J. M. Bowers. *Studies in Computer-Supported Cooperative Work - theory, practice and design*. North-Holland, 1991.

[35] M. Koch and T. Gross. Computer-supported cooperative work - concepts and trends. In *Proceedings of The 11th International Conference of the Association Information and Management (AIM2006)*, 2006.

[36] C. A. Ellis, S. J. Gibbs, and G. Rein. Grougroup - some issues and experiences. *Communications of the ACM*, 34(1):39–58, January 1991.

[37] J. Rama and J. Bishop. A survey and comparison of cscw groupware applications. In *Proceedings of the The Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists on IT Research in Developing Countries (SAICSIT2006)*, pages 198–205, 2006.

[38] M. Koch. Cscw and enterprise 2.0 - towards an integrated perspective. In *Proceedings of The 21st Conference Bled eConference*, pages 416–427, 2008.

[39] J. M. Ng, E. Chan, H. H. S. Ip, K. Y. Kwok, Y. K. Lee, and H. H. Tsang. A multimedia conferencing system for co-operative medical diagnosis. In *Proceedings of The 6th Annual IEEE Symposium on Computer-Based Medical Systems (CBMS1993)*, pages 94–99, 1993.

[40] P. Zeng, Y. Hao, W. Shao, and Y. Liu. Towards a network video conference system based-on group cooperation mode for multidisciplinary collaborative product design. In C. Xiong, H. Liu, Y. Huang, and Y. Xiong, editors, *Intelligent Robotics and Applications*, volume 5315 of *Lecture Notes in Computer Science*, pages 311–320. Springer Berlin Heidelberg, 2008.

[41] A. R. Dennis, J. F. George, L. M. Jessup, J. F. Nunamaker Jr., and D. R. Vogel. Information technology to support electronic meetings. *Management Information Systems Quarterly*, 12(4):591–624, December 1988.

[42] L. Chidambaram and B. Jones. Impact of communication medium and computer support on group perceptions and performance - a comparison of face-to-face and dispersed meetings. *MIS Quarterly*, 17(4):465–491, December 1993.

[43] A. Etzioni and O. Etzioni. Face-to-face and computer-mediated communities - a comparative analysis. *The Information Society*, 15(4):241–248, October-December 1999.

[44] J. S. Valacich, S. Sarker, J. Pratt, and M. Groomer. Computer-mediated and face-to-face groups: Who makes riskier decisions? In *Proceedings of The 35th Hawaii International Conference on System Sciences (HICSS2002)*, pages 133–142, 2002.

[45] K. Burke and L. Chidambaram. Do mediated contexts differ in information richness - a comparison of collocated and dispersed meetings. In *Proceedings of The 29th Annual Hawaii International Conference on System Sciences*, pages 92–101, 1996.

[46] L. Branson, T. S. Clausen, and C. Sung. Group style differences between virtual and f2f teams. *American Journal of Business*, 23(1):65–70, 2008.

[47] H. P. Andres. A comparison of face-to-face and virtual software development teams. *Team Performance Management: An International Journal*, 8(1/2):39–48, 2002.

[48] W. A. Hatem, A. S. Kwan, and J. C. Miles. A comparison of face to face and computer mediated collaboration. In *Proceedings of The eg-ice Workshop on Intelligent Computing in Engineering*, July 2011.

[49] R. Anson and B. E. Munkvold. Beyond face-to-face - a field study of electronic meetings in different time and place modes. *Journal of Organizational Computing and Electronic Commerce*, 14(2):127–152, 2004.

[50] M. Billinghurst, H. Kato, K. Kiyokawa, D. Belcher, and I. Poupyrev. Experiments with face-to-face collaborative ar interfaces. *Virtual Reality*, 6(3):107–121, October 2002.

[51] L.L. van der Leeden. The influence of gaze direction on social presence of the remote participant. In *Proceedings of The 13th Twente Student Conference on IT*, volume 13, 2010.

[52] M. Ashdown and P. Robinson. The escritoire - remote collaboration in a task space. In *Proceedings of The ACM SIGMM Workship on Experiental Telepresence*, pages 73–75, 2003.

[53] R. D. Arvey. Why face-to-face business meetings matter. White paper, National University of Singapore, 2009.

[54] R. Jain, P. Kim, and Z. Li. Experiential meeting system. In *Proceedings of The ACM SIGMM Workshop on Experiential Telepresence (ETP2003)*, pages 1–12, 2003.

[55] G. Zurita and N. Baloian. Handheld-based electronic meeting support. In *Proceedings of The 11th International Workshop on Groupware (CRIWG2005)*, pages 341–350, 2005.

[56] D. A. Mejia, A. L. Moran, and J. Favela. Supporting informal co-located collaboration in hospital work. In *Proceedings of The 13th International Workshop on Groupware (CRIWG2007)*, pages 255–270, 2007.

[57] K. Kiyokawa. *Emerging Technologies of Augmented Reality - Interfaces and Design*, chapter Human Communication in collaborative augmented reality systems, pages 236–260. Idea Group Publishing, 2007.

[58] A. Butz, T. Hollerer, S. Feiner, B. MacIntyre, and C. Beshers. Enveloping users and computers in a collaborative 3d augmented reality. In *Proceedings of The 2nd IEEE and ACM International Workshop on Augmented Reality (IWAR1999),*, pages 35–44, 1999.

[59] C. Coffin, S. Bostandjiev, J. Ford, and T. Hollerer. Enhancing classroom and distance lealearn through augmented reality. In *Proceedings of The World Conference on Educational Multimedia, Hypermedia and Teelcommunications (Ed-Media2010)*, pages 1140–1147, 2010.

[60] J. Rekimoto and Y. Ayatsuka. Adding another communication channel to reality - an experience with a chat-augmented conference. In *Proceedings of The 16th Conference on Human Factors in Computing Systems (CHI1998)*, pages 271–272, 1998.

[61] J. Wither, S. DiVerdi, and T. Hollerer. Annotation in outdoor augmented reality. *Computers & Graphics*, 33(6):679–689, June 2009.

[62] A. Montero, T. Zarraonandia, I. Aedo, and P. Diaz. Uses of augmented reality for supporting educational presentations. In *Proceedings of The 13th IEEE Internationak Conference on Advanced Learning Technologies (ICALT2013)*, pages 426–428, July 2013.

[63] A. Hakeem and M. Shah. Ontology and taxonomy collaborated framework for meeting classification. In *Proceedings of The 17th International Conference on Pattern Recognition (ICPR 2004)*, volume 4, pages 219–222, 2004.

[64] J. M. Davis. Planning and leading productive meetings. http://www.flexstudy.com/catalog/schpdf.cfm?coursenum=96026 (Last accessed: December 7, 2013).

[65] N. C. Romano and J. F. Nunamaker Jr. Meeting analysis - findings from research and practice. In *Proceedings of the 34th Annual Hawaii International Conference on System Sciences (HICSS2001)*, volume 9, 2001.

[66] J. Schlegel. Running effective meetings - types of meetings. http://www.salary.com/running-effective-meetings-6/ (Last accessed: Decemebr 8, 2013).

[67] E. Prytz, S. Nilsson, and A. Jonsson. The importance of eye-contact for collaboration in ar systems. In *Proceedings of The 9th IEEE International Symposium on Mixed and Augmented Reality (ISMAR2010)*, pages 119–126, October 2010.

[68] S. Morreale. Assessing motivation to communicate: willingness to communicate and personal report of communication apprehension. Technical report, National Communication Association, Washington D.C., USA, 2007.

[69] W. Karwowski. *Handbook of Human Factors and Ergonomics*, chapter The Discipline of Human Factors and Ergonomics, pages 3–37. John Wiley & Sons, Inc., 4 edition, 2012.

[70] Gavriel Salvendy. *Handbook of Human Factors and Er.* John Wiley & Sons Inc., 4 edition, 2012.

[71] W. Huang, L. Alem, and M. A. Livingston, editors. *Human Factors in Augmented reality Environments*. Springer, 2013.

[72] M. Nakanishi. *Augmented Reality*, chapter Human Factor Guideline for Applying AR-based Manuals in Industry, pages 129–156. Intech, 2010.

[73] M. A. Livingston. Evaluating human factors in augmented reality systems. *IEEE Computer Graphics and Applications*, 25(6):6–9, November 2005.

[74] K. Nowak. Defining and differentiating copresence, social presence and presence as transportation. In *Proceedings of The HCI International Conference on Human-Computer Interaction*, pages 686–690, 2001.

[75] E. Sung and R. E. Mayer. Five facets of social presence in online distance education. *Computers in Human Behavior*, 28(5):1738–1747, September 2012.

[76] A. Tang, F. Biocca, and L. Lim. Comparing differences in presence during social interaction in augmented reality versus virtual reality environments

- an exploratory study. In *Proceedings of The 7th Annual International Workshop on Presence*, pages 204–208, 2004.

[77] K. Kreijns, P. A. Kirschner, W. Jochems, and H. van Buuren. Measuring perceived social presence in distributed learning groups. *Education and Information Technologies*, 16(4):365–381, December 2011.

[78] C. J. Knudsen. Video mediated communication (vmc) - producing a sense of presence between individuals in a shared virtual reality. In *Proceedings of The International Symposium on Educational Conferencing (ISEC 2002)*, 2002.

[79] R. M. Ryan and E. L. Deci. Intrinsic and extrinsic motivations: Classic definitions and new directions;. *Contemporary Educational Psychology*, 25(1):54–67, 2000.

[80] J. C. McCroskey. Reliability and validity of the willingness to communicate scale. *Communication Quarterly*, 40(1):16–25, 1992.

[81] H. Reinders and S. Wattana. Learn english or die: The effects of digital games on interaction and willingness to communicate in a foreign language. *Digital Culture & Education*, 3(1):3–29, April 2011.

[82] itseez. Opencv. http://opencv.org/, 1999.

[83] Qt digia. Qt. http://qt.digia.com/ (Last accessed: December 3, 2013).

[84] R. Azuma. A survey of augmented reality. *Presence: Teleoperators and Virtual Environments*, 6(4):355–385, August 1997.

[85] J. Carmigniani and B. Furht. *Handbook of Augmented Reality*, chapter Augmented Reality: An Overview, page 3. Springer, 2011.

[86] D. L. Jorgensen. *Participant Observation: A Methodology for Human Studies (Applied Social Research Methods)*. Number 15 in Applied Social Research Methods. SAGE Publications, 1989.

[87] A. Dunser and E. Hornecker. An observational study of children interacting with an augmented story book. In *In Proceedings of The 2nd International Conference of E-Learning and Games (Edutainment 2007)*, pages 305–315, 2007.

[88] M. Prasad and T. A. Hammond. Observational study on teaching artifacts created using tablet pc. In *Proceedings of The 30th Conference on Human Factors in Computing Systems (CHI2012)*, pages 301–316, 2012.

[89] C. Tang and S. Carpendale. An observational study on information flow during nurses' shift change. In *Proceedings of The 25th Conference on Human Factors in Computing Systems (CHI2007)*, pages 219–228, 2007.

[90] P. Bordia. Face-to-face versus computer-mediated communlgation - a synthesis of the experimental literature. *The Journal of Business Communication*, 34(1):99–120, January 1997.

[91] R. J. Garmston and V. A. von Frank. *Unlocking Group Potential do Improve Schools*, chapter Building Effective Groups, pages 15–23. Corwin, 2012.