

Mutual awareness in co-localized and distant collaborative tasks using shared interfaces

Paper ID: 11989

Abstract. Shared interface allowing several users in co-presence to interact simultaneously on digital data on a single display is an uprising challenge in Human Computer Interaction. Its development is motivated by the advent of large displays such as wall screens and tabletops. It affords fluid and natural digital interaction without hindering human communication and collaboration. It enables mutual awareness, making participant conscious of each other activities.

In this paper, we are interested in Mixed Presence Groupware (MPG), when two or more remote shared interfaces are connected for distant collaboration. Our contribution strives to answer to the question: Can the actual technology convey sufficient context's elements of the remote site to enable efficient collaboration between two distant groups?

We propose HERMES, an experimental platform we hope lessen the gap between co-present and distant interaction. HERMES is combining a multiuser tactile interactive tabletop, a video-communication system enabling eye-contact with real size distant user visualization and a spatialized sound system for speech transmission. A robust computer vision module for distant users' gesture visualization completes the platform. We discuss first experiments using HERMES for a collaborative task (mosaic completion) in term of distant mutual awareness. Although HERMES does not provide the same presence feeling in distant and or co-localized situation, a first and important finding emerges: distance does not hinder efficient collaboration anymore.

1 Introduction

Shared interfaces allowing multiuser interaction in co-presence on a single device is a challenging Human Computer Interaction (HCI) research field. Though, from the Xerox Parc Colab project in 1986 dedicated to informal group meetings [1][2], the domain produces relatively few literature in comparison to distant personal workstation based collaborative systems.

The relevance of shared interface has been yet precisely identified. Shared interface is complementary to common personal digital devices, such as Personal Computer (PC), Personal Digital Assistant (PDA)... and Personal Phone. It does not hinder interaction between people when accessing to the digital world. It conveys more conviviality and does not draw the attention of the participants from their interaction. In contrast, the one-person/one-computer paradigm of personal interfaces *de facto* tends to impede direct human communication whenever data interaction is needed: a user has to focus his attention on his personal device

to find data, whereas his partner is waiting... to find a place beside him to see the sought information. With shared interfaces which use large displays, such as scene walls or tabletops, data are reachable by every participant at anytime. No disruptive turn taking occurs anymore. Interactions are explicit, visible (hand gesture, pen pointing) and thus understandable by everyone: shared interfaces are in line with day-to-day people interaction. They even help bridging the gap between physical world (human to human interaction) and digital world (human to computer and human to computer to human interaction).

We are especially interested in distant collaboration between groups, and the so called Mixed Presence Groupware (MPG) [3], connecting two (or more) distant shared interfaces. The creative domain such as co-design, architecture, urbanism, etc. is particularly concerned by MPG: it often involves remote teams manipulating graphical objects (images, drawings, plans, etc.) for which shared interfaces, such as tabletops, propose adapted natural interactions (pen based, tactile or gesture interactions).

However, distant collaboration has to preserve as far as possible the fluidity and the mutual awareness provided by co-presence. Following Tang *et al.* [3], we focus on remote gesture visualization of distant users, as it conveys major information facilitating distant communication such as intentionality (who is intending to do what), action identity (who is doing what) and pointing [4][5].

We propose HERMES, a platform combining a multiuser tactile interactive tabletop, a video-communication system enabling eye-contact with real size distant user visualization and a spatialized sound system. A robust computer vision module for distant users' gesture visualization completes the set-up.

The main question this paper strives to answer is: may the collaboration between two distant groups be as efficient as the co-present collaboration, when the technical tools conveying elements of the distant context (real size distant user video, remote gesture visualization) are now available and improved? We present and discuss a collaborative application designed to investigate collaboration efficiency and presence feeling in distant and co-localized situation.

The remain of this paper is organized as follows: Section 2 presents a state of the art on shared interfaces, distant collaboration and mutual awareness. In Section 3 we specify our objectives. Section 4 describes the HERMES platform whereas Section 5 details the experiment. Findings are resumed in Section 6. Conclusion and future work are in section 7.

2 Related work

The development of shared interfaces relies on the advent of new display devices, such as large wall screens or tabletops, and on the development of platforms assuming multiple independent inputs. Thus, in a first attempt to address multiuser application on a single display in co-presence, the Xerox Parc Colab Project (1987-1992)[1] proposed a PC network allowing private work as well as control of a shared digital white board. In 1993, Pederson *et al.* [2] extended the Xerox Liveboard concept and proposed Tivoli, an electronic white board application for supporting informal workgroup meetings, using a large screen pen-based

interactive display allowing up to three simultaneous pen interactions. Tivoli strove to provide its users with simplicity of use, easily understood functionalities, and access to computational power to augment informal meeting practices. However, multiuser interaction was not the main focus of [1] [2].

In 1991, also at Xerox Parc, Bier *et al.* [6] developed MMM, a multiuser editor managing up to three mice simultaneously on the same computer. In 1998, Greenberg *et al.* [7] proposed a Shared Notes system investigating how people move from individual to workgroup through the use of both personal digital assistants (PDAs) and a shared public display. Later in 1999, Stewart *et al.* [8] defined the Single Display Groupware concept (SDG), which stemmed from their work on children groups. Their experiments on a shared drawing application on PC showed that multi-mouse implementation was largely preferred by children to single mouse version because it provides more fun and more activity. Hourcade *et al.* [9] proposed a Multiple Input Devices (MID) offering a Java toolkit for simultaneous use of multiple independent devices for Windows98. Tse *et al.* [10] extended the idea and proposed a toolkit for rapidly SDG prototyping on Microsoft.Net. Microsoft Research India renews the challenge on multi-mouse interaction for computers to respond to educational needs in rural primary schools in developing countries where very few computers are available per student (one PC for ten students) [11]. The authors experiment children dedicated application allowing up to 5 or more simultaneous mice used.

The Calgary University GroupLab has investigated groupware interaction for long time. Gutwin [12] analyzed workspace awareness for distributed groupware. In line with Gutwin, Tang *et al.* [13] extended the SDG concept to MPG, connection of two or more SDG at distant places for remote and co-localized collaboration. The authors focused on presence disparity, describing that people do not interact similarly with their co-localized and distant partners. They proposed and analyzed distant groupware embodiments, such as telepointers or remote gesture visualization in distant communication and collaboration [3].

We focus on tabletop displays as a horizontal surface encouraging group members to work in a socially cohesive and conducive way. It affords seamless role-changing and more equitable decision-making and information access [14]. TableTop, as a shared input/output device, is an emerging interdisciplinary domain involving augmented reality, user interface technologies, multi-modal and multi-user interaction, CSCW, and information visualization (see [15]).

In 2001, Dietz *et al.* [16] presented MERL Diamond Touch, a multi-user touch table for which each tactile interaction is user identified. Enabling up to 8 user simultaneous interactions, bi-manual or multi-touch interaction per user is however hindered by the hardware, which does not provide the loci of the contacts, but the bounding box encompassing these user's contacts. In contrast, Rekimoto's Smartskin [17] is a real multi-touch device, but without user identification. Note that several multi-touch devices have emerged: Bérard's Magic Table [18], Wilson's TouchLight [19], the low-cost Multi-touch sensing surface of Han [20] and Philips's Entertaible [21] among numerous examples. By now, the Diamond Touch is the only commercially available multitouch tabletop.

3 Motivations

We want to address distant groupware collaboration and to design a platform which preserves most of the characteristics of face-to-face interaction. We base our approach on Gutwin’s workspace awareness analysis [12] who organized previous works of Endsley on situation awareness [22], Segal on consequential communication [23] and Clark *et al.* on conversational grounding [24] among others.

Situational awareness is knowledge of a dynamic environment, it is maintained through perceptual information gathered from the environment and others’ activities. It is peripheral to the primary group activity [12]. It depends on perception, comprehension and prediction of the environment and of others’ actions [22]. It relies on non intentional informational sources such as *consequential communication*, *artifacts manipulation*, and on *intentional communication*. *Consequential communication* is information that emerges from a person’s activity [23]. It is non intentional and most of it is conveyed by the visual channel: position, posture, head, arms, hands’ movement, *etc.* *Artifacts* are a second source of information about the current actions in progress because their characteristic sound depending on the action (moving, stacking, dividing, and so on) gives salient feedbacks of their use. Finally, *intentional communication*, through conversation and gestures (deictic, manifesting or visual evidence actions), completes the perceptual information gathered about the environment. Situational awareness is the way people maintain up-to-date mental models of complex and dynamic environments. It helps to be aware of the state of task objects and of one another’s activities. It facilitates to plan what to say or to do and to coordinate speech and actions.

Visual information helps also people to communicate about the task to be done by ensuring that their message is properly understood. It enables the expansion of their common ground during the session and facilitates mutual understanding between participants. There is significantly less talk about the talk, or about the task to be done [24].

We aim at conceiving a platform which preserves at most situational awareness for remote groupware collaboration. We choose tabletops first because it is a now available shared interface, and above all because it is probably the most common tool used for group meetings and human interaction. Interaction and communication around a table are observed to be more equally distributed between participants in contrast to white boards which often induce role disparity as the person at the board is given to be the leader’s meeting [14].

To facilitate intentional communication and presence feeling, we choose to use a video-communication system providing real size visualization of the distant users and eye-contact by means of a spy camera (see section 4 for details). As Tang *et al.* [3][13], we add a computer vision module allowing to capture the local gesture of the participant on or above the table and to transmit it to the distant site for overlay on the distant desktop image. This remote gesture visualization module, similar but probably more robust than VideoArms [13] (see section 4), combined with the video-communication system, participates in conveying most of the visual information needed to feed situational awareness.

Thus, HERMES, the designed platform, should provide people with coordination, as participants can see each other through video communication and each others' action through remote gesture visualization on the tabletop. It should grant action identity as participants are able to perceive who is doing what. They also can anticipate and predict which action the distant participants intend to do, and which digital object they are about to grasp as they see arms and hands above the table at the distant site. Most of the common social rules are therefore preserved, as involuntary conflicts about availability of objects are avoided.

Finally, intentional communication is partially enhanced as participant also can point to digital object to show something or explain an action or an idea.

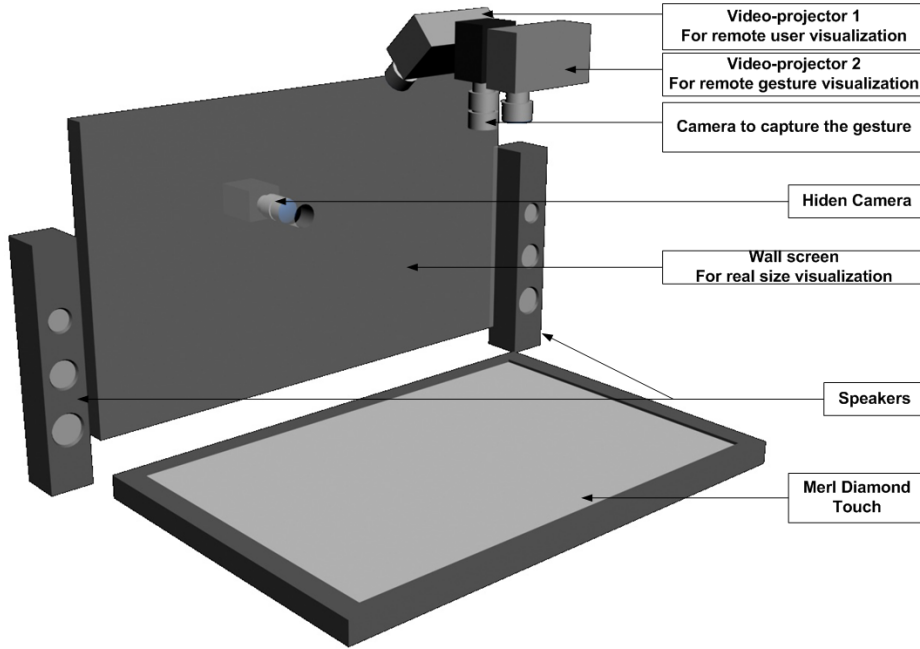


Fig. 1. HERMES is a platform combining a Diamond Touch, a video communication system, a spatialized audio system and a computer vision module.

4 HERMES platform

HERMES is a platform combining a multiuser tactile interactive tabletop, a video-communication system enabling eye-contact with real size distant user visualization, a computer vision module for remote gesture visualization of the distant users and a spatialized sound system. (see Fig. 1 and Fig. 2).

We use Merl Diamond Touch [16], which is hitherto the only available shared tabletop for simultaneous multiuser interaction. This device is a passive tactile surface on which the desktop image is projected from a ceiling mounted video-projector (video-projector 2 of Fig. 1).

The video communication system uses a spy camera hidden behind a rigid wood-screen and peeping through an almost unnoticeable 3mm wide hole. A second video-projector (video-projector 1 of Fig. 1) beams on the wall screen the video of the distant site captured by the symmetric remote spy camera (see Fig. 2). Eye-contact is guaranteed by approximately placing the camera peephole at the estimated eyes'height of a seated person and beaming the distant video on the screen such that the peephole and the distant user's eyes coincide. Fine tuning of hole's design and video-projector beam's orientation is performed to avoid camera's dazzle.

The computer vision module uses a camera placed at the ceiling and pointing at the tabletop. The module consists of a segmentation process detecting any object above the table by comparing, at almost the frame rate, between the captured camera image and the actual known desktop image projected on the tabletop, up to a geometric and color distortion. In output, it produces an image mask of the detected objects (hands, arms, or any object) extracted from the camera image. The mask is compressed using RLE (Run Length Encoding) and is sent through the network to the distant site. There, the image mask is decompressed and then overlaid with the current desktop image before projection on the tabletop. We use semi-transparency to let the user see the desktop "under" the arms of his/her distant partner.

In the current version of HERMES (see Fig. 2), the video system is separated from the rest of the architecture and runs with direct cable connexion, to avoid lag problems due to network. HERMES also support a network-based video-conference system.

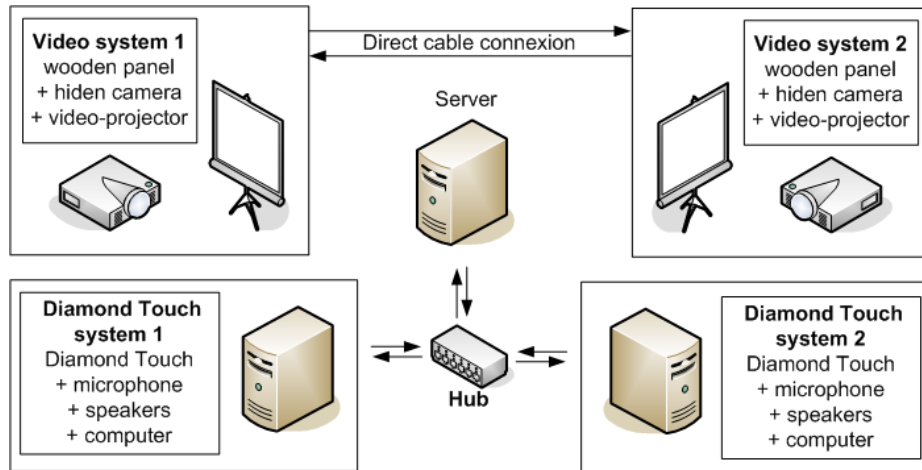


Fig. 2. HERMES Architecture: shared application and remote gesture analysis is implemented at each distant site on a single work station; video communication is running on a separate network.

Fig. 3 shows users' gesture and their visualization on the remote site. The left bottom image shows the overlay of the detected hand of site 2 (upper right

image); the right bottom image shows the overlay of the detected arms of the site 1 (upper left image).

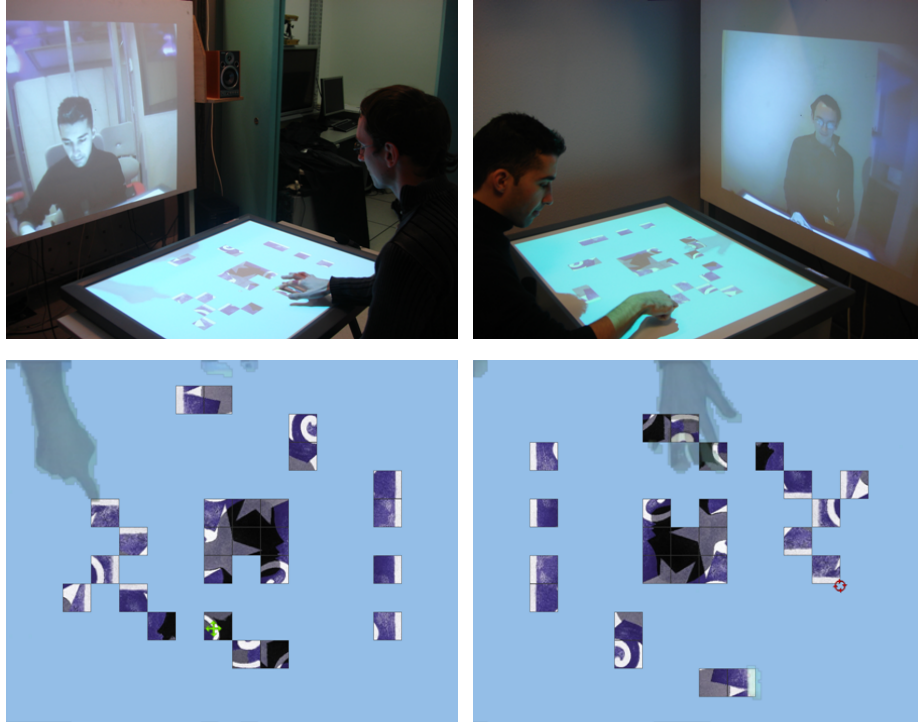


Fig. 3. Remote gesture visualization: view of both distant tables in action (upper line) and view of both overlaid desktops (lower line). The application presented in the pictures concerns mosaic completions.

The computer vision module improves four technical weaknesses of VideoArms [3]. 1) it detects any object on or over the table without needing any learning stage or *a priori* knowledge about the object to detect; 2) no restriction on the projected images is imposed (videoArms needs dark tones images); 3) it is robust to external lightning changes (variation in the daylight or in the artificial lightning); 4) calibration is automatic. The computer vision algorithm provides 32 image masks per second when running alone for a SXGA (1280x1024) image desktop and a medium camera resolution (384x288) (we use a Sony EVI-D70 whose pan, tilt and zoom functions facilitates the camera view control). When combined with the shared Java application, the image masks are refreshed on each site at between 12 and 17 Hz on a dual-core Intel Xeon 3.73GHz (Netburst architecture) with 2GB of RAM.

As camera capture and desktop image are not synchronized, a delay may occur and cause some echoes on the image mask detection, which is particularly obvious when the computer vision frame rate drops low (under 14hz).

5 User study

We aim at investigating how distance affects interaction and collaboration when most of visual information needed for mutual awareness is provided. We conceived a digital mosaic completion task as experimental application. Collaborative mosaic completion performed by two users in both co-localized or distant situations are compared, in order to understand the effects of distance in term of user experience and objective criteria such as activity progression, resource access and use, awareness, coordination, and collaboration. We are also interested by the impact of object orientation on the coordination according to the collaborative configuration (side-by-side or face-to-face).

In comparison with puzzles, mosaics are composed of squared pieces. On-table textual puzzle completions has been investigated by Kruger *et al.* [25] who observed three roles of the piece orientation: *understanding* (ex: reading), *coordination* (an implicit private space is created when a piece is oriented toward a particular user) and *communication* (voluntary orientation of a piece toward a user is used to raise his attention). Their main finding was that users significantly touch more often pieces oriented toward themselves than those oriented toward another user. Pieces perpendicularly oriented are considered as public. The authors focused mainly on document orientation but not on the document localization. We are interested in extending Kruger *et al.*'s results to mosaic completion in co-localized and distant digital situation. We think that the localization of the pieces on the table may also contribute to design implicit private areas, as Kruger *et al.* showed for piece orientation. Finally, we aim at studying the influence of the image type of the mosaic on the completion processes.

In the co-localized situation, the users are sitting side-by-side in front of the table (Diamond Touch) as this configuration seems the most natural. In remote configuration, the users are virtually sitting face-to-face, on both sides of the table, using the HERMES platform. We thought that this arrangement was more convenient to distant communication as it was compatible to video-communication. This follows Tang recommendation [26], as face-to-face collaboration is more comfortable for verbal and non-verbal interaction. Furthermore, it was compliant with the situated-literal approach suggested by Gutwin's [12] for shared workspace conception: the situated-literal configuration consists of displaying distant visual awareness information *literally* (not symbolic) at the workspace place where it is originated from. It is in line with the way people use their existing skills with mechanism of feedthrough¹, consequential information and gestural communication [12].

The experiment we conducted was designed to be series of 6 mosaic completions, 3 in the co-localized situation and 3 in the distant situation. The mosaics are composed of 5x5 squared pieces. For each situation, 3 different mosaic types (abstract, figurative and textual) are completed by the pairs of participants. A

¹ As Dix *et al.* remarks [27], when artifacts are manipulated, they give off information which is *feedback* for the person performing the action and *feedthrough* for the persons who are watching.

textual mosaic represents a text (here a poem): the "right" orientation of each piece can be easily inferred as it contains words and typos. A figurative mosaic represents a scene or a portrait: the "right" orientation of each piece is more ambiguous and can necessitate to assembly many pieces before being deduced. An abstract mosaic represents an abstract painting or a fractal: the only orientation constraint is that all the pieces have the same final orientation.

To solve the mosaics, a Java application has been designed to run locally on a Diamond Touch and on the HERMES platform. Because of the Diamond Touch technical limitations, the application supports multiuser manipulation of pieces but only one-finger interaction per user. Only two action types on the mosaic pieces are allowed: moving or rotating a piece. A mosaic piece can be moved by the users along an invisible grid by touching it near its center and dragging it from one place to another. It can also be rotated but in 90 degrees steps as its edges have to remain parallel to the table sides. The user has to touch around one of the 4 piece corners, and to perform a rotational motion. A visual feedback is given to let the user identify the current selected action (a cross pointer for dragging, and a round arrow for rotation).

During the mosaic completions, the pairs of subjects were filmed and all their actions were recorded (piece number, situation and orientation on the table of the touched piece, action type - rotation or moving).

A total of 24 participants took part in the study, randomly put in pairs. There were one female pair, six male pairs and five mixed pairs. All participants were postgraduates, had normal or corrected to normal vision. During the experiments, the order of the mosaic completions were counterbalanced in situations (co-localized and distant) and in mosaic types (abstract, figurative and textual). The pairs of participants completed an individual training period before the 6 collaborative mosaic completions.

6 Findings

6.1 Objective evaluation

Comparison of mosaic types: Figurative mosaics are completed more quickly ($M=362s$, $SD=182s$) than textual mosaics ($M=435s$, $SD=394s$). Abstract mosaic completions are the most time consuming ($M=565s$, $SD=394s$). The completion times of the three mosaic types were compared using a Friedman Anova and a significant difference was observed ($F(2)=30.3$, $p<0.001$). Post-hoc comparisons were performed which revealed significant differences between the completion time of each pair of mosaic types. The same observations were made when considering the two conditions (co-localized vs distant) separately.

Textual mosaics are completed with less rotations ($M=29$, $SD=5$) than figurative mosaics ($M=59$, $SD=31$). Abstract mosaic completions requires more rotations ($M=86$, $SD=52$). The three mosaic types were compared using a Friedman Anova and a significant difference was observed ($F(2)=33.1$, $p<0.001$). Post-hoc comparisons were performed which revealed significant differences between each pair of mosaic types. The same observations were made when considering the two conditions (co-localized vs distant). These differences reflect the difficulty to find the right orientation of each piece of abstract mosaic pieces.

Comparison between co-localized and distant situations: For all the kind of mosaic (Text, figurative or abstract), a Wilcoxon test reveals no significant differences between mosaic completions times in co-presence ($M=397s$, $SD=221s$) and in remote configuration ($M=441s$, $SD=237s$).

Concerning the particular class of textual mosaics, no significant difference exists between resolution times in remote configuration ($M=456s$, $SD=210s$) and in co-presence ($M=368s$, $SD=149s$). No significant differences were observed between number of rotations in co-presence ($M=53$, $SD=23$) and in remote configuration ($M=56$, $SD=26$) as well as between number of movings in co-presence ($M=132$, $SD=43$) and in remote configuration ($M=146$, $SD=56$).

Mosaic completion processes in co-localized and distant situations: In the side-by-side/co-localized situation, the subjects exploit a larger surface on the table than in the face-to-face/distant situation (see Fig. 4). Moreover, the strategies used to complete the mosaics are different. In the side-by-side/co-localized situation, each user tends to work independently before merging their assemblage: two distinct areas are visible on the right image of Fig. 4. In distant/face-to-face situation the users tends to work together on the same zone of the table (only one main area on the left image of Fig. 4).

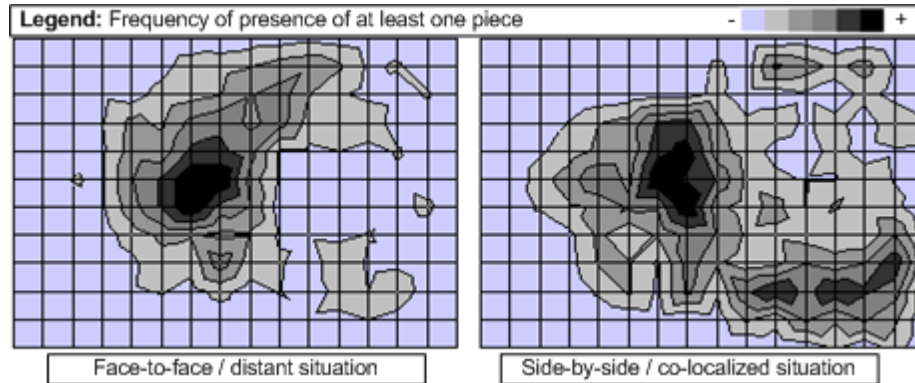


Fig. 4. Dominant localizations of the pieces on the table during the mosaic completions.

The dominant orientation of the pieces for textual mosaics (see Fig. 5) is also different in side-by-side/co-localized and face-to-face/distant configuration since in the first case, the orientation of the pieces is not conflicting whereas in the second case the subjects have to negotiate to find the more comfortable orientation. With no surprise, in the side-by-side situation, the pieces are oriented toward the two subjects. In the face-to-face situation, two strategies were equally used by the subjects: orientate the pieces toward one of the subject or perpendicularly. One can note that pieces oriented toward a particular subject are mostly situated around his right hand (*i.e.* on the right of the table for pieces oriented toward the downside of the table and on the left of the table for pieces oriented toward the upside of the table).

Discussion: From these first results from objective criteria, no significant difference can be observed between distant and co-localized situations when consider-

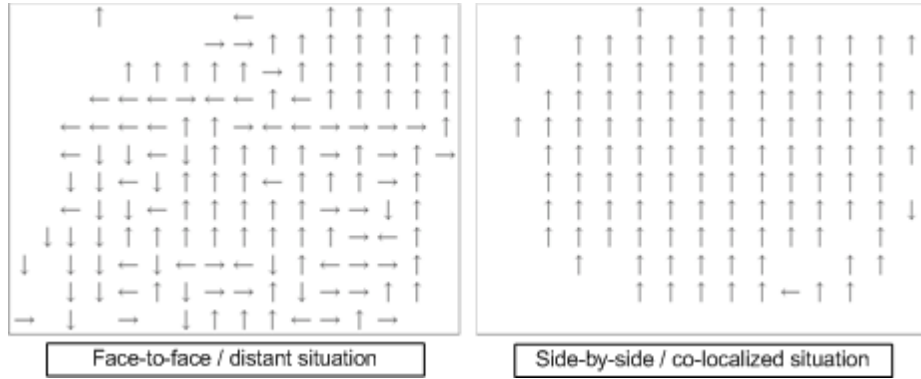


Fig. 5. Dominant orientations of the pieces on the table during the 1st third of the textual mosaic completions. An uparrow expresses that pieces are the most often oriented to be read from the downside of the table, and so on.

ing completion time, type of mosaic and piece manipulation (number of rotations and moves). Consequently, the actual technology conveys sufficient context's elements of the remote sites to enable efficient collaboration between two distant groups without hindering the completion of the task.

However, distant and co-localized configurations differ in term of coupling. Coupling [28] qualifies the degree people are working together. It can evolve to loose coupling when each person can work without needing information from another person, to tight coupling when they collaborate. This observed difference may be due more to the spatial configuration (side-by-side *vs* face-to-face) than to the co-localization or distant parameter. Indeed, in side-by-side configuration, people tend to impede each other's interaction on piece access and manipulation. This probably favors a first independent phase during which each user tries to complete an independent part of the puzzle. Furthermore, such coupling may be biased by the fact that our mosaic application does not facilitate independent work merging. One can not move and rotate a group of pieces. This interaction is actually boredom as it has to be done one piece at a time. Thus, loose coupling may be hindered in the face-to-face configuration in favor of an early collaboration.

6.2 User experience

Subjects were asked to comment the mosaic completions by focusing on the application and on the differences between the distant and co-localized situations.

Mosaic application: To the 1st question ("What do you think of the mosaic application?"), the most common remarks made by the 24 subjects are:

- It is difficult to rotate the pieces (19 subjects).
- Moving many pieces simultaneously should be managed (14 subjects).
- Rotating many pieces simultaneously should be managed (9 subjects).

These remarks confirm that some users have completed sub-parts of the mosaic and would have needed tools to merge these parts.

Configuration: To the 2nd question ("Could you compare mosaic completion in co-localized and in remote situations?"), the most common remarks made are:

- When concentrated on the mosaic completion, the video-communication system is not useful, but the sound system is (14 subjects).
- The major trouble is reading the textual mosaic in face-to-face (distant) situation (13 users).
- The distant gesture representation is interesting since it brings communicational and intentional information (7 users).
- The same users also remark that the other user's hands are not always easy to see (7 users).
- 6 users find the mosaic completions more pleasant in co-localized situation, 3 in distant situation and 2 found the two situations identical. The other users did not compare the two situation in that way.

As our experiments were focused on the mosaic completion, people do not need to talk to each other before interacting. The participants feel useless the video-communication system.

Only 6 over 24 users openly prefer the co-localized situation, the others are indifferent at worst case. Thanks to remote gesture visualization, co-localized and distant collaboration are both worthwhile experience. Distant collaboration is not seen as a poor ersatz of co-present interaction.

The reading problems due to orientation of textual pieces in the face-to-face situation are more critical. It is not explicitly due to the distant configuration.

Subjective opinions: We have selected a few comments from the subjects which are particularly representative of how the experiment was felt:

- "It is better in the remote configuration, more efficient, arms do not cross: we do not bother each other"
- "There is no presence feeling with the video system. I felt as if I played alone and the robot sometimes helps me."
- "In the distant situation, we are forced to speak, other's actions are more difficult to perceive."
- "We are more efficient in the side-by-side configuration, it is easier to read."

The users sometimes mix remarks about distance factor and "spatial configurations" (face-to-face *vs.* side-by-side configurations) since distant and co-localized situations are not symmetric. Although experiments have shown that co-localized and distant collaboration are comparable in terms of efficiency, presence feeling may be radically different (see the second comment about the robot). This can be seen as the fact that we respect almost all the Gutwin's recommendation about workspace awareness (we provide tools for intentional and consequential communication) but the artifacts and feedthrough. Remote gesture visualization lacks of incarnation. The presence feeling of the distant participant may be enhanced with sound effects of piece manipulation.

Further experiments are necessary to conclude on the influence of each parameter on the collaboration strategy.

7 Conclusion and future work

In this paper, to address the design of a collaborative multiuser platform, we propose HERMES which combines a multiuser tactile interactive tabletop, a video-communication system enabling eye-contact with real size distant user visualization, a spatialized sound system and a computer vision module for distant users' gesture visualization.

Preliminary evaluation of HERMES highlights important issues for the design of collaborative tabletop interfaces. Although HERMES does not provide the same presence feeling in distant and or co-localized situations, an important result emerges: distance does not hinder efficient collaboration anymore. Further experiments are necessary to conclude on the influence of each parameter on the collaboration strategy. The second finding, and it is not surprising, is that the orientation of object during collaborative tasks is important, particularly with textual material.

HERMES needs a few technical improvements to enhance the presence feeling: improve the quality of video and sound, increase the opacity of hand representation and take into account more feedthrough such as rubbing noises. Concerning the mosaic application, the rotation problems encountered by the users could be easily solved by increasing the surface dedicated to the rotations on the pieces. To avoid asymmetric behaviors from the users on each distant sites, the fact that users can be right or left handed has to be thoroughly considered. Finally, the application should support containers to move and/or rotate set of pieces at the same time, to enable fitting of mosaic parts independently completed. This last point should induce more loose coupling in the collaborative task.

A new series of experiments is planned when these technical improvements would be integrated. The role of the video will be evaluated with objective criteria, as well as the importance of hand visualization and their mutual influence. The impact of the collaborative configuration (face-to-face *vs* side-by-side) according to the situation (distant *vs* co-localized) will be also studied.

References

1. M. Stefik, D. Bobrow, G. Foster, S. Lanning, and D. Tatar, "Wysiwiw revised: Early experiences with multi-user interfaces," in *ACM Transactions on Office Information Systems*, vol. 5:2, pp. 147–167, April 1987.
2. E. Pederson, K. McCall, T. Moran, and F. Halasz, "Tivoli: an electronic whiteboard for informal workgroup meeting," in *interCHI'93*, pp. 391–398, 1993.
3. A. Tang, C. Neustaedter, and S. Greenberg, "Videoarms: embodiments in mixed presence groupware," in *Proceedings of the BCS-HCI British HCI Group Conference*, 2006.
4. R. Kraut, S. Fussell, and J. Siegel, "Visual information as a conversational resource in collaborative physical tasks," *Human Computer Interaction*, vol. 18, pp. 13–49, 2003.
5. D. Kirk, D. Stanton, and T. Rodden, "The effect of remote gesturing on distance instruction," *CSCL'05*, 2005.
6. E. Bier and S. Freeman, "Mmm: A user interface architecture for shared editors on a single screen," in *UIST'91*, pp. 79–86, 1991.

7. S. Greenberg and M. Boyle, "Moving between personal devices and public displays," in *Workshop on Handheld CSCW*, November 1998.
8. J. Stewart, B. B. Bederson, and A. Druin, "Single display groupware: a model for co-present collaboration," in *CHI'99*, pp. 286–293, 1999.
9. J. P. Hourcade and B. B. Bederson, "Architecture and implementation of a java package for multiple input devices (mid)," *Tech Report HCIL-99-08*, 1999.
10. E. Tse and S. Greenberg, "Sdgtoolkit: A toolkit for rapidly prototyping single display groupware," in *CSCW'02*, Novembre 2002.
11. U. Pawar, J. Pal, and K. Toyoma, "Multiple mice for computers in education in developing countries," in *Conference on Information and Communication Technologies for Development*, 2006.
12. C. Gutwin and S. Greenberg, "A descriptive framework of workspace awareness for real-time groupware," *Computer Supported Cooperative Work, Special Issue on Awareness in CSCW*, vol. 11, pp. 411–446, 2002.
13. A. Tang, M. Boyle, and S. Greenberg, "Understanding and mitigating display and presence disparity in mixed presence groupware.," *Journal of Research and Practice in Information Technology*, vol. 37, no. 2, 2005.
14. Y. Rogers and S. E. Lindley, "Collaborating around vertical and horizontal large interactive displays: which way is best?," *Interacting with Computers*, vol. 6, pp. 1133–1152, 2004.
15. *Tabletop 2006, Adelaide, Australia*, IEEE Computer Society, 2006.
16. P. H. Dietz and D. Leigh, "DiamondTouch: A Multi-User Touch Technology," in *UIST'01*, pp. 229–226, November 2001.
17. J. Rekimoto, "Smartskin: An infrastructure for freehand manipulation on interactive surfaces," 2002.
18. F. Bérard, "The Magic Table: Computer-Vision Based Augmentation of a Whiteboard for Creative Meetings," in *PROCAM'03*, (Nice, France), June 2003.
19. A. Wilson, "Touchlight : an imaging touch screen and display for gesture based interaction," *ICMI'04*, 2004.
20. J. Han, "Low-cost multi-touch sensing through frustrated total internal reflection," in *Symposium on User Interface Software and Technology*, (Seattle), pp. 115 – 118, 2005.
21. G. Hollemans, T. Bergman, V. Buil, K. van Gelder, M. Groten, J. Hoonhout, T. Lashina, E. van Loenen, and S. van de Wijdeven, "Entertaible: multi user multi-object concurrent input," in *UIST'06*, 2006.
22. M. Endsley, "Toward a theory of situation awareness in dynamic systems," *Human Factors*, vol. 37(1), pp. 32–64, 1998.
23. L. Segal, "Designing team workstations: The choreography of teamwork, in local applications," *Local Applications of the Ecological Approach to Human-Machine Systems*, 1995.
24. H. Clark and S. Brennan, "Grounding in communication," in *Readings in Groupware and Computer Supported Cooperative Work: Assisting Human- Human Collaboration* (M.-K. Publishers, ed.), pp. 222–233, 1991.
25. R. Kruger, S. Carpendale, S. Scott, and S. Greenberg, "How people use orientation on tables: comprehension, coordination and communication," in *GROUP'03*, (Sanibel Island, USA), pp. 369–378, ACM Press, 2003.
26. J. Tang, "Findings from observational studies of collaborative work," *International Journal of Man-Machine Studies*, vol. 34, no. 2, pp. 143–160, 1991.
27. A. Dix, J. Finlay, G. Abowd, and R. Beale, *Human-Computer Interaction*. 1993.
28. T. salvador, J. Scholtz, and J. Larson, "The denver model for groupware design," *ACM SIGCHI Bulletin archive*, vol. 28, pp. 52 – 58, 1996.