PEPR eNSEMBLE (PC1) – PhD proposal

Acronym and title: DYTelepres - Supporting Collaborative Dynamics in Telepresence Systems

Host institution and direction: The PhD student will be hosted at IMT Atlantique (Brest campus) within the INUIT team of Lab-STICC. The proposed director is Cédric Fleury, as he plans to defend his HDR before the beginning of the thesis.

Context:
Videoconferencing and telepresence have long been a way to enhance communication among remote users. They improve turn-taking, mutual understanding, and negotiation of common ground by supporting non-verbal cues such as eye-gaze direction, facial expressions, gestures, and body langue [3, 6, 10]. They are also an effective solution to avoid the "Uncanny Valley" effect [7] that can be encountered when using avatars.
However, such systems are often limited to basic setups in which each user must seat in front of a computer equipped with a camera. Other systems, such as Multiview [9] or MMSpace [8], handle groups, but still only group-to-group conversations are possible. This leads to awkward situations in which colleagues in the same building stay in their office to attend a videoconference meeting instead of attending together, or participants are forced to have side conversations via chat. More recent work investigates setups that allow users to move into the system and interact with share content. t-Rooms [5] displays remote users on circular screens around a tabletop. CamRay [1] handles video communication between two users interacting on remote wall-sized displays. GazeLens [4] integrates a remote user in a group collaboration around physical artifacts on a table.

Challenges and scientific objectives:
The current telepresence systems do not support different moments in the collaboration, such as tightly coupled and loose collaboration, subgroup collaboration, spontaneous or side discussions. Supporting such dynamics in collaboration is a major challenge for the next telepresence systems. The first objective of this PhD work is to understand the social mechanisms involved in these collaborative dynamics. The second objective is to design solutions for mediating them through technology in remote collaboration.

Approach:
We first plan to study collaboration dynamics in co-located situations without technology mediation. In particular, we will observe groups of co-located users interacting in large interactive spaces, such as rooms equipped with large screens, tabletops, or AR devices. Based on our findings, we will design telepresence systems that use different technologies or devices to support various phases of the collaboration in remote settings. For example, we can imagine video that appears or disappears appropriately to manage tightly coupled or loose collaboration, or video split and displayed at different positions of the screens to handle subgroup collaboration. Additional devices, such as smartphones, could also be used on the fly to make side conversations possible. Besides video, we will also consider the use of spatialized sound [2]. It can be valuable for determining remote user positions, managing subgroup discussions without disturbing others or providing privacy for side conversations. Finally, lessons learned from the development of these prototypes will inform the design of future telepresence systems.

Positioning within the PEPR eNSEMBLE:
This project fits inside the **PC1** and is totally in line with the **3rd priority theme** as it aims to support transitions between various phases of the collaboration. However, it can also contribute to the **2nd priority theme** as it

will propose guidelines on how to represent remote collaborators in telepresence systems according to the different moments of the collaboration.

Organization of the PhD project:
The PhD work will be organized according to the following steps:

- First 3 months: the PhD student will analyze the state-of-the-art and identify collaborative situations that involve multiple phases in collaboration, as well as transitions between these phases. These situations will be explored in the rest of the thesis.
- Next 6 months: based on the findings from the previous study, the student will design and evaluate a telepresence system.
- Next 12 months: the student will design and evaluate telepresence systems based on the findings of the previous study.
- Next 9 months: the proposed telepresence system will be redesigned based on the results of the first evaluation. The student may also consider other collaborative scenarios and assess the proposed system in these new scenarios.
- Last 6 months: the student will concentrate on writing the thesis manuscript and finalizing papers.

Partners:

- Cédric Fleury, MCF in computer science/HCI at IMT Atlantique, Lab-STICC, INUIT team: he will provide expertise on telepresence systems for large interactive spaces, such as wall-sized displays or large augmented reality spaces.
- Michel Beaudouin-Lafon, Prof. in computer science/HCI at Univ. Paris-Saclay, LISN, ExSitu Inria team: he will provide expertise on computer-supported cooperative work and design of interactive systems.
- Olivier Gladin, research engineer at Inria Saclay, LISN, ExSitu Inria team: although he will not be directly involved in the PhD supervision, he will provide a technical support on telepresence systems and remote collaboration, as well as video and audio capturing and streaming.

All partners are involved in the **EquipEx+ Continuum** and the PhD student will benefit from the wall-sized displays, AR/VR devices, and telepresence systems available in the platforms WILD, WILDER and ATOL.

References:

[1] H. Bai, P. Sasikumar, J. Yang, and M. Billinghurst. "A User Study on Mixed Reality Remote Collaboration with Eye Gaze and Hand
[1] I. Avellino, C. Fleury, W. Mackay and M. Beaudouin-Lafon. "CamRay: Camera Arrays Support Remote Collaboration on Wall-Sized Displays". *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI'17)*. 2017.

[2] L. Fyfe, O. Gladin, C. Fleury, M. Beaudouin-Lafon. "Combining Web Audio Streaming, Motion Capture, and Binaural Audio for Collaborative Telepresence". *Proceedings of the 4th Web Audio Conference (WAC)*. 2018.

[3] E. A. Isaacs and J. C. Tang. "What Video Can and Can't Do for Collaboration: A Case Study." *Proceedings of the ACM International Conference on Multimedia* (MULTIMEDIA'93). 1993.

[4] K.-D. Le, I. Avellino, C. Fleury, M. Fjeld, A. Kunz. "GazeLens: Guiding Attention to Improve Gaze Interpretation in Hub-Satellite Collaboration". *Proceedings of the Conference on Human-Computer Interaction (INTERACT'19)*. 2019.

[5] P. K. Luff, N. Yamashita, H. Kuzuoka, and C. Heath. "Flexible Ecologies And Incongruent Locations." *Proceedings of the Conf. on Human Factors in Computing Systems*. (CHI '15). 1995.

[6] A. F. Monk and C. Gale. "A Look Is Worth a Thousand Words: Full Gaze Awareness in Video-Mediated Conversation." In: *Discourse Processes* 33.3, 2002, pp. 257–278.

[7] M. Mori, K. F. MacDorman and N. Kageki, "The Uncanny Valley [From the Field]", *IEEE Robotics & Automation Magazine*, vol. 19, no. 2, pp. 98-100, 2012.

[8] K. Otsuka, "MMSpace: Kinetically-augmented telepresence for small group-to-group conversations". *Proceedings of 2016 IEEE Virtual Reality (VR'16)*. 2016.

[9] A. Sellen, B. Buxton, and J. Arnott. "Using Spatial Cues to Improve Videoconferencing." *Proceedings of the ACM Conference on Human Factors in Computing Systems* (CHI '92). 1992.

[10] E. S. Veinott, J. Olson, G. M. Olson, and X. Fu. "Video Helps Remote Work: Speakers Who Need to Negotiate Common Ground Benefit from Seeing Each Other." *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '99). 1999.