



HAL
open science

Modeling and real-time feedback of collaboration in virtual reality using multimodal signals

Aurélien Léchappé, Cédric Fleury, Mathieu Chollet, Cédric Dumas

► To cite this version:

Aurélien Léchappé, Cédric Fleury, Mathieu Chollet, Cédric Dumas. Modeling and real-time feedback of collaboration in virtual reality using multimodal signals. ECCE 2023: European Conference on Cognitive Ergonomics, Swansea University, Sep 2023, Swansea, United Kingdom. hal-04494775

HAL Id: hal-04494775

<https://hal.science/hal-04494775>

Submitted on 7 Mar 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Modeling and real-time feedback of collaboration in virtual reality using multimodal signals

Aurélien Léchappé

LS2N, Lab-Sticc, IMT Atlantique, aurelien.lechappe@imt-atlantique.fr

Cédric Fleury

Lab-Sticc, IMT Atlantique, cedric.fleury@imt-atlantique.fr

Mathieu Chollet

School of Computing Science, University of Glasgow ; LS2N, IMT Atlantique, mathieu.chollet@glasgow.ac.uk

Cédric Dumas

LS2N, IMT Atlantique, cedric.dumas@imt-atlantique.fr

Issues in surgery can be related to collaboration problems such as miscommunication and misunderstanding. Analysis of these collaborative situations can be done by experts, but it is a long and complex process. Moreover, this approach does not allow for real-time analysis and feedback on the collaboration. Some recent approaches aim to assess and model the quality of collaboration using measures that can be computed in real-time (e.g., non-verbal, gaze, physiological data). Considering some of the challenges related to the evaluation and modeling of collaboration, we are working on a tool that aims to compute metrics in real-time to report on the quality of collaboration within a group. This would allow us to act before a critical situation emerges. We first apply this tool to a collaboration situation in a virtual environment to test their effectiveness and their effects on the situation before applying them in a real situation as surgery.

Keywords: Collaboration, Virtual reality, Collaboration cues, Multimodal signals

1 CONTEXT

Operating rooms (OR) have been extensively studied among the critical systems domain to prevent adverse events and improve surgical practice, tools or training [14]. Our previous work studied robotic surgery from the surgeon perspective [13] and from the team perspective [3]. A surgical team is typically a small team composed of an expert surgeon, surgeon assistant, anesthetist, nurse anesthetist and circulating nurse. In surgery, a number of issues are related to miscommunication and misunderstanding among them. These issues as well as a lack of other non-technical skills such as decision-making, leadership and situation awareness, increase risk of surgical complications [8]. Human factors (HF) specialists can manage to analyze working situations and identify issues than can be remediated, this is a long process based on postures, speech and activity analysis. Moreover, HF approach in the OR is not a routine procedure. Our research focuses on the detection of collaboration issues among the team, for debriefing purposes or to provide feedback on collaboration in real-time and detect issues such as misunderstanding or coordination problems.

Since the beginning of seminal research on multimodal human-computer interaction [7], much progress has been made in capturing users' activities to assess various individual and team processes. Indeed, the number of sensors and the general computing power available to researchers have increased substantially. The capture and processing

of complex social behavior data such as gaze, gestures, speech and para-verbal data are now possible in real-time. This has contributed to the increase in research on modeling and detecting collaboration using multimodal data over the past decade.

For efficient collaborative activities, collaborators must have an updated representation of the collaborative situation. To do this, they construct together a shared mental model of the situation. This representation has several names in the literature (common ground [1], common frame of reference [4], etc.) and refers to the sum of shared information about the understandings, knowledge, beliefs and assumptions between two or more people during an activity [1]. This mutual knowledge is necessary. Collaborators must continuously update their representations of the situation during the activity. Meier et al. [6] propose a framework to describe different aspects of the collaborative process involved during collaborative activity. Some of them are involved in maintaining common ground. These aspects are *communication*, *coordination*, *joint information processing*, *interpersonal relationship* and *motivation*. In this model, each aspect has several dimensions that all play a role in the activity, such as *sustaining mutual understanding* and *dialogue management* for communication; *information pooling* and *reaching consensus* for joint information processing. This framework is used by studies that attempt to detect collaboration using multimodal signals such as speech, activity, physiological or gaze [2,5,10,11].

In a literature review on modeling co-located collaboration, Praharaaj et al. [9] propose some high-level indices to evaluate collaboration based on Meier et al.'s dimensions. These indices are synchrony, equality, individual accountability, intra-individual variability, information sharing, mutual understanding, and reciprocal interaction, each measured using low-level metrics. In a later paper, Praharaaj et al. [10] propose an overview of collaboration indicators from audio, based on the indices postulated above. They propose parameters which variation would be indicators of the collaboration quality: dominance, rapport, expertise, active participation, roles, and knowledge co-construction. These parameters are calculated using low-level indices like speaking time distribution, turn-taking, overlap, discourse analysis or synchrony in audio features. An other work of Schneider et al. [12] presents a survey of uses of multimodal data to capture different collaboration constructs. Regarding collaboration processes, authors report indicators most present in the literature to assess collaboration. They report verbal indicators such as participation in discussion, similarity of discourse, equality of contribution as relevant indicators of collaboration quality. Authors also mention that episodes of synchrony (physiological, posture) between participants and joint visual attention can be used to predict collaboration quality. Schneider et al. [12] provide a list of collaboration indicators used and measured in studies across a wide variety of contexts and tasks. The authors raise the question of the transferability of multimodal collaboration analysis to other domains.

Thanks to the advances in multimodal data research described above, and those first predictors of collaboration found in the literature, we are building an experimental platform to aggregate and update in real time data whom model a partial view of the activity.

2 GOAL

Our approach is to apply and test the effectiveness of the collaboration indicators presented above. We conduct this in a small group collaboration situation in a controlled virtual environment to measure collaboration problems in different ways. Using a virtual reality collaborative task has the advantage of collecting precise behavioral data but more importantly, enable fine control of the collaboration situation and the deployment of a variety of feedbacks in

response to users' actions and collaborative states. The application of measures and augmentations of collaboration in virtual reality allows us to test their effectiveness and their effects on the situation before applying them in a real situation.

We conducted two experiments in which two people had to collectively complete a puzzle-solving task in a virtual environment (Figure 1). The first study aimed to evaluate the impact of adding collaborative cues in the virtual environment on the quality of collaboration and co-presence. The second study focused on the use of non-verbal and gaze measures to report the level of collaboration during the activity through the evaluation of some collaboration behaviors and processes. We collected and synchronized audio, video, gaze, and activity data for each collaborator in the virtual environment in real time. Using this data, we derived and computed some metrics such as speech time distribution, turn-taking, overlapping and joint visual attention events during activity execution.

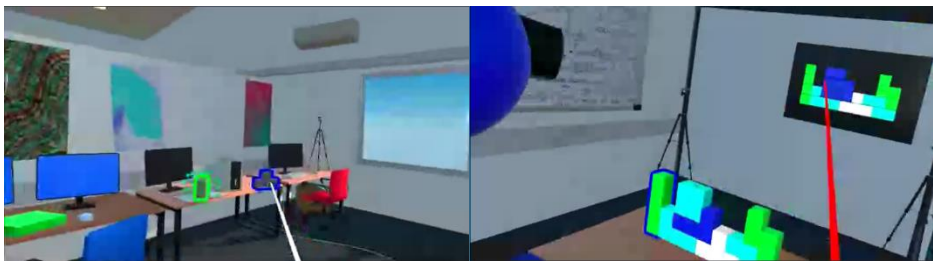


Figure 1: Virtual environment and collaborative task

3 CHALLENGE

The analysis of collaboration raises several challenges. In most collaborative situations, it is difficult to **distinguish what is a good or lower collaboration**. Signals indicating the quality of collaboration can vary depending on factors such as the task, team composition, or individual elements related to collaborators, making it complex to standardize the assessment and recognizing what constitutes good collaboration.

A second challenge is **measuring collaboration**. One of the most common solutions is the intervention of an expert to assess collaboration during the activity. However, this method of assessment is limited because it is not repeatable and can only assess a specific collaboration situation. Another solution is to start from a known situation. In both cases, a relevant analysis of the quality of the collaboration and of the processes underway during the activity is performed, but these approaches are not compatible with real-time feedback on the level of collaboration. To do this, many indicators are proposed in the literature to evaluate collaboration processes. However, these indicators may be linked to factors that depend on the context of the collaboration. Considering these indicators in contexts where they are not relevant could have negative effects on measurement of the collaboration.

Another challenge is to automatically **identify and measure relevant indicators** for reporting the level of collaboration in the collaborative situation. This would involve considering parameters such as the composition of the team, the expertise of each person, the roles of the collaborators or their proximity. It will also be necessary to consider the task performed and the interactions that they involve between the collaborators. This would imply the design of collaboration profiles / collaborative situation profiles depending on these parameters, to select relevant indicators to evaluation the collaboration.

4 CONCLUSION

With respect to the severity of problems and limitations of assessing the quality of collaboration in surgery, it would be useful to have real-time feedback on the quality of collaboration. Keeping in mind the challenges presented, we are working on a tool that compute real-time collaboration indicators to report on the quality of collaboration and to prevent emergence of incidences. Preliminary results of the tool using non-verbal and gaze measures to distinguish good from poor collaboration are encouraging.

REFERENCES

- [1] Herbert H. Clark and Susan E. Brennan. 1991. Grounding in communication. In *Perspectives on socially shared cognition*. American Psychological Association, Washington, DC, US, 127–149. DOI:<https://doi.org/10.1037/10096-006>
- [2] Yong Dich, Joseph Reilly, and Bertrand Schneider. 2018. Using Physiological Synchrony as an Indicator of Collaboration Quality, Task Performance and Learning. In *Artificial Intelligence in Education (Lecture Notes in Computer Science)*, Springer International Publishing, Cham, 98–110. DOI:https://doi.org/10.1007/978-3-319-93843-1_8
- [3] Julien Lefrançois, Christine Vidal-Gomel, Élise Olivier, and Jérôme Rigaud. 2022. Enjeux croisés de formation des acteurs et de résilience en chirurgie urologique assistée par un robot. *Savoirs* 60, 3 (2022), 31–48.
- [4] Anthony Loiselet and J.-M Hoc. 2001. La gestion des interférences et du référentiel commun dans la coopération: implications pour la conception. *Psychologie Française* 46, (January 2001), 167–179.
- [5] Roberto Martinez-Maldonado, Yannis Dimitriadis, Alejandra Martinez-Monés, Judy Kay, and Kalina Yacef. 2013. Capturing and analyzing verbal and physical collaborative learning interactions at an enriched interactive tabletop. *Intern. J. Comput.-Support. Collab. Learn.* 8, 4 (December 2013), 455–485. DOI:<https://doi.org/10.1007/s11412-013-9184-1>
- [6] Anne Meier, Hans Spada, and Nikol Rummel. 2007. A rating scheme for assessing the quality of computer-supported collaboration processes. *Computer Supported Learning* 2, 1 (March 2007), 63–86. DOI:<https://doi.org/10.1007/s11412-006-9005-x>
- [7] Laurence Nigay and Joëlle Coutaz. 1993. A design space for multimodal systems: concurrent processing and data fusion. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems (CHI '93)*, Association for Computing Machinery, New York, NY, USA, 172–178. DOI:<https://doi.org/10.1145/169059.169143>
- [8] Jason C. Pradarelli, Steven Yule, and Douglas S. Smink. 2020. Evaluating Non-Technical Skills in Surgery. In *Health Services Research*, Justin B. Dimick and Carrie C. Lubitz (eds.). Springer International Publishing, Cham, 125–135. DOI:https://doi.org/10.1007/978-3-030-28357-5_12
- [9] Sambit Praharaaj, Maren Scheffel, Hendrik Drachsler, and Marcus Specht. 2021. Literature Review on Co-Located Collaboration Modeling Using Multimodal Learning Analytics—Can We Go the Whole Nine Yards? *IEEE Trans. Learning Technol.* 14, 3 (June 2021), 367–385. DOI:<https://doi.org/10.1109/TLT.2021.3097766>
- [10] Sambit Praharaaj, Maren Scheffel, Marcel Schmitz, Marcus Specht, and Hendrik Drachsler. 2022. Towards Collaborative Convergence: Quantifying Collaboration Quality with Automated Co-located Collaboration Analytics. In *LAK22: 12th International Learning Analytics and Knowledge Conference*, ACM, Online USA, 358–369. DOI:<https://doi.org/10.1145/3506860.3506922>
- [11] Bertrand Schneider, Kshitij Sharma, Sebastien Cuendet, Guillaume Zufferey, Pierre Dillenbourg, and Roy Pea. 2018. Leveraging mobile eye-trackers to capture joint visual attention in co-located collaborative learning groups. *Intern. J. Comput.-Support. Collab. Learn.* 13, 3 (September 2018), 241–261. DOI:<https://doi.org/10.1007/s11412-018-9281-2>
- [12] Bertrand Schneider, Gahyun Sung, Edwin Chng, and Stephanie Yang. 2021. How Can High-Frequency Sensors Capture Collaboration? A Review of the Empirical Links between Multimodal Metrics and Collaborative Constructs. *Sensors* 21, 24 (December 2021), 8185. DOI:<https://doi.org/10.3390/s21248185>
- [13] Chloe Topolski, Cédric Dumas, Jérôme Rigaud, and Caroline Gl Cao. Why is circular suturing so difficult?
- [14] Charles Vincent and René Amalberti. 2016. *Safer Healthcare*. Springer International Publishing, Cham. DOI:<https://doi.org/10.1007/978-3-319-25559-0>