







participants. Vicon Nexus and Pupil Capture have to be validated if they are recognizing the subjects and are properly estimating the gaze. In case of too many false measurements, the calibration procedure should be repeated. Once the motion and gaze tracking are reliable, the internal recording (in Vicon Nexus and Pupil Capture) can start. After starting the internal data recording, all the scripts and application for LSL data streaming and synchronization has to be started. It is necessary to check if all the devices are sending streams on LSL network. When all the data are streamed and recorded properly, the experiment can start. Once the experiment is finished, all applications and scripts have to be stopped, and it is necessary to check if all the output files are properly saved and closed.

#### IV. CONCLUSION

In this work, we presented the experimental setup for acquiring a multimodal dataset containing eye gaze and body motion during human-human interaction. The setup involves a Vicon Motion tracking system and two Pupil-Labs gaze-tracking glasses. The acquired data are saved in raw format in manufacturer's recording software. For synchronizing the data obtained from different measuring equipment we developed applications and scripts that are streaming the data to LSL network and capture that data at the same place where timestamps of different data sources are matched. Our next steps will involve preparation of a dataset for the experimental setup that is described in Section III. that includes annotating specific events and fixations during interaction so that temporal correlations between important events can be modeled. Further on we will focus on improving the models for upper body humanoid robot control based on the derived model.

#### ACKNOWLEDGMENT

This work was partially supported by MNTR project III44008 and TR35003, EU H2020 project ACTICIPATE and

FCT project RBCog-Lab research infrastructure.

#### REFERENCES

- [1] M. Gallotti, M.T. Fairhurst, C.D. Frith, "Alignment in social interactions", *Consciousness and cognition*, 48, pp. 253-261., 2017
- [2] N. F. Duarte, M. Raković, J. Marques, J. Santos-Victor, "Action Alignment from Gaze Cues in Human-Human and Human-Robot Interaction" In Proceedings of the European Conference on Computer Vision (ECCV), 2018.
- [3] N. F. Duarte, M. Raković, J. Tasevski, M.I. Coco, A. Billard, J. Santos-Victor, "Action anticipation: Reading the intentions of humans and robots", *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 4132–4139, Oct 2018. <https://doi.org/10.1109/LRA.2018.2861569>
- [4] D. Preissmann, C. Charbonnier, S. Chagué, J.P. Antonietti, J. Llobera, F. Ansermet, P. J. Magistretti, "A motion capture study to measure the feeling of synchrony in romantic couples and in professional musicians.", *Frontiers in psychology*, vol. 7, 2016.
- [5] E. van der Kruk, M. Reijne, "Accuracy of human motion capture systems for sport applications; state-of-the-art review.", *European journal of sport science*, nol.18, no. 6, pp. 806-819, 2018.
- [6] G.E. Robertson, G.E. Caldwell, J. Hamill, G. Kamen, S. Whittlesey, "Research methods in biomechanics", Human kinetics, 2018
- [7] C. Mandery, Ö. Terlemez, M. Do, N. Vahrenkamp, T. Asfour, "The KIT whole-body human motion database", *In 2015 IEEE International Conference on Advanced Robotics (ICAR)*, pp. 329-336, 2015
- [8] B. Burger, A. Puupponen, T. Jantunen, "Synchronizing eye tracking and optical motion capture: How to bring them together.", *Journal of Eye Movement Research*, vol. 11, no. 2, 2018.
- [9] Qualisys. <http://www.qualisys.com>
- [10] Vicon. <http://www.vicon.com>.
- [11] Optitrack, <https://optitrack.com/>
- [12] A. Duchowski, "Eye tracking methodology", Springer, 2007.
- [13] M. Kassner, W. Patera, A. Bulling, "Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction", In: Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing: Adjunct publication, ACM, pp. 1151–1160, 2014
- [14] L. Bishop, W. Goebel, "Mapping visual attention of ensemble musicians during performance of "temporally-ambiguous" music." *In Conference of Music & Eye-Tracking*. 2017.
- [15] F. Marandola, "Eye-Hand synchronization and interpersonal interaction in xylophone performance: A comparison between African and Western percussionists", *Journal of Eye Movement Research*, vol.11, no.2, 2019,
- [16] D.L. Mills, "Internet time synchronization: the network time protocol", *IEEE Tran. on communications*, Vol. 39, no. 10., pp 1482-1493, 1991
- [17] C. Kothe, "Lab streaming layer (LSL)", Accessed on April 26., 2019., <https://github.com/scn/labstreaminglayer>