Abstract Dual eye-tracking (DUET) is at the confluents of cognitive (and social) psychology and computer science. DUET is a novel methodology to explore the socio-cognitive processes underlying collaboration. The basic aims of DUET are to better understand, through gaze indicators, a socially distributed cognitive system and to support collaboration by real time gaze display of collaborators. We are interested in finding whether and how patterns of eye-movements can reflect the cognition underlying collaboration. This paper concentrates on the major motivations, methodological challenges and the future aspects of DUET.

Dual Eye-Tracking: Lessons Learnt

Kshitij Sharma, Patrick Jermann, Marc-Antoine Nüssli, Pierre Dillenbourg
CRAFT, École Polytechnique Fédérale de Lausanne
<firstname>.<lastname>@epfl.ch

1. Motivations

Gaze and speech relation: Griffin and Bock (2000) proposed the notion of eye-voice span (900 ms approximately) as the duration between last fixation on an object and the word onset about that object. Allopenna, et. al. (1998) showed that the listeners' voice-eye span between hearing a verbal reference and looking at the object of reference is between 500 and 1000 ms. Richardson, et. al. (2007) found that speaker's voice leads the corresponding listener's gaze by 2 seconds.

Measuring collaboration and performance: The cross-recurrence of two gaze streams was used to measure the collaboration quality by Richardson, et. al. (2007) in a listening comprehension task and by Nüssli (2011) in a pair program comprehension task. Nüssli, et. al. (2009) fed features calculated from raw gaze to machine learning algorithms to predict success in collaborative Raven and Bongard puzzle. In a shared map study Cherubini et. al. (2008) computed the level of misunderstanding as the distance between the referrer's and the partner's gaze points. Gaze patterns are also indicative of expertise in a particular task. In a collaborative Tetris task, Jermann et. al. (2010) used gaze patterns to predict the team composition. The preliminary results are encouraging but there are still restrictions on the generalizability of results because gaze patterns are extremely task specific.

2. Methodological Challenges

Activity Dependence: Gaze patterns are extremely task specific. Gaze can be described at different levels of aggregation. At low (fixation, saccadic length) and medium levels (dispersion, transitions in different parts of stimulus) there are no general relationships between variables and success/expertise. The back and forth gaze transitions between two objects has different meanings in Raven matrices (rule is induced by the help of speech at the end of interaction) and Bongard puzzle (induction is done at the beginning of interaction) (Nüssli, et. al., 2011). High gaze dispersion indicates information gathering in Raven matrices (Nüssli, et. al., 2011) and reflection over the map in concept mapping task (Sangin, et. al., 2008). At a high and task specific level of aggregation there are relations. In debugging a program we cannot get the gaze indicators of subtasks by merely aggregating raw gaze data, we need a more semantic aggregation (Nüssli, 2011). In a nutshell, we need to identify a match between the granularity of task decomposition and level of aggregation of gaze data.

Segmenting Interaction: Collaboration contains sequence of actions and communicative moves. To build a model of the interaction quality, it is essential to put temporal markers in the flow of interaction. Sharma, et. al. (2012) defined the segments of interaction by detecting stable (looking at a small set of objects) and together (looking at the same set of objects) moments. Simply put, for automatic assessment of quality of interaction, we need to find out how to automatically segment the interaction based on gaze data. It is important to keep in mind that different level of gaze aggregation can help in finding evidences for different granularities of task decomposition.

3. Applications

The most sought application of DUET is the Gaze Aware Tool to support the collaboration by mediating interaction through gaze input. Ishii and Kobyayash (1992), and Monk and Gale (2002) designed systems displaying the face of the collaborators. Stan and Brennen (2004) showed that displaying partners' gaze while debugging a program helped finding bugs. In another experiment Brennan, et. al. (2008) showed that displaying partners' gaze in "Os-in-Q" search helped the collaborators in more effective labor division.

4. Conclusion

We presented DUET as a research field at the intersection of psychology and computer science. Understanding socially distributed cognition and to measure collaboration quality have attracted most of the efforts. The task dependency of gaze imposes the first methodological challenge of finding a level of raw gaze aggregation to be indicative of a particular level of task decomposition. The second challenge has its roots in the first challenge itself. As the gaze is task dependent, and so the task decomposition, the challenge is to find a way to find different interaction phases using gaze data.

References

Allopenna, P.D., Magnuson, J.S, and Tanenhaus, M.K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models* 1,* 2,* 3,* 4,* 5. *Journal of memory and language*, 38(4),

Brennan, S.E., Chen, X., Dickinson, C.A., Neider, M.B., and Zelinsky, G.J. (2008). Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition*, 106(3), 1465–1477.

Cherubini, M., Nüssli, M.-A., and Dillenbourg, P. (2008). Deixis and gaze in collaborative work at a distance (over a shared map): a computational model to detect misunderstandings. *In Proceedings of ETRA*, ACM.

Griffin, Z.M., and Bock, K. (2000). What the eyes say about speaking. *Psychological science*, 11(4).

Ishii, H. and Kobayashi, M. (1992). Clearboard: a seamless medium for shared drawing and conversation with eye contact. *In Proceedings of CHI*, ACM.

Monk, A.F., and Gale, C. (2002). A look is worth a thousand words: full gaze awareness in video-mediated conversation. *Discourse Processes*, 33(3), 257–278.

Nüssli, M.-A., Jermann, P., and Li, W. (2010). Using dual eye-tracking to unveil coordination and expertise in collaborative Tetris. *In Proceedings of the British HCI, British Computer Society.*

Nüssli, M.-A. (2011). Dual-Eye Tracking Methods for the Study of Remote Collaborative Problem Solving. PhD thesis, EPFL.

Nüssli, M.-A., Jermann, P., Sangin, M., and Dillenbourg, P., (2009). Collaboration and abstract representations: towards predictive models based on raw speech and eye-tracking data. *In Proceedings of CSCL.*

Richardson, D. C., Dale, R., and Kirkham, N. Z., (2007). The art of conversation is coordination. *Psychological Science*, 18(5).

Sangin, M., Molinari, G., Nüssli, M.-A., and P. Dillenbourg, P. (2008). How learners use awareness cues about their peer's knowledge?: insights from synchronized eye-tracking data. *In Proceedings of the ICLS*.

Sharma, K., Jermann, P., Nüssli, M.-A., and Dillenbourg, P. (2012). Gaze evidences for different activities in program understanding. *In Proceedings of the PPIG.*

Stein, R., and Brennan, S.E. (2004). Another person's eye gaze as a cue in solving programming problems. *In Proceedings of the ICMI, ACM.*