ASCILITE SHAPING THE FUTURE OF TERTIARY EDUCATION
Show Me The Learning
Adelaide, Nov 27-30, 2016

Mining video data: tracking learners for orchestration and design

Kate Thompson
School of Education and Professional Studies
Griffith University

Sarah Howard
Early Start Research Institute
University of Wollongong

Jie Yang
SMART Infrastructure Facility
University of Wollongong

Jun Ma
SMART Infrastructure Facility
University of Wollongong

Learning spaces influence how we act, however there is a lack of systemic research addressing the impact of environments on teaching and learning. In this paper, we introduce a hybrid tracking technique in which a colour model is combined with algorithms to identify human positions, and applied to video data. The aim of identifying patterns of movement that could be used to indicate successful collaboration in open plan learning spaces. We apply the method to a previously analyzed dataset, to demonstrate how multiple analytic techniques can be used to build a complex understanding of learner movement in relation to collaboration and learning. We conclude with suggestions of the ways in which the results could be used by instructors to inform orchestration of complex learning environments, as well as directions for future research.

Keywords: learning analytics, learner tracking, learning spaces

Introduction

Learning spaces influence how we act, in ways we may not notice (Amedeo, Golledge & Stimson, 2009). Given the substantial investment by universities in the design and construction of new learning spaces, it seems reasonable to expect changes in the ways learners are considered, designed for, and facilitated. However, there is a lack of systemic research addressing the impact of environments on teaching and learning (Brooks, Walker & Baepler, 2014). Many new learning spaces adopt an open-plan configuration. To understand how learners use these types of spaces, we need to know what it means for learning, and in turn, for teaching. With growth in the use of 'big data' (e.g. Macfadyen & Dawson, 2010), which can be broad, or deep, new methods are being developed to do just that. In this paper, we present an investigation of learners' movements in an open-plan space as they complete a complex learning task, using a hybrid tracking technique in which a colour model is combined with algorithms to identify human positions, and applied to video data. We present background to the methodological approach and the context in which this analysis takes place. The methods developed for tracking learners are described and the results presented. We discuss the results in relation to the findings of previous analyses of the same dataset, and suggest ways in which the results could be used by instructors to inform orchestration of complex learning environments, as well as directions for future research.

Background

In this paper, we argue that in order to understand, design for, and teach in new learning spaces, we need to work in the intersection of several core methodological and theoretical approaches to understanding learning and teaching. A body of work that examines the facilitation of such learning environments is *orchestration* (Prieto, Dlab, Gutiérrez, Abdulwahed, & Balid, 2011). Orchestration is used to describe a teacher's management of a classroom in which students have access to a range of technological devices (Dillenbourg, Javela & Fischer, 2009). There is growing interest in the study of learning and teaching in spaces that include digital and non-digital tools and the development of automated or semi-automated methods (learning analytics, multimodal learning analytics) to understand activities in these environments (see e.g. Martinez-Maldonado, Goodyear, Kay, Thompson & Carvalho, 2016; Thompson, Ashe, Carvalho, Goodyear, Kelly & Parisio, 2013). To study complex learning environments, such as new university open-plan learning spaces, multiple measures of learner activity are needed. In multimodal learning analytics (Blikstein, 2013), multiple modes of activity – gesture, gaze, as well as discourse, movement and the creation of artefacts are considered, and a more systemic view of a learning situation is adopted as the results of the analyses are recombined in order to develop a model of understanding (Thompson, 2013). Recent work is developing the relationship between using real-time learning analytics for orchestration of learning across physical and digital spaces (Martinez-Maldonado, 2016). *Actionable science,* from recent work in ecology, encourages scientists to work directly with policy makers, so that science has a better chance of influencing policies (Beardsley, 2011). The ability of learning analytics techniques to be actionable relies on core interaction between researchers and instructors in order to develop tools that meet the needs of learners, and can be integrated into practice (Martinez-Maldonado, Pardo, Mirriahi, Yacef, Kay & Clayphan, 2016).

There is ongoing discussion about appropriate ways for universities to design for, and assess, graduate attributes such as teamwork, interpersonal communication, problem solving, critical thinking and creativity (Frawley, Dyson, Tyler & Wakefield, 2015). Computer-supported collaborative learning is a common activity in many higher education learning situations. Identifying indicators of successful and unsuccessful collaboration using learning analytic approaches is more common where there is a digital component to the learning environments, where the evidence of student activity, such as eportfolios (Aguiar, Ambrose, Chawla, Goodrich & Brockman, 2014), or online discussions (e.g. Wise, Zhao & Hausknecht, 2014) are digital and can be collected and analyzed. Yet, we know very little about identifying productivity, when groups use a combination of physical and digital tools in a learning space (Goodyear, Jones & Thompson, 2013). Typically, the application of learning analytics to physical spaces has focused on gaze (e.g. Schneider & Pea, 2015) or attention (Raca & Dillenbourg, 2015). In these studies, however, students remain at desks, and it is their response to the teacher, or other students in a fixed group that is analyzed (e.g. Raca, Tormey & Dillenbourg, 2013). Open-plan learning spaces are potentially more complex than this. Across many studies an important aspect of collaboration has been *convergence*, whether there is a particular physical location of gaze (e.g. Schneider & Pea, 2015; Raca & Dillenbourg, 2015) or a concept in dialogue (Jeong & Chi, 2007). We argue that by being able to track learners in their movement in an open-plan learning space, that their physical location could also be used as an indicator of the productivity of their collaboration, and indicate to teachers whether intervention is needed. A learner tracking tool was developed to help us to gain a birds-eye-view of patterns of movement around a room, in order to be used in combination with other measures of learning and assessment.

Methods

Human tracking has been investigated in many settings in the last decade. For instance, the tracking techniques have been applied to areas such as public surveillance, gaming, human-computer interactions and robotics. As a high-level computer vision task, the aim of human tracking is to establish the coherent relations of human beings, given consecutive video streaming frames. For major computer-vision based applications, accurate tracking is the fundamental work enabling identification of activity and behavior. Previous work reported on in learning analytics has focused on the use of heatmaps to achieve an understanding of the patterns of users around technology (Martinez-Maldonado et al., 2016).

This dataset has been reported on previously (see Thompson et al., 2013 for further detail). Masters students were given a design task to be completed in a 5-week period. The group analyzed collaborated on this task in online and a face-to-face environment. The group consisted of four students (Damien, Eileen, Gabrielle, and Lavina, pseudonyms), and met online (Skype) four times and face-to-face three times (these were recorded as well as their work in Google Drive). One segment of their first meeting was analyzed in this analysis, although all face-to-face sessions have been analyzed previously (Thompson et al., 2013). The students' grades in this group were comparable to the rest of the class for the individual components. However, this group received the highest grade in the class for the collaborative component, which led us to identify the collaboration in this group as successful. The physical environment (Figure 3) contained writeable walls, onto which computers could be projected. The furniture could be moved to suit the needs of the learners in the room. The task required students to work in groups of four to discuss and collaboratively design educational design patterns and a pattern language. In the previous analysis, automated discourse analysis was used to examine the ways in which learners used the tools, interpreted the assignment brief, and designed their own roles for this task. Changes in the use of keywords from the assignment brief, and contained in the final assignment were used to show the shift in focus in this group, and to identify who was responsible for saying them. The gaze was also recorded, and used to identify the focus of attention within the group and compare this between the first fifteen minutes of the first session and the middle of the final session. The first excerpt is also the focus of our proof of concept below. In what follows, we describe the development of the algorithms for learning tracking and the application to this data. Many algorithms have been developed for human tracking. Figure 1 shows the main features of three core models adopted in human tracking.

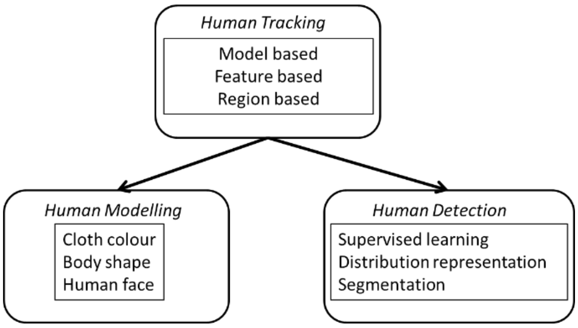

Figure 1. Functional diagram for human tracking.

As observed from Figure 1, human tracking is traditionally presented using a top-down approach, which consists of three modules. *Human modelling* module characterizes a group of people (or individual people) of interest. Firstly, the target objects (human beings) need to be represented in a specific format; then the module applies different features to characterize the objects, such as cloth color, body shape, or even the faces (Rincon, Makris, Urunuela, & Nebel, 2011, Feng, Guan, Xu & Tan, 2009, Ramanan, Forsyth & Zisserman, 2007). *Human detection* module involves the identification of the human beings in the given video frames. The identification module can either be provided in the initialization stage only or be integrated into the tracking algorithm. A variety of algorithms can be employed for identification, such as, supervised learning, distribution representation and segmentation (Yang, Bouzerdoum, & Phung 2010). *Human Tracking* module brings human modeling and human representation together to look for the target object of interest. According to the target representations, existing human tracking techniques can be classified as model-based, feature-based, and region-based tracking. Model-based tracking generally detects the human model in a video sequence. The commonly-used models are human body and face. Feature-based tracking employs various features, such as skin color, texture, and edge. Finally, region-based methods track the moving target in the forms of blobs, or body parts. We refer the readers to (Hu, Tan, Wang & Maybank, 2004) for a more comprehensive survey on human object tracking.

The performance of the tracking method using only one type of target representation is easily influenced by environmental noise including illumination changes, image blur, or camera movements. To improve the tracking performance, in this paper, a hybrid tracking technique is employed by combing the Hue Saturation Value (HSV) color feature, Mean Shift algorithm (Cheng, 1995) and Kalman filter (Kulkarni & Vargantwar, 2014). More precisely, the HSV color feature is employed to describe colors in terms of their shade and brightness. Then we apply the Mean Shift algorithm on the HSV feature to identify the exact human positions in the current frame. Next the Kalman filter is further employed to search for human in the next frame. The employed tracking technique is also illustrated in Figure 2.

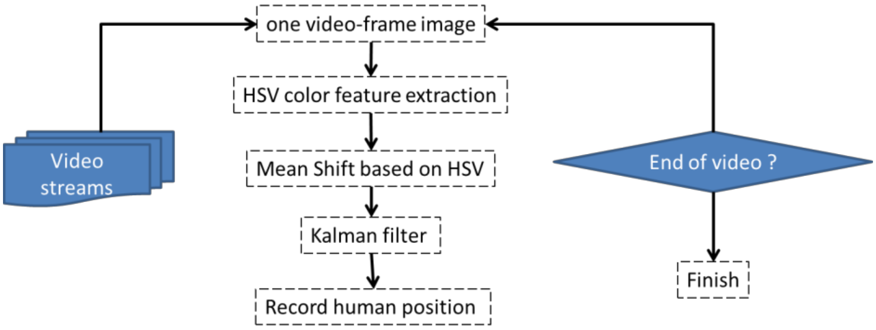

Figure 2. The employed human tracking technique

Results and discussion

The previous analysis of this group focused on the roles that emerged within the group (leader, synthesizer, online tool specialist, and coordinator), with evidence from the discourse as well as gaze of how these roles developed over the five weeks (Thompson et al., 2013). The current results demonstrate movement of the students in their very first encounter (see Figure 3). The student represented by the red line (at the laptop) becomes the online tool specialist. The student represented by the green line becomes the coordinator. The student represented by the blue line becomes the group leader, and the student sitting at the keyboard who does move during this segment becomes the synthesizer. In this first segment, the previous analysis showed that the students represented by the green and blue lines were the focus of other team members' attention, but were not in control of the tools that drew the others' gaze, so we assumed that this was due to their verbal contribution. When we examined patterns in their gaze in the original analysis, it was clear that the learners were focused on different objects, with little convergence. This is consistent with their movement, with different patterns of movement also clear in Figure 3.

These results indicate four distinctly different patterns of movement in the learning space indicative of the group's collaborative process. Of the four students, two have clearly wider-ranging patterns of movement, while the other two were more stationary (one responsible for the displays on the whitewalls, the other a focus of attention due to verbal contributions). Such patterns of movement have implications for the types of resources introduced by teachers into the learning environment. For example, the sitting student has not moved because of the nature of their task. This may be a function of the keyboard they are using and placement of other surfaces they can use to work. If they are stationary, it is possible that their contribution to the discussion is limited. An awareness of movements, in relation to tasks and roles performed by students may have implications for the effectiveness of learning designs, group roles, how teachers frame group tasks and configure the learning space. Understanding how the space was used, its relation to other information we have about the group, such as how roles and social interactions change over time, informs teachers' knowledge of how to effectively configure and facilitate groups in new learning spaces. What we can see from this is that at this early stage in their teamwork, the students were not collaborating around a shared representation physically. They were not gathered around the table, or the whitewall. From the beginning, their movement was separate, their movements were dispersed throughout the space and the roles that they adopted in the group reflected these behaviours. A similar representation of their final face-to-face meeting would indicate very little movement, with all gathered around the far whitewall, stationary on chairs as they focused on their individual roles in creating the shared assignment.

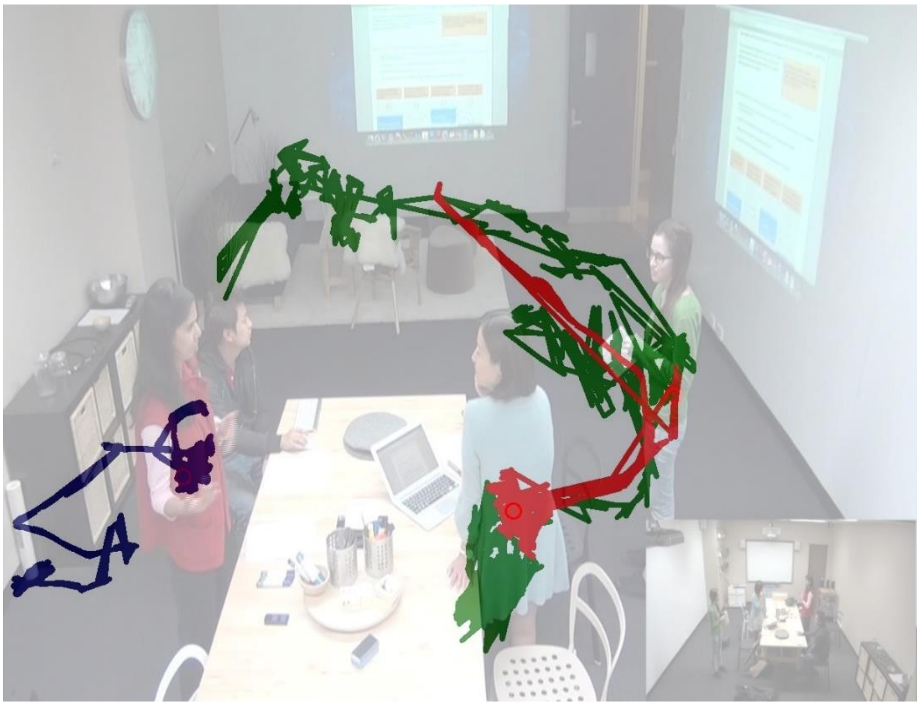

Figure 3: results of the application of the learning tracking algorithm

Future work will better represent these changes over time, and will include more complex learning spaces. Here we have presented analysis of a single group, to test our approach. However, the classroom is a much more complex and dynamic space. Our methods will need to be further developed to handle multiple groups and a teacher. As we examine different groups and in a variety of contexts, we will be able to identify patterns of movement and use that can be firstly used to prompt further and deeper investigation, and secondly used to be able to identify patterns of behaviour to help with orchestration in such spaces. To support better learning designs, fine-tune and provide feedback to students in real-time, it is necessary to start to develop findings and methods in a predictive fashion.

We introduced this paper with reference to the recent investment by universities in new learning spaces, as well as the importance of graduate attributes, such as teamwork, interpersonal skills, and creativity. These findings, while preliminary, suggest that the motion of learners could be indicative of roles and progress through a task in a way useful to orchestration by a teacher. They point to students' dispersed behaviour patterns, guided by their task and role, which is made possible by the open-learning space. In a more traditional classroom, students would be situated in relation to the work, likely including table and/or digital device. The more traditional space is where teachers would have experience and would frame their learning designs. Identifying behaviours and patterns that may influence or relate to learning in new spaces can inform and support teachers' new learning designs, guide decisions on resources to include in these spaces and expectations for student's roles and collaborations.

References

Aguiar, E., Ambrose, G. A., Chawla, N. V., Goodrich, V., & Brockman, J. (2014). Engagement vs Performance: Using electronic portfolios to predict first semester engineering student persistence. *Journal of Learning Analytics,* 1(3), 7-33. https://doi.org/10.18608/jla.2014.13.3

Amedeo, D., R. G. Golledge, and R. J. Stimson. 2009. Person Environment Behavior Research: Investigating Activities and Experiences in Spaces and Environments. New York, NY: Guilford.

Brooks, Walker & Baepler, 2014. Overview. New Directions for teaching and learning

Cheng, Y (1995). "Mean shift, mode seeking, and clustering", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(8): 790-799. https://doi.org/10.1109/34.400568

del Rincon, J. M., Makris, D., Urunuela, C. O. & Nebel, J. C. (2011). Tracking human position and lower body parts using Kalman and particle filters constrained by human biomechanics. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 41(1):26–37. https://doi.org/10.1109/TSMCB.2010.2044041

Dillenbourg, P., Jarvela, S., & Fischer, F. (2009). The evolution of research on computer-supported collaborative learning: From design to orchestration. In N. Balacheff, S. Ludvigsen, T. Jong, A. Lazonder, & S. Barnes (Eds), *Technology-enhanced learning,* pp. 3-19. New York: Springer.

Feng, S., Guan, Q., Xu, S. & Tan, F. (2009). Human tracking based on mean shift and Kalman filter. *In Proceedings of International Conference on Artificial Intelligence and Computational Intelligence*, volume 3, pages 518–522.

Frawley, J. K., Dyson, L. E., Tyler, J. & Wakefield, J. (2015). Building graduate attributes using student-generated screencasts. In *Australasian Society for Computers in Learning and Tertiary Education,* pp. 100-111, Perth, Australia.
Goodyear, P., Jones, C., Thompson, K. (2014). Computer-supported collaborative learning: Instructional approaches, group processes and educational designs. In J. M. Spector, M. D. Merrill, J. Elen, M. J. Bishop (Eds.), Handbook of research on educational communications and technology, (pp. 439-451). New York: Springer. https://doi.org/10.1007/978-1-4614-3185-5_35
Hu, W., Tan, T., Wang, L. & Maybank, S. (2004). A survey on visual surveillance of object motion and behaviors. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 34(3):334–352. https://doi.org/10.1109/TSMCC.2004.829274
Jeong, H. & Chi, M. T. H. (2007). Knowledge convergence and collaborative learning. *Instructional Science,* 35, 287-315. https://doi.org/10.1007/s11251-006-9008-z
Kulkarni, A., Vargantwar, M. (2014) Video Based Tracking with Mean-Shift and Kalman Filter, *International Journal of Engineering Research & Technology* (IJERT), 3(5): 1271-1274.
Macfadyen, L., & Dawson, S. (2010). Mining LMS data to develop an "early warning system" for educators: A proof of concept. Computers and Education, 54, 588-599. https://doi.org/10.1016/j.compedu.2009.09.008
Martinez-Maldonado, R., Goodyear, P., J.Kay, Thompson, K., and Carvalho (2016) An Actionable Approach to Understand Group Experience in Complex, Multi-surface Spaces. *SIGCHI Conference: Human Factors in Computing Systems, CHI 2016, 2062-2074.* https://doi.org/10.1145/2858036.2858213
Martinez-Maldonado, R., (2016). Seeing Learning Analytics Tools as Orchestration Technologies for Learning Activities Across Physical and Digital Spaces. In *International Conference on Learning Analytics and Knowledge, (LAK 2016)*, Edinburgh, UK. 2016, pages 70-73.
Prieto, L. P., Dlab, M. H., Gutierrez, I., Abdulwahed, M., & Balid, W. (2011). Orchestrating technology enhanced learning: A literature review and a conceptual framework. *International Journal of Technology Enhanced Learning,* 3(6), pp. 583-598. https://doi.org/10.1504/IJTEL.2011.045449
Raca, M. & Dillenbourg, P. (2014). Classroom social signal analysis, *Journal of Learning Analytics,* 1(3), 176-178. https://doi.org/10.18608/jla.2014.13.16
Raca, M., Tormey, R. & Dillenbourg, P. (2013). Student motion and its potential as a classroom performance metric. *3rd International Workshop on Teaching Analytics (IWTA).*
Ramanan, D., Forsyth, D. A. & Zisserman, A. (2007). Tracking people by learning their appearance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(1):65–81.
Schneider, B. & Pea, R. (2015). Does seeing one another's gaze affect group dialogue? A computational approach. *Journal of Learning Analytics,* 2(2), 107-133. https://doi.org/10.18608/jla.2015.22.9
Thompson, K. (2013). Using micro-patterns of speech to predict the correctness of answers to mathematics problems: An exercise in multimodal learning analytics. International Conference on Multimodal Interaction, ICMI 2013, New York City, New York: ACM Digital Library.
Thompson, K., Ashe, D., Carvalho, L., Goodyear, P., Kelly, N., Parisio, M. (2013). Processing and Visualizing Data in Complex Learning Environments. American Behavioral Scientist, 57(10), 1401-1420.
Wise, A. F., Zhao, Y., & Hausknecht, S. N. (2014) Learning analytics for online discussions: Embedded and extracted approaches. *Journal of learning analytics,* 1(2), 48-71. https://doi.org/10.18608/jla.2014.12.4
Yang, J., Bouzerdoum, A., Phung, S. L. (2010). Dimensionality reduction using compressed sensing and its application to a large-scale visual recognition task, *International Joint Conference on Neural Networks* (IJCNN 2010), 1-8.

Acknowledgements

We gratefully acknowledge the financial support of the Australian Research Council, through grant FL100100203 for the original data collection.

Please cite as: Thompson, K., Howard, S., Yang, J. & Ma, J. (2016). Mining video data: tracking learners for orchestration and design. In S. Barker, S. Dawson, A. Pardo, & C. Colvin (Eds.), *Show Me The Learning*. *Proceedings ASCILITE 2016 Adelaide* (pp. 560-565). https://doi.org/10.14742/apubs.2016.836

Note: All published papers are refereed, having undergone a double-blind peer-review process.