Understanding shared autonomy of collaborative humans using motion capture system for simulating team assembly

Tadele Belay Tuli, Martin Manns, and Michael Jonek

University of Siegen
PROTECH - Institute of Production Technology
FAMS - Chair for Production Automation and Assembly
57076 Siegen, Germany
Tel.: +49 271 740 - 2267
Homepage: https://protech.mb.uni-siegen.de/fams
sekretariat-fams@mb.uni-siegen.de

Abstract. In virtual production planning, simulating human motions helps to improve process planning and interaction efficiency. However, simulating multiple humans sharing tasks in a shared workplace requires understanding how human workers interact and share autonomy. In this regard, an Inertial Measurement Unit based motion capture is employed for understanding shifting roles and learning effects. Parameters such as total time, distance, and acceleration variances in repetition are considered for modeling collaborative motion interactions. The results distinguish motion patterns versus the undertaken interactions. This work may serve as an initial input to model interaction schemes and recognize human actions behavior during team assembly. Furthermore, the concept can be extended toward a human-robot shared autonomy.

Keywords: shared autonomy, role shifting, human motion capture, team interaction, manual assembly.

1 Introduction

In virtual production planning, simulating human motions for team-work requires understanding how humans physically interact and share autonomy. When human partners perform tasks together, corporate communications may occur through force interaction, audio-visual perception, or individual motion behavior. Moreover, such communication combined with the humans' cognitive capability helps create smooth transition during role changes in case of repetitive tasks [15].

During collaborative tasks, human workers may change roles to be efficient and overcome temporary challenges that arise due to visibility, reachability, stability, and capability. In this context, the leader-follower principle can be served provided that the leader excels at the desired task by guiding the partner. Furthermore, the leader-follower concept has been applied to guide robots to imitate what the human partner has demonstrated e.g., [1].

The knowledge of multi human motion behavior could also be used for transferring human skills into robot models that has been extended into activity recognition and role adaptation. In the human-robot collaboration (HRC) scheme, understanding multi human motion behaviors during physical interaction can be useful to power robots' cognitive capability that allows them to adapt to dynamic tasks for complementing human workers. In virtual manufacturing, it may further enhance HRC process planning and interaction efficiency. However, simulating multiple humans comprising physical contact is difficult. Various methods that are able to predict humans motion and recognize activities have been presented e.g., [7, 3]. In an assembly context, it is unclear, e.g., how two persons change roles and adapt tasks in dynamic situations.

Therefore, in order to investigate human interaction during team works in manual assembly processes, a simultaneous multi human motion capturing approach is proposed. In this regard, identical wearable motion tracking systems are employed to simulate digital human models (DHM) using a virtual shop floor environment (i.e., using the MOSIM MMI framework [14]). Spatial and temporal analysis has been conducted to investigate parameters that measure or indicate shifting roles during physical interaction in collaborative works in smart assembly. Results are analyzed considering learning effect and shifting roles.

2 Related works

2.1 Predicting shifting roles

According to the report of MarketsandMarkets™[8], the use of collaborative robots is steadily increasing yearly by more than ten percent. In the future, this may lead to a wide application range that requires human-robot shared working space. Such a situation may create a conflict of role shifting due to the autonomy of both partners (e.g., human vs. human, human vs. robot or robot vs. robot) [20]. In humans' collaborative tasks, assuming an assembly of a huge object, there is a high probability of role shift that may account to various circumstances. The core idea is that, this kind of role shifting can be learned, and after some time, will be normalized. The underlying challenges of human motion prediction between two adult humans have been described in various works. For example, according to [19], action prediction depends on the workers' physical and functional similarity. Action prediction for two workers who are quite different physically and functionally is challenging. The partner who perform a difficult task by observing may process a conflict [2, 19]. In such situations, predicting the shifting role may be desired to resolve shifting role conflicts and further to establish mutual caring. For instance, a partner who is losing contact with an object with a hand may exhibit twisting motions attempting to control the object. In such a situation, informative role shifting or sharing is required.

The concept of role shifting can be applied in HRC, for example, to create a symbiotic human-robot partnership during collaborative object handling [17]. Furthermore, it may help to transfer learned skills to robots using machine learning techniques e.g. [13] and for creating a semantically represented knowledge base e.g. [18].

2.2 Quality measures for role and intention differentiation

Motion capture systems are being employed to acquire human motion behaviors in shop-floor environments. Features that help recognize human activities and roles may include joint position and orientation in spatial and temporal spaces. Identifying human activities and the underlying action may support the decision to adopt the task lead when a shifting of roles occurs. Rule based activity recognition and short term long memory based action prediction have been presented to contextualize human activities and anticipations [7, 3]. However, the quality of motion recognition is essential to correctly discriminate the workers' role during collaborative tasks .

Quality of motion capture can be affected by the hardware and software of a motion capture system [12, 10]. The presented techniques for motion capturing includes optical sensors (e.g. cameras)[4], combination of multi-view camera and Inertial Measurement Unit(IMU) [9]. According to [9], an approach to combine IMU and calibrated video has been proposed to solve lighting and background dynamics in addition to solving the limitations of video or IMU data sets. However, the effect of accelerated motions and force analysis has been excluded. Wearable IMU based motion capture systems have been considered for their occlusion-free accuracy and robustness for different applications [16]. According to the review of [5], wearable motion tracking systems are regarded as cost-effective, accurate, non-invasive, and portable.

3 Design of Experiment

A wearable motion capture system is employed for capturing two persons who are working together. The task is to assemble aluminum (Al) profiles with 2.5 m length to a Bosch TS/2+ system in a lab environment. The workers grasp the object from end-to-end and then carry it to the target position, and finally assemble using screws. Two similar IMU-based bodysuits from Xsens ®(MVN Link) are employed. These bodysuits are desired for sampling motion data at high speed (e.g., 240 fps), because dynamic actions may affect the statistical analysis (cf. [11]). Indicative features for role and autonomy shifting are obtained by reasoning the variation of spatial and temporal data.

The assembly process is defined as follows: First, the participants grasp the extruded Al profile (i.e., 40x40 mm^2, 4E, LP, Length = 2.5 m) object from end-to-end and carry it to the assembly position. At the assembly position, an interior joint is fixed. Once the partners reach the assembly position, they assemble the object from each side by pushing the object into the slot. The slot has a constraint from both sides, which confirms the final positioning. Finally, they tighten the screw and complete the task. The test comprises two scenarios as described below;

3.1 Experiment scenarios:

Test one: Speechless operation - In this scenario, the main task is to investigate how humans share autonomy and shift leading roles. At the same time, the scenario is selected to investigate the learning effect over time. The participants have no previous knowledge about the process and operation procedures. Also, they are instructed not to communicate orally with each other during the operation. Instead, the fixed starting and final positions are shown to them. The operation is repeated fifteen times. The joint slot is small and hardly visible to the partners from the opposite side, leading to potential conflicts.
Test two: Instructive operation - in this scenario, the workers are instructed to communicate with each other. They are told to discuss what and how to achieve the goal by exchanging required information about the process. Such an approach is required to contextualize the way humans communicate for sharing autonomy in real-time problems. In the process, a random obstacle is introduced to destruct the learning effect and increase negotiation activities. This operation takes place fifteen times.

3.2 Detecting role shifts in collaborative tasks

Decoupling features that are associated with role shift may also require a complex method combining various techniques. However, motion behavior can be analyzed in spatial and temporal spaces for understanding how roles are shifting, and autonomy is shared. The motion behavior in this regard considers the position and acceleration variations using time scale. At a fixed time interval, both workers' similarities of position and acceleration are analyzed to identify the partner, who is leading, and derive the associated reasons why the role shift takes place. The initiating factors for shifting roles such as obstacles, lack of visibility, loss of attention, and lack of process knowledge are considered for detecting the potential instant of shifting roles. The success criteria are measured based on the video analysis by experts. When the leader shifts to the follower role, there is a significant acceleration spikes with a time a difference. Moreover, the acceleration patterns and the initiating factors are verified based on the actual video record.

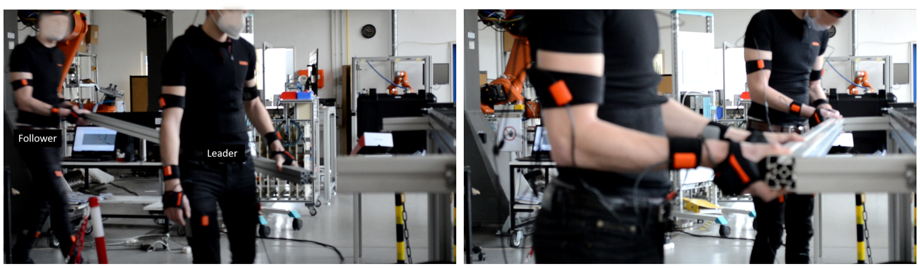
3.3 Learning effect in shared autonomy

Before extracting dynamic features of role in a collaborative task, keyframe annotation is applied to label each repetitive cycle's start and final frames. Then, the selected frames are exported for further analysis. Here, keyframes that are not part of the measurement are excluded. Each cycle of operation is represented as a separate file. Then, each cycle operation is analyzed regarding the similarities of parameters such as total time, total distance and the variance of acceleration of body joints. The similarity analysis of the two workers' motion is required to investigate the learning effect and behavior of collaboration. The similarity analysis could also provide a hint when a conflict in collaboration takes place.

According to [6], the similarity of two vectors can be measured using various methods, including Partial Curve Mapping (PCM), Curve Length, and Dynamic Time Warping (DTW). In the current work, curve length and total time are applied to investigate similarities of the workers' motion with repetitions.

4 Result and Discussion

The motion capture data that has been generated in a wearable IMU system is analyzed in spatial and temporal spaces. Further decomposition of motions into spatial features is used to extract workers' behavior in negotiation and finding assembly solutions without prior knowledge. In this regard, the workers share not only autonomy of part handling but also role leadership. For instance, the leading role has been shifted among the workers during walking and inserting into a position, as illustrated in Fig. 1.



(a) (b)

Fig. 1. Shifting roles during collaborative task; (a) leader-follower basis carry object and walk, and (b) role negotiation during physical interaction to find assembly position that is constrained by joints

Spatial constraints such as obstacles, guided positioning, and low tolerance geometries are challenging to perform with strict procedures and instructions. Instead, it demands interactive negotiation through physical interfaces. There could be a sliding of roles from one worker to the other frequently to find optimal assembly process. In such scenarios, human workers may able to adapt and plan in real-time operation. The ultimate goal is extending this skill into simulation environment. In this regard, reasoning knowledge base can be developed for each process from motion capture data.

The shifting roles are estimated based on the comparison of the acceleration variation as it is depicted in Fig. 2a and Fig. 2b. The instant of higher acceleration may indicate role conflict in which the partner attempts to overtake the lead. Accordingly, the shifting of roles is detected e.g., at the obstacle, the first participant is leading and then waiting for the second partner to walk over the

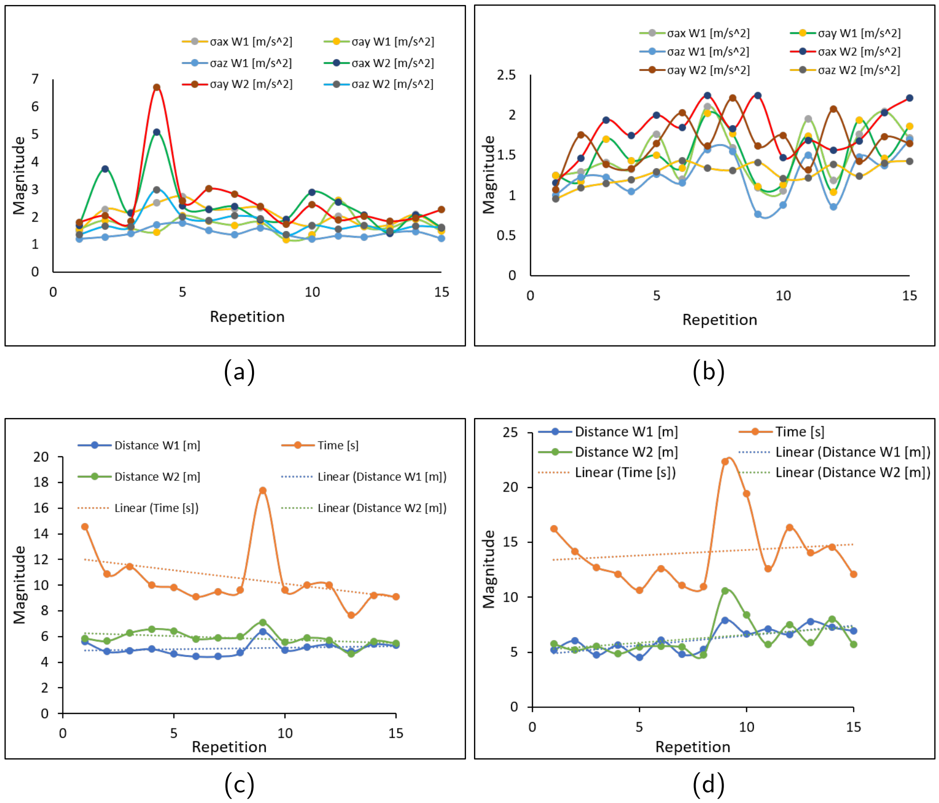


Fig. 2. (a, b) The standard deviation of pelvis joint linear acceleration for each cycle of operations for test one and two; (c, d) Learning effect analysis based on total distances, and duration

obstacle before the next steps. In the 4^{th} cycle of operation (see Fig. 2a), a random obstacle is introduced. The second worker who was following had trouble following and instead took a leading role to avoid an obstacle. The spread of the acceleration is a good indicator of the scenario. The motion capture data is parameterized versus the sequence of repetitions to understand how the learning and role adaptation takes place. The standard deviation (e.g., $\sigma_{a(x,y,z)}$) describes how the body joint accelerations (e.g., pelvis) vary. In this context, the lower value of the standard deviation shows both workers' body joint motion compliance. The role shifts during an obstacle avoidance is accurately detected using pelvis joint. However, the acceleration variance of a single joint alone is not sufficient for robust role shift detection. To further elaborate, it is necessary to investigate in future works motions from different joints, particularly foot, hand and head movements.

Similarly, the total distance covered by the body joints indicates how the motion varies in each sequence of repetitions. The linear regression of total time and distance are used to analyze the learning effect (see Fig. 2c and Fig. 2d). The total distance in test one is almost constant (slope $\leq 5\%$). Apparently, in test two, the obstacles are causing the workers to chose a dynamic path. Besides, at the ninth cycle, the workers are instructed to exchange leading roles (see Fig. 2c and Fig. 2d). The total time is also decreasing in test one (slope=-20%) except the ninth cycle where as in test two, it is linearly increasing (slope = 10%).

5 Conclusion and future outlook

The current work presented approaches that serve to recognize shifting roles and shared autonomy for collaborative tasks. Shift of roles may affect the performance efficiency, process plan and workers interaction. The ultimate goal is to transfer learned skills to robots for creating efficient HRC. Further investigation that employs feedback systems has to be carried out for improving instant detection of role shifts.

Acknowledgment

The authors would like to acknowledge the financial support by the Federal Ministry of Education and Research of Germany within the ITEA3 project MOSIM (grant number: 01IS18060AH), and by the European Regional Development Fund (EFRE) within the project SMAPS (grant number: 0200545).

References

1. Billard, A., Calinon, S., Dillmann, R., Schaal, S.: Robot Programming by Demonstration. In: Siciliano, B., Khatib, O. (eds.) Springer Handbook of Robotics, pp. 1371–1394. Springer, Berlin, Heidelberg (2008), https://doi.org/10.1007/978-3-540-30301-5_60
2. Blake, R., Shiffrar, M.: Perception of Human Motion. Annual Review of Psychology 58(1), 47–73 (Dec 2006), https://www.annualreviews.org/doi/10.1146/annurev.psych.57.102904.190152, publisher: Annual Reviews
3. Carrara, F., Elias, P., Sedmidubsky, J., Zezula, P.: LSTM-based real-time action detection and prediction in human motion streams. Multimedia Tools and Applications 78(19), 27309–27331 (Oct 2019), https://doi.org/10.1007/s11042-019-07827-3
4. Elhayek, A., Kovalenko, O., Murthy, P., Malik, J., Stricker, D.: Fully Automatic Multi-person Human Motion Capture for VR Applications. In: Bourdot, P., Cobb, S., Interrante, V., kato, H., Stricker, D. (eds.) Virtual Reality and Augmented Reality. pp. 28–47. Lecture Notes in Computer Science, Springer International Publishing, Cham (2018)
5. Filippeschi, A., Schmitz, N., Miezal, M., Bleser, G., Ruffaldi, E., Stricker, D.: Survey of Motion Tracking Methods Based on Inertial Sensors: A Focus on Upper Limb Human Motion. Sensors (Basel, Switzerland) 17(6) (Jun 2017), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5492902/

6. Jekel, C.F., Venter, G., Venter, M.P., Stander, N., Haftka, R.T.: Similarity measures for identifying material parameters from hysteresis loops using inverse analysis. International Journal of Material Forming (may 2019), https://doi.org/10.1007/s12289-018-1421-8
7. Liu, Z., Liu, Q., Xu, W., Liu, Z., Zhou, Z., Chen, J.: Deep Learning-based Human Motion Prediction considering Context Awareness for Human-Robot Collaboration in Manufacturing. Procedia CIRP 83, 272–278 (Jan 2019), http://www.sciencedirect.com/science/article/pii/S2212827119306948
8. Ltd, M.R.P.: Collaborative Robot (Cobot) Market by Payload, Component (End Effectors, Controllers), Application (Handling, Assembling & Disassembling, Dispensing, Processing), Industry (Electronics, Furniture & Equipment), and Geography – Global Forecast to 2026, https://www.marketsandmarkets.com/Market-Reports/collaborative-robot-market-194541294.html
9. Malleson, C., Collomosse, J., Hilton, A.: Real-Time Multi-person Motion Capture from Multi-view Video and IMUs. International Journal of Computer Vision (Dec 2019), https://doi.org/10.1007/s11263-019-01270-5
10. Manns, M., Fischer, K., Du, H., Slusallek, P., Alexopoulos, K.: A new approach to plan manual assembly. International Journal of Computer Integrated Manufacturing pp. 1–14 (Apr 2018), https://www.tandfonline.com/doi/full/10.1080/0951192X.2018.1466396
11. Manns, M., Mengel, S., Mauer, M.: Experimental Effort of Data Driven Human Motion Simulation in Automotive Assembly. Procedia CIRP 44, 114–119 (2016)
12. Manns, M., Otto, M., Mauer, M.: Measuring motion capture data quality for data driven human motion synthesis. Procedia CIRP 41, 945–950 (2016)
13. Mohammed, O., Bailly, G., Pellier, D.: Acquiring Human-Robot Interaction skills with Transfer Learning Techniques. In: Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction. pp. 359–360. HRI '17, Association for Computing Machinery, New York, NY, USA (Mar 2017), https://doi.org/10.1145/3029798.3034823
14. MOSIM: Github repository codes of MOSIM framework (Feb 2021), https://github.com/Daimler/MOSIM, original-date: 2020-10-22T14:24:54Z
15. Reed, K.B., Peshkin, M.A.: Physical Collaboration of Human-Human and Human-Robot Teams. IEEE Transactions on Haptics 1(2), 108–120 (Jul 2008), conference Name: IEEE Transactions on Haptics
16. Roetenberg, D., Luinge, H., Slycke, P.: Xsens MVN: Full 6DOF Human Motion Tracking Using Miniature Inertial Sensors p. 7 (2009)
17. Tirupachuri, Y., Nava, G., Rapetti, L., Latella, C., Darvish, K., Pucci, D.: Recent Advances in Human-Robot Collaboration Towards Joint Action. arXiv:2001.00411 [cs] (Jan 2020), http://arxiv.org/abs/2001.00411, arXiv: 2001.00411 version: 1
18. Tuli, T.B., Kohl, L., Chala, S.A., Manns, M., Ansari, F.: Knowledge-Based Digital Twin for Predicting Interactions in Human-Robot Collaboration. Västerås, Sweden (Sep 2021)
19. Vesper, C.: How to support action prediction: Evidence from human coordination tasks. In: The 23rd IEEE International Symposium on Robot and Human Interactive Communication. pp. 655–659 (Aug 2014), iSSN: 1944-9437
20. van Zoelen, E.M., Barakova, E.I., Rauterberg, M.: Adaptive Leader-Follower Behavior in Human-Robot Collaboration. In: 2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN). pp. 1259–1265 (Aug 2020), iSSN: 1944-9437