Evaluation of the Application of Smart Glasses for Decentralized Control Systems in Logistics

Thomas Kirks[1], Jana Jost [1] and Tim Uhlott[1]

Abstract— **The approach of this paper is to investigate the application of smart glasses for human-robot interaction for scenarios like warehouse logistics, transportation systems and logistics in general. Following various research questions we conduct an experiment with multiple tests to gather information on the accuracy of localization with the Microsoft HoloLens and the feasibility of its application. The tests can also be taken into account when testing other wearables for the same purpose. An adequate accuracy enables the human worker to be integrated in decentralized control systems better and build teams of the diverse entities in human-robot interaction and cooperation.**

I. INTRODUCTION

Due to the rising amount of individual products and just-in-time delivery flexible industrial systems are needed. Such complex cyber-physical production systems (CPPS) arise in the Industry 4.0 and can be seen as socio-technical systems in which humans and technologies are working close together [1]. First robotic systems occur which focus on the acceptance of the human by offering natural interaction [2]. While the technologies adapt to the human, the information about the worker is often very limited. Especially information referring to the location of the human worker is mandatory. Only then the different technologies can adjust their routes if interaction is needed and the overall system can provide the human worker with only those data which he needs.

II. BACKGROUND

A. Localization and Navigation in Warehouses

Localization is a crucial task for efficient processes in warehouses and production facilities. Due to just-in-time and lot size one production all goods have to arrive at various production steps in a very short time window. Further, large storage areas are reduced by small buffer places so that intermediate bearing is decreased. For optimally routing goods through the warehouse, the positions of all entities have to be known. Nowadays, flexible transportation is mainly be done by automated transport vehicles (ATV) or the human worker. Often the ATV and the human work together and an organization as a heterogeneous fleet is needed.

Fig. 1. Navigation methods for ATV according to [4]

Navigation for ATV can be divided into the three subtasks: position estimation (localization), calculation of the movement (path planning) and tracing the path (path control) [3]. Further, the variety of navigation techniques (see Fig. 1) can be classified according to [4] into methods using a physical guideline and ones following a virtual one. The first can be realized as an optical, magnetic or inductive lane. The ATV is equipped with proper sensors for continuously tracking the lane. Depending on the solution the installation costs may vary, while all of them lack flexibility [5]. Methods using a virtual guideline are far more flexible. They can be realized as ground markers arranged in a grid, as off-floor markers used with laser-navigation or as active sender e.g GPS [3]. Localization based on active senders can be done by different methods e.g. “received signal strength indicator” [6].

Most of the above mentioned navigation methods are just suitable for an ATV and cannot be used to help the human worker navigate through the warehouse. Usually he is equipped with wearables which offer at least an internal measurement unit (IMU) and can calculate the odometry.

B. Human-Robot Interaction

Nowadays, humans encounter different robots in a variety of working fields of production and logistics facilities. Regardless of the degree of autonomy of the robot, in every application some form of interaction between robots and humans takes place [7]. Although human-robot interaction (HRI) is often referred to as a subdiscipline of human-computer interaction (HCI) or computer-supported cooperative work (CSCW) [8], the complexity and dynamics of such systems differ from general HCI [9]. There is a variety of taxonomies which classify HRI depending on different characteristics see [8], [10] and [9]. Mainly all taxonomies

*This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 688117 (SafeLog).

[1]Thomas Kirks, Jana Jost and Tim Uhlott are with the Department Automation and Embedded Systems, Fraunhofer Institute for Materialflow and Logistics, 44227 Dortmund, Germany thomas.kirks@iml.fraunhofer.de jana.jost@iml.fraunhofer.de tim.uhlott@iml.fraunhofer.de

DOI: 10.1109/ITSC.2019.8917159 ©2019 IEEE

distinguish the degree of autonomy of the robot as well as the role of the worker. Since the classification only uses a general role (e.g. supervisor) of the worker, individual abilities each one offers are not taken into account. These different abilities are very important for an efficient HRI. Especially, since today there are already various types of social robots which maintain a cognitive model of the human, behave in a social way and learn from the human behavior as well [11].

C. Decentralized Control Systems Organization

Due to just-in time delivery and lot size one the flexibility of current and future warehouses and production facilities is guaranteed by decentralized control systems which offer easy deployment of ATVs as well as other robotic systems [12]. All this entities are represented by agents which contain the functionalities and characteristics of their physical entity as well as behaviors and the ability to exchange data with one another in specific forms of protocols [13]. For sharing information and teaming up, the human can also be integrated in such a multi-agent system by representing him with all his abilities and properties as an agent [2]. In addition it makes sense to model multi-agent systems as a hybrid of conventional MAS and service-oriented architectures as stated by Iñigo-Blasco et al. [14].

III. ADAPTIVE INTERACTION IN LOGISTICS ENVIRONMENTS

In logistics, especially warehouse logistics, in a conventional sense, there is continuous conveyor technology but also autonomous transport vehicles sustaining the material flow. This is state-of-the-art technology for decades. For flexible warehouses, these ATVs use free navigation and have at some point the necessity to interact with humans sharing the same environment for collision avoidance, state transmission or also direct interaction when working together hand in hand. The interaction can take place along the entire supply chain, referring to goods inbound, quality assurance, storage, picking, packing and goods outbound. Sometimes the robots support the human worker, for example in picking scenarios when the robot takes items to a picking station and has to hand over the item to the worker. To relieve the worker, the robot can lift the load handling device (LHD) and the items on top to an ergonomic height, too. Here, the robot has to know the appropriate height. Also when approaching the worker, the robot has to know the convenient distance to meet to make the worker feel comfortable. For these scenarios localization, identification and the team integration play a major role. Kirks et al. [2] showed how to retrieve absolute pose of the worker for interaction, we also want to use this information and provide it to other non-locatable objects in warehouse logistics. An ATV, simply following an optical line, does not know its absolute pose in an environment and still can fulfil transport tasks. In the interaction between robot and human (referring to distance and height adaption), the poses of both entities need to be known to ensure convenient and effective team work.

We conduct an experiment in which we run specific tests to find out about the applicability of the HoloLens as an adequate interface to human-robot interaction, where it can accurately support the team work between human and robot especially for the purpose of sensing and perception of the human from the robot's point of view. In the tests we consider multiple factors that may influence the overall accuracy. Hence, the following questions drive the configuration of the setup and tests accordingly:

1) Does the sense of rotation of the human influence the tracking quality of the HoloLens?
2) Does the quality of tracking improve when a pose reference code is seen more often?
3) Does the vertical distance between HoloLens and reference code influence the tracking quality?
4) Does the quality decrease over time (after the reference code had been recognized for the last time)?
5) Does rotational movement influence the quality more than the translational movement?
6) For line following (optical) navigation systems: Does the front facing line have influence on the tracking?
7) Is the HoloLens localization and tracking accurate enough for coarse and fine localization?

To answer these questions we will describe the use case, design a test setup and procedure and elaborate the findings in the following paragraphs.

A. Use-Case Overview

The examined use case focusses on the order-picking process in warehouse logistics, where we differ between goods-to-person and person-to-goods scenario. In the goods-to-person scenario the worker is located at a picking station and ATVs take goods from the storage to this station, where the worker accepts these and packs them into packages or small load carriers. Here, multiple robots are used to provide items in sequence and reduce process delays. In the person-to-goods scenario the worker moves to the storage racks, locates the items to pick in the shelves, picks them and puts them into carriers which the supporting ATV has loaded - in this case one ATV satisfies the process stability. The interaction in both cases needs the ATV to know its location to either follow the picker or find the picking station for load handover. Furthermore, for adaption to the worker it is helpful to adjust the LHD to the physical properties (e.g. height) of the worker to provide an ergonomic handover of goods. For retrofit purposes, where the ATV can not locate itself in relation to the worker, it can drive to stations and storage areas using deprecated navigation methods but the technology may not be able to find its absolute position in the shop floor. Alike for cost reduction, when we want to deploy cost efficient ATVs that can not locate themselves - because absolute positioning demands a more expensive global positioning system. Hence, we want to evaluate the method of providing an accurate pose to the ATV by the use of the HoloLens - a device the picker might use already for other purposes -, which is able to locate itself, track itself in the process, identify other objects and by an agent implementation also provide a calculated absolute pose for this object (e.g. the ATV). To be able to use the proposed system the accuracy of this method has to be determined. Therefore, the considered use case is set up with a human worker wearing a HoloLens, an ATV that is not able to calculate its own absolute pose, but is able to simply follow a line and drive to picking stations or storage shelves. HoloLens (representative for the human worker) and ATV are represented by software agents, thus being able to communicate and share information (for example the absolute pose) in a multi-agent system [2].

B. Experimental Setup

For the experiment that will investigate different factors to satisfy the requirements of use case, we mounted the HoloLens on a tripod, which is fixed on an ATV for robust test repeatability (see Fig. 2). The ATV is moving according to different tests and we record position data of the localisiation system of the HoloLens. To determine the absolute accuracy we use a reference localisation system called OptiTrack [15]. This is a motion capturing system that is able to track specific markers in its area of interest and determine their poses with an accuracy in sub millimetre range. This data is also recorded at the same time the data of the HoloLens is captured. In that way we gain comparison of both systems on the same time basis. We defined an area of 6 by 6 meters, where the ATV (and the mounted HoloLens equipped with an OptiTrack marker) can move around. More precisely, the ATV is driving on a defined track using a line following navigation method. On two opposite locations we placed calibrated markers on the ground, which the HoloLens uses to locate itself in the same coordinate system as OptiTrack. The ATV has a tripod mounted on top and the HoloLens is fixed on the tripod and facing the ground ahead. The ATV can heighten and lower its top lid to adjust the level of the HoloLens. This represents various heights of people wearing the HoloLens in a range of 70 cm to 190 cm working on the ground or walking on the shop floor. Within this area it is possible to track the HoloLens at all times.

C. Agent Implementation

To point out the relation between worker and ATV and to understand the advantages of such a system, we want to explain the multi-agent system as follows. There is an agent running on the HoloLens representing the human with its features and abilities (to integrate the user in the MAS and transfer the calculated absolute pose to vehicle) and an agent representing the ATV (to access functionality of lifting LHD to adjust height and read and write the actual absolute pose). Both agents communicate with each other. This implementation allows for the calculation of the relation of the absolute poses of the two entities and, since the convenient interaction height for item hand over is given by the features of the human agent on the HoloLens, it allows for direct individual height adaption of the LHD of the robot.

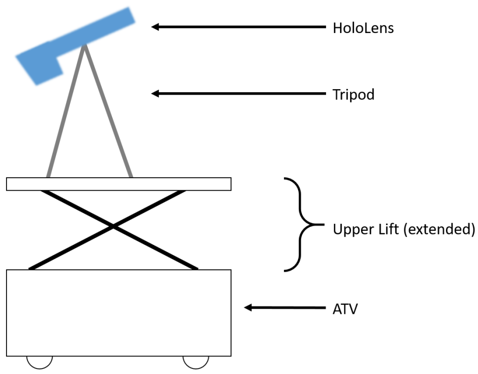

Fig. 2. ATV with extended LHD and tripod with mounted HoloLens.

D. Conduction of Experiment

To verify or refute the statements from the beginning of this section we developed multiple tests and recorded data for evaluation accordingly. These tests are listed in Table I. We vary the conditions rotational direction (CW = clockwise; CCW = counter clockwise), the visibility of the code for pose determination (visible = the code is visible multiple times during the test drive for pose correction; start = the code is only visible at the start for initial pose determination), the drive mode (track = the ATV is following the circular path; turning = the ATV is turning on the spot; still = the ATV stands still) and the LHD (low = the LHD is lowered; high = the LHD is extended; high/low = the height is altered from low to high and vice versa - a continuous motion of the LHD). Before the tests started we calibrated the reference system for the whole measurement volume. The OptiTrack software states an accuracy of less than 0.7 mm by offering 120 fps. The data of the HoloLens can just be derived every at a rate of 10 fps.

In tests 1 to 8 we let the ATV follow an optical path in a circular shape inside the tracking volume of the OptiTrack system. Here we vary direction of rotation, visibility of the reference code and the position of the LHD. In tests 9 and 10 the ATV stands still, recognizes the reference code once in the beginning and we vary the position of the LHD. In test 11 and 12 the ATV is turning clockwise or counter clockwise on the spot while only recognizing the reference code on start-up. Test 13 is the same as test 4, only we added more features to the ground using multiple highly detailed random pictures along the path. Test 14 is equal to test 10, only we took more data over time. Test 15 is the repeated test 14, only we lifted the ATV on a 80 cm high table (motion from LHD retracted to LHD extended and vice versa) and started the test.

TABLE I
TEST MATRIX

Test	Rotation	Code	Drive Mode	Load Handling Device
1	CW	visible	track	low
2	CCW	visible	track	low
3	CW	start	track	low
4	CCW	start	track	low
5	CW	visible	track	high
6	CCW	visible	track	high
7	CW	start	track	high
8	CCW	start	track	high
9	-	start	still	low
10	-	start	still	high/low
11	CW	start	turning	low
12	CCW	start	turning	low
13	CW	start	track	low
14	-	start	still	high/low
15	-	start	still	high/low

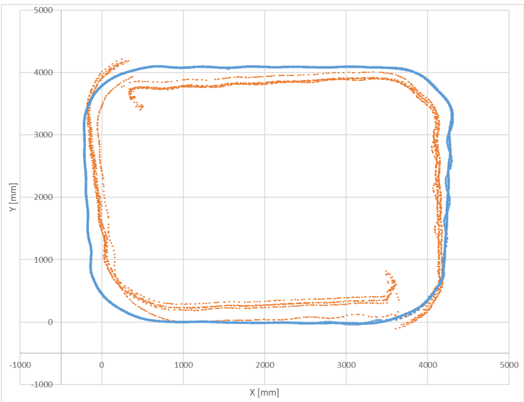

Fig. 3. Multiple reference markers in the process (OptiTrack blue; HoloLens red).

Related to the last question, the tests showed in some cases distance of more than +/- 8 mm between HoloLens and reference system which is not accurate enough for fine localization e.g. item distinction, but sufficient for coarse localization.

E. Results and Discussion

Referring to section III question 1 the direction of rotation does not influence the accuracy since test 3 (CW; standard deviation (SD) X = 12.5 mm and Y = 14.0 mm) and test 4 (CCW; SD X = 14.9 mm and Y = 16.8 mm) show a maximum difference of 2 mm comparing the standard deviations.

To answer question 2 test 1 (reference marker is seen at two locations for each run) and test 3 (marker is only seen at the start) are evaluated. As one can see in Fig. 3 the recordings of test 1 do not match very well (distances of more than 200 mm in some regions) and around the reference markers there are larger distances. In contrast to Fig. 4, which shows the results of test 3, one can see that the distance between reference system and HoloLens is quite low (euclidean distance approximately 50 mm). The error for test 1 is 285 (RMSE) and for test 3 the error is 32 (RMSE). Therefore, it is not necessary or is - like our results show - contra-productive to use more reference markers.

For determining the influence of the height of the HoloLens (question 3), we have conducted the tests 10, 14 and 15 and evaluated for each test case the height of the HoloLens during testing (e.g. see Fig. 5) as well as the absolute distance between reference system and HoloLens (e.g. see Fig. 6). Throughout all the test recordings, one can see that the data of the HoloLens height are very stable if the LHD is retracted (e.g. see readings 551 to 601 in Fig. 5 and Fig. 6) or extended fully (e.g. see readings 2181 to 2201 in Fig. 5 and Fig. 6) and there is no movement of the LHD. Further, while the ATV is lowering its LHD again the accuracy is increasing (e.g. see readings 1 to 501 in Fig. 5 and Fig. 6). There are more outliers if the LHD is almost at its lowest point as well as once it is just starting to heighten. Despite the outliers and although the overall accuracy is very low at the retracted LHD position, shortly while reaching the lowest point or leaving it there is a large improvement in the accuracy. Since, the LHD might be slightly unstable when moving, an influence of the height - especially when comparing the total retracted with the total extended position - can be found. While reaching or leaving the lowest position the HoloLens might recognize more features which then lead to the sudden changes in the accuracy. Having a look at the distance values themselves, one can see that the total deviation between reference system and HoloLens value is just 2 cm. For working with a robot with an extended height, these 2 cm might not have any influence on the ergonomic position for the worker. More important is that the height can be used for adapting the robot to the individual needs of each human worker.

To answer question 4 we evaluated test 9. we had the HoloLens record data over a timespan of 5 minutes while the ATV was neither moving nor the LHD was moved. The reference code was only recognized once at the beginning of the test. The mean value for X-direction is 2473.28 mm (SD=1.28), for Y-direction 4442.29 (SD=0.75) and for the

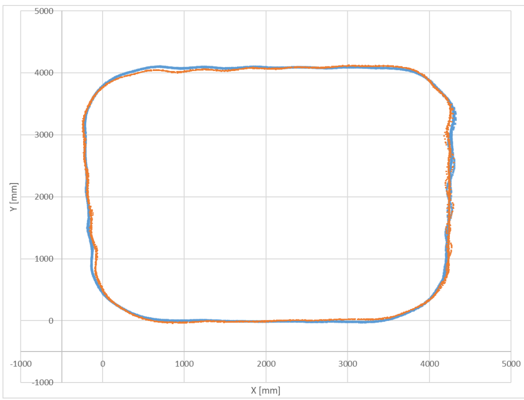

Fig. 4. Reference marker only at start (OptiTrack blue; HoloLens red).

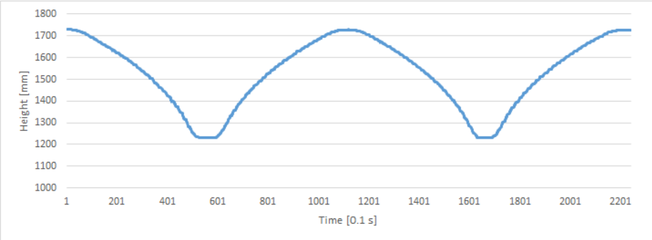

Fig. 5. HoloLens height for test 15.

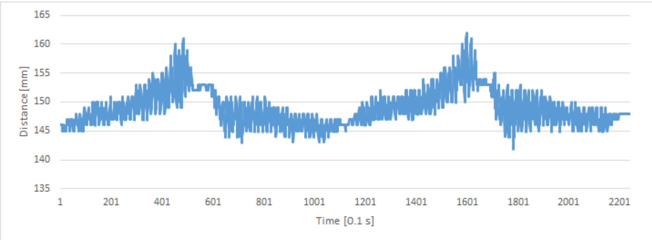

Fig. 6. Absolute Distance for test 15.

height 604.27 (SD=1.49). The standard deviations show only minimal changes over the whole time span. As one can see there is a small change over time concerning the absolute pose (see Fig. 7 for X-values, 8 for Y-values and 9 for height values). There is a continuous increase for X-values (4 mm), a continuous decrease fo Y-values (3 mm) and a continuous increase for height values (6 mm) in 5 min of the test run.

Concerning question 5 we evaluate tests 1 to 8, where the ATV is line-following the path clockwise and counter clockwise. The assumption that the accuracy is reduced when cornering is refuted. In the beginning of our evaluation, we compared the distance between OptiTrack data and HoloLens data (see Fig. 10). Just before cornering, one of the X- or Y-values were increasing (e.g. reading from 101 to 221), while cornering the accuracy retained the same level but showing outliers. Soon we figured out that although we tried to match both coordinate systems through markers as good as possible, there were some translational and rotational displacements. That is why we tried to find a common origin we centralized the HoloLens and OptiTrack pointclouds ($centered_H$, $centered_O$) using their calculated mean ($mean_H$, $mean_O$). The next step is to calculate the rotation. Through inverting the transposition of tensors of the centralized HoloLens pointcloud, then dot multiplying it with the centralized OptiTrack pointclouds, results in a matrix, factorizeable with singular value decomposition:

$$H = transpose(centered_H) * centered_O \tag{1}$$

$$U, S, Vt = svd(H) \tag{2}$$

Ignoring the scaling S (because the point clouds should not be stretched in each other), the rotation can be calculated using the transpose of the factorized rotations U and Vt:

$$R = Vt^T * U^T \tag{3}$$

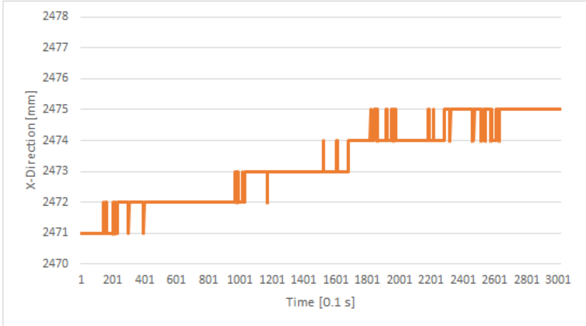

Fig. 7. Change of X-values over time.

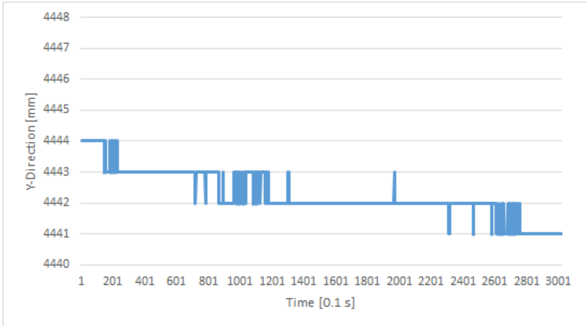

Fig. 8. Change of Y-values over time.

Using the calculated rotation R, the translation t will be calculated as below:

$$t = -R * mean_H^T + mean_O^T \tag{4}$$

One can see the data after transformation in Fig. 12. The data during a curve (e.g. readings 221 to 321) do not differ from data on a straight line (e.g. readings 321 to 501) and question 5 could not be confirmed. Since we have just tested on a parcours with four curves which have almost all the same size, further tests with more curves of different sizes might be needed to refute question 5.

We could neither confirm nor refute question 6. Although we have recorded data throughout the different scenarios (sometimes with the line for the following of the ATV visible, sometimes not visible) for determining the influence, there

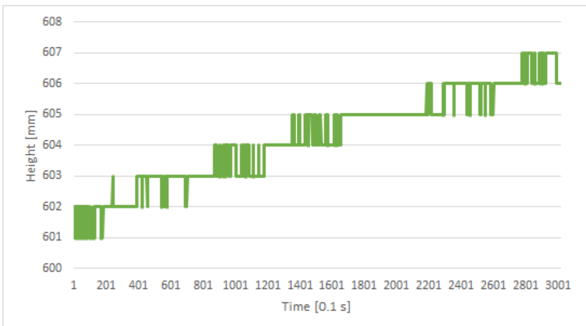

Fig. 9. Change of height values over time.

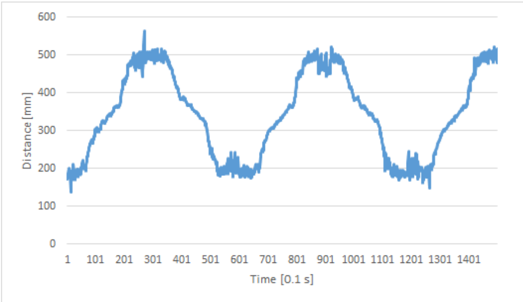

Fig. 10. Distance between OptiTrack and HoloLens - X-direction.

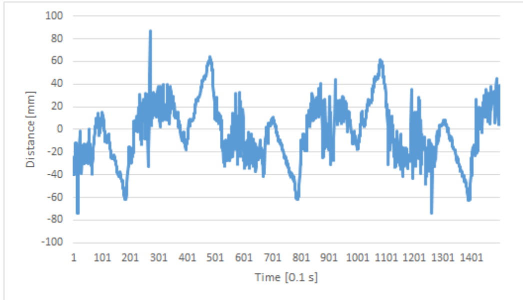

Fig. 12. Distance between OptiTrack and HoloLens after transformation - X-direction..

might be various dependencies of the condition e.g. rotation, height. Therefore, we concluded the need of another test scenario. In one test case we could use the data of the optical guided vehicle whereas in the other test case we would need data of a vehicle which is free navigating and therefore no optical line on the ground will be visible. At that moment, our vehicle with the ability to lift its LHD is not able to navigate freely. Therefore, we have to postpone the testing.

IV. CONCLUSION

The overall results of the tests we have run in the experiment show a feasible application of the HoloLens as an adequate wearable for HRI for adaptive decentralized control systems. The results for questions 1 to 5 state that there is no significant reason why rotational movement, direction of rotation, height variation and duration of use should lead to inappropriate decrease of accuracy for tracking the human. Referring to question 6 we see a slight problem, when using track guided ATVs. When using ATVs with free navigation this should not be a problem. Regarding question 7, the accuracy of the HoloLens localization method is sufficient for coarse localization since other localization systems e.g. Ubisense [16] partially do not offer better accuracy. The usage for fine localization is limited, therefore other methods e.g. markers on objects of interest are needed. In conclusion the HoloLens is useful for gaining the humans pose, tracking it and providing a calculated pose to other entities that can not create or gain their own pose - but are able to receive externally calculated location information (e.g. via services).

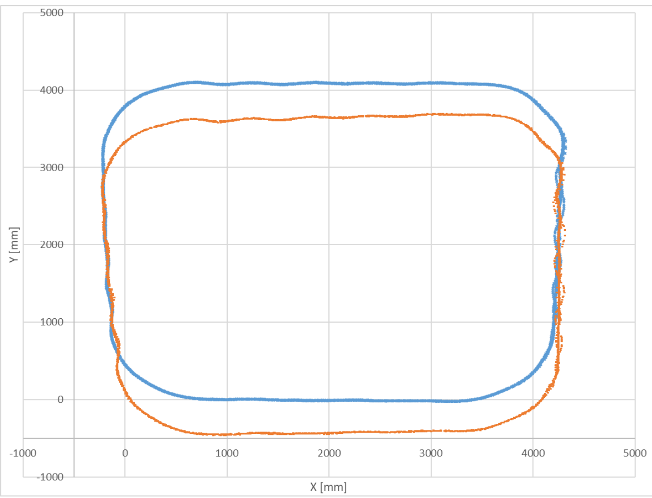

Fig. 11. Recorded data of Optitrack (blue) and non transformed data of HoloLens (red).

V. OUTLOOK

Error propagation might be critical, without referencing with markers over time. Hence, when calculating the pose from reference markers, then moving through an environment, later providing a calculated pose to entities that cannot gain their pose by themselves, then again calculating the pose of another HoloLens from this pose and so on, errors propagated in the process steps due to the influences of quality of pose recognition, tracking and pose handover. For that reason, we want to conduct another experiment, where we run the mentioned steps multiple times and find out about the increase in error for the factors of influence. Additionally, we plan to setup a more complex parcours (relating to question 5) with more curves and a longer overall path to investigate accuracy changes. Furthermore, we plan to conduct a real world scenario experiment with a worker and the ATV using the implemented agents and standard picking process with the help of HoloLens and the described methods for localization. In this process we also want to retrieve feedback regarding ergonomic height adaption of ATV in case study (referring to human factors and stress levels). Finally and since the experiments up to now took place in controlled indoor environments, we would like to test the methods also in an outdoor scenario, for example a port, where human workers drive trucks to specific locations, get off the truck and have to handle goods in inbound or outbound scenarios.

REFERENCES

[1] M. t. Hompel, M. Putz, and A. Nettsträter, "Social Networked Industry" (Whitepaper), 2019.

[2] T. Kirks, J. Jost, T. Uhlott, and M. Jakobs, "Towards Complex Adaptive Control Systems in Intralogistics", in The 21st IEEE International Conference on Intelligent Transportation Systems, IEEE, 2018.

[3] G. Ullrich, "Fahrerlose Transportsysteme: Eine Fibel", Vieweg+Teubner, Wiesbaden, 2014.

[4] Verein Deutscher Ingenieure, “VDI Norm 2510: Fahrerlose Transportsysteme”, Beuth, 1992.
[5] M. ten Hompel, T. Schmidt, and L. Nagel, “Materialflusssysteme Förder- und Lagertechnik”, Springer, 2007.
[6] H. Nikookar and R. Prasad, “Introduction to Ultra Wideband for Wireless Communications”, Springer, Dordrecht, Netherlands, 2009.
[7] M. A. Goodrich and A.C. Schultz, “Human-Robot Interaction: A Survey. Foundations and trends in human-computer interaction” Now Publishers Inc., 2008.
[8] H. A. Yanco and J. Drury, ”A Taxonomy for Human-Robot Interaction”, in 2002 AAAI fall symposium, A. C. Schultz, Eds. Nort Falmouth, Massachusetts, Menlo Park, California: AAAI Press, 2002, pp. 111–119.
[9] J. Scholtz, “Theory and Evaluation of Human Robot-Interaction”, in Proceedings of the 36th Hawaii International Conference on System Sciences (HICCS), 2003.
[10] S. Thrun, “Toward a Framework of Human-Robot Interaction”, in Human-Computer Interaction, vol. 19 No. 1, pp. 9–24, June 2004.
[11] C. Breazeal, “Toward sociable robots”, in Robotics and Autonomous Systems 42, 2003, pp. 14–23.
[12] A. Kamagaew, J. Stenzel, A. Nettstraeter, and M. ten Hompel, “Concept of cellular transport systems in facility logistics,” in The 5th International Conference on Automation, Robotics and Applications, Dec 2011, pp. 40–45.
[13] S. Poslad, “Specifying protocols for multi-agent system interaction,” ACM Transactions on Autonomous and Adaptive Systems, vol. 2, no. 4, Nov., Article 15 ,2007.
[14] P. Iñigo-Blasco, F. Diaz-del-Rio, M. C. Romero-Ternero, D. Cagigas-Muñiz, and S. Vicente-Diaz. 2012. “Robotics software frameworks for multi-agent robotic systems development.” Robot. Auton. Syst. 60, 6 (June 2012), 803-821.
[15] Optitrack, “Optitrack Motion Capture System”, 2019. [Online]. Available: http://optitrack.com/. [Accessed: 10-Apr-2019]
[16] Ubisense, “Ubisense”, 2019. [Online]. Available: https://www.ubisense.net/. [Accessed: 10-Apr-2019]