Creating Realistic Human Model Motion by Hybrid Motion Capturing Interfaced with the Digital Environment

Jochen Bönig[1*], Jerome Perret[2], Christian Fischer[1], Holger Weckend[1], Florian Döbereiner[1], and Jörg Franke[1]

[1]Institute FAPS
University of Erlangen-Nuremberg
Erlangen, 91058, Germany

[2]Haption GmbH, Aachen
Haption S.A.
Soulgé-sur-Ouette, 53210, France

ABSTRACT

A key challenge in production engineering projects is to achieve significant time and cost savings through early validation of manual assembly operations. Nevertheless, the use of digital human models for dynamic analyses is not very prevalent because of the high modeling complexity in the digital environment: with existing simulation tools, the worker's motions are either unrealistic or too time-consuming to program. Hence, further research is needed for developing a time-saving and realistic human motion simulation. In this article, we present an experimental setup for the early validation of manufacturing tasks through interactive simulation. We use a new hybrid motion-capture system interfaced with the digital environment, which facilitates the generation of realistic human model motion in real time. The software platform used is DELMIA V5. The article describes the relationship between optical and inertial tracking, and how the drift of the inertial sensors can be compensated by using a kinematic chain with a human model. Sequences of postures can be saved, both for the human model and tools, and later replayed synchronously. Finally, we detail the use of our setup in a real-world scenario within automotive manufacturing. This article acts as a practical contribution to simulation-based Manufacturing Ergonomics and Human Factors, illustrating the effectiveness of state-of-the-art technology for viable cost and time savings.

1. INTRODUCTION

In the industry, the product engineering process (PEP) underlies a continuous improvement process. Global aim is speeding up the PEP as well as forward displacement respectively parallelization of single parts of the process which again results in a reduction of the time to market. The basis therefore is certainly the concurrent engineering [1]. To realize such approaches software solutions are required that support the PEP. One of the key factors is design-for-assembly. Beyond the reduction of the time to market, it enables to carry out in a virtual environment assembly validations which were previously based on physical prototypes. A consequential advantage is the possibility to mount the product virtually during the design phase. Thus, serious drawbacks in the product field can be eliminated very early in the design phase, which in turn leads to a considerable saving of costs. Going one step further, it is possible to take the human factor into account within early stages of the engineering process. Virtual studies of manual assembly procedures can be complemented by digital human models, using standardized working environments and tools. By doing this, the virtual validation becomes even more meaningful.

Virtual validation of manual assembly operations is not very prevalent, because of two reasons. On the one hand, it takes about 300 seconds to generate one second of keyframe human model animations[13]. On the other hand, the result is a non-natural movement. [2] By connecting digital human models via motion capturing to an actor, it is possible to control the human model (then called "avatar") in real time [3]. In addition, real assembly tools and simplified mock-ups are often used, so that the actor can have some feedback on his actions. One approach to deal with such mixed reality situations is Smart Hybrid Prototyping [4]. The combination of mixed reality and optical motion capturing leads often to static or dynamic occlusion, in consequence of physical mock-ups (PMU) in the capture volume (CV). The detection of targets by cameras is sometimes blocked by the PMU during manual assembly. To avoid occlusion we decided to use a hybrid motion capture system, which uses the more accurate optical as master

* Corresponding author: Tel.: (0049) 9131-8528314; Fax: (0049) 9131-302528; E-mail: jochen.boenig@faps.uni-erlangen.de

system and, during occlusion, an inertial sensor system. The problem of inertial tracking is the drift of inertial sensors. However, this drift can be compensated by using a kinematic chain with a human model. The digital factory software DELMIA V5 offers many standard analysis tools for the virtual validation of manual assembly tasks. Therefore, DELMIA V5 is very suitable as basis software. Based on CAD data, the validation system has the advantage to use only one data structure. Sophisticated data preparation, like in VR systems, is no longer required. Detected errors and ergonomic findings during simulation can easily be corrected or optimized in the CAD system. To combine validation software and hybrid motion capturing, the Real Time Interface for DELMIA (RTID) from Haption is used.

The movements of the avatar generated by the motion capture system can be saved in catalogues of the software environment DELMIA V5. As a consequence, the companies have the possibility to save their assembly sequences in a realistic way and to reuse them in future analyses. These catalogues have to be compiled systematically in order to identify or retrieve clearly the necessary static postures or movement sequences. If a certain repertoire of catalogues exists, it can be used without the need of the motion capture system for every validation. The integration of the human factor into an early stage of validation is hence workable.

This case study demonstrates a method to create realistic animations for digital human models in real time within a mixed reality setup while avoiding occlusion through PMU's inside the capture volume. Additionally we introduce a way to store the generated animations for further use and therefore enhance the efficiency of the developed system.

2. EXPERIMENTAL SETUP

The experimental setup consists of a hardware and a software part. The transmitter and receiver of the capture data are described in the hardware chapter. It is focused on the infrared cameras and passive targets for optical tracking and the WLAN antenna respectively the dongles and the sensors for inertial tracking. On the software side, the data flow is described from the controller via server to the client and in detail on the three single systems connected via Ethernet.

2.1. HARDWARE SETUP

The hybrid motion capture system from Advanced Realtime Tracking combines optical tracking and inertial sensors to get the position of a target referring to a global coordinate system. The person in the CV wears a hybrid target suit of 17 targets. Each of these targets consists of four reflectors and one inertial sensor. For the optical tracking, the cameras send out a synchronized infrared-(IR)-flash. The retro-reflecting spherical markers reflect the IR-radiation into the direction of the incoming light. At least two cameras must receive the reflected signal to calculate the six degrees-of-freedom-(6DOF)-orientation of the targets [3]. If less than two cameras receive a signal, the inertial sensor takes place. The sensor measures rotations along three axes, which are transferred over a WLAN antenna connected to a computer [5]. Inertial tracking has the advantage that occlusion cannot occur, given that sensor and transducer are integrated and from a so called Inside-In system [6]. The used sensor unit consists respectively of a three-axis micro electro mechanical system (MEMS) sensor for acceleration, a three-axis MEMS gyroscope for angle speed and a three axis magneto-resistive sensor for terrestrial magnetic field. The influence of temperature on the sensors is leveled by an integrated thermo element [7]. For data transmission and communication between the single sensors a 2.4 Hz RF transceiver and a 16 bit microcontroller are utilized. The functionality of the gyroscope is described in [8], the manner of functioning of the acceleration sensor in [9]. For the data transmission to the controller a WLAN antenna is used. Over USB hub, two ten-channel dongles are connected with the antenna. Every inertial sensor is mounted with his serial number to the dongle for communication. The sensors are mounted to the two dongles with a channel distance of ten and a splitting of eight to seven. The channel distance of ten guaranties a good performance, since the dongles cannot share a channel. Additionally a third tracking technology is used for the hand finger system. A data glove is integrated working on the principle of strain gauges. The data transfer is realized via Bluetooth directly to the Server.

2.2. SOFTWARE SETUP

Information processing is allocated to three platforms (see Figure 1), because of the high need of processing power. The distributed architecture consists of ART Controller, Client and Server. The data from the glove is directly transferred to the Server, whereas the information signals from cameras and inertial sensors respectively WLAN antenna are sent to the Controller. On the Controller the data is processed, bundled and transferred to the receiving software DTrack2 of ART. IPSI Server synchronizes the data of DTrack2 over the Plug-In RTID with DELMIA V5. The visualization signals are sent via DELMIA to desktop monitor, HMD and 3D projector. The communication of the distributed architecture works over Ethernet. The Controller receives for example the tracking data and changes it into readable coordinates, the data is provided on both the server and client for disposal. Hereby IPSI demands the data automatically and provides it for other application. To save resources of Client and Controller the Server undertakes the physics and collision computation using the IPSI Server software communicating with DELMIA via Ethernet.

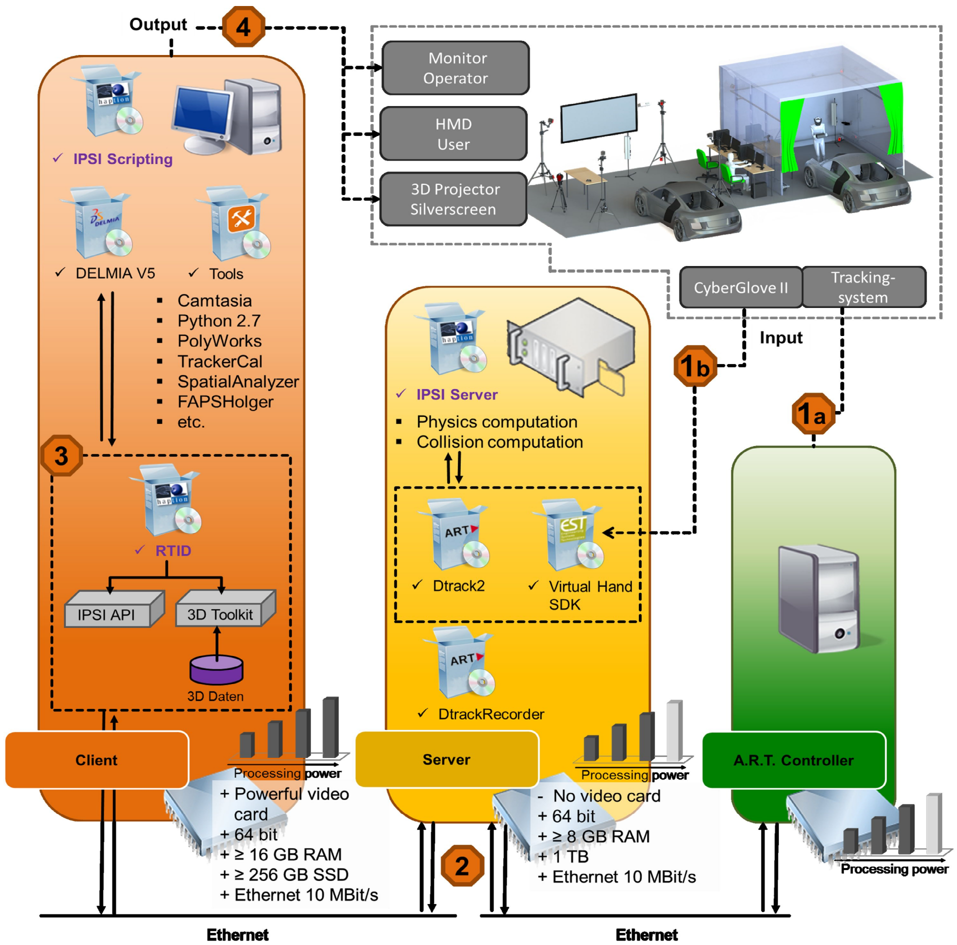

Figure 1. Soft- and hardware setup.

3. SIMULATION WORKFLOW

Before the physical activity of an actor can be gathered, a virtual environment needs to be compiled with DELMIA. This virtual environment should reflect the real conditions of the assembly operations to be examined. The modules as well as their working tools are deposited in the product and resource list of DELMIA. The uploaded elements will be positioned in the virtual environment so as to readjust the initial situation of the assembly operation. In doing so the necessary scope for the assembly must not exceed the extent of the capture volume. At this point it is helpful to construct the floor space of the process according to the dimensions of the capture volume. Consequently while positioning the elements it can already be assured that the actor does not leave the tracking area during the assembly. If the intention is to work virtually only then the compilation of the virtual environment is completed at this particular time. Hence the initialization of the motion capture system can be started. With the introduced system it is possible to work not only virtually. It is also feasible to use additional physical mock-ups to start the assembly operation. Hence the actor can perform the mounting procedure directly at the real object while his movements are being tracked. To enable this operation the position of the real objects in the capture volume needs to be adjusted to the position of the virtual objects. This is also called the registration of objects. A good accuracy is indispensable to obtain the illusion of the coexistence of real and virtual environment [10].

To realize the registration of objects the original coordinate systems of the real and the virtual environment need to coincide. In order to achieve this, a retrievable point needs to be defined in the real environment. This retrievable point represents the coordinate system of the real environment. After having finished the compilation of the virtual environment, the simulation of the movements can be started via Haption RTID. Subsequently to the fixing of the target coordinate systems to the virtual human model, the actor can commence the assembly operation and the performed movements can be recorded with RTID. It requires always two persons to operate the motion capture system. One person represents the actor, which puts the target set of the A.R.T tracking system, the cyber glove and the head-mounted display on so as to perform the necessary mounting procedures in the capture volume. The other person represents the operator of the system: he is responsible for the adjustment of the computer, the initiation of the movements and he supervises the perfect functionality of the system. While the actor puts the necessary hardware on, the operator starts the system in a certain succession to display the movements of the actor in DELMIA and to connect them with the virtual human model. To achieve this aim the following practice needs to be observed:

LOADING THE VIRTUAL ENVIRONMENT:

The virtual scenario is loaded via the standard DELMIA toolbar as process environment. If the registration between virtual and real objects is not yet conducted, this procedure has to be done now. Again it is important to ensure the execution of the room calibration as well as the definition of the spawn point for the digital human model in accordance to the tracking system. The next step is the generation of an initial state. As a result, all object positions including the manikin are saved in a process file and can be reset to their defined start-positions.

CONFIGURATION OF THE DIGITAL HUMAN MODEL:

Before transferring the motion data to DELMIA, the configuration files of the used manikin have to be selected during adding the manikin to simulation. The files contain information about the used data input systems and the mapping between target coordinate systems and extremities of the human model. Besides the loading of the configuration file, the target reference has to be defined as manikin base. If collision detection is requested additionally, the collision partners have to be selected and therefore meshed with voxels for the detection algorithm.

DEFINITION OF THE BOUNDARY CONDITIONS:

After manikin configuration, the simulation involved products and resources have to be defined. Therefore the options *active* and *movable* are available. Only active objects are available for any kind of interaction, e.g. for collision detection. Movable objects have to be defined as such ones.

STARTING THE REAL-TIME SIMULATION:

After having finished the simulation generation, the real time visualization of the tracked coordinate data is activated. The positions of the single targets in the capture volume are now displayed as moving frames in the virtual environment. If the actor in the capture volume moves, the movements are transferred in real time to the corresponding frames of the virtual environment.

ATTACHING THE TARGET FRAMES TO THE MANIKIN:

The actor moves in the capture volume to bring the frames in correlation with the respective limbs of the manikin. In this case, particular emphasis needs to be put on the positioning of the hand-targets. After the frames are correctly positioned, the actor must maintain the current stance, while the desktop operator performs the attachment of the frames to the manikin. Thus, the target frames are attached to the manikin and the movements of the actor are transmitted in real time to the virtual human model (see Figure 2).

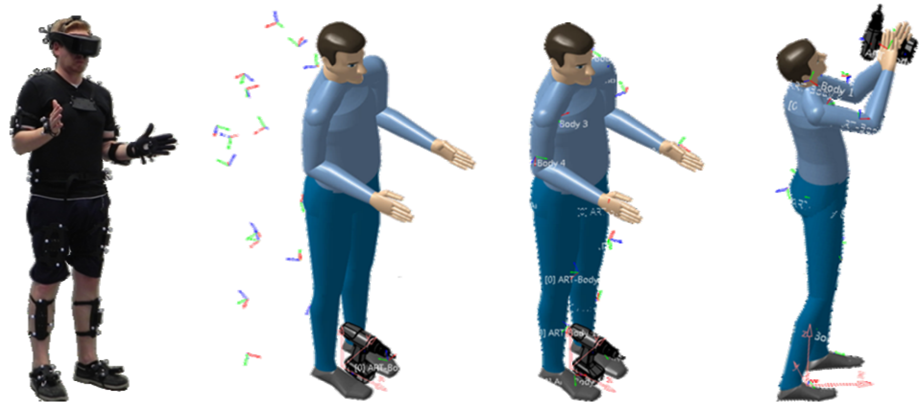

Figure 2. Attaching the target coordinate systems to the manikin.

FIRST PERSON PERSPECTIVE FOR HEAD-MOUNTED DISPLAY:

Following to the attachment, the perspective must be changed to the viewing angle of the manikin. Then the head of the manikin must be hidden to avoid misstatement. In case the head of the manikin is not hidden, it is possible that the inside of the manikin head is briefly visible during rapid movements. This can lead to severe disorientation of the actor.

After the start of simulation and the adjustment of viewing angle perspective, recordings of the movements of the actor can be created with little effort. The recording frame-rate can be adjusted up to the speed of the physics simulation. Nevertheless, a value of 15 frames per second is sufficient for a high-quality animation. The recorded movements are saved in a process file and deposited in the tree below the manikin as so-called tasks. The hereby obtained task can be used directly for the analysis of the assembly process. Another possibility is to isolate the recorded movements and to save them separately in catalog structures. In these catalogs, they are always available for reuse.

Initially, the described practice seems very complex. However, provided that the virtual scenario has already been created and the registration of all physical objects is concluded, the initial set-up time for the system lies at about 15 minutes. This includes the start-up time of the hardware as well as the fitting of the hybrid target suit to the actor. Subsequently it is possible to record any of the actor's movements in real time, rendering the time consuming manual generation of human model animations unnecessary.

4. HYBRID (OPTICAL/INERTIAL) TRACKING

The actor is equipped with a set of targets, which combines optical markers and accelerometers. As discussed earlier, the tracking system delivers very accurate positions and orientations for the optical markers, but is subject to occlusions. On the contrary, the data coming from the accelerometers have a low resolution and a tendency to drift, due to the two steps of numerical integration needed in order to generate angles. Last but not least, the accelerometers deliver only orientations, and no positions.

When occlusions occur, they are usually neither immediate nor total. It means that some individual markers will still be visible, although the tracking system fails to identify them as a complete target. In an ideal system, the positions of individual markers could be combined with the data coming from the accelerometers. This can be done with simple filtering methods, and in this case a Kalman filter would be a very good choice. However, in our case we don't have access to the raw data; indeed the DTrack2 software preprocesses the data before streaming it out, and delivers either the position and orientation coming from the optical markers, or the angles integrated from the accelerometers, depending on whether there is an occlusion or not. If applied directly to the avatar, it has the following effect: whenever an occlusion occurs, the affected segment freezes in position, but continues to rotate around its center.

In order to overcome that difficulty, we introduced a special filter in the avatar control software. We took advantage of our real-time physics simulation software, which includes a force-driven solver for kinematic chains [11]. In this approach, the position/orientation data for each segment of the avatar is compared with the previous location, producing position/orientation errors. The errors are multiplied by control gains (stiffness and damping) in order to compute force/torque values, which are then inserted into the kinematic solver. The solver applies a numerical integration scheme very comparable to the Jacobian transpose method and obtains new values for the joint angles, which tend to minimize the errors. Our filtering algorithm consists in setting to zero the value of the position gains when an occlusion is detected, so that the force applied to the affected segment is zero. Because of the drift in the accelerometers, it is necessary to compute an offset at the same time, in order to ensure a smooth transition. When the occlusion vanishes, the orientation offset is set back to zero and the gains are reset to their normal values (figure 3).

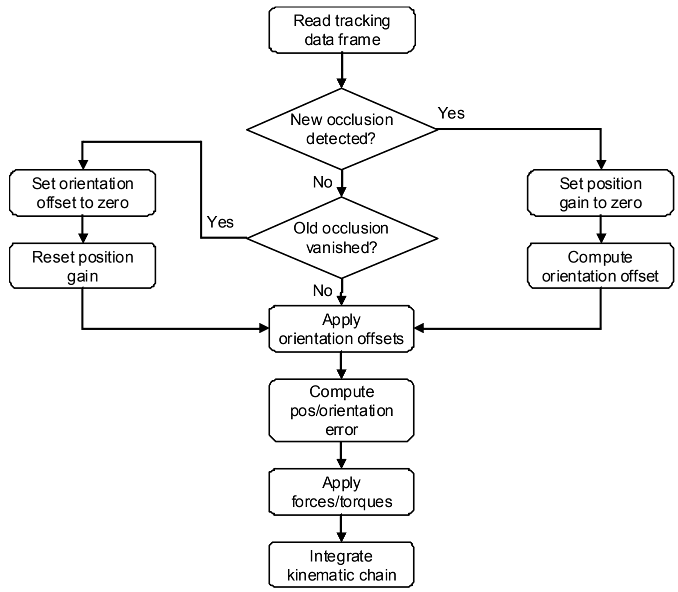

Figure 3. Hybrid tracking filtering algorithm.

Since not all segments are affected by occlusions simultaneously, the kinematic chain always receives some forces, which are sufficient to drive the joints of the avatar. Typical cases are the hands and feet of the actor, which are occluded by the PMU. As long as the optical targets on the arms and legs can be seen, the actors keeps full control of the avatar's hands and feet, both in position and orientation.

5. CATALOGS

Despite the advantages of a motion capture system, the user does not want to start it each time during assembly planning and virtual validation. Hence it would be nice to archive some typical body postures and movements for later usage. This solution is feasible within the simulation software DELMIA V5 which offers the saving of avatar's postures in catalogues. Also his movements generated by the motion capture system can be saved in catalogues of the software environment DELMIA V5. As a consequence, the companies have the possibility to durably get their specific assembly sequences in a realistic way and to use them in future analyses. These catalogues have to be structured systematically in order to identify or to retrieve clearly the necessary static postures or movement sequences [12]. If a certain repertoire of catalogues exists, it can be consulted without the need of the motion capture system for every hedging. The integration of the human factor into an early stage of hedging is hence workable. The user obtains a realistic and natural movement of the digital human model within DELMIA V5 without a big time consumption for teaching. Reasons for this are the realistic acceleration of body parts and a higher resolution of single postures for a sequence (15 frames per second) compared to the teaching of a digital human model. In such a way the production engineer can also consider difficult assembly situations in which the worker strangely behaves because of obstacles in the assembly path or heavy tools. Here the need of company specific libraries for postures gets clear as they regard the case oriented assembly cells, environments and products. Several assembly sequences of the catalogue can be connected in an arbitrary order and extended by single postures. If more assembly situations will be necessary the catalogues can be modularly extended by applying the motion capture system again.

6. Experiments

The experimental setup is built up for manual assembly validation. Our experiments are done on an automotive specific environment. The assembly task is placing a power electronic unit in a car front. The occlusion problem in mixed reality situations is solved with the hybrid motion capturing system (see Figure 4). This allows us to conduct the experiments using a real car body. Up until now it was necessary to construct complicated PMU's to replicate the assembly situation and simultaneously avoid occlusions. With the experimental setup we can place any type of assembly setup inside the capture volume and record the actor's movements while he performs the necessary assembly tasks using the actual components and tools.

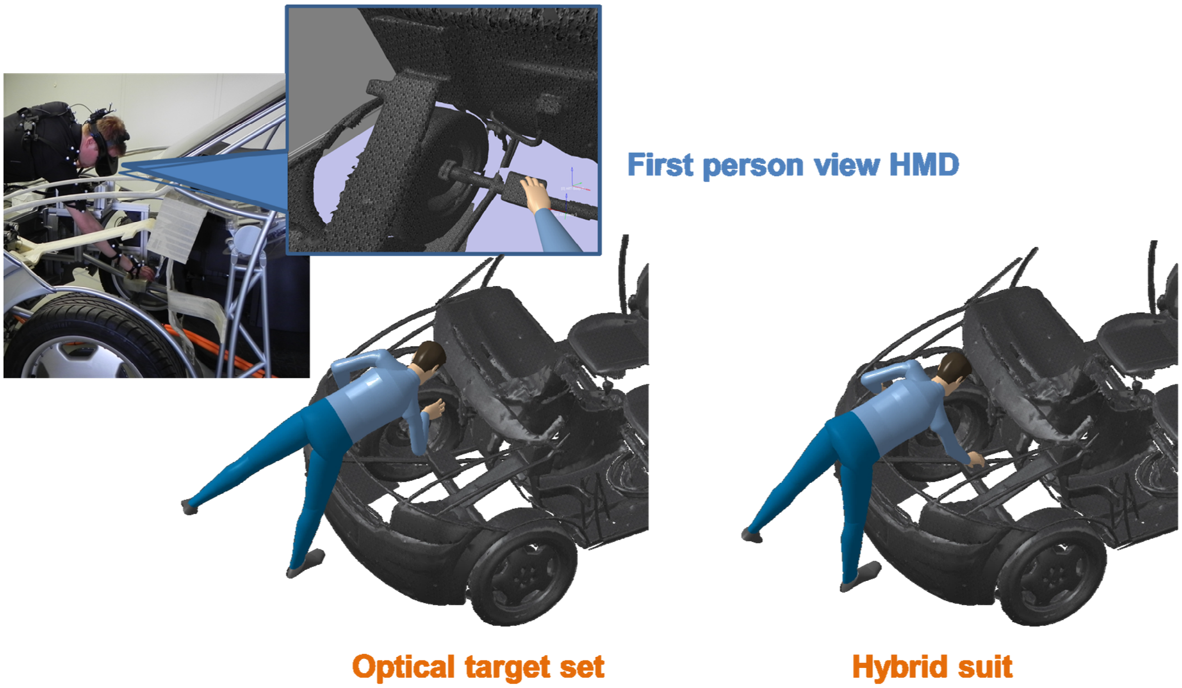

Figure 4. Test scenario.

The picture shows that the purely optical system does not represent the realistic human arm position. The cameras do not detect the hand target anymore, whereby the arm posture of the last detected position is displayed. The hybrid system provides the correct data and thus a realistic picture of the human posture. Consequently we are able to investigate different kinds of assembly problems in real time and with realistic results referring to the motion of the digital human model.

7. Summary

In this study, creating realistic human model motion by hybrid motion capturing is introduced. DELMIA V5 is the basis system for manual assembly simulation. This has several basic advantages over standard VR systems. One key benefit is the direct use of CAD data for simulation. Elaborated data conversion of updates is waived for the virtual assembly analysis. Another benefit lies in real time recording of an actor's movements. With this feature it is possible to create realistic animations of human motion without the need for time consuming manual teaching. Moreover, the generated animations can be saved in catalogues for further use, thus eliminating the necessity to conduct multiple recordings for similar assembly tasks. Last but not least the developed system utilizes inertial sensors to substitute capture data, should the optical tracking be impaired by occlusions. This permits the use of actual components, tools and working environments for digital assembly validation, as a consequence rendering the construction of task specific physical mock-up's unnecessary.

ACKNOWLEDGEMENTS

This research and development project was funded by the German Federal Ministry of Education and Research (BMBF) within the grant 16N12036. The authors are deeply grateful for this support and are responsible for the contents of this publication.

REFERENCES

[1] *H.-J. Bullinger and A. Bading, Forschungs- und Entwicklungsmanagement: Simultaneous engineering, Projektmanagement, Produktplanung, rapid product development. Stuttgart: Teubner, 1997.*

[2] *J. Bönig, C. Fischer, V. Marquardt, S. Matzka, J. Franke: "Methodical Integration of Assembly Specific Influences concerning High-Voltage Components into the Virtual Validation Process", In: J. Franke (ed.) 2nd International Electric Drives Production Conference. EDPC 2012. Nuremberg, October 16–17, 2012, pp. 50–57. IEEE Service Center, Piscataway, NJ (2012)*

[3] *J. Bönig, C. Fischer, M. Brossog, M. Bittner, M. Fuchs, H. Weckend, J. Franke: "Virtual validation of the manual assembly of a power electronic unit via motion capturing connected with a simulation tool using a human model", In: M. Abramovici, R. Stark (eds.) Smart product engineering. Proceedings of the 23rd CIRP Design Conference, Bochum, Germany, March 11th-13th, 2013, pp. 463–472. Springer, Berlin, London (2013)*

[4] *M. Auricht, B. Beckmann-Dobrev, R. Stark: "Frühzeitige multimodale Absicherung virtueller Prototypen. Evaluation am Beispiel einer PKW-Heckklappe", ZWF 107(5), 327–331 (2012)*

[5] *Trivisio Prototyping GmbH, Bedienungsanleitung Colibri - Inertialer Bewegungssensor, Trier: Trivisio Prototyping GmbH, 2011.*

[6] *M. Riedl, Potential eines virtuellen Fahrerplatzmodells in der Fahrzeugkonzeptentwicklung, München: TU München Lehrstuhl für Ergonomie, 2012.*

[7] *Trivisio Prototyping GmbH, Colibri Wireless Specification Sheet, Trier: Trivisio Prototyping GmbH, 2013.*

[8] *F. Cacchione, "Mechanical characterization and simulation of fracture processes in polysilicon Micro Electro Mechanical Systems (MEMS)," Dipartimento di Ingegneria Strutturale, Polytechnikum Mailand, Mailand, 2007.*

[9] *G. Vijila, S. Vijayakumar, and A. M. A. Gupta, "Design and Analysis of 3D Capacitive Accelerometer for Automotive Applications."*

[10] *R. T. Azuma: A Survey of Augmented Reality. In: Presence: Teleoperators and Virtual Environments 6, 4 (August 1997), S. 355–385, 1997.*

[11] *A. Rennuit, A. Micaelli, X. Merlhiot, C. Andriot, F. Guillaume, N. Chevassus, D. Chablat and P. Chedmail, "Passive control architecture for virtual humans", International Conference on Intelligent Robots and Systems - IROS, 2005*

[12] *T. Gudehus: Entwicklung eines Verfahrens zur ergonomischen Bewertung von Montagetätigkeiten durch Motion-Capturing. Kassel Univ. Press, Kassel, 2009.*

[13] *H. Hoffmann; M. Dangelmaier; R. Schirra: Effiziente Produktionsplanung in der virtuellen Realität mit Menschmodellen und Motion Capturing. In: D. Spath: Fraunhofer-Institut für Arbeitswirtschaft und Organisation -IAO-, Stuttgart: Technologiemanagement in der Praxis: Forschen und Anwenden, 25 Jahre Fraunhofer IAO, Fraunhofer IRB Verlag, Stuttgart, 2006.*