Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 00 (2016) 000–000

www.elsevier.com/locate/procedia

Changeable, Agile, Reconfigurable & Virtual Production

Presenting a Novel Motion Capture-based Approach for Walk Path Segmentation and Drift Analysis in Manual Assembly

Philipp Agethen[a,*], Michael Otto[a], Felix Gaisbauer[a], Enrico Rukzio[b]

[a]Daimler AG, Wilhelm-Runge-Str. 11, 89081 Ulm, Germany
[b]Ulm University, James-Franck-Ring, 89081 Ulm, Germany

* Corresponding author. Tel.: +49-731-5052426; fax: +49-711-3052156294. E-mail address: philipp.agethen@daimler.com

Abstract

Automotive industry is currently facing the challenge to cope with the market demand for mass-customization whilst remaining competitive. In production planning, this trend towards product-diversification leads to a rising complexity, since growing numbers of variants are hitting mixed-model assembly lines. Due to these changing preconditions, traditional planning models and respective simulations tend to decreasingly reflect reality. Actual manual assembly processes can deviate significantly from their corresponding plans due to simplified assumptions of simulation models, methods and tools. In order to contribute to a better prediction quality of planning models, this paper investigates walk paths in real assembly situations with regard to their deviation from corresponding plans. A novel algorithm set for walk path reconstruction and neural network based classification of work tasks is introduced. Therewith, data gathered by a mobile tracking setup can be automatically segmented and subsequently assigned to the process plans. This novel approach enables an assessment of predetermined assembly times by comparing reference to real walk paths. The method's technical performance is verified in laboratory evaluation scenarios and its applicability is proven in a productive automotive final assembly line during operation.

Keywords: Time determination; monitoring; motion capture; automotive assembly; neuronal networks; walk path; segmentation; operator drift

1. Introduction

In the automotive industry, manual assembly planning is becoming increasingly complex since the demand for product variants with shorter life-cycles is growing continually. As a consequence of this development, traditional planning methods tend to decreasingly reflect reality since simplifying assumptions are leading to higher uncertainties within these models. For instance, previous work indicates that walk paths (see [1]) and process times (i.e. [2,3]) of real assembly operators can deviate significantly from their corresponding process plans. In this context, it is crucial to review planning models with regard to their applicability in order to unveil inconsistencies between planned and real processes, especially when considering walk paths in mixed-model assembly flow lines. Therefore, an efficient method for investigating actual task execution times of manual assembly processes is a valuable tool towards a better planning performance.

This paper presents a novel approach to analyze the disparity between planned and actual assembly times of manual processes based on the operator's walk paths.

Therefore, an optical tracking system is utilized to unobtrusively record the routes of each assembly operator within a particular work place. These trajectories are subsequently processed and segmented via a neural network approach. These resulting subtrajectories are being matched and assigned to the process plan afterwards. Therewith, the actual assembly times can be compared with predetermined process plans for each cycle.

The remainder of the paper is structured as follows: First, the state of art in the context of assembly time determination is reviewed. Second, a non-intrusive approach is presented being able to compare predetermined and actual assembly times for each product using multiple distributed depth cameras. Subsequently, the applicability and technical performance of the proposed method is verified within two holistic evaluations,

2212-8271 © 2016 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the scientific committee of the Changeable, Agile, Reconfigurable & Virtual Production Conference.

the former verifying the neuronal network algorithm classification performance and the latter assessing the overall applicability of this novel spatial disparity (drift) calculation approach. The paper concludes with an outlook on further optimization potentials and the resulting impact on future planning paradigms.

2. Assembly time in mixed-model assembly flow lines

Time determination systems are commonly used in industrial production systems for organization, planning and efficiency appraisals [4]. The resulting data sets are utilized among others for development, calculation, incentive systems, production program planning, production sequence and manufacturing planning. Within manual assembly planning they are commonly used to assign times to assembly processes, which are previously defined in workflow descriptions. The latter enables planners to optimize geometric and time-based interplay of the operator, workpieces, resources, material, energy and information within a working system [5].

Time management systems have brought up multiple methods for work time determination, which can be clustered in target time determination and actual time determination:

2.1. Determination of reference assembly times and predetermined motion time systems

Target times are defined as times, which have been derived from previously captured actual times [6]. This includes the category of "predetermined motion time systems" (PMTS), which are based on large studies with determined and fixed influence factors. They cluster basic work tasks in a tabular form [7] and are used to assign reference assembly times or so-called target times to workflow descriptions. Multiple standards exist to plan assembly times:

In Europe [8] commonly used PMT-systems are "methods-time measurement" (MTM) [9], MTM-UAS or various company proprietary systems such as "C-values". In NAFTA region, MODAPTS [10] technique is wide-spread in automotive industry as well. Depending on the manufacturing type, repetitiveness, cycle time and type of workplace, the used planning method abstraction level highly differs. For example, highly repetitive, monotonous mass-production with short cycle times can be planned via MTM-1 on a low abstraction level, whereas for custom-made products higher-aggregated and therefore simplified planning methods, such as MTM-UAS or MEK are applied.

Limitations: Whilst target time and predetermined motion time systems inherit advantages in terms of simplicity, feasibility, planning speed and harmonization, they also have some drawbacks in contrast to actual time determination methods.

Analyzing human work is especially difficult since it is highly flexible, statistically distributed and varies to a large-scale [2,3]. Modeling statistic variances can help enhancing model and simulation quality, dealing with variance and optimizing the overall planning quality. Using reference assembly times, this variance is neglected. This could lead to unexpected errors during operation in production as well as to an overexertion of assembly operators. Therefore, reference assembly times have to be compared to actual shop-floor data in order to verify their validity.

2.2. Determination of actual assembly times

In contrast to PMTS and target time determination, actual times are defined to be real times which humans or resources need to execute certain process steps [6]. Various research has already been carried out in the field of determination of real assembly times (see [4,7,8]), ranging from methods of self-documentation over direct to indirect measuring principles. According to Deuse and Busch [7], direct measuring approaches, in which a third person or sensor is gathering temporal data from the shop-floor, are usually applied when analyzing manual assembly lines. In most cases, the execution time is determined using stopwatches [8], whereas recent approaches (see [11]) utilize sensors like inertial measurement units or RFID-Tags (i.e. [12,13]) being attached to parts or the human body.

Limitations: Even though these methods are frequently used in practice, they neglect important influence factors. Generally, the manual documentation of process times using stopwatches lacks of objectivity since the unambiguous determination of particular actions (i.e. screwing) in human motion is error-prone and depends on the respective person (see 4.1). Additionally, commonly used RFID approaches enable binary process monitoring on abstract level only. These approaches only monitor binary process acknowledgements during production.

In contrast to the state of the art, in the following chapter a method will be introduced which produces highly detailed information on spatiotemporal relations within a production environment to gain deeper insights on the processes itself and to overcome drawbacks of the previously mentioned systems.

3. Markerless Motion Capture for assembly time analysis

This paper presents a non-intrusive approach to automatically determine the disparity between planned and actual assembly times of manual processes based on the operator's walk paths. In order to implement such a system, a common definition of trajectories is given in the following.

3.1. Concept

Following Buchin et al. (see [14]) a discrete trajectory τ is defined as "a mapping from a series of time stamps $t_0, t_1, \dots, t_n$ to the plane (or a higher-dimensional space)." For any timestamp t_i, the location in the plane at time t_i is denoted by $\tau(t_i)$. For any two times $t_i, t_j \in \{t_0, \dots, t_n\}$ with $t_i \le t_j$, the subtrajectory of τ from time t_i to time t_j is defined as $\tau\left[t_i, t_j\right]$ [14].

Based on this definition, Figure 1 depicts the concept of analyzing the spatial distance between the reference and real walk path in order to deduce the underlying time gap in the course of time. Assuming that the operator drift x_{Diff} represents the vector between a corresponding point pair being projected onto the conveying vector, the delta-time t_{Diff} results from equation 1.

$$\Delta t_{Diff} = \frac{x_{Diff}}{v_{Conveyor}} = \frac{\left(\tau_{Plan}(t_i) - \tau_{Real}(t_j)\right)_{Conveyor}}{v_{Conveyor}} \quad (1)$$

As automotive end-assembly is typically carried out in paced assembly lines, consisting of connected conveyor belts with a constant speed, $v_{Conveyor}$ is known in advance. Therefore, the spatial distance $\left(\tau_{Plan}(t_i), \tau_{Real}(t_j)\right)$ between a corresponding point pair contains the cumulated time difference stemming from the disparity between plan and reality. Moreover, when observing two consecutive cycles, it is possible to determine the particular temporal gap for each car within the production sequence.

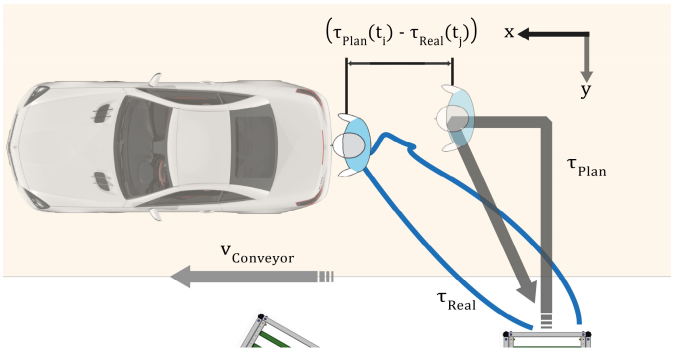

Figure 1. Spatial difference between the planned and real walk paths in a manual assembly line.

When implementing a temporal assessment approach based on walk paths, two main tasks can be identified: First, the implementation of a robust walk path reconstruction system and second, the determination of corresponding point pairs between planned and actual trajectories.

3.2. Walk path reconstruction

The presented method utilizes an optical Motion Capture (MoCap) system (see [15]) consisting of multiple distributed Kinect depth cameras. Using this approach stakeholders are enabled to robustly gather anonymized skeletal data from each assembly operator [15]. Skeletal data represents an estimation of the human body's posture and position in three-dimensional space over time. Due to the redundant sensors, this system is able to cover large working areas up to several workplaces. Moreover, since no additional markers or sensors have to be attached to the operator, it can virtually be ruled out, that the non-intrusive working principle affects the operator's movements or productivity.

Based on this tracking system, the gathered skeletal data are used to obtain the trajectory τ_{Real}, describing the operator's two-dimensional motion on the plane during the observed time frame. For this purpose, the center of mass is approximated within every frame by the center of the hip, spine and shoulder joints (see [16]) and subsequently projected onto the ground plane [1]. Finally, the trajectory is reconstructed by linking the resulting vectors for each point in time $t_i \in \{t_0, \ldots, t_n\}$ to a concatenated walk path [1].

3.3. Segmentation

Reference walk paths τ_{Plan} are usually planned using PMTS [7]. As pointed out in chapter 2.2, however, these predetermined tasks times can deviate from their actual execution times. As a result, the planned time intervals of assembly operations can significantly differ from reality on the shop-floor. For example, on one occasion an assembly operator might complete working on a particular car ten seconds after the scheduled time because he was held up with the previous task due to unexpected excess work, whilst on another occasion he might be able to start ten seconds before. Consequently, the timestamps t_i and t_j describing a corresponding point pair in τ_{Plan} and τ_{Real} have to be determined independently in order to deduce the delta-time according to equation 1 – which linearly corresponds to the spatial operator's drift due to fixed $v_{Conveyor}$ speed.

An accurate method of obtaining these coinciding planned and real timestamps t_i and t_j in both trajectories is to detect distinctive actions (i.e. screwing, fetching components or returning to a certain position). In this paper the transitions between walking and interacting with the environment are chosen as the criterion to detect corresponding point pairs since the spatial starting and ending points of assembly tasks are independent of execution times and temporal offsets between plan and reality.

Based on this approach, a novel classifier is implemented using a neural network, which is thereby segmenting the walk path into a set of subtrajectories. The network utilizes several spatiotemporal criteria presented by Buchin et al. [14]: Velocity, angular velocity and acceleration. Since the classification based on those three input parameters cannot be assumed to be linear and the exact functional relation is not known a priori, a multi-layer neural network is chosen for this problem. The implemented two-layered feed forward neural network hereby consists of three inputs neurons - one hidden layer including three neurons and one output neuron, which serves as classification output. Due to the properties of the training data (pairs are known a priori), the supervised backpropagation learning method (see [17]) is chosen.

Based on these extracted subtrajectories and the reference walk paths, which are pre-segmented by definition, subsequently, the first and last element of each path describing operator walking are used to establish corresponding point pairs, ultimately solving equation 1.

4. Evaluation

In order to determine the performance and applicability of the proposed algorithm set, both in a laboratory environment and on the shop-floor, several experiments have been carried out.

4.1. Assessment method validation in laboratory setting

Initially, the proposed segmentation approach is examined with regard to its temporal accuracy. Therefore, a representative experimental setup is designed, that is based on characteristic processes within an assembly line.

Experimental setup: As depicted in Figure 2 the setup consists of two racks and one table. Each rack contains two boxes including a single screw on difference shelves. The pickup-points for screws 1 and 3 are at a height of 1 m, whereas screws 2 and 4 are placed at 0.6 m. Each rack has a distance of 2.5 m to the table, spanning an angle of approximately 30°. A workpiece with four screwing points is placed on the table at a height of 0.6 m. Moreover, six Kinect cameras are used during all experiments, being evenly arranged in a circle around the interaction area.

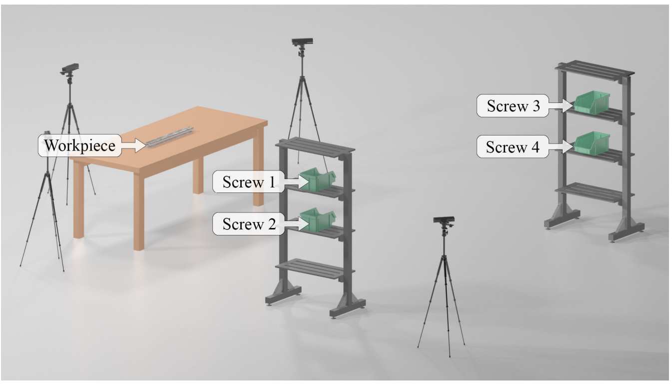

Figure 2. Experimental setup: Two racks with each two screws are placed next to a table with a workpiece on it, while six Kinects are tracking the user's movements.

Design of experiments: During the evaluation a representative set of assembly tasks are carried out twice by five production planning employees, thus generating five training and five reference sets. Within the scenario, a user is instructed to mount four components on a single workpiece representing the car in the assembly line. The operator has to successively fetch each screw, carry it to the workpiece and finally tighten it. Subsequently, these steps are repeated in reverse order, in order to disassemble the components, thus reestablishing the initial situation.

Having determined the training and reference sets, each trajectory is manually segmented in a post-processing step. The former is used to train the neural network, whereas the latter serves as baseline for assessing the proposed approach. Since several preliminary tests unveiled that the temporal determination of transitions between interacting and walking varies considerably when observing video footage of tasks, both data sets are manually segmented in a post-processing step by a group of five people. By using redundant subtrajectories from more than one participant as a baseline and input for the neural network, the subjective bias can be minimized.

Finally, the neural network is trained using an implementation of the backpropagation algorithm for 10 000 cycles at a learning rate of 0.05 with a series of approximately 90 000 input vectors consisting of velocity, acceleration and angular velocity. In order to obtain universally valid input values they are normalized using their 95th percentile. Having initialized the classifier, the algorithm is applied to the five reference sets. The results are ultimately compared with the average of the manually segmented transitions representing the baseline.

Results: Figure 3 depicts the results of this experiment. The upper section illustrates the segmented walk paths in a bird's eye view. The red lines represent the subtrajectories being assigned to "walking", whereas the black paths show areas in which the user was interacting. The lower left side depicts the distribution of absolute time differences in the manually segmented reference data set. The distribution of absolute time differences between the results of the proposed approach and the baseline can be seen on the lower right side.

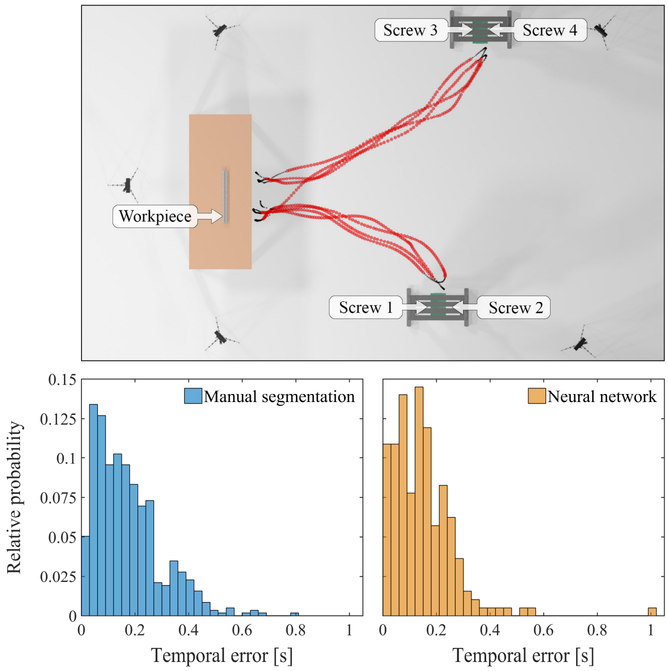

Figure 3. Top: Bird's eye view of walk paths, Bottom: Histogram of temporal errors. Left: Manual segmentation, Right: Neural network based segmentation.

Comparing the temporal divergences in the reference data set with the results generated by the neural network (see Table 1), the temporal segmentation accuracy is regarded to be sufficient since no significant differences between the benchmark and the proposed approach can be determined.

Table 1. Overall temporal divergences between the manually segmented baseline and the proposed neural network based approach.

Approach	Manually segmented	Neural network
Mean temporal error	0.17 s	0.15 s
90th percentile	0.35 s	0.27 s
Max	0.80 s	1.01 s

Furthermore, the evaluation unveils that the classifier occasionally suffers from false-positive "walking" recognitions for a short series of frames. However, these defective datasets can be suppressed automatically using a low-pass filter in combination with continuity criteria.

Having either manually or automatically annotated center of mass vectors of the respective experiment, the data has been matched to the planned correspondences in terms of start- and end-point of each transition between "walking" and "interaction". Highly accurate automatic transition detection between these two states allowed semi-automatic mapping between planned and real data. After applying the low-pass filter on the results gathered from the neural network classification, in average only one out of 20 transitions were false-positive recognitions and had to be removed manually for matching the correct work-task process.

4.2. On-site shop-floor evaluation of assessment method

The described system has been used in real production environments, in order to assess the applicability of the proposed system within the industrial environment of an automotive final assembly line

Experimental setup: The mobile motion capture setup using multiple Kinect v2 sensors has been set up in an operative shop-floor situation at a Mercedes Benz Cars assembly plant. The system was consisting of 5 cameras that covered a 6 m x 6 m large tracking frustum. The sensors were arranges, so that all of them were oriented the center of one workplace within the mixed-model paced assembly line. With this system, a given real mixed-model production sequence has been observed and captured for more than 5 hours in order to gain anonymized skeletal data. The latter was converted into trajectories in a post-processing step and afterwards annotated by the neural network using the described methods.

The work tasks in the particular station were conducted overhead with the car body bottom being at an average height of 2 m. The station work tasks included fetching, screwing and safety acknowledge tasks. The conducted assembly processes vary significantly for each product within the sequence, whereas corresponding PMTS planning predicts that the overall work load averages during the observed time frame.

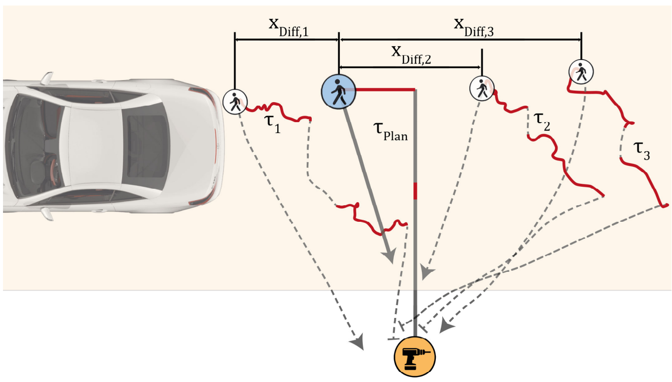

Figure 4. Three captured vs. planned trajectories: Results of the proposed segmentation an annotation approach deducing the spatial difference x_{Diff} between plan and reality.

Results: Figure 4 depicts the pre-planned walk path for the captured station with three exemplary annotated and segmented walk paths. Task with a classification of "interaction" are denoted as red lines. Captured walk paths are depicted as dashed, grey lines.

These three different examples in Figure 4 underline the findings from the previous experiments, since the classifier is able to reliably distinguish between walking and interacting. Based on these subtrajectories, the distinctive points and corresponding point pairs are established using the last transition (interacting-walking) within each cycle. The resulting spatial differences (or operator drift) x_{Diff} are exemplarily illustrated in Figure 4 . It can be seen that, that the overall walk path length depends on the current operator drift. Moreover, the drift situation can be visualized for each car within the production. Therefore, for each captured trajectory the total drift is calculated and visualized in Figure 5. This graphic depicts the course of x_{Diff} for 50 product cycles. It can be seen, that the disparity between plan and reality has a cyclical character, whereas the overall workload has an average above zero, meaning that the operator by trend does not drift out of his station and is able to complete all assigned work tasks in time. Nevertheless, it can be reasonably stated, that in this particular time frame real assembly processes deviate from their corresponding plans since Figure 5 shows spatial difference up to ± 1.5 m over a considerable timespan.

Furthermore, the evaluation unveils that the operator drift in this particular station is limited, since one of the work tasks includes screwing using a stationary tool, which has to be fetched in each cycle at the same position in the station.

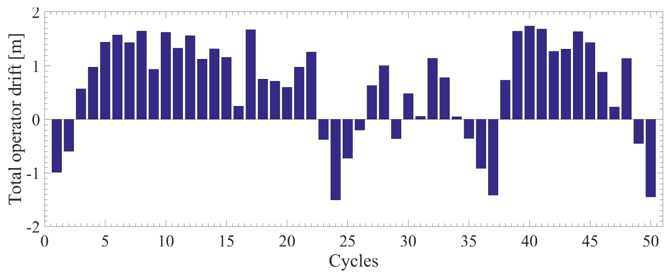

Figure 5. Cumulated total operator drift over multiple cycles times, respectively products.

Markerless motion capture system showed technical limitations, since the tracking technology sporadically suffers from false-positive recognitions due to occasional occlusion within the industrial environment. However, these defective data sets could be filtered out manually in a post-processing step.

5. Conclusion

In this paper a novel, non-intrusive approach to assess real and pre-planned walk paths is presented. Using the proposed algorithm set, production planners are enabled to assess temporal and spatial disparities between real production sequences and their corresponding plans.

In order to obtain reliable tracking data, a markerless, scalable motion capture system is proposed being able to track operator's motion. Having calculated the walk path trajectory for the operator out of the raw three dimensional skeletal data, a segmentation on distinctive actions is carried out. This process is achieved with the help of a novel classification approach using pre-trained, multi-layer neural networks. Therewith, input parameters are deducted from the operator's walking trajectory, such as velocity, angular velocity, acceleration. Subsequently, the outcomes of the two-layered feed forward neural network are utilized to determine correspondences between the reference walk path and real trajectory in order to analyze the temporal discrepancies between the process plans and reality.

In order to get insights into the technical performance of the proposed approach, a two-staged evaluation has been carried out: First, the applicability of the neural network based segmentation algorithm within a laboratory environment is presented. The resulting findings show that automatic segmentation is sufficiently to classify processes solely based on the walk path. The mean segmentation error of the neural network is 0.15 s, whereas mean manual segmentation error is 0.17 s.

Second, an on-site shop-floor evaluation has been carried out in a real automotive final assembly line production site. The proposed and pre-trained neural network is applied to data sets being gathered in an operating final assembly line. By bringing the method to a real shop-floor surrounding, the total operator drift can be calculated in the course of time.

6. Outlook

Future work will optimize and extend the novel classification algorithm to robustly segment more distinct work tasks and to classify the automotive final assembly processes at a more detailed level.

Finally, planning paradigms have to be discussed in general. Nowadays, industrial standard uses PMTS whilst novel planning approaches are technically already available. However, changing these planning paradigms to a completely new approach, leads to novel planning processes on the one hand, but also to the introduction of novel industrial regulations and clearance processes in automotive industry on the other hand. This will be discussed in further publications in detail.

7. Acknowledgements

We would like to thank the German Federal Ministry of Education and Research for co-funding this research within the ARVIDA project (ARVIDA project, http://www.arvida.de/, grant no. 01IM13001N).

References

[1] Agethen P, Otto M, Mengel S, Rukzio E. Using Marker-less Motion Capture Systems for Walk Path Analysis in Paced Assembly Flow Lines. Procedia CIRP 2016.

[2] Baines T, Hadfield L, Mason S, Ladbrook J. Using empirical evidence of variations in worker performance to extend the capabilities of discrete event simulations in manufacturing. Simulation Conference, 2003. Proceedings of the 2003 Winter, vol. 2, 2003, p. 1210–6. doi:10.1109/WSC.2003.1261552.

[3] Folgado R. Heterogeneity and Variability on Human-Centered Assembly Systems. PhD Thesis. Instituto Superior Técnico, Universidade de Lisboa, 2012.

[4] Schlick CM, Bruder R, Luczak H. Arbeitswissenschaft. Berlin, Heidelberg: Springer Berlin Heidelberg; 2010.

[5] DIN EN ISO 6385. Ergonomic principles in the design of work systems 2004.

[6] REFA. REFA Methodenlehre der Betriebsorganisation, Datenermittlung. München: Fachbuchverlag Leipzig; 1997.

[7] Deuse J, Busch F. Zeitwirtschaft in der Montage. In: Lotter B, Wiendahl H-P, editors. Montage in der industriellen Produktion, Berlin, Heidelberg: Springer Berlin Heidelberg; 2012, p. 79–107.

[8] Picker, Christoph. Prospektive Zeitbestimmung für nicht wertschöpfende Montagetätigkeiten. Lehrstuhl Für Fertigungsvorbereitung 2007. doi:10.17877/DE290R-883.

[9] MTM Association for Standards and Research. Website MTM Association n.d. http://www.mtm.org/ (accessed May 30, 2016).

[10] MODAPTS - The Language of Work n.d. http://www.modapts.org/ (accessed May 30, 2016).

[11] Stoppuhr im Ärmel - Mediendienst Januar 2012 - Thema 3. Fraunhofer-Gesellschaft n.d. http://www.fraunhofer.de/de/presse/presseinformationen/2012/januar/stoppuhr-im-aerml.html (accessed May 30, 2016).

[12] Zhang Y, Jiang P, Huang G, Qu T, Zhou G, Hong J. RFID-enabled real-time manufacturing information tracking infrastructure for extended enterprises. Journal of Intelligent Manufacturing 2012;23:2357–2366.

[13] Shibata T, Tsuda T, Araki S, Fukuda K. RFID-based production process monitoring solutions. NEC Tech J 2006;1:77–81.

[14] Buchin M, Driemel A, van Kreveld M, Sacristán V. Segmenting trajectories: A framework and algorithms using spatiotemporal criteria. Journal of Spatial Information Science 2011;2011:33–63.

[15] Otto M, Agethen P, Geiselhart F, Rukzio E. Towards ubiquitous tracking: Presenting a scalable, markerless tracking approach using multiple depth cameras. Proc of EuroVR 2015 (European Association for Virtual Reality and Augmented Reality) 2015.

[16] Gabel M, Gilad-Bachrach R, Renshaw E, Schuster A. Full body gait analysis with Kinect. 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2012, p. 1964–7. doi:10.1109/EMBC.2012.6346340.

[17] Dinesha V. NeuronDotNet - Neural Networks in C#. Sourceforge 2013. https://sourceforge.net/projects/neurondotnet/ (accessed May 22, 2016).