Automatic Capturing and Analysis of Manual Manufacturing Processes with Minimal Setup Effort

Sönke Knoch
German Research Center for Artificial Intelligence (DFKI)
Saarland Informatics Campus, Saarbrücken, Germany
soenke.knoch@dfki.de

Frederic Kerber
German Research Center for Artificial Intelligence (DFKI)
Saarland Informatics Campus, Saarbrücken, Germany
frederic.kerber@dfki.de

Vladimir Pavlov
German Research Center for Artificial Intelligence (DFKI)
Saarland Informatics Campus, Saarbrücken, Germany
vladimir.pavlov@dfki.de

Shreeraman Ponpathirkoottam
German Research Center for Artificial Intelligence (DFKI)
Saarland Informatics Campus, Saarbrücken, Germany
shreeraman.ponpathirkoottam@dfki.de

UbiComp/ISWC '16 Adjunct, September 12–16, 2016, Heidelberg, Germany
ACM 978-1-4503-4462-3/16/09.
http://dx.doi.org/10.1145/2968219.2971368

Abstract
The ongoing development of industrial manufacturing towards more individualization and smaller lot sizes opens up a new range of challenges. Not only do the processes in the factories need adaptation, but the workers need more support as well. We showcase a system that is able to address both aspects: an instrumentation of a manual workplace provides direct feedback for planning engineers, while at the same time acquiring data that is helpful for giving the worker feedback. Within this demo we focus on bi-manual picking and assembly processes observed by a lightweight optical recognition system enhanced by ultrasonic sensors, but also give an outlook on other possible modules.

Author Keywords
Industrie 4.0; manual assembly; timekeeping; production planning; ergonomics

ACM Classification Keywords
H.4.m [Information Systems Applications]: Miscellaneous; J.6 [Computer-aided Engineering]: Computer-aided Manufacturing (CAM)

Introduction
Recent developments in the manufacturing industry, termed Industrie 4.0 [2], aim at bridging the gap between real and virtual worlds and foster the utilization of the potential which

the Internet of Things provides. Further decreases of lot sizes and increasing customer demands for individual products require new approaches, such as multi-variant manufacturing lines with very short to nonexistent without setup times, which simultaneously introduce a new level of complexity. Consequently, production managers increasingly need IT support to keep track of the manufacturing process. Additionally, workers demand more assistance, e.g. to cope with often-changing variants or multiple available materials.

Our system supports both groups: for the production planners, more realistic timing information than provided by today's routine methods, such as Methods-Time Measurement (MTM), provides a direct benefit. For the workers, live feedback about their movement patterns, e.g. shown in a digital worker guidance system, can possibly help to improve their working process. Long-term learning effects could thereby result in a permanent improvement, even if the system is not installed anymore. Although both types of information could already be derived from highly instrumented environments or by instrumenting the worker with a sensor suit, the special aspect of our system is its lightness. It enables utilization whenever and particularly wherever it is needed. As soon as the problem is resolved, there is no need to further capture the process, which is important in terms of privacy and employee acceptance.

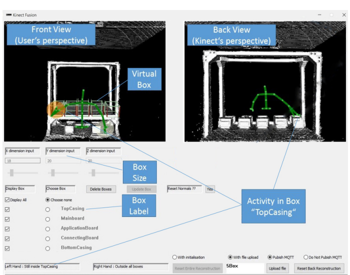

Figure 1: Assembly station editor to easily create a model representation of the workplace.

Related Work

The process of (automatic) capturing of manual manufacturing processes is closely related to the field of human activity analysis (see [1] for an overview). To track the physical activities and movement patterns of workers, different approaches can be taken. At one extreme, full-body sensor suits (e.g. [6]) could be employed. Although they might provide precise information, they are also a possible hindrance for the wearer. In the middle of the spectrum, systems that use a combination of wearables and environmental sensors can be found. For example, Stiefmeier et al. made use of such a hybrid system to track the activities in a car production scenario [4]. At the other extreme, approaches that solely rely on environmental instrumentation can be found. Mostly, vision-based systems are employed here (see e.g. [5] for an overview). Based on the captured images, movements can be analyzed and depending on the setup, skeletal information can also be retrieved.

Concept

A modular concept is used to provide the required components without complicating the system (and its setup phase) with unnecessary parts. All communication follows an event-driven approach based on the publish-subscribe paradigm. Through complex event processing, information from several modules can be combined on a meta level to aggregate atomic sensor events and add semantics. Due to limited space, we focus on the components that will be an integral part of our demo, and outline the possible extensions of our system only at the end.

Assembly Station Editor

A prerequisite for our modules dealing with material arrangements, grasp distances or tool handling is to have a digital representation of our assembly station as well as positioning information for the tools and materials. As stated above, we focus on a lightweight system that does not require long setup phases or use by specialists. Therefore, we created a simple assembly station editor (see Figure 1) that even laymen can use to build a virtual model of the assembly station and its components such as the material boxes or tools. In a first step, a 3D model of the assembly station itself is created through an automatic capturing process. Around each individual component, a virtual box can be placed by simply clicking on a point within the model.

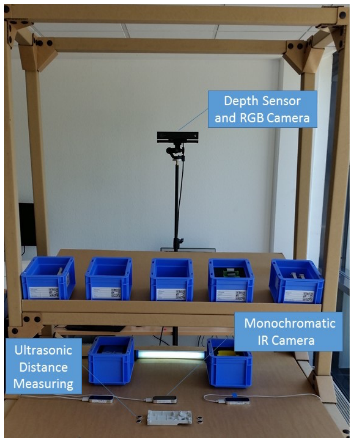

Figure 2: Instrumented cardboard assembly station.

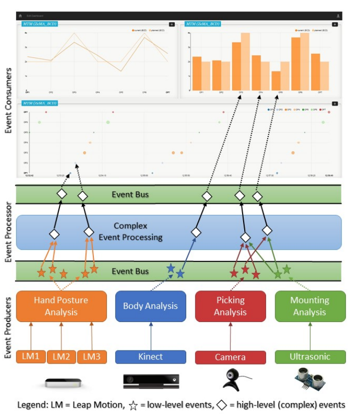

Figure 3: Overview of our event-based architecture.

Handling Analysis
The handling analysis module can detect which virtual box a worker is reaching into and with which hand, by utilizing a single depth camera. Ultrasonic sensors placed on both sides of the workpiece position provide further information about time periods spent assembling the product.

Body Analysis
By means of the depth camera, a body posture analysis is also supported. With live feedback, workers can be made aware of behavior with possibly bad (long-term) effects. As the system is designed for short term use, we see it as temporary help for the workers to think about their movements instead of acting as guardians for them.

Additional Modules
The *Hand Posture Analysis* module recognizes the hand posture when material is grasped. It also detects which hand was involved and distinguishes between grabbing and picking, e.g. taking single or multiple screws. The *Picking Analysis* module uses a camera to observe material boxes. It can again be detected which box was targeted and depending on the material, it can also be analyzed whether the correct amount was taken out. Our complex event processing enables us to combine several low-level events into suitable high-level events. We provide this information via an event bus so that it is available to downstream components, such as for *visualization applications*.

Implementation

Figure 3 shows an overview of our event-based architecture. We distinguish event producers (sensors), our event processing layer and event consumers (applications). The architecture's loose coupling of modules allows an easy integration and removal of modules, while event rules can be modified without affecting the remaining system.

Instrumenting the Assembly Station
To follow our idea of a lightweight system, we decided to focus on technologies that do not require instrumentation of workers and are able to sense activities contact-free. Thus, solely marker-free optical and ultrasonic sensors were used on the event producer level. To demonstrate our system, we set up a cardboard assembly station (see Figure 2) and offer assembly instructions for a small product that can be assembled with only a few steps. The assembly station is observed by a single depth camera (Microsoft Kinect 2) and further equipped with two ultrasonic sensors attached to a .NET Gadgeteer embedded system to track the working area in detail. For the capturing process of the assembly station in our editor, we utilized KinectFusion. The *Hand Posture Analysis* is based on Leap Motion sensors and the data collected by them is published over TECS[1], Thrift-based communication middleware available for multiple programming languages. The respective data processing was then implemented in Java. The *Picking Analysis* module employs a standard off-the-shelf USB camera and uses OpenCV for picture processing.

Body Analysis
The *Body Analysis* has been realized with the Kinect 2 sensor observing 25 fulcrums at 30 fps. Activities are detected by observing the position of a sphere with a radius of 10 cm centered around each hand position, respectively.

Hand Posture Analysis
To analyze a hand's posture, seven feature points are observed: one per finger, plus the palm and wrist positions. The system learns a gesture such as picking or grabbing in a recording phase. During recognition, the Cosine Similarity is used to compare recorded and detected actions.

[1] http://tecs.dfki.de/, last accessed July 24th, 2016

Picking Analysis

Similar to the Kinect system, the user arranges one or multiple rectangles directly in the camera's image. All activities within these labeled rectangles are then recognized. Thus, multiple areas can be observed depending on the width of the respective camera's angle. We utilize background subtraction based on the Gaussian mixture distribution and filter out shadows to avoid false positives. The background model is thereby based on the last 500 frames and detects an activity if the alteration within the box is higher than 50%, allowing a stable hand detection.

Event Processing

Based on the information our modules sense, events are generated; e.g. if someone picks an item from a box, we generate a (complex) event. All events are provided in JSON format using the MQTT-Publish-Subscribe protocol. The complex event processing (CEP) engine APAMA[2] processes all events that occur on the MQTT level. In APAMA, event patterns can be defined and processed that form the input for upper-level applications.

Example Visualization

Based on the Collaborative Electronic Performance Board (CEPBoard) by Pavlov et al. [3], a system that allows the creation and maintenance of dashboards, we created a dashboard in the event consumer layer (see Figure 3). The upper two widgets provide a comparison between MTM and our measured time values. It enables the analyst to detect deviations between the MTM catalogue times and the times measured by the system. The widget below shows the latest operation times in a bubble chart. At a glance, the analyst can see if operations were left out, executed in order, or detected several times (e.g. due to wrong pick operations).

[2]www.softwareag.com/us/products/apama_webmethods/analytics/, last accessed July 24th, 2016

Acknowledgments

This research was funded in part by the German Federal Ministry of Education and Research (BMBF) under grant number 01IS16022E (project BaSys4.0), 01IS13015 (project SmartF-IT), and 01IS12050 (project ProMo4ProM).

REFERENCES

1. Jake K. Aggarwal and Michael S. Ryoo. 2011. Human Activity Analysis: A Review. *ACM Comput. Surv.* 43, 3, Article 16 (April 2011), 43 pages.

2. Henning Kagermann, Johannes Helbig, Ariane Hellinger, and Wolfgang Wahlster. 2013. *Recommendations for Implementing the Strategic Initiative INDUSTRIE 4.0: Securing the Future of German Manufacturing Industry; Final Report of the Industrie 4.0 Working Group*. Forschungsunion.

3. Vladimir Pavlov, Sönke Knoch, and Matthieu Deru. 2015. CEPBoard: Collaborative Electronic Performance Board and Editor for Production Environments in Industry 4.0. In *Proc. PerDis '15*. ACM, 239–240.

4. Thomas Stiefmeier, Clemens Lombriser, Daniel Roggen, Holger Junker, Georg Ogris, and Gerhard Troester. 2006. Event-Based Activity Tracking in Work Environments. In *Proc. IFAWC '06*. 1–10.

5. Daniel Weinland, Remi Ronfard, and Edmond Boyer. 2011. A Survey of Vision-Based Methods for Action Representation, Segmentation and Recognition. *Computer Vision and Image Understanding* 115, 2 (Feb. 2011), 224–241.

6. Felix Wenk and Udo Frese. 2015. Posture from Motion. In *Proc. IROS '15*. IEEE.