Manufacturing Letters 21 (2019) 45–49

Contents lists available at ScienceDirect

Manufacturing Letters

journal homepage: www.elsevier.com/locate/mfglet

A self-aware and active-guiding training & assistant system for worker-centered intelligent manufacturing

Wenjin Tao [a,*], Ze-Hao Lai [a], Ming C. Leu [a], Zhaozheng Yin [b], Ruwen Qin [c]

[a] Department of Mechanical and Aerospace Engineering, Missouri University of Science and Technology, Rolla, MO 65409, USA
[b] Department of Computer Science, Missouri University of Science and Technology, Rolla, MO 65409, USA
[c] Department of Engineering Management and Systems Engineering, Missouri University of Science and Technology, Rolla, MO 65409, USA

ARTICLE INFO

Article history:
Received 15 January 2019
Received in revised form 4 July 2019
Accepted 11 August 2019
Available online 14 August 2019

Keywords:
Intelligent manufacturing
Deep learning
Augmented reality
Cyber-physical system
Smart manufacturing

ABSTRACT

Training and on-site assistance is critical to help workers master required skills, improve worker productivity, and guarantee the product quality. Traditional training methods lack worker-centered considerations that are particularly in need when workers are facing ever-changing demands. In this study, we propose a worker-centered training & assistant system for intelligent manufacturing, which is featured with self-awareness and active-guidance. Multi-modal sensing techniques are applied to perceive each individual worker and a deep learning approach is developed to understand the worker's behavior and intention. Moreover, an object detection algorithm is implemented to identify the parts/tools the worker is interacting with. Then the worker's current state is inferred and used for quantifying and assessing the worker performance, from which the worker's potential guidance demands are analyzed. Furthermore, onsite guidance with multi-modal augmented reality is provided actively and continuously during the operational process. Two case studies are used to demonstrate the feasibility and great potential of our proposed approach and system for applying to the manufacturing industry for frontline workers.

1. Introduction

Cyber-Physical Systems (CPS) have allowed the traditional manufacturing to enter into a new era, which is currently further boosted by Artificial Intelligence (AI) technologies, such as machine learning and deep learning, towards intelligent manufacturing [1,2]. To meet the fast-growing consumer demands for highly-customized, high-quality products, manufacturers must make their manufacturing systems more flexible and efficient and, meanwhile, ensure that workers in the systems are agile and highly skilled. Workforce training and on-site assistance is essential to help workers learn desired skills, improve worker productivity, reduce the rate of rejects, and guarantee the product quality. Therefore, how to train and assist the workforce flexibly, efficiently and effectively is one of the critical factors contributing to a company's market success.

Traditionally, operational instructions are provided in a lecture-based manner or a mentor-based manner. However, these methods have some limitations. For example, the lecture-based training can simultaneously teach lots of workers but is lack of immediate interaction. While it is more interactive and can have real-time communications, the mentor-based training is more costly and inefficient. For further evaluation of the worker's performance and optimization of the operational workflow, a time-motion study is often applied. Nevertheless, it requires a direct and continuous observation of the task and manual analysis for each operational step, which is time-consuming and lack of flexibility [21].

To provide the instructional information more interactively and immersively, Augmented Reality (AR) technologies have been widely deployed in industry, and it has been proven to be an excellent interface for presenting multi-media information to workers [13–17]. However, existing AR methods often use pre-defined scripts to control how the instructional information is provided and they lack worker-centered considerations that are particularly in need when workers are facing ever-changing demands.

The limitations of existing methods have motivated us to develop a training & assistant system that can effectively improve the workforce outcomes. It is worker-centered, i.e., every element in the system is to assist the worker in achieving the best operational result. To realize worker-centered training, a necessary task is to perceive the worker's states, such as behavior and intention. There exist different kinds of sensors that have been used for this

* Corresponding author at: 311 Engineering Research Lab, Missouri University of Science and Technology, 500 W. 16th Street, Rolla, MO 65409, USA. Tel.: 573 466 3528.

E-mail address: w.tao@mst.edu (W. Tao).

https://doi.org/10.1016/j.mfglet.2019.08.003
2213-8463/

purpose [3–7]. To recognize worker activities, various methods have been applied [8–12,23]. While being aware of a worker's states during the training, necessary instructional information can be introduced to guide the worker's training with AR techniques.

This project aims to develop a self-aware, active-guiding training & assistant system for worker-centered intelligent manufacturing by exploring advanced sensing technologies, AI methods, and AR techniques. Specifically, as shown in Fig. 1, we have designed a multi-modal sensing system to sense the worker via different modalities, and have developed efficient and robust deep learning algorithms that analyze sensor data to recognize worker states. This awareness of worker state allows the system to understand the worker both physically and mentally, thus creating a basis for intelligent decision making. Finally, we have created multi-modal AR instructions that are generated according to the training decision made and provided to meet the worker's needs.

2. Worker state awareness

2.1. Multi-modal sensing system

To comprehensively perceive the worker, we developed a multi-modal sensing system illustrated in Fig. 1(b). The system is composed of both ambient and wearable sensors with different modalities. Each sensor has a unique capability in collecting specific information about the worker.

Various ambient sensors were used to capture the worker's activities in the workplace. Optic cameras were used to capture RGB images. Depth cameras such as a Microsoft Kinect or Lidar (light detection and ranging) sensors were applied to obtain data in the 3D space. Infrared cameras can detect the worker in a dark environment. Pressure sensing mats were developed to capture the standing states. Ambient sensing can collect a large amount of data without interfering the worker's activity.

Nevertheless, the complex setup and occlusion issue are main challenges in implementing ambient sensing. To compensate for these limitations, wearable sensing was applied. A smart Eyewear containing cameras was worn to perceive the surroundings from the first-person view of the worker. IMU (Inertial Measurement Unit) sensors were used to capture the movement of the worker body. sEMG (surface Electromyography) sensors were utilized to obtain the muscle activities. ECG (Electrocardiogram) sensors were used to monitor the worker's heart activities. EEG (electroencephalogram) sensors were used to collect electrical events of the human brain. All of the data were synchronized and sent to the local workstation via different transmission protocols.

2.2. Worker behavior and intention understanding

With data obtained from the above multi-modal sensing system, we developed deep learning models, such as convolutional neural networks (CNN) and recurrent neural networks (RNN), to understand the worker behavior and intention from both spatial and temporal perspectives (e.g., walking towards a workstation, turning a screwdriver, etc.). The worker intention comprises mental activities related to specific tasks such as having confidence in, or feeling confused for, a specific operation. Specifically, we explored designing different models, including vision-based, IMU-based, sEMG-based, and EEG-based deep learning models, to recognize the worker activity and mental intention.

A single sensing modality cannot guarantee robust perception under various circumstances. Therefore, we developed data fusion algorithms to take advantage of multi-modal sensing. Different sensing modalities were fused to augment individual speculations for making the final inference. The optimal fusion method was identified by comparing their overall performance.

2.3. Interacting part/tool detection

Most activities of a worker involve worker-object interactions. Detecting objects the worker is interacting with is important for understanding activities of the worker and for providing instructional information to help the worker locate desired objects. In this study, object detection algorithms, such as R-CNN [18], were implemented to recognize the interacting parts or tools in real time (e.g., Fig. 2(a)). To establish the dataset for training the algorithms, we designed a data collecting system to take pictures of the objects automatically (see Fig. 2(b)). Manually collecting data of some objects from all kinds of scales and viewpoints is difficult or inefficient. Thus, we developed a data synthesis approach to generate data directly from CAD models (see Fig. 2(c)). The CAD model of

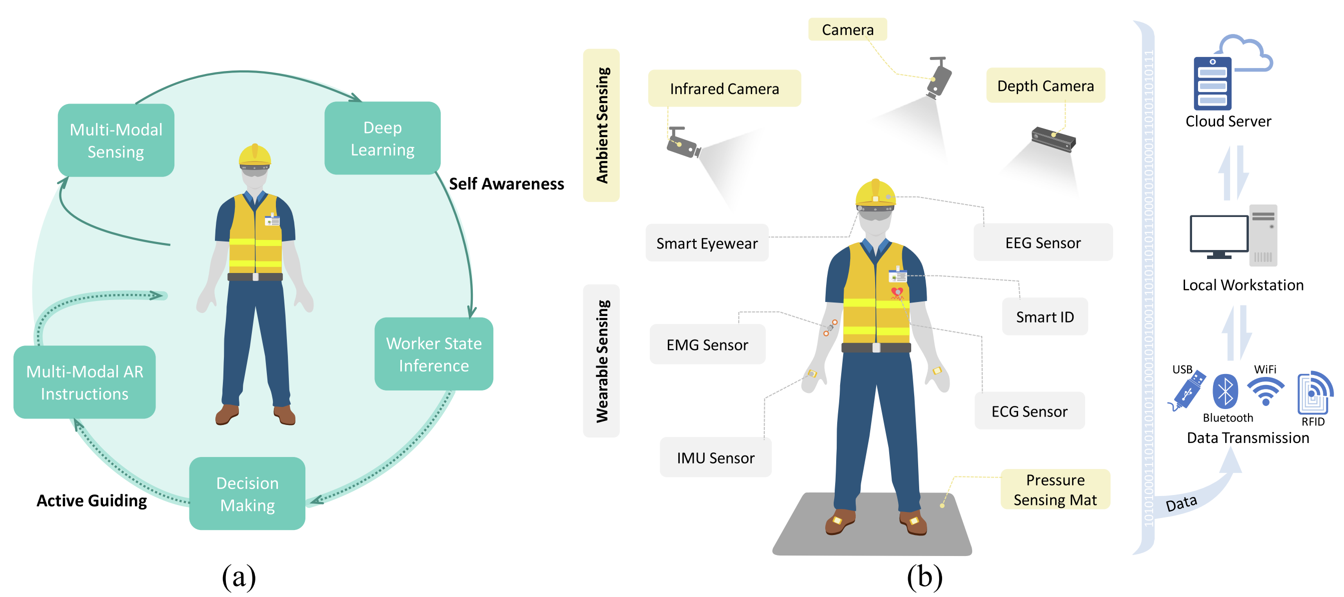

Fig. 1. (a) Overview of the proposed worker-centered training & assistant system; (b) Schematic of the proposed multi-modal sensing system.

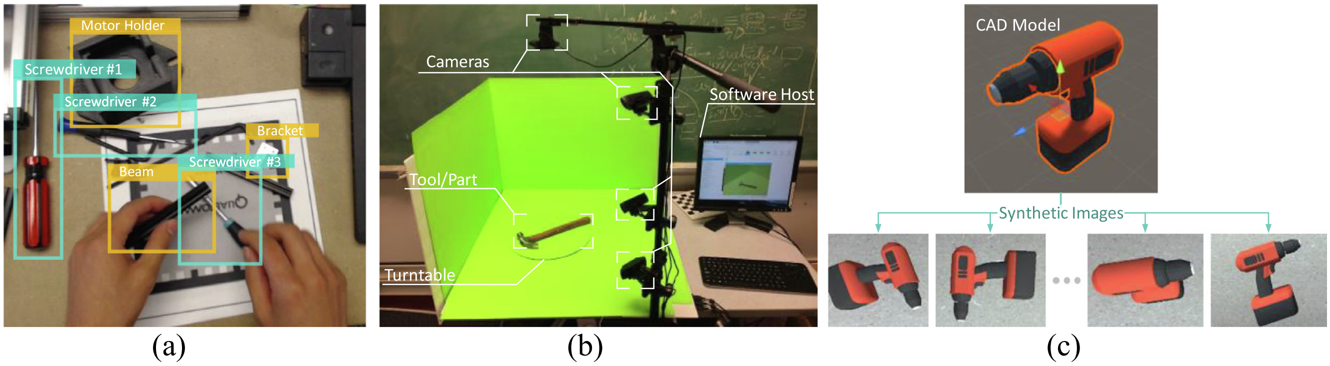

Fig. 2. (a) Part/Tool detection results with highlighted bounding boxes; (b) Our developed data collecting system; (c) Image data synthesis rendered from a CAD model.

an object was designed from CAD software or 3D scanning data and then imported to virtual scenarios. The model was rendered with different poses, obtained by setting the camera from various distances and perspectives, to simulate the variations in the physical world. With the synthesizing method, a large amount of data were generated with labels annotated efficiently.

3. Active guidance for worker

3.1. Multi-modal guidance with augmented reality

Augmented Reality (AR) technologies have been applied to manufacturing training, mainly for simulating costly or dangerous processes beforehand. In this study, we developed an instructional system with multi-modal AR to provide timely, onsite guidance for the worker.

The working scenario was captured with a first-person camera to perceive the physical world. The camera pose was estimated in real time to allow the generated virtual information to be superimposed upon the real world with intuitive mapping, which effectively eliminates the discomfort that virtual information may bring to the worker. A monitor-based or glasses-based AR interface was applied to provide graphics that are rendered for the worker and can be overlaid on the physical world. Finger-wearable haptic rings were used to give the worker a realistic feedback of the sense of touch. An auditory display was included to give the worker a timely sound feedback such as a vocal warning. An assistant laser pointer with two degrees of freedom was designed to help the worker search for the desired tools or parts. All modalities of guidance were integrated to achieve a comprehensive and complementary operational assistance. If the required parts/tools do not appear in the workspace or are not detected, the instructional information still can be provided via the visual or audio interfaces.

3.2. Demand analysis and guiding strategies

Estimating the worker's potential demands for assistance (i.e., assistant information that can instruct workers to optimize their current operational workflows, e.g., how well the current operation is performed and what the next operation is) and then providing guidance accordingly is crucial to achieving the functionality of active guiding. After the worker's states are perceived, including 1) what the worker is doing, 2) what the worker's mental intention is, and 3) what the desired tools/parts are, all the information is integrated to determine the effective assistance that can meet the worker's demand, such as instructional information to conduct the current step or a reminder warning to fix a previous illegal operation.

Furthermore, a guiding strategy was developed in order to provide instructions appropriately. A worker's performance was evaluated in comparison to experienced workers, and a "Demanding Score" is defined to represent the level of demanding for assistant information. Specifically, the time taken of each operational step can be obtained using the deep learning approaches mentioned above. If a particular action takes more time than average, the Demanding Score is increased. Then, if the Demanding Score is higher than a threshold, the needed assistant information will be actively added with the above developed multi-modal guidance system. For example, graphics information will be displayed via the monitor or AR glasses, and the laser pointer will point to the desired tool/part for the next step if the worker is in a confused state. In addition, the training progress of each worker is logged, and it can be retrieved by worker identification techniques, such as RFID tag or facial recognition. Also, the timing, i.e., how to provide the guidance at the right time, is critical in the training process. It should be timely enough but not disturb the ongoing operation.

4. Case study

The proposed self-aware and active-guiding training & assistant system has been progressively validated. In this section, two case studies in manufacturing assembly are presented.

4.1. Multi-modal recognition of worker activity

In this case study, we developed a multi-modal approach for worker activity recognition in manufacturing assembly tasks using Inertial Measurement Unit (IMU) signals obtained from a Myo armband and videos from a visual camera (see Fig. 3(a)). A worker activity dataset of six assembly activities has been established, as shown in Fig. 3(b). These activities are: grabbing tool/part, hammering nail, using power-screwdriver, resting arm, twisting screwdriver, and using wrench.

For IMU signals, we designed two modalities in both frequency and spatial domains. For the camera data, two more modalities were included at the video frame and video clip levels. Accordingly, four deep learning networks were built to cope with data from the different modalities. Then, all the individual networks were fused to output the final inference. Various fusion methods were evaluated including the maximum fusion, average fusion and weighted fusion. The developed approach has been evaluated and shown to achieve promising recognition accuracy in experiments, i.e., 97% and 100% in the leave-one-out and half-half experiments, respectively [9,19,22].

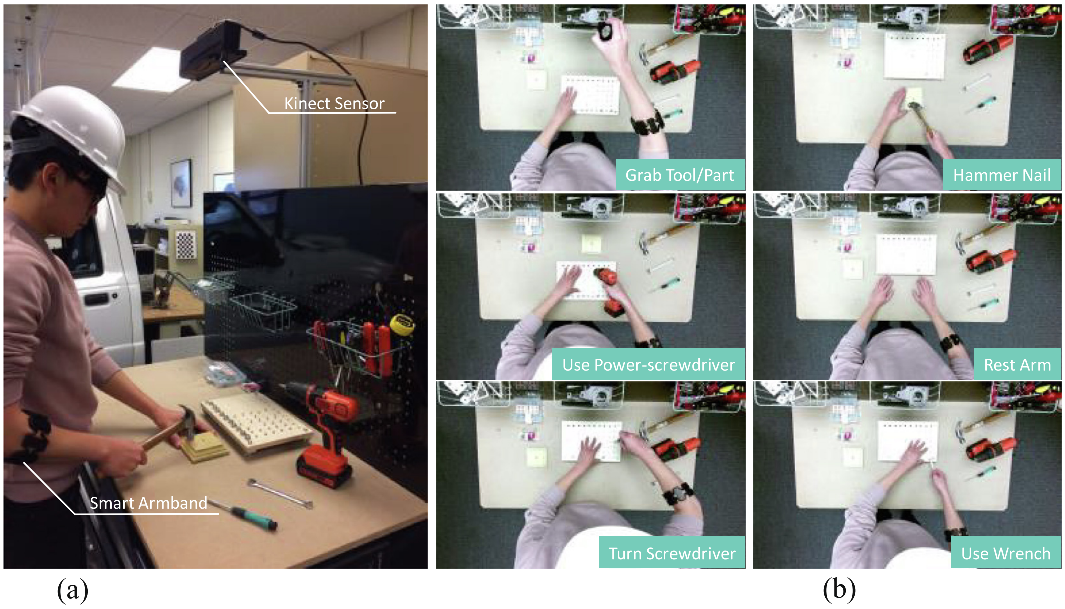

Fig. 3. (a) Experimental setup and (b) examples of the six worker activities.

4.2. Comparison of AR and manual guidance in a mechanical assembly training task

In this case study, we applied multi-modal AR guidance in a training task, i.e., assembling the spindle subassembly of a desktop carving machine (see Fig. 4(a, b)). To assess its effectiveness compared with traditional manual guidance, we recruited 20 subjects without any prior experience on the assembly task. They were divided into two groups and asked to conduct the task with manual and AR instructions, respectively. Then their performances were compared in terms of the completion time and number of errors (see Fig. 4(c, d)). The AR method has shown superiority over the manual one. This has demonstrated the feasibility and potential of applying the AR method to the industry for the frontline workers [20].

5. Conclusions

In this ongoing research, we have proposed a novel worker-centered training & assistant system for intelligent manufacturing. This system has the self-awareness of the worker's state and can provide active guidance to the worker as needed. Compared to traditional approaches, our proposed system starts with the worker's experience, considers more of the worker's learning effect, and has more interactions with the worker. The worker's state is perceived with multi-modal sensing and deep learning methods, and is used to analyze and determine the potential guiding demands. Then active instructions with augmented reality are provided to suit the worker's needs. The case studies have shown the feasibility and promise of applying the proposed system for training and assisting frontline workers. Also, our proposed self-aware and

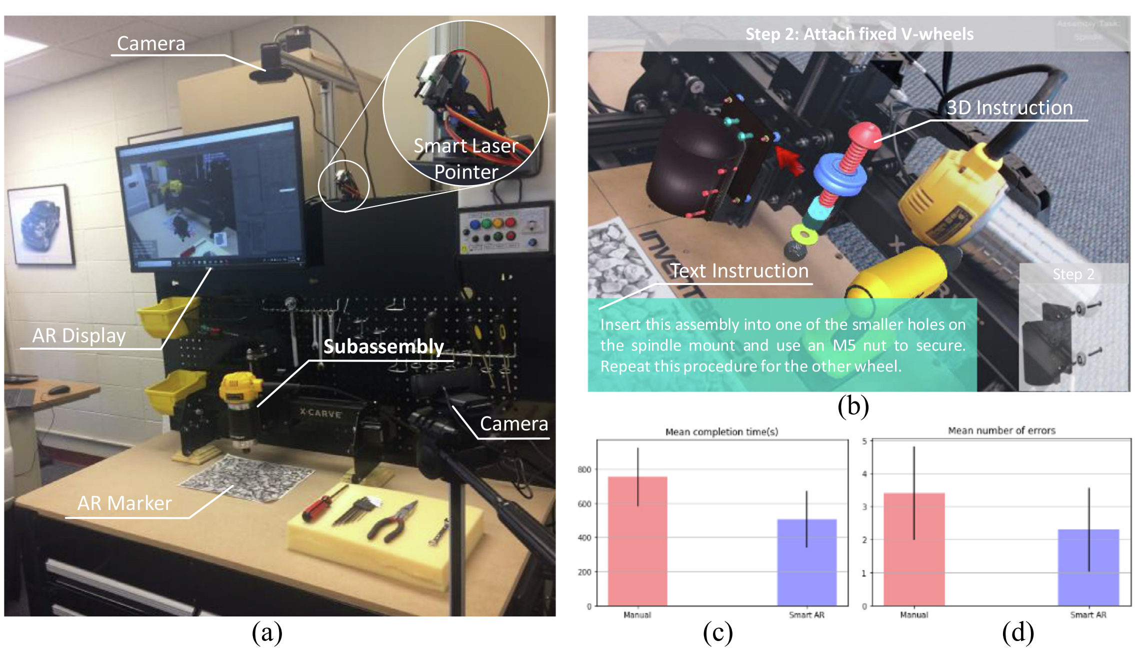

Fig. 4. (a) Experimental setup; (b) AR display content; Performance comparison between manual and AR guidance: (c) completion time and (d) number of errors.

active-guiding training & assistant system has constructed a framework for further studies in worker-centered intelligent manufacturing.

Declaration of Competing Interest

None.

Acknowledgments

This research work is supported by the National Science Foundation grant CMMI-1646162 and NRI-1830479, and also by the Intelligent Systems Center at Missouri University of Science and Technology. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

[1] Lee J, Bagheri B, Kao HA. A cyber-physical systems architecture for industry 4.0-based manufacturing systems. Manuf Lett 2015;3:18–23.
[2] Zhong RY, Xu X, Klotz E, Newman ST. Intelligent manufacturing in the context of industry 4.0: a review. Engineering 2017;3(5):616–30.
[3] Ward JA, Lukowicz P, Troster G, Starner TE. Activity recognition of assembly tasks using body-worn microphones and accelerometers. IEEE Trans Pattern Anal Mach Intell 2006;28(10):1553–67.
[4] Akhavian R, Behzadan AH. Smartphone-based construction workers' activity recognition and classification. Autom Constr 2016;71:198–209.
[5] Khosrowpour A, Fedorov I, Holynski A, Niebles JC, Golparvar-Fard M. Automated worker activity analysis in indoor environments for direct-work rate improvement from long sequences of RGB-D images. In: Construction Research Congress 2014: Construction in a Global Network. p. 729–38.
[6] Al-Amin M, Tao W, Qin R, Leu MC. Sensor data based models for workforce management in smart manufacturing. IISE Annu Conf Expo 2018;2018:1955–60.
[7] Li P, Meziane R, Otis MJD, Ezzaidi H, Cardou P. A smart safety helmet using IMU and EEG sensors for worker fatigue detection. In Robotic and Sensors Environments (ROSE). 2014 IEEE International Symposium on (pp. 55-60). IEEE; 2014.
[8] Wang J, Chen Y, Hao S, Peng X, Hu L. Deep learning for sensor-based activity recognition: A survey. Pattern Recogn Lett 2018.
[9] Tao W, Lai ZH, Leu MC, Yin Z. Worker activity recognition in smart manufacturing using IMU and sEMG signals with convolutional neural networks. Procedia Manuf 2018;26:1159–66.
[10] Yang J, Nguyen MN, San PP, Li X, Krishnaswamy S. July. Deep convolutional neural networks on multichannel time series for human activity recognition. In Ijcai 2015;15:3995–4001.
[11] Tao W, Leu MC, Yin Z. American Sign Language alphabet recognition using Convolutional Neural Networks with multiview augmentation and inference fusion. Eng Appl Artif Intell 2018;76:202–13.
[12] Liu Y, Nie L, Liu L, Rosenblum DS. From action to activity: sensor-based activity recognition. Neurocomputing 2016;181:108–15.
[13] Yew AWW, Ong SK, Nee AYC. Towards a griddable distributed manufacturing system with augmented reality interfaces. Rob Comput Integr Manuf 2016;39:43–55.
[14] Makris S, Karagiannis P, Koukas S, Matthaiakis AS. Augmented reality system for operator support in human–robot collaborative assembly. CIRP Ann-Manuf Technol 2016;65(1):61–4.
[15] Tao W, Lai ZH, Leu MC, Yin Z. Manufacturing assembly simulations in virtual and augmented reality. Augmented, Virtual, and Mixed Reality Applications in Advanced Manufacturing 2019.
[16] Doshi A, Smith RT, Thomas BH, Bouras C. Use of projector based augmented reality to improve manual spot-welding precision and accuracy for automotive manufacturing. Int J Adv Manuf Technol 2017;89(5–8):1279–93.
[17] Uva AE, Gattullo M, Manghisi VM, Spagnulo D, Cascella GL, Fiorentino M. Evaluating the effectiveness of spatial augmented reality in smart manufacturing: a solution for manual working stations. Int J Adv Manuf Technol 2018;94(1–4):509–21.
[18] Ren S, He K, Girshick R, Sun J. Faster r-cnn: Towards real-time object detection with region proposal networks. In: Advances in neural information processing systems. p. 91–9.
[19] Tao W, Lai ZH, Leu MC, Yin Z. Multi-modal recognition of worker activity for human-centered intelligent manufacturing. Submitted to J Manuf Syst 2019.
[20] Lai ZH, Tao W, Leu MC, Yin Z. Smart augmented reality instructional system for mechanical assembly towards worker-centered intelligent manufacturing. Submitted to J Manuf Syst 2019.
[21] Gilbreth FB. 1911. Motion study: A method for increasing the efficiency of the workman. D. Van Nostrand Company.
[22] Al-Amin M, Tao W, Doell D, Lingard R, Yin Z, Leu MC, Qin R. Action recognition in manufacturing assembly using multimodal sensor fusion. 25th International Conference on Production Research Manufacturing Innovation: Cyber Physical Manufacturing, August 9–14, 2019, 2019.
[23] Tao W, Lai ZH, Leu MC, Yin Z. American Sign Language Alphabet Recognition Using Leap Motion Controller. Proceedings of the 2018 Institute of Industrial and Systems Engineers Annual Conference (IISE 2018), 2018.