# Posture-related data collection methods for construction workers: A review

**Abstract**: Construction workers' posture-related data is closely connected with their safety, health, and productivity performance and has drawn the attention of researchers in construction management and other fields. Accordingly, many data collection methods have been developed and applied to collect posture-related data. Despite the importance, there lacks a review of previous data collection methods in the construction industry. This paper fills the research gap by reviewing previous methods to collect posture-related data for construction workers via 1) summarizing working principles and applications of posture-related data collection in construction management, which demonstrates the extensive use of motion sensors and Red-Green-Blue (RGB) cameras in posture-related data collection, 2) comparing the above methods based on data quality and feasibility on construction sites, which reveals the reason why motion sensors and RGB cameras have been prevalent in previous studies, 3) revealing research gaps of posture-related data collection tools and applications, and providing possible future research directions.

**Keywords**: behavior-based safety (BBS); computer vision; construction worker; deep learning; motion sensor; occupational safety and health (OSH); pose estimation.

## 1 Introduction

The construction industry plays a significant role in the economy of developed and developing countries [1–4]. Despite its significance, the global construction industry has shown poor safety, health, and productivity performance. In terms of safety, the construction industry is one of the most dangerous fields to work in, wherein one in five workers deaths have been caused by construction [5]. In terms of health, the construction industry is a high-risk industry for work-related musculoskeletal disorders (WMSDs) due to the highly physically demanding tasks which expose construction workers to a number of risk factors such as overexertion, awkward posture, and repetitive motions [6,7]. With regard to productivity, the construction industry has been falling behind, since the annual increase in global labor productivity rate for construction was only one-third of that in manufacturing over the past two decades [8]. In conclusion, the construction industry has unsatisfactory performance and there is an urgent need to improve the safety, health, and productivity of the construction industry.

Construction safety, health, and productivity performance are closely related to working posture-related data. In construction safety management, unsafe behaviors are the major cause (over 80%) of accidents [9], and construction workers' postures have been used as the predictor of unsafe behaviors to prevent potential unsafe risks from developing into accidents [10,11]. With regards to health management, awkward posture is one of the eight risk factors related to WMSDs [12], and is included in several ergonomic assessment scales [13–15]. For productivity, postures could reflect working status, such as "effective work" or "contributory work", and contribute to individual productivity assessment [16]. Conclusively, workers' posture-related data is important for improving construction safety, health, and productivity performance.

Despite its importance, collecting workers' posture-related data on construction sites remains a challenging task due to the following reasons. First, unlike industrial production-lines, construction activities are relatively random in nature [15]. Second, construction workers' working area is not fixed, which makes it challenging to collect data continuously. Third, harsh working environments of construction sites bring more challenges to the maintenance of data collection devices.

Knowing the value as well as the challenges, over the years, researchers have endeavored to improve data collection tools and safety, health and productivity management methods based on posture-related data. This study intends to provide a comprehensive review of existing methods for the collection and application of posture-related data in construction from the perspectives of data reliability and feasibility. Specifically, this review aims to fulfill the following objectives: 1) categorizing previous studies relevant to workers' posture-related data and appraising their advantages and disadvantages, 2) comparing data reliability and their feasibility, and 3) revealing research gaps and suggesting possible future research directions.

This study encompasses multiple types of posture-related terminologies, which are generally referred to as *posture-related data*. The terminologies are explained as follows. *Pose* and *posture* are static concepts, which serve to denote the configuration of the human body at specific moments. In the following discussion, by "pose" we refer to the positions of the key joints and by "posture" we refer to the semantic description of a human body configuration. *Pose* includes both 2D pose and 3D pose, which represent the 2D and 3D coordinates of major human body joints, respectively. *Posture* represents either whole-body postures (e.g., squatting and sitting) or the posture of body segments (e.g., straight/bent wrist). Moreover, human body shapes, body orientation, and postural stability, are also included in this review. As for dynamic cases, "action" and "activity" are used to describe the dynamics of a human body in a certain period of time. Action and activity are frequently used interchangeably. In the following discussion, the study follows the definitions proposed by Turaga et al., i.e. "action" refers to the simple motions patterns of a single person in a short duration of time, such as tens of seconds; "activity" is the "complex sequence of actions performed by several humans who could be interacting with each other in a constrained manner" in longer durations [17]. Moreover, this review also includes studies related to joint kinematic data, such as joint 3D locations and joint 3D accelerations.

The rest of the paper is arranged as follows. First, section 2 introduces the literature search strategies and respective results. In section 3, the collected studies are classified into three categories according to posture-related data collection methods, namely, manual methods, contact-sensor-based methods, and non-contact-sensor-based methods. Each type of method is introduced from the perspectives of working principles, advantages, and disadvantages. In section 4, the applications of posture-related data collection methods in safety, health, and productivity management are elaborated. Based on their working principles and application scenarios, section 5 compares the posture-related data collection methods according to data reliability and feasibility. Section 6 discusses the research gaps in accuracy, intrusiveness, and the application of posture-related data collection methods. Finally, potential future research directions are suggested based on the research gaps.

## 2 Research method

A two-stage review method was adopted to identify refereed journal and conference articles. In the first stage, a search query was used on Scopus to identify potential candidates. The search query was designed to include “construction industry,” “construction worker,” or “construction workers,” and “pose,” “poses,” “posture,” or “postures” in the title, abstract, and keywords, as shown in Table 1. The search was limited to research articles, conference articles, and literature reviews written in English. 139 papers were identified in the first stage. In the second stage, a manual screening of the articles was conducted to filter the studies not directly related to construction workers’ posture-related data. Finally, 57 articles were included in the review.

*Table 1 Literature searching query*

| Title/Abstract/Keywords | Document type | Language |
|---|---|---|
| (“construction industry” OR “construction worker*”) AND (“pose*” OR “motion*”) | Research articles/ Conference articles/ Literature reviews | English |

Figure 1 and Figure 2 present the annual publication numbers of the reviewed articles categorized by application scenarios and data collection methods, respectively. The total number of annual publications increased between 1999 and 2019, especially after 2010. According to Figure 1, most of the reviewed articles, 35 out of 57, focused on improving construction workers’ health issues with posture-related data. Thirteen articles used posture-related data in identifying unsafe behaviors. Four articles assessed labor productivity with posture-related data. The remaining articles were entirely focused on developing posture-related data collection methods or assessing their accuracy. In consideration of the data collection methods, this study divided them into three categories, i.e. manual methods including self-report and manual observation, contact-sensor-based methods including wearable motion sensors and marker-based motion capture systems, and non-contact-sensor-based methods mainly based on computer vision. Manual methods and contact-sensor-based methods have been widespread in the past two decades, and the use of the contact-sensor-based methods reached its peak in 2017. Non-contact-sensor-based methods appeared after 2013 when computer vision-based methods were mature enough to support data collection from images and video clips.

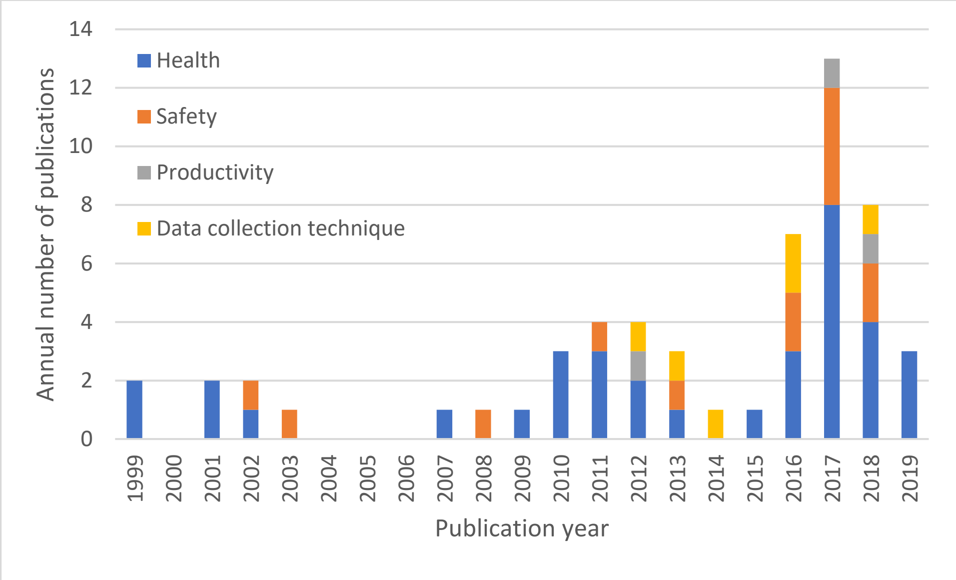


*Figure 1 Annual number of publications categorized by application scenarios*

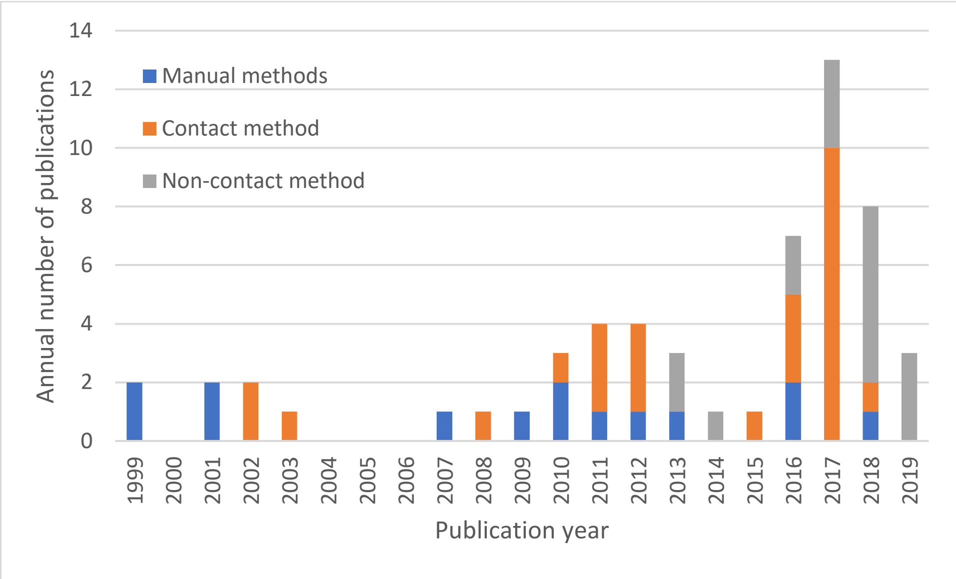


*Figure 2 Annual number of applications categorized by data collection methods*

Figure 3 depicts the journals and disciplines of the reviewed research. The disciplines reported are based on the journal categories from the Clarivate Journal Citation Report. The diversity of the categories demonstrates the multidisciplinary nature of the reviewed topic. “Construction & Building Technology” and “Industrial Engineering” accounted for most of the reviewed articles. Within these two categories, Automation in Construction, Applied Ergonomics, and Journal of Construction Engineering and Management published more articles than the other journals. Furthermore, the category, "Occupational Health," showed the popularity in applying posture-related data when solving construction workers' health problems.

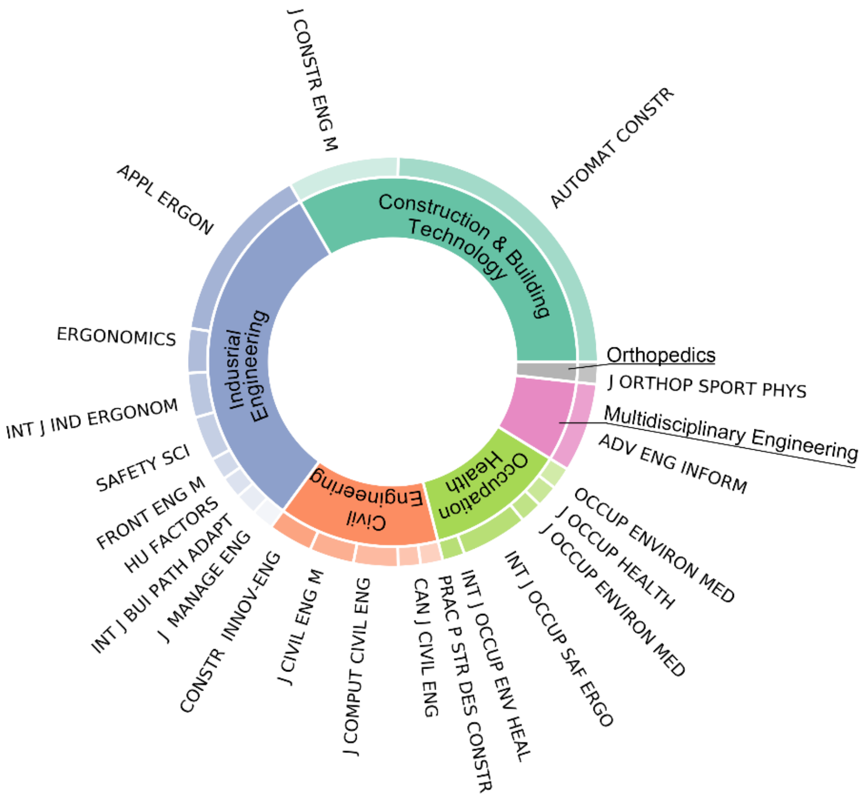


*Figure 3 Journals and disciplinary of the reviewed papers*

## 3 Posture-related data collection methods in the construction industry

From the perspectives of working principles and data formats, this section summarizes previous posture-related data collection methods for construction workers and discusses their advantages and disadvantages. Posture-related data collection methods can be divided into two categories: manual methods and automatic methods. Automatic methods usually involve sensors, which include contact sensors and non-contact sensors [18]. Accordingly, this section classifies posture-related data collection methods into manual methods, contact-sensor-based methods, and non-contact-sensor-based methods.

### 3.1 Manual methods

Manual methods acquire posture-related data manually. Self-report and observation were the two popular manual methods used in different studies. It should be noted that though the outputs of these self-report and observation data were primarily ergonomic assessments, the data collection process incorporated posture-related data, such as joint angles and body orientations. Therefore, this review further details the self-report and observation methods in the following section.

#### 3.1.1 Self-report

Self-report methods involved asking the workers to recall their postures during construction tasks and answer questions about their postures. Questionnaires or interviews were used to collect self-repot data, which are logistically easy to conduct and have a low initial cost. These studies looked into posture-related data from multiple perspectives including identifying awkward postures, manual material handling and balance stability [19–21].

Despite their wide application, there exist several shortcomings of self-report methods. First, self-report methods result in subjective data, thus calling into question its reliability. Second, self-report methods cannot collect data continuously, which makes the ergonomic analysis of prolonged activities quite difficult. Finally, with an increasing number of participants, self-report methods become labor and time consuming, making it impossible to collect "big data" from construction sites.

#### 3.1.2 Observation

Systematic observation is "an objective and well-ordered method for close examination of some aspects of behavior to obtain reliable data unbiased by observer interpretation" [22]. Working postures are recorded and analyzed using a variety of methods including pen and paper-based observation methods, videotaping, and computer-aided analysis [23]. For example, Louhevaara assessed postural workload based on field observation, while Lee and Han recorded the working postures with video first and later collected postures through frame sampling [24,25].

Systematic observation methods typically specify which and how variables should be recorded in order to make the results more objective and comparable. At least nine different techniques have been developed in ergonomics to collect posture-related data for assessing physical strain at work [23]. For instance, Pose, Activity, Tools, and Handling (PATH), an observation-based sampling method specially designed for construction tasks, uses seven-digit codes to record construction postures. Four digits describe the postures of a worker's back, arm, leg, and hand load, while the remaining three digits describe the construction activity, tools being used and the grip of the tool [15]. PATH has been successfully applied in the construction industry to record the working postures of laborers, carpenters, ironworkers, plasterers, and tilers, thus facilitating the ergonomic risk assessment [26].

In summary, systematic observation includes specific rules of data coding and recording, which increases the accuracy and the level of detail of the collected data. However, since the categorization of postures relies on the observers' experience, errors resulted from subjective judgment cannot be avoided [27].

## 3.2 Contact-sensor-based methods

Contact-sensor-based methods use body attached sensors or markers to collect construction workers' posture-related data. The sensors could measure or calculate joint kinematic data, such as joint acceleration, joint position, and joint angle.

### 3.2.1 Inertial Measurement Unit (IMU) and Electro-goniometers

IMU is a sensor system using measurement systems, e.g., gyroscopic sensors and accelerometers, to estimate relative position, velocity and acceleration [28]. When attached to workers' joints, IMU could collect joint kinematic data for behavioral analysis. For example, an IMU attached to the worker's waistline could measure three-axis accelerations for postural stability assessment and safety hazards identification [29,30]. In addition to IMU, electro-goniometers could also measure joint kinematics. For example, a Lumbar Motion Monitor System, which includes a portable tri-axial electro-goniometer attached to the worker's back, could continuously document lumbar region postures [31].

Multiple IMUs attached to key human body joints constitute an IMU system, and could collect construction workers' full-body posture-related data for behavioral analysis. For example, an IMU system, which employed eight IMUs covering the upper/lower back, arms, and upper/lower legs, could estimate joint angles and identify postures such as standing up, stooping, and squatting [32]. Another system used an IMU-based suit with 17 IMU sensors to collect 3D poses consisting of 28 joint center locations [33]. Due to the high dimensionality of 3D pose data in such cases, dimension reduction techniques, such as the Bag of Features (BoF) and motion tensor decomposition, were used to compress the full-body 3D pose data [33,34]. Then classification algorithms, such as Supported Vector Machine (SVM), were trained to categorize the compressed 3D pose data [33,34].

#### 3.2.2 Marker-based motion capture system

Marker-based motion capture systems are used for 3D motion capture and analysis in laboratories. A marker-based motion capture system usually consists of cameras and reflective markers. The cameras are equipped with infrared sensors. The reflective markers are placed on designated locations of a human body which are then tracked by the camera. The system estimates the movement trajectory of each marker. A marker-based motion capture system could complete several tasks, such as analyzing the gait of construction workers, as well as assessing postural stability [35,36]. Marker-based motion capture systems are usually highly accurate. For instance, a Vicon-460 system can provide an overall accuracy of 63±5 μm for the most favorable parameter setting [37].

Compared with manual methods, contact-sensor-based methods possess the following advantages. First of all, the sensors could measure and record joint trajectories automatically without manual intervention. Second, the collected data is objective and accurate. Furthermore, the data could be collected continuously at a high frequency. However, if applied on construction sites, the contact sensors would need to be tied tightly to the construction workers' trunks and limbs, which may lead to discomfort and annoyance. Besides, additional costs might arise from recharging and maintaining the sensors.

### 3.3 Non-contact-sensor-based methods

Non-contact-sensor-based methods could collect data in a non-invasive way. They usually use images or videos of construction sites, which contain visual information related to working postures. Reviewed articles indicate depth cameras and Red-Green-Blue (RGB) cameras to be the two most popular tools for non-contact-sensor-based methods.

### 3.3.1 Depth camera

Depth cameras generate range images or 3D point clouds. In a range image, each pixel corresponds to a numerical value representing the distance from the camera, i.e. the depth of the pixel. 3D point clouds consist of the 3D coordinates of the points on the external surfaces of the scanned objects. The following is a brief introduction to three types of depth cameras, including structured light depth cameras, stereo depth cameras, and time-of-light (ToF) depth cameras.

Structured light depth cameras project light patterns on to a scene and extract depth information by analyzing the distortion of the observed patterns [38]. Kinect V1, is a typical structured light depth camera, which relies on infrared light patterns to estimate depth. In the construction industry, it was used to estimate the joint location trajectories of construction workers [9].

Stereo depth cameras perceive depth by simulating the human binocular vision system. A stereo depth camera captures images with at least two image sensors and calculates depth by estimating disparities between matching key points in the images. A previous study developed a stereo depth camera consisting of two common smartphones, and applied it to estimate 3D poses [39]. Another study applied commercially available stereo cameras to detect the 2D locations of human body joints and compute the 3D positions of the joints using triangulation [40].

ToF depth cameras or sensors determine depth by computing the time taken by light to get back to the camera. The time taken is then multiplied by the speed of light to obtain the depth. Kinect V2 contains a ToF camera, where the light is infrared. Kinect V2 was used in construction to collect 3D poses for unsafe behavior detection [10,11]. LiDAR sensors and radar sensors are also ToF depth cameras, which use laser lights or radio waves to calculate depth. Previous studies have used them to recognize human postures or estimate human poses according to 3D point clouds [41,42].

### 3.3.2 RGB camera

Widespread surveillance cameras on the construction sites provide vast amounts of information for pose estimation. However, unlike depth images containing the depth information of each pixel, the images captured by RGB cameras contain only 2D information, making it a challenge to collect posture-related data from RGB images. Based on the 2D information of an image, researchers utilized hand-crafted features or learned features to detect workers, recognize postures, and estimate 2D or 3D poses. Section 3.3.3 provides a summary of relevant algorithms.

### 3.3.3 Computer vision algorithms for posture-related data collection from depth cameras and RGB cameras

The reviewed algorithms that extract posture-related data from construction images or videos are classified into four categories based on the outputs, including worker detection, posture classification, pose estimation, and action recognition.

*Worker detection* algorithms aim to find workers in an RGB image, which could answer the question "are there any construction workers in the image?". The reviewed studies selected regions of interest first by using sliding detection windows or detecting moving objects from a series of RGB images, then extracted hand-crafted features from the selected regions, such as histograms of oriented gradients (HoG) and color features [43,44]. Machine learning algorithms, such as SVM and K-Nearest Neighbor (KNN), were applied to train classifiers based on the extracted features to differentiate construction workers from other objects [43,44].

*Worker posture classification* algorithms take a step further and classify the postures of the detected workers from images. In some previous studies, the first step of posture recognition was worker detection [16,45]. Bai el al. applied a similar strategy to [43] to detect workers, i.e., using motion features to detect moving objects and using color features to identify workers from the detected moving objects [16]. After worker detection, a silhouette was created for each worker, which was then thinned to generate a skeleton. An Artificial Neural Network (ANN) was designed to classify the skeletons into effective, contributory, and ineffective categories. In terms of depth images, depth information was employed to detect workers. Ray and Teizer computed the median image from a set of depth images to subtract the background and search the largest bounding boxes for clusters of connected pixels for the detection of workers [45]. The depth values of the pixels surrounded by the bounding box were then rescaled and reshaped into a vector. Finally, linear discriminant analysis (LDA) was applied to classify postures into standing, squatting, sitting, stooping, bending and crawling [45].

*Worker pose estimation* algorithms include 2D pose estimation and 3D pose estimation. *2D pose estimation* is a classic task in computer vision, which aims at "obtaining 2D pixel positions of human body joints from an image" [46]. The output of 2D pose estimation is 2D skeletons consisting of the 2D coordinates of human body joints. *3D pose estimation* is "the task of producing a three-dimensional figure that matches the spatial position of the depicted person" given an image of a human being [47]. The results of 3D pose estimation are 3D skeletons consisting of 3D coordinates of human body joints.

In terms of 2D pose estimation, a two-branch Convolutional Neural Network (CNN) was applied to estimate the 2D skeletons of construction workers from RGB site images, where the first branch detected body parts and the second branch predicted the corresponding body part association [48]. However, 2D poses are view variant and view-invariant features are required for ergonomic posture classification according to 2D poses [48]. 3D joint locations of human bodies are view invariant. To collect 3D poses, Liu et al. used two cameras to record working scenarios at the same time from different point of views [39]. Then Hue-Saturation-Value (HSV) color features and optical flow were applied to track 2D body joints from each of the image sequences captured by the two cameras. Then, scale-invariant feature transform (SIFT) and speeded up robust features (SURF) were used to match the 2D skeletons in the two image sequences, and finally, paired body joints were triangulated to compute the 3D joint positions. Other methods estimated 3D poses from monocular RGB images with CNN, such as [49] and [50]. Zhang et al. applied a multi-stage CNN to estimate the 3D poses of workers from construction site video frames [49]. In each stage, a 2D joint predictor generated belief maps of human body joints, then a probabilistic 3D pose model estimated a 3D pose based on the 2D belief maps. After that, the estimated 3D pose was projected back onto the image plane to generate a new set of 2D belief maps. Next, a fusion layer fused the two sets of 2D belief maps, which were passed to the next stage for 2D joint location prediction. After six stages, the probabilistic 3D pose model generated a 3D pose according to the final set of 2D joint belief maps. Yu et al. used another CNN architecture to estimate the 3D poses of construction workers from site images [50]. The network was twofold. The first part was a CNN, named Stacked Hourglass network, to estimate 2D poses from the images; the second part was a separate neural network which inferred 3D poses based on 2D poses and bone length constraints.

*Work action recognition* algorithms focus on identifying dynamic actions, given a video clip or an image sequence. Previous studies employed various feature extraction methods and learned classifiers on the features [51–53]. In terms of depth video, 3D poses were typically used in feature extraction [51,52]. For example, Han et al. estimated 3D poses in each depth video frame first and then obtained 3D pose sequences [9,40]. Afterwards, kernel principal component analysis (kernel PCA) was used to reduce the dimension of a 3D pose sequence by mapping the sequence onto a 3D space and generating a 3D trajectory. For action recognition, the studies classified the actions by measuring the similarity between the trajectories of a new action and those of the prior actions. Khosrowpour et al. also generated features based on 3D poses from depth video frames [51]. Joint angles were used as the pose features of each frame. Then a BoF approach was used to generate the features for the video clips. For action recognition, SVM was used to classify the BoF representations of the actions. BoF has also been used to extract features from RGB video clips for action recognition. Yang et al. first generated dense trajectories from RGB video clips by densely sampling and tracking the features with dense optical field, then extracted features by computing HoG, HoF (Histogram of Optical Flow), and MBH (Motion Boundary Histograms) within a space-time volume around the trajectory [53]. Then codebooks and BoF representation were calculated. Finally, SVM was also applied to classify actions based on the BoF representations.

Above action recognition algorithms were further tested on video clips containing a single action of one construction worker. However, typical construction site surveillance includes continuous construction actions of random order and duration. To solve this problem, Khosrowpour et al. modelled and inferred duration-variable work action series with a Hidden Markov Model (HMM) [51]. For multi-worker action recognition from site surveillance videos, a two-stream CNN architecture was designed in [54]. The first step was to track workers and create temporally and spatially cropped videos. Then, spatial streams and temporal streams of individual workers were extracted, where spatial streams were RGB images and temporal streams were warped optical flow fields. Next, two-stream CNN recognized construction actions from the spatial stream and the temporal stream, respectively. Finally, the work recognition results from the two streams were fused.

Table 2 summarizes the above algorithms. Before 2016, most of the algorithms used hand-crafted features to represent images or videos and then applied machine learning algorithms to classify the images or videos according to the features. Hand-crafted features are suitable for small-sized homogeneous datasets [55]. For example, some reviewed studies used color features to detect workers from RGB images, which assumed that all construction workers wore safety vests [43,44]. However, the detection method might fail if the worker does not wear a safety vest. Compared with hand-crafted features, learnt features in deep learning performs better on heterogeneous and large datasets [55]. After 2016, deep learning networks, such as CNN, came into use for pose estimation and action recognition, which learn the features through a training process and could directly use the raw data. For example, CNN could successfully identify the construction workers despite them not wearing safety vests [54].

In terms of dataset, Table 2 compares above studies based on data collection environments, participants, data categories, and the total number of training examples. For real construction site applications, a dataset should consider the variety of working environments (indoors and outdoors), trades, and working postures or actions. However, nearly half of the previous studies were conducted in indoor environments, limiting the application in outdoor construction fields.

In addition, the generalization ability of the algorithms used in the studies is questionable. Poor generalization ability could be caused by the difference between the distributions of the training data and test data [56]. According to Table 2, only four studies included more than ten workers and three studies included more than three construction trades, meaning that the training datasets are less diversified. When used on real construction sites, the algorithms might be faced with unforeseen data and generate biased predictions.

*Table 2 Comparison of computer vision algorithms for construction worker pose estimation*

| Task | Input | Algorithm | Dataset | | | | Performance evaluation | Ref | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Environments | Participant | Labeling | No. of data samples | | | |
| Worker detection | RGB image | 1. Motion features for foreground blobs recognition<br>2.HOG and SVM for human body identification<br>3.HOC (RGB/HSV) and KNN for construction worker detection | Field | 5 workers | Positive and negative examples of standing workers | 2700 images | 99.00% (Precision) | [43] | (Park and Brilakis, 2012) |
| | RGB image | 1.Multi-scale sliding detection windows<br>2.HoG and HoC (HSV) as features<br>3.SVM for worker/non-worker classification | Field | Multiple workers on 5 construction sites | | 8000 images | 98.83% (Accuracy) | [44] | (Memarzadeh et al., 2013) |
| Posture classification | RGB image | 1.Motion segmentation algorithm for identifying moving objects<br>2.Color variance for extracting workers<br>3.Region growing technique for creating silhouette<br>4.Fast parallel algorithm for creating 2D skeletons<br>5. ANN for posture classification | Field | 2 rebar workers | 3 types of postures | 2000 images | 99.00% and 81.00% for each worker (Accuracy) | [16] | (Bai et al., 2012) |
| | Depth image (Infrared) | 1.Extracting person using pixel depth<br>2.LDA for posture classification | Laboratory | 1 student | 4 postures | 22226 images | 0.33% (highest error rate) | [45] | (Ray and Teizer, 2012) |
| 2D Pose estimation | RGB image | Two-branch CNN<br>- 1st branch for body part detection<br>- 2nd branch for body part association | Field | - | - | - | - | [48] | (Yan et al., 2017) |
| 3D Pose estimation | Depth image (Stereo) | 1. HSV features and optical flow for 2D skeleton estimation<br>2. SIFT and SURF for 2D skeleton matching and triangulation for 3D skeleton reconstruction | Laboratory | 1 student | 2 actions | 1000 images | 3.80 cm (Average error of bone length measurements) | [39] | (Liu et al., 2016) |
| | RGB image | Multi-stage CNN architecture combining Convolutional Pose Machine and a probabilistic 3D joint estimation mode | Field | 11 people | 3D joint locations of daily actions | Human 3.6M dataset | 8.93 cm (Mean per joint position error) | [49] | (Zhang et al., 2018) |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | RGB image | CNN for 2D pose estimation and 3D pose reconstruction | Field | 11 people | 3D joint locations of daily actions | Human 3.6M dataset | 3.90 cm (Mean per joint position error) | [50] | (Yu et al., 2019) |
| | Depth video (Infrared) | 1. 3D pose estimation<br>2. Kernel PCA for dimension reduction of 3D pose sequences<br>3. Temporal spatial similarity check for action detection | Laboratory | 1 student | 50 trials of ladder climbing (25 for ascending and 25 for descending) performed by 1 participant | 50 video clips | 98% (accuracy), 100% (precision), 96% (recall) | [9] | (Han et al., 2013) |
| | Depth video (Stereo) | 1.2D pose estimation, 2D pose matching and triangulation for 3D skeleton reconstruction<br>2. Kernel PCA for dimension reduction of 3D pose sequences<br>3. Temporal spatial similarity check for action detection | Laboratory | 1 student | 25 trials of ladder climbing, each of which is comprised of an ascending, a reaching-far-to-side, and a descending action | 25 video clips | 88% (precision), 88% (recall) | [40] | (Han et al., 2013) |
| Action recognition | Depth video (Infrared) | 1.Bag-of-poses for representing short pose sequence and SVM for classification<br>2.HMM for long pose sequence classification | Laboratory | NA | 7 actions of 1 trade | 11 video clips | 76.00% (accuracy) | [57] | (Khosrowpour et al., 2014) |
| | RGB video | 1.Densely sampling and tracking feature points on multiple spatial scales for dense trajectory generation<br>2.Motion Boundary Histogram as feature descriptor<br>3.SVM for classification | Field | 4-25 workers for each action | 11 types of actions of 5 trades | 1176 video clips | 59.00% (accuracy) | [53] | (Yang et al., 2016) |
| | RGB video | 1.CNN for extracting and learning frame-level features<br>2.LSTM for fusing and learning spatiotemporal features from image sequences | Laboratory | 1 student | 4 actions in ladder climbing | 200 video clips | 92% (accuracy) | [58] | (Ding et al., 2018) |
| | RGB video | Two-stream CNN:<br>1.MDNet for single worker tracking<br>2.FlowNet2.0 to extract spatial and temporal features<br>3.Temporal Segment Networks for recognizing activities from spatial and temporal streams<br>4.Fusing results of two streams | Field | Multiple workers on a construction site for 41 days | 16 tasks of 4 trades | 1055 video clips | 80.50% (accuracy) | [54] | (Luo et al., 2018) |

## 3.4 Summary

Table 3 summarizes the advantages, disadvantages, and outputs of the reviewed posture-related data collection methods. Manual methods, including self-report and systematic observation, describe postures with natural languages or standardized coding. Manual methods are easy to implement, but the accuracy of the results largely depends on subjective judgment. Besides, as data collection and data recording rely on manual work, these methods are not feasible for continuous data collection. Last but not least, these methods are labor and time consuming.

Contact-sensor-based methods provide precise posture-related data, including postures, joint kinematic data, and 3D poses. Furthermore, the data can be collected continuously and accurately. However, the sensors and markers need to be attached to workers' bodies, which is intrusive and might instigate irritation, consequently making them unwilling to wear these sensors and markers.

Compared with contact sensors, non-contact sensors are less invasive, as workers would not need to wear any sensors. In terms of data accuracy, non-contact-sensor-based methods cannot provide highly accurate joint kinematic data, such as joint acceleration. Although acceleration is the second derivative of displacement, current non-contact-sensor-based methods are not able to measure joint coordinates that are accurate enough for derivation. Considering the applicability on construction sites, non-contact sensors might have limited use on construction site environments. For example, infrared depth cameras are extremely sensitive to the environment. Sunlight and far distances between workers and cameras severely affect the depth estimation [59,60]. In contrast, stereo cameras and monocular RGB cameras are more suitable for outdoor construction sites, but occlusions and low light conditions might affect their accuracy. In reality, construction workers might be occluded, and some workspaces may have poor lighting available. Importantly, the performance of the reviewed algorithms under occlusion or in low light conditions is unknown.

*Table 3 Advantages, disadvantages and outputs of posture-related data collection methods for construction workers*

| **Method** | | **Advantage** | **Disadvantage** | **Outputs** |
|---|---|---|---|---|
| Manual methods | Self-report<br>Systematic observation | Easy to implement | Subjective judgment<br>Not continuous<br>Labor and time consuming | Postures |
| Contact-sensor-based methods | Wearable sensors | Automatic<br>Continuous<br>Accurate | Intrusive<br>Uncomfortable | Joint kinematics<br>Postures<br>3D poses |
| | Marker-based motion capture system | | | Joint kinematics<br>3D poses |
| Non-contact-sensor-based methods | Depth camera | Less intrusive<br>Automatic<br>Continuous | Infrared-based depth cameras are sensitivity to sunlight<br>Less accurate than contact sensors | 3D poses |
| | RGB camera | Less intrusive<br>Automatic<br>Continuous<br>Less sensitivity to sunlight | Less accurate than depth camera | Posture features<br>2D poses<br>3D poses |

# 4 The application of the posture-related data collection methods in the construction industry

Construction workers' posture-related data is closely related to their safety, health, and productivity. This section reviewed the studies for using posture-related data to enhance/monitor safety, health and productivity.

## 4.1 Safety hazards identification

Posture-related data has been applied to improve safety performance through detecting safety violations, evaluating fall risks and identifying work site hazards.

Safety violation detection was modelled as a posture classification problem or an action recognition problem in the reviewed studies. To convert safety violation detection into a posture classification problem, a representative frame was selected from a video clip of an action, followed by labelling them either as safe or unsafe behavior [10,11]. The advantage of this strategy is so that the representative frame could be detected before the safety violation occurs. One of the limitations is that this strategy fails to consider the temporal information. For example, it might be difficult to differentiate "climbing up" and "climbing down" according to a climbing posture. Action recognition algorithms consider both spatial and temporal features.

Using such algorithms, Han et al. and Ding et al. detected three types of unsafety behaviors in ladder climbing with action recognition algorithms, demonstrating the potential of the algorithms in detecting safety violations from construction site videos [9,58]. However, the reviewed action recognition algorithms were tested on small-sized datasets consisting of only four types of ladder climbing actions of two workers. Thus, the performance of the algorithms in detecting other safety violations remains to be tested. Moreover, these safety violation detection methods require predefined rules or templates of safety violation behaviors, limiting their practical usage, since it is challenging to define the rules and templates for all unsafe behaviors. Besides, with an increase in the number of target safety violations, more features might be used for classification, requiring more computing resources and increasing the algorithm latency.

Besides detecting unsafe behavior, fall risk is a leading cause of injuries on construction sites. The identification of fall risks has become a popular issue in construction site management. Postural stability is often a contributing factor in injuries due to falls. Previous studies have evaluated postural stability through questionnaires and IMUs [21,29]. Written questionnaires were used to collect the construction workers' perceptions of postural stability [21]. To facilitate the automatic and objective estimation of postural stability, IMUs were used by Jebelli et al. to estimate two clinical postural stability measurements: the average velocity of center of pressure and resultant acceleration [29].

IMU acceleration data is also related to work site hazards. Kim et al. used the acceleration data to analyze the patterns of bodily responses and found that abnormal bodily responses were associated with locations of safety hazards [30]. The disadvantage of the approach was that the method could only perform post-accident analysis and cannot provide early warnings to prevent accidents.

## 4.2 Work-related musculoskeletal disorder risk assessment

Working postures, durations, and work-rest schedules are closely related to WMSDs, which are very common in construction workers and could lead to extremely adverse effects on construction workers' health [6,61]. Posture-related data could help to assess the workloads of different working tasks and mitigate the risk of fatigue and injuries [62,63]. Additionally, posture-related data can be fed into ergonomic or biomechanical models/algorithms for WMSDs risk assessment.

#### 4.2.1 Posture-related data in ergonomic assessment tools

Ergonomic assessment tools defined the rules of coding posture-related data and subsequent rating ergonomic risks based on the data. Various ergonomic assessment tools have been applied in the construction industry. For example, Quick Exposure Check for musculoskeletal risks (QEC), an ergonomic assessment tool, records the occurrence, frequency, and repetitiveness of working postures as well as the weight handled during the tasks. It was applied in a study to explore the effect of using self-compacting concrete on concrete workers' ergonomic risks [64]. Another ergonomic assessment tool, Ovako Working Pose Analysis System (OWAS), codes the postures of back, arms, and legs and classifies whole-body postures into four ergonomic risk levels. Previous studies have employed OWAS to investigate working postures during construction activities and compare the workload between aged and young construction workers [24,25,65,66]. To automate the ergonomic assessment process, some studies have automated posture data collection. However, the major challenge was that some ergonomic assessment tools, such as OWAS, only include the qualitative and semantic descriptions of postures, whereas the output of automatic data collection methods is generally numeric. To solve the problem, Zhang et al. defined the postures in OWAS with joint angle ranges [49]. On the other hand, there exist some other ergonomic tools studies that defined postures quantitatively (i.e. joint angles) such as ISO 11226 and Rapid Entire Body Assessment (REBA) [32,62]. Given the quantitative definition of ergonomic postures, IMUs, depth cameras, and 3D pose estimation algorithms were used to estimate 3D joint locations and calculate joint angles; then, the ergonomic risks of the corresponding postures were assessed according to the quantified rules according to the joint angles [32,45,49,62]

In addition to whole-body ergonomic assessment, previous studies also paid attention to the ergonomics of specific body parts. Contact-sensor-based methods were widely used in related studies to collect the kinematic data of certain body parts, such as the trunk and hand postures [12,19,67–71]. Furthermore, to facilitate timely intervention of non-ergonomic postures, an IMU-based system was developed to monitor the head, neck and trunk positions in real-time and a smartphone application processed the data and provided feedback for long-duration awkward postures [63].

To summarize, automatic posture-related data collection methods change the ergonomic assessment methods in the construction industry from being time-consuming and biased to objective and automatic. However, some ergonomic risk factors, such as the weight being handled, cannot be measured using the aforementioned methods. For the purpose, biomechanical analysis was used to integrate posture-related data and external forces, as discussed in the following section.

#### 4.2.2 Posture-related data in biomechanical analysis

Biomechanical analysis includes detailed analysis by calculating the joint forces or torques based on 3D poses. Given construction workers' poses and external forces, software packages, such as 3D Static Strength Prediction Program (3DSSPP) and OpenSim, can estimate joint forces or torques [72,73]. Through reviewing the videotapes of construction workers or collecting working poses from virtual reality, previous studies simulated typical work poses in 3DSSPP to calculate joint and muscle forces or torques, which, in turn, were used to calculate ergonomic exposures or assess the benefits of ergonomic interventions [74–77].

The development of automatic 3D pose data collection method makes it possible to collect data for biomechanical analysis under real conditions. Both RGB cameras and depth cameras have been applied to collect 3D pose data from construction sites, which are then fed into biomechanical models for joint force or torques calculation [50,78]. These methods could lead to accurate and detailed field-based ergonomic risk assessment methods.

### 4.3 Productivity evaluation

Posture-related data could be used in productivity evaluation through analyzing work status, such as working or resting. In a study by Bai, et al., posture classification algorithms were applied to evaluate productivity by classifying postures during tying rebar in a bridge construction project into three categories: effective work, contributory work, and ineffective work [16]. To evaluate the productivity of more diversified construction tasks, action recognition algorithms were employed so that both spatial and temporal features could be used to differentiate the actions. For example, Khosrowpour et al. classified actions into breaking, idling, walking, cutting and measuring, holding, picking up, and putting down [57]. Luo et al. further developed an action recognizer that could be used to recognize 16 construction activities, which were classified into three modes for productivity analysis, including productive mode, semi-productive mode, and non-productive mode [54].

## 5 Comparison of posture-related data collection methods for construction workers

The application of posture-related data collection methods in construction management could be considered as data streams, which are generated with various tools, processed as various data formats, and used in multiple scenarios. Figure 4 illustrates this concept of data streams, wherein the first two columns depict data collection methods, the third column represents data formats, and the last two columns highlight application scenarios. The width of each stream represents the number of articles of each type.

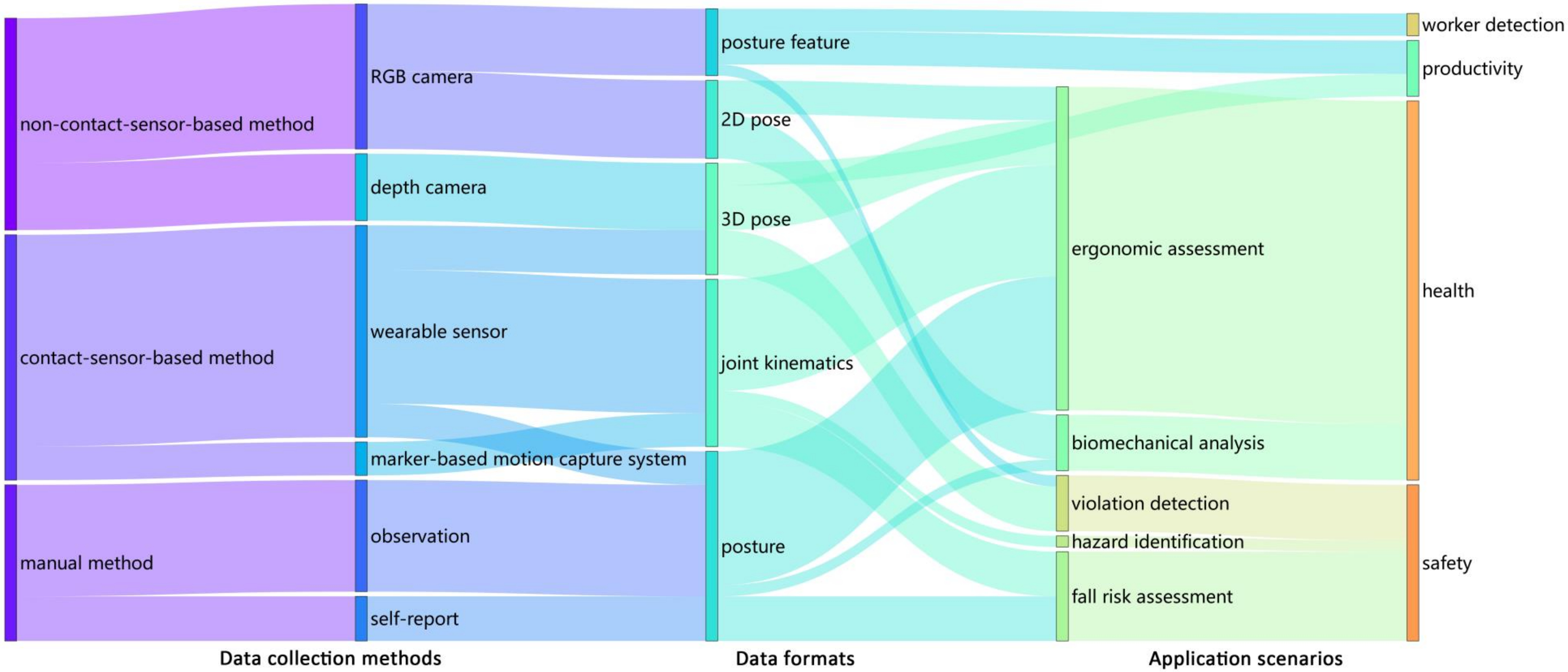


Figure 4 Data streams of applying posture-related data in construction site management

Figure 4 illustrates that contact-sensor-based methods have been most widely used followed by non-contact-sensor-based methods. Among the data collection tools, wearable sensors and RGB cameras are the most used tools. It could also be observed that wearable sensors and RGB cameras generated more types of data for different end objectives as compared to other tools. Postures and 3D poses are the most frequently used data formats, while 3D poses have been used for more application scenarios.

The following sections compare the posture-related data collection methods from the perspectives of data quality and feasibility factors, including intrusiveness, working environments, and cost. While data quality determines the accuracy and detailedness of the assessment results, intrusiveness, working environment, and cost are relevant factors to the feasibility of the assessment methods on the construction sites.

## 5.1 Data quality

Data quality is the fitness of "its intended uses in operations, decision making, and planning" [79]. In the case of posture-related data collection for construction workers, the following assessment criteria are selected: accuracy, consistency, and timeliness.

*Data accuracy* represents the degree to which data correctly describe the ground truth [80]. In the reviewed posture-related data collection methods, automatic methods achieve higher accuracy than manual observation problems. For contact sensors, a marker-based motion capture system provided an overall accuracy of 63±5 μm [37]; IMU can provide orientation accuracy of ±1º for dynamic conditions under different orientations [81]. The kinematic data collected by contact sensors could accurately reflect joint movements and bodily responses to the environment and is suitable for studying construction workers' behaviors with clinical approaches requiring highly accurate data. Postural stability evaluation with IMU and gait analysis with marker-based motion capture system are examples of such clinical approaches [29,35]. For non-contact sensors, the maximum error of Kinect V1 in the estimation of the joint centers were in the range of 2 cm to 4 cm [82]; while the maximum error of RGB camera-based methods for the same task is about 3 cm [83].

*Data consistency* means "the absence of difference when comparing two or more representations of a thing against a definition" [84]. Manual methods are of low consistency since the classification of posture-related data relies on manual judgments. For example, the same postures may be classified into different ergonomic risk levels by different observers [27]. Automatic methods, on the contrary, can collect posture-related data according to predefined rules, which helps to maintain a high consistency.

*Data timeliness* is the degree to which data represents reality from the required point in time [84]. Manual methods perform worse for this attribute since neither self-report nor systematic observation provides continuous results, whereas automatic methods can process and provide results in near real time. The computation time is highly related to the complexity of task and the computer configurations. The reported computation time is less than 0.20s for 3D pose estimation from a construction site image and about 13.36s for action recognition [54,62]. Short latency allows timely feedback and a shorter waiting period for decision making, such as the identification of safety violations and introducing subsequent interventions.

It is worth mentioning that some of the data formats can be transformed into another format and hence can be used for other end objectives. Figure 5 represents the relations among the data formats. 3D joint locations could be estimated as the double integral of 3D joint acceleration. Given the bone length constrains, 3D joint location and 3D joint angles are inter-calculable. 2D joint locations could be generated by projecting 3D joint locations on a specific plane. Finally, given the classification rules, postures could be recognized according to poses. In short, if a method could collect 3D joint locations, it could generate 3D poses, 2D poses, and postures. Further, if a method could provide acceleration data, it could generate all the other data formats. To recapitulate, contact-sensor-based methods can provide most types of posture-related data followed by RGB cameras.

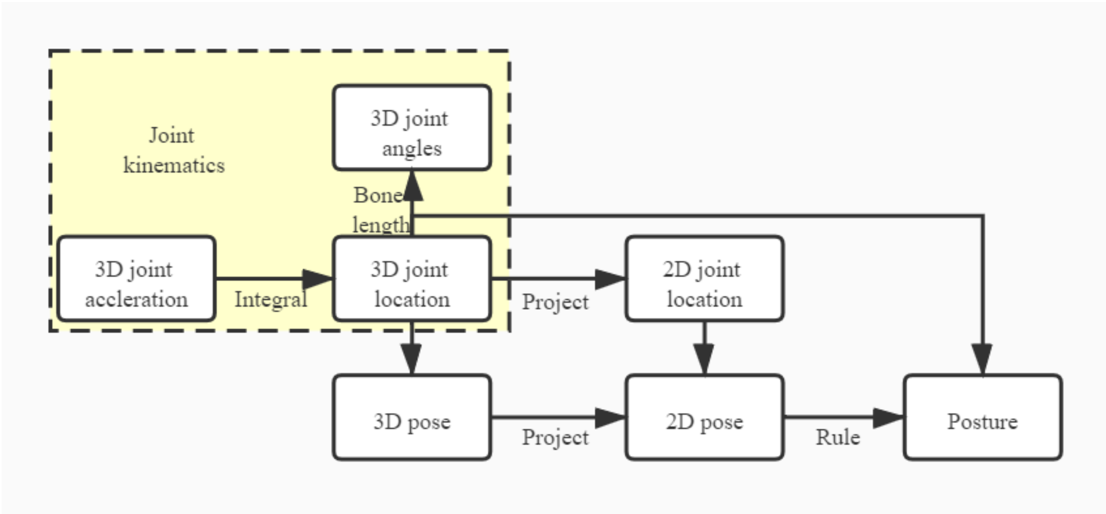


*Figure 5 Data transferability of different data formats*

## 5.2 Feasibility factors

### 5.2.1 Intrusiveness

Intrusiveness refers to the negative effects of the posture-related data collection methods on the workers' normal working operations. Earlier posture-related collection methods are intrusive to some extent. The intrusiveness of interview and self-report is positively correlated to data collection frequency and quality. To minimize intrusiveness, some researchers collected data twice a day; however, this resulted in sparse data that may exclude vital information [85]. To increase data richness, one has to collect data on a more frequent basis, which will interrupt the regular working of construction workers. Contact sensors are intrusive in a sense that they need to be tied tightly to workers' body segments, which will make workers feel uncomfortable [62]. In addition, some sensors need to be calibrated frequently, which also limits their application on real construction sites [62]. Non-contact-sensor-based methods, on the contrary, could collect data continuously without any sensors attached to the workers. However, electronic monitoring and surveillance in the workplace might not always be beneficial because of workers' feeling of constantly being watched/monitored, data privacy issues, reduced creative behaviors and more attention paid on quantity rather than quality [86,87].

### 5.2.2 Working environment

Construction sites are complex and dynamic, requiring a posture-related data collection method to be capable of adapting such an environment. As aforementioned, some data collection methods are not suitable for certain environmental conditions. For example, infrared-based depth cameras cannot be used outdoors because they are prone to direct sunlight [10]. In addition, marker-based motion capture systems are not suitable for construction sites, because they rely on special cameras that necessitate each tag, attached to a body joint, to be in direct line of sight of four cameras for an accurate 3D location assessment. RGB camera-based methods seem to be more feasible, which could capture images or videos in both outdoor and indoor environments, as well as in both near and far views. However, since the view of one camera cannot cover the whole site, a camera layout plan is needed for large-scale surveillance, as proposed in a relevant article [88]. Manual data collection methods, though intrusive and inaccurate, do not have specific requirements on construction site environments.

#### 5.2.3 Cost

Manual-based methods, such as self-report, interview, and manual observation, have little hardware cost but involve high labor and time costs. Automatic methods, on the contrary, are less labor and time consuming but require the purchasing and maintenance of hardware. Table 4 summarizes the price of wearable motion sensor systems, marker-based motion capture systems, depth cameras and RGB or closed-circuit television (CCTV) cameras. Each equipment type includes three example products.

For wearable motion sensor systems, commercially available products have been developed, which usually consist of IMU sensors, data dongles, drivers, software, and accessories. Such a system is user-friendly. The users could easily capture 3D poses by following the instructions without getting into complex system configurations or algorithm development. The price of each IMU ranges from 206 to 1000 USD. Resultantly, collecting data for numerous workers at a same time will cost a lot.

The price of a marker-based motion capture system depends upon the number of cameras required. Dense camera arrangement ensures that each marker is visible to at least four cameras for 3D localization and results in high accuracy. Camera resolution, frame-rate and synchronization method (wire/wireless) also influence the cost. The price of an eight-camera-system varies from 20,000 to 100,000 USD, which is much higher than that of wearable motion sensor systems.

For cameras, the price of a depth camera varies between 177 and 350 USD. While algorithms are required to extract 3D human body skeletons from depth images, the three depth cameras listed in Table 4 come with software development kits, which allow users to obtain 3D joint positions with a few lines of code. The price of the CCTV cameras ranges from 24 to 200 USD. These cameras could provide both RGB videos and infrared videos. However, special algorithms are needed to obtain 3D pose data using these [89,90]. In addition to cameras, the cost of a complete CCTV system also includes hard disks, cables, and installation fees. According to a vendor's quotation, the total price of installing a CCTV system including eight high-resolution cameras on construction sites is roughly about 12,000 USD and may vary significantly based on site conditions and system requirements.

*Table 4 The price of cameras or sensors applied in pose data collection*

| **Automatic pose data collection method** | **Example Product** | **Price [USD]** | **Reported application** | **References** |
|---|---|---|---|---|
| Wearable motion sensor system | 3-Space$^{TM}$ MoCap Starter Bundle (17 IMUs, sensor straps, 3 wireless dongles, drivers and software) | 3499 (About 206 per IMU) | 3D skeleton data of one person | [81] |
| | Xsens MVN | 600 per IMU | | [91] |
| | Motion Node | Over 1000 per IMU | | [92] |
| Marker-based motion capture system | OptiTrack (8-camera system)<br>Nokov (8-camera system)<br>Vicon (8-camera system) | About 20,000<br>About 40,000<br>About 100,000 | 3D skeleton data of single person with millimeter accuracy | [93]<br>[94]<br>[95] |
| Depth camera | Kinect for Windows V2 | 248 | 3D skeleton data of at most six people in 0.5~4.5 m | [96] |
| | Intel RealSense Depth Camera D435 | 177 | 3D skeleton data in 10 m | [97] |
| | TVico | 350 | 3D skeleton data in 0.6~5.0m | [98] |
| CCTV camera (including RGB camera) | Hikvision DS-2CE56C0T-IT3<br>Dahua 2PM Eyeball<br>Hikvision DS-2DE3304W-DE | About 70<br>About 34<br>205 | RGB videos for large area surveillance. 3D skeleton data could be extracted from the captured video frames with computer vision algorithms. | [99]<br>[100]<br>[101] |

The comparison reveals that the hardware cost of the marker-based motion capture systems is the highest, followed by the wearable systems. Depth cameras and CCTV cameras are cheaper in comparison. As for CCTV camera-based pose data collection methods, deep learning neural networks are usually trained to estimate the 2D or 3D poses from RGB images. These neural networks are usually data-hungry, which requires human resources to collect and label training data. As such, the labor cost of CCTV-camera-based methods is initially very high due to data labeling and network training requirements; however, subsequently, the cost decreases. Sensor-based methods, on the contrary, have a stable labor cost, including putting on/off sensors, recharging and replacing non-working sensors. For CCTV camera systems, they have a lower initial unit cost than that of wearable motion sensors. Besides, one CCTV camera could cover multiple workers, while one wearable motion sensor system can be used to collect the pose data of one worker only. When equipment failures occur, cameras are easier to be replaced than sensors, since the new sensor needs to be calibrated, synchronized, and integrated with previous sensors.

### 5.3 Summary

Table 5 compares the feasibility factors and the data quality of the posture-related data collection methods. The black boxes represent that the data collection methods have been used in previous studies to generate the respective data whereas the gray boxes represent that although the data collection methods can generate the data formats, they have yet been utilized. Over all, manual methods are inaccurate, and may result in excessive costs especially in the long term, but are suitable for various working environments. On the other hand, contact-sensor-based methods, including motion capture systems and IMUs could provide multiple types of accurate data but are highly intrusive. In addition, motion capture systems are pragmatic for indoor environments only. Non-contact-sensor-based methods are less intrusive however, suitable working environments for these depend on the type of camera being used. Their short-term cost arises from purchasing cameras and training algorithms. The long-term cost, however, would be lower than contact-sensor-based methods due to lower maintenance costs.

In conclusion, the choice of data collection methods depends upon the end objectives, construction site environment and budget constraints. For example, if the target is to obtain a qualitative evaluation of the working performance of a certain worker in a short period of time, manual methods could be used due to their easy implementation. For quantitative and detailed performance evaluation of a small number of workers, contact sensors could be used, such as using IMU. Non-contact sensors are suitable to study multiple workers over a long period of time. Specifically, infrared cameras are suitable when there is no direct sunlight, such as indoors whereas RGB cameras could be used when the lighting conditions are good.

*Table 5 The comparison of data quality and feasibility factors of posture-related data collection methods*

| Method | | Intrusiveness | Working Environments | Short-term cost | Long-term cost | Accuracy | Consistency | Timeliness | Transferability | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | Posture | 2D skeleton | 3D skeleton | Joint kinematics |
| Manual method | Observation | Low | Outdoor/Indoor | Low | High | Low | Low | Not real-time | | | | |
| | Self-report | Related to frequency | Outdoor/Indoor | | | | | | | | | |
| Contact-sensor-based methods | Motion capture system | High | Indoor | High | Middle | 2mm | High | 120-360 fps | | | | |
| | IMU | High | Outdoor/Indoor | | | 1 degree | High | 200 fps | | | | |
| Non-contact-sensor-based methods | Depth camera | Low | Indoor (Infrared cameras) | High | Low | 5.5 cm | High | 30 fps | | | | |
| | RGB camera | Low | Outdoor/Indoor | | | 3.9 cm | High | Subject to algorithm complexity and computer configuration | | | | |

## 6 Research gaps and possible future research directions

This section aims to identify the research gaps related to posture-related data collection tools and their applications in construction management. The research gaps related to the data collection tools are discussed in section 6.1 first. Then, the research gaps related to the application of posture-related data collection in construction management are discussed in section 6.2.

### 6.1 Research gaps related to posture-related data collection tools

#### 6.1.1 Mitigating the intrusiveness of contact-sensor-based methods

Contact-sensor-based methods are able to provide multiple types of posture-related data with high accuracy. Additionally, contact-sensor-based methods are the only choice to measure acceleration in the reviewed methods. However, their application on construction sites is limited due to their intrusiveness: workers have to wear multiple sensors in order to collect whole-body poses. Motion sensors with less intrusiveness might help solve the problem. Recently, a highly transparent and stretchable sensor has been developed to detect motions [102]. If attached to a worker's body, these soft and thin sensors would be able to collect posture-related data in a less intrusive way when compared to traditional wearable sensors.

Another solution to reduce intrusiveness is to use other sensors already carried by the workers, such as mobile phones, instead of asking them to wear additional sensors. Mobile phones have been used in a previous study to measure construction workers' motions, but the phones need to be tied tightly to the workers' limbs. Considering that mobile phones are usually larger than motion sensors, wearing mobile phones would be more intrusive than wearing motion sensors. A plausible solution might be using kinematic data collected with a mobile phone while in workers' pockets. Interested researchers could refer to the study by Kwapisz et al. [103], who identified daily activities such as walking, jogging, climbing stairs, sitting, and standing using mobile phone kept in one's pocket. In this way, workers do not need to wear any additional sensors. Nevertheless, the limitation is that a mobile phone could only collect the data at a certain body joint. Moreover, whether it can be used to differentiate construction working postures/actions or not remains to be tested.

#### 6.1.2 Increasing the accuracy of collection posture-related data from monocular RGB images

Monocular RGB cameras are widely used in the reviewed studies to collect workers' posture-related data from images because of their low cost and intrusiveness. Their performance could be further enhanced by considering the following recommendations.

**a. Dataset**

Machine learning and deep learning have been widely used in existing pose data collection methods. Dataset is a critical factor affecting the accuracy and generalization ability of the trained algorithms. A dataset consisting of various trades, workers, and activities contribute to the generalization ability of the trained algorithms. For example, a pose estimation algorithm might provide inaccurate results if it was trained on a rebar worker’s data, but was applied to a scaffold worker. Table 2 reveals that the number of data items in the reviewed pose datasets are quite limited as per the requirements of deep learning methods. A commonly used 3D pose dataset in computer science, known as Human 3.6M, consists of more than 3.6 million images from the daily life of 13 participants [104], which are much larger than the other datasets described in Table 2.

Another issue is the lack of standards regulating the content and format of posture-related datasets. According to Table 2, the datasets were established for different purposes with different formats from different trades of workers, making it difficult to compare the accuracy among the studies. If standards are established, researchers will be able to compare the performances of their approaches in a much more objective way.

**b. Surveillance from a far distance**

Construction sites are generally large outdoor areas. If captured from a far distance, an RGB picture will be of low resolution and fail to provide enough information to support pose estimation. Previous studies, however, were trained on the data collected from a short distance. For instance, Human 3.6M was built based on the data collected within an area of 4m x 3m

[104]. Future studies may consider training algorithms to estimate poses from low-resolution images. This could be achieved by using super-resolution algorithms proposed in computer science domain to recover high-resolution images from low-resolution images [105]. As a result, low-resolution images could be resolution-enhanced first, and then respective algorithms could be applied for pose estimation.

**c. Dark and Light environments**

Construction sites may present a variety of lighting environments. Some construction workers work in daylight or well-lit conditions, while others work in dark environments. In such cases, motion sensors, which are light-independent, could work well. On the contrary, RGB cameras can only work in well-lit conditions, while infrared-based depth cameras cannot work in direct sunlight. The combination of RGB camera-based methods and depth camera-based methods could be explored to see whether this could solve the issue. In addition, since LiDAR and radar use their own signals to illuminate the target, they might be more robust to scene lighting. Considering the previous success in posture recognition and pose estimation with LiDAR and radar [41,42], future research could test their applicability on construction sites to collect posture-related data.

**d. High angle shot**

In surveillance systems on construction sites, cameras are usually installed at a higher elevation than the workers, looking down upon the workers. However, the majority of the reviewed studies captured construction field images or video clips from an eye-level camera. The shape of the projections of a 3D skeleton from a high angle camera is different from that of an eye-level view. As a result, the trained algorithm could only be used for images captured from eye-level cameras and may lead to inaccurate results for images or video clips captured from high-angle cameras, indicating that most of the previous algorithms might fail in actual construction environments. To counter this issue, view-invariant features could compute the abstraction of images that are independent of the viewpoints. In the construction industry, Yan et al. used bone length ratio and joint angle ratio as view-invariant features to reconstruct 3D skeletons from the pixel locations of 2D skeletons and joints [48]. Theoretically, the method is applicable to any viewpoints if the key joints are visible. However, the method was only validated in a laboratory environment, where the camera could only be slightly higher than eye-level, which is comparably lower than the camera's height on actual construction site. As a result, whether the method is still applicable to the construction site images captured from a high angle or not, remains unknown. Future studies could retrain the algorithms with the 3D pose data and the images collected from real construction sites.

**e. Occlusion**

Although computer vision has been used widely in previous studies because RGB cameras are non-invasive, economical, and could cover the whole construction site with a few cameras, occlusion is still a challenge to this method. Construction sites are dynamic and complex, and workers are usually occluded in the captured images or video clips. As a result, estimating poses or recognizing postures under occlusion should be studied further. In computer vision, some algorithms have been proposed to estimate poses under occlusion [106]. Future research should apply/modify such algorithms to solve occlusion problems especially for construction sites.

Interestingly, to counter issues related to occlusion, in addition to visual signals, radio signals have been applied to human pose estimation. Zhao et. al proposed a through-occlusion human pose estimation method with radio signals in the Wi-Fi range, which could traverse walls and reflect off the human body [107]. However, its applicability to construction sites needs further investigation. First, metallic structures, which are widely prevalent on construction sites could block radio frequency signals and affect the pose estimation precision. Second, the method might fail in inter-person occlusion, which usually exists in crowded construction workspaces. Third, the operating distance of the radio signal used in the aforementioned study was about 13 m. On construction sites, the distances could be way much larger. Accordingly, radio signals with longer operating distances should be tested for construction site applications.

Another possible solution is drone cameras. Drone cameras can roam construction sites and capture workers' postures from multiple views, which might be helpful to solve the occlusion problems and 3D pose reconstruction from multiple images. Previous studies have successfully applied drone cameras to detect workers and estimate 2D poses [108,109]. However, capturing images or videos with drones is also faced with the issues of high angle shots and long distances. Super-resolution and de-occlusion algorithms might help address the problems [105,106].

## 6.2 Research gaps related to the application of posture-related data in construction management

### 6.2.1 Multifunctional posture-related data collection

Posture-related data plays a vital role in construction site management from the aspects of safety, health, and productivity. Current posture-related data collection methods, such as the computer-vision based 3D pose data collection method, could satisfy a number of needs including the identification of unsafe behavior, ergonomic assessment, and labor productivity evaluation. Previous studies have focused only on one application, rather than using a comprehensive approach. Future research should focus on an integrated analysis of safety, health, and productivity using a single data collection system.

### 6.2.2 Environment awareness integration

For construction workers, the surroundings, including tools, materials, and environments, could provide information related to behaviors. For example, if a worker is working on a steel bar mesh, he or she might be a rebar worker. Such awareness of the surroundings could further benefit site management. Considering unsafe behavior identification, the awareness of surroundings could provide us with the status of workers, nearby tools, and machinery, which could better assist in detecting unsafe behaviors. Similarly, for ergonomic analysis, tool or material identification could help us to estimate external loads for biomechanical analysis. In a previous study, the external loads were estimated based on manual records [50]. If the tools and materials held by workers could be automatically identified, the external loads could be estimated in a more automatic method, aiding better ergonomic analysis together with posture-related data.

### 6.2.3 Privacy problems in data collection

Although the importance of construction workers' posture-related data has been proved in previous research, there might exist privacy problems in collecting construction workers' posture-related data on construction sites. Workers might feel apprehensive about working under the surveillance of sensors, cameras, or observers. To the authors' best knowledge, studies exploring this issue are very scarce. Accordingly, research is warranted to fill this gap.

Another problem is data ownership. A construction project usually involves multiple stakeholders. As a result, who owns the data, with whom the data can be shared, and what type of information has to be censored, are becoming critical issues. Previous studies have proposed a multi-server information-sharing approach on a private cloud to address the above issues for Building Information Modelling (BIM)-based projects, which may provide a possible solution to posture-related data ownership and privacy. However, the consent of workers still remains an unresolved issue [110].

## 7 Conclusion

This paper presented a comprehensive review of existing studies on posture-related data collection in the construction industry. The descriptive statistics of 57 articles (covered in this study) show that construction workers' posture-related data has been a hot topic of research in the last decades, which drew the attention of researchers from multiple disciplines, including construction management, industrial engineering, computer science, and occupational health. The articles were then reviewed based on data collection methods and application scenarios. The reviewed studies were first categorized into three classes including manual methods, contact-sensor-based methods, and non-contact-sensor-based methods. For each data collection method, the working principles, and examples of its application in the construction industry were introduced. Then, the articles were classified based on their applications i.e. safety, health, and productivity. Further, a comparison was made to assess the performance of each posture-related data collection method. The performance was assessed based on data quality and feasibility for construction sites, which shows the advantages of motion sensors and RGB-image-based pose estimation. Then, research gaps and possible future research directions, especially for sensor-based methods and the RGB-image-based methods, are provided. Additionally, recommendations have been proposed for some of the limitations.

Overall, this study serves as a comprehensive reference for academia and any practitioners interested in posture-related data collection for construction workers. For the purpose of academia, this study summarizes the performances of posture-related data collection methods in construction management. Besides, research gaps and suggestions are presented for future research to consider. For the construction industry, this study provides a detailed summary and comparison of available posture-related data collection tools, as well as their application scenarios. Thus, industrial practitioners will be able to identify the most suitable data collection methods as per their requirements.

Despite an extensive review, this study, however, possesses the following limitations. First of all, the application of posture-related data in robotics is not considered in the research. Though this idea has been proposed in a study, it is used in a simulation case instead of on-site application [111]. Secondly, the parameters of pose data collection tools, such as accuracy, frequency, and cost, came only from the reviewed studies, following a protocol described in Section 2. However, there might exist other data collection tools with a better performance which could not be integrated into this review.

**Acknowledgment**

The authors acknowledge the support of the Research Grants Council of Hong Kong [PolyU 15210720, PolyU 152047/19E] and the Innovation and Technology Commission of Hong Kong [ITP/020/18LP].

## Reference


[1] T. Wang, Value added of U.S. construction as a percentage of GDP 2018, Statista. (2019). https://www.statista.com/statistics/192049/value-added-by-us-construction-as-a-percentage-of-gdp-since-2007/ (accessed March 6, 2020).

[2] Australian Bureau of Statistics, Australian National Accounts: National Income, Expenditure and Product, (2019). https://www.abs.gov.au/ausstats/abs@.nsf/Latestproducts/5204.0Main Features22018-19?opendocument&tabname=Summary&prodno=5204.0&issue=2018-19&num=&view= (accessed March 6, 2020).

[3] National Bureau of Statistics of China, National data, 2019. http://www.stats.gov.cn/tjsj/zxfb/201907/t20190731_1683002.html (accessed March 6, 2020).

[4] DMG Events, Construction Market in India, Mumbai, 2015. https://www.arbinternational.es/ARCHIVO/documentos/sectorial/1431507298_construction-market-report.pdf.

[5] Commonly Used Statistics | Occupational Safety and Health Administration, (no date). https://www.osha.gov/data/commonstats (accessed July 21, 2020).

[6] D. Wang, F. Dai, X. Ning, Risk Assessment of Work-Related Musculoskeletal Disorders in Construction: State-of-the-Art Review, Journal of Construction Engineering and Management. 141 (2015) 04015008. doi:10.1061/(ASCE)CO.1943-7862.0000979.

[7] X. Wang, X.S. Dong, S.D. Choi, J. Dement, Work-related musculoskeletal disorders among construction workers in the United States from 1992 to 2014, Occupational and Environmental Medicine. 74 (2017) 374–380. doi:10.1136/oemed-2016-103943.

[8] F. Barbosa, J. Mischke, M. Parsons, Improving construction productivity, McKinsey. (2017). https://www.mckinsey.com/industries/capital-projects-and-infrastructure/our-insights/improving-construction-productivity (accessed May 18, 2019).

[9] S. Han, S. Lee, F. Peña-Mora, Vision-Based Detection of Unsafe Actions of a Construction Worker: Case Study of Ladder Climbing, Journal of Computing in Civil Engineering. 27 (2013) 635–644. doi:10.1061/(ASCE)CP.1943-5487.0000279.

[10] H. Guo, Y. Yu, Q. Ding, M. Skitmore, Image-and-Skeleton-Based Parameterized Approach to Real-Time Identification of Construction Workers' Unsafe Behaviors, Journal of Construction Engineering and Management. 144 (2018) 04018042. doi:10.1061/(ASCE)CO.1943-7862.0001497.

[11] Y. Yu, H. Guo, Q. Ding, H. Li, M. Skitmore, An experimental study of real-time identification of construction workers' unsafe behaviors, Automation in Construction.

82 (2017) 193–206. doi:10.1016/j.autcon.2017.05.002.

[12] N.D. Nath, R. Akhavian, A.H. Behzadan, Ergonomic analysis of construction worker's body postures using wearable mobile sensors, Applied Ergonomics. 62 (2017) 107–117. doi:10.1016/j.apergo.2017.02.007.

[13] D. Al Madani, A. Dababneh, Rapid entire body assessment: A literature review, American Journal of Engineering and Applied Sciences. 9 (2016) 107–118. doi:10.3844/ajeassp.2016.107.118.

[14] L. McAtamney, E. Nigel Corlett, RULA: a survey method for the investigation of work-related upper limb disorders, Applied Ergonomics. 24 (1993) 91–99. doi:10.1016/0003-6870(93)90080-S.

[15] B. Buchholz, V. Paquet, L. Punnett, D. Lee, S. Moir, PATH: A work sampling-based approach to ergonomic job analysis for construction and other non-repetitive work, Applied Ergonomics. 27 (1996) 177–187. doi:10.1016/0003-6870(95)00078-X.

[16] Y. Bai, J. Huan, S. Kim, Measuring Bridge Construction Efficiency Using the Wireless Real-Time Video Monitoring System, Journal of Management in Engineering. 28 (2012) 120–126. doi:10.1061/(ASCE)ME.1943-5479.0000061.

[17] P. Turaga, R. Chellappa, V.S. Subrahmanian, O. Udrea, Machine Recognition of Human Activities: A Survey, IEEE Transactions on Circuits and Systems for Video Technology. 18 (2008) 1473–1488. doi:10.1109/TCSVT.2008.2005594.

[18] S. Kumar Bohidar, K. Patel, P.K. Sen, Robotic Sensor: Contact and Non-contact Sensor, International Journal of Advance Research In Science And Engineering (IJARSE). 4 (2015). https://pdfs.semanticscholar.org/c401/069cbca1f465555c60da98806a3d6cb59279.pdf (accessed January 14, 2020).

[19] A.J. Caban-Martinez, K.M. Santiago, J. Stillman, K.J. Moore, D.A. Sierra, J. Chalmers, M. Baniak, M.M. Jordan, Physical Exposures, Work Tasks, and OSHA-10 Training Among Temporary and Payroll Construction Workers, Journal of Occupational and Environmental Medicine. 60 (2018) E159–E165. doi:10.1097/JOM.0000000000001267.

[20] L.A.M. Elders, A. Burdorf, Interrelations of risk factors and low back pain in scaffolders, Occupational and Environmental Medicine. 58 (2001) 597–603. doi:10.1136/oem.58.9.597.

[21] A. DiDomenico, R.W. McGorry, Y.H. Huang, M.F. Blair, Perceptions of postural stability after transitioning to standing among construction workers, Safety Science. 48 (2010) 166–172. doi:10.1016/j.ssci.2009.07.006.

[22] American Psychological Association, Systematic observation – APA Dictionary of Psychology, American Psychological Association. (2019).

https://dictionary.apa.org/systematic-observation (accessed August 6, 2019).

[23] G. Li, P. Buckle, Current techniques for assessing physical exposure to work-related musculoskeletal risks, with emphasis on posture-based methods, Ergonomics. 42 (1999) 674–695. doi:10.1080/001401399185388.

[24] V. Louhevaara, Is the physical work load equal for ageing and young blue-collar workers?, International Journal of Occupational Safety and Ergonomics. 24 (1999) 559–564. doi:10.1016/S0169-8141(98)00061-4.

[25] T.-H. Lee, C.-S. Han, Analysis of Working Postures at a Construction Site Using the OWAS Method, International Journal of Occupational Safety and Ergonomics. 19 (2013) 245–250. doi:10.1080/10803548.2013.11076983.

[26] S. Tak, B. Buchholz, L. Punnett, S. Moir, V. Paquet, S. Fulmer, H. Marucci-Wellman, D. Wegman, Physical ergonomic hazards in highway tunnel construction: Overview from the Construction Occupational Health Program, Applied Ergonomics. 42 (2011) 665–671. doi:10.1016/j.apergo.2010.10.001.

[27] A. Golabchi, S. Han, A.R. Fayek, A fuzzy logic approach to posture-based ergonomic analysis for field observation and assessment of construction manual operations, Canadian Journal of Civil Engineering. 43 (2016) 294–303. doi:10.1139/cjce-2015-0143.

[28] G. Dudek, M. Jenkin, Inertial Sensors, GPS, and Odometry, in: B. Siciliano, O. Khatib (Eds.), Springer Handbook of Robotics, Springer Berlin Heidelberg, Berlin, Heidelberg, 2008: pp. 477–490. doi:10.1007/978-3-540-30301-5_21.

[29] H. Jebelli, C.R. Ahn, T.L. Stentz, Fall risk analysis of construction workers using inertial measurement units: Validating the usefulness of the postural stability metrics in construction, Safety Science. 84 (2016) 161–170. doi:10.1016/j.ssci.2015.12.012.

[30] H. Kim, C.R. Ahn, K. Yang, Identifying Safety Hazards Using Collective Bodily Responses of Workers, Journal of Construction Engineering and Management. 143 (2017). doi:10.1061/(ASCE)CO.1943-7862.0001220.

[31] J.A. Hess, L.D. Kincl, K. Davis, The impact of drywall handling tools on the low back, Applied Ergonomics. 41 (2010) 305–312. doi:10.1016/j.apergo.2009.08.001.

[32] E. Valero, A. Sivanathan, F. Bosche, M. Abdel-Wahab, Analysis of construction trade worker body motions using a wearable and wireless motion sensor network, Automation in Construction. 83 (2017) 48–55. doi:10.1016/j.autcon.2017.08.001.

[33] A. Alwasel, A. Sabet, M. Nahangi, C.T. Haas, E. Abdel-Rahman, Identifying poses of safe and productive masons using machine learning, Automation in Construction. 84 (2017) 345–355. doi:10.1016/j.autcon.2017.09.022.

[34] J. Chen, J. Qiu, C. Ahn, Construction worker's awkward posture recognition through supervised motion tensor decomposition, Automation in Construction. 77 (2017) 67–81.

doi:10.1016/j.autcon.2017.01.020.

[35] S.S. Chiou, C.S. Pan, A. Bhattacharya, Kinematics and kinetics of gait on stilts: Identification of risk factors associated with construction stilt use, Ergonomics. 51 (2008) 1814–1829. doi:10.1080/00140130801961885.

[36] P. Simeonov, H. Hsiao, J. Powers, D. Ammons, T. Kau, A. Amendola, Postural stability effects of random vibration at the feet of construction workers in simulated elevation, Applied Ergonomics. 42 (2011) 672–681. doi:10.1016/j.apergo.2010.10.002.

[37] M. Windolf, N. Götzen, M. Morlock, Systematic accuracy and precision analysis of video motion capturing systems—exemplified on the Vicon-460 system, Journal of Biomechanics. 41 (2008) 2776–2780. doi:10.1016/j.jbiomech.2008.06.024.

[38] H. Sarbolandi, D. Lefloch, A. Kolb, Kinect range sensing: Structured-light versus Time-of-Flight Kinect, Computer Vision and Image Understanding. 139 (2015) 1–20. doi:10.1016/j.cviu.2015.05.006.

[39] M. Liu, S. Han, S. Lee, Tracking-based 3D human skeleton extraction from stereo video camera toward an on-site safety and ergonomic analysis, Construction Innovation. 16 (2016) 348–367. doi:10.1108/CI-10-2015-0054.

[40] S. Han, S. Lee, A vision-based motion capture and recognition framework for behavior-based safety management, Automation in Construction. 35 (2013) 131–141. doi:10.1016/j.autcon.2013.05.001.

[41] G. Diraco, A. Leone, P. Siciliano, Human posture recognition with a time-of-flight 3D sensor for in-home applications, Expert Systems with Applications. 40 (2013) 744–751. doi:10.1016/j.eswa.2012.08.007.

[42] A. Sengupta, F. Jin, R. Zhang, S. Cao, mm-Pose: Real-Time Human Skeletal Posture Estimation using mmWave Radars and CNNs, IEEE Sensors Journal. (2020) 1–1. doi:10.1109/jsen.2020.2991741.

[43] M.-W. Park, I. Brilakis, Construction worker detection in video frames for initializing vision trackers, Automation in Construction. 28 (2012) 15–25. doi:10.1016/j.autcon.2012.06.001.

[44] M. Memarzadeh, M. Golparvar-Fard, J.C. Niebles, Automated 2D detection of construction equipment and workers from site video streams using histograms of oriented gradients and colors, Automation in Construction. 32 (2013) 24–37. doi:http://dx.doi.org/10.1016/j.autcon.2012.12.002.

[45] S.J. Ray, J. Teizer, Real-time construction worker posture analysis for ergonomics training, Advanced Engineering Informatics. 26 (2012) 439–455. doi:10.1016/j.aei.2012.02.011.

[46] Z. Wang, G. Liu, G. Tian, A Parameter Efficient Human Pose Estimation Method Based on Densely Connected Convolutional Module, IEEE Access. 6 (2018) 58056–58063.

doi:10.1109/ACCESS.2018.2874307.

[47] J. Martinez, R. Hossain, J. Romero, J.J. Little, A Simple Yet Effective Baseline for 3d Human Pose Estimation, in: 2017 IEEE International Conference on Computer Vision (ICCV), IEEE, 2017: pp. 2659–2668. doi:10.1109/ICCV.2017.288.

[48] X. Yan, H. Li, C. Wang, J. Seo, H. Zhang, H. Wang, Development of ergonomic posture recognition technique based on 2D ordinary camera for construction hazard prevention through view-invariant features in 2D skeleton motion, Advanced Engineering Informatics. 34 (2017) 152–163. doi:10.1016/j.aei.2017.11.001.

[49] H. Zhang, X. Yan, H. Li, Ergonomic posture recognition using 3D view-invariant features from single ordinary camera, Automation in Construction. 94 (2018) 1–10. doi:10.1016/j.autcon.2018.05.033.

[50] Y. Yu, H. Li, X. Yang, L. Kong, X. Luo, A.Y.L. Wong, An automatic and non-invasive physical fatigue assessment method for construction workers, Automation in Construction. 103 (2019) 1–12. doi:10.1016/j.autcon.2019.02.020.

[51] A. Khosrowpour, J.C. Niebles, M. Golparvar-Fard, Vision-based workface assessment using depth images for activity analysis of interior construction operations, Automation in Construction. 48 (2014) 74–87. doi:10.1016/j.autcon.2014.08.003.

[52] S. Han, M. Achar, S. Lee, F. Peña-Mora, Empirical assessment of a RGB-D sensor on motion capture and action recognition for construction worker monitoring, Visualization in Engineering. 1 (2013) 6. doi:10.1186/2213-7459-1-6.

[53] J. Yang, Z. Shi, Z. Wu, Vision-based action recognition of construction workers using dense trajectories, Advanced Engineering Informatics. 30 (2016) 327–336. doi:10.1016/j.aei.2016.04.009.

[54] X. Luo, H. Li, D. Cao, Y. Yu, X. Yang, T. Huang, Towards efficient and objective work sampling: Recognizing workers’ activities in site surveillance videos with two-stream convolutional networks, Automation in Construction. 94 (2018) 360–370. doi:10.1016/j.autcon.2018.07.011.

[55] M.A. Khan, M. Sharif, T. Akram, M. Raza, T. Saba, A. Rehman, Hand-crafted and deep convolutional neural network features fusion and selection strategy: An application to intelligent human action recognition, Applied Soft Computing. 87 (2020) 105986. doi:10.1016/j.asoc.2019.105986.

[56] Y. Chung, P.J. Haas, E. Upfal, T. Kraska, Unknown Examples & Machine Learning Model Generalization, (2018) 1–11. http://arxiv.org/abs/1808.08294.

[57] A. Khosrowpour, J.C. Niebles, M. Golparvar-Fard, Vision-based workface assessment using depth images for activity analysis of interior construction operations, Automation in Construction. 48 (2014) 74–87. doi:10.1016/j.autcon.2014.08.003.

[58] L. Ding, W. Fang, H. Luo, P.E.D. Love, B. Zhong, X. Ouyang, A deep hybrid learning

model to detect unsafe behavior: Integrating convolution neural networks and long short-term memory, Automation in Construction. 86 (2018) 118–124. doi:10.1016/j.autcon.2017.11.002.

[59] J. Suarez, R.R. Murphy, Using the Kinect for search and rescue robotics, in: 2012 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR), IEEE, 2012: pp. 1–2. doi:10.1109/SSRR.2012.6523918.

[60] Y. He, B. Liang, Y. Zou, J. He, J. Yang, Depth Errors Analysis and Correction for Time-of-Flight (ToF) Cameras, Sensors. 17 (2017) 92. doi:10.3390/s17010092.

[61] W. Umer, M.F. Antwi-Afari, H. Li, G.P.Y. Szeto, A.Y.L. Wong, The prevalence of musculoskeletal symptoms in the construction industry: a systematic review and meta-analysis, International Archives of Occupational and Environmental Health. 91 (2018) 125–144. doi:10.1007/s00420-017-1273-4.

[62] Y. Yu, X. Yang, H. Li, X. Luo, H. Guo, Q. Fang, Joint-Level Vision-Based Ergonomic Assessment Tool for Construction Workers, Journal of Construction Engineering and Management. 145 (2019) 04019025. doi:10.1061/(ASCE)CO.1943-7862.0001647.

[63] X. Yan, H. Li, A.R. Li, H. Zhang, Wearable IMU-based real-time motion warning system for construction workers' musculoskeletal disorders prevention, Automation in Construction. 74 (2017) 2–11. doi:10.1016/j.autcon.2016.11.007.

[64] R. Rwamamara, P. Simonsson, Self-Compacting Concrete Use for Construction Work Environment Sustainability, Journal of Civil Engineering and Management. 18 (2012) 724–734. doi:10.3846/13923730.2012.723399.

[65] K.W. Li, C. Lee, Postural Analysis of Four Jobs on Two Building Construction Sites: an Experience of Using the OWAS Method in Taiwan, Journal of Occupational Health. 41 (1999) 183–190. doi:10.1539/joh.41.183.

[66] D.P. Gilkey, T.J. Keefe, P.L. Bigelow, R.E. Herron, K. Duvall, J.E. Hautaluoma, J.S. Rosecrance, R. Sesek, Low back pain among residential carpenters: Ergonomic evaluation using OWAS and 2D compression estimation, International Journal of Occupational Safety and Ergonomics. 13 (2007) 305–321. doi:10.1080/10803548.2007.11076731.

[67] P.M. Ludewig, T.M. Cook, Translations of the Humerus in Persons With Shoulder Impingement Symptoms, Journal of Orthopaedic & Sports Physical Therapy. 32 (2002) 248–259. doi:10.2519/jospt.2002.32.6.248.

[68] X. Shen, I. Awolusi, E. Marks, Construction Equipment Operator Physiological Data Assessment and Tracking, Practice Periodical on Structural Design and Construction. 22 (2017) 04017006. doi:10.1061/(ASCE)SC.1943-5576.0000329.

[69] W. Lee, E. Seto, K.-Y. Lin, G.C. Migliaccio, An evaluation of wearable sensors and their placements for analyzing construction worker's trunk posture in laboratory

conditions, Applied Ergonomics. 65 (2017) 424–436. doi:10.1016/j.apergo.2017.03.016.

[70] W. Umer, H. Li, G.P.Y. Szeto, A.Y.L. Wong, Identification of Biomechanical Risk Factors for the Development of Lower-Back Disorders during Manual Rebar Tying, Journal of Construction Engineering and Management. 143 (2017) 04016080. doi:10.1061/(ASCE)CO.1943-7862.0001208.

[71] M.F. Antwi-Afari, H. Li, D.J. Edwards, E.A. Paern, D.-G. Owusu-Manu, J. Seo, A.Y.L. Wong, Identification of potential biomechanical risk factors for low back disorders during repetitive rebar lifting, Construction Innovation-England. 18 (2018). doi:10.1108/CI-05-2017-0048.

[72] The University of Michigan, 3D Static Strength Predication Program, Ver. 2.0, User's Manual, (1995). https://c4e.engin.umich.edu/tools-services/3dsspp-software/ (accessed April 18, 2018).

[73] S.L. Delp, F.C. Anderson, A.S. Arnold, P. Loan, A. Habib, C.T. John, E. Guendelman, D.G. Thelen, OpenSim: Open-Source Software to Create and Analyze Dynamic Simulations of Movement, IEEE Transactions on Biomedical Engineering. 54 (2007) 1940–1950. doi:10.1109/TBME.2007.901024.

[74] C.S. Pan, S.S. Chiou, Analysis of biomechanical stresses during drywall lifting, International Journal of Industrial Ergonomics. 23 (1999) 505–511. doi:10.1016/S0169-8141(98)00016-X.

[75] P.S. Dasgupta, L. Punnett, S. Moir, S. Kuhn, B. Buchholz, Does drywall installers' innovative idea reduce the ergonomic exposures of ceiling installation: A field case study, Applied Ergonomics. 55 (2016) 183–193. doi:10.1016/j.apergo.2016.02.004.

[76] J. Seo, S. Lee, J. Seo, Simulation-Based Assessment of Workers' Muscle Fatigue and Its Impact on Construction Operations, Journal of Construction Engineering and Management. 142 (2016) 04016063. doi:10.1061/(ASCE)CO.1943-7862.0001182.

[77] A. Golabchi, S. Han, J. Seo, S. Han, S. Lee, M. Al-Hussein, An Automated Biomechanical Simulation Approach to Ergonomic Job Analysis for Workplace Design, Journal of Construction Engineering and Management. 141 (2015) 04015020. doi:10.1061/(ASCE)CO.1943-7862.0000998.

[78] J. Seo, R. Starbuck, S. Han, S. Lee, T.J. Armstrong, Motion Data-Driven Biomechanical Analysis during Construction Tasks on Sites, Journal of Computing in Civil Engineering. 29 (2015) B4014005. doi:10.1061/(ASCE)CP.1943-5487.0000400.

[79] T.C. Redman, Data and Data Quality Defined, in: Data Quality: The Field Guide, Digital, Web, Boston, 2001: p. 241.

[80] R. Mahanti, Data Quality : Dimensions, Measurement, Strategy, Management, and Governance., ASQ Quality, Milwaukee, Wisconsin, 2018.

[81] Yost Labs, 3-Space™ Blue Tooth, (2019). https://yostlabs.com/product/3-space-

bluetooth/ (accessed January 10, 2020).
[82] A. Mobini, S. Behzadipour, M. Saadat Foumani, Accuracy of Kinect's skeleton tracking for upper body rehabilitation applications, Disability and Rehabilitation: Assistive Technology. 9 (2014) 344–352. doi:10.3109/17483107.2013.805825.
[83] Y. Yu, H. Li, X. Yang, Yantao Yu, Heng Li, Xincong Yang, 3D Posture Estimation from 2D Posture Data for Construction Workers, in: M. Al-Hussein (Ed.), Isarc2019, The International Association for Automation and Robotics in Construction, Banff, AB, Canada, 2019: pp. 26–34. doi:10.22260/ISARC2019/0004.
[84] N. Askham, D. Cook, M. Doyle, H. Fereday, M. Gibson, U. Landbeck, R. Lee, C. Maynard, G. Palmer, J. Schwarzenbach, The six primary dimensions for data quality management, The Global Data Management Community. (2013) 6. https://www.whitepapers.em360tech.com/wp-content/files_mf/1407250286DAMAUKDQDimensionsWhitePaperR37.pdf.
[85] H. Guo, Y. Yu, T. Xiang, H. Li, D. Zhang, The availability of wearable-device-based physical data for the measurement of construction workers' psychological status on site: From the perspective of safety management, Automation in Construction. 82 (2017) 207–217. doi:10.1016/j.autcon.2017.06.001.
[86] P.J. Holland, B. Cooper, R. Hecker, Electronic monitoring and surveillance in the workplace, Personnel Review. 44 (2015) 161–175. doi:10.1108/PR-11-2013-0211.
[87] K. Ball, Workplace surveillance: an overview, Labor History. 51 (2010) 87–106. doi:10.1080/00236561003654776.
[88] X. Yang, H. Li, T. Huang, X. Zhai, F. Wang, C. Wang, Computer-Aided Optimization of Surveillance Cameras Placement on Construction Sites, Computer-Aided Civil and Infrastructure Engineering. 33 (2018) 1110–1126. doi:10.1111/mice.12385.
[89] C.-H. Chen, D. Ramanan, 3D Human Pose Estimation = 2D Pose Estimation + Matching, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2017: pp. 5759–5767. doi:10.1109/CVPR.2017.610.
[90] X. Zhou, Q. Huang, X. Sun, X. Xue, Y. Wei, Towards 3D Human Pose Estimation in the Wild: A Weakly-Supervised Approach, in: 2017 IEEE International Conference on Computer Vision (ICCV), IEEE, 2017: pp. 398–407. doi:10.1109/ICCV.2017.51.
[91] MVN, MVN Analyze, (2019). https://www.xsens.com/products/mvn-analyze (accessed January 10, 2020).
[92] MotionNode, MotionNode Miniature Inertial Measurement Unit, (2019). https://www.motionnode.com/ (accessed January 10, 2020).
[93] OptiTrack, OptiTrack - Build Your Own Motion Capture System, (2019). https://optitrack.com/systems/ (accessed January 10, 2020).
[94] Nokov Mocap, About Nokov Motion Capture System, (2019).

https://www.nokov.com/en/about (accessed January 10, 2020).

[95] Vicon, Vicon Visualization Tool | Build Your Ideal Motion Capture System, (2019). https://www.vicon.com/visualization/ (accessed January 10, 2020).

[96] Microsoft, Kinect for Windows v2 Xbox Sensor with Kinect Adapter for Windows, (2019). https://support.xbox.com/en-HK/xbox-on-windows/accessories/kinect-for-windows-v2-setup#c8652d85386a4baba138b73ac11e0899 (accessed January 10, 2020).

[97] Intel® RealSense™, Depth Camera D435 – Intel® RealSense™ Depth and Tracking Cameras, (2019). https://www.intelrealsense.com/depth-camera-d435/ (accessed January 10, 2020).

[98] TVico, TVico - Interactive Android™ Box, (2019). https://tvico.io/ (accessed January 10, 2020).

[99] Hikvision, DS-2CE56C0T-IT3 1MP CMOS EXIR Turret Camera Key Features, 2019. https://www.hikvision.com/hk/products/Turbo-HD-Products/Turbo-HD-Cameras/Value-Series/DS-2CE56C0T-IT3/ (accessed January 10, 2020).

[100] Dahua Technology USA, Eyeball Cameras – Dahua Technology USA Inc, (2019). https://us.dahuasecurity.com/product-category/network-products/fixed-cameras/pro/eyeball-cameras-pro/ (accessed January 10, 2020).

[101] Hikvision, DS-2DE3304W-DE 3MP Network Mini PTZ Dome Camera Key Features, 2019. https://www.hikvision.com/en/products/IP-Products/PTZ-Cameras/Pro-Series/DS-2DE3304W-DE/ (accessed January 10, 2020).

[102] X. Jing, H. Li, H.-Y. Mi, Y.-J. Liu, P.-Y. Feng, Y.-M. Tan, L.-S. Turng, Highly transparent, stretchable, and rapid self-healing polyvinyl alcohol/cellulose nanofibril hydrogel sensors for sensitive pressure sensing and human motion detection, Sensors and Actuators B: Chemical. 295 (2019) 159–167. doi:10.1016/j.snb.2019.05.082.

[103] J.R. Kwapisz, G.M. Weiss, S.A. Moore, Activity recognition using cell phone accelerometers, ACM SIGKDD Explorations Newsletter. 12 (2011) 74. doi:10.1145/1964897.1964918.

[104] C. Ionescu, D. Papava, V. Olaru, C. Sminchisescu, Human3.6M: Large Scale Datasets and Predictive Methods for 3D Human Sensing in Natural Environments, IEEE Transactions on Pattern Analysis and Machine Intelligence. 36 (2014) 1325–1339. doi:10.1109/TPAMI.2013.248.

[105] Z. Wang, J. Chen, S.C.H. Hoi, Deep Learning for Image Super-resolution: A Survey, (2019). http://arxiv.org/abs/1902.06068.

[106] F. Huo, E.A. Hendriks, Multiple people tracking and pose estimation with occlusion estimation, Computer Vision and Image Understanding. 116 (2012) 634–647. doi:http://dx.doi.org/10.1016/j.cviu.2011.12.006.

[107] M. Zhao, T. Li, M.A. Alsheikh, Y. Tian, H. Zhao, A. Torralba, D. Katabi, Through-

Wall Human Pose Estimation Using Radio Signals, in: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, IEEE, 2018: pp. 7356–7365. doi:10.1109/CVPR.2018.00768.

[108] R. Ashour, T. Taha, F. Mohamed, E. Hableel, Y.A. Kheil, M. Elsalamouny, M. Kadadha, K. Rangan, J. Dias, L. Seneviratne, G. Cai, Site inspection drone: A solution for inspecting and regulating construction sites, in: 2016 IEEE 59th International Midwest Symposium on Circuits and Systems (MWSCAS), IEEE, 2016: pp. 1–4. doi:10.1109/MWSCAS.2016.7870116.

[109] S. Penmetsa, F. Minhuj, A. Singh, S. Omkar, Autonomous UAV for Suspicious Action Detection using Pictorial Human Pose Estimation and Classification, ELCVIA Electronic Letters on Computer Vision and Image Analysis. 13 (2014) 18. doi:10.5565/rev/elcvia.582.

[110] J. Zhang, Q. Liu, Z. Hu, J. Lin, F. Yu, A multi-server information-sharing environment for cross-party collaboration on a private cloud, Automation in Construction. 81 (2017) 180–195. doi:10.1016/j.autcon.2017.06.021.

[111] M. Kurien, M.-K. Kim, M. Kopsida, I. Brilakis, Real-time simulation of construction workers using combined human body and hand tracking for robotic construction worker system, Automation in Construction. 86 (2018) 125–137. doi:10.1016/j.autcon.2017.11.005.