Human Factors and Systems Interaction, Vol. 52, 2022, 130–139
https://doi.org/10.54941/ahfe1002148

Instrumenting the Human into Safety 4.0

Saed Amer, Dana Alhashmi, Ravindra Goonetilleke, and Ahmad Mayyas

Khalifa University of Science and Technology, P O Box 127788, Abu Dhabi, UAE

ABSTRACT

Managing the workers' health and safety faces many challenges due to the dependency on human assessments especially when it comes to human monitoring and detecting non-conformance. Conventionally, HSE decision-making is achieved by collecting information from the worker himself or by an HSE officer making it mostly subjective and difficult to quantify and share. The team proposes a constant and continuous approach to empirically monitor the workers using machine vision capabilities along with smart decision-making tools to detect, recognize and classify human behaviors. The input of the system is coherent and effective while the output is unbiased, quantifiable, and communicable, the needed ingredients to integrate the human factors into Industry 4.0. The scope of this work focuses on the worker's health and safety by setting another building block in the Safety 4.0 vision. The proposed system consists of multiple integrated components including continuous video streaming devices, machine vision components, computer logic capability, communication schemes, and locally executed alarms. The system was tested on a simulated environment using a human factors simulation platform then was validated with actual environments with workers acting with HSE non-conformance while performing different tasks. The results show the system's ability to recognize the human location, posture and speed then compare it to the HSE guidelines. The results also show that the system was able to provide fast responses by giving warnings, reporting an incident to the management, or shutting the process down if an injury is recognized. Finally, the system generates data and reports that are ready to be transmitted onto the Internet of Things.

Keywords: Human factors engineering, Human factors simulation, Image processing, Musculoskeletal disorders, Autonomous HSE monitoring, Human posture

INTRODUCTION

Warehouses are an integral part of any supply chain and inventory management giving jobs to many workers. For example, in the United States alone, there are around 1,002,809 warehouse workers. Due to the sheer number of warehouse workers and the heavy physical tasks required by the warehouse workers, many physical injuries occur. According to the U.S. Bureau of Labor Statistics, the average rate of documented injuries in the warehousing business in 2018 was 5.1 for every 100 full-time workers. This rate accounted for over 8% of the 2.8 million non-fatal occupational injuries and illnesses recorded by private sector companies including cases that resulted during job transfers (Fady Attia).

Realizing the high level of injuries in warehouse distribution facilities, the U.S. occupational safety and health administration of more widely known as OSHA set of rules and regulations that warehouses must follow to ensure workers' safety such as the OSHA regulation 1910.176 for material handling and storage in a warehouse (United States Department of Labor). Countries do regular inspections on warehouse factories and set fines and punishments to companies that do not abide by these rules. For example, in October 2018 the municipality in the emirate of Abu Dhabi within the United Arab Emirates issued 44 fines and 199 warnings on companies failing to secure worker safety in only 3 days. Therefore, many companies hire safety supervisors or foremen to ensure workers are doing their jobs safely and efficiently. Safety and warehouse supervisors require high costs that include wages and other human resources costs. Furthermore, the number of supervisors is unfeasible to maintain constant surveillance on the workforce especially in huge warehouse distribution facilities with multiple workstations, therefore, the injuries within the workforce are still subsided (Ahmed Lari et al., 2020).

Many of the occupational, health and safety programs aim to reduce the risk that is identified as the sources of harm caused by the non-compliance to the ergonomic guidelines (Aje et al., 2018). Overcoming such a dilemma is approached with three main stages that are correlated and interdependent. The first stage suggests a proper detection of the discrepancies between the job requirements and human performances. The second stage requires proper assessments to quantify the level of deviation from the correct human to task balance. The third stage ensures that the gap between the human capabilities and the job requirements is reduced. As confirmed, the ergonomics gap develops many health and safety implications which include organizational, psychological, and/or physiological issues. Many of the physiological implications are classified as musculoskeletal disorders (MSDs) which can be seen as inflammations or deteriorations in different body structures. Detecting the risk of ergonomics misbalance is challenging due to the inconsistency of the signs that suggest mishaps such as repetitive tasks, overloading the worker, incorrect body postures, and trespassing dangerous zones (Ha et al., 2009). Another challenge is the subjective nature of risk detection due to the dependence on the human assessor to look for and observe the pattern of signs that need to be interpreted into risk.

Still, until today, MSDs cause many levels of illnesses and injuries to workers. Nonetheless, many other losses are to be blamed on MSDs including reduction of productivity and quality deterioration. Over half of the recorded work-related diseases in Europe are identified as MSDs. Other HSE issues are responsible for economic losses that reached up to forty percent for Europe (Bevan, 2015). According to the US Bureau of Labor, more than one-third of all accidents and health-related diseases in the USA are blamed on work-related MSDs (Gerr et al., 2013). Conventional approaches to address such issues include self-monitoring where the workers are tasked to continuously perform predefined assessments on themselves to ensure safety. Officer supervision is also considered to enforce HSE regulations. More effectively, human factors measurements are carried out with direct contact with the worker to study and detect ergonomics noncompliance (Bevan, 2015).

Hittanagi et al. propose a system that collects useful information from digital videos using a fixed camera with a tracking system that detects human movement based on the optical flow estimation with various combinations of enhanced computer vision and image processing technology. The system uses an algorithm in morphological manipulation with a median filter to remove noises and applies thresholds to remove unwanted objects. Also, the blob analysis is used to determine the type of object limits. the proposed system was able to detect and track the movements obtained from live streaming videos. The work also demonstrates optical flow and gray scaling, estimation of optical flow, image segmentation into blobs, and blob analysis (Hittanagi, 2020).

Ding et al. suggest the implementation of recognition techniques for human posture that employ features extraction and recognition then teach the system to search for them. The work proposes an algorithm that improves the accuracy of recognizing human posture. Their system looks for angles and distances to be identified as local relationships between each joint and the joints' universal position. The rule learning method is used to classify human posture. The results show that different human postures were effectively recognized by the algorithm. Also, obtained from their results, the implemented rule-based learning method has higher recognition capability than the ones obtained using traditional machine learning and CNN methods (Ding et al., 2020).

The project's primary purpose is to establish a virtual superintendent system that monitors workers in warehouse distribution facilities. The system keeps constant surveillance on the human workers and continuously looks for any problematic action that may threaten the workers' health and safety or disrupt the workflow within the warehouse. The system must then send warning signals to both the workers and supervisors based on the problematic actions done by the workers. The system must provide mitigation measures when injuries occur at the workplace. Also, the system will reduce HSE monitoring costs by alleviating the dependency on human foremen to supervise HSE conformance. Moreover, the system will record the worker's data such as productivity, posture, and attendance and keep the information in a secured database. Finally, the system will establish communication channels for the Internet of Things (IoT) among the different collaborative processes internally and externally. The main requirements of developing an autonomous surveillance system are the ability to monitor the worker's performance, detect all moving objects, send an emergency signal to workers and supervisors when a nonconformance is detected, and finally, identify the worker's coordinates, posture, and speed.

METHODOLOGY

The system employs four main components that include a video capturing system, a Human Factors Engineering (HFE) software, a computer vision tool, and a computer logic capability. Mathworks Matlab® is utilized in this work to perform the needed programming algorithms used to identify the workers' behaviors. The videos obtained by a camera are fed into Matlab®

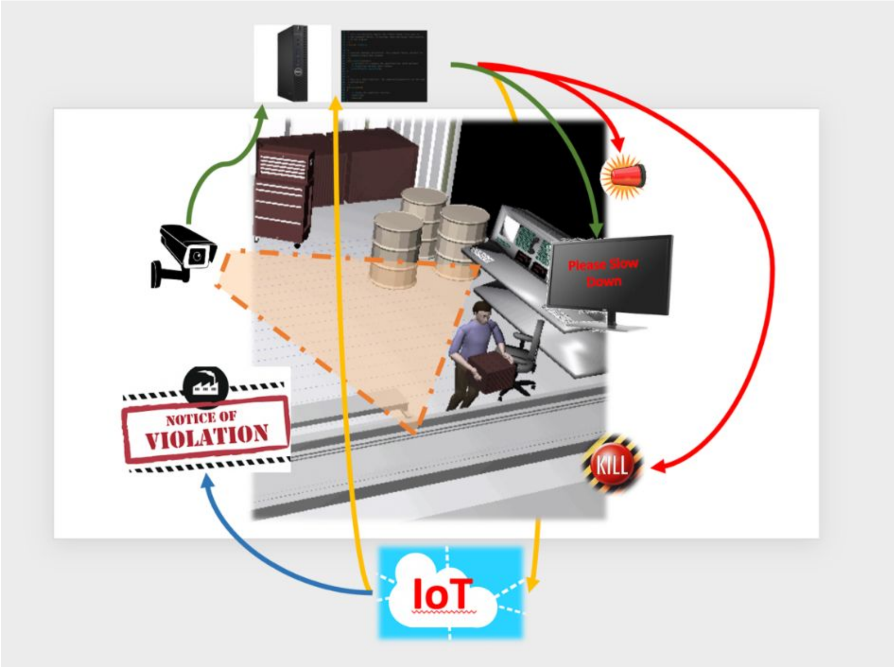

Figure 1: The system's layout model.

to be analyzed in the image processing toolbox. The HFE simulation software considered in this work is Siemens Jack® which is used to create accurate scenarios that commonly occur in the workplace, to identify the factors that cause injuries to the workers. Figure 1 presents the architecture of the proposed system which starts with a camera that streams live videos of the work area to an image processing tool for object identification and motion detection. The tool starts the process using Blob (binary large object) analysis, a fundamental technique for interpreting data visualizations by analyzing clusters of consistent images. This step is a preferred choice in cases where the objects being inspected possess discernible qualities that separate them from the background, such as moving objects in an otherwise static environment. allows for the detection of a person as a whole - not separately, like segregated limbs ((Troncale, 2000)). Under ideal conditions, if the size of the bounding box increases more than the threshold then we raise an alert to show something abnormal happened.

The video is inputted into the program, then partitioned into several picture frames that are converted to greyscale by eliminating the hue and saturation of their true colors while retaining their luminance. The indexed images are inspected for image regions separate from the masked environment, also known as the Background Region, whereupon they are isolated and tagged as being part of the Foreground Region using threshold values. As shown in Figure 2, the refined frame images are converted into a binary format and subjected to threshold computation. It checks the object limit properties, that is to say finding the coordinates of four pixels with minimum and maximum x-y-values, from which the width; given by x_{min} and x_{max}, and the

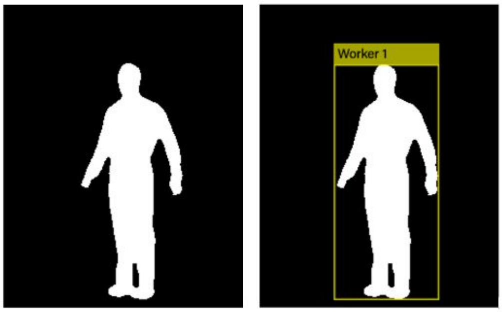

Figure 2: The bounding box (BBox).

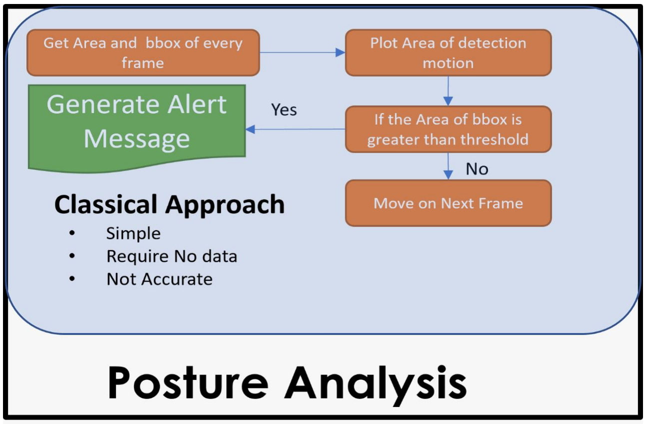

Figure 3: Posture analysis flow chart.

height; given by y_{min} and y_{max}, create the bounding box or BBox. If multiple objects are detected, they are inspected separately and split into individual Blobs, each marked by a BBox drawn over them and labeled by the object's ID (Hercik et al., 2020).

The light reflection on surfaces can cause the program to confuse pixel color change with motion, so IDs of non-existent workers may be registered in the log. This can be managed with heightened sensitivity for background environment separation. Figure 3 shows the BBox method for posture analysis which is achieved by implementing four phases:

1. The video frame is converted, smoothed, and refined
2. The BBox is plotted for any detected motion
3. Emergency or abnormal box sizes are monitored
4. If a case is detected, an alert is generated; if not, the program moves on to the next video frame

The Kalman filter algorithm is an extremely useful method of assigning measurable values to tracked objects in motion. It follows two steps; the first being scanning data points of clusters (the object's Blobs), where a centroid (mean point) is calculated for each cluster. Then, in the second iteration, the previous data points are reassigned to the cluster with the closest centroid and by alternating between these two steps, a convergence will eventually be reached when no apparent change can be deduced between cluster assignments. Moreover, a commonly used approach to compute the distance between a centroid point in a cluster and its surrounding data is the Euclidean distance formula as shown in Equation 1 (Zhu, 2009).

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \tag{1}$$

In simpler terms, the centroid will be used as the center of the object and will be used to identify the location of the object when mapped to the image frame and is the pixel location as x_c and y_c in the binary plane. Displayed as a coordinate point, the x_c and y_c values are calculated by adding all blob's pixels and dividing them by the number of the total pixels. This process can be defined mathematically in the below equations (Bourmaud et al., 2016).

$$x_c = \frac{1}{n} \sum_{k=0}^{n} xi \tag{2}$$

$$y_c = \frac{1}{n} \sum_{k=0}^{n} yi \tag{3}$$

Where n represents the number of pixels in the Blob and xi, yi are the coordinates of the pixels.

SIMULATION

Human Factors simulation is a powerful tool that can offer a way of analyzing environments in which physical prototyping is unattainable. Similarly, this project was conducted during the COVID-19 pandemic when ready access to the Human Factors lab and on-site congregation between team members was infeasible due to the tight restrictions. A safe and time-efficient alternative method of building a model environment for a distribution warehouse was found in simulation modeling, which was used to train the software system to look for incorrect work practices and abnormal human postures. Jack® was used to create several animated scenarios of workers in an order-picking workstation adjacent to an assembly line. Variations in the worker's pace in completing a single task were systematically controlled alongside a different collection of postures; each comprised an example of a work scenario. Jack® also has added features with an anthropometric database to perform human factors analysis. This helped in compartmentalizing postures into broad categories that included forms such as stretching, bending down, lying on the ground, and walking as presented in Figure 4.

As per the six primary warehouse processes, which are: receiving, put-away, storing, picking, packing, and shipping; a typical sequence of operations concentrated on the duties of an order picker working at a distribution warehouse was mapped out (Tarczynski, 2013). An expected work scenario would have a 5000 loop of the process flow which would span over an

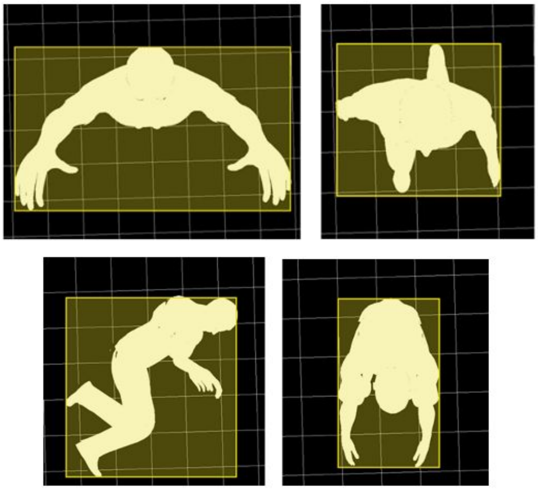

Figure 4: Model postures for stretching and walking lying on the ground and bending down.

Figure 5: Worker squatting in an order packaging scenario.

8-hour shift, which translates to 3 loops within an average of 2 minutes in simulated rendition.

As shown in Figure 5, a production cell was established with a computer desk, a station for material assembly, one for container pickup, and one for placing the finished product on a conveyor belt. The methodology of generating the simulation for the tasks was applied in three main steps using Jack®. These steps consisted of building the environment model, creating a simulation for tasks' sequence, and performing the ergonomic analysis. As the video of the worker is streamed into the system, a database is being filled with vital data that represent the worker's attendance, position, posture, time, and speed. The system allows for an organized output text file to be created and extracted into a readable format. The record is divided by each video frame, with Centroids x_c and y_c representing (x, y) coordinates, respectively, of the

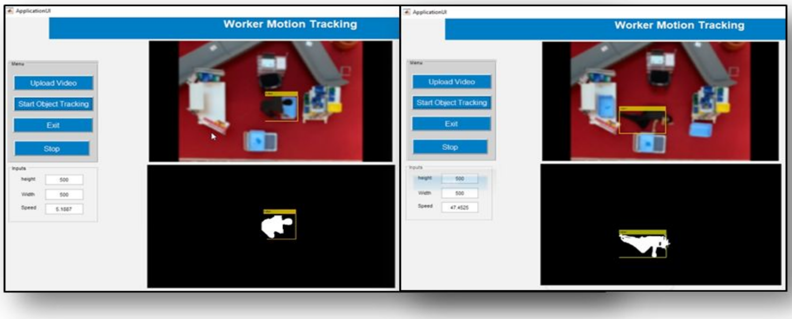

Figure 6: System performance with a live video feed.

BBox's center point. And "X_1, Y_1, X_2, Y_2" values as the BBox coordinates; (x_1, y_1), (x_1, y_2), (x_2, y_2), and (x_2, y_1).

System training is carried out by recording the data in an ideal scenario. The ideal scenario is orchestrated with the correct steps in the right order and schedule. The postures attained during this scenario are tested using the Human Factors simulation software for validations. The positions of the worker at the right times are also considered in the ideal scenario. The database collected during running the ideal scenario will be used as a reference to be compared to the current scenario. In other words, the system depends on measurable and unbiased assessments to detect any nonconformances. Nonetheless, a level of tolerance is maintained to make up for the acceptable human variations.

RESULTS AND DISCUSSION

The output file obtained from multiple scenarios declared the system's ability to detect the human and record the location size and speed of the worker. The figure below shows the GUI of the proposed system and the outcome obtained while running live testing. Normally, a neutral pose should not exceed 60,000 Pixel^2, as the worker moved through the workstation, most values were not exceeding a noticeable limit. Figure 4 illustrates that bending down to perform a pick produces a BBox that is stretched beyond the accepted BBox size which represents a threshold for posture conformance. the fainting scenario is expressed with a BBox size that was more than doubled a BBox size which is beyond a limit that is predefined to represent a worker's fall. In such cases, the system gives a clear indication that an abnormality.

Based on the team's analysis of the results the team set up the system to have a preset output range in which the warning system is triggered. There are different thresholds identified for the types of warnings that were obtained from mapping the BBox size, posture, and speed to the ones observed from the worker. For example, bending down will have an abnormal warning level and fainting will have an emergency warning level. The results exported in the output file were compared to what could be seen through real-life observations. The two values that accurately captured the two abnormal scenarios were recorded and compared to the average mean of the BBox size for the video frames that displayed both postures.

CONCLUSION

The developed system is able to input live videos and differentiate the workers from their surroundings by using image processing capabilities. The workers are given unique IDs; and their location, posture, and speed are monitored and recorded automatically by using bounding box analysis in Matlab. Parameters are given in a standard manner, ranging from normal, abnormal, and emergency alarming by general ergonomics and work area boundaries. The system is valid through accomplished video streaming and simulation series of scenarios that depicted the workers performing typical tasks done in the warehouse that can be deemed reckless or harmful workplace. The system was calibrated and tested on recreated live scenarios and on simulated scenarios that simulate warehouse workers using Siemens Jack®. This was done to ensure that the system works as intended. The system can be used to replace human foremen within a warehouse distribution facility and eliminate the liabilities associated with the human foreman. This includes the inability to keep constant surveillance of the workers, subjective analysis of the workers, and yearly costs.

However, the system still requires tailoring to be fitted in the industry which requires new reference scenarios then calibration with a new simulation and testing in the production floor. Furthermore, the system needs to be tested on multiple workers to identify each worker alone with the different postures.

REFERENCES

A. Troncale, "Moving theory into practice: Digital Imaging for Libraries and Archives," *Microform & Imaging Review*, vol. 29, no. 3, 2000.

C. E. P. Fady Attia, "5 insightful statistics related to Warehouse Safety," *Pallet Racking Inspection, Protection & Repair Solutions*. [Online]. Available: https://www.damotech.com/blog/5-insightful-statistics-related-to-warehouse-safety. [Accessed: 13-Feb-2022].

C. Ha, Y. Roquelaure, A. Leclerc, A. Touranchet, M. Goldberg, and E. Imbernon, "The French Musculoskeletal Disorders Surveillance Program: Pays de la Loire Network," *Occupational and Environmental Medicine*, vol. 66, no. 7, pp. 471–479, 2009.

F. Ahmed Lari, D. A. Abdulla Lari, and M. A. Darwish Abdulla Lari, "Construction of structural model for Strategic Foresighting and public policy," *International Journal of Sustainable Construction Engineering and Technology*, vol. 11, no. 2, 2020.

F. Gerr, N. B. Fethke, L. Merlino, D. Anton, J. Rosecrance, M. P. Jones, M. Marcus, and A. R. Meyers, "A prospective study of musculoskeletal outcomes among manufacturing workers," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 56, no. 1, pp. 112–130, 2013.

G. Bourmaud, R. Mégret, A. Giremus, and Y. Berthoumieu, "From intrinsic optimization to iterated extended Kalman filtering on lie groups," *Journal of Mathematical Imaging and Vision*, vol. 55, no. 3, pp. 284–303, 2016.

G. Tarczynski, "Warehouse real-time simulator how to optimize order picking time," *SSRN Electronic Journal*, 2013.

O. O. Aje, B. Smith-Campbell, and C. Bett, "Preventing musculoskeletal disorders in factory workers: Evaluating a new eight minute stretching program," *Workplace Health & Safety*, vol. 66, no. 7, pp. 343–347, 2018.

R. Hercik, Z. Machacek, J. Koziorek, J. Vanus, M. Schneider, and W. Walendziuk, "Continuity detection method in binary image signal," *Elektronika ir Elektrotechnika*, vol. 26, no. 6, pp. 4–9, 2020.

R. Hittanagi, "Vehicle detection, tracking and counting for traffic surveillance using digital image processing," *International Journal for Research in Applied Science and Engineering Technology*, vol. 8, no. 9, pp. 1061–1066, 2020.

S. Bevan, "Economic impact of musculoskeletal disorders (MSDS) on work in Europe," *Best Practice & Research Clinical Rheumatology*, vol. 29, no. 3, pp. 356–373, 2015.

S. Bevan, "Economic impact of musculoskeletal disorders (MSDS) on work in Europe," Best Practice & Research Clinical Rheumatology, vol. 29, no. 3, pp. 356–373, 2015.

"United States Department of Labor," *1910.176 - Handling materials - general. | Occupational Safety and Health Administration*. [Online]. Available: https://www.osha.gov/laws-regs/regulations/standardnumber/1910/1910.176. [Accessed: 13-Feb-2022].

W. Ding, B. Hu, H. Liu, X. Wang, and X. Huang, "Human posture recognition based on multiple features and rule learning," *International Journal of Machine Learning and Cybernetics*, vol. 11, no. 11, pp. 2529–2540, 2020.

Y.-Fang Zhu, "Background subtraction and color clustering-based moving objects detection," *2009 International Conference on Information Engineering and Computer Science*, 2009.