Predicting Sagittal Plane Lifting Postures From Image Bounding Box Dimensions

Runyu L. Greene, **Yu Hen Hu**, **Nicholas Difranco**, **Xuan Wang**, University of Wisconsin-Madison, USA, **Ming-Lun Lu**, National Institute for Occupational Safety and Health, Cincinnati, Ohio, USA, **Stephen Bao**, **Jia-Hua Lin**, Washington Department of Labor and Industries, Olympia, USA, and **Robert G. Radwin**, University of Wisconsin-Madison, USA

Objective: A method for automatically classifying lifting postures from simple features in video recordings was developed and tested. We explored if an "elastic" rectangular bounding box, drawn tightly around the subject, can be used for classifying standing, stooping, and squatting at the lift origin and destination.

Background: Current marker-less video tracking methods depend on a priori skeletal human models, which are prone to error from poor illumination, obstructions, and difficulty placing cameras in the field. Robust computer vision algorithms based on spatiotemporal features were previously applied for evaluating repetitive motion tasks, exertion frequency, and duty cycle.

Methods: Mannequin poses were systematically generated using the Michigan 3DSSPP software for a wide range of hand locations and lifting postures. The stature-normalized height and width of a bounding box were measured in the sagittal plane and when rotated horizontally by 30°. After randomly ordering the data, a classification and regression tree algorithm was trained to classify the lifting postures.

Results: The resulting tree had four levels and four splits, misclassifying 0.36% training-set cases. The algorithm was tested using 30 video clips of industrial lifting tasks, misclassifying 3.33% test-set cases. The sensitivity and specificity, respectively, were 100.0% and 100.0% for squatting, 90.0% and 100.0% for stooping, and 100.0% and 95.0% for standing.

Conclusions: The tree classification algorithm is capable of classifying lifting postures based only on dimensions of bounding boxes.

Applications: It is anticipated that this practical algorithm can be implemented on handheld devices such as a smartphone, making it readily accessible to practitioners.

Keywords: biomechanical models, spine, biomechanics, anthropometry, work physiology, gait, posture, job risk assessment, manual materials handling, spine, low back

Address correspondence to Robert G. Radwin, Department of Industrial and Systems Engineering, University of Wisconsin-Madison, 1550 Engineering Drive, Madison, WI 53706, USA; e-mail: rradwin@wisc.edu.

HUMAN FACTORS
Vol. 61, No. 1, February 2019, pp. 64–77
DOI: 10.1177/0018720818791367
Article reuse guidelines: sagepub.com/journals-permissions

INTRODUCTION

Back pain is among the most common workplace injuries and most costly to workers compensation, averaging over 20% of employer costs (Bhattacharya, Schulte, & Anderson, 2012). The U.S. Bureau of Labor Statistics (BLS) reported that in 2016, back injuries were a leading injury that resulted in days away from work or job transfer or restriction in a variety of industries (BLS, 2018). In 2015, musculoskeletal injuries such as sprains and strains caused by overexertion in lifting accounted for 31% of all the cases of nonfatal occupational injuries and illnesses requiring days away from work, remaining relatively unchanged from previous years (BLS, 2016). In 2014, the number of cases affecting the lumbar region accounted for 1,580 of the total 7,750 musculoskeletal injuries in the state of Wisconsin, or just over 20% (National Institute for Organizational Safety & Health [NIOSH], 2014).

Manual materials handling and lifting are contributing factors of low-back pain and injuries (Barondess et al., 2001; Bernard et al., 1997). Extreme working postures are a risk factor for low-back disorders (Bernard et al., 1997; da Costa & Vieira, 2010; Marras et al., 1995). Numerous factors have been identified as greatly impacting the force on the lower vertebrae (L4/L5 compression) during a lift, such as critical joint angles of the hip and knee that increase moment loading (Hwang, Kim, & Kim, 2009). These joint angles are good indicators of the amount of stress placed on the lower back during a lift because they affect the moment at the low back vertebrae. Of the different lifting postures, stoop lift is often associated with greater lower back moments resulting from the combination of knee and trunk angles (straight knees and bent trunk). Previous studies have reported association between stoop lift and low back disorders (e.g., Holmström, Lindell, & Mortiz, 1992; Myers et al., 1999).

Video recordings have become an indispensable tool in occupational ergonomics. A survey of Certified Professional Ergonomists conducted by Dempsey, McGorry, and Maynard (2005) reported that video cameras were the most commonly used tool (91.6%) and that they were rated "very useful" by most of the participants. With the advantages of objectivity, low cost, and little interference with work, video task evaluation tools have been drawing increasing attention. Furthermore, since video recordings are commonly used in industrial applications, new exposure analysis tools based on video for automatically extracting information should gain wide acceptance by both practitioners and workers.

Motion-tracking tools, including technologies that require markers, are limited in their ability to measure parameters such as body angles in industrial practice. Camera placement is often challenging in the workplace, while marker placement takes time to attach and can interfere with the work. Although many of these methods offer laboratory precision, such accuracy and precision may not be needed for occupational health and safety practice. Marker-less methods are similarly subject to variations in the observation and estimation techniques for projected angles and leads to less accuracy and high intra- and interobserver variability (Plantard, Shum, Le Pierres, & Multon, 2016). To combat these issues, some researchers have developed error-correction models that improve the accuracy of collected data, though it is still difficult to evaluate the effectiveness of these models (Spector, Lieblich, Bao, McQuade, & Hughes, 2014).

A variety of studies have explored computer vision–based methods for analyzing human motion and activities. Wang and Suter (2007) proposed a framework for recognizing human activities based on silhouette data extracted from videos. Marfia and Roccetti (2016) proposed a method that identifies stoop lifts based on lower back location in relation to the ground detected by the computer vision using silhouette information. Similarly, Seo, Yin, and Lee (2016) used silhouette information extracted using computer vision to identify postures in videos. Li and Lee (2011) developed a method that extracts skeletal models of workers from 2D videos with joint locations that can provide abundant information for ergonomic applications. Mehrizi et al. (2017; Mehrizi, Peng, Tang, et al. 2018; Mehrizi, Peng, Xu, et al., 2018) demonstrated a computer vision marker-less motion capture method to assess 3D pose, back moments, and joint kinematics of lifting tasks. These methods possess great potential for application in task evaluation. However, they are dependent on a controlled environment for information extraction and their ability to yield accurate conclusions in complex industrial settings where poor lighting, occlusion, and dynamic backgrounds are common.

Practical methods should be sufficiently robust to obtain reliable information without being compromised from various sources of noise. Advances in computer vision video technology have been applied in ergonomics for evaluating repetitive motion tasks (Chen, Hu, Yen, & Radwin, 2013), exertion frequency, and duty cycle (Akkas et al., 2016) and visualizing repetitive motion task factors (Greene, Azari, Hu, & Radwin, 2017). This approach extracts key spatiotemporal features of motion in a semi-automatic manner from conventional video recordings whereby the analysts interactively select a region of interest such as a hand or arm relative to a stationary region. This method relaxes the need for high-precision tracking rather than imposing an a priori model on the tracked activity. Consequently, the analysis complexity is greatly reduced and may be more tolerable of the numerous variations encountered in field video recordings of occupational tasks.

The current study explores how a practical computer vision approach can be applied to lifting tasks. While previous methods have focused on obtaining precise joint angles, a tool capable of classifying lifting postures that pose different levels of load to the low back is more pragmatic for ergonomics practice. We devise a practical method that is insensitive to challenging workplace conditions, such as poor illumination, poor vantage points, and obstructions, by relaxing the need for high precision and emphasizing extracting essential features from a video that are useful for assessing stress and strain on the lower back. This study explores if the features of a simple "elastic" rectangular bounding box continuously tracking the subject can be used for classifying three types of postures assumed while lifting or lowering an object at the origin or destination (i.e., standing, stooping, and squatting). The ultimate objective is to develop computer vision software for automatically classifying lifting postures based solely on bounding box dimensions without needing to measure joint angles or fit the data to skeletal models.

METHODS

Objectives and Overview

Our approach builds on the concept of extracting simple key features from the video image rather than fitting complex skeleton models used in motion tracking. Background subtraction, a video-processing technique that identifies moving regions by comparing every frame to the stationary background, is used to identify the moving foreground (i.e., the worker), thus not needing to track specific body parts. A bounding box is drawn tightly surrounding the foreground (Figure 1). Using this technique, we have devised a method for identifying the instant an object is first moved at the lift origin of the lift or last stops moving after the object is released at the lift destination. The dimensions of the bounding box and the image it encloses at the origin and destination of the lift are measured (i.e., height and width).

Mannequin poses of various lifts for varying locations were generated using the 3D Static Strength Prediction Program Version 7.0.0 (3DSSPP) from the University of Michigan (2017a). A bounding box was drawn around each mannequin, and the corresponding box height (H) and width (W) were used as a training set to develop a decision tree algorithm to identify the lifting posture of a subject based on bounding box dimensions. In this paper, a specific lifting posture is classified as one of three states: squatting, stooping, or standing.

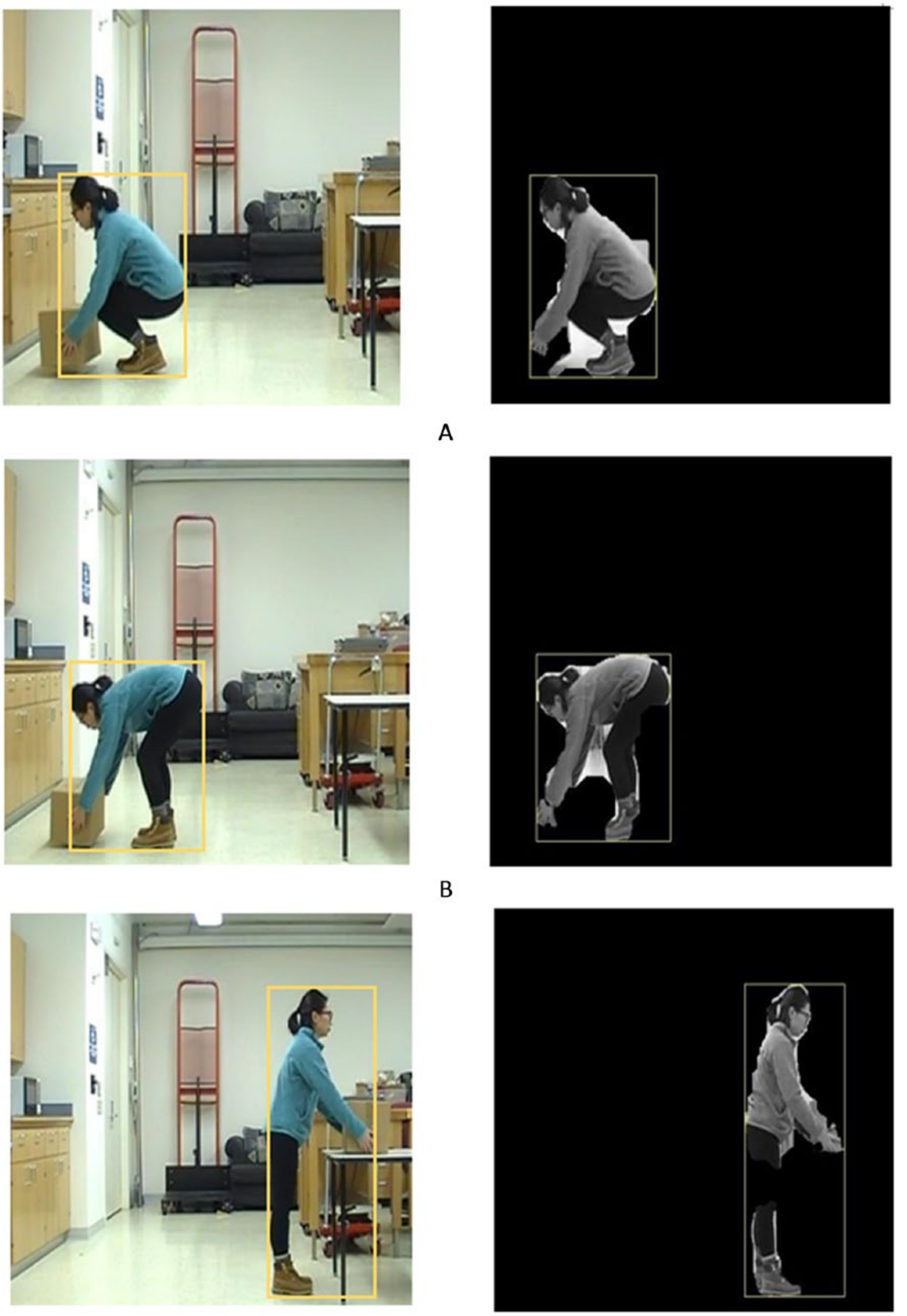
The resulting decision tree algorithm was tested using videos of actual industrial lifting tasks. The instant when an object in a video is lifted or released was identified, a bounding box was drawn around the subject, and the aspect ratio was entered into the algorithm for classifying the lifting posture assumed. Independently, the lifting posture was classified based on the subject's hip and knee angles and compared against the algorithm classifications to measure its accuracy.

Algorithm Development

The 3DSSPP program allows the manipulation of several factors of the lifting task and subject, such as the gender and anthropometry. A right-facing sagittal mannequin posed in the standing, stooping, and squatting positions, with corresponding bounding boxes, is illustrated in Figure 2. Previous research has demonstrated that lifting posture states can be classified from angle measurements of the hip and knee (Anderson, Chaffin, Herrin, & Matthew, 1985; Burgess-Limerick & Abernathy, 1997; Dysart & Woldstad, 1996). The definitions of torso and knee angles are illustrated in Figure 3.

A squat is classified when the included knee angle (α_k) is $\alpha_k < 130°$, and a stoop is classified when the torso angle (α_t) is $\alpha_t > 40°$ flexion from the vertical. If neither of these conditions are true, the lifting posture is classified as a stand. If both $\alpha_k < 130°$ and $\alpha_t > 40°$, the lifting posture is classified as a squat since stoop lifts are intended to be done with relatively straight legs. This classification comes from the notion that a squatting subject may reach farther out to lift an object, requiring greater hip flexion while the knees are still bent. These classification rules are illustrated in Figure 4. The angles used to classify these lifting postures for the models can be determined from data provided in 3DSSPP.

To consider the range of lifting locations, the manual lifting Threshold Limit Value (TLV) documentation developed by the American Conference of Governmental Industrial Hygienists (ACGIH) was used (ACGIH, 2017). The TLV classifies distinct zones (three horizontal and four vertical). The horizontal zones are defined by the horizontal hand distance from the center of the body for 30 cm, 60 cm, and 80 cm. The vertical zones are based on the anthropometry of the subject and use key landmarks, including mid-shin height, knuckle height when the arms are completely extended, and 8 cm below shoulder height. These zones expand from the floor to the upper reach limits.



A

B

C

Figure 1. (Left) Original video frames and (right) background subtracted computer vision images of a subject performing (A) squatting, (B) stooping, and (C) standing lifts during a lab-simulated task. In the task, the subject picks up an object from the shelf on the left of the video frame, transfers and lowers it, and places it on the ground in the right portion of the video frame. A rectangular bounding box encloses the subject in each image, independent of the object lifted.

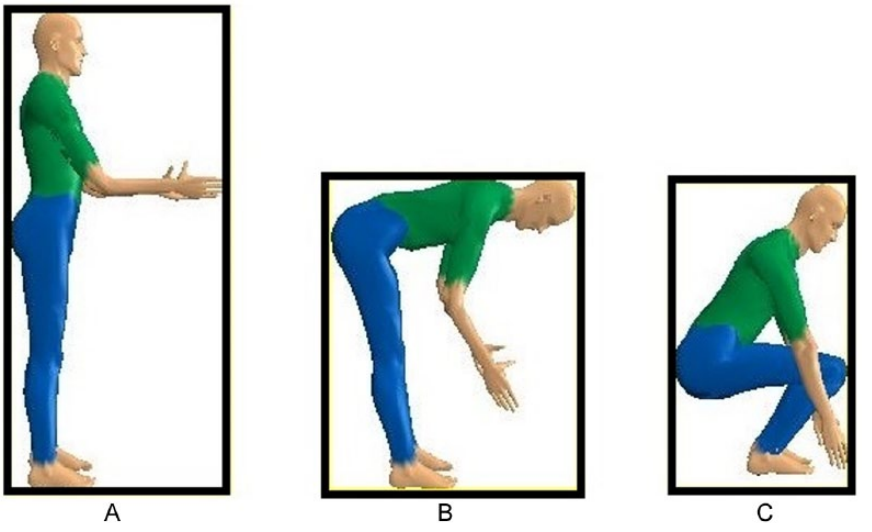


Figure 2. Examples from 3DSSPP of (A) standard standing, (B) stoop, and (C) squat lifts and their associated bounding box dimensions.

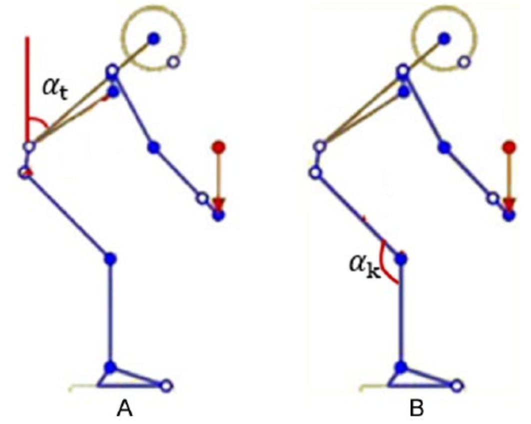


Figure 3. 3DSSPP models demonstrating (A) torso angle (α_t) and (B) knee angle (α_k) measured to classify lifting postures.

To reflect the differences in lifting posture between anthropometries at key zone boundary points, the 20 intersections of the horizontal and vertical zone markers on the TLV diagram were computed for each anthropometry using values from the 2012 Anthropometric Survey of U.S Army Personnel, and the hand location was entered into 3DSSPP to find the resulting bounding lifting posture state (Gordon et al., 2014). A representation of the 20 zone intersections is shown in Figure 5.

The lifting postures were generated by inputting the hand locations according to the mannequin's anthropometry and applying the posture prediction feature of 3DSSPP (Chaffin, 2008; Hoffman, Reed, & Chaffin, 2007). The bounding box dimensions of captured images in pixels were measured using University of Wisconsin MVTA software (Yen & Radwin, 1995). To determine the lifting state, the angles of the torso and upper leg of the subject were read from the "Body Segment Angles" window in 3DSSPP. The goal was to record measurements for attainable lifting tasks with hand positions in each of the 12 TLV zones and consider each of the three different lifting postures within each zone for different anthropometries to demonstrate variability of lifting posture states between models of varying sizes and gender. This involved utilization of subjects from six anthropometry classifications available in 3DSSPP, including 5th, 50th, or 95th percentile height males and females. To account for subjects of varying anthropometries, the mannequin standing height was used for normalizing the bounding box H and W. The respective dimensions H and W were ratios of the bounding box height to standing height and bounding box width to standing height, in pixels.

After excluding those locations that were unreachable due to anthropometry, bounding box data were generated for 105 distinct lifting postures generated by 3DSSPP across six different anthropometries and classified into the three

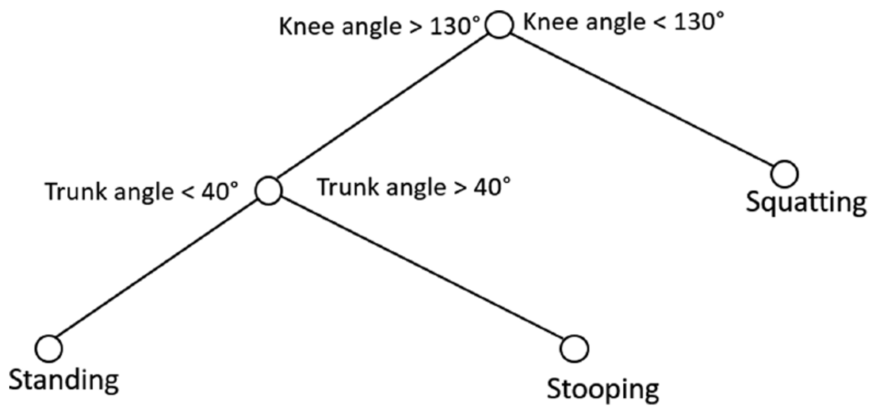


Figure 4. Rules of classifying standing, squatting, and stooping lifting postures.

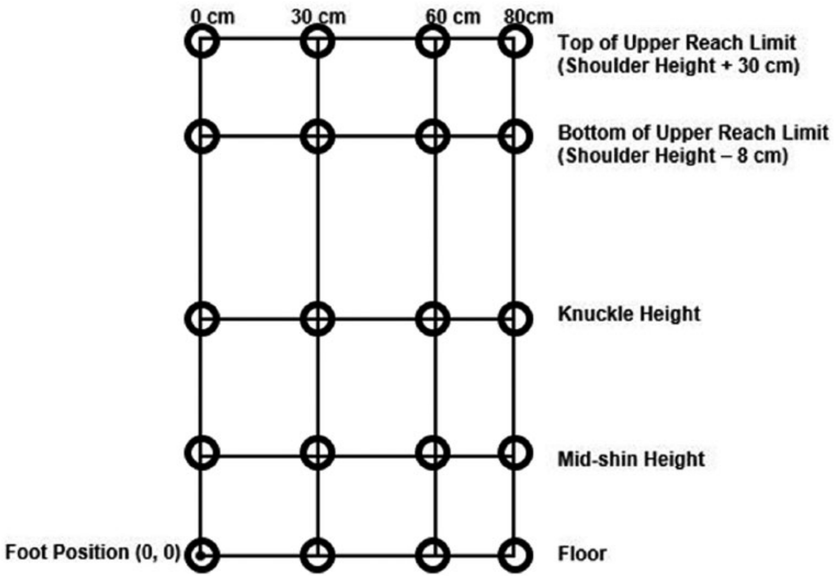


Figure 5. Representation of intercepts of three horizontal and four vertical zones.

lifting states. Of these, 43 were classified as squats, 13 as stoops, and 49 as stands. These locations reflect the variation of parameters across differing lifting states and demonstrate how parameters vary for a lift based on the lifting posture definitions used in this study. Since in the workplace it is not always possible to locate a camera angle perpendicular to the subject, the mannequin for each of the 105 tasks was rotated by 30° with respect to the sagittal plane, and bounding boxes were drawn and measured, creating 105 additional conditions.

An example of how the bounding box varies when the camera angle changes from a perpendicular view of the sagittal plane to a view rotated by ±30° is shown in Figure 6. Since the width of the bounding box is the projection of the horizontal distance from the farthest left to right point on the mannequin image in the video plane, the width when rotated 30°, $W_{30} = \cos(30°)\ W = .866\ W$ (Figure 7). Since the bounding box dimensions are identical when the camera is rotated 30° clockwise or counterclockwise, measurements were only taken when the camera was rotated clockwise in

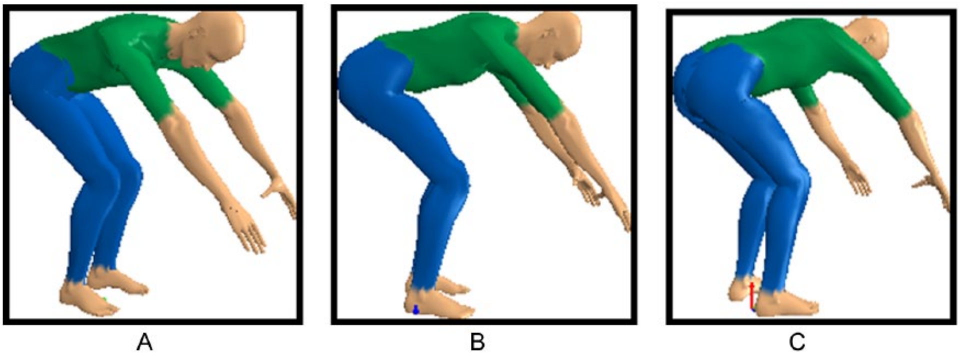


Figure 6. Examples of mannequin performing the same lifting posture with (A) the camera rotated counterclockwise by 30°, (B) perpendicular to sagittal plane, and (C) rotated clockwise by 30° and their associated bounding boxes.

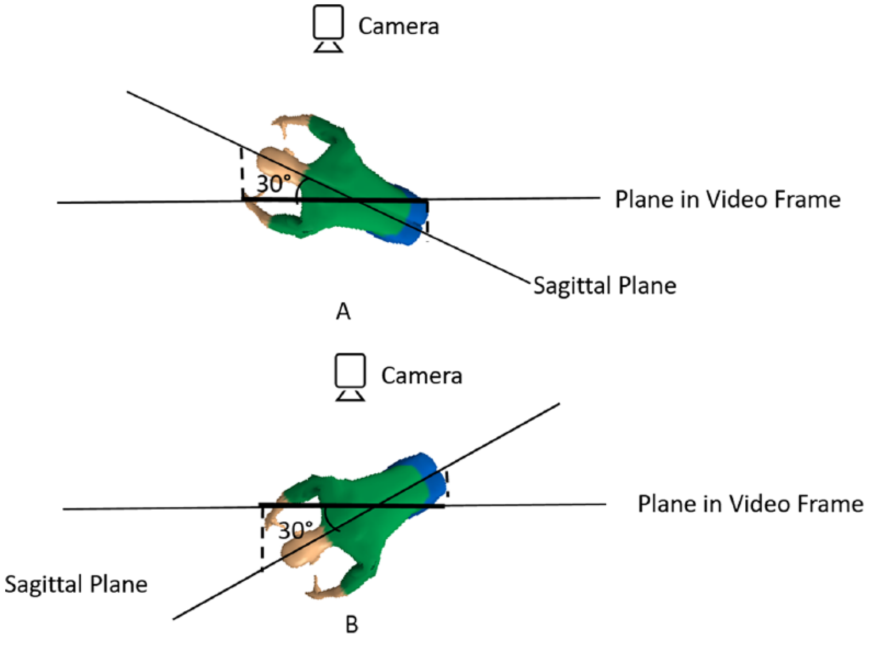


Figure 7. Demonstration of bounding box width as the projection of the mannequin's sagittal plane length on the plane on video frame, and the bounding box widths are identical for when the mannequin is rotated (A) counterclockwise and (B) clockwise.

pixels using the MVTA software. The standing height in pixels was also measured for each anthropometry, and the bounding box dimensions of each additional condition were normalized as ratios to the standing height of their associated anthropometry. Summary data for the bounding box dimensions of the 210 cases are provided in Table 1.

Because the 3DSSPP posture prediction algorithm prefers squat over stoop for the majority of entered hand locations, an imbalanced data set resulted in which the number of stoop lifts was significantly less than the other classes. Random oversampling with replacement was applied to avoid biasing the algorithm toward the more represented classes. An additional 66 data points were resampled from the stoop class in both the sagittal plane and rotated 30°, resulting in 92 stoop lifts, which was the average number of

TABLE 1: Average ± Standard Deviation of Box Dimensions, Normalized to Subject Standing Height

		Box Height (H)		Box Width (W)	
Lifting Posture	Cases	Sagittal Plane	30° Rotation	Sagittal Plane	30° Rotation
Squat	86	0.54 ± 0.17	0.50 ± 0.08	0.58 ± 0.09	0.49 ± 0.11
Stoop	26	0.78 ± 0.07	0.81 ± 0.07	0.65 ± 0.06	0.61 ± 0.06
Stand	98	1.02 ± 0.06	1.02 ± 0.07	0.43 ± 0.09	0.36 ± 0.09

cases in squat and standing lifts. In total, there were 276 data points in the training set used to develop the decision tree. After randomly reordering the data set, a classification and regression tree (CART) algorithm (Breiman, Friedman, Stone, & Olshen, 1984) was used to generate a tree model that uses the bounding box H and W to classify the data into the lifting posture categories (Mathworks, 2017). The accuracy of the tree model was evaluated using a 10-fold cross validation method (Breiman et al., 1984).

Algorithm Validation

A total of 76 static images of lifting postures encompassing a variety of manual material handling tasks were randomly selected from videos recorded in industrial settings to test the algorithm's ability to classify lifting postures (Bao, Howard, Spielholz, & Silverstein, 2016; Lu, Waters, Krieg, & Werren, 2014; Safetyvideopreviews, 2012; University of Michigan Center for Ergonomics, 2014, 2017b). These videos were recorded in various manufacturing plants and warehouses. Some examples of the tasks included lifting and stacking metal panels, lifting metal cylinders onto hooks on a conveyor, lifting and moving stacks of metal frames, lifting crates from a stack, and lifting a large, heavy stack of paper. The inclusion criteria for video frames are (a) the full body of the subject is visible at the origin and destination without twisting, (b) the load is not occluded by other objects, and (c) the camera is perpendicular to the subject's sagittal plane. Since only 10 squat lifts were found, to attain an equal number of lifting postures of each class in the test data set, 10 lifting postures were randomly selected from the stooping and standing, respectively, resulting in a data set of 30 lifting postures.

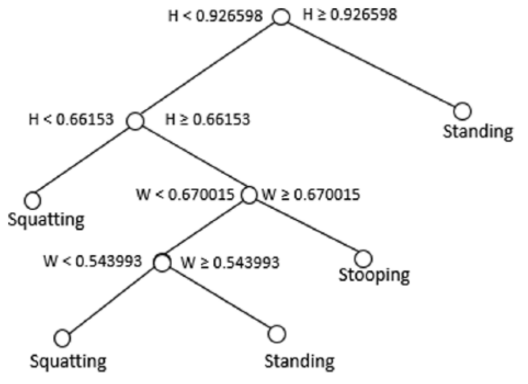


Figure 8. Decision tree generated from training data set, where H is the bounding box height normalized by standing height and W is the bounding box width normalized by standing height.

For each video frame containing the lifting postures in the validation data set, the bounding box height and width of the subject were measured using MVTA software (Yen & Radwin, 1995) and then normalized to ratios relative to the subject's standing height. The knee angle and trunk angle were also measured in MVTA to determine the classification of each lifting posture according to the definitions previously used.

This research complied with the American Psychological Association Code of Ethics and was approved by the Institutional Review Board (IRB) at the University of Wisconsin-Madison. Informed consent was obtained from each participant in accordance with the individual collaborating IRB.

RESULTS

Decision Tree Model

The resulting decision tree representation of the bounding box height and width for each state

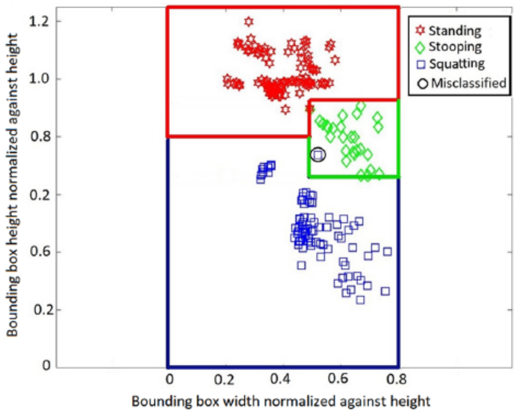


Figure 9. Spatial representation of decision tree classification results for the training set based on the bounding box height (H) and width (W), each normalized by the subject's stature.

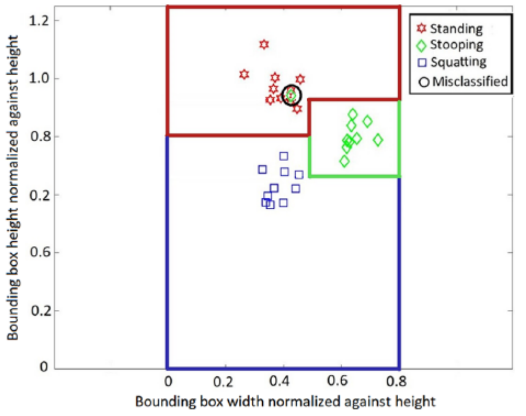


Figure 10. Spatial representation of decision tree classification results of the test set based on the bounding box height (H) and width (W), each normalized by the subject's stature.

TABLE 2: Confusion Matrix

	Algorithm Classification		
Ground Truth Classification	Squatting	Stooping	Standing
Squatting	85	1	0
Stooping	0	92	0
Standing	0	0	98

is shown in Figure 8. The spatial relationships for the decision tree demonstrate how the algorithm would classify postures based on the H and W of a bounding box are shown in Figure 9.

The resubstitution error obtained from applying the algorithm back against the training data set was 0.36%. The error rate of the tree model was 0.0108 using the 10-fold cross-validation method.

A confusion matrix was created containing the number of lifting posture states correctly and incorrectly classified (Table 2). There was only one error when testing the algorithm against the training data, misclassifying the lifting posture as stooping while the lifting posture was squatting. The sensitivity and specificity of squat, stoop, and standing, respectively, were 98.8% and 100.0%, 100.0% and 99.5%, and 100.0% and 99.5%.

Algorithm Validation

The validation misclassified 3.33% of the test set cases, in which 0 of 10 squats, 1 of 10 stoops, and 0 of 10 stands were misclassified. In the misclassified case, the stoop lifting posture was classified as standing. The sensitivity and specificity were 100.0% and 100.0% for squat, 90.0% and 100.0% for stoop, and 100.0% and 95.0% for standing, respectively. The distribution of normalized bounding box heights and widths of the test set and classification results using the decision tree algorithm generated with the training set are presented in Figure 10.

DISCUSSION

This research demonstrated the feasibility of classifying lifting postures based on only dimensions of bounding boxes tightly drawn around the subject, which are associated with the position of the hands and feet relative to the hip and knee angles, without direct angle measurements, instrumentation, markers, or fitting the image to a skeletal model. The results from the analysis demonstrate a relatively high rate of success for the algorithm correctly classifying a lifting posture. The resubstitution error of 0.36% obtained through applying the decision tree algorithm back to the original training set appears to be extremely small. This is because resubstitution error is a very optimistic measure of error rate, and we have provided other measures of error to have a more comprehensive idea about the performance of the algorithm. The 10-fold cross-validation error of the decision tree algorithm was 1.08%. To further understand the algorithm's ability to classify cases, a set of 30 actual lifting tasks were entered to test the algorithm, and only one case was misclassified. Because the algorithm contained low training and validation errors for the wide range of samples used, there is high confidence that this algorithm could be applied to identify the lifting posture state for arbitrary lifts.

Given that the full decision tree generated multiple permutations, the branches obtained for the final tree had shown up in most iterations. When compared with the alternative decision trees generated using the three parameters in addition to the bounding box height to width ratio, the training errors remained practically unchanged. It was determined that the height of the bounding box was effective as a distinguishing parameter as it was used at the first level of branching for all decision trees generated.

In the test set, one stoop lift was classified as standing. The trunk flexion angle of the subject did not significantly exceed the threshold for stoop lifts ($\alpha_T = 41.67°$), and the legs were straight ($\alpha_k = 180°$). Consequently, the height of the resulting bounding box exceeded the threshold for standing, misclassifying it as a stand. Although this misclassified a lower risk lifting posture (standing) as a higher risk lifting posture (stooping), when trunk flexion angle is small, trunk flexion angle contributes less to the moment on the low back.

In addition to measuring bounding boxes when the camera's viewing angle was perpendicular to the mannequin's sagittal plane, bounding box dimensions were measured when the camera was rotated $\pm 30°$ horizontally, for a 60° range off of the sagittal plane to account for when the camera is not perfectly aligned. The algorithm performed well despite a camera offset. It might be further improved by incorporating more camera angles, and the exact range of tolerance before lifting postures start to confound should be further studied to ensure appropriate limits. Future work will investigate methods of accommodating videos taken from various views so that the analysis is more tolerable of variations that may be encountered in field video recordings.

This research utilized the 12 TLV lifting zones and three anthropometries to account for the variability of lifting postures in relation to hand position for a given subject size and normalizing the bounding box dimensions as a ratio relative to standing height. The benefit of using 3DSSPP to train the algorithm is that a great number of lifting postures were easily generated using different TLV lifting zone boundary intersection locations. Since the lifting posture prediction feature of 3DSSPP automatically positions the subject according to an inverse kinematics algorithm based on the location of the hands in relation to the feet (University of Michigan Center for Ergonomics, 2017a), the resulting lifting postures in this research are reproducible. The lifting postures generated by 3DSSPP, however, are not actual lifts and therefore may not be representative of actual lifting postures due to variations in individual characteristics, workers' preferences, training, and environmental factors, including the nature of the object, obstructions, traction, and the workers' apparel (University of Michigan Center for Ergonomics, 2017a). Although the bounding box dimensions obtained from the test set in the current study show similar spatial distribution as those of the 3DSSPP-generated lifting postures and the test of the algorithm on industrial tasks produced satisfactory results, the algorithm still needs to be tested on a wide variety of lifting tasks. The current study does not consider the effect of load on posture.

Another future improvement is to identify additional easily recordable parameters that could help further distinguish lifting states and assess their risk to the lower back. One parameter that might be important is the position of the feet relative to the hands and the horizontal corners of the bounding box. There is potential that this parameter would correlate with the degree of hip and knee flexion, though it would require an adjustment of the hand position measurements since 3DSSPP determines those relative to the feet being the origin of the coordinate plane. To accomplish this, we are currently working on methods of detecting the hand and feet at the origin and destination of the lift from videos using computer vision.

This research demonstrates that it is possible to classify lifting postures as squat, stoop, or stand based on the bounding box dimensions of a subject with a significant degree of accuracy, which has several potential applications in assessing risks of manual lifting tasks. The findings from this analysis can be used to guide development of future computer vision algorithms that obtain data from continuous video recordings of actual occupational lifting tasks, with the overall goal of improving recommendations for occupational lifting safety.

We have developed computer vision software based on video background subtraction that continuously tracks the subject's body, applies the described bounding box, and identifies origins and destinations of lifts. This program can perform these analyses automatically with reasonable accuracy. We intend to incorporate the posture classification algorithm described in this paper into this computer vision program to automatically classify dynamic working postures for a series of consecutive frames in a video. In addition to standing, squatting, and stooping lift postures, the algorithm can easily identify other related states during work, including walking (the bounding box moves without changing dimensions), lifting overhead (the bounding box height grows larger than the stature), and standing still (the bounding box does not change in dimensions or move).

Separating the human body from other elements in a video image based on the different patterns of movement, our computer vision–based background subtraction algorithm can distinguish the human body regardless of variations in shape and size of the object being lifted. Additionally, since the computer vision algorithm draws the bounding box by detecting the silhouette of the subject and then drawing the box tightly around the silhouette, even if the object being handled blocks part of the body, the shape of the silhouette detected does not change, and therefore, the bounding box is not affected either. This feature is useful for providing reliability for applications in the field, where variations of objects being handled are common and may obstruct the body linkages.

The method discussed in the current paper offers a practical solution to identifying working postures. Research has associated additional important risk factors in lifting in addition to working postures. For example, Marras et al. (1993) reported association between trunk dynamics and risks of low-back disorders. Lavender, Andersson, Schipplein, and Fuentes (2003) found lifting speed as a significant factor for spine loading in L5/S1. Additionally, the weight of the object being handled also directly affects L5/S1 moment (Lavender et al., 2003) and is related to low-back disorders (e.g., Eriksen, Natvig, & Bruusgaard, 1999; Kerr et al., 2001). We expect to incorporate these additional factors with our computer vision–based approach to provide a comprehensive lifting analysis. The computer vision background subtraction algorithm and the posture classification method described in the current paper may be useful for continuously calculating the frequency of lifting tasks and the duration that each lifting posture is assumed during work, and this can greatly facilitate the analysis and evaluation of lifting tasks. We are using the same bounding box methodology for estimating parameters utilized in calculating the NIOSH lifting equation for evaluating the risks involved with lifting tasks, including horizontal, vertical, and distance measures. Together, these continuous evaluations may prove useful for offering additional information for continuously evaluating the risks associated with lifting rather than just the NIOSH equation alone.

We anticipate that given the simplicity of the resulting algorithm, it could be executed with minimal processing time, requiring low processor power. With the advantages of tolerance of noise and low computing requirements, we anticipate that such methods can be implemented on a handheld device such as a smartphone, making it readily accessible to ergonomic practitioners in the industry.

ACKNOWLEDGMENTS

This study was funded, in part, by a grant from the National Institute for Occupational Safety and Health (NIOSH/CDC), R01OH011024 (Radwin). The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of NIOSH/CDC.

KEY POINTS

- A practical approach suitable for computer vision applications was devised using the features of a simple “elastic” rectangular bounding box for classifying three types of lifting postures (standing, stooping, and squatting) without needing to measure joint angles or fit the data to skeletal models.
- A bounding box was drawn around images of a worker lifting, and the box height (H) and width (W) were used as a training set to develop a decision tree algorithm to identify the lifting postures of a subject based only on bounding box dimensions.
- The algorithm had a relatively high rate of success for correctly classifying lifting postures.
- The simplicity of the resulting method should make it insensitive to challenging workplace conditions, such as poor illumination and poor vantage points, while requiring relatively low processing power for fast execution.

ORCID ID

Robert G. Radwin https://orcid.org/0000-0002-7973-0641

REFERENCES

Akkas, O., Lee, C. H., Hu, Y. H., Yen, T. Y., & Radwin, R. G. (2016). Measuring elemental time and duty cycle using automated video processing. *Ergonomics*, *59*, 1514–1525.

American Conference of Governmental Industrial Hygienists. (2017). *TLV®/BEI® introduction*. Retrieved from http://www.acgih.org/tlv-bei-guidelines/tlv-bei-introduction

Anderson, C. K., Chaffin, D. B., Herrin, G. D., & Matthew, L. S. (1985). A biomechanical model of the lumbosacral joint during lifting activities. *Journal of Biomechanics*, *18*(8), 571–584.

Bao, S., Howard, N., Spielholz, P., & Silverstein, B. (2006). Quantifying repetitive hand activity for epidemiological research on musculoskeletal disorders–Part II: Comparison of different methods of measuring force level and repetitiveness. *Ergonomics*, *49*(4), 381–392.

Barondess, J. A., Cullen, M. R., de Lateur, B., Deyo, R., Donaldson, K., & Drury, C. (2001). *Musculoskeletal disorders and the workplace: Low back and upper extremities*. Washington, DC: National Academy of Sciences.

Bernard, B. P., Putz-Anderson, V., Burt, S. E., Cole, L. L., Fairfield-Estill, C., Fine, L. J., & …Tanaka, S. (1997). *Musculoskeletal disorders and workplace factors; a critical review of epidemiologic evidence for work-related musculoskeletal disorders of the neck, upper extremity, and low back*. Retrieved from https://www.cdc.gov/niosh/docs/97-141/pdfs/97-141.pdf?id=10.26616/NIOSHPUB97141

Bhattacharya, A., Schulte, P., & Anderson, V. (2012). *Workers' compensation costs in wholesale and retail trade sectors*. Retrieved from https://www.cdc.gov/NIOSH/docs/2013-147/pdfs/2013%E2%80%93147.pdf

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. Washington, DC: CRC Press.

Burgess-Limerick, R., & Abernathy, B. (1997). Toward a quantitative definition of manual lifting postures. *Human Factors*, *39*, 141–148.

Chaffin, D. B. (2008). Digital human modeling for workspace design. *Reviews of Human Factors and Ergonomics*, *4*(1), 41–74.

Chen, C. H., Hu, Y. H., Yen, T. Y., & Radwin, R. G. (2013). Automated video exposure assessment of repetitive hand activity level for a load transfer task. *Human Factors*, *55*, 298–308.

da Costa, B. R., & Vieira, E. R. (2010). Risk factors for work-related musculoskeletal disorders: A systematic review of recent longitudinal studies. *American Journal of Industrial Medicine*, *53*(3), 285–323.

Dempsey, P. G., McGorry, R. W., & Maynard, W. S. (2005). A survey of tools and methods used by certified professional ergonomists. *Applied Ergonomics*, *36*(4), 489–503.

Dysart, M. J., & Woldstad, J. C. (1996). Posture prediction for static sagittal-plane lifting. *Journal of Biomechanics*, *29*(10), 1393–1397.

Eriksen, W., Natvig, B., & Bruusgaard, D. (1999). Smoking, heavy physical work and low back pain: A four-year prospective study. *Occupational Medicine*, *49*(3), 155–160.

Gordon, C. C., Blackwell, C. L., Bradtmiller, B., Parham, J. L., Barrientos, P., Paquette, S. P., & …Kristensen, S. (2014). *2012 anthropometric survey of U.S. Army personnel: Methods and summary statistics*. Natick, MA: U.S. Army Natick Soldier Research, Development and Engineering Center.

Greene, R. L., Azari, D. P., Hu, Y. H., & Radwin, R. G. (2017). Visualizing stressful aspects of repetitive motion tasks and opportunities for ergonomic improvements using computer vision. *Applied Ergonomics*, *65*, 461–472.

Hoffman, S. G., Reed, M. P., & Chaffin, D. B. (2007). *Predicting force-exertion postures from task variables* (No. 2007-01-2480). Warrendale, PA: SAE.

Holmström, E. B., Lindell, J., & Moritz, U. (1992). Low back and neck/shoulder pain in construction workers: Occupational workload and psychosocial risk factors. Part 2: Relationship to neck and shoulder pain. *Spine*, *17*(6), 672–677.

Hwang, S., Kim, Y., & Kim, Y. (2009). Lower extremity joint kinetics and lumbar curvature during squat and stoop lifting. *BMC Musculoskeletal Disorders*, *10*, 15.

Kerr, M. S., Frank, J. W., Shannon, H. S., Norman, R. W., Wells, R. P., Neumann, W. P., . . . Ontario Universities Back Pain Study Group. (2001). Biomechanical and psychosocial risk factors for low back pain at work. *American Journal of Public Health*, *91*(7), 1069.

Lavender, S. A., Andersson, G. B., Schipplein, O. D., & Fuentes, H. J. (2003). The effects of initial lifting height, load magnitude, and lifting speed on the peak dynamic L5/S1 moments. *International Journal of Industrial Ergonomics*, *31*(1), 51–59.

Li, C., & Lee, S. (2011). Computer vision techniques for worker motion analysis to reduce musculoskeletal disorders in construction. In *Computing in civil engineering* (pp. 380–387). Reston, VA: ASCE.

Lu, M., Waters, T., Krieg, E., & Werren, D. (2014). Efficacy of the revised NIOSH lifting equation for predicting risk of low back pain associated with manual lifting: A one-year prospective study. *Human Factors*, *56*, 73–85.

Marfia, G., & Roccetti, M. (2016). A practical computer based vision system for posture and movement sensing in occupational medicine. *Multimedia Tools and Applications*, *76*(6), 8109–8129.

Marras, W. S., Lavender, S. A., Leurgans, S. E., Fathallah, F. A., Ferguson, S. A., Gary Allread, W., & Rajulu, S. L. (1995). Biomechanical risk factors for occupationally related low back disorders. *Ergonomics*, *38*(2), 377–410.

Marras, W. S., Lavender, S. A., Leurgans, S. E., Rajulu, S. L., Allread, S. W. G., Fathallah, F. A., & Ferguson, S. A. (1993). The role of dynamic three-dimensional trunk motion in occupationally-related low back disorders. *Spine*, *18*(5), 617–628.

Mathworks. (2017). *Decision trees*. Retrieved from https://www.mathworks.com/help/stats/classification-trees-and-regression-trees.html

Mehrizi, R., Peng, X., Tang, Z., Xu, X., Metaxas, D., & Li, K. (2018). *Toward marker-free 3D pose estimation in lifting: A deep multi-view solution*. Presented at 13th IEEE International Conference on Automatic Face & Gesture Recognition.

Mehrizi, R., Peng, X., Xu, X., Zhan, S., Metaxas, D., & Li, K. (2018). A computer vision based method for 3D posture estimation of symmetrical lifting. *Journal of Biomechanics*, *69*, 40–46.

Mehrizi, R., Xu, X., Zhang, S., Pavlovic, V., Metaxas, D., & Li, K. (2017). Using a marker-less method for estimating L5/S1 moments during symmetrical lifting. *Applied Ergonomics*, *65*, 541–550.

Myers, A. H., Baker, S. P., Li, G., Smith, G. S., Wiker, S., Liang, K. Y., & Johnson, J. V. (1999). Back injury in municipal workers: A case-control study. *American Journal of Public Health*, *89*(7), 1036–1041.

National Institute for Organizational Safety & Health. (2014). *Number, incidence rate, and median days away from work for nonfatal occupational injuries and illnesses involving days away from work for musculoskeletal disorders by part of body and ownership, Wisconsin, 2014*. Retrieved from https://www.bls.gov/iif/oshwc/osh/case/wi2014_pob.pdf

Plantard, P., Shum, H., Le Pierres, A. S., & Multon, F. (2016). Validation of an ergonomic assessment method using Kinect data in real workplace conditions. *Applied Ergonomics*, *65*, 562–569.

Safetyvideopreviews. (2012). *Manual material handling/safe lifting*. Retrieved from https://www.youtube.com/watch?v=rrI2n8qehrY&t=76s

Seo, J., Yin, K., & Lee, S. (2016). Automated postural ergonomic assessment using a computer vision-based posture classification. In *Construction Research Congress 2016* (pp. 809–818). Reston, VA: ASCE.

Spector, J. T., Lieblich, M., Bao, S., McQuade, K., & Hughes, M. (2014). Automation of workplace lifting hazard assessment for musculoskeletal injury prevention. *Annals of Occupational and Environmental Medicine*, *26*, 15.

University of Michigan Center for Ergonomics. (2014). *Paper flopping—Job modification*. Retrieved from https://www.youtube.com/watch?v=61cu5qvH0kM&index=54&list=PLn5IJRj74S88rnFFV6ObxS6nFdDXUFiGW

University of Michigan Center for Ergonomics. (2017a). *3DSSPP software*. Retrieved from https://c4e.engin.umich.edu/tools-services/3dsspp-software/.2017a

University of Michigan Center for Ergonomics. (2017b). *Stacking, facing line2 CE*. Retrieved from https://www.youtube.com/watch?v=MxTgvuhVAJA&t=55s

U.S. Bureau of Labor Statistics. (2016). *2015 nonfatal occupational injuries and illnesses: Cases with days away from work*. Retrieved from https://www.bls.gov/news.release/osh2.nr0.htm

U.S. Bureau of Labor Statistics. (2018). *Head, back, and hand injuries resulting in days away from work or job transfer or restriction*. Retrieved from https://www.bls.gov/opub/ted/2018/head-back-and-hand-injuries-resulting-in-days-away-from-work-or-job-transfer-or-restriction.htm

Wang, L., & Suter, D. (2007). Recognizing human activities from silhouettes: Motion subspace and factorial discriminative graphical model. In *Computer vision and pattern recognition* (pp. 1–8). Piscataway, NJ: IEEE.

Yen, T. Y., & Radwin, R. G. (1995). A video-based system for acquiring biomechanical data synchronized with arbitrary events and activities. *IEEE Transactions on Biomedical Engineering*, *42*(9), 944–948.

Runyu L. Greene is a PhD student in industrial and systems engineering at the University of Wisconsin-Madison. She has BS and MS degrees in industrial and systems engineering from the University of Wisconsin-Madison.

Yu Hen Hu is a professor in electrical and computer engineering at the University of Wisconsin-Madison. He has a BS degree from the National Taiwan University and a PhD from the University of Southern California.

Nicholas Difranco is a human factors engineer at Medtronic (Framingham, MA). He has BS and MS degrees in biomedical engineering from the University of Wisconsin-Madison.

Xuan Wang is a PhD student in electrical and computer engineering at the University of Wisconsin-Madison. She has a MS degree in electrical and computer engineering from the University of Wisconsin-Madison.

Ming-Lun Lu is a certified professional ergonomist and the manager for the Musculoskeletal Health Cross-Sector Research Program at the National Institute for Occupational Safety and Health. He holds a PhD in industrial hygiene with a concentration on ergonomics from the University of Cincinnati (OH).

Stephen Bao is a senior epidemiologist with Washington State Department of Labor and Industries. He has BS and MS degrees from Shanghai Jiaotong University, China and a licentiate and PhD degree from Luleå University of Technology (Sweden).

Jia-Hua Lin is a research ergonomist with the SHARP (Safety and Health Assessment of Research

for Prevention) program within the Washington State Department of Labor and Industries. He graduated from University of Wisconsin-Madison with MS and PhD degrees in industrial engineering.

Robert G. Radwin is Duane H. and Dorothy M. Bluemke Professor in the College of Engineering at the University of Wisconsin-Madison, where he advances new methods for research and practice of human factors engineering and occupational ergonomics. He has a BS degree from New York University Polytechnic School of Engineering and MS and PhD degrees from the University of Michigan.

Date received: March 29, 2018
Date accepted: June 30, 2018