Technical Correspondence

Load Asymmetry Angle Estimation Using Multiple-View Videos

Xuan Wang, Yu Hen Hu, *Fellow, IEEE*, Ming-Lun Lu, and Robert G. Radwin, *Senior Member, IEEE*

***Abstract*—A robust computer vision-based approach is developed to estimate the load asymmetry angle defined in the revised NIOSH lifting equation. The angle of asymmetry enables the computation of a recommended weight limit for repetitive lifting operations in a workplace to prevent lower back injuries. An open-source package OpenPose is applied to estimate the two-dimensional (2-D) locations of skeletal joints of the worker from two synchronous videos. Combining these joint location estimates, a computer vision correspondence and depth estimation method is developed to estimate the 3-D coordinates of skeletal joints during lifting. The angle of asymmetry is then deduced from a subset of these 3-D positions. Error analysis reveals unreliable angle estimates due to occlusions of upper limbs. A robust angle estimation method that mitigates this challenge is developed. We propose a method to flag unreliable angle estimates based on the average confidence level of 2-D joint estimates provided by OpenPose. An optimal threshold is derived that balances the percentage variance reduction of the estimation error and the percentage of angle estimates flagged. Tested with 360 lifting instances in a NIOSH-provided dataset, the standard deviation of angle estimation error is reduced from 10.13° to 4.99°. To realize this error variance reduction, 34% of estimated angles are flagged and require further validation.**

***Index Terms*—Asymmetry angle, manual lifting, NIOSH lifting equation, three-dimensional (3-D) skeletal joints estimation, video monitoring.**

I. Introduction

Overexertion during manual lifting is a leading cause of lower back pain and related health issues that cost the industry billions of dollars annually [1]. To protect workers from overexertion, the National Institute for Occupational Safety and Health (NIOSH) developed the revised NIOSH lifting equation (RNLE) [2] It has become the most widely used tool to assess the risk of lower back pain associated with lifting and lowering tasks in the workplace [3].

The RNLE computes the recommended weight lifted (RWL) and the lifting index as the ratio of the load weight and the RWL [2]. The RWL is derived from measurements dependent on the worker's body postures and movements during lifting. Direct-reading methods of these spatial and temporal factors include wearable motion sensors such as inertia motion units and visible markers (see [4]–[7]). However, these methods are intrusive to the normal work routines, often difficult to synchronize signals in the context of work activities, and observation duration is limited by battery life [8].

Video monitoring is a nonintrusive approach to acquire the measurements for the RNLE [9]–[12]. Earlier methods are based on manual scoring from the videos [9]; or fitting the video data into a biomechanical human body model [10], [11], [12]. Mehrizi *et al.* [10] reported computer vision algorithms to estimate three-dimensional (3-D) posture for symmetrical lifting tasks. In [11], a deep neural network is developed to predict lower back joint load and the risk estimate for low back disorders. In these works, cameras are carefully calibrated, and models are trained using acquired experiment data to develop the classifier [10] or the in-house 3-D joint location estimation algorithm [11], [12].

In this study, we present a robust computer vision workflow to estimate the asymmetry angle of asymmetric manual lifting.

Specifically, the proposed system is tasked to process two video clips taken synchronously by two video cameras of a manual lifting operation. For each video, it detects a keyframe when the lifting operation starts. Then, it applies an open-source human pose estimation software package OpenPose [14] to extract 2-D coordinates of skeletal joints of the subject performing the lifting operation. Next, a procedure called the structure from motion (SfM) is applied to estimate the relative camera pose. Manual marking of a set of corresponding corner feature points is used to improve accuracy. After estimation of the camera poses, 3-D coordinates of skeletal joints are estimated using triangulation. Finally, a robust estimation formula is proposed to estimate the angle of asymmetry using the estimated 3-D coordinates of skeletal joints. We also leveraged the confidence score of estimated 2-D coordinates of skeletal joints provided by OpenPose to predict the angle estimation error. We tested this workflow on a dataset acquired by NIOSH for a different purpose [13]. This dataset consists of 360 lifting operations. The mean value of angle estimation error was $-0.48°$ with a standard deviation of 10.14°. If one excludes those angle estimates with a confidence score below 0.5, the mean angle estimation error became $-1.12°$ with a standard deviation of 4.99°.

This approach exhibits several unique features: first, it leverages an open-source software package to estimate 2-D image coordinates of body skeletal joints for each camera. No retraining using the experiment data in this work was performed. second, cameras poses in the experiment are not available. A structure for motion procedure and manually selected matching feature points are used to estimate camera poses. Finally, a robust angle estimation procedure using the estimated 3-D coordinates of skeletal joints was proposed. The empirical relation between angle estimation error and the confidence score of 2-D skeletal joint estimates was leveraged to predict unreliable angle estimates. By rejecting these outliers, the overall angle estimation accuracy is significantly improved.

An important technical innovation of this work is the development of an end-to-end workflow that incorporates generic computer vision software modules such as the OpenPose to provide a robust estimation of the angle of asymmetry. Existing approaches [10]–[12] developed

Manuscript received May 15, 2020; revised November 15, 2020 and May 30, 2021; accepted August 10, 2021. Date of publication October 7, 2021; date of current version November 12, 2021. This work was supported by the NIOSH research Contract 5D30118P01249. This article was recommended by Associate Editor K. Li. *(Corresponding author: Yu Hen Hu.)*

Xuan Wang and Yu Hen Hu are with the Department of Electrical and Computer Engineering, University of Wisconsin–Madison, Madison, WI 53706 USA (e-mail: xwang554@wisc.edu; yhhu@wisc.edu).

Ming-Lun Lu is with the Taft Laboratories, National Institute for Occupational Safety and Health, Division of Applied Research and Technology, Cincinnati, OH 45226 USA (e-mail: uzl5@cdc.gov).

Robert G. Radwin is with the Industrial and Systems Engineering, University of Wisconsin–Madison, Madison, WI 53706 USA (e-mail: rradwin@wisc.edu).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/THMS.2021.3112962.

Digital Object Identifier 10.1109/THMS.2021.3112962

customized computer vision algorithms and software that are hard to adapt to varying practical scenarios. In this work, we demonstrated the promises of applying generic pose estimation modules for monitoring manual lifting operations. As numerous pose estimation algorithms being developed in the computer vision community, our approach can leverage state-of-the-art pose estimation modules to deliver robust and accurate estimates of 2-D and 3-D coordinates of skeletal joints. As such, ergonomic researchers may focus on inferring lifting equation parameters from these estimated poses for real-time monitoring. This is different from traditional approaches that focus on developing customized pose-estimation packages using data gathered from laboratory experiments. Note that the generic pose estimation modules may be retrained using ground truth labels to further enhance their performance for the specific dataset. This retraining step will *customize* a generic pose estimation package for a specific application scenario. Different retraining for different datasets will need to be performed when the system is to be applied to a different work environment.

A potential challenge of applying a generic pose estimation module is the need to assess potential estimation error. In this work, we analyzed causes that contribute to excessive angle estimation error and leverage a confidence score assessed by the generic pose estimation package to infer the angle estimation error. Our efforts showed that an exploratory investigation of estimation error is an integral part of the proposed workflow to enhance the reliability of the outcome.

The rest of this article, the definition of load asymmetry angle and the computer vision-based pose estimation method are discussed in Section II. The dataset used for the estimation method is described in Section III. Results and discussion are reported together in Section V, followed by the conclusion and future works in Section VI.

II. Backgrounds and Related Works

A. Load Asymmetry Angle

In the RNLS, the RWL is expressed as the product of a load constant (nominal weights, about 23 kg) and six multipliers, including horizontal, vertical, distance, angle, frequency, and coupling multipliers [15]. Among them, the angle multiplier AM is defined as follows:

$$AM = 1 - -0.0032A \quad 0 \leq A \leq 135. \tag{1}$$

where A is the load asymmetry angle in a unit of degrees. As shown in Fig. 1, the angle of asymmetry is defined as the angle between the asymmetry line and the mid-sagittal line [15].

From (1), an increment of 15° in the value of A reduces RWL by 1.1 kg. Although the required accuracy of A is not specified in [15], we may use (1) to assess whether the angle estimation error is acceptable in practice.

From Fig. 1, the asymmetry angle may be estimated given the 3-D coordinates of the worker's hands and ankles. In [16], body asymmetry is characterized using five angles during a lifting task. These angles are defined on horizontal vectors along the directions of the grips, shoulders, pelvis, and feet. Hence, measurements of 3-D coordinates of these body skeletal joints are required. Dempsey and Fathallah [17] discussed the distinction between load asymmetry and trunk asymmetry.

B. Video-Based Posture Estimation

Previously, we developed a video-processing algorithm [18] to estimate RNLE parameters including the lifting frequency and the hand location for lifting the object. In this algorithm, we use motion segmentation to track the worker's movement. When a lifting operation begins, the object being lifted will change from stationary to moving, leaving a ghost shadow of the object after subtracting the stationary background. By detecting the presence of this ghost shadow (ghosting effect), the frame index when the lifting starts can be determined. This method will be employed in this work to detect the onset of a lifting instance.

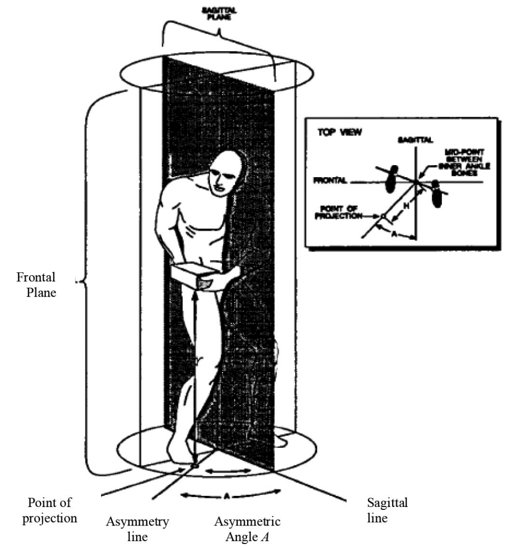


Fig. 1. Graphic representation of asymmetry angle [15].

C. Deep Neural Network Trained Human Pose Estimation

Several deep neural networks trained human pose estimation algorithms have been developed recently and made available as open-source software packages [13], [19]. Among them, OpenPose [14] has been known to give accurate results and is widely adopted in activity recognition, and gait analysis applications. OpenPose uses part affinity fields as a nonparametric representation of body parts and delivers 2-D coordinates of 25 key points including major skeletal joints. We use OpenPose in this work because it can process the videos of our dataset without any customization.

III. Method

A. 2-D Skeletal Joint Coordinate Estimation

The dataset used in this work consists of two synchronous video clips taken from two video cameras from the opposing sides of a subject performing a lift operation. In each trial, the subject walks toward a shelf, pickups the object from the shelf, turns around, and walks to the destination to drop off the object. One camera is stationary capturing the entire cycle of the lifting operation. The other camera was panned horizontally by an operator to track the movement of the subject. No camera intrinsic parameters (e.g., focal lengths) and extrinsic parameters (e.g., poses and positions) are available.

Using a lifting instance detection algorithm [18] developed in our lab, we detect the key video frames from each video clip when the subject moves the object and ready to turn, yet still having both feet on the floor. For each key video frame, OpenPose will estimate the 2-D image coordinates (in # of pixels) of 25 skeletal joints of the subject.

B. Estimating 3-D Coordinates of Skeletal Joints

Given the 2-D skeletal joint coordinates estimated using OpenPose from the respective key video frames, our next goal is to estimate the

3-D coordinates of these skeletal joints. Since the camera poses are not available, our approach is to first calibrate the cameras using a SfM (from two views) procedure [20], [21]. Then, we estimate the 3-D coordinates of skeletal joints from the corresponding 2-D coordinates of both views using triangulation [20].

SfM is a computer vision technique that simultaneously estimates camera poses and the 3-D coordinates of a set of corresponding feature points extracted from two or more views (cameras) of the same scene. It consists of the following steps.

a) Extract a set of visually distinct feature points from both keyframes.
b) Establish correspondence matching of feature points between the keyframes.
c) Using the epipolar constraint to estimate the fundamental matrix **F**.
d) Estimate (relative) camera pose (**R**, **t**) from the matrix **F**, where **R** is a 3-D rotation matrix and **t** is the linear translation of camera centers.
e) Estimate 3-D coordinates of the matching feature points.

In this work, in step a), we applied the SURF feature detector to estimate the set of feature points. In step b), we manually selected a set of matching corner feature points corresponding to static objects visible from both keyframes. Steps c) and d) are realized using MATLAB computer vision toolbox functions *estimateFundamentalMatrix* and *cameraPose*. The translation vector **t** is a univector whose magnitude cannot be estimated using only two views. We did not perform step e) because manually picked feature points in step b) are not the desired skeletal joints. Instead, we apply triangulation (MATLAB command *triangulate*) to the 2-D coordinates of skeletal joints using the estimated camera pose and deduce corresponding 3-D coordinates relative to the reference camera. These 3-D coordinates are subject to the same yet unknown scaling factor since we only have two views. In Section III.C, we shall explain that this unknown scaling factor does not affect the estimation of the angle of asymmetry as it depends on the relative orientation between pairs of skeletal joints rather than their absolution 3-D coordinates.

The decision of applying the OpenPose package to elicit 2-D coordinates of skeletal joints from each view without retraining directly impacts the method described earlier. Specifically, the accuracy of these 2-D coordinates of skeletal joints may not be sufficiently accurate to be used in SfM [step b)] for estimating the camera poses.

C. Robust Load Asymmetry Angle Estimation

Given the 3-D coordinates of wrists and hip joints, we may proceed to estimate the asymmetry line defined in Fig. 1. The projection on the floor of a unit vector along the direction between left and right wrists is perpendicular to the symmetry line. Similarly, the projection on the floor of a unit vector between the two hip joints is perpendicular to the sagittal line. The angle difference between these two-unit vectors thus is the load asymmetry angle [15].

Let P_{LW}, P_{RW}, P_{LH}, and P_{RH} be the horizontal (*x* and *y*) components of the estimated 3-D coordinates of the left-, right-side wrists and the left-, right-hip joints of the worker during lifting. One may compute a *wrist* direction vector as follows:

$$\mathbf{v} = \begin{bmatrix} v_1 & v_2 \end{bmatrix}^T = P_{RW} - P_{LW}. \tag{2}$$

Then, the wrist angle $\theta_W = \tan^{-1}(v_2/v_1)$. Similarly, one may compute the hip direction angle θ_H. Finally, the asymmetry angle A may be estimated as follows:

$$A = \theta_W - \theta_H. \tag{3}$$

Note that the **v** vector is dependent only on the *relative* horizontal world coordinates (*x* and *y*) of the related skeletal joints. Therefore, there is no need to perform training to align the 3-D coordinates estimated in this section with those provided by the ground truth motion capture (MoCap) device. Moreover, it is invariant to scaling the 3-D coordinates of the skeletal joints. Hence, the coordinates estimated in Section III.B may be used without modification.

Recall that $d\tan\theta/d\theta = 1 + \tan^2\theta$, we have (for θ_W or θ_H)

$$d\theta = \frac{d\tan\theta}{1+\tan^2\theta} = \frac{(v_1 dv_2 - v_2 dv_1)/v_1^2}{1+(v_2/v_1)^2} = \begin{bmatrix} \cos\theta & -\sin\theta \end{bmatrix} \cdot \frac{d\mathbf{v}}{|\mathbf{v}|}$$

Hence
$$\|d\theta\| \le \|d\mathbf{v}\|/\|\mathbf{v}\|. \tag{4}$$

Equation (4) says the angle estimation error will be smaller than that of the relative skeletal joint coordinate estimation error. When the two estimated skeletal joints are wide apart (larger value of $\|\mathbf{v}\|$), the relative joint estimation error will be smaller. Since (4) applies to both θ_W and θ_H in (3), the estimation of asymmetry angle A is robust against the coordinate estimation errors of the skeletal joints $d\mathbf{v}$.

IV. Laboratory Experiment and Dataset [13]

A. Experiment Setup

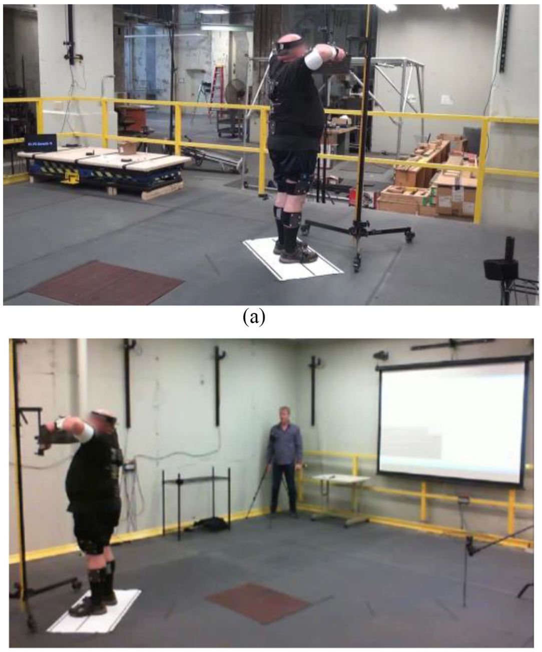
The experiment data were adapted from a study conducted by NIOSH [13] previously. In this study, human body postures in different symmetrical lifting tasks were recorded. A total of 12 different originating lifting hand locations are used according to the American Conference of Governmental Industrial Hygienists Threshold Limit Values for lifting [22].

For each of the 12 initial hand locations, a subject repeats a lifting task three times. Learning effects were mitigated by instructing the subjects to perform random ordered trials. The subject will walk from a starting point toward the lifting station, lift the object with both hands, and turn-around walk back to the drop-off station near the starting point. A wire basket was used as the lifting object. The wire texture would reduce (to some extent) the visual obstruction reduction of the hands holding the object. The basket was set on a 12 × 12 cm platform to facilitate the subjects to perform lifting tasks naturally. The height of the platform was adjusted according to the designated 12 initial locations during lifting.

A MoCap system OptiTrack (Model Flex 13, Innovative Sports, Inc., Chicago, USA) is used to track marker clusters attached to 13 positions on the body of the subject. This MoCap system claims an average accuracy of 0.7 mm in the 3-D coordinate system when calibrated.

Two cameras were used to record the video data. Both cameras were mounted on tripods and were synchronized with the MoCap system. One camera was a web camera (Microsoft 1080p LifeCam, 640 × 480 pixels, 30 fps). It is placed 4 meters away from the starting point, at an eye-level height, with a fixed viewing angle perpendicular to the subject's walking path. The other camera was a camcorder (Sony, 1280 × 720 pixels, 30 fps). The second camera is located across the walking path of the subject and is controlled by a staff [see Fig. 2(b)] to pan horizontally following the subject walking to and back from the lifting station. An example of corresponding views of the two cameras is shown in Fig. 2(a) and (b).

Since the two cameras have different resolutions, we scale video frames of the first camera (1280 × 720) so that the same subject has about the same height (in units of pixels) as that in the second camera. This *ad hoc* scaling operation helps improve the accuracy of skeletal joints' 3-D coordinate estimations. All subsequent processing steps (including individual camera calibration) were performed after the video frames of the first camera were scaled.



(a)

(b)

Fig. 2. (a) Sony camera view. (b) Life camera view [photo credit: CDC/NIOSH].

Each camera is initially calibrated by capturing a short clip of a calibration checkerboard. MATLAB camera calibration app is applied to obtain intrinsic parameters, including focal length, of each camera. The camera poses, however, are not available since the video clips of the checkerboard were taken before the experimentation at different camera poses.

B. Subjects

Ten subjects were recruited. These subjects were employees in the division of the Applied Research and Technology office of NIOSH in Cincinnati, OH, USA. Inclusion criteria and exclusion criteria were applied to screen the subjects. Written consents were obtained according to the NIOSH-approved IRB study protocol.

C. Experiment Protocol

The path the subject is directed to follow during each lifting trial is marked on the floor, with the initial position, the lifting location, and the finishing line identified. The subject is instructed to line up toes to each of these lines when performing such tasks. The subject will walk from the initial position toward the lifting station following the line and will lift the basket in the front with both hands. Then, the subject will turn around, carrying the basket to a shelf to release the object, and walk to the finishing line. The subjects will perform these steps at their own pace and turning at their preference. The distance between these specific locations is not more than 20 steps. An experiment will not be recorded until the subject became familiarized with the required steps. Each trial lasted about 15 s.

The frame number of the beginning of lifting (BOL) for the ground truth data was established manually by two researchers independently. BOL is defined as the instant when the basket started to move. The ground truth (MoCap) asymmetry angles then are estimated by feeding corresponding MoCap-annotated 3-D skeletal joint positions into (2) and (3).

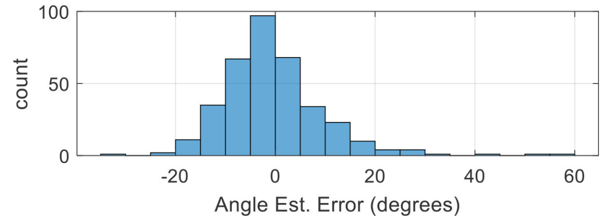
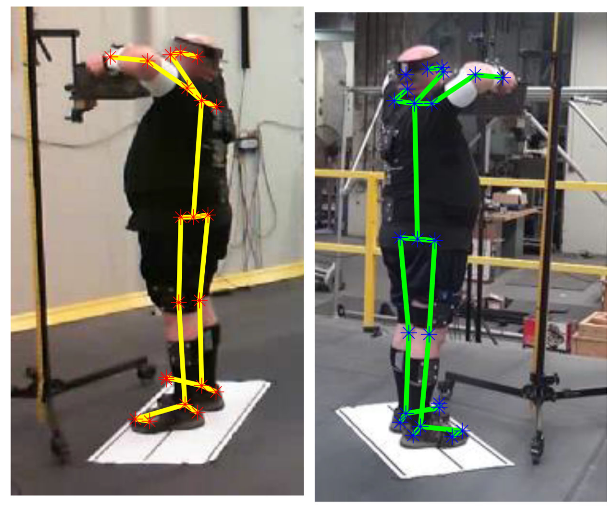


Fig. 3. Distribution of angle estimation errors (mean = 0.48°, Std = 10.14°).



(a)

(b)

Fig. 4. OpenPose [14] 2-D skeleton joints estimation for the same scene but captured from two views. (a) Left arm is visible. (b) Right arm is visible [photo credit: CDC/NIOSH].

V. Results and Error Analysis

We apply the estimated angle of asymmetry of the 360 lifting instances in the NIOSH dataset and compared it to the computed ground-truth values. The error distribution is shown in Fig. 3.

After reviewing the corresponding videos and the 2-D pose estimates, it is found that those large errors of estimated angles are often due to self-occlusion. Since the cameras are placed to face the sagittal plane of the tester, only one side of the body is exposed to a camera. As shown in Fig. 4, each camera may view only one side of the tester. When a body part is not found in the view, OpenPose reports a predicted 2-D position of the joint and enters a confidence score of 0. Using these predicted 2-D joint positions to estimate corresponding 3-D coordinates is likely to cause large errors of the estimated asymmetry angles.

Out of the 360 lifting instances in the dataset, the number of instances that some of the joints are occluded and not detected by OpenPose as listed in Table I. Note that the wrist joint is prone to occlusion. When this happens, we opted to use the elbow joints (of both hands) to estimate the vector $\mathbf{v}$ in (2). If an elbow is also occluded, the fallback choice is to use the shoulder joints.

We also hypothesize that the weighted averaged confidence scores of the 2-D joint position estimates may be correlated to the angle estimation error. In Fig. 5, a scatter plot of absolute values of angle estimation error versus the weighted averaged confidence score of all 2-D joint

TABLE I
PERCENTAGE INSTANCES THE JOINT IS NOT DETECTED BY OPENPOSE

	Wrist missing	Elbow missing
Right side	16.94%	4.72%
Left side	28.61%	15.56%

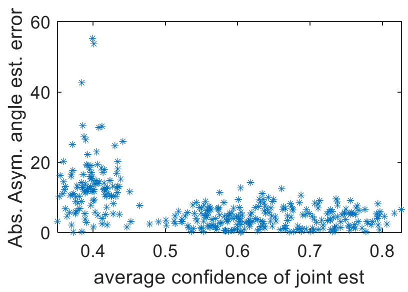


Fig. 5. Absolute values of angle estimation error as a function of averaged confidence scores of 2-D joint estimates.

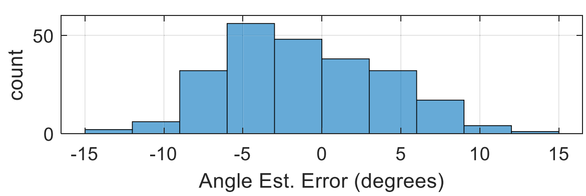


Fig. 6. Distribution of angle estimation errors after excluding unreliable estimates (mean = –1.12°, Std = 4.99°).

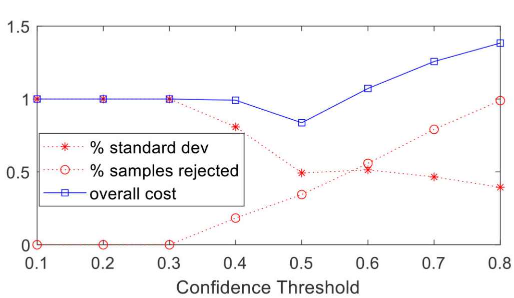


Fig. 7. Choosing optimal threshold value of averaged confidence score to minimize the overall cost.

location estimates is provided. Note that when the weighted averaged confidence score is below 0.5, the angle estimation error increases dramatically. Therefore, one may decide that if the weighted averaged confidence score is smaller than 0.5, the corresponding asymmetry angle estimate may be deemed unreliable, meaning potentially the error may be large. When this algorithm is deployed in the warehouse to monitor the lifting operations, unreliable angle estimates may be flagged to be reviewed manually. Note that the angle estimation errors due to self-occlusion are relatively less accurate. If these angle estimates deemed to be unreliable are excluded, the resulting error distribution is shown in Fig. 6. Note STD is reduced from 10.14° to 4.99°.

We search the range between 0.1 to 0.8 of the averaged confidence score and find 0.5 yields the optimal solution. This optimization process is summarized in Fig. 7.

VI. CONCLUSION

In this research, an algorithm to estimate the load asymmetry angle for the RNLE was developed. This algorithm does not require customized training using local datasets and provides an assessment of the estimated angle being reliable or not. We identify that self-occlusion is the main source of estimation errors. It may be mitigated with additional cameras properly placed to avoid blind spots of viewing. Future work will focus on general-purpose training robust learning-based 3-D pose estimation algorithms to provide accurate 3-D coordinates of body poses in a harsh work environment and integration of estimating other RNLE parameters in one estimation process.

ACKNOWLEDGMENT

The authors would like to thank the National Institute for Occupational Safety and Health (NIOSH) research team members Menekse Barim, Marie Hayden, and Dwight Werren for their assistance with the dataset used for this study. The findings and conclusions in this study are those of the authors and do not necessarily represent the official position of the NIOSH, the Centers for Disease Control and Prevention (CDC). Mention of any company or product does not constitute an endorsement by NIOSH, CDC.

REFERENCES

[1] Liberty Mutual Insurance, "Liberty mutual workplace safety index," 2018. [Online]. Available: https://business.libertymutualgroup.com/business-insurance/Documents/Services/Workplace%20Safety%20Index.pdf

[2] T. R. Waters, V. Putz-Anderson, A. Garg, and L. J. Fine, "Revised NIOSH equation for the design and evaluation of manual lifting tasks," *Ergonomics*, vol. 36, no. 7, pp. 749–776, 1993.

[3] B. Lowe, P. Dempsey, E. Jones, and National Institute for Occupational Safety and Health, "Assessment methods used by certified ergonomics professionals," in *Proc. Human Factors Ergonom. Soc. Annu. Meeting*, 2018, pp. 838–842.

[4] P. G. Dempsey, "Usability of the revised NIOSH lifting equation," *Ergonomics*, vol. 45, no. 12, pp. 817–828, 2002.

[5] G. Li and P. Buckle, "Current techniques for assessing physical exposure to work-related musculoskeletal risks, with emphasis on posture-based methods," *Ergonomics*, vol. 42, no. 5, pp. 674–695, 1999.

[6] B. Juul-Kristensen *et al.*, "Assessment of work postures and movements using a video-based observation method and direct technical measurements," *Appl. Ergonom.*, vol. 32, no. 5, pp. 517–524, 2001.

[7] S. Kim and M. A. Nussbaum, "Performance evaluation of a wearable inertial motion capture system for capturing physical exposures during manual material handling tasks," *Ergonomics*, vol. 56, no. 2, pp. 314–326, 2013.

[8] A. Patrizi, E. Pennestr, and P. P. Valentini, "Comparison between low-cost marker-less and high-end marker-based motion capture systems for the computer-aided assessment of working ergonomics," *Ergonomics*, vol. 59, no. 1, pp. 155–162, 2016.

[9] X. Xu, C.-C. Chang, G. S. Faber, I. Kingma, and J. T. Dennerlein, "The validity and interrater reliability of video-based posture observation during asymmetric lifting tasks," *Human Factors*, vol. 53, no. 4, pp. 371–382, 2011.

[10] R. Mehrizi, X. Peng, X. Xu, S. Zhang, D. Metaxas, and K. Li, "A computer vision-based method for 3D posture estimation of symmetrical lifting," *J. Biomech.*, vol. 69, pp. 40–46, 2018.

[11] R. Mehrizi, X. Peng, D. N. Metaxas, X. Xu, S. Zhang, and K. Li, "Predicting 3-D lower back joint load in lifting: A deep pose estimation approach," *IEEE Trans. Human–Mach. Syst.*, vol. 49, no. 1, pp. 85–94, Feb. 2019.

[12] R. Mehrizi, X. Peng, X. Xu, S. Zhang, and K. Li, "A deep neural network-based method for estimation of 3D lifting motions," *J. Biomech.*, vol. 84, pp. 87–93, 2019.

[13] M. S. Barim, M. L. Lu, S. Feng, G. Hughes, M. Hayden, and D. Werren, "Accuracy of an algorithm using motion data of five wearable IMU sensors for estimating lifting duration and lifting risk factors," in *Proc. Human Factors Ergonom. Soc. Annu. Meeting*, 2019, pp. 1105–1111.

[14] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, "OpenPose: Realtime multi-person 2D pose estimation using part affinity fields," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 1, pp. 172–186, 2019.

[15] T. R. Waters, V. Putz-Anderson, and A. Garg, *Applications Manual for the Revised NIOSH Lifting Equation*. DHHS (NIOSH) Publication Number 94-110, 1994.

[16] A. Delisle, M. Gagnon, and P. Desjardins, "Kinematic analysis of footstep strategies in asymmetrical lifting and lowering tasks," *Int. J. Ind. Ergonom.*, vol. 23, no. 5-6, pp. 451–460, 1999.

[17] P. G. Dempsey and F. A. Fathallah, "Application issues and theoretical concerns regarding the 1991 NIOSH equation asymmetry multiplier," *Int. J. Ind. Ergonom.*, vol. 23, no. 3, pp. 181–191, 1999.

[18] X. Wang, Y. H. Hu, M. L. Lu, and R. G. Radwin, "The accuracy of a 2D video-based lifting monitor," *Ergonomics*, vol. 62, no. 8, pp. 1043–1054, 2019.

[19] D. Pavllo, C. Feichtenhofer, D. Grangier, and M. Auli, "3D human pose estimation in video with temporal convolutions and semi-supervised training," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 7753–7762.

[20] R. Szeliski, *Computer Vision: Algorithms and Applications*. Berlin, Germany: Springer, 2010.

[21] N. Snavely, S. M. Seitz, and R. Szeliski, "Photo tourism: Exploring photo collections in 3D," in *Proc. ACM Siggraph*, vol. 25, 2006, pp. 835–846.

[22] T. J. Armstrong, "ACGIH TLV for hand activity level," in *Biomechanics in Ergonomics*. Boca Raton, FL, USA: CRC Press, 2007, pp. 377–390.