Journal of Biomechanics 69 (2018) 40–46

Contents lists available at ScienceDirect

Journal of Biomechanics

journal homepage: www.elsevier.com/locate/jbiomech
www.JBiomech.com

A computer vision based method for 3D posture estimation of symmetrical lifting

Rahil Mehrizi [a], Xi Peng [b], Xu Xu [e], Shaoting Zhang [f], Dimitris Metaxas [b], Kang Li [a,b,c,d,*]

[a] Department of Industrial & Systems Engineering, Rutgers University, Piscataway, New Jersey, USA
[b] Department of Computer Science, Rutgers University, Piscataway, New Jersey, USA
[c] Department of Biomedical Engineering, Rutgers University, Piscataway, New Jersey, USA
[d] Department of Orthopaedics, Rutgers New Jersey Medical School, Newark, New Jersey, USA
[e] Edward P. Fitts Department of Industrial and Systems Engineering, North Carolina State University, Raleigh, North Carolina, USA
[f] Department of Computer Science, University of North Carolina, Charlotte, North Carolina, USA

ARTICLE INFO

Article history:
Accepted 8 January 2018

Keywords:
Computer vision
Marker-less motion capture
Joint kinematics assessment
Lifting
Discriminative approach

ABSTRACT

Work-related musculoskeletal disorders (WMSD) are commonly observed among the workers involved in material handling tasks such as lifting. To improve work place safety, it is necessary to assess musculoskeletal and biomechanical risk exposures associated with these tasks. Such an assessment has been mainly conducted using surface marker-based methods, which is time consuming and tedious. During the past decade, computer vision based pose estimation techniques have gained an increasing interest and may be a viable alternative for surface marker-based human movement analysis. The aim of this study is to develop and validate a computer vision based marker-less motion capture method to assess 3D joint kinematics of lifting tasks. Twelve subjects performing three types of symmetrical lifting tasks were filmed from two views using optical cameras. The joints kinematics were calculated by the proposed computer vision based motion capture method as well as a surface marker-based motion capture method. The joint kinematics estimated from the computer vision based method were practically comparable to the joint kinematics obtained by the surface marker-based method. The mean and standard deviation of the difference between the joint angles estimated by the computer vision based method and these obtained by the surface marker-based method was 2.31 ± 4.00°. One potential application of the proposed computer vision based marker-less method is to noninvasively assess 3D joint kinematics of industrial tasks such as lifting.

1. Introduction

Lifting is one of the common manual material handling tasks performed in the workplaces. It is considered as one of the main risk factors for low back disorders (Kuiper et al., 1999; da Costa and Vieira, 2010; Nimbarte et al., 2010). In order to improve work place safety, it is necessary to analyze musculoskeletal and biomechanical risk exposures associated with lifting by measuring the joint kinematics and assessing the critical joint stress through biomechanical analysis.

The most common method for measuring joint kinematics is the surface-marker based motion capture method. This method uses reflective surface markers and optical motion capture systems to track the body movements: the surface markers are attached near the joints of the subject and the 3D positions of each joint are estimated using the 3D coordinates of these surface markers. This method is limited since it requires expensive optical motion capture equipment; and attaching surface markers to human body is time consuming and tedious.

Non-invasive computer vision-based marker-less motion capture systems have recently gained an increasing interest for pose estimation. A variety of computer vision algorithms have been proposed for 3D human pose estimation (Gavrila 1999; Mündermann et al., 2006; Poppe, 2007; Holte et al., 2012). These algorithms can be categorized into two types of approaches, generative (model-based) and discriminative (model-free) approaches. Generative approaches use an a priori model of the subject including body dimensions and kinematic trees; and perform a local search around an initial pose estimate. Disadvantages of this type of approaches are the need of initialization and high computational cost.

* Corresponding author at: Department of Industrial & Systems Engineering, Rutgers University, Piscataway, New Jersey, USA.
E-mail address: kl419@soe.rutgers.edu (K. Li).

https://doi.org/10.1016/j.jbiomech.2018.01.012
0021-9290/

Discriminative approaches use training data to predict the pose directly and are computationally efficient compared to the generative approaches.

Most of these computer vision algorithms for human pose estimation are designed for activity recognition; and their accuracy and robustness are not sufficient for 3D joint kinematics assessment and biomechanical analysis. There are only few studies proposing computer vision based marker-less methods for joint kinematics assessment. Drory et al. (2017) presented a discriminative computer vision based method for marker-less estimation of the full body kinematics for a cyclist. Their method is capable of estimating the pose in 2D space accurately, however; its performance is not tested for the 3D body pose estimation. Several studies (Mündermann et al., 2005; Saboune and Charpillet, 2005; Corazza et al., 2006; Ceseracciu et al., 2014; Sandau et al., 2014) proposed generative computer vision based methods for joint kinematics assessment during gait, which have been validated against the joint kinematics obtained by the surface marker-based method. These studies used eight cameras to capture images from different views around the subjects and converted the captured images data to 3D surface meshes by applying background subtraction and constructing visual hulls. Pose estimation was then performed by fitting the predefined 3D body models to the 3D surface meshes. In order to achieve the results comparable to the surface marker-based system, these studies required a substantial number of cameras. The data processing of points fitting was computationally expensive. Little research has been conducted for assessing 3D joint kinematics for work-related activities including lifting using computer vision-based marker-less methods although manual lifting is a common task in the workplace and associated injuries are substantial.

This paper proposes a computer vision-based marker-less motion capture method for estimating and assessing full-body kinematics and validates the method for symmetrical lifting. In this method, we integrate a discriminative approach for pose estimation with morphological constraints to achieve the accuracy and robustness sufficient for 3D joint kinematics assessment. We tested our method on three types of manual lifting with different starting and ending heights. The joint kinematics estimated from the computer vision based method were practically comparable to the joint kinematics obtained by the surface marker-based method. Therefore, the proposed computer vision method can be potentially considered as a simpler and faster substitute for the surface marker-based systems.

2. Methods

2.1. Data acquisition

The data set (Mehrizi et al., 2017) consists of 12 healthy males (age 47.50 ± 11.30 years; height 1.74 ± 0.07 m; weight 84.50 ± 12.70 kg) performing various symmetric lifting tasks in a laboratory at self-selected speed while being filmed by both camcorder and a synchronized motion tracking system that directly measured the body movement. They lifted a plastic crate ($39 \times 31 \times 22$ cm) weighing 10 kg and placed it on a shelf without moving their feet. They performed three vertical lifting ranges from floor to knuckle height (FK), knuckle to shoulder height (KS) and floor to shoulder height (FS).

45 Reflective markers (Cappozzo et al., 1995) were attached to the lifters' body segments and 3D positions of markers during the lifting tasks were measured by a motion tracking system (Motion Analysis, Santa Rosa, CA) with a sampling rate of 100 Hz. The raw 3D coordinate data were filtered with a fourth-order Butterworth low-pass filter at 8 Hz. Two digital camcorder (GR-850U, JVC, Japan) with 720×480 pixel, synchronized with the motion tracking system resolution also recorded the lifting from two views, 90° (side view) and 135° positions.

2.2. Computer vision method for joint kinematics assessment

In this study, we propose a computer vision based marker-less method using a discriminative approach for joint kinematics assessment. The workflow of the proposed method consists of three steps including feature extraction, 3D pose reconstruction, and angle calculation, which are summarized in Fig. 1. The input of this system is the video (image sequence) recorded by digital camcorders.

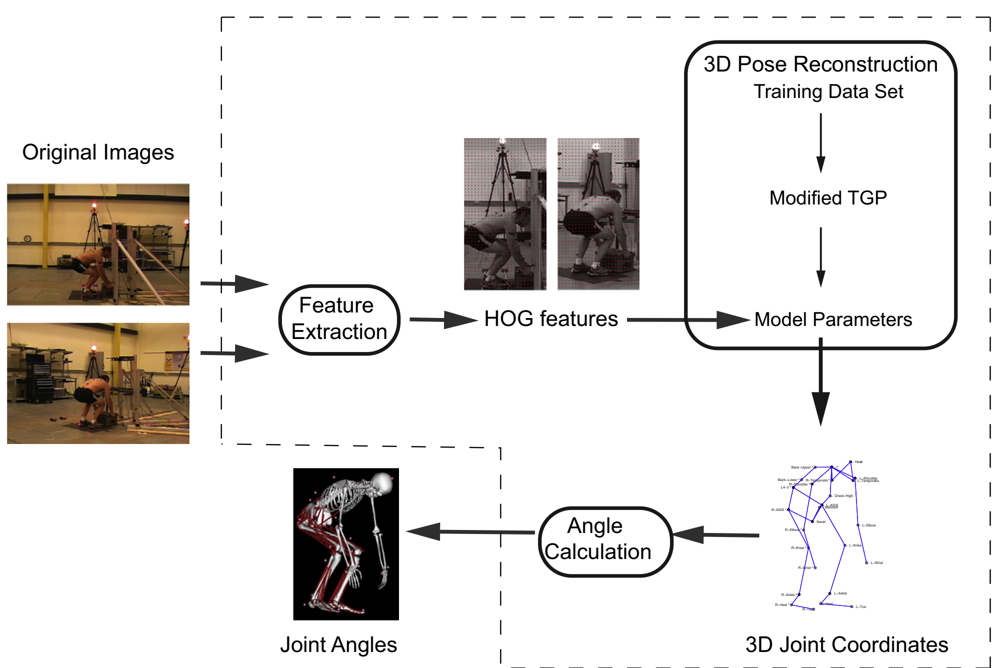

Fig. 1. Workflow of the proposed computer vision based marker-less method.

2.2.1. Feature extraction

In the first step, the recorded images are converted to image features. The descriptor of the image feature used in this study is Histogram Orientation Gradient (HoG) (Dalal and Triggs, 2005), which characterizes the distribution of gradient directions of images and has been widely used for object recognition and human detection (Suard et al., 2006; Zhu et al., 2006; Yang and Ramanan, 2011).

In order to extract the HoG features, the input image is divided into cells, i.e. small groups of connected pixels. A histogram of gradient orientation is calculated for pixels within the cell. The histograms are contrast-normalized by calculating a measure of intensity across a larger region of an image and then using this measure to normalize all the cells within this region. The descriptors of the HoG features are the concatenation of the histograms for the input image. Since we want to use captured images from both views in our data for pose estimation, the HoG features extracted for each view are concatenated to make a longer image descriptor. The HoG feature extraction is now implemented as a standard function in Matlab. It took around 0.05 s per image (480×720 pixels) to complete the HoG feature extraction for a PC with a 3.4 GHz i7 processor and 8 GB of RAM.

2.2.2. 3D Pose Reconstruction

A modified algorithm based on Twin Gaussian Process (TGP) (Bo and Sminchisescu, 2010) is proposed to extract the 3D pose from each frame of the videos. TGP (Bo and Sminchisescu, 2010) is a discriminative approach, which maps directly from image features to a human pose as follows. Assume that we have a set of inputs e.g. image feature observations ($R = (r_1, r_2, \ldots, r_N)$) and a set of the corresponding outputs e.g. 3D pose measures ($X = (x_1, x_2, \ldots, x_N)$). Since the Gaussian distribution of the inputs (N_R) is known, we can calculate the offset between the estimated Gaussian distribution of outputs (N_X), including the unknown pose for a testing image, and the homologous input distribution using Kullback-Leibler divergence function (D_{KL}). Then, for a testing image, we can estimate the corresponding 3D pose measure (x^*) by minimizing Kullback-Leibler divergence:

$$x^* = \arg\min(D_{KL}(N_X || N_R))$$

The TGP method can be applied on either a monocular image or multi view images for predicting whole-body 3D poses. Note that the 3D pose measures can be a set of estimated virtual markers defining the body joints. In this study, we used the 3D coordinates of 45 estimated 'virtual' markers as the 3D pose measures, which matched with the definitions proposed by (Cappozzo et al., 1995).

Since the errors of predicted 3D poses from the TGP method (Bo and Sminchisescu, 2010) could be extremely high and may not be good for joint kinematics assessment, we propose a modified TGP method by introducing the morphology constraints of the upper limb segment lengths to the original TGP algorithm, since lifting tasks involve significant upper limb motion. These morphology constraints enforce the consistency of upper limb segment lengths, i.e., the Euclidian distances between shoulder-elbow and elbow-wrist should be constant over all frames. With these constraints, the differences between predicted segment lengths and the constant length values are added as penalties to the Kullback-Leibler divergence cost function, which should be minimized. By adding these constraints to the algorithm, the search space is much smaller and thus more accurate pose estimation can be achieved.

The TGP algorithm requires a training dataset to train the model parameters used for calculating the Kullback-Leibler divergence. Bo and Sminchisescu (2010) have examined the effect on pose estimation results of two type of training, jointly training and separately training. Jointly training trained the model parameter on the motions of all subjects (the whole data set) while separately training trained the model on the motions of individual subject separately. They have reported no significant difference between the two types of training on pose estimation. Since jointly training uses a bigger training data set and is computationally more expensive, separately training was employed to train the model for each subject in this study. Accordingly, the k-fold cross validation technique was used to test our modified TGP method for each subject,

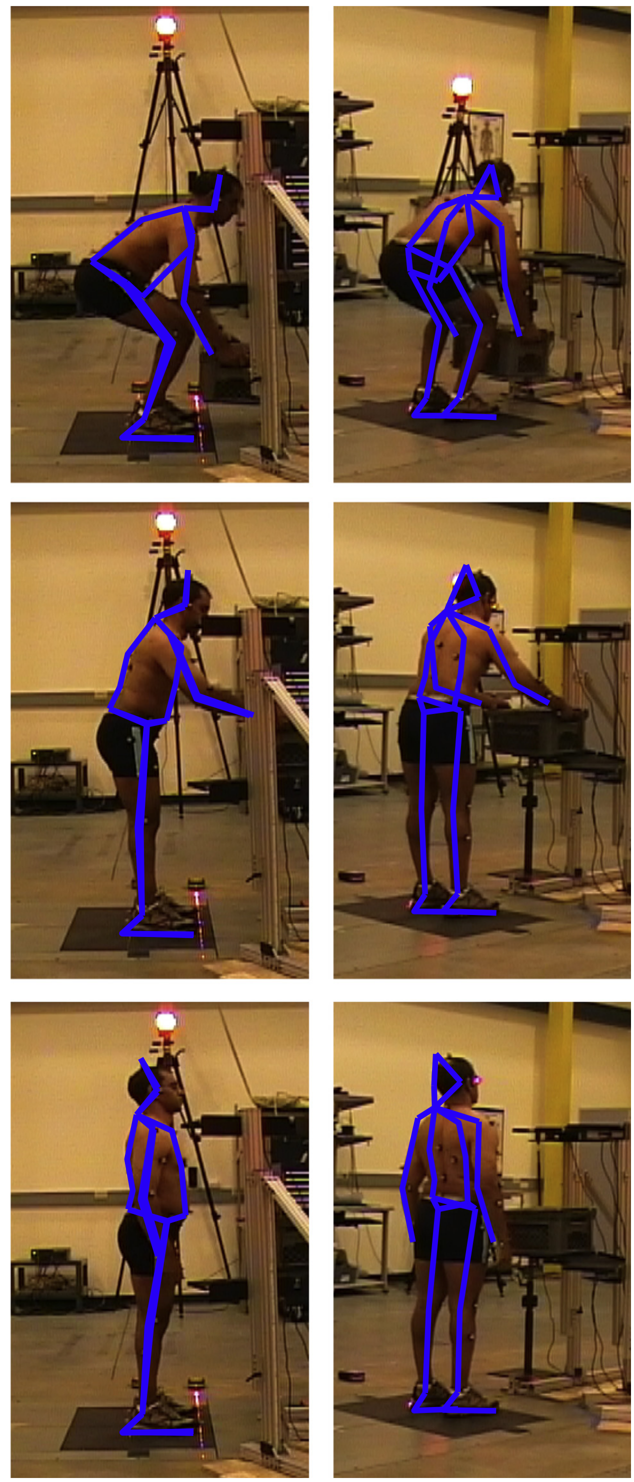

Fig. 2. Estimated full-body 3D posture from the proposed computer vision base marker-less method for selected frames of a representative subject performing a FK lifting for both views (90 and 135°). Activities from top to down: grabbing the box from the floor, putting the box on the shelf, in the standing position.

Table 1
Mean ± SD (mm) of Euclidean distance between the estimated 'virtual' marker positions obtained from our proposed method and the corresponding experimental surface marker positions.

Subject	FK	KS	FS	Average
1	14.57 ± 8.71	10.42 ± 9.67	8.72 ± 7.51	11.24 ± 9.47
2	9.63 ± 7.29	7.90 ± 6.79	11.83 ± 6.75	9.79 ± 7.63
3	7.79 ± 7.06	9.30 ± 7.00	7.89 ± 6.99	8.32 ± 7.58
4	7.64 ± 6.23	10.31 ± 7.96	6.96 ± 5.61	8.30 ± 7.19
5	7.45 ± 5.32	6.82 ± 5.91	6.41 ± 3.89	6.89 ± 5.73
6	6.77 ± 4.39	10.02 ± 4.56	6.84 ± 4.21	7.88 ± 4.94
7	7.09 ± 4.76	10.95 ± 6.77	8.30 ± 5.46	8.78 ± 6.27
8	8.34 ± 5.73	7.99 ± 6.36	9.22 ± 6.71	8.52 ± 7.02
9	12.66 ± 10.41	15.96 ± 15.81	12.32 ± 11.14	13.65 ± 14.74
10	7.08 ± 3.10	8.35 ± 6.69	6.86 ± 3.51	7.43 ± 5.56
11	5.10 ± 3.07	5.54 ± 3.73	5.29 ± 2.74	5.31 ± 3.52
12	6.81 ± 3.67	9.55 ± 5.69	6.83 ± 3.76	7.73 ± 4.88
Average	8.41 ± 7.35	9.42 ± 9.11	8.12 ± 7.16	9.52 ± 7.95

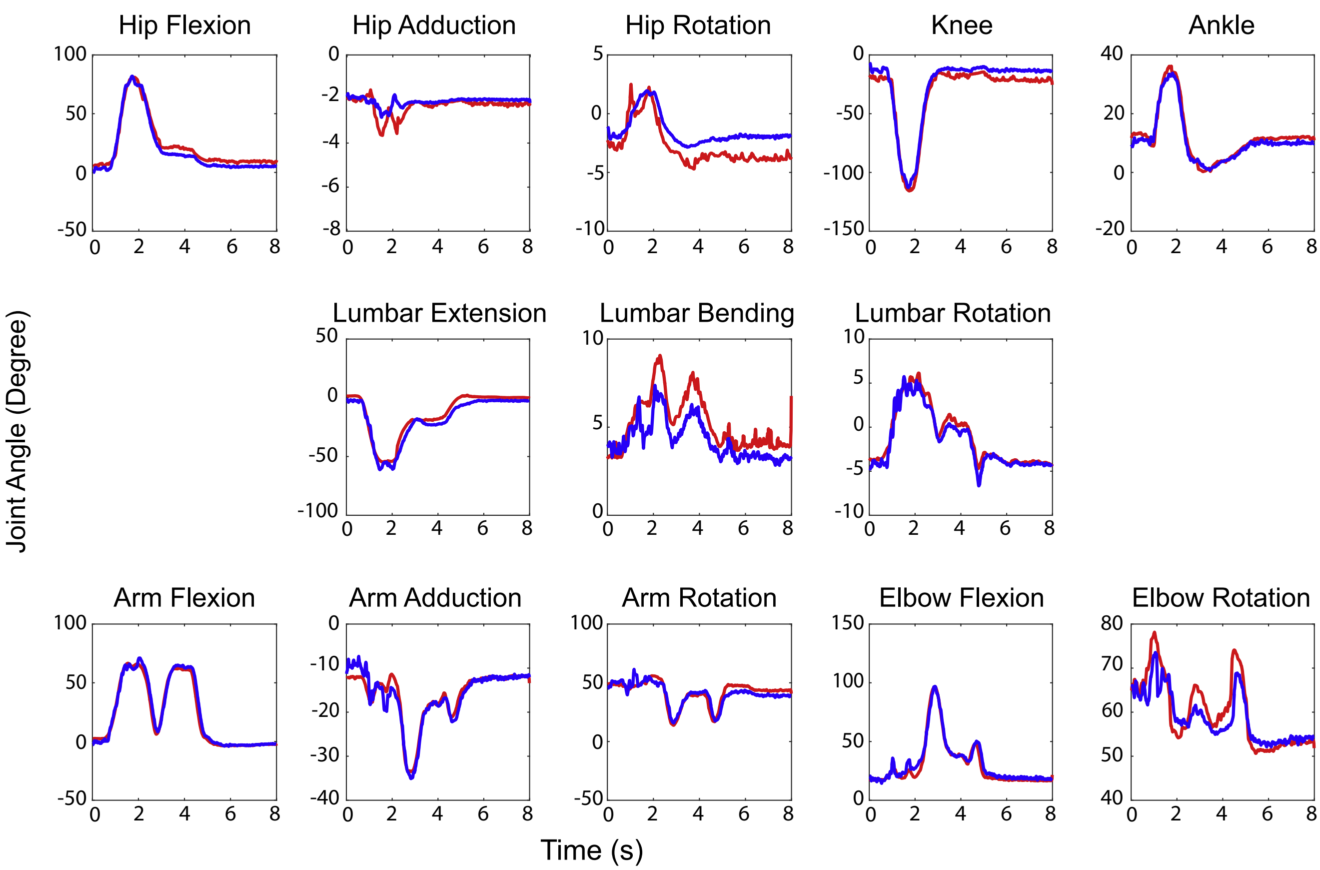

Fig. 3. Joint angles calculated from the marker-based (red line) and the proposed computer vision base marker-less method (blue line) for a representative subject performing a FK lifting. The lifting task begins in a straight upright position and ends in the same position. Vertical axis shows the joint angles in degree and horizontal axis shows the time in second. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

i.e., we divided the videos and corresponding marker based motion data into four partitions, and then used three of them to train the model and the remaining partition for testing each time. By repeating the test four times, we obtained the results for the whole dataset.

2.2.3. Angle calculation

Once the 3D pose measures (3D coordinates of 45 estimated 'virtual' markers) were estimated at each frame, we used OpenSim (Delp et al., 2007) to perform inverse kinematics to calculate the joint kinematics. Both our proposed modified TGP algorithm and the surface marker based method output a set of marker trajectories: 45 estimated 'virtual' marker trajectories for the modified TGP algorithm (our proposed computer vision based marker-less method) and 45 surface marker trajectories (the surface marker based method).

The use of OpenSim for computing joint kinematics from these estimated 'virtual' or surface marker motions required the creation of a subject-specific musculoskeletal model via scaling of a generic model. The musculoskeletal model developed by Hamner et al. (Hamner et al., 2010) was chosen as the generic model. An internal Marker Set defined by (Cappozzo et al., 1995) was placed on the model. The scaling procedure utilized the Scale Model tool in OpenSim, which adjusted the body segment lengths of a generic

Table 2
Mean ± SD of the joint angle differences between the proposed marker-less method and the surface marker-based methods.

Subject	FK Mean ± SD (Degree)	KS Mean ± SD (Degree)	FS Mean ± SD (Degree)	Average
1	2.90 ± 2.21	2.83 ± 2.08	2.76 ± 2.18	2.83 ± 2.16
2	2.26 ± 1.52	2.38 ± 2.19	2.78 ± 1.72	2.48 ± 1.85
3	2.01 ± 1.62	2.53 ± 2.03	2.49 ± 1.90	2.35 ± 1.87
4	2.18 ± 1.65	2.32 ± 1.50	2.73 ± 2.35	2.41 ± 1.89
5	1.53 ± 0.96	1.71 ± 1.30	1.86 ± 1.48	1.70 ± 1.27
6	1.49 ± 0.81	2.07 ± 1.47	2.07 ± 1.39	1.88 ± 1.29
7	2.00 ± 1.12	2.53 ± 1.96	2.34 ± 1.43	2.29 ± 1.56
8	5.92 ± 3.41	2.26 ± 1.66	2.55 ± 1.90	3.58 ± 2.96
9	2.32 ± 1.56	2.99 ± 2.50	2.24 ± 1.36	2.52 ± 1.90
10	2.09 ± 1.22	2.45 ± 1.97	1.86 ± 1.15	2.13 ± 1.51
11	1.40 ± 0.72	1.80 ± 1.20	1.77 ± 1.08	1.66 ± 1.03
12	1.62 ± 0.84	2.09 ± 1.67	2.17 ± 1.21	1.96 ± 1.31
Average	2.31 ± 2.01	2.33 ± 1.87	2.30 ± 1.68	2.31 ± 4.00

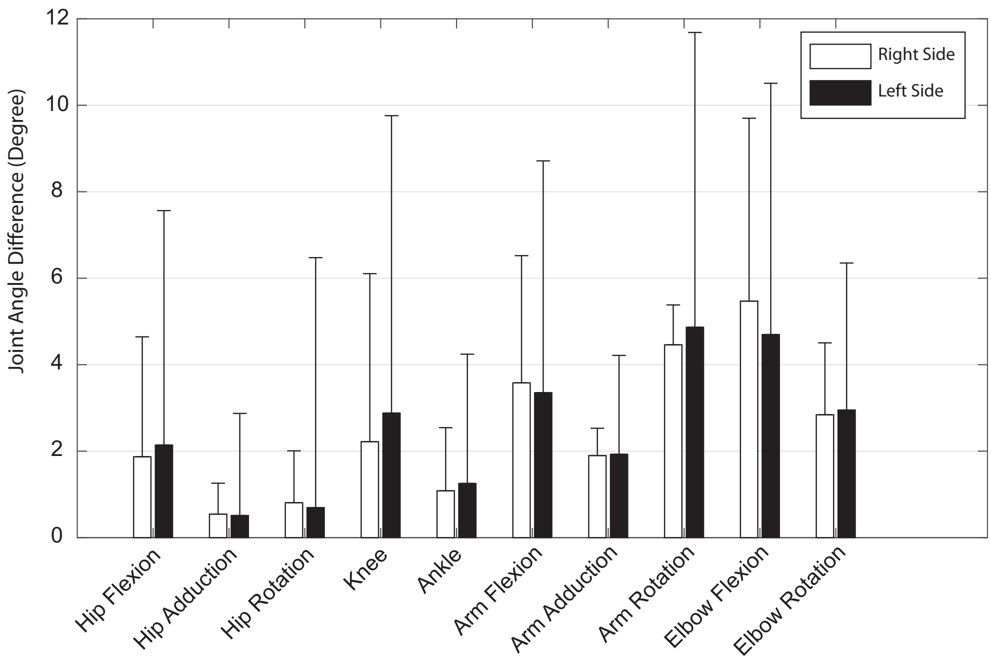

Fig. 4. Comparison of the mean joint angle differences of the left- and right- side of the body. Boxes represent the mean values and bars show the standard deviation.

model to best match the internal markers and the corresponding experimental surface markers captured during a standing posture or the corresponding 'virtual' markers estimated from the computer vision based marker-less method. After scaling, the inverse kinematics tool embedded in OpenSim (Delp et al., 2007) was used to solve for the joint angles such that the discrepancy between experimentally measured surface markers and virtual markers on the model was minimized.

2.3. Data analysis and validation

The proposed computer vision based marker-less method was validated against the surface marker-based method. For each frame of the video, the average Euclidean distance between the estimated 'virtual' marker positions and the corresponding experimental surface marker positions was calculated as the 3D pose measure error for the frame. The 3D pose reconstruction error was calculated as the average of the 3D pose measure errors over all the frames. We also calculated the joint angle difference between the marker-less and marker-based method. Paired t-tests were performed to examine whether the joint angles estimated from the marker-less and marker-based method were significantly different. All the preceding data analyses were performed using MATLAB® programs (The MathWorks, Boston, MA).

3. Results

Our proposed computer vision based marker-less method successfully estimated the 3D human pose as illustrated in Fig. 2. The grand mean ± SD of the 3D pose reconstruction error with the modified TGP algorithm was 9.52 ± 7.95 mm (Table 1).

For all of the lifting trials, there is a good agreement of joint angle trajectories estimated by the computer vision based marker-less and surface marker-based methods as illustrated by Fig. 3. The joint angle differences between the two methods are presented in Table 2. The mean ± standard deviation of the joint angle differences over the whole data set was 2.31 ± 4.00°. The joint angle differences for three lifting styles are similar (2.31 ± 2.01° for FK, 2.33 ± 1.87° for KS, and 2.30 ± 1.68° for FS), but the

Table 3
Mean ± SD of the joint angle differences, and p-values of paired t-tests between the computer vision based marker-less method and the surface marker-based method. Hip, knee, ankle, arm, and elbow angles differences are calculated based on the average of the differences on the left- and right- side of the body.

Joint	Angle	Mean ± SD (Degree)	p-value
Hip	Flexion	2.01 ± 2.86	0.643
	Adduction	0.53 ± 0.67	0.060
	Rotation	0.75 ± 1.07	0.899
Knee	Flexion	2.55 ± 4.07	0.498
Ankle	Flexion	1.17 ± 1.57	0.260
Lumbar	Extension	1.87 ± 2.77	0.253
	Bending	0.57 ± 0.69	0.150
	Rotation	0.74 ± 1.02	0.051
Arm	Flexion	3.47 ± 5.39	0.723
	Adduction	1.91 ± 2.32	0.029
	Rotation	4.66 ± 6.32	0.084
Elbow	Flexion	5.08 ± 6.38	0.071
	Rotation	2.90 ± 3.20	0.140

joint angle differences vary across subjects, from 1.66 ± 1.03° to 3.58 ± 2.96°.

For the symmetrical lifting, the joint angle differences between the computer vision based marker-less method and the surface marker-based method for the left- and right- side joints of the body are comparable (Fig. 4).

Joint angle differences between the computer vision based marker-less method and the surface marker-based method vary across body joint angles as shown in Table 3. The maximum difference was observed for the elbow flexion with the mean ± standard deviation of 5.08 ± 6.38° (3.38% of the range motion). No significant difference was observed in joints angle difference between the two methods for most joint angles except the arm adduction. However, there is no systematic error of the joint angle estimation as illustrated by Fig. 5 based on the distribution of the joint angle differences.

4. Discussion

The present study proposed and validated a computer vision based marker-less method for assessing 3D joint kinematics during symmetrical lifting. The whole body 3D pose during symmetrical lifting was successfully reconstructed by the proposed computer vision based marker-less method. The success of our proposed method in 3D pose reconstruction results from the integration of morphological constraints and the discriminative computer vision approach. The average 3D reconstruction error for the original TGP (Bo and Sminchisescu, 2010) without integrating with the morphological constraint, was much higher (40 mm vs. 9.52 mm, unpublished result). The integration of morphological constraints prevented the infeasible solutions with large reconstruction errors. Thus, our proposed method was able to achieve a low reconstruction error, which enabled us to estimate 3D joint angles with reasonable accuracies. The 3D joint kinematics estimation from the proposed method were practically comparable to the 3D joint kinematics from the surface marker-based method.

The performance of our proposed method may be activity-dependent and be affected by self-occlusion and object-occlusion. We have shown that the joint angle estimation accuracy varies across body joints for the proposed computer vision based marker-less method. Among all of the joints, the elbow flexion angle exhibited the highest difference between the proposed marker-less method and the surface marker-based method. This may be attributed to the self-occlusion and object-occlusion during lifting. When the subject placed the box on the shelf, lower arms could be blocked by the shelf (Fig. 2) and the arms could be occluded by the torso. However, the joint angle estimation accuracy for the left- and right- side of the body were very similar, which may be caused by that all of the lifting motions were symmetrical in this study.

Our proposed method can estimate 3D joint kinematics accurately with a low cost setup. Many computer vision based approaches (Bodor et al., 2003; Mikić et al., 2003; Oreifej and Liu, 2013) have been developed and successfully applied for human

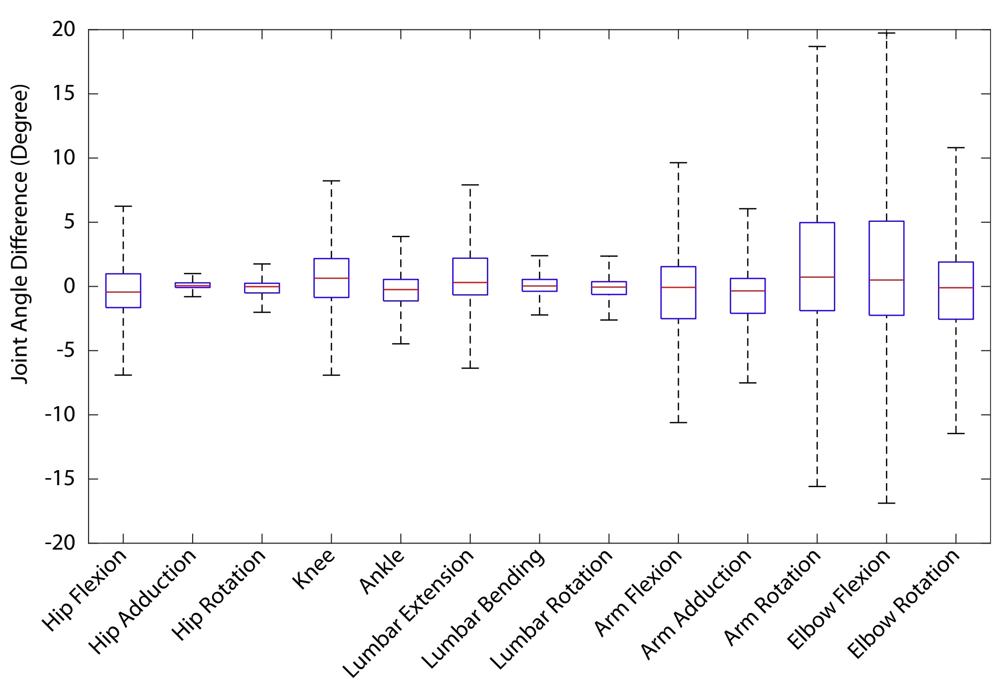

Fig. 5. Distribution of the joint angles difference. Middle line is the median and boxes represents 25 and 75 percentile and bars show the minimum and maximum values. Whisker size is set to two percent.

tracking and activity recognition based on a small number of optical cameras, however, their accuracy for the 3D joint kinematics estimation may not be sufficient for biomechanics research. On the other hand, the studies (Mündermann et al., 2005; Saboune and Charpillet, 2005; Corazza et al., 2006; Ceseracciu et al., 2014; Sandau et al., 2014) capable of achieving suitable accuracy for the 3D joint kinematics estimation, usually use a large number of cameras, which is not always practical in the workplace. Our proposed method seem to work well with a low cost setup, which includes two low-resolution optical cameras. This may result from the image features used in the methods. The HoG features were required to be normalized to the contrast, which can minimize the effect of image quality and account for changes in illumination. Recently depth sensors such as Kinect have been also used for joint kinematics assessment (Dutta, 2012; Diego-Mas and Alcaide-Marzal, 2014; Plantard et al., 2016). However, there are two major disadvantages of using Kinect for workplace activity assessment. First, depth sensors can be only used in a short range of distance from the depth sensors, which may not allow them to be used in large space workplaces and fields (Weerasinghe et al., 2012). Second, depth sensors are sensitive to the environment illumination and would be difficult to use in outdoor environments (Andersen et al., 2012). Since the proposed method uses a simple experimental setup with only two regular optical cameras, it has the potential to be a viable noninvasive alternative of the surface marker based method, especially for outdoor environments.

The ultimate goal of our research is to provide a field noninvasive biomechanical analysis and joint kinematics assessment tool by taking the advantages of advanced computer vision techniques. The present study is a starting point of the research along this direction. There are three limitations about this study that should be further investigated in the future research. First, the proposed method was mainly developed for symmetrical lifting tasks. Whether and how well this method can be extended for measuring the 3D joint kinematics of other common workplace activities, such as asymmetrical lifting, pushing, and pulling would be worth to investigate. Second, the effects of clothing and image quality on the accuracy of the 3D pose estimation have not been investigated. Finally, the proposed method requires training for individual subjects in order to personalize subject-specific models. A more generalized method with a less restricted training requirement will be more favorable.

Conflict of interest statement

The authors have no any financial or proprietary interests in the materials described in the article.

Acknowledgements

This work was supported in part by the NSF (CNS 1229628, CMMI 1334389, IIS 1451292, IIS 1555408, and IIS 1703883). The human motion data was collected at Liberty Mutual Research Institute for Safety when the third author was in Harvard School of Public Health - Liberty Mutual postdoctoral program.

References

Andersen, M.R., Jensen, T., Lisouski, P., Mortensen, A.K., Hansen, M.K., Gregersen, T., and Ahrendt P., (2012). Kinect depth sensor evaluation for computer vision applications. Electrical and Computer Engineering Technical Report ECE-TR-6.

Bo, L., Sminchisescu, C., 2010. Twin gaussian processes for structured prediction. Internat. J. Comput. Vision 87 (1–2), 28–52.

Bodor, R., Jackson, B., and Papanikolopoulos, N., (2003). Vision-based human tracking and activity recognition. Proc. of the 11th Mediterranean Conf. on Control and Automation

Cappozzo, A., Catani, F., Della Croce, U., Leardini, A., 1995. Position and orientation in space of bones during movement: anatomical frame definition and determination. Clin. Biomech. 10 (4), 171–178.

Ceseracciu, E., Sawacha, Z., Cobelli, C., 2014. Comparison of markerless and marker-based motion capture technologies through simultaneous data collection during gait: proof of concept. PloS One 9 (3), e87640.

Corazza, S., Muendermann, L., Chaudhari, A., Demattio, T., Cobelli, C., Andriacchi, T. P., 2006. A markerless motion capture system to study musculoskeletal biomechanics: visual hull and simulated annealing approach. Annal. Biomed. Eng. 34 (6), 1019–1029.

da Costa, B.R., Vieira, E.R., 2010. Risk factors for work-related musculoskeletal disorders: a systematic review of recent longitudinal studies. Am. J. Indust. Med. 53 (3), 285–323.

Dalal, N., and Triggs, B., (2005). Histograms of oriented gradients for human detection. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), IEEE

Delp, S.L., Anderson, F.C., Arnold, A.S., Loan, P., Habib, A., John, C.T., Guendelman, E., Thelen, D.G., 2007. OpenSim: open-source software to create and analyze dynamic simulations of movement. IEEE Transact. Biomed. Eng. 54 (11), 1940–1950.

Diego-Mas, J.A., Alcaide-Marzal, J., 2014. Using Kinect™ sensor in observational methods for assessing postures at work. Appl. Ergon. 45 (4), 976–985.

Drory, A., Li, H., Hartley, R., 2017. A learning-based markerless approach for full-body kinematics estimation in-natura from a single image. J. Biomech. 55, 1–10.

Dutta, T., 2012. Evaluation of the Kinect™ sensor for 3-D kinematic measurement in the workplace. Appl. Ergon. 43 (4), 645–649.

Gavrila, D.M., 1999. The visual analysis of human movement: a survey. Comput. Vision Image Understand. 73 (1), 82–98.

Hamner, S.R., Seth, A., Delp, S.L., 2010. Muscle contributions to propulsion and support during running. J. Biomech. 43 (14), 2709–2716.

Holte, M.B., Tran, C., Trivedi, M.M., Moeslund, T.B., 2012. Human pose estimation and activity recognition from multi-view videos: comparative explorations of recent developments. IEEE J. Selected Top. Signal Process. 6 (5), 538–552.

Kuiper, J.I., Burdorf, A., Verbeek, J.H., Frings-Dresen, M.H., van der Beek, A.J., Viikari-Juntura, E.R., 1999. Epidemiologic evidence on manual materials handling as a risk factor for back disorders: a systematic review. Int. J. Ind. Ergon. 24 (4), 389–404.

Mehrizi, R., Xu, X., Zhang, S., Pavlovic, V., Metaxas, D., Li, K., 2017. Using a markerless method for estimating L5/S1 moments during symmetrical lifting. Appl. Ergon.

Mikić, I., Trivedi, M., Hunter, E., Cosman, P., 2003. Human body model acquisition and tracking using voxel data. Int. J. Comput. Vision 53 (3), 199–223.

Mündermann, L., Anguelov, D., Corazza, S., Chaudhari, A.M., and Andriacchi, T.P., (2005). Validation of a markerless motion capture system for the calculation of lower extremity kinematics. Proc. American Society of Biomechanics, Cleveland, USA

Mündermann, L., Corazza, S., Andriacchi, T.P., 2006. The evolution of methods for the capture of human movement leading to markerless motion capture for biomechanical applications. J. NeuroEng. Rehabil. 3 (1), 1.

Nimbarte, A.D., Aghazadeh, F., Ikuma, L.H., Harvey, C.M., 2010. Neck disorders among construction workers: understanding the physical loads on the cervical spine during static lifting tasks. Industrial health 48 (2), 145–153.

Oreifej, O., and Liu, Z., 2013. Hon4d: Histogram of oriented 4d normals for activity recognition from depth sequences. in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition

Plantard, P., Shum, H.P., Le Pierres, A.-S., Multon, F., 2016. Validation of an ergonomic assessment method using Kinect data in real workplace conditions. Appl. Ergon.

Poppe, R., 2007. Vision-based human motion analysis: an overview. Comput. Vision Image Understand. 108 (1), 4–18.

Saboune, J., and François Charpillet. (2005). "Markerless human motion capture for gait analysis. arXiv preprint cs/0510063.

Sandau, M., Koblauch, H., Moeslund, T.B., Aanæs, H., Alkjær, T., Simonsen, E.B., 2014. Markerless motion capture can provide reliable 3D gait kinematics in the sagittal and frontal plane. Med. Eng. Phys. 36 (9), 1168–1175.

Suard, F., Rakotomamonjy, A., Bensrhair, A., and Broggi, A., (2006). Pedestrian detection using infrared images and histograms of oriented gradients. Intelligent Vehicles Symposium, 2006 IEEE, IEEE.

Weerasinghe, I.T., Ruwanpura, J.Y., Boyd, J.E., Habib, A.F., 2012. Application of Microsoft Kinect sensor for tracking construction workers. Construct. Res. Congress 2012: Construct. Challenges in a Flat World.

Yang, Y., and Ramanan, D., (2011). Articulated pose estimation with flexible mixtures-of-parts. Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, IEEE

Zhu, Q., Yeh, M.-C., Cheng, K.-T., and Avidan, S., (2006). Fast human detection using a cascade of histograms of oriented gradients. Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on, IEEE