Published in final edited form as:

Ergonomics. 2015 December ; 58(12): 2057–2066. doi:10.1080/00140139.2015.1051594.

The Accuracy of Conventional 2D Video for Quantifying Upper Limb Kinematics in Repetitive Motion Occupational Tasks

Chia-Hsiung Chen¹, David Azari², Yu Hen Hu¹, Mary J. Lindstrom³, Darryl Thelen⁴, Thomas Y. Yen², and Robert G. Radwin^{2,*}

¹Department of Electrical and Computer Engineering, University of Wisconsin-Madison

²Department of Industrial and Systems Engineering, University of Wisconsin-Madison

³Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison

⁴Department of Mechanical Engineering, University of Wisconsin-Madison

Abstract

Objective—Marker-less 2D video tracking was studied as a practical means to measure upper limb kinematics for ergonomics evaluations.

Background—Hand activity level (HAL) can be estimated from speed and duty cycle. Accuracy was measured using a cross correlation template-matching algorithm for tracking a region of interest on the upper extremities.

Methods—Ten participants performed a paced load transfer task while varying HAL (2, 4, and 5) and load (2.2 N, 8.9 N and 17.8 N). Speed and acceleration measured from 2D video were compared against ground truth measurements using 3D infrared motion capture.

Results—The median absolute difference between 2D video and 3D motion capture was 86.5 mm/s for speed, and 591 mm/s² for acceleration, and less than 93 mm/s for speed and 656 mm/s² for acceleration when camera pan and tilt were within ± 30 degrees.

Conclusion—Single-camera 2D video had sufficient accuracy (< 100 mm/s) for evaluating HAL.

Practitioner Summary—This study demonstrated that 2D video tracking had sufficient accuracy to measure HAL for ascertaining the American Conference of Government Industrial Hygienists Threshold Limit Value® for repetitive motion when the camera is located within ± 30 degrees off the plane of motion when compared against 3D motion capture for a simulated repetitive motion task.

Keywords

Repetitive motion; hand activity level; marker-less video tracking; work related musculoskeletal disorders; exposure assessment

*Corresponding Author Robert G. Radwin, PhD, 1550 Engineering Drive, University of Wisconsin-Madison, Madison, WI 53706, 608-263-6596, 608-262-8454, radwin@discovery.wisc.edu.

1. Introduction

Upper extremity musculoskeletal injuries are common in hand intensive hand work involving repetitive motions and exertions (Fan, et al., 2009; Harris, et al., 2011; Hegmann, et al., 2009; Silverstein, et al., 2010). The annual Workplace Safety Index (Liberty Mutual, 2014) reports that upper limb *repetitive motion* injuries were among the top ten causes of disabling in 2012 and accounted for \$1.8 Billion in workers compensation costs alone.

A multi-institutional consortium supported by the CDC (NIOSH) recently completed a landmark prospective study of upper extremity work-related musculoskeletal disorders in a wide variety of industries involving video data collection on 3,287 workers for (Burt, et al., 2011; Fan, et al., 2009; Garg, et al., 2009; Garg, et al., 2012; Gerr, et al., 2014; Harris, et al., 2011; Kapellusch et al., 2014; Wurzelbacher et al., 2010). This data set holds great promise for establishing dose-response relationships among physical stresses and health outcomes. The methods used for exposure assessment, mostly involved either direct measurements using instruments attached to a worker's hands or arms, or indirect observations. Direct measurement for assessing workplace tasks remains challenging because attaching sensors on working hands is time consuming, invasive, produces large quantities of data, requires training and expertise, and may interfere with normal working operations (Radwin, Lin & Yen, 1994; Lin & Radwin, 1998; Yen & Radwin, 2000). A more practical method is needed for assessing repetitive motion in the workplace that can be readily used by industry practitioners.

While over the past few decades there was significant advancement in visual tracking in the computer science field, few of those techniques are applicable to evaluating repetitive motion exposure in the industrial environment. This is because tracking accuracy is often limited by poor illumination, space constraints, and visual obstructions in the workplace. Three dimensional visual tracking is desirable, although most solutions require prior knowledge of the model (Choi and Christensen, 2012), or require placement of special visual markers on the target (Armstrong, et al., 2002), which is not desirable and most often times not achievable due to interference and safety concerns. Although progress has been made in recovering 3D motion using a single camera (Davison, et al., 2007), currently the most widely available industrial worker videos were filmed and stored in traditional 2D format. These videos are usually taken from non-ideal camera vantage points, and contain poor resolution, quality, and significant noise.

We have developed novel video processing software for automatically tracking hand motion and extracting motion kinematics using a cross correlation-based feature extraction template-matching algorithm to track the motion (Chen, et al., 2013) that does not require sensors or instruments. A direct measurement of hand speed makes it possible to evaluate the American Conference of Government Industrial Hygienists (2001) Threshold Limit Value[®] (TLV[®]) for hand activity level (HAL).

The HAL scale is reported in integers from 0 to 10 and is based on the movement frequency and duty cycle (Latko, et al., 1997; Radwin, et al., 2015). In Akkas et al. (2015) we developed a new equation for computing HAL directly from tracked hand speed and duty

cycle rather than relying on estimates of frequency for an automated instrument to directly measure HAL. We have found that 100 mm/s increase in speed increased HAL by approximately 0.5 units. Consequently, if 2D video tracking were sufficiently accurate within those bounds, it would be an acceptable means for evaluating HAL for repetitive jobs.

The objective of the current study was to test the accuracy of estimating motion kinematics automatically and unobtrusively from single-camera marker-less video by tracking a single region of interest (ROI) on the hands or arms. Specifically we estimated differences between 2D video and infrared 3D motion tracking for a simulated repetitive motion task. Since it is not often possible to locate the video camera directly perpendicular to the plane of motion in a workplace setting, we also investigated accuracy when the camera pan and tilt angles are varied.

2. Methods

2.1 Participants

Ten young, healthy volunteers (18 to 32 years) were recruited from the University of Wisconsin–Madison. There were 5 males and 5 females; all were right-hand dominant, except one male. The protocol was reviewed and approved by the Institutional Review Board of University of Wisconsin–Madison prior to recruitment and all participated with informed consent.

2.2 Apparatus and Experimental Procedures

A laboratory mock-up of a simple load transfer task was constructed to simulate repetitive motion activities typically performed in an industrial setting and is depicted in Figure 1. The apparatus consisted of a 6 RPM turntable driven by an electric motor and attached to a chute to collect the weighted bottles. The weighted plastic bottles were filled with calibrated quantities of lead shot.

The frequency and duty cycle were paced using three audio cues to achieve various hand activity levels (HAL) as described by Latko et al (1997). The first tone signaled to reach, grasp, and move the bottle from the dispenser to the turntable. The second tone signaled the subject to return the hand to the rest position. The third tone signaled the start of a rest period where the subject was instructed to remain at rest until the cycle repeated. A cycle for the task was depicted in Figure 1. Duty cycle was measured as the percent of hand exertion time to the total cycle time of the activity (i.e. exertion time/cycle time). Frequency was the rate of exertions in exertions per second or Hz.

The experimental conditions are described in Table 1. The order was randomized for each participant. Participants were required to perform nine sets of 10 to 30 cycles each, lasting between 60 and 200 seconds. Practice was provided before each condition and minutes of rest were given in between conditions in order to prevent fatigue.

A single JVC Everio GZ-MS230BUS video camera was positioned three meters to the right of the participant and set to record at 30 frames per second (fps). Each trial was videoed and

stored as a single 720×480 resolution MPEG video file. Five Optotrak 3020 infrared sensors sampled at 60 Hz each were placed on each subject, of which three of them were on the back of the hand and two were on the lower arm. The sensor location on the lower arm was utilized in this analysis since it contained the least occlusion.

2.3 Video Tracking Algorithm

Infrared 3D motion tracking markers were located in the same point on the hand as the regions of interest (ROI) for ground truth measurements and the resulting speed and acceleration measures were compared against those obtained from 2D conventional video. Rectangular ROIs were 20 pixels $\times$ 20 pixels and marked manually in the initial video frame to identify the focal area where activities are to be tracked, such as a point on the hand or arm. The ROI was located on the forearm near the infrared marker. The small markers did not interfere with tracking and their visibility permitted us to closely align them with the ROI.

A template matching tracking algorithm was implemented to track the ROI motion trajectory over subsequent video frames. To elaborate, denote $\mathbf{r}_i$ to be a vector of intensity values of all pixels within the ROI at the i^{th} frame, and $\mathbf{r}_{i+1}(\mathbf{w})$ to be a candidate ROI at the $(i+1)^{th}$ frame with a displacement of $\mathbf{w} \in \Omega$ in a pre-specified search area. Currently Ω is forced to integer pixel only, and is statically defined to a fixed-sized search region. The cross-correlation between $\mathbf{r}_i$ and $\mathbf{r}_{i+1}(\mathbf{w})$ is defined as the angle between these two vectors ($\|\mathbf{r}_i\|$ is the magnitude of the vector $\mathbf{r}_i$, $\mathbf{r}_i^T$ is the transpose of $\mathbf{r}_i$):

$$R(\mathbf{w}) = \mathbf{r}_i^T \mathbf{r}_{i+1}(\mathbf{w}) / (\|\mathbf{r}_i\| \cdot \|\mathbf{r}_{i+1}(\mathbf{w})\|)$$

The displacement $\mathbf{w}^* = \arg. \max_{\mathbf{w}} R(\mathbf{w})$ determines the updated position of the ROI at the $(i+1)^{th}$ frame, $\mathbf{r}_{i+1} = \mathbf{r}_{i+1}(\mathbf{w}^*)$. The vector sequence of $\mathbf{w}^*$ is recorded as the displacement between these two frames, and the new position of the ROI yields the motion trajectory at the $(i+1)^{th}$ frame. The algorithm was programmed in C# with the open source OpenCVSharp (.Net wrapper for the OpenCV) vision library.

The speed and acceleration values were calculated using the framework shown in Figure 2. The Optotrak 3D location data (x, y, z) was interpolated, Butterworth low-pass filtered, and down-sampled from 60Hz to 30Hz (x', y', z') in order to match the video sampling rate. The raw video location data (x_{vid}, y_{vid}) and the corresponding down-sampled 3D location data (x', y', z') were compared by their corresponding speed and acceleration (Figure 2).

The video raw location values x_{vid}, y_{vid} were low-pass filtered using a Butterworth filter and scaled to bridge the pixel-based measurement and the physical measurement. The scaling constant was determined by calculating the ratio between the mean 2D Optotrak distance in centimeters and the video tracking distance in pixels between the upper and lower 15 percentile x -coordinate data. Different scaling constants were selected for each subject to address potential pan/tilt/zoom/location changes of the video camera among recording sessions.

The velocity (v_x, v_y) and acceleration (a_x, a_y) at each time stamp were calculated as follows:

$$\begin{aligned} \nu_{p,i} &= (p_{i+1} - p_{i-1}) / 2\Delta \\ a_{p,i} &= (p_{i+1} - 2 \times p_i + p_{i-1}) / \Delta^2, \end{aligned}$$

where $p \in \{x, y\}, i \in \{2, 3, \dots\}$. In our experiment, $\Delta = 1/30$ s, which is the sampling period of the video. The magnitudes of speed and acceleration were determined by:

$$\begin{aligned} \nu_{xy, \text{video}} &= \sqrt{(\nu_x)^2 + (\nu_y)^2} \\ a_{xy, \text{video}} &= \sqrt{(a_x)^2 + (a_y)^2}. \end{aligned}$$

The Optotrak ground truth speed and acceleration were calculated similar to the video, using down-sampled location values x', y', z' . Both 2D and 3D speed and acceleration data were calculated. The 3D version was calculated using all three components of the location data (x', y', z'), while the 2D version only utilized x, y components of the location data (x', y').

2.4 Simulation on Camera Placed at Non-Optimal Location

In order to study the error induced by camera placement, we compared and examined the difference between the video tracking values and the 3D infrared marker data projected onto various 2D planes corresponding to different pan/tilt combinations with respect to the sagittal plane camera placement. This was possible without actually varying the 2D camera location because the Optotrak tracked the ROI in 3D.

We simulated in Matlab what would have been captured by a video camera from varying views by utilizing the 3D Optotrak data, and then comparing the original video-tracking captured data with respect to those produced by the simulated camera. The down-sampled 3D Optotrak data was first panned and tilted to an angle within ± 30 degree. The pan-tilt-roll rotation angles are shown in Figure 3 (a), where $\theta_1, \theta_2, \theta_3$ represents pan, tilt, and roll angle about $+x, +y$, and $+z$ axis, respectively. The x - y - z rotation transformation is given by

$$\mathbf{R} = \begin{bmatrix} c_2 c_3 & -c_2 s_3 & s_2 \\ s_1 s_2 c_3 + s_3 c_1 & -s_1 s_2 s_3 + c_3 c_1 & -s_1 c_2 \\ -c_1 s_2 c_3 + s_3 s_1 & c_1 s_2 s_3 + c_3 s_1 & c_1 c_2 \end{bmatrix},$$

where $s_i = \sin \theta_i$, and $c_i = \cos \theta_i$. To transform marker coordinates from x - y - z to x' - y' - z' as in Figure 3 (b), we multiplied the position vector by the transpose of the matrix $\mathbf{R}$,

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \mathbf{R}^T \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

The speed and acceleration in the new projected plane were then calculated based on the rotated, depth-dropped data. We calculated and recorded the median absolute difference of

the speed and acceleration between the video tracking and those of the projected plane for each pan-tilt angle pairs.

2.5 Statistical Methods

The effects of HAL and load on the median error in speed and acceleration between video and the 2D or 3D Optotrak were assessed using linear mixed effects models (R Core Team, 2014). Linear mixed effects models are regression models where additional error terms are included to account for the nested structure of the data (9 experimental runs within each subject) (West, et al., 2014). Approximate maximum likelihood estimation was used (Pinheiro, et al., 2014). Each of the models included a random effect for subject and fixed effects for load, HAL and the 2-way, load-by-HAL interaction term. This term allows the effect of load to vary for each level of HAL. Median error was transformed to the square root scale before analysis to obtain constant variance over the range of the responses.

3. Results

Representative time series plots of speed magnitude for the 3D Optotrak (i.e. ground truth) measurement compared to the 2D sagittal plane Optotrak measurement and the marker-less video measurement are shown in Figure 4. Load and HAL were varied to obtain a range of velocities and accelerations. By varying weight and HAL, the median speed increased as the weight increased, with median speed ranging from 298.1 mm/s to 348.4 mm/s for the Optotrak 3D, and 209.5 mm/s to 248.5 mm/s for the Optotrak 2D. Similar time series plots of acceleration for the 3D Optotrak ground truth measurement compared to the 2D sagittal plane Optotrak ground truth measurement and the marker-less video measurement are shown in Figure 5. A similar trend was observed for the median acceleration, where the median acceleration ranged from 1625.0 mm/s² to 1996.2 mm/s² for the Optotrak 3D, and 1195.1 mm/s² to 1503.8 mm/s² for the Optotrak 2D.

The percent change in median error due to load was approximately 6% for speed and 13% for acceleration. HAL had a significant effect on median error for acceleration ($p < .0001$) but not for speed ($p > 0.1304$) for both 2D and 3D comparisons. Pairwise comparisons between the levels of HAL for acceleration were significant ($p < .0283$) except for HAL 4 vs HAL 5 for Video vs Optotrak 3D ($p = 0.2204$). The increase in median error from HAL = 2 to HAL = 4 for acceleration is substantial for both the 2D and 3D comparisons 58% and 46% percent respectively.

Summary statistics are shown in Table 2. The estimated values for the median error in speed and acceleration for different HAL values are shown in Figure 6. The estimated median error for speed and acceleration was greater for the video versus Optotrak 3D compared to the video versus Optotrak 2D.

The median error across various camera positions of speed and acceleration measurements for the video tracking with respect to the projection of the rotated 3D Optotrak is shown in Figure 7. In all cases, the median error was smallest at positive pan angles near the center with 0 pan/tilt, and largest towards the edges and corners, where motion is distorted in the 2D viewing plane. Inspection of Figure 7 indicates a slightly higher sensitivity to changes in

pan angle rather than tilt angle for this particular task. For the video tracking, the median error starts from 59.8 mm/s and 500.9 mm/s² for speed and acceleration, respectively; increase to (< 72 mm/s and < 531 mm/s²) within 6 degrees off center, (< 75 mm/s and < 575 mm/s²) at 12 degrees rotation in both directions, and tops at (93 mm/s and 656 mm/s²) at the edges.

4. Discussion

Akkas et al. (2015) estimated that every 100 mm/s increase in RMS speed increased HAL by approximately 0.5 units. The median absolute difference between video tracking and the 2D Optotrak (55.3 mm/s) as well as the difference between video marker-less tracking and the 3D Optotrak (86.5 mm/s) fell within that range. The 3D Optotrak data contained out of plane motion not captured by the 2D video or 2D depth-dropped Optotrak data, resulting in a smaller median absolute difference among the 2D datasets. Therefore we conclude for a representative laboratory load transfer task that video tracking was sufficient for estimating HAL.

The video measurement compared to the Optotrak simulation showed greater median error and distortion for all pan and tilt angles starting at 0 degrees. This discrepancy is likely due to force of integer pixel on the video tracking system, in addition to different sampling rates and resolution between the two systems. The Optotrak was accurate up to 0.1 mm every 1/60th second, while the video tracking location was only accurate to one pixel every 1/30th second. The camera distance dictated an average pixel ratio of about 1:3.4, or 1 pixel representing 3.4 millimeters of motion. As a result, small deviations in the 2D pixel position induced several tens of mm/s median error in speed, and several hundreds of mm/s² median error in acceleration. The minimum speed and acceleration median error (59.8 mm/s, 500.9 mm/s² respectively) support this notion. The change in median error for different pan/tilt combinations with respect to the minimum at no rotation was up to 93 mm/s and 656 mm/s² for speed and acceleration, respectively.

Our analysis shows that within ± 30 degrees of camera pan and tilt, the median error between the video tracking and the Optotrak was less than 93 mm/s for speed, within the 100 mm/s error limit for estimating HAL (Akkas, et al., 2015). Note that the transformation changed the orientation but did not account for translations between frames. This should be sufficient for our analysis since we are only interested in speeds and accelerations, which do not depend on the location of the origin. In this analysis we also did not look into roll angle since they would not affect the magnitude of speed and acceleration on the image plane.

Our previous studies (Chen, et al., 2013; Akkas, et al., 2015) focused on methods to estimate HAL rating from the acquired motion kinematics, while this paper focused on estimation of accuracy of the extracted video-based motion kinematics itself. The proof-of-concept investigation by Chen et al. (2013) focused on describing a method to extract the frequency and duty cycle from the acquired spatiotemporal data, and the Akkas et al. (2015) study focused on devising a new set of evaluation criteria (hand speed and duty cycle) for estimating the HAL rating using the acquired spatiotemporal data. The current study focused on estimation of accuracy of the video-based method compared to the 3D Optotrak

measurement and estimation of error if the camera were to be placed at a non-optimal position.

This study demonstrated the feasibility of measuring video kinematics and its potential for an exposure assessment methodology based on single digital video recording and marker-less tracking. In the actual workplace environment, tracking moving body parts may be more challenging. Tracking failure may occur due to lighting variation, highly occlusion, and posture variations. It may not always be possible to locate the camera at the optimal vantage point for video-based assessment, which may affect the speed, and acceleration assessment depending on the amount of the out-of-plane movements. We anticipate that the use of 2D video will simplify kinematic measurements and be less susceptible to these error sources.

5. Summary and Conclusions

1. The median absolute difference between video tracking and the 2D Optotrak (55.3 mm/s) as well as the difference between video marker-less tracking and the 3D Optotrak (86.5 mm/s) fell within a 100 mm/s range, resulting in error within ± 0.5 HAL.
2. The median error between the video tracking and the Optotrak for camera pan and tilt within ± 30 degrees was less than 93 mm/s for speed, within the 100 mm/s error limit for estimating HAL.
3. Marker-less 2D video tracking is a practical means for measuring upper limb kinematics from conventional single camera digital video for use in workplace ergonomics evaluations

Acknowledgements

This research was funded in part by a grant from the National Institutes of Health, 1R21 EB01458301 (Radwin). The authors wish to thank Mr. Steven Nelms for assistance building the repetitive motion task laboratory apparatus.

References

- Akkas O, Azari DP, Chen C-H, Hu YH, Ulin SS, Armstrong TJ, Rempel D, Radwin RG. A hand speed and duty cycle equation for estimating the ACGIH hand activity level rating. *Ergonomics*. 2015; 58(2):184–194. [PubMed: 25343278]
- ACGIH Worldwide. Hand Activity Level TLV®. Cincinnati, OH: 2001.
- Armstrong, B.; Verron, T.; Heppe, L.; Reynolds, J.; Schmidt, K. RGR-3D: simple, cheap detection of 6-DOF pose for teleoperation, and robot programming and calibration. In *Robotics and Automation, 2002; Proceedings. ICRA'02. IEEE International Conference on*; IEEE. (2002); p. 2938-2943.
- Burt S, Crombie K, Jin Y, Wurzelbacher S, Ramsey J, Deddens J. Workplace and individual risk factors for carpal tunnel syndrome, *Occupational and Environmental Medicine*. 2011; 68:928–933. [PubMed: 21613639]
- Chen C-H, Hu YH, Yen TY, Radwin RG. Automated video exposure assessment of repetitive hand activity level for a load transfer task. *Human Factors*. 2013; 55(2):298–308. [PubMed: 23691826]
- Choi C, Christensen HI. Robust 3D visual tracking using particle filtering on the special Euclidean group: A combined approach of keypoint and edge features. *The International Journal of Robotics Research*. 2012; 31(4):498–519.
- Davison AJ, Reid ID, Molton ND, Stasse O. MonoSLAM: Real-time single camera SLAM. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*. 2007; 29(6):1052–1067.

- Fan ZJ, Silverstein BA, Bao S, Bonauto DK, Howard NL, Spielholz PO, et al. Quantitative exposure-response relations between physical workload and prevalence of lateral epicondylitis in a working population. *American Journal of Industrial Medicine*. 2009; 52(6):479–490. [PubMed: 19347903]
- Garg, A.; Kapellusch, J. Consortium Pooled Data Job Physical Exposure Assessment; Paper presented at the 17th World Congress in Ergonomics; (2009);
- Garg A, Kapellusch J, Hegmann K, Wertsch J, Merryweather A, Deckow-Schaefer G, Malloy EJ, WISTAH Hand Study Research Team. The Strain Index (SI) and Threshold Limit Value (TLV) for Hand Activity Level (HAL): risk of carpal tunnel syndrome (CTS) in a prospective cohort, *Ergonomics*. 2012; 55(4):396–414.
- Gerr F, Fethke NB, Merlino L, Anton D, Rosecrance J, Jones MP, Marcus M, Meyers A.R. A prospective study of musculoskeletal outcomes among manufacturing workers: I. Effects of physical risk factors. *Human Factors*. 2014; 56(1):112–30. [PubMed: 24669547]
- Harris C, Eisen E, Goldberg R, Krause N, Rempel D. 1st place, PREMUS best paper competition: workplace and individual factors in wrist tendinosis among blue-collar workers - the San Francisco study. *Scand J Work Environ Health*. 2011
- Hegmann, KT.; Thiese, MS.; Ott, U.; Oostema, S.; Garg, A.; Kapellusch, J., et al. Prospective Cohort Study of Upper Extremity MSDs Among 17 Diverse Employers; Paper presented at the 17th World Congress in Ergonomics; (2009);
- Kapellusch JM, Garg A, Milwaukee, Hegmann KT, Thiese MS, Malloy EJ. The Strain Index and ACGIH TLV for HAL: Risk of Trigger Digit in the WISTAH Prospective Cohort. *Human Factors*. 2014; 56(1):98–111. [PubMed: 24669546]
- Latko WA, Armstrong TJ, Foulke JA, Herrin GD, Rabourn RA, Ulin SS. Development and evaluation of an observational method for assessing repetition in hand tasks. *American Industrial Hygiene Association Journal*. 1997; 58(4):278–285. [PubMed: 9115085]
- Liberty Mutual. 2014 Liberty Mutual Workplace Safety Index, Liberty Mutual Research Institute for Safety. Hopkinton, MA: 2014.
- Lin ML, Radwin RG. Agreement between a frequency-weighted filter for continuous biomechanical measurements of repetitive wrist flexion against a load and published psychophysical data. *Ergonomics*. 1998; 41(4):459–475. [PubMed: 9557587]
- Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team. *nlme: Linear and Nonlinear Mixed Effects Models*. 2014
- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: 2014.
- Radwin RG, Lin ML, Yen TY. Rapid Communication. Exposure assessment of biomechanical stress in repetitive manual work using frequency-weighted filters. *Ergonomics*. 1994; 37(12):1984–1998. [PubMed: 7828580]
- Radwin RG, Azari DP, Lindstrom MJ, Ulin SS, Armstrong TJ, Rempel D. A frequency-duty cycle equation for the ACGIH hand activity level. *Ergonomics*. 2015; 58(2):173–183. [PubMed: 25343340]
- Silverstein B, Fan Z, Bonauto D, Bao S, Smith C, Howard N, et al. The natural course of carpal tunnel syndrome in a working population. *Scandinavian Journal of Work Environment and Health*. 2010; 36(5):384–393.
- West BT, Welch KB, Galecki AT, Gillespie BW. *Linear mixed models: A practical guide using statistical software*. 2014
- Wurzelbacher S, Burt S, Crombie K, Ramsey J, Luo L, Allee S, et al. A comparison of assessment methods of hand activity and force for use in calculating the ACGIH(R) hand activity level (HAL) TLV(R). *J Occup Environ Hyg*. 2010; 7(7):407–416. [PubMed: 20446152]
- Yen TY, Radwin RG. Comparison between using spectral analysis of electrogoniometer data and observational analysis to quantify repetitive motion and ergonomic changes in cyclical industrial work. *Ergonomics*. 2000; 43(1):106–132. [PubMed: 10661695]

Figure 1.

Laboratory repetitive transfer task where a participant (a) reaches for a bottle; (b) grasps and moves the bottle; (c) releases the bottle and returns to the starting position; and (d) rests until next cycle. The plots in (e) and (f) represent the respective speed and acceleration of the arm ROI in (a)-(d) starting at time = 7 s.

Figure 2.

Overview of the verification framework for the kinematics data acquired from video tracking.

Figure 3. Illustration of the rotation on the coordinate system, where (a) shows the original coordinate system defined by the Optotrak, and (b) shows the rotated coordinate system in the dashed line.

Figure 4. Representative record of ROI speed v. time for the 3D Optotrak ground truth measurement, 2D sagittal plane Optotrak measurement, and marker-less video measurement.

Figure 5.
Representative record of ROI acceleration v. time for the 3D Optotrak ground truth measurement, 2D sagittal plane Optotrak measurement, and marker-less video measurement.

Figure 6.
Estimated values for the median error for varying HAL.

Figure 7.
The median errors for the speed and acceleration between different pan/tilt angles represented in heat maps.

Table 1

Pace (Frequency and Duty Cycle) and Load Conditions

Paced HAL	Frequency (Hz)	Duty Cycle (%)	Load (kg)
2	0.25	10	0.2
2	0.25	10	0.9
2	0.25	10	1.8
4	0.50	20	0.2
4	0.50	20	0.9
4	0.50	20	1.8
5	0.50	50	0.2
5	0.50	50	0.9
5	0.50	50	1.8

Table 2

Summary statistics for 2D/3D Optotrak and video tracking across all cases.

	Data Type	Median	(25 th , 75 th) percentile	Maximum
Speed (mm/s)	2D Optotrak	231.1	(119.8, 355.7)	1045.1
	3D Optotrak	332.9	(168.8, 505.7)	1405.1
	Video	245.2	(117.1, 405.1)	1210.7
	Video - 2D Optotrak	55.3	(22.7, 116.3)	863.1
	Video - 3D Optotrak	86.5	(38.4, 162.1)	984.9
Acceleration (mm/s²)	2D Optotrak	1410.6	(830.6, 2215.8)	10931.6
	3D Optotrak	1883.8	(1191.3, 2789.0)	11273.8
	Video	1621.7	(888.0, 2595.4)	9927.3
	Video - 2D Optotrak	515.2	(230.5, 973.0)	7964.3
	Video - 3D Optotrak	591.6	(270.4, 1082.1)	7950.7