ABSTRACT

Free stall comfort is reflected in various indicators, including the ability for dairy cattle to display unhindered posture transition movements in the cubicles. To ensure farm animal welfare, it is instrumental for the farm management to be able to continuously monitor occurrences of abnormal motions. Advances in computer vision have enabled accurate kinematic measurements in several fields such as human, equine and bovine biomechanics. An important step upstream to measuring displacement during posture transitions is to determine that the behavior is accurately detected. In this study, we propose a framework for detecting lying to standing posture transitions from 3D pose estimation data. A multi-view computer vision system recorded posture transitions between Dec. 2021 and Apr. 2022 in a Swedish stall housing 183 individual cows. The output data consisted of the 3D coordinates of specific anatomical landmarks. Sensitivity of posture transition detection was 88.2% while precision reached 99.5%. Analyzing those transition movements, breakpoints detected the timestamp of onset of the rising motion, which was compared with that annotated by observers. Agreement between observers, measured by intra-class correlation, was 0.85 between 3 human observers and 0.81 when adding the automated detection. The intra-observer mean absolute difference in annotated timestamps ranged from 0.4s to 0.7s. The mean absolute difference between each observer and the automated detection ranged from 1.0s to 1.3s. There was a significant difference in annotated timestamp between all observer pairs but not between the observers and the automated detection, leading to the conclusion that the automated detection does not introduce a distinct bias. We conclude that the model is able to accurately detect the phenomenon of interest and that it is equitable to an observer.

INTRODUCTION

All cubicles in a dairy barn are usually identical, however a natural variability exists both in animal size relative to the cubicle (Dirksen et al., 2020), and in individual motion patterns and locomotor activity (Shepley et al., 2020). A factor of stall comfort, which affects lesion prevalence and lying-time, is how easy it is for the cow to get up and down in the cubicle (Zambelis et al., 2019). Ease of movement during posture transition was highlighted as an evaluation criteria for stall quality regarding cow comfort by Lidfors (1989), who noted that cows in cubicles were more regularly seen performing abnormal motions (such as sideways lunging or horse-like rising) than on pasture. Ceballos et al. (2004) analyzed the kinematics of posture transitions and found that cows used less longitudinal space when rising in a cubicle than on an open pack. Given the evidence for the link between restrictive movements and signs of reduced welfare (Beaver et al., 2021), the quality of posture transitions is included as an indicator in welfare assessment schemes such as Welfare Quality (Blokhuis et al., 2013).

Assessing ease of posture transition per-se, rather than indirect signs of reduced comfort such as hock lesions (Dirksen et al., 2020) or reduced lying time (Shewbridge Carter et al., 2021), is more challenging, and practical objective methods are needed (Brouwers et al., 2023). Visual observations noting the occurrence of abnormal behaviors are commonplace in farm management and welfare assessment schemes. Alternatively, ease of movement can be assessed quantitatively by measuring the displacement of anatomical landmarks.
throughout bouts of posture transition (Ceballos et al., 2004). Drawbacks exist for both approaches. The visual method relies on time-consuming, sporadic human observations. Although Zambelis et al. (2019) found excellent agreement between observers (kappa of 0.93 for getting-up movement ease) a degree of subjectivity always exists in visual scoring of animal movements (Chaplin & Munksgaard, 2001; Vasseur, 2017). The acquisition of 3D kinematics data by Ceballos et al. (2004) relied on fitting motion-capture reflectors on cows, requiring lengthy preparation and exposure of the equipment to damage. These limitations might be a reason behind the low sample size (n = 5 cows with at least 2 bouts per cow) in the latter study.

Considering the variability in cow sizes and kinematic profiles, and the need for objective methods to assess ease of movement, we propose a framework to detect lying-to-standing (LTS) posture transitions from 3D pose estimation data. As a step in validating the potential of this method, the aim of this study was to measure the performance of a feature extractor in detecting the onset of LTS posture transitions compared with the human eye.

**MATERIALS AND METHODS**

The study presented here was approved by the ethical committee Uppsala djurförsöksnämnd under approval 5.8.18–13069/2021. Consideration for the 3Rs in animal research regards the use of already collected, non-invasive, video material.

**LOCATION AN ANIMALS**

Recordings were obtained at the Swedish Livestock Research Centre’s dairy barn (Uppsala, Sweden). The herd comprises Swedish Holstein and Swedish Red cattle housed indoors with access to pasture 120 d a year, between May and September. Video was recorded on 30 separate days (midnight to midnight), sampled for convenience, between 2021 and 12–08 and 2022–04–28. As the barn is lit at all times, recordings were obtained at all times of day. An average of 51 cows were present simultaneously in the pen, with individuals being added and removed throughout, for a total of 183 different individuals having visited the pen during the study period. Seven RGB cameras (G3 Bullet, Ubiquiti) were placed around an area approximately 1/4 of the pen, located closest to the sorting gate to the milking robot, and oriented toward the rows of cubicles so that all cubicles in the study ward, including forward lunge room defined as the 60cm beyond the head rail, were visible by at least 2 cameras. The study ward comprised the 12 cubicles (Cubicle divider cc1800 with rigid head bar, Delaval) for which video coverage was optimal, out of 66 total in the pen. The cameras were installed on fixed metal rails, part of the barn’s infrastructure, between 2.8 and 3.6 m high. The location of each camera, as well as the stall layout are shown in Figure 1.

Cows had access to feeding troughs with *ad-libitum* mixed feed as well as 2 rotary brushes, and concentrate dispensed both at the milking robot and at concentrate dispensers. Passage through the milking robot’s sorting gate was compulsory for access to the feed. Milking was done by one milking robot (VMS V300, Delaval) which cows had access to on a voluntary basis. Cows were brought to the robot by farm staff if they had not been milked in over 12h.

**KEY-POINT ACQUISITION IN 3 DIMENSIONS**

This study used a 3D pose estimation software (Sony Multi-Camera System, Sony Nordic). The software estimates the 3D pose by finding cross-view correspondences across inferred 2D poses of the same object on synchronized views. It then creates a track for each object based on spatial continuity in the 3D location. The initial synchronization is achieved by reading the timestamp of each frame, and relating the first full-second transition for a common timestamp across all video recordings as the initial synchronized frame. The initial frame synchronization is provided as an input to the multi-camera system. Synchronization is maintained using the estimated time of arrival of each frame in the processing buffer.

The 2D object detector and pose estimator use convolutional neural networks to detect cows and specific anatomical landmarks on RGB images, in the form of
The landmarks used in this study were limited to the center-top of the poll, highest point at the withers, spine at the 13th thoracic vertebra, and the top of the sacrum taken immediately behind the uppermost part of the ilium (referred to respectively as head, withers, t13 and sacrum).

The output data consists of one key-point for each anatomical landmark with X, Y and Z coordinates for each object and given frame. Figure 2 shows the estimated 3D position of the key-points, linked to create a visual structure, for 2 objects during a LTS transition, as well as the video frames used to generate them.

**DETECTION OF POSTURE TRANSITIONS**

The recordings were sampled visually by one observer with the aim of finding 1000 sequences containing LTS transitions. When a cow was observed fully getting up from a lying positions, the timestamp was annotated, and a video sequence corresponding to a window of ±15 s around the annotated timestamp was extracted. In the final data set, an arbitrary 979 sequences were eventually identified. These sequences were then processed with the 3D pose estimation software.

When the cow rises, the line formed by linking the sacrum and t13 key-points increases its angle compared with the horizontal plane, as the cow’s back is at an angle with the ground. By calculating the difference between the sacrum height and withers height, and following this difference through time, we identified peaks corresponding to LTS motions. When a peak above 0.4 (in the coordinates’ arbitrary spatial reference system) was detected, the frame was considered to be within a potential rising motion. The mean withers Z position in the 120 frames located 330 frames after the peak was then compared with the mean withers Z position in the last 120 frames of the sequence. If the ratio of height difference after and before the peak was higher than 140%, the track was classified as a LTS motion. Figure 3 illustrates this by showing the vertical position of the key points. At 16 s, there is an important difference in the withers (orange) and sacrum (green) heights. This difference points toward a potential rising bout. Calculating the difference in withers position between the 5s and 27s marks, we determine that the animal has transitioned from a low, lying posture to a high, standing posture.

In these 979 sequences, this method initially detected 493 LTS motions for which the cow was tracked at each consecutive frame. For the remainder (486 sequences), the tracks was interrupted for several frames and the motion was captured in several separate tracks. Detec-
tions were stitched together if they fit the following criteria:

- The tracks are found in the same 30-s sequence.
- The second track starts after the first track vanishes, and within an interval of 30 frames.
- The Euclidian distance in the 3D pose estimator’s coordinate system between the last point in the vanishing track and the first point of the starting one is lower than 0.2.

There was no limit on the number of tracks appended together to form one single track, as long as the above conditions were fulfilled. The resulting stitched track was kept if it contained more than 700 frames, and discarded otherwise.

Using this method, an additional 370 rising sequences were detected by applying the height difference rule to stitched tracks, giving a total of 863 predicted positives. For the remaining 116 sequences, either the animal was not detected by the pose estimation software, the posture transition detector failed to identify the occurrence, or the motion was split between different tracks that were not relatable due to noise or an interruption across more than 30 frames. Visual inspection of the predicted LTS motions revealed 4 false positives.

Twenty-two true positives were discarded from the data set, because the posture transition was initiated before the start of the video snippet and thus not captured in its entirety.

**SIGNAL PROCESSING**

Each series of raw coordinates was processed to attenuate noise. A low pass filter with a cut-off frequency of 10Hz was applied, to remove high-frequency noise resulting from key-point jittering. This cut-off was chosen based on the recommendations by Hamäläinen et al. (2011) and Riaboff et al. (2020) for noise removal on animal activity data. The filter was applied separately to each key-point and its respective X, Y and Z coordinates’ time series. The filter was implemented in Python 3.9 using the function “butter” from the SciPy package (Virtanen et al., 2020). Figure 3 illustrates the filtered Z coordinates time series during a rising sequence.

From the processed signal, consisting of the coordinates of each key-point in 3 dimensions, we detected the timestamp at which the cow starts rising. Considering solely the kinematic features available through the 4 key-points, this is most clearly reflected by the change in the withers’ position, as rising on the elbows will cause the withers to rise upwards slightly, visible by an increase in the withers Z (vertical) coordinate. When doing so, the cow aligns its back along the length of the cubicle, which is reflected in a change of the withers’ Y coordinate (axis perpendicular to the cubicle’s length). Although, from a behavior perspective, there is more to the LTS transition than solely the withers’ movement, the system was blind to all but the position of 4 anatomical landmarks. The withers were chosen for the stability of the key-point (low jittering) and for their consistent motion pattern in the LTS transition across sequences. To detect the exact onset of rising motions, we used linearly penalized segmentation (Pelt), implemented the python library “Ruptures” (Truong et al., 2020). Pelt was applied to the bivariate series of the Y (lateral, perpendicular to the cubicles) and Z (height) positions of the withers to identify breakpoints in the time series. No restrictions were set on the number of breakpoints to be detected. A baseline height (Z coordinate) was calculated for each sequence as the median wither height in the sequence’s first 30 frames. The break points detected by Pelt were iterated through. If the median wither height in the 30 frames following the breakpoint was higher than the baseline, the breakpoint was then considered to be the start of the rising motion. If not, we iterated to the next breakpoint and applied the same logic.
Data processing, feature extraction and analyses were carried out in Python 3.9.3 using packages Numpy 1.21.5 (Harris et al., 2020) and Scipy 1.9.1. (Virtanen et al., 2020).

**VALIDATION EXPERIMENT**

To evaluate the performance of the tool in detecting the occurrence of LTS bouts, we compared the timestamps automatically detected to those annotated by 3 human observers, considered the gold standard for behavioral observations. Observers were provided with the following definition: “The cow is lying down and rises on its breastbone and elbows, which causes the withers to rise visibly above the rest of the back.” This definition is based on that of Lidforss (1989) but it adds the withers’ position as an indicator. The animals were seen to initiate the movement by centering their elbows under the body, this in turn causes the withers to rise slightly. This motion of the withers was used to determine the exact onset of the rising motion. The description was accompanied by illustrations taken from Schnitzer (1971) and Cermak (1988) as well as an ethogram describing the sequence of movements in the LTS transition, in which the movement to label was explicitly identified. This ethogram described the stages of the posture transition based on Lidforss (1989) and on Schnitzer (1971). Observers all received the same training, where the ethogram was explained and examples were showcased; they reviewed 5 videos of different cows rising, and agreed on the exact frame to label as the onset of the rising motion. These 5 videos were taken from the original data set and used solely for training observers.

The validation data set was sampled randomly from the 471 complete LTS sequences captured in a single track. In total, 60 unique LTS sequences were annotated by at least 1 observer. This number was determined a-priori as no prior data was available on observer variability in posture transition detection. These sequences were the original 30 s synchronized video snips from which the key-points were detected. The video was available to the observers from all 7 cameras used for key-point detection, plus one additional ceiling mounted camera. Observers were free to choose the camera offering the best view of the animal performing the bout. Every observer was provided with a total of 55 randomly selected video clips. Of these 55 sequences, 30 were common to all observers and 10 were unique to each observer (40 different sequences per observer). The remaining 15 sequences were randomly resampled from the prior 40 and re-annotated by the same observer, to measure intra-observer reliability.

All sequences were blinded, with a different label each time the sequence appeared.

**STATISTICAL ANALYSIS**

The mean absolute difference (MAD) in annotate timestamp was calculated between each observer to quantify intra-observer reliability as,

$$MAD(i) = \frac{1}{15} \sum_{s=1}^{15} |\Delta_s|,$$

where $i = 1, 2, 3$ (observers), $\Delta_s = t_{s,1} - t_{s,2}$ with $t_{s,1}$ and $t_{s,2}$ being the time stamp of the s:th sequence provided at 1st and 2nd assessment occasion, respectively, by observer i. Also, the inter-rater MAD was calculated as

$$MAD(i, j) = \frac{1}{15} \sum_{s=1}^{15} |\Delta_{s(i,j)}|,$$

where $\Delta_{s(i,j)} = |t_{s,i} - t_{s,j}|, \forall i = 1, 2, 3$, with $t_{s,i}$ being the time stamp of the s:th common sequence by i:th observer. Mean difference (MD), e.g., $MD(i) = \frac{1}{15} \sum_{s=1}^{15} \Delta_s$, were calculated and normality of $\Delta_s$ and $\Delta_{s(i,j)}$ was assessed visually on a qq-plot. MD and MAD indicate inter-observer systematic bias, and dispersion, respectively.

The following mixed effects models were fitted using statsmodels.formula.api. Mixedlm (Seabold & Perktold, 2010) in Python 3.9, to evaluate the observer effect and intra-class correlation (ICC) with or without the automated detection:

$$t_{s,i,r} = \beta_0 + \beta_1 I(i = 2) + \beta_2 I(i = 3) + u_i + \varepsilon_{s,i,r}, \quad [1]$$

$$t_{s,i,r} = \beta_0 + \beta_1 I(i = 2) + \beta_2 I(i = 3) + \beta_3 M + u_i + \varepsilon_{s,i,r}, \quad [2]$$

where $\beta_0$ is the (fixed) intercept, $u_i \sim N(0, \sigma_u^2)$ is a random sequence effect, $\beta_1$ and $\beta_2$ are fixed observer effects, $\beta_3$ is a fixed effect corresponding to the automated detection taken as an additional observer (referred to as the “Model” or $M$), and $\varepsilon_{s,i} \sim N(0, \sigma_e^2)$ is a (random) error term. The sequence number is indicated by the subscript $s$, $I_i$ are the observers and $r$ is the index for repeated sequences annotated 1 to 2 times by the same observer. The observer effects were tested using ANOVA. ICC

$$ICC = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}$$

as a measure of inter-observer agreement were calculated. A post-hoc pairwise $t$-test with Bonferroni correction for 6 tests was then computed to test the pairwise differences between observers. The annotated timestamps were not normalized because a 1 s difference between observers,
for example, has the same practical meaning in this context regardless of whether the annotation is done at the 4 s mark or the 12 s mark.

The performance of the algorithm was assessed in the same way, by treating the algorithm as an additional observer and seeing if it differed from the human observers. The differences were calculated between the algorithms’ detection (denote $T^d$) and the observer annotation, $T^o$. Bland-Altman plots were prepared for each observer pair $(T_i, T_j) = \left(\{t_{i,j}\}, \{t_{j,s}\}\right)$, and also comparing $T^d$ with $T^o$, with a view to checking for the absence of a pattern and points beyond 1.96 standard deviations. MAD(H, M), and MD(H, M) were calculated.

### RESULTS

A total of 836 rising bouts were detected out of 979 visually selected sequences equating to a sensitivity of 88.5% or a false negative rate of 11.5%. Four sequences

![Bland-Altman plots](image)

**Figure 4.** Bland-Altman plots comparing the timestamp of onset of cows’ rising motions annotated by human observers to that predicted by the model. 3D pose estimation provided the coordinates of cows’ anatomical landmarks. Detecting breakpoints in the key-point motion enabled to detect the onset of rising.

<table>
<thead>
<tr>
<th>Inter-observer agreement (MAD ± σ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observer 1</td>
</tr>
<tr>
<td>Model</td>
</tr>
<tr>
<td>Observer 1</td>
</tr>
<tr>
<td>Observer 2</td>
</tr>
</tbody>
</table>

**Table 1.** Inter observer agreement. The table contains the mean absolute difference (MAD) ± σ between the annotations of all pairs of observers (including the model). Note that pairs between observers calculate the MAD on 30 sequences, whereas pairs with the model include an additional 10 annotations, unique to each observer.
were wrongly classified as rising motions giving a precision of 99.5% or false positive rate of 0.5%.

Model (1) comparing only human observers, gave ICC = 0.85. There was a significant observer effect in predicting the annotated timestamps of LTS onset ($P < 0.001$) according to the ANOVA. When the model (2) was fitted to assess performance of the prediction, the ICC decreased to 0.81, remaining at a similarly satisfactory level of agreement. There was however no significant difference between the predicted timestamp (“model”) and each observer’s annotations according to the post-hoc pair-wise $t$-test with Bonferroni correction of the type-1 error at $\alpha = 0.0083$. There was a significant difference between all observer pairs: $P(T_1, T_2) = 0.0016$; $P(T_1, T_3) < 0.001$; $P(T_2, T_3) = 0.0018$.

Mean absolute differences $T_{s, MH}$ are summarized in Table 1. These values indicate good inter-observer agreement and good agreement between humans and machine. The magnitude of $T_{s, MH}$ is identical to that of $T_{s, MM}$, meaning that $T_M$ could be used in further research, as the model does not deviate from the observers more than they do from one-another. Figure 5 shows the timestamp annotated by each observer (including the model and repeat sequences) for each sequence.

Intra-observer reliability was assessed using the mean absolute difference in seconds, and consistency using the standard deviation ($\sigma$). Observer 1 had a MAD of $0.55 \pm 0.88$ s ($\mu \pm \sigma$). Observer 2 had $0.68 \pm 1.47$s, and observer 3 had $0.36 \pm 0.48$. Pooled standard error was $0.27$s. The standard deviation is preferred here to the standard error, to quantify the variability in the differences between and within-observers in annotated timestamp, independently of the number of samples. These results indicate very good intra-observer reliability, under 1 s on average.

Finally, we compared the annotations to the automated detections visually using the Bland-Altman plot in Figure 4. The upper-left plot shows most points to be centered around 0, without signs of consistent bias from the model. More importantly, the spread was similar when comparing observers to the algorithm and observers together.

**DISCUSSION**

The ICC values show a good agreement between automated model detection and human observers in detecting the onset of cows’ rising motions, according to previous research on the use of ICC as a reliability metric in animal motion scoring (Kaler et al., 2009). The ANOVA demonstrated a significant observer effect, strengthening the claim that observations of cows’ movements are prone to individual variations. The
post-hoc test showed a significant difference in annotated timestamp between all pairs of observers, but the difference between the model and the observers was not significant. We conclude from this that the model’s detection lies somewhere in-between the observers’ annotations. The mean difference of −0.006s between observers and the model, in Figure 4, and the proximity of the points to 0 show that there is no systematic bias introduced by the automated detection. This latter finding is also supported by Figure 5, showing the timestamp annotated by each observer at each sequence, in which there is no evidence of the detection being consistently divergent from human annotations, as the triangular points (model) are not systematically above or below the circular ones (observers). We also see that the predictions do not tend to be further from the annotations than the annotations are from each other.

This agreement is a crucial step in validating the capability of 3D computer vision to accurately identify this specific kinematic feature in bovine behavior. Notably, the findings suggest that the model’s performance does not considerably differ from human observers when compared with the variability among human observers. This suggests that the model does not introduce a distinct source of error in the detection process. While there are discrepancies between the model and human observations, the magnitude of these divergences is not meaningful in comparison to the overall duration of the LTS transition.

Some limitations are important to mention however. One such limitation is the likely over-representation of specific individuals. The animals were filmed in a limited area of the barn and we can expect a degree of site fidelity from the animal (Vázquez Diosdado et al., 2018) leading to some individuals being over-represented. As there was no individual detection, correcting for individual was not possible. It is also unlikely that all recorded bouts were spontaneous; some may have been triggered by human intervention or by the presence of agonistic individuals. Bout motivation could introduce changes in kinematic patterns and velocity, and potentially impact the accuracy of the automated detection.

Limitations also exists regarding external validity as the study was conducted in a single cubicle design, under a limited period of time, and using manually selected video sequences. This manual selection work upstream to automated processing is an important limitation which drove the high sensitivity and specificity. The same system should be tested on continuous recordings. To counterbalance this limitation however, the posture transition is an evident behavior, with a large difference in key-point height before and after, which would easily be captured even with noisy keypoints by simply following the height of the cow’s back.

The scope of this study was determined retrospectively; the decision to compare the automated detection to manual annotations was made after collecting the data and visually identifying LTS motions. The inclusion criteria were based on data quality and not experimental considerations. The exclusion of 22 longer bouts discarded important information with implications for the most vulnerable individuals when it comes to stall comfort, as a long pause during the posture transition is associated with adverse welfare outcomes (Zambelis et al., 2019).

The study’s gold standard was human observation, which is known to be variable across observers, due to individual subjectivity. Although there is a bias incorporated in the model, this bias is consistent across observations. The accuracy of the model could be improved by both altering the ethograms to make them more “machine-learnable” (Brouwers et al., 2023), and by diversifying the data. Importantly, although human observations are biased, humans are rarely “completely off,” especially when the phenomenon at hand, such as posture transition, is evident. Algorithms on the other hand sometimes produce unexpected results, and monitoring and understanding their occurrence is essential for practical application. For instance, a difference of 6s is found between the model and observer 2 in sequence 31 (Figure 5). Upon visual inspection of this sequence, the algorithm picked up on the onset of the adjustment movements, particularly lengthy in this very sequence, making up the initial part of the posture transition. The second observer on the other hand, noted the moment the fast rising motion occurred. This is not an error of either method but a misalignment in the interpretation of the behavior. Referring to the description of the behavior provided to the observers, and quoted in the methods section, the timestamp automatically detected is closer to the phenomenon of interest.

Most significant for this research, automated detection via computer vision offers an objective method for detecting specific motions, which is desirable for studies of behavior and motion patterns. Judging by the advances in equine kinematic research, markerless computer vision constitutes both a robust and practical data acquisition tool to measure the displacement of anatomical landmarks, offering similar accuracy to motion capture, albeit for specific motions (Lawin et al., 2023). Reliably identifying the motion of interest is only a step in the study of posture transition kinematics, which contain welfare indicators (Zambelis et al., 2019), the measure of which can be automated (Brouwers et al., 2023). Future studies using this technology aim at implementing individual recognition, which could contribute to a pool of sensor data at individual level. Yet, in the absence of individual identification,
this technology is still able to deliver meaningful information either at herd or at cubicle level. The automated detection through 3D computer vision could, after further validation, serve as a new gold standard for the task of detecting LTS transitions (and other movements) similarly to how interpreting accelerometer data has become standard in behavior classification of ruminants (Riafbof et al., 2022).

CONCLUSION

In summary, our results demonstrate good agreement between human observations and automated detection of cows’ rising motions. Notably, they indicate that the model introduces no more bias than human observers. This finding validates the use of multi-view 3D pose estimation for detecting the onset of rising motions in bovine behavior, albeit in the conditions of a single farm. Automating the task with computer vision presents an opportunity to scale up bovine kinematic measurements and behavior monitoring, and apply objective methods to their study.

Notes

The authors thank the Swedish Research Council (Formas) for funding this research, the personnel of the Swedish Dairy Research Center at Lövsta for their outstanding support, and Sony Nordic for their extensive collaboration. The authors declare that Sony Nordic has contributed to this research in kind and in staff hours. Sony Nordic provided the technology to generate the 3D pose. Conceptualization, study design, statistical analysis and presentation of the results were decided by researchers at the Swedish University of Agricultural Sciences. Sony Nordic contributed in drafting the methods section regarding key-point acquisition in 3 dimensions, and in revising the final manuscript. This study was not conducted with the purpose of supporting a commercial claim.

REFERENCES


ORCIDs

Adrien Kroese • https://orcid.org/0009-0001-5780-7345
Moudud Alam • https://orcid.org/0000-0002-3183-3756
Elin Hernlund • https://orcid.org/0000-0002-5769-3958
David Berthet • https://orcid.org/0009-0006-3106-0622
Lena-Mari Tamminen • https://orcid.org/0000-0001-6781-4533
Nils Fall • https://orcid.org/0000-0001-5597-2358
Niclas Högberg • https://orcid.org/0000-0002-2672-7924