Early and non-intrusive lameness detection in dairy cows using 3-dimensional video

K. ABDUL JABBAR,* M. F. Hansen,* M. L. Smith,* and L. N. Smith*

*Centre for Machine Vision, Bristol Robotics Laboratory, University of the West of England, Bristol BS16 1QY, United Kingdom
ABSTRACT

Lameness is a major issue in dairy herds and its early and automated detection offers animal welfare benefits together with high potential commercial savings for farmers. Current advancements in automated detection have not achieved a sensitive measure for classifying early lameness. A novel proxy for lameness using 3-dimensional (3D) depth video data to analyse the animal’s gait asymmetry is introduced. This dynamic proxy is derived from the height variations in the hip joint during walking. The video capture setup is completely covert and it facilitates an automated process. The animals are recorded using an overhead 3D depth camera as they walk freely in single file after the milking session. A 3D depth image of the cow’s body is used to automatically track key regions such as the hooks and the spine. The height movements are calculated from these regions to form the locomotion signals of this study, which are analysed using a Hilbert transform. Our results using a 1-5 locomotion scoring (LS) system on 22 Holstein Friesian dairy cows, a threshold could be identified between LS 1 and 2 (and above). This boundary is important as it represents the earliest point in time at which a cow is considered lame, and its early detection could improve intervention outcome thereby minimising losses and reducing animal suffering. Using a linear Support Vector Machine (SVM) binary classification model, the threshold achieved an accuracy of 95.7% with a 100% sensitivity (detecting lame cows) and 75% specificity (detecting non-lame cows).

Key words: 3D computer vision, early lameness detection, gait asymmetry, locomotion analysis
1. Introduction

Lameness in dairy cows is acknowledged as being one of the most serious problems that affect an animal's welfare and thus, farm productivity (De Mol et al., 2013). Willshire & Bell (2009) reported that lameness in the UK’s national herd accounted for financial losses of up to £127.8 million in the year 2009. Regardless of its causes, early detection and prompt treatment minimises losses and reduces animal suffering (Cha, Hertl, Bar, & Gröhn, 2010; Leach, Tisdall, Bell, Main, & Green, 2012). Until now, measurement and analysis of weight distribution or walking pattern as the animal walks on force plates or the use of body sensors (accelerometers) are the most established conventional gait analysis methods. However, due to high expense, implementation complexity (Chapinal, de Pasillé, Rushen, & Wagner, 2010; Maertens et al., 2011) and high vulnerability to damage and loss of the recording equipment while collecting the data; such systems have never been implemented on a large scale, or on a regular basis, in dairy farming. Automated vision based methods for lameness detection are in their infancy and are based almost entirely on a single static measurable trait (i.e. estimating the animal’s back curvature/posture to predict gait soundness, Poursaberi, Bahr, Pluk, Van Nuffel, & Berckmans, 2010). However, although well established in the literature, there is unreliability in using back arching, as reported by Poursaberi et al. (2011), whereby some lame cows do not present an arched back, while conversely some healthy cows do show an arched back. Both Viazzi et al. (2014) and Van Hertem et al. (2014) developed automated lameness detection systems based on the measurements of the back arch, using 3-dimensional (3D) video. Although such systems are applicable for commercial farm implementations, no published research has shown a method that focused on early lameness classification that is suitable for daily use on a commercial farm, as we present here.

2.0 Method
Because many quadrupeds (including cows) walk in a symmetrical manner, gait symmetry has been the principal indicator in many conventional methods. However, it has been reported that gait asymmetry may occur for reasons other than lameness (e.g. udder fill; Flower, Sanderson, & Weary (2006) or a slippery floor causing the cows to take short and careful steps; van der Tol et al. (2005)). For similar reasons, a levelled concrete surface (with micro-grooves to improve the grip as the animals walk) was used while recording the data after the milking session. However, from a wider perspective, monitoring locomotion is generally useful for the farmers because it may reveal other well-being issues (Van Nuffel et al., 2015) - e.g. mastitis; Van Nuffel et al. (2015) or sole ulcers; Flower et al. (2006).

In a symmetrical (healthy) gait, the animal’s feet are expected to be on the ground for the same amount of time and the footfalls within each pair of legs are evenly spaced in time. As a consequence, the left and right side of the body perform the same motion half a stride out of phase (Hildebrand et al., 1985; Remy, Buffinton, & Siegwart, 2009). However, in the case of a lame animal, the limbs tend to exhibit a certain asymmetry as the animal walks, which could be used as an indicator for a certain lameness stage. In dairy cows it is known that lameness significantly worsens the vertical symmetry (i.e. symmetry of the weight distribution between the right and left legs) as the animals walk on force plates (Thorup et al., 2014). Thus, the contralateral limb movements of lame animals are expected to show asymmetry as the animal walks. However, prior investigations have mainly focused on measuring the kinematic differences of these limbs on force plates, which is -as mentioned earlier- a complex method to implement on commercial dairy farms. Instead, by using 3D video from the top of the herd, here we investigate the height movement variations of the hip joints to study gait asymmetry.

It is hypothesized that a dynamic measure over a full gait cycle, observing the regular movements of each footfall, will assist in detecting early stage lameness. Standard 2-
dimensional (2D) video imagery when used in this way presents numerous problems which are difficult to overcome (Van Hertem et al., 2014). These include segmentation of the foreground from the background, occlusions and sensitivity to lighting variance. Recent advances in acquisition technology have allowed deployment of cheap and accurate 3D sensors, capable of video recording, which helps overcome those issues associated with 2D capture, and assists in the extraction of robust features. By incorporating Hildebrand’s work on locomotion and results from force plate methods, a novel extrapolation from 3D video data was developed to extract motion in terms of height variation symmetry, thus, objectively analysing an animal's locomotion.

From an implementation perspective, dairy farmers tend to prefer any system offering the least possible interference in the daily routine of the herd. Farmers also prefer a capturing setup where minimal human involvement is required to achieve maximum accuracy and this points to the need for an automated mechanism. One of the major subjectivity concerns in many conventional and manual methods is the presence of a human observer, which is known to affect the cow’s behaviour (Breuer, Hemsworth, Barnett, Matthews, & Coleman, 2000; Grandin, 2010; Reader, Green, Kaler, Mason, & Green, 2011). The accuracy of the lameness scoring is highly contingent on the animal’s behaviour, which in cows is liable to variation in the presence of observers. Therefore, in order to be able to study pain-related behaviour in the most reliable manner; the data capturing system has to be completely covert (human involvement during the procedure should not be required). By using an overhead view (i.e. from above the herd), our capturing system is completely covert, thus, enabling objective results to be obtained. Our approach also facilitates full automation and provides a hardware configuration which is less prone to damage and the presence of complex and noisy image backgrounds.
The locomotion data presented here is an initial part of a large ongoing data collection project at Bridge Farm, Glastonbury, United Kingdom, where more than 200 Holstein Friesian dairy cows are housed. All cows were milked twice a day. A custom race has been built next to the milking parlour which forces the cows to walk unconstrained in single file underneath the 3D camera. This race was in regular use as an exit from the milking parlour for several months to allow the animals to adapt to the changes, before collecting the data. The data consists of 23 3D recorded sessions from 22 cows, using a standard depth-sensor camera (ASUS Xtion PRO LIVE, ASUSTeK Computer Inc., Taipei, Taiwan). All cows have visible brand numbers and are tagged with standard Half Duplex (HDX) electronic tags for identification purposes. A Radio Frequency Identification (RFID) reader (Agrident ASR700 Controller, Agrident B.V., Meterik, Limburg, Netherlands) was used to read the tags as the cows walked in the race. A single camera was used through-out the entire data collection to capture the animals from an overhead position. Both the camera and the RFID reader were connected to a computer (Windows 7, i5, 8GB RAM). As we are studying a sensitive lameness stage, it is important that we observe as many possible cycles of the locomotion’s resulting signals. Following several tests at different Field of Views (FOVs); the 3D data presented here is captured at a height of 3.69 m off the ground. This was the maximum height achieved to acquire as many footfalls as possible without causing heavy distortions in the depth data (pixel resolution at this setting is 3.6 mm × 3.6 mm). The horizontal FOV was around 6 m. This has allowed the capture of at least two full gait cycles i.e. eight footfalls on average. The average acquired frames for one cow’s locomotion was 70. This also means that we were able to perform the analysis as the cow’s body leaves the frame (i.e. when the hooks are still visible). The camera operated at 30 frames s⁻¹.
To provide conventional manual scoring, an experienced local observer has scored each cow using the locomotion/lameness score (LS) system provided by (Sprecher, Hostetler, & Kaneene, 1997).

3. Results & Discussion

As presented in Table 1, the animals were scored in an open field as they walked freely from the cow race. This was performed immediately after (~5-7 min) the evening milking session when the 3D recordings were made, in order to minimise any variations that might occur given a longer time frame (e.g. injury). At the time of scoring, two additional standard 2D digital video cameras (one looking to the side, the other looking at the rear of the animals) were used to assist with reviewing the manual locomotion scores and identifying the cows using the brand number. The observer watched the recorded 2D videos and gave a final score for each cow with a clear brand number. The data was organised manually; the desired (manually scored with a brand number) cows were located in the RFID logs, and the timestamps of these readings were then used to locate the cows in the 3D recorded data. Each cow used in this data has been scored at least three times over the period of three weeks (with the exception of the severely lame cows i.e. LS 4 and 5 in Table 1), from the 20th May 2015 to the 2nd June 2015. Because early lameness is being investigated, only cows that repeatedly received manual scores of either 1, 2 or 3 across the three sessions were used. This provides a reliable data-set of cows scored at LS 1, 2 and 3 that can be used confidently to establish a sensitive early lameness threshold. The scored cows were extracted from the recorded 3D data as separate ONI files (labelled with the unique brand number), each cow's locomotion represents a single ONI file which was then processed in MATLAB (R2015b, The MathWorks Inc., MA, USA).
The pre-processing steps of the 3D data involve subtracting the background (an image of the cow race when there is no cow present) and applying a height threshold to eliminate surrounding object pixels and discarding extraneous information by filtering-out the noisy areas from the subtracted depth image. The resulting image was then smoothed using a symmetric Gaussian low-pass filter to remove quantization artefacts in the raw image. This processed 3D image is used to extract the height measurements from key Regions of Interests (ROIs), to compare the changes in the 3D surface as the cow progresses under the camera. Our algorithm is able to extract high curvedness (convex) features of the animal’s hooks and spine from the processed 3D image, by applying the curvedness measure as first proposed by Koenderink & van Doorn (1992):

\[
C_{(x,y)} = \sqrt{\frac{\kappa_1^2 (x,y) + \kappa_2^2 (x,y)}{2}}
\]

(1)

\[
C_{\text{max}} = \max(C)
\]

(2)

\[
\bar{C}_{(x,y)} = \frac{C_{(x,y)}}{C_{\text{max}}}
\]

(3)

where \(\bar{C}\) is the curvedness measure of the 3D shape. It represents the normalised magnitude of the combined principle curvatures (\(\kappa_1 + \kappa_2\)). The principal curvatures (in differential geometry) are calculated from the Gaussian and mean curvatures of the surface. They correspond to the orthogonal axes which reflect a point on the object’s surface. By thresholding the curvedness, the most prominent convex features (which corresponds to peaks) are visible - as shown in Fig. 1. The scapula or shoulders are very difficult to extract at the current camera height. However, we found that the peaks were a reliable feature to extract the hooks in order to track the hind limb movements. These peaks are typically represented by a region of 10-20 pixels allowing the local maxima of this region to be located. For
increased robustness to noise, the algorithm calculated a weighted average using a 2D Gaussian convolution window over each thresholded region to find the pixel with the highest curvedness value. Thus, we are able to robustly locate the hooks’ ROIs by tracking the outermost peak points as the animal walks. Using this approach, it was found that the spine represents the largest connected object given in a binary converted image of the curvedness threshold. Figure 1 illustrates the image processing pipeline described above. This process was repeated for each frame in the data. An overall detection rate (number of successfully processed frames where all features were correctly tracked /all frames) of 85.7% on the first attempt for the automated features extraction algorithm, for both the hooks and the spine features. All frames were manually observed to ensure correct features extraction. An interactive tool for manual intervention allowed the correction of any obvious misdetections, in order to correct ROIs for accurate feature points. This test allowed us to identify some of the most common problems in our data (i.e. changes in the spine’s curvedness which leads to a separated spine ROI or the pins been identified as hooks when the whole body alignment changes). Upon modifying the algorithm, a better automatic performance is achieved for hooks and spine features (96.1% and 100%, respectively).

A dynamic measure of height changes for each ROI was applied by calculating the median and maximum variations. It was found that maximum height variations were more suitable for this analysis as they are more sensitive to small changes, especially in cows with early stage lameness. These measures are normalised by removing the global locomotion variations from the surface of the cow. A middle ROI (near the sacrum bone) was located between the right and left hook to remove the effect of the cow’s overall movement towards and away from the camera by subtracting the sacrum ROI variations from both hooks’ ROIs. The resulting signals were then filtered using a moving-average digital filter to remove noise (mainly due to the high distance of the camera position) and a sine wave was fitted (using a
least-squares cost function) to the mean in each estimated period. In a healthy cow, as shown in Fig. 2, the right-left locomotion signals may not start equally out of phase but shift to become equally out of phase for the right-left hooks (i.e. the movements of the right-left hind limbs) at a certain time in the locomotion, representing a full cycle of footfalls. This is mainly because the animals enter the FOV freely, i.e. the starting footfall (limb) is unknown and it varies across the data. Because of their lateral sequenced gait (hind-left, fore-right, hind-right, fore-left) as they walk, the phase difference given one full cycle between the out of phase maxima and minima peaks (from the locomotion signals) usefully indicates how symmetrical the height variations are. Thus, it is a key proxy that can be used to track, measure and rank the symmetry between the movement of the right and left hind limbs and to subsequently, establish distinguished patterns between locomotion scores. Because of the nature of these sinusoidal signals, i.e. single cycle sinusoids (mono-components), the Hilbert transform is a suitable technique to estimate the instantaneous varied frequency between right and left signals. This transform converts the locomotion signals from the time-domain into analytic signals in which the phase and magnitude of the original data can be analysed directly. Here the magnitude and phase will change in synchronization with the original sinusoidal signal and the differences between right and left can be calculated. Figure 2 shows the signal processing steps of this study, as described above. Figure 3 shows examples of various locomotion signals from our data for LS 1-3. The difference in the amplitude changes are noticeably in lame cows, indicating either hook has moved higher/lower as compared to the other. This supports the previous findings, that cows standing with discomfort in one limb, remove weight from that limb and shift it primarily to the contralateral limb (Neveux, Weary, Rushen, von Keyserlingk, & de Passillé, 2006), resulting in significantly higher height variations (maxima peak) in the contralateral limb as compared to the lame limb (minima
peak) in a given cycle. Thus, a smaller phase difference is observed in a lame as compared to a healthy cow.

The resulting locomotion signals of this study correlate well with the manual locomotion scores which are heavily reliant on the limb movements, even though the limbs themselves are not visible from the view point of the 3D camera. Subsequently, we were able to extract a novel proxy by measuring the resulting symmetry from the height movements as we are closely observing the dynamics of each hind limb across all frames, as the animals walk freely. This has allowed us to anticipate objective lameness trends at early stages.

The symmetry patterns derived from the phase difference of close locomotion scores i.e. LS 1, 2 and 3 are noticeably changing across the majority of the examined data. Our results in Table 1 show a clear difference in the overall mean phase difference of the right-left signals in LS 1, 2 and 3. Here lameness reduces the overall mean difference due to uneven peaks in the locomotion signals resulting from asymmetric height movements. However, in severe lameness scores, due to very limited data (only two cows in locomotion scores 4 and 5), although the mean differences sit within the early lameness threshold, they fall outside the trend observed in scores when more data and sessions are available. It is important to mention that collecting more data in LS 4 and 5 is very difficult.

<table>
<thead>
<tr>
<th>Manual locomotion scores</th>
<th>Algorithm scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cows ²</td>
<td>N³</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
However, this does not affect the main purpose of this study, as we are able to observe a sensitive trend for early lameness. A significant statistical difference is shown using one-way ANOVA between all five groups ($P<0.05$). Student t-tests (unpaired two-sample t-tests, given unknown variance) reveal a significant difference between the data in LS 1 and each other level, as shown in Table 1. The same test shows a significant difference for LS 1 vs LS 2 and 3 combined, LS1 vs all other levels ($P < 0.0004$, $P < 0.000009$ respectively). Thus, a sensitive pattern was observed in the mean phase differences as the lameness level increases. We suggest a threshold from this data at a mean phase difference of 0.09 (by subtracting the full standard deviation from the mean phase difference of LS 1). However, this could result in a small overlap between LS 1 and 2 which could be further refined given more data. At this early stage lameness threshold (i.e. LS 1 vs. all lameness levels), we used a supervised learning (liner SVM) classification model to assess the system’s sensitivity (100%), specificity (75%) and overall accuracy (95.7%). The sensitivity represents the ability to detect lame cows from LS 2 to 5, and the specificity represents the ability to detect the non-lame cows in LS 1. The binary classification model’s confusion matrix is shown in Table 2.

<table>
<thead>
<tr>
<th>True class</th>
<th>Predicted class</th>
<th>Lame</th>
<th>Healthy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lame</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Healthy</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Confusion matrix for the early lameness threshold for all cows using a linear SVM classification. This strict binary classification is established between LS 1 (Healthy) and LS 2, 3, 4 and 5 (Lame). An accuracy of 0.95 is achieved using this classification at a very sensitive lameness stage, n = 23.

4. Conclusions:

Preliminary results of a non-intrusive 3D video data capturing setup have been presented that allow regular daily data capture on a large scale in commercial dairy farms. Our algorithm is
able to detect lameness trends at early stages. The extracted novel proxy from the 3D data is a dynamic symmetry measure which reflects the locomotion soundness by tracking the movements of the spine and hind limbs. The presented results show patterns that enable us to distinguish between close locomotion scores; i.e. LS 1, 2 and 3 on 22 dairy cows. Based on these results, we are able to identify an early lameness threshold on a 1-5 scoring system. We believe that our study strides towards an accurate, automated and objective locomotion assessment without the need for human involvement. One of the major advantages of our system is that we are able to capture data after each milking session on a daily basis, thus, small developing lameness trends could be incorporated and detected potentially even before a human observer could. Future work will focus on improving the robustness of the algorithms using further captured data and by analysing the individual cow’s variation.
ACKNOWLEDGMENTS

This research is funded by the University of the West of England (Bristol, UK). Special thanks to the observer, E. F. Wright and the staff of Kingshay Farming and Conservation Ltd. and Bridge Farm (Glastonbury, UK); for their extensive help and input during data collection at the farm.
REFERENCES


Fig. 1 (greyscale)
Fig. 1 (online color)
Fig. 2

Maximum Depth Change (cm)

Maximum Normalised Depth Change (cm)

Maximum Normalised Depth Change (cm)

Amplitude

Frames
Fig. 3

Locomotion Score 1  
Locomotion Score 2  
Locomotion Score 3

Maximum Normalised Depth Change (cm)

Frames
Fig. 1. Automated 3D depth image processing pipeline and features extraction for the hooks and the spine from a single 3D cow image. The first image is a raw depth image from the camera in the race; followed by the same image with the background removed, height threshold applied and smoothed to prevent limiting any curvature information; followed by the curvedness data calculation image with high peaks shown; followed by a binary converted image of the curvedness threshold to track the spine; followed by the features (ROIs) selection image. The distinctly curved (highest convex regions i.e. spine, hooks and pins) are clearly visible. This data is used to extract the ROIs in each frame.

Fig. 2. Locomotion signals and their Hilbert transform derived from height variation measurements. This figure shows the signal processing steps in a descending order. The measurements are taken at 30 frames per second. The first figure represents raw maximum depth changes in cm in each ROI across all frames, right hook ROI (solid), left hook ROI (dashed), sacrum ROI (x-dotted); followed by normalized measurements for the right and left hooks ROIs after subtracting the sacrum ROI measurements from each hook ROI; followed by a filtered, smoothed sinusoidal fitted signals which represent the locomotion signals of this study; followed by the wrapped Hilbert transform (*-dotted) for the difference between the right hook ROI and the left hook ROI.

Fig. 3. Examples of filtered sinusoidal locomotion signals for three different lameness scores. Right hook ROI (solid) and left hook ROI (dashed). The left column represents cows with locomotion score 1 (healthy); the middle column represents cows with locomotion score 2; the right column represents cows with locomotion score 3. All locomotion scores presented in this figure are according to Sprecher et al. (1997) scoring system.