# Vertical back movement of cows during locomotion: Detecting lameness with a simple image processing technique

Ibrahim Akin, Yilmaz Kalkan and Yalcin Alper Ozturan

## SUPPLEMENTARY FILE

### Supplementary materials and methods

The images were processed in three main steps:

i: Image pre-processing; ii: Image processing; iii: Decision.

*i. Image pre-processing:* The images were processed by using software (Matlab R2018a, MathWorks Inc). In this step, stored images (Figure 1a) were prepared for calculations and decision process with the effort of clear, low-sized "black & white" (B&W) images with a minimum number of image errors. Firstly, the colored (RGB–Red-Green-Blue) images were clipped to get the desired region of interest (ROI, back arch of cows) (Figure 1b). Images were converted to 1080×1920×1 size, including only the red dimension, due to easy image extraction in front of the red background (Figure 1c). This 1080x1920x1-sized image was filtered using a 5x5 median filter to obtain a smoother image (Figure 1d). Then, the images were converted to a complete B&W image for the ROI extraction (Figure 1e; the cow is black, and the background is white). Another search was done on images to detect and correct the noise and pixel errors by finding the small and distinct white areas. Thus, the distinct white dots that artifacts may cause in the image were also converted to black to obtain a complete black cow image (Figure 1f).

*ii. Image Processing/Analysis:* Each complete black cow image was processed to obtain a clear back posture image of the cows. Therefore, to identify the ROI (back arch of cows), the head (Figure 2a, $L_h$) and tail (Figure 2a, $L_t$) parts were ignored in each image. Thus, the problems that arise from the movement of the head and tail are eliminated and the highest point of the cow's back was achieved (Figure 2b, $L_b$).

This kind of black and white image is being stored in 2D matrices including 0's (black) and 1's (white) only (Figure 3). Therefore, the highest point of the cow's back arch was calculated in the computer by a simple counting process of zeros in each column.

The number of black pixels was calculated by counting the number of 0’s in each column of each matrix. This number also corresponds to the height of each column as well. Then, the maximum of each column gives the maximum height of the overall image (cow back arch). This number was stored in memory and the same procedure was repeated until the end of the same cow’s movement was finished. For example, if we had N images of the same cow while it was walking, N-many numbers, which represent the maximum height of each image, were recorded in memory. These N-recorded data were stored for use in the decision step of ALDS.

*iii. Decision:* In the previous steps, the image was processed and the highest points in each image were stored. This procedure was applied to all cows and a data matrix is generated as follows:

$$\boldsymbol{C} = \begin{bmatrix} f_{1,1} & \cdots & f_{1,N_1} \\ \vdots & \ddots & \vdots \\ f_{M,1} & \cdots & f_{M,N_M} \end{bmatrix}$$

where M is the total number of cows (in our case M=63), $N_i$ is the number of images for the i$^{th}$ cow. Each row of this matrix was allocated for data of different cows as:

$$\boldsymbol{C}_i = \left[f_{i,1}, f_{i,2}, \ldots, f_{i,N_i}\right];\ i = 1,2, \ldots, M$$

If there was no cow inside the image, the maximum height (or the maximum number of black pixels) tended to be zero; and if a cow was visible in the image, the height was increased and changed from image to image. Similarly, when the cow was not invisible in the image, the height decreased and reached zero as expected. An example plot of stored $\boldsymbol{C_i}$’s for the same cow with LS=5 is given in Figure 4a.

The mean value of these curves depends on the cow’s size (height) and the length of the recorded data ($N_i$’s). Hence, the data had to be normalized for a correct decision-making procedure. In this step, thresholding and normalization were applied as well.

Normalization was applied for each cow separately as $\boldsymbol{D}_i = \boldsymbol{\alpha}_i \boldsymbol{C}_i$ $(i = 1,2, \ldots, M)$ where the $\alpha_i$ is the normalization coefficient for i$^{th}$ cow calculated as:

$$\alpha_i = \frac{\sqrt{\boldsymbol{C}_i \boldsymbol{C}_i^T}}{N_i};\ i = 1,2, \ldots, M$$

where ‘T’ represents vector-transpose operation, M is the total number of cows, $N_i$ is the number of images for the i[th] cow. Using recorded **C**i’s, $\alpha_i$ was calculated for the i[th] cow as:

$$\alpha_i = \frac{\sqrt{\boldsymbol{C}_i \boldsymbol{C}_i^T}}{N_i} = \frac{\left( [f_{i,1}, f_{i,2}, \dots, f_{i,N_i}] \begin{bmatrix} f_{i,1} \\ f_{i,2} \\ \vdots \\ f_{i,N_i} \end{bmatrix} \right)^{1/2}}{N_i} = \frac{\sqrt{\sum_{n=0}^{N_i} (f_{i,n})^2}}{N_i};\ i = 1,2,\dots,M$$

Hence, after normalization, the new data is:

$$\boldsymbol{D}_i = \frac{\sqrt{\sum_{n=0}^{N_i} (f_{i,n})^2}}{N_i} \boldsymbol{C}_i = \alpha_i [f_{i,1}, f_{i,2}, \dots, f_{i,N_i}];\ (i = 1,2,\dots,M)$$

Then thresholding was applied to eliminate non-desired regions. A threshold level ($\beta_i$) was chosen separately for each cow using mean and standard deviation (std) values of $\mathbf{D}_i$ as:

$$\beta_i = mean(\boldsymbol{D}_i) - \frac{std(\boldsymbol{D}_i)}{2};\ i = 1,2,\dots,M$$

$\boldsymbol{D}_i$ was a 1x$N_i$ size vector hence, it includes $N_i$ data. By using threshold value ($\beta_i$) a new vector ($\widehat{\boldsymbol{D}_i}$) was arranged by deleting some data in $\boldsymbol{D}_i$ ($\alpha_i f_{i,n};\ i = 1,2..,M\ and\ n = 1,2,\dots,N_i$) using the algorithm as:

1. Take $\boldsymbol{D}_i = \alpha_i [f_{i,1}, f_{i,2}, \dots, f_{i,N_i}]$
2. Take n=1, m=1
3. If $\alpha_i f_{i,n} \geq \beta_i$
   a. $g_{i,k} = f_{i,n}$
   b. n=n+1, k=k+1
4. Else
   a. n=n+1
5. If n ≤ $N_i$, go to Step 3
6. Else
   a. $K_i$= k; (final length of the new vector)
   b. New vector $\widehat{\boldsymbol{D}_i}$ (size “1x$K_i$”) is ready
   c. Stop

The effect of this procedure can be seen in Figure 4b. After thresholding, a new vector ($g$) was generated from $f$ values as $\widehat{\boldsymbol{D}_i} = [g_{i,1}, g_{i,2}, \dots, g_{i,K_i}]$; $(i = 1,2, \dots, M), (K_i \leq N_i)$. Using the new vector $\widehat{\boldsymbol{D}_i}$ the locomotion score (LS) for i<sup>th</sup> cow was determined by the following decision algorithm:

1. Take $\widehat{\boldsymbol{D}_i} = [g_{i,1}, g_{i,2}, \dots, g_{i,K_i}]$
2. Calculate the mean of $\widehat{\boldsymbol{D}_i} \leftrightarrow \mu_i$
3. Calculate the standard deviation of $\widehat{\boldsymbol{D}_i} \leftrightarrow \sigma_i$
4. Calculate lameness factor $\theta_i$ as $\leftrightarrow \theta_i = \left(\frac{\sigma_i}{\mu_i}\right) * 100$
5. Decide the new computerized automatic lameness score (ALDS)

a. If $\theta_i < \rho_1 \rightarrow ALDS = 1$ (normal)

b. If $\rho_2 > \theta_i \geq \rho_1 \rightarrow ALDS = 2$

c. If $\rho_3 > \theta_i \geq \rho_2 \rightarrow ALDS = 3$

d. If $\rho_4 > \theta_i \geq \rho_3 \rightarrow ALDS = 4$

e. If $\theta_i \geq \rho_4 \rightarrow ALDS = 5$ (severely lame)

6. Stop

In step 4, the calculated lameness factor ($\theta_i$) is very small; hence it is scaled by 100 to convert it into a more interpretable value. Here, $\rho_1, \rho_2, \rho_3, \rho_4$ are threshold values for the LS procedure and must be chosen carefully for each farm. As explained in section 2.2, 12 cows were reserved to determine these threshold values for the calibration of the proposed system.

Initially, colored images were clipped to obtain the region of interest (ROI), which in this case is the back arch of cows. Images were converted to a single-channel black and white format (1080×1920×1) to facilitate easy extraction against a red background. A 5x5 median filter was applied to smooth the images. Further processing involved the removal of noise and pixel errors by identifying and correcting small white areas, ensuring clear black cow images. Each black cow image underwent processing to extract a clear back posture. The head and tail parts of the cow were ignored in each image to eliminate potential movement artifacts, enabling the identification of the highest point of the cow's back. The black and white images were stored in 2D matrices comprising only 0s (black) and 1s (white). The height of the cow's back arch was determined by counting the number of black pixels in each column, with the maximum height recorded for each image. These recorded heights were stored for subsequent analysis in the decision-making step. The processed images and recorded heights were used to generate a data matrix for all cows. Each cow's data was normalized to account for differences in size and recorded data length. Thresholding was applied using mean and standard deviation values to eliminate non-desired regions and create a new vector. This new vector represented the processed data after thresholding. Using this vector, the locomotion score (LS) for each cow was determined based on a decision algorithm. The calculated lameness factor was scaled by 100 to ensure interpretability. Threshold values for the LS procedure (ρ_1, ρ_2, ρ_3, ρ_4) were carefully chosen for the farm and used to assign the new computerized automatic lameness score (ALDS).

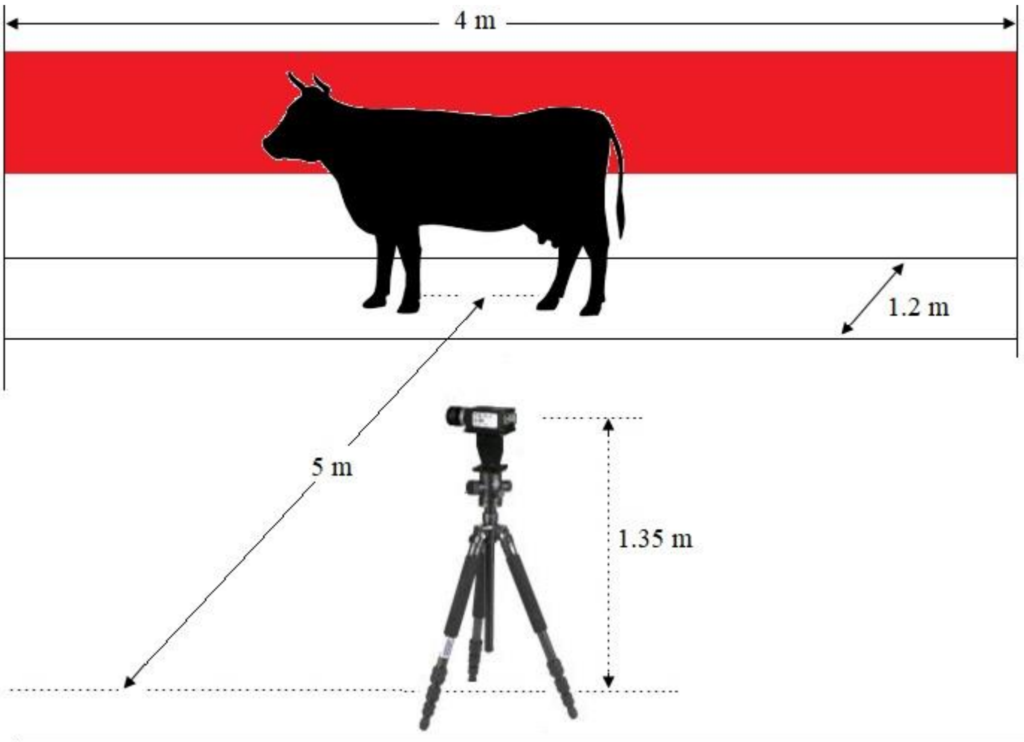
**Supplementary Figure S1:**
The walking path at the exit of the milking parlor.

**Supplementary Figure S2:**
Camera and walking path configuration.

**Supplementary Figure S3:**
Diagram illustrating the summary of the study's materials and methods.

**Supplementary Figure S1:**

**Supplementary Figure S2:**



**Supplementary Figure S3:**