Analysis of individual classification of lameness using automatic measurement of back posture in dairy cattle

S. Viazzi,* C. Bahr,* A. Schlageter-Tello,† T. Van Hertem,‡ C. E. B. Romanini,* A. Pluk,* I. Halachmi,‡ C. Lokhorst,† and D. Berckmans*¹

*Division of Measure, Model & Manage Bioresponses (M3-BIORES), KU Leuven, Kasteelpark Arenberg 30, B-3001 Leuven, Belgium

†Wageningen UR Livestock Research, PO Box 65, NL-8200AB Lelystad, the Netherlands

‡Institute of Agricultural Engineering, Agricultural Research Organization (ARO), Volcani Center, PO Box 6, Bet Dagan IL-50250, Israel

ABSTRACT

Currently, diagnosis of lameness at an early stage in dairy cows relies on visual observation by the farmer, which is time consuming and often omitted. Many studies have tried to develop automatic cow lameness detection systems. However, those studies apply thresholds to the whole population to detect whether or not an individual cow is lame. Therefore, the objective of this study was to develop and test an individualized version of the body movement pattern score, which uses back posture to classify lameness into 3 classes, and to compare both the population and the individual approach under farm conditions. In a data set of 223 videos from 90 cows, 76% of cows were correctly classified, with an 83% true positive rate and 22% false positive rate when using the population approach. A new data set, containing 105 videos of 8 cows that had moved through all 3 lameness classes, was used for an ANOVA on the 3 different classes, showing that body movement pattern scores differed significantly among cows. Moreover, the classification accuracy and the true positive rate increased by 10 percentage units up to 91%, and the false positive rate decreased by 4 percentage units down to 6% when based on an individual threshold compared with a population threshold.

Key words: lameness detection, dairy cattle, back arch, image processing

INTRODUCTION

Lameness, which can be defined as a deviation in gait to reduce pain (Scott, 1989), is the third most important health-related cause of economic loss in

the dairy industry after fertility and mastitis (Booth et al., 2004). The loss can be measured in terms of decreased milk yield (Green et al., 2002; Archer et al., 2010), reduced reproductive performance (Sprecher et al., 1997; Garbarino et al., 2004), increased culling risk (Barkema et al., 1994; Booth et al., 2004), and increased production costs (Cha et al., 2010). Lameness is also a major welfare problem (Brujinis et al., 2012) because it is a painful condition (Whay, 1997; Whay et al., 1998; Flower et al., 2008) that reduces mobility and prevents normal behavior (Singh et al., 1993). Early detection is therefore important for effective treatment and prevention.

Currently, diagnosis of lameness at an early stage in dairy cows relies on visual observation by the farmer, which is based on subjective scores called locomotion scores. Visual locomotion scoring is the most common method to obtain a prevalence rate (Flower and Weary, 2009). However, the visual scoring method is limited because it is sensitive to variations within and between observers (Thomsen et al., 2008; Channon et al., 2009). Visual locomotion scoring assesses alterations in the cow's gait and posture through qualitative parameters. Because differences exist in the walking patterns of each individual cow (Pluk et al., 2012), the effect of lameness on the measured variables cannot be considered identical for all cows. As are all living organisms, cows are complex, individually different, and time-variant dynamic (CITD) systems (Berckmans, 2008) and show lameness in different ways. The anatomy and posture of cows, for example, can affect a range of gait variables, such as stride length, back arch, position of the head, or speed (Flower and Weary, 2009).

Today, decision-making on commercial dairy farms is sensory-based and individual by measuring traits such as activity (Liu and Spahr, 1993), neck activity (Holman et al., 2011), individual feed intake (Halachmi et al., 1998), and milk properties (Tomaszewski, 1993). However, an individual approach has yet to be applied to detection of lameness.

Received June 4, 2012.

Accepted August 16, 2012.

¹Corresponding author: daniel.berckmans@biw.kuleuven.be

In the last 10 yr, a major research effort has been made to develop systems that automatically and objectively detect lameness. Different techniques have been applied, such as performance plus behavior analysis (Kramer et al., 2009), kinetic and kinematic analysis, and image processing. Kinematic analysis measures the geometry of movement, without considering the forces that cause the movement, and calculates different aspects of gait such as stride length, stance, and swing duration (Flower et al., 2005). Kinetic methods such as ground reaction force measurements (Rajkondawar et al., 2002) and a 4-balance weighting of each leg (Pastell et al., 2006) offer a means to assess lameness locomotion by evaluating load distribution to quantify the effects of lesions in the locomotor system. StepMetrix (BouMatic, Madison, WI), the only such product currently available commercially, is based on these kinetic principles; it measures the vertical ground reaction forces and position of the load exerted by cows while they are walking.

Computer vision techniques, on the other hand, offer the possibility to extract other variables such as step overlap (Pluk et al., 2010) and back posture (Poursaberi et al., 2010). They can be easily incorporated into existing farm system, are relatively inexpensive, and provide continuous information without the need for manipulating or placing a sensor on each animal.

However, all of these studies are based on a population-based model, that defines thresholds and criteria at the group level rather than focusing on the individual differences between animals.

Deviations in gait are only partly caused by pain associated with injuries on the hooves and limbs (Flower et al., 2008). According to Greenough et al. (1981), deviations can also be the result of factors such as damage in the locomotor system. Natural differences among cows, such as udder size, can further affect changes in walking gait and therefore result in individual differences among animals.

The hypothesis of this paper is that the variables to detect lameness are conditioned by large variation among individual cows and that, to be accurate, an automatic lameness detection system must account for this variation. Therefore, the hypothesis of this study is that an individual-based model that assigns thresholds for each animal would outperform a population-based model that defines a threshold for the whole herd.

The main objective of this study was to develop an image-based detection system that would allow us to apply the body movement pattern (**BMP**) on an individual level and that uses back posture to classify lameness (Poursaberi et al., 2011). The second objective was to compare both the population and individual approaches under farm conditions.

MATERIALS AND METHODS

Experimental Setup

Animals and Housing. The experimental data were gathered on a commercial dairy farm located in Yifat, Israel. The herd size of the farm was 951 lactating Holstein-Friesian cows with an average milk production of 11,500 kg/yr per cow. The cows were divided in 11 groups according to farm production requirements. The cows were milked 3 times a day in a 2 × 32 side-by-side parallel milking parlor. The last 2 milking groups, comprising 90 cows in total, were used for the experiment. After milking, which is the most suitable time to conduct gait assessment in dairy cattle (Flower et al., 2006), the cows were recorded while passing through a 4-m-wide, 7-m-long corridor with a concrete floor. To force the cows to create a single file, the alley was narrowed to 1.5 m wide with a mobile demountable fence of light tubular rails that were removed after every video session to avoid any further intrusion into the farm's routine and any interference with cow traffic (Figure 1).

Video Recordings. Videos were acquired from October 26 through December 21, 2011. During the first month, videos were recorded using a Canon 60D camera equipped with a Canon 17-85 IS USM lens (Canon Inc., Tokyo, Japan). During the second month, a Nikon D700 camera with a Nikon 28/105 AF Nikkor lens (Nikon Inc., Tokyo, Japan) was used. The cameras were positioned on a tripod at a height of 1.35 m perpendicular to the route of cow passage and 6.75 m away from the alley, which permitted a side view of the cows (Figure 2). Color videos were captured in QuickTime H.264 compressed format, with a frame rate of 25 fps at a resolution of 1,920 × 1,080 pixels. In total, the final database included 1,024 videos of cows with one expert score per video spread over a period of 14 d.

Scoring Methods. A veterinary expert on lameness visually scored all video recordings, using the 5-point locomotion score of Flower and Weary (2006). This scoring method is based on the observation of 5 gait attributes: flatness of the back, steadiness of head carriage, tracking up, asymmetry of the gait, and reluctance to bear weight. The criteria used to categorize cow locomotion are shown in Table 1. To increase the expert's reliability (Engel et al., 2003; O'Callaghan et al., 2003), the 5-point scoring scale was simplified into a 3-category scale. Scores 1 and 2 classified cows as not lame; score 3 classified cows as lame; and scores 4 and 5 classified cows as severely lame. The visual locomotion score was chosen as a gold standard (Dohoo et al., 2003) because of its practical application: it can be

Figure 1. Drawing of the experimental setup. The cows walked in a single file through the 7-m-long corridor that was narrowed to 1.5 m with a demountable fence of light tubular rails. (The corridor needed to be removed when the tractor passed over the cattle grid to bring food to the pens.) The camera was placed perpendicular to the route of cow passage. Color version available in the online PDF.

used onsite without the need for technical equipment, it can be applied easily to a large number of animals, and it is relatively fast compared with other methods that require, for example, the use of analgesic (Rushen et al., 2007) or the use of hoof testers equipped with a pressure gauge (Dyer et al., 2007).

Lameness Detection Algorithm

BMP. The BMP (Figure 3) was presented by Poursaberi et al. (2011). The highest point (**R**) in the total curvature of the animal's back was used as a starting point to calculate the BMP. Two ellipses were fitted to the left and right sides of point **R** and their orientations, θ_1 and θ_2 , were calculated. The front ellipse

represents the shape of the back around the shoulder, and the hind ellipse represents the shape of the back around the hip.

The intersection between the 2 lateral axes of both ellipses was determined and the resulting angle, θ_3 , was calculated; L_2 is the vertical distance between this intersection point and **R**. A line that connects the position of the cow's muzzle with the longitudinal axis of the front ellipse was drawn; L_1 is the vertical distance between the cow's muzzle and the longitudinal axis of the front ellipse. The angle θ_4 , between the horizontal line that connects L_1 with the longitudinal axis of the front ellipse and the line that connects the muzzle of the cow with the longitudinal axis of the front ellipse, was calculated.

Figure 2. Example of a video recording. A cow leaving the milking parlor was recorded while passing through the corridor. Color version available in the online PDF.

Table 1. The 5-point locomotion score of Flower and Weary (2006)

Score	Description	Criteria
1	Smooth and fluid movement	<ul style="list-style-type: none">• Flat back• Steady head carriage• Hind hooves land on or in front of fore hooves• Symmetrical gait
2	Imperfect locomotion but ability to move freely and not diminished	<ul style="list-style-type: none">• All legs bear weight equally• Flat or mildly arched back• Steady head carriage• Hind hooves do not track-up perfectly• Slightly asymmetric gait
3	Capable locomotion but ability to move freely is compromised	<ul style="list-style-type: none">• All legs bear weight equally• Arched back• Steady head carriage• Hind hooves do not track-up• Asymmetric gait
4	Ability to move freely is obviously diminished	<ul style="list-style-type: none">• Slight limp can be discerned• Obviously arched back• Head bobs slightly• Hind hooves do not track-up• Asymmetric gait• Animal is reluctant to bear weight on at least one limb but still uses that limb in locomotion
5	Ability to move is severely restricted and animal must be vigorously encouraged to move	<ul style="list-style-type: none">• Extremely arched back• Obvious head bob• Poor tracking-up with short strides• Asymmetric gait• Inability to bear weight on one or more limbs

The linear combination of these 4 angles (θ_1 , θ_2 , θ_3 , θ_4) and 2 distances (L_1 , L_2) was used to calculate the BMP:

$$\text{BMP} = w_1 \times \frac{\theta_2}{\theta_1} + w_2 \times \frac{\theta_4}{\theta_3} + w_3 \times \frac{L_1}{L_2},$$

where (weight) w_1 describes the relation between the front and hind ellipse, w_2 includes the relation between both ellipses and the position of the head, and w_3 relates the head position to the curvature of the back. To calculate a BMP value for each cow, the average of the BMP calculated every 5 frames was used.

Classifier. Decision tree learning (Quinlan, 1986) is a simple but powerful classifier, and it was chosen to classify the BMP output into the 3 lameness classes (not lame, lame, and severely lame) according to the given scores. Decision tree learning is widely used as a predictive model that maps observations about an item to conclusions about the item’s target value. In these tree structures, each branching node represents a choice between two or more alternatives. Every branching node is part of a path to a leaf node. In this research, the leaf node represented the particular classification of the lameness score, based on the given BMP values. Each of the 2 branching nodes has an associated weight that represents the threshold used to classify the BMP score into 1 of the 3 classes, in both the individual and population data sets. These

thresholds were obtained by maximizing the information gain of each data set.

Data Analysis

Data Sets. From the video database, a first data set (**D1**) contained 223 randomly selected videos of recorded cows and was used to evaluate the classifier based on a population threshold. A second data set (**D2**) was created that contained 105 videos of 8 cows (**D2_{Cow1}**) that, based on the expert scores, had moved through all 3 lameness classes within the recording period.

These data sets were used to test the hypothesis that the BMP for each cow is individually different and to evaluate the improvement of the classifier when an individual threshold is applied. The distribution of lameness classes based on the expert scores and used as gold standard is shown in Table 2.

Classification Evaluation. The software Weka 3.6.1 (Hall et al., 2009) was used to evaluate the classifier performance for each data set. This software calculated different metrics such as confusion matrix, true and false positive rates, precision, and accuracy.

The confusion matrix is a square matrix, **CM**, with the dimension n , corresponding to the number of target classes. It contains information about actual and predicted classifications of a classifier. In this research, the rows of the confusion matrix of order 3 showed the lameness classes scored by the expert and the columns

Table 2. Distribution of lameness classes in the data sets

Data set ¹	Not lame	Lame	Severely lame
D1	133	44	46
D2	33	37	35
D2 _{Cow1}	3	3	8
D2 _{Cow2}	5	2	5
D2 _{Cow3}	4	6	4
D2 _{Cow4}	4	5	3
D2 _{Cow5}	4	6	2
D2 _{Cow6}	7	4	2
D2 _{Cow7}	4	6	4
D2 _{Cow8}	2	5	7

¹Data set 1 contained 223 randomly selected videos of recorded cows and was used to evaluate the classifier based on a population threshold; data set 2 contained 105 videos of 8 cows (D2_{Cow1} to D2_{Cow8}) that, based on the expert scores, had moved through all 3 lameness classes within the recording period.

illustrated the lameness classes predicted by the classifier.

The true positive rate is the proportion of instances correctly classified as class c among all instances clas-

sified as that class by the expert. In the confusion matrix, the true positive rate is calculated by dividing the diagonal element by the sum of the relevant row:

$$\text{True positive rate} = \frac{CM_{cc}}{\sum_{k=1}^n CM_{ck}}.$$

The false positive rate is the proportion of instances classified as a particular class c , but belonging to another class, among all instances not belonging to that class. In the confusion matrix, the false positive rate is calculated by subtracting the column sum of the class by the diagonal element and dividing this result by the row sum of all other classes:

$$\text{False positive rate} = \frac{\sum_{j=1}^n CM_{jc} - CM_{cc}}{\sum_{j,k=1}^{n,n} CM_{jk} - \sum_{k=1}^n CM_{ck}},$$

Figure 3. Parameters θ_1 , θ_2 , θ_3 , θ_4 , L_1 , and L_2 as extracted from the cow and used to classify lameness. The highest point (R) in the total curvature of the animal's back was used as a starting point to calculate the body movement pattern. Color version available in the online PDF.

where j , k , and c are used as indices to represent the elements of the confusion matrix. The first subscript element represents the row index and the second one the column index.

The precision is the proportion of instances correctly classified as class c among all those that were classified as this particular class by the algorithm. In the confusion matrix, precision is calculated by dividing the diagonal element by the sum over the relevant column:

$$\text{Precision} = \frac{CM_{cc}}{\sum_{j=1}^n CM_{cj}}.$$

Accuracy is the number of correctly classified instances among all the instances.

Classification Evaluation: Group Analysis. The decision tree learning classifier was applied to data set D1, using a k -fold cross validation to avoid overfitting of the training data set. This method partitions the data set into k subsamples: a single subsample is retained as validation data, whereas the remaining samples are used as training data. The cross-validation process is repeated k times, with each of the subsamples used once as validation data. The number of folds was set to 10 because this value usually results in the best error estimation (Witten and Frank, 2005).

Classification Evaluation: Individual Analysis. The same classifier was trained for each D2_{Cow*i*} data set, using individual thresholds for each of the 8 cows. Therefore, 8 decision trees were obtained with Weka 3.6.1 (Hall et al., 2009). These decision trees had the same structure (2 branches and 3 leaves) but different thresholds to choose the path to the leaves and therefore to the classes of lameness. Because the number of instances per individual cow did not allow cross-validation, classification performance was evaluated on the training data set itself. The outcome can therefore only be interpreted as an upper-bound estimation of the classifier's performance. The accuracy and true and false positive rates of each individual cow data set were calculated. The weighted sum of these metrics was based on the number of instances per individual cow data set and used to compare the individual-based threshold to the population-based threshold. The comparison was conducted by evaluating the metrics on the training data set D2 itself, which is the total of all D2_{Cow*i*} data sets.

Statistical Analysis

The Statistics Toolbox Software (version 7.1) from Matlab (R2009a, MathWorks, Natick, MA) was used for statistical analysis. The Kruskal-Wallis test, applied

Table 3. Confusion matrix of the decision tree classifier using 10-fold cross validation on data set 1

Classified by the expert	Classified by the algorithm		
	Not lame	Lame	Severely lame
Not lame	119	6	8
Lame	19	11	14
Severely lame	6	1	39

in this study to check for significant differences ($P < 0.05$) between cows' BMP values, is a nonparametric test that assumes the BMP values come from a continuous distribution but not necessarily from a normal distribution (Kruskal and Wallis, 1952).

Following the assumption that walking patterns are different among individual animals and that the effect of lameness on the BMP is thus not identical for all cows, the analysis was carried out on an individual level. For each class (not lame, lame, and severely lame), each of the 8 cows was compared with the population composed by the other instances of not lame, lame, and severely lame cows.

RESULTS

Group Analysis

Table 3 and Table 4 illustrate the results of the decision tree classifier applied to data set D1 by using 10-fold cross-validation. As can be seen from the confusion matrix (Table 3), 119 not lame, 11 lame, and 39 severely lame instances were correctly classified. Overall, 169 out of 223 instances were correctly classified, an accuracy of 76%.

The results indicate that accuracy is only high when detecting the not lame instances of cows. The precision for cows classified as not lame was 83%, whereas it was only 61% for lame and 64% for severely lame cows (Table 4). Moreover, the true positive rate for lame was significantly lower (25%) compared with that for the other 2 classes, indicating that this class is the most difficult to detect.

Table 4. Result of the decision tree classifier using 10-fold cross validation on data set 1

Classified by the expert	True positive rate	False positive rate	Precision
Not lame	0.89	0.28	0.83
Lame	0.25	0.04	0.61
Severely lame	0.85	0.12	0.64
Weighted average	0.83	0.22	0.76

Table 5. Confusion matrix of the decision tree classifier using 10-fold cross validation and binary classification on data set 1

Classified by the expert	Classified by the algorithm	
	Not lame	Lame
Not lame	121	12
Lame (lame + severely lame)	22	68

When a binary classification (lame and not lame) was used instead of 3 classes, severely lame and lame cows were both classified as lame. As expected, the metrics of the classifier improved. As can be seen in Table 5 and Table 6, the precision of the binary classification was 85%, the true positive rate was 85%, the false positive rate was 18%, and the overall accuracy was 85%.

Individual Analysis

In Figure 4, the BMP scores are shown in box plots grouped per cow and class. Results show that the “not lame” and “lame” classes as well as the “lame” and “severely lame” classes often overlap. If the classes are compared for each cow, however, almost no overlap is seen, making the BMP suitable for classification.

Using the Kruskal-Wallis test, each cow was tested for each class of lameness compared with the respective class population, and the P -value was calculated. The results are given in Table 7. A significant difference ($P < 0.05$) was found for 2 cows in the “not lame” class (cow 3 and cow 4), for 2 cows in the “lame” class (cow 4 and cow 7), and for 3 cows in the “severely lame” class (cow 1, cow 4, and cow 5). In total, 5 cows of 8 were significantly different ($P < 0.05$) from the population. The Kruskal-Wallis test confirmed high variability between cows.

Comparing Population- and Individual-Based Classifications

In Table 8, the results of the classifier for each individual cow D2_{Cow*i*} data set are presented. In the bottom row, the total result is a weighted sum of all individual cows based on the number of instances per data set. To compare the classification using a threshold for the population to the classification using a threshold for the individual, the same classifier was trained and evaluated on data sets D1 and D2. The results are presented in Table 9; as can be seen, data set D2 was representative for data set D1.

When comparing Table 8 with Table 9, it can be seen that the individual threshold increased both the accuracy and the true positive rate by 10 percentage units to 91% while decreasing the false positive rate by 4 percentage units to 6%.

DISCUSSION

To date, back posture of a dairy cow is promising variable to be measured automatically by using visual techniques (Pluk et al., 2010; Poursaberi et al., 2010) because the accuracy based on optima thresholds exceeds 90% (Poursaberi et al., 2010). It is important to note that the BMP study is still in an early stage of development. The proposed algorithm is not yet fully automatic because it still requires a manual procedure to detect when the hooves are placed on the ground and because this procedure is only applicable for cows walking at a steady pace. Moreover, the algorithm cannot yet detect which limb is affected. As the BMP appears to offer a promising approach for studying lameness in dairy cattle, future studies will enhance the algorithm in terms of automation and dependencies.

In this research, an individualized BMP version was applied under farm conditions. The results obtained using a threshold for the whole population revealed an overall classification accuracy of 76%, a true positive rate of 83%, and a false positive rate of 22% when 3 classes were used. The accuracy and true positive rate increased to 85% and the false positive rate decreased to 18% when a binary classification was used.

According to the official specifications of StepMetrix (BouMatic), a commercially available product for lameness detection based on the measurement of ground reaction forces, the accuracy of that system was claimed to be 85%, without reporting sensitivity and specificity (BouMatic, 2012). However, a study performed to demonstrate the machine’s sensitivity and specificity for assessing hoof lesions in dairy cows concluded that the sensitivity of this system varied from 20.4 to 35.2%, and the specificity varied from 85.4 to 95.4% (Bicalho et al., 2007).

The BMP algorithm was improved when using an individual approach rather than a population-based one. A population-based approach, in fact, does not take into account the high variability among cows or other factors that can affect cow locomotion such as age, udder size, and environment (Boelling and Pollott, 1998). A population approach maximizes the accuracy of the average of the population only theoretically, because

Table 6. Results of the decision tree classifier using 10-fold cross validation and binary classification on data set 1

Classified by the expert	True positive rate	False positive rate	Precision
Not lame	0.91	0.24	0.85
Lame (lame + severely lame)	0.76	0.09	0.85
Weighted average	0.85	0.18	0.85

Figure 4. The body movement pattern (BMP) score was calculated and grouped per cow (for the 8 cows in data set 2) and gait score. In the figure, an example of the overlap between cow 4 (not lame) and cow 5 (lame) is shown.

almost no cow fulfills the specifications of a population average in reality. In other words, if one cow is not within a certain standard deviation of its scoring group, it will be misclassified.

The use of an individual approach can compensate for this inaccuracy by considering the high variability among cows. The hypothesis, that significant differences exist between individual cows' walking patterns, has already been formulated concerning the measurement of touches and release angles of fetlock joints in dairy cows (Pluk et al., 2012). In this study, it was demonstrated that significant differences also exist with respect to individual BMP. In fact, 62.5% of the

cows were significantly different ($P < 0.05$) from the average of their population. The classification based on an individual threshold increased the accuracy by 10 percentage units up to 91% compared with that based on the population threshold.

The BMP method is based on the idea that the cow shifts its weight toward the healthy limbs as soon as lameness causes discomfort or pain in one leg. As a result, the cow tends to increase the curvature of its back. Because the shifting of weight also causes an unusual weight distribution of the walking gait, the cow lowers its head in response (Poursaberi et al., 2011). This compensative walking behavior in the appearance of

Table 7. The *P*-values computed by Kruskal-Wallis test comparing the body movement pattern mean of one cow to the population

Cow ¹	Not lame	Lame	Severely lame
D2 _{Cow1}	0.4638	0.9515	0.0478*
D2 _{Cow2}	0.3695	0.4361	0.2912
D2 _{Cow3}	0.0461*	0.1448	0.7556
D2 _{Cow4}	0.0003*	0.0052*	0.0005*
D2 _{Cow5}	0.2212	0.7023	0.0495*
D2 _{Cow6}	0.0780	0.7301	0.7759
D2 _{Cow7}	0.1066	0.0352*	0.2885
D2 _{Cow8}	0.6150	0.1076	0.5447

¹Data set 2 contained 105 videos of 8 cows (D2_{Cow1} to D2_{Cow8}) that, based on the expert scores, had moved through all 3 lameness classes within the recording period.

**P* < 0.05.

the arched back has high inter-cow variability because of breed, morphology, age, and individual tolerance to pain (Flower and Weary, 2009). When using an individual approach, all of these influences are considered implicitly because the threshold adapts to each individual, making the system more precise and robust.

Changes in the lameness score for an individual cow can be detected by continuous monitoring of the BMP of the cow throughout its life. The system is able to calculate the parameters automatically every time the animal passes through the recording area after milking. Variation in these parameters can automatically trigger an early warning that informs and alerts the farmer about an individual animal's condition. Thus, lameness can be treated at an early stage, reducing treatment costs and improving the welfare of the animal.

CONCLUSIONS

The BMP algorithm for monitoring lameness in dairy cows was tested, under farm conditions, on 223 randomly selected cow videos and correctly classified

Table 8. Results of the decision tree classifier for each of the 8 cows in data set 2, using the data set for both training and evaluation

Cow ¹	True positive rate	False positive rate	Accuracy
D2 _{Cow1}	0.93	0.01	0.93
D2 _{Cow2}	0.89	0.22	0.89
D2 _{Cow3}	0.86	0.11	0.86
D2 _{Cow4}	0.91	0.05	0.91
D2 _{Cow5}	1.00	0	1.0
D2 _{Cow6}	1.00	0	1.0
D2 _{Cow7}	0.86	0.08	0.86
D2 _{Cow8}	0.88	0.05	0.88
Total	0.91	0.06	0.91

¹Data set 2 contained 105 videos of 8 cows (D2_{Cow1} to D2_{Cow8}) that, based on the expert scores, had moved through all 3 lameness classes within the recording period. The total result is the weighted sum based on the number of instances per data set.

Table 9. Results of the decision tree classifier on data sets D1 and D2 using the data sets for training and evaluation

Data set ¹	True positive rate	False positive rate	Accuracy
D2	0.81	0.10	0.81
D1	0.80	0.12	0.80

¹Data set 1 contained 223 randomly selected videos of recorded cows and was used to evaluate the classifier based on a population threshold; data set 2 contained 105 videos of 8 cows (D2_{Cow1} to D2_{Cow8}) that, based on the expert scores, had moved through all 3 lameness classes within the recording period.

76% of cow instances. The high variability between individual cows made it difficult to discriminate between contiguous classes. When a threshold was used for each individual cow, the overall accuracy and true positive rate increased, by 10 percentage units to 91% compared with the population-based model, and the false positive rate was decreased, by 4 percentage units to 6%. Moreover, the classification rate was high (>85%) in all 3 lameness classes. An individual-based threshold approach can be used to overcome individual variability in the development of an automatic lameness detection system based on image analysis.

ACKNOWLEDGMENTS

The authors thank farm personnel from Kibbutz Yifat (Israel) for their cooperation in the project. Thanks are due to Aaron Antler (Agricultural Research Organization, Bet Dagan, Israel) for building the experiment setup on the farm. The authors also thank Doron Bar, Dani Amram, and Roni Meyer from SCR Engineers Israel (Netanya) for their technical help. The work of Aaron Antler and that of SCR Engineers Israel were partly funded by the Israeli Agricultural Ministry Chief Scientist Fund (project numbers 459-4426-10; 459-4369-10; 459-4398-951). This study was part of the Marie Curie BioBusiness FP7-PEOPLE-ITN-2009-2014.

REFERENCES

- Archer, S. C., M. J. Green, and J. N. Huxley. 2010. Association between milk yield and serial locomotion score assessments in UK dairy cows. *J. Dairy Sci.* 93:4045–4053.
- Barkema, H. W., J. D. Westrik, K. A. S. Vankeulen, Y. H. Schukken, and A. Brand. 1994. The effects of lameness on reproductive performance, milk production and culling in Dutch dairy farms. *Prev. Vet. Med.* 20:249–259.
- Berckmans, D. 2008. Precision livestock farming (PLF)—Preface. *Comput. Electron. Agric.* 62:1.
- Bicalho, R. C., S. H. Cheong, G. Cramer, and C. L. Guard. 2007. Association between a visual and an automated locomotion score in lactating Holstein cows. *J. Dairy Sci.* 90:3294–3300.
- Boelling, D., and G. E. Pollott. 1998. Locomotion, lameness, hoof and leg traits in cattle: I. Phenotypic influences and relationships. *Livest. Prod. Sci.* 54:193–203.

- Booth, C. J., L. D. Warnick, Y. T. Grohn, D. O. Maizon, C. L. Guard, and D. Janssen. 2004. Effect of lameness on culling in dairy cows. *J. Dairy Sci.* 87:4115–4122.
- BouMatic. 2012. StepMetrix brochure. Accessed Aug. 14, 2012. http://www.boumatic.com/assets/files/euro/brochures/StepMetrix_EN.pdf.
- Brujinis, M. R. N., B. Beerda, H. Hogeveen, and E. N. Stassen. 2012. Assessing the welfare impact of foot disorders in dairy cattle by a modeling approach. *Animal* 6:962–970.
- Cha, E., J. A. Hertl, D. Bar, and Y. T. Groehn. 2010. The cost of different types of lameness in dairy cows calculated by dynamic programming. *Prev. Vet. Med.* 97:1–8.
- Channon, A. J., A. M. Walker, T. Pfau, I. M. Sheldon, and A. M. Wilson. 2009. Variability of Manson and Leaver locomotion scores assigned to dairy cows by different observers. *Vet. Rec.* 164:388–392.
- Dohoo, I., W. Martin, and H. Stryhn. 2003. Pages 85–120 in *Veterinary Epidemiologic Research*. 1st ed. AVC Inc., Charlottetown, PEI, Canada.
- Dyer, R. M., N. K. Neerchal, U. Tasch, Y. Wu, P. Dyer, and P. G. Rajkondawar. 2007. Objective determination of claw pain and its relationship to limb locomotion score in dairy cattle. *J. Dairy Sci.* 90:4592–4602.
- Engel, B., G. Bruin, G. Andre, and W. Buist. 2003. Assessment of observer performance in a subjective scoring system: Visual classification of the gait of cows. *J. Agric. Sci.* 140:317–333.
- Flower, F. C., D. J. Sanderson, and D. M. Weary. 2005. Hoof pathologies influence kinematic measures of dairy cow gait. *J. Dairy Sci.* 88:3166–3173.
- Flower, F. C., D. J. Sanderson, and D. M. Weary. 2006. Effects of milking on dairy cow gait. *J. Dairy Sci.* 89:2084–2089.
- Flower, F. C., M. Sedlbauer, E. Carter, M. A. G. von Keyserlingk, D. J. Sanderson, and D. M. Weary. 2008. Analgesics improve the gait of lame dairy cattle. *J. Dairy Sci.* 91:3010–3014.
- Flower, F. C., and D. M. Weary. 2006. Effect of hoof pathologies on subjective assessments of dairy cow gait. *J. Dairy Sci.* 89:139–146.
- Flower, F. C., and D. M. Weary. 2009. Gait assessment in dairy cattle. *Animal* 3:87–95.
- Garbarino, E. J., J. A. Hernandez, J. K. Shearer, C. A. Risco, and W. W. Thatcher. 2004. Effect of lameness on ovarian activity in postpartum Holstein cows. *J. Dairy Sci.* 87:4123–4131.
- Green, L. E., V. J. Hedges, Y. H. Schukken, R. W. Blowey, and A. J. Packington. 2002. The impact of clinical lameness on the milk yield of dairy cows. *J. Dairy Sci.* 85:2250–2256.
- Greenough, P., F. MacCallum, and A. Weaver. 1981. *Lameness in Cattle*. 2nd ed. J. B. Lippincott Co., Philadelphia, PA.
- Halachmi, I., Y. Edan, E. Maltz, U. M. Peiper, U. Moallem, and I. Brukental. 1998. A real-time control system for individual dairy cow food intake. *Comput. Electron. Agric.* 20:131–144.
- Hall, M., E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten. 2009. The Weka data mining software: An update. *SIGKDD Explor. Newsl.* 11:10–18.
- Holman, A., J. Thompson, J. E. Routly, J. Cameron, D. N. Jones, D. Grove-White, R. F. Smith, and H. Dobson. 2011. Comparison of oestrus detection methods in dairy cattle. *Vet. Rec.* 169:47.
- Kramer, E., D. Cavero, E. Stamer, and J. Krieter. 2009. Mastitis and lameness detection in dairy cows by application of fuzzy logic. *Livest. Sci.* 125:92–96.
- Kruskal, W., and A. Wallis. 1952. Use of ranks in one-criterion variance analysis. *J. Am. Stat. Assoc.* 47:583–621.
- Liu, X., and S. L. Spahr. 1993. Automated electronic activity measurement for detection of estrus in dairy cattle. *J. Dairy Sci.* 76:2906–2912.
- O'Callaghan, K. A., P. J. Cripps, D. Y. Downham, and R. D. Murray. 2003. Subjective and objective assessment of pain and discomfort due to lameness in dairy cattle. *Anim. Welf.* 12:605–610.
- Pastell, M., H. Takko, H. Grohn, M. Hautala, V. Poikalainen, J. Praks, I. Veermae, M. Kujala, and J. Ahokas. 2006. Assessing cows' welfare: Weighing the cow in a milking robot. *Biosystems Eng.* 93:81–87.
- Pluk, A., C. Bahr, T. Leroy, A. Poursaberi, X. Song, E. Vranken, W. Maertens, A. Van Nuffel, and D. Berckmans. 2010. Evaluation of step overlap as an automatic measure in dairy cow locomotion. *Trans. ASABE* 53:1305–1312.
- Pluk, A., C. Bahr, A. Poursaberi, W. Maertens, A. van Nuffel, and D. Berckmans. 2012. Automatic measurement of touch and release angles of the fetlock joint for lameness detection in dairy cattle using vision techniques. *J. Dairy Sci.* 95:1738–1748.
- Poursaberi, A., C. Bahr, A. Pluk, A. Van Nuffel, and D. Berckmans. 2010. Real-time automatic lameness detection based on back posture extraction in dairy cattle: Shape analysis of cow with image processing techniques. *Comput. Electron. Agric.* 74:110–119.
- Poursaberi, A., C. Bahr, A. Pluk, I. Veermae, E. Kokkin, V. i. Poikalainen, and D. Berckmans. 2011. Online lameness detection in dairy cattle using body movement pattern (BMP). Pages 732–736 in *Proc. 11th Int. Conf. Intelligent Systems Design and Applications (ISDA 2011)*.
- Quinlan, J. R. 1986. Induction of decision trees. *Mach. Learn.* 1:81–106.
- Rajkondawar, P. G., U. Tasch, A. M. Lefcourt, B. Erez, R. M. Dyer, and M. A. Varner. 2002. A system for identifying lameness in dairy cattle. *Appl. Eng. Agric.* 18:87–96.
- Rushen, J., E. Pombourcq, and A. M. de Passille. 2007. Validation of two measures of lameness in dairy cows. *Appl. Anim. Behav. Sci.* 106:173–177.
- Scott, G. B. 1989. Changes in limb loading with lameness for a number of Friesian cattle. *Br. Vet. J.* 145:28–38.
- Singh, S. S., W. R. Ward, K. Lautenbach, and R. D. Murray. 1993. Behavior of lame and normal dairy cows in cubicles and in a straw yard. *Vet. Rec.* 133:204–208.
- Sprecher, D. J., D. E. Hostetler, and J. B. Kaneene. 1997. A lameness scoring system that uses posture and gait to predict dairy cattle reproductive performance. *Theriogenology* 47:1179–1187.
- Thomsen, P. T., L. Munksgaard, and F. A. Togersen. 2008. Evaluation of a lameness scoring system for dairy cows. *J. Dairy Sci.* 91:119–126.
- Tomaszewski, M. A. 1993. Record-keeping systems and control of data flow and information retrieval to manage large high-producing herds. *J. Dairy Sci.* 76:3188–3194.
- Whay, H. R. 1997. Pain in the lame cow. *Ir. Vet. J.* 50:603–609.
- Whay, H. R., A. E. Waterman, A. J. F. Webster, and J. K. O'Brien. 1998. The influence of lesion type on the duration of hyperalgesia associated with hindlimb lameness in dairy cattle. *Vet. J.* 156:23–29.
- Witten, I. H., and E. Frank. 2005. Data mining: Practical machine learning tools and techniques. Page 286 in *Data Management Systems*. 2nd ed. Morgan Kaufmann, San Francisco, CA.