A novel automated system to acquire biometric and morphological measurements and predict body weight of pigs via 3D computer vision¹

Arthur F. A. Fernandes,^{†,✉} João R. R. Dórea,[†] Robert Fitzgerald,[‡] William Herring,[‡] and
Guilherme J. M. Rosa^{†,1,2}

[†]Department of Animal Sciences, University of Wisconsin-Madison, Madison, WI 53706; [‡]Genus PLC, 100 Bluegrass Commons Blvd., Ste. 2200, Hendersonville, TN 37075; and ¹Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison, Madison, WI 53706

ABSTRACT: Computer vision applications in livestock are appealing since they enable measurement of traits of interest without the need to directly interact with the animals. This allows the possibility of multiple measurements of traits of interest with minimal animal stress. In the current study, an automated computer vision system was devised and evaluated for extraction of features of interest, as body measurements and shape descriptors, and prediction of body weight in pigs. From the 655 pigs that had data collected 580 had more than 5 frames recorded and were used for development of the predictive models. The cross-validation for the models developed with data from nursery and finishing pigs presented an R^2 ranging from 0.86

(random selected image) to 0.94 (median of images truncated on the third quartile), whereas with the dataset without nursery pigs, the R^2 estimates ranged from 0.70 (random selected image) to 0.84 (median of images truncated on the third quartile). However, overall the mean absolute error was lower for the models fitted without data on nursery animals. From the body measures extracted from the image, body volume, area, and length were the most informative for prediction of body weight. The inclusion of the remaining body measurements (width and heights) or shape descriptors to the model promoted significant improvement of the predictions, whereas the further inclusion of sex and line effects were not significant.

Key words: depth image, image analysis, Microsoft Kinect, precision farming, swine

© The Author(s) 2018. Published by Oxford University Press on behalf of the American Society of Animal Science. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com.

J. Anim. Sci. 2019.97:496–508

doi: 10.1093/jas/sky418

INTRODUCTION

It is of great importance in pig production to define efficient management strategies in order to achieve optimum pig marketing weight and grade size in the most cost-efficient fashion. However, optimization of feeding and management in growing-finishing pig units is difficult (de Lange et al., 2001). The accurate prediction

of pig growth curve and understanding the factors that affect it are valuable tools to assist producers in improving the efficiency of pork production (de Lange et al., 2001; Gous et al., 2006). Frequent measures of pig body weight (BW) can enable precise estimation of individual animal growth curve and assess intergroup and intragroup variability. However, consecutive manual measurements of BW are not feasible in practice since it is labor-intensive, costly, and may increase animal stress leading to reduced animal performance and even animal loss (Grandin and Shivley, 2015; Faucitano and Goumon, 2018). To acquire frequent measures of BW, one alternative is via the use of automatic scales installed in each animal pen.

¹We gratefully acknowledge financial support from the Coordination for the improvement of High Education Personnel (CAPES), Brazil.

²Corresponding author: grosa@wisc.edu

Received August 10, 2018.

Accepted October 25, 2018.

However, these scales represent a physical intervention in the environment and the animals need an adaptation period (Kongsro, 2014). Moreover, such scales can be expensive. On the other hand, computer vision systems (CVS) offer a noninvasive approach, via the use of strategically positioned cameras. Thus, the use of CVS could be a powerful technology to assess real-time growth curves with minimal negative impact. Recently, CVS based on depth cameras have been proposed to predict pig BW (Kongsro, 2014; Condotta et al., 2018). However, these applications involve some level of manual processing either for selection of the best images out of the group of captured frames, image segmentation, features extraction, among others. Since manual processing is not feasible for a large-scale application, full automation is a key factor in order to deploy an optimized CVS at commercial settings. Therefore, the objectives of this study were as follows: 1) to develop an automated CVS for depth video processing, and extraction of phenotypes of interest and 2) to evaluate the ability of the developed models to predict pig BW at commercial farm conditions.

MATERIALS AND METHODS

The data set of animal BW and recordings of the weighing process were supplied by Pig Improvement Company (PIC, a Genus company, Hendersonville, TN). PIC follows rigorous animal-handling procedures that are in compliance with federal and institutional regulations regarding proper animal care practices (FASS, 2010).

Local, Animals, and Devices

The study was conducted in a pig multiplier facility in the United States. A total of 655 animals were randomly selected including males and females from 3 different terminal lines. From those pigs, 37 were nursery animals with average BW of 32 kg, and 618 were late finishing pigs with average BW of 120 kg. The video acquisition was performed using a Kinect V2 sensor (Microsoft, Redmond, WA) which has an RGB camera (resolution of 1920×1080 pixels), depth sensors (resolution of 512×424 pixels), and microphone array. The video emulation, processing, feature extraction, and training of predictive models were performed on a Windows 10 computer equipped with an Intel i7-7700k CPU (Intel, Santa Clara, CA) and an NVIDIA Quadro M4000 GPU (NVIDIA, Santa Clara, CA).

Data Acquisition

In this study, groups of pigs that were housed in the same pen were moved to a separate area where they were individually weighed. The difference from a traditional weighing process was the introduction of the CVS for image acquisition in the area preceding the scale and the fact that all pigs from a pen were weighed. The electronic scale used was an EziWeigh5i (Tru-Test, Mineral Wells, TX), which has a measured standard error of $\pm 1\%$ of the load. Before weighing the pigs, the scale was calibrated using blocks of known weight. The period each pig stayed in the prescale area was random and dictated by the time required to measure and annotate the weight of the previous pig in the line.

The recording process was initialized before the first pig of each group was weighed. Thus, each group had its own video with respective animal order and weight. The RFID sensor (radio-frequency identification) from each animal was used in order to obtain the pig identification (ID) and order in which they entered the prescale area. The pig ID was also manually annotated. This information was used to evaluate the CVS ability to recognize animals and keep track of the order information.

In a second moment, the videos were automatically processed for image segmentation and features extraction for each animal. To reproduce an actual implementation of the CVS, the video and the segmentation algorithm were processed in two parallel independent routines. The first was the video emulation of an actual Kinect camera via the Kinect for Windows SDK v2.0 (Microsoft, 2014). The second was the CVS containing the segmentation algorithm that retrieves and process frames from a connected Kinect camera. The CVS presented in this study corresponds to a library of custom codes written in MATLAB (Release 2017b) (The MathWorks, 2017). The connection between the CVS and the emulated video was made available by internal functions from the Kinect for Windows SDK (Microsoft, 2014). For the efficient establishment of this connection, the Kinect for Windows SDK custom C++ code was encapsulated within a MATLAB MEX function format following the directions in the Kin2 toolbox (Terven and Córdova-Esparza, 2016). Therefore, it was possible to connect the emulated Kinect device and retrieve the sensor intrinsic parameters, as the depth sensor effective focal length (f). To minimize the noise on the set of segmented images, the segmentation algorithm also classified each frame as acceptable or not based on animal posture (Figure 1). If the frame was accepted, then its features were extracted and saved for evaluation. The segmentation process was split into 3 steps: 1) first segmentation

Figure 1. Diagram of the steps comprising the computer vision algorithm devised and implemented: 1) First segmentation, **A** is an example of a frame without the pig, **B** is a frame with a pig, and **C** and **D** are their respective histograms. **A** is used to acquire the distance to the floor and then discarded, whereas **B** is passed to the next step. 2) Second segmentation: **A** discards current frame if the pig is touching the border and/or if it is not with a straight posture, **B** otherwise removes the pig head and tail from the image and passes the image to the next step. 3) Rotates and centralizes the image, and then performs the estimation of body traits and shape descriptors. The estimated features for the current pig frame are finally saved. After the current frame is fully processed or discarded, the system captures the next available frame from the connected Kinect device.

(recognize if there is a pig in the image); 2) second segmentation (segment the body of the animal in the image), and 3) feature extraction (extract the features of interest related to BW). Details on each of these steps are provided below.

First Segmentation

This step corresponds to the identification of the scene presented on the image. Basically, it differentiates whether the scene presented in the image contains a pig. To identify if there was a pig on a frame, the depth matrix was processed as an

intensity image. In this format, the intensity value in each pixel corresponds to the actual distance from the camera. Then, each pixel in the intensity image can be classified into background or foreground (objects in the scene). For this classification process, an adaptive threshold with a sensitivity of 0.4 was applied (Bradley and Roth, 2007). The largest foreground object identified was considered as the pig (Figure 2C). Before the first pig was present, the algorithm performed an estimation of the distance between camera and floor based on the mode of the group of pixels farthest from the camera:

$$fdist = mode(maxPeak), \quad (1)$$

where *fdist* is the estimated distance from the floor to the camera and *maxPeak* is the cluster of pixels above the 90th percentile of pixel distance (Figure 1C). After the identification of the pig on the image, the CVS passes the frame to the second segmentation. Otherwise, after the estimation of the distance from the camera to the floor, the program starts over with the next available frame.

Second Segmentation

After a pig was identified on the scene, a series of criteria were checked in order to accept the frame as valid. This checking is important in order to remove frames in which pigs were not well positioned or the image processing algorithm did not perform well as intended. The first criterion to select the frame was to identify if the pig was not connected to the border of the scene. This criterion was evaluated via the morphological opening of the logical image derived from the adaptive threshold (Figure 2D). In this step, every object connected to the border of the region of interest was removed from the logical image. For example, if a pig was pressed against a wall or door, the correspondent frame was discarded. Also, the pig body position was evaluated by an outline of the pig dorsal area and estimation of its spine position. For estimation of the spine position, only the back area that was between the shoulder and rump of the pig was used. Shoulder and rump were identified via an adaptation of Hough transform to identify round

objects (Hough, 1962; Atherton and Kerbyson, 1999). The Hough transform is the method initially developed for identification of lines in the picture plane of a grayscale image. The method consists of the identification of edge points and the assignment of an indicator value to each accumulator corresponding to a specific line position and orientation. The highest peaks in the accumulator array correspond to the strongest lines (Szelisk, 2011). The pig head and tail were identified and removed as the regions in front of the shoulder and after the rump (Figure 2E). Spine “backbones” were estimated as the median points across sections of the thoracic and lumbar area estimated from the dorsal view. A total of 11 points, approximately every other backbone on the thoracic (14 to 15 vertebrae) and lumbar (6 to 7 vertebrae) regions, were estimated. To ensure a more robust prediction of the spine curvature, a cubic curve was fitted over those points and used as the predicted back spine. A pig was considered as having a straight position if the spine curvature approximated a straight line by a linear equation with a coefficient of determination R^2 of 0.95 or higher. If a frame achieved these criteria, it was passed to the feature extraction step. Otherwise, the current frame was discarded and the program started over with the next available frame.

Feature Extraction

Finally, for any accepted frame, the algorithm rotates and centers the pig in the image and then

Figure 2. Image processing within the first and second segmentation. **A:** Depth frame from a Kinect sensor. **B:** Resulting image after adaptive threshold applied. **C:** Mask of the object selected as the pig. **D:** Cleaning and smoothing of the mask via image opening. **E:** Identification of shoulders and rump via Hough transform and removal of remaining head and tail as the regions adjacent to the shoulder and rump. **F:** Rotated and centralized pig with the delimited predicted spine (vertical line) and widths (horizontal lines).

performs the feature extraction and saved the original frame, as a black and white mask of the pig and the corresponding extracted features of interest. The set of extracted features was divided into 2 classes: body measurements and shape descriptors.

The body measurements were area, volume, length, 11 widths (W1 to W11), and 11 heights (H1 to H11) at equidistant locations across the back of the animal, from the shoulders to the rump. The length, width, and area of an object on an image can be estimated by the sum of its pixels. However, the estimation of the real measurement is also affected by the distance from the camera to the object (i.e., the same object will have a smaller area in the formed image if positioned farther from the camera). To estimate the real area of any pixel, the trigonometric principle of image magnification can be used:

$$m = f / d_o, \quad (2)$$

$$a_p = 1 / m, \quad (3)$$

$$a_o = \sum_{p=1}^n a_p, \quad (4)$$

$$v_o = \sum_{p=1}^n a_p * h_p, \quad (5)$$

where m is the magnification factor, f is the effective focal length that is intrinsic to the camera, and d_o is the distance the object is from the camera (equation 2). The area of a pixel a_p in the metric scale, on the object plane, can be approximated in this case from the magnification factor as shown in equation 3. Thus, the area of any object a_o can be estimated as the sum of the area of the pixels a_p that form its image (equation 4). The index p ($p = 1, \dots, n$) refers to the n pixels that form the image representation of the pig. The pixel heights h_p can be estimated by the difference between the distance to the camera for the pixel p in the pig dorsal area and the distance from the camera to the floor. Therefore, the animal volume v_o is the sum of the volumes specific to each pixel that constitutes the image representation of the pig (equation 5). It is important to point out that this volume is not the real volume of the animal body, but an apparent volume since only the top view of the animal was obtained, and then projected to the floor.

The shape descriptors used in the present study were 1) eccentricity, which is a measurement of roundness, and was estimated as the ratio between the foci and the major axis of the ellipsis that has the same second moments as the pig, and 2) the first 20 Fourier descriptors. Fourier descriptor is a class of global image descriptors normally used for shape analysis and image matching (Bowman et al., 2000;

Zhang and Lu, 2002; Burger and Burge, 2016). For robustness, each segmented pig image was rotated, centralized, and transformed to the polar plane representation previous to the estimation of its discrete Fourier transform in 1 dimension. Thus, the Fourier descriptors for the polar image can be estimated via a discrete Fourier transformation as follows:

$$FD_k = \sum_{j=0}^{l-1} X_j w_l^{jk}, \quad (6)$$

where FD_k ($k = 0, 1, \dots$) is each of the Fourier descriptors for the signal X of length l (the number of columns of the polar image) and $w_l = e^{-(2\pi i)/l}$, with $i = \sqrt{-1}$ the imaginary constant. Since FD_0 (the first Fourier descriptor) corresponds to the sum over the signal, i.e., the area of the object in the image, all Fourier descriptors were also divided by FD_0 for invariance on animal size.

After processing a given frame, the algorithm moves to the next available frame. To count the number of pigs and assign each video frame to the correct animal, the following criterion was used: If there was a sequence of frames without a pig or if there was a pig appearing in the opposite location from the previous recorded, the next valid frame was assigned to the next animal. Thus, for each video session, the number of pigs and the number of frames for each pig were tracked and frames with their features were automatically assigned to the respective pig and the dataset generated from each video session was saved for subsequent statistical analysis.

Statistical Analysis

The dataset of extracted image features and measured BWs was merged and manipulated for statistical analysis in R (R Core Team, 2016) via custom code from the authors. Since the number of frames generated per animal during the image acquisition can be large (up to 500 frames), strategies to reduce data dimension were evaluated. Each reduced dataset had just 1 data point for each variable for each animal. Before the frames selection to build the reduced dataset, outlier frames from each individual animal were removed. Eight strategies to reduce data dimension were evaluated as follows: 1) random selection of a single image from each pig (Random); 2) selection of the features from the image with highest estimated animal area (Max Area); 3) volume (Max Volume); 4) length (Max Length); 5) computing the average for every feature across all frames for each pig (Average); 6) computing the truncated average with exclusion on 20%

of the extreme data for each animal (Truncated Average); 7) computing the median for every feature across all frames for each pig (Median); and 8) computing the truncated median for every feature across of the subset at the third quartile of the extracted variables (Third Quartile). Notice that the truncated average and third-quartile methods are calculating averages across a reduced number of images. For an animal with few images, for example, 5 original images, this subset will have 3 images only, after the average or median is calculated. These reduced datasets were used to predict the BW through linear models using the following equation:

$$\mathbf{y} = \sum_{j=1}^k \beta_j \mathbf{x}_j + \mathbf{e} \quad (7)$$

where $\mathbf{y}$ is the vector of observations, β_j is the coefficient for the j predictor variable, $\mathbf{x}_j$ is the vector of observations for the j predictor variable, and $\mathbf{e}$ is the vector of residual effects. For all the reduced datasets, 10 permutations on a 5-fold cross-validation (CV) were used to access the stability of the estimations. Also, within each step of the CV, a stepwise regression approach with the Akaike Information Criterion (AIC) used as model selection criterion was applied to access the importance of the predictor variables. As such, 400 different model selection scenarios were considered (8 reduced datasets $\times$ 10 permutations $\times$ 5 CVs). The stepwise regression was implemented using the *stepAIC* function of the MASS package in R (Venables and Ripley, 2002; R Core Team, 2016).

After evaluating the methods for data reduction and frequency of appearance of variables, 5 final models to predict pig BW were evaluated on the best data reduction scenario. These final models

differ by the systematic inclusion of selected variables. The models included as follows: 1) Volume, Area, and Length (VAL); 2) VAL, Eccentricity, and selected Polar Fourier descriptors (VALS); 3) VAL and selected heights and widths (VALB); 4) A model including all variables in VALS and VALB (VALBS); and 5) VALBS plus sex and line effects (VALBSSL). To evaluate the prediction quality of these models, a 5-fold CV over 10 permutations was performed.

RESULTS

The Kinect sensor records depth data at a rate of 30 frames per second (fps), whereas the devised CVS performed at a rate of 8 fps. Therefore, if a pig stayed under the sensor for 30 s (which was the average time for weighing an animal), it had a total of 240 frames evaluated by the CVS (30 s $\times$ 8 fps). However, the number of recorded frames for each animal was heterogeneous (Figure 3) because the recording time for each pig was varied. It is important to point out that the data were collected in a commercial setting, without much control, and because of this some animals had just a few frames collected. The most common reasons for missing frames were as follows: 1) animals passing too quickly over the area; 2) animals that did not stay at an appropriated position; and 3) animals that passed together under the camera resulting in a scene occlusion and consequently fewer to no valid frames for those animals. For the purpose of model training, animals with less than 5 frames recorded were removed from the dataset. Thus, a total of 580 pigs had their data analyzed in the present study. From this total, 21 were nursery animals with the

Figure 3. Histogram of the number of frames (N) selected per pig. The black vertical line marks the number of pigs that had a count of 5 frames.

average BW of 31.5 kg and 559 finishing pigs within the average BW of 119.8 kg (Figure 4).

Analysis Including Nursery Data

A 5-fold CV was performed over 10 permutations for each reduced dataset; thus, a total of 400 model selections were performed (10 permutations on 5-fold CV and 8 different reduced datasets). The mean absolute error (MAE) and the coefficient of determination (R^2) for the selected models were estimated in the CV. A total of 265 different models were selected across the 400 different model selection scenarios (Figure 5). Also, most of the selected models had 14 variables, whereas the single model that appeared the most had 13 variables. Since there was a high correlation between body measures that are close to each other (e.g., H6 and H7), the several different models presented subtle switches between similar variables (Figure 6). It is worth to note that the model that contained all possible 25 variables was never selected and the most complex models had 19 variables. From the set of variables selected in each validation round for each of those reduced datasets, we can observe that body area, volume, and length were present in every model (Figure 5). Also, apart from those variables, the 3 models that appeared the most also presented in common 2 widths (W1 and W8) and 2 heights (H3 and H7).

For each step of the CV and permutation, the MAE and R^2 for the selected model were estimated (Figure 7). This allows for a robust comparison between the possible ways to reduce the dataset. For all different reduced datasets, the R^2 was high with averages ranging from 0.87 for a randomly selected image for each animal to 0.925 for the

average of the image variables at the third quartile for each animal. However, there is a separation between using the information from a single image against summarizing the information from several images of the same animal. Thus, the results for the reduced datasets that are originated from averages or median of the full sets of images for each animal performed better than the random selection of a single image, or the selection of the image that presented maximum area, length, or volume estimate for the given animal.

Analysis Without Nursery Animals

In this section, the nursery animals were removed from the dataset with the intent to evaluate if it is possible to improve the BW prediction of the heavier animals alone. Initially, a similar analysis to the previous section was conducted. That is, to access the difference in data manipulation previously to fitting a linear model for prediction of BW (Figure 8A and B). The results were similar to those from the data including nursery animals. The main difference was the reduction in the magnitude of the R^2 . Previously, the calculated R^2 values ranged from 0.88 for the data set composed of a random image for each animal to 0.92 for the truncated mean and third quartile datasets (Figure 7B). However, after removing the data from nursery animals, the R^2 ranged from 0.74 for the randomly selected image dataset to 0.81 for the truncated mean and third quartile (Figure 8B). However, this increase of R^2 when including the data from nursery animals did not translate to lower MAE (Figures 7A and 8A) since the median MAE for both scenarios is within the same intervals.

Figure 4. Histogram of live body weight (kg) distribution for nursery and finishing pigs, with respective averages and standard deviations (SD).

Figure 5. Evaluation of stepwise model selection over the cross-validations for the 8 different reduced dataset methods when including the data on nursery and finishing pigs, and body measures importance where W1 to W11 and H1 to H11 refer to the 11 widths and heights measured. (A) Number of times (Count) a specific model (group of body measurement) was selected. (B) Histogram of the number of different models (Count) per number of variables (body measurements) selected in the model (N). (C) The frequency that each body measurement was selected across all selected models.

It was also assessed the importance of using only body volume, area and length, the inclusion of the remaining body measurements (heights and widths), shape descriptors, sex, and line effects (Figures 8C and 7D). The selected widths were W1, W2, W5, W6, and W8, and the selected heights were H3, H4, H9, H10, and H11, whereas for the shape descriptors the remaining variables were eccentricity and the polar Fourier descriptors of order 1, 2, 4, 10, 11, 12, and 13. These variables were selected based on the variable appearance results for the datasets with and without nursery animals. There was a clear improvement of including body measurements or shape descriptors compared with using only the information of volume, area, and length. However, there was no visible improvement when including both shape and body measurements or

the effect of sex and line after accounting for all body measurement variables. A graph with predicted BW vs. measured BW on the third quartile dataset for the finishing pigs is presented for the model fitVALBS, which included both body measurements and shape descriptors (Figure 9). The correlation between the predicted and measured BW was of 0.88, and the MAE was of 3.6%, which corresponds to 4.36 kg for animals with an average BW of 120 kg.

DISCUSSION

The current study presents an autonomous framework for real-time video segmentation and extraction of image features for prediction of BW of pigs in commercial farms. To evaluate the effect

Figure 6. Correlations (Cor) between body weight (BW), volume, area, widths (W1 to W11), heights (H1 to H11), and length for nursery and finishing animals.

Figure 7. Results for the different models across cross-validation for the dataset with nursery and finishing pigs. (A) Box plots for mean absolute error (MAE), as a percentage of the average body weight. (B) The coefficient of determination (R^2) of the different models across cross-validation sets. The hinges represent the 1.5 interval interquartile (approximately 95% confidence interval), and red dots outlier results.

of including animals of different age classes, analyses with the full dataset (including younger animals) and with a subset with only finishing pigs were performed. The inclusion of data from nursery pigs increased the range of BW to a minimum of 27 kg and a maximum of 156 kg, whereas for the dataset without nursery pigs the range was from 83 to 156 kg. As a consequence, the R^2 from the models including younger animals were higher with an average of 0.92 (Figure 7), whereas for the models without the nursery pigs the average R^2 was of 0.80 (Figure 8). This higher R^2 estimate is actually an artifact due to the increase in the data range, with a consequent increase on the correlation between the

predictor's body measures and BW followed by an increase of the R^2 statistics. However, as presented in the results, this increase in R^2 is not reflected in the MAE. Actually, the best predictions for the test dataset in the CV including only finishing pigs had an MAE of 3% while the best models when including the nursery pigs had an MAE of 3.5%. Therefore, in a linear model setting, the inclusion of the wide range of the data across the growth period is forcing a model that is actually predicting the average weight at a given age. This may not be optimal for commercial pig operations, in which predictions of BW within age classes may be necessary.

Figure 8. Results for the different models across cross-validation for the dataset with only finishing pigs. The hinges represent the 1.5 interval interquartile (approximately 95% confidence interval) and red dots outlier results. Mean absolute error (MAE) as a percentage of the average body weight and coefficient of determination (R^2) for comparison between data reduction (A and B) and for inclusion of variables in the model (C and D). **fitVAL** stands for a model with volume, area, and length; **fitVALS** includes also shape descriptors; **fitVALB** includes also body measures; **fitVALBS** includes both shape descriptors and body measures; **fitVALBSSL** includes shape descriptors, body measures, sex, and line information.

Figure 9. Regression of body weight (BW) on the predicted BW for the model including volume, area, length, selected widths (W1, W2, W5, W6, and W8), heights (H3, H4, H9, H10, and H11), eccentricity and the polar Fourier descriptors of order 1, 2, 4, 10, 11, 12, and 13, the correlation (r) between selected and predicted and the mean absolute error (MAE) in percentage for the third quartile dataset including only finishing pigs.

In previous work, [Kongsro \(2014\)](#) demonstrated the application of 3D cameras for the estimation of pig BW with a reported residual mean square error (RMSE) of 4.8% (3.38 kg) and an R^2 of 0.99 for 71 animals with BW ranging from 30 to 140 kg. Recently, [Condotta et al. \(2018\)](#) performed a similar study, with more individuals (234 pigs), and reported a standard error of 3.13 kg and R^2 of 0.99 for animals of 10 to 125 kg. The findings in the current study for the dataset with nursery animals were within the previously reported range. In another study, [Pezzuolo et al. \(2018\)](#) evaluated not only the possibility of using a Kinect sensor to predict BW but also for measurement of biometric traits for pigs within 6 and 46 kg. In their study, the biometric traits measured from the images had high correlations with the manual measurements (from 0.77 for back height to 0.93 for heart girth). The main contribution of our work is the full automation of the whole process of image selection, processing, and

analysis. To the best of the researchers' knowledge, the only other work in the literature that presented results from a fully autonomous image analysis for prediction of pig BW was [Kashiha et al. \(2014\)](#). In their work, however, the images were obtained using a surveillance camera (grayscale image) positioned at a specific distance from the pen floor. Images from standard digital cameras have the disadvantage of producing images that are strongly influenced by environmental conditions. These light differences plus differences in background conditions make it difficult for practical applications of these cameras with their image segmentation algorithms. Therefore, the algorithms developed by [Kashiha et al. \(2014\)](#) are optimum for white animals on a dark floor, and thus, they are not robust for implementation in commercial farms with many different backgrounds, illumination conditions, and animal coat colors. On the other hand, the depth image is invariant with regard to the background color and suffers only from extreme light conditions that can increase the noise at the pixel level. Also, with direct measurement of the distance from the sensor to the floor, it is possible to retrieve information of animal measurements in the metric scale (i.e., transform from pixels to centimeters) without the need for setting the CVS at a fixed known distance.

Even though the presented approach of reducing the dataset was efficient in removing images of animals with a bad position, achieving similar prediction error to previous works, there is still variation within the set of images for a given animal, after the removal of outlier images ([Figure 10](#)). Possible reasons for this variation across images from the same animal may be due to the following: 1) There is natural variation in the images retrieved by the algorithm since the animals are moving; 2) Some of the retrieved images may be nonoptimum, which

can result in a cascade of reduced quality in segmentation and feature extraction; and 3) The internal variability of the Kinect depth sensor that has a resolution of 2 mm and a variation on the order of 6 mm ([Yang et al., 2015](#)). Thus, the image processing can still be further improved. Another possibility is of outlier animals in the group of finishing pigs. This is because within animals with similar BW, there are animals with a leaner body composition and others with more fat deposition. In fact, several factors can influence fat deposition, such as nutrition and genetics. For example, [Reyer et al. \(2017\)](#) studied the genetics of body fat mass of pigs from the same breeding line and observed that animals took from 124 to 167 d to reach 110 kg of BW, with a lean mass ranging from 56% to 66%, and subcutaneous fat thickness from 4.5 to 14 mm. These differences in body composition may imply in differences in shape and biometric measures for animals of the same weight class, which will increase the error in predicting BW with the current methods. Therefore, additional studies are needed in order to further improve the prediction quality of BW for finishing pigs.

In the present application, the possibility of using shape descriptors in the linear model in order to account for the variation in animal position and segmentation was evaluated. Body measurements as height and length among others were used previously as predictor variables for BW since they have a positive correlation with BW. But, shape descriptors as Fourier transforms are a more general group of image features that are used in order to classify images of different objects and shape recognition ([Bowman et al., 2000](#); [Zhang and Lu, 2002](#)) and are not necessarily linearly correlated with BW. The incorporation of shape descriptors to the linear model increased the predictive ability in

Figure 10. Averages (dots) and interquartile interval (vertical lines) for body length (the most variable biometric trait) and volume (the least variable biometric trait) measured across the frames for each animal vs. body weight for finishing animals.

comparison to the model containing only volume, area, and length information. However, the inclusion of shape descriptors to the model containing all selected body measurements did not show a significant improvement in MAE, even though there was an increase in R^2 (Figure 8C and D). Thus, other approaches need to be devised in order to account for this systematic variation due to animal positioning. One possibility is to use more rigorous thresholds on animal position. However, this can reduce the number of images retrieved per animal. Another approach would be to include a classification step after the image segmentation or the use of nonlinear models that can account for these small differences in animal posture. This may be a more efficient approach and may be able to account also for differences in body conformation across animals of the same age group.

In this study, an entire pen of pigs was moved to a weighing and managing area. Another practice, more common in commercial farms, is to position a scale inside the pig pen and weight just a fraction of the animals. This allows the farmer to have an estimate of the pen average weight and variability. Ultimately, the presented approach should be adapted to be used with a camera inside a pig pen, so all pigs have weight predictions, without the need of moving or stressing the animals. This is a topic for further research and development so that a full CVS applicative can be made available to the pig industry.

Currently, with the increasing size of farms and the need for increasing productivity, farmers are faced with the challenge of efficiently managing more animals. This raises the demand for automated tools for monitoring animal growth and health. Thus, CVS applications have been proposed in pig production for tracking of animal activity level (Tillett et al., 1997; Lind et al., 2005; Kashiha et al., 2013), behavior (Viazzi et al., 2014; Nasirahmadi et al., 2015; Lao et al., 2016), and gait and locomotion (Kongsro, 2013; Stavarakis et al., 2015). The methods presented here can be integrated with other applications, the most direct being the prediction of gait and walking kinematics (Stavarakis et al., 2015) that used similar sensors. Also, the proposed CVS can be easily modified to function in conjunction with animal tracking algorithms for the collection of data in barn condition without the need for human interaction. This provides the possibility for several measurements of pig BW and biometric measurements across time. With this information, precise prediction of the growth curve for each animal can be obtained, allowing the farmer to better define marketing

strategies (Tokach and Henry, 2008; Conte et al., 2012; Cadéro et al., 2018). This would be interesting since within the traditional management system, it is not practical to weigh pigs several times due to the labor involved and intensive stress caused to the animals. Moreover, the integration with other methods as described above can provide the farmer with detailed information not only on a pen (average) level but also for each pig. A successful application should be integrated also with an intuitive user interface that allows farmers to remotely access the information obtained. By having this information easily accessible, farmers will be able to have a live report of the farm performance status for timely management decisions, e.g., to better pinpoint animals with health issues.

CONCLUSION

A fully automated system using the 3D camera for online extraction of body measurements of interest and prediction of pig BW was presented. The main differences between the current approach and the previous studies are the absence of manual manipulation of the images and the inclusion of measurement of shape descriptors. Therefore, the presented approach moves towards the implementation of a CVS for the acquisition of biometric traits and BW on commercial farms. The current implementation is not able to identify animal ID and still requires the need for moving the animals. For a successful commercial application, there is still the need to incorporate means to correct track individual ID in a large group of animals without the need for human interaction with the pigs.

The truncated average and truncated median on the third quartile were the dataset reduction strategies that achieved lower MAE. Incorporation of sex and line on the model did not improve predictions compared with the models that accounted for shape descriptors and body measurements. Overall, this result shows that it is possible to achieve high predictive accuracy with only the information retrieved from the images.

LITERATURE CITED

- Atherton, T. J., and D. J. Kerbyson. 1999. Size invariant circle detection. *Image Vis. Comput.* 17:795–803. doi:10.1016/S0262-8856(98)00160-7
- Bowman, E. T., K. Soga, and T. W. Drummond. 2000. Particle shape characterisation using Fourier analysis. *CUED/D-Soils/TR315*. 1:20.
- Bradley, D., and G. Roth. 2007. Adaptive thresholding using the integral image. *J. Graph. Tools.* 12:13–21. doi:10.1080/2151237X.2007.10129236

- Burger, W., and M. J. Burge. 2016. Digital image processing: an algorithmic introduction using java. 2nd ed. In: D. Gries and F. B. Schneider, editors. Springer London, London. doi:10.1007/978-1-4471-6684-9
- Cadéro, A., A. Aubry, L. Brossard, J. Y. Dourmad, Y. Salaün, and F. Garcia-Launay. 2018. Modelling interactions between farmer practices and fattening pig performances with an individual-based model. *Animal* 12:1277–1286. doi:10.1017/S1751731117002920
- Condotta, I. C. F. S., T. M. Brown-Brandl, K. O. Silva-Miranda, and J. P. Stinn. 2018. Evaluation of a depth sensor for mass estimation of growing and finishing pigs. *Biosyst. Eng.* doi:10.1016/j.biosystemseng.2018.03.002
- Conte, S., P. G. Lawlor, N. O'Connell, and L. A. Boyle. 2012. Effect of split marketing on the welfare, performance, and carcass traits of finishing pigs. *J. Anim. Sci.* 90:373–380. doi:10.2527/jas.2010-3453
- FASS. 2010. Guide for the care and use of agricultural animals in research and teaching. 3rd ed. Fed. Anim. Sci. Soc., Champaign, IL.
- Faucitano, L., and S. Goumon. 2018. Transport of pigs to slaughter and associated handling. In: *Advances in pig welfare*. Woodhead Publishing, Cambridge, MA. p. 261–293. doi:10.1016/B978-0-08-101012-9.00009-5
- Gous, R., T. R. Morris, and C. Fisher. 2006. Mechanistic modelling in pig and poultry production. CABI, Wallingford, Oxfordshire, UK.
- Grandin, T., and C. Shivley. 2015. How farm animals react and perceive stressful situations such as handling, restraint, and transport. *Animals (Basel)*. 5:1233–1251. doi:10.3390/ani5040409
- Hough, P. V. C. 1962. Method and means for recognizing complex patterns. United States: N. p., 1962. <https://www.osti.gov/biblio/4746348>
- Kashiha, M. A., C. Bahr, S. Ott, C. P. H. Moons, T. A. Niewold, F. Tuytens, and D. Berckmans. 2013. Automatic monitoring of pig activity using image analysis. Springer, Cham. p. 555–563. doi: 10.1007/978-3-319-02895-8_50
- Kashiha, M. A., C. Bahr, S. Ott, C. P. H. Moons, T. A. Niewold, F. O. Ödberg, and D. Berckmans. 2014. Automatic weight estimation of individual pigs using image analysis. *Comput. Electron. Agric.* 107:38–44. doi:10.1016/j.compag.2014.06.003
- Kongsro, J. 2013. Development of a computer vision system to monitor pig locomotion. *Open J. Anim. Sci.* 3:254–260. doi:10.4236/ojas.2013.33038
- Kongsro, J. 2014. Estimation of pig weight using a Microsoft Kinect prototype imaging system. *Comput. Electron. Agric.* 109:32–35. doi:10.1016/j.compag.2014.08.008
- de Lange, C. F. M., B. J. Marty, S. Birkett, P. Morel, and B. Szkotnicki. 2001. Application of pig growth models in commercial pork production. *Can. J. Anim. Sci.* 81:1–8. doi:10.4141/A00-006
- Lao, F., T. Brown-Brandl, J. P. Stinn, K. Liu, G. Teng, and H. Xin. 2016. Automatic recognition of lactating sow behaviors through depth image processing. *Comput. Electron. Agric.* 125:56–62. doi:10.1016/j.compag.2016.04.026
- Lind, N. M., M. Vinther, R. P. Hemmingsen, and A. K. Hansen. 2005. Validation of a digital video tracking system for recording pig locomotor behaviour. *J. Neurosci. Methods* 143:123–132. doi:10.1016/j.jneumeth.2004.09.019
- Microsoft. 2014. Kinect SDK for Windows. Redmond, WA. <https://developer.microsoft.com/en-us/windows/kinect>
- Nasirahmadi, A., U. Richter, O. Hensel, S. Edwards, and B. Sturm. 2015. Using machine vision for investigation of changes in pig group lying patterns. *Comput. Electron. Agric.* 119:184–190. doi:10.1016/j.compag.2015.10.023
- Pezzuolo, A., M. Guarino, L. Sartori, L. A. González, and F. Marinello. 2018. On-barn pig weight estimation based on body measurements by a Kinect v1 depth camera. *Comput. Electron. Agric.* 148:29–36. doi:10.1016/J.COMPAG.2018.03.003
- R Core Team. 2016. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Reyer, H., P. F. Varley, E. Murani, S. Ponsuksili, and K. Wimmers. 2017. Genetics of body fat mass and related traits in a pig population selected for leanness. *Sci. Rep.* 7:9118. doi:10.1038/s41598-017-08961-4
- Stavarakakis, S., W. Li, J. H. Guy, G. Morgan, G. Ushaw, G. R. Johnson, and S. A. Edwards. 2015. Validity of the Microsoft Kinect sensor for assessment of normal walking patterns in pigs. *Comput. Electron. Agric.* 117:1–7. doi:10.1016/j.compag.2015.07.003
- Szelisk, R. 2011. Computer vision: algorithms and applications. In: D. Gries and F. B. Schneider, editors. Springer-Verlag, London.
- Terven, J. R., and D. M. Córdova-Esparza. 2016. Kin2. A Kinect 2 toolbox for MATLAB. *Sci. Comput. Program.* 130:97–106. doi:10.1016/j.scico.2016.05.009
- The MathWorks. 2017. MATLAB Release 2017b. The MathWorks, Natick, MA.
- Tillett, R. D., C. M. Onyango, and J. A. Marchant. 1997. Using model-based image processing to track animal movements. *Comput. Electron. Agric.* 17:249–261. doi:10.1016/S0168-1699(96)01308-7
- Tokach, M., and S. Henry. 2008. When, where and why: marketing decisions based on packer matrixes and pig flow. In: *Swine Profitability Conference*. Manhattan, KS. p. 49–62.
- Venables, W. N., and B. D. Ripley. 2002. *Modern applied statistics with S*. Springer-Verlag, New York.
- Viazzi, S., C. Bahr, T. Van Hertem, A. Schlageter-Tello, C. E. B. Romanini, I. Halachmi, C. Lokhorst, and D. Berckmans. 2014. Comparison of a three-dimensional and two-dimensional camera system for automated measurement of back posture in dairy cows. *Comput. Electron. Agric.* 100:139–147. doi:10.1016/j.compag.2013.11.005
- Yang, L., L. Zhang, H. Dong, A. Alelaiwi, and A. El Saddik. 2015. Evaluating and improving the depth accuracy of kinect for windows v2. *IEEE Sens. J.* 15:4275–4285. doi:10.1109/JSEN.2015.2416651
- Zhang, D., and G. Lu. 2002. Shape-based image retrieval using generic Fourier descriptor. *Signal process. Image Commun.* 17:825–848.