Computers and Electronics in Agriculture 153 (2018) 169–176

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

Estimating pig weights from images without constraint on posture and illumination

Kyungkoo Jun[a,*], Si Jung Kim[b], Hyun Wook Ji[c]

[a] Dept. of Embedded Systems Engineering, Incheon National University, South Korea
[b] Howard R Hughes College of Engineering, UNLV, United States
[c] Easy Bio Inc., South Korea

ARTICLE INFO

Keywords:
Pig
Weight estimation
Image processing
Neural network
Machine learning

ABSTRACT

This paper proposes an image based pig weight estimation method different from the previous works in three ways. The first difference is no constraint on pig posture and image capture environment, reducing the stress of the pigs. The second one is that the features obtained from 2D images are used without depending on 3D depth information. And the third is that our estimation model is constructed by exploiting the recent advances in machine learning. Besides the pig area size which had been a major feature parameter for the estimation, two new features, curvature and deviation, are introduced because those are related with the postures, thus being able to quantify the weight adjustment. A set of experiments are conducted to investigate how the performance is affected by the combination of the features and the neural network configurations. By using 477 training and 103 test images, the average estimation error of 3.15 kg was achieved, and the coefficient of determination of the model was $R^2 = 0.79$.

1. Introduction

Monitoring pig weights in a regular basis is one of the important tasks in pig raising farms. The changes in weights provide direct means to assess the health and growth state of the pigs. The weight is also a key factor to determine whether the pigs reach a state of maturity for market. However, the weight measuring work is labor-intensive and stressful for not only farm workers but also the pigs. In a worse case, when the farm workers drive the pigs to a weighing equipment, both the workers and the pigs can be exposed to injury. To facilitate this job, a narrow corridor is commonly used; the pigs are herded to walk through a path which is one-pig wide and has an underneath scaling device. However, this weighing corridor approach does not completely remove the pig stress.

Computer vision algorithms using 2D or 3D cameras have been widely applied to automate farming jobs such as monitoring animal feeding, location, condition, and aggressive or reproductive behaviors (Marchant et al., 1999; Azzaro et al., 2011; Tasdemir et al., 2011; Yilmaz et al., 2013; Kulikov et al., 2014; Lee et al., 2016; Spoliansky et al., 2016; Nasirahmadi et al., 2015, 2016, 2017). Even fish raising farms have introduced optical sensors and machine vision algorithms to manage the quality of fish products (Saberioon et al., 2017). Besides normal images, the use of ultrasound and thermal images have been examined to overcome the drawbacks of the cameras which are affected by ambient lighting (Stajnko et al., 2008; Duff et al., 2010; Halachmi et al., 2013). Typically, 3D image based methods have been studied for various applications because of the advantages from depth information (Menesatti et al., 2014; Salau et al., 2014; Weber et al., 2014; Kuzuhara et al., 2015; Vázquez-Arellano et al., 2016). To this end, the adoption of low cost 3D depth cameras such as Microsoft Kinect has been increasing in agricultural and livestock applications (Kawasue et al., 2013; Viazzi et al., 2014; Rosell-Polo et al., 2015).

To overcome the problems caused by the manual weighing, digital image based approaches were attempted to estimate or measure body features of various livestock (Costa et al., 2013; Pallottino et al., 2015; Porto et al., 2015; Mortensen et al., 2016; Oczak et al., 2014; Guo et al., 2017; Pezzuolo et al., 2018a). In the case of the pigs, images were captured from the top and sometimes from the side at the same time by using 2D or 3D depth cameras (Brandl and Jørgensen, 1996; Schofield et al., 1999; Wu et al., 2004; McFarlane et al., 2005; Kollis et al., 2007; Parsons et al., 2007; Wang et al., 2008; Kashiha et al., 2014; Shi et al., 2016; Pezzuolo et al., 2018b). The images were then processed by image processing algorithms in order to estimate body measures from extracted features. Depth images could produce point clouds which improved the accuracy of estimation and measurement.

Commercial automatic weighing systems such as Weight-Detect by

* Corresponding author.
E-mail address: kjun@inu.ac.kr (K. Jun).

https://doi.org/10.1016/j.compag.2018.08.006
Received 18 March 2018; Received in revised form 1 July 2018; Accepted 2 August 2018

PLF Agritech, eYeScan by Fancom, Pigwei by Ymaging, optiSCAN by Holscher Leuschner, Growth Sensor by GroStat, qscan by Innovent and WUGGL by WUGGL have been introduced (Vranken and Berckmans, 2017). The proprietary algorithms for the image processing separate pig areas from background, and determine shape and volume features, which are then used to calculate various body measures. The regression models correlate these measures with weight estimation.

The image based weight estimation consists of a set of sequential steps. The pig segmentation, which separates pig region from image background is, the first step. Hough transform (Hough, 1962) and a-priori knowledge about the pig shape were exploited to improve the segmentation results (Kashiha et al., 2014). It still had the limitation that it did not work on non-straight postures and was suitable for estimating the average weight rather than individual ones.

The accuracy of the weight estimation largely depends on the features that are used as input to the prediction models, which have been mostly the linear or non-linear regression models in previous works. For instance, the use of the boundary length of the pig's shape extracted from 2D images and the average distance of the pig pixels to the edges was examined (Wongsriworaphon et al., 2015). Despite the introduction of the new features, it still required the straight posture, the consistency of the illumination, and human intervention to mark the pig areas manually.

In addition to the optical sensors and the computer vision algorithms, a neural network approach was introduced to estimate pig weights (Wang et al., 2008). The network was trained to correlate pig image features with corresponding weights. Obtaining the training images was restrictive to the pigs. Since the images should be captured only when the pigs had straight posture, the pigs were forced to be kept in a narrow corridor. Also, the corridor was installed with a cover to prevent ambient sunlight from affecting the images. The network model was relatively simple compared with the recent complex and advanced deep learning models.

Rather than 2D images, 3D images were exploited to improve the estimation accuracy because they could provide depth information (Kongsro, 2014; Banhazi and Dunn, 2016; Shi et al., 2016; Guo et al., 2017, Pezzuolo et al., 2018b). Pig heights were estimated from the depth information of a stereo vision camera (Shi et al., 2016). The heights along with the pig shape information were then used in a linear regression model for the weight estimation. The pig heights could be estimated from the Kinect camera images and were used to supplement 2D area information (Kongsro, 2014). The pig volume was calculated from the depth information and used as one of the search keys in a database that lists the volume with the corresponding weights (Banhazi and Dunn, 2016). Point clouds, which were generated from 3D images, were used to measure body features (Guo et al., 2017; Pezzuolo et al., 2018b). The point clouds are more versatile than 2D images because the body curvature such as heart girth can be approximated.

Despite the introduction of the additional 3D information, the straight posture of pigs in the images was still required. And, the lighting condition has become more important because some infrared based depth cameras such as Kinect are susceptible to sunlight, limiting its application to indoor use. Hence, it is still necessary to manually select those images that satisfies the posture and the lighting requirements in order to estimate the weights.

The posture constraint that these image based methods have sometimes requires the use of stressful environment such as narrow space in order to keep the pig bodies straightened. However, the need for such narrow space is ironic considering that one of the motivations of the image based works is to remove the narrow corridor that gives the stress to the pigs. Considering the limitations from the posture and the indoor use of the depth cameras, the advantages of the image based works over the manual scaling diminish, although non-contact weighing is a significant benefit.

This paper proposes a novel image based method that has no posture constraint and depends only on the 2D features for the weight estimation. Pigs are kept in a roomy space and are free to have non-straight postures during the image capture. Our method can work by using 2D cameras, only if the image segmentation of the pigs from the background is successful. It makes our work more applicable to outdoor farm environments in which infrared 3D depth cameras have difficulties due to sunlight. These advantages enable us to weigh pigs at pens where they grow, reducing the stress and required efforts.

Those improvements are achieved by developing a novel set of image processing and feature retrieval methods along with the help of recent advances in machine learning. One of our contributions is the definition of two features that represent the shape and locality of pig segmentations in the images; both are vital information to the weight estimation. Those features along with sizes of the pig segmentations are given as inputs to our estimation model which are implemented as a fully connected neural network.

The paper is organized as follows. Section 2 describes our proposed work in detail. Involved image processing steps are explained and feature extraction is discussed, which is then followed by the design of the fully connected network. Section 3 presents the results from experiments which are different in the number of used features and the structure of the neural network, giving a few hints about how the different combinations affect the accuracy. Section 4 concludes the paper.

2. Image based weight estimation method

The overall flow of the proposed method is similar to that of the previous works. In short, features are retrieved from images and then given to a neural network as input to estimate weights. However, having no constraint on the posture, the image processing steps become challenging typically in terms of the feature extraction. In addition, the need for new features arises in order to describe posture, which in turn tends to increase the size of the neural network. Fortunately, recent advances in machine learning fields enable us to design and implement such expanded neural networks.

An outdoor environment in which pig images are captured is configured as follows. A camera is installed over a pen at a fixed location overlooking the center of the pen. Note that, although the installed camera was an infrared based depth camera, the depth information was ignored because sunlight distorted the depth unreliably. It will be discussed in detail in Section 3. Hence, the captured images are treated as 2D images. top view and gray; some of the images are shown in Fig. 1. There is one pig at a time inside the pen. The pen is larger than the narrow corridor, allowing pigs to change their posture freely without stress. As a result, the captured images contain various posture of pigs; the body shapes of some pigs are relatively straight while others twist, head positions and orientations are different. Also, the pig locations are not same.

Illumination of the environment is not consistent. Since the pen is located outdoor, even under a shade, it is not free from brightness change due to indirect sunlight. Even for the Kinect images, it seems that the ambient IR from sunlight influences the IR dot patterns which the Kinect actively emits to measure the depth, spoiling the measurement inconsistently. The captured images, hence, have different levels of luminance. Although such inconsistency of the images poses challenges for image processing, the image capturing in the outdoor pen is beneficial for pigs and farm workers as well; they do not need to herd the pigs into unfamiliar indoor places. A video that shows the environment is available (Pig weighing environment, (n.d.)) and the details of the images will be more discussed in the next section.

Given the captured images, three features are retrieved from each image. One is the size of a pig area, and the others are related with pig posture. Hence, segmentation that can separate pig areas from other background objects is needed. Before discussing the features in detail, how the segmentation distinguishes the pig areas from the background is explained.

The segmentation begins with image binarization which is then

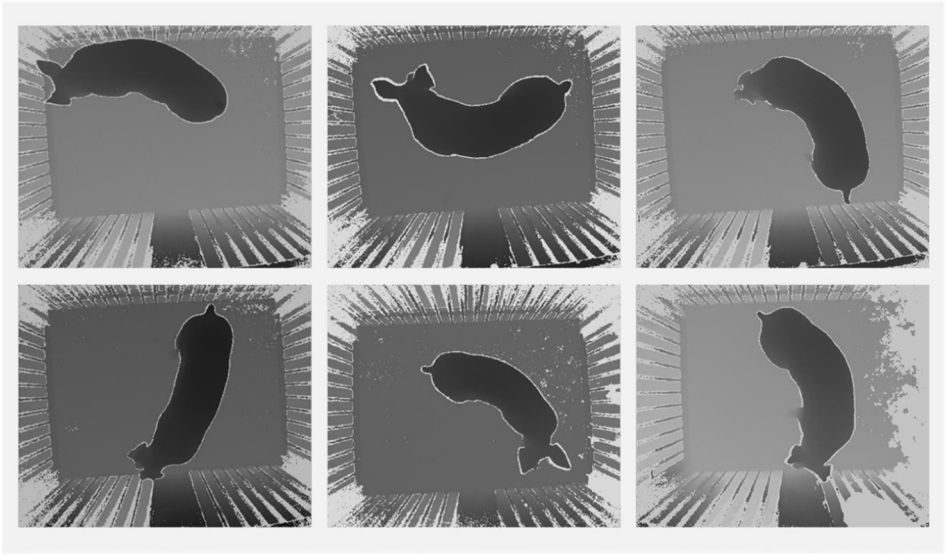

Fig. 1. The top view 2D images show that the pigs can have different postures moving freely and the light condition is not controlled.

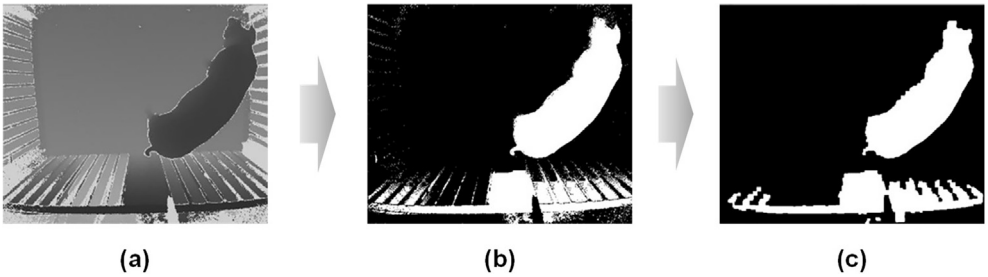

Fig. 2. Segmentation starts with binarization. (a) input image $\boldsymbol{I_G}$, (b) binarized image $\boldsymbol{I_B}$, (c) the result after noise removal.

followed by morphological operations. The binarization converts a gray input image I_G of Fig. 2(a) to a binary image I_B of Fig. 2(b) by using the Otsu threshold (Otsu, 1979). However, only the central region of the images, which excludes fences, is considered when calculating the threshold. It is because the fences that reflect sunlight irregularly, thus creating noise, distort the threshold calculation. In our work, the region is set to be the rectangular area at (100,100) with the width of 310 and the height of 170.

The binarization result I_B contains noises. Hence, a series of morphological operations are needed to remove them. Two operations are applied sequentially, an open followed by a close, both by using a 3×3 kernel, removing noises in the background and within the pig areas as well. The result is shown in Fig. 2(c), which shows the segmented pig area but still contains non-pig objects such as the fences.

Next step is then to remove the head parts of the pig segmentations. Since the pigs in the images have different head positions and orientations, the resulting head shapes, which are 2D, are diverse. For example, the shapes of raised heads are quite different from those of lowered heads. Thus, the removal of the head parts can reduce the dimensionality of the pig shapes. It is achieved by another morphological open operation but with a different kernel. A circular kernel of $s \times s$ size shown Fig. 3 is used. When applied to the 2D pig images, the open operation removes bulged parts such as heads and tails. The size of parts that are removed by the operation is determined by s; the larger it is, the more pixels of the bulged areas. Their neighboring regions are also removed which leaves the smaller pig segmentations.

It is hardly possible to choose a value of s that works for all cases because the pig segmentations have different shapes and sizes. Alternatively, instead of selecting a single value for s, a set of different values are used, having the same effect that the head-removal operation

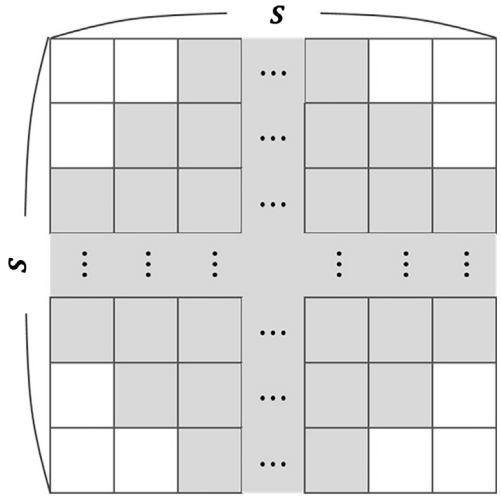

Fig. 3. Circular kernel with the size of $\boldsymbol{S \times S}$ is used by a morphological open operation to remove bulged parts of the pig shapes such as heads and tails.

is performed multiple times each with different values, resulting in multiple results. For example, using n different kernels, $k_1, k_2, \cdots, k_n$ each of which has the size of s_k, $(k = 1, \cdots, n)$, there are n results, I_M^k, $(k = 1, \cdots, n)$. Fig. 4 shows the case of $n = 4$ and $s_1 = 55$, $s_2 = 60$, $s_3 = 65$, $s_4 = 70$ where four I_M^k, $(k = 1, \cdots, 4)$ are created. Upon close look, the pig segmentation in I_M^4 is smaller than that of I_M^1 and such difference is attributed to $s_4 > s_1$.

The results after removing bulged parts, I_M^k, include not only the pig segmentations but also other areas, which do not belong to the pigs, like

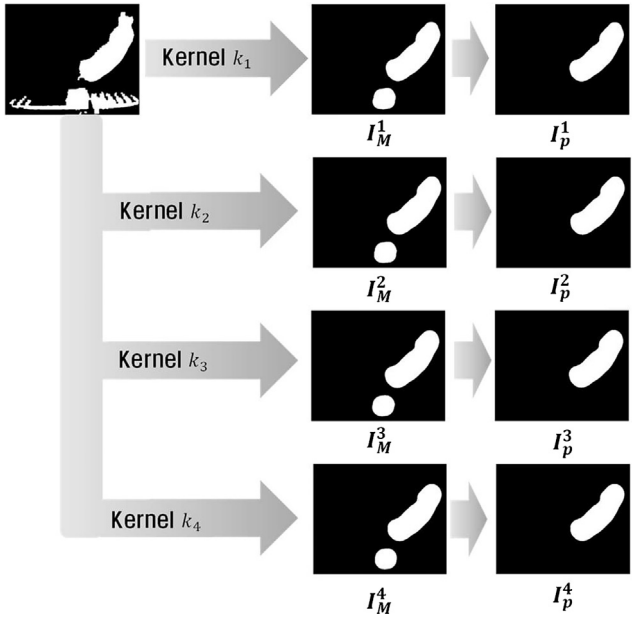

Fig. 4. A morphological open operation, being applied to the leftmost binary image, removes bulged parts such as the head and the tail from the pig shape; the operation, which uses four different kernels $\boldsymbol{k}_1$, $\boldsymbol{k}_2$, $\cdots$,$\boldsymbol{k}_4$, generates corresponding results $\boldsymbol{I_M^k}$, ($\boldsymbol{k} = 1,\cdots,4$). By selecting the largest object each from $\boldsymbol{I_M^k}$, the pig segmentation results in $\boldsymbol{I_p^k}$,($\boldsymbol{k} = 1,\cdots,4$) are shown.

the small circles at the bottom. Those are created due to shades, sunlight or the fences. Selecting only the pig segmentations among them can be achieved by finding the largest one. Since the pigs are the largest objects in the images, other areas except the largest one obviously correspond to non-pig areas. For this, the contour analysis is performed. Let $C_i^k (i = 1, 2,\cdots,m)$ the closed contour areas obtained after applying the contour analysis to I_M^k. Let C_r^k the pig segmentation, then r is determined as follows.

$$r = \underset{i=1,\cdots,m}{argmax}\Lambda(C_i^k) \tag{1}$$

where $\Lambda(C_i)$ is the size of C_i^k.

The resulting pig segmentations, C_r^k, $(k = 1,\cdots,n)$, each for the kernel of s_k, are used as masks when applying a bitwise AND operation to the binarization image I_B, removing all non-pig areas. As a result, the final segmentation results, I_p^k s of Fig. 4 are binary and have non-zero pixels only in the pig areas.

From I_p^k, the features are extracted and are used to estimate the weights. The first feature, c^k for I_p^k, is the size of the pig segmentation. Due to its linear and proportional relationship with the weights, it provides good approximations for the weights as it did in the previous works. The second one, δ^k for I_p^k, is *curvature* which represents how much a pig shape bends or twists. And the third, ϕ for all of I_p^k, is *deviation*; how apart the pig segmentations are from the center of I_p^k in average. Different from c^k by which the weights can be directly estimated, δ^k and ϕ have an indirectly role to adjust the weights according to the postures. The process of the feature calculation is discussed in detail. Let I_p as the set of I_p^k and ψ^k the set of 2D coordinates (x, y) of non-zero pixels of I_p^k as follows.

$$I_p = \{I_p^k |\ k = 1,\cdots,n\} \tag{2}$$

where n is the number of the used circular kernels.

$$\psi^k = \{(x, y)|\ I_P^k(x, y) > 0,\ I_p^k \in I_p\} \tag{3}$$

where $I_p^k(x, y)$ is the value of pixel at the coordinate (x, y) of I_p^k.

Then, the size of the pig area in I_p^k, c^k becomes

$$c^k = |\psi^k|,\ k = 1,\ \cdots,n \tag{4}$$

The curvature δ^k is calculated by measuring how symmetrical the pig segmentations are. The rationale behind is that the less symmetrical a pig shape is, the more curved it is. Thus the degree of symmetry can be estimated as follows. Imagine a straight line that crosses the center of the pig region in its length direction, dividing the shape in two halves. Then the ratio of the number of pixels over the line to the number under the line is an approximation about the curvature. Intuitively, ratios other than one imply that corresponding shapes are curved; the more apart from one, the more curved.

The dividing lines can be determined by using principal component analysis (PCA), which finds a basis along which values have the largest variance. Since ψ^k is the set of the pixel coordinates forming the pig region, the application of PCA to ψ^k can find a new basis, B^k which stretches in the length direction of the pig region. Thus, B^k is, in fact, the dividing line.

Having B^k become a new x axis, the coordinates of ψ^k are transformed into ψ_{tr}^k as follows.

$$\psi_{tr}^k = \{(x^{'}, y^{'})|(x^{'}, y^{'}) = PCA(x, y), (x, y) \in \psi^k\} \tag{5}$$

where $PCA(x, y)$ is the conversion of (x, y) with the basis B^k as the x axis.

The transformation, then, moves the pig region in such a way that, in the new coordinate bases, it is centered at (0,0) and aligned parallel to the new x axis in its length direction as shown in Fig. 5(a) before the transformation and (b) after.

For the definition of the curvature δ^k, Let χ_+^k and χ_-^k the set of the pixels over and under the new x axis, respectively, as follows.

$$\chi_+^k = \{(x, y)|(x, y) \in \psi_{tr}^k and y > 0\} \tag{6}$$

$$\chi_-^k = \{(x, y)|(x, y) \in \psi_{tr}^k and y < 0\} \tag{7}$$

Then, δ^k corresponding the pig segmentation in I_p^k is defined as

$$\delta^k = \frac{max(|\chi_+^k|, |\chi_-^k|)}{min(|\chi_+^k|, |\chi_-^k|)} \tag{8}$$

Understanding the property of $\delta^k \geq 1$, the closer to 1, the more symmetrical the shape is, implying that the pig posture is straight. On the contrary, as δ^k increases, the pigs are likely to have curved or bended posture, making it necessary to adjust the weight prediction accordingly. It is necessary to calculate $\delta^k s$, $(k = 1,\cdots,n)$for each I_p^k, because the corresponding pig shapes are different depending on the kernel sizes.

The third feature is the deviation of the pig shape from the image center. It was introduced according to the observation that the pig segmentations; as they are more apart from the center of the images, they have larger back sizes than their actual sizes. This oversizing effect is due to the pigs not being positioned at the center, which leads to the camera likely capturing their flanks. These images are then detected mistakenly as back during the segmentation. These errors in the segmentation increase c^k's, the number of the corresponding pixels, resulting in over-estimation.

The deviation can be measured by using the mean Euclidean distance between the image center and all the segmented pixels of the pigs. Let (x_c, y_c) the center coordinate of the images, which is the same for all. Given the coordinates of the pig segmentations, ψ^k, $(k = 1,\cdots,n)$, the deviation ϕ is defined as follows.

$$\phi = \sum_{k=1}^{n} \frac{\sum_{(x,y)\in\psi^k} \sqrt{(x_c - x)^2 + (y_c - y)^2}}{|\psi^k|} \tag{9}$$

The bigger ϕ is, the more adjustment to weight prediction is needed because of the inflated back size. Note that the deviation ϕ is a single value while the other features have multiple values, each for different kernel sizes. It is because ϕ has little difference among I_p^k.

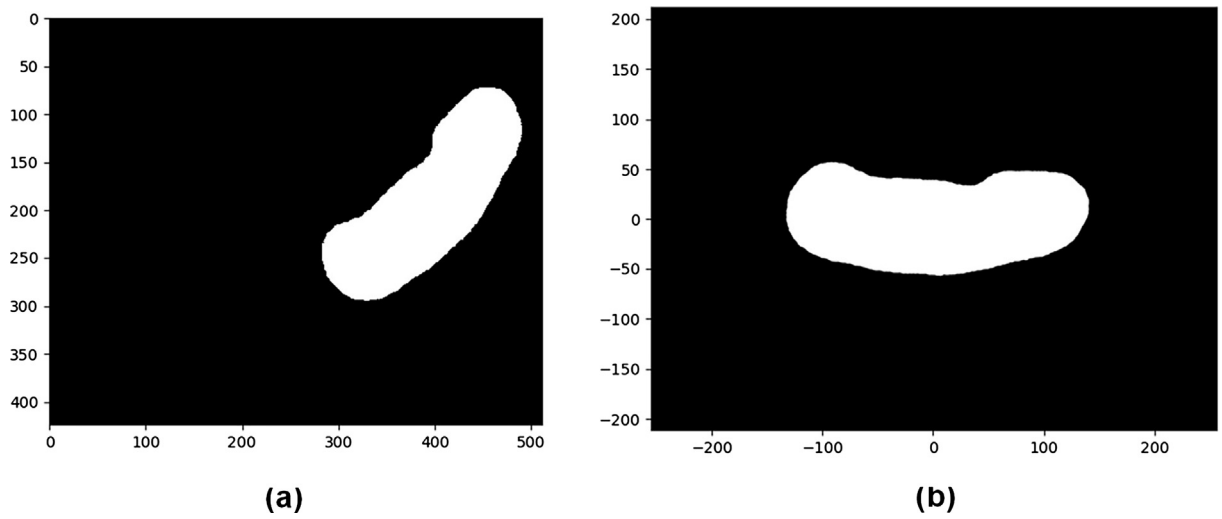

Fig. 5. PCA transformation of a pig shape: a pig shape (a) is transformed into (b) in which the shape is aligned parallel to x axis in its length direction and centered at (0,0).

For each of the images, there is a feature vector $\boldsymbol{F}$ having the above mentioned features as its elements as follows.

$$\boldsymbol{F} = (c^1, \delta^1, \cdots, c^n, \delta^n, \phi) \tag{10}$$

where n is the number of the kernels. In other words, given an input image I_G, $2 \times n + 1$ dimensional feature vector $\boldsymbol{F}$ is retrieved. In our work, since four kernels $n = 4$ are used, $\boldsymbol{F}$ thus becomes a vector of size 9.

Having the features as input, a fully connected neural network in Fig. 6 is used, shortly the neural network hereafter, to predict corresponding pig weights. Note that some of the edges between the nodes are omitted in order to keep the figure simple. The input layer has the same number of nodes as the dimension of the feature vector. The output layer has a single node, which outputs its prediction value for weights. There exist three hidden layers each of which has a set of ξ_1, ξ_2, ξ_3 nodes. There are not any nonlinear activation functions such as rectified linear unit (ReLU) (Glorot et al., 2011) or softmax (Bishop, 2006) because the linear relationship between the features and the weights are distorted.

For the training of the neural network, a loss function ℓ is defined as the sum of the mean squared prediction error plus a regularization factor, which is the sum of the squared weights of the neural network, in order to prevent the overfitting as follows.

$$\ell = \frac{\sum (\hat{y} - y)^2}{N_t} + \alpha \cdot L \tag{11}$$

where y is the actual weight, $\hat{y}$ the prediction, N_t the size of the training samples, α the regularization constant, which is a numerical value that controls the contribution of the regularization factor to the loss function, and L the regularization factor.

The training proceeds in the direction of minimizing the loss function ℓ by updating the weights and the biases of the nodes with the help of the gradient descent. To avoid overfitting, L is set to be the sum of the L2 norms of the weights and the biases.

3. Performance evaluation

Having a set of 513 pig images and the corresponding actual weight labels, the performance of the proposed method is evaluated. The images and their weights were recorded over several days in July, August and November 2016 in a farm owned by Easy Bio, a bioengineering company of South Korea. The pigs in the images are all

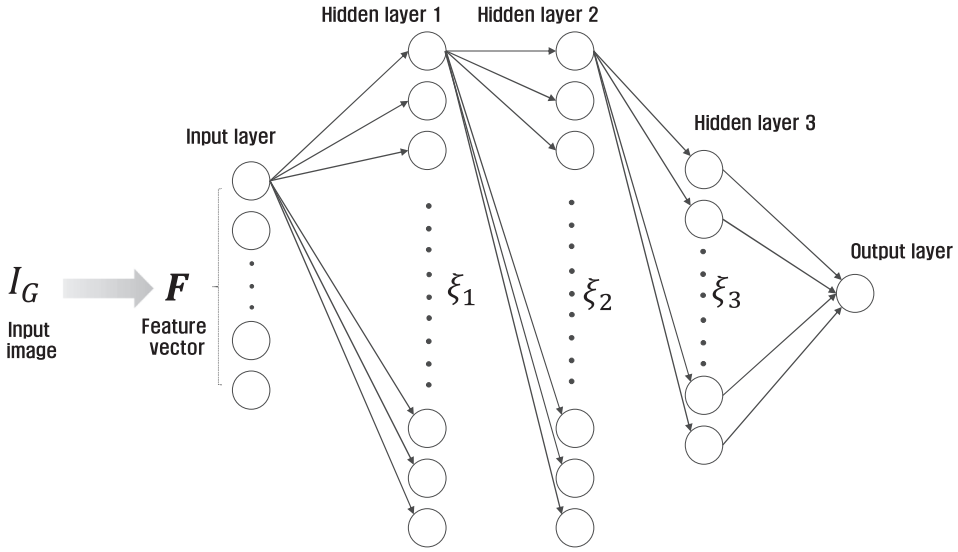

Fig. 6. Fully connected neural network for the pig weight estimation.

different ones. The number of the pigs in July and August was 262 and the November pigs was 251. In particular, there was no duplication of the pigs between July/August and November because all the July/August pigs had already been sold for the market before November. The actual weights were measured up to the tenth precision in kilogram by using a weighing device, which was installed underneath the pen. Considering the weight variance according to the pig movement, reading the scale was the responsibility of a skilled person. The weight distribution is in the range of 70– 120 kg. All the pigs are the same breed.

A video showing how the images were captured is available (Pig weighing environment, (n.d.)). In fact, a Kinect camera was used because our original plan was to develop a 3D image based estimation method. However, it turned out that, in the outdoor environment, the depth information captured by the Kinect camera was not consistent due to sunlight. As a result, the current method uses only the 2D features. Nevertheless, the 3D images, even if the depth value is not reliable, have been helpful in the pig segmentation.

The images are randomly divided into two separate sets: one training set of 410 images and one test set of 103, which corresponds to 20% of the entire images. The test set is never used during the training session. For the neural network training, the learning rate is set to 0.0001, which controls how much the weights of our neural network model are adjusted along the downward slope of the loss gradient for every mini-batch of training samples, and the regularization constant of 0.000015.

The performance of seven configurations of our model are compared. Table 1 shows the configurations which are different in the sizes of $\boldsymbol{F}$ and the structure of the neural network. The configuration names at the leftmost column of the table have a form of $cfg\text{–}F_n/L_n/N_n$ where F_n is the size of $\boldsymbol{F}$, L_n the number of layers in the neural network, and N_n the number of the nodes at the first hidden layer. The middle column shows the number of nodes at each layer, starting from an input layer. The last number of the list is 1 for all because it corresponds to the output layer. The rightmost column shows the vector elements of F and their order. For example, a configuration of cfg-4/5/32 uses four-element $\boldsymbol{F}$, which consists of only c^ks, while cfg-9/*/* employ all nine-feature $\boldsymbol{F}$.

For performance evaluation, there are three metrics that assess the difference between an actual weights y and a predicted value $\hat{y}$ by our model. The first metric is mean absolute error (MAE) of the differences between y and $\hat{y}$. The second one is root mean squared error (RMSE) for the same differences and the third is coefficient of determination R^2. Note that MAE and RMSE are similar, but both of them are used each for its unique role. While MAE provides an average of absolute differences, RMSE, which uses square of errors, produces a weighted average that can emphasize the existence of large errors.

Fig. 7(a) shows MAEs when the test data set is given as input to the different configurations. The configuration of cfg-4/5/32 shows the largest MAE. Note that it is the only configuration that uses a single feature, c^k. It is an indirect proof that unconstrained posture of the pigs complicated the weight estimation beyond what could be achieved with only c^k by the previous works. On the contrary, the configurations which use the posture related features, either δ^k or ϕ, or both, such as cfg-5/5/32, cfg-8/5/32, and cfg-9/5/32 showed better MAEs.

Table 1
Configurations of our model for performance evaluation.

Configuration	Num. of nodes at layers	Feature vector F
cfg-4/5/32	4-32-32-16-1	(c^1, c^2, c^3, c^4)
cfg-5/5/32	5-32-32-16-1	$(c^1, c^2, c^3, c^4, \phi)$
cfg-8/5/32	8-32-32-16-1	$(c^1, \delta^1, c^2, \delta^2, c^3, \delta^3, c^4, \delta^4)$
cfg-9/5/32	9-32-32-16-1	$(c^1, \delta^1, c^2, \delta^2, c^3, \delta^3, c^4, \delta^4, \phi)$
cfg-9/5/64	9-64-32-16-1	$(c^1, \delta^1, c^2, \delta^2, c^3, \delta^3, c^4, \delta^4, \phi)$
cfg-9/6/128	9-128-64-32-16-1	$(c^1, \delta^1, c^2, \delta^2, c^3, \delta^3, c^4, \delta^4, \phi)$
cfg-9/7/256	9-256-128-64-32-16-1	$(c^1, \delta^1, c^2, \delta^2, c^3, \delta^3, c^4, \delta^4, \phi)$

Another MAE improvement is attributed to the increased size of the neural networks. Comparing cfg-9/5/32 with cfg-9/5/64 which are same in the size of $\boldsymbol{F}$ but differ in the size of the neural networks, it is clearly observed that the increased size of the neural network improved MAE. However, when observing the results of all cfg-9/*/*, the relatively smaller neural network of cfg-9/5/64 is better than the bigger ones, cfg-9/6/128 and cfg-9/7/256. Thus it should be also mentioned that the complexity of the neural networks does not always bring improvement.

Fig. 7(b) shows that all RMSE of the configurations are larger than the corresponding MAE. It is because errors are squared before being averaged in RMSE. Hence, it causes the difference between RMSE and MAE to increase when there are large magnitude errors. However, the relative magnitudes of RMSE are similar to the results of MAE; cfg-4/5/32 shows the worst error, cfg-9/5/64 the best, and the relative differences among the configurations are alike. From the results of MAE and RMSE, it is shown that the configuration of cfg-9/5/64 is best at not only the average error but also the magnitude of errors.

Coefficient of determination, R^2 provides a measure of how well the models predict the actual weights. Similar to the results of MAE and RMSE, Fig. 8 shows that cfg-4/5/32 is the least appropriate model, while cfg-9/5/64 is the best model having R^2 of 0.79. The R^2 results again prove that the posture related features contribute to the improvement of the model suitability. Note that only cfg-4/5/32 which uses none of the posture features has R^2 below 0.7. Even though our work achieved 0.79 of R^2 at best which is below that of the previous works, 0.9 or even 0.99, it should be considered that there have been no attempt to control the image capture environment or to manually select a subset of the images that met the posture constraints.

Fig. 9 shows the weight prediction results along with the actual weights when using the configuration of cfg-9/5/64, which outperforms the other configurations. There are 103 pairs of weights for the test data set, which are plotted in the ascending order of the actual weights. The predictions show a tendency that they match their corresponding actual weights within limited 3.3 kg differences in average. Upon a closer look, the predictions are bigger when the actual weights are in the range of 90–105 kg, and smaller over 105 kg. It is because the morphological processing removes more pixels from large sized pig segmentations, decreasing c^k further than smaller ones.

Fig. 10 shows the frequency distribution of the absolute estimation errors from the 103 test cases. The error distribution has the mean of 3.15 kg with a standard deviation of 2.28. There are five instances which are larger than 7 kg, corresponding to 4% of the test data. The test images that caused such large differences have in common excessive exposure to sunlight which made the pig segmentation challenging. The segmentation results in such cases failed to separate pigs correctly; non-pig regions were included or only parts of pig areas were detected.

Such failures often occur when the pigs are too close to the sunlight exposed pen borders, resulting in larger segmentation sizes than the actual sizes. Such cases can be avoided by capturing the images only when the pigs are present in the central region, or selecting only the images in which the segmented pig areas do not contact the borders. Also, the thresholds for the maximum and minimum pig areas can filter out abnormally larger or smaller cases. These issues will be included in our future works.

4. Conclusions

The proposed image based pig weight estimation method is different from the previous works in several ways. There was no constraint on the pig posture and the image capture environment, used only 2D features for the weight estimation model, and made use of the recent advances in the machine learning.

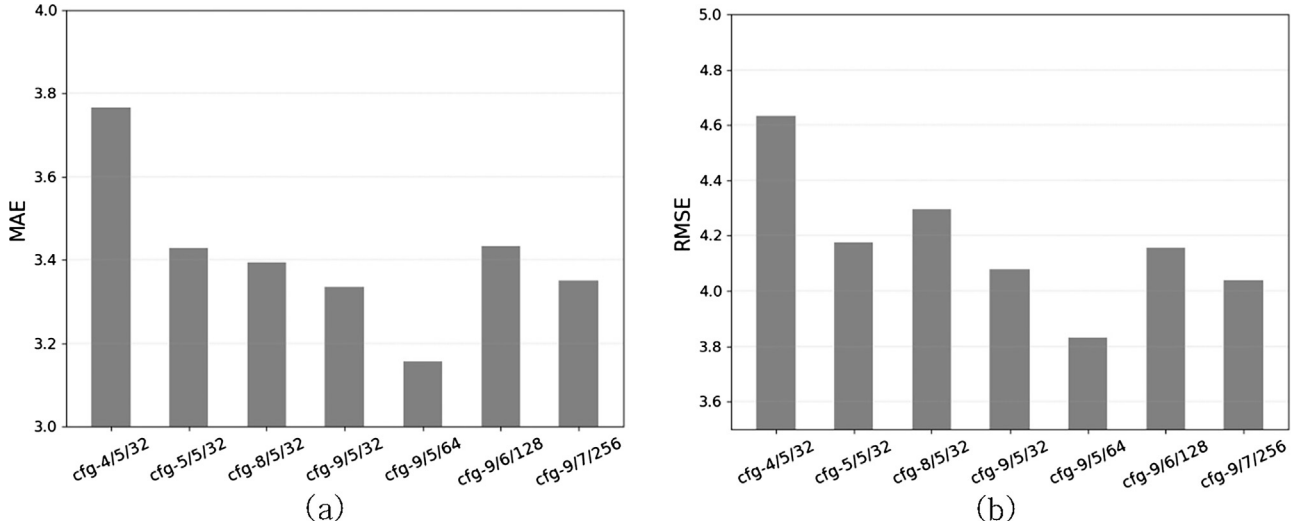

Fig. 7. (a) Mean absolute error (MAE) results and (b) root squared mean error (RMSE) results from the test data set.

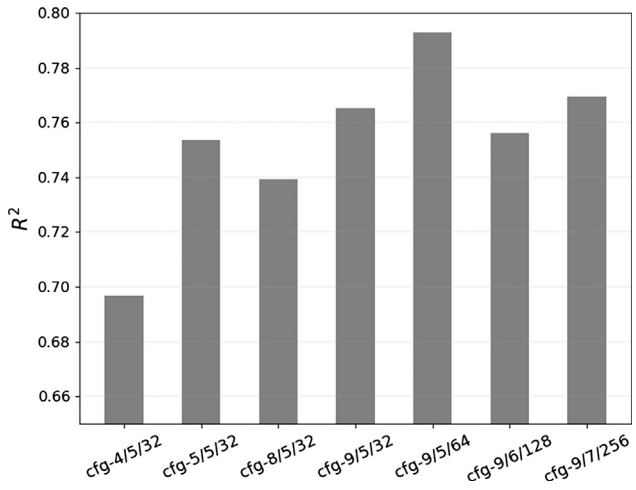

Fig. 8. Coefficient of determination R^2 from the test data set.

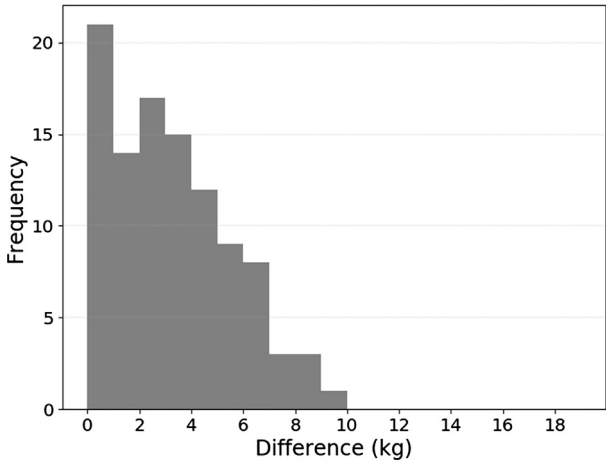

Fig. 10. Histogram of absolute estimation errors.

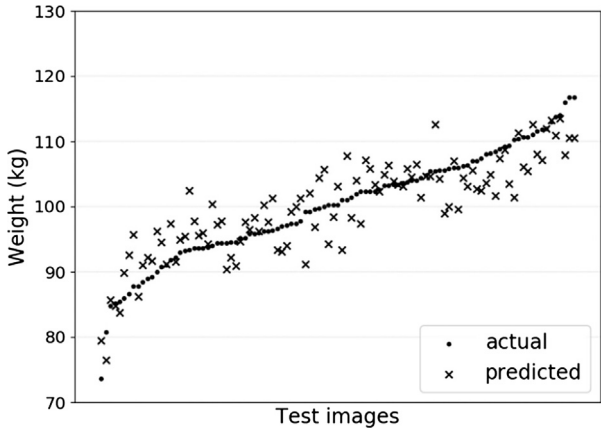

Fig. 9. The actual weights (dots) and the corresponding predicted weights ('x' marks) for the 103 test images. A pair of the dot and x marker corresponds to a single test image.

To meet challenges caused by these changes, the novel features were introduced in order to describe the posture: the curvature and the deviation. Thanks to these features, pigs did not need to be placed in narrow spaces or forced to be in a specific posture during image capture, which helps to reduce stress of weighing process. Due to these contributions, It was feasible to achieve the error with the mean of 3.15 kg and the model with $R^2 = 0.79$. These results are not impressive, but notable if considering the unconstrained posture, no human intervention in image selection, and use of only 2D feature, under which condition the previous works were not even able to work.

Although our objective was to develop a non-stressful method, it was necessary to isolate the pigs into the pen in order to collect the training data for the neural network. However, once the model is obtained, the isolation is not mandatory as long as the top view images like the training samples can be obtained. It would be possible by installing cameras at the places under which the pigs move around. However, such images may have the problems that multiple pigs appear, or only the partial bodies are captured. Our future works will address such problems by developing the image processing methods in order to use the images which are captured in stress-free environment.

Acknowledgement

This work was supported by Incheon National University Research Grant in 2017. We thank all the staff in Easy Bio Inc., who supported and helped the study.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the

online version, at https://doi.org/10.1016/j.compag.2018.08.006.

References

Azzaro, G., Caccamo, M., Ferguson, J., Battiato, S., Farinella, G., Guarnera, G., Puglisi, G., Petriglieri, R., Licitra, G., 2011. Objective estimation of body condition score by modeling cow body shape from digital images. J. Dairy Sci. 94 (4), 2126–2137.

Banhazi, T., Dunn, M. 2016. U.S. Patent No. 9,311,556. Washington, DC: U.S. Patent and Trademark Office.

Bishop, Christopher M., 2006. Pattern Recognition and Machine Learning. Springer.

Brandl, N., Jørgensen, E., 1996. Determination of live weight of pigs from dimensionsmeasured using image analysis. Comput. Electron. Agric. 15 (1), 57–72.

Costa, A., Ismayilova, G., Borgonovo, F., Leroy, T., Berckmans, D., Guarino, M., 2013. The use of image analysis as a new approach to assess behavior classification in a pig barn. Acta. Vet. Brno. 82, 25–30.

Duff, A.M., Carter, J.A., Hughes, C.A., Gates, K.N., Ellason, C.S., Stewart, W.S., Ribeiro, F.R.B., 2010. The relationship of real-time ultrasound body composition measurements, body weight and hip height with body condition score in mature Boer crossbred does. J. Dairy Sci. 93 (1), 459.

Glorot, X., Bordes, A., Bengio, Y. 2011. Deep sparse rectifier neural networks. In: Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics. pp. 315–323.

Guo, H., Ma, X., Ma, Q., Wang, K., Su, W., Zhu, D., 2017. LSSA_CAU: an interactive 3d point clouds analysis software for body measurement of livestock with similar forms of cows or pigs. Comput. Electron. Agric. 138, 60–68.

Halachmi, I., Klopcic, M., Polak, P., Roberts, D.J., Bewley, J.M., 2013. Automatic assessment of dairy cattle body condition score using thermal imaging. Comput. Electron. Agric. 99, 35–40.

Hough, P.V. 1962. U.S. Patent No. 3,069,654. Washington, DC: U.S. Patent and Trademark Office.

Kashiha, M., Bahr, C., Ott, S., Moons, C.P., Niewold, T.A., Ödberg, F.O., Berckmans, D. 2014. Weight estimation of pigs using top-view image processing. In: International Conference Image Analysis and Recognition. Springer, Cham. pp. 496–503.

Kawasue, K., Ikeda, T., Tokunaga, T., Harada, H., 2013. Three-dimensional shape measurement system for black cattle using KINECT sensor. Int. J. Circ. Syst. Signal Process. 7 (4), 222–230.

Kongsro, J., 2014. Estimation of pig weight using a Microsoft Kinect prototype imaging system. Comput. Electron. Agric. 109, 32–35.

Kollis, K., Phang, C.S., Banhazi, T.M., Searle, S.J., 2007. Weight estimation using image analysis and statistical modelling: a preliminary study. Appl. Eng. Agric. 23 (1), 91–96.

Kulikov, V.A., Khotskin, N.V., Nikitin, S.V., Lankin, V.S., Kulikov, A.V., Trapezov, O.V., 2014. Application of 3-D imaging sensor for tracking minipigs in the open field test. J. Neurosci. Meth. 235, 219–225.

Kuzuhara, Y., Kawamura, K., Yoshitoshi, R., Tamaki, T., Sugai, S., Ikegami, M., Kurokawa, Y., Obitsu, T., Okita, M., Sugino, T., Yasuda, T., 2015. A preliminarily study for predicting body weight and milk properties in lactating Holstein cows using a three-dimensional camera system. Comput. Electron. Agric. 111, 186–193.

Lee, J., Jin, L., Park, D., Chung, Y., 2016. Automatic recognition of aggressive behavior in pigs using a kinect depth sensor. Sensors (Basel, Switzerland) 16 (5), 631.

McFarlane, N., Wu, J., Tillett, R., Schofield, C., Siebert, J., Ju, X., 2005. Shape measurements of live pigs using 3-D image capture. Anim. Sci. 81 (3), 383–391.

Marchant, J., Schofield, C., White, R., 1999. Pig growth and conformation monitoring using image analysis. Anim. Sci. 68 (1), 141–150.

Menesatti, P., Costa, C., Antonucci, F., Steri, R., Pallottino, F., Catillo, G., 2014. A lowcost stereovision system to estimate size and weight of live sheep. Comput. Electron. Agric. 103, 33–38.

Mortensen, A.K., Lisouski, P., Ahrendt, P., 2016. Weight prediction of broiler chickens using 3D computer vision. Comput. Electron. Agric. 123, 319–326.

Nasirahmadi, A., Richter, U., Hensel, O., Edwards, S., Sturm, B., 2015. Using machine vision for investigation of changes in pig group lying patterns. Comput. Electron. Agric. 119, 184–190.

Nasirahmadi, A., Hensel, O., Edwards, S.A., Sturm, B., 2016. Automatic detection of mounting behaviors among pigs using image analysis. Comput. Electron. Agric. 124, 295–302.

Nasirahmadi, A., Edwards, S.A., Sturm, B., 2017. Implementation of machine vision for detecting behavior of cattle and pigs. Livest. Sci. 202, 25–38.

Oczak, M., Viazzi, S., Ismayilova, G., Sonoda, L.T., Roulston, N., Fels, M., Bahr, C., Hartung, J., Guarino, M., Berckmans, D., 2014. Classification of aggressive behavior in pigs by activity index and multilayer feed forward neural network. Biosyst. Eng. 119, 89–97.

Otsu, N., 1979. A threshold selection method from gray-level histograms. IEEE Trans. Syst. Man Cybern. 9 (1), 62–66.

Pallottino, F., Steri, R., Menesatti, P., Antonucci, F., Costa, C., Figorilli, S., Catillo, G., 2015. Comparison between manual and stereovision body traits measurements of Lipizzan horses. Comput. Electron. Agric. 118, 408–413.

Parsons, D., Green, D., Schofield, C., Whittemore, C., 2007. Real-time control of pig growth through an integrated management system. Biosyst. Eng. 96 (2), 257–266.

Pezzuolo, A., Guarino, M., Sartori, L., Marinello, F., 2018a. A feasibility study on the use of a structured light depth-camera for three-dimensional body measurements of dairy cows in free-stall barns. Sensors (Basel, Switzerland) 18 (2), 673.

Pezzuolo, A., Guarino, M., Sartori, L., González, L.A., Marinello, F., 2018b. On-barn pig weight estimation based on body measurements by a Kinect v1 depth camera. Comput. Electron. Agric. 148, 29–36.

Pig weighing environment. (n.d.). Retrieved February 12, 2018, from < https://drive.google.com/file/d/0B0WnTFsOha_JemZpMmltT25Fa0E/view?usp=sharing > .

Porto, S.M.C., Arcidiacono, C., Anguzza, U., Cascone, G., 2015. The automatic detection of dairy cow feeding and standing behaviours in free-stall barns by a computer vision based system. Biosyst. Eng. 133, 46–55.

Rosell-Polo, J.R., Cheein, F.A., Gregorio, E., Andújar, D., Puigdomènech, L., Masip, J., Escolà, A., 2015. Chapter three - advances in structured light sensors applications in precision agriculture and livestock farming. Adv. Agron. 133, 71–112.

Saberioon, M., Gholizadeh, A., Cisar, P., Pautsina, A., Urban, J., 2017. Application of machine vision systems in aquaculture with emphasis on fish: state-of-the-art and key issues. Rev. Aquacult. 9 (4), 369–387.

Salau, J., Haas, J.H., Junge, W., Bauer, U., Harms, J., Bieletzki, S., 2014. Feasibility of automated body trait determination using the SR4K time-of-flight camera in cow barns. SPRINGERPLUS 3.

Schofield, C.P., Marchant, J.A., White, R.P., Brandl, N., Wilson, M., 1999. Monitoring pig growth using a prototype imaging system. J. Agric. Eng. Res. 72 (3), 205–210.

Shi, C., Teng, G., Li, Z., 2016. An approach of pig weight estimation using binocular stereo system based on LabVIEW. Comput. Electron. Agric. 129, 37–43.

Spoliansky, R., Edan, Y., Parmet, Y., Halachmi, I., 2016. Development of automatic body condition scoring using a low-cost 3-dimensional Kinect camera. J. Dairy Sci. 99 (9), 7714–7725.

Stajnko, D., Brus, M., Hocevar, M., 2008. Estimation of bull live weight through thermographically measured body dimensions. Comput. Electron. Agric. 61 (2), 233–240.

Tasdemir, S., Urkmez, A., Inal, S., 2011. Determination of body measurements on the Holstein cows using digital image analysis and estimation of live weight with regression analysis. Comput. Electron. Agric. 76 (2), 189–197.

Vázquez-Arellano, M., Griepentrog, H.W., Reiser, D., Paraforos, D.S., 2016. 3-D imaging systems for agricultural applications—a review. Sensors 16 (5), 618.

Viazzi, S., Bahr, C., Van Hertem, T., Schlageter-Tello, A., Romanini, C.E.B., Halachmi, I., Lokhorst, C., Berckmans, D., 2014. Comparison of a three-dimensional and twodimensional camera system for automated measurement of back posture in dairy cows. Comput. Electron. Agric. 100, 139–147.

Vranken, E., Berckmans, D., 2017. Precision livestock farming for pigs. Animal Front. 7 (1), 32–37.

Wang, Y., Yang, W., Winter, P., Walker, L., 2008. Walk-through weighing of pigs using machine vision and an artificial neural network. Biosyst. Eng. 100 (1), 117–125.

Weber, A., Salau, J., Haas, J.H., Junge, W., Bauer, U., Harms, J., Suhr, O., Schonrock, K., Rothfuss, H., Bieletzki, S., Thaller, G., 2014. Estimation of backfat thickness using extracted traits from an automatic 3D optical system in lactating HolsteinFriesian cows. Livestock Sci. 165, 129–137.

Wongsriworaphon, A., Arnonkijpanich, B., Pathumnakul, S., 2015. An approach based on digital image analysis to estimate the live weights of pigs in farm environments. Comput. Electron. Agric. 115, 26–33.

Wu, J., Tillett, R., McFarlane, N., Ju, X., Siebert, J.P., Schofield, P., 2004. Extracting the three-dimensional shape of live pigs using stereo photogrammetry. Comput. Electron. Agric. 44 (3), 203–222.

Yilmaz, O., Cemal, I., Karaca, O., 2013. Estimation of mature live weight using some body measurements in Karya sheep. Trop. Anim. Health Prod. 45 (2), 397–403.