Calf Robust Weight Estimation Using 3D Contiguous Cylindrical Model and Directional Orientation from Stereo Images

Ryo Nishide
Kobe University
Hyogo, Japan
nishide@port.kobe-u.ac.jp

Ayumi Yamashita
Kobe University
Hyogo, Japan
yamashita@cs25.scitec.kobe-u.ac.jp

Yumi Takaki
Kobe University
Hyogo, Japan
yumi@people.kobe-u.ac.jp

Chikara Ohta
Kobe University
Hyogo, Japan
ohta@port.kobe-u.ac.jp

Kenji Oyama
Kobe University
Hyogo, Japan
oyama@kobe-u.ac.jp

Takenao Ohkawa
Kobe University
Hyogo, Japan
ohkawa@kobe-u.ac.jp

ABSTRACT

Calving interval is often used as an indicator for fertility of beef cattle, however, maternal abilities are also required because the value of breeding cows depends on how efficiently the healthy and growing calves are produced. The calf's weight has been used as an indicator of maternity ability since the past few decades. We propose a method to estimate body weight by modeling the shape of calf using 3D information extracted from the stereo images. This method enables to predict the swelling of the cattle's body by creating a 3D model, which cannot be obtained solely from a 2D image. In addition, it is possible to estimate robust weight regardless of different shooting conditions toward cattle's posture and orientation. An image suitable for estimation is selected from motion images taken by the camera installed in the barn, and 3D coordinates are calculated by the images. Then, only the body is developed with a 3D model as it has the highest correlation with the body weight. Considering that the side of cattle's body may not be exactly perpendicular to the camera's shooting direction, a symmetric axis is extracted to find the inclination of cattle body from the camera in order to generate a 3D model based on the symmetric axis. 3D contiguous cylindrical model is used for the body of a cattle which has a rounded shape. In order to manipulate the shapes of the cylindrical surface, the circle and ellipse fittings are applied and compared. The linear regression equation of the volume of the cylindrical model and the actually measured body weight are used to estimate the cattle weight. As a result of modeling with the proposed method using the actual camera images, the correlation coefficient between the body weight and the model volume was at the best value, 0.9107. Even when experimentally examined with the different 3D coordinates obtained from other types of camera, the MAPE (Mean Absolute Percentage Error) was as low as 6.39%.

SoICT 2018, December 6–7, 2018, Danang City, Viet Nam

ACM ISBN 978-1-4503-6539-0/18/12…$15.00
https://doi.org/10.1145/3287921.3287923

CCS CONCEPTS

• **Computing methodologies** → *Camera calibration*; *Epipolar geometry*; *Image segmentation*; *Object detection*; *Image processing*; • **Applied computing** → *Agriculture*;

KEYWORDS

weight estimation, calf, cattle, cow, stereo camera, circle fitting, ellipse fitting, three-dimensional reconstruction, depth camera.

ACM Reference Format:
Ryo Nishide, Ayumi Yamashita, Yumi Takaki, Chikara Ohta, Kenji Oyama, and Takenao Ohkawa. 2018. Calf Robust Weight Estimation Using 3D Contiguous Cylindrical Model and Directional Orientation from Stereo Images. In *The Ninth International Symposium on Information and Communication Technology (SoICT 2018), December 6–7, 2018, Danang City, Viet Nam.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3287921.3287923

1 INTRODUCTION

The production of dairy and beef cattle has been facing several issues in Japan, such as the lack of successors in the aging society, the rise of retail price due to the substantial reduction of cattle, and the increase of breeding costs due to the change of global demand. Thus, the decline of supply capacity and weakening of production base are the top priority issues to be addressed.

Generally, calving interval is often used as an indicator for fertility of beef cattle. However, the value of breeding cows depends on how efficiently the healthy and growing calves are produced. Even though the subsequent reproduction rate is high, the total fertility rate may not necessarily be high as expected, as problems may occur during the following growth trait of calves. In order to gain high subsequent reproduction rate, maternal abilities are also required as well as ideal calving interval. Maternal abilities refer to lactation and nursing ability, and calves growth weight has been used as an indicator of maternal ability since the past few decades[1]. The objective of this work is to facilitate the calves' weight estimation in order to grasp their maternal abilities and improve productivity.

There are two typical ways to estimate the weight of livestock; direct and indirect measurements. Direct measurement uses the scale and needs to move each livestock to ride on the scale, though it may be stressful for both the livestock and the breeders. On the other hand, for indirect measurement, several methods have

been proposed to use 2D images to estimate the weight[2–6]. For example, [2] proposed a method to estimate the weight using the image shot directly from above the cattle, and counting the number of pixels which represents the part of the cattle's body in the image. On the other hand, [7] uses weight prediction model from a 3D camera to extract morphological traits of hip height, hip width, and rump length of dairy cows. However, these methods require an overhead camera which must be placed in a high location above the cattle, and that the cattle must be shot perpendicularly to the ground or in a certain angle from above. In some occasions, an overhead camera may not be feasible to deploy in the environment.

The authors have also been working on indirect measurement to estimate weight by a 3D model of cattle's shape using stereo images taken from the side[8, 9]. Generally, stereo images are generated from two horizontally placed cameras that are used for obtaining the depth information in the form of a disparity map[10–12]. This enables to predict the swelling of the body which cannot be obtained by 2D images, besides, it does not require a unique environment to shoot the cattle from above. On the other hand, however, the side of cattle body may not be exactly perpendicular from the camera's shooting direction. As it is necessary to detect how much the cattle body is inclined from the perpendicular angle, we employ a symmetric axis which can express the inclination of the body. This enables to generate a 3D model of cattle accurately even though the body is slightly inclined, and provides robust weight estimation regardless of the posture and facing direction of cattle to some extent. Furthermore, our past findings have revealed the possibility that ellipse fitting can provide more accurate modeling rather than circle fitting[9]. However, reducing the high computational cost to perform ellipse fitting was an issue to be solved. As an update of our work, a least square method is used to perform ellipse fitting, and the fitting result is implemented into our model.

The images suitable for weight estimation are selected manually from videos taken inside the barn, and 3D coordinates are calculated from those images. A symmetric axis of a body part of cattle is extracted by circle/ellipse fitting with point cloud data plotted on the vertical and depth direction on a plane. Then, a 3D contiguous cylindrical model is used because the body of a cattle has a rounded shape. The weight is estimated using a linear regression equation with the 3D contiguous cylindrical model volume and the cattle's measured weight. For experiments, the applications of the usage of the symmetric axis, and the accuracy of modeling method using ellipse fitting and circle fitting are compared. Thus, using the linear regression equation, the weight is estimated. In order to verify the adaptability to other camera devices, depth camera which obtains the 3D coordinates using infrared is also examined.

Sec. 2 provides basic information about cattle and its body of which we aim to detect, and Sec. 3 describes the method to generate the 3D model of the body. Sec. 4 verifies the efficiency of our method from the results of the experiment, and Sec. 5 concludes the paper while indicating the direction to future works.

2 CATTLE BODY PARTS FOR EXTRACTION

2.1 Fundamental Information of Cattle Body

Calves are usually fed in a group, and the newborn calves of Japanese Black Cattle weigh approximately 30 kg. In the first 3 months,

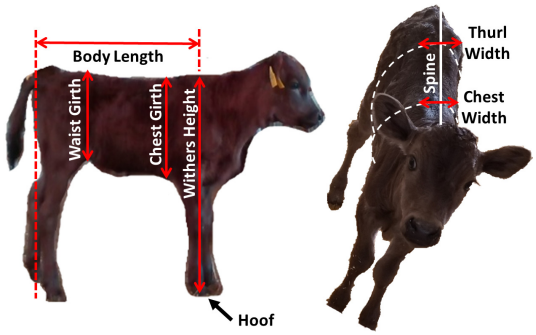


Figure 1: Names of Cattle Body Parts.

Figure 2: Sample Image Used for Analysis.

calves are generally raised together with their mother. In our experiment, smaller calves are preferable, but we chose calves up to 200 kg because a large number of learning data are necessary for weight estimation.

Fig. 1 left describes the name and range of each cattle body part. *Chest Girth* is known to be closely related to the cattle's weight, then following with the *Waist Girth*. Therefore, it seems reasonable to state that the width from the chest to the abdomen has the highest correlation to the cattle's weight, then the *Thurl Width*, *Body Length*, *Chest Width*, and so on. On the other hand, *Withers Height* and weight are known to be irrelevant. Thus, we focus on modeling from the chest to the abdomen which is highly correlated to the cattle weight.

2.2 Extraction of Body Parts

2.2.1 Detection of Cattle. Fig. 1 right is the image of cattle taken from above. It shows that the cattle body is symmetrical around the spine, and has a rounded shape. Therefore, if the shape of either of the side is detected, then the entire body can be modeled. In our analysis, we use images of cattle taken from the side, as shown in Fig. 2.

The following two methods are used together with extracting the cattle area. The first is a background subtraction method[13] using the background image taken in advance with no cattle in the foreground. The second is a method using parallax values obtained from the stereo matching method[10–12]. Specifically, we create a mask image, which is obtained from background subtraction

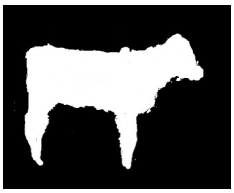


Figure 3: Binary Image.

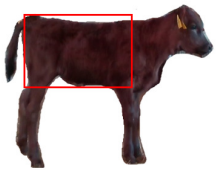


Figure 4: Body Range.

methods using background images and images of cattle taken at the same place. Then, we generate a disparity image from the stereo image and create another mask image extracted only from the parts with disparity value which is close to the cattle area. The common part of these two mask images is taken to extract the cattle area. We expect that applying these two methods may remove the remaining noise for each process. Moreover, the image is converted into a binary image in which the cattle area is white and the remaining part is black, and the largest white area is extracted as a cattle area while eliminating the other small white area as noise. The extracted image of the cattle area is shown in Fig. 3.

2.2.2 Extraction of Cattle Body. Chest girth is measured from the shoulder. It is difficult to decide the location of the shoulder from the image, so the exact length of chest girth is impossible to measure. In this work, instead of uniformly deciding the chest girth, the 3D coordinates from the chest to the buttocks are obtained and the thickness of the cattle's stomach is measured. Furthermore, the body is modeled using a contiguous cylindrical model, and the weight is obtained by correlating to the model[8, 9].

The range of *Body Length* and *Withers Height* can be verified in Fig. 1. Body length is from the shoulder to the buttocks. To obtain the body length data from the static image, it is necessary to locate the shoulder. As the shoulder has few distinctive features, the withers height is detected first with a perpendicular line to the ground, then the size from the line to the tip of the buttocks is determined as the body length. Withers height is the length from the shoulder top (withers) to the ground (the hoof in contact with the ground). In fact, we detect the hoof from the image and set the withers height with a perpendicular line from the hoof to the shoulder top. The hoof is determined from the mask image of the cattle used to detect the leg.

After the left and right ends of the body are determined by the location of legs, the top and bottom ends of the body are decided. The top and bottom ends are decided by the density of white pixels in a mask image in the sideways range of body. Fig. 4 shows the body range in a red rectangle.

3 CATTLE BODY MODELING

3.1 Overview of Proposed Method

The flow for estimating the cattle weight is illustrated in Fig. 5. Stereo camera images are used as an input, however, in attempting to verify the applicability of our weight estimation with other camera, a depth camera has also been deployed.

In order to estimate the weight from images, (1) the cattle and its body is extracted from the image, and (2) the body is modeled using 3D point cloud data extracted from the body range. The images of

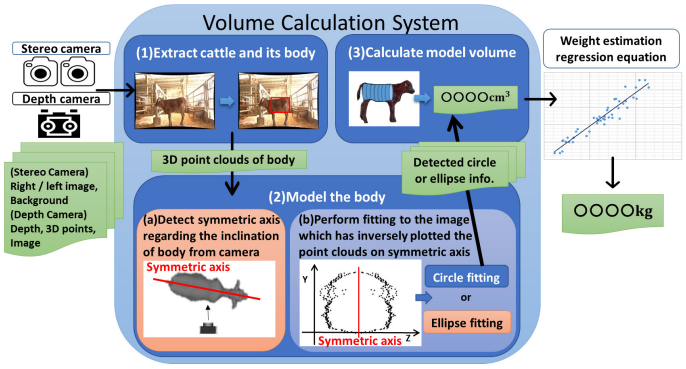


Figure 5: Procedure for Weight Estimation.

cattle taken from the side are used in our work, however cattle may not be exactly perpendicular from the camera's shooting direction. Therefore, (a) symmetric axis has been employed to detect the inclination of the body, in order to ensure the robustness of our method. Then, (b) the point is plotted inversely using the symmetrical axis obtained by inclination detection, and circle or ellipse fitting is performed. Finally, (3) the model volume is calculated, and the weight is estimated by the regression equation of body weight from the model volume.

Specifically, (a) body inclination is detected in the following procedures; partition the body to equal size width, perform circle fitting to a point group of each partition, and detect the inclination of the body using the circles' center. Then, (b) the following procedures generate a cattle body model; divide while sliding the body (slide division), perform circle/ellipse fitting with a symmetrically plotted point cloud, and generate a contiguous cylindrical model. The details of each process are described in the following section.

3.2 Modeling Cattle Body from 3D Coordinates

3.2.1 Necessity of Modeling. 3D point cloud data is generated only when the disparity is obtained from the stereo matching method. Basically, when the camera image size is $(x \times y)$ pixel, approximately $(x \times y)$ 3D point cloud data exist. When the image becomes large, the number of points increases accordingly. The shape and contour of the object become visible as the number of points increases, however, it will be difficult to determine which part is an error, or which part should be focused on. As our experimental subject, the Japanese Black Cattle does not have many distinctive characteristics in color, the 3D point cloud data seem to contain many errors of matching between images. Therefore, it is difficult to calculate the precise 3D volume data of the trunk from the 3D point cloud data. A method is necessary to approximate the point cloud data to the 3D model structure.

The shape of cattle is complicated, and various 3D models of different sizes are necessary to create an elaborate cattle model. Here, we aim to generate modeling in a generally fixed 3D structure (i.e. cylinder, cuboid).

3.2.2 Dividing the Body. This work employs a cylindrical model as the cattle body has a rounded shape. The cylindrical model from the side has a rectangular shape. The red rectangular area in Fig. 4 can

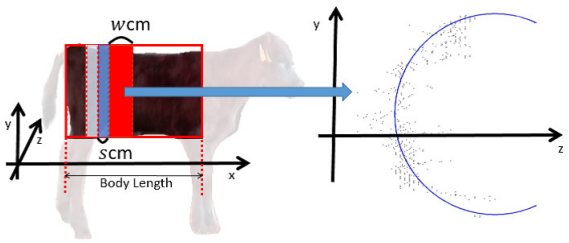


Figure 6: Slide Division and Mapping in yz-plane.

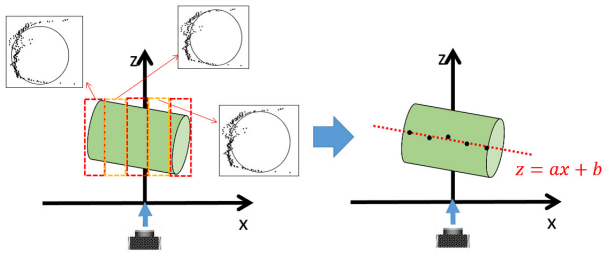


Figure 7: Method to Find the Symmetric Axis of the Body.

be interpreted as a cylindrical model viewed from the side. Thus, using the 3D point cloud data obtained from stereo matching, fitting is performed to create a model with a height y and a width z on a yz-plane, and the radius of the cylindrical model is determined.

Considering the cattle body in Fig. 4 to be sliced vertically, the vertical length of the body varies depending on its sliced location. In order to create a model of the body as similar to the real one, it is necessary to designate the different size of girth depending on its location. Thus, a body model is sliced into several partitions in order to express various sizes of girth.

Let x be the horizontal direction, y be the vertical direction, and z be the depth direction. The body length is the length of the body in the horizontal direction. Slide division method is applied to divide the body in a fixed thickness size (w cm) by sliding the cutting range (s cm) little by little (for example, in a w = 5cm thickness, slide s = 1cm each). The image of slide division method is shown in Fig. 6.

Thus, the body is divided into several partitions in x-direction. The point cloud in each partition is assigned as a set, and each set is mapped on the yz-plane. For each point cloud, as one z value is determined for one y value, the average of z is obtained for each y-direction. Finally, the detailed body shape can be obtained.

3.2.3 Fitting Method. The shape of cattle is detected from point cloud in yz-plane. Circle fitting can be applied assuming that the body is round, however, the body is much more like an ellipse rather than a circle[9]. Especially, smaller calves are skinny with a flat belly, which makes their body elliptical, and thus, errors may likely occur if the circle fitting is applied. To verify such concerns, we perform modeling by using two methods, circle fitting and ellipse fitting, in order to compare the accuracy through experiments.

Circle fitting is performed using a generalized *Hough transform*. *Hough transform* is a technique to detect straight lines and circles from images[14]. Searching entirely for the center and radius of the circle may impose a high computational cost. However, the cost may be reduced by limiting the search range by assigning the conditions such as the radius size and the center location of the circle. For example, as the point cloud is inferred only from the body of cattle, the diameter of a circle may not exceed of a certain size even when the points may include errors. Therefore, the search range may be reduced to a circle diameter which is smaller than the y-direction width of the point cloud. In this way, not only that it can reduce computational cost, but also the outliers which are obviously irregular can be removed.

Among various ways for ellipse fitting[15], a *least squares method* has been chosen in our work. The *Least squares method* is a method to solve the least squares of errors, as in ellipse for example, in which the center, major axis, minor axis and an inclination are determined from variables a_1 to a_5 in equation (1).

$$\sum_{i=1}^{I} \left(x_i^2 + a_1 x_i y_i + a_2 y_i^2 + a_3 x_i + a_4 y_i + a_5 \right)^2 = 0. \quad (1)$$

I refers to the number of points, and the i'th coordinate is (x_i, y_i) in which $(i = 1, 2, \ldots, I)$. The center, major axis, minor axis and the inclination are obtained uniformly, and the outliers cannot be removed during calculation as in *Hough transform*. In this paper, instead of controlling the size of the ellipse, the width of the body (location of point cloud) is considered when determining the area of the ellipse. The detail is described in Sec. 3.2.6.

A circle can be regarded as an ellipse with the same length of the major and minor axis and will be expressed as an ellipse from here on. *Hough transform* is applied for circle fitting, and *least squares method* is applied for ellipse fitting. Using these two types of fitting, a contiguous cylindrical model is generated individually. The efficiencies of the models are then compared with the experiments.

3.2.4 Obtaining the Symmetric Axis. Sec. 2.2.1 denotes that the body of a cattle is almost symmetrical around the spine. Therefore, if the 3D point cloud is rotated around the symmetric axis (on the spine), then the body structure can be reconstructed. The images used for analysis are taken from the side of cattle, and the symmetric axis should exist perpendicular to the camera's shooting direction. However, cattle do not always appear perpendicularly to the camera, so if the symmetric axis (location of the spine) is misaligned, the body will be reconstructed inaccurately compared with the actual size.

The method to correct the inclination of cattle toward the camera is examined by assigning the symmetric axis from the 3D point cloud. Fig. 7 is an image of xz-plane where the camera is at the origin. The cylinder is assumed to be the cattle body. First, divide the point cloud of the body into I number of points in x-direction. Next, perform circle fitting with the point clouds included in the divided area, and detect each center point (x_i, z_i) where $(i = 1, 2, \ldots, I)$. Then, solve the linear regression equation for the detected I number

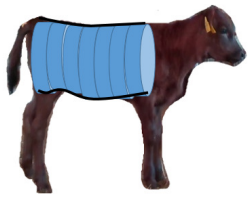


Figure 8: Contiguous Cylindrical Model.

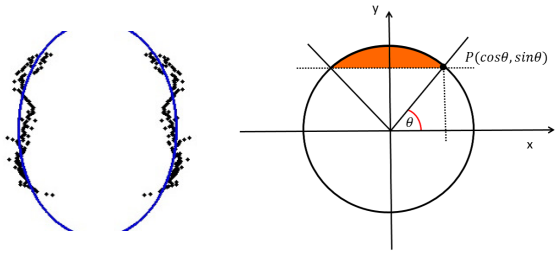


Figure 9: Ellipse Fitting and Corner Area Removal.

of points, and find the linear expression of the inclination body toward the camera. Finally, extract the symmetric axis of the body.

3.2.5 Contiguous Cylindrical Modeling. From the image plotted with the average point cloud data, and with the projected point cloud data around the symmetric axis, perform fitting as described in Fig. 6. From the ellipse constructed by fitting, the radius of the cylindrical model would be obtained. Fig. 8 is the image of the contiguous cylindrical model of the body. The actual data will have more divisions and cylinders, and the outer surface of the body will be much more smoothened.

3.2.6 Extracting Body Area in Ellipse Fitting. Fig. 9 left shows an example of detected point cloud data and ellipse fitting. Note that the point cloud is concentrated to the center, and not extended to the far end of the major axis. The far end can be considered that it is outside the body range. Therefore, when calculating the area of the ellipse, among all the points in point cloud, the maximum and minimum points in y-direction are extracted, and the top and bottom portions are removed. Fig. 9 right illustrates the maximum point P, and the orange portion to be removed.

4 EVALUATION

4.1 Experiment Settings

The target of the experiment is newborn calves which weigh less than 250 kg. at the Food Resources Education and Research Center, Graduate School Agricultural Science, Kobe University. The two network cameras, AXIS M1145 (Fig. 10 left) and Corega CG-NCBU031A (Fig. 10 right), are placed in parallel to construct stereo cameras. The cameras are facing towards the passage where only the calves pass by. There are 184 images extracted from the video taken on January 13 and January 19, 2017. Using those data, the linear regression equation of the actual measurement of weight and the volume, and the correlation coefficient are obtained. Moreover,

AXIS M1145.

Corega CG-NCBU031A.

Figure 10: Two Network Cameras for Experiment.

regarding the images for the same calves taken simultaneously, the results are obtained by the calculation of average volume. The efficiency of the method is examined by comparing the correlation coefficient of each method. Meanwhile, using the linear regression equation, the weight W_i^{estimate} is estimated from volume V_i, then the error$_i$ of W_i^{estimate} from the actual measurement of weight W_i^{actual} is calculated using equation (2), and the average is obtained (*MAPE*).

$$\text{error}_i = \frac{\left|W_i^{\text{estimate}} - W_i^{\text{actual}}\right|}{W_i^{\text{actual}}} \times 100. \quad (2)$$

The accuracy for circle fitting and ellipse fitting described in Sec. 3.2.3 is compared. On the other hand, the accuracy is also compared whether or not the inclination of a body stated in Sec. 3.2.4 is applied. If the inclination of a body is not applied, then the entire point cloud data is plotted in the yz-plane without dividing the body, and the symmetric axis is located along the center $(x_{\text{center}}, y_{\text{center}})$ of the circle fitting with the entire point cloud data. The following three methods are examined.

(a) Circle fitting with symmetric axis regarding the inclination of the body,
(b) Ellipse fitting with symmetric axis disregarding the inclination of the body,
(c) Ellipse fitting with symmetric axis regarding the inclination of the body.

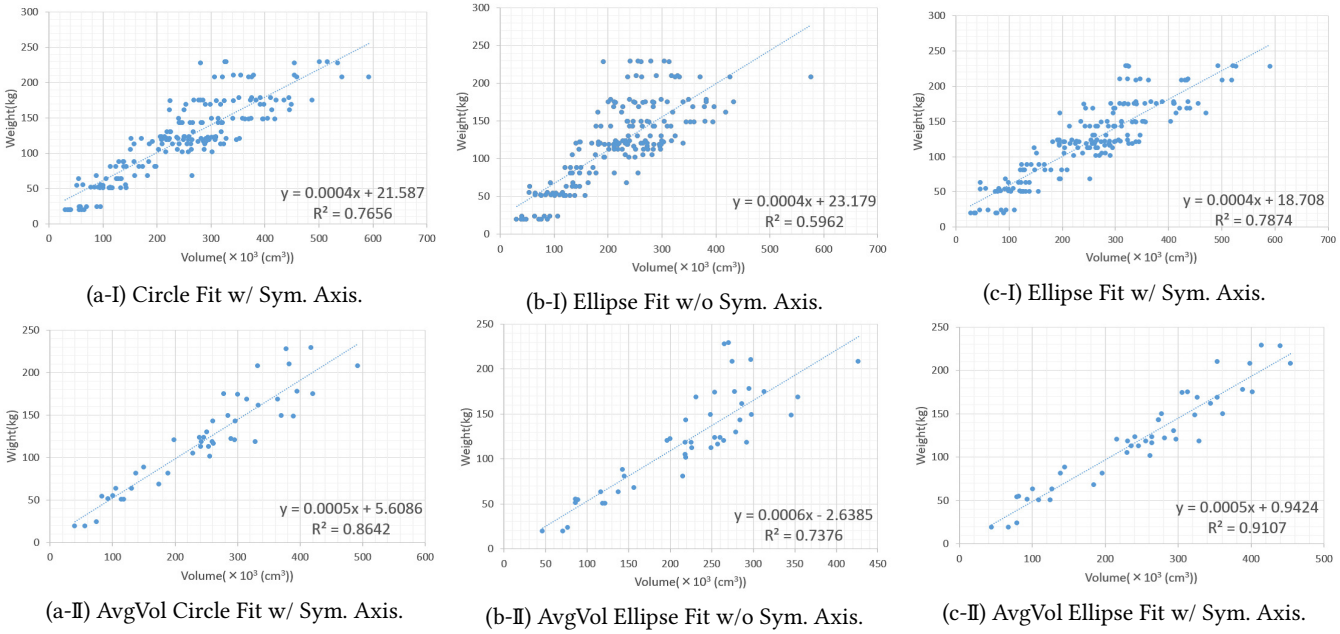
The experiment settings are as follows:

- The width to slice the cattle body is $w = 5$cm, and the length to repeatedly slide is $s = 1$cm.
- Number of division to determine the symmetric axis is $I = 10$.
- The threshold used to determine the top and bottom body area is 70%.

4.2 Volume and Weight Estimation

4.2.1 Graph Analysis. This section provides information about the graphs in Fig. 11.

- In (a-I) and (a-II), the symmetric axis regarding the inclination of the body is used, and the result of circle fitting drawn from the image by inverting the point cloud data is examined.
- In (b-I) and (b-II), the result of ellipse fitting from the image plotted with the entire point cloud data disregarding the inclination of the body is described.
- In (c-I) and (c-II), the symmetric axis regarding the inclination of body is used, and the result of ellipse fitting drawn from the image by inverting the point cloud data is provided.



(a-Ⅰ) Circle Fit w/ Sym. Axis. (b-Ⅰ) Ellipse Fit w/o Sym. Axis. (c-Ⅰ) Ellipse Fit w/ Sym. Axis.

(a-Ⅱ) AvgVol Circle Fit w/ Sym. Axis. (b-Ⅱ) AvgVol Ellipse Fit w/o Sym. Axis. (c-Ⅱ) AvgVol Ellipse Fit w/ Sym. Axis.

Figure 11: Body Volume and Weight Evaluation Based on the Types of Fitting Method and Usage of Symmetric Axis.

Table 1: Correlation Coefficient and *MAPE*

Fitting Method	Cor. Coef. (R^2 =)	*MAPE*(%)
Circle Fit w/ Sym. Axis	0.7656	22.91
Ellipse Fit w/o Sym. Axis	0.5962	26.45
Ellipse Fit w/ Sym. Axis	0.7874	22.83
AvgVol Circle Fit w/ Sym. Axis	0.8642	17.85
AvgVol Ellipse Fit w/o Sym. Axis	0.7376	23.05
AvgVol Ellipse Fit w/ Sym. Axis	0.9107	15.81

(a-Ⅰ), (b-Ⅰ) and (c-Ⅰ) show the plotted area, in which the calculated volume of the model is on the horizontal axis, and the measured weight is on the vertical axis. On the other hand, (a-Ⅱ), (b-Ⅱ) and (c-Ⅱ) show the result of the average volume obtained individually for each cattle with multiple images.

4.2.2 Correlation Analysis of Volume and Weight. Tab. 1 shows the summarized result of the graphs. The correlation between the volume estimated with each respective method and the actual measurement of weight and *MAPE* of the weight estimated from linear regression are shown. Using the average volume obtained individually for each cattle with multiple images, the correlation between the average volume and the actual weight is also shown.

4.3 Discussion of Experiment

First, we have compared the case of using the average volume of a cattle shot from a camera at the same time with the case of using the volume separately. The results have been improved for all methods. Each stereo image may include errors from the stereo matching method, fitting, and so on, but it can be mitigated by calculating the average. However, it may not always be the best solution to obtain average volume, because the volume of cattle varies extremely in some cases. Therefore, other methods must also be considered.

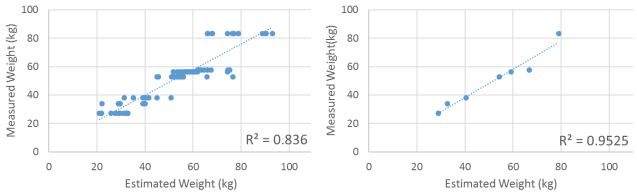
In our past works[8, 9], only the ideal images with cattle shot approximately perpendicular to the camera range have been selected manually and used. Circle fitting was performed to those selected images in which the *MAPE* was 30.76% and the *MAPE* of the average volume was 21.46%. Even considering that the results of the past works were biased by its manual selection process, most of the results in Tab. 1 have improved significantly by employing symmetric axis and/or ellipse fitting in this work.

Next, we have examined the two fitting methods in Sec. 3.2.3. The results in (a-Ⅰ, Ⅱ) and (c-Ⅰ, Ⅱ) from Sec. 4.2.1 are obtained in the same environment using different fitting methods. The correlation coefficient for ellipse fitting which used the least squares method was slightly higher. This indicates that the ellipse fitting was able to model the cattle body closer to its actual shape.

Finally, the results with or without using the symmetric axis have been compared. The results show that the correlation coefficient was higher for the case when the symmetric axis was used. Especially, the results were remarkable for large calves which weigh over 100kg, probably because the cattle positioned slightly not perpendicular to the camera was successfully aligned.

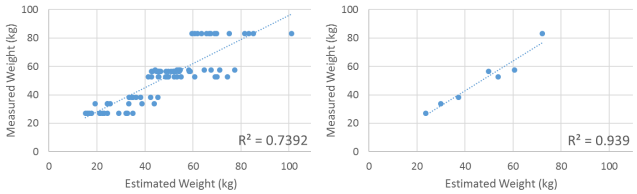
4.4 Weight Estimation Using Other Camera

4.4.1 Deploying Depth Camera. The reason of the drawbacks of using a stereo camera is that it requires calibration in order to obtain 3D coordinates. Calibration must be conducted each time when the location of camera changes, which extremely affects the accuracy of 3D point cloud data. It is desirable to calibrate before the camera is deployed and to be handled very cautiously when moving the camera, however, this may not be an easy task in the



(A) Weight from Volume. (B) Weight from Avg. Volume.

Figure 12: Graph of Circle Fitting with Depth Camera.



(A) Weight from Volume. (B) Weight from Avg. Volume.

Figure 13: Graph of Ellipse Fitting with Depth Camera.

practical environment. Searching for a solution, we have considered utilizing the depth camera which does not require calibration and examined the possibility to estimate weight.

First of all, the features of the stereo camera and depth camera are compared. In order to calculate the depth with a stereo camera, the stereo matching method is employed using left and right images and calibration is required for pre-processing. For depth camera, depth computation is performed with infrared, and thus the shadows do not affect the result[16]. However, the depth cannot be obtained at a place under the strong sunlight such as in the outdoor environment. For the performance in 3D point cloud data, stereo matching depends on the calibration accuracy, while depth camera does not require pre-processing and the location of points are highly accurate in the indoor environment.

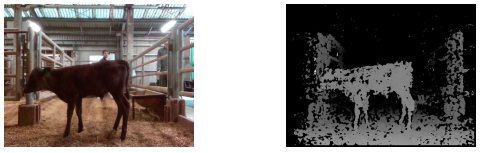
Intel RealSense 3D Camera ZR300[1] is used for depth camera. Even though the resolution is low, the depth accuracy seems high as calibrations are not necessary. This camera can obtain 3D coordinates from depth measurement in real-time, but it cannot save video data. Thus, static image data are captured instead of video data, and the binary files of 3D coordinates and depth are generated. Calves body are modeled from these data, then the volume is calculated, and finally, the weight is estimated.

The data taken with a depth camera on January 10, 2018 is used for evaluation. The experiment was held at the same place as Sec. 4.1, but the target was on 7 newborn calves which weighed less than 100 kg. In Sec. 4.2, the fitting considering the inclination of the body was effective, so the weight was estimated with two linear regression equations used in Fig. 11 (a-Ⅱ) and (c-Ⅱ), and thus, we will examine here whether the same linear regression equations can be applied to estimate the weight even for the data taken by different cameras and depth values obtained by different techniques.

[1]https://click.intel.com/media/ZR300-Product-Datasheet-Public-002.pdf

Table 2: *MAPE* of the Experiment with Depth Camera

Fitting Method	Cor. Coef. (R^2 =)	*MAPE*(%)
Circle Fit w/ Sym. Axis	0.8360	11.42
Ellipse Fit w/ Sym. Axis	0.7392	16.28
AvgVol Circle Fit w/ Sym. Axis	0.9525	6.39
AvgVol Ellipse Fit w/ Sym. Axis	0.9390	8.42



Color image. Color depth.

Figure 14: Successful Color Depth Extraction.

Color image. Color depth.

Figure 15: Unsuccessful Color Depth Extraction.

4.4.2 Result of Estimation. The weight estimation results using the image from the depth camera are shown in this section. Fig. 12 provides the estimated weight result from the model generated with circle fitting. On the other hand, Fig. 13 is the estimated weight result performed with ellipse fitting. In *A*, the vertical axis is the measured weight, the horizontal axis is the estimated weight using the same linear regression equation from Sec. 4.2, and the correlation coefficient is also provided. Using the average volume obtained from multiple images of the same cattle, *B* shows the weight estimation from the average volume.

Tab. 2 shows the correlation coefficient and *MAPE* of estimated weight using linear regression equation for circle/ellipse fitting and the actual measured weight. Fig. 14 is the successful example of extracting color depth, in which the left is the color image used for analysis, and the right is the color depth image. Fig. 15 is an example that failed to extract the depth of body, and thus, the data was not suitable and not used for analysis.

4.4.3 Discussion. Tab. 2 shows that the maximum value of *MAPE* was 16.28%. Weight estimation using the same linear regression equation with a different camera was highly accurate. Comparing ellipse and circle fittings, ellipse fitting had a significant result than circle fitting in Sec. 4.2, however, ellipse fitting had a larger error (*MAPE*) when the depth camera was used. This is probably because the center point of circle fitting was used when the symmetric axis is calculated. Considering that ellipse fitting is suitable for calves of which girths are not shaped as large as the circle, circle fitting has been used when generating point cloud for ellipse fitting. Some points were plotted to enlarge the circle than the ideal body size. The depth camera only had data of few calves weighing less

than 100 kg, which might have considerably affected the error rate, however, we believe that the efficiency has been verified as that weight can be estimated with *MAPE* as low as 8.42%.

On the other hand, the detectability of the depth of the body is highly affected by the environment as shown in Fig. 15. As the depth is computed using infrared, the radiation may have been absorbed due to the effect of solar rays and the black color calves. Even inside the barn, depth data was hardly extracted during the daytime as it was affected by the bright sunlight. However, depth camera is highly practical because it does not require calibration, and the placement and relocation can be performed with ease. Therefore, a depth camera can be useful if there are not much sunlight and the shadows of the body do not appear in the image.

5 CONCLUSION

In this paper, the weight estimation method using the volume of the 3D contiguous cylindrical model of cattle body has been proposed. Two network cameras were deployed to work as a stereo camera, and the image appropriate for analysis was chosen from the video taken with the camera. The experiment was conducted to verify the effectiveness of modeling the body from images to estimate the weight. For alternative methods, depth camera was also examined using the same linear regression equation to estimate the weight, and the results have shown for applicability with other cameras.

The accuracy of weight estimation was improved by applying the inclination of the symmetric axis of cattle. Even though the body of cattle was slightly inclined from the camera's shooting direction, a highly accurate result was obtained. Therefore, the proposed method is useful to perform a robust weight estimation disregarding the posture and facing direction of cattle. However, small cattle have flat belly compared to large cattle, and thus when performing circle fitting to designate the symmetric axis, the trunk of the body was modeled larger than the actual size. Even though the ellipse fitting was applied to model the trunk, circle fitting was conducted to generate the point cloud data from image, which resulted in generating large trunk. Thus, when designating the symmetric axis of the trunk, the result may improve using multiple directional images together with the images from the side of cattle.

It is also important to improve the accuracy of 3D coordinates, as they are used to obtain the radius of the model by fitting. However, there are various errors in the 3D point cloud, such as in the calibration of the camera, or in the calculation of parallax values from block matching. As a solution, we have also examined modeling with 3D coordinates obtained from infrared, instead of using a stereo matching method. The depth camera showed the possibility to obtain accurate 3D coordinates, though the solar rays hindered to obtain the depth value even in the indoor cowshed environment. The depth camera does not require calibration and is easy to install, but further investigation is required to determine the filming conditions in order to stably obtain the depth value.

An efficient weight estimation was enabled by averaging the volume of the same cattle, thus using multiple images are desirable to be used. However, the images sometimes provided a completely different volume estimation results, therefore calculating the average volume individually is not always the best solution to the problem. Instead of calculating the volume for each stereo image individually, using multiple images in the modeling process can be expected to work effectively. For example, as determining of the location of the symmetric axis extremely affects the results, using multiple images to relatively matching with the location of the symmetric axis may significantly improve the volume calculation.

In our experiment, each cattle was induced to move in front of the camera, in which the workload is the same as guiding the cattle one by one to board on the scale. Thus, it is desirable to estimate the weight without using human labor but to take the video of cattle in their natural habitat in the barn. As the video recording environment has already been equipped inside the barn, a method is necessary to automatically extract analyzable images from the video data, and apply it in the practical environment.

ACKNOWLEDGMENTS

This research was supported by the subsidized projects (2016-2017) of Japan Livestock Technology Association, the Council for Information Utilization of Beef Cattle Improvement, the Ministry of Agriculture, Forestry and Fisheries, and JST CREST Grant Number JPMJCR1682, Japan. Note that any implications in this work are not their official opinions.

REFERENCES

[1] K. Meyer. 1992. Variance components due to direct and maternal effects for growth traits of Australian beef cattle. *Livestock Production Science*, 31, 179–204.
[2] M. Ogata, Y. Nakahashi, R. Shibuya, and K. Kuchida. 2011. Possibility and its precision of body weight prediction by image analysis method for Japanese Black calves. *Research Bulletin Obihiro University*, 32, 14–19.
[3] C. Enevoldsen and T. Kristensen. 1997. Estimation of Body Weight from Body Size Measurements and Body Condition Scores in Dairy Cows. *Journal of Dairy Science*, 80, 1988–1995.
[4] S. Tasdemir, A. Urkmez, and S. Inal. 2011. Determination of body measurements on the Holstein cows using digital image analysis and estimation of live weight with regression analysis. *Computers and Electronics in Agriculture*, 76, 189–197.
[5] U. Paputungan, L. Hakim, G. Ciptadi, and H. F. N. Lapian. 2013. The estimation accuracy of live weight from metric body measurements in ongole grade cows. *Journal of the Indonesian Tropical Animal Agriculture*, 38, 149–155.
[6] A. J. Heinrichs, G. W. Rogers, and J. B. Cooper. 1992. Prediction body weight and wither height in Holstein heifers using body measurements. *Journal of Dairy Science*, 75, 3576–3581.
[7] X. Song, E. A. M. Bokkers, P. P. J. van der Tol, P. W. G. Groot Koerkamp, and S. van Mourik. 2018. Automated body weight prediction of dairy cows using 3-dimensional vision. *Journal of Dairy Science*, 101, 5, 4448–4459.
[8] A. Yamashita, T. Ohkawa, K. Oyama, C. Ohta, R. Nishide, and T. Honda. 2017. Estimation of Calf Weight from Fixed-Point Stereo Camera Images Using Three-Dimensional Successive Cylindrical Model. *Proceedings of the 5th IIAE International Conference on Intelligent Systems and Image Processing 2017*, 247–254.
[9] A. Yamashita, T. Ohkawa, K. Oyama, C. Ohta, R. Nishide, and T. Honda. 2018. Calf Weight Estimation with Stereo Camera Using Three-Dimensional Successive Cylindrical Model. *Journal of the Institute of Industrial Applications Engineers*, 6, 1, 39–46.
[10] J. J. Orteu. 2009. 3-D computer vision in experimental mechanics, optics and lasers in engineering. *Optics and Lasers in Engineering*, 47, 282–291.
[11] D. Scharstein, R. Szeliski, and R. Zabin. 2002. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision*, 47, 7–42.
[12] Z. Zhang. 2000. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 22, 1330–1334.
[13] Y. Benezeth, P. M. Jodoin, B. Emile, H. Laurent, and C. Rosenerger. 2008. Review and Evaluation of Commonly-Implemented Background Subtraction Algorithms. *19th International Conference on Pattern Recognition*, 1–4.
[14] P. V. C. Hough. 1962. Method and means for recognizing complex patterns, U.S. Patent No. 3 069 654, Filed Mar. 25th., 1960, Issued Dec. 18th., 1962.
[15] D. K. Prasad, M. K. H. Leung, and C. Quek. 2013. An unconstrained, non-iterative, least squares based geometric Ellipse Fitting method. *Pattern Recognition*, 46, 1449–1465.
[16] E. J. Fernandez-Sanchez, J. Diaz, and E. Ros. 2013. Background Subtraction Based on Color and Depth Using Active Sensors. *Sensors*, 13, 8895–8915.