RAPID AND AUTOMATED BODY MEASUREMENT OF CATTLE BASED ON STATISTICAL SHAPE MODEL

Yuzhi Bao[1], Hexiao Lu[1,2], Jianhuan Wu[1], Jie Lei[1], Jialong Zhang[2], Xinying Luo[1], Hao Guo[1,2,*]

[1]College of Land Science and Technology, China Agricultural University,
Beijing 100083, China - (2020321010228, luhexiao)@cau.edu.cn,
huanhuan9982@163.com, (sy20223213383, sy20213213175, guohaolys)@cau.edu.cn
[2]College of Information and Electrical Engineering, China Agricultural University,
Beijing 100083, China - 1250391455@qq.com

KEY WORDS: Precision livestock farming, Body measurement, Statistical shape model, Point cloud

ABSTRACT:

The current methods of non-contact livestock body measurement directly deal with the low-quality point cloud data of livestock, which have low robustness and lack practicality. On the one hand, the success rate of keypoint detection for livestock body measurement is low. Due to the severe occlusion and noise in the point cloud data, body measurements of some data cannot be performed. On the other hand, the key frames need to be manually selected from the point cloud sequence during processing. Inspired by the work of 3D reconstruction based on animal statistical shape models, we implement the construction and learning of the statistical shape model of real cattle. Given the establishment of the statistical shape model of cattle, a 3D reconstruction and body measurement approach of real cattle based on low-quality point cloud data is proposed. Nine indicators are calculated and the overall estimation MAPE (Mean Absolute Percentage Error) is 10.27%. The whole process of the body measurement algorithm proposed in our paper can be extended to other quadrupeds.

1. INTRODUCTION

Accurate monitoring of the livestock body is vital for farmers and breeders to comprehend the growth status, production, reproduction and breeding of livestock. Manual measurement cannot meet the growing demand for intelligence (Shuai et al., 2020, Bartol et al., 2021). The automated growth monitoring of livestock is of great significance for the sustainable development of animal husbandry.

The detection, tracking and analysis of animals have diverse applications in biology, neuroscience, ecology, agriculture and recreation. Although widely used, the field of computer vision focuses more on modeling the human body, such as estimating human pose and analyzing human behavior. However, it is not directly feasible to extend or apply this work to animals. The main reason is that compared with humans, animals are obviously not as cooperative as humans. Inappropriate interventions can have a dramatic impact on animal well-being, so there are far fewer 3D scan datasets for animals than for humans such as (Anguelov et al., 2005, Weiss et al., 2011, Bogo et al., 2014). Therefore, existing research on the shape and pose of animals lags behind that on the human body by a large margin.

1.1 Methods based on multi-view stereo vision

In recent years, scholars have carried out a lot of research on the application of stereo vision in the field of animals. (Wu et al., 2004) used six high-resolution cameras to obtain images of the top, side, and rear views of each pig, and developed a stereo imaging system that reconstructed the 3D shape of live pigs. (Pezzuolo et al., 2018) proposed a photogrammetry method based on Structure from Motion (SfM) and applied it to the 3D modeling and measurement of the pig body. The above animal reconstruction methods based on multi-view stereo technology require conditions such as illumination and camera synchronization, and the poses of animals are limited.

1.2 Methods based on depth camera

(Kongsro, 2014) used the Kinect camera to collect depth map images of pigs and estimated the weight of the pig from images. (Wang et al., 2018) proposed a portable automatic measurement system for pig body size, in which two Xtion depth cameras were utilized to capture point cloud data from two viewpoints, and the body measurement was realized by segmentation and pose normalization. (Ruchay et al., 2020) designed a vision system consisting of three Kinect v2 cameras to acquire cattle data for automatic body measurement. The above approaches have high operating efficiency, but also have high demands on animal cooperation and multi-camera synchronization, and the reconstruction accuracy is not high.

1.3 Methods based on 3D template

(Cashman and Fitzgibbon, 2012) predefined a dolphin template to learn a low-dimensional model of its deformation by manually extracting keypoints and manually segmenting. The model was optimized to minimize reprojection errors of keypoints and silhouettes. The method works for dolphins, but there are limitations in fitting non-template objects. (Vicente and Agapito, 2013) obtained the template of the corresponding object from the reference image and used the deformation of the template to fit the input image. The resolution of the reconstruction result obtained by this method is low. (Kanazawa et al., 2016) learned separate animal models of cats and horses and presented a volumetric deformation framework to deform 3D templates through user interaction. Template-based reconstruction techniques do not explicitly model joints, and the result of template deformation is still a rough shape, resulting in poor reconstruction. The

* Corresponding author

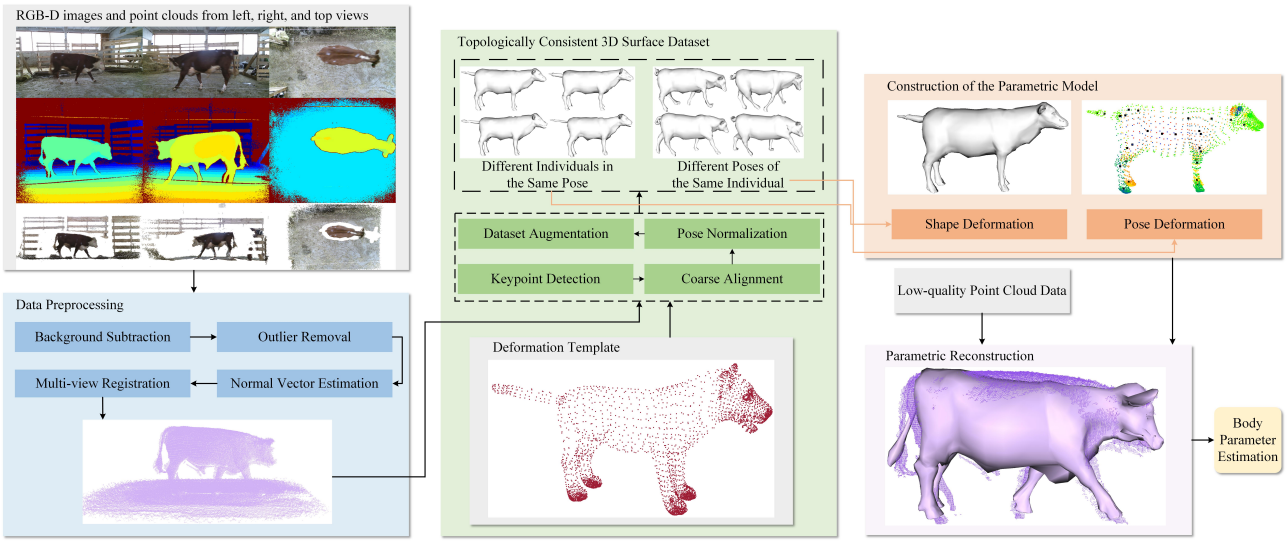


Figure 1. The pipeline of our approach.

majority of the three aforementioned methods solely concentrate on the shape features of outcomes, and fail to ensure the accuracy and reasonableness of topological results.

1.4 Methods based on statistical shape models

Currently, it is a research hotspot to use deep learning and other technologies to regress parameters of animal parametric models from RGB images or videos to realize 3D reconstruction of animal body surfaces. Inspired by the parametric model of the human body, (Zuffi et al., 2017) proposed a Skinned Multi-Animal Linear model (SMAL) to capture variations in shapes and poses among various quadruped toy figurines. Based on SMAL, researchers have sequentially proposed species-specific parametric models. (Li et al., 2021) defined the horse model hSMAL and applied it to video-based lameness detection. (Biggs et al., 2019) proposed a system to recover 3D models of various quadrupeds from videos. (Zuffi et al., 2019) proposed an end-to-end model SMALST that integrated the SMAL model into a regression network to reconstruct 3D animal shapes with texture information from a single image of a field scene. (Biggs et al., 2020) generated a new parametric model SMBLD including limb scaling. The 3D reconstruction of the animal body surface is generally developing in a data-driven direction. The SMAL model is used as the deformation template in this paper.

In this work, we take full advantage of statistical shape models to obtain topologically consistent 3D surface datasets. Our contribution is to propose a method for parametric reconstruction of cattle from low-quality point cloud data. Our process is generalizable and can be broadened to encompass other quadrupeds.

2. METHODS

2.1 Overview

Fig.1 shows a pipeline of the proposed method. A series of preprocessing is first performed on the original data. Based on the prior model, the template mesh is fitted to the scan data using constraints such as keypoints. A pose normalization method is proposed, and all fitted meshes are unified into a standard pose. On the topologically consistent 3D surface dataset, parametric models of animal shape and pose are constructed and learned. Finally, based on low-quality observation data, animal parametric reconstruction and body parameter measurement are carried out.

2.2 Experimental dataset and preprocessing

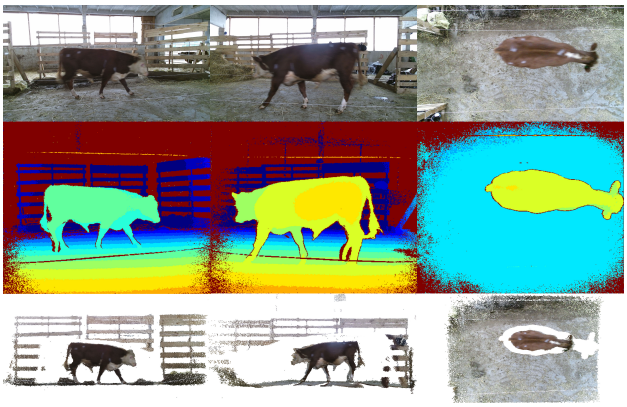


Figure 2. RGB-D images and point cloud data of cattle from three different views.

As shown in Fig.2, the original dataset is captured synchronously by three Microsoft Kinect v2 cameras and consists of RGB-D images and point clouds from left, right, and top views of 103 cattle, as well as manually measured body references. The dataset, provided by (Ruchay et al., 2020), was obtained on a rectilinear passway through an automated data acquisition system. Each data is mainly composed of standing cattle and background environment. Three laptops respectively control three depth cameras and are on the same network. Because the time on the laptop is synchronized, the minimum time interval among three camera devices is selected to generate the dataset to obtain the best matching result for the point cloud data. The dataset also includes transformation matrices among cameras for registering data obtained from three viewpoints into a unified coordinate system. The cattle are constantly moving freely during the acquisition process, so there are differences in the data poses of the cattle. Since a series of obstacles such as sensor quality, animal movement and on-site environment, missing data, outliers and noise are unavoidable.

To obtain reliable point cloud data, background subtraction, outlier removal, normal vector estimation and multi-view registration are performed. Fig.3 shows the data preprocessing result of a low-quality point cloud of a cattle.

Figure 3. The result of data preprocessing of a cattle.

2.3 Topologically consistent 3D surface dataset

The template mesh is fitted to the point cloud data and processed by pose normalization. The meshes of different poses are normalized to the canonical pose. Then the topologically consistent surface dataset is output.

2.3.1 Parametric model As shown in Fig.4, the SMAL model $\boldsymbol{\mathcal{T}} = \mathcal{M}(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\gamma})$ (Zuffi et al., 2017) is selected for its generalization, where $\boldsymbol{\mathcal{T}}$ is a 3D mesh with 3889 vertices and 7774 faces, $\boldsymbol{\beta}$, $\boldsymbol{\theta}$ and $\boldsymbol{\gamma}$ are pose, shape and translation parameters, respectively. The pose parameter $\boldsymbol{\beta}$ is a 33×3 dimensional vector representing the rotations of the 33 joints in the parametric model. The shape parameter $\boldsymbol{\theta}$ that determines the body shape characteristics of the model is a 41-dimensional vector, which is obtained by PCA on the 3D animal toy dataset. The displacement in three orthogonal directions relative to the initial position is controlled by the translation parameter $\boldsymbol{\gamma}$.

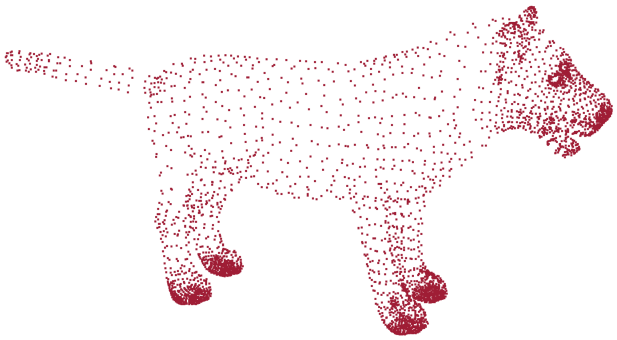


Figure 4. An illustration of the SMAL model.

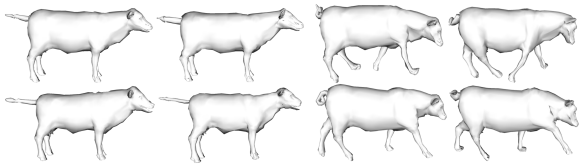
2.3.2 Keypoint detection and coarse alignment In model fitting, high-quality corresponding points are essential for registration. DeepLabCut (Mathis et al., 2018) is used for 2D keypoint detection. 2D is mapped to 3D according to the intrinsic and extrinsic parameters of the camera. Then keypoints $\boldsymbol{\mathcal{S}}_k = \{\boldsymbol{s}_i^k\}_{i=1}^{27}$ on scan data and $\boldsymbol{\mathcal{T}}_k = \{\boldsymbol{t}_i^k\}_{i=1}^{27}$ on template mesh are obtained. Considering that keypoints on the template $\boldsymbol{\mathcal{T}}$ only need to be predefined once, the corresponding points $\boldsymbol{\mathcal{T}}_k$ are marked manually.

To prevent being trapped in a local optimum while fitting, the sum of squared errors of $\boldsymbol{\mathcal{S}}_k$ and $\boldsymbol{\mathcal{T}}_k$ is minimized as eq(1) (Horn et al., 1988):

$$E(a, \boldsymbol{\mathcal{R}}, \boldsymbol{b}) = \sum_{i=1}^{27} \left\| a\boldsymbol{\mathcal{R}}\boldsymbol{s}_i^k + \boldsymbol{b} - \boldsymbol{t}_i^k \right\|^2, \tag{1}$$

where a is the scale factor, $\boldsymbol{\mathcal{R}}$ is the rotation matrix and $\boldsymbol{b}$ is the translation vector. The transformed data is $\boldsymbol{\mathcal{S}}^{'}$ and its keypoints are $\boldsymbol{\mathcal{S}}_k^{'} = \{\boldsymbol{s}_i^{'k}\}_{i=1}^{27}$.

2.3.3 Pose normalization and dataset augmentation To reduce the impact of the animal's free movement on fitted meshes, $\boldsymbol{\beta} = \boldsymbol{0}$ is set to realize normalization. Blender is used to build the skeleton, the data of 190 different individuals in the same pose and 99 different poses of the same individual are obtained as Fig.5.



(a) Different individuals in the same pose. (b) Different poses of the same individual.

Figure 5. Topologically consistent 3D surface dataset.

2.4 Construction of the parametric model

2.4.1 Shape deformation Principal component analysis is used for dimensionality reduction to characterize body shapes of different cattle. Vertex coordinates are arranged as vectors to form a shape space. The PCA analysis is performed on the set of points in the shape space. Orthogonal bases are extracted to express the shape subspace, so that a set of shape parameters $\boldsymbol{\theta} = \{\theta_1, ..., \theta_{|\theta|}\}$ is used to characterize the shapes of different individuals.

All vertices $\{\boldsymbol{p}_i^f \mid i = 1, ..., N_v; f = 1, ..., F\}$ of topologically consistent meshes of F different cattle with the same pose are input. N_v is the number of vertices on each individual. $\{\boldsymbol{p}_i^f\}$ is performed row stacking to get $\boldsymbol{P} \in R^{F \times N_v}$. The shape space $\boldsymbol{P}$ is normalized by orientation and decentralized, and the average shape $\overline{\boldsymbol{\mathcal{T}}}$ and the covariance matrix $\boldsymbol{\mathcal{D}}$ of the normalized shape space are computed. $\boldsymbol{\mathcal{D}}$ is subjected to eigenvalue decomposition. According to the size of the eigenvalues, the eigenvectors are arranged in descending order. The eigenvectors corresponding to the first 23 eigenvalues are taken to form the shape base $\boldsymbol{\mathcal{V}} = \{\boldsymbol{\mathcal{V}}_1, ..., \boldsymbol{\mathcal{V}}_{|\theta|}\}$. Linear blending is performed to approximate the shapes of different individuals, as eq(2):

$$f(\boldsymbol{\theta}) = \overline{\boldsymbol{\mathcal{T}}} + \boldsymbol{\mathcal{V}}\boldsymbol{\theta}, \tag{2}$$

where $f(\boldsymbol{\theta})$ is the shape mapping function. Fig.6 shows the average shape $\overline{\boldsymbol{\mathcal{T}}}$ of the statistical shape model of the real cattle. Different shapes can be obtained by adjusting the value of the shape parameter $\boldsymbol{\theta}$.

2.4.2 Pose deformation Linear blend skinning (LBS) is a commonly used skinning method that binds 3D model meshes to bones. A bone vertex weight map establishes a connection between each vertex and the bones. Given $\boldsymbol{\beta} = \{\beta_1, ..., \beta_{|\beta|}\}$, each vertex $\boldsymbol{v}_i$ is processed by LBS as eq(3):

$$\boldsymbol{v}_i = \sum_{m=1}^{|\beta|} \boldsymbol{w}_{im}(\boldsymbol{R}_m \boldsymbol{p}_i + \boldsymbol{t}_m), \boldsymbol{w}_{im} \geq 0 \wedge \sum_{m=1}^{|\beta|} \boldsymbol{w}_{im} = 1 \tag{3}$$

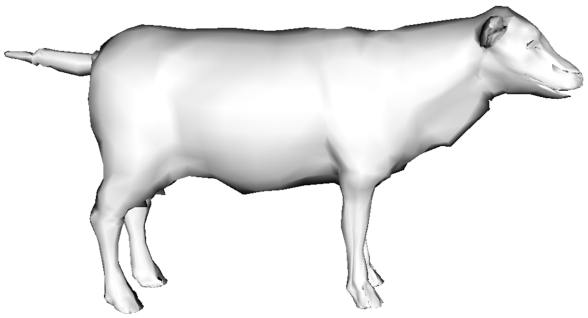


Figure 6. The average shape of the statistical shape model of the real cattle.

Where $\boldsymbol{w}_{im}$ is the weight of the $m\text{-}th$ bone to the $i\text{-}th$ vertex, $\boldsymbol{p}_i$ is the coordinate of the $i\text{-}th$ vertex in the standard pose, $\boldsymbol{R}_m$ and $\boldsymbol{t}_m$ are the rotation and translation matrices of the $m\text{-}th$ bone, respectively.

SSDR (Smooth Skinning Decomposition with Rigid Bones) (Le and Deng, 2012) is used to solve $\boldsymbol{w}_{im}$. Given $\mathcal{J}$ target poses (i.e., datasets of different poses of the same individual), meshes $\boldsymbol{\mathcal{T}}$ in a resting state (including $\mathcal{N}$ vertices) and $|\beta|$ bone joint points. Vertices of target poses are $\{\boldsymbol{v}_i^j \mid i = 1, .., \mathcal{N}; j = 1, ..., \mathcal{J}\}$. The algorithm is as eq(3):

$$\min_{\boldsymbol{w},\boldsymbol{R},\boldsymbol{t}} E = \min_{\boldsymbol{w},\boldsymbol{R},\boldsymbol{t}} \sum_{j=1}^{\mathcal{J}} \sum_{i=1}^{\mathcal{N}} \|\boldsymbol{v}_i^j - \sum_{m=1}^{|\beta|} \boldsymbol{w}_{im}(\boldsymbol{R}_m^j \boldsymbol{p}_i + \boldsymbol{t}_m^j)\|^2 \tag{4}$$

where $\boldsymbol{R}_m^j$ and $\boldsymbol{t}_m^j$ are rotation and translation matrices of the $m\text{-}th$ bone in the $j\text{-}th$ pose, respectively. The statistical shape model of the real cattle is $\boldsymbol{\mathcal{T}} = \mathcal{F}(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\gamma})$.

2.5 Parametric reconstruction

The process of 3D reconstruction is transformed into the problem of minimizing the loss function. The animal shape and pose prior (Biggs et al., 2020) are used for regularization. Let $\boldsymbol{\mu}_{\boldsymbol{\beta}}$ and $\boldsymbol{C}_\beta$ denote the mean and covariance matrices of the pose prior. The constraint E_{pose} is expressed by the Mahalanobis distance as eq(5):

$$E_{pose}(\boldsymbol{\beta}) = (\boldsymbol{\beta} - \boldsymbol{\mu}_{\boldsymbol{\beta}})^\top \boldsymbol{C}_\beta^{-1} (\boldsymbol{\beta} - \boldsymbol{\mu}_{\boldsymbol{\beta}}) \tag{5}$$

the shape prior constraint E_{shape} is similar to the pose prior as eq(6):

$$E_{shape}(\boldsymbol{\theta}) = (\boldsymbol{\theta} - \boldsymbol{\mu}_{\boldsymbol{\theta}})^\top \boldsymbol{C}_{\boldsymbol{\theta}}^{-1} (\boldsymbol{\theta} - \boldsymbol{\mu}_{\boldsymbol{\theta}}) \tag{6}$$

where $\boldsymbol{\mu}_{\boldsymbol{\theta}}$ and $\boldsymbol{C}_{\boldsymbol{\theta}}$ are the mean and covariance of the shape prior, respectively. The local joint rotation constraint E_{rotate} reduces the rotation of joints on the X-axis and Z-axis to meet the needs of cattle's motion poses. The keypoint constraint $E_{keypoint}$ minimizes the sum of distances between keypoints to optimize the shape and pose of the template mesh $\boldsymbol{\mathcal{T}}$ as eq(7):

$$E_{keypoint}(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\gamma}, s) = \sum_{i=1}^{N_k} \|s\boldsymbol{v}_i^k(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\gamma}) - \boldsymbol{s}_i^{'k}\|^2 \tag{7}$$

where N_k is the number of keypoints, s is the scale factor and $\boldsymbol{v}_i^k(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\gamma})$ is the keypoint of $\boldsymbol{\mathcal{T}}$ under pose $\boldsymbol{\beta}$, shape $\boldsymbol{\theta}$, and translation $\boldsymbol{\gamma}$. $\boldsymbol{s}_i^{'k}$ and $\boldsymbol{v}_i^k(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\gamma})$ are corresponding points. The data constraint E_{data} is expressed as the sum of the distances of all corresponding points of $\boldsymbol{\mathcal{T}}$ and $\boldsymbol{\mathcal{S}}^{'}$ and measures how close $\boldsymbol{\mathcal{T}}$ is to $\boldsymbol{\mathcal{S}}^{'}$ as eq(8):

$$E_{data}(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\gamma}, s) = \sum_{i=1}^{N_c} \|s\boldsymbol{v}_i^c(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\gamma}) - \boldsymbol{s}_i^{'c}\|^2 \tag{8}$$

where N_c is the number of corresponding points, and $\boldsymbol{s}_i^{'c}$ is the nearest neighbor point of $\boldsymbol{v}_i^c(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\gamma})$ on $\boldsymbol{\mathcal{S}}^{'}$. The loss function $E_{\mathcal{P}}(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\gamma}, s)$ can be expressed as eq(9):

$$\begin{aligned} E_{\mathcal{P}}(\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{\gamma}, s) = {} & w_{pose}E_{pose}(\boldsymbol{\beta}) + w_{shape}E_{shape}(\boldsymbol{\theta}) + \\ & w_{rotate}E_{rotate}(\boldsymbol{\beta}) + w_{keypoint}E_{keypoint}(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\gamma}, s) + \\ & w_{data}E_{data}(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\gamma}, s) \end{aligned} \tag{9}$$

where w_{pose}, w_{shape}, w_{rotate}, $w_{keypoint}$ and w_{data} are five weights. With the goal of minimizing $E_{\mathcal{P}}(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\gamma}, s)$, the gradient descent algorithm Adam (Kingma and Ba, 2014) is used to optimize $\boldsymbol{\beta}$, $\boldsymbol{\theta}$, $\boldsymbol{\gamma}$ and s, and the best fitting mesh for low-quality observation data is is shown in Fig.7.

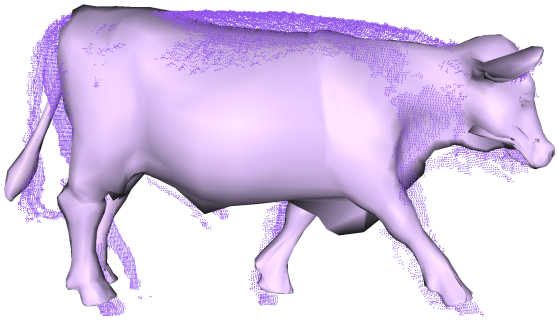


Figure 7. An illustration of the reconstruction effect.

2.6 Body parameter estimation

Nine body parameters chest width (CW), ilium width (IW), hip joint width (HJW), oblique body length (OBL), hip length (HL), withers height (WH), hip height (HH), heart girth (HG) and chest depth (CD) are evaluated. Evaluation indicators mean absolute error (MAE) and mean absolute percentage error (MAPE) are selected as eq(10):

$$\begin{aligned} MAE &= \frac{1}{N_E} \sum_{i=1}^{N_E} |\hat{y}_i - y_i|, \\ MAPE &= \frac{100\%}{N_E} \sum_{i=1}^{N_E} \left| \frac{\hat{y}_i - y_i}{y_i} \right|, \end{aligned} \tag{10}$$

where $\hat{y}_i$ is the estimated value, y_i is the real value, and N_E is the number of experimental data.

3. RESULTS AND DISCUSSION

Table1 and Table2 show the results without pose normalization and pose normalization, respectively. The results after pose normalization show some improvement, with an overall MAPE of 10.27%. The overall MAPE without pose normalization is 10.93%. Judging from the measurement results of body parameters after pose normalization, the MAPE of CW, HH, HG, and CD is all less than 10%, indicating that this method has a high accuracy of measuring these body parameters. The MAPE of IW, HJW, OBL, HL, and WH is greater than 10%. There are

two reasons why pose normalization improves a little over no normalization. One is that the training samples of the SMAL model are animal toys, and the shape of the cattle toy data is different from that of the real cattle. The second is the lack of some body shapes in the training data of the parametric model. These all lead to poor local fitting results when the template mesh is fitted to the cattle observation data, whether normalized or not. Therefore, in the future, it is necessary to use a richer dataset to train the parametric model, which will enhance the generalization ability of the model and ultimately improve the measurement accuracy of body parameters.

Table 1. MAE and MAPE results for nine body parameters without pose normalization.

Indicators	CW	IW	HJW	OBL	HL
MAE(cm)	4.41	4.64	5.70	21.95	4.17
MAPE(%)	9.86	11.14	11.52	13.11	9.06
Indicators	WH	HH	HG	CD	
MAE(cm)	20.31	12.21	9.93	7.09	
MAPE(%)	16.82	9.72	5.59	11.56	

Table 2. MAE and MAPE results for nine body parameters with pose normalization.

Indicators	CW	IW	HJW	OBL	HL
MAE(cm)	4.29	4.57	5.75	22.08	5.49
MAPE(%)	9.59	10.97	11.61	13.15	11.86
Indicators	WH	HH	HG	CD	
MAE(cm)	15.28	9.82	9.63	5.86	
MAPE(%)	12.61	7.83	5.21	9.60	

Fig.8 shows our results in comparison to those of (Du et al., 2022). The overall MAPE of (Du et al., 2022) is 12.82%, where CW, HG, HJW, HL and IW are less accurate than ours. The main reason is that the cattle are scanned while they are in motion, causing the width keypoints to deviate from the same cross-section of the body. Especially when the data is severely missing, it is difficult to get accurate results. In contrast, our method is more robust. But the MAPE of CD, HH, OBL and WH is still high, the root cause is the lack of corresponding body shapes in the training data.

Existing methods for estimating body parameters directly on low-quality point clouds are less robust and lack practicality. Due to the high failure rate of keypoint detection, severely occluded areas and noisy points, the body measurement of some data is limited and cannot be executed. Our proposed body parameter estimation method successfully tackles the challenge of measurement when data quality is low. Moreover, pose normalization can address the difficulty of inconsistent body measurements when animals move freely. Experiments show that our approach has high accuracy and robustness. Each reconstruction takes about 2 minutes, which is much faster than our previous work (Luo et al., 2023).

4. CONCLUSION

Due to the insufficient cooperation of animals and the self-occlusion of quadrupeds, it is difficult to obtain high-precision animal 3D scan datasets. Currently, the lack of robustness and practicality in measuring animal bodies on low-quality point cloud data is a prevalent issue. Aiming at the above problems, based on the statistical shape model, this paper combines the pose normalization method to obtain a topologically consistent 3D surface dataset. The real cattle parametric model is constructed and learned, and the geometric difference of the cattle body is decomposed into two parts: shape and pose. Based on the statistical shape model of cattle, we propose a method for 3D reconstruction and body measurement of real cattle based on low-quality point clouds. The overall pipeline of our proposed body parameter estimation method for cattle can be extended to other quadrupeds.

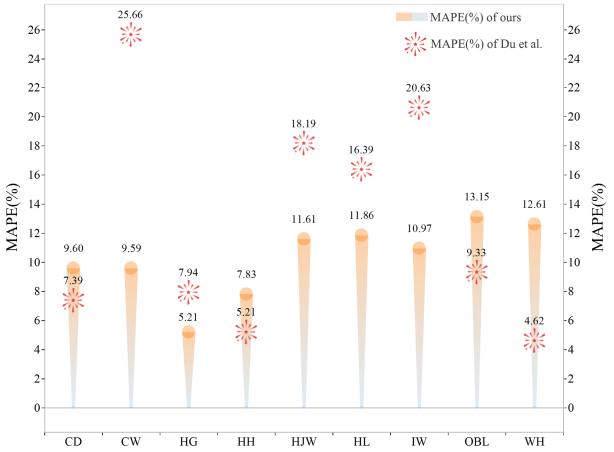


Figure 8. Comparisons with the state-of-the-art method on the same dataset.

ACKNOWLEDGEMENTS

The authors wish to thank everyone who participated in the experiment. Thanks to Ruchay Alexey and his research group for providing us with experimental materials. This research is supported by the National Key Research and Development Program of China (2021YFD1300502-01) and the National Natural Science Foundation of China (grant numbers 42071449, 41601491).

REFERENCES

Anguelov, D., Srinivasan, P., Koller, D., Thrun, S., Rodgers, J., Davis, J., 2005. Scape: shape completion and animation of people. *ACM SIGGRAPH 2005 Papers*, 408–416.

Bartol, K., Bojanić, D., Petković, T., Pribanić, T., 2021. A review of body measurement using 3D scanning. *Ieee Access*, 9, 67281–67301.

Biggs, B., Boyne, O., Charles, J., Fitzgibbon, A., Cipolla, R., 2020. Who left the dogs out? 3d animal reconstruction with expectation maximization in the loop. *Computer Vision – ECCV 2020*, Springer International Publishing, Cham, 195–211.

Biggs, B., Roddick, T., Fitzgibbon, A., Cipolla, R., 2019. Creatures great and smal: Recovering the shape and motion of animals from video. *Computer Vision–ACCV 2018: 14th Asian Conference on Computer Vision, Perth, Australia, December 2–6, 2018, Revised Selected Papers, Part V 14*, Springer, 3–19.

Bogo, F., Romero, J., Loper, M., Black, M. J., 2014. Faust: Dataset and evaluation for 3d mesh registration. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 3794–3801.

Cashman, T. J., Fitzgibbon, A. W., 2012. What shape are dolphins? building 3d morphable models from 2d images. *IEEE transactions on pattern analysis and machine intelligence*, 35(1), 232–244.

Du, A., Guo, H., Lu, J., Su, Y., Ma, Q., Ruchay, A., Marinello, F., Pezzuolo, A., 2022. Automatic livestock body measurement based on keypoint detection with multiple depth cameras. *Computers and Electronics in Agriculture*, 198, 107059.

Horn, B., Hilden, H., Negahdaripour, S., 1988. Closed-Form Solution of Absolute Orientation using Orthonormal Matrices. *Journal of the Optical Society of America A*, 5, 1127-1135.

Kanazawa, A., Kovalsky, S., Basri, R., Jacobs, D., 2016. Learning 3d deformation of animals from 2d images. *Computer Graphics Forum*, 35number 2, Wiley Online Library, 365–374.

Kingma, D. P., Ba, J., 2014. Adam: A Method for Stochastic Optimization. *CoRR*, abs/1412.6980.

Kongsro, J., 2014. Estimation of pig weight using a Microsoft Kinect prototype imaging system. *Computers and Electronics in Agriculture*, 109, 32–35.

Le, B. H., Deng, Z., 2012. Smooth Skinning Decomposition with Rigid Bones. *ACM Trans. Graph.*, 31(6).

Li, C., Ghorbani, N., Broomé, S., Rashid, M., Black, M. J., Hernlund, E., Kjellström, H., Zuffi, S., 2021. hSMAL: Detailed horse shape and pose reconstruction for motion pattern recognition. *arXiv preprint arXiv:2106.10102*.

Luo, X., Hu, Y., Gao, Z., Guo, H., Su, Y., 2023. Automated measurement of livestock body based on pose normalisation using statistical shape model. *Biosystems Engineering*, 227, 36-51.

Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., Bethge, M., 2018. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience*, 21(9), 1281-1289.

Pezzuolo, A., Milani, V., Zhu, D., Guo, H., Guercini, S., Marinello, F., 2018. On-Barn Pig Weight Estimation Based on Body Measurements by Structure-from-Motion (SfM). *Sensors*, 18(11), 3603.

Ruchay, A., Kober, V., Dorofeev, K., Kolpakov, V., Miroshnikov, S., 2020. Accurate body measurement of live cattle using three depth cameras and non-rigid 3-D shape recovery. *Computers and Electronics in Agriculture*, 179, 105821.

Shuai, S., Ling, Y., Shihao, L., Haojie, Z., Xuhong, T., Caixing, L., Aidong, S., Hanxing, L., 2020. Research on 3D surface reconstruction and body size measurement of pigs based on multi-view RGB-D cameras. *Computers and Electronics in Agriculture*, 175, 105543.

Vicente, S., Agapito, L., 2013. Balloon shapes: Reconstructing and deforming objects with volume from images. *2013 International Conference on 3D Vision-3DV 2013*, IEEE, 223–230.

Wang, K., Guo, H., Ma, Q., Su, W., Chen, L., Zhu, D., 2018. A portable and automatic Xtion-based measurement system for pig body size. *Computers and electronics in agriculture*, 148, 291–298.

Weiss, A., Hirshberg, D., Black, M. J., 2011. Home 3d body scans from noisy image and range data. *2011 International Conference on Computer Vision*, IEEE, 1951–1958.

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. *Computers and Electronics in Agriculture*, 44(3), 203–222.

Zuffi, S., Kanazawa, A., Berger-Wolf, T., Black, M. J., 2019. Three-d safari: Learning to estimate zebra pose, shape, and texture from images" in the wild". *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 5359–5368.

Zuffi, S., Kanazawa, A., Jacobs, D. W., Black, M. J., 2017. 3d menagerie: Modeling the 3d shape and pose of animals. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 6365–6373.