https://doi.org/10.2352/ISSN.2470-1173.2020.12.FAIS-174

Cattle Identification and Activity Recognition by Surveillance Camera

Haike Guan, Naoki Motohashi, Takashi Maki, Toshifumi Yamaai
Applied AI Research & Development Center, Innovation R&D Division, Ricoh Co. LTD.; Ebina, Kanagawa/Japan

Abstract

In cattle farm, it is important to monitor activity of cattle to know their health condition and prevent accidents. Sensors were used by conventional methods to recognize activity of cattle, but attachment of sensors to the animal may cause stress. Camera was used to recognize activity of cattle, but it is difficult to identify cattle because cattle have similar appearance, especially for black or brown cattle.

We propose a new method to identify cattle and recognize their activity by surveillance camera. The cattle are recognized at first by CNN deep learning method. Face and body areas of cattle, sitting and standing state are recognized separately at same time. Image samples of day and night were collected for learning model to recognize cattle for 24-hours.

Among the recognized cattle, initial ID numbers are set at first frame of the video to identify the animal. Then particle filter object tracking is used to track the cattle. Combing cattle recognition and tracking results, ID numbers of the cattle are kept to the following frames of the video. Cattle activity is recognized by using multi-frame of the video. In areas of face and body of cattle, active or static activities are recognized. Activity times for the areas are outputted as cattle activity recognition results.

Cattle identification and activity recognition experiments were made in a cattle farm by wide angle surveillance cameras. Evaluation results demonstrate effectiveness of our proposed method.

Introduction

In cattle farm, it is important to monitor activity of cattle to know their health condition and prevent accidents. Sensors were used by conventional methods to recognize activity of cattle, but attachment of sensors to the animal may cause stress. The sensors may easily fail or be destroyed by collision.

Cameras were used to recognize activity of cattle, but it is difficult to identify cattle when cattle have similar appearance, especially for black or brown cattle, as shown in Figure 1.

Related Work

GPS sensors were used by conventional methods to recognize activities of cattle [1], but attachment of sensors to the animal may cause stress. The sensors may easily fail or be destroyed by collision. These sensors can not make difference between movement of head and whole body of the cattle either.

Cameras were used to recognize activity of cattle by conventional methods. Cattle were identified by using patterns on their bodies [2], but it is difficult to identified cattle that do not have pattern on their body, such as black or brown cattle. Cattle position detection and cattle tracking by camera method was proposed to recognize activities of cattle [3][4]. But this method can not identify specific cattle's activities, when there are multiple cattle in the place.

Image analysis method was proposed to recognize activities of cattle [5]. Method by detecting cattle from their shape of body and tracking them to recognize their activities was proposed [6]. Multiple time-series image frames were used to recognize activities of cattle [7][8][9]. But it is difficult for these methods to identify specific cattle's activities when there are multiple cattle in same place.

Proposed Method

We propose a new method to identify cattle and recognize their activity by surveillance camera. The cattle are recognized at first by CNN deep learning method. Face and body areas of cattle, sitting and standing state are recognized separately at same time. Image samples of day and night were collected for learning model to recognize cattle for 24-hours.

Among the recognized cattle, initial ID numbers are set at first frame of the video to identify the animal. Then particle filter object tracking method is used to track the cattle. Combing cattle recognition and tracking results, ID numbers of the cattle are kept to the following frames of the video. Cattle activity is recognized by using multi-frame of the video. In areas of body and face of cattle, active or static activities are recognized. Activity times for the areas are outputted as cattle activity recognition results.

Cattle identification and activity recognition experiments were made in a cattle farm by wide angle surveillance cameras. Evaluation results demonstrate effectiveness of our proposed method.

Cattle Recognition and Identification

In our proposed method, as shown in Figure 2, cattle recognition is made at first to locate the cattle. In this step a single image that is inputted from video of surveillance camera is used for the recognition.

In cattle recognition, as shown in Figure 3, recognition is made by recognizing whole body area and face area.

Sit and stand states of cattle body area are also recognized. We use CNN deep learning method to do the recognition [10]. Face and whole body areas, sit, stand states, are recognized at same time. In Figure 3, yellow rectangle means face area. Blue rectangle means whole body area with stand state. Green rectangle means whole body area with sit state.

Face and Body Recognition

To recognize activity of cattle, we recognize activity of face area and body area separately. From activity condition of whole body, we can know health condition of the cattle. From activity of face area, we can know if the cattle are eating or chewing. If activity time of eating is too short, the cattle maybe not healthy. Face and body area recognition is made by using RefineDet deep

Figure 1. Cattle image of surveillance camera

learning neural network [10]. Data of cattle's face area and body area are used to train model.

Sit and Stand State Recognition

Body area are also separated as sit and stand states. Target areas of cattle are divided to 3 type, face area, body area of sit state and body area of stand state. Training data of face, body of sit state and body of stand state are used to train the cattle recognition model. In a test single image, 3 type area can be recognized at same time.

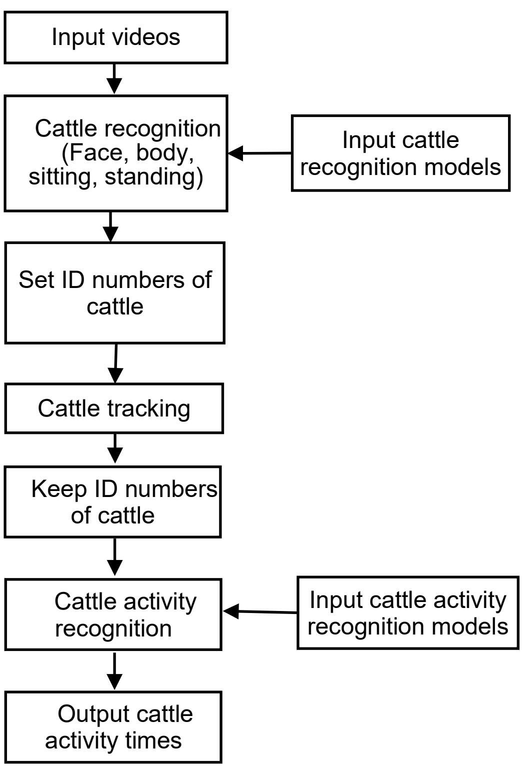


Figure 2 Flow chart of proposed method

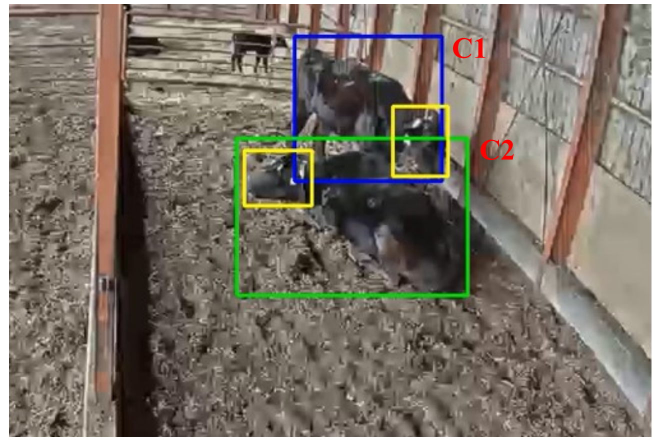


Figure 3. Cattle recognition of face area and body area of sit and stand states.

Green color of rectangle means whole body of sit state, blue color means whole body of stand state, and yellow color means face area.

Cattle Identification

In our proposed method, identification number of cattle is set at first frame of the video. As shown in Figure 3, cattle are set as C1 and C2 which means ID number of the cattle. The ID number is kept by combining cattle track and cattle recognition results.

Cattle Tracking

In our proposed method, object tracking is used to keep ID number of the cattle. After setting ID number of cattle at first frame, if object tracking is success, the ID number is kept. If cattle tracking failed, cattle recognition results are used to restart the cattle ID number.

Because the cattle's shape changes very frequently, particle filter tracking method is used in our method. State of the target is defined by Equation (1), where S is state of the target. x, y is position, $v_x$and v_x are speed, H_x, H_y are size of the target. M is multiple factor. Propagation function of Equation (2) is expanded to Equation (3) to Equation (9)

$$S(x, y, v_{x,} v_y, H_x, H_y, \mathrm{M}) \tag{1}$$

$$\boldsymbol{S}_k = f_k(\boldsymbol{S}_{k-1}, n_{k-1}) \tag{2}$$

$$x_k = x_{k-1} + v_{xk-1} * \Delta t \tag{3}$$

$$y_k = y_{k-1} + v_{yk-1} * \Delta t \tag{4}$$

$$v_{xk} = v_{xk-1} \tag{5}$$

$$v_{yk} = v_{yk-1} \tag{6}$$

$$M_k = M_{k-1} \tag{7}$$

$$H_{xk} = H_{xk-1}(1 + M_{k-1}) \tag{8}$$

$$H_{yk} = H_{yk-1}(1 + M_{k-1}) \tag{9}$$

Measurement equation of Equation (10) is written as Equation (11), which is color histogram of target area.

$$\boldsymbol{z}_k = \boldsymbol{h}_k(\boldsymbol{S}_k, \boldsymbol{n}_k) \quad (10)$$

$$p_y^u = f\sum_{i=1}^{I} k\left(\frac{\|y-x_i\|}{M}\right)\delta[h(x_i)-u] \quad (11)$$

$$k(r) = \begin{cases} 1-r^2 : r<1 \\ 0 : otherwise \end{cases} \quad (12)$$

K(r) is kernel used to effect of edge area. Gaussian random noise is added to each item in Equation (13) to Equation (19). r_i (i=1, …, 7) is random noise.

$$x_k = x_{k-1} + v_{xk-1} * \Delta t + r_1 \quad (13)$$

$$y_k = y_{k-1} + v_{yk-1} * \Delta t + r_2 \quad (14)$$

$$v_{xk} = v_{xk-1} + r_3 \quad (15)$$

$$v_{yk} = v_{yk-1} + r_4 \quad (16)$$

$$M_k = M_{k-1} + r_5 \quad (17)$$

$$H_{xk} = H_{xk-1}(1 + M_{k-1}) + r_6 \quad (18)$$

$$H_{yk} = H_{yk-1}(1 + M_{k-1}) + r_7 \quad (19)$$

$$\rho[p,q] = \sum_{u=1}^{m} \sqrt{p^{(u)} q^{(u)}} \quad (20)$$

Similarity of Equation (20) is calculated, which is used as similarity between target area and model area.

$$w_k = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{d^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(1-\rho[p_S(k),q])}{2\sigma^2}} \quad (21)$$

where,

$$d = \sqrt{1-\rho[p,q]} \quad (22)$$

$$S = \sum_{k=0}^{N} w_k S_k \quad (23)$$

State parameter $\boldsymbol{S}$ is calculated by Equation (23). $\boldsymbol{S}$ is tracking results.

If similarity $\rho$of Equation (20) is bigger than a threshold value, cattle tracking is successful. The cattle ID number is kept.

Identification Number Keep and Restart

If the cattle are tracked successfully, the ID do not change. State condition parameters $\boldsymbol{S}$ are saved. If cattle tracking failed, cattle recognition is made again. If similarity ρ of Equation (20) between recognized result area and saved successfully tracked results area just before failed frame, is bigger than a threshold value, the cattle ID number is restarted.

Action recognition

In cattle recognition and tracking results of face area and body area, activity recognition is made separately. N frames are used for activity recognition. Activity is divided as static activity and active activity. Feature vector is calculated from the N frames' images to do activity recognition.

Feature Points Detection

To calculate feature vector, feature points are selected from target area at first. Target area is face area and body area of cattle. The target area is divided to blocks. In each block of the target area, contrast feature of $\boldsymbol{t}$, is calculated from Equation (24) and Equation (25). If $\boldsymbol{t}$ is bigger that a threshold value, the block is selected to calculate feature vector.

$$S_{ij} = \int \frac{\partial^2 f(x,y)}{\partial i \partial j} dxdy \qquad i,j \in x,y \quad (24)$$

$$t = S_{xx}S_{yy} - S_{xy}^2 \quad (25)$$

In each block of target area that are selected as feature points, motion vector of the points is calculated by Equation (26) and Equation (27). f_1 and $f_2$are intensity of pixels in 2 continue frames.

$$f_2(x,y) = f_1(x - u\Delta t, y - v\Delta t) \quad (26)$$

$$u\frac{\partial f(x,y)}{\partial x} + v\frac{\partial f(x,y)}{\partial y} + \frac{\partial f(x,y)}{\partial t} \approx 0 \quad (27)$$

$$u = \frac{S_{yt}S_{xy} - S_{xt}S_{yy}}{S_{xx}S_{yy} - S_{xy}^2} \quad (28)$$

$$v = \frac{S_{xt}S_{xy} - S_{yt}S_{xx}}{S_{xx}S_{yy} - S_{xy}^2} \quad (29)$$

Motion vector (u, v) of 1 block in the target area are calculated by Equations of (28) and Equation (29).

For N frames image that are used to recognized activity, N-1 motion vector are calculated for 1 block.

Movement Vector Feature Calculation

Average motion vector of (u, v) are used to recognize activity of cattle. For 1 target area average motion vector (u, v) are calculated by Equation (28) and Equation (29).

Active and Static Recognition

Active and static activity is recognized in the proposed method. N frames are used to do recognition. Average motion vectors (u, v) is calculated. If motion vector is bigger than a threshold value, it is recognized as activity. If the threshold value is smaller than the threshold, it is recognized as static activity.

Activity Recognition Results Combination

Action recognition results of face and body area are obtained separately. If recognition results of body area is sit, recognition results are static sit or active sit. If recognition results of body is stand, recognition results are static stand or active stand. Face area recognition results are static or active.

Experiment results

Experiment was made by using 3 surveillance cameras as shown in Figure 4. 24 hours' video data was recorded and used for the experiment. 3 cattle were used as surveillance target. Experiment setup is illustrated in Figure 4.

Cattle Recognition Model Training

12 hour's data were used for training cattle recognition model, which was selected from 24 hour's data by 1 hour's interval. 3 surveillance cameras' video data were used to pick up training data. Training data included day and night data. Every 2 minutes 1 frame were picked up for labeling. 1100 frame image data were labeled for training.

6 hour's video data of 3 surveillance cameras were used for picking up validation data, which included day and night time data. Every 2 minutes 1 frame were picked up and 550 frames were picked for validation.

For both training and validation data, in each frame, face area and body area were labeled by rectangle. For body area, sit or stand state were also labeled.

The labeled leaning data and validation data were used to train cattle recognition model. RefineDet network was used to train the cattle recognition model [10].

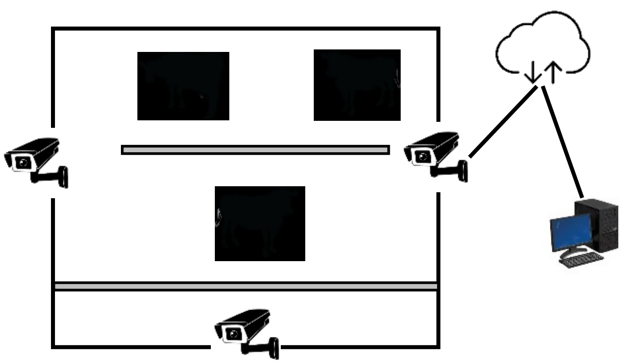


Figure 4. Experiment setup for cattle activity recognition experiments. 3 surveillance cameras were set up. Video data is saved to computer by network.

Cattle Recognition and state Recognition Experiment

6 hour's video data of day and night were used to test cattle recognition model. Cattle recognition results of day time is shown in Figure 4. Blue rectangle means cattle body area of stand state recognition results. Green rectangle means cattle body area of sit state recognition results. Yellow rectangle means cattle face area recognition results. Cattle recognition results of night time are shown in Figure 5.

Cattle Activity Recognition Experiment Results

Cattle activity recognition was made by using cattle recognition and cattle tracking results. In face area and body area, 25 frames image are used for recognizing activity. Figure 6 illustrates cattle activity recognition results of day time. Figure 7 illustrates cattle activity recognition results of day time. Red color rectangle means active activity. Other color rectangle means static activity.

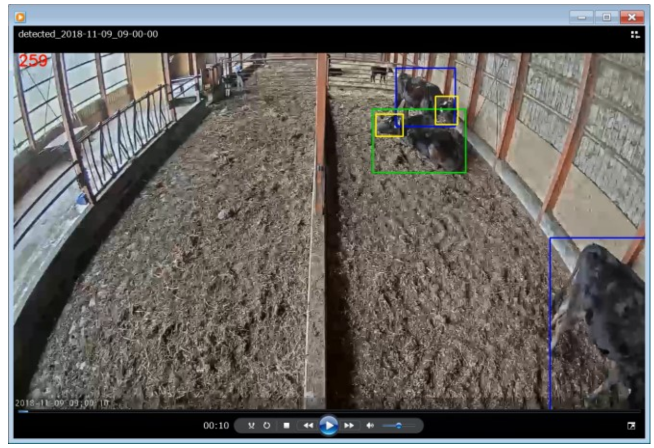


Figure 4. Cattle recognition result of day time. Green color of rectangle means whole body and sit state, blue color means whole body and stand state, and yellow color means face area.

Figure 5. Cattle recognition result of night time. Green color of rectangle means whole body and sit state, blue color means whole body and stand state, and yellow color means face area.

Figure 6. Cattle activity recognition results of day time. Red rectangle means active activity. Rectangle of other color means static activity

Figure 7. Cattle activity recognition results of night time. Red rectangle means active activity. Rectangle of other color means static activity.

Cattle Activity Recognition Evaluation Results

2 hours day time and 1 hour night time video data were used to evaluate cattle activity recognition results. Ground truth video were labeled by people by 1 second unit. Cattle recognition results were outputted by 1 second unit. Body area recognition results were used for evaluation. Evaluation results is illustrated in Table 1.

Table 1. Cattle activity recognition evaluation results

		Recognition rate (%)	False positive rate (%)
Day time	Active	68.8	4.1
	Static	87.4	2.9
Night time	Active	60.0	1.8
	Static	80.8	18.4

Conclusion

We propose a new method to identify cattle and recognize their activity by surveillance camera. The cattle are recognized at first by CNN deep learning method. Face and body areas of cattle, sitting and standing state are recognized separately at same time. Image samples of day and night time were collected for learning model to recognize cattle for 24-hours.

Among the recognized cattle, initial ID numbers are set at first frame of the video to identify the animal. Then particle filter object tracking is used to track the cattle. Combing cattle recognizing and tracking results, ID numbers of the cattle are kept to the following frames of the video. Cattle activity is recognized by using multi-frame of the video. In areas of face and body of cattle, active or static activities are recognized. Activity times for the areas are outputted as cattle activity recognition results.

Cattle identification and activity recognition experiments were made in a cattle farm by wide angle surveillance cameras. Evaluation results demonstrate effectiveness of our proposed method.

References

[1] Torben Godsk and Mikkel Baun Kjargaard, "High Classification Rates for Continuous Cow Activity Recognition Using Low-Cost GPS Positioning Sensors and Standard Machine Learning Techniques", 11th Industrial Conference, ICDM 2011, pp. 174-122, 2011

[2] Thi Thi Zin, Member, IAENG, Cho Nilar Phyo, Pyke Tin, Hiromitsu Hama, Ikuo Kobayashi,"Image Technology based Cow Identification System Using Deep Learning", Proceedings of the International MultiConference of Engineers and Computer Scientists 2018 Vol I, pp320-323, 2018

[3] Trung-Kien Dao, Thi-Lan Le, David Harle, Paul Murray, Christos Tachtatzis, Stephen Marshall, Craig Michie, Ivan Andonovic," AUTOMATIC CATTLE LOCATION TRACKING USING IMAGE PROCESSING", 23rd European Signal Processing Conference, pp2681-2685

[4] Alberto Rivas, Pablo Chamoso, Alfonso González-Briones and Juan Manuel Corchado, "Detection of Cattle Using Drones and Convolutional Neural Networks", Sensor 2018, pp1- 15, 2018

[5] Gu Jingqiu, Wang Zhihai, Gao Ronghua, Wu Huarui, "Cow behavior recognition based on image analysis and activities", Int J Agric & Biol Eng , Vol. 10 No.3, pp165-174, 2017

[6] Abozar Nasirahmadi, Sandra A. Edwardsa, Barbara Sturma ,"Implementation of machine vision for detecting behaviour of cattle and pigs", Livestock Science 202 (2017), pp25-38, 2017

[7] Eui-Ju Heo, Sung-Jin Ahn, and Kang-Sun Choi, "Real-Time Cattle Action Recognition for Estrus Detection", KSII TRANSACTIONS ON INTERNET AND INFORMATION SYSTEMS VOL. 13, NO. 4, Apr. 2019, pp2148-2162, 2019

[8] Sung-Jin Ahn, Dong-Min Ko, Eui-Ju Heo, and Kang-Sun Choi, "Real-Time Cow Action Recognition based on Motion History Image Feature", 2018 IEEE International Conference on Consumer Electronics (ICCE), pp1-3, 2018

[9] Thi Thi Zin ,"Markov Chain Techniques for Cow Behavior Analysis in Video-based Monitoring System", Proceedings of the International MultiConference of Engineers and Computer Scientists 2018 Vol I IMECS 2018, pp339-342, 2018

[10] Shifeng Zhang, Longyin Wen, Xiao Bian, Zhen Lei1, Stan Z. Li, "Single-Shot Refinement Neural Network for Object Detection", Proceedings CVPR IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 4203–4212, 2018.

Author Biography

Haike Guan received his PhD in information processing from Tokyo Institute of Technology (1995). Since then he has worked in Ricoh Company LTD. and currently is a senior specialist researcher in Innovation Research and Development Division. His work has focused on research and development in image processing, pattern recognition, computer vision, human activity recognition and related fields.

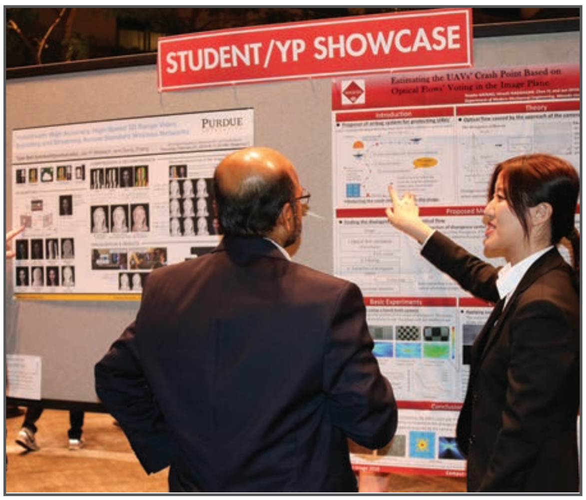
JOIN US AT THE NEXT EI!

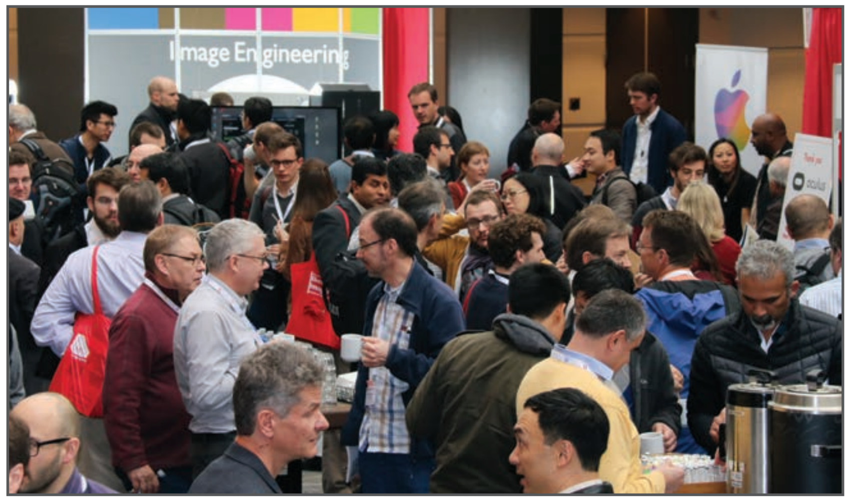
IS&T International Symposium on

Electronic Imaging

SCIENCE AND TECHNOLOGY

Imaging across applications . . . Where industry and academia meet!





• SHORT COURSES • EXHIBITS • DEMONSTRATION SESSION • PLENARY TALKS •
• INTERACTIVE PAPER SESSION • SPECIAL EVENTS • TECHNICAL SESSIONS •

www.electronicimaging.org

IS&T
imaging.org