Automatic monitoring method of cow ruminant behavior based on spatio-temporal context learning

Yujuan Chen[1,2,3], Dongjian He[1,2,3], Huaibo Song[1,2,3*]

(1. *College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling, Shaanxi 712100, China*;
2. *Key Laboratory of Agricultural Internet of Things, Ministry of Agriculture, Yangling, Shaanxi 712100, China*;
3. *Shaanxi Key Laboratory of Agricultural Information Perception and Intelligent Service, Yangling, Shaanxi 712100, China*)

Abstract: Automatic monitoring of cow rumination has great significance in the development of modern animal husbandry. In order to solve the problem of high real-time requirement of ruminant behavior monitoring, a tracking method based on STC (Spatio-Temporal Context) learning was carried out. On the basis of cow's mouth region extraction, the spatial context model between target object and its local surrounding background was built based on their spatial correlations by solving the deconvolution problem, and the learned spatial context model was used to update the STC learning model for the next frame. Tracking in the next frame was formulated by computing a confidence map as a convolution problem that integrates the STC learning information, and the best object location could be estimated by maximizing the confidence map. Then the target scale was estimated based on the confidence evaluation. Finally, accurate tracking of the mouth movement trajectory was realized. To verify the effectiveness of the proposed method, the performance of the algorithm was tested using 20 video sequences. Besides, the tracking results were compared with the Mean-shift algorithm. The results showed that the average success rate of STC learning monitoring algorithm was 85.45%, which was 9.45% higher than the Mean-shift algorithm, the detection rate of STC learning monitoring algorithm was 18.56 s per video, which was 22.08% higher than that of the Mean-shift algorithm. The results showed that the fast tracking method based on STC learning monitoring algorithm is effective and feasible.
Keywords: dairy cow, rumination, intelligent monitoring, STC learning, Mean-shift
DOI: 10.25165/j.ijabe.20181104.3509

1 Introduction

Ruminant activity is an important index reflecting the health of animal rumens. When cows suffer from diseases, their rumination time will decrease significantly. The influence of a variety of diseases will affect the rumination time uniquely[1]. For example, for cows with rumen tympanites, abomasum displacements or digestive tract diseases, such as enteritis, the rumination time decreases the most, followed by postpartum paralysis and reproductive diseases[2]. In addition to parity, the number of days spent being milked also has an effect on the rumination time[3]. Monitoring cow rumination has great significance in the development of modern animal husbandry. At present, the behavior of cow rumination is primarily observed by farmers with high labor costs and work intensity, which often fail to meet the management needs of large-scale farming in addition to other problems. Some scholars focused on automated ways of recording the daily behavior of animals with rumens. Monitoring rumination behavior is still limited by the impact of complex systems and a variety of physical factors. For example, noises caused by the collision, friction and shaking among the sensors will affect the accuracy and reliability of the data. Therefore, the

automatic tracking and monitoring of cow ruminating behavior based on video analysis has important significance, and has attracted increasing attention.

Nowadays, the widely used target tracking algorithms are Mean-shift algorithm, inter-frame differentiation detection method, Cam-shift algorithm, optical flow method and so on. For Mean-shift target tracking algorithm, the target color histogram is always used as the search feature, and the real position is achieved by iterative Mean-shift vectors to make the algorithm converge to the target, so as to achieve the goal of object tracking. However target color histogram features in the feature description are slightly deficient, and lack of spatial information, and the target tracking accuracy has been greatly affected[4-6]. Detection method of inter-frame differentiation is one of the most widely used methods on moving target detection and segmentation. For this method, the moving regions in the image are extracted by using closed value based on time difference between two adjacent frames and the third frame. Since the method cannot extract the whole region of the object, it is easy to cause holes in the moving target, which leads to the difficulty of extracting information of the connected target domain[7-10]. The optical flow method has been used to calculate the spatial and temporal gradient of the image sequences for estimating the motion fields under the condition of appropriate smoothness constraints. The feature point of optical flow method has the characteristics of less amount of calculation, high speed and low flexibility, but it is difficult to extract shape of moving object accurately by feature matching. The optical flow method does not require the prior knowledge of any scene information and can detect the moving object, it can also be used to detect moving objects in background motion and camera motion

conditions, but its anti-noise performance is poor, and cannot detect moving targets' contours accurately[11-13]. The Cam-shift algorithm is easy to converge to the non-target area when tracking the color changes and large area color interference, but the background is too complex to converge to the target area, which would lead to the failure of object tracking[14-16]. The STC learning algorithm can efficiently use information between frames and local context information of the target to calculate and estimate target position of the likelihood confidence map by considering the local context background information of target area. The STC learning algorithm can improve the computing speed by using the Fourier transform in spatial context model, which can obtain good real-time performance and robustness[17,18].

Rumination can be divided into following four stages: regurgitation, where processed cuds return from the rumen to mouth, chewing, saliva mixing, and swallowing. Based on the cattle-chewing phenomenon, the videos were all captured as continuous frames on the basis of the constructed platform used for capturing bovine ruminant video. By using moving object tracking algorithm based on STC learning algorithm, we could track the moving of the cow's mouth accurately, and draw the curve of mouth movement of cows so as to realize the intelligent monitoring of dairy ruminant behavior, and the results could provide new reference for the dairy ruminant behavior monitoring.

2 Materials and methods

2.1 Materials

All the videos were collected from five different cows cultivated on a commercial dairy farm in Yangling, Shaanxi Province, China on July 2013. The videos were captured by an integrated IP camera (DS-2DM1-714, Hikvision Inc., Hangzhou, China) at a resolution of 704×574 pixels, the color video format was set to PAL (Phase Alternation Line) with a frame rate of 25 frames/s and a code rate of 2048 kb/s. The test cows were American Holstein cows in mid lactation. Considering that cows always ruminate in prone states in most cases, videos were all taken when cows were in a prone state, and it would be more typical and representative[2]. The length of the six captured videos were all 16'49". To verify the effectiveness of the algorithm, the original

video data were divided into 20 video segments. The length of each video segment was approximately 30". In these video segments, the captured video frame rate in this study was 10 frames/s. Thus it is convenient to realize a real-time detection of ruminant behaviors. The videos contain complicated cow behaviors, such as raising head, turning head, rumination and static state, making it difficult to accurately monitor the cow's ruminant behavior. The statistics of the dairy cows' video information obtained were shown in Table 1.

Table 1   Video information of cow's rumination behaviors

No.	Length of videos	Time	Total frames	Behaviors
1	16'50"	Day	10100	Raising head, turning head, rumination and static state
2	16'49"	Day	10090	Raising head, turning head, rumination and static state
3	16'49"	Day	10090	Raising head, turning head, rumination and static state
4	16'49"	Day	10090	Raising head, turning head, rumination and static state
5	16'49"	Night	10090	Raising head, turning head, rumination and static state
6	16'49"	Night	10090	Raising head, turning head, rumination and static state

2.2 STC learning method

The STC learning algorithm[19] is a simple but effective method for target tracking. The main idea is modeling for the tracking target and the spatio-temporal relation of its local context through a Bayesian framework. Then the spatio-temporal relationship and the focus and attention features of the biological vision system are integrated to evaluate the confidence map of the position of the target in the new frame, the maximum confidence position is the target position of the new frame. The STC learning algorithm was usually used in various target tracking and detecting[20-22].

2.2.1 Cow ruminant tracking process based on STC learning algorithm

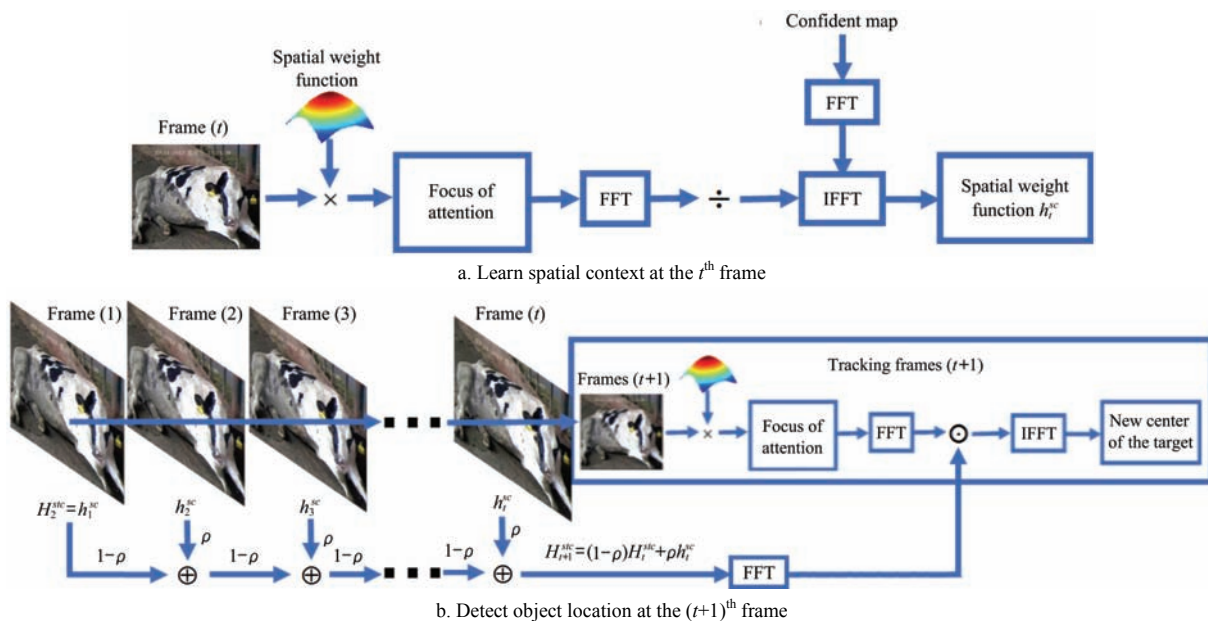The tracking process of dairy ruminant behavior using the STC learning algorithm is shown in Figure 1.

a. Learn spatial context at the tth frame

b. Detect object location at the $(t+1)^{th}$ frame

Figure 1   Basic flow chart of cow ruminant behavior based on STC learning algorithm[19]

The tracking problem is formulated by computing a confidence map which estimates the object location likelihood, as is shown in Equation (1):

$$c(x) = \sum_{c(z)\in X^c} P(x\,|\,c(z),o)P(c(z)\,|\,o) \tag{1}$$

where, $x \in R^2$ is an object location and o denotes the object presented in the scene. In the current frame, we have the object location x^*. The context feature set is defined as $X^c=\{c(z)=(I(z),z)|z\in\Omega_c(x^*)\}$, where $I(z)$ denotes image intensity at location z and $\Omega_c(x^*)$ is the neighborhood of location x^*. The conditional probability $P(x|c(z),o)$ models the spatial relationship between the object location and its context information which helps resolve ambiguities when the image measurements allow different interpretations, and $P(c(z)|o)$ is a context prior probability which models appearance of the local context.

The spatial context model is represented by a conditional probability equation, as is shown in Equation (2):

$$P(x|c(z),o)=h^{SC}(x-z) \tag{2}$$

where, $h^{SC}(x-z)$ is a function with respect to the relative distance and direction between object location x and its local context location z, thereby encoding the spatial relationship between an object and its spatial context.

The context prior probability is simply modeled by Equation (3):

$$P(c(z)|o)=I(z)w_\sigma(z-x^*) \tag{3}$$

where, $I(\cdot)$ is image intensity and presents appearance of context, and $w_\sigma(\cdot)$ is a weighted function defined by Equation (4):

$$w_\sigma(z-x^*)=ae^{-\frac{|z-x^*|^2}{\sigma^2}} \tag{4}$$

where, a is a normalization constant that restricts $P(c(z)|o)$ in Equation (3) to range from 0 to 1, and it satisfies the definition of probability and σ is a scale parameter.

The target position confidence map is modeled as Equation (5):

$$c(x)=be^{-\left|\frac{x-x^*}{\alpha}\right|^\beta}=\sum_{z\in\Omega_c(x^*)}h^{SC}(x-z)I(z)w_\sigma(z-x^*)$$
$$=h^{SC}(x)\otimes(I(x)w(x-x^*)) \tag{5}$$

where, b is a normalization constant, α is a scale parameter, and β is a shape parameter. Equation (5) can be transformed to the frequency domain in which the Fast Fourier Transform (FFT) algorithm can be used for fast convolution. That is,

$$F(be^{-\left|\frac{x-x^*}{\alpha}\right|^\beta})=F(h^{SC}(x))\,e\,F(I(x)w(x-x^*)) \tag{6}$$

where, F denotes the FFT function and $\odot$ is the element-wise product. Therefore, we can get Equation (7):

$$h^{SC}(x)=F^{-1}\left(\frac{F(be^{-\left|\frac{x-x^*}{\alpha}\right|^\beta})}{F(I(x)w(x-x^*))}\right) \tag{7}$$

where, F^{-1} denotes the inverse FFT function.

Each frame calculates the spatial context h_t^{sc} of the current frame according to the target state, and the STC model is updated by Equation (8):

$$H_{t+1}^{stc}=(1-\rho)H_t^{stc}+\rho h_t^{sc} \tag{8}$$

where, ρ is a learning parameter.

Calculate target position confidence map in the $(t+1)^{th}$ frame by Equation (9):

$$c_{t+1}(x)=F^{-1}(F(H_{t+1}^{stc}(x))\,e\,F(I_{t+1}(x)w_{\sigma_t}(x-x^*))) \tag{9}$$

The object location x_{t+1}^* in the $(t+1)^{th}$ frame is determined by maximizing the new confidence map, as is shown in Equation (10):

$$x_{t+1}^*=\arg\max_{x\in\Omega_c(x_t^*)}c_{t+1}(x) \tag{10}$$

We propose the scale update scheme as Equation (11)[19]:

$$\begin{cases} s_t'=\sqrt{\dfrac{c_t(x_t^*)}{c_{t-1}(x_{t-1}^*)}}, \\ \overline{s_t}=\dfrac{1}{n}\sum_{i=1}^n s_{t-i}', \\ s_{t-1}=(1-\lambda)s_t+\lambda\overline{s_t}, \\ \sigma_{t+1}=s_t\sigma_t, \end{cases} \tag{11}$$

where, $c_t(\cdot)$ is the confident map that is computed by Equation (9), and s_t' is the estimate scale between two consecutive frames. To avoid oversensitive adaptation and reduce noise introduced by estimation error, the estimated target scale s_{t+1} is obtained through filtering in which $\overline{s_t}$ is the average of the estimated scales from n consecutive frames, and $\lambda>0$ is a fixed filter parameter (similar to ρ in Equation (8)).

2.2.2 Obtain the tracking curve by using the normalized distance of tracking target

An example of the tracking curve is shown in Figure 2. The abscissa represents the number of frames, the ordinate represents normalized distance of the central point of each target tracking region, small red crosses represent the marked chewing times.
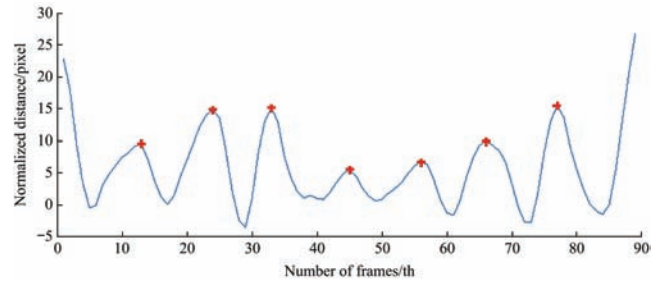

Figure 2　Tracking curve of cow's mouth movement area

2.2.3 Cubic interpolation spline interpolation of ruminant movement

It can be found in Figure 2 that there is a certain jitter in the tracking curve when the STC learning algorithm is used to track the target and the target contour is not smooth enough. In order to realize the automatic monitoring of cow ruminant behavior, the cubic interpolation spline method was used in this study to fit it, and the automatic calculation of the rumination number was achieved.

A smooth target tracking curve according to Figure 2 using cubic interpolation spline is shown in Figure 3. The abscissa represents the number of frames, the ordinate represents the normalized distance. The blue point represents the value of the normalized distance variable, and the red curve represents tracking curve of smooth sample fitting. Each sharp peak is a chewing process in Figure 3.

2.3 Mean-shift algorithm

Mean-shift algorithm is a kind of non-parametric estimation method. It is realized by an iterative process. First, we calculate the mean offset value of the current point and move the point to the mean offset value. Next, we set that point as a new starting point and repeat that process until meeting certain criteria. The Mean-shift algorithm is widely used in such applications as clustering, tracking[23-26], segmentation[27,28], and image smoothing.

For example, Mean-shift was applied to image de-noising algorithm by Zhao[29] and vehicle tracking by Li[30].   Mean-shift

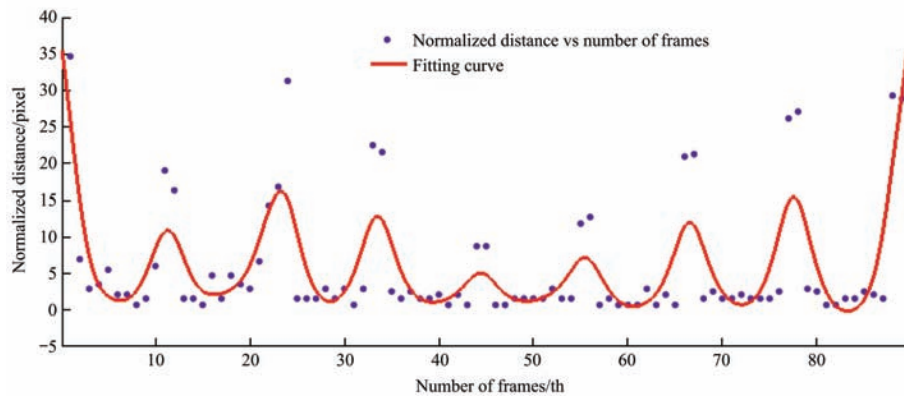algorithms were also used in image segmentation[31], as well as supermarket customer movement tracking[32].

Figure 3   Fitting curve of ruminant movement based on three interpolation spline interpolation

2.4   Evaluating indicator

The error rate expression is shown in Equation (12).   The monitoring success rate is $V=1-R_1(R_2)$.   To ensure the accuracy of the data extracted, the 6 segment video data were divided into 20 small segments to test the data.

$$R_1(R_2)=(|n_1-n_0|)/n_0 \tag{15}$$

where, R_1 represented the monitoring error rate of Mean-shift algorithm. R_2 represented the monitoring error rate of STC learning algorithm. n_0 represented artificial observation of cow chewing times, n_1 represented chewing times by Mean-shift algorithm, and n_2 represented chewing times by STC learning algorithm.

3   Results and discussion

3.1   Test results

The test results are shown in Table 2.   The units of n_0, n_1 and n_2 are times. T_1, T_2 were monitoring time of the Mean-shift method and STC learning method, unit (s).

The contrast of test error results was shown in Table 2, the standard deviation of the monitoring error of Mean-shift algorithm was 11.12%.   The monitoring error rate of STC learning algorithm was 9.92%.   It can be seen from Table 2, under the same monitoring environment and conditions, the average monitoring success rate of Mean-shift algorithm was 78.07%, and 85.45% for STC learning algorithm.   The results showed that the STC learning algorithm has higher detection accuracy.   The average time of each video clip detected by Mean-shift algorithm was 23.82 s, and that of the STC learning algorithm was 18.56 s, improving 22.08%, which showed the STC learning algorithm has faster monitor speed.

The contrast of test error results was shown in Figure 4.   The horizontal axis represented the number of videos, and the vertical represented the error rate.   The average monitoring success rate of

Mean-shift algorithm was 78.07%, and that of STC learning algorithm was 85.45%.   The error rate of mean-shift was extremely lower than the STC algorithm at Video 14 and Video 19, the reason of which was that in these two videos, the cow was turning or raising its head slightly.   Besides, the tracking region we selected also has influence on the detection accuracy.   The results showed that the STC learning algorithm has higher detection accuracy.   So the STC learning algorithm applied to cow ruminant behavior monitoring is more suitable compared with Mean-shift algorithm.

Table 2   Contrast of test error results

| No.	n_0	n_1	R_1/%	T_1	n_2	R_2/%	T_2
1	34	28	17.65	19.00	33	2.94	18.47
2	34	29	14.71	18.27	30	11.76	18.67
3	33	26	21.21	19.77	30	9.09	19.40
4	37	23	37.88	22.90	27	27.03	18.25
5	38	28	26.32	21.12	42	10.53	17.75
6	32	27	15.63	20.30	41	28.13	17.50
7	31	32	3.23	23.56	30	3.23	17.70
8	37	26	29.73	20.89	31	16.22	20.52
9	31	24	22.58	20.46	26	16.13	21.14
10	31	28	9.68	63.63	28	9.68	18.06
11	37	29	21.62	26.35	43	16.22	16.32
12	33	26	21.21	23.13	35	6.06	16.80
13	32	17	46.86	25.33	30	6.25	16.21
14	32	31	3.13	25.72	43	34.36	22.32
15	37	28	24.32	24.19	28	24.32	18.06
16	36	25	30.56	24.12	32	11.11	17.38
17	30	22	26.67	18.27	29	3.33	15.39
18	36	28	22.22	20.51	31	13.89	20.62
19	30	28	6.67	17.18	40	33.33	22.72
20	30	19	36.67	21.63	28	6.67	18.01
Average			21.93	23.82		14.51	18.56
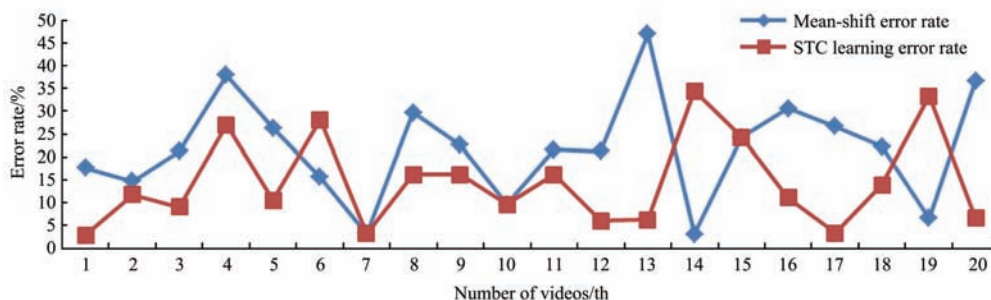
Figure 4   Contrast result of test error rate

The error rate of the two monitoring methods was fitted by R^2, and the results are shown in Figure 5. The horizontal axis represented the number of videos, the vertical axis represented the success rate, and $f(x)$ was the linear model poly.

The average time of the two monitoring methods was contrasted in Figure 6. The abscissa represented the number of videos, the ordinate represented time (s). The average time of each video clip detected by Mean-shift algorithm and STC learning algorithm was 23.82 s, 18.56 s, respectively, which was improved by 22.08% for the STC learning algorithm. The results showed that the STC learning method was faster and more efficient.
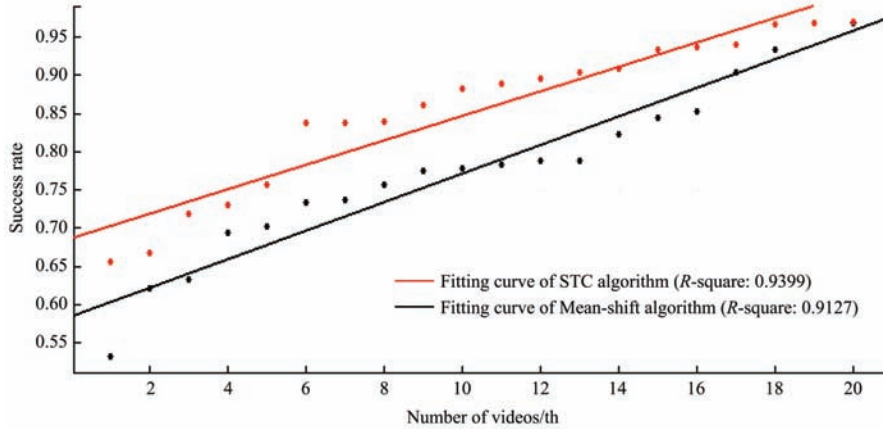
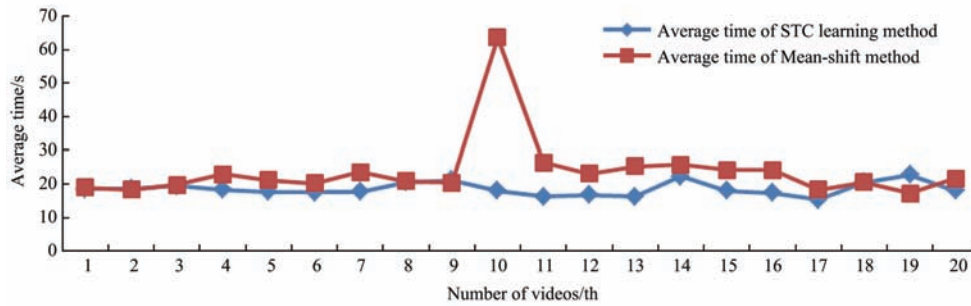Figure 5   Contrast of error rate fitted by R^2

Figure 6   Contrast of average time of Mean-shift method and STC learning method

3.2 Discussion

As the test environment was complex, the monitoring processes were under the interference of cow behaviors, such as raising head, turning head and other movements. As a result, certain errors could exist in the detection processes. The analyses are as follows:
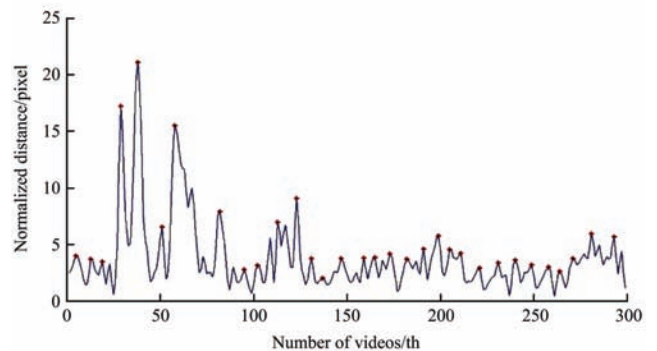
Cut out video clips in second segments which include cow twist action and compare the monitoring results of Mean-shift algorithm and STC learning algorithm. 3 typical frames of a cow twist action were shown in Figure 7a. The mouth trajectory curves of the two methods are shown in Figures 7b and 7c.

a. Cow twist action
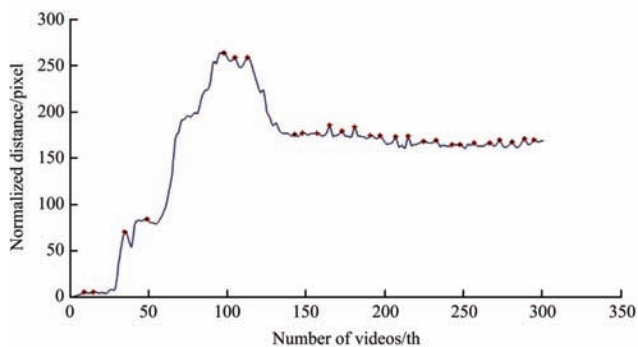
b. Trajectory curve of Mean-shift algorithm corresponding to cow's twist action

c. Trajectory curve of STC learning algorithm corresponding to cow's twist action

Figure 7   Trajectory curve corresponding to cow's twist action

In this video clip, the cows chewed 35 times, the number of chewing times tracked by using the Mean-shift algorithm was 27 times, and the running time was 28.02 s.   The number of chewing times tracked by using the STC learning algorithm was 27 times, the running time was 15.80 s.

Cut out video clips in second segments which contain cow shaking head movement and compare the monitoring results of Mean-shift algorithm with STC learning algorithm.   Three typical frames of a cow shaking head movement are shown in Figure 8a. The mouth trajectory curves of the two methods are shown in Figures 8b and 8c.

In this video clip, the cows chewed 33 times, the number of chewing times tracked by using the Mean-shift algorithm was 25 times, and the running time was 27.99 s.   The number of chewing times tracked by using the STC learning algorithm was 35 times, the running time was 16.85 s.

Cut out video clips in second segments which contain cow raising head movement and compare the monitoring results of Mean-shift algorithm with STC learning algorithm.   Three typical frames of a cow shaking head movement are shown in Figure 9a. The mouth trajectory curves of the two methods are shown in Figures 9b and 9c.

a. Cow shacking head movement
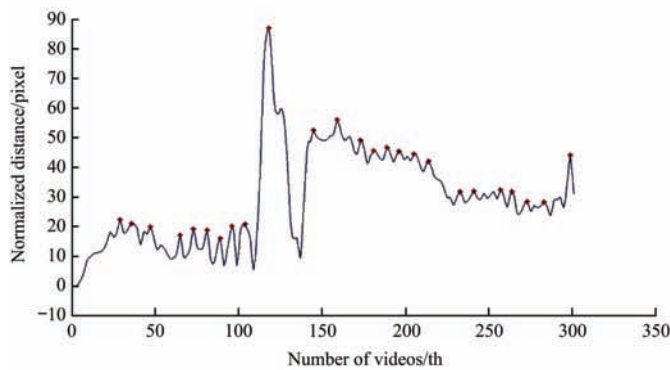
b. Trajectory curve of Mean-shift algorithm corresponding to cow's shacking head movement
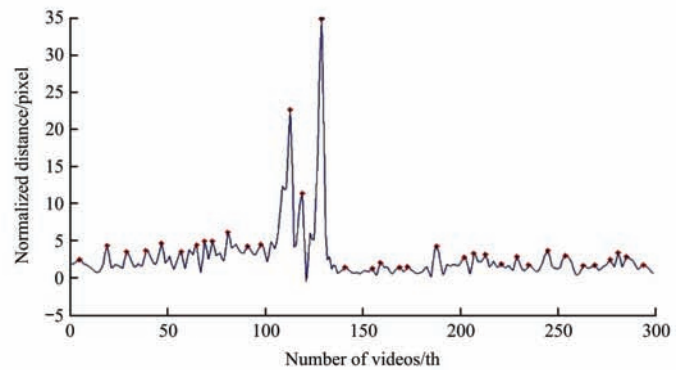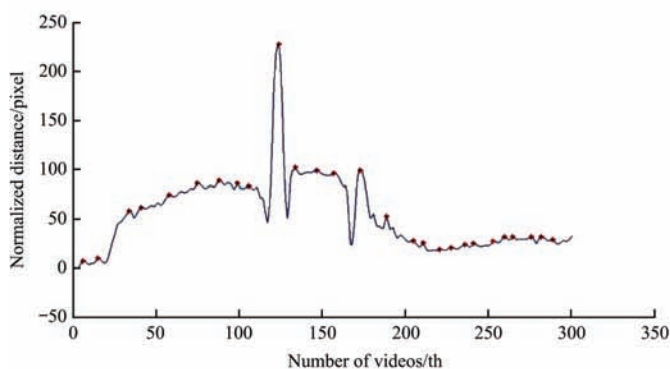
c. Trajectory curve of STC learning algorithm corresponding to cow's shacking head movement
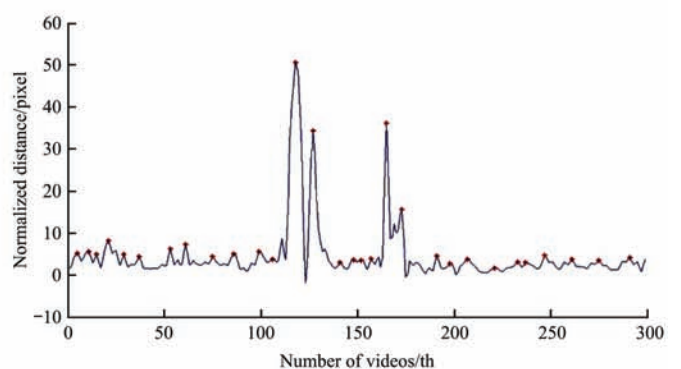
Figure 8   Trajectory curve corresponding to cow's shacking head movement

a. Cow raising head movement

b. Trajectory curve of Mean-shift algorithm corresponding to cow's raising head movement

c. Trajectory curve of STC learning algorithm corresponding to cow's raising head movement

Figure 9   Trajectory curve corresponding to cow's raising head movement

In this video clip, the cows chewed 35 times, the number of chewing times tracked by using the Mean-shift algorithm was 27 times, and the running time was 28.02 s.   The number of chewing times tracked by using the STC learning algorithm was 32 times, the running time was 18.80 s.

Combined with three kinds of cow ruminant behavior monitoring results under interference above, the method based on STC learning algorithm has better robustness, faster speed and higher accuracy compared with Mean-shift algorithm.

4   Conclusions

Ruminant activity is an important index reflecting the health of cows.   In order to realize the high efficiency of real-time monitoring of cow's rumination, the automatic monitoring method by using the STC learning algorithm was studied and compared with the Mean-shift algorithm.   The main conclusions are as follows:

1) Aiming at the dairy cow scale farming environment, this paper proposed an intelligent monitoring method of cow ruminating behavior based on the STC learning algorithm. 20 videos segments were used to conduct the experiment, and the average detection accuracy of dairy ruminant behavior was 85.45%. Compared with the Mean-shift algorithm, the STC learning algorithm had better detection rate, indicating that the method is feasible to monitor cow ruminant behavior.

2) The average time of each video clip detected by Mean-shift algorithm and STC learning algorithm was 23.82 s, 18.56 s, respectively, which was improved by 22.08% for the STC learning algorithm.   The results showed that the STC learning method was faster and more efficient.

3) To ensure the accuracy of dairy ruminant behavior in this study, manual selection of cow mouth region was used in the initial frame.   Further research will be focused on how to achieve the automatic selection of cow mouth area.   In addition, this study was carried out aiming at a single cow, and the research for cow group ruminant behavior still warrants further research.

[References]

[1]   Yan X, Dong G, Xu W, Liu A, Jose G, Wang Y, et al.   Analysis of influence factors on cow's rumination and activity in Beijing.   Acta Veterinaria Et Zootechnica Sinica, 2016; 47(5): 955–961. (in Chinese)

[2]   Shao D.   Researches on variation of the rumination and its influencing factors in lactating cows. 2015; Jilin University. (in Chinese)

[3]   Bao Y, Chen X, Zhang L.   Analysis of differential diagnosis of cow ruminant disease.   Agricultural Development and Equipments, 2016; 22(1): 164–159. (in Chinese)

[4]   Wu Y, Xie X, Han Z.   Research on target tracking algorithm based on particle filter and Mean-Shift.   International Conference on Frontiers of Mechanical Engineering and Materials Engineering, 2013; pp.1050–1053.

[5]   Comaniciu D, Ramesh V, Meer P.   Kernel-based object tracking.   IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003; 25(5): 564–577.

[6]   Chen W, Zhang X, Luo S.   Video vehicle tracking based on improved Mean-Shift algorithm.   Advanced Materials Research, 2011; 179-180: 1408–1411.

[7]   Chen D.   The research and design of target detecting and tracking system based on DSP.   Yanshan University, 2016, (in Chinese)

[8]   Huang J, Gao Y, Li Y, Chen J, Jiang J.   Algorithm research on motion area segmentation based on improved frame differences.   IEEE Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), 2015; pp.1152–1155.

[9]   Cui N, Zhang L, Wang X, Yang F, Lu B.   Application of adaptive high-degree cubature Kalman filter in target tracking.   Acta Aeronautica et Astronautica Sinica, 2015; 36(12): 3885–3895. (in Chinese)

[10]   Li H, Fan H.   Research of several background modeling based on background subtraction.   Industrial Control Computer, 2012; 25(7): 62–64. (in Chinese)

[11]   Zuo Q, Qi Y.   A novel spatial–temporal optical flow method for estimating the velocity fields of a fluid sequence.   Visual Computer, 2015; 33(3): 1–10.

[12]   Barron J L, Fleet D J, Beauchemin S S.   Performance of optical flow techniques.   International Journal of Computer Vision, 1994; 12(1): 43–77.

[13]   Sugisaka M, Sato S.   Improvement of the precision of the gradient method and object tracking using optical flow.   Artificial Life and Robotics, 2004; (4): 182–184.

[14]   Xia J, Rao W, Huang W, Lu Z.   Automatic multi-vehicle tracking using video cameras: An improved CAMShift approach.   KSCE Journal of Civil Engineering, 2013; 17(6): 1462–1470.

[15]   Wang Z, Yang X, Xu Y, Yu S.   CamShift guided particle filter for visual tracking.   Pattern Recognition Letters, 2008; 30(4): 407–413.

[16]   Wang J, Liu Y, Wu M.   Cam Shift tracking algorithm based on speed-up robust features.   Journal of Computer Applications, 2013; 33(2): 499–502. (in Chinese)

[17]   Zhang K, Zhang L, Yang M.   Real-time compressive tracking. Computer Vision-ECCV 2012.   Spring Berlin Heideberg, 2012; pp.864–877.

[18]   Zhang K, Zhang L, Yang M.   Fast compressive tracking.   IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014; 36(10): 2002–2015.

[19]   Zhang K, Zhang L, Liu Q, Zhang D, Yang M.   Fast visual tracking via dense spatio-temporal context learning.   Computer Vision-ECCV 2014. Springer International Publishing, 2014; pp.127–141.

[20]   Lu J, Chen Y, Zou Y, Zou G.   Long-term tracking based on spatio-temporal context.   Journal of Shanghai Jiaotong University (Science), 2017; 22(4): 504–512.

[21]   Wang C, Shen Y, Wang Y, Zhu Z.   Gesture tracking and recognition based on spatio-temporal context.   Computer Engineering and Applications, 2016; 52(9): 202–207. (in Chinese)

[22]   Li N, Wu X, Xu D, Guo H, Feng W.   Spatio-temporal context analysis within video volumes for anomalous-event detection and localization. Neurocomputing, 2015; 155: 309–319.

[23]   Kamal A. H. M, Montse Parada.   Translation based estimation technique to handle occlusion while using mean-shift in tracking.   Research Journal of Applied Sciences, 2009; 4(4): 129–133.

[24]   Chen K, Song K, Choi K, Guo Y.   Optimized meanshift target reference model based on improved pixel weighting in visual tracking.   Journal of Electronics (China), 2013; 30(3): 283–289.

[25]   Lu R, Huang X, Xu W, Shen L.   Meanshift tracking with kalman filter and rotation-invariant features.   Applied Mechanics and Materials, 2013; 2617(380): 1824–1828.

[26]   Jatoth R, Gopisetty S, Hussain M.   Performance analysis of Alpha Beta filter, kalman filter and meanshift for object tracking in video sequences. International Journal of Image, Graphics and Signal Processing (IJIGSP), 2015; 7(3): 24–30.

[27]   Hou P, Xu J, Zhao J, Zhan X, Fan G.   A novel model based on LBP and meanshift for UAV image segmentation.   Applied Mechanics and Materials, 2014; 701–702: 270–273.

[28]   Mohcine B, Benayad N.   Object detection and segmentation using adaptive meanshift blob tracking algorithm and graph cuts theory.   Image Processing and Communications Challenges 5.   Springer International Publishing, 2014: 143–151.

[29]   Zhao H.   Image denoising algorithm based on multi-scale Meanshift. Journal of Jilin University: Engineering and Technology Edition, 2014; 44(5): 1417–1422. (in Chinese)

[30]   Li B, Zeng Z, Chen J.   Vehicle classification and tracking based on particle swarm optimization and meanshift.   Advanced Materials Research, 2010; 121–122: 417–422.

[31]   Wu J, Song S, An W.   Mean Shift method for multi-channel image segmentation.   Packaging Engineering, 2015; 36(21): 89–94. (in Chinese)

[32]   Wang X, Wu W, Qian Y.   Trajectory clustering based customer movement tracking in a supermarket.   CAAI Transactions on Intelligent Systems, 2015; (2): 187–192. (in Chinese)