Intelligent monitoring method of cow ruminant behavior based on video analysis technology

Chen Yujuan¹,², He Dongjian¹,², Fu Yinxi³, Song Huaibo¹,²*

(1. College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling, Shaanxi 712100, China; 2. Ministry of Agriculture Key Laboratory for Agricultural Internet of Things, Yangling, Shaanxi 712100, China; 3. Department of Foreign Languages, Northwest A&F University, Yangling, Shaanxi 712100, China)

Abstract: To overcome the limitations of traditional dairy cow's rumination detection methods, a video-based analysis on the intelligent monitoring method of cow ruminant behavior was proposed in this study. The Mean Shift algorithm was used to track the jaw motion of dairy cows accurately. The centroid trajectory curve of the cow mouth motion was subsequently extracted from the video. In this way, the monitoring of the ruminant behavior of dairy cows was realized. To verify the accuracy of the method, six videos, a total of 99,000 frames were selected. The test results demonstrated that the success rate of this method was 92.03%, despite the interference of behaviors, such as raising or turning of the cow’s head. The results demonstrate that this method, which monitors the ruminant behavior of dairy cows, is effective and feasible.

Keywords: dairy cow, rumination, intelligent monitoring, video analysis, animal behavior

DOI: 10.25165/j.ijabe.20171005.3117


1 Introduction

Ruminant activity is an important index reflecting the health of animals with rumens. When cows suffer from disease, rumination time decreases significantly. The influence of a variety of diseases affect the rumination time uniquely[1]. For example, in cows with rumen tympanites, abomasum displacements or digestive tract diseases, such as enteritis, the rumination time fell the most followed by postpartum paralysis and reproductive diseases[2]. Besides, infectious diseases also have an effect on the rumination time[3]. For cow farmers, the earlier the access to potential health problems of dairy cows, the lower the cost of solving the corresponding problems. Therefore, 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 who have high labor costs and work intensity and often fail to meet the management needs of large-scale farming in addition to other problems. Therefore, several scholars focus on automated ways of recording the daily behavior of animals with rumens. Huang et al.[4] proposed a simple webcam to record animal behaviors. Davis et al.[5] utilized a GPS Herd Activity and Well-Being Kit (GPS HAWK) to monitor cattle behavior and the effect of sample interval on travel distance. Swain et al.[6] evaluated an active transponder system to monitor spatial and temporal location of cattle within patches of a grazed sward. An evaluation of three-dimensional accelerometers of monitoring and classifying behavior patterns in cattle was presented by
Robert et al.\cite{7} To distinguish between a variety of cattle foraging activities, an accelerometer-based activity monitor was used by Yoshitoshi et al.\cite{8} White et al.\cite{9} proposed an evaluation of two-dimensional accelerometers to monitor behavior of beef calves after castration. Ding et al.\cite{10} proposed a method based on the IGER recorder to record the round-the-clock feeding and the behavior of the Yak in autumn and winter. Liu et al.\cite{11} presented a real-time extraction method of dairy cattle’s moving process based on a mixed Gaussian model. A method based on video analysis of respiratory frequency and abnormalities detection of dairy cows was developed by Zhao et al.\cite{12}, where the success rate of respiratory abnormalities was 89.06%. Yao et al.\cite{13} built a dairy ruminant information acquisition system based on ANT (Advanced Net Tools). Takuji et al.\cite{14} proposed a proprietary ruminant feeding behavior detection method based on a telemetry tap device. Schirmann et al.\cite{15} designed a method of automatic monitoring of Holstein cow rumination based on HR-Tags collar. A ruminant time-recording method based on a rumination monitoring system (RMS) was proposed by Byskov et al.\cite{16} Reith et al.\cite{17} proposed a rumination time-recording method as well, while it was based on collar-mounted sensor technology. 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.

Rumination can be divided into the 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 objective of this study was to propose an automatic method to video-monitor the ruminant behavior of a cow. Firstly, according to the video acquisition platform which was built for cattle, continuous frames of the cattle were captured. Then, the Mean Shift algorithm was implemented to achieve accurate tracking of the trajectory of cow mouth regions. Lastly, the cow mouth movement trajectory was acquired to realize intelligent monitoring of cow rumination behavior. So, the research results could be the first step of building an automatic cow’s health monitoring system, and could also promote the level of intelligent equipment of breeding industry.

2 Materials and methods

2.1 Materials

Video data were collected, both in the daytime and night, from five different cows on a commercial dairy farm in Yangling, Shaanxi Province on July 2013 by the camera (DS-2DM1-714 integrated IP camera, Hikvision Inc., Hangzhou, China) at a resolution of 704 pixel×574 pixel, the color video format was set to Phase Alternation Line (PAL) with a frame rate of 25 frames per second and with the code rate of 2048 kb/s. The test subjects were American Holstein cows in mid lactation. Considering that cows ruminate in a prone state in most cases, videos taken when cows were in a prone state were collected as it would be more typical and representative\cite{12}. The lengths of the six captured videos were all 16'49". To verify the effectiveness of the algorithm, the original video data were divided into 40 video segments for the convenience of analysis. The length of each video segment was approximately 1'. In these video segments, the captured video frame rate set in this study was 10 frames per second. It is thus not difficult to realize a real-time detection of ruminant behaviors. The videos contain cow behaviors, such as raising head, turning head, rumination and static state. The statistics of the dairy cow’s video information obtained are shown in Table 1.

<table>
<thead>
<tr>
<th>No.</th>
<th>Time</th>
<th>Length of videos</th>
<th>Total frames</th>
<th>Behaviors</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Day</td>
<td>16'49&quot;</td>
<td>10 090</td>
<td>Raising head, turning head, rumination and static state</td>
</tr>
<tr>
<td>2</td>
<td>Day</td>
<td>16'49&quot;</td>
<td>10 090</td>
<td>Raising head, turning head, rumination and static state</td>
</tr>
<tr>
<td>3</td>
<td>Day</td>
<td>16'49&quot;</td>
<td>10 090</td>
<td>Raising head, turning head, rumination and static state</td>
</tr>
<tr>
<td>4</td>
<td>Day</td>
<td>16'49&quot;</td>
<td>10 090</td>
<td>Raising head, turning head, rumination and static state</td>
</tr>
<tr>
<td>5</td>
<td>Night</td>
<td>16'49&quot;</td>
<td>10 090</td>
<td>Raising head, turning head, rumination and static state</td>
</tr>
<tr>
<td>6</td>
<td>Night</td>
<td>16'49&quot;</td>
<td>10 090</td>
<td>Raising head, turning head, rumination and static state</td>
</tr>
</tbody>
</table>

2.2 Methods

2.2.1 Mean Shift algorithm

Mean Shift algorithm is a kind of non-parametric
estimation method. It is realized by an iterative process. Firstly, the mean offset value of the current point was calculated and the point was moved to the mean offset value. Next, the point was set as the new starting point and that process was repeated until meeting certain criteria. Mean Shift is widely used in such applications as clustering, tracking\textsuperscript{[18-21]}, segmentation\textsuperscript{[22,23]}, and image smoothing. For example, Mean Shift was applied to image de-noising algorithm by Zhao\textsuperscript{[24]} and vehicle tracking by Li et al.\textsuperscript{[25]} Mean Shift algorithms were also used in image segmentation\textsuperscript{[26]}, as well as supermarket customer movement tracking\textsuperscript{[27]}.

(1) Establishment of the target model and candidate target model

Assume that \( y_0 \) denotes the center position of the moving target area, \( y \) is the central location of the candidate target region. \{\( x_i \); \( i = 1, 2, 3, \ldots, m \}\) represents a collection of all the pixels in the target area, \( m \) represents the total number of eigenvalues. The probability density estimation of the eigenvalues of the target model is shown in Equation (1)\textsuperscript{[28]}:

\[
\hat{q}_n = C \sum_{i=1}^{n} k \left( \frac{y_n - x_i}{h} \right) \delta \left( b(x_i) - u \right)
\]  

(1)

The probability density estimation formula of the target candidate model \( \hat{p}_u \) is shown in Equation (2):

\[
\hat{p}_u = C_h \sum_{i=1}^{n} k \left( \frac{y - y_n}{h} \right) \delta \left( b(x_i) - u \right)
\]  

(2)

In Equations (1) and (2), \( u \in \{1, 2, \ldots, n_n\} \) represents the number of color, \( k(x) \) is the contour function of kernel function; \( y \) is the width of the kernel function; \{\( y_i \); \( i = 1, 2, 3, \ldots, m \}\) represents a collection of all the pixels in the target area; \( b(x_i) \) represents the color value of the pixel at \( x_i \); \( \delta(x) \) is Kronecker delta function; \( \delta(b(x_i) - u) \) is used to judge whether the value of the pixel \( x_i \) in the target area belongs to the interval eigenvalues of \( u \) or not; If the judgment is true, its value is 1, otherwise it is 0. \( C \) and \( C_h \) are the normalization coefficient.

The equation of \( C \) and \( C_h \) are shown in Equations (3) and (4):

\[
C = \frac{1}{n} \sum_{i=1}^{n} k \left( \frac{y_n - x_i}{h} \right)
\]  

(3)

\[
C_h = \frac{1}{n} \sum_{i=1}^{n} k \left( \frac{y - y_n}{h} \right)
\]  

(4)

To make sure that \( \sum_{u=1}^{m} \hat{q}_u = 1 \).

(2) Determination of similarity measure

To compare the similarity between the two models, the Bhattacharyya (BH) coefficient is usually used as a measuring standard of similarity, which is abbreviated as BH coefficient.

Assume that \( \hat{q}_n \) and \( \hat{p}_u \) represent the target template and the candidate target template, respectively. The computation method of similarity \( \hat{\rho}(y) \) between these two models is shown in Equation (5):

\[
\hat{\rho}(y) = \rho(\hat{p}, \hat{q}) = \sum_{u=1}^{m} \sqrt{\hat{p}_u \hat{q}_u}
\]  

(5)

where, \( d(y) \) represents the distance between the two model histograms, and this is expressed in Equation (6):

\[
d(y) = \sqrt{1 - \hat{\rho}(y)}
\]  

(6)

(3) Positioning of moving objects

Comaniciu et al.\textsuperscript{[28]} introduced the theory of converting calculation of the maximum value of similarity function in order to calculate the maximum value of probability density function, and realized the goal of positioning the target by Mean Shift vector in detail. The steps are as follows:

Step 1: Calculate the candidate target model \( \hat{p}(y_0) \).

\( \hat{\rho}(y) \) is expanded at \( p(\hat{y}_0) \) by a Taylor series expansion. Assuming the BH coefficient is approximated to Equation (7):

\[
\rho(p(\hat{y}_0), \hat{q}) = \frac{1}{2} \sum_{u=1}^{n} p_u(y_0) \hat{q}_u + \frac{1}{2} C_h \sum_{i=1}^{n} w_i k \left( \frac{y - y_n}{h} \right)
\]  

(7)

In Equation (7), \( w_i \) represents pixel weight, and the accessor method is shown in Equation (8):

\[
w_i = \sum_{u=1}^{n} \frac{\hat{q}_u}{p_u(\hat{y}_0)} \delta \left( b(x_i) - u \right)
\]  

(8)

Step 2: Calculate the center of the candidate target \( y_1 \), which is listed in Equation (9):

\[
y_1 = \frac{\sum_{i=1}^{n} x_i w_i \left( \frac{y_0 - x_i}{h} \right) / \sum_{i=1}^{n} w_i \left( \frac{y_0 - y_n}{h} \right)}{\sum_{i=1}^{n} \left( \frac{y_0 - x_i}{h} \right) / \sum_{i=1}^{n} \left( \frac{y_0 - y_n}{h} \right)}
\]  

(9)
Step 3: Stop calculating if \( \| y_i - y_0 \| < \varepsilon \) or reaching the maximum number of iterations (\( \varepsilon \) is the permitted error value of the iteration stopping criterion). Otherwise, use \( y_i \) to replace \( y_0 \) and go back to Equation (8), and continue to find the candidate target location which meets the criteria.

In short, the vector of Mean Shift algorithm is always pointing to the direction of the largest density increment in a data space, so it drifts from the beginning to the direction where the largest similarity between the two models exists.

2.2.2 Cow mouth trajectory tracking based on Mean Shift

When the Mean Shift tracking algorithm is used in dairy cattle rumination behavior, the tracked motion target area should be initialized at first. The motion target area can be selected by automatic and manual methods. Automatic selection methods automatically choose a target and add the initial tracking box in image sequences. However, the algorithm is complex, and it is difficult to ensure its accuracy. Therefore, manual marking method was used in this work. The mouse was utilized to select the tracked moving target, and that is the repeatedly chewing mouth area of the cattle.

Next, the histogram of the candidate target model in the latter frame was calculated; this is so the degree of similarity between two models based on BH coefficient could be obtained. Then, the maximum of BH coefficient was selected and the tracking window drift towards the densest direction were made after iteration. Finally, the true location of a moving target could be determined. The concrete implementation steps are as follows:

Step 1: Read the video image and then use the mouse to select cow mouth tracking frame. Take \( y_0 \) as a starting point for the current frame. As shown in Figure 1, the cow mouth was in the red box and the red dot in the red box was the center of the moving target area, which is \( y_0 \).

Step 2: Calculate \( \hat{q}_m \) of the cow mouth intercept according to Equation (1) and the candidate target model \( \hat{p}_m \) according to Equation (2). Search for the tracked object that may contain the cow mouth candidate region from the second frame and assume that the center of the candidate target area was \( y \).

Figure 1  Tracking box of cow

Step 3: Calculate the similarity \( \rho(y) \) between the two models according to Equation (1). The results were shown in Figure 2. Figure 2a was the first frame; Figure 2b was the second frame; Figure 2c was comparison of the histogram similarity between the first and second frame. In Figure 2c, the horizontal coordinate represented the color range, and the vertical coordinate represented the probability of similarity. Red was the histogram of the target area in the first frame, and blue was the histogram of the second target area.

Figure 2  Similarity comparisons of two consecutive frames

Step 4: Compare the similarity between the former and the latter frames.

1) If \( \rho(y_0) > \rho(y_i) \), then \( y_1 \leftarrow (y_i + y_0)/2 \), which
means if the similarity of next frame is smaller than the previous one, the new target location will be the average value of them. Next, recalculate the similarity of $\rho'(y_i)$.

2) If $||y_1 - y_0|| \leq \varepsilon$, find the target location and end the iterative search. Otherwise, $y_0 \leftarrow y_1$, jump to Step2.

The candidate regions obtained after the above steps were shown in the red square area of Figure 2b. The red dot was the center position of the candidate region. From Figure 2b, the cow mouth region was in the tracking box, which indicated that the tracking result was good.

2.2.3 Results of cow mouth trajectory tracking

Part of the video frames was captured from video No.1 and the Mean Shift algorithm was used to track targets. The results were shown in Figure 3. Figure 3a showed the first six frames got from Video No. 1. Figure 3b showed the tracking areas’ histogram similarity comparison results of the first six frames. Red was the target areas’ histograms of the previous frame and blue was the target areas’ histograms of the latter frame. The first image of Figure 3b was the first frame, so it only had its own histogram data. Figure 3c was the cow mouth tracking results; Figure 3d was the cow mouth target area center point trajectory of the first six frames. The horizontal axis represented the number of frames, and the vertical represented the distance of center points between the latter and the former frame. The origin was the center of target area in the first frame. When cows would chew the cud, their jaw would move up and down, and then the center of the tracking box would also change up and down. Depending on the situation, the tracking curve can be drawn. It can be seen from Figure 3b that the former and the latter two frames’ histograms of the target area were basically coincident, which showed that the method had a high matching degree of similarity. In Figure 3c, cattle mouths were all in the tracking box, which showed that this method had good tracking results and could be carried out to monitor dairy ruminant behavior.

![First six frames of Video No.1](image1)

![Comparison of the former and latter frames’ histogram](image2)

![Detection of mouth region of cow’s (Enlarged)](image3)

![Trajectory curve of the given six frames](image4)

Figure 3  Trajectory tracking results of cow mouth
3 Results and discussion

3.1 Results

The formula of error rate is shown in Equation (10).

\[ R = \frac{|n_1 - n_0|}{n_0} \]

where, \( R \) represented the monitoring error rate. The monitoring success rate was \( V = 1 - R \). To ensure the accuracy of the data extracted, the six video data were divided into 40 small segments to test the data. The test results are shown in Table 2. \( n_0 \) represented artificial observation of cow chewing times, and \( n_1 \) represented chewing times by the algorithm proposed. The units of \( n_0 \) and \( n_1 \) are times. As it can be seen in Table 2, the difference between detected chewing times and the true chewing times was small. The average monitoring error rate was 7.97%, and the success rate was 92.03%.

<table>
<thead>
<tr>
<th>No.</th>
<th>( n_0 )</th>
<th>( n_1 )</th>
<th>( R )%</th>
<th>No.</th>
<th>( n_0 )</th>
<th>( n_1 )</th>
<th>( R )%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>67</td>
<td>60</td>
<td>10.45</td>
<td>21</td>
<td>65</td>
<td>62</td>
<td>3.12</td>
</tr>
<tr>
<td>2</td>
<td>68</td>
<td>60</td>
<td>11.76</td>
<td>22</td>
<td>64</td>
<td>68</td>
<td>6.25</td>
</tr>
<tr>
<td>3</td>
<td>68</td>
<td>64</td>
<td>5.88</td>
<td>23</td>
<td>68</td>
<td>62</td>
<td>8.00</td>
</tr>
<tr>
<td>4</td>
<td>70</td>
<td>67</td>
<td>4.29</td>
<td>24</td>
<td>66</td>
<td>55</td>
<td>7.47</td>
</tr>
<tr>
<td>5</td>
<td>64</td>
<td>61</td>
<td>5.00</td>
<td>25</td>
<td>61</td>
<td>65</td>
<td>14.71</td>
</tr>
<tr>
<td>6</td>
<td>70</td>
<td>58</td>
<td>17.14</td>
<td>26</td>
<td>61</td>
<td>62</td>
<td>4.55</td>
</tr>
<tr>
<td>7</td>
<td>67</td>
<td>60</td>
<td>10.45</td>
<td>27</td>
<td>60</td>
<td>65</td>
<td>5.89</td>
</tr>
<tr>
<td>8</td>
<td>67</td>
<td>60</td>
<td>10.45</td>
<td>28</td>
<td>59</td>
<td>45</td>
<td>10.48</td>
</tr>
<tr>
<td>9</td>
<td>66</td>
<td>59</td>
<td>10.61</td>
<td>29</td>
<td>60</td>
<td>64</td>
<td>4.41</td>
</tr>
<tr>
<td>10</td>
<td>55</td>
<td>62</td>
<td>12.73</td>
<td>30</td>
<td>52</td>
<td>47</td>
<td>17.74</td>
</tr>
<tr>
<td>11</td>
<td>63</td>
<td>61</td>
<td>3.17</td>
<td>31</td>
<td>51</td>
<td>49</td>
<td>3.92</td>
</tr>
<tr>
<td>12</td>
<td>64</td>
<td>60</td>
<td>6.25</td>
<td>32</td>
<td>51</td>
<td>49</td>
<td>3.92</td>
</tr>
<tr>
<td>13</td>
<td>50</td>
<td>54</td>
<td>8.00</td>
<td>33</td>
<td>50</td>
<td>45</td>
<td>1.00</td>
</tr>
<tr>
<td>14</td>
<td>67</td>
<td>62</td>
<td>7.47</td>
<td>34</td>
<td>49</td>
<td>54</td>
<td>10.20</td>
</tr>
<tr>
<td>15</td>
<td>34</td>
<td>29</td>
<td>14.71</td>
<td>35</td>
<td>49</td>
<td>53</td>
<td>8.16</td>
</tr>
<tr>
<td>16</td>
<td>66</td>
<td>69</td>
<td>4.55</td>
<td>36</td>
<td>50</td>
<td>51</td>
<td>2.00</td>
</tr>
<tr>
<td>17</td>
<td>68</td>
<td>64</td>
<td>5.89</td>
<td>37</td>
<td>51</td>
<td>45</td>
<td>11.76</td>
</tr>
<tr>
<td>18</td>
<td>67</td>
<td>60</td>
<td>10.48</td>
<td>38</td>
<td>50</td>
<td>47</td>
<td>6.00</td>
</tr>
<tr>
<td>19</td>
<td>68</td>
<td>65</td>
<td>4.41</td>
<td>39</td>
<td>51</td>
<td>53</td>
<td>3.92</td>
</tr>
<tr>
<td>20</td>
<td>62</td>
<td>51</td>
<td>17.74</td>
<td>40</td>
<td>52</td>
<td>54</td>
<td>3.85</td>
</tr>
</tbody>
</table>

The average monitoring error rate was 7.97%.

### Table 2 Test error results

Figures 5a-5d were the tracking results of dairy cattle ruminant behavior of four video segments. Small fluctuations in the figures were the cow chewing actions, and the continuous chewing process was the chewing cud process. There would be an interval between two chewing cud processes and that is the regurgitation process.

3.2 Discussion

At present, the behavior of cow rumination is mainly observed by farmers, who have high labor cost, high work intensity, and fail to meet the management needs of large-scale farming and other problems. The method proposed in this paper solved this problem efficiently by tracking the center of a cow’s mouth region. The frame-rate set could realize real-time detection of ruminant behaviors.

As the test environment was complex, and the monitoring processes were under the interference of cow behaviors, such as their raising head, turning head and other movements; as a result, a certain error could exist in the detection processes. The analyses were as follows:

In the sixth part of the divided segments, the monitored dairy cow had head shaking movements which resulted in a big error for ruminant monitoring. Three typical frames of a cow shaking its head were shown in Figure 6a. It be seen that the cow shakes its head with a large movement within a small-time interval. The trajectory curve of the mouth was shown in Figure 6b. The head shaking or raising head movement would resulted in the curve of a rapid change, that is because the tracked motion target area was cow mouth, and the curve drawn, depend on the points of the center of tracked motion target area. When a cow shakes its head rapidly, the tracked motion target area would change rapidly, so the points of the curve change rapidly. The existence of
the action led to a big error of the monitoring, but it also indicated that this method could be used to monitor the abnormal movements of the cow head region.

The monitoring error rate of the twentieth segmented video was 17.74%. This was due to a more obvious lateral head action in the video. When the cow was in a lateral head state, the curve was like that of the chewing behavior, which led to big errors in the monitoring. Three typical frames of cow in lateral head state were shown in Figure 7a, and the corresponding output curve was shown in Figure 7b. Since the changes of the curve were not obvious, it was easy to cause the error.
In this paper, the ways to decompose videos and choose track regions were through human operation. In further work, an automatic method will be considered to decompose videos and choose track regions.

4 Conclusions

Intelligent monitoring of the behaviors of dairy cows is an inevitable demand of the development of modern animal husbandry. To realize the intelligent monitoring of cows’ behavior under the raising environment, the Mean Shift algorithm was used to test the 40 segments video data in this study. The primary results are as follows:

1) Aiming at the dairy cow scale farming environment, an intelligent monitoring method of cow ruminating behavior based on the Mean Shift target tracking algorithm was proposed. Forty videos segments were used to conduct the experiment, and the average detection accuracy of dairy ruminant behavior was 92.03%, indicating that the method is feasible to monitor cow ruminant behavior.

2) As cow is a living body, the actions of which, such as raising its head and turning its head, will affect the tracking results of the ruminations to a certain extent.

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.

Acknowledgments

This work is supported by the National Key Technology R&D Program of China (No. 2017YFD0701603), and the Natural Science Foundation of China (No. 60975007). The authors would like to thank all of the authors cited in this article and anonymous referees for their helpful comments and suggestions.

[References]


