TRANSACTIONS OF THE CHINESE SOCIETY FOR AGRICULTURAL MACHINERY
Vol. 47, No. 5, 2016

doi:10.6041/j.issn.1000-1298.2016.05.039

Real-time Target Detection for Moving Cows Based on Gaussian Mixture Model

Liu Dong Zhao Kaixuan He Dongjian
(*College of Mechanical and Electronic Engineering*, *Northwest A&F University*, *Yangling*, *Shaanxi* 712100, *China*)

Abstract: Target detection is the basic work for analyzing the behavior of the cows using video analysis technology. It is difficult to extract the moving cows accurately and real-timely with the existing target detection methods because of the complex background environment. In this study, a series of improvement measures were proposed based on Gaussian mixture model to meet the system requirements. A dynamic background modeling method with penalty factor was proposed for the mathematical model of the background which can overcome the high model complexity. A two-class classification algorithm based on chromaticity distortion and brightness distortion was proposed to avoid the influence of the shaded area in the foreground detection algorithm. Local update method was proposed to avoid missing the target if it stays for a long time. In order to verify the validity of the algorithm, four evaluation parameters were introduced to test the detection algorithm including model complexity, false detection rate of foreground, false detection rate of background and processing time. Experimental results show that model complexity was 50.85% lower than the classical method. False detection rate of foreground and false detection rate of background were 18.18% and 7.52%, which had 19.50 and 13.37 percent lower than the classical Gaussian mixture model. Processing time of average single frame was 29.25% lower. Statistics indicate that the algorithm proposed in this study can improve the detection performance and it is an extension to classical Gaussian mixture model.

Key words: moving cows; video analysis; target detection; shadow detection; Gaussian mixture model

0 Introduction

Intelligent video-surveillance is an important topic in computer vision, which has been widely used in precision livestock farming, such as lameness detection[1-4], breath detection[5-6], body condition score[7-9], disease surveillance[10] and behavioral analysis[11-13] etc. Perceiving animal information and behaviors with intelligent video-surveillance system are fundamental to precision livestock farming. Accurate and real-time animal targets detecting in complex environment is the basic premise for information perceiving.

In recent years, international and domestic academics had got a number of research results, but problems still exist. The complex background environment makes the experiments done under the harsh terms and usually leads to unsatisfactory results. Infrared thermography technology had been used to avoid environmental impact, but which imposed costs[14]. LI, et al[15], established the experiment system with artificial light source in indoor to acquire stable background. In such conditions, the basic theory and application methods can accurately distinguish between target and background. ZHAO, et al[16], constructed background model by local background stitching method under complex environment, which got a good result but difficult to fulfill continuous inspection. Large animals in livestock production, such as cows, pigs etc., usually been raised out of doors, which has the characteristics of environmental gradient and mutation, such as the mutation of lighting equipment state, the gradient of the sun's brightness and angle, the effect of the wind and birds invasion. All of these cause a lot of obstacles to distinguish objects and background, which may make a bad effect to animal behaviors perceiving. Therefore, it is important and necessary to develop a

Received date: 2015-10-11 Accepted date: 2016-01-21
Supported by National Natural Science Foundation of China (Grant No. 61473235)
Corresponding author: He Dongjian, Professor. E-mail: hdj168@nwsuaf.edu.cn. Tel: +86-15332491226.

new method which is robust to the farming environment for accurate and real-time detecting target.

Gaussian mixture model (GMM)[17-20] can establish dynamic background model to fulfill continuous inspection, but problems still exist in the condition of complex background, such as high complexity of model and computation, not adaptive to establish optimal number of components and the foreground melting problem. The improved methods had been proposed to solve such problems. For instance, ZHANG et al[21], regarded the points with high relativity in a neighborhood as a "block", then modeled the mean of "block" instead of "pixels" to reduce the algorithm complexity. But that method caused blurred edges and jagged edges at the expense of losing useful information for time efficiency. The contradiction of GMM between veracity and timeliness is an open problem, and the key is how to distinguish and eliminate redundant information. LIN, et al[22], proposed a screening approach that removed the models whose weighting are less than the initial value in the processing of iteration. But the selection of initial value was still decided by experiment, which is easy to cause the foreground melting problem with larger initial value and the dynamic is not obvious with smaller initial value. In this paper, we studied moving cows under the complex background environment and proposed a method to identify the redundant models to speed up the GMM iteration without losing any useful information. The proposed methods aim at solving the difficulties in target detection under open environment, the effects of shadows and the foreground melting problem.

1 Background modeling method

1.1 Problems of GMM in complex background modeling

Classical GMM construct color distribution model based on the probability distributions of pixels in time domain, which have been widely used in intelligent security and other fields. GMM establishes M components (general 3 - 5) of components for each pixel. The recent samples collection of one pixel at t time is $\boldsymbol{x}_t = (x^{(t)}, x^{(t-1)}, \cdots, x^{(t-T)})$. The Gaussian distribution model for $\boldsymbol{x}_t$ is given by:

$$P(\hat{\pi}_m, \hat{\mu}_m, \hat{\sigma}_m^2 | \boldsymbol{x}_t, B+F) = \sum_{m=1}^{M} \hat{\pi}_m N(x^{(t)}; \hat{\mu}_m, \hat{\sigma}_m^2)$$

$$\left(0 < \hat{\pi}_m < 1, \ \sum_{m=1}^{M} \hat{\pi}_m = 1\right) \quad (1)$$

Where, $\boldsymbol{x}_t$ is the recent collected sample at t time; B and F represent the background pixels and the foreground pixels; M is the number of Gaussian model; $N(x^{(t)}; \hat{\mu}_m, \hat{\sigma}_m^2)$ is the m-th Gaussian function; $\hat{\pi}_m$ is the weight of the m-th Gaussian function; $\hat{\mu}_m$ is the mean of the m-th Gaussian function; $\hat{\sigma}_m^2$ is the variance of the m-th Gaussian function.

Because the new sample could be either pixel in background or pixel in foreground, both B and F are included in the Eq. (1). When sampling result $\boldsymbol{x}_t$ is given, the expectation maximization (EM) algorithm[23] is commonly used to search for the recursion equation, which is given by:

$$\hat{\pi}_m \leftarrow \hat{\pi}_m + \alpha(o_m^{(t)} - \hat{\pi}_m) \quad (2)$$

$$\hat{\mu}_m \leftarrow \hat{\mu}_m + o_m^{(t)}(\alpha / \hat{\pi}_m)(x^{(t)} - \hat{\mu}_m) \quad (3)$$

$$\hat{\sigma}_m^2 \leftarrow \hat{\sigma}_m^2 + o_m^{(t)}(\alpha / \hat{\pi}_m)[(x^{(t)} - \hat{\mu}_m)^2 - \hat{\sigma}_m^2] \quad (4)$$

Where, α is the learning rate to determine the renewal speed of parameters; $o_m^{(t)}$ is the ownership parameters to determine whether the sample belongs to the m-th Gaussian model according to the "3σ-principle" or not.

If new samples belong to the foreground, a new component will be established with a small weight $\hat{\pi}_m$. Then the weights are updated with Eq. (2), normalized and sorted descending to insure the weights sum up to 100%. Finally, the background model is established by the top 90% of the sum of weights, given by:

$$P(x^{(t)} | \boldsymbol{x}_t, B) = \sum_{m=1}^{B} \hat{\pi}_m N(x^{(t)}; \hat{\mu}_m, \hat{\sigma}_m^2) \quad (5)$$

in which, $B = \arg\min_b \left(\sum_{m=1}^{b} \hat{\pi}_m > 90\% \right)$.

Obviously, classical GMM figures out the weight, the mean and the variance for every pixel with square, root and exponent arithmetic, which has high algorithm complexity especially for RGB images with three channels. In addition, it is lack of theory foundation of selecting the number of components and intercepting "90%" as the background model from the total components.

Dairy farms are usually in the open environment, which are easy to be affected by environmental factors.

Besides, in the camera view, the actual required the number of components in different areas vary. If there are too small components to describe the background, it is easy to take the target as background falsely which is called "under-fitting". On the contrary, if there are too many components to describe the background, it is difficult to avoid the influence from shaky leaves and shaking camera, which is called "over-fitting".

In this paper, a stable background is obtained from a fixed camera. To overcome flaws of classical GMM under complex background, a self-restriction method was proposed based on minimum message length criterion, which can select the number of components adaptively and remove the redundant components.

1.2 The self-restriction method based on "minimum number of samples"

Assuming that the number of components is fixed, the weights $\hat{\pi}_m$ can be read as the probability of samples from the m-th component. The weights obey multinomial distribution in statistic sense, and the likelihood function is given by:

$$\mathcal{L}_1(\boldsymbol{\pi}, M | \boldsymbol{x}) = \lg\left(\prod_{m=1}^{M} \hat{\pi}_m^{n_m}\right)$$
$$\left(0 < \hat{\pi}_m < 1, \sum_{m=1}^{M} \hat{\pi}_m = 1\right) \tag{6}$$

Where, n_m is the number of samples belong to the m-th component.

Then the maximum likelihood estimate follows by introducing the Lagrange multiplier λ to Eq. (6):

$$\frac{\partial}{\partial \hat{\pi}_m}\left(\mathcal{L}_1 + \lambda\left(\sum_{m=1}^{M} \hat{\pi}_m - 1\right)\right) = 0 \tag{7}$$

After getting rid of λ we get:

$$\hat{\pi}_m = \frac{n_m}{\sum_{m=1}^{M} n_m} \tag{8}$$

It can be rewritten in recursion form that is the Eq. (2). Obviously, in the condition of fixed number of components, each component will be assigned several samples saturatedly by the maximum likelihood estimation, which usually results in error model and can't be adaptive to establish optimal number of model. A solution is proposed in this paper based on the truth as follows: the samples collection $\boldsymbol{x}_t = [x^t, x^{t-1}, \cdots, x^{t-T}]$ consist of the background, the target and the noise, moreover, the number of noise N are far less than the others.

The constraint factor N for complexity is introduced into the likelihood function, given by:

$$\mathcal{L}_2(\boldsymbol{\pi}, M | \boldsymbol{x}) = \lg\left(\prod_{m=1}^{M} \hat{\pi}_m^{n_m - N}\right) \tag{9}$$

Where, N represents the noise from each component with a small value (e.g. 30). Eq. (9) can be rewritten as:

$$\mathcal{L}_2(\boldsymbol{\pi}, M | \boldsymbol{x}) = \lg\left(\prod_{m=1}^{M} \hat{\pi}_m^{n_m}\right) + \lg\left(\prod_{m=1}^{M} \hat{\pi}_m^{-N}\right) \tag{10}$$

Where, the second item on the right side is called the model complexity constraint, then the maximum likelihood estimate follows by introducing the Lagrange multiplier λ to Eq. (10):

$$\frac{\partial}{\partial \hat{\pi}_m}\left(\mathcal{L}_2 + \lambda\left(\sum_{m=1}^{M} \hat{\pi}_m - 1\right)\right) = 0 \tag{11}$$

After getting rid of λ we get:

$$\hat{\pi}_m^{(t+1)} \leftarrow \hat{\pi}^{(t)} + \alpha(o_m^{(t)} - \hat{\pi}_m^{(t)}) - \alpha N/T \tag{12}$$

Where, T is the sample size.

To reduce the model complexity, the model complexity constraint gives each component N "negative-samples", that is, the weights $\hat{\pi}_m$ are initialized to be negative. When there are enough samples to "support" the m-th component ($n_m - N > 0$), the weight of the m-th component is going to be positive. Otherwise, while the "negative-weight" is detected, the corresponding component is called "error component" and been abandoned during the process of iteration. The improved background updating equation for GMM is given by:

$$\begin{cases} \hat{\pi}_m \leftarrow \hat{\pi}_m + \alpha(o_m^{(t)} - \hat{\pi}_m) - \alpha N/T \\ \hat{\mu}_m \leftarrow \hat{\mu}_m + o_m^{(t)}(\alpha/\hat{\pi}_m)(x^{(t)} - \hat{\mu}_m) \\ \hat{\sigma}_m^2 \leftarrow \hat{\sigma}_m^2 + o_m^{(t)}(\alpha/\hat{\pi}_m)[(x^{(t)} - \hat{\mu}_m)^2 - \hat{\sigma}_m^2] \end{cases} \tag{13}$$

2 Target detection algorithm

The classical GMM is described as follows: the new sample makes a match between the components one by one with the "2.5σ-principle". If the new sample is classified as background whenever matches with any component, the background model is updated by Eq. (13). Otherwise, the sample is classified as foreground, then a new component with small weight and large variance is going to be established to replace the original one whose weight is the smallest. The mean of the new component is the pixel value[24-25].

While the dairy farms are in the open environment, the classical algorithm can't distinguish the shadow from the target. Because the classical algorithm is a two-class classification algorithm constitutionally, and the shadow area usually beyond the range of "2.5σ" which is easy to be misclassified as foreground. To solve this problem, two features are selected to evaluate the brightness bias and chromaticity bias from the shadow. And a two-class classification algorithm is designed to distinguish between the target and the shadow. The two features are given by:

$$\beta_{RGB} = \arg\min_{\beta}(x_{RGB} - \beta\mu_{RGB})^2 = \frac{x_R\mu_R + x_G\mu_G + x_B\mu_B}{\mu_R^2 + \mu_G^2 + \mu_B^2} \tag{14}$$

$$D_{RGB} = \sqrt{(x_R - \beta\mu_R)^2\ (x_G - \beta\mu_G)^2\ (x_B - \beta\mu_B)^2} \tag{15}$$

Where, β_{RGB} is the mean shift between the sample and the component; D_{RGB} is the orthogonal distance between the sample and the component; x_R, x_G and x_B are the pixel values of 3-channel images; μ_R, μ_G and μ_B are the mean of components which are matched by the current sample.

The more β_{RGB} close to 1, the more the current sample close to background, and the rules are given by:

$$x_i = \begin{cases} S & (0.5 < \beta_{RGB} < 1 \text{ and } D_{RGB} < 4\sigma\beta_{RGB}) \\ F & (\text{Else}) \end{cases} \tag{16}$$

Where, S is the abbreviation of shadow; F is the abbreviation of foreground.

β_{RGB} locating in the range of 0.5 and 1 means that the current pixel value is smaller than the mean of background model. The threshold value of D_{RGB} values 4σ is the relaxation of "3σ-principle", which conforms to the intuitive understanding that shadow is close to the mean of background model but darker.

In the Eq. (13), α is the learning rate that determines the renewal speed of background. The large default value leads to quick background updating that is easy to misclassify the target from background. On the contrary, the small default value can't reach the purpose of dynamic modeling. This contradiction is inevitable relying on adjusting parameters. For instance, if the cows remain still for a long time, the algorithm continues training until the target disappear. A solution is proposed in this paper based on the relative invariance of cows' body in moving, which regards the position of moving cows' body as the judgment basis and takes local updating strategy. A minimum circumscribed rectangle and a maximum inscribed rectangular are calculated to mark cows in every frame. When the center distance of inscribed rectangular between adjacent frames is less than a threshold value (3 - 5 pixels), the target stops moving. Then, the areas outside of circumscribed rectangle are updated, otherwise all areas are updated. Besides, if the target remains still for more than 50 frames, all areas are updated to avoid failure detection when the target remains still in the border of the camera view.

3 Experiments and analysis

Video data were collected on a commercial dairy farm in Yangling, Shaanxi Province in August 2013. Side-view images were acquired while cows were passing through a 2 m wide, 7 m long aisle with a concrete floor to the water trough. To avoid any intrusion into the farm's routine and interference with cow traffic, the camera (DS-2DM1-714 integrated IP camera, Hikvision Inc., Hangzhou, China) was positioned on the supported beam of a feeding shed at a height of 1.8 m and 35 m away from the alley. The CCD sensor was paralleled to the corridor, and the focus length was set to 45 mm to assure that the width of the field of view was twice of the length of one single cow. Color video format was set to PAL with frame rate of 25 frames per second and code rate of 2 048 kb/s at a resolution of 704 pixel × 574 pixel.

To analyze the performance of the algorithm, four dynamic scenes were used to compare the improved algorithm with the original one. The dynamic scenes include 49 226 frames at sunset, 744 frames in a sunny day, 276 frames on a rainy day and 601 frames at night with infrared extended sources. Three evaluation indicators were used, including the target false positive rate (V_{TFPR}), the background false positive rate (V_{BFPR}) and the processing time (V_{PT}). V_{TFPR} is the percentage of pixels belong to the target but misidentified as the background; V_{BFPR} is the percentage of pixels belong to the background but misidentified as the target. And they were calculated as follows:

$$V_{\mathrm{TFPR}} = \frac{|A_1 - A_2|}{A_1} \times 100\% \tag{17}$$

$$V_{\mathrm{BFPR}} = \frac{|F - A_2|}{F} \times 100\% \tag{18}$$

Where, A_1 is the real target areas segmented manual through Photoshop. A_2 is the detected target areas inside of the minimum circumscribed rectangle. F is the total detected areas.

3.1 The robustness of environment changes

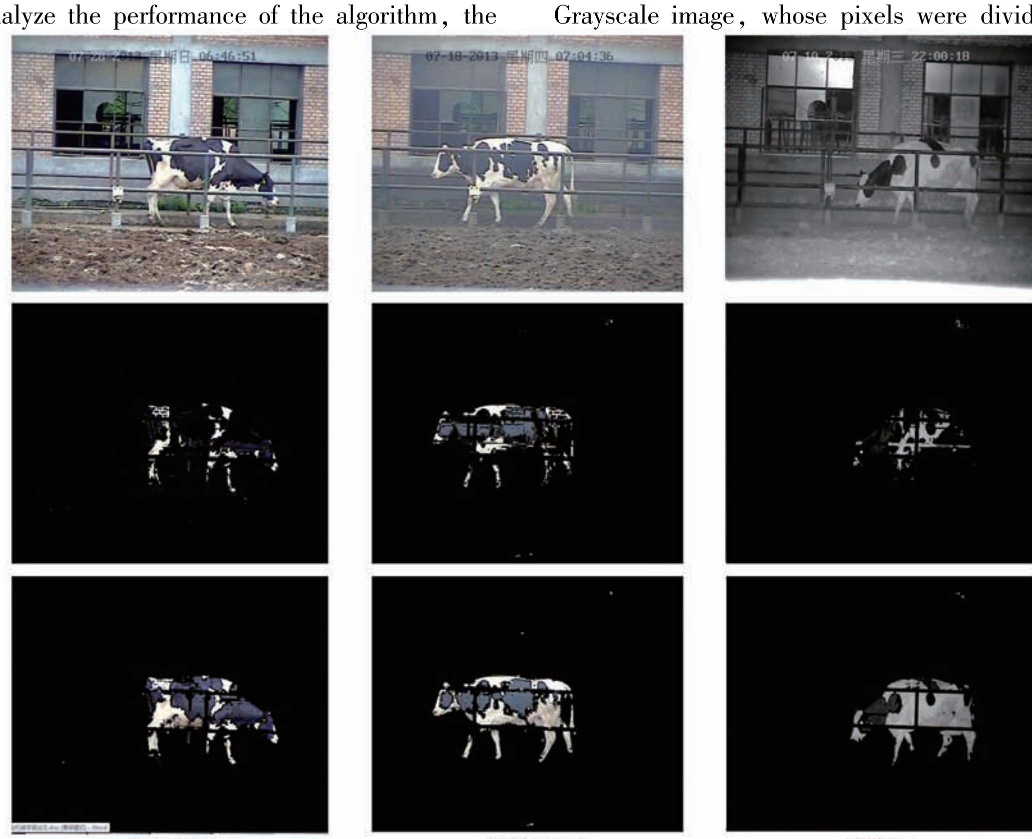
Detection results of the classical GMM and improved GMM under different conditions (sunny day, rainy day and night) were shown in Fig. 1 from left to right, including the original image, detection results by classical GMM and detection results by improved GMM from top to bottom. Obviously, the improved GMM detected much more complete targets, especially at night with single-channel of the priori information under infrared light source. The experimental results showed that the improved GMM could model the background environment accurately.

To analyze the performance of the algorithm, the changing trends of V_{TFPR} and V_{BFPR} in the process of sunset were recorded. And 32 frames including the cows' whole body were manually selected. The changing trends were showed in Figs. 2 and 3. The average target false positive rate was reduced from 37.68% to 18.18% and the average background false positive rate was reduced from 20.89% to 7.52%, indicating that the algorithm in this paper was better than the classical GMM. The background false positive rate expressed the accuracy of the model to describe the background. As shown in Fig. 3, the background false positive rate of classical GMM kept rising, indicating the poor adaptation to environmental gradient. But the general trend of improved GMM leveled off.

As a result, algorithm in this paper showed a better robustness to environmental gradient and more accurate detection than the classical GMM.

3.2 Analysis of model complexity

Generally, classical GMM describes a pixel with 3 – 5 components. In this paper, five components were introduced to give a full consideration to the complex areas. Besides, in order to compare the effect of constraints, the model complexity was visualized by Grayscale image, whose pixels were divided into five



(a) Sunny day (b) Rainy day (c) Night

Fig. 1 Result of moving cows segmentation

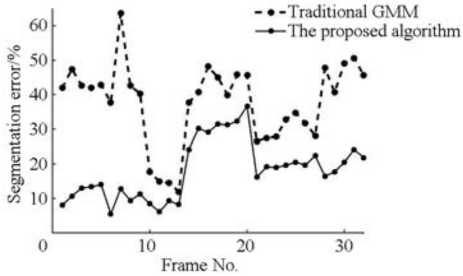


Fig. 2 Segmentation error of foreground

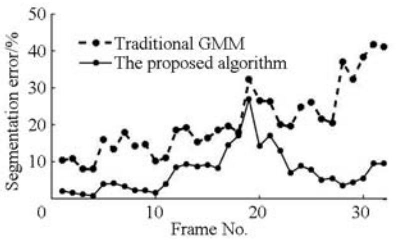


Fig. 3 Segmentation error of background

grades to represent the number of components. As shown in Fig. 4b, where there were darker, where the complexity were higher. Fig. 4c showed the statistical histogram of model complexity from where a cow walked through the camera visual field from left to right including 744 frames. It is evident from Fig. 4c that components and the complexity had the tendency to decline after the cow walking through, which was related to the dynamic background modeling. Comparing with the fixed components (five-components) of classical GMM, the minimum model complexity and the maximum model complexity were reduced by 56.3% and 45.4% respectively of all the 744 frames. The introduction of the penalty factor made the model complexity reduce by 50.85% on average, greatly improved the efficiency of algorithm. What is more, the areas needed five components to describe were consistent with the target indicating that the improved GMM only eliminated the redundant model and maximumly retained useful information.

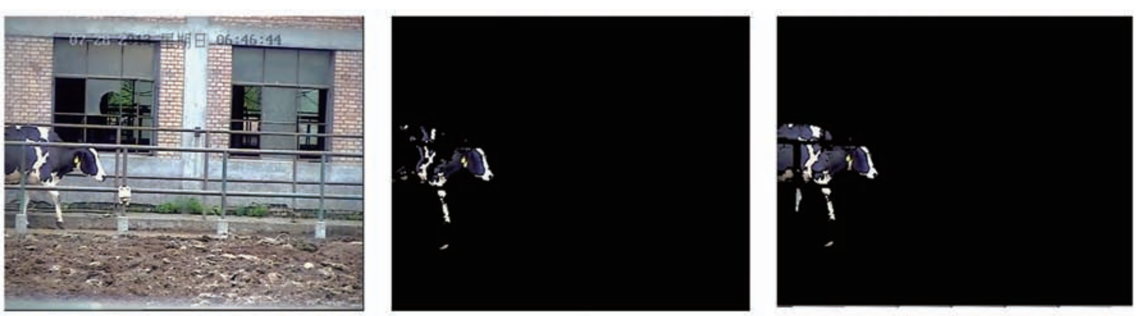
Big improvement can be observed in reducing processing time after testing 1 875 frames. The average processing time for single frame reduced by 29.25%, from 0.017 1 s to 0.012 1 s and was measured on an Inter(R) Core(TM) i5-4200M CPU @2.50 GHz PC. It is the foundation of body condition score, lameness detection etc.

(a) Original image (b) Complexity of the model (c) Statistical histogram of model complexity

Fig. 4 Visual model complexity

3.3 The foreground melting problem

To verify the effect of the local updating method to the foreground melting problem, we selected a video in which a cow walked into the camera view and stayed still for 20 s in the position. Fig. 5 showed the detection results of classical GMM and improved GMM. Obviously, local updating method could solved the foreground melting problem effectively.



(a) Original image (b) Result from classical GMM (c) Result from improved GMM

Fig. 5 Segmentation result after target staying for a long time

4 Conclusions

(1) Compared with classical GMM, the average model complexity reduced by 50.85% and the processing time reduced by 29.25% after introducing the constraint factor.

(2) The designed two-class classification algorithm based on the brightness bias and chromaticity bias was feasible to avoid the influence from the shadow. And the average target false positive rate reduces by 19.50% and the average background false positive rate reduces by 13.37%.

(3) The local updating method, which makes use of the relative invariance of cows' body in moving, solves the foreground melting problem.

References

[1] PLUYM L M, MAES D, VANGEYTE J, et al. Development of a system for automatic measurements of force and visual stance variables for objective lameness detection in sows: SowSIS[J]. Biosystems Engineering, 2013, 116(1): 64 – 74.

[2] FAN B, ONTERU S K, MOTE B E, et al. Large-scale association study for structural soundness and leg locomotion traits in the pig [J]. Genetics Selection Evolution, 2009, 41(3): 1 – 9.

[3] CONTE S, BERGERON R, GONYOU H, et al. Measure and characterization of lameness in gestating sows using force plate, kinematic, and accelerometer methods[J]. Journal of Animal Science, 2014, 92(12): 5693 – 5703.

[4] LIU Bo, ZHU Weixing, YANG Jianjun, et al. Extracting of pig gait frequency feature based on depth image and pig skeleton endpoints analysis [J]. Transactions of the CSAE, 2014, 30(10): 131 – 137. (in Chinese)

[5] JI Bin, ZHU Weixing, LIU Bo, et al. Video analysis for tachypnea of pigs based on fluctuating ridge-abdomen [J]. Transactions of the CSAE, 2011, 27(1): 191 – 195. (in Chinese)

[6] ZHAO Kaixuan, HE Dongjian, WANG Enze. Detection of breathing rate and abnormity of dairy cattle based on video analysis[J]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(10): 258 – 263. (in Chinese)

[7] LIU J, JIANG W, GUO Y. Study on application of image recognition technique in the cow body condition score[C] // 2011 IEEE 13th International Conference on Communication Technology (ICCT), 2011:373 – 376.

[8] WANG Long. Application of dairy cow body condition scoring based on machine vision [D]. Shanghai: Donghua University, 2014. (in Chinese)

[9] LI Ji'nan, SUN Peng, QIN Chunfu, et al. Research advance of body condition score in the management of feeding the dairy cows[J]. China Animal Husbandry & Veterinary Medicine, 2013, 10(10): 115 – 119. (in Chinese)

[10] ZHU Weixing, PU Xuefeng, LI Xincheng, et al. Automatic identification system of pigs with suspected case based on behavior monitoring[J]. Transactions of the CSAE, 2010, 26(1): 188 – 192. (in Chinese)

[11] PORTO S M C, ARCIDIACONO C, ANGUZZA U, et al. A computer vision-based system for the automatic detection of lying behaviour of dairy cows in free-stall barns[J]. Biosystems Engineering, 2013, 115(2): 184 – 194.

[12] LAO Fengdan, TENG Guanghui, LI Jun, et al. Behavior recognition method for individual laying hen based on computer vision [J]. Transactions of the CSAE, 2012, 28(24): 157 – 163. (in Chinese)

[13] YAN Li, SHEN Mingxia, XIE Qiuju, et al. Research on recognition method of lactating sows' dangerous body movement[J]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(1): 266 – 272. (in Chinese)

[14] WANG Cheng, WANG Runsheng. IR target detection under open country scene with a complex background [J]. Infrared and Laser Engineering, 2000, 1(1): 5 – 8. (in Chinese)

[15] LI Jianqiao, ZHANG Xiaodong, ZOU Meng, et al. 3-D surveying and gaits analysis of Chinese mitten crab in smooth road[J]. Transactions of the Chinese Society for Agricultural Machinery, 2012, 43(Supp.): 335 – 338. (in Chinese)

[16] ZHAO K X, HE D J. Target detection method for moving cows based on background subtraction [J]. International Journal of Agricultural & Biological Engineering, 2015, 8(1): 42 – 49.

[17] ZIVKOVIC Z. Improved adaptive gaussian mixture model for background subtraction [C] // IEEE International Conference on Pattern Recognition, 2004: 28 – 31.

[18] STAUFFER C, GRIMSON W E L. Adaptive background mixture models for real-time tracking [C] // IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1999.

[19] ZIVKOVIC Z, HEIJDEN F V D. Efficient adaptive density estimation per image pixel for the task of background subtraction [J]. Pattern Recognition Letters, 2006, 27(7): 773 – 780.

[20] STAUFFER C, GRIMSON W E L. Learning patterns of

activity using real-time tracking[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2000, 22(8): 747-757.

[21] ZHANG Yanping, BAI Yunqiu, ZHAO Yong, et al. Moving object detection based on improved Gaussian mixture models [J]. Computer Engineering and Applications, 2011, 46(34): 155-157. (in Chinese)

[22] LIN Qing, XU Zhu, WANG Shitong, et al. Moving objects detection of adaptive gaussian mixture models on HSV[J]. Computer Science, 2010, 10(10): 254-256. (in Chinese)

[23] LI Xuchao. Expectation maximization method for parameter estimation of image statistical model [J]. Journal of Image and Graphics, 2012, 17(6): 619-629. (in Chinese)

[24] CHEN Zujue, CHEN Xiaojun, HE Hong. Moving object detection based on improved mixture gaussian model [J]. Journal of Image and Graphics, 2007, 12(9): 1585-1589. (in Chinese)

[25] BAI Xiangfeng, LI Aihua, LI Xilai, et al. A novel background Gaussian mixture model [J]. Journal of Image and Graphics, 2011, 16(6): 983-988. (in Chinese)

doi:10.6041/j. issn. 1000-1298. 2016. 05. 039

基于混合高斯模型的移动奶牛目标实时检测

刘　冬　赵凯旋　何东健
(西北农林科技大学机械与电子工程学院, 陕西杨凌 712100)

摘要: 针对奶牛养殖场背景复杂和环境多变导致现有的目标检测算法无法满足鲁棒性和实时性需求的问题,基于递归背景建模思想,在混合高斯模型中引入惩罚因子,提出了一种动态背景建模方法,采用局部更新策略,以降低模型复杂度和解决前景消融问题;提出基于色度偏差和亮度偏差的二分类算法,避免目标物阴影区域的影响。对不同天气及环境变化剧烈情况下获取的奶牛视频样本进行实验。结果表明,与混合高斯模型相比,平均模型复杂度降低了50.85%,前景误检率和背景误检率分别降低了19.50和13.37个百分点,单帧运行时间降低了29.25%,检测准确率更高、实时性更好,且解决了前景消融问题,能满足在复杂背景和环境条件下实时提取奶牛目标的要求。

关键词: 移动奶牛; 视频分析; 目标检测; 阴影检测; 混合高斯模型
中图分类号: S24;TP319　　**文献标识码**: A　　**文章编号**: 1000-1298(2016)05-0288-07

Real-time Target Detection for Moving Cows Based on Gaussian Mixture Model

Liu Dong　Zhao Kaixuan　He Dongjian
(*College of Mechanical and Electronic Engineering*, *Northwest A&F University*, *Yangling*, *Shaanxi* 712100, *China*)

Abstract: Target detection is the basic work for analyzing the behavior of the cows using video analysis technology. It is difficult to extract the moving cows accurately and real-timely with the existing target detection methods because of the complex background environment. In this study, a series of improvement measures were proposed based on Gaussian mixture model to meet the system requirements. A dynamic background modeling method with penalty factor was proposed for the mathematical model of the background which can overcome the high model complexity. A two-class classification algorithm based on chromaticity distortion and brightness distortion was proposed to avoid the influence of the shaded area in the foreground detection algorithm. Local update method was proposed to avoid missing the target if it stays for a long time. In order to verify the validity of the algorithm, four evaluation parameters were introduced to test the detection algorithm including model complexity, false detection rate of foreground, false detection rate of background and processing time. Experimental results show that model complexity was 50.85% lower than the classical method. False detection rate of foreground and false detection rate of background were 18.18% and 7.52%, which had 19.50 and 13.37 percent lower than the classical Gaussian mixture model. Processing time of average single frame was 29.25% lower. Statistics indicate that the algorithm proposed in this study can improve the detection performance and it is an extension to classical Gaussian mixture model.

Key words: moving cows; video analysis; target detection; shadow detection; Gaussian mixture model

收稿日期: 2015-10-11　修回日期: 2016-01-21
基金项目: 国家自然科学基金项目(61473235)
作者简介: 刘冬(1990—),男,博士生,主要从事视频分析方法研究,E-mail: xnld@nwsuaf.edu.cn
通信作者: 何东健(1957—),男,教授,博士生导师,主要从事生物图像分析及识别、智能化检测与控制研究,E-mail: hdj168@nwsuaf.edu.cn

引言

智能视频监控是计算机视觉的一个重要研究领域，已经广泛应用于精准畜牧领域，例如：跛行检测[1-4]、呼吸检测[5-6]、体况评定[7-10]、疾病检测[10]、行为识别[11-13]等。使用视频监控技术感知动物运动行为已经成为精准畜牧业的研究热点。从视频中准确、实时地检测出运动目标是自动感知动物行为的基础。

国内外研究者近年来在智能感知动物行为领域取得了一定的成果。然而，由于复杂背景环境造成目标物提取困难，使得许多研究对环境条件要求苛刻，故有学者利用红外成像仪获取运动图像以便检测出目标，该方法不易受到外界环境影响[14]，但红外成像仪价格昂贵，不利于推广。为减少环境影响，李建桥等[15]在室内搭建实验平台，采用人造光源，获得了稳定的实验环境，常用的分析理论和方法均能相对准确地分离出背景与目标物。ZHAO 等[16]采用局部背景拼接的方法在复杂环境下建立了背景模型，并取得了较好的提取结果，但该方法对环境变化敏感，不利于长时间检测。现代畜禽生产中的奶牛、肉牛、羊等大型动物，由于其体形庞大且活动量大，饲养场所多为室外开放环境，具有环境因素渐变或突变的特点，例如：一天内太阳光亮度和角度渐变、照明设备开关的突变、风吹草动以及入侵飞鸟等。这些因素会对背景模型的建立和目标物的提取造成很大的误差，进而对动物行为感知研究造成影响。因此，现有的检测技术不能直接平移到开放环境中去，需要一种适应开放环境的移动目标检测方法，其中背景模型的准确建立是关键。

利用混合高斯模型（Gaussian mixture model，GMM）[17-20]可以建立动态背景模型，满足长时间观测的实际需求。但对于复杂环境背景建模问题，仍存在模型复杂度高、计算量大、无法根据局部背景差异建立最优模型个数以及无法解决目标物长时间停留造成的消融问题。目前，国内外提出了各种 GMM 改进方法，主要针对模型复杂、运算量大的问题提出了不同的简化方法。例如张燕平等[21]为解决经典 GMM 运算量大的问题，引入分块建模思想，将邻域内相关性较强的点视为“块”，对“块”的均值进行建模，以降低算法复杂度。但该方法以损失有用信息为代价换取时间效率，造成边缘锯齿化以及目标模糊化。GMM 的时效性与准确性矛盾是一个开放性问题，如何判定冗余信息，是快速、准确建立准确背景模型的关键。林庆等[22]提出一种模型筛检策略，在 GMM 迭代过程中某一模型的权值小于初始值时，被认为是多余模型并删除。但初值的选取依然缺乏理论依据，过大造成前景消融问题；过小则改进效果不明显。故本文在 GMM 基础上，以移动奶牛为研究对象，研究冗余模型判定方法，在不损失有用信息前提下加快 GMM 迭代速度，解决开放环境下的干扰问题、目标物阴影区域的影响及前景消融问题，为复杂环境下的移动奶牛实时提取提供一种新的方法。

1　背景建模方法

1.1　经典 GMM 复杂背景建模问题分析

经典 GMM 根据视频中每个像素在时间域上的概率分布信息来构建颜色分布模型，在智能安防等领域得到广泛应用。GMM 对每个像素建立了 M（一般为 3 ~ 5）个高斯分布模型。一个像素点在 t 时刻的近期采样值为 $\boldsymbol{x}_t=(x^{(t)},x^{(t-1)},\cdots,x^{(t-T)})$。对样本集 $\boldsymbol{x}_t$ 建立高斯分布模型，即

$$P(\hat{\pi}_m,\hat{\mu}_m,\hat{\sigma}_m^2|\boldsymbol{x}_t,B+F)=\sum_{m=1}^{M}\hat{\pi}_m N(x^{(t)};\hat{\mu}_m,\hat{\sigma}_m^2)$$
$$\left(0<\hat{\pi}_m<1,\ \sum_{m=1}^{M}\hat{\pi}_m=1\right)\tag{1}$$

式中　B——背景像素值
F——前景像素值
M——高斯模型个数
$N(x^{(t)};\hat{\mu}_m,\hat{\sigma}_m^2)$——第 m 个高斯函数
$\hat{\pi}_m$——第 m 个高斯模型的权值
$\hat{\mu}_m$——第 m 个高斯模型的均值
$\hat{\sigma}_m^2$——第 m 个高斯模型的方差

新添加的样本可能是背景像素也可能是前景像素，因此式（1）包含了 B 和 F。当得到新的样本集 $\boldsymbol{x}_t$，采用 EM 算法[23]（最大期望估计）求得背景模型递归更新方程为

$$\hat{\pi}_m\leftarrow\hat{\pi}_m+\alpha(o_m^{(t)}-\hat{\pi}_m)\tag{2}$$
$$\hat{\mu}_m\leftarrow\hat{\mu}_m+o_m^{(t)}(\alpha/\hat{\pi}_m)(x^{(t)}-\hat{\mu}_m)\tag{3}$$
$$\hat{\sigma}_m^2\leftarrow\hat{\sigma}_m^2+o_m^{(t)}(\alpha/\hat{\pi}_m)[(x^{(t)}-\hat{\mu}_m)^2-\hat{\sigma}_m^2]\tag{4}$$

式中　α——学习率，决定了参数更新速度
$o_m^{(t)}$——权属参数

若新样本遵循 3σ 原则并属于第 m 个模型，$o_m^{(t)}$ 取 1，否则取 0。

当前帧为前景目标时，GMM 会将前景像素量化为一个新的高斯模型，其权值 $\hat{\pi}_m$ 较小。按照式（1）更新后，对权值归一化并按由大到小的顺序排序，以确保权值累加和等于 1。使用权值累加和的前 90% 来表示背景模型，即

$$P(x^{(t)}|\boldsymbol{x}_t,B)=\sum_{m=1}^{B}\hat{\pi}_m N(x^{(t)};\hat{\mu}_m,\hat{\sigma}_m^2)\tag{5}$$

其中 $B=\arg\min_{b}\left(\sum_{m=1}^{b}\hat{\pi}_m>90\%\right)$

可以看出，首先，经典 GMM 对每一帧图像的每一个像素都要计算 M 个权重 $\hat{\pi}_m$、均值 $\hat{\mu}_m$ 和方差 $\hat{\sigma}_m^2$，其中有平方、开方以及指数运算，尤其对于 3 通道 RGB 视频图像的运算量大大增加。其次，背景模型的“90%”截取方法和模型个数 M 的取值无充分理论依据，选取不当将会降低目标检测的准确率。

奶牛养殖场为野外开放环境，容易受到外界环境因素影响，摄像机视野中，不同区域实际所需模型数不同。若取值偏小，模型对当前环境的描述处于“欠拟合”状态，目标物与背景接近时，容易将目标物误检为背景；若取值偏大，又会造成“过拟合”问题，模型数冗余，易将复杂背景中轻微摇晃的树叶误检为前景，并且无法避免摄像机的轻微抖动造成的影响。

运动奶牛视频采集实验是在摄像机固定的条件下进行的，大部分背景区域基本固定。故本文基于最小信息长度准则，提出一种自适应模型复杂度约束方法，既能自适应地选取模型个数，又能去除冗余模型，有效克服经典 GMM 在复杂背景下的缺点。

1.2　基于“最小样本”的模型复杂度约束

假设模型个数 M 固定，混合高斯模型的权值 $\hat{\pi}_m$ 可以看作样本来自第 m 个模型的概率，在统计意义上其服从多项分布，似然函数为

$$\ell_1(\boldsymbol{\pi},M|\boldsymbol{x})=\lg\left(\prod_{m=1}^{M}\hat{\pi}_m^{n_m}\right)$$
$$\left(0<\hat{\pi}_m<1,\ \sum_{m=1}^{M}\hat{\pi}_m=1\right)\tag{6}$$

式中　n_m——样本集 $\boldsymbol{x}_t$ 中第 m 个高斯模型的样本数

将拉格朗日乘子 λ 引入式(6)，则最大似然估计可以写为

$$\frac{\partial}{\partial\hat{\pi}_m}\left(\ell_1+\lambda\left(\sum_{m=1}^{M}\hat{\pi}_m-1\right)\right)=0\tag{7}$$

求解拉格朗日方程，消去 λ 后得

$$\hat{\pi}_m=\frac{n_m}{\sum_{m=1}^{M}n_m}\tag{8}$$

将式(8)写为递归形式即为式(2)所示的权值更新方程。可见，固定模型个数 M 条件下，极大似然估计法给每一个模型分配若干样本直到模型饱和，该方法无法避免“错误模型”，且无法依据不同区域自适应建立最优模型的个数。本文基于如下事实对上述过程进行改进：1 个像素点构成的样本 $\boldsymbol{x}_t$ 是由背景、目标物和噪声构成，且噪声样本数 N 远远小于其他两者。

对似然函数式(6)添加复杂度约束因子 N，即

$$\ell_2(\boldsymbol{\pi},M|\boldsymbol{x})=\lg\left(\prod_{m=1}^{M}\hat{\pi}_m^{n_m-N}\right)\tag{9}$$

式中，N 取较小的数(例如 30)，表示剔除每个模型样本中的噪声数，式(9)可改写为

$$\ell_2(\boldsymbol{\pi},M|\boldsymbol{x})=\lg\left(\prod_{m=1}^{M}\hat{\pi}_m^{n_m}\right)+\lg\left(\prod_{m=1}^{M}\hat{\pi}_m^{-N}\right)\tag{10}$$

式(10)中，右边第 2 项为模型复杂度约束项，引入拉格朗日乘子 λ，式(10)的最大似然估计为

$$\frac{\partial}{\partial\hat{\pi}_m}\left(\ell_2+\lambda\left(\sum_{m=1}^{M}\hat{\pi}_m-1\right)\right)=0\tag{11}$$

求得权重递归方程为

$$\hat{\pi}_m^{(t+1)}\leftarrow\hat{\pi}^{(t)}+\alpha(o_m^{(t)}-\hat{\pi}_m^{(t)})-\alpha N/T\tag{12}$$

式中　T——像素点在 t 时刻的近期采样数

模型复杂度约束项对每个模型赋予了 N 个“负样本”，即权值 $\hat{\pi}_m$ 初始化为负值。当有足够多样本来“支撑”第 m 个模型时，即满足 $n_m-N>0$ 时，则第 m 个模型的权值为正，此时 N 取最小样本数。在 GMM 算法迭代过程中，当检测到某一模型的权值为负值时，认为该模型是由噪声引起的“错误模型”，则抛弃这个模型，由此可以自适应选取最优模型个数，从而实现模型复杂度的约束。改进的 GMM 背景更新方程可以写为

$$\begin{cases}\hat{\pi}_m\leftarrow\hat{\pi}_m+\alpha(o_m^{(t)}-\hat{\pi}_m)-\alpha N/T\\ \hat{\mu}_m\leftarrow\hat{\mu}_m+o_m^{(t)}(\alpha/\hat{\pi}_m)(x^{(t)}-\hat{\mu}_m)\\ \hat{\sigma}_m^2\leftarrow\hat{\sigma}_m^2+o_m^{(t)}(\alpha/\hat{\pi}_m)[(x^{(t)}-\hat{\mu}_m)^2-\hat{\sigma}_m^2]\end{cases}\tag{13}$$

2　目标提取算法

经典 GMM 前景检测算法中，当新的一帧到来时，每个位置的像素分别与该位置对应的 M 个高斯模型匹配验证，若能与任意一个模型按照 2.5σ 原则匹配上，则认为当前像素为背景，同时按照式(13)更新背景模型；若与所有高斯模型都不匹配，则认为当前像素为前景，并建立新的模型来替换已有的权重最小的模型，新模型的权值设为较小值，均值设为当前像素值，方差设为较大的初始值[24-25]。

奶牛场开放环境中，上述前景检测算法不能很好地辨认出奶牛阴影，原因在于上述算法实质为 1 个二分类算法，阴影区域的亮度偏差和色度偏差处于 2.5σ 以外，被误检为前景。为解决该问题，本文选择 2 个特征来衡量阴影区域的亮度偏差和色度偏差，对前景再进行 1 次二分类，2 个特征分别为

$$\beta_{RGB}=\arg\min_{\beta}(x_{RGB}-\beta\mu_{RGB})^2=\frac{x_R\mu_R+x_G\mu_G+x_B\mu_B}{\mu_R^2+\mu_G^2+\mu_B^2} \tag{14}$$

$$D_{RGB}=\sqrt{(x_R-\beta\mu_R)^2\ (x_G-\beta\mu_G)^2\ (x_B-\beta\mu_B)^2} \tag{15}$$

式中　β_{RGB}——与背景模型期望的偏差

D_{RGB}——与背景模型期望的距离

x_R、x_G、x_B——当前像素 R、G、B 通道像素值

μ_R、μ_G、μ_B——与当前像素相匹配模型的 R、G、B 通道期望值

β_{RGB}越接近 1 表示当前像素与背景越接近。分类规则公式表示为

$$x_i=\begin{cases}S & (0.5<\beta_{RGB}<1\ 且\ D_{RGB}<4\sigma\beta_{RGB})\\ F & (其他)\end{cases} \tag{16}$$

式中　S——奶牛阴影　　F——移动奶牛

若 $0.5<\beta_{RGB}<1$，表示当前像素值比背景模型期望值小，亮度偏暗。D_{RGB}的阈值取 4σ 时，实际上放宽了背景阈值，符合阴影部分像素与背景模型期望值接近的直观理解。

式(13)中学习率 α 决定了背景更新的速度，预设值过大则造成背景更新过快，将移动缓慢的物体误分类为背景；预设值过小又达不到动态背景建模的目的。奶牛在摄像机视野中停留时间过长，会被逐渐训练成背景直至消失，仅仅靠调整参数无法彻底解决该问题[25]。在奶牛运动过程中，牛身部分相对不变，故将牛身位置信息作为是否停留的判断依据，采用局部更新的方法：根据前一帧检测结果用一个最大内接矩形和一个最小外接矩形标记前景目标，若后一帧最大内接矩形重心与前一帧的最大内接矩形的重心重合或小于一个阈值(2～5 个像素)，则表明前景目标停止移动，选择只更新外接矩形以外的区域，否则更新整个图像。此外，若检测超过 50 帧前景无移动，则更新整个图像，可有效避免目标物在摄像机视野边界停留造成的检测失效问题。

3　实验结果与分析

实验视频采集于陕西杨凌科元克隆股份有限公司的规模奶牛养殖场，拍摄对象为美国荷斯坦奶牛通过挤奶间与饮水槽之间宽 2 m、长 7 m 的走廊时的侧视图。为避免拍摄工作对奶牛活动的影响，摄像机(海康威视 DS－2DM1－714 型网络球形摄像机)安装于距离走廊 35 m、高 1.8 m 的支撑架上。CCD 传感器平行于走廊，焦距为 45 mm 以确保摄像机视野宽度大于牛身长度的 2 倍。采集视频为 PAL 制式，帧率 25 帧/s，码率 2 048 kb/s，分辨率 704 像素 × 574 像素。采集时间为 2013 年 7 月。

为了验证本文方法对环境渐变或突变的鲁棒性，人工截取一天傍晚日落过程 49 226 帧以及晴天 744 帧、雨天 276 帧、夜间(红外光源)601 帧视频，对改进算法与经典 GMM 提取移动奶牛的结果进行对比实验。

采用前景误检率(Target false positive rate，V_{TFPR})、背景误检率(Background false positive rate，V_{BFPR})和运行时间(Processing time，V_{PT})3 个评价指标。V_{TFPR}和 V_{BFPR}分别为属于目标的像素点被误判为背景的百分比和属于背景的像素点被误判为前景的百分比，计算公式为

$$V_{\mathrm{TFPR}}=\frac{|A_1-A_2|}{A_1}\times 100\% \tag{17}$$

$$V_{\mathrm{BFPR}}=\frac{|F-A_2|}{F}\times 100\% \tag{18}$$

式中　A_1——人工标注获得的奶牛面积

A_2——检测的奶牛面积

F——检测的全部前景面积

通过 Photoshop 进行手工分割获取 A_1，通过自动检测奶牛最小外接矩形，将矩形区域内目标像素数作为检测的奶牛面积 A_2。

3.1　对环境变化的鲁棒性分析

采用本文方法对晴天、雨天以及夜间环境下实验样本检测，结果如图 1 所示，其中从上到下依次为原始图像、经典 GMM 提取结果、本文算法提取结果。由图可见，不同环境条件下本文算法提取结果的奶牛目标更加完整，尤其在夜间(红外光源)只有单通道图像先验信息条件下，本文算法能够对当前环境背景准确建模。

对环境变化剧烈的日落过程奶牛目标检测进行实验，选取 32 帧具有有效牛身信息的提取结果进行统计，GMM 建模和本文建模方法的 V_{TFPR}、V_{BFPR}变化如图 2、3 所示。分析图 2、3 可知，本文算法提取效果优于经典 GMM。平均前景误检率由 37.68% 降低到 18.18%，降低了 19.50 个百分点。平均背景误检率由 20.89% 降低到 7.52%，降低了 13.37 个百分点。图 3 中的背景误检率实际上表达了所建背景建模的准确性，可以看出传统混合高斯模型的背景误检率呈现不断上升的趋势，表明随环境变化(日落过程)背景模型对环境的适应性变差。本文建模方法背景误检率趋于平稳。

上述结果表明，与经典 GMM 方法相比，本文方法对环境变化具有良好的鲁棒性，且背景模型更切合实际环境，前景检测算法准确性更高。

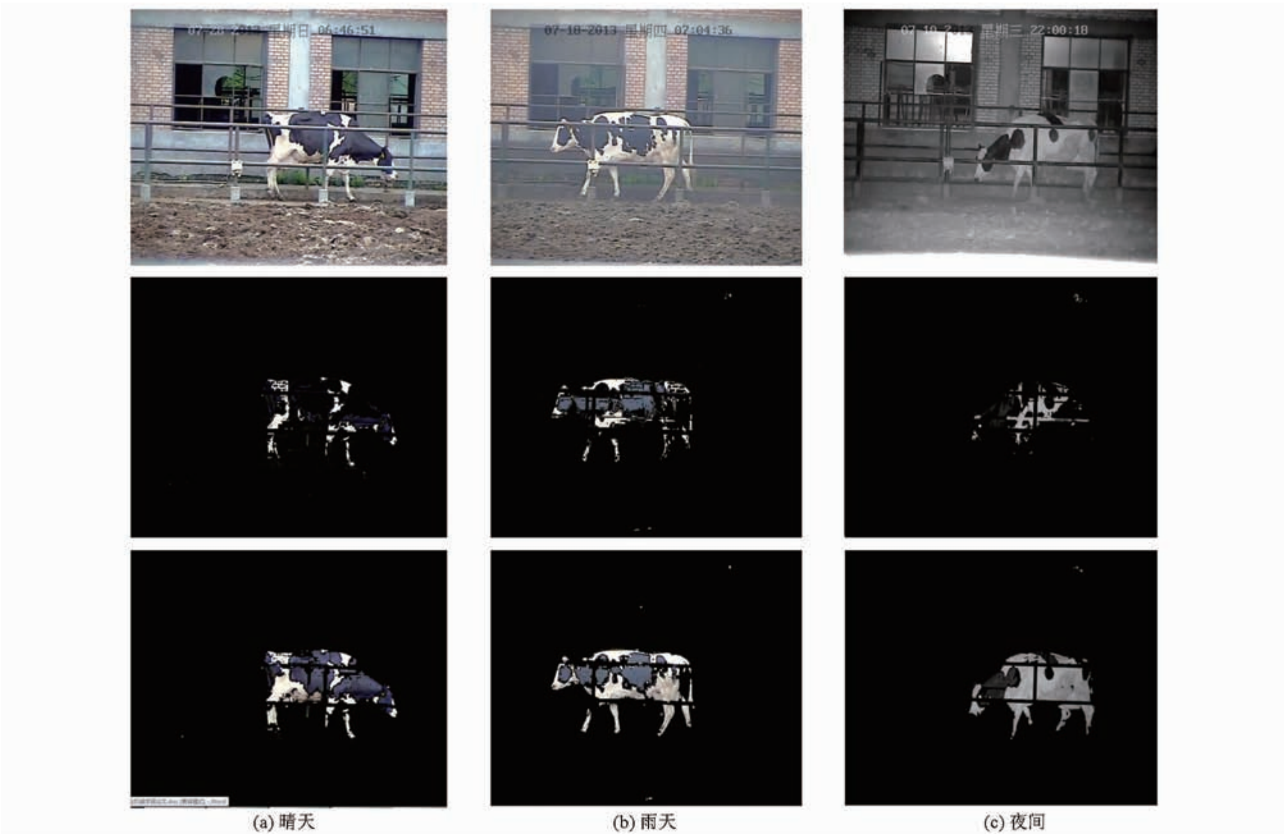


图1　移动奶牛提取结果

Fig. 1　Result of moving cows segmentation

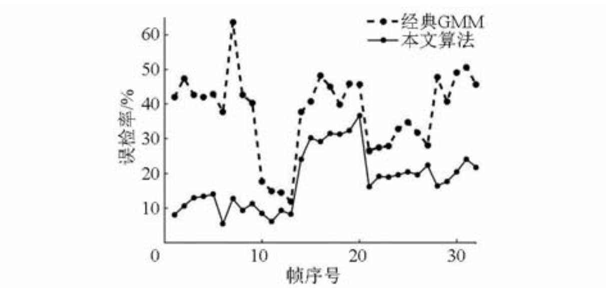


图2　前景误检率

Fig. 2　Segmentation error of foreground

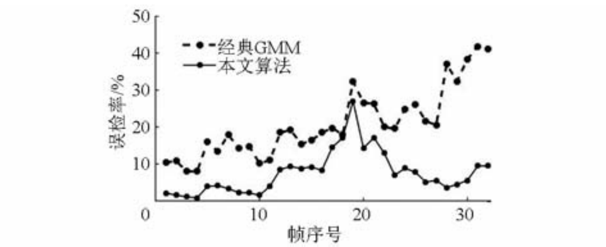


图3　背景误检率

Fig. 3　Segmentation error of background

3.2　模型复杂度分析

通常,经典GMM选用3~5个高斯函数来描述一个像素点的背景,本实验选用5个高斯函数以充分考虑部分点的复杂性。为了直观对比模型复杂度约束效果,采用灰度图表示每个像素点所使用的高斯函数个数,将灰度(0~255)等分为5个灰阶,用不同的灰度表示采用高斯函数的个数,像素点颜色越暗的区域使用的高斯函数个数越多,即模型复杂度越高。本文对奶牛从左向右缓步进入摄像机视野的744帧图像统计其模型的复杂度,如图4所示。图4a为第600帧原图,图4b为图4a可视化后的模型复杂度,表示每个像素位置使用的高斯函数个数。

图4c统计了使用不同高斯模型个数(1~5)的像素占全体像素的百分比。由图4可看出:①经过复杂度约束后,74%的区域仅用了1~3个高斯模型便可描述,且奶牛经过的区域,复杂度有降低的趋势,这与动态背景建模有关。②与固定模型数(5个)相比,统计出744帧视频图像处理过程中的最小模型复杂度和最大模型复杂度分别降低了56.3%和45.4%。表明本文引入约束惩罚项可使GMM的模型复杂度平均降低50.85%,大大提高了算法效率。③有17%的区域需要5个模型数来描述,与目标奶牛占全部像素的比例基本一致,表明本文方法最大程度保留了有用信息。

为验证本文算法的实时性能,在配置为Inter(R) Core(TM) i5-4200M CPU @2.50 GHz计算机上对1 875帧视频图像进行奶牛提取实验。与经典GMM相比,平均单帧处理时间从0.017 1 s降低到0.012 1 s,降低了29.25%。移动目标的高效提取为体况评价、步态分析等研究奠定了基础。

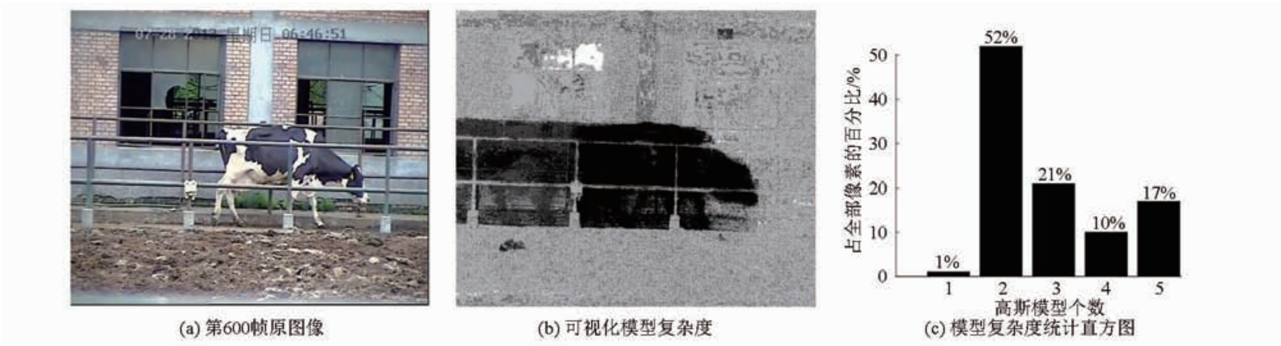


图 4　可视化模型复杂度

Fig. 4　Visual model complexity

3.3　前景消融问题

在 GMM 建模中，若奶牛在视野中停留时间过长，则会被逐渐训练成背景直至消失。为验证本文方法解决前景消融问题的效果，从原始实验样本中节选有奶牛停留的片段，视频中奶牛从左向右进入摄像机视野，在图 5a 所示位置停留了约 20 s。图 5b 和图 5c 分别为经典 GMM 和本文算法在奶牛停留了 15 s 后的提取结果。可以看出，采用局部更新法有效解决了奶牛长时间停留造成的前景消融问题。

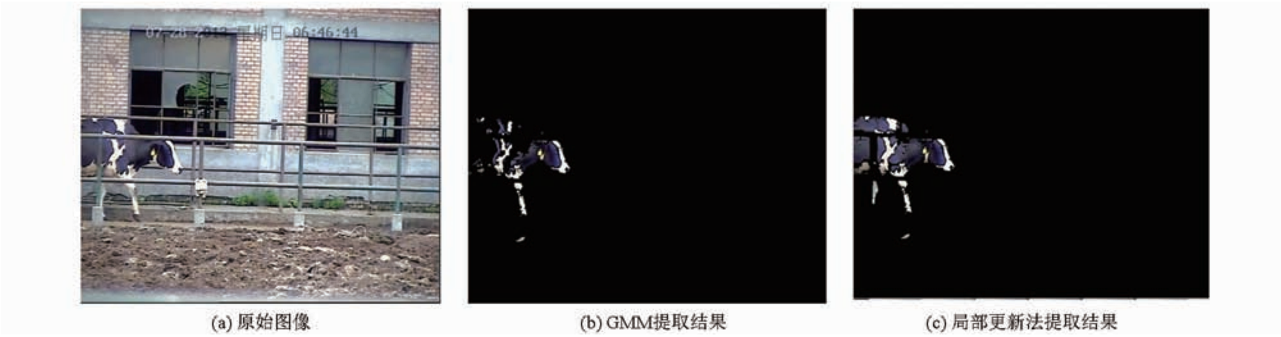


图 5　长时间停留后提取结果

Fig. 5　Segmentation result after target staying for a long time

4　结论

（1）引入了复杂度约束因子改进的混合高斯背景建模方法，与经典 GMM 方法相比，模型复杂度平均降低了 50.85%。

（2）在前景检测算法中引入亮度偏差和色度偏差，对奶牛目标进行二分类，减少了目标阴影造成的误检。且前景误检率和背景误检率分别降低了 19.50 和 13.37 个百分点，单帧运行时间平均降低了 29.25%。

（3）改进模型利用牛身位移尺度不变性，采用局部更新策略，有效解决了经典 GMM 中前景目标长时间停留引起的消融问题。

参　考　文　献

1　PLUYM L M, MAES D, VANGEYTE J, et al. Development of a system for automatic measurements of force and visual stance variables for objective lameness detection in sows: SowSIS[J]. Biosystems Engineering, 2013, 116(1): 64 – 74.

2　FAN B, ONTERU S K, MOTE B E, et al. Large-scale association study for structural soundness and leg locomotion traits in the pig[J]. Genetics Selection Evolution, 2009, 41(3): 1 – 9.

3　CONTE S, BERGERON R, GONYOU H, et al. Measure and characterization of lameness in gestating sows using force plate, kinematic, and accelerometer methods[J]. Journal of Animal Science, 2014, 92(12): 5693 – 5703.

4　刘波，朱伟兴，杨建军，等. 基于深度图像和生猪骨架端点分析的生猪步频特征提取[J]. 农业工程学报, 2014, 30(10): 131 – 137.
LIU Bo, ZHU Weixing, YANG Jianjun, et al. Extracting of pig gait frequency feature based on depth image and pig skeleton endpoints analysis[J]. Transactions of the CSAE, 2014, 30(10): 131 – 137. (in Chinese)

5　纪滨，朱伟兴，刘波，等. 基于脊腹线波动的猪呼吸急促症状视频分析[J]. 农业工程学报, 2011, 27(1): 191 – 195.
JI Bin, ZHU Weixing, LIU Bo, et al. Video analysis for tachypnea of pigs based on fluctuating ridge-abdomen[J]. Transactions of the CSAE, 2011, 27(1): 191 – 195. (in Chinese)

6　赵凯旋，何东健，王恩泽. 基于视频分析的奶牛呼吸频率与异常检测[J]. 农业机械学报, 2014, 45(10): 258 – 263.

ZHAO Kaixuan, HE Dongjian, WANG Enze. Detection of breathing rate and abnormity of dairy cattle based on video analysis[J]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(10): 258－263. (in Chinese)
7 LIU J, JIANG W, GUO Y. Study on application of image recognition technique in the cow body condition score[C]∥2011 IEEE 13th International Conference on Communication Technology (ICCT), 2011:373－376.
8 王龙. 机器视觉在奶牛体况评定中的应用[D]. 上海: 东华大学, 2014.
WANG Long. Application of dairy cow body condition scoring based on machine vision[D]. Shanghai: Donghua University, 2014. (in Chinese)
9 李吉楠,孙鹏,覃春富,等. 体况评分在奶牛饲养管理上应用的研究进展[J]. 中国畜牧兽医, 2013, 10(10): 115－119.
LI Ji'nan, SUN Peng, QIN Chunfu, et al. Research advance of body condition score in the management of feeding the dairy cows [J]. China Animal Husbandry & Veterinary Medicine, 2013, 10(10):115－119. (in Chinese)
10 朱伟兴,浦雪峰,李新城,等. 基于行为监测的疑似病猪自动化识别系统[J]. 农业工程学报, 2010, 26(1): 188－192.
ZHU Weixing, PU Xuefeng, LI Xincheng, et al. Automatic identification system of pigs with suspected case based on behavior monitoring[J]. Transactions of the CSAE, 2010, 26(1): 188－192. (in Chinese)
11 PORTO S M C, ARCIDIACONO C, ANGUZZA U, et al. A computer vision-based system for the automatic detection of lying behaviour of dairy cows in free-stall barns[J]. Biosystems Engineering, 2013, 115(2): 184－194.
12 劳凤丹,滕光辉,李军,等. 机器视觉识别单只蛋鸡行为的方法[J]. 农业工程学报, 2012, 28(24): 157－163.
LAO Fengdan, TENG Guanghui, LI Jun, et al. Behavior recognition method for individual laying hen based on computer vision [J]. Transactions of the CSAE, 2012, 28(24): 157－163. (in Chinese)
13 闫丽,沈明霞,谢秋菊,等. 哺乳母猪高危动作识别方法研究[J]. 农业机械学报, 2016,47(1):266－272.
YAN Li, SHEN Mingxia, XIE Qiuju, et al. Research on recognition method of lactating sows' dangerous body movement[J]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(1): 266－272. (in Chinese)
14 王程,王润生. 野外复杂背景下红外图像的目标检测[J]. 红外与激光工程, 2000, 1(1): 5－8.
WANG Cheng,WANG Runsheng. IR target detection under open country scene with a complex background [J]. Infrared and Laser Engineering, 2000, 1(1): 5－8. (in Chinese)
15 李建桥,张晓冬,邹猛,等. 中华绒螯蟹平面运动三维观测和步态分析[J]. 农业机械学报, 2012, 43(增刊):335－338.
LI Jianqiao, ZHANG Xiaodong, ZOU Meng, et al. 3-D surveying and gaits analysis of Chinese mitten crab in smooth road[J]. Transactions of the Chinese Society for Agricultural Machinery, 2012, 43(Supp.): 335－338. (in Chinese)
16 ZHAO K X, HE D J. Target detection method for moving cows based on background subtraction[J]. International Journal of Agricultural & Biological Engineering, 2015, 8(1): 42－49.
17 ZIVKOVIC Z. Improved adaptive gaussian mixture model for background subtraction[C]∥IEEE International Conference on Pattern Recognition, 2004:28－31.
18 STAUFFER C, GRIMSON W E L. Adaptive background mixture models for real-time tracking[C]∥IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1999.
19 ZIVKOVIC Z, HEIJDEN F V D. Efficient adaptive density estimation per image pixel for the task of background subtraction[J]. Pattern Recognition Letters, 2006, 27(7): 773－780.
20 STAUFFER C, GRIMSON W E L. Learning patterns of activity using real-time tracking[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2000, 22(8): 747－757.
21 张燕平,白云球,赵勇,等. 应用改进混合高斯模型的运动目标检测[J]. 计算机工程与应用, 2011, 46(34): 155－157.
ZHANG Yanping, BAI Yunqiu, ZHAO Yong, et al. Moving object detection based on improved Gaussian mixture models[J]. Computer Engineering and Applications, 2011, 46(34): 155－157. (in Chinese)
22 林庆,徐柱,王士同,等. HSV 自适应混合高斯模型的运动目标检测[J]. 计算机科学, 2010, 10(10): 254－256.
LIN Qing, XU Zhu, WANG Shitong, et al. Moving objects detection of adaptive gaussian mixture models on HSV[J]. Computer Science, 2010, 10(10): 254－256. (in Chinese)
23 李旭超. 图像统计模型参数估计中的期望最大值算法[J]. 中国图象图形学报, 2012, 17(6): 619－629.
LI Xuchao. Expectation maximization method for parameter estimation of image statistical model[J]. Journal of Image and Graphics, 2012, 17(6): 619－629. (in Chinese)
24 陈祖爵,陈潇君,何鸿. 基于改进的混合高斯模型的运动目标检测[J]. 中国图象图形学报, 2007, 12(9): 1585－1589.
CHEN Zujue, CHEN Xiaojun, HE Hong. Moving object detection based on improved mixture gaussian model[J]. Journal of Image and Graphics, 2007, 12(9): 1585－1589. (in Chinese)
25 白向峰,李艾华,李喜来,等. 新型背景混合高斯模型[J]. 中国图象图形学报, 2011, 16(6): 983－988.
BAI Xiangfeng, LI Aihua, LI Xilai, et al. A novel background Gaussian mixture model[J]. Journal of Image and Graphics, 2011, 16(6): 983－988. (in Chinese)