

Research Article

A Multiobjective Piglet Image Segmentation Method Based on an Improved Noninteractive GrabCut Algorithm

Feilong Kang , Chunguang Wang , Jia Li , and Zheyong Zong

College of Mechanical and Electrical Engineering, Inner Mongolia Agricultural University, Hohhot 010018, China

Correspondence should be addressed to Chunguang Wang; jdwcg@imau.edu.cn

Received 3 May 2018; Accepted 31 May 2018; Published 3 July 2018

Academic Editor: Chen Gong

Copyright © 2018 Feilong Kang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the video monitoring of piglets in pig farms, study of the precise segmentation of foreground objects is the work of advanced research on target tracking and behavior recognition. In view of the noninteractive and real-time requirements of such a video monitoring system, this paper proposes a method of image segmentation based on an improved noninteractive GrabCut algorithm. The functions of preserving edges and noise reduction are realized through bilateral filtering. An adaptive threshold segmentation method is used to calculate the local threshold and to complete the extraction of the foreground target. The image is simplified by morphological processing; the background interference pixels, such as details in the grille and wall, are filtered, and the foreground target marker matrix is established. The GrabCut algorithm is used to split the pixels of multiple foreground objects. By comparing the segmentation results of various algorithms, the results show that the segmentation algorithm proposed in this paper is efficient and accurate, and the mean range of structural similarity is [0.88, 1]. The average processing time is 1606 ms, and this method satisfies the real-time requirement of an agricultural video monitoring system. Feature vectors such as edges and central moments are calculated and the database is well established for feature extraction and behavior identification. This method provides reliable foreground segmentation data for the intelligent early warning of a video monitoring system.

1. Introduction

Pig farming is an important industry in China's agricultural economy [1]. Modern pig farming implements intelligent and digital management of the breeding industry by establishing large-scale modernized pig farms. Reducing the working intensity of the breeder and reducing the number of breeders are the most important factors to increase production, reduce the enterprise cost, and increase farmers' profits.

Artificial intelligence based on machine learning has important theoretical value and significant application potential in pig behavior recognition. Research on the precise segmentation of pig foreground targets is the basic work of advanced research such as artificial intelligence on target tracking and behavior recognition. This identifies piglets' movements, resting, or feeding activities and promptly determines whether piglets have been squeezed for long periods of time, thus issuing an early warning to alert keepers in time to rescue the piglets and improve their survival rate.

In complex environments, effective foreground detection is still the hotspot and challenge of the research. The mainstream background modeling detection methods, such as background difference, frame difference, optical flow, mixed Gaussian model, or background subtraction algorithm [2–7], cannot effectively detect a stationary foreground target. Each method has its limitations, and there is no common way to accurately handle all complex scenarios, especially if these include illumination changes, heavy shadows, or a foreground target that does not move for a long time.

The segmentation tracking algorithm that was proposed by and Schofield in 1995 is based primarily on initial segmentation and background estimation [2]; image difference is used to detect the intermediate background fusion method. This method requires the foreground target to move all the time; otherwise it is easy to lose the target. Perner in 2001 [3] detects pig position in grayscale videos by using background subtraction and threshold segmentation. This method is used by Lind for tracking pig behavior under the

effects of different drugs [4]. Viazzi proposes that there must be a knock, bite, or squeeze in the contact behavior lasting for more than 5 seconds and uses background subtraction to identify the attack behavior [5]. Gang provides a target detection method based on a mixed Gaussian model and two-dimensional wavelet transform [6]. This method is still based on background modeling; therefore detection of less mobile foreground targets is still invalid. Peter developed a very convenient pig behavior video monitoring system based on a mobile phone [7]. The tracking algorithm is divided into 2 steps: first, a map is established from a library of image segmentation data on pigs; second, a 5-dimensional Gaussian model is used to identify a pig's position and shape. The system can track 3 pigs simultaneously for up to 8 minutes without losing the target. However, the system also has a large number of identification problems in a complex piggery environment under the influence of dust and dirt.

The key to successful background modeling is to update the background model at all times. The main methods to do this include statistical average, median filter, Gaussian modeling, and mixed Gaussian background modeling. The existing background modeling method is inadequate for situations where there are illumination changes and less mobile foreground targets. When the light suddenly changes, most of the pixels will be detected as part of the foreground, which will affect the detection effect. Aiming to address the problem of light change, less mobile targets, and fast background changes, the foreground target detection method is more effective [8–13].

Bin Shao et al. developed a real-time image-processing system in 2008 to detect the movement and monitor the comfort of pigs [8]. In this paper, the global threshold method is adopted, while image noise is removed by the morphological method. The minimum European distance is used to distinguish the thermal comfort of pigs.

Navarro developed a system in 2009 to detect piglets in narrow circles [9]. The nine piglets were marked with color, and the system automatically tracked the piglets' position, so as to analyze their current status: whether they were squeezed into a group, were active, were suckling, etc. The biggest problem of the system is that expert experience is required to analyze and mark the target in the image in advance.

In 2013, Kashiha proposed a water consumption tracking analysis method based on image-processing algorithms [10] and in 2014 proposed an image detection method based on morphology and ellipse matching [11]. In 2014, Chang et al. proposed a detection method based on the foreground target [12]. The emphasis of the detection method is transferred to the algorithm research of the foreground target. Guo in 2015 proposed a precise segmentation method based on the multithreshold segmentation method [13]. The initial segmentation is divided by the maximum entropy global threshold, the distance curve from the center to the edge is calculated, and the circular radius is obtained. The accuracy of the threshold segmentation is improved based on the quadratic segmentation in each circular region. This method can segment the foreground target more accurately than other global segmentation algorithms.

Vincent et al. proposed the watershed method [14]. This method is fast but cannot control the number, size, and compactness of the superpixels. Boykov et al. proposed the Graph Cut algorithm [15]. In this paper, the optimal path is found by the maximum flow/minimum cut algorithm, and the optimal segmentation of the interactive target image is completed. Rother et al. proposed a GrabCut algorithm based on the Graph Cut [16]. A good segmentation result can be obtained by using a small amount of user interaction. In 2017, Sun et al. proposed a method of pig image segmentation based on an improved Graph Cut algorithm [17]. This method is based on the interaction of regional block watershed combined with the improved Graph Cut segmentation algorithm for image foreground and background, although interactive segmentation is unable to meet the requirements of a real-time video surveillance system.

This paper will focus on the image segmentation of multitarget piglets. Piglets are characterized by a wide range of activities, which are fast-moving and involve stacking and extrusion, which is difficult for image segmentation. It is automatically judged when piglets have been pressed for a long time, thus issuing a timely warning to feeding staff to rescue the piglet and improve the overall survival rate of piglets.

2. Methods

The proposed segmentation method adopts the combinational algorithm, including histogram equalization to enhance image contrast, bilateral filtering to preserve edge and reduce noise, and adaptive threshold segmentation method to automatically calculate the local threshold to complete foreground target extraction. This method uses the morphological processing method to simplify the image pixels and filter out the grille, cracks, and other background interference and establish the pixels' tag array, which may belong to the foreground and background as an input matrix of the GrabCut algorithm. Using a GrabCut algorithm to segment foreground target pixels of pigs can effectively improve the accuracy of recognition. By calculating edge, central moment, and other feature vectors of an elliptic fitting contour, which are used for feature extraction and behavior identification, the video monitoring system algorithm frame diagram is shown in Figure 1.

2.1. Bilateral Filtering. Bilateral filtering is a nonlinear filter that can achieve the effect of maintaining edge and smoothing [18–21]. The weight of the bilateral filtering not only considers the Euclidean distance of the pixel, but also considers the similarity between the center pixel and the neighborhood pixel. Bilateral filtering is also a weighted average method, representing the intensity of a pixel with a weighted average of ambient pixel brightness. Space weight usually uses the weighting calculation method of Gaussian filter by calculating the distance between two pixels [22]; the formula is as follows:

$$g(i, j) = \frac{\sum_{k,l} f(k, l) \omega(i, j, k, l)}{\sum_{k,l} \omega(i, j, k, l)} \quad (1)$$

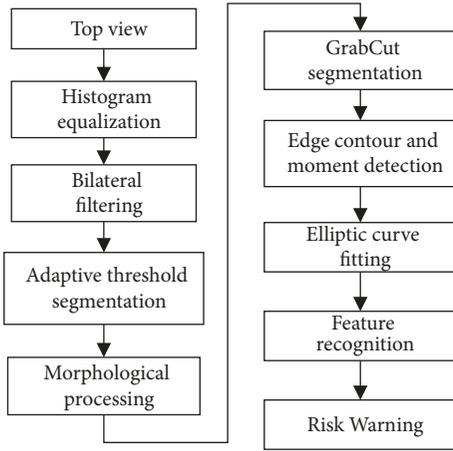


FIGURE 1: Algorithm frame diagram of the video monitoring system.

The weighting coefficient $\omega(i, j, k, l)$ depends on the kernel product of the domain and range.

The domain kernel is expressed as follows:

$$d(i, j, k, l) = \exp\left(-\frac{(i-k)^2 + (j-l)^2}{2\sigma_d^2}\right) \quad (2)$$

The range kernel is expressed as follows:

$$r(i, j, k, l) = \exp\left(-\frac{f(i, j) - f(k, l)^2}{2\sigma_r^2}\right) \quad (3)$$

When the two kernels are multiplied, the bilateral filtering weight functions are generated as follows:

$$\begin{aligned} \omega(i, j, k, l) \\ = \exp\left(-\frac{(i-k)^2 + (j-l)^2}{2\sigma_d^2} - \frac{f(i, j) - f(k, l)^2}{2\sigma_r^2}\right) \end{aligned} \quad (4)$$

2.2. Adaptive Threshold Segmentation. In the case of uneven illumination or distribution of gray values, if global threshold segmentation is used, the segmentation effect is often not satisfactory. The strong illumination regions are separated only by global threshold, while areas that are shaded or weakly illuminated are not segmented. In image filter processing, average filtering, Gaussian filtering, and median filtering use different rules to calculate the current pixel as the center of the neighborhood grayscale average. This principle can be applied to the threshold segmentation algorithm (an improved adaptive threshold segmentation method [23]) by setting the parameter to calculate adaptive and variable thresholds. The adaptive threshold of each pixel point is different and is obtained by calculating a weighted average of the pixel area.

In adaptive threshold processing, the size of the filter operator is determined by the size of the segmented object. If the size of the filter is too small, the calculated local threshold will not be ideal. The width of the filter operator must be

greater than the width of the object being recognized. The larger the size of the filter operator is, the better the result will be as a reference for the threshold value of each pixel. This paper uses the mean filter operator as the parameter of the adaptive threshold segmentation algorithm. First, the image is filtered; the result is denoted as F ; then the adaptive threshold reference matrix is obtained. $T = (1 - \text{ratio}) * F$. The segmentation results are obtained by using the adaptive threshold reference matrix.

2.3. Noninteractive GrabCut Algorithm. The Graph Cut algorithm is one of the classical algorithms of combinatorial graph theory. In recent years, many scholars have applied this algorithm to images and video segmentation, which has achieved good results. The Graph Cut algorithm is a kind of image segmentation technique based on a graph cutting algorithm. It requires human interaction markers and foreground and background pixels as input. The algorithm is set up by empowering the graph based on the various degrees of similar background and foreground pixels and by solving the minimum cutting to distinguish the foreground and background. The GrabCut algorithm is an image segmentation method based on the Graph Cut algorithm. The energy function is defined as follows:

$$E(\underline{\alpha}, k, \underline{\theta}, z) = U(\underline{\alpha}, k, \underline{\theta}, z) + V(\underline{\alpha}, z) \quad (5)$$

$$U(\underline{\alpha}, k, \underline{\theta}, z) = \sum_n D(\alpha_n, k_n, \underline{\theta}, z_n) \quad (6)$$

$$V(\underline{\alpha}, z) = \gamma \sum_{(m,n) \in C} [\alpha_m \neq \alpha_n] \exp(-\beta \|z_m - z_n\|^2) \quad (7)$$

$$\begin{aligned} D(\alpha_n, k_n, \underline{\theta}, z_n) = & -\log p(z_n | \alpha_n, k_n, \underline{\theta}) \\ & -\log \pi(\alpha_n, k_n) \end{aligned} \quad (8)$$

The U function represents the area data item of the energy function. The foreground and background mixture Gaussian models are used to indicate the probability that a pixel is a foreground or background pixel. V function represents the boundary of the energy function and the discontinuous penalty of neighborhood pixels between m and n . If the difference between two neighborhood pixels is very small, the possibility that they belong to the same foreground and background is very big. Conversely, the two pixels are likely to be edge and separated to foreground and background. The mixed Gaussian model is used to calculate the probability that each pixel belongs to the background or the foreground, and the image segmentation result is obtained by optimizing the energy function.

The traditional interactive GrabCut algorithm requires the user to provide an image with a marker rectangle containing the foreground target to extract the foreground in the image. In order to satisfy noninteractive video monitoring requirements, this paper takes the image of adaptive threshold segmentation results and obtains the array of tags that may be the foreground target and the background. Through multiple eroded treatments, the pixels in the image belong only to the foreground target, and pixel value of 255 is used to

mark the foreground target pixels. Elements that may belong to the background through multiple expansion are obtained. A combination of foreground target and the background of two markers in the image is used as the input image for the GrabCut algorithm. At the same time, the foreground target minimum envelope rectangles are used as input rectangles for the GrabCut algorithm. Because there are also some background pixels in the tag rectangle, the marking matrix is not completely correct. The Gaussian mixture model does not require that all training data are correct. The GrabCut algorithm uses the properties of the Gaussian mixture model, so that even if part of the classification is not correct, the final result is correct by iterative steps.

3. Experiments

The experimental data collection site is the pig farm of Inner Mongolia Agricultural University Experimental Base at the Inner Mongolia Agriculture and Animal Husbandry Science and Technology Park. It is located at 668 km of National Road 110, Tumd Right Banner, Baotou City, Inner Mongolia Autonomous Region. It is at 110.5° east longitude and 40.5° north latitude. The pig farm is a closed structure, with a building area of 290.83 m², 29.7 m east, north-south width 9 m. The ceiling is 2.4 m high, and the walls are insulated. Each circle of the experiment base is 3 m wide and 4 m long. The base is equipped with an automatic feeding system and a drinking water pipe. The ground uses a mesh grille. We set up the camera in the pig house, taking into account the angle of collection. Camera installation should avoid occlusion and missing pictures caused by different angles. The shooting distance should be kept at 1.5 m. The length of time is daylight hours (7 to 9 a.m.). The fixed position of the camera can effectively avoid shaking, so as to improve the stability of the video sample collection.

Video image is with 1440 * 1440 resolution, computer configuration for Intel (R) Core (TM) i7-4790 @ 3.6 GHz CPU, 16 GB of memory, graphics NV9500GT, 512 MB of memory, Windows operating system 10, in VS2012 programming environment with Open CV function library development of video monitoring system.

4. Results and Discussion

The experiment collected 160 hours of video images and randomly selected 200 typical images for a noninteractive algorithm test. Figure 2(a) is one of the original images. Figure 2(b) shows the result of histogram equalization. After histogram equalization, the contrast of the piglets is enhanced, but the image is susceptible to noise, shadow, and light changes. The histogram equalization mainly solves the problem of the low contrast of the piglets' images, enlarging the grayscale level of the output image to the specified degree so that the detail of the image becomes clearer. The effect of noise on the image after equalization is considerable: noise in the dark area may be amplified and become visible, while the bright area may lose information. Compared with the traditional mean filtering and Gaussian filtering methods, bilateral filtering preserves the edge of the target image based

on the filtered noise, which is the most important feature of bilateral filtering (Figure 2(c)). After filtering out the background of a brick wall and part of the grid, the edge of the foreground subject has been well preserved. The task of removing the grid is completed by image segmentation.

Figure 2(d) shows the result obtained by using the adaptive threshold segmentation method. The foreground object can be separated without human interaction. The filtering operator of the adaptive threshold segmentation algorithm is not ideal. As the size increases, the foreground object becomes more and more complete. The adaptive threshold segmentation can overcome the unevenness and shadow of the light, so that the target area can be segmented completely. The adaptive threshold segmentation may not completely remove the background, and other objects, such as iron pipes and grilles, which are similar to the gray value of the piglets, need to be processed by image morphological methods. After being eroded 2 times by morphological treatment (Figure 2(e)), the noise and background were removed. There is also part of a white screen in the upper right corner of the picture, which is very close to the piglet and can be temporarily retained.

Figure 2(f) is eroded 5 times; this shows the foreground of the target in the image pixels marked, with the result of the expansion after 5 times (Figure 2(g)) and the composite tag matrix (Figure 2(h)) as the input of the GrabCut algorithm images. Figure 2(i) shows the noninteractive GrabCut segmentation image as a result. It can be seen that the method can adaptively use the accurate segmentation of the pig target and effectively realize interactive real-time segmentation where the target and the background have a detailed division.

4.1. Comparison of Segmentation Effects. The traditional interactive GrabCut method, interactive watershed algorithm, OTSU threshold segmentation method, adaptive threshold segmentation algorithm, and the piglet image segmentation method based on the improved noninteractive GrabCut algorithm were compared and analyzed.

By analyzing the segmentation effect (Figure 3), the traditional GrabCut algorithm is more sensitive to local noise, it is time-consuming and unsatisfactory for extracting edges, and there is difficulty in performing multiobjective segmentation tasks, resulting in large error segmentation and poor segmentation results. The interactive watershed algorithm can distinguish the target area after marking the foreground object multiple times, but there is a large number of oversegmentation phenomena and the segmentation error is large. The interactive segmentation method fails to meet the noninteractive requirements of the video surveillance system. Although the adaptive OTSU global threshold segmentation algorithm is simple to calculate [24], it does not produce good results for nonuniform illumination pictures. Adaptive threshold can overcome the effects of sudden light changes and environmental changes. However, it cannot completely separate the foreground target from the grid and does not achieve the expected results. The proposed method adopts prior methods such as histogram equalization, bilateral filtering, adaptive threshold segmentation, and morphological

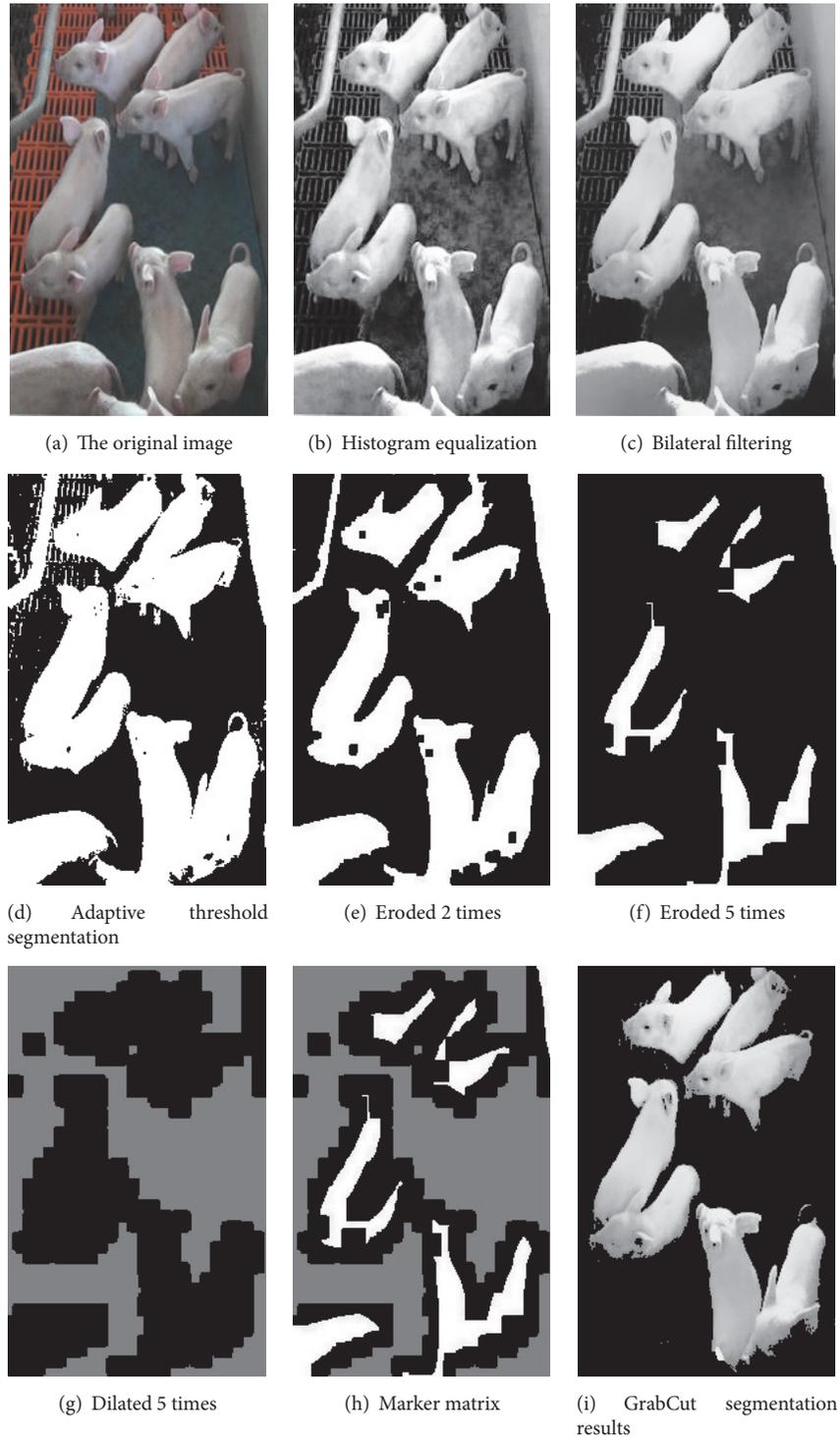


FIGURE 2: The segmentation of images of piglets by a noninteractive GrabCut algorithm.

processing to simplify the image and to extract the foreground target pixels. This greatly reduces the computational complexity of the GrabCut algorithm and not only realizes the automatic effective segmentation of multiple foreground targets, but also meets the real-time requirements of the video surveillance system.

All piglets' contours are detected on the segmentation of the pig image with the noninteractive GrabCut algorithm, and the small area of each contour is filtered out. In Figure 4, in the top view, the ellipse is more suitable to fit the pigs' shape, and the center moment of the contour can be calculated. The piglets can be tracked dynamically according

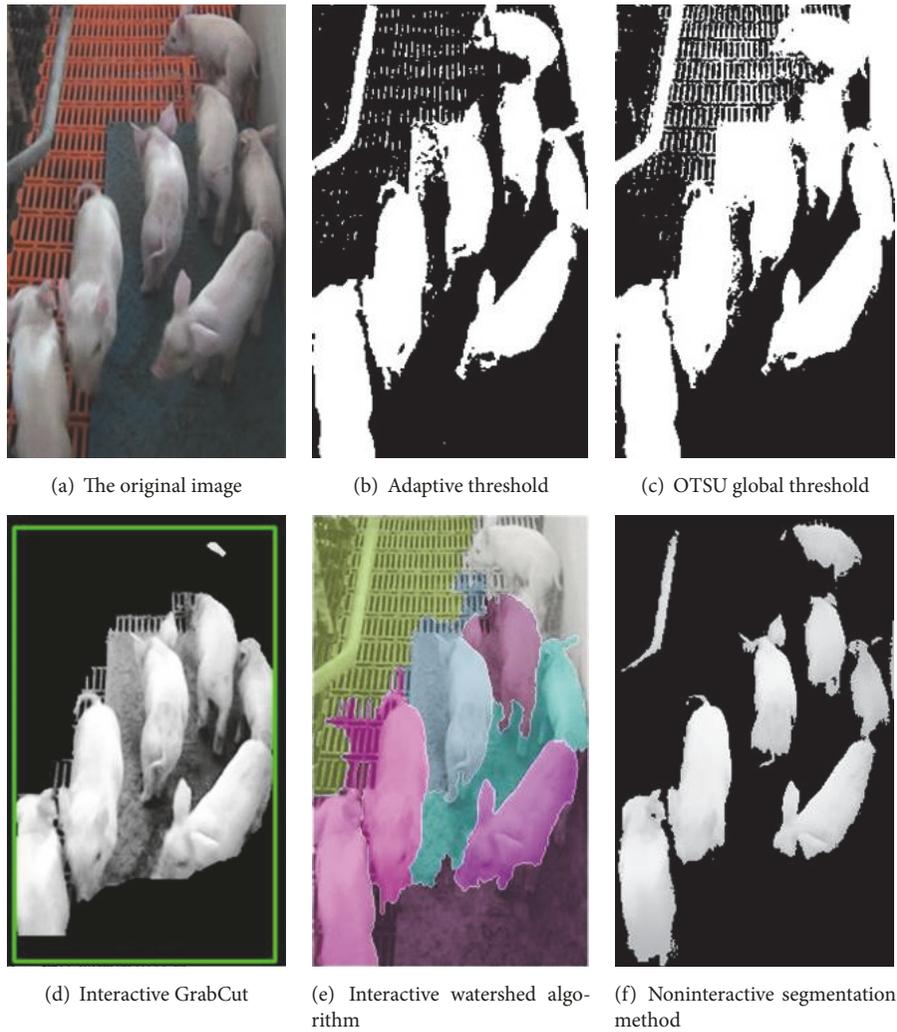


FIGURE 3: Segmentation effect of each segmentation method.

to the barycentric coordinates. By calculating the Euclidean distance between the central moments, the range of piglet movement can be dynamically observed and the minimum value of the Euclidean distance under normal conditions can be calculated. When the situation is dangerous, the minimum Euclidean distance is seriously below average. The system could send alerts to farmers in dangerous situations to prevent the piglets from being squeezed for a long time, thus increasing the survival rate of the piglets.

The results obtained by the segmentation method of the foreground object proposed in this paper can also be further studied in many aspects, for example, counting the time of contact between the ellipse and the trough and estimating the time of eating or drinking by the piglet. Calculating the linear relationship between the test area and the weight of the pig according to the size of the contour, so as to further estimate the weight of the piglet, according to the displacement changes of ellipse and the centroid, the pig's status is observed for a long time. It is clear that the research on the precise segmentation of the foreground target is the

basic work of artificial intelligence advanced research such as target tracking and behavior identification.

4.2. Comparison and Analysis of SSIM Indicators. Wang et al. proposed an image quality assessment method based on structural distortion, called the structural similarity (SSIM) method [25]. Structure similarity is an index to measure the similarity of two images, which reflects the similarity between the segmented image and the standard reference image. The structural similarity ranges from 0 to 1. When the two images are exactly the same, the SSIM value is equal to 1. The larger the similarity value, the higher the segmentation accuracy and the better the segmentation effect. The index can measure the distortion of the image structure and has certain versatility. The formula is

$$\text{SSIM}(S, T) = \frac{2\mu_S\mu_T + C_1}{\mu_S^2 + \mu_T^2 + C_1} * \frac{2\sigma_{ST} + C_2}{\sigma_S^2 + \sigma_T^2 + C_2} \quad (9)$$

$$\mu_S = \frac{1}{Z} \sum_{i=1}^Z S_i \quad (10)$$

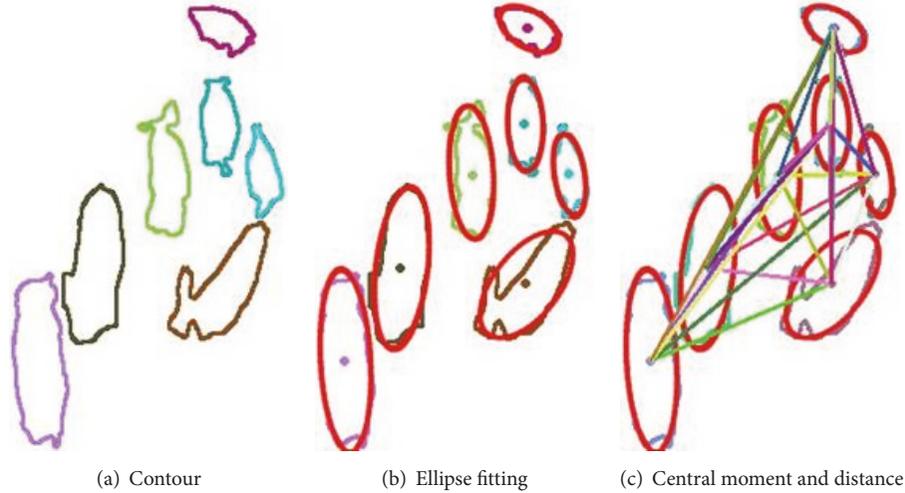


FIGURE 4: Results of the piglet ellipse fitting contour.

$$\mu_T = \frac{1}{Z} \sum_{i=1}^Z T_i \quad (11)$$

$$\sigma_s = \sqrt{\frac{1}{Z-1} \sum_{i=1}^Z (S_i - \mu_s)^2} \quad (12)$$

$$\sigma_T = \sqrt{\frac{1}{Z-1} \sum_{i=1}^Z (T_i - \mu_T)^2} \quad (13)$$

$$\sigma_{ST} = \frac{1}{Z-1} \sum_{i=1}^Z (S_i - \mu_s)(T_i - \mu_T) \quad (14)$$

In the formula, μ_s and μ_T denote gray mean values of the segmentation effect image S and the standard reference image T, respectively, which are the brightness metrics of the image structure information; σ_s and σ_T represent the standard deviation of S and T, which is a measure of the contrast of the image structure information; σ_{ST} is covariance between S and T, which is a measure of the image structure information; Z is the total number of pixels.

In order to compare and analyze SSIM indexes of different segmentation methods, standard segmentation needs to be obtained as evaluation criteria for segmentation effects. In the experimentally acquired images, 200 acquired images are selected, manually segmented images are used as reference images, and manual interactive segmentation is performed. The method can overcome interference items such as illumination, noise, and sundries, and the edge information can be preserved more completely.

Figure 5 shows the structural similarities obtained by using different segmentation algorithms for 200 images to analyze the SSIM similarity evaluation between different segmentation methods and reference images. As can be seen from Figure 5, the average range of structural similarity of the segmentation method in this paper is [0.88, 1]. The algorithm proposed in this paper can provide reliable foreground

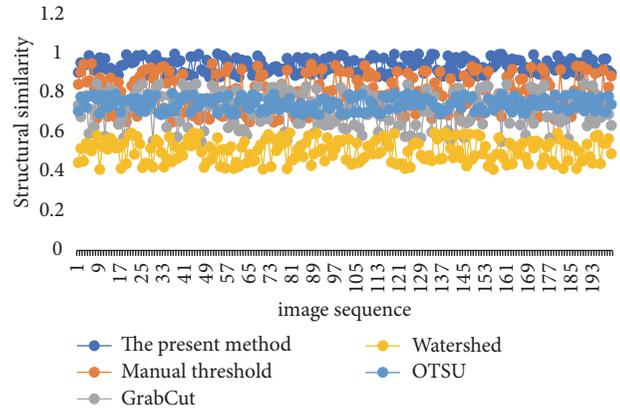


FIGURE 5: The structural similarity of different segmentation algorithms.

segmentation data for intelligent behavior identification for video surveillance systems.

4.3. Comparison of Segmentation Efficiency. To compare and analyze the segmentation efficiency of different segmentation algorithms, we collected 200 images of pigs randomly selected and divided into two groups, which included 10 test images. The average execution time of five different segmentation algorithms of each group is shown in Table 1.

The results of Table 1 show that the proposed algorithm is stable and effective, and the average program running time of different segmentation algorithms is calculated as follows: GrabCut algorithm 2062 ms and the proposed method 1606 ms. The proposed method uses histogram equalization, bilateral filtering, adaptive threshold segmentation, and the morphological processing method to simplify the image pixels and to extract the foreground target, sharply reducing the amount of calculation time of the GrabCut algorithm. The average time taken by the GrabCut algorithm is 22.1% lower than for conventional methods, so the efficiency is greatly

TABLE 1: Segmentation efficiency of different segmentation algorithms.

| Num. | Algorithm | Time (unit: ms) | | |
|------|------------------------|-----------------|---------|---------|
| | | Group 1 | Group 2 | Average |
| 1 | The proposed methods | 1498 | 1714 | 1606 |
| 2 | Threshold segmentation | 768 | 588 | 678 |
| 3 | GrabCut | 2361 | 1763 | 2062 |
| 4 | Watershed | 1839 | 1977 | 1908 |
| 5 | OTSU global threshold | 403 | 491 | 447 |

improved. In this paper, the segmentation method, combined with the advantages of the algorithm, not only realizes the effective segmentation of the automatic multiforeground target, but also satisfies the real-time requirements of the video monitoring system.

5. Conclusions

In this paper, we propose an image segmentation method based on an improved noninteractive GrabCut algorithm. Through histogram equalization and bilateral filtering (which can preserve edge and reduce noise), the adaptive threshold segmentation method is used to automatically calculate local thresholds and complete foreground target extraction, using the morphological processing method to filter out the grille, cracks, and other background interference pixels. The foreground target marker matrix is built, combining GrabCut algorithms to segment multiple pig targets. The average elapsed time is 1606 ms, which effectively improves the segmentation accuracy, and the foreground target structure similarity (SSIM) is in the average range of [0.88, 1].

The segmentation method proposed in this paper not only realizes the effective segmentation of noninteractive multiforeground targets, but also satisfies the real-time requirement of a video monitoring system. The research on the precise segmentation of the foreground objects is the basic work of the advanced research on target tracking and behavior recognition, which can be combined with artificial intelligence and machine learning techniques to further deepen the real-time analysis of pig behavior research and environmental monitoring.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was funded by “Twelve-Five” National Science and Technology Support Program (Project 2014bad08b05-04).

References

- [1] W. Huang and X. Shunlai, “Current situation and development of China’s pig industry,” *Livestock and Poultry Industry*, vol. 9, pp. 4–8, 2011.
- [2] J. A. Marchant and C. P. Schofield, “Extending the snake image processing algorithm for outlining pigs in scenes,” *Computers and Electronics in Agriculture*, vol. 8, no. 4, pp. 261–275, 1993.
- [3] P. Perner, “Motion tracking of animals for behavior analysis,” in *International Workshop on Visual Form*, pp. 779–786, Springer-Verlag, 2001.
- [4] N. M. Lind, M. Vinther, R. P. Hemmingsen, and A. K. Hansen, “Validation of a digital video tracking system for recording pig locomotor behaviour,” *Journal of Neuroscience Methods*, vol. 143, no. 2, pp. 123–132, 2005.
- [5] S. Viazzi, G. Ismayilova, M. Oczak et al., “Image feature extraction for classification of aggressive interactions among pigs,” *Computers and Electronics in Agriculture*, vol. 104, no. 2, pp. 57–62, 2014.
- [6] G. J. Tu, H. Karstoft, L. J. Pedersen, and E. Jørgensen, “Segmentation of sows in farrowing pens,” *IET Image Processing*, vol. 8, no. 1, pp. 56–68, 2014.
- [7] P. Ahrendt, T. Gregersen, and H. Karstoft, “Development of a real-time computer vision system for tracking loose-housed pigs,” *Computers and Electronics in Agriculture*, vol. 76, no. 2, pp. 169–174, 2011.
- [8] B. Shao and H. Xin, “A real-time computer vision assessment and control of thermal comfort for group-housed pigs,” *Computers and Electronics in Agriculture*, vol. 62, no. 1, pp. 15–21, 2008.
- [9] J. M. Navarro-Jover, M. Alcañiz-Raya, V. Gómez et al., “An automatic colour-based computer vision algorithm for tracking the position of piglets,” *Spanish Journal of Agricultural Research*, vol. 7, no. 3, pp. 535–549, 2009.
- [10] M. Kashiha, C. Bahr, S. A. Haredasht et al., “The automatic monitoring of pigs water use by cameras,” *Computers and Electronics in Agriculture*, vol. 90, no. 90, pp. 164–169, 2013.
- [11] M. Kashiha, C. Bahr, S. Ott et al., “Automatic weight estimation of individual pigs using image analysis,” *Computers and Electronics in Agriculture*, vol. 107, no. 3, pp. 38–44, 2014.
- [12] G. J. Tu, H. Karstoft, L. J. Pedersen, and E. Jørgensen, “Foreground detection using loopy belief propagation,” *Biosystems Engineering*, vol. 116, no. 1, pp. 88–96, 2013.
- [13] Y.-Z. Guo, W.-X. Zhu, P.-P. Jiao, C.-H. Ma, and J.-J. Yang, “Multi-object extraction from topview group-housed pig images based on adaptive partitioning and multilevel thresholding segmentation,” *Biosystems Engineering*, vol. 135, pp. 54–60, 2015.
- [14] L. Vincent and P. Soille, “Watersheds in digital spaces: an efficient algorithm based on immersion simulations,” *IEEE*

- Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 6, pp. 583–598, 1991.
- [15] Y. Boykov and M. P. Jolly, “Interactive Graph Cuts for Optimal Boundary & Region Segmentation of Objects in N-D Images,” *IEEE International Conference Computer Vision*, vol. 1, pp. 105–112, 2001.
- [16] C. Rother, V. Kolmogorov, and A. Blake, “GrabCut—interactive foreground extraction using iterated graph cuts,” in *Proceedings of the ACM SIGGRAPH (SIGGRAPH '04)*, pp. 309–314, August 2004.
- [17] L. Sun, Y. Li, Y. Zou, and Y. Li, “Pig image segmentation method based on improved graph cut algorithm,” *Transactions of the Chinese Society of Agricultural Engineering*, vol. 33, no. 16, pp. 196–202, 2017.
- [18] C. Tomaci and R. Mabduchi, “Bilateral filtering for gray and color images,” in *Proceedings of the 6th IEEE International Conference Computer Vision*, pp. 839–846, Bombay, India, January 1998.
- [19] B. Weisst, “Fast median and bilateral filtering,” *Acm Transactions on Graphics*, vol. 25, no. 3, pp. 519–526, 2006.
- [20] J. Dorsey, “Fast bilateral filtering for the display of high-dynamic-range images,” in *Proceedings of the Conference on Computer Graphics and Interactive Techniques*, pp. 257–266, ACM, 2002.
- [21] S. Paris and F. Durand, “A fast approximation of the bilateral filter using a signal processing approach,” *International Journal of Computer Vision*, vol. 81, no. 1, pp. 24–52, 2009.
- [22] Q. Jin, I. Grama, C. Kervrann, and Q. Liu, “Nonlocal means and optimal weights for noise removal,” *SIAM Journal on Imaging Sciences*, vol. 10, no. 4, pp. 1878–1920, 2017.
- [23] D. B. G. Roth, “Adaptive thresholding using the integral image,” *Journal of Graphics Gpu & Game Tools*, vol. 12, no. 2, pp. 13–21, 2007.
- [24] N. Otsu, “A threshold selection method from gray-level histograms,” *IEEE Transactions on Systems Man & Cybernetics*, vol. 9, no. 1, pp. 62–66, 2007.
- [25] Z. Wang, A. C. Bovik, and E. P. Simoncelli, “Structural approaches to image quality assessment,” *Handbook of Image and Video Processing*, vol. 8, pp. 961–974, 2005.

