

Design of UAV Cooperative Target Tracking System Based on Computer Vision Algorithm

Yucao Tang

University of California, Irvine, California, 94086, United States
*Corresponding author: yucaot@uci.edu

Abstract: In this paper, a video intelligent monitoring and tracking system is designed based on Computer Vision algorithm. The system is developed and implemented with OpenCV computer vision library, Visual Studio and MFC framework. Firstly, the overall framework of the system is designed, and the second is to analyze the preprocessing of video data. Then, the target is extracted by the Gaussian mixture background difference method, and the target is separated from the background by thresholding. Then, the target is tracked by Camshift and Kalman filters and an alarm prompt is issued. The final system implementation results show that the moving target extraction based on the Gaussian mixture background difference method, and through binary processing, can be accurately and reliably separated from the background, and the target detection effect is remarkable; the alarm mechanism can be triggered accurately and timely, and the alarm prompt information can be issued.

Keywords: computer vision algorithm; video surveillance; tracking

1. Introduction

Under the trend of newer video data recording and transmission technology, people's need for security monitoring continues to increase^[1]. To a large extent, it promotes the long-term and stable development of video monitoring systems. Video surveillance has its own unique advantages. The collected video data can help people observe all the conditions within the monitoring range in real time to prevent sudden accidents and dangerous events^[2]. However, the existing video surveillance systems mostly use manual monitoring methods to continuously monitor, not only the loss of human resources is too large, but also the long-term monitoring can easily cause the on-duty personnel to be distracted and miss important data information. Accordingly, this paper introduces computer vision technology for video intelligent monitoring, autonomously evaluates abnormal conditions based on video data content, and prompts alarms^[3].

2. The framework of video intelligent monitoring and tracking system

The video intelligent monitoring and tracking system based on computer vision algorithm is composed of two parts: hardware and software^[5]. The software framework takes the host as the carrier, uses the camera to collect the video image data information within the monitoring range, performs preprocessing and analysis, and uses OpenCV visual processing^[6]. The library compiles the code to detect the collected data information, and after detecting and discovering the moving object, it is evaluated in the I value area, and the alarm mechanism is triggered^[7]. The video image is presented on the PC side, and is compiled in the human-computer interaction interface using Visual Studio, so that the PC side can watch and monitor in real time and observe the changing state of the parameters. Hardware includes computers, cameras, network equipment, storage media, and so on.

2.1 Video capture module

The front-end video acquisition module is based on an integrated digital camera, LVDS SerDes chip, processor chip DM368, DDR2 SDRAM, and Flash Ethernet controller. Its functions are shown in Figure 1.

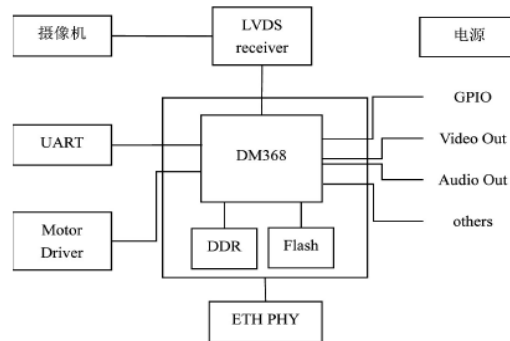


Figure 1: Functional Framework

DM368 can provide a variety of complete external interfaces, audio CODEC can be integrated inside the chip at the audio level^[8], and only a few peripheral components can be directly connected to the input output device to collect and output audio data. The front-end video capture module supports multiple interfaces such as UART, PWM, etc., thus providing sufficient space for its function expansion. Through storage facilities such as mobile hard disks, the front-end storage function can be realized. The PWM output can effectively control the front-end gimbal, and flexibly and appropriately adjust the visualization angle of the system^[9].

2.2 Video transmission module

Based on the DM368 of the front-end video acquisition module, the video image data information is compressed, and the code is written, and it is encapsulated into an IP data packet according to the relevant protocol^[10]. Through the DM368 back end using the Ethernet control chip as the carrier, the front-end network communication related functions can be realized. The main purpose of the video intelligent monitoring and tracking system is security. When designing, it is necessary to fully consider the security and stability of information transmission to ensure the reliability and accuracy of system data and information. The front end selects a private network to access people, and is assisted by EPON and security gate data penetration technology, which can not only meet the needs of mass data transmission of video images, but also ensure the quality and safety of transmission communication. As an important component of the video intelligent monitoring system, the mobile video acquisition terminal can transmit video image data information in real time without wiring, which is very convenient and reliable.

2.3 Video Management Analysis Module

The video image management analysis module can implement unified optimization management and control for video intelligent monitoring and tracking system resources. The front-end video acquisition module has a fixed face recognition module in the video encoding and decoding chip, which can analyze and process video image information for the first time, but the deep intelligent analysis still needs to rely on the video image management and analysis module to complete. The video management analysis module mainly includes central management server, network video recording, access gateway, intelligent video processing, etc. Among them, as the core component of the video management analysis module, the central management server can optimize the management and control of front-end, server, application modules, customers, etc.; the report combining intelligent video processing and front-end intelligent analysis mainly adopts behavior pattern recognition, target detection and analysis, automatic Intelligent tracking and other related technologies, based on the comparison and analysis of intelligent video analysis results and intelligent video processing background data identification system effective information, feedback processing countermeasures to the central management server in real time.

2.4 The principle of multi-UAV cooperative target tracking

When multiple UAVs cooperatively track a single ground target, the three-dimensional position of the target measured by the cameras in each UAV and the state of the UAV are estimated by an information fusion model to obtain the estimated value of the target state, which is then obtained by the cooperative control strategy. For each UAV control command, according to the different ways of collaborative information fusion, the information fusion input is the UAV measurement vector and state vector.

3. Research on UAV target recognition technology based on computer vision

3.1 Preface

At present, UAV target recognition technology has a very broad application prospect in UAV military and civilian fields. The algorithm of UAV target image segmentation, feature extraction, target recognition and other algorithms are systematically researched and compared and analyzed, and the simulation experiment of the UAV target image recognition algorithm is carried out to verify the feasibility of the algorithm. Automatic obstacle avoidance technology provides a better technical foundation. Target recognition technology refers to the technology that uses radar and computers to identify distant targets. UAV target recognition technology is the basis of UAV automatic obstacle avoidance technology. At present, UAV target recognition technology has played a very important role in UAV application fields such as aerial photography, disaster relief, power inspection, and target search. effect. Therefore, the research on UAV target recognition technology has very important application value.

3.2 UAV target image recognition algorithm process

The target image recognition algorithm flow of the UAV is shown in Figure 2.

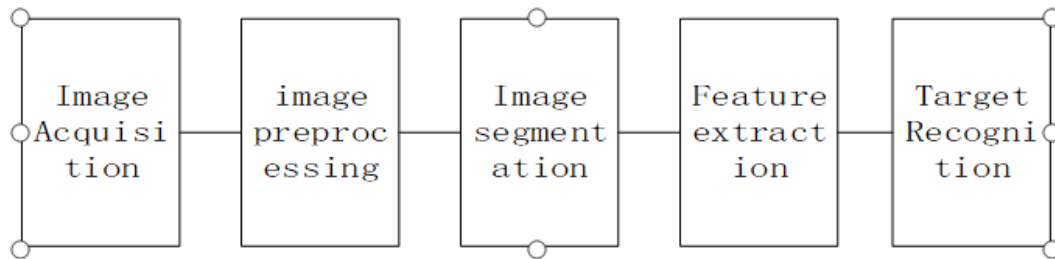


Figure 2: The algorithm flow of target image recognition of UAV

After the image sensor obtains the target image of the UAV, the collected target image must be pre-processed such as noise filtering and image enhancement. In order to separate the target image of the UAV to be recognized from the collected target image, It is necessary to perform the target image segmentation of the UAV on the filtered and image-enhanced images. After the target image of the UAV is segmented, feature extraction is performed, so that the essential features of the target of the UAV can be extracted from the image information. After the feature extraction The classification features are classified and the target of the UAV is judged, that is, the classification judgment, and finally the target of the UAV is identified and tracked.

3.3 Research on the target image recognition algorithm of UAV

3.3.1 Image preprocessing

(1) image filtering

Image filtering is to remove the noise in the original image. There are many methods of image filtering, such as mean filtering, median filtering, etc. Although the mean filtering can reduce the noise, the edges of the filtered image become blurred. Median filtering can not only reduce noise, but also make the edge of the image clearer, especially for salt and pepper noise, the filtering effect is better.

(2) image enhancement

Image enhancement is divided into two categories: image enhancement based on spatial domain and image enhancement based on frequency domain: the algorithm based on frequency domain is to modify the transform coefficient value of the image in a certain transform domain of the image; image enhancement based on spatial domain The enhancement algorithm is to directly operate on the gray level of the image, such as histogram correction, grayscale transformation and other algorithms. Histogram equalization can make the image clearer, enhance the grayscale contrast between the target and the background, and expand the grayscale difference between the target and the background, so that the edge details of the image can be detected more easily

3.3.2 Image segmentation

Image segmentation refers to the image processing technology that divides the image into several meaningful regions or parts according to the needs. The quality of the segmentation directly affects the effect of subsequent image processing [3]. After the image is quantized, it is necessary to perform image segmentation. The algorithm of image segmentation is mainly divided into the following three categories: 1) segmentation method based on threshold; 2) segmentation method based on edge; 3) segmentation method based on region. Region segmentation methods include region splitting and merging, and region growing.

(1) Threshold segmentation method

The 17-value image segmentation method is an effective and simple method. The 17-value method is characterized by simple operation and high efficiency, and is widely used. To increase the robustness of image segmentation, image segmentation needs to have automatic 17-value selection.

(2) Edge segmentation method

The edge is the most basic feature of the image, which can reflect the discontinuity of the grayscale of the image. Image edge detection is an important basis for image segmentation, image feature extraction, and image texture analysis. This kind of algorithm is simple and has a good effect on image segmentation in simple scenes. However, in complex scenes, incomplete target boundaries or false edges are prone to occur. There are many other problems, it is difficult to segment the target and it is sensitive to noise. Commonly used image edge detection algorithms are: Roberts, sobel, prewitt and canny and other operators for edge detection.

1) Roberts operator edge detection algorithm

The Roberts operator is a first-order differential operator that uses the template operator of the mask structure to perform spatial convolution operations. The Roberts operator template is as follows:

$$\begin{array}{cc}
 \begin{array}{|c|c|}
 \hline
 +1 & 0 \\
 \hline
 0 & -1 \\
 \hline
 \end{array} &
 \begin{array}{|c|c|}
 \hline
 0 & +1 \\
 \hline
 -1 & 0 \\
 \hline
 \end{array} \\
 G_x & G_y
 \end{array}$$

The characteristic of Roberts operator is that it has no smoothing and filtering links. It directly finds the edge through differential calculation, so it is more sensitive to subtle noise;

2) Sobel operator edge detection algorithm

The Sobel operator template is as follows:

$$G_x = \begin{bmatrix} -1 & 1 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & -1 \end{bmatrix}, G_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

The Sobel operator is a first-order differential operator that uses the template operator of the mask structure to perform spatial convolution operations. The Sobel operator has better noise suppression ability than Roberts operator because of increasing the weight of the central image point.

3) Prewitt operator edge detection algorithm

The Prewitt edge detection operator is performed in the image space r by using the template in the two directions of the horizontal edge and the vertical edge to perform the neighborhood convolution operation on the image. The operator is usually expressed by the following formula:

The template of the Prewitt operator is as follows:

$$G_x = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}, G_y = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$$

The Prewitt operator regards the pixels whose gray value is greater than or equal to the threshold value as edge points. Because the gray value of many noise points is very large, it is easy to misjudge the noise points as edge points, and it is easy to lose the relatively large amplitude value. small edge points.

(3) Area division method

The basic idea of region growing is to check its neighbors, determine whether they are similar, and assign similar neighbors to a region. This similarity criterion can be color, grayscale, texture, or other properties. Although the segmentation effect of this kind of algorithm is good, the operation is complicated and the real-time performance is poor.

3.3.3 Image feature extraction

Commonly used image features are: color feature, texture feature, spatial relationship feature and shape feature. The advantage of the extraction method based on color features is that it has a certain robustness, and its color component features are not affected by the change of viewing angle and the size of the object. . The advantage of the extraction method based on texture features is that it has a certain robustness to noise, but the disadvantage is that it may be affected by the image resolution, resulting in a large deviation. The advantage of the extraction method based on spatial features is that it can strengthen the ability to distinguish the content and description of the image, but the disadvantage is that it is more sensitive to changes such as rotation and scaling of the image. The advantage of the extraction method based on shape features is that it is more effective in the identification of objects of interest.

3.3.4 Target recognition

At present, the methods of target recognition at home and abroad mainly include: classic statistical recognition method, knowledge-based target recognition method, model-based statistical recognition method, information fusion-based target recognition method, and statistical learning-based target recognition method recognition methods.

(1)Based on the classical statistical identification method

The method trains a large number of samples, observes the distribution of target characteristics statistically, and uses feature matching classification technology to identify targets in the scene definition domain in the pattern space distance. This method cannot correctly identify the target in a complex background or when the target changes.

(2)Knowledge-based target recognition method

This method mainly focuses on the identification of available knowledge sources, the effective organization and verification of knowledge when applying to new scenarios, and the understanding and explicit expression of rules. The combination of domain knowledge and processing power determine the performance of the recognition algorithm.

(3)Model-based statistical identification method

The method utilizes target, environment, background and sensor models to identify the target, and matches the actual and predicted features according to the model features.

(4)Target recognition method based on information fusion

The method is to integrate various sensors with different functions to improve the system's ability to recognize objects in complex scenes. The target recognition method mainly focuses on the research of the algorithm and the development of the model.

(5)object recognition method based on statistical learning

The method obtains some statistical features by learning a large number of target and non-target samples, and builds a recognition model based on these features to identify the target. This method can distinguish samples better and has better robustness.

3.4 Simulation experiment and analysis

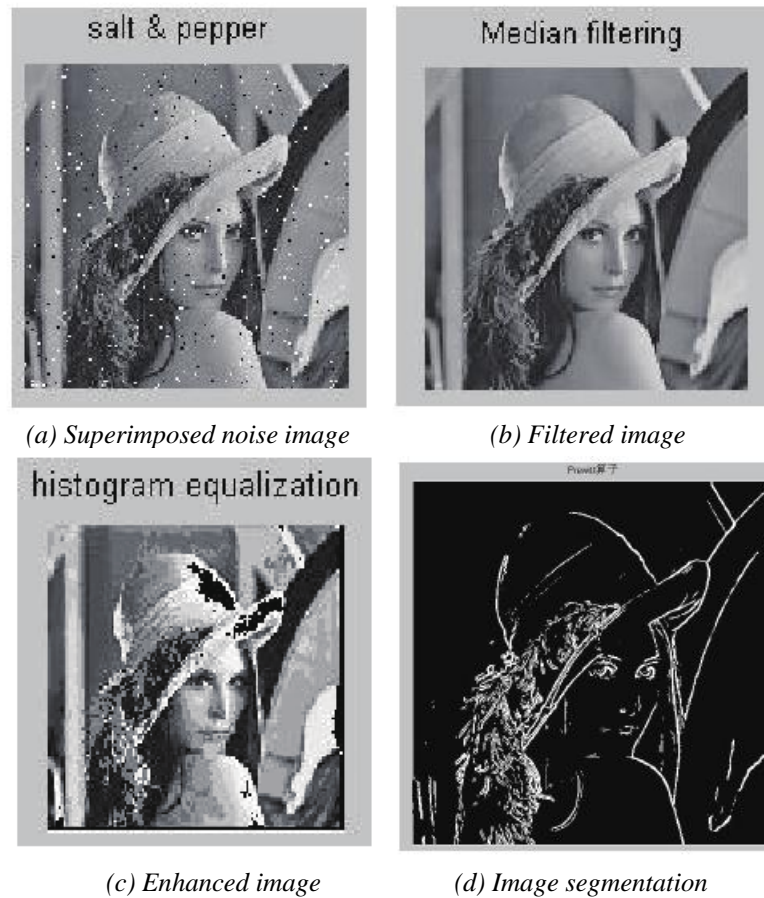


Figure 3: Target image simulation experiment

Figure 2 is the target recognition experiment using the lena image with salt and pepper noise, Figure 3 (a) is the lena image superimposed with salt and pepper noise, Figure 3 (b) is the median filter of the lena image with salt and pepper noise, after filtering, not only a large amount of salt and pepper noise is filtered out, but also many important details of the image are preserved, and the image is relatively clear. Figure 3(c) is the histogram equalization of the median filtered image, in order to enhance the grayscale contrast between the target and the background and expand the grayscale difference between the target and the background, so that the image can be more easily detected. After the histogram equalization of the image, the image is obviously enhanced, and the details of the image are more prominent and clear. Figure 3(d) is an image obtained by adopting the edge-based image segmentation method for the enhanced image, and a satisfactory edge detection and segmentation effect is obtained.

4. Conclusion

This paper studies the existing UAV target recognition technology in detail, systematically researches and compares the algorithms of image segmentation technology, feature extraction technology, and target recognition technology, and conducts a UAV target image recognition algorithm. The simulation experiment provides a technical basis for the UAV automatic obstacle avoidance technology, which will have a wide range of application value in the military and civilian fields.

References

- [1] Xiao Guideng. Research on video intelligent monitoring and tracking system based on computer vision algorithm [J]. Automation Technology and Application, 2021, 40(6):5.
- [2] Xu Zhengshuo. Thoughts on automatic navigation of unmanned aerial vehicles based on computer vision technology [J]. 2022(3).

- [3] Li Yuxia, He Lei, Xu Yongxin. *Design of teaching experimental platform for "Introduction to Remote Sensing of UAVs" based on image recognition [J].* 2022(5).
- [4] Xu Bo. *Research on video target tracking based on deep learning [D]. Dalian University of Technology, 2020.*
- [5] Zhang Haifeng. *Design of high-precision differential positioning UAV system based on RTK [J]. Surveying and Mapping and Spatial Geographic Information, 2022, 45(2):4.*
- [6] Yu Xiaowen, Tang Tao, Zou Yiwei. *Construction of an intelligent monitoring system for drilling operations based on computer vision technology [J]. China Petroleum and Chemical Standards and Quality, 2021, 41(21):2.*
- [7] Tan Jianhao, Zhang Siyuan, Wang Yaonan, et al. *A UAV target tracking method and system based on color histogram similarity: CN110766723A[P]. 2020.*
- [8] Zhang Yu. *ROS-based UAV target detection and visual positioning system [D]. Shandong University, 2020.*
- [9] Peng Cong, Chen Li. *A system and method for high-altitude parabolic detection, tracking and warning based on computer vision: CN112800953A[P]. 2021.*
- [10] Xin Shaou, Chen Ke, Song Zhenlin, et al. *Multi-UAV cooperative target tracking algorithm based on pure orientation [J]. Electronic Design Engineering, 2021, 29(24):6.*