Research and Application of Moving Target Tracking Algorithm Based on Computer Vision

Dong Fei

Shaanxi Railway Institute, Weinan, Shaanxi, China, 714000

Keywords: tracking algorithm; image segmentation algorithm; machine vision

Abstract: Computer vision technology has developed rapidly in the current computer world, and the processing methods for video images have also been significantly improved. Based on the computer theory based on Marr, this paper discusses and studies the tracking of moving objects in video. Taking the motion change of the monitored vehicle as an example, the Robert algorithm is used to track the changes of the vehicle during driving, and the multi-association template matching method is used to improve the accuracy of the algorithm. It is proved in the experiment that the algorithm in this paper has better precision in the tracking of the vehicle and can meet the current needs for tracking the moving target.

1. Introduction

Before the computer appeared, the most important channel for humans to obtain information from the outside world was the human eye, which relied on the human visual system [1-2]. With the development of the industrial revolution, the emergence of computers, the human eye is more and more unable to keep up with the speed of the times, so it is necessary to rely on the power of science. The emergence of computer vision technology has greatly expanded the inadequacy of the human visual system. With the help of various current sensing devices, human tracking of various dynamic targets has become possible.

"Smart Life" is the new wave of technology pursued by the current world and one of the most popular technology topics [3-4]. In urban life, the essential means of transportation for people to travel is a wide variety of vehicles. At present, the development of urbanization in China, the traffic pressure of the city has increased greatly, various traffic accidents have emerged in an endless stream, and traffic has been frequent. Therefore, the smart life of the city is more urgent for the demand of intelligent transportation. In the field of intelligent transportation, the tracking and video analysis of road traffic has become the key research object of intelligent transportation [5-6]. Under such a background, this paper sets the goal of the movement as the research object of the moving car in the urban road as the research object of the dynamic target, so as to hope to provide useful help to the relevant researchers.

2. Visual tracking theory

After years of computer development, people have made great strides in the study of dynamic target tracking. However, when a 3D image is mapped into a 2D image in a computer, there is a problem of loss of information, noise points, and definition of a motion concept by a person. There will also be problems in the future development of computer vision systems, such as non-rigid features, unclear targets, occlusions, complex environments, and changes in illumination. Therefore, for computer tracking algorithms, its robustness and accuracy are the requirements of future development [7-8].

2.1. Application of fractal in visual tracking

Euclid has been studying and exploring the world for a long time after the establishment of geometry, but in practice, many natural phenomena cannot be explained by Euclid geometry in the last century. In the 1970s Mandelbrot proposed fractal geometry [9-10]. According to his point of view, in space, the dimension is not necessarily an integer, and there is a dimension of the score. In this fractal nonlinear system, the research analysis does not simplify the object and study the law of the object. As a result, in the natural phenomenon, many geometric problems that cannot be explained are solved after the fractal geometry is proposed. On the basis of fractal geometry, Tan et al. compress the fractal image coding of computers and obtain good research results in face recognition. With the development of science, the combination of fractalism and visual tracking will have far-reaching effects [11].

2.2. Bio-vision based tracking

The "cocktail party" phenomenon is the result of the dreams of current computer vision researchers. In the 1980s, Marr first integrated image processing, human psychology, biology and medical results from the perspective of computer information processing, and proposed the first complete computer vision system framework in computer history. The structure of this framework is a bottom-up type, modular and one-way. The borrowing of computers on biological vision systems has enabled computers to have mechanisms that mimic biological vision. The current use of computers to establish artificial neural network models is another important development direction for computers to solve visual problems.

Marr's theory of computer vision is an information processing theory that combines the knowledge of subjects such as physics, biology, and computer image processing. Under this visual theory, there are mainly three stages in the generation of visual images.

(1) Generate a two-dimensional sketch. The initial two-dimensional sketch of the image uses the edge segments, lines, spots, and endpoints of the image to describe the brightness, and then uses the virtual lines to describe the geometric relationship between them, so that a certain range of scales 2D sketch can be obtained.

(2) Generate a 2.5-dimensional diagram. After the initial sketch is obtained in the previous step, it can be processed to obtain a description of the geometric features of the object, such as the direction and distance of the target object, their continuity, reflectivity, etc., and the lighting is estimated roughly.

(3) Generate a three-dimensional model. Further processing the sketch of the previous step, realizing the three-dimensional structure of the object, establishing a central coordinate system centering on the object, and obtaining the spatial structure of the image.

In the visual tracking of images, there are two ways to use more, one is bottom-up, and the other is top-down. This article uses the first method. Combined with Marr's visual theory, the generation of visual images in this paper can be divided into the following three stages: 1. Layer vision; 2. Intermediate vision; 3. High-level vision. These three segments correspond to the hierarchical phase of Marr's theory.

Bottom-up image tracking is to obtain the position, velocity, and acceleration of the motion. Therefore, when performing image processing, the first thing to do is to lock the target of the motion and determine the motion information of the target, that is, the position, speed, and direction. This paper studies the tracking algorithm with the goal of urban sports, and implements the tracking framework based on Marr's theory. The tracking of the vehicle can be divided into the following three stages: 1. locking of the vehicle; 2. extraction and identification of the locked vehicle; 3. obtaining information such as the position, direction, speed and the like of the target vehicle.

3. Tracking algorithm for moving targets

3.1. Kalman filter target tracking algorithm

The Kalman filter target tracking algorithm is one of the more commonly used algorithms in the currently used tracking algorithms. The algorithm first appeared in the 1960s. When the era combined the filtering theory with the state space model, a recursive estimation algorithm, namely Kalman filtering theory, was obtained. The basic idea of this theory is to use the measured estimates at the previous time point and the values measured at the current time point to continuously change the variables of the model to maintain the validity of the model. After years of development, this kind of algorithm has been widely used in the field of computer image tracking. In this algorithm, the state of the target at time K is a small amount of x_k, then the formula for the target to be initialized at time t_0 is: $\widehat{x_k} = \phi_{k,k-1} x_{k-1}$; at time t_k update the state of the target: $\widehat{x_k} = \phi_{k,k-1}\widehat{x_{k-1}} + K_k(Z_k - H_k \phi_{k,k-1}\widehat{x_{k-1}})$. Here $K_k Z_k$ represents the state vector of m × n dimensions and n × n dimensions in the measurement matrix of the image. However, for the actual exercise time and the number of video frames recorded by the computer, since the time is very short, the target is difficult to appear very large changes in the adjacent two frames. Therefore, we can set a certain interval time to unit time, and the target maintains a uniform motion state in this unit time, so that a target transfer matrix can be obtained.

3.2. Color probability model

In the image, there is a huge amount of information light column points. From the perspective of human vision, it is essentially to use the color of the world and the changes in brightness to perceive the world. Using this principle, we can use the three primary colors in the computer to get the RGB color space. In computer vision, color is closely related to the scene and the target. This color relationship is different from the geometric relationship. This is because the color is rarely interfered by factors such as viewing angle, and the color features are used. The tracking of the target can be said to have good stability. Therefore, in order to track the target, in order to make the tracking accurate and fast, the selection of color features is very important. Usually when choosing color features, we use the color space of RGB and the color space of HIS.

3.3. Image segmentation algorithm

An important method of segmenting moving objects in a video is to segment the image. When segmenting a moving vehicle, the edge detection operator is used to detect the edge of the vehicle, which improves the accuracy of the segmentation. Based on this algorithm, the maximum segmentation vehicle segmentation strategy is used to enhance the image segmentation. The experimental results show that the algorithm has good performance and effect.

The OTSU algorithm is the maximum inter-class variance method proposed by the principle of discriminant analysis and least squares. This algorithm is a method of threshold segmentation. The algorithm will be described below.

Divide the threshold of a pixel into two parts, the target C_0 and the background C_1. Set the gray value of the image within a certain range $\{0,1,\ldots,l-1\}$, the pixel n_i gray level is i, the pixel is represented as: $N = \sum_{i=0}^{l-1} n_i$, The probability that grayscale i appears is

$$p_i = \frac{n_i}{N}, P_i \geq 0, \sum_{i=0}^{l-1} p_i = 1 \quad (1)$$

Select the threshold for image segmentation to segment the target and background. $C_0 = \{0,1,\ldots,t\}; C_1 = \{t+1, t+2, \ldots, l-1\}$; the probability of occurrence of the target and background is as follows

$$P_0(t) = \sum_{i=1}^{t} p_i, P_1(t) = \sum_{i=t+1}^{l-1} p_i \quad (2)$$

Their average values are as follows:

$$\mu_0(t) = \sum_{i=0}^{t} i p_i / P_0(t), \mu_1(t) = \sum_{i=t+1}^{l-1} i p_i / P_1(t) \quad (3)$$

The gray level of the image is:

$$\mu_T = \sum_{i=0}^{l-1} i p_i \quad (4)$$

Using the formulas 2, 3, and 4 above to calculate the variance between classes, the result is:

$$O_B^2(t) = P_0(t)[\mu_0(t) - \mu_T]^2 + P_1(t)[\mu_1(t) - \mu_T]^2 \quad (5)$$

Then define the variance within the class and calculate:

$$O_w^2(t) = P_0(t)\sum_{i=0}^{t}[(i-\mu_0)^2 * \frac{p_i}{P_0(t)}] + P_1(t)\sum_{i=t+1}^{l-1}[(i-\mu_1)^2 * \frac{p_i}{P_1(t)}] \quad (6)$$

The total variance is defined by the interclass and intraclass variances:

$$O_r^2 = O_h^2 + O_w^2 \quad (7)$$

For second-order statistics, $O_w^2(t)$, O_r^2 is independent, and a judgment is made at this time:

$$\eta(t) = \sigma_B^2/\sigma_T^2 \quad (8)$$

Under the above conditions, both types of t values are optimal thresholds, which is the biggest criterion of $\eta(t)$:

$$max\ \eta(t) = \max\eta(t^*)\ (0 < t < l-1) \quad (9)$$

4. Performance analysis of target tracking algorithm

Among the various visual tracking algorithms currently available, two types of algorithms are most commonly used. One is a motion-based algorithm and the other is a model-based algorithm. In this paper, the second model-based algorithm is used to achieve the target tracking by using the matching function. There are two types of matching templates: The first type of target-based method uses the local features of the target to lock, such as using angles and colors to match. In some environments with more complex environments, the effect is better than the boundary match. The most used method in this way is to use angular features for positioning. Tracking the corner points of the target to reflect the dynamic characteristics of the target is one of the most commonly used methods in this way; the template of the second target area incorporates the features of the target area such as centroid, perimeter, area, and color. Among the selection factors of the feature. Polanna et al. used a rectangular box on the selection tracking based on the characteristics of the target area instead of the single-person boundary. When occlusion occurs, the tracking can be maintained by comparing their centroid speeds.

5. Conclusion

Tracking of moving targets has now become a hot research direction in the field of computer vision. It has been widely used in the current military, transportation, and daily production and life, and has attracted the attention of researchers. In this paper, the traditional dynamic recognition algorithms and patterns are reviewed and elaborated. Based on Marr’s theory and Kalman filter tracking algorithm, the computer vision target tracking algorithm is implemented. This plays a very important role in the city's handling of vehicle traffic problems. In this paper, when the vehicle is segmented, the Robert edge detection operator is used to decompose the features of the target, and then the gray level in the image is used to improve the importance of image edge detection, and the practicality and accuracy of the largest class variance is improved to adapt to the complex environment. This provides a new and stable algorithm for tracking urban road vehicles.

References

[1] Luo Haibo, Xu Lingyun, Hui Bin, et al. Research Status and Prospects of Target Tracking Methods Based on Deep Learning [J]. Infrared and Laser Engineering, 2017(5).

[2] Zheng Liang. Research on Trajectory Tracking of Rhythmic Gymnastics Based on Computer Vision [J]. Modern Electronic Technology, 2017(19): 94-98.

[3] Ding Xiaofeng, Shang Zhenhong, Liu Hui, et al. Multi-template Target Tracking Algorithm Based on Mean Shift [J]. Computer Engineering and Applications, 2017(6).

[4] Ma Xiaohong, Yin Xianglei. Overview of Target Tracking Methods Based on Correlation Filters [J]. Electronic Technology Application, 2018, v.44; No. 480(06):9-13+20.

[5] Improvement of JPEG Algorithm [J]. Wang Lixia, Li Bajin. Computer Development and Application. 2004(11)

[6] Mao Jie. Video Target Tracking Algorithm Based on Wireless Sensor Network [J]. Control Engineering, 2018, 25(1).