

Object Tracking and Behavior Recognition Based on Gray Prediction

Zhang Fang

College of Information, Zhejiang Sci-Tech University, Hangzhou, 310018, China

email: zhang_fang2014@163.com

Keywords: Object Detection; Object Tracking; Gray Prediction Model; Behavior Recognition

Abstract. For existing problems in visual tracking and behavior recognition, we propose a novel method to track and recognize based on gray prediction in this paper. We firstly use the background subtraction method to detect moving target and use cross entropy method to process image binaryzation. After that, the morphology filter is used to eliminate the noise, in this way we extract the template whose size is determined by the contour segmentation. The improved gray prediction employs GM(1,1) model to reduce the prediction scope, such that human body's motion trajectory can be recognized. Through the tracking and recording of human moving trajectory, we can identify moving human behavior, like jumping, tumble and squat. Experimental results prove that the proposed algorithm can recognize jumping, tumble, squat and other common human motion behavior correctly as well as very robust.

Introduction

Recently, analysis and recognition of human action based on computer vision is a very hot research field, which focus on the motion detection of the image sequence containing human figures [1-2], object tracking[3-4] and the understanding and recognition of human movement behavior[5-6], the core research is detecting and tracking human movement from single or multiple video sequences, obtaining human motion data. In this paper, we use crossing difference image, binaryzation image based on cross entropy, processing it with mathematical morphology method, getting human position after connection detection, proposing the improved gray prediction model GM(1,1) to reduce the prediction scope, so that we can judge human motion trajectory and recognize human behavior by tracking and recording human motion trajectory.

Object Detection

The main process of object detection is background subtraction, binaryzation, morphological processing, and connectivity analysis.

Background subtraction

The formula of background image difference algorithm had been listed as follows:

$$D(x, y) = |F(x, y) - BK(x, y)| \quad (1)$$

Where, $D(x,y)$ is the gray value of difference image, $F(x,y)$ is the gray value of the current frame image, $BK(x,y)$ is the gray value of background image.

Binaryzation

In order to extract the moving region, we need an appropriate threshold. In this paper we use the cross entropy to calculate the difference image's threshold, which is used for images' binaryzation. When we use P and Q to represent images before and after binaryzation respectively, the optimal threshold is required which can keep minimum information difference between difference image and binaryzation image. The cross entropy is a concept which measures the difference of two probability distribution P and Q . $P = \{p_1, p_2, \dots, p_N\}$, $Q = \{q_1, q_2, \dots, q_N\}$, the cross entropy between P and Q is:

$$D(P, Q) = \sum_{i=1}^N p_i \ln \frac{p_i}{q_i} \quad (2)$$

Mathematical morphology depositing

There are noises and internal cavity in the image after motion region detection, which will adversely affect the connection region detection. Hence, before the connection analysis, we use the mathematical morphology depositing and filling the hole inside the goal area.

Connection analysis

After morphological operations, image small noise point has been removed, small gap has been connected and the internal cavity has been filled, then our algorithm analysis the connectivity on the preserved images. After connection analysis, some connected region appears. In addition to some real target area, some small connected area will also produce, they are not the real target area, but the set of some connected noise point. In this paper, we calculate the whole pixel value of those produced area, the area will be accepted as the target region, where we extract the human profile from, only if the value is larger than a threshold, thus the adverse effects of the noise region will be eliminated.

Object Tracking based on Optimized GM(1,1) Model

Assuming the original sequence is $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$, one time accumulated sequence is $X^{(1)} = (x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n))$, so we call the gray differential equation[7]

$$x^{(0)}(k) + az^{(1)}(k) = b \quad (3)$$

as the original form of GM(1,1) model. Where

$$x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i), \quad k=1, 2, \dots, n. \quad (4)$$

Optimized background value is

$$\begin{cases} z^{(1)}(k+1) = \frac{1}{12}[5x^{(1)}(k) + 8x^{(1)}(k+1) - x^{(1)}(k+2)] \\ z^{(1)}(n) = \frac{1}{12}[-x^{(1)}(n-2) + 8x^{(1)}(n-1) + 5x^{(1)}(n)] \end{cases} \quad k=1, 2, \dots, n-2. \quad (5)$$

If $\hat{a} = (a, b)^T$ is parameter sequence, and

$$Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}, \quad B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix} \quad (6)$$

Then the least square parameter estimation of GM(1,1) model $x^{(0)}(k) + az^{(1)}(k) = b$ meet $\hat{a} = (B^T B)^{-1} B^T Y$, so we get estimate parameter. The response time formula of whitening equation is

$$\hat{x}(t) = (x^{(1)}(n) - \frac{b}{a})e^{-a(t-n)} + \frac{b}{a}, \quad t=1, 2, \dots, n. \quad (7)$$

The response time formula of gray differential is

$$\hat{x}^{(1)}(k) = (x^{(1)}(n) - \frac{b}{a})e^{-a(k-n)} + \frac{b}{a}, \quad k=1, 2, \dots, n. \quad (8)$$

Predicting value is

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k)\alpha, \quad k=1, 2, \dots, n. \quad (9)$$

The target position in the image is composed of vertical and horizontal coordinates, so we establish ordinate and abscissa of GM(1,1) model with ordinate and abscissa of historical frame respectively while tracking.

Human Behavior Identify

As a hot research field, human behavior understanding and description has been extensively concerned in recent years, which refers to the analysis and recognition of human movement pattern, and describes with natural language. In this paper, we match the testing sequence with the calibrated reference sequence which represents typical behavior. We firstly transform image sequence into a set of static shape model by using behavior recognition method based on template matching technique, and then compare it with the stored behavior specimens in the recognition processing.

Experiment Result and Analysis

In this paper, the experimental sequence was filmed with video camera, and then it was divided into segments, which was cut into images afterwards. Then we will do a series of image processing, including isolating target from the background, tracking and predicting target, judging tumble and analyzing other abnormal behavior. By using the algorithm we proposed, Fig1 was judged as jumping sequence, Fig2 was judged as squat sequence and Fig3 was judged as tumble sequence.

The research of human behavior understanding complies with several basic behavior, e.g. feature extraction and motion representation, behavior recognition and senior behavior and scene understanding. Feature extraction and motion representation, which based on the bottom and middle level processing, such as target detection, classification and tracking, extracts feature of target image from target motion information, behavior recognition matches the motion feature extracted from the input sequence with the reference sequence, to judge what kind of behavior model the current motion is. In this paper, we use crossing difference image, binaryzation image based on cross entropy, processing it with mathematical morphology method, getting human original position after connection detection, eliminating adverse effect of background. By calculating and predicting the body center of gravity, we roughly judge the human body position in the next frame. Then we search the area to find out the human body, judging by gravity change to determine whether human behavior is abnormal or not. The proposed template matching technology has the advantages of low computational complexity and simple realization process.

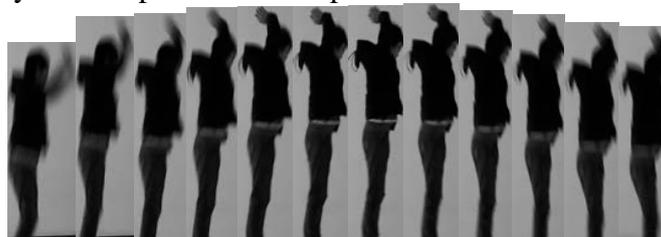


Fig.1 Jumping Sequence



Fig.2 Squat Sequence



Fig.3 Tumble Sequence

Conclusion

In this paper, we detect human motion image which segment the image sequence videoed by DV. We have achieved high correct rate of analysis result in a proper speed. By using predicting method, we reduced searching region, calculating complexity, to meet real time tracking requirement. The template matching technique has the advantages of low computational complexity and easily to implement, but it is sensitive to changes in the movement of time interval and noise.

References

- [1] Tong Lingling, Dai Feng, Zhang Dongming, et al. Encoder combined video moving object detection[J]. Neurocomputing, 2014, 139: 150-162.
- [2] Pan Jing, Pang Yanwei, Zhang Kun, et al. Energy-saving object detection by efficiently rejecting a set of neighboring sub-images[J]. Signal Processing, 2013, 93(8): 2205-2211.
- [3] Lopes Nuno Vieira, Couto Pedro, Jurio Aranzazu, et al. Hierarchical fuzzy logic based approach for object tracking[J]. Knowledge-Based Systems, 2013, 54: 255-268.
- [4] Zhang Xiaoqin, Hu Weiming, Chen Shengyong, et al. Graph-embedding-based learning for robust object tracking[J]. IEEE Transactions on Industrial Electronics, 2014, 61(2):1072-1084.
- [5] Gavrilu D. The visual analysis of human movement: a survey. Computer Vision and Image Understanding[J], 1999, 73(1):82-98.
- [6] Bregler C. Learning and recognizing human dynamics in video sequences[C]. Proc IEEE Conference on Computer Vision and Pattern Recognition, Puerto Rico, 1997, 568-574.
- [7] Dang Yaoguo, Liu Sifeng, Liu Bin. The GM models that $x^{(1)}(n)$ be taken as initial value[J]. Chinese Journal of Management Science, 2005, 13(1): 132-135.