Dynamic Focus Capture Method and Its Application

Shoujin Wang[1, a *], Yang Cao[1, b] and Jingang Shi[1, c]

School of Information & Control engineering, Shenyang Jianzhu University, Shenyang, China

[a]23917240@qq.com, [b]34101755@qq.com, [c]8269086@qq.com

Keywords: Dynamic focus capture; Human-computer interaction; Intelligent monitoring; Motion capture

Abstract. With the development of computer science and technology, motion capture is increasingly applied in many fields such as film production, research analysis, human-computer interaction, intelligent monitoring and so on. The main task of the motion capture is to analyze and process the data from the sensor, and abstract to recognize the action. How to achieve fast and accurate implementation of the action capture has been a hot research in recent years.

Background Sampling Method

The background in this paper should be defined as camera environment that a pending motion captured object locates. For the proper operation of the entire system, it is important to select the background image. Ideally, the background should not be changed with the moving of pending motion-captured object, dividing image with background difference under ideal background is the most accurate. However, in practice, background captured by the camera is sure to impact by the motion of pending motion-captured object [1, 2]. For example: set the room as the background, a person to be the pending captured object. When the person moves within the camera shooting range, it will cause changes in the overall image brightness, the focal length of the shooting and shadows. At present, for better shooting results, most cameras will be automatically adjusted according to shooting images, such as adjusting the sensitivity and lens focal length [3, 4]. This is going against with the immobilization of background shooting, shooting parameters should be fixed as much as possible. The changing of shadow moving object is inevitable, light direction adjustment can alleviate this problem.

In some cases, the background image itself needs to be updated appropriately. For example, during prolonged outdoor shooting, outdoor light will change gradually over time; in this case, the background image should be updated at regular intervals.

Because all detected objects will be withdrawn every time when background being updated. And updating background frequently is relatively troublesome. By using background adaptation algorithm can be programmed to automatically update the background. Specific ideas are as follows: Storage record for a period of time, analyze each frame of the video frame-by-pixel and generate a new background image pixel by pixel [5, 6]. New background pixel should meet the following criteria: 1. The RGB color value of the pixel of recurring or similar appears in the same position for recording, and more than a certain percentage (e.g. related pixels within the ceiling of the room should be kept in the same color). 2. The color of the pixel should be close to the corresponding pixel of last background image.

Image Segmentation and Preprocessing

Image Preprocessing. In order to improve the efficiency and reliability of the system, we must process the preprocessed image. The method used in this paper is gray image processing. In the RGB model, if R = G = B, the color shows a gray color, wherein the value of R = G = B is called the gray values, so each pixel of the grayscale image only requires one byte to store ash value (also known as the intensity value, brightness value), the range of grayscale is 0-255. There are four general methods of color images gray scaling [7, 8].

Average Method. Take the average of the three-component color image as the gray value.

$$f(i,j)=(R(i,j)+G(i,j)+B(i,j))/3 \quad (1)$$

If not specified, this paper adopts the third method for gradation processing which is the average value of grayscale image method.

After the acquiring the grayscale image, the processing of the images can be converted to the corresponding data structures for image processing, which consists of a two-dimensional matrix data structure to implement, store the grayscale value of each pixel of the gray scale. The process of preprocessed image that being described later is based on this data structure for processing in actual.

Image Segmentation

After gray pretreatment of background and pending FIG, we will use the background subtraction method to segment the target object be tested. Background subtraction is also known as background subtraction. The principle of background subtraction is to different the current frame and the background image to obtain the target area, this method can identify and extract the moving target better compare with frame difference method, is the most common used method in motion segmentation . But it needs to build a background image which must be free of moving targets, and should be able to update frequently to adapt the current changes in the background.

If without taking into account the influence of noise n (x, y, t), the video frame image I (x, y, t) can be regarded as composed of background image b (x, y, t) and moving target m (x, y, t):

$$I(x, y, t)=b(x, y, t)+m(x, y, t) \quad (2)$$

From (2-5) can be concluded, the motion target m(x,y,t):

$$m(x, y, t)=I(x, y, t)-b(x, y, t) \quad (3)$$

In practice, due to the influence of noise, the formula (2-3) can not get a real moving target, but the difference image d (x, y, t) that is consisted of noise and movement of the target area, namely:

$$d(x, y, t)=I(x, y, t)-b(x, y, t)-m(x, y, t) \quad (4)$$

It needs one judge principle to have further processing to get moving targets, the most commonly used method is threshold segmentation method:

$$m(x,y,t)=\begin{cases} I(x,y,t) & d(x,y,t)>T \\ 0 & d(x,y,t)\le T \end{cases} \quad (5)$$

T is a threshold. It takes about 50 in the actual situation is appropriated.

Moving Target m (x, y, t) will be stored in the two-dimensional matrix sequence, will be called for focus analysis.

The Acquisition of Dynamic Focus

Image Filtering. The target image that being image segmentation can be further screened to capture the target body more accurately. Take identifying hand movements for an example, we should focus on the identification of color pixels as the hand target, omitted factors sleeves, clothes and other interference; if only detect to the target object in an area, we should limit the effective target image range. Filter specific steps are as follows:

Regional Filtering

$$f(x,y,t)=\begin{cases} m(x,y,t) & m(x,y,t)\in F \\ 0 & m(x,y,t)\notin F \end{cases} \quad (6)$$

Gray Filtering

$$f(x,y,t)=\begin{cases} m(x,y,t) & m(x,y,t)\in G \\ 0 & m(x,y,t)\notin G \end{cases} \quad (7)$$

f (x, y, t) represents the filtered image, m (x, y, t) represents the target image motion. F for the screening area, G target gradation ranges.

This process can be superimposed image filters used in accordance with the actual situation, in order to ensure the accuracy of capture, and reduce system overhead.

Binary Image. Binary image refers to each pixel of the image on the only two possible values or gradation state, often in black and white, B & W, monochrome image represents binary image. Binary image all pixels 0 and 1 can only take from these two values. Its representation is shown in Fig. 1.

In this article, the binary image uses a two-dimensional matrix consisting of 0's and 1's representation. The two values should be corresponded to the closed and open, closed means the pixel in the background, open means the pixel is to be detected.

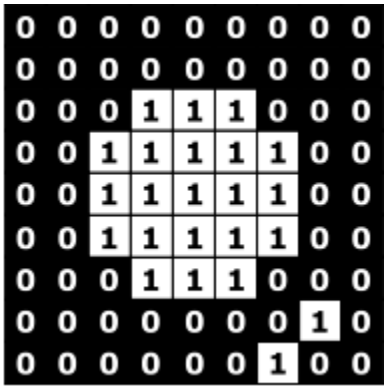

Figure 1. Internal representation of binary image

In the dynamic capture technology, the moving target image that after filtered will be converted to a binary image, stored in the form of two-dimensional matrix. For each element of the matrix inside, the foreground is represented with to "1", the background is "0".

Centroid Algorithm. Centroid means center of a substance mass, is considered a hypothetical point of this mass is concentrated in the physical system. Similarly, the concept of the center of mass can be introduced into the binary image processing. If the binary image represents a collection of all the G "1" of the pixel (p) as a whole, formed part of the geometric center can also find out, it may be referred to the geometric center of the binary image " centroid ", denoted by point C (x, y), formed part of the area is regarded as a binary image," quality ", referred to as M (unit: pixels). By centroid algorithm, we can measure out a binary image centroid point C and the quality M.

Binary image centroid calculation can be used in the following equation:

$$x = \frac{\sum_{i}^{m}\sum_{j}^{n} iI(i,j)}{\sum_{i}^{m}\sum_{j}^{n} I(i,j)} \quad y = \frac{\sum_{i}^{m}\sum_{j}^{n} jI(i,j)}{\sum_{i}^{m}\sum_{j}^{n} I(i,j)},$$

$$I(i,j) = \begin{cases} 1, & (i,j) \hat{I} \ Target \\ 0, & (i,j) \hat{I} \ Background \end{cases} \tag{8}$$

I (i, j) is a binary image. (I, j) are the coordinates of pixels in the image, m, n is the number of rows and columns of the image.

From the formula (11), the calculation time centroid point C of the pixel is proportional to the degree of complexity of a binary image, when dealing with large binary image less efficient.

Thus, the focus has to be to capture a moving object is calculated, C (x, y) represents the position of the focus, M represents a quality focus, the image area the size of the actual significance of the detected moving object.

In this system, a data structure is defined for focus is as follows:

```
Focus Point
 {
        Px     (Integer, horizontal axis of the focus)
        Py     (Integer, vertical axis of the focus)
        M      (Integer, mass of the focus)
  }
```

In summary, the moving target image is compressed into a one-dimensional representation and stored in FocusPoint structure. Passed by the camera image is converted to a continuous focus sequence and stored in FocusPoint structure array.

The Acquisition of Multifocal

Understood by the section, for the image containing a moving target focus acquisition process can be divided into image filtering, image binarization, the focus of a three-step calculation. Among them, for the same pending image, we will have a different focus if changing the image filtering conditions. Different focus can be obtained from a same image without interfering with each other, which created the possibility to simultaneously track multiple targets.

In modern computers, the conception of parallel and concurrent are no longer strange to us. Parallel in the operating system refers to a set of programs by independent asynchronous execution speed is not equal on the overlap time (same time occurrence) to distinguish concurrency. Concurrency means in the same period, two or more program execution time is superimposed on the (macroscopically the same time, the micro is still the order of execution) . The system uses multi-threading technology to calculate multiple dynamic focuses in the same pending image and for the user is "simultaneously" in macro, which plays an important role in motion determination and application systems.

Motion Detection

The essence of the motion detection in this article is the conclusion from the obtained focus array analysis. It is benefited from the simplicity of focus data structure to the original picture, the focus of analysis will be quite easy to understand. The following actions will take detecting "from left to right" as an example to explain the operation of the detection system works:

List the most recently calculated focus, denoted by fp (x), (0 <x <t). Among them, t is the number of elements set the focus, x is sorted by time (in the past -> present in chronological order) the focus position with respect to the sequence.

For each focus fp (x), (0 <x <t), whenever fp (x) .M> m, referred to as "reliable", otherwise referred to as "unreliable." Statistics were "reliable" and "unreliable" number. The significance of the condition can be recognizable object large enough to be included valid.

For each focus fp (x), (x> 1), each time fp (x) .Px-fp (x-1) .Px> k, referred to as "reliable", otherwise referred to as "unreliable.” Statistics were "reliable" and "unreliable" number. The meaning of the condition is judging the significance of the focus moving trend.

Synthesize each condition, obtain the conclusion of whether "from left to right" movement of the target has occurred.

Application Operation Process

System operation flow is shown in Fig. 2 Firstly, we have a background sample for background subtraction for image segmentation. After sampling the background, keep the camera stationary, the system should be constantly revised in the background according to the actual operating system

environment. Then the camera continuously input images, each frame of the image before processing should be pretreated to enhance efficiency after processing. After calculating the target under test and store data related to focus, to focus data series as a basis for motion detection. Finally, if a predetermined operation is detected, then we trigger the corresponding action.

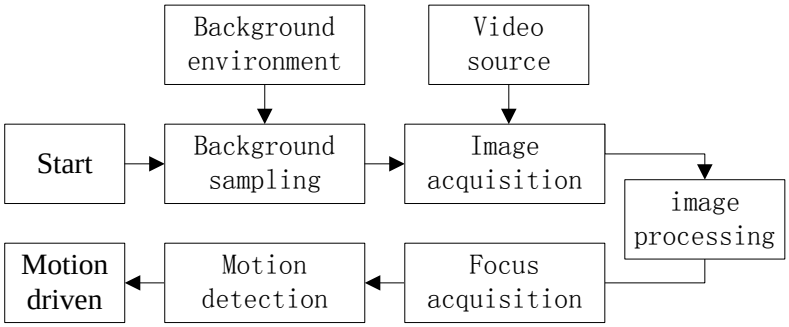

Figure 2. System operation flow chart

Conclusions

This paper describes the dynamic focus capture method and explains from background sampling method, image segmentation and pre-processing, dynamic focus acquisition and motion detection as four aspects. Firstly, it introduces the background sampling methods. Background sampling has important basic effect for the background operation of the system, the accurate background image can directly affect the success rate of subsequent motion detection; then it introduces the process of image segmentation and preprocessing, the process is used for simplified image structure, improve operational efficiency; the concept of dynamic focus is the emphasis of this article as well as the core of this system; motion detection method is based on a dynamic focus, discusses the advantages and disadvantages of this method.

References

[1] Li Zhaokun. Human computer interaction system based on motion recognition technology [D]. Suzhou: Soochow University, 2014:12~17

[2] Min Lin, Bin Li, Qiao-Hong Liu.Identification of Eye Movements from Non-front--al Face Images for Eye-controlled Systems[J]International Journal of Automation & C--omputing,2014,issue No.5: 29~30

[3] Sha youth. Video based traffic statistics algorithm research [D]. Wuhan: Huazhong University of Science and Technology, 2013:44~45

[4] Wei Xiaopeng, Liu Rui, Zhang Qiang. Review of techniques for motion capture data processing [J].Computer Aided Drafting, Design and Manufacturing,2012, issue No.1: 5~7

[5] Jiang Mingyan, the video image processing in traffic statistics algorithm [D]. Chengdu: University of Electronic Science and technology, 2013:37~38

[6] J. Xiao, Y.T. Zhuang, F. Wu. Feature Visualization and Interactive Segmentation of 3D Human Motion [J].Journal of Software, 2008, 19(8): 1995-2003.

[7] Y.D Yang, L.L. Wang, A.M. Hao. Motion String: A Motion Capture Data Representation for Behavior Segmentation [J]. Journal of Computer Research and Development, 2008, 45(3): 527-534.

[8] Pan Hong, Xiao Jun, Wu Fei, Guo Tongqiang. 3D Human Motion Retrieval Based on Key-Frames [J].Journal of Computer-Aided Design & Computer Graphics, 2009, 21(2): 214-222.