

Classification of Image Gesture Segmentation Methods

Lin Ling, Dechao Wan

Faculty of Artificial Intelligence and Big Data, Yibin University, Yibin 644000, China

Abstract

Gesture segmentation is a key step in image gesture recognition. This paper summarizes and classifies gesture segmentation technology in detail, which is of great significance to the implementation of gesture recognition technology.

Keywords

Image; gesture segmentation; gesture recognition.

1. Introduction

In addition to the palm part, the images captured by Kinect also contain some other irrelevant backgrounds. Therefore, in order to improve the accuracy of subsequent gesture recognition, it is necessary to separate the palm part from the image, that is, gesture segmentation. In the multi-feature fusion gesture recognition method proposed by Weng et al^[1], the Bayesian skin color model is used for gesture segmentation. Scholars have explored a variety of gesture segmentation methods, which can be roughly divided into the following three categories: segmentation by skin color^[2], segmentation by motion model^[3], and segmentation by threshold^[4].

2. Skin color model method

The skin color-based gesture segmentation method utilizes the skin color feature of the hand to directly extract the pixels similar to the skin color in the image as the gesture part. This method can be combined with hand skin color, texture, multi-fusion information, etc. for detection. Among them, the method based on hand skin color is widely used in gesture recognition because it only pays attention to color information and does not consider changes in hand posture, spatial position, etc., face recognition, etc. The main steps of skin color segmentation are as follows: the selection of color space, the establishment of model, and the selection of pixel points. Color spaces mainly include the following three.

The RGB color space is a basic color space model, which relies on the relationship between the three components of R, G, and B to determine the skin color pixels. Using this color space is simple and intuitive, but it is easily affected by lighting. The HSV color space, also known as the hue saturation model, mainly relies on the H and S components to determine the skin color pixels, and the V component, which has the greatest impact on illumination changes, is not the main factor. Therefore, based on the HSV color space, there is a certain robustness in simple skin color detection. sex. YCRCB is YUV, Y is the color luminance component, and CR and CB are the density offset components of red and blue, respectively. CR and CB chromaticity can maintain good robustness to ordinary lighting, so the skin color detection algorithm based on YCRCB color space is the most used one, and is often used to process digital photography systems or continuous films. Using the gesture segmentation method based on the skin color model, it is necessary to exclude other skin color parts other than gestures in the output image, such as areas such as arms and faces, otherwise the accuracy of subsequent gesture recognition may be affected.

3. Motion model method

The gesture segmentation method based on the motion model requires that during the execution of dynamic gestures, the hand is in a moving state, while other backgrounds and environments are in a static state. At this time, gesture segmentation can be performed using the motion information of the hand. The gesture segmentation of the common motion model method includes the following three.

The inter-frame difference method uses the adjacent frames in the acquired image sequence to perform a difference operation, and removes the background to extract the gesture part according to a predetermined threshold. This method is simple to implement and has fast detection speed, but requires stable hand movement speed. Too fast or too slow will lead to incomplete hand regions segmented by differential operations, and even cannot effectively segment gestures.

After the background difference method is established, the difference between the collected image and the background model is used to differentiate each frame of the collected image, so as to realize the positioning of the gesture. For the establishment of the background model, there are three commonly used methods at present, namely Gaussian background modeling, median method and mean method. The choice of modeling method is also an important factor affecting the effect of background difference. This method requires that the background cannot be changed greatly. Once the background of the image changes to a certain extent, the gesture cannot be completely segmented by the difference operation of the original background model. The optical flow method is to compare and analyze the acquired image sequence between two consecutive frames, analyze the correlation between the two frames of images and the changes that occur between each pixel, and obtain the motion information in the image through calculation. When using this method for analysis, each pixel in the image has a motion vector tangent to the motion trajectory, which corresponds to the point projection on the target object one-to-one, and then gestures are detected based on the velocity vector. The optical flow method can also accurately detect gestures in complex environmental backgrounds, but the algorithm is complex, the amount of calculation is large, and the brightness is required to be constant.

4. Threshold segmentation method

The threshold segmentation method is a classic method in gesture segmentation. Its core idea is to set a threshold and compare all the pixels in the grayscale image, and separate the gesture part from the background environment in the image according to the comparison result. Calculated as follows:

$$g(x, y) = \begin{cases} 255 & f(x, y) \geq T \\ 0 & f(x, y) \leq T \end{cases}$$

where T is the segmentation threshold, g is the gray value function after threshold segmentation, and f is the gray value of each pixel in the unprocessed image. Through the analysis of the formula, it can be known that different threshold settings have a direct impact on the final image segmentation effect. There are four commonly used threshold segmentation methods, namely the region growing algorithm, the maximum between-class variance (OTSU) method, the histogram method, and the minimum mean square error method. The threshold segmentation method is the most widely used technique in gesture segmentation because it is easy to implement, requires little computation, and has strong applicability.

After the Kinect sensor appeared, the threshold segmentation method was more favored by researchers. The main function of Kinect is to collect images. The collected depth images have

different color depths, and the different color depths represent different distances between objects and the camera. The closer the object within the range is to the camera; the whiter the relevant area in the image, that is, the larger the gray value, the farther the object is from the camera. At the same time, the images collected by Kinect will not be affected by changes in lighting, background and other factors. In practical application scenarios, the hand can be segmented from the complex background as long as the hand is kept at a certain distance from the rest of the body, and the segmentation effect is very good.

5. Summary

In this paper, the gesture segmentation methods are classified and explained in detail, which provides important theoretical support for gesture recognition technology.

References

- [1] Weng C, Yang L, Zhang M, et al. Robust Hand Posture Recognition Integrating Multi-cue Hand Tracking[C]. Proceedings of the International Conference on Entertainment for Education Digital Techniques & Systems, 2010.
- [2] Wang Long, Liu Hui, Wang Bin, et al. Gesture recognition method combining skin color model and convolutional neural network [J]. Computer Engineering and Applications, 2017, 53(006): 209-214.
- [3] Xu Bing, Niu Yanxiong, Lv Jianming. Object Detection and Segmentation Algorithms in Complex Dynamic Scenes [J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(2): 310-317.
- [4] Yang Xiuguo. Research and analysis of image threshold segmentation method [D]. East China Normal University, 2009.