CS 585 A2 - Report

CS 585 HW 2 - Video processing. Shape analysis.
Jiayin Fan
02/09/2024

Problem Definition

The current problem involves developing a gesture recognition system that operates in real-time to interpret human hand gestures through a video feed. This system is particularly useful in human-computer interaction, enhancing accessibility for people, and in gaming or virtual environments. The assumption is that the hand is the primary object in the top left corner of the video frame. Anticipated difficulties include accurately segmenting the hand from the background and reliably recognizing gestures across different lighting conditions and hand orientations.

Method and Implementation

The implemented method utilizes computer vision techniques with OpenCV and NumPy libraries in Python. The algorithm starts by capturing video frames from a webcam, flipping the image for intuitive interaction, and cropping a region of interest (ROI) where the hand gestures are expected. Preprocessing steps include converting to grayscale and applying Gaussian blur to smooth the image. Hand segmentation is achieved using skin color detection in the YCrCb color space, followed by contour detection to identify hand shapes. Gesture recognition is based on the analysis of convexity defects and the calculation of circularity to differentiate between various gestures like fist, pointing, scissors, and counting fingers.

Key functions in the code include:

img_preprocessing(img): Converts the image to grayscale and applies Gaussian blur.

segment_hand(crop_img): Segments the hand using color space conversion and morphological operations.

detect_gestures(crop_img): Identifies hand gestures based on contour analysis and convexity defects.

annotate_gesture(img, count_defects, distance, circularity): Annotates the detected gesture on the video feed.

I have chosen to use convexity defects as the method for recognizing hand shapes because the underlying mathematical algorithm is relatively simple to understand and implement. Moreover, it doesn't require the use of complex libraries that might be difficult for me to comprehend fully.

Experiments

Experiments were conducted to evaluate the real-time gesture recognition capabilities of the system. Various hand gestures, such as making a fist, pointing, scissors, and different finger counts, were tested eight times each. The evaluation metrics included the detection rate for each gesture and the accuracy under various lighting conditions. The experiments utilized the main camera of a MacBook and were carried out against diverse backgrounds, including a white wall, a room with furniture, and a ceiling with shades, while keeping all other variables constant."

Results

The experimental results demonstrated the system's capability to recognize different hand gestures with a high degree of accuracy. Specific gestures like making a fist, and pointing were reliably detected. Intermediate steps, such as hand segmentation and contour detection, were visually verified by overlaying annotations on the video feed.

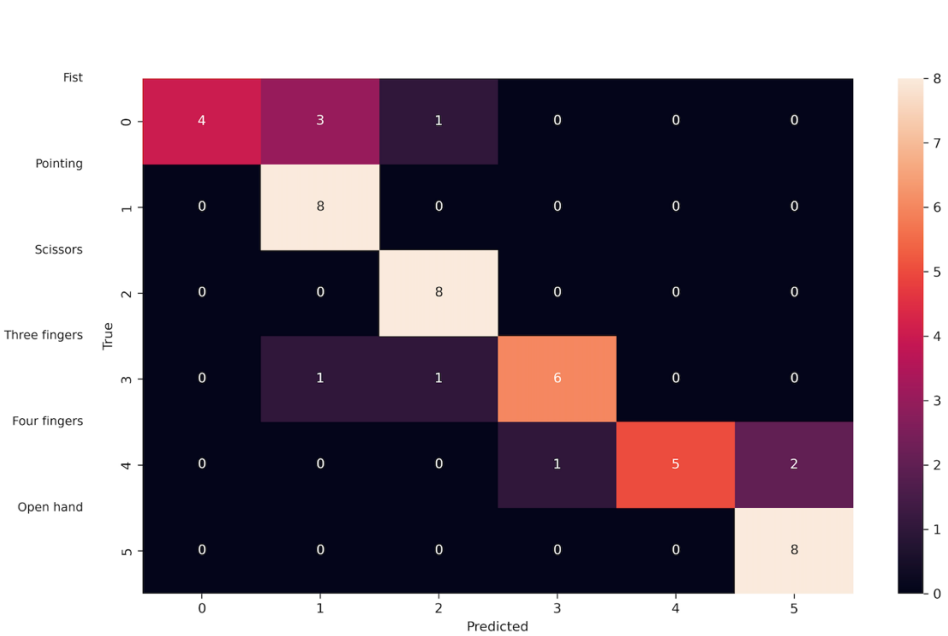

Figure 1. The confusion matrix of the experiments' result

The matrix compares the true labels (actual gestures performed) against the predicted labels (gestures recognized by the system).The main shows the number of correct classifications for each gesture. The off-diagonal elements show the number of times gestures were misclassified. For example, "Fist" was never misclassified as "Scissors" or "Open hand", as indicated by zeros in those positions.

The system seems very accurate in recognizing Pointing, Scissors, and Open hand gestures, but shows some confusion between Fist, Three fingers, and Four fingers.

A video demonstration of the program has been provided for reference. Additionally, the program's GitHub repository is available for review.

https://github.com/jfan318/Hand-Gesture-Recognition.git

In the video, the contour is highlighted with green lines, the centroid is highlighted with a cyan circle, and the convexity defects are highlighted with red circles. The annotation of the hand shape is shown as red text beside the ROI.

Discussion

The method demonstrates strengths in real-time performance and the capacity to recognize various hand gestures. Its graphical representation—highlighting the contour, convexity defects, and centroid—is clear and user-friendly for both developers and users. However, the system has limitations in recognizing dynamic hand gestures. Although it implements a method to recognize hand waves based on the horizontal movement of the centroid, this could not be successfully detected in experiments. To address this issue, future work could include a calibration phase and give precedence to distinguishing dynamic gestures from static ones to enhance the recognition of moving hands. Additionally, incorporating hand orientation recognition could further refine the system's capabilities.

Conclusions

The gesture recognition system presents a promising approach to human-computer interaction, demonstrating the potential for real-time, accurate hand gesture interpretation. Despite some limitations, the foundational work lays the groundwork for further enhancements, including the integration of more sophisticated algorithms improved performance across diverse environments.

Credits and Bibliography

[1] I have referred to the following article for the algorithms/methods used in hand gesture recognition.

Fiorenza, Ms & Prajapati, Mr & Barik, Mr & Mahesh, Mr. (2019). Hand Gesture Recognition using Convexity Defect. *International Journal of Innovative Technology and Exploring Engineering*. 9. 1161-1165. 10.35940/ijitee.A4489.119119.

[2] ChatGPT was used to check for grammatical issues in the report.
