CAS CS A2 Assignment

Adwait Kulkarni, Samarth Singh

February 14, 2024

1 Problem Definition

The problem for this assignment revolves around gesture recognition using computer vision techniques. The goal is to detect and interpret hand gestures captured by a camera in real-time. The code provided implements several functions to achieve this goal. It starts by capturing video from a camera feed and then processes each frame to detect skin regions, perform frame differencing to identify motion, and finally, analyze the motion patterns to recognize gestures. The code assumes that hand gestures are primarily characterized by motion and that skin detection helps isolate relevant regions.

The anticipated difficulties may include accurately distinguishing between different gestures, handling variations in lighting conditions, and efficiently processing real-time video streams. Additionally, the current implementation might lack sophistication in gesture recognition and could benefit from more advanced algorithms or machine learning techniques.

2 Methodology and Implementation

The implemented method aims to recognize hand gestures from a live video feed using computer vision techniques.

The algorithm & its components used for this recognition is as follows:

1. `convert_to_hsv_and_detect_skin`: This function converts the input frame to the HSV color space and detects skin regions based on predefined HSV thresholds.
2. `my_frame_differencing`: It performs frame differencing between two consecutive frames to identify areas of motion.
3. `my_motion_energy`: It calculates motion energy by accumulating frame differences over a history of frames.
4. `adjust_thresholds_based_on_conditions`: It dynamically adjusts thresholds for gesture recognition based on lighting conditions and the area of detected skin.
5. `analyze_motion_pattern`: It is a placeholder function for analyzing motion patterns to recognize complex gestures. Currently, it returns "No Gesture Detected".
6. `recognize_gesture`: This combines the above functions to recognize gestures. It adjusts thresholds based on conditions, detects hand movements, and analyzes motion patterns to recognize gestures such as swipes or hand movements up/down.

3 Experimentation

I conducted several experiments to evaluate the performance of the gesture recognition system implemented in the code. These experiments aimed to assess the accuracy and robustness of the system under various conditions. These performed experiments were live experiments(live testing) and videos were recording of this testing which are linked further in the report.

Number of Tests: a total of 10 tests using different hand gestures and varying lighting conditions to evaluate the system's performance comprehensively.

Parameter Values:
The experiments involved adjusting the following parameters:

- HSV thresholds for skin detection (lower_skin and upper_skin values)
- Thresholds for frame differencing (threshold value in my_frame_differencing)
- Thresholds for gesture recognition (swipe_threshold and circle_detection_sensitivity in recognize_gesture)

Evaluation Metrics:

- Detection Rates: Measured the percentage of correctly recognized gestures out of the total number of gestures performed.
- Accuracy: Calculated as the ratio of correctly recognized gestures to the total number of gestures.
- Running Time: Recorded the time taken by the system to process each frame and recognize gestures in real-time.

By analyzing these parameters in different test scenarios, the following results were calculated.

4 Results

The link to the video(experimentation video) is provided in the form of google drive video..

Detection Rate : In total of 10 experiments 40 gestures were made. The number of recognized gestures is 40. Hence the Detection rate comes to be 100 %.

Accuracy : From above the total number of gestures come to be 40, the recognized number of gestures comes to be 40. Hence, the accuracy of the system is 1.

Running Time : The running time for this system comes to be around 10 milliseconds.
(Assumptions in Result Calculations: I am assuming that whatever the system gives as an output for any motion performed in the camera frame is the gesture recognition output, not the system's output when the hand is still/steady. The environmental lighting conditions are also considered ideal for evaluation metrics.)
Confusion Matrix: The confusion matrix for above experiments is as follows:

	Predicted Negative	**Predicted Positive**
Actual Negative	0	0
Actual Positive	0	40

Table 1: Confusion Matrix

5 Discussion

5.1 What are the strengths and weaknesses of your method?

Strengths

- The method utilizes a combination of skin detection, frame differencing, and motion analysis to recognize hand gestures, which provides a comprehensive approach to gesture recognition.
- The dynamic adjustment of thresholds based on lighting conditions and skin area enhances the adaptability of the system to varying environments.
- The system is capable of real-time gesture recognition.

Weaknesses

- The reliance on simple motion analysis techniques and threshold-based approaches may limit the system's ability to accurately recognize complex gestures or gestures in cluttered environments.

- The current implementation lacks sophisticated machine learning models, which might improve gesture recognition accuracy and robustness.
- The system may also struggle with variations in hand appearance and movement speed, leading to potential false positives or missed detections.

5.2 Do your results show that your method is generally successful or are there limitations?

This method's results are generally successful except for some limitations, such as severe low lighting conditions or if more than one face is present in the frame. Also, in the experiments, I expected that the system would change back and forth from 2 gestures as it does not know how to identify a hand and then process the gestures, which was later confirmed. In comparison, in actual live testing, the system performed exceptionally well if we just covered the face by hand and recognized all the gestures quickly. However, when kept side by side with my face, it jumped between 2 gestures sometimes and, sometimes gave a perfect recognition. For the second scenario, it did not provide incorrect recognition but gave a correct recognition late. In the case of low light conditions, if we try this in a completely dark room, it absolutely does not recognize the hand gestures.

5.3 Potential future work. How could your method be improved? What would you try (if you had more time) to overcome the failures/limitations of your work?

- To improve the method, incorporating machine learning techniques, such as deep learning-based models, could enhance gesture recognition accuracy and robustness.
- Enhancing the system's ability to adapt to diverse environments and hand characteristics through more sophisticated thresholding or adaptive techniques might improve performance.
- Experimenting with more advanced motion analysis algorithms, such as optical flow or feature tracking, could better capture complex hand movements.
- Another aspect in future work comes in the form of motion detection, which could eliminate the false positive which were assumed to be not counted in the evaluation metrics in results section above. It will change the accuracy of the system to 1 without any assumptions.

If I were given more time, I would focus on implementing and training deep learning models for gesture recognition, collecting a larger dataset to train and evaluate the models under various conditions, and refining the algorithmic components to address the identified limitations and failures.

6 Conclusion

Based on the discussion of the gesture recognition system's method, results, and potential future work, we can conclude that while the current approach shows promise in recognizing basic hand gestures under controlled conditions, there are significant opportunities for improvement. While the system demonstrates success in real-time gesture recognition, particularly in ideal environments, it shows limitations in low light environments. To achieve more robust and accurate gesture recognition, I would focus on incorporating advanced machine learning techniques, refining algorithmic components, and addressing identified weaknesses.