Hand Segmentation Method Using Multi-frame Images

ECE5470 : Final Project by **Jilong Wu** (jw859), **Xi He** (xh243)
College of Engineering, Cornell University - Fall 2014

Overview

This project is implementing a hand segmentation algorithm that can be used for many applications such as hand recognition. Instead of using a naïve hand segmentation only based on skin color detection, our proposed segmentation method is using multiple frames to assist in obtaining a more accurate hand region. This method will be implemented on a sequence of images (videos). The proposed algorithm is first trained for a set of training videos whose background noise are classified into 4 different levels. In the last step, a set of testing images will be evaluated using the set of optimal parameters.

Algorithm Impplementation

Figure 1 Original after HSV thresholding

Figure 2 Morphological filter applied

Figure 3 Frame Chnage

Figure 4 AND Operation

1. The vision algorithm is applied on each frame of the input. For each frame, it first converts it from RGB to HSV image and apply the HSV threshold on the frame(Figure 1)
2. Then the morphological filtering is appliedto remove small objects from the foreground.(Figure 2)
3. Since the algorithm assumes static background with only moderate hand movement in the frames. it detects the change by comparing each pixel of the frame with the same pixels in the previous frame. If the change is bigger than the threshold value defined, it would mark this pixel as a change pixel, specifically 255 if bigger than threshold 0 if not.Then the AND operation on HSV filtered image and frame changes is applied. (Figure 3 and Figure 4).
4. We use moment of the ANDed frame to calculate the approximate location of center of mass of the hand.
5. Using the center of mass we apply our hand search algorithm on the ANDed frame to locate the hand region of interest. (Figure 5)
6. Lastly, we apply the region of interest on the HSV filtered image to get the final segmentation (Figure 6).

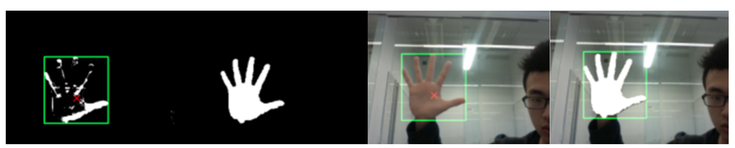

Figure 5 Region of interest through search

Figure 6Final hand segmentation

Figure 7 Region of interest +original

Figure 8 Hand segmentation+original

Program Flowcharts

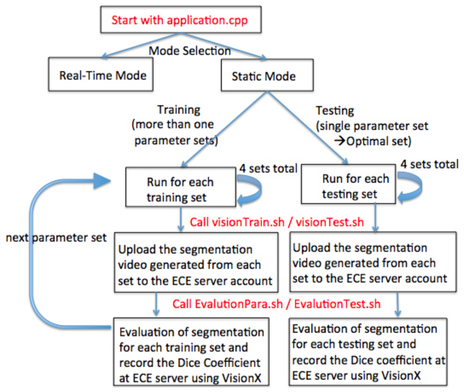

Figure 9 Flowchart of the proposed method

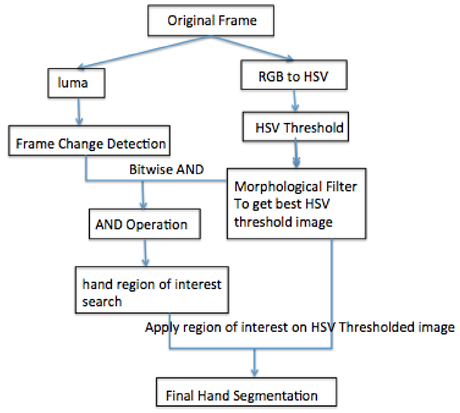

Figure 10 Flowchart of the vision algorithm

Figure 9 shows the flowchar of the whole process of this project from algorithm implementation to evaluation process. This procedure is used for both training process and testing process.

Experiment Design

In general, the experiment has two phases:
1. The first phase is to use some training video sets to train the proposed method with different sets of parameters. Then we plot the performance of each set of parameters and pick the set of parameter which produces the best performance.
2. In the second phase, we have another testing video sets. Using the optimal parameters we obtained in the training phase, we obtain the segmentation performance of the testing sets and evaluate our proposed method.

Figure 11 A sample frame of segmentation results on training sets with original parameters(noise level 1 to 4 from left to right)

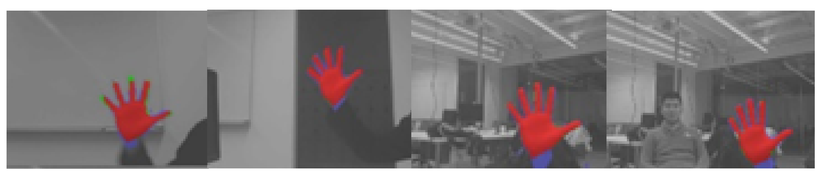

Figure 12Final image comparison results using vrdiff for testing sets with optimized parameters (noise level 1 to 4 from left to right)

Performance Results

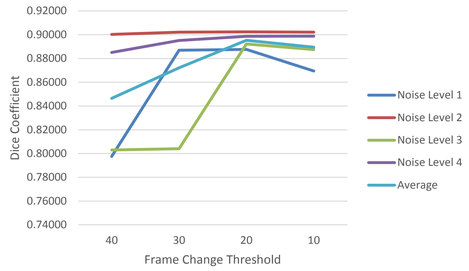

Figure 13 Training results on frame change threshold values with optimal threshold value plicked: thresh=20, Dice Coefficient =89.52%

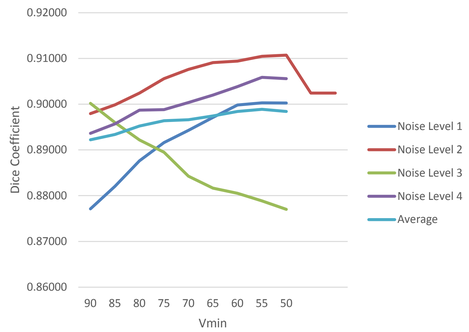

Figure 14 Training results onV values with optimal V values picked: V=55-255 Dice Coefficient = 89.84%

Parameter	Noise1	Noise2	Noise3	Noise4	avg
H:0-20 S:45-150 V:55-255 Threshold:20	0.88895	0.89576	0.86669	0.85853	0.87748

Figure 15 Results of Optimized segmentation method on testing sets. The first four from legyt to right are the accruracy for diffewrent noise levels

Conclusion

In this project, we have implemented a hand segmentation method using multi-frame images. After careful research and studying, we successfully implemented the proposed method on training examples on various noise backgrounds under the assumption that with static background with certain illumination condition and moderate hand movement and no similar color objects around the hand.

Using the optimal parameters obtained from training process, the final result on training examples can be as high as 90.85% and for testing example it can be as high as 87.75%(Dice Coefficient). It turns out the algorithm works well and consistently under the similar condition as the training sets.