ECE-5554 Computer Vision: Problem Set 3

Murat Ambarkutuk
murata@vt.edu

Mechanical Engineering Department,
Virginia Polytechnic Institute and State University

November 27, 2015

Answer Sheet

1 Programming Problem (Activity Recognition)

1.1 Motion History Images (MHI)

To successfully recognize activitis on image sequences, first background subtraction should be performed. In the data set given the camera is fixed and the scene is stationary, so one can assume that pixels have different values at different time steps depicts a new object or a person. That assumption enables one to extract foreground pixels from that of background by simply subtracting consecutive frames.

$$D(x, y, t) = I(x, y, t) - I(x, y, t-1)$$

However intuitive, just subtracting may introduce some noise as foreground due to the nature of imaging systems. Thus, a threshold was employed on $D(x, y, t)$ to disregard minor changes resulting from the imaging system.

$$D(x, y, t) = (I(x, y, t) - I(x, y, t-1)) > \tau$$

where $D(x, y, t)$ is a binary image sequence depicting the moving pixels in the consecutive frames. The threshold value τ is chosen by testing different values and it is 1500 in the following examples images.

After acquiring $D(x, y, t)$, motion history image, $H_t(x, y, t)$, can be constructed the formulation stated below.

$$H_t(x, y, t) = \begin{cases} \tau, & \text{if } D(x, y, t) = 1 \\ max(0, H_t(x, y, t-1)), & \text{otherwise} \end{cases}$$

where τ can be referred as "memory term". As τ increases, the action will be carried to further future. Conversely, as it decreases, the effects of the action will fade away faster from the motion history images. Since each portion of the actions matter same for the recognition, τ is set the total number of frames in the sequence. Three example motion history images is listed below. (Figures-1, 2, 3) Also please refer "allMHIs.mat" file to see full results.

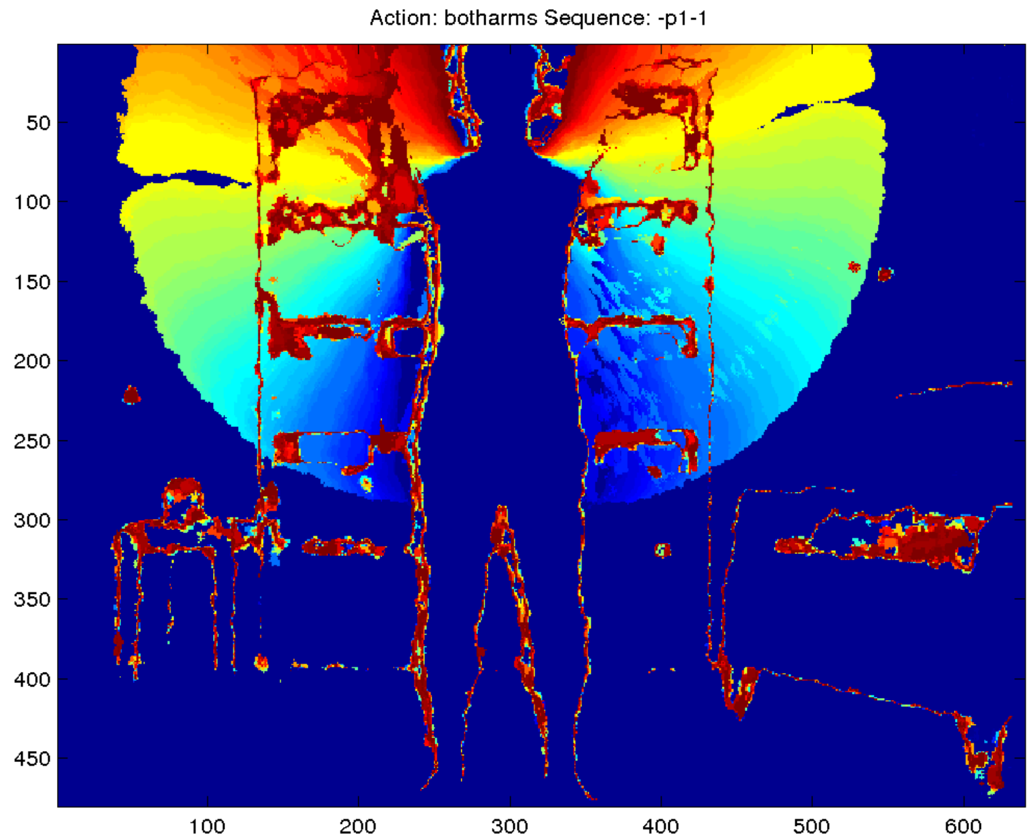

Figure 1: Motion History Image:1, Image Sequence: Both Arms Up, Sequence: 1

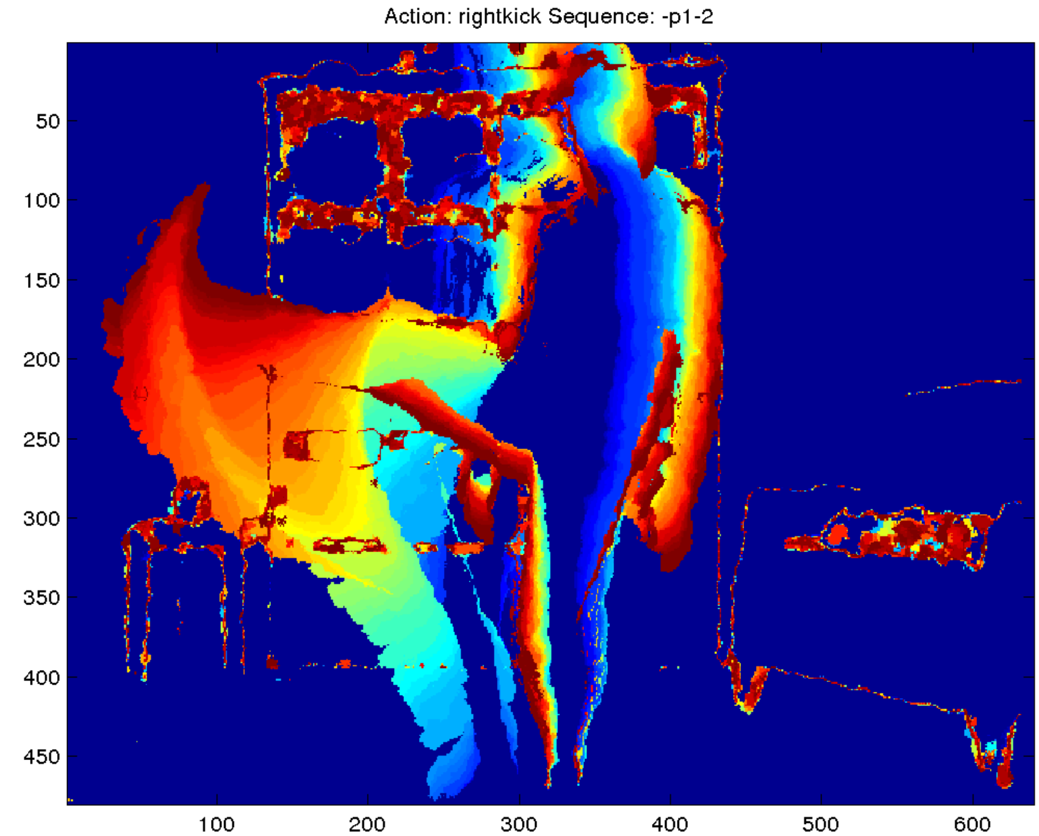

Figure 2: Motion History Image:2, Image Sequence: Right Kick, Sequence: 2

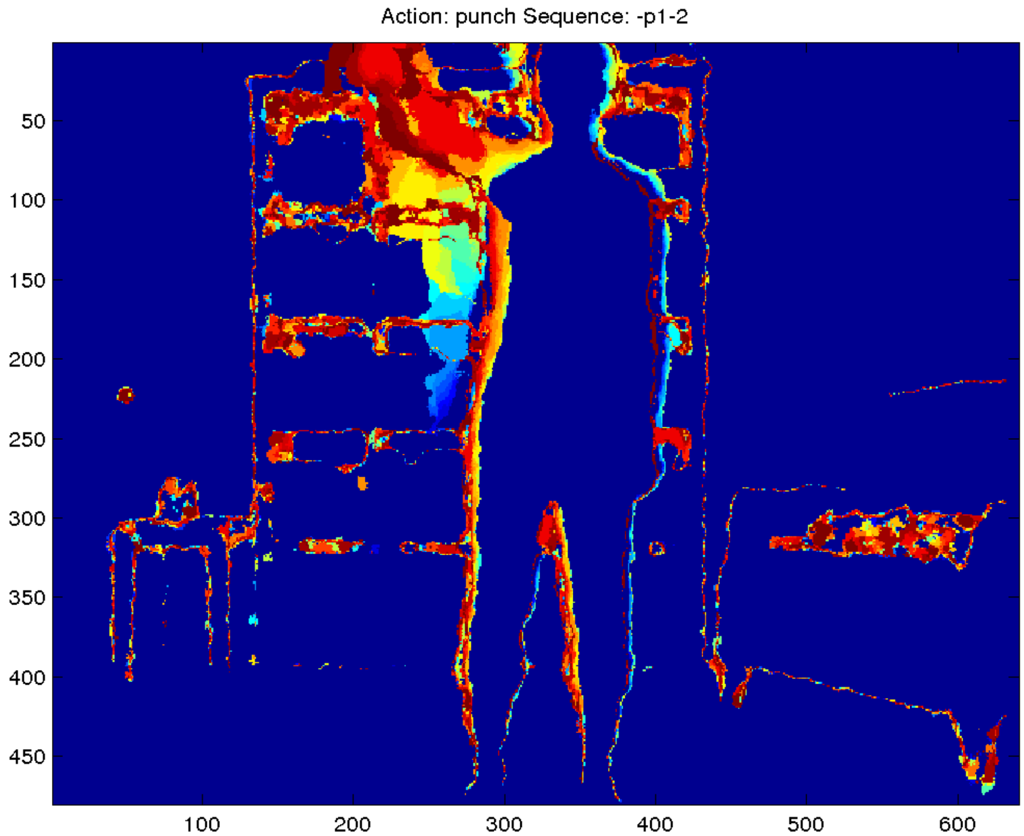

Figure 3: Motion History Image:3, Image Sequence: Punch, Sequence: 2

1.2 Hu Moments

In each MHI frame depicting the all the major changes in the frame, 7 Hu moments calculated to utilize as features. Then, each moment was compiled into one matrix as a feature space. Please refer "huVectors.mat" file to see results.

1.3 Activity Prediction

Activity prediction is achieved with kNN classifier. As distance metric normalized euclidian distance is employed. Figure-4 and Figure-5 show the best matches for punch MHI and crouch MHI, respectively. Table-1.3 shows the overall performance of the *Leave-one-out Cross Correlation* process.

Confusion Matrix

	Both Arms Up	Crouch	Left Arm Up	Punch	Right Kick
Both Arms Up	3	0	1	0	0
Crouch	0	1	3	0	0
Left Arm Up	0	1	3	0	0
Punch	0	0	0	4	0
Right Kick	0	1	0	0	3

Table 1: Confusion Matrix

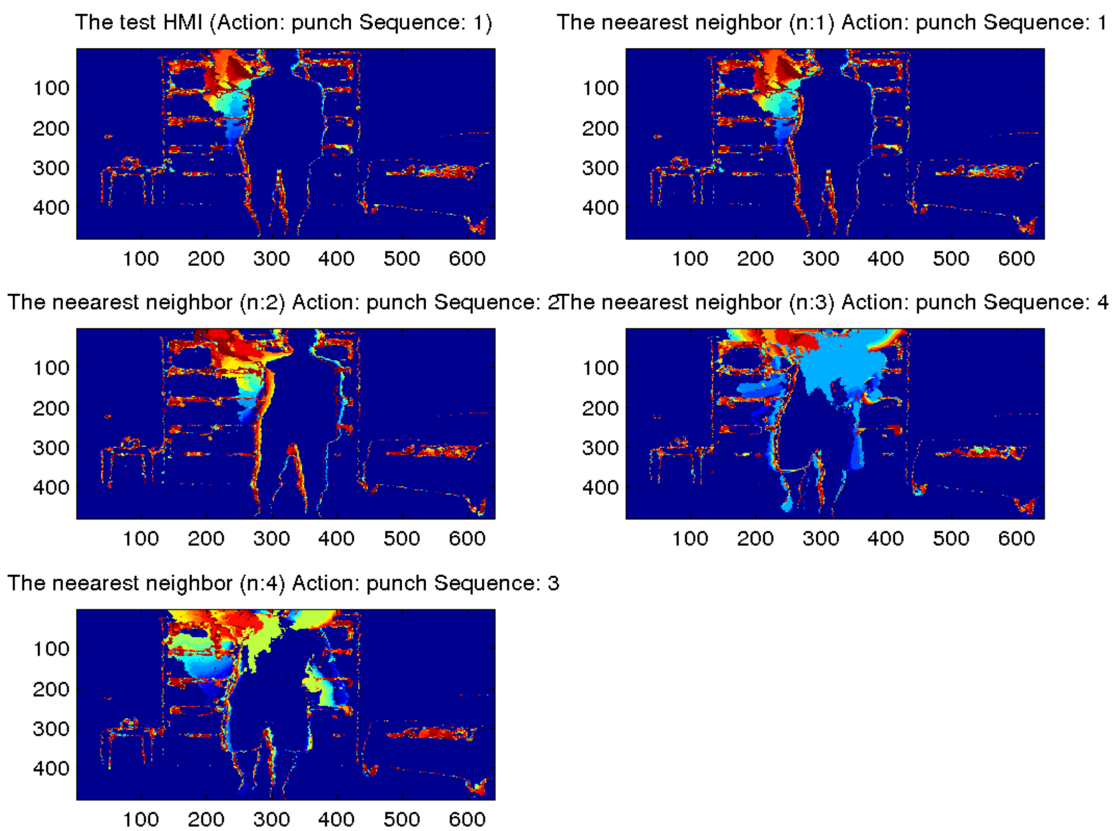

Figure 4: Top left is the query MHI (Punch), the rest is the best resulted matches

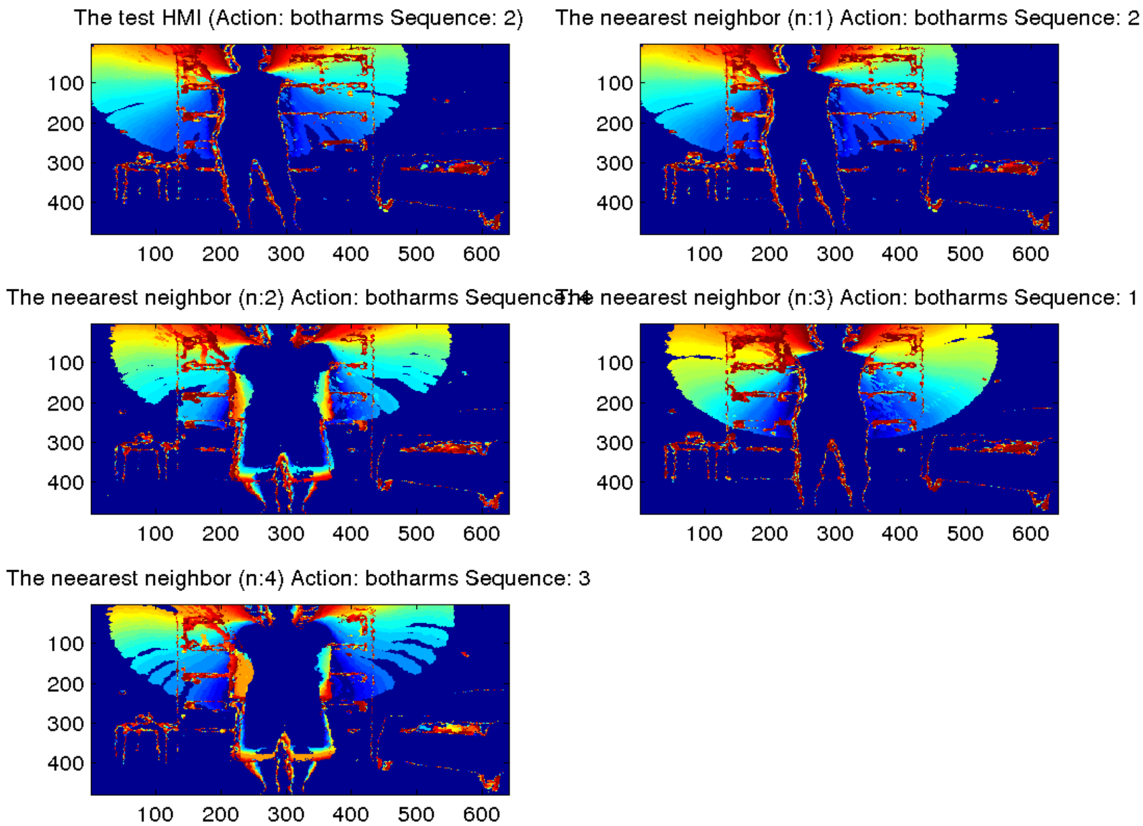

Figure 5: Top left is the query MHI (Both Arms Up), the rest is the best resulted matches

2 Extra Point

2.1 Depth Segmentation

Since the given data set is a depth map, the pixels correlates with the pyhsical distances of the objects to the camera. That cue is exploited by using depth segmentation before the $D(x, y, t)$ constructed by thresholding the depth maps.

$$I_{segmented} = I(x, y, t) > \tau$$

The following figure (Figure-6) the results of the depth segmentation with threshold $\tau = 40000$ value.

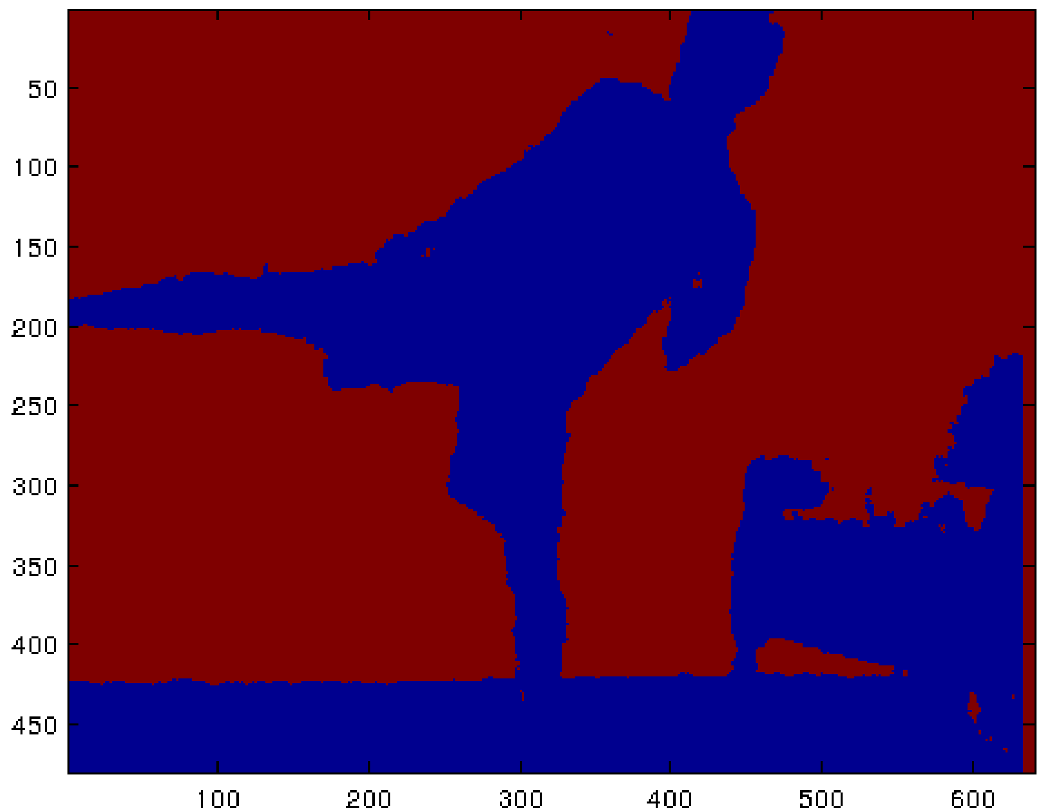

Figure 6: Example result of depth segmentation process

After thresholding, $min()$ function performed to consecutive frames to calculate $D(x, y, t)$. The figure below (Figure-7) is an example resulting MHI frame. Also, Table-2.1 shows the resulting confusion matrix. Please note that depth segmentation process has increased the overall performance of the algorithm by correcting the failure case in the “Crouch” segments. On the other hand, it also should be considered that in “Punch” segment, depth segmentation has showed 1 misclassification where the proposed technique can show the best performance.

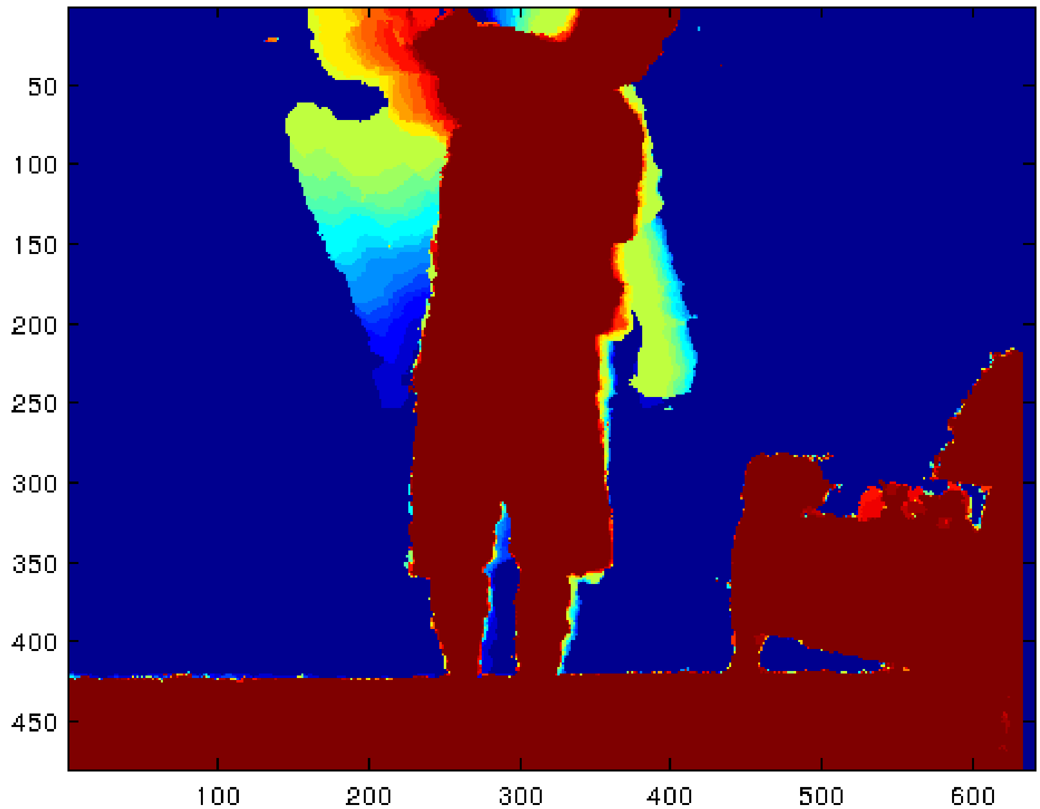

Figure 7: Example MHI Frame Resulted After Depth Segmentation

	Both Arms Up	Crouch	Left Arm Up	Punch	Right Kick
Both Arms Up	4	0	0	0	0
Crouch	0	3	0	1	0
Left Arm Up	0	0	4	0	0
Punch	0	0	1	3	0
Right Kick	0	0	0	0	4

Table 2: Confusion Matrix