

COMP4421 (Fall 2018)

Assignment #3

Name: Kristian Suhartono    StudentID : 20313207    E-mail: [ksuhartono@connect.ust.hk](mailto:ksuhartono@connect.ust.hk)

I. Exercises

1. Huffman Coding

| Original Source |        |       | Source Reduction |      |               |     |               |    |               |    |               |   |
|-----------------|--------|-------|------------------|------|---------------|-----|---------------|----|---------------|----|---------------|---|
| Intensity       | Prob   | Code  | 1                |      | 2             |     | 3             |    | 4             |    | 5             |   |
| 1               | 0.25   | 10    | 0.25             | 10   | 0.25          | 10  | <b>0.3125</b> | 11 | <b>0.4375</b> | 0  | <b>0.5625</b> | 1 |
| 3               | 0.25   | 01    | 0.25             | 01   | 0.25          | 01  | 0.25          | 10 | 0.3125        | 11 | 0.4375        | 0 |
| 9               | 0.1875 | 00    | 0.1875           | 00   | 0.1875        | 00  | 0.25          | 01 | 0.25          | 10 |               |   |
| 7               | 0.125  | 110   | 0.125            | 110  | <b>0.1875</b> | 111 | 0.1875        | 00 |               |    |               |   |
| 2               | 0.0625 | 1110  | <b>0.125</b>     | 1111 | 0.125         | 110 |               |    |               |    |               |   |
| 12              | 0.0625 | 11111 | 0.0625           | 1110 |               |     |               |    |               |    |               |   |
| 15              | 0.0625 | 11110 |                  |      |               |     |               |    |               |    |               |   |

| Intensity | Frequency | Compressed(bits) | Original Size | Compressed Size |
|-----------|-----------|------------------|---------------|-----------------|
| 1         | 4         | 2                | 16            | 8               |
| 3         | 4         | 2                | 16            | 8               |
| 9         | 3         | 2                | 12            | 6               |
| 7         | 2         | 3                | 8             | 6               |
| 2         | 1         | 4                | 4             | 4               |
| 12        | 1         | 5                | 4             | 5               |
| 15        | 1         | 5                | 4             | 5               |
| Total     |           |                  | 64            | 42              |

Compression Ratio:  $64 / 42 = 1.523809524 = 1.5238$

## 2. Adaboost Classifier

T = 1:

Sample weights:

|        |        |        |        |        |        |        |        |        |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0.1111 | 0.1111 | 0.1111 | 0.1111 | 0.1111 | 0.1111 | 0.1111 | 0.1111 | 0.1111 |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|

Error : 0.2222

Classifier Weight:  $1 / 2 * \ln((1 - 0.4444) / 0.4444) = 0.1117$

T = 2:

Sample weights before normalization:

|        |        |        |        |        |        |        |        |        |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0.0994 | 0.0994 | 0.0994 | 0.0994 | 0.0994 | 0.1111 | 0.1111 | 0.1111 | 0.1111 |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|

Sum: 0.9414

Sample weights after normalization:

|        |        |        |        |        |        |        |        |        |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0.1056 | 0.1056 | 0.1056 | 0.1056 | 0.1056 | 0.1180 | 0.1180 | 0.1180 | 0.1180 |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|

Error : 0.6708

Classifier Weight:  $1 / 2 * \ln((1 - 0.6708) / 0.6708) = -0.3559$

T = 3:

Sample weights before normalization:

|        |        |        |        |        |        |        |        |        |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0.1056 | 0.1507 | 0.1507 | 0.1056 | 0.1056 | 0.1111 | 0.1111 | 0.1111 | 0.1586 |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|

Sum: 1.1101

Sample weights after normalization:

|        |        |        |        |        |        |        |        |        |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0.0951 | 0.1358 | 0.1358 | 0.0951 | 0.0951 | 0.1001 | 0.1001 | 0.1001 | 0.1429 |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|

Error : 0.2903

Classifier Weight:  $1 / 2 * \ln((1 - 0.2903) / 0.2903) = 0.4470$

T = 4:

Sample weights before normalization:

|        |        |        |        |        |        |        |        |        |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0.0711 | 0.1016 | 0.1016 | 0.0951 | 0.0951 | 0.1001 | 0.0749 | 0.0749 | 0.1069 |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|

Sum: 0.8213

Sample weights after normalization:

|        |        |        |        |        |        |        |        |        |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0.0865 | 0.1237 | 0.1237 | 0.1157 | 0.1157 | 0.1219 | 0.0912 | 0.0912 | 0.1301 |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|

Error : 0.7683

Classifier Weight:  $1 / 2 * \ln((1 - 0.7683) / 0.7683) = -0.5994$

T = 5:

Sample weights before normalization:

|        |        |        |        |        |        |        |        |        |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0.0865 | 0.1237 | 0.1237 | 0.2107 | 0.2107 | 0.1219 | 0.0912 | 0.0912 | 0.1301 |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|

Sum: 1.1897

Sample weights after normalization:

|        |        |        |        |        |        |        |        |        |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0.0727 | 0.1040 | 0.1040 | 0.1771 | 0.1771 | 0.1024 | 0.0767 | 0.0767 | 0.1094 |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|

Error : 0.5419

Classifier Weight:  $1 / 2 * \ln((1 - 0.5419) / 0.5419) = -0.0840$

a. Two best classifiers at T = 3 and T = 1, so H3 and H1

- b. Final classifier response:  
[1, 1, -1, 1, -1, -1, -1, 1, -1]

## II. Programming Tasks

## 1. Digit Segmentation

The algorithm:

- Preprocessing the image to remove noise
- Thresholding the input to get binary images
- Horizontal and vertical projection to get the digits (finding the peaks in each projection)
- Clearing of excess empty spaces
- Resizing of image to 20x20
- Padding to 28x28
- Centering of image based on the center of gravity

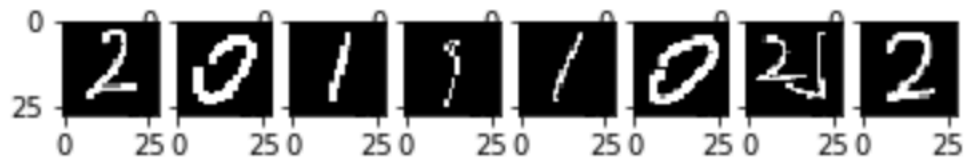
a. Image 1:



b. Image 2:



c. Image 3:



## 2. Classification via Adaboost

Classifier used: Decision Tree

- Accuracy of Image 1:  $15/16 = 93.75\%$
- Accuracy of Image 2:  $11/12 = 91.67\%$
- Accuracy of Image 3:  $11/16 = 68.75\%$