Program Assignment #2

- Due Date: Dec. 4 (Tue) 23:59
- Submission: Source code & summary report

Design a handwritten character and word recognition system as follows.

    1. Preprocessing

      - Words and characters segmentation

        • Separate an individual character from a given file.

        • In the case of a word, include the process of separating the word into characters.

      - (Optional) Rotation

        • Adjust the given characters to the same angle.

      - (Optional) Thickness

        • Adjust the given characters to the same thickness.

      - If there are other preprocessing tasks that you think would be necessary, design and implement them. It is optional, you do not need to implement optional preprocessing tasks.

    2. Feature extraction

      - Implement your own feature extraction method to distinguish each alphabet.

    3. Classifier

      - Recognize the segmented characters and words.

      - In the case of a word, recognize full character strings for the word.

Notation

    - It is most desirable for your program to work as a whole, i.e., it can accept both the character and the word. (Additional points)

    = When a character is an input, your program works as a character-based recognizer. Furthermore, your program works as a word-based recognizer when a word is an input.

    - It is allowed to implement character and word recognizers, separately.

Data description

    - Training data

      • There are 5 character image files that alphabets a, b, c, d, e are written respectively.

      • Each file consists of 60 characters.

      • Choose the scale of data according to your preference: <u>grayscale or binary scale</u> (black and white)

    - Validation data

      • There are 5 character image files and 4 word image files. (All grayscale files)

      • Each character image file consists of 30 characters.

      • Word image files consist of 80, 45, 90, 30, and 30 words, respectively.

Ex)

Report

- The report should be structured as follows. (Within 10 pages including cover page)

1. Describe your preprocessing, feature vectors, and classifier methods.

2. Demonstrate the recognition rates for character and word recognition using valid data.

3. Summarize how to run your program for evaluation.

- Describe the theoretical content briefly and focus on the actual implementation method.

Precaution

1. Implement environment: MATLAB

   - You should include 'main.m' function in your program for evaluation.

2. If your report exceeds 10 pages, there would be penalties.

3. File name: ID_name.zip (Source code and report should be compressed in zip format)

   ex) 20180000_HeejinChoi.zip

4. TA will evaluate final performance of your program using the test data that is not provided.

   → To avoid overfitting

5. Rate: Character recognition rate (60%), Word recognition rate (30%), Report (10%)

   PA2 score = 100 – (Performance rank * 1 – 1)

   ex) First place: 100 points / Second place: 99 points / Third place: 98 points / …

Good Luck!