RECOGNIZATION AND RECOMMENDATION OF HANGUL FONTS USING COMPUTER VISION BASED TECHNIQUES

CHUNG SEIJUN ???????? OUM ?????? KIM ??????? CHOO ??????

Computer Vision 2019-1
Department of Computer Science
Seoul National University
(Dated: April 26,2019)

Fonts are a set of pre-designed ... I don't know what should go in to an abstract. PLEASE FILL ME IN

Objective: This project recognizes different Hangul fonts based on Computer Vision Techniques and provide akin fonts that avoids legal issues.

1. INTRODUCTION

From the monospaced fonts on a VT100 terminal to the Seoul fonts by the Seoul city government, the range of available fonts to users have expanded, due to the rise of tools for font creation.

Fonts are predefined collections of letters, varying in design and size. Different fonts are utilized based on the purpose of the user, as an intermediary of the user's purposes and emotions.

However, as more and more fonts are being introduced to the market, the quest for the appropriate font is either an evergoing battle with vast images, or an effortless defeat, known as the default font. Even when one has a desired font with its image in hand, the search is analogous to finding a needle in the haystack. Also, finding the appropriate font is not just a matter of choice, as it can spur legal disputes as well.[?]

As of now, there is no way to search the desired font from an input of image one has found in books, magazines, or from a street sign;

Compared to Hangul fonts, English fonts consist of 26 upper and 26 lower case letters, a total of 52 characters. The limited number of characters, and their simple form makes English font recognition not as challenging and previous researches have prodcued high recognition rate(add reference?),

However, the Hangul font consists of 2,350(KS X 1001) or 11,172(KPS 9566) characters, each composed of onset, nucelus, and codas. Due to this complicated format, research on font recogntion of Hangul characters is not active, compared to other Latin alphabets.

2. PAPER SURVEY

2.1. Classical Computer Vision

We shall utilize many techniques learned from our class and materials.

2.2. Optical Font Recognition Using Typographical Featurs

This paper[?] aims to identify typeface, weight, slope and the size of text from an image block without any knowledge of the content of the text.

2.3. Large-Scale Visual Font Recognition

This paper[?] addresses the large-scale visual font recognition (VFR) problem, which aims at automatic identification of the typeface.

3. KEY FEATURES

3.1. Hangul componenet decomposition

Cosidering its outline and location, decompose a Hangul glyph into its onset, necleus and coda, using stastistical-structural information.

3.2. Feature Detection

A Study on Typology for Hangul Fonts Identification and Classification in terms of character's typeface, character size and character slope and etc. based on Computer Vision Techniques.

3.3. Machine Learning

Mass of Hangul character fonts recogntion with CNN¡ Hideen Markov Model with reference to paper[?] or other materials.

3.4. Estimating Surface Normal Vector of Hangul Fonts (Optional)

4. ALGORITHM

```
detect_letters("input.img")
print("Hello, World!")
```

5. EXPECTED RESULTS

We aim for the following goals:

1. Extract distinct features of Korean characters based on Computer Vision Techniques
2. Recognize the provided font
3. Recommend most similar fonts

6. DATASET

The experimental data set used in this paper is **PHD08**[**?**], a large-sacle Korean character database. **PHD08** has 2,187 samples for each of 2,350 Korean letters, consisting of a total 5,139,450 samples.

References

[1] The Font Bureau, Inc. v. NBC Universal, Inc. et al https://dockets.justia.com/docket/new-york/nyedce/1:2009cv04286/296847

[2] Abdelwahab Zramdini, et al. *Optical Font Recognition Using Typographical Features*. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLI-GENCE, 1988.

[3] Guang Chen, et al. *Large-Scale Visual Font Recognition*. CVPR, 2014.

[4] Chris Tensmeyer, et al. *Convolution Neural Networks for Font Classification*. CVPR, 2017.

[5] D. Ham, D.Lee, I.Jung and I. Oh. *Construction of Printd Hangul Character Database PHD08* Journal of the Korea Contents Association, vol.8, no.11, pp. 33-40, Nov. 2008.