Генерация стуктурных признаков для классификации букв в цифровых изображениях

Липкина Анна Львовна

Студент

Факультет ВМК МГУ имени М. В. Ломоносова, Москва, Россия

E-mail: lipkina96@mail.ru

Задача распознавания печатных текстов остается актуальной, несмотря на большое число существующих OCR систем, поскольку востребованы недорогие компактные алгоритмы для частных приложений в типографике (распознавание шрифтов по изображениям текста), в оцифровке документов (таблицы, чеки, накладные) и т.п. Существует множество подходов к решению данной задачи: например, структурный анализ изображения или применение нейросетевых классификаторов [4–5].

В данной работе предлагается подход к генерации признаков для классификации букв, опирающийся на структурный анализ изображения буквы.

У каждой буквы есть всего лишь несколько фиксированных очертаний (*графем*) [2], по которым можно быстро понять, что это за буква. В качестве такого очертания в данной работе рассматривается *непрерывное скелетное представление* графемы [1]. Целью исследования является проверка гипотезы о том, что скелетное представление является достаточным для распознавания букв в большинстве шрифтов [3].

Идея построения признаков для распознавания заключается в выделении ключевых вершин степеней 1 и 3 в скелетном представлении графемы и последующем построении признаков на основе их положения (рис. 1). Прямоугольный участок плоскости, в который заключена графема, разбивается на n равных частей горизонтальными прямыми и на m равных частей — вертикальными. В каждом из получившихся mn прямоугольников считается отдельно количество ключевых вершин степени 1 и степени 3. В итоге для каждой графемы получается признак из $2nm$ элементов (рис. 2). По этим признакам строится словарь классификации: «признак — список букв, имеющие данный признак».

Для проведения эксперимента по построению такого словаря были выбраны следующие значения: $m = n = 3$. В результате эксперимента была оценена возможность построения классификатора по

составу и расположению вершин 1 и 3 степени скелетного графа.

Если в качестве основы для построения словаря брать графемы только из одного шрифта, то точность распознавания других графем в других шрифтах оказывается не очень высокой — порядка 70% – 80%. Метрика качества считалась следующим образом: для каждой новой графемы проверяется, попало ли её признаковое описание в словарь или нет. В случае успеха, в списке слов по данном признаку проверяется наличие (попадание) буквы, которую изображает графема. В качестве метрики качества брался процент таких попаданий в словарь.

Если при каждом непопадании графемы в словарь фиксировать и отображать это в словаре, то таким образом можно пополнить нашу базу знаний о том, какие графемы могут быть у букв. Поэтому далее был проведен эксперимент по построению словаря по представлениям графем в 5 различных шрифтах. Точность распознавания на других шрифтах в этом случае составила не менее 85%.

Стоит отметить, что этот метод обладает двумя достоинствами: легкость встраивания новых знаний о графемах и независимость распознавания графемы от размера и шрифта.

Предложенный метод можно использовать, например, для предварительного отбора похожих эталонов. Но использование только вершин скелетного графа не исчерпывает предлагаемый подход. В частности скелетный граф позволяет строить признаки, основанные на длине, ориентации и кривизне ребер скелета. Их применение будет исследовано в дальнейшем.

Иллюстрации

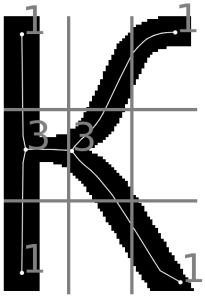

Рис. 1. Графема буквы «К», её скелетное представление и выделенные ключевые вершины степеней 1 и 3

1	0	1
0	0	0
1	0	1

0	0	0
1	1	0
0	0	0

Рис. 2. Признаки для графемы буквы «К», построенные по количеству ключевых вершин степени 1 (слева) и степени 3 (справа)

Литература

1. Местецкий Л. М. Непрерывная морфология бинарных изображений: фигуры, скелеты, циркуляры. М.: ФИЗМАТЛИТ, 2009.
2. Гордон Ю. Книга про буквы от Аа до Яя. М. Изд-во Студии Артемия Лебедева, 2017.
3. ParaType. Цифровые шрифты. ParaType, 2008.
4. Страница с примерами распознавания графем компанией ABBYY
   `https://habrahabr.ru/company/abbyy/blog/228251/`
5. Страница с примерами методов распознавания текста
   `https://habrahabr.ru/post/220077/`