

Мера сходства элементов штрихового представления рукописного текста на основе Фурье-дескриптора

Феоктистов Дмитрий Дмитриевич^{1*}

feoktistovdd@my.msu.ru

Местецкий Леонид Моисеевич^{1,2}

mestlm@mail.ru

¹Москва, МГУ

² Москва, ВШЭ

В работе рассматривается подход к распознаванию русскоязычных рукописных текстов, основанный на штриховой сегментации цифровых изображений текстовых документов. Штриховая сегментация состоит в представлении изображения текста в виде последовательности каллиграфических элементов письма, называемых штрихами: овалов, палочек, крючков и т.п. Штриховая сегментация осуществляется методами, основанными на непрерывном медиальном представлении изображения рукописного текста [1].

Распознавание текста предлагается свести к количественной оценке сходства и различия отдельных штрихов, а также кортежей штрихов. Такой подход основывается на следующей гипотезе: все почерки являются вариацией некоторого эталонного «правильного» варианта написания каллиграфических элементов. Действительно, обучение письму производится с помощью прописей, в которых в течение многих лет, на протяжении жизни нескольких поколений, не меняются правила написания штрихов, из которых строятся буквы и слова. В основе подхода лежит разработка меры сходства и различия штрихов, полученных в результате штриховой сегментации. Целью является построение дискриминативной метрики на пространстве штрихов [2]. В нашей работе исследуется метод построения такой метрики на основе так называемых Фурье-дескрипторов [3, 4] штрихов.

Алгоритм, описанный в [1], представляет каждый штрих в виде ломаной $L = \{(x_j, y_j)\}_{j=1}^l$. На нее можно посмотреть, как на последовательность точек в комплексной плоскости $\{x_j + iy_j\}_{j=1}^l$, т.е. как на сигнал, для получения векторного представления которого предлагается использовать дискретное преобразование Фурье. Для первичного определения того, является ли предлагаемое векторное представление дискриминативным была рассмотрена задача классификации штрихов набора панграмм, изображенного на рисунке 1.

В первую очередь была произведена проверка того, как ведет себя отношение внутриклассового расстояния к межклассовому в евклидовой метрике в зависимости от количества используемых коэффициентов Фурье. Как видно из рисунка 2, при всех значениях наблюдается выполнение гипотезы компактности, при этом при увеличении количества коэффициентов наблюдается деградация разделимости классов, т.к. появляются коэффициенты, отвечающие за шумы. Из этого можно сделать вывод, что для описания штрихов достаточно использовать небольшое количество коэффициентов Фурье.

аэрофотоъёмка ландшафта уже
выселила земли богатей и
процветающих крестьян.
всё ускоряющаяся эволюция
компьютерных технологий
предъявила жёсткие требования
к производителям как собственно
вычислительной техники, так и
периферийных устройств.
завершён ежегодный съезд
эрудированных школьников, мечтающих
глубоко проникнуть в тайны
физических явлений и химических
реакций.
съели же ещё этих легких
французских булок да выпей чаю.
флегматичная эта верблюдница
исцелит у подвезда засыхающий
горький шиповник.
микровавилонница попросту забыла
ред ключевых множителей
и т.п.

Рис. 1. Изображение текста, для штрихов которого рассматривалась задача классификации

Рис. 2. Результаты эксперимента, проверяющего гипотезу компактности для различного количества коэффициентов Фурье

Для проверки предложенной гипотезы решалась задача классификации штрихов методом k ближайших соседей, качество измерялось на кросс валидации с пятью разбиениями выборки с помощью F1-масо метрики, т.к. она учитывает дисбаланс классов. Для каждого количества коэффициентов Фурье выбиралось лучшее значение k по результатам кросс валидации, после чего изучались результаты этого алгоритма. Как видно из рисунка 3 существенной разницы между использованием 5–23 коэффициентов Фурье нет: и значение метрики, и ее устойчивость примерно одинаковы. Значит, мы можем использовать малое количество признаков, например, 5 в случае евклидова расстояния и 7 в случае косинусного. Отметим, что идею с выбором расстояния следует развить, т.к. косинусное расстояние дает большие значения F1-масо, хоть и менее устойчивые. Это говорит о том, что может существовать более дискриминативное расстояние, чем рассмотренные. Для этой цели предлагается использовать алгоритмы обучения метрик.

Таким образом, в работе предложен способ векторизации штрихов с помощью дискретного преобразования Фурье. Экспериментально показано, что для

Рис. 3. Результаты эксперимента, проверяющего гипотезу компактности для различного количества коэффициентов Фурье

такого представления штрихов выполнена гипотеза компактности, также показано, что для оптимального описания достаточно 5–7 коэффициентов. При дальнейшем развитии работы будет исследована возможность применения алгоритма Фурье для расстояния более дискриминативного, чем евклидово и косинусное. После чего найденное расстояние будет применено к задаче поиска ключевых слов в рукописном контексте.

Работа поддержана грантом РНФ № 22-68-00066.

- [1] Местецкий Л. М. Штриховая сегментация рукописного текста // Математические методы распознавания образов: Тезисы докладов 21-й Всероссийской конференции с международным участием, г. Москва 2023 г, М.: Российская академия наук, 2023. — С. ??–??.
- [2] Пронина Н. М., Местецкий Л. М. Классификация штрихов рукописного текста на основе расстояния Фреше // 33-я Международная конференция по компьютерной графике и машинному зрению «ГрафиКон 2023», г. Москва, ИИУ РАН, Russia, 19-21 сентября 2023.
- [3] Zahn C. T. and Roskies R. Z. Fourier descriptors for plane closed curves. // IEEE Transactions on Computers. C-21(3):269–281, 1972.
- [4] Yip R., Tam P. and Leung D. Application of elliptic fourier descriptors to symmetry detection under parallel projection. // Pattern Analysis and Machine Intelligence, IEEE Transactions on, 16:277–286, 1994.