

Королев С.Н., аспирант

Научный руководитель Нифонтов Ю.А., доц., канд. техн. наук

АЛГОРИТМ КОРРЕЛЯЦИОННО-ЧАСТОТНОГО РАСПОЗНАВАНИЯ ЗАШУМЛЕННЫХ ТЕКСТОВЫХ ИЗОБРАЖЕНИЙ

Рассматривается метод оптического распознавания текстового изображения, подверженного сильным искажениям шумового и морфологического характера. В частности, в ходе экспериментов исходное (идеальное) изображение подвергалось воздействию так называемого спекл-шума (мультипликативный шум, искажающий форму букв) и аддитивного гауссовского шума, искажающего фон изображения. На рис. 1 приведен пример текстового изображения, на котором тестировался разработанный алгоритм классификации.

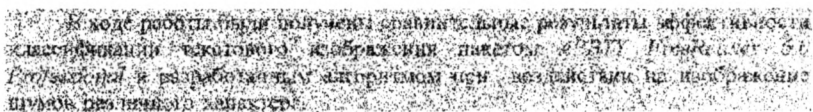


Рис. 1. Пример текстового изображения для классификации

Задача осложнена тем, что зачастую реальные текстовые изображения, получаемые с фото- и видеокamer, используемых, например, в аппаратуре распознавания номеров движущихся автомашин, вагонов железнодорожного состава или в аппаратуре автоматического распознавания факсов (при некачественной линии связи), помимо аддитивных и мультипликативных помех шумового характера имеют сложный вид искажений, такой как межсимвольная интерференция, которую невозможно преодолеть обычной линейной фильтрацией, и возможно лишь частично ослабить комплексом из морфологических операций. В связи с этой проблемой известные методы сегментации текстового изображения (т.е. разбиения его на однородные области – строки, слова и буквы) в условиях данной задачи являются неэффективными. Разработанный метод распознавания основан на алгоритме расчета взаимных корреляционных функций набора эталонных изображений букв алфавита и анализируемого изображения в частотной области, используя при этом разбиение изображения только на строки и слова.

Задача решается путем анализа изображений слов и эталонных изображений букв в частотной области, т.е. путем расчета их двумерных ДПФ распределений, вычисления их свертки (для всех эталонов букв в цикле, для каждого знакоместа анализируемого слова) и последующего расчета обратного ДПФ от полученной свертки.

Коэффициент взаимной корреляции вычисляется как максимальное значение двумерной функции, получаемой после вышеперечисленных действий. Вид этой функции показан на рис.2.

Т.к. некоторые буквы похожи на другие, и это сходство еще более усугубляется при действии помех, то полученное двумерное тело функции распределения может иметь несколько локальных максимумов, вносящих значительную неопределенность при определении местоположения буквы в слове.

Результатом работы первой части алгоритма классификации являются предполагаемые координаты расположения букв-эталонов и значения корреляций в этих точках, т.е. набор гипотез о местоположении букв в слове, из которых необходимо выбрать истинные значения. Далее в действие вступает вторая часть алгоритма, заключающаяся в сложной сортировке полученных ранее результатов и принятии решения в пользу определенной буквы одним из трех способов – метод отбора символа с максимальной корреляцией (максимум максиморум), метод максимального правдоподобия (учитывается энергетика букв), метод максимального правдоподобия в комбинации с априорной статистикой частоты появления букв в тексте.

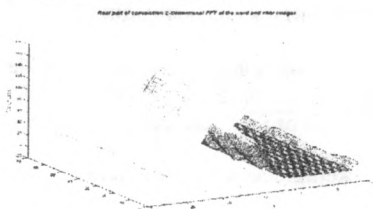


Рис. 2. Вещественная часть инверсного двумерного ДПФ от свертки частотных преобразований распознаваемого слова и буквы-эталона

В ходе экспериментов были получены сравнительные результаты эффективности классификации текстового изображения пакетом ABBYY FineReader 6.0 Professional и разработанным алгоритмом при воздействии на изображение шумов различного характера. Результаты экспериментов сведены в таблицу. В качестве интегральной оценки качества распознавания принят такой параметр, как вероятность правильного распознавания (число правильно распознанных символов).