

УДК 004.93

ОБНАРУЖЕНИЕ ТЕКСТА ПРИ ПОМОЩИ НАПРАВЛЕННОГО ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ

Голиков И.Ю.

к.т.н., доцент каф. ИС Николенко А.А.

Одесский Национальный Политехнический Университет, УКРАИНА

АННОТАЦИЯ. В статье предложен алгоритм, повышающий точность определения текстовых областей при помощи направленного вейвлет-преобразования, что может быть использовано для дальнейшей разработки системы оптического распознавания текстов.

Введение. Распознавание образов – обширная тема для дискуссии в различных областях современного общества. Сфера применения велика: начиная от распознавания текстовых конструкций, заканчивая моделированием систем распознавания 3D-объектов. Одной из актуальных задач в данной научной сфере является проблема распознавания символов. Решение этой задачи можно использовать в системах для распознавания автомобильных номеров в режиме реального времени, аннотирования изображений, перевода рукописных работ в электронный вид и т.п.

Для решения задач распознавания используются различные подходы, однако наибольшую популярность приобрели методы, не требующие предварительной сегментации, такие как скрытые модели Маркова и свёрточные нейронные сети [1]. В последнее время в исследовательских работах по оптическому распознаванию символов часто используются нейросетевые классификаторы.

Цель работы. Целью работы ставится разработка алгоритма, позволяющего повысить точность определения текстовых областей при помощи направленного вейвлет-преобразования.

Основная часть работы.

Условно задачу распознавания текста можно разбить на несколько этапов:

1. Выделение векторов признаков для обучения.
2. Уменьшение размерности векторов выделенных признаков.
3. Обучение нейронных сетей на полученных векторах признаков.

Одной из главных задач при обработке изображения является нахождение эффективного представления, позволяющего отобразить его в компактном виде. Вейвлет-преобразование является действенным методом для выделения признаков объектов в задачах распознавания, зарекомендовавшим свою эффективность в спектральном анализе сигналов. Наиболее используемым вейвлетом является вейвлет Хаара. Его использование обусловлено относительной простотой алгоритма (однократный и имеет квадратичную сложность $O(N^2)$).

В реалиях современных исследований и наращивания аппаратных мощностей представляется возможность использовать более сложные алгоритмы, позволяющие решать проблемы, возникающие при использовании вейвлета Хаара, такие как неправильное распознавание наклонных печатных текстов и неправильное распознавание повернутого текста.

Предлагается алгоритм, позволяющий повысить эффективность выделения признаков и их представления в более компактной форме. Использование непрерывных двумерных вейвлетов позволяет использовать новый параметр: угол поворота. Проведен эксперимент, демонстрирующий повышение информативности признаков распознаваемого текста. Информативность оценивалась со структурной точки зрения в количестве пикселей на выходе.

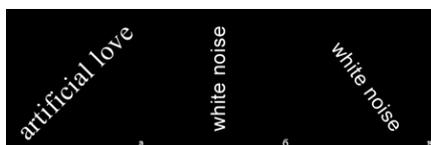


Рис. 1 – Тестовые изображения с углами: а) 45°, б) 90°, в) 127°

В качестве входа использовалось изображение с разрешение 256*256 формата *.jpeg, для его обработки использовался комплексный непрерывный двумерный вейвлет Морле [2]. Эксперимент проводился с фиксированным масштабом и не изменялся для каждого из шагов эксперимента. Предварительно зафиксировав угол, с которым был наклонен текст (рис.1), производилось преобразование с шагом угла $\pi/12$.

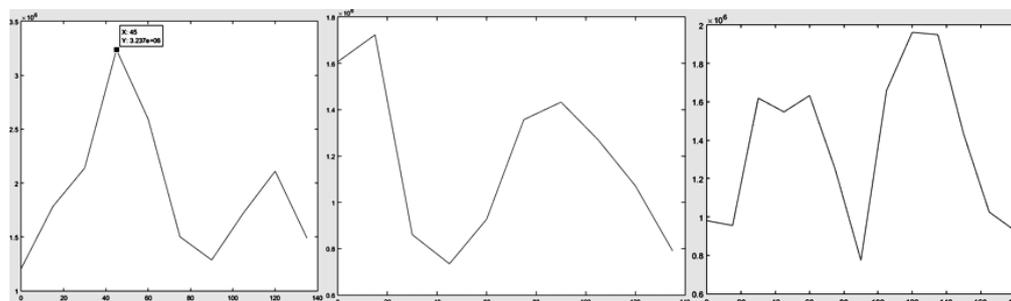


Рис. 2 – Графики зависимостей количества пикселей (y) от угла вейвлета (x) для случаев тестовых изображений : а) 45° , б) 90° , в) 127°

По результатам эксперимента были получены изображения модулей вейвлет-преобразования, для которых производились дальнейшие расчеты. Средствами Matlab были получены числовые значения количества пикселей на изображениях. Получены графики зависимостей количества пикселей от угла для трех рассматриваемых случаев (рис.2). Из графиков видно, что с изменением угла направленности вейвлета меняется также и информативность признаков, причем для каждого конкретного случая наблюдаются различные зависимости. Эксперимент (а) доказывает теоретическое предположение о том, что наиболее информативным является угол, на который изначально был повернут текст. Такой вывод позволит в дальнейшем моделировать системы распознавания, в которых в зависимости от входных параметров будет производиться проецирование осей в соответствии с найденным углом, что существенно повысит точность распознавания текстовых областей на изображении. Эксперимент (в) проводился для изображения текста, повернутого в диапазоне $[120; 135]$ градусов (127°). Из графика видно, что максимальная информативность достигается на заданном диапазоне. Это позволяет говорить о том, что даже при отсутствии точного значения угла максимальное количество пикселей распределяется в окрестностях начального угла поворота, что позволяет выбрать приблизительное значение угла для поворота и дальнейшего распознавания. При видимых преимуществах метод имеет существенный недостаток: при распознавании изображения, повернутого на 90° , происходит неверное распознавание и максимальная информативность признаков соответствует 0° (рис.2б), поэтому оговаривается, что метод применим только для наклонных текстов.

Выводы. Был проведен ряд экспериментов, демонстрирующих преимущества использования направленного вейвлет-преобразования для нахождения текстовых областей и сделаны соответствующие выводы: угол поворота вейвлет-преобразования, соответствующий углу поворота текста на изображении, несет в себе наибольшую структурную информативность; увеличение числа выделяемых пикселей на 70% по сравнению с вейвлетом с углом поворота 0° (рис.2) для наклонных текстов, что позволяет в несколько шагов определить верный наклон текста и существенно увеличить точность определения текстовых областей на изображении. В дальнейшем, алгоритм можно использовать для систем оптического распознавания текста и продемонстрировать повышение эффективности выделения текстовых областей (например, используя критерий Прэтта [3]), а значит и системы в целом.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. LeCun, Y. Gradient based learning applied to document recognition / Y. LeCun, L. Bottou, Y. Bengio, P. Haffner. – 1998. – Vol. 86, Issue 11. – P. 2278-2324.
2. Смоленцев Н. К. Основы теории вейвлетов. Вейвлеты в MATLAB. – М.: ДМК Пресс, 2014. – 628 с.
3. Прэтт У. . Цифровая обработка изображений. / пер. с англ. Д. С. Лебедева. – М.: Мир, 1982. 792 с.