

МІНІСТЕРСТВО ОСВІТУ І НАУКИ УКРАЇНИ
ОДЕСЬКИЙ НАЦІОНАЛЬНИЙ ПОЛІТЕХНІЧНИЙ УНІВЕРСИТЕТ

ІНСТИТУТ КОМП'ЮТЕРНИХ СИСТЕМ

МАТЕРІАЛИ ДЕВ'ЯТОЇ
МІЖНАРОДНОЇ НАУКОВОЇ КОНФЕРЕНЦІЇ
СТУДЕНТІВ ТА МОЛОДИХ ВЧЕНІХ



ПРИСВЯЧЕНА 55-РІЧЧЮ
ІНСТИТУТУ КОМП'ЮТЕРНИХ СИСТЕМ

“Сучасні інформаційні технології 2019”

“Modern Information Technology 2019”



NetCracker®



23-24 травня

Одеса
«Екологія»
2019

УДК 004.93

ИСПОЛЬЗОВАНИЕ МЕТОДА НАПРАВЛЕННЫХ ВЕЙВЛЕТОВ ДЛЯ ЛОКАЛИЗАЦИИ ТЕКСТОВЫХ ОБЛАСТЕЙ НА ИЗОБРАЖЕНИИ

Соловьёва К.В.

к.т.н., доцент каф. ИС Николенко А.А.

Одесский Национальный Политехнический Университет, УКРАИНА

АННОТАЦИЯ. В работе рассматривается возможность повышения точности локализации текстовых областей на изображении за счёт применения метода направленных вейвлетов. Предлагается алгоритм, демонстрирующий возможность обнаружения текста, расположенного под произвольным углом.

Введение. Задачи нахождения и распознавания текста на изображении возникли достаточно давно и до сих пор не являются однозначно решёнными, в связи с чем существует большое количество методов, связанных с локализацией и классификацией текстовых областей. Наибольшую трудность вызывает задача локализации – определения положения текстовой области на изображении. На сегодняшний день не существует универсального метода для решения этой задачи, поэтому на практике использование тех или иных методов сильно зависит от типов изображений, рассматриваемых в конкретной ситуации.

Цель работы. Целью работы является повышение точности локализации текстовых областей на изображении, расположенных под произвольным углом к осям посредством применения метода направленных вейвлетов.

Основная часть работы. В результате проведения анализа существующих методов локализации текстовых областей было замечено, что очень небольшое число методов позволяет находить на изображении текст разного масштаба, расположенный под углом [1].

В рамках работы проблема определения угла поворота решается применением двумерного направленного непрерывного вейвлет-преобразования (2-D CWT), которое позволяет за счёт изменения параметров масштаба, сдвига и угла поворота «подчёркивать» перепады между текстовой и не текстовой областями с большей или меньшей степенью в зависимости от начального размера шрифта и направления текста относительно горизонтальной оси.

Для того чтобы воспользоваться возможностью 2-D CWT извлекать из анализируемого изображения признаки, связанные с различными углами поворота текста, необходимо правильно подобрать анализирующую вейвлет-функцию. Можно выделить изотропные (ненаправленные) и анизотропные (направленные) вейвлеты. Если вейвлет является анизотропным, в анализе прослеживается зависимость от угла, и двумерное CWT действует как локальный фильтр для изображения при определённых параметрах масштаба, сдвига и угла наклона [2,3].

Для решения задачи локализации текста, расположенного под произвольным углом, с помощью 2-D CWT предлагается следующий алгоритм:

1. Сформировать набор из N углов для анализа изображения, который бы равномерно охватывал диапазон от 0 до 2π .
2. Выбрать анизотропный вейвлет и подобрать (экспериментально) параметр масштаба, при котором вейвлет наилучшим образом реагирует на заданный размер шрифта.
3. Для каждого угла из сформированного набора вычислить 2-D CWT (результат – набор из N двумерных матриц коэффициентов вейвлет-преобразования).
4. Просуммировать абсолютные значения коэффициентов для каждого из N результатов преобразования и найти результат m с максимальным значением этой суммы. Соответствующий угол α , для которого считалось CWT – угол с максимальным значением «отклика».
5. Провести пороговую бинаризацию и применить операцию дилатации к результату преобразования m так чтобы текстовая область представляла собой замкнутый компонент.
6. Выделить полученную область в минимальный описывающий её прямоугольник, направленный под полученным углом α , обрезать изображение по его контуру и повернуть на

а. Полученный сегмент изображения – результат локализации – можно подавать на распознавание.

В ходе работы ради эксперимента, с целью демонстрации повышения точности локализации текста при получении информации об угле поворота надписи, было сгенерировано 8 различных изображений размером 512x512 с текстовыми строками одного размера шрифта, расположенными под различными углами на однородном чёрном фоне. Для анализа изображений был выбран анизотропный вейвлет Гаусса при фиксированном параметре масштаба. Вычисления производились в среде Matlab с использованием пакета Wavelet Toolbox.

2-D SWT вычислялось для фиксированного количества углов, а именно от 0 до $7\pi/8$ с шагом $\pi/8$ (8 направлений). Пример работы описанного выше алгоритма представлен на рисунке 1.

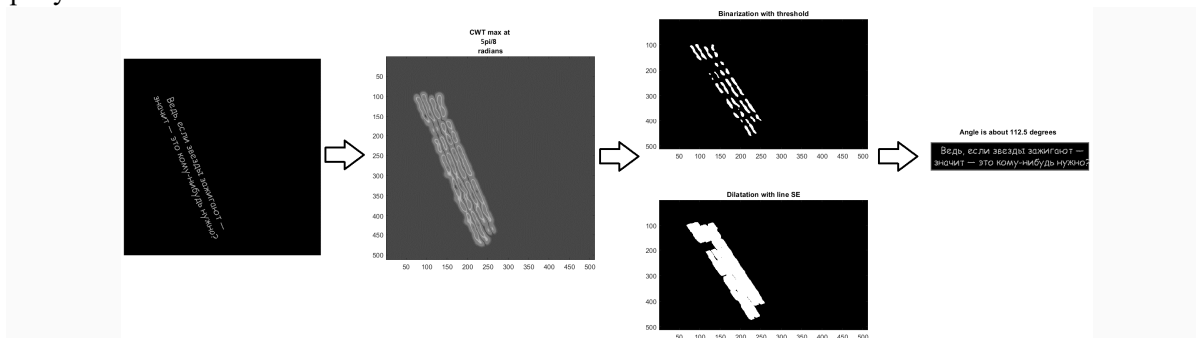


Рис. 1 – Пример работы алгоритма для изображения с текстом, направленным под углом 112°

В качестве критерия оценки точности локализации текстовой области было принято отношение площади пересечения области построенного ограничивающего прямоугольника (A) и замкнутой области, содержащей текст (B), к сумме (объединению) этих площадей:

$$p = \frac{A \cap B}{A \cup B} \quad (1)$$

Для каждого из сгенерированных изображений (соответственно содержащих текстовые строки, расположенные под разными углами) в результате применения описанного алгоритма были получены сегменты-результаты локализации и была посчитана оценка точности p_{1i} по формуле 1. Было найдено среднее значение $p_{1cp.} = 0,76$. Далее та же оценка p_{2i} была посчитана для сегментов, полученных в результате описания потенциальных текстовых областей прямоугольниками, параллельными горизонтальной оси (предположив, что информации об угле поворота нет). Среднее значение в этом случае составило $p_{2cp.} = 0,36$.

Выводы. В ходе работы было продемонстрировано, как с помощью двумерного непрерывного вейвлет-преобразования возможна локализация текста под произвольным углом, а также как наличие информации об угле поворота надписи влияет на точность локализации. Полученный результат свидетельствует о том, что при наличии информации об угле поворота, в условиях однородного фона и фиксированного масштаба, система выделяет текст более чем в 2 раза точнее (в сегменте изображения, поступающем в дальнейшем на распознавание, доля нетекстовых участков сокращается в среднем в 2 раза), что позволит избежать определённого количества ошибок на стадии распознавания. В дальнейшем планируется автоматизировать процессы выбора масштаба, подходящего для обнаружения того или иного размера шрифта на изображении, а также адаптировать алгоритм к изображениям с неоднородным фоном.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Болотова Ю. А. Обзор алгоритмов детектирования текстовых областей на изображениях и видеозаписях / Ю.А. Болотова, В.Г. Спицын, П.М. Осина // Компьютерная оптика. – 2017. – Т. 41, № 3. – С. 441-452.
2. Гонсалес Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс. – М.: Техносфера, 2012. — 1104 с.
3. Коноплев А. О. Вейвлет-анализ двумерных изображений / А. О. Коноплев // Вестник МГУЛ – Лесной вестник. – 2007. – №2. – С. 77-82.