

ОБНАРУЖЕНИЕ ТЕКСТА НА ИЗОБРАЖЕНИЯХ С НЕОДНОРОДНЫМ ФОНОМ С ПОМОЩЬЮ АЛГОРИТМА STROKE WIDTH TRANSFORM

Сергеев Я. С.

Научный руководитель – д.т.н., проф. Машталир С.В.

Харковский национальный университет радиоэлектроники
(61166, Харків, просп. Науки, 14, каф. Информатики, тел. (057) 702-14-19)
e-mail: yaroslav.serhieiev@nure.ua

This paper proposes a solution to the text characters detection and symbols extraction problem in images with a non-uniform background using the SWT algorithm. This method allows to recognize and select characters with an accuracy of 85%, regardless of size, font family and language.

В настоящее время очень актуальной является проблема автоматизации распознавания текста. Данная задача решается в оцифровке уже имеющихся печатных текстов, тем самым сохраняя много времени, например, или для перевода текста на изображениях с неоднородным фоном (рекламные афиши, плакаты и т.д.), делая взаимодействие человека с внешним миром проще.

Задача распознавания текста на изображениях с неоднородным фоном состоит из нескольких этапов. На первом этапе на изображении необходимо выделить текстовые области. В данной работе анализ текстовых областей основывается на такой характеристике как толщина штриха символа [1, 2]. Она основана на предположении, что в пределах одной текстовой области (надпись, строка, слово) толщина штриха символа и соотношение высоты/ширины символа неизменно. Метод состоит из следующих этапов: предварительная обработка изображения, SWT-преобразование, фильтрация и объединение компонент. Предварительная обработка изображения происходит следующим образом: исходное изображение преобразовывается в оттенки серого, далее полученное изображение сглаживается фильтром Гаусса апертурой 3×3 либо 5×5 . После сглаживания к полученному изображению применяется детектор границ Канни. SWT-преобразование изображения реализуется по следующему алгоритму. Шаг 1. Выбирается пиксель контурного изображения. Его координаты сохраняются в памяти как P1. В зарезервированном изображении с размытым контуром находится пиксель, соответствующий этим координатам, его окрестности проверяются на наибольшую (если текст светлее фона) либо наименьшую (в противном случае) интенсивность. Координаты этого пикселя также сохраняются в память (как P2). Шаг 2. Проводится луч из пикселя с координатами P1 в направлении пикселя с координатами P2. Если построенный луч пересекает любой необработанный пиксель контура, либо достигает

границы изображения он заполняется пикселями определенной интенсивности, рассчитываемой исходя из длины получившегося отрезка. Шаг 3. Повторяются шаги 1...2 до тех пор, пока каждый пиксель не размытого контурного изображения не будет принадлежать какому-либо построенному лучу. В результате данного этапа обработки получаем так называемое SWT-изображение. При этом текстовые области явно отличаются от любых других областей, что можно увидеть ниже, на рис. 1



Рис 1. Построение SWT-изображения: а - исходное изображение, б - инвертированное контурное изображение, в - SWT-изображение.

Фильтрация полученного SWT-изображения включает в себя проверку контуров на наличие общих точек. Для выделенных символов проверяется пропорция отношения его высоты к ширине, и, если это значение превышает 1.5, то символ исключается из дальнейшего процесса обработки. В конечном итоге получаем объединенные в компоненты символы, готовые к последующему процессу классификации либо распознавания. При этом фон с изображения удаляется, остается только массив текстовых областей. Тестирование показало, что количество обнаруженных и впоследствии выделенных символов находится в диапазоне 75-90% от общего количества символов, однако результат сильно зависит от качества исходного изображения, так как при сильной зашумленности исходного изображения сглаживающий фильтр может размыть контуры символов, сделав их непригодными для дальнейшего распознавания.

Таким образом, предложенный метод позволяет обнаружить и выделить в среднем 85 % символов на изображениях независимо от языка и размера шрифта. Это позволяет использовать его в самых разнообразных областях, тем самым доказывая его эффективность в системах автоматического распознавания текста.

Список використаних джерел:

1. Cong, Y. Detecting Texts of Arbitrary Orientations in Natural Images [Electronic resource]. / [Y. Cong, B. Xiang, L. Wenyu and others]. — Lab of Neuro Imaging and Department of Computer Science, UCLA 2012.
2. Epshtein, B. Detecting Text in Natural Scenes with Stroke Width Transform [Electronic resource] / B. Epshtein, E. Ofek, Y. Wexler. — Microsoft Corporation, 2009.