

Сегментация изображений текстовых полей с фиксированным шаблоном в видеопотоке при распознавании на мобильном устройстве

Б.И. Савельев¹, Д.В. Полевой^{1,2}

¹Национальный исследовательский технологический университет «МИСиС»

²Институт системного анализа ФИЦ УИ РАН

Благодаря быстро развивающимся технологиям, мобильные устройства стали мощными и легкодоступными, что позволяет использовать их в качестве устройства для получения изображения. Неудобство интерфейса для ввода текстовой информации можно компенсировать, например, с помощью технологий оптического распознавания.

При распознавании кадров видеопотока, основными этапами являются: получение и первичная обработка изображений, локализации документа с учетом ориентации, наведение полей, распознавание текстовых строк и межкадровое интегрирование результатов распознавания.

Использование в качестве источника изображений камер мобильных устройств порождает ряд проблем: эффекты освещения (тени, блики и т. п.) [1]; шум печати, в частности, непропечаткой (разрывами слитных черт символов), "слипанием" соседних символов, пятнами, ложными точками на фоне и т. п.

В работе рассматривается задача сегментации на отдельные символы проективно исправленных изображений полей документа. Рассмотрены существующие подходы нахождения порога бинаризации [2], выравнивания фона изображения и "обжимания" символов по порогу бинаризации. В работе предложен комбинированный алгоритм сегментации полей документа для полутонового изображения. Данный алгоритм был успешно использован для сегментации полей в реализованной на практике системе оптического распознавания этикеток при помощи мобильных устройств.

Литература

- [1] Арлазаров В.Л., Корольков Г.В., Славин О.А. Линейный критерий в задачах OCR // В сб. трудов ИСА РАН "Развитие безбумажных технологий в организациях", М.: Эдиториал УРСС, 1999, С. 28-46.
- [2] Otsu N. A Threshold Selection Method from Gray-Level Histograms // IEEE Transactions on Systems, Man, and Cybernetics. 1979. Vol. 9. № 1. P. 62–66.