

Отзыв

на выпускную квалификационную работу Чэнь Юйцзин
«Распознавание встроенных субтитров в видеопотоке»

В настоящее время видеоконтент в сети Internet снабжается, как правило, внешними субтитрами. Эти субтитры накладываются на видеопоток непосредственно при его воспроизведении. Однако существует достаточно много видео, в которых субтитры являются неотъемлемой частью этого видео. Эти субтитры называются встроенными. При обработке таких видео встает задача извлечения и распознавания этих субтитров с целью их дальнейшего использования. Поэтому работа Чэнь Юйцзин, посвященная разработке метода распознавания внедренных субтитров, представляется достаточно актуальной.

Задача распознавания субтитров, решается в работе в два этапа. На первом этапе необходимо обнаружить на изображении области текста, относящегося к субтитрам, а на втором этапе — распознать локализованное изображение субтитров с помощью какой-либо системы распознавания текста.

Чэнь Юйцзин провел анализ методов локализации субтитров на изображениях. На основе этого анализа им был предложен метод локализации субтитров, учитывающий особенности их изображений в видеопотоке. Этот метод извлекает из видеопотока отдельные кадры. Изображение извлеченного кадра сначала размывается для подавления шумов для подавления высокочастотных шумов. После чего это изображение бинаризуется с помощью адаптивного порога и на нем ищутся контурные точки. Для выделения областей на контурном изображении к нему применяется морфологического расширения. Локализация областей осуществляется с помощью метода связных компонент. Поскольку центры областей, относящиеся к субтитрам, обладают определенной периодичностью в горизонтальном направлении и концентрируются в определенной области по вертикали, то их можно отфильтровать по предложенному в работе алгоритму. В результате фильтрации выделяются области субтитров. Для последующего распознавания на вход распознающего текст алгоритма, передаются изображения выявленных областей. Для временной привязки субтитров к кадрам изображения Чэнь Юйцзин использует тот факт, что изображение каждого субтитра присутствует в нескольких последовательных кадрах. Алгоритм, предложенный автором, сравнивает изменение числа контурных точек в области субтитров на двух последовательных кадрах. Если это изменение превышает некоторый порог, то считается в этом кадре субтитр закончился. Таким образом формируется временные метки существования каждого субтитра.

Чэнь Юйцзин реализовал предложенный метод на языке Python с использованием библиотеки обработки изображений OpenCV, а также библиотеки распознавания символов PyTesseract. Экспериментальное исследование реализованного метода проводилось на относительно коротких видео с различными типами внедренных субтитров. При этом определялись – доли пропущенных строк, точность локализации и точность распознавания. Результаты работы алгоритма сравнивались с результатами работы известного приложения Videocr, выполняющего те же функции. Результаты сравнения показали, что предложенный автором алгоритм смог достичь более высоких показателей.

За время подготовки работы Чэнь Юйцзин, продемонстрировал хорошую теоретическую подготовку, показал умение самостоятельно разбираться в методах обработки изображений, а также умение творчески применять полученные знания на практике с использованием современных программных средств.

Считаю, что работа Чэнь Юйцзин заслуживает оценки “Отлично”.

Руководитель,
Доцент кафедры КММС, к.т.н.



В.М. Гришкин