

РАСШИРЕНИЕ К БРАУЗЕРУ ДЛЯ ИЗВЛЕЧЕНИЯ ТЕКСТА ИЗ ИЗОБРАЖЕНИЙ

Рассматривается плагин Project Naptha, позволяющий считывать текст с картинок, открытых в окне браузера.

ВВЕДЕНИЕ

Сегодня информация может быть представлена как в текстовом виде, так и в виде изображений и мультимедийных файлов. Извлечение информации из таких файлов является одним из приоритетных направлений в области разработки в сфере распознавания образов.

I. ПЛАГИН ПРОЕКТ НАПТНА

Основная идея разработанной игры, заключается в поиске и «захвате» специальных точек на карте, а так же Project Naptha — расширение для браузера Google Chrome, для автоматического распознавания текста на изображениях, загружаемых в браузер. Расширение представляет собой систему оптического распознавания символов, реализованную на языке JavaScript.

Алгоритм работы данного расширения следующий. На первом этапе определяется местоположение блоков с текстом на изображении. Для реализации этого механизма используется проект компании Microsoft – Stroke Width Transform (SWT). Алгоритм выделения текста, основан на том, что за текст принимается группа объектов, имеющих примерно одинаковую толщину линий (рис. 1). Достоинством использованного алгоритма является то, что способен распознавать текст независимо от его начертания.

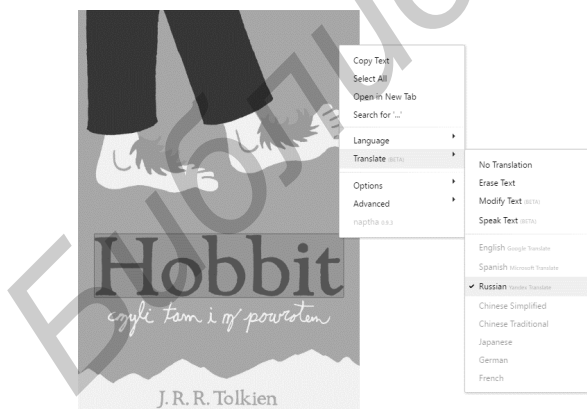


Рис.1. Контекстное меню расширения

После выбора блока текста на картинке и выбора в контекстном меню пункта «Copy Text»,

этот блок посылается на сервер с Ocrad OCR — прикладной сервис с открытым кодом для распознавания текста на изображениях. Данный ресурс позволяет распознать текст, находящийся в выделенном блоке. Полученный текст можно корректировать и использовать для других целей, например для перевода (рис. 2).



Рис.2. Результат работы команды «Translate»

Для удобства пользователя, расширение распознает только те изображения на открытой странице, на которые наведен указатель мыши.

II. ВЫВОД

Приложение имеет высокое быстродействие и при работе в фоновом режиме не оказывает влияние на скорость работы браузера. Запуск расширения происходит только при наведении указателя мыши на картинку, что помогает экономить ресурсы компьютера по сравнению с полным сканированием всех картинок при загрузке страницы. Достоинством приложения является то, что оно позволяет не только распознавать текст на изображениях в браузере, но также предоставляет возможность редактировать, удалять и переводить текст непосредственно на изображении.

1. Project Naptha – Режим доступа: <https://projectnaptha.com/>
2. Расширение к браузеру для извлечения текста из изображений – Режим доступа: <https://habrahabr.ru>

Янчук Иван Сергеевич, магистрант кафедры ИТАС.

Научный руководитель: Навроцкий Анатолий Александрович, заведующий кафедрой автоматизированных систем обработки информации, кандидат физико-математических наук, доцент, navrotsky@bsuir.by.