

РОЗПІЗНАВАННЯ ДРУКОВАНОГО ТЕКСТУ З КАРТИНКИ І КОНВЕРТАЦІЯ ЙОГО В ДРУКОВАНИЙ ФОРМАТ

Сіденко Д.А.

Науковий керівник – доц. Вовк О.В.

Харківський національний університет радіоелектроніки
(61166, Харків, пр. Науки, 14, каф. Медіасистем та технологій,
тел. (057) 702-13-78)

e-mail: daryna.sidenko@nure.ua, тел. (050) 010-88-37

Typing a lot of text can take a long time, even if it's just automatic reprinting of information, for example, from a picture, when there is a photo of printed text material. Therefore, quite often there is a need to automate this process. This article will discuss how to automatically recognize printed text from a picture and convert it to print format.

Набір великої кількості тексту може зайняти досить багато часу, навіть якщо це всього лише автоматичний передрук інформації. Тому досить часто виникає необхідність автоматизувати цей процес. Розглянемо автоматичне розпізнавання друкованого тексту з картинки і конвертацію його в друкований формат.

Картинки або фото обробляються для того, щоб текст, відображений на них, автоматично був переведений в текстовий формат. Технічно процес відбувається наступним чином: користувач завантажує картинку на сервер, або переносить її в програму, софт обробляє зображення, використовуючи особливі алгоритми, і видає у вигляді файлу або у вікні програми сфотографований текст в друкованому вигляді.

В даний час розроблено досить багато таких різноманітних програм. Вони відрізняються за функціоналом зовсім незначно, але можуть істотно відрізнятися за якістю обробки. Деякі програми допускають досить багато помилок в розпізнаванні тексту, тоді як інші – розпізнають все практично ідеально. Якість розпізнавання залежить від початкового якості фото, але за інших рівних умов більшу роль відіграють алгоритми роботи і широта бази, яка використовується або онлайн-сервісу. Кожна програма здатна працювати тільки з тими символами, які були занесені в її базу. Якщо мова йде про не надто популярні і візуально нестандартні мови, то знайти відповідний софт може бути непросто. Чим складніше форматування або розташування букв на фотографії, тим складніше програмі правильно розпізнати текст, і тим більше буде помилок. Неточності можуть виникнути на стадії визначення місця розташування друкованих знаків на зображенні.

При визначенні букви програма використовує певний «алгоритм» порівнянь з її основними рисами – розташуванням і розміром елементів (деякі утиліти також враховують сусідні розпізнані букви і лексичну сполучуваність). Завдяки цій особливості, навіть якщо невелика частина букви стерлася або змінена, вона все ще може бути розпізнаною.

Найкраще визначається текст з відсканованих зображень документів, скріншотів. Більш-менш нормально може бути визначений і сфотографований на камеру текст. Найгірше розпізнаються матеріали з фото поганої якості, зробленого під кутом, особливо якщо має місце складне форматування. Художні шрифти не розпізнаються.

Переваги використання таких сервісів очевидні. Основна з них – це суттєва економія часу на обробку матеріалу. І хоча текст, одержуваний на виході, може мати дуже низьку якість, редагування його все одно займе, в більшості випадків, менше часу, ніж передрукування «з нуля».

Переваги:

- значна економія часу, що минає на автоматичну передрук тексту;
- уникнення помилок у формулах і складних символічних побудовах, які можуть бути при ручному передруці матеріалу;
- розпізнавання і перенесення тексту, який ви не можете надрукувати на своїй клавіатурі (наприклад, текст з арабською в'яззю, ієрогліфами і іншими символами, які відсутні на традиційній російсько-англійській клавіатурі).

Недоліки:

- обмеженість бази мов;
- великі складності бувають при роботі з текстом зміщеного типу, тобто таким, який містить як російські, так і англійські символи.;
- втрата форматування або нездатність розпізнати особливості початкового оформлення тексту;
- просто низька якість розпізнавання, пов'язане з невірною налаштуванням або неякісною опрацюванням самого софта;
- помилки розпізнавання, пов'язані з початковим низькою якістю фото;
- серйозні проблеми з розпізнаванням виникають у випадках, коли текст розташований блоками, розподіленими по фото нерівномірно або навіть в дві колонки - краще за все цей софт розпізнає суцільний текст;
- якість розпізнавання може падати у міру додавання все нових і нових фото за один сеанс;
- іноді процес обробки зображення може йти дуже довго, так як його швидкість падає в міру збільшення тривалості навантаження на програму;
- стандартне для майже всіх програм, дуже низька якість розпізнавання «складних» алфавітів, наприклад, ієрогліфів або арабської в'язі. Помилки, причому досить великі, в цьому випадку неминучі;
- неправильне розпізнавання символічних груп.

В даний час практично не існує таких програм, які працюють дійсно якісно. Текст, що отриманого після обробки фото в такому редакторі, все одно потрібно ретельно редагувати. І така редактура може займати не менше часу, ніж початковий передрук. Найбільш серйозною проблемою вважається низька якість розпізнавання тексту в цілому. Так як саме через нього доводиться виконувати тривалу вичитку і редактуру матеріалу.