

Приложение на OCR софтуер с отворен код при разпознаване на графични символи от машиностроителни чертежи

Красимир Каменов

OCR application of open source software in recognizing graphic symbols of engineering

drawings: OCR techniques typically used for the processing of text documents on the paper, resulting in the formation of text file needed for future processing. In this paper explored the use of OCR method for processing graphic documents representing drawings of machine parts used in mechanical transmissions and gearboxes.

Key words: *mechanical drawing, object recognition, optical character recognition (OCR), engineering drawing*

ВЪВЕДЕНИЕ

OCR е автоматизиран метод за конвертиране на текстов документ от хартиен носител до компютърен, най- често текстов файл, даващ възможност за по-нататъшна корекция, съхранение и публикуване. При повечето OCR системи първоначално хартиеният документ се сканира, а полученият файл е растрерно изображение в някой от популярните формати: PDF, JPEG или BMP.

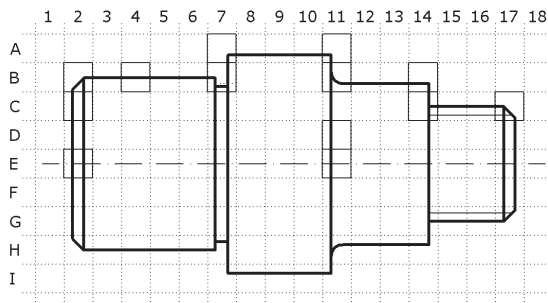
В настоящият доклад се разглежда възможностите за използването на OCR метод в необичайно приложение- за обработка не на текстови, а графични документи, представляващи чертежи на машиностроителни детайли, които се използват в механични предавки и редуктори.

ПРОБЛЕМ И РЕШЕНИЕ

Обикновено OCR методите се използват за обработка на текстови документи върху хартия, в резултат на което се получава текстов файл, необходим за бъдеща обработка(корекции на текста, съхраняване като текстов или подобен файл, разпечатване). Идеята на настоящия доклад е да се проведе експеримент и проверят резултатите от използването на OCR софтуер за разпознаване на специфични геометрични елементи от сканирани чертежи на детайли, използвани в редукторостроенето – валове и зъбни колела. За да се осъществи процеса на разпознаване е необходимо да се създаде език с който да се опишат отделните графични символи, получени в резултат на подходящо направена сегментация.

За осъществяване на експеримента е използван софтуер от проекта *Tesseract*, който представлява свободната версия на бившия комерсиален продукт за разпознаване на текст на фирмата *Hewlett Packard*. Изборът на софтуера е направен от лицензионни съображения (лиценз *Apache 2.0*) и наличието на възможност за обучение в нов език, който е една от целите на настоящата работа.

На Фиг. 1 е показано опростено изображение на вал, сегментирано посредством мрежа с квадратни клетки. Редовете и колоните на мрежата са озаглавени за улесняване на описанието на отделните клетки от мрежата. Клетките, оградени с непрекъснати линии(квадрати) представляват уникалните геометрични елементи (символите), от които чрез подходящи връзки е съставено цялото изображение на вала. Общият брой на клетките е $9 \times 18 = 162$, от които запълнените със символи са 72. Броят на уникалните символи е 13, от които посредством подходящо комбиниране се получава цялото изображение. Тези графични символи представляват елементарните геометрични обекти, които ще се използват при разпознаването на чертежи на ротационни детайли тип вал.



Фиг. 1

В

Табл. 1 са показани част от символите, съставлящи езика за геометрично описание на ротационни машинни детайли. Те са подбрани след подробно изследване на различно сегментирани изображения в зависимост от големината на клетките на мрежата и разположението и спрямо изображението на няколко конструкции на валове и оси. Промяната в големината на клетките от мрежата може да бъде използвана за сегментация на няколко нива в зависимост от крайните цели на процеса. В най-лявата колона от таблицата са дадени означенията на съответния графичен символ, където индексът съвпада с поредния номер на символа в групата символи, различаващи се по ъгъла на ротация. Всеки от ъглите в градуси, показани на първия ред, представлява ъгъл между равнините обхващащи материал в дадения символ, както е показано на Фиг. 2 символа А1, при който материалът е щрихован частично.



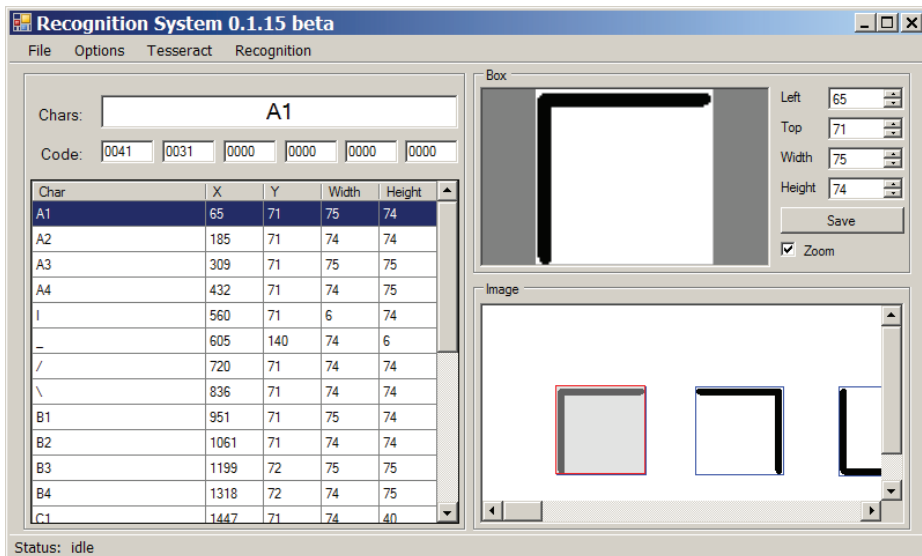
Фиг. 2

Табл. 1

Означение\Ъгъл	90°	180°	270°
A1..A4	Г Г Г Г		
B1..B4	Г Г Г Г		
C1..C4		Г Г Г Г	
D1..D2			Г Г
E1..E4			Г Г Г Г

Разположението на мрежата е от голямо значение за формата на символите, които всъщност представляват отделни изрязани участъци от цялостното изображение. Тя определя и необходимостта от по-голям обхват на процеса на обучение, за да се осигури разпознаване на символи с форма различаваща се от показаната в таблицата, но близка по своята геометрична същност. Разликата е продиктувана от случайното или подбрано и оптимизирано в някаква степен

разположение на сегментиращата мрежа върху изображението. В зависимост от сложността на алгоритъма за сегментиране, може да се получи голям набор от различаващи се символи, които от своя страна удължават процеса на обучение и предизвикват нарастване на базата от геометрични и други характеристики, използвани за окончателното разпознаване на конструктивните и технологични елементи на конкретно разглеждания детайл.



Фиг. 3

Генерираният от Tesseract .box файл, визуализиран от показаната на Фиг. 3 програма и необходим при по-нататъшните процедури за разпознаване може да се види в лявата част на прозореца. Този файл представлява текстово описание на геометричните параметри на символите- координатите X и Y на горния ляв ъгъл, , ширината и височината на квадратчето(бокса). В първата колона на таблицата вляво(DataGridView) е условното наименование на символа, необходимо за текстовото му описание. Разработеният за тази цел софтуер дава възможност за уточняване на тези параметри, водеща до постигане на по- добра представителност на дадения символ, необходима при по-нататъшното обучение. Показаната извадка от символи е примерна и илюстрира принципите за възможна употреба на Tesseract при разпознаване на нетекстови символи, в случая елементарни графични обекти, съставляващи геометрията на ортогонална проекция, представяща ротационен детайл от тип вал. Конструкцията на вала е умишлено опростена, като са избегнати шпонкови канали, шлицы, центрови отвори, освободения и др., които обикновено присъстват в реално изображение на вал.

Процедурата по разпознаване продължава с обучение, което трябва да обхване възможно по-голям брой сегментирани чертежи, при което се обогатява базата данни от възможните символи като геометрични модификации. Процесът на обучение не е предмет на настоящия доклад. Обучението води до повишаване на процента на вярно разпознатите символи от различни чертежи, начертани с различни размери и различно качество.

Разработеният софтуер, е не само графичен интерфейс на Tesseract, пригоден за разпознаване на нетекстови символи, но и цялостна система за разпознаване на

съставните елементи на проекции на машиностроителни детайли. Предназначен е за работа с ОС Windows 7 и XP, с инсталиран .Net Framework 3.5. Разработката му е осъществена посредством интегрираната среда с отворен код SharpDevelop 4.3. Графичната визуализация се реализира чрез MagickNet, библиотека с отворен код, предоставяща възможност за работа с над 100 файлови формата. Разработени са както класове, обслужващи графичното представяне на символите, съставлящи конкретния език, така и класове обслужващи сегментирането и разпознаването на елементарните геометрични обекти, на които се разгражда цялостното изображение. Постигната е сравнително добра производителност на системата, която би могла да се подобри чрез усъвършенстване на някои класове.

ЗАКЛЮЧЕНИЕ

Проведените с помощта на разработеният софтуер експерименти по създаване език за описание на геометрични елементи от машиностроителни чертежи на ротационни детайли доказва възможността за употреба на OCR софтуер в необичайна за него роля. Използването на Tesseract за тази цел обаче налага разработване на допълнителни софтуерни решения, включващи графичното представяне на чертежите и съдържащите се в тях елементи. Самият механизъм за разпознаване зависи от правилното сегментиране, където се налага разработката и приложението на нови и усъвършенствани алгоритми. Пряката употреба на Tesseract за целите на разпознаване на машиностроителни чертежи, без допълнителните програми е невъзможно.

ЛИТЕРАТУРА

- [1] Bunke H., P.S.P Wang. Handbook of Character Recognition and Document Image Analysis. Singapore, Uto-Print: World Scientific Publishing Co., 2000.
- [2] Bishop C., Pattern recognition and machine learning, Springer, 2006.
- [3] Ray S., Tesseract OCR Engine, Google Inc, OSCON, 2007.

За контакти:

Гл. ас. Красимир Каменов, Катедра "Машинознание, машинни елементи и инженерна графика", Русенски университет "Ангел Кънчев", тел.: 082-888 461, e-mail: kkaменов@uni-ruse.bg

Докладът е рецензиран.