

Методи за конвертиране „растер-вектор” при графични документи- предимства и недостатъци

Красимир Каменов

Methods of raster to vector conversion for graphic documents - advantages and disadvantages:

Preliminary vectorization subprocess is performing rough handling binary raster image and leads to segments and polylines with nonzero thickness This paper is precisely this sub-process and methods used to carry it and aims to help the selection of appropriate methods, meeting the specific requirements of the developed recognition system design.

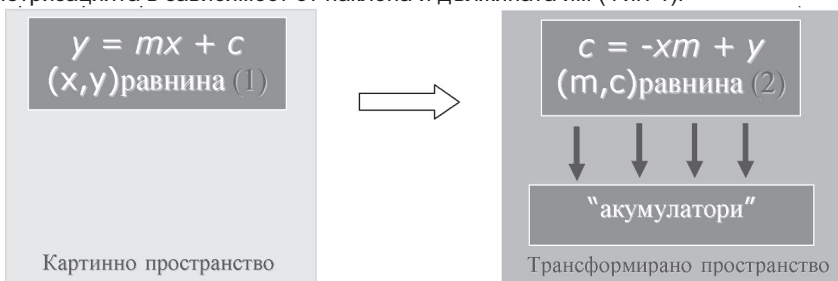
Keywords: vectorisation, recognition system

ВЪВЕДЕНИЕ

Векторизацията е процес на преобразуване на растерно изображение във векторно. Растерното изображение е съставено от пиксели, а векторното от насочени линейни сегменти. Този процес се използва широко както в системите за разпознаване на текстови документи (OCR), така и при разпознаване на графични обекти. Предварителната векторизация е подпроцес, който извършва грубата обработка на бинарното растерно изображение и води до получаване на отсечки и полилинии с ненулева дебелина. Предмет на настоящия доклад е именно този подпроцес и методите, които се използват за осъществяването му. Поради развитието на изследванията и експериментите в тази област, през последните 20 години се появиха голям брой методи за предварителна обработка, които за улеснение на класификацията им, могат да се разделят на следните обобщени групи- методи, използващи трансформации на Хаф (Hough), контурни методи, изтъняващи методи, методи на подвижния граф (run-graph) и методи на решетката от шаблони. Почти всички от разработените досега методи и съответстващите им алгоритми могат да се причислят към изброените групи. В основата си се състоят се от три последователни етапа- претърсване за характерни точки, трасиране на получените точки до сегменти и апроксимация за получаване на полилинии (полигонизация). Основната разлика между класифицираните по-горе групи се крие в първите два етапа. Настоящият доклад има за цел да помогне за избора на подходящи методи, удовлетворяващи специфичните изисквания на разработваната система за разпознаване на технически документи.

Методи, базирани на Хаф (Hough) трансформации

Най-простият случай за трансформации на Хаф е линейно преобразование за откриване на прави линии, а в по-сложните се търсят нелинейни обекти като окръжности и елипси или части от тях. Един от начините, при който е възможно приложението на Хаф трансформации за разпознаване на отсечки е параметризацията в зависимост от наклона и дължината им (Фиг. 1).



Фиг. 1

На всяка права от (x,y) равнината съответства една точка от (m,c) равнината. Равнината (m,c) се разделя на правоъгълни участъци - "акумулатори", които „акумулират“ за всеки черен пиксел от (x,y) равнината всички пиксели, лежащи по дължината на линията от (2). Когато линията от (2) е начертана за всеки черен пиксел, клетките на които съответства, нараства. След изчисляване за всички пиксели в картинното пространство, линиите се разпознават като пикове (върхове) в трансформираното пространство. Като се вземе предвид шума, всеки връх, по-висок от предварително избрана стойност, определя линия по формула (1).

Тъй като върховете се формират в (m,c) равнината за точки, които принадлежат на прекъснати или зашумени линии от оригиналното изображение, методът може да се използва за разпознаване на линии в леко зашумени изображения. Въпреки това, тъй като ъгловият коефициент и отрезът на правата се изчисляват през определено разстояние, прецизността не може да бъде колкото на първоначалното изображение, затова качеството на разпознатите линии е много по-точно за наклонени линии.

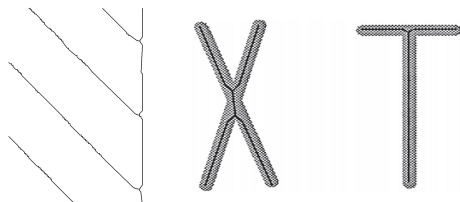
Изтъняващи методи

Тези методи използват процедура за изтъняване като подпроцес при откриване на точки от средните оси за да се получи скелет с еднопикселова дебелина преди да се извърши следващия трасиращ подпроцес.

Изтъняването [4] прилага набор морфологични операции - от входното растерно изображения се получава скелет, чиято топологична структура е идентична с оригиналното изображение и може много по-лесно да се обработва и анализира от първоначалното изображение. Скелетът се дефинира като формиран от центровете на окръжности с максимален радиус, разположени в обектите. На основата на така дефинираното понятие за скелет, изтъняващите алгоритми се разделят на три групи - итерационна ерозия на границите, трансформация на разстоянията и адекватен скелет.

Итерационните изтъняващи методи използват идеята за постепенно свиване на контура (или премахване на външния слой пиксели) на линейните обекти, като челото на вълната се движи отвън към вътрешността докато останат само пикселите на скелета. Изтъняването, разгледано на ниво пиксел, представлява итерационна ерозия на границите на линейния обект. Главна процедура е придвижването на прозорец с размери 3×3 пиксела върху изображението и намиране с помощта на правила и ограничения положението на центъра на прозореца. В резултат на всяко сканиране отбелязаните точки се премахват. Процесът се повтаря докато няма повече пиксели за отстраняване.

Главен проблем на алгоритмите за изтъняване е времето за обработка, което е кубична зависимост от резолюцията на изображението. Освен това тези алгоритми водят до изкривявания при X и T-образните пресичания на линиите, както е показано на фиг. 2.



Фиг. 2

Дистанционната трансформация на бинарното изображение се дефинира чрез заместване на всеки пиксел с число, показващо минималното разстояние до бяла точка. Разстоянието между две точки се дава с броя пиксели по най-късата 4-свързана верига между двете точки. След като е намерена дистанционната функция се използва локалния и максимум за намиране на скелета. Той би трябвало да бъде най-малкия набор от точки, за да може да се реконструира изображението точно. Времето необходимо за обработка е квадратна функция от резолюцията на изображението, което е много по-малко в сравнение с кубичната функция на итерационните алгоритми.

Комбинирайки точките от скелета с тези получени чрез итерационните алгоритми на изтъняване се дефинира т.нар. „адекватен скелет“. Комбинацията довежда до получаване на скелет с максимална дебелина 2 пиксела. След това се изтънява до необходимата дебелина от 1 пиксел. Алгоритъмът получава като резултат по-точни скелети в сравнение с обикновените изтъняващи итерационни алгоритми, но за сметка на увеличеното време за допълнителни обработки.

Като цяло, целта на изтъняването е да се намали обемът от данни, така че да остане само топологичната форма, която да е инвариантна спрямо ориентацията и размера. Полученият резултат обикновено изисква и допълнителна обработка. Повечето алгоритми за изтъняване предоставят възможност за поддържане на свързаност. Въпреки това, основните им недостатъци са голямата сложност на изчисленията, загуба на част от информацията (дебелината на линиите), изкривяванията в пресечните точки и др. Те биха могли да се използват и при разпознаване на технически чертежи, но основното им приложение е в областта на OCR, където дебелините на линиите не са от значение и размерите на символите са относително малки.

Скелетите, произведени от изтъняващите процедури са все още в растерен формат и трябва да се векторизират чрез по-нататъшна обработка. Еднопикселивият скелет се свързва във верига от трасиращ подпроцес. След това се прилага полигонизация, която превръща веригата в полилиния, която съдържа само критичните точки.

Контурни методи

С цел понижаване изчислителната тежест при изтъняващите алгоритми, друга група от алгоритми се опитват да намалят обема от данни преди откриване на медиалните оси. Основната идея е да се намери формата - ръбовете (или контура) на линейния обект, първо, а след това да се намерят средните точки на двойка точки от два паралелно противоположни ръба [3]. Тези алгоритми трасират медиалните точки едновременно. Това е различно от алгоритмите за изтъняване, при които започва трасирането едва след като са открити всички медиални точки. Основните интензивни изчислителни операции са разпознаването на ръбовете и полигонизацията. Разходът на време за разпознаване на ръб е линейна зависимост от броя пиксели или квадратна от резолюцията на изображението, т.е. е по-бърза от алгоритмите за изтъняване. Освен това може много лесно да бъде получена дебелината на линията, което е важно за по-нататъшното тълкуване на чертежа.

Проблеми при всички разновидности на тези алгоритми възникват при обработката на пресечните точки. Изследванията върху този метод доказват, че не е подходящ за употреба при криви линии или такива с многократно пресичане.

Методи, базирани на движещия се граф (run graph)

Методът на движещия се граф [2] се дефинира като полуекторно представяне на растерното изображение преди разпознаването на линии и сегментацията. Счита се, че това е достатъчно за структурно представяне на линейни обекти, ефективно за извличане на линия, запазване на информацията и лекота при обработката.

Движението има посока, която може да бъде вертикална или хоризонтална. Посоката съвпада с максималната последователност от черни пиксели. Движението се дефинира с четворката

$$R = \{d, c_d, b_d, e_d\}, b_d \leq e_d$$

Където d е посоката на движение (0 за хоризонтална, 1- за вертикална), c_d - ортогонална координата, т.е. координата на движение перпендикулярно на настоящото или номер на реда при хоризонтално или колоната при вертикално, b_d е начална координата на движението, а e_d е крайната координата на движението.



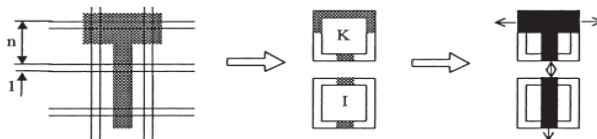
Фиг. 3

Няколко случая на представяне на движещ се граф са показани на фиг. 3. R1 е пример за хоризонтално движещ се граф, R2 за вертикално. Движенията A и B са съседни, когато имат една и съща посока, разликата от ортогоналните им координати е 1 и разликата между максималната начална координата и минималната крайна е ≤ 1 .

Предимство на метода е еднократното разглеждане на всеки пиксел, водещо до квадратна зависимост на времето за обработка от резолюцията. Като недостатък може да се изтъкне затруднението при обработка на пресечните точки, водещо до грешната им векторизация при смяната на посоката на движение на графа или поради зашуменост на изображението. Този метод също е неподходящ за обработка на криви линии.

Методи, базирани на мрежата от шаблони

Методът е предложен най-напред за разпознаване на специфични фигури (символи) в диаграми [1]. Основната идея е да се раздели даденото изображение с помощта на предварително зададена мрежа и да се разпознае характерна фигура само чрез проверка на разпределението на черните пиксели по границата на всяка клетка от мрежата (Фиг.4). Посредством тези фигури се създава контролна карта. Накрая се извършва извличане на линейни сегменти чрез анализ на така получената контролна карта.



Фиг. 4

При този метод е много сложно да се избере подходящ размер на клетките от мрежата. Големите размери предизвикват появата на много неразпознати клетки, които изискват повече време за обработка. Обратно, намаляването на размерите на клетките също предопределя сложност на извличането на линии. Това също води и

до затруднения при разпознаването на по-сложни обекти като дъги и прекъснати линии. Поради това, че се изследват само границите на клетките от мрежата, могат да се получат некоректни свързвания на прекъснати линии, в случаите, когато прекъсването е по-малко от размера на клетката.

ЗАКЛЮЧЕНИЕ

Векторизацията е основен процес при преобразуването на растерни изображения във векторни, който се прилага както при графични, така и при текстови документи. При машиностроителните чертежи от особена важност е запазването на дебелината на линиите като важен атрибут на отделните геометрични обекти и коректното представяне на точките на пресичане. Тази информация е от съществено значение за по-нататъшния анализ и обработка. Въпреки големия брой методи и алгоритми за реализацията им, няма универсално решение за всички възможни случаи, които могат да възникнат при извличането на информация от машиностроителни чертежи. Практическото използване на разгледаните методи трябва да бъде съобразено със специфичните изисквания във всеки конкретен случай, като много често се налага комбинирането им.

ЛИТЕРАТУРА

[1] Lin X., Shimotsuji S., Minoh M., Sakai T. Efficient diagram understanding with characteristic pattern detection. Computer Vision, Graphics and Image Processing, 1985.

[2] Boatto L et al. An Interpretation System for Land Register Maps. IEEE Computer, 1992.

[3] Jimenez J, Navalon JL. Some experiments in image vectorisation. IBM Journal of Research and Development, 1982.

[4] Tamura H. A Comparison of line thinning algorithms from digital geometry viewpoint. Proceedings of the 4th International Conference on Pattern Recognition, Kyoto, Japan, 1978.

[5] Dori D. Orthogonal Zig-Zag: an Algorithm for Vectorizing Engineering Drawings Compared with Hough Transform. Advances in Engineering Software, 1997.

За контакти:

Гл.ас. Красимир Каменов, Катедра “Машинознание, машинни елементи и инженерна графика”, Русенски университет “Ангел Кънчев”, тел.: 082-888 461, e-mail: kkamenov@ru.acad.bg

Докладът е рецензиран.