

도면 자동 벡터화를 위한 선의 굵기 인식이 가능한 세선화의 전처리 기법

전일수¹ · 원남식¹ · 부기동¹

A Preprocessing Scheme of Thinning Capable of Lines' Thickness Recognition for the Automated Vectorizing of Maps

Ilsoo Jeon¹ · Namsik Won¹ · Kidong Bu¹

요 약

세선화 결과에서 원래 선의 굵기 정보는 도면 자동 벡터화 시스템을 구현하는데 유용하게 사용될 수 있다. 본 연구는 세선화 결과에서 원래 선의 굵기 정보를 표현할 수 있는 세선화의 전처리 기법을 제안하였다. 제안된 기법에서는 입력 도면에서 선을 구성하는 각 화소들에 대해서 그것이 주변화소들로부터 둘러싸인 정도를 나타내는 깊이를 계산하고, 세선화 결과에서 골격선의 깊이 정보를 보고 원래 선의 굵기를 알 수 있게 하였다. 제안된 기법을 구현하여 등고선에 대해 실험한 결과, 골격선의 깊이 정보로부터 그 선이 주곡선인지 계곡선인지 쉽게 구별할 수 있었다.

주요어: 세선화, 세선화의 전처리, 도면 벡터라이징, 선굵기

ABSTRACT

Information representing the thickness of the original lines from the thinning results can be used efficiently in order to implement the automated vectorizing system. This paper proposes a preprocessing scheme of thinning which can show the information of the original lines' thickness from the thinning result. In the proposed scheme, the depth of each pixel constructing the lines was calculated, which was represented by the number of layers composed of neighboring pixels surrounding the original pixel. And then the original lines' thickness could be recognized through the depth information of the skeleton from the thinning results. We implemented the proposed scheme and experimented on a contour map. Using the depth information of the skeleton, we could easily distinguish each line of the contour either an intermediate or an index contour.

KEYWORDS: Thinning, Preprocessing of Thinning, Vectorizing of Maps, Line Thickness

1999년 5월 30일 접수 Received on May 30, 1999

¹ 경일대학교 컴퓨터공학과 (isjeon@bear.kyungil.ac.kr, nswon@bear.kyungil.ac.kr, kdbu@bear.kyungil.ac.kr)
Dept. of Computer Engineering, Kyungil University

서 론

GIS는 지리적 자료를 수집하여 입력, 저장, 분석, 출력을 할 수 있는 컴퓨터 응용 시스템으로, 최근 이의 응용이 날로 증가함에 따라 이에 대한 관심이 고조되고, 또한 여러 분야에서 GIS와 관련된 연구 및 제품들이 발표되고 있다.

GIS를 구축하기 위해서 가장 많은 시간과 노력을 필요로 하는 것은 도면 정보의 입력이다. 도면 정보의 입력을 위해 각종 주제도를 스캔한 래스터 영상으로부터 자동 혹은 반자동으로 벡터화하는 시스템을 사용하면 효율적으로 정보를 입력할 수 있다. 자동 벡터화 시스템에서는 벡터화에 필요한 정보의 양을 줄이고, 또한 형태 분석을 쉽게 하기 위해 벡터 정보를 생성하기 전에 대개 세선화 과정을 거친다.

기존의 세선화 알고리즘의 입력으로는 주로 이진 영상이 사용되며, 세선화 결과 또한 이진 영상으로 나타나므로 그것으로부터 입력 시 선의 굵기를 알 수 없다. 그런데 세선화 결과에서 선의 굵기 정보가 나타난다면, 그것은 자동 벡터화 시에 유용한 정보가 된다. 예를 들어, 등고선을 세선화한 결과에서 등고선의 굵기 정보가 나타나면, 그것으로부터 그 등고선이 주곡선인지 계곡선인지 쉽게 알 수 있다. 그리고 행정경계와 도로망 등이 같이 표현된 주제도에서 그것들을 표현하는 선의 굵기가 다르다면, 세선화 결과에서 선의 굵기 정보를 이용하여 자동으로 행정 경계와 도로망을 분리해 낼 수 있다.

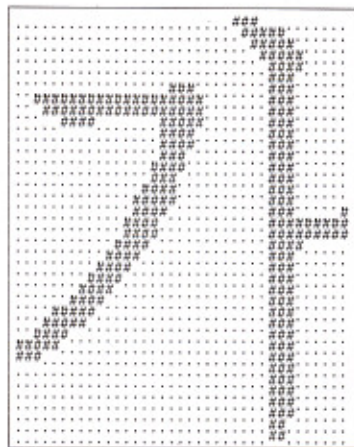
본 연구는 세선화 결과에서 선의 굵기 정보를 표현하기 위한 방안으로 기존의 세선화 알고리즘에 전처리 형식으로 한 과정을 추가하여 이를 가능하게 하기 위한 기법을 제안한다. 제안된 기법에서는 선을 구성하는 각 화소들에 대해서 그것이 주변화소들로부터 둘러싸인 정도를 먼저 계산하고, 그 정보를 보존하여 세선화가 종료될 때 결과로 나타나게 한다. 이와같이 제안된 방법을 구현하여 실험한 결과

를 고찰하고 결론과 함께 향후 연구 방향을 제시하였다.

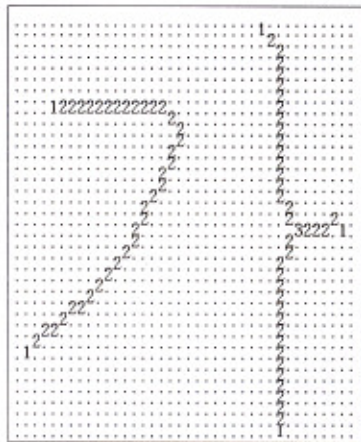
세선화

세선화는 입력 영상의 형태를 잘 반영하면서 특징 추출에 별로 영향을 주지 않는 화소를 제거하여 두께가 1인 화소로써 그 특징 골격을 형성하는 작업을 말한다. 세선화는 입력 영상으로부터 구조적인 특징 추출을 위한 전처리 기법으로 많이 이용되어 왔다. 이는 세선화를 함으로써 처리할 정보의 양이 줄고 또한 패턴 인식이나 분석이 훨씬 쉬워지기 때문이다. 지리 도면 자동 벡터화를 위해 세선화를 할 경우, 세선화에 의해 추출된 골격선의 품질이 도면 자동 벡터화에 영향을 미치게 되므로, 적절하고도 좋은 세선화 알고리즘을 사용해야 자동 벡터화 시스템을 만들기 용이할 것이다.

세선화는 도면 자동 벡터화 시스템뿐만 아니라 문자 인식, 지문 인식, X-Ray 자동 영상 분석, 인쇄회로기판 자동 분석 등 다양한 분야에 적용이 되므로 세선화에 관한 기존의 연구(Chen과 Hsu, 1988; Datta와 Parui, 1994; Mandalia 등, 1992; Suzuki와 Abe, 1987; Zang과 Suen, 1984)는 많이 있다. 이경호 등(1997)은 지리 도면의 자동 벡터화를 위해 세선화 알고리즘들의 성능을 평가하였으며, 등고선도, 지적도, 상하수도 각각에 적합한 알고리즘을 고찰하였다. 그림 1은 세선화 결과의 한 예로, 사용된 세선화 알고리즘은 원남식과 손윤구(1995a, 1995b)가 제안한 WPTA이다. 기존의 세선화 알고리즘은 세선화 결과가 이진 정보 형태이나, WPTA는 각 화소들에 대해 8방향의 연결 상태를 나타내는 연결값을 이용하였고, 세선화 결과는 각 화소의 연결값에 기반한 수치로 표현된다.



(a) Input image of '가'



(b) Thinning result

FIGURE 1. Input image of '가' and its thinning result

제안된 전처리 기법

도면의 세선화 결과로 남는 골격선은 대부분 선의 중심부에 위치한 화소들로 구성된다. 선이 굵을수록 골격선을 둘러싸던 주변화소들의 두께는 두껍다. 선을 구성하는 각 화소에 대해 그 화소를 둘러싸고 있는 주변 화소들의 두께를 그 화소의 깊이라 정의하고, 화소 p 의 깊이를 $depth_p$ 로 표현한다. 그리고 임의의 두 화소간의 떨어진 정도를 거리로 정의한다. 이때 두께와 거리의 단위는 화소이다.

그림 2는 화소 p 에 대해 거리가 2 이내에 있는 모든 화소들을 나타내고 있다. 그림 2에서 $p_0 \sim p_7$ 은 거리 1의 위치에 있는 화소들이고, 가장 자리에 위치하며 약간 검게 표시된 $p_8 \sim p_{23}$ 은 거리 2의 위치에 있는 화소들이다.

p_{22}	p_{21}	p_8	p_9	p_{10}
p_{21}	p_7	p_0	p_1	p_{11}
p_{20}	p_6	p	p_2	p_{12}
p_{19}	p_5	p_4	p_3	p_{13}
p_{18}	p_{17}	p_{16}	p_{15}	p_{14}

FIGURE 2. Pixels within distance 2 from the pixel p

각 흑화소의 깊이를 계산하는 방법은 그 흑화소에 대해 거리 1의 위치에 있는 화소들 중 적어도 백화소가 1개 이상 존재하면 그 화소의 깊이는 1이 되고, 거리 1의 위치에 백화소가 하나도 없으면(모두 다 흑화소이면) 거리 2의 위치에 있는 화소들에 대해 적어도 하나의 백화소가 존재하는지를 조사한다. 만약 거리 2의 위치에 적어도 백화소가 하나 존재하면 그 화소의 깊이는 2가 되고, 그렇지 않으면 거리 3의 위치에 있는 화소들을 같은 방법으로 조사하여 그 화소의 깊이를 계산한다. 이러한 방법으로 각 화소의 깊이를 계산하면 모든 흑화소들은 적어도 깊이가 1 이상이 되며, 깊이가 1인 경우는 한 흑화소에 대해 그것이 흑화소로 완전히 둘러싸이지 못할 때이며, 가장 자리에 위치한 화소들이 그 경우에 해당한다. 그리고 가장자리 바로 안쪽 화소들의 깊이는 2가 되며, 안쪽으로 들어갈수록 각 화소의 깊이는 1씩 증가하게 된다.

도면에서 선을 구성하는 한 화소의 깊이를 효율적으로 계산하는 알고리즘을 C 언어의 함수 형식으로 그림 3에 나타내었다. 그림 3에 기술된 알고리즘은 깊이를 계산해야 하는 모든 화

소에 대해 병렬 처리가 가능하고, 각 화소의 깊이 계산 시간은 각 화소가 소속된 선의 굵기에 비례한다. 입력 도면의 가로와 세로 중 크기가 작은 쪽의 화소수가 n 이라면 한 화소의 깊이를 계산하는 제안된 알고리즘의 수행시간은 $O(n^2)$ 이 된다. 그림 4는 영어의 'A'자에 대해 제안된 깊이 계산 알고리즘을 적용하여 그 글자를 구성하는 각 화소의 깊이 정보를 나타내고 있다.

```

int Depth (int i, j)
(
// 매개변수 i, j는 행과 열을 각각 의미
int width;
width = 1;
for ( ; ; width++) {
if ( pij로부터 거리 width 위치에 있는 화소 중 적어도 하나가 도면 범위를 벗어남 )
return width;
if ( pij로부터 거리 width 위치에 적어도 하나의 백화소가 존재 )
return width;
}
)
    
```

FIGURE 3. Depth calculation algorithm for the pixel p_{ij}

기존 대부분의 세선화 알고리즘에서 입력과 출력 이미지는 이진 정보의 형태로 표현된다. 즉, 입력력 이미지는 패턴을 구성하는 흑화소와 그렇지 않은 백화소로 표현되고, 알고리즘 내에서 처리시에는 흑화소는 1로, 백화소는 0으로 표현하고 사용한다. 논의의 편의상 향후 그 화소가 흑화소인지 백화소인지를 표현하는 1 혹은 0을 그 화소의 화소값으로 정의한다. 화소값으로 이진 정보를 사용하는 알고리즘들에서 제안된 화소의 깊이를 계산하는 과정을 추가하여 흑화소들의 깊이를 먼저 계산할 수 있다. 그리고 각 흑화소에 대해서 화소값으로 1 대신 계산된 깊이를 사용하면 세선화 결과로 남는 골격선의 각 화소값은 1 대신 그 화소의 깊이가 수치 형태로 표현된다.

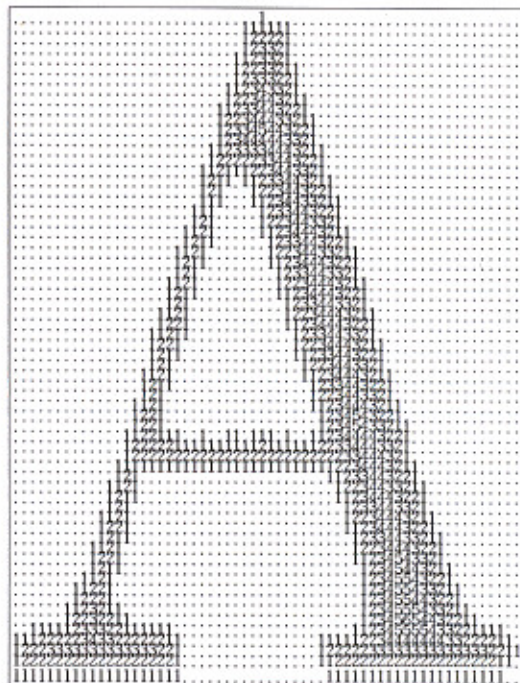


FIGURE 4. Each pixel's depth about the character 'A'

원남식과 손윤구(1995a, 1995b)가 제안한 WPTA에서는 이진 정보를 입력으로 받아 백화소는 -1로, 흑화소는 그 화소로부터 거리 1의 위치에 있는 8 방향의 이웃 화소들 중 흑화소의 수를 그 화소값으로 사용하였다. 그러므로 WPTA에서 흑화소는 화소값으로 1 대신 0 ~ 7의 값을 가진다. 기존의 세선화 알고리즘들에서 세선화 과정에서 사용되는 화소값은 모두 십진수 한 자리수이므로 그림 5에서처럼 화소값은 십진수 일의 자리에서 표시하고, 깊이는 십진수 일의 자리 이상에서 표현하면, 하나의 정수로 화소값과 깊이를 다 표현되게 할 수 있다.

십의 자리 이상	일의 자리
깊이	화소값

FIGURE 5. Expression of both a depth and a pixel's value within an integer

제안된 기법이 기존의 어떠한 세선화 알고리즘에도 전처리 기법으로 적용될 수 있도록 하기 위하여 깊이 정보는 깊이를 나타내는 값에다 10을 곱하고, 거기에서 기존 세선화 알고리즘에서 사용하는 화소값을 더하여 그 화소를 표현하는 값을 확장화소값으로 정의한다. 이렇게 하면 확장화소값 내에는 깊이 정보와 화소값이 공존하게 되므로 각 화소의 깊이 정보를 저장하기 위해서 별도의 공간을 사용할 필요가 없다. 식 (1)은 화소 p 의 화소값 $_p$ 과 깊이 $_p$ 로부터 확장화소값 $_p$ 을 계산하는 식이다.

$$\text{확장화소값}_p = \text{깊이}_p * 10 + \text{화소값}_p \dots\dots\dots (1)$$

확장화소값으로부터 깊이를 계산하기 위해서는 확장화소값을 10으로 나눈 몫을 취하면 되고, 화소값은 확장화소값을 10으로 나눈 나머지를 취하면 된다. 식 (2)는 화소 p 의 확장화소값 $_p$ 으로부터 그 화소의 깊이 $_p$ 를 계산하는 식이고, 식 (3)은 화소 p 의 확장화소값 $_p$ 으로부터 그 화소의 화소값 $_p$ 을 계산하는 식이다. 식 (2), (3)에서 사용된 '%'는 나머지 계산 연산자이고, '/'는 정수 나누기 연산자이다.

$$\text{깊이}_p = \text{확장화소값}_p / 10 \dots\dots\dots (2)$$

$$\text{화소값}_p = \text{확장화소값}_p \% 10 \dots\dots\dots (3)$$

기존의 세선화 알고리즘에서 본 연구에서 제안된 기법을 전처리 형식으로 사용하여 각 화소에 대해 확장화소값을 먼저 계산하고, 기존의 세선화 알고리즘 내에서 각 화소의 화소값 사용 시에 식 (2)를 사용하여 그 화소의 깊이를 계산하여 사용하면 세선화 알고리즘의 수행 결과로 나타나는 골격선에는 각 화소들의 깊이가 나타난다.

실험 및 고찰

제안된 방법을 C 언어로 구현하여 실험하였으며, 실험에 사용된 세선화 알고리즘은

WPTA이다. 그림 6은 입력에 사용된 등고선 도면을 스캔한 영상을 나타낸다. 그림 7은 그림 6의 영상에서 백화소는 '.', 흑화소는 '#' 글자로 변환한 그림이다. 그림 8은 그림 7을 입력으로하여 제안된 기법을 적용한 결과이며, 그림 8에서 골격선에 나타난 수치는 각 화소의 깊이를 표시하고 있다. 그림 8에서 알 수 있듯이, 가장 바깥쪽과 안쪽에서 두 번째 등고선의 골격선을 이루는 화소들의 깊이는 주로 2 혹은 3이고, 그 외 나머지 등고선의 골격선을 이루는 화소들의 깊이는 1 혹은 2로 이루어져 있음을 알 수 있다. 이들 깊이 정보의 통계를 이용하면 원래의 등고선이 주곡선인지 계곡선인지를 쉽게 알 수 있고, 또한 그 정보는 등고선 자동 벡터화에 유용하게 사용될 수 있을 것이다. 그리고 여러 개의 주제도가 겹쳐서 표시된 도면에서 각 주제도 별로 선의 굵기가 서로 다를 때, 골격선의 깊이 정보를 이용하면 자동으로 각 주제도를 분리하여 벡터화할 수도 있을 것이다.



FIGURE 6. Contour map



FIGURE 7. Converted picture as '.' and '#' for each pixel in the Fig. 6

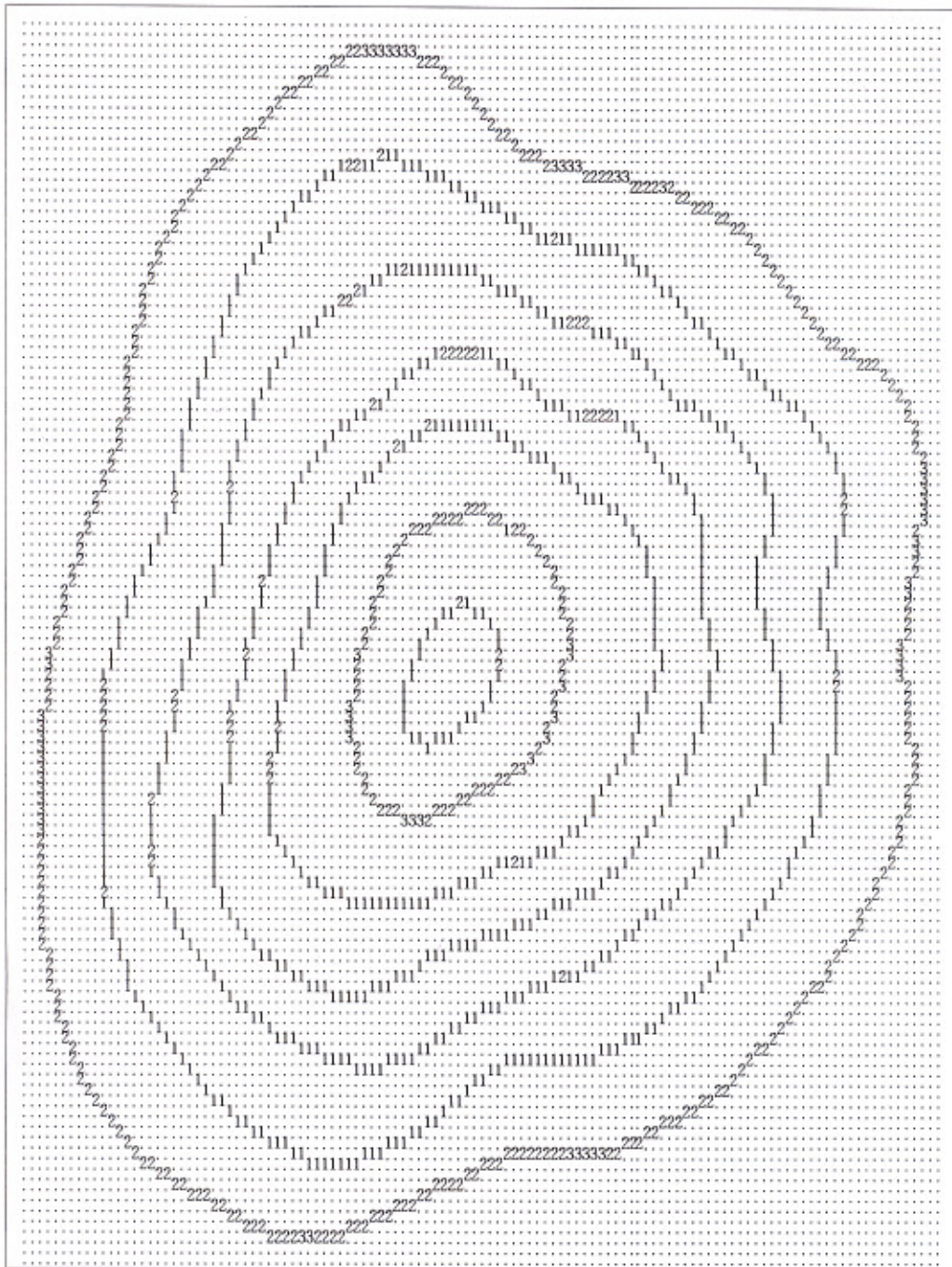


FIGURE 8. Thinning result using the proposed scheme from the Fig. 7

결 론

지리 도면의 자동 백터화 시에 세션화된 결과로부터 원래 선의 굵기 정보를 알 수 있으면 그것은 유용한 정보가 된다. 본 연구는 세션화 결과에서 원래 선의 굵기 정보를 표현하기 위한 방안으로 기존의 세션화 알고리즘에 전처리 형식으로 한 과정을 추가하여 이를 가능하게 하기 위한 기법을 제안하였다.

제안된 기법에서는 입력 도면에서 선을 구성하는 각 화소들에 대해서 그것이 주변화소들로부터 둘러싸인 정도, 즉 깊이를 계산하고, 세션화 후 골격선의 화소들에 대해서 그 깊이가 수치의 형태로 나타나게 하여 그 깊이를 보고 입력선의 굵기를 알 수 있도록 하였다.

제안된 기법을 구현하여 실험한 결과, 등고선의 경우에 골격선의 깊이 정보를 보고 그 선이 주곡선인지 계곡선인지 쉽게 알 수 있었다. 또한 각 주제도를 나타내는 선의 굵기가 각각 다른 여러 개의 주제도가 혼합된 도면의 경우에도 골격선의 깊이 정보를 보면 그 선이 어떤 주제도를 구성하는 선인지를 쉽게 알 수 있다. 그러므로 본 연구에서 제안된 기법은 그러한 혼합 도면으로부터 주제도별 자동 백터화 시스템 개발에 유용하게 사용될 수 있을 것이다.

향후 연구 과제로는 여러 가지 주제도가 혼합된 지리 도면으로부터 각각의 주제도를 자동으로 분리하고, 아울러 이들을 백터화하는 알고리즘을 개발하는 것과 등고선 자동 백터화 시스템을 개발하는 것이다. **KAGIS**

참고문헌

- 원남식, 손윤구. 1995a. 4-인접 연결값을 이용한 병렬 세션화 알고리즘. 정보과학회논문지 22(7):1047-1056.
- 원남식, 손윤구. 1995b. 8-이웃 연결값에 의한 병렬 세션화 알고리즘. 정보처리논문지 2(5):2960-2970.
- 이경호, 김경호, 조성배, 최윤철. 1997. 지리도면의 자동 백터화를 위한 영상 세션화 알고리즘의 체계적인 성능평가. 정보처리논문지 4(12):2960-2970.
- Chen, Y. S. and W.H. Hsu. 1988. A modified fast parallel algorithm for thinning digital patterns. Pattern Recognition Letters 7:99-106.
- Datta, A. and S.K. Parui. 1994. A robust parallel thinning algorithm for binary images. Pattern Recognition 27(9):1181-1192.
- Mandalia, A.D., A.S. Pandya. and R. Sudhaker. 1992. Modified fast parallel thinning algorithm for noisy handprinted characters. '92 Proceedings of the 2nd Singapore International Conf. on Image Processing. Singapore. pp. 7-11.
- Suzuki, S. and K. Abe. 1987. Binary picture thinning by an iterative parallel two-subcycle operation. Pattern Recognition 20(3):297-307.
- Zang, T.Y. and C.Y. Suen. 1984. A fast parallel algorithm for thinning digital patterns. Communication of the ACM 27(6):236-239. **KAGIS**