

소축적 지도제작을 위한 데이터 감축 기법에 관한 연구 A Study on the Data Reduction Techniques for Small Scale Map Production

곽강을* 이호남** 김명배**

Kwak, Kang Youl · Lee, Ho Nam · Kim, Myung Bae

요 旨

수치지도 제작의 일반화 기법은 대축척 실측 지도를 이용한 소축척 편집 지도의 자동 제작과 수치 지도 데이터베이스 구축 등의 활용 분야에 필수적 요소이다. 본 연구는 수작업으로 처리되고 있는 편집도의 제작을 자동화하기 위한 방안으로 디지털 환경하에서의 일반화 기법에 대한 연구로서 Zero-Crossing 알고리즘을 국립지리원 발행 1/5,000 지도 6도엽에 포함되어 있는 전남 여수 일대의 해안선에 적용, 다양한 단순화 허용 범위에 의해 수치 지도를 제작하여 편집도와 도해적인 표현 방법에 대한 문제점을 비교 분석하였다. 또한, Zero-Crossing 알고리즘에 의한 단순화의 효율성을 입증하기 위해 Douglas-Peucker 알고리즘을 이용한 단순화 결과와 비교 분석을 하였으며, 단순화의 허용 범위별로 일반화가 이루어진 수치 지도와 수작업으로 제작된 편집도와의 비교를 통해 일반화된 Line에 대한 평가를 도출하고, 단순화 허용 범위의 선택에 따른 효율성을 판단하기 위해 일반화 처리 결과의 데이터 용량을 분석하였다.

ABSTRACT

This paper is concentrated on map generalization in digital environment for automated multi-scale map production using conventional hardcopy maps. Line generalization is urgently required process to prepare small scale digital map database when large scale map databases are available. This paper outlines a new approach to the line generalization when preparing small scale map on the basis of existing large scale digital map. Line generalizations are conducted based on zero-crossing algorithm using six sheets of 1/5,000 scale YEOSU area which produced by National Geographic Institute. The results are compared to Douglas-Peucker algorithm and manual method. The study gives full details of the data reduction rates and alternatives based on the proposed algorithm.

1. 서 론

최근 국가적인 차원에서 정보화 사업으로 추진되고 있는 GIS를 효율적으로 활용하기 위해서는 수치 지도의 제작 기술과 소프트웨어 및 데이터베이스를 비롯하여 이에 대한 평가 등의 핵심 기반 기술이 준비되어야 한다. 수치 지도의 제작에 있어서 각종 지형 자료는 지상 측량과 항공사진 또는 인위성 영상을 이용한 해석 사진 측량 및 기존 지도의 입력 등 다양한 기법을 활용하여 수집되고 있다. 이렇게 수집된 자료는 사

용 목적에 따라 필요 부분의 표현을 극대화하기 위해 특정 지형의 삭제 또는 간결화가 요구되며, 이러한 일련의 데이터 감축 과정은 일반화를 통해서 이루어질 수 있다.^{1,3,12)}

지도는 표현하고자 하는 지형의 정보를 완벽하게 일치시킬 수 있다면 가장 이상적이라 할 수 있으나 편집도는 물론 실측도의 제작에도 이러한 표현은 거의 불가능하다. 이러한 경우 축척에 따라 정보의 취사선택이 이루어지게 되며, 해당 축척별 용도에 따른 이용상 중요한 것, 목표가 되는 것은 모두 표현하게 되며 중요하지 않은 것은 생략을 하는 일반화 과정을 거치게 된다.^{2,11)}

*중앙항업(주) 지리정보연구원 연구원

**명지대학교 대학원 박사과정

지도의 일반화는 실측된 지형 자료를 이용하여 사용자의 요구에 맞게 편집하는 역할을 담당하는 부분으로서, 지도 제작의 관점에서는 주로 대축척으로 표현되어 있는 실측도로부터 소축척 지도를 편집 제작하고자 할 경우 매우 중요한 전처리 과정이다.⁷⁾ 일반화는 선형성 지형 요소를 수치화 하는 과정에서 발생하는 중복 입력의 오류를 배제할 수 있으며, 래스터 데이터의 벡터화 처리 등에 매우 유용하다는 점에서 그 중요성을 찾을 수 있다.^{13,14)}

일반화는 모든 경우에 적용되는 명확한 규칙이 없거나 선택이 불가능할 경우 매우 어려운 작업으로서 작업자의 주관적인 견해로 문제를 해결하기 때문에 해당 영역에 대한 지형학적 지식은 물론 예술적인 감각까지도 필요하게 된다.⁶⁾ 지도의 일반화는 컴퓨터를 사용한 자동화의 개념이 부각되면서 더욱 어려운 문제로 나타나고 있고, 지도의 이용에 대한 확장 개념의 등장과 더불어 일반화는 지난 수 십년간 수치지도 제작 분야에 있어서 가장 해결하기 어려운 난제 중의 하나로 인식되고 있다.⁹⁾

특히 기존의 종이 지도에서의 축척에 따른 경직성을 탈피하기 위한 방법으로 수치지도 제작이 요구되고,⁵⁾ 지도 제작의 진산화에 따른 새로운 개념의 데이터 등장 및 그래픽 자료의 효율적인 방법에 대한 개선의 필요성이 부각되고 있고, 신속한 지리 정보의 생성 및 재편집을 위한 자동화 방안으로도 지도 일반화에 대한 요구가 급증되고 있다.^{8,10)}

또한, 편집하고자 하는 대상 지도가 단순히 정보 전달만을 강조하는 실용 매체일 뿐만 아니라 사용자의 요구에 맞게 지도의 크기나 축척, 주제별 내용 등이 충분히 포함되고 예술성이 가미된 편집 물로서의 역할을 기대하고자 할 경우 일반화 처리에 대한 연구는 조속히 해결해야 할 과제이다.⁴⁾

따라서 본 연구에서는 Line의 단순화를 래스터 데이터 구조의 수치 지도에 적용하기 위하여 Marr와 Hildreth가 제시한 Zero-Crossing 기법에 대한 해석 프로그램을 작성하여 일반화를 시도하였으며, 타당성을 규명하기 위해 기존의 Line 단순화에 가장 많이 사용되고 있는 Douglas-Peucker 기법과 비교 분석을 실시하였다. 또한, 단순화 허용범위를 설정하여 개발된 프로그램에 적용하여 도출된 결과와, 수작업으로 제작된 기존지도와의 비교분석을 통해 적용된 알고리즘의 효율성을 제

시하고자 하였다.

2. Zero-Crossing 알고리즘

기존의 일반화 알고리즘은 주로 벡터 데이터 구조에서 적용이 되어 왔으나, 이 경우 데이터 구조는 래스터 데이터를 선형화 하는 과정이 요구되며, 따라서 선형화 과정에서도 데이터의 감축이 시도 된다고 볼 수 있다. 이러한 오류를 최소화 하고 래스터를 기반으로 하는 GIS에 활용하기 위해서는 래스터 데이터에서의 직접적인 일반화에 대한 필요성이 요구되고 있다.

2.1 Zero-Crossing 알고리즘의 개념

연속된 영역 내에서의 임의의 화상 함수 $F(i, j)$ 의 윤곽선의 Laplacian은 다음과 같이 정의된다.

$$G(x, y) = -\nabla^2 F(x, y) \quad (2-1)$$

여기서, Laplacian은

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \quad (2-2)$$

이며, 만일 $F(x, y)$ 가 진폭에 있어서 상수 또는 선형적 변화를 나타낸다면 Laplacian $G(x, y)$ 은 0(zero)이 된다. 이 때, $F(x, y)$ 의 변화율이 현저하다면, $F(x, y)$ 의 변곡점에서 $G(x, y)$ 는 부호가 변화를 일으키게 된다. 따라서 $G(x, y)$ 의 Zero-Crossing은 윤곽선이 존재함을 의미하는 것이다. 식 (2-1)에서 (-)부호는, 윤곽선의 진폭이 좌에서 우로 또는 밑에서 위로 증가하기 위해서는 $G(x, y)$ 의 zero-crossing이 양의 경사도를 가져야 하기 때문에 부여되는 것이다. Marr & Hildreth는 Laplacian의 적용에 앞서 이루어지는 Gaussian-shaped smoothing에 있어서의 Laplacian of Gaussian(LoG) 윤곽선 검출 오퍼레이터를 제안했다. 연속된 영역에서의 LoG 기울기는,

$$G(x, y) = -\nabla^2 \{F(x, y) \odot H_s(x, y)\} \quad (2-3)$$

이며, 따라서 이차 미분의 선형화와 convolution의 선형화에 의해서 LoG response는 다음과 같이 정의될 수 있다.

$$G(x, y) = F(x, y) \odot H(x, y) \quad (2-4)$$

여기서 \odot 는 convolution operator이다.

2.2 해석 방법

체인 코드는 래스터 데이터를 저장하기 위한 데이터 구조로서, Freeman에 의해 유도되었기 때문에 Freeman 체인 코드로 명명되기도 한다. 체인 코드는 지형의 불연속성으로 인해 곡률에 대한 표현으로서 사용될 수 없을 수도 있기 때문에 곡률을 위한 유용한 표현을 획득하기 위해서는 체인 코드에서의 불연속성을 제거해야 한다. Pavlidis와 Eccels는 다음의 방정식을 이용하여 체인 코드에서의 불연속성을 제거할 수 있다고 하였다.

$$s(i) = (|e(i) - e(i-1)| + 11) \bmod 8 - 3 \quad (2-5)$$

여기서, $e(i) i=1, 2, 3, \dots, N$ $s(i) i=1, 2, 3, \dots, N$ 이다. 이 경우 체인 코드로 구성된 Gaussian function은 다음과 같이 나타낼 수 있다.

$$f(t, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp(-t^2/(2\sigma^2)) \quad (2-6)$$

식 (2-6)을 이차 미분의 형태로 나타내면

$$f''(t, \sigma) = \frac{df'(t, \sigma)}{dt} = C(\sigma^2 - t^2) \exp(-t^2/(2\sigma^2)) \quad (2-7)$$

으로 나타낼 수 있다. 여기서 C는 상수이다. 만일 C가 상수가 아니라면 sigma의 극값에서 zero crossing이 되지 않는다. 이산 구조인 경우 convolution은,

$$X_n = \sum S_k \odot h_{n-k} \quad (2-8)$$

로서 표현할 수 있다. 그러나, 이를 연속 구조로 나타내기 위해서는 다음과 같은 convolution의 합의 형태를

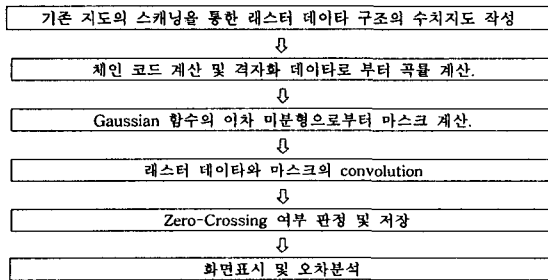


그림 1. 연구 수행을 위한 데이터 처리 흐름도

취해야 한다.

$$F(x, \sigma) = f(x) \times g(x, \sigma) \quad (2-9)$$

$$F(x, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \int f(u) \exp(-t^2/(2\sigma^2)) du$$

이상과 같은 이론에 근거하여 본 연구에서는 다음과 같은 과정의 처리 절차로 연구를 수행하였다.

3. 실험 및 결과 분석

본 논문의 연구대상지역은 그림 2와 같은 전남 여수 일대의 해안선 일부이며, 이 지역을 표본으로 하여 다양한 축척으로 수치화 하였다.

연구 지역은 국립지리원 발행 실측도인 축척 1/5,000, 1/25,000 지형도와 편집도인 1/50,000 지형도 및 1/250,000지세도를 이용하여 동일 지역에 대해 입력하였고, 그림 2는 축척별로 입력된 연구 지역을 출력한 것으로서 축척별로 입력된 점의 수와 1/5,000 지형도와

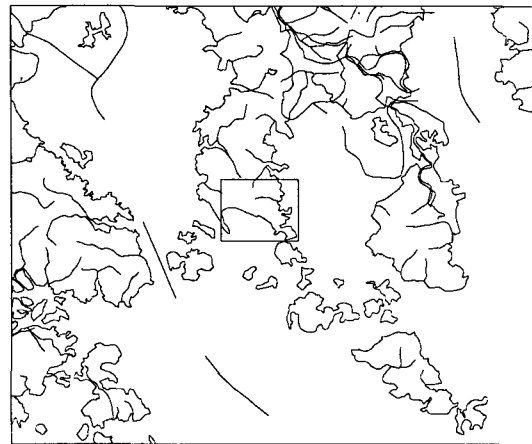


그림 2. 연구 지역

표 1. 일반화 전의 축척별 입력 지도의 데이터 비교

축척	Point	데이터 감축률(%)
1/5,000	3403	0
1/25,000	408	88.01
1/50,000	341	89.98
1/250,000	106	96.89

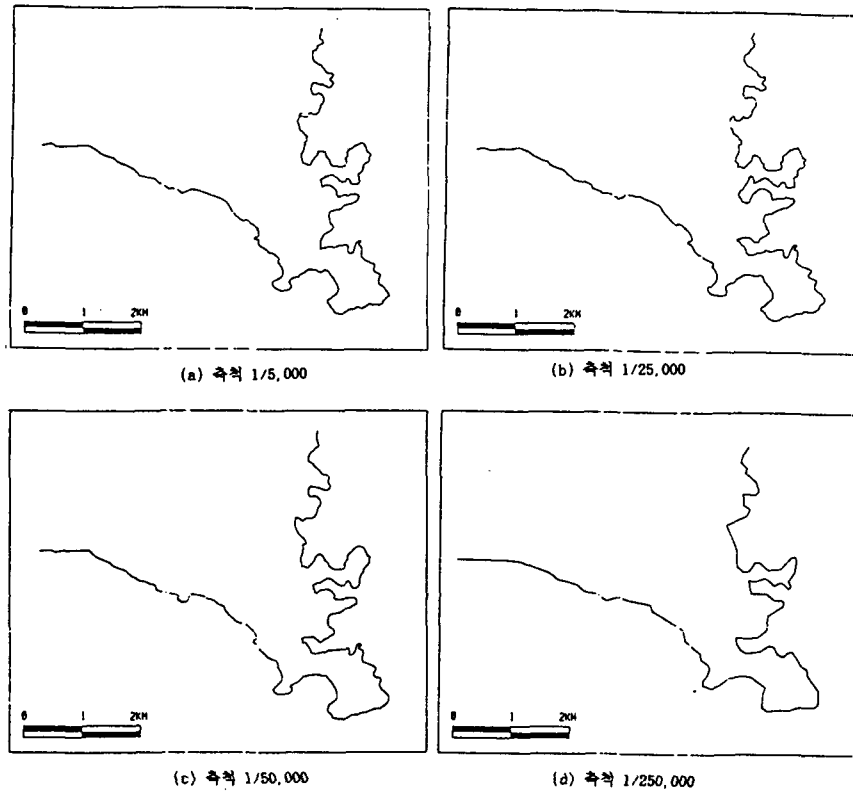


그림 3. 연구대상 지역의 축척별 입력 지도

비교하여 축척별 점수의 감축율은 표 1과 같다.

그림 3에서 볼 수 있듯이 실측도인 축척 1/5,000과 1/25,000의 지도에서는 지형 특성의 반영 정도가 거의 동일함을 알 수 있으나, 1/50,000이나 1/250,000과 같은 편집도는 실측도에 비해 지형 상의 오차가 많음을 알 수 있었으며, 이는 편집도의 피할 수 없는 특징으로 판단된다. 또한, 지형의 요철과 같은 지형 특성의 반영 정도가 지역마다 다른 도해적인 표현상의 문제점을 내포하고 있음을 알 수 있었다. 이는 일정한 기준 없이 작업자의 주관적 견해에 의해 수작업으로 단순화가 수행되었기 때문이라고 판단된다.

Line 일반화를 위한 기존의 방법에는 여러 가지가 있으나 대부분의 경우 데이터의 감축에 기본적 개념을 두고 있다. 그러한 기법의 대표적인 예로서 Douglas-Peucker 알고리즘을 들 수 있으며, 그림 4는 이러한 연구 지역의 특성 분석을 위해 1/5,000 지도를 이용하여 구축된 데이터를 각각 허용 범위 10 m, 25 m, 50 m, 100 m로 설정한 후 Douglas-Peucker 알고리즘으로 일반화한 지도이고, 표 2는 데이터 감축 후의 구성점에 대한

표 2. Douglas-Peucker 알고리즘을 이용한 일반화 결과

허용 범위 (m)	점의 수	데이터 감축률(%)
0	3403	-
10	255	92.5
25	136	96.0
50	90	97.4
100	55	98.4

점의 수와 감축율을 나타내고 있다. 그림 4에서 볼 수 있듯이 지도 구성 점의 수는 상당히 감축되었으나 이는 1/5,000 지형도를 구성하고 있는 대부분의 해안선이 돌기형의 요철부로 이루어졌기 때문에 이러한 점들이 감축되면서 점의 수가 감소된 것으로 볼 수 있으며, 실제의 지형 특성을 반영하기 위해서는 데이터의 저장량이 상대적으로 증가할 수밖에 없음을 예측할 수 있다.

또한, 지형 특성의 유지측면에서 수작업에 의한 때보다도 오히려 신뢰성을 확보할 수 없는 단점으로 인해

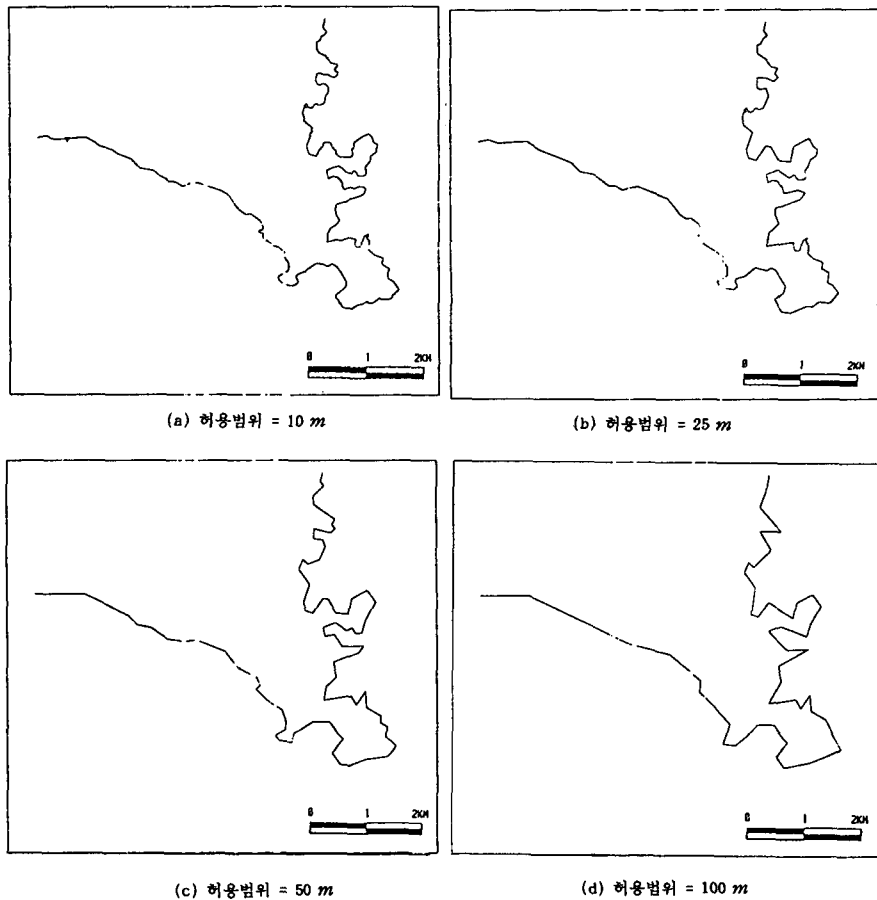


그림 4. Douglas-Peucker 알고리즘을 이용한 일반화 결과

상당한 개선이 요구되고 있다. 따라서 본 연구에서는 상기와 같은 일반화 문제를 해결하기 위한 방안으로 Marr & Hildreth(1980)가 제시한 래스터 데이터 구조에서 데이터를 감축하는 Zero-Crossing 알고리즘을 이용하여 일반화 후의 최종 성과의 시인성 및 지형 특성의 유지에 관한 문제점을 해결하고자 하였다.

래스터 데이터는 도면을 스캐닝하는 과정에서 지형 지물의 유무가 화상의 농도에 의해 결정되는 각각의 화소로 구성되어 있으며, 스캐닝된 지도 데이터로부터 얻고자 하는 것은 Line의 윤곽선 즉 외곽 부분 또는 중심선에 관한 것이다. 특히 체인 코드로 구성된 래스터 데이터에 있어서 윤곽선 부분을 검출하기 위한 필터링의 기법으로 일반적으로 식 (2-6)과 같은 Gaussian 함수가 활용되고 있으며, 에지를 검출하여 Zero-Crossing을 찾기 위해 식 (2-7)과 같이 Gaussian 함수의 이차 미분식이 이용되고 있다. 특히, Gaussian 함수의 sigma는

Line의 smoothing 정도를 결정하는 필터의 크기로서 본 연구에서는 단순화의 허용 범위를 의미하고 있다. 그림 5는 본 연구에서 지형 특성의 반영 정도와 데이터 감축률을 향상시키기 위해서 1/5,000 지형도를 대상으로 단순화의 허용 범위를 5, 10, 15, 20의 네 단계로 구분하여 Zero-Crossing 알고리즘에 적용하여 일반화를 시도한 결과이고, 표 3은 일반화 이후의 감축된 데이터량을 나타내고 있다. 단순화의 허용 범위를 4단계로 달리해 가면서 단순화를 시도한 결과 단순화 이후의 지형은 다음의 그림 5에서 볼 수 있듯이 실제의 지형과 거의 동일함을 알 수 있었으며, 지형 특성의 반영 정도도 월등히 향상되었음을 알 수 있었다. 일반화 허용 범위 별로 감축된 점의 수는, 허용 범위를 5로 하였을 때 363개, 10인 경우 242개, 15의 경우 189개이고 20의 경우는 총 154개로 구성 점의 수가 크게 감축되어 있음을 알 수 있었다. 이러한 점수의 감축에 다른 데이터의 감

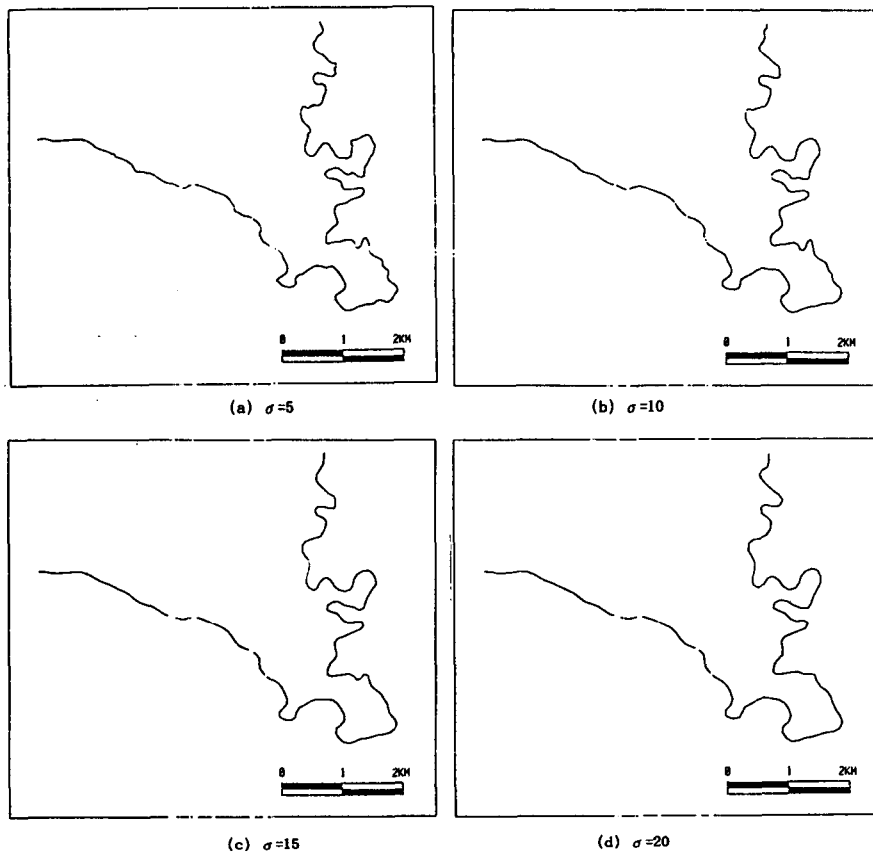


그림 5. Zero-Crossing 알고리즘을 이용한 일반화 결과

축률을 보면, 단순화 이전 1/5,000 지형도의 point수가 3,403이었던 것이 단순화 이후 허용 범위를 20으로 하였을 때 데이터 감축률은 95%를 상회하고 있다. 이는 Douglas Peucker 알고리즘에 의한 결과에 비해서 감축률은 낮으나 지형 특성의 반영 정도를 고려해 볼 때 보다 향상된 결과를 도출할 수 있음을 알 수 있었다. 또한 단순화의 허용 범위에 따른 지형의 오차는 불과 수

m 이내의 수준으로 감소하는 등 수작업에 의한 결과나 Douglas Peucker 알고리즘에 의한 결과에 비해 도해적으로 지형의 표현 능력이 향상되어 있음을 알 수 있었다.

4. 결 론

특정 지역에 적용된 제한된 연구 성과 이기는 하나 일반화의 허용범위와 비 선형성 지도 자료에 대한 일반화의 알고리즘 및 물리적인 저장 구조등에 대한 복합적인 연구가 시행될 경우 래스터 데이터를 이용해야만 하는 일반화의 보다 효율적인 방법론이 제시될 것으로 기대하며, 이상과 같은 연구 수행에 의해서 얻은 결과를 간단히 요약하면 다음과 같다.

첫째, 지형의 특성과 기본 형상을 유지하면서 단순화를 수행하기 위해서는 Zero-Crossing 알고리즘이 효과적이었음을 알 수 있었다.

표 3. 단순화 후의 데이터 비교 (축척 1/5,000)

구 분	Points	데이터 감축률 (%)	
단순화 전	3,403	0	
단순화 후	$\sigma=5$	363	89.33
	$\sigma=10$	242	92.89
	$\sigma=15$	189	94.45
	$\sigma=20$	154	95.47

둘째, Douglas Peucker 알고리즘에 의한 단순화 결과에 비해 지형의 표현을 정확하고 부드러운 곡선 형태로 나타낼 수 있었으며, 도해적인 표현 능력이 향상된 결과를 도출할 수 있었다.

셋째, 허용 범위별로 수행된 단순화 이후의 각각의 데이터 용량은 단순화 이전에 비해 상당 수준 감축되었으며, 기존의 알고리즘에 의해 단순화된 지형도의 데이터 용량과 유사한 정도로 감축되었으나, 지형 특성의 반영 정도는 월등히 향상 되었음을 알 수 있었다.

參考文獻

1. Bernd M.P., "Computer-Assisted Generalization-An Important Software-Tool in GIS", '89 ISPRS, Vol. IV, 1992, pp. 664-672.
2. Buttenfield B.P., McMaster R.B., "Map Generalization: Making rules for knowledge representation", Longman Scientific & Tech. 1992, pp 21-39.
3. David J. Maguire. *et al.*, "Geographical Information Systems", Vol.1, John Wiley, 1991, pp. 111-117.
4. Kilpelainen T., "Multiple Representation & Knowledge - based Generalization of Topographic Data", '92 ISPRS, Washington D.C., 1992, pp. 954-964.
5. Monmonier M.C., "Computer-Assisted Cartography", Prentice-Hall, 1982, pp. 170-175.
6. Philip W., "Computer-Assisted Cartographic Generalization and its practical Application with PHOCUS", '92 ISPRS Proceeding, 1992, pp. 139-142.
7. Robinson A.H., *et al.*, "Element of Cartography", John Wiley, 1991, pp. 251-254.
8. Stoms D.M., "Effects of Habitat Map Generalization in Biodiversity Assessment", PE & RS, Vol. 11, 1992, pp. 1587-1591.
9. Thapa K., "Automatic line Generalization Using Zero-Crossings", PE&RS, Vol. 4, 1988, pp. 512-517.
10. Zhilin Li & Stan Openshaw, "Algorithms for automated line generalization based on a natural principle of objective generalization", IJGIS, Vol. 6, No. 5, 1992, pp. 373-389.
11. 김감래, 이호남, 김종훈, "디지털환경하의 목적지향형 지도 일반화", 대한토목학회 학술발표개요집, 1992, pp. 320-323.
12. 김감래, 이호남, 박인해, "지도 일반화에 따른 단순화 알고리즘의 평가에 관한 연구", 한국측지학회지 제 10 권, 제 2 호, 1992, pp. 63-71.
13. 김감래, 이호남, "선형성 지형자료의 일반화에 대한 효율적인 알고리즘에 관한 연구", 한국측지학회지, 1994, pp. 43-52.
14. 김감래, 이호남, 광강울, "Line 단순화를 위한 Zero-crossing 알고리즘의 적용에 관한 연구", 대한토목학회 학술 발표회집, 1994, pp. 451-454.