

# 임계값 설정에 따른 선형 단순화 알고리즘의 반응 특성 연구

## Performance Assessments of Three Line Simplification Algorithms with Tolerance Changes

이재은<sup>1)</sup> · 박우진<sup>2)</sup> · 유기윤<sup>3)</sup>  
Lee, Jae eun · Park, Woo jin · Yu, Ki yun

### Abstract

The result of the line simplification algorithm varies with the choice of algorithms, the change in tolerance and the selection of target objects. Three of the algorithms used in this study are Sleeve-fitting, Visvalingam-Whyatt, and Bend-simplify. They were applied to the three kinds of objects which were buildings, rivers, and roads with the five degrees of the tolerance. Through this experiments the vector displacement, the areal displacement, and the angular displacement were measured and the qualitative analysis was performed with the trend line of the errors. The experimental results show that errors were differ from tolerance values, and characteristics of line simplification algorithms based on changes of tolerance were understood.

Keywords : Map generalization, Line simplification algorithms, Tolerance, Performance assessments

### 초 록

선형 단순화 알고리즘의 결과물은 단순화 알고리즘의 선택, 임계값의 선택, 대상 객체의 선택 등에 의해 달라진다. 본 연구에서는 Sleeve-fitting, Visvalingam-Whyatt, Bend-simplify 알고리즘의 임계값을 5단계로 조절하여 건물, 하천, 도로 객체에 적용하고, 단순화 결과물에 대해 거리오차, 면적오차, 각오차를 측정하였다. 그리고 오차값들의 추세를 통해 정성적인 분석을 실시하였다. 실험 결과 선형 단순화 알고리즘에서 임계값의 조절에 따른 오차값의 변화를 알 수 있었고, 선형 단순화 알고리즘들의 임계값의 변화에 따른 특징을 파악할 수 있었다.

핵심어 : 지도 일반화, 선형 단순화 알고리즘, 임계값, 성능평가

## 1. 서 론

지도 제작의 가장 근본적인 목적은 현실세계를 축소시켜서 나타낸다는 것이다. 이 과정에서 필연적으로 크기의 축소 과정이 필요하다. 지도 일반화 기술은 이 과정에서 지도가 전달하고자 하는 현실세계의 정보를 보다 정확하게 전달하게 하기 위한 기술이다. 즉, 현실세계를 지도상에 나타낼 때, 그 표현 방법에 관한 법칙이나 방법들을 포

함한다. 하지만 이러한 방법들을 수동으로 적용할 경우에는 많은 어려움이 따르게 된다.

이러한 어려움을 극복하기 위해, 여러 가지 자동화 방법들이 개발되었는데, 그 중 현재까지 가장 널리 연구되고 활용되어지고 있는 방법은 선형 단순화 방법이라고 할 수 있다. 최소한의 자료 처리와 저장 공간으로 현실세계의 선형에 대한 정확한 묘사를 위한 연구들이 진행되는 과정에서, 그 결과는 다양한 선형 단순화 알고리즘의 개

1) 서울대학교 건설환경공학부 박사과정(E-mail:perfectsolo@snu.ac.kr)  
2) 서울대학교 건설환경공학부 박사과정(E-mail:woojin1@snu.ac.kr)  
3) 교신저자 · 정회원 · 서울대학교 건설환경공학부 부교수(E-mail:kiyun@snu.ac.kr)

발로 이어졌다. 선형 단순화는 지도 일반화 기술 중에서 가장 핵심이라고 할 수 있는 부분으로, 현실세계를 축소하여 나타낼 때, 지도상의 모든 선형 객체에 대해서 불필요한 절점의 삭제 또는 중요한 영역의 보존 등의 방식으로 객체를 단순화 시켜주는 방법이다.

선형 단순화의 있어서 임계값은 결과에 크게 영향을 줄 수 있는 요소이다. 선형 단순화 알고리즘들은 다양한 방식으로 임계값을 입력할 수 있는데, 임계값의 수준에 따라 단순화의 수준도 변하게 된다. 따라서 임계값은 사용자가 원하는 단순화 결과물의 수준에 따라 적절하게 입력될 필요가 있으며, 임계값의 변화에 따른 각 알고리즘의 성능 평가도 이루어져야 한다. 단일 알고리즘에 대하여 임계값을 조절한 것에 대한 성능평가에 대한 연구는 다수 존재하고 있지만, (Shahriari 등, 2002; Cheung 등, 2006; Shi 등, 2006) 여러 알고리즘에 대해 임계값의 조절에 따른 알고리즘의 성능 평가에 대한 연구는 미흡한 편이다.

본 연구에서는 여러 선형 단순화 알고리즘들을 임계값을 변화시키면서 다양한 객체에 적용시키고, 그 결과를 분석하고자 하였다. 이를 위하여, Sleeve-fitting, Visvalingam-Whyatt, Bend-simplify 알고리즘을 선택하여, 수치지도로부터 건물, 하천, 도로 객체를 각각 20개씩 추출하여 실험 데이터로 선정하고, 알고리즘들의 임계값을 5단계로 설정하여 각각의 객체에 단계별로 선형 단순화 알고리즘을 적용하였다. 각 객체에 대해 얻어진 단순화 결과물에 대한 분석을 위해, 단순화 결과물에 대해 거리 오차, 면적오차, 각오차를 측정하였다. 이 과정에서 측정된 모든 오차값을 알고리즘의 종류, 객체의 종류, 오차값에 따라 분류한 후, 오차값들의 평균을 그래프의 형태로 표현하여 분석을 시도하였다.

## 2. 선형 단순화 알고리즘

선형 단순화 과정에는 여러 가지 알고리즘들이 사용된다. 가장 대표적인 것으로 1973년에 발표되어 현재까지 가장 널리 사용되는 Douglas-Peucker 알고리즘이 있다. 이 알고리즘에 대한 연구는 지난 수십 년간 계속되어 왔으며, 알고리즘의 성능에 대한 평가도 다수의 논문에서 다루고 있다. 따라서 이 알고리즘은 본 연구에서는 다루지 않았다. 본 연구에서는 Douglas-Peucker와 유사한 절점 제거 방식의 알고리즘인 Sleeve-fitting 알고리즘, 그리고 절점들이 이루는 영역들을 제거하여 전체적인 선형 형상의 유지에

초점을 맞춘 Visvalingam-Whyatt 알고리즘, Bend-Simplify 알고리즘에 대한 실험 및 평가를 수행하였다.

### 2.1 Sleeve-fitting 알고리즘

Sleeve-fitting 알고리즘은 먼저 알고리즘의 적용 대상이 되는 선형을 Sleeve라는 직사각형 모양의 단위로 나눈다. 각각의 Sleeve는 시작점  $P_j$ 로부터 연속적으로 이어진  $P_{j+k}$  까지의 점을 포함하게 된다. 이 때 Sleeve의 방향은  $P_j$  와  $P_{j+k}$ 를 이은 직선에 평행해야 한다. 첫 번째 Sleeve의 경우, 시작점은 첫 점인  $P_1$ 이 된다. 끝점인  $P_s$ 는  $P_1$ 과  $P_s$  사이에 있는 점을 모두 포함할 수 있는 Sleeve가 생성되도록  $s$ 값이 설정되어야 한다. 첫 번째 Sleeve가 형성된 후 첫 번째 Sleeve의 끝점이었던  $P_s$ 를 시작점으로 하는 두 번째 Sleeve를 생성한다.(Zhao 등, 1997) 이와 같은 과정을 각각의 Sleeve를 따라 대상 선형의 시작점으로부터 끝점까지 반복한다. 이렇게 생성된 Sleeve의 중심선이 원래 대상선형으로부터 단순화 된 선형이 된다. Sleeve-fitting 알고리즘의 과정은 그림1에 나타나 있다.

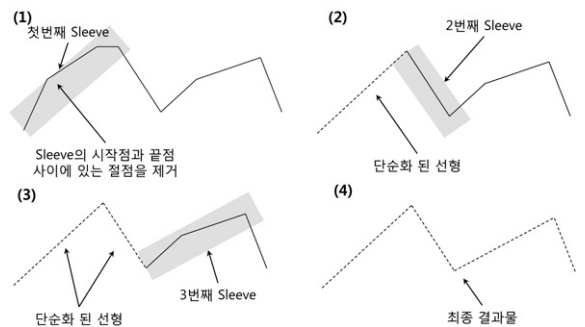


그림 1. Sleeve-fitting 알고리즘

### 2.2 Visvalingam-Whyatt 알고리즘

Visvalingam-Whyatt 알고리즘은 대상 선형 전체에 대해서 인접한 세 점으로 구성된 삼각형의 면적인 유효면적을 계산하게 된다.(Visvalingam 등, 1993) 모든 유효면적을 계산하고 나면, 가장 작은 유효면적의 꼭짓점을 제거한다. 알고리즘이 반복적으로 수행되면서, 각각의 유효면적은 다시 계산되고 다시 최소 면적을 갖는 유효면적의 꼭짓점을 제거하는 과정을 반복하게 된다. 알고리즘의 종료 조건은 최소 유효면적으로 설정하여 계산된 최소 유효면적이 임계값보다 크면 반복을 종료하는 것으로 설정했다. 알고리즘의 수행 과정은 그림2에 나타나 있다.

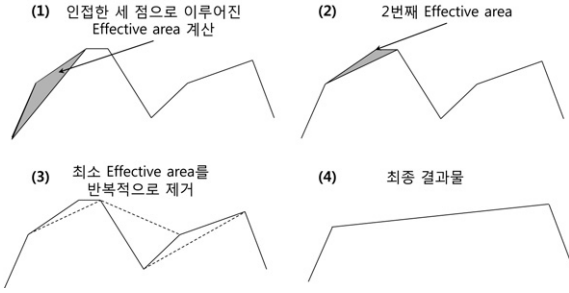


그림 2. Visvalingam-Whyatt 알고리즘

### 2.3 Bend-simplify 알고리즘

Bend-simplify 알고리즘은 Visvalingam-Whyatt 알고리즘과 마찬가지로 영역을 기반으로 한 단순화 알고리즘이다 (Wang 등, 1998). 이 알고리즘은 선형을 변곡점 단위로 나눠지는 Bend 단위로 분할한 후, 각각의 Bend들의 특징을 파악하고, 불필요한 Bend의 경우, 그 Bend를 제거 하는 방식이다. 먼저 Bend 단위로 선형을 분할하기 위해서는 각 점에서 두 선분이 이루는 각도를 계산하여, 이 각도가 같은 방향성을 보이는 선형을 하나의 Bend로 인식한다. 이렇게 Bend 단위를 인식한 후에는, 각 Bend의 밑선의 길이를 계산하고, 이 밑선의 길이가 주어진 임계값보다 작을 경우 이 Bend를 제거한다. 알고리즘의 수행 과정은 그림 3에 나타나 있다.

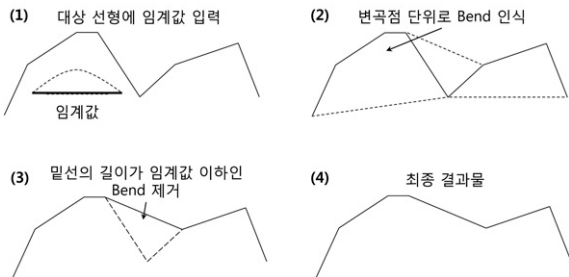


그림 3. Bend-simplify 알고리즘

## 3. 실험 및 결과

### 3.1 실험데이터

본 연구에서는 대전 지역의 수치지도 2.0으로부터 건물 외곽선, 하천 경계, 도로 경계선 각 20개를 추출하여 실험 데이터로 사용하였다. 사용된 실험 데이터들은 그림 4와 같다.

건물					
하천					
도로					

그림 4. 사용된 실험 데이터

### 3.2 실험 방법

위 데이터들에 대한 실험은 다음과 같이 진행되었다. 앞서 언급한 세 가지 선형 단순화 알고리즘(Sleeve-fitting, Visvalingam-Whyatt, Bend-simplify)에 대하여 해당 알고리즘마다 각각 5단계로 임계값을 조절하면서 건물, 하천, 도로 각각의 객체마다 알고리즘을 적용하였다. 임계값 조절에 있어서, 각 알고리즘마다 임계값이 의미하는 바를 다를 수 있다. 따라서 알고리즘별 결과들이 최대한 유사한 수준으로 단순화 되도록 임의의 임계값으로 반복적인 실험을 통해 임계값을 결정하였다.

알고리즘 별 임계값의 설정은, Sleeve-fitting의 경우 오차량 자체가 임계값의 영향을 크게 받기 때문에 2m로부터 2m간격으로 5단계로 설정하였다. Visvalingam-Whyatt 알고리즘은, 임계값을 2m<sup>2</sup>단위로 주었을 때는 오차값의 변화가 거의 없어서 알고리즘의 특성을 파악하기가 어려웠다. 의미 있는 오차값의 변화를 관측하기 위해 임계값을 5m<sup>2</sup>부터 5m<sup>2</sup>단위로 설정하여 실험하였다. Bend-simplify 알고리즘 역시, Sleeve-fitting과 같이 2m간격의 임계값으로 실험해 본 결과, 오차값의 변화가 미미하여 알고리즘의 특성을 파악하는 데 적합하지 않다는 결론을 얻었다. 따라서 Bend-simplify 알고리즘에서도 5m<sup>2</sup>단위의 임계값으로 실험을 수행하였다.

먼저 3가지의 알고리즘별로 임계값을 변화시켜가면서 건물, 하천, 도로 객체 각 20개에 대해 알고리즘을 적용하고 난 뒤, 그 결과물로서 얻어진 점 데이터들에 대한 거리오차, 면적오차, 각오차를 측정하였다.

표 1. 알고리즘별 임계값 설정

임계값/ 알고리즘	Sleeve -fitting	Visvalingam -Whyatt	Bend -simplify
초기 값	2m	5m <sup>2</sup>	5m
간격	2m	5m <sup>2</sup>	5m

거리오차의 경우, McMaster(1986)가 제안한 방법을 사용하였다. 이 방법은 단순화되기 전의 선형의 절점에서부터 단순화 후의 선분까지의 수직거리를 측정하는 방법이다. 본 연구에서는 알고리즘간의 비교를 위해 측정된 수직거리의 최댓값으로 거리오차를 계산하였다. 면적오차는 전체적인 선형 모양의 변화를 측정하기 위한 척도로, 이 방법 역시 McMaster(1986)가 제안한 방법을 사용하였다. 먼저 단순화 된 선형의 밀선과 단순화 된 선형이 이루는 면적을 계산한다. 계산된 면적을 모두 더한 후, 더해준 면적을 단순화 된 선형의 길이로 나누어서 면적오차를 측정한다. 이 과정을 수식으로 나타내면 식 (1)과 같다.

$$areal\ displacement = \frac{\sum_{j=1}^m s_j}{\sum_{i=1}^n l_i} \quad (1)$$

식 (1)에서  $m$ 은 단순화 된 선형과 밀선이 이루는 모든 영역의 개수이고,  $s_j$ 는 각각의 영역들의 면적이다. 그리고  $n$ 은 원래 선형의 x,y 좌표 쌍의 개수,  $l_i$ 는 원래 선형의 길이를 의미한다. 각오차 역시 McMaster(1986)가 제안한 방법을 활용하였다. 각오차의 측정은 연속적인 직선에서의 일어나는 각 변화량의 절대값을 계산하는 방법으로 이루어지는데, McMaster(1986)의 방법은 변화의 비율, 즉, 단순화 이후의 선형에서 연속적인 직선간의 각도 차이를 모두 더한 값을 단순화 이전의 선형에서의 각도 차이를 모두 더한 값으로 나눈 뒤, 100을 곱하여 그것의 비율을 계산하는 것으로 각오차를 측정하였다. 본 연구에서는 비율은 사용하지 않고, 단순히 나눈 값 그 자체를 각오차로 간주하였다. 식 (2)는 각오차를 측정하는 방법이다.

$$angular\ displacement = \frac{\sum_{j=1}^n |a_i - a_j|}{\sum_{i=1}^n |a_i - a_{i+1}|} \quad (2)$$

식 (2)에서  $n$ 은 선형의 절점의 개수,  $a_i - a_j$ 는 단순화 된 선형의 직선간의 각도차이를 나타내고,  $a_i - a_{i+1}$ 은 단순화 전의 선형의 직선간의 각도 차이를 나타낸다. 측정된 거리오차, 면적오차, 각오차 값들의 평균을 구해 알고리즘별, 객체별로 그래프를 그리고, 임계값의 변화에 따른 오차량의 증가를 분석하였다.

### 3.3 실험 분석

위의 일련의 실험 과정을 거친 후 얻어진 데이터들을 정리한 후, 그래프의 형태로 나타내었다. 한 개의 알고리즘에서 건물, 하천, 도로 3가지의 객체에 대한 거리오차, 면적오차, 각오차 3가지의 오차를 측정하여, 총 9개의 그래프를 얻었다. 그래프에는 각 알고리즘별, 객체별로 측정된 오차값들의 평균값을 꺾은선 그래프의 형태로 나타냈다. 이렇게 얻어진 그래프를 통해, 임계값의 조절에 따른 알고리즘별, 객체별 변화 특성을 알아보았다. 이러한 분석을 수행한 결과, 사용한 알고리즘 및 객체별로 나타나는 오차값들의 특성을 파악할 수 있었다.

#### 3.3.1 Sleeve-fitting 알고리즘에 대한 분석

Sleeve-fitting 알고리즘의 결과는 그림 5와 같다. 전체적인 오차값의 분포에서 다른 두 가지 알고리즘보다 큰 값을 보이는 양상을 보였다.

특히, Sleeve-fitting에서의 거리오차값은 입력된 임계값 이상이 나오지 않았으며, 임계값의 변화와 매우 밀접한 상관성을 보였다. 면적오차값의 분포는 인공물과 자연물에 따라 차이가 나타났다. 인공물이라고 할 수 있는 건물과 도로에서는 최대 오차값의 크기가 3m미만이었지만, 자연물인 하천에서는 최대 오차값의 크기가 4m이상으로 나왔다. 실험결과로 봤을 때, Sleeve-fitting 알고리즘은 면적오차 면에서는 인공물에 더 적합하다고 볼 수 있다. 각오차의 경우에는 건물객체에서의 오차값은 작지만 오차값들이 흩어져서 분포하는 형태인 반면, 하천과 도로에서의 오차값은 더 크지만 전체적으로 오차값들이 평균선 주위에 밀집하여 분포하는 모습을 보였다. 이러한 점으로 보아, Sleeve-fitting 알고리즘은 각오차만을 놓고 본다면 하천이나 도로 같은 선형으로 이루어진 객체에 적합한 알고리즘이라고 할 수 있다.

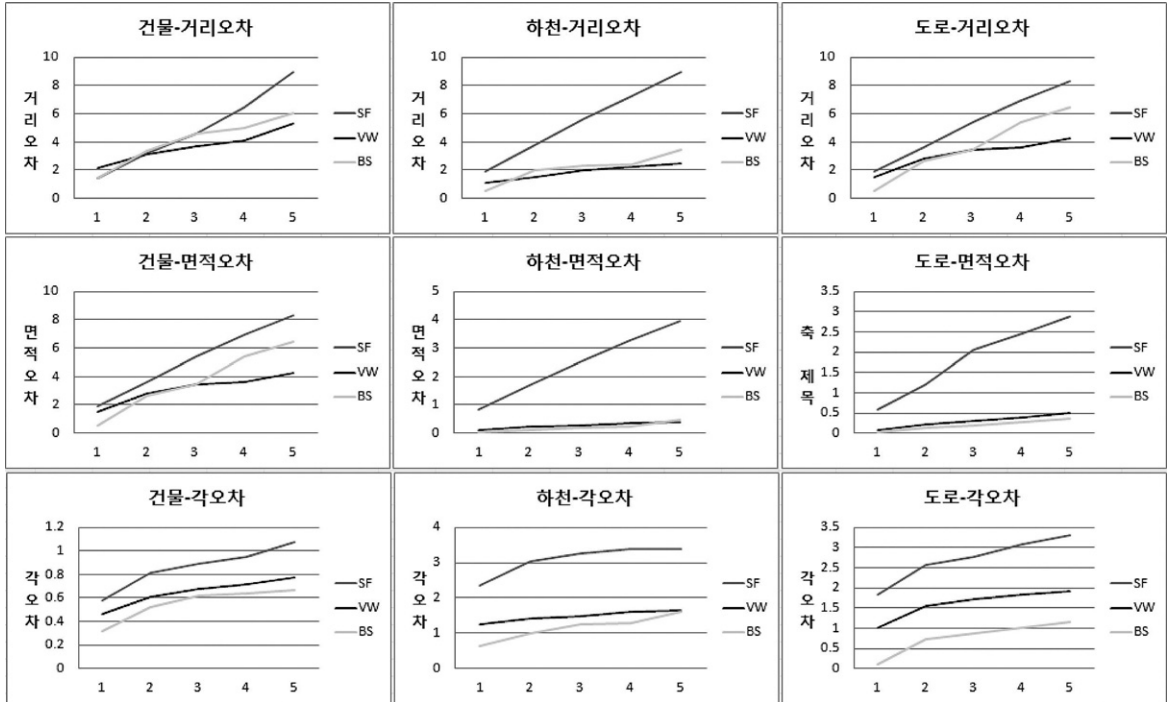


그림 5. 각 알고리즘별, 객체별로 나타난 거리오차, 면적오차, 각오차

### 3.3.2 Visvalingam-Whyatt 알고리즘에 대한 분석

Visvalingam-Whyatt 단순화 결과물에 대한 오차 측정 결과는 그림 6과 같다. 이 결과에서는 Sleeve-fitting과 다른 점을 볼 수 있었다. 전체적으로 모든 오차값들이 감소하였으며, 특히 Sleeve-fitting의 거리오차에서 볼 수 있었던 임계값의 증가에 따른 오차량의 증가가 뚜렷하게 나타나지 않았다. 면적오차값은 거리오차값과 마찬가지로 Sleeve-fitting에서의 오차값보다 작게 나왔지만, 인공물과 자연물에서의 차이나 그래프의 모양 등은 Sleeve-fitting과 유사한 모습을 보였다. 각오차에서도 마찬가지로 오차값의 차이만 있을 뿐, 그래프와 오차값들의 분포 형태는 대체적으로 Sleeve-fitting과 유사한 모습을 보였다.

### 3.3.3 Bend-simplify 알고리즘에 대한 분석

Bend-simplify 알고리즘에서의 오차값의 분포는 그림 7과 같다. 역시 비슷한 방식인 Visvalingam-Whyatt 알고리즘에서의 분포와 비슷한 양상을 보였다. 먼저 거리오차값은 전체적인 분포값의 형태는 Visvalingam-Whyatt 알고리즘과 비슷하게 나타났다. 면적오차값의 분포에서도 전체적으로는 Visvalingam-Whyatt 알고리즘과 유사한 양상을 보였고, 다른 두 알고리즘에서 공통적으로 나타난 인공물에서

와 자연물에서의 차이도 관찰할 수 있었다. 각오차값은 그래프의 모양과 오차값의 분포 양상 모두에서 Visvalingam-Whyatt 알고리즘과 거의 유사한 양상을 보였다.

## 4. 결론

본 연구에서는 세 가지 선형 단순화 알고리즘을 임계값을 조절하여 적용한 결과에 대한 성능 평가를 실시하였으며 다음과 같은 결론을 내릴 수 있었다.

첫째, Sleeve-fitting은 임계값의 증가에 따라 오차값의 변화의 폭이 크고, 거리오차에서 입력된 임계값 이상의 값이 나오지 않는다. 가령, 2m의 임계값을 입력했을 때, 오차값은 2m를 넘지 않는 범위 내에 분포하게 된다. 따라서 원하는 결과물의 조건에 따라 임계값을 조절하는 것이 가능할 것이라고 생각된다.

둘째, 세 알고리즘 모두에서 거리오차값이 임계값의 변화에 가장 민감한 변화양상을 보였다. 따라서 임계값의 변화에 따른 단순화 알고리즘의 성능 평가를 위해서는 거리오차값을 사용하는 것이 가장 적절하다고 생각된다.

셋째, 면적오차값이 세 알고리즘 모두에서 자연물에서의 오차값들이 일관적이지 못하게 분포하고 오차량이 큰

것으로 볼 때, 면적오차의 측정은 인공물을 대상으로 할 때 더 의미가 있는 것으로 보인다.

셋째, 각오차값은 세 알고리즘 모두에서 건물 객체에서의 오차값이 더 작게 나타났다. 즉, 면으로 구성되어 있는 객체에서의 단순화는 선으로 구성되어 있는 객체에서의 단순화에 비해 각오차 값에는 크게 영향을 끼치지 않는다고 해석할 수 있다.

본 연구를 통해 선형 단순화 알고리즘에서 임계값의 조절에 따른 오차값의 변화를 파악하고자 하였다. 본 연구에서는 제한적인 알고리즘과 제한적인 실험 데이터를 사용하였지만, 보다 다양한 알고리즘과 다양한 객체에 대한 실험이 이루어진다면 이러한 연구를 바탕으로 객체의 특성에 따라 적절한 알고리즘과 임계값을 자동으로 선택하는 자동화 기법도 나올 수 있으리라고 생각한다.

### 감사의 글

본 연구는 국토해양부 첨단도시개발사업-차세대 Hyper-Live Map 기반기술 개발과제의 연구비지원 (11첨단 도시G10)에 의해 수행되었습니다.

### 참고문헌

Cheung, C.K. and Shi, W. (2006), Positional error modeling for line simplification based on automatic shape similarity analysis in GIS, *Computers&Geoscience*, Vol. 32, pp. 462-475.

McMaster, R.B. (1986), A stastical analysis of mathematical measures for linear simplification, *The American Cartographer*, Vol. 13, No. 2, pp. 103-116.

Shahriari, N. and Tao, V. (2002), Minimizing positional errors in line simplification using adaptive tolerance values, *Symposium on Geospatial Theory*, Ottawa, pp. 153-166.

Shi, W. and Cheung, C. (2006), Performance evaluation of line simplification algorithms for vector generalization, *The Cartographic Journal*, Vol. 43, No. 1, pp. 27-44.

Visvalingam, M. and Whyatt, J.D. (1993), Line generalisation by repeated elimination of points, *The Cartographic Journal*, Vol. 30 No.1, pp. 46-51.

Wang, Z. and Müller, J.C. (1998), Line generalization based on analysis of shape characteristics, *Cartography and Geographic Information systems*, Vol. 25, No. 1, pp. 3-15.

Zhao, Z. and Saalfeld, A. (1997), Linear-time sleeve-fitting polyline simplification algorithms, *ACSM/ASPRS'97 Technical papers*, pp. 214-223.

(접수일 2012. 08. 03, 심사일 2012. 08. 14, 심사완료일 2012. 08. 25)