

Douglas-Peucker 단순화 알고리즘 개선에 관한 연구 A Study on The Improvement of Douglas-Peucker's Polyline Simplification Algorithm

황 철 수

Hwang, Chul-Sue

요 旨

Douglas-Peucker 알고리즘의 원리를 충실하게 반영한 단순 tree 구조의 단순화 기법은 단순화 지표가 실제 계층적 자료구조에 명확히 내재되는 장점을 갖는다. 그러나 단순 tree 방법은 단순화 지표의 계층성이 항상 보장되지 못할 가능성을 안고 있다. 그것은 Douglas-Peucker 알고리즘의 원리가 선형 사상의 국지적 특성을 충실하게 반영하지 못하는 전역적 기법이기 때문이다. 본 연구에서는 이러한 계층적 오류를 극복하기 위해 블록다각형 탐색기법을 활용하여 형태적 대표점을 찾아 이를 기초로 계층적 자료구조를 갖는 단순화 기법(CALS)을 구현하였다. CALS에 의한 방법은 단순 tree 방법에서 발생한 중상위 계층의 오류를 보정하는 효과가 있기 때문에 단순 tree 구조에 비해 단순화의 공간적 정확도를 향상시킨다.

ABSTRACT

A Simple tree-structured line simplification method, which exactly follows the Douglas-Peucker algorithm, has a strength for its simplification index to be involved into the hierarchical data structures. However, the hierarchy of simplification index, which is the core in a simple tree method, may not be always guaranteed. It is validated that the local property of line features in such global approaches as Douglas-Peucker algorithm is apt to be neglected and the construction of hierarchy with no thought of locality may entangle the hierarchy. This study designed a new approach, CALS(Convex hull Applied Line Simplification), a) to search critical points of line feature with convex hull search technique, b) to construct the hierarchical data structure based on these critical points, c) to simplify the line feature using multiple trees. CALS improved the spatial accuracy as compared with a simple tree method. Especially CALS was excellent in case of line features having the great extent of sinuosity.

1. 서 론

컴퓨터 지도학 또는 GIS에서 단순화(simplification) 기법은 폭넓은 의미에서 자료의 중요한 특성들을 결정하고 이를 가능한 한 보존하며 불필요한 상세함(detail)을 제거하는데 목적을 둔다. 그것은 대축척에서 취득된 지도 사상(map features)의 자료를 그보다 소축척으로 표현하고자 할 때 기기(device)의 제한된 해상력으로 원자료의 상세도를 유지하지 못 할 뿐만 아니라 불필요한 처리 시간을 요구하여 효율성을 저하시키고, 고밀도의 자료가 오히려 중요한 정보에 대한 인식과 판단을 방해할 우려가

있기 때문이다. 최근에는 고해상도를 갖는 대축척의 수치지도로부터 상대적으로 소축척의 자료를 추출하여 데이터베이스의 중복을 피하고 비용을 절감하기 위한 목적에서도 단순화 기법은 주목을 받고 있다.¹⁾

지도 사상에 대한 단순화 기법 연구는 선형 사상에 대해서 가장 활발하게 이루어져 왔다. 지금까지 보고된 약 100여개의 선형 사상 단순화 알고리즘은 가능한 한 원래의 기하학적 특성을 최대한 보존하여 원자료의 공간적 정확도 혹은 해상력을 최대한으로 유지하려는 방법론과 가능한 처리 속도를 빠르게 하려는 방법으로 정리된다. 또한 단순화 기법의 유형은 크게 제거(elimination), 원활화(smoothing), 변위(displacement), 강조(enhancement)로 나눌 수 있는데 보통 절차상 제거가 우선시 되며 대부분

서울대학교 국토문제연구소 박사후연수연구원

의 연구가 여기에 집중되어 왔다.²⁾ 특히 그동안 개발된 기법들을 기하학적 특성에 따라 1) 단순 '국지적 단순화(local features)', 2) Reumann & Witkam 기법 · Opheim 기법 · Lang 기법 등의 '확장된 국지적 단순화(extended local features)', 3) Douglas-Peucker 기법 · 프랙탈응용 기법 등의 '전역적 단순화(global features)'으로 분류할 수 있다.^{3,5)} 국내에서는 기존 알고리즘을 이용해 축척을 달리하는 지도 자료를 생성하거나 단순화 알고리즘을 개선시키기 위한 연구가 있었지만 아직 활발하게 연구가 진행되고 있지 않다.^{6,9)}

본 연구는 지도 사상 가운데 선형 사상(linear features)을 대상으로 한 단순화 기법 가운데 지난 25년 동안 가장 일반적으로 이용되고 있는 Douglas-Peucker 알고리즘의 특성을 분석하여 단순화의 공간적 정확도(spatial accuracy)란 측면에서 이를 개선할 수 있는 기법을 개발하는데 그 목적을 둔다. 본 연구를 통해 최근 활발히 논의되고 있는 수치지도간의 축척 변환을 위한 통합적 기법을 개발하는데 그 기초를 제공하는데 기여하고자 한다.

본 연구에서는 수치지도의 선형 사상(line feature)을 대상으로 대축척 자료로부터 그보다 소축척 자료를 추출하는데 적용할 수 있는 단순화 기법(Convex hull Applied Line Simplification, CALS)을 개발하였다. 즉, 단순화 방법 가운데 공간적 정확도가 높고 명료한 축척 변환 지표로 갖고 있는 Douglas-Peucker 방법에 국지적 특성을 고려할 수 있는 형태적 대표점을 찾는 루틴을 첨가하여 이를 계층적 자료구조로 변환시킨 후 이를 이용해 단순화를 진행하는 방법을 이용하였다.

그리고 CALS의 효용성을 평가하기 위해 기존의 대표

적 방법인 Douglas-Peucker 알고리즘을 이용한 단순 Tree 방법과 CALS에 대해서 각각 선의 길이나 변위(displacement)의 크기, 그리고 실제 처리 시간을 측정하였다. 그리고 CALS에 의한 방법을 평가하기 위해 지형적 특성에 따라 개발한 실험적 선 자료와 항공사진에서 해석입체도화기를 통해 수치적으로 도화한 자료에 대해 CALS로부터 추출된 자료를 위의 측정 방법으로 비교하였다.

2. Douglas-Peucker 알고리즘의 원리

Douglas-Peucker 알고리즘¹⁰⁾은 선형사상(line feature)을 구성하고 있는 전체 점(vertices)들을 동시에 고려하는 전역적 단순화 기법(global methods)으로써 단순화 정도를 결정하는 최소기준치의 임계띠(tolerance band, ϵ)를 설정하여 필요 없는 점들을 제거하는 방식을 사용한다. 즉 선형사상에 임계띠를 씌운 후 선분의 양 끝점(v_0, v_n)을 연결한 직선(edge, e)에 대해서 그 사이에 존재하는 점(v_{0+1}, \dots, v_{n-1})들이 직선에 내린 수선의 길이($\delta(v_k, e_{ij})$)를 측정하여 길이가 가장 긴 점을 구한다. 그리고 그 수선의 길이가 임계띠보다 크면 그 점을 다시 새로운 끝점 가운데 하나로 설정(v_0, v_k)하여 위와 같은 과정을 반복(재귀호출)하는 방법으로 임계띠보다 큰 수선을 갖는 점들을 선택하여 선을 단순화시킨다. 도식 1은 Douglas-Peucker 알고리즘의 pseudo-code를 나타낸 것이다.¹¹⁾

이러한 알고리즘을 통해 구체적인 프로그램을 작성하면 다음과 같은 과정에 따라 선형사상이 단순화된다(그림 1). A 단계에서 부동점(anchor point, P1)과 유동점(floating point, P40)을 이은 선분의 양쪽에 임계치 만큼

```

Simplify ( $P_{ij}, \epsilon$ ) {
  If  $j > i+1$ , then {
    Find  $k \in (i, j)$  with  $\delta(v_k, e_{ij}) \geq \delta(v_s, e_{ij}), s \in (i, j)$ ;
    If  $\delta(v_k, e_{ij}) > \epsilon$ , then {
      Add  $v_k$  to the output polyline vertex set  $V$ ;
      Simplify ( $P_{ik}, \epsilon$ ); // recursively handle left sub-polyline
      Simplify ( $P_{kj}, \epsilon$ ); // recursively handle left sub-polyline
    }
  }
}

```

도식 1. Pseudo-code for Douglas-Peucker Algorithm

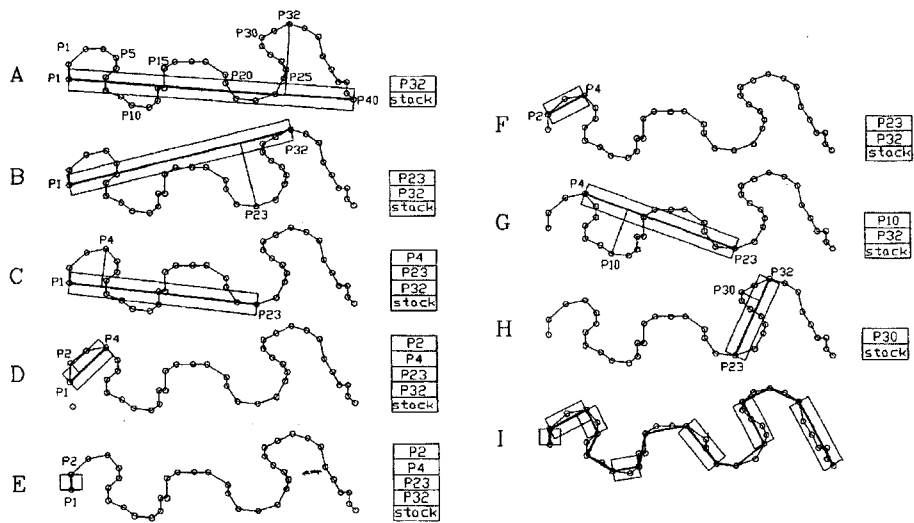


그림 1. Douglas-Peucker 알고리즘을 이용한 프로그램 개발 과정

의 임계치를 씌우고, 중간에 있는 모든 점들(P2-39)이 선분 P1-P40에 내린 수선의 길이를 비교하여 그 최대 길이를 갖는 점을 구한다. 그 점(P32)은 임계치 외부에 존재하므로(즉 임계치를 초과하므로) '제거'에서 제외되며 P32는 스택(stack)에 저장된다. B 단계에서 임계치는 P32를 유동점으로 P1과 P32 사이에 걸쳐지고 A 단계에서의 작업을 반복하여 P23을 구하고 스택에 저장한다. 물론 스택의 특성으로 P23은 P32 위에 위치한다. 이와 같은 방법으로 임계치는 계속 이동하고 C 단계와 D 단계에서 P4와 P2가 각각 스택에 저장된다. E 단계에서는 선의 출발위치에 이르렀기 때문에 비교할 중간점이 없기 때문에 이제 스택에 저장된 점들을 이용하여 탐색이 진행된다.

이렇게 탐색이 반복적으로 진행되어 부동점과 유동점 사이에 중간점이 없거나 중간점들이 모두 임계치 미만인 경우 스택에 저장된 점을 유동점으로 이용하게 된다. F 단계에서 스택의 맨 위에 저장된 P2를 부동점으로 그 아래 저장된 P4를 유동점으로 임계치를 씌웠을 때 중간에 있는 점(최대 수선 길이를 갖는 점)이 임계치 내부에 존재하므로 즉, 임계치 미만이므로 그 점은 제거된다. G 단계에서는 다시 P4를 부동점으로 스택의 맨 위에 저장된 P23을 유동점으로 탐색을 진행하여 P10을 구하고, 이때 수선 길이가 임계치를 초과하므로 P10을 스택에 저장한다. 위와 같은 탐색과 스택의 이용이 반복되어 I 단계에서 보여주는 것과 같이 선사상이 단순화된다.

3. 단순화 과정과 자료구조의 결합

선형 사상의 단순화를 위해 계층적 자료구조를 이용하는 연구는 계층의 수가 정해진 단순한 strip Tree로부터 Buttenfield(1985)에 의해 Douglas-Peucker 알고리즘을 응용한 방법으로 발전하였다.¹²⁾ 그러나 당시까지 개발된 알고리즘은 상당히 복잡하여 실제로 단순화에 이용되기보다 strip Tree에 대한 기하학적 특성에 대한 해석을 통해 단순화 과정에 요구되는 최소기준치(tolerance parameters)의 설정에 필요한 정보를 찾는 데 중점이 되었다.

3.1 단순화 지표의 추출과 Tree 구조 개발

Tree 자료 구조를 사용해 안정되게 Douglas-Peucker 알고리즘을 구현할 수 있는 방법은 Cromley(1991)의 '선형사상의 계층적 단순화' 연구에서 비롯되었다.¹³⁾ 그는 Douglas-Peucker 알고리즘에서 단순화의 지표가 되는 수선의 길이를 계층적인 Tree 구조에 각각의 node에 삽입하여 이를 기초로 단순화를 진행하여 단순화 프로시저의 효율성을 향상시켰다.

단순 Tree 구조의 원리를 이용해 단순화 단계에 따라 그 결과가 자료구조로 조직되는 과정을 분석하면 다음과 같다(그림 2, 3). 우선, 71개의 점으로 구성된 원래의 선형 사상을 대상으로 P1과 P71을 각각 부동점과 유동점으로 잡아 수선의 길이가 최대인 점 P49를 탐색하여 이를 Tree의 root node에 좌표와 수선길이를 저장한다(B).

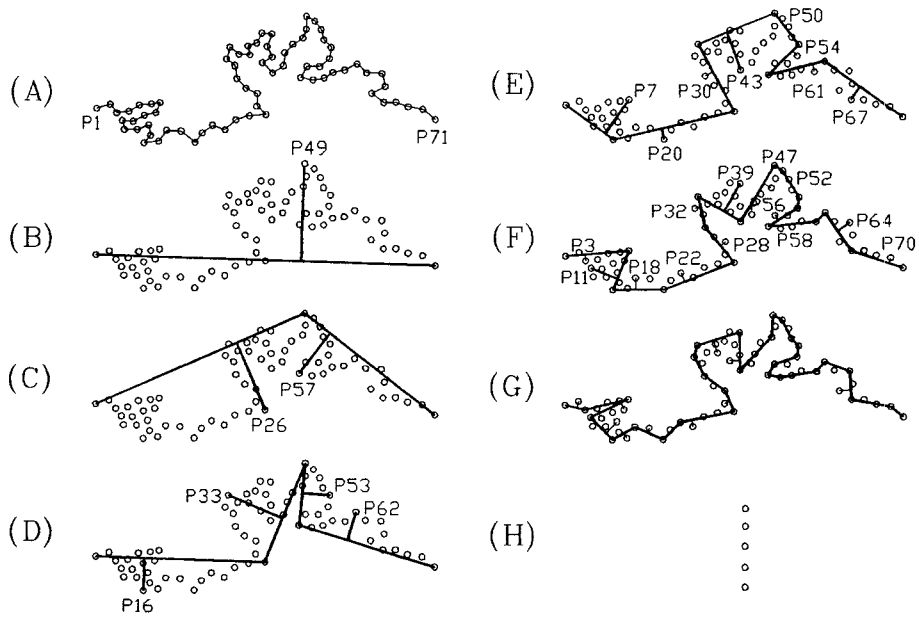


그림 2. Tree 구조를 이용한 Douglas-Peucker 알고리즘

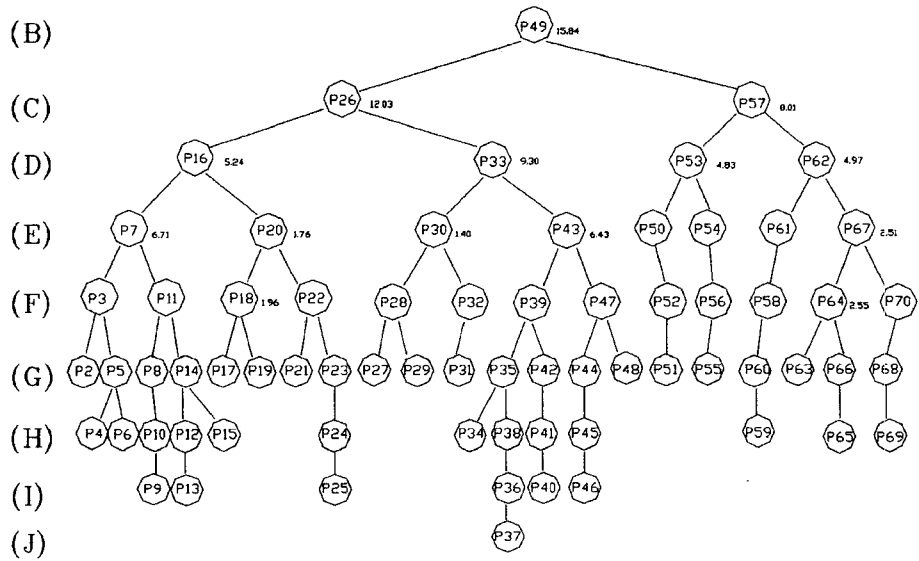


그림 3. 단순화 지표가 삽입된 계층적 Tree 구조

P49를 기준으로 왼쪽의 sub-polyline(P1-P49)와 오른쪽의 sub-polyline(P49-P27)에 대해 각각 수선 길이가 최대인 P26과 P57를 구하여 이진트리의 left child node와 right child node에 삽입한다. 이때 상위의 node에 각 child node를 가리키는 포인터를 첨가한다(C). 선형 사상을 구성하는 모든 점들이 Tree의 node로 구성될 때까지 동일

한 과정을 반복한다(D ... J).

이와 같이 구성된 Tree를 이용하여 단순화시키는 방법은 각 node마다 저장하고 있는 수선의 길이를 임의의 단순화 지표와 비교시키는 방법을 통해 진행한다. 본 연구에서는 계층간 순회순서(traversal order)로 중위순회(inorder) 방법을 이용하였다. 예를 들어, 지표의 수치를

7.0으로 정했다면 7.0 이상의 수선의 길이를 갖는 점을 추출한다. 즉 처음에 P49 node와 비교하면 그 값이 15.84로 7.0을 초과하므로 차하의 왼쪽 node인 P26 node로 그 비교는 진행된다. P26 node 역시 기준치를 초과하므로 다시 차하의 P16 node가 비교된다. 여기서 그 값이 기준치에 미치지 못하므로 같은 계층에 있으면서 P26 node의 오른쪽 차하 node인 P20 node가 비교되며, 같은 방식으로 P33, P30, P43, P57, P53, P62 node가 차례로 비교된다. 그 결과 단순화 지표의 수치를 7.0으로 정했을 때 P26, P33, P49, P57이 추출되고 최종적으로 시작점(P1)과 끝점(P71)이 첨가되어 단순화가 완료된다.

3.2 단순 Tree 구조의 오류

본 연구에서는 단순 Tree 구조에 따라 Douglas-Peucker 알고리즘을 적용하여 계층적으로 선형 사상을 단순화할 때 한 가지 개념적 모순의 가능성을 발견하였다. 그것은 단순화 결과의 정확도에 영향을 미칠 수 있는 요소로 작용한다.

계층적 단순화는 기준치와 계층적으로 저장된 각각의 node를 비교할 때 그 진행과정이 Tree의 계층적 경로(hierarchical path)를 따라 진행된다. 이때 어느 node가 기준치보다 미만일 경우 그보다 차하인 node들은 비교 대상에서 제외되어 그 만큼 프로세스의 효율성을 높인다. 이와 같은 비교 방식에 의한 단순화 방법은 기존의 방법론에 비해 개념적 정의가 명확하다는 장점을 갖고 있다.

다시 말해, 단순 Tree 구조에 따라 Douglas-Peucker 알고리즘을 적용하는 방법은 자료의 저장 형태가 계층성을 갖는다는 ‘기본적 전제’에서 출발하는 것이고, 그것은 각각의 node가 지니는 단순화 지표, 즉 수선의 길이가 자료 구조에 따라 체계적으로 정리되어 저장됨으로써 가능하다. 그런데 여기서 실제로 이러한 기본적 전제가 항상 보장되지 못한다는 사실을 확인하였다. 즉, 차하의 node가 갖는 값이 그보다 차상의 node가 갖는 값에 비해 더 클 경우가 발생할 수 있다. 이 경우 단순화 결과를 추출하는데 그 절차상차상의 node가 기준치보다 작은 값을 가지고 있을 때 차하의 node는 비교 대상에서 제외된다. 그러나, Douglas-Peucker 알고리즘에서 수선의 길이가 갖는 의미는 선형 사상의 기하학적 특성을 대변하여 단순화 결과의 정확도를 좌우한다. 앞의 경우 차상

의 node에 비해 차하의 node가 전체 선형 사상의 기하학적 특성을 더 반영할 수 있다. 실제로 그림 3에서 ‘P16-P7 node’, ‘P20-P18 node’, ‘P67-P64 node’간의 계층성이 어긋난 것을 볼 수 있다.

이와 같은 계층성의 어긋남이 만약 하위 계층에서만 비롯된다면 그 단순화 결과에 대한 기하학적 정확도에 심각한 영향을 미치지 않을 것이고 그것은 모든 단순화 방법에 정도의 차이는 있지만 존재하는 현상으로 판단할 수 있다. 그러나 이런 어긋남이 Tree의 중상위 계층에서 발생하였다면 기하학적 왜곡이 심각할 수 있어 자연히 그 결과에 대한 신뢰도가 떨어지기 마련이다.

기존의 연구는 대부분 단순화의 정확도나 효율성에 초점을 맞춰 진행되었고, 위와 같은 어긋남에 대해서는 논의 없었다. 그렇다면 ‘계층적 단순화 방법’에서 계층상의 어긋남이 발생하는 원인은 무엇인가? 첫째, 전역적 단순화 방법으로 분류되는 Douglas-Peucker 알고리즘은 절차상 단순화가 선형 사상의 전역적 특성(global features)을 고려하는데서 출발하기 때문에 국지적 특성(local features)을 충분히 반영하지 못할 가능성을 내포하고 있다. 둘째, 이렇게 국지적 특성이 고려되지 않은 상태에서 계층적 자료구조에 단순화 지표를 삽입함으로써 자료구조의 계층성에 어긋남이 발생할 수 있다. 결과적으로 Douglas-Peucker 알고리즘과 계층적 자료구조의 장점을 이용하면서 최대한 계층적 어긋남을 극복하기 위해서는 그 자료구조에 선형 사상의 국지적 특성이 반영된 방법이 요구된다.

4. 국지적 특성을 고려한 CALS 개발

본 연구에서는 단순한 수학적 계산을 이용하는 ‘국지적 단순화’나 ‘확대한 국지적 단순화’와 달리 선형 사상에서 이른바 ‘형태적 대표점(critical points)’를 추출하는 방법으로 선형 사상의 국지적 특성을 이끌어 냈다. 즉 선형 사상을 구성하고 있는 점들을 연결할 때 가장 외곽에 위치하는 점, 이른바 convex hull을 구성하는 점을 찾아내는 방법을 개발하고 이를 기초로 형태적 대표점을 추출하였다. 본 연구에서 형태적 대표점을 찾기 위해 convex hull 방법을 시도하는데는 Hershberger & Snoeyink(1992)의 연구가 기본 원리로 참조되었다.¹⁴⁾ 이들은 선형자료를 구성하는 점들을 대상으로 convex hull을 구성할 경우 Douglas-Peucker 알고리즘을 통해 선택되는

점들은 모두 convex hull을 이루고 있는 점들에 포함된다. 이는 사실을 입증하였다.

4.1 형태적 대표점 추출 기법 개발

선형 사상을 구성하고 있는 점들 가운데 형태적인 대표점을 추출하기 위해 본 연구에서는 우선 모든 대상점들에 대해 convex hull을 탐색하는 기법을 개발하였다. 모든 convex hull을 구성하는 점들에는 공통적으로 X, Y 좌표의 최대값과 최소값을 갖는 점들이 항상 존재하며, 시계 반대 방향으로 '최대 X', '최대 Y', '최소 X', '최소 Y' 순서로 연결된다. 본 연구에서는 이와 같은 convex hull의 성질을 이용하여 convex hull을 구성하고 있는 점들을 탐색하였다. 즉, 위의 4개 점들에 대해서 차례로 두 점을 선택하여 서로 연결한 다음, 그 두 점 사이에 존재하는 점들을 하나씩 그 선분과 비교하여 오른쪽에 위치하고 있으면 그 점을 convex hull을 구성하는 점으로 인식하는 방법이다.

Convex hull을 찾는 구체적인 과정은 다음과 같다(그림 4). (A) ① 49개 점(P1-P49)으로 구성된 선형 사상에 대해 행.열의 크기가 같은 matrix를 씌워 각각의 점이 matrix의 어느 위치-셀(cell)-에 존재하는가를 확인한다. 여기서 행.열의 크기는 자료의 수에 따라 적절하게 조절하

는데 본 연구에서는 각 셀이 평균적으로 4개 정도의 점을 갖도록 행.열의 크기를 맞추었다. 여기서 각 셀에 포함되는 점의 수는 임의로 조절이 가능하지만 실험을 통해 처리시간을 측정해보면 셀당 평균 4개 혹은 8개의 점을 분포시킬 때 상대적으로 빠르게 진행되었다. 그것은 셀당 너무 많은 점을 분포시킬 경우 셀당 처리해야 할 경우의 수가 증가하고 너무 적은 점을 분포시키면 전체 matrix를 구성하는 셀의 개수가 증가하여 마찬가지로 프로세스가 비효율적으로 진행되기 때문이다. 한편, matrix의 범위는 점들의 x, y 좌표 중에 최대, 최소인 점, 즉 P28, P11, P1, P43을 기준으로 한다. 그리고 x축 방향으로 열(column)의 번호를 또 y축 방향으로 행(row)의 번호를 부여하고 이를 참조해([열][행]) matrix에서의 위치를 알 수 있다. 예를 들어, matrix의 左下에 존재하는 셀을 기준으로 matrix[0][0], 그 위의 셀은 matrix[0][1] 등으로 위치를 참조하며 따라서 P36의 경우는 matrix[2][0]에 위치한다.

② 선형 사상을 구성하는 점들에 대해서 '최대 x', '최대 y', '최소 x', '최소 y'의 점들을 구하여(P28, P11, P1, P43), 앞서 기술한 방법으로 시계 반대 방향으로 탐색을 진행시켜 나간다. 우선 P28과 P11을 잇는 선분이 통과하는 셀들에 포함되어 있는 점들을 대상으로 그 선

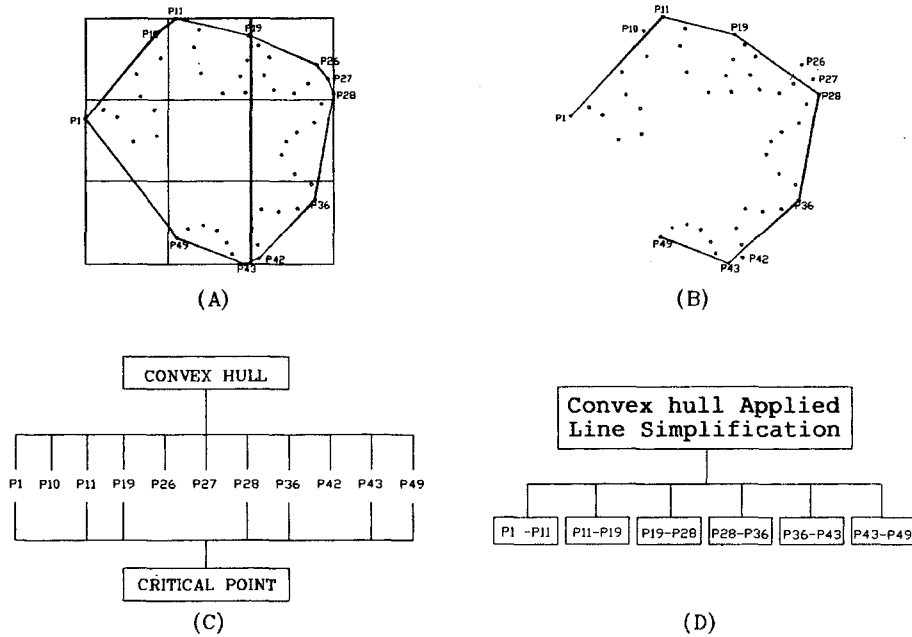


그림 4. 선형 사상의 형태적 대표점 추출

분 오른쪽에 위치한 점들을 찾는다.

③ 위에서 찾은 점들 중에 선분에서 가장 멀리 떨어져 있는 점, P26을 중간점으로 ②의 과정을 반복한다. 즉 P28과 P26 그리고 P26과 P11를 연결한 선분들을 기준으로 이로부터 오른쪽에 가장 멀리 위치한 점들을 찾아 나간다. 이렇게 재귀적(recursive) 처리를 통해서 구한 P1-P10-P11-P19-P26-P27-P28-P36-P42-P43- P49를 연결 하면 선형 사상을 구성하고 있는 모든 점들을 내부에 포함하는 convex hull이 이루어진다.

(B)(C) convex hull을 구성하고 있는 점들 중에 P10-P11, P26-P27-P28, P42-P43은 선형 사상을 구성하는 점들로서 서로 이웃하는 연속적인 성격을 갖고 있다.

(D) 이런 연속적 점들 간에는 그 변위의 차이가 크지 않기 때문에, 선행하는 점들을 생략하고 마지막 점들(P11, P28, P43)만으로도 선형 사상의 국지적 특성을 반영하는 형태적 특성을 갖는 대표점을 충분히 파악할 수 있고, 이를 기초로 조직되는 계층적 자료구조는 그 복잡성이 현저하게 줄어든다. 또한 선행 점을 생략함으로써 발생할 수 있는 계층상의 어긋남의 가능성은 최하위 계층에서 일어나기 때문에 단순화 결과에 크게 영향을 주지 않는다.

4.2 국지적 Tree 구축

선형 사상의 형태적 대표점을 통해 구축되는 복합적 자료구조는 전체적인 선형 사상에 대해 구축하는 계층적 자료구조에서 발생할 수 있는 계층상의 어긋남-특히, 중

상위 계층-을 최대한 억제할 수 있다. 이와 같은 형태적 대표점은 단순화 정도에 상관없이 최소한 유지해야 하는 점으로 인식할 수 있다. 또한 이 대표점은 축척의 변동으로 지도의 선형 사상이 크게 변동 혹은 왜곡될 경우 기하학적 ‘위기적 변화(catastrophic change)’를 유발하는 축척을 확인할 수 있는 기초로 활용될 수도 있다.¹⁵⁾

그림 4의 (D)는 원래 49개 점으로 구성된 선형 사상을 계층적으로 조직하는데 최소한 유지해야 할 점으로 7개 점을 이용하고 있다. 그림 5는 단순 Tree 구조에서 이용한 선형 사상(그림 3)에 대해 국지적 인자가 중간 단계에 삽입한 결과이다. 단순 Tree 구조에서 발생하였던 중상위 계층에서의 계층성의 어긋남이 모두 제거되었다. 다만, 하위 계층에서 어긋남(P44-P45)이 하나 발생한 것을 확인할 수 있다. 따라서 형태적 대표점을 중간에 포함시켜 단순화 Tree를 구성하는 이 방법은 단순화 결과 상대적으로 기하학적 공간적 정확도를 크게 감소시킬 수 있는 중상위 계층의 오류를 제거하는데 그 핵심이 있다. 따라서 P44-P45에서 볼 수 있듯이 완벽하게 Tree의 하위 계층에서 발생하는 오류는 제거할 수 없지만 공간적 정확도란 측면에서 그 영향이 미미하기 때문에 전반적인 정확도를 개선시킬 수 있다.

5. CALS 방법의 평가

본 절에서는 단순 Tree 구조를 이용한 계층적 단순화

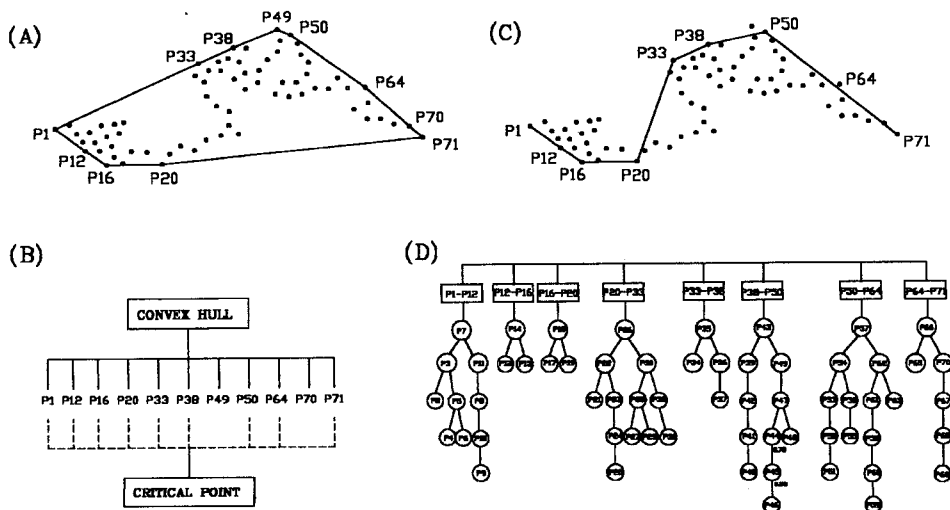


그림 5. 국지적 특성을 고려한 자료구조

방법에서 계층성의 어긋남, 특히 중상위 계층의 어긋남을 제거하기 위해 본 연구에서 제안한 ‘볼록다각형 생성 기법을 응용한 선형 단순화(Convex hull Applied Line Simplification, 이하 CALS라 칭함)’의 평가를 위해 단순화 결과의 위치적 정확도를 척도로 Douglas-Peucker 알고리즘을 그대로 적용한 단순 Tree 구조 방법과 비교하였다.

5.1 축척변환 지표와 단순화 정확도 측정

상대적으로 대축척인 자료원으로부터 원하는 소축척의 자료를 추출하기 위해서 CALS나 단순 Tree 방법 모두 적절한 단순화 지표 혹은 축척 지표를 요구하는데 이를 구하기 위해 Töpfer의 법칙을 응용하였다. Töpfer의 법칙은 상대적으로 대축척인 자료원으로부터 원하는 소축척의 자료를 추출하기 위해 자료의 수를 결정하는데 지난 30여년 동안 일반적으로 사용되고 있는 경험 법칙이다.¹⁶⁾ 즉 대축척 지도에서 나타나는 사상의 수와 상대적으로 소축척인 지도에서 나타나는 사상의 수에는 다음과 같은 관계가 있음을 나타내는 것이다.

$$N_f = N_a * \sqrt{M_f / M_a}$$

N_a : 대축척(M_a)에서 사상의 수

N_f : 소축척(M_f)에서 사상의 수

본 연구에서는 이와 같은 법칙을 응용하여 대축척에서 선형 사상을 구성하고 있는 점의 수를 참조하여 이보다 소축척으로 그 선사상을 표현하고자 할 때 유지해야 하는 점의 수를 구하였다. 그 다음 이 기준에 근사하도록 CALS과 단순 Tree 방법에서 지표를 조절하여 소축척의 자료를 추출하였다. 이때 Tree에 저장된 단순화 지표(수선의 길이)와 비교하기 위한 기준값을 사용자가 입력하거나 자동으로 적절한 단위만큼 증가시킬 수 있다. 본 연구에서는 임의로 증가치를 설정하여 프로그램을 반복(iteration)하여 Töpfer의 계산식에서 산출된 자료수에 가장 근사하도록 기준값을 선정하였다. 따라서 증가치의 정밀도는 Töpfer의 법칙에서 구한 소축척 자료의 수와 실제 추출된 자료의 수의 일치정도에 영향을 미칠 수 있다. 또한 Tree의 node에서 동일한 지표값이 나타날 경우 Töpfer의 법칙에서 계산된 점의 수와 약간의 차이가 발생할 수 있다.

한편 CALS를 평가하기 위해 단순화 결과의 공간적 정확도(spatial accuracy)를 측정하였다. 선형 단순화의 정확

도는 크게 선이 갖는 고유한 특성(길이, 각도, 굴곡도 등)을 정량적으로 측정하는 방법과 원래 사상과 단순화된 사상 사이에 생긴 변위(displacement)를 측정하는 방법으로 분류할 수 있다.¹⁸⁾ 본 연구에서는 이 가운데 원래 선형 사상과 단순화된 사상을 중첩할 때 생기는 다각형의 총면적을 계산(areal displacement)하는 방식을 채택하였다(그림 6). 이와 함께 실제 계산상의 효율성을 평가하기 위해 두 가지 방법에 대해 CPU 처리속도를 비교하였다.

5.2 실험 및 평가

본 절에서는 선형사상을 지형적 특성(geomorphic qualities)에 따라 구분한 실험적 자료와 실제 해석입체도를 통해 얻은 자료에 대해 CALS를 적용하여 이를 평가하였다. 전자의 경우 선형 사상을 ‘굴곡도’와 ‘곧은 정도’ 그리고 각의 ‘예리한 정도’나 ‘완만한 정도’ 등을 기준으로 현재 발간된 여러 가지 형태의 지도를 임의로 선정하여 디지털타이핑(manual digitizing) 절차를 통해 개발하였다(표 1, 그림 7).

선형 사상의 유형을 6가지로 구분하여 단순화의 정확도를 평가한 결과, 대체적으로 CALS로부터 추출한 자료가 단순 Tree 방법을 이용한 자료에 비해 변위가 적게 나타났고, 그 처리 시간에 있어서는 양자 사이에 별다른 차이가 없었다.

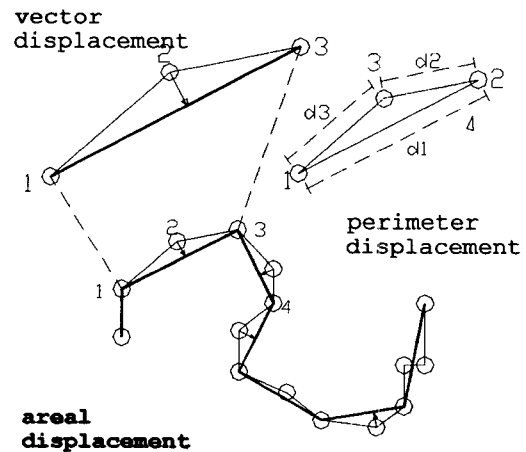


그림 6. 단순화 기법의 정확도 평가: 변위측정

표 1. 실험적 선형사상의 분류

	굴곡진 선	곧은 선	폐곡선
완만한 선	line 1	line 3	line 5
예리한 선	line 2	line 4	line 6

형태적인 면에서 전반적으로 굴곡이 있는 line 1과 line 2-원래 축척이 1:5,000에 대해서 단순 Tree 방법과 CALS로 1:25,000과 1:50,000에 해당하는 자료를 추출했을 때 변위의 크기를 비교해보면, CALS가 단순 Tree 방법에 비해 각각 1.2%, 1.9%, 10.5%, 6.7% 정확도가 향상되었음을 알 수 있다(표 2, 3). 또한, CALS에는 형태적 대표점을 찾는 루틴이 첨가되었음에도 line 1에서는 단순 Tree 방법보다 더 효율적이었다. 그것은 단순 Tree

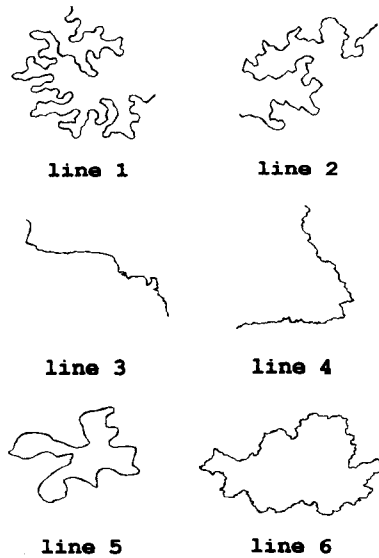


그림 7. 실험적 선형 사상

표 2. line 1의 정량적 평가

	1:5,000	1:25,000		1:50,000	
	자료원	단순tree	CALS	단순tree	CALS
점의 수	133	58	58	42	42
선의 길이(m)	6848.9	6460.5	6497.3	6138.0	6158.7
변위 크기(m ²)		426.4	421.1	595.2	583.8
처리 시간(초)		0.479	0.440	0.443	0.420
Töpfer		60		42	

표 3. line 2의 정량적 평가

	1:5,000	1:25,000		1:50,000	
	자료원	단순tree	CALS	단순tree	CALS
점의 수	148	66	66	46	46
선의 길이(m)	5165.2	4941.1	4951.3	4713.1	4724.5
변위 크기(m ²)		204.9	183.4	329.8	307.6
처리 시간(초)		0.531	0.531	0.495	0.495
Töpfer		66		47	

방법으로 단순화를 수행하려 할 때 계층 구조가 복잡하여 이를 구성하는데 더 많은 계산이 요구되고 또한 계층의 깊이(depth of Tree)가 깊어져 탐색하기 위한 시간이 그만큼 지연되기 때문이다.

선형 사상이 전체적으로 곧은 특성을 갖고 있는 line 3과 line 4는 1:25,000 지형도에 나타난 도로 사상을 입력한 것이다. 원자료의 축척으로부터 1:100,000과 1:200,000에 해당하는 자료를 두 가지 방법으로 추출하였을 때, CALS는 단순 Tree 방법에 비해 변위의 크기 면에서 line 3의 1:100,000 자료에서 4.2% 향상되었고 나머지 세 경우에는 각각 2.7%, 1.2%, 1.9%씩 정확도가 낮아졌다(표 4, 5).

그리고 폐곡선에 대한 평가를 위해, 완만하지만 굴곡이 크게 형성된 line 5와 각이 예리한 반면 굴곡은 크지 않은 line 6을 각각 1:5,000의 토양도와 행정구역도에서 선택하여, 1:50,000, 1:100,000, 1:200,000의 자료로 추출하였다. line 5의 경우 CALS가 단순 Tree 방법에 비해 위치적 정확도가 10.6% 와 7.3% 향상되었음을 알 수 있다. 또, line 6의 경우 1:50,000 자료에서 0.9% 감소한 반면 1:200,000 자료에서는 2.7% 향상되었다(표 6, 7).

위와 같은 비교를 통해서 전체적으로 선형 사상에 굴곡이 있는 경우에 CALS로부터 추출한 자료는 단순 Tree 방법에 비해 단순화의 위치적 정확도가 크게 향상됨을 확인하였다. 그것은 단순 Tree의 계층적 선형 사상

표 4. line 3의 정량적 평가

	1:25,000	1:100,000		1:200,000	
	자료원	단순tree	CALS	단순tree	CALS
점의 수	103	53	53	36	36
선의 길이(m)	18634.2	18373.9	18351.8	18072.0	18015.9
변위 크기(m ²)		1651	1582	2860	2901
처리 시간(초)		0.348	0.440	0.330	0.421
Töpfer		52		36	

표 5. line 4의 정량적 평가

	1:25,000	1:100,000		1:200,000	
	자료원	단순tree	CALS	단순tree	CALS
점의 수	121	62	62	43	43
선의 길이(m)	22750.0	21990.0	22003.1	21475.7	21495.3
변위 크기(m ²)		2088	2149	4233	4392
처리 시간(초)		0.421	0.440	0.403	0.424
Töpfer		61		43	

표 6. line 5의 정량적 평가

	1:5,000		1:50,000		1:100,000	
	자료원	단순tree	CALS	단순tree	CALS	
점의 수	141	44	44	32	32	
선의 길이(m)	1822.7	1779.2	1782.9	1752.0	1758.7	
변위 크기(m ²)		276.7	247.3	572.1	530.2	
처리 시간(초)		0.440	0.476	0.421	0.458	
Töpfer		45		32		

표 7. line 6의 정량적 평가

	1:5,000		1:50,000		1:200,000	
	자료원	단순tree	CALS	단순tree	CALS	
점의 수	293	92	92	47	47	
선의 길이(m)	3422.7	3165.0	3152.8	2975.4	2969.3	
변위 크기(m ²)		552.7	557.5	1070.0	1041.5	
처리 시간(초)		0.934	0.916	0.897	0.879	
Töpfer		93		46		

의 단순화 방법에서 계층적 어긋남이 단순화 결과의 위치적 정확도에 영향을 미치고 있음을 암시한다. 반면에 계층적 어긋남 특히 중상위 계층에서 발생하는 어긋남이 제거된 CALS 방법에서는 위치적 정확도가 향상되었다. 한편 처리 시간을 비교한 결과 CALS에 형태적 대표점을 찾는 루틴이 첨가되었지만 대체로 큰 차이를 발견할 수 없었고, 단순 Tree의 계층적 단순화에서 계층의 구성이 복잡하고 계층의 깊이가 깊은 경우에 있어서는 CALS가 오히려 프로세스가 빨리 진행되었다.

다음은 CALS의 실용적 이용 가능성을 살펴보기 위해 해석입체도화기를 이용하여 사진 축척이 1:20,000인 항공 사진을 기초로 각각 1:5,000, 1:10,000, 1:20,000의 등고선 자료(고도 250 m, 350 m 등고선 2개, 그림 8)를 수치도화한 다음, 1:5,000 수치도화자료로부터 CALS를 적용하여 1:10,000과 1:20,000 자료를 추출하여 양자를 비교하였다. 도화에는 ASP 2000 analytical stereoplotter (Kappa Rotation:±15°, 해상력:1 μm)를 이용하였고, 항공 사진은 1988년 10월에 촬영된 1:20,000 흑백 사진으로 사진번호는 안동지구 24-57, 24-58이며, 도화자는 경력 10년 이상된 숙련자이다.

표 8과 9에서 1:10,000으로 도화한 자료의 수가 Töpfer 법칙에 따른 수보다 상대적으로 적게 나타나는 것을 감안하더라도 CALS로부터 추출된 자료에 비해 위치적 정확도가 많이 떨어지는 것을 알 수 있다. 더욱이

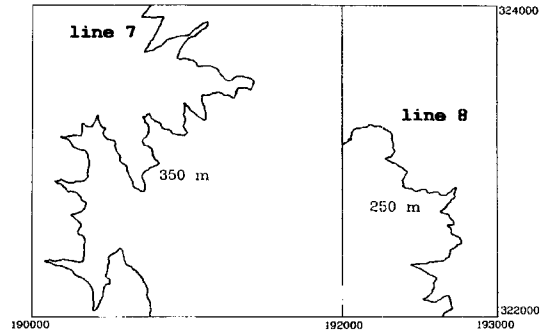


그림 8. 해석도화기를 이용해 추출한 등고선 자료 (1:5,000)

표 8. line 7의 정량적 평가

	1:5,000		1:10,000		1:20,000	
	자료원	도화자료	CALS	도화자료	CALS	
점의 수	621	353	435	275	296	
선의 길이(m)	8005.1	7831.7	7995.7	7803.4	7963.2	
변위 크기(m ²)		21651.4	1009.8	24400.3	3201.4	
Töpfer		439		311		

표 9. line 8의 정량적 평가

	1:5,000		1:10,000		1:20,000	
	자료원	도화자료	CALS	도화자료	CALS	
점의 수	289	159	194	123	127	
선의 길이(m)	2942.5	2818.1	2936.5	2856.0	2920.6	
변위 크기(m ²)		7476.4	407.9	9557.0	1332.8	
Töpfer		204		145		

line 8에 대한 실험에서는 Töpfer의 계산식에서 산출된 점의 수(145)보다 오히려 도화자료의 수(123)에 맞게 CALS에 의한 점의 수를 조절하여 비교하였는데, 자료수가 비슷함에도 불구하고 선의 길이나 변위의 크기를 살펴볼 때 CALS에서 취득한 자료가 정확도가 우수함을 확인할 수 있다. 이와 같이 도화자료와 CALS를 통해 추출한 자료의 직접 비교한 결과는 어느 정도 예상을 할 수 있지만, 대축척 자료에서 자동적인 방법에 의해 소축척 자료를 추출하기 위한 향후 연구에 정확도란 측면에서는 그 기법의 적용에 무리가 없음을 시사하고 있다는 데 의의가 있다.

6. 결 론

본 연구는 대표적 선형 사상에 관한 단순화 기법으로

인정받고 있는 Douglas-Peucker 알고리즘의 분석을 통해 알고리즘의 특성상 선형 사상의 국지적 특성이 충분히 반영되지 못할 가능성을 밝혔다. 그리고 이와 같은 문제를 형태적 대표점을 탐색하여 이를 기초로 단순화 Tree를 구축함으로써 개선할 수 있는 기법을 개발하였다. 본 연구를 통하여 얻은 결과를 요약하면 다음과 같다.

첫째, 단순 Tree는 선형 사상을 단순화하는데 Douglas-Peucker 알고리즘을 계층적 자료구조로 구현함으로써 단순화 과정을 개념적으로 명확히 하고 효율성을 향상시켰다. 그러나 본 연구에서는 단순 Tree 방법에서 핵심적이라 할 수 있는 ‘단순화 지표의 계층성’이 항상 보장되지 못한다는 점을 발견하였고, Douglas-Peucker 알고리즘이 국지적 특성을 충분히 반영하지 못한다는 특성에서 그 이유를 찾았다.

둘째 중상위 계층에서 발생하는 계층성의 어긋남을 보정하기 위해 선형 사상의 국지적 특성을 반영할 수 있는 형태적 대표점을 convex hull 탐색기법을 응용하여 탐색한 다음 이를 기초로 계층적 자료구조를 완성시켰다. CALS에서 수행한 단순화 결과는 단순 Tree 방법에 비하여 대체적으로 공간적 정확도가 향상되었다. 특히, 선형 사상의 유형에 따른 검증을 통해서 CALS은 특히 굴곡이 많은 자료에 대해 단순화 후의 공간적 정확도를 크게 향상시킨다는 사실을 밝혔다. 이로부터 선형 사상의 굴곡의 정도가 커짐에 따라 단순 Tree 방법에서 계층적 어긋남이 발생할 가능성이 높고 이것은 단순화 결과의 위치적 정확도에 영향을 미치고 있음을 확인하였다. 또한 CALS은 처리 시간에 있어서도 단순 Tree 방법에 비해 크게 뒤지지 않았고, 오히려 계층적 구조가 복잡한 경우에 있어서는 CALS가 더욱 신속한 처리속도를 보였다.

셋째, 도화자료(1:5,000)로부터 CALS를 이용하여 취득한 자료들(1:10,000과 1:20,000)은 항공사진으로부터 해당 축척으로 직접 도화한 자료들보다 공간적 정확도가 확실히 높다. 따라서 대축척의 자료를 수치지도 데이터 베이스로 구축하여 이보다 소축척인 자료를 신속하게 취득함으로써 축척에 따라 자료를 입력함으로써 발생하는 비용과 자료의 중복을 제거할 수 있는 가능성을 제시하였다.

참고문헌

1. 국방과학연구소, 표준지도 생성기법 연구, 응용연구과제

- 3단계 보고서, 1998, 71-84.
2. João, E. M., Causes and Consequences of Map Generalisation, Taylor & Francis: Bristol, 1998, 35-58.
3. McMaster, R. B., “The Integration of Simplification and Smoothing Algorithms in Line Generalization”, Cartographica, Vol. 26, No. 1, 1989, 101-121.
4. Robinson, A. H., Morrison, J. L., Muehrcke, P. C. Kimmerling A. J., and Guptill, S. C., Elements of Cartography, (6th ed.) John Wiley & Sons: New York, 1995, 462-472.
5. Opheim, H., “Fast Data Reduction of a Digitized Curve”, Geo-Processing, Vol. 2, 1982, 33-40.
6. 황철수, “다축척 수치지도 구축을 위한 선형 사상의 일반화에 관한 연구”, 지리학논총, 제21호, 1993, 17-34.
7. 김감래, 이호남, “선형성 지형자료의 일반화에 대한 효율적인 알고리즘에 관한 연구”, 한국측지학회지, 제12권, 제1호, 1994, 43-52.
8. 박경렬, “일반화를 통한 소축척 수치지도 제작방안”, 한국측지학회 '98 세미나 발표회집, 1998, 63-93.
9. 김두일, 김종석, “선형사상에 따른 단순화 알고리즘의 반응 특성 연구”, 대한지리학회지, 제33권, 제4호, 1998, 623-634.
10. Douglas, D. H. and Peucker, T. K., “Algorithms for The Reduction of The Number of Points Required to Represent a Digitized Line or its Caricature”, The Canadian Cartographer, Vol. 10, No. 2, 1973, 112-122.
11. Saalfeld, A., “Topologically Consistent Line Simplification with the Douglas-Peucker Algorithm”, Cartography and Geographic Information Science, Vol. 26, No. 1, 1999, 7-18.
12. Buttenfield, B., “Treatment of the Cartographic Line”, Cartographica, Vol. 22, No. 2, 1985, 1-26.
13. Cromley, R. C., “Hierarchical Methods of Line Simplification”, Cartography and Geographic Information Systems, Vol. 18, No. 2, 1991, 125-131.
14. Hershberger, J. and Snoeyink, J., “Speeding Up the Douglas-Peucker Line Simplification Algorithm”, Proceedings of the 5th International Symposium on Spatial Data Handling, Charleston: SC, 1992, 134-143.
15. Muller, J. C., “Digital Terrain Modelling”, in Maguire, D. J., Goodchild, M. F., and Rhind, D. W. (eds.) Geographical Information Systems: Principles and Applications, Longman:London, 1991, 269-297.
16. Dutton, G., “Scale, Sinuosity, Point Selection in Digital Line Generalization”, Cartography and Geographic Information Science, Vol. 26, No. 1, 1999, 33-53.
17. Töpfer, F. and Pillewizer, W., “The Principle of Selection”, The Cartographic Journal, Vol. 3, 1966, 10-16.

18. McMaster R. B., "Automated Line Generalization",
Cartographica, Vol. 24, No. 2, 1987, 74-111.
19. Breward, R. W., "A Mathematical Approach to The

Storage of Digitised Contours", The Cartographic
Journal, Vol. 9, 1972, 82-86.