수치지형도 일반화를 위한 도로 네트워크 데이터의 선택 기법 연구

The Selection Methodology of Road Network Data for Generalization of Digital Topographic Map

박우진¹⁾ · 이영민²⁾ · 유기윤³⁾

Park, Woo Jin · Lee, Young Min · Yu, Ki Yun

Abstract

Development of methodologies to generate the small scale map from the large scale map using map generalization has huge importance in management of the digital topographic map, such as producing and updating maps. In this study, the selection methodology of map generalization for the road network data in digital topographic map is investigated and evaluated. The existing maps with 1:5,000 and 1:25,000 scales are compared and the criteria for selection of the road network data, which are the number of objects and the relative importance of road network, are analyzed by using the Töpfer's radical law and Logit model. The selection model derived from the analysis result is applied to the test data, and the road network data of 1:18,000 and 1:72,000 scales from the digital topographic map of 1:5,000 scale are generated. The generalized results showed that the road objects with relatively high importance are selected appropriately according to the target scale levels after the qualitative and quantitative evaluations.

Keywords : Digital Topographic Map, Road Network, Map Generalization, Selection and Elimination, Töpfer's radical law, Logit Model

초 록

지도 일반화 기법을 이용하여 대축척 지도자료로부터 소축척 지도자료를 생산하기 위한 방법론 개발은 수치지형도의 제작, 갱신 등의 관리에 있어서 매우 중요하다. 본 연구에서는 수치지형도의 도로와 같은 네트워크 형태의 객체를 일반화하기 위한 하나의 단계인 선택 기법을 제안, 적용하였다. 이를 위해, 기존의 1:5,000 축척과 1:25,000 축척의 수치지형도를 상호 비교하여 도로 네트워크 객체의 선택과 관련된 기준(선택 객체의 개수, 상대적 중요도)들을 Töpfer의 radical 법칙과 Logit 모형을 이용하여 분석하였다. 여기서 분석된 결과를 바탕으로 하여 테스트 데이터에 대해 선택 모델을 적용하여 1:5,000 수치지형도 도로중심선 레이어로부터 일반화된 1:18,000, 1:72,000 축척의 네트워크 데이터셋을 도출하였다. 일반화된 결과에 대하여 정성적, 정량적 평가를 실시한 결과, 상대적으로 높은 중요도를 가진 네트워크 객체들이 목표 축척수준에 맞게 적절히 선택된 결과를 나타내었다.

핵심어 : 수치지형도, 도로 네트워크, 지도 일반화, 선택 및 삭제, Töpfer의 radical 법칙, Logit 모형

1. 서 론

웹이나 모바일 환경에서의 지도서비스, 위치기반서비스에

있어서 도로, 하천, 철도와 같은 네트워크 데이터는 중요도가 매우 높은 데이터라고 할 수 있다. 네트워크 데이터를 다양한 축척수준으로 디스플레이를 하거나 경로를 탐색하는데

1) Member, Department of Civil & Environmental Engineering, Seoul National University(E-mail: woojin1@snu.ac.kr)

2) Department of Civil & Environmental Engineering, Seoul National University(E-mail: daldanka@snu.ac.kr)

3) Corresponding Author, Member, Department of Civil & Environmental Engineering, Seoul National University(E-mail: kiyun@snu.ac.kr)

있어서 다양한 축척으로 구축된 네트워크 데이터는 매우 유용하다. 다양한 축척수준의 지도 데이터셋을 구축하기 위해서는 원 지도자료에 지도 일반화(map generalization) 기법을 적용할 필요가 있다. 네트워크 형태의 선형 데이터에 대한 지도 일반화는 선택(selection), 재분류(reclassification), 단순화(simplification) 등의 기법을 적용해야 한다(Choe and Kim, 2007). 본 연구에서는 수치지형도의 도로 네트워크 객체에 대한 선택 기법을 제안하고 이를 실제 수치지형도에 적용하고자 한다.

선택 연산자(selection operator)는 해당 축척수준에서 지형 지물의 크기가 작거나 중요도가 낮아 해당 축척수준의 지도 상에서 굳이 표현할 필요가 없는 객체들을 선택하여 삭제시키는 과정이다. 여기에서는 어떤 객체를 몇 개 정도로 남길 것인가가 가장 핵심적인 문제이다. 어떤 객체를 선택하고 삭제할 것인가는 객체의 기하학적 특성, 의미론적 특성, 객체들 간의 분포 특성 등을 고려하여 객체의 중요도를 측정함으로써 이루어질 수 있다. 몇 개의 객체를 선택하고 삭제할 것인가는 지도의 목표 축척수준, 목적, 생산자의 의도 등에 따라 결정되는데, 일반적으로 Töpfer의 radical 법칙, 최소도화객체(Minimum Mapping Unit, MMU)⁴⁾ 크기를 이용한다.

Töpfer가 주장한 radical 법칙에서는 지도의 축척수준과 객체 개수와의 관계를 여러 국가 지형도를 분석하여 귀납적으로 규명하였다. Töpfer와 Pillewizer(1966)는 지도 제작 상 심볼들의 과장을 고려한 적정 객체 개수에 대한 연구를 발전시켰다.

도로망도 일반화 기법에 대해서는 Thomson과 Richardson(1999)이 그래프 기반 측정으로 arc의 상대적 중요도를 측정해서 데이터 축소를 위한 정보로 활용한 바 있다. 즉, 경로탐색에서 빈번하게 이용되는 arc에 높은 중요도를 부여하였으며 제거되어야 할 도로의 적정 수는 Töpfer의 radical 법칙을 이용하였다. Machaness와 Mackechnie(1999) 역시, 그래프 기반의 군집화 기술을 이용해서 네트워크를 단순화하는 기법을 제안하였으며, 그래프 기반 측정기법을 통해 node와 arc의 중요도를 정량화하였다. Liu 등(2010)은 도로 네트워크 일반화를 위한 도로 선택을 하는데 있어서 도로의 4가지 정보(통계적, 계량적, 위상적, 주제적)를 보로노이(Voronoi) 다이어그램과 스트로크(stroke)에 기초하여 산정한 후, 선택 기준으로 활용하는 연구를 수행하였다. Touya(2010)는 도로 네트워크를 선택하는데 있어서 지리적 문맥과 구조적 특성을 분석하

여 이를 참고자료로 활용하는 선택 기법을 제안하였으며, 비도시 지역의 도로는 그래프 이론, 도시 지역의 도로는 블록 병합 알고리즘을 적용하여 도로를 선택하였다. 박우진 등(2012)은 Töpfer의 radical 법칙의 심볼과장계수와 심볼형태계수를 적절히 조절하여 도로 데이터의 선택 및 삭제 수준을 조절하는 기법을 연구한 바 있다.

위의 연구사례들을 살펴보면, 기존의 네트워크 객체 선택 기법은 대부분 객체의 다양한 특성 정보(형상적, 위상적, 주제적)를 종합적으로 고려하여 목표 축척수준에서 도화되거나 삭제될 객체를 선택한다. 이 과정에서 각 특성에 대한 임계조건에 만족하지 않는 객체를 삭제하거나, 객체의 상대적 중요도를 계산한 후, 중요도가 낮은 객체를 삭제하는 방식을 주로 쓰고 있다. 이때, 객체의 선택 과정에서 고려되는 여러 파라미터들은 선택에 미치는 영향력에 차이가 있을 수 있다. 또한 파라미터 간에 높은 상관관계가 나타날 경우, 비합리적인 결과가 도출될 수도 있다. 따라서 파라미터 간의 상관성 또는 선택 결과에 미치는 상대적 영향력을 분석하여 이를 선택 규칙에 적용할 필요가 있다. 그러나 선택에 미치는 각 파라미터들의 상대적 영향력에 대한 정량적인 분석 및 적용에 대한 연구는 아직 미진한 상태이다.

Fig. 1. Detailed flow chart of selection methodology

4) 최소도화객체(Minimum Mapping Unit, MMU) : 주어진 축척에서 지도 상에서 표현할 수 있는 객체의 최소 크기. 주로 객체의 면적, 또는 객체의 폭 등으로 표현됨.

따라서, 본 연구에서는 네트워크 객체의 선택에 미치는 요소들과 영향력을 정량적으로 분석하기 위해 기존의 수치지형도 1:5,000과 1:25,000 축척에 대한 도로중심선 데이터를 중첩하여 축척별 도로 네트워크 객체 간의 매칭쌍을 탐색하고 매칭된 객체들에 대한 형상적 특성값들을 통계적으로 분석하였다. 여기서 도출된 분석결과를 바탕으로 하여 네트워크 객체 선택의 기준이 되는 속성값들에 대해 확률함수 형태의 선택 모형을 도출하였다. 또한 이를 테스트용 수치지형도에 적용하여 목표 축척수준에 맞게 도로 네트워크 객체를 선택 또는 삭제하였다. 마지막으로, 목표 축척수준으로 선택된 결과 데이터에 대해서 정량적, 정성적인 평가를 실시하였다. 본 연구에서 제안한 방법론의 상세 흐름도는 다음 그림과 같다(Fig. 1).

2. 대상데이터 및 목표 축척수준

국토지리정보원에서는 수치지형도를 다양한 축척(1/1,000, 1/2,500, 1/5,000, 1/25,000)으로 제작, 관리하고 있으며 수치지형도의 제작 및 갱신은 항공측량, 지리조사, 현지보완측량을 통해 디지털라이징 방식으로 이루어진다. 본 연구에서의 대상데이터는 수치지형도 Ver. 2.0의 도로중심선(A002) 데이터이며, 도로경계(A001) 데이터를 보조데이터로 활용하였다(NGII, 2010).

지도 일반화를 통해 소축척 지도를 생성하기 위해서는 지도자료를 일반화하고자 하는 목표 축척수준을 설정해야 한다. 축척수준의 결정은 원지도자료의 축척, 지도의 목적, 자주 사용되는 축척, 업데이트 주기, 자동화 정도 등을 고려하여 결정하여야 한다. 또한 목표 축척은 가장 기본적인 수준에서부터 점진적으로 높은 수준까지 만들어져야 한다. 본 연구에서의 목표 축척수준은 수치지형도를 웹 또는 모바일 환

경에서 지도서비스 형태로 제공하는 것을 전제로 하여 산정하였다. Google Map과 같은 인터넷 지도 서비스에서는 지도 자료의 웹서비스를 위한 타일 스키마(tile scheme)를 정의하고 있는데 여기에는 타일의 ID, 크기, 해상도를 비롯하여 여러 단계의 확대/축소 레벨에 따른 축척수준을 설정해 놓고 있다(MicroImages, Inc, 1999). Google Map에서 제공하는 타일 스키마에서의 줌인 및 줌아웃 레벨별 축척수준은 아래의 표와 같다(Table 1).

본 연구에서는 위의 Google Map 타일 스키마를 바탕으로 하여 다축척 네트워크 데이터셋의 목표 축척수준을 1:5,000(Level 17), 1:18,000(Level 15), 1:72,000(Level 13)의 3단계로 설정하였다. 이러한 목표 축척수준은 모델 일반화(model generalization) 프로세스의 결과물이라고 할 수 있으며, 이를 바탕으로 Level 18부터 Level 13까지 6단계로 확대 또는 축소할 수 있는 다축척 네트워크 데이터를 생성하는 것을 목표로 하였다.

목표 축척수준 1:5,000에 대해서는 기구축된 수치지형도와 같은 축척수준이므로 기존 자료를 그대로 활용하였으며 1:18,000과 1:72,000 축척수준은 1:5,000 축척의 수치지형도에 지도 일반화를 적용해서 새로 구축하고자 하였다. 다음의 표는 본 연구에서 설정한 목표 축척수준과 그에 대한 특징을 정리한 표이다(Table 2).

Table 2. Target scale level and characteristics

Target scale level	Zoom level	Characteristics
1:5,000	Level 18	One or two of administrative dong are visible on the mobile display. Raw data of digital topographic map of 1:5,000 is used.
	Level 17	
1:18,000	Level 16	Four or five of administrative dong are visible on the mobile display. Digital topographic map of 1:5,000(raw data) is generalized to 1:18,000 scale.
	Level 15	
1:72,000	Level 14	One or two of administrative gu on the mobile display are visible Digital topographic map of 1:5,000(raw data) is generalized to 1:72,000 scale.
	Level 13	

Table 1. Tile schema of Google Map (zoom and scale level)

Zoom level	Scale level
20	1:564
19	1:1,128
18	1:2,257
17	1:4,514
16	1:9,028
15	1:18,056
14	1:36,112
13	1:72,224
12	1:144,448
...	...

Fig. 2. Road centerline of digital topographic map of 1:5,000 (left) and 1:25,000 (right) scale for target area of training data (all over Giheung-gu, Yongin-si)

3. 도로 네트워크 객체의 선택 기준에 대한 통계적 분석

도로 네트워크 객체의 선택 모델을 도출하기 위한 '선택 기준 통계적 분석' 과정을 훈련용 데이터에 적용하여 결과를 분석하였다. 대상지역은 도심지역, 비도심지역, 고속도로와 지방도, 소로 등이 적절히 혼합되어 분포하고 있는 용인시 기흥구 지역으로 선정하였으며 대축척 지도자료는 수치지형도 1:5,000(도엽번호: 37709085, 37709095), 소축척 지도자료는 수치지형도 1:25,000(도엽번호: 377093) 데이터로 구성하였다 (Fig. 2).

3.1. 전처리

수치지형도 도로중심선 데이터는 도로의 모든 교차점을 기준으로 하여 개별적인 객체로 나누어져 있으며, 교차점 사이에서도 여러 객체로 쪼개어진 경우가 많다. 이렇게 되면 도로 객체의 네트워크적인 다양한 특성값을 이용하여 일반화하는데 있어서 매우 분절된 형태의 결과가 나타날 수 있기 때문에 단일한 속성을 가진 도로객체들을 하나로 이어주는 전처리 과정이 필요하다. 뿐만아니라 도로폭, 차선수 등의 속성값이 누락되어 있거나 잘못 기입된 부분이 다소 포함되어 있었다. 이에 본 연구에서는 도로명주소 전자지도의 도로구간 데이터를 보조자료로 하여 도로중심선을 'stroke' 단위로 재구조화하고, 속성값 문제를 보완하기 위한 전처리 과정을 적용하였다.

'stroke'는 교차점에서 동일한 방향성과 유사한 속성을 가지

는 두 선형을 하나의 stroke로 묶는 과정을 반복함으로써 네트워크 상의 모든 에지들을 재분류하는 방법론이다(Thomson and Richardson, 1999). 도로 네트워크 선형을 stroke 개념을 사용하여 재구조화하게 되면, 네트워크 선형의 위상적, 기하학적 관계를 파악하는데 있어서 보다 유용하다(Fig. 3).

Fig. 3. Road network and the concept of stroke (Thomson and Richardson, 1999)

본 연구에서는 우선, 누락되거나 잘못 기입된 속성값(차로수, 도로폭, 도로구분 등)들에 대해서 도로명주소 전자지도의 도로구간 데이터로부터 동일한 도로객체의 속성값을 추출하여 도로중심선 데이터의 속성값을 수정하였다. 그런 다음, 인접한 도로객체들 간의 연결성과 속성적 유사성을 이용하여 stroke 재구조화를 실시하였다. 즉, 인접한 두 도로객체 간의 편각이 30° 이내이고 차선수의 차이가 일정개수(본 연구에서는 실험적으로 3개를 적용) 이하인 경우에 동일한 stroke 객체로 묶어주는 방식으로 모든 도로객체를 stroke로 재구조화하였다.

도로구분은 명목적 척도로 저장된 속성데이터이므로 통계적 분석을 위해서 수치형 척도로 변환할 필요가 있었다. 수

치지형도 1:5,000의 지형지물 속성목록에 따르면 도로구분을 고속국도, 일반국도, 지방도, 면리간도로, 소로, 미분류(특별시/광역시도, 시도, 군도)로 구분하고 있다(NGII, 2010). 본 연구에서는 정량적 분석을 위해 가장 하위단계인 '소로'에 1, '면리간도로'에 2, '미분류'에 3, '지방도'에 4, '일반국도'에 5, 그리고 가장 상위단계인 '고속국도'에 6의 값을 각각 부여하였다.

3.2. 객체 매칭 및 특성값 측정

이 단계에서는 축척별 지도자료에 대하여 각 지도객체 간의 매칭쌍을 탐색하여 두 축척의 지도자료에서 동시에 존재하는 객체와 대축척 지도에서만 존재하는 객체로 구분한다. 즉, 대축척 지도자료를 일반화한 결과물이 소축척 지도자료와 거의 같다고 가정했을 때, 일반화 과정에서 삭제된 객체와 삭제되지 않은 객체를 구분하는 것이다. 두 축척수준에서의 도로객체에 대한 매칭쌍 추출 프로세스는 다음 그림과 같다(Fig. 4).

Fig. 4. Extraction process for matching pair of road object between different scales

우선, 1:5,000 도로경계선 폴리곤과 1:25,000 도로중심선 데이터를 중첩하여 1:25,000 도로중심선을 포함하고 있는 1:5,000 도로경계선 폴리곤 객체를 추출한다. 추출된 도로경계선 객체들에 대응되는 1:5,000 도로중심선 객체들을 '선택된 객체' 그룹, 그 이외의 도로중심선 객체를 '삭제된 객체' 그룹으로 분류하였다. 분류 결과, 전체 도로중심선 객체(총연장 143920.14m)에 대하여 59%의 '선택된 객체'(총연장 85075.15m) 그룹과 41%의 '삭제된 객체'(총연장 58844.99m) 그룹으로 분류되었다. Fig. 5는 축척별 도로중심선 데이터 간의 도로객체 매칭 결과를 나타낸다. 그림에서 진한 회색 점선은 1:25,000 수치지형도 상의 객체, 검은색 실선은 1:5,000 수치지형도에서 '선택된 객체', 옅은 회색 실선은 1:5,000 수치지형도 상의 '삭제된 객체'를 의미한다.

Fig. 5. Extraction result of matching pair of feature for road centerline data between different scales

두 그룹에 속한 도로객체들의 특성값들을 계산하는데 있어서 속성적 특성값(도로구분 등 명목적 특성), 형상적 특성값(도로폭, 차선수, 길이, stroke degree⁵⁾, 굴곡도⁶⁾ 등), 관계적 특성값(주요 공공기관과의 이격거리, 일정 면적 이상의 공원과 이격거리, 지하철, 철도역, 터미널 등 주요 교통시설과의 이격거리 등) 등을 적용할 수 있다. 본 연구에서는 도로길이, 도로폭, 도로구분을 통계분석을 위한 도로객체의 특성값으로 선정하였다. 그 이외의 변수들에 대해서는 통계적으로 유의하지 못하거나 다른 변수와의 다중공선성(multicollinearity)⁷⁾이 발생하여 모델에서 제외시켰다.

3.3. 특성값 통계적 분석

분류된 그룹 내에서 전체 객체 개수와 선택된 객체 개수를 Töpfer의 radical 법칙에 적용하여 도로객체 클래스에 대한 심볼계수를 계산하였다(박우진 등, 2012). 본 연구에서 적용한 Töpfer의 radical 법칙은 다음 식과 같다(Eq. 1).

$$n_f = n_a \times C \times \sqrt{\frac{M_a}{M_f}} \quad (1)$$

여기서 n_f 는 결과 지도의 객체 수, n_a 는 입력 지도의 객체 수, C 는 심볼계수, M_a 는 입력 지도의 축척계수, M_f 는 결과 지도의 축척계수이다. 분석결과, 도로객체 클래스에 대한 심

5) Stroke degree : 네트워크가 stroke 단위로 구조화되어 있을 때, 하나의 stroke 내에 존재하는 교차점의 개수

6) 굴곡도(sinuosity) : 곡선의 시작점과 끝점 간의 실제거리와 최단거리의 비율로 굽은 정도를 측정하는 것

7) 다중공선성(multicollinearity) : 회귀분석에서 독립변수들 간에 강한 상관관계가 나타나는 문제이며 회귀분석의 전제조건에 위배되기 때문에 좋지 않은 분석 결과를 도출함.

볼계수는 1.065로 나타났다. 여기서 객체수를 선밀도(전체 면적에서의 선형의 길이)라는 측면에서 보면, n_f 와 n_a 에 각각 결과지도와 입력지도의 도로객체의 길이로 바꾸어 계산할 수 있는데, 이 경우 심볼계수는 1.322로 나타났다. 이는 1:5,000 수치지형도로부터 1:25,000 축척수준에 맞게 도로객체 클래스에 선택 연산자를 적용하는 데 있어서, Töpfer의 radical 법칙에 의해 도출된 선택 개체수보다 약간 더 많은 개체가 선택 될 수 있도록 해야 한다는 것을 의미한다.

또 하나의 통계적 분석으로 객체의 다양한 특성값들과 위의 '선택된 객체'와 '삭제된 객체' 그룹 간의 상관분석을 실시하여 특성값들 중 선택된 그룹과의 상관성을 분석하였다. 특성값들과 그룹 간의 관계를 분석하여 그룹을 분류하는 방법론으로는 다중회귀모형, 다기준 의사결정 방법론, 기계학습 또는 데이터마이닝 분야의 분류 알고리즘을 이용한 방법 등을 적용할 수 있다. 본 연구에서는 분석 방법 중 계량경제학 분야에서 주로 쓰이는 정성적 반응 회귀모형을 적용하였다. 정성적 반응 회귀모형은 피회귀변수(종속변수)가 이분적 변수(예를 들면, 상품의 구입의사 여부, 안전에 대한 찬성 또는 반대 여부 등이며 1 또는 0의 값으로 표현됨)이고, 그와 관련된 다양한 독립변수가 있을 때, 각 독립변수의 값에 따라 피회귀변수가 1이 될 확률을 구하는 확률모형이다. 이 분석방법에는 선형확률모형(Linear Probability Model, LPM), Logit 모형, Probit 모형, Tobit 모형 등이 있다(Gujarati and Porter, 2009). 이 중, 본 연구에서는 Logit 모형을 적용하여 각 객체의 특성값들과 구분된 그룹 간의 상관성을 분석하였다.

Logit 모형은 독립변수의 선형결합을 이용하여 사건의 발생가능성(발생확률)을 예측하는데 사용되는 통계기법으로, 로지스틱 회귀분석(logistic regression)이라고도 불린다. 종속변수가 1이 될 확률을 표현하는데 있어서 LPM이 직선 형태의 확률함수를 가지는데 반해서, Logit 모형은 누적확률함수 형태에 가까운 비선형 확률함수를 가지기 때문에 종속변수의 설명력이 높다는 장점이 있다(Fig. 6).

Fig. 6. Probability function of LPM and Logit model

또한 Logit 모형은 독립변수로 연속형 변수와 범주형 변수를 동시에 사용할 수 있어 도로중심선의 특성값인 도로폭, 도로길이, 도로구분 등을 독립변수로 적용하는 데에도 적절하다고 할 수 있다.

Logit 모형에 대한 계산식은 다음과 같다(Eq. 2, Eq. 3).

$$P_i = E(Y_i = 1|X_i) = \frac{1}{1 + e^{-(\beta_1 + \beta_2 X_i)}} \quad (2)$$

$$L_i = \ln\left(\frac{P_i}{1 - P_i}\right) = \beta_1 + \beta_2 X_i \quad (3)$$

이때, Y_i 는 종속변수, X_i 는 독립변수, P_i 는 종속변수가 1이 될 확률, β 는 각 독립변수에 대한 회귀계수, L_i 는 로그 승산비(odds ratio)이다.

본 연구에서는 독립변수로 전단계에서 측정한 도로 stroke 객체의 길이(Length), 평균 폭(Width), 도로 구분(Class) 이 세 요소를 적용하였다. 훈련용 데이터에 Logit 모형을 적용하여 도출된 분석 결과는 다음과 같다(Fig. 7).

각 변수에 대한 t검정을 수행한 결과 모두 통계적으로 유의한 결과를 나타내었으며 모델의 전체적인 유의도 역시 LR(Likelihood Ratio)테스트를 통해 검증한 결과, 높은 유의성을 나타내었다.

$$L_i = -7.5165 + 0.1088 \times Width + 0.0021 \times Length + 2.2213 \times Class$$

$se = (0.5764)$	(0.0290)	(0.0007)	(0.2182)
$t = (-13.0401)$	(3.7460)	(3.1291)	(10.1811)

Fig. 7. Result of applying Logit model to training data

4. 선택 기법 적용

위 단계에서 도출된 분석결과를 이용하여 1:5,000 수치지형도의 도로중심선 데이터로부터 1:18,000, 1:72,000 축척 수준의 도로중심선 데이터를 추출하기 위한 선택 기법을 적용하였다. 분석결과를 적용할 테스트 데이터는 수원시 팔달구 지역에 해당하는 수치지형도 1:5,000 4개 도엽(도엽번호: 37709081, 37709082, 37709091, 37709092)으로 구성하였다 (Fig. 8). 훈련용 데이터와 마찬가지로 테스트 데이터에는 도심, 비도심 지역이 적절히 혼합되어 있으며 다양한 위계를 가진 도로객체가 혼재해 있는 지역이다.

Fig. 8. Road centerline of digital topographic map of 1:5,000 scale for target area of test data (all over Paldal-gu, Suwon-si)

테스트 데이터에 대한 전처리 프로세스로 stroke 재구조화를 실시하였으며 재구조화 방법론은 3.1절에서 훈련용 데이터에 적용한 그것과 동일하게 적용하였다. 여기에 3.3절에서 도출된 도로폭, 도로길이, 도로구분과 선택된 객체 그룹간의 관계식을 Eq. 2에 대입하여 테스트 데이터의 도로중심선에 대한 선택확률함수 (Selection Probability Function, 이하 SPF)를 도출하였다. 이 선택확률함수를 모든 도로중심선 객체에 적용하여 SPF 값을 측정하였다. 다음 그림은 도로중심선 클래스에 대해 측정된 SPF 값에 대한 히스토그램을 나타낸다 (Fig. 9). SPF 값의 분포는 대체적으로 중간 수준(0.38~0.45, 0.49~0.7)의 값들이 높은 빈도수를 나타내었고 그 이외의 값들에서는 몇 개의 소그룹(0.05~0.1 0.73~0.74 등)을 이루면서 분포하는 것으로 나타났다.

위에서 산정된 SPF 값을 바탕으로 하여 목표 축척수준에 맞게 낮은 확률값을 가지는 객체는 '삭제'하고 높은 확률값을 가지는 객체는 '선택'해야 한다. 이를 위해 본 연구에서는 Töpfer의 radical 법칙(Eq. 1)에 상기 통계분석 과정에서 도출된 선밀도에 대한 심볼계수를 적용하고 원자료의 축척계수와 목표 축척수준의 축척계수를 대입하여 각 목표 축척수준에 대한 목표 선형 총길이를 계산한다. 그리고 SPF 값과 선형길이에 대한 누적 함수를 그려서, 목표 총길이가 도출되기 위해 적용되어야 할 SPF 값의 임계치를 누적 함수 상에서 역산하여 계산하였다. 아래의 그림은 SPF와 선형길이에 대한 누적함수와 목표 선형 총길이를 이용하여 SPF 임계치를 산정하는 과정을 나타낸다(Fig. 10).

Fig. 9. Histogram of SPF value for road network of test data

Fig. 10. Threshold estimation using total length of target line and the cumulative function of SPF and length of line

다음 표는 목표 축척수준에 따른 목표 선형 총길기와 SPF 값의 임계치를 나타낸다(Table 3).

Table 3. Total length of target line and threshold of SPF according to the target scale level

Target scale level	Total length of target line (m)	Threshold of SPF
1:5,000	500083.64	0
1:18,000	348435.87	0.502041
1:72,000	174217.93	0.604848

아래의 그림은 위에서 도출한 SPF 임계치를 적용하여 목표 축척수준에 따라 네트워크 객체를 선택 또는 삭제한 결과를 나타낸다(Fig. 11, Fig. 12).

Fig. 11. Results of selection and elimination of network object according to the target scale level of 1:18,000

Fig. 12. Results of selection and elimination of network object according to the target scale level of 1:72,000

5. 정성적, 정량적 평가

선택된 결과를 정성적(시각적)인 관점에서 살펴보면, 목표 축척수준에 따라 상대적으로 중요도가 높은 도로객체가 선택되고 골목길과 같은 중요도가 낮은 도로객체가 삭제된 결과를 나타내고 있다. 대체적으로 네트워크의 연결성이 보존되어 있으나 경우에 따라서 네트워크 객체 간의 연결이 끊어지거나 끝이 약간 꺾어진 형태의 stroke 객체도 다소 발견되는데 이는 도로중심선 객체의 형태가 교차로점에서 끊어지지 않고 ‘┐’자로 꺾어진 형태로 되어있는 객체들로 인해 stroke 객체가 제대로 생성되지 않았기 때문으로 보인다. 또한 도로중심선 원데이터의 속성값에 오기입된 값들(도로폭이 0 또는 과대한 값으로 기입되는 경우)이 다소 포함되어 있어 stroke 객체가 제대로 연결되지 않은 경우도 있었다.

다음으로, 선택된 결과에 대한 정량적 분석을 위해 목표 축척수준이 소축척이 될수록 어떤 속성값을 가진 객체들이 중점적으로 선택되거나 삭제되었는지를 살펴보았다. 아래의 그림들은 각 목표 축척수준에 따라 선택된 네트워크 데이터에 대하여 네트워크 객체의 폭, 길이, 위계에 대한 객체수 그래프를 나타낸다(Table 4, Table 5, Table 6).

결과표에 대한 분석 결과, 축척수준이 소축척으로 갈수록 작은 폭(0m~10m), 짧은 길이(0m~1000m), 낮은 위계(미분류, 면리간도로, 소로)의 속성을 가진 네트워크 객체들이 집중적으로 삭제된 것을 확인할 수 있다. 이를 통해 Logit 모형 형태의 선택확률함수를 이용한 본 연구의 선택 모델이 상대적으로 낮은 속성값(작은 폭, 짧은 길이, 낮은 위계)을 가진 네트

위크 객체를 제거하고 높은 속성값(큰 폭, 긴 길이, 높은 위계)을 가진 네트워크 객체를 선택하는데 있어서 효과적이라고 볼 수 있다.

Table 4. Width and number of objects for each target scale

Width \ Target Scale	1:5,000	1:18,000	1:72,000
0m~10m	2,766	248	42
10m~20m	50	45	26
20m~30m	12	12	12
30m~40m	2	2	2

Table 5. Length and number of objects for each target scale

Length \ Target Scale	1:5,000	1:18,000	1:72,000
0m~500m	2,609	91	9
500m~1000m	155	150	10
1000m~1500m	32	32	29
1500m~2000m	14	14	14
2000m~2500m	5	5	5
2500m~3000m	4	4	4
3000m~3500m	6	6	6
3500m~	5	5	5

Table 6. Class and number of objects for each target scale

Class \ Target Scale	1:5,000	1:18,000	1:72,000
Narrow path	65	0	0
Road between myeon and ri	404	0	0
Unclassified road	2,349	295	70
General national road	9	9	9
National expressway	3	3	3

6. 결론

본 연구에서는 수치지형도를 원하는 목표 축척으로 일반화하기 위한 하나의 단계로, 도로 네트워크 데이터를 목표 축척 수준에 맞게 선택 또는 삭제하는 방법론을 제안하고 적용하였다. 이를 위해, 기존의 1:5,000 축척과 1:25,000 축척의 수치

지형도를 상호 비교하고 여기에 Töpfer의 radical 법칙과 Logit 모형을 적용하여 도로 네트워크 객체의 선택과 관련된 기준들을 분석하였다. 여기서 분석된 결과를 바탕으로 하여 테스트 데이터에 대해 선택 모델을 적용하였으며, 1:5,000 수치지형도 도로중심선 레이어로부터 일반화된 1:18,000, 1:72,000 축척의 네트워크 데이터셋을 도출하였다. 일반화된 결과에 대하여 정성적, 정량적 평가를 실시한 결과, 상대적으로 높은 속성값(큰 폭, 긴 길이, 높은 위계)들을 가진 네트워크 객체들이 목표 축척수준에 맞게 적절히 선택된 결과를 나타내었다. 본 연구를 통해 지도 일반화를 통한 다양한 축척수준의 도로 데이터를 생성하는 과정을 자동화함으로써, 지도생산 비용을 절감하고 균질한 품질의 지도데이터 구축을 보다 용이하게 하는 효과가 있을 것으로 판단된다. 보다 다양한 객체 클래스에 대한 선택 및 삭제 모델을 적용하는 부분, 위치기반서비스를 위한 다양한 지도 일반화 연산자에 대한 연구들에 대해서는 추후 연구과제로 남기기로 한다.

감사의 글

본 연구는 국토교통부 첨단도시개발사업 - Realtime Digital Map 생성 및 활용기술 개발과제의 연구비지원(11첨단 도시G10)에 의해 수행되었습니다.

References

- Choe, B. and Kim, Y. (2007), Framework and Workflows for Spatial Database Generalization, *Transactions in GIS*, Vol. 11, No. 1, pp. 101-114.
- Gujarati, D. N. and Porter, D. C. (2008), *Basic Econometrics 5th edition*, McGraw-Hill, New York, pp. 541-573.
- Liu, X., Zhan, F. B., and Ai, T. (2010), Road selection based on Voronoi diagrams and "strokes" in map generalization, *International Journal of Applied Earth Observation and Geoinformation*, 12S, S194-S202.
- Mackaness, W. A. and Mackechnie, G. A. (1999), Automating the Detection and Simplification of Junctions in Road Networks, *GeoInformatica*, Vol. 3, No. 2, pp. 185-200.
- MicroImages, Inc. (1999), Google Maps Structure, <http://www.microimages.com/documentation/TechGuides/76googleMapsStruc.pdf> (last data accessed:

16 May 2013)

- National Geographic Information Institute (2010), *Attribute Lists of Terrains and Features of Digital Map ver. 2.0*.
- Park, W., Ji, B., and Yu, K. (2012), Control methodology of map generalization scale using Töpfer's radical law, *Proceedings on 2012 Spring Conference of Korea Society of Surveying, Geodesy, Photogrammetry, and Cartography*, 26-27 April, Gangneung, South Korea, pp. 309-311. (in Korean with English abstract)
- Thomson, R. C. and Richardson, D. E. (1999), The 'good continuation' principle of perceptual organization applied to the generalization of road network, In: *Proceedings of the 8th International Symposium on Spatial Data Handling*, Vancouver, Canada, pp. 1215-1223.
- Töpfer, F. and Pillewizer, W. (1966), The Principles of Selection, *The Cartographic Journal*, Vol. 3, No. 1, pp. 10-16.
- Touya, G. (2010), A Road Network Selection Process Based on Data Enrichment and Structure Detection, *Transactions in GIS*, Vol. 14, No. 5, pp. 595-614.

(Received 2013. 05. 20, Revised 2013. 06. 18, Accepted 2013. 06. 26)