

공간데이터베이스 다중 추상화시스템 구축에 관한 연구

최 병 남*

Designing for a system of multi level spatial database abstraction

Byong-Nam Choe*

요 약

GIS을 활용하는 다양한 응용분야에서 공간데이터베이스의 상세한 정도(추상화 수준)는 그 목적에 따라 달라진다. 추상화 수준이 달라지면 공간자료의 기하학적인 단순화뿐 아니라, 객체(인스턴스), 공간 및 속성 자료, 위상관계, 의미관계 등이 변화되거나 유지되어야 한다. 이와 관련하여 본 연구에서는 원시 공간데이터베이스를 추상화하여 목표데이터베이스를 생성하는 방법을 제시하였다. 공간데이터베이스를 추상화하는 연산자를 재정의하고, 어떤 객체에 어떤 연산자들을 어떤 순서로 적용하여야 하는지에 대한 기준을 제시하고, 이 기준에 따라 워크플로우를 제시하였으며, 프로토타입시스템을 개발하여 타당성을 검토했다.

주요어 : 공간데이터베이스, 공간데이터베이스 추상화, 연산자, 워크플로우

ABSTRACT : Abstraction of geographic data within a spatial database context must deal with geometrical simplification, modification or maintenance of the integrity of features, spatial and aspatial data, topology within class, and relationships between classes. This research is to propose a method to abstract a spatial database into a high level database. This study refines operators to carry out the modifications required by the abstraction process in the spatial database. Then, a set of operator sequences (workflows) is suggested to specify operators required to abstract a given feature class and to arrange them in order. Finally, a prototype system is developed, based on rule management with a graphic user interface. When the abstraction process is implemented sequentially we demonstrate that a preferential ordering of operations improves efficiency and reduces loss and distortion in the information.

keywords : spatial database, spatial database abstraction, operator, workflow

* 국토연구원 GIS연구센터 연구위원

1. 서론

환경, 자원관리, 농업, 조정, 교통, 토지 이용계획 등의 분야에서 GIS을 활용하기 위해 구축하는 공간데이터베이스의 상세한 정도는 그 목적에 따라 달라진다. 따라서 상세한 수준의 원시 공간데이터베이스를 이용해 GIS 응용시스템 목적에 적합한 목표데이터베이스를 생성한다면 구축 비용과 시간을 절감할 수 있을 것이다. 이와 관련하여 공간데이터베이스를 구성하는 지리자료의 추상화(geographic data abstraction)가 자동으로 수행되어야 하는데, 이는 GIS분야에서 주요 연구 이슈 중 하나이다(Jones et al. 1996, Devogele et al. 1996). GIS분야에서 다중수준 자료 추상화(multiple levels of data abstraction)는 다축척 지도제작 개념에서부터 발전되어 왔다. 초기 다축척 지도제작은 지도를 제작하는 사람에 의해서 수작업으로 이루어졌다. 1960년대부터 단순화 알고리즘에 대한 연구가 시작되었고, 1980년대에 들어 알고리즘을 이용한 자동지도일반화시스템에 대한 연구와 개발이 시작되었다(Lee 1993). 1990년대에 들어서는 다축척 GIS 구축 혹은 서로 다른 축척의 공간데이터베이스 통합이라는 관점의 연구가 시작되었다.

지리자료 추상화는 공간자료의 기하학적인 일반화(generalization)뿐만 아니라 비공간자료, 위상관계, 다른 객체들 간의 관계 등의 변화를 포함하는 프로세스이다. 이러한 변화는 여러 단계로 구성되며, 각 단계는 연산자로 정의된다. 하나의 객체

에 대한 추상화는 하나 이상의 연산자가 적용되기 때문에 연산자 적용순서가 중요하다. 적용순서가 적절하지 못할 경우 정보의 완결성에 영향을 미치고, 시스템의 처리효율성이 낮아진다.

본 연구의 목적은 상세한 수준의 공간데이터베이스를 추상화하여 덜 상세한 공간데이터베이스를 생성하는 워크플로우 모델을 제시하는데 있다. 이를 위해 추상화 과정에서 발생하는 변화를 반영한 연산자를 재정의하였고, 객체 클래스에 적용할 연산자를 선택하는 기준과 선택된 연산자를 적용하는 순서 기준에 따라 기본 워크플로우를 제시하였다. 그리고 프로토타입시스템을 구축하여 실제 적용해봄으로써 제시한 방법의 타당성을 검토했다.

본 연구에서는 상세한 수준의 원시공간데이터베이스를 다양한 GIS 응용시스템 목적에 적합한 목표공간데이터베이스로 생성하는 과정을 공간데이터베이스 다중 추상화(multi level spatial database abstraction)라 한다. 지도 일반화는 공간자료의 기하학적인 단순화에 초점이 있으나, 공간데이터베이스 추상화는 데이터베이스의 구조와 내용의 변화를 포함한다. 공간데이터베이스의 추상화 수준(level of abstraction) 변화를 대상으로 한다는 관점에서 다축척 지도를 이용해서 소축척 지도를 만드는 지도 일반화(map generalization)와 구별되는 용어를 사용한다. 객체와 객체들 간의 공간관계 정의는 OGC(Open GIS Consortium)(1998)와 ISO(International Organization for Standardization)/TC211(1996)에서 표준으로 정한 추상 기하모델(abstract geometry model)에 따른다.

2. 관련 연구

대축척 지도에서 중요하지 않은 객체를 제거하고, 복잡한 선 및 경계선을 단순화하고, 여러 개 면 객체를 결합하여 소축척 지도를 만드는 과정을 지도 일반화라고 한다. 숙련된 지도제작자의 수작업으로부터 시작된 일반화는 관련기술의 발전에 따라 개념과 방법이 자동 일반화로 진보되었다. 이는 객체의 기하학적인 표현을 단순화하는 지도제작 일반화(cartographic generalization)와 추상화 수준이 서로 다른 지리정보 모델을 다루는 모델 지향적 일반화(model-oriented generalization)로 구분되는데(Brassel and Weibel 1988), 이들의 가장 큰 차이는 데이터베이스 조작 여부이다(Müller 1995).

지도제작 일반화에서는 표현축척이 달라지게 되었을 때 표현되는 자료의 종류 및 질, 표시속도 등이 적절하게 이루어질 수 있도록 지리자료를 변화시키는 알고리즘과 연산자에 대한 연구가 이루어졌다. 표현되는 자료의 종류 및 질과 관련하여 지도일반화 연산자가 제시되었으며(Beard, 1991, McMaster, 1992, Lee, 1995), 표시속도와 관련해서는 데이터에서 점을 제거하는 알고리즘(McMaster, 1992)과 자료구조 및 인덱스(Li, 1995, Becker, 1991)등이 제시되었다. 지도제작 일반화를 위한 연산자 연구는 1960년대 후반부터 이루어져 왔으며 McMaster(1992)에 의해서 체계화되었다.

데이터베이스 조작을 대상으로 하는 모델 지향 일반화는 주로 다중축척 GIS 구축이나 축척이 다른 데이터베이스의 통합

과 관련된 연구이다. 모델 지향 일반화에서는 일반화 수준이 달라질 때 필요한 데이터의 선택, 데이터의 변화 혹은 생성, 공간자료의 기하/위상관계 변화 등과 관련된 이슈를 다루고 있다. 이와 관련하여 규칙(Sester, 1998, Richardson, 1994, Jones, 1996), 연산자(Beard, 1991, Jones, 1996, Peng, 1996, Dettori, 1996), 다중축척을 지원하는 모델(Devoegele, 1996, Rigaux, 1994, Rigaux, 1995, Barnett, 1996) 등이 제시되었다.

일반화는 시작부터 결과물이 생성될 때까지의 과정 또한 중요하다. 일반화 과정에 대한 연구는 최상의 결과를 얻기 위한 연산자의 적용순서, 사용자 개입의 최소화와 사용자가 개입하는 경우에 적절한 유도방법 등을 다루고 있다(McMaster, 1992). 이와 관련하여 데이터의 계층구조를 이용한 일반화 과정모델(Robinson, 1995, Jones, 1996, Ruas, 1996), 연산자의 적용순서(Peng, 1996, Lee, 1995, Regnaud, 1999), 그래픽 사용자 인터페이스 제공 방법(Weibel, 1991) 등이 제시되었다.

위에서 살펴 본 것처럼 지도제작 일반화나 모델 지향적 일반화와 관련하여 연산자가 제시되고 일반화 과정에 대한 연구가 있었다. 지금까지 제시된 연산자와 일반화 과정과 관련된 연구는 다음과 같은 특징과 한계를 가지고 있다. 첫째, 연산자 정의는 기하학적인 측면과 위상관계 측면을 포함하고 있으나, 공간데이터베이스 추상화에서 발생하는 스키마 변화는 고려되지 않았다. 둘째, 공간데이터베이스 추상화와 관련하여 연산자의 선택과 적용순서에 대한 기준이 명확하지 않고, 이에 따른 기본 워크플로우가 제시되지 않았

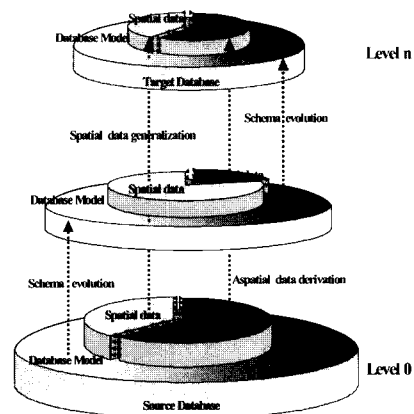
다. 셋째, 단일객체를 대상으로 한 일반화이며 여러 객체를 동시에 고려하지 않고 있다.

3. 다중 공간데이터베이스 추상화 전략

공간데이터베이스의 추상화 수준(상세한 정도)이 변화하면 객체, 공간 및 비공간 자료, 클래스 사이의 관계 등이 변화한다. 예를 들어, 클래스 Building을 덜 상세한 수준으로 추상화할 경우 면적이 작은 독립 건물은 제거되고, 인접해 있는 건물은 결합되어 건물집단 지역(Buildingblock)이 된다. 이것은 기존의 클래스 Building으로부터 새로운 클래스 Buildingblock이 생성되고, 새 클래스의 공간자료(모여 있는 건물의 경계를 연결한 건물집단지역의 경계)와 비공간자료(건물집단지역에 포함된 건물의 동수, 면적, 유형 등)가 파생되었음을 의미한다. 또한 클래스 Buildingblock과 다른 클래스 사이의 공간 및 의미관계(spatial and semantic relationships)가 형성된다.

이러한 공간데이터베이스 추상화 과정에서 원시데이터베이스와 목표데이터베이스의 클래스 사이의 대응관계(correspondence relationship)뿐만 아니라 객체클래스 내 객체(인스턴스)들 사이의 대응관계도 형성된다. 이 대응관계는 원시데이터베이스가 목표데이터베이스로 추상화되는 골격과 내용을 형성한다. 클래스 대응관계는 데이터베이스 스키마 변화를 결정하고, 객체 대응관계는 데이터베이스 내용 변화를

결정한다. 이러한 현상을 그림 1과 같이 데이터베이스 모델(스키마 변화), 공간자료 및 비공간자료(내용 변화)로 구분하여 설명할 수 있다. 즉 공간데이터베이스의 추상화 수준이 변할 경우 모델 수준에서는 클래스, 위상관계, 공간관계, 의미관계 등이 재정의(schema change)되고, 데이터 수준에서는 공간자료의 기하학적인 단순화(spatial data generalization)가 일어나고, 비공간자료는 그대로 유지되거나 새로운 클래스 정의에 적합하게 생성된다(aspatial data derivation). 특히 공간데이터베이스 추상화 과정에서 객체 유형(feature type)의 변경이나 새로운 객체 클래스가 생성될 경우 스키마 변화가 일어나고, 공간자료 및 비공간자료는 각각 독립적으로 혹은 상호 영향을 주면서 생성된다. 이와 같은 과정에서 위상관계, 공간관계, 의미관계 등의 일관성을 유지하는 것이 중요하다 (Dettori and Puppo 1996).



[그림 1] 공간데이터베이스 추상화 개념적 틀

공간데이터베이스의 추상화 수준을 변경시키기 위해서는 공간자료 및 비공간자료, 스키마 등의 변화 프로세스를 포함하는 연산자가 먼저 정의되어야 한다. 다음으로 객체 클래스의 추상화에 적절한 연산자가 선택되어야 하고, 선택된 연산자는 좋은 결과를 산출할 수 있도록 적절하게 순서화 되어 워크플로우를 이루어야 한다.

4. 공간데이터베이스 추상화 연산자와 워크플로우

4.1 연산자

본 연구에서는 ESRI(1996)가 제시한 10개 연산자중 위에서 말한 공간데이터베이스 스키마 및 내용 변화에 대한 개념을 반영하기가 용이한 Preselection, Elimination, Simplification, Aggregation, Collapse, Classification 을 재정의하여 사용하였다.

1) Preselection

목표데이터베이스에 포함시키고자 하는 객체클래스들을 원시데이터베이스에서 선택하는 연산자이다. Preselection 적용으로 객체클래스의 객체(인스턴스)나 속성을 선택할 수 있고, 공간 및 비공간 자료는 변화하지 않으며, 클래스 내 객체들의 위상관계와 클래스간의 공간관계 및 의미관계는 유지되어야 한다. 이 연산자 적용으로 선택된 객체 클래스와 공간관계 및 의미관계가 있는 다른 객체 클래스도 목표

데이터베이스에 포함된다. 선정된 객체클래스는 다른 연산자의 적용대상이 된다. 선택의 기준은 응용시스템의 목적이나 공간자료의 출력 축척이 될 수도 있다. 결과를 시각적으로 인식하기 위해서 출력 축척이 객체클래스들을 선정하는 기준이 될 수 있으나 전적으로 종속되지는 않는다.

2) Elimination

Elimination은 응용시스템 목적에 적합하지 않은 객체(인스턴스)를 목표데이터베이스의 객체 클래스에서 제외시키는 연산자이다. 비공간자료나 혹은 다른 객체 클래스와의 공간관계를 조건으로 부여하고, 그 조건을 만족하는 객체만을 제거한다. 적용 후 객체 유형은 변화하지 않고 클래스 내 객체들의 위상관계와 클래스간의 공간관계 및 의미관계는 유지된다.

3) Aggregation

공간적으로 인접해 있거나 이웃해 있는 객체들을 결합해서 새로운 공간 객체를 생성하는 연산자이다. 따라서 새로운 클래스의 공간 및 비공간자료도 생성되는데, 비공간자료는 공간자료의 변화에 따라 파생된다. Aggregation은 면 혹은 점 객체인 경우에 적용되며 결과는 반드시 면 객체이다. 따라서 적용 후 클래스 내 객체들의 위상관계는 새롭게 생성되며, 클래스간의 공간관계 및 의미관계는 제거되고 독립적인 클래스를 형성한다.

4) Collapse

면 객체를 선 혹은 점 객체로 변화시키는 연산자이다. Collapse 연산자가 적용될 경우 비공간자료는 변하지 않지만, 객체 유형이 변화되어 공간자료의 변화가 수반된다. Collapse가 적용된 후 클래스 내에서 위상관계는 새롭게 생성되며, 클래스 사이의 공간관계 및 의미관계는 그대로 유지된다.

5) Classification

Classification은 선(edge), 면(face) 유형의 객체에 적용된다. 객체 유형이 face인 경우 원시 속성값을 목표데이터베이스에 적합하도록 재분류하여 동일한 값을 갖는 인접한 객체들의 공통경계를 제거하여 새로운 객체를 생성한다. 객체 유형이 edge인 경우 속성값을 재분류하여 서로 연결된 객체들이 동일한 속성값을 갖는 경우에 새로운 객체로 생성된다. Classification을 적용하기 위해서 항목의 척도에 따라서 속성을 재분류하는 것이 필요한 경우도 있다. 등간, 등비 척도의 비공간자료는 적절한 연산자를 적용하여 계산된 값을 적용한다. Classification이 적용된 후 클래스 내에서 위상관계는 재생성되며, 클래스 사이의 공간관계 및 의미관계는 제거되고 독립적인 클래스를 형성한다. 그러나 node and edge feature 사이에서 여러 개의 edge가 연결되는 junction node와 단순히 edge와 edge가 연결된 connected node중에서 junction node와 edge사이에 공간관계는 종속관계이기 때문에 유지되어야 한다.

6) Simplification

객체의 형상을 최대한 유지하면서 불필요하게 상세한 기하학적인 형태를 단순화하는 연산자로 선 혹은 면 유형의 객체에 적용된다. 적용 후 객체 유형은 변화하지 않으며, 비공간자료는 변하지 않고 공간자료만 달라진다. 이 과정에서 클래스 내 객체들의 위상관계와 클래스간의 공간관계 및 의미관계는 유지된다.

4.2 연산자 선택과 적용순서 기준

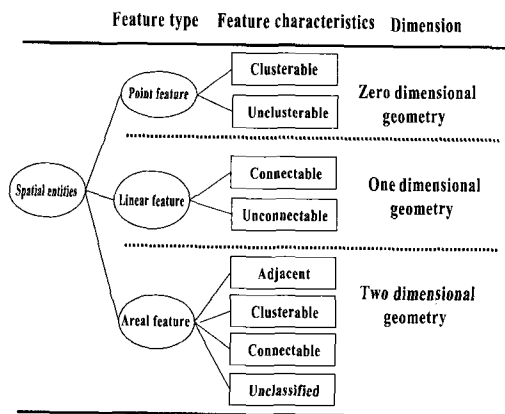
1) 연산자 선택기준

객체 클래스의 추상화에는 하나 이상의 연산자가 적용되기 때문에 객체 클래스 추상화에 적용 가능한 연산자를 선택해야 한다. Lee(1995)는 자료일반화(data generalization)를 위해 고려해야할 사항을 제시하고 Vegetation, Hydrography, Individual buildings, Developed areas 등의 워크플로우를 제시했다. 그러나 연산자를 선정하는 기준과 적용하는 순서를 정하는 기준이 명확하지 않고, 고려사항을 적용하는 우선순위가 없다.

본 연구에서는 객체 클래스의 유형과 특성을 객체 클래스 추상화에 적용할 연산자 선택기준으로 한다. 첫째, 모든 객체 클래스는 객체 유형(예 : point, lineal, areal)을 갖고 있는데, 연산자 정의는 객체 유형과 밀접한 관계가 있다. 이 관계에 따라 객체 클래스에 적용할 수 있는 연산자와 적용할 수 없는 연산자가 있다. 예를 들어 점 객체에는 Simplification을 적용할 수 없고, 선 객체에는 Collapse를 적용

할 수 없다. 둘째, 객체 클래스의 추상화 수준이 높아지면서 공간범위가 변한다. 이 공간범위 변화는 서로 군집화 (clustering)하여 경계를 이루거나(여러 개 점 객체들이 하나의 면 객체로 군집화), 서로 연결되어 선형을 이루거나(connectivity), 인접한 경계 확장(adjacency)으로 나타나기도 한다. 본 연구에서는 이를 객체 특성이라고 하는데, 객체특성에 따라 객체 클래스의 추상화에 적용할 수 없는 연산자와 적용할 수 있는 연산자가 있다. 예를 들어 도로는 연결되어 있어 Classification을 적용 가능하지만 connectivity가 없는 등고선에는 Classification을 적용할 수 없다. 또한 객체 유형이 점인 전봇대에 Aggregation을 적용한 결과에 어떤 의미를 부여하기 어렵다. 이와 같은 경우에는 아주 특수한 경우가 아니면 일반적으로 Aggregation을 적용하지 않는다.

결국 연산자 정의, 객체 유형, 객체특성 사이에는 상호의존성이 존재하고, 이 관계에 따라 객체 클래스에 적용할 연산자를 선택할 수 있다. 그림 2는 객체 유형과 특성의 관계를 나타낸 것이다.



[그림 2] 객체 유형과 특성 분류

2) 연산자 적용순서

객체 클래스의 추상화에 연산자가 적용되는 순서에 따라 처리 결과와 효율성이 달라진다. 따라서 연산자를 적용하는 순서가 중요한데, 적용순서를 정하는 기준으로 다음과 같이 두 가지를 고려할 수 있다. 첫째, 추상화된 데이터베이스의 일관성, 완결성, 정확성 등이 확보되어야 한다(ESRI 2000). 이와 관련하여 일관성은 연산자의 제약조건으로 반영되고, 정확성은 적용하는 알고리즘과 조건에 의존적이기 때문에 연산자의 적용순서와 관계가 없다. 완결성은 데이터베이스에 누락 (omission) 오류가 없음을 나타내는 것이다. 따라서 완결성 수준을 높이기 위해 정보의 손실 혹은 왜곡 가능성을 최소화하는 순서로 연산자를 적용한다. 둘째, 연산자 적용순서에 따라 각 연산자에 적용되는 대상과 데이터 양이 다르기 때문에 추상화에 소요되는 시간과 자원이 달라진다. 따라서 추상화의 효율성을 최대화 하는 순서로 연산자를 적용한다.

연산자 적용순서를 정하는 기준으로 처리 효율성의 최대화와 정보손실 및 왜곡의 최소화는 상반관계에 있다. 모든 공간 자료는 어느 정도는 부정확한 부분을 포함하고 있지만, 일관성, 완결성, 정확성 등의 확보는 GIS에서 가장 중요한 이슈이다(NCGIA 1991). 따라서 처리 효율성 최대화보다 완결성을 우선 고려하여 다음과 같이 연산자 적용순서를 정한다.

객체 클래스에 Elimination과 Aggregation

을 적용할 경우 Elimination을 먼저 적용하는 것이 처리 효율성은 높으나, 자료 손실 혹은 왜곡의 가능성이 높아지기 때문에 Aggregation을 먼저 적용한다. face 객체의 경우도 같은 이유로 Classification을 Elimination보다 먼저 적용한다. 그 다음에 Collapse, Simplification, Classification (edge-type), Elimination은 어떤 순서로 적용해도 완결성에 영향을 미치지 않기 때문에 처리 효율성을 최대화 하는 순서로 적용한다. Elimination을 다른 연산자들 보다 먼저 적용하는 경우 처리대상이 줄어들기 때문에 처리시간이 빨라진다. 그리고 Simplification보다 Collapse를 먼저 적용하는 것이 처리해야할 데이터량을 줄일 수 있기 때문에 처리시간을 줄일 수 있다.

위의 내용을 요약하면 전체적인 연산자 적용 우선순위는 Aggregation 혹은 Classification(face) - Elimination - Collapse 혹은 Classification(edge) - Simplification이 바람직하다.

4.3 기본 워크프로우

연산자를 선정하는 기준(객체의 유형과 특성과 선정된 연산자를 적용하는 순서기준(자료 손실 및 왜곡의 최소화, 처리 효율성의 최대화)을 적용하여, 본 연구에서는 표 1과 같이 기본 워크플로우를 제시하였다.

점 객체에는 선 혹은 면 객체에만 적용 가능한 Collapse, Classification, Simplification은 적용할 수 없고, Aggregation과 Elimination을 적용할 수 있다. 그러나 군집성(cluster)을 갖지 않는 점 객체에 Aggregation을 적용이 하는 것이 의미적인 한계가 있어서 Elimination만을 적용한다(Workflow1). 군집성을 갖는 경우에는 적용순서 기준에 따라 먼저 Aggregation을 적용하고, 그 결과에 대해서 Elimination를 적용한다(Workflow2).

선 객체의 경우 면 혹은 점 객체에만 적용되는 Aggregation과 Collapse는 적용할 수 없고 Simplification, Elimination, Classification을 적용한다. 선 객체가 연결성(connectivity)을 갖는 경우는 Elimination - Classification - Simplification 순으로 적용되고(Workflow3). 그렇지 않은 경우에는 Elimination - Simplification 순서로 적용된다(Workflow4).

객체 클래스가 face 유형인 경우는 Classification - Elimination - Simplification 순서로 적용되고(Workflow5), 연결성을 갖는 면 객체(예: roads, rivers)는 Elimination - Collapse - Simplification 순서로 적용된다(Workflow6). 군집성을 갖는 면 객체는 Aggregation - Elimination - Collapse - Simplification 순서로 적용되고(Workflow7), 이외의 면 객체는 Elimination - Collapse - Simplification 순서로 적용된다(Workflow8).

<표 1> 기본 워크플로우

Feature			Base workflow	
Class name	Type	Characteristic	Workflows	operators
Roadfacility, Height, Station	Point	Unclusterable	Workflow1	Elimination
Tree	Point	Clusterable	Workflow2	Aggregation Elimination
Roadcenter, Linearstream, Subway	Edge	Connectable	Workflow3	Elimination Classification Simplification
Contour	Line	Unconnectable	Workflow4	Elimination Simplification
Parcel, Township	Face	Adjacent	Workflow5	Classification Elimination Simplification
Widthstream	Polygon	Connectable	Workflow6	Elimination Collapse Simplification
Building	Polygon	Clusterable	Workflow7	Aggregation Elimination Collapse Simplification
Bridge, Reservoir, Roadarea	Polygon	Unclassified	Workflow8	Elimination Collapse Simplification

5. 공간데이터베이스 다중 추상화 사례 및 평가

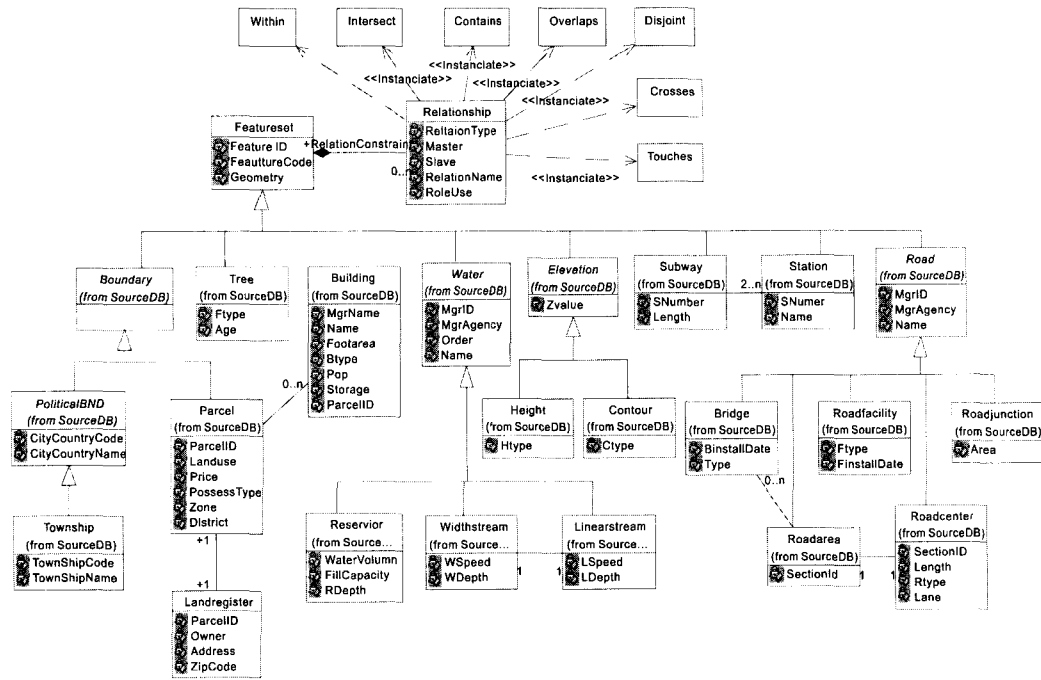
5.1 공간데이터베이스 추상화 구현사례

본 연구에서는 국립지리원에서 제작한 축척 1/1000 수치지형도¹⁾를 이용해 원시 데이터베이스(그림 3, 4 참조)를 구축하고, 기본 워크플로우(표 1)를 적용해 목표 데이터베이스(그림 5, 6 참조)를 얻었다. 시스템 개발은 ArcInfo8 GIS(ESRI 2000)에서 제공하는 기본 공간자료 연산 기능을 사용했다.

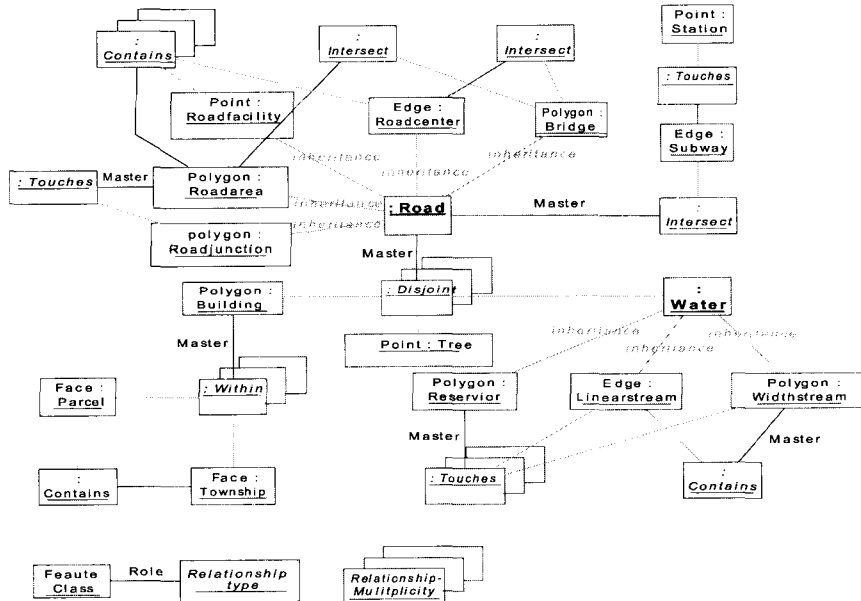
실험은 공간데이터베이스 추상화 과정이 비교적 복잡한 Workflow2, 3, 5, 7를 대상으로 하였다. 본 실험에서 각 연산자의 파라미터 중 일부는 USGS(1994)(축척 1/24000)가 제시한 것을 적용하였고, USGS가 제시하지 않은 나머지는 본 연구에서 반복적으로 적용해 얻은 경험 값을 사용하였다. 먼저 Preselection 연산자를 이용해 클래스 Tree, Roadcenter, Building, Parcel을 선택하였으며, 규칙에 따라 이 클래스들과 관계를 가지고 있는 클래스도 선택되었다.

1) 건물, 도로, 등고, 가로수, 맨홀 등을 포함하고 있으며 1996년에 제작됨.

최병남

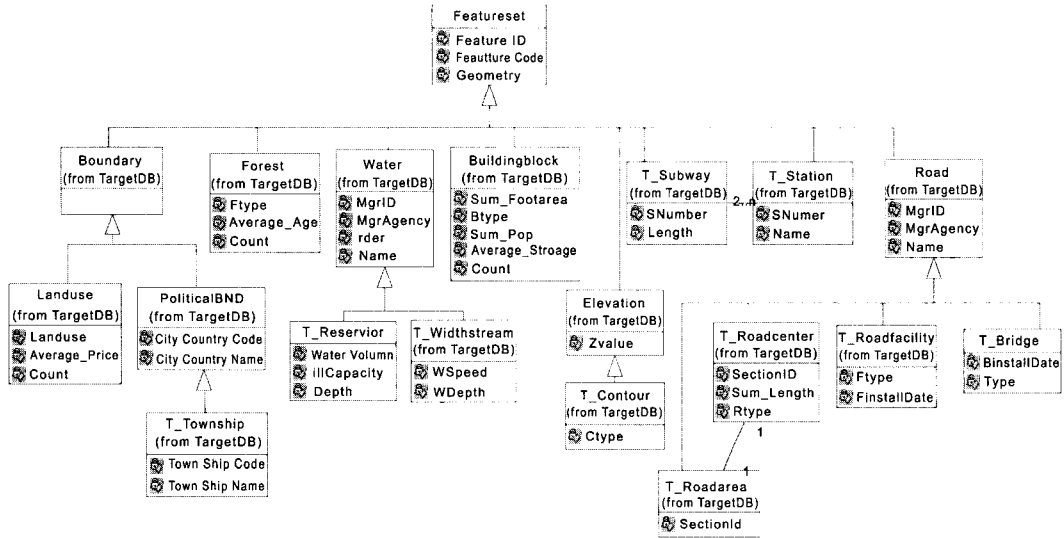


[그림 3] 원시데이터베이스 데이터 모델

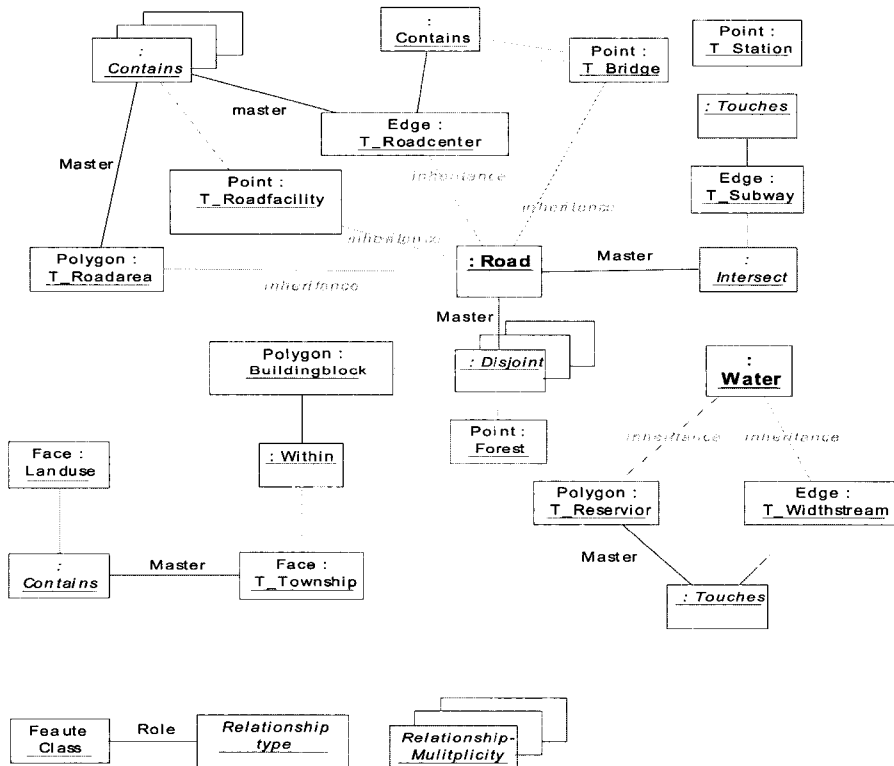


[그림 4] 원시데이터베이스 공간관계

공간데이터베이스 다중 추상화시스템 구축에 관한 연구



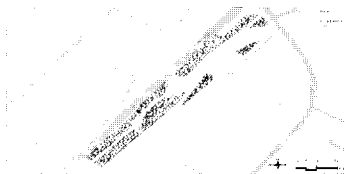
[그림 5] 목표데이터베이스 데이터 모델



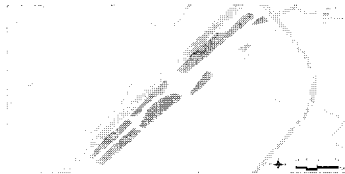
[그림 6] 목표데이터베이스 공간관계

1) Trees (Point, Cluster) → Workflow2 → Forest

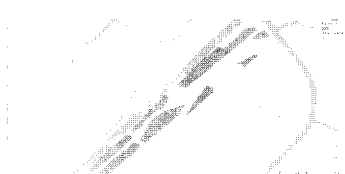
Tree는 군집성을 가지고 있기 때문에 Workflow2(Aggregation - Elimination)를 적용하였으며, 그 결과 클래스 Forest가 생성되었다. 이는 클래스 Tree에 Aggregation이 적용된 결과이다. Forest의 비공간자료인 Average_Age, Count는 클래스 Tree의 비공간자료 Age와 객체 수로부터 파생되었다. 이 과정에서 barrier constraint(Road-Disjoint-Tree)는 Forest와 Road의 중첩을 방지한다(그림 5, 6, 7). Workflow2에 대한 비교-Workflow2는 Elimination - Aggregation 순서이다.



Workflow2 적용 전



비교 Workflow2 적용 후

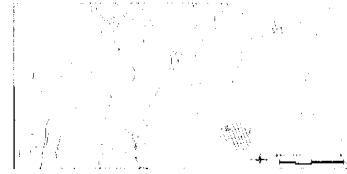


Workflow2 적용 후

[그림 7] Workflow2

2) Roadcenter (Edge, Connectivity) → Workflow3 → T_Roadcenter

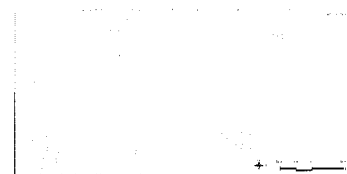
네트워크 구조를 갖는 Roadcenter는 연결성을 가지고 있기 때문에 Workflow3(Elimination - Classification - Simplification)를 적용하였다. Elimination 적용 조건에 적합하지 않는 객체는 제거되고, Classification을 적용하여 새로운 객체인 주요도로(T_Roadcenter)가 생성되었다. 이 과정에서 비공간자료 Length는 Sum_Length를 파생하고, 의미관계는 유지되고 공간관계(Roadarea-Contains-Roadcenter)는 갱신(T_Roadarea-Contains-T_Roadcenter)되었다(그림 5, 6, 8). Workflow3에 대한 비교-Workflow3는 Classification - Elimination - Simplification 순서이다.



Workflow3 적용 전



비교 Workflow2 적용 후



Workflow2 적용 후

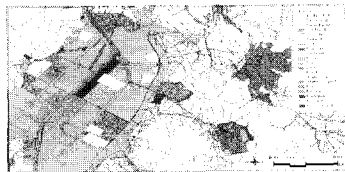
[그림 8] Workflow3

3) Parcel (Face, Adjacency) → Workflow5 → Landuse

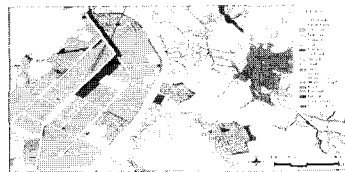
Parcel은 인접성을 갖는 face 유형 객체이기 때문에 Workflow5(Classification - Elimination - Simplification)를 적용하였다. Classification적용으로 새로운 클래스 Landuse가 생성되었고, 비공간자료 Average_Price와 Count는 각각 Price와 필지수로부터 파생되었다. Landregister와 의미관계는 제거되었고, 공간관계(Township-Contains-Parcel)는 T_Township-Contains-Landuse로 변경되었다(그림 5, 6, 9). Workflow5에 대한 비교-Workflow5는 Elimination - Classification - Simplification 순서이다.

4) Building (Polygon, Cluster) → Workflow7 → Buildingblock

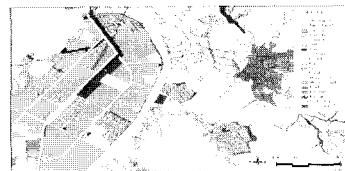
Building은 군집성을 갖고 있는 면 객체이기 때문에 Workflow7(Aggregation - Elimination - Collapse - Simplification)이 적용되었다. Aggregation 조건이 만족되는 건물물이 합해져서 새로운 클래스 Buildingblock이 생성되었다. 비공간자료 Sum_Footarea, Sum_Pop, Average_Storage가 생성되었고, 공간관계 (Road-Disjoint-Building)는 (Road-Disjoint-Buildingblock)으로 갱신되었다(그림 5, 6, 10). Workflow7에 대한 비교-Workflow7는 Elimination - Aggregation - Collapse - Simplification 순서이다.



Workflow5 적용 전

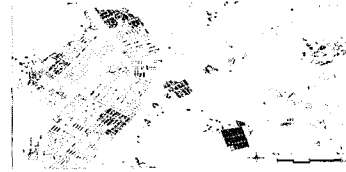


Workflow5 적용 후

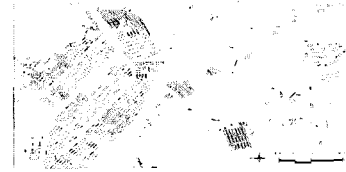


비교 Workflow5 적용 후

[그림 9] Workflow5



Workflow7 적용 전



Workflow7 적용 후



비교 Workflow7 적용 후

[그림 10] Workflow7

5.2 결과 평가

공간데이터베이스 추상화 결과는 원시 자료를 변형(transform) 및 전환(convert)시켜 생성한 것이다. 따라서 원시 자료의 가공이 적절히 수행되었는지, 생성된 결과를 공간분석과 같은 목적에 사용할 수 있는 질적 수준을 가지고 있는지를 적절한 방법으로 평가해야 한다. 본 연구에서는 추상화된 공간데이터베이스의 일관성, 완결성, 정확성(NGIA 1991), 효율성을 평가하였다.

일관성은 추상화된 공간데이터베이스의 논리적 모순이 없음을 의미하는 것으로 의미관계와 공간관계로 구분하여 평가했다²⁾. 의미관계 일관성 평가를 위해 원시 스키마와 생성된 스키마를 비교하여 ‘유지 혹은 제거되어야 할 의미관계들이 목표데이터베이스에서 유지 혹은 제거되었는가’를 확인하였으며, 그 결과 일관성이 유지되었다. 또한 공간관계 일관성 평가를 위해 공간자료 타입에 의한 논리적 부정관계에 따라 관련 클래스를 대상으로 불일치 발생여부를 확인하였으며, 그 결과 불일치 경우는 없었다.

정확성은 오류(error)의 반대 개념으로 위치 정확성과 주제 정확성³⁾으로 구분할 수 있다. 추상화된 공간데이터베이스의 위치 정확성은 사용한 알고리즘과 조건에 종속적이다. 본 연구에서 위치 정확성을 축척 1/25000 지형도와 1/1000 원시

데이터베이스를 축척 1/25000으로 추상화한 결과와 중첩하여 시각적으로 비교하였다. 기존 종이지도는 표현의 명료성과 가식성을 높이기 위하여 객체가 이동 혹은 과장되어 있다는 것을 알 수 있었으며, 추상화된 공간자료의 위치 정확성이 더 높음을 시각적으로 확인하였다.

완결성은 데이터베이스에서 누락(omission) 오류가 없음을 나타내는 것으로 추상화 결과에 대한 자료 손실 및 왜곡의 정도를 공간과 주제 차원에서 측정하여 완결성을 평가하였다. 공간 완결성(spatial completeness)은 동일한 조건에서 자료 손실이나 왜곡을 최소화해야 높아진다는 측면에서 연산자 순서를 달리 하였을 때 나타나는 공간적 분포영역, 객체(인스턴스) 수, 면적, 길이를 비교하여 평가하였다. Workflow2, 7은 비교-Workflow2, 7보다 공간자료의 손실이 적게 일어났고(객체수의 변화), 공간적인 분포 영역 또한 더 넓게 나타났다(클래스의 분포범위). Workflow3과 비교-Workflow3의 출력결과는 동일하게 보이지만 클래스의 객체 수에서 차이가 난다. 이는 객체의 타입(edge)의 특성으로 Classification 대상이 아닌 객체들이 먼저 제거되어야 원하는 형태대로 추상화가 되기 때문이며, 따라서 비교-Workflow3은 올바르게 추상화되었다고 보기 어렵다. 비교-Workflow5 수행 결과는 조각화에 의해 객체 수가 증가하였으며, 이는 자

2) 위상구조 일관성은 추상화 후 목표데이터베이스에 적절한 위상구조가 생성되었는지를 확인함으로써 평가할 수 있는데, 이는 기반시스템에서 제공하고 있기 때문에 본 연구에서는 제외하였음.

3) 주제 정확성(thematic accuracy)은 데이터베이스에 저장된 속성 값의 정확성으로 자료의 특성에 따라 다르기 때문에 평가대상에서 제외했음.

첫 주제의 성격에 맞지 않게 추상화될 우려가 있다. 주제 완결성은 '원시데이터베이스 클래스의 비공간자료 값의 범위가 목표데이터베이스에서 유지되는가'를 평가하였다. 원시 클래스 Building의 건물 유형은 범주형 속성 값 6가지가 존재했는데, 기본 워크플로우에 의한 추상화 후에는 6가지 유형이 유지되었으나, 비교-Workflow7의 경우는 4 가지 주제 특성이 유지되었다. 이는 객체의 집단화 과정에서 사칙연산이 가능한 비공간자료는 재계산되고, 범주형 자료는 재분류되어 새로운 클래스의 비공간자료가 생성되는 유도과정에서 연산자 적용순서 때문에 생긴 결과이다. 따라서 자료의 손실이나 왜곡이 최소화되는 기본 워크플로우를 적용하는 경우에 완결성이 높다고 할 수 있다.

효율성은 연산자 순서를 달리 하였을 때 소요되는 처리 시간을 측정하여 평가하였다. Workflow2, Workflow5, Workflow7는 비교-워크플로우 보다 처리시간이 더 걸렸고, Workflow3는 비교-워크플로우 보다 덜 걸렸다. 적용순서 기준에서 언급한 것처럼 효율성과 완결성은 충돌하는 경우가 발생하는데, 이 경우 효율성보다는 완결성을 우선 적용해야 완결성을 확보할 수 있다.

6. 결 론

상세한 수준의 원시 공간데이터베이스를 이용해 GIS 응용시스템 목적에 적합한 목표데이터베이스를 추상화하기 위해

서는 스키마와 내용을 변경시켜야 한다. 공간데이터베이스 추상화에는 여러 연산자가 적용되고, 적용되는 연산자 순서를 바꾸면 그 결과가 다르게 나타난다. 따라서 연산자의 적용순서는 추상화된 공간데이터베이스의 일관성, 완결성, 정확성에 영향을 미친다. 본 연구를 통해 연산자의 정의도 중요하지만 연산자의 적용순서 또한 중요하다는 것을 알 수 있었다.

본 연구를 통해 다음과 같은 결론을 얻을 수 있었다. 첫째, 상세한 수준의 공간데이터베이스로부터 다양한 GIS응용 목적 각각에 적합한 공간데이터베이스를 자동으로 혹은 사용자 요구에 따라 생성할 수 있다. 둘째, 공간데이터베이스 추상화 과정이 체계적으로 정의되었기 때문에 원천 데이터베이스가 변경되었을 때 목표데이터베이스를 자동으로 갱신할 수 있다. 따라서 데이터베이스 구축 비용절감은 물론 데이터 일치성을 확보할 수 있다. 셋째, 대축척 지도를 소축척으로 일반화하는 방법으로도 활용할 수 있다. 넷째, 이 연구 결과는 지도 일반화는 물론 다양한 GIS 응용 목적에 적합한 다중 수준 공간데이터베이스 구축을 위한 CASE Tool을 개발하는데 중요한 사안이 될 수 있을 것으로 기대한다.

본 연구의 한계로 다음과 같은 사항들이 있으며 이는 앞으로 더 연구되어야 할 것들이다. 첫째, 본 연구에서 객체 유형을 point (point, node), linear (line, edge), areal (polygon, face)로 구분하고, 공간관계 및 의미관계를 연산자에 반영하여 연산자를 정의하고 기본 워크플로우를 제시

하였다. 좀 더 일반화된 워크플로우를 정립하기 위해 객체 유형과 관계성을 더 세분화하는 연구가 이루어져야 한다. 둘째, 데이터베이스 모델 중심의 본 연구 결과에 지도학적인 표현을 위한 typification, exaggeration, displacement, aesthetic refinement (ESRI 1996)이 통합되어야 한다. 이를 위해 데이터베이스 모델에 객체의 기하학적인 표현 요소들을 관계성에 반영하는 연구가 이루어져야 할 것이다. 셋째, 공간 DB 추상화가 제대로 되었는지를 판단하기 위해 그 과정과 결과를 객관적으로 검증하는 방법이 제시되어야 한다.

참고문헌

- Barnett, L. and Carlis, J.V., 1996, "A Roads data model: a necessary component for feature-based map generalization", In Proceedings of the 4th ACM International Symposium Geographic Information Systems (ACMGIS'96), pp.58-67.
- Beard, K. and Mackaness, W., 1991, "Generalization operations and supporting structure", In Proceedings of the 10th International Symposium on Computer-Assisted Cartography Auto-Carto 10 (Bethesda: American Congress for Surveying and Mapping), Vol. 10, pp.29-45.
- Becker, B. Six, H. W. and Widmayer, P., 1991, "Spatial priority search: an access technique for scaleless maps", In Proceedings of the Association for Computing Machinery Special Interest Group on Management of Data 1991 Annual Conference (ACM SIGMOD), pp.128-137.
- Brassel, K. E. and Weibel, R., 1988, "A review and conceptual framework of automated map generalization", International Journal of Geographical information systems (IJGIS88), Vol. 2(3), pp. 229-244.
- Dettori, G., and Puppo, E., 1996, "How generalization interacts with the topological and metric structure of maps", In Proceedings of the 7th International Symposium on Spatial Data Handling (SDH96), pp.9A-27-9A38.
- Devogele, T., Trevisan, J., and Raynal, L., 1996, "Building a multi-scale database with scale-transition relationships", In Proceedings of the 7th International Symposium on Spatial Data Handling (SDH96), pp.337-351.
- ESRI, 1996, Automation of map generalization the cutting-edge technology, <http://arconlie.esri.com>.
- ESRI, 2000, Map generalization in GIS: practical solutions with workstation ArcInfo software, <http://www.esri.com/library/whitepapers/pdfs/generalization.pdf>.
- ISO/TC211/WG 1, 1996, ISO 15046-3 Conceptual Schema Language (Working Draft 1.0). SO/TC211 N222 Document, ISO/TC211 Working Group 1, pp.31.
- Jones, C. B. Kidner, D. B. Luo, L. Q. Bundy, L. and Ware, J. M., 1996, "Database design for a multi-scale spatial information system International Journal of Geographical Information Systems (IJGIS96), Vol. 10(8), pp.901-920.
- Lee, D., 1993, "From master database to multiple cartographic representation",

- In Proceedings of the 16th International Cartographic Conference, Cologne, Germany, 3-9 May, Vol. 2, pp.1075-85.
- Lee, D., 1995, "Experiment on formalizing the generalization process", In GIS And Generalization, edited by J. C. Müller, J. P. Lagrange, and R. Weibel (London: Taylor& Francis), pp.219-234.
- Li, Z., 1995, "An examination of algorithm for the detection of critical points on digital cartographic lines", The Cartographic Journal, Vol. 32, pp.121-122.
- McMaster, R. B. and Shea, K. S., 1992, Generalization in Digital Cartography (The Association of American Geographers).
- Müller, J. C., Weibel, R., Lagrange, J. P., and Salge, F., 1995, "Generalization: state of the art and issues", In GIS And Generalization, edited by J. C. Müller, J. P. Lagrange, and R. Weibel (London: Taylor& Francis), pp.3-17.
- NCGIA, 1991, Technical issues in GIS, edited by Micheal F. Goodchild, and Karen K. Kemp, (National Center for Geographic Information and Analysis University of California, Santa Barbara), pp.45-1-45-11.
- OGC (Open GIS Consortium), 1998, The Open GIS Guide, Introduction to Interoperable Geoprocessing, <http://www.opengis.org/techno/guide.html>.
- Peng, W. and Tempfli, K., 1996, An object-oriented design for automated database generalization, In Proceedings of the 7th International Symposium on Spatial Data Handling (SDH96), pp.199-213.
- Regnault, N., Edwardes, A., and Barrault, M., 1999, Strategies in building generalization: modelling the sequence, constraining the choice, In Workshop on Progress in Automated Map Generalization (ACI99), http://www.geo.unizh.ch/~barrault/WSHOP_ACI/workshop.htm.
- Richardson, D. E., 1994, Generalization of spatial and thematic data using inheritance and Classification and Aggregation hierarchies, In Proceedings of the 6th International Symposium on Spatial Data Handling (SDH94), pp.957-972.
- Rigaux, P. and Scholl, M., 1994, Multiple representation modelling and querying, In Geographic Information Systems, International Workshop on Advanced Information Systems (IGIS'94), pp.59-69.
- Rigaux, P. and Scholl, M., 1995, Multi-scale partitions: application to spatial and statistical databases, In Proceedings of the 4th International Symposium on Advances in Spatial Databases (SDD95), pp.170-183.
- Robinson, G. J., 1995, "A hierarchical top-down bottom-up approach to topographic map generalization", In GIS And Generalization, edited by J. C. Müller, J. P. Lagrange, and R. Weibel (London: Taylor& Francis), pp.235-245.
- Ruas, A. and Plazanet, C., 1996, Strategies for automated generalization, In Proceedings of the 7th International Symposium on Spatial Data Handling (SDH96), pp. 319-336.
- Sester, M., Anders, K. H. and Walter, V., 1998,

- “Linking objects of different spatial data sets by Integration and Aggregation”, *GeoInformatica*, Vol. 2(4), pp. 335-358.
- USGS(US Geological Survey), 1994, Instructions for Stereocomposition of Map Manuscripts Scribed at 1:24,000, Topographic Division, USGS March 1964, Reston, Virginia.
- Weibel, R., 1991, “Amplified intelligence and rule-based systems” , In *Map Generalization: Making Rules for Knowledge Representation*, edited by B. P. Buttenfield, and R. B. McMaster (London: Longman), pp. 172-186.