

유전자알고리즘을 이용한 지도제작 오류 처리 프로그램 개발

The Development of Mapping Error Handling Program using Genetic Algorithm

김현덕* · 박기석** · 박경렬*** · 최석근****

Kim, Hyun-Duck · Park, Ki-Surk · Park, Kyung-Yul · Choi, Seok-Keun

- * 충북대학교 대학원 토목공학과 석사과정(hyunduck0@git.co.kr)
- ** 충북대학교 대학원 토목공학과 박사과정(kspark@git.co.kr)
- *** 공간정보기술 대표이사 박경렬(kypark@git.co.kr)
- **** 충북대학교 공과대학 토목공학과 부교수(choi@chungbuk.ac.kr)

요지

소축척 수치지도 및 지형도 제작을 위한 자동화 처리 과정에서는 기하학적 및 논리적인 오류가 발생하고 있고, 기하학적 오류는 많은 부분에서 자동화 처리가 가능하나 논리적 오류는 논리적으로 자동 판단이 곤란한 경우가 많기 때문에 대부분 수작업으로 이루어 지고 있는 실정이다.

따라서, 본 연구에서는 논리적 오류문제를 해결하기 위하여 폴리곤 겹침에 대해서 도로와 건물 폴리곤간의 겹침위치를 탐색하고, 이를 처리하기 위해 유전자알고리즘을 이용하여 프로그램을 개발하였다. 그 결과 지도제작 과정에서의 논리적 오류에 대한 자동화가 가능하도록 하였고, 그로 인한 지도제작 오류를 최소화할 수 있었다.

1. 서론

21세기 정보화 시대에 국토공간 정보화 사업을 위하여 국가 GIS 구축 사업이 수행되고 있으며, 이를 위해 1/1,000, 1/5,000 및 1/25,000 등의 수치지도 및 종이지도를 제작하고 있다.

이와 같이 축척이 다른 지도를 제작하기 위해서는 국가의 막대한 예산이 소요되고, 지도제작 기간이 길어짐에 따라 최신 정보 제공과 시간·경제적인 많은 문제점을 나타내고 있다. 국내·외에서 소축척지도제작은 대축척지도를 이용한 일반화(Generalization) 과정을 통하여 자동제작하고자 하는 연구가 수행되고 있고, 컴퓨터 기술의 발달로 그 가능성이 높아졌으나, 표현 방법과 자동화 등 여러 가지 면에서 많은 문제점들이 나타나고 있다.

이러한 자동지도제작과정에서는 기존 수치지도나 지형도가 가지고 있었던 단순 도

형 자료로 인한 정보 부족과 자동화 처리 과정에서 발행하는 기하학적·논리적 오류가 발생하고 있고, 기하학적 오류는 이미 많은 부분에서 자동화 처리가 이루어졌으나, 논리적 오류는 논리적으로 자동 판단이 곤란한 경우가 많기 때문에 대부분 수작업으로 이루어 지고 있어 자동화 처리가 요구되는 실정이다.

따라서, 본 연구는 자동화처리과정에서 논리적 오류로 나타나는 폴리곤 겹침에 대해서 도로와 건물 폴리곤간의 겹침오류를 검색하고, 이를 처리할 수 있는 유전자알고리즘을 제안하여 자동처리프로그램을 개발함으로써 지도제작오류 최소화, 지도제작 자동화를 향상, 지도제작비 절감 등에 기여할 수 있는 프로그램을 개발하는데 그 목적이 있다.

2. 유전자알고리즘 이론

유전자 알고리즘은 잠재되어 있는 모집단(population)을 초기화하고, 적합한 선택과정을 통해 다음 세대에 생존할 개체들을 선별한다. 선별된 개체들에 대해 교차(crossover), 돌연변이(mutation) 등의 유전 연산을 시행하고, 그 결과로 산출된 개체들에 대해 적합도를 이용하여 해 공간을 탐색하고 최적해를 찾는다.

유전자 알고리즘은 복잡한 해 공간의 탐색 능력이 우수하여 변수와 제약이 많은 대형 공간문제를 해결하는데 적합하고, 또한 모형에 대한 유연성 및 강건성이 높아 제약 첨가나 목적함수의 변경이 용이하다는 장점을 가지고 있다. 따라서, 지도 데이터의 경우 방대한 폴리곤 객체의 위치 예측 및 최적화 구성에 유전자 알고리즘이 적합하다.

```
begin
  t ← 0
  initialize P(t)
  evaluate P(t)
  while(not termination-condition) do
    begin
      t ← t + 1
      Select P(t) from P(t-1)
      alter P(t)
      evaluate P(t)
    end
  end
end
```

그림 1. 일반적인 유전자 알고리즘 구성도

그러나, 본 연구에서 다루고자 하는 논리적 오류의 해결 같은 무수한 결과가 나올 수 있는 복잡한 해 공간을 탐색해야 하는 문제에 있어서는 기존의 유전자 알고리즘과 비교해 보다 정확한 선별, 효과적인 탐색, 다양한 해 공간에 대한 검토가 가능하도록 검색시 대략적인 위치 이동 직선을 가진다.

문제를 해결하기 위한 임의의 해를 염색체라 부르고, 이 염색체상의 각 인자를 유전자라 한다. 생물학에서는 유전자가 염기

하나와 대응되지 않지만, 유전 알고리즘에서는 인자 하나를 유전자라 한다. 생물학에서 유전자형(genotype)은 유전자의 조합이고, 표현형(phenotype)은 관찰되는 형질을 말한다.

유전자는 문제의 해로 적당한 표현으로 정의할 때, 그에 따르는 데이터 구조가 선택되어야만 한다. 예로 실수와 정수의 값을 가질 경우 그에 해당하는 해를 표현하여야 한다. 그 외의 다른 정보를 포함할 경우 오히려 탐색 공간이 넓어지고 수행 능력을 저해한다.

유전자 알고리즘에 적용되는 연산자는 초기화, 돌연변이, 그리고 교배 등이다.

초기화 연산자는 유전자 알고리즘이 어떻게 시작되는지 결정하고, 새로운 해들을 진화시킬 유전적 재료로 구성된다. 돌연변이 연산자는 우수하지 않거나 전혀 다른 형질을 이용함으로써 근접할 수 없는 해를 찾거나 전혀 다른 개체를 생성하게 한다. 교배 연산자는 두 부모유전자로부터 하나의 새로운 개체들을 생산하게 하고, 이 개체들은 부모 유전자로부터 유전된다.

목적함수는 단지 하나의 개체가 다른 개체들과 비교해 얼마나 더 좋은지에 대한 하나의 척도만을 제공한다. 일반적으로 적합도는 유전자 알고리즘의 교배 여부를 판단하여 개체들의 적합성을 결정하는데 사용된다. 이 적합도에 따라 유전자 알고리즘이 해집단을 수렴시키는데 큰 영향을 미친다.

3. 오류 검색 알고리즘 개발

도로 및 건물은 각각의 폴리곤 객체로 되어 있으므로 각 객체간의 논리적 오류를 판단해야 한다. 이 판단의 기준은 각 폴리곤 객체가 직선으로 되어 있으므로 도로 폴리곤 내의 한번에 해당하는 직선과 건물 폴리곤 내의 한번에 해당하는 직선을 가지고 두 직선의 교차 점을 판단하여 겹침을 확인할 수 있다. 먼저 해당 폴리곤의 최소영역(MBR)을 검색후 교차점 검색을 통해 탐색하도록 하였다.

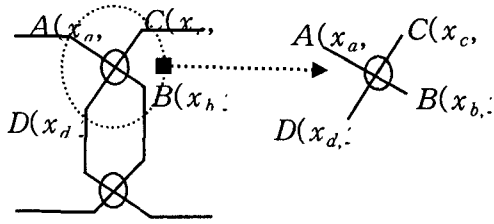


그림 2. 폴리곤 겹침

<그림 2>에서 폴리곤 점을 이용하여 각 점의 기울기를 구하고, 기울기가 같으면 평행이기 때문에 기울기가 같도록 선행조건을 고려하여야 한다.

m_{ab} (A점과 B 점의 기울기),

m_{cd} (C점과 D 점의 기울기) 일때,

$m_{ab} - m_{cd} \neq 0$ (선행조건)

선행조건이 만족하면 (식 1)과 같이 직선 방정식을 만들 수 있다.

$$\frac{(X - X_a)}{(X_b - X_a)} = \frac{(Y - Y_a)}{(Y_b - Y_a)} = t(\text{임의의실수}) \quad \text{-----}(1)$$

$$\frac{(X - X_c)}{(X_d - X_c)} = \frac{(Y - Y_c)}{(Y_d - Y_c)} = s(\text{임의의실수})$$

각각을 X와 Y의 식으로 만들고, t와 s를 (식 2)를 이용하여 결정한다.

$$t = \frac{(x_c - x_a)(y_d - y_c) + (x_d - x_c)(y_a - y_b)}{(x_b - x_a)(y_d - y_c) - (x_d - x_c)(y_b - y_a)} \quad \text{-----}(2)$$

$$s = \frac{(x_c - x_a)(y_b - y_a) + (x_b - x_a)(y_c - y_d)}{(x_b - x_a)(y_d - y_c) - (x_d - x_c)(y_b - y_a)}$$

(식 2)에서 t와 s를 구한 값이 0이 되지 않으면 두 직선은 교차점을 가지고 있다고 판단할 수 있다. 두 폴리곤에 기준인 도로 폴리곤은 고정시키고, 건물 폴리곤을 최소 이동 거리에 맞게 이동시켜야 한다.

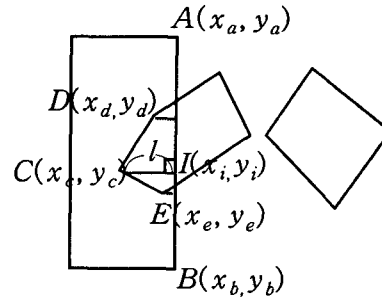


그림 3. 겹친 폴리곤의 최대 거리

<그림 3>에서 폴리곤 객체를 이동 할때 최소의 거리를 가질수 있도록 하기 위하여 검색 과정에서 찾은 폴리곤의 C, D, E점이 도로 폴리곤과 겹치는 건물 폴리곤의 점들이다.

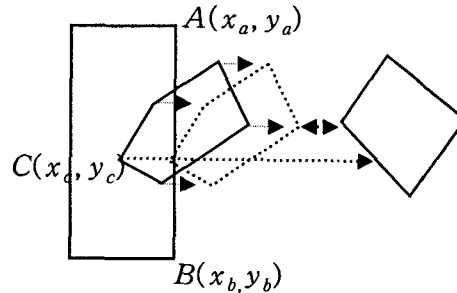


그림 4. C점에서 AB에 수직인 직선

이들 D, C, E점을 <그림 4>와 같이 A점과 B점에 수직이 되는 직선을 그리고, 교점과 거리가 가장 먼 C점을 선택하여 직선을 그린다.

A, B점의 직선방정식은 (식 3)과 같다.

$$Y = mX - mx_a + y_a \quad \text{-----}(3)$$

여기에서, $m = \frac{(y_b - y_a)}{(x_b - x_a)}$ 이다.

C점에서 직선AB에 수직이 되는 직선 방정식은 (식 4)와 같다.

$$Y = -mX + mx_c + y_c \quad \text{-----}(4)$$

(식 4)의 직선이 폴리곤 객체가 이동하는 경로가 되고, 위의 식을 이용하여 교점 $I(X_i, Y_i)$ 를 결정한다.

$$X_i = \frac{1}{2} \left\{ (x_c + x_a) + \frac{(y_c - y_a)}{m} \right\} \quad \text{-----}(5)$$

$$Y_i = \frac{1}{2} \{ m(x_c - x_a) + (y_c + y_a) \}$$

점 $C(x_c, y_c)$ 와 교점 $I(x_i, y_i)$ 의 직선의 거리 l 을 결정하고, 거리 l 이 폴리곤 객체의 최소 이동 거리가 된다.

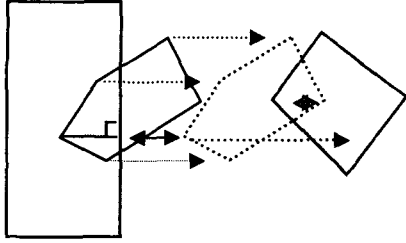


그림 5. 최소거리 이상 이동

<그림 5>와 같이 폴리곤 객체의 거리 만큼 전체 폴리곤 점들을 최소 이동거리로 이동시킨다. 물론 최소 이동거리는 건물간의 완전한 떨어짐이 아니기 때문에 최소 이격 거리를 가지고 이동한다. 최소 이격 허용거리는 사람이 인식할 수 있는 임계값인 0.20mm으로 하였다.

유전자 알고리즘 적용을 위해 각각의 $v(x,y)$ 염색체와 유효한 염색체를 발생시키는 직선방정식을 이용하여 각 폴리곤 객체간의 거리를 가지고 적합도를 평가하는 함수는 각각의 거리의 합이 최소가 되게 한다. 이때 건물 폴리곤 이동에 따르는 건물간의 겹침으로 발생하는 거리는 제외시킨다. 두 개의 shp 파일을 이용하여 하나는 고정된 도로 폴리곤으로, 다른 하나는 움직임이 가능한 건물 폴리곤으로 구성된 지도 데이터를 각 객체를 정렬한다. 먼저 검색 프로그램을 도로 폴리곤과 건물 폴리곤의 겹침점을 리스트 구성하고, 각 리스트내에 유전자를 포함하게 한다.

유전자가 각각의 x,y 좌표를 가지고 있으면 x,y 좌표 범위는 해당 shp 파일의 영역과 검색 프로그램에서 구축된 직선의 방정식 내로 제안한다. 개체군과 최대 세대수는 실험적인 경험적 방법을 이용하여 겹침점 리스트의 총 카운트에 15배 산정하였고, 교배확률은 100%, 2% 돌연변이 확률을 이용하였다.

4. 프로그램 설계 및 개발

4.1 프로그램 설계

일반적인 지도제작과정에서의 논리적 오류는 도로선형을 확대과정에서 많이 발생하며, 이로 인한 도로 및 건물은 각각의 폴리곤 객체로 되어 있으므로 먼저 각 객체간의 논리적 오류를 판단해야 한다. 이를 위해 도로 폴리곤 내의 한변에 해당하는 직선과 건물 폴리곤 내의 한변에 해당하는 두 직선의 교차 점을 판단하여 해당 폴리곤의 최소영역(MBR)을 검색한 후 교차점을 검색하도록 하였다. 한 점에 대해서 교차가 이루어 지면 두 폴리곤은 겹침이므로 더 이상 반복하지 않는다.

<그림 6>은 겹침 폴리곤 판단에 대한 전체적인 흐름도를 나타낸다.

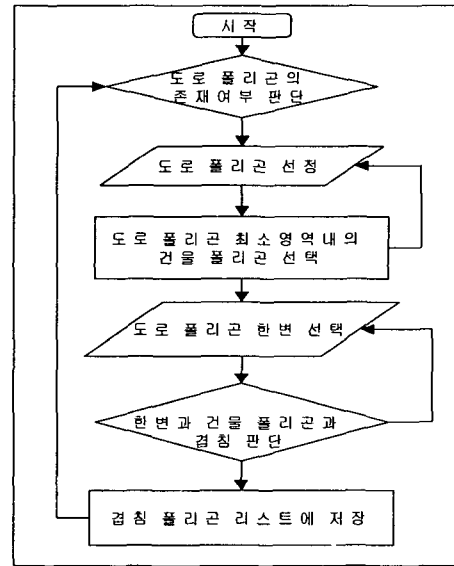


그림 6. 겹침 폴리곤 판단 흐름도

그러나, 건물 폴리곤의 이동으로 인하여 또 다른 건물 폴리곤과 겹치는 논리적 오류가 발생한다. 이 경우 이동하는 전체 폴리곤을 축소하거나 한쪽을 고정하고 이동하는 부분만 축소는 방법이 있으며, 본 연구에서는 <그림 7>과 같이 처리하였다.

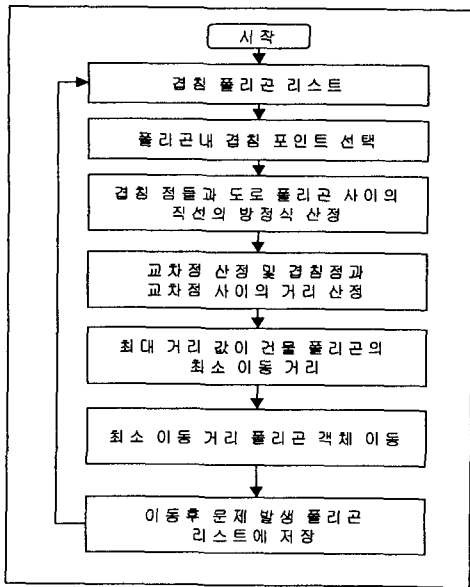


그림 7. 검침 폴리곤의 처리 흐름도

4.2 프로그램 개발

개발 프로그램은 검색과 자료처리 등 두 과정으로 이루어진다. 검색 프로그램은 윈도우 2000을 기본 OS 환경으로 개발언어는 C/C++ 이고, 개발툴은 visual studio 6.0을 이용하였다. 검색 프로그램에 사용되는 데이터는 shp 데이터로 도형 및 속성 정보까지 다룰 수 있어 향후 많은 정보를 추가시키는데 용이하여 사용하였다.

검색 프로그램은 기본적인 shp 데이터를 로드할 수 있고, 이를 각 폴리곤 객체에 대한 검침 점을 찾을 수 있다.

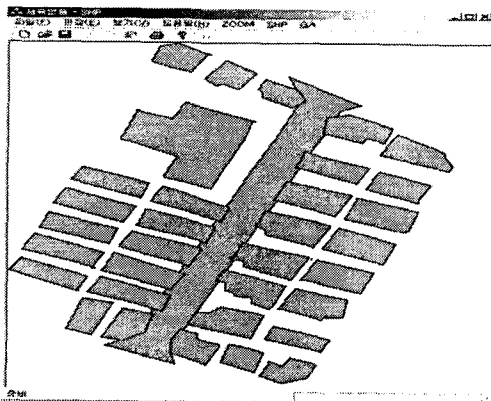


그림 8. 겹쳐진 shp 데이터 로드화면

<그림 8>은 폴리곤 겹침에 있는 shp 데이터를 로드한 것이며, 폴리곤 겹침에 대한

뚜렷한 형상을 확인검침을 판단한다.

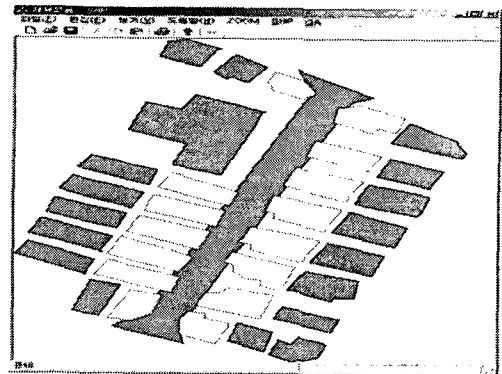


그림 9. 교차 폴리곤 화면

<그림 9>의 검침 폴리곤에 대해 선처리 과정을 거친 후 유효 범위를 보다 세밀하게 줄인 다음 유전자 알고리즘을 이용하여 건물 폴리곤 객체를 이동시킨다.

<그림 10>은 건물 폴리곤이 이동하여 최적화가 되는 모습이나, 이동된 개체의 폴리곤이 다른 폴리곤 객체에 대해 영향을 줄 때 생기는 문제점과 고려해야 할 경우가 아직 정확히 매칭 되지 않아 개체가 잘못 움직이는 것을 볼 수 있다.

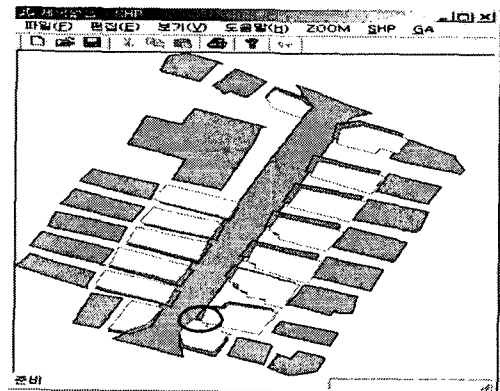


그림 10. 자료처리 결과

이와 같이, 본 연구에서 개발한 프로그램을 실제 지형도에 적용하였으며, 대상지역은 폐합하지 않는 선형 도로가 있는 지역을 선정하였다.

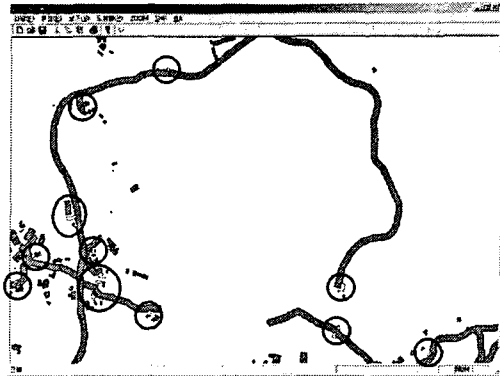


그림 11. 연구대상지역 오류 탐색 결과

<그림 11~12>는 본 연구에서 개발한 오류탐색 프로그램을 이용하여 1/25,000 자동 지도 제작 시에 대부분 수동작업으로 처리되는 도로 과장화를 대상으로 시범연구를 수행하였다. 그림 11.은 도로(폴리곤)의 확대에 의해 발생하는 건물(겹침 폴리곤)을 찾아 표시해 주고 자료 처리한 결과이고, 그림 12는 그림 11에서 탐색된 자료를 본 연구에서 개발한 오류처리 프로그램으로 수정한 결과이다.

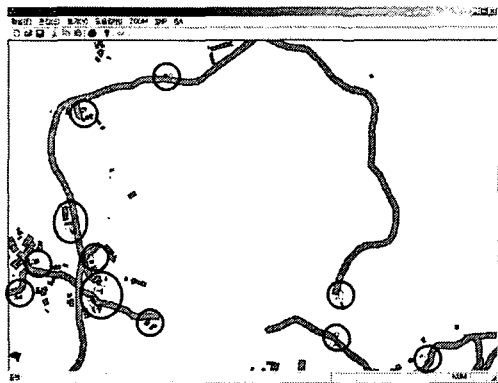


그림 12. 실제 지형도 오류처리결과

5. 결론

본 연구는 지도 제작에 있어서 발생하는 논리적 오류문제를 해결하기 위하여 유전자 알고리즘을 적용한 결과 다음과 같은 결론을 얻을 수 있었다.

1. 지도제작시 발생하는 논리적 오류 검색 및 자료처리 프로그램을 개발함으로써 지도제작 오류와 수작업을 최소화할 수 있었고, 자동화율을 향상할 수 있었다.
2. 유전자알고리즘을 이용한 지도제작 자동화처리 프로그램을 개발함으로써 지도의 미적요소와 시간·경제적인 문제를 해결할 수 있었다.