Graph spectral 기법을 이용한 N:M 대응 폴리곤쌍 탐색
Identification of N:M corresponding polygon pairs using a graph spectral method

허용[1)] · 유기윤[2)]
Huh, Yong · Yu, Ki Yun

[1)] 서울대학교 공과대학 건설환경공학부 박사과정(E-mail : hy7808@snu.ac.kr)
[2)] 정회원, 서울대학교 공과대학 건설환경공학부 부교수(E-mail : kiyun@snu.ac.kr)

Abstract

Combined with the indeterminate boundaries of spatial objects, n:m correspondences makes an object-based matching be a complex problem. In this study, we model the boundary of a polygon object with fuzzy model and describe their overlapping relations as a weighted bipartite graph. Then corresponding pairs including 1:0, 1:1, 1:n and n:m relations are identified using a spectral singular value decomposition.

▸Keywords : Polygon dataset, N:M correspondence, Fuzzy boundary, Spectral SVD analysis

1. 서 론

Bel Hadj Ali(2000)는 두 폴리곤 집합 A와 B의 개별 폴리곤을 그래프의 노드로 표현하고 어떤 폴리곤 객체 A_i와 B_j 사이에 중첩이 발생하면 에지 $e_{Ai,\,Bj}$가 생성되는 이분 그래프를 이용하여 대응쌍을 탐지하였다. 하지만 이 방법은 중첩 여부에 의하여 에지의 생성 여부가 0과 1로 결정되므로 두 자료의 위치 및 형상 오차에 민감한 단점이 있다. 본 연구에서는 Zhan (1984)의 기법에 따라 폴리곤의 경계를 퍼지 모형으로 표현하고 중첩이 발생할 경우 퍼지 함수의 교점을 이용하여 0부터 1사이의 연속값으로 가중치를 부여하였다. 이를 이용하여 가중 이분 그래프를 생성하고 그래프 분광 분석을 이용하여 대응쌍을 탐색하였다.

2. 연구방법 및 내용

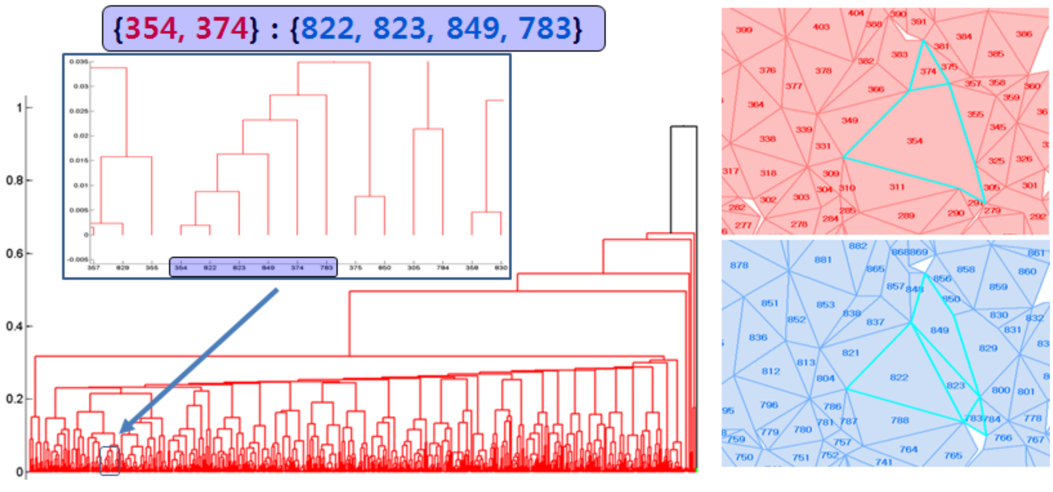
그래프의 분할은 어떤 그래프의 라플라시안 행렬의 Fiedler 벡터를 이용하여 얻을 수 있다. 이 벡터의 어떤 두 원소값 사이의 차이는 각각의 원소값에 대응되는 두 노드 사이의 에지의 가중치에 반비례한다. 즉, 그래프 상에서 높은 인접성을 가지는 노드들은 해당 노드에 대응되는 Fiedler 벡터의 원소값들이 유사한 특성을 가지게 되며, 이들 원소값을 대상으로 군집화 기법을 적용하면 그래프 분할을 수행할 수 있게 된다. Dhillon(2001)의 Spectral Singular Value Decomposition(SVD) 알고리즘은 일반적인 그래프 분할을 두 자료집합 사이의 공군집화(co-clustering)로 확장한다. 가중 이분 그래프의 인접성 행렬 A와 각 노드의 밀집도를 의미하는 대각 행렬 $D_1(i,i)=\sum_j A_{ij}$과 $D_2(j,j)=\sum_i A_{ij}$을 이용하면 식 (1)과 같이 generalized eigenproblem을 얻게 된다.

$$\begin{bmatrix} D_1 & -A \\ -A^T & D_2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} \tag{1}$$

$u = D_1^{1/2}x$과 $v = D_2^{1/2}y$를 이용하면 식 (1)은 식 (2)와 같이 정리된다. $A_n = D_1^{-1/2}A\,D_2^{-1/2}$는 행렬 A의 정규화된 형태이므로 u와 v는 각각 정규 인접성 행렬 A_n의 좌우 특이벡터이고, $1-\lambda$는 특이값이 된다. 식 (2)의 경우 $1-\lambda$가 특이값이 되므로 식(1)의 Fiedler 벡터는 A_n와 A_n^T의 두 번째 특이벡터를 이용하여 $\left[D_1^{-1/2}u_2; D_2^{-1/2}v_2\right]$과 같이 얻어진다.

$$\begin{aligned} D_1^{1/2}x \quad - \quad D_1^{1/2}Ay \; &= \lambda D_1^{1/2}x \\ -D_2^{1/2}A^Tx \quad + \quad D_2^{1/2}y \quad &= \lambda D_2^{1/2}y \end{aligned} \;\Rightarrow\; \begin{aligned} D_1^{-1/2}A\,D_2^{-1/2}\,v &= (1-\lambda)u \\ D_2^{-1/2}A^TD_1^{-1/2}u &= (1-\lambda)v \end{aligned} \;\Rightarrow\; \begin{aligned} A_nv &= (1-\lambda)u \\ A_n^Tu &= (1-\lambda)v \end{aligned} \tag{2}$$

Dhillon(2001)은 u_2, v_2 이외에 복수의 특이벡터를 이용하여 l차원의 공간에서의 분류를 수행하였다. 만약 $U = [u_2, \ldots, u_{l+1}]$, $V = [v_2, \ldots, v_{l+1}]$이라면 $Z = \left[D_1^{-1/2}U; D_2^{-1/2}V\right]$을 통하여 각각의 노드별 인접성은 기하학적 좌표와 좌표 사이의 인접성으로 표현될 수 있다. 따라서 높은 중첩 기반의 연결성을 가지는 폴리곤 객체들의 l차원의 공간에서 정의되는 좌표집합 Z의 대응 좌표들이 서로 기하학적으로 근접하게 된다. 좌표집합 Z의 상호 거리를 이용하여 덴드로그램으로 표현하였을 때, 하나의 노드로부터 분리되는 sub-tree의 leaf들에 대응되는 폴리곤들은 높은 인접성을 가지므로 대응관계를 구성하는 폴리곤 집합이 된다. [그림 1]은 각각 400여개의 폴리곤을 가상 생성한 두 개의 폴리곤 집합을 대상으로 적용한 결과로 2:4 대응과 같은 n:m 대응쌍을 탐지할 수 있었다.



[그림 1] Spectral SVD를 이용한 폴리곤 집합의 인접성 덴드로그램과 N:M 대응

감사의 글

본 연구는 국토해양부 첨단도시기술개발사업 - 지능형국토정보기술혁신 사업과제의 연구비지원(07국토정보C04)에 의해 수행되었습니다.

참고문헌

Bel Hadj Ali, A. (2000), Mesure entre entite´s surfaciques-Application a` la qualification des liens d'appariement, Bulletin d'Information de l'IGN, Vol, 71, pp. 33-43.

Dhillon, I. S. (2001), Co-clustering documents and words using bipartite spectral graph partitioning, 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD, San Francisco, California, pp. 269-274.

Zhan, F. B. (1984), Approximate analysis of binary topological relations between geographic regions with indeterminate boundaries, SoftComputing, Vol. 2, pp. 28-34.