

모바일 사용자를 위한 효율적인 스카이라인 계산¹⁾

김지현, 김 명

이화여자대학교 컴퓨터공학과

e-mail: jhrosa@ewhain.net, mkim@ewha.ac.kr

Efficient Skyline Computation for Mobile Users

Jihyun Kim, Myung Kim

Dept. of Computer Science & Engineering, Ewha Womans University

요 약

스카이라인 계산은 사용자에게 의미 있는 정보를 추천하기 위해 많이 사용되고 있다. 일반적으로 스카이라인은 대용량 데이터 집합을 대상으로 하기 때문에 사전에 계산해 놓고 사용한다. 그러나 사용자가 이동하는 경우에는 이와 같이 계산해 놓은 스카이라인이 사용자의 위치 조건을 반영한 것이 아니므로 부적절할 수 있다. 본 연구에서는 사용자가 이동하는 조건을 고려하여 사용자에게 실시간으로 최적화된 스카이라인을 지속적으로 제공하는 기법을 제안한다. 제안한 기법은 실험을 통해 그 효율성을 검증하였다.

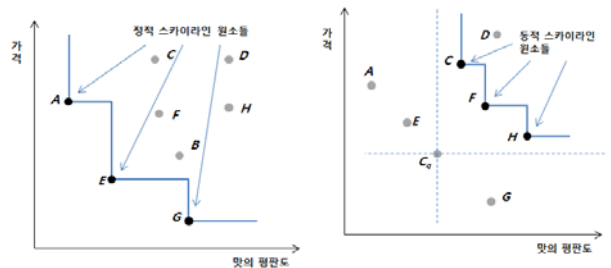
1. 서론

스마트폰과 같은 모바일 기기의 사용의 증가함에 따라 이동 환경에서 신속하고 정확한 정보를 제공하는 추천 서비스에 대한 필요성이 급증하였다. 이러한 추천 서비스를 위해 많이 사용되고 있는 연산으로 스카이라인(skyline) 계산을 들 수 있다. 스카이라인 계산은 다중 특성을 고려하여 사용자에게 최적인 데이터를 제공하는 질의 계산으로, 최근에 다양한 기법들이 개발되었다 [1, 2, 3, 4, 5, 7, 8].

스카이라인은 다차원 점들로 구성된 데이터 집합으로부터 사용자의 선호도를 고려하였을 때, 집합내의 다른 점들에 지배되지 않는 점들로 구성된 가장 큰 부분 집합이다. 예를 들어, 그림 1(a)를 보자. 이 그림의 점들은 8개의 레스토랑 정보를 2차원 속성 (맛 평판도와 가격) 으로 나타낸다. 그림에서와 같이 레스토랑 A, E, G 는 다른 레스토랑보다 우수한 레스토랑들로 이들이 스카이라인을 이룬다. 단, 가격과 맛의 평판도는 낮은 점수가 더 우수하다고 가정하자.

이제 모바일 사용자의 경우를 생각해 보자. 사용자가 이동하는 환경에서는 사용자의 위치 정보를 고려해야 한다. 위치를 제외한 다른 속성들이 우수하여 스카이라인에

포함될 수 있다고 해도, 그 레스토랑이 사용자의 현 위치에서 아주 멀리 있다면 그러한 레스토랑을 추천하는 것은 이동 중인 사용자에게 별 도움이 되지 않을 것이다. 그림 1(b)에서는 사용자가 C_q 지점에 현재 위치하고 있다고 가정하고 계산한 스카이라인을 보인다. 이 스카이라인에는 레스토랑 C, F 와 H 가 포함되며 정적인 스카이라인과 다르다는 것을 알 수 있다.



(a) 정적 스카이라인 (b) 동적 스카이라인

그림 1. 스카이라인 예제

본 연구에서는 이러한 점을 고려하여 이동하는 사용자에게 지속적으로 적절한 스카이라인을 계산해 주는 알고리즘을 제안하고자 한다. 논문은 다음과 같이 구성된다. 2절에서는 기존 연구를 소개하고, 3절에서는 제안하는 스카이라인 계산 기법을 소개하며, 4절에서는 이 기법들을 분석 및 평가하고, 5절에서는 결론을 맺는다.

1) 이 논문은 2013년도 정부(교육부)의 재원으로 한국연구재단의 기초연구사업 지원을 받아 수행된 것임. (2013R1A1A2013124).

2. 관련 연구

스카이라인 계산은 [1]에 의해 처음으로 소개되었으며, 그 이후로 다양한 연구들이 진행되고 있다 [2, 3]. 스카이라인 계산은 정적 속성들만을 고려한 계산과 동적 속성 (이동 환경 고려) 포함한 계산들로 구분될 수 있다. 정적 속성만을 고려한 경우는 필터를 이용하여 스카이라인에 속할 가능성이 없는 데이터들을 사전에 제거하거나 [2], 데이터들을 정렬하여 비교횟수를 줄이거나 [3], 분할 정복 기법을 사용하여 데이터 간의 비교를 국부적으로 진행시키거나 [5], 정렬된 데이터들을 스캔해 나가면서 특정 조건에 만족하면 데이터 스캔을 멈추는 정지선 기법들이 제안되었다 [3].

이동 환경을 고려하여 스카이라인을 계산한 연구들도 있다. 사용자의 위치를 고려하지 않고 대상 객체의 고정 속성인 정적 속성만을 고려하여 계산하는 기법들이 [2, 3, 4]에 제시되어 있다. 이동 환경에서의 스카이라인 질의 계산은 공간적으로 고정되어 있다는 가정 하에 R-트리나 각도 파티션 기법, 그리드 기법을 이용하여 계산하거나 사전에 유효 영역을 결정하여 이용하는 기법 [5, 6]을 사용하였다. 또한 모바일 상황에서의 품질이 우수한 스카이라인을 추천하기 위한 스카이라인 계산 기법 [7, 8]들이 있다.

기존의 이러한 이동 환경을 고려한 연구들은 사용자의 위치를 고려하여 최적의 스카이라인을 실시간으로 제공하는 것은 못한다. 본 연구에서는 사용자의 위치 조건을 고려하여 사용자 주변의 특정 거리 이내의 데이터들로부터 스카이라인을 계산하여 제공하는 것을 목표로 한다. 본 연구는 사용자가 이동할 때마다 스카이라인을 새로 계산하는 것이 아니라, 이전의 위치에서 계산한 스카이라인을 갱신해 가면서 새로운 위치에서의 스카이라인을 신속하게 계산해주는 장점이 있다.

3. 제안하는 스카이라인 계산 기법

본 연구에서는 사용자가 이동 중인 상황을 고려하여 사용자 질의가 연속적으로 발생하는 경우, 사용자의 위치에 최적화된 신속하고 우수한 스카이라인 계산 기법을 제안한다. 제안하는 기법은 두 단계로 구성된다. 첫 단계에서는 구역별 스카이라인 계산을 하고 둘째 단계에서는 사용자의 질의를 실시간으로 처리한다. 제안하는 알고리즘들은 다음과 같이 기술될 수 있다.

알고리즘 1: 구역별 스카이라인 계산

사용자가 이동할 수 있는 지역을 여러 파티션으로 분할하고 각 파티션 단위로 스카이라인을 구해 놓는다. 이 때 스

카이라인 계산에는 위치 정보는 포함하지 않는다. 이 스카이라인 계산에는 본 연구팀이 개발한 스카이라인 계산을 위한 전처리 기법 [5, 7]들을 사용한다. 이 기법들은 필터를 이용하여 스카이라인을 신속하게 계산해 준다.

알고리즘 2: 사용자 위치를 고려한 동적 스카이라인 계산

사용자의 출발점을 기준으로 사전에 정해진 특정 거리 (ϵ) 이내에 포함된 데이터 집합에 대해 스카이라인을 계산하여 제공한다. 사용자가 특정 거리 (δ) 이상으로 이동하게 되면, 새로운 위치가 정해지고, 그 위치를 중심으로 반경 (ϵ) 이내의 데이터를 기준으로 스카이라인을 실시간으로 갱신하여 제공한다.

그림 2를 사용하여 알고리즘2를 설명하고자 한다. 사용자가 그림 2(a) 상의 지도에서 이동하고 있다고 가정하자. 사용자는 Q, Q', Q'' 방향으로 이동하고 있다고 하자. 여기에서 Q 를 중심으로 한 사각형의 반경은 ϵ 에 해당한다. $|Q - Q'| = \delta$ 에 해당한다. 사용자가 이와 같이 δ 이상의 거리를 이동하게 되면 Q' 를 중심으로 하는 새로운 사각형 내의 점들을 대상으로 스카이라인을 구해주는 것이다. 여기서 새로운 위치에서의 스카이라인 계산에는 기존에 구해 놓은 그리드 셀 별 스카이라인과 Q 를 중심으로 한 사각형의 스카이라인을 활용하여 신속하게 구하는 것이다.

제안하는 기법은 사용자가 특정 거리 이내를 벗어나는 경우에 있어서도 스카이라인을 다시 전체 계산하는 것이 아니라 질의 영역에 한번 계산된 스카이라인이 포함될 경우에는 계산된 스카이라인을 이용할 수 있는 장점이 있다. 또한 그리드 내의 새로운 원소가 추가되는 경우에도 데이터 집합에 대해 스카이라인을 전체 계산하는 것이 아니라 추가되는 원소와 해당 그리드 셀을 위해 구해 놓은 스카이라인만을 비교하여 스카이라인을 갱신할 수 있어 효율적이다.

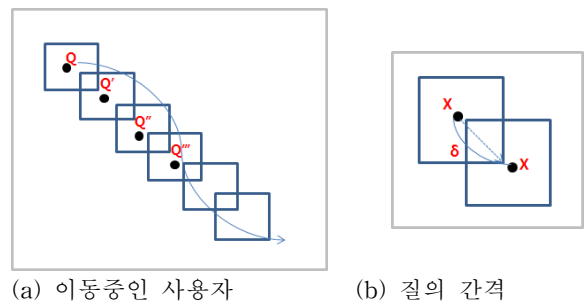


그림 2. 이동 환경에서의 스카이라인 계산 예제

4. 알고리즘 분석과 성능 평가

본 연구에서는 빠르게 이동 중인 다수의 사용자들을 대상으로 질 높은 추천 데이터를 신속하게 제공해 주는

알고리즘을 개발하였다. 추천되는 데이터는 이동 중인 사용자의 주변에 위치한 데이터 집합에서 추출된 스카이라인 데이터이다. 제안한 스카이라인 계산 알고리즘은 사용자가 이동하면서 사용자를 중심으로 특정 거리 이내의 데이터 집합에 대해 스카이라인을 신속하게 제공해 준다. 즉, 이 알고리즘은 추천 서비스를 하려는 대상 데이터 집합이 수시로 변경된다고 해도 이를 즉시 반영하여 스카이라인이 계산될 수 있도록 하며, 클라이언트 측에서 계산된 스카이라인을 이동 중인 사용자의 다음 위치에서 계산할 스카이라인 계산에 수정하여 재사용할 수 있다는 점이 효율적이다.

5. 결론

본 연구에서는 이동 중인 사용자에게 신속하게 스카이라인을 계산하여 제공해 주는 기법을 제안하였다. 제안한 알고리즘은 사용자의 현재 위치와 반경 정보를 고려하여 사용자 주변에서 최적인 스카이라인을 지속적으로 계산하여 제공한다. 연속적으로 계산되는 스카이라인들을 다음 스카이라인 계산에 재활용하여 계산 속도를 높였다. 처리 속도가 빨라서 다수의 사용자들을 지원할 수 있으며, 클라이언트 서버 환경에서 실제 사용이 가능한 기법으로 평가된다.

참고문헌

- [1] S. Borzsonyi, D. Kossmann, and K. Stocker, "The skyline operator," In *ICDE*, pp. 421 - 430, 2001
- [2] J. Chomicki, P. Godfrey, J. Gryz, and D. Liang, "Skyline with presorting," In *ICDE*, pp. 717-719, 2003.
- [3] I. Bartolini, P. Ciaccia, and M. Patella, "SaLSa: Computing the skyline without scanning the whole sky," In *CIKM*, pp. 405-414, 2006.
- [4] A. Vlachou, C. Doulkeridis, and Y. Kotidis, "Angle-based space partitioning for efficient parallel skyline computation," In *ACM SIGMOD*, pp. 227-238, 2008.
- [5] J. Kim, and M. Kim, "Skyline Computation Using Adaptive Filters," *International Journal of Computer Science and Information Technology & Security(IJCSITS)*, Vol 2, No. 2, pp. 431-434, Apr. 2012.
- [6] Z. Hung, H. Lu, B. Ooi and A.Tung, "Continuous Skyline queries for moving objects," *IEEE TKDE*, Vol. 18, pp. 377-391, 2006.
- [7] J. Kim, and M. Kim, "An Extended Skyline Computation Scheme for Recommendation Services in Mobile Environments," *Journal of KIISE:Computing Practice and Letters*, Vol. 18, No. 7, pp. 558-562, Jul., 2012.
- [8] J. Kim, and M. Kim, "Skyline Computation permitting Dynamic Determination of Query Regions," *International Journal of Computer Science and Information Technology & Security (IJCSITS)*, Vol. 2, No. 5, pp. 939-943, Oct. 2012.