



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2019-0079347
(43) 공개일자 2019년07월05일

(51) 국제특허분류(Int. Cl.)

G06F 16/00 (2019.01)

(52) CPC특허분류

G06F 16/245 (2019.01)

G06F 16/2228 (2019.01)

(21) 출원번호 10-2017-0181476

(22) 출원일자 2017년12월27일

심사청구일자 2017년12월27일

(71) 출원인

서강대학교산학협력단

서울특별시 마포구 백범로 35 (신수동, 서강대학교)

(72) 발명자

정성원

서울특별시 양천구 오목로 300, 604호(목1동, 현대하이페리온2 204동)

유형철

서울특별시 마포구 대흥로 48, 405호(대흥동, 프라임원룸텔)

(74) 대리인

정부연

전체 청구항 수 : 총 15 항

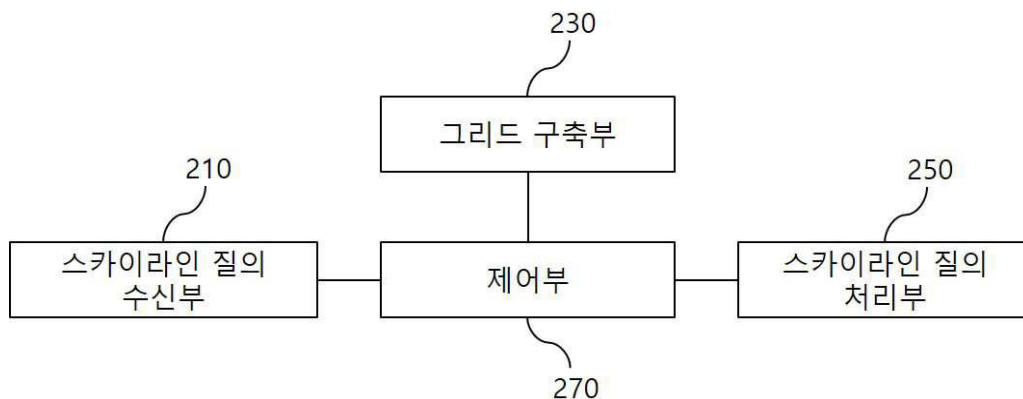
(54) 발명의 명칭 **그리드 기반의 스카이라인 질의처리장치 및 방법, 이를 기록한 기록매체**

(57) 요약

본 발명은 그리드 기반의 스카이라인 질의처리장치 및 방법에 관한 것으로 스카이라인 질의 상의 적어도 일부 속성을 원점에서 멀어질수록 증가하는 단위를 가지고 데이터를 포함할 수 있는 단위 블록들로 차원화 하여 그리드를 생성하는 그리드 생성부 및 상기 그리드 상에서 상기 단위 블록들 중 적어도 지배당하지 않고 해당 데이터를 반드시 포함하는 단위 블록을 결정하여 상기 생성된 그리드를 갱신하는 그리드 갱신부를 포함한다. 따라서, 본 발명은 스카이라인을 위한 새로운 형태의 인덱스 구조를 통해 효율적인 질의 처리가 가능하다.

대표도 - 도2

130



이 발명을 지원한 국가연구개발사업

과제고유번호 201739036

부처명 과학기술정보통신부

연구관리전문기관 정보통신기술진흥센터

연구사업명 정보통신기술인력양성

연구과제명 적응형 블록체인 플랫폼 기술 개발 및 전문 인력 양성

기 여 율 1/1

주관기관 서강대학교산학협력단

연구기간 2017.06.01 ~ 2017.12.31

명세서

청구범위

청구항 1

스카이라인 질의 상의 적어도 일부 속성을 원점에서 멀어질수록 증가하는 단위를 가지고 데이터를 포함할 수 있는 단위 블록들로 차원화 하여 그리드를 생성하는 그리드 생성부; 및

상기 그리드 상에서 상기 단위 블록들 중 적어도 지배당하지 않고 해당 데이터를 반드시 포함하는 단위 블록을 결정하여 상기 생성된 그리드를 갱신하는 그리드 갱신부를 포함하는 그리드 기반의 스카이라인 질의처리장치.

청구항 2

제1항에 있어서, 상기 그리드 생성부는

상기 원점에서 멀어질수록 각 차원의 분할점이 등차수열 형태로 증가하도록 차원화 하는 것을 특징으로 하는 그리드 기반의 스카이라인 질의처리장치.

청구항 3

제1항에 있어서, 상기 그리드 갱신부는

상기 그리드 상에서 상기 단위 블록들 중 적어도 지배당하지 않고 해당 데이터를 반드시 포함하는 단위 블록 각각이 적어도 하나의 서브-단위 블록을 포함하도록 갱신하는 것을 특징으로 하는 그리드 기반의 스카이라인 질의처리장치.

청구항 4

제1항에 있어서,

상기 갱신된 그리드를 통해 맵-리듀스(Map-Reduce) 연산을 수행하는 맵-리듀스 연산부를 더 포함하는 것을 특징으로 하는 그리드 기반의 스카이라인 질의처리장치.

청구항 5

제4항에 있어서, 상기 맵-리듀스 연산부는

상기 갱신된 그리드 상의 모든 단위 블록에 대해 각도 기반 분할(angular partitioning)을 통해 적어도 하나의 그룹으로 분할하는 것을 특징으로 하는 그리드 기반의 스카이라인 질의처리장치.

청구항 6

제4항에 있어서, 상기 맵-리듀스 연산부는

맵(map) 연산 단계에서 상기 갱신된 그리드에 대해 비스카이라인(non-skyline) 데이터를 프루닝(pruning)한 후 콤바이너(combiner) 연산 단계로 전달하는 것을 특징으로 하는 그리드 기반의 스카이라인 질의처리장치.

청구항 7

제6항에 있어서, 상기 맵-리듀스 연산부는

상기 컴바이너 연산 단계에서 상기 그리드 상의 각 단위 블록을 행 우선 순서로 검색하면서 지역적 스카이라인을 결정 한 후 리듀스(reduce) 연산 단계로 전달하는 것을 특징으로 하는 그리드 기반의 스카이라인 질의처리장치.

청구항 8

제7항에 있어서, 상기 맵-리듀스 연산부는

상기 리듀스 연산 단계에서 상기 그리드 상의 각 단위 블록을 행 우선 순서로 검색하면서 전역적 스카이라인을 결정하는 것을 특징으로 하는 그리드 기반의 스카이라인 질의처리장치.

청구항 9

그리드 기반의 스카이라인 질의처리장치에서 수행되는 스카이라인 질의처리방법에 있어서,

- (a) 적어도 하나의 속성을 포함하는 스카이라인 질의를 수신하는 단계;
- (b) 상기 스카이라인 질의 상의 적어도 일부 속성을 기초로 그리드를 구축하는 단계; 및
- (c) 상기 구축된 그리드를 통해 맵-리듀스 연산을 수행하여 상기 스카이라인 질의를 처리하는 단계를 포함하는 그리드 기반의 스카이라인 질의처리방법.

청구항 10

제9항에 있어서, 상기 (b) 단계는

- (b-1) 스카이라인 질의 상의 적어도 일부 속성을 원점에서 멀어질수록 증가하는 단위를 가지고 데이터를 포함할 수 있는 단위 블록들로 차원화 하여 그리드를 생성하는 단계; 및
- (b-2) 상기 그리드 상에서 상기 단위 블록들 중 적어도 지배당하지 않고 해당 데이터를 반드시 포함하는 단위 블록을 결정하여 상기 생성된 그리드를 갱신하는 단계를 포함하는 것을 특징으로 하는 그리드 기반의 스카이라인 질의처리방법.

청구항 11

제10항에 있어서, 상기 (b-1) 단계는

상기 원점에서 멀어질수록 각 차원의 분할점이 등차수열 형태로 증가하도록 차원화 하는 단계인 것을 특징으로 하는 그리드 기반의 스카이라인 질의처리방법.

청구항 12

제10항에 있어서, 상기 (b-2) 단계는

상기 그리드 상에서 상기 단위 블록들 중 적어도 지배당하지 않고 해당 데이터를 반드시 포함하는 단위 블록 각각이 적어도 하나의 서브-단위 블록을 포함하도록 갱신하는 단계인 것을 특징으로 하는 그리드 기반의 스카이라인 질의처리방법.

청구항 13

제9항에 있어서, 상기 (c) 단계는

상기 생성된 그리드 상의 모든 단위 블록에 대해 각도 기반 분할(angular partitioning)을 통해 적어도 하나의 그룹으로 분할하여 맵-리듀스 연산을 수행하는 단계인 것을 특징으로 하는 그리드 기반의 스카이라인 질의처리 방법.

청구항 14

제9항에 있어서, 상기 (c) 단계는

(c-1) 상기 갱신된 그리드에 대해 비스카이라인(non-skyline) 데이터를 프루닝(pruning)한 후 컴바이너(combiner) 연산 단계로 전달하는 맵(map) 연산 단계;

(c-2) 상기 그리드 상의 각 단위 블록을 행 우선 순서로 검색하면서 지역적 스카이라인을 결정 한 후 리듀스(reduce) 연산 단계로 전달하는 컴바이너 연산 단계; 및

(c-3) 상기 그리드 상의 각 단위 블록을 행 우선 순서로 검색하면서 전역적 스카이라인을 결정하는 리듀스 연산 단계를 포함하는 것을 특징으로 하는 그리드 기반의 스카이라인 질의처리방법.

청구항 15

그리드 기반의 스카이라인 질의처리장치에서 수행되는 스카이라인 질의처리방법을 기록하는 컴퓨터 수행 가능한 기록매체에 있어서,

적어도 하나의 속성을 포함하는 스카이라인 질의를 수신하는 과정;

상기 스카이라인 질의 상의 적어도 일부 속성을 기초로 그리드를 구축하는 과정; 및

상기 생성된 그리드를 통해 맵-리듀스 연산을 수행하여 상기 스카이라인 질의를 처리하는 과정을 포함하는 기록 매체.

발명의 설명

기술 분야

[0001] 본 발명은 스카이라인 질의처리 기술에 관한 것으로, 보다 상세하게는 스카이라인을 위한 새로운 형태의 인덱스 구조를 통해 효율적인 질의 처리가 가능한 그리드 기반의 스카이라인 질의처리장치 및 방법, 이를 기록한 기록 매체에 관한 것이다.

배경 기술

[0003] 스카이라인 질의는 the maxima of a set of vectors 문제를 데이터베이스 분야로 확장한 최초의 연구이다. 이후, 초기 스카이라인 질의는 인덱스를 사용하지 않는 BNL, SFS 같은 기법에서부터, B^+ -tree 및 R^* -tree 등 기존에 개발된 인덱스 기법을 이용한 단일 머신 기반의 shooting star 기법 등으로 발전하였다. 빅 데이터 시대가 도래하면서 데이터가 점점 대용량화 되어 스카이라인 질의는 대용량 데이터에 대해서도 빠른 응답을 요구하게 되었다.

[0004] 한국등록특허 제10-1344649(2013.12.17)호는 해시 기반의 스카이라인 질의 처리 방법 및 그 장치에 관한 것으로, 스카이라인 질의를 처리하기 위한 전체 데이터에 대한 해시 테이블을 형성하여 이를 통해 후보 집합을 우선 생성함으로써 보다 사용자에게 유리한 구조를 제공할 수 있고, 해시 테이블을 통해 질의의 해가 될 수 있는 후보군 데이터 집합을 구함으로써 스카이라인 질의가 $O(n)$ 의 비용으로 처리될 수 있어 질의에 대한 빠른 응답을 제공할 수 있다.

[0005] 한국등록특허 제10-1554569(2015.09.15)호는 스카이라인 질의 시스템 및 방법에 관한 것으로, 쿼드 트리를 사용하여 데이터의 분포에 대한 특성을 파악하고, 파악한 특성에 따라 스카이라인에 포함될 수 있는 데이터만을 분산시킬 수 있고, 각각의 머신에서 처리한 결과를 최종적으로 하나의 머신에서 처리하는 종래의 분산 처리 알고리즘과 달리 필요한 일부분의 데이터를 복사함으로써 연산이 하나의 머신에 몰리는 경우를 없앴기 때문에 종래의 기술보다 더 빠르게 스카이라인을 계산할 수 있다.

선행기술문헌

특허문헌

[0007] (특허문헌 0001) 한국등록특허 제10-1344649(2013.12.17)호
(특허문헌 0002) 한국등록특허 제10-1554569(2015.09.15)호

발명의 내용

해결하려는 과제

[0008] 본 발명의 일 실시예는 스카이라인을 위한 새로운 형태의 인덱스 구조를 통해 효율적인 질의 처리가 가능한 그리드 기반의 스카이라인 질의처리장치 및 방법을 제공하고자 한다.

[0009] 본 발명의 일 실시예는 2단계 그리드를 이용하여 맵-리듀스 연산을 수행함으로써 스카이라인의 특징을 활용할 수 있는 그리드 기반의 스카이라인 질의처리장치 및 방법을 제공하고자 한다.

[0010] 본 발명의 일 실시예는 2단계 그리드를 이용하면서도 1번의 맵-리듀스 연산만을 사용함으로써 디스크 I/O를 줄일 수 있는 그리드 기반의 스카이라인 질의처리장치 및 방법을 제공하고자 한다.

과제의 해결 수단

[0012] 실시예들 중에서, 그리드 기반의 스카이라인 질의처리장치는 스카이라인 질의 상의 적어도 일부 속성을 원점에서 멀어질수록 증가하는 단위를 가지고 데이터를 포함할 수 있는 단위 블록들로 차원화 하여 그리드를 생성하는 그리드 생성부 및 상기 그리드 상에서 상기 단위 블록들 중 적어도 지배당하지 않고 해당 데이터를 반드시 포함하는 단위 블록을 결정하여 상기 생성된 그리드를 갱신하는 그리드 갱신부를 포함한다.

[0013] 상기 그리드 생성부는 상기 원점에서 멀어질수록 각 차원의 분할점이 등차수열 형태로 증가하도록 차원화 할 수 있다.

[0014] 상기 그리드 갱신부는 상기 그리드 상에서 상기 단위 블록들 중 적어도 지배당하지 않고 해당 데이터를 반드시 포함하는 단위 블록 각각이 적어도 하나의 서브-단위 블록을 포함하도록 갱신할 수 있다.

[0015] 상기 그리드 기반의 스카이라인 질의처리장치는 상기 갱신된 그리드를 통해 맵-리듀스(Map-Reduce) 연산을 수행하는 맵-리듀스 연산부를 더 포함할 수 있다.

[0016] 상기 맵-리듀스 연산부는 상기 갱신된 그리드 상의 모든 단위 블록에 대해 각도 기반 분할(angular partitioning)을 통해 적어도 하나의 그룹으로 분할할 수 있다.

[0017] 상기 맵-리듀스 연산부는 맵(map) 연산 단계에서 상기 갱신된 그리드에 대해 비스카이라인(non-skyline) 데이터를 프루닝(pruning)한 후 콤바이너(combiner) 연산 단계로 전달할 수 있다.

[0018] 상기 맵-리듀스 연산부는 상기 콤바이너 연산 단계에서 상기 그리드 상의 각 단위 블록을 행 우선 순서로 검색하면서 지역적 스카이라인을 결정 한 후 리듀스(reduce) 연산 단계로 전달할 수 있다.

[0019] 상기 맵-리듀스 연산부는 상기 리듀스 연산 단계에서 상기 그리드 상의 각 단위 블록을 행 우선 순서로 검색하면서 전역적 스카이라인을 결정할 수 있다.

[0020] 실시예들 중에서, 그리드 기반의 스카이라인 질의처리장치에서 수행되는 스카이라인 질의처리방법에 있어서,

(a) 적어도 하나의 속성을 포함하는 스카이라인 질의를 수신하는 단계, (b) 상기 스카이라인 질의 상의 적어도 일부 속성을 기초로 그리드를 구축하는 단계 및 (c) 상기 생성된 그리드를 통해 맵-리듀스 연산을 수행하여 상기 스카이라인 질의를 처리하는 단계를 포함한다.

- [0021] 상기 (b) 단계는 (b-1) 스카이라인 질의 상의 적어도 일부 속성을 원점에서 멀어질수록 증가하는 단위를 가지고 데이터를 포함할 수 있는 단위 블록들로 차원화 하여 그리드를 생성하는 단계 및 (b-2) 상기 그리드 상에서 상기 단위 블록들 중 적어도 지배당하지 않고 해당 데이터를 반드시 포함하는 단위 블록을 결정하여 상기 생성된 그리드를 갱신하는 단계를 포함할 수 있다.
- [0022] 상기 (b-1) 단계는 상기 원점에서 멀어질수록 각 차원의 분할점이 등차수열 형태로 증가하도록 차원화 하는 단계를 포함할 수 있다.
- [0023] 상기 (b-2) 단계는 상기 그리드 상에서 상기 단위 블록들 중 적어도 지배당하지 않고 해당 데이터를 반드시 포함하는 단위 블록 각각이 적어도 하나의 서브-단위 블록을 포함하도록 갱신하는 단계를 포함할 수 있다.
- [0024] 상기 (c) 단계는 상기 생성된 그리드 상의 모든 단위 블록에 대해 각도 기반 분할(angular partitioning)을 통해 적어도 하나의 그룹으로 분할하여 맵-리듀스 연산을 수행하는 단계를 포함할 수 있다.
- [0025] 상기 (c) 단계는 (c-1) 상기 갱신된 그리드에 대해 비스카이라인(non-skyline) 데이터를 프루닝(pruning)한 후 컴바이너(combiner) 연산 단계로 전달하는 맵(map) 연산 단계, (c-2) 상기 그리드 상의 각 단위 블록을 행 우선 순서로 검색하면서 지역적 스카이라인을 결정 한 후 리듀스(reduce) 연산 단계로 전달하는 컴바이너 연산 단계 및 (c-3) 상기 그리드 상의 각 단위 블록을 행 우선 순서로 검색하면서 전역적 스카이라인을 결정하는 리듀스 연산 단계를 포함할 수 있다.
- [0026] 실시예들 중에서, 컴퓨터 수행 가능한 기록매체는 그리드 기반의 스카이라인 질의처리장치에서 수행되는 스카이라인 질의처리방법을 기록하는 컴퓨터 수행 가능한 기록매체에 있어서, 적어도 하나의 속성을 포함하는 스카이라인 질의를 수신하는 과정, 상기 스카이라인 질의 상의 적어도 일부 속성을 기초로 그리드를 구축하는 과정 및 상기 생성된 그리드를 통해 맵-리듀스 연산을 수행하여 상기 스카이라인 질의를 처리하는 과정을 포함한다.

발명의 효과

- [0028] 개시된 기술은 다음의 효과를 가질 수 있다. 다만, 특정 실시예가 다음의 효과를 전부 포함하여야 한다거나 다음의 효과만을 포함하여야 한다는 의미는 아니므로, 개시된 기술의 권리범위는 이에 의하여 제한되는 것으로 이해되어서는 아니 될 것이다.
- [0029] 본 발명의 일 실시예에 따른 그리드 기반의 스카이라인 질의처리장치 및 방법은 단계 그리드를 이용하여 맵-리듀스 연산을 수행함으로써 스카이라인의 특징을 활용할 수 있다.
- [0030] 본 발명의 일 실시예에 따른 그리드 기반의 스카이라인 질의처리장치 및 방법은 2단계 그리드를 이용하면서도 1번의 맵-리듀스 연산만을 사용함으로써 디스크 I/O를 줄일 수 있다.

도면의 간단한 설명

- [0032] 도 1은 본 발명의 일 실시예에 따른 그리드 기반의 스카이라인 질의처리 시스템을 설명하는 도면이다.
- 도 2는 도 1에 있는 스카이라인 질의처리장치를 설명하는 블록도이다.
- 도 3은 도 1에 있는 그리드 구축부를 설명하는 블록도이다.
- 도 4는 도 1에 있는 스카이라인 질의처리장치에서 스카이라인 질의를 처리하는 과정을 설명하는 순서도이다.
- 도 5는 도 2에 있는 스카이라인 질의 처리부에서 맵-리듀스 연산을 수행하는 과정을 설명하는 순서도이다.
- 도 6는 스카이라인을 설명하는 예시도이다.
- 도 7은 도 2에 있는 그리드 구축부에서 구축하는 그리드의 2단계 구조를 설명하는 예시도이다.
- 도 8은 도 2에 있는 스카이라인 질의 처리부에서 수행되는 맵-리듀스 연산 중 각도 기반의 분할을 통해 단위 블

록 그룹을 생성하는 과정을 설명하는 예시도이다.

도 9은 도 1에 있는 스카이라인 질의처리장치에서 수행되는 스카이라인 질의처리 과정을 설명하는 예시도이다.

발명을 실시하기 위한 구체적인 내용

- [0033] 본 발명에 관한 설명은 구조적 내지 기능적 설명을 위한 실시예에 불과하므로, 본 발명의 권리범위는 본문에 설명된 실시예에 의하여 제한되는 것으로 해석되어서는 아니 된다. 즉, 실시예는 다양한 변경이 가능하고 여러 가지 형태를 가질 수 있으므로 본 발명의 권리범위는 기술적 사상을 실현할 수 있는 균등물들을 포함하는 것으로 이해되어야 한다. 또한, 본 발명에서 제시된 목적 또는 효과는 특정 실시예가 이를 전부 포함하여야 한다거나 그러한 효과만을 포함하여야 한다는 의미는 아니므로, 본 발명의 권리범위는 이에 의하여 제한되는 것으로 이해되어서는 아니 될 것이다.
- [0034] 한편, 본 출원에서 서술되는 용어의 의미는 다음과 같이 이해되어야 할 것이다.
- [0035] "제1", "제2" 등의 용어는 하나의 구성요소를 다른 구성요소로부터 구별하기 위한 것으로, 이들 용어들에 의해 권리범위가 한정되어서는 아니 된다. 예를 들어, 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소도 제1 구성요소로 명명될 수 있다.
- [0036] 어떤 구성요소가 다른 구성요소에 "연결되어" 있다고 언급된 때에는, 그 다른 구성요소에 직접적으로 연결될 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성요소가 다른 구성요소에 "직접 연결되어" 있다고 언급된 때에는 중간에 다른 구성요소가 존재하지 않는 것으로 이해되어야 할 것이다. 한편, 구성요소들 간의 관계를 설명하는 다른 표현들, 즉 "~사이에"와 "바로 ~사이에" 또는 "~에 이웃하는"과 "~에 직접 이웃하는" 등도 마찬가지로 해석되어야 한다.
- [0037] 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한 복수의 표현을 포함하는 것으로 이해되어야 하고, "포함하다" 또는 "가지다" 등의 용어는 실시된 특징, 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것이 존재함을 지정하려는 것이며, 하나 또는 그 이상의 다른 특징이나 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.
- [0038] 각 단계들에 있어 식별부호(예를 들어, a, b, c 등)는 설명의 편의를 위하여 사용되는 것으로 식별부호는 각 단계들의 순서를 설명하는 것이 아니며, 각 단계들은 문맥상 명백하게 특정 순서를 기재하지 않는 이상 명기된 순서와 다르게 일어날 수 있다. 즉, 각 단계들은 명기된 순서와 동일하게 일어날 수도 있고 실질적으로 동시에 수행될 수도 있으며 반대의 순서대로 수행될 수도 있다.
- [0039] 본 발명은 컴퓨터가 읽을 수 있는 기록매체에 컴퓨터가 읽을 수 있는 코드로서 구현될 수 있고, 컴퓨터가 읽을 수 있는 기록 매체는 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록 장치를 포함한다. 컴퓨터가 읽을 수 있는 기록 매체의 예로는 ROM, RAM, CD-ROM, 자기 테이프, 플로피 디스크, 광 데이터 저장 장치 등이 있다. 또한, 컴퓨터가 읽을 수 있는 기록 매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어, 분산 방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수 있다.
- [0040] 여기서 사용되는 모든 용어들은 다르게 정의되지 않는 한, 본 발명이 속하는 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가진다. 일반적으로 사용되는 사전에 정의되어 있는 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 것으로 해석되어야 하며, 본 출원에서 명백하게 정의하지 않는 한 이상적이거나 과도하게 형식적인 의미를 지니는 것으로 해석될 수 없다.
- [0042] 도 6은 스카이라인을 설명하는 예시도이다.
- [0043] 도 6을 참조하면, 스카이라인 질의는 도 6에 있는 그림과 같은 형태의 질의에 해당할 수 있다. 두 개의 데이터 A, B가 주어졌을 때, A의 모든 속성이 B보다 나쁘지 않으며(도 6에선 원점에서 멀지 않으며), A의 적어도 하나의 속성이 B보다 좋다면(도 6에선 A가 B보다 원점에 가깝다면) A는 B를 지배할 수 있다. 예를 들어, 고객이 고려하고 있는 속성인 '가격'과 '해변과의 거리' 모두에 대해 $p_2(630)$ 보다 우위에 있는 $p_1(610)$ 은 $p_2(630)$ 를 지배한다.
- [0044] 그리고 어떤 데이터에도 지배되지 않는 데이터들의 집합을 스카이라인이라 한다. 예를 들어, $p_5(650)$ 와 $p_7(670)$ 처럼 서로 지배하지 않는 경우는 모두 스카이라인에 해당한다. 고객이 '가격'을 더 중시할 지, '해변과의 거

리'를 중시할 지 알 수 없기 때문이다. 이와 같이 스카이라인 질의는 고객이 고려하는 다양한 속성에 대해 적어도 고객이 원하지 않을 만한 데이터를 사전에 처리해 주기 때문에 추천 시스템 및 의사 결정 문제에 활용될 수 있다.

- [0046] 도 1은 본 발명의 일 실시예에 따른 그리드 기반의 스카이라인 질의처리 시스템을 설명하는 도면이다.
- [0047] 도 1을 참조하면, 그리드 기반의 스카이라인 질의처리 시스템(이하, 스카이라인 질의처리 시스템이라 한다.)(100)은 사용자 단말(110), 스카이라인 질의처리장치(130) 및 데이터베이스(150)를 포함할 수 있다.
- [0048] 사용자 단말(110)은 스카이라인 질의를 입력하고 관련 응답을 확인할 수 있는 컴퓨팅 장치에 해당할 수 있고, 스마트폰, 노트북 또는 컴퓨터로 구현될 수 있으며, 반드시 이에 한정되지 않고, 태블릿 PC 등 다양한 디바이스로도 구현될 수 있다. 사용자 단말(110)은 스카이라인 질의처리장치(130)와 네트워크를 통해 연결될 수 있고, 사용자 단말(110a) 내지 사용자 단말(110c)을 포함하는 복수의 사용자 단말(110)은 스카이라인 질의처리장치(130)와 동시에 연결될 수 있다.
- [0049] 스카이라인 질의처리장치(130)는 스카이라인 질의를 수신하고 그리드 기반의 맵-리듀스 연산을 통해 스카이라인 질의를 처리하여 사용자 단말(110)에 질의처리 결과를 제공할 수 있는 컴퓨터 또는 프로그램에 해당하는 서버로 구현될 수 있다. 스카이라인 질의처리장치(130)는 사용자 단말(110)과 블루투스, WiFi 등을 통해 무선으로 연결될 수 있고, 네트워크를 통해 사용자 단말(110)과 데이터를 주고 받을 수 있다.
- [0050] 스카이라인 질의처리장치(130)는 데이터베이스(150)를 포함하여 구현될 수 있고, 데이터베이스(150)와 독립적으로 구현될 수 있다. 데이터베이스(150)와 독립적으로 구현된 경우 스카이라인 질의처리장치(130)는 데이터베이스(150)와 유선 또는 무선으로 연결되어 데이터를 주고 받을 수 있다.
- [0051] 데이터베이스(150)는 스카이라인 질의처리를 위해 필요한 다양한 정보들을 저장하는 저장장치이다. 데이터베이스(150)는 사용자 단말(110)로부터 수신한 적어도 하나의 스카이라인 질의 또는 스카이라인 질의를 처리하여 얻은 결과를 저장할 수 있고, 반드시 이에 한정되지 않고, 스카이라인 질의처리장치(130)가 스카이라인 질의를 처리하는 과정에서 다양한 형태로 수집하거나 가공한 정보들을 저장할 수 있다.
- [0052] 데이터베이스(150)는 특정 범위에 속하는 정보들을 저장하는 적어도 하나의 독립된 서버-데이터베이스들로 구성될 수 있고, 적어도 하나의 독립된 서버-데이터베이스들이 하나로 통합된 통합 데이터베이스로 구성될 수 있다. 적어도 하나의 독립된 서버-데이터베이스들로 구성되는 경우에는 각각의 서버-데이터베이스들은 블루투스, WiFi 등을 통해 무선으로 연결될 수 있고, 네트워크를 통해 상호 간의 데이터를 주고 받을 수 있다. 데이터베이스(150)는 통합 데이터베이스로 구성되는 경우 각각의 서버-데이터베이스들을 하나로 통합하고 상호 간의 데이터 교환 및 제어 흐름을 관리하는 제어부를 포함할 수 있다.
- [0053] 도 2는 도 1에 있는 스카이라인 질의처리장치를 설명하는 블록도이다.
- [0054] 도 2를 참조하면, 스카이라인 질의처리장치(130)는 스카이라인 질의 수신부(210), 그리드 구축부(230), 스카이라인 질의 처리부(250) 및 제어부(270)를 포함할 수 있다.
- [0055] 스카이라인 질의 수신부(210)는 사용자 단말(110)로부터 적어도 하나의 속성을 포함하는 스카이라인 질의를 수신할 수 있다. 예를 들어, 도 6에서의 스카이라인 질의는 '해변에서 가깝고 가격이 싼 호텔'에 해당할 수 있다.
- [0056] 일 실시예에서, 스카이라인 질의 수신부(210)는 질의와 관련된 적어도 하나의 속성과 해당 속성에 관한 조건을 포함하는 스카이라인 질의를 수신할 수 있다. 스카이라인 질의 수신부(210)는 적어도 하나의 속성을 입력하고 해당 속성에 관한 조건을 선택할 수 있는 인터페이스를 사용자 단말(110)에 제공할 수 있다.
- [0057] 그리드 구축부(230)는 스카이라인 질의 수신부(210)를 통해 수신한 스카이라인 질의 상의 적어도 일부 속성을 기초로 그리드를 구축할 수 있다. 스카이라인 질의처리장치(130)는 그리드 구축부(230)를 통해 구축된 그리드를 기초로 스카이라인 질의처리를 위한 맵-리듀스 연산을 수행할 수 있다. 일 실시예에서, 그리드 구축부(230)는 데이터베이스(150)에 저장된 데이터 중에서 특정 개수의 데이터를 샘플링(sampling)할 수 있고, 샘플링된 데이터를 이용하여 그리드를 구축할 수 있다. 그리드 구축부(230)에 대해서는 도 3에서 보다 자세히 설명한다.
- [0058] 스카이라인 질의 처리부(250)는 그리드 구축부(230)를 통해 구축된 그리드를 기초로 맵-리듀스 연산을 수행하여 스카이라인 질의를 처리할 수 있다. 일 실시예에서, 스카이라인 질의 처리부(250)는 스카이라인 질의를 처리한 결과를 사용자 단말(110)에 제공할 수 있다. 스카이라인 질의 처리부(250)에서 수행되는 맵-리듀스 연산에 대해

서는 도 5에서 보다 자세히 설명한다.

- [0059] 일 실시예에서, 스카이라인 질의 처리부(250)는 스카이라인 질의처리 결과를 2차원 좌표 상의 위치로서 제공하거나 또는 3차원 공간 상의 위치로서 제공할 수 있다. 또한, 스카이라인 질의 처리부(250)는 2차원 좌표 또는 3차원 공간에 대응하는 지도 또는 지형 위에 표시된 형태로 제공할 수 있다.
- [0060] 제어부(270)는 스카이라인 질의처리장치(130)의 전체적인 동작을 제어하고, 스카이라인 질의 수신부(210), 그리드 구축부(230) 및 스카이라인 질의 처리부(250) 간의 제어 흐름 또는 데이터 흐름을 관리할 수 있다.
- [0062] 도 3은 도 1에 있는 그리드 구축부를 설명하는 블록도이다.
- [0063] 도 3을 참조하면, 그리드 구축부(230)는 그리드 생성 모듈(310), 그리드 갱신 모듈(330) 및 제어 모듈(350)를 포함할 수 있다.
- [0064] 그리드 생성 모듈(310)은 스카이라인 질의 상의 적어도 일부 속성을 원점에서 멀어질수록 증가하는 단위를 가지고 데이터를 포함할 수 있는 단위 블록들로 차원화 하여 그리드를 생성할 수 있다. 여기에서, 속성에 대한 차원화는 각 속성을 독립적인 축에 대응시키는 것에 해당할 수 있다. 보다 구체적으로, 그리드 생성 모듈(310)은 스카이라인 질의 상의 적어도 일부 속성을 각 속성별로 독립된 축에 대응시키고 원점에서 증가하는 단위를 가지는 속성 값을 기초로 배치된 데이터들을 포함할 수 있는 단위 블록들로 구성된 그리드를 생성할 수 있다.
- [0065] 일 실시예에서, 그리드 생성 모듈(310)은 원점에서 멀어질수록 각 차원의 분할점이 등차수열 형태로 증가하도록 차원화 할 수 있다. 그리드 생성 모듈(310)은 각 차원의 분할점을 등차수열 형태로 증가하도록 차원화 함으로써 각 단위 블록의 크기가 상이하게 할 수 있다. 각 단위 블록의 크기는 상이하더라도 특정 값의 등차를 가지는 등차수열 형태로 증가하므로 그리드 생성 모듈(310)은 각 단위 블록을 사전에 정해진 크기로 생성할 수 있다.
- [0066] 그리드 갱신 모듈(330)은 그리드 생성 모듈(310)에 의해 생성된 그리드 상에서 단위 블록들 중 적어도 지배당하지 않고 해당 데이터를 포함하는 단위 블록을 결정하여 생성된 그리드를 갱신할 수 있다. 그리드 갱신 모듈(330)은 그리드 생성 모듈(310)에 의해 생성된 그리드의 하위 그리드로서 갱신된 그리드를 연결할 수 있다.
- [0067] 일 실시예에서, 그리드 갱신 모듈(330)은 샘플링된 데이터를 기초로 생성된 그리드를 갱신할 수 있다. 보다 구체적으로, 그리드 갱신 모듈(330)은 그리드 구축부(230)에 의해 샘플링된 데이터 및 그리드 생성 모듈(310)에 의해 생성된 그리드를 기초로 단위 블록들 중 적어도 지배당하지 않고 샘플링된 데이터 중 일부를 포함하는 단위 블록을 결정하여 생성된 그리드를 갱신할 수 있다.
- [0068] 일 실시예에서, 그리드 갱신 모듈(330)은 그리드 생성 모듈(310)에 의해 생성된 그리드 상에서 단위 블록들 중 적어도 지배당하지 않고 해당 데이터를 반드시 포함하는 단위 블록 각각이 적어도 하나의 서브-단위 블록을 포함하도록 갱신할 수 있다. 그리드 생성 모듈(310)은 원점에서 멀어질수록 증가하는 단위를 가지도록 차원화 하여 그리드를 생성하였기 때문에 지배당한 단위 블록의 서브-단위 블록의 수는 해당 단위 블록을 지배하는 단위 블록의 서브-단위 블록의 수보다 많을 수 있다.
- [0069] 제어 모듈(350)은 그리드 구축부(230)의 전체적인 동작을 제어하고, 그리드 생성 모듈(310) 및 그리드 갱신 모듈(330) 간의 제어 흐름 또는 데이터 흐름을 관리할 수 있다.
- [0071] 도 4는 도 1에 있는 스카이라인 질의처리장치에서 스카이라인 질의를 처리하는 과정을 설명하는 순서도이다.
- [0072] 도 4를 참조하면, 스카이라인 질의처리장치(130)는 스카이라인 질의 수신부(210)를 통해 적어도 하나의 속성을 포함하는 스카이라인 질의를 수신할 수 있다(단계 S410). 스카이라인 질의처리장치(130)는 그리드 구축부(230)를 통해 스카이라인 질의 상의 적어도 일부 속성을 기초로 그리드를 구축할 수 있다(단계 S430). 스카이라인 질의처리장치(130)는 스카이라인 질의 처리부(250)를 통해 구축된 그리드를 통해 맵-리듀스 연산을 수행하여 스카이라인 질의를 처리할 수 있다(단계 S450).
- [0074] 도 5는 도 2에 있는 스카이라인 질의 처리부에서 맵-리듀스 연산을 수행하는 과정을 설명하는 순서도이다.
- [0075] 도 5를 참조하면, 스카이라인 질의 처리부(250)는 맵(map) 연산 단계에서 갱신된 그리드에 대해 비스카이라인(non-skyline) 데이터를 프루닝(pruning)한 후 컴바이너(combiner) 연산 단계로 전달할 수 있다(단계 S510).

스카이라인 질의 처리부(250)는 컴바이너 연산 단계에서 그리드 상의 각 단위 블록을 행 우선 순서로 검색하면서 지역적 스카이라인을 결정한 후 리듀스(reduce) 연산 단계로 전달할 수 있다(단계 S530). 스카이라인 질의 처리부(250)는 리듀스 연산 단계에서 그리드 상의 각 단위 블록을 행 우선 순서로 검색하면서 전역적 스카이라인을 결정할 수 있다(단계 S550).

- [0076] 스카이라인 질의처리장치(130)는 사용자 단말(110)로부터 수신한 스카이라인 질의를 기초로 데이터베이스(150)에 저장된 데이터들을 이용하여 그리드 구축부(230)를 통해 2단계 그리드를 구축할 수 있다. 이 때, 그리드 구축부(230)에 의해 구축된 그리드는 각도 기반 분할에 의해 R개의 그룹으로 분할될 수 있다. 스카이라인 질의처리장치(130)는 스카이라인 질의 처리부(250)를 통해 스카이라인 질의처리를 위한 맵-리듀스 연산을 수행할 수 있다. 맵-리듀스 연산은 크게 맵(map) 연산, 컴바이너(combiner) 연산 및 리듀스(reduce) 연산으로 구분될 수 있다.
- [0077] 맵-리듀스 연산 중 맵(map) 연산 단계에서는 그리드 구축부(230)의 그리드 생성 모듈(310)에 의해 생성된 그리드에 대해 스카이라인이 될 수 없는 데이터를 프루닝(pruning)한 후 컴바이너 연산 단계로 전달할 수 있다.
- [0078] 맵-리듀스 연산 중 컴바이너(combiner) 연산 단계에서는 동일 머신에서 처리된 각 맵 연산 결과를 입력으로 받을 수 있고, 해당 입력을 이용하여 지역적인 스카이라인을 계산한 후 리듀스 연산 단계로 전달할 수 있다. 보다 구체적으로, 그리드 구축부(230)의 그리드 생성 모듈(310)에 의해 생성된 그리드 상의 각 단위 블록을 행 우선 순서로 검색(예를 들어, 도 7을 기준으로 $b^1, b^2, \dots, b^9$ 순서로 검색)하면서 각 단위 블록에 대한 스카이라인을 결정할 수 있다. 결정된 스카이라인은 이후 각 차원별로 아직까지 방문하지 않은 인접한 단위 블록으로 전달되어 필터(filter) 데이터로 사용될 수 있다.
- [0079] 그리드 구축부(230)의 그리드 생성 모듈(310)에 의해 생성된 그리드의 특정 단위 블록에 대한 필터 데이터는 적어도 하나의 차원에 대해선 해당 단위 블록의 데이터보다 좋은 값을 가지기 때문에 해당 차원에 대해선 스카이라인 질의를 위한 비교를 수행하지 않을 수 있다.
- [0080] 그리드 구축부(230)의 그리드 생성 모듈(310)에 의해 생성된 그리드의 모든 단위 블록에 대한 방문이 끝나면 지역적인 스카이라인이 결정될 수 있다. 이후 모든 데이터는 각 그룹을 담당하는 리듀스에는 후보 데이터로, 그 외의 그룹을 담당하는 리듀스에는 필터 데이터로 전달될 수 있다. 이 때, 타 그룹의 지배되지 않은 단위 블록에 영향을 줄 수 없는 필터 데이터는 전달하지 않을 수 있다.
- [0081] 맵-리듀스 연산 중 리듀스(reduce) 연산 단계에서는 그리드 구축부(230)에 의해 구축된 그리드 상의 각 단위 블록을 행 우선 순서로 검색하면서 전역적 스카이라인을 결정할 수 있다. 이 때, 각 머신에서 수행되는 리듀스는 각 단위 블록에 대한 전역적 스카이라인을 결정하는 과정에서 입력된 후보 데이터만이 아니라 필터 데이터까지 고려할 수 있다. 결정된 스카이라인은 결과로서 출력될 수 있고, 해당 스카이라인은 아직까지 방문하지 않은 인접한 단위 블록에 필터 데이터로서 전달될 수 있다.
- [0083] 도 7은 도 2에 있는 그리드 구축부에서 구축하는 그리드의 2단계 구조를 설명하는 예시도이다.
- [0084] 도 7을 참조하면, 스카이라인 질의처리장치(130)는 그리드 구축부(230)를 통해 첫 번째 그리드(730)와 두 번째 그리드(750)로 구성된 2단계 구조를 갖는 그리드(710)를 구축할 수 있다. 그리드 구축부(230)는 그리드 생성 모듈(310)을 통해 첫 번째 그리드(730)를 생성할 수 있다. 그리드 구축부(230)는 그리드 갱신 모듈(330)을 통해 첫 번째 그리드(730)로부터 두 번째 그리드(750)를 생성할 수 있고, 두 번째 그리드(750)를 첫 번째 그리드(730)에 반영하여 첫 번째 그리드(730)를 갱신할 수 있다.
- [0085] 일 실시예에서, 그리드 구축부(230)는 첫 번째 그리드(730)와 같이 원점에서 멀어질수록 각 차원의 분할점이 등차 1의 등차수열 형태로 증가하도록 차원화 할 수 있다. 그리드 생성 모듈(310)에 의해 생성된 첫 번째 그리드(730)는 각 단위 블록의 크기가 상이하지만, 각각의 크기는 사전에 정해진 형태에 해당할 수 있다. 그리드 갱신 모듈(330)은 원점에 가까운 데이터가 없는 경우를 대비하기 위하여 두 번째 그리드(750)를 생성하여 첫 번째 그리드(730)에 반영할 수 있다. 그리드 갱신 모듈(330)은 첫 번째 그리드(730)에 포함된 단위 블록 중 데이터를 보유하고, 다른 단위 블록에 지배되지 않은 단위 블록들로 구성된 두 번째 그리드(750)를 생성할 수 있다.
- [0086] 두 번째 그리드(750)는 모든 영역에 대한 단위 블록들을 전부 포함할 필요가 없다. 예를 들어, (a) 그리드(710)에서 데이터 p_4 를 보유한 b^9 단위 블록은 b^9 가 b_5 에 의해 지배되므로, 두 번째 그리드(750)에 포함되지

않는다. 또한, 그리드에 포함된 모든 단위 블록의 크기를 사전에 알 수 있으므로, 등차수열을 이용한 간단한 계산으로 특정 데이터가 속한 단위 블록을 트리 탐색 등의 추가적인 비용없이 구할 수 있다.

[0088] 도 8은 도 2에 있는 스카이라인 질의 처리부에서 수행되는 맵-리듀스 연산 중 각도 기반의 분할을 통해 단위 블록 그룹을 생성하는 과정을 설명하는 예시도이다.

[0089] 도 8을 참조하면, 스카이라인 질의처리장치(130)는 스카이라인 질의 처리부(250)를 통해 맵-리듀스 연산을 수행할 수 있고, 맵-리듀스 연산 중 그리드 구축부(230)의 그리드 생성 모듈(310)에 의해 생성된 첫 번째 그리드에 대하여 모든 단위 블록을 각도 기반 분할(angular partitioning approach)을 통해 R(도 8에서 R은 3)개의 그룹으로 분할할 수 있다. 각도 기반 분할을 통해 R개로 분할된 각 그룹은 리듀스 연산 단계에서 각각의 리듀스 연산에 의해 처리될 수 있다.

[0091] 도 9은 도 1에 있는 스카이라인 질의처리장치에서 수행되는 스카이라인 질의처리 과정을 설명하는 예시도이다.

[0092] 도 9를 참조하면, M1, M2 및 M3는 각 머신에서 수행되는 맵 연산(910)을, C1, C2 및 C3는 컴바이너 연산(930)을, R1, R2 및 R3는 리듀스 연산(950)에 해당할 수 있다. 스카이라인 질의 처리부(250)는 그리드 구축부(230)에 의해 구축된 그리드를 각도 기반 분할을 통해 R개의 그룹으로 분할할 수 있고, 모든 맵 연산(910)으로 전달할 수 있다.

[0093] M1, M2 및 M3는 사각형, 삼각형 및 원으로 표시된 포인트들을 입력받을 수 있다. M1은 각 사각형 포인트와 함께 독립적으로 호출될 수 있다. M1은 사각형 포인트가 속한 단위 블록과 서브-단위 블록을 얻을 수 있고, 해당 단위 블록과 서브-단위 블록이 프루닝 되었는지 체크할 수 있다. 만약 프루닝 된 경우에는 해당 사각형 포인트를 제거할 수 있다. 따라서, 색으로 칠해진 단위 블록과 서브-단위 블록에 있는 포인트들은 제거될 수 있다. 스카이라인 질의 처리부(250)는 맵 연산(910) 단계에서 모든 맵 연산(910)에 대해 상기와 같은 연산을 동일하게 수행할 수 있다. 스카이라인 질의 처리부(250)는 각 맵 연산(910)에서 제거되지 않은 포인트들을 컴바이너 연산(930)으로 전달할 수 있다.

[0094] C1은 지역적 스카이라인 포인트를 각 단위 블록에 대해 산출할 수 있다. C1은 지역적 스카이라인 포인트를 동일 그룹에 속한 단위 블록에게 전송할 수 있다. 전송된 포인트들은 리듀스 연산(950)에서 후보 포인트가 될 수 있다. C1은 지역적 스카이라인 포인트 중에서 필터 포인트를 결정할 수 있고, 결정된 필터 포인트를 모든 다른 그룹으로 전송할 수 있다. R2에서 왼쪽 및 아래쪽 필터 포인트들은 각각 C1과 C3로부터 전송받은 필터 포인트들이다. 스카이라인 질의 처리부(250)는 모든 컴바이너 연산(930)에 대해 상기와 같은 연산을 수행할 수 있고, 필터 포인트와 후보 포인트는 리듀스 연산(950)인 R1 내지 R3의 입력으로 전달할 수 있다.

[0095] 필터 포인트와 후보 포인트는 리듀스 연산(950)의 입력으로 사용될 수 있다. 필터 포인트는 스카이라인이 아닌 포인트들을 제거하는데 사용될 수 있다. 도 9에서 필터 포인트로부터 시작된 라인들은 필터링 프로세스에 해당할 수 있다. 스카이라인 질의 처리부(250)는 마지막으로 남아 있는 포인트들 중에서 전역적 스카이라인 포인트들을 산출할 수 있다. R1 내지 R3와 같은 모든 리듀스 연산(950)들은 독립적으로 호출될 수 있고 스카이라인 질의 처리부(250)는 각 리듀스 연산(950)으로부터 산출된 전역적 스카이라인 포인트들을 병합하여 사용자 단말(110)에 최종 결과로서 제공할 수 있다.

[0097] 상기에서는 본 발명의 바람직한 실시예를 참조하여 설명하였지만, 해당 기술 분야의 숙련된 당업자는 하기의 특허 청구의 범위에 기재된 본 발명의 사상 및 영역으로부터 벗어나지 않는 범위 내에서 본 발명을 다양하게 수정 및 변경시킬 수 있음을 이해할 수 있을 것이다.

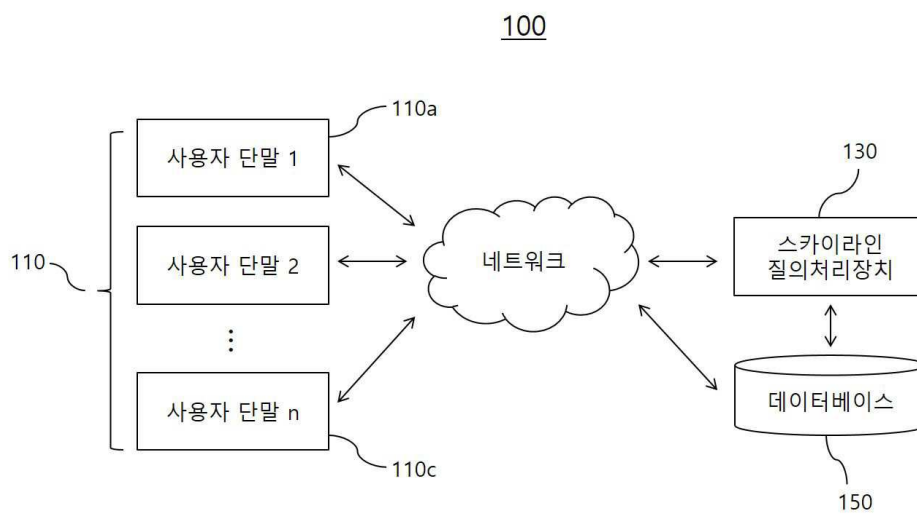
부호의 설명

[0099] 100: 그리드 기반의 스카이라인 질의처리 시스템
110: 사용자 단말 130: 스카이라인 질의처리장치

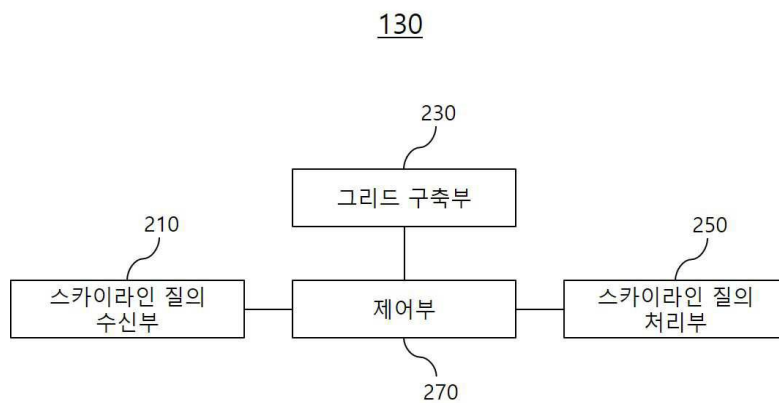
150: 데이터베이스
 210: 스카이라인 질의 수신부 230: 그리드 구축부
 250: 스카이라인 질의 처리부 270: 제어부
 310: 그리드 생성 모듈 330: 그리드 갱신 모듈
 350: 제어 모듈
 910: 맵 연산 930: 컴바이너 연산
 950: 리듀스 연산

도면

도면1



도면2

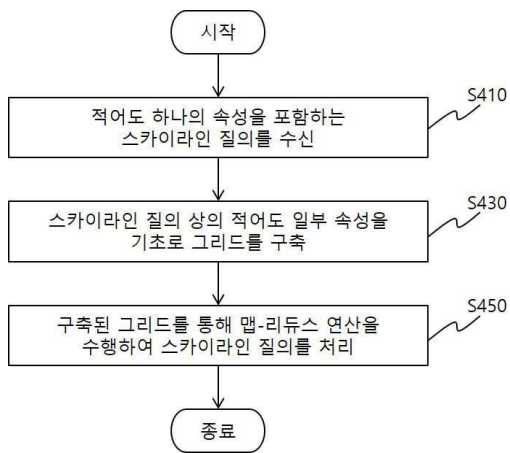


도면3

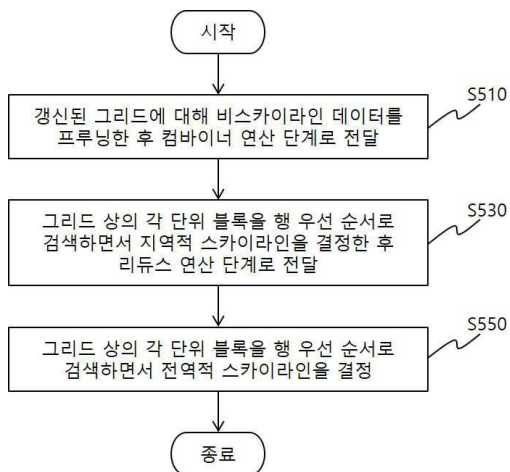
230



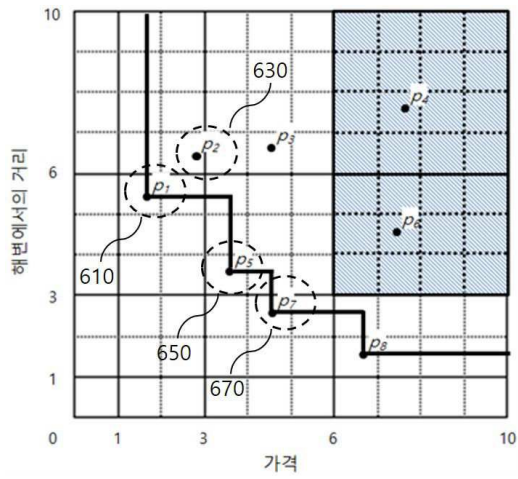
도면4



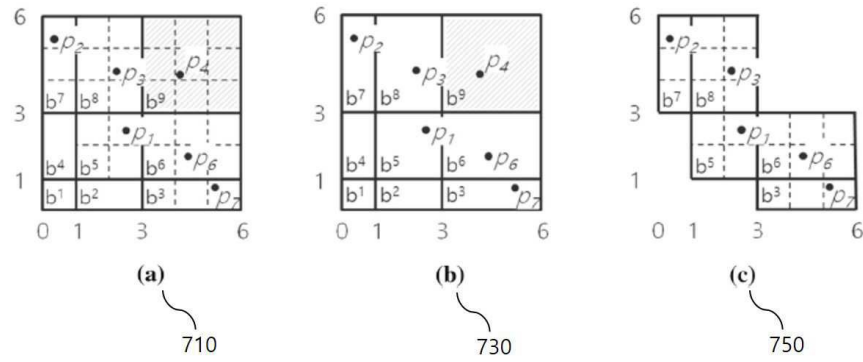
도면5



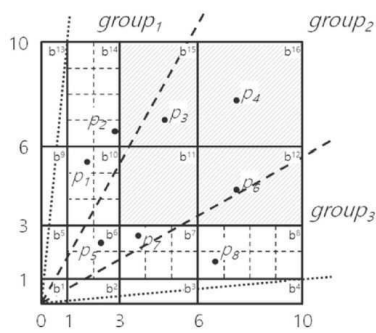
도면6



도면7



도면8



도면9

