



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2019-0079354
(43) 공개일자 2019년07월05일

(51) 국제특허분류(Int. Cl.)

G06F 16/00 (2019.01)

(52) CPC특허분류

G06F 16/245 (2019.01)

G06F 16/2246 (2019.01)

(21) 출원번호 10-2017-0181486

(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)

조범준

서울특별시 마포구 마포대로 53, A동 807호(도화동, 마포트라펠리스)

(74) 대리인

정부연

전체 청구항 수 : 총 16 항

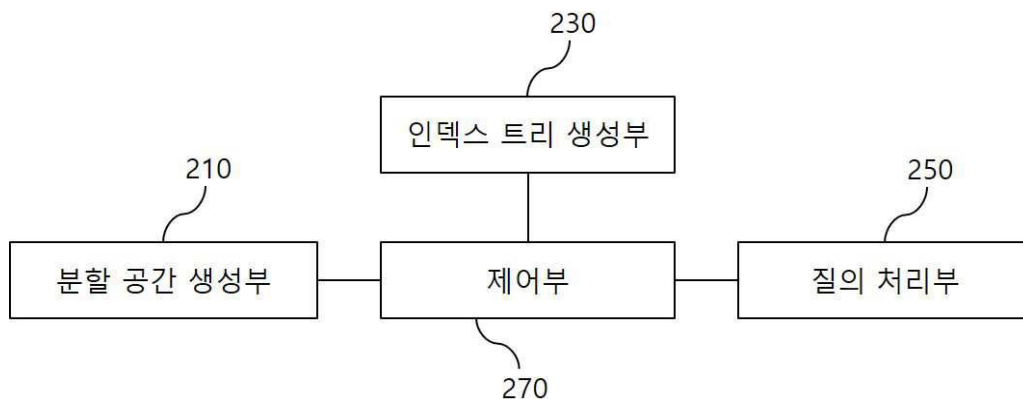
(54) 발명의 명칭 분할 공간 기반의 공간 데이터 객체 질의처리장치 및 방법, 이를 기록한 기록매체

(57) 요약

본 발명은 분할 공간 기반의 공간 데이터 객체 질의처리장치 및 방법에 관한 것으로, 적어도 하나의 공간 데이터 객체를 포함하는 데이터 공간을 분할하여 복수의 분할 공간들을 생성하는 분할 공간 생성부 및 상기 복수의 분할 공간에 관한 정보 및 해당 분할 공간에 포함된 공간 데이터 객체를 모두 포함하는 최소 경계 사각형(Minimum Boundary Rectangle, MBR) 정보를 기초로 인덱스 트리를 생성하는 인덱스 트리 생성부를 포함한다. 따라서, 본 발명은 데이터 공간을 효율적으로 분할하고 적절한 테이블 스키마를 사용하여 효율적인 질의 처리를 할 수 있다.

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

적어도 하나의 공간 데이터 객체를 포함하는 데이터 공간을 분할하여 복수의 분할 공간들을 생성하는 분할 공간 생성부; 및

상기 복수의 분할 공간에 관한 정보 및 해당 분할 공간에 포함된 공간 데이터 객체를 모두 포함하는 최소 경계 사각형(Minimum Boundary Rectangle, MBR) 정보를 기초로 인덱스 트리를 생성하는 인덱스 트리 생성부를 포함하는 분할 공간 기반의 공간 데이터 객체 질의처리장치.

청구항 2

제1항에 있어서, 상기 분할 공간 생성부는

상기 복수의 분할 공간들 각각에 관해 해당 공간 데이터 객체의 밀집도가 특정 기준 이하가 될 때까지 해당 분할 공간을 해당 분할 공간의 중심점을 기준으로 재귀적으로 재분할 하는 것을 특징으로 하는 분할 공간 기반의 공간 데이터 객체 질의처리장치.

청구항 3

제1항에 있어서, 상기 인덱스 트리 생성부는

상기 분할 공간에 관한 정보를 기초로 생성된 키(Key) 값 및 상기 최소 경계 사각형 정보를 기초로 생성된 밸류(Value) 값을 포함하는 행 데이터로 구성된 테이블을 기초로 상기 인덱스 트리를 생성하는 것을 특징으로 하는 분할 공간 기반의 공간 데이터 객체 질의처리장치.

청구항 4

제3항에 있어서, 상기 인덱스 트리 생성부는

상기 분할 공간이 생성될 때마다 각각의 축에 대해 원점에서 가까운 방향의 분할 공간을 0, 먼 방향의 분할 공간을 1로 표시하고 각 축에 대한 비트(bit)를 연결하여 상기 키 값을 생성하는 것을 특징으로 하는 분할 공간 기반의 공간 데이터 객체 질의처리장치.

청구항 5

제4항에 있어서, 상기 인덱스 트리 생성부는

상기 분할 공간에 대한 재귀적인 재분할이 발생하는 경우 이전 분할 정보 및 현재 발생한 재분할 공간에 대한 정보를 연결하는 방식으로 표시하는 것을 특징으로 하는 분할 공간 기반의 공간 데이터 객체 질의처리장치.

청구항 6

제3항에 있어서, 상기 인덱스 트리 생성부는

내부 노드에 관한 상기 행 데이터를 저장하는 인덱스 테이블 및 리프 노드에 관한 상기 행 데이터를 저장하는 데이터 테이블을 생성함으로써 상기 인덱스 트리를 생성하는 것을 특징으로 하는 분할 공간 기반의 공간 데이터

객체 질의처리장치.

청구항 7

제1항에 있어서,

상기 인덱스 트리를 이용하여 공간 데이터 객체 질의를 처리하는 질의 처리부를 더 포함하는 것을 특징으로 하는 분할 공간 기반의 공간 데이터 객체 질의처리장치.

청구항 8

제7항에 있어서, 상기 질의 처리부는

상기 인덱스 트리를 이용하여 질의 포인트와 질의 반경을 포함하는 범위 질의 또는 질의 포인트와 최근접 이웃 수를 포함하는 kNN 질의를 처리하는 것을 특징으로 하는 분할 공간 기반의 공간 데이터 객체 질의처리장치.

청구항 9

분할 공간 기반의 공간 데이터 객체 질의처리장치에서 수행되는 공간 데이터 객체 질의처리방법에 있어서,

(a) 적어도 하나의 공간 데이터 객체를 포함하는 데이터 공간을 분할하여 복수의 분할 공간들을 생성하는 단계; 및

(b) 상기 복수의 분할 공간에 관한 정보 및 해당 분할 공간에 포함된 공간 데이터 객체를 모두 포함하는 최소 경계 사각형(Minimum Boundary Rectangle, MBR) 정보를 기초로 인덱스 트리를 생성하는 단계를 포함하는 분할 공간 기반의 공간 데이터 객체 질의처리방법.

청구항 10

제9항에 있어서, 상기 (b) 단계는

상기 분할 공간에 관한 정보를 기초로 생성된 키(Key) 값 및 상기 최소 경계 사각형 정보를 기초로 생성된 벨류(Value) 값을 포함하는 행 데이터로 구성된 테이블을 기초로 상기 인덱스 트리를 생성하는 단계인 것을 특징으로 하는 분할 공간 기반의 공간 데이터 객체 질의처리방법.

청구항 11

제10항에 있어서, 상기 (b) 단계는

상기 분할 공간이 생성될 때마다 각각의 축에 대해 원점에서 가까운 방향의 분할 공간을 0, 먼 방향의 분할 공간을 1로 표시하고 각 축에 대한 비트(bit)를 연결하여 상기 키 값을 생성하는 단계인 것을 특징으로 하는 분할 공간 기반의 공간 데이터 객체 질의처리방법.

청구항 12

제11항에 있어서, 상기 (b) 단계는

상기 분할 공간에 대한 재귀적인 재분할이 발생하는 경우 이전 분할 정보 및 현재 발생한 재분할 공간에 대한 정보를 연결하는 방식으로 표시하는 단계인 것을 특징으로 하는 분할 공간 기반의 공간 데이터 객체 질의처리방법.

청구항 13

제10항에 있어서, 상기 (b) 단계는

내부 노드에 관한 상기 행 데이터를 저장하는 인덱스 테이블 및 리프 노드에 관한 상기 행 데이터를 저장하는 데이터 테이블을 생성함으로써 상기 인덱스 트리를 생성하는 단계인 것을 특징으로 하는 분할 공간 기반의 공간 데이터 객체 질의처리방법.

청구항 14

제9항에 있어서,

(c) 상기 인덱스 트리를 이용하여 공간 데이터 객체 질의를 처리하는 단계를 더 포함하는 것을 특징으로 하는 분할 공간 기반의 공간 데이터 객체 질의처리방법.

청구항 15

제14항에 있어서, 상기 (c) 단계는

상기 인덱스 트리를 이용하여 질의 포인트와 질의 반경을 포함하는 범위 질의 또는 질의 포인트와 최근접 이웃수를 포함하는 kNN 질의를 처리하는 단계인 것을 특징으로 하는 분할 공간 기반의 공간 데이터 객체 질의처리방법.

청구항 16

분할 공간 기반의 공간 데이터 객체 질의처리장치에서 수행되는 공간 데이터 객체 질의처리방법을 기록하는 컴퓨터 수행 가능한 기록매체에 있어서,

적어도 하나의 공간 데이터 객체를 포함하는 데이터 공간을 분할하여 복수의 분할 공간들을 생성하는 과정; 및
상기 복수의 분할 공간에 관한 정보 및 해당 분할 공간에 포함된 공간 데이터 객체를 모두 포함하는 최소 경계사각형(Minimum Boundary Rectangle, MBR) 정보를 기초로 인덱스 트리를 생성하는 과정을 포함하는 기록매체.

발명의 설명

기술 분야

[0001] 본 발명은 공간 데이터 객체 질의처리 기술에 관한 것으로, 보다 상세하게는 데이터 공간을 효율적으로 분할하고 적절한 테이블 스키마를 사용하여 효율적인 질의 처리를 할 수 있는 분할 공간 기반의 공간 데이터 객체 질의처리장치 및 방법에 관한 것이다.

배경 기술

[0003] 공간 데이터와 같은 다차원 데이터에 대한 효과적인 유사 질의어(similarity query) 처리의 핵심은 공간적으로 인접한 오브젝트들을 물리적으로도 가까운 위치에 저장하고, 검색시 필요한 영역에 속한 데이터만 접근하는 것이다. 그러나, 하둡(Hadoop)과 같이 일반적으로 빅데이터를 다루기 위해 사용되는 클라우드 컴퓨팅 기반의 프레임워크들은 오브젝트들 간의 공간적 인접성을 고려하지 않고 있기 때문에 질의 처리시 대량의 긍정 오류(false positive)를 포함한 데이터를 검색하게 된다.

[0004] 공간 데이터 객체 질의는 범위 질의(Range Query) 및 kNN 질의(k-Nearest Neighbor Query)를 포함할 수 있고, HBase와 같은 NoSQL 데이터베이스 관리 시스템을 이용한 인덱싱 방법은 분산 파일 시스템에서 빠른 데이터 임의 접근과 효율적인 데이터 갱신을 위한 효과적인 프레임워크를 제공할 수 있다. 범위 질의는 질의 포인트 q와 질

의 반경 r 을 입력으로 받아 질의 포인트 q 로부터의 거리가 r 보다 작은 데이터 객체 집합을 결과로서 반환한다. kNN 질의는 질의 포인트 q 와 최근접 이웃의 수 k 를 입력으로 받아 질의 포인트 q 로부터 가장 근접한 k 개의 데이터 객체 집합을 결과로서 반환한다.

- [0005] HBase에서의 색인 생성은 1차원 행 키(Rowkey)만 지원하도록 설계된 점이 문제이다. 따라서, 선형화(linearization) 기법은 일반적으로 HBase 시스템에서 공간 데이터를 저장하는데 사용된다. 이러한 방법은 데이터 공간을 그리드(grid) 모양의 셀(cell)로 세분화하고 z -순서를 사용하여 셀을 순차적으로 정렬한다. 공간 데이터의 행 키는 해당 셀의 z -순서 번호를 사용하여 생성될 수 있다.
- [0006] 도 4는 선형화 기법을 사용하여 범위 질의를 처리하는 과정을 설명하는 예시도이다. 도 4를 참조하면, 범위 질의(Range Query)(410)가 실행되면 질의 범위 내의 최소 및 최대 z -순서 값이 계산될 수 있다. 다음으로 행 키가 z -순서 범위 내에 있는 행을 검색하여 질의가 처리될 수 있다. 그러나, 공간적 근접성을 완벽하게 보장할 수 없기 때문에 선형화 기법의 결과에는 종종 많은 긍정 오류(False Positive)(430)가 포함될 수 있다.
- [0007] 도 5는 다차원 인덱싱 레이어를 사용하여 범위 질의를 처리하는 과정을 설명하는 예시도이다. 도 5를 참조하면, 범위 질의(510)가 실행되면 먼저 인덱스 레이어(Index Layer)(530)를 검색하고 관련 영역과 연관된 행 키가 있는 행만을 접근함으로써 긍정 오류를 줄일 수 있다. 그러나, 이러한 방법은 공간 분할의 단위가 크기 때문에 질의 처리 시 여전히 많은 양의 긍정 오류에 대한 접근이 필요할 수 있다.
- [0008] 한국공개특허 제10-2016-0004781(2016.01.13)호는 빅데이터 프레임워크를 활용한 온톨로지 질의 처리방법에 관한 것으로, 테이블에 저장되는 각 노드들에 클래스 타입을 접두어로 저장하도록 하고, 클래스(Class)와 프로퍼티(Property)들을 각각의 테이블 컬럼 패밀리로 정의하도록 스키마를 세분화하여 저장하며, HBase 기반 RDF 저장소의 개선된 테이블 스키마를 기반으로 질의를 분석하여 HBase 필터를 적용해 검색해야 하는 데이터의 양을 획기적으로 줄여 효율적으로 질의를 처리할 수 있도록 하고, 기존 질의 처리기에서 처리하지 못한 중첩 트리플 패턴 질의를 처리할 수 있다.
- [0009] 한국등록특허 제10-1117709(2012.02.10)호는 공간 분할 트리의 최소 데이터-불균등 커버를 이용한 다차원 히스토그램 방법 및 이를 실행하기 위한 프로그램이 저장된 기록매체에 관한 것으로, 주어진 공간을 다양한 크기의 공간들로 분할하여 형성한 공간 분할 트리 내 각 분할된 공간 내 데이터 객체의 불균등에 기초하여 최소 데이터-불균등 커버를 판단한 후 이에 기반하여 히스토그램의 버킷을 생성함으로써, 종래의 다차원 히스토그램 방법과 달리 데이터 객체가 균등하게 분포되지 않은 상황에서도 영역 질의의 선택도에 대한 추정값 계산의 정확성을 확보하는 효과가 있다.

선행기술문헌

특허문헌

- [0011] (특허문헌 0001) 한국공개특허 제10-2016-0004781(2016.01.13)호
(특허문헌 0002) 한국등록특허 제10-1117709(2012.02.10)호

발명의 내용

해결하려는 과제

- [0012] 본 발명의 일 실시예는 데이터 공간을 효율적으로 분할하고 적절한 테이블 스키마를 사용하여 효율적인 질의 처리를 할 수 있는 분할 공간 기반의 공간 데이터 객체 질의처리장치 및 방법을 제공하고자 한다.
- [0013] 본 발명의 일 실시예는 재귀적인 사분할을 통해 분할된 복수의 분할 공간 및 공간 데이터 객체의 최소 경계 사각형을 기초로 인덱스 트리를 생성할 수 있는 분할 공간 기반의 공간 데이터 객체 질의처리장치 및 방법을 제공하고자 한다.
- [0014] 본 발명의 일 실시예는 효율적인 인덱스 트리를 이용하여 유사 질의를 처리함으로써 긍정 오류를 효과적으로 줄일 수 있는 분할 공간 기반의 공간 데이터 객체 질의처리장치 및 방법을 제공하고자 한다.

과제의 해결 수단

- [0016] 실시예들 중에서, 분할 공간 기반의 공간 데이터 객체 질의처리장치는 적어도 하나의 공간 데이터 객체를 포함하는 데이터 공간을 분할하여 복수의 분할 공간들을 생성하는 분할 공간 생성부 및 상기 복수의 분할 공간에 관한 정보 및 해당 분할 공간에 포함된 공간 데이터 객체를 모두 포함하는 최소 경계 사각형(Minimum Boundary Rectangle, MBR) 정보를 기초로 인덱스 트리를 생성하는 인덱스 트리 생성부를 포함한다.
- [0017] 상기 분할 공간 생성부는 상기 복수의 분할 공간들 각각에 관해 해당 공간 데이터 객체의 밀집도가 특정 기준 이하가 될 때까지 해당 분할 공간을 해당 분할 공간의 중심점을 기준으로 재귀적으로 재분할 할 수 있다.
- [0018] 상기 인덱스 트리 생성부는 상기 분할 공간에 관한 정보를 기초로 생성된 키(Key) 값 및 상기 최소 경계 사각형 정보를 기초로 생성된 밸류(Value) 값을 포함하는 행 데이터로 구성된 테이블을 기초로 상기 인덱스 트리를 생성할 수 있다.
- [0019] 상기 인덱스 트리 생성부는 상기 분할 공간이 생성될 때마다 각각의 축에 대해 원점에서 가까운 방향의 분할 공간을 0, 먼 방향의 분할 공간을 1로 표시하고 각 축에 대한 비트(bit)를 연결하여 상기 키 값을 생성할 수 있다.
- [0020] 상기 인덱스 트리 생성부는 상기 분할 공간에 대한 재귀적인 재분할이 발생하는 경우 이전 분할 정보 및 현재 발생한 재분할 공간에 대한 정보를 연결하는 방식으로 표시할 수 있다.
- [0021] 상기 인덱스 트리 생성부는 내부 노드에 관한 상기 행 데이터를 저장하는 인덱스 테이블 및 리프 노드에 관한 상기 행 데이터를 저장하는 데이터 테이블을 생성함으로써 상기 인덱스 트리를 생성할 수 있다.
- [0022] 상기 분할 공간 기반의 공간 데이터 객체 질의처리장치는 상기 인덱스 트리를 이용하여 공간 데이터 객체 질의를 처리하는 질의 처리부를 더 포함할 수 있다.
- [0023] 상기 질의 처리부는 상기 인덱스 트리를 이용하여 질의 포인트와 질의 반경을 포함하는 범위 질의 또는 질의 포인트와 최근접 이웃 수를 포함하는 kNN 질의를 처리할 수 있다.
- [0024] 실시예들 중에서, 분할 공간 기반의 공간 데이터 객체 질의처리방법은 분할 공간 기반의 공간 데이터 객체 질의 처리장치에서 수행되는 공간 데이터 객체 질의처리방법에 있어서, (a) 적어도 하나의 공간 데이터 객체를 포함하는 데이터 공간을 분할하여 복수의 분할 공간들을 생성하는 단계 및 (b) 상기 복수의 분할 공간에 관한 정보 및 해당 분할 공간에 포함된 공간 데이터 객체를 모두 포함하는 최소 경계 사각형(Minimum Boundary Rectangle, MBR) 정보를 기초로 인덱스 트리를 생성하는 단계를 포함한다.
- [0025] 상기 (b) 단계는 상기 분할 공간에 관한 정보를 기초로 생성된 키(Key) 값 및 상기 최소 경계 사각형 정보를 기초로 생성된 밸류(Value) 값을 포함하는 행 데이터로 구성된 테이블을 기초로 상기 인덱스 트리를 생성하는 단계일 수 있다.
- [0026] 상기 (b) 단계는 상기 분할 공간이 생성될 때마다 각각의 축에 대해 원점에서 가까운 방향의 분할 공간을 0, 먼 방향의 분할 공간을 1로 표시하고 각 축에 대한 비트(bit)를 연결하여 상기 키 값을 생성하는 단계일 수 있다.
- [0027] 상기 (b) 단계는 상기 분할 공간에 대한 재귀적인 재분할이 발생하는 경우 이전 분할 정보 및 현재 발생한 재분할 공간에 대한 정보를 연결하는 방식으로 표시하는 단계일 수 있다.
- [0028] 상기 (b) 단계는 내부 노드에 관한 상기 행 데이터를 저장하는 인덱스 테이블 및 리프 노드에 관한 상기 행 데이터를 저장하는 데이터 테이블을 생성함으로써 상기 인덱스 트리를 생성하는 단계일 수 있다.
- [0029] 상기 분할 공간 기반의 공간 데이터 객체 질의처리방법은 (c) 상기 인덱스 트리를 이용하여 공간 데이터 객체 질의를 처리하는 단계를 더 포함할 수 있다.
- [0030] 상기 (c) 단계는 상기 인덱스 트리를 이용하여 질의 포인트와 질의 반경을 포함하는 범위 질의 또는 질의 포인트와 최근접 이웃 수를 포함하는 kNN 질의를 처리하는 단계일 수 있다.
- [0031] 실시예들 중에서, 기록매체는 분할 공간 기반의 공간 데이터 객체 질의처리장치에서 수행되는 공간 데이터 객체 질의처리방법을 기록하는 컴퓨터 수행 가능한 기록매체에 있어서, 적어도 하나의 공간 데이터 객체를 포함하는 데이터 공간을 분할하여 복수의 분할 공간들을 생성하는 과정 및 상기 복수의 분할 공간에 관한 정보 및 해당

분할 공간에 포함된 공간 데이터 객체를 모두 포함하는 최소 경계 사각형(Minimum Boundary Rectangle, MBR) 정보를 기초로 인덱스 트리를 생성하는 과정을 포함한다.

발명의 효과

- [0033] 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.
- [0034] 본 발명의 일 실시예에 따른 분할 공간 기반의 공간 데이터 객체 질의처리장치 및 방법은 재귀적인 사분할을 통해 분할된 복수의 분할 공간 및 공간 데이터 객체의 최소 경계 사각형을 기초로 인덱스 트리를 생성할 수 있다.
- [0035] 본 발명의 일 실시예에 따른 분할 공간 기반의 공간 데이터 객체 질의처리장치 및 방법은 효율적인 인덱스 트리를 이용하여 유사 질의를 처리함으로써 긍정 오류를 효과적으로 줄일 수 있다.

도면의 간단한 설명

- [0037] 도 1은 본 발명의 일 실시예에 따른 분할 공간 기반의 공간 데이터 객체 질의처리 시스템을 설명하는 도면이다.
- 도 2는 도 1에 있는 공간 데이터 객체 질의처리장치를 설명하는 블록도이다.
- 도 3은 도 1에 있는 공간 데이터 객체 질의처리장치에서 공간 데이터 객체 질의를 처리하는 과정을 설명하는 순서도이다.
- 도 4는 선형화 기법을 사용하여 범위 질의를 처리하는 과정을 설명하는 예시도이다.
- 도 5는 다차원 인덱싱 레이어를 사용하여 범위 질의를 처리하는 과정을 설명하는 예시도이다.
- 도 6은 공간 데이터 객체 질의처리장치에서 수행되는 공간 분할 과정을 설명하는 예시도이다.
- 도 7은 도 2에 있는 인덱스 트리 생성부에서 생성하는 인덱스 트리 노드의 구조를 설명하는 예시도이다.
- 도 8은 도 2에 있는 인덱스 트리 생성부에서 생성하는 인덱스 트리 노드에 대한 테이블 구성을 설명하는 예시도이다.
- 도 9는 공간 데이터 객체 질의처리장치에서 수행되는 범위 질의 처리 과정을 설명하는 예시도이다.

발명을 실시하기 위한 구체적인 내용

- [0038] 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.
- [0039] 한편, 본 출원에서 서술되는 용어의 의미는 다음과 같이 이해되어야 할 것이다.
- [0040] "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.
- [0041] 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.

- [0042] 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.
- [0043] 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.
- [0044] 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.
- [0045] 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.
- [0047] 도 1은 본 발명의 일 실시예에 따른 분할 공간 기반의 공간 데이터 객체 질의처리 시스템을 설명하는 도면이다.
- [0048] 도 1을 참조하면, 분할 공간 기반의 공간 데이터 객체 질의처리 시스템(이하, 공간 데이터 객체 질의처리 시스템이라 한다.)은 사용자 단말(110), 공간 데이터 객체 질의처리장치(130) 및 데이터베이스(150)를 포함할 수 있다.
- [0049] 사용자 단말(110)은 공간 데이터 객체 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 공간 데이터 객체 질의처리장치(130)와 네트워크를 통해 연결될 수 있고, 사용자 단말(110a) 내지 사용자 단말n(110c)을 포함하는 복수의 사용자 단말(110)은 공간 데이터 객체 질의처리장치(130)와 동시에 연결될 수 있다.
- [0050] 공간 데이터 객체 질의처리장치(130)는 공간 분할을 통해 생성한 인덱스 트리를 이용하여 공간 데이터 객체 질의를 수신하여 처리할 수 있는 컴퓨터 또는 프로그램에 해당하는 서버로 구현될 수 있다. 공간 데이터 객체 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.
- [0051] 공간 데이터 객체 질의처리장치(130)는 데이터베이스(150)를 포함하여 구현될 수 있고, 데이터베이스(150)와 독립적으로 구현될 수 있다. 데이터베이스(150)와 독립적으로 구현된 경우 공간 데이터 객체 질의처리장치(130)는 데이터베이스(150)와 유선 또는 무선으로 연결되어 데이터를 주고 받을 수 있다.
- [0052] 데이터베이스(150)는 공간 데이터 객체 질의처리를 위해 필요한 다양한 정보들을 저장할 수 있는 저장장치이다. 데이터베이스(150)는 사용자 단말(110)로부터 수신한 적어도 하나의 범위 질의 또는 kNN 질의에 관한 정보 및 해당 질의를 처리하여 얻은 결과를 저장할 수 있고, 반드시 이에 한정되지 않고, 공간 데이터 객체 질의처리장치(130)가 공간 데이터 객체 질의를 처리하는 과정에서 다양한 형태로 수집하거나 가공한 정보들을 저장할 수 있다.
- [0053] 일 실시예에서, 데이터베이스(150)는 공간 데이터들을 저장하고 관리할 수 있는 HBase에 해당할 수 있다. 여기서, HBase는 아파치 HBase(Apache HBase)로서 하둡(Hadoop) 플랫폼을 위한 공개 비관계형 분산 데이터베이스에 해당할 수 있다. HBase는 구조화된 대용량의 데이터에 빠른 임의접근을 제공하는 구글의 빅 테이블과 비슷한 데이터 모델을 가지고, HDFS(Hadoop Distributed File System)의 데이터에 대한 실시간 임의 읽기/쓰기 기능을 제공할 수 있다.
- [0054] 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 서브-데이터베이스들로 구성되는 경우에는 각각의 서브-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 서브-데이터베이스들을 하나로 통합하고 상호 간의 데이터 교환 및 제어 흐름을 관리하는 제어부를 포함할 수 있다.

[0056] 도 2는 도 1에 있는 공간 데이터 객체 질의처리장치를 설명하는 블록도이다.

[0057] 도 2를 참조하면, 공간 데이터 객체 질의처리장치(130)는 분할 공간 생성부(210), 인덱스 트리 생성부(230), 질의 처리부(250) 및 제어부(270)를 포함할 수 있다.

[0058] 분할 공간 생성부(210)는 적어도 하나의 공간 데이터 객체를 포함하는 데이터 공간을 분할하여 복수의 분할 공간들을 생성할 수 있다. 여기에서, 데이터 공간은 공간 데이터 객체가 정의되는 다차원 공간에 해당할 수 있다. 공간 데이터 객체는 각 차원에 해당하는 속성 값을 포함하여 정의될 수 있고, 분할 공간 생성부(210)는 공간 데이터 객체가 가질 수 있는 각 속성 값들의 범위를 이용하여 데이터 공간을 정의할 수 있다.

[0059] 예를 들어, 분할 공간 생성부(210)는 데이터 공간이 2차원 좌표계를 가지고 각 축의 길이가 원점으로부터 100의 길이를 가지는 정사각형 공간에 해당하는 경우, 해당 공간의 중심점인 (50, 50)의 위치를 기준으로 해당 중심점을 지나는 각 축에 평행한 직선을 기초로 4개의 분할 공간으로 사분할 할 수 있다. 사분할 된 각 분할 공간은 모두 정사각형에 해당할 수 있고, 동일한 크기를 가질 수 있다.

[0060] 일 실시예에서, 분할 공간 생성부(210)는 복수의 분할 공간들 각각에 관해 해당 공간 데이터 객체의 밀집도가 특정 기준 이하가 될 때까지 해당 분할 공간을 해당 분할 공간의 중심점을 기준으로 재귀적으로 재분할 할 수 있다. 여기에서, 공간 데이터 객체의 밀집도는 특정 분할 공간 내에 존재하는 공간 데이터 객체의 수에 해당할 수 있다. 예를 들어, 분할 공간 내에 존재하는 공간 데이터 객체의 수가 4 이하가 될 때까지 재분할을 하는 경우 분할 공간 생성부(210)는 특정 분할 공간이 해당 분할 공간 내에 9개의 공간 데이터 객체를 포함하고 있다면 해당 분할 공간을 4개의 분할 공간으로 재분할 할 수 있고, 재분할된 분할 공간 중에서 4개 이상의 공간 데이터 객체를 포함하고 있는 분할 공간이 존재하는 경우 해당 분할 공간에 대해서는 재분할 과정을 다시 수행할 수 있다.

[0061] 인덱스 트리 생성부(230)는 분할 공간에 관한 정보 및 해당 분할 공간 데이터 객체를 포함하는 최소 경계 사각형(Minimum Boundary Rectangle, MBR) 정보를 기초로 인덱스 트리를 생성할 수 있다. 여기에서, 최소 경계 사각형은 해당 분할 공간 데이터 객체를 모두 포함하는 사각형 중에서 넓이가 가장 최소인 사각형에 해당할 수 있다. 인덱스 트리 생성부(230)는 특정 분할 공간 및 해당 분할 공간에 존재하는 공간 데이터 객체들 모두를 포함하는 최소 경계 사각형 각각에 대한 정보를 이용하여 인덱스 트리를 생성할 수 있다.

[0062] 일 실시예에서, 인덱스 트리 생성부(230)는 분할 공간에 관한 정보를 기초로 생성된 키(Key) 값 및 최소 경계 사각형 정보를 기초로 생성된 밸류(Value) 값을 포함하는 행 데이터로 구성된 테이블을 기초로 인덱스 트리를 생성할 수 있다. 예를 들어, 인덱스 트리 생성부(230)는 (키, 밸류) 값으로 구성된 행 데이터를 생성하여 HBase 테이블의 한 행으로 저장할 수 있다.

[0063] 일 실시예에서, 인덱스 트리 생성부(230)는 분할 공간이 생성될 때마다 각각의 축에 대해 원점에서 가까운 방향의 분할 공간을 0, 먼 방향의 분할 공간을 1로 표시하고 각 축에 대한 비트(bit)를 연결하여 키 값을 생성할 수 있다. 예를 들어, 인덱스 트리 생성부(230)는 분할을 통해 생성된 분할 공간이 있는 경우 해당 분할에서 x축에 대해서는 원점으로부터 먼 방향이고 y축에 대해서는 원점에서 가까운 방향의 분할 공간이라면 각각 '1', '0'으로 표시될 수 있고 이를 연결하여 최종적으로 '10'으로 표시될 수 있다.

[0064] 일 실시예에서, 인덱스 트리 생성부(230)는 분할 공간에 대한 재귀적인 재분할이 발생하는 경우 이전 분할 정보 및 현재 발생한 재분할 공간에 대한 정보를 연결하는 방식으로 표시할 수 있다. 예를 들어, 인덱스 트리 생성부(230)는 분할을 통해 생성되고 '10'으로 표시된 분할 공간에 대해 재분할이 발생한 경우, 재분할된 4개의 분할 공간들은 각각 '00', '01', '10' 및 '11'로 표시될 수 있고, 이전 분할 정보인 '10'에 연결되어 최종적으로 '1000', '1001', '1010' 및 '1011'로 표시될 수 있다.

[0065] 일 실시예에서, 인덱스 트리 생성부(230)는 내부 노드에 관한 행 데이터를 저장하는 인덱스 테이블 및 리프 노드에 관한 행 데이터를 저장하는 데이터 테이블을 생성함으로써 인덱스 트리를 생성할 수 있다. 인덱스 테이블은 내부 노드에 관한 행 데이터를 저장할 수 있고, 내부 노드에 관한 행 데이터는 분할 공간 생성부(210)에 의

해 생성된 분할 공간에 관한 정보를 기초로 생성된 키 값, 해당 내부 노드를 루트로 하는 트리 내에 존재하는 공간 데이터 객체의 수, 해당 내부 노드의 자식 노드 정보 및 각 자식 노드와 연관된 MBR 정보를 포함할 수 있다. 데이터 테이블은 리프 노드에 관한 행 데이터를 저장할 수 있고, 리프 노드에 관한 행 데이터는 각 리프 노드와 연관된 분할 공간에 관한 정보를 기초로 생성된 키 값, 해당 리프 노드와 연관된 분할 공간 내에 존재하는 공간 데이터 객체에 관한 정보를 포함할 수 있다.

[0066] 질의 처리부(250)는 인덱스 트리 생성부(230)를 통해 생성된 인덱스 트리를 이용하여 공간 데이터 객체 질의를 처리할 수 있다. 여기에서, 인덱스 트리는 Q-MBR(Quadrant-based MBR) 트리에 해당할 수 있다. Q-MBR 및 이를 이용한 Q-MBR 트리에 대해서는 도 6 내지 8에서 보다 자세히 설명한다.

[0067] 일 실시예에서, 질의 처리부(250)는 인덱스 트리 생성부(230)를 통해 생성된 인덱스 트리를 이용하여 질의 포인트와 질의 반경을 포함하는 범위 질의 또는 질의 포인트와 최근접 이웃 수를 포함하는 kNN 질의를 처리할 수 있다. 보다 구체적으로, 질의 처리부(250)는 인덱스를 탐색하는 동안 Q-MBR 트리를 너비 우선 순서(Breadth First Search, BFS)에 따라 탐색하고 다음 반복(iteration)을 로드하기 위한 행 키(Rowkey)를 계산할 수 있다.

[0068] 일 실시예에서, 질의 처리부(250)는 범위 질의 처리를 위한 인덱스 검색을 위해 N, I 및 D라는 세 집합을 사용할 수 있다. 여기에서, N은 현재의 반복(iteration)에서 탐색을 위해 로드된 노드를 저장할 수 있다. I는 다음 반복(iteration)에서 인덱스 테이블에서 로드될 행 키 집합에 해당할 수 있다. D는 데이터 테이블에서 로드할 행 키 집합에 해당할 수 있다.

[0069] 질의 처리부(250)는 루트 노드의 행 키를 I에 삽입한 다음 인덱스 테이블에서 노드를 로드하여 탐색을 시작할 수 있다. 질의 처리부(250)는 인덱스 트리의 최대 깊이를 알고 있는 경우 루트 노드에 대한 단일 행 키만 삽입하는 대신 하위 레벨 노드에 대한 복수의 행 키를 I에 삽입할 수 있다.

[0070] 질의 처리부(250)는 현재 노드가 내부 노드인 경우 자식 노드의 MBR과 질의 포인트 q 간의 최소 거리를 계산할 수 있고, 거리가 질의 반경 r 이내에 있는 자식의 행 키를 I에 삽입할 수 있다. 질의 처리부(250)는 N에 있는 모든 노드를 탐색한 후 I에 저장된 행 키를 사용하여 추가 노드들을 프리페치(prefetch)하고 다음 반복(iteration)을 위해 N에 결과를 저장할 수 있다.

[0071] 질의 처리부(250)는 인덱스 탐색 과정에서 리프 노드에 도달하면 리프 노드의 행 키를 D에 저장할 수 있고, 질의 처리 마지막 단계에서 관련 공간 데이터 객체를 로드할 수 있다. 질의 처리부(250)는 I에 노드가 없는 경우 리프 노드에 저장된 공간 데이터 객체를 D에 로드할 수 있고, 공간 데이터 객체와 질의 포인트 q 간의 거리를 계산하여 범위 질의에 응답할 수 있다.

[0072] 질의 처리부(250)는 질의 포인트 q와의 거리가 질의 반경 r보다 작거나 같은 공간 데이터 객체만을 결과 집합에 삽입할 수 있다. 질의 처리부(250)는 결과 집합에 속한 공간 데이터 객체에 관한 정보를 질의 처리 결과로서 사용자 단말(110)에 제공할 수 있다.

[0073] 일 실시예에서, 질의 처리부(250)는 kNN 질의 처리를 위해 N과 Q의 두가지 우선순위 큐(Queue)를 유지할 수 있다. 여기에서, N은 질의 포인트에서 MBR까지의 최소 거리의 오름차순으로 노드를 저장할 수 있고, Q는 kNN의 후보 집합을 내림차순으로 저장할 수 있다.

[0074] 질의 처리부(250)는 kNN 질의 두 단계로 처리할 수 있다. 첫 번째 단계에서 질의 처리부(250)는 근사 범위 r을 사용하여 리프 노드의 행 키를 찾기 위해 Q-MBR 트리를 탐색할 수 있다. 여기에서, 근사 범위는 원하는 이웃 수를 찾기 위해 충분한 수의 공간 데이터 객체를 보장하는 최소 거리에 해당할 수 있다.

[0075] 질의 처리부(250)는 인덱스 탐색 과정에서 N에 저장된 노드들을 순차적으로 검사하여 최소 거리가 r보다 작은 자식 노드를 가지고 있는지를 결정할 수 있다. 질의 처리부(250)는 자식 노드의 최소 거리가 r보다 작은 경우 자식 노드의 행 키를 리스트 I에 삽입할 수 있다. 질의 처리부(250)는 범위 r은 초기 단계에서 매우 크게 설정할 수 있고, 현재 노드보다 가까운 노드의 객체 수가 k를 초과하면 N에서 노드의 최대 거리로 갱신할 수 있다.

[0076] 질의 처리부(250)는 현재 노드가 리프 노드인 경우 현재 노드의 행 키를 리스트 I에 저장하여 함께 로드할 수 있다. 질의 처리부(250)는 N에 더 이상 노드가 없으면 행 키가 I에 저장된 노드를 로드할 수 있다. 질의 처리부(250)는 N에 노드가 없고 I에 행 키가 없을 때 첫 번째 단계를 종료할 수 있다. 질의 처리부(250)는 인덱스 검색 후 I에 저장된 리프 노드의 공간 데이터 객체를 로드하고 거리를 평가하여 질의 결과 집합을 식별할 수 있다. 질의 처리부(250)는 평가된 공간 데이터 객체를 Q에 저장할 수 있고, 평가할 공간 데이터 객체가 더 이상 없을 때 질의 결과 집합을 반환할 수 있다.

- [0077] 제어부(270)는 공간 데이터 객체 질의처리장치(130)의 전체적인 동작을 제어하고, 분할 공간 생성부(210), 인덱스 트리 생성부(230) 및 질의 처리부(250) 간의 제어 흐름 또는 데이터 흐름을 관리할 수 있다.
- [0079] 도 3은 도 1에 있는 공간 데이터 객체 질의처리장치에서 공간 데이터 객체 질의를 처리하는 과정을 설명하는 순서도이다.
- [0080] 도 3을 참조하면, 공간 데이터 객체 질의처리장치(130)는 분할 공간 생성부(210)를 통해 적어도 하나의 공간 데이터 객체를 포함하는 데이터 공간을 분할하여 복수의 분할 공간들을 생성할 수 있다(단계 S310). 일 실시예에서, 분할 공간 생성부(210)는 복수의 분할 공간들 각각에 관해 해당 공간 데이터 객체의 밀집도가 특정 기준 이하가 될 때까지 해당 분할 공간을 해당 분할 공간의 중심점을 기준으로 재귀적으로 재분할 할 수 있다.
- [0081] 공간 데이터 객체 질의처리장치(130)는 인덱스 트리 생성부(230)를 통해 분할 공간에 관한 정보 및 해당 분할 공간에 포함된 공간 데이터 객체를 모두 포함하는 최소 경계 사각형 정보를 기초로 인덱스 트리를 생성할 수 있다(단계 S330). 일 실시예에서, 인덱스 트리 생성부(230)는 분할 공간에 관한 정보를 기초로 생성된 키(Key) 값 및 최소 경계 사각형 정보를 기초로 생성된 밸류(Value) 값을 포함하는 행 데이터로 구성된 테이블을 기초로 인덱스 트리를 생성할 수 있다.
- [0082] 일 실시예에서, 인덱스 트리 생성부(230)는 분할 공간이 생성될 때마다 각각의 축에 대해 원점에서 가까운 방향의 분할 공간을 0, 먼 방향의 분할 공간을 1로 표시하고 각 축에 대한 비트(bit)를 연결하여 키 값을 생성할 수 있다. 일 실시예에서, 인덱스 트리 생성부(230)는 분할 공간에 대한 재귀적인 재분할이 발생하는 경우 이전 분할 정보 및 현재 발생한 재분할 공간에 대한 정보를 연결하는 방식으로 표시할 수 있다. 일 실시예에서, 인덱스 트리 생성부(230)는 내부 노드에 관한 행 데이터를 저장하는 인덱스 테이블 및 리프 노드에 관한 행 데이터를 저장하는 데이터 테이블을 생성함으로써 인덱스 트리를 생성할 수 있다.
- [0083] 공간 데이터 객체 질의처리장치(130)는 질의 처리부(250)를 통해 인덱스 트리 생성부(230)에 의해 생성된 인덱스 트리를 이용하여 공간 데이터 객체 질의를 처리할 수 있다(단계 S350). 일 실시예에서, 질의 처리부(250)는 인덱스 트리를 이용하여 질의 포인트와 질의 반경을 포함하는 범위 질의 또는 질의 포인트와 최근접 이웃 수를 포함하는 kNN 질의를 처리할 수 있다.
- [0085] 도 6은 공간 데이터 객체 질의처리장치에서 수행되는 공간 분할 과정을 설명하는 예시도이다.
- [0086] 도 6을 참조하면, 공간 데이터 객체 질의처리장치(130)는 분할 공간 생성부(210)를 통해 원점(611)을 기준으로 수직으로 교차하고 각각 80의 최대 크기를 갖는 두개의 축에 의해 형성되는 데이터 공간(610)을 정의할 수 있다. 분할 공간 생성부(210)는 데이터 공간(610)을 공간의 중심점 (40,40)을 기준으로 사분할하여 복수의 분할 공간들을 생성할 수 있다. 분할 공간 생성부(210)는 해당 공간 데이터 객체의 밀집도가 특정 기준 이하가 될 때까지 해당 분할 공간을 재귀적으로 재분할 할 수 있다. 도 6에서, 밀집도는 4에 해당할 수 있다. 즉, 분할 공간 생성부(210)는 단일 분할 공간 내부에 존재하는 공간 데이터 객체의 수가 4이하가 될 때까지 재귀적으로 공간을 재분할 할 수 있다.
- [0087] 공간 데이터 객체 질의처리장치(130)는 데이터 공간(610)을 분할한 후 각 사분면에 존재하는 공간 데이터 객체에 대한 최소 경계 사각형(MBR)(613)을 생성할 수 있다. 공간 데이터 객체 질의처리장치(130)는 분할 공간(631) 및 MBR(633)에 관한 정보를 기초로 Q-MBR(Quadrant-based MBR)(630)을 생성할 수 있고, HBase 데이터베이스에 저장할 수 있다. 공간 데이터 객체 질의처리장치(130)는 Q-MBR(630)을 HBase 테이블에 저장하여 계층적 인덱스 트리의 빌딩 블록으로 사용할 수 있다.
- [0088] 분할 공간(631)에 관한 정보는 업데이트 비용을 줄이기 위해 행 키로 사용될 수 있다. 각 열 값은 별도의 키(key)-밸류(value) 쌍 형식으로 저장될 수 있고, 행 키를 업데이트하면 해당 키-밸류 쌍에 많은 수의 삽입이 발생할 수 있다. 따라서, 자주 업데이트되는 MBR(633) 정보는 행 키로 사용할 수 없다. 최악의 경우는 업데이트가 발생할 때 그룹화된 공간 데이터 객체에 대해 새로운 행 키를 만들고 모든 키-밸류 쌍을 다시 생성하는 경우에 해당할 수 있다. 따라서, MBR(633) 정보는 상대적으로 업데이트하기 어려운 열에 저장될 수 있고, 공간 질의 처리 중에 거리 계산에 사용될 수 있다.

- [0090] 도 7은 도 2에 있는 인덱스 트리 생성부에서 생성하는 인덱스 트리 노드의 구조를 설명하는 예시도이다.
- [0091] 도 7을 참조하면, 공간 데이터 객체 질의처리장치(130)는 인덱스 트리 생성부(230)를 통해 Q-MBR을 생성하고 관리할 수 있는 인덱스 트리를 생성할 수 있다. 여기에서, 인덱스 트리는 Q-MBR 트리에 해당할 수 있고, HBase 테이블 또는 메모리 인덱스로 구현될 수 있다. Q-MBR 트리는 쿼드 트리 구조와 유사할 수 있다.
- [0092] 일 실시예에서, 인덱스 트리 생성부(230)는 내부 노드(Internal Node)에 관한 행 데이터를 저장하는 인덱스 테이블 및 리프 노드(Leaf Node)에 관한 행 데이터를 저장하는 데이터 테이블을 생성함으로써 인덱스 트리를 생성할 수 있다. 내부 노드는 인덱스 테이블(Index Table)에 저장될 수 있고, 노드의 사분면 정보(Quadrant), 자식 노드의 MBR(MBRs of children) 및 서브 트리에 포함된 공간 데이터 객체의 수(Number of objects)로 구성될 수 있다. 리프 노드는 데이터 테이블(Data Table)에 저장될 수 있고, 사분면(Quadrant), 리프 노드의 공간 데이터 객체의 수(Number of objects) 및 공간 데이터 객체 목록으로 구성될 수 있다.
- [0094] 도 8은 도 2에 있는 인덱스 트리 생성부에서 생성하는 인덱스 트리 노드에 대한 테이블 구성을 설명하는 예시도이다.
- [0095] 도 8을 참조하면, 리프 노드 영역 내에 존재하는 공간 데이터 객체에 대한 포인터 대신 사분면 정보를 사용하여 데이터 테이블에 저장된 공간 데이터 객체에 직접 접근할 수 있다. 도 8에 포함된 테이블의 구성은 도 6에 표시된 공간 데이터 객체에 대한 계층적인 Q-MBR 인덱스 구조의 예를 포함하고 있다. 각각의 Q-MBR은 인덱스 트리의 리프 노드에 해당할 수 있고, 부모 노드의 Q-MBR은 자식 노드의 Q-MBR들을 포함하는 사분면과 MBR로 표현될 수 있다.
- [0096] 인덱스 테이블의 'Meta family' 항목의 '#objects' 값은 해당 노드를 루트 노드로 하는 트리 내에 존재하는 공간 데이터 객체의 수를 나타낸다. 예를 들어, 'Rowkey' 항목 값이 'Root'인 노드에 대해 '#objects' 값이 15이므로 해당 트리 내에 존재하는 공간 데이터 객체의 수는 15개에 해당한다.
- [0098] 도 9는 공간 데이터 객체 질의처리장치에서 수행되는 범위 질의 처리 과정을 설명하는 예시도이다.
- [0099] 도 9를 참조하면, 공간 데이터 객체 질의처리장치(130)은 질의 처리부(250)를 통해 범위 질의를 처리할 수 있다. 질의 처리부(250)는 R0의 행 키를 L에 삽입할 수 있고, 이를 N에 로드함으로써 시작할 수 있다. 질의 처리부(250)는 두 개의 자식 노드인 R2 및 R3이 질의 범위에 중첩되므로 R2 및 R3의 행 키를 첫 번째 반복(iteration)에서 I에 삽입할 수 있고, 인덱스 테이블로부터 함께 로드할 수 있다. 동일하게, R9 및 R6의 행 키가 삽입되어 두 번째 반복(iteration)에서 로드될 수 있다. 질의 처리부(250)는 R9 및 R6가 리프 노드이므로 다음 반복(iteration)에서 범위 질의에 대한 응답을 위해 R9 및 R6의 공간 데이터 객체를 검사할 수 있다. 결과 집합 R은 2개의 공간 데이터 객체 p1과 p2를 포함하고 있고, 탐색할 노드가 없기 때문에 질의 처리부(250)는 탐색을 종료할 수 있다.
- [0101] 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.

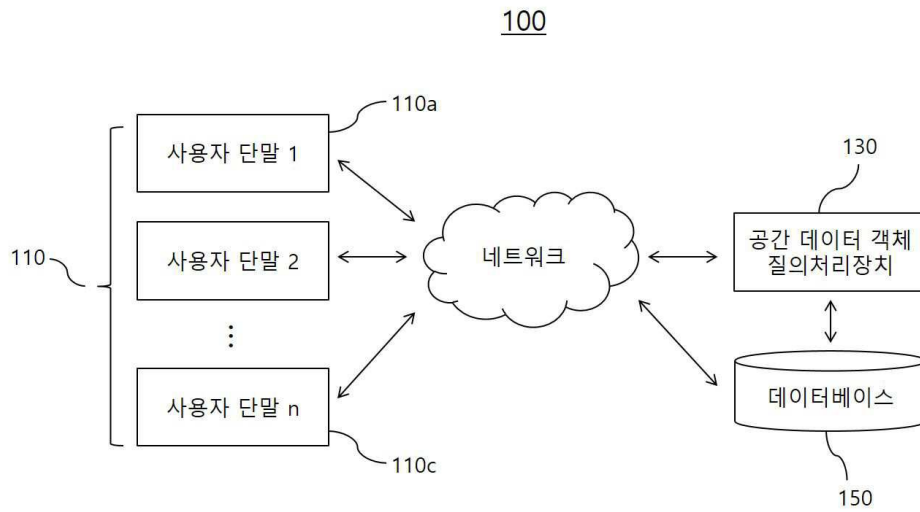
부호의 설명

- [0103] 100: 분할 공간 기반의 공간 데이터 객체 질의처리 시스템
 110: 사용자 단말 130: 공간 데이터 객체 질의처리장치
 150: 데이터베이스
 210: 분할 공간 생성부 230: 인덱스 트리 생성부
 250: 질의 처리부 270: 제어부

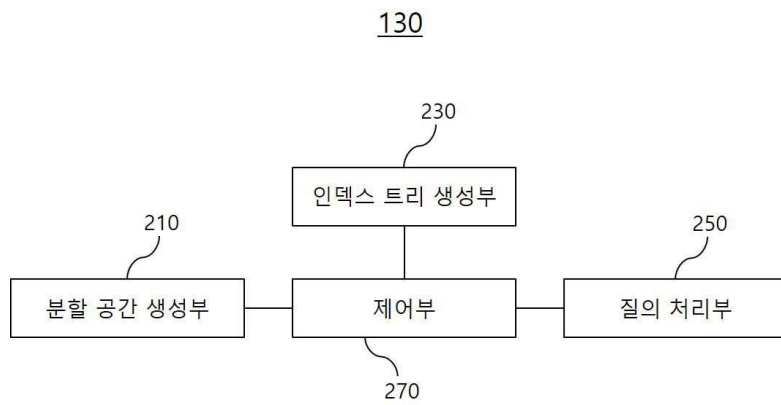
410: 범위 질의 430: 긍정 오류
 510: 범위 질의 530: 인덱스 레이어
 610: 데이터 공간 611: 원점
 613: 최소 경계 사각형 630: Q-MBR
 631: 분할 공간 633: MBR

도면

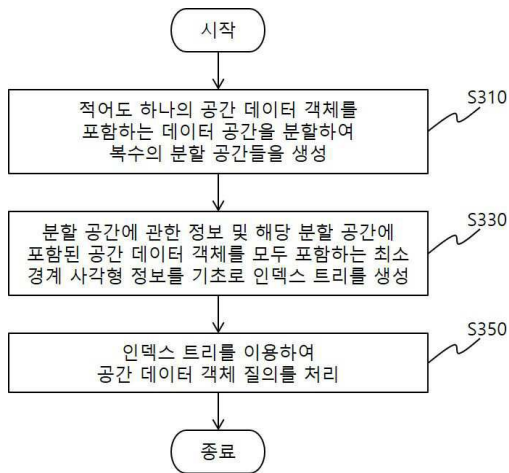
도면1



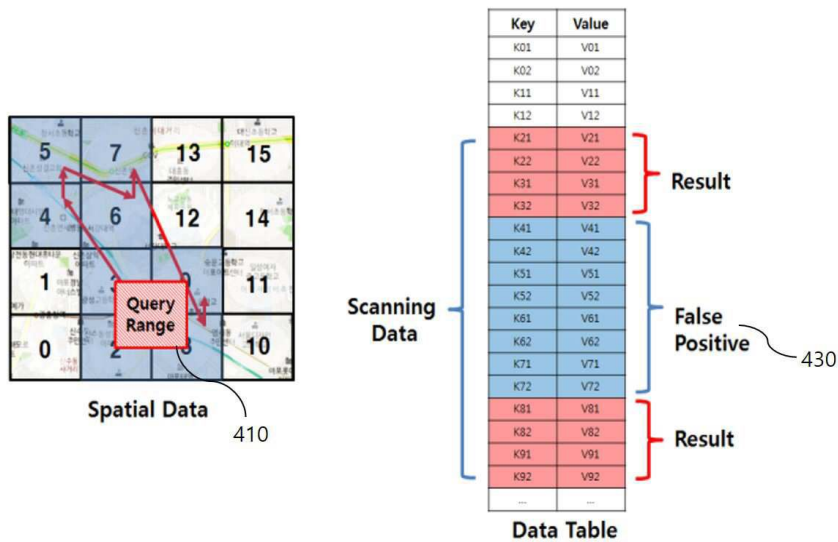
도면2



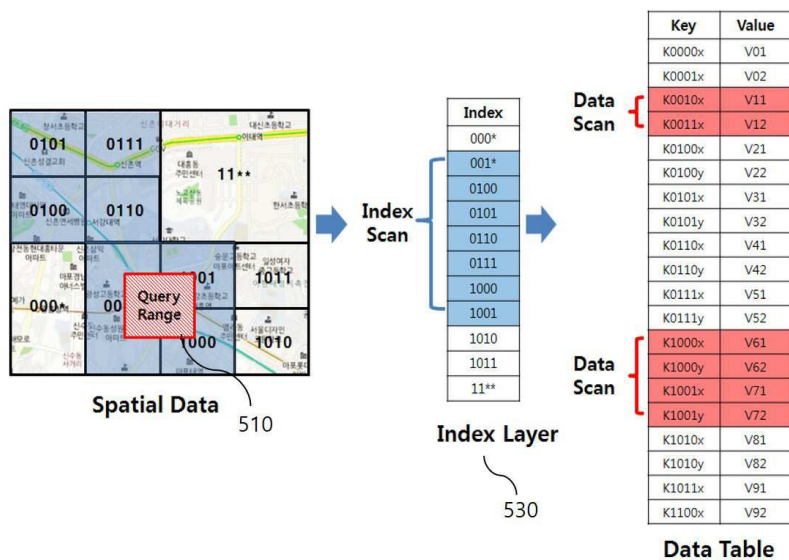
도면3



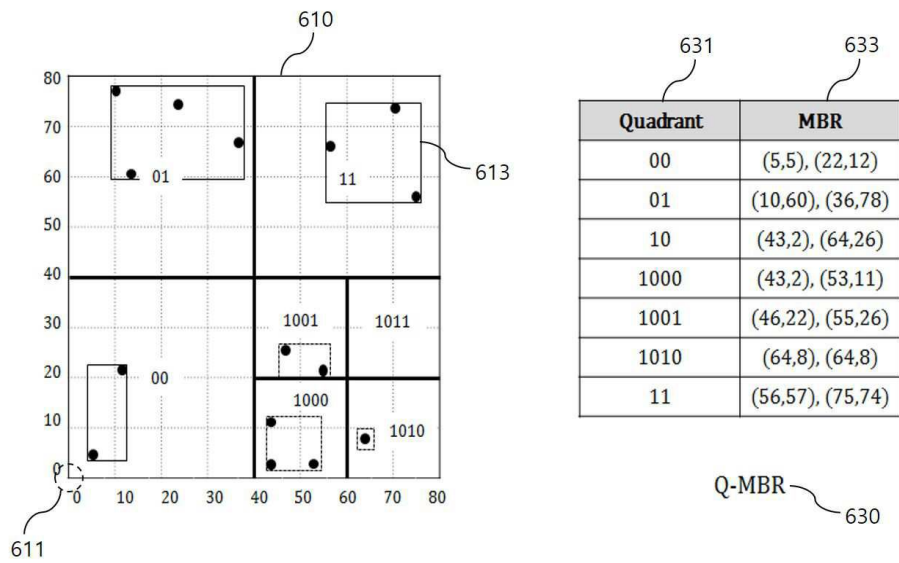
도면4



도면5



도면6



도면7

Type	Component	Description
Internal Node	Quadrant	Binary values of the quadrant information.
	MBRs of children	Coordinates of the lower left and upper right corners of the MBR.
	Number of objects	Number of objects in the sub-tree rooted at this node.
Leaf Node	Quadrant	Binary values of the quadrant information.
	Number of objects	Number of objects in this node.

도면8

Internal nodes in index table

Row key (Quadrant)	Meta family	MBR family			
	# objects	00	01	10	11
Root	15	(5,5), (22,12)	(10,60), (36,78)	(43,2), (64,26)	(56,57), (75,74)
10	6	(43,2), (53,11)	(46,22), (55,26)	(64,8), (64,8)	

Leaf nodes in data table

Row key (Quadrant)	Data Family			
	d1	d2	d3	d4
00	p ₁	p ₂		
01	p ₃	p ₄	p ₅	p ₆
1000	p ₇	p ₈	p ₉	
1001	p ₁₀	p ₁₁		
1010	p ₁₂			
11	p ₁₃	p ₁₄	p ₁₅	

도면9

