



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2014-0014819
(43) 공개일자 2014년02월06일

(51) 국제특허분류(Int. Cl.)
G06F 19/00 (2011.01) G01N 21/25 (2006.01)
(21) 출원번호 10-2012-0081875
(22) 출원일자 2012년07월26일
심사청구일자 2012년07월26일

(71) 출원인
인하대학교 산학협력단
인천광역시 남구 인하로 100, 인하대학교 (용현동)
(72) 발명자
김덕환
서울 양천구 목동로 230, 103동 1004호 (목동, 목동아이파크아파트)
유재환
인천 연수구 능히대로79번길 30, 103동 202호 (옥련동, 럭키송도아파트)
(74) 대리인
특허법인 정안

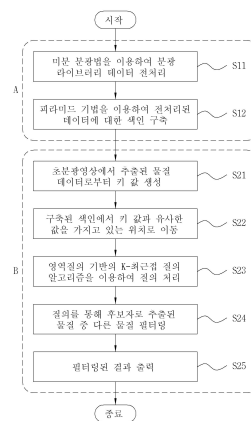
전체 청구항 수 : 총 8 항

(54) 발명의 명칭 피라미드 기법 기반 색인 및 질의를 통한 초분광영상의 물질탐지 방법

(57) 요약

본 발명은 피라미드 기법 기반 색인 및 질의를 통한 초분광영상의 물질탐지 방법에 관한 것으로, 본 발명에 따른 초분광영상의 물질탐지방법은 미분 분광법을 이용하여 분광 라이브러리 데이터 전처리하는 단계와, 피라미드 기법을 사용하여 상기 전처리된 분광 데이터에 대한 색인을 구축하는 단계와, 초분광영상으로부터 추출된 물질 데이터로부터 키 값을 생성하는 단계와, 상기 색인 내에서 상기 물질 데이터의 키 값과 유사한 분광 데이터의 키 값이 존재하는 위치로 이동하는 단계와, 상기 위치로 이동 후 그 위치를 기준으로 소정 추출범위 내에서 특정 질의 알고리즘에 따라 질의를 수행하여 상기 분광 라이브러리로부터 후보자 물질을 추출하는 단계와, 상기 질의를 통해 추출된 후보자 물질 중 다른 물질을 필터링하는 단계와, 상기 필터링을 통해 남은 최종 후보자를 출력하는 단계를 포함한다.

대표도 - 도1



특허청구의 범위

청구항 1

미분 분광법을 이용하여 분광 라이브러리 내 분광 데이터를 전처리하는 단계와,
 피라미드 기법을 사용하여 상기 전처리된 분광 데이터에 대한 색인을 구축하는 단계와,
 초분광영상으로부터 추출된 물질 데이터로부터 키 값을 생성하는 단계와,
 상기 색인 내에서 상기 물질 데이터의 키 값과 유사한 분광 데이터의 키 값이 존재하는 위치로 이동하는 단계와,
 상기 위치로 이동 후 그 위치를 기준으로 소정 추출범위 내에서 특정 질의 알고리즘에 따라 질의를 수행하여 상기 분광 라이브러리로부터 후보자 물질을 추출하는 단계와,
 상기 질의를 통해 추출된 후보자 물질 중 다른 물질을 필터링하는 단계와,
 상기 필터링을 통해 남은 최종 후보자를 출력하는 단계를 포함하는 것을 특징으로 하는 피라미드 기법 기반 색인 및 질의를 통한 초분광영상의 물질탐지 방법.

청구항 2

제1항에 있어서,
 상기 전처리 단계는,
 1차 미분 분광법과 2차 미분 분광법을 이용하여 상기 분광 데이터를 파장 영역 간의 변화량을 변환하는 것을 특징으로 하는 피라미드 기법 기반 색인 및 질의를 통한 초분광영상의 물질탐지 방법.

청구항 3

제1항에 있어서,
 상기 색인 구축 단계는,
 상기 전처리된 분광 데이터를 구성하는 밴드의 개수에 근거하여 상기 전처리된 분광 데이터를 소정 개의 피라미드로 분할하고 그 분할된 피라미드 중 어느 하나의 피라미드 번호를 선택하는 단계와,
 상기 선택된 피라미드에서 상기 전처리된 분광 데이터가 배치될 높이를 정하는 단계와,
 상기 선택된 피라미드에 배치될 키 값을 구하는 단계를 포함하는 것을 특징으로 하는 피라미드 기법 기반 색인 및 질의를 통한 초분광영상의 물질탐지 방법.

청구항 4

제1항에 있어서,
 상기 키 값 생성 단계는,
 상기 추출된 물질 데이터를 상기 미분 분광법을 이용하여 전처리하고 그 전처리된 데이터에 피라미드 기법을 적용하여 키 값을 생성하는 것을 특징으로 하는 피라미드 기법 기반 색인 및 질의를 통한 초분광영상의 물질탐지 방법.

청구항 5

제1항에 있어서,
 상기 특정 질의 알고리즘은 피라미드 기법을 이용한 영역질의 기반의 K-최근접 질의 알고리즘인 것을 특징으로 하는 피라미드 기법 기반 색인 및 질의를 통한 초분광영상의 물질탐지 방법.

청구항 6

제1항에 있어서,

상기 필터링 단계는,

적분 분광 정합 필터를 사용하여 필터링을 수행하는 것을 특징으로 하는 피라미드 기법 기반 색인 및 질의를 통한 초분광영상의 물질탐지 방법.

청구항 7

제1항에 있어서,

상기 최종 후보자 출력단계,

상기 최종 후보자 물질이 소정 개미만이면 상기 추출범위를 지정해진 범위로 확장하여 후보자 물질을 재검색하는 것을 특징으로 하는 피라미드 기법 기반 색인 및 질의를 통한 초분광영상의 물질탐지 방법.

청구항 8

제7항에 있어서,

상기 추출범위는 초기 범위의 ± 5 밴드로 제한하는 것을 특징으로 하는 피라미드 기법 기반 색인 및 질의를 통한 초분광영상의 물질탐지 방법.

명세서

기술분야

[0001] 본 발명은 피라미드 기법 기반의 색인 및 K-최근접 질의를 이용하여 필요한 데이터에만 접근하여 검색시간을 줄이고, 정확도 높은 후보자를 추출하여 물질 탐색의 정확도를 개선할 수 있는 피라미드 기법 기반 색인 및 질의를 통한 초분광영상의 물질탐지 방법에 관한 것이다.

배경기술

[0002] 현재 대부분의 초분광영상(Hyperspectral Image) 관련 연구는 유효밴드 추출, 전처리, 표적탐지 및 물질 분류 등과 같은 영상처리 분야에서 진행되고 있다. 하지만 높은 정확도를 갖는 물질 분류와 물질 탐색을 하기 위해서는 지상에서 측정된 분광 라이브러리 데이터와 초분광영상을 이용한 분광 정합(Spectral matching), 색인 구축(indexing) 및 후보자 필터와 같은 연구들이 필요하다.

[0003] 분광라이브러리는 지표에 존재하는 다양한 물질에 대한 분광반사율 자료를 모아놓은 일종의 데이터베이스라고 할 수 있다. 분광라이브러리의 분광반사율 자료는 지상에서 근거리로 측정되기 때문에 다른 측정기와 비교했을 때 노이즈가 적고 대부분 400~2500nm의 파장 범위에서 약 1~10nm의 파장 폭을 갖고 있다. 따라서 다양한 분야에서 물질 비교 및 특성을 분석하기 위한 참조 데이터로 널리 사용되고 있다.

[0004] 100~200nm의 넓은 파장 폭과 3~10개의 밴드를 가지고 있는 다중분광 영상(Multi spectral image)과 달리 초분광영상은 400~2500nm 파장 범위에서 약 10nm의 파장 폭을 갖고, 수백 개의 밴드 즉 수백 개의 차원으로 구성된 고차원 데이터이다.

[0005] 이러한 데이터를 Kd-Tree, R-tree, X-tree 등과 같은 다차원 색인 방법을 사용하여 색인을 구축하게 되면 차원의 저주(Curse of dimensionality)라는 문제가 발생하게 된다. 차원의 저주란 차원의 수가 증가함에 따라 검색 시간이나 기억공간에 대한 요구가 지수 승으로 증가하여 순차적 접근 방법보다 처리시간이 오래 걸리게 되는 현상이다. 이와 더불어 아무런 필터링 없이 질의 알고리즘을 수행했을 시 입력된 물질과는 전혀 관련 없는 결과가 나올 확률이 크다.

[0006] 고차원 데이터의 색인 구축과정에서 발생하는 차원의 저주 문제를 해결하기 위해 피라미드 기법(Pyramid-Technique)이 제안되었다. 피라미드 기법은 다차원의 데이터를 1차원의 데이터로 변환하고 이를 B+ 트리(tree)를 이용하여 색인하는 방법이다. 최근에는 피라미드 기법을 이용하여 분광 라이브러리를 위한 색인을 구축하는 방법이 제안되고 있다. 그러나 이 방법은 분광 라이브러리 색인에는 좋은 효율을 보일 수 있지만, 초분광영상에 사용하기에는 적합하지 않다. 분광 데이터는 같은 물질이라도 태양의 각도, 습도, 날씨 등의 환경적인 문제들로 인하여 전체 밴드에 대한 상대적인 반사율은 같으나, 절대 반사율이 다른 문제가 발생할 수 있다. 따라서,

이런 데이터들을 종래의 피라미드 기법과 같이 영역이 고정되어있는 색인 알고리즘에 사용했을 시 물질 탐지가 불가능 할 수도 있으며, 서로 다른 물질이 특정밴드와 반사율 값이 같아 후보자로 추출될 수도 있는 문제점이 있다.

선행기술문헌

비특허문헌

- [0007] (비특허문헌 0001) Dimitris Manolakis, David Marden, Gary A. Shaw, "Hyperspectral image Processing for Automatic Target Detection Applications," Lincoln Laboratory Journal, vol.14, pp.79-116, 2003.
- (비특허문헌 0002) David Landgrebe, "Hyperspectral Image Data Analysis as a High Dimensional Signal Processing Problem," IEEE Signal Processing Magazine, vol.19. no.1, pp.17-28, Jan. 2002.
- (비특허문헌 0003) Stefan Berchtold, Christian Bohm, Daniel A.Keim, Hans-Peter Kriegel, "A cost model for nearest neighbor search in high-dimensional data space," In Proceedings of PODS, pp.78-86, 1997.
- (비특허문헌 0004) Piotr Indyk, Rajeev Motwani, "Approximate Nearest Neighbors: Towards Removing the curse of Dimensionality," In Proceedings of the thirtieth annual ACM symposium on Theory of computing, pp.605-613, 1998.
- (비특허문헌 0005) Stefan Berchtold, Christian Bohm, Hans-Pter Kriegel, "The Pyramid-Technique: Towards Breaking the Curse of Dimensionality," In Proceedings of Conference ACM SIGMOD, pp.142-153, 1998.
- (비특허문헌 0006) Jia Li, Cheng Wang, "Indexing Method for Hyperspectral Data Fast Retrieval by Pyramid Technique," In CSSE Proceedings of the 2008 International Conference, vol.04, pp.519-523, 2008.
- (비특허문헌 0007) Jae-Hwan Yu, Deok-Hwan Kim, "Spectral matching using Range Queries based on Pyramid-Technique in Hyperspectral Image Library," In Proceedings of Spring KoCon Conference Jeju, vol.9, no.1 pp. 153-154, 2011.

발명의 내용

해결하려는 과제

- [0008] 본 발명은 종래기술의 문제점을 해결하기 위하여 안출된 것으로, 피라미드 기법을 사용하여 초분광영상의 색인을 구축하고, 그 구축된 색인을 이용하여 초분광영상의 물질을 검색하는 피라미드 기법 기반 색인 및 질의를 통한 초분광영상의 물질탐지방법을 제공하기 위한 것이다.
- [0009] 또한, 본 발명은 미분 분광법을 이용한 전처리를 통해 분광 라이브러리에서대상 물질을 추출하고, 피라미드 기법 기반의 K-최근접 질의를 통해 후보자를 검출하는 피라미드 기법 기반 색인 및 질의를 통한 초분광영상의 물질탐지방법을 제공하기 위한 것이다.
- [0010] 또한, 본 발명은 질의 처리 과정에서 적분 분광 정합 필터를 통해 거짓값을 필터링하여 물질탐지의 정확도를 개선할 수 있는 피라미드 기법 기반 색인 및 질의를 통한 초분광영상의 물질탐지방법을 제공하기 위한 것이다.

과제의 해결 수단

- [0011] 상기한 과제를 해결하기 위하여, 본 발명에 따른 피라미드 기법 기반 색인 및 질의를 통한 초분광영상의 물질탐지방법은 미분 분광법을 이용하여 분광 라이브러리 내 분광 데이터를 전처리하는 단계와, 피라미드 기법을 사용하여 상기 전처리된 분광 데이터에 대한 색인을 구축하는 단계와, 초분광영상으로부터 추출된 물질 데이터로부터 키 값을 생성하는 단계와, 상기 색인 내에서 상기 물질 데이터의 키 값과 유사한 분광 데이터의 키 값이 존

재하는 위치로 이동하는 단계와, 상기 위치로 이동 후 그 위치를 기준으로 소정 추출범위 내에서 특정 질의 알고리즘에 따라 질의를 수행하여 상기 분광 라이브러리로부터 후보자 물질을 추출하는 단계와, 상기 질의를 통해 추출된 후보자 물질 중 다른 물질을 필터링하는 단계와, 상기 필터링을 통해 남은 최종 후보자를 출력하는 단계를 포함하는 것을 특징으로 한다.

[0012] 상기 전처리 단계는, 1차 미분 분광법과 2차 미분 분광법을 이용하여 상기 분광 데이터를 파장 영역 간의 변화량을 변환하는 것을 특징으로 한다.

[0013] 상기 색인 구축 단계는, 상기 전처리된 분광 데이터를 구성하는 밴드의 개수에 근거하여 상기 전처리된 분광 데이터를 소정 개의 피라미드로 분할하고 그 분할된 피라미드 중 어느 하나의 피라미드 번호를 선택하는 단계와, 상기 선택된 피라미드에서 상기 전처리된 분광 데이터가 배치될 높이를 정하는 단계와, 상기 선택된 피라미드에 배치될 키 값을 구하는 단계를 포함하는 것을 특징으로 한다.

[0014] 상기 키 값 생성 단계는, 상기 추출된 물질 데이터를 상기 미분 분광법을 이용하여 전처리하고 그 전처리된 데이터에 피라미드 기법을 적용하여 키 값을 생성하는 것을 특징으로 한다.

[0015] 상기 특정 질의 알고리즘은 피라미드 기법을 이용한 영역질의 기반의 K-최근접 질의 알고리즘인 것을 특징으로 한다.

[0016] 상기 필터링 단계는, 적분 분광 정합 필터를 사용하여 필터링을 수행하는 것을 특징으로 한다.

[0017] 상기 최종 후보자 출력단계, 상기 최종 후보자 물질이 소정 개미만이면 상기 추출범위를 지정해진 범위로 확장하여 후보자 물질을 재검색하는 것을 특징으로 한다.

[0018] 상기 추출범위는 초기 범위의 ± 5 밴드로 제한하는 것을 특징으로 한다.

발명의 효과

[0019] 본 발명은 피라미드 기법을 사용하여 초분광영상의 색인을 구축하므로, 초분광영상과 분광 라이브러리 색인 과정에서 발생하는 차원의 저주 문제를 해결할 수 있다.

[0020] 또한, 본 발명은 미분 분광법을 이용한 전처리를 진행하므로 특정 밴드의 선별 정확도를 개선할 수 있다.

[0021] 또한, 본 발명은 피라미드 기법 기반의 K-최근접 질의를 통해 후보자를 검출하므로, 지정된 범위 외에도 검색이 가능하다.

[0022] 또한, 본 발명은 질의 처리 과정에서 적분 분광 정합 필터를 통해 거짓값을 필터링하여 물질탐지의 정확도를 개선할 수 있다.

도면의 간단한 설명

[0023] 도 1은 본 발명에 따른 피라미드 기법 기반 색인 및 질의를 통한 초분광영상의 물질탐지방법을 도시한 흐름도이다.

도 2는 도 1의 전처리 과정을 설명하기 위한 분광 반사 곡선과 1차 미분 분광 반사 곡선을 도시한 일 예이다.

도 3은 도 1의 전처리 과정을 설명하기 위한 2차 미분 분광 반사 곡선을 도시한 일 예이다.

도 4는 도 1의 피라미드 기법을 이용한 색인 구축 단계에서 피라미드 번호 선택 및 피라미드 내 높이 선택을 도시한 일 예이다.

도 5는 도 1의 질의 처리 과정에서 B+ 트리에서의 질의 영역을 확장하는 과정을 도시하는 일 예이다.

도 6은 본 발명과 관련된 상대적인 값으로 재생성된 콘크리트 적분 분광 반사 곡선을 도시한 그래프이다.

발명을 실시하기 위한 구체적인 내용

[0024] 이하, 첨부된 도면들을 참조하여 본 발명의 실시예를 상세하게 설명한다.

[0025] 도 1은 본 발명에 따른 피라미드 기법 기반 색인 및 질의를 통한 초분광영상의 물질탐지방법을 도시한 흐름도이다.

[0026] 도 1에 도시된 바와 같이, 본 발명에 따른 초분광영상의 물질탐지방법은 분광 라이브러리 색인 구축 과정(A)과

질의 처리 과정(B)으로 구분된다.

- [0027] 분광 라이브러리 색인 구축 과정(A)는 전처리 단계와 색인 구축 단계로 구분된다.
- [0028] 먼저, 미분 분광법을 이용하여 분광 라이브러리 내에 포함된 분광 데이터를 전처리한다(S11). 이러한 전처리 과정을 통해 분광 라이브러리 데이터(분광 데이터)들을 이전 밴드 반사율과 현재 밴드 반사율의 변화량으로 변환하므로, 보다 정확하게 특징 밴드를 선별할 수 있다.
- [0029] 피라미드 기법을 이용하여 상기 전처리된 분광 데이터에 대한 색인을 구축한다(S12). 피라미드 기법을 이용한 색인 구축은 분광 라이브러리에 포함된 각각의 분광 데이터를 1차원 데이터로 구성된 키(key) 값으로 변환한다. 그 키 값을 이용하여 상기 분광 데이터를 B+ 트리에 배치한다. 즉, 변환된 키 값을 이용하여 분광 라이브러리의 색인을 구축한다.
- [0030] 분광 라이브러리 색인 구축 후, 초분광영상에서 물질 데이터가 추출되면 다섯 단계의 질의 처리 과정(B)을 통해 상기 구축된 색인을 사용하여 분광 라이브러리로부터 후보자를 추출한다.
- [0031] 질의 처리 과정(B)의 첫 번째는, 초분광영상으로부터 추출된 물질 데이터로부터 키 값을 만든다(S21). 이때, 키 값은 색인 구축 과정과 동일한 방법으로 생성된다. 다시 말해서, 미분 분광법을 사용하여 상기 추출된 물질 데이터를 전처리하고, 그 전처리된 물질 데이터에 피라미드 기법을 적용하여 키 값을 생성한다.
- [0032] 상기 생성된 키 값을 이용하여 B+ 트리 내에서 상기 추출된 물질 데이터의 키 값과 유사한 값을 가지고 있는 질의점으로 이동한다(S22). 다시 말해서, 구축된 색인 내에 상기 추출된 물질 데이터의 키 값과 유사한 분광 데이터의 키 값이 존재하는 위치로 이동한다.
- [0033] 상기 질의점으로 이동 후 상기 질의점을 기준으로 소정 범위 내에서 특정 질의 알고리즘에 따라 질의를 수행한다(S23). 여기서, 특정 질의 알고리즘으로는 영역질의 기반의 K-최근접 질의 알고리즘이 사용된다. 이러한 질의 처리를 통해 분광 라이브러리로부터 후보자 물질을 추출한다.
- [0034] 상기 질의 알고리즘에 의해 후보자로 추출된 물질 중 다른 물질을 필터링한다(S24). 여기서, 물질 탐지 정확도를 높이기 위하여 질의 알고리즘을 통해 추출된 후보자를 다시 전체 차원을 비교하여 잘못 추출된 물질을 필터링한다. 여기서, 상기 잘못 추출된 물질을 필터링하기 위해 적분 분광 정합 필터가 사용된다.
- [0035] 상기 추출된 후보자 중 상기 필터링을 통해 제거되고 남은 최종 후보자를 결과로 출력한다(S25).

[0036] 상기한 분광 라이브러리 색인 구축 과정(A)을 좀더 상세하게 설명한다. 분광 라이브러리 색인 구축과정(A)은, 미분 분광법을 이용한 전처리 단계(S11)와 피라미드 기법을 이용한 색인 구축 단계(S12)로 구성된다.

[0037] 먼저, 전처리 단계에서는 미분 분광법을 이용하여 참조 데이터로 사용되는 분광 라이브러리 내에 포함된 물질들의 분광 데이터를 전처리한다. 상기 미분 분광법은 분광 데이터(반사율)를 밴드(파장 영역)간의 변화량으로 변환하는 기법으로, 1차 미분 분광법과 2차 미분 분광법이 있다.

[0038] 1차 미분 분광법은 반사율, 광휘도, 흡수율 등으로 만들어진 반사 곡선을 변화량으로 나타내는 미분 분광 반사 곡선으로 변환하는 방법이다. 즉, 1차 미분 분광법은 다양한 방법으로 만들어진 분광 반사 곡선을 이전 밴드과 현재 밴드의 반사율을 이용하여 이전 밴드와 현재 밴드의 반사율 변화량으로 나타내는 미분 분광 반사 곡선으로 변환시키는 것이다. 상기 미분 분광 반사 곡선은 절대 반사율의 차이가 제거되고, 이전 밴드에 대한 변화량을 갖게 된다. 분광 라이브러리 데이터를 1차 미분 데이터로 변환하기 위해 [수학식 1]을 사용한다.

수학식 1

[0039]
$$d_n^{1st} = (r_{n+1} - r_n) / (\lambda_{n+1} - \lambda_n)$$

[0040] 여기서, $n \in \{1, 2, \dots, b-1\}$ 이고, b 는 밴드 값(번호)이다.

[0041] r_{n+1} 과 r_n 은 다음 밴드와 현재 밴드의 반사율(reflectance)이고, λ_{n+1} 과 λ_n 는 다음 밴드와 현재 밴드의 파장(wavelength)이다. 1차 미분 데이터는 다음 밴드와 현재 밴드의 반사율 차를 다음 밴드와 현재 밴드의 파장 차로 나누는 것으로 구해진다. [수학식 1]을 이용하면 측정된 분광 데이터에서 다음 밴드와 현재 밴드의 반사율 변화량으로 변환되게 된다. 도 2의 (a)에 도시된 분광 반사 곡선을 1차 미분 분광 반사 곡선으로 변환하면 도 2의 (b)와 같다. 도 2(b)는 도 2(a)에 도시된 분광 반사 곡선에 대비하여 두 반사 곡선이 절대 반사율 차이가 없 어지고, 이전 밴드에 대한 변화량으로 변환되면서 거의 일치하게 된다.

[0042] 2차 미분 분광법은 측정된 분광 데이터에서 [수학식 1]을 이용하여 만들어진 변화량을 다시 한번 더 미분을 하여 변화량에 대한 변화량으로 변환하는 것이다. 2차 미분 데이터는 변화량에 대한 변화량으로 분광 반사 곡선이 만들어지기 때문에 특징 밴드 선택 시 변화량이 큰 밴드를 좀 더 확실하게 얻을 수 있다. 즉, 특징 밴드를 선별 의 정확성을 향상시킬 수 있다.

[0043] 2차 미분 데이터는 [수학식 2]를 통해 연산한다.

수학식 2

$$d_r^{2nd} = \frac{(d_{r+1}^{1st} - d_r^{1st})}{[0.5 \times (\lambda_{r+2} - \lambda_r)]}$$

[0045] 여기서, $n \in \{1, 2, \dots, b - 2\}$ 이다.

[0046] [수학식 2]는 [수학식 1]을 이용하여 데이터를 1차 미분 데이터로 변환한 후, 다음 밴드의 1차 미분 데이터와 현재 밴드의 1차 미분 데이터의 차를 현재 밴드에서 2번째 앞에 있는 밴드의 파장과 현재 밴드의 파장의 차에 0.5를 곱한 값으로 나누는 것으로 나타내고 있다. 여기서, 다음 밴드의 1차 미분 데이터와 현재 밴드의 1차 미 분 데이터의 차에 가중치를 주어 변화량을 좀 더 증가시키기 위해 상기 파장의 차에 0.5를 곱한다. 또한, 파장 값에서 2번째 앞에 있는 밴드의 파장이 사용되는 이유는 2차 미분은 현재와 다음밴드와 2번째 앞에 있는 밴드의 값을 이용하기 때문이다. [수학식 2]를 이용하여 도 2(a)에 대한 2차 미분 분광 반사 곡선은 도 3과 같다. 도 3에 도시된 2차 미분 분광 반사 곡선은 도 2(b)에 도시된 것과 같이 점선과 실선이 거의 동일하게 그려진 다는 것을 확인할 수 있다. 도 2(b)와 비교해보면 변화량에 대한 변화량으로 나타나기 때문에 이전 밴드와 현재 밴드 에 대한 변화량을 좀 더 확실히 나타나는 것을 확인할 수 있다.

[0047] 상기 미분 분광법에 의해 전처리된 분광 데이터로부터 피라미드 기법을 사용하여 색인을 구축한다. 피라미드 기 법은 차원의 저주 문제를 해결하기 위하여 D차원(다차원)의 데이터에서 특징을 나타내는 차원을 이용하여 1차원 의 값으로 변환시키고, B+ 트리를 이용하여 색인하는 방법이다.

[0048] 이와 같이, 피라미드 기법은 특징을 나타내는 1개의 차원(밴드)만을 이용하기 때문에 차원의 저주를 해결할 수 있다. 피라미드 기법을 이용하여 D차원의 데이터를 1차원 값으로 변환하는 과정은 두 단계로 이루어진다. 첫 번 째 단계는 D차원의 데이터를 2D개의 파라미드로 분할하고 분할된 피라미드 중 어느 하나의 피라미드 번호를 선 택한다. 예를 들어, 분광 데이터가 3개의 밴드로 구성되어 있다면 피라미드는 도 4의 (a)에 도시된 바와 같이 6 개가 된다. 두 번째 단계는 도 4의 (b)에 도시된 바와 같이 선택된 피라미드에서 데이터가 배치될 높이를 정하 고, 여러 개의 조각들로 분할한다. 이 과정에서 분할된 조각은 B+ 트리에서 하나의 데이터 페이지로 사용하게 된다.

[0049] D차원의 데이터 v 가 주어졌을 때 피라미드 기법을 이용하여 1차원의 데이터로 변환하는 과정은 총 세 단계로 이 루어진다. 첫 번째 단계는 키 값이 배치될 피라미드 번호를 선택하는 단계이다. 먼저, D차원의 데이터를 2D개의 피라미드로 분리한다. 피라미드를 분리한 후 [수학식 3]을 이용하여 피라미드 번호를 선택한다.

수학식 3

$$i = \begin{cases} j_{max} & v_{j_{max}} < 0.5 \\ j_{max} + d & v_{j_{max}} \geq 0.5 \end{cases}$$

[0050]

[0051] 여기서, j_{max} 는 중심점 0.5에서 거리가 제일 먼 반사율을 가지고 있는 차원의 번호이며, $v_{j_{max}}$ 는 제일 먼 반사율을 가지고 있는 차원의 반사율이다. $j_{max} + d$ 는 선택된 차원의 반사율이 0.5이상일 때 선택된 차원에서 전체 차원을 더한다. [수학식 3]을 이용하여 특징을 나타내는 차원을 선택하는 방법은 다음과 같다. 각 차원의 반사율 중에서 피라미드의 중심점인 0.5를 기준으로 거리가 제일 먼 값을 갖고 있는 차원을 피라미드의 번호로 선택한다. 만약 선택된 차원의 값이 0.5이상의 값을 갖는다면 선택된 차원에 전체 차원을 더하여 피라미드 번호로 선택하게 된다.

[0052] 두 번째 단계는 선택된 피라미드에서 데이터(전처리된 분광 데이터)가 배치될 높이를 정하는 단계이다. 데이터가 배치될 높이는 [수학식 4]를 이용하여 구한다.

수학식 4

$$h_v = |0.5 - v_{iMODd}|$$

[0053]

[0054] 여기서, v_{iMODd} 는 [수학식 3]에서 0.5 이상의 반사율을 갖고 있는 차원을 원래의 차원으로 되돌리기 위한 것이다. 데이터의 중심점인 0.5를 기준으로 선택된 피라미드의 반사율의 절대적인 차 값을 이용하여 높이가 정해진다. 즉, 0.5와 반사율의 거리로 높이가 정해지게 된다.

[0055] 마지막 세 번째 단계는 [수학식 3]과 [수학식 4]를 이용하여 색인에 사용될 키 값을 구한다. $B^+ - Tree$ 에서 사용될 키 값은 다음 [수학식 5]를 이용하여 구한다.

수학식 5

$$PV_v = i + h_v$$

[0056]

[0057] $B^+ - Tree$ 의 키 값은 [수학식 3]을 이용하여 선택된 피라미드 번호와 [수학식 4]를 이용하여 구해진 높이의 합으로 만들어지게 된다. 즉, 키 값은 피라미드 번호와 높이로 구해지게 된다. 마지막으로 [수학식 5]를 이용해 만들어진 키 값은 B^+ 트리를 사용하여 검색, 추가, 삭제가 이루어지게 된다.

[0058] 이하, 도 1의 질의 처리 과정(B)을 좀더 상세하게 설명한다.

[0059] 본 발명에서는 각 차원의 최소값과 최대값으로 지정된 범위만 질의 수행이 가능한 피라미드 기법을 이용한 영역 질의 알고리즘의 단점을 해결하기 위해, 피라미드 기법을 이용한 영역질의 기반의 K-최근접 질의 알고리즘을 제안한다.

[0060] 종래의 피라미드 기법을 이용한 영역질의 알고리즘은 물질이라도 측정할 당시의 환경에 의한 오차로 인해 다른 차원에서 키 값이 생성되거나, 다른 페이지에 데이터가 배치되었을 때 후보자 추출이 불가능 할 수 있다.

[0061] 그러나, 본 발명에서 제안하는 질의 알고리즘은 초기에 지정된 범위 외 다른 차원 혹은 다른 페이지도 검색이 가능하다.

[0062] 또한, 서로 다른 물질이 같은 키 값을 갖고, 후보자로 추출되는 문제가 발생할 수 있다. 즉, 다른 물질임에도

불구하고, 특징을 나타내는 밴드의 값이 동일하고, 반사율이 비슷하여 같은 키 값을 갖는 물질이 있을 수 있기 때문이다. 따라서 본 발명에서는 참 값과 거짓 값을 분류할 수 있는 적분 정합 필터를 추가하여 후보자를 추출할 때마다 다른 물질을 필터링하도록 한다.

[0063] 분광 데이터의 후보자 추출에는 노이즈와 오차를 고려한 질의 영역이 필요하다. 따라서, 본 발명의 K-최근접 질의 알고리즘은 피라미드 영역이 아닌 B+ 트리에서 질의 영역을 결정하도록 하였다. 또한, K개의 후보자를 추출함으로써 과도한 후보자 추출을 방지하였고, 반대로 후보자가 추출되지 않을 경우를 대비하여 K개에 따라 질의 영역을 증가시킴으로써 다른 차원 혹은 다른 리프 노드에 배치되어 있는 데이터에도 접근이 가능하게 한다. B+ 트리의 리프 노드들은 링크리스트로 연결되어 있기 때문에 B+ 트리에서의 질의 영역의 확장은 리프 노드에 존재하는 모든 데이터에 접근이 가능하여, 노이즈와 오차 영향을 적게 받는다.

[0064] 그리고, 후보자 추출 시 모든 차원의 데이터와 비교가 가능하도록 B+ 트리의 리프 노드를 키 슬롯과 데이터 슬롯으로 분리한다. 상기 키 슬롯에는 분광 라이브러리의 대표 값인 키 값을, 데이터 슬롯에는 분광 라이브러리의 데이터에 접근이 가능한 포인터를 입력시킨다.

[0065] 상기 질의 처리 과정은 다섯 단계로 이루어진다. 첫 번째 단계는 피라미드 기법에 의해 생성된 초분광영상에서 추출된 물질 데이터의 키 값을 이용하여 질의 기준위치가 되는 B+ 트리의 리프 페이지로 이동한다. 상기 키 값에 따라 B+ 트리의 리프 페이지로 이동한다.

[0066] 두 번째 단계는 리프 페이지에서 키 값을 중심으로 임의의 범위 $\pm\Delta r$ 만큼 질의 영역을 확장한다. 이렇게 결정된 질의 영역은 후보자 추출 범위로 사용되게 된다.

[0067] 세 번째 단계는 상기 데이터 슬롯의 포인터에 링크된 분광 라이브러리의 데이터들로부터 적분 분광 정합 필터를 이용하여 k개의 후보자를 추출한다. 두 번째 단계에서 선택된 질의 영역 내에 다른 물질이 존재 할 수도 있기 때문에 적분 분광 정합 필터를 이용하여 다른 물질을 걸러낸다.

[0068] 네 번째 단계는 질의 알고리즘을 수행한 후 후보자의 개수가 k개가 안됐을 경우 도 5에 도시된 바와 같이 Δr 만큼 질의 영역을 확장시켜 재검색을 한다. 재검색 후, 후보자의 개수가 k개가 되지 않으면, k개가 될 때까지 범위를 증가시켜 재검색하는 과정을 반복 수행한다. 하지만 후보자의 개수가 k개가 될 때까지 전체 데이터를 비교해 볼 수는 없다. 범위가 계속 증가하여 전체 데이터를 비교하게 되면 결국 순차적 접근 방법과 같아지고, 범위 증가 계산으로 인해 속도가 더 느려질 수 있다. 따라서, 이런 문제를 방지하기 위해 최대 범위를 초기 범위의 ± 5 밴드로 제한하였다.

[0069] 다섯 번째 단계는 최종적으로 추출된 k개의 후보자들의 정합율을 이용하여 순위를 계산하고 결과로 출력한다.

[0070] 앞서 설명한 바와 같이, 영역질의 기반의 K-최근접 질의 알고리즘의 세 번째 단계에서는 초분광영상에서 물질 데이터가 주어졌을 때 질의 처리를 통해 추출된 후보자 중 거짓 값을 걸러내기 위해 적분 분광 정합 필터를 이용하여 필터링을 수행한다.

[0071] 본 발명에서 제안하는 적분 분광 정합 필터를 이용한 필터링은 세 단계로 이루어지게 된다.

[0072] 첫 번째 단계는 [수학식 1]을 이용하여 만들어진 1차 미분 데이터를 [수학식 6]을 이용하여 물질의 상대적인 값으로 재생성한다.

수학식 6

$$I_z = \sum_{n=1}^x d_n^{1st}$$

[0073]

[0074] 여기서, b는 밴드 번호 값이고, $z \in \{1, 2, \dots, b-2\}$ 이고, $x = b-2$ 이다. [수학식 6]을 이용하여 물질의 상대적인 값으로 재생성한 적분 분광 반사 곡선은 도 6과 같다.

[0075] 도 6과 도 2(a)를 비교하면 도 6의 두 분광 반사 곡선이 거의 일치한다는 것을 확인할 수 있다. 그 이유는 1차 미분 데이터를 이용하여 절대 값의 차가 제거되고, 변화량에 따른 상대적인 값으로 재생성했기 때문이다. 다음은 [수학식 6]을 이용하여 재생성된 적분 데이터를 각각의 밴드 값에 대한 정합율을 구한다. 그 정합율은 [수학식 7]과 같이 정의한다.

수학식 7

[0076]
$$d_z^m = I_z^d / I_z^L$$

[0077] 여기서, I_z^d 는 입력된 초분광영상의 물질데이터의 적분값을 의미하며, I_z^L 는 후보자로 추출된 분광라이브러리의 적분값을 의미한다. 후보자로 추출된 분광 라이브러리와 입력된 초분광영상 데이터가 각각의 밴드 값에 따라 얼마나 같은 반사율 값을 갖고 있는지를 구할 수 있다. 이렇게 계산된 정합율에 [수학식 8]을 적용하여 동일한 밴드와 다른 밴드로 분류하게 된다.

수학식 8

[0078]
$$t_z = \begin{cases} 1, & 0.7 \leq d_z^m \leq 1.3 \\ 0, & otherwise \end{cases}$$

[0079] 구해진 각각의 밴드 값에 대한 정합율이 0.7에서 1.3 사이의 값을 갖는다면, 동일한 밴드로 분류하고, 그렇지 않다면 다른 밴드로 분류한다. 분류된 밴드를 [수학식 9]를 이용하여 총 정합율을 계산한다.

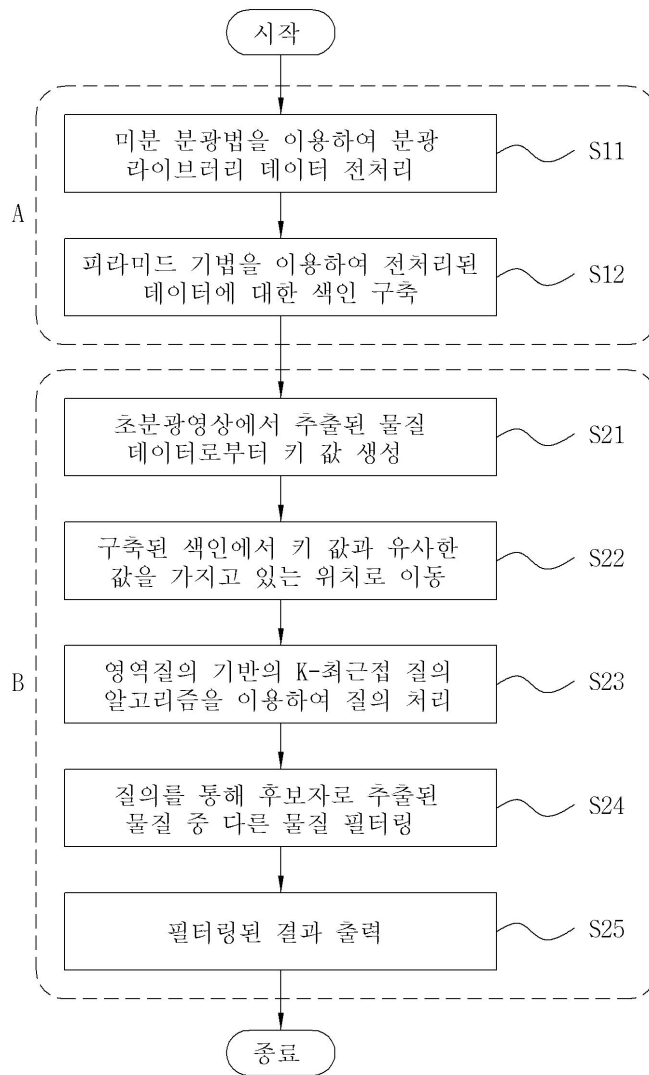
수학식 9

[0080]
$$g = \left(\sum_{z=1}^x t_z \right) / x$$

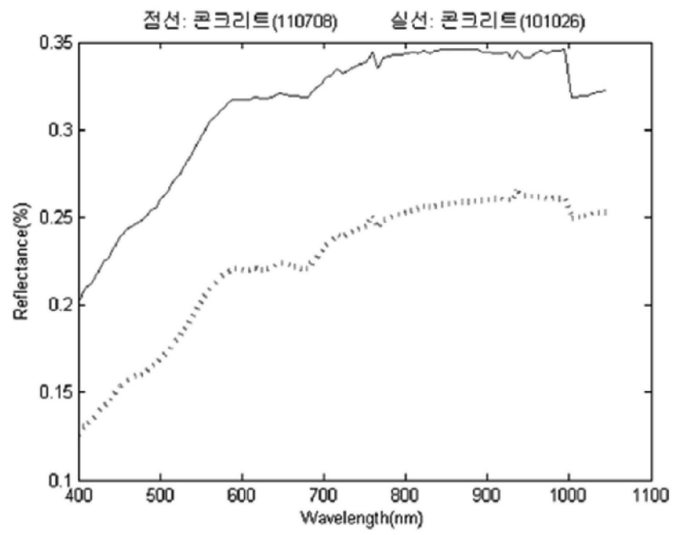
[0081] 총 정합율은 [수학식 8]을 이용하여 구해진 t_z 의 총 합을 x 로 나누어 계산할 수 있다. 즉, 전체 밴드 중에 몇 개의 밴드가 동일 하느냐를 계산하는 식이다. 마지막으로 이렇게 구해진 g 가 70%이상의 값을 갖는다면, 추출된 후보자와 입력된 초분광영상 물질 데이터는 같은 데이터라 판단하고, K-최근접 질의의 최종 후보자로 추가한다.

도면

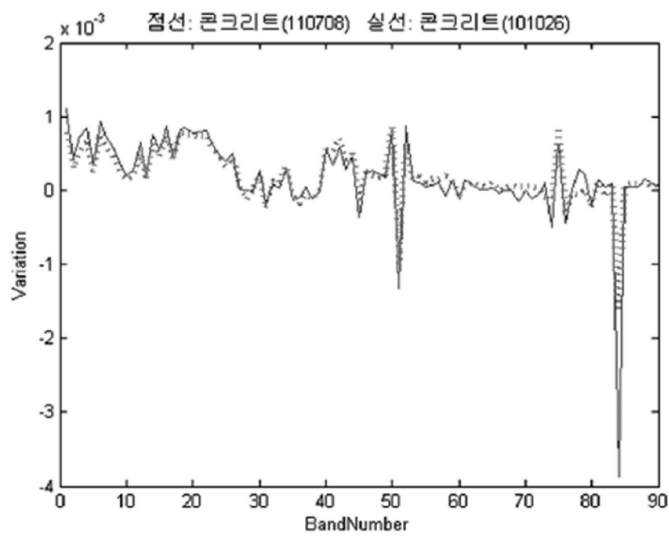
도면1



도면2

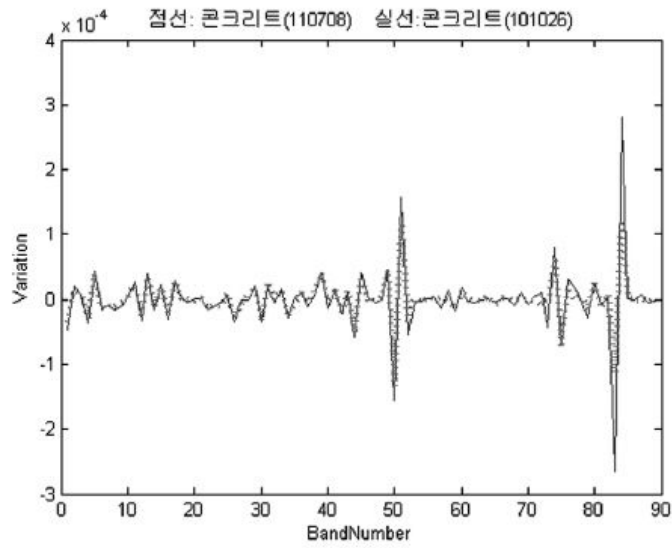


(a)

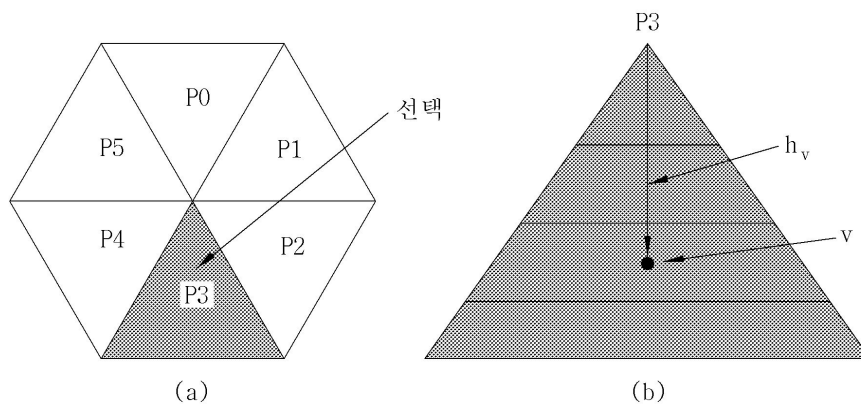


(b)

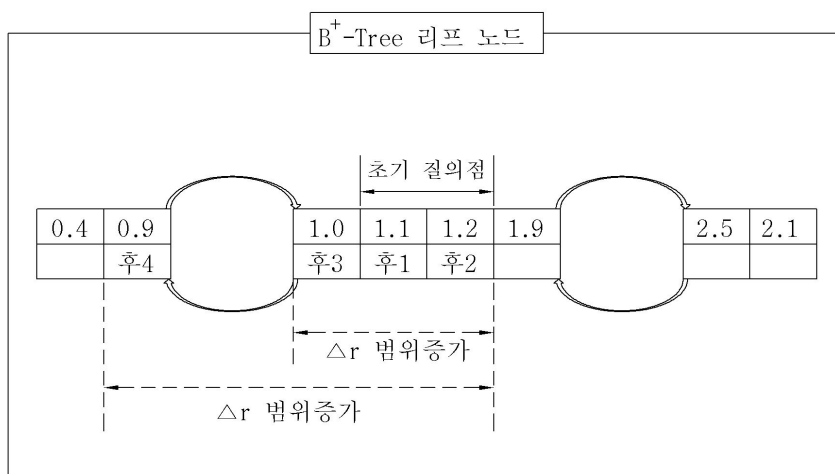
도면3



도면4



도면5



도면6

