

초분광 영상 라이브러리에서 피라미드 색인 기법의 영역 질의를 이용한 스펙트럴 매칭

Spectral matching using Range Queries based on Pyramid-Technique in Hyperspectral Image Library

유재환, 김덕환*

인하대학교 전자공학과

Jae-Hwan Yu, Deok-Hwan Kim*

Dept. Electronic Engineering, Inha University*

요약

초분광 영상은 기존의 다중분광 영상보다 많은 밴드를 통해 넓은 범위의 파장 영역에 대한 반사율을 담고 있는 고차원 데이터이다. 이와 같은 고차원 데이터를 기존의 R-Tree, X-Tree와 같은 다차원 색인 방법 사용하게 되면 차원의 저주(Curse of Dimensionality)라는 문제가 발생한다. 본 논문에서는 차원의 저주 문제를 해결하기 위해 피라미드 기법을 사용하여 초분광 영상 라이브러리의 색인을 구축하였다. 피라미드 기법은 D차원의 데이터를 2D차원의 피라미드에 사상하고, B+-트리를 이용하여 1차원적으로 색인하는 방법이다. 실험 결과 스펙트럴 매칭을 위한 영역질의 방법이 후보자 추출 시간, 데이터 접근 빈도 측면에서 순차적 접근 방법보다 좋은 성능을 나타냈다.

1. 서론

최근 초분광 영상(Hyperspectral Image)을 이용한 유효밴드 추출, 전처리, 목표탐지 및 물질 분류 등의 영상 처리 연구가 활발히 진행되고 있다.

초분광 영상은 기존의 다중분광 영상보다 많은 밴드를 통해 넓은 범위의 파장 영역에 대한 반사율을 담고 있는 고차원 데이터이다. 이와 같은 고차원의 데이터를 기존의 T-Tree, R-Tree, X-tree와 같은 다차원 색인 방법을 사용하게 되면 차원의 저주(Curse of Dimensionality)라는 문제가 발생하게 된다. 차원의 저주란 차원의 수가 증가함에 따라 검색시간이나 기억공간에 대한 요구가 지수 승으로 증가하여 처리 시간이 순차적 접근 방법보다 오래 걸리게 되는 현상이다. 차원의 저주를 해결하기 위한 방법으로 다차원 데이터를 1차원의 데이터로 변환하는 피라미드 기법(Pyramid-technique)이 제안되었다[1].

본 논문에서는 초분광 영상 라이브러리 색인 구축 과정에서 발생하는 차원의 저주를 해결하기 위하여 피라미드 기법을 사용하여 색인을 구축하였다. 스펙트럴 매칭을 위한 영역 질의와 영역질의 기반의 K-최근접 질의를 이용하여 순차적 접근 방법과 비교하는 실험을 진행 하

였다.

2. 피라미드 기법

피라미드 기법은 D차원의 데이터를 1차원의 값으로 변환하고, B+-트리를 이용하여 색인하는 방법이다. 피라미드 기법을 이용하여 1차원 값으로 변환하는 과정은 두 단계로 이루어진다. 첫 단계는 D차원의 데이터를 2D개의 피라미드로 분할한다. 여기서 각각의 피라미드는 데이터의 중심점(0.5)을 기준으로 2D개의 피라미드로 분할된다. 두 번째 단계는 2D개의 피라미드를 다시 여러 개의 조각들로 분할한다. 분할된 조각은 B+-트리에서 하나의 데이터 페이지가 된다.

D차원의 데이터 v 가 주어지면 세 단계를 걸쳐 1차원의 데이터로 변환 된다.

$$i = \begin{cases} j_{\max} & \text{if } (v_{j_{\max}} < 0.5) \\ (j_{\max} + d) & \text{if } (v_{j_{\max}} \geq 0.5) \end{cases} \quad (1)$$

첫 단계는 식1을 이용하여 v 가 속해있는 피라미드 번호를 구한다. 각각의 차원 값을 비교하여 중심점(0.5)에서 가장 멀리 떨어져 있는 값이 속해있는 차원이 피라미드 번호 P_i 로 선택된다.

$$h_v = |0.5 - v_{i \text{ MOD } d}| \quad (2)$$

* 교신저자: deokhwan@inha.ac.kr

- 본 연구는 한국과학기술원 영상정보특화연구센터를 통한 방위사업청과 국방과학연구소의 연구비 지원으로 수행됨 (계약번호 UD100006CD)
- 본 과제(연구)는 지식경제부와 한국산업기술진흥원의 전략기술인력양성사업으로 수행된 결과임

두 번째는 피라미드에서 v 의 높이를 구하는 단계다. 식 2를 이용하여 중심점(0.5)과 선택된 차원의 값 v_i 의 차를 이용하여 높이를 구한다.

$$PV_v = (i + h_v) \quad (3)$$

마지막으로 식 3을 이용하여 1차원 값인 PV_v 를 구한다. PV_v 는 B^+ -트리의 키 값으로 사용된다.

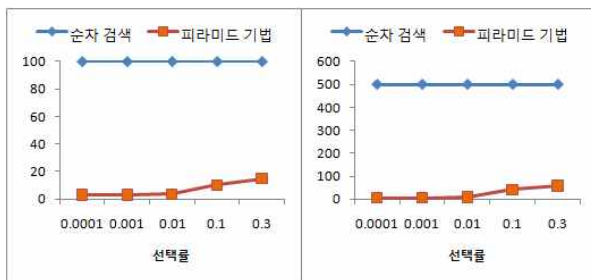
3. 질의 처리

피라미드 기법에서 널리 사용되는 질의로는 영역질의와 영역질의 기반의 K -최근접 질의가 사용된다. 영역질의는 질의 점 Q 와 거리 R 이 주어졌을 때, 질의 점을 기준으로 R 만큼의 거리 안에 존재하는 후보 값들을 모두 추출 방법이다.

K -최근접 질의는 영역 질의와 같은 방식으로 영역안의 후보 값들을 K 개를 찾아내는 방법이다. 만약 영역 안의 후보 값들이 K 개가 되지 않는다면 영역을 증가시켜 K 개가 될 때까지 후보자를 찾아낸다.

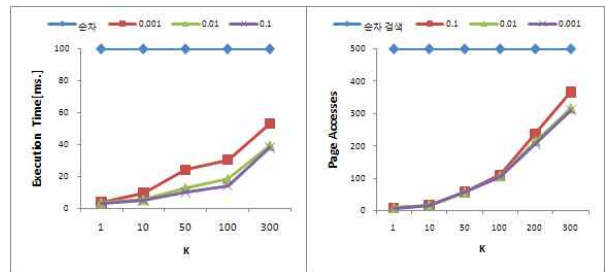
4. 실험

실험에는 초분광 영상 측정기 CASI-1500과 지상 분광 측정기 FieldSpec-3을 이용하여 350nm부터 2500nm의 영역에서 1nm씩 총 2150밴드를 가지고 있는 라이브러리 데이터 500개가 사용되었다. 스펙트럴 매칭을 위해 순차적 접근 방법과 피라미드 기법을 이용한 영역질의, K -최근접 질의의 검색 시간, 접근 빈도수를 각각 비교하였다.



▶▶ 그림 1. 영역질의 성능비교

그림 1은 순차적 접근 방법과 피라미드 기법을 이용한 영역질의의 검색시간과 접근 빈도수를 각각 보여준다. 영역질의는 범위가 증가할수록 검색시간과, 접근 빈도수가 늘어나는 것을 볼 수 있다. 그러나 스펙트럴 매칭실험에서 낮은 성능을 보이는 범위 0.3이라도 순차적 접근 방법보다 검색 시간은 85%, 접근 빈도수는 약 88% 줄어 피라미드 기법을 이용한 영역질의가 순차적 접근 방법보다 성능이 좋음을 알 수 있다.



▶▶ 그림 2. K -최근접 질의 성능 비교

그림 2는 순차적 접근 방법과 피라미드 기법을 이용한 K -최근접 질의의 검색시간과 접근빈도수를 각각 보여준다. K -최근접 질의는 K 개가 증가할수록 검색 시간과 접근 빈도수가 증가하였다. 그러나 스펙트럴 매칭실험에서 순차적 접근 방법과 영역 0.001, K 가 300이었을 때 검색 시간은 47%, 접근 빈도수는 27% 감소하였다.

5. 결론 및 향후 연구방향

본 논문에서는 피라미드 기법을 이용하여 초분광 영상 라이브러리의 색인을 구축하였고, 영역질의와 영역질의 기반의 K -최근접 질의를 이용하여 검색 시간과 접근 빈도 횟수를 측정하는 실험을 진행 하였다. 실험 결과 순차적 접근 방법에 비해 피라미드 기법이 좋은 성능을 보인다는 것을 확인할 수 있었다.

향후 연구로는 초분광 영상의 정확한 색인을 위하여 보다 정확한 특징을 나타내는 차원을 키 값으로 갖는 빠른 색인 알고리즘을 개발하고자 한다.

■ 참고 문헌 ■

- [1] Berchtold, S., Böhm, C., Kriegel, H. "The Pyramid-Technique: Towards Breaking the Curse of Dimensionality" In SIGMOD Conference pp.142-153
- [2] Jia Li, Cheng Wnag, "Indexing Method for Hyperspectral Data Fast Retrieval by Pyramid Technique" CSSE Proceedings of the 2008 International Conference- Volume 04 pp.519-523