

계층적 3차원 컨볼루션 희소코딩을 이용한 확률적 볼륨렌더링*

최준영⁰, Tran Minh Quan, 정해진, 정원기
전기전자컴퓨터공학부, 울산과학기술원
(juny0603, quantm, goodhen2, wkjeong)@unist.ac.kr

Probabilistic Volume Rendering using Hierarchical 3D Convolutional Sparse Coding

JunYoung Choi⁰, Tran Minh Quan, Haejin Jeong, Won-Ki Jeong
School of Electrical and Computer Engineering, Ulsan National Institute of Science and Technology

요약

본 논문은 [1]에서 소개한 고차원 특성벡터 기반 볼륨렌더링 시스템에 계층적 3차원 컨볼루션 희소코딩(Hierarchical 3D Convolutional Sparse Coding: H-CSC)과 확률적 다중전이함수를 적용하여 시각화 품질을 크게 향상시키는 방법을 소개한다. 또한 [2], [3]의 방법과 제안한 방법을 비교하는 유저스터디를 진행하여, 제안한 방법의 사용성과 시각화 품질을 검증하였다.

1. 서론

3차원 볼륨 데이터의 효과적인 시각화를 위해서 다양한 전이함수에 대한 연구가 진행되었다. 최근 [1]에서는 인공지능의 자율학습 방법(3D CSC[4])과 지도학습 방법(인공신경망)을 활용한 효과적인 볼륨렌더링 방법이 소개되었다. [1]에서 제안한 방법은 입력 볼륨으로 학습한 3차원 CSC 기반의 고차원 특성벡터와 사용자 입력을 이용하여 볼륨의 구조를 자동으로 분리할 수 있는 장점이 있었다. 반면에 고정된 크기의 필터들을 사용하였기 때문에 각각의 원소가 정해진 크기의 지역만큼의 정보만 학습할 수 있어 주어진 데이터의 정보를 충분히 담지 못하는 문제가 있었다. 본 논문에서는 계층적 3차원 CSC를 적용하여 이전의 한계를 극복하고 주목할 만한 성능 향상을 이루었다.

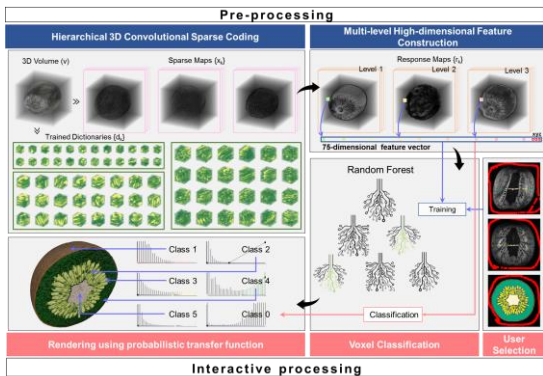


그림 1: 계층적 3차원 CSC를 이용한 확률적 볼륨렌더링 개요도

2. 계층적 3차원 CSC를 이용한 확률적 볼륨렌더링

2.1. 계층적 3차원 컨볼루션 희소코딩

계층적 3차원 CSC의 핵심 아이디어는 CSC에서 학습하는 지역 정보가 필터의 크기로 정의된다는 것이다. 본 논문에서는 세 종류의 필터 크기($7^3, 15^3, 31^3$)를 사용하였으며, 그림 2는 알파벳 데이터에서 다중 스케일로 추출한 필터를 나타낸다. 가장 작은 크기의 (a)에서는 알파벳의 모서리와 같은 지역적 특징을 학습하였고, 중간 크기의 (b)에서는 알파벳의 일부분을 학습하였다. 가장 큰 (c)는 알파벳 전체 형태를 학습할 수 있었다. 즉, 다양한 크기의 필터를 사용하여 사전을 구성하면 주어진 데이터에서 훨씬 더 많은 정보를 학습할 수 있고, 이는 결과적으로 복셀의 분류 정확도의 향상을 가져온다.

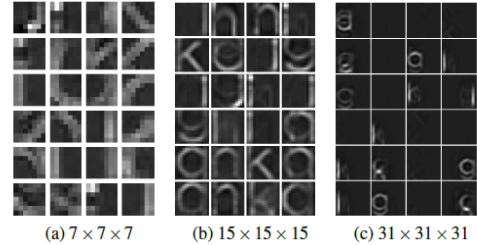


그림 2: Alphabet 데이터에서 추출한 멀티 스케일 필터

또다른 중요한 아이디어는 서로 다른 크기의 원소간에 종속성을 만드는 것이다. 이 개념은 이전 단계에서의 뉴런 활성화가 다음 단계의 컨볼루션 레이어에 대한 입력으로 사용되는 컨볼루션 뉴럴 네트워크(CNN)의 구조와 유사하다.

이 두 아이디어를 통해 계층적 3차원 CSC를 제안한다. 세 종류의 크기($7^3, 15^3, 31^3$)의 사전(d_k^1, d_k^2, d_k^3)을 사용하고, 이전 단계의 사전을 통해 복원한 볼륨은 다음 단계의 입력으로 주어진다. 이는 아래 식과 같다.

$$v^1 \approx \sum_{k=1}^{24} d_k^1 * x_k^1 = \sum_{k=1}^{24} r_k^1 = a^1, \quad v^2 = a^1$$

$$v^2 \approx \sum_{k=1}^{24} d_k^2 * x_k^2 = \sum_{k=1}^{24} r_k^2 = a^2, \quad v^3 = a^2$$

$$v^3 \approx \sum_{k=1}^{24} d_k^3 * x_k^3 = \sum_{k=1}^{24} r_k^3 = a^3$$

이렇게 추출해낸 계층적 특성을 통해 각 복셀 당 계층적 고차원 특성 벡터를 구성하게 된다.

* 구두발표논문
* 본 논문은 요약논문 (Extended Abstract) 으로서, 본 논문의 원본 논문은 IEEE Visualization Conference 2017에 발표될 예정이다
* 본 논문은 2016년도, 2017년도 정부(교육부)의 재원으로 한국과학창의재단(2016년도, 2017년도 학부생 연구프로그램)의 지원을 받아 수행된 연구임

2.2. 확률적 전이함수

[1]에서 복셀 분류 후에 데이터 값에 기반한 1차원 전이함수를 통해 각 복셀에 시각적 정보를 부여하던 방식과 달리, 본 방법은 인공지능 분류기가 부여한 복셀이 각 라벨에 속할 확률을 기반으로 시각적 정보를 부여하는 확률적 전이함수를 제안한다. 사용자는 각 라벨에 속할 확률(0~1)을 시각적 정보(색상, 투명도)로 매핑시키는 확률적 전이함수를 통해 각 라벨의 확률을 기반으로 색상과 투명도를 지정 ($TableColor(n, v, Prob_n)$, $TableAlpha(n, v, Prob_n)$)할 수 있고, 각 복셀 (v)이 가지는 확률에 따라 이들 정보가 합쳐지게 되어 최종 색상($v.Color$)과 투명도($v.Alpha$)를 아래 식과 같이 결정한다.

$$v.Color = \sum_n TableColor(n, v, Prob_n) \times v.Prob_n$$

$$v.Alpha = \sum_n TableAlpha(n, v, Prob_n) \times v.Prob_n$$

최종적으로 이 색상과 투명도들을 관선 투사법을 통해 투사하여 최종 시각화 결과를 만들어 낸다. 이 방법은 각 라벨에 속할 확률을 통해 시각적 정보를 보간함으로써 블록화 현상과 같은 문제를 해결할 수 있다.

3. 사용자 평가

본 방법의 사용성과 시각화 품질을 검증하기 위해서 30명의 참가자에게 기존의 방법([2], [3])과 본 방법을 주어진 시나리오에 따라 사용하게 하고, 과정과 결과를 분석하였다.

3.1. 실험 설계

참가자들은 [2]의 3차원 전이함수(P1)와 페인팅 기반의 볼륨렌더링(P2)을 사용하게 되며, 3가지 태스크(T1, T2, T3)를 진행하게 된다. T1과 T2는 알파벳 a와 b가 포함되어 있는 데이터이며, 모범답안과 복셀이 99% 이상 일치하게 시각화 하여야 하고 소요시간이 측정된다. T3는 CT 촬영을 통해 얻은 치아 데이터이며, 주어진 시간(10분) 동안 모범답안과 최대한 유사하게 시각화 하여야 한다.



그림 3: 왼쪽부터 T1, T2, T3에 대한 입력, 모범 답안

3.2. 실험 결과

P1을 사용했을 때, T1, T2에 대한 평균 소요시간은 63.8초, 129.1초 이고, P2를 사용했을 때, T1, T2에 대한 평균 소요시간은 78.7초, 71.2초로 나타났다. 이 결과를 통해 P1은 태스크의 난이도가 높아짐에 따라 시간이 더 오래 걸렸지만, P2는 태스크의 난이도가 높아져도 일정 시간 안에 주어진 태스크를 완료할 수 있음을 알 수 있다.

T3에 대해 [2], [3] 방법과 본 논문의 방법의 정확도의 평균은 0.450, 0.737, 0.841로 나타났으며, 일원분산분석과 Fisher's LSD 사후 검증 방법을 통해 본 논문의 방법이 [2], [3]과 의미 있는 차이가 있음을 확인했으며, 정확도가 가장 높게 나타남을 통해, 본 논문 방법의 성능을 검증 하였다.

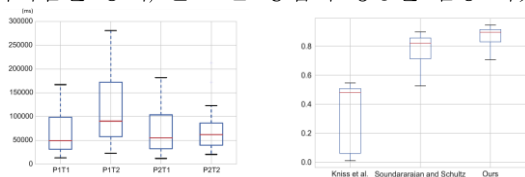


그림 4: 태스크 1, 2에 대한 프로그램 1, 2의 걸린 시간(왼쪽), 태스크 3에 대한 [2], [3], 본 논문의 방법의 정확도(오른쪽)

4. 시각화 결과

다양한 데이터에 대해 [2], [3]의 방법과 본 논문의 방법의 시각화 결과를 비교하여 그림 5에 나타내었다. 본 논문의 방법은 [2], [3]의 방법에 비해 효과적인 노이즈의 감소를 보여주었으며, 데이터의 각 구조들을 더 정확하게 분리하였다.

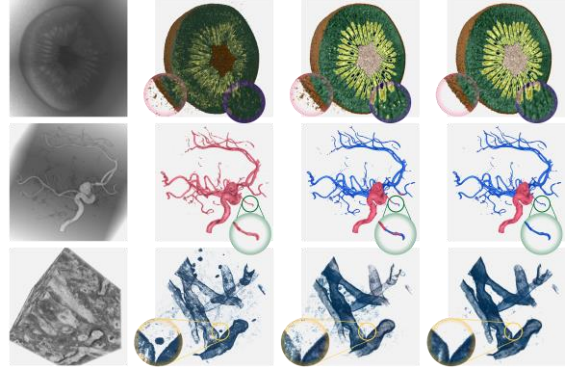


그림 5: 왼쪽부터 입력 데이터, [2]방법, [3]방법, 본 논문의 방법의 시각화 결과. 위에서부터 256x256x256 크기의 MRI-Kiwi, MRT-Aneurysm, EM-Mouse 데이터

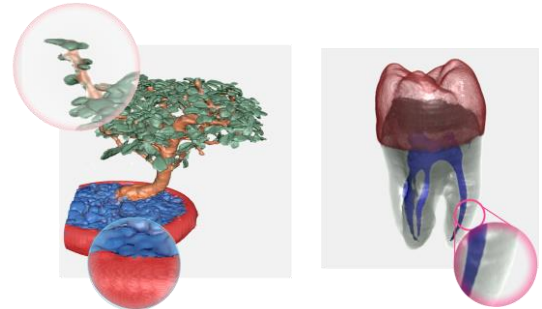


그림 6: CT-Bonsai(왼쪽), CT-Tooth(오른쪽) 데이터의 본 논문의 방법을 사용하여 시각화한 결과

5. 결론

본 논문에서는 고품질의 시각화 결과를 얻기 위해 [1]의 시스템을 계층적 3차원 CSC로 확장하고, 확률적 전이함수를 도입하였으며, 유저스터디를 통해 성능을 검증하였다.

향후 더 큰 볼륨데이터를 처리하기 위해 이 방법을 대규모 병렬 컴퓨팅 시스템으로 확장할 것이며, 매쉬 분류, 표면 재구성과 같은 계층적 3차원 CSC를 통해 성능 향상을 기대할 수 있는 다양한 그래픽스 분야에 확대 적용할 계획이다.

참고문헌

- [1] 최준영, Tran Minh Quan, 정원기. "고차원 특성을 활용한 확률적 볼륨 렌더링." *한국컴퓨터그래픽스학회 학술대회*, (2016.7): 66-67.
- [2] J. Kniss, G. Kindlmann, and C. Hansen. Multidimensional transfer functions for interactive volume rendering. *IEEE Transactions on Visualization and Computer Graphics*, 8(3):270-285, July 2002.
- [3] K. P. Soundararajan and T. Schultz. Learning Probabilistic Transfer Functions: A Comparative Study of Classifiers. *Computer Graphics Forum*, 34(3):111-120, June 2015.
- [4] T. M. Quan and W.-K. Jeong. Compressed sensing dynamic MRI reconstruction using GPU-accelerated 3D convolutional sparse coding. *In Proceedings of Medical image computing and computer-assisted intervention (MICCAI)*, pp. 484-492, 2016.