Mesh CNN vs Mesh R CNN vs 3D Shape Segmentation

Mesh CNN

Data Processing

1. Converting to .obj format(obj형식으로 변환)

파일 형식이 다른 3D 객체가 있으면 Meshlab을 사용하여 쉽게 변환

Meshlab은 모든 표준 파일 형식(.off, .obj, .ply…)을 지원

cmd에 meshlabserver -i /path/to/model.off -o /path/to/model.obj로 변환가능

1. Simplifying to the same number of faces(동일한 수의 면으로 단순화)

모든 mesh를 대략 같은 수의 면들로 단순화 할 것을 제안

이미지 CNN을 훈련하기 전에 이미지의 크기를 같은 크기로 조정하는 것, 네트워크를 더 빨리 만들고 매개변수를 덜 필요하게 할 것이다.

1. Non-manifold inputs(많지 않은 인풋)

MeshCNN은 여러가지 삼각망사를 지원

입력으로 ‘나쁜’삼각망사를 사용하고 여러가지 빈틈없는 Mesh를 아웃풋으로 사용이르게 하다.

Segmentation

1. Ground-Truth Segmentation Files

분할 교육에 사용되는 텍스트 파일은 \*.seg(에지 분할 ID)와 \*.seseg(소프트 에지 분할 ID) 두 가지가 있다.

모든 메쉬는 동일한 이름을 가진 해당 eseg 및 seseg 파일을 가져야 한다(예: 01.obj는 01.eseg 및 01.eseg를 가지고 있다).

1. Preparing a Different Segmentation Dataset

각 메쉬는 대략 동일한 수의 가장자리로 단순화한 후, 원래의 고해상도 분할 파일에서 저해상도 메쉬(eseg)를 위한 새로운 지면-투명 분할 파일이 생성된다.

seseg는 새로운 eseg 파일에서 생성할 수 있다.

Code Options

모든 일반 옵션은 base\_options.py에 정의되어 있다.

훈련 전용 옵션이 train\_options.py에 있고 추론(테스트) 옵션이 test\_options.py에 있다.

Misc

회귀 작업에 MeshCNN을 사용하려면 분할 네트워크 끝에 불필요한 ReLU가 있다는 점에 유의하십시오.

이는 분할 작업에 큰 차이는 없어 보이지만 제거하지 않으면 회귀(양수 값만 예측)의 성능에 타격을 줄 것이다.

Mesh R CNN

https://arxiv.org/pdf/1906.02739.pdf

3D Shape Segmentation

기존의 CNN은 학습된 convolution 필터, pooling작용 및 활성화 기능을 사용하여 이미지 콘텐츠를 분류하는 완전히 연결된 연결된 계층에 의해 처리 가능한 기능으로 이미지 데이터를 변환

일반적인 3D 딥러닝 접근방식은 3D에서 convolution 작동이 잘 정의되지 않기 때문에 기존의 2D방법을 유클리드 이외의 영역으로 확장

3D딥러닝과 관련된 가장 어려운 연구주제 중 하나는 형상 편집과 모델링과 같은 컴퓨터 그래픽 애플리케이션을 지원하는 것이 핵심이기에 3D 형상의 의미론적 세분화다.

Segment 3D 형상에 대한 당면 과제로는 다양한 topologe처리, 소음이 많은 기하학적 구조와 다양한 해상도 처리, 다양한 segment에 대한 의미 표현 모델링등이다.

3D 세분화는 다중 뷰, 볼륨 측정 또는 본질적인 딥러닝 기반 접근방식을 통해 수행할 수 있다. 다중 뷰 및 부피 측정 접근방식은 각각 2D또는 3D그리드와 같은 유클리드 구조를 사용하여 2D CNN으로 3D모양을 처리한다.

특히, 다중 뷰 접근방식은 모델을 중심으로 다른 관점에서 찰영한 렌더링 깊이 영상의 집합을 사용하여 3D 모델의 표현을 단순화하므로, 분할이 3D모델 폴리곤 밀도와 무관하게 된다.

다중 뷰 접근방식은 데이터가 2D로 투영될 때 기하학적 정보가 손실될 수 있기 때문에 3D 형태의 기하학적 특성을 완전히 이용할 수 없다.

부피 측정 접근법은 물체의 기하학적 디테일을 무색하게 할 수 있는 복셀을 사용하여 3D 형상에 근사하다.

내재적 접근방식은 점 기반 접근방식과 convolution 기반 접근방식으로 더욱 나눌 수 있다.

점 기반 접근방식은 형상 정점에 직접 형상 추출기를 정의하는 반면, convolution 기반 접근방식은 전통적인 convolution 연산을 그리드 같은 구종에서 삼각형 메쉬로 확장한다.

점 기반 접근법은 대부분 형상의 각 꼭지점을 독립적으로 처리하고 지역 정보를 느슨하게 활용한다.

기존의 convolution 기반 접근법에 의해 사용되는 추가 구조는 형태 표현 복잡성을 증가시켜 고밀도 폴리곤 모델의 처리를 하지 못한다.

일반적으로 3D 세분화 접근방식은 통제된 시나리오에서 수집된 데이터셋에 대한 성능을 검증하며, 대부분 포토그램 측정 기법을 사용하여 재구성된 3D 모델에 대해 수행된 평가가 부족하다.

본 논문에서는 3D 형상을 중심으로 다양한 관점에서 3D 증강 뷰를 구축함으로써 장점을 모두 유지하는 새로운 3D 세분화 접근방식을 제안

3D 증강 뷰는 2D 일반 그리드에 3D 형상 부분을 투영하는 것으로, 그리드의 각 셀은 해당 투영된 부분의 깊이와 정규에 대한 정보를 인코딩한다.

이를 통해 학습할 파라미터 수를 크게 줄이고 다양한 망사로 도형을 3D분할할 수 있다.

우리는 공개된 데이터셋의 합성 3D 모양과 스마트폰으로 캡처하고 포토그램 측정 기법을 사용하여 재구성하는 객체의 3D 모양에 대한 우리의 접근 방식을 평가한다.

결과는 제안된 접근 방식이 대안적 접근방법에 의해 사용되는 매개변수의 1%만을 사용함으로써 최첨단 정확도를 달성할 수 있음을 보여준다.

2.1 문제 공식

정점 x ∈ X로 구성된 3D 형상을 가진 는 라벨 공간 L = {1, . . . , L}을 통해 확률분포 p(x)를 출력하는 신경망 접근 p(x) = ΓΘ(x)를 설계한다. 여기서 L은 분할 라벨 수이다.

X의 출력 분할은 로 계산되며 여기서 h(x)는 정점 x의 세그먼트 클래스를 정의하는 라벨이다.

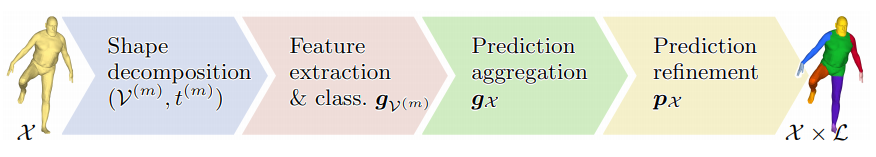


그림 1. 우리의 접근 개요 다양한 관점에서 3D 증강 뷰는 3D 형상(모양분해)으로 계산

점 별 형상(즉, 좌표 및 지표면 정규)은 이러한 3D 뷰에서 추출하여 분할 예측을 얻기 위해 분류

예측은 원래 형상에 다시 투영되고 집계되며, 국소 예측 일관성을 위해 조건부 무작위 필드를 통해 정제

여기서 h(x)는 꼭지점 x의 세그먼트 클래스를 정의하는 테이블이다.

형상분해는 입력된 3D 모양을 3D 증강 뷰 또는 3D 뷰로 변환

각 3D뷰는 각 꼭지점의 클래스를 예측하는 기능 추출 및 분류 네트워크, 즉 ViewNet에 의해 처리

예측 집합은 각 3D 뷰의 ViewNet 예측을 원래 3D 모양에 재프로젝트한다.

예측 정교화는 원래 모양에 대한 문맥 정보를 사용하여 클래스 예측을 개선한다.

그림 1은 블록 다이어그램을 보여준다.

2.2 형상분해

입력 형태를 여러 구성 요소로 분해하여 3D 형상 표현을 단순화한다.

형상 분해는 형상 정점을 군집화하거나 기하학적 원시 요소를 사용하거나 다른 관점에서 범위 스캔을 생성하여 수행할 수 있다.

우리는 후자의 3D 표현, 분해능 및 꼭지점 위상에 관계없이 3D 형태를 처리하기 위해 유사한 접근 방식을 사용

정점 X = (), , n=1,…,N의 삼각망 형태로 X를 주어 M의 다른 관점으로부터 3D 뷰를 구축하여 X를 단순화

Let I(u,v;) = (u,v,d(u,v)) 는 m번째 관점 에서 캡처되는 범위 스캔이며, 여기서 (u,v)는 픽셀의 좌표, d(u,v)는 3D 형상의 깊이 값, m=1,…,M

를 정점 X의 좌표에 범위 스캔의 좌표(u,v,d(u,v))를 등록하여 정점 = ()을 얻는 m번째 3D 뷰가 되도록 한다.

3D 뷰의 면은 2D 영상의 일반적인 그리드 패턴을 사용하여 깊이 값을 연결함으로써 얻는다.

각 꼭지점 에 대해 우리는 표면 정상 을 계산하여 3D 뷰의 신호를 로 정의한다.

3D 뷰와 입력 3D 모양의 관계는 대응 함수 로 정의되며, 3D 모양의 해당 정점에 m번째 3D 뷰의 정점을 할당한다.

그림 2. 3D 증강 뷰의 예



왼쪽 : FAUST 데이터 집합의 합성 3D 모양

오른쪽 : 형상 분해 모듈에서 추출한 3D 증강 뷰의 예

3D 뷰는 동일한 정점 밀도를 가지며 낮은 분해능에서도 기본 형상을 캡처

2.3 형상 추출 및 분류

형상 추출 및 분류 모듈은 M 3D 뷰를 병렬로 처리하여 공유 가중치를 가진 일련의 깊이 신경 네트워크, 즉 ViewNets를 통해 형상을 학습

형식적으로 각 ViewNet은 비선형 파리메트릭 함수로, 정점형 피쳐 를 입력으로 취하고 확률 분포 을 출력하는 것이다. 여기서 L은 분할 등급과 의 수이다.

을 의 모든 정점의 pdfs를 포함하는 행렬로 한다.

ViewNet모듈은 내인성 convolution(IC), 완전연결(FC) 및 softmax 계층의 구성으로 정의

FC와 softmax는 표준 레이어인 반면, IC 레이어는 3D 뷰에서 convolution 연산을 수행하기 위해 전통적인 유클리드 CNN에서 사용되는 convolution 레이어를 대체

IC레이어를 사용하는 에서 convolution은 x의 로컬 인접지역의 신호를 고정된 그리드에 매핑하는 일련의 가중치 기능과 로컬 좌표 프레임의 형태로 추가 정보가 필요하다.

2.4 예측집계

각 3D 뷰에서 유추된 예측은 원래의 입력에 분할 결과를 전송하기 위해 3D 형상 X에 다시 투영되고 집계

이를 ProjNet이라고 명칭함

ProjNet은 의 ViewNet 예측 를 입력으로, 임의의 m의 대응 함수 를 입력으로 하는 풀링 작업을 채택하여, X에 정의된 단일 신뢰 맵를 생산한다.

풀링 연산을 정의되며, 여기서 는 정점 에 상대적인 3D 뷰 지수 집합이며, 는 번째 3D 뷰의 정점 에 연관된 분할 등급에 대한 확률 분포다.

2.5 정밀 예측

ProjNet의 출력은 포인트 와이즈 예측이다.

각 꼭지점의 라벨 예측은 주변 정점과 독립적으로 추정되므로 국소 라벨 불일치가 발생할 가능성이 높다.

또한 입력 3D 형상의 일부 정점이 3D 뷰 중 하나에 투영되지 않았을 수 있으므로 X에 정의되지 않은 라벨 예측으로 정점이 나타날 수 있다.

따라서 우리는 구조적이고 밀도 높은 예측을 도출하기 위해 상황별 정보를 이용하는 표면 기반 조건 무작위 필드(CRF)접근법을 사용하여 로컬 라벨 일관성을 부여

각 꼭지점 에 대해 은 라벨 을 할당하는 랜덤 변수가 되게 하고 은 X의 N정점과 연관된 랜덤 변수의 집합이 되게 한다.

Y와 관련된 CRF 에너지는 로 정의된다.

여기서 단항는 의 정점 에 대한 할당 비용을 정량화하고 쌍항 는 은 정점 의 공동 할당 비용을 정량한다.

는 꼭지점 을 L에 할당하는 비용을 측정하기 때문에 단항은 로 정의한다.

대신 쌍방향 전위는 다음과 같은 세 가지 가우스 커널의 가중 합으로 정의된다.

여기서 는 꼭지점 사이의 측지의 거리, 는 X의 ID 함수, 는 라벨 호환성 용어다.

은 국소 공간 일관성을 선호하며, 는 유사한 특성을 가진 정점에 유사한 라벨을 할당하도록 촉진

세번째 커널 는 참신하고 대칭을 차이를 보여주는 것을 소개하기위해 도입

대칭 부분은 서로 멀리 위치할 가능성이 높기 때문에, 우리는 유사한 라벨을 부착하기 위해 먼 지점을 피하기 위해서 를 설계함.

CRF 학습 가능한 매개변수 세트는 로 정의된다.

그림 3은 CRG가 인간의 3D 모양에 대해 학습한 매개변수(즉, )의 예를 통해 세그먼트 사이의 관계를 학습하는 방법을 보여준다.

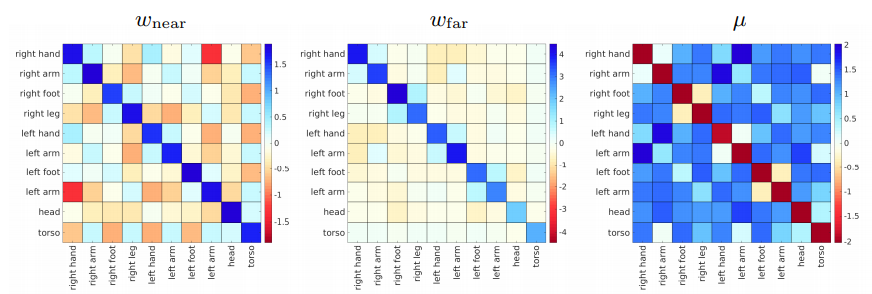


그림 3. 조건부 무작위 필드(CRF)의 예에서 인간의 3D 형상의 경우 가중치()를 학습

에서 우리는 머리 중량이 머리와 오른발/오른팔 사이의 관계가 아니라 머리와 몸통 사이에 강한 관계가 있음을 암시한다는 것을 관찰할 수 있다.

마찬가지로 몸통무게는 몸통과 발/손 사이보다는 모통과 팔/다리 사이에 강한 관계가 있음을 시사한다.

X에 대한 y의 가장 가능성이 높은 pdf 구성은 Eq.1 에 정의된 에너지 E(y)를 최소화함으로써 얻는다.

CRF 분포의 정확한 추론은 난해하기 쉬우므로 평균 필드 근사치를 사용

대략적인 평균 필드 추론에 대한 반복 알고리즘은 알고리즘의 각 단계를 CNN 계층으로 바꾸어 RNN으로 구현