

Wasserstein Center 손실을 이용한 스케치 기반 3차원 물체 검색☆

Sketch-based 3D object retrieval using Wasserstein Center Loss

지 명 근¹ 전 준 철^{1*} 김 남 기¹
Myunggeun Ji Junchul Chun Namgi Kim

요 약

스케치 기반 3차원 물체 검색은 다양한 3차원 물체를 사람이 손으로 그린 스케치를 질의(query)로 사용하여 물체를 편리하게 검색하는 방법이다. 본 논문에서는 스케치 기반 3차원 물체 검색을 위해 스케치 CNN(Convolutional Neural Network)과 Wasserstein CNN 모델에 Wasserstein Center 손실을 적용하여 물체의 검색 성공률을 향상시키는 새로운 방법을 제안한다. 제안된 Wasserstein Center 손실이란 각 물체의 클래스(category)의 중심을 학습하고, 동일한 클래스의 특징과 중심 간의 Wasserstein 거리가 작아지도록 만드는 방법이다. 이를 위하여 제안된 3차원 물체 검색은 다음의 단계로 수행된다. 첫 번째로, 3차원 물체의 특징은 3차원 물체를 여러 방향에서 촬영된 2차원 영상의 특징을 CNN을 이용하여 추출하고, 각 영상 특징의 Wasserstein 중심을 계산한다. 두 번째로, 스케치의 특징은 별도의 스케치 CNN을 이용하여 추출하였다. 마지막으로, 추출한 3차원 물체의 특징과 스케치의 특징을 본 논문에서 제안한 Wasserstein Center 손실을 이용하여 학습하고 스케치 기반의 3차원 물체 검색에 적용하였다. 본 논문에서 제안한 방법의 우수성을 입증하기 위하여 SHREC 13과 SHREC 14의 두 가지 벤치마크 데이터 집합을 이용하여 평가하였으며, 제안된 방법이 기존의 스케치 기반 검색방법들과 비교하여 모든 측정 기준에서 우수한 결과를 나타냄을 확인할 수 있었다.

☞ 주제어 : 합성곱 신경망, 영상 검색, 딥 러닝, 스케치 기반 3차원 물체 검색

ABSTRACT

Sketch-based 3D object retrieval is a convenient way to search for various 3D data using human-drawn sketches as query. In this paper, we propose a new method of using Sketch CNN, Wasserstein CNN and Wasserstein center loss for sketch-based 3D object search. Specifically, Wasserstein center loss is a method of learning the center of each object category and reducing the Wasserstein distance between center and features of the same category. To do this, the proposed 3D object retrieval is performed as follows. Firstly, Wasserstein CNN extracts 2D images taken from various directions of 3D object using CNN, and extracts features of 3D data by computing the Wasserstein barycenters of features of each image. Secondly, the features of the sketch are extracted using a separate Sketch CNN. Finally, we learn the features of the extracted 3D object and the features of the sketch using the proposed Wasserstein center loss. In order to demonstrate the superiority of the proposed method, we evaluated two sets of benchmark data sets, SHREC 13 and SHREC 14, and the proposed method shows better performance in all conventional metrics compared to the state of the art methods.

☞ keyword : Convolutional Neural Network, Image retrieval, Deep Learning, Sketch-based 3D object retrieval

1. 서 론

터치스크린 기술이 발전함에 따라 태블릿 컴퓨터 및 스마트 폰과 같은 컴퓨터 시스템과 상호 작용할 수 있는

방법인 스케치는 키워드를 검색어로 사용하는 이전의 연구보다 직관적이고 사용자 입력이 편리하여 다양한 2차원 또는 3차원 물체 검색에 활용되고 있다 [1, 2, 3]. 그러나 기존에 제안된 딥러닝(deep learning) 방법을 포함한 방법들의 경우 정확한 검색률에 있어서 만족스러운 결과를 얻지 못하고 있는 실정이다. 본 연구에서는 3차원 물체를 스케치를 이용하여 검색하는 과정에서 CNN과 새로 제안하는 Wasserstein Center 손실 함수를 적용한 스케치 기반 모델 검색 방법을 제안한다. 본 논문의 연구 주제와 같이 CNN을 이용하여 유사도를 측정하는 문제에는 Contrastive

¹ Department of Computer Science, Kyonggi University, Gyeonggi-do, 443-760, Korea.

* Corresponding author (jchun@kyonggi.ac.kr)

[Received 31 October 2018, Reviewed 1 November 2018, Accepted 17 November 2018]

☆ 본 연구는 2018학년도 경기대학교 대학원 연구원장학생 장학금 지원에 의하여 수행되었음.

손실[4], Triplet 손실[5], Center 손실[6] 등의 방법이 있다. 본 논문에서 제안하는 Wasserstein Center 손실은 Center 손실의 거리 측정 방법을 개선한 방법이다[7]. Wasserstein Center 손실은 각 물체의 클래스 중심을 학습하고, 동일한 클래스의 특징과 중심 간의 Wasserstein 거리가 작아지도록 유도하는 방법이다. Wasserstein Center 손실을 스케치 기반 3차원 물체 검색에서 적용하기 위해 제안된 방법은 크게 두 개의 구조로 구성된다. 첫 번째로 3차원 물체의 특징을 추출하는 Wasserstein CNN이다. 이 부분은 3차원 물체의 여러 방향에서 촬영한 영상의 특징을 CNN을 이용하여 추출하고, 이 특징에 Wasserstein 중심을 계산하여 3차원 물체의 특징을 추출하는 부분이다. 두 번째로 스케치의 특징을 CNN을 이용하여 추출하는 Sketch CNN이다. Wasserstein CNN과 Sketch CNN의 출력을 Wasserstein Center 손실을 이용하여 학습하게 된다.

본 논문은 다음과 같이 구성된다. 2장에서는 본 논문에서 제시된 방법과 관련된 연구를 소개한다. 3장에서는 Wasserstein 거리와 Wasserstein 중심, Center 손실의 개념을 소개하고, 3차원 물체의 특징을 추출하기 위한 Wasserstein CNN 구조를 설명하며, 아울러 Wasserstein Center 손실을 이용한 스케치 기반 3차원 물체 검색방법을 제안한다. 4장에서는 제안된 방법에 의한 실험 결과를 소개하며, 끝으로 5장에서는 결론 및 향후 연구 방향에 관하여 기술한다.

2. 관련 연구

3차원 물체에 비해 2차원 스케치는 매우 추상적이고 사람에게 따라 주관적으로 그려지기 때문에 일반적으로 동일 물체를 그린 스케치는 서로 큰 차이가 있다. 최근에는 스케치 기반의 3차원 물체 검색에 대한 광범위한 연구가 진행되었다. 스케치와 3차원 물체의 2차원 투영을 기술하기 위해 지역 기반 Zernike 모멘트, 외곽 기반 푸리에 기술자, 이심률 특징 및 원형성 특징을 포함하는 통합 기술자 ZFEC이 제안되었다[2]. 또 다른 연구에서는 정의된 2차원 투영의 3차원 물체 실루엣을 3차원 물체를 표현하는 데 사용하였다[8]. 스케치 및 3차원 물체를 설명하기 위해 3가지 종류의 기술자로 에지 로컬 방향 히스토그램(HELO), 경사지향 히스토그램(HOG) 및 푸리에 기술자를 사용하며, 맨해튼 거리와 KD-트리는 스케치를 3차원 물체의 2차원 투영과 매칭시키는 데 사용하였다.

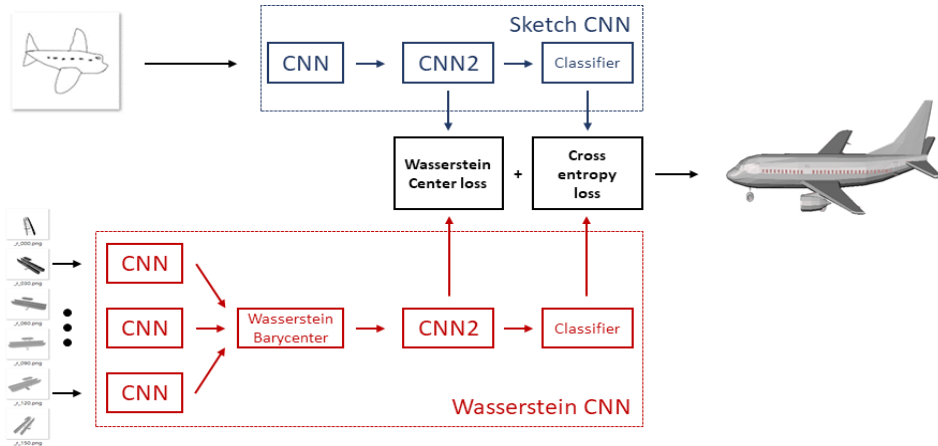
한편, 직접 설계한 특징 외에도 학습 기반 특징을 이용

하는 스케치 기반 3차원 물체 검색방법도 연구되었다. Eitz는 Gabor local line-based features (GALIFs), bag-of-feature (BOF) 히스토그램을 이용하여 스케치와 3차원 물체의 2차원 투영의 특징을 추출하고, 스케치와 2차원 투영의 BOF 히스토그램의 최소 거리를 검색을 위한 유사성으로 사용하였다[1]. Furuya 등은 스케치와 3차원 물체의 2차원 투영을 기술하기 위해 BF-SIFT 특징을 제안했다[9].

최근에는 심층 신경망을 이용한 스케치 기반의 3차원 물체 검색에 관한 연구가 활발히 진행되고 있다. Wang은 스케치와 3차원 물체의 특징을 추출하는 두 개의 삼 CNN을 이용하여 3차원 물체를 검색하는 방법을 제시하였다[3]. Xie는 3차원 물체에서 여러 개의 영상을 선택하여, CNN을 이용하여 영상의 특징을 추출하고, 추출된 특징의 Wasserstein 중심을 계산하여 3차원 물체와 스케치를 매칭시키는 방법을 제안하였다[10].

제안하는 방법과 관련된 기존의 입력 데이터 간의 유사도를 이용하는 손실 함수에 관한 연구는 다음과 같다. Hadsell은 입력이 두 개의 쌍으로 구성될 때 같은 클래스 쌍은 거리가 가까워지고, 다른 클래스일 때는 일정 거리 이상이 되게 하는 Contrastive 손실을 제시하였다[4]. Schroff는 얼굴 인식을 위해 입력이 같은 클래스와 다른 클래스를 모두 포함한 3개의 쌍으로 구성되어 같은 클래스 간의 거리를 줄이고, 다른 클래스 간의 거리를 늘리는 Triplet 손실을 제안하였다[5]. Wen은 얼굴 인식을 위해 각 클래스의 중심을 학습하고, 동일한 클래스와 중심 간의 거리를 줄이는 Center 손실을 제안하였다[6]. He는 스케치를 이용하여 3차원 물체를 검색하기 위해 앞의 Triplet 손실과 Center 손실을 합친 Triplet Center 손실을 제안하였다[11].

그러나 Contrastive 손실은 데이터 쌍을 잘 구성하지 못하면 학습 속도가 떨어지는 문제가 있고, Triplet 손실은 데이터를 3개의 쌍으로 구성해야 하므로 학습 시간이 오래 걸린다는 단점이 있다. 따라서 본 논문에서는 3차원 물체의 좋은 특징을 얻기 위하여 3차원 물체의 여러 방향에 대한 영상의 특징을 CNN을 이용하여 추출하고, 추출한 특징을 Wasserstein 중심을 계산하여 3차원 물체의 특징을 얻는 방법과, 스케치의 특징을 CNN을 이용하여 추출하고, 3차원 물체와 스케치의 특징을 거리 측정 방식을 개선한 Center 손실인 Wasserstein center 손실을 사용하여 학습시키는 방법을 제안하고 구현 및 평가 하였다.



(그림 1) Wasserstein Center 손실을 이용한 스케치 기반 3차원 물체 검색의 개요
 (Figure 1) The overview of Sketch-based 3D object retrieval using Wasserstein Center Loss

3. Wasserstein Center 손실을 이용한 스케치 기반 3차원 물체 검색

스케치 기반 3차원 물체 검색을 위하여 본 논문에서는 그림 1과 같이 3차원 물체의 특성추출을 위한 Wasserstein CNN과 스케치 CNN으로부터 Wasserstein Center 손실을 이용하여 3차원 물체를 검색하는 방법을 제안한다. 3.1과 3.2장에서는 Wasserstein 거리와 Wasserstein 중심을 소개하며, 3.3장에서는 3차원 물체의 특징을 추출하기 위한 Wasserstein CNN의 특성을 설명하고, 3.4장에서는 Center 손실을 소개한다. 3.5장에서는 스케치 기반 3차원 물체 검색에 적용된 Wasserstein Center 손실에 관하여 기술하였다.

3.1 Wasserstein 거리

Wasserstein 거리[12]란 확률 분포의 거리를 측정하는 여러 방법 가운데 하나로 Kantorovich-Rubinstein Metric 또는 Earth Mover's Distance(EMD)라고 불린다. Wasserstein 거리는 컴퓨터 비전 [13] 및 기계 학습 [7]에서 확률 분포의 거리 측정에 널리 사용되었다.

$p \in \mathbb{R}^{r \times 1}$, $q \in \mathbb{R}^{s \times 1}$ 을 확률 분포라고 가정할 때, 확률 분포 p 와 q 사이의 운송 계획은 다음 식 1과 같다.

$$R(p, q) = \{ T \in \mathbb{R}_+^{r \times s}; T1 = p, T^T 1 = q \} \quad (1)$$

위의 식 1에서 T 는 전송 계획이고, 1은 모든 요소가 1인 열 벡터이다. 이때 p 와 q 사이의 Wasserstein 거리 $D(p, q)$ 는 다음과 같이 정의할 수 있다.

$$D(p, q) = \min_{T \in R(p, q)} \langle M, T \rangle \quad (2)$$

위 식 2에서 $M \in \mathbb{R}^{r \times s}$ 은 p 와 q 의 쌍 거리 행렬로, 그라운드 매트릭스라고 한다. 이때 $\langle M, T \rangle$ 은 M 과 T 의 내적을 의미한다. Wasserstein 거리 $D(p, q)$ 는 p 의 질량을 q 로 전송하는 최적의 전송 계획 비용을 의미한다.

많은 경우 식 2는 유일한 해를 가지지 않을 수 있어서 엔트로피 정규화 항을 더한 아래 식 3[14]를 사용한다.

$$D(p, q) = \min_{T \in R(p, q)} \langle M, T \rangle + \gamma \langle T, \log T \rangle \quad (3)$$

식 3에서 $\langle T, \log T \rangle$ 는 음의 엔트로피이고, γ 는 정규화 매개 변수이다. 위 식의 최적 해는 다음 식 4와 같이 표현한다.

$$\hat{T} = \text{diag}(u) K \text{diag}(v) \quad (4)$$

식 4에서 $K = e^{-M/\gamma}$ 이고, 벡터 u 와 v 는 Sinkhorn의 알고리즘[15]을 사용했다.

3.2 Wasserstein 중심

Wasserstein 중심이란 Wasserstein 거리를 이용하여 계산한 확률 분포 집합의 중심점을 말한다. 확률 분포 집합이 $p_i \in \mathbb{R}^{r \times 1}, i = 1, 2, \dots, n$ 일 때, 이 집합의 중심 p_b 는 아래 식과 같이 정의된다[16].

$$\operatorname{argmin}_{p_b} \sum_{i=1}^n \lambda_i D(p_b, p_i) \quad (5)$$

식 5의 $D(p_b, p_i)$ 는 p_b 와 p_i 의 Wasserstein 거리이며, λ_i 는 가중치이다. Wasserstein 중심 p_b 는 다음과 같이 반복적으로 p_b^t 를 계산하여 구할 수 있다.

$$p_b^t = \prod_{i=1}^n (K^T a_i^t)^{\lambda_i} \quad (6)$$

$$c_i^{t+1} = \frac{p_b^t}{K^T a_i^t}; a_i^{t+1} = \frac{p_i}{K^T a_i^t}$$

식 6의 p_b^t 는 Wasserstein 중심 p_b 의 t 번째 반복이고, c_i^{t+1}, a_i^{t+1} 는 보조 변수이다[17].

3.3 Wasserstein CNN

앞서 제시된 그림 1의 Wasserstein CNN 구조와 같이 3차원 물체의 여러 방향에서 렌더링 된 이미지의 특징을 CNN을 이용하여 추출하고, 추출된 특징의 Wasserstein 중심을 계산하여 3차원 물체의 특징을 추출한다[10]. 본 연구에서는 CNN2를 이용하여 스케치에서 추출된 특징과 동일한 공간으로 사상시켜준다. 마지막으로 분류기(classifier)를 이용하여 특징을 분류한다.

3.4 Center 손실

Center 손실은 주로 얼굴 검증 문제에서 교사 학습인 Softmax 손실을 보정하기 위하여 제안되었다. Center 손실은 같은 클래스의 특징에서 중심을 학습하고, 각각의 특징과 그 중심의 거리를 최소화한다. Center 손실은 다음 식 7과 같다.

$$L_C = \frac{1}{2} \sum_{i=1}^m \|x_i - c_{y_i}\|_2^2 \quad (7)$$

위 식 7에서 x_i 는 입력된 특징이고, $c_{y_i} \in \mathbb{R}^d$ 는 클래스 y_i 의 중심이고, d 는 특징의 차원을, m 은 특징의 개수를 나타낸다. 전체 데이터에 대해 중심을 계산하는 것은 너무 많은 계산이 필요하므로 미니 배치 별로 중심을 학습한다.

3.5 Wasserstein Center 손실

Wasserstein center 손실은 위의 Center 손실에서 중심과 특징 사이의 거리를 유클리드 거리 대신 Wasserstein 거리를 사용한 손실 함수이다[7]. Wasserstein Center 손실은 다음 식 8과 같다.

$$L_{WC} = \sum_{i=1}^m D(x_i, c_{y_i}) \quad (8)$$

위 식 8에서 x_i 는 Wasserstein CNN을 이용하여 추출된 3차원 물체의 특징과 Sketch CNN을 이용하여 추출된 스케치의 특징이다. $c_{y_i} \in \mathbb{R}^d$ 는 y_i 번째 클래스의, 3차원 물체의 특징과 스케치 특징의 공통 중심이다. x_i 에 대한 L_{WC} 의 기울기와 c_{y_i} 의 갱신 방정식은 다음 식 9와 같다.

$$\frac{\partial L_{WC}}{\partial x_i} = D(x_i, c_{y_i}) \quad (9)$$

$$\Delta c_j = \frac{\sum_{i=1}^m \delta(y_i = j) \cdot D(x_i, c_{y_i})}{1 + \sum_{i=1}^m \delta(y_i = j)}$$

위 식 9에서 δ 는 조건이 참일시 1, 거짓일시 0이다. 제안하는 방법에서는 Wasserstein center 손실과 함께 Cross entropy 손실을 사용한다[18]. 두 손실을 합한 전체 손실은 다음 식 10과 같다.

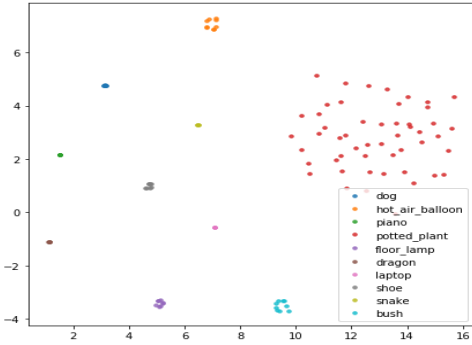
$$L = L_{WC} + L_C \quad (10)$$

$$= \sum_{i=1}^m D(x_i, c_{y_i}) + \sum_{i=1}^m y_i \log(y'_i)$$

위 식 10에서 L_C 는 Cross entropy 손실이고, y'_i 는 Classifier 구조에서 예측된 3차원 물체와 스케치의 클래스이다.

다음 그림 2는 제안된 Wasserstein Center 손실을 이용하여 학습된 특징을 임의의 10개의 클래스(개, 열기구, 피아노, 화분, 플로어 램프, 용, 노트북, 신발, 뱀, 덩불)를

선택하여 t-SNE[19]을 이용하여 시각화한 예시이다. 실험 결과 유사한 특성을 갖는 물체들이 효과적으로 군집을 형성하는 것을 확인할 수 있다.



(그림 2) Wasserstein Center 손실을 이용하여 학습된 특징 (임의의 10개 클래스 : 개, 열기구, 피아노, 화분, 플로어 램프, 용, 노트북, 신발, 뱀, 덩굴) 시각화 (Figure 2) A visualization of learned features (randomly selected 10 classes : dog, hot air balloon, piano, potted plant, floor lamp, dragon, laptop, shoe, snake, bush) by Wasserstein center loss

4. 실험 결과

이 장에서는 제안된 방법인 Wasserstein Center 손실을 이용한 스케치 기반 3차원 물체 검색결과에 대한 평가 수행하였다. 실험을 위하여 SHREC 13[2] 데이터 셋과 SHREC 14[20] 데이터 셋을 이용하여 기존에 소개된 최신의 스케치 기반 3차원 물체 검색 기술과 그 결과를 비교하고, 검색의 정확도를 평가하였다.

4.1 데이터 셋 및 실험 설정

실험 및 평가는 SHREC 13과 SHREC 14 데이터 셋을 대상으로 진행되었다. SHREC 13 데이터 셋은 스케치 기반 3차원 물체 검색을 위하여 제작된 데이터 셋으로, 사람이 그린 7,200개의 스케치와 3차원 물체 1,258개를 총 90개 클래스로 구성하고 있다. 스케치 데이터는 클래스별로 학습 시 50장, 테스트 시 30장씩 사용하였다.

SHREC 14 데이터 셋은 SHREC 13 데이터 셋의 크기

(표 1) 실험 환경 (Table 1) Experimental Environments

구분	세부 환경
CPU	AMD Ryzen 7 1700 3GHz
GPU	NVIDIA GeForce GTX 1080 Ti
RAM	32.00 GB
OS	Ubuntu 16.04
Language	Python 3.5
Develop Tool	Jupyter Notebook
Library	Pytorch

를 확장시킨 데이터 셋으로 13,680개의 스케치와 3차원 물체 8,987개가 총 171개 클래스로 구성된다. SHREC 13 실험과 동일하게 스케치 데이터는 클래스별로 학습 시 50장, 테스트 시 30장씩 사용하였다.

본 논문에서 사용한 실험 환경은 표 1과 같으며, 제안하는 구조에서 3차원 물체를 12방향에서 렌더링한 측면 영상이 사용되었다. Wasserstein CNN과 Sketch CNN의 CNN 구조는 모두 Resnet-18[21]가 사용되었다. Wasserstein CNN에서는 측면 영상을 Resnet-18 구조의 CNN에 입력한 뒤 Resnet-18 구조의 ‘average pool’ 레이어에서 512개의 출력을 받아 Wasserstein 중심을 계산하여 CNN2에 입력한다. Sketch CNN의 CNN에서도 스케치를 Resnet-18 구조에 입력한 뒤 ‘average pool’ 레이어의 512개의 출력을 받아 CNN2에 입력한다. Wasserstein CNN과 Sketch CNN 모두 CNN2는 512-300-100의 출력을 가지는 레이어들로 구성되었으며, 모든 레이어에 ReLU 활성화 함수가 사용되었다. Classifier는 데이터 셋의 클래스 개수와 같은 수의 출력을 한다.

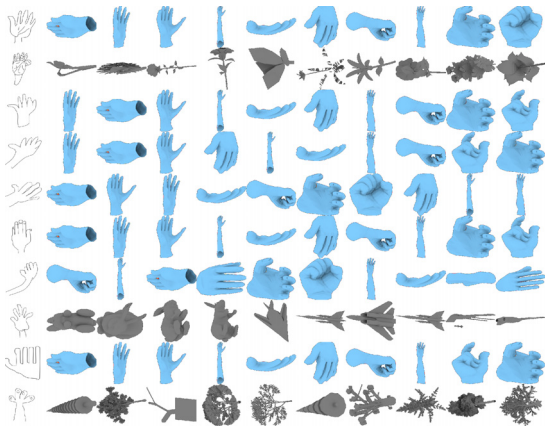
제안된 방법의 평가는 두 데이터 셋 모두 정확도-재현율(Precision-Recall) 곡선(PR-Curve), nearest neighbor (NN), first tier (FT), second tier (ST), E-measure (E), discounted cumulated gain (DCG) and mean average precision (mAP)을 평가 지표로 하였다.

4.2 SHREC 13 데이터 셋 평가

제안된 방법을 SHREC 13 데이터 셋에 적용하여 성능을 평가하기 위하여 기존에 소개된 3차원 물체 검색 기술들(푸리에 기술자(FDC)[8], 에지 기반 푸리에 스펙트럼 기술자(EFSD)[8], 뷰 클러스터를 이용한 스케치 기반 검색(SBR-VC)[8], 교차 도메인 다양체 랭킹법(CDMR)[9],

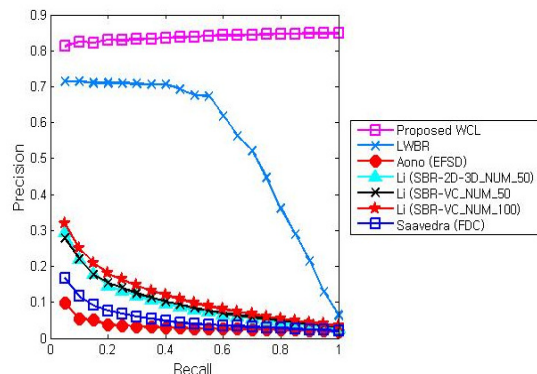
삼 네크워크(Siamese)[3], 학습된 wasserstein 중심 표현(LWBR)[10], 깊은 상관 메트릭 학습(DCML)[22], Triplet-Center 손실(TCL)[11])과 본 논문에서 제안한 Wasserstein Center 손실을 이용한 방법(WCL)을 PR-Curve와 NN, FT, ST, E, DCG, mAP를 이용하여 비교했다.

그림 3은 임의의 ‘손’ 클래스의 스케치로 검색을 한 결과 예제이다. 잘못 분류한 결과가 있으나, 형태는 유사함을 보이는 것으로 제안하는 구조는 입력한 스케치와 유사한 형태의 3차원 물체를 잘 검색하는 것을 알 수 있다.



(그림 3) SHREC 13 데이터 셋에 대한 검색 예제. 회색은 잘못 검색된 클래스(검색 클래스 : 손)

(Figure 3) Retrieval examples on SHREC 13 dataset. Mismatch highlighted in gray (Retrieval classes : hand)



(그림 4) SHREC 13 데이터 셋의 PR-Curve 결과 비교
(Figure 4) The precision-recall curves in SHREC 13 dataset

그림 4는 LWBR, EFSD, SBR-VC, FDC과 본 논문에서 제안한 WCL을 PR-Curve를 이용하여 비교해 본 결과이다. 이 실험 결과에서 보는 바와 같이 1 이하의 전체 재현율에서 기존의 방법보다 높은 정확도를 보였다. 이는 다른 방법들과 비교했을 때 검색 순위가 낮은 경우에도 검색 결과가 우수한 것을 의미한다.

(표 2) SHREC 13 데이터 셋의 NN, FT, ST, E, DCG, mAP 결과 비교 (%)

(Table 2) Comparison of NN, FT, ST, E, DCG, and mAP results in SHREC 13 dataset (%)

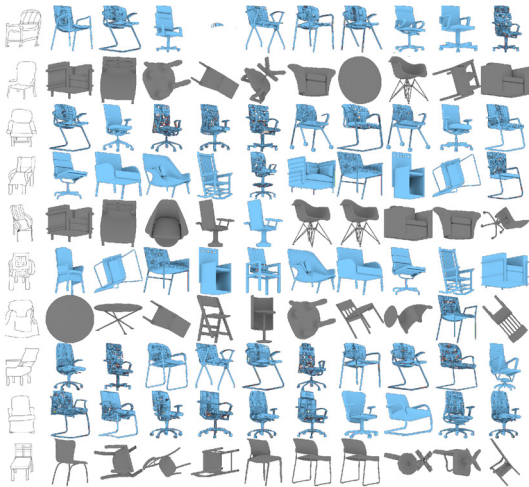
Methods	NN	FT	ST	E	DCG	mAP
CDMR [9]	27.9	20.3	29.6	16.6	45.8	25.0
SBR-VC [8]	16.4	9.7	14.9	8.5	34.8	11.6
Siamese [3]	40.5	40.3	54.8	28.7	60.7	46.9
DCML [22]	65.0	63.4	71.9	34.8	76.6	67.4
LWBR [10]	71.2	72.5	78.5	36.9	81.4	75.2
TCL [11]	76.3	78.7	84.9	39.2	85.4	80.7
WCL	80.7	82.4	86.3	40.0	87.4	83.9

표 2는 CDMR, SBR-VC, Siamese, DCML, LWBR, TCL과 본 논문에서 제안한 WCL을 NN, FT, ST, E, DCG, mAP를 이용하여 비교해 본 결과이다. 이 결과에서는 모든 평가 지표가 다른 최신의 방법보다 높았으며, 특히 기존의 방법 중 제일 성능이 좋았던 TCL 방법보다 모든 평가 지표가 0.8~4.7% 높았다.

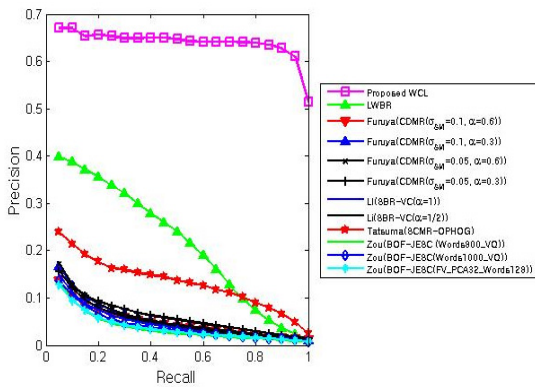
4.3 SHREC 14 데이터 셋 평가

SHREC 14 데이터 셋을 이용한 성능 평가에서는 기존에 소개된 3차원 물체 검색 기술들(경사지향 히스토그램의 중첩 피라미드(SCMR-OPHOG)[20], BOF 결합 기반의 확장된 형태 문맥 기법(BOF-JESC)[20], CDMR[9], SBR-VC[8], Siamese[3], DCML[22], LWBR[10], TCL[11])과 제안한 Wasserstein Center 손실을 이용한 방법을 PR-Curve와 NN, FT, ST, E, DCG, mAP를 이용하여 비교하였다.

그림 5는 임의의 ‘안락의자’ 클래스의 스케치로 검색을 한 결과 예제이다. 분류 오류로 제시되는 결과의 경우에도 합리적인 유사성이 있음 확인 할 수 있다. 예를 들어 안락의자의 스케치이지만 검색결과 일반 의자인 것과 같이 형태는 두 물체간의 유사함을 보이는 것을 알 수 있다. 이는 제안하는 구조가 입력한 스케치와 유사한 형태의 3



(그림 5) SHREC 14 데이터 셋에 대한 검색 예제. 회색은 잘못 검색된 클래스(검색 클래스 : 안락의자)
 (Figure 5) Retrieval examples on SHREC 14 dataset. Mismatch highlighted in gray(Retrieval classes : armchair)



(그림 6) SHREC 14 데이터 셋의 PR-Curve 결과 비교
 (Figure 6) The precision-recall curves in SHREC 14 dataset

차원 물체를 보다 효과적으로 검색하는 것으로 판단 할 수 있다.

그림 6은 LWBR, CDMR, SBR-VC, SCMR-OPHOG, BOF-JESC와 본 논문에서 제안한 WCL을 PR-Curve를 이용하여 비교해 본 결과이다. 이 결과에서는 1 이하의 전체 재현율에서 기존의 방법보다 정확도를 보였다. 이는 다른 방법들에 비해 검색 순위가 낮은 결과도 보다 잘 검

색이 됨을 의미한다.

표 3은 CDMR, SBR-VC, Siamese, DCML, LWBR, TCL 과 본 논문에서 제안한 WCL을 NN, FT, ST, E, DCG, mAP를 이용하여 비교해 본 결과이다. 이 결과에서는 모든 평가 지표가 다른 최신의 방법보다 높았으며, 특히 기존의 방법 중 제일 성능이 좋았던 TCL 방법보다 모든 평가 지표가 2.5~16.5% 높은 결과를 보였다.

(표 3) SHREC14 데이터 셋의 NN, FT, ST, E, DCG, mAP 결과 비교 (%)

(Table 3) Comparison of NN, FT, ST, E, DCG, and mAP results in SHREC14 datasets (%)

Methods	NN	FT	ST	E	DCG	mAP
CDMR (6)	10.9	5.7	8.9	4.1	32.8	5.4
SBR-VC (5)	9.5	5.0	8.1	3.7	31.9	5.0
Siamese (27)	23.9	21.2	31.6	14.0	49.6	22.8
DCML(21)	27.2	27.5	34.5	17.1	49.8	28.6
LWBR (7)	40.3	37.8	45.5	23.6	58.1	40.1
TCL (8)	58.5	45.5	53.9	27.5	66.6	47.7
WCL	61.0	61.3	67.9	33.4	73.6	64.2

5. 결 론

본 논문에서는 스케치 기반 3차원 물체 검색을 위한 Wasserstein Center 손실을 제안하였다. 본 논문에서는 Wasserstein CNN과 Sketch CNN에 Resnet을 이용하여 3차원 물체와 스케치의 특징을 추출하고, CNN2 구조를 이용하여 서로를 사상하였다. 사상한 특징과 예측한 클래스를 추출하고, 추출한 결과를 Wasserstein Center 손실과 Cross entropy 손실을 이용하여 학습하였다. 제안한 방법은 검색 결과가 정확한 라벨이 아니더라도 비슷한 모양을 찾는 것을 보아 3차원 물체와 스케치의 형태를 잘 학습한 것을 알 수 있었다. 제안한 방법을 SHREC 13과 SHREC 14 데이터 셋으로 평가한 결과 최신의 방법들과 비교했을 때 기존의 방법들에 비하여 정확한 검색 결과율을 보임을 확인할 수 있었다.

향후 본 논문에서 제안한 구조는 사용자가 스케치 입력력을 더 효과적으로 입력하는 입력 인터페이스의 개발과 스케치를 이용해 검출된 3차원 물체를 증강현실 시스템에 적용하여 증강물체로 활용하는 응용연구에 적용할 예정이다.

참고문헌(Reference)

- [1] M. Eitz, R. Richter, T. Boubekur, K. Hildebrand, and M. Alexa, "Sketch-based shape retrieval," *ACM Transactions on Graphics*, vol. 31, no. 4, pp. 1 - 10, 2012. <https://doi.org/10.1145/2185520.2335382>
- [2] B. Li, Y. Lu, A. Godil, T. Schreck, B. Bustos, A. Ferreira, T. Furuya, M. J. Fonseca, H. Johan, T. Matsuda, R. Ohbuchi, P. B. Pascoal, and J. M. Saavedra, "A comparison of methods for sketch-based 3D shape retrieval," *Computer Vision and Image Understanding*, vol. 119, pp. 57 - 80, 2014. <https://doi.org/10.1016/j.cviu.2013.11.008>
- [3] Fang Wang, Le Kang, and Yi Li, "Sketch-based 3D shape retrieval using Convolutional Neural Networks," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015. <https://doi.org/10.1109/cvpr.2015.7298797>
- [4] R. Hadsell, S. Chopra, and Y. LeCun, "Dimensionality Reduction by Learning an Invariant Mapping," 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 2 (CVPR'06), 2006. <https://doi.org/10.1109/cvpr.2006.100>
- [5] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015. <https://doi.org/10.1109/cvpr.2015.7298682>
- [6] Y. Wen, K. Zhang, Z. Li, and Y. Qiao, "A Discriminative Feature Learning Approach for Deep Face Recognition," *Lecture Notes in Computer Science*, pp. 499 - 515, 2016. https://doi.org/10.1007/978-3-319-46478-7_31
- [7] A. Rolet, M. Cuturi, and G. Peyré. "Fast dictionary learning with a smoothed wasserstein loss," *International Conference on Artificial Intelligence and Statistics*, Cadiz, Spain, pp. 630 - 638, 2016. <http://www.jmlr.org/proceedings/papers/v51/rolet16.pdf>
- [8] B. Li, Y. Lu, A. Godil, T. Schreck, M. Aono, H. Johan, J. M. Saavedra, and S. Tashiro. "Shrec'13 track: Large scale sketchbased 3D shape retrieval," *Eurographics Workshop on 3D Object Retrieval*, Girona, Spain, pp. 89 - 96, 2013. <https://dx.doi.org/10.2312/3DOR/3DOR13/089-096>
- [9] T. Furuya and R. Ohbuchi. "Ranking on cross-domain manifold for sketch-based 3D model retrieval," *International Conference on Cyberworlds*, Yokohama, Japan, pp. 274 - 281, 2013. <https://doi.org/10.1109/cw.2013.60>
- [10] J. Xie, G. Dai, F. Zhu, and Y. Fang, "Learning Barycentric Representations of 3D Shapes for Sketch-Based 3D Shape Retrieval," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017. <https://doi.org/10.1109/cvpr.2017.385>
- [11] He, Xinwei, et al. "Triplet-Center Loss for Multi-View 3D Object Retrieval," *arXiv preprint arXiv:1803.06189*, 2018. http://openaccess.thecvf.com/content_cvpr_2018/CameraReady/1632.pdf
- [12] V. I. Bogachev and A. V. Kolesnikov, "The Monge-Kantorovich problem: achievements, connections, and perspectives," *Russian Mathematical Surveys*, vol. 67, no. 5, pp. 785 - 890, 2012. <https://doi.org/10.1070/rm2012v067n05abeh004808>
- [13] Y. Rubner, C. Tomasi, and L. J. Guibas. "The Earth Mover's Distance as a metric for image retrieval," *International Journal of Computer Vision*, vol. 40, no. 2 pp. 99 - 121, 2000. <https://doi.org/10.1023/a:1026543900054>
- [14] M. Cuturi. "Sinkhorn distances: Lightspeed computation of optimal transport," *Advances in Neural Information Processing Systems*, Lake Tahoe, Nevada, USA, pp. 2292 - 2300, 2013. <https://papers.nips.cc/paper/4927-sinkhorn-distances-light-speed-computation-of-optimal-transport.pdf>
- [15] R. Sinkhorn, "Diagonal Equivalence to Matrices with Prescribed Row and Column Sums," *The American Mathematical Monthly*, vol. 74, no. 4, p. 402, 1967. <https://doi.org/10.2307/2314570>
- [16] J.-D. Benamou, G. Carlier, M. Cuturi, L. Nenna, and G. Peyré, "Iterative Bregman Projections for Regularized Transportation Problems," *SIAM Journal on Scientific Computing*, vol. 37, no. 2, pp. A1111 - A1138, 2015. <https://doi.org/10.1137/141000439>
- [17] N. Bonneel, G. Peyré, and M. Cuturi, "Wasserstein barycentric coordinates," *ACM Transactions on Graphics*,

- vol. 35, no. 4, pp. 1 - 10, 2016.
<https://doi.org/10.1145/2897824.2925918>
- [18] P.-T. de Boer, D. P. Kroese, S. Mannor, and R. Y. Rubinstein, "A Tutorial on the Cross-Entropy Method," *Annals of Operations Research*, vol. 134, no. 1, pp. 19 - 67, 2005.
<https://doi.org/10.1007/s10479-005-5724-z>
- [19] L. van der Maaten and G. Hinton. "Visualizing highdimensional data using t-SNE," *Journal of Machine Learning Research*, vol. 9, pp. 2579 - 2605, 2008.
<http://www.jmlr.org/papers/volume9/vandermaaten08a/vandermaaten08a.pdf>
- [20] B. Li, Y. Lu, C. Li, A. Godil, T. Schreck, M. Aono, M. Burtscher, H. Fu, T. Furuya, H. Johan, J. Liu, R. Ohbuchi, A. Tatsuma, and C. Zou. "Extended large scale sketch-based 3D shape retrieval," *Eurographics Workshop on 3D Object Retrieval*, Strasbourg, France, pp. 121 - 130, 2014.
<http://dx.doi.org/10.2312/3dor.20141058>
- [21] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
<https://doi.org/10.1109/cvpr.2016.90>
- [22] S. Ferradans, G.-S. Xia, G. Peyré, and J.-F. Aujol, "Static and Dynamic Texture Mixing Using Optimal Transport," *Scale Space and Variational Methods in Computer Vision*, pp. 137 - 148, 2013.
https://doi.org/10.1007/978-3-642-38267-3_12

◎ 저 자 소 개 ◎



지 명 근(Myunggeun Ji)

2017 B.S. in Computer Science, Kyonggi University, Suwon, Korea
2017~Present : M.S. Student in Computer Science, Kyonggi University, Suwon, Korea
Research Interests : Computer Vision, Augmented Reality
E-mail : jmg2968@gmail.com



전 준 철(Junchul Chun)

1984 B.S. in Computer Science, Chung-Ang University, Seoul, Korea
1986 M.S. in Computer Science(Software Engineering), Chung-Ang University, Seoul, Korea
1992 M.S. in Computer Science and Engineering (Computer Graphics), The Univ. of Connecticut, USA
1995 Ph.D. in Computer Science and Engineering (Computer Graphics), The Univ. of Connecticut, USA
2001.02~2002.02 Visiting Scholar, Michigan State Univ. Pattern Recognition and Image Processing Lab.
2009.02~2010.02 Visiting Scholar, Univ. of Colorado, Wellness Innovation and Interaction Lab.
Research Interests : Augmented Reality, Computer Vision, Human Computer Interaction
E-mail : jchun@kgu.ac.kr



김 남 기(Namgi Kim)

1997 B.S. in Computer Science and Engineering, So-Gang University, Seoul, Korea
2002 M.S. in School of Computing, KAIST, Daejeon, Korea
2005 M.S. in School of Computing, KAIST, Daejeon, Korea
Research Interests : Communication system, Network
E-mail : ngkim@kgu.ac.kr