

렌더링 비교 뉴럴넷 기반 가구 조립 설명서 부품의 6D 자세 추정

*박재우, **강이삭, ***조남익

서울대학교 전기정보공학부 뉴미디어통신연구소(INMC)

*bjw0611@snu.ac.kr, **isaackang@ispl.snu.ac.kr, ***nicho@snu.ac.kr

*Jaewoo Park, **Isaac Kang, ***Nam Ik Cho

Seoul National University Electrical Computer Engineering, INMC

요 약

본 논문에서는 뉴럴넷 기반 렌더링 비교 방식을 사용하여 가구 조립 설명서에 표기된 부품의 자세를 추정하는 방법을 제안한다. 제안하는 방법은 부품의 자세를 임의로 가정한 후, 가정한 자세로 투사한 부품의 영상과 설명서의 부품 영상을 비교하여 두 영상의 부품을 일치시키는 자세 변화를 추정하는 방식으로 진행된다. 또한, 설명서에 반복적으로 모델을 적용하여 부품의 자세를 점차적으로 정확하게 보정하는 방식을 사용하였으며, 네트워크의 구성 및 자세 추정에 사용되는 목표 함수를 다양하게 실험하여 성능을 비교하였다. 본 연구에선 IKEA 의 Stefan 의자 조립 설명서의 부품 데이터셋으로 실험을 진행하였으며, 해당 데이터셋에 대하여 제안하는 방법이 정확하게 자세를 보정함을 확인하였다.

1. 서론

영상 내 물체의 자세 추정 문제는 물체의 CAD 및 카메라 변수가 주어질 때 물체 중심의 3차원 좌표를 카메라 중심의 3차원 좌표로 변환하는 행렬을 추정하는 것을 목표로 한다. 이때, 물체 중심의 3차원 좌표가 변환된 후 카메라 변수를 이용하여 투영되었을 때 영상내 물체의 모습과 일치하는 조건을 만족해야 한다. 또한, 강체인 물체가 3차원 공간상에 있다고 가정하여 변환에 사용되는 변수의 자유도가 회전과 이동에 각각 3으로 고정되기 때문에 6D 자세 추정 문제로도 언급된다. 정확한 물체 자세 추정 방법은 로봇이 영상 내의 물체를 잡아야 하는 상황이나, 영상 내의 물체에 굴곡에 맞게 패턴을 입력하여 하는 증강 현실 등의 다양한 상황에서 필요한 핵심적인 기술이다. 그러나, 물체의 자세 추정 문제는 변수의 자유도가 높고 가려진 부분에 대한 예측 및 물체의 대칭성까지 고려해야 하여 어려운 문제로 고려되고 있다.

기존에는 물체의 자세를 추정하기 위해 깊이 영상(Depth map)을 입력으로 받아서 물체 CAD의 점 집합(Point cloud)과 비교하여 보정하는 ICP(Iterative Closest Point)[1] 알고리즘이 가장 보편적으로 사용되었다. 그러나 깊이 영상 없이 RGB 영상만 주어지거나, 영상에 목표로 하는 물체 이외에 배경이나 다른

물체의 점 집합이 많이 포함되거나, 물체가 다른 물체에 가려져 있는 상황에서는 강인하게 작동하지 못하는 한계를 보였다. 이에 최근 딥러닝 기술이 넓은 범위의 영상 인식 및 검출 문제에 뛰어난 성능을 보임에 따라, 깊이 영상과 영상 이미지를 같이 사용하여 자세를 추정하는 방식들이 제안되었으나, 깊이 영상을 찍는 카메라의 낮은 해상도 문제와 깊이 영상 정보가 주어지지 않는 경우 사용이 불가능한 문제가 있다. 따라서, 깊이 영상을 사용하지 않고 RGB 영상만을 자세 추정 뉴럴넷에 이용하여 뛰어난 성능을 보인 연구들이 최근 발표되고 있다. 영상만 이용하는 뉴럴넷 기반의 물체 자세 추정 방식은 크게 4가지로 분류할 수 있는데, 먼저 물체별로 3D 검출상자(Bounding box)의 여덟 꼭짓점 혹은 물체의 특징점(Key point)들을 물체 중심 3D 좌표에 정의하고, 해당 점들이 입력 영상에 투영된 2D 위치를 추정하도록 학습한 후, 추정된 좌표와 사전에 정의한 3D 좌표가 일치하도록 PnP(Perspective-n-Point) 알고리즘[22]을 기반으로 변환행렬을 추정하는 방식들이 제안되었다 [2, 3, 4, 5, 6, 7, 8, 9]. 두번째 방식은 물체의 3D 표면에 UV-map 을 정의한 후, 뉴럴넷을 이용하여 입력 영상에서 물체의 UV-map 을 추정하고 마찬가지로 PnP + RANSAC(Random Sampling Consensus) [22]을 기반으로 영상에서 추정한 값과 3D 물체 표면에 정의한 UV-map 이 일치하도록 변환하는 행렬을 찾는 방식이다 [10, 11,

12, 13, 14]. 그러나, 위의 방식들은 검출 대상인 특징점이나 UV-map 이 일부분 가려지는 경우 자세 추정 오차가 크게 발생하는 문제가 있다. 세 번째 방식은 물체를 사전에 여러 방향에서 찍어서 물체의 후보 모습들을 지정한 후, 현재 추정하고자 하는 물체의 모습이 어느 방향에서 취득된 영상과 가장 가까운지 분류하고 해당 방향을 물체의 자세로 추정하는 방식이다 [15, 16, 17]. 이 방식은 물체의 자세를 정해진 구간으로 나누어 분류하기 때문에 구간 내 오차가 발생하여 추정 결과가 정확하지 않다. 또한, 위의 세 가지 방식 모두 물체의 특징점, UV-map 그리고 각도 별 물체의 모습과 같이 추정해야 할 물체의 정보를 사전에 지정하고 학습해야 하므로, 새로운 물체에 대한 추정은 불가능하다는 문제가 있다.

위와 같은 문제를 해결하기 위해, 3D 물체에 사전에 정의한 정보를 사용하지 않고, 실시간으로 물체의 현재 자세를 렌더링(Rendering)하여 입력 영상의 물체와 비교하여 변환행렬을 직접적으로 추정하는 네번째 방식이 제안되었다 [18, 19]. 이들은 앞선 두 가지 방식과 달리 추정 대상이 물체의 사전 정보가 아닌 변환행렬 자체이고, CAD 가 주어지면 물체의 종류에 상관없이 렌더링이 가능하기 때문에, 다양한 모양의 물체의 자세 추정을 학습하면 변환행렬을 구성하는 변수의 규칙을 학습하여 처음 보는 물체에 대해서도 자세 추정이 가능하다는 장점이 있다. 또한, 자세 추정을 반복적으로 적용하여 점차 정확하게 추정하는 방식을 이용하여 BOP Challenge 와 같은 벤치마크 자세 추정 대회에서도 타 방식에 비해 뛰어난 정확도를 나타내었다. 따라서, 본 논문에서도 위의 네번째 방식을 기반으로 IKEA 의 가구 조립 설명서에 표기된 조립 부품의 자세를 추정하는 뉴럴넷을 실험하였다. 이를 위해, 가구 부품을 다양한 자세로 투영하여 생성한 논문의 데이터셋을 이용하여 학습하였으며 [20], 기존 렌더링 비교 논문에 제시된 목적 함수 외에 추가적으로 다른 목적 함수를 실험하여 결과를 비교하였다.

2. 실험 방법

2.1 데이터셋

본 논문에서는 [20]에서 제안한 IKEA 사의 Stefan 의자 조립에 사용되는 부품으로 구성된 데이터셋을 사용하였다. 기존 Stefan 조립 설명서는 부품의 자세 정보가 주어지지 않아 학습 및 실험에 사용할 수 없기 때문에 [20]에서와 같이 Stefan 을 구성하는 부품의 CAD 를 지정한 자세가 되도록 변환하고 렌더링하여 정답 자세를 아는 데이터셋을 생성한 후, 이를 학습과 성능 평가용으로 나누어 이용하였다. 정확하게는 총 731 장의 렌더링 영상을 학습에 사용하였고 85 장을 성능 평가에 사용하였으며, Figure 1. 예시와 같이 조립도 영상은 한 장에 조립되는 부품들이 여러 개 표기되어 있으므로 부품 개수로 확인하였을 때 각각 총 1935 개의 부품 영상을 학습에, 총

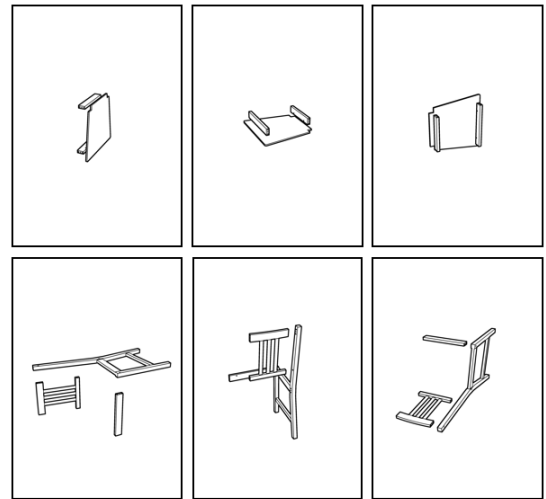


Figure 1. 사용한 데이터셋 영상 예시

225 개의 부품 영상을 성능 평가에 사용하였다. 또한, 조립 영상의 현실성을 높이기 위하여 조립되는 부품간 위치 관계를 기존 Stefan 과 동일하게 사용하되 투영한 각도 및 조립 순서를 다양하게 변경하여 구성하였다.

2.2 자세 인식기 작동 방식

본 논문에서는 [18]과 [19]에서 제안된 물체 자세 인식 방법을 IKEA Stefan 부품에 적용하여 실험을 진행하였다. 위의 두 논문은 모두 물체의 현재 자세를 렌더링한 영상과 자세를 추정하고자 하는 물체가 있는 영상을 비교하여 점차 그 차이를 줄이는 변환행렬을 추정하는 것을 목표로 하며, 이를 렌더링 비교(Render-and-compare) 방식이라 한다. 위 방식은 크게 세 가지로 구성되어 진행되는데, 각각 렌더링 단계, 자세 변화 추정 단계, 그리고 자세 보정 단계로 이루어지며, 각각의 단계별 입력과 출력을 Figure 2. 에 표기하였다. 먼저 렌더링 단계는 물체의 CAD 와 변환행렬을 입력으로 받아서 물체를 변환행렬에 맞게 투영했을 시의 영상을 생성하는 단계로, 이 때, 카메라의 초점거리(Focal length)와 주점(Principal point)과 같은 변수는 모두 고정되게 주어진다 가정하며, 비대칭계수(skew coefficient)는 0 으로 가정한다. 본 논문에서는 렌더링 단계에 pytorch3d[23]의 함수를 이용하여 실험하였으며, 텍스처는 전체가 균일한 검은색으로 설정하였다. 이어서 렌더링 한 영상과 추정하고자 하는 설명서 영상을 쌓아서 자세 추정기에 입력한다. 이 때, 흰 색 배경에 검은색 선으로 구성된 조립 설명서의 영상에서 선을 검출하여 사용하기 위해 설명서의 밝기 값을 역전하여 합쳐서 아래 그림과 같이 사용하였다.

이 후, 통합한 영상을 자세 추정기에 입력하여 현재 자세의 변환행렬(그림 상의 R_i, t_i)에서 추정하고자 하는 물체의 자세로 보정하기 위한 자세 보정 변환행렬(그림 상의 $\Delta R_i, \Delta t_i$)를 추정한다. 이어서, 기존 변환행렬에 추정한 자세 보정 변환행렬을

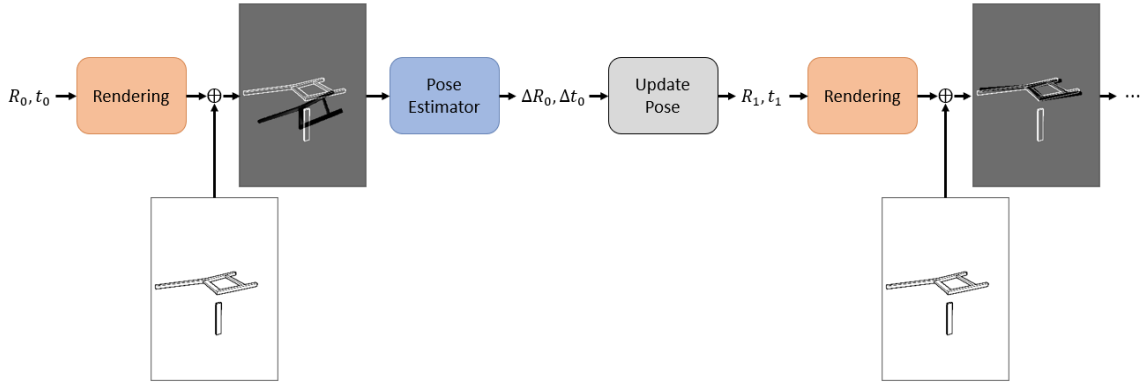


Figure 2. Render-and-compare 기반 자세 추정 방식

반영하여 수정된 변환행렬 (그림 상의 R_{i+1}, t_{i+1})을 구한다. 위 작업을 총 N 번 반복하여 최종적인 변환행렬을 구하는 방식으로 진행되며, 본 논문에서는 1 번만 반복하여 진행하였다.

2.3 자세 인식기 출력 형식

회전행렬 R_i 와 변이행렬 t_i 는 얽혀 있는(Entangled) 관계이기 때문에, 동일한 값을 가지는 변이행렬도 회전행렬에 따라 다른 이동 결과를 보이고, 회전행렬은 주기에 따른 모호성이 있기 때문에 회전행렬과 변이행렬을 바로 추정하기는 어렵다. 따라서, 본 논문에서도 회전행렬과 변이행렬을 뉴럴넷에서 바로 출력하기보다 간접적인 값을 먼저 추정한 후, 후처리 작업을 통해 구하였다. 먼저, 회전을 나타내는 두 가지 다른 출력 방식을 채택하여 각각 실험하여 비교하였는데, 첫번째로 [12, 13, 18, 19]와 같이 많은 논문에서 회전행렬 추정 시 사용되는 사원수(Quaternion)를 출력하고 회전행렬로 변환하는 방식과 [21]에서 회전 공간에서의 모호성을 해소하기 위해 제안된 두 회전축 추정 방식을 실험하였다. 두 회전축 추정 방식은 자세 추정기에서 회전에 관한 두 3차원 벡터 e_1 과 e_2 를 출력할 때,

$$e'_1 = \frac{e_1}{\|e_1\|_2} \quad (1)$$

$$e'_3 = \frac{e'_1 \times e_2}{\|e_2\|_2} \quad (2)$$

$$e'_2 = e'_3 \times e'_1 \quad (3)$$

와 같이 3 개의 축(e'_1, e'_2, e'_3)을 구하여 회전행렬을 구성한다.

변이행렬 t_i 또한 [19]에서 제안한 풀린(Disentangled) 좌표 출력 방식을 채택하였다. 풀린 좌표 예측 방식은 변이행렬 값을 직접적으로 예측하지 않고 간접적으로 예측하는 방식을 사용한다. 이는 먼저 영상 속에서 물체가 몇 픽셀 위, 아래 또는 좌우로 이동해야 하는지, 그리고 물체의 크기가 현재 렌더링 된 것에 비하여 크기를 어느정도 비율로 변경해야 하는지 예측한 후 위 값 들을 변이행렬로 변환하는 과정을 거친다. 이 때 i 번째 자세 추정에서 픽셀단위의 상하이동 추정 값을 v_i^y , 좌우이동 추정

값을 v_i^x 라 하고 주어진 초점거리를 f 라 하면,

$$\begin{pmatrix} v_i^x \\ v_i^y \end{pmatrix} = \begin{pmatrix} x_{i+1} - x_i \\ y_{i+1} - y_i \end{pmatrix} = \begin{pmatrix} f \left(\frac{t_{i+1}^x}{t_{i+1}^z} - \frac{t_i^x}{t_i^z} \right) \\ f \left(\frac{t_{i+1}^y}{t_{i+1}^z} - \frac{t_i^y}{t_i^z} \right) \end{pmatrix} \quad (4)$$

와 같이 나타낼 수 있으며, 이 때, i 번째 자세 추정에 대해 t_i^x, t_i^y, t_i^z 는 각각 변이행렬의 X, Y, Z 값이고 (x_i, y_i) 는 물체의 영상내 위치를 나타낸다. (4)의 수식에 물체의 크기 변화 비율을 추정한 v_i^z 를 반영하여 변이행렬 t_{i+1} 에 대하여 정리하면,

$$\begin{pmatrix} t_{i+1}^x \\ t_{i+1}^y \\ t_{i+1}^z \end{pmatrix} = \begin{pmatrix} \frac{v_i^x}{f} t_i^z + t_i^x \\ \frac{v_i^y}{f} t_i^z + t_i^y \\ t_i^z \end{pmatrix} v_i^z \quad (5)$$

와 같이 정리하여 변이행렬을 구할 수 있다. 이에 따라 변이행렬을 구하는 문제를 회전행렬과 풀린 좌표이면서 픽셀 단위의 이동(v_i^x, v_i^y)과 물체의 크기 변화 비율(v_i^z)을 추정하는 직관적인 문제로 변형하여 학습하였다. 네트워크 출력 값에서 변이행렬을 구하는 과정은 (5)의 수식을 이용하였다.

2.4 목적 함수

변환행렬을 추정하는 문제의 목적함수 종류는 크게 두 가지로 나뉘는데, 1) 물체의 CAD 점들을 추정한 변환행렬로 이동한 위치와 정확한 변환행렬로 이동한 위치와의 차이, 2) 추정한 변환행렬과 정확한 변환행렬 간의 거리이다. 본 논문에서는 두 가지 목적함수에 따른 결과를 각각 구하여 비교하였다. 기존에 일반적으로 사용되던[18, 19] 점간 간격 목적 함수(Point matching loss)는 아래와 같다:

$$L_i^{pm} = \frac{1}{N} \sum_j \| (R x_j + t) - (\hat{R}_i x_j + \hat{t}_i) \|_1 \quad (6)$$

이 때, R, t 는 각각 물체의 정확한 자세를 나타내는 변환 행렬이고, \hat{R}_i, \hat{t}_i 는 i 번째 추정된 물체의 자세를 나타내며, x_j 는 물체를



Figure 3. 부품 자세 추정 결과

나타내는 CAD 상의 총 N 개의 모든 3D 점들을 각각 의미한다.

변환행렬 간의 거리는 회전행렬 \hat{R}_i 과 R 의 측지 거리(Geodesic distance)와 2.3 절에서 정의한 (v_i^x, v_i^y, v_i^z) 의 추정 값과 정답 간의 거리를 같이 사용하였다. 정확하게는 회전 행렬 간의 측지 거리와 (v_i^x, v_i^y, v_i^z) 간의 거리를

$$L_i^{geodesic} = \|\log(\hat{R}_i R^T)\|_F \quad (7)$$

$$L_i^v = \left\| \frac{v_i^x - \hat{v}_i^x}{W} \right\|_1 + \left\| \frac{v_i^y - \hat{v}_i^y}{H} \right\|_1 + \|v_i^z - \hat{v}_i^z\|_1 \quad (8)$$

로 지정하여 사용하였으며, 이 때 (v_i^x, v_i^y, v_i^z) 는 i 번째 자세 추정에서의 정답을, $(\hat{v}_i^x, \hat{v}_i^y, \hat{v}_i^z)$ 는 추정된 값을 나타내고, W 와 H 는 영상의 너비와 높이를 나타낸다.

3. 실험 결과 및 분석

3.1 평가 방법

본 논문에서는 물품 자세 추정 방식을 평가할 때 사용되는 대표적인 평가 지수인 ADD 점수와 대칭성이 있는 물품의 자세 추정 정확도를 평가할 때 사용되는 ADD-S 점수를

$$ADD = \frac{1}{N} \sum_{x_j \in \mathcal{N}} \|(Rx_j + t) - (\hat{R}_i x_j + \hat{t}_i)\|_2 \quad (9)$$

$$ADD - S = \frac{1}{N} \sum_{x_j \in \mathcal{N}} \min_{x_k \in \mathcal{N}} (\|(Rx_k + t) - (\hat{R}_i x_j + \hat{t}_i)\|_2) \quad (10)$$

와 같이 사용하여 ADD(-S) 점수가 물체 지름의 10% 미만인 비율을 구하여 실험결과를 도출하였다. 본 실험에 사용된 Stefan 부품 중 1,2,3 번은 대칭성을 띄고 있으므로 ADD-S 점수를, 나머지 부품 5 종류는 ADD 점수를 기반으로 평가하였다.

3.2 결과분석

기존 [18, 19]에서와 같이 부품이 검출되어 있다는 가정하에 단위행렬을 회전변환 초기자세로 설정하여 실험을 진행하였으며, ResNet18 모델을 기본 구조로 사용하였다. Figure 3. 은 두 회전축 출력방식에 점 간 거리 목적함수 학습의 추정 예시이다. 가장 왼쪽이 입력 설명서 영상, 두번째가 정답 자세, 세번째가 초기 자세, 네번째가 보정한 자세 결과를 나타낸다. 또한, 아래 Table 1. 은 회전변환의 출력 방식 별 성능과 목적 함수 별 성능이다.

Table 1. 출력 형식별, 목적 함수 별 자세 추정 결과

출력 형식, 목적 함수	사원수, L_i^{pm}	$(e'_1, e'_2, e'_3), L_i^{geodesic} + L_i^v$	$(e'_1, e'_2, e'_3), L_i^{pm}$
부품 1	86.67%	93.33%	100.00%
부품 2	90.00%	64.00%	88.00%
부품 3	84.00%	56.00%	84.00%
부품 4	95.00%	40.00%	90.00%
부품 5	80.00%	65.71%	94.29%
부품 6	82.86%	65.71%	88.57%
평균	86.67%	60.89%	88.44%

실험 결과는 두 회전축 출력 방식이 사원수 방식보다 성능이 높았고, 목적 함수는 측지거리와 v 간 거리를 사용하는 경우 보다 점 간 거리를 사용하는 방식이 성능이 더 높았다.

4. 결론

본 논문에서는 렌더링 비교 방식을 이용하여 가구 조립 설명서의 부품의 자세를 인식하는 실험을 진행하였으며, 회전 변환 출력 방식과 목적 함수 별 성능을 비교하였다.

감사의 말

이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(2021R1A2C2007220). This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (2021R1A2C2007220).

참고문헌

- [1] Besl, Paul J., and Neil D. McKay. "Method for registration of 3-D shapes." *Sensor fusion IV: control paradigms and data structures*. Vol. 1611. International Society for Optics and Photonics, 1992.
- [2] Hu, Yinlin, et al. "Segmentation-driven 6d object pose estimation." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2019.
- [3] Pitteri, Giorgia, et al. "On object symmetries and 6d pose estimation from images." *2019 International Conference on 3D Vision (3DV)*. IEEE, 2019.
- [4] Rad, Mahdi, and Vincent Lepetit. "Bb8: A scalable, accurate, robust to partial occlusion method for predicting the 3d poses of challenging objects without using depth." *Proceedings of the IEEE International Conference on Computer Vision*. 2017.
- [5] Hu, Yinlin, et al. "Single-stage 6d object pose estimation." *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2020.
- [6] Peng, Sida, et al. "Pvnet: Pixel-wise voting network for 6dof pose estimation." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2019.
- [7] Oberweger, Markus, Mahdi Rad, and Vincent Lepetit. "Making deep heatmaps robust to partial occlusions for 3d object pose estimation." *Proceedings of the European Conference on Computer Vision (ECCV)*. 2018.
- [8] Song, Chen, Jiaru Song, and Qixing Huang. "Hybridpose: 6d object pose estimation under hybrid representations." *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2020.
- [9] Fu, Mingliang, and Weijia Zhou. "DeepHMap++: Combined projection grouping and correspondence learning for full DoF pose estimation." *Sensors* 19.5 (2019): 1032.
- [10] Park, Kiru, Timothy Patten, and Markus Vincze. "Pix2pose: Pixel-wise coordinate regression of objects for 6d pose estimation." *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2019.
- [11] Wang, He, et al. "Normalized object coordinate space for category-level 6d object pose and size estimation." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2019.
- [12] Wang, Gu, et al. "GDR-Net: Geometry-Guided Direct Regression Network for Monocular 6D Object Pose Estimation." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021.
- [13] Zakharov, Sergey, Ivan Shugurov, and Slobodan Ilic. "Dpod: 6d pose object detector and refiner." *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2019.
- [14] Li, Zhigang, Gu Wang, and Xiangyang Ji. "Cdpn: Coordinates-based disentangled pose network for real-time rgb-based 6-dof object pose estimation." *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2019.
- [15] Kehl, Wadim, et al. "Ssd-6d: Making rgb-based 3d detection and 6d pose estimation great again." *Proceedings of the IEEE international conference on computer vision*. 2017.
- [16] Tekin, Bugra, Sudipta N. Sinha, and Pascal Fua. "Real-time seamless single shot 6d object pose prediction." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018.
- [17] Sundermeyer, Martin, et al. "Augmented autoencoders: Implicit 3D orientation learning for 6D object detection." *International Journal of Computer Vision* 128.3 (2020): 714-729.
- [18] Li, Yi, et al. "Deepim: Deep iterative matching for 6d pose estimation." *Proceedings of the European Conference on Computer Vision (ECCV)*. 2018.
- [19] Labbé, Yann, et al. "Cosypose: Consistent multi-view multi-object 6d pose estimation." *European Conference on Computer Vision*. Springer, Cham, 2020.
- [20] 강이삭, 조남익 "UV-map 기반의 신경망 학습을 이용한 조립 설명서에서의 부품의 자세 추정", 한국방송공학회 하계학술대회, 2020 July.
- [21] Zhou, Yi, et al. "On the continuity of rotation

representations in neural networks." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2019.

[22] Fischler, Martin A., and Robert C. Bolles. "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography." *Communications of the ACM* 24.6 (1981): 381-395.

[23] Ravi, Nikhila, et al. "Accelerating 3d deep learning with pytorch3d." *arXiv preprint arXiv:2007.08501* (2020).