

정육로봇을 위한 자유형상 물체의 자세 추정 및 절단경로 추적

박연출¹, 박세준², 이용준^{3*}

¹LG전자 로봇선행연구소, ²SK 정보보호담당, ³극동대학교 해킹보안학과

Pose Estimation of Free-form Object and Cutting Path Tracking for Butcher Robot

Yeon-Chool Park¹, Se-Joon Park², Yong-Joon Lee^{3*}

¹Advanced Robotics Lab., LG Electronics

²Information Security Group, SK

³Department of Hacking Security, Far East University

요약 본 논문은 육류 가공 산업에서 로봇을 이용한 소 근육 분리의 자동화에 관한 연구로 실시간으로 자유형상(Free-form)의 변형 가능한 소 근육을 인식하고 로봇 팔을 이용한 절단 시 절단면의 절단 경로(Cutting Path)를 추출하고 동적으로 근육 표면과 절단 경로의 변형을 추적하는 비전 시스템을 제안한다. 본 논문은 3D Point cloud로부터 Spin-image를 생성하여 소 근육 인식에 활용하는 프레임워크를 제안하였고, 올바른 자세 추정을 위해 추출된 특징점을 통계적 기법을 통해 Outlier를 제거하는 방법, 가상의 절단 경로 실린더와 곡률을 이용하여 가장 깊은 절단면을 찾고 추적하는 방법을 제시한다. 궁극적으로 지향하는 시스템은 본 과정을 통해 분리 절단면을 실시간으로 추적하여 정육 로봇이 고기를 손상시키지 않고 근육을 올바르게 분리할 수 있도록 돕는 데 있다.

Abstract This research studies the automation of bovine muscle separation using a robot in the meat processing industry. We propose a vision system that recognizes deformable bovine muscles in real-time, extracts the cutting path on the surface using a robotic manipulator, and dynamically tracks the deformation of the muscle surface and cutting path. The contributions of this paper are threefold. First, it defines a framework to recognize the bovine muscle using a spin image generated from a 3D point cloud. Second, it presents a method to remove the outliers from the feature points extracted for correct posture estimation through statistical techniques. Last, it proposes a method of finding and tracking the deepest cutting-guide path using a virtual cylinder and curvature information. The system's ultimate goal is to track the separation cuts in real-time, and through this process, help a robot butcher separate the muscle correctly without damaging the meat.

Keywords : Free-Form, Deformable, Recognition, Pose Estimation, Tracking, Manipulation

1. 서론

본 논문은 육류 가공 산업에서 로봇을 이용한 소 근육 분리의 자동화에 관한 연구로, 로봇 팔을 이용하여 소 근육 절단 시 자유 형상(free-form)의 소 근육을 인식하

고, 절단에 의한 근육 표면의 동적 변형에 적응하면서 실시간으로 절단 가이드 경로(cutting guide path)를 추정 및 추적하는 실시간 비전 시스템을 개발하는 것을 목표로 한다.

비전 분야에서 Free-form/Deformable 물체 인식

*Corresponding Author : Yong-Joon Lee(Far East Univ.)

email: 2020032@kdu.ac.kr

Received November 4, 2021

Accepted January 7, 2022

Revised December 29, 2021

Published January 31, 2022

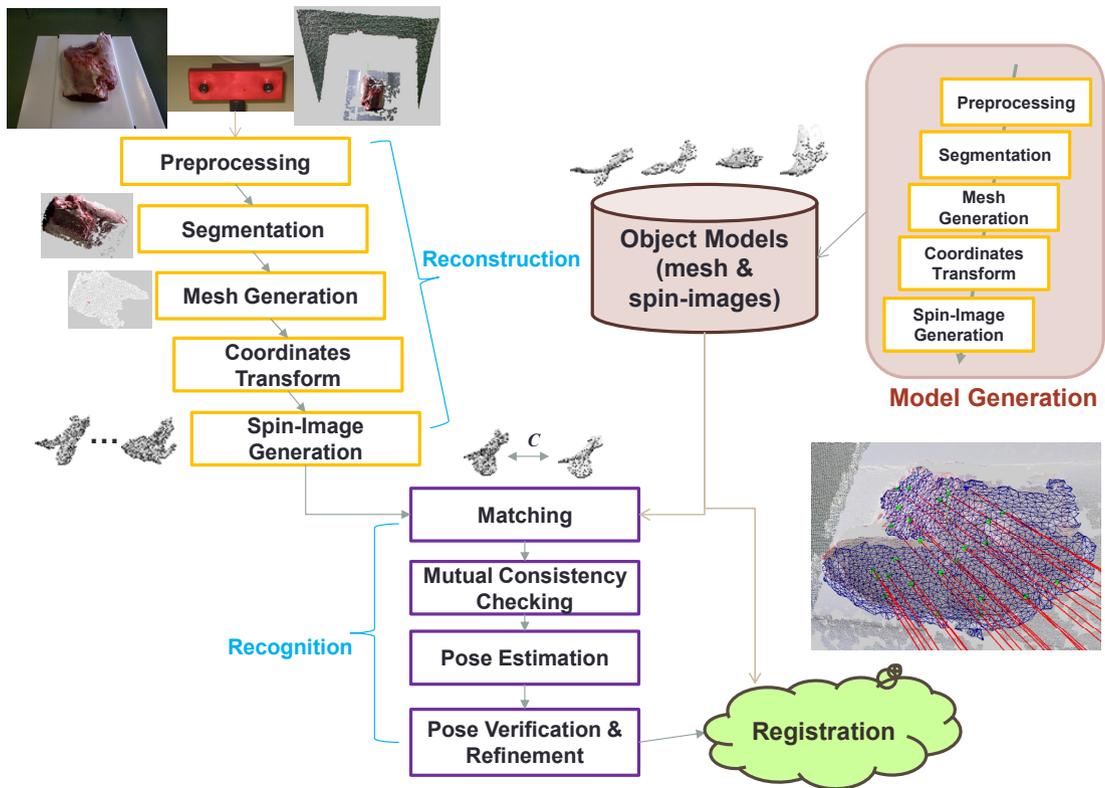


Fig. 1. Overall architecture of free-form object recognition and pose estimation

및 추적은 오랫동안 가장 어려운 연구 주제 중 하나로 여겨져 왔다. 특히 우리가 목표로 하는 소 근육과 같은 물체는 절단시 매우 비선형적인 움직임을 보이는 Free-form 물체이므로 인식, 자세 추정 및 추적 기술을 연구하기에 매우 도전적인 과제 중 하나이다.

Free-form 물체의 인식에 관한 대표적인 연구로는 Andrew Johnson에 의해 처음 제시된 임의의 3차원 표면 형태를 2차원 객체 지향 좌표계로 변환하여 물체를 인식하는 Spin-image가 있다[1]. Spin-image는 임의의 3차원 표면을 2차원 벡터로 변환할 수 있기에 형태 특징 정보를 그대로 유지할 수 있고, 자세 변화에 강인하고 매칭 속도가 빨라 Free-form 물체 인식에 널리 활용되어 왔다.

Körtgen은 3D Shape Contexts를 이용하여 3D shape를 매칭하고 검색하는 방법을 제안하였는데 3D Shape Context는 Spin image와 유사하게 물체 표면의 Surface point를 중심으로 Semi-local 특성을 식별자(descriptor)로 활용하는 방식이다[2]. 또 다른 지역 식별자로는 Radu Rusu에 의해 Fast Point Feature

Histogram(FPFH)이 제시되었는데 3차원 좌표 거리와 법선벡터의 각도차를 구하고 이를 16개의 bin(bin)에 누적하여 히스토그램(histogram)을 구하는 방식이다. 이 히스토그램을 식별자로 활용하여 표면 정합을 수행하는 방식이다[3].

Hui Chen과 Bir Bhanu는 Local Surface Patch 식별자를 제안하였고[4], 이를 이용하여 3D point cloud로부터 Free-form 물체를 인식하며, Joel Vidal은 Bertram Drost에 의해 제안된 Point Pair Feature [5]를 이용하여 3D Free-form 물체를 인식하는 방법을 제안하였다[6].

최근에는 딥러닝 기법을 이용하여 3차원 Point cloud의 물체 종류를 인식하는 방법들이 고안되고 있는데 PointNet/PointNet++이 대표적인 예이며 [7,8], 이후 Sparse tensor를 활용하여 7차원 공간(space-time-chroma)에서 Semantic Segmentation을 수행하는 연구 [9]와 Point cloud에 Transformer 기법을 적용하여 3차원 물체를 인식하는 연구로 이어지고 있다[10].

본 논문은 소 근육을 인식하고 자세를 추정하기 위해

서 RGBD 센서를 이용하여 Free-form surface를 3차원으로 복원하고 Spin-image를 생성하여 사전 제작된 3차원 모델과 정합, 실시간으로 소 근육의 자세를 추정하는 방식을 최초로 제안하였고, 추정된 자세 정보를 이용하여 해당 물체의 절단 경로(cutting path)를 실시간으로 추정 및 추적하는 방법을 제시한다.

구체적인 본 논문의 기여는 다음과 같다.

- 소 근육을 사전 제작된 Spin-image 모델과 비교하여 인식 및 자세를 추정하고, 모델에 정의된 실린더 형상의 절단 가이드 경로를 기반으로 최적의 절단 경로를 찾고 추적하는 자유 형상 물체(소 근육)의 자세 추정 및 절단 경로 추정 프레임워크를 최초로 제안한다.
- 모델로부터 얻어진 Spin-image와 Point cloud로부터 얻어진 Spin-image를 비교해 최상의 대응쌍을 구하는 과정에서 Spin-image의 겹침(overlap) 모양과 각 bin의 강도(intensity)를 함께 고려한 통계적 유사도 측정 방법을 제시한다.
- 다수이면서 최적의 3차원 대응점을 구하기 위해 Voting에 기반하여 특이값(outlier)을 제거하는 확률적 상호 대응쌍 일관성 검증 방법을 제시한다.
- 절단 경로를 추정하기 위해 가상의 3차원 실린더를 연결하고, 실린더 내에서 지역적 곡률이 가장 큰 지점을 찾아 3차원 곡선 피팅(curve fitting)한 후 Non-rigid ICP 기법에 의해 정합하고 경로를 갱신하는 방법을 제안한다.

본 논문의 구성은 다음과 같다. 2장에서는 Spin-image를 이용한 3차원 Free-form 물체 인식 및 자세 추정 방법에 대하여 제안하고, 3장에서는 절단면의 경로를 검출 및 추적하기 위하여 실린더 형태의 가상 절단 경로를 생성하는 방법을 제시하며, 4장에서는 결론 및 향후 연구 계획에 대하여 언급한다.

2. 소 근육 인식 및 자세 추정

2.1 소 근육 인식 및 자세 추정 시스템

제안된 소 근육 인식 및 자세 추정 시스템은 Fig. 1과 같이 구성된다. 시스템의 전체 흐름은 다음과 같다. 첫째, RGBD 센서를 이용하여 Point cloud를 획득하고 영상에서 테이블의 법선벡터(surface normal) 정보를 이용하여 테이블로부터 고기 부분 개체를 분리(segmentation)

한다. 여기서 물체 표면 패치는 다각형 메쉬(mesh)로 변환되고 모든 방향 포인트는 Spin-image로 변환된다. Spin-image는 좌표에 위치한 각 bin의 포인트 수를 누적하여 계산된 일종의 2D Hash 테이블이다. 우리가 특징 정보로 Spin-image를 선택한 이유는 특정 위치에서 바라본 여타 3차원 포인트의 위치와 밀도를 2차원 영상으로 투영해서 모든 지점에서의 상대적 형상을 자세 변화에 강인(pose invariant)하게 지정한 범위 내에서 확인할 수 있다는 점, 그리고 이로 인해 모든 지점의 2차원 영상을 갖고 있다면 2차원 영상만으로 물체의 3차원 형상을 추정할 수 있다는 점과 Point cloud와 달리 정렬되어 있으므로 이로 인해 얻을 수 있는 계산 속도 향상의 이득에 있다.

모든 방향 포인트에 대한 모든 Spin-image가 생성되면 인식과 자세추정을 위해 특정 위치에서 사전 제작된 3차원 물체 모델의 Spin-image와 일치 여부를 판별해야 한다. 여기서 객체 모델에서 생성된 Spin-image는 대상 표면과 동일한 시퀀스를 갖는다. 그런 다음 상관관계수 계산 프로세스에서 지정된 비용 함수를 기반으로 3차원 물체 모델과 소 근육(bovine muscle)의 3차원 표현(representation) 사이의 최상의 대응 Spin-map을 선택한다. 그러나 여전히 잘못된 대응점이 있을 수 있으므로 본 논문에서 제안하는 상호 대응 일관성(Mutual Correspondence Consistency) 척도 확인 방법을 통해 다시 확인한다. 이 과정에서 RANSAC 접근법 [11]을 이용하여 모든 대응 중에서 4개의 대응이 선택되며, 정합 비용 함수가 특정 임계값을 충족하면 반복을 종료한다. 그리고 반복이 끝나면 자세 변환 행렬을 추정한다. 마지막으로 반복 ICP 알고리즘 [12]에 의해 소 근육의 자세를 정밀하게 재추정하는 refinement 정합 과정을 거쳐 최종 변환 매트릭스가 생성되면 표면 추적 모듈로 전달하여 절단 경로(cutting path)를 감지할 수 있도록 한다.

2.2 3차원 Free-form 물체 인식 및 자세 추정

이 절에서는 소 근육을 인식하고 자세를 추정하는 과정 중 특징정보인 Spin-image를 생성하는 방법과, 생성된 Spin-image를 매칭하여 유의미한 대응점을 찾아내는 방법, 그리고 자세를 추정하고 정합하는 과정을 상세히 설명한다.

2.2.1 Spin-image 생성

우선 Spin-image를 생성하기 위해서는 법선 벡터와 정렬된 포인트(point)가 필요하므로 메쉬 생성과정이 선

행적으로 필요하다. 본 논문에서는 메쉬를 생성하기 위해 Ohtake의 메쉬 생성 방법을 사용한다[16]. Ohtake의 메쉬 생성 방식은 Mesh reconstruction, De-noising, Decimation, Sharp feature restoration, Topology cleaning 스테이지로 구성된다.

Spin-image는 Andrew E. Johnson이 처음 소개한 2차원 표면 표현 기술이며 3차원 장면에서 물체 표면(surface) 매칭 및 물체 인식에 사용된다[1]. Spin-image는 뷰어 지향(Viewer-oriented) 좌표계가 아닌 객체 지향(Object-oriented) 좌표계에서 모든 기하학적 표면의 전역적 속성을 인코딩할 수 있다는 장점이 있다. 객체 지향 좌표계는 표면 또는 객체의 특정 위치를 원점으로 고정하는 좌표계이며 뷰어 지향 좌표계는 표면을 향한 관찰자의 시점을 기반으로 하는 좌표계이다. 객체 지향 좌표계를 사용하면 표면이나 객체에 대한 설명이 시점과 무관하여 시점이 변경되더라도 값이 변경되지 않는다는 장점을 갖는다.

Spin-image 생성의 핵심 요소는 Oriented point를 사용한다는 것이다[13]. Oriented points O 는 위치와 방향성을 모두 갖는 3차원 표면 위의 점 집합을 말한다. 또, 이러한 Oriented Points를 갖는 표면은 Vertex를 갖는 다각형 메쉬 M^O 로 표현된다. 표면 메쉬의 Vertex에서 Oriented point O 는 표면 Vertex (p 로 표시)와 표면 법선 벡터 n 으로 구성되는 3차원 점으로 정의되고, 이 p, n 을 정의하면 로컬 좌표계에 해당하는 2차원 basis (p, n)를 공식화 할 수 있다. 여기서 n 에 수직인 접선 평면 P 을 구할 수 있고, 음이 아닌 수직 거리 L 에 의해 원통 좌표계가 생성된다(Fig. 2).

$$S_o : R^3 \rightarrow R^2$$

$$S_o(x) \rightarrow (\alpha, \beta)$$

$$= (\sqrt{\|x-p\|^2 - (n \cdot (x-p))^2}, n \cdot (x-p))$$
(1)

그 다음 이를 이용해 생성되는 Spin-map, S_o 는 3차원 점 x 를 2차원 좌표 (α, β) 로 투영한 것으로 위에서 언급한 (p, n) 원통 좌표계로부터 구할 수 있다 (Eq. (1)).

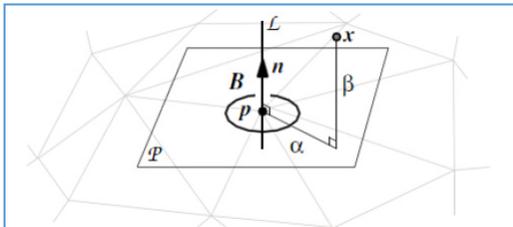


Fig. 2. Cylindrical coordinate system of Spin-image and its (p, n) 2-D basis ([13]참조)

다음 단계는 Spin-image를 나타낼 2차원 배열 (Spin-map)을 만드는 것이다. 이를 위해 2차원 점을 배열의 이산 빈(bin)에 축적하는 방식을 취한다. 2차원 점이 빈에 누적될 때마다 테이블에서 주변 셀을 증가시켜 2차원 배열을 갱신한다. 이 시점에서 데이터의 노이즈를 고려해야 한다. 따라서 2차원 점의 기여도는 Bilinear interpolation에 의해 2차원 배열의 4개의 주변 빈으로 분산된다. 표면을 전체적으로 표현하기 위해 표면 메쉬에 정의된 모든 Oriented point에 대해 이 과정을 반복하여 Spin-image를 생성한다.

2.2.2 매칭 및 일관성 검증

다음 단계로 사전 제작된 여러 모델의 Spin-image들과 런타임시 생성된 대상물체의 Spin-image들을 매칭하여 유사도를 측정한다.

Andrew E. Johnson이 제안한 원래의 Spin-image 매칭 방법은 가려짐이 많은 복잡한 환경에서 물체를 찾기 위한 거리 측정 방법으로 Spin-image 간의 중첩된 영역의 대응만을 정합에 사용했다. 그러나 겹친 빈의 상관관계만 가지고는 충분한 식별력을 제공하지 않기에 실제로는 유사한 Spin-image 사이에서 중첩된 빈이 충분하더라도 상관계수가 높게 나오지 않는다. 불충분한 조명, 노이즈, 고기의 두께 차이 등이 이러한 문제를 야기시킨다.

따라서 이러한 문제를 해결하려면 매칭시 Spin-image의 겹침(overlap) 모양과 각 빈의 강도(intensity)를 함께 다루는 접근 방식을 고려해야 한다.

우리는 이 목적을 달성하기 위해 두 가지 접근 방식을 사용하였다: 첫째, Stochastic distance, 둘째, Bhattacharyya distance[14].

a) Stochastic Distance

일반적으로 물체 인식에 Bayesian 확률을 활용할 때는 $P(feature_{object} | feature_{model})$, 즉 model의 특징정보가 주어졌을 때 물체의 특징정보가 발견될 확률을 사용한다. 그러나 베이지안 확률에서 $P(B|A)$ 는 $P(B|\bar{A})$ 항을 포함하지 않는다 (Eq. (2)). 즉, 입력 영상에는 특정 특징정보가 존재하지 않음에도 모델에서는 해당 특징정보가 발견될 경우의 확률을 고려하지 않는다. 이것은 모델이 대응 가능한 모든 경우의 수를 고려하여 생성되지 않는다면 확률이 항상 1에 가까운 값이 나올 수 있음을 의미한다. 이로 인해 Spin-image 매칭에 곧바로 적용할 수가 없다.

$$P(B | A) = \frac{P(A | B)P(B)}{P(A | B)P(B) + P(A | \bar{B})P(\bar{B})} \quad (2)$$

따라서, 본 논문에서는 Spin-image 매칭에 적합하도록 - 모델은 갖고 있지 않은 특징을 인식대상은 갖는 경우를 고려하여 - 조건부 확률 $\psi_{m,s}$ 를 재정의하였다 (Eq. (3)).

식을 축약하면 $P(s \cap m) / P(s \cup m)$ 과 같고, 여기서 s 는 입력 Spin-image의 단일 빈을, m 은 모델 Spin-image의 단일 빈을 의미한다.

$$\begin{aligned} \psi_{m,s} &= \frac{P(s | m)P(m)}{P(s | m)P(m) + P(s | \bar{m})P(\bar{m}) + P(m | \bar{s})P(\bar{s})} \\ &= \frac{P(s \cap m)}{P(s) + P(m) - P(s \cap m)} = \frac{P(s \cap m)}{P(s \cup m)} \end{aligned} \quad (3)$$

최종적으로 매칭에 활용되는 식은 겹침(overlap) 유사도 Eq. (3)과 강도(intensity) 상관계수의 곱으로 이루어진다. 수정된 방정식 $P_{m,s}$ 는 다음과 같다.

$$P_{m,s} = \psi_{m,s} \times \rho_{m,s} \quad (4)$$

여기서 강도(intensity)를 입력값으로 사용하는 상관계수 $\rho_{m,s}$ 는 Eq. (5)와 같으며, m_i 는 i 번째 모델 빈(bin)의 누적값, s_i 는 입력 Spin-image의 i 번째 빈의 누적값, \bar{m} , \bar{s} 는 각각 모델과 입력 영상 Spin-image의 평균 강도(intensity)를 의미한다.

$$\begin{aligned} \rho_{m,s} &= \frac{\sum_{i=1}^n (m_i - \bar{m})(s_i - \bar{s})}{(n-1)\sigma_m\sigma_s} \\ &= \frac{\sum_{i=1}^n (m_i - \bar{m})(s_i - \bar{s})}{\sqrt{\sum_{i=1}^n (m_i - \bar{m})^2 (s_i - \bar{s})^2}} \end{aligned} \quad (5)$$

b) Bhattacharyya distance

거리 측정을 위해 사용한 두 번째 거리 측정 방법은 Bhattacharyya 계수이다. Bhattacharyya 계수는 겹침(overlap) 정도를 이용하여 두 통계 샘플 간의 상대적 근접성을 판별하는 척도이다. 기본 수식은 Eq. (6)과 같다.

$$\beta = \sum_{i=1}^n \sqrt{\sum \alpha_i \cdot \sum \beta} \quad (6)$$

여기서 i 는 Partition이 있는 경우 각 Partition의 색인을 의미하지만, 본 논문에서는 n 은 Spin-image 전체 빈

(bin)의 수 512x512이고, i 는 각 빈(bin)의 색인을 의미한다. 즉, $\sum \alpha_i$, $\sum \beta_i$ 는 입력 장면(input scene) Spin-image와 모델 Spin-image 각각의 i 번째 bin에 있는 정규화된(normalized) 누적값을 의미한다. 따라서 이 공식은 두 샘플의 겹침이 많을수록 값이 더 크게 되고, 겹침이 전혀 없으면 0이 되는 특성을 갖는다.

c) 상호 대응쌍 일관성 검증

이제 Spin-image 매칭을 통해 대응쌍들을 찾았다면 남은 과정은 매칭 과정에서 발생한 모호한 매칭을 제거하는 것이다. 매칭 과정에서 다수의 잘못된 대응쌍들이 발생한다. 이는 신뢰할 수 있는 일대일 대응을 보장하지 않는다는 것을 의미하고, 자세 추정의 오류를 야기시킨다. 이는 인식을 위해 Spin-image를 사용할 때의 단점 중 하나이다. 또 이로 인해 조합 폭발(combinatorial explosion)에 빠질 수 있다. 따라서 [15]는 거리 기반의 일관성 검증 알고리즘을 제안하였고, 거리값으로 일관성을 검증하고 두 쌍 (또는 세 쌍)의 대응을 찾아서 클러스터로 그룹화하는 방식을 사용하였다.

그러나 이 방식에는 두 가지 단점이 있다. 계산 부담과 적은 수의 대응(correspondence)만 살아남는 것이다. 때때로 우리 실험에서는 이 방식을 적용했을 경우 겨우 2~3 개의 쌍만 살아 남았다. 또한 회전 모호성으로 인해 발생하는 문제를 필터링하지 못한다.

따라서 우리는 대신 Voting에 기반한 확률론적 접근 방식을 제안한다. 제안된 확률적 일관성 검사 방법의 원칙은 대응쌍이 일관성 있게 양방향으로 인접 지역을 가리켜야 한다는 것이다. 즉, 대응쌍이 동일한 인접지역으로 모여야 한다는 것을 전제로 한다.

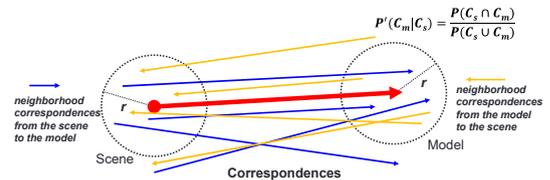


Fig. 3. Concept of consistency checking method

이 접근 방식의 장점은 단순하지만 안정적이고 빠르며 많은 대응쌍을 유지시킬 수 있다는 점이다. Fig. 3은 접근 방식의 아이디어를 설명한다. 여기서의 기본 원칙은 상호간에 같은 쌍, 같은 지점을 가리켜야 한다는 것이다. 따라서 방정식은 Eq. (3)과 같이 어느 한쪽 방향만 조건을 만족했을 때 확률값이 1에 가까워지는 문제점을 방지

Table 1. Mutual consistency checking of pair-wise correspondences

Algorithm
1. Sort all the correspondences according to the distance value calculated in matching sequence.
2. C_i for all correspondences
2.1 Search all neighborhoods for each correspondence using Euclidean distance threshold t .
2.2 H_k for all neighborhoods in C_i
A. Calculates the consistency probability P'_i of C_i
B. Representative correspondence $R_i = \underset{i}{\operatorname{argmax}}(P'_i)$
C. Remove other correspondences except the representative correspondence.
D. If a representative correspondence is selected in a group, the correspondence is copied as same as the number of neighbors in the group to preserve the amount of correspondences.

하기 위해 $P(B|\tilde{A})$ 를 포함해야 하며, Eq. (7)처럼 표현할 수 있다.

$$P'(c_m|c_s) = \frac{P(c_s \cap c_m)}{P(c_s \cup c_m)} \quad (7)$$

여기서 C_m 은 임의의 대응쌍을 기준으로 모델의 일정 반경 r 내에 존재하는 모든 이웃하는 대응쌍의 개수이고, C_s 는 임의의 대응쌍을 기준으로 인식대상 물체의 일정 반경 r 내에 존재하는 모든 이웃하는 대응쌍의 개수를 의미한다. C_m 은 쌍의 출발점이 소근육 3D 모델에서 시작하고 C_s 는 대응 쌍의 출발점이 인식대상 물체의 표면에서 시작된다. 그림에서 적색 화살표는 대표 대응쌍을 의미한다. 여기서 확률값이 1에 가깝다는 것은 양방향에서 출발점의 모든 인접하는 이웃들이 대표 대응쌍의 종착점 주변에 동일하게 이웃하게 모여있다는 것을 의미한다.

제안된 방법의 알고리즘은 Table 1과 같다.

d) 정합

인식 및 포즈 추정의 마지막 단계는 정합(registration)이다. 본 논문에서는 대응점을 사용하여 포즈를 추정하기 위해 RANSAC 기법 [1]을 사용한다. 3D 공간에서 변환 행렬을 구하려면 최소 4개의 대응쌍이 필요하다. 따라서 4개의 점을 무작위로 선택하고 이를 사용하여 모델 프레임에서 객체 프레임으로의 변환행렬을 계산한다. 이러한 단계는 MSE(Mean Squared Error)가 사전 정의된 공차 임계값을 충족하는 특정 상태로 수렴될 때까지 반복된다. 원래 매개변수 공간 \mathcal{z} 에서 직접 변환 (R,T)을 계산하는 대신 [17]에 제시된 Quaternion 기반 방법을 사용하였다. 여기서 변환 행렬은 벡터로 표현된다.

마지막으로 자세 정제(pose refinement) 프로세스는 일반적인 ICP 알고리즘을 사용하는데[2], Fig. 4는 기본 원리를 설명한다. 최종 정합을 위한 변환행렬은 [17]에

제시된 단위 Quaternion 기반 등록 방법을 사용하여 구한다.

Fig. 5는 등록 개선 시퀀스의 예를 보여준다. 빨간 점선의 원 주변에서 점진적 개선을 쉽게 확인할 수 있다.

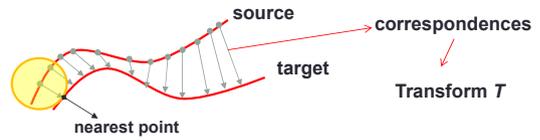


Fig. 4. Concept of ICP

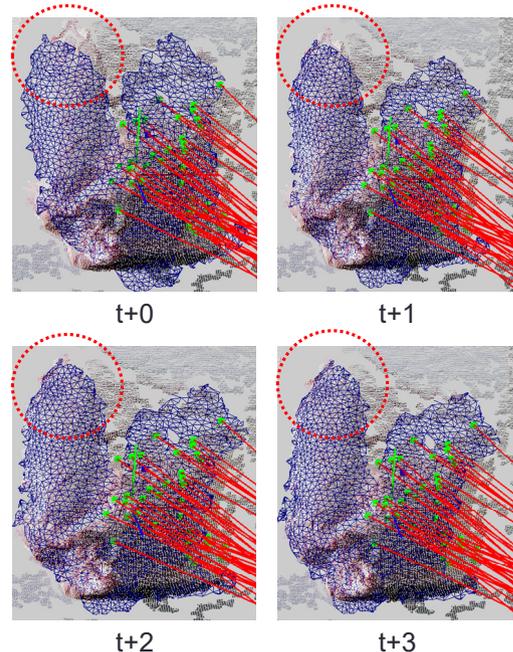


Fig. 5. Pose refinement process

3. 절단 가이드 경로 추정

Fig. 6에서 볼 수 있듯이 소근육 절단의 첫 번째 과정은 두 근육 사이의 절단면(움푹 들어간 골짜기)을 찾고 해당 위치의 가상 절단면 라인을 따라 소 부위를 옆구리 살과 옆구리 뒷살 두 부분으로 분리하는 것이다. 이 장에서는 변형 상황에서 로봇의 절단 작업에 사용될 절단 경로(cutting path)를 추적하는 시스템을 제안한다. 이를 위해 먼저 몇 가지 원칙을 미리 정의해야 한다. 원칙은 다음과 같다: 첫째, 절단 가이드 경로를 실시간으로 감지하고 추적해야 한다. 이 원칙에 따라 우리의 목표는 가이드 경로 추적을 위한 모든 프로세스에서 500ms 미만의 연산시간을 임계점으로 잡았다. 둘째, 로봇의 조작 동작을 단순하고 일관적이며 깔끔하게 만들기 위해 가이드 경로는 안정적이고 최적의 이동 경로로 수립되어야 한다. 셋째, 실제 작업에서 절단 경로를 따라 깨끗하게 절단되어야 하므로 절단 가이드 경로가 충분히 자연스럽게 부드러워야 한다.

3.1 초기 절단 가이드 경로 생성

Fig. 6은 수작업으로 고기를 절단하는 과정 및 절단 경로(cutting path)의 예시이다.



Fig. 6. Cutting process and virtual cutting path

절단면 추정에 들어가기에 앞서 우선 모델 제작시 초기 가이드 경로를 미리 정의한다. 이를 위해 Fig. 7과 같이 컬러 메쉬를 사용한다. 여기에서 가이드 경로는 일련의 Vertex를 포함하는 세 부분으로 구성되며, 시작 위치는 노란색 Vertex, 끝 위치는 보라색 Vertex, 중간은 빨간색 Vertex로 표현되는데 노란색 점과 보라색 점 사이의 연결을 의미한다. 흰색으로 표시된 나머지 지점은 안내 경로로 간주되지 않는다. 이렇게 컬러로 각 Vertex를 달리 표현한 것은 모델에서 별도의 식별자(descriptor)를 사용하지 않고 컬러로써 해당 Vertex를 쉽게 찾고 구분하기 위함이다.

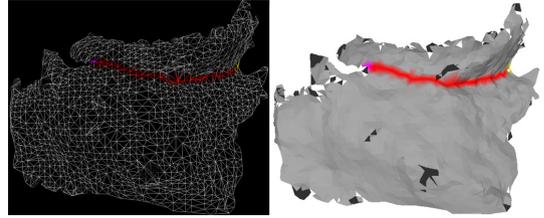


Fig. 7. Pre-defined cutting-path in a model which is represented by mesh; left: wireframe representation, right: surface representation. Yellow Vertex: start position, purple Vertex: end position, red vertices: intermediate path between start position and end position

초기 가이드 라인을 그릴 때는 MeshLab [18]과 같은 메쉬 처리 도구를 사용한다. Fig. 8에 표시된 예에서는 총 28개의 Vertex가 사용되었다.

3.2 실시간 자유변형 표면 추적

$t+1$ 이후의 객체 추적은 일반적으로 실시간으로 객체의 방향과 이동을 추적해야 한다. Free-form / Deformable Object 추적의 경우, 점 또는 메쉬의 변형을 통해 모양의 변형을 추적해야 한다. 그러나 이러한 방법은 여전히 많은 리소스와 연산 시간이 필요하다. 실제로 우리는 이런 종류의 방법을 사용하여 500ms의 처리 시간 제한을 달성할 수 없었다. 따라서 우리는 이전 인식 단계에서 소근육의 정확한 자세와 카메라가 설치된 로봇 팔의 주행 거리 측정이 충분히 정확하다고 가정하고 시간 소모를 최소화하기 위해 색상 히스토그램 일치 기반의 한 영역 추적을 사용한다. 이는 정확한 물체의 자세는 반복되는 정합 과정에서 구해지므로 대략적인 위치만 알아도 되기 때문이다.

컬러 히스토그램 일치 및 정합, 메쉬 생성, 이 세 단계는 표면 추적 단계의 핵심 파이프라인이다. 이는 실시간으로 반복된다.

Fig. 8은 추적 시퀀스의 예를 보여준다. 왼쪽 사진은 메쉬로 변환된 오브젝트 표면을 나타내고, 오른쪽 사진은 분할에 의해 변형된 고기 표면의 메쉬를 나타낸다.

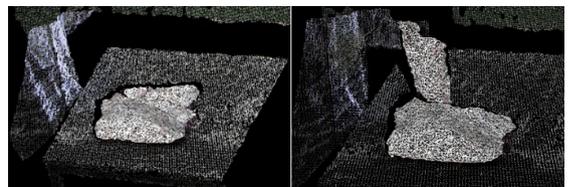


Fig. 8. Surface Tracking

3.3 Non-rigid ICP 기반 절단 경로 추적

이 단계의 목표는 가장 깊은 절단면 골짜기 지점을 찾고 절단 중에도 연속 시퀀스에서 절단 경로를 지속적으로 추적하는 것이다. 여기서 전체 과정을 세 단계로 나눌 수 있다: 첫째, 가장 깊은 계곡 지점의 검출; 둘째, 절단 경로의 정합과 추적; 셋째, 절단 경로 정제(cutting path refinement).

각 단계를 상세 설명하자면, 대상 표면 R^3 에서 N_p 개의 입력 Point cloud 포인트 집합 $P = \{\vec{p}_i\}$ 와 모델에 정의된 N_x 개의 절단 가이드 경로 포인트 집합 $X = \{\vec{x}_j\}$ 가 주어진다고 가정해 보자. 그리고 N_p 포인트를 갖는 평균 곡률값 집합 $K = \{\kappa_i\}$ 도 제공되며, 이

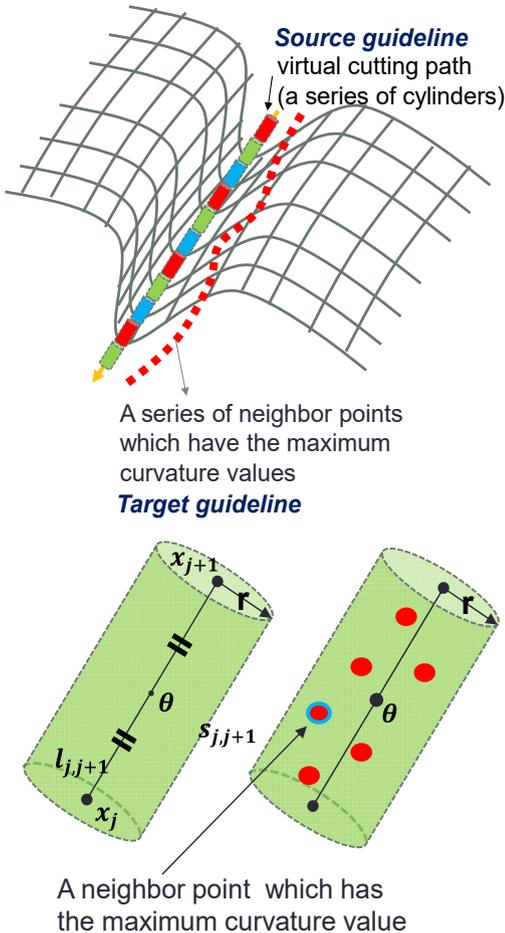


Fig. 9. Virtual cutting path, Left: virtual guide path which consists of a series of synlinders, Right: virtual cylinder

값은 가이드 경로를 감지하기 전에 계산할 수 있다[21]. 가이드 경로는 메쉬(mesh)에 있는 정점(vertex)의 연결된 리스트이다. 따라서 우선 시작 위치부터 종단 위치까지 순차적으로 정점(vertex)을 정렬해야 한다.

결과적으로 두 가이드 경로 점 x_j, x_{j+1} 사이의 링크 $l_{j,j+1}$ 에는 반경 r 과 중심점 $\Theta = \{\theta_j\}$ 를 갖는 가상 실린더 $s_{j,j+1}$ 가 있다. 여기서, $N_p = N_x$ (Fig. 9 참조).

다음 네 단계는 절단 경로 추적 과정이다. 추적 단계에서는 발산이 없다면 종료 규칙 없이 계속 반복된다.

- 1) 실린더에서 가장 가까운 점 계산: $Y_k = C(P_k, \Theta)$,
where $Y = \{(\vec{p}_0, \kappa_0), \dots, (\vec{p}_j, \kappa_j)\}$.
- 2) 가장 가까운 지점 중에서 최대 곡률 값을 갖는 지점을 찾음: $\psi_k = \underset{Y_{k,j}}{\operatorname{argmax}}(x_j)$.
- 3) Non-rigid ICP를 사용하여 affine 변환 계산:
 $T = (\Theta_0, \Psi_k)$.
- 4) Transform 적용: $\Theta_{k+1} = T(\Theta_0)$
(cost: $O(N_p)$)

이 단계에서 Non-rigid 표면에 대한 Affine Transform 행렬은 다음과 같이 계산할 수 있다[19].

$$E(X) := E_d(X) + \alpha E_s(X) + \beta E_l(X) \quad (8)$$

여기서 $E_d(X)$ 는 distance term이고, $E_s(X)$ 는 가중치가 α 인 stiffness term이며, $E_l(X)$ 는 각각 가중치가 β 인 Landmark 항이다. 필요한 경우 Landmark term을 무시할 수 있다.

비용 함수를 확장하면 완전한 2차 함수는 다음과 같다.

\tilde{E} 를 반복적으로 최소화하면 E의 국소 최솟값이 발견된다. 미분을 0으로 설정하고 결과적인 선형 연립 방정식을 풀면 최적해를 구할 수 있다.

$$\begin{aligned} \tilde{E}(X) &= \left\| \begin{bmatrix} \alpha M \otimes G \\ WD \\ \beta D_L \end{bmatrix} X - \begin{bmatrix} 0 \\ WU \\ U_L \end{bmatrix} \right\|_F^2 \\ &= \|AX - B\|_F^2 \end{aligned} \quad (9)$$

3.4 가상 절단 경로 정제

위 단계에서 생성된 가이드 경로는 지역 내에서 진동할 수 있으므로 우리의 목표인 절단 작업에는 정확도 측면에서 여전히 충분하지 않다. 따라서 마지막 단계로 절

단 경로 정제(cutting path refinement) 프로세스가 필요하다. 이를 위해 이 분야에서 일반적으로 사용되는 다항식 곡선 피팅(curve fitting) 방법을 사용한다.

처음에는 직선 (즉, 1차 다항식)에서 k차 다항식으로 일반화시킨다[20].

$$y = a_0 + a_1x + \dots + a_kx^k, \quad (10)$$

잔차는 다음과 같이 주어진다.

$$R^2 \equiv \sum [y_i - (a_0 + a_1x_i + \dots + a_kx_i^k)]^2 \quad (11)$$

다음으로, n개의 점 (x_i, y_i) 와 다항식 계수 a_0, \dots, a_k 가 주어지면, 곡선 피팅을 위한 연립방정식을 세울 수 있다 (Eq. (12)).

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^k \\ 1 & x_2 & x_2^2 & \dots & x_2^k \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^k \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{bmatrix} \quad (12)$$

축약하면 Eq. (13)과 같이 표현할 수 있고, 역행렬을 구하면 벡터 해 a 를 구할 수 있다 Eq. (14).

$$y = Xa \quad (13)$$

$$a = (X^T X)^{-1} X^T y \quad (14)$$

Fig. 14는 정제 프로세스의 결과를 나타낸다. 무지개색 선이 안내 경로이다. 부드러운 곡선으로 잘 피팅된 것을 확인할 수 있다. 이 예에서 우리는 5차 다항식을 사용하고 월드(world) 좌표 (x_i, y_i) 및 (x_i, z_i) 를 입력으로 사용했다. 절단 경로는 3차원 곡선이어야 하므로 (x_i, y_i) 를 먼저 곡선 피팅(curve fitting)하고, (x_i, z_i) 를 순차적으로 피팅하여 3차원 곡선 피팅을 완료한다.

3.5 실험결과

개발된 알고리즘은 구조광 3D 비전 센서로 Kinect를 사용하여 테스트하였다. Scene에 대한 RGB-D 정보를 제공한다. 옆구리살과 옆구리 뒷살 두 개의 소 근육에 의해 형성된 고기 부분의 3D 모델은 Fig.10과 같다. 제시된 알고리즘의 각 단계는 크기, 소 종 및 연령이 다양한 여러 육류에서 테스트되었다.

Fig. 11~14는 각 과정별 결과를 단계별로 보여준다.

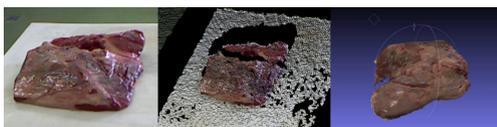


Fig. 10. Example of bovine muscle, 2D input image, 3D Point cloud, reconstructed 3D mesh model

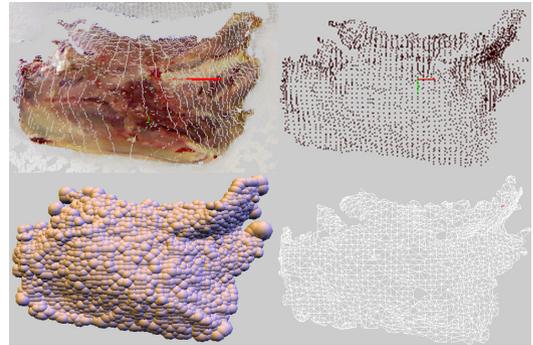


Fig. 11. From the left-top, input scene of bovine muscle, right: segmented meat surface, bottom-left: ball representation for searching neighborhoods and outlier removal, bottom-right: generated mesh

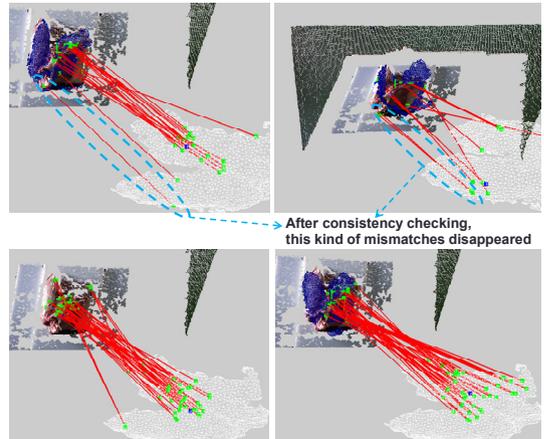


Fig. 12. Examples of mismatches and best matches

4. 결론

본 논문은 소 근육 분리의 자동화에 관한 연구로 실시간으로 소 근육을 RGBD 센서를 이용하여 3차원으로 복원하고 Spin-image를 생성하여 사전 제작된 3차원 모델과 정합, 실시간으로 소 근육의 자세를 추정하는 방식을 최초로 제안하였고, 추정된 자세 정보와 가상의 실린더와 곡률(curvature) 정보를 이용하여 해당 물체의 절단 경로(cutting path)를 실시간으로 추정 및 추적하는 방법을 제시하였다. Spin-image 상관관계를 사용하여 관찰된 모델을 정합하기 위한 3차원 대응점을 찾는데 활용하였고, 절단면의 변형 가능한 메쉬를 실시간으로 업데이트하는 방법도 제시하였다. 또 RGBD와 실제 소 근육을 사용하여 실험을 진행하였고 가능성을 확인하였다.

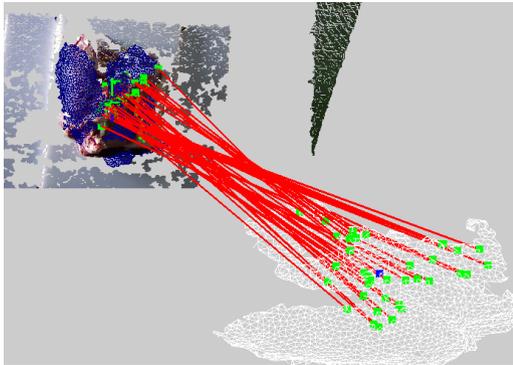


Fig. 13. Best matches extracted by consistency checking algorithm

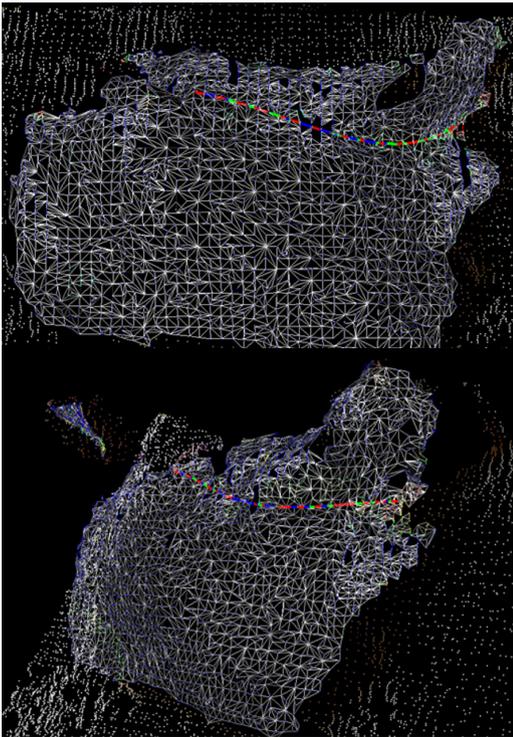


Fig. 14. Refined cutting guide-path: rainbow colored lines represent the cutting guide path

우리의 다음 과제는 소 근육 분리 작업의 자동화를 달성하기 위해 알고리즘을 Visual Servoing 시스템과 통합하고 딥러닝 기법을 활용하여 Free-form 물체의 인식 및 자세 추정의 강인성 및 정확도를 향상시키는 데 있다.

References

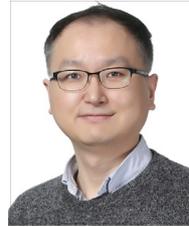
- [1] A. Johnson, *Spin-images: a representation for 3-D surface matching*, Ph.D. dissertation, The Robotics Institute, Carnegie Mellon University, August 1997.
- [2] Körtgen M, Park G-J, Novotni M, Klein R, "3D shape matching with 3D shape contexts", *Proceedings of 7th central European seminar on computer graphics*, 2003.
- [3] R. B. Rusu, N. Blodow and M. Beetz, "Fast Point Feature Histograms (FPFH) for 3D registration". *IEEE International Conference on Robotics and Automation*, pp. 3212-3217, 2009. DOI: <http://dx.doi.org/10.1109/ROBOT.2009.5152473>
- [4] Hui Chen and Bir Bhanu, "3D free-form object recognition in range images using local surface patches", *Proceedings of the 17th International Conference on Pattern Recognition*, Vol. 3, pp. 136-139, 2004. DOI: <http://dx.doi.org/10.1016/j.patrec.2007.02.009>
- [5] B. Drost, M. Ulrich, N. Navab and S. Ilic, "Model globally, match locally: Efficient and robust 3D object recognition", *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp.998-1005, 2010. DOI: <http://dx.doi.org/10.1109/CVPR.2010.5540108>
- [6] Vidal, Joel et al. "A Method for 6D Pose Estimation of Free-Form Rigid Objects Using Point Pair Features on Range Data", *Sensors*, Basel, Switzerland, Vol. 18, No. 8, p.2678, August 2018. DOI: <http://dx.doi.org/10.3390/s18082678>
- [7] Charles R. Qi, Hao Su, Kaichun Mo, Leonidas J. Guibas, "PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.652-660, 2017. DOI: <http://dx.doi.org/10.1109/CVPR.2017.16>
- [8] Charles R. Qi Li Yi Hao Su Leonidas J. Guibas, "PointNet++: deep hierarchical feature learning on point sets in a metric space", *Proceedings of the 31st International Conference on Neural Information Processing Systems*, pp.5105-5114, 2017.
- [9] C. Choy, J. Gwak and S. Savarese, "4D Spatio-Temporal ConvNets: Minkowski Convolutional Neural Networks", *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3070-3079, 2019. DOI: <http://dx.doi.org/10.1109/CVPR.2019.00319>
- [10] Nico Engel, Vasileios Belagiannis, Klaus Dietmayer, "Point Transformer," *Access*, IEEE, Vol. 9, pp.134826-134840, Sept. 2021. DOI: <http://dx.doi.org/10.1109/ACCESS.2021.3116304>
- [11] Martin A. Fischler and Robert C. Bolles, "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis & Automated

Cartography". *Comm. of ACM*, Vol. 24, No. 6, pp.381-395, June 1981.
DOI: <http://dx.doi.org/10.1145/358669.358692>

- [12] Besl, Paul J.; N.D. McKay, "A Method for Registration of 3-D Shapes". *IEEE Transaction on Pattern Analysis and Machine Intelligence*, IEEE Computer Society, Vol. 14, No. 2, pp.239-256, 1992.
DOI: <http://dx.doi.org/10.1109/34.121791>
- [13] A. Johnson and M. Hebert, "Surface Matching for Object Recognition in Complex Three-Dimensional Scenes", *Image and Vision Computing*, Vol. 16, pp.635-651, 1998.
DOI: [http://dx.doi.org/10.1016/S0262-8856\(98\)00074-2](http://dx.doi.org/10.1016/S0262-8856(98)00074-2)
- [14] Djouadi, A.; Snorrason, O.; Garber, F., "The quality of Training-Sample estimates of the Bhattacharyya coefficient". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 12, No. 1, pp.92-97, 1990.
DOI: <http://dx.doi.org/10.1109/34.41388>
- [15] A. Johnson and M. Hebert, "Using Spin Images for Efficient Object Recognition in Cluttered 3D Scenes", *IEEE transactions on Pattern Analysis and Machine Intelligence*, Vol. 21, No. 5, May 1999.
DOI: <http://dx.doi.org/10.1109/34.765655>
- [16] Yutaka Ohtake, Alexander Belyaev, and Hans-Peter Seidel, "An integrating approach to meshing scattered point data", *Proceedings of the 2005 ACM symposium on Solid and physical modeling (SPM '05)*, ACM, New York, NY, USA, pp.61-69, 2005.
DOI: <http://dx.doi.org/10.1145/1060244.1060252>
- [17] Raouf Benjema and Francis Schmitt, "A Solution for the Registration of Multiple 3D Point Sets Using Unit Quaternions", *Proceedings of the 5th European Conference on Computer Vision*, ECCV '98, Hans Burkhardt and Bernd Neumann (Eds.), Vol. 2, Springer-Verlag, London, UK, UK, pp.34-50, 1998.
DOI: <http://dx.doi.org/10.1007/BFb0054732>
- [18] "MeshLab," Available From: <https://github.com/cnr-isti-vclab/meshlab> (accessed Dec. 20, 2021)
- [19] Brian Amberg, Sami Romdhani, Thomas Vetter. "Optimal Step Nonrigid ICP Algorithms for Surface Registration", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, CVPR 2007, Minneapolis, Minnesota, USA. IEEE Computer Society, pp.18-23, June 2007.
DOI: <http://dx.doi.org/10.1109/CVPR.2007.383165>
- [20] "Least Squares Fitting-Polynomial", Available From: <http://mathworld.wolfram.com/LeastSquaresFittingPolynomial.html> (accessed Dec. 20, 2021)
- [21] "Principal curvatures", Available From: http://www.math.union.edu/~jauregui/principal_curvatures.pdf (accessed Dec. 20, 2021)

박 연 출(Yeon-Chool Park)

[정회원]



- 2004년 2월 : 송실대학교 컴퓨터학과 박사
- 2005년 4월 ~ 2009년 3월 : 성균관대학교 ISRI 연구교수
- 2009년 5월 ~ 2010년 6월 : 스웨덴 Umeå Univ. Post-Doc.
- 2011년 4월 ~ 2013년 2월 : 프랑스 Pascal Institute Post-Doc.
- 2013년 9월 ~ 현재 : LG전자 CTO 책임연구원

〈관심분야〉

로봇비전, 인공지능, 딥러닝

박 세 준(Se-Joon Park)

[정회원]



- 1996년 2월 : 송실대학교 수학과 (이학사)
- 1998년 2월 : 송실대학교 컴퓨터공학과 (공학석사)
- 2004년 8월 : 송실대학교 컴퓨터공학과 (공학박사)
- 2006년 1월 ~ 현재 : SK 정보보호담당 수석연구원

〈관심분야〉

인공지능보안, 개인정보보호, 정보보호

이 용 준(Yong-Joon Lee)

[중신회원]



- 2005년 2월 : 송실대학교 컴퓨터학과 박사 (공학박사)
- 2010년 2월 ~ 2016년 3월 : 한국인터넷진흥원 수석연구원
- 2016년 4월 ~ 2020년 3월 : 국방보안연구소 연구관
- 2021년 4월 ~ 현재 : 극동대학교 해킹보안학과 교수

〈관심분야〉

인공지능보안, 국방보안, 해킹보안