

SIFT 및 HSV 특징 추출 기반 폐기물 객체 유사도 측정 모델

고준혁¹, 최혁순², 김진아³, 문남미⁴

호서대학교 컴퓨터공학과 석사과정¹

호서대학교 컴퓨터공학과 학사과정²

서울미디어대학원대학교 인공지능응용소프트웨어학과 교수²

호서대학교 컴퓨터공학과 교수³

junhyeok970306@gmail.com, hyuksoon2001@gmail.com,
kkim.jinah00@gmail.com, nammee.moon@gmail.com

The SIFT and HSV feature extraction-based waste Object similarity measurement model

JunHyeok Go¹, Hyuk soon Choi², Jinah Kim³ Nammee Moon⁴

Dept. of Computer Science, Hoseo University^{1,2,4}

Dept. of AI Software Engineering, Seoul Media Institute of Technology³

요약

폐기물을 처리하는데 있어 배출과 수거에 대한 프로세스 자동화를 위해 폐기물 객체 유사도 판별이 요구된다. 이를 위해 본 연구에서는 폐기물 데이터셋에서 SIFT(Scale-Invariant Feature Transform)와 HSV(Hue, Saturation, Value)기반으로 두 이미지의 공통된 특징을 추출해 융합하고, 기계학습을 통해 이미지 객체 간의 유사도를 측정하는 모델을 제안한다. 실험을 위해 수집된 폐기물 데이터셋 81,072장을 활용하여 이미지를 학습시키고, 전통적인 임계치 기반 유사도 측정과 본 논문에서 제시하는 유사도 측정을 비교하여 성능을 확인하였다. 임계치 기반 측정에서 SIFT와 HSV는 각각 0.82, 0.89(Acc)가 측정되었고, 본 논문에서 제시한 특징 추출 방법을 사용한 기계학습의 성능은 DT(Decision Tree)와 SVM(Support Vector Machine) 모두 0.93 (Acc)로 4%의 정확도가 향상되었다.

1. 서론

특징점 추출 알고리즘은 최근에도 이미지의 유사도를 측정하는 방법으로 사용하는 연구들이 진행되고 있다[1,2,3]. 특징점 추출 알고리즘인 SURF, SIFT, ORB을 이용해 이미지의 회전, 왜곡, 스케일 변화에서의 매칭 성능을 비교한 연구가 존재한다[3]. 폐기물 객체 또한 카메라 각도에 따라 왜곡, 스케일 변화가 존재하기 때문에 SIFT를 통해 이미지의 특징을 추출하였다. 하지만 SIFT는 회색조에서 특징을 추출하므로 색상 정보가 무시될 수 있고, 이미지 전체에서 특징을 추출하기 때문에 배경에 대한 특징도 같이 추출될 수 있는 문제가 존재한다.

본 논문에서는 문제 해결을 위해 HSV 색상 정보를 추출하여 SIFT 특징점과 융합하고, 배경을 제거한 이미지 객체의 특징점을 더하는 두가지 방법을 제안한다. 제안된 방법을 통해 추출된 특징점을 DT, SVM 모델을 이용한 실험을 진행하여 그 성능을 검증하였다.

2. 관련 연구

2.1 SIFT 특징점 추출

디스크립터(Descriptor) 특징 추출 방법은 객체의 관심 지점인 키포인트(keypoint)와 그 주변 특징 벡터인 디스크립터를 이용한 특징 추출 방법이다[4]. SIFT 특징점 추출 방법은 디스크립터를 활용한 특징 추출 방법 중 하나로 회전, 크기 변화에 불변한 특성을 가지고 있다[4]. 하지만, SIFT 특징점 추출 방법은 회색조에서 추출되므로, 색상정보를 포함하지 않기 때문에 추가적인 특징 추출 방법을 사용해야 한다.

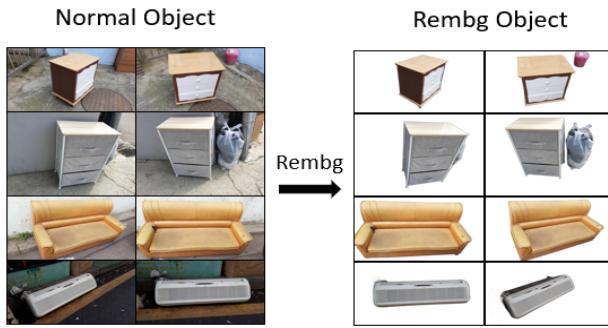
2.2 HSV 히스토그램 특징 추출

HSV 히스토그램 특징 추출은 색조, 채도, 밝기 정보를 가지는 히스토그램을 특징점으로 사용하여 색 분포를 추출하는 방법이다. 본 논문에서는 SIFT 디스크립터 특징 벡터와 HSV 히스토그램 특징 벡터를 융합하여 이미지상의 공통된 특징을 추출한다.

3. 폐기물 데이터셋 구성

데이터셋은 폐기물 데이터로 객체 유사도를 학습하기 위해 여러 각도로 촬영된 이미지를 사용하였다. 폐기물 이미지 81,072장을 이용하여 같은 객체로 구성된 이미지 40,536쌍과 서로 다른 객체로 구성된 이미지 40,536쌍을 제작하여 총 81,072쌍의 학습 데이터셋을 제작하였다. 또한, 학습 데이터, 검증 데이터, 테스트 데이터를 각각 7:2:1로 나누어 학습하였다.

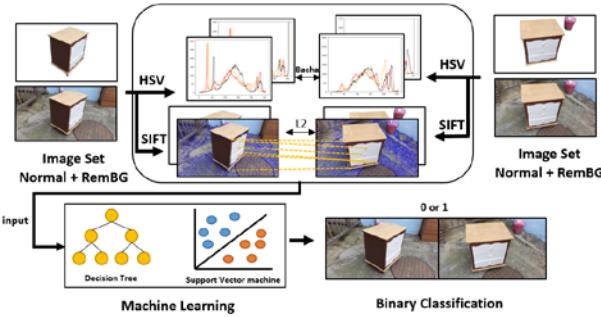
3.1 데이터 전처리



(그림 1) 배경을 제거한 폐기물 데이터 예시

SIFT, HSV 특징 추출은 이미지 전체에서 특징을 추출하기 때문에, 배경의 특징이 같이 추출될 수 있다. Rembg는 U2net 기반의 배경 제거 라이브러리로 다양한 연구에서 사용되어지고 있다[5,6]. 본 연구 또한 배경을 제거하는 목적으로 (그림 1)과 같이 Rembg를 사용하였으며, 배경을 포함한 이미지의 특징 벡터와 배경을 제거한 Rembg 이미지의 특징 벡터를 융합하여 유사도 측정 모델의 입력으로 사용하였다.

4. 유사도 측정 모델



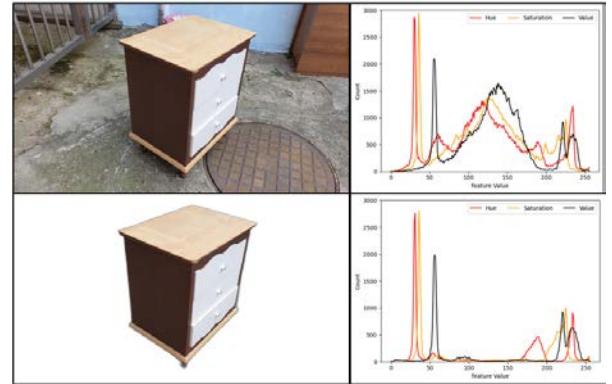
(그림 2) SIFT, HSV 특징 추출 기반 유사도 측정 프로세스

본 논문에서 제시하는 SIFT, HSV 특징 추출 기반 유사도 측정 모델의 프로세스는 (그림 2)와 같다.

이미지 셋을 원본 이미지와 Rembg 이미지로 만들고 SIFT 특징점 추출, HSV 히스토그램 분석을 통해 각각 특징을 추출하고 융합한다. 추출된 특징들은 유사도를 측정하기 위해 사용되며, DT, SVM 모델의 입력으로 사용되어 유사도를 예측한다.

4.1 HSV 기반 특징 추출

이미지의 색상, 채도, 명도 특징을 추출하기 위해 HSV 색상 공간에서 히스토그램을 만들어 색상 분포를 확인하였다. Rembg 이미지의 경우 배경이 현실에 거의 존재하지 않은 최소값으로 고정되어 이를 삭제하는 후처리를 진행하였다. (그림 4)는 HSV 공간에서 색상, 채도, 명도 특징을 추출한 히스토그램이다. Rembg 이미지에서 배경의 색상 특징이 제거되어 노이즈가 감소한 것을 확인 할 수 있다.



(그림 4) 폐기물 데이터 HSV 히스토그램 시각화

4.2 SIFT 기반 특징점 추출

SIFT 디스크립터 특징 추출 알고리즘은 스케일 공간 구조, DOG(Difference of Gaussian)연산, 키포인트 검출, 디스크립터 벡터 추출 순서로 진행된다[4].

폐기물 데이터에서 크기 변화에 강한 키포인트를 검출하기 위해 이미지 피라미드를 활용하여 폐기물 데이터의 스케일 공간을 구성하고 DOG 연산을 통해 외각선을 검출한다.

$$\mathbf{D}(\mathbf{x}) = \mathbf{D} + \frac{\partial \mathbf{D}^T}{\partial \mathbf{x}} \mathbf{x} + \frac{1}{2} \mathbf{x}^T \frac{\partial^2 \mathbf{D}}{\partial \mathbf{x}^2} \mathbf{x} \quad (1)$$

$$\hat{\mathbf{X}} = -\frac{\partial^2 \mathbf{D}^{-1}}{\partial \mathbf{x}^2} \frac{\partial \mathbf{D}}{\partial \mathbf{x}} \quad \mathbf{D}(\hat{\mathbf{X}}) = \mathbf{D} + \frac{1}{2} \frac{\partial \mathbf{D}^T}{\partial \mathbf{x}} \hat{\mathbf{X}} \quad (2)$$

DOG 연산으로 검출된 외각선에서 키포인트의 후보군을 찾고 식(1)을 이용하여 정확한 키포인트의 위치를 탐색한다. 식(1)은 테일러 급수 2차 전개를 이용한 식으로 D는 DOG 이미지, X는 키포인트 후보인 (x, y, octave)를 의미한다. 식(2)를 통해 위치조정 값 $\hat{\mathbf{X}}$ 도출하고, $|\mathbf{D}(\hat{\mathbf{X}})|$ 의 값이 0.03이하인 약한 키포인트는 삭제하는 후처리를 진행한다.

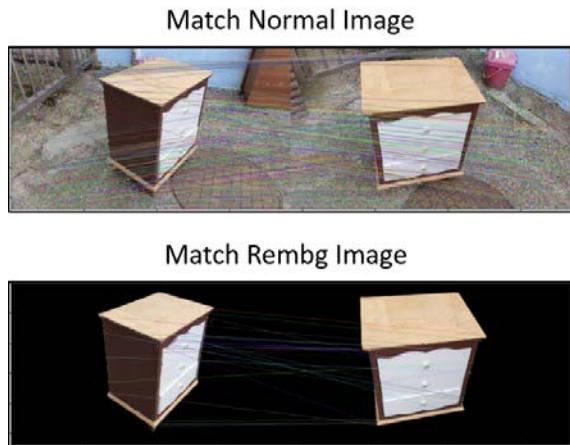
이미지 회전에 강한 특징점을 검출하기 위해 키포인트 주변 픽셀의 상대적인 그래디언트 방향 정보를 가지는 디스크립터를 생성한다. 스케일 공간에서의 키포인트와 디스크립터 생성을 통해 폐기물 이미지에서 크기와 회전 변화에 강한 특징 벡터를 추출한다.

4.3 이미지 특징점 매칭

유사도를 측정하고자 하는 이미지 쌍을 SIFT, HSV 분석을 각각 진행하고, 검출된 특징 사이의 거리를 각각 측정한다.

$$\sum_i (\mathbf{A}_i - \mathbf{B}_i)^2 \quad (3)$$

SIFT 특징점 사이의 거리는 유clidean 거리식(3)을 이용하여 측정하고, 최근접 이웃 방식을 이용해 이웃된 특징점의 거리가 각각 40%, 50%, 60%, 70% 이하인 매칭점들의 개수를 각각 추출한다. (그림 5)는 이미지 쌍에서 60% 임계치에서 매칭된 특징점을 시각화한 이미지이다. SIFT 매칭 역시 Rembg Image 매칭에서 배경에 대한 노이즈가 제거된 것을 확인할 수 있다.



(그림 5) SIFT 특징점 매칭 시각화

HSV 히스토그램 비교 알고리즘은 바타차야 거리 측정 방법을 사용했으며, 두 이미지의 히스토그램을 비교하여 식(4)와 같이 구하게 된다.

$$d(H_1, H_2) = \sqrt{1 - \frac{1}{\sqrt{H_1 H_2 N^2}} \sum_i \sqrt{H_1(I) \cdot H_2(I)}} \quad (4)$$

SIFT의 매칭된 특징점 크기(4,)와 HSV 히스토그램 거리 측정 값 벡터크기(1,)를 더하여 (5,)크기의 특징점 벡터를 생성하게 되며, 원본 이미지와 배경 제거한 이미지의 특징점을 융합하여 총 (10,)크기의 특징 벡터를 유사도 측정을 위한 기계학습의 입력으로 사용한다.

5. 실험

5.1 실험 환경

모든 실험은 Intel i9-11900K CPU, NVIDIA GeForce RTX 3090 GPU로 진행되었으며 Linux 20.04.2 LTS의 운영체제 버전과 CUDA version 11.10의 워크스테이션에서 진행되었다.

5.2 임계치 기반 유사도 분석

<표 1>은 학습데이터와 테스트 데이터를 이용해서 원본 이미지와 Rembg 이미지에 대해 가장 높은 성능을 기록하는 임계치의 acc를 측정한 결과를 나타내는 표이다. 임계치 기반으로 데이터의 acc를 측정했을 때, SIFT 특징점에서는 0.5, 0.6 임계치에서 0.82의 정확도를 보였고, HSV은 0.89의 acc로 SIFT보다 HSV에서 높은 성능을 보였다. 또한, 배경을 제거하지 않은 이미지가 전체적으로 높은 성능을 보였다.

<표 1> 임계치를 이용한 유사도 분석 결과

Feature	Threshold	Train_acc	Test_acc	Rembg
SIFT	0.4	0.79	0.78	X
SIFT	0.5	0.82	0.82	X
SIFT	0.6	0.82	0.82	X
SIFT	0.7	0.79	0.79	X
SIFT	0.4	0.67	0.69	O
SIFT	0.5	0.69	0.69	O
SIFT	0.6	0.69	0.69	O
SIFT	0.7	0.66	0.65	O
HSV		0.89	0.87	X
HSV		0.77	0.77	O

5.2 기계 학습을 이용한 유사도 분석

이미지쌍에서 추출한 특징 벡터를 기계 학습의 입력값으로 사용한다. 기계학습 알고리즘은 scikit learn에서 구현된 DT와 SVM을 사용하였다. HSV와 SIFT를 융합한 특징 벡터의 효용성을 실험하기 위해, SIFT 특징 벡터를 사용한 성능과 HSV 특징 벡터를 융합하여 예측한 성능을 비교하였다.

DT의 하이퍼파라미터는 실험을 통해 <표 2>로 설정하였으며, 실험 결과는 <표 3>와 같다.

<표 2> Decision Tree의 하이퍼파라미터 설정

Class weight	Balanced	Min samples leaf	11
Criterion	Gini	Min samples split	5
Max depth	9	Splitter	Best
Max feature	sqrt		

<표 3> Decision Tree의 학습 결과

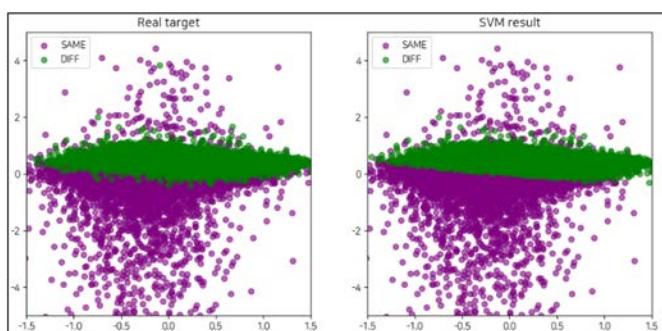
Feature	Background	Acc	Recall	Precision
SIFT	Normal	0.83	0.90	0.78
SIFT+HSV	Normal	0.91	0.94	0.89
SIFT	Rembg	0.76	0.68	0.81
SIFT+HSV	Rembg	0.81	0.88	0.77
SIFT	Normal + Rembg	0.84	0.90	0.78
SIFT+HSV	Normal + Rembg	0.93	0.95	0.92

SVM의 하이퍼파라미터는 실험을 통해 <표 4>로 설정 후 학습을 진행하였다.

<표 4> Support Vector machine 하이퍼파라미터 설정

C	1000	kernel	rbf
gamma	1.0		

PCA를 이용한 차원 축소로 SVM을 시각화 하여 결정 경계를 확인하였다. SIFT, HSV 특징 추출이 다른 이미지쌍에서는 0에 가까운 숫자로 추출되기 때문에, SVM 시각화 그래프 (그림 6) 역시 다른 이미지쌍은 0에 가깝게 나타나는 것을 확인할 수 있다.



(그림 6) PCA 차원 축소를 이용한 SVM 시각화
(좌) 실제 데이터의 분포 (우) SVM 예측한 분포

<표 5> Support Vector machine 학습 결과

Feature	Background	Acc	Recall	Precision
SIFT	Normal	0.82	0.94	0.75
SIFT+HSV	Normal	0.91	0.94	0.89
SIFT	Rembg	0.76	0.68	0.82
SIFT+HSV	Rembg	0.81	0.90	0.76
SIFT	Normal + Rembg	0.83	0.90	0.78
SIFT+HSV	Normal + Rembg	0.93	0.96	0.901

DT의 학습 결과 <표 3> 와 SVM 학습 결과 <표 5> 모두 SIFT 특징점 추출만 사용한 모델보다 HSV 히스토그램 분석을 같이 사용했을 때, 최소 5%, 최대 10%의 성능 향상을 보여줬으며, 일반적인 임계치 기반 유사도 분석보다 높은 성능을 보였다. 또한, 배경을 제거한 이미지의 특징점을 단독으로 사용했을 경우에는 원본 이미지 보다 낮은 성능을 보였으나, 원본 이미지와 배경을 제거한 이미지를 같이 사용하였을 때 높은 성능을 보였다.

최종적으로 본 연구에서 제안하는 SIFT 특징점과 HSV 히스토그램 분석 특징 벡터를 같이 사용하고, 배경을 제거한 이미지와 원본 이미지의 특징을 융합한 모델이 DT, SVM 모델에서 둘다 0.93 Acc 으로 가장 높은 성능을 보였다.

6. 결론

본 연구에서는 폐기물 이미지 데이터의 SIFT 특징 점 추출과 HSV 히스토그램 분석을 이용하여 이미지 쌍의 공통된 특징을 추출하고, 기계학습의 입력으로 사용하는 방법을 제시하였다. 임계치 기반 유사도 분석과 기계학습 DT, SVM를 비교하였으며, 배경을 제거한 이미지의 특징 벡터를 융합하는 방법을 추가적으로 사용하였다. 임계치 기반 유사도 분석과 비교하여 acc 기준 4% 이상의 성능이 향상되었음을 확인하였다.

향후에는, 추가적인 연구로 인공 신경망을 이용한 유사도 모델을 비교 및 설계하여 보다 높은 성능의 모델을 제작할 수 있을 것이다.

ACKNOWLEDGEMENT

이 논문은 2023년도 정부(산업통상자원부)의 재원으로 공공혁신수요기반신기술사업단의 지원을 받아 수행된 연구임 (No. G02P18960001202)

참고문헌

- [1] Liu, Y., Tian, J., Hu, R., Yang, B., Liu, S., Yin, L., & Zheng "Improved feature point pair purification algorithm based on SIFT during endoscope image stitching." Frontiers in Neurorobotics 16, 2022
- [2] Gupta, Surbhi, Munish Kumar, & Anupam Garg "Improved object recognition results using SIFT and ORB feature detector." Multimedia Tools and Applications 7, 34157-34171 2019
- [3] Karami, Ebrahim, Siva Prasad, and Mohamed Shehata "Image matching using SIFT, SURF, BRIEF and ORB: performance comparison for distorted images." arXiv preprint arXiv:1710.02726 2017
- [4] David G. Lowe "Distinctive Image Features from Scale-Invariant Keypoints" International Journal of Computer Vision (IJCV) 60 91-110 2004
- [5] Fitria, M., Candra, Y., Al-Assad, M. H., Roza, S., & Dawood, R "A Deep Learning-Based Model for Classifying Sweetness Level of Sky Rocket Melon: A Preliminary Result." 2023 2nd International Conference on Computer System, Information Technology, and Electrical Engineering (COSITE). IEEE 2023.
- [6] Fan, M., Lu, W., Niu, W., Peng, X., & Yang, Z. "A Large-Scale Invariant Matching Method Based on DeepSpace-ScaleNet for Small Celestial Body Exploration." Remote Sensing 14.24 2022.