

HSV 임베딩 기반 3D 객체 색상 군집화와 Visual Grounding 향상

박민호¹, 임종민², 차수빈³, 오인호⁴, 김광수**

¹성균관대학교 소프트웨어학과 학부생

²성균관대학교 소프트웨어학과 박사과정

³성균관대학교 인공지능학과 박사과정

⁴성균관대학교 소프트웨어학과 석사과정

**성균관대학교 소프트웨어학과 교수

e-mail : ^{1,3}(parkminho00000, chasobin99)@gmail.com,

^{2,4,**}(jm.lim, inho2007, kim.kwangsu)@skku.edu

Color Clustering of 3D Objects via HSV Embedding for Improved Visual Grounding

Minho Park¹, Jongmin Lim², Subin Cha³, Inho Oh⁴, Kwangsoo Kim*

^{1,2,4,*}Dept of Computer Science and Engineering, Sungkyunkwan University

³Dept of Artificial Intelligence, Sungkyunkwan University

요약

본 연구는 HSV-DBSCAN에서 제안된 HSV 색상 임베딩이 실제 3D Visual Grounding 성능 향상에 기여하는 이유를 실증적으로 분석한다. HSV 색공간을 활용할 경우, 객체의 색상 특성에 따라 적절한 구성 요소(hue, saturation, value)를 기준으로 자연스러운 색상 군집이 형성된다. 유채색은 주로 hue 값 기준으로, 무채색은 saturation과 value 기준으로 밀집되며, 이는 객체 간 색상 구분에 있어 더 뚜렷한 경계를 제공한다. 본 논문에서는 다양한 색상 분포를 갖는 3D 장면을 기반으로 이러한 군집화 특성을 시각화하고, HSV 기반 임베딩이 3D 객체의 색상 구분에서 RGB 대비 우위를 가짐을 정량적으로 분석한다.

1. Introduction

3D Visual Grounding(3DVG)은 자연어 쿼리를 기반으로 3D 장면 내 특정 객체를 식별하는 기술로, 로봇 공학, 증강현실, 자율주행 등 다양한 분야에 활용된다. 일반적으로 3D 객체를 표현하는 point cloud는 RGB 색상 정보를 포함하지만, RGB 값은 색상 간 유사성을 충분히 반영하지 못해 객체 구분이 불명확하다.

본 연구는 3D 객체의 HSV 색 공간 임베딩이 색상 특성에 따라 구조적 군집을 형성하는 현상에 주목한다. 유채색은 hue를 중심으로, 무채색은 saturation과 value를 기준으로 각각 밀집되는 분포를 보이며, 이는 색상 기반 객체 구분에 명확한 경계를 제공한다. 이러한 군집화 특성을 시각적 및 정량적으로 분석하고, HSV 임베딩이 3D 객체 표현에 구조적 이점을 제공함을 실증적으로 제시한다.

2. Related Work

HSV-DBSCAN[1]은 HSV 색 공간에서 DBSCAN을 활용하여 색상 기반 객체 탐색을 수행하는 Zero-shot 3DVG 기법이다. RGB 기반 표현의 조명 민감도와 왜곡 문제를 보완하며, 사전 학습 없이 색상 정보를 활용할 수 있다는 장점이 있다. 그러나 HSV 임베딩이 성능 향상에 기여하는 구조적 근거와 군집화 특성에 대한 분석이 부족하다.

3. Methods

본 연구에서는 각 3D 객체의 point cloud로부터 평균 RGB 색상을 계산한 뒤 HSV 색 공간으로 변환하여 색상 임베딩 벡터를 생성한다. 시각적으로는 RGB와 HSV 공간에서 임베딩 분포를 시각화하여 동일 색상 범주에 속한 객체들이 어떻게 분포하고 군집을 형성하는지 비교한다.

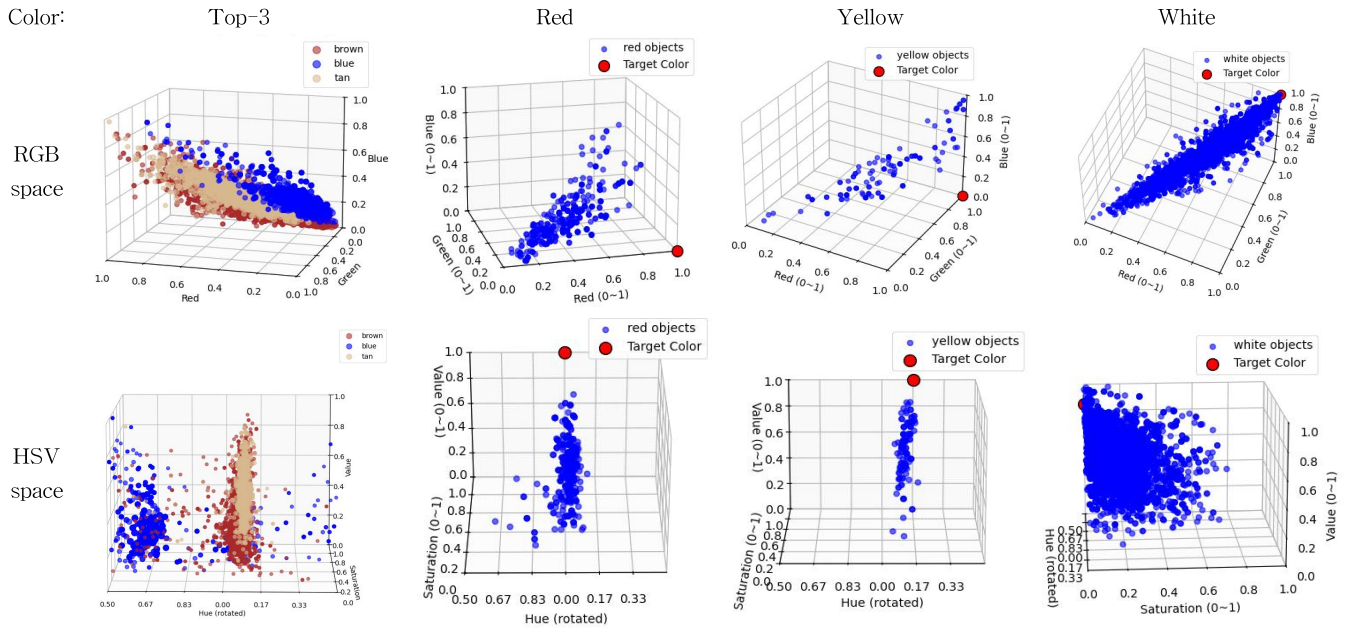
정량적 평가는 각 객체의 임베딩 벡터와 대응하는 target 색상 벡터 간의 유클리드 거리의 평균 및 표준편차를 계산하여 수행한다.

4. Experiments

HSV 임베딩의 색상별 군집화 효과와 3DVG 성능 기여를 평가하기 위해, 색상 범주 내 임베딩 군집화 분석과 색상 기반 색상 유사도 정량 분석 실험을 수행하였다. 실험에는 ScanRefer[2, 3] train 데이터셋에서 색상 표현이 포함된 쿼리 중심의 장면을 선별하여 사용하였다.

4.1. 색상 범주 내 임베딩 군집화 분석

그림 1은 Top-3 색상 및 일부 색상에 대한 객체 평균 색상 벡터를 RGB 및 HSV 공간에 시각화한 것이다. 각 점은 객체를 나타내며, KDE[4]를 통해 색상 분포 밀도를 강조하였다. RGB 공간에서는 색상이 불규칙하게 분포하



(그림 1) RGB / HSV 공간 색상 벡터 분포 시각화 비교

며 유사한 색상 간의 경계가 불분명하게 나타난다. 반면, HSV 공간에서는 유채색 객체들이 Hue 값을 기준으로 뚜렷한 클러스터를 형성하고, 무채색 객체들은 Saturation과 Value 축의 조합에 따라 분리되어 실제 색상 인식과 일치하는 구조적 분포를 나타낸다.

4.2 3DVG 색상 유사도 정량 분석

HSV 임베딩이 색상 유사도를 효과적으로 반영하는지를 평가하기 위해, 객체 색상과 실제 쿼리 색상 간의 거리 분포를 분석하였다. 표 1에 나타나듯, Hue를 기준으로 한 HSV 공간에서는 대상 색상 주변에 데이터가 더욱 밀집되며, RGB 공간에 비해 거리 분포의 평균과 표준편차가 모두 더 작게 나타났다. 이는 HSV 기반 표현이 색상 간 유사성을 정밀하게 포착할 수 있음을 시사한다.

색상 예시	RGB dist		HSV dist	
	평균	표준편차	평균	표준편차
red	0.6931	0.1090	0.0680	0.0746
yellow	0.5765	0.1175	0.1006	0.0417
green	0.3894	0.1155	0.2982	0.1228
pink	0.5774	0.2031	0.1263	0.0724
전체	0.4515	0.1296	0.1655	0.0676

<표 1> RGB vs HSV 색상 거리 분포 비교

5. Conclusion

본 연구는 HSV 색공간 기반 임베딩이 RGB 대비 3D VG에서 색상 표현에 더 적합함을 실증적으로 보였다. 객체의 평균 색상과 대상 색상 간의 거리 분포를 분석한 결과, HSV 공간에서는 RGB보다 평균 거리와 표준편차가

작아, 대상 색상 주변에 객체들이 더 밀집된 양상을 보였다. 이는 HSV 임베딩이 색상 기반의 군집화를 유도하여 명확한 구분 경계를 형성함을 의미한다. 이러한 특성은 색상 유사도 판단이 중요한 3DVG task에서 HSV 기반 표현이 효과적인 선택임을 뒷받침한다.

ACKNOWLEDGMENT

이 논문은 2025년도 정부(개인정보보호위원회)의 재원으로 한국인터넷진흥원의 지원을 받아 수행된 연구임 (RS-2023-00231200, 자율주행 환경에서 AI 학습 가능한 개인 영상정보 프라이버시 보존 기술개발)

참고문헌

- [1] 박민호, 임종민, 차수빈, 오인호, 김광수, “HSV-DBSC AN: Zero-shot 3D Visual Grounding에서의 모델 비의존적 색상 매칭 기법,” 2025년도 한국멀티미디어학회 춘계학술발표대회, under review, 2025.
- [2] A. Dai, A. X. Chang, M. Savva, M. Halber, T. Funkhouser, and M. Nießner, “ScanNet: Richly-annotated 3D Reconstructions of Indoor Scenes,” Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), 2017.
- [3] D. Z. Chen, A. X. Chang, and M. Nießner, “ScanRefer: 3D Object Localization in RGB-D Scans Using Natural Language,” Proc. Eur. Conf. Comput. Vis. (ECCV), 2020, 202 - 221.
- [4] Parzen, E., On Estimation of a Probability Density Function and Mode, The Annals of Mathematical Statistics, 1962, 1065 - 1076.