

ChromaTag: Delta E 색차 제어 기반의 투명 마커 포즈 추정

박유진, 김은경, 박천음*

국립한밭대학교

20191906@edu.hanbat.ac.kr, ekim@hanbat.ac.kr, parkce@hanbat.ac.kr

ChromaTag: Delta E Chromatic Difference Control Based on Transparent Marker Pose Estimation

Yu Jin Park, Eunkyung Kim, Cheoneum Park*

Hanbat National University

요약

본 연구는 흑백 ArUco 마커를 인간의 시각적 방해 최소화하는 투명 ChromaTag로 제작하고, 이를 이용해 6-DOF 포즈를 추정하는 프레임워크를 제안한다. 제안 방법은 인간 시각 인지 모델인 ΔE 색차를 기반으로 투명 마커를 합성하고, YOLO-Pose 모델을 통해 마커의 4개 코너점을 직접 검출하여 포즈를 추정한다. 모델 학습을 위해, 기존 A4 문서 YOLO 데이터셋 기반으로 총 12,710개의 학습데이터를 구축하며, ΔE 난이도에 따른 가중 샘플링을 적용한다. 실험 결과, 제안 방법은 포즈 추정 태스크에서 mAP@[.5:.95] 기준 0.99 성능을 달성하였다. 동일한 실험 조건에서 기존 ArUco 마커의 OKS mAP@[.5:.95]가 0.984를 기록한 것과 비교할 때, ChromaTag는 낮은 ΔE 조건에서도 포즈 추정 정확도와 키포인트 검출 안정성을 유지하여 투명 마커 시스템의 실용적 가능성을 입증하였다.

I. 서론

증강현실(AR)과 로봇틱스 분야에서 객체의 정확한 6-DOF 포즈 추정 은 널리 사용하는 기술이다. 기존의 ArUco나 AprilTag 같은 가시적 마커 는 높은 인식률을 제공하지만, 뚜렷한 흑백 패턴으로 인해 시각적 방해 를 유발하고 한계가 있다[1, 2]. 이러한 문제를 해결하기 위해 투명 마커에 대 한 연구가 활발히 진행되고 있다. InfraredTags [3]는 적외선 형광 물질을 활용한 방법을 제시했으나, 특수한 IR 카메라와 필터링 장비가 필요하다. iMarkers [4]는 자외선 기반 방법을 제안했지만 특수 장비에 의존성이 존 재한다. 이러한 접근법들은 추가 장비로 인한 비용과 확장성에 대한 어려 움이 있다. 이에 본 연구는 별도의 하드웨어 없이 표준 RGB 카메라 환경 에서 인간의 시각 인지 모델을 역이용하는 ChromaTag 프레임워크([6]과 다름)를 제안한다.

II. 본론

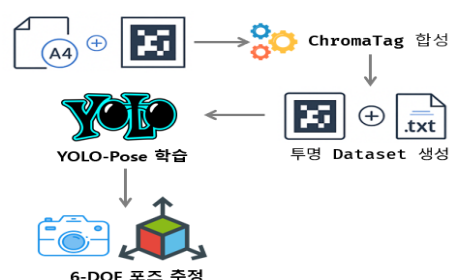


Figure 1. ChromaTag 프레임워크 파이프라인 개요

본 논문에서 제안하는 ChromaTag 프레임워크는 [Fig.1]과 같다. 이는 인간의 시각 시스템이 특정 색차(ΔE) 변화에 둔감하다는 점을 이용, 사람 이 인지하기 어려운 색상 신호를 마커에 인코딩하고 YOLO-Pose 모델을

통해 검출 및 포즈를 추정하는 방식이다.

본 프레임워크는 인간의 시각 인지 모델과 유사한 CIE 2000 ΔE 색차 공식을 활용한다. 이 공식은 밝기(L)와 두 색상 채널(a, b)을 분리하는 CI ELAB 색 공간을 기반으로 한다. 인간의 눈은 밝기(L)의 변화에는 매우 민감하지만, 색상(a, b)의 미세한 변화에는 상대적으로 둔감하다[5].

이를 기반으로, 제안 방법은 마커가 합성될 영역의 L 값을 원본 배경과 유사하도록 유지하면서, 마커의 패턴 정보를 인간이 잘 인지하지 못하는 a와 b 채널에 인코딩한다. 최종 픽셀 색상은 사용자가 설정한 목표 target ΔE (예: 2.0)와 A+C 필드(색차 분배 지도)를 결합하는 것으로 결정된다. A와 C 필드는 다음과 같다.

- A-Field (저주파 위장): 마커의 전반적인 영역을 담당한다. 가우시안 차분(Difference of Gaussians, DoG) 필터를 사용하여 마커의 날카로운 경계선(고주파)을 제거하고, 부드러운 저주파 구조만 남긴다. 이를 통해 마커의 존재감을 최소화하는 위장 효과를 극대화한다.
- C-Field (코너 강화): 모델이 코너를 찾도록 보조 신호를 코너에 추가 하는 것으로, 특정 방향성을 가진 미세한 줄무늬 패턴인 가보르(Gabor) 필터를 네 개의 코너에 추가한다. 사람이 인지하기 어렵지만, 모델 은 학습을 통해 코너의 위치 예측이 가능하게 된다.

각 픽셀에 적용될 실제 목표 색차 local ΔE 는 target $\Delta E \times A+C$ Field (x,y)로 계산되며, 이 값을 만족하도록 a*, b* 채널을 미세 조정하여 최종 색상을 결정한다.

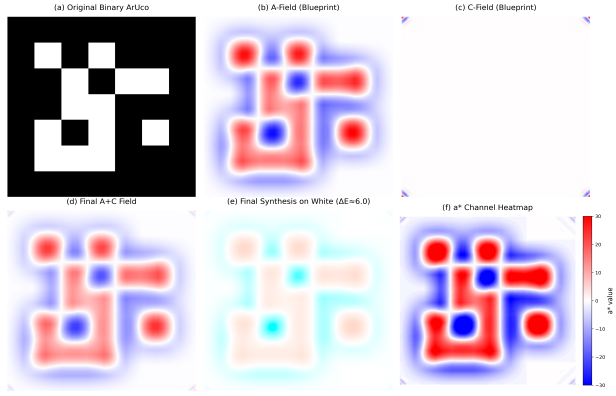


Figure 2. ChromaTag A+C 변조 필드 합성 (a) ArUco 마커. (b) A-Field. (c) C-Field. (d) A+C 필드. (e) 합성 ($\Delta E \approx 6.0$), (f) 결과물의 a^* 채널 히트맵(시각화를 위해 신호 세기는 실제보다 강조됨)

[Fig.2]는 ChromaTag의 A+C 변조 과정을 시각화한다: (a)의 원본 마커로부터 (b) DoG 기반 A-Field와 (c) 가보르 기반 C-Field를 결합하여 (d) 최종 변조 필드를 생성하며, 이를 통해 (e) 합성된 결과는 색상 채널에 (f) 명확한 신호를 인코딩하면서도 시각적으로 인지하기 어렵다.

생성된 변조 필드(d)를 깨끗한 배경에 적용한 최종 합성 결과물(e)은 원본의 명확한 형태가 사라지고 배경과 자연스럽게 융화된 패턴을 보인다. 마지막으로 a^* 채널 히트맵(f)은 해당 패턴이 시각적으로는 인지하기 어려우나, 색상 채널에는 데이터로서 명확히 인코딩되었음을 나타낸다. 공정한 성능 비교를 위해 동일한 학습 파이프라인에서 ArUco와 직접 비교 실험을 수행하였다.

A4 문서 YOLO 데이터셋 기반으로 ΔE 난이도별 가중 샘플링을 적용하여 총 12,710개 합성 데이터를 구축하였다. 최적화된 DoG/Gabor 파라미터로 A+C 변조를 수행하였다. 대용량 데이터셋의 실시간 ΔE 계산을 위해 colormath 라이브러리의 CPU 기반 연산을 GPU 병렬 처리로 최적화하여 처리 속도 향상을 달성하였다.

모든 데이터는 마커 당 4개의 키포인트를 갖는 YOLO Pose 형식으로 레이블링 된다. Yolo11m-pose 모델을 100 epoch 학습하여 테스트 데이터셋으로 검증한 결과는 [Table 1], [Table 2]와 같다. 성능 검증을 위해 동일한 YOLO11m-pose 파이프라인에서 ArUco와 직접 비교하였고, ChromaTag는 OKS mAP@[.5:.95] 0.995 vs 0.984로 우수한 성능을 보였다. 특히 ChromaTag는 모든 ΔE 구간에서 안정적인 정밀도/재현율을 유지하여 투명 마커의 강건성을 확인하였다.

방법	ΔE 범위	Pose OKS mAP@[.5:.95]	Precision	Recall
Ours	ΔE 1~2 (Visible)	0.995	~1.0	~1.0
Ours	ΔE 2~3 (Subtle)	0.992	~1.0	~1.0
Ours	ΔE 3~4 (Invisible)	0.985	~1.0	~1.0
ArUco Marker		0.984	~1.0	0.973

Table 1. ChromaTag vs ArUco 성능 비교(YOLO11m-pose, 동일 조건)

[Fig.3]은 제안 모델의 결과를 시각적으로 보여준다. ΔE 가 낮아져 육안 식별이 어려운 조건(좌측 열)에서도 2D 코너 검출(2행)이 안정적이며, 이를 기반으로 정확한 6-DOF 포즈 추정(3행)이 가능함을 확인할 수 있다

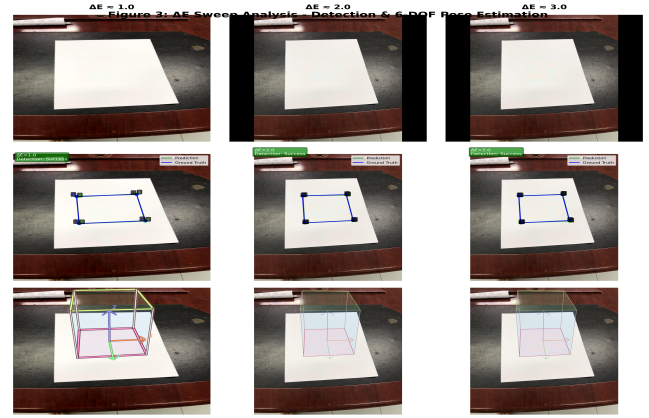


Figure 3. ΔE 난이도에 따른 테스트 데이터셋 기반 ChromaTag의 시각적 은닉성 및 단계별 포즈 추정 결과

III. 결론

본 연구는 ΔE 색차 제어와 A+C 변조를 결합하여 표준 RGB 카메라만으로 동작하는 투명 마커 시스템 ChromaTag를 제안하고, 그 가능성을 보였다. A4 합성 데이터셋 기반 실험에서 달성한 약 0.99의 높은 포즈 mAP 성능은, 특수 장비 없이도 포즈 추정이 가능함을 보인다. 이에 따라, 본 논문에서 제안한 ChromaTag는 사용자 경험을 해치지 않는 비침해 AR을 구현할 수 있으며, 물리적 객체에 정밀한 3D 디지털 정보를 증강시키 것으로 상호작용 시스템 발전에 기여할 것이라 사료된다.

참 고 문 헌

- [1] Garrido-Jurado, S. et al., "Automatic generation and detection of highly reliable fiducial markers under occlusion," Pattern Recognition, 47 (6), 2280–2292, 2014. DOI: 10.1016/j.patcog.2014.01.005.
- [2] E. Olson, "AprilTag: A robust and flexible visual fiducial system," 2011 IEEE International Conference on Robotics and Automation, Shanghai, China, 2011, pp. 3400–3407, doi: 10.1109/ICRA.2011.5979561.
- [3] Mustafa Doga Dogan, Ahmad Taka, Michael Lu, Yunyi Zhu, Akshat Kumar, Aakar Gupta, Stefanie Mueller, CHI 2022, DOI: 10.1145/3491102.3501951.
- [4] Ali Tourani et al., "Unveiling the Potential of iMarkers: Invisible Fiducial Markers for Advanced Robotics," arXiv:2501.15505 (2025).
- [5] Mullen KT, Plant GT. Colour and luminance vision in human optic neuritis. Brain. 1986 Feb;109 (Pt 1):1–13. doi: 10.1093/brain/109.1.1. PMID: 3942852.
- [6] DeGol, J., Bretl, T., & Hoiem, D. (2017). Chromatag: A colored marker and fast detection algorithm. In Proceedings of the IEEE international conference on computer vision (pp. 1472–1481).