

영상 압축 환경에서의 압축 강도에 따른 VGGT 성능 변화 분석

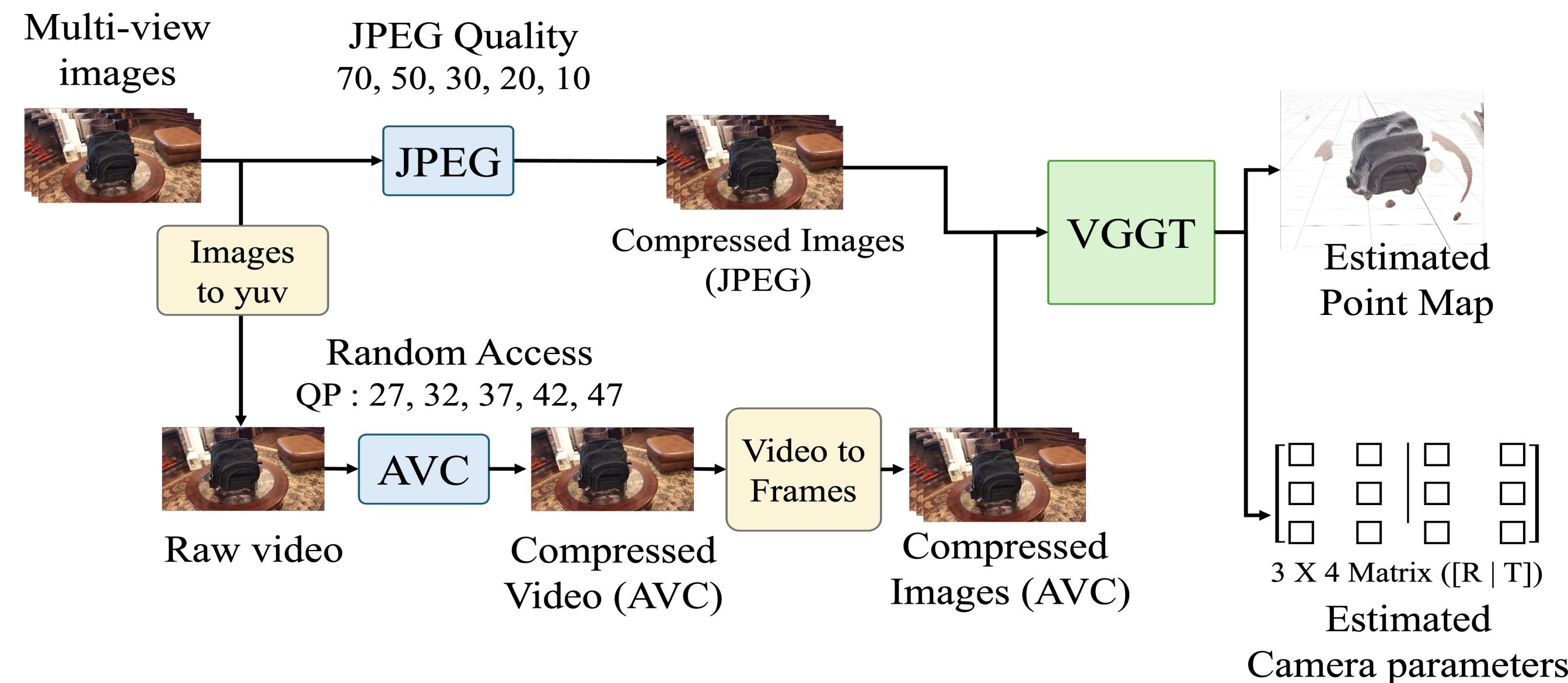
Performance Analysis of VGGT under Various Compression Levels in Lossy Compression

*김제희, *이윤호, *김은지, *문채원, **정진우, *박상효[†]
 *경북대학교, **한국전자기술연구원

Introduction

- 최근 VGGT(Visual Geometry Grounded Transformer)와 같은 3D 비전 모델은 단일 이미지로 카메라 파라미터, 깊이 맵, 포인트 맵을 추론하며 뛰어난 장면 복원 능력을 보여준다.
- 그러나 실제 응용에서는 전송 대역폭, 저장 공간 등의 제약으로 손실 압축 영상이 입력으로 사용되는 경우가 빈번하다. 이로 인한 압축 아티팩트가 모델의 장면 해석 성능 저하를 초래할 가능성 있으나, 현재까지 관련 연구가 없어 실제 응용 안정성을 확보할 수 없다.
- 본 연구는 VGGT 모델을 대상으로 입력 이미지의 압축 강도 변화가 성능에 미치는 영향을 체계적으로 분석하였다.
- CO3Dv2 데이터셋을 기반으로 대표적인 이미지 코덱과 동영상 코덱인 JPEG 및 AVC 코덱을 다양한 압축 강도로 적용하여 실험을 수행하였으며, 포인트 맵의 포인트 수, Chamfer Distance, AUC를 평가지표로 사용하였다.
- 이를 통해 압축 강도에 따른 모델의 재구성 품질 및 성능 저하 양상을 정성적·정량적으로 분석하였다.

Experiments



- CO3Dv2 데이터셋에서 9개의 카테고리 선정.
- JPEG 코덱은 python openCV 라이브러리를 사용하여 Quality 70, 50, 30, 20, 10으로 설정.
- AVC 코덱은 ffmpeg을 사용하여 이미지 형식을 yuv420p로 변환 후 하나의 시퀀스로 이어 붙여 QP를 27, 32, 42, 47로 설정하여 압축 후 프레임 추출.
- 압축된 영상을 1장씩 VGGT의 입력으로 사용하여 포인트 맵, 카메라 포즈를 예측.

Results

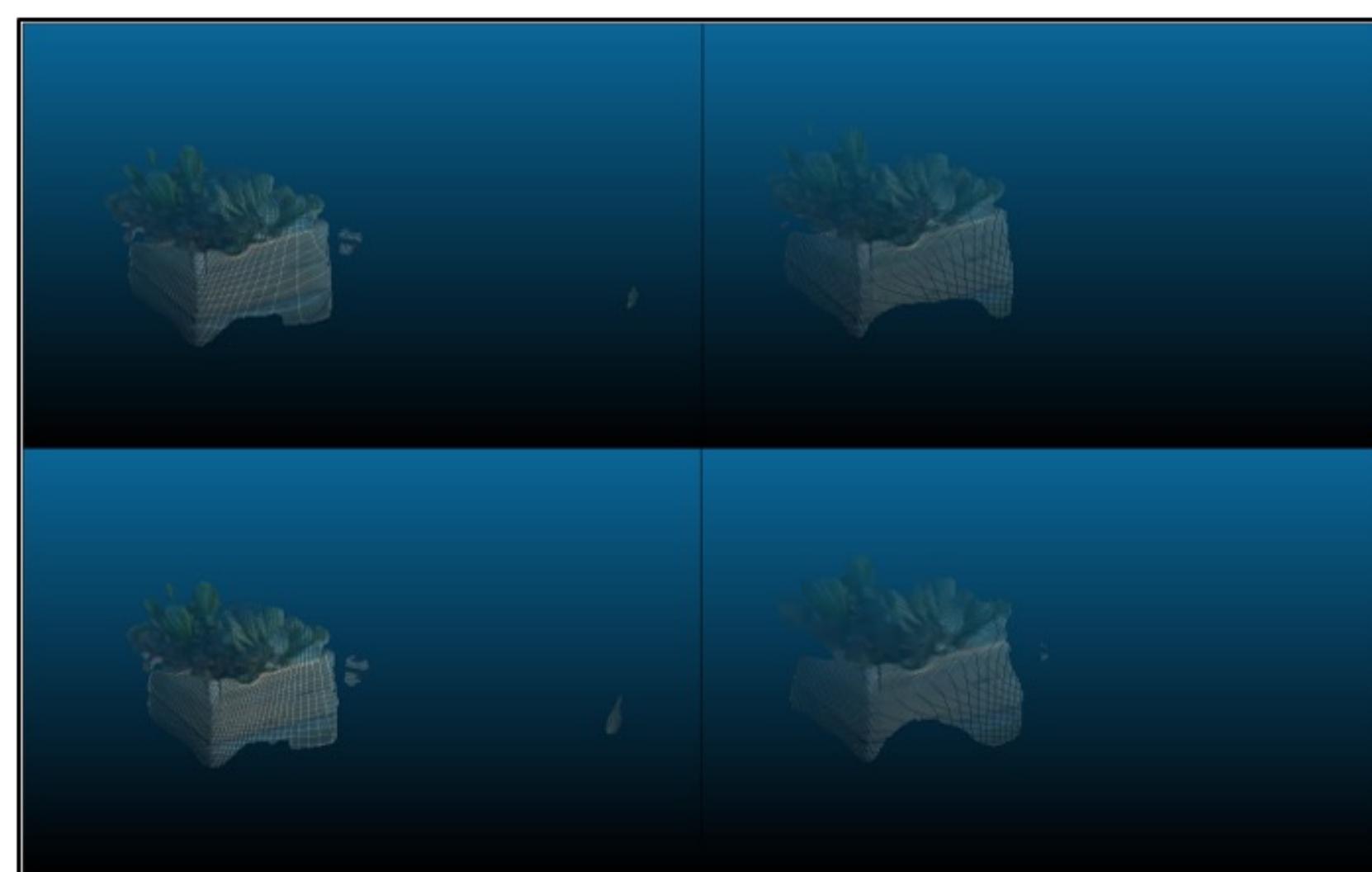


그림 1. plant에 대한 VGGT의 포인트 맵 추론 결과.
(위: JPEG Quality 70, 10, 아래: AVC QP 27, 47)

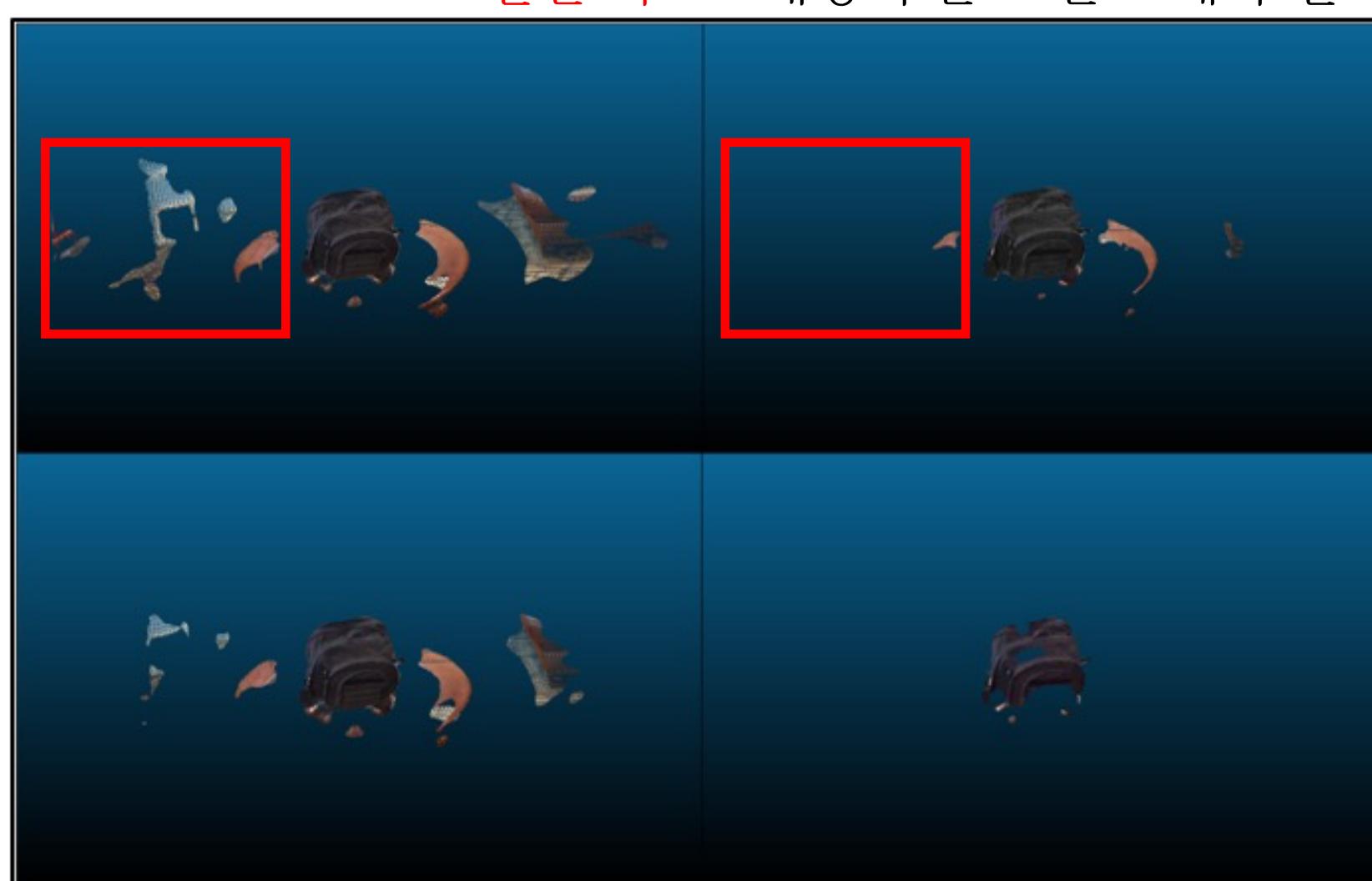


그림 2. backpack에 대한 VGGT의 포인트 맵 추론 결과.
(위: JPEG Quality 70, 10, 아래: AVC QP 27, 47)

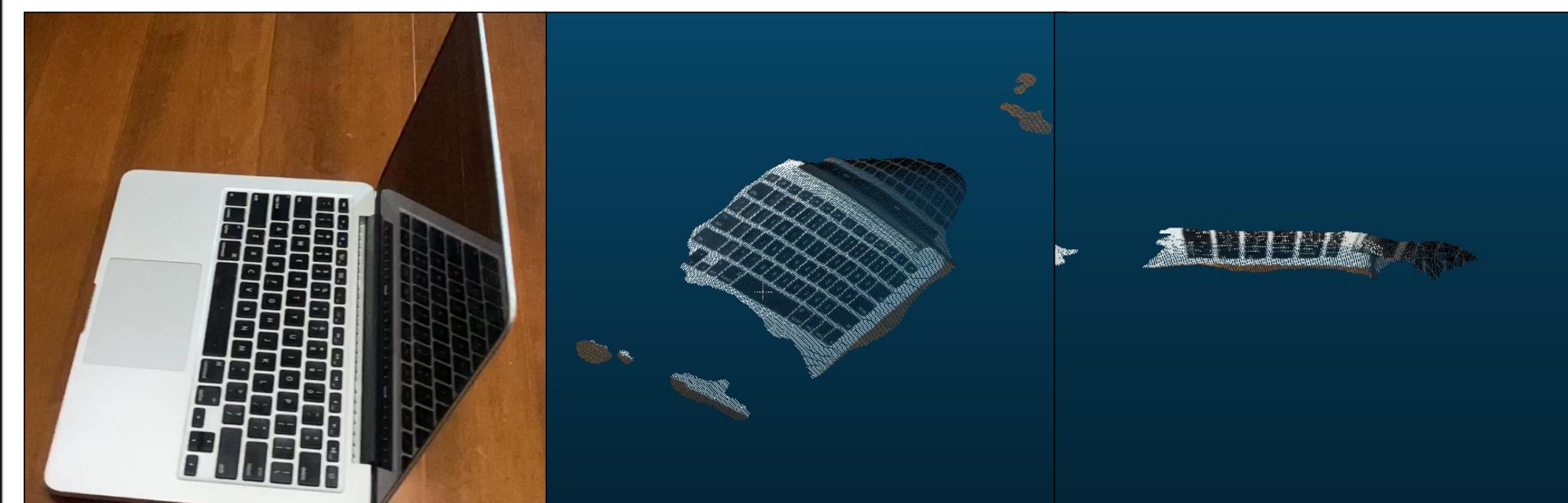


그림 3. Laptop 입력 이미지와 VGGT의 포인트 맵 추론 결과.
(순서대로 입력 이미지, 예측 포인트 맵 상단 시점, 옆면 시점)

Codec	압축강도	포인트 수	PSNR ↑
Raw image	-	26614.91	-
AVC	QP 27	25234.78	38.49
AVC	QP 32	22855.32	36.35
AVC	QP 37	20012.51	34.04
AVC	QP 42	16038.15	31.56
AVC	QP 47	15355.68	28.68
JPEG	Quality 70	26173.84	47.33
JPEG	Quality 50	24491.97	40.46
JPEG	Quality 30	22597.27	38.50
JPEG	Quality 20	20148.48	35.95
JPEG	Quality 10	16450.08	31.59

표1. 압축 강도 별 예측된 Point 수와 입력 영상의 PSNR

Codec	압축강도	All category ↓	Laptop ↓
Raw image	-	3.4717	3.3954
AVC	QP 27	3.4721	3.3869
AVC	QP 32	3.4923	3.4029
AVC	QP 37	3.5279	3.4177
AVC	QP 42	3.5616	3.4067
AVC	QP 47	3.5884	3.3267
JPEG	Quality 70	3.4801	3.3850
JPEG	Quality 50	3.5059	3.3896
JPEG	Quality 30	3.5075	3.4375
JPEG	Quality 20	3.5226	3.4529
JPEG	Quality 10	3.5400	3.4095

표2. 압축 강도 별 Chamfer Distance (포인트 맵의 유사도)

Codec	압축강도	AUC@30 ↑	AUC@15 ↑	AUC@5 ↑	AUC@3 ↑
Raw image	-	0.8737	0.8453	0.7595	0.6848
AVC	QP 27	0.8586	0.8198	0.6933	0.6025
AVC	QP 32	0.8565	0.8147	0.6765	0.5786
AVC	QP 37	0.8524	0.8031	0.6479	0.5383
AVC	QP 42	0.8413	0.7844	0.6044	0.4914
AVC	QP 47	0.8300	0.7539	0.5437	0.4156
JPEG	Quality 70	0.8753	0.8464	0.7615	0.6881
JPEG	Quality 50	0.8781	0.8474	0.7635	0.6881
JPEG	Quality 30	0.8718	0.8434	0.7521	0.6749
JPEG	Quality 20	0.8715	0.8425	0.7432	0.6609
JPEG	Quality 10	0.8766	0.8349	0.7249	0.6329

표3. 압축 강도 별 AUC (카메라 포즈 추정 평가)

- [그림1,2, 표1] 포인트 맵의 시각적 품질은 대체로 유지되었으나, 압축 강도가 높을 수록 배경 부분의 포인트 수가 감소하고 AVC QP 42 이상부터 중심 객체의 포인트 예측 성능도 저하되었다. JPEG의 경우 시각적 품질이 낮아져도 배경 부분의 포인트수가 감소할 뿐, 중심 객체의 포인트 예측엔 유의미한 변화가 없었다.
- [그림3, 표2] 포인트 맵의 유사도를 평가하기 위해 Chamfer Distance를 분석한 결과, 전반적으로 강한 압축일수록 포인트 맵 유사도가 낮아졌으며, 반사되는 부분이 있는 Laptop 카테고리는 모델의 구조적 한계로 비정상적인 결과가 관찰되었다. 이러한 비정상적 추론 결과는 카메라 포즈 추정에서도 나타났다.
- [표3] 카메라 포즈 추정의 경우 AVC는 전반적인 압축 강도 증가에 따라 전반적인 성능 저하가 나타난 반면, JPEG은 AUC@30, AUC@15 기준에서 유의미한 저하가 관찰되지 않았다. 이는 VGGT가 기반으로 하는 비전 트랜스포머의 전역적 구조 인식 특성으로 인해 JPEG 코덱의 세부 시각 정보 손실에 상대적 강건함을 보인 결과로 추정된다.

Conclusion

- 본 연구에서는 CO3Dv2 데이터셋을 기반으로 대표적인 이미지 코덱인 JPEG과 대표적인 동영상 코덱인 AVC 코덱을 적용하여 다양한 압축 강도의 데이터셋을 구성하고, 이를 기반으로 현재 최고 성능의 3D 비전 모델인 VGGT의 성능 변화를 분석하였다.
- 정성적 평가에서는 AVC 코덱의 경우 QP 42 이상부터 중심 물체의 포인트 예측 성능 저하가 나타났으나 JPEG의 경우 큰 변화를 보이지 않았다.
- 압축 강도가 높아질수록 추정된 포인트가 일관적으로 감소하였고, 추론된 포인트 맵 유사도 역시 저하되었다.
- 카메라 포즈 예측은 AVC의 경우 모든 임계값에서 성능 저하가 나타났지만, JPEG은 AUC@30, AUC@15에 대해서 성능 저하가 나타나지 않았다.
- 종합적으로 **AVC 코덱은 QP 42 이상, JPEG 코덱은 Quality 20 이하부터 성능 저하가 급격히 발생하였다.** 본 결과는 일정수준 이하의 압축 환경에서는 3D 비전 모델의 성능이 유지될 수 있음을 시사하나, **강한 압축 환경에서도 강건한 모델 설계의 필요성을 보여준다.**