

기계 대상의 피쳐 부호화를 위한 에지 보존 가이드 필터링 기법

서주연¹, 장호원¹, 김연희², 이주영², 정세윤², 고현석^{1*}

¹ 한양대학교, ² 한국전자통신연구원

hyeseo@hanyang.ac.kr, jhwon238@hanyang.ac.kr, kimyounhee@etri.re.kr,
leejy1003@etri.re.kr, jsy@etri.re.kr, *hyunsuk@hanyang.ac.kr

Edge-preserving Guided Filtering for Feature Coding for Machines (FCM)

Juyeon Seo¹, Howon Jang¹, Younhee Kim², Jooyoung Lee², Se Yoon Jeong²,
and Hyunsuk Ko^{1*}

¹ Hanyang University, ² Electronics and Telecommunications Research Institute (ETRI)

요 약

최근 머신 비전(machine vision) 응용의 확산으로 중간 피쳐(feature)를 효율적으로 전송하기 위한 부호화 기술의 중요성이 커지고 있다. 이를 위해 MPEG에서는 FCM 표준화를 진행하고 있다. 본 논문에서는 내부 코덱(inner codec)에서의 부호화 효율을 향상시키기 위해, 가이드 필터(guided filter)를 활용한 에지 보존 전처리 기법을 제안한다. 제안 기법은 저활성 영역을 효과적으로 억제하면서도 경계 정보를 보존하여 압축 효율을 향상시킨다. 실험 결과, FCTMv8.1 대비 전체 평균 2.01%의 BD-rate 이득을 달성했으며, 특히 객체 추적 작업에서 평균 4.28%의 BD-rate 이득이 확인되었다. 또한 인코더 및 디코더 처리 시간 변화가 미미하여 제안 기법이 연산 효율을 유지하면서도 성능을 향상시킴을 확인하였다.

I. 서 론

최근 인공지능 기술의 발전으로 객체 탐지, 분할, 추적과 같은 머신 비전(machine vision) 응용 분야가 빠르게 확장되고 있다. 이에 따라 중간 피쳐(feature)를 기기 간에 효율적으로 전송하기 위한 기술의 중요성이 높아지고 있다. 그러나 기존의 전통적 영상 부호화 기술은 인간의 시각 특성에 최적화되어 설계되었기 때문에, 기계 학습 모델이 활용하는 피쳐의 분포나 의미적 구조를 충분히 반영하지 못하는 한계가 있다.

이러한 문제를 해결하기 위해 MPEG(Moving Picture Experts Group)에서는 FCM(Feature Coding for Machines) 표준화를 추진하고 있으며 [1, 2], 피쳐 맵의 전송 효율 향상을 위한 다양한 압축 및 전처리 기법이 연구되고 있다. 최근에는 양방향 필터(Bilateral filter) 기반 전처리를 인코더 측에 적용하여 부호화 성능을 향상시키는 방법이 제안된 바 있다 [3]. 그러나 양방향 필터는 지역 윈도우 내 모든 픽셀 쌍에 대해 연산을 수행해야 하므로 계산 복잡도가 높아 실제 적용 시 부담이 되는 문제가 존재한다.

본 논문에서는 이러한 한계를 해결하기 위해, 양방향 필터를 보다 낮은 계산 복잡도로 대체할 수 있는 가이드 필터(guided filter) 기반 전처리 기법을 제안한다. 제안 방법은 저활성 영역의 불필요한 세부 정보를 억제하고 의미적 경계를 보존함으로써, 중간 피쳐의 효율적 전송을 가능하게 한다.

II. 본 론

제안하는 에지 보존 가이드 필터링 기법은 FCTM(Feature Compression Test Model) 인코더의 내부 코덱(inner codec) 입력 단계에서 적용된다. 해당 기법은 압축 효율 향상을 목적으로 인코더 측에서만 수행되며, 디코더 측에서 추가적인 신호화(signaling) 정보를 필요로 하지 않는다. 전체 파이프라인은 그림 1에 나타나 있으며, packed frame에 대해 가이드 필터가 적용된 후 내부 코덱으로 전달되는 구조를 가진다.

본 논문에서는 가이드 필터의 여러 변형 중 self-guided filter 방식을 사용한다. 이는 입력 피쳐 자체를 가이드 이미지로 활용하여 필터를 적용하는 방식으로, 외부 정보 없이도 국소적 구조의 특성을 효과적으로 반영할 수 있다는 장점을 갖는다. 특히 self-guided filter는 국소적 선형 모델에 기반하여 평탄한 영역을 부드럽게 억제하고, 에지 정보는 유지하는 특성이 있어 내부 코덱에서 사용되는 VTM 부호화 효율 향상과 머신 비전 작업 성능의 유지 측면에서 모두 유리하다.

가이드 필터는 로컬 윈도우 내에서 수식 1과 같은 선형 모델을 가정한다.

$$q_i = a_k l_i + b_k, \quad i \in w_k \quad (1)$$

여기서 I 는 입력 피쳐이자 가이드 이미지이며, q_i 는 필터링된 출력을 의미한다. 또한 $i \in w_k$ 는 중심 위치 k 를 갖는 로컬 윈도우 w_k 내의 위치 i 를 나타낸다. 그림 1의 y_q 와 y_f 가 각각 I 와 q 에 해당되며, 선형 계수 a_k 와 b_k 는 윈도우 평균 및 분산으로부터 수식 2와 같이 계산된다.

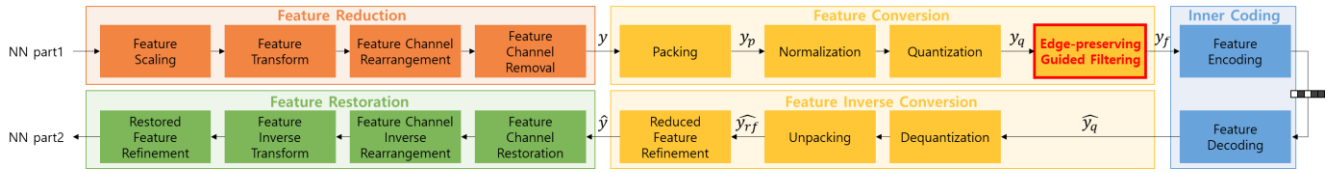
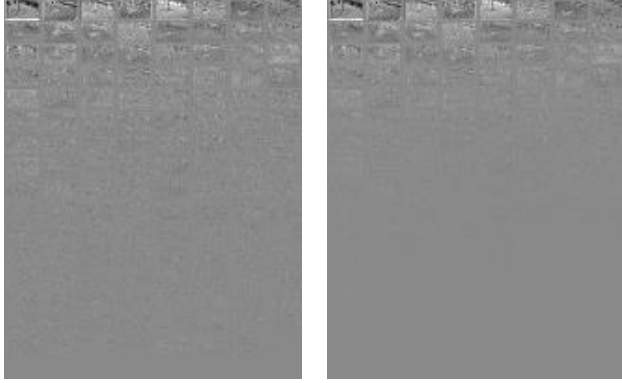


그림 1. 제안 방법을 포함한 FCTM 전체 파이프라인



(a) 필터 미적용

(b) 필터 적용

그림 2. Self-guided filter 적용 전후의 packed frame 비교 예

$$a_k = \frac{cov_{w_k}(I, I)}{\sigma_{I,k}^2 + \epsilon}, \quad b_k = (1 - a)\mu_{I,k} \quad (2)$$

본 논문에서는 윈도우 반경 $w_k = 12$, 정규화 항 $\epsilon = 0.1$ 을 사용하였다. 이를 통해 저활성 영역의 잡음성 세부 성분은 효과적으로 평탄화되고, 머신 비전 모델의 성능에 중요한 의미적 경계는 보존된다. 그림 2는 self-guided filter 적용 전후의 packed frame을 비교한 예시이다. 필터 적용 후 중후반 채널에서 나타났던 노이즈성 세부 구조가 크게 감소하고, 초반부 채널에 존재하는 디테일 구조는 유지되는 것을 확인할 수 있다.

또한 제안 방법은 양방향 필터에 비해 계산 복잡도가 크게 낮다는 이점을 갖는다. 가이드 필터는 누적합 기반 박스 필터링(box filtering)을 사용하므로 전체 연산 복잡도는 $O(N)$ 수준으로 유지된다. 반면, 양방향 필터는 윈도우 내 모든 픽셀 쌍의 거리를 계산해야 하므로, 윈도우 반경 r 을 사용할 때 전체 연산 복잡도는 일반적으로 $O(Nr^2)$ 수준에 이른다. 즉, 제안 방법은 양방향 필터 대비 연산량을 크게 줄이면서도 유사한 에지 보존 효과를 제공할 수 있다.

제안 방법의 성능을 검증하기 위해 CTTC(Common Test and Training Conditions) [4] 환경에서 객체 검출(object detection), 객체 추적(object tracking), 인스턴스 분할(instance segmentation), 의미론적 분할(semantic segmentation) 작업을 대상으로 실험을 수행하였다. 사용된 데이터셋은 OpenImageV6, SFU, TVD, HIEVE, PandaSet이며, 성능 평가는 FCTMv8.1을 기준으로 BD-rate와 인코더 및 디코더 처리 시간의 상대적 변화를 나타내는 인코딩 비율 변화(EncR) 및 디코딩 비율 변화(DecR)를 측정하였다.

표 1의 결과에서 확인할 수 있듯이, 제안 방법은 FCTMv8.1 대비 전체 평균 2.01%의 BD-rate 이득을 보였으며, 특히 객체 추적 작업에서는 4.28%의 BD-rate 이득을 기록하여 필터링에 의한 성능 개선이 뚜렷하게 확인되었다. 또한 인코더와 디코더의 처리 시간 변화는

표 1. FCTMv8.1 대비 제안 방법의 성능 비교

Task	Dataset	BD-rate	EncR	DecR
Instance segmentation	OpenImageV6	3.51%	100.51%	105.84%
Object detection	OpenImageV6	-1.52%	100.85%	108.66%
	SFU (Class A/B)	-1.17%	98.82%	104.88%
	SFU (Class C)	-0.40%	99.06%	110.00%
	SFU (Class D)	-3.61%	98.86%	103.45%
Object tracking	TVD (Overall)	-9.47%	103.41%	100.51%
	HIEVE (1080p)	-2.24%	96.75%	101.99%
	HIEVE (720p)	-1.14%	99.56%	101.10%
Semantic segmentation	PANDAM1	-6.67%	99.62%	104.32%
	PANDAM2	3.43%	99.29%	105.88%
	PANDAM3	1.00%	99.73%	103.81%
OVERALL		-2.01%	100.25%	105.55%

각각 전체 평균 약 100.25% 및 105.55% 수준으로 거의 증가하지 않아 연산 부담이 크지 않음을 확인하였다.

이러한 결과는 self-guided filter 기반 전처리가 FCM 환경에서 연산 효율을 유지하면서도 경계 보존 능력과 저복잡도를 동시에 달성할 수 있는 효과적인 접근임을 보여준다.

III. 결 론

본 논문에서는 내부 코덱 단계의 부호화 효율을 향상시키기 위해 에지 보존 가이드 필터링 기법을 제안하였다. 제안 기법은 낮은 복잡도를 유지하면서 저활성 영역을 효과적으로 억제하고 에지 정보를 보존하여 FCTMv8.1 대비 평균 2.01%, 객체 추적 작업 기준 4.28%의 BD-rate 이득을 보였다. 이러한 결과는 제안 방법이 연산 부담을 크게 증가시키지 않으면서도 실질적인 압축 성능 향상을 제공하는 도구로 활용될 수 있음을 보여준다.

ACKNOWLEDGMENT

이 논문은 2026년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No. 2020-0-00011, (전문연구실)기계를 위한 영상부호화 기술)

참 고 문 헌

- [1] "Call for Proposals on Feature Compression for Video Coding for Machines," ISO/IEC JTC 1/SC 29/WG 2, N00282, Apr. 2023.
- [2] "Call for Proposals responses report for Feature Compression for Video Coding for Machines," ISO/IEC JTC 1/SC 29/WG 2, N00318, Oct. 2023.
- [3] T. Ji and S. Deshpande, "[FCM] CE4.1.2: Bilateral filtering encoder only," ISO/IEC JTC 1/SC 29/WG 4, m74024, Oct. 2025.
- [4] "Common Test and Training Conditions for FCM," ISO/IEC JTC 1/SC 29/WG 4, N0705, Jul. 2025.