Check for updates

특집논문 (Special Paper)
방송공학회논문지 제30권 제3호, 2025년 5월 (JBE Vol.30, No.3, May 2025)
https://doi.org/10.5909/JBE.2025.30.3.355
ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

FCM을 위한 비신경망 기반 피쳐 변환 기술

김 연 희[a)‡], 정 세 윤[a)], 이 주 영[a)]

Non-NN based Feature Transform Methods for FCM

Younhee Kim[a)‡], Se Yoon Jeong[a)], and Jooyoung Lee[a)]

요 약

이 논문은 기계 학습 기반 비디오 처리 환경에서 효율적인 특징 정보 전송을 위한 FCM(Feature Coding for Machines) 기술 중, 비신경망 기반 피쳐 변환 기술(Non-NN-based Feature Transform)에 대해 다룬다. FCM은 클라우드-엣지 분할 컴퓨팅을 활용 시나리오로 하며, 중간 특징 데이터를 효과적으로 압축·전송하기 위한 기술이다. 본 논문에서는 특히 PCA(주성분 분석) 기반의 두 가지 주요 기술인 PreBM PCA와 Online PCA를 소개하고 분석한다. PreBM PCA는 사전 샘플 데이터로부터 추출한 주성분 벡터를 활용하여 연산 부담을 줄이면서 높은 기계 학습 성능을 유지하는 기술이고, Online PCA는 실시간 데이터 변화에 적용 가능하며, 네트워크 구조와 분할 지점(split point)에 제약 없이 적용할 수 있는 FCM을 위한 기술이다. 본 논문은 각 기술의 구조, 구현 방법, 성능을 비교·분석하고 향후 FCM 표준 내에서의 활용 가능성과 발전 방향을 제시한다.

Abstract

This paper explores Non-NN-based feature transform techniques for Feature Coding for Machines (FCM), which aims to efficiently compress and transmit intermediate features extracted at split points in machine learning models. Designed for split computing between edge devices and the cloud, This work focuses on two major PCA-based approches: PreBM PCA and Online PCA. PreBM PCA reduces computational overhead using predefined basis vectors, while Online PCA adapts to real-time data changes and is applicable across various network architectures and split points. These techniques offer advantages such as low complexity, high versatility, and preservation of inference accuracy, making them strong candidates for future inclusion in the FCM standard. The paper analyzes their structure, implementation, and performance, and discusses their applicability and future directions.

Keyword : Feature Coding for Machines (FCM), Principal Component Analysis (PCA), PreBM PCA, Online PCA, Split Computing

a) 한국전자통신연구원 미디어연구본부(Media Research Division, Electronics and Telecommunications Research Institute)
‡ Corresponding Author : 김연희(Younhee Kim)
E-mail: kimyounhee@etri.re.kr
Tel: +82-42-860-5407
ORCID: https://orcid.org/0000-0003-0658-6762

※ This work was supported by the Institute of Information and Communications Technology Planning and Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2020-0-00011, Video Coding for Machine).
· Manuscript March 28, 2025; Revised May 8, 2025; Accepted May 8, 2025.

Copyright ⓒ 2025 Korean Institute of Broadcast and Media Engineers. All rights reserved.
"This is an Open-Access article distributed under the terms of the Creative Commons BY-NC-ND (http://creativecommons.org/licenses/by-nc-nd/3.0) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited and not altered."

Ⅰ. 서 론

최근 인공지능(AI) 및 기계 학습 기반 영상 처리 기술이 발전함에 따라 비디오 부호화 기술은 새로운 도전에 직면하고 있다. 기존의 비디오 압축 기술(예: HEVC, VVC)은 인간의 시각적 품질을 유지하는 데 초점을 맞추었으나, 점점 기계 학습 태스크의 성능 유지를 위한 데이터 전송 및 부호화 효율성을 높이는 부호화 기술에 대한 요구가 높아지고 있다.

FCM(Feature Coding for Machines)은 이러한 요구를 반영하여, 기계 학습 모델에서 생성되는 중간 특징 데이터를 효율적으로 부호화하고 전송하는 기술을 표준화하는 것을 목표로 한다[1,2]. FCM은 ISO/IEC JTC 1/SC 29/WG4에서 표준화가 진행 중으로 VCM(Video Coding for Machines)과는 달리, 네트워크 모델을 두 개의 부분(NN Part 1과 NN Part 2)으로 분할하여, 단말에서 NN Part 1을 처리하고, 서버에서 NN Part 2를 처리하는 클라우드 기반 연산과 엣지 디바이스 간의 데이터 전송 및 분할 컴퓨팅(Split Computing) 방식을 활용 시나리오로 상정하고 있다[3,4,5,6].

FCM에 제안된 기술은 피쳐 축소 부분에 신경망 기반 기술을 사용하는지 여부에 따라 신경망 기반 축소(NN-based reduction) 방식과 비신경망 기반 축소(Non-NN-based reduction) 방식으로 나뉜다. 마찬가지로, 내부 부호화기(inner codec) 역시 신경망 기반 기술의 적용 여부에 따라 신경망 기반 내부 부호화(NN-based inner coding) 방식과 비신경망 기반 내부 부호화(Non-NN-based inner coding) 방식으로 구분된다. 현재 FCM의 참조 소프트웨어인 FCTM(FCM Test Model)에서는 신경망 기반 축소와 비신경망 기반 내부 부호화 방식을 채택하고 있다. 내부 부호화 방식으로는 최신 비디오 압축 표준인 VVC(Versatile Video Coding)를 채택하여 사용하고 있으며 VTM(Versatile Video Coding Test Model) 버전 23.3을 사용하고 있다. 즉 중간 피쳐를 신경망 기반 방법으로 축소한 후 비디오 데이터 형식으로 변환해 기존 비디오 압축 코덱인 VVC를 활용하여 인코딩이 가능하도록 한다.

본 논문에서는 비신경망 기반 축소(Non-NN-based reduction) 접근 방식에 대한 기술들을 중심으로 살펴보고자 한다. FCM에 제안된 비신경망 기반 축소(Non-NN-based reduction) 접근 방식의 대부분은 PCA(주성분 분석, Principal Component Analysis) 기법을 활용하고 있다. PCA 기반의 특성 변환은 객체 검출, 추적 등과 같은 비전 과제에서 차원 축소와 함께 학습 성능을 유지하거나 개선하는 데 기여한 바 있으며, 여러 선행 연구에서도 그 유효성이 입증된 바 있다. 특히 제한된 연산 자원 환경에서는 PCA를 통한 전처리가 계산 효율성과 일반화 성능 측면에서 실질적인 이점을 제공할 수 있다[7,8,9]. PCA는 데이터의 주요 정보를 보존하면서 차원을 효과적으로 축소할 수 있는 강력한 압축 기법으로, FCM 표준화 과정에서 여러 연구 기관과 기업들이 제안한 핵심적인 접근 방식 중 하나이다[10,11].

기존의 영상 압축 기술은 인간이 시각적으로 인식하는 품질을 유지하는 데 중점을 두고 있지만, FCM에서는 영상이 아닌 기계 학습 모델의 정확도를 유지하면서 데이터 전송량을 최소화하는 것이 목표이다. PCA는 이러한 목표를 달성하기 위한 핵심 기술 중 하나로, 다음과 같은 장점이 있다.

- 차원 축소를 통한 효율적인 데이터 압축: PCA는 고차원의 특징 데이터를 선형 변환하여 저차원으로 축소하면서도, 데이터의 중요한 정보를 유지할 수 있다.
- 컴퓨팅 자원 절약: 기존의 비디오 압축 기법과 비교하여, PCA는 변환 행렬을 사용하여 계산량을 줄일 수 있어 인코더 및 디코더의 복잡도를 줄일 수 있다.
- 기계 학습 성능 유지: PCA 기반 압축은 데이터의 중요 정보를 보존하는 방식으로 수행되기 때문에, 영상 품질이 아니라 객체 검출, 추적 등의 기계 학습 성능을 유지하는 데 유리하다.

본 논문에서는 PCA 기반의 FCM 제안 기술들을 분석하고, FCM 표준 내에서 PCA 기반 기술의 활용 가능성과 향후 발전 방향을 제시하고자 한다.

Ⅱ. 본 론

FCM에 제안된 대표적인 PCA 기반 피쳐 부호화 기술은 ECS(Energy Compensation Scaling) 기법을 적용한 PreBM

PCA 기술과 Online-PCA 기술을 꼽을 수 있다. FCM CfP(Call for Proposal) 응답 기술로 제출된 이 두 기술은 CfP 이후 149차 MPEG 회의 때까지 CE2에서 계속 추가 실험을 진행하였다[12].

1. PreBM PCA 기술

1.1 개요

PreBM(Predefined Bases vectors and Mean values) PCA 기술은 기계 학습 모델의 중간 특징 데이터를 효율적으로 부호화하기 위해, 미리 정의된 PCA 주성분 벡터(Predefined Basis Vectors)와 평균(Mean Values)을 활용하여 데이터 압축 및 변환을 수행하는 기술이다. 본 기술은 FCM 표준화 과정에서 제안되었으며, 특징맵(Feature map)의 차원을 줄이고 저장 및 전송 효율을 높이는 것을 목표로 한다[10,13].

PreBM의 핵심 개념은 PCA를 기반으로 한 변환 기법을 적용하되, 실시간 연산을 줄이기 위해 사전 정의된 주성분 벡터 및 평균값을 활용하는 것이다. PreBM PCA 기술의 특징은 다음과 같다[14,15,16]:

- 사전 정의된 주성분 벡터 사용: 미리 학습된 주성분 벡터를 사용하여, 실시간 변환 과정에서 PCA 주성분 벡터를 구하는 연산을 수행할 필요가 없다.
- 평균값 보정: 입력 특징맵에서 평균값을 미리 계산하여 저장하며, 이를 디코딩 과정에서 복원함으로써 데이터의 손실을 최소화한다.
- 에너지 보상 스케일링(ECS, Energy Compensation Scaling): PCA 변환 과정에서 발생하는 특징 크기 변화로 인한 성능 저하를 보정하는 기법으로, 디코딩 과정에서 추가적인 스케일링을 적용한다.

1.2 요소 기술

1.2.1 PreBM 생성

- 다양한 데이터셋(예시: OpenImagesV6, HiEve 비디오 데이터셋)에서 성능 평가를 위해 사용할 데이터를 제외한 데이터셋에서 샘플 데이터를 수집하여 부호화 대상 특징맵을 추출한다[3].
- 수집한 샘플들로부터 특징맵을 추출, 병합하여 PCA 주성분 벡터를 도출한다.
- 도출한 벡터에서 특정 개수의 주성분 벡터만을 선정하여 저장한다.
- 특징맵의 전반적인 평균을 보정할 수 있도록 수집한 샘플 특징 맵의 평균값을 계산하여 저장한다.
- 이 과정은 인코딩 및 디코딩 전에 한 번만 수행되므로, 실시간 연산 부담을 줄일 수 있다.
- 생성된 PreBM은 인코더와 디코더가 공유하므로 전송 데이터를 줄일 수 있다.

1.2.2 인코딩

PreBM 기술의 인코딩은 아래와 같은 과정은 거쳐 이루어지고 그림 1은 이를 도식화하였다.

- ECS(Energy Compensation Scaling) 매개변수를 추출한다.
- 다계층 특징맵 병합: 픽셀 재배열(Pixel shuffling)을 통해 다계층 특징맵을 병합하여 단일 층의 특징맵을 생성한다.
- PCA 변환 적용: 사전 정의된 주성분 벡터를 활용하여 특징맵을 변환한다. 이때 기저 벡터의 수를 제한하여 출력 계수 차원을 축소하게 된다.
- 양자화 및 패킹: 변환 계수 데이터를 10비트 정밀도로 양자화하고, 효율적인 압축을 위해 정렬된 형식으로 패킹한다.
- VVC(Versatile Video Coding, 비디오 압축 표준) 인코딩: 압축된 특징 데이터를 VTM(VVC Test Model)을 활용하여 비디오 프레임 형식으로 인코딩한다.

PreBM PCA의 성능 핵심 요소 기술 중 하나는 ECS (Energy Compensation Scaling) 기법으로 이는 PCA 변환 후 특징맵 신호의 크기가 변하는 것을 보정해 주는 기술이다[17]. 그림 2는 PCA 차원 축소 과정에서 특징(Feature)의 분포가 어떻게 변하는지를 보여준다. 그림 2에서 알 수 있듯이 역변환 후 특징값(feature)의 최대값은 원래 특징값의 최대값보다 작아진다. 하여 인코더에서는 원본 특징(feature)으로부터 향후 디코더에서 역변환된 특징(feature)을 원본 특징의 스케일과 일치하도록 조정할 수 있는 매개변수를 추출한다. 이 스케일링 값(S_i)은 식 (1)을 이용하여

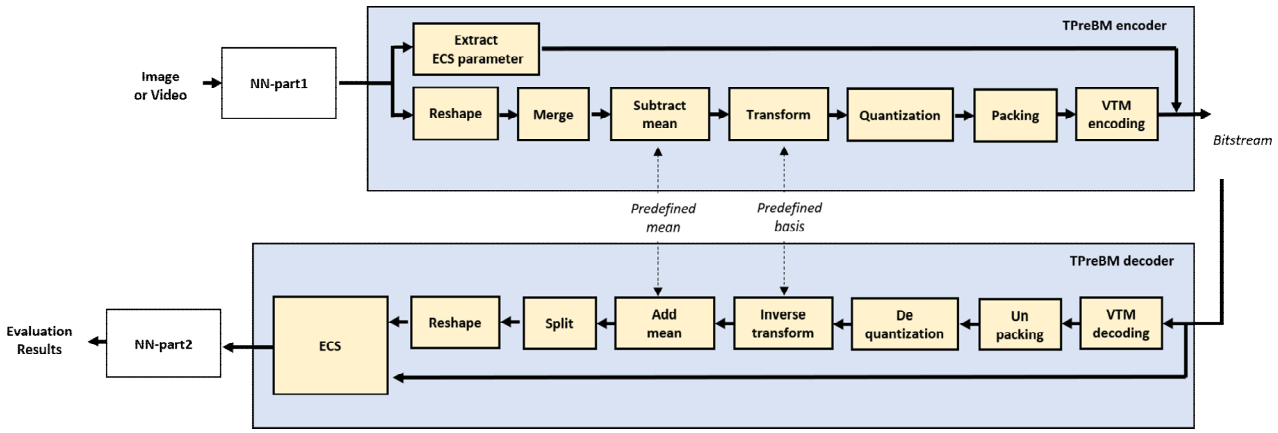


그림 1. PreBM PCA를 이용한 기계 학습 모델의 중간 특징 데이터 부호화기와 복호화기의 구조도[15]
Fig. 1. Overall pipeline of PreBM PCA encoder and decoder[15]

구한다. i는 다계층 특징맵의 계층을 지시하고, ori과 rec는 원본과 복원값을 각각 지시한다. 즉 인코더에서는 원본 특징의 각 계층의 최대값(max_{Li}^{ori})과 평균값(μ_{Li}^{ori})을 구하여 전송한다. 디코더에서 식 (1)을 이용하여 스케일링 값(s_i)을 계산해 복원된 특징맵을 확대해 준다.

$$s_i = \frac{\sigma_{Li}^{ori}}{\sigma_{Li}^{rec}} \cong \frac{max_{Li}^{ori} - \mu_{Li}^{ori}}{max_{Li}^{rec} - \mu_{Li}^{rec}} \quad (1)$$

PreBM PCA 기술은 다계층 특징맵을 변환하여 단일 계층으로 만든 후 전체에 PCA를 적용하는 점이 특징이다. 이

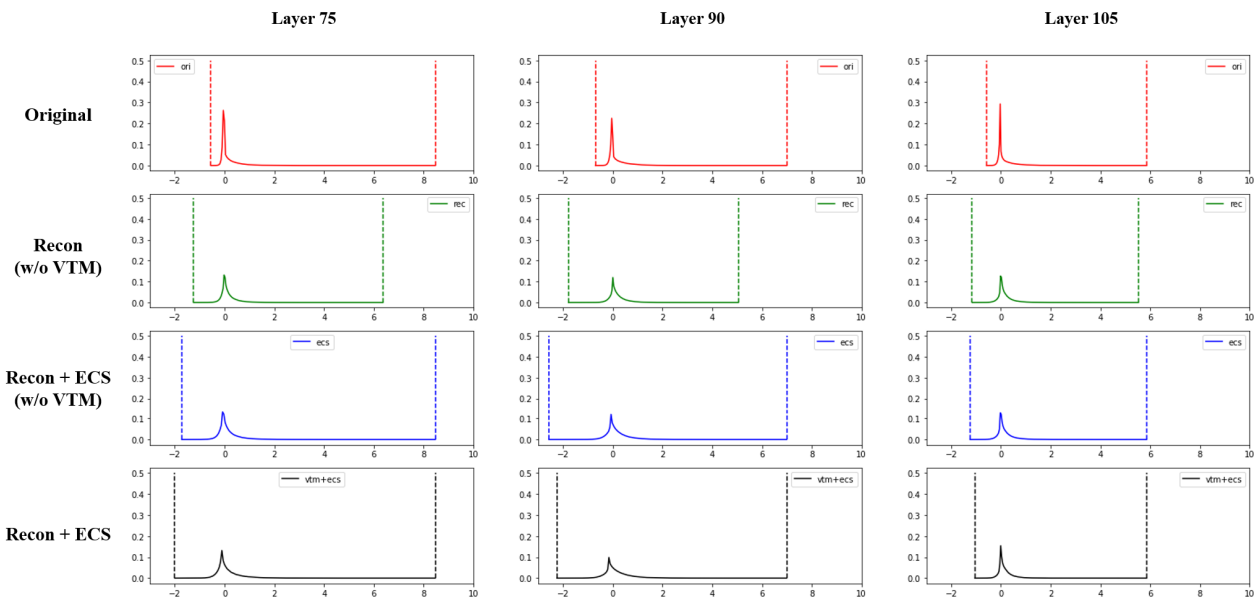


그림 2. PreBM PCA 과정을 통한 FCM 중간 분할 특징맵의 분포[7]
Fig. 2. The feature distribution change by PreBM PCA transformation[7]

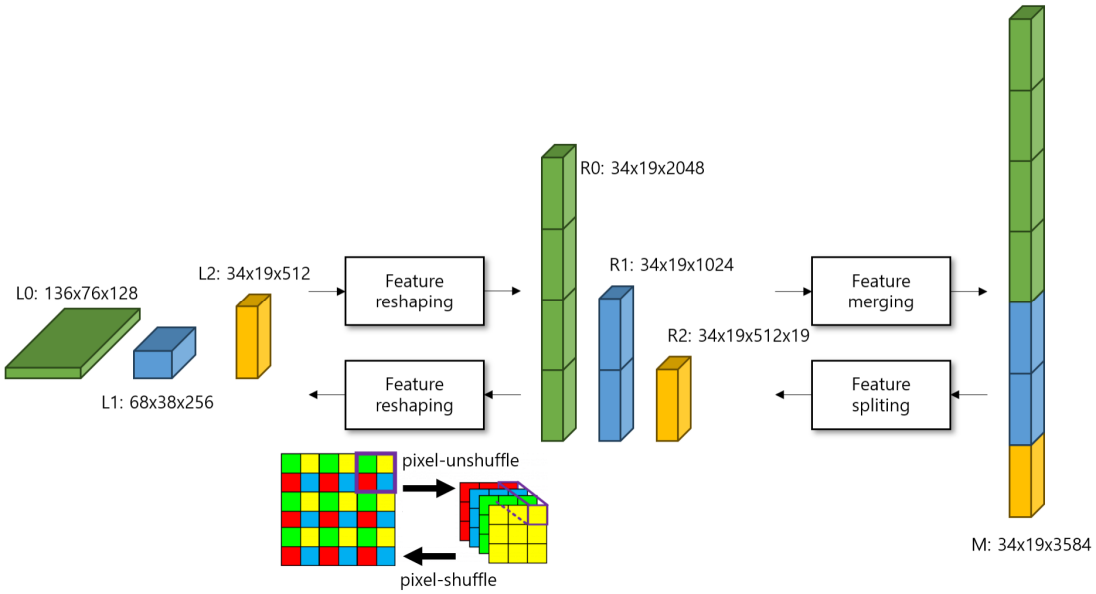


그림 3. PreBM PCA기술 내에서 다계층 특징맵을 단일 특징맵으로 변형하는 과정[10]
Fig. 3. Pixel-shuffling and merging multi-layer features[10]

는 채널 및 계층간 중복 정보를 효율적으로 압축하고자 함이다. 다계층 특징맵을 변환하는 자세한 과정은 그림 3을 통해 알 수 있다[18].

1.2.3 디코딩

디코딩 과정에서는 다음과 같은 순서로 원본 데이터를 복원한다:

- 헤더 분석 및 VTM 디코딩: 압축된 비디오 프레임을 분석하여 원본 특징 데이터를 복원한다.
- 패킹 해제 및 역양자화: 패킹된 데이터를 원래 형식으로 변환하고, 양자화된 데이터를 역양자화하여 연산 가능한 형태로 복구한다.
- PCA 역변환 수행: 저장된 주성분 벡터를 활용하여 원본 특징맵을 복원한다.
- 특징맵 분할 및 픽셀 역재배열(pixel-unshuffle): 변환된 단일 층 특징맵을 원래 다층 구조로 변환한다.
- ECS 적용: 인코딩 과정에서 발생한 신호 크기 변화를 보정하여 최종적인 특징맵을 출력한다.

1.3 성능

PreBM 기술은 FCM CfP 응답 기술이나 이후 CE2에서 계속 실험을 진행하다가 제149차 MPEG 회의에서 중단하

표 1. FCTM4.0에 통합한 PreBM PCA의 성능
Table 1. The summary of PreBM PCA performanc compared to FCTM4.0 and Remote inference anchor

Task / Dataset		FCTM4.0			Remote inference
		BD-rate (%)	EncT (%)	DecT (%)	BD-rate (%)
Instance Segmentation	OpenImageV6	237.12	350.27	27.93	-83.18
Object Tracking	TVD	N/A	316.89	28.95	N/A
	HiEVE (1080p)	26.80	408.80	33.22	-17.17
	HiEVE (720p)	329.64	359.02	24.88	-13.26
Object Detection	SFU (ClassAB)	339.78	320.48	24.58	11.61
	SFU (ClassAB)	434.20	290.99	16.10	-36.89
	SFU (ClassAB)	602.70	213.46	21.43	-18.23
	OpenImageV6	1000.51	248.60	19.04	240.51
Overall		N/A	341.075%	25.943%	N/A

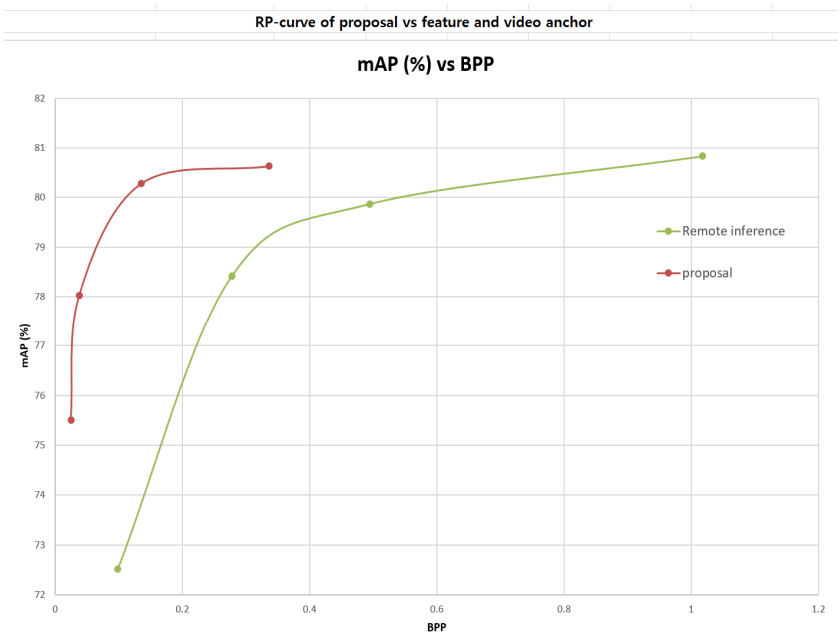


그림 4. 인스턴스 분할 OpenImageV6 데이터셋에서의 Remote anchor 대비 PreBM PCA 성능[18]
Fig. 4. Rate-Performance (RP) curve comparing the instance segmentation performance of the PreBM PCA approach with the Remote Anchor baseline on the OpenImages V6 dataset[18]

기로 결정되었다[19,20]. 표 1의 성능은 PreBM PCA 기술을 FCTM4.0에 통합하여 FCTM4.0 앵커(Anchor)와 비교한 실험 결과이다[18].

PreBM PCA 기술은 디코더 속도에 대해서 FCTM4.0 대비 큰 복잡도 감축 결과를 보이며, 부호화 성능에 대해서는 2개의 데이터셋을 제외한 모든 테스트 데이터셋에서 VVC를 이용한 부호화 결과(Remote Inference Anchor) 대비 BD-rate 개선 결과를 보인다. 그림 4는 인스터스 분할 임무에 대해 VVC 부호화 결과(Remote Inference Anchor) 대비 OpenImageV6 데이터셋에서의 성능(RP) 그래프이다.

2. Online PCA 기술

2.1 개요

Online PCA 기술은 FCM 기술 제안 요청에 대한 응답 기술로 처음 제안되었다[11,21]. 본 기술은 입력 데이터에 적응적으로 PCA 주성분을 계산하여 피쳐 데이터를 부호화하는 기술로 주기적으로 PCA 주성분을 업데이트하기에 실시간 비디오 데이터가 지속적으로 유입되는 환경에 적합한 부호화 방식이다[22,23,24,25]. 또한 네트워크 및 분할 지점에 대한 종속성이 없이 피쳐 부호화가 가능하여 FCM의 중요 활용 사례인 연산 오프로딩(computational offloading) 구현이 가능하다[26].

Online PCA는 다음과 같은 특징을 가진다:

- 범용성: 부호화 대상인 피쳐의 네트워크 구조나 분할 지점에 상관없이 다양한 모델에 적용 가능하다.
- 점진적 업데이트: 새로운 데이터가 들어올 때마다 주기적으로 주성분을 갱신하여, 매 프레임마다 주성분을 계산하지 않아도 된다.
- 저장 공간 절약: 대규모 데이터에서 필요한 차원만 유지하여 스토리지 사용을 최소화한다.
- 빠른 연산 속도: 전체 데이터를 한 번에 처리하는 것이 아니라 주기적으로 업데이트하기 때문에 연산량이 줄어든다.

2.2 요소 기술

Online PCA 기술의 인코딩은 아래와 같은 과정은 거쳐 이루어지고 그림 5는 이를 도식화하였다[25].

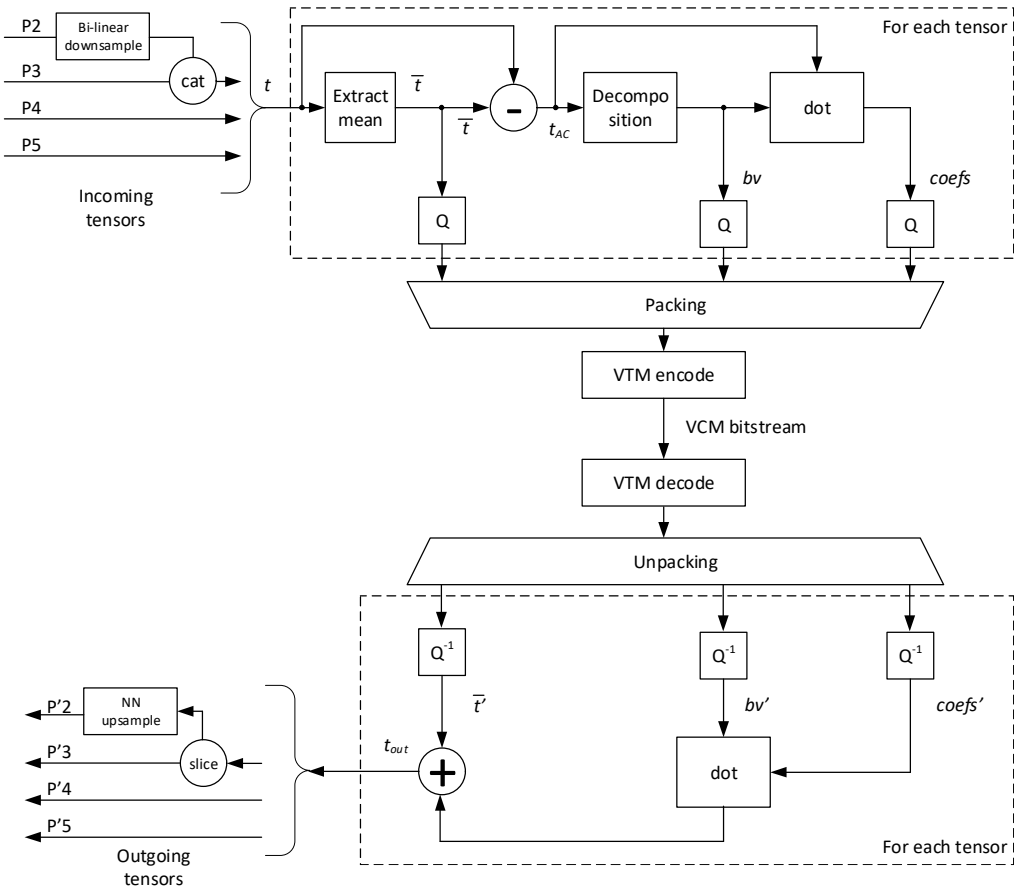


그림 5. Online PCA 인코딩 및 디코딩 과정[25]
Fig. 5. Online PCA encoding and decoding[25]

- PCA 주성분 분석: 특징맵 계층 데이터의 채널별 주성분 추출을 프레임마다 계산하는 것이 아니라 특정 간격의 프레임(예시: 인트라 주기)에서만 수행한다.
- PCA 변환 적용: 추출한 주성분을 이용하여 채널별 PCA 변환을 적용하여 상응하는 계수를 저장한다.
- 양자화 및 패킹: 변환 계수 데이터를 10비트 정밀도로 양자화하고, VVC 압축을 위해 정렬된 형식으로 패킹한다.
- VVC 인코딩: 압축된 특징 데이터를 VTM을 활용하여 비디오 프레임 형식으로 인코딩한다. VVC에서 인코딩하는 해상도를 시퀀스 단위로 결정하는 방법을 채용하기에 기저 벡터(basis vectors)의 추출 빈도와 관계없이, 매 프레임 기저 벡터 및 채널 평균값은 별도의 서브픽처(subpicture)로 인코딩된다.

그림 4는 Online PCA를 제안한 기고서에서 가져온 인코더 및 디코더의 구조이다[25]. Online PCA는 특징맵 데이터가 큰 경우 두 개의 계층을 하나의 계층으로 결합함으로써 PCA 수행 전 데이터를 축소한다. 객체 검출 및 분할 네트워크인 Faster R-CNN 모델의 분할 지점(P Layer)의 경우 P2 레이어는 이중 선형 보간법(bilinear downsampling)을 사용하여 다운샘플링된 후, 채널 차원에서 P3 레이어와 결합(concatenation)된다. 그 후 세 개의 계층(P2/P3, P4, P5)이 순서대로 인코딩된다. 디코더에서는 최근접 이웃 보간법(nearest-neighbour interpolation)을 사용하여 복원된 P′2/P′3 결합텐서로부터 업샘플링하여 P′2를 복원한다. 객체 추적 네트워크인 JDE 모델의 분할 지점(split points)(DN53 및 ALT1)에서는 3개의 계층으로 구성되기에 계층 간 결합 과정을 수행하지는 않는다.

표 2. Online PCA에서 사용하고 있는 PCA 주성분 사용 개수[25]
Table 2. Number of PCA components used in Online PCA[25]

Task	Dataset	Sequence	Components
Instance Segmentation	OpenImage	OpenImageV6	60, 30, 30
Object Detection	OpenImage	OpenImageV6	60, 30, 30
	SFU	Traffic_2560x1600_30_val	60, 30, 30
		Kimono_1920x1080_24_val	70, 35, 35
		ParkScene_1920x1080_24_val	60, 30, 30
		Cactus_1920x1080_50_val	50, 25, 25
		BasketballDrive_1920x1080_50_val	80, 40, 40
		BQTerrace_1920x1080_60_val	80, 40, 40
		BasketballDrill_832x480_50_val	60, 30, 30
		BQMall_832x480_60_val	60, 30, 30
		PartyScene_832x480_50_val	120, 60, 60
		RaceHorses_832x480_30_val	80, 40, 40
		BasketballPass_416x240_50_val	80, 40, 40
		BQSquare_416x240_60_val	80, 40, 40
		BlowingBubbles_416x240_50_val	90, 45, 45
		RaceHorses_416x240_30_val	60, 30, 30
Object Tracking	TVD	TVD01	75, 175, 50
		TVD02	75, 125, 50
		TVD03	25, 50, 25
	HIEVE	HIEVE2	55, 55, 30
		HIEVE13	50, 75, 50
		HIEVE16	35, 55, 35
		HIEVE17	65, 90, 90
		HIEVE18	55, 80, 80

Online PCA는 채널별(Channel-wise)로 PCA 변환을 수행한다. 이는 다수의 채널로 구성되면 채널 내 채널 간 중복성을 제거하기 위한 접근 방식이다. 또한 정보량을 줄이면서도 성능을 최대한 유지할 수 있는 적절한 PCA 주성분 개수는 사전 실험을 통해 표 2와 같이 설정된다.

PCA 변환 정보는 그림 6에서 보듯이 VTM으로 부호화한 후 디코더에 전달된다. 기저 벡터(basis vectors)의 생성 빈도와 관계없이, 매 프레임 기저 벡터 및 채널 평균은 별도의 서브픽처(subpicture)로 VVC 인코딩되며, 여러 프레임에서 사용되는 기저 벡터는 고품질을 유지하기 위해 프레임 QP 대비 QP 오프셋(Offset) -18을 적용하여 인코딩된다. 그림 6은 SFU 데이터셋의 RaceHorses 시퀀스 첫 번째 프레임에 대한 VVC 인코더의 입력으로 구성된 데이터의 서브픽처 배치를 도식화한 것이다. 그림 5에서 볼 수 있듯이, Online PCA 부호화 데이터는 세 개의 영역(서브픽처, sub-pictures)으로 인코딩된다.

- 채널별 평균값(channel-wise mean values)이 포함되며, 이는 계층별로 배치한다.

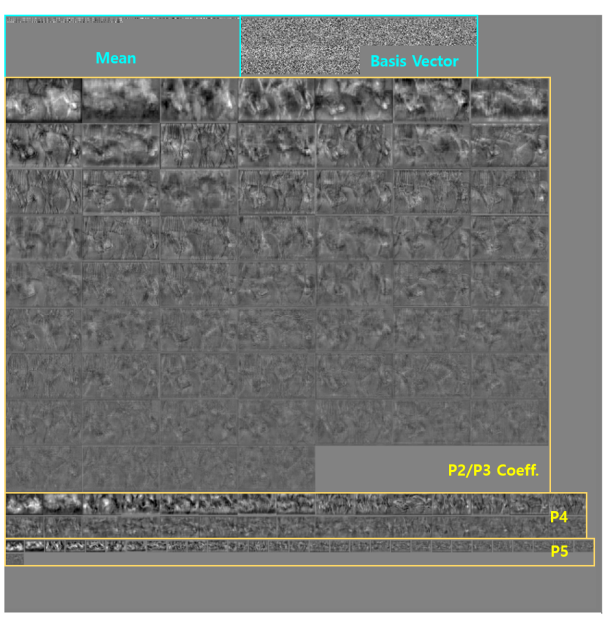


그림 6. Online PCA 기술에서 피쳐 부호화 정보를 VVC로 부호화할 때 서브픽처의 구성[25]
Fig. 6. VVC sub-pictures for Online PCA[25]

- 채널별 기저 벡터(channel-wise basis vectors)가 포함되며, 이 또한 계층별로 배치한다.
- 각 계층의 PCA 계수 세트(coefficients)가 배치된다.

2.3 성능

Online PCA 기술의 성능 실험은 여러 데이터셋(Open-Images, SFU, HiEve 등)을 이용하여 실험되었으며, 제148차 MPEG 회의에서 보고된 최신 성능 결과는 다음과 같다[25]. Online PCA 기술은 FCTM의 선택 가능한 툴(Tools) 중 Refinement on fused domain 옵션과 함께 사용할 수 없기에 비활성화되었다. 이는 현재 채택된 FCM에서 VVC 복원 후의 특성과 Online PCA의 VVC 복원 후 특성이 다르기 때문에 옵션이 적용 가능하지 않다.

아래 표 3의 성능은 Online PCA 기술을 FCTM4.0에 통합하여 FCTM4.0 앵커(Anchor) 및 영상을 VVC로 압축 전송하는 Remote inference 앵커와 비교한 실험 결과이다[20]. Online PCA 기술의 성능은 특정 네트워크 및 특정 분할 지점의 피쳐맵을 대상으로 사전 학습한 신경망 기반 피쳐 축소 방식의 FCTMv4.0 대비 낮은 BD-rate을 보인다. 그러나 피쳐 대신 영상을 VVC 압축 전송받아 원격에서 기계 성능을 측정하는 Remote Inference Anchor 성능 대비 일부 객체 검출(Object Detection) 및 객체 추적(Object Tracking) 임무에서 BD-rate 절감 효과를 보인다. 연산 속도 측면에서도 일부 데이터셋에 대해 낮은 복잡도 결과를 보인다[25].

III. 결 론

본 논문에서는 기계 학습을 위한 비디오 처리 환경[26,27]에서 효율적인 특징 정보 부호화를 위한 Feature Coding for Machines(FCM) 기술 중, 비신경망 기반 피쳐 변환 기법(Non-NN-based Feature Transform)에 대해 기술하였다. 특히 PCA 기술을 기반으로 한 두 기술인 PreBM과 Online PCA 기술을 소개하고 분석하였다. PreBM PCA 기술은 사전에 샘플 데이터로부터 계산해놓은 주성분 벡터와 평균을 활용하여 연산 부담을 줄이면서 높은 성능을 유지하는 중간특징정보 부호화 기술이다. 반면 Online PCA 기술은 다양한 네트워크 모델 및 분할 지점의 특징정보를 부호화할 수 있는 장점을 지닌 기술이다. 주기적으로 PCA 주성분 벡터를 업데이트함으로써 실시간 비디오 환경에서 객체 검출, 객체 추적, 인스턴스 세그멘테이션 등 다양한 영상 처리 수행에 적용 가능하다. PCA는 차원 축소를 통해 데이터의 중요 정보를 보존하면서 압축률을 높이고, 연산 자원을 절감함으로써 단말과 서버 간의 효율적인 분산 처리 및 통신을 지원하는 데 적합하다. 이러한 특성은 클라우드 기반 연산 환경과 엣지 컴퓨팅이 결합되는 FCM 시나리오에서 특히 유용하기에 앞으로 다양한 PCA 기반 응용 기술 및 최적화 전략이 제안될 것으로 기대된다. 본 논문에서 살펴본 비신경망 기반 기법은 네트워크 구조나 분할 지점에 상관없이 다양한 모델에 적용 가능하며, 연산 자원이 제한된 환경에서 유리한 장점을 가지고 있어 FCM의 최종적인 지향점에

표 3. FCTM4.0에 통합한 Online PCA의 성능[25]
Table 3. The summary of Online PCA performanc compared to FCTM4.0 and Remote inference anchor[25]

Task / Dataset		FCTM4.0			Remote inference
		BD-rate (%)	EncT (%)	DecT (%)	BD-rate (%)
Instance Segmentation	OpenImageV6	2459.31	N/A	N/A	48.13
Object Tracking	TVD	1437.60	N/A	N/A	3.50
	HiEVE (1080p)	585.60	463.16	86.32	298.95
	HiEVE (720p)	192.40	361.19	63.18	-38.23
Object Detection	SFU (ClassAB)	179.92	539.47	53.91	-28.71
	SFU (ClassAB)	N/A	1640.49	181.33	N/A
	SFU (ClassAB)	1642.84	944.17	180.58	-6.46
	OpenImageV6	828.89	1760.89	384.06	49.77
Overall		N/A	N/A	N/A	N/A

유효한 접근 방식 중 하나로 향후 기술적 한계를 극복하기 위한 후속 연구가 진행되길 기대한다.

참 고 문 헌 (References)

[1] ISO/IEC JTC 1/SC 29/WG 2, N00282, “Call for Proposals on Feature Compression for Video Coding for Machines”, Apr. 2023.

[2] ISO/IEC JTC 1/SC 29/WG 2, N00318, “Call for Proposals responses report for Feature Compression for Video Coding for Machines”, Oct. 2023.

[3] ISO/IEC JTC 1/SC 29/WG 4, N00625, “Common test and training conditions for FCM”, Jan. 2025.

[4] ISO/IEC JTC 1/SC 29/WG 4, N00628, “Preliminary WD of feature coding for machines”, Jan. 2025.

[5] ISO/IEC JTC 1/SC 29/WG 4, N00627, “Software manual of FCTM”, Jan. 2025.

[6] ISO/IEC JTC 1/SC 29/WG 4, N00626, “Algorithm description of FCTM”, Jan. 2025.

[7] W. Zhang, J. Li, and C. Wang, “Adaptive convolutional neural network-based principal component analysis for multi-modal facial recognition,” Advances in Mechanical Engineering, Vol. 16, No. 2, 2024.
doi: https://doi.org/10.1177/16878132251325420

[8] H. Liu and M. Zhao, “Moving Object Detection Method Based on the Fusion of Online Robust Principal Component Analysis and Image Registration,” Neural Processing Letters, Vol. 56, No. 55, 2024.
doi: https://doi.org/10.1007/s11063-024-11463-w

[9] A. Razzaque and A. Badholia, “PCA based feature extraction and MPSO based feature selection for gene expression microarray medical data classification,” Measurement: Sensors, Vol. 31, 2024.
doi: https://doi.org/10.1016/j.measen.2023.100945

[10] Y. Kim, et al., “[FCVCM] Response to CfP on FCVCM: A transform-based feature coding method with energy compensation scaling,” ISO/IEC JTC 1/SC 29/WG 2, m65704, Oct. 2023.

[11] C. Rosewarne and R. Nguyen, “[FCVCM] Response to FCVCM Call for Proposals from Canon,” ISO/IEC JTC 1/SC 29/WG 2, m65218, Oct. 2023.

[12] ISO/IEC JTC 1/SC 29/WG 4, N00430, “FCM CE 2”, Oct. 2023.

[13] M. Lee, et al., “[VCM Track 1] Response to CfE: A transformation-based feature map compression method,” ISO/IEC JTC 1/SC 29/WG 2, m60788, Oct. 2022.

[14] Y. Kim, et al., “[FCM] CE2.3.5 PreBM PCA,” ISO/IEC JTC 1/SC 29/WG 4, m66461, Jan. 2024.

[15] Y. Kim, et al., “[FCM] CE2.3.6 PreBM PCA +ECS,” ISO/IEC JTC 1/SC 29/WG 4, m66464, Jan. 2024.

[16] Y. Kim, et al., “[FCM] CE2.2.1 FCTM + ECS,” ISO/IEC JTC 1/SC 29/WG 4, m67618, Apr. 2024.

[17] Y. Kim, et al., “[FCM] CE4.1: ECS,” ISO/IEC JTC 1/SC 29/WG 4, m68758, Jul. 2024.

[18] Y. Kim, et al., “[FCM] CE2.4: PreBM + ECS,” ISO/IEC JTC 1/SC 29/WG 4, m70046, Oct. 2024.

[19] F. Racapé, C. Rosewarne, and Y. Kim, “[FCM] CE2: Summary report of the core experiment on Non-NN-based feature transform,” ISO/IEC JTC 1/SC 29/WG 4, m69974, Jan. 2025.

[20] C.Rosewarne and Y. Zhang “BoG on Feature Coding for Machines,” ISO/IEC JTC 1/SC 29/WG 4, m71631, Jan. 2025.

[21] C. Rosewarne and R. Nguyen, “[VCM Track 1] Response to CfE on Video Coding for Machine from Canon,” ISO/IEC JTC 1/SC 29/WG 2, m60788, Oct. 2022.

[22] C. Rosewarne and R. Nguyen, “[FCM] CE2.3.1: Online PCA,” ISO/IEC JTC 1/SC 29/WG 4, m66340, Jan. 2024.

[23] C. Rosewarne and R. Nguyen, “[FCM] CE2.1.1 and CE2.1.2: Online PCA,” ISO/IEC JTC 1/SC 29/WG 4, m67469, Apr. 2024.

[24] C. Rosewarne and R. Nguyen, “[FCM] CE2.1: Online PCA,” ISO/IEC JTC 1/SC 29/WG 4, m68703, Jul. 2024.

[25] C. Rosewarne and R. Nguyen, “[FCM] CE2.1: Online PCA,” ISO/IEC JTC 1/SC 29/WG 4, m69974, Oct. 2024.

[26] ISO/IEC JTC 1/SC 29/WG 2, N00190, “Use cases and requirements for Video Coding for Machines”, Apr. 2022.

[27] Ayoung Kim et al., “Suboptimal video coding for machines method based on selective activation of in-loop filter,” ETRI Journal, Vol. 46, No. 3, pp.538-549, June 2024.
doi: https://doi.org/10.4218/etrij.2023-0085

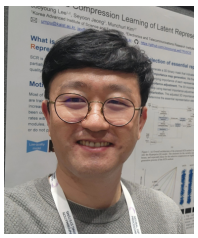
저 자 소 개

김 연 희

- 2000년 : 아주대학교 정보및컴퓨터공학과 학사
- 2002년 : 아주대학교 정보및컴퓨터공학과 석사
- 2009년 : George Mason University Computer Science 박사
- 2009년 ~ 현재 : 한국전자통신연구원 미디어부호화연구실 책임연구원
- ORCID : http://orcid.org/0000-0003-0658-6762
- 주관심분야 : 영상처리/압축, 컴퓨터 비전, FCM, MPEG

정 세 윤

- 1995년 : 인하대학교 전자공학과 학사
- 1997년 : 인하대학교 전자공학과 석사
- 2014년 : KAIST 전기및전자공학과 박사
- 1996년 ~ 현재 : 한국전자통신연구원 미디어부호화연구실 책임연구원
- ORCID : https://orcid.org/0000-0002-1675-4814
- 주관심분야 : 실감 방송, 비디오 코딩, 컴퓨터 비전



이 주 영

- 2003년 : 아주대학교 미디어학부 학사
- 2006년 : KAIST 전산학부 석사
- 2024년 : KAIST 전기및전자공학부 박사
- 2006년 ~ 현재 : 한국전자통신연구원 미디어부호화연구실 책임연구원
- ORCID : https://orcid.org/0000-0003-0753-0699
- 주관심분야 : 인공지능, 컴퓨터 비전, 생성 모델, 이미지/비디오 압축