졸업논문 소개

FCM을 위한 선택적 학습전략을 적용한 다중스케일 특징 압축 기법 연구

한규웅 / 한국항공대학교 Media Communication Lab.

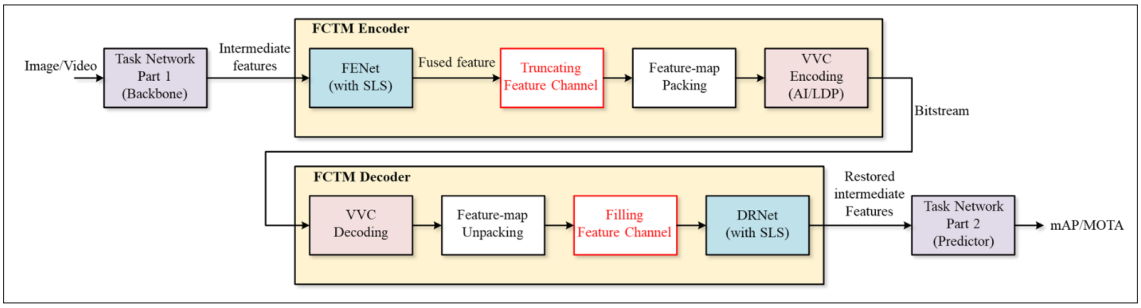
최근 자율주행, 스마트 시티, 감시 비디오, IoT(Internet of Things) 등 다양한 분야에서 신경망 기반 머신비전(machine vision) 응용이 확산됨에 따라, 기계가 소비하고 처리하는 비디오 데이터가 급격히 증가하고 있다. 또한, 높은 임무 수행 성능을 달성하기 위해 신경망 모델과 추론의 복잡도도 증가하고 있다. 이러한 문제를 해결하기 위해 MPEG(Moving Picture Experts Group)에서는 기계 소비를 위한 고효율 비디오 압축 표준인 VCM(Video Coding for Machines)과 FCM(Feature Coding for Machines)을 개발하고 있다. VCM은 입력 비디오 자체를 비전 임무를 고려하여 압축 전송하는 반면, FCM은 분석 네트워크를 통하여 입력 비디오로부터 추출된 특징(feature)을 압축 전송하고 서버에서 비전 임무를 수행하는 환경을 가정한다.

현재, FCM은 기술제안요청서(CfP: Call for Proposals)의 응답기술 평가를 바탕으로 시험모델인 FCTM(Feature Compression Test Model)을 개발하면서 기술검증 및 표준 기술 개발을 위한 핵심실험(CE: Core Experiments)을 진행하고 있다.

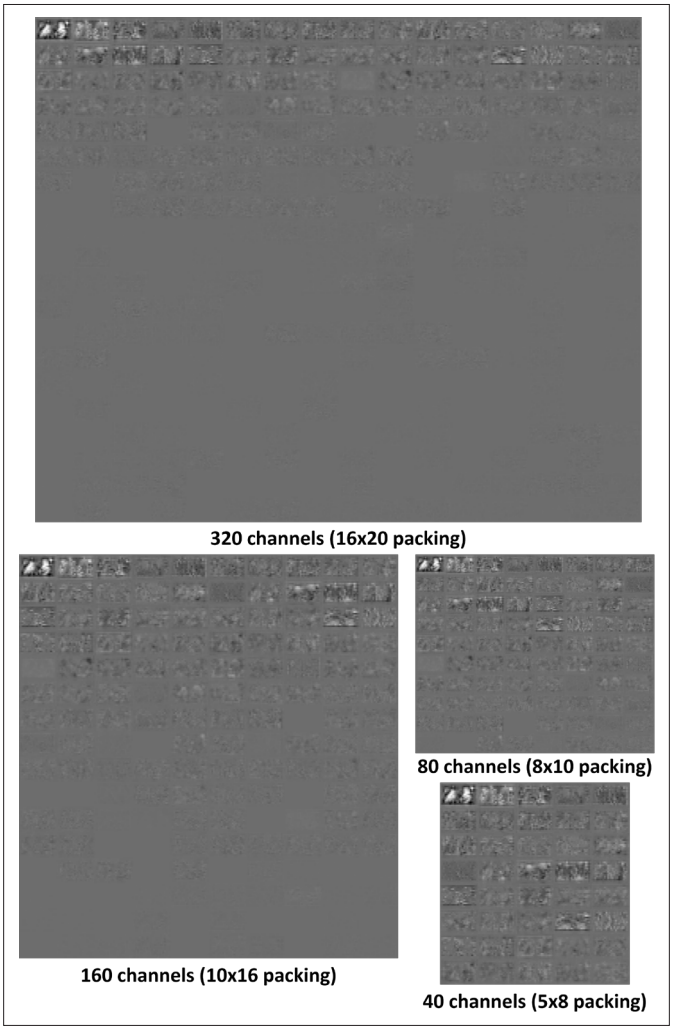
본 연구에서는 선택도를 조절하여 개선된 선택적 학습전략(SLS: Selective Learning Strategy)과 QP 적응적 채널절삭 기법(QACT: QP-Adaptive Channel Truncation)을 적용한 FCTM의 압축성능 개선기법을 제안한다. 〈그림 1〉은 본 제안기법을 적용한 FCTM의 프레임워크를 나타내며, 〈그림 2〉는 SLS와 QACT 기법을 적용한 FCTM의 특징맵을 나타낸 것이다. 제안기법은 SLS를 통해 특징맵을 특징 채널별 중요도 순서대로 정렬하여 특징맵의 공간적 상관성을 증가시키고, 주어진 비트율(QP)에 따라 정렬된 특징맵을 순차적으로 절삭하여 해상도를 조정함으로써 효율적인 특징 압축을 가능하게 한다.

제안기법을 FCTM에 구현하였으며, FCM CTTC(Common Training and Test Conditions)에 따라 BD(Bjøntegaard delta)-rate 성능을 확인하였다. 〈표 1〉은 제안기법의 성능을 나타낸다. FCTM v4.0 대비 모든 머신비전 임무에서 BD-rate 절감의 성능향상을 보였으며, 전체적으로 17.3%의 비트율 절감 성능향상을 보였다.

또한, 본 연구에서는 FCM CE1에서 연구되고 있는 advMSFC(advanced Multi-Scale Feature Compression) 복잡도 및 성능개선 기법으로 구조적 복잡도를 개선하고 MSFR(Multi-Scale Feature Restoration)의 특



<그림 1> SLS와 QACT를 적용한 FCTM 프레임워크



<그림 2> SLS와 QACT를 적용한 특징맵

<표 1> SLS와 QACT를 적용한 FCTM의 성능

QACT with SLS vs. FCTM-v4.0				
Task	Test dataset	BD-rate [%]	EncR [%]	DecR [%]
Instance Segmentation	OpenImageV6	-9.13%	89.983%	102.813%
Object Detection	OpenImageV6	-24.41%	88.107%	111.830%
	SFU (Class A/B)	-4.31%	60.492%	105.858%
	SFU (Class C)	-18.17%	65.298%	108.563%
	SFU (Class D)	-27.48%	66.777%	105.172%
Object Tracking	TVD (OVERALL)	-35.31%	62.234%	99.571%
	HIEVE (1080p)	-1.58%	78.617%	100.196%
	HIEVE (720p)	-18.04%	75.342%	100.004%
	OVERALL	-17.30%	72.149%	105.634%

징 복원성능을 향상시킬 수 있는 HCC(Half-Channel Concatenation) 기법과 다중스케일 특징 간의 최소자승 오차(MSE: Mean Square Error)를 이용한 학습 손실함수 개선기법을 제안한다. FCM CTTC에 따른 실험결과 FCTM v3.2 대비 객체 검출 작업에서 제안기법의 성능저하가 있지만, 객체 분할과 객체 추적 작업에 대해서는 최대 86.38%의 비트율 절감 성능향상을 보였다.

본 연구에서는 선택적 학습전략(SLS)과 QP 적응적 채널절삭(QACT) 기법을 적용하여 FCM의 참조 소프트웨어인 FCTM의 성능개선 기법을 제시하였다. 또한, 현재 FCM의 핵심실험에서 연구되고 있는 advMSFC의 성능개선을 위한 개선기법으로 다중스케일 특징 간의 MSE를 이용한 손실함수 개선, HCC를 이용한 MSFR 모듈의 구조적 개선기법을 제안하였다. 실험결과 제안한 방법들은 각각 FCTM v4.0, FCTM v3.2 대비 17.3%와 86.38%의 BD-rate 절감의 성능개선을 확인하였다. 본 연구에서 제안한 기법들은 효율적인 특징 압축을 가능하게 하는 다중스케일 특징 압축 기법으로 FCM 표준의 유력한 후보기술이 될 수 있을 것으로 기대된다.

한 규 웅

- 2023년 2월 : 한국항공대학교 항공전자정보공학부 학사
- 2025년 2월 : 한국항공대학교 항공전자정보공학과 석사
- 주관심분야 : 비디오 부호화, VCM, FCM, 딥러닝