

일반논문 (Regular Paper)

방송공학회논문지 제31권 제2호, 2026년 3월 (JBE Vol.31, No.2, March 2026)

<https://doi.org/10.5909/JBE.2026.31.2.302>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

FCM을 위한 통계적 특성 기반 비선형 변환 및 스케일링 방법

이 성 은^{a)}, 서 예 원^{a)}, 한 희 지^{a)}, 정 순 흥^{b)}, 추 현 곤^{b)}, 최 해 철^{a)†}

Statistic-Guided Nonlinear Transform and Scaling for Feature Coding for Machines

Seongeun Lee^{a)}, Yewon Seo^{a)}, Heeji Han^{a)}, Soon-heung Jung^{b)}, Hyon-Gon Choo^{b)}, and Haechul Choi^{a)†}

요 약

신경망 기반 시각 인식 모델은 높은 성능을 제공하지만 연산 복잡도로 인해 에지 디바이스에서의 단독 추론이 제한되며, 이에 따라 에지 디바이스와 서버가 나누어 추론하는 분할 추론 구조가 사용되고 있다. 분할 추론 환경에서 중간 특징의 효율적인 전송을 위해 국제 표준화 단체 Moving Picture Experts Group은 Feature Coding for Machines (FCM) 표준화를 진행하고 있다. FCM은 균등 양자화 적용으로 인해 특징 분포의 통계적 특성을 충분히 반영하지 못하며, 내부 코덱의 부호화 과정에 의해 Quantization Parameter (QP)가 증가할수록 특징 분포 왜곡이 크게 발생하는 한계가 있다. 본 논문에서는 이러한 한계를 개선하기 위해 비균등 양자화를 모사하는 통계적 특성 기반 비선형 변환과 QP 기반 스케일링 방법을 제안하여 부호화 효율을 향상시키고 특징 분포 왜곡을 개선한다. 실험 결과, 제안 방법은 FCM의 테스트 모델인 Feature Compression Test Model v8.1 대비 평균 3.14%의 BD-rate 향상을 달성하였다.

Abstract

Deep neural network -based visual recognition models achieve high performance, but standalone inference on edge devices is limited due to high computational complexity. To address this issue, split inference architectures are widely used, where intermediate features are transmitted between edge devices and servers. For efficient transmission of intermediate features, the Moving Picture Experts Group (MPEG) is standardizing Feature Coding for Machines (FCM). However, FCM relies on uniform quantization, which does not sufficiently reflect the statistical characteristics of feature distributions, and feature distortion becomes more severe as the quantization parameter (QP) increases in the inner codec. This paper addresses these limitations by proposing a statistical feature - guided nonlinear transform that emulates non-uniform quantization, along with a QP-based scaling method. The proposed approach improves coding efficiency while mitigating feature distribution distortion. Experimental results show that the proposed method achieves an average BD-rate reduction of 3.14% compared to Feature Compression Test Model (FCTM) v8.1.

Keyword : Video Coding, Feature Coding, Computer vision, Deep Learning

Copyright © 2026 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. 서론

딥러닝 기반 시각 인식 모델은 자율주행, 감시, 스마트 시티와 같은 실시간 응용 환경에서 높은 임무 성능을 달성하고 있다. 그러나 이러한 성능은 대규모 신경망 구조와 높은 연산 복잡도를 전제로 하며, 연산 자원이 제한된 에지 디바이스에서 전체 추론을 수행하기에는 한계가 있다.

이러한 문제를 완화하기 위해, 신경망의 일부 연산을 에지 디바이스에서 수행하고 나머지를 서버에서 처리하는 분할 추론 구조^{[1][2]}가 제안되어 왔다. 분할 추론은 에지 디바이스의 연산 부담을 줄일 수 있으나, 신경망의 중간 계층에서 생성된 특징 데이터를 서버로 전송해야 한다. 특히 신경망의 중간 특징은 데이터량이 크기 때문에, 이를 효율적으로 압축할 수 있는 기술이 필요하다.

이러한 배경에서 중간 특징 데이터를 압축하여 전송 효율을 높여려는 연구가 활발히 이루어지고 있다. 국제 표준화 단체 Moving Picture Experts Group (MPEG)은 이를 표준화하기 위해 Feature Coding for Machines (FCM) 프로젝트를 진행하고 있다. FCM은 영상 신호를 부호화하는 기존 비디오 코덱과 달리 신경망 내부에서 생성되는 특징 데이터를 기계가 효율적으로 소비할 수 있도록 압축하는 것을 목표로 한다. 이를 위해 MPEG FCM은 Feature Compression Test Model (FCTM)^[3]을 시험 모델로 정의하고, 다양한 부호화 도구에 대한 검증과 개선을 진행하고 있다.

그러나 FCM은 특징값을 일정한 간격으로 정수화하는 균등 양자화^[4]를 기반으로 설계되어 있어, 실제 신경망 특징이 가지는 비균등한 통계적 분포를 충분히 반영하지 못

한다. 이로 인해 엔트로피 부호화 효율이 제한되며, 중요한 특징 구간에서 양자화 손실이 발생할 수 있다. 또한 FCTM의 내부 코덱으로 사용되는 Versatile Video Coding (VVC)^[5]은 양자화 계수(Quantization Parameter, QP)에 따라 양자화 단계 크기가 변화하므로, QP가 증가할수록 부호화된 특징의 동적 범위(Dynamic Range)가 축소되고 특징의 분포 왜곡이 커지게 된다. 이러한 분포 왜곡은 후속 특징 복원 단계에서 임무 성능 저하로 이어질 수 있다.

본 논문에서는 이러한 문제를 해결하며 부호화 효율을 높이기 위해, 첫째, 신경망 특징의 통계적 특성을 반영하는 비선형 변환을 균등 양자화기 전에 적용한다. 이 처리는 비균등 양자화를 모사할 수 있기 때문에 특징값에 따른 임무 성능의 영향별로 양자화 손실을 비균등하게 조정할 수 있다. 또한 비선형 변환은 특징값이 자주 분포하는 중앙 영역을 더욱 밀집시켜 중앙 구간에 대응되는 심볼의 발생 확률을 높이며, 이를 통해 엔트로피 부호화 효율을 향상시킨다. 둘째, 내부 코덱에 의해 발생하는 분포 왜곡을 보정하기 위한 QP 기반 스케일링 방법을 제안한다. 이 분포 왜곡은 QP에 따라 증가하는 경향이 있으므로 스케일링 강도를 QP로 조절함으로써 부호화된 특징의 동적 범위를 원신호에 가깝도록 복원시킨다. 이러한 두 방법의 결합을 통해 본 논문은 신경망의 임무 성능 정확도를 고려하며 특징 압축 효율을 개선한다.

II. 관련 연구

1. Feature Coding for Machines

기존 영상 부호화 표준은 인간 시각을 기준으로 설계되어 시각적 화질 보존을 목표로 하지만 신경망이 소비하는 특징 데이터는 시각적 품질보다는 신경망의 임무 성능이 우선이다. 또한 신경망 특징 데이터는 통계적 분포가 자연 영상 신호와 다를 수 있다. 이로 인해 기존 영상 코덱을 그대로 신경망 특징 데이터 부호화에 적용할 경우 압축 효율 저하나 임무 성능 손실이 발생할 수 있다.

이러한 문제를 해결하기 위해 MPEG은 기계 소비 환경에 최적화된 특징 압축을 목표로 FCM 표준화를 추진하고 있다. FCM은 신경망에서 생성되는 중간 특징을 새로운 부

a) 국립한밭대학교 지능미디어공학과(Hanbat National University)

b) 한국전자통신연구원(ETRI)

‡ Corresponding Author : 최해철(Haechul Choi)

E-mail: choihc@hanbat.ac.kr

Tel: +82-42-821-1149

ORCID: <https://orcid.org/0000-0002-7594-0828>

※ 본 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구이며(No. RS-2020-II200011, (전문연구실)기계를 위한 영상 부호화 기술), 2026년도 교육부 및 대전광역시의 재원으로 대전RISE센터의 지원을 받아 수행된 지역혁신중심 대학지원체계(RISE)의 결과입니다.(2026-RISE-06-002)

· Manuscript January 29, 2026; Revised March 10, 2026; Accepted March 10, 2026.

호화 대상으로 정의하고 압축률과 임무 수행 정확도 간의 균형을 유지하는 것을 목표로 한다^[6]. 이를 통해 분할 추론 환경에서의 네트워크 사용 부담을 완화한다.

FCM 표준화 과정에서는 다양한 기술 제안을 검증하기 위해 시험 모델인 FCTM을 개발하고, 정기적인 Core Experiment^[7]를 통해 FCTM에 적용된 기술들의 성능을 확인하고 있다. FCTM은 분할 추론 구조를 기반으로, 에지 측에서 생성된 중간 특징을 압축하여 서버로 전송하고 복원된 특징을 후속 신경망의 입력으로 제공하는 역할을 수행한다. 이러한 구조에서 FCTM은 특징 전송 효율과 복원 정확도를 결정하는 핵심 요소로 작용한다.

그림 1은 FCTM의 전체 부호화 파이프라인을 나타낸다. FCTM은 특징의 차원을 축소하는 Feature Reduction, 내부 코덱에서 처리 가능한 형태로 변환하는 Feature Conversion, 특징을 비트스트림으로 압축하는 Feature Encoding, 그리고 서버 측에서 비트스트림을 복원하는 Feature Decoding, 특징을 복원하는 Feature Inverse Conversion 및 Feature Restoration 단계로 구성된다^[3]. 이러한 단계적 파이프라인을 통해 FCTM은 중간 특징의 데이터량을 효과적으로 줄이면서도 다양한 시각 인식 임무에서 안정적인 특징 복원을 지원하도록 설계되어 있다.

Feature Conversion 단계에서는 특징을 정수값으로 변환하기 위해 균등 양자화^[4]가 사용되며, Feature Inner Coding 단계에서는 내부 코덱으로 VVC^[5]가 적용된다. VVC는 예측, 변환, 양자화, 확률 기반 엔트로피 부호화 등을 통해 특징을 비트스트림으로 압축한다. VVC는 QP에 따라 양자화 구간 크기가 달라지는 구조를 가지므로, QP가 증가할수록 큰 양자화 오류를 일으킨다. 이는 예측, 변환

등의 다양한 부호화 도구에서 누적 및 전파되어 복호화된 특징의 동적 범위가 축소되는 경향이 나타날 수 있다.

2. 양자화

특징 압축에서 양자화는 연속적인 값을 제한된 정수값으로 변환하는 단계로 압축 효율과 복원 정확도에 직접적인 영향을 미친다. 기존 연구에서는 입력 데이터의 분포 특성과 활용 목적에 따라 다양한 양자화 방법이 사용되어 왔으며 대표적으로는 균등 양자화, 비균등 양자화, 그리고 compander 기반 양자화가 널리 활용되고 있다.

균등 양자화는 그림 2-1과 같이, 전체 입력 범위를 동일한 간격으로 나누고, 각 구간을 하나의 대푯값으로 대응시키는 가장 단순한 양자화 방식이다. 모든 입력값은 자신이 속한 구간의 대푯값으로 변환되며 구간 간격이 일정하기 때문에 구현이 간단하고 계산 비용이 낮다는 장점이 있다. 이러한 특성으로 인해 균등 양자화는 영상 및 오디오 압축을 포함한 다양한 신호 처리 시스템에서 기본적인 양자화 방식으로 사용되어 왔다. 그러나 입력 데이터의 분포가 특정 구간에 집중되어 있는 경우에도 모든 구간에 동일한 정밀도를 적용하므로 데이터의 통계적 특성을 충분히 반영하지 못할 수 있다.

비균등 양자화는 그림 2-2와 같이, 입력 데이터의 분포나 중요도를 고려하여 양자화 구간의 간격을 다르게 설정하는 방식이다. 이 방법은 값이 자주 등장하거나 중요한 구간에서 더 세밀한 양자화 구간을 할당하고, 값이 자주 등장하지 않거나 상대적으로 덜 중요한 구간에는 넓은 양자화 구간을 적용함으로써 전체 양자화 오차를 줄인다. 다만 비균등 양

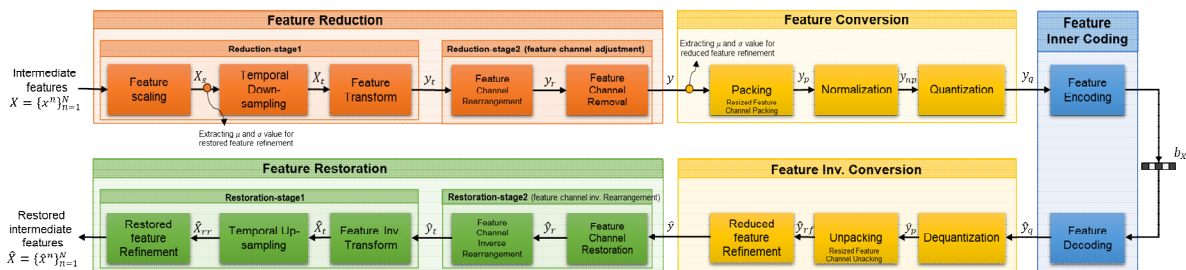


그림 1. FCTM 파이프라인^[3]

Fig. 1. FCTM Pipeline^[3]

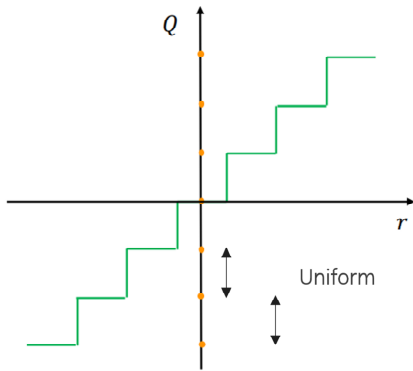


그림 2-1. 균등 양자화^[8]
 Fig. 2-1. Uniform quantization^[8]

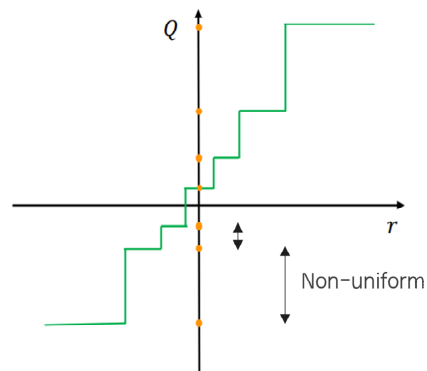


그림 2-2. 비균등 양자화^[8]
 Fig. 2-2. Non-uniform quantization^[8]

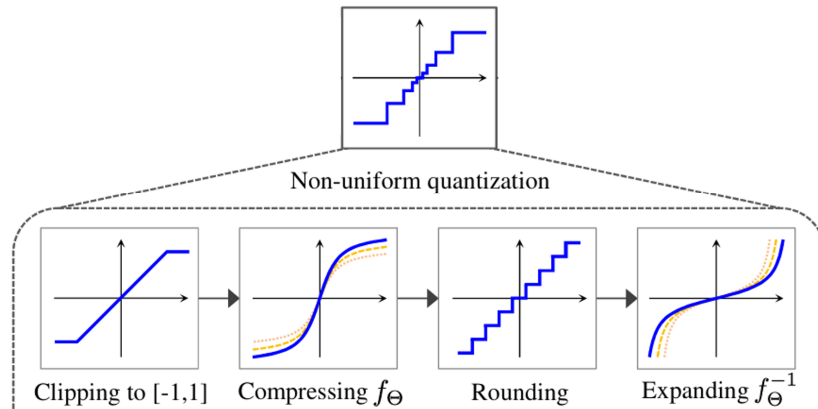


그림 3. Compander 기반 양자화^[9]
 Fig. 3. Compander-based quantization^[9]

자화는 입력 데이터의 분포를 사전에 정확히 파악해야 하며, 분포가 변하는 환경에서는 최적의 양자화 구간을 유지하기 어렵다는 한계가 있다.

Compander 기반 양자화^[9]는 그림 3과 같이, 비선형 변환과 균등 양자화를 결합한 방식이며 비균등 양자화와 유사한 효과를 단순한 구조로 구현할 수 있다. 이 방식에서는 입력 데이터를 비선형 함수로 먼저 변환하여 분포를 완화한 뒤 균등 양자화를 수행하고, 역변환을 적용하여 원래의 분포로 되돌린다. 이를 통해 균등 양자화의 단순성과 연산 효율을 유지하면서도 입력 데이터의 통계적 특성을 반영할 수 있다.

이와 같이 각 양자화 방식은 장단점이 있으며, 입력 데이터의 분포 특성과 압축 목적에 따라 서로 다른 방식이 사용

될 수 있다. 특히 신경망 중간 특징과 같이 비균등한 분포를 가지며 값에 따라 신경망 임무 정확도에 미치는 영향이 상이한 경우, 단순한 균등 양자화만으로는 압축 효율과 임무 정확도를 동시에 만족시키기 어렵기 때문에 분포 특성을 고려한 양자화 방법에 대한 연구가 지속적으로 이루어지고 있다.

III. 제안 방법

FCTM은 내부 코덱에서 특징을 압축하기 위해 균등 양자화를 사용한다. 앞서 기술한 바와 같이, 균등 양자화는 구현이 단순하고 안정적이라는 장점이 있으나 입력 특징의 통계적 분포를 고려하지 않고 특징값을 동일한 중요도로

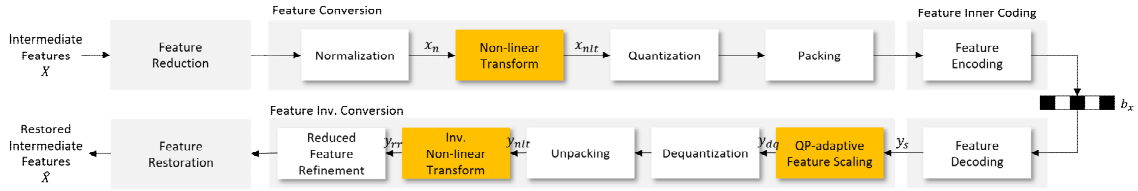


그림 4. 제안 방법의 파이프라인
Fig. 4. Pipeline of the proposed method

간주한다는 한계를 가진다. 신경망의 중간 특징은 값이 중앙 부근에 집중된 비균등 분포를 보이는 경우가 많으며, 특징값에 따라 임무 정확도에 미치는 영향이 서로 다를 수 있다. 이러한 특성은 균등 양자화 과정에서 충분히 반영되지 않는다.

또한 FCTM의 내부 코덱으로 사용되는 VVC는 QP에 따라 양자화 단계 크기가 조절되기 때문에 QP가 증가할수록 복호화된 특징의 동적 범위가 축소되고 분포가 중앙으로 압축되는 현상이 발생한다. 이러한 분포 변화는 특징 복원 과정에서 정확도 저하를 유발할 수 있다.

본 논문에서는 특징값의 임무 중요도 차이를 고려하지 못하는 균등 양자화의 한계와 내부 코덱의 QP에 의한 분포 축소 현상을 완화하기 위해 FCTM의 파이프라인에 통계적 특성 기반 비선형 변환과 QP 기반 스케일링을 결합한 방법을 제안한다. 제안 방법의 파이프라인은 그림 4와 같다.

1. 특징 신호의 비선형 변환

신경망 중간 특징의 분포는 일반적으로 중앙 영역에 값이 밀집된 비균등 형태를 보인다. 그러나 값이 밀집된 구간

이 항상 임무 수행에 중요한 정보를 포함하는 것은 아니다. 이를 확인하기 위해 본 연구에서는 FCTM의 x_n 특징값을 0.2 간격의 등간 구간으로 분할한 뒤, 각 구간에 속하는 특징값을 0으로 치환하여 임무 성능 변화를 측정하는 중요도 분석 실험을 수행하였다. 0.2 간격의 등간 분할은 특징값이 밀집된 중앙 영역을 독립적으로 분석함과 동시에 전체 범위를 균등하게 분할하여 밀집 구간과 비밀집 구간 간 성능 기여도를 비교하기 위해 적용하였다. 그림 5-1과 그림 5-2는 중요도 분석 실험 결과이다. 각 표에서 파란색으로 표시된 구간은 밀집 영역에 해당하며, 빨간색으로 표시된 구간은 특징값을 0으로 치환 시 가장 큰 성능 저하가 나타나는 구간을 의미한다. 본 분석 결과는 설명의 명확성을 위해 대표 시퀀스를 중심으로 제시하였으며, 제안하는 비선형 변환 설계는 특정 시퀀스에 종속되지 않도록 CTTC에서 사용되는 SFU 데이터셋에서 관찰된 공통적인 분포 경향을 바탕으로 도출하였다.

실험 결과, 값이 가장 밀집된 중앙 영역의 값을 0으로 치환하였을 때에도 성능 저하가 관찰되었으나, 일부 비밀집 구간을 0으로 치환하였을 때 더 큰 성능 감소가 나타나는 경우가 확인되었다. 이는 구간별 특징 개수 대비 성능 저하

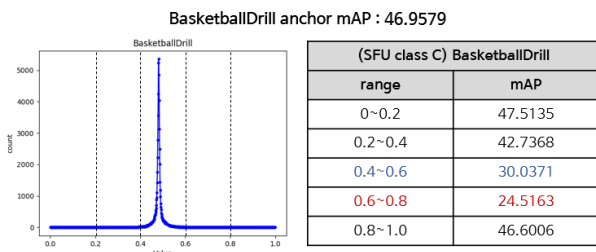


그림 5-1. BasketballDrill 시퀀스의 특징 분포 및 구간별 성능
Fig. 5-1. Feature distribution and interval-wise task performance for the BasketballDrill sequence

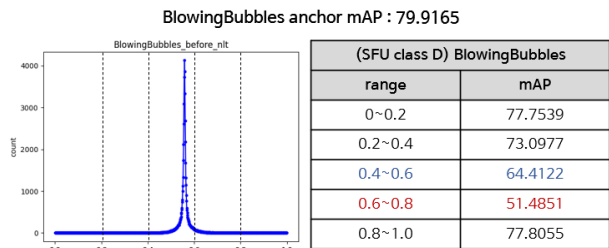


그림 5-2. BlowingBubbles 시퀀스의 특징 분포 및 구간별 성능
Fig. 5-2. Feature distribution and interval-wise task performance for the BlowingBubbles

의 상대적 크기를 고려할 때, 각 구간별 임무 성능에 대한 영향이 동일하지 않음을 보여준다. 특징 분포의 밀도와 성능 민감도는 반드시 비례하지 않으며, 상대적으로 적은 개수의 특징값이라도 임무 수행에 결정적인 영향을 미칠 수 있다.

이러한 관찰을 바탕으로 본 연구에서는 임무 수행에 덜 중요한 구간의 특징은 더 강하게 압축하고, 중요한 구간의 정보는 약하게 압축하도록 특징 분포를 조정하는 비선형 변환을 적용한다. 구체적으로, 중요도가 상대적으로 낮은 중앙 영역의 특징값에는 큰 양자화 구간을 적용하여 강하게 압축하고, 임무 수행에 중요한 구간의 특징값에는 작은 양자화 구간을 적용하여 정보 손실을 최소화한다.

비균등 양자화를 모사하기 위해, 비선형 변환은 균등 양자화 이전에 적용하여 입력 특징의 분포 형태를 변환한다. 비선형 변환 및 비선형 역변환 과정에서는 비선형 함수를 적용하기 위해, 식 (1), (2)와 같이 0부터 1 사이의 값을 갖는 특징을 -1과 1 사이의 범위로 변환하는 전처리하는 과정 $f_{pre}(x)$ 와 변환된 특징을 원래 범위로 복원하는 후처리 과정 $f_{post}(x)$ 을 포함한다.

본 절에서는 전처리 후처리 함수의 입력을 일반 변수 x 로 표기하고, 스케일링 이후의 특징값은 x' 로 표기한다. 모든 함수의 연산은 특징 텐서의 각 원소 단위로 적용된다.

$$f_{pre}(x) = 2x - 1 \quad (1)$$

$$f_{post}(x) = \frac{x+1}{2} \quad (2)$$

비선형 변환은 특징 x_n 에 대해 전처리 후, 식 (3)과 같이 inverse μ -law 함수^[10]를 사용한다. 이 함수는 파라미터 μ 를 통해 비선형 강도를 조절할 수 있어, 특징의 통계적 특성에 따라 변환 기울기를 유연하게 제어할 수 있으며, 중앙 영역을 상대적으로 압축하고 중앙에서 일부 떨어진 구간의 표현 범위를 확장하는 비선형 특성을 가진다. 본 연구에서는 이러한 특성을 활용하여 중앙 영역의 분포를 압축하고,

중앙 영역이 큰 양자화 스텝을 갖는 비균등 양자화 효과를 유도한다. 이로 인해 중요도가 낮은 구간의 값들은 양자화 과정에서 더 많은 값이 하나의 양자화 구간에 대응되며, 중요도가 높은 구간의 값들은 변환 이후에도 비교적 넓은 값 범위를 유지하므로 양자화 이후에도 임무 수행에 필요한 특징 정보가 안정적으로 보존된다. 이와 동시에, 분포가 밀집된 중앙 구간의 값들은 하나의 양자화 구간에 더 많이 대응되며, 엔트로피 부호화 과정에서 중앙 구간의 심볼 발생 확률이 증가하고 평균 부호 길이가 감소하여 전체 부호화 효율이 향상된다. 식 (4)는 비선형 변환의 전체 과정을 나타낸다.

$$f_{inv.\mu_{nt}}(x') = sign(x') \frac{(1 + \mu_{nt})^{|x'|} - 1}{\mu_{nt}} \quad (3)$$

$$x_{nt} = f_{post}(f_{inv.\mu_{nt}}(f_{pre}(x_n))) \quad (4)$$

비선형 변환이 적용된 특징은 복원 과정에서 대응되는 비선형 역변환을 통해 원래의 분포에 가깝게 복원한다. 비선형 역변환은 특징 y_{nt} 에 대해 전처리 후, 식 (5)와 같이 비선형 변환에서 사용된 함수의 역함수인 μ -law 함수^[10]를 사용하여, 압축된 특징 분포를 변환 이전의 형태에 가깝게 되돌린다. 이를 통해 부호화 효율을 향상시키면서 후속 신경망이 요구하는 특징 표현의 안정성을 유지할 수 있다. 식 (6)은 비선형 역변환의 전체 과정을 나타낸다.

$$f_{\mu_{nt}}(x') = sign(x') \frac{\ln(1 + \mu_{nt}|x'|)}{\ln(1 + \mu_{nt})} \quad (5)$$

$$y_{rr} = f_{post}(f_{\mu_{nt}}(f_{pre}(y_{nt}))) \quad (6)$$

비선형 변환의 강도는 비선형 파라미터 μ_{nt} 에 의해 결정된다. 식 (7)과 같이, 입력 특징의 통계적 특성에 따라 비선형 형성을 적응적으로 설정하기 위해 비선형 계수 μ_{nt} 는 입력 특징의 표준편차로 결정되며, 프레임 단위로 적응적으로 산출된다. 여기서 입력 특징 x 는 적용 단계에 따라 비선형

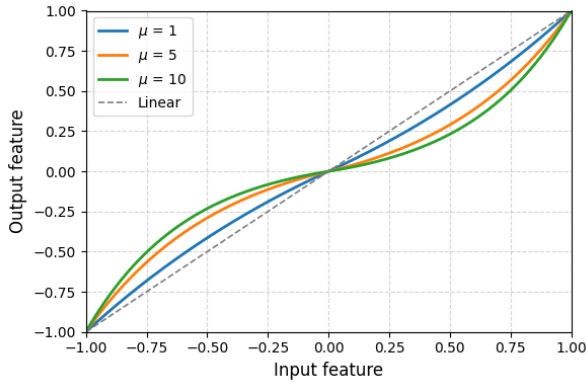


그림 6-1. 비선형 계수에 따른 비선형 변환 함수의 기울기
 Fig. 6-1. Slope of the nonlinear transform function with respect to the nonlinear parameter

변환 단계에서는 x_n , 비선형 역변환 단계에서는 y_{nt} 에 해당한다. 표준편차는 특징값이 평균으로부터 얼마나 분산되어 있는지를 나타내는 통계량으로, 표준편차를 기반으로 변환 강도를 조절함으로써 추가적인 학습이나 복잡한 최적화 과정 없이도 입력 분포 변화에 유연하게 대응할 수 있다. 그림 6-1과 그림 6-2는 μ_{nt} 값에 따른 부호화기에서 사용하는 비선형 변환 함수의 기울기와 복호기에서 사용하는 비선형 역변환 함수의 기울기를 나타낸다. 이를 통해 분포가 넓은 경우에는 더 강한 비선형 변환을 적용하여 중앙 부근의 분포를 강하게 압축하고, 분포가 좁은 경우에는 약한 비선형 변환을 적용하여 과도한 왜곡을 방지한다.

$$\mu_{nt} = \sigma_x \tag{7}$$

2. 특징 신호의 스케일링

내부 코덱에서 수행되는 양자화 과정에 의한 양자화 에러는 누적 및 전파되어 특징 데이터의 분포에 변화를 유발할 수 있다. 특히 FCTM의 내부 코덱으로 사용되는 VVC는 QP에 따라 양자화 구간 크기가 달라지므로, 복원된 특징이 어떻게 변화하는지 살펴보았다. 본 절에서 제시하는 분포 분석 및 파라미터 도출 결과는 설명의 명확성을 위해 대표 시퀀스를 중심으로 제시하였다. 제안하는 스케일링 설계는 특정 시퀀스에 종속되지 않으며, CTTC의 각 데이터셋에서

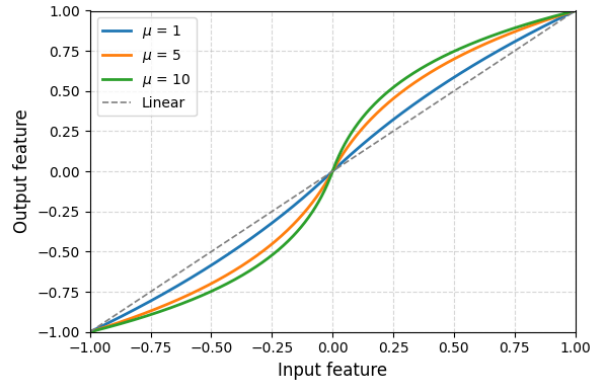


그림 6-2. 비선형 계수에 따른 비선형 역변환 함수의 기울기
 Fig. 6-2. Slope of the inverse nonlinear transform function with respect to the nonlinear parameter

선정한 두 개의 시퀀스에서 공통적으로 관찰된 분포 축소 경향을 기반으로 도출하였다. 그림 7은 다양한 QP 조건에서 내부 코덱 부호화 전후의 특징 분포를 비교한 실험 결과이다. 이 그림에서, QP가 증가함에 따라 분포의 양측 꼬리 영역이 줄어들고 특징값이 중앙으로 집중되는 경향을 직관적으로 확인할 수 있다.

즉, QP가 증가할수록 복원된 특징의 최소값과 최대값의 차이가 줄어들며, 특징의 동적 범위가 점진적으로 감소한다. 표 1은 HiEve-13 시퀀스의 QP 변화에 따라 내부 코덱 부호화 전후 특징의 동적 범위가 어떻게 달라지는지를 정량적으로 보여주며, QP가 커질수록 복원된 특징의 최소값은 증가하고 최대값은 감소하는 경향을 확인할 수 있다. 결론적으로 내부 코덱의 부·복호화 후 특징의 분포는 원본에 비해 중앙 영역으로 집중되면서 분포의 폭이 감소하며, 그 왜곡은 QP가 커질수록 증가한다.

이와 같은 분포 열화는 특징값의 범위 감소뿐만 아니라 통계적 특성의 변화로 이어진다. 이러한 변화는 이후 복원 과정에 영향을 미치며, 동적 범위가 축소된 상태에서는 원래의 분포를 충분히 복원하기 어렵다. 특히 QP가 높은 조건에서는 양자화로 인한 왜곡이 누적 및 전파되어, 후속 신경망에서 사용하는 특징의 안정성과 정확도를 저하시킬 수 있다. 따라서 내부 코덱 이후 단계에서 QP에 따라 달라지는 분포 열화 현상을 보정하는 과정이 필요하다.

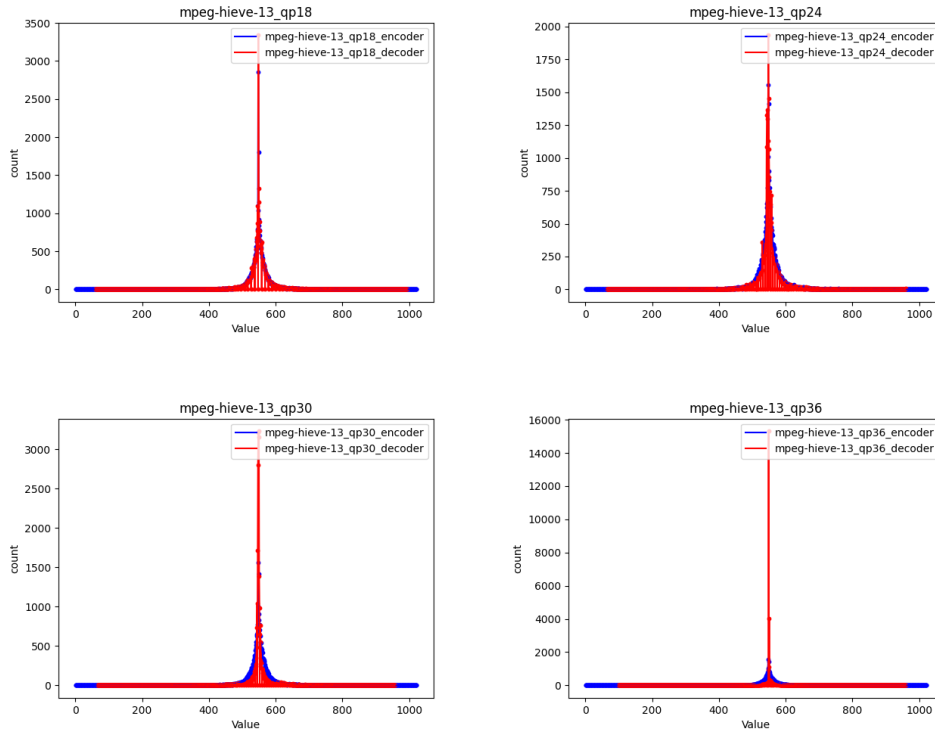


그림 7. HiEve-13 시퀀스의 QP 변화에 따른 내부 코덱 복호화 분포 열화

Fig. 7. Degradation of reconstructed feature distributions with varying QP values in the inner codec for the HiEve-13 sequence

표 1. HiEve-13 시퀀스의 QP 변화에 따른 특징의 동적 범위

Table 1. Dynamic range of features at under different QPs for the HiEve-13 sequence conditions

QP	Original feature input to VVC			Reconstructed feature output from VVC			ΔDR
	min	max	DR	min	max	DR	
7	0	1023	1023	64	988	924	99
12	0	1023	1023	64	986	922	101
17	0	1023	1023	64	991	927	96
22	0	1023	1023	64	960	896	127
27	0	1023	1023	54	966	912	111
32	0	1023	1023	67	960	893	130
37	0	1023	1023	88	960	872	151
42	0	1023	1023	132	960	828	195
47	0	1023	1023	204	949	745	278

이를 해결하기 위해 본 연구에서는 QP에 따른 특징 분포의 축소 경향을 모델링하고, 내부 코덱에서 복원된 특징의 동적 범위를 보정하는 스케일링 방법을 제안한다. VVC는 QP가 6 증가할 때마다 양자화 구간 크기가 두 배로 증가하

는 구조를 가지므로, 이 구조를 기반으로 QP에 따른 분포 축소비율을 스케일링 계수로 표현할 수 있다. 제안하는 스케일링 함수는 식 (8)과 같이, QP가 증가할수록 보정 강도가 점진적으로 커지며, 이를 통해 높은 QP 조건에서도 복

원된 특징 분포가 원본 분포에 가깝게 유지되도록 한다.

$$S(QP) = 2^{\alpha \left(\frac{QP-4}{6} \right)} + \beta \tag{8}$$

스케일링은 식 (9)와 같이, 특징의 평균을 기준점으로 이동시킨 뒤, QP에 따라 결정된 스케일링 계수를 적용하여 분포의 폭을 확장하고, 이후 다시 평균을 복원함으로써 분포의 중심 위치는 유지한 채 동적 범위만 보정한다. 이 과정을 통해 특징 분포의 중앙 집중 현상을 완화하고, 후속 복원 단계에서 원래 분포 특성을 더 안정적으로 복원할 수 있다.

$$y_{dq} = S(QP)(y_s - \mu_x) + \mu_x \tag{9}$$

스케일링 계수의 파라미터인 식 (8)의 α 와 β 는 데이터 기반으로 추정한다. 여러 영상 시퀀스를 대상으로 다양한 QP 조건에서 내부 코덱의 부호화 전 특징과 복호화 특징의 동적 범위를 측정하고 그 비율을 통해 QP에 따른 분포 축소

소 정도를 정량적으로 분석하였다. 측정된 비율과 스케일링 함수의 예측값이 유사해지도록 최소 제곱 기반의 커브 피팅(Curve Fitting)을 수행하였으며, 이를 통해 스케일링 강도를 조절하는 파라미터를 결정하였다. 그림 8-1과 그림 8-2는 다양한 QP 조건에서 측정한 분포 축소 비율과 제안한 스케일링 함수의 커브 피팅 결과를 나타낸다.

커브 피팅 결과, $\alpha=0.05$, $\beta=0.02$ 가 QP 변화에 따른 분포 축소 경향을 안정적으로 모델링하는 것으로 나타났다. 보정 항 β 를 포함하지 않은 경우 평균 피팅 오차는 3.92%였으며, 보정 항 β 를 포함한 경우 평균 피팅 오차는 3.54%로 감소하였다. 이는 보정 항 β 를 추가함으로써 피팅 정확도가 향상되었음을 의미한다. 단일 파라미터 α 만 사용하는 경우에도 안정적인 경향을 보였으나 보정 항 β 를 포함함으로써 평균 오차를 더욱 감소시킬 수 있었다. 최종적으로 본 연구에서는 $\alpha=0.05$, $\beta=0.02$ 를 스케일링 함수에 적용하여, QP 변화에 따른 특징 분포 축소를 효과적으로 보정한다.

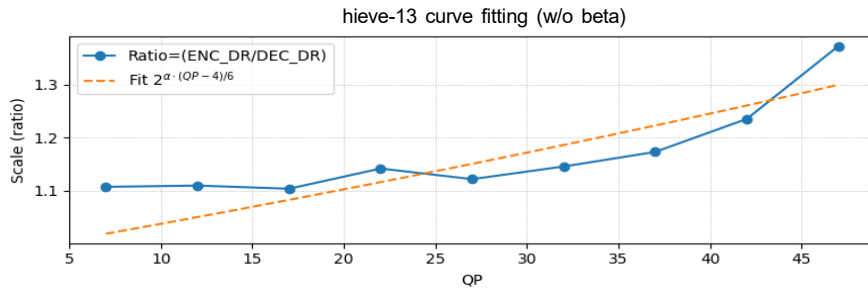


그림 8-1. HiEve-13 시퀀스의 스케일링 파라미터 α 값 피팅 그래프
 Fig. 8-1. Curve fitting result of the scaling parameter α for the HiEve-13 sequence

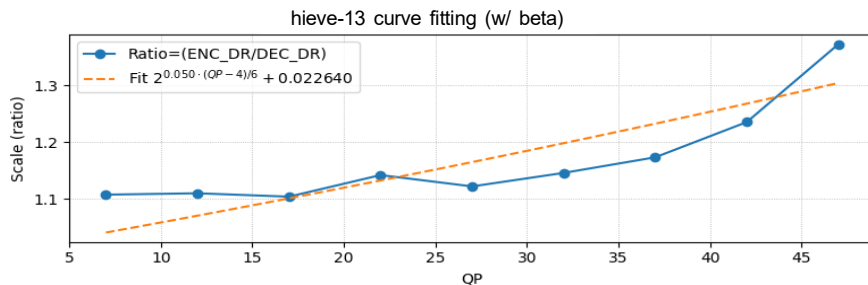


그림 8-2. HiEve-13 시퀀스의 스케일링 파라미터 α , β 값 피팅 그래프
 Fig. 8-2. Curve fitting results of the scaling parameters α and β for the HiEve-13 sequence

IV. 실험 결과

본 연구에서 제안 방법은 FCTM v8.1을 기반으로 구현하고, FCM 표준화에서 공정한 실험 및 평가를 위해 사용되는 Common Training and Test Conditions (CTTC)^[11]를 준수하여 실험을 수행하였다. 이에 따라 모든 실험은 동일한 신경망 구조^{[13][14][15]}, 데이터셋, 그리고 내부 코덱 설정 하에서 진행되었다. 제안 방법은 FCTM v8.1^[12]과 원격 추론 (remote inference) 방식과의 비교를 통해 평가하였다. 원격 추론 방식은 특징 압축 없이 예지 디바이스에서 생성된 중간 특징을 그대로 서버로 전송하여 추론을 수행하는 방식이다. SFU Class A/B 시퀀스의 경우 FCTM v8.1이 원격 추론 대비 낮은 임무 성능을 보이는 경향이 관찰되었으며, 본 연구의 모든 방법은 FCTM v8.1을 기반으로 구현되었기 때문에 해당 시퀀스에서는 원격 추론 대비 성능 차이가 동일한 경향으로 나타날 수 있다.

평가 대상 임무는 Instance Segmentation, Object detection, 그리고 Object Tracking으로 구성되며, 성능 지표로는 mAP@0.5^[16], mAP@0.5-0.95^[17], 그리고 MOTA^[18]를 사용하였다. 압축 효율 비교에는 Bjøntegaard Delta-rate (BD-rate)^[19]를 사용하였다. BD-rate가 음수인 경우는 동일한 임무 성능을 더 낮은 비트율로 달성했음을 의미하며, 이는 압축 효율이 향상되었음을 나타낸다. 또한 연산 복잡도를 확인하기 위해 FCTM v8.1 대비 인코딩 및 디코딩 시간 비율(EncR, DecR)을 측정하였으며, 비트율과 임무 성능 간의 관계를 시각적으로 분석하기 위해 Rate-Performance Curve (R-P Curve) 곡선을 함께 제시한다.

제안 방법의 효과를 명확히 분석하기 위해 실험은 비선형 변환만 적용하는 경우, 스케일링만 적용한 경우, 그리고 두 방법을 함께 적용한 경우로 나누어 수행하였다.

1. 비선형 변환

비선형 변환의 효과를 검증하기 위해, 비선형 계수를 고정값으로 설정한 경우와 입력 특징의 통계적 특성에 따라 적응적으로 설정한 경우를 비교하였다. 표 2-1은 비선형 계수를 고정값으로 적용했을 때의 실험 결과를 나타낸다. 고

정 계수를 적용한 경우 일부 데이터셋에서는 성능 개선이 관찰되었으나, 전체적으로는 데이터셋별 분포 차이를 충분히 반영하지 못해 성능이 불안정하게 나타났다. 그 결과, FCTM v8.1 대비 전체 평균 BD-rate 0.09%의 성능 하락이 발생하였다.

표 2-2는 입력 특징의 통계적 특성에 따라 비선형 계수를 적응적으로 적용했을 때의 실험 결과를 나타낸다. 이 경우 Instance Segmentation을 제외한 모든 데이터셋에서 일관된 성능 향상이 관찰되었으며, FCTM v8.1 대비 전체 평균 BD-rate 2.92%의 이득을 달성하였다. 이는 데이터셋에 따라 신경망 중간 특징의 분포 형태가 서로 다르다는 점을 고려할 때, 비선형 변환의 강도를 분포 특성에 맞게 조절하는 것이 효과적임을 보여준다.

고정된 비선형 계수는 이러한 분포 차이를 반영하지 못해, 일부 경우에는 과도한 분포 압축으로 중요한 특징 정보가 손실되거나, 반대로 변환 효과가 충분하지 않은 문제가 발생할 수 있다. 반면, 제안 방법은 특징의 표준편차를 이용해 분포의 확산 정도를 정량적으로 반영하고 이에 따라 비선형 변환의 강도를 조절한다. 분포가 넓은 경우에는 강한 비선형 변환을 적용하여 값들을 중앙 부근으로 집중시키고, 분포가 상대적으로 좁은 경우에는 약한 변환을 적용하여 불필요한 정보 손실을 방지한다.

이러한 통계적 특성 기반의 적응적 비선형 변환은 특징 분포를 엔트로피 부호화에 유리한 형태로 변환하면서도 임무 수행에 중요한 구간의 정밀도를 유지하여 전반적인 부호화 효율과 임무 성능을 동시에 향상시키는 결과로 이어진다.

BD-rate 개선의 원인을 명확히 분석하기 위해 FCTM v8.1과 제안하는 비선형 변환 방법의 평균 비트 감소율 및 평균 임무 성능 차이를 비교하였다. 전체 QP 조건에 대해 비트 감소율은 0.43%로 나타났으며, 평균 임무 성능은 0.05 향상되었다. 이는 임무 성능이 소폭 향상되었음을 의미하며, 동시에 비트 감소 효과가 상대적으로 더 크게 기여하였음을 보여준다. 따라서 본 실험에서 관찰된 BD-rate 개선은 성능 향상과 함께 평균 부호 길이 감소가 복합적으로 작용한 결과로 해석할 수 있다.

표 2-1. 고정값($\mu_{nl}=0.05$) 비선형 변환 실험 결과

Table 2-1. Experimental results of the nonlinear transform with a fixed parameter ($\mu_{nl}=0.05$)

Task	Dataset	Nonlinear transform(fixed) vs FCTM v8.1	Nonlinear transform(fixed) vs remote inference
		BD-rate	BD-rate
Instance Segmentation	OpenImageV6	7.07%	-95.25%
Object detection	OpenImageV6	-0.29%	-94.94%
	SFU Class A/B	0.74%	18.38%
	SFU Class C	4.16%	-87.05%
	SFU Class D	0.96%	-87.93%
Object Tracking	TVD(OVERALL)	-7.62%	-96.03%
	HIEVE(1080p)	-1.96%	-95.16%
	HIEVE(720p)	-2.33%	-92.66%
OVERALL		0.09%	-78.83%

표 2-2. 통계적 특성 기반 적응적 비선형 변환 실험 결과

Table 2-2. Experimental results of the adaptive nonlinear transform based on statistical feature characteristics

Task	Dataset	Nonlinear transform(std) vs FCTM v8.1			Nonlinear transform(std) vs remote inference
		BD-rate	EncR	DecR	BD-rate
Instance Segmentation	OpenImageV6	1.47%	100.863%	96.941%	-95.64%
Object detection	OpenImageV6	-7.33%	98.451%	94.054%	-95.46%
	SFU Class A/B	-7.58%	97.905%	88.916%	12.19%
	SFU Class C	-0.67%	102.511%	86.353%	-88.69%
	SFU Class D	-4.15%	99.306%	84.282%	-88.23%
Object Tracking	TVD(OVERALL)	-1.36%	95.392%	84.547%	-96.01%
	HIEVE(1080p)	-3.20%	93.978%	85.471%	-95.50%
	HIEVE(720p)	-0.56%	94.668%	84.630%	-92.51%
OVERALL		-2.92%	99.071%	89.182%	-79.98%

2. 스케일링

스케일링은 내부 코텍의 부호화 과정으로 인해 발생하는 분포 열화를 완화하여 복원된 특징의 통계적 범위를 안정적으로 유지하는 것을 목표로 한다.

표 3은 QP 기반 스케일링을 적용했을 때의 성능을 FCTM v8.1 및 원격 추론 방식과 비교한 결과를 나타낸다. 제안한 스케일링은 Instance Segmentation을 제외한 모든 데이터셋에서 일관된 성능 향상이 관찰되었으며, FCTM v8.1 대비 전체 평균 1.28%의 BD-rate 이득을 보였다. 이는 스케일링을 통해 내부 코텍의 양자화로 인한 정보 손실이

효과적으로 보정되었음을 보여준다.

그림 9는 스케일링 적용 전후의 특징 분포를 비교한 결과를 보여준다. 스케일링이 적용된 경우, QP가 증가하더라도 복원된 특징의 최소값과 최대값 범위가 더 안정적으로 유지되는 것을 확인할 수 있다. 내부 코텍 부호화 전후 특징의 동적 범위 차이는 스케일링 적용 전 평균 143.1에서 적용 후 36.7로 감소하였다. 이는 스케일링을 통해 내부 코텍 부호화 과정에서 발생하는 분포 왜곡이 완화되었음을 의미한다. 이러한 스케일링은 후속 복원 과정에서 특징 분포의 안정성을 높이며, 결과적으로 임무 정확도 향상에 기여한다.

표 3. 스케일링 실험 결과

Table 3. Experimental results of the scaling

Task	Dataset	Scaling vs FCTM v8.1			Scaling vs remote inference
		BD-rate	EncR	DecR	BD-rate
Instance Segmentation	OpenImageV6	3.16%	98.209%	104.997%	-95.50%
Object detection	OpenImageV6	-6.08%	101.555%	88.609%	-95.37%
	SFU Class A/B	-0.28%	99.313%	87.992%	20.85%
	SFU Class C	-2.45%	100.997%	83.457%	-89.30%
	SFU Class D	-0.13%	98.483%	84.456%	-87.65%
Object Tracking	TVD(OVERALL)	-0.54%	94.359%	79.652%	-95.99%
	HIEVE(1080p)	-1.29%	94.906%	79.970%	-95.85%
	HIEVE(720p)	-2.64%	94.467%	77.602%	-92.64%
OVERALL		-1.28%	98.819%	88.194%	-78.93%

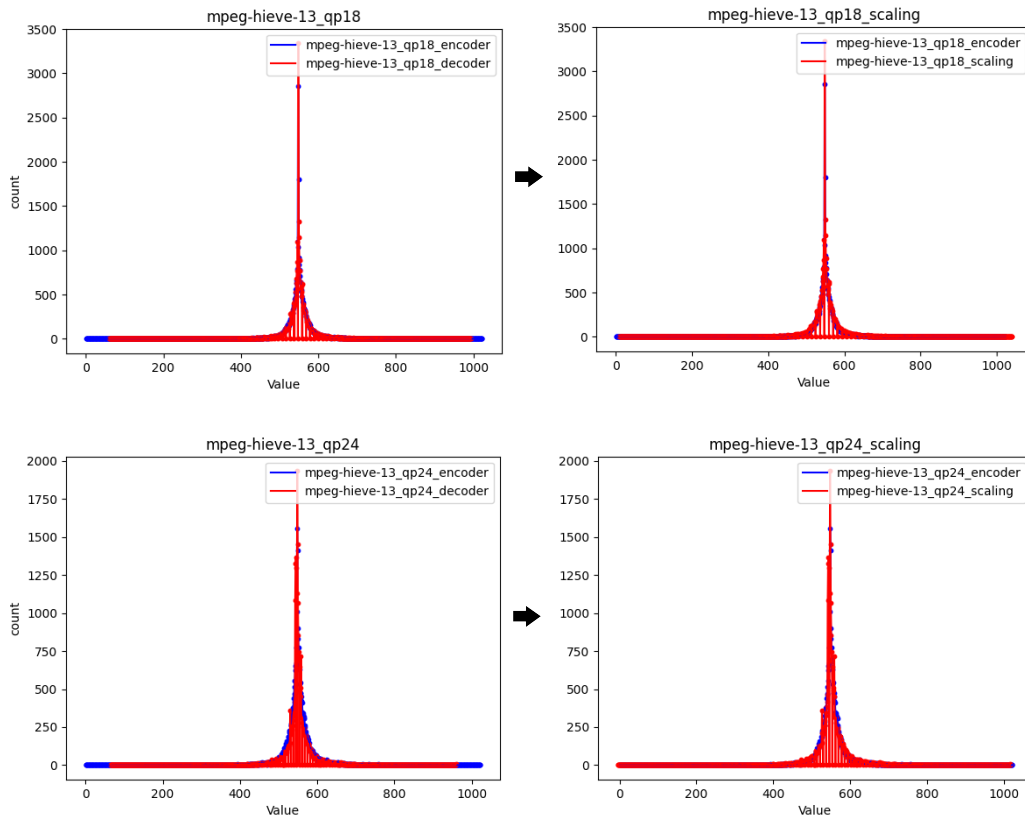


그림 9. HiEve-13 시퀀스의 특징 분포 비교 : (좌) 스케일링 적용 전, (우) 스케일링 적용 후

Fig. 9. Feature distributions of the HiEve-13 sequence : (left) before scaling and (right) after scaling

3. 비선형 변환 및 스케일링

마지막으로 비선형 변환과 스케일링을 함께 적용한 경우

의 성능을 분석하였다. 비선형 변환은 특징값의 임무 중요도 차이를 반영하여, 중요도가 낮은 구간에는 넓은 양자화 구간을 적용하고 중요한 구간에는 세밀한 양자화를 적용함

으로써 임무 성능을 고려한 부호화 효율을 향상시킨다. 또한 임무 성능에 대한 영향이 상대적으로 작은 값들이 중앙 구간에 집중된 분포 특성으로 인해, 중앙 구간의 심볼 발생 확률이 증가하면서 엔트로피 부호화 효율도 함께 개선된다. 스케일링은 내부 코덱의 부호화 과정으로 인해 복원된 특징 분포가 축소되는 현상을 완화하여, 이후 복원 과정에서 더 안정적인 특징 분포를 유지하도록 한다.

두 방법을 함께 적용할 경우, 내부 코덱 이후의 특징 분포가 스케일링을 통해 안정화된 상태에서 비선형 역변환이 수행된다. 이로 인해 단일 방법을 적용했을 때보다 비선형 역변환이 더 안정적으로 작동하며, 그 결과 압축 효율과 임

무 성능에서 일관된 성능 향상이 나타난다.

표 4는 비선형 변환과 스케일링을 함께 적용한 제안 방법의 성능을 FCTM v8.1 및 원격 추론 방식과 비교한 결과를 나타낸다. Instance Segmentation 데이터셋에서는 비선형 변환 및 스케일링 단독 적용 시와 마찬가지로 성능 감소 경향이 관찰되었다. 이는 segmentation 임무가 객체 경계 및 픽셀 단위 정밀도에 민감한 구조적 특성을 가지기 때문에, 특징 비선형 변환 및 스케일링 과정에서 세밀한 공간 정보가 일부 손실되었기 때문으로 분석된다. 그러나 특정 임무에서의 성능 하락에도 불구하고 제안 방법은 FCTM v8.1 대비 전체 평균 3.14%의 BD-rate 이득을 달성하여, 비

표 4. 제안하는 비선형 변환 및 스케일링을 결합한 실험 결과

Table 4. Experimental results of the proposed method combining nonlinear transform and scaling

Task	Dataset	Proposal (Nonlinear transform+Scaling) vs FCTM v8.1			Proposal vs remote inference
		BD-rate	EncR	DecR	BD-rate
Instance Segmentation	OpenImageV6	1.92%	100.818%	97.991%	-95.64%
Object detection	OpenImageV6	-8.06%	99.539%	101.040%	-95.49%
	SFU Class A/B	-9.11%	97.663%	88.377%	8.62%
	SFU Class C	-0.55%	101.934%	83.573%	-88.72%
	SFU Class D	-4.03%	98.880%	88.668%	-88.24%
Object Tracking	TVD(OVERALL)	-1.79%	94.108%	80.366%	-96.03%
	HIEVE(1080p)	-2.72%	92.992%	81.186%	-95.56%
	HIEVE(720p)	-0.76%	94.210%	80.733%	-92.53%
OVERALL		-3.14%	98.824%	90.002%	-80.45%

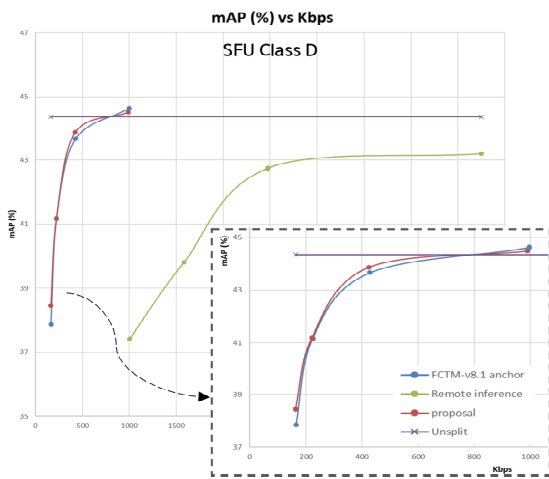


그림 10-1. SFU Class D에 대한 제안하는 비선형 변환 및 스케일링의 R-P Curve

Fig. 10-1. R-P Curves of the proposed nonlinear transform and scaling method for the SFU Class D

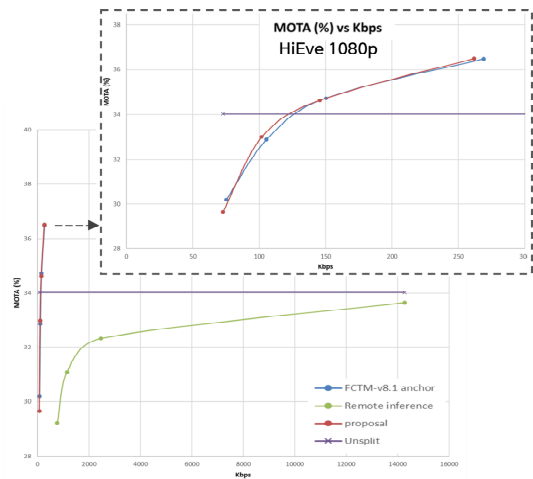


그림 10-2. HiEve 1080p에 대한 제안하는 비선형 변환 및 스케일링의 R-P Curve

Fig. 10-2. R-P Curves of the proposed nonlinear transform and scaling method for the HiEve 1080p

선형 변환 또는 스케일링을 단독으로 적용한 경우보다 더 큰 압축 효율 향상을 보였다. 또한 EncR 및 DecR 결과에서 확인할 수 있듯이 제안 방법의 인코딩 및 디코딩 시간은 FCTMv8.1 대비 유의미한 증가 없이 유지하였다. 이는 비선형 변환 및 스케일링 방법에 따른 연산 복잡도 증가가 제한적이며, 추가적인 지연 없이 향상된 압축 효율을 제공한다는 점을 시사한다.

그림 10-1과 그림 10-2는 일부 데이터셋에 대한 R-P Curve를 나타낸 것이다. 가로축은 비트율(kbps)을, 세로축은 해당 임무의 성능 지표를 나타낸다. 비선형 변환과 스케일링을 함께 적용한 경우, 곡선이 FCTM v8.1 및 원격 추론 대비 전반적으로 좌상단으로 이동하는 경향을 보였다. 이는 비선형 변환을 통해 동일한 임무 성능을 더 낮은 비트율로 달성하고, 스케일링을 통해 동일한 비트율에서 더 높은 임무 성능을 달성하였음을 의미한다. 결과적으로, 비선형 변환과 스케일링을 함께 적용한 제안 방법은 부호화 효율과 복원 안정성을 동시에 확보함으로써 개별 방법 대비 더 큰 성능 향상을 달성하였다.

V. 결론

본 논문에서는 FCM 환경에서 중간 특징을 압축할 때 발생하는 두 가지 한계를 분석하고, 이를 개선하기 위한 비선형 변환과 스케일링 방법을 제안하였다. FCTM은 균등 양자화를 기반으로 설계되어 있어 신경망 중간 특징이 가지는 비균등한 통계적 분포뿐만 아니라 특징값에 따라 임무 성능에 미치는 영향이 서로 다르다는 점을 충분히 반영하지 못한다. 또한 내부 코덱의 부호화 과정에서는 QP 증가에 따라 복원된 특징 분포의 동적 범위가 축소되는 문제가 발생한다. 이러한 한계는 부호화 효율 저하와 복원 품질 및 임무 정확도 감소로 이어질 수 있다.

제안한 비선형 변환은 입력 특징의 통계적 특성과 중요도 분석 결과를 바탕으로 임무 수행에 상대적으로 덜 중요한 구간은 더 강하게 압축하고 중요한 구간의 표현 정밀도는 유지하도록 특징 분포를 조정하였다. 특히 특징 분포는 임무 수행에 덜 중요한 구간에 집중되어 있으므로 이를 강하게 압축함으로써 심볼 발생 확률을 집중시킬 수 있다. 그

결과, 엔트로피 부호화 과정에서 평균 부호 길이가 감소하며 전반적인 압축 효율이 향상된다. 또한 비선형 계수를 입력 특징의 표준편차에 따라 적응적으로 설정함으로써 데이터 특성 변화에 유연하게 대응할 수 있음을 보였다.

스케일링은 내부 코덱의 부호화로 인해 발생하는 특징 분포 축소 현상을 완화하는 데 목적이 있다. QP 증가에 따라 복원된 특징의 동적 범위가 감소하는 현상을 분석하고, 이를 보정하기 위한 스케일링 함수를 도입함으로써 복원된 특징 분포가 내부 코덱 이전의 통계적 범위에 가깝게 유지되도록 하였다. 이러한 스케일링은 후속 복원 과정에서 특징 분포의 안정성을 높여 임무 수행 정확도 향상에 기여한다.

실험은 FCM의 CTC를 준수하여 수행되었으며 Instance Segmentation, Object detection, 그리고 Object Tracking 임무를 대상으로 제안 방법의 성능을 분석하였다. 그 결과, FCTM v8.1을 기준으로 비선형 변환 단독 적용 시 평균 BD-rate 2.92%, 스케일링 단독 적용 시 평균 BD-rate 1.28%의 성능 향상을 확인하였고, 두 방법을 함께 적용한 경우에는 평균 BD-rate 3.14%의 성능 향상을 달성하였다. 이는 비선형 변환을 통해 향상된 부호화 효율이 스케일링을 통해 안정화된 특징 분포 위에서 더 효과적으로 나타났기 때문이다.

제안 방법은 기존 FCTM 부호화 구조를 유지하면서도 특징 분포의 통계적 특성과 내부 코덱의 양자화 특성을 함께 고려함으로써, 압축 효율과 임무 성능을 동시에 개선할 수 있음을 보여준다. 향후에는 비선형 변환과 스케일링 파라미터의 추가적인 정교화를 통해 더 일반화된 특징 압축 프레임워크로 발전시킬 수 있을 것으로 기대된다.

참고 문헌 (References)

- [1] S. Liu and C. Rosewarne, "AHG on Video Coding for Machines," ISO/IEC JTC 1/SC 29/WG 4, m73899, Oct. 2025.
- [2] C. Rosewarne and Y. Zhang, "AHG on Feature Coding for Machines," ISO/IEC JTC 1/SC 29/WG 4, m73872, Oct. 2025.
- [3] "Algorithm description of FCTM," ISO/IEC JTC 1/SC 29/WG 4 N0704, July 2025.
- [4] S. Lee et al., "[FCM] Non-CE4: Quantization modification", ISO/IEC JTC 1/SC 29/WG 4 m70203, Kemer, Nov. 2024.

- [5] B. Bross et al., "Overview of the Versatile Video Coding (VVC) Standard and Its Applications," IEEE Transactions on Circuits and Systems for Video Technology, Vol.31, No.10, pp.3736 - 3764, Oct. 2021.
doi: <https://doi.org/10.1109/TCSVT.2021.3101953>
- [6] Convenor of ISO/IEC JTC 1/SC 29/AG 03, "Press release of 144th MPEG meeting," ISO/IEC JTC 1/SC 29/AG 03 N0129, October 2023.
- [7] H. Jeong and H. Y. Kim, "[FCM][CE4] Summary report," in ISO/IEC JTC 1/SC 29/WG 4 m74448, Geneva, Oct. 2025.
- [8] A. Gholami et al., "A Survey of Quantization Methods for Efficient Neural Network Inference," arXiv:2103.07156, 2021.
doi: <https://doi.org/10.48550/arXiv.2103.07156>
- [9] K. Yamamoto, "Learnable Companding Quantization for Accurate Low-Bit Neural Networks," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp.5029 - 5038, 2021.
doi: <https://doi.org/10.1109/CVPR46437.2021.00499>
- [10] Cisco Systems, "Waveform Coding Techniques," Cisco Technical Note, [Online]. Available: <https://www.cisco.com/c/en/us/support/docs/voice/h323/8123-waveform-coding.pdf>
- [11] "Common Test and Training Conditions for FCM," ISO/IEC JTC 1/SC 29/WG 4 N0705, June-July. 2025.
- [12] "Feature Compression Test Model (FCTM) v8.1.," [Online]. Available: <https://git.mpeg.expert/MPEG/Video/fcm/fctm.git>
- [13] R. Girshick, "Fast R-CNN," Proceedings of the IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, pp.1440 - 1448, 2015.
doi: <https://doi.org/10.1109/ICCV.2015.169>
- [14] K. He et al., "Mask R-CNN," Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, pp.2961 - 2969, 2017.
doi: <https://doi.org/10.1109/ICCV.2017.322>
- [15] Z. Wang et al., "Towards Real-Time Multi-Object Tracking," 2020. [Online]. Available: <https://github.com/Zhongdao/Towards-Realtime-MOT>
- [16] M. Everingham et al., "The Pascal Visual Object Classes (VOC) Challenge," International Journal of Computer Vision, Vol.88, No.2, pp.303 - 338, June 2010.
doi: <https://doi.org/10.1007/s11263-009-0275-4>
- [17] T.-Y. Lin et al., "Microsoft COCO: Common Objects in Context," European Conference on Computer Vision (ECCV), 2014.
doi: https://doi.org/10.1007/978-3-319-10602-1_48
- [18] K. Bernardin and R. Stiefelwagen, "Evaluating Multiple Object Tracking Performance: The CLEAR MOT Metrics," EURASIP Journal on Image and Video Processing, 2008.
doi: <https://doi.org/10.1155/2008/246309>
- [19] Bjøntegaard G., "Calculation of average PSNR differences between RD-curves," in ITU-T SG 16 Q.6 document VCEG-M33, 13th VCEG meeting, Austin, Texas, USA, Apr. 2001.

저 자 소 개



이 성 은

- 2024년 : 국립한밭대학교 정보통신공학 (학사)
- 2026년 : 국립한밭대학교 지능미디어공학 (석사)
- ORCID : <https://orcid.org/0009-0004-4009-8775>
- 주관심분야 : 비디오 부호화, 특징 부호화, 컴퓨터 비전, 딥러닝



서 예 원

- 2025년 : 국립한밭대학교 정보통신공학 (학사)
- 2025년 ~ 현재 : 국립한밭대학교 지능미디어공학 석사과정
- ORCID : <https://orcid.org/0009-0002-3443-4774>
- 주관심분야 : 비디오 부호화, 특징 부호화, 컴퓨터 비전, 딥러닝

저 자 소 개



한 희 지

- 2018년 : 국립한밭대학교 정보통신공학 (학사)
- 2020년 : 국립한밭대학교 멀티미디어공학 (석사)
- 2026년 : 국립한밭대학교 지능미디어공학 (박사)
- ORCID : <https://orcid.org/0000-0001-8187-2553>
- 주관심분야 : 비디오 부호화, 특징 부호화, 컴퓨터 비전, 딥러닝



정 순 흥

- 2001년 2월 : 부산대학교 전자공학 (학사)
- 2003년 2월 : KAIST 전기및전자공학 (석사)
- 2016년 2월 : KAIST 전기및전자공학 (박사)
- 2003년 ~ 2005년 : LG전자 단말연구소 주임연구원
- 2019년 ~ 2020년 : 인디애나대학교 방문연구원
- 2005년 ~ 현재 : 한국전자통신연구원 책임연구원
- ORCID : <https://orcid.org/0000-0003-2041-5222>
- 주관심분야 : 실감미디어, 컴퓨터 비전, 영상부호화, 영상처리



추 현 군

- 1998년 2월 : 한양대학교 전자공학과(공학사)
- 2000년 2월 : 한양대학교 전자공학과(공학석사)
- 2005년 2월 : 한양대학교 전자통신전파공학과(공학박사)
- 2015년 ~ 2017년 : 한국전자통신연구원 디지털홀로그래피연구실장
- 2017년 ~ 2018년 : Warsaw University of Technology, Poland 방문연구원
- 2023년 ~ 현재 : 한국전자통신연구원 실감미디어연구실장
- ORCID : <https://orcid.org/0000-0002-0742-5429>
- 주관심분야 : Computer Vision, 3D imaging and holography, 3D depth imaging, 3D broadcasting system



최 해 철

- 1997년 : 경북대학교 전자공학 (학사)
- 1999년 : 한국과학기술원 전기및전자공학 (석사)
- 2004년 : 한국과학기술원 전기및전자공학 (박사)
- 2004년 ~ 2010년 : 한국전자통신연구원 선임연구원
- 2010년 ~ 현재 : 국립한밭대학교 지능미디어공학과 교수
- ORCID : <https://orcid.org/0000-0002-7594-0828>
- 주관심분야 : 비디오 부호화, 컴퓨터 비전, 딥러닝