

특집논문 (Special Paper)

방송공학회논문지 제30권 제3호, 2025년 5월 (JBE Vol.30, No.3, May 2025)

<https://doi.org/10.5909/JBE.2025.30.3.377>

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

MPEG FCM에서의 효과적인 비활성화 채널 복원 방법에 관한 연구

이동민^{a)}, 정원식^{b)}, 서정일^{a)†}

Research on Effective Restoration Method for Inactive Channel in MPEG FCM

Dongmin Lee^{a)}, Won-Sik Cheong^{b)}, and Jeongil Seo^{a)†}

요약

인공지능이 발전함에 따라 기계가 시각 데이터를 처리하는 경우가 증가하고 있다. 이에 인간의 시각 인지 체계에 최적화된 영상 데이터 대신, 기계가 영상을 인지하는 데 최적화된 영상 데이터를 다루는 연구가 진행되는 추세이다. 이에 MPEG에서는 특징 맵을 효율적으로 압축하고 처리하는 기술인 FCM을 연구하고 있으며, 현재 표준화를 진행하고 있다. 본 연구는 FCM의 세부 기술 중 하나인 QA-RBICR 기술에 대한 개선을 진행하여 성능을 향상하였다. 본 연구는 기존 QA-RBICR에서 비활성화된 채널을 복원하는 알고리즘을 변경하여 EICR이라는 새로운 네 가지 방식을 제시하였으며, 그 중 Mean interpolate 방식은 SFU 데이터셋에서 Class C와 Class D 각각 0.72%, 3.28%의 BD-rate 향상을 달성하였고, HiEve-1080p 데이터셋에서 0.97%의 BD-rate 향상을 달성하였다.

Abstract

As artificial intelligence advances, machines are increasingly processing visual data. Therefore, it has become increasingly important to utilize image data optimized for machine perception rather than for human vision. To address this trend, MPEG is currently researching and standardizing FCM, a technology aimed at efficiently compressing and processing feature maps. This study focuses on improving the QA-RBICR technique, a specific component of FCM. We propose modifications to the inactive channel restoration algorithm used in QA-RBICR and introduce four new EICR methods. Among these, the Mean interpolate method achieved BD-rate improvements of 0.72% and 3.28% for Class C and Class D on the SFU dataset, respectively, and a BD-rate improvement of 0.97% on the HiEve-1080p dataset.

Keyword : Video coding for machines, Feature coding for machines, Video coding, RBICR

a) 동아대학교 컴퓨터공학과(Dept. of Computer Engineering, Dong-A University)

b) 한국전자통신연구원(Electronics and Telecommunications Research Institute, ETRI)

† Corresponding Author : 서정일(Jeongil Seo)

E-mail: jeongilseo@dau.ac.kr

Tel: +82-51-200-7796

ORCID: <https://orcid.org/0000-0001-5131-0939>

※ 본 연구는 과학기술정보통신부의 재원으로 정보통신기획평가원(No.2020-0-00011, 기계를 위한 영상 부호화 기술, No.2023-0-00076, SW중심대학(동아대학교))의 지원을 받아 수행된 결과임.

· Manuscript May 1, 2025; Revised May 7, 2025; Accepted May 7, 2025.

I. 서 론

최근 인공지능 기술의 급속한 발전은 다양한 분야에 혁신을 가져왔으며 객체 탐지, 객체 추적, 객체 분할과 같은 컴퓨터비전 분야와 영상 처리 분야에서 두드러진 성과를 보이고 있다.

하지만 인공지능의 시각 처리는 발전하고 있는 데 반해 이에 사용되는 대부분의 영상 데이터는 기본적으로 인간의 시각 인지 체계에 최적화되어 있다는 한계가 존재한다. 이는 기존의 영상 데이터는 인공지능 및 기계 학습 모델의 시각적 특징 추출 및 처리 효율성 관점에서 최적화되어 있지 않다는 한계를 가짐을 의미하며, 인공지능이 영상 데이터를 처리하는 데 있어 성능 저하를 초래할 수 있음을 뜻한다.

이에 MPEG(Moving Picture Experts Group)은 Video Coding for Machines(VCM)과 Feature Coding for Machines(FCM)라는 이름으로 기계 시각에 최적화된 영상 압축 기술에 대한 연구를 진행하고 있으며, 현재 표준화 작업을 진행하는 중이다. 이 중 FCM은 VCM에서 파생된 기술로, 특징 맵을 효율적으로 압축하는 기술이다. 이때 FCM의 기술을 구현하여 소프트웨어로 만든 것을 Feature Compression Test Model(FCTM)이라 한다^[1].

본 연구에서는 FCTM의 알고리즘 중 하나인 QP-Adaptive Range-based Inactive Channel Removal with Selective Learning Strategy(QA-RBICR)^[2]에 주목하였다. QA-RBICR은 채널 기반 중요도 순으로 정렬된 특징 맵에서 비활성 채널을 삭제하여 비트레이트를 효과적으로 절감하는 기술이며, 우수한 성능을 바탕으로 MPEG 149차 회의에서 공식적으로 채택되어 FCTM-v6.1에 탑재되었다. FCTM 인코더에서 특징 맵의 일부 채널을 삭제한 뒤 FCTM 디코더 측에서 이를 복원하는 기술은 MPEG 회의에 여러 번 기고된 바 있다. 그러나 이러한 기술 중, 현재 FCTM에 적용된 기술은 QA-RBICR이 유일하다.

현재까지 기고된 특징 맵의 채널을 삭제한 뒤 복원하는 알고리즘들은 모두 FCTM 디코더 측에서 채널을 복원할 때, 보유한 정보를 적극적으로 활용하지 못한다는 한계가 존재한다. 이는 QA-RBICR에서도 마찬가지이며, 이에 본 연구는 채널 삭제 및 복원을 다루는 기술 중 유일하게 FCTM에 적용된 기술인 QA-RBICR에 대해 비활성 채널

삭제 후 복원 시의 알고리즘에 대한 개량을 수행하였다. 개량한 알고리즘으로 SFU 데이터셋 및 HiEve-1080p 데이터셋에 대한 성능 실험을 진행한 결과 성능 향상을 보일 수 있었다.

본 논문은 2장에서 FCM의 개요 및 FCM에서의 특징 맵 관련 기존 연구를, 3장에서 QA-RBICR에 대해 본 연구에서 제안하는 알고리즘을 소개한다. 이후 4장에서는 알고리즘을 적용했을 때의 성능 변화를 보이고, 마지막으로 5장에서 연구에 대한 결론 및 추후 연구를 논한다.

II. 관련 연구

1. Feature Coding for Machines

FCM은 효율적으로 특징 맵을 압축하고, 압축한 특징 맵을 다시 복원하는 기술이다. FCM을 소프트웨어로 구현한 FCTM은 신경망 네트워크에서 특징 맵을 추출하는 Neural Network part 1(NN part 1), 특징 맵을 활용하여 컴퓨터비전 작업을 처리하는 Neural Network part 2(NN part 2) 사이의 분할점으로 작용한다.

FCTM은 NN part 1에서 추출된 특징 맵을 압축하는 Feature reduction 모듈, 압축된 특징 맵을 비트스트림으로 변환하기 이전에 전처리를 수행하는 Feature conversion 모듈, 특징 맵을 비트스트림으로 변환하거나 비트스트림을 다시 특징 맵으로 변환하는 Feature inner coding 모듈, 압축된 특징 맵을 다시 복원하기 이전에 전처리를 수행하는 Feature inverse conversion 모듈, 압축된 특징 맵을 NN part 2에 입력하기 위해 복원하는 Feature restoration 모듈로 이루어져 있다.

본 논문에서 다루는 QA-RBICR은 Feature reduction 모듈과 Feature restoration 모듈에 탑재된 기술로, 압축된 특징 맵에 대해 비활성 채널을 삭제하는 후처리를 진행함으로써 비트레이트를 절감하는 기술이다.

2. QP-Adaptive Channel Truncation with Selective Learning Strategy

QP-Adaptive Channel Truncation with Selective

Learning Strategy(QACT)^[3,4,5]는 MPEG 144차 회의에서 처음 기고되었으며, 148차 및 149차 회의에서 Core Experiment 4 : Feature conversion and adjustment(FCM CE 4) 관련 기술로 기고된 연구이다. QACT는 기존의 Feature reduction 모듈에서 사용하는 Feature fusion and Encoding Network(FENet)^[6]에 대해, Selective Learning Strategy(SLS)를 바탕으로 새롭게 학습된 FENet을 사용한다. SLS를 바탕으로 학습된 FENet이 적용된 Feature reduction 모듈에서, 압축된 특징 맵은 채널의 중요도 순서로 정렬된다.

이후 QACT는 정렬된 특징 맵의 상위 n 개 채널을 제외한 나머지 채널은 사용하는 비트에 비해 성능에 큰 영향을 주지 않는다고 판단하여 삭제하는데, 이를 통해 비트레이트를 절감하게 된다. 이때 n 은 영상마다 지정된 양자화 정도인 Quantization Parameter(QP)마다 다르며, OpenImage 데이터셋에선 [320, 160, 80, 40]이며 SFU, TVD, HiEve 데이터셋에선 [320, 240, 180, 135]이다. n 의 범위가 320부터 시작하는 이유는 QACT가 기고되었을 당시에는 Feature reduction 모듈의 출력 특징 맵이 320개 채널이었기 때문이다. 현재 FCTM-v6.1은 LightFCTM^[7]이 적용되었기 때문에 Feature reduction 모듈의 출력 특징 맵은 192개의 채널을 가진다.

QACT는 여러 채널로 이루어진 특징 맵에서 중요한 채널은 따로 존재한다는 사실에 기반해 중요하지 않은 채널을 삭제함으로써 비트레이트를 절감한다는 목표로 진행된 연구이다. 그러나 QACT는 삭제하는 채널의 수를 절대적으로 고정하기 때문에, 다양한 특징 맵에 대해 그 특성을 반영하지 못한다는 한계가 존재한다.

3. Feature Channel Rearranging and Truncation based on Importance

Feature Channel Rearranging and Truncation based on Importance(FCRT)^[8,9,10]는 MPEG 147차 회의에서 처음 기고되고, 이후 148차 회의에선 CE 4에 기고된 기술이다. FCRT는 QACT와 비슷하게, 채널의 중요성을 바탕으로 특징 맵을 정렬한 뒤 상위 256개의 채널만 남겨두고 나머지 채널은 삭제하는 기술이다.

FCRT에서의 중요도 기반 채널 정렬은 FENet에 적용된 Gain Unit^[11]을 바탕으로 진행된다. FENet에서 Gain Unit의 스케일 벡터는 320개의 스칼라값을 가지며, 채널들의 중요도를 학습하여 각 채널을 스케일링하게 된다. FCRT는 이러한 Gain Unit의 스케일 벡터를 이용해 각 채널을 해당하는 중요도 순으로 내림차순하여 특징 맵을 정렬한다. 정렬된 특징 맵은 Feature conversion 모듈에서 하나의 프레임으로 패킹했을 때, 정렬하지 않은 특징 맵을 패킹했을 때 보다 높은 공간적 중복성을 가진다. 이를 바탕으로 FCRT는 패킹된 프레임을 Versatile Video Coding(VVC)를 사용해 비트스트림으로 만들 때의 압축 효율을 향상시킨다.

채널의 중요도를 내림차순으로 정렬한 특징 맵에서, 인덱스가 낮은 채널은 상대적으로 중요한 정보를, 인덱스가 큰 채널은 중요하지 않은 정보를 담게 된다. 이에 FCRT에선 VVC에서의 코딩 효율성을 향상시킬 뿐만 아니라 절대적인 비트레이트를 감축하기 위해 상위 256개의 채널을 제외한 나머지 채널은 삭제하게 된다. 이때 삭제된 채널은 FCTM 디코더 측에서 복원되며, 모든 값을 0으로 채워 생성한 새로운 채널로 대체된다.

FCRT에서 채널을 삭제하는 알고리즘은 QACT와 마찬가지로, 특징 맵에서 중요한 채널은 별도로 존재한다는 아이디어를 바탕으로 구상된 기술이다. 이는 FCTM에서, 특징 맵에서 중요한 채널을 활용하는 것이 중요한 사실이라는 것을 시사한다.

4. Range-based Inactive Channel Removal

Range-based Inactive Channel Removal(RBICR)^[12,13,14]은 MPEG 147차 회의에서 FCM CE 4 분야에 처음 기고된 기술이며, 이후 148차 및 149차 회의에서 CE 4에 기고된 기술이다. RBICR은 특징 맵의 채널들을 활성 채널과 비활성 채널로 분류한 뒤, 비활성 채널을 삭제하여 비트레이트를 절감하는 기술로, 이때 비활성 채널은 각 채널의 Range, 즉 범위에 기반해 분류된다.

특징 맵 $x = \{x_i | i = 0, 1, \dots, n\}$ 에 대해 채널 x_i 의 최솟값을 $\min(x_i)$, 채널 x_i 의 최댓값을 $\max(x_i)$ 라고 했을 때 그 채널의 범위를 $range(x_i) = \max(x_i) - \min(x_i)$ 라고 한다. 특징 맵의 채널 개수가 n 일 때, 채널 범위의 집

합 $X = \{range(x_i) | i = 0, 1, \dots, n\}$ 에 대해 식 (1)의 *range_cut*이 RBICR에서 비활성 채널을 판별하는 기준이 된다.

$$range_cut = \frac{mean(X)}{1.5} \quad (1)$$

range_cut 미만의 $range(x_i)$ 를 가지는 채널 x_i 는 비활성 채널로 판별되어 삭제되고, 활성 채널만 남은 특징 맵이 FCTM 인코더의 나머지 부분을 통과한다. 이때 FCTM 인코더는 어떤 채널이 비활성 채널이었는지를 나타낸 이진 배열 *inactive_channel_arr*를 특징 맵과 함께 전송한다. 이후 삭제된 비활성 채널은 FCTM 디코더 측에서 특징 맵의 차원을 복구하기 위해 새롭게 생성된다. 이때, 삭제되었던 비활성 채널은 활성 채널로만 이루어진 텐서의 평균값으로 채워 생성한 채널로 복원하게 된다.

RBICR은 상위 n 개의 채널만 전송하는 QACT, FCRT와는 달리 *range*라는 기준을 바탕으로 채널의 수를 유동적으로 조절하기 때문에, 유의미한 정보를 가지는 채널들을 최대한 활용할 수 있다는 장점이 있다. 그러나 RBICR은 FCTM 디코더 측에서 비활성 채널을 복원하기 위해 *inactive_channel_arr*라는 추가적인 정보를 보냄에도 불구하고 이를 능동적으로 활용하지 않고 단순히 텐서의 평균값으로만 비활성 채널을 복원한다는 한계가 존재한다.

RBICR의 아이디어는 중요한 채널은 따로 존재하며, 나머지 채널은 삭제한 뒤 복원했을 때 비트레이트 대비 mAP와 같은 성능 지표에서 우수한 성능을 보인다는 것이다. 이는 QACT 및 FCRT의 기술적 의의와 유사하며, 세 기술이 가지는 공통된 의의는 특징 맵에서 정렬된 순서에서 상위 n 개 채널 안에 속하거나, 채널의 범위 값이 기준을 넘는 등 특정한 기준을 충족한 채널을 활용하는 것이 중요하다는 점을 시사한다.

5. QP-Adaptive Range-based Inactive Channel Removal with Selective Learning Strategy

QA-RBICR은 QACT와 RBICR의 융합 기술로, MPEG

149차 회의에서 FCM CE 4에 기고된 기술이다. QA-RBICR에서는 우선 SLS가 적용된 FENet을 통과한 특징 맵의 채널들이 중요도별로 정렬된다. 이후 정렬된 특징 맵에 대해, QP에 따라 비활성 채널의 기준을 설정한 뒤 비활성 채널이 삭제된다. 이때 채널의 활성 여부를 분류하는 기준은 표 1의 값을 활용한 식 (2)와 같다.

표 1. QP에 따른 QA-RBICR Threshold

Table 1. QA-RBICR Threshold according to QP

QP Range	QA-RBICR Threshold
$QP < 14$	12
$14 \leq QP < 19$	6
$19 \leq QP < 23$	3
$23 \leq QP$	1.5

$$range_cut = \frac{mean(X)}{QA-RBICR\ Threshold} \quad (2)$$

QA-RBICR은 객체 탐지, 객체 분할, 객체 추적에서 기존 FCTM-v5.0 대비 큰 성능 향상을 보였다. 이에 QA-RBICR은 현재 FCTM-v6.1에서 기본 설정으로 채택되어 사용되고 있다.

그러나 QA-RBICR은 비활성 채널을 복원하는 알고리즘이 RBICR과 동일하기에, 텐서의 평균값만을 활용한다는 RBICR의 한계를 그대로 답습했다는 문제가 있다. 이에 본 연구에서는 QA-RBICR에서 사용하는 비활성 채널 복원 알고리즘을 개량함으로써 성능 향상을 이루고자 하였다.

III. 제안 방법

기존의 QA-RBICR은 FCTM 디코더 측에서 비활성 채널을 복원할 시, 텐서의 평균값으로 채운 채널을 생성하여 비활성 채널을 복원하는 방법을 사용한다. 하지만 이는 활성 채널의 정보를 효율적으로 사용하지 않는다고 할 수 있다. 그림 1은 FCTM 디코더 측에서 QA-RBICR을 통해 비활성 채널이 어떤 방식으로 복원되는지를 나타낸다.

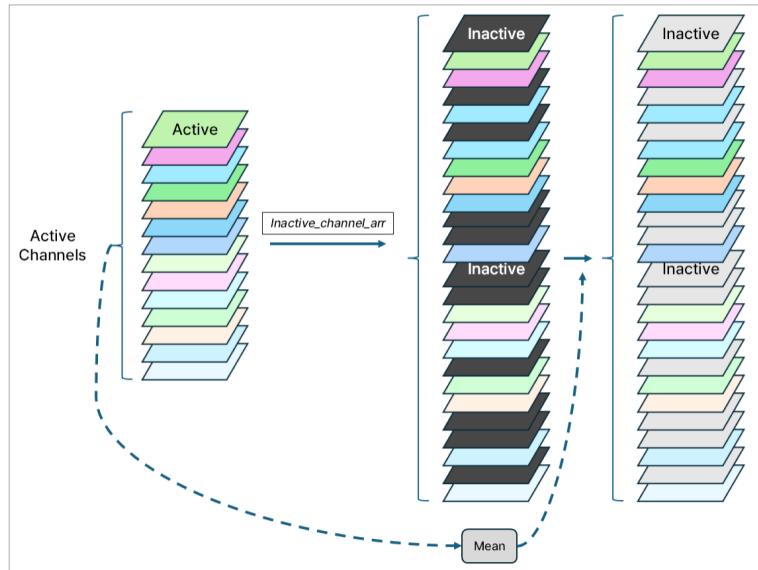


그림 1. QA-RBICR에서의 비활성 채널 복원 방법
 Fig. 1. Method of inactive channel restoration in QA-RBICR

FCTM의 디코더 측은 FCTM 인코더 측으로부터 활성 채널로 이루어진 텐서와 QA-RBICR을 적용하기 이전 특징 맵에서 어떤 채널이 비활성 채널이었는지를 0과 1로 나타낸 이진 배열 *inactive_channel_arr*를 받는다. 이를 바탕으로 그림 1과 같이 만약 활성 채널을 복원할 차례라면

텐서에 저장된 활성 채널을 그대로 사용하며, 비활성 채널을 복원할 차례라면 텐서의 평균값을 채워 새롭게 생성한 채널로 기존의 비활성 채널의 자리를 대체한다.

이때 QA-RBICR이 가지는 *inactive_channel_arr*라는 정보는 SLS가 적용된 FENet의 출력 특징 맵이 중요도

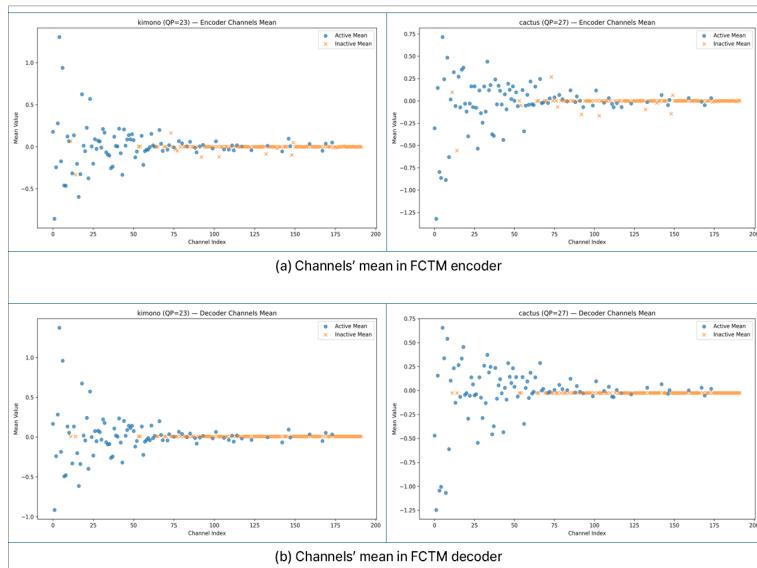


그림 2. FCTM 인코더 및 FCTM 디코더에서의 채널들의 평균값
 Fig. 2. Channel's mean values in FCTM encoder and FCTM decoder

순으로 정렬되어 나온다는 점에서 매우 유용한 정보이다. 왜냐하면 *inactive_channel_arr*에서 인덱스가 낮은 활성 채널은 중요도가 높다는 것을 의미하기 때문에, 해당하는 활성 채널을 더욱 적극적으로 활용할 수 있는 여지가 존재하기 때문이다.

그러나 기존 QA-RBICR은 이러한 유용한 정보를 가지고 있음에도 불구하고, 단순히 텐서의 평균값으로만 비활성 채널을 복원한다. 이는 보유한 정보를 제대로 활용하지 못한 복원 방법일뿐더러, 실제로 비활성 채널을 복원했을 때 기존 비활성 채널과 다르게 복원된다는 문제가 존재한다. 그림 2의 (a)는 FCTM 인코더 측에서 비활성 채널을 삭제하기 이전 채널들의 평균값을, (b)는 FCTM 디코더 측에서 비활성 채널을 복원했을 때 채널들의 평균값을 나타낸 산점도 그래프이다.

그림 2의 (a)에서 비활성 채널은 대체로 평균값이 0에 가까운 모습을 보이나, 몇몇 비활성 채널은 0 이외의 평균을 가지는 모습을 확인할 수 있다. 그러나 FCTM 디코더 측에서 텐서의 평균값으로 비활성 채널을 생성하여 복원한 (b)의 경우, (a)와는 달리 하나의 값으로만 비활성 채널이 구성되어 기존의 비활성 채널 분포와는 다른 획일화된 비활성

채널 분포를 띠게 된다.

이에 본 연구에선 주어진 활성 채널의 정보를 적극적으로 활용하여 비활성 채널을 복원하는 Effective Inactive Channel Restoration(EICR)을 제안한다. EICR은 기존 방식과는 달리, *inactive_channel_arr*을 보다 적극적으로 활용하여 비활성 채널을 복원한다. 그림 3은 EICR에서 비활성 채널이 어떤 식으로 복원되는지를 나타낸다.

그림 3에서, EICR은 비활성 채널을 복원할 때 QA-RBICR과는 달리 활성 채널을 직접적으로 활용한다. 기존 방식이 단순히 활성 채널로 구성된 텐서의 평균값으로만 비활성 채널을 복원하였다면, EICR은 가까운 인덱스의 활성 채널의 값을 가져와 사용하는 등 활성 채널 그 자체를 직접 활용한다. 이는 QACT, FCRT, RBICR이 가지는 기술적 의의인 “특정 기준을 충족한 채널을 적극적으로 활용하는 것이 중요하다”를 수용한 것이다.

이때 활성 채널을 활용하는 방법은 크게 네 가지 방식이 있으며, 각각 Duplicate 방식, Interpolate 방식, Mean duplicate 방식, Mean interpolate 방식이다. 각 방식이 비활성 채널을 복원하는 과정은 그림 4와 같다.

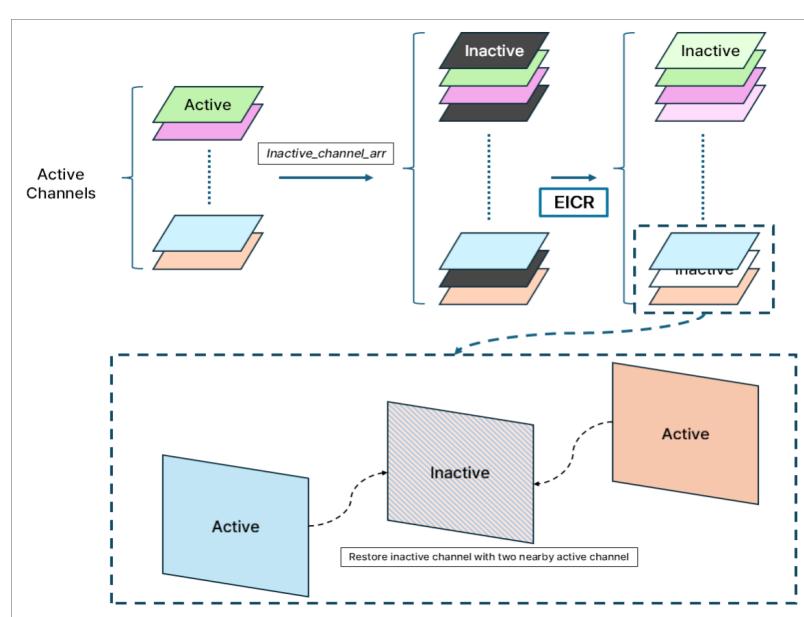


그림 3. EICR에서의 비활성 채널 복원 방법
Fig. 3. Method of inactive channel restoration in EICR

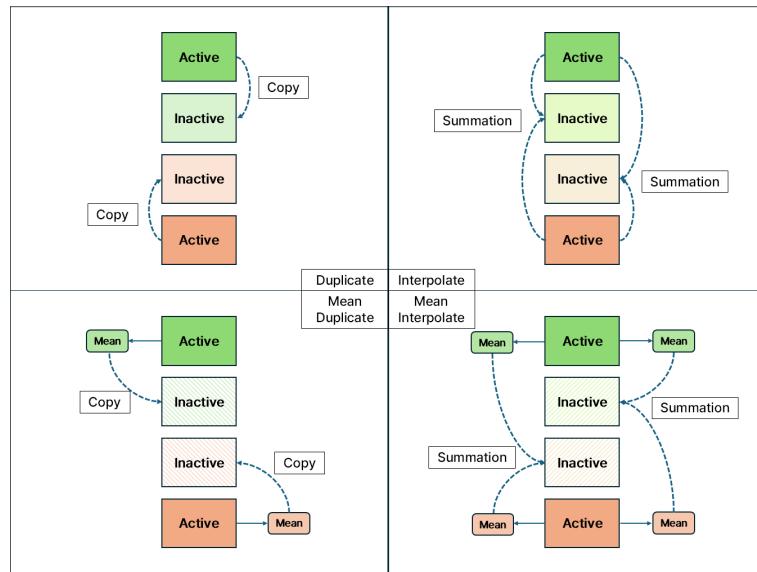


그림 4. EICR의 네 가지 방식
 Fig. 4. Four methods of EICR

EICR은 비활성 채널을 복원할 때 *inactive_channel_arr*을 활용한다. 비활성 채널 x_i 기준 가장 가까운 활성 채널들을 활용하는데, 이때 복원할 비활성 채널의 인덱스 기준 인덱스가 낮은 채널 $x_l (0 \leq l < i)$ 중 가장 가까운 활성 채널을 *left_channel*, 인덱스가 큰 채널 $x_r (i < r \leq 192)$ 중 가장 가까운 활성 채널을 *right_channel*이라 하자. 또 i 와 *left_channel*의 인덱스의 차이를 *ld*, i 와 *right_channel*의 인덱스의 차이를 *rd*라고 하자. 이를 바탕으로 식 (3), (4)를 계산할 수 있다.

$$left_w = 1 / (ld + 1) \quad (3)$$

$$right_w = 1 / (rd + 1) \quad (4)$$

Duplicate 방식은 가장 가까운 활성 채널을 복제하여 삭제된 비활성 채널을 복원하는 방식이다. 비활성 채널 x_i 기준 가장 가까운 활성 채널인 *left_channel*과 *right_channel* 중 더 가까운 활성 채널을 선택한 뒤 이를 복제하고, *ld* 또는 *rd*를 활용해 채널의 값들을 줄여 비활성 채널을 복원한다.

이때 채널의 값들을 줄이는 이유는, 단순히 채널을 복제하는 경우 특정 맵의 다양성이 줄어들기 때문이다. 식 (3), (4), (5)은 **Duplicate** 방식에서 복원되는 비활성 채널의 생성 과정이다.

$$duplicate = \begin{cases} if(left_w \leq right_w) : left_channel \times left_w \\ if(left_w > right_w) : right_channel \times right_w \end{cases} \quad (5)$$

Interpolate 방식은 가장 가까운 활성 채널 2개를 활용하여 새롭게 채널을 생성하여 비활성 채널을 복원하는 방식이다. 비활성 채널 x_i 기준 가장 가까운 활성 채널인 *left_channel*과 *right_channel*에 대해 *left_channel*에는 *ld*를, *right_channel*에는 *rd*를 활용해 채널의 값들을 줄인다. 이후 두 채널 *left_channel*과 *right_channel*을 더해 새로운 채널 *interpolate*를 생성한 뒤, *ld*와 *rd*를 활용해 채널의 값을 키운다. 식 (3), (4), (6)은 **Interpolate** 방식에서의 비활성 채널 생성 과정이다.

Mean duplicate 방식은 가장 가까운 활성 채널의 평균값으로 새롭게 채널을 생성하여 비활성 채널을 복원하는 방

$$interpolate = \frac{(left_w \times left_channel) + (right_w \times right_channel)}{(left_w + right_w)} \quad (6)$$

$$\text{mean_duplicate} = \begin{cases} \text{mean}(\text{left_channel}) \times \text{left_w} & \text{if } (\text{left_w} \leq \text{right_w}) \\ \text{mean}(\text{right_channel}) \times \text{right_w} & \text{if } (\text{left_w} > \text{right_w}) \end{cases} \quad (7)$$

$$\text{mean_interpolate} = \frac{(\text{left_w} \times \text{mean}(\text{left_channel}) + (\text{right_w} \times \text{mean}(\text{right_channel})))}{(\text{left_w} + \text{right_w})} \quad (8)$$

식이다. 이 방식에선 Duplicate 방식과 동일하게 가장 가까운 채널 *left_channel* 또는 *right_channel*을 선정한 후, 그 활성 채널의 평균값을 계산한다. 가장 가까운 활성 채널의 평균값을 *mean_duplicate*이라 했을 때, *ld* 또는 *rd*를 활용해 값을 줄인다. 식 (3), (4), (7)은 *mean_duplicate*의 계산 과정이다.

Mean interpolate 방식은 가장 가까운 활성 채널 2개의 평균값을 활용해 비활성 채널을 복원하는 방식이다. 비활성 채널 x_i 기준 가까운 활성 채널인 *left_channel*과 *right_channel*에 대해, 두 채널의 평균값을 계산한다. 이후 *left_channel*의 평균값에는 *ld*를, *right_channel*의 평균값에는 *rd*를 통해 값의 크기를 줄인 뒤 두 평균값을 더한 값을 *mean_interpolate*를 얻는다. 이후 *ld*와 *rd*를 활용해 값을 키운다. 식 (3), (4), (8)은 *mean_interpolate*의 계산 과정이다.

Mean duplicate 방식과 Mean interpolate 방식의 경우 하나의 값을 *mean_duplicate* 및 *mean_interpolate*를 얻게 된다. 이 값을 바탕으로 채널 하나의 가로 길이를 *w*, 세로 길이를 *h*라 했을 때 $\text{np.full}(w, h, \text{mean_duplicate})$ 또는 $\text{np.full}(w, h, \text{mean_interpolate})$ 를 활용해 새로운 채널을 생성한다. 이렇게 새롭게 생성된 채널로 비활성 채널을 복원한다.

이러한 네 가지 EICR 방식은 기존의 QA-RBICR과는 다르게 SLS가 적용된 활성 채널과 *inactive_channel_arr*을 적극적으로 활용하여 비활성 채널을 복원하게 된다.

표 2. SFU 데이터셋에서의 EICR과 Anchor 및 Remote Inference 사이의 BD-rate 비교
Table 2. BD-rate comparison between EICR and Anchor and Remote Inference in SFU dataset

	Duplicate		Interpolate		Mean Duplicate		Mean Interpolate	
	vs Anchor	vs Remote Inference	vs Anchor	vs Remote Inference	vs Anchor	vs Remote Inference	vs Anchor	vs Remote Inference
Class A/B	3.37%	-49.59%	3.33%	-47.49%	1.10%	-49.84%	0.70%	-50.37%
Class C	-1.34%	-86.19%	-0.51%	-86.13%	-1.23%	-86.18%	-0.72%	-86.05%
Class D	-4.67%	-88.84%	-2.37%	-88.57%	-0.51%	-88.37%	-3.28%	-88.69%

IV. 실험 결과

사용한 QP와 영상 데이터셋은 FCM Common Test and Training Conditions^[15]를 바탕으로 설정하였다. 데이터셋은 객체 탐지를 평가하기 위한 SFU 데이터셋을 사용하였으며, FCTM 버전은 FCTM-v6.1, VTM 버전은 VTM-23.3을 사용하였다. 표 2는 EICR의 성능과 기준 방법론의 성능 (QA-RBICR, Remote Inference)을 BD-rate 관점에서 비교한 표이다. 이때 성능이 양수라면 같은 mAP를 달성하기 위해 더 많은 비트가 필요하다는 뜻으로 성능이 악화함을 의미하고, 이와 반대로 성능이 음수라면 같은 mAP를 달성하기 위해 더 적은 비트가 필요하다는 뜻으로 성능이 개선됨을 의미한다.

표 3, 표 4는 각각 EICR 방식들과 Anchor에서의 mAP를 기록한 표와, EICR 방식들에서의 mAP와 Anchor에서의 mAP 사이의 차이를 절대적 수치로 기재한 표이다. 실험을 진행한 결과, 표 2에서 전반적으로 Class A/B에서는 성능 저하가, Class C 및 Class D에서는 성능 향상이 발생하였음을 확인하였다. 특히 Duplicate 방식과 Mean interpolate 방식은 Class D에서의 BD-rate가 4.67%, 3.28% 향상되는 등 기준 방법론 및 다른 EICR 방식보다 우세하였다.

네 가지 방식 중 가장 유의미한 성능 향상을 나타낸 방식은 Mean interpolate 방식이었다. Class D에서 BD-rate를 3.28% 향상됨과 동시에, 다른 방식들과는 달리 Class A/B에서의 성능 저하 현상이 BD-rate 기준 약 0.70%로 약하게

표 3. SFU 영상별 Anchor와 EICR의 mAP

Table 3. mAP of Anchor and EICR in SFU dataset

Dataset	QP	BPP/kbps	mAP (Anchor)	mAP (Duplicate)	mAP (Interpolate)	mAP (Mean Duplicate)	mAP (Mean Interpolate)
Traffic	4	1188.087	64.811	64.639	64.575	64.800	64.703
	13	489.018	63.799	63.798	63.744	63.795	63.795
	16	348.836	61.810	61.367	61.601	61.649	61.646
	19	240.109	59.844	59.674	59.941	59.958	60.122
Kimono	13	266.554	77.904	77.941	77.973	77.904	77.904
	17	147.404	79.306	79.105	79.203	79.195	79.195
	19	109.777	76.315	75.854	76.373	76.315	76.759
	21	80.553	74.647	73.960	73.942	74.580	74.638
ParkScene	16	265.344	59.351	60.010	59.569	59.923	60.034
	19	171.805	58.746	58.710	58.401	58.665	58.719
	21	125.289	56.929	57.219	57.137	56.882	57.031
	24	70.935	53.410	53.528	53.597	53.424	53.415
Cactus	22	267.023	65.779	65.597	65.711	65.842	65.700
	24	163.208	65.902	65.519	65.706	65.766	65.870
	26	109.786	66.654	66.198	66.365	66.355	66.467
	28	70.928	64.462	64.753	64.975	64.567	64.372
BasketballDrive	14	830.977	52.426	52.427	52.514	52.386	52.361
	16	625.002	51.627	51.537	51.497	51.498	51.623
	19	378.111	49.585	49.384	49.563	49.509	49.455
	21	260.887	47.970	48.098	48.100	48.027	48.044
BQTerrace	4	1714.125	60.025	60.003	59.983	60.116	60.140
	13	490.910	55.865	55.901	55.890	55.918	55.898
	16	316.138	51.985	51.726	51.935	51.965	52.022
	19	197.559	45.350	45.219	45.464	45.314	45.318
BasketballDrill	7	2144.470	50.852	50.715	50.739	50.873	50.868
	15	828.511	49.801	49.770	49.779	49.806	49.734
	20	341.089	46.942	46.878	46.691	46.972	46.913
	23	183.229	42.947	42.909	42.665	42.762	42.851
BQMall	10	2257.660	70.775	70.899	70.978	70.845	70.825
	14	1358.039	69.901	69.968	69.934	69.977	70.037
	22	311.967	62.617	62.760	62.686	62.630	62.650
	25	154.393	56.396	56.482	56.607	56.569	56.723
PartyScene	10	1120.000	87.620	87.692	87.796	87.662	87.642
	15	541.064	84.972	85.015	85.122	84.984	85.019
	18	320.845	81.057	81.290	81.220	81.189	81.281
	22	146.849	75.024	74.799	74.558	74.857	75.079
RaceHorsesC	22	268.830	53.311	53.252	53.159	53.570	53.349
	23	209.871	53.626	53.919	53.692	54.086	53.844
	28	61.263	51.469	51.321	51.117	51.309	51.253
	31	32.699	48.566	48.855	48.990	48.795	48.723
BasketballPass	13	1771.332	36.616	36.688	36.718	36.651	36.773
	20	609.328	34.816	35.157	35.214	35.038	35.034
	23	327.740	31.294	31.674	31.525	31.484	31.696
	26	160.219	29.052	28.802	28.830	28.842	28.926
BQSquare	18	507.687	36.066	36.229	36.088	36.139	36.247
	20	349.213	35.704	36.181	36.222	35.598	35.809
	23	191.527	32.682	32.771	32.471	32.581	32.792
	26	100.223	27.186	27.326	27.259	27.221	27.139
BlowingBubbles	9	1565.769	83.715	83.761	83.772	83.783	83.738
	18	391.889	82.180	82.203	82.235	81.940	82.140
	21	218.049	81.067	81.082	81.236	80.987	80.942
	23	142.012	79.088	78.854	79.025	78.889	78.966
RaceHorsesD	21	298.641	50.617	51.019	50.579	50.890	50.615
	23	187.019	50.278	50.617	50.642	50.678	50.556
	28	54.761	48.141	48.208	47.938	47.986	47.832
	31	29.258	44.680	44.910	45.257	44.672	45.345

표 4. SFU 영상별 Anchor와 EICR의 mAP 차이

Table 4. mAP differences of Anchor and EICR in SFU dataset

Dataset	QP	mAP (Anchor)	mAP differences in Duplicate	mAP differences in Interpolate	mAP differences in Mean Duplicate	mAP differences in Mean Interpolate
Traffic	4	64.811	-0.172	-0.236	-0.011	-0.108
	13	63.799	-0.002	-0.056	-0.005	-0.004
	16	61.810	-0.443	-0.209	-0.161	-0.164
	19	59.844	-0.170	0.096	0.114	0.277
Kimono	13	77.904	0.038	0.069	0.000	0.000
	17	79.306	-0.201	-0.103	-0.111	-0.111
	19	76.315	-0.461	0.058	0.000	0.444
	21	74.647	-0.687	-0.705	-0.067	-0.009
ParkScene	16	59.351	0.659	0.218	0.572	0.683
	19	58.746	-0.036	-0.345	-0.081	-0.027
	21	56.929	0.290	0.208	-0.047	0.102
	24	53.410	0.118	0.187	0.014	0.005
Cactus	22	65.779	-0.182	-0.068	0.063	-0.079
	24	65.902	-0.383	-0.196	-0.136	-0.032
	26	66.654	-0.456	-0.288	-0.299	-0.187
	28	64.462	0.291	0.513	0.105	-0.090
BasketballDrive	14	52.426	0.002	0.089	-0.040	-0.065
	16	51.627	-0.090	-0.130	-0.129	-0.004
	19	49.585	-0.201	-0.023	-0.077	-0.130
	21	47.970	0.129	0.131	0.058	0.075
BQTerrace	4	60.025	-0.022	-0.042	0.091	0.115
	13	55.865	0.036	0.025	0.053	0.033
	16	51.985	-0.259	-0.050	-0.020	0.037
	19	45.350	-0.131	0.114	-0.036	-0.033
BasketballDrill	7	50.852	-0.136	-0.112	0.021	0.016
	15	49.801	-0.031	-0.022	0.006	-0.067
	20	46.942	-0.064	-0.251	0.030	-0.029
	23	42.947	-0.038	-0.282	-0.185	-0.096
BQMall	10	70.775	0.124	0.202	0.069	0.049
	14	69.901	0.066	0.033	0.076	0.136
	22	62.617	0.144	0.070	0.013	0.033
	25	56.396	0.086	0.211	0.173	0.327
PartyScene	10	87.620	0.071	0.176	0.042	0.022
	15	84.972	0.043	0.150	0.012	0.047
	18	81.057	0.232	0.163	0.132	0.224
	22	75.024	-0.226	-0.466	-0.168	0.054
RaceHorsesC	22	53.311	-0.059	-0.152	0.259	0.038
	23	53.626	0.292	0.066	0.460	0.218
	28	51.469	-0.148	-0.352	-0.160	-0.215
	31	48.566	0.290	0.425	0.230	0.158
BasketballPass	13	36.616	0.073	0.103	0.035	0.157
	20	34.816	0.341	0.398	0.222	0.219
	23	31.294	0.380	0.231	0.190	0.402
	26	29.052	-0.250	-0.222	-0.209	-0.126
BQSquare	18	36.066	0.162	0.021	0.072	0.181
	20	35.704	0.478	0.519	-0.106	0.105
	23	32.682	0.089	-0.211	-0.102	0.109
	26	27.186	0.140	0.073	0.035	-0.047
BlowingBubbles	9	83.715	0.046	0.057	0.068	0.023
	18	82.180	0.023	0.055	-0.240	-0.040
	21	81.067	0.015	0.169	-0.080	-0.126
	23	79.088	-0.234	-0.063	-0.199	-0.122
RaceHorsesD	21	50.617	0.402	-0.038	0.273	-0.003
	23	50.278	0.339	0.364	0.399	0.277
	28	48.141	0.068	-0.202	-0.155	-0.309
	31	44.680	0.230	0.577	-0.008	0.666

나타나는 것을 확인할 수 있었다.

SFU 데이터셋에서 클래스마다 동일 비트레이트 대비 mAP의 평균값을 비교한 결과는 표 5와 같다. 이때 Mean interpolate 방식에서 BD-rate에서 성능이 열화된 것과 달리 Class A 및 Class B의 mAP 평균값이 향상되었음을 확인하였다.

표 5. EICR과 QA-RBICR의 mAP 비교

Table 5. mAP comparison between EICR and QA-RBICR

	Duplicate	Interpolate	Mean Duplicate	Mean Interpolate
Class A	-0.314%	-0.162%	-0.025%	+0.0003%
Class B	-0.127%	-0.028%	-0.007%	+0.060%
Class C	+0.066%	-0.014%	+0.102%	+0.093%
Class D	+0.294%	+0.234%	+0.025%	+0.175%

BD-rate에서 성능이 저하될 때 mAP 평균값에서의 성능이 향상되는 현상이 관측되는 이유는 BD-rate를 계산하는 식에는 비트레이트와 mAP가 함께 고려되기 때문이다. 실험에서, 큰 비트레이트에서 mAP가 낮은 경우가 존재하여 BD-rate에 큰 영향을 미치는 모습을 관찰할 수 있었다.

모든 EICR 방식에서, Class C, D에서 성능 향상을 이루 어낸 것과는 달리 Class A와 B에서는 의미 있는 성능 향상을 거두지 못했다. 특히 Class A의 경우 모든 방식에서 기존 방법론보다 성능이 저하되거나 의미 있는 성능 향상을 거두지 못했다. 이에 대한 원인을 분석하기 위해, mAP에 영향을 미칠 법한 변수들을 바탕으로 상관계수를 조사하였다. 표 6은 Mean Interpolate 방식에서, 변수별로 mAP에

표 6. Mean Interpolate에서의 mAP 차이에 대한 각 변수의 상관계수

Table 6. Correlation coefficients for each variable for mAP difference

	Resolution	Target	QP	Active
Correlation coefficient	-0.1216	-0.0744	0.039491	-0.05198

표 7. HiEve-1080p 데이터셋에서의 EICR과 Anchor 및 Remote Inference 사이의 BD-rate 비교

Table 7. BD-rate comparison between EICR and Anchor and Remote Inference in HiEve-1080p dataset

	Duplicate		Interpolate		Mean Duplicate		Mean Interpolate	
	vs Anchor	vs Remote Inference	vs Anchor	vs Remote Inference	vs Anchor	vs Remote Inference	vs Anchor	vs Remote Inference
HiEve-1080p	2.00%	-93.18%	2.11%	-93.20%	0.54%	-93.25%	-0.97%	-93.34%

끼치는 영향을 상관계수로 분석한 표이다.

Mean Interpolate에서 상관계수를 분석하였을 때, Resolution(해상도)과 Target(탐지 객체 수)에서의 상관계수가 다른 변수(QP, Active 채널 수)에 비해 높은 값을 가졌다. 즉, 해상도가 높고 탐지할 객체 수가 많은 경우 성능에 악영향을 미치는 정도가 높다는 것이다.

이를 바탕으로, Class A/B에서의 성능 열화는 데이터셋의 특성 때문에 발생한다고 판단하였다. Class A/B의 경우 각각 해상도가 2560x1600, 1920x1080으로, 해상도가 832x480, 416x240인 Class C, Class D보다 고해상도이며, 특히 Class A는 Class B, C, D의 평균 탐지 객체 수인 9.84의 약 10배인 100이라는 탐지 객체 수를 가지고 있다. 이러한 특성 때문에 Class A/B, 특히 Class A에서 EICR의 성능 열화가 발생한다고 추측하였다.

추가적으로, SFU 데이터셋 이외의 데이터셋에서의 성능을 평가하고자 HiEve-1080p 데이터셋에서 EICR 방식들의 성능을 측정하였다. 표 7은 HiEve-1080p에서의 EICR 방식들과 Anchor 및 Remote Inference 사이의 BD-rate를 비교한 표이다.

HiEve-1080p 데이터셋에서도 SFU 데이터셋에서와 마찬가지로 Mean Interpolate 방식에서 성능 향상을 발견할 수 있었다. 표 8, 표 9은 각각 EICR 방식들과 Anchor에서의 MOTA를 기록한 표와, EICR 방식들에서의 MOTA와 Anchor에서의 MOTA 사이의 차이를 절대적 수치로 기재한 표이다.

HiEve-1080p에서, Duplicate와 Interpolate, Mean Duplicate의 경우 성능이 악화하였으나 Mean Interpolate은 미미하지만 성능 향상을 발견할 수 있었다. 하지만 본 연구를 진행한 실험 환경에서 HiEve-720p, TVD 데이터셋에서의 실험 결과를 도출해 낼 수 없었기에, 객체 추적에 대한 성능을 일반화하지는 못하였다.

표 8. HiEve-1080p 영상별 Anchor와 EICR의 MOTA
Table 8. MOTA of Anchor and EICR in HiEve-1080p dataset

Dataset	QP	BPP/kbps	MOTA (Anchor)	MOTA (Duplicate)	MOTA (Interpolate)	MOTA (Mean Duplicate)	MOTA (Mean Interpolate)
HIEVE-13	12	313.825	61.98	61.67	61.48	61.64	61.63
	20	91.295	58.28	58.14	57.97	58.53	58.51
	23	53.476	54.59	54.73	54.67	55.01	55.06
	26	32.171	47.32	47.64	47.72	47.60	47.69
HIEVE-16	12	1263.605	21.02	21.12	21.01	20.84	20.97
	20	390.447	20.49	20.34	20.51	20.38	20.42
	23	194.027	18.77	18.58	18.65	18.56	18.68
	26	85.637	14.45	14.53	14.50	14.53	14.47

표 9. HiEve-1080p 영상별 Anchor와 EICR의 MOTA 차이
Table 9. MOTA differences of Anchor and EICR methods in HiEve-1080p dataset

Dataset	QP	MOTA (Anchor)	MOTA differences in Duplicate	MOTA differences in Interpolate	MOTA differences in Mean Duplicate	MOTA differences in Mean Interpolate
HIEVE-13	4	61.97556	-0.310	-0.496	-0.333	-0.344
	13	58.27942	-0.140	-0.310	0.250	0.235
	16	54.58707	0.144	0.087	0.427	0.469
	19	47.31964	0.325	0.397	0.284	0.367
HIEVE-16	13	21.02408	0.099	-0.015	-0.187	-0.051
	17	20.48809	-0.147	0.024	-0.105	-0.072
	19	18.77032	-0.193	-0.123	-0.209	-0.094
	21	14.44732	0.081	0.057	0.079	0.024

V. 결 론

본 연구에서는 FCTM-v6.1에 적용된 기술인 QA-RBICR에 대해, 기존 기술과는 다른 방법으로 비활성 채널을 복원하여 BD-rate, mAP 등의 성능을 향상하고자 하였다. 텐서의 평균값으로 채워 만든 채널을 이용해 비활성 채널을 복원하는 기존 방법론과는 달리, 채널을 삭제하기 전의 텐서에서 어떤 채널이 활성 채널 또는 비활성 채널이었는지에 대한 정보를 바탕으로 비활성 채널을 복원하는 EICR을 제시하였다.

본 논문은 EICR의 네 가지 방식인 Duplicate, Interpolate, Mean duplicate, Mean interpolate에 대해 기존 Anchor 및 Remote Inference와의 성능 비교 실험을 진행하였으며, 평균적으로 성능이 우수하였던 Mean interpolate 방식의 경우 SFU 데이터셋 기준 Class A/B에서 BD-rate 기준 약 0.70%

의 성능 저하가 발생하였으나, Class C와 Class D에 대해서는 각각 0.72%, 3.28%의 BD-rate 향상을 얻었다. 또 각 클래스에서 mAP의 평균을 비교하였을 때는 Class A, B, C, D에서 각각 0.0003%, 0.06%, 0.093%, 0.175%의 mAP 평균값 향상을 얻었다. HiEve-1080p 데이터셋에서는 Mean Interpolate 방식이 0.97%의 BD-rate 향상을 얻을 수 있었다.

그러나 제시한 모든 방법론에 대해 Class A에서 의미 있는 성능 향상을 끌어내지 못한 것이 본 연구의 한계이다. 이 한계를 극복하기 위해선 EICR에서 고해상도 및 탐지할 객체의 수가 많다는 점이 왜 다른 번수들보다 악영향을 미치는 정도가 큰지 조사할 필요가 있으며, 이를 바탕으로 성능을 더 개선하기 위한 방법론을 강구할 필요가 있다. 또, SFU 데이터셋과 HiEve-1080p 이외의 데이터셋에서도 실험을 진행하여 일반적인 성능을 평가할 필요가 있다. 또

Duplicate, Interpolate 등 EICR의 방식들에서, 특징 맵에 대한 세밀한 분석에 기반하여 비활성 채널을 복원하지 않고 복사 및 보간과 일반적으로 사용하는 Scaling 방식을 사용하여 FCTM에 최적화된 알고리즘을 제시하지 못했다는 한계가 존재한다.

추후 앞서 말한 상관관계에 대한 조사 및 추가적인 기술 연구를 바탕으로 현재 제시한 방법론에서 더 발전한 기술을 개발할 예정이다. 또 EICR에서 비활성 채널을 복원할 때의 방법론과 Scaling 방식을 추가로 연구하여 보다 최적의 방안을 모색하고자 하며, MPEG 회의마다 FCTM에 추가 및 수정되는 사항을 계속 추적하며 기술 연구를 이어가고자 한다.

참 고 문 헌 (References)

- [1] “Algorithm description of FCTM”, ISO/IEC JTC 1/SC 29/WG 04 N0626, 2025.01.
- [2] Gyu-Woong Han, In-Keun Yoo, and Jae-Gon Kim (Korea Aerospace University), Se Yoon Jeong, Jooyoung Lee, and Younhee Kim (ETRI), Fabien Racapé, Hyomin Choi, Mateen Ulhaq (InterDigital), “[FCM] CE4.2.3a: Combination of QACT (CE4.2.1) and RBICR (CE4.2.2) with Selective Learning Strategy”, ISO/IEC JTC 1/SC 29/WG 4 m71203, 2025.01.
- [3] Yong-Uk Yoon, Gyuwoong Han, Jae-Gon Kim (Korea Aerospace University), Se Yoon Jeong, Jooyoung Lee, Younhee Kim, and Jungwon Kang (ETRI), “[FCVCM] Response to FCVCM CfP: Advanced Multi-Scale Feature Compression with QP-Adaptive Feature Channel Truncation”, ISO/IEC JTC 1/SC 29/WG 4 m65703, 2023.10.
- [4] Gyu-Woong Han and Jae-Gon Kim (Korea Aerospace University), Se Yoon Jeong, Jooyoung Lee, and Younhee Kim (ETRI), “[FCM] CE4-related: QP-Adaptive Channel Truncation with Selective Learning Strategy”, ISO/IEC JTC 1/SC 29/WG 4 m70152, 2024.11.
- [5] Gyu-Woong Han, In-Keun Yoo, Jae-Gon Kim (Korea Aerospace University), Se Yoon Jeong, Jooyoung Lee, and Younhee Kim (ETRI), “[FCM] CE4.2.1: QP-Adaptive Channel Truncation with Selective Learning Strategy”, ISO/IEC JTC 1/SC 29/WG 4 m71202, 2025.01.
- [6] Hyewon Jeong, Seunghwan Jang, Dalhong Lim, Hui Yong Kim (Kyung Hee University), Jooyoung Lee, Se Yoon Jeong, Younhee Kim (ETRI), “[FCM] CE 1.1.8. L-MSFC-v2 with fine-tuning”, ISO/IEC JTC 1/SC 29/WG 4 m66341, 2024.01.
- [7] Dalhong Lim, Hyewon Jeong, Hui Yong Kim (KHU), Jooyoung Lee, Se Yoon Jeong, Younhee Kim (ETRI), “[FCM] CE1.1.4: LightFCTM”, ISO/IEC JTC 1/SC 29/WG 4 m71247, 2025.01.
- [8] Heeji Han, Haechul Choi (Hanbat National University), Soon-heung Jung, Jin Young Lee, Sangwoon Kwak, Won-Sik Cheong, Hyon-Gon Choo (ETRI), “[FCM] Feature Channel Rearranging and Truncation based on Importance”, ISO/IEC JTC 1/SC 29/WG 4 m68939, 2024.07.
- [9] Heeji Han, Haechul Choi (Hanbat National University), Soon-heung Jung, Jin Young Lee, Sangwoon Kwak, Won-Sik Cheong, Hyon-Gon Choo (ETRI), “[FCM] CE4.2.2a report: Feature Channel Rearrangement and Truncation”, ISO/IEC JTC 1/SC 29/WG 4 m70190, 2024.11.
- [10] Heeji Han, Haechul Choi (Hanbat National University), Soon-heung Jung, Jin Young Lee, Sangwoon Kwak, Won-Sik Cheong, Hyon-Gon Choo (ETRI), “[FCM] CE4.2.2b report: Feature Channel Rearrangement”, ISO/IEC JTC 1/SC 29/WG 4 m70191, 2024.11.
- [11] Z. Cui, J. Wang, S. Gao, T. Guo, Y. Feng and B. Bai, “Asymmetric Gained Deep Image Compression With Continuous Rate Adaptation,” 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 2021.
doi: <https://doi.org/10.1109/CVPR46437.2021.01039>
- [12] Juan Merlos, Fabien Racapé, Hyomin Choi, Mateen Ulhaq (InterDigital), “[FCM] Range-based channel removal and frame packing post feature reduction”, ISO/IEC JTC 1/SC 29/WG 4 m68965, 2024.07.
- [13] Fabien Racapé, Juan Merlos, Hyomin Choi, Mateen Ulhaq (InterDigital), “[FCM] CE4.2.1: Range-based channel removal and frame packing post feature reduction”, ISO/IEC JTC 1/SC 29/WG 4 m70058, 2024.11.
- [14] Fabien Racapé, Hyomin Choi, Mateen Ulhaq (InterDigital), “[FCM] CE4.2.2: Range-based channel removal and frame packing post feature reduction”, ISO/IEC JTC 1/SC 29/WG 4 m71173, 2025.01.
- [15] “[Common Test and Training Conditions for FCM]”, ISO/IEC JTC 1/SC 29/WG 04 N0548, 2024.07.

저자소개

이동민



- 2020년 3월 ~ 현재 : 동아대학교 컴퓨터공학과 학사과정
- ORCID : <https://orcid.org/0009-0007-2009-8753>
- 주관심분야 : 영상 처리, 딥러닝, 컴퓨터비전

정원식



- 1992년 2월 : 경북대학교 전자공학과 (공학사)
- 1994년 2월 : 경북대학교 대학원 전자공학과 (공학석사)
- 2000년 2월 : 경북대학교 대학원 전자공학과 (공학박사)
- 2000년 5월 ~ 현재 : 한국전자통신연구원 책임연구원
- ORCID : <https://orcid.org/0000-0001-5430-2969>
- 주관심분야 : 이머시브 미디어 기술, 기계를 위한 영상 부호화, 딥러닝기반 신호처리, 멀티미디어 표준화

서정일



- 1994년 2월 : 경북대학교 공학사
- 1996년 2월 : 경북대학교 공학석사
- 2005년 8월 : 경북대학교 공학박사
- 1998년 3월 ~ 2000년 10월 : LG반도체 선임연구원
- 2000년 11월 ~ 2023년 2월 : 한국전자통신연구원 실감미디어연구실장
- 2023년 3월 ~ 현재 : 동아대학교 컴퓨터공학과 부교수
- ORCID : <https://orcid.org/0000-0001-5131-0939>
- 주관심분야 : 멀티미디어, 오디오/비디오 부호화, 딥러닝, 컴퓨터비전