

일반논문 (Regular Paper)

방송공학회논문지 제30권 제2호, 2025년 3월 (JBE Vol.30, No.2, March 2025)

<https://doi.org/10.5909/JBE.2025.30.2.198>

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

동적 메쉬 기반 비디오의 텍스쳐 맵 부호화를 위한 율 제어 방법

최 한 솔^{a)}, 심 동 규^{a)†}

Rate Control Method for Encoding Texture Map of Dynamic Mesh-based Video

Hansol Choi^{a)} and Donggyu Sim^{a)†}

요 약

본 논문에서는 동적 메쉬 기반 비디오의 텍스쳐 맵 압축 과정에서 비트가 제한될 때 시공간적으로 일관적인 화질을 유지하기 위한 텍스쳐 맵 부호화의 율 제어 방법을 제안한다. 제안하는 방법은 먼저 시간적 화질의 일관성을 향상시키기 위하여 텍스쳐 맵의 시간적 연속성에 따라서 꾹쳐 단위로 가변적인 목표 비트를 할당한다. 연속성의 유무가 서로 다른 꾹쳐가 유사한 화질로 압축될 수 있는 목표 비트를 영상의 비트율에 따라 도출하며, 각 비트율에서 실험적으로 추정된 평균 비트를 이용하여 이전 꾹쳐의 율-왜곡 성능을 참조하는 기준의 방법 대비 효율적인 비트 할당을 수행한다. 또한 제안하는 방법은 공간적 화질의 일관성을 향상시키기 위하여 CTU 수준의 율 제어 과정에서 의미적으로 유사한 CTU의 $R-\lambda$ 모델 파라미터를 참조한다. 제안하는 방법은 현재 표준화가 진행중인 동적 메쉬 압축 표준인 V-DMC의 참조 소프트웨어인 TMM v8.0에 구현되었으며, 시퀀스와 꾹쳐 수준에서 각각 평균 0.02%, 2.6%의 목표 비트 오차율을 보이며 3D 정점의 복원된 휘도 성분과 렌더링된 2D 이미지의 휘도 성분의 시간적 화질 변동으로서 각각 0.5dB와 0.4dB의 표준 편차를 달성한다.

Abstract

In this paper, we propose a rate control method for texture map of dynamic mesh-based video to maintain spatiotemporally consistent image quality under bit limit. The proposed method first allocates variable target bits to each picture according to the temporal continuity of the texture map to improve temporal image quality consistency. The picture-level target bits, which enable pictures with different continuities to be compressed with similar image quality, are derived under sequence-level bit limit, and the experimentally estimated average bits at each sequence-level target bit are used to allocate bits efficiently compared to the existing methods that refer to the rate-distortion performance of the previous picture. In addition, the proposed method refers to the $R-\lambda$ model parameters of semantically similar CTU in the CTU-level rate control process to improve spatial quality consistency. The proposed method was implemented in TMM v8.0, the reference software of V-DMC, and achieves an average bit error rate of 0.02% and 2.6% at the sequence and picture levels, respectively, and achieves standard deviations of 0.5dB and 0.4dB as temporal quality variations of reconstructed luminance of 3D points and of rendered 2D image, respectively.

Keyword : Mesh, Texture map, Compression, Rate control, V-DMC

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

3D 공간 데이터를 기반으로 실제와 같은 몰입형 경험을 제공하는 3D 실감 미디어는 게임, 영화 등 문화 산업을 넘어서, 원격 회의와 원격 의료와 같은 실생활의 편의성을 향상시키는 다양한 분야에 걸쳐 활용될 수 있도록 지속적인 연구가 진행되고 있다. 메쉬는 3D 공간 데이터를 표현하는 대표적인 포맷 중 하나로 3D 공간상의 정점과 정점 간의 연결 정보로 구성되어 객체의 형상과 표면을 표현한다. 포인트 클라우드와 복셀 대비 적은 양의 데이터로 물체의 표면을 자연스럽게 표현이 가능하다는 장점이 부각되어 전통적인 3D 모델링 응용뿐만 아니라 최신 AR/VR 응용에 이르기까지 가장 널리 사용되는 3D 공간 데이터의 포맷 중 하나로 자리 잡고 있다^[1,2]. 3D 실감 미디어가 실사와 비슷하게 표현될수록 사용자의 몰입도와 상호작용 정확도가 향상되어 현실감 있는 경험을 제공할 수 있다. 이를 위해서는 3D 공간상에서 기하학적인 구조와 시간적 변화의 정밀한 표현이 중요하지만 실제와 유사하게 3D 실감 미디어를 제작할 수록 필요로 하는 데이터량은 기하급수적으로 증가할 것이다. 대규모 메쉬 데이터의 거대한 데이터량은 네트워크 대역폭이나 저장 장치에 제약이 있는 환경에서의 작업에 제한을 발생시킨다. 이러한 한계를 극복하기 위하여 메쉬 데이터를 보다 효율적으로 저장 및 전송하기 위한 압축 기술에 대한 연구가 지속적으로 수행되고 있다. 메쉬를 압축하는 기술은 시간적 변형이 없는 정적 메쉬를 압축하는 기술을 시작으로 발전해왔다. 정적 메쉬 압축 기술은 정점 간 연결 정보를 그래프 알고리즘으로 압축한 뒤 연결 정보를 이용한 중복성 제거를 통해 정점의 3D 좌표와 정점 단위의 추가적인 속성 정보인 색상 정보, 법선 벡터, 반사율 등을 압축한다^[3,4,5,6,7]. 메쉬 데이터의 전송이 보편화됨에 따라서

사용자의 환경 제약을 고려하여 해상도 축면에서 점진적인 전송 및 복원이 가능하도록 기술이 확장되었다^[8,9]. 최근에는 2D 컨텐츠를 주로 사용하던 다양한 응용들에서 더욱 향상된 실감도를 제공하기 위해 메쉬 기반 3D 동적 컨텐츠의 활용이 고려되면서, 시간에 따라 변형되는 동적 메쉬를 효과적으로 압축하기 위한 방법이 활발하게 연구되고 있다^[10,11,12,13,14].

동적 메쉬를 효과적으로 압축할 수 있는 기술의 표준화를 위해 국제 표준화 단체인 Moving Picture Experts Group (MPEG)은 Video-based Dynamic Mesh Coding (V-DMC) ^[15]를 이름으로 2020년부터 동적 메쉬 압축 기술에 대한 표준화를 진행중이다. V-DMC는 먼저 메쉬를 해상도 관점에서 점진적인 방법으로 압축한 뒤 정점 또는 표면의 속성 정보를 압축한다. 메쉬의 다양한 속성 정보들 중에서 색상 정보는 3D 실감 미디어 응용에서 현실감을 제공하기 위한 중요한 역할을 하며, V-DMC는 색상 정보가 정점 단위의 속성으로서 존재하는 경우와 텍스쳐 맵의 형태로 존재하는 경우에 대한 압축 기술을 모두 포함한다. 텍스쳐 맵은 2D 평면으로 투영된 메쉬 표면의 색상 정보를 담고 있는 이미지로, 적은 개수의 정점으로도 복잡한 색상과 세부적인 표면 디테일을 효과적으로 표현할 수 있는 장점으로 다양한 실감 미디어 응용에서 주목받고 있다. 텍스쳐 맵은 기존의 2D 비디오 코덱^[16,17,18,19]을 이용하여 압축되며 코덱 내부의 시공간적 중복성 제거 기술을 통해서 동적 메쉬 기반 컨텐츠의 텍스쳐 맵이 효과적으로 압축된다. 동적 메쉬 기반 컨텐츠의 전송이 요구되는 실감 미디어 응용에서 사용자의 경험을 저하시키지 않으면서 안정적인 서비스를 제공하기 위해서는, 사용자의 네트워크 대역폭과 저장소 크기의 제약에 따라 적절하게 비트를 제어하면서 컨텐츠의 품질을 유지할 수 있는 율 제어 기술이 필수적이다. 동적 메쉬 기반 컨텐츠의 색상 정보가 텍스쳐 맵 형식일 경우, 고품질의 컨텐츠에서 일반적으로 색상 정보가 메쉬보다 훨씬 큰 해상도와 데이터 크기를 갖는다. 따라서 텍스쳐 맵은 동적 메쉬 기반 컨텐츠의 압축된 총 데이터량을 결정하는 주요 요인이 될 수 있으며, 전체 제한 비트량을 달성하기 위해서 텍스쳐 맵의 율 제어 기술의 성능이 중요한 역할을 할 수 있다.

V-DMC가 텍스쳐 맵을 압축하기 위해 사용하는 기존 2D 비디오 코덱의 인코더는 자연 영상에 최적화되어 설계된

a) 광운대학교 컴퓨터공학과(Department of Computer Engineering, Kwangwoon University)

‡ Corresponding Author : 심동규(Donggyu Sim)
E-mail: dgsim@kw.ac.kr
Tel: +82-2-940-5470

ORCID: <https://orcid.org/0000-0002-2794-9932>

※ 본 연구는 정보통신기획평가원의 ICT 혁신인재 4.0 사업(IITP-2025-RS-2022-00156215) 및 2024년도 광운대학교 우수연구자 지원 사업에 의한 연구결과로 개발한 기술입니다.

· Manuscript February 3, 2025; Revised March 16, 2025; Accepted March 17, 2025.

율 제어 기능을 지원한다^[20]. 기존 2D 비디오 코덱의 율 제어 기능은 제한된 비트 조건에서 복원 화질을 최대화하기 위해, 영상의 각 세부 단위에 대하여 자연 영상의 시간적 연속성을 가정할 경우 참조 관계에 따라 예상되는 정보량에 비례하는 비트율로 압축한다. 반면 텍스쳐 맵은 시간적 연속성 측면에서 자연 영상과 상이한 특성을 갖으며, 이는 본래 3D 공간상에서 색상 정보가 갖는 시간적 연속성이 2D 공간으로 투영되는 과정에서 다양한 정도로 손실될 수 있기 때문이다. 텍스쳐 맵을 압축하는 과정에서 기존의 율 제어를 수행할 경우, 텍스쳐 맵의 시간적 연속성 변화에 따른 정보량 변화를 예측하지 못하여 화질 관점에서 변동이 발생할 뿐만 아니라 제한된 비트율 조건을 만족하지 못하는 한계가 존재한다.

본 논문은 동적 메쉬 기반 비디오의 텍스쳐 맵을 제한된 비트를 사용하여 시공간적으로 일관된 화질로 압축하기 위한 율 제어 방법을 제안한다. 제안하는 방법은 영상에서의 시간적 연속 강도와 율-왜곡 성능의 비례하는 관계를 이용하여, 텍스쳐 맵의 시간적 연속 강도가 높은 경우에는 상대적으로 적은 비트를 할당하고 반대로 시간적 연속 강도가 낮은 경우에는 상대적으로 높은 비트를 할당함으로써 픽쳐 단위의 시간적으로 일정한 화질을 목표로 비트 할당을 수행한다. 픽쳐 단위의 시간적 연속성을 저복잡도로 판단하기 위해 텍스쳐 맵과 대응되는 복원된 메쉬 간 시간적 연속성의 관계를 이용하며, 복원된 메쉬의 시간적 변화 정도를 내포하는 메쉬 인코더의 정보를 지표로 텍스쳐 맵의 시간적 연속성을 판단한다. 텍스쳐 시간적 연속성은 일정 수준 이상인 경우와 연속성이 존재하지 않는 경우로 크게 나뉘며, 일정 수준 이상일 경우에도 연속의 정도는 시간에 따라 다양한 수준으로 변화할 수 있다. 픽쳐 단위의 가변적인 시간적 연속성에 따른 최적의 비트를 예측하기 위한 방법으로, 시간적 연속성의 유무에 따라 동일한 화질로 압축될 수 있는 각각의 평균적인 비트량을 시퀀스 목표 비트율에 따라 실험적으로 추정하였고, 시간적 연속성이 존재하는 경우 연속 강도에 반비례하게 비트 추정치를 증감한다. 각 픽쳐는 연속성 유무 및 연속 강도에 따라 도출된 비트 추정치에 비례하게 최종 목표 비트를 할당받는다. 제안하는 픽쳐 수준의 비트 할당 방법은 기존의 R- λ 모델에 기반하여

GOP 단위로 이전 픽쳐의 율-왜곡 성능을 참조하는 방법 대비, 텍스쳐 맵의 율-왜곡 성능의 변동에 강인한 비트 할당 효율성을 보인다. 픽쳐에 할당된 비트 제약 내에서 CTU 수준의 율 제어를 수행하기 위하여 기존의 방법처럼 CTU 단위의 R- λ 모델에 기반한 율 제어를 진행한다. CTU 수준에서 또한 가변적인 시간적 연속성에 따른 율-왜곡 성능의 공간적인 변화가 발생하며, CTU의 율-왜곡 성능을 예측하여 서로 유사한 화질로 압축될 수 있도록 비트를 할당한다. CTU의 압축 과정에서 공간 도메인에서의 블록 단위 예측 성능은 CTU의 율-왜곡 성능을 결정하는 중요한 요소이므로, 제안하는 방법은 현재 CTU와 예측 성능이 유사할 것으로 예상되는 CTU를 직전에 압축된 동일한 시간적 연속 수준을 갖는 픽쳐에서 탐색하여 R- λ 모델 파라미터를 참조한다. 참조 CTU 탐색 과정에서, 현재 CTU와 임의의 참조 CTU 각각에 3차원 공간상에서 유사한 위치의 메쉬 표면의 색상 정보가 투영되었을 수록 시공간적 연속성이 유사하며, 따라서 두 CTU의 예측 성능이 유사할 것으로 가정한다. 각 CTU에 색상 정보가 투영된 메쉬 표면의 중심 좌표를 기준으로 CTU를 선택하며, 메쉬 표면의 중심 좌표는 텍스쳐 맵 생성 과정에서 도출되는 UV 좌표를 이용하여 저복잡도로 계산한다. 제안하는 방법은 텍스쳐 맵의 가변적인 시간적 연속성에 따른 율-왜곡 성능의 변화를 예측하여 비트 오차율 관점에서 시퀀스와 픽쳐 레벨에서 각각 평균 0.02%, 2.6%의 오차율을 달성하였다. 또한 픽쳐와 CTU 수준에서의 효율적인 비트 분배를 수행하여 3D 정점과 렌더링된 2D 이미지의 복원된 휘도 성분에 대하여 각각 0.5dB, 0.4dB의 PNSR의 표준 편차를 보여 객관적 화질의 시간적 일관성이 향상됨을 보이고, 12.3%와 8.7%의 BD-rate 향상 결과를 통해 평균적인 화질 또한 향상됨을 보인다.

본 논문은 다음과 같이 구성된다. 2장에서는 동적 메쉬 압축 방법과 R- λ 모델 기반의 율 제어 방법을 설명한다. 3장에서는 제안하는 동적 메쉬 비디오의 텍스쳐 맵의 압축을 위한 율 제어 방법을 설명한다. 4장에서는 제안하는 율 제어 방법의 성능을 제시하며 5장에서는 결론을 맺는다.

II. 동적 메쉬 압축 방법과 R- λ 모델 기반 영상 울 제어 방법

1. V-DMC의 동적 메쉬 압축 방법

V-DMC는 정점의 좌표와 정점 간 연결 관계가 시간에 따라 변화할 수 있는 동적 메쉬를 효율적으로 압축하기 위해 그림 1의 구조를 기반으로 동적 메쉬를 압축한다. 메쉬 압축 과정의 주요 특징으로는, 메쉬의 점진적인 전송 및 복원을 지원하기 위하여 메쉬를 다수 개의 해상도 단계로 분리하여 압축하며, 각각의 시간적 변화를 효율적으로 압축하기 위하여 시간적으로 앞서 복원된 메쉬를 참조하여 압축한다. 또한 인코더는 메쉬의 압축률과 품질 간 균형을 유연하게 조절하기 위한 손실 압축을 수행할 수 있다. 메쉬를 손실 압축할 경우 복원된 메쉬의 표면 구조가 원본 메쉬와 달라질 수 있기 때문에, 텍스쳐 맵 등 메쉬의 추가 속성 정보를 압축할 때에는 복원된 메쉬에 대응되는 속성 정보를 원본 속성 정보로부터 생성하여 압축한다.

V-DMC의 인코더가 메쉬를 압축하는 구체적인 과정은 다음과 같이 수행된다. 먼저 메쉬를 저해상도 메쉬와 해상도를 단계적으로 향상시킬 수 있는 고주파 성분으로 분리한다. 저해상도 메쉬는 원본 메쉬의 정점 수를 감소시키고 감소된 정점 간 연결관계를 생성함으로써 도출된다. 단순

화 과정에서 손실된 고주파 성분은 저해상도 메쉬의 표면을 균등하게 세분화하여 도출된 정점과 원본 정점 간 좌표의 차이로 보상되며 V-DMC에서는 이를 변위 벡터라고 부른다. 고주파 성분을 도출하는 과정은 재귀적으로 수행되어 해상도를 점진적으로 향상시킬 수 있는 고주파 성분이 단계적으로 도출된다. 인코더에서 목표하는 압축률 및 품질의 균형에 따라서 원본 메쉬의 단순화 정도와 세분화 단계를 결정할 수 있다.

전처리 단계에서 도출된 저해상도 메쉬와 변위 벡터는 각각 다음 과정으로 압축된다. 저해상도 메쉬는 정적 부호화 방법과 시간적 참조 부호화 방법 중 한 가지 방법으로 무손실 압축된다. 현재 압축하고자 하는 저해상도 메쉬가 시간적으로 직전에 압축된 저해상도 메쉬와 정점 간 연결 관계가 유사하다고 판단될 경우 시간적 참조 부호화 방법을 수행하고, 반대의 경우에는 정적 부호화 방법을 수행한다. 정적 부호화 방법은 MPEG EdgeBreaker^[21], Google DRACO^[22]와 같은 기준에 정적 메쉬를 효율적으로 부호화하는 기술을 이용한다. 시간적 참조 부호화 방법은 시간적으로 직전에 압축되어 복원된 저해상도 메쉬와 정점 단위 움직임 벡터만을 압축하며 정점 간 연결 정보는 이전 저해상도 메쉬의 연결정보를 그대로 사용한다.

변위 벡터는 세분화된 메쉬의 정점 당 존재하며 저해상도 메쉬의 정점 대비 상대적으로 조밀하게 존재하고 공간적 중복성을 갖는다. 이러한 특징에 기반하여 변위 벡터에

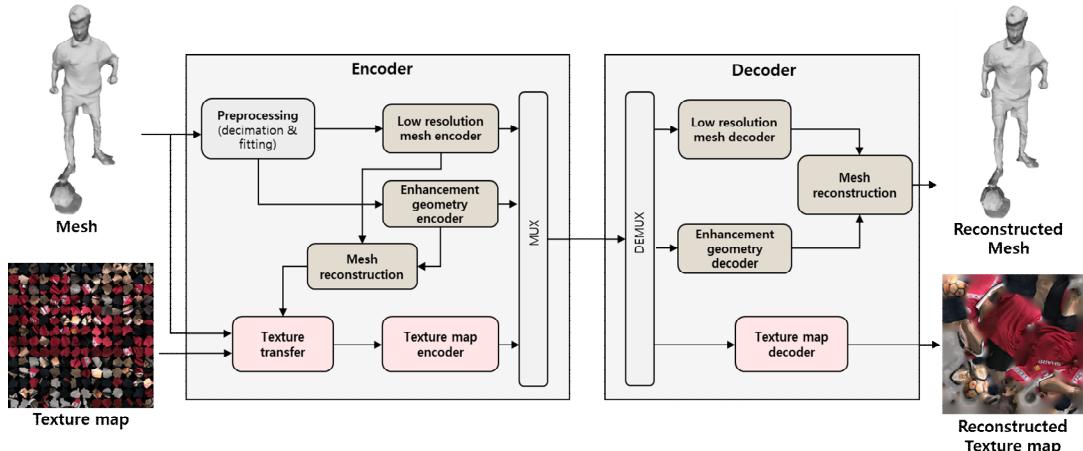


그림 1. V-DMC 인코더와 디코더 블록도
Fig. 1. Block diagram of V-DMC encoder and decoder

선형 웨이블릿 기반의 리프팅 변환을 수행하여 저주파 성분과 고주파 성분으로 분해하고 주파수 대역별 차등적 양자화를 수행하고, 2D 이미지로 패킹하여 기존 비디오 코덱을 이용하여 압축한다. 비디오 코덱 내부의 시간적 예측 방법을 통해서 시간적으로 인접한 메쉬의 변위 벡터 간 시간적 중복성을 효과적으로 제거할 수 있다. 변위 벡터의 압축률 및 화질의 균형 조절을 위해 비디오 코덱 외부 또는 내부의 양자화 단계에서 양자화 파라미터 값을 조정할 수 있다. 디코더에서는 저해상도 메쉬 복원 및 세분화하고, 변위 벡터를 복원하여 세분화된 각 정점에 합산하는 과정을 통해 메쉬를 복원한다.

메쉬의 손실 압축 과정에서 메쉬의 단순화, 정점 좌표 양자화, 변위 벡터 양자화를 원인으로 복원된 메쉬에는 원본 메쉬 대비 왜곡이 존재할 수 있다. 따라서 메쉬의 추가적인 색상 정보인 텍스쳐 맵을 압축하는 과정에서, 복원된 메쉬의 표면에 대응되는 색상을 도출하여 텍스쳐 맵을 재생성하는 과정이 실행된다. 텍스쳐 맵을 재생성하는 과정으로, 먼저 원본 메쉬 표면의 샘플링된 위치에 대응되는 색상 값을 원본 텍스쳐 맵으로부터 도출한다. 복원된 메쉬 표면의 각 샘플링된 위치의 색상 값을, 원본 메쉬 표면의 샘플링된 정점을 중 공간적으로 가까운 정점들의 색상 값을 이용하여 예측하고 UVatlas^[23] 방법을 이용하여 다시 텍스쳐 맵으로 패킹한다. UVatlas는 주어진 해상도의 텍스쳐 맵에 메쉬 표면의 색상 값을 투영하는 과정에서 발생할 수 있는 손실을 최소화하도록 메쉬 표면을 분할하여 패킹한다. 텍스쳐 맵의 압축률 및 화질의 균형을 조절하기 위해서는, 텍스쳐 맵을 재생성하는 과정에서의 텍스쳐 맵 해상도를 조정하거나 비디오 코덱 내부의 양자화 파라미터를 조정할 수 있다.

2. 영상 압축을 위한 R- λ 모델 기반의 율 제어 방법

전통적인 영상 압축 기술은 일반적으로 자연 영상을 효율적으로 압축하기 위한 목적으로 연구되었으며 뛰어난 압축 효율성을 바탕으로 점차 활용 범위가 확장되고 있다. 최근에는 동적 메쉬의 압축뿐만 아니라 포인트 클라우드 기반 비디오, 뉴럴 네트워크의 임의의 레이어의 출력 정보 등

의 다양한 종류의 데이터를 비디오 기반으로 압축하는 방법에서 전통적인 비디오 코덱이 활용되고 있으며, 각각의 데이터 압축에서 요구되는 압축 효율을 충족하는 데 중요한 역할을 하고 있다^[24,25].

율 제어 기술은 제한된 대역폭의 저장 매체나 네트워크 환경을 고려해야 하는 다양한 멀티미디어 응용에서 필수적으로 요구되는 핵심 기능으로 영상 압축 기술과 함께 발전해왔다. 영상 율 제어 방법은 영상의 비트를 결정하는 주요 요소를 무엇으로 간주하는지에 따라 대표적으로 R-Q 모델 기반 방법^[26], R-p 모델 기반 방법^[27], R- λ 모델 기반 방법^[20]으로 분류할 수 있다. 각 방법은 양자화 파라미터, 잔차 신호가 변환 및 양자화된 계수의 값, 비트율과 왜곡의 중요도 비율을 의미하는 λ 를 영상의 비트를 결정하는 주요 요소로 간주한다. 그 중 가장 높은 성능을 보이는 R- λ 모델 기반의 방법에 관련하여 다양한 방향으로 연구가 진행되어 왔으며^[20,28,29,30], HEVC 표준의 참조 소프트웨어^[31] 또한 R- λ 모델 기반의 율 제어 방법을 채택하였다. 본 절에서는 먼저 영상의 압축 원리와 율-왜곡 최적화 개념을 설명하고 이어서 R- λ 모델 기반의 율 제어 방법을 설명한다.

2.1 영상 압축 원리와 율-왜곡 최적화

영상 압축 방법은 시공간적 중복성 제거와 인지 화질 관점에서 중요도가 낮은 정보의 감축에 기초한다. 일반적으로 영상에서 장면이 변화함에 따른 픽셀 값의 변화는 공간적으로 가까운 픽셀들끼리 유사하며, 한 픽쳐 내에서도 공간적으로 가까운 픽셀들끼리 픽셀 값이 유사하다는 특성이 존재한다. 따라서 픽쳐를 블록 단위로 분할하여 예측, 변환, 양자화, 산술 코딩의 순서로 압축을 수행한다. 먼저 블록 단위로 시공간적 예측을 수행하여 신호의 중복성을 제거하여 잔차 신호만을 남긴다. 예측에 대한 잔차 신호에 대하여 인지적 시각 특성을 고려하여 주파수별 차등적 손실을 주위하여 주파수 변환 도메인에서 양자화를 수행한다. 양자화는 압축률과 왜곡을 결정하는 주요 단계이며 양자화 세기를 의미하는 양자화 파라미터의 값을 인코더에서 지정할 수 있다. 이후 양자화된 계수 값에 산술 부호화를 수행하여 비트스트림이 생성된다.

영상 압축의 효율과 처리 지연 간의 균형을 조절하기 위해 영상 내 프레임간 참조 제약을 여러 수준으로 두고자

All Intra (AI), Low Delay (LD), Random Access (RA) 환경이 지정되었다. AI 환경은 영상 내 모든 픽쳐를 독립적으로 압축하여 임의 접근성이 높지만 압축 효율은 상대적으로 낮다. LD와 RA 환경은 앞서 코딩된 픽쳐를 참조하여 일반 영상에서 AI 환경 대비 훨씬 높은 압축 효율을 보인다. LD 환경에서는 압축 효율을 높이면서 낮은 지연을 유지하기 위해 시간적으로 앞선 픽쳐만을 참조할 수 있다. RA 환경에서는 디스플레이 순서와 압축 순서를 달리하여 시간적으로 이후 픽쳐도 참조가 가능하여 세 가지 환경 중에서 가장 압축 효율이 높다.

앞서 설명한 손실 압축 과정을 통해 발생하는 화질 왜곡과 비트율은 서로 반비례 관계이기 때문에 주어진 영상에 대하여 압축률을 높일수록 화질 왜곡 또한 증가한다^[32]. 인코더에서는 가능한 모든 파라미터 조합에 대하여 비트율과 왜곡을 계산하고, 동일한 비트율일 경우에는 더 작은 왜곡을 도출하는 파라미터 조합을 선택한다. 따라서 영상의 울-왜곡 관계는 각 비트율에서 최소 왜곡인 경우들에 의해 정의되고 수식 1로 표현된다^[33]. 수식 1의 D 는 왜곡을, R 는 비트율을 의미하며, C 와 K 는 각 영상의 압축 성능을 의미한다.

$$D(R) = CR^{-K} \quad (1)$$

인코더에서는 영상의 울-왜곡 관계에 따라 선택 가능한 비트율과 왜곡 조합들 중에서 한 가지 조합을 선택하기 위하여 비트율과 왜곡 간 중요도 차이를 고려한다. 따라서 영상 인코딩은 비트율과 왜곡 간 중요도 차이를 반영한 비용 함수의 값이 최소가 되는 코딩 파라미터 조합을 결정하는 과정으로 정의되고 수식 2로 표현된다^[32]. 수식 2의 D 는 영상에 손실 압축을 수행할 때 발생하는 화질의 왜곡을 의미하며, R 은 영상을 압축한 후 비트율을 의미한다. 반비례 관계인 왜곡과 비트율을 종합적으로 고려하였을 때의 비용인 $(D + \lambda R)$ 을 손실 함수라고 하고, $\hat{\theta}$ 는 손실 함수의 값이 최소가 되도록 하는 인코딩 파라미터 조합이다. λ 는 왜곡 대비 비트율의 중요도를 의미하며 왜곡과 비트량 간 균형을 조절하는 역할이다. λ 값에 따라 임의 영상의 울-왜곡 관계에서 가능한 한 조합의 울, 왜곡이 결정되며, λ 의 값이 증가할수록 왜곡보다 비트율이 증가함에 따라서 손실 함수 값에 미치는 영향이 커지기 때문에 낮은 비트율로 압축된다.

그림 2는 영상에서 λ 와 비트율의 일반적인 반비례 관계를 시각화한다.

$$\hat{\theta} = \arg \min_{\theta} (D + \lambda R) \quad (2)$$

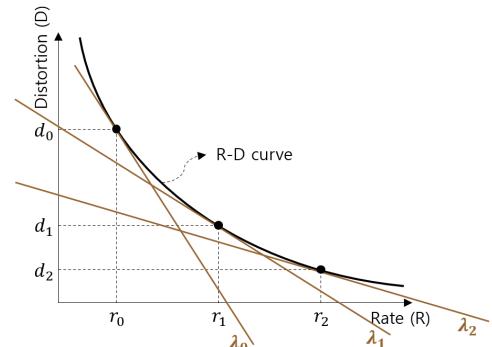


그림 2. 영상에서 λ 에 따른 따른 울과 왜곡
Fig 2. Rate and distortion according to λ

수식 1이 나타내는 울-왜곡 관계식은 λ 와 비트율의 반비례 관계를 나타내는 수식 3으로 정리될 수 있다. 수식 3의 α 와 β 는 수식 1의 C 와 K 로부터 도출되며 해당 영상의 울-왜곡 특성을 의미한다. 비트율과 λ 의 관계는 R- λ 모델 기반의 울-왜곡 제어 기술의 핵심적인 개념으로, 비트 할당과 비트율 제어 단계 모두에서 사용된다.

$$\lambda = -\frac{\Delta D}{\Delta R} = CKR^{-K-1} = \alpha R^{\beta} \quad (3)$$

λ 값에 의해 제어되는 비트량은 영상의 픽셀 값 잔차 신호에서 발생한 비트량과 영상 분할, 예측, 변환, 양자화 등의 각 압축 단계에서 전송되는 모든 파라미터에서 발생한 비트량을 포함한다. 영상의 전체 비트량을 구성하는 요소 중 잔차 신호는 가장 큰 비중을 차지하며, 잔차 신호의 비트량을 결정짓는 주요 요소는 양자화 단계의 양자화 파라미터이다. 따라서 각 λ 값에 대하여 최소의 비용을 달성하는 QP값은 비례 관계이며 [34]는 λ 와 QP의 관계를 수식 4로 모델링 하였다.

$$QP = 4.3281 \times \ln \lambda + 14.4329 \quad (4)$$

2.2 R- λ 모델 기반의 울 제어 수행 과정

R- λ 모델 기반 울 제어 방법은 영상의 R- λ 관계를 예측하여 영상을 목표하는 비트율로 압축하기 위한 λ 값을 결정한다. 이 과정에서 영상 평균 왜곡을 최소화하기 위하여 영상을 구성하는 각 세부 단위마다 최적화된 λ 값을 도출하며 이는 수식 5로 표현된다. 울 제어 기능은 크게 비트 할당 단계와 비트율 제어 단계로 구성된다. 비트 할당 단계는 평균 화질 최대화 관점에서 비트 사용을 최적화하기 위하여 영상의 총 할당 비트를 GOP, 픽쳐, CTU 등으로 계층적 분배를 수행한다. 비트율 제어 단계는 비트 할당 단계에서 결정된 비트율로 압축되기 위한 비용 함수의 λ 값을 결정을 수행한다.

$$\hat{\lambda} = \arg \min_{\lambda} \left(\sum_{i=1}^M d_i \right) \quad s.t. \quad \sum_{i=1}^M r_i \leq R \quad (5)$$

비트 할당 단계는 영상의 계층적인 인코딩 순서에 따라서 현재 인코딩하고자 하는 단위의 목표 비트를 상위 계층의 목표 비트로부터 할당한다. 먼저 한 GOP의 목표 비트량은 영상의 목표 비트량에서 분배된다. 인코더에서는 AI, LD, RA 중 참조 구조를 선택하고 GOP마다 참조 구조가 반복된다. 이에 따라 GOP 간 압축 성능을 동일하게 가정하고 GOP 내 픽쳐 개수에 비례하게 각 GOP의 비트를 최초 할당한다. 이후 각 GOP의 인코딩을 시작하기 이전에, 수식 6을 통해 이전 GOP까지 발생한 비트 달성을 오차의 일부를 보상하여 GOP의 최종 목표 비트를 도출한다. 수식 6의 $\text{AvgBit}_{\text{Pic}}$ 은 시퀀스 수준의 목표 비트를 시퀀스의 총 픽쳐 개수로 나눈 픽쳐의 평균 목표 비트량을 의미하고, $\text{PicNum}_{\text{Coded}}$ 와 $\text{ActualBit}_{\text{Coded}}$ 는 현재까지 인코딩이 수행된 픽쳐의 개수와 실제 발생 비트의 총 합을 의미하며, $\text{PicNum}_{\text{GOP}[i]}$ 는 GOP 내 픽쳐의 개수를 의미한다.

픽쳐 레벨의 비트 분배 단계에서는 GOP 내 픽쳐 참조 구조에 따라 GOP 평균 화질이 최대화될 수 있도록 각 픽쳐에 비트를 차등적으로 분배한다. 시간적으로 연속되는 자

연 영상에서 참조 구조 내 각 픽쳐의 참조 여건은 픽쳐의 압축 성능에 영상을 미치는 주된 요소이며, 다른 픽쳐들에 의해 참조되는 횟수는 GOP 전체 성능에 미치는 영향력을 결정한다. 따라서 두 요소를 기준으로 각 픽쳐에 할당할 목표 비트량을 결정한다. 이 때, 픽쳐의 울-왜곡 특성과 GOP 평균 화질에 미치는 영향력이 고려된 비트 수준이 참조 구조별로 각 픽쳐에 대하여 정의되어 있으며, 동일한 비트 수준의 픽쳐는 동일한 비율의 목표 비트량을 할당 받는다. 픽쳐의 비트 수준은 LD와 RA 각각에 대하여 GOP 크기별로 정의되어 있으며, 그림 3의 (a)와 (b)는 LD-P 환경에서 GOP 크기가 4, 8인 경우를, 그림 3의 (c)는 RA 환경에서 GOP 크기가 8이고 인트라 픽쳐의 간격이 16인 경우를 나타낸다. GOP의 참조 구조에 무관하게 인트라 프레임은 가장 높은 비트 수준인 0 값이 부여된다. LD-P 참조 환경에서는 각 픽쳐의 인코딩 과정에서 발생한 왜곡은 이후 픽쳐들

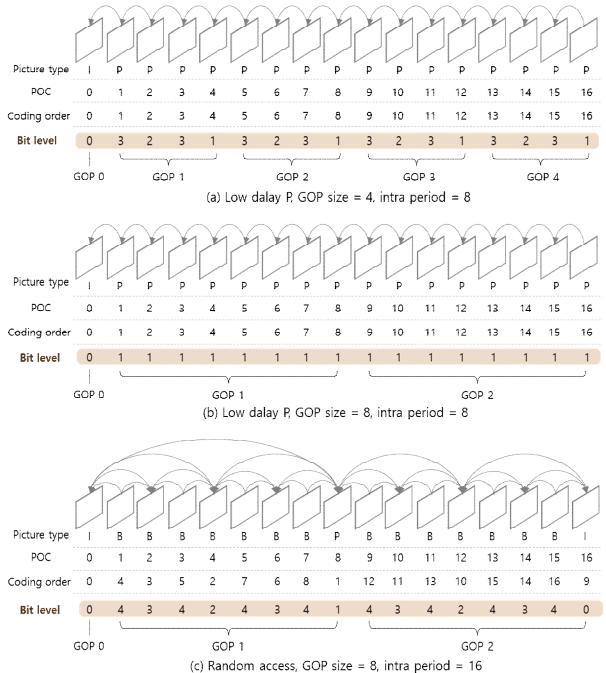


그림 3. GOP의 구조와 크기에 따른 픽쳐별 비트 수준

Fig. 3. Bit level of picture according to size and structure of GOP

$$\text{TargetBit}_{\text{GOP}[i]} = \left(\text{AvgBit}_{\text{Pic}} + \frac{\text{AvgBit}_{\text{Pic}} \times \text{PicNum}_{\text{Coded}} - \text{ActualBit}_{\text{Coded}}}{\text{SW}} \right) \times \text{PicNum}_{\text{GOP}[i]} \quad (6)$$

의 참조에 의해 순차적으로 전파된다. 이러한 순차적인 참조 구조에서 압축 왜곡 전파에 의한 평균 성능 저하를 최소화하기 위하여, GOP 크기가 4 픽쳐로 작을 때는 GOP 주기마다 비트 수준을 상대적으로 높게 설정한다. 계층적인 참조 구조를 갖는 RA 환경에서는 시간적 계층이 낮을수록 이후 픽쳐들에 의해 참조되는 횟수가 많기 때문에 높은 비트 수준을 설정한다.

위 과정을 통해 각 픽쳐에 설정된 비트 수준에 따라서, GOP 내 각 픽쳐의 초기 목표 비트와 초기 λ 값을 $R\cdot\lambda$ 모델을 이용하여 수식 7과 수식 8을 통해 계산한다. 초기 목표 비트와 초기 λ 값은 수식 8의 최적화 과정을 통해 결정되며, 수식 7을 통해 예상되는 각 픽쳐의 비트량 합이 GOP 목표 비트에 가장 근접하도록 하는 λ 값과 그 때 수식 7을 통해 도출된 비트량으로 결정된다. 수식 7에서 $PicLevel[j]$ 는 j번째 픽쳐의 비트 수준을, $InitTargetBit_{Pic[j]}$ 는 j번째 픽쳐의 초기 목표 비트를, $(\hat{\lambda}_{GOP[i]} \times \lambda Ratio_{PicLevel[j]})$ 는 비트 수준별 초기 λ 값을 의미한다. $\lambda Ratio_{PicLevel[j]}$ 는 GOP 평균 화질에 미치는 영향력에 기반하여 각 비트 수준에 따라 정의된 가중치 값이고, $\hat{\lambda}_{GOP[i]}$ 는 최적화 과정을 통해 결정되는 각 픽쳐의 공통 λ 값을 의미한다. $\alpha_{PicLevel[j]}$ 와 $\beta_{PicLevel[j]}$ 는 이전 GOP에서 비트 수준별로 적응적으로 업데이트된 픽쳐 수준의 $R\cdot\lambda$ 모델 파라미터 값이며, 각 비트

수준의 예상되는 율-왜곡 성능을 의미한다.

수식 9는 GOP 내 픽쳐 단위의 목표 비트를 최종적으로 결정하는 과정을 나타낸다. 현재 압축하고자 하는 픽쳐보다 앞서 압축된 픽쳐들에서 발생한 목표 비트에 대한 오차의 누적치를 일부 현재 픽쳐의 목표 비트에 반영하기 위하여, GOP의 남은 비트인 $BitLeft_{GOP[i]}$ 중에서 수식 7을 통해 도출한 초기 목표 비트량인 $InitTargetBit_{PicLevel[j]}$ 에 해당되는 비율로 최종 목표 비트가 결정된다.

GOP의 목표 비트를 각 픽쳐로 분배하는 과정과 동일한 목적과 원리로, 각 픽쳐의 목표 비트를 픽쳐 내 각 CTU로 분배한다. 각 CTU의 초기 목표 비트는 $R\cdot\lambda$ 모델을 이용하여 수식 10의 과정으로 도출되며, 초기 λ 값은 픽쳐의 목표 비트율로부터 도출된 λ 값을 사용한다. 수식 10의 $InitTargetBit_{CTU[k]}$, $\lambda_{Pic[j]}$, $PixelNum_{CTU[k]}$ 는 각각 픽쳐 내 k번째 CTU의 초기 목표 비트량, 초기 λ 값, 픽셀 개수를 의미한다. k번째 CTU의 예측된 율-왜곡 성능을 의미하는 파라미터 $\alpha_{PicLevel[j]CTU[k]}$, $\beta_{PicLevel[j]CTU[k]}$ 는 현재 픽쳐와 비트 수준이 같으면서 직전에 압축된 픽쳐의 k번째 CTU에서 업데이트된 파라미터를 참조한다. 이전 CTU 까지 압축을 수행하여 누적된 비트 오차를 남은 CTU들의 최종 목표 비트를 도출하는 과정에 반영하여 수식 11을 통해 각 CTU의 최종 목표 비트가 도출된다. 수식

$$InitTargetBit_{PicLevel[j]} = \left(\frac{\hat{\lambda}_{GOP[i]} \times \lambda Ratio_{PicLevel[j]}}{\alpha_{PicLevel[j]}} \right)^{\frac{1}{\beta_{PicLevel[j]}}} \times PixelNum_{PicLevel[j]} \quad (7)$$

$$\hat{\lambda}_{GOP[i]} = \arg \min_{\lambda_{GOP[i]}} \text{abs} \left(TargetBit_{GOP[i]} - \sum_{j=0}^{GOPsize-1} InitTargetBit_{Pic[j]} \right) \quad (8)$$

$$TargetBit_{Pic[j]} = BitLeft_{GOP[i]} \times \frac{InitTargetBit_{PicLevel[j]}}{InitTargetBitSum_{Pic[j:PicNum_{GOP[i]}-1]}} \quad (9)$$

$$InitTargetBit_{CTU[k]} = \left(\frac{\lambda_{Pic[j]}}{\alpha_{PicLevel[j]CTU[k]}} \right)^{\frac{1}{\beta_{PicLevel[j]CTU[k]}}} \times PixelNum_{CTU[k]} \quad (10)$$

$$TargetBit_{CTU[k]} = TargetBit_{Pic[j]} \times \left(\frac{InitTargetBit_{CTU[k]}}{InitTargetBitSum_{CTU[k:CTUNum_{Pic[j]}-1]}} \right) - \frac{(allocatedBitsLeft_{Pic[j]} - BitsLeft_{Pic[j]})}{SW} \quad (11)$$

11의 $TargetBit_{pic[j]}$ 는 현재 CTU가 포함된 픽쳐의 목표 비트를, $InitTargetBit_{CTU[k]}$ 는 현재 CTU에 할당된 초기 비트 값을, $InitTargetBitSum_{CTU[k:CTUNumPic[j]-1]}$ 는 현재 픽쳐에서 남은 CTU의 초기 비트 값의 합산 값을, $allocatedBitsLeft_{Pic[j]}$ 와 $BitsLeft_{Pic[j]}$ 는 현재 CTU를 압축하는 시점에서 당초 계획했던 픽쳐 수준의 잔여 비트와 실제 잔여 비트를, SW 는 비트 에러 보상률을 조정하기 위한 파라미터를 의미한다.

계층적 비트 할당의 최종 단인 CTU를 할당된 목표 비트로 압축하기 위한 인코딩 파라미터 λ 와 QP 값을 예측하는 과정을 수행하며 이를 비트율 제어라고 한다. 먼저 수식 12를 통해서 R- λ 모델을 사용하여 CTU 목표 비트율에 대응되는 λ 값을 예측한다. 수식 12의 $\alpha_{CTU[k]}$, $\beta_{CTU[k]}$ 는 수식 10의 $\alpha_{PicLevel[j]CTU[k]}$, $\beta_{PicLevel[j]CTU[k]}$ 와 동일하며, 이전 픽쳐로부터 예상한 현재 CTU의 율-왜곡 성능을 나타낸다. 도출된 λ 값으로부터 양자화 과정에서 사용할 QP 값을 수식 4를 통해서 도출한다.

$$\lambda_{CTU[k]} = \alpha_{CTU[k]} \times TargetBpp_{CTU[k]}^{\beta_{CTU[k]}} \quad (12)$$

비트율 제어 단계에서 도출한 λ 와 QP를 이용하여 각 CTU와 픽쳐의 인코딩을 수행한 후, 각 CTU와 픽쳐에서 발생한 실제 비트와 왜곡을 이용하여 R- λ 모델 파라미터를 업데이트 한다. 수식 13과 14는 비트 수준별로 픽쳐 수준의 R- λ 모델 파라미터를 업데이트하는 과정을 나타낸다. 수식 15와 16은 비트 수준별로 CTU 수준의 R- λ 모델 파라미터를 업데이트하는 과정이다. 수식 13~16에서 λ , $ActualBits$, $ActualDistortion$ 은 각각 현재 인코딩이 완료된 픽쳐 또는 CTU의 인코딩 과정에서 사용된 λ 의 값과 실제로 발생한 비트와 왜곡의 값을 의미한다.

$$\alpha_{PicLevel[j]} = \lambda_{Pic[j]} \times ActualBits_{Pic[j]}^{1 + \frac{\lambda_{Pic[j]} \times ActualBits_{Pic[j]}}{ActualDistortion_{Pic[j]}}} \quad (13)$$

$$\beta_{PicLevel[j]} = -\frac{\lambda_{Pic[j]} \times ActualBits_{Pic[j]}}{ActualDistortion_{Pic[j]}} - 1 \quad (14)$$

$$\alpha_{PicLevel[j]CTU[k]} = \lambda_{CTU[k]} \times ActualBit_{CTU[k]}^{1 + \frac{\lambda_{CTU[k]} \times ActualBit_{CTU[k]}}{ActualDistortion_{CTU[k]}}} \quad (15)$$

$$\beta_{PicLevel[j]CTU[k]} = -\frac{\lambda_{CTU[k]} \times ActualBit_{CTU[k]}}{ActualDistortion_{CTU[k]}} - 1 \quad (16)$$

HEVC는 앞서 설명한 블록 단위의 시공간적 중복성 제거 기반의 압축 과정을 통해 텍스쳐 맵을 전반적으로 효과적으로 압축한다. 하지만 텍스쳐 맵의 재생성 과정에서 메쉬 표면 세그먼트가 투영되는 2D 공간상의 위치가 시간적으로 달라질 수 있음에 따라 시간적 불연속이 발생할 수 있으며, 시간적 불연속이 발생한 영역에서는 시간적 예측이 효과적으로 동작하지 못하여 상대적으로 율-왜곡 성능이 낮다. 따라서 시간적 연속성이 존재하는 영역과 동일한 화질로 압축되기 위해 상대적으로 더 높은 비트가 필요하다. HEVC의 기존 율 제어 기술은 이러한 텍스쳐 맵이 갖는 율-왜곡 성능의 변동에 대한 고려를 하지 못하기 때문에, 텍스쳐 맵을 정해진 비트로 압축하는 과정에서 효율적인 비트 할당과 정확한 비트율 제어를 수행하지 못하는 한계가 존재한다.

III. 제안하는 텍스쳐 맵을 제어 방법

V-DMC의 인코더에서 메쉬를 손실 압축하는 경우 복원된 메쉬의 표면에 대응되는 색상 정보를 원본 색상 정보로부터 예측한 뒤 텍스쳐 맵으로 투영하여 압축한다. 복원된 메쉬 표면의 색상 정보를 텍스쳐 맵으로 투영하는 과정에서, 메쉬 표면의 구조가 시간적으로 변화함에 따라 색상 정보가 투영되는 텍스쳐 맵 상의 위치가 변화할 수 있다. 이로 인해 색상 정보가 본래 3차원 공간에서 갖던 시간적 연속성이 텍스쳐 맵으로 투영된 후 감소할 수 있으며, 메쉬 표면 구조의 시간적 변화가 일정 수준 이상으로 클 경우 텍스쳐 맵의 시간적 연속성이 급격하게 감소한다. 화면 간 예측이 수행 가능한 환경에서 텍스쳐 맵이 갖는 시간적 연속성의 변동은 율-왜곡 성능의 시간적 변동을 발생시키며, 텍스쳐 맵의 임의의 화질로 압축되기 위해 필요한 비트량은 일반

적으로 시간적 연속 정도와 반비례하는 관계를 보인다. HEVC의 기존 율 제어 기술은 자연 영상의 시간적 연속을 전제로 설계되었기 때문에 텍스쳐 맵의 시간적 연속성 변화에 따른 율-왜곡 성능의 변화를 예측할 수 없으므로, 텍스쳐 맵의 화질을 고려하면서 목표하는 비트로 압축하는데 한계가 있다.

이를 해결하기 위해서 본 논문에서는 텍스쳐 맵의 시간적 연속성 변화에 따른 율-왜곡 성능 변화를 고려하여 텍스쳐 맵을 시공간적으로 일관된 화질로 압축할 수 있는 율 제어 방법을 제안한다. 제안하는 율 제어 방법은 기존 R- λ 모델 기반의 율 제어 방법과 같이 크게 계층적 비트 할당 단계와 비트율 제어 단계로 구성된다. 그림 4는 V-DMC의 텍스쳐 맵 인코더 내부에서 제안하는 텍스쳐 맵 율 제어 방법의 전체적인 수행 과정을 나타내며, 율 제어 과정에서 사용하는 메쉬의 압축 정보를 함께 나타낸다. 먼저 그림 4의 텍스쳐 맵 인코더에 입력된 각 텍스쳐 맵에 대하여 제안하는 핵심 단위의 연속성 유무 판별 과정을 수행한다. 텍스쳐 맵의 시간적 연속성을 저복잡도로 판단하기 위한 방법으로, 텍스쳐 맵과 대응되는 메쉬 표면 구조의 시간적 연속

성의 연관성을 이용하여 텍스쳐 맵에 대응되는 저해상도 메쉬의 화면간 예측 수행 여부를 지표로 시간적 연속 유무를 판단한다. 계층적 비트 할당 단계에서는 먼저 그림 4의 제안하는 연속성 적응적인 핵심 단위 비트 할당 단계가 수행되며, 핵심 단위의 시간적 연속 유무에 무관하게 시간적으로 일정한 화질로 압축할 수 있는 핵심 단위의 목표 비트 할당을 수행한다. 시간적 연속성 유무에 따라 율-왜곡 성능에 차이가 존재하는 핵심들을 유사한 화질로 압축할 수 있는 목표 비트를 주어진 시퀀스 수준의 목표 비트율에 따라 도출한다. 임의의 시퀀스 수준의 목표 비트율에서 시간적 연속성이 존재하는 핵심과 존재하지 않는 핵심이 화질이 유사하기 위한 평균적인 비트량을 실험적으로 도출하였고, 도출된 평균 비트량에 비례하는 목표 비트를 할당한다. 시간적 화질 일관성을 더욱 향상시키기 위한 방법으로, 시간적 연속성이 존재하는 핵심의 경우 시간적 연속 정도에 반비례하게 목표 비트를 세부적으로 증감한다. 시간적 연속 정도를 저복잡도로 판단하기 위한 방법으로 대응되는 저해상도 메쉬의 모션 비트량과 시간적 연속 정도가 반비례함을 가정하고 시간적 연속 정도를 판단한다. 그림 4의 제안

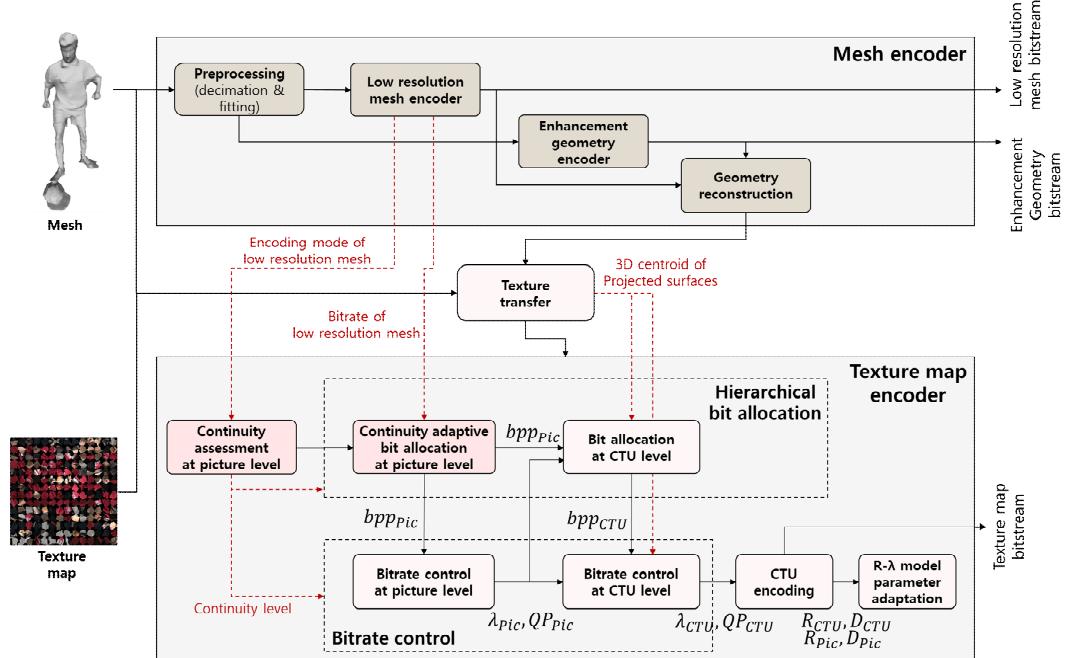


그림 4. 제안하는 텍스쳐 맵 율 제어 방법의 블록도
 Fig. 4. Block diagram of the proposed texture map rate control method

하는 CTU 수준 비트 할당 단계에서는, 픽쳐에 할당된 목표 비트가 픽쳐 내 각 CTU를 공간적으로 유사한 화질로 압축 할 수 있도록 분배되며 기존의 $R-\lambda$ 모델을 활용한다. 픽쳐에 시간적 연속성이 존재하지 않는 경우에 또한 CTU 단위로 정확한 $R-\lambda$ 모델 파라미터 예측을 수행하고자 제안하는 율-왜곡 성능 예측 방법을 수행한다. 제안하는 CTU 단위의 $R-\lambda$ 모델 파라미터 예측 방법은 현재 CTU와 율-왜곡 성능이 유사할 것으로 예상되는 참조 CTU를 선택하고자, 현재 픽쳐와 시간적 연속 유무가 동일한 픽쳐의 참조 CTU들 중에서 투영되어있는 메쉬 표면의 3D 공간상에서의 중심 좌표가 현재 CTU와 가장 유사한 CTU를 참조한다. 그림 4의 CTU 수준 비트율 제어 단계에서 또한 제안하는 방법으로 예측된 CTU의 $R-\lambda$ 모델 파라미터를 사용하여 보다 정확한 λ 와 QP 값의 예측을 수행한다. 이후 수행되는 CTU 단위의 인코딩 과정과 $R-\lambda$ 모델 파라미터 업데이트 과정은 기존의 방법과 동일한 과정으로 수행된다.

1. 텍스쳐 맵과 복원된 메쉬의 시간적 연속성 관계

텍스쳐 맵의 시간적 연속성은 메쉬의 표면 구조를 기준으로 텍스쳐 맵 상에 색상 정보를 투영할 위치를 결정하는 과정에서 결정된다. 인코더에 입력된 텍스쳐 맵은 함께 입력된 원본 메쉬에 대응되기 때문에, 손실 압축을 통한 왜곡이 존재하는 메쉬 표면에 대응되는 색상 정보로 변경하여 압축을 수행한다. 이 때 텍스쳐 맵 재생성 과정에 따라서, 복원된 메쉬 표면의 색상을 UVAtlas 알고리즘을 이용하여 텍스쳐 맵으로 패킹한다. 패킹 과정에서 색상 정보의 손실이 최소화되도록 메쉬의 표면을 다수 개의 세그먼트로 분할하여 이미지로 투영하기 때문에, 3D 공간에서 색상 정보가 갖는 시공간적 연속성이 텍스쳐 맵 도메인에서 유지되지 않을 수 있다. 결론적으로 각 색상 정보가 패킹되는 퍼셀 위치는 메쉬 표면의 구조에 의존적으로 결정되며, 두 메쉬 프레임의 복원된 메쉬 표면의 구조가 유사할수록 패킹된 텍스쳐 맵에서 유사도 또한 높다.

두 메쉬 프레임의 복원된 메쉬의 표면 구조 유사도는 저해상도 메쉬의 표면 구조 유사도와 유관하며, 이는 저해상도 메쉬를 세분화하고 고주파 성분을 더하는 과정을 통해 복원된 메쉬가 도출되기 때문이다. 두 메쉬 프레임의 저해

상도 메쉬의 표면 구조 유사도는 저해상도 메쉬의 시간적 예측 수행 여부를 지표로 판단이 가능하다. 저해상도 메쉬를 압축하는 단계에서, 저해상도 메쉬의 표면 구조가 이전 프레임과 일정 수준 이상 유사할 경우에 시간적 예측이 선택되기 때문이다. 그림 5는 저해상도 메쉬의 압축 과정에서 정적 예측이 수행된 경우와 화면 간 예측이 수행된 각각의 경우에, 메쉬에 대응되는 텍스쳐 맵의 직전 프레임과의 유사도를 비교한다. 그림 5의 T번째 메쉬는 저해상도 메쉬에 정적 예측이 수행된 경우로, T번째와 T-1번째 텍스쳐 맵의 유사도가 적은 것을 보인다. 반면 그림 5의 T+1번째 메쉬는 저해상도 메쉬의 압축 과정에서 T번째 저해상도 메쉬를 참조한 경우로, 결과적으로 T+1번째와 T번째의 텍스쳐 맵의 유사도가 일정 수준 이상으로 높은 것을 보인다.

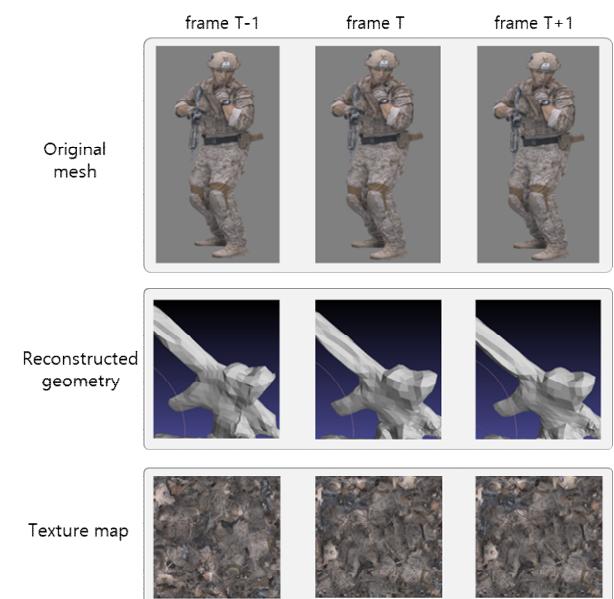


그림 5. 복원된 메쉬 표면의 구조 유사도에 따른 텍스쳐 맵 유사도 비교
Fig. 5. Comparison of texture map similarity and the structural similarity of the reconstructed mesh surface

저해상도 메쉬가 시간적 예측을 수행한, 인접한 텍스쳐 맵의 유사도가 일정수준 이상으로 높은 경우에, 메쉬 정점 단위의 움직임 복잡도에 따라 텍스쳐 맵 유사도의 세부적인 수준이 결정된다. 이는 정점 단위의 움직임이 복잡할수록 메쉬의 표면 구조에 차이가 증가하며 텍스쳐 맵 패킹 결과에 변화가 커진다. 메쉬 정점 단위의 움직임 복잡도는

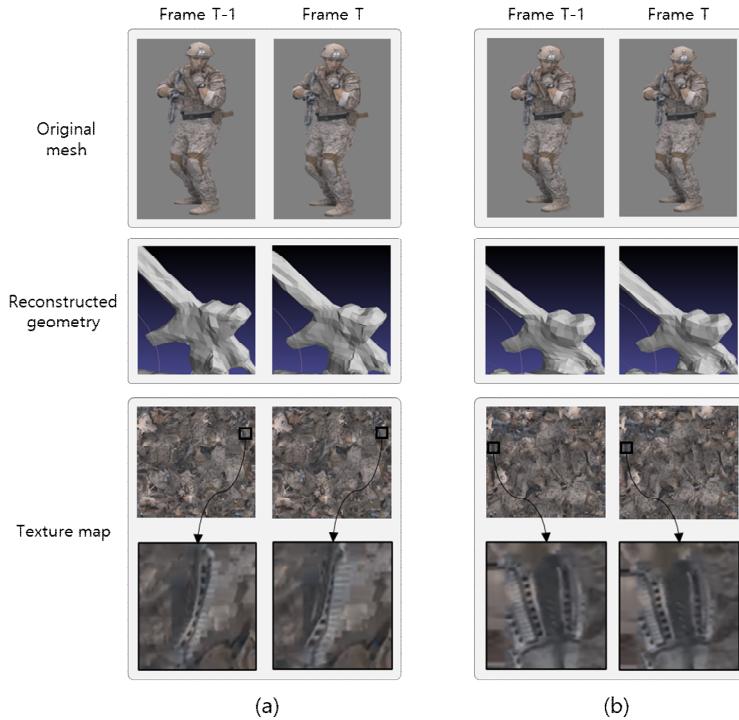


그림 6. 정점 움직임 복잡도에 따른 텍스쳐 맵의 시간적 연속 정도
 Fig 6. Temporal continuity strength of the texture map depending on the vertex motion

저해상도 메쉬의 시간적 예측에 따른 모션 비트율을 지표로 판단할 수 있다. 그림 6은 저해상도 메쉬의 모션 비트율이 상대적으로 큰 경우와 작은 경우에, 각 경우에서 텍스쳐 맵과 직전 텍스쳐 맵 간 유사도를 비교한다. 그림 6의 (a)는 상대적으로 모션 비트율이 큰 경우를, 그림 6의 (b)는 상대적으로 모션 비트율이 작은 경우를 나타낸다. 두 경우 모두 텍스쳐 맵이 직전 텍스쳐 맵과 유사도가 일정 수준 이상 높지만, 세부적으로는 저해상도 메쉬의 모션 비트율이 낮은 (b)가 (a)보다 직전 텍스쳐 맵과의 유사도가 세부적으로 더 높은 것을 보인다.

2. 제안하는 텍스쳐 맵 울 제어 방법

제안하는 텍스쳐 맵 울 제어 방법의 전체적인 수행 과정으로 먼저 꾹쳐와 CTU 순서로 계층적인 목표 비트 할당을 수행하며, 비트 할당의 가장 하위 단위인 CTU의 목표 비트를 달성하기 위한 λ 와 QP 값을 도출하여 인코딩을 수행한다. 제안하는 꾹쳐 단위의 목표 비트 할당 방법은 그림 7의

과정으로 크게 네 단계로 수행된다. 첫 번째 단계는 텍스쳐 맵 인코더에 입력된 텍스쳐 맵 시퀀스에 대하여 꾹쳐 단위로 직전 프레임과의 시간적 연속성의 유무를 판단하는 과정을 수행한다. 시간적 연속성 유무 판단을 낮은 복잡도로 수행하기 위하여 텍스쳐 맵에 대응되는 메쉬를 압축하는 과정에서 결정된 저해상도 메쉬의 압축 모드를 지표로서 사용한다. 두 번째 단계는 꾹쳐의 시간적 연속성의 유무에 무관하게 일정한 화질로 압축될 수 있는 비트량의 예상치를 시퀀스 수준의 비트율에 따라서 통계적으로 도출된 평균치에 기반하여 도출한다. 세 번째 단계에서는 시간적 연속성이 존재하는 꾹쳐에 한해서 시간적 연속 정도에 따라서 비트량의 예상치를 조정한다. 꾹쳐의 시간적 연속 정도를 낮은 복잡도의 방법으로 도출하기 위하여, 꾹쳐에 대응되는 저해상도 메쉬의 모션 비트량과의 반비례 관계를 이용하여 도출하였다. 마지막 단계는 각 꾹쳐의 최종적인 목표 비트를 도출하며, 시퀀스 수준의 가용 비트를 각 꾹쳐의 예상 비트량에 비례하게 할당한다. 제안하는 꾹쳐 단위의 목표 비트 할당 방법의 각 단계의 자세한 수행 과정을 다음

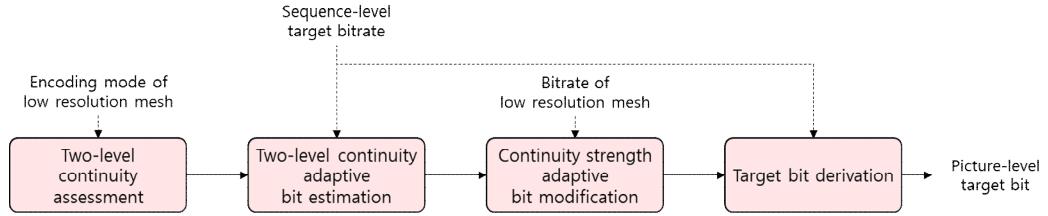


그림 7. 제안하는 픽쳐 단위의 목표 비트 할당 방법
Fig. 7. Proposed target bits allocation methods at the picture level

과 같이 설명한다.

픽쳐 단위 시간적 연속성 판단 단계에서는 각 픽쳐의 시간적 연속성 유무를 판단함으로써 율-왜곡 성능의 수준을 두 가지 수준으로 구분한다. 픽쳐의 시간적 연속성 유무를 저복잡도로 판단하기 위한 방법으로서, 대응되는 저해상도 메쉬가 시간적 예측 방법을 이용하여 압축된 경우에 텍스쳐 맵에 시간적 연속성이 존재한다고 판단한다. 저해상도 메쉬를 압축하는 과정에서 시간적 예측 방법이 선택된 경우는 저해상도 메쉬의 표면 구조에 일정 수준 이상의 시간적 연속성이 존재함을 의미하며, 따라서 복원된 메쉬의 표면 구조 또한 일정 수준 이상의 시간적 연속성이 존재함을 의미한다. 색상 정보가 텍스쳐 맵에 투영되는 위치는 복원된 메쉬 표면의 구조에 따라서 결정되기 때문에, 결과적으로 복원된 메쉬 표면 구조에 시간적 연속성이 존재한다고 판단되는 경우 텍스쳐 맵 또한 시간적 연속성이 존재함으로 판단한다. 시간적 연속성이 존재하지 않는 픽쳐는 P-픽쳐 일지라도 화면 간 예측의 성능이 높지 못하여 인트라 예측 픽쳐와 유사한 발생 비트량을 보이기 때문에, 시간적 연속성이 존재하지 않는 P-픽쳐와 I-픽쳐는 율-왜곡 성능이 유사할 것으로 간주된다.

시간적 연속성에 따른 목표 비트 도출을 위한 첫 번째 단계로, 각 픽쳐에 시간적 연속성 유무에 따라서 두 가지 비트 수준으로 목표 비트를 할당하며, 기존 $R-\lambda$ 모델 기반 율 제어 방법과는 다르게 시퀀스 비트율에 따라 실험적으로 추정한 각 비트 수준의 평균 비트를 이용한다. 기존 $R-\lambda$ 모델 기반 율 제어 방법은 참조 픽쳐 유무에 따라 각 픽쳐의 비트 수준을 결정하며, 픽쳐가 시간 축에서 점진적으로 변화함을 가정하기 때문에 동일 비트 수준이 정의된 픽쳐들 간 율-왜곡 성능이 거의 동일할 것으로 가정한다. 따라서

앞선 GOP에서 픽쳐의 $R-\lambda$ 모델 파라미터를 비트 수준별로 참조하여 현재 GOP 내 픽쳐들의 목표 비트를 $R-\lambda$ 모델을 통해 예측한다. 텍스쳐 맵의 경우 제안하는 픽쳐 분류 절차에 따라 시간적 연속성이 존재함으로 판단된 픽쳐들 간에도 연속 정도의 차이가 존재하며 이에 따른 율-왜곡 성능의 차이가 발생한다. 텍스쳐 맵의 시간적 연속 정도의 차이가 큰 폭으로 존재하는 이유로, 복원된 메쉬의 표면 구조가 시간적으로 변화함에 따라서 발생하는 텍스쳐 맵에서의 시간적 변화가 픽쳐 내 대부분의 영역에서 발생하기 때문이다. GOP 단위로 비트 수준별 $R-\lambda$ 모델 파라미터를 직전 GOP로부터 예측하는 기존 방법은 텍스쳐 맵에 대해서는 $R-\lambda$ 모델 파라미터 예측 성능이 감소할 수 있다.

시간적 연속 강도에 따라 변하는 텍스쳐 맵의 율-왜곡 성능에 개인화된 픽쳐 단위 비트 할당을 수행하기 위하여, 시간적 연속성이 존재하지 않는 픽쳐와 시간적 연속성 존재하는 픽쳐가 동일한 화질로 압축되기 위해 필요한 평균적인 비트를 실험적으로 추정하여 픽쳐 단위 목표 비트를 결정하는 과정에서 사용한다. 일반적인 영상의 율-왜곡 특성에 따르면, 시퀀스 수준의 비트율이 증가할수록 율-왜곡 성능이 서로 다른 두 픽쳐가 동일한 화질로 압축되기 위해 필요한 비트 간 상대적인 차이가 감소하며, 제안하는 방법은 이를 가정하여 시퀀스 비트율에 따라 시간적 연속성이 존재하는 픽쳐와 존재하지 않는 픽쳐가 동일한 화질로 압축될 수 있는 각각의 평균적인 비트를 실험적으로 도출하였다. 사전에 실험 영상들에 대하여 양자화 파라미터를 임의의 값으로 고정하고 압축을 수행하는 과정을 여러 수준의 양자화 파라미터에 대하여 수행함으로써 시퀀스 수준의 다양한 비트율에 대한 압축을 수행하였으며, 이 때 양자화 파라미터를 고정할 경우 각 픽쳐가 유사한 화질로 압축됨을 가

정하였다. 실험을 통해 관측된 결과를 바탕으로, 시퀀스 수준의 비트율과 픽쳐의 시간적 연속성 존재 유무에 따라 평균적으로 발생한 비트의 관계를 수식 18로 정의하였으며, 수식 18을 통해 도출되는 평균 비트 값은 절대 값이 아닌 비트 수준 간 상대적인 값으로 도출된다. 수식 18의 bpp_{seq} 는 시퀀스 수준의 목표 비트율을, $BitLevel[j]$ 는 j 번째 픽쳐의 비트 수준을 의미한다. $BitLevel[j]$ 가 1인 경우는 시간적 연속성이 존재하는 픽쳐이며, 평균 비트의 상대적인 값이 1로 도출된다. $BitLevel[j]$ 가 0인 경우는 시간적 연속성이 존재하지 않는 픽쳐이며, 상대적인 평균 비트는 시퀀스 수준의 목표 비트율에 반비례하게 도출된다. 시퀀스 수준의 목표 비트율과 상대적인 평균 비트 간 반비례 관계를 나타내는 파라미터는, 텍스쳐 맵에 대응되는 저해상도 메쉬의 움직임 복잡도를 의미하는 모션 정보 비트량인 M_{AVG} 가 일정 수준 미만인 경우와 일정 수준 이상인 경우에 대하여 각각 실험적으로 도출되었다.

앞서 도출된 시간적 연속 유무에 따른 비트 기대치의 상대적인 값은 비트 수준별로 평균 비트를 기반으로 도출되었다. 따라서 시간적 연속성이 존재하는 비트 수준 1의 픽쳐의 경우, 시간적 연속 강도와 픽쳐의 율-왜곡 성능이 비례함을 반영해주기 위해서, 수식 17을 통해 연속의 강도에 비례하게 목표 비트의 상대적 값을 감소시킨다. 수식 17의 $RTargetBit'_{BitLevel[j]}(bpp_{seq})$ 는 수식 18을 통해 도출한 시간적 연속 유무에 따른 비트 기대치의 상대적인 값이며, $ContinuityWeight[j]$ 는 시간적 연속 강도에 따른 가중치를 의미한다. 각 픽쳐의 시간적 연속 강도는 대응되는 저해상도 메쉬를 압축하여 발생한 모션 비트량을 지표로 판단

하여, 이는 메쉬 표면의 변화 정도와 텍스쳐 맵의 시간적 연속성의 반비례 관계에 기반하여 저복잡도로 텍스쳐 맵의 시간적 연속성을 판단하기 위함이다. 수식 19는 상대적 목표 비트의 값을 증감하기 위한 가중치를 도출하는 과정을 나타낸다. 수식 19의 $InterLRMeshBit_j$ 는 j 번째 픽쳐에 대응되는 저해상도 메쉬의 압축 과정에서 발생한 모션 비트량을 의미하고, $InterLRMeshAvgBit$ 는 메쉬 정점의 움직임 복잡도가 평균인 경우에 대응되는 평균 모션 비트량을 의미하고, $IntraLRMeshBit$ 는 시간적 연속성이 없는 픽쳐에 대응되는 저해상도 메쉬의 평균 비트량을 의미한다. 수식 18을 통해 도출된 상대적인 목표 비트는 시간적 연속 강도가 평균적인 $InterLRMeshAvgBit$ 값일 경우에 해당되며, 현재 픽쳐의 시간적 연속 강도에 대응되는 목표 비트로 조정하기 위한 가중치를 도출하기 위하여 현재 저해상도 메쉬의 모션 비트량인 $InterLRMeshBit_j$ 와 평균적인 모션 비트량인 $InterLRMeshAvgBit$ 과의 차이의 상대적인 값을 수식 19를 통해 도출한다.

결과적으로 각 픽쳐의 목표 비트는 수식 21을 통해 도출되며, 수식 17~19를 통해 도출된 각 픽쳐의 상대적 목표 비트의 비율만큼을 시퀀스에 남은 비트로부터 할당 받는다. 수식 21의 $RTargetBit'_{BitLevel[j]}$ 는 j 번째 픽쳐의 상대적 목표 비트이고, 수식 20을 통해 도출되는 $RTargetBitSum_{seq}$ 는 j 번째 픽쳐를 포함하여 영상 내 압축하지 않은 남은 픽쳐들의 상대적 목표 비트의 합을 의미하고, $BitLeft_{seq}$ 는 영상 수준에서 남은 비트를 의미한다. 제안하는 픽쳐 수준의 비트 할당 방법을 수행함으로써, 영상 내 픽쳐들은 시간적 연속성의 유무 및 연속 강도의 세밀한 차이에 무관하게 주어진 시퀀스

$$RTargetBit'_{BitLevel[j]}(bpp_{seq}) = RTargetBit_{BitLevel[j]}(bpp_{seq}) \times ContinuityWeight[j] \quad (17)$$

$$RTargetBit_{BitLevel[j]}(bpp_{seq}) = \begin{cases} \frac{1}{165.8} & BitLevel[j] = 1 \\ \frac{(38.3 + bpp_{seq})}{22.6} & BitLevel[j] = 0 \& M_{AVG} < th \\ \frac{22.6}{(11.1 + bpp_{seq})} & BitLevel[j] = 0 \& M_{AVG} \geq th \end{cases} \quad (18)$$

$$ContinuityWeight[j] = \frac{(InterLRMeshBit_j - InterLRMeshAvgBit)}{IntraLRMeshAvgBit} \quad (19)$$

수준의 비트 조건하에 유사한 화질로 압축될 수 있다.

$$RTargetBitSum_{seq} = \sum_{j=0}^{PictureLeftSeq-1} RTTargetBit_{BitLevel[j]} \quad (20)$$

$$TargetBit_{Pic[j]} = BitLeftSeq \times \frac{RTargetBit_{Pic[j]}}{RTargetBitSum_{seq}} \quad (21)$$

제안하는 CTU 수준의 목표 비트 할당 방법은 기존 R- λ 모델 기반 율 제어 방법과 같이 R- λ 모델을 이용하여 픽쳐 내 각 CTU에 대하여 픽쳐 목표 비트율로부터 도출된 람다 값을 사용하여 압축할 경우 예상되는 비트를 목표 비트로 지정한다. 기존 R- λ 모델 기반 율 제어 방법은 픽쳐 간의 시간적 연속성을 전제로 직전에 압축된 동일 비트 수준의 픽쳐에서 동일한 위치의 CTU의 R- λ 모델 파라미터를 참조하는 반면, 텍스쳐 맵의 경우에는 픽쳐의 시간적 연속성이 가변적이기 때문에 동일 위치 CTU 간 율-왜곡 성능이 유사함을 가정할 수 없다. 텍스쳐 맵이 생성되는 과정에서 기하 정보의 시간적 변화에 따라서 메쉬 표면의 색상 정보가 투영되는 텍스쳐 맵 상의 위치가 시간에 따라 달라진다. 이에 따라 임의 위치의 CTU에서 시간에 따른 픽셀 값의 변화와 시공간적 연속성에 변화가 발생하여 인코딩 과정에서의 CTU의 블록 단위 예측 성능이 시간적에 따라 변할 수 있다. CTU의 예측 성능은 율-왜곡 성능에 핵심적인 영향을 미치는 요인으로, 제안하는 방법은 각 CTU에 대하여 예측 성능이 유사할 것으로 예상되는 CTU를 직전에 압축된 동일 비트 수준의 픽쳐에서 탐색하여 해당 CTU의 R- λ 모델 파라미터를 참조하여 CTU 단위의 율 제어를 수행한다.

그림 8은 제안하는 CTU 단위의 R- λ 모델 파라미터 예측

을 통한 CTU 수준의 율 제어의 수행 과정을 나타낸다. 메쉬 표면의 색상이 투영되지 않은 CTU는 비트 할당 과정에서 제외한다. 그림 8에서 나타내는 절차의 순서대로 제안하는 CTU 단위의 율 제어 과정은 각 CTU에 대하여 기존의 R- λ 모델 기반의 율 제어 방법과 동일하게 R- λ 모델에 기반하여 목표 비트 도출과 λ 및 QP 값 도출을 수행하며, 보다 정확한 R- λ 모델 파라미터를 예측하기 위하여 제안하는 R- λ 모델 파라미터 예측 방법을 선행한다.

그림 8의 제안하는 R- λ 모델 파라미터 예측 방법은 각 CTU의 예측 성능이 유사할 것으로 예상되는 CTU를 메쉬 객체에서 동일한 표면의 색상이 투영된 경우로 가정한다. 텍스쳐 맵은 메쉬 표면의 색상 정보를 고해상도로 표현하기 위해, 예를 들어 사람 모델 객체의 경우 HD 해상도에서 4K 해상도에 이르는 범위를 갖는다. 텍스쳐 맵을 기존의 2D 비디오 코덱으로 압축하는 과정에서, 본 논문에서는 64 x 64 픽셀 크기의 CTU를 고려하여 한 CTU 안에 색상 정보가 투영된 메쉬 표면의 면적이 메쉬 객체의 전체 면적에서 차지하는 비율이 매우 작다. 따라서 서로 다른 픽쳐의 임의의 두 CTU에 색상 정보가 투영된 메쉬 표면이 한 객체 내에서 상대적으로 유사한 위치의 표면일수록, 각 CTU의 픽셀 값의 시공간적 연속 특성이 서로 유사할 확률이 높으며 이에 따라 예측 성능 또한 유사할 확률이 높다고 가정한다.

제안하는 R- λ 모델 파라미터 예측 방법의 첫 번째 단계로 각 CTU에 대하여 텍스쳐 맵 변환 과정에서 해당 CTU에 투영된 메쉬 표면 조각의 중심 좌표를 도출한다. 텍스쳐 맵을 변환하는 과정에서, 메쉬 표면에서 샘플링된 포인트마다 색상정보가 텍스쳐 맵으로 투영된 픽셀 좌표를 갖으며 이를 UV 좌표라고 부른다. 각 CTU에 대해, 메쉬 표면에서 샘플링된 포인트의 UV 좌표가 해당 CTU의 영역 내에

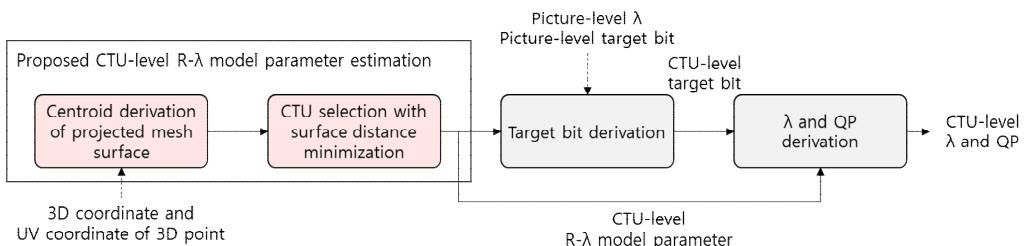


그림 8. 제안하는 CTU 단위의 R- λ 모델 파라미터 예측을 통한 율 제어 과정

Fig. 8. CTU-level rate control process with the proposed R- λ model parameter prediction method

속하는 경우 이들 포인트의 평균 좌표를 도출하며 해당 과정은 수식 22로 표현된다. 수식 22의 CTU_m 은 꽂쳐 내 임의의 CTU를 의미하며, (x_p, y_p, z_p) 는 메쉬의 표면에서 샘플링된 총 n 개의 포인트 중 임의의 p 를 인덱스로 갖는 포인트의 3차원 상의 좌표를 의미하고 (u_p, v_p) 는 인덱스가 p 인 포인트의 UV 좌표를 의미한다. 따라서 수식 22는, n 개의 포인트들 중에서 UV 좌표가 CTU의 영역 내에 포함되는 경우의 포인트들의 3차원 상의 좌표를 평균내는 과정을 나타낸다.

제안하는 $R\text{-}\lambda$ 모델 파라미터 예측 방법의 두 번째 단계로 각 CTU에 대하여 해당 CTU에 투영된 표면과 의미적으로

동일한 표면이 투영된 참조 CTU를 이전에 압축된 시간적 연속 유무가 동일한 꽂쳐에서 탐색한다. 의미적으로 동일한 표면을 저복잡도로 탐색하기 위하여, 현재 CTU와 참조 CTU에 투영된 메쉬 표면의 중심 좌표 간 거리가 최소가 되는 참조 CTU를 선택하며 해당 과정은 수식 23으로 표현된다. 수식 23의 $SurfaceCentroid_{CTU_{curr}}$ 와 $SurfaceCentroid_{CTU_m}$ 는 각각 현재 CTU와 참조를 위한 후보 CTU에 대하여 수식 22를 통해 도출된 투영된 평면의 중심 좌표이며, 중심 좌표 간 차이가 가장 작은 후보 CTU를 선택하는 과정을 나타낸다.

그림 9의 (a)와 (b)는 각각 임의의 꽂쳐에 시간적 연속성

$$SurfaceCentroid_{CTU_m} = \frac{1}{n} \sum_{p=1}^n (x_p, y_p, z_p) \quad \text{where } (u_p, v_p) \in CTU_m \quad (22)$$

$$CTU_{pred} = \arg \min_{CTU_m} |SurfaceCentroid_{CTU_{curr}} - SurfaceCentroid_{CTU_m}| \quad (23)$$

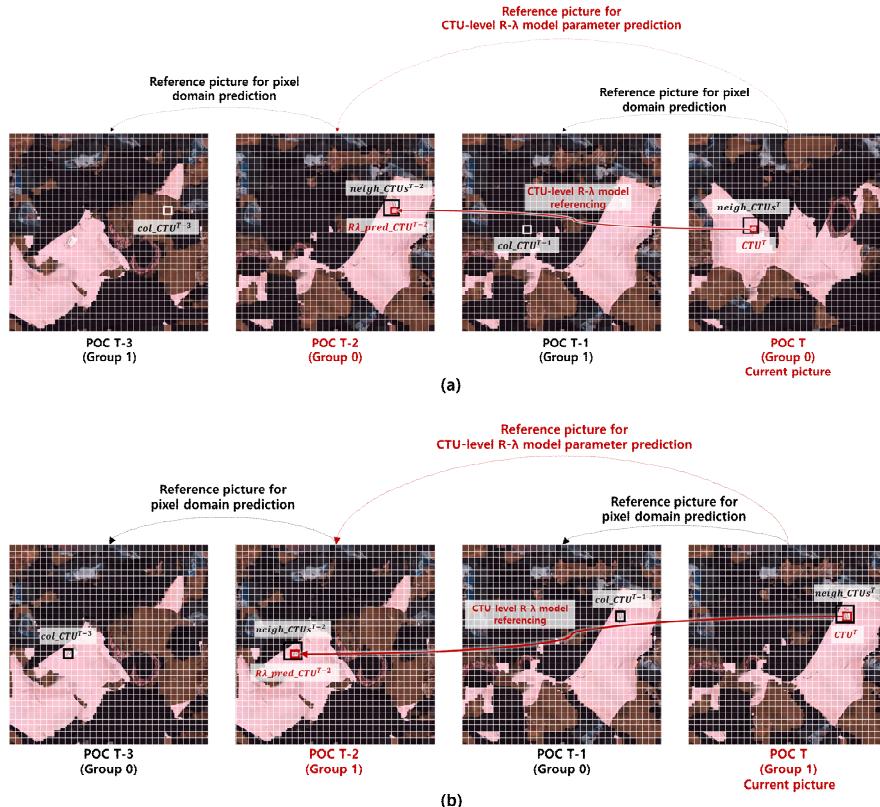


그림 9. 투영된 메쉬 표면 간 최소 거리를 가지는 두 CTU 각각의 시공간적 중복성 비교
 Fig. 9. Comparison of Spatiotemporal Redundancy of Two CTUs with the Minimum Distance of the Projected Mesh Surface

이 존재하지 않는 경우와 존재하는 경우에, 픽쳐 내 임의의 CTU가 투영된 메쉬 표면 간 최소 거리를 갖는 CTU와 유사한 시공간적 연속성을 갖음을 나타낸다. 예시의 텍스쳐 맵의 해상도는 2048×2048 픽셀이고 CTU의 크기는 64×64 픽셀이다. 그림 9에서 CTU^T 는 현재 프레임 내 임의의 CTU이며, $R\lambda_{pred_CTU}^{T-2}$ 는 CTU^T 의 $R\lambda$ 모델 파라미터를 예측하기 위해 선택된 CTU이다. col_CTU^{T-1} 과 col_CTU^{T-3} 은 각각 CTU^T 와 $R\lambda_{pred_CTU}^{T-2}$ 의 블록 단위 압축과정에서 시간적으로 참조할 수 있는 동일 위치 CTU이며, $neigh_CTUs^T$ 와 $neigh_CTUs^{T-2}$ 는 각각 CTU^T 와 $R\lambda_{pred_CTU}^{T-2}$ 의 인접한 복원 영역을 나타낸다.

시간적 연속성의 경우, 픽쳐의 시간적 연속성이 존재할 때는 CTU와 $R\lambda$ 모델 파라미터 참조 CTU 각각의 참조 프레임에서 동일 위치의 CTU와의 시간적 연속성이 모두 강함을 보이며, 픽쳐의 시간적 연속성이 존재하지 않는 경우에는 두 CTU 모두 시간적 연속성이 약하게 존재하여 시간적 연속 정도가 유사하다고 판단할 수 있다.

CTU 수준의 목표 비트 할당이 완료되면 목표 비트에 따른 CTU의 λ 와 QP 값을 각각 수식 12와 수식 4를 통해 도출한 후 인코딩을 수행한다. 수식 12를 이용하여 목표 비트율로부터 λ 값을 도출하는 과정에서 사용하는 $R\lambda$ 모델 파라미터는 해당 CTU의 목표 비트 결정 과정에서 참조한 CTU의 $R\lambda$ 모델 파라미터를 동일하게 사용한다. 메쉬 표면의 색상 정보가 투영되지 않아서 비트가 할당되지 않은 CTU의 경우에는, 해당 CTU의 λ 값은 고정된 최대 값으로 설정된다. 각 CTU와 픽쳐의 인코딩을 마치면 해당되는 비트 수준의 CTU와 픽쳐 레벨의 $R\lambda$ 모델 파라미터를 업데이트한다.

IV. 실험 결과

본 장에서는 제안하는 텍스쳐 맵을 제어 방법의 성능을 평가하며 V-DMC 레퍼런스 소프트웨어 버전 8 (TMM v8.0)^[35] 텍스쳐 맵 인코더의 기존을 제어 방법과 성능을

비교하여 제안하는 방법의 우수성을 보인다. 4.1절에서는 제안하는 방법과 기존 방법의 성능을 평가한 동일한 실험 환경 및 조건을 제시하고 4.2절에서는 제안하는 방법과 기존 방법의 성능을 비트 오차율, 압축 효율, 객관적 화질 변동 측면에서 비교한다.

1. 실험 환경 및 조건

본 논문의 제안하는 윤 제어 방법은 TMM v8.0을 기반으로 구현하였으며, 기존 방법과의 성능 비교를 위하여 TMM v8.0의 텍스쳐 맵 인코더인 HM v16.21^[31]의 기존 윤 제어 방법과 성능을 비교하였다. 제안하는 방법과 기존 방법의 성능을 동일한 실험 환경에서 비교하고자 다음과 같이 공통의 실험 환경을 구성하였다. 성능 측정을 위한 실험 시퀀스로 TMM v8.0의 공통 실험 조건 (Common Test Condition, CTC)^[36]에서 지정한 8 가지의 시퀀스를 사용하였으며, 표 1은 각 시퀀스에 대하여 CTC에 정의된 다섯 가지 비트율 구간인 R1~R5에서 텍스쳐 맵의 원본 해상도, 재생성 후 해상도와 양자화 파라미터를 나타낸다. 양자화 파라미터를 제외한 텍스쳐 맵의 인코딩 파라미터와 메쉬의 인코딩 파라미터를 CTC에 따라 설정하였으며, 텍스쳐 맵의 양자화 파라미터는 각 윤 제어 과정에서 결정된다. 제안하는 방법과 기존 방법의 텍스쳐 맵 인코더의 시퀀스 수준 목표 비트율을 동일하게 설정하고자, 각 시퀀스와 각 비트율 구간 R1~R5에 대하여 CTC를 따라 압축했을 때 발생하는 텍스쳐 맵의 비트율을 텍스쳐 맵 인코더의 시퀀스 수준의 목표 비트율로 설정하였다. 표 2는 시퀀스와 비트율 구간별로 설정한, 텍스쳐 맵 인코더의 시퀀스 수준 목표 비트율을 나타낸다. 메쉬와 텍스쳐 맵의 시간적 예측 구조로 LD 환경을 지정하였으며 GOP 크기와 인트라 픽쳐 간격을 각각 32 픽쳐로 설정하였다.

2. 목표 비트 달성을 평가

본 절에서는 제안하는 윤 제어 방법의 목표 비트 달성 오차율을 시퀀스와 픽쳐 수준에서 평가한 결과를 제시한다. 목표 비트 달성 오차율은 인코딩 대상의 윤-왜곡 성능을 적절하게 예측하여 목표 비트율로 압축할 수 있는 λ 의 값을

표 1. TMM v8.0의 공통 실험 조건에 정의된 테스트 시퀀스

Table 1. Test sequences defined in the common test condition of TMM v8.0

Category	Sequence	Frame count	Rate ID	Original texture map resolution	Transferred texture map resolution	Texture map quantization parameter	Visualization
A	Longdress	300	R1	2048 × 2048	1024 × 1024	41	
			R2		1024 × 1024	34	
			R3		2048 × 2048	38	
			R4		2048 × 2048	36	
			R5		2048 × 2048	32	
	Soldier	300	R1	2048 × 2048	2048 × 2048	33	
			R2		2048 × 2048	31	
			R3		2048 × 2048	29	
			R4		2048 × 2048	27	
			R5		2048 × 2048	26	
B	Basketball_player	300	R1	2048 × 2048	2048 × 2048	41	
			R2		2048 × 2048	35	
			R3		2048 × 2048	29	
			R4		2048 × 2048	26	
			R5		2048 × 2048	22	
	Dancer	300	R1	2048 × 2048	1024 × 1024	36	
			R2		2048 × 2048	36	
			R3		2048 × 2048	30	
			R4		2048 × 2048	27	
			R5		2048 × 2048	23	
C	Football	300	R1	4096 × 4096	1024 × 1024	37	
			R2		2048 × 2048	36	
			R3		4096 × 4096	39	
			R4		4096 × 4096	36	
			R5		4096 × 4096	34	
	Levi	150	R1	4096 × 4096	2048 × 2048	35	
			R2		2048 × 2048	29	
			R3		4096 × 4096	33	
			R4		4096 × 4096	30	
			R5		4096 × 4096	27	
	Mitch	300	R1	4096 × 4096	4096 × 4096	44	
			R2		4096 × 4096	42	
			R3		4096 × 4096	40	
			R4		4096 × 4096	39	
			R5		4096 × 4096	37	
	Thomas	300	R1	4096 × 4096	4096 × 4096	40	
			R2		4096 × 4096	39	
			R3		4096 × 4096	37	
			R4		2048 × 2048	30	
			R5		2048 × 2048	28	

표 2. 각 시퀀스와 각 비트율 구간의 목표 비트율

Table 2. Target bitrate for each sequence and each bitrate range

Sequence name	Rate ID	Target bitrate (kbps)	Sequence name	Rate ID	Target bitrate (kbps)
Longdress	R1	922	Football	R1	1228
	R2	2212		R2	2573
	R3	3138		R3	3771
	R4	3939		R4	5088
	R5	6210		R5	6456
Soldier	R1	1139	Levi	R1	594
	R2	2073		R2	1277
	R3	2836		R3	1744
	R4	3910		R4	2502
	R5	5197		R5	3406
Basketball_player	R1	634	Mitch	R1	783
	R2	1202		R2	1156
	R3	2337		R3	1763
	R4	3386		R4	2238
	R5	5611		R5	3560
Dancer	R1	604	Thomas	R1	929
	R2	1238		R2	1144
	R3	2400		R3	1811
	R4	3404		R4	2419
	R5	5587		R5	3868

표 3. 제안하는 방법과 기존 방법의 시퀀스 수준 비트 오차율

Table 3. Sequence-level bit error rate of the proposed method and the existing method

Sequence name	Rate ID	Sequence bit error (%)		Sequence name	Rate ID	Sequence bit error (%)	
		TMM v8.0 RC	Proposed			TMM v8.0 RC	Proposed
Longdress	R1	0.04	0.03	Football	R1	0.05	0.07
	R2	0.03	0.01		R2	0.03	0.02
	R3	0.02	0.00		R3	0.01	0.02
	R4	0.02	0.02		R4	0.01	0.00
	R5	0.02	0.01		R5	0.01	0.02
Soldier	R1	0.36	0.04	Levi	R1	0.27	0.03
	R2	0.06	0.01		R2	0.18	0.03
	R3	0.01	0.03		R3	0.19	0.01
	R4	0.01	0.01		R4	0.13	0.00
	R5	0.01	0.01		R5	0.23	0.00
Basketball_player	R1	0.03	0.03	Mitch	R1	25.24	0.05
	R2	0.02	0.03		R2	8.17	0.03
	R3	0.02	0.02		R3	3.62	0.02
	R4	0.01	0.01		R4	2.14	0.02
	R5	0.01	0.01		R5	0.25	0.01
Dancer	R1	0.06	0.06	Thomas	R1	5.79	0.04
	R2	0.03	0.02		R2	3.64	0.03
	R3	0.04	0.02		R3	2.22	0.02
	R4	0.02	0.01		R4	0.03	0.02
	R5	0.01	0.01		R5	0.01	0.01

표 4. 제안하는 방법과 기존 방법의 픽처 수준 비트 오차율

Table 4. Picture-level bit error rate of the proposed method and the existing method

Sequence name	Rate ID	Picture bit mean error (%)		Sequence name	Rate ID	Picture bit mean error (%)	
		TMM v8.0 RC	Proposed			TMM v8.0 RC	Proposed
Longdress	R1	5.8%	4.6%	Football	R1	6.6%	6.1%
	R2	6.3%	3.2%		R2	6.0%	3.0%
	R3	6.0%	3.4%		R3	6.1%	2.5%
	R4	6.6%	2.8%		R4	5.8%	2.5%
	R5	6.4%	1.9%		R5	5.5%	2.4%
Soldier	R1	11.8%	1.7%	Levi	R1	11.0%	4.7%
	R2	16.6%	3.6%		R2	9.2%	4.2%
	R3	15.8%	2.0%		R3	8.4%	1.9%
	R4	14.2%	2.9%		R4	7.6%	2.8%
	R5	13.5%	1.9%		R5	7.0%	2.4%
Basketball_player	R1	6.4%	3.2%	Mitch	R1	234.5%	2.2%
	R2	6.1%	3.3%		R2	534.1%	1.7%
	R3	5.9%	1.5%		R3	70.3%	1.4%
	R4	5.3%	1.7%		R4	54.1%	1.6%
	R5	5.4%	1.7%		R5	41.4%	1.6%
Dancer	R1	6.3%	3.1%	Thomas	R1	151.2%	2.0%
	R2	6.6%	2.8%		R2	121.1%	4.5%
	R3	6.4%	1.7%		R3	66.8%	2.1%
	R4	5.7%	1.4%		R4	30.2%	3.5%
	R5	5.3%	1.5%		R5	21.8%	2.8%

얼마나 정확하게 예측했는지를 평가하는 지표이다. 표 3과 표 4는 각각 제안하는 방법과 기존 방법의 목표 비트 달성을 시퀀스 단위와 픽처 단위에서 측정한 결과를 제시한다. 시퀀스 수준의 에러율과 픽처 수준의 평균 에러율을 각각 수식 24와 25를 통해 계산하였다. 수식 24의 $TargetBit_{Seq}$ 와 $ActualBit_{Seq}$ 는 각각 시퀀스 수준의 목표 비트와 실제 발생 비트를 의미하며, 수식 25의 $RTargetBit_{BitLevel[j]}$ 와 $ActualBit_{Pic}[i]$ 는 각각 i 번째 픽처의 목표 비트와 실제 발생 비트를 의미한다. 제안하는 방법은 시퀀스 수준과 픽처 수준에서 각각 1% 내외, 4% 이내의 에러율을 보인다. 기존 방법은 Soldier, Levi, Mitch, Thomas 시퀀스에 대하여 시퀀스 수준과 픽처 수준에서 모두 상대적으로 높은 에러율을 보인다. 전체 8개의 시퀀스 중에서 해당 4개의 시퀀스는 R1~R5의 모든 비트율 구간

에서 기본 기하 정보의 화면 간 예측을 수행 비율이 약 30% 이상으로, 텍스쳐 맵의 시간적 연속성의 변화가 빈번한 특징을 갖는다. 그 중 Thomas와 Mitch 시퀀스는 기본 기하 정보의 화면 간 예측 수행 여부가 시간 축에서 가장 빈번하게 변하는 특징을 갖는다. 기존 율 제어 방법은 텍스쳐 맵 인코더에서 모든 P-픽처를 한 종류의 비트 수준으로 간주하여 R- λ 모델 파라미터를 공유하기 때문에, 이와 같이 유효곡 성능에 큰 차이가 존재하는 픽처가 번갈아 등장하는 빈도가 높을수록 목표 비트 달성 오차율이 증가하게 된다. 그럼 10은 Levi 시퀀스의 R1을 예시로 제안하는 방법과 기존 방법의 픽처 단위 비트 오차 발생의 양상을 보여준다. 그럼 10의 실험 결과를 도출하기 위하여 제안하는 방법과 기존 방법에 동일하게 텍스쳐 맵의 시퀀스 수준의 목표 비트를 14,857,845 비트로 설정하여 율 제어를 수행하였다.

$$BitErrorRate_{Seq} = abs\left(\frac{TargetBit_{Seq} - ActualBit_{Seq}}{TargetBit_{Seq}}\right) \times 100 \quad (24)$$

$$AvgBitErrorRate_{Pic} = \sum_{i=0}^{PictureCount-1} abs\left(\frac{TargetBit_{Pic}[i] - ActualBit_{Pic}[i]}{TargetBit_{Pic}[i]}\right) \times 100 \quad (25)$$

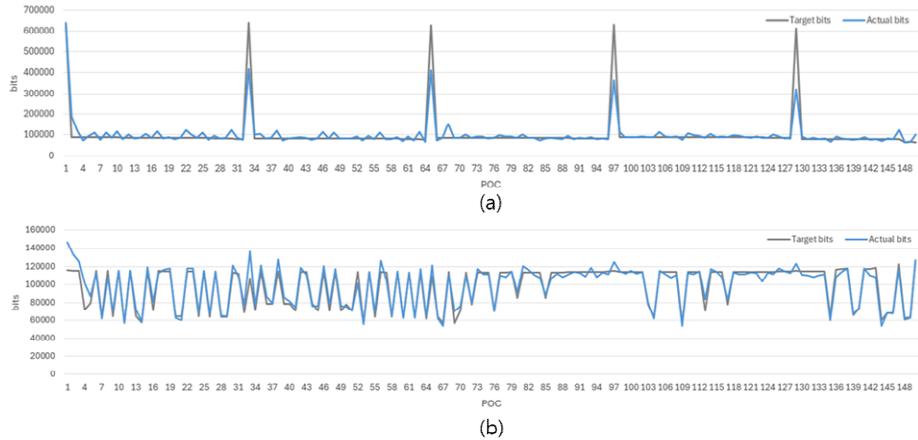


그림 10. Levi 시퀀스의 R1 조건에서 제안하는 방법과 기존 방법의 픽쳐 단위 비트 오차 비교
Fig. 10. Comparison of picture-level bit error between the proposed method and the existing method under the R1 condition of the Levi sequence

을 제어 과정에서 이전 각 GOP와 픽쳐에서 발생한 비트 에러는 다음 각 GOP와 픽쳐의 목표 비트를 설정하는 단계에서 일부 값만 보상되기 때문에, 픽쳐 단위로 실제 할당된 비트의 합은 시퀀스 단위의 당초 목표 비트와 차이가 존재 할 수 있다. 기존 방법에서 픽쳐 단위로 할당된 비트의 총 합과 실제로 소비한 비트는 각각 15,110,062 비트와 15,009,848 비트이다. 제안하는 방법은 픽쳐 단위로 할당된 비트의 총 합이 14,698,105 비트이며 실제로 소비한 비트는 총 14,861,752 비트로, 제안하는 방법이 기존 방법 대비 시퀀스 수준의 당초 목표 비트를 정확하게 달성하였음을 보인다. 그림 10의 결과에서 픽쳐 단위 시간적 연속 특성이 서로 다른 POC 0~32 구간과 POC 85~101 구간에 대하여 제안하는 방법과 기존 방법의 비트 달성을 다음과 같 이 분석하였다. POC 0~32 구간은 시간적 연속성이 존재하는 15장의 픽쳐와 시간적 연속성이 존재하지 않는 18장의 픽쳐가 교차하여 등장하는 구간이다. 따라서 해당 구간 내 픽쳐들의 율-왜곡 성능의 편차가 큰 픽쳐가 빈번하게 교차하여 등장하는 구간으로, 각 픽쳐의 율-왜곡 성능을 의미하는 $R-\lambda$ 모델 파라미터 α 의 표준 편차가 약 0.01인 구간이다. 해당 구간에서는 연속되는 픽쳐들의 $R-\lambda$ 모델 파라미터가 서로 일정 수준 이상으로 유사하기 때문에, 기존 방법 또한 낮은 비트 달성을 보인다.

파라미터를 예측하는 과정에서 이전에 압축된 픽쳐들 중 현재 픽쳐와 시간적 연속 유무가 동일한 픽쳐의 $R-\lambda$ 모델 파라미터를 참조하기 때문에, 시간적 연속성의 변동이 큰 구간에서도 $R-\lambda$ 모델 파라미터를 기준 방법 대비 정확하게 예측할 수 있음을 보인다. POC 85~101 구간은 시간적 연속성이 유사한 픽쳐가 연속적으로 등장하는 구간으로, 각 픽쳐의 율-왜곡 성능을 의미하는 $R-\lambda$ 모델 파라미터 α 의 표준 편차가 약 0.01인 구간이다. 해당 구간에서는 연속되는 픽쳐들의 $R-\lambda$ 모델 파라미터가 서로 일정 수준 이상으로 유사하기 때문에, 기존 방법 또한 낮은 비트 달성을 보인다.

3. 압축 효율 및 객관적 화질 일관성 평가

제안하는 방법의 압축 효율을 평가하기 위하여 기존 윤제어 방법과 비교하여 3D 정점의 복원된 색상 정보와 렌더링된 2D 이미지의 색상 정보에 대하여 Bjøntegaard-Delta rate (BD-rate)^[37,38]를 측정하였다. BD-rate는 동일한 화질에서 기존 방법 대비 제안하는 방법의 비트 변화율을 의미 한다. 제안하는 방법과 기존 방법 모두 시퀀스 수준의 목표 비트 대비 평균 1% 내외의 유사한 비트로 압축되었으므로, 제시된 BD-rate 결과는 제안하는 방법의 효율적인 비트 사용을 통한 화질 향상의 의미를 포함한다. 표 5는 제안하는 윤제어 방법을 TMM v8.0 텍스쳐 맵 인코더의 기존 윤

표 5. 3D 정점과 렌더링된 2D 이미지의 복원된 기하 및 색상 정보에 대한 BD-rate

Table 5. BD-rate for the reconstructed geometry and color of 3D points and rendered 2D images

Category	Sequence name	BD-rate for reconstruction of 3D points					BD-rate for reconstruction of rendered 2D image	
		D1	D2	Luma	Chroma Cb	Chroma Cr	Geom	Luma
A	Longdress	0.0%	0.0%	-6.8%	-5.6%	-5.8%	0.0%	-5.9%
	Soldier	-0.1%	-0.1%	-13.8%	13.7%	12.1%	-0.1%	-11.3%
B	Basketball_player	0.0%	0.0%	-8.6%	-7.0%	-7.5%	0.0%	-4.9%
	Dancer	0.0%	0.0%	-8.0%	-7.4%	-8.5%	0.0%	-4.7%
C	Football	0.0%	0.0%	-9.1%	-13.5%	-12.3%	0.0%	-4.6%
	Levi	-0.1%	-0.1%	-8.8%	-7.7%	-7.5%	-0.1%	-4.1%
	Mitch	-2.6%	-2.5%	-29.2%	-42.2%	-39.2%	-2.7%	-21.4%
	Thomas	-1.7%	-1.6%	-26.0%	-57.2%	-31.5%	-1.7%	-20.2%
Cat-A average		-0.1%	-0.1%	-10.3%	4.1%	3.2%	-0.1%	-8.6%
Cat-B average		0.0%	0.0%	-8.3%	-7.2%	-8.0%	0.0%	-4.8%
Cat-C average		-1.1%	-1.1%	-18.3%	-30.2%	-22.6%	-1.1%	-12.6%
Overall average		-0.4%	-0.4%	-12.3%	-11.1%	-9.2%	-0.4%	-8.7%

표 6. 제안하는 방법과 기존 방법의 PSNR의 표준 편차 비교

Table 6. Comparison of standard deviation of PSNR between the proposed and the existing method

Category	Sequence name	Rate ID	Standard deviation of PSNR for reconstruction of 3D points (Luma)		Standard deviation of PSNR for reconstruction of rendered 2D image (Luma)	
			TMM v8.0 RC	Proposed	TMM v8.0 RC	Proposed
A	Longdress	R1	1.1	0.5	0.7	0.4
		R2	1.0	0.5	0.6	0.5
		R3	1.1	0.5	0.7	0.4
		R4	1.0	0.5	0.7	0.4
		R5	0.9	0.5	0.6	0.4
	Soldier	R1	2.4	0.3	1.2	0.4
		R2	2.2	0.4	1.3	0.3
		R3	1.9	0.3	1.1	0.3
		R4	1.6	0.3	0.9	0.3
		R5	1.3	0.3	0.7	0.3
B	Basketball_player	R1	1.2	0.5	0.5	0.3
		R2	1.0	0.4	0.4	0.3
		R3	0.7	0.3	0.3	0.2
		R4	0.7	0.3	0.3	0.2
		R5	0.5	0.3	0.3	0.2
	Dancer	R1	1.0	0.4	0.5	0.3
		R2	1.0	0.3	0.4	0.3
		R3	0.8	0.3	0.4	0.3
		R4	0.7	0.3	0.4	0.2
		R5	0.5	0.3	0.3	0.3
C	Football	R1	1.0	0.8	0.8	0.8
		R2	0.8	0.8	0.8	0.8
		R3	0.8	0.7	0.7	0.7
		R4	0.7	0.7	0.7	0.7
		R5	0.7	0.6	0.6	0.6
	Levi	R1	1.3	0.9	1.1	1.1
		R2	1.2	0.9	0.6	0.5
		R3	1.2	1.1	0.7	0.5
		R4	1.2	1.2	0.6	0.5
		R5	1.3	1.3	0.7	0.6
	Mitch	R1	0.6	0.2	0.4	0.2
		R2	0.7	0.2	0.5	0.2
		R3	0.7	0.2	0.6	0.2
		R4	0.8	0.3	0.6	0.2
		R5	0.9	0.3	0.7	0.3
	Thomas	R1	0.7	0.1	0.5	0.1
		R2	0.7	0.1	0.6	0.1
		R3	0.8	0.1	0.6	0.1
		R4	0.5	0.1	0.4	0.1
		R5	0.5	0.1	0.4	0.1

제어 방법과 비교하여 3D 정점의 복원된 색상 정보와 렌더링된 2D 이미지의 색상 정보에 대한 BD-rate 성능을 나타낸다. 제안하는 방법은 기존 방법과 비교하여 3D 정점의 복원된 휘도 성분과 렌더링된 2D 이미지의 휘도 성분에 대하여 각각 12.3%와 8.7%의 BD-rate 향상을 달성한다. 제안하는 방법은 텍스쳐 맵의 시간적 연속성이 가변적이라는 특성에 따라서 시간적 참조 구조에 종속적인 비트 할당을 수행하지 않고 텍스쳐 맵의 시간적 연속성에 따라 차등적인 비트 할당을 수행함으로써 영상 평균 화질의 향상을 달성하였다.

표 6은 제안하는 방법과 기존 방법의 객관적 화질의 시간적 일관성을 비교하고자 3D 정점과 렌더링된 2D 이미지의 각 휘도 성분의 픽쳐 단위 Peak Signal-to-Noise Ratio (PSNR)의 표준 편차를 수식 26과 27을 통해 계산하여 제시한다. 수식 26의 $3DPointPSNR_{Pic}[i]$ 는 i 번째 픽쳐에서 복원된 3D 정점의 색상 정보에 대하여 원본 3D 정점의 색상 정보를 비교하여 계산된 PSNR을 의미하며, Avg3DPointPSNR_{pic}는 시퀀스 내 모든 픽쳐의 3DPointPSNR_{pic}[i] 값의 평균 값을 의미한다. 수식 27의 ImagePSNR_{pic}[i]는 i 번째 픽쳐에서 복원된 3D 정점의 색상 정보를 특정 다수 개의 시점의 평면으로 투영한 이미지 도메인에서, 복원된 색상 정보와 원본 색상 정보를 비교하여 계산된 PSNR을 의미한다.

표 6에 제시된 결과를 통해 제안하는 방법이 기존 방법보다 픽쳐 단위 PSNR의 낮은 표준 편차를 달성하였음을 확인할 수 있으며, 제안하는 방법이 텍스쳐 맵의 가변적인 윤곽 성능에 따라 일관된 객관적 화질로 압축될 수 있도록 적절한 비트를 할당하였음을 의미한다. 그림 11은 제안하는 방법이 기존 방법 대비 높은 객관적 화질 일관성을 달성을 함을 시각적으로 나타내고자 하며, Soldier 시퀀스의 R2 조건에서 측정한 픽쳐 단위 PSNR을 시간에 따른 그래프로 나타낸다. 그림 12와 그림 13은 그림 11에서 PSNR 결과를 제시한 총 300장의 픽쳐들 중, 각각 113~117번째 픽쳐 구간과 126~130번째 픽쳐 구간에 대하여 복원된 색상 정보

$$\sigma_{3DPointPSNR} = \sqrt{\frac{1}{PictureCount} \sum_{i=0}^{PictureCount-1} (3DPointPSNR_{Pic}[i] - \text{Avg3DPointPSNR}_{Pic})^2} \quad (26)$$

$$\sigma_{ImagePSNR} = \sqrt{\frac{1}{PictureCount} \sum_{i=0}^{PictureCount-1} (\text{ImagePSNR}_{Pic}[i] - \text{AvgImagePSNR}_{Pic})^2} \quad (27)$$



그림 11. Soldier 시퀀스의 R2 조건에서 제안하는 방법과 기존 방법의 3D 정점의 복원된 휘도 성분의 PSNR 변화 비교

Fig. 11. Comparison of PSNR variation of the reconstructed luminance component of 3D points between the proposed method and the existing method under the R2 condition of the Soldier sequence

를 렌더링하여 제안하는 방법의 시공간적 화질 변동이 기존 방법 대비 상대적으로 안정적임을 보인다. 그림 12 와 그림 13의 (a)는 기존 방법의 결과를, (b)는 제안하는 방법의 결과를 나타낸다. 그림 12는 제안하는 방법의 객관적 화질이 기존 방법보다 높은 구간을 나타내며, 그림 13은 반대로 기존 방법이 제안하는 방법보다 객관적 화질이 높은 구간을 나타낸다. 그림 12에서 POC 113의 텍스쳐 맵은 시간적 연속성이 존재하지 않으며 POC 114~117의 텍스쳐 맵은 시간적 연속성이 존재하는 경우이다. 제안하는 방법은 POC 113의 텍스쳐 맵에 더 많은 비트를 할당하여 주/객관적 화질의 일관성을 달성하였고, 기존 방법은 POC 113~117의 텍스쳐 맵에 균일한 비

트를 할당하여 POC 113의 텍스쳐 맵에서 심한 화질 저하가 발생하였으며, POC 114의 텍스쳐 맵은 직전 텍쳐로부터의 $R-\lambda$ 모델 파라미터 예측 성능이 낮기 때문에 화질 저하가 발생하였다. POC 113~114의 텍스쳐 맵에서 발생한 화질 저하는 이후 텍스쳐 맵들의 압축 성능 저하에도 영향 미치는 경향을 확인할 수 있다. 그림 13은 기존 방법이 제안하는 방법보다 실제 발생 비트 및 객관적 화질이 높음에도 불구하고, 기존 방법 대비 제안하는 방법의 주관적 화질 저하가 적게 발생하였음을 나타낸다. 이는 제안하는 CTU 단위의 $R-\lambda$ 모델 파라미터 예측 방법을 통해서 한 텍쳐 내에서 CTU 단위로 효율적인 비트 할당이 수행되었음을 나타낸다.

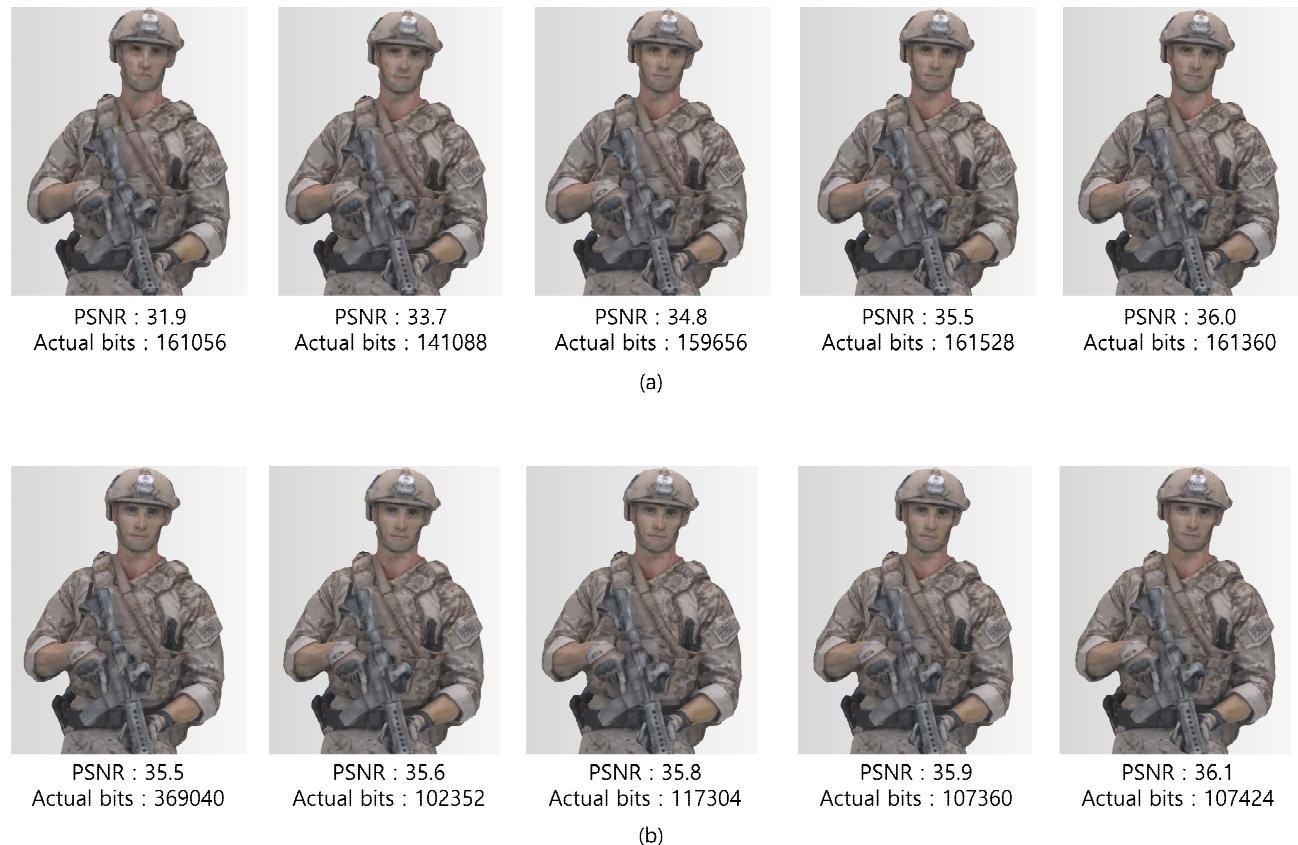


그림 12. 제안하는 방법과 기존 방법의 Soldier 시퀀스의 R2 조건에서 시공간적 화질의 일관성 비교 (POC 113~117)
 Fig. 12. Visualization for comparing the consistency of image quality between the proposed method and the existing method under the R2 condition of Soldier sequence

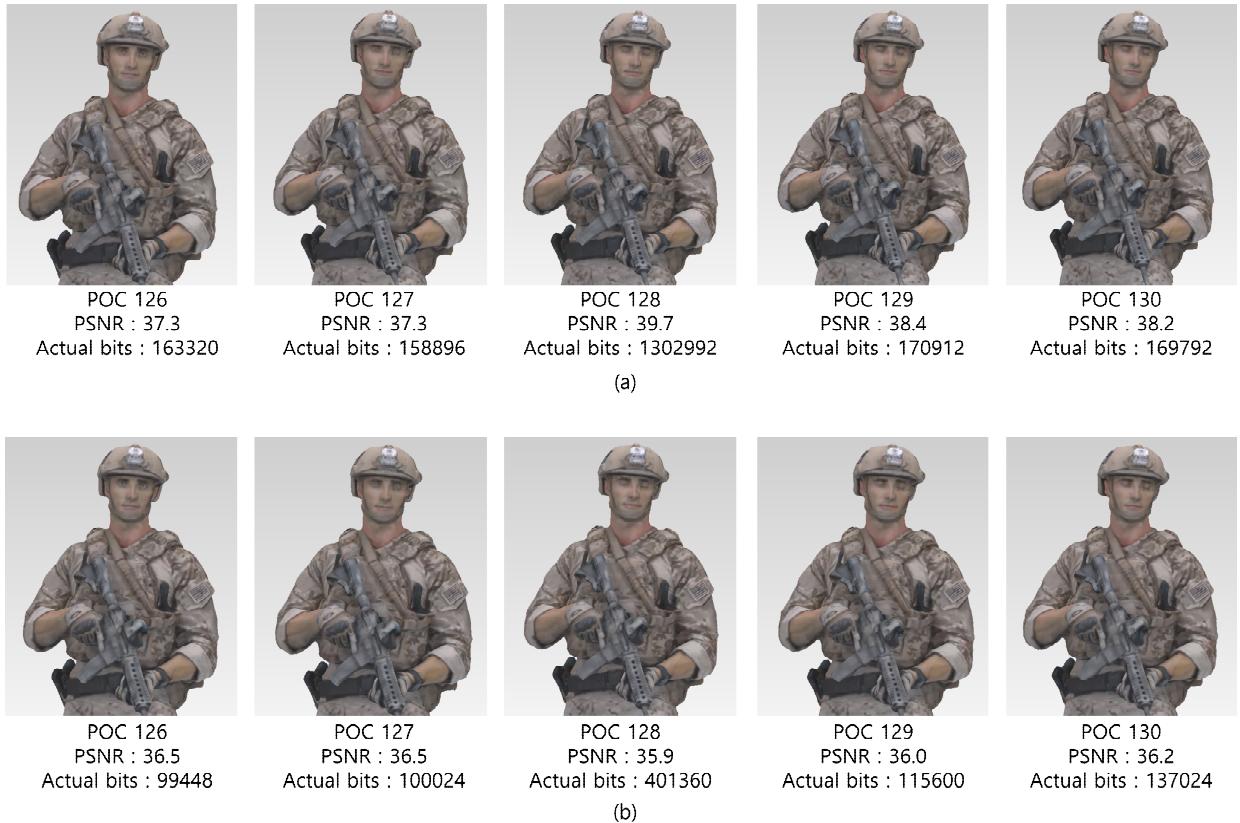


그림 13. 제안하는 방법과 기존 방법의 Soldier 시퀀스의 R2 조건에서 시공간적 화질의 일관성 비교 (POC 126~130)

Fig. 13. Visualization for comparing the consistency of image quality between the proposed method and the existing method under the R2 condition of Soldier sequence (POC 126~130)

4. 제안하는 방법의 연산량 분석

본 절에서는 제안하는 빠져 수준의 목표 비트 할당 방법과 CTU의 $R-\lambda$ 모델 파라미터 예측 방법을 수행하는 과정에서 요구되는 연산량을 분석한다. 제안하는 빠져 수준의 목표 비트 할당을 수행하는 과정은 시간적 연속성 유무와 시퀀스 수준의 비트율에 따른 목표 비트량의 상대적인 값을 도출하는 과정, 시간적 연속 정도에 따라서 목표 비트량의 상대적인 값을 조정하는 과정, 절대적인 목표 비트량을 계산하는 과정으로 이루어지며, 메쉬를 압축하는 과정에서 도출된 정보를 활용하여 총 $O(1)$ 의 낮은 연산량을 요구한다. 먼저 각 빠져마다 수행되는 시간적 연속 유무를 판단하는 과정은 빠져마다 대응되는 저해상도 메쉬의 화면간 예측 수행 여부를 1회 판단하는 연산량을 요구한다. 시퀀스

비트율에 따른 빠져 목표 비트량의 상대적인 값을 도출하는 과정 또한, 각 빠져의 시간적 연속 유무에 따라 정의된 비선형 수식을 1회 계산한다. 시간적 연속 정도와 시간적 연속 정도에 따른 목표 비트량의 상대적인 값을 조정하는 과정은 각각 선형 방정식을 통해 계산된다. 마지막 과정으로, 도출된 각 빠져의 목표 비트량의 상대적인 값에 해당하는 비율만큼을 시퀀스 수준에서 남은 비트량으로부터 선형 방정식을 통해 절대적인 수치의 목표 비트를 할당 받는다. 기존의 빠져 수준의 목표 비트 할당 방법은 $R-\lambda$ 관계식을 통해 도출된 각 빠져의 예상 비트량의 합이 시퀀스 총 가용 비트량에 근접하도록 GOP 수준의 λ 값을 조정해가며 반복적 기법을 통해 비트를 할당한다. 따라서 반복 횟수만큼 빠져마다 예상 비트량을 $R-\lambda$ 관계식을 통해 계산해야 되지만, 제안하는 빠져 수준의 목표 비트 할당 방법은 해당 과정

을 수행하지 않기 때문에 제안하는 꾹쳐 수준의 방법을 수행함으로써 전체 인코딩 수행 시간에 유의미한 변동이 발생하지 않을 것으로 예상된다.

제안하는 CTU 수준의 $R\cdot\lambda$ 모델 파라미터 예측 방법은 CTU마다 투영된 표면의 중심 좌표를 계산하는 과정과 중심 좌표 간 거리가 최소가 되는 CTU를 탐색하는 절차로 이루어지며, 기존의 율 제어 방법 대비 추가적인 연산량을 요구한다. 꾹쳐 내 모든 CTU에 대하여 각 CTU에 투영된 표면의 중심 좌표를 계산하는 과정에서, 먼저 메쉬 객체의 전체 표면에서 샘플링된 n 개의 포인트들에 대하여 UV 좌표를 CTU 영역 크기로 나누어서 해당 포인트가 투영되는 CTU를 찾는 과정이 수행하며 이는 한 꾹쳐 당 $O(n)$ 의 계산 복잡도를 갖는다. 그리고 꾹쳐 내 m 개의 CTU에 대하여 각 CTU에 투영된 포인트들의 3차원 공간상의 평균 좌표를 계산하며 이는 꾹쳐 당 $O(m)$ 의 계산 복잡도를 갖는다. 꾹쳐 내 각 CTU에 대하여 투영된 표면 간 최소 거리를 갖는 참조 CTU를 탐색하는 과정은, 참조 프레임에서 색상 정보가 한 픽셀 이상 존재하는 모든 CTU에 대하여 거리 비교를 수행하기 때문에 CTU마다 최대 $O(m)$ 의 계산 복잡도를 갖으며 한 꾹쳐 당 최대 $O(m^2)$ 의 계산 복잡도를 갖는다.

V. 결 론

본 논문에서는 동적 메쉬 기반의 텍스쳐 맵을 제한된 비트에서 압축할 때 시공간적으로 일관적인 화질로 압축하기 위한 율 제어 방법을 제안하였다. 텍스쳐 맵을 압축하기 위해 사용되는 기존 2D 비디오 코덱의 율 제어 방법은 영상의 시간적 연속성을 가정하고 설계되었기 때문에, 색상 정보 투영 과정에 의한 시간적 불연속 발생할 수 있는 텍스쳐 맵을 목표하는 비트와 일관적인 화질로 압축하는데 한계가 있다. 제안하는 방법은 텍스쳐 맵의 시간적 연속성을 판단하여 각 프레임을 일관적인 화질로 압축할 수 있는 비트를 할당한다. 시간적 연속성의 변화에 따른 꾹쳐 단위의 율-왜곡 특성 변화에 강인한 비트 할당을 수행하고자, 제한된 시퀀스 수준의 비트 조건에서 시간적 연속 정도가 서로 다른 각 꾹쳐가 유사한 화질로 압축될 수 있는 평균적인 비트를 실험적인 도출하여 사용하였다. 또한 제안하는 방법은 한

프레임을 공간적으로 일관적인 화질로 압축하기 위하여 CTU 단위의 율-왜곡 성능을 이전 프레임의 의미적으로 유사한 CTU를 참조하여 예측한다. 제안하는 방법은 율-왜곡 성능의 정확한 예측을 통해 비트 오차율 관점에서 시퀀스와 꾹쳐 레벨에서 각각 평균 0.02%, 2.6%의 오차율을 달성하였다. 또한 꾹쳐와 CTU 수준에서 효율적인 비트 할당을 통하여, 3D 정점과 렌더링된 2D 이미지의 복원된 흐도 성분에 대하여 각각 12.3%와 8.7%의 BD-rate 향상을 달성하였고, 각각 0.5dB, 0.4dB의 PNSR의 표준 편차를 보여 객관적 화질의 시간적 일관성 향상을 달성하였다.

참 고 문 헌 (References)

- [1] C. Chu, J. Park, H. Kim, J. Park, S. Lim and B. Koo, "Recent Trends of 3D Reconstruction Technology," *Electronics and Telecommunications Trends*, vol. 22, no. 4, pp. 1-11, Aug. 2007.
doi: <https://doi.org/10.22648/ETRI.2007.J.220401>
- [2] S. Lee, B. Hwang, S. Lim, S. Yoon, T. Kim, K. Kim, D. Kim, C. Park, "Recent Trends and Prospects of 3D Content Using Artificial Intelligence Technology," *Electronics and Telecommunications Trends*, vol. 34, no. 4, pp. 15-22, Aug. 2019.
doi: <https://doi.org/10.22648/ETRI.2019.J.340402>
- [3] M. Deering, "Geometry compression," In *Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques*, New York, NY, United States, pp. 13 - 20, Sep. 1995.
- [4] H. Lopes, G. Tavares, J. Rossignac, A. Szymczak and A. Safanova, "Edgebreaker: A simple compression for surfaces with handles," In *Proceedings of the 7th ACM Symposium on Solid Modeling and Applications*, New York, NY, United States, pp. 289 - 296, Jun. 2002.
doi: <https://doi.org/10.1145/566282.566324>
- [5] C. Touma and C. Gotsman, "Triangle mesh compression," In *Proceedings of the 1998 Graphics Interface Conference*, Vancouver, British Columbia, Canada, pp. 26 - 34, Jun. 1998.
- [6] C. Courbet and C. Hudelot, "Taylor prediction for mesh geometry compression," *Computer Graphics Forum*, vol. 30, no. 1, pp. 139 - 151, Jan. 2011.
doi: <https://doi.org/10.1111/j.1467-8659.2010.01838.x>
- [7] M. Isenburg and P. Alliez, "Compressing polygon mesh geometry with parallelogram prediction," In *Proceedings of the Conference on Visualization*, Boston, MA, USA, pp. 141 - 146, Oct. 2002.
doi: <https://doi.org/10.1109/VISUAL.2002.1183768>
- [8] H. Hoppe, "Progressive meshes," In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, New York, NY, United States, pp. 99 - 108, Aug. 1996.
doi: <https://doi.org/10.1145/3596711.3596725>
- [9] R. Pajarola and J. Rossignac, "Compressed progressive meshes," *IEEE Transactions on Visualization and Computer Graphics*, vol. 6, no. 1,

- pp.79 - 93. Aug. 2002.
doi: <https://doi.org/10.1109/2945.841122>
- [10] N. Kwon, J. Byeon, H. Choi and D. Sim, "SHVC-based Texture Map Coding for Scalable Dynamic Mesh Compression," *Journal of Broadcast Engineering*, vol. 28, no. 3, pp. 314-328, May. 2023.
doi: <https://doi.org/10.5909/JBE.2023.28.3.314>
- [11] M. Kim, J. Byeon and D. Sim, "Valence based Lifting Wavelet Transform for Video-based Dynamic Mesh Compression," *Journal of Broadcast Engineering*, vol.29, no.1, pp. 42-56, Jan. 2024.
doi: <https://doi.org/10.5909/JBE.2024.29.1.42>
- [12] J. Zhang and C. Owen, "Octree-based animated geometry compression," *Computers and Graphics*, vol. 31, no. 3, pp.463 - 479. Jun. 2007.
doi: <https://doi.org/10.1016/j.cag.2006.12.002>
- [13] M. Bici and G. Akar "Improved prediction methods for scalable predictive animated mesh compression," *Journal of Visual Communication and Image Representation*, vol. 22, no. 7, pp.577 - 589, Oct. 2011.
doi: <https://doi.org/10.1016/j.jvcir.2011.07.006>
- [14] D. Graziosi and K. Hayashi, "Dynamic Mesh Coding Using Orthogonal Atlas Projection," 2024 Picture Coding Symposium (PCS), Taichung, Taiwan, pp. 1-5. Jun. 2024.
doi: <https://doi.org/10.1109/PCS60826.2024.10566351>
- [15] WD 1.0 of V-DMC, ISO/IEC JTC/SC29/WG7, MDS22184, Oct.2022
- [16] G. Sullivan, J. Ohm, W. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *Institute of Electrical and Electronics Engineers (IEEE) Transactions on circuits and systems for video technology*, vol.22, no.12, pp. 1649-1668, Dec. 2012.
doi: <https://doi.org/10.1109/TCSVT.2012.2221191>
- [17] B. Bross, J. Chen, S. Liu, and Y.-K. Wang, "Versatile Video Coding (Draft 10)," JVET-S2001, July. 2020.
- [18] M. Lee, H. Song, J. Park, B. Jeon, J. Kang, J. Kim, Y. Lee, J. Kang and D. Sim, "Overview of versatile video coding (H. 266/VVC) and its coding performance analysis," *IEIE Transactions on Smart Processing & Computing*, vol. 12, no. 2, pp. 122-154, Apr. 2023.
doi: <https://doi.org/10.5573/IEIESPC.2023.12.2.122>
- [19] J. Lee, J. Park, H. Choi, J. Byeon and D. Sim, "Overview of VVC," *Broadcasting and Media Magazine*, vol. 24, no. 4, pp.10-25, Oct. 2019.
- [20] B. Li, H. Li, L. Li and J. Zhang, " λ Domain Rate Control Algorithm for High Efficiency Video Coding," *IEEE Transactions on Image Processing*, vol. 23, no. 9, pp.3841-3854, Sep. 2014.
doi: <https://doi.org/10.1109/TIP.2014.2336550>
- [21] J. Rossignac, "Edgebreaker: Connectivity compression for triangle meshes," *IEEE Trans. Vis. Comput. Graphics*, vol. 5, no. 1, pp. 47-61, Jan. 1999.
doi: <https://doi.org/10.1109/2945.764870>
- [22] Draco, <https://github.com/google/draco>
- [23] UVAtlas, <https://github.com/microsoft/UVAtlas>
- [24] V-PCC Codec Description, document ISO/IEC JTC1/SC29/WG07N 00010, MPEG WG7, Online, Oct. 2020.
- [25] Algorithm Description of FCTM, ISO/IEC JTC 1/SC 29/WG 4 N00460, MPEG WG4, Online, Jan. 2024.
- [26] H. Choi, J. Nam, J. Yoo, D. Sim and I. Bajic, "Rate Control Based on Unified RQ Model for HEVC," document Rec. JCTVC-H0213, San Jose, CA, USA, Feb. 2012.
- [27] Z. He, Y. Kim and S. Mitra, "Low-delay rate control for DCT video coding via ρ -domain source modeling," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 8, pp. 928 - 940, Aug. 2001.
doi: <https://doi.org/10.1109/76.937431>
- [28] I. Marzuki, J. Lee and D. Sim, "Optimal CTU-level rate control model for HEVC based on deep convolutional features," *IEEE Access*, vol. 8, pp. 165670-165682, Sep. 2020.
doi: <https://doi.org/10.1109/ACCESS.2020.3022408>
- [29] W. Lim and D. Sim, "A perceptual rate control algorithm based on luminance adaptation for HEVC encoders," *Signal, Image and Video Processing*, vol. 14, no. 5, pp. 887-895. Jan. 2020.
doi: <https://doi.org/10.1007/s11760-019-01620-3>
- [30] I. Marzuki, J. Lee, W. Wiratama, D. Sim, "Deep convolutional feature-driven rate control for the HEVC encoders," *IEEE Access*, vol. 9, pp. 162018-162034, Dec. 2021.
doi: <https://doi.org/10.1109/ACCESS.2021.3132294>
- [31] HM, <https://vcgit.hhi.fraunhofer.de/jvet/HM>
- [32] A. Ortega and K. Ramchandran, "Rate-distortion methods for image and video compression," *IEEE Signal Process. Mag.*, vol. 15, no. 6, pp. 23 - 50, Nov. 1998.
doi: <https://doi.org/10.1109/79.733495>
- [33] G. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," *IEEE Signal Process. Mag.*, vol. 15, no. 6, pp. 74 - 90, Nov. 1998.
doi: <https://doi.org/10.1109/79.733497>
- [34] B. Li, J. Xu, D. Zhang and H. Li, "QP Refinement According to Lagrange Multiplier for High Efficiency Video Coding," In 2013 IEEE International Symposium on Circuits and Systems (ISCAS), Beijing, China, pp. 477-480, May. 2013.
doi: <https://doi.org/10.1109/ISCAS.2013.6571884>
- [35] TMM, <http://mpegx.int-evry.fr/software/MPEG/dmc/mpeg-vmesh-tm>
- [36] Common Test Conditions for V-DMC, document ISO/IEC JTC1/SC29/WG07 N00555, MPEG WG7, Online, Jan. 2023.
- [37] G. Bjontegaard, Calculation of Average PSNR Differences Between RD-Curves, document Rec. VCEG-M33, Austin, TX, USA, Apr. 2001.
- [38] V-DMC codec description, document ISO/IEC JTC1/SC29/WG07 N01011, MPEG 3DG, Kemer, Turkey, Nov. 2024.

저자소개



최한솔

- 2018년 2월 : 광운대학교 컴퓨터공학과 학사
- 2020년 2월 : 광운대학교 컴퓨터공학과 석사
- 2020년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 박사과정
- ORCID : <https://orcid.org/0000-0001-7947-0723>
- 주관심분야 : 영상신호처리, 영상압축, 컴퓨터비전



심동규

- 1993년 2월 : 서강대학교 전자공학과 공학사
- 1995년 2월 : 서강대학교 전자공학과 공학석사
- 1999년 2월 : 서강대학교 전자공학과 공학박사
- 1999년 3월 ~ 2000년 8월 : 현대전자 선임연구원
- 2000년 9월 ~ 2002년 3월 : 비로비전 선임연구원
- 2002년 4월 ~ 2005년 2월 : University of Washington Senior research engineer
- 2005년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 교수
- ORCID : <https://orcid.org/0000-0002-2794-9932>
- 주관심분야 : 영상신호처리, 영상압축, 컴퓨터비전