

<https://doi.org/10.7236/JIIBC.2020.20.1.113>
JIIBC 2020-1-16

계층적 깊이 영상으로 표현된 다시점 영상에 대한 H.264/AVC 비디오 압축 표준에 관한 연구

A Study on H.264/AVC Video Compression Standard of Multi-view Image Expressed by Layered Depth Image

지인호*

Innho Jee*

요약 다시점 비디오는 하나의 장면을 여러 대의 카메라를 이용하여 여러 시점에서 촬영한 것이다. 따라서 사용자의 요구에 의해서 원하는 시점의 비디오를 제공할 수 있는 장점을 갖는다. 본 논문에서는 보다 향상된 기법을 활용하여 계층적 깊이 영상 구조 표현의 압축 성능을 향상시켰다. 계층적 깊이 영상을 H.264 기술로 부호화시켜 압축된 데이터 크기를 확인하고, 복원된 각 영상의 품질 성능을 알아보았다. H.264/AVC 기술은 쉽게 비디오와 관련된 콘텐츠에 대한 H.264 기술로 확장될 수 있다. 그래서 깊이 정보를 포함하는 다시점 영상을 효과적으로 압축할 수 있는 계층적 깊이 영상 구조라는 새로운 콘텐츠에 적용하는 방법을 제안하였다. 다시점 비디오 영상의 막대한 데이터 양이 감소되며, 고품질의 영상을 제공되고, 에러 복원 기능이 강화하는 장점이 있는 것을 보여주었다.

Abstract The multi-view video is a collection of multiple videos capturing the same scene at different viewpoints. Thus, there is an advantage of providing for user oriented view pointed video. This paper is suggested that the compression performance of layered depth image structure expression has improved by using more improved method. We confirm the data size of layer depth image by encoding H.264 technology and the each performances of reconstructed images. The H.264/AVC technology has easily extended for H.264 technology of video contents. In this paper, we suggested that layered depth structure can be applied for an efficient new image contents. We show that the huge data size of multi-view video image is decreased, and the higher performance of image is provided, and there is an advantage of for stressing error restoring.

Key Words : Multi-view Image, Layer Depth Image, H.264/AVC, 3-D Warping

1

1. 서 론

다시점 비디오(Multi-view video)는 한 장면을 여러

위치의 시점에서 다수의 카메라로 동시에 촬영한 것으로, 이 비디오 콘텐츠를 이용하면 사용자들이 요구하는 다양한 시점의 영상을 제공 할 수 있다. 그러나 많은 시점을

*정회원, 홍익대학교 소프트웨어융합학과
접수일자: 2019년 12월 14일, 수정완료: 2020년 1월 14일
게재확정일자: 2020년 2월 7일

Received: 14 December, 2019 / Revised: 14 January, 2020 /
Accepted: 7 February, 2020
Corresponding Author: ijee@hongik.ac.kr
Dept. of Software and Communications Engineering, Hongik
University, KOREA

표현하기 위해서는 그 만큼의 카메라의 수가 필요하므로 비용적인 측면과 기술적인 측면에서 어려움이 존재한다. 그래서 주어진 시점들을 이용하여 사용자의 요구에 맞는 시점을 생성하는 기술이 필요하며, 이 기술을 구현하기 위해서는 각 시점에서의 깊이 영상(depth image) 정보가 유용하게 사용될 수 있다. 깊이 영상을 이용하면, 촬영된 이외의 시점을 보간(interpolation)할 수 있으며 또한 다시점 영상에 대하여 입체 화면의 구성이 가능하게 한다. 지금도 많은 연구가 진행되고 있는 스테레오 정합 기법은 깊이 영상을 생성하는 전통적 기술에 해당하며, 최근에는 깊이 카메라를 사용하는 방법이 있다⁹⁾.

결과적으로 3차원 TV (3D TV), 자유시점 TV (Free viewpoint TV) 등의 분야에서 다시점 비디오가 보다 효과적으로 활용되기 위해서는 컬러 영상뿐만 아니라 깊이 영상의 사용이 필수적이다. 그러나 다시점 비디오는 카메라 수만큼의 컬러 영상들과 깊이 영상들이 존재하므로 막대한 데이터의 양을 갖게 되어서, 그 양을 줄이는 새로운 압축 부호화 기술의 개발이 필수적이게 된다⁹⁾. 최근, 계층적 깊이 영상(Layered Depth Image)의 개념을 이용하여 깊이 정보를 포함하는 다시점 비디오를 효과적으로 압축 부호화하는 기술이 주목받고 있다. 이 방법은 여러 시점의 깊이 정보와 워핑(warping) 함수를 사용하여 다시점의 컬러와 깊이 영상을 합성하여 하나의 데이터 구조로 만드는 방법이다. 본 논문에서는 계층적 깊이 영상에서 실제 거리를 이용하고 오버랩(overlap) 문제를 해결하고, 그리고 컬러 변환을 기반으로 복원을 위해서 추가되는 데이터의 양을 감소시켜 압축 성능을 향상시키는 방법을 제안하였다^{9)[10]}.

H.264는 비디오 압축 기술을 새로운 차원으로 보내준다. H.264에서는 I-프레임 인코딩을 위하여 새로운 첨단 예측 방식이 도입되고, 이 방식은 I-프레임의 비트 사이즈를 상당히 줄일 수 있으며, 프레임의 각 매크로 블록 안에 있는 더 작은 픽셀 블록들을 연속적으로 예측할 수 있게 함으로써 높은 품질을 유지할 수 있다.

본 논문의 구성으로, 2장에서는 계층적 깊이 영상 표현에 의한 다시점 영상에 대한 효과적인 압축 기술을 설명하고, 3장에서는 H.264/AVC 비디오 압축 표준 기술을 설명하였다. 그리고 4장에서는 제안된 다시점 영상에 대한 압축과 H.264 부호화 기술 실험의 결과에 대한 검토를 제시하였다. 마지막 5장에서는 본 연구에 대한 결론 및 향후 연구과제를 제시하였다.

II. 다시점 컬러 영상과 깊이 영상으로부터 계층적 깊이 영상 생성

1. 계층적 깊이 영상의 개념

계층적 깊이 영상은 복잡한 기하학적 정보를 갖는 3차원 물체나 장면을 영상기반 렌더링 기법을 이용하여 표현하는 방법 중의 하나이다. 그래서 이 방법은 여러 시점에서 얻은 다수의 깊이 영상을 합성하여 하나의 데이터 구조를 생성한다. 각 계층적 깊이 영상 화소는 색상정보 외에 화소와 카메라 사이의 거리를 나타내는 깊이 정보와 계층적 깊이 영상의 렌더링을 지원하는 추가적인 특성 정보를 가지고 있다. 또한, 계층적 깊이 영상은 각 화소의 위치마다 다수의 계층을 가지므로, 이 정보를 이용하면 새로운 카메라 위치에서 3차원 물체를 효과적으로 렌더링 할 수 있다. 그림 1^[10]은 이러한 특징을 갖는 계층적 깊이 영상의 생성 방식을 나타내고 있다. 그림 1^[10]에서 보는 바와 같이, 계층적 깊이 영상을 생성하기 위해서 기준시점에서 물체를 향해 광선을 투과시킨다. 이때, 광선 A는 물체와 두 곳에서 만나게 되며, 광선 B는 네 곳에서 만난다. 따라서 광선 A가 통과하는 계층적 깊이 영상 화소에는 2개의 계층이 생성되며, 광선 B가 통과하는 화소에는 4개의 계층이 형성된다. 따라서 모든 화소 위치마다 같은 수의 계층이 형성되는 것이 아니라, 광선과 물체의 교차점의 갯수에 따라 생성되는 계층의 수가 달라진다. 그리고 각 교차점에서 색상과 깊이 정보를 모두 저장하여 하나의 구조로 만들면 계층적 깊이 영상이 생성된다^{[2][5][8]}.

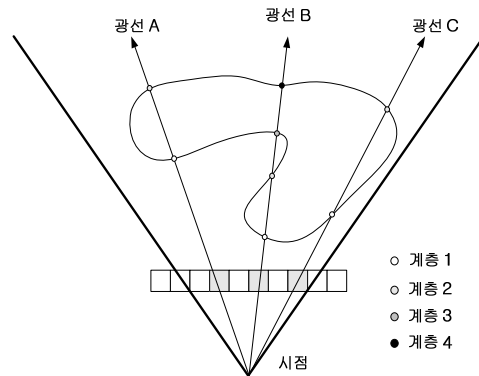


그림 1. 계층적 깊이 영상 생성의 개념도

Fig. 1. The diagram of layered depth image generation

2. 실사 영상으로부터의 계층적 깊이 영상의 생성과 데이터 구조

실사 영상에서의 카메라 행렬은 주어진 카메라 매개변수로부터 계산된다. 그래서 수정된 카메라 행렬과 3차원 워핑식은 다음과 같다^{[2][3]}.

$$C'_i = A_i \cdot E_i, \quad C'_j = A_j \cdot E_j \quad (1)$$

여기서, C'_i 는 i 번째 카메라의 카메라 행렬이다. 그리고 A 는 카메라의 내부 매개변수를 나타내는 행렬이고, E 는 외부 매개변수를 나타내는 행렬이다.

$$A = \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \quad E = \begin{bmatrix} R_{11} & R_{12} & R_{13} & T_1 \\ R_{21} & R_{22} & R_{23} & T_2 \\ R_{31} & R_{32} & R_{33} & T_3 \end{bmatrix}$$

A 와 E 의 행렬의 정의는 다음과 같다. 여기서 α_x , α_y 는 초점거리이고, s 는 비틀어짐(skew)정도를 나타내며, (x_0, y_0) 은 카메라 중심의 위치이다. 그리고 $R_{3 \times 3}$ 은 회전행렬, $T_{3 \times 1}$ 는 이동행렬이다. 이때, C'_i 는 4×4 행렬이 되어야 하기 때문에 $A \cdot E$ 에 [0001]인 4번째 행을 추가한다.

다시점의 컬러 영상과 깊이 영상을 사용하여 계층적 깊이 영상을 생성하는 예를 그림 2^[10]에 나타내었다^[11].

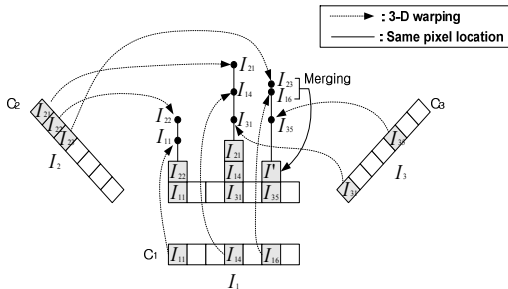


그림 2. 다시점의 컬러 영상과 깊이 영상으로부터 계층적 깊이 영상의 생성

Fig. 2. The multi-view color image and the generation of layered depth image from depth image

먼저, 그림 2^[10]에서는 다수의 깊이 영상을 이용한 계층적 깊이 영상 생성을 위해 다른 위치의 세 개의 카메라 $I_i(i=1,2,3)$ 가 존재하고 C_1, C_2, C_3 의 영상을 갖는다. 그리고 I_{11} 는 첫 번째 카메라 I_1 의 C_1 에서 첫 번째 화소를 나타낸다. I_{21}, I_{31} 도 마찬가지이다. 두 번째로 사용자의 의

해서 기준 영상이 선정된다. 여기에서는 C_1 을 기준 영상으로 선택한다. 그리고 나머지 영상들은 기준 영상으로 3차원 워핑을 수행한다. 세 번째로 컬러와 깊이 값을 갖는 워핑 된 화소들은 깊이 값에 근거하여 각 화소의 위치에 저장된다. 만약, 워핑 된 화소들 사이의 깊이 값 차이가 미리 정해 놓은 임계값 보다 적으면, 이것들은 하나의 화소로 합병된다. 그림 2^[10]에서 'l'가 이 경우에 해당되며, 깊이 값은 평균 된다. 반대로 임계값 보다 크게 되면 새로운 계층이 생성된다^{[9][10]}.

컴퓨터 그래픽스를 이용해 생성된 모델이 아닌 실제 물체에 대해서는 광선이 실제 물체를 투과할 수 없기 때문에 앞서 언급한 방식을 그대로 적용할 수가 없다. 따라서 실사 영상으로부터 계층적 깊이 영상을 생성하기 위해서는, 여러 시점에서 깊이 영상을 획득하여 이를 합성하는 방식을 사용한다.

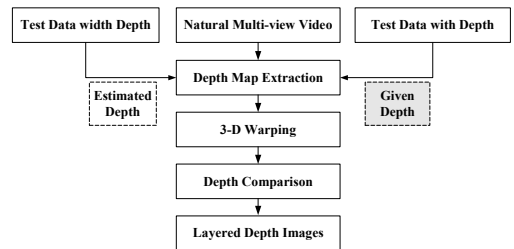


그림 3. 실사 다시점 비디오로부터 계층적 깊이 영상의 생성 과정
 Fig. 3. The generation procedure of layered depth image from real multi-view video

그림 3^[10]에서 계층적 깊이 영상을 생성하는데 필요한 정보인 깊이 정보를 제공하는 데이터는 Microsoft Research(MSR)에서 데이터가 유일하다. MSR에서는 평행한 8대의 카메라로부터 획득한 다시점 비디오를 제공하고 있으며, MSR에서 배포하는 테스트 데이터는 카메라별 색상뿐 아니라 깊이 영상 및 카메라 매개변수를 포함하고 있다^[14]. 이러한 실사 다시점 비디오 시퀀스로부터 컬러와 깊이 영상을 갖는 계층적 깊이 영상을 생성하는 전 과정을 나타낸 것이다^[11]. 다시점 비디오로부터 깊이 정보를 얻은 후, 기준시점으로 시점 변환을 수행하는 3차원 워핑이 수행된다. 그 다음 기준시점과 기준시점으로 변환된 여러 시점들 간의 중복성을 제거하면서 계층적 깊이 영상의 데이터 구조를 형성한다^{[9][10]}.

3. 계층적 깊이 영상으로부터 다시점 영상의 복원

워핑 영상처리는 비선형의 변환으로 변환 전과 변환 후의 화소들이 일대일 대응을 보장하지 못한다. 다시점 영상으로 복원 과정에서는 역 워핑이 수행되므로 홀(hole)과 오버랩(overlap)문제가 발생한다. 따라서 재생된 시점 영상들의 품질을 보장하기 위해서는 홀과 오버랩의 문제를 해결하여야만 한다^{[9][10]}. 계층적 깊이 영상으로부터 다시점 영상을 복원하는 과정을 그림 4^[10]에 나타내었다. 우선적으로 계층적 깊이 영상은 각각의 워핑된 영상들로 분해된다. 그리고 역 워핑을 통해서 원래의 시점으로 이동하게 된다. 역 워핑식은 식(1)에서 구해진 것을 이용하여 얻어진다. 워핑에서와 마찬가지로 폐색 영역으로 인한 홀(hole)문제와 그리고 오버랩문제가 발생하게 된다. 오버랩문제는 부호화 과정과 동일하게 처리한다. 그래서 깊이 값에 근거하여 전경은 남기고 배경은 버리면 된다. 다음 단계로 홀 문제를 해결하기 위해서 보간(interpolation)을 수행한다. 주변 화소를 이용하여 홀 부분을 채운 보간 시점 영상은 원래의 시점 영상과의 차이가 줄어들게 된다. 보간 과정은 계층적 깊이 영상을 생성하는 과정에서도 수행된다. 그 이유는 두 영상 간 오차를 전송하는 추가 데이터(residual data)를 생성하기 위해서이다. 따라서 보간은 압축의 효율을 증가시키는 한 요인이 된다. 마지막 단계에서는 이렇게 계층적 깊이 영상 생성 단계에서 만들어진 추가 데이터를 더해서 우수한 품질의 복원된 다시점 영상을 생성하게 된다^{[11][2][4]}.

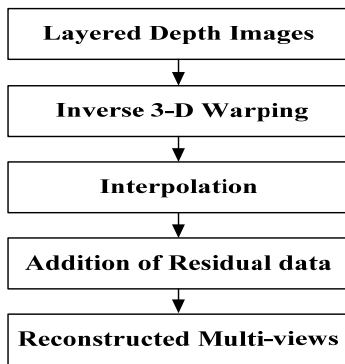


그림 4. 다시점 영상들의 재생과정
Fig. 4. Reconstruction procedure of multi-view images

III. H.264/AVC 비디오 압축표준

H.264는 ITU-T(국제전기통신연합), ISO(국제표준화 기구), IEC(국제전기표준회의)가 공동으로 추진하는 차

세대 비디오 압축표준으로 MPEG-4 파트 10 또는 MPEG-4 AVC(Advanced Video Codec)으로 불린다. MPEG-2에 비해 3배 이상, MPEG-4에 비해 2배 이상의 데이터 압축능력을 제공하는 기술이다^{[11][13]}.

국내 DMB방송은 물론 블루레이(Blu-ray) 및 HD-DVD와 같은 차세대 고화질 동영상을 위한 압축 코덱으로, 기존 동영상 압축 기술에 비해 2 ~ 3배의 압축 효율과 함께 화질도 크게 개선됐다^[13].

비디오 압축은 디지털 비디오 파일이 효율적으로 전송되고 저장될 수 있도록 중복된 비디오 데이터를 줄이고 제거하는 것이다. 프로세스에는 소스 비디오 알고리즘을 적용하여 즉시 전송과 저장이 가능한 압축파일을 생성하는 과정이 포함된다. 압축된 파일을 재생하기 위해서는 역 알고리즘을 적용하여 원본 비디오와 실질적으로 동일한 내용을 보여주는 비디오를 생성한다. 파일을 압축, 전송, 해제 및 디스플레이 하는데 소요되는 시간이 주어진 동일한 프로세싱 능력에서는 압축 알고리즘이 고도화될수록 응답속도가 높아진다^[13].

1. 비디오 코딩 레이어

VCL(Video Coding Layer)는 영상 자료 콘텐츠를 효율적으로 표현하는 것으로 정해져 있다. H.264의 VCL은 MPEG-2 비디오 압축 기술과 비슷한데, 변환 코딩을 위해 시간적 예측 및 공간적 예측의 복합으로 구성된다. 그림 5^{[11][12][13]}은 매크로 블록에 대한 것이다. 입력 비디오는 여러 개의 블록으로 쪼개진다. 연결된 첫 번째의 화상이나 랜덤 액세스 점이 일반적으로 'Intra'코드가 된다. 다시 말해, 화상 그 자체를 포함하는 것으로 다른 정보를 사용하지 않는다. Intra 프레임에 있는 각각의 블록은 이전에 코드화된 블록의 이웃하는 샘플들을 공간적으로 사용함으로써 예측되어진다. 어떻게 어떤 이웃하는 샘플들을 Intra 예측에 사용할 것인지를 인코딩 처리 과정에서 채택된다. 동시에 인코더와 디코더에서 Intra 정보를 전송한다. 시퀀스의 남은 화상들이나 랜덤 액세스 점 사이에 남아 있는 모든 것들이 'Inter'코딩에 사용된다. Inter 코딩은 이전에 디코딩된 다른 화상으로부터 예측된다(모션 보상). Inter 예측(모션 평가)에 대한 인코딩 처리 과정은 채택된 모션 데이터, 절충된 기준 화상과 블록의 모든 샘플에 적용된 공간 배치로 구성된다. 사이드 정보로 전송된 모션 데이터는 Inter 예측 신호로 제공되기 위해 인코더와 디코더에 동시에 사용된다.

원래 화상과 예측된 블록 사이의 차이점인 Intra와 Inter 중 하나는 예측에 사용되고, 나머지는 변환된다. 변

환 계수는 스케일화 되고 양자화되는데, 양자화된 변환 계수는 엔트로피 코딩되어 Intra 프레임이나 Inter 프레임 예측에 대한 사이드 정보와 함께 전송된다. 인코더는 다음 블록이나 다음 화상에 대한 예측을 수행하기 위해 디코더에 저장된다. 그러므로 양자화된 변환 계수는 디코더 측면에서 같은 방법으로 역 스케일화 되고 역변환된다. 결과적으로 디코더 예측 오차가 생기게 된다. 이런 디코더 예측 오차는 이전 화상에 더해지고, 더해진 결과는 디코더 비디오 출력으로 제공되는 디블록킹 필터에 공급된다.

또한, DCT 계수를 양자화한 후의 신호를 엔트로피 부호화할 때 주변 정보에 따라 적응적으로 부호 표를 선택하는 방식으로 영상의 특성에 맞게 부호화하여 압축 효율을 높인다. 이러한 기술들을 이용하는 H.264/AVC 표준은 MPEG-4 비디오 부호화 표준에 비해 압축 성능이 약 2배 정도 좋다^[13].

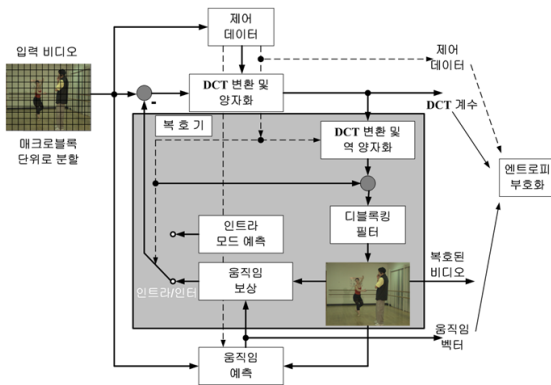
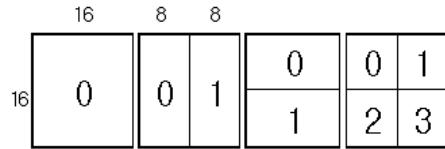


그림 5. H.264의 기본 코딩 구조
 Fig. 5. The basic encoding structure of H.264

2. P 슬라이스에서의 모션 보정

Intra 매크로 블록코딩 타입에, 다양한 예측적 또는 모션 보정 코딩 유형은 P 슬라이스 매크로 블록으로 지정된다. 각각의 P 형태 매크로 블록은 모션 설명을 위해 사용되는 고정된 크기의 블록들과 일치한다. 16×16, 16×8, 8×16, 8×8 샘플의 휘도 블록 크기를 가진 분할된 부분들은 각각 Intra-16×16, Intra-16×8, Intra-8×16, Intra-8×8의 P 매크로 블록 형태와 일치된 syntax로 제공된다. Intra-8×8 매크로 블록 모드가 선택된 경우에 각 8×8 서브 매크로 블록에 대한 syntax 요소에 첨가해 전송된다. 이런 syntax 요소는 일치된 서브 매크로 블록들인 8×8, 8×4, 4×8, 4×4 샘플의 휘도 블록 크기를 갖는 모션 보정 예측을 사용하여 코드화된다. 그림 6^[12]는 블록의 분할을 설명한다^[13].

1 매크로 블록 분할 16×16, 16×8, 8×16, 8×8의 경우



2 매크로 블록 서브 분할 8×8, 8×4, 4×8, 4×4의 경우

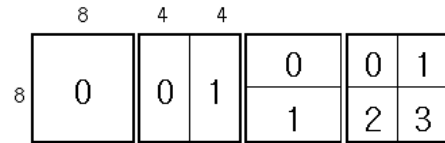


그림 6. 모션 보정에 대한 매크로 블록의 분할
 Fig. 6. The division of macro blocks for motion compensation

3. B 슬라이스에서의 모션 보정

이전의 비디오 코딩 표준과 비교해 보면, 그림 7^[12]는 B 슬라이스의 개념을 H.264로 일반화되었다. 다른 화상들은 모션 보정 예측을 위한 B 화상을 기준으로 할 수 있다. 이런 예측은 멀티 화상 버퍼링의 메모리 관리 제어 동작에 의존하게 된다. 그리하여 B 슬라이스와 P 슬라이스 사이의 근본적인 차이점은 B 슬라이스는 어떤 의미에서는 예측 신호를 갖는 2개의 명확한 모션 보정 예측값의 평균치를 사용하는 매크로 블록이나 블록들로 코드화된다. 일반적으로 B 슬라이스들은 각기 첫 번째와 두 번째의 기준 화상 버퍼로 언급되는 2개의 명백한 기준 화상 버퍼들을 이용한다. 실제로 각각 기준 화상 버퍼에 위치한 어떤 화상들은 멀티 화상 버퍼에 대한 결과이다. 그리고 동작 형태는 이미 잘 알려진 MPEG-2의 B 화상 예측 방법과 비슷하다^[13].

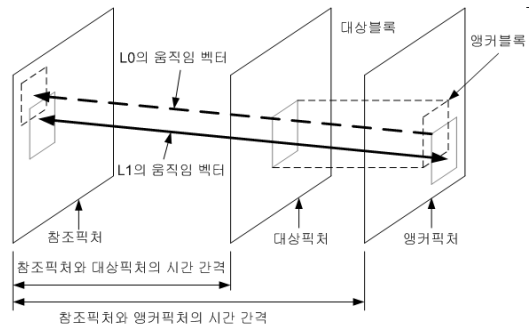


그림 7. B 픽처 시간 예측 모드
 Fig. 7. The time prediction mode of B picture

4. 엔트로피 코딩

H.264는 엔트로피 코딩에 2가지 방법이 지원된다. 엔트로피 코딩 방법은 양자화된 변환 계수 값을 제외하고 모든 syntax 요소들에 대해 싱글 무한 확장 코드 워드를 사용한다. 그리하여 각 syntax 요소에 대한 다른 VLC 표를 설계하는 대신에 단지 싱글 코드 워드 표를 매핑하여 데이터 통계에 따라서 주문을 받는다. 싱글 코드 워드 표는 매우 단순하고 규칙적인 디코딩 특성을 가진 exp-Golomb 코드이다.

양자화된 변환 계수 값을 전송하기 위해 소위 CAVLC (Context-Adaptive Variable Length Coding)이라고 불리는 좀더 복잡한 방법을 사용한다. 대략적으로 여러 가지 syntax 요소에 대한 VLC 표는 스위칭 되고, 이미 전송된 syntax 요소들에 의존한다. VLC 표가 제한적인 통계에 부합되도록 잘 설계되었기 때문에 엔트로피 코딩의 성능은 단지 싱글 VLC 표를 사용하는 것과 비교하여 개선된다. 만약 엔트로피 코딩에서 그림 8^[12]은 CABAC (Context-Adaptive Binary Arithmetic Coding)을 사용한다면 효율이 좀 더 개선될 수 있다.

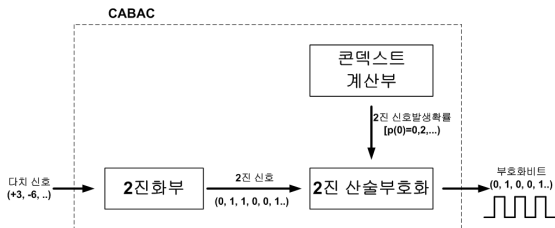


그림 8. CABAC
Fig. 8. CABAC

IV. 실험 및 결과

계층적 깊이 영상의 압축 성능과 H.264/AVC의 기술에 대한 성능을 평가하기 위해서 두 가지 실험을 진행하였다. 첫 번째 실험에서는 계층적 깊이 영상을 H.264 부호화 기술에 적용시켜 영상의 압축 정도와 압축된 데이터를 다시 복원해 온 영상과의 품질을 비교해 보았다. 마지막으로 워핑 된 계층적 깊이 영상을 역 워핑 하여 복원 영상 성능 실험을 하였다.

1. 계층적 깊이 영상을 H.264 압축과 복원 성능 실험
본 논문에서는 계층적 깊이 영상 표현을 통한 H.264/AVC

압축 성능과 이 압축에서 복원되었을 때의 다시점 영상들의 품질을 PSNR(Peak Signal to Noise)로 성능 측정하였다. 그리고 실험에서 사용된 다시점 비디오는 Microsoft Research(MSR)에서 제공한 "Breakdancers"와 "Ballet"를 사용하였다^{[6][7]}. 두 비디오는 깊이 정보와 카메라 변수가 제공되고 있으며, 비디오 품질은 XVGA급이며, 15 fps이고, 영상의 크기는 1024*768 사이즈이다. 그리고 1 차원 원호의 8대의 카메라에 의해서 촬영되었다.

그림 9과 10은 Breakdancers 다시점 비디오와 Ballet 다시점 비디오의 첫 번째 프레임의 나타낸 것으로 8대 카메라에서 생성된 각 시점의 컬러 영상과 깊이 영상을 보여준다.

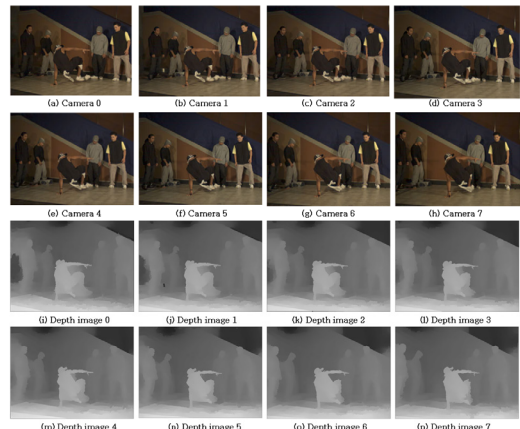


그림 9. Breakdancers의 다시점 영상
Fig. 9. The multi-view image of Breakdancers

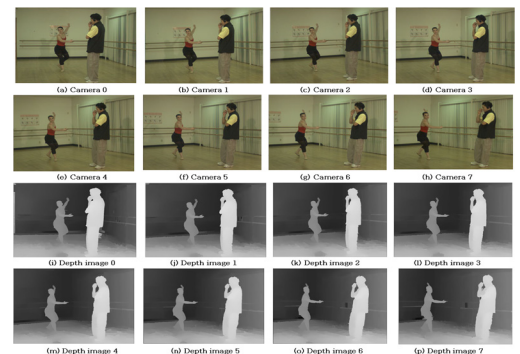


그림 10. Ballet의 다시점 영상
Fig. 10. The multi-view image of Ballet

2. H.264/AVC 압축으로부터 복원된 다시점 영상 품질에 관한 실험

이번 실험에서는 그림 9 - 10처럼 하나의 YUV로 만든 Breakdancers 와 Ballet 영상을 H.264 부호화시켜 압축한 영상의 데이터 크기를 알아보았다. 그리고 그 압축데이터를 다시 복원하여 원 영상과 복원된 각각의 color 영상과 depth 영상의 신호대잡음비 정도를 실험하였다. H.264 압축은 비교적 적은 데이터 값으로 나왔고, 복원된 각각의 영상의 품질은 좋은 성능을 나타내었다.

표 1. Breakdancers와 Ballet의 color영상복원의 PSNR성능
Table 1. The PSNR of reconstruction color image of Breakdancers and Ballet

	Breakdancers		Ballet	
	1 st Frames	2 nd Frames	1 st Frames	2 nd Frames
PSNR of Decoded camera 0 (color) [dB]	39.8	39.5	41.5	41.5
PSNR of Decoded camera 1 (color) [dB]	40.5	40.9	42.3	42.3
PSNR of Decoded camera 2 (color) [dB]	39.7	39.8	40.8	41.8
PSNR of Decoded camera 3 (color) [dB]	40.5	39.6	42.2	42.1
PSNR of Decoded camera 4 (color) [dB]	40.9	40.8	40.3	39.9
PSNR of Decoded camera 5 (color) [dB]	39.9	40.9	40.2	40.3
PSNR of Decoded camera 6 (color) [dB]	39.8	39.9	41.4	41.8
PSNR of Decoded camera 7 (color) [dB]	39.7	39.7	41.3	42.1

표 2. Breakdancers와 Ballet의 depth 영상복원의 PSNR성능
Table 2. The PSNR of reconstruction depth image of Breakdancers and Ballet

	Breakdancers		Ballet	
	1 st Frames	2 nd Frames	1 st Frames	2 nd Frames
PSNR of Decoded camera 0 (depth) [dB]	45.1	46.3	47.0	46.8
PSNR of Decoded camera 1 (depth) [dB]	46.0	46.3	48.6	47.9
PSNR of Decoded camera 2 (depth) [dB]	46.4	45.0	46.8	47.3
PSNR of Decoded camera 3 (depth) [dB]	45.6	46.7	46.0	46.8
PSNR of Decoded camera 4 (depth) [dB]	47.8	45.0	47.8	48.1
PSNR of Decoded camera 5 (depth) [dB]	45.0	47.3	47.0	47.2
PSNR of Decoded camera 6 (depth) [dB]	45.41	46.7	48.0	47.8
PSNR of Decoded camera 7 (depth) [dB]	46.8	45.8	47.9	47.5

V. 결 론

깊이 정보를 포함하는 다시점 영상은 각 시점마다 컬러 영상과 깊이 영상을 포함하므로, 데이터의 크기는 엄청난 양이 된다. 최근 이것을 효과적으로 압축하는 기술로 3차원 워핑을 이용한 계층적 깊이 영상 표현이 소개되었다. 계층적 깊이 영상은 각 시점의 영상을 기준시점으로 워핑 변환하고, 동일 위치의 워핑 된 각 시점의 화소끼리 거리 값 비교를 통해서 중복되는 화소를 제거하는 기술이다. H.264는 전송률에 있어서나 또는 잡음 및 효율성에 있어서, 현재의 MPEG-2 비디오와 같이 현존하는 표준과 비교해 볼 때 비트율 저장의 2가지 요인에서 중대한 개선을 달성하였다. 이와 같은 H.264는 기존의 MPEG-4에 비해 약 40% 정도, MPEG-2에 대해서는 약 60% 정도 이상의 월등한 압축 효율을 보여주는 향상된 압축 성능을 가진 새로운 최신의 영상 압축 기술이다. 그리고 계층적 깊이 영상을 H.264 부호기를 통해 압축된 데이터를 다시 복원했을 때의 성능을 PSNR의 수치로 확인할 수 있었다. 또 다른 실험으로는 임계값을 조정하여 3차원 워핑 된 영상을 Filling 보간 실험을 하였고, 보간 된 영상과 추가 데이터 영상을 더해진 보간 실험을 하였다. 계층적 깊이 영상 표현에 대한 H.264 실험 결과에서, 복원했을 때의 영상이 PSNR을 통해 우수함을 볼 수 있었다. H.264 부호기를 거친 다시점 비디오는 추가적인 압축이 가능하다. 계층적 깊이 영상의 각 계층 간의 높은 유사성과 프레임 간의 유사성을 이용하여 동영상 부호화를 적용하는 것이다. 그래서 계층적 깊이 영상을 H.264 부호화에 적용하는 기술의 개발과 연구가 필요하다.

References

- [1] S. U. Yoon, and Y. S. Ho, "Multiple Color and Depth Video Coding Using a Hierarchical Representation," IEEE Trans. on Image Processing, Vol. 17, pp. 1450-1460, Nov. 2007.
- [2] J. Shade, S. J. Gortler, L. W. He, and R. Szeliski, "Layered Depth Images," in Proc. SIGGRAPH'98, pp. 231-242, Jul. 1999
- [3] S. U. Yoon, S. Y. Kim, E. K. Lee, and Y. S. Ho, "A Framework for Multi-view Video Coding Using Layered Depth Images," Lecture Notes Comput. Sci. (LNCS), Vol. 3767, pp. 431-442, Nov. 2005.
- [4] S. U. Yoon, S. Y. Kim, E. K. Lee, and Y. S. Ho, "A Framework for Representation and Processing of Multi-view Video Using the Concept of Layered Depth

- Image," J. VLSI Signal Process. Syst., Vol. 46, No. 2-3, pp. 432-441, Mar. 2007.
- [5] S. U. Yoon, S. Y. Kim, and Y. S. Ho, "Preprocessing of Depth and Color Information for Layered Depth Image Coding," Lecture Notes Comput. Sci. (LNCS), Vol. 3333, pp. 622-699, Nov. 2004.
- [6] C. L. Zitnick, S. B. Kang, M. Uyttendaele, S. Winder, and R. Szeliski, "High-quality Video View Interpolation Using a Layered Representation," ACM Trans. Graph, Vol. 23, No. 3, pp. 600-608, Aug. 2014.
- [7] Interactive Visual Media Group at Microsoft Research [Online]. Available: <http://research.microsoft.com/vision/InteractiveVisualMediaGroup/3DVideoDownload/>
- [8] S. U. Yoon, E. K. Lee, S. Y. Kim, Y. S. Ho, K. Yun, S. Cho, and N. Hur, "Inter-camera Coding of Multi-view Video Using Layered Depth Image Representation," Lecture Notes Comput. Sci. (LNCS), Vol. 4261, pp. 432-441, 2006.
- [9] Joong-Hee Lim, Min-Tae Kim, "Efficient Compression and Authentication Technique of Multi-view Image with Color and Depth Information", IWIT, 2009.
- [10] Joong-Hee Lim, "Efficient Compression and Authentication Technique of Multi-view Image with Color and Depth Information by Layered Depth Image Representation", Hong-ik Univ., 2008.
- [11] "Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264/ISO/ICE 14496-10 AVC" in Joint Video Team (JVT) of ISO/ICE MPEG and ITU-T VCEG, JVT-G050, 2003.
- [12] Jae-Chang Jung, "H.264/AVC Video Compression Standard and Applications of Industry", Han-Yang Univ. 2015.
- [13] Yo-Sung Ho, Seung-Hwan Kim, "Algorithm Understanding and Program Analysis of Algorithm H.264/AVC", DooYang Press Sep. 30, 2009.
- [14] C. L. Zitnick, S. B. Kang, M. Uyttendaele, S. Winder, and R. Szeliski, "High-quality Video View Interpolation Using a Layered Representation," Proc. of ACM SIGGRAPH, pp. 600-608, Aug. 2014.
- [15] Wan-Beom Lee, "A Study on Color Image Compression Using Downscaling Method and Sub-sampling Method", Journal of the Korea Academic-Industrial cooperation Society, Vol. 20, No. 20, pp. 20-25, Feb. 2019.
<http://dx.doi.org/10.5762/KAIS.2019.20.2.20>
- [16] Jin-Su Kim, Jae-Yeong Cho, and Nam-Jae Park, "Block Chain Based CCTV Image Forgery · Modulation Verification Mechanism", Journal of KIIT, Vol. 17, No.

8, pp. 107-114, Aug. 2019.

<http://dx.doi.org/10.14801/jkiit.2019.17.8.107>

저 자 소 개

지 인 호(정회원)



- 1980년 2월 : 서울대학교 전자공학과 공학사
- 1983년 8월 : 서울대학교 전자공학과 공학석사
- 1995년 6월 : Polytechnic School of Engineering at New York University, USA, 전기 및 컴퓨터공학과, 공학박사
- 1982년 ~ 1988년 : 국방과학연구소 선임연구원
- 2004년 ~ 2005년 : University of Maryland at College Park, USA, 연구교수
- 2012년 ~ 2013년 : Temple University at Philadelphia, USA, 연구교수
- 1995년 ~ 현재 : 홍익대학교 소프트웨어융합학과 교수
- 주관심분야 : DSP and 3D Image Processing, Multimedia Security, Multimedia Signal Processing

※ 본 논문은 2017년도 홍익대학교 학술연구진흥비에 의하여 지원되었음