

<http://dx.doi.org/10.7236/IIBC.2014.14.4.29>

IIBC 2014-4-5

계층적 깊이 영상으로 표현된 다시점 영상의 효과적인 압축 기술

Effective Compression Technique of Multi-view Image expressed by Layered Depth Image

지인호*

Inn-Ho Jee*

요 약 다시점 비디오는 카메라 수만큼의 컬러영상들과 깊이 영상들이 존재하므로 막대한 양의 데이터를 갖게 되어 서, 그 양을 줄이는 새로운 압축 기술 개발이 필수적이다. 최근 계층적 깊이 영상(Layered Depth Image : LDI)의 개념을 이용하여 깊이 정보를 포함하는 다시점 비디오를 효과적으로 압축 부호화하는 기술이 주목 받고 있다. 이 방법은 여러 시점의 깊이 정보와 워핑(warping) 함수를 사용하여 다시점의 컬러와 깊이 영상을 합성하여 하나의 데이터 구조를 만드는 방법이다. 본 논문에서는 계층적 깊이 영상에서 실제 거리를 이용하고 오버랩(overlap) 문제를 해결하고 그 리고 컬러 변환을 기반으로 복원을 위해서 추가되는 데이터의 양을 감소시켜 압축 성능을 향상시키는 방법을 제안하였다.

Abstract Since multi-view video exists a number of camera color image and depth image, it has a huge of data. Thus, a new compression technique is indispensable for reducing this data. Recently, the effective compression encoding technique for multi-view video that used in layered depth image concepts is a remarkable. This method uses several view point of depth information and warping function, synthesizes multi-view color and depth image, becomes one data structure. In this paper we use actual distance for solving overlap in layered depth image that reduce required data for reconstructing in color-based transform. In experimental results, we confirmed high compression performance and good quality of reconstructed image.

Key Words : Multi-view, Multi-media, Multi-view video, Warping, Layered depth image, Overlap

1. 서 론

멀티미디어 콘텐츠의 발전은 가상현실과 3차원 오디오 비주얼 기술의 실현을 가능하게 하고 있다. 이러한 새로운 3차원 멀티미디어 콘텐츠는 다시점 비디오(multi-view video)의 표현, 처리, 압축, 재현 등의 기술을 기반으로 하고 있다. 다시점 비디오는 하나의 장면을

여러 시점에서 다수의 카메라로 촬영한 다중 동영상의 집합이다. 그림 1은 다시점 비디오의 생성 과정을 나타낸 것이다. 8대의 카메라가 서로 다른 시점에서 동일한 장면을 촬영을 하고, 촬영된 비디오 콘텐츠들이 하나의 집합으로 구성된다. 이렇게 생성된 다시점 비디오는 재생시 사용자의 요구에 따라 여러 시점을 제공할 수 있고 또한 입체 영상을 제공할 수 있다. 그래서 3차원 TV, 자유 시

*정회원, 홍익대학교 컴퓨터정보통신공학과
접수일자 : 2014년 7월 10일, 수정완료 : 2014년 8월 7일
게재확정일자 : 2014년 8월 8일

Received: 10 July, 2014 / Revised: 7 August, 2014

Accepted: 8 August, 2014

**Corresponding Author: ijee@hongik.ac.kr

Dept. of Computer & Information Communications, Engineering.,
Hongik University, Korea

점 TV, 감시 카메라 영상, 멀티미디어 교육등에서 응용될 수 있다. 최신의 3차원 기술에서는 깊이 영상(depth image) 또는 깊이 맵(depth map)을 포함하는 다시점 비디오의 요구가 많아지고 있다. 여러 시점을 제공하는 다시점 비디오는 사용자에게 의해서 존재하지 않는 시점을 요구받을 수 있다. 근본적으로 이런 문제를 해결하기 위해서는 많은 카메라를 사용하여 많은 시점 영상들을 생성하는 것이다. 그러나 이 경우 많은 비용뿐만 아니라 생성된 다시점 비디오 데이터를 저장하고 전송하는 등의 과정에서 많은 어려움이 존재하게 된다. 그래서 요구받은 시점의 영상이 인접한 시점 영상들을 이용하여 생성되는 보간(interpolation) 방법을 고려하게 되었다. 보간 기법을 이용하여 새로운 시점 영상을 생성하기 위해서는 깊이 영상이 유용한 정보를 제공할 수 있다. 이러한 이유가 다시점 비디오에 깊이 영상을 포함하는 대표적인 이유가 된다. 그 밖에도 입체 비디오 생성 등의 이유로 깊이 영상의 필요성이 증대되고 있다.

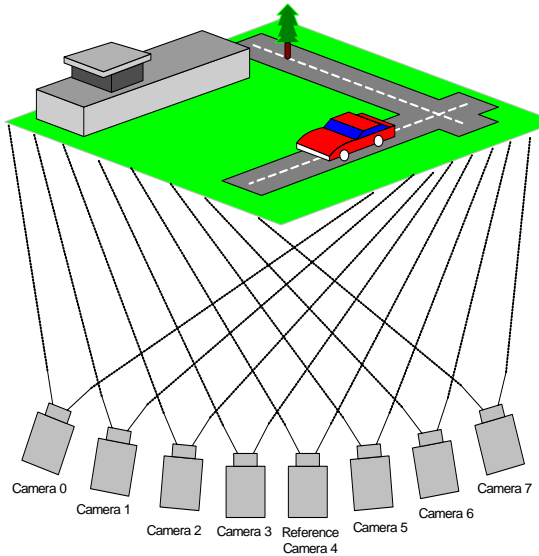


그림 1. 다시점 비디오의 생성 과정
Fig. 1. Generation scheme of multi-view video

깊이 영상은 카메라로부터 장면 내에 존재하는 객체와의 거리를 그레이 레벨(gray level) 영상으로 나타낸 것이다. 이것은 양안 시차를 이용한 스테레오 알고리즘이나 깊이 카메라로부터 생성이 가능하다. 그림 2는 다시점 비디오에서 스테레오 알고리즘을 이용하여 깊이 영상

을 생성하는 과정을 나타낸 것이다. 다양한 스테레오 알고리즘들이 제시되고 있으며, 최근에는 에너지를 기반으로 하는 알고리즘은 그 성능이 우수한 것으로 알려져 있다.

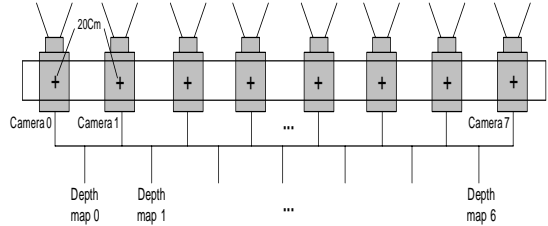


그림 2. 깊이 영상의 생성 과정
Fig. 2. Generation scheme of depth image

본 논문의 구성으로 2장에서는 다시점 컬러 영상과 깊이 영상으로부터 계층적 깊이 영상 생성하는 방법을 제시하였다. 3장에서는 제안된 계층적 깊이 영상의 개념을 설명하였다. 그리고 4장에서는 계층적 깊이 영상으로부터 다시점 영상의 복원 과정을 제시하였다. 5장에서는 계층적 깊이 영상과 복원 성능 실험 결과를 제시하였다. 마지막 6장에서는 본 연구에 대한 결론 및 추후 연구 과제를 제안하였다.

II. 다시점 컬러 영상과 깊이 영상으로부터 계층적 깊이 영상 생성

1. 계층적 깊이 영상의 생성

동일 시점으로 이동한 영상들은 화소들의 중복성을 판단하기 위해서 깊이 영상으로부터의 깊이 값을 비교하게 된다. 다음의 그림 3는 이 과정을 나타낸 것이다^[10].

서로 다른 위치의 3개 카메라 $I_i (i=1,2,3)$ 에서 촬영된 영상이 C_1, C_2, C_3 라고 할 때, I_{11} 은 첫 번째 카메라 I_1 에서 촬영된 영상 C_1 의 첫 번째 화소를 나타낸다. I_{21}, I_{31} 도 마찬가지로이다. 그리고 C_1 을 기준 영상으로 선택하면, 나머지 C_2, C_3 영상은 기준 영상과 같은 시점으로 3차원 위평을 수행하게 된다. 계층적 깊이 영상을 생성하기 위해서는 기준 영상 C_1 과 위평된 C_2, C_3 영상의 동일 위치 화소들은 깊이 값을 비교하게 된다. 즉, 각 깊이 영상들에서 동일한 위치의 화소 값(깊이 값)의 차를 구하고 임계값 I_0 와 비교를 하게 된다. 그림에서 I_{11} 과 I_{22} 의 깊이

값의 차가 임계값 보다 큰 값이므로 각각 독립된 화소들로 저장된다. 그리고 I_{31} , I_{14} , I_{21} 간의 값이 값 차도 임계값 보다 큰 값을 가져서 모두 독립적으로 저장된다. 또한 I_{35} 와 I_{16} 과의 값이 값 차, I_{35} 와 I_{23} 과의 값이 값 차도 임계값 보다 큰 값을 가지고 있어 독립적으로 저장된다. 그러나 I_{16} 과 I_{23} 과의 값이 값 차이는 임계값 보다 작으며, 따라서 이 두 화소들은 중복된 화소로 간주되어 하나의 화소로 합병된다. 그림에서 I' 이 해당되며, 값이 값은 평균된다. 결과적으로 생성된 계층적 깊이 영상 화소들은 서로 다른 계층으로 구성된 것을 확인할 수 있다.

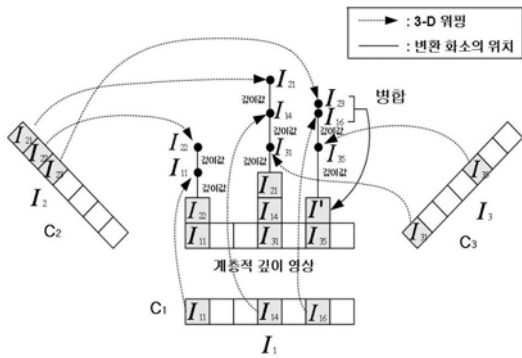


그림 3. 다시점의 컬러 영상과 깊이 영상으로부터 계층적 깊이 영상의 생성
 Fig. 3. Layered depth image generation from multi-view color and depth image

본 논문에서는 기존의 방법과 다르게 계층적 깊이 영상을 통해서 중복된 화소들을 보다 정교하게 찾기 위해서는 깊이 영상의 화소 값이 아닌 실제 거리를 환산하여서 이용하면 가능하다. 또 한 가지 방법으로 전경 화소를 이용하는 것이다. 3차원 워핑 과정에서 여러 화소들이 기준 시점의 하나의 화소 좌표로 변환되는 오버랩(overlap) 문제가 발생된다. 이 경우 배경보다는 전경의 화소가 변환 이동되도록 하는 것이다. 상대적으로 깊이 값이 작은 것이 전경이 되고 깊이 값이 큰 것이 배경이 된다. 이 방법은 각 영상들 간의 유사성이 높아져, 중복된 화소가 많아지게 되어서 압축 성능을 증가시키는 효과를 갖는다^[10].

그림 4은 계층적 깊이 영상을 생성하는 전 과정을 나타낸 것이다^{[11][3]}. 다시점 영상으로부터 깊이 정보를 얻은 후, 기준 시점으로 시점 변환을 수행하는 3차원 워핑을 수행한다.

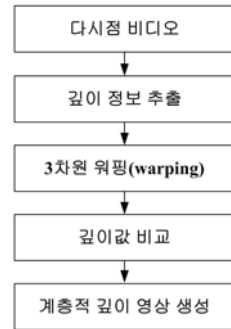


그림 4. 다시점 비디오로부터 계층적 깊이 영상의 생성 과정
 Fig. 4. Flow chart of layered depth image from multi-view video

그 다음 기준 시점과 기준 시점으로 변환된 여러 시점들 간의 값의 값을 비교한다. 그리고 비교된 깊이 값에 의해서 중복성을 제거하면서, 압축된 계층적 깊이 영상의 데이터 구조를 형성한다.

III. 제안된 계층적 깊이 영상의 개념

본 논문에서는 계층적 깊이 영상의 압축 성능을 향상시키기 위해서 세 가지의 방법을 제안하였다.

첫 번째로 깊이 값의 비교를 깊이 영상의 화소 값이 아닌 실제 거리를 계산하여서 비교하였다. 8비트 그레이 레벨(gray level)의 깊이 영상의 범위는 0~255이다. 다시점 비디오는 스튜디오 환경에서는 생성이 되므로, 최소 거리로부터 최대거리까지의 거리 간 차이는 깊이 영상만큼 크지 않다. 8비트 그레이 레벨의 깊이 영상에서 임계값 T를 9로 하면, 현재의 깊이 값이 x일 때, x+9이하일 때는 하나의 계층으로 처리되고, 그 이상 일 경우는 새로운 계층이 형성된다. 실제 거리의 범위가 0~99라고 가정하고 x의 값을 120이라고 가정하면 120~128까지의 범위가 하나의 계층으로 처리된다. 이 범위를 실제 거리 값의 범위로 환산하면 47.06~50.59가 되고 그 범위는 3.5정도이다. 따라서 실제 거리에서는 임계값이 3.5로 작아지게 되어서 압축률이 떨어지게 된다. 그리고 깊이 영상에서는 정수범위에서 거리를 비교하지만 실제 거리에서는 소수점 이하까지도 비교하게 되는 것이다. 그러나 시각적으로 소수점 이하의 거리는 큰 의미를 가지지 못한다. 따라서 계층의 생성과 통합을 결정하는 깊이 비교를 범위가 작은 실제 거리 값으로 수행하면 좁아진 범위만큼의

정확도와 압축효율을 얻을 수 있다.

두 번째로 3차원 워핑 과정에서 현 시점의 복수 좌표들이 기준 시점의 동일 좌표로 변환되는 오버랩(overlap) 문제가 발생된다. 그림 5는 워핑 과정에서 발생하는 오버랩의 예를 보여준다.

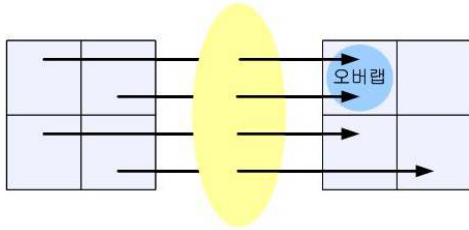


그림 5. 오버랩(overlap)의 예
Fig. 5. Example of overlap

배경 화소와 전경 화소가 동일한 위치로 워핑이 되는 경우, 둘 중에 하나의 화소를 선택하여야만 한다. 동일한 위치의 전경은 배경 앞에 위치하므로, 전경은 해당 화소 위치에 남겨지고 배경은 버려지는 것이 올바른 워핑이 될 것이다. 전경의 거리는 배경의 거리보다 짧기 때문에 깊이 영상을 이용하여 전경 화소만 이동되도록 할 수 있다. 계층적 깊이 영상 표현에서 오버랩에서 전경만 워핑된 영상들은 서로 상호간의 유사성이 높아지게 된다. 유사한 화소들이 많으면, 주어진 임계값에 의해서 계층의 증가를 막게 되어서 압축 성능을 증가시키는 효과를 갖는다.

세 번째로 본 논문에서 제안한 계층적 깊이 영상의 전체적인 데이터 구조는 그림 6과 같다. 컬러 정보를 R, G, B 채널로 저장하면 각 채널간의 상관관계가 커서 계층적 깊이 영상에서 중복성이 제거된 화소들은 원래의 컬러로 복원되기가 어렵다. 따라서 본 논문에서는 Y, Cr, Cb 형태로 컬러 변환을 수행한 후 계층적 깊이 영상으로 표현하므로 Y, Cr, Cb로 저장된다. Auxiliary data는 8비트만을 가지고 해당 DP가 다시점 영상들 중에서 몇 번째 카메라 영상인가의 정보를 제공해서 복원 시 중요한 정보로 사용된다.

또한 residual data는 복원 시 필요한 데이터로, 폐색 영역의 데이터와 오차 데이터로 구성된다. 이 데이터의 정밀도는 복원된 다시점 영상의 성능을 좌우한다. 기준 영상의 경우 residual data가 0이므로 별도로 저장할 필요가 없다. 그래서 배열은 인덱스 1~NumofCamera-1이

된다. 그리고 압축 성능을 높이기 위해서 residual data의 8비트 Y성분은 4~7비트로 양자화 하여 저장하였다. 그리고 Cr과 Cb는 부 표본화(sub sampling)를 수행하여서 데이터 양을 1/4로 축소하여 저장하였다. 또한 LDI에는 워핑을 수행할 수 있는 각 카메라 행렬이 저장되어 복원 시 사용된다.

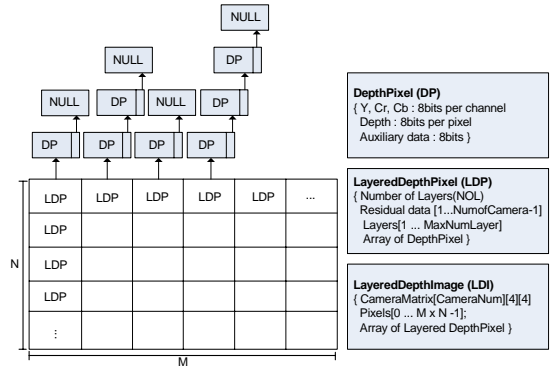


그림 6. 제한된 계층적 깊이 영상의 데이터 구조
Fig. 6. Data structure of constrained layered depth image

VI. 계층적 깊이 영상으로부터 다시점 영상의 복원

1. 다시점 영상의 복원 과정

계층적 깊이 영상으로부터 다시점 영상을 복원하는 과정을 그림 7에 나타내었다. 우선적으로 계층적 깊이 영상은 각각의 워핑된 영상들로 분해된다. 그리고 역 워핑을 통해서 원래의 시점으로 이동하게 된다. 워핑에서와 마찬가지로 폐색영역으로 인한 홀(hole)문제와 그리고 오버랩 문제가 발생하게 된다. 오버랩 문제는 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 위해서 이다. 따라서 보간은 압축의 효율을 증가시키는 한 요인이 된다. 마지막 단계에서는 이렇게 계층적 깊이 영상 생성 단계에서 만들어진 추가 데

이터를 더해서 우수한 품질의 복원된 다시점 영상을 생성하게 된다^{[1][2][4]}.

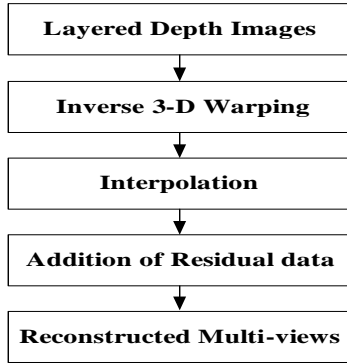


그림 7. 다시점 영상들의 재생과정
 Fig. 7. Reconstruction procedure of multi-view images

2. 복원 다시점 영상을 위한 보간법과 추가 데이터

영상의 기하학 처리에 해당하는 워핑 변환은 오버랩 문제와 홀(hole) 문제가 발생하게 된다. 홀 문제는 각 시점 영상들의 모든 화소가 기준 시점 영상으로 워핑 되지 않으며, 기준 시점 영상의 특정 화소들은 각 시점 영상들의 화소에는 존재하지 않기 때문이다. 다시점 영상으로의 복원 과정에서 워핑과 역 워핑을 거치므로 무수히 많은 홀이 발생하고, 이는 복원되는 다시점 영상의 품질을 저하시키게 된다. 그러므로 이를 극복하기 위해서 보간법을 사용하였다. 본 논문에서는 폐색 영역과 인접 화소와의 높은 유사성에 근거하여, 홀 화소 주변의 화소들을 평균하여 홀 화소에 삽입하였다. 화소 보간에 대한 식은 다음 식(1)로 정리된다.

$$I(x_h, y_h) = \frac{1}{\text{Total pixel}} \sum_{x=-x_0}^{x_0} \sum_{y=-y_0}^{y_0} IW(x, y), (x, y \neq 0) \quad (1)$$

여기서, $I(x_h, x_h)$ 는 보간 될 화소를 나타내며 (x_h, x_h) 는 $(x = 0, y = 0)$ 과 동일좌표이다. 그리고 $(-x_0 \cdots x_0), (-y_0 \cdots y_0)$ 는 주변화소의 범위를 나타내고, $IW(x, y)$ 는 역 워핑 된 영상의 화소를 말하며 홀 화소는 제외된다.

보간 과정을 통해서 원래의 다시점 영상들과의 유사성이 높아졌으므로, 그 차이 값이 감소하므로 복원을 위한 추가해야 할 데이터양이 감소하는 장점을 갖게 된다.

추가 데이터는 계층적 다시점 영상을 생성할 때 저장된다. 따라서, 부호화 단계에서 미리 복호화 과정을 수행하여서 원래의 다시점 영상들과 복원된 다시점 영상들 간의 차이를 계산하게 된다. 이렇게 만들어진 추가 데이터의 양은 계층적 다시점 영상의 압축 성능을 좌우하게 된다. 가능한 적은 양이 되어야만 압축 성능은 좋아지게 된다. 그래서 Y영상은 4~7비트로 양자화 하며, 그리고 Cr과 Cb는 데이터양을 1/4로 축소하였다. 추가 데이터의 양은 다음으로 계산된다.

- Y 성분 추가 데이터 양
 = 양자화 비트 수 * 다시점 영상의 가로 크기 * 다시점 영상의 세로 크기
- Cr과 Cb 성분의 추가 데이터 양
 = (다시점 영상의 가로 크기 * 다시점 영상의 세로 크기)/2
- 추가 데이터 양
 = Y 성분 추가 데이터 양 + Cr과 Cb성분의 추가 데이터양

Y 성분은 밝기 성분으로 영상의 품질에 대한 민감도가 높으므로 복원 영상에 대한 품질을 좌우하게 된다. 그리고 색상 성분인 Cr과 Cb는 영상 품질에 대한 민감도가 적어서 데이터양을 많이 축소하였다.

V. 계층적 깊이 영상과 복원 성능 실험

본 논문에서 계층적 깊이 영상 표현을 통한 압축 성능과 이 압축에서 복원되었을 때의 다시점 영상들의 품질을 PSNR(PSNR(Peak Signal To Noise Ratio))로 성능을 측정하였다. 실험을 위해서 사용된 다시점 비디오는 Microsoft Research(MSR)에서 제공한 “Breakdancers”와 “Ballet”를 사용하였다^[7]. 두 비디오는 깊이 정보와 카메라 변수가 제공되고 있으며, 비디오 품질은 XGA급이며, 전송률은 15fps 이고, 1 차원 원호의 8대의 카메라로부터 촬영되었다. 그림 8과 그림 9는 Breakdancers와 Ballet 다시점 비디오의 첫 번째 프레임에 나타낸 것으로 8대 카메라에서 생성된 각 시점의 컬러 영상과 깊이 영상을 보여준다.

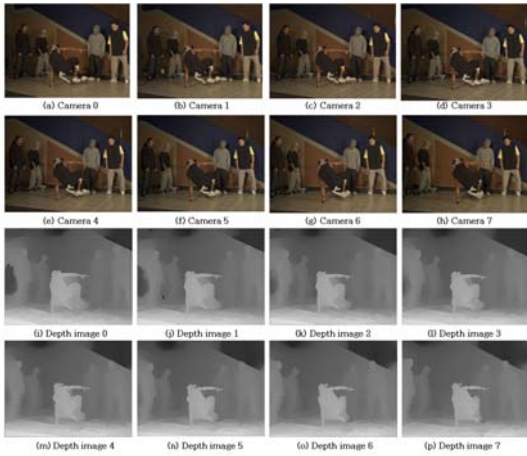


그림 8. Breakdancers의 다시점 영상
Fig. 8. Multi-view image of Breakdancers

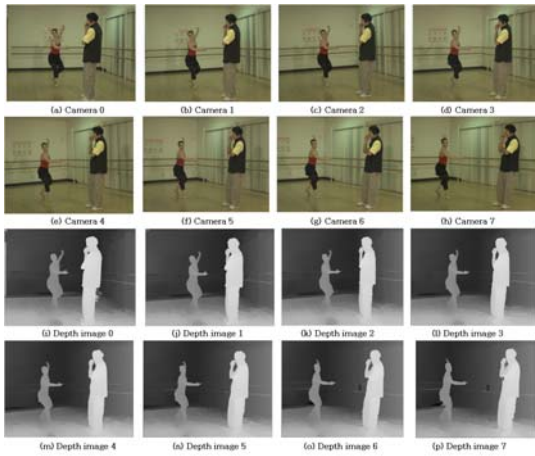


그림 9. Ballet의 다시점 영상
Fig. 9. Multi-view image of Ballet

PSNR은 두 영상의 차이를 정량적으로 나타낸 것으로 다음과 같이 정의된다.

$$PSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right) \quad (2)$$

여기서, 평균제곱 오차 MSE(Mean Square Error)는 다음과 같다.

$$MSE = \frac{1}{mn} \sum_{k=0}^{m-1} \sum_{l=0}^{n-1} \| I(k,l) - \hat{I}(k,l) \|^2 \quad (3)$$

$I(i,j)$ 는 원 영상을 나타내고 $\hat{I}(i,j)$ 는 복원 영상을 나타낸다.

1. 계층적 깊이 영상의 생성을 통한 압축 실험

첫 번째 실험에서는 다양한 임계값에 근거하여 다시점 비디오를 계층적 깊이 영상으로 표현하였을 때의 압축 정도를 실험하였다. 임계값이 클수록 워핑된 화소들이 병합될 확률이 높아져서 압축률도 높아지게 된다.

3차원 워핑 실험에서는 중심에 해당하는 4번 카메라를 기준으로 하고 나머지의 카메라 영상들이 4번 카메라로 좌표 변환 된다. 오버랩 문제를 고려하지 않으면 먼저 전경부분이 3차원 워핑 되고 그 다음으로 동일한 좌표에 배경이 3차원 워핑 된다. 따라서 해당 좌표에는 배경만 있고 전경의 객체들이 사라진 영상을 생성하게 된다. 이것은 계층적 깊이 영상의 압축 성능의 저하뿐만 아니라, 이것을 복원하는 경우에 사라진 전경의 객체에 대한 더 많은 추가적인 데이터를 요구하게 된다. 따라서 오버랩 문제의 해결은 압축 효율을 증가시키는 하나의 작업이 된다.

표 1은 실험 비디오에 대한 임계값과 거리 비교 그리고 오버랩 문제 해결 여부에 따른 계층적 깊이 영상의 데이터 크기를 비교한 것이다. 8개의 컬러 영상과 8개의 깊이 영상으로 구성된 한 프레임의 크기는 25, 166 [Kbytes]이다. 표 1에서는 추가 데이터(Residual data)를 포함하지 않았고, 임계값(Threshold)은 1~9로 증가시키면서 데이터의 크기를 측정하였다. 임계값이 커질수록 데이터의 크기가 작아지는 것을 알 수 있다. 그러나 임계값 5.0 이후에는 데이터 크기가 완만하게 감소하는 것을 확인할 수 있다. 거리비교(distance)는 깊이 영상의 화소값을 이용하는 것과 실제 거리로 환산에서 이용하는 것을 나타냈다. 깊이 영상의 범위는 0~255이지만, Breakdancers의 실제 거리 범위는 44~120이며 Ballet의 경우는 42~130이다^[7]. 그래서 실제 거리로 환산하게 되면 정확도는 향상되고 범위는 좁아져서 데이터의 압축 효율을 높일 수 있음을 확인하였다. 마지막으로 오버랩 문제를 전경만 남도록 처리하면, 4번 영상으로 워핑된 다시점 영상들의 유사성이 높아져서 계층적 깊이 영상에서 계층의 생성을 줄이게 되어서 그 크기가 작아진 것을 확인할 수 있다.

표 1. 실험 비디오 프레임들의 데이터 크기
 Table 1. Data size of experimental video frames

Condition			Breakdancers		Ballet	
threshold	distance	overlap	1 st Frames	2 nd Frames	1 st Frames	2 nd Frames
1	depth map	no	20,931	21,160	23,402	23,533
1	real	no	13,961	13,739	17,055	17,018
1	real	yes	13,596	13,374	15,684	15,628
3	depth map	no	15,290	15,192	16,872	16,886
3	real	no	11,926	11,721	13,513	13,553
3	real	yes	11,592	11,385	12,697	12,702
5	depth map	no	13,716	13,533	14,836	14,872
5	real	no	11,156	10,978	12,582	12,598
5	real	yes	10,852	10,676	11,819	11,835
7	depth map	no	12,988	12,754	13,971	14,028
7	real	no	10,857	10,706	12,224	12,209
7	real	yes	10,572	10,419	11,518	11,507
9	depth map	no	12,523	12,281	13,415	13,458
9	real	no	10,671	10,575	11,990	11,987
9	real	yes	10,399	10,303	11,341	11,345

2. 계층적 깊이 영상으로부터의 복원된 다시점 영상들의 품질에 관한 실험

복원을 하기 위해서는 역 3차원 워핑과 보간 과정이 수행되어야 한다. 역 워핑 결과 영상에서는 홀 화소에 해당하는 흰색 부분이 워핑 영상 보다 더 많이 존재하는 것을 확인할 수 있다. 그러나 보간을 거치게 되면 사라진 것을 확인할 수 있으며, 이로 인해 추가 데이터의 양을 감소시킬 수 있다.

그림 10과 11은 보간 된 다시점 영상에 추가 데이터를 더하여 얻어진 최종 복원 영상이다. 추가 데이터에서 Y 성분은 5비트로 양자화 되었고, Cr과 Cb는 1/4씩 축소되

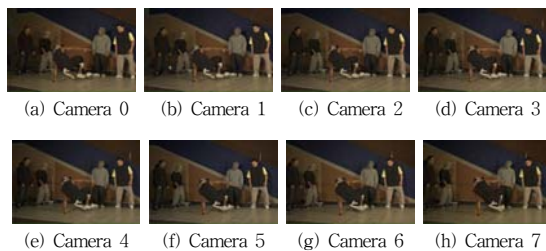


그림 10. LDI 구조에서 각각의 영상으로 복원된 Breakdancers의 영상들

Fig. 10. Reconstructed Breakdancers's images from each LDI scheme

었다. 그래서 66%와 69%로 압축된 계층적 영상으로부터 다시점 영상을 복원 한 것이다.

표 2는 복원된 다시점 영상들의 PSNR을 나타낸 것으로 임계값은 9로 고정하고 양자화 계수를 4~7까지 변화를 주었다. 주어진 방법에서 계층적 깊이 영상의 크기가 다르지만 추가되는 데이터의 양은 양자화 계수에 의해 동일하다. B 방법과 C방법에서는 PSNR의 차이가 크지 않지만, 표2에서 B방법의 데이터 양이 더 크므로 C방법이 더 우수함을 확인 할 수 있다.

표 2. 복원된 다시점 영상의 성능
 Table 2. Performances of reconstructed multi-view image

	Breakdancers						Ballet					
	A		B		C		A		B		C	
	BPP	SNR	BPP	SNR	BPP	SNR	BPP	SNR	BPP	SNR	BPP	SNR
Th=9,Q=4	5.71	32.8	5.14	33.1	5.05	33.0	6.01	33.7	5.56	33.4	5.35	33.5
Th=9,Q=5	5.95	36.7	5.36	37.1	5.27	37.1	6.23	38.1	5.78	37.6	5.57	37.8
Th=9,Q=6	6.16	38.5	5.58	39.1	5.49	39.1	6.45	40.8	5.99	40.0	5.79	40.4
Th=9,Q=7	6.38	39.0	5.79	39.7	5.71	39.7	6.67	41.6	6.21	40.8	6.01	41.2
Th=3,Q=4	6.61	32.9	5.54	33.1	5.43	33.1	7.11	33.9	6.05	33.4	5.78	33.5
Th=3,Q=5	6.83	36.8	5.76	37.2	5.65	37.2	7.33	38.4	6.27	37.7	6.00	37.9
Th=3,Q=6	7.04	38.6	5.98	39.3	5.87	39.3	7.55	41.1	6.49	42.3	6.22	40.6
Th=3,Q=7	7.26	39.1	6.19	39.9	6.09	39.9	7.77	42.0	6.71	41.0	6.44	41.4

A : 깊이 영상의 화소값을 비교하여 LDI 압축
 B : 오버랩 문제의 전경과 배경을 고려하지 않은 LDI 압축
 C : 오버랩 문제에서 전경만을 변환한 LDI 압축

마지막으로 그림 12은 제안 방법 C의 압축 성능을 위하여, BPP(Bit Per Pixel)와 PSNR과의 관계를 나타낸 것이다. 임계값은 9로 설정하였다. 두 실험 비디오 모두 5~6BPP에서 33dB~40dB의 성능을 나타내고 있어, 30%이상 압축이 되어도 화질에 큰 영향이 없음을 보여준다.

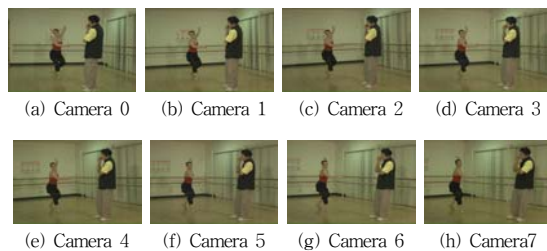
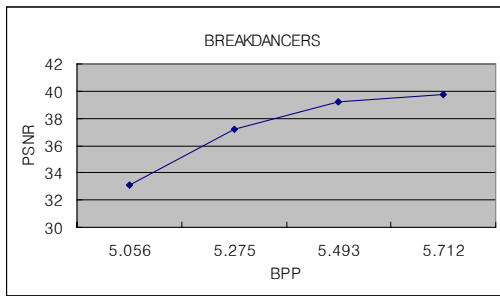
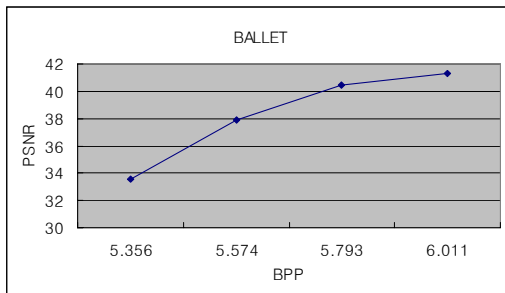


그림 11. LDI 구조에서 각각의 영상으로 복원된 Ballet 영상들

Fig. 11. Reconstructed Ballet's images from each LDI scheme



(a)



(b)

그림 12. BPP에 따른 PSNR 성능
Fig. 12. PSNR performance vs BPP

V. 결론

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 표현을 제안하였다. 계층적 깊이 영상은 각 시점의 영상을 기준 시점으로 워핑 변환하고, 동일 위치의 워핑된 각 시점의 화소끼리 거리 값 비교를 통해서 중복되는 화소를 제거하는 기술이다.

본 논문에서는 보다 향상된 기법을 활용하여 계층적 깊이 영상 구조 표현의 압축 성능을 향상시켰다. 우선, 깊이 영상의 화소 값을 비교하지 않고 실제 거리로 환산하여 비교하여 정확도를 높였으며 비교 거리의 범위도 감소시켰다. 두 번째로 워핑 과정에서 배경과 전경이 동일한 화소 위치에 워핑 되는 것을 막고 전경만 워핑이 되도록 하여 워핑된 영상들 간의 유사성을 높였다. 세 번째로 계층적 깊이 영상의 전체적인 데이터 구조를 효과적으로 변경하였다. YCrCb신호로 컬러모델 변환하여 저장하

로 컬러간의 상관성을 줄여서 복원 시 비디오의 품질을 향상시켰다. 그리고 추가 데이터의 양을 감소시키기 위해서 보간을 수행하고 Y성분은 양자화하고 Cr과 Cb 성분은 1/4로 부 표본화를 수행하였다.

실험 결과를 통해서, 1~9까지의 다양한 임계값에서 압축 효율은 향상된 것을 확인할 수 있었다. 그리고 보간과 추가 데이터의 사용으로 복원된 다시점 영상들의 품질 저하도 적은 것을 BPP를 통해서 확인할 수 있다.

향후 과제로 다시점 비디오에 대한 시간적 중복성을 제거하는 비디오 압축 부호화를 추가하여 압축 효율을 더욱 향상 시킬 수 있다.

References

- [1] Seung-Uk Yoon, and Yo-Sung Ho "Multiple color and depth video coding using a hierarchical representation," Vol.17.11 Nov. 2007, pp. 1450-1460.
- [2] J. Shade, S. J. Gortler, L. W. He, and R. Szeliski, "Layered depth images," in Proc. SIGGRAPH'98, Jul. 1998, pp. 231 - 242.
- [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. 2004.
- [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] Martin Vetterli and Jelena Kovacevic, Wavelets and Subband Coding, Prentice-Hall, Inc. 1995.
- [10] Jong-Hong Shin and Innho Jee, "H.264 encoding technique of multi-view video expressed by layered depth image," Journal of Internet, Broadcasting and Communication, vol. 14, no. 2, pp. 43-51, April. 2014.
- [11] Jung-Ho Kim, Soon Chul Kwon, Kwang Chul Son, Chae Bong Sohn, and Seung Hyun Lee, "Effect of 2Dimension and 3Dimension Images on Human Factors," International Journal of Internet, Broadcasting and Communication vol. 6, no. 1, pp. 11-13, 2014.
- [12] Won-ill Kim, "A study on the smoothing method for efficient video stream transmission on IPTV network," International Journal of Internet, Broadcasting and Communication vol. 3, no. 2, pp. 16-25, 2011.

저자 소개

지 인 호(정회원)



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

※ 본 논문은 2012년도 홍익대학교 학술연구진흥비에 의하여 지원되었음