

3차원 워핑 방법을 이용한 효율적인 다시점 깊이영상 부호화 방법

나상태, 오관정, 이천, 호요성
 광주과학기술원 정보통신공학과

I. 서론

공간상의 거리정보를 나타내는 깊이영상은 색상영상에 비해 상대적으로 적은 비트 소모가 예상되지만, 다시점의 깊이영상을 부호화 하기 위해서는 많은 데이터량을 필요로 한다. 본 논문에서는 효과적인 다시점 깊이영상 부호화를 위하여 3차원 워핑 방법을 이용하여 부호화하고자 하는 시점의 가상영상을 만들어 부호화시에 부가적인 정보로 이용함으로써 효율을 높이는 방법을 제안한다.

II. 제안한 다시점 깊이영상 부호화 방법

3차원 워핑 방법을 이용하여 부호화하고자 하는 시점의 영상을 합성하여 부호화에 이용하는 방법을 제안한다. 깊이영상은 영상 자체가 이미 깊이정보를 포함하고 있기 때문에 카메라 파라미터를 제외한 부가적인 정보 없이 3차원 워핑이 가능하다.

3차원 워핑을 통한 깊이영상 합성은 먼저, 2차원 깊이영상의 좌표를 3차원 공간좌표로 투영하기 위해 투영식인 식(1)을 이용한다.

$$P_{WC} = R \cdot A^{-1} \cdot P_{ref} \cdot D + t \quad (1)$$

$$P_{target} = A \cdot R^{-1} \cdot P_{WC} - t \quad (2)$$

여기서, P 는 2차원 참조영상과 목표영상, 또는 3차원 공간에서의 좌표정보를 나타내며 R 은 회전변수, A 는 카메라의 고유변수, D 는 깊이정보, t 는 이동변수를 각각 나타낸다. 투영된 3차원 공간상의 좌표를 식(2)를 이용하여 목표영상에 재투영 시킨 후 참조영상의 화소위치와 합성하고자 하는 목표영상의 화소위치가 서로 부합되는 곳에 참조영상의 화소값을 전달함으로써 적절한 영상이 합성된다.

이때 비폐색영역으로 인한 홀이 발생하며 이를 채우기 위한 방법은 다음과 같다. 만약 합성하고자 하는 영상의 좌우시점 영상을 모두 참조영상으로 사용할 수 있을 경우에는 좌우시점 영상 모듈을 이용하여 같은 목표깊이영상을 합성한다. 이때 좌측시점 영상을 참조영상으로 이용하여 합성을 한 경우에는 합성된 영상의 좌측에 홀이 발생하게 되며, 우측시점 영상을 이용하여 합성을 한 경우에는 우측에 홀이 발생하게 된다. 그러므로, 좌측에 홀이 발생한 영상을 기준으로 하였을 때 홀 부분의 화소값을 우측에 홀이 있는 영상을 이용하여 채움으로써 적절하게 홀을 채울 수 있다. 이때 여전히 남아있는 홀은 주위의 이용가능한 8개의 화소값을 이용하여 채운다. 만약 한쪽시점 방향의 참조영상만이 이용가능한 경우에는 주위 8개의 이용 가능한 화소값들의 중간값을 이용하여 홀을 채운다.

두 번째로 배경에 의해 전경이 가려지는 현상이 발

생하며, 이를 해결하기 위한 방법은 다음과 같다. 깊이영상에서 각각의 화소값은 깊이정보를 가지고 있고 깊이정보가 작을수록 즉 카메라로부터 가까이 있는 객체일수록 화소값이 크다. 그러므로, 목표영상에서 하나의 좌표에 두개 이상의 좌표가 투영될 경우 화소값을 비교하여 큰값을 전경으로 간주하여 선택한다.

합성된 영상은 깊이영상 부호화시에 부가적인 정보 영상으로 쓰여 부호화 효율을 높여준다. 즉 복호화된 I-시점의 영상으로 P-시점과 B-시점의 영상을 합성한 다음 P-시점 부호화시에 부가적인 정보로 이용하며, 다시 복원된 P-시점의 영상으로 B-시점의 영상을 합성하여 I-시점을 이용하여 합성한 B-시점의 영상을 이용한 홀 제거 후 B-시점 영상 부호화시에 부가적인 정보로 이용함으로써 부호화 효율을 높일 수 있다

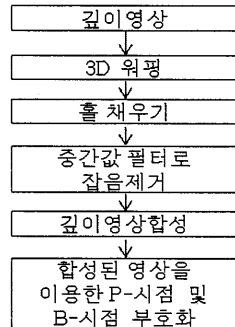


그림 1. 깊이영상 부호화 방법의 순서도

III 실험 결과 및 결론

참조영상의 이용이 가능한 P-시점과 B-시점의 영상을 이용한 실험결과를 표1에 나타내었다. 그 결과 비디오 부호화 표준인 H.264를 기반으로 제작된 다시점 비디오 부호화 소프트웨어인 JMVM에 비해 더 높은 이득이 있었다. 그 이유는 3차원 워핑 방법과 적절한 홀 채우기 방법, 그리고 전경과 배경의 알맞은 표현을 이용하여 생성된 가상시점의 영상들이 부가적인 정보영상으로 효과적으로 이용되었기 때문이다. 제안한 방법은 JMVM에 비해 P-시점에서 최대 약0.87dB 정도 PSNR 개선이 있었다.

표 1. JMVM방법에 대한 제안한 방법의 이득

Views	Ballet		Breakdancers	
	P	B	P	B
Bit saving [%]	11.42	7.61	16.37	14.71
PSNR [dB]	0.725	0.452	0.865	0.728

3차원 워핑 방법을 이용한 효율적인 다시점 깊이영상 부호화 방법

나상태, 오관정, 이천, 호요성
광주과학기술원 정보통신공학과

Efficient Multi-view Depth Video Coding using 3D Warping Technique

Sang-Tae Na, Kwan-Jung Oh, Cheon Lee, Yo-Sung Ho
Gwangju Institute of Science and Technology

stna@gist.ac.kr, kjoh81@gist.ac.kr, leecheon@gist.ac.kr, hoyo@gist.ac.kr

Abstract: 3 차원 공간상의 객체와 카메라 사이의 거리정보를 나타내는 깊이영상의 특징은 3 차원 복원 또는 3 차원 워핑(warping)과 같은 응용분야에 활용이 가능하며, 이를 이용하여 다시점 비디오 부호화 방법에서 부호화 효율을 높이기 위해 사용될 수 있으며, 3 차원 TV, 자유시점 TV 에도 이용될 수 있다. 이때 사용되는 깊이영상은 단일 또는 그 이상시점의 영상들을 필요로 한다. 그러므로 단일시점뿐만 아니라 다시점 깊이영상을 전송하거나 저장하기 위한 부호화 방법이 필수적이다. 이 논문에서는 효과적인 다시점 깊이영상 부호화 방법으로, 부호화하고자 하는 시점의 깊이영상을 참조영상과 카메라 파라미터를 이용하여 3 차원 워핑 방법으로 합성한 후, 그 시점을 부호화 할 때 합성된 영상을 부가적인 정보영상으로 이용하여 부호화 효율을 높이고자 한다. 특히, 합성할 때 발생하는 홀(hole)을 채우기 위한 방법과 배경이 전경을 가리게 되는 현상을 해결하기 위한 방법 또한 제안하였다. 제안한 방법에 의해, 최고 0.86dB 의 PSNR 개선이 있었다.

Keywords: 3 차원 워핑(warping) 방법, 다시점 깊이영상 부호화, 홀(hole) 채우기.

I. 서론

최근 멀티미디어와 가상현실에 대한 관심이 증가하면서 현실에 존재하는 장면을 좀더 사실적으로 표현하기 위한 노력들이 여러분야에서 연구되고 있다. 그 중의 한 분야로써 기존의 2차원 색상정보에 더해 3차원 깊이정보를 이용한 연구들이 진행되고 있다. 2차원 색상정보와 깊이정보를 이용하여 3차원의 깊이를 느낄 수 있게 해주는 3차원TV, 또는 깊이정보를 이용하여 가상시점 영상들을 만들어서 사용자가 자유롭게 시점을 선택하여 볼 수 있는 자유시점TV가 그 예라 할 수 있다. 또한 깊이정보는 부호화 효율을 향상시키기 위해서 이용될 수 있다. 그 예로서 다시점 비디오 부호화를 들 수 있다. 다시점 비디오는 한 장면을 여러 시점에서 다수의 카메라로 촬영한 것으로 사용자가 원하는 시점의 영상을 제공하는 것을 목적으로 한다. 그러나 카메라 수만

큼의 비디오가 존재하므로 데이터량이 매우 많아서 이를 저장하고 전송하기 위해서는 효과적인 부호화 방법이 필수적이다. 한가지 방법으로 깊이정보를 이용하여 부호화하고자 하는 시점의 영상을 합성한 후에 그 합성된 영상을 부가적인 정보영상으로 이용함으로써 부호화 효율을 높일 수 있다[1]. 이때 사용되는 깊이영상은 단일 또는 그 이상의 깊이영상을 필요로 한다.

깊이영상은 3차원 공간상의 객체와 카메라 사이의 거리정보를 8bits만을 이용하여 흑백의 단위로 나타내었으므로, 색상영상에 비해 훨씬 단순한 화소값의 분포를 갖는다. 그렇기 때문에 부호화하는데 쓰이는 비트스모울 또한 상대적으로 적지만, 여전히 상당한 비트가 요구된다. 특히, 다시점의 깊이영상을 부호화하기 위해서는 많은 데이터량을 필요로 한다. 그러므로, 효과적인 부호화 방법을 필요로 한다.

본 논문에서는 효과적인 다시점 깊이영상 부호화 방법을 제안한다. 먼저, 깊이영상 또한 색상영상의 특징을 그대로 가지고 있기 때문에 현재 영상부호화의 표준인 H. 264, 또는 다시점의 영상을 부호화 할 때는 시간축 방향뿐만 아니라 시점방향에도 밀접한 상관성이 있기 때문에 Fraunhofer-HHI에서 제작한 다시점 비디오 부호화 참조 소프트웨어인 JMVM이 사용되었다[2]. 두 번째로 깊이영상만이 가지고 있는 특징을 이용하여, 부호화하고자 하는 시점의 깊이영상을 합성하여, 그 시점을 부호화하는데 부가적인 정보영상으로 이용하는 것과 이때 발생하는 홀(hole)을 채우기 위한 방법 그리고, 전경이 배경에 의해서 가려지는 현상을 해결하기 위한 방법들을 제안한다.

II. 다시점 깊이영상의 응용분야 및 기존의 연구

깊이영상은 2차원 이미지의 각각의 좌표점들에 대한 3 차원 거리정보를 나타내기 때문에 이를 이용하여, 3차원 복원 또는 3차원 워핑(warping)등에 이용되어, 다시점 영상의 부호화할 때 부가적인 정보로써 효율을 높이기 위해 또는 자유시점 TV등의 응용분야에 활용이 가능하다. 이러한 깊이영상을 부호화하기 위해 시간방향뿐만 아니라 시점방향의 상관성이 고려된 JMVM부호화기를 이용할 수 있다.

1. 다시점 깊이영상의 응용분야.

시점합성예측 방법은 다시점 비디오를 부호화하기 위해 고안된 새로운 방법이다. 그림 1에서 볼 수 있듯이 부호화 하고자 하는 P 또는 B-시점의 영상을 이미 부호화 및 복호화된 이웃영상과 깊이영상, 카메라 파라미터를 이용하여 합성한 후에 P 또는 B-시점의 영상을 부호화 할 때 합성된 각각의 영상을 부가적인 정보로 이용함으로써 부호화 효율을 높일 수 있다. 시점합성예측 방법은 이웃 시점방향간의 거리정보는 그 시점에서의 깊이정보에 의존한다는 전제하에 발전되었다. 그러므로, 효율적인 시점합성예측 방법을 위해서는 효율적인 깊이 정보 부호화 방법을 필요로 한다[1].



그림 1. 시점합성예측 방법

깊이 영상은 또한 자유시점TV에서 이용될 수 있다. 자유시점TV는 기존의 이미 정해진 시점에서 수동적으로 시청하는 TV가 아닌 사용자가 임의의 시점을 자유롭게 선택할 수 있는 TV이다. 이를 위해 모든 시점의 영상을 전송하거나 저장하는 것은 거의 불가능하다. 그러므로, 몇몇 시점의 영상과 깊이영상만을 이용하여 임의의 가상시점을 생성하여 마치 모든 시점의 영상이 전달된 것과 같은 효과를 만들 수 있다[3].

2. 기존의 깊이영상 부호화 연구

다시점 깊이영상을 부호화 하기 위해서 현재 비디오 부호화 표준인 H.264를 이용하여 각각의 시점을 독립적으로 부호화 할 수 있다. 하지만, 다시점 비디오와 마찬가지로 다시점 깊이영상 또한 같은 장면에 대한 여러시점에서의 깊이정보를 가지고 있다. 즉 시간축 방향의 잉여성뿐만 아니라 시점방향에서의 잉여성분을 줄여서 부호화 효율을 높일 수 있다. 그러므로, 다시점 비디오 부호화 소프트웨어인 JMVM을 다시점 깊이영상 부호화에도 똑같이 적용 될 수 있다. 즉, 부호화에 쓰이는 계층적 B구조를 이용할 뿐만 아니라, 시점방향간의 상관성도 같이 이용된다면, 부호화 효율을 더욱 효과적으로 높일 수 있다[4].

Ⅲ. 제안한 다시점 깊이영상 부호화 방법

3차원 워핑 방법을 이용하여 부호화 하고자 하는 시점의 영상을 합성하여 부호화에 이용하는 방법을 제안한다. 3차원 워핑을 통한 깊이영상 합성은 먼저, 2차원 깊이영상의 좌표를 3차원 공간좌표로 투영하기 위해 투영식인 식(1)을 이용한다.

$$P_{WC} = R \cdot A^{-1} \cdot P_{ref} \cdot D + t \quad (1)$$

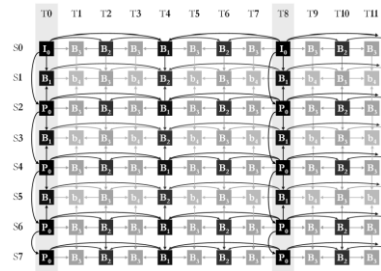


그림 2. 계층적 B구조를 이용한 JMVM 부호화 구조

$$P_{target} = A \cdot R^{-1} \cdot P_{WC} - t \quad (2)$$

여기서, P 는 2차원 참조영상과 목표영상, 또는 3차원 공간에서의 좌표정보를 나타내며 R 은 회전변수, A 는 카메라의 고유변수, D 는 깊이정보, t 는 이동변수를 각각 나타낸다. 투영된 3차원 공간상의 좌표를 합성하고자 하는 2차원 목표영상의 적합한 위치에 식(2)를 이용하여 재투영 시킨다. 이렇게 재투영을 시킨 후 참고영상의 화소위치와 합성하고자 하는 목표영상의 화소위치가 서로 부합되는 곳에 참고영상의 화소값을 전달함으로써 적절한 영상이 합성된다.

이때, 비폐색영역으로 인한 홀이 발생할 뿐만 아니라, 전경이 배경에 의해 가려지는 현상이 일어나게 된다.

먼저, 발생된 홀을 채우기 위한 방법은 다음과 같다. 만약 합성하고자 하는 영상의 좌우시점 영상을 모두 참조영상으로 사용할 수 있을 경우에는 좌우시점 영상 모두를 이용하여 같은 목표깊이영상을 합성한다. 이때 좌측시점 영상을 참조영상으로 이용하여 합성을 한 경우에는 합성된 영상의 좌측에 홀이 발생하게 되며, 우측시점 영상을 이용하여 합성을 한 경우에는 우측에 홀이 발생하게 된다. 그러므로, 좌측에 홀이 발생한 영상을 기준으로 하였을 때, 홀 부분의 화소값을 우측에 홀이 있는 영상을 이용하여 채움으로써 적절하게 홀을 채울 수 있다. 이때 여전히 남아있는 홀은 주위의 이용가능한 8개의 화소값의 중간값을 이용하여 채운다. 만약 한쪽시점 방향의 참조영상만이 이용 가능할 경우에는 주위 8개의 화소값중 이용 가능한 화소값들의 중간값을 이용하여 홀을 채운다.

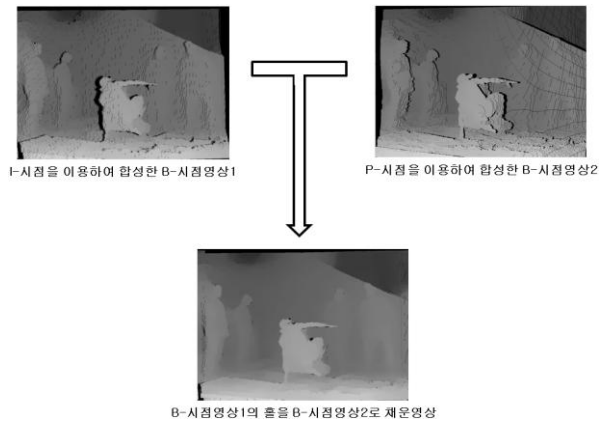


그림 3. 좌우시점 참조영상을 이용한 홀 채우기 방법

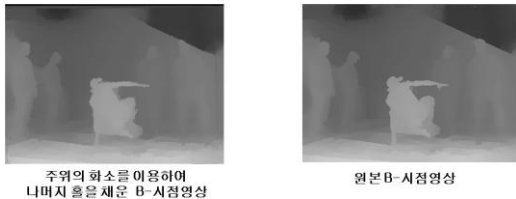


그림 4. 합성된 영상과 원본영상 비교

그림 2에서는 좌우시점의 참조영상을 이용하여 합성한 영상과 이를 이용하여 홀을 채운 결과를 보여주고, 그림 3에서는 그림 2에서 여전히 남아있는 홀을 주위의 화소값을 이용하여 제거한 후 원본 영상과 비교한 것을 보여주는 그림이다.

두 번째로 참조영상의 시점에서는 보이는 배경부분이지만, 합성된 영상에서는 전경에 의해 가려져야 하는 부분이 오히려 합성시에 전경을 가리고 나타나는 현상을 해결하기 위한 방법은 다음과 같다. 깊이영상에서 각각의 화소값은 깊이정보를 가지고 있고 깊이정보가 작을수록 즉 카메라로부터 가까이 있는 객체일수록 화소값이 크다. 그러므로, 보이지 않아야 할 부분 즉 배경 부분이 보이는 현상이 일어나는 것을 방지 하기 위해서 깊이영상의 화소값을 이용한다. 목표깊이영상에서 참조영상으로부터 3차원 좌표를 통해 투영된 좌표가 두번 이상 같게 나타날 경우 그 좌표에서의 화소값을 서로 비교하여 그 값이 더 작은것을 배경의 화소값으로 간주하고, 큰값을 선택하여 그 좌표에서의 적절한 화소값으로 간주한다.

이렇게 합성된 영상은 깊이영상 부호화시에 부가적인 정보영상으로 쓰여 부호화 효율을 높여준다. 즉 부호화를 거쳐 복호화된 I-시점의 영상으로 P-시점과 B-시점의 영상을 합성한 다음 이 합성된 영상을 먼저 P-시점 부호화 시에 부가적인 정보로 이용하며, 다시 복호화를 거쳐 복원된 P-시점의 영상으로 B-시점의 영상을 합성하여 I-시점을 이용하여 합성한 B-시점의 영상을 이용하여 홀을 제거한 후 B-시점 영상 부호화시에 부가적인 정보로 이용함으로써 부호화 효율을 높일 수 있다. 그림 4에서는 깊이영상 부호화 방법의 순서도를 나타내며, 그림 5에서는 JMVM에서 8개의 시점에 대해서 이용되는 시점방향의 구조도와 제안된 시점방향의 구조도를 비교하는 그림이다.

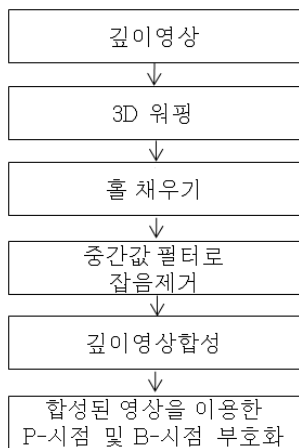


그림 5. 깊이영상 부호화 방법의 순서도

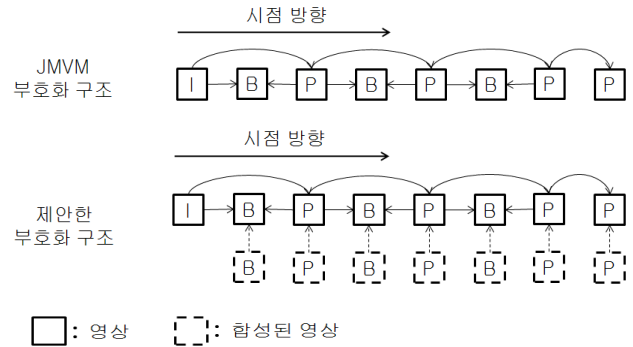
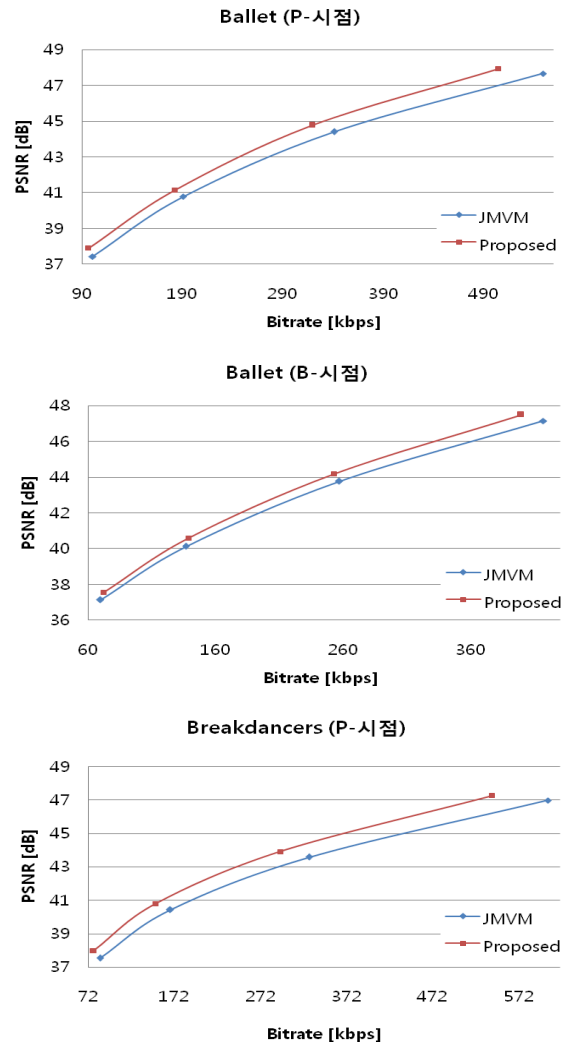


그림 6. 제안한 방법과 JMVM 방법에서의 부호화 구조 비교

IV. 실험결과 및 토의

이 논문에서 제안한 방법의 성능을 평가하기 위해서 사용된 실험영상은 XGA(1024 x 768)급 사이즈의 Breakdancers와 Ballet의 영상이다. 총 100프레임의 영상을 4번과 5번, 6번의 시점에 각각 I-시점, B-시점, P-시점을 적용하였다. JMVM 1.0을 이용하여 부호화 및 복호화를 하였으며, 다양한 양자화 레벨(22, 27,32,37)이 이용되었다.



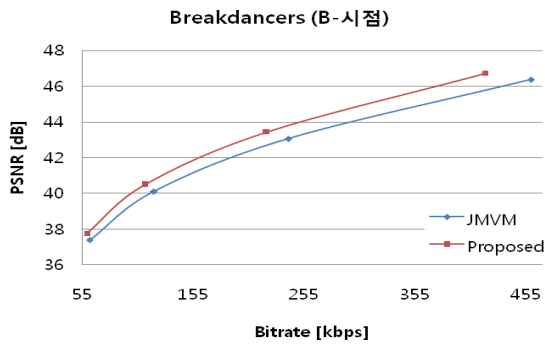


그림 7. 실험영상들의 비트율-왜곡 곡선

표 1. JMVM 방법과 제안한 방법의 효율비교 “ Ballet (P-시점) ”

QP	Bitrate [kbps]		PSNR [dB]	
	JMVM	Proposed	JMVM	Proposed
22	549.91	504.85	47.65	47.93
27	341.54	319.39	44.40	44.76
32	191.03	182.28	40.77	41.14
37	100.72	96.30	37.43	37.90

표 2. JMVM 방법과 제안한 방법의 효율비교 “ Ballet (B-시점) ”

QP	Bitrate [kbps]		PSNR [dB]	
	JMVM	Proposed	JMVM	Proposed
22	416.93	399.73	47.14	47.48
27	256.93	252.61	43.76	44.18
32	136.96	138.89	40.13	40.60
37	69.71	72.02	37.13	37.55

표 3. JMVM 방법과 제안한 방법의 효율비교 “ Breakdancers (P-시점) ”

QP	Bitrate [kbps]		PSNR [dB]	
	JMVM	Proposed	JMVM	Proposed
22	606.03	540.88	46.98	47.26
27	329.01	295.37	43.58	43.90
32	167.42	150.41	40.43	40.79
37	85.90	78.56	37.53	37.99

표 4 JMVM 방법과 제안한 방법 효율비교 “ Breakdancers (B-시점) ”

QP	Bitrate [kbps]		PSNR [dB]	
	JMVM	Proposed	JMVM	Proposed
22	459.83	418.64	46.37	46.73
27	240.97	221.62	43.06	43.45
32	119.86	111.65	40.11	40.49
37	62.12	59.51	37.39	37.76

표 5 제안한 방법의 이득

Views	Ballet		Breakdancers	
	P	B	P	B
Bit saving [%]	11.42	7.61	16.37	14.71
PSNR [dB]	0.725	0.452	0.865	0.728

실험결과에서 볼 수 있듯이 Breakdancers 의 전체 이득이 Ballet의 전체 이득에 비해 높다는 것을 확인할 수 있다. 그 이유는 Ballet에서의 전경과 배경사이의 변화가 Breakdancers의 그것보다 크기 때문이다. 즉, Ballet에서 전경과 배경사이의 거리가 멀어서 각각의 시점사이의 홀이 더욱 커지고 이 커진 홀은 시점을 합성할 때 부정적인 영향을 미치며, 합성된 영상이 부가적인 정보로 쓰일 때 Breakdancers에 비해 좋지 않은 결과가 나오기 때문이다. 그 결과 Breakdancers의 P시점에서는 최대 약 0.87dB, Ballet에서는 그보다 낮은 0.73dB의 PSNR 이득이 있었다.

V. 결론

본 논문에서는 3차원 워핑 방법을 이용한 효율적인 다 시점 깊이영상 부호화 방법을 제안하였다. P-시점 또는 B-시점 부호화시에 깊이영상이 가지고 있는 깊이정보를 이용하여 워핑 방법을 통해 가상의 영상을 생성하여 P 또는 B-시점부호화 할 때 부가적인 정보영상으로 이용하여 부호화 효율을 높일 수 있다. 가상의 영상을 합성할 때 발생하는 홀을 채우기 위해 이웃시점의 영상 또는 이웃 화소값을 이용하였고, 전경이 배경에 의해 가려지는 현상을 해결하기 위해 참조영상에서의 여러좌표가 목표영상에서 같은 좌표에 투영될 때, 그 좌표에서의 화소값이 가장 큰 값이 전경임을 유추하여 이용하였다. 제안한 방법을 JMVM과 비교해 볼 때 최고 0.87dB정도 향상된 화질의 영상을 얻을 수 있었다.

감사의 글

본 연구는 광주과학기술원(GIST) 실감방송연구센터(RBRC)를 통한 정보통신부대학IT연구센터(ITRC)와 교육인적자원부 두뇌한국21(BK21) 정보기술사업의 지원에 의한 것입니다.

참고문헌

- [1] S. Ince, E. Martinian, S. Yea, A. Vetro, " Depth Estimation for View Synthesis in Multiview Video Coding," *Proc. of IEEE 3D TV Conference*, May, 2007
- [2] ISO/IEC MPEG & ITU-T VCEG JVT-U207, "Joint Multiview Video Model (JMVM)," October, 2006.
- [3] A. Smolic, K. Mueller, P. Merkle, C. Fehn, P. Kauff, P. Eisert, T. Wiegand, " 3D Video and Free Viewpoint Video-Technologies, Applications and MPEG Standards," *Proc. of ICME 2006*, pp.2161-2164, July, 2006.
- [4] P. Merkle, A. Smolic, K. Muller, T. Wiegand, " Efficient Compression of Multi-view Depth data Based on MVC," *Proc. of IEEE 3D TV Conference*, May, 2007.