

## 자유시점 TV와 3차원 비디오 구현을 위한 가상시점 영상합성 기술

오관정<sup>o</sup>, 호요성  
광주과학기술원 정보통신공학과  
{kjoh81, hoyo}@gist.ac.kr

### 요약

다시점 비디오 시스템은 자유시점 TV와 3차원 비디오 구현을 위한 대표적인 3차원 비디오 시스템으로, 가상시점에 대한 영상합성 기술은 다시점 비디오 부호화 기술과 함께 3차원 비디오 시스템 구현을 위한 핵심기술이다. 일반적으로 가상시점에 대한 영상합성은 깊이영상, 3차원 워핑, 카메라 매개변수를 통해 합성하는 DIBR(depth image based rendering) 기술이 널리 사용되고 있다. 본 논문에서는 DIBR 기술을 기반으로 하여, 참조 영상간의 색상차로 인한 합성영상의 부자연스러움을 보완하기 위해 깊이영상 기반의 히스토그램 정합 방법을 제안한다. 또한, 비페색 영역에 대한 효율적인 빈 곳 채우기를 위해 깊이영상 기반의 채움(inpainting) 기술을 제안한다. 컴퓨터 모의실험을 통해 제안하는 가상시점 영상합성 기술이 객관적인 화질과 주관적인 화질에서 모두 우수함을 확인했다.

### 1. 서론

영상처리 기술의 발달과 실감형 영상 기술에 대한 필요성에 따라 다양한 3차원 영상 시스템에 대한 연구가 활발히 진행되고 있다. 유럽에서는 2002년에 ATTEST (advanced three-dimensional television system technologies) 프로젝트를 통해 3차원 TV에 관한 기초 기술을 연구했으며, 2004년부터는 20여개의 연구기관이 컨소시엄을 구성하여 3DTV 프로젝트를 수행하고 있다. 미국에서는 NASA와 MIT에서 3차원 영상과 홀로그래피 기술을 연구하고 있고, 일본에서는 NIST (national institute of information and communications technology)의 UMRC (universal media research center)에서 3차원 TV 및 초실감 방송 (ultra realistic)에 관련된 연구가 진행 중이다.

다시점 비디오 부호화와 가상시점 합성기술은 이러한 3차원 비디오 시스템의 핵심 기술로 다시점 비디오 부호화의 경우 MPEG (moving picture experts group)과 JVT (joint video team)을 중심으로 MPEG-2 MVP (multiview video profile), MPEG-4 MAC (multiple auxiliary component), 최근 MVC (multiview video coding)까지 다양한 다시점 비디오 부호화 관련 연구가 진행되었다 [1].

가상시점 합성기술의 경우 컴퓨터 비전과 그래픽스 관련 분야에서 오랫동안 연구되어져 왔고, 최근 MPEG에서 표준화를 준비중인 "FTV (free viewpoint television) and 3D Video"에서 다시점 영상과 깊이영상을 기반으로 하는 가상시점 영상합성이 EE2 (exploration and experiment)에서 연구되어지고 있다 [2].

가상시점 영상합성은 좌, 우 참조 시점의 영상을 이용하여 가운데 시점에 대한 중간영상을 합성하는 것이 일반적이다. 이때 영상합성에 이용되는 참조 영상들 사이에 조명이나 카메라 빛 노출의 차이로 인해 색상차이가 존재하는 경우, 합성된 영상에는 색상차이로 인한 부자연스러움이 나타난다. 이를 해결하기 위해 본 논문에서는 깊이 영상을 고려한 히스토그램 정합 방법을 제안한다. 또한 영상 합성에 이용되는 깊이영상이 완전하지 못하기 때문에 이로 인해 합성영상에는 두 참조화면으로부터 모두 채워질 수 없는 빈 곳(hole)이 생긴다. 일반적으로 이러한 빈 영역은 비페색영역에 발생하고 영상 보간 기술이나 채움(inpainting) 기술을 통해 채워진다. 그러나 이러한 기술들은 객체의 경계와 같이 전경과 배경에 맞닿아 있는 빈 영역에 대해서는 그 효율성이 떨어지므로, 본 논문에서는 이를 보완하기 위해 깊이영상을 고려한 채움 기술을 제안한다.

## 2. 제안하는 가상시점 영상합성 기술

그림 1은 본 논문에서 제안하는 가상시점 영상합성 기술의 흐름도이다.

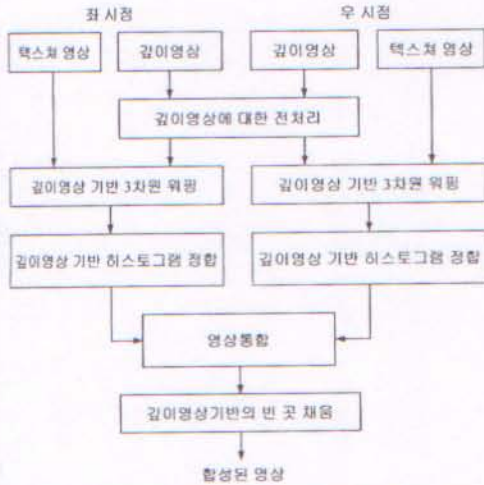


그림 1. 제안하는 가상시점 영상합성 기술 흐름도

깊이영상에 대한 전처리는 영상합성에 이용되는 깊이영상의 초기 오류값을 보정하는 역할을 한다. 현재 영상합성에 이용되는 깊이영상은 다양한 예측 기술들을 통해 만들어진 것으로 적지 않은 초기 오류 영역을 포함하고 있다. 이를 보정하기 위해 영상 팽창과 침식을 이용한다. 그림 2에서 보듯이, 원 깊이영상에서 존재했던 얼굴에 대한 오류영역이 보정되었음을 확인할 수 있다.

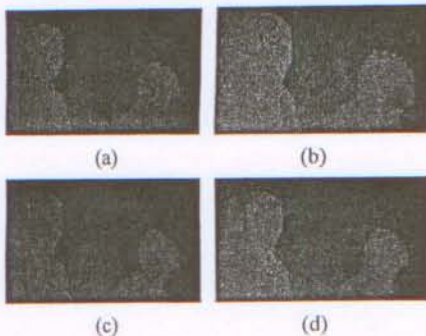


그림 2. 영상 팽창/침식을 이용한 깊이영상의 오류 보정: (a) 원영상, (b) 팽창 후, (c) 침식 후, (d) 보정 후

다음은 3차원 워핑을 이용하여 가상시점의 초기 영상을 만든다. 3차원 워핑이란 그림 3과 같이 깊이영상과 카메라 매개변수를 이용하여 현재 시점의 영상을 다른 시점으로 투영시키는 작업을 말한다.

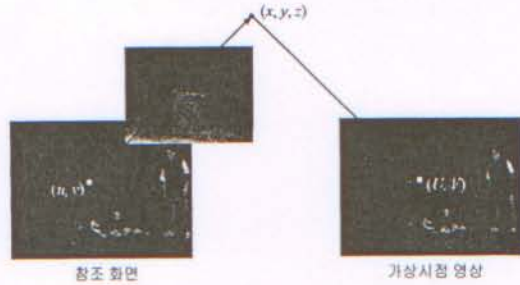


그림 3. 영상합성에 이용되는 3차원 워핑

3차원 워핑은 참조화면이 3차원 공간상에 투영된 후 다시 가상시점으로 투영되는 두 번의 투영을 통해 이뤄진다. 식 (1)과 식 (2)는 참조 화면에 대한 3차원 공간상의 투영과 3차원 공간상에서 가상시점으로의 투영에 대한 공식이다.

$$(x, y, z)^T = R_{ref} A_{ref}^{-1} (u, v, 1)^T d_{u,v} + t_{ref} \quad (1)$$

$$(l, m, n)^T = A_{ver} R_{ver}^{-1} \{ (x, y, z)^T - t_{ver} \} \quad (2)$$

여기서  $A, R, t$ 는 카메라 매개변수이고  $d$ 는 깊이영상의 깊이값을 나타낸다. 참조화면의 좌표  $(u, v)$ 는 가상시점에서  $(l/m, m/n)$ 이 된다.

텍스처 영상에 3차원 워핑을 직접 적용하면 그림 3에서와 같이 절삭 오류로 인한 실선 모양의 빈 곳이 생기게 된다. 이러한 문제를 해결하기 위해, 일반적으로 그림 4와 같이 깊이영상에 3차원 워핑을 적용하고, 이를 미디언 필터링 한 후 상응하는 텍스처 영상의 화소를 불러오는 방식을 이용한다.

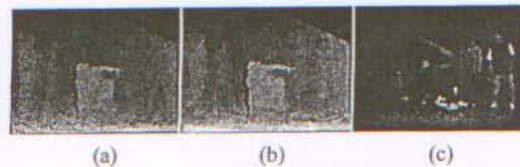


그림 4. 깊이영상 기반의 워핑: (a) 워핑된 깊이 영상, (b) 미디언 필터링, (c) 워핑된 텍스처 영상

3차원 워핑을 좌우 참조화면에 적용하면 합성하고자 하는 중간시점에 대한 두개의 3차원 워핑된 영상이 얻어진다. 일반적으로 3차원 워핑된 영상은 참조화면에는 보이지 않지만 가상시점에서는 보이는 비폐색영역을 제외하고는 모든 영역에서 값을 갖는다. 그러나 텍스처 영상과 깊이영상의 객체의 경계가 일치하지 않는 경우에, 그림 5와 같이 합성된 영상의 배경 영역에 전경의 잔상이 남는 경우가 발생한다. 본 논문에서는 빈 영역의 팽창을 통해 이러한 잔상을 제거한다. 확장된 빈 영역은 후에 영상 통합이나 빈 곳 채우기 과정에서 보정된다.



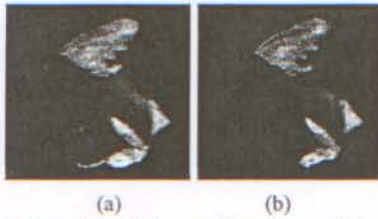


그림 5. 잔상 제거: (a) 팽창 전, (b) 팽창 후

다음 단계는 두 참조 화면간에 존재하는 색상차 보정을 위한 깊이영상 기반의 히스토그램 정합이다. 히스토그램 정합[3]은 색상차 보정을 위한 대표 기술로 두 영상간의 히스토그램을 분석하여 색상차를 보정하는 방법이다. 본 논문에서는 그림 6과 같이 다른 빈 영역을 갖는 두 참조 화면에 대해 같은 빈 영역을 갖도록 하여 히스토그램 정합의 정확도를 높인다.



그림 6. 빈 영역 맞추기 작업: (a) 우 영상, (b) 좌 영상, (c) 같은 빈 영역을 갖도록 변형된 영상

같은 빈 영역을 갖는 두 참조 화면에 대해 식 (3)을 이용하여 히스토그램을 계산하고, 식 (4)를 이용하여 누적 히스토그램을 얻어낸다.

$$h_L[v] = \frac{1}{w \times h} \sum_{m=0}^{h-1} \sum_{n=0}^{w-1} \delta[v, y_L(m, n)] \quad (3)$$

$$\text{with } \delta[a, b] = \begin{cases} 1, & \text{if } a = b \\ 0, & \text{else} \end{cases}$$

$$C_L[v] = \sum_{i=0}^v h_L[i] \quad (4)$$

여기서  $w$ 와  $h$ 는 영상의 가로와 세로 크기이다.

두 참조 화면의 누적 히스토그램의 평균 누적 히스토그램으로 가상시점에 대한 누적 히스토그램을 정의하고, 이를 중심으로 좌우 참조 화면의 누적 히스토그램을 정합하여 두 참조 화면간의 색상차를 보정한다. 참조 화면과 가상시점에 대한 누적 히스토그램을 정합하기 위해 식 (5)와 같이 각 화소값에 대한 누적 분포를 비교해가며, 히스토그램의 각 화소값에 대한 값들을 계산한다.

$$M[v] = \begin{cases} u, & \text{for } C_1[u] < C_2[v] \leq C_1[u+1] \quad C_1[v] \leq C_2[v] \\ u+1, & \text{for } C_1[u] < C_2[v] \leq C_1[u+1] \quad C_1[v] > C_2[v] \end{cases} \quad (5)$$

여기서  $C_1$ 와  $C_2$ 은 각각 가상시점과 좌시점에 대한 누적 히스토그램을 의미한다.

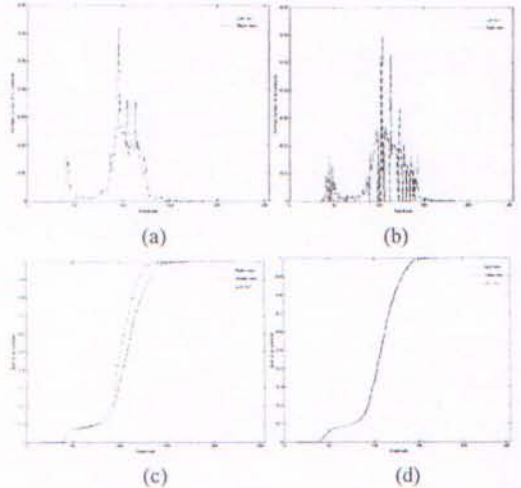


그림 7. 참조 영상 사이의 색상차 보정: (a) 보정 전 히스토그램, (b) 보정 후 히스토그램, (c) 보정 전 누적 히스토그램, (d) 보정 후 누적 히스토그램

그림 7은 히스토그램 정합의 예이다. 색상차까지 보정된 두 합성영상은 식 (6)을 통해 통합된다.

$$I_V(u, v) = \alpha I_B(u, v) + (1 - \alpha) I_A(u, v) \quad (6)$$

여기서  $I_B$ 는 영상합성에서 대부분의 영역을 차지하는 기준 참조화면이고,  $I_A$ 는 기준 참조화면에서 채워지지 않은 영역을 보충하는 보조 참조화면으로 정의된다.  $I$ 는 통합된 영상을 의미한다.  $\alpha$ 는  $I_B$ 가 빈 영역이 아니면 1이고 빈 영역이면 0이다. 즉,  $I_B$ 에 존재하지 않는 영역만이  $I_A$ 에서 합성된다.  $I_B$ 와  $I_A$ 는 가상시점과의 거리에 따라 보다 가까운 시점이  $I_B$ 로 결정된다.

영상합성의 마지막 단계는 영상통합을 통해서도 채워지지 않는 남아있는 빈 영역을 채우는 것이다. 이를 위해 일반적으로 영상 보간법이나 채움 기술 [4]이 이용된다. 그러나 이러한 방법들은 빈 영역의 주변 화소들을 이용하여 합성하기 때문에 전경과 배경을 모두 인접하고 있는 빈 영역에 대해서는 그 효율이 떨어진다. 일반적으로 전경과 배경 모두에 인접한 빈 영역의 경우 비폐색 영역이므로 배경에 존재하는 빈 영역으로 간주되어 배경에 속하는 인접화소들로부터 채워져야 한다.

따라서 본 논문에서는 깊이값을 고려한 채움 기술을 통해 이러한 문제를 해결한다. 제안하는 방법은 그림 8과 같이 빈 영역의 주변 화소의 깊이값을 비교하여 그 깊이값의 차이가 큰 경우에는 깊이값이 큰 위치의 주변 화소를 깊이값이 작은 위치의 화소로 대체한다. 즉, 전경과 배경 모두에 인접한 빈 영역의 인접한 화소를 인위적으로 배경에 속하는 화소들만 존재하도록 바꾸는 것이다.



(a) (b)

그림 8. 빈 영역의 인접 화소에 대한 처리: (a) 처리 전, (b) 배경에 속하는 인접 화소만 갖도록 처리

3. 실험 결과 및 분석

본 논문에서 제안한 영상합성 기술의 성능 실험을 위해 Ballet와 Breakdancers 영상에 대해 영상을 합성했다. 제안하는 기술은 현재 MPEG의 FTV/3D Video 그룹에서 사용중인 Nagoya 대학의 영상합성에 대한 참조 소프트웨어[5]를 기반으로 구현되었고, 각 기술이 영상합성에 어떠한 영향을 미치는지 주관적 화질과 객관적 화질을 통해 비교했다. 그림 9, 그림 10, 그림 11은 각 방법에 대한 합성영상의 예를 보여주고 있다.



(a) (b)

그림 9. 전처리에 대한 합성영상: (a) 전, (b) 후



(a) (b)

그림 10. 색상차 보정에 대한 합성영상: (a) 전, (b) 후



(a) (b)

그림 11. 빈영역 채우기에 대한 합성영상: (a) 전, (b) 후

표 1. 제안하는 영상합성 기술에 대한 화질 비교

방법	Ballet		Breakdancers	
	적용 전	적용 후	적용 전	적용 후
전처리	31.77dB	31.79dB	31.73dB	31.64dB
색상차 보정	31.77dB	31.59dB	31.73dB	31.88dB
빈영역 채우기	31.77dB	32.49dB	31.73dB	31.75dB

표 1은 깊이 영상의 전처리, 색상차 보정, 빈 영역 채우기에 대한 합성영상의 화질을 비교하여 보여주고 있다. 실험을 통해 알 수 있듯이 본 논문에서 제안하는 영상합성 기술은 객관적 화질과 주관적 화질면에서 기존의 방식보다 좋은 성능을 보였다.

4. 결론

본 논문에서는 DIBR 기반의 영상합성 기술에서 깊이영상의 초기 오류값을 보정하는 전처리 기술과 참조 영상간의 색상차로 인한 합성영상의 부자연스러움을 줄여주는 참조 영상간의 색상차 보정기술을 제안했다. 또한, 깊이 영상 기반의 빈 영역 채우기 기술을 제안하여, 전경과 배경 사이에 존재하는 비페색 영역을 효율적으로 채웠다. 실험을 통해 제안하는 영상합성 방법이 객관적 화질과 주관적 화질면에서 기존의 방법에 비해 좋은 성능을 보임을 확인했다.

감사의 글

본 연구는 광주과학기술원 실감방송연구센터를 통한 대학IT연구센터(ITRC)의 지원에 의한 것입니다.

참고문헌

- [1] ISO/IEC JTC1/SC29/WG11, "Survey of algorithms used for multi-view video coding (MVC)," Doc. N6909, Jan. 2005.
- [2] ISO/IEC JTC1/SC29/WG11, "Description of exploration experiments in 3-D video coding," Doc. N10173, Oct. 2008.
- [3] U. Fecker, M. Barkowsky, and A. Kaup, "Improving the prediction efficiency for multi-view video coding using histogram matching," Proc. of PCS, April 2006.
- [4] A. Telea, "An image inpainting technique based on the fast marching method," Journal of Graphics Tools, vol. 9, no. 1, pp. 25-36, 2004.
- [5] ISO/IEC JTC1/SC29/WG11, "Reference softwares for depth estimation and view synthesis," Doc. M15377, April 2008.