

움직임 예측을 이용한 깊이 영상 탐색 방법

이상범, 호요성
광주과학기술원

Depth Map Estimation Using Motion Estimation for 3D Video Coding

Sang-Beom Lee, Yo-sung Ho
Gwangju Institute of Science and Technology (GIST)

요 약

본 논문에서는 시간적 상관도를 향상시키는 깊이 비디오 탐색 방법을 제안한다. 제안하는 방법은 좌영상과 우영상을 이용하여 정합 오차를 계산하여 깊이값을 탐색하는 기존의 방법에 시간적 가중치 함수를 추가하여 깊이 비디오의 시간적 상관도를 높인다. 특히, 제안하는 방법은 시간적 가중치 함수를 적용하기 위해 움직임 예측 과정을 수행하여 움직이는 객체를 포함한 모든 영역에 대해 시간적 상관도를 향상시킨다. 실험 결과를 통해, 제안하는 방법이 깊이 비디오의 시간적 상관도를 향상시켜 합성 영상의 화질을 향상시켰고, 또한, 깊이 비디오의 부호화 효율을 증대시켰음을 확인할 수 있었다.

Keywords: depth video estimation, temporal consistency, multiview video, 3DTV

1. 서론

다시점 영상이란 동시에 다수의 카메라를 이용하여 동일한 3차원 장면을 여러 시점에서 촬영한 것을 말한다. 다시점 영상은 사용자가 마치 촬영된 장소에 있는 듯한 느낌을 받도록 하기 때문에 차세대 방송 시스템으로 각광받고 있다 [1]. 다시점 영상을 이용하게 되면 시청자가 원하는 시점을 선택적으로 재생할 수 있을 뿐만 아니라, 인접한 2시점, 또는 다시점 영상 전체를 이용하여 스테레오 모니터, 다시점 모니터 등의 3차원 디스플레이를 통해 시청자에게 실감나는 3차원 영상을 제공할 수 있다.

다시점 카메라로 장면을 촬영하게 되면 카메라의 물리적인 간격 때문에 시점을 이동할 때, 비교적 큰 시차를 만들게 된다. 큰 시차는 시청자로 하여금 불편함을 유발할 수 있기 때문에, 실제 촬영된 시점 사이에 가상의 중간시점 영상을 합성하여 부드러운 시점 전환을 가능하게 해야 한다.

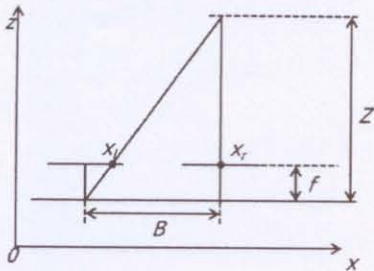
중간시점 영상을 합성하기 위해서는 깊이 영상을 필요로 하는데, 여기서 깊이 영상이란 영상 내에 존재하는 객체들의 3차원 거리 정보를 나타내는 영상을

말하며, 각 화소값은 해당 화소의 깊이 정보를 나타낸다. 깊이맵의 정확도는 합성된 중간시점 영상의 화질을 좌우하기 때문에 정확한 깊이맵을 생성하는 것은 매우 중요하다.

최근, 국제 표준화 그룹인 Moving Picture Experts Group (MPEG) 내의 3차원 비디오 부호화 그룹에서는 다시점 영상 및 깊이 영상 부호화의 중요성을 인지하여, 깊이 영상을 탐색할 수 있는 소프트웨어를 요청했다 [2]. 이에 대한 응답으로, 그래프 컷 기반의 깊이 영상 탐색 소프트웨어가 만들어졌고 이를 이용한 각종 실험이 진행 중이다 [3]. 이 소프트웨어는 비교적 좋은 화질의 깊이 영상을 제공하지만, 매 프레임마다 독립적으로 깊이 영상을 탐색하기 때문에 깊이 영상의 시간적 상관도가 낮다. 시간적 상관도 저하 문제는 깊이 영상을 이용하여 중간영상을 합성했을 때, 장면이 흔들리는 잡음을 발생시켜 시청자에게 시각적 불편함을 유발한다. 또한, 깊이 영상 자체의 부호화 과정에서 시간적 예측 효과를 저하시키기 때문에 깊이 영상의 부호화 효율을 떨어뜨린다.

본 논문에서는 깊이 영상 탐색 과정에서 발생하는 시간적 상관도 저하 문제를 해결하여 깊이 영상의 부

호화 효율을 높이는 방법을 제안한다. 제안하는 방법은 깊이 영상을 탐색할 때 이전 프레임에서 탐색한 깊이값을 고려한다. 또한, 움직임 예측을 이용하여 배경 영역과 움직이는 객체를 분리한 다음, 각각의 영역에 대해 시간적 상관도를 향상시킨다.



(그림 1) 변이와 깊이의 상관 관계

II. 다시점 깊이 비디오 탐색 방법

II.1 변이(disparity)와 깊이(depth)의 상관 관계

그림 1은 변이와 깊이의 상관 관계를 나타낸다. 그림 1에서 특정 3차원 점이 우측 영상 평면(image plane)에 사상(mapping)되고 이 점이 사상된 점의 우측 영상 평면 내에서의 위치는 (x_r, y_r) 라고 가정할 수 있다. 같은 방법으로 이 점의 좌측 영상 평면에서의 사상된 좌표는 (x_l, y_l) 라고 가정할 수 있다. 그렇다면, 변이 d 와 깊이 Z 의 관계는 다음과 같이 나타낼 수 있다.

$$Z = \frac{Bf}{d} = \frac{Bf}{x_l - x_r} \quad (1)$$

여기서 B 는 카메라 간 거리, f 는 각 카메라의 초점 거리를 나타낸다. 위 식을 통해, 다시점 비디오의 시점 간 상관도를 이용하여 변이를 탐색할 수 있다면, 우리는 3차원 장면의 실제 깊이도 탐색 가능하다는 것을 알 수 있다.

II.2 변이 계산

MPEG의 3차원 비디오 부호화 그룹은 깊이 비디오 탐색 및 시점 합성 소프트웨어를 제작하기 위한 노력을 지속했다. 그들은 깊이 비디오 탐색 및 시점 합성을 위한 다양한 기술들을 제안하고 채택함으로써 소프트웨어를 구현하고 배포하게 되었다. 이 가운데 깊이 비디오 탐색 소프트웨어는 변이 계산 과정, 그

래프 컷 기반의 오차 최소화 과정, 변이-깊이 변환 과정과 같이 크게 세 부분으로 분류될 수 있다.

깊이 비디오 탐색의 첫 번째 단계에서는 가운데 시점의 모든 화소에 대한 정합 오차를 계산한다. 다시점 영상은 최소 세 개 이상의 시점을 가지기 때문에, 우리는 스테레오 영상을 이용할 때와 다르게 좌 영상, 우영상 각각에 대한 정합 오차를 동시에 계산할 수 있다. 이렇게 함으로써, 스테레오 정합에서의 문제점인 폐색 영역 문제를 쉽게 해결할 수 있다. 다시점 영상을 이용한 깊이 비디오 탐색 방법의 정합 함수는 다음과 같이 정의된다.

$$E_{sim}(x, y, d) = \min\{E_L(x, y, d), E_R(x, y, d)\} \quad (2)$$

$$E_L(x, y, d) = |I_C(x, y) - I_L(x + d, y)| \quad (3)$$

$$E_R(x, y, d) = |I_C(x, y) - I_R(x - d, y)| \quad (4)$$

여기서 $I(x, y)$ 는 좌표 (x, y) 에서의 화소값을 나타낸다.

두 번째 단계는 그래프 컷 기반의 오차 최소화 과정이다. 그래프 컷 알고리즘은 주변 화소들의 정합 오차를 고려하여 화소들의 변이값을 할당하기 때문에 보다 신뢰할만한 변이값을 구할 수 있게 된다.

세 번째 단계는 변이-깊이 변환 과정이다. 깊이 영상은 가장 가까운 곳의 깊이를 255, 가장 먼 곳의 깊이를 0으로 하는 8비트 흑백 영상으로 표현될 수 있다. 다시 말해, 3차원 장면의 가장 가까운 곳과 가장 먼 곳 사이를 256개의 균일한 평면으로 분할할 수 있다. 그렇다면 좌표 (x, y) 에 대응하는 실제 깊이값인 Z 는 다음의 수식을 통해 8비트 흑백 영상의 화소값으로 변환될 수 있다.

$$v = \left\lfloor 255 - \frac{255(Z - Z_{near})}{Z_{far} - Z_{near}} + 0.5 \right\rfloor \quad (5)$$

여기서 Z_{far} 와 Z_{near} 는 각각 가장 가깝고 가장 먼 깊이값을 나타낸다.

III. 제안하는 깊이 비디오 탐색 방법

III.1 움직임 예측을 이용한 시간적 가중치 함수

깊이 영상 탐색 방법은 매 프레임마다 독립적으로 깊이 영상을 탐색하기 때문에 시간적 상관도가 떨어지는 문제점을 갖고 있다. 따라서, 제안하는 방법은 현재 프레임의 깊이 영상을 탐색할 때, 이전 프레임의 깊이값을 참조하는 새로운 정합 함수를 사용하여 깊이 영상의 시간적 상관도를 향상시킨다. 새로운 정

합 함수에는 이전 프레임에서의 깊이값을 고려하는 가중치 함수가 추가된다. 시간적 상관도를 고려한 새로운 정합 함수는 다음과 같이 나타낼 수 있다 [4].

$$E_{new}(x, y, d) = E_{sim}(x, y, d) + E_{temp}(x, y, d) \quad (6)$$

$$E_{temp}(x, y, d) = \lambda |d - D_{prev}(x + \Delta x, y + \Delta y)| \quad (7)$$

여기서 λ 는 가중치 함수의 기울기를 나타내고, $D_{prev}(x, y)$ 는 이전 프레임의 좌표 (x, y) 의 깊이값을 나타낸다.

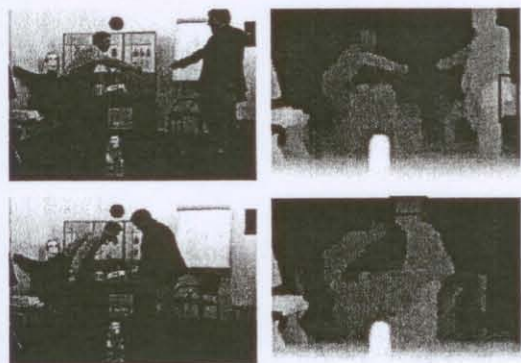
제안하는 방법은 움직이는 객체 분리를 위해서 영상 내의 16x16 크기의 블록에 대해서 평균 절대차(mean absolute difference, MAD)를 계산한 다음, 임계치를 이용하여 해당 블록이 움직임이 있는지 없는지 여부를 판단한다. 움직이는 객체에 대해 시간적 가중치 함수를 적용하기 위해, 제안하는 방법은 움직임 예측 기술을 사용한다. 앞서 언급한 가중치 함수의 기울기 λ 는 다음과 같이 정의될 수 있다.

$$\lambda = \begin{cases} 1 & \text{if } MAD_k < Th \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

여기서 MAD_k 는 좌표 (x, y) 를 포함하는 k번째 블록의 MAD를 나타내고 Th 는 임계치를 나타낸다.

III.2 깊이값의 오차 전파 방지

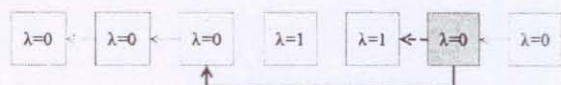
깊이값 오차 최소화 과정에서 그래프 컷 알고리즘을 사용하면 깊이 영상 경계 주변에 깊이값 오차가 발생한다. 이러한 문제점은 전경의 객체가 장면 밖에서 갑자기 안으로 출현할 때 혹은 영상의 가장자리 부근으로 접근할 때 가장자리 주변의 배경 영역이 전경의 객체의 깊이값을 가지면서 발생한다. 그림 2는 가장자리 주변의 깊이값 오차를 보여준다.



(그림 2) 깊이 영상의 깊이값 오차

만약 오차가 존재하는 영상 가장자리 주변 영역이 배경으로 검출될 경우, 시간적 가중치 함수는 잘못된 깊이값을 참조하게 되어 이 오차가 다른 프레임으로 잘못 전파된다.

이러한 시간적 상관도 향상 방법의 심각한 문제점을 해결하기 위해, 영상 가장자리의 오차 전파를 방지하는 방법이 요구된다. 따라서, 제안하는 방법은 만약 λ 가 0일 경우에는 시간적 가중치 함수가 동일한 부분의 λ 가 0인 이전 프레임을 참조한다. 다시 말해, 이전 프레임들 가운데 λ 값이 0인 가장 최근 프레임을 참조하도록 한다. 그림 3은 시간적 가중치 함수를 적용할 때 이전 프레임의 참조 방법을 나타낸다. 그림 3에서 보면 굵은 선은 가장자리 영역에 대한 참조 방법, 점선은 일반적인 참조 방법을 나타낸다.



(그림 3) 가중치 함수의 이전 프레임 참조 방법

IV. 실험 결과 및 분석

제안하는 방법의 성능 평가를 위해, MPEG의 3차원 비디오 부호화 그룹에서 테스트 영상으로 사용 중인 "Book Arrival", "Lovebird1", "Newspaper" 영상을 사용했다.

깊이 정보의 정확도 평가를 위해서, 각 영상의 깊이 영상을 탐색하였고, 이를 이용하여 영상을 합성한 후, 원본 영상과 비교하였다. 프레임 수는 100프레임으로 정했다.

표 1은 합성 영상의 PSNR을 나타낸다. 표 1을 통해 제안하는 시간적 상관도 향상 방법이 깊이 영상의 정확도를 그대로 유지시켰다는 것을 확인하였다.

(표 1) 합성 영상의 평균 PSNR

테스트 영상	시점	기존의 방법	제안하는 방법	Δ PSNR (dB)
Book Arrival	8	34.3986	34.4529	0.0542
	9	35.5694	35.5472	-0.0222
Lovebird1	6	30.9870	30.9866	-0.0004
	7	30.4126	30.4214	0.0087
Newspaper	4	24.3732	24.3695	-0.0037
	5	25.4442	25.4389	-0.0054
평균		31.4685	31.5019	0.0334

제안하는 방법이 깊이 영상의 부호화에 미치는 영향을 평가하기 위해서, MPEG 3차원 비디오 부호화 그룹에서 배포한 깊이 영상 탐색 소프트웨어로 얻은 깊이 영상과 제안하는 방법을 통해 얻은 깊이 영상 각각을 H.264 부호화기를 이용하여 부호화하고 이를 비교하였다. 프레임 수는 100 프레임으로 정하였고, GOP 구조는 IPPP... 구조를 이용하였다. 표 2와 그림 4는 깊이 비디오 각각을 부호화한 결과를 나타낸다. 부호화 결과를 통해 제안하는 방법이 평균 약 2.56dB의 화질 향상 혹은 42.55%의 비트율 감소 효과를 가져온다는 것을 확인하였다 [5].

[표 2] 깊이 비디오 부호화 결과

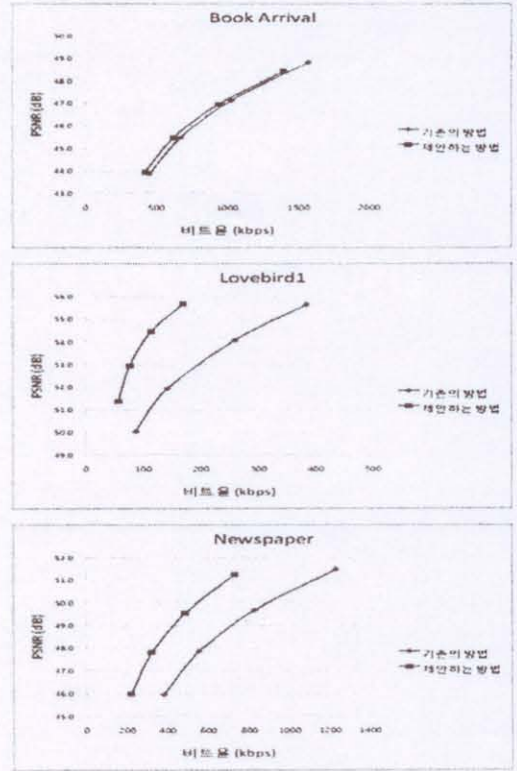
테스트 영상	QP	기존의 방법		제안하는 방법	
		PSNR (dB)	비트율 (kbps)	PSNR (dB)	비트율 (kbps)
Book Arrival	22	48.83	1573.54	48.43	1396.91
	25	47.13	1032.20	46.96	941.86
	28	45.48	677.70	45.46	625.13
	31	43.86	456.22	43.92	423.30
Lovebird1	22	55.64	384.62	55.66	170.26
	25	54.07	260.00	54.44	114.63
	28	51.92	142.81	52.96	77.36
	31	50.06	87.58	51.37	57.33
Newspaper	22	51.47	1229.67	51.25	730.68
	25	49.67	829.03	49.53	488.31
	28	47.83	556.19	47.78	323.41
	31	45.95	383.23	45.97	221.28

V. 결론

본 논문에서는 깊이 영상의 탐색 과정에서 시간적 상관도를 고려하여 깊이 영상의 부호화 효율을 높이는 방법을 제안하였다. 제안하는 방법은 기존의 정합 함수에 가중치 함수를 추가했고, 움직임 예측을 통해 모든 영역에 대해서 시간적 상관도를 높였다. 제안하는 방법은 기존의 방법에 비해 평균 약 2.56dB의 화질을 향상시켰으며, 42.55%의 비트율을 감소시켰다. 결과적으로 제안하는 방법이 깊이 영상의 화질을 그대로 유지하면서도 깊이 영상의 부호화 효율을 높였음을 알 수 있었다.

감사의 글

본 연구는 지식경제부 및 정보통신산업진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음 (NIPA-2009-(C1090-0902-0017))



[그림 4] 깊이 비디오 부호화 R-D 곡선

참고문헌

- [1] A. Smolic, K. Muller, P. Merkle, C. Fehn, P. Kauff, P. Eisert, T. Wiegand, "3D Video and Free Viewpoint Video ? Technologies, Applications and MPEG Standards," in Proc. of IEEE International Conference on Multimedia and Expo, pp. 2161-2164, 2006.
- [2] ISO/IEC JTC1/SC29/WG11, "Call for Contributions on 3D Video Test Material," N9595, Jan. 2008.
- [3] ISO/IEC JTC1/SC29/WG11, "Reference Software of Depth Estimation and View Synthesis for FTV/3DV," M15836, Oct. 2008.
- [4] S. Lee, C. Lee, Yo-Sung Ho, "Temporal Consistency Enhancement of Background for Depth Estimation," 3D Systems and Applications (3DSA), pp.S08-03(1~4), 2009.
- [5] ITU-T SG16 Q.6, "An Excel Add-in for Computing Bjontegaard Metric and Its Evolution," VCEG-AE07, Marrakech, MA, Jan. 2007.