

블록 초점 정보의 보간을 통한 새로운 초점 기반 2D-to-3D 변환

*한찬희 *김준호 **강현수 *김진수 *최해철 *이시웅

*한밭대학교 정보통신전문대학원 **충북대학교 정보통신공학부

*chani@hanbat.ac.kr

Novel Focus-based 2D-to-3D Conversion using Interpolation of Block Focus

*Han, Chan-Hee *Kim, June-Ho **Kang, Hyun-Soo *Kim, Jin-Soo *Choi, Hae-Chul

*Lee, Si-Woong

*Graduate School of Information & Communications, Hanbat National University

** School of Information & Communications, Chungbuk National University

요약

3차원 동영상은 수많은 응용분야에서 차세대 멀티미디어 콘텐츠로 큰 주목을 받고 있지만 2차원에서 3차원 콘텐츠로의 과도기인 현 시점에서 3차원 콘텐츠의 부족현상은 관련 산업분야의 큰 문제점으로 떠오르고 있다. 2D-to-3D 변환 기술은 높은 비용 없이 기존의 방대한 2차원 콘텐츠를 3차원 콘텐츠로 재사용하는 것이 가능하기 때문에 3차원 콘텐츠의 부족 문제를 해결할 수 있는 기술로 큰 관심을 끌고 있다. 본 논문에서는 블록 단위 초점 정보의 보간을 통한 새로운 초점 기반 3차원 변환 기법을 제안한다. 기존의 화소 단위 초점 측정치를 그대로 깊이 정보로 이용하는 경우나 분할 영역 단위 초점 측정치를 깊이 정보로 활용하는 경우는 이웃 화소간 깊이정보의 불연속성이 발생할 수 있지만 제안방식은 블록 초점 측정치의 보간수행으로 인해 이웃 화소 사이에서 뿐만 아니라 블록경계 혹은 영역경계에서도 특별한 스무딩필터 처리 없이도 화소간 깊이정보의 불연속성이 발생하지 않게 된다. 실험결과를 통해 제안한 방식이 기존의 방식들 보다 더 자연스러운 깊이 정보 추출 성능을 보여준다는 것을 알 수 있다.

1. 서론

2차원 디스플레이의 기술 개발이 최근 어느 정도 포화 상태에 다름에 따라, 3차원 디스플레이 기술이 차세대 디스플레이로 큰 주목을 받고 있으며, 사용자의 욕구 또한 2차원 영상에서 3차원 영상으로 변화하고 있다. 3차원 입체영상기술은 차세대 멀티미디어 분야의 핵심 기술로 자리 잡을 것이며, 그 응용분야 또한 정보통신, 방송, 영화, 의료, 게임, 가상현실 등 매우 다양할 것으로 예상할 수 있다. 인간이 입체감을 느끼는 것은 물체를 볼 때 두 눈이 일정 깊이 떨어져 있어 좌안과 우안에 차이가 있는 다른 영상이 각각 맺히게 되는데 이것을 양안 시차(binocular disparity)라 하며, 뇌에서는 양안 시차가 있는 두 영상을 종합적으로 판단하여 하나의 영상으로 인지하게 되기 때문이다. 이러한 원리를 기반으로 동일한 장면에 대해 스테레오 혹은 다시점 영상을 생성하여 디스플레이 하면 3차원 입체감을 느낄 수 있다.



그림 1 스테레오 카메라에 의한 3차원 입체감 인지

현재까지 다양한 형태의 3차원 입체영상기술이 연구, 발표 되었지만 보편화를 위해서는 아직 해상도, 시야각, 관찰자의 피로 등 해결해야 할 많은 문제점이 존재하고 있다. 그 중에서도 가장 큰 문제점은 3차원 디스플레이용 콘텐츠의 부족이다. 3차원 입체영상은 주로 스테레오 카메라나 영상편집 툴을 사용하여 획득하지만 스테레오 카메라를 이용한 제작비용의 한계, 영상편집 툴에 의한 시간소요가 문제가 되어 아직까지 방송이나 영화와 같은 영상매체 시장에서는 2차원 콘텐츠로 제작하는 경우가 더 많다. 이러한 환경 하에서 3차원 디스플레이 시장의 활성화를 위해 기존의 방대한 2차원 콘텐츠를 높은 추가비용 없이 3차원 콘텐츠로 확보할 수 있는 2차원 동영상의 3차원(2D-to-3D) 변환 기술이 크게 주목 받고 있다.

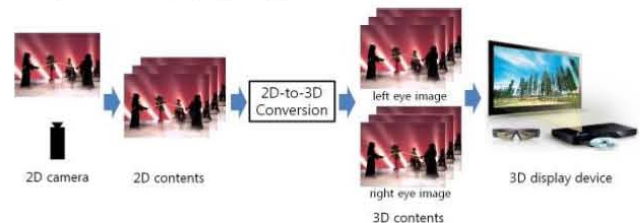


그림 2 2D-to-3D 변환에 의한 3차원 입체감 인지

인간이 입체감을 인지할 수 있는 요소로는 앞서 언급한 양안 시차

이외에도 몇 가지가 존재한다. 실제로 인간은 한쪽 눈만으로도 입체감 즉, 현재 위치와 대상들 간의 상대적인 깊이를 인지하는 것이 가능하며, 그 요인으로는 장면의 기하학적 구성, 폐색 구간, 객체의 크기나 질감, 인간의 경험적 사고, 객체의 흐릿한 정도, 객체의 움직임 등[1]이 있다. 이러한 요소들은 2차원 영상 내에서 존재할 수 있으며 이를 3차원 변환을 위한 단안 단서(monocular cue)라 한다. 이러한 단서들 중 영상 각 위치의 흐릿한 정도를 측정하여 깊이정보를 추출하는 방법들 [2,3]을 초점 기반 깊이 추출(depth-from-focus)이라 하며 텍스처 기반(depth-from-texture) 혹은 경계 기반(depth-from-edge)이라고도 한다. 그 기본개념은 그림 3에서와 같이 카메라로부터 가까운 영역의 텍스처는 먼 영역보다 상세하고 특히 초점이 맞춰진 영역의 경계는 그렇지 않은 영역의 경계보다 더 날카롭다는 것이다.



그림 3 2D-to-3D 변환에 의한 3차원 입체감 인지

[2]에서는 초점 연산자를 정의하여 영상의 각 영역에서 초점 정도를 계산한다. 연산자의 종류에는 local scale 기반 연산, sobel mask 기반 연산, 표준편차 기반 연산 등이 있다. 화소 단위로 초점 측정치를 그대로 깊이 정보로 이용하는 경우나 분할 영역 단위로 초점 측정치를 깊이 정보로 활용하는 경우는 이웃 화소간 거리정보의 불연속성이 발생할 수 있다. 이를 해결하기 위해 [2]에서는 비대칭 가우시안 스무딩 필터를 적용한다. 하지만 필터의 크기가 너무 작으면 경계영역이나 객체 내부에서 깊이 정보가 부정확해 질 수 있고 필터의 크기가 커질수록 연산량이 매우 많아지는 단점이 있다. 본 논문에서는 이러한 문제점을 극복하기 위해 블록 단위 초점 정보의 보간을 통한 새로운 초점 기반 2D-to-3D 변환 기법을 제안한다. 제안방식은 블록 초점 측정치의 보간수행으로 인해 이웃 화소 사이에서 뿐만 아니라 블록경계 혹은 영역경계에서도 특별한 스무딩 필터 처리 없이도 화소간 거리정보의 불연속성이 발생하지 않게 된다.

2. 블록 단위 초점 측정치의 보간을 이용한 깊이 추출

제안방식에서는 우선 영상을 일정한 블록크기로 분할 한 후 각 블록에서의 초점정도를 계산한다. 이때 그림 4와 같이 이웃한 블록들 내부의 화소들도 일부 포함시킨다. 이는 현재 초점 정도가 측정되는 블록에서 그 블록뿐만 아니라 주변상황까지도 고려하기 위함이다.

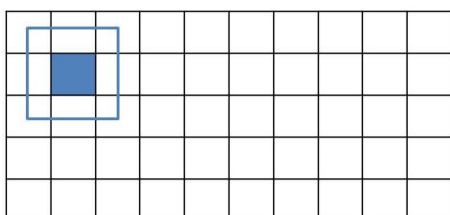


그림 4 블록 별 초점정도의 측정영역

각 블록의 초점정도가 계산된 후에는 그림 5와 같이 각 블록 중앙에 계산된 초점량을 삽입한 후 영상 전체적으로 보간을 수행한다.

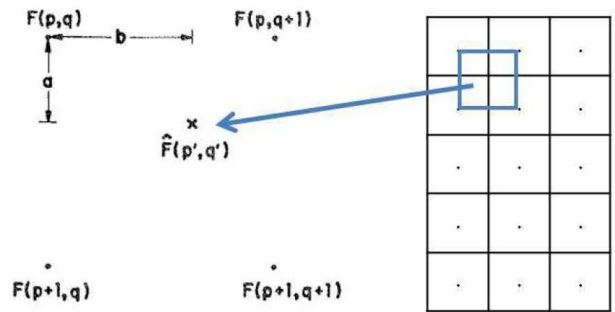


그림 5 블록 초점량의 보간

보간될 화소위치에서의 초점량은 다음 식(1)과 같다. 여기서, $F'(p',q')$ 이 보간될 화소위치에서의 초점량이다.

$$F'(p',q') = (1-a)[(1-b)F(p,q) + bF(p,q+1)] + a[(1-b)F(p+1,q) + bF(p+1,q+1)] \quad (1)$$

영상 모든 위치에서 보간 수행으로 초점량이 계산된 후에는 깊이 값으로 스케일링(scaling)을 수행한다.

3. 실험결과

그림 6(b)와 (c)는 각각 화소 단위와 분할 영역 단위 초점 측정치를 필터링하여 얻은 깊이 정보들을 보여주고 있으며, 그림 6(d)는 제안방식의 결과를 보여주고 있다. 실험결과를 통해 제안방식의 경우 기존의 방식들에 비해 이웃 화소 간의 불연속성이 발생하지 않고 객체 내부와 경계부분에서의 깊이 표현이 더 자연스러운 것을 확인할 수 있다.



(a) 원영상 (b) 화소 단위 필터링 방식 (c) 분할 영역 단위 필터링 방식 (d) 제안 방식

그림 6 제안방식의 주관적 성능 비교

표 1 제안방식의 복잡도 성능 비교

	복잡도
제안방식	1
화소 단위 방식	1.1
영역 단위 방식	2.4

표 1은 각 방식의 복잡도를 나타낸다. 여기서 복잡도는 최종 깊이 맵 추출까지의 수행시간으로 하였으며, 제안방식의 복잡도를 1로 하였을 때 기존방식들의 복잡도를 비율로 나타내었다.

복잡도 성능 비교 결과를 통해 화소 단위 초점 측정치를 필터링하는 경우와는 복잡도 측면에서 큰 차이를 보이지 않는다는 것을 알 수 있다. 이는 실험에서 정한 크기로 스무딩 필터를 적용하는 복잡도나 제안방식의 보간수행에 소요되는 복잡도가 큰 차이가 없기 때문이다. 하지만 화소 단위 방식이 제안방식의 주관적 성능을 달성하기 위해서는 더욱 큰 크기의 스무딩 필터 적용이 필요하므로 유사한 주관적 성능 하에서는 복잡도가 상당히 클 것으로 예상된다. 반면 분할 영역 단위 초점 측정치를 필터링하는 경우에는 필터링 적용뿐만 아니라 영역 분할에 소요되는 시간이 있기 때문에 다른 방식들에 비해 복잡도가 상당히 크다는 것을 알 수 있다.

제안방식의 주관적 성능과 복잡도 비교 결과를 통해 제안방식이 낮은 복잡도에서도 우수한 주관적 성능을 달성할 수 있음을 확인할 수 있다.

4. 결론

본 논문에서는 블록 단위 초점 정보의 보간을 통한 새로운 초점 기반 3차원 변환 기법을 제안하였다. 제안방식에서는 영상을 일정한 블록크기로 분할하여 각 블록에서 초점정도를 계산한 후 블록 중앙에 계산된 초점값을 삽입하고 그 외 위치에서는 보간을 수행하였다. 기존의 화소 단위 초점 측정치를 그대로 깊이 정보로 이용하는 경우나 분할 영역 단위 초점 측정치를 깊이 정보로 활용하는 경우는 이웃 화소 간 깊이정보의 불연속성이 발생하는 문제를 해결하기 위해 스무딩 필터를 적용하였다. 하지만 필터의 크기가 너무 작으면 객체 내부로 갈수록 깊이 정보가 부정확해 질 수 있고 필터의 크기가 커질수록 복잡도가 매우 커지는 단점이 있었다. 제안방식은 블록 초점 측정치의 보간 수행으로 인해 이웃 화소 사이에서 뿐만 아니라 블록경계 혹은 영역경계에서도 특별한 스무딩필터 처리 없이도 화소간 깊이정보의 불연속성이 발생하지 않게 된다. 실험결과를 통해 제안한 방식이 기존의 방식들 보다 깊이 정보의 주관적 성능과 복잡도 비교에서 더 우수하다는 것을 알 수 있다.

5. 참고문헌

- [1] L. Zhang, C. Vazquez, and S. Knorr, "3D-TV content creation: Automatic 2D-to-3D video conversion," *IEEE Transactions on Broadcasting*, Vol. 57, No. 2, pp. 372 - 383, 2011.
- [2] W. J. Tam, A. Soung Yee, J. Ferreira, S. Tariq, and F. Speranza, "Stereoscopic image rendering based on depth maps created from blur and edge information," *Stereoscopic Displays and Applications XII*, Vol. 5664, pp.104-115, 2005.
- [3] C. C. Cheng, C. T. Li, and L. G. Chen, "A 2D-to-3D Conversion System Using Edge Information," *IEEE International Conference on Consumer Electronics(ICCE)*, Vol. 56, pp.1739-1745, 2010.