

신뢰 확산을 이용한 영역분할 기반의 다시점 깊이맵 추정 기법

이상범, 오관정, 호요성
광주과학기술원 정보통신공학과
e-mail: {sblee, kjoh81, hoyo}@gist.ac.kr

Segmentation-based Multi-view Depth Map Estimation Using Modified Belief Propagation

Sang-Beom Lee, Kwan-Jung Oh and Yo-Sung Ho
Gwangju Institute of Science and Technology (GIST)

요 약

본 논문에서는 다시점 영상을 이용한 새로운 깊이맵 추정 기법을 제안한다. 제안하는 방법은 영역분할(segmentation) 기법을 통해 얻어진 세그먼트(segment)와 자체적응(self-adaptation) 정합함수를 이용하여 초기 깊이맵을 추정하며, 기존의 방법들이 가지고 있던 입력 영상의 제한성을 해결하기 위해 3차원 워핑 기법을 적용하여 깊이값을 탐색한다. 특히, 3차원 워핑 기법은 변이맵에서 깊이맵으로 변환하는 과정에서 발생하는 깊이값의 오차를 줄인다. 정제(refinement) 과정에서는 세그먼트 단위의 신뢰확산(belief propagation) 기법을 이용하여 최종 깊이맵을 획득한다.

1. 서론

디지털 기술의 발전과 방송 통신의 융합으로 방송 매체가 다양해짐에 따라, 디지털 기술의 특성을 이용한 방송 관련 부가 서비스들이 새롭게 선보이고 있다. 지금까지의 방송은 TV 화면에 평면적인 화상 및 음성 등의 한정된 정보만을 전송하였다. 하지만 TV의 디지털화 및 방송 네트워크의 지능화와 고도화가 진행되면서 이동 방송, 데이터 방송, 휴대 인터넷 방송, IPTV 및 T-Commerce 등의 다양한 방송 멀티미디어 서비스가 제공되기 시작했다. 현재 방송 중인 TV의 발전 방향은 고화질과 화면 크기의 확대로 가고 있지만 TV 화면 자체가 2차원이기 때문에 TV를 통해서 입체감을 느낄 수는 없다. 다시 말해, 실제 원거리에 떨어져 있는 물체를 TV를 통해 현실감 있게 접하고자 하는 시청자들의 요구 사항을 제대로 만족시키지 못하고 있는 실정이다.

최근 컴퓨터 그래픽스, 몰입형(immersive) 디스플레이, 디지털 영상통신 기술이 급속히 발전함에 따라, 우리는 시뮬레이션을 이용하여 현실 세계를 재창조하고 이를 경험할 수 있게 되었다. 이 가운데 다시점 영상(multi-view video) 재현이 가능한 몰입형 디스플레이 장치는 현실 세계를 재구성한 콘텐츠로부터 현실감 있는 느낌을 사용자에게 제공할 수 있다. 여기서 다시점 영상이란 똑같은 3차원 장면을 같은 시간에 여러 시점에서 다수의 카메라로 촬영한 것을 말하며, 입체 TV, 자유 시점 TV, 감시 카메라 영상 등 다양한 영상에 응용될 수 있다

[1].

현재 방송 중인 TV의 발전 방향은 고화질과 화면 크기의 확대로 가고 있지만 TV 화면 자체가 2차원이기 때문에 TV를 통해서 입체감을 느낄 수는 없다. 시청자에게 입체감과 현실감을 제공하기 위한 요소 중의 하나가 바로 시각 정보이므로 다시점 영상은 시청자에게 원하는 시점의 영상을 제공하여 깊이감과 입체감을 제공해 준다. 현재 3차원 TV는 많은 관심을 받고 있으며 점차 현재 사용하고 있는 2차원 TV를 대체할 것으로 기대를 받고 있다.

3차원 TV는 시청자에게 원하는 시점의 영상을 제공하는 것을 목적으로 한다. 부드러운 시점 전환을 위해서는 3차원 TV의 핵심기술 중 하나인 중간영상 생성 방법(intermediate view reconstruction, IVR)이 필수적인데, 중간영상을 생성하기 위해서는 카메라와 객체 간의 거리 정보를 가지는 깊이맵(depth map)을 이용해야 한다.

깊이맵의 정확도는 중간영상의 화질과 밀접한 연관이 있기 때문에 정확한 깊이맵을 추정하는 것이 매우 중요하다. 하지만 깊이맵을 얻기 위한 기존의 스테레오 정합 알고리즘[2]은 깊이값 불연속점을 가지는 화소 주변에서 가장 많은 오차를 발생시킨다. 이러한 오차는 중간영상 생성 과정에서 객체의 경계를 불분명하게 만들기 때문에 객체의 경계에 민감한 시청자로 하여금 불편함을 줄 수 있다. 또한 기존의 방법들은 입력 영상의 제한성, 즉, 병렬 카메라 구성에서 촬영된 영상 혹은 교정(rectification)된 영상만을 사용해야 한다는 문제점을 지니고 있다.

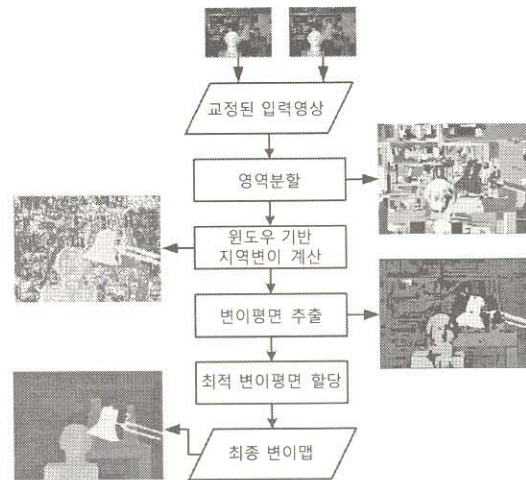
본 논문에서는 새로운 깊이맵 추정 기법을 제안한다. 제안하는 방법은 깊이값 불연속점에서의 오차를 줄이기 위해 영역분할(segmentation) 기반의 깊이맵 추정 기법을 사용한다. 영역분할 기법을 통해 얻어진 각 세그먼트(segment)들을 3차원 워핑을 적용하여 깊이값을 탐색한다. 3차원 워핑을 이용한 깊이값 탐색 방법은 변이맵(disparity map)에서 깊이맵으로 변환하는 과정에서 발생하는 깊이값의 오차를 줄일 수 있다. 깊이값 탐색 과정에서는 세그먼트 단위의 새로운 자체적응(self-adaptation) 정합함수를 이용하여 초기 깊이맵을 추정한다. 또한, 기존의 스테레오 정합 알고리즘들은 깊이맵을 추정할 때 좌영상 혹은 우영상 하나만을 사용할 수 있었지만 제안하는 방법은 깊이맵 추정 시에 좌영상과 우영상 모두 사용하여 깊이값을 탐색하므로 보다 정확한 깊이값을 얻을 수 있다. 정제(refinement) 과정에서는 최근 가장 많이 쓰이고 성능 면에서도 가장 우수한 평가를 받는 신뢰확산(belief propagation) 기법을 사용한다.

2. 깊이맵 추정 기법

2.1 영역분할 기반 깊이맵 추정 기법

기존의 스테레오 정합 알고리즘은 깊이값 불연속점 주변에서 많은 오차가 발생한다. 앞서 언급했듯이, 이러한 오차는 중간영상 생성시 객체의 경계가 중복되거나 불분명해지는 문제점을 야기한다. 최근에는 이와 같은 문제점을 해결하기 위해 영역분할을 이용하는 깊이맵 추정 기법이 제안되었다 [3]. 이 방법은 영상 내의 비슷한 화소들끼리 묶어 영상을 세그먼트들로 분할한 다음, 세그먼트 각각에 대해 하나의 깊이값을 할당하는 방식을 취하고 있다. 영역분할 기반의 깊이맵 추정 기법은 분할된 영역 내부에서는 깊이값이 미세하게 변화하며 깊이값의 불연속점은 세그먼트의 경계에서 발생한다고 가정한다.

영역분할 기반의 스테레오 정합 알고리즘은 그림 1에서 도식화한 것처럼 일반적으로 네 단계의 연속적인 절차를 거친다 [3]. 첫째, 영역분할 알고리즘을 적용하여 비슷한 화소들끼리 하나의 세그먼트로 묶는다. 둘째, 초기 변



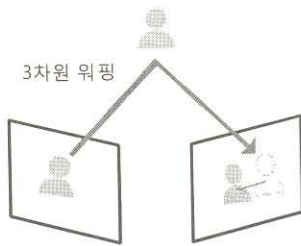
(그림 1) 영역분할 기반 스테레오 정합 알고리즘의 일반적인 블록 다이어그램

이맵 구성을 위해 윈도우 기반의 지역변이 탐색 방법을 적용한다. 셋째, 변이평면을 추출하기 위해 평면 조정 방법을 사용한다. 넷째, 추출된 변이평면을 이용하여 최적의 변이평면을 할당한다. 이와 같은 방법은 기존의 방법들에 비해 객체의 경계가 명확해질 뿐만 아니라, 보다 정확한 변이맵을 추정할 수 있다.

2.2 3차원 워핑을 이용한 깊이맵 추정 기법

앞서 언급했듯이, 기존의 스테레오 정합 알고리즘은 병렬 카메라 구성에서 얻은 영상 혹은 교정(rectification) 과정을 거친 영상만을 입력으로 가질 수 있다. 이는 변이를 구하기 위해 주변 영상에서 수평 방향으로만 탐색을 하기 때문인데, 이같은 방법은 다양한 카메라 구성, 즉, 병렬 카메라 구성뿐만 아니라 원호 카메라 구성 등을 갖는 다시점 영상에 대한 깊이맵을 추정하는 데에는 한계가 있다.

이와 같은 문제점을 보완하기 위해 3차원 워핑을 이용한 깊이맵 추정 기법이 제안되었다 [4]. 이 방법은 깊이맵을 구하고자 하는 시점에서의 영상의 특정 세그먼트를 3차원 워핑을 통하여 주변 영상으로 투영(projection)한 다음 깊이값을 탐색하는 방법을 이용한다. 이같은 방법을 이용하면 카메라 구성과 독립적으로 깊이맵 추정 방법을 사용할 수 있게 된다. 그림 2에 3차원 워핑을 이용한 깊이값 탐색 방법을 도식화하였다.



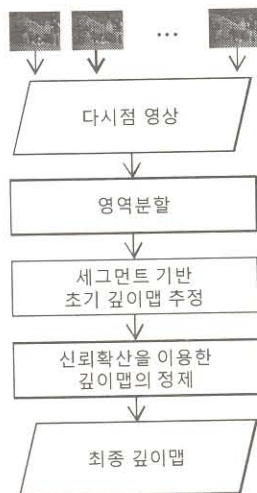
(그림 2) 3차원 워핑을 이용한 깊이값 탐색 방법

2.3 신뢰확산을 통한 정제 방법

초기 깊이맵 탐색 과정을 통해 얻어진 깊이맵은 많은 오차를 갖게 된다. 왜냐하면 영상 내부의 배경에 대한 깊이값을 탐색하는 경우, 배경 내에 존재하는 화소들의 색상 차이가 그리 크지 않기 때문에 잘못된 깊이값에서도 최소의 오차가 발생하여 이를 정확한 깊이값으로 인식할 수 있기 때문이다. 이와 같은 문제점을 해결하기 위해 현재까지 그래프 컷 (graph cut), 동적 프로그래밍 (dynamic programming), 신뢰확산 (belief propagation) 등의 많은 정제 (refinement) 방법들이 제안되었는데 현재 신뢰확산 방법이 가장 좋은 성능으로 평가받고 있다.

3. 다시점 영상의 깊이맵 추정 기법

본 논문에서는 다시점 영상에 적용 가능한 깊이맵 추정 기법을 제안한다. 그림 3은 제안하는 방법의 블록 다이어그램을 나타낸다.



(그림 3) 제안하는 방법의 블록 다이어그램

먼저 영역분할 기법을 이용하여 영상을 세그먼트들로 세분화한다. 또한, 제안하는 방법에서는 세그먼트 단위로 초기 깊이맵을 탐색한다. 기존의 방법은 화소 단위로 깊이값을 탐색하는데, 스테레오 정합 테스트 영상에 비해 상대적으로 큰 다시점 영상을 화소 단위로 탐색하게 되면 많은 오차를 포함할 뿐만 아니라 복잡도도 증가하게 되기 때문이다. 초기 깊이맵을 탐색하는 과정에서는 3차원 워

핑 기법과 자체적응 함수를 사용한다. 마지막 단계인 정제 과정에서는 세그먼트 단위의 신뢰확산 방법을 사용하는데 정제 과정을 반복하여 깊이맵의 오차를 줄인다. 결과적으로 제안하는 방법을 통해 정제된 깊이맵을 추정한다.

3.1 영역분할 방법

제안하는 방법의 첫번째 단계에서는 깊이맵을 구하고자 하는 시점에서의 영상을 유사한 색상 정보를 갖는 세그먼트들로 세분화한다. 앞서 언급한 것처럼, 깊이값의 불연속점은 세그먼트의 경계 부분에서 발생하는데 이 가정을 만족시키기 위해서는 세그먼트의 크기가 작도록 세분화하는 것이 바람직하다. 제안하는 방법에서는 이와 같이 세그먼트가 깊이값의 불연속점을 포함하지 않는 조건을 만족시키기 위해 충분한 Felzenszwalb의 그래프 기반의 영역분할 알고리즘[5]을 사용한다.

3.2 세그먼트 기반 초기 깊이맵 추정 방법

제안하는 방법은 영역분할 방법을 통해 얻어진 세그먼트를 단위로 깊이값을 탐색한다. 가장 많이 알려진 정합 함수는 SD(Squared intensity Difference)와 AD(Absolute intensity Difference)이지만 이 함수들은 카메라간 색상의 불일치에 민감하므로 이에 강인한 정합 함수가 필요하다. 따라서 제안하는 방법에서 사용하는 자체적응 함수[3]는 기존의 AD 함수에 경사도 맵(gradient map)을 이용한 함수를 추가한다. 제안하는 방법은 경사도 맵의 강인성을 높이기 위해 기존의 방법이 $+x$ 방향과 $+y$ 방향의 경사도 맵을 사용한 것을 확장하여 4방향 모두에 대한 경사도 맵을 사용한다. 세그먼트 단위의 자체적응 함수는 다음과 같다.

$$C(x, y, d) = (1 - \omega) \times C_{MAD}(x, y, d) + \omega \times C_{MGRAD}(x, y, d) \quad (1)$$

여기서 ω 는 가중치를 나타내며 $C_{MGRAD}(x, y, d)$ 와 $C_{MAD}(x, y, d)$ 는 각각 다음과 같이 정의된다.

$$C_{MGRAD}(x, y, d) = \frac{1}{M} \sum_{(x', y') \in S_k} \{ |\nabla_x I_1(x, y) - \nabla_x I_2(x', y')| + |\nabla_y I_1(x, y) - \nabla_y I_2(x', y')| + |\nabla_{-x} I_1(x, y) - \nabla_{-x} I_2(x', y')| + |\nabla_{-y} I_1(x, y) - \nabla_{-y} I_2(x', y')| \} \quad (2)$$

$$C_{MAD}(x, y, d) = \frac{1}{M} \sum_{(x', y') \in S_k} |I_1(x, y) - I_2(x', y')| \quad (3)$$

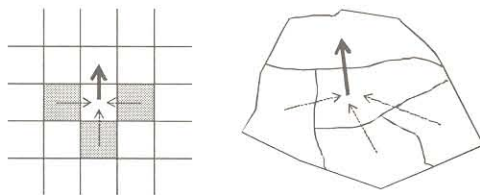
여기서 M 은 세그먼트 내의 화소의 개수, S_k 는 해당 세그먼트, (x', y') 는 3차원 워핑된 주변 영상에서의 위치, $\nabla_x, \nabla_y, \nabla_{-x}, \nabla_{-y}$ 는 각각 $+x$ 방향, $+y$ 방향, $-x$ 방향, $-y$ 방향의 경사도 맵을 의미한다.

다시점 영상에서는 하나의 세그먼트에 대한 깊이값 탐

색 과정에서 좌영상과 우영상 모두에 대해 정합 함수를 적용하여 보다 정확한 깊이값을 얻을 수 있다. 양쪽 영상에 대한 깊이값 탐색 과정에서 얻을 수 있는 가장 큰 장점은 한쪽 영상만을 사용할 때 발생하는 폐색영역(occlusion)에 관한 문제점을 해결할 수 있다는 것이다.

3.3 세그먼트 기반의 신뢰확산 방법

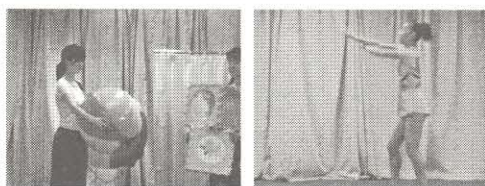
정제 과정에서 초기 깊이맵 탐색 과정에서 발생한 오차를 줄일 수 있다. 제안하는 방법은 현재 가장 높은 성능을 보이고 있는 신뢰확산 방법[6]을 사용하여 초기 깊이맵을 정제한다. 기존의 화소 단위의 신뢰 확산 방법이 상하좌우의 인접 화소들에게 메시지를 보내어 주변 화소들의 깊이값을 고려하는 방법에 착안하여, 제안하는 방법에서는 인접 세그먼트들에게 메시지를 보내는 세그먼트 단위의 신뢰확산 방법을 이용하도록 한다. 그림 4에 화소 단위 및 세그먼트 단위의 신뢰 확산 방법을 도식화하였다. 그림에서 알 수 있듯이, 현재 세그먼트에서 주변 세그먼트로 메시지를 보낼 때 (굵은 화살표), 주변 세그먼트에서 현재 세그먼트로의 메시지(점선 화살표)들을 이용한다.



(a) 화소 단위 (b) 세그먼트 단위
(그림 4) 신뢰 확산 방법

4. 실험 결과 및 분석

본 논문에서 제안하는 깊이맵 추정 기법의 성능을 평가하기 위해, 두 개의 테스트 영상을 사용하였다. 그림 5는 테스트 영상을 나타낸다.

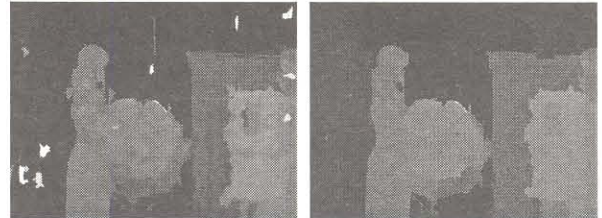


(a) Akko&Kayo (b) Rena
(그림 5) 테스트 영상

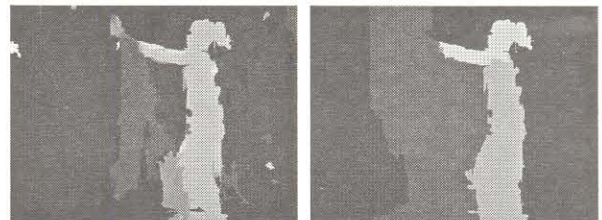
본 실험에서의 영역분할 알고리즘은 Felzenszwalb의 공개 소스[7]를 이용했다. 색상 차이의 임계값은 0으로 하였고, 각 세그먼트의 최소 크기는 60으로 설정했다. 이렇게 하여, 한 세그먼트에서 두 객체의 경계를 포함하지 않는 영역분할 기법을 효과적으로 적용할 수 있다.

깊이값 탐색 과정에서 가중치 w 를 0.5로 설정하였을

때 가장 오차가 작은 초기 깊이맵을 얻을 수 있었다. 또한, 3차원 워핑을 이용할 때, NTT에서 제공하는 투영 테스트 프로그램을 이용했다. 제안하는 방법을 이용하여 얻은 최종 깊이맵을 그림 6, 그림 7에 나타내었다.



(a) 초기 깊이맵 (b) 정제된 최종 깊이맵
(그림 6) "Akko&Kayo" 시점 28에 대한 실험 결과



(a) 초기 깊이맵 (b) 정제된 최종 깊이맵
(그림 7) "Rena" 시점 28에 대한 실험 결과

위의 결과에서 알 수 있듯이, 제안하는 방법을 사용하여 얻은 깊이맵은 기존의 깊이맵 추정 기법으로 획득한 깊이맵에 비해 객체의 경계가 비교적 명확함을 알 수 있었다. 또한, 초기 깊이맵과 비교하였을 때, 정제 과정을 통해서 깊이맵의 오차를 상당 부분 줄일 수 있었다.

5. 결론

본 논문에서는 영역 분할의 다시점 깊이맵 추정 기법을 제안하였다. 제안한 방법은 영역분할 기반의 깊이맵 추정 기법을 사용하였고, 다시점 영상에 대한 깊이맵 추정에 맞도록 기존의 스테레오 정합 알고리즘을 확장하였다. 제안하는 방법은 영역분할을 통해 얻어진 세그먼트 기반으로 3차원 워핑 기법을 적용하여 깊이값을 탐색하는 방법을 사용하였다. 또한, 영상 간 색상의 불일치에 강인한 자체적응 함수를 세그먼트 기반 탐색에 적합하도록 적용하였다. 깊이맵 정제 과정에서는 신뢰확산 방법을 이용하여 초기 깊이맵에서 발생한 오차들을 줄일 수 있었다. 결과적으로 제안하는 방법이 기존의 스테레오 정합 알고리즘을 확장하여 양질의 깊이맵을 추정할 수 있었다.

감사의 글

본 연구는 광주과학기술원(GIST) 실감방송연구센터(RBRC)를 통한 정보통신부대학 IT연구센터(ITRC)와 교육인적자원부 두뇌한국21(BK21) 정보기술사업단의 지원에 의한 것입니다.

참고문헌

- [1] ISO/IEC JTC1/SC29/WG11 M11259, "FTV (Free Viewpoint Television): Achievements and Challenge," Oct. 2004.
- [2] D. Scharstein, R. Szeliski, "A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms," International Journal of Computer Vision, vol. 47, no. 1, pp. 7-42, 2002.
- [3] A. Klaus, M. Sormann, K. Karner, "Segment-based Stereo Matching Using Belief Propagation and a Self-Adapting Dissimilarity Measure," International Conference of Pattern Recognition, vol. 3, pp. 15-18, 2006.
- [4] C. L. Zitnick, S. B. Kang, M. Uyttendaele, S. Winder, R. Szeliski, "High-quality Video View Interpolation Using a Layered Representation," SIGGRAPH, pp. 600-608, 2004.
- [5] P. F. Felzenszwalb, D. P. Huttenlocher, "Efficient Graph-Based Image Segmentation," International Journal of Computer Vision, vol. 59, no. 2, pp. 167-181, Sep. 2004.
- [6] P. F. Felzenszwalb, D. P. Huttenlocher, "Efficient Belief Propagation for Early Vision," International Journal of Computer Vision, vol. 70, no. 1, Oct. 2006.
- [7] Efficient Graph-Based Image Segmentation, <http://people.cs.uchicago.edu/~pff/segment/>, 2007.