

분산 값을 이용한 결합 양측 업샘플링

*이동우 **김만배

강원대학교 IT대학 컴퓨터정보통신공학과

*dw@kangwon.ac.kr, **manbae@kangwon.ac.kr

Joint Bilateral Upsampling using Variance

*Dong-Woo Lee **Manbae Kim

Dept. of Computer and Communications Eng., IT College, Kangwon National University

요약

최근 3D에 대한 관심이 집중되면서 고품질의 3D영상을 얻기 위해 고품질의 깊이 영상이 필요하고 이를 구현하기 위한 연구가 활발히 진행되고 있다. 깊이 영상을 얻기 위해서 Time-of-Flight(ToF)방식의 깊이 센서가 활용되고 있는데 이 깊이 센서는 실시간으로 깊이 정보를 획득할 수 있지만 낮은 해상도와 노이즈가 발생한다는 단점이 있다. 따라서 깊이 센서로 생성된 저해상도 깊이맵을 고해상도로 변환해야 한다. 주로 깊이 영상의 해상도를 높이기 위해서 Joint Bilateral Upsampling(JBU)방식이 사용되고 있다. 따라서 본 논문은 JBU 방식을 보강하여 블록단위로 분산에 따른 참조 영상의 가중치를 다르게 두어 깊이 영상의 품질을 향상시키는 방법을 제안한다.

1. 서론

최근 3차원 영상 산업이 각광받으면서 고품질의 3D 콘텐츠에 대한 요구가 증가하고 있다. 또한 어느 자리에서도 실감 있는 입체감을 느낄 수 있는 영상을 원하고 있다. 입체감을 나타내기 위해서는 깊이 정보가 필요하다. 물체의 깊이 정보를 획득하는 방법으로는 크게 수동적 방식과 능동적 방식으로 구분 할 수 있다. 수동적 방식의 대표적인 예로는 여러 카메라로부터 얻은 영상을 이용하는 스테레오 매칭 등이 있다. 이 방식은 텍스처가 없는 영역, 반복되는 패턴이 있는 영역, 가려진 영역 등에서 문제가 발생한다. 능동적 방식으로는 레이저를 이용하여 물체를 스캔하는 방식, ToF 센서 방식 등이 있다. 능동적 방식은 수동적 방식과는 반대로 텍스처가 있는 영역에서 정확도가 떨어진다. 그리고 ToF 방식의 깊이 센서를 통해서 얻은 깊이 영상은 해상도가 낮고, 반사가 심한 물체나 빛의 양이 많은 야외 등에서 노이즈가 발생한다는 단점이 있다. 하지만 깊이 영상을 실시간으로 획득할 수 있기 때문에 다양하게 활용이 되고 있다. 수동적 방식과 능동적 방식의 단점들은 보완하고 장점들은 수용하기 위해서 두 가지의 방식을 결합하여 깊이 영상의 품질 및 해상도를 개선하는 2D-plus-Depth구조에 대한 연구가 활발히 진행되고 있다[1].

따라서 본 논문에서는 ToF 방식의 깊이 센서를 통해 얻은 저해상도의 깊이 영상에 CCD카메라에서 얻은 고해상도의 컬러 영상을 참조하여 고해상도의 깊이 영상으로 보강하는 방법을 제안한다. 본 논문의 구성은 다음과 같다. 2절에서는 Joint Bilateral Upsampling(JBU)방식을 소개한다. JBU방식을 보강하여 블록단위로 분산에 따른 참조 영상의 가중치를 다르게 두어 깊이 영상의 품질을 향상시키는 방식을 3절

에서 제안하며, 4절에서 실험결과를 설명한다. 마지막으로 5절에서 결론을 맺는다.

2. 관련 연구

2.1 Bilateral Filter

Bilateral Filter는 에지의 정보 손실은 적게 하면서 노이즈를 줄이는데 사용되는 대표적인 영상 필터링 기법이다. 필터링 시 depth 등의 정보를 참조해 가중치를 적용하는 것으로 식 (1)와 같다.

$$J_p = \frac{1}{k_p} \sum_{q \in \Omega} I_q f(\|p - q\|) g(\|I_p - I_q\|) \quad (1)$$

식 (1)의 p는 중심 픽셀의 위치이고, q는 참조 픽셀의 위치이다. 두 픽셀의 거리를 유클리디안 거리를 사용해서 식 (2)와 같이 계산한다.

$$\|p - q\| = \sqrt{(p_x - q_x)^2 + (p_y - q_y)^2} \quad (2)$$

중심 픽셀과 인접 픽셀간의 거리가 가까울수록 가중치가 커지며, 거리가 멀수록 가중치가 작아진다. 가중치는 가우시안 분포를 이용하여 식 (3)와 같이 계산한다.

$$f(\|p - q\|) = \exp\left(-\frac{(\|p - q\|)^2}{2\sigma_f^2}\right) \quad (3)$$

2.2 Joint Bilateral Upsampling

Joint Bilateral Upsampling은 Bilateral Filter를 응용한 것으로 필터링 시 depth 등의 정보를 참조할 때, 저해상도의 깊이 영상이 아닌 원본 컬러 영상을 참조하는 방식이며 식 (4)처럼 나타낼 수 있다[2].

$$\bar{D}_p = \frac{1}{k_{p \downarrow}} \sum_{q \downarrow \in \Omega} D_{q \downarrow} f(\|p \downarrow - q \downarrow\|) g(\|\bar{I}_p - \bar{I}_q\|) \quad (4)$$

$$g(\|\bar{I}_p - \bar{I}_q\|) = \exp\left(-\frac{(\|\bar{I}_p - \bar{I}_q\|)^2}{2\sigma_g^2}\right) \quad (5)$$

다운샘플링 된 깊이 영상에 원본 컬러 영상의 픽셀 차이 값의 비중에 참조하여 깊이 영상을 업샘플링한다. 컬러 영상을 참조함으로써 기존 보간법보다 경계 영역이 더 잘 보존이 된다. 하지만 컬러 영상의 색상 차이가 뚜렷하지 않은 영역에 한해선 처리가 잘못 되어 부정확한 깊이 영상을 얻을 수 있다.

3. 제안 알고리즘

Joint Bilateral Upsampling에서 가중치를 계산할 때 가우시안 분포를 사용하게 되는데 동일한 값으로 구하여 색상 차이가 뚜렷하지 않은 영역에 한해 처리가 잘못 되는 점을 고려해 블록 단위로 분산 값을 구하여 σ_D 값을 결정한다.

분산을 구하는 식은 다음과 같다.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n X_i \quad (6)$$

$$var = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (7)$$

여기서, x 는 깊이영상의 깊이 값이며, n 은 블록의 크기이다. 식 (7)에서 구한 분산 값을 이용하여 가우시안 분포의 σ_D 값으로 사용한다. 분산 값이 작으면 σ_D 값을 증가, 아니면 σ_D 값을 감소시킨다. σ_D 값의 범위는 분산 값에 따라 다르게 설정한다.

σ_D 를 구하는 식은 다음과 같으며 여기서, m 은 σ_D 의 최대 크기를 임의로 지정해 준 것이다.

$$\sigma_D = m - \frac{(var_D - var_{D_{min}}) \times (m - 1)}{var_{D_{max}} - var_{D_{min}}} \quad (8)$$

σ_D 를 가우시안 함수에 적용하는 식은 다음과 같다.

$$g(\|\bar{I}_p - \bar{I}_q\|) = \exp\left(-\frac{(\|\bar{I}_p - \bar{I}_q\|)^2}{2\sigma_D^2}\right) \quad (9)$$

4. 실험 결과 및 분석

본 실험에 사용된 영상은 Middlebury Stereo 사이트[3]에서 제공하는 영상을 실험 깊이맵으로 채택하였다. 제안된 고해상도 깊이맵 영상 D_H 의 객관적 화질을 측정하기 위해 식 (10), (11)을 이용하여 PSNR를 측정한다. 원 깊이영상 D_L 를 다운 샘플링 하여 얻은 저해

상도 깊이영상 D_L 를 제안한 방법으로 확대된 깊이영상 D_H 를 얻는다.

$$MSE = \frac{1}{N} \sum_{i,j \in I} (D_H^O(i,j) - D_H(i,j))^2 \quad (10)$$

$$PSNR = 10 \cdot \log\left(\frac{255^2}{MSE}\right) \quad (11)$$

표 1은 Bilinear, JBU 그리고 제안한 방법으로 각각 업샘플링한 결과와 원 깊이영상과의 PSNR 비교 결과를 보여준다. 표 1에서 보듯이 Bilinear, JBU 그리고 제안한 방법의 평균 PSNR은 각각 29.83 dB, 29.92 dB, 30.13 dB였다. 따라서 제안하는 방법의 평균 PSNR은 Bilinear, JBU보다 0.3 dB, 0.21 dB 높았다.

	Bilinear	JBU	Proposed
Aloe	27.84	27.72	28.02
Baby	33.38	33.37	33.54
Bowling	29.72	29.79	30.15
Flowerpots	25.85	25.96	26.28
Midd	31.82	32.16	32.16
Monopoly	30.37	30.49	30.65
평균	29.83	29.92	30.13

표 1. 각 알고리즘 별 PSNR

그림 1의 (a),(b)는 고정된 σ_g 값에 따른 Joint Bilateral Upsampling의 결과영상이고 (c)는 제안한 방법으로 σ_D 값을 적용하여 Upsampling한 결과영상이다. 그림 1(b)에서 보듯이, σ_g 값이 크면 원본 컬러영상의 픽셀을 좀 더 많이 참조하게 되면서 시각적 왜곡이 발생할 수 있다.



그림 1. (a) JBU (low σ) (b) JBU (high σ) (c) Proposed

5. 결론

본 논문은 Variance를 이용한 고해상도 깊이맵 생성을 제안하였다. 기존의 JBU에서 동일한 σ_g 값으로 고해상도 깊이맵을 생성하였을 때보다 Variance를 이용하여 각각 다른 σ_g 값을 적용하였을 때 좋은 결과 값을 얻을 수 있었다. 시각적으로 경계 영역이 선명해지는 것을 확인할 수 있었다. 제안한 방법 역시 시각적 왜곡의 문제점을 앓고 있고, 이와 같은 원인은 원본 컬러 영상의 무늬, 재질 등과 같은 요소에 영향을 크게 받아 경계 영역이 제대로 검출되지 못 하는 문제점으로 이를 보완하기 위한 방안들을 계속 연구할 계획이다.

감사의 글

본 논문은 지식경제부 및 산업기술평가관리원의 산업융합 원천기술개발사업(정보통신)의 지원을 받았음 [KI002058, 대화형 디지털 홀로그래프 통합서비스 시스템의 구현을 위한 신호처리요소기술 및 하드웨어 IP 개발]

참고문헌

- [1] J. Zhu, L. Wang, R. Yang, J. Davis, "Fusion of Time-of-Flight Depth and Stereo for High Accuracy Depth Maps," IEEE Proc. CVPR, 2008.
- [2] J. Kopf, MF. Cohen, D. Lischinski and M. Uyttendaele, "Joint bilateral upsampling," ACM SIGGRAPH, 2007.
- [3] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *Int. J. Comput. Vision*, 47(1-3), pp. 7-42, 2002.