

컬러 영상 화소 분류를 이용한 깊이 영상의 홀을 채우는 기법

이건원, *한종기

세종대학교, *세종대학교

geonwonlee@sju.ac.kr, hjk@sejong.edu

Hole Filling Technique for Depth Map using Color Image Pixel Clustering

Geon-Won Lee, *Jong-Ki Han

Sejong University *Sejong University

요 약

실감미디어의 수요가 높아짐에 따라, 실감 미디어 콘텐츠 제작에 반드시 필요한 깊이영상에 대한 중요성이 커지고 있다. 다시점 영상으로부터 계산된 깊이 영상은 물체 주위와 배경 영역에 홀을 가지고 있다. 이러한 깊이영상에서의 홀을 채울 때, 이에 대응하는 컬러영상의 색상 특성을 고려하는 방법을 제안한다. 본 논문에서는 컬러 영상의 화소들을 색상 유사성을 이용하여 클래스로 분류하고, 홀의 깊이정보를 예측할 때 같은 클래스의 유효한 깊이값 만을 사용하는 방법을 소개한다. 제안하는 방법을 사용하면 깊이영상의 홀을 효율적으로 채워 넣을 수 있다. 실감미디어 제작에 있어 제안하는 방법을 사용한다면, 사실감 있는 깊이 정보를 얻을 수 있다.

1. 서론¹

5G 네트워크 통신이 보급화 되고, 네트워크 통신의 대역폭이 늘어남에 따라, 많은 데이터를 사용하는 실감미디어의 수요 또한 늘어나고 있다. 실감 미디어에서 몰입감을 주는 요소로는 시점의 자유도, 깊이 정보 등이 있다. 여기서 깊이 정보는 영상 내 사물이 시점으로부터 얼마나 멀리 위치해 있는지를 알려준다. 이러한 깊이정보는 물리적인 센서 (Kinect, LiDAR 등), 혹은 다시점 촬영 환경에서의 서로 연관이 있는 시점의 영상으로부터 깊이 영상으로 획득된다[1][2]. 하지만 위 방법으로 얻어진 깊이영상에는 센서의 부정확함, 다시점 카메라 간의 사각지대의 발생으로 인한 홀(Hole)이 존재한다.

깊이 영상의 홀을 제거하기 위하여 많은 선행연구들이 진행돼 왔으나 한계를 가지고 있다. 홀이 발생하는 사각지대는

전경이 아닌 배경이라는 특성을 이용하는 방법[3], 홀 주변의 화소들의 유사성을 이용하는 방법[4] 등이 제안됐다. 위 두 방법은 홀 영역 내 다른 물체를 고려하지 못하기 때문에 예측된 깊이정보가 부정확하다.

본 논문에서는 깊이영상의 홀을 효율적으로 채워넣기 위하여, 대응하는 컬러 영상의 화소 분류를 이용 깊이영상 내 홀을 채우는 방법을 제시한다. 깊이영상에 대응하는 컬러영상을 유사한 특성을 가지는 클래스들로 분류 후, 홀을 채울 시 홀 위치와 동일한 클래스를 가진 깊이 영상내 유효한 화소 만을 이용하는 방법을 제안한다.

본 논문의 구성은 다음과 같다. 2 절에서는 제안하는 홀 채우기 방법에 대해 서술하고, 3 절에서는 제안하는 방법의 효과를 실험을 통해서 검증한다. 마지막으로 4 절에서는 본 논문에 대한 결론을 맺는다.

¹ 연락저자: 한종기

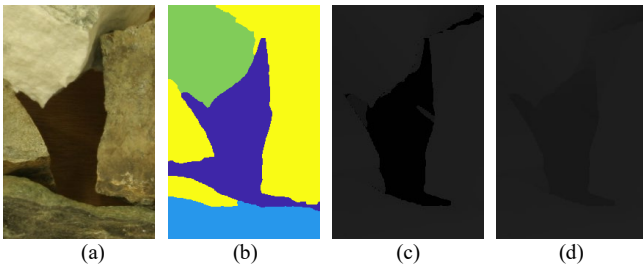


그림 1. 하나의 패치 영역에 대한 제안하는 알고리즘의 예시. (a) 컬러영상 패치, (b) 클래스 단위로 분할된 패치, (c) 깊이영상 패치, (d) 제안하는 방법으로 홀을 채운 패치.

2. 제안하는 홀 채우기 방법

2 절에서는 깊이 영상의 홀을 이에 대응하는 컬러 영상의 화소 분류를 이용하여 채우는 방법을 제안한다. 제안하는 방법은 크게 두가지 단계로 구성된다. 첫 번째 단계는, 홀 영역에 대한 직사각형 모양의 컬러영상, 깊이영상 패치를 획득하고, 컬러영상 패치에 대해 화소 분류를 수행하여 클래스화 하는 것이다. 그리고 두 번째 단계는 각 화소의 클래스 정보를 사용하여, 깊이 영상 내 홀과 동일한 클래스를 가지는 유효한 화소 값 들을 이용하여 홀을 채우는 것이다.

첫 번째로 영상의 홀 영역을 포함하는 패치영역을 구성하고, 이 패치영역에 해당하는 컬러영상 패치, 깊이영상 패치를 획득한다. 획득한 컬러 영상 패치를 k-mean 알고리즘[6]을 사용하여 유사한 색상정보를 가지는 영역별로 분류 후 클래스화 한다. 각 클래스는 폐색된 고유한 영역 하나를 일컫는다.

그 다음, 얻은 클래스 정보를 사용하여 깊이영상 패치내 홀을 예측한 값으로 채운다. 하나의 홀 화소에 대한 예측 값은, 홀 화소 위치 p 에 대해서 같은 클래스를 가지며, 유효한 (깊이정보를 가지는) 화소 집합 $S(p)$ 내 화소 q 들을 가중치 합하여 얻는다.

$$D(p) = \sum_{q \in S(p)} w_{(p,q)} D(q) \quad (1)$$

식 (1)은 깊이 영상 패치 내 홀 위치 p 에서의 새로이 예측된 깊이정보를 유도하는 방법을 보여준다. D 는 깊이영상 패치를 일컫는다. $D(p)$ 는 p 위치의 새로이 예측된 깊이 값이고, 각 화소 q 들은 각각의 가중치 w 를 가지며, $w_{(p,q)}$ 는 p 위치 홀의 깊이 값을 예측할 때 사용하는 q 위치의 유효한 깊이 값에 대한 가중치이다. 이는 위치 p 과 q 사이의 거리 가까울수록 더 큰 값을 가진다.

그림 1 을 보면, 제안하는 알고리즘을 사용하였을 때, 컬러영상 패치 영역 (a)는 (b)와 같이 구분된 클래스 영역으로 나뉘지고, 깊이영상 패치 (c)의 홀 영역은, (b)의 클래스 정보를

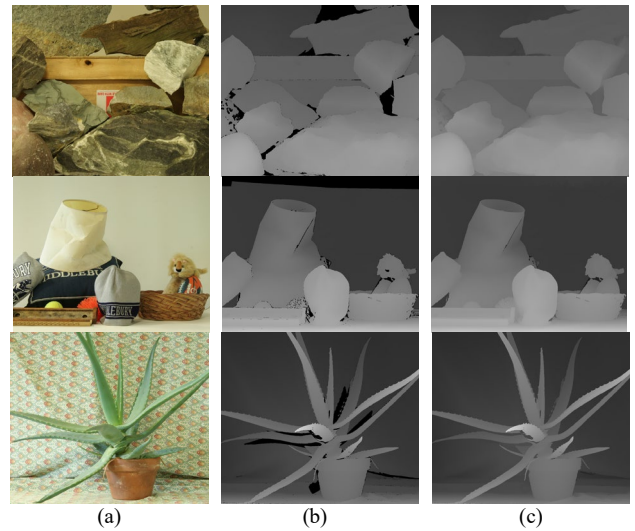


그림 2. 제안하는 방법에 대한 실험 결과. (a) 컬러영상, (b) 홀이 존재하는 깊이영상, (c) 제안하는 알고리즘을 사용한 홀 채우기 결과.

사용하여 동일한 클래스(색상)이며 유효한 깊이정보를 사용하여 (d)와 같이 채워진다.

3. 실험결과

본 논문에서 제안하는 알고리즘은 Matlab 2019b 환경에서 구현하였고, 실험에 사용한 데이터 세트는 Middlebury Stereo Datasets[5]에서 3 개의 세트를 선택하여 사용하였다. [5]의 데이터 세트는 스테레오 컬러영상과 이에 대응하는 깊이영상으로 구성되었다.

그림 2 에서 제안하는 알고리즘에 대한 실험 결과를 볼 수 있다. 제안하는 알고리즘은 다양하게 발생한 홀 영역을 적절히 채우고 있음을 볼 수 있다. 그림 2 의 두번째 데이터 세트의 경우, 전경인 물체 뒤에 발생한 홀 영역의 경우, 배경의 화소 위치들이 같은 클래스로 분류되기 때문에, 전경 물체의 깊이정보를 왜곡하지 않고 배경에 대한 깊이정보가 적절히 채워지고 있음을 볼 수 있다. 그리고 그림 2 의 첫번째 데이터 세트의 경우, 복잡한 사물내 발생한 깊이영상의 홀 영역의 경우도, 적절하게 채워지고 있음을 볼 수 있다. 따라서 제안하는 방법으로 깊이영상의 홀을 채웠을 때 효과가 있음을 알 수 있다.

4. 결론

본 논문에서, 컬러영상의 화소 분류 정보를 사용하여 이에 대응하는 깊이영상의 홀을 채우는 방법을 제안했다. 깊이 영상의

홀을 채우기 전에, 홀 영역을 포함하는 컬러영상의 패치를 획득 후 패치 내 화소를 유사한 색상 값을 가지는 클래스로 분류하였다. 이후 깊이영상에서 발생하는 홀이 패치 내 자신과 같은 클래스를 가지는 유효한 깊이 화소 만을 가지고 채워지도록 알고리즘을 설계하였다. 제안하는 알고리즘을 사용하면 컬러영상의 색상 특성을 고려하여 적절한 예측 깊이 정보를 계산할 수 있다. 제안하는 알고리즘을 사용함으로써 실감미디어 콘텐츠의 품질 향상을 기대한다.

distance measure," in IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 34, no. 4, pp. 1907-1916, Aug. 2004

ACKNOWLEDGMENT

This work was partly supported by the National Research Foundation of Korea (NRF) under Grant NRF-2018R1A2A2A05023117 and partly by Institute for Information & communications Technology Promotion (IITP) under Grant 2017-0-00486 funded by the Korea government (MSIT).

참고문헌

- [1] Y. W. Kuan, N. O. Ee and L. S. Wei, "Comparative Study of Intel R200, Kinect v2, and Primesense RGB-D Sensors Performance Outdoors," in IEEE Sensors Journal, vol. 19, no. 19, pp. 8741-8750, 1 Oct.1, 2019
- [2] M. Wien, J. M. Boyce, T. Stockhammer and W. Peng, "Standardization Status of Immersive Video Coding," in IEEE Journal on Emerging and Selected Topics in Circuits and Systems, vol. 9, no. 1, pp. 5-17, March 2019
- [3] P. Ndjiki-Nya et al., "Depth Image-Based Rendering With Advanced Texture Synthesis for 3-D Video," in IEEE Transactions on Multimedia, vol. 13, no. 3, pp. 453-465, June. 2011.
- [4] Z. Wang, J. Hu, S. Wang and T. Lu, "Trilateral constrained sparse representation for Kinect depth hole filling", Pattern Recognition Letters, vol. 65, pp. 95-102, 2015.
- [5] D. Scharstein and C. Pal, "Learning Conditional Random Fields for Stereo," 2007 IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, MN, pp. 1-8, 2007.
- [6] S. Chen and D. Zhang, "Robust image segmentation using FCM with spatial constraints based on new kernel-induced