Journal of the Korea Industrial Information Systems Research Vol. 24 No.5, Oct. 2019 : 1-8
http://dx.doi.org./10.9723/jksiis.2019.24.5.001 ISSN:1229-3741

표면 모델링을 통한 깊이 영상 내 측정 실패 화소 보정 방법+

(Correction Method for Measurement Failure Pixels in Depth Picture using Surface Modeling)

이 동 석[1], 권 순 각[2]*
(DongSeok Lee and SoonKak Kwon)

요 약 본 논문에서는 깊이 영상에서 깊이 카메라의 일시적인 오류로 인해 측정이 되지 않은 깊이 화소의 값을 보정하는 방법을 제안한다. 깊이 영상의 한 블록 내에서 측정이 정상적으로 된 화소의 좌표와 깊이 값을 이용하여, 해당 블록의 깊이 값과 제일 오차가 적은 평면과 곡면을 모델링한다. 그 후 각각의 모델링된 표면을 통해 추정된 깊이 값과 원래 측정된 깊이 값을 비교하여 오차를 계산한다. 그 후 오차가 제일 작게 계산된 표면을 선택한 후, 측정에 실패한 깊이 화소를 선택된 모델링 표면을 통해 깊이 값을 추정함으로써 보정한다. 모의실험 결과 제안된 방법을 통한 보정방법은 5×5 영역의 중간 값을 이용한 보정방법에 비해 보정 성능이 평균 20% 개선되었음을 확인하였다.

핵심주제어: 깊이 카메라, 깊이 영상 보정, 표면 모델링

Abstract In this paper, we propose a correcting method of depth pixels which are failed to measure since temporary camera error. A block is modeled to plane and sphere surfaces through measured depth pixels in the block. Depth values in the block are estimated through each modeled surface and a error for the modeled surface is calculated by comparing the original and estimated pixels, then the surface which has the least error is selected. The pixels which are failed to measure are corrected by estimating depth values through selected surface. Simulation results show that the proposed method increases the correction accuracy by an average of 20% compared with the correction method of 5×5 median method.

Keywords: Depth camera, Depth picture correction, Surface modeling

1. 서 론

색상 대신 카메라와의 거리를 측정하여 영상화하는 깊이 카메라를 통해 깊이 영상을 촬영함으로써 기존 색상영상에서는 얻지 못하던 객체와의 거리 정보를 획득할 수 있다. 객체 거리 단독으로 또는 기존 색상 정보와 결합하여 객체의 새로운 정보를 획득함으로써 새로운 영상 응용이 가능해진다. 이는 깊이 영상을 통한 실내의 3D 지도 작성 (Pomerleau et al., 2011; Sturm, et al., 2012), 터치 인터페이스 구현 (Wilson, 2010; Lee and Kwon, 2019) 등만 아니라 얼굴 인식 (Fanelli et al., 2013), 객체 추적 (Siddiqui and

* Corresponding Author: skkwon@deu.ac.kr
+ 이 논문은 2019년도 BB21+사업에 의하여 지원되었음. 또한, 2019학년도 동의대학교 교내연구비에 의해 연구되었음(201902270001).
Manuscript received July 28, 2019 / revised August 13, 2019 / accepted August 16, 2019
1) 동의대학교 컴퓨터소프트웨어공학과, 제1저자
2) 동의대학교 컴퓨터소프트웨어공학과, 교신저자

Medioni, 2010) 등에 활용되고 있다. 또한 최근 ToF 방식의 카메라가 스마트 폰에 탑재되는 등 깊이 영상의 활용은 앞으로도 더욱 확대될 것으로 예측된다.

깊이 카메라를 통해 깊이 영상을 촬영할 때 깊이 카메라의 측정 성능 한계로 인해 카메라와의 거리를 측정하지 못하는 화소가 발생할 수 있다. 이러한 화소는 깊이 영상을 통해 객체의 정보를 얻고자 하는 활용에 있어서 정확도를 떨어뜨리는 요인이 되기 때문에 원래 값과 비슷한 값으로 보정해줄 필요가 있다. Silberman et al.(2012)는 Levin et al.(2004)가 제안한 흑백 영상의 색상화 알고리즘을 응용하여 깊이가 측정되지 않은 화소를 채워넣는 방법을 적용하였다. 하지만 이 방법은 보정을 위해 색상 영상이 필요하다는 단점이 있다. Miao et al.(2012)는 깊이 영상 내 깊이 측정이 되지 않은 화소를 포함한 잡음을 경계선 검출 기반의 필터와 라플라스 연산자를 이용하여 보정하는 방법을 제안하였다. Xue et al.(2017)은 화소를 보정하기 위해 색상을 재구성하는데 이용되는 저랭크 정규화 (Low rank regularization)를 깊이 영상에 맞게 변형한 낮은 경사도 정규화 (Low gradient regularization) 방법을 제안하였다. 하지만 이러한 방법들은 색상 영상의 재구성 알고리즘을 응용하여 깊이 영상 내 주변 화소들 간의 연속성을 이용하여 화소를 보정하는 방법으로써, 각 객체가 구성하는 표면의 특징은 이용하지 않는다는 한계가 있다.

본 논문에서는 깊이 영상을 평면 (Lee and Kwon, 2018) 또는 구면 (Kwon and Lee, 2016)으로 모델링하여 측정에 실패한 화소를 보정하는 방법을 제안한다. 깊이 영상은 표면을 표현하는 영상으로 볼 수 있다. 이 때, 표면을 작은 영역으로 나눈다면, 한 영역내의 표면은 평면 또는 구면 등의 일부로 나타낼 수 있다. 이러한 점을 이용하여 깊이 영상 내 측정이 되지 않은 화소에 대해 주변 깊이 화소를 이용하여 평면 또는 구면을 모델링하고, 이를 통해 해당 화소를 보정하는 방법이다. 먼저 깊이 영상을 정사각형 블록으로 분할한 후, 블록 내 측정이 된 화소를 이용하여 해당 구역의 평면이나 구면을 모델링한다. 다음으로 모델링된 화소와 원래 화소의 오차를 비교하여 오차가 적은 표면을 선택한다. 그 후 모델링된 표면을 이용하여 깊이가 측정되지 않은 화소의 깊이 값을 보정한다. 본 논문에서는 제안된 방법과 5×5 영역의 중간 값을 이용한 보정 방법과 비교하여 정확도를 비교한다. 제안된 방법을 통해 객체의 표면 특징을 이용하여 깊이 측정에 실패한 화소를 정확하게 보정할 수 있다.

본 논문의 구성은 다음과 같다. 2장에서는 깊이 영상 내 깊이 화소 값과 좌표를 이용하여 평면과 구면을 모델링 하는 방법을 제안한다. 3장에서는 깊이 측정에 실패한 화소가 포함된 깊이 영상을 제안된 방법을 통하여 보정함으로써 기존 방법과 정확도를 비교한다. 4장에서는 본 논문의 결론을 내린다.

2. 표면 모델링을 통한 깊이 영상 내 측정 실패 화소 보정 방법

깊이 영상을 촬영할 때, 깊이 카메라가 거리 측정에 실패하여 화소 값이 없는 잡음이 깊이 영상에 포함된다. 이러한 깊이 측정에 실패하는 원인으로는 깊이 카메라의 일시적인 오류로 인한 측정 실패와 객체들 간의 서로 가리어짐 현상으로 인한 원인으로 분류할 수 있다. Fig. 1은 이러한 깊이 측정이 실패한 화소가 포함된 깊이 영상의 예이다.

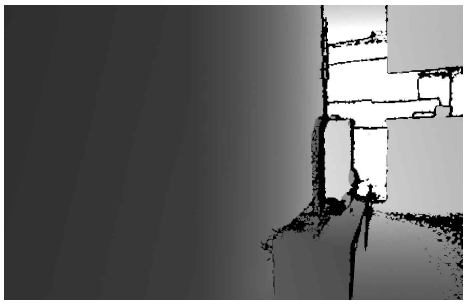


Fig. 1 Example of Measurement Failure Pixels in Depth Picture

본 논문에서는 측정에 실패한 화소의 깊이 값을 블록단위로 보정하는 방법을 제안한다. 먼저 블록 내 측정에 성공한 화소를 이용하여 해당 블록의

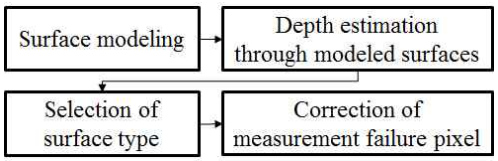


Fig. 2 Flowchart of Proposed Method

표면을 평면과 구면으로 모델링한다. 그 후 모델링된 표면들과 측정된 깊이 화소들을 비교하여 오차가 제일 작은 표면을 선택한 후, 선택된 표면을 통해 추정된 깊이 값으로 측정에 실패한 화소를 보정한다. 제안하는 방법의 흐름도는 Fig. 2와 같다.

2.1 깊이 화소의 3차원 좌표 변환

카메라는 실제 세계의 3차원 좌표를 2차원으로 투영하는 일종의 좌표 변환계로 볼 수 있다. 핀홀 카메라 모델은 이러한 관점에서 영상이 원점으로 투영되는 모델이다. 핀홀 카메라 모델에서는 Z축이 카메라 촬영 방향으로 정의된 카메라 좌표계를 적용한다. 이 때 영상이 투영되는 평면은 원점으로부터 Z축으로 초점거리 f만큼 떨어진 XY평면에 위치한다. Fig. 3은 핀 홀 카메라 모델의 개념도를 나타낸 것이다.

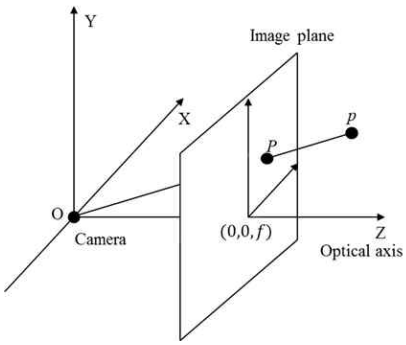


Fig. 3 Picture Projection from Pinhole Camera Model

Fig. 3에서 카메라 좌표계의 한 점 $p \equiv (x,y,z)$는 영상 면 위의 한 점 $P \equiv (u,v)$로 투영된다. 이 때 두 좌표간의 관계는 식 (1)으로 표현된다.

$$f : u = z : x \rightarrow u = f\frac{x}{z} \qquad (1)$$
$$f : v = z : y \rightarrow v = f\frac{y}{z}$$

여기서, p의 Z축 좌표인 z는 해당 점이 투영된 영상 좌표계에서 화소 P의 깊이 값 d로 표현된다. 이를 이용하여 카메라 좌표계의 한 점 (x,y,z)과 (u,v)과의 관계는 식 (2)로 표현할 수 있다.

$$\begin{aligned} x &= du/f \\ y &= dv/f \\ z &= d \end{aligned} \qquad (2)$$

2.2 깊이 정보를 이용한 블록 단위 평면 모델링

평면을 모델링하기 위해 깊이 영상을 $N \times N$ 크기의 블록들로 분할한다. 그 후 각 블록 내 깊이 값을 가지는 화소 각각의 좌표를 식 (2)을 통해 3차원 좌표로 변환한다. 그 후 표면을 모델링하기 위해 평면의 3차원 좌표계에서의 특성을 이용한다. 3차원 좌표계 상에서의 평면은 식 (3)으로 표현할 수 있다. 식 (3)에서 a_p, b_p, c_p는 평면의 인자이다.

$$a_p x + b_p y - z + c_p = 0 \qquad (3)$$

블록 내 화소들의 3차원 좌표 $p_i \equiv (x_i, y_i, z_i)$ $(1 \le i \le N^2)$들을 식 (3)에 대입함으로써 식 (4)의 행렬식을 만들 수 있다. 식 (4)에서 $N \times 3$ 행렬 **A**와 $N \times 1$ 행렬 **B**는 점들의 좌표와 관련된 행렬이고, 3×1 행렬 **R**은 평면의 인자에 관련된 행렬이다.

$$\mathrm{AR} = \mathrm{B}$$
$$\mathrm{A} = \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ & \cdots & \end{bmatrix} \quad \mathrm{B} = \begin{bmatrix} z_1 \\ z_2 \\ \cdots \end{bmatrix} \quad \mathrm{R} = \begin{bmatrix} a_p \\ b_p \\ c_p \end{bmatrix} \qquad (4)$$

이 때 주어진 점들 각각의 거리의 제곱의 합으

로 표현되는 오차가 제일 작은 평면을 찾기 위해 식 (5)로 구할 수 있는 **A**의 의사역행렬 $\mathbf{A}^{+}$를 이용한다. 그 후 식 (6)을 계산하여 오차가 제일 작은 평면의 인자를 구할 수 있다. 식 (6)은 **A**가 세 점을 가지는 정방행렬일 때, $\mathbf{A}^{-1}$를 통해 세 점을 포함하는 평면의 인자 **R**을 구하는 방법과 유사하다. 식 (6)을 통해 주어진 점들과의 오차가 제일 작은 평면의 인자를 구함으로써 평면을 모델링한다.

$$\mathrm{A}^{+} = \left(\mathrm{A}^{\mathrm{T}}\mathrm{A}\right)^{-1}\mathrm{A}^{\mathrm{T}} \tag{5}$$

$$\mathrm{R} = \mathrm{B}\mathrm{A}^{+} \tag{6}$$

2.3 깊이 정보를 이용한 블록 단위 구면 모델링

3차원 좌표계 상에서의 구면은 식 (7)로 표현할 수 있다.

$$(x-a_s)^2 + (y-b_s)^2 + (z-c_s)^2 - r^2 = 0 \tag{7}$$

여기서, a_s, b_s, c_s, r_s 는 구면의 인자이다. 그 후 블록 내의 점 p_i들을 식 (7)에 대입하면 식 (8)의 행렬식이 도출된다.

$$\mathrm{AR} = \mathrm{B}$$
$$\mathrm{A} = \begin{bmatrix} (x_1)^2 + (y_1)^2 & x_1 & y_1 & z_1 & 1 \\ (x_2)^2 + (y_2)^2 & x_2 & y_2 & z_2 & 1 \\ & & \cdots & & \end{bmatrix},$$
$$\mathrm{B} = \begin{bmatrix} {z_1}^2 \\ {z_2}^2 \\ \cdots \end{bmatrix} \quad \mathrm{R} = \begin{bmatrix} -1 \\ 2a_s \\ 2b_s \\ 2c_s \\ -(a_s^2 + b_s^2 + c_s^2 - r_s^2) \end{bmatrix} \tag{8}$$

그 후 평면 모델링 방법과 유사하게 식 (6)을 통해 주어진 점들과의 오차가 제일 작은 인자로 구성된 행렬 **R**을 구함으로써 구면을 모델링한다.

2.4 모델링 표면을 통한 깊이 측정 실패 화소 보정

평면 및 곡면을 모델링한 후에 잡음 제거를 위한 표면을 선택하기 위하여 두 표면에 대한 오차를 계산한다. 이 때 오차는 각 모델링된 표면에서 추정된 깊이와 정상적으로 측정된 화소의 깊이 값의 제곱 합을 비교하여 오차가 적은 표면을 선택한다. 블록 내 화소 $P_i \equiv (u_i, v_i)$에서 모델링된 평면을 통해 추정되는 깊이 값 $\bar{d}_{p_i}$은 식 (9)를 통해 계산한다(Lee and Kwon, 2018).

$$\bar{d}_{p_i} = -c_p / \left(a_p u_i / f + b_p v_i / f - 1\right) \tag{9}$$

또한 P_i에서 모델링된 구면을 통해 추정되는 $\bar{d}_{s_i}$는 식 (10)로 표현되는 이차식의 해를 구함으로써 계산할 수 있다(Kwon et al., 2016).

$$\begin{aligned} &\left((u_i/f)^2 + (v_i/f)^2 - 1\right)(\bar{d}_{s_i})^2 - \\ &\left(2a_s u_i/f + 2b_s v_i/f + 2c\right)\bar{d}_{s_i} - \\ &\left({a_s}^2 + {b_s}^2 + {c_s}^2 - r^2\right) = 0 \end{aligned} \tag{10}$$

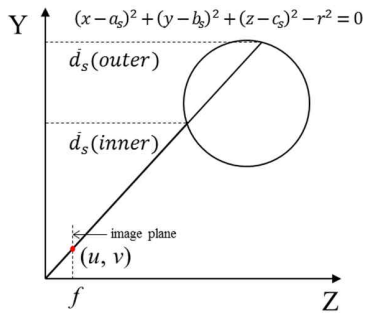


Fig. 4 Two Modes in Depth Pixel Estimation using Spherical Modeling

여기서 $\bar{d}_{s_i}$의 해는 두 개인데, 이는 모델링된 구면으로부터 추정될 수 있는 깊이 값이 Fig. 4처럼 2개가 될 수 있다는 것을 의미한다. 따라서 구면 모델링 모드는 안쪽 면 모드와 바깥쪽 면 모드로 구분하여 추정해야 한다.

P_i에서 실제로 측정된 깊이 d_i와 평면 및 구면 모델링을 통해 추정된 깊이 $\bar{d}_{p_i}$, $\bar{d}_{s_i}(inner)$, $\bar{d}_{s_i}(outer)$의 차이를 각각의 오차라고 했을 때,

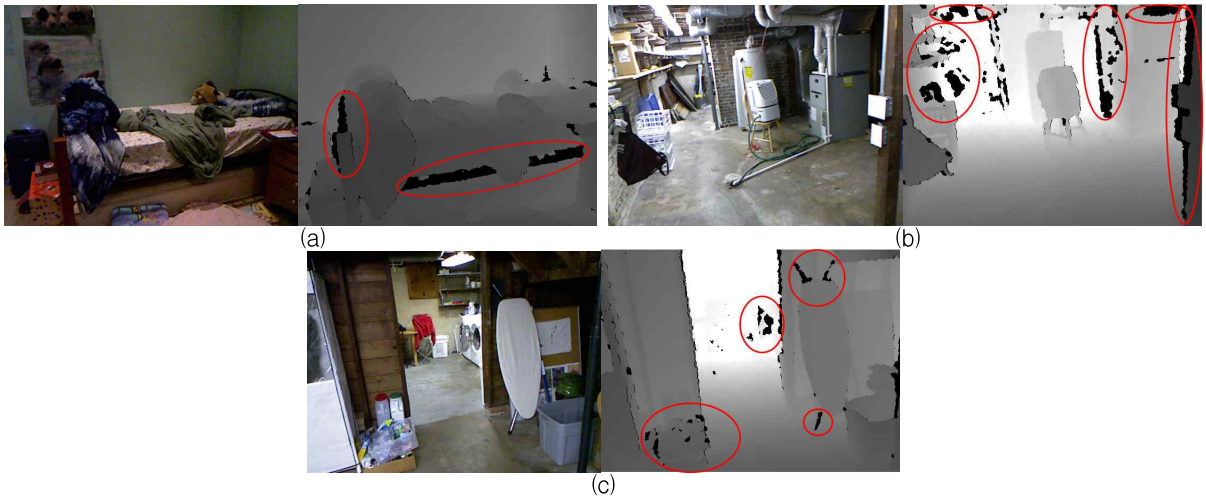


Fig. 5 Depth Pictures for Simulation. (a) Bedroom, (b) Basements, and (c) Laundry room

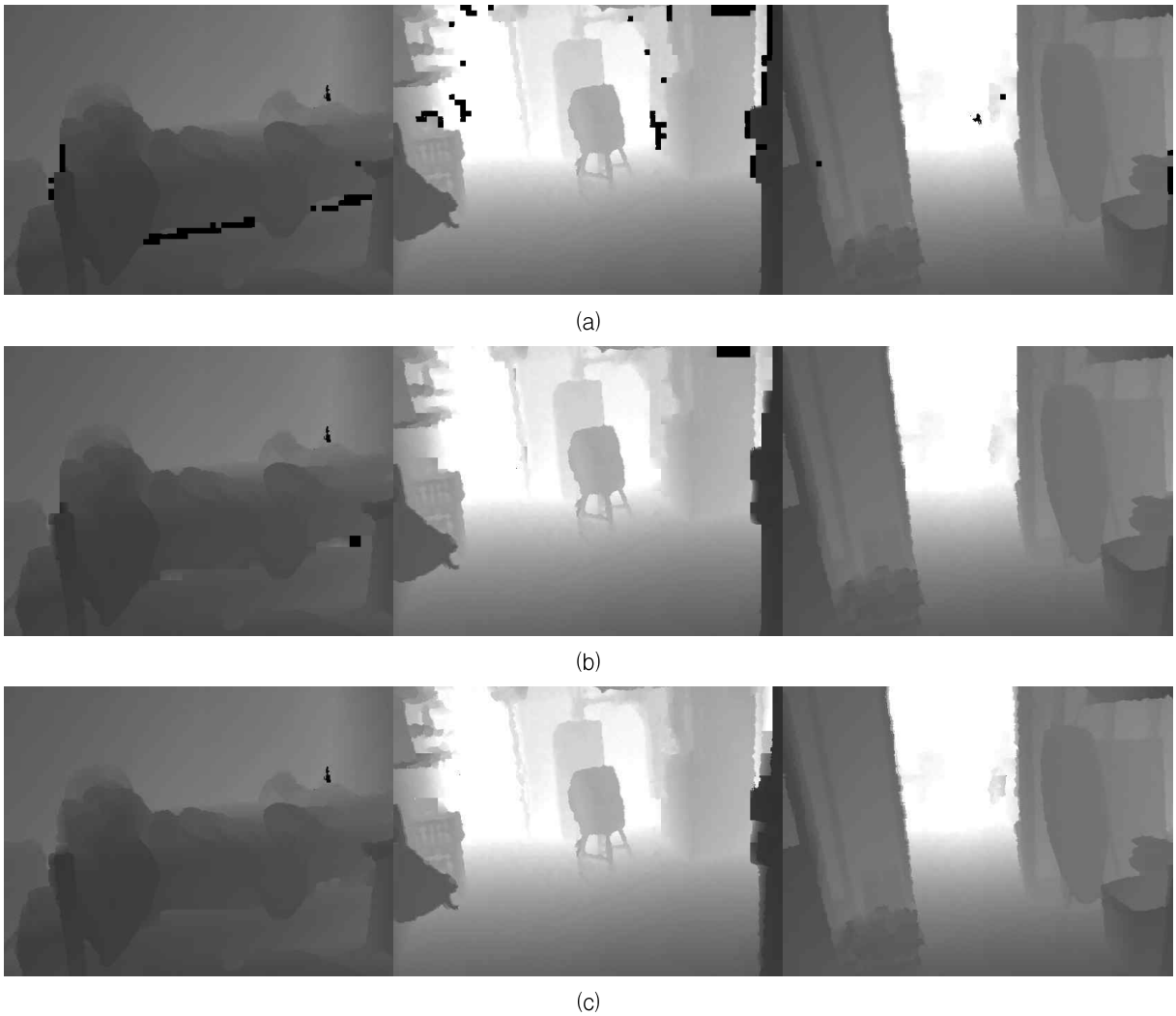


Fig. 6 Correction Result according to Block Size. (a) N=8, (b) N=16, and (c) N=32

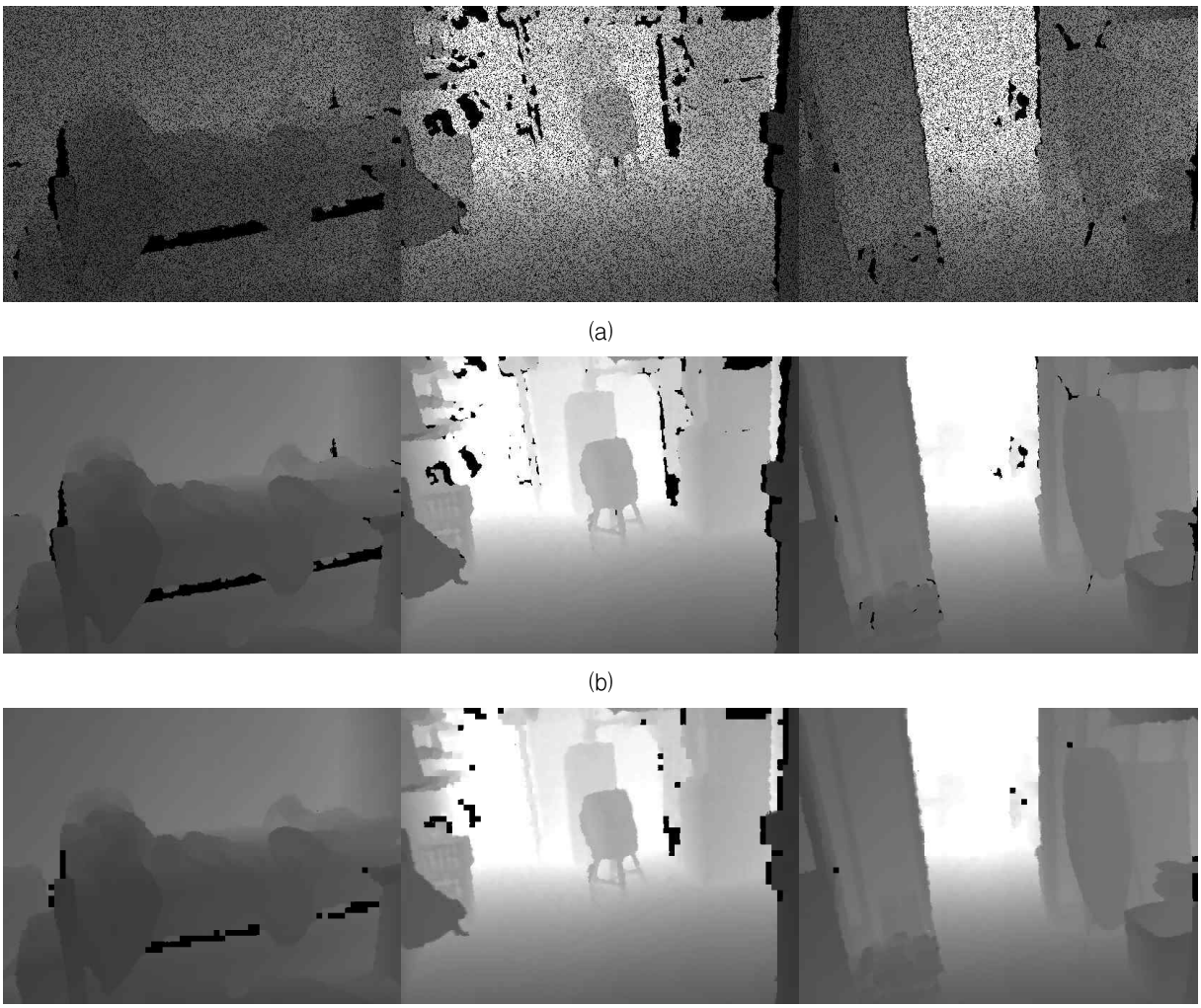
18	0	19	18	0
0	18	19	0	16
17	17	0	17	16
18	17	16	16	0
0	17	16	16	15

median value in 5×5 block = 17

Fig. 7 Correction Method through Median Value

Table 1 Comparison of Accuracies between Measurement Failure Pixels Correction Methods

Soruce picture	MSE		Improved accuracy
	5×5 median filtering method	Proposed method (8×8)	
bedroom	103.31	80.47	22.1%
basement	2186.08	1602.17	26.7%
laundry room	905.89	882.37	2.6%



(a)

(b)

(c)

Fig. 8 Correcting Measurement Failure Pixels. (a) Depth Picture including Generated Noises, (b) Correction Picture through Median Values, and (c) Correction Picture through Proposed Method

블록 내 깊이 화소들의 실제 측정 깊이와 추정된 깊이 값 간의 오차의 제곱 합이 제일 작게 계산되는 표면을 선택한다. 그 후 블록 내 깊이 측정에 실패한 화소에 대해 선택된 표면을 추정한 식 (9) 또는 식 (10)을 통해 해당 화소를 보정한다.

3. 실험결과

본 논문에서 제안한 방법을 실험하기 위해 실험영상으로 Fig. 5와 같이 측정에 실패한 화소가 포함된 깊이 영상 (Silberman et al., 2012)을 사용하였다. 이 영상은 Microsoft사의 Kinect로 촬영되었으며, 해상도는 640×480이며, 식 (2)에서의 f는 585.6이다. 이 때 영상 내 검정색으로 표시된 화소가 측정이 실패된 화소이며, 해당 화소가 밀집된 곳에 대해서 원으로 표시를 했다.

Fig. 6은 블록 크기 N을 8, 16, 32으로 변경하면서 실험한 결과이다. 블록의 크기가 작을 경우, 잡음 제거의 정확도가 높지만, 일부 영역의 경우 블록의 모든 화소가 깊이 값을 가지지 않는 경우가 발생하여 보정을 하지 못하는 경우가 발생하였다. 반대로 블록의 크기가 클수록 보정된 화소의 개수가 증가했지만 보정 정확도가 떨어지는 경우가 발생하였다. 또한 객체 내부에서는 비교적 정확하게 보정이 되었지만, 객체 경계부분에서는 보정의 정확도가 떨어졌다.

정상적으로 측정된 화소 중 임의의 화소의 깊이 값을 제거한 후, 제안된 방법을 통해 제거된 화소를 보정하여 원래의 깊이 값과 보정된 깊이 값을 비교함으로써 보정 정확도를 측정하였다. 그 후 제거된 화소의 위치에서 보정된 깊이 값과 원래의 깊이 값 간의 오차를 중앙값을 이용한 보정 방법과 정확도를 비교하였다. 중앙값을 이용한 보정 방법은 Fig. 7과 같이 보정하고자 하는 화소를 중앙으로 한 5×5 영역 내 깊이가 측정된 화소 값들의 중앙값을 이용하여 화소를 보정하는 방법이다.

Fig. 8은 각 보정 방법들을 적용하여 잡음이 추가된 깊이 영상을 보정한 결과이다. 이 때 중앙값을 이용한 방법과 비교했을 때, 좀 더 많은 화소가 보정이 된 것을 볼 수 있다. Table 1은 보정의 정확도를 측정한 결과이다. 보정 정확도를 5×5 중간값 방법과 비교하였다. 제안된 방법에서는 블록 크기가 작을수록 보정이 좀 더 정확하게 됨을 볼 수 있다. 5×5 중앙값을 이용하여 화소를 보정하는 방법과 비교했을 때 블록 크기가 8일 때, 보정 정확도가 평균 20%정도 개선됨을 볼 수 있다.

4. 결 론

본 논문에서는 깊이 영상에서 일시적으로 값이 측정되지 않는 화소를 표면 모델링을 통해 보정하는 방법을 제안하였다. 제안된 방법을 통해 블록 단위로 표면의 종류를 판단하고, 판단된 표면으로 모델링하여 측정에 실패한 깊이 화소를 보정하였다. 정확도 실험 결과 블록의 크기가 8일 때, 5×5 중간값 방법에 비해 보정 정확도가 평균 20%정도 개선이 되었지만, 보정이 되지 않는 화소가 어느 정도 존재한다는 단점이 있다. 이러한 단점을 보완하기 위해, 제안된 방법과 다른 알고리즘을 결합하여 이러한 문제점을 해결할 필요성이 있다. 제안된 방법을 통해 깊이 영상을 활용한 객체 인식 등에서 그 정확도를 향상시킬 수 있을 것으로 기대된다.

References

Fanelli, G., Dantone, M. and Van Gool, L. (2013). Real Time 3D Face Alignment with Random Forests-based Active Appearance Models, *Proceeding of the IEEE International Conference and Workshops on Automatic Face and Gesture Recognition*, Apr. 22-26, Shanghai, China, pp. 1-8.

Kwon, S. K., Lee, D. S. and Park, Y. H. (2016). Depth Video Coding Method for Spherical Object, *Journal of the Korea*

Industrial Information Systems Research, 21(6), 23-29. https://doi.org/10.9723/JKSIIS.2016.21.6.023

Lee, D. S. and Kwon, S. K. (2018). Intra Prediction of Depth Picture with Plane Modeling, *Symmetry*, 10(12), 1-16. https://doi.org/10.3390/sym10120715

Lee, D. S. and Kwon, S. K. (2019). Virtual Touch Sensor using a Depth Camera, *Sensor*, 19(4), 1-13. 2019. https://doi.org/10.3390/s19040885

Levin, A., Lischinski, D. and Weiss,Y. (2004). Colorization using Optimization, *Proceedings of ACM SIGGRAPH*, Aug. 8-12, Los Angeles, California, pp. 689-694.

Miao, D., Fu, J., Lu, Y., Li, S. and Chen, C. W. (2012). Texture-assisted Kinect Depth Inpainting, *Proceeding of the IEEE International Symposium on Circuits and Systems*, May 20-23, Seoul, Korea, pp.604-607.

Pomerleau, F., Magnenat, S., Colas, F., Liu, M. and Siegwart, R. (2011). Tracking A Depth Camera: Parameter Exploration for Fast ICP, *Proceeding of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Sep. 25-30, San Francisco, California, pp. 3824-3829.

Siddiqui, M. and Medioni, G. (2010). Human Pose Estimation from A Single View Point, Real-time Range Sensor, Proceeding of the IEEE Computer Society *Conference on Computer Vision and Pattern Recognition Workshops*, Jun. 13-18, San Francisco, California, pp. 1-8.

Silberman, N., Hoiem, D., Kohli, P. and Fergus. R. (2012). Indoor Segmentation and Support Inference from Rgbd Images, *Proceeding of the European Conference on Computer Vision*, Oct. 7-13, Florence, Italy, pp. 746-760.

Sturm, J., Engelhard, N., Endres, F., Burgard, W. and Cremers, D. (2012). A Benchmark for The Evaluation of RGB-D Slam Systems, *Proceeding of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct. 7-12, Vilamoura, Portugal, pp. 573-580.

Wilson, A. D. (2010). Using A Depth Camera as A Touch Sensor, *Proceeding of the ACM International Conference on Interactive Tabletops and Surfaces*, Nov. 7-10, Saarbrucken, Germany, pp. 69-72.

Xue, H., Zhang, S. and Cai, D. (2017). Depth Image Inpainting: Improving Low Rank Matrix Completion with Low Gradient Regularization, *IEEE Transactions on Image Processing*, 26(9), 4311-4320. https://doi.org/10.1109/TIP.2017.2718183

이 동 석 (DongSeok Lee)
- 정회원
- 동의대학교 컴퓨터소프트웨어공학과 공학석사
- 동의대학교 컴퓨터소프트웨어공학과 박사과정
- 관심분야 : 멀티미디어 신호처리, 영상 인식

권 순 각 (SoonKak Kwon)
- 정회원
- 경북대학교 전자공학과 공학사
- KAIST 전기및전자공학과 공학석사
- KAIST 전기및전자공학과 공학박사
- 동의대학교 컴퓨터소프트웨어공학과 교수
- 관심분야 : 멀티미디어처리, IoT