

평면 모델링을 통한 깊이 영상 부호화의 개선

(Improvement of Depth Video Coding by Plane Modeling)

이 동 석¹⁾, 권 순 각^{2)*}

(Dong-Seok Lee and Soon-Kak Kwon)

요 약 본 논문에서는 평면을 모델링하여 깊이 정보를 보정하고 부호화를 개선하는 방법을 제안한다. 먼저 보정하고자 하는 화소를 중심으로 수평, 수직 방향으로 최소 자승법을 이용하여 평면을 모델링한 후 예측 오차에 근거하여 예측된 평면이 적합한지 판단한다. 그 후 모델링된 평면상의 깊이 화소 값으로 보정한다. 제안된 방법을 통해 평면으로 이루어진 깊이 영상뿐만 아니라 다양한 깊이 정보를 가지는 깊이 영상에 대해서도 보정이 가능하다. 제안된 방법을 적용하여 부호화 성능을 나타내는 엔트로피 척도를 측정 한 결과, 부호화 성능이 최대 80.2% 개선된 것을 확인하였다.

핵심주제어 : 깊이 영상 보정, 잡음 보정, 깊이 부호화

Abstract In this paper, we propose a method of correcting depth image by the plane modeling and then improving the coding performance. We model a plane by using the least squares method to the horizontal and vertical directions including the target pixel, and then determine that the predicted plane is suitable from the estimate error. After that, we correct the target pixel by the plane mode. The proposed method can correct not only the depth image composed the plane but also the complex depth image. From the simulation result that measures the entropy power, which can estimate the coding performance, we can see that the coding performance by the proposed method is improved up to 80.2%.

Key Words : Depth Image Correction, Noise Correction, Depth Video Coding

1. 서 론

기존에는 색상 영상을 통해 객체를 인식하고, 객체에 대한 정보를 획득하여 정보를 처리하는 경우가 일반적이었다[1-2]. 하지만 최근 깊이 측정 기술의 발달로 인해 깊이 카메라나 깊이 측정 센서의 보급이 널리 퍼지고 있고, 이에 따라 측

정된 깊이 정보를 이용하여 객체를 인식하여 이에 따른 정보를 처리하는 연구가 많아지고 있다 [3-6]. 하지만 깊이를 측정할 때, 측정 오차로 인해 잡음을 포함하는 깊이 영상이 촬영된다[7-8]. 이 잡음은 깊이를 이용한 정보 처리에 있어서 정확도를 저하시키는 요인이 된다. 깊이를 이용한 정보 처리에서 정확도를 향상시키기 위해서는 깊이 영상의 잡음을 제거하는 과정이 필수적이다.

깊이 영상을 이용한 기존 보정 방법은 평면 모델에서의 잡음 모델을 이용하여 잡음을 제거하거나, 비 평면 모델에서 나타나는 잡음 모델을 이용하여 잡음을 제거하는 방법이 연구되었다

* Corresponding Author : skkwon@deu.ac.kr

† 이 논문은 2016년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 지역산업선도인력양성사업 성과임(No. 2016909955).

Manuscript received Oct. 13, 2016 / accepted Oct. 19, 2016

1) 동의대학교 컴퓨터소프트웨어공학과 석사과정

2) 동의대학교 컴퓨터소프트웨어공학과 교수

[9-10]. 또한 색상 영상과 정합을 통해 깊이 오차를 제거하는 방법도 제안되었다[11]. 하지만 잡음 모델을 이용한 방법은 깊이가 카메라의 깊이 측정 방법에 따라 잡음 모델이 달라지기 때문에 범용적으로 이용하기엔 무리가 있다. 또한 색상 영상을 이용하여 깊이를 보정하는 방법은 환경에 영향을 받는 색상 영상의 특성상 정확한 깊이 보정을 기대하기 어렵다.

본 논문에서는 평면 모델링을 이용하여 깊이 영상을 보정하는 방법을 제안한다. 이는 촬영하는 영역을 여러 개의 평면으로 분할 할 수 있다는 것을 이용한 것이다. 먼저 보정하고자 하는 화소를 중심으로 수평, 수직 방향으로 평면의 단면이 되는 직선을 모델링한다. 이 때 최소자승법을 이용하여 평면의 단면인 직선의 방정식 계수를 찾는다. 그 후 예측오차를 이용하여 모델링한 평면이 적합한지를 판단한다. 그 후 수직, 수평 방향의 보정 값이 값을 이용하여 최종 보정 값을 계산하여 깊이를 보정한다.

본 논문에서 제안된 방법을 통해 카메라의 종류에 상관없이 깊이를 보정할 수 있다. 뿐만 아니라 현재 깊이 측정 센서를 장착하는 모바일 제품이 증가하고 있는데, 본 논문에서 제안된 보정 방법을 적용하여 이러한 제품들의 깊이 측정 정확도를 향상시킬 수 있다.

2. 평면 모델링을 이용한 깊이 영상에서의 잡음 제거

2.1 평면 모델링과 최소자승법을 통한 깊이 값 예측

깊이 영상에서 굴곡이 없는 평면을 촬영한다고 가정하면 평면은 Fig. 1 (a)와 같이 투영된다. 이 평면을 Y축과 수직인 평면으로 절단하면 촬영된 평면은 Fig. 1 (b)과 같이 ZX평면 좌표계에서 직선으로 나타난다.

영상 좌표계에서의 한 좌표 (x, y) 위의 화소 p 가 깊이 D 를 가질 때, 이 화소로 투영된 원 카메라 좌표계에서의 좌표는 $(Dx/f, Dy/f, D)$ 이다 [12]. 이 좌표를 절단면에서의 직선의 방정식

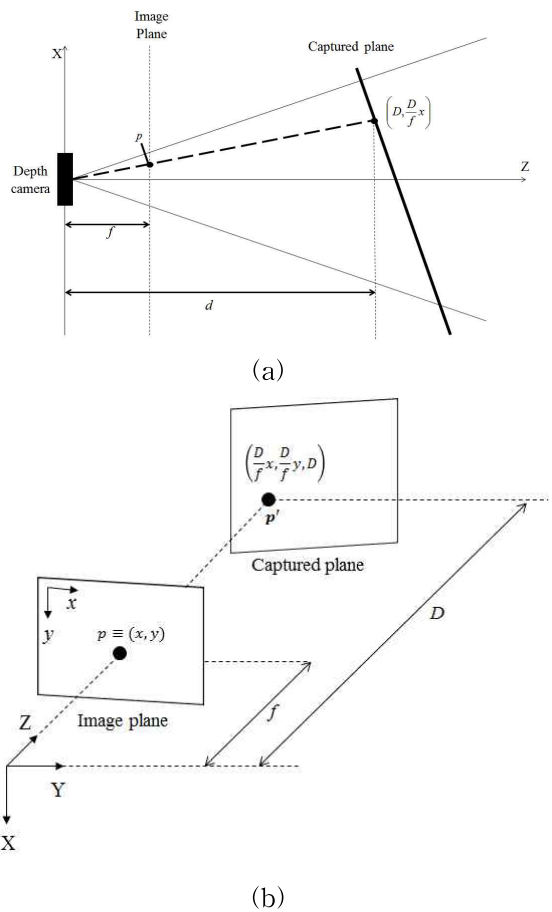


Fig. 1 Projection of captured plane:
 (a) Projection of p on plane and
 (b) Cross section cutting plane parallel to Y axis.

$X=aZ+b$ 에 대입하여 D 에 대해 정리하면 식 (1)의 형태가 된다. 이는 깊이 값 D 가 x 에 대해 분수함수의 관계를 가진다는 것을 뜻한다.

$$D = \frac{bf}{x - af} \tag{1}$$

평면을 깊이 카메라를 통해 깊이 영상을 촬영하면 이상적인 경우에는 Fig. 2에서의 점선과 같이 깊이가 촬영되어야 한다. 하지만 실제로는 깊이 측정 오차로 인해 잡음이 포함된 깊이가 Fig. 2에서의 실선과 같이 촬영된다[13].

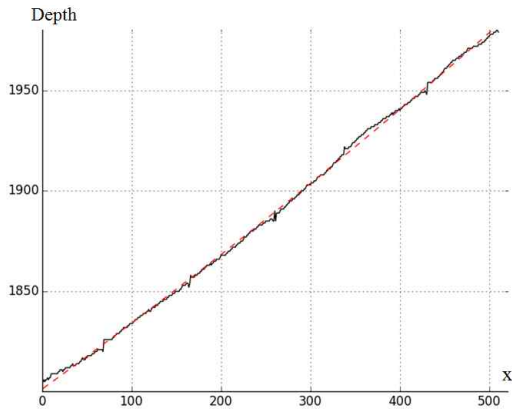


Fig. 2 Depth difference between ideal and measured values when plane is captured.

평면을 깊이 카메라로 촬영할 때, 측정된 깊이 \tilde{D} 는 잡음으로 인해 그 위치에서의 이상적인 깊이 D 와 다를 수 있다. 측정된 \tilde{D} 를 보정하기 위해 평면을 모델링하여 이상적인 깊이 D 를 찾아 보정한다. 이를 위해 Fig. 1 (b)에서 절단면상의 화소들 p_i ($0 \leq i \leq N$)을 이용하여 최소자승법을 통해 식 (1)의 인자 a, b 를 예측할 수 있다.

최소자승법을 적용하기 위해 $a=1/af, \beta=-b/a, u=xD$ 로 치환하고, 식 (1)에 p_i 의 영상 좌표계에서의 좌표 (x_i, y_i) 와, p_i 에서 측정된 깊이 D_i 를 대입하면 식 (2)와 같이 정리된다.

$$D_i = \alpha u_i + \beta \quad (2)$$

그 후 이상적인 깊이와 측정 깊이 간 오차의 제곱합 $(\tilde{D}_i - D_i)^2$ 을 최소로 하는 a, β 를 찾기 위해 식 (3)과 같이 편미분을 수행한 후 p_i 들을 이용하여 a, β 를 찾는다.

$$\begin{aligned} \frac{\delta}{\delta \alpha} \sum_{i=0}^N (\tilde{D}_i - D_i)^2 &= 2 \sum_{i=0}^N (\tilde{D}_i - \alpha u_i - \beta)(-u_i) = 0 \\ \frac{\delta}{\delta \beta} \sum_{i=0}^N (\tilde{D}_i - D_i)^2 &= 2 \sum_{i=0}^N (\tilde{D}_i - \alpha u_i - \beta)(-1) = 0 \end{aligned} \quad (3)$$

식 (3)을 통해 찾은 a, β 를 식 (2)에 대입한 후, x_i 를 식 (2)에 대입하면 p_i 의 이상적인 깊이 값 D_i 를 찾을 수 있다.

위 방법을 이용하여 수평방향 뿐만 아니라 수직방향에 대해서도 보정값을 구할 수 있다. 또한 여러 방향에서 구한 보정값을 이용하여 좀 더 정확한 보정값을 찾을 수 있다.

2.2 깊이 값 예측을 통한 깊이 영상 보정

촬영 영역이 되는 객체들은 한 개 이상의 평면과 곡면으로 분할할 수 있다. 또한 곡면에 대해서도 곡률이 작은 곡면은 평면으로 간주할 수 있다. 이를 이용하여 깊이 카메라로 촬영한 깊이 정보에 대해 평면 모델링을 통해 정확한 깊이 값을 예측할 수 있다.

보정하고자 하는 화소 $p(x,y) \equiv (x,y)$ 에 대해 좌우 각각 N 개의 화소 $p_{(x \pm k,y)}$ ($1 \leq k \leq N$)를 이용한다. 그 후 이들 화소에서 측정된 깊이들을 식 (3)에 대입하여 식 (2)에서의 a, β 를 찾은 후, x 를 식 (2)에 대입하여 D_x 를 구한다.

이 때 곡면이 큰 표면이나 객체 간 경계부분 등의 경우에는 평면 모델링을 통한 깊이 값 예측이 정확하지 않다. 이러한 경우를 판단하기 위해 최소자승법을 적용할 때 예측오차 E 를 식 (4)와 같이 구하여 예측의 정확도를 추측할 수 있다. 이 때, E 가 T ($0 < T \leq 1$) 이하이면, D_x 는 평면 모델링을 이용한 예측을 이용하여 깊이 값을 추측하지 않고, 범위 내 깊이의 평균을 이용한다.

$$\begin{aligned} E &= \frac{S_t - S}{S_t} \\ S_t &= \sum_{k=-N}^N (\tilde{D}_{(x+k,y)} - \bar{D})^2, \\ \bar{D} &= \frac{1}{2N+1} \sum_{k=-N}^N \tilde{D}_{(x+k,y)}, \\ S &= \sum_{k=-N}^N (\tilde{D}_{(x+k,y)} - D_x(x+k,y))^2 \end{aligned} \quad (4)$$

수직 방향에 대해서도 같은 방법을 통해 D_y 를 구한다. 그 후 수평방향으로 예측한 깊이 값 D_x 와, 수직 방향으로 예측한 깊이 값 D_y 를 평균하여 평균 예측 값 $D=(D_x+D_y)/2$ 를 구하여 화소 $p_{(x \pm k,y)}$ 의 깊이 값을 보정한다.

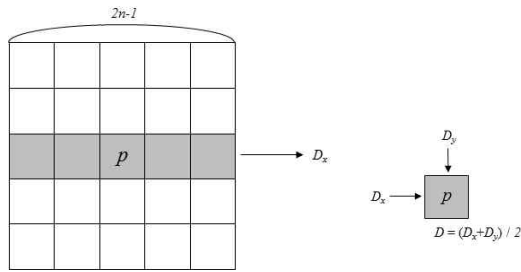


Fig. 3 Correction depth image through modeling of plane.

3. 모의실험결과

3.1 결정변수와 환경요인의 설정

본 논문에서 제안된 보정 방법을 이용하여 여러 영상에 대해 보정을 수행하고, 보정된 영상에 대해 부호화 효율을 계산하였다. 이 때 실험에 이용한 촬영영상은 Fig. 4와 같다. 평면에 대하여 깊이 값을 보정한 결과와 촬영되는 객체가 많아 깊이 값이 복잡한 환경에서 촬영한 영상에 대해 보정을 수행하였다. 이 때 깊이 카메라로는 Kinect v2를 이용하였고, 깊이 영상의 해상도는 512X424를 이용하였다. 또한 부호화를 수행하는데 이용하는 영역 N 은 3을 적용하였고, T 는 0.80을 적용하였다.

부호화 방법은 수평 방향으로 DPCM을 수행하여 부호화하였다. 또한 부호화 효율을 측정하기 위해 부호화된 영상의 엔트로피 파워를 측정하였다. 엔트로피 파워는 측정 대상과 동일한 엔트로피를 가지는 환경에서 모든 신호의 빈도가 동일한 백색잡음의 출력으로 정의된다. 부호화가 된 영상 X 의 엔트로피를 $h(X)$ 라 한다면, 엔트로피 파워 $E(X)$ 는 식 (5)과 같이 정의된다. 이 때 f_i 는 신호 i 의 확률로써, 영상에서는 그 신호가 나타난 빈도수를 전체 픽셀수로 나눈 것이다.

$$E(X) = \frac{1}{2\pi e} e^{2h(X)} \quad (5)$$

$$h(X) = - \sum_i f_i \ln(f_i)$$



(a)



(b)

Fig. 4 Source image for simulation:

(a) Simple plane image and (b) Image including objects.

엔트로피 파워가 낮다는 것은 깊이 영상의 특성에 맞는 부호화 방법을 적용하였을 때 압축 효과를 기대할 수 있다는 의미이다. 따라서 보정 전 영상과 보정 후 영상에 대해 엔트로피 파워를 비교하여 부호화 효율을 측정할 수 있다.

3.2 모의실험결과

먼저 Fig. 4에 대한 영상들에 대해 보정을 수행하기 전과 수행 후 깊이 영상을 비교하였다. 이 때 보정 전후의 영상에서 깊이 값을 비교하였다. Fig. 5와 Fig. 6은 Fig. 4를 보정하기 전의 깊이 영상과 보정한 후의 깊이 영상을 하나의 y 값에서 x값에 따른 깊이 값을 비교한 것이다.

Fig. 5과 Fig. 6에서 점선은 보정 전의 깊이 값이고, 실선은 보정 후의 깊이 값이다. 이 때 평면으로 이루어진 깊이 영상에서는 Fig. 5와 같이 보정 후에 깊이 값이 평면에 거의 일치하게 보정되었다. 인물이 포함된 영상에 대해서도 Fig. 6과 같이 깊이 값에 포함된 잡음이 상당수 제거됨을 볼 수 있다.

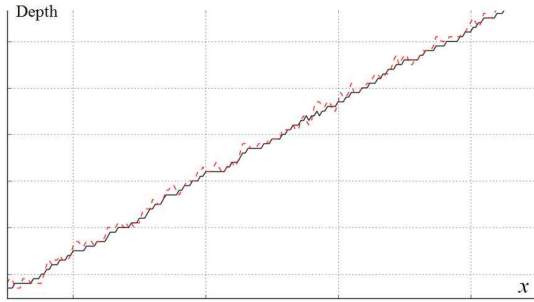


Fig. 5 Depth comparison of before and after correction of Fig. 4 (a).

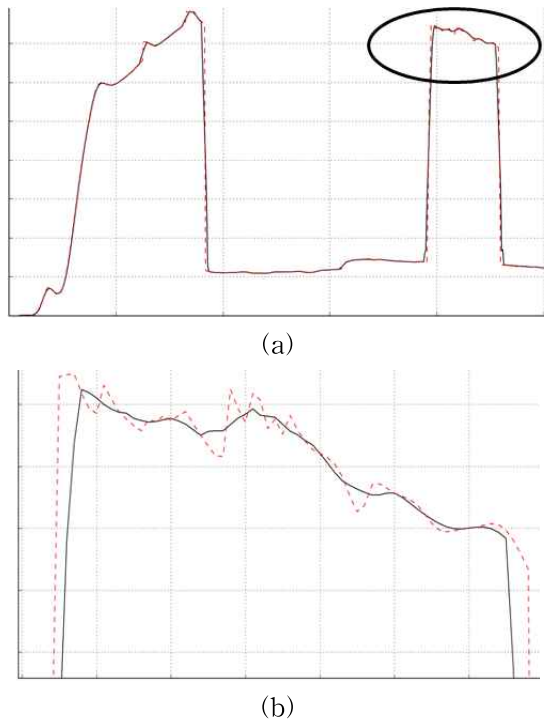


Fig. 6 Depth comparison of before and after correction of Fig. 4 (b):
(a) All area and (b) Magnifying part of (a).

또한 보정 전과 보정 후의 영상에 대해 수평 DPCM을 통한 부호화를 수행한 후의 엔트로피 파워를 측정하였다. 이 때, Fig. 4 (a) 영상에 대하여서는 엔트로피 파워가 80.2% 감소하였고, Fig. 4 (b)에 대해서는 45.66% 감소하였다. 이는 이 보정방법을 통한 부호화가 평면 영역이 많은 영상에서 더욱 효과적임을 알 수 있다.

Table 1 Measurement of entropy power between before and after correction.

Source Image	Entropy power	
	Before correction	After correction
Fig. 4 (a)	4.272	0.845
Fig. 4 (b)	262.924	142.874

보정에 이용하는 화소의 개수 N 을 달리하여 엔트로피 파워를 비교하였다. 그 결과 Fig. 4 (a) 처럼 평면영역이 많은 영상에 대해서는 N 이 클수록 보정의 효율이 좋았다. 하지만 Fig. 4 (b) 처럼 표면이 복잡한 깊이 영상에 대해서는 N 이 클수록 오히려 보정의 효율이 떨어지는 것을 확인할 수 있었다. 또한 N 이 과도하게 클 경우 보정이 과도하게 일어남으로 인해 깊이 정보가 상당수 손실되었다. 이는 Fig. 7에서 N 에 따른 보정 결과에서 확인할 수 있다. Fig. 7에서 실선은 N 이 3일 때이며, 점선은 N 이 10일 때, 점과 점선이 섞인 선은 N 이 30일 경우이다.

Table 2 Measurement of entropy power according to N .

Source Image	N		
	3	10	30
Fig. 4 (a)	0.845	0.627	0.620
Fig. 4 (b)	142.874	173.149	261.202

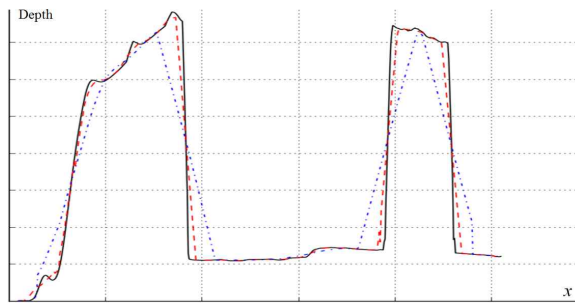


Fig. 7 Depth comparison according to n after correcting Fig. 4 (b).

4. 결론

본 논문에서는 평면 모델링을 이용하여 깊이 영상을 보정하는 방법을 제안하였다. 먼저 보정하고자 하는 화소를 중심으로 수평, 수직 방향으로 평면의 단면이 되는 직선을 모델링하였다. 이때 최소자승법을 이용하여 직선 방정식의 계수를 찾아냈다. 그 후 예측오차를 이용하여 모델링한 평면이 적합한지를 판단하였다. 그 후 수직, 수평 방향의 보정 깊이 값을 측정하여 두 값을 보간하여 최종 보정 값을 찾았다. 실험결과 평면을 촬영한 영상에서는 보정의 효율이 좋음을 확인할 수 있었다. 또한 복잡한 깊이 값을 갖는 깊이 영상에 대해서도 보정이 비교적 잘 되었음을 확인할 수 있었다.

본 논문에서 제안된 방법을 통해 정확한 깊이 값을 찾아내어 좀 더 정밀한 깊이정보를 통한 제스처 인식이나 촬영 영역에 대한 모델링 정보를 획득할 수 있다. 뿐만 아니라 본 논문을 적용함으로써 깊이 영상 부호화의 효율이 좋아진다는 것이 실험 결과 밝혀졌다. 이를 통해 앞으로 깊이 영상 처리에 있어서 널리 응용될 것으로 기대된다.

References

[1] S. Y. Kim, C. Y. Yoon, and E. J. Yu, "A Study on the Development of Learning

Contents of Augmented Reality by Perception Rate and Speeding," *The Journal of Internet Electronic Commerce Research*, Vol. 14, No. 4, pp. 313-333, 2014.

[2] S. Y. Kim and S. M. Lee, "Implementation of an Image Board Remote Control System Using PDA Based on Embedded Linux in Wireless Internet," *The Journal of Information Systems*, Vol. 17, No. 1, pp. 155-171, 2008.

[3] R. M. Salinas, R. Medina-Carnicer, F. J. Madrid-Cuevas, and A. Carmona-Poyato, "Depth Silhouettes for Gesture Recognition," *Pattern Recognition Letters*, Vol.29, No.3, pp.319-329, 2008.

[4] P. Suryanarayan, A. Subramanian, and D. Mandalapu, "Dynamic Hand Pose Recognition Using Depth Data," *Proceeding of 20th International Conference on Pattern Recognition*, pp. 3105-3108, 2010.

[5] M. Zollhoefer, M. Martinek, G. Greiner, M. Stamminger, and J. Suessmuth, "Automatic Reconstruction of Personalized Avatars From 3d Face Scans," *Computer Animation and Virtual Worlds*, Vol. 22, No. 2-3, pp. 195 - 202, 2011.

[6] J. Tong, J. Zhou, L. Liu, Z. Pan, and H. Yan, "Scanning 3D Full Human Bodies Using Kinects," *IEEE Transactions on Visualization and Computer Graphics*, Vol. 18, No. 4, pp. 643 - 650, 2012.

[7] B. Choo, M. Landau, M. Devore, and P. A. Beling, "Statistical Analysis-Based Error Models for the Microsoft Kinect™ Depth Sensor," *Sensors*, Vol. 14, No. 9, pp. 17430-17450, 2014.

[8] C. V. Nguyen, S. Izadi, and D. Lovell, "Modeling Kinect Sensor Noise for Improved 3d Reconstruction and Tracking," *3D Imaging, Modeling, Proceeding of 2012 Second International Conference on Processing, Visualization and Transmission*,

pp. 524-530, 2012.

[9] Y. S. Kim, B. M. Kang, H. S. Lim, O. Choi, K. C. Lee, J. D. K. Kim, and C. Y. Kim, "Parametric Model-Based Noise Reduction for ToF Depth Sensors," Proceeding of SPIE 8290, Three-Dimensional Image Processing (3DIP) and Applications II, pp. 82900A-82900A-8, 2012.

[10] A. Belhedi, A. Bartoli, V. Gay-Bellile, S. Bourgeois, P. Sayd, and K. Hamrouni, "Non-Parametric Depth Calibration of a Tof Camera," Proceeding of 2012 19th IEEE International Conference on Image Processing, pp. 549-552, 2012.

[11] D. Herrera, J. Kannala, and J. Heikkilä, "Joint Depth and Color Camera Calibration with Distortion Correction," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 34, No. 10, pp. 2058-2064, 2012.

[12] R. Hartley and A. Zisserman, "Multiple View Geometry in Computer Vision," Cambridge University Press, Cambridge, UK, 2008.

[13] S. K. Kwon and D. S. Lee, "Obtainment of Background Image Using Depth Information," International Organization of Scientific Research Journal of Engineering, Vol. 5, No. 8, pp. 43-46, 2015.



이 동 석 (Dong-Seok Lee)

- 정회원
- 동의대학교 컴퓨터소프트웨어공학과 공학사
- 동의대학교

컴퓨터소프트웨어공학과 석사과정

- 관심분야 : 멀티미디어 신호처리, 영상 인식



권 순 각 (Soon-Kak Kwon)

- 정회원
- 경북대학교 전자공학과 공학사
- KAIST 전기및전자공학과 공학 석사

- KAIST 전기및전자공학과 공학박사
- 동의대학교 컴퓨터소프트웨어공학과 교수
- 관심분야 : 멀티미디어처리, IOT