

구형 객체의 깊이 영상 부호화 방법

(Depth Video Coding Method for Spherical Object)

권순각^{1)*}, 이동석²⁾, 박유현³⁾

(Soon-Kak Kwon, Dong-Seok Lee, and Yoo-Hyun Park)

요약 본 논문은 구형 객체가 촬영된 깊이 영상에서 깊이 정보를 통하여 제일 근접한 구를 찾아내어 깊이 영상을 부호화하는 방법을 제안한다. 블록단위로 분할된 깊이 영상에 대해 최소자승법을 통해 촬영된 구형 객체와 제일 근접한 구의 형태를 찾는다. 그 후 찾아낸 구의 형태로 깊이 값을 예측하고, 측정된 깊이 값과의 오차를 통해 깊이 영상을 부호화한다. 또한, 블록 내의 부호화된 각 깊이 화소들과 찾아낸 구의 인자 정보를 같이 부호화한다. 제안된 방법으로 구형 객체에 대해 기존 DPCM 방법보다 최대 81% 이상의 부호화 효율 향상이 이루어졌다.

핵심주제어 : 깊이 영상 부호화, 구 모델링, 표면 모델링

Abstract In this paper, we propose a method of depth video coding to find the closest sphere through the depth information when the spherical object is captured. We find the closest sphere to the captured spherical object using method of least squares in each block. Then, we estimate the depth value through the found sphere and encode the depth video through difference between the measured depth values and the estimated depth values. Also, we encode factors of the estimated sphere with encoded pixels within the block. The proposed method improves the coding efficiency up to 81% compared to the conventional DPCM method.

Key Words : Depth Video Coding, Sphere Modeling, Surface Modeling

1. 서론

깊이 영상을 이용하여 기존 색상 영상에서는 얻을 수 없었던, 객체의 위치 정보와 거리 정보를 획득하여, 이를 통한 새로운 객체의 정보를 획득할 수 있다. 이러한 깊이 영상의 특성을 통

한, 깊이 영상을 이용한 새로운 응용분야에 대한 연구가 아래와 같이 이루어지고 있다.

색상 영상에서의 환경에 민감하고, 객체내의 여러 다른 색상 정보가 존재한다는 단점을 깊이 카메라를 통해 극복하여 좀 더 정확한 객체 검출을 수행하는 연구가 이루어졌다[1-2]. 깊이 영상의 거리를 이용하여 평면의 정보를 이용하여 이를 통해 영상의 왜곡과 잡음을 제거하는 연구도 이루어졌다[3-4]. 또한 깊이 영상의 거리정보를 이용하여 배경영역에 터치를 인식하는 연구[5-6]와, 이를 이용하여 여러 이벤트를 제공하는 연구[7]가 이루어졌다. 이뿐만 아니라 인물의 형태를 인식하여 사람의 얼굴을 인식하는 연구[8-10]도

* Corresponding Author : skkwon@deu.ac.kr

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

Manuscript received Dec. 6, 2016 / accepted Dec. 14, 2016

1) 동의대학교 컴퓨터소프트웨어공학과, 교신저자

2) 동의대학교 컴퓨터소프트웨어공학과

3) 동의대학교 컴퓨터소프트웨어공학과

여러 이루어졌다.

이러한 깊이 영상의 응용분야의 확대로 인해 깊이 영상의 부호화에 대한 필요성이 증가되었다. 먼저 Depth Lookup Table를 이용하여 깊이 영상을 부호화하는 방법이 연구되었다[11]. 또한 객체의 경계 정보를 깊이 영상 부호화에 이용하는 방법이 제안되었다[12-13]. 깊이 영상을 히스토그램기반으로 분석하여 깊이 영상을 부호화 하는 방법도 제안되었다[14]. 이처럼 깊이 영상에 나타나는 특징을 이용하여 영상 부호화 효율을 높이는 연구가 많이 이루어졌다. 하지만 지금까지 이루어진 깊이 영상 부호화 연구는 대체로 색상 영상 부호화에 대해 보조적으로 이루어지거나, 색상 영상과 연계하여 깊이 영상을 부호화하는 방법에 한정되었다.

깊이 영상의 깊이 값은 촬영 영역의 표면의 정보로 볼 수 있다. 이 때, 표면에 대해 간단한 도형의 표면으로 모델링하여 이를 통해 깊이 영상을 부호화할 수 있다. 이에 본 논문에서는 구형 객체의 깊이 영상을 부호화 하는 방법을 제안하고자 한다. 먼저 깊이 영상 내 블록에 대해 제일 근접한 구를 최소자승법을 이용하여 찾아낸 후 이를 통해 깊이 영상을 부호화한다. 이러한 방법을 이용하여 깊이 영상 내 구형의 객체를 효율적으로 부호화할 수 있다. 제안된 방법을 통해 깊이 영상을 이용하는 여러 응용 연구 분야에 적용할 수 있다.

2. 구형 객체 모델링과 영상 부호화

깊이 영상에서 구형 객체에 대해 객체의 표면을 이루는 구면을 추측하는 단계는 Fig. 1과 같다. 이는 $N \times N$ 크기로 분할된 깊이 영상에서 한 블록 내의 깊이 정보를 통해 해당 블록의 깊이 정보로 표현되는 표면과 제일 근접한 구면을 모델링하는 것이다.

깊이 영상 내 구형 객체의 표면은 Z축이 카메라의 광축으로 나타내는 3차원 카메라 좌표계에서 식 (1)로 모델링된다.

$$(x-a)^2 + (y-b)^2 + (z-c)^2 = r^2 \quad (1)$$

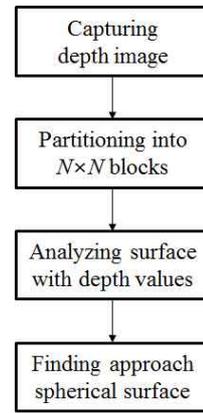


Fig. 1 Flowchart of spherical surface modeling

이러한 구 표면은 깊이 카메라를 통해 카메라 초점 거리 f 에 위치하는 가상의 영상 평면의 (h, w) 좌표에 위치한 화소에 깊이 값 d 가 투영된다. Fig. 2는 이러한 구 표면의 투영을 보인다.

이 때 식 (1)을 실제로 깊이 영상에 적용하기 위해서는 2차원 좌표인 영상 좌표계와 3차원 좌표인 카메라 좌표계 간 변환을 수행하여야 한다. 이 때, 깊이 영상은 각 화소마다 카메라 좌표계에서의 Z축 좌표인 깊이 값을 화소로 가진다. 이 깊이 값을 이용하여 영상 좌표계에서 카메라 좌표계로 변환을 수행할 수 있다. 2차원 영상 좌표계에서의 화소의 위치 (h, w) 과 카메라 좌표계에서의 좌표 (x, y, z) 의 관계는 식 (2)와 같다. 이 때 초점거리 f 는 식 (3)을 통해 구할 수 있다. 식 (3)에서 height와 width는 깊이 영상의 수직, 수평 방향의 해상도이다. 또한 fov_v , fov_h 는 각각 카메라의 수직 시야각, 수평 시야각으로 카메라 내부 인자 중 하나이다.

$$x = \frac{d}{f}w, y = \frac{d}{f}h, z = d \quad (2)$$

$$f = \frac{height}{2} \frac{1}{\tan(fov_v/2)} \quad (3)$$

$$= \frac{width}{2} \frac{1}{\tan(fov_h/2)}$$

구 표면을 나타내는 식인 식 (1)에 식 (2)를 대입하여 화소 좌표계의 좌표 (h, w) 로 나타낼 수 있다. 이 식을 깊이 d 로 정리하여 식 (4)을

얻는다.

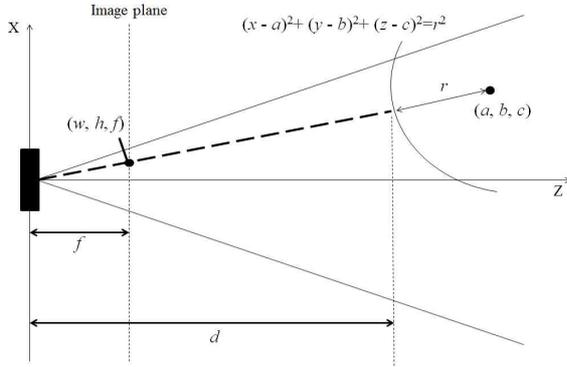


Fig. 2 Spherical surface projected onto image plane

$$\left(\frac{1}{f}\right)^2 (w^2 + h^2 + f^2)d^2 - \frac{2}{f}(aw + bh + cf)d + (a^2 + b^2 + c^2 - r^2) = 0 \quad (4)$$

d 의 이차식으로 나타나는 식 (4)의 근을 구함으로써 식 (5)와 같이 d 를 구할 수 있다.

$$d = -\frac{\alpha \pm \sqrt{\alpha^2 + \beta(a^2 + b^2 + c^2 - r^2)}}{\beta} f \quad (5)$$

$$\alpha = aw + bh + cf$$

$$\beta = w^2 + h^2 + f^2$$

식 (5)에서의 d 는 식 (1)로 나타나는 구면의

$$\mathbf{X}_n = \begin{bmatrix} a_n \\ b_n \\ c_n \\ r_n \end{bmatrix}, \quad \mathbf{F}_n = \begin{bmatrix} e_n(h_1, w_1) \\ e_n(h_1, w_2) \\ \dots \\ e_n(h_N, w_N) \end{bmatrix},$$

$$\mathbf{J}(\mathbf{F}_n) = \begin{bmatrix} \frac{\delta}{\delta a_n} e_n(h_1, w_1) & \frac{\delta}{\delta b_n} e_n(h_1, w_1) & \frac{\delta}{\delta c_n} e_n(h_1, w_1) & \frac{\delta}{\delta r_n} e_n(h_1, w_1) \\ \dots & \dots & \dots & \dots \\ \frac{\delta}{\delta a_n} e_n(h_N, w_N) & \frac{\delta}{\delta b_n} e_n(h_N, w_N) & \frac{\delta}{\delta c_n} e_n(h_N, w_N) & \frac{\delta}{\delta r_n} e_n(h_N, w_N) \end{bmatrix}$$

$$\frac{\delta}{\delta a_n} e_n(h, w) = \frac{f}{w^2 + h^2 + f^2} \left(-w \pm \frac{w(aw + bh + cf) - a(w^2 + h^2 + f^2)}{\sqrt{(aw + bh + cf)^2 - (w^2 + h^2 + f^2)(a^2 + b^2 + c^2 - r^2)}} \right) \quad (6)$$

$$\frac{\delta}{\delta b_n} e_n(h, w) = \frac{f}{w^2 + h^2 + f^2} \left(-h \pm \frac{h(aw + bh + cf) - b(w^2 + h^2 + f^2)}{\sqrt{(aw + bh + cf)^2 - (w^2 + h^2 + f^2)(a^2 + b^2 + c^2 - r^2)}} \right)$$

$$\frac{\delta}{\delta c_n} e_n(h, w) = \frac{f}{w^2 + h^2 + f^2} \left(-f \pm \frac{f(aw + bh + cf) - c(w^2 + h^2 + f^2)}{\sqrt{(aw + bh + cf)^2 - (w^2 + h^2 + f^2)(a^2 + b^2 + c^2 - r^2)}} \right)$$

$$\frac{\delta}{\delta r_n} e_n(h, w) = f \left(\mp \frac{r}{\sqrt{(aw + bh + cf)^2 - (w^2 + h^2 + f^2)(a^2 + b^2 + c^2 - r^2)}} \right)$$

인자 a, b, c, r 가 주어졌을 때의 깊이 값으로, 이상적인 깊이 값으로 볼 수 있다. 이상적인 깊이 값 d 와 (h, w) 에서 실제로 측정된 깊이 값 \tilde{d} 의 차이가 제일 작을 때의 a, b, c, r 를 구함으로써 해당 블록에서 최적의 구 표면을 결정할 수 있다. 이 때 모델링된 표면의 (h, w) 에서의 근사된 곡면의 깊이와 실제로 측정된 깊이의 오차 $e \equiv d - \tilde{d}$ 가 최소가 되도록 구의 인자를 결정하기 위해 최소자승법을 적용한다. 이 때, e 는 인자에 대해 비선형식으로 나타나므로 가우스-뉴턴법을 적용한다.

가우스-뉴턴 법에서 n 단계에서 블록 내 각 화소에서의 d 와 \tilde{d} 의 차로 이루어진 행렬 \mathbf{F}_n 과 \mathbf{F}_n 에서의 자코비안 행렬 $\mathbf{J}(\mathbf{F}_n)$, 인자 값을 나타내는 \mathbf{X}_n 는 식 (6)과 같다.

가우스-뉴턴법을 적용하여 식 (7)과 같이 다음 단계의 인자 값인 \mathbf{X}_{n+1} 를 구한다.

$$\mathbf{X}_{n+1} = \mathbf{X}_n - \mathbf{F}_n / \mathbf{J}_n \quad (7)$$

$$= \mathbf{X}_n - (\mathbf{J}_n^T \mathbf{J}_n)^{-1} \mathbf{J}_n^T \mathbf{F}_n$$

이를 P 회 반복하여 주어진 깊이 값으로 이루어진 표면과 제일 근접한 곡면을 구할 수 있다. 이 때, 초기 인자 $\mathbf{X}_1 = [a_1, b_1, c_1, r_1]^T$ 를 정함에 있어 a_1 와 b_1 값은 블록 중심의 좌표인 $h_{N/2}$ 와 $w_{N/2}$ 를 식 (2)에 대입하여 구할 수 있고, r_1 는 임의의 초

기값 r_0 를 대입한다. c_l 는 블록 중심의 깊이 값 $\tilde{d}(h_{N/2}, w_{N/2})$ 에서 r_l 를 더한 수치를 초기 값으로 한다.

그 후 구해진 인자를 토대로 해당 블록이 구 표면의 일부인지를 확인한다. 이를 위해 식 (8)과 같이 가우스-뉴턴법을 P 회 적용한 뒤 얻어진, 오차를 나타내는 행렬 \mathbf{F}_P 행렬 내의 모든 인자의 제곱 합을 더한 값 ϵ 를 구한다. 이 값이 문턱 값 E 이내일 경우 구 표면의 일부로 인정하여 부호화 과정을 수행한다. 그렇지 않을 경우엔 구면이 아닌 다른 표면으로 판단하여 해당 방법을 적용하지 않는다.

$$\epsilon = \sum_{k=1}^{N^2} (\mathbf{F}_P)_k \quad (8)$$

구면으로 판정된 블록 내 각 화소 (h_i, w_j)에서의 $\tilde{d}(h_i, w_j)$ 와 모델링 과정을 수행하여 찾은 인자에서의 깊이 값 $\bar{d}(h_i, w_j)$ 의 차이를 이용하여 각 화소를 부호화한다. 그 후 각 블록에서 블록 내 부호화된 화소와 함께 구의 인자 값도 부호화하여 전체 영상을 부호화할 수 있다.

3. 모의실험 및 결과

3.1 결정변수와 환경요인의 설정

본 논문에서는 Microsoft사의 Kinect V2 깊이 카메라를 이용하여 모델링을 통한 구형 객체 표면 추측의 정확도와 이를 이용한 깊이 영상 부호화 성능을 측정하였다. 이 때 Kinect V2는 ToF(Time of Flight) 방식으로 깊이를 측정하며 FOV(Field of View)는 수평 방향으로 70.6°, 수직방향으로 60°이다. 또한 깊이 영상 해상도는 512x424이다. 가우스-뉴턴 법을 적용함에 있어 r_0 는 100을 적용하였고, 가우스-뉴턴 법을 반복하는 횟수 P 는 100회를 적용하였다.

원본 영상은 Fig. 3을 촬영한 깊이 영상을 사용하였다. 이 때 구 모델링 및 구 부호화의 효율성을 측정하는 실험에서는 Fig. 3 (a)의 깊이 영

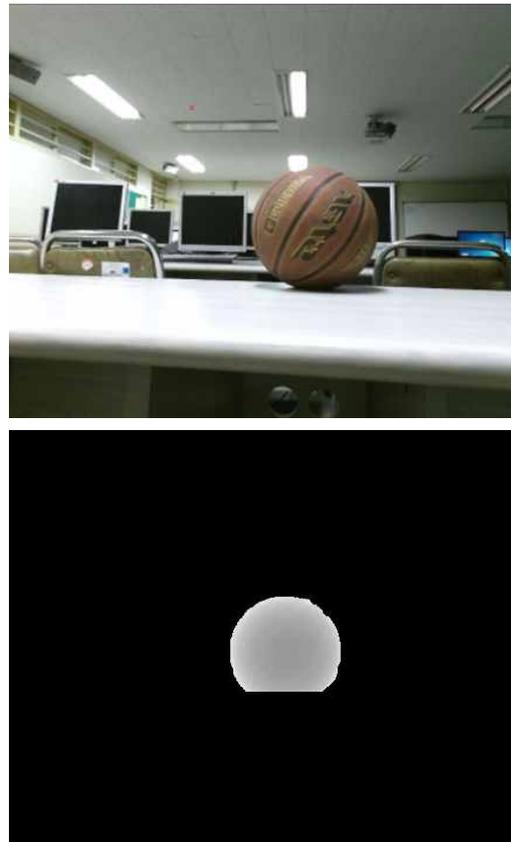


Fig. 3 Sample image of spherical object: (a) Color image and (b) Depth image

상에서 배경을 제거한 깊이 영상 Fig. 3 (b)를 이용하였다. 이 때 객체가 없는 깊이 영상을 먼저 촬영한 뒤 객체를 촬영하여 배경에서 촬영영상을 빼는 방법을 통해 객체 영상을 얻었다[15].

3.2 모의실험결과

Fig. 4는 Fig. 3 (b)의 깊이 영상을 이용하여 제안된 방법을 적용하여 구한, 구 표면의 깊이 영상이다. 이 때 블록의 크기 N 이 작아질수록 제안된 방법을 통해 부호화를 수행할 수 있는 영역이 늘어남을 알 수 있다.

Fig. 5는 원본 깊이 영상과 제안된 방법을 통해 모델링 된 구 표면의 깊이 값을 나타낸 것이다. 이 때 점선은 실제로 측정된 구형 객체의 깊이 값이고, 실선은 제안된 방법을 통해 구한 구 표면의 깊이 값이다. 이 때 깊이 카메라의 측정 오차로 인해

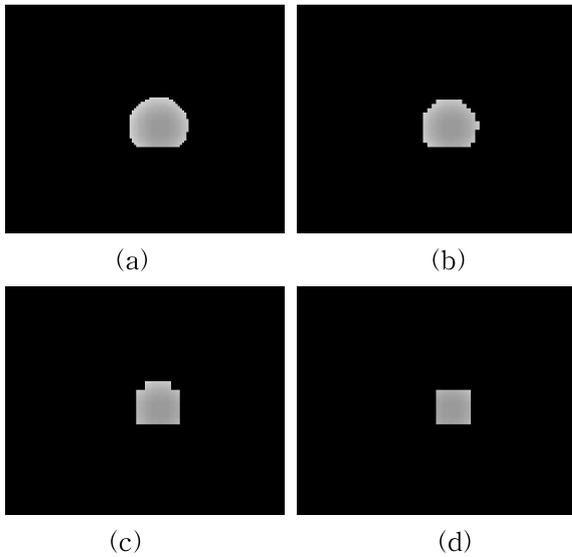


Fig. 4 Estimated spherical surface for proposed method: (a) $N=4$, (b) $N=8$, (c) $N=16$ and (d) $N=32$

불균일한 깊이 값을 가지는 촬영 깊이 값에 대해 제안된 방법을 통해 근접한 구 표면을 찾아냈음을 알 수 있다.

제안된 방법을 통해 추측된 구 표면의 깊이 값에 대해 MSE를 블록의 크기에 따라 측정하였다. 그 결과 Table 1와 같이 N 이 작아짐에 따라 MSE가 줄어들음을 확인할 수 있었다.

Table 1 MSE in applying proposed method

	N			
	4	8	16	32
MSE	20.125	20.329	20.351	20.513

또한 제안한 방법의 부호화 효율을 예측하기 위해서 원 영상을 제일 기본적인 DPCM으로 부호화한 영상과 엔트로피 파워를 식 (8)을 이용하여 측정하여 비교하였다. 엔트로피 파워는 같은 엔트로피 환경을 가지는 환경에서 모든 신호의 빈도가 동일한 백색잡음의 출력으로 정의된다. 부호화에서는 측정 대상과 동일한 신호 개수를 가지고, 각 신호의 개수가 모두 동일할 때, 한 신호 종류의 세기로 볼 수 있다. 식 (8)에서 f_i 는 신

호 i 의 확률이다.

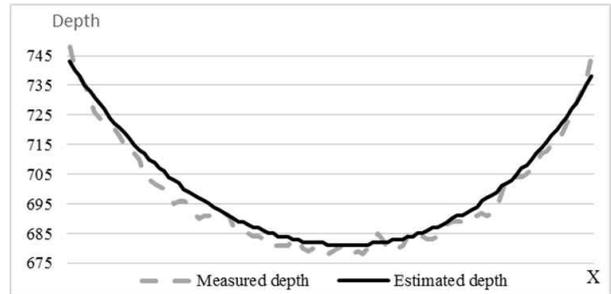


Fig. 5 Comparison between measured depth values and estimated depth values

$$N(X) = \frac{1}{2\pi e} e^{2h(X)} \tag{9}$$

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

이 때 부호화를 수행하지 않는 영상과 제안된 부호화 방법을 수행한 후의 엔트로피 파워를 Table 2와 같이 비교하였다. 이 때, 단순 DPCM을 통해 부호화를 한 경우에 비해 본 논문에서 제안된 방법을 통해 부호화를 한 경우 엔트로피 파워가 줄어들어 부호화 효율이 높아짐을 알 수 있다. 또한 블록의 크기가 커지면, 엔트로피 파워 측면에서 효율이 낮아지지만, 부호화 측면에서 부가 정보를 보내야하는 상황을 고려하면, 큰 블록 크기가 전체 효율 측면에서 좋을 수 있을 것이다. 즉, 블록 내 예측 효율이 약간 떨어지더라도, 각 블록에서 구의 인자 값을 전송하는 횟수가 줄어들 수 있다.

Table 2 Entropy power in applying proposed method

Coding method	DPCM	Proposed method			
		$N=4$	$N=8$	$N=16$	$N=32$
Entropy power	8.615	1.654	1.710	1.762	1.903

4. 결론

본 논문에서는 구형 객체에 대해 깊이 영상을 부호화 하는 방법을 제안하였다. 먼저 블록 내 화소를 이용하여 최소자승법을 적용하여 촬영된 객체와 제일 근접한 구를 찾았다. 그 후 찾아낸 구의 표면의 정보를 이용하여 깊이 값을 부호화 하였다. 그 결과 촬영된 구형 객체에 대해 부호화 효율이 뛰어나함을 확인할 수 있었다. 하지만 객체의 경계선 부분에서는 배경 영역의 깊이 값이 함께 포함되기 때문에 본 방법을 적용하기 어려운 부분이 있었다. 따라서 경계부분에 대해서는 적응적인 부호화 전략이 필요할 것으로 보인다. 본 논문에서 제안된 구면 표면의 정보를 모델링하여 깊이 영상을 부호화하는 새로운 깊이 영상 부호화 방법을 통해 깊이 영상 처리 분야뿐만 아니라 기존 영상 처리 방법[16-17]에도 널리 활용될 것으로 기대된다.

References

- [1] Lee D. S. and Kwon, S. K., "A Recognition Method for Moving Objects Using Depth and Color Information," *Journal of Korea Multimedia Society*, Vol. 19, No. 4, pp. 681-688, 2016.
- [2] Preis, J., Kessel, M., Werner, M., and Linnhoff-Popien, C., "Gait Recognition with Kinect," *Proceeding of the First Workshop on Kinect in Pervasive Computing*, pp. 1-4, 2012.
- [3] Kwon S. K. and Lee, D. S., "Correction of Perspective Distortion Image Using Depth Information," *Journal of Korea Multimedia Society*, Vol. 18, No. 2, pp. 106-112, 2015.
- [4] Lee D. S. and Kwon, S. K., "Improvement of Depth Video Coding by Plane Modeling," *Journal of the Korea Industrial Information Systems Research*, Vol. 21, No. 5, pp. 11-17, 2016.
- [5] Lee, D. S. and Kwon, S. K., "Touch Pen Using Depth Information," *Journal of Korea Multimedia Society*, Vol. 18 No. 11, pp. 1313-1318, 2015.
- [6] Kwon S. K. and Lee, D. S., "Recognition Method of Multiple Objects for Virtual Touch Using Depth Information," *Journal of the Korea Industrial Information Systems Research*, Vol. 21, No. 1, 27-34, 2016.
- [7] Lee D. S. and Kwon, S. K., "Video Event Control System by Recognition of Depth Touch," *Journal of the Korea Industrial Information Systems Research*, Vol. 21, No. 1, 35-42, 2016.
- [8] Suryanarayan, P., Subramanian, A., and Mandalapu, D., "Dynamic Hand Pose Recognition Using Depth Data," *Proceeding of 20th International Conference on Pattern Recognition*, 2010, pp. 3105-3108.
- [9] Zollhoefer, M., Martinek, M., Greiner, G., Stamminger, M., and Suessmuth, J., "Automatic Reconstruction of Personalized Avatars from 3D Face Scans," *Computer Animation and Virtual Worlds*, Vol. 22, No. 2-3, 2011, pp. 195 - 202.
- [10] Tong, J., Zhou, J., Liu, L., Pan, Z. and Yan, H., "Scanning 3D Full Human Bodies Using Kinects," *IEEE Transactions on Visualization and Computer Graphics*, Vol. 18, No. 4, 2012, pp. 643 - 650.
- [11] Jäger, F., "Simplified Depth Map Intra Coding With an Optional Depth Lookup Table," *Proceeding of 2012 International Conference on 3D Imaging*, 2012, pp. 1-4.
- [12] Liu, S., Lai, P., Tian, D., and Chen, C. W., "New Depth Coding Techniques with Utilization of Corresponding Video," *IEEE Transactions on Broadcasting*, Vol. 57, No. 2, 2011, pp. 551-561.
- [13] Oh, B. T., Lee, J., and Park, D. S., "Depth Map Coding Based on Synthesized View Distortion Function," *IEEE Journal of Selected Topics in Signal Processing*, Vol. 5, No. 7, 2011, pp. 1344-1352.
- [14] Park S. H. and Yoo, J. S., "Depth Map

Pre-Processing Using Gaussian Mixture Model and Mean Shift Filter,” Journal of the Korea Institute of Information and Communication Engineering, Vol. 15, No. 5, pp. 1155-1163, 2011.

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

[16] Kim, S. Y., Yoon, C. Y., and Yu, E. J., “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.

[17] Kim S. Y. and Lee, S. M., “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.



권순각 (Soon-Kak Kwon)

- 정회원
- 경북대학교 전자공학과 공학사
- KAIST 전기및전자공학과 공학석사
- KAIST 전기및전자공학과 공학박사
- 동의대학교 컴퓨터소프트웨어공학과 교수
- 관심분야 : 멀티미디어처리, IOT



이동석 (Dong-Seok Lee)

- 정회원
- 동의대학교 컴퓨터소프트웨어공학과 공학사
- 동의대학교 컴퓨터소프트웨어공학과 석사과정
- 관심분야 : 멀티미디어 신호처리, 영상 인식



박유현 (Yoo-Hyun Park)

- 정회원
- 부산대학교 전자계산학과 이학사
- 부산대학교 전자계산학과 이학석사
- 부산대학교 전자계산학과 이학박사
- 동의대학교 컴퓨터소프트웨어학과 부교수
- 관심분야 : 인터넷시스템, 클라우드 컴퓨팅