

깊이영상을 이용한 사람의 키 추정 방법

(A Method of Estimating the Human Height Using Depth Images)

김 흥 준¹⁾, 박 유 현²⁾, 권 순 각^{3)*}

(Kim Heung-Jun, Park Yoo-Hyun, and Kwon Soon-Kak)

요 약 본 논문에서는 깊이센서 카메라를 이용하여 획득된 깊이영상으로부터 사람의 키를 측정하는 방법을 제안한다. 깊이영상을 이용하면 색상영상에 비해 정확한 키의 측정이 가능하다. 본 논문에서는 수직 방향으로 사람의 중앙부를 검출하고, 중앙 위치에서 측정된 키의 값을 누적하는 방법을 제안한다. 실험을 통해 제안한 방법이 기존의 방법 보다 성능이 우수한 것을 확인하였다.

핵심주제어 : 깊이 카메라, 깊이 영상, 신장 측정

Abstract In This Paper, We Propose a Method to Estimate the Human Height from the Depth Image Obtained using a Depth Camera. Using the depth Image, Accurate Measurement for Human Height is Possible Compared with Color Image. This Paper Presents a Method to Detect the Center of a Person in the Vertical Direction and to Accumulate the Measured Height Values at the Center Position. Simulation Results Show that the Proposed Method has Better Performance than the Conventional Methods.

Key Words : Depth Camera, Depth Image, Measurement of Human Height

1. 서 론

최근 영상보안 시스템에서 신원 확인을 위한 신체 정보(biometry)로는 얼굴, 걸음걸이, 키 등이 있고 그 외 옷이나 모자 등의 정보를 추가로 이용하는 방법들이 연구되고 있으며, 그 중 키(height)는 대상의 신원을 확인하는데 중요한 정보가 될 수 있다[1-2]. 또한 매장에서 고객의 신체적 특징 중에 키를 획득하여 마케팅에 활용하

는 연구 또한 진행되고 있다[3-4]. 사람의 키를 추정하기 위해 색상 카메라를 사용하는 기존연구로 영상 내에서 특정 물체의 높이를 기준으로 키를 추정하는 방법[5], 실제 높이를 붉은색 마커로 표시하고, 영상에서 마커를 추출한 뒤 키를 추정하는 방법[6], 카메라의 위치정보와 소실점을 이용한 방법[7], 후보 객체들의 가로세로 비율, 크기 등의 조건을 이용한 방법[8], 워너 최적화 기법을 이용한 방법[2] 등이 있다. 색상 카메라를 사용하는 경우 원거리에서도 키 추정이 가능한 장점이 있으나 오차율이 높은 단점이 있다.

반면, 깊이센서 카메라를 이용하는 경우 깊이 정보(depth) 취득범위로 인해 원거리에서의 키 추정은 불가능한 단점이 있지만, mm 단위의 깊이정보를 제공하기 때문에 색상 카메라에 비해 정확한 키 추정이 가능한 장점이 있다. 깊이센서

* Corresponding Author : skkwon@deu.ac.kr

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

Manuscript received Feb, 6, 2017 / accepted Mar, 21, 2017

1) 동의대학교 컴퓨터소프트웨어공학과, 제 1저자
2) 동의대학교 컴퓨터소프트웨어공학과,
3) 동의대학교 컴퓨터소프트웨어공학과, 교신저자

카메라를 이용한 기존연구로 키넥트를 이용한 실내에서의 키 추정 방법[1]에서는 키넥트에서 제공하는 관절 점 정보와 깊이센서 카메라가 설치된 높이 정보를 이용하여 정지 상태에 있는 사람의 키를 추정 하였다. 그 결과 기존의 방법에 비해 키 추정 오차율이 크게 감소하였으나, 측정 대상의 머리 숙임이나 작은 움직임에도 센서가 반응하여 실시간 키 추정에 있어 오차를 발생키는 문제가 있었다.

본 논문에서는 깊이센서 카메라를 이용하여 정지 중인 사람의 키를 1cm, 보행중인 사람의 키를 2cm 오차 이내로 추정하는 방법을 제안한다. 키를 추정하기 위해서는 두 깊이영상에서 배경과 객체를 구분하는 이진화 단계가 선행되어야 한다. 먼저, 배경 깊이영상을 촬영하여 축적하고, 이후 깊이영상의 각 화소에 대하여 촬영된 깊이영상의 깊이 값과 배경영상의 깊이 값을 비교하여 이진화를 수행한다[9-10]. 본 논문에서는 사람의 키를 측정하는 것이 목적이므로 일정 크기 이상의 객체를 사람 객체로 가정한다. 사람 객체로부터 중심선을 찾은 뒤 객체 내에서 중심선의 양 끝에 해당하는 정수리와 발바닥 화소를 검출한다. 검출된 정수리와 발바닥 화소 사이의 실제 거리는 두 화소를 카메라를 원점으로 하는 3차원 좌표계로 변환하여 계산할 수 있다[11]. 계산된 정수리와 발바닥 화소 사이의 거리는 사람의 키에 해당한다. 이때 카메라와 사람의 거리에 따라 오차가 발생하므로 실험을 통해 보정 식을 구하고 보정 값을 더해 오차를 보상하여 키를 추정한다.

깊이센서 카메라는 초당 30프레임의 속도로 깊이영상을 획득하므로 실시간으로 추정되는 사람의 키는 사람의 작은 움직임에도 값이 변화한다. 본 논문에서는 이러한 문제를 해결하기 위해 실시간으로 추정된 키의 값을 누적하여 보행중인 사람에 대해서도 키 추정이 가능해졌으며 정확도 또한 향상되었다.

실험단계에서 정지 및 보행상태의 사람을 대상으로 키를 추정하고 기존의 방법과 비교하여 제안한 방법의 성능이 우수함을 확인하였다.

본 논문의 구성은 다음과 같다. 2장에서 실시간으로 획득되는 깊이영상으로부터 키를 추정하는 방법을 설명하고, 3장에서 누적된 키 추정 값

을 이용한 정확도 개선 방법을 설명한다. 그리고 4장에서는 제안한 방법으로 실험한 결과를 기존의 방법과 비교하여 성능을 평가하고 5장에서 결론을 맺는다.

2. 깊이정보를 이용한 키 추정 방법

2.1 배경영상 축적 및 이진화

배경과 객체를 구분하기 위해서 먼저 배경 깊이영상을 촬영하여 축적하고, 이후 깊이영상의 각 화소에 대하여 촬영된 깊이영상의 깊이 값과 배경영상의 깊이 값을 비교하여 이진화를 수행한다[9-10]. 본 논문에서는 배경 깊이영상을 100프레임 축적하여 각 화소의 최소 값을 배경영상으로 이용하였다. Fig. 1은 차영상과 이진화된 영상을 나타낸다. 이때 Fig. 1에서 이진화 영상 (b)에

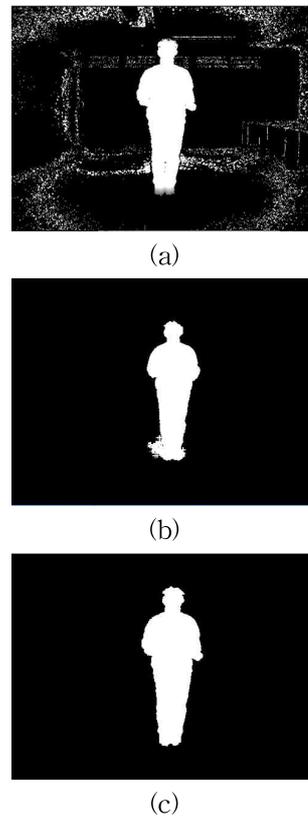


Fig. 1 Difference image and binary image. (a) Difference Image, (b) Binary Image and (c) improved Binary Image

서 발 주변의 바닥이 객체로 이진화되는 문제가 발생한다. 이 경우 키 추정 시 오차를 발생시킬 수 있기 때문에 차 영상에서 깊이 차이가 100mm 이하인 화소는 0mm로 설정하여 해결하였다. 본 논문에서는 사람의 키를 측정하는 것이 목적이므로 일정 크기이상의 객체를 사람 객체로 가정한다.

2.2 정수리 및 발바닥 화소 검출

Fig. 2는 깊이영상을 이용한 사람의 키 추정 방법을 순서도로 나타낸 것이다.

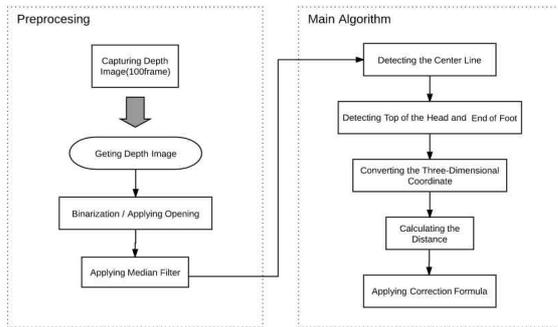


Fig. 2 Flowchart for a method of estimating the human height in using depth image.

중심선과 정수리를 찾는 방법을 Fig. 3에 나타내었다. 얼굴의 중심은 몸의 중심과 일치하므로 얼굴의 중심을 찾는다. 먼저 얼굴의 중심을 찾기 위한 y축 위치가 얼굴영역에 있도록 gap 값을 결정한다. 본 논문에서는 gap 값을 25로 설정하였다. (0, gap)위치에서 수평방향으로 탐색하여 배경영역에서 객체영역으로 바뀌는 점을 left, 객체영역에서 배경영역으로 바뀌는 점을 right로 찾고, 두 점의 중심을 중심점 mid로 한다. 머리카락 영역은 키에 해당하지 않고 잡음이 많이 발생하므로 정수리는 머리 끝에서 중심선을 따라 수직방향 아래로 3만쯤 아래에 위치한 (mid, 3) 화소로 보정한다. 발바닥은 객체 중심선을 따라 아래로 수직방향 끝에 해당하는 화소로 한다. 본 논문에서는 정수리와 발바닥 화소를 사용하기 때문에 깊이영상에 사람의 전신이 모두 포함되어 있어야

정확한 키 추정이 가능하다.

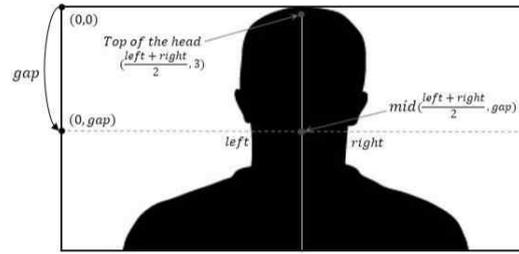


Fig. 3 Detection method of centerline and top of the head.

2.3 키 추정 및 보정

검출된 정수리와 발바닥 화소 사이의 실제 거리를 계산하기 위해서는 두 화소를 카메라를 원점으로 하는 3차원 좌표계로 변환한다.[11] 3차원 좌표로 변환된 정수리 화소를 $H(x_c, y_h, z_h)$, 발바닥 화소를 $F(x_c, y_f, z_f)$ 라 할 때 두 점사이의 실제거리는 아래 식 (1)을 통해 계산된다.

$$height = |y_f - y_h| \quad (1)$$

단, 깊이센서 카메라를 바닥과 수평하지 않게 설치한 경우 깊이정보에 해당하는 z축도 고려해야 하므로 아래 식 (2)를 통해 계산된다.

$$height = \sqrt{(y_f - y_h)^2 + (z_f - z_h)^2} \quad (2)$$

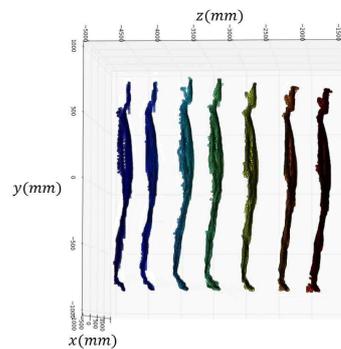


Fig. 4 Size of the object according to the distance from the camera.

그러나 계산된 결과는 카메라와 사람의 거리에 멀어질수록 따라 값이 커지는 오차가 발생한다. Fig. 4는 거리에 따른 객체의 크기변화 분석을 위해 깊이정보를 3차원 좌표로 변환하고 python의 matplotlib 모듈을 사용하여 3차원 공간상에 나타낸 것이다. Fig. 4에서 z축은 카메라와 사람사이의 거리를 의미하며 1500mm에서 3500mm까지는 객체가 일정하게 커지며 4000mm 부터는 다시 작아진다. 본 논문에서 사용한 깊이카메라는 Microsoft 사의 Kinect V2로 깊이 값 취득 범위는 8000mm이나 4000mm 부터는 오차가 많이 발생하므로 1500mm ~3900mm 사이를 키 추정 범위로 제한하였다. 키 추정 범위 내에서는 거리에 따라 일정하게 객체가 증가하므로 아래 보정 식 (3)을 적용한다.

$$height_c = (height - 0.05 \times depth_c + 185) / 10 \quad (3)$$

식 3은 고정된 크기의 물체를 카메라로부터 조금씩 먼 곳으로 옮기며 계산된 거리에 따른 오차로부터 도출되었다. 여기서 $height_c$ 는 보정된 키 추정 값, $height$ 는 보정전의 키 추정 값, $depth_c$ 는 카메라로부터 객체의 중심까지의 거리이다.

Fig. 5는 제안된 방법을 통해 정지상태의 사람

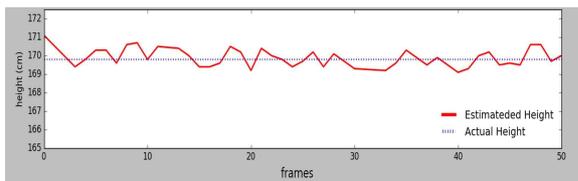


Fig. 5 Human height estimation result in stop state.

의 키를 50프레임 동안 추정하여 그래프로 나타낸 것이다. 50프레임 동안 추정 결과 최대 오차는 +1.3cm, 평균 오차율은 0.392cm이다.

3. 누적된 키 추정 값을 이용한 정확도 개선 방법

Fig. 5에서 각 프레임마다 정확한 키가 추정되지 않고, 약간의 측정 오차가 발생하였다. 오차가 발생한 이유는 깊이센서 카메라의 측정 오차와 사람의 미세한 움직임 때문이다. 이번 장에서는 프레임마다 키 추정 값을 누적하여 정확도를 개선하는 방법을 설명한다.

Fig. 6은 프레임마다 키 추정 값을 누적하고, 프레임마다 누적평균을 그래프에 나타낸 것이다. 이때 사람은 깊이센서 카메라의 촬영범위 밖에 위치해 있다가 0~30프레임 동안 촬영범위 안으로 이동하였고, 촬영범위 안으로 완전히 들어온 31프레임부터는 정지해 있었다. 사람이 촬영범위 안에 완전히 들어와 있지 않았던 0~30프레임에서는 올바른 키 추정이 되지 않았기 때문에 키 추정 값의 누적평균 또한 영향을 받아 오차가 크게 발생하였다.

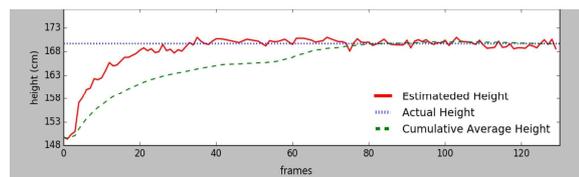


Fig. 6 Cumulative average graph(outside-stand).

Table 1 The change of the human height estimation value in the stop and walking state.

Object	Estimated Height (cm)								
	Stand			Walk			Walk louder		
	Min	Max	Amplitude	Min	Max	Amplitude	Min	Max	Amplitude
1	169.1	171.1	2.0	168.9	172.3	3.4	167.6	172.1	4.5
2	161.7	163.4	1.7	161.0	163.7	2.7	160.7	163.7	3.0
3	176.8	178.7	1.9	176.0	178.8	2.8	175.3	178.9	3.6
4	175.2	177.0	1.8	175.0	177.2	2.2	174.7	177.2	2.5
5	177.0	179.4	2.4	176.9	179.4	2.5	175.8	179.5	3.7

오차를 줄이기 위해서는 사람이 촬영범위 안에 완전히 들어와 있지 않았던 0~30프레임은 평균 계산에서 제외해야 한다. 이를 위해서는 사람이 촬영범위 안에 완전히 들어와 있는 경우와 그렇지 않은 경우를 구분할 수 있어야 한다.

Table 1에 사람이 촬영범위 안에 완전히 들어와 있는 상태에서 정지 및 보행상태의 사람을 촬영하고 키 추정 값의 변화 폭을 표로 나타내었다. Table 1에서 키 추정 값의 최대 변화 폭은 4.5cm이다. 따라서 누적된 키 추정 값에서 최댓값 보다 5cm 이상의 작은 값은 올바르지 않은 값으로 간주하고 누적평균 계산에서 제외한다.

Fig. 7은 키 추정 값의 최대 변화 폭을 이용하여 촬영범위 밖의 추정 값을 누적평균에서 제외한 누적평균 그래프이다.

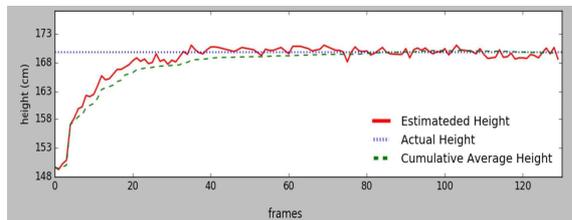


Fig. 7 Improved cumulative average graph (outside-stand).

4. 실험 결과

이번 장에서는 제안한 방법의 성능을 시험하고 기존의 방법과 성능을 비교해본다. 실험에 사용한 깊이센서 카메라는 Microsoft 사의 Kinect v2이며 바닥으로부터 120cm 높이에 바닥과 수평하게 설치하였다.

Table 2 Kinect v2 specifications.

Classification		Specification
Color	Resolution	1920 × 1080
	fps	30fps
Depth	Resolution	512 × 424
	fps	30fps
Depth Angle	Horizontal	70°
	Vertical	60°
Range of Depth		0.5 ~ 8.0m

Table 2에 Kinect v2의 성능을 표로 나타내었다. Fig. 8에 정지 및 보행상태(작은 보폭, 큰 보폭)의 사람을 대상으로 50frame 동안 키를 추정한 결과를 그래프로 나타내었다. 프레임이 누적될수록 실제 키에 가깝게 추정됨을 알 수 있다.

Table 3 The result of human height estimation using proposed method.

Object	Actual Height	Estimated Height (cm)			Error (cm)		
		Stand	Walking	Large Walking	Stand	Walking	Large Walking
1	174.7	175.1	175.9	176.2	+0.4	+1.2	+1.5
2	160.0	159.4	160.2	161.2	-0.6	+0.2	+1.2
3	169.8	169.1	170.0	170.4	-0.7	+0.2	+0.6
4	152.1	151.6	151.9	152.6	-0.7	-0.3	+0.5
5	172.6	173.0	173.5	173.9	+0.4	+0.9	+1.3
6	156.2	156.2	157.3	157.5	0	+1.1	+1.3
7	178.3	178.7	180.0	180.2	+0.4	+1.7	+1.9
8	161.2	161.6	162.4	162.7	+0.4	+1.2	+1.5
9	183.0	182.2	182.4	182.4	-0.8	-0.6	-0.6
10	177.5	178.2	178.6	179.2	+0.7	+1.1	+1.7
Average errors					0.51	0.85	1.21

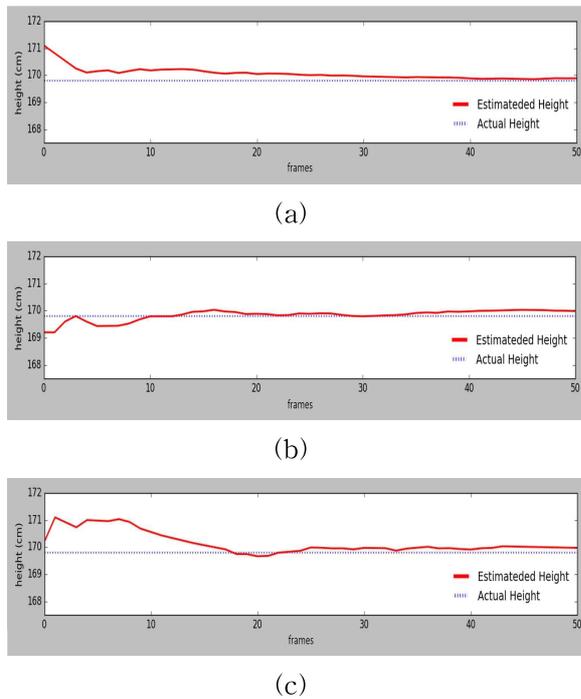


Fig. 8 The result of human height estimation using proposed method.
 (a) Stop state, (b) walking state (c) A large walking state

제안하는 방법의 성능을 평가하기 위해 5명을 대상으로 50프레임 동안 키를 추정하였다. Table 3은 제안한 방법으로 사람의 키를 추정한 결과이다. 정지 및 보행상태(작은 보폭, 큰 보폭)의 사람에 대해 각각 0.51cm, 0.85cm, 1.21cm의 평균 오차를 보였다. 깊이 카메라를 통한 키 추정은 물리적인 방식의 실제 키 측정과 다르게, 머리카락 부분도 키로 추정하기 때문에 헤어스타일의 영향을 받는다. 남성의 경우 올림머리, 여성의 경우 파마머리일 때 실제 키보다 크게 측정되었다.

Table 4 Comparison of average error of human height in stop state.

Method	Average Error (cm)
H. Choi & H. Shin[6]	4.92
S. Kim & J. Song[1]	1.12
Proposed Method	0.51

Table 4에 정지상태에서 키 추정 평균오차를 기존의 방법과 비교하여 나타내었다. 색상 카메라를 사용하는 H. Choi&H. Shin의 방법에 비해 큰 성능 차이를 보이고, 깊이 센서를 사용하는 S. Kim&J. Song의 방법에 비해서도 제안하는 방법의 성능이 더 좋았다. S. Kim&J. Song의 방법은 Microsoft의 Kinect SDK에서 제공하는 관절점 사이의 거리를 보정하여 키를 추정하기 때문에, 본 논문에서 제안한 정수리와 발바닥을 이용한 키 추정 방법에 비해 정확도가 낮다.

Table 5 Comparison of average error of human height in walking state.

Method	Average Error (cm)
S. Kim & J. Song[8]	4.28
Winer Optimization Method[2]	1.43
Proposed Method	0.85

Table 5에 보행상태에서 키 추정 평균오차를 기존의 방법과 비교하여 나타내었다. 보행상태의 사람에 대해서도 기존방법에 비해 제안하는 방법의 성능이 우수함을 확인하였다.

5. 결론

본 논문에서는 깊이영상을 이용해 사람의 키를 추정하는 방법을 제안하였다. 제안한 방법은 먼저, 사람 객체로 부터 중심선을 찾은 뒤 객체 내에서 중심선의 양 끝에 해당하는 정수리와 발바닥 화소를 검출한다. 검출된 정수리와 발바닥 화소 사이의 실제 거리를 이용하여 키를 추정한다. 이후 카메라와 사람의 거리에 따른 오차를 보정한다. 실시간으로 추정되는 사람의 키는 사람의 작은 움직임에도 값이 변화하는 문제가 있다. 이러한 문제를 해결하기 위해 실시간으로 추정된 키의 누적 평균을 이용하였다. 그 결과 보행중인 사람에 대해서도 키 추정이 가능해졌으며 정확도

또한 크게 향상되었다. 실험을 통해 기존의 색상 영상과 깊이영상을 이용한 키 추정 방법에 비해 성능이 우수함을 확인하였다.

본 논문에서 제안한 방법은 대상의 신원을 파악하는 영상보안 분야, 고객의 신체적 특징을 파악하는 마케팅 분야, 다양한 신호처리 분야[12, 13]에 활용될 수 있을 것이다.

References

- [1] Momeni-k, M., Diamantas, S. C., Ruggiero, F. and Siciliano, B., "Height Estimation from a Single Camera View," In VISAPP(1), pp. 358-364, 2012.
- [2] Jeon, S. H., Song, J. K., Park, J. S. and Yoon, B. W., "Video Based Pedestrian Height Estimation Using Winer Optimization," Journal of Korea Multimedia Society, Vol. 19, No. 2, pp. 264-270, 2016.
- [3] BenAbdelkader, C., and Yacoob, Y., "Statistical Body Height Estimation from a Single Image," Proc. of Int. Conf. on 8th Automatic Face & Gesture Recognition, pp. 1-7, 2008.
- [4] Park, S. W., Kim, T. E. and Choi, J. S., "Robust Estimation of Heights of Moving People Using a Single Camera," Proceedings of the International Conference on IT Convergence and Security 2011, Vol. 120, pp. 389-405, 2011.
- [5] Kim, S. M., Song, J. K., Yoon, B. W. and Park, J. S., "Height Estimation of Pedestrian Based on Image," Journal of Korea Institute of Electronic Communication Sciences, Vol. 9, No. 9, pp. 1035-1042, 2014.
- [6] Choi, H. J. and Shin, H. S., "Modeling for Webcam-Based Height Measurement System," Journal of the Korean Institute of Electrical Engineers, pp. 1419-1420, 2015.
- [7] Jeon, T. J., Kim, J. S., Lee, K. J., Bae, H. B. and Lee, S. Y., "Human Height Estimation Algorithm Using the Single RGB Camera," Journal of the Institute of Electronics Engineers of Korea, pp. 477-487, 2014.
- [8] Kim, S. M., Song, J. K., Yoon, B. W. and Park, J. S., "Height Estimation Using Kinect in the Indoor," Journal of the Korea Institute of Electronic Communication Sciences, Vol. 9, No. 3, pp. 343-350, 2014.
- [9] Kim, H. S. and Kwon, S. K., "Presentation Method Using Depth Information," Journal of Korean Society Of Broad Engineers, Vol. 18, No. 3, pp. 409-415, 2013.
- [10] Lee, D. S. and Kwon, S. K., "Recognition Method of Multiple Objects for Virtual Touch Using Depth Information," Journal of Korea Industrial Information Systems Research, Vol. 21, No. 1, pp. 27-34, 2016.
- [11] Lee, D. S. and Kwon, S. K., "Correction of Perspective Distortion Image Using Depth Information," Journal of Korea Multimedia Society, Vol. 18, No. 2, pp. 106-112, 2015.
- [12] 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," Journal of Internet Electronic Commerce Research, Vol. 14, No. 4, pp. 313-333, 2014.
- [13] Kim, S. Y. and Lee, S. M., "Implementation of an Image Board Remote Control System Using PDA Based on Embedded Linux in Wireless Internet," Journal of Information Systems, Vol. 17, No. 1, pp. 155-171, 2008.



김 흥 준 (Kim Heung-Jun)

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



박 유 현 (Park Yoo-Hyun)

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



권 순 각 (Kwon Soon-Kak)

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