

키넥트를 이용한 손 영역 검출의 정확도 개선

김희애¹ · 이창우^{2*}

Detection Accuracy Improvement of Hand Region using Kinect

Heeae Kim¹ · Chang Woo Lee^{2*}

¹Department of Computer Information Engineering, Kunsan National University, Kunsan 573-701, Korea

²Department of Computer Information Engineering, Kunsan National University, Kunsan 573-701, Korea

요 약

최근 마이크로소프트사의 키넥트(Kinect)를 이용하여 객체를 추적하거나 인식하는 기술에 관한 연구가 활발히 진행되고 있다. 이와 같은 환경에서 손을 검출하고 추적하는 기술은 휴먼-인터페이스 개발의 가장 기본적인 기술이다. 본 논문에서는 키넥트를 이용하여 단색으로 제한되지 않은 환경에서 검출된 손의 경계를 개선하는 방법을 제안한다. 제안된 방법은 키넥트 센서의 깊이영상(Depth Image)에서 추출된 손의 영역과 칼라영상에서 피부색(Skin Color)의 검출 결과를 결합하여 손 영역의 정확도를 개선한다. 제안된 방법으로 손 영역의 검출을 수행한 결과 깊이 영상만 사용한 방법 보다 손 영역을 검출하는 정확도를 높일 수 있음을 보였다. 제안된 방법이 수화나 제스처 인식 시스템에 적용된다면 정확도 개선에 많은 기여를 하리라 기대한다.

ABSTRACT

Recently, the researches of object tracking and recognition using Microsoft's Kinect are being actively studied. In this environment human hand detection and tracking is the most basic technique for human computer interaction. This paper proposes a method of improving the accuracy of the detected hand region's boundary in the cluttered background. To do this, we combine the hand detection results using the skin color with the extracted depth image from Kinect. From the experimental results, we show that the proposed method increase the accuracy of the hand region detection than the method of detecting a hand region with a depth image only. If the proposed method is applied to the sign language or gesture recognition system it is expected to contribute much to accuracy improvement.

키워드 : 키넥트, 깊이영상, 손 영역 검출, 피부색

Key word : Kinect, Depth Image, Hand Region Detection, Skin Color

접수일자 : 2014. 09. 25 심사완료일자 : 2014. 10. 15 게재확정일자 : 2014. 10. 29

* **Corresponding Author** Chang Woo Lee(E-mail:leecw@kunsan.ac.kr, Tel:+82-63-469-4861

Department of Computer Information Engineering, Kunsan National University, Kunsan 573-701, Korea

Open Access <http://dx.doi.org/10.6109/jkiice.2014.18.11.2727>

print ISSN: 2234-4772 online ISSN: 2288-4165

©This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
Copyright © The Korea Institute of Information and Communication Engineering.

I. 서 론

최근 인간과 컴퓨터의 상호작용의 한 수단으로 손 영역의 검출과 인식에 관한 연구가 활발히 진행되고 있다 [1, 2]. 입력영상으로부터 손 영역의 추출을 위한 소프트웨어적인 수단으로 피부색모델을 이용한 검출 방법 [3]이나, 2차원, 혹은 3차원 가상모델[4], 그리고 특징검출에 의한 손 영역 검출방법[5]이 제안되었다. 또한 하드웨어적 수단으로는 마이크로소프트사의 키넥트 센서 [6-9]나, 스테레오카메라를 이용한 검출 방법[10-12], 등이 제안되었다.

키넥트는 2010년 마이크로소프트사에서 Xbox 360의 입력장치로 시장에 출시한 제품으로 깊이 맵(Depth Map)을 획득하기 위한 IR 프로젝터와 센서, 그리고 컬러영상 획득을 위한 RGB 센서로 구성되어있다[6, 11, 13]. 키넥트는 게임 인터페이스로서의 사용보다 컴퓨터 비전 분야의 수단으로 연구가 활발하게 진행되었으며, 특히 센서의 보정(Calibration)에 관한 연구[7], 모션인식에 관한 연구[8], 삼차원 복원에 관한 연구[9] 등이 활발히 진행되고 있다.

현재까지 진행된 손 영역 검출 방법으로 키넥트를 이용한 연구들에서는 손 영역이 센서로부터 최근접 지역에서 최대 크기의 영역이라는 가정으로 손 영역을 검출하고 있으며, 일부 방법에서는 피부색모델과의 중첩영역을 손으로 검출하는 연구결과가 발표되었다[1]. 제스처 인식시스템이나, 수화 인식시스템, 등에서 보다 정확한 손 영역의 검출이 선행된다면 시스템의 성능 향상 효과를 기대할 수 있지만, 기존의 손 영역을 검출하는 연구들은 손 영역의 경계에 대한 후처리과정에 관심을 보이지 않았다.

본 논문에서는 키넥트 센서의 깊이영상과 컬러영상의 피부색 정보를 결합하여 검출된 손 영역의 정확한 경계 도출을 위한 방법을 제안한다. 제안된 논문의 구성은 다음과 같다. 먼저 2장에서 키넥트의 적외선 센서와 컬러영상 센서로부터 획득한 영상을 이용하여 손 영역을 검출하는 과정을 자세히 설명하고, 보다 정확한 손 영역 검출을 위한 영역확장 기법에 대하여 설명한다. 3장에서는 제안된 방법의 우수성을 증명하기 위해 배경이 단순한 경우와 복잡한 경우에 대한 실험과정을 설명하며, 마지막 4장에서 결론을 맺는다.

II. 본 론

마이크로소프트사의 키넥트는 깊이정보를 획득하기 위해 레이저 광선을 방출하는 IR 프로젝터와 물체에 반사되어 돌아오는 레이저 광선을 획득하는 IR 센서, 그리고 컬러영상을 획득하기 위한 RGB센서로 구성되어 있다[6]. 본 논문에서는 그림 1에서 보는 바와 같이 키넥트 센서의 입력을 이용하여 손 영역을 검출하고, 보다 정확한 손 영역을 검출하기 위한 영역 확장 방법을 제안한다.

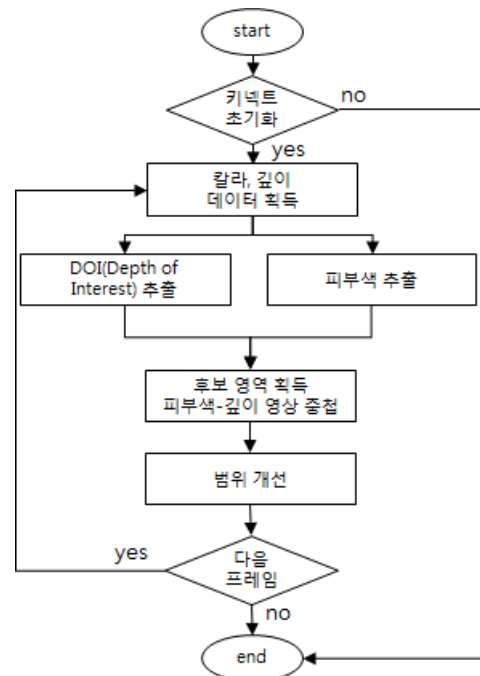


그림 1. 제안된 방법의 흐름도
Fig. 1 Flowchart of the Proposed Method

제안된 방법은 먼저 키넥트 장치로부터 컬러영상과 깊이정보를 포함한 깊이영상을 획득한다. 획득된 깊이 영상에서 손 영역의 후보영역을 검출하기 위해 센서에 근접한 최대 영역이 손이라는 가정으로 깊이영상을 이진화한다. 또한 컬러영상으로부터 피부색 모델을 이용하여 손 영역의 후보영역을 검출한다. 이렇게 형성된 두 개의 초기 손 후보 영상을 정합하여 손 영역을 검출한다. 마지막 단계에서 보다 정확한 손 영역의 검출을 위한 영역확장을 수행한다.

그림 2에서 보는 바와 같이 키넥트 센서에서 획득된 깊이영상과 칼라영상은 각각 센서들의 배치관계로 인해 그 화각이나 해상도 면에서 차이를 보인다. 그림 2의 (a)와 (b)는 각각 근거리와 원거리의 객체의 깊이 영상을 칼라영상에 중첩(overlap)한 영상으로 두 영상간의 차이를 보여준다.

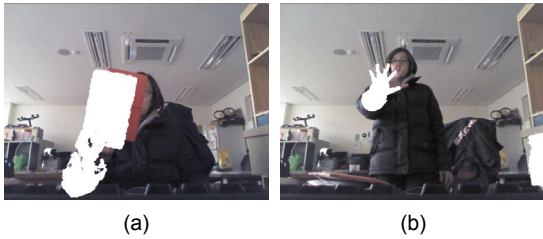


그림 2. 깊이영상과 칼라영상의 위상차이 (a) 근거리 (b) 원거리
Fig. 2 Discrepancy between Depth and Color Images (a) Near (b) Far

일반적으로 두 영상을 1대1로 정합할 수 없기 때문에 두 영상의 보정과정을 거치게 된다. 그림 3에서는 이를 보정하기 위해 OpenNI[14]를 사용하여 영상의 보정을 수행한다.

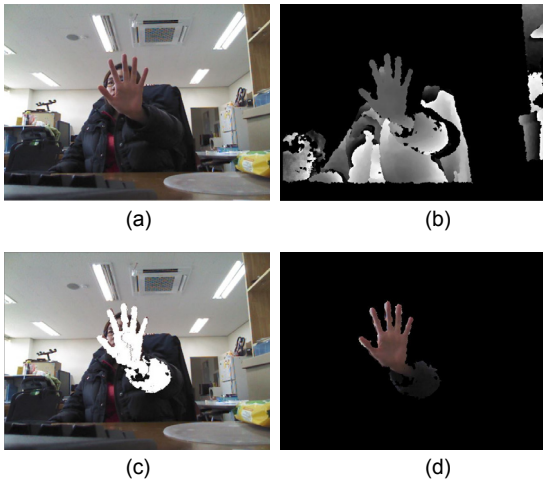


그림 3. 보정 전후(a) 칼라영상 (b) 깊이영상 (c) 깊이영상이 중첩된 칼라영상 (d) 칼라영상과 깊이 영상의 차이
Fig. 3 Before and After Calibration (a) Color Image (b) Depth Image (c) Color Image Overlapping Depth Image (d) Discrepancy between Depth and Color Image

그림 3의 (a)와 (b)는 각각 키넥트 센서를 이용한 깊이영상과 칼라영상의 예이고, (c)와 (d)는 두 영상의 보정 이후의 정합영상이다. 특히 그림 3의 (d)는 정합이후의 손 영역의 검출 정확도가 낮다는 예시으로써, 깊이센서로 손을 인식할 때 가장 가깝고 큰 물체가 손이라고 가정하고 배경을 제거하면 소매 등이 포함된 거친 형태의 손 영역을 얻을 수 있다. 이 영역에 피부색을 검출한 결과를 결합하면 보다 정확한 손 영역을 얻게 된다.

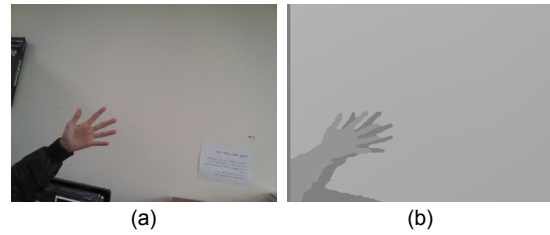


그림 4. 배경의 복잡도가 낮은 입력영상의 예 (a) 입력영상 (b) 깊이영상
Fig. 4 Example of an Input Image with Simple Background (a) Input Image (b) Depth Image

후보 영역 영상은 단순히 깊이영상과 피부색영상을 중첩한 영상이므로 피부색 영역임에도 포함되지 않는 부분이 존재할 수 있다. 본 논문에서는 손 영역 검출의 정확도를 높이기 위하여 깊이영상 또는 피부색영상의 데이터가 유효한 영역의 모든 픽셀을 대상으로 수행한다. 대상 픽셀의 주변 5*5영역에서 피부색 픽셀이 하나 이상 존재하고, 이 영역의 색 평균이 피부색 모델의 범위 이내라면 해당 픽셀이 피부색일 가능성이 높다고 판단한다.

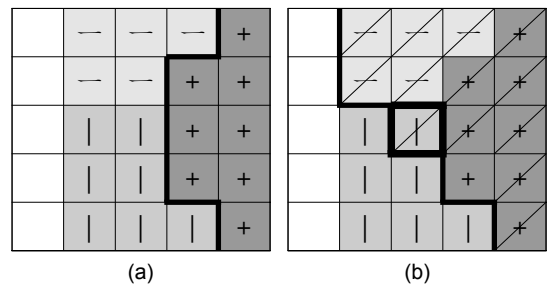


그림 5. 검출 영역의 정확도 개선 개념도(a) 개선 전 영역 (b) 개선 후 영역
Fig. 5 Overview of the Improvement of Detection Accuracy (a) Before Region Growing (b) After Region Growing

그림 5는 개선할 영역의 예시를 나타내는 개념도이다. ‘-’로 표시된 영역은 피부색영역, ‘!’는 깊이영역, ‘+’는 중첩 영역을 각각 도시한 그림이다. 제안된 방법을 적용한 후에는 그림 5의 (b)에 ‘/’으로 표시된 것처럼 포함되지 않았던 피부색 영역이 추가된 결과영상을 도시한 것이다.

III. 실험결과

실험을 위해 사용된 시스템은 Intel Core i7 3.4GHz, 16G RAM을 사용하였고, Windows 8.1 운영체제와 Visual C++ 2012 컴파일러 환경에서 실험하였다. 마이크로소프트사의 PC버전 키넥트 장치를 이용하여 영상을 획득하였으며 OpenCV 2.4.9버전의 라이브러리[15]를 이용하였다. 키넥트 센서의 근거리모드(Near mode)를 사용하고 깊이 영상에서는 미리 설정한 반경 이내의 객체를 검출하였다. 검출된 결과는 RGB 카메라와 깊이 센서의 위상 차이에 의해 서로 어긋나 있으므로, OpenNI 라이브러리를 이용하여 보정 과정을 거쳤다. 피부색 모델의 색상 범위는 조명과 주변의 영향을 많이 받기 때문에 그 범위를 동적으로 변경할 수 있도록 하였다.

그림 6의 (a)는 단순한 배경으로부터 깊이 범위 0.4~0.7m 이내의 물체를 검출한 결과이다. 손 이외에도 소매부분이 포함된 것을 볼 수 있다. 그림 6의 (b)는 피부색 범위를 $75 < C_b < 131$, $135 < C_r < 138$ 로 설정한 후 손 영역을 검출한 결과이다.

본 실험의 정확도를 판별하기 위하여, 기준영상(Ground Truth Image)을 수작업으로 준비하였다. 기준영상에서 손 영역의 픽셀 수와 제안된 방법에 의한 손 영역에서 검출된 픽셀 수를 구한 후 이 둘의 비율을 계산하였다. 결과는 중첩되는 비율과 중첩되지 않는 범위의 비율로 나뉘는데, 정확하게 검출된 픽셀의 수를 참 긍정(True Positive, TP)이라하고, 중첩되지 않는 영역에서 기준영상의 손 영역임에도 불구하고 실험결과에서 손 영역으로 검출되지 못한 영역의 픽셀 수를 거짓 부정(False Negative, FN)의 결과를 계산하였다. 이와 동시에 기준영상에서 손 영역이 아닌데도 불구하고 실험결과에서 손 영역으로 검출된 영역의 픽셀 수인 거짓 긍정(False Positive, FP)을 계산하였다. 또한, 기준영상

에서 손영역이 아닌 영역에 대해 실험 결과 영상에서도 손 영역이 아니라고 검출된 참 부정(True Negative, TN)의 값을 계산하였다. 이 수치들을 이용하여 제안된 방법의 검출의 정확도를 계산하기 위해 수식 (1)과 (2)에서와 같이 정확률(Precision Rate)과 재현율(Recall Rate)을 계산하였다.

$$\text{정확률} = \frac{TP}{TP + FP} \quad (1)$$

$$\text{재현율} = \frac{TP}{TP + FN} \quad (2)$$



그림 6. 깊이영상과 피부색영상으로부터 손 검출결과 (a) 검출된 깊이영상 (b)피부색검출영상
Fig. 6 Immediate Result of Hand Detection from Depth and Color Images (a) Detected Depth Image (b) Result of Skin-color Detection

그림 6의 각 결과와 이를 기준영상의 손 영역을 비교한 결과, 그림 6의 (a)는 참 긍정이 전체의 97.5%에 육박하나, 거짓 긍정이 127.0%로 손 이외의 영역이 두 배 이상 검출된 결과를 보였다.

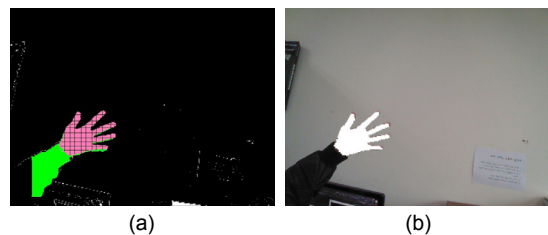


그림 7. 제안된 방법의 실험결과 (a) 중첩영상 (피부색 + 깊이) (b) 제안된 방법의 결과
Fig. 7 Experimental Result of the Proposed Method (a) Overlapped Image (Skin Color Image + Depth Image) (b) Final Result

그림 6의 (b)는 참 긍정이 86.3%이었으나, 제한된 배경임에도 불구하고 빛의 영향으로 배경에서 손이 아닌 영역인 거짓 긍정이 74.3% 이상 검출되었다. 따라서 깊이영상과 칼라영상 각각에서 손영역을 검출하는 방법은 많은 제약과 오류를 유발시켰다. 그러나 제안된 방법은 그림 7에서 보는 바와 같이 깊이 영상과 칼라영상을 결합하여 참 긍정이 90.2%, 거짓 긍정이 6%만 검출되었다. 그림 8은 다양한 배경환경에서 제안된 방법의 실험결과를 도시하였다.

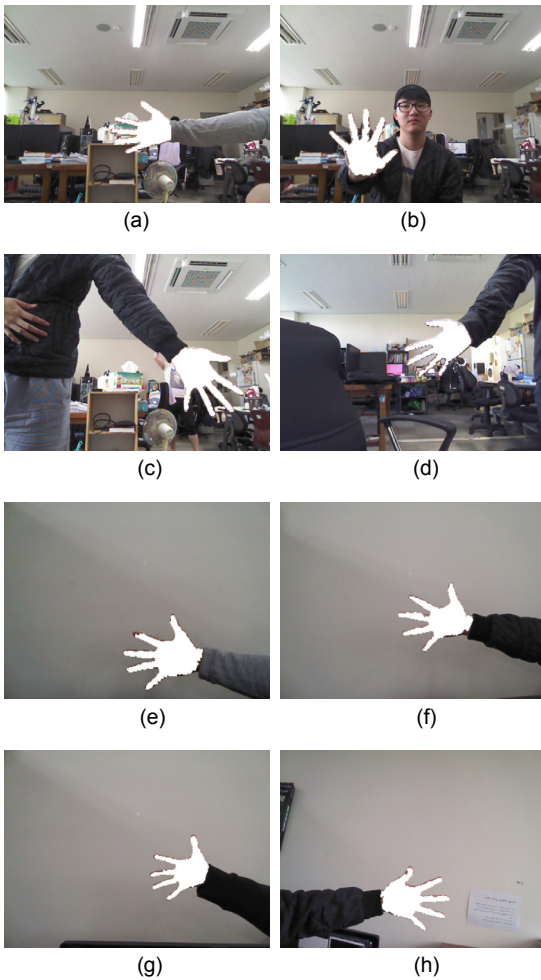


그림 8. 다양한 배경에서 실험결과(%) (a) 92.0 (b) 86.9 (c) 89.9 (d) 86.3 (e) 89.8 (f) 89.9 (g) 91.7 (h) 88.9
Fig. 8 Experimental Result in Simple and Cluttered Background (a) 92.0 (b) 86.9 (c) 89.9 (d) 86.3 (e) 89.8 (f) 89.9 (g) 91.7 (h) 88.9

그림 8의 (a)-(d)까지는 복잡한 배경에서 손 영역을 검출한 결과이고, (e)-(g)까지는 단순한 배경에서 손 영역을 확장하여 검출한 결과이다. 한다. 그림 8의 각 숫자는 참 긍정을 표시한 것이다.

제안된 방법의 일반성을 확보하기 위해 단순한 배경에서 획득한 영상 250장과 복잡한 배경에서 획득한 실험영상 250장을 대상으로 정확률과 재현율을 측정하였다. 표 1에서는 제안된 방법이 다양한 환경에서 실험한 결과를 도시한다. 깊이 정보만을 사용한 경우, 피부색만을 이용한 경우와 제안된 방법에 의한 손 영역 검출의 정확도를 비교한 것이다. 표 1에서 보는 바와 같이 참긍정의 비율은 깊이영상에서 제일 높게 나왔다. 그러나 손이 아닌 영역을 검출한 거짓 긍정률이 상대적으로 높다는 사실을 인지한다면 좋은 결과라 할 수 없다. 제안된 방법이 정확률면에서 많은 향상을 보였다.

표 1. 정확률과 재현율 평균(%)

Table. 1 Average of Precision and Recall Rate (%)

		정확률	재현율
단순배경	피부색	66	87.3
	깊이	41.8	97.4
	제안된 방법	94.1	88.4
복잡배경	피부색	23	83.4
	깊이	30.6	97.9
	제안된 방법	95.1	84.6

IV. 결 론

본 논문에서 키넥트의 깊이 센서로 인식한 손의 정확도를 피부색 정보를 결합하여 향상시킬 수 있는 방법을 제안하였다. 깊이영상의 부족한 정보를 컬러영상에서 보충하고, 손 영역의 정확도를 높이기 위한 영역 확장 방법을 제시하였다. 제안된 방법의 다양한 실험 결과는 단순한 배경에서 94.1%, 복잡한 배경에서 95.1%의 정확률을 보였다. 제안된 연구의 결과를 수화인식시스템이나 손가락의 모양이 필요한 제스처 인식시스템에 적용한다면 해당 시스템의 성능을 크게 향상시킬 수 있다. 향후과제로 에지 정보를 결합하여 에지 방향으로 에지의 경계영역까지 손 영역을 확장하는 방법을 고안할 계획이다.

REFERENCES

- [1] H. Park, J. Choi, J. Park, K. Moon, "A Study on Hand Region Detection for Kinect-Based Hand Shape Recognition," *Journal of Broadcasting Engineering*, vol. 18, no. 3, pp. 393-400, May 2013.
- [2] J. Y. Chang, M. W. Ryu, S. C. Park, "Technology Trends of Range Image Based Gesture Recognition," *Electronics and Telecommunications Trends*, vol. 29, no. 1, pp. 11-55, Feb. 2014.
- [3] S. Bilal, R. Akmeiliawati, M. J. E. Salami, A. A. Shafie, and E. M. Bouhabba, "A hybrid method using haar-like and skin-color algorithm for hand posture detection, recognition and tracking," in *Proceeding of the International Conference on Mechatronics and Automation*, pp. 934-939, 2010.
- [4] S. Park, S. Yu, J. Kim, S. Kim, and S. Lee, "3D hand tracking using Kalman filter in depth space," *EURASIP Journal on Advances in Signal Processing*, vol. 36, pp. 1-18, 2012.
- [5] J. Kim, J. Baek, E. Kim, "A part-based rotational invariant hand detection," in *Proceeding of the 2013 International Conference on Fuzzy Theory and Its Applications*, pp. 127-129, 2013.
- [6] Wikipedia Online Documentation [Internet]. Available: <http://en.wikipedia.org/wiki/Kinect>.
- [7] K. Berger, K. Ruhl, C. Brummer, Y. Schroder, A. Scholz, and M. Magnor, "Markerless motion capture using multiple color-depth sensors," in *Proceeding of the Vision, Modeling and Visualization(VMV) 2001*, pp. 317-324, 2011.
- [8] T. Dutta, "Evaluation of the Kinect sensor for 3-d kinematic measurement in the workplace," *Applied Ergonomics*, 2011.
- [9] T. Stoyanov, A. Louloudi, H. Andreasson, and A. J. Lilienthal, "Comparative evaluation of range sensor accuracy in indoor environments," Proceedings of the European Conference on Mobile Robots (ECMR), Sweden, pp. 7-9, September 2011.
- [10] J. P. Wachs, M. Kölsch, H. Stern, and Y. Edan, "Vision-based hand-gesture applications," *Communications of the ACM*, vol. 54, pp. 60-71, 2011.
- [11] A. Conci, E. Nunes, J. J. Pantrigo, A. Sanchez, "Comparing color and texture-based algorithms for human skin detection", *Computer Interaction*, vol. 5, pp. 168-173, 2008.
- [12] V. A. Oliveira, A. Conci, "Skin Detection using HSV color space," In: *Sibgrapi*, No. 12, 2009.
- [13] K. Khoshelhan and S. O. Elberink, "Accuracy and resolution of kinect depth data for indoor mapping applications," *Sensor*, vol. 12, no. 2, pp. 1437-1454, Feb. 2012.
- [14] OpenNI Online Documentation [Internet]. Available: <http://www.openni.org/>
- [15] OpenCV Online Documentation [Internet]. Available: <http://opencv.org/>



김희애(Heeae Kim)

2011년 ~ 현재 : 국립군산대학교 컴퓨터정보공학과 학부과정생
※관심분야: 영상처리



이창우(Chang Woo Lee)

2004년-- 현재 : 군산대학교 컴퓨터정보공학과 부교수
※관심분야 : 패턴인식, 컴퓨터비전, 상황인식