

동심원 추적 알고리즘을 사용한 손가락 동작 인식

황동현 · 장경식*

Finger-Gesture Recognition Using Concentric-Circle Tracing Algorithm

Dong-Hyun Hwang · Kyung-Sik Jang*

School of Computer Science and Engineering, Korea University of Technology and Education, Cheonan 330-708, Korea

요 약

본 논문에서는 저가의 웹 카메라를 사용하여 영상을 입력받아 손 부분 영상을 추출한 후 동심원 추적 알고리즘을 사용하여 손가락 동작을 인식하는 기법을 제안한다. 제안하는 알고리즘은 기존의 인터페이스처럼 손에 별도의 센서를 부착하지 않음으로 신체에 불편함을 주지 않는다. 또한 저가인 웹 카메라를 사용해서 비용적인 측면에서 활용성을 증가 시켰다. 동심원 추적 알고리즘을 사용하여 펴진 손가락의 개수뿐 아니라, 손가락 접힘 여부 정보를 효율적으로 추출할 수 있다. 본 논문에서 제안한 알고리즘을 사용하여, 손가락 동작을 평균 95.48%의 정확도로 인식할 수 있음을 확인했으며, 손을 사용한 HCI 및 원격 제어 명령어 입력수단으로 활용가능성을 확인하였다.

ABSTRACT

In this paper, we propose a novel algorithm, Concentric-Circle Tracing algorithm, which recognizes finger's shape and counts the number of fingers of hand using low-cost web-camera. We improve algorithm's usability by using low-price web-camera and also enhance user's comfortability by not using a additional marker or sensor. As well as counting the number of fingers, it is possible to extract finger's shape information whether finger is straight or folded, efficiently. The experimental result shows that the finger gesture can be recognized with an average accuracy of 95.48%. It is confirmed that the hand-gesture is an useful method for HCI input and remote control command.

키워드 : 손가락 동작 인식, 손 추적, 인간-컴퓨터 상호작용, 컴퓨터 비전

Key word : Finger-gesture Recognition, Hand Tracing, Human-Computer Interaction, Computer Vision

Received 06 August 2015, Revised 10 August 2015, Accepted 25 August 2015

* Corresponding Author Kyung-Sik Jang(E-mail:ksjang@koreatech.ac.kr, Tel:+82-41-560-1352)

School of Computer Science and Engineering, Korea University of Technology and Education, Cheonan 330-708, Korea

Open Access <http://dx.doi.org/10.6109/jkiice.2015.19.12.2956>

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. 서론

최근 인간-컴퓨터 상호작용 분야는 다양한 제스처 인식에 관한 연구가 진행되고 있다. 그중 손은 인간이 가지고 있는 생체 중 가장 많이 사용되는 생체로서 직관적으로 물체를 가리키고 조작하는 것이 가능하다. 이에 따라 손 제스처 인식은 키보드와 마우스와 같은 입력 인터페이스를 대체할 새로운 인터페이스로 각광 받고 있다. 그중 비전 기반 인식 시스템은 손에 별도의 센서를 부착하지 않기 때문에 손의 자연스러운 움직임에 지장을 주지 않고 조작 시 사용자에게 불편함을 주지 않는다는 이점이 있다[1].

최근에는 비전 기반으로 손의 제스처를 인식하는 연구가 활발하게 진행되어지고 있다. 먼저 COG(Center of Gravity)를 이용하여 손가락의 개수를 카운트하는 알고리즘은 빠르고 정확한 인식률을 가지고 있지만 인식할 수 있는 제스처가 제한되어 있다[2, 3]. 그 다음으로는 Hu Moments, Fourier Descriptor(FD)와 Convex Hull과 Convexity Defects를 이용하여 손가락의 개수를 카운트하는 알고리즘이 있다[4].

이 알고리즘은 손가락의 끝점을 정확히 찾을 수 있기에 이를 이용하여 다양한 제스처를 인식하는 것이 가능하다. 하지만 인식률이 앞서의 알고리즘에 비해 떨어진다. Centroid를 기점으로 손의 외곽선(Contour)을 구성하는 포인트와의 거리를 구하여 일정 Threshold를 넘는 값을 이용하여 손가락의 개수를 구하는 알고리즘이 있다[5]. 이는 대략적인 손가락의 위치를 알 수 있기 같은 손가락 개수여도 다른 포즈에 대한 인식이 가능하고 반만 접힌 손가락에 대해서도 인식을 하여 다양한 제스처를 지원하는 장점이 있지만 인식 속도가 느리다.

본 논문에서는 가격이 상대적으로 저가인 웹 카메라를 사용하고 동심원 추적 알고리즘을 사용하여 손가락의 동작을 인식 할 수 있는 방법을 제안한다. 제안하는 동심원 추적 알고리즘은 COG와 동심원을 이용하여 다른 제스처 인식 알고리즘에 비하여 높은 인식률 및 속도를 가지고 있다. 또한 기존의 COG기반 알고리즘이 제한된 제스처만을 인식할 수 있었던 반면 제안하는 알고리즘에서는 두 개의 동심원을 이용하여 손가락을 반만 접는 제스처에 대해서도 인식이 가능 하도록 하여 다양한 제스처를 인식할 수 있도록 하였다.

II. 동심원 추적 알고리즘

본 논문에서 제안하는 알고리즘은 세 부분으로 구성되어 있다. 카메라로부터 영상을 입력 받아 손 인식에 적합하게 영상을 처리 후 손 영역을 추출하는 전처리부분, 추출된 ROI(Region of Interest)에 대하여 COG 및 손바닥의 반지름을 구하고 Convex Hull을 이용하여 최장 손가락의 길이 계산 후, Scale이 변환함에 따른 동심원 크기 보상 작업 및 동심원 추적 알고리즘을 수행하는 부분, 마지막으로 완전히 펼쳐진 손가락의 개수, 반 정도 접힌 손가락의 개수, 손가락이 두 개인 경우 두 손가락 사이의 각도 등의 제스처의 특징을 얻어내는 부분이다. 그림 1은 본 논문에서 제안하는 시스템의 전체적인 구조도이다.

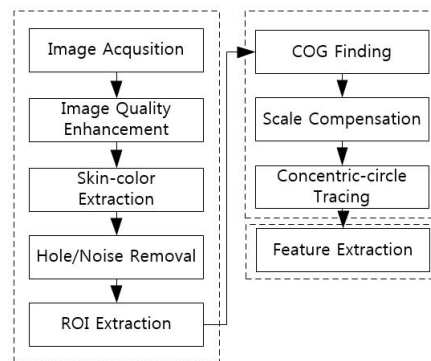


Fig. 1 Structure of system

2.1. 전처리

제안하는 알고리즘에서는 이진 영상을 사용하기 때문에 손 영역만을 분리한 이진 영상을 요구로 한다. 이를 위해 같은 입력받은 이미지에 대하여 전처리를 통해 손 이미지를 얻는다. 전처리 과정은 이미지 품질 향상, 홀/노이즈 제거, 피부색 추출, ROI추출의 과정으로 이루어져 있다.

이미지 품질 향상은 카메라로부터 영상을 입력받고 이미지를 좌우 변환 하여 실제 손의 움직임과 이미지의 좌, 우 움직임을 동일하게 한다. 입력 받은 이미지는 손의 영상부분이 명확하고 다양한 컬러정보를 가지고 있기 때문에 피부색 추출 단계 후 외곽선이 거친 형태의 피부색 이미지가 추출되어진다. 따라서 식 (1)과 같이 Normalized Box Filter를 이용한 블러처리를 통해 영상

의 컬러정보의 편차를 감소시킨다.

$$K = \frac{1}{K_{width} \times K_{height}} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad (1)$$

피부색 추출을 위해서 전 단계에서 처리된 영상에서 사람의 피부색 특성을 이용하여 피부색 영상을 추출하게 된다. 일상적으로 많이 사용하는 RGB 색 공간은 밝기 정보에 매우 민감하다는 문제점이 있다. 따라서 순수한 색상 정보만을 이용하기 위해서 YCbCr, HSV, R/G Ratio 등의 색 공간으로 변환하여 밝기정보를 제외한 순수한 색상 정보로 피부색을 추출할 수 있도록 해야 한다. 이중 YCbCr은 피부색 추출 시 다른 색 공간과 비교하였을 때 가장 정확하게 피부색을 추출한다[6]. YCbCr 색 공간에서 피부색을 식 (2)[7]의 범위 안에 속하는 화소로 정의하고, 피부색인 화소는 255의 값을 주고 그렇지 않은 화소는 0의 값을 가진 이진영상을 만든다. 이렇게 하여 얻어진 영상은 그림 2의 (b)와 같다.

$$\begin{aligned} 133 \leq C_b \leq 173 \\ 77 \leq C_r \leq 127 \end{aligned} \quad (2)$$

그림 2의 (b)를 보면 영상에 노이즈 및 홀이 있음을 확인할 수 있다. 본 논문에서는 노이즈 및 홀을 제거하기 위해서 형태학 연산을 사용하였다. 형태학 연산 중 침식(Erosion)과 팽창(Dilation) 연산을 이용하였고 먼저 노이즈를 제거하기 위해 침식 후 팽창 연산인 열기(Open) 연산을 수행한 후 홀을 제거하기 위한 팽창 후 침식 연산인 닫기(Close) 연산을 수행하였다. 연산을 수행한 후 그림 2의 (c)와 같이 잡음과 홀이 제거된 영상을 얻을 수가 있다.

형태학 연산 처리 후 영상에 피부색과 유사한 작은 객체가 남아있을 수 있다. 본 논문에서는 우선 모든 객체의 가장 바깥쪽에 나타나는 외곽선을 Suzuki85 알고리즘을 통해 검출한다[8]. 검출된 외곽선은 Chain Code로 표현되어 있다. 그린 정리 방식으로 검출된 외곽선들의 면적을 계산한 후 가장 큰 면적을 갖는 외곽선을 손 영상으로 판단한다. 손 영상을 제외한 다른 객체를 없애기 위해서 0으로 채워진 영상공간에 손 영상의 외곽선을 그리고 외곽선 내부 화소 값을 전부 255로 채우

는 작업을 수행한다. 이 과정을 통해서 최종적으로 손만 남은 ROI 영상을 추출 할 수 있으며 결과는 그림 2의 (d)와 같다.

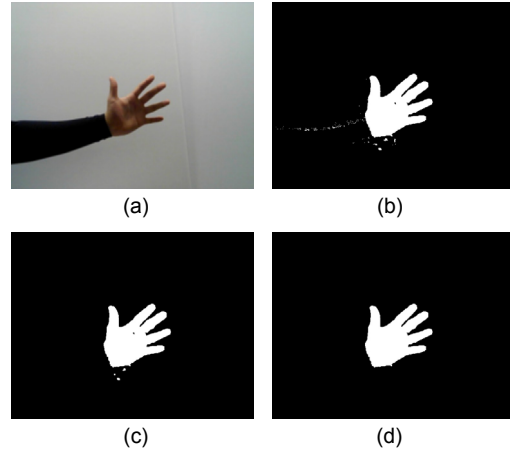


Fig. 2 Image preprocessing (a) Blurred image (b) Binarized skin color image (c) After morphological (d) ROI image

2.2. 동심원 추적 알고리즘을 적용한 특징점 추출

본 논문에서 제안하는 동심원 추적 알고리즘을 적용하여 제스처의 특징점을 추출하는 과정은 우선 COG를 찾고 손의 크기 변화에 따른 동심원의 스케일 보상 후 동심원 추적 알고리즘을 사용하여 특징점을 추출하는 과정으로 구성되어 있다.

손 영상에서 손바닥의 COG을 찾기 위해서 본 논문에서는 거리변환(Distance Transform) 기법을 사용하였다. 거리변환기법은 255값을 가진 해당 화소에서 0값을 가진 화소까지의 최단거리 값을 해당 화소의 값으로 저장하는 기법이다. 거리변환 영상의 가장 큰 값을 가진 화소의 위치가 손바닥의 중심좌표가 되고 해당 화소의 값을 손바닥의 반지름으로 판단한다. COG와 손바닥의 반지름으로 판단하는 반지름의 원은 그림 3과 같다.

동심원 추적 알고리즘에 사용할 동심원의 반지름 길이의 기준을 구해야 한다. 기존의 타 COG기반 알고리즘은 손바닥의 반지름과 임의의 상수를 곱한 값을 취하였지만 이는 사람에 따라 고유한 손의 형태를 반영하지 않았기 때문에 적합하지 않다. 본 논문에서는 최장 손가락 길이와 손바닥의 반지름을 이용하여 사용자의 신체에 최적화된 동심원의 반지름 길이를 계산한다. 다섯 손가락이 펼쳐진 상태인 경우 Convex Hull을 구성하는

점과 손바닥의 COG 사이의 거리를 식 (3)을 이용하여 구한다.

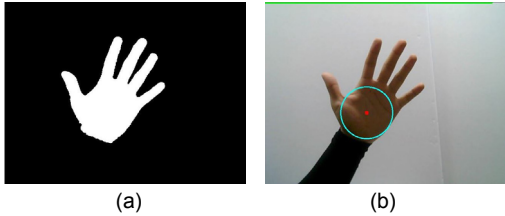


Fig. 3 Acquire ROI and COG image (a) ROI image (b) COG image

$$D_i = \sqrt{(C_x - H[i]_x)^2 + (C_y - H[i]_y)^2} \quad (3)$$

여기서 $C_{x,y}$ 는 COG 좌표의 x, y 값을 나타내며 영상에서 COG를 추출하기 위해서 본 논문에서는 $H[i]_{x,y}$ 는 Convex Hull을 구성하는 i 번째 좌표의 x, y 값을 나타낸다. 이렇게 구해진 i 개의 거리 값 중 가장 최댓값을 최장 손가락 길이로 판단한다. 최장 손가락 길이는 손바닥이 앞, 뒤로 움직일 때에 따라 즉 손의 스케일에 따라 변하게 되어 있다. 제스처를 취하고 있을 때에는 다섯 손가락이 다 펼쳐져 있지 않음으로 최장 손가락 길이를 다시 계산 할 수 없는 문제점을 가지고 있다. 본 논문에서는 프레임 간 손바닥의 반지름의 비율을 이용하여 식 (4)처럼 최장 손가락 길이를 보정한다.

$$MFD_{cur} = MFD_{prev} \times \frac{R_{cur}}{R_{prev}} \quad (4)$$

MFD_{cur} 는 현재 프레임의 최장 손가락 길이를 의미하며 MFD_{prev} 는 전 프레임의 최장 손가락 길이 그리고 R_{cur}, R_{prev} 은 각각 현재 프레임과 이전프레임의 손바닥의 반지름을 의미한다.

본 논문에서는 동심원 추적 알고리즘을 위한 2개의 동심원을 구성한다. 동심원의 반지름은 식 (5)와 같다.

$$\begin{aligned} C_{r1} &= \frac{3}{5}(MFD - R) + R \\ C_{r2} &= \frac{1}{3}(MFD - R) + R \end{aligned} \quad (5)$$

R 와 MFD 는 각각 COG 탐색, 스케일 보상 단계에서 구한 손바닥의 반지름과 최장 손가락 길이를 의미한다. 위의 식을 통해 구한 2개의 반지름을 이용하여 동심원을 구성한다.

본 논문에서 제시하는 제스처 인식을 위한 특징점은 완전히 퍼져 있는 손가락의 개수, 반 접혀있는 손가락의 개수로 정의한다. 이를 구하는 방법으로 먼저 앞서 구성한 동심원에 대하여 외곽선을 추출한다. 추출한 외곽선을 구성하는 좌표를 따라 손 영상의 화소 값이 0에서 255로 바뀌는 좌표를 저장하고 손가락의 개수를 1씩 카운트한다. 이와 같은 동심원-추적 알고리즘을 두 개의 동심원에 대하여 각각 수행한다. 수행 결과는 그림 4와 같다.

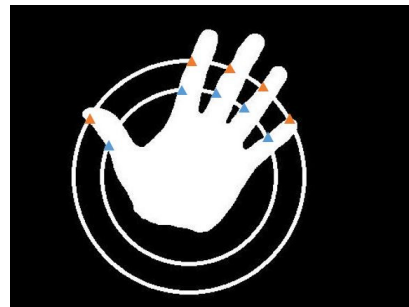


Fig. 4 Result of feature extraction using proposed algorithm

이를 통해서 완전히 퍼진 손가락의 개수와 반만 접힌 손가락의 개수를 구할 수 있다. 완전히 퍼진 손가락의 개수 FC_s 는 C_{r2} 를 반지름으로 구성한 동심원을 동심원 추적 알고리즘을 통해 구한 손가락의 개수로 한다. 접힌 손가락의 개수는 식 (6)을 통해 구할 수 있다.

$$FC_f = FC_{r2} - FC_{r1} \quad (6)$$

FC_f 는 접힌 손가락의 개수를 의미하고 FC_{r1} 과 FC_{r2} 는 각각 반지름 C_{r1} 과 C_{r2} 를 기반으로 동심원 추적 알고리즘을 수행하여 구한 손가락의 개수이다.

III. 실험 및 결과

본 논문에서 제시하는 동심원 추적 알고리즘을 통해

인식 가능한 특징점(FC_s, FC_f)을 이용한 제스처를 손가락 개수를 카운트하는 것과 손가락을 접는 동작이 포함된 것으로 그림 5, 6과 같이 임의로 선정하였다.

실험에 사용한 PC 및 웹 카메라의 사양은 Intel사의 Core i7 3.5GHz CPU, 메모리 8GB를 탑재한 윈도우 기반의 PC 및 200만 화소에 640X480 해상도로 30fps 동영상 촬영이 가능한 웹 카메라를 사용하였다.

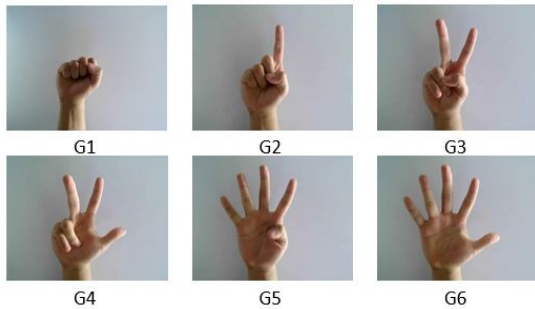


Fig. 5 Sample gesture with finger-counting

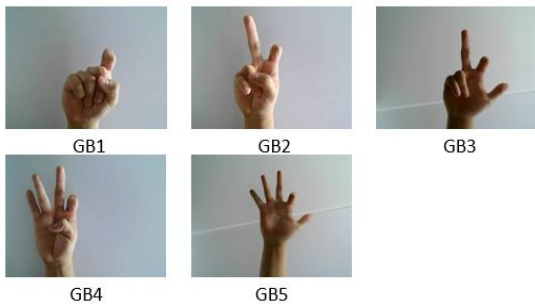


Fig. 6 Sample gesture with finger-folding

본 논문에서는 제스처 인식 성능 평가를 위해서 5명의 피 실험자를 대상으로 그림 5, 6에 제시한 제스처를 각각 30회씩 총 150회를 취하도록 하였을 때 인식률을 계산하였다. 표 1에서 확인할 수 있는 것처럼 완전히 펼쳐진 손가락의 개수만을 특징점으로 가지고 있는 제스처는 96.56%의 높은 인식률을 보이고 있다. 그리고 반접힌 손가락의 개수도 특징점으로 가지고 있는 제스처는 펼쳐져 있는 손가락의 개수가 많을수록 상대적으로 인식률이 떨어지는 것을 확인 할 수 있다. 접힌 손가락을 사용한 제스처의 평균 인식률은 94.4%이고 전체 제스처의 평균 인식률이 95.48%로 이는 기존의 COG를 이용한 알고리즘과 Hu moments를 이용한 알고리즘의

각각의 인식률인 91%, 94%보다 높은 인식률을 가지는 것을 확인할 수 있다[2].

동심원 추적-알고리즘을 사용해서 제스처의 특징점을 찾는데 소요하는 단계별 수행시간을 300회 측정하여 평균을 구한 값은 표 2와 같다. 그리고 우리가 제안하는 방식과 비교 대상 알고리즘의 인식률 및 평균 수행시간을 정리한 결과는 표 3과 같다.

Table. 1 Result of gesture recognition (recognition accuracy, %)

Category	Type	FC_s	FC_f	Accuracy
Gesture with finger-counting	G1	0	0	100
	G2	1	0	96.67
	G3	2	0	98.00
	G4	3	0	94.70
	G5	4	0	93.33
	G6	5	0	96.67
Gesture with finger-folding Ref.[2]	GB1	0	1	100
	GB2	1	1	96.67
	GB3	2	1	93.33
	GB4	3	1	92.00
	GB5	4	1	90.00

Table. 2 Average execution time per stage (time, ms)

Stage	Average Execution Time
Skin-color extraction	15
Concentric-circle tracing	0.053
Etc.	0.547
Total	15.6

Table. 3 Recognition accuracy and average execution time of each algorithm (recognition accuracy, % | time, ms)

Method	Accuracy	Execution Time
Proposed method	95.48	15.6
COG Based Ref. [2]	91.00	22.06
Hu moments Ref. [4]	94.00	4600
FD Ref. [4]	96.00	1700

실시간 제스처 인식 시스템을 위해서는 최대 응답 지연시간을 45ms 이내로 유지해야한다고 알려져 있다 [9]. 본 논문에서 제안 하는 알고리즘의 처리시간은 평균 15.6ms로 실시간 제스처 인식 시스템에 사용 가능함을 확인 하였다. 그리고 COG, Hu moments, Fourier Descriptor를 이용한 다른 제스처 인식 알고리즘의 각각의 평균 수행시간이 22.06ms, 4600ms, 1700ms인 것과 비교하여 우리가 제안하는 알고리즘의 속도가 매우 빠른 것을 확인할 수 있었다[2, 4]. 또한 전체 과정 중 피부색 추출과정에 대부분의 수행 시간을 소요하고 실제 동심원 추적 알고리즘을 수행하는데 소요하는 시간은 0.053ms이기 때문에 본 논문에서 제시하는 알고리즘은 매우 빠르게 손가락의 제스처를 인식하는 것을 확인할 수 있었다.

IV. 결 론

본 논문에서 제시한 알고리즘을 통해서 완전히 펼쳐져있는 손가락의 개수와 접힌 손가락의 개수를 이용해서 기존의 COG기반 제스처 인식 알고리즘과 Convex Hull과 Convexity Defects를 이용한 제스처 인식 알고리즘에서 식별할 수 없는 특징점을 가진 제스처를 높은 정확도로 인식할 수 있다. 또한 알고리즘의 평균 수행 시간 또한 기존에 제안된 알고리즘과 비교하였을 때 빠른 것을 확인할 수 있었고 실시간 제스처 인식 시스템을 위한 최대 응답 지연시간인 45ms 이내를 만족하였기 때문에 사용자는 알고리즘 수행 시간에 의한 지연을 느끼기 어렵다[9].

본 논문의 알고리즘을 이용하여 무균 상태에서 영상 조작을 요구하는 의학 분야, 몸이 불편하여 기존의 키보드, 마우스와 같은 입력 인터페이스를 사용하기 어려운 사람들을 위한 입력 인터페이스, 컴퓨터 엔터테인먼트 시스템의 몰입도를 높일 수 있는 인터페이스로 활용이 가능하다. 본 논문에서 제안한 알고리즘은 손가락이 붙어 있는 경우를 구분하는 것이 불가능한 한계점이 있다. 추후에는 손가락이 붙어 있는 경우의 경우에도 제

스처 식별이 가능하도록 알고리즘을 구현하고 손가락 사이의 각도에 의한 객체 제어가 가능할 수 있도록 알고리즘을 개선해야할 것이다.

REFERENCES

- [1] J. H. Kim, U. J. Ahn, M. H. Lee, J. D. Lee, Human-Computer Interaction basic principle and User-friendly development process, 1st ed. Seoul, KR: Hongrung Pub., ch.9, pp140-141, 2014.
- [2] Malima, Asanterabi, Erol Özgür, and Müjdat Çetin. "A fast algorithm for vision-based hand gesture recognition for robot control." *Signal Processing and Communications Applications, 2006 IEEE 14th. IEEE*, 2006.
- [3] Lee, Daeho, and Youngtae Park. "Vision-based remote control system by motion detection and open finger counting." *Consumer Electronics, IEEE Transactions on* 55.4, pp. 2308-2313, 2009.
- [4] Nagarajan, S., T. S. Subashini, and V. Ramalingam. "Vision based real time finger counter for hand gesture recognition." *International Journal of Technology* 2.2, 2012.
- [5] Panwar, Meenakshi, and Pawan Singh Mehra. "Hand gesture recognition for human computer interaction." *Image Information Processing (ICIIP), 2011 International Conference on. IEEE*, 2011.
- [6] Singh, Sanjay Kr, et al. "A robust skin color based face detection algorithm." *Tamkang Journal of Science and Engineering* 6.4, pp. 227-234, 2003.
- [7] Basilio, Jorge Alberto Marcial, et al. "Explicit image detection using YCbCr space color model as skin detection." *Applications of Mathematics and Computer Engineering*, pp. 123-128, 2011.
- [8] Suzuki, Satoshi. "Topological structural analysis of digitized binary images by border following." *Computer Vision, Graphics, and Image Processing* 30.1, pp. 32-46, 1985.
- [9] Sheridan, Thomas B., and William R. Ferrell. "Remote manipulative control with transmission delay." *IEEE Transactions on Human Factors in Electronics* 4.1, pp. 25-29, 1963.



황동현(Dong-Hyun Hwang)

2010 ~ 한국기술교육대학교 컴퓨터공학부 학사과정

※관심분야 : Human Computer Interaction, Computer Vision



장경식(Kyung-Sik Jang)

1987 고려대학교 공학사

1989 KAIST 전기전자공학과 석사

1998 Tokyo Institute of Technology 전기전자공학 박사

1989 ~ 1997 KT 연구개발단

1998 삼성전자 시스템LSI

1999 ~ 한국기술교육대학교 컴퓨터공학부 교수

※관심분야 : Embedded System, Sensor Network, Vision Sensor