

# 강인한 손가락 끝 추출과 확장된 CAMSHIFT 알고리즘을 이용한 자연스러운 Human-Robot Interaction 을 위한 손동작 인식

## A Robust Fingertip Extraction and Extended CAMSHIFT based Hand Gesture Recognition for Natural Human-like Human-Robot Interaction

이 래 경, 안 수 용, 오 세 영\*  
(Lae-Kyoung Lee<sup>1</sup>, Su-Yong An<sup>1</sup>, and Se-Young Oh<sup>1</sup>)

<sup>1</sup>Pohang University of Science & Technology

**Abstract:** In this paper, we propose a robust fingertip extraction and extended Continuously Adaptive Mean Shift (CAMSHIFT) based robust hand gesture recognition for natural human-like HRI (Human-Robot Interaction). Firstly, for efficient and rapid hand detection, the hand candidate regions are segmented by the combination with robust  $YCbCr$  skin color model and haar-like features based adaboost. Using the extracted hand candidate regions, we estimate the palm region and fingertip position from distance transformation based voting and geometrical feature of hands. From the hand orientation and palm center position, we find the optimal fingertip position and its orientation. Then using extended CAMSHIFT, we reliably track the 2D hand gesture trajectory with extracted fingertip. Finally, we applied the conditional density propagation (CONDENSATION) to recognize the pre-defined temporal motion trajectories. Experimental results show that the proposed algorithm not only rapidly extracts the hand region with accurately extracted fingertip and its angle but also robustly tracks the hand under different illumination, size and rotation conditions. Using these results, we successfully recognize the multiple hand gestures.

**Keywords:** skin color model, extended CAMSHIFT, hand gesture recognition, fingertip, HRI

### I. 서론

최근 R-learning 기반의 교육용 로봇이나 지능형 휠체어 로봇과 같은 재활 보조로봇과 같이 인간에게 친숙하고 같이 생활하면서 도움을 주는 반려자의 존재로써 로봇의 역할이 부각되고 있다. 이에 발맞춰 인간과 로봇간의 원활한 상호작용을 위한 자연스럽게 다양한 형태의 인간 친화적인 인터페이스에 대한 대중적 요구와 이를 이용한 응용서비스 및 관련 연구가 활발히 진행되고 있다[15]. 특히 최근 가상현실, 3D TV, 스마트 폰과 같은 통합형 멀티미디어 기기나, 지능형 시스템 및 응용 서비스의 대중화로 인해, 기존 키보드나 마우스, 음성 인식기반 소통 방법과 같이 환경적인 제약이 많은 방식에 대한 개선의 요구가 급격히 증대되고 이를 대체하기 위한 다양한 방법들이 제안되고 있는 상황이다.

기존의 가장 일반화되어 있는 사용자 인터페이스의 한 영역으로서 영상 기반 손 동작 및 형태 인식은 단일 카메라에서 다수, 3D 카메라 등을 이용한 복합 시스템으로의 확장이 가속화 되고 있다. 특히 그 대표적 사례로, 최근에는 Kinect와 같은 3차원 카메라의 대중화와 더불어 GPU와 같은 영상 처리 성능 개선 기술의 혁신적 발전으로 인한 공간적, 표현적 제약성을 극복하고 저비용, 대중화를 위한 무 표식(markerless) 기반에 손과 손가락 추출 및 관련 동작 인식이 일반화되고, 방향 지시, 클릭, 손가락 필기체 인식 등의 다양

한 영상기반의 손 동작 인식에 대한 연구가 집중적으로 진행되고 있다[2,3,14]. 이와 동시에 기존의 단순히 손 모습, 동작, 부분적 몸 동작 등의 제한적인 표현 방식에서 벗어나 좀더 자연스럽고 다양한 의미의 부여, 전달을 위한 방법으로 전신 동작 인식과 손가락 동작 인식에 대한 연구 또한 최근 활발하게 진행되고 있는 상황이다[1].

이에 본 논문에서는 단일 카메라 기반 환경에서 강인한 손 추적 알고리즘과 손가락 정보 인식 방법을 융합하여 2차원의 손동작 정보를 바탕으로 보다 자연스럽게 다양한 화자의 의사를 전달하기 위한 HRI 방법을 제안하고자 한다. 특히 연속적인 필터링 과정(cascade filtering procedure)을 제한함으로써 다양한 배경과 환경 조건하에서 손 영역 검출의 강인성과 연산 속도 개선 효과를 동시에 가지도록 하였다. 이와 동시에 검출된 손 영역으로부터 효과적인 손가락 끝 추출 방법을 제안하여, 향후 가상현실이나 3D기반 응용서비스로의 활용 가능성을 보인다. 그리고 다양한 손동작 인식을 위해 외부 환경 조건 변화에 강인한 확장된 CAMSHIFT 추적 알고리즘을 제안하고, 이를 통해 얻어진 2차원의 손동작 경로 패턴과 속도 값을 바탕으로 콘덴세이션(condensation) 알고리즘과의 융합을 통한 안정적인 손동작 인식 방법을 제안하였다.

본 논문의 전체적인 구성은 다음과 같다. 먼저 II 장에서 제안하고자 하는 HRI 시스템 구성 및 전체 알고리즘 흐름도를 바탕으로, 개선된 손 영역 검출 방법과 강인한 손가락 끝 추출 및 엄지손가락 판별 방법을 설명한다. 그리고, III 장에서는 검출된 초기 손 영역을 바탕으로 효율적이고 안정적인 손 추적을 위해 제안된 확장된 CAMSHIFT 알고리즘과 이를 통해 추적된 손의 동작 경로 및 관련 정보를 활용한 손동작

\* 책임저자(Corresponding Author)

논문접수: 2011. 12. 13., 수정: 2012. 1. 9., 채택확정: 2012. 2. 11.

이래경, 안수용, 오세영: 포항공과대학교 전자전기공학과

(laeklee@postech.ac.kr/grasshop@postech.ac.kr/syoh@postech.ac.kr)

\* 본 연구는 지식경제부 및 정보통신산업진흥원의 "IT명품인재양성사업"의 연구결과로 수행되었음 (C1515-1121-0003).

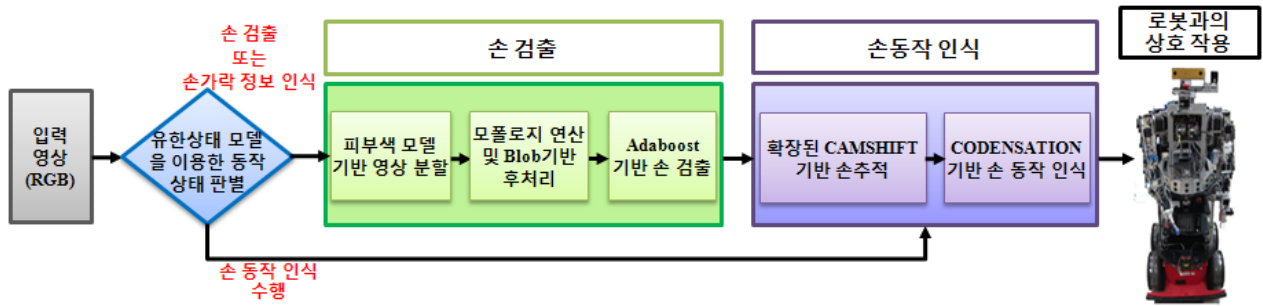


그림 1. 제안된 손 검출, 손 동작 기반의 HRI 시스템 구성도.  
Fig. 1. The proposed hand detection, gesture recognition based HRI system.

인식 방법을 제안한다. 다음 IV 장에서는 제안된 알고리즘들에 대한 다양한 성능 평가 및 해당 결과들에 대한 내용을 기술하고, 끝으로 V 장에서는 제안된 알고리즘에 대한 요약을 통해 본 논문에 대한 결론을 내리며 마무리 짓는다.

II. 개선된 손 영역 검출과 강인한 손가락 끝 추출

본 연구는 그림 1과 같은 HRI 시스템 구조를 바탕으로 배경 조건에 대한 추가 설정이나 제약 없는 일반적인 실내 환경에서 320x240 해상도의 RGB 입력영상을 바탕으로 수행되었다. 그림 2는 본 논문에서 제안하는 손가락 정보 추출 및 안정적인 손 추적 과정을 통한 손동작 인식 알고리즘의 세부 흐름도를 나타낸다.

1. 피부색 모델과 Haar-like 특징을 융합한 개선된 손 영역 검출

보다 자연스럽게 강인한 손 검출과 손동작 인식을 수행하기 위해, 본 논문에서는 그림 3과 같은 연속적인 3단계 검출/필터링 과정을 제안함으로써 안정적인 손 영역을 추출이 가능하도록 하였다. 해당 방법의 경우, 기존의 색상기반 물체 분할 방법이 가지는 장점들(사전 학습과정의 불필요, 알고리즘의 단순화, 실시간 처리)과 Viola & Jones의 특징 정보를 이용한 adaboost 방법[4,5]의 빠른 검출 속도와 높은 인식률의 장점들을 동시에 가지는 단순하면서도 효율적인 융합 방법

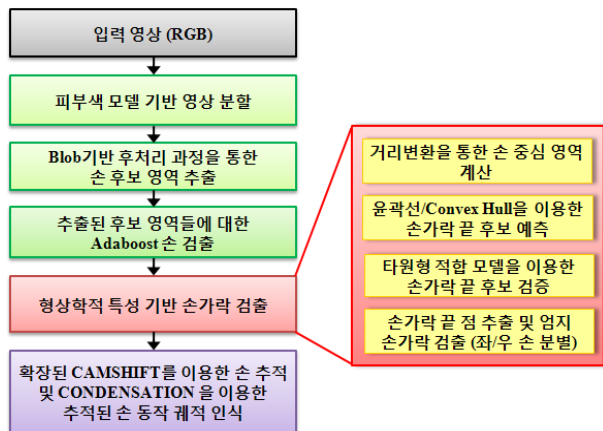


그림 2. 제안된 손가락 정보를 이용한 손동작 인식 알고리즘 흐름도.

Fig. 2. The flow chart of the proposed hand gesture recognition algorithm with finger information.

이다. 본 논문에서는 이러한 상호 보완적인 두 방법들의 융합 구조를 제안함으로써, 대상체 검출 정확도의 향상과 더불어 검출 속도의 개선 효과를 동시에 추구하고자 하였다. 이를 위한 첫 단계로, 색상 정보기반의 손 후보 영역 분할과정과 이렇게 분할된 손 후보 영역들에 대해 후 처리 과정(post-processing)를 거침으로써 주어진 영상에서의 안정적인 1차 손 후보 blob 리스트들의 추출이 이루어진다. 끝으로 추출된 blob 리스트의 영역들에 대한 adaboost 손 검출기를 이용한 검출과정의 수행을 통해 최종적인 손 영역 검출이 완료된다.

1.1 피부색 모델 기반의 손 영역 분할과 후처리 과정

본 논문에서는 다양한 손 동작에 따른 조명 변화에 강인하고 효과적인 피부색 모델의 선정을 위해 기존의 여러 색상기반 피부색 검출 방법들[6,7]에 대한 사전 비교 실험을 통해 얻어진 평가 결과를 바탕으로 안정적인 피부색 기반의 검출 방법을 선택하도록 하였다. 그 결과로 적은 연산량과 낮은 실패율을 가지면서, 조명 변화에 강인한 특성을 가지는 Hsu 방법[6]을 채택하였다. 해당 방법은 표 1과 같이 RGB 영상을  $YCbCr$  색상 모델로 변환한 후, 전체 색상 성분 중 조명 변화에 민감한 Y성분을 제외한  $C_b, C_r$  색상 성분들에 대한 색상 조정과정과 타원 모델 기반의 피부색 필터링 과정을 거침으로써 최종적인 피부색 영역의 검출이 이루어진다. 일반적으로 앞선 과정들을 통해 얻어진 이진화 영상의 경우, 잡음과 더불어 유사한 피부색 영역에 대한 분할 결과들 또한 포함하고 있다. 이에 대해 모폴로지 연산을 통한 잡음 제거 과정과 잔존하는 유사한 피부색 물체들을 제거하기 위한 blob 기반의 연산과정(크기/형상 기반 필터링)으로 구성된 후처리 과정의 수행을 통해 보다 안정적인 손 후보 영역들의 추출이 가능하도록 하였다. 그림 4에서는 다수의 피부색 물체들이 존재하고 조명 변화가 존재하는 입력 영상들에 대한 피부색 영역(적색 부분)의 추출 영상과, 이후 후처리 과정을 통해 추출된 최종적인 손 후보 영역들(녹색 사각 경계선)에 대한 결과들을 보여준다.

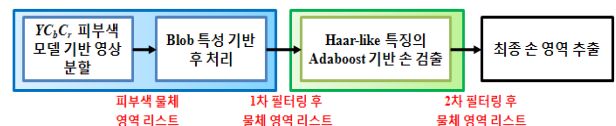


그림 3. 환경 변화에 강인한 손 검출 알고리즘 흐름도.

Fig. 3. The flow chart of the proposed improved hand detection under various light conditions.

표 1. 제안된 피부색 모델 기반 손 영역 추출 알고리즘.

Table 1. The proposed skin-color model based hand region segmentation.

1: 색상 공간 변환 ( $RGB \rightarrow YC_bC_r$ )

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.5 \\ 0.5 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} \quad (1)$$

2: 색상 상관 관계 보정 및 피부색 필터링

2-1: 색상 상관 관계 보정:  $C_b, C_r$  성분에 대한 비선형 변환

$$\bar{C}_b(Y) = \begin{cases} 108 + \frac{(K_l - Y) * (118 - 108)}{K_l - Y_{min}}, & \text{if } Y < K_l \\ 108 + \frac{(Y - K_h) * (118 - 108)}{Y_{max} - K_h}, & \text{if } Y > K_h \end{cases} \quad (2a)$$

$$\bar{C}_r(Y) = \begin{cases} 154 - \frac{(K_l - Y) * (154 - 144)}{K_l - Y_{min}}, & \text{if } Y < K_l \\ 154 - \frac{(Y - K_h) * (154 - 132)}{Y_{max} - K_h}, & \text{if } Y > K_h \end{cases}$$

$$Wc_i(Y) = \begin{cases} WLC_i + \frac{(Y - Y_{min}) * (Wc_i - WLC_i)}{K_l - Y_{min}} & \text{if } Y < K_l \\ WHC_i + \frac{(Y_{max} - Y) * (Wc_i - WHC_i)}{Y_{max} - K_h} & \text{if } Y > K_h \end{cases} \quad (2b)$$

$$C'_i(Y) = \begin{cases} (C_i(Y) - \bar{C}_i(Y)) \frac{Wc_i}{Wc_i(Y)} + C_i(K_h) & \text{if } K_h < Y < K_l \\ C_i(Y) & \text{if } Y \in [K_l, K_h], \end{cases}$$

where  $i$  in  $Wc_i(Y)$  is  $b$  or  $r$ ,  $Wc_b=46.97$ ,  $WLC_b=23$ ,  $WHC_b=14$ ,  $Wc_r=38.76$ ,  $WLC_r=20$ ,  $WHC_r=10$ ,  $K_l=125$ ,  $K_h=188$ ,  $Y_{min}=16$ , and  $Y_{max}=235$ .

2-2: 피부색 필터링: Hsu의 타원형 모델 적용

$$\begin{aligned} x &= (\cos \theta) * (C'_b - c_x) + (\sin \theta) * (C'_r - c_y), \\ y &= (-\sin \theta) * (C'_b - c_x) + (\cos \theta) * (C'_r - c_y), \end{aligned} \quad (3)$$

$$\frac{(x - ec_x)^2}{a^2} + \frac{(y - ec_y)^2}{b^2} \leq 1,$$

where  $c_x = 109.38$ ,  $c_y = 152.02$ ,  $\theta = 2.53$  in radians,  $ec_x = 1.60$ ,  $ec_y = 2.41$ ,  $a = 25.39$ , and  $b = 14.03$ .

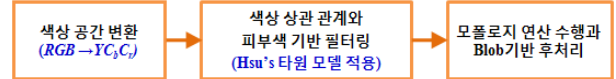
3: 모폴로지 연산과 Blob 기반 후처리 과정

- 모폴로지 열림/닫힘 연산을 통한 배경 잡음 제거
- Blob 기반 후처리를 통한 손 후보 영역 추출

결과: 검출된 피부색 손 후보 영역  $R_{hand}$

1.2 Haar-like 기반 Adaboost 기반 손 영역 검출

앞서 제안된 방법을 통해 얻어진 결과에는 그림 4(c)나 그림 5(d)와 같이 손 영역 이외에 영역들(얼굴, 목, 그림/배경 등)이 함께 존재하게 된다. 이런 유사 피부색 물체들과의 효과적인 분별을 위해서, 본 논문에서는 그림 5와 같이 기존 얼굴 검출 영역에 많이 사용되는 Viola & Jones의 adaboost 학습 방법을 손 검출에 적용함으로써 보다 빠르고 강한 검출 성능을 확보하도록 하였다. 이를 위해 그림 5(b)에 보이는 것처럼 총 15개의 haar-like 특징들(기본 특징: 8개, 확장형 특징:



(a) The proposed skin-color model based segmentation process.



(b) Multiple skin-color object segmentation results.



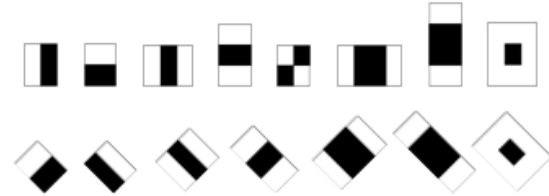
(c) Skin-colored objects detection under various light conditions.

그림 4. 제안된  $YC_bC_r$  피부색 모델 기반의 영상 분할.

Fig. 4. The proposed  $YC_bC_r$  skin-color model based segmentation.



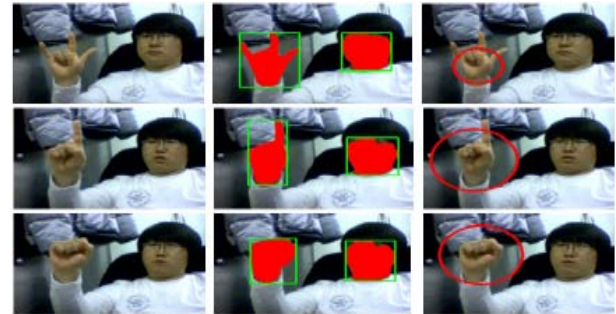
(a) Haar-like features based hand detection process.



(b) The proposed haar-like features set for hand detector.



(c) Adaboost based hand detection result.



(d) Improved hand detection results (skin color + haar-like feature).

그림 5. 제안된 Haar-like 특징 기반 손 검출 및 개선된 검출 알고리즘 기반 검출 결과.

Fig. 5. The proposed haar-like feature based adaboost hand detection and improved hand detection results.

7개)을 이용하여 손에 대한 학습 과정을 수행한 뒤, 학습 결과를 바탕으로 주어진 입력 영상에 대한 손 검출 과정을 수행하게 된다(그림 5(c), (d)). 기존 adaboost 기반 물체 검출의 경우, 일반적으로 영상 전체 영역에 대한 탐색과 검출 과정을 수행하는 반면, 본 논문에서는 검출된 손 후보 blob 리스트들의 존재 영역에 국한하여 탐색/검출 과정이 적용됨으로



써, 기존 방식에 비해 검출 속도 개선(최소 20 % 상승)효과와 더불어 오인식 감소(최대 8~9 %)효과를 동시에 가질 수 있을을 여러 실험들을 통해서 확인할 수 있었다. 그림 5에서는 피부색 모델과 adaboost 모델을 융합하는 과정과 제안된 방식의 손 검출 결과들을 보여주고 있다.

2. 다양한 손동작에 강인한 손가락 끝 추출 및 엄지 손가락 판단 방법

본 논문에서는 기존의 보편적인 손 형상, 모션 정보뿐만 아니라 사람들간의 의사 소통 방식과 같이 손가락 개수 및 방향 등의 손가락 정보의 추출, 융합을 통해 보다 자연스럽고 직관적인 의미 전달과 표현이 가능한 방법을 제안하고자 한다. 이를 위해서 특정 마크나 부가 장치 없이 일반 저가의 단일 카메라 환경에서 실시간 연산이 가능하고, 손의 고 자유도 특성에 따른 형상/자세 변화에 강인한 손가락 끝 추출 방법과 엄지 손가락의 판별 방법을 제안한다.

기존의 대표적인 손가락 추출방법으로는 그림 6과 같이 손의 경계(contour)영역을 따라 얻어진 기울기 정보(K-Curvatures)를 기반으로 손가락 사이의 굴곡 정보를 이용한 검출 방법[8]과 그림 7과 같은 Convex Hull 기반의 Convexity Defects를 이용한 방법[9]이 있다. 이외에도 능동 형상 모델(active shape modeling)과 에지(edge)정보를 이용해 손가락 형상 모델과의 매칭 확률을 이용한 검출 방법도 있다. 일반적으로 기울기 정보를 이용한 전자의 경우, 기울기 차이가 뚜렷한 형상에 대해서는 보편적으로 좋은 결과를 보여주나 다양한 포즈로 인해 발생하는 크기 변화나 굴곡화된 형상은 물



그림 6. 기울기 정보를 이용한 손가락 끝 추출 결과.  
Fig. 6. Curvature based fingertip extraction results.



그림 7. 제안된 손가락 끝과 엄지 손가락 추출 과정 및 결과.  
Fig. 7. Proposed fingertip & thumb identification detection flow chart and results.

론, 손 이외에 손목, 팔꿈치 등 부가적인 영역이 포함된 상황에 대해서는 다수의 오 검출 결과를 보인다. 그로 인해 정확한 손가락 끝 검출을 위해서는 추가 필터링 과정과 정량화된 허용 기울기 값의 선정에 대한 문제점을 가지고 있다. 반면 Convex Hull 기반 방법의 경우, 전자에 비해 계산량이 적고, 비교적 노이즈에 강인한 방법인 반면, 정확한 손가락의 추정 및 해당 손가락의 방향 정보를 얻기 위해서는 추가 후처리 과정을 반드시 필요로 하는 단점을 가진다.

본 논문에서 제안하는 강인한 손가락 끝 추출 방법은 후자의 방법을 기반으로 그림 7(a)와 같은 6단계의 세부 과정들로 구성된다. 첫번째 단계로, 앞서 손 영역 추출 과정에서 얻어진 최종적인 손 영역 혹은 피부색 기반 필터링 과정을 통해 얻어진 손 영상(Step #1)에 대한 거리 변환(distance transform)연산과 그 결과 영상의 특징을 이용해 손의 중심점을 추출하는 과정(Step #2)이 수행된다. 이후 손 영역의 경계 정보의 추출과정과 Convex Defects 계산을 통해, 손가락 끝 후보 지점들을 추출하는 과정들(Step #3, #4)을 수행하게 된다. 여기서 추출된 손가락 끝 후보 영역들이 손의 크기, 모양, 자세 변화에 따라서 가지는 많은 오류 정보들을 제거하기 위해 이웃한 손가락 끝 후보 점들간의 상호 거리, defects정보에 기반한 후보 지점들간의 융합/필터링 과정을 거치게 된다. 이렇게 얻어진 후보 지점들 중에 보다 안정적인 손가락 끝점을 예측하기 위해, 손가락 첫 마디가 가지는 타원형의 형태학적 특성을 토대로 개선된 타원 근사화 모델(enhanced direct least-square fitting of ellipse)과의 적합도를 평가하게 된다. 이러한 결과를 바탕으로 각 손가락 끝점과 손 중심점과의 벡터 각도 차 정보를 이용한 추가 필터링 과정을 거침으로써, 근사화된 최적의 손가락 끝 지점들을 예측하는 과정을 마치게 된다 (Step #5). 끝으로 예측된 손가락 끝 지점들에 대한 타원 중심 좌표와 그 기울기 정보를 이용하여 인접한 손 경계 정보와의 교점 계산을 통해 최종적인 손가락 끝 지점의 위치를 추출할 수 있다(Step #6).

대부분의 손동작 인식의 경우, 정밀한 손가락 끝점에 대한 위치 정보보다는 비교적 안정적인 손가락 끝점의 존재 위치와 정확한 각도 정보만으로도 충분히 그 활용 가치를 가지기 때문에, 소요 연산 시간 및 적용 어플리케이션의 필요 정밀도 수준에 따라 마지막 단계인 최종 손가락 끝 좌표 추출 단계는 선택적인 적용이 가능한 특성을 가진다. 본 논문에서는 이후 손동작 인식 실험 과정에 있어, 연산 시간의 단축을 위해 근사화된 타원 중심 좌표를 실제 손가락 끝점 좌표로 대체하여 사용하도록 하였다. 또한 손 혹은 손가락을 이용한 동작 인식에 있어 엄지 손가락은 그 존재 여부에 따라 좌우 손의 구분 역할 이외에, 손가락을 통한 의미 표현, 다양한 모션 구분 용도 등에서 그 검출의 중요성을 내포하고 있다. 이를 위해 본 논문에서는 손의 중심 좌표를 기준으로 추출된 손의 방향 정보와 손가락 끝 좌표 간의 벡터 각(반 시계 방향 기준)을 이용하여 강인한 엄지 손가락의 검출이 가능하도록 하였다. 그림 7(b)의 첫째 행 영상들은 7(a)에서 설명한 손가락 끝 좌표의 추출 단계별 결과들을 보여주며, 7(b)의 둘째 행에서는 여러 입력 영상들에 대한 엄지손가락 추출 결과와 그 세부 항목들을 보여준다.

**III. 확장된 CAMSHIFT 알고리즘을 이용한 손 추적과  
손동작 경로 정보를 활용한 손동작 인식**

**1. 확장된 CAMSHIFT 알고리즘을 이용한 이동 물체 추적**

보다 정확한 손동작 인식의 수행을 위해서는 안정적인 손 추적 과정이 반드시 선행되어야 한다. 이를 위해 본 논문에서는 다양한 손동작에 따라 빈번하게 발생하는 얼굴과 같은 유사 색의 물체 혹은 배경과의 겹침 문제(occlusion)을 해결하고, 배경 및 조명 변화에 강한 추적 성능을 가지는 확장된 CAMSHIFT 알고리즘을 제안한다. 그림 8에서 보여지는 바와 같이 제안된 확장형 CAMSHIFT 알고리즘은 HSV색상 공간에서 색상(hue)정보 이외에 채도(saturation)성분을 추가적으로 사용하는 다채널 색상 히스토그램 모델의 채택하여 적은 연산량으로 낮은 색상(hue)값에서의 색상 분별력 개선 효과를 가지도록 하였다. 그리고 안정적인 물체 추적 성능을 위해 적응형 색상 히스토그램 기반의 역 투영 방법을 이용한 추적 대상체의 히스토그램 모델 갱신 방법을 제안하였다. 이렇게 제안된 추적 대상체의 색상 히스토그램 갱신과정은 그림 9와 같은 과정들의 순차적 반복을 통해 이루어진다.

먼저, 사용자의 초기 설정 혹은 이전 검출 결과에 따라 얻어진 추적 윈도우 영역을 동일한 크기의 M개의 블록으로 분할하는 과정을 수행한다. 그리고 각 분할된 영역들에 대한 누적 히스토그램의 계산과 식 (4)를 통해 얻어진 기준 히스토그램과의 오차 값 계산을 통한 개별 블록들에 대한 분류 및 가중치 할당 과정이 수행된다. 여기서 개별 블록들은 식 (4)의 범위에 따라 완전 유효 항목(perfect inlier), 부분적 유효 항목(weighted inlier), 거짓 항목(outlier)의 세 항목으로 구분되고 (그림 9(b)에서 각각 흰색 블록, 파란 블록, 적색 블록으로 표현), 이들 중 유효 항목들(부분적 유효+완전 유효 항목)의 히스토그램과 그 가중치를 바탕으로 현재 시간(t)의 평균 히스토그램 분포( $\hat{H}_t$ )를 계산하게 된다. 그리고 식 (5)에서 주어진 것과 같은 추적 대상체의 색상 히스토그램 모델 (H)의 갱신 과정을 통해서, 색상/ 휘도 변화에 강한 추적 특성을

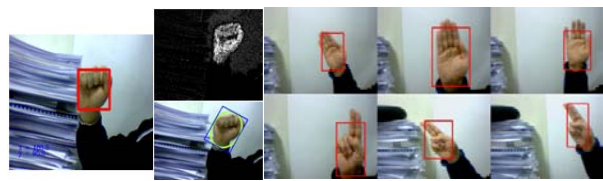
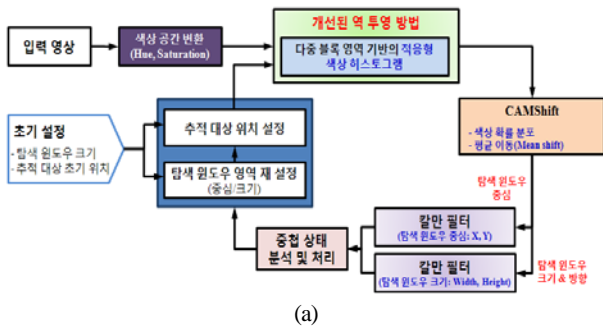


그림 8. 제안된 확장된 CAMSHIFT 알고리즘 및 손 추적결과.  
Fig. 8. Proposed Extended CAMSHIFT flow chart and hand tracking results for various hand pose.

가지도록 하였다(일반적으로  $\tau=0.3\sim 0.5$ 사이 값 적용).

$$E'_{block\ i} = \frac{1}{S_i} \sum_{x \in B_i} [H'_{ij} - H'_{ij}{}^{t-1}]^2 \quad (4)$$

$$H_t = (1 - \tau) * \hat{H}_t + \tau H_{t-1} \quad (5)$$

한편, 유사 색상의 물체 혹은 배경과의 겹침 상황에 대한 판단과 그에 대한 상황 처리를 위해, 그림 8와 같이 위치 정보와 속도 정보로 구성된 상태 벡터  $\mathbf{x}_t = (x(t), y(t), v_x(t), v_y(t))^T$ 와 측정벡터  $\mathbf{z}_t = (x(t), y(t))^T$ 를 가지는 2개의 칼만 필터[11]와 식 (6)과 같은 Filter's Residue개념의 결합 구조를 채택하였다.

$$r(k) = \sqrt{(x(k) - \hat{x}(k))^2 + (y(k) - \hat{y}(k))^2} \quad (6)$$

해당 구조를 바탕으로, CAMSHIFT를 통해 예측된 대상체의 중심 위치와 크기(폭, 높이/각도)를 함께 보정함으로써, 타 물체와의 겹침 발생시 안정적인 추적이 가능하도록 하였다. 여기서  $\hat{x}(k)$ ,  $\hat{y}(k)$ 는 현재 이미지상의 물체 위치 예측 정보를,  $x(k)$ 와  $y(k)$ 는 관측된 위치 정보를 의미한다. 그림 8는 제안된 확장된 CAMSHIFT의 전체 흐름도와 주어진 입력 영상에서의 개선된 역 투영 결과와 손 추적 결과들을 보여주고 있다. 그림 9는 다중 블록 기반의 적응형 컬러 히스토그램 갱신 과정의 흐름도와 다중 히스토그램 적용 결과 및 그에 따른 역 투영 결과를 보여준다. 그리고 그림 10에서는 크기/이동/회전 변화가 존재하는 연속적 손동작 영상으로부터의 추적 결과들을 보여준다.

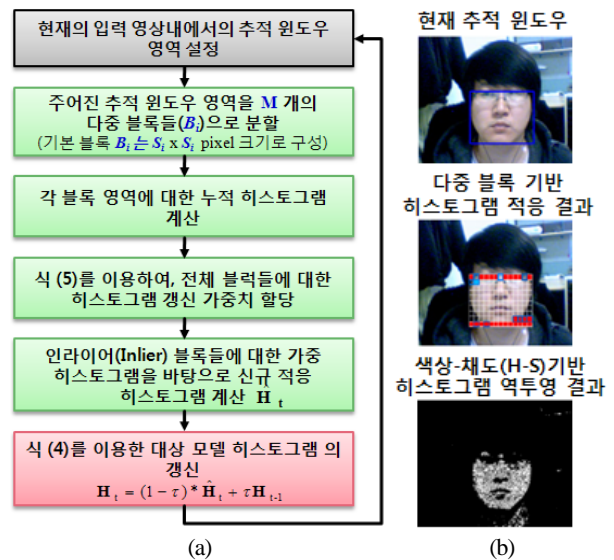


그림 9. 연속적 입력 영상에 대한 다중 블록 기반 적응형 히스토그램 갱신 과정.

Fig. 9. Multi-block region based adaptive target histogram model update process for sequential input images.



그림 10. 확장된 CAMSHIFT기반 손 추적 결과.  
Fig. 10. Extended CAMSHIFT based hand tracking results.



2. 손동작 경로 정보를 이용한 CONDENSATION 기반 손동작 인식 알고리즘

앞서 본 논문에서는 손동작 인식 방법을 손의 형태, 위치 변화 정보에 대한 인식[2]과 수화와 같이 손동작이 가지는 2차 혹은 3차원적 동작 패턴에 대한 인식[12]으로 구분하였다. 특히, 단순한 손의 형상 및 위치 변화 인식에 치중된 전자의 방법 대신에 보다 일반화되고 자연스러운 손의 2차원적 동작 특성에 중점을 두고, 여기에 손가락 개수/방향 정보를 융합함으로써, 다양한 손동작의 패턴을 통한 의사전달이 가능한 알고리즘을 제안하고자 한다. 이를 위해서 앞서 제안된 안정적인 손 추적의 결과를 바탕으로 얻어진 손 중심 좌표와 속도 정보로부터 생성된 연속적인 경로 패턴 데이터들을 학습, 인식하는 과정을 제안한다. 이에 효과적인 2차원 경로 데이터들의 추출을 위해 광류 정보(optical flow) [12]를 이용한 모션 속도성분을 바탕으로 현재 상황에 대한 유한 상태 모델(finite state model)에 따라 동작의 시작, 진행과 종료 시점을 판단하고, 상황에 맞게 선택적인 동작 경로 데이터 추출/저장 과정을 수행하도록 하였다. 이런 과정들을 통해 동일한 기준 패턴의 동작 모델에 대해서도 사용자와 취해지는 동작의 특성에 따라 불규칙한 경로 데이터들이 빈번하게 입력되고, 이를 기준 동작 모델 정보와의 효과적인 매칭 과정을 수행하기 위해 Black & Jepson이 제안한 콘텐츠제이션[12] 알고리즘을 채택함으로써, 추출된 입력 경로 및 속도 데이터들에 대한 확률적 판단에 기반한, 안정적인 인식과정의 수행이 가능하도록 하였다. 해당 알고리즘의 경우, 입력 경로 데이터를 곡선으로의 근사화하기 위한 N 개의 가중치 샘플을 이용하는 방법으로 각 샘플들은 상태 정보와 확률에 비례하는 가중치로 구성되는 구조를 가지고, 학습된 모델에 대한 주어진 입력 데이터의 적합도를 확률적 수치로 표현한다. 본 논문에서는 사전에 임의의 손동작 모델들을 학습시키고 난 뒤, 이를 입력 동작 모델과의 유사도를 계산함으로써 해당 동작에 대한 인식 성능을 평가하도록 하였다.

IV. 실험 결과 및 고찰

제안한 방법들의 효과적인 검증을 위해 먼저 개별 방법들에 대한 성능평가 수행 이후, 융합된 전체 알고리즘의 평가를 수행하도록 하였다. 동일한 실험 조건 구축을 위해 각 실험에 맞도록 제작된 영상 DB를 바탕으로 기존 방법과의 비교를 통한 제안된 방법들의 성능 검증 및 평가를 진행하였다.

1. 개선된 손 검출 알고리즘의 성능 평가

제안된 방법에 대한 효과적인 검증을 위해 주어진 입력 영상들에 대한 기존 adaboost 검출기만을 이용한 결과와 제안



그림 11. 다양한 손 입력 영상에 대한 검출 성능 비교.  
Fig. 11. The comparisons for hand detection performance according to different hand poses.

표 2. 손 검출 능력 비교 결과.

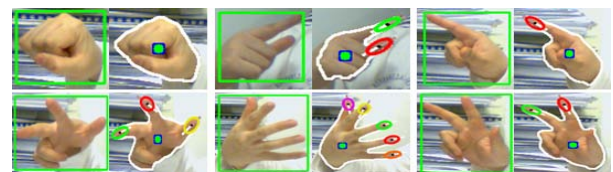
Table 2. Hand detection performance comparison result.

	Adaboost 모델	제안된 융합모델 (Skin Color & Adaboost)
검출 성공률	85 %	91 %
평균 계산시간	14.54 ms	8.63 ms

된 연속 융합 모델을 이용한 방법과의 상호 비교를 통해 성능 평가를 수행하였다. 실험 조건의 단일화를 위해, 그림 11에서와 같은 영상 내에 단일 손이 존재하는 상황을 가정하여 다양한 포즈, 조명 변화를 기준으로 검출 정확도와 소요 계산 시간을 주요 평가 지표로 사용하였다. 제안된 방법의 경우, 입력 영상에 대해 적색으로 표시된 색상기반 영상 분할 결과를 바탕으로 녹색 사각부분의 피부색 후보 영역을 추출하게 된다. 그리고 이 영역 정보를 바탕으로 adaboost 기반 손 검출기에 의해 최종 손 영역이 추출되어 붉은색 원 영역으로 표시된다. 특히 표 2의 결과와 같이, 제안된 융합 모델이 여러 실험 영상들에 대해 배경이나 조명 변화에 강인한 검출 성능을 가지고 동시에 손 영역 검출에 필요한 계산시간의 단축 결과를 증명할 수 있었다. 이는 앞선 연속된 필터링 과정들을 통해, adaboost 검출기의 오 검출 발생 확률을 줄이는 동시에, 슬라이딩 윈도우의 스캔 영역 범위를 추출된 손 후보 영역으로 제한함으로써 불필요한 연산의 최소화를 통해 얻어지는 결과임을 확인할 수 있었다.

2. 손가락 끝 추출 및 엄지손가락 검출 방법의 성능 평가

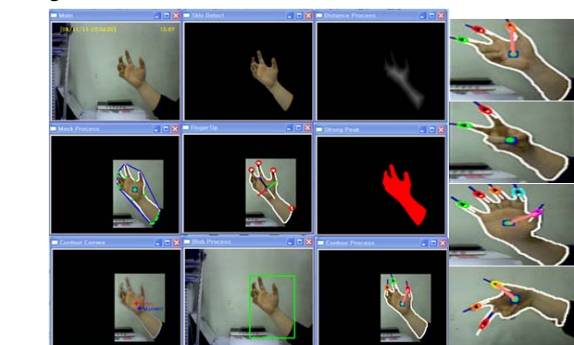
제안된 손가락 끝 추출 방법에 대한 실험을 위해, 총 340



(a) Fingertip and its orientation detection with detected hand region.



(b) Fingertip and thumb identification results for sequential hand gestures.



(c) Fingertip extraction result for overall arm regions.

그림 12. 강인한 손가락 끝 추출 및 엄지손가락 검출 결과.  
Fig. 12. Robust fingertip extraction and thumb identification results.

표 3. 손가락 끝 점 및 엄지 손가락 검출 결과.

Table 3. Fingertip extraction and thumb identification result.

	손가락 끝점	엄지손가락 검출
검출 성공률	95.88 % (326)	97.06 % (330)
오 검출률	4.12 % (14)	2.94 % (10)
평균 계산시간	6.86 ms	

프레임으로 구성된 다양한 자세의 정적 영상과 연속적 동작 변화가 존재하는 실험용 영상을 제작하여, 그에 따른 손가락 끝 추출 및 엄지손가락 판별에 대한 성능 평가를 수행하도록 하였다. 주요 성능 평가 실험들은 앞서 제안된 손 영역 추출 방법에 따른 최적화된 손 영역이 주어졌던 경우와 정확한 손 영역이 주어지지 않은 두 가지 경우로 나누어 진행하였다. 후자의 경우, 사용된 실험 영상 내에 다수 포함된 손목 이외의 팔 부분이 존재하는 영상을 바탕으로 성능 평가 실험을 수행하였고, 그림 12과 표 3와 같은 실험 결과를 얻을 수 있었다. 그림 12(a)의 경우는 최적화된 손 영역이 주어졌던 상황에 대한 손 중심 추출 및 손가락 끝 검출 결과를 보여준다 (녹색/파란색 원: 손 중심, 타원: 손가락 끝점 및 그 방향). 전체적으로 다양한 포즈/각도의 손 영상에 대해서 안정된 손가락 끝점 추출 및 그 각도 예측 결과를 보여준다. 그림 12(b)와 (c)의 경우, 최적화된 손 영역이 주어지지 않는 상황에 대한 정적 영상과 연속적 동작 자세 변화 영상에 대한 검출 결과들을 나타낸다. 특히, 손목, 팔과 같은 부가적인 손 영역들이 존재하고 정확한 대상 영역에 대한 정보가 주어지지 않는 경우, 최소한의 손 영역을 찾은 뒤 제안된 손가락 끝 추출 및 엄지손가락 판별에 대한 각 세부 단계별 수행 결과들을 보여준다(그림 12의 (b), (c)에서 주황색 선은 손 중심과 엄지손가락 연결선을 의미).

3. 확장된 CAMSHIFT 기반 물체 추적 성능 평가

제안한 확장된 CAMSHIFT 기반의 물체 추적 성능을 평가하기 위해 조명 변화와 다양한 외란(불연속 색상, 크기/형상 변화)이 존재하는 상황에서 손 이외의 대상체(얼굴, 원통형 물체 등)에 추적 실험들을 수행하였다. 신뢰성 있는 평가를 위해 기존 CAMSHIFT 알고리즘과 비교와 더불어 손동작 수행시 자주 발생하는 겹침 상황들(유사 색상 이동물체와의 겹침, 동일 형상 물체와의 겹침, 비슷한 색상의 배경과의 겹침 등)에 대해서도 유사 모의 실험을 통해 제안된 알고리즘의 우수성을 검증하였고, 그 결과는 그림 13, 14와 같은 결과 영상들을 통해서 나타난다. 그림 13은 다양한 추적 대상체들의 초기 위치를 사용자가 지정해준 상황에 대해서, 연속적 입력 영상들에 대한 추적 성능을 보인다. 얼굴 추적 영상의 경우, 다양한 외란 요소들(안경/목 영역)이 존재함에도 불구하고 다양한 움직임에서 안정적인 추적 결과(파란 사각형 영역)와 그에 따른 다중 블록기반 히스토그램 판별 결과들을 동시에 보여준다. 그리고 손과 원통형 물체의 경우에 대해서도, 그림에서 보이는 것과 같이 회전/크기/조명 변화가 존재하는 상황에서도 안정적인 추적 성능을 보여줌을 확인할 수 있었다.

그리고 앞서 언급한 다양한 겹침 상황들에 대한 제안된 추적 방법의 성능 평가 결과는 그림 14와 같다. 먼저 그림 14(a)의 경우, 추적 대상은 얼굴 영역이며 추적 과정 중 유사 색상의 이동물체인 손으로 인해 발생하는 겹침 현상에 대한

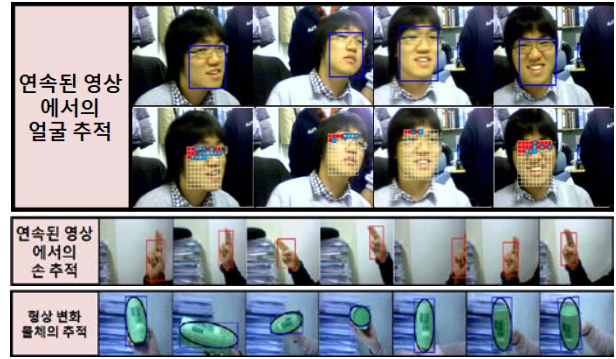
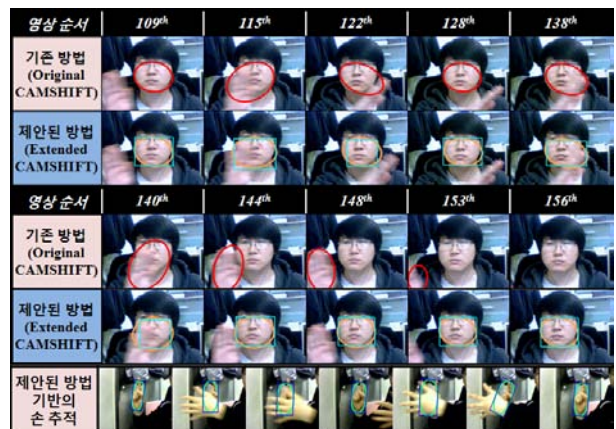
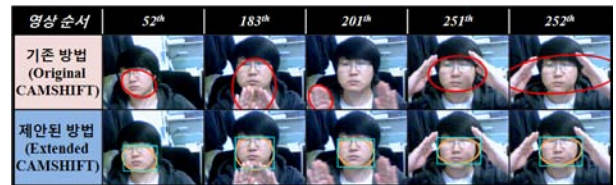


그림 13. 확장된 CAMSHIFT 기반 다양한 물체 추적 결과.  
Fig. 13. Extended CAMSHIFT based various object tracking results.



(a)



(b)

그림 14. 다양한 겹침 조건에 대한 확장된 CAMSHIFT 기반 물체 추적 결과.

Fig. 14. Extended CAMSHIFT based various object tracking with several occlusion conditions.

추적 실험 결과 영상들을 보여주고 있다. 효과적인 성능 비교를 위해, 기존 알고리즘과의 동일 초기 추적 조건을 바탕으로 해당 영상들에 대한 두 방법에 대한 추적 결과들 중 일부 결과에 대한 도식화(기존 방법-적색 타원, 제안 방법-파란색 사각형)를 통해 비교하도록 하였다.

기존 방법의 경우, 유사 물체와의 겹침 발생 이전 상황에서는 비교적 안정적인 추적 성능을 보이나, 몇 차례 겹침 발생 이후 (115번째 영상, 122번째 영상, 138번째부터 144번째 영상까지) 다소 불안정한 추적 결과를 나타내다가, 결국 140번째 영상 이후부터는 목표 대상체의 추적 정보를 잃어버리고 다른 유사 물체를 추적하는 결과를 보임을 알 수 있었다. 그리고 156번째 영상 이후로는 추적 대상체에 대한 정보 소실로 인해 더 이상 추적 과정의 수행이 불가능하였다. 이와



는 반대로 제안된 방법의 경우, 실험 중에 여러 차례 겹침 현상의 발생에도 불구하고 안정적이고 강인한 추적 결과를 보임을 증명하였고, 이후 영상들에 대해서도 성공적인 물체 추적의 결과를 보였다.

또한 하나의 고정 추적 대상체인 손에 대해 다른 이동 손과의 겹침이 발생하는 경우(동일/유사물체와의 겹침 조건)에 대해서도 그림 14(a)에서 보는 것과 같이, 겹침 발생시 부분적 추적 결과의 변화가 발생하였으나, 겹침 상황 발생 이후 추적 대상체 크기나 위치 예측에 있어 안정적인 추적 결과를 보임을 확인하였다. 그리고 비슷한 색상의 배경과의 겹침 조건에 대한 유사한 상황을 만들기 위해, 앞선 첫번째 실험과는 반대로 추적 대상체는 이동 중인 얼굴 영역으로 두고, 유사 색상의 물체가 추적 대상물 주변으로 접근하고 일정 기간 정지해있는 조건에 대해서 그 성능을 테스트하였다. 그림 14(b)에 나타난 결과와 같이, 기존 알고리즘은 유사 색상의 정적 물체 혹은 배경으로 인한 추적 실패에 따른 발산으로 인해, 대상체 크기 및 위치 예측 모두 실패하는 결과를 보이는 반면, 제안된 방법의 경우 안정적인 추적 결과를 보임을 알 수 있었다. 비록 기존 방법과 비교했을 때, 연산 수행 시간의 부분적 증가가 발생하였지만, 최대 50 fps 범위까지 처리 속도가 보장되고 추적 대상 물체에 대한 다양성과 배경 조건에 대한 강인성을 동시에 가짐으로써, 안정적인 추적 성능을 가진 실시간 추적 방법임을 여러 실험들을 통해 입증할 수 있었다.

#### 4. 손동작 경로를 이용한 손동작 인식 성능 평가

앞선 실험 과정을 통해 이미 그 성능이 검증된 확장된 CAMSHIFT 기반 손 추적 정보를 이용하여, 그림 15과 같이 6개의 사전 학습된 손동작 경로 패턴 모델을 기준으로 콘텐츠 세션 알고리즘을 이용한 인식 성능 평가를 수행하였다. 본 실험에서는 등록된 패턴들을 기준으로 총 100회의 다양한 동작 변화(속도의 가 감속, 손 궤적의 부분 변화, 손 모양 변화)를 통해 사용자 동작의 인식 성공률을 측정하였고 그 결과는 표 4와 같이 평균 84 %의 신뢰할 수 있는 인식률을 나타내었다. 다소 차이가 있었지만 안정적인 물체 추적 능력을 가진 방법을 이용함으로써, 손 모양 조건의 변화에 따른 인식률의 변화는 미미한 반면, 손동작 궤적의 불명확성 과 급격한 속도 변화 요소의 경우 입력 데이터에 직접적인 영향을 줌으로 인해 전체적인 인식률 변화에 비교적 큰 영향을 미침을 알 수 있었다. 해당 실험들을 통해 제안된 방법이 비교적 안정적인 인식 성능을 가짐을 확인할 수 있었지만, 향후 인식하고자 하는 동작들의 증가와 더불어 유사 입력 패턴들의

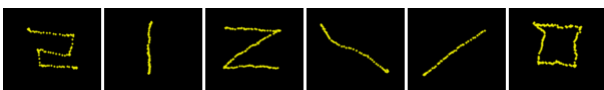


그림 15. 손동작 인식을 위해 사전 인식된 손동작 경로 모델.  
Fig. 15. Predefined hand gesture trajectories for gesture recognition.

표 4. 손동작 인식 결과.

Table 4. Hand gesture recognition result.

	손동작 모델 수	인식률 (%)
실험 결과	6	84 %

존재에 따른 안정적인 인식률 보장을 위한 개선 방안이 요구 될 것이다. 이를 위해 효과적인 입력데이터의 정량화 과정과 전처리 과정이 추가되어야 하며, 동시에 신경망과 같은 보다 안정적이고 개선된 패턴 인식 알고리즘의 적용 또한 고려되어야 할 것이다. 그리고 여기서는 추가적인 실험 결과를 포함하지 않았지만, 앞서 제안된 손가락 개수 및 손가락 방향, 엄지 손가락의 유무 정보를 같이 이용할 경우, 같은 손동작이라도 다양한 의미 표현으로 확장, 전달할 수 있음은 따로 구체적인 검증이 필요치 않으며, 이를 통해 보다 자연스럽고 다양한 의미의 의사전달 방법으로 사용될 수 있는 가능성을 보여줄 수 있다.

## V. 결론

본 논문에서는 HRI 분야로의 응용을 위한 자연스러운 손동작 인식 방법을 제안하였다. 특히, 단일 카메라의 입력 영상으로부터 피부색 색상 모델과 adaboost 학습을 융합한 개선된 손 영역 검출 방법을 제안하고, 동시에 검출된 손 영역으로부터 강인한 손가락 끝점 추출 및 엄지 손가락의 판별 방법을 제안하였다. 이를 통해 손동작으로부터 다양한 손가락 정보들을 획득할 수 있었고, 이를 확장된 CAMSHIFT 기반 손동작 궤적 정보와의 융합을 통해 보다 자연스럽고 다양한 의사 표현 및 전달이 가능하도록 하였다. 끝으로 효과적인 동작 상태 판단에 따른 입력된 동작 패턴 데이터의 인식을 위해 콘텐츠 세션 기법을 적용함으로써 안정적인 인식과정이 수행되도록 하였다. 제안된 방법들에 대한 다양한 실험들을 통해 제안된 방법들의 우수한 성능을 검증할 수 있었으며, 이러한 결과를 토대로 로봇이나 지능형 휠체어 동작 제어 혹은 휴머노이드 로봇의 손/손가락 제어와 같은 HRI 분야로의 직접적인 응용뿐만 아니라 멀티미디어 기기의 제어와 같은 타 분야로의 확장 응용 또한 가능할 것으로 사려된다.

## 참고문헌

- [1] D. D. Nguyen, T. C. Pham, and J. W. Jeon, "Fingertip detection with morphology and geometric calculation," *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1460-1465, St. Louis, USA, Oct. 11-15, 2009.
- [2] L. K. Lee and S. S. Kim, "Recognition of hand gesture to human-computer interaction," *Korean Institute of Intelligent Systems (in Korean)*, vol. 11, no. 1, pp. 28-32, 2001.
- [3] L. K. Lee, S. Y. An, and S. Y. Oh, "A robust fingertip extraction and hand tracking for human-robot interaction," *ICROS Annual conference*, pp. 138-142, 2011.
- [4] P. Viola and M. J. Jones, "Robust real-time face detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137-154, 2004.
- [5] R. Lienhart and J. Maydt, "An extended set of haar-like features for rapid object detection," *IEEE International Conference on Image Processing*, pp. 900-903, 2002.
- [6] R. L. Hsu, M. Abdel-Mottaleb, and A. K. Jain, "Face detection in color images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 696-706, 2002.
- [7] M. J. Jones and J. M. Rehg, "Statistical color models with application to skin detection," Tech. Rep. Tech. Rep. CRL 98/11, Cambridge Research Laboratory, Dec. 1998.



- [8] S. Malik, "Real-time hand tracking and finger tracking for interaction," CSC2503F Project Report, Dec. 2003.
- [9] C. Davatzikos and J. L. Prince, "Convexity analysis of active contour problems," *Image Vision Computing*, vol. 17, pp. 27-36, 1999.
- [10] G. R. Bradski, "Computer vision face tracking for use in a perceptual user interface," Intel Technology Journal Q2, 1998.
- [11] Applied Optimal Estimation, A. Gelb, ed. MIT Press, 1992.
- [12] M. J. Black and A. Jepson, "A probabilistic framework for matching temporal trajectories: condensation-based recognition of gestures and expressions," *Proc. of European Conference Computer Vision*, vol. 1, pp. 909-924, 1998.
- [13] S. Baker and I. Matthews, "Lucas-kande 20 years on: a unifying framework: Part 1," Technical Report CMU-RI-TR-02-16, Carnegie Mellon University Robotics Institute, 2002.
- [14] S. W. Shon, J. Beh, C. J. Yang, H. Wang, and H. S. Ko, "Robot user control system using hand gesture recognizer," *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 17, no. 4, pp. 368-374, 2011.
- [15] K. E. Kom and K. B. Sim, "Study of emotion recognition based on facial image for emotional rehabilitation biofeedback," *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 16, pp. 957-962, 2010.



### 이 래 경

1999년 부산대학교 전기공학과 졸업.  
2001년 부산대학교 대학원 전기공학과 졸업(공학석사). 2009년~현재 포항공과대학교 전자전기공학과 박사과정 재학 중. 관심분야는 이동로봇, HRI, SLAM.



### 안 수 용

2006년 포항공과대학교 전자전기공학과 졸업. 2006년~현재 포항공과대학교 대학원 전자전기공학과 통합과정 재학 중. 관심분야는 이동로봇, SLAM.



### 오 세 영

1974년 서울대학교 전자공학과(공학사).  
1978년 Case Western Reserve Univ. 전자공학과(공학석사). 1981년 Case Western Reserve Univ. 전자공학과(공학박사).  
1981년~1984년 Univ. of Illinois 조교수.  
1985년~1988년 Univ. of Florida 조교수.  
1988년~현재 포항공과대학교 전자전기공학과 정교수. 관심분야는 신경망, 진화연산, 이동로봇, 지능자동차, 지능시스템.