

<http://dx.doi.org/10.7236/JIIBC.2013.13.4.91>

JIIBC 2013-4-13

단일 형상의 이동 방향 인식에 의한 손 동작 리모트 컨트롤러 구현

Implement of Finger-Gesture Remote Controller using the Moving Direction Recognition of Single

장명수*, 이우범**

Myeong-Soo Jang, Woo-Beom Lee

요약 본 논문에서는 단일 카메라를 사용하여 사용자 제스처 명령을 사용자의 손가락 개수 및 이동 방향을 기반으로 인식하는 손동작 리모트 컨트롤러 인터페이스를 구현한다. 구현한 손동작 리모트 컨트롤러 인터페이스는 손 영역의 추출을 위해서 RGB 컬러모델로부터 YCbCr 컬러모델로의 색 변환에 의한 색차정보를 이용하고, 추출된 손 영역의 중심에서 이중 원형 추적(Double Circle Tracing)에 의한 교차정보를 이용하여 손가락 개수 및 위치 정보를 계산한다. 특히, 제안하는 방법에서 지속적 사용자 명령은 제스처 단일 형상의 방향 이동에 의해서 반복 수행이 가능하고, 손가락 위치 정보는 사용자 경험(UX: User eXperience)에 기반 하여 사용자 명령을 증폭 수행한다. 또한 Intel社의 OpenCV 라이브러리를 기반으로 C++언어를 사용하여 구현된 손동작 컨트롤러 인터페이스의 성능 평가를 위해서는 상용 비디오 플레이어 프로그램에 적용하여 각 컨트롤 모드별 인식률을 측정한 결과 평균 89%의 인식률을 보였다.

Abstract A finger-gesture remote controller using the single camera is implemented in this paper, which is base on the recognition of finger number and finger moving direction. Proposed method uses the transformed YCbCr color-difference information to extract the hand region effectively. The number and position of finger are computer by using a double circle tracing method. Specially, a user continuous-command can be performed repeatedly by recognizing the finger-gesture direction of single shape. The position information of finger enables a user command to amplify a same command in the User eXperience. Also, all processing tasks are implemented by using the Intel OpenCV library and C++ language. In order to evaluate the performance of the our proposed method, after applying to the commercial video player software as a remote controller. As a result, the proposed method showed the average 89% recognition ratio by the user command-mode.

Key Words : Finger-Gesture, Double Circle Tracing, Moving Direction Recognition, User Experience, Remote Controller

1. 서 론

컴퓨터 비전 기반의 손 동작 인식을 이용하는 HCI 방

식은 채널 변경이나 볼륨 컨트롤과 같은 지속적인 명령의 반복에 수행에 적합하다. 따라서 스마트 TV나 콘솔 게임기 등의 기기들에서 기존의 물리적 컨트롤러 장치들

*준회원, 상지대학교 컴퓨터정보공학부

**정회원(교신저자), 상지대학교 컴퓨터정보공학부

접수일자 : 2013년 3월 7일, 수정완료 : 2013년 6월 29일

게재확정일자 : 2013년 8월 16일

Received: 7 March, 2013 / Revised: 29 June, 2013 /

Accepted: 16 August, 2013

**Corresponding Author: beomlee@sangji.ac.kr

School of Computer Information Engineering, Sangji University, Korea

대신하여 사용자의 제스처 인식에 기반을 두는 리모트 컨트롤러 인터페이스 개발은 중요한 연구 분야로 주목받고 있다. 그러나 손은 관절로 구성되어 있어 동일한 동작의 표현에도 동작을 행하는 사람에 따라서 다양한 표현이 가능하기 때문에 인식문제에 있어서 어려움이 있으며, 현재까지 그 해법에 관한 많은 연구가 진행되고 있다^{[1][2]}.

이러한 연구 가운데에서 비착용형 방식의 제스처 인식 방법은 카메라의 성능에 영향을 많이 받는 단점에도 불구하고, 추가적인 장치 없이 인식이 가능하며 구현에도 추가적인 비용이 발생하지 않기 때문에 활발히 연구되고 있다^{[1][3][8]}.

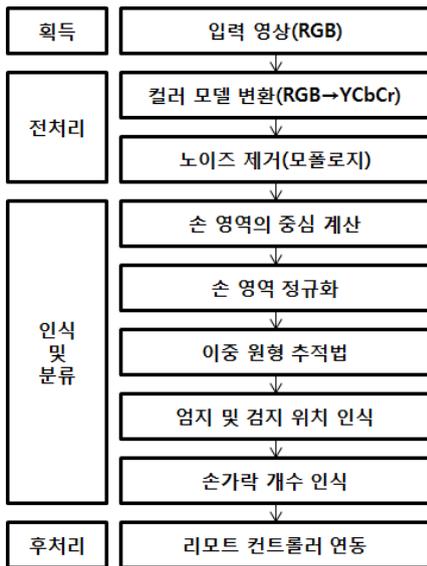


그림 1. 본 논문에서 제안하는 리모트 컨트롤러의 처리과정
Fig. 1. Remote Controller Processing Tasks in Our Approach

따라서 본 논문에서는 카메라의 성능에 대한 영향을 줄이기 위해 형상 변화에 의한 특징정보를 추출하는 방법이 아닌 단일 손가락 제스처 형상의 이동 방향을 인식하는 리모트 컨트롤러를 구현한다.

단일 형상의 손가락 위치 및 이동 방향 인식을 위해서 YCbCr 컬러모델로부터 색차정보를 이용하여 손 영역을 추출한다. 그리고 추출된 손 영역에서 중심점을 계산하여 정규화한 후 반지름이 다른 두 개의 원과의 교차정보를 이용하는 이중 원형 추적법을 이용하여, 엄지와 검지의 유무 및 위치 정보를 검출하고 검지의 이동 방향을 인식하여 리모트 컨트롤러의 기능과 연동한다.

본 논문에서 구현하는 실시간 리모트 컨트롤러는 비교적 단순하고 저비용의 구현 가능하고, 특히 사용자 경험(UX: User eXperience) 기반의 기능을 포함하기 때문에 직관적인 인터페이스의 구현이 가능하다. 그리고 구현된 리모트 컨트롤러는 상용 미디어 플레이어에 적용하여 그 성능의 효율성을 보인다.

II. 손 영역 추출

RGB 컬러모델에서의 영상처리는 R, G, B 컬러요소 값을 각각 처리해야 하기 때문에 정보 처리량이 많고, 밝기 변화에 민감하여 비효율적이다. 반면에 YCbCr 컬러모델은 밝기(Y: Luminance)와 색차신호(Cb: 푸른색의 강도, Cr: 붉은색의 강도)에 기반을 두는 색 표현 방식으로 색차 신호만으로 영상처리를 하기 때문에 정보 처리량도 감소하고 밝기 변화에 의한 영향을 효과적으로 감소시킬 수 있다^{[4][5]}.

따라서 본 논문에서는 손동작 인식을 위한 전처리 과정으로서 손 영역 추출을 위하여 원 영상을 YCbCr 컬러모델로 변환하여 사용한다. YCbCr 컬러모델로의 변환을 위해서는 아래 식 (1)을 이용한다^{[4][5]}.

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

그림 2의 (c)와 같이 RGB 컬러모델 영상에서 실험에 의해 구해진 손 영역에 해당하는 임계값으로 손 영역을 추출한 결과 영상은 밝기에 의한 영향으로 인해 주변 잡음이 심하게 나타나는 것을 볼 수 있다. 그러나 RGB 컬러모델을 YCbCr 컬러모델로 변환한 후에 획득된 손 영역의 결과 영상인 (d)를 보면 밝기에 의한 영향이 효과적으로 감소되는 것을 볼 수 있다.

추출한 손 영역 영상으로부터 잡음을 줄이기 위해서는 영상 모폴로지 변환(morphological transformation)을 사용한다. 모폴로지 변환의 팽창(dilation)과 침식(erosion) 연산은 영상 내의 잡음 제거 또는 구성 요소들의 결합 또는 분리 등을 위하여 많이 사용된다^{[6][7]}.

$$I_E(x, y) = \underset{(x', y') \in K}{MIN} I_S(x+x', y+y') \quad (2)$$

$$I_D(x, y) = \underset{(x', y') \in K}{MAX} I_S(x + x', y + y') \quad (3)$$

식 (2)와 (3)에서 $I_S(\cdot)$ 는 입력 영상의 화소값을 의미하며, $I_E(\cdot)$ 과 $I_D(\cdot)$ 는 각각 침식과 팽창 변환된 영상을 나타낸다. 그리고 $MIN(\cdot)$ 와 $MAX(\cdot)$ 는 각각 형태소 K 에 상응하는 화소들 가운데 최대값과 최소값을 추출하는 함수이다.

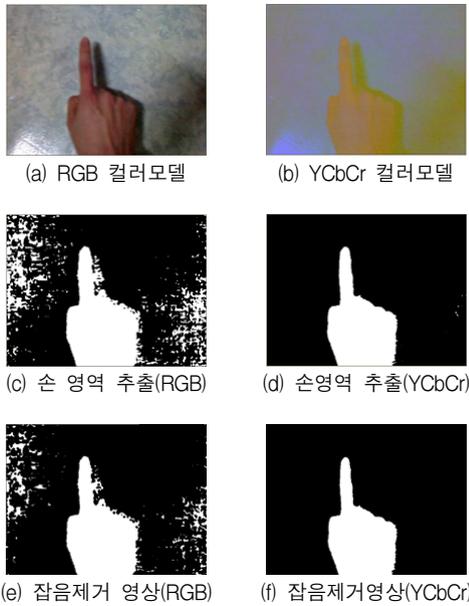


그림 2. RGB와 YCbCr 컬러모델 비교
Fig. 2. RGB vs. YCbCr Color Models

III. 손가락 인식

추출된 손 영역으로부터 사용자 제스처를 인식하기 위해서는 먼저 본 논문에서 특징으로 정의한 손가락 개수의 분석이 수행되어야 한다. 본 논문에서 손가락 개수의 인식은 손 영역의 중심에서 정해진 반지름을 갖는 원과의 교차정보를 통해 이루어지며 이를 하기 위해서는 먼저 손 영역의 중심 $C(x, y)$ 를 구한다.

손 영역의 중심은 아래 식 (4)와 같이 손 영역의 X와 Y축 좌표의 위치 값을 각각 더한 후에 전체 화소 수로 나눈으로써 정의한다.

$$C(x, y) = \left(\frac{\sum_{i=1}^N x_i}{N}, \frac{\sum_{i=1}^N y_i}{N} \right) \quad (4)$$

식 (4)에서 x_i, y_i 는 각각 영상에서의 i 번째 화소의 x, y 좌표 값을 의미하고, N 은 손 영역을 구성하고 있는 전체 화소수를 나타낸다. 그림 3과 4에서 손 내부의 빨간색 점은 식 (4)에 의해 손 영역으로부터 구해진 손 영역 중심을 나타낸다.

손 영역의 중심 값을 구한 다음에는 손가락 추출에 필요한 동일한 추적 원의 반지름 설정을 위하여 손 영역의 크기에 의한 정규화를 수행한다.

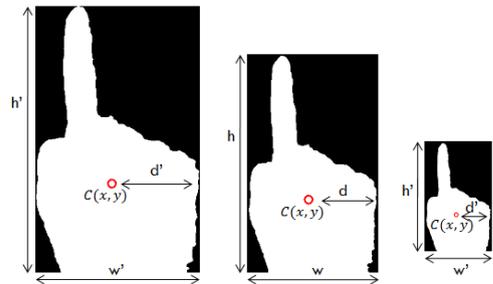


그림 3. 정규화 과정의 예
Fig. 3. Example of Normalization Processing

정규화는 손 영역의 중심에서부터 손 영역의 우측 끝까지의 거리를 이용하며, 이 거리를 이용하는 이유는 손가락 한 개를 펼 영상과 다섯 개 모두를 펼 영상에서 그 값의 차이가 없기 때문이다. 정규화에 의한 영상의 높이 (Height)와 폭(Width)은 아래 식 (5)와 같이 정의한다.

$$h = h' \frac{d}{d'}, w = w' \frac{d}{d'} \quad (5)$$

식 (5)에서 h, w 는 각각 정규화 과정을 거친 후 영상에서의 높이와 폭을 의미하고, h', w' 는 각각 정규화 과정 전 영상에서의 높이와 폭을 의미한다. d 는 정규화를 위한 지정 상수 값으로 손 영역의 중심에서 우측 끝까지의 거리에 해당하며, d' 는 정규화 과정 전 영상에서 손 영역의 중심에서의 우측 끝까지의 거리를 의미한다. 그림 3은 식 (5)에 의해 손 영역에 대한 확대 및 축소 정규화 과정의 예이다. 이상의 정규화 과정은 단일 카메라에서 거리에 의한 손 영역의 크기 영향을 효과적으로 감소

시킬 수 있다.

정규화 과정 후에 손가락 인식을 위해서는 원형 추적법을 이용한다. 그러나 단일 원과의 교차정보만을 추출하게 되면 잡음과 원이 교차되는 의도하지 않은 상황이 발생할 수 있기 때문에 본 논문에서는 반지름의 길이가 다른 두 개의 원을 사용한다.

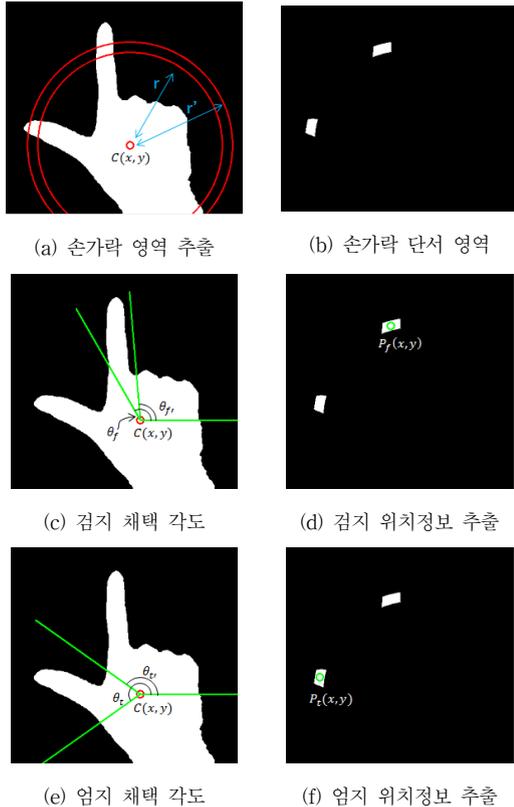


그림 4. 이중 원형 추적에 의한 손가락 단서 영역 추출
Fig. 4. Finger Clue Region Extraction by Double Circle Tracing Method

본 논문에서 제안하는 이중 원형 추적법(Double Circle Tracing)은 두 원 사이의 손가락에 해당하는 영역을 모두 추출하여 임계값 이하의 영역들은 제거함으로써 잡음에 의한 오검출되는 경우를 없앨 수 있다. 그림 4의 (a)와 (b)는 반지름의 길이가 다른 두 개의 원(각각 반지름이 r, r' 인 원)을 이용한 이중 원형 추적에 의해서 손가락 단서(Clue) 영역을 추출한 예이다. 손가락 단서 영역이 추출된 후에는 이 단서 영역으로부터 단일 형상의 이동 방향 인식에 중요한 엄지 및 검지의 위치 정보를 추

출한다.

손가락과 원의 교차정보를 추출할 때에 교차정보가 정해진 그림 4의 (c)와 같이 검지 채택 각도 $\theta_f \sim \theta_f'$ 안에 속한다면 이 손가락은 검지로 분류하고, (e)와 같이 엄지 채택 각도 $\theta_e \sim \theta_e'$ 안에 속한다면 이 손가락은 엄지로 분류한다. 그리고 검지 및 엄지의 위치 정보 $P_f(\cdot)$ 와 $P_e(\cdot)$ 는 이중 원형 추적에 의한 교차 흔적(Blob) 가운데에서 검지 영역과 엄지 영역의 중심값으로 식 (4)와 동일한 방법으로 계산한다.

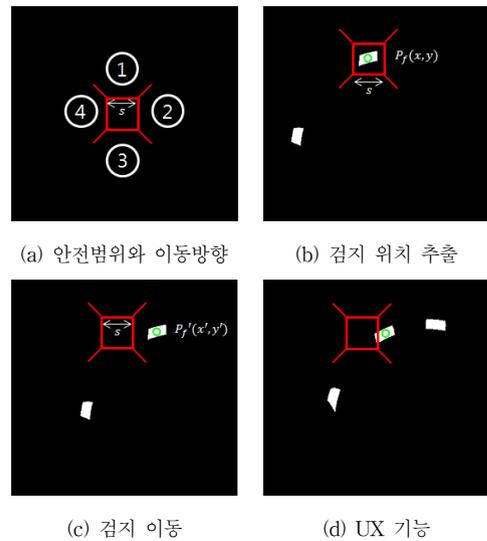


그림 5. 단일 형상의 이동 방향 인식
Fig. 5. Moving Direction Recognition of Single Shape

단일 형상의 이동 방향을 인식하기 위해서는 그림 5와 같이 검지 위치 정보인 $P_f(\cdot)$ 의 이동을 분석한다. 그림 5의 (a)에서 붉은 색의 사각형 영역은 안전 범위(Safe Zone)를 나타낸다. 안전 범위는 단일 형상의 방향 인식에서 계속 움직이는 손동작 때문에 의도하지 않은 명령들의 입력 문제를 보완할 수 있다. 따라서 검지의 이동은 안전 범위를 벗어난 경우에만 인지되며, 그 방향은 그림 5의 (a)에서처럼 4 방향으로 구분된다.

검지의 4-이동 방향 인식은 아래 식 (6)과 표 1의 검지 이동 분류표에 의해서 수행된다.

$$D = AMAX(dx, dy), \text{ when } |D| > \frac{s}{2} \quad (6)$$

where, $dx = x' - x, \quad dy = y' - y$

표 1. 검지 이동 인식표
Table 1. Forefinger Moving Recognition Table

D		부호	
		양(+)	음(-)
선택값	dx	②	④
	dy	③	①

식 (6)에서 s 는 안전 범위를 지정하는 상수 값이고, (x, y) 와 (x', y') 는 그림 5에 나타낸 좌와 같이 각각 검지의 현재 위치 정보 $P_f(x, y)$ 와 이동 후의 위치 정보 $P_f'(x', y')$ 를 나타내는 좌표이다. 그리고 $AMAX(\cdot)$ 는 절대값이 큰 인수 값을 반환하는 함수이다.

검지의 이동 방향이 인식되면, 인식한 방향에 따른 이벤트를 발생시킨다. 이때 일정 시간마다 지속적인 이벤트를 발생시켜 지속적인 명령의 반복수행이 가능하도록 구현한다. 이 지속적인 명령의 반복수행은 검지의 위치가 다시 안전 범위 안으로 들어가거나 엄지가 인식되지 않을 때 또는 제스처 모양이 변하기 전까지 지속되며 검지의 방향이 이동되어 다른 방향으로 인식되면 그 방향에 따른 이벤트가 발생하고 그에 해당하는 지속적인 명령을 반복한다.

특히 엄지가 계속 인식되고 있을 때 손가락의 개수가 3개로 증가하면 이에 해당하는 이벤트가 발생하고 해당 명령을 반복 수행한다. 그러나 이 경우에는 인식 방향은 같고 제스처만 변경되는 상황이기 때문에 사용자 경험에 기반하여 동일 기능의 증폭 명령을 수행한다.

그림 5의 (d)는 사용자 경험 기반을 적용한 예로 엄지가 계속 인식되고 있으면서 단순히 손가락이 하나 증가한 경우로서 기존에 저장해 둔 검지의 위치는 아직 저장되어 있다. 따라서 초기 명령으로 피드백 되지 않고 해당 기능을 증폭 수행한다.

IV. 제스처 분류 및 인식

본 논문에서 구현하는 손 동작 리모트 컨트롤러의 기능을 수행하기 위해서는 먼저 사용자 명령에 따른 제스

처를 그림 6과 같이 정의하고 표 2과 같이 그 기능을 상세화 한다.

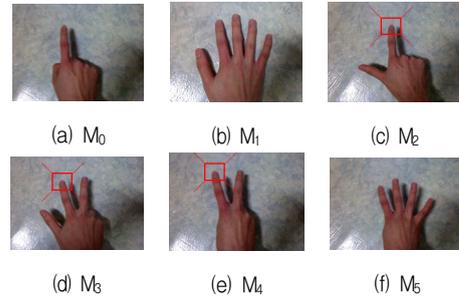


그림 6. 손동작 리모트 컨트롤러를 위한 제스처 인식 모드
Fig. 6. Gesture Recognition Mode for Finger-Gesture Remote Controller

표 2. 손동작 리모트 컨트롤러의 기능표
Table 2. Function Table for Finger-Gesture Remote Controller

방향 모드	East	West	South	North
M ₀	Standby Status			
M ₁	Play or Pause Toggle key			
M ₂	FF	Rewind	Vol. D	Vol. U
M ₃	x n(5)	x n(5)	x n(5)	x n(5)
M ₄	Next	Prev.	None	None
M ₅	Mute key			

그림 6에서 (a) M₀는 기본 대기 모드로 검지의 위치를 계속 찾지만 저장은 하지 않는다. (b) M₁는 동영상 플레이어가 정지 상태일 때는 재생상태로, 재생상태일 때는 일시정지 상태로 변환하는 토글 키(Toggle key) 기능을 수행한다. 그리고 단일 형상의 이동 방향 인식에 의해서 (c) M₂는 검지의 이동방향에 따라 위쪽은 소리를 키우고 아래쪽은 소리를 줄이고 오른쪽은 동영상을 앞으로 감고 왼쪽은 뒤로 감는 기능을 수행한다.

특히 사용자 경험 기반을 적용한 (d) M₃는 M₂의 기능을 5배 증폭된 값으로 해당 기능을 수행한다. (e) M₄는 검지의 이동방향에 따라 오른쪽은 다음파일을 재생시키는 기능을 왼쪽은 이전파일을 재생시키는 기능을 수행하고, (f) M₅는 음소거 기능을 수행한다.

이상에서 정의된 명령 제스처는 그림 7에 나타낸 인식 과정에 의해서 분류된다. 그림 7에서 명령 인식은 M₀에

서 시작하고, Err.는 예외처리(Exception Handling)를 의미한다. N_f 는 인식된 손가락의 개수를 나타내는 변수이고, T_b 는 엄지 손가락의 유무를 나타내는 이진 변수이다. 또한 T_n 은 명령 모드 인식을 위한 중간 인식 단계를 의미한다.

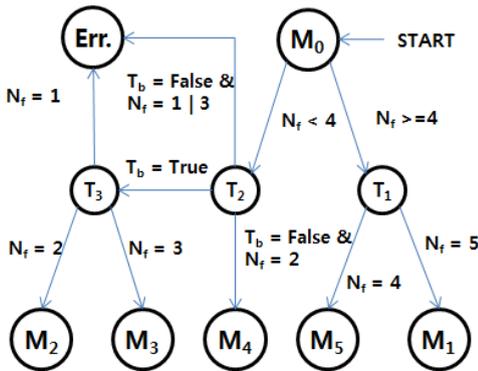


그림 7. 제스처 명령 모드 분류 체계
Fig. 7. Gesture Command Mode Classification Scheme

V. 실험

본 논문에서 구현한 실시간 리모트 컨트롤러 인터페이스의 성능평가를 위해서 Intel社의 컴퓨터비전 라이브러리인 OpenCV^[9]를 기반으로 MS Windows 7 환경 하에서 C++ 언어를 사용하여 각 처리 단계를 구현하였다. 손동작 입력 장치로는 일반적으로 사용되는 저가의 웹캠을 사용하였다. 실험에 상용된 입력 영상은 640x480 화소(pixel)이며 정규화에 사용된 d 값은 80 화소이다. 그리고 이중 원형 추적을 위한 원형 반경은 각각 150 화소와 170 화소의 원을 이용하였으며, 검지 채택 각도와 엄지 채택 각도는 각각 $95^{\circ} \sim 120^{\circ}$ 와 $150^{\circ} \sim 255^{\circ}$ 로 설정하였다. 그리고 단일 형상의 이동 방향 인식에서 안전 범위 상수 s 는 10 화소로 하였다.

성능평가 방법으로는 상용 미디어 플레이어 프로그램에 구현한 실시간 리모트 컨트롤러 모듈을 적용하여 비디오 플레이어 컨트롤을 위한 각 인식 제스처 모드의 인식률을 분석하였다. 그 결과 표 3과 같이 평균 89%의 인식률로 성공적인 인터페이스 기능을 보였다. 그러나 실

험에서 M_3 와 M_4 의 인식률이 떨어지는 것을 볼 수 있었는데 이는 검지와 엄지를 각도로 찾는 방법으로 인해 손목이 돌아가면서 발생하는 검지와 엄지의 오인식 때문이다.

표 3. 리모트 컨트롤 모드별 인식률

Table 3. Recognition Ratio of Remote Controller Modes

모드	M_1	M_2	M_3	M_4	M_5	평균
인식률	92%	90%	85%	87%	91%	89%

VI. 결론

본 논문에서는 단일 형상의 이동 방향 인식을 이용하여 구현이 단순하고 직관적인 손동작 리모트 컨트롤러를 구현하였다. 구현된 손동작 리모트 컨트롤러 인터페이스는 사용자의 편의성뿐만 아니라, 저비용 구현이 가능하며 인식 성능이 높다. 또한 5가지 손가락 제스처에 대해서 제안된 명령만을 실행하도록 설계하였으나, 단일 형상의 이동 방향 인식을 적용하여 12가지로 기능 확장이 가능하다. 특히 지속적인 명령의 반복 수행과 손가락 개수의 변화에 따른 UX적 명령 증폭이 가능하다. 그러나 가변적 환경에서의 인식 성능이 떨어지는 문제점이 있었다.

향후 가변적 환경에서의 인식 성능을 개선하고 제스처 동작 위치 제한 등의 문제가 해결된다면, 스마트 홈 환경에서 한 대의 카메라를 이용하여 다수의 가전제품을 컨트롤할 수 있는 컨트롤러로 다양한 활용이 가능할 것으로 기대된다.

References

[1] R. Z. Khan and N. A. Ibraheem, "Survey on Gesture Recognition for Hand Image Posture", Computer and Information Science Vol. 5(3), pp. 110-121, 2012.

[2] Freeman, W. T., Weissman, C. D., "Television control by hand gestures", IEEE International Workshop on Automatic Face and Gesture Recognition, pp.179-183, 1995.

- [3] Hasan, M. M., Mishra, P. K. "Brightness factor matching for gesture recognition system using scaled normalization", International Journal of Computer Science & Information Technology (IJCSIT), Vol. 3(2), pp.35-46, 2011.
- [4] Neelamani, R., Queiroz, R. de., Fan, Z., et al. "JPEG Compression History Estimation for Color Images". IEEE Trans. Image Process., Vol. 15(6), pp. 1365 - 1378, 2006.
- [5] Guangxin Li, Image Fusion Based on Color Transfer Technique, Image Fusion and Its Applications, Dr. Yufeng Zheng (Ed.), InTech, 2011.
- [6] R. M. Haralick, et al, "Image analysis using mathematical morphology", IEEE Transactions on Pattern Analysis and Machine Intelligence 9(4), pp. 532-550, 1987.
- [7] G. Bradski and A. Kaehler, Learning OpenCV, O'REILLY, pp.176-180, 2008.
- [8] S. Mitra and T. Acharya, "Gesture Recognition: A Survey", IEEE Trans. on SMC Vol. 37(3), pp. 311-324, 2007.
- [9] Intel. Open source computer vision library: Reference manual. 2001.
- [10] M. S. Jang, W. B. Lee, "Implement of the Video Remote Controller By using the Finger Gesture Interface", KISPS Fall Conference Proceedings. Vol. 13(2), pp. 215-217, 2012.
- [11] C. Y. Choi, W. B. Lee, "Implement of Hand Gesture Interface using Ratio and Size Variation of Gesture Clipping Region", The Journal of The IIBC, Vol. 13(1), pp. 121-127, Feb. 2013.
- [12] X. Shen, etc., "Dynamic hand gesture recognition: An exemplar-based approach from motion divergence fields", Image and Vision Computing 30, pp. 227-235, 2012.

※ 이 논문은 2012년도 상지대학교 교내 연구비 지원에 의한 것임.

저자 소개

장 명 수(준회원)



- 2011년 : 상지대학교 컴퓨터정보공학부 (학사)
- 2013년 : 상지대학교 컴퓨터정보공학과 (석사과정)
- 관심분야 : 영상처리, 한방의료영상진단기, 컴퓨터비전

이 우 범(정회원)



- 1995년 : 영남대학교 컴퓨터공학과 학사
- 1997년 : 영남대학교 컴퓨터공학과 석사
- 2000년 : 영남대학교 컴퓨터공학과 박사
- 2000년 3월~2004년 2월 : 대구과학대학교 컴퓨터공학과 교수
- 2004년 3월~2007년 2월 : 영남대학교 전자정보공학부 컴퓨터전공 특임교수
- 2007년 3월~현재 : 상지대학교 컴퓨터정보공학부 교수
- 관심분야 : 뉴로비전컴퓨팅, IT융합의료기(한방의료영상), 제스처인식