

모멘트특징량을 리용한 손짓모양검출방법

정일식, 김정철

사람-컴퓨터호상작용(HCI)은 사람들의 생활에 컴퓨터를 리용한 많은 처리기술들이 보급되는것과 더불어 매우 중요한 부분으로 되고있다. 살림방과 침실, 사무실 등 그 어디에나 여러가지 기능들을 수행하는 지령들을 필요로 하는 전자제품들이 설치되어있다.

오늘날 거의 모든 전자제품들에는 그 사용자를 위한 원격조종기술이 리용되고있다.[1] 실제로 텔레비전을 볼 때 통로절환과 색조절, 음량조절 등을 비롯하여 심지어 전등을 켜거나 끄는 생활세부에도 원격조종기술들이 포함되어있다. 그렇지만 이러한 원격조종기들은 그 물리적모양의 다양성과 풍부한 기능, 단추배열의 혼잡과 기타 조건(눈이 나쁜 사람들 특히 늙은이들인 경우 시력장애) 등으로 하여 필요한 조종을 위한 정확한 명령건선택의 곤란과 같은 여러 문제점들을 산생시키고있다. 이러한 문제점을 극복하기 위하여 보다 편리하고 효과적인 기술들이 개발되고있는데 그것은 손짓모양식별에 기초한 원격조종기술이다. 이 기술을 리용한 체계들은 사용자의 손짓모양을 검출 및 식별하고 그 모양을 전자제품 혹은 전자장치들을 조종하도록 미리 정의된 지령들로 해석하고 동작하게 한다. 그렇지만 일반적으로 전자제품들은 가정용 혹은 실내용이 대부분으로서 이것들은 조명이 약하거나 조명변화가 심한 환경에서 동작하는것만큼 이러한 조명조건하에서 안정한 손모양검출을 진행하는 것이 매우 중요하다.

본문에서는 모멘트특징량에 기초한 특징검출을 리용하여 백열등, 형광등과 같은 약한 조명조건에 대하여 안정한 손짓모양을 검출하기 위한 한가지 방법을 제거하였다.

1. 턱값려파에 의한 손령역검출

실시간동영상스트림으로부터 얻어진 화상에서 화상의 배경을 제거한 후에 피부화상령역분할을 진행하여 손령역을 검출한다. 이를 위하여 본문에서는 턱값려파에 의한 비피부령역제거방법을 제안하였다.

우선 얻어진 RGB형식의 화상을 YCbCr형식으로 변환한 다음 비피부성분들을 제거하기 위하여 턱값에 의한 려파를 진행한다. 턱값은 미리 준비한 자체의 피부색자료모임을 리용하여 결정하는데 피부령역을 표시하는 화소들의 특징들을 찾기 위해 많은 화상들중에서 몇개의 화상들을 우연적으로 선택하고 YCbCr형식으로 변환한다. 다음 변환된 화상의 매 화소들에 대하여 그것의 관련값(x, y 자리표값, 세기값 등)들의 분포를 결정한다. 화소들의 색분포를 볼 때 피부조각들과 비피부조각들은 전체 령역에 걸쳐 자기의 색값들을 가지는데 특히 Y 성분에 대하여 피부조각의 색값은 110~165인 반면에 비피부조각은 16~208이다. Y 성분의 이 성질은 화상의 밝기값으로는 피부령역을 려파할수 없고 다른 2개의 나머지

성분들 즉 Cb, Cr으로 러파할수 있다는것을 의미한다. 그림 1은 YCbCr공간에서 피부조각들과 비피부조각들을 표시하는 표본점들의 Cb-Cr색분포를 보여주고있다.

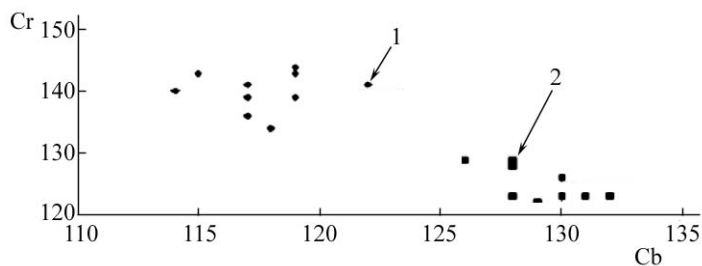


그림 1. 피부화소들과 비피부화소들의 Cb-Cr색분포

1-피부화소, 2-비피부화소

그림 1에서 알수 있는바와 같이 피부와 비피부조각들의 Cb, Cr값들은 그래프에서 명백히 구별되어 분포된다. 그러므로 러파기의 턱값은 쉽게 결정될수 있다. 백열등과 형광등 빛조건에서 임의의 다른 화상들에 대하여 실험을 반복한 결과 예상한대로 턱값들은 표 1에서 보여준것처럼 아주 차이났다. 조명조건이 나쁜 다른 화상들에 대하여서도 실험을 진행한 결과 턱값러파에 의해서 손령역이 명백히 구별되어 검출되었다.

표 1. 백열등빛과 형광등빛조건에 대한 턱값

백열등빛	형광등빛
$100 \leq Cb \leq 122$	$114 \leq Cb \leq 128$
$132 \leq Cr \leq 150$	$140 \leq Cr \leq 158$

그런데 손령역을 턱값러파할 때 잡음과 이지러짐이 생긴다.(그림 2)

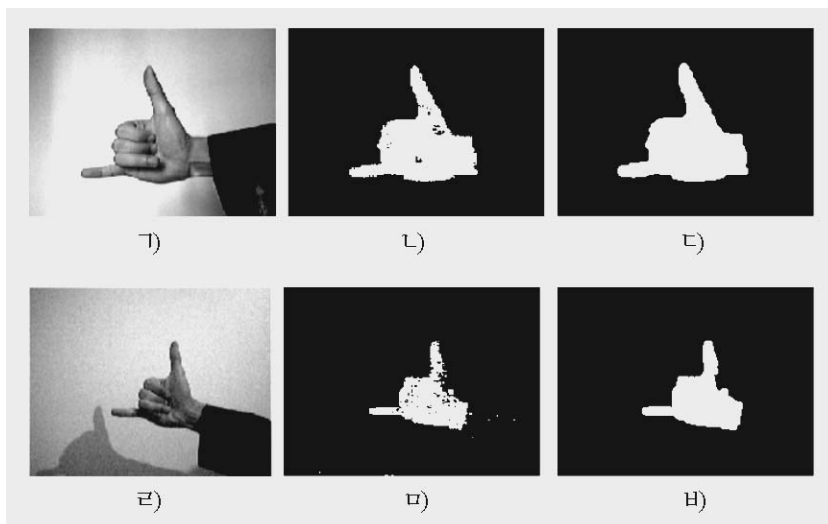


그림 2. 턱값러파를 리용한 화상의 피부색령역분리

가) 보통 조명조건때 원화상, 나) 보통 조명조건때 분리된 피부령역,
다) 보통 조명조건때 평활화한 피부령역, 라) 나쁜 조명조건때 원화상,
마) 나쁜 조명조건때 분리된 피부령역, 바) 나쁜 조명조건때
평활화한 피부령역

이러한 잡음과 이지러짐은 형태학적러파(Morphological Filtering) 즉 확장법(Dilation)과 침식법(Erosion)을 결합하여 제거할수 있다.[3]

제안한 방법의 주요한 우점은 변환처리때 조명의 영향을 없앨수 있다는것이다.

2. 모멘트특징량에 의한 특징검출

모멘트특징량에 기초한 특징추출방법은 물체인식을 비롯한 여러 인식분야에서 《구별력있는》 특징추출을 위한 가장 효과적인 방법들중의 하나로 알려져있다.[2] 이 방법은 2진화된 객체화상으로부터 많은 자체의 특성들을 유도해내는데 이 특성들은 회전과 척도변환, 평행이동에 대하여 불변이다.

정의 $f(i, j)$ 를 크기가 $M \times N (i=1, 2, \dots, M, j=1, 2, \dots, N)$ 인 수자식화상의 한 점이라고 할 때 2차원모멘트들과 $f(i, j)$ 의 $(p+q)$ 차원중심모멘트들은 다음과 같다.

$$m_{pq} = \sum_{i=1}^M \sum_{j=1}^N i^p j^q f(i, j) \quad (1)$$

$$U_{pq} = \sum_{i=1}^M \sum_{j=1}^N (i - \bar{i})^p (j - \bar{j})^q f(i, j) \quad (2)$$

여기서

$$\bar{i} = \frac{m_{10}}{m_{00}}, \quad \bar{j} = \frac{m_{01}}{m_{00}}$$

이다.

여기로부터 표준화된 중심모멘트 η_{pq} 는 다음과 같이 표시할수 있다.

$$\eta_{pq} = \frac{U_{pq}}{U_{00}^r}$$

여기서 $r = (p+q)/2+1, p+q=2, 3, \dots$ 이다.

이 2차원 및 3차원모멘트들로부터 7개의 모멘트특징량들은 다음과 같이 유도된다.

$$\phi_1 = \eta_{20} + \eta_{02} \quad (3)$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (4)$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad (5)$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (6)$$

$$\phi_5 = (\eta_{30} - 3\eta_{12})^2 (\eta_{30} + \eta_{12}) [(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})^2 (\eta_{21} + \eta_{03}) [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (7)$$

$$\phi_6 = (\eta_{20} - \eta_{02}) [(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11} (\eta_{30} + \eta_{12}) (\eta_{21} + \eta_{03}) \quad (8)$$

$$\phi_7 = (3\eta_{21} - \eta_{03}) (\eta_{30} + \eta_{12}) [(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] - (\eta_{30} - 3\eta_{12}) (\eta_{21} + \eta_{03}) [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (9)$$

회전 및 척도변환, 평행이동, 잡음추가 등을 포함한 손짓에 대한 시험화상들을 그림 3에 보여주었다. 그리고 식 (3)–(9)를 리용하여 계산한 모멘트특징값들을 표 2에 보여주었다.

표 2. 손짓에 대한 모멘트특징값들

	A1	A2	A3	A4
ϕ_1	0.340 28	0.319 44	0.340 28	0.350 41
ϕ_2	0.043 403	0.043 403	0.043 403	0.044 013
ϕ_3	0.023 148	0.023 148	0.023 148	0.007 648
ϕ_4	0.002 572	0.002 572	0.002 572	3.54×10^{-5}
ϕ_5	-5.56×10^{-6}	-5.56×10^{-6}	-5.56×10^{-6}	-0.000 487
ϕ_6	-0.000 15	-0.000 15	-0.000 15	-0.023 819
ϕ_7	1.91×10^{-5}	1.91×10^{-5}	1.91×10^{-5}	2.03×10^{-7}

표 2로부터 알수 있는바와 같이 손짓의 첫 3개에 대해서 모멘트특징량이 같은 값을 가진다. 오직 1개의 모멘트값 즉 ϕ_1 만이 척도변화에 대하여 6.1%의 오차를 보여주었다. 기타 모멘트값들은 3개의 그림에 대하여 거의 동일하다. 다만 마지막 경우 즉 잡음을 포함한 손짓에 대해서 결과가 매우 차이난다. 이것은 모멘트특징량이 잡음에 아주 민감하다는것을 보여주고있는데 바로 이것이 이 방법의 결함이다.

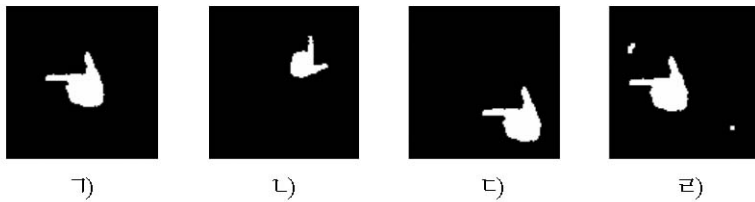


그림 3. 손짓에 대한 회전, 척도변환, 평행이동 및 잡음추가

가) 표준, 나) 회전 및 척도변환, 다) 평행이동, 라) 잡음추가

이상의 손짓에 대한 응용실험에서 알수 있는바와 같이 모멘트특징량검출방법은 회전 및 척도변환, 평행이동변환들에 대해 아주 효율적이므로 이것을 리용하여 여러 손짓모양들을 분류하는 간단한 분류기를 작성할수 있다.

3. 손짓식별방법

사용자의 손짓을 전자장치를 조종하기 위하여 미리 설정한 지령으로 해석할수 없다면 아무런 의미가 없게 된다. 따라서 손짓령역검출뿐아니라 그 의미를 식별하는 손짓분류를 진행하여야 한다. 그 분류과정은 신경망학습을 통하여 실현할수 있다.

설계된 신경망은 BP신경망으로서 사용자손짓의 표본모임에 대한 불변모멘트를 입력벡토르로 하고 지령모임들을 대응한 출력벡토르로 하여 그것들사이의 어떤 함수를 근사화할수 있을 때까지 반복적으로 학습시켰다. 또한 분류할수 있는 손짓수가 제한되어있기때문에 신경망의 깊이를 세층으로 제한하였다.

신경망분류에서는 20가지 손짓을 표현하는 995장의 표본화상들로 학습을 진행하여 분류정확도가 90%이상 되는 7가지 손짓을 분류하였다.

맺 는 말

실내조명환경(백열등빛, 형광등빛 등)과 조명이 약한 환경에서 조명변화에 대하여 안정한 손짓영역검출방법을 제안하였다. 즉 모멘트특징량에 의한 특징검출을 진행하고 BP신경망으로 20가지 손짓을 포함한 995장의 표본화상들에 대한 분류를 진행하여 분류정확도가 90%이상인 손짓 7개를 분류하였다.

참 고 문 헌

- [1] T. Baudel et al.; Commun. ACM, **36**, 7, 28, 1993.
- [2] P. Premaratne; CSSIP-DSTO Internal Publication, 53~60, 2003.
- [3] R. C. Gonzalez et al.; Digital Image Processing using MATLAB, Pearson Prentice-Hall, 101~120, 2004.

주체103(2014)년 7월 5일 원고접수

A Method of Hand Gesture Detection based on Moment Invariants

Jong Il Sik, Kim Jong Chol

This paper presents a novel method of detection and classification of hand gestures based on moment invariants to be invariant to translation, scaling and rotation, which is non-sensitive low lightning condition.

Key words: hand gesture detection, skin detection, moment invariant