# 에지 방향성 히스토그램과 HMM을 이용한 제스처 인식에 관한 연구

이기준*

## A Study on Gesture Recognition using Edge Orientation Histogram and HMM

Kee-Jun Lee*

요 약

본 논문에서는 에지 방향성 히스토그램과 주성분 분석을 통해서 얻어진 특징 정보를 저차원 제스처 심볼로 구성하여 제스처를 인식하는 알고리즘에 대해 기술한다. 제안된 방법은 기존의 기하학적인 특징 기반 방법이나 외관기반 방법에 비해 많은 계산 량이 요구 되지 않고 최소한의 정보를 사용하고도 높은 인식률을 유지 할 수 있기에 실시간 시스템 구축에 매우 적합하다. 또한 제스처 인식 시 발생하는 잘못된 인식이나 인식 오차를 줄이기 위해 객체 공간상에 투영된 모델 특징 값을 은닉마르코프 모델의 입력 기호로 이용되기 위해서 군집화 알고리즘을 통해 특정한 상태 기호로 구성한다. 이렇게 함으로써 임의의 입력 제스처는 확률 값이 가장 높은 해당 제스처 모델로 인식하게 된다.

ABSTRACT

In this paper, the algorithm that recognizes the gesture by configuring the feature information obtained through edge orientation histogram and principal component analysis as low dimensional gesture symbol was described. Since the proposed method doesn't require a lot of computations compared to the existing geometric feature based method or appearance based methods and it can maintain high recognition rate by using the minimum information, it is very well suited for real-time system establishment. In addition, to reduce incorrect recognition or recognition errors that occur during gesture recognition, the model feature values projected in the gesture space is configured as a particular status symbol through clustering algorithm to be used as input symbol of hidden Markov models. By doing so, any input gesture will be recognized as the corresponding gesture model with highest probability.

키워드

주성분 분석, 에지 방향성 히스토그램, 은닉 마르코프 모델

Key word

PCA(Principal Component Analysis), EOH(Edge Orientation Histogram), HMM(Hidden Markov Models)

* 정회원 : 광주보건대학 보건교육정보과(leekj@ghc.ac.kr)

접수일자 : 2011. 09. 19
심사완료일자 : 2011. 10. 08

## Ⅰ. 서 론

컴퓨터 기술의 발달과 함께 정보 시스템이 복잡하게 되면서 인간과 정보 시스템 사이에 자연스럽게 정보를 교환할 수 있는 지적 인터페이스에 관한 관심이 날로 커지고 있으며 특히 컴퓨터 비전기술을 기반으로 하는 다양한 대화형 시스템을 구축하기 위한 연구들이 활발히 진행되고 있다. 특히 인간은 일상생활에서 제스처, 표정과 같은 비언어적인 수단을 이용하여 수많은 정보를 전달한다. 따라서 자연스럽고 지적인 인터페이스를 구출하기 위해서는 제스처와 같은 비언어적인 통신 수단에 대한 연구가 매우 중요하다.

이러한 분위기 속에서 보다 인간이 사용하기 쉬운 새로운 사용자 인터페이스 제작 기술로서 제스처 인식 기술이 주목을 받고 있다[1]. 제스처란 손가락, 손, 팔, 머리 또는 몸의 신체적인 움직임 포함하는 의미 있는 신체의 동작을 말하는데 여기에는 움직이고 있는 동작이나 정지되어 있는 상태까지 포함한다[2]. 제스처 인식은 크게 센서에 기반한 제스처 인식, 영상 기반의 제스처 인식으로 구분된다[3]. 센서 기반의 제스처 인식은 움직임의 위치와 방향을 얻기 위한 방법으로 기계식 방식, 관성 방식, 자기 방식, 광학 방식 등이 사용되며 이와 같은 센서 기반의 방식은 사람의 몸에 수신기를 부착하기 때문에 동작을 제한하는 단점이 있지만 정교한 제스처나 포즈 인식이 가능하다. 영상 기반의 제스처 인식은 주로 비디오카메라와 컴퓨터 비전 기술을 이용하여 정보를 획득하는 방식으로 영상 그 자체를 이용하거나 카메라의 입력 영상으로부터 기하학적인 특징을 추출하는 방식이다. 영상 그 자체를 사용하는 경우 계산양이 적어 실시간 인식에 적합하나 일반성이 결여되고 기하학적 특징을 이용하는 경우는 정적인 제스처를 인식하기 위해 이용되고 있지만 많은 계산 양을 필요로 하기 때문에 실시간 인식에는 부적합하다고 논의되어지고 있다.

그래서 본 논문에서는 입력 시퀀스의 각 영상으로부터 신체 영역을 분리한 후 제스처의 형태의 정보인 실루엣 영상을 생성한 후 에지 방향성 히스토그램과 주성분 분석을 이용하여 신체의 형상 변화를 축소된 특징 매개변수 공간에서 인식하는 방법에 대해서 기술한다. 입력된 영상은 전처리 과정을 통해 실루엣 영상으로 변환되며 이 실루엣 영상은 에지 방향성 히스토그램을 통해 36개의 특징으로 생성되고, 이 특징을 이용하여 주성분 분석법이라는 통계적 방법에 의해 인체의 외관 특징들을 표현할 수 있는 저 차원 벡터공간으로 투영된다. 각 동작들은 이 공간 내에서 순차적으로 연속된 점의 궤적으로 표현되고, 미리 학습된 모델 궤적과 입력 영상의 궤적을 비교함으로써 동작인식이 이루어지게 된다. 아울러 보다 안정된 인식을 위해 최종 인식결과를 HMM을 사용하여 판단하였다.

## Ⅱ 전체 시스템 구성도

본 논문에서 제안하는 제스처 인식 시스템의 알고리즘은 카메라를 통해 얻어진 영상에서 신체 영역만을 추출하고 영상의 크기 정규화 및 노이즈를 제거한 후 실루엣 영상을 생성 후 에지 방향성 히스토그램을 이용하여 특징을 생성한 후 PCA와 HMM를 이용하여 제스처 인식을 수행한다. 그림 1은 본 논문에서 제안하는 알고리즘의 순서도를 나타낸다.

그림 1. 시스템 전체 구성도
Fig. 1 Flowchart of the proposed System

## Ⅲ. 전처리

### 3.1 배경제거

카메라를 통하여 얻은 영상 시퀀스는 단순 배경에서 취득한 것이며, 일반 환경에서 취득한 영상에는 제스처 인식에 필요 없는 많은 배경들이 포함되어 있다. 그러나 제스처 인식에 필요한 것은 신체 영역(전경)이므로 우선 배경과 신체 영역을 분리하는 작업이 필요하고 이를 위해서는 먼저 배경 모델을 생성해야 한다. 그러나 조명의 밝기가 일정하지 않고 수시로 변하기 때문에 같은 카메라로 일정 시간 동안 똑같은 배경을 촬영한다고 할지라도 모두 동일하지 않아 안정적인 배경 모델을 얻는데 어려움이 따른다.

본 논문에서는 조명 변화로 인한 배경의 밝기 변화를 측정하여 시간 요소($t$)를 고려해서 일정 시간 $T_i$동안 배경 영상 $I_t$를 취득한 다음 영상 영역 $R$ 내에 있는 각 픽셀($x$)들을 분석하여 조명이 가장 밝았을 때의 화소 값 $P_{\max}(x)$와 가장 어두울 때의 화소 $P_{\min(x)}$을 얻는다. 이 두 화소 값의 차이 $D(x)$는 조명의 변화로 나타날 수 있는 밝기의 임계치이다. 이 3가지의 요소를 이용해 배경 모델(Background Model)을 구성한다. 이와 같은 내용은 식(1-4)와 같다[4].

$$BM=\{P_{\max}(x),P_{\min}(x),D(x)\}_{x\in R} \quad (1)$$

$$P_{\max}(x)=MaxI_t(x),(1\le t\le T_i) \quad (2)$$

$$P_{\max}(x)=MinI_t(x),(1\le t\le T_i) \quad (3)$$

$$D(x)=P_{\max}(x)-P_{\min}(x) \quad (4)$$

일단 배경 모델이 만들어지면, 이진 영상 $B(x)$는 식(5)에서 보여주는 것처럼 입력 영상 $I(x)$가장 밝은 화소 값 $P_{\max}(x)$와 가장 어두운 화소값 $P_{\min(x)}$의 차분 연산을 통해 얻은 차이 값이 임계치 $D(x)$보다 크면 최대값(255)을 그렇지 않으면 최소값(0)의 화소값을 갖게 한다.

$$B(x)=\begin{Bmatrix}255 & if|P_{\max}(x)-I(x)|\ or\ |P_{\min}(x)-I(x)|>D(x)\\ 0 & otherwise\end{Bmatrix} \quad (5)$$

식(5)는 조명으로 인해 생길 수 있는 밝기 차이는 무시하고 제스처의 포즈 변화의 움직임의 차이를 갖는 영역만 분리하는 기준이 된다. 식(2)의 결과로 얻어진 이진 영상에는 배경 모델에서 설정한 밝기 값의 임계치를 벗어나는 조명의 변화로 인해 배경임에도 불구하고 전경 영역으로 분리되어 1픽셀의 작음 점들이 포함될 수 있다. 따라서 이 잡음을 없애기 위해 모폴로지(Morpology) 연산을 통해서 잡음을 제거하였다. 한 번의 침식연산을 수행하고 이때 객체의 줄어드는 현상이 발생하기에 다시 팽창 연산을 사용하여 복원하였다.

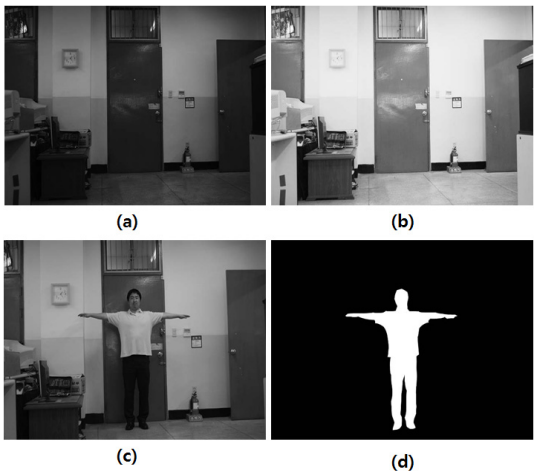


그림 2. (a) 최소 밝기의 영상 (b) 최대 밝기의 영상 (c) 입력영상 (d) 입력영상을 배경 모델 파라메타를 이용해 추출한 실루엣 영상
Fig. 2 (a) The minimum brightness of the image (b) The maximum brightness of the image (c) Input Image (d) Silhouette image extracted using the background model parameters of input image

### 3.2 에지 방향성 히스토그램 생성

앞의 절에서 설명한 방법을 통하여 얻어진 정규화 된 제스처 영상들을 이용하여 고유공간을 구성하게 되면 제스처 이미지의 크기가 640 * 480 이기에 307,200차원이라는 계산 량이 필요하게 된다. 본 논문에서는 이 차원의 크기를 줄이기 위해 에지 방향성 히스토그램을 이용하여 신체 영역이 담고 있는 36개의 특징을 생성하는 방법을 제안한다. 에지 방향성 히스토그램은 ROI(Region of Interest)영상에서 포함하고 있는 36개의 방향에 분포한 에지의 양을 표현하는 방법이다.

36방향의 에지를 얻기 위하여 하나의 영상에서 다음과 같은 에지의 $x$축과 $y$축에 대한 에지 영상 $G_x$와 $G_y$를 추출한다.

$$G_x = Sobel(I_{ROI}),\ G_y = Sobel(I_{ROI}) \quad (6)$$

그리고 식(7-8)을 통하여 픽셀$(i,j)$의 에지의 방향과 크기를 계산한다.

$$\theta_{i,j} = \arctan(G_x(i,j)/G_y(i,j)) \quad (7)$$

$$m_{i,j} = \sqrt{G_x(i,j)^2 + G_y(i,j)^2} \quad (8)$$

각 영상의 ROI의 에지 방향성 히스토그램 특징에 해당하는 각 픽셀$(i,j)$의 에지크기 $m_{i,j}$을 각 방향성별로 모두 합산을 통해 추출한다. 그림 3은 방향성 히스토그램을 생성한 모습이다.

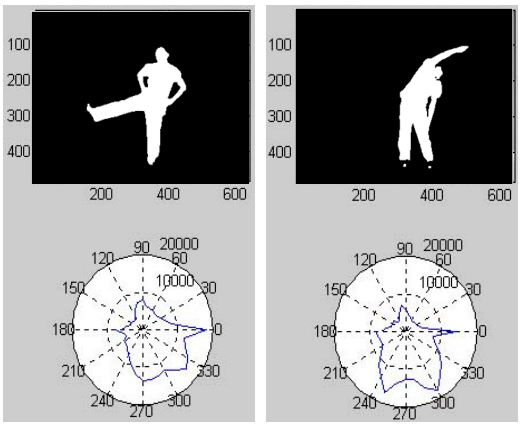

그림 3. 방향성 히스토그램
Fig. 3 Orientation Histogram

### 3.3 영상 군집화

시각적인 방법으로 얻은 영상으로부터 특징 값을 추출하여 신체의 포즈와 모션을 효과적으로 표현하고 인식하는 연구는 전역 모션 정보로부터 특징을 추출하는 방법과 머리, 손, 발과 같이 신체의 특징의 부위에 의미가 있는 부분 모션 정보로부터 특징을 추출하는 방법으로 나누어진다.

전역 특징 정보를 이용하는 방법으로, MITLab에서 사용했던 MHI(Motion History Image)와 MEI(Motion Energy Image)가 있는데 이는 Hue 모멘트 벡터를 특징정보로 사용하였다[6]. 이 방법은 모션 정보가 누적되기 때문에 우발적인 객체의 모션이 있을 경우 오 인식 될 수 있고 완전한 모션 패턴이 항상 주어져야 한다는 제약이 따른다. 따라서 이러한 단점을 보완하고자 MHI 영상으로부터 실루엣 윤곽선을 구하고, 그 정보로부터 에지 방향성 히스토그램을 통해 특징정보를 그룹화 하였다. 그림 4는 팔벌려 뛰기 모션을 그룹화한 데이터 집합이다. 여기서 $EH$는 에지 방향성 히스토그램을 뜻한다.

그림 4. 실루엣 영상에서의 에지 방향성 히스토그램 추출 및 DB구성
Fig. 4 Extraction of EOF in Silhouette image and DB configuration

## Ⅳ. 제스처 인식

### 4.1 주성분 분석을 이용한 공간생성

주성분 분석법은 고차원의 입력 데이타 집합을 저 차원의 의미 있는 데이타 집합으로 줄일 수 있다. 제스처 영상 데이터의 경우 하나의 동작을 구성하는 프레임(frame)의 수가 많고 특징을 추출하기 비교적 어렵기 때문에 빠른 인식속도와 효과적인 특징 추출이 가능한 방법을 적용해야 한다. 따라서 2장에서 추출한 연속적인 동작의 선형적 특징을 36개의 에지 방향성 히스토그램을 이용하여 저 차원 벡터로 표현하는 방법에 대해 기술한다.

앞 절에서 구한 특징 벡터를 $x$라 하고, 식(9)와 같이 표현 할 수 있다. 이 벡터의 고유공간을 계산하기 위해서

는 먼저 모든 특징 벡터의 평균 벡터를 구하여 각 특징 벡터의 차를 구해야한다. 평균 벡터 $c$와 새로운 특징 집합 $X$는 식(10-11)과 같다.

$$Robust\ Point = \sum_{n=1}^{6} BLOG(x, y, n) \tag{9}$$

$$c = (1/N)\sum_{i=1}^{N} x_i \tag{10}$$

$$X = [x_1 - c, x_2 - c, \ldots\ldots\ldots\ldots, x_N - c]^T \tag{11}$$

그런 다음, 식(12)을 만족하는 고유벡터를 구하기 위해 공분산 행렬 $Q$에 대한 고유치 λ와 고유벡터 $e$를 구하면 된다.

$$Q \overset{\Delta}{=} XX^T, \lambda_i e_i = Qe_i \tag{12}$$

고유 벡터를 구하기 위해 고유치 분해를 사용하지 않고 특이치 분해(Singular Value Decomposition)이용함으로써 특징 집합 $X$의 공분산 행렬에 대한 고유벡터를 쉽게 얻을 수 있었다. 이렇게 얻어진 고유공간에 평균 벡터 $c$를 뺀 특징 집합 $X$를 식(13)을 이용하여 투영시켰다.

$$f_j = [e_1, e_2, e_3, \ldots\ldots\ldots, e_k]^T (x_n - c) \tag{13}$$

이와 같이 얻어진 저 차원 벡터 공간을 고유공간이라고 부른다. 앞서 설명한 바와 같이 주성분 분석은 몇 개의 주성분 벡터를 유도하여 이를 통해 차원의 축소와 자료의 요약을 주목적으로 하고 있다. 따라서 전체 객체 영상을 적절하기 설명하기 위하여 보유해야 할 주성분의 수를 결정해야 한다. 식(14)는 우리가 필요한 개수를 선택하는 방법이다.

$$\left(\sum_{i=1}^{k} \lambda_i / p\right) \times 100 \geq 70 \tag{14}$$

여기서 $\lambda_i$는 $i$번째 고유값, $p$는 전체 고유값 개수라고 할 때 전체 분산 중 주성분 $C_i$가 설명할 수 있는 비율은 $\lambda_i/p$이다. 즉 $k$개 주성분들이 설명할 수 있는 누적 비율이 70%이상일 때의 개수를 선택하게 된다. 그림 5은 실험에 사용한 제스처들의 특징 집합으로부터 구한 고유치 개수에 따른 주성분의 누적기여도를 보여주고 있다.

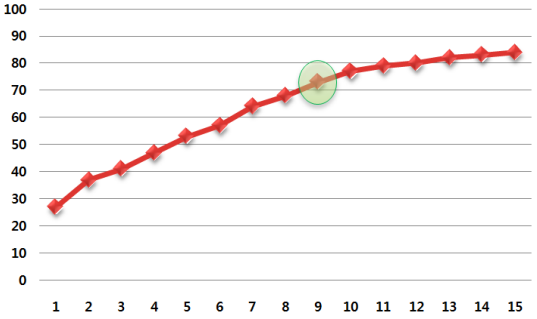


그림 5. 고유치의 개수에 따른 누적 기여도
Fig. 5 Cumulative contribution depending on the number of eigenvalue

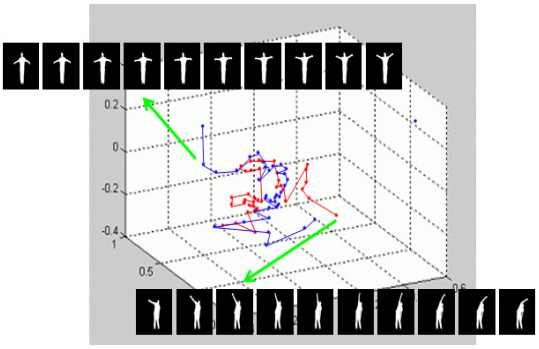

그림 6. 영상 시퀀스를 제스처 공간으로의 투영
Fig. 6 The result of projecting image sequences to gesture space

그림 6은 옆구리 동작과 팔 벌리기 동작이 포함된 영상 시퀀스를 제스처 공간에 투영한 결과를 나타낸 것이다.

### 4.2 HMM을 이용한 제스처 인식

제스처 공간에 투영된 점들은 클러스트링 알고리즘에 의해서 몇 개의 제스처 패턴으로 분류되어질 수 있다.

이렇게 분류된 제스처 패턴들에 대해 특정한 심볼을 부여함으로써, 제스처 시퀀스는 심볼 시퀀스로 형상화되어지고 이를 은익 마르코프 모델의 입력으로 사용한다. 앞 절에서 구한 저차원의 특징 데이터 값들을 기호로 바꾸기 위해서는 클러스터링 알고리즘을 이용해 몇 개의 제스처 군집(Cluster)으로 나누고, 각 군집에 대해 특정 기호를 할당한다. 그리고 각 클러스터의 중심 좌표 값은 코드 북으로 저장되어 새로운 특징 값이 들어왔을 때 기호를 할당하는 기준이 된다. 제스처들의 심볼 집합이 입력으로 들어오면 Baum-Welch 알고리즘[7]에 의해 인식에 필요한 3가지 파라메터$(\pi, A, B)$을 식(15 ~ 17)로 추정할 수 있다. 상태 천이 확률 $a_{i,j}$는 은닉 마르코프 모델의 상태가 $i$로부터 $j$로 변화하는 확률을 의미한다. 그리고 확률 $b_{i,j}$는 출력 심볼 $y$가 상태 $i$로부터 $j$로 천이되면서 관측될 수 있는 확률이고, $\pi_i$는 초기 상태 확률 값을 나타낸다.

$$\overline{\pi_i} = \gamma_1(i) \tag{15}$$

$$\overline{\alpha_{i,j}} = \frac{\sum_{t=1}^{T-1} \xi_t(i,j)}{\sum_{t=1}^{T-1} \gamma_t(i)} \tag{16}$$

$$\overline{b_{i,j}}(k) = \frac{\sum_{t=1, s.t. y_t = k}^{T} \gamma_t(j)}{\sum_{t=1}^{T} \gamma_t(j)} \tag{17}$$

학습 과정을 통해 각 제스처에 해당되는 3가지 파라메터$(\pi, A, B)$값들이 결정되는데 이를 제서처 모델이라한다. 인식하고자 하는 새로운 심볼 집합($Y$)이 주어지면 은닉 마르코프 모델에서는 각 제스처 모델 $\lambda_i$에 대한 확률 값은 전방 변수인 $\alpha_t(i)$와 후방 변수인 $\beta_t(i)$를 이용하여 식(18)과 같이 구하고 가장 높은 확률 값을 갖는 모델로 인식하게 된다.

$$P(Y|\lambda_i) = \sum_i \sum_j \alpha_t(i) a_{ij} b_{ij}(y_{t+1}) \beta_{t+1}(j) \tag{18}$$

## V. 실험결과 및 결론

실험에 이용된 영상은 정면에 위치한 비전 카메라를 통하여 들어오는 그레이 영상을 입력 영상으로 사용하였다. 학습에 사용된 제스처는 다리 벌리기, 다리 펴기, 옆구리 운동, 팔 벌리기, 일반 걷기, 팔 올리기등 일반적인 체조의 움직임을 이용하였다. 그림 7은 제스처 인식에 사용한 각 포즈에 대한 실루엣 영상 집합 일부를 보여주고 있다.

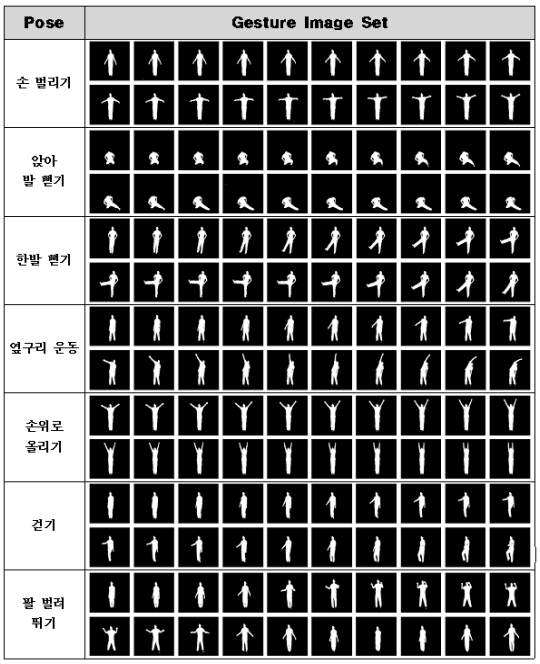


그림 7. 실험에 사용한 영상 시퀀스들
Fig. 7 Images of Gesture used for the study

주성분 분석을 이용한 인식에 가장 문제점은 고유공간을 구성하는 시간과 제스처 인식을 수행하는 속도가 빠르지 않기에 실시간 적용에는 무리가 있다. 그래서 제스처 인식에 대한 고유공간을 구성하기 위해 640 * 480으로 촬영한 제스처 영상을 크기 정규화를 거쳐 320 * 240영상으로 변환하였다. 각각의 포즈는 50프레임으로 구성되어 있으며, 한 사람당 350프레임(7개의 포즈 * 50프레임)의 영상이 사용되었고, 실험에 참가한 사람은 8명이므로 총 2800프레임을 이용하여 모델을 구성하였다. 그리고 본 논문에서는 제스처의 모든 영상의 데이터

를 사용하지 않고 2장에서 제시한 방법을 사용하여 얻어진 실루엣 영상을 사용하였다. 제스처의 실루엣 영상만으로도 제스처의 특징을 상당 부분 가지고 있기에 이 특징을 에지 방향성 히스토그램의 입력데이터로 사용하였다. 또한 제스처 실루엣 영상이 320 * 240 = 76,800개의 데이터를 가지고 있는데 이 데이터를 에지 방향성 히스토그램을 이용하여 36개 특징 데이터만 추출하게 된다. 즉 전처리 과정에서 76,800개의 데이터를 35개로 축소시키게 되는 것이다.

표 1은 주성분 분석의 고유 공간 구성 시간과 기존의 주성분 분석과 제안한 방법의 인식률을 비교 나타내었다 실험환경은 인텔 듀얼 CPU 1.86GHz, 2GB RAM에서 실험하였고, 그 결과 인식률의 큰 변화 추이 없이 상당한 속도 개선의 효과를 만들었다.

표 1. PCA와 EOH를 적용한 PCA의 처리시간 및 인식률
Table. 1 PCA processing the recognition rate using PCA and EOH

| 동작 | 모델 영상 개수 | PCA | | PCA using EOH | |
|---|---|---|---|---|---|
| | | Training Time(sec) | 인식률 | Training Time(sec) | 인식률 |
| 손 벌리기 | 400 | 112.5 | 94.8% | 0.127 | 92.50% |
| 앉아 발뻗기 | 800 | 380.7 | 95.0% | 0.436 | 91.00% |
| 한발 뻗기 | 1200 | 672.7 | 94.5% | 0.597 | 91.50% |
| 옆구리 운동 | 1600 | 1053.9 | 94.7% | 0.917 | 90.70% |
| 손위로 올리기 | 2000 | 1580.5 | 93.7% | 2.531 | 90.50% |
| 걷기 | 2400 | 2300.6 | 92.3% | 3.047 | 90.10% |
| 팔벌려 뛰기 | 2800 | 3220.5 | 91.3% | 5.262 | 90.00% |

그리고 표 2는 각 제스처에 대한 매칭 방법별 매칭 성공률에 대한 분석 결과를 나타내었다. 표 2에서 나타나듯이 HMM을 이용한 매칭 방법이 기존의 최소 거리 매칭 방법보다는 매칭 성공률이 높음을 알 수 있었다. 그러나 각 포즈에서 비슷한 동작(손 벌리기와 손위로 올리기)이 발생하는 부분에서는 잘못된 매칭이나 매칭 실패가 기존의 거리계산 방법에서는 빈번하게 발생되어 인식률이 심각하게 저하되었고 제안한 방법에서도 인식오류가 발생하였다. 그 이유를 분석한 결과 모델을 구성했던 영상 시퀀스들과 모델과 동일한 속도로 동일한 동작을 취한 영상 시퀀스들에 대해서는 거의 대부분 올바르게 인식되었지만, 일부(손 위로 올리기)가 다른 시퀀스(손 벌리기)로 인식하는 오류가 발생하였다. 이는 제안한 방법이 손이나 발의 정확한 위치를 계측하여 하는 방식이 아닌 외관 기반 정보를 이용하여 모델을 구성하였기에 비슷한 동작에서 오 인식이 발생된 것으로 사료된다.

표 2. 매칭 방법에 따른 인식률
Table. 2 Successful Matching Rate According to the Matching Methods

| 구분 | 인식률(%) | |
|---|---|---|
| 동작 | 기존의 거리계산 | HMM |
| 손 벌리기 | 82% | 89% |
| 앉아 발뻗기 | 75% | 92% |
| 한발 뻗기 | 57% | 97% |
| 옆구리 운동 | 70% | 95% |
| 손위로 올리기 | 68% | 85% |
| 걷기 | 98% | 98% |
| 팔벌려 뛰기 | 68% | 96% |
| 평균 | 74.00% | 93.14% |

본 논문에서는 EOH와 주성분 분석법을 결합하여 기존의 영상 데이터 전부를 사용하는 방법이 아닌 제스처의 형태의 정보인 실루엣 영상을 EOH의 입력으로 사용하여 얻어진 특징 데이터를 주성분 분석에 활용함으로써 차원을 낮추면서도 인식률을 90%이상 유지하는 방법을 제시하였고, 은닉 마르코프 모델을 적용하여 매칭 인식률을 향상시켰다. 이 방법들의 특징은 손이나 발의 정확한 위치를 계측하여 특징 정보를 얻는 것이 아니고, 영상에서 쉽게 계산이 가능한 특징 값들을 이용하여 구한 모션 히스토리 정보를 인식과정에 사용한다는 점이다. 그러므로 구체적인 동작을 인식하면서 많은 계산량이 요구되지 않기 때문에 실세계에서 구현이 용이하고 실시간 시스템 구축에 적합하다.

## 참고문헌

[1] Andrea Corradini, "Dynamic Time Warping for Off-line Recognition of a Small Gesture Vocabulary," RATFG-RTS,2001.

[2] 석홍일, 이성환, “3차원 손 모델링 기반의 실시간 손 포즈 추적 및 손가락 동작 인식” 정보과학회논문지 제35권 12호, pp.780-788, 2008. 12.
[3] M.Turk, ″Computer Vision in the Interface,″ Commumications of the ACM,Vol.47, No.1, 2004.
[4] Ismail Haritaoglu, David Harwood and Larry S. Davis, ″W4: Who? When? Where? What? A Real-time System for Detecting and Tracking People,″ Third Face and Gesture Recognition Conference, pp.222-227, 1998.
[6] James W. Davis, Aaron F. Bobic. ″The Representation and Recognition of Action using Temporal Templates″, MIT Media Lab Technical Report 402, 1997.
[7] Caelli. T and McCane.B, ″Components analysis of Hidden Markov Models in Computer vision″, Image Analysis and Processing, 2003, pp.510-515, September 17-19, 2003.
[8] Jong-Min Kim, and Maeng-Kyu Song, ″Three Dimensional Gesture Recognition Using PCA of Stereo Images and Modified Matching Algorithm″, IEEE Fuzzy Systems and Knowledge Discovery Vol.4, pp.116-120, Oct. 2008.

## 저자소개

**이기준(Kee-Jun Lee)**

1994년 2월 : 조선대학교 전자계산학과(이학사)
1997년 2월 : 조선대학교 전산통계학과(이학석사)
2001년 8월 : 조선대학교 전산통계학과(이학박사)
2002년 3월～현재 : 광주보건대학 보건교육정보과 부교수
※관심분야 : 영상처리, 내용기반검색, 멀티미디어