

HMM을 이용한 보행자 인식

(HMM-Based Human Gait Recognition)

신 봉 기[†] 석 흥 일^{††}

(Bong-Kee Sin) (Heung-II Suk)

요약 최근, 사람을 인식하는데 있어 걸음걸이가 기존에 사용되어 오던 많은 생체인식을 보완할 만한 것으로 등장하였다. 본 연구는 보행자 실루엣의 동적 특징과 은닉 마르코프 모델(HMM)을 이용한 보행자 인식 방법을 제안한다. 보행자의 보행 모델은 무한 순환 구조의 HMM 두 가지를 사용하였다. 하나는 자기 조직화 지도(SOM)를 벡터 양자화기로 하는 이산 HMM방식이고, 다른 하나는 주성분 분석(PCA) 공간으로 변환된 특징 벡터를 이용하는 연속 HMM방식이다. 실험 결과 HMM이 몇 가지 변수의 조정에 대해 일관성 있는 성능 변화를 보이며 최고 88.1%의 인식률을 기록하였다. 또한 기존 연구 결과와 비교하여 볼 때 특징과 제안 구조의 모델은 보행자 인식에 충분한 적용 가능성이 있으며, 나아가 걸음걸이가 생체 인식으로 이용되기에는 좋은 지표가 될 수 있을 것으로 판단된다.

키워드 : 보행자 인식, 생체인식, 은닉 마르코프 모델

Abstract Recently human gait has been considered as a useful biometric supporting high performance human identification systems. This paper proposes a view-based pedestrian identification method using the dynamic silhouettes of a human body modeled with the Hidden Markov Model(HMM). Two types of gait models have been developed both with an endless cycle architecture: one is a discrete HMM method using a self-organizing map-based VQ codebook and the other is a continuous HMM method using feature vectors transformed into a PCA space. Experimental results showed a consistent performance trend over a range of model parameters and the recognition rate up to 88.1%. Compared with other methods, the proposed models and techniques are believed to have a sufficient potential for a successful application to gait recognition.

Key words : pedestrian recognition, biometric, Hidden Markov Model

1. 서론

홍채, 지문 등을 이용한 인식은 이들을 열기 위한 과정에서 거부감을 줄 수 있으며, 얼굴을 이용한 인식은 가능한 정면의 모습을 찍어야 한다는 단점을 가지고 있다. 그러나 걸음걸이를 이용한 인식은 개인에게 특정 자세를 요구하지도 않으며, 심지어는 이를 이용한 인식이 행해지고 있다는 것을 그 사람이 모르게 할 수 있다. 사람들은 잘 아는 사람이라면 얼굴을 알아볼 수 없는 먼 거리에 있더라도 그의 걸음걸이만으로 인식할 수 있다는 것을 알고 있으며, 이는 Johansson의 Moving Light

Display(MLD)를 이용한 연구[1]에 의해 밝혀진바 있다.

최근 컴퓨터 시각분야에서는 사람의 걸음걸이를 인식에 이용하려는 많은 연구가 있으며, 이들 연구는 크게 '모델 기반 방법'과 모델이 없는 '특징 기반 방법'으로 나뉘어 진다[2,3]. 두 방법 모두 1) 특징 추출, 2) 특징 대응(feature correspondence), 3) 고급 처리(high-level processing)라는 공통된 구조를 따르지만, 특징 대응을 생성하는데 있어서 주요 차이점이 있다. '모델 기반 방법'은 신체의 각 구조를 2D 리본, 3D 원통 등을 이용하여 표현한 미리 정의된 2D 또는 3D의 모델에 대해, 입력 2D 이미지를 비교하여 특징 대응이 자동적으로 생성 되도록 하는 방식이다. 반면, 모델이 없는 '특징 기반 방법'은 연속된 프레임간에 위치, 속도, 모양, 색상 등을 기반으로 특징 대응을 생성하는 방식이다. 본 연구에서는 사람의 걸음걸이에 대한 구조적 특징을 표현하기 위해서 '특징 기반 방법'을 이용한다.

본 연구에서는 입력 이미지에서 사람을 찾기 위해서

· 이 논문은 2003년도 한국학술진흥재단의 지원에 의하여 연구되었음 (KRF-2003-002-D00300)

† 정 회 원 : 부경대학교 컴퓨터공학과 교수
bkshin@pknu.ac.kr

†† 학 생 회 원 : 부경대학교 컴퓨터공학과
heungil@gmail.com

논문접수 : 2005년 5월 6일

심사완료 : 2006년 3월 15일

배경 이미지와의 차이를 이용하고, 사람의 실루엣으로 프로파일 벡터를 생성하며, 프로파일 벡터열을 이용하여 걸음에 대한 특징을 표현한다. 그러나 비슷한 외양을 가진 사람이 있을 수도 있으므로 사람의 구조적인 특징을 나타내는 실루엣만으로는 인식을 하는데 한계가 있을 것이다. 이를 보완하기 위해 걸음걸이의 동적 특징을 고려할 필요가 있으며, 걸음걸이는 이전 자세에 큰 영향을 받는 특징이 있으므로 '마르코프 사슬(Markov chain)'로 표현할 수 있다. 은닉 마르코프 모델(HMM)[4]은 '마르코프 사슬'의 하나로서 변형과 잡음이 많은 시계열 데이터 분석에 아주 적합한 모델이다. 본 연구에서는 이를 모델로 채택하였다. 인식에 사용될 프로파일 벡터는 걸어가는 사람의 수직 중심축으로부터 좌·우측 실루엣까지의 거리로 구성하였다. 이 것은 Kale 등[5]의 방법과 직접 비교가 가능한 특징이며 본 실험의 결과 더 우수한 성능을 나타내었다. 그리고, 이산 HMM(DHMM)과 연속 HMM(CHMM) 각각을 모델로 하여 서로 다른 두 가지의 인식기를 설계하였다. DHMM을 이용한 인식을 위해서는 SOM을 양자화기로 하여 프로파일 벡터열을 코드열로 바꾸었으며, CHMM을 이용한 인식에서는 고차원의 프로파일 벡터를 주성분 분석을 하여 차원을 낮춤으로서 훈련 데이터 부족으로 인한 과적응(overfitting)이 발생하지 않도록 하였다.

본 논문에서 제안한 방법의 특징은 사람의 구조적인 특징과 걸음걸이의 동적인 특징을 모두 고려한 것이다. 구조적인 특징은 좌·우측 실루엣으로 표현되고, 걸음걸이의 동적인 특징은 HMM에 의해 잘 나타내어진다. 특히, HMM을 무한 순환 고리 구조로 표현하였기 때문에 걸음걸이 주기내의 어느 위치에서 데이터열이 시작하여도 인식을 할 수 있다는 장점이 있다. 또한, 입력 비디오에서 사전 분할(segmentation)이 없어도 되며 입력의 길이가 길면 길수록 성능이 정확해질 수 있다는 장점이 있다.

본 논문은 다음과 같이 구성된다. 2절에서는 관련 연구를 알아보고, 3절에서는 시스템 전체의 전반적인 개요를 설명한다. 4절에서는 걸음의 특징을 나타내는 프로파일 벡터를 추출하는 방법을 설명하며, 5절에서는 인식을 위한 시스템의 구성, 6절에서는 실험 결과를 설명한다. 마지막으로 7절에서는 본 연구에 대한 결론을 내리고, 앞으로의 연구 방향을 제시한다.

2. 관련 연구

이 절에서는 사람의 걸음걸이를 이용한 인식과 밀접한 관련이 있는 추적, 검출 및 사람의 움직임에 대한 인식을 다룬 연구들을 살펴본다. 이에 대한 전반적인 소개와 기술은 J. Aggarwal 등[2], C. Cedras 등[6], D.M.

Gavrila 등[3]의 조사 논문에 잘 분류, 설명되어 있다.

사람의 행동이나 특정 자세를 인식하기 위한 연구에서는 HMM이 자주 사용된다. J. Yamato 등은 테니스의 자세를 인식하기 위해 이산 HMM을 사용하였으며 [7], A. Kale 등은 이전화된 이미지에서 실루엣의 폭을 이용해 특징 벡터를 생성한 뒤, 생성된 벡터를 이용해 k-means 클러스터링을 수행하여 5개의 기본 자세를 선택하였다. 이들 5개의 기본 자세와 입력 특징 벡터와의 유클리드 거리를 값으로 하는 5차원 벡터를 생성하고, HMM을 인식기로 하여 89.09%의 성능을 얻었다[5,8]. 한편, J.J Little 등은 광류(optical flow)에 의한 방식을 사용하여 개개인의 걸음걸이에 대한 움직임 및 주파수와 위상(phase)의 특징을 찾고, 이를 인식에 이용하였다 [9]. R. Collins 등에 의한 실험에서는 형태정합을 이용하여 관찰하는 위치에 의존적인 방식을 소개하였으며 [10], C. BenAbdelkader 등은 추출된 이미지간의 상관 관계를 이용하여 자기유사도(self-similarity)를 계산하고, PCA와 k-NN 방법을 이용하여 벡터의 차원을 낮추어 인식하는 방법을 제안하였다[11].

3. 시스템 구성 개요

본 연구에 대한 전반적인 시스템의 구성은 그림 1에 나타나 있다. 시스템은 크게 입력 이미지에서 사람을 검출하는 부분과, 검출된 사람의 걸음걸이에 대한 프로파일 벡터를 생성하여 특징을 추출하는 부분, 추출된 특징 벡터를 이용해 인식기의 파라미터를 훈련하고, 이를 이용해 인식을 수행하는 부분으로 이루어져 있다. 현재 두 가지의 HMM 인식기는 서로 독립이며 별도의 인식기로 생각할 수 있다.

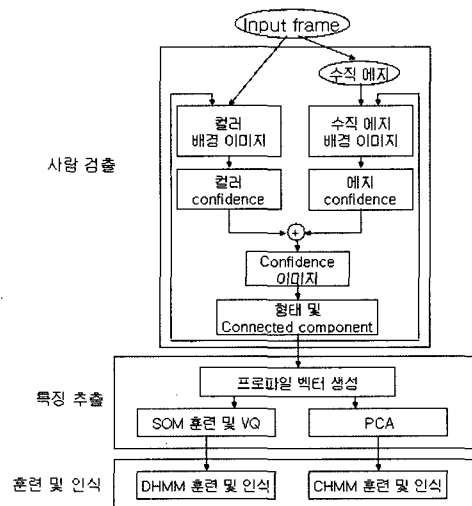


그림 1 시스템 기능 및 구조도

4. 보행자 프로파일 추출

사람의 걸음걸이를 볼 때 그 사람의 걸어가 는 모습을 잘 나타낼 수 있는 방법 중 하나는 그 사람의 실루엣이다. 본 연구에서는 이미지에서 사람의 실루엣을 찾고, 그 실루엣으로부터 걸음걸이의 구조적인 특성을 표현하는 프로파일 벡터를 생성한다. 움직이는 사람을 찾기 위해서는 J. Sabri 등이 사용한 방식[12]을 이용한다.

t 시자의 입력 비디오 프레임은 $I' = \{I'_{ij} | 1 \leq i \leq h, 1 \leq j \leq w\}$ 와 같이 표현되며, w 와 h 는 각각 프레임의 가로, 세로의 픽셀수이다. 배경 이미지 $B = \{b_{ij} | b_{ij} = \sum_t w_t I'_{ij}, I'_{ij} \notin F\}$

와 같이 생성되며, F 는 이미지상에서 보행자로 인식된 픽셀들의 집합을 나타낸다. 여기서, w_t 는 가중치로 $w_t = \alpha(1-\alpha)^{t-1}$ 와 같이 계산되고, 이는 첫 번째 입력 프레임부터 t 시자의 프레임까지의 픽셀값을 모두 고려하여 지수적으로 가중된 픽셀값을 배경 이미지의 픽셀값으로 할당하는 기능을 한다. $\alpha \in [0, 1]$ 는 배경 이미지 갱신에 대한 학습률이다.

프로파일 벡터 생성을 위한 영상처리 과정은 다음과 같다.

- A. 새로운 입력 이미지에 대해 가중평균 방법을 이용하여 컬러 이미지와 수직 에지 이미지에 대한 각각의 배경 이미지를 생성한다. 이미지의 각 픽셀에 대해 배경 평균 픽셀값은 식 (1)의 방법으로, 표준편차는 식 (2)의 방법으로 갱신한다.

$$\mu_t = \alpha x_t + (1-\alpha)\mu_{t-1} \quad (1)$$

$$\sigma_t^2 = \alpha(x_t - \mu_t)^2 + (1-\alpha)\sigma_{t-1}^2 \quad (2)$$

식 (1)과 (2)에서 x_t 는 t 번째 입력 이미지의 픽셀값이고, μ_t 와 σ_t 는 t 번째 입력 이미지까지의 평균과 표준 편차를 나타낸다.

- B. 배경 이미지와 새로운 입력 이미지와의 차이 $D_{ij} = I_{ij} - B_{ij}$ 를 구하고 보행자 픽셀로서의 신뢰도 C_{ij} 를 식 (3)과 같이 계산한다[12].

$$C_{ij} = \begin{cases} 0 & D_{ij} < m\sigma_{ij} \\ \frac{D_{ij} - m\sigma_{ij}}{M\sigma_{ij} - m\sigma_{ij}} & m\sigma_{ij} \leq D_{ij} \leq M\sigma_{ij} \\ 1 & D_{ij} > M\sigma_{ij} \end{cases} \quad (3)$$

여기서 $0 < m < M$ 이고 $M\sigma_{ij}$ 와 $m\sigma_{ij}$ 는 각각 최대한계와 최소한계를 말한다.

- C. 컬러 이미지에 대한 신뢰도와 에지 이미지에 대한

신뢰도를 결합하고, 신뢰도가 너무 낮은 경우는 잘못 찾은 픽셀로 간주하고 제거한다.

- D. 형태 연산과 연결 성분 분석을 적용하여 사람의 영역(blob) H 를 결정한다.

프로파일은 H 에 대한 수평 중심으로부터 좌·우측 실루엣까지의 거리로 표현된다. 수평 중심은 식 (4)와 같이 계산되며, 실루엣은 H 의 가장 좌측과 가장 우측의 픽셀이다.

$$M_c = \frac{\sum_{(i,j) \in H} j}{|H|} \quad (4)$$

위 식에서 M_c 는 수평 중심을 나타내고, $|H|$ 는 H 에 속한 픽셀의 개수를, (i, j) 는 픽셀의 좌표를 의미한다. 벡터를 표현함에 있어서 모든 실루엣 픽셀을 고려하게 되면 벡터의 차원이 너무 커지게 되며, H 의 크기가 가변적이므로 벡터의 크기도 가변적으로 되게 된다. 따라서, 그림 2와 같이 외접 사각형을 상하로 등분하여 최하 단점을 뺀 40차원의 프로파일 벡터를 생성한다. 벡터의 원소는 앞쪽 발끝을 시작으로 시계 방향으로 구성하고, 외접 사각형의 높이를 이용하여 벡터의 값을 정규화한다.

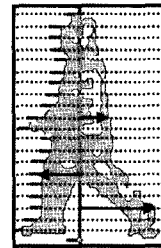


그림 2 프로파일 벡터 생성을 위한 실루엣 측정

5. 보행자 모델

본 연구에서는 보행자의 동적인 특성을 고려하기 위해 DHMM과 CHMM을 인식/보행자 모델로 사용한다. DHMM의 입력 데이터는 SOM을 이용해 프로파일 벡터를 양자화한 코드열이고, CHMM의 입력 데이터는 프로파일 벡터에 대해 주성분 분석을 적용하여 차원을 낮춘 벡터이다. 우선, DHMM의 입력 데이터를 위한 SOM을 이용한 벡터 양자화를 먼저 설명하고, CHMM의 입력 데이터 생성을 위한 주성분 분석을 설명한다.

5.1 Kohonen 자기조직화 지도

사람의 걸음걸이는 특정 자세들의 반복으로 이루어진다는 것을 고려하여, 자기조직화 지도(SOM) 출력층의 구조를 그림 3과 같이 고리 모양으로 한다. 이 때 출력

층의 각 노드는 하나의 자세를 표현하는 40차원 벡터이다. 그리고, 각각의 이미지에서 추출된 프로파일 벡터를 이용하여 각각의 노드에 대한 가중치 벡터를 훈련한다. 훈련된 가중치 벡터를 코드 벡터로 하여 프로파일 벡터와의 유사성을 비교하고 가장 유사한 코드 벡터가 프로파일 벡터를 대신한다. 즉, 식 (5)에 나타난 것과 같이 코드 벡터와 프로파일 벡터간의 거리를 계산함으로써 가장 작은 거리를 가지는 코드 벡터의 아이디가 프로파일 벡터에 대한 코드(codeword)로 결정된다.

$$codeword = \arg \min_k \|x - w_k\| \quad (5)$$

여기서, x 는 프로파일 벡터이고, w_k 는 훈련된 코드 벡터이며, k 는 코드 벡터의 번호이다.

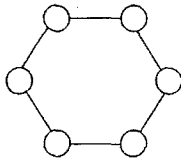


그림 3 SOM의 출력층 구조

5.2 주성분 분석(PCA)

주성분 분석은 벡터 성분간의 분산 및 상관 관계를 고려하여 벡터의 차원을 줄이되, 정보의 손실을 최소화하는 방법이다. 벡터 데이터에 대해 분산이 큰 방향이 작은 방향보다 상대적으로 많은 정보를 포함하고 있으므로, 데이터를 분산이 큰 방향의 축으로 변환하여 데이터의 차원을 줄인다. 본 연구에서는 프로파일 벡터 데이터에 주성분 분석을 적용하여 차원을 낮춘 후 CHMM을 훈련한다.

5.3 HMM

걸음걸이를 이용한 인식에 있어서 걸음걸이의 구조적인 특징만으로는 좋은 인식률을 얻기는 어렵다. 왜냐하면, 신체의 구조적인 특징이 비슷한 사람이 있을 가능성도 있기 때문이다. 그러나 사람의 걸음걸이를 자세히 들여다보면 구조적인 특징이 비슷하다 할지라도 개개인마다 다른 동적 특성이 있다는 것을 알 수 있다. 이런 동적 특성은 HMM의 상태 변화로 표현될 수 있다.

HMM은 상태의 주요한 변화를 '마르코프 과정'으로 모델링한다. 상태 변화 시퀀스는 겉으로 드러나지 않으며(hidden), 직접 관측이 가능한 데이터를 통해서 간접적으로 알 수 있다. HMM은 초기 상태의 확률 분포 π , 상태간의 전이 확률 분포 $A = \{a_{ij}\}$, 어떤 상태에서 특정 심볼을 출력할 확률 분포 $B = \{b_{jk}\}$ 를 파라미터로 가지며, $\lambda = (\pi, A, B)$ 로 표기한다. 심볼 출력 확률

은 DHMM에서는 특정 값을 가지나 CHMM에서는 하나의 일반적인 혼합 가우스(mixture-Gaussian) 분포 함수로 표현한다.

본 연구에서 이용하는 HMM의 구조는 사람의 걸음걸이에 주기성이 있다는 것을 감안하여 무한순환 고리 구조로 구성한다(그림 4). DHMM의 파라미터는 SOM 벡터 양자화기로 양자화한 코드열 데이터를 사용하여 훈련하며, CHMM의 파라미터는 주성분 분석으로 차원을 낮춘 벡터로 훈련한다.

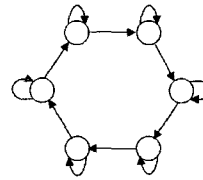


그림 4 HMM 구조

HMM은 등록된 보행자마다 각각의 모델($\lambda = (A, B, \pi)$)을 만들어 훈련하므로, 개개인의 걸음걸이의 특성을 표현하는 수단이 된다. 즉, 어떤 사람의 걸음걸이에 대한 특성 벡터가 입력으로 주어질 경우 그 사람의 모델이 입력 특성 벡터에 대해 가장 높은 우도(likelihood)를 출력하게 된다.

$$\hat{i} = \arg \max_i P(\mathbf{X} | \lambda_i) \quad (6)$$

각 모델의 우도는 전진(forward) 알고리즘을 이용하여 계산한다[4]. 식 (6)의 \mathbf{X} 는 입력 벡터열이고, λ_i 는 i 번째 사람에 대한 HMM 모델을 나타내며, \hat{i} 는 가장 높은 우도를 보인 모델의 번호를 의미한다.

6. 실험

6.1 데이터

실험에 사용된 데이터는 캐나다 Calgary 대학의 데이터를 사용하였으며[13], 6명의 사람을 각각 7번씩 촬영하여 총 42개의 비디오로 되어 있다. 비디오는 초당 15 프레임으로 촬영되었으며, 각 비디오의 평균 프레임 수는 85이고, 약 6초간, 평균 6걸음의 길이이다. 이미지의 크기는 320×160 픽셀이고 24비트 컬러이다.

6.2 샘플 영상 및 처리 과정

영상 처리의 과정을 나타내고 있는 그림 5를 보면 (a)와 (d)는 앞 절에서 설명한 방법으로 생성된 배경 이미지이며, (b)와 (e)는 새로운 입력 이미지와 그의 수직 에지 이미지이다. 배경 이미지와 입력이미지간의 차이에 대한 신뢰도(confidence) 이미지가 (c)와 (f)에 각각 나와 있으며, 이 둘을 결합하고, 형태 연산 및 연결 성분

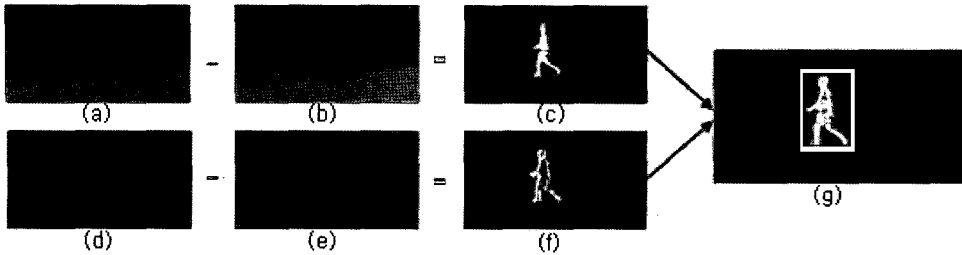


그림 5 영상처리 과정. (a)컬러 배경 이미지, (b)입력 이미지, (c)컬러 이미지 차이에 대한 신뢰도, (d)수직 에지 배경 이미지, (e)입력 이미지에 대한 수직 에지 이미지, (f)에지 이미지 차이에 대한 신뢰도, (g)검출된 보행자 실루엣과 외접 사각형

분석을 한 결과가 (g)이다.

6.3 프로파일 표현

그림 6은 프로파일 벡터를 어떻게 구성하는 지를 보여주고 있다. 그림의 가운데 부분에 있는 수직선은 무게 중심을 지나는 수직축을 표현한다. 이 중심에서부터 실루엣까지의 수평 거리를 차례로 측정하여 프로파일 벡터를 만든다.

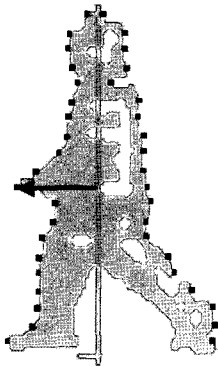


그림 6 프로파일 벡터의 구성

6.4 데이터 코딩

6.4.1 SOM 양자화

DHMM의 입력 벡터는 SOM을 양자화기로 사용한 코드열이다. 각 비디오 시퀀스의 프레임에서 추출된 프로파일 벡터를 훈련 데이터로 사용해서 SOM을 훈련시켰으며, 훈련된 SOM의 가중치 벡터의 형상은 그림 7에 나와 있다. 훈련된 가중치 벡터는 벡터 양자화의 코드 벡터로 사용되었다.

벡터 양자화 과정에서는 입력 이미지의 프로파일 벡터를 가장 유사한 코드 벡터의 코드로 대응한다. 그림 8에 한 사람의 걸음 반주기에 해당하는 프레임들에 대한 벡터 양자화 결과가 나타나있다. 프로파일 벡터에 특정 코드를 할당하는 양자화 과정에서 정보의 손실이 수반

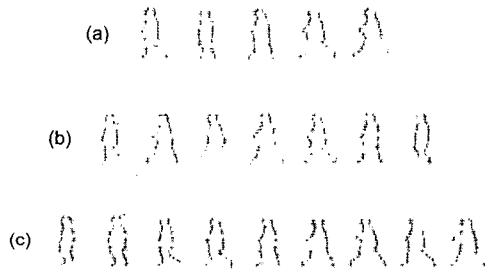


그림 7 코드북의 각 코드 벡터 가시화. (a) 5 코드북, (b) 7 코드북, (c) 9 코드북

된다. 이 정보의 손실은 프로파일 벡터와 코드 벡터와의 오차에 해당하는 것으로 그림 8의 양자화 과정에서 나타난 오차 값의 변화를 그림 9에 나타내었다. 이를 살펴보면 보폭의 크기가 최대 되었을 때 오차의 값이 특히 크게 나타나는 것을 볼 수 있는데, 사람마다 서로 다른 보폭 크기를 가지는 것을 몇 개의 고정된 자세로 표현하면서 나타나는 결과로 생각된다.

6.4.2 주성분 분석(PCA)

CHMM의 입력으로 40차원의 프로파일 벡터를 그대로 사용하려면 매우 많은 양의 훈련 데이터가 필요하다. 데이터가 부족하면 HMM이 제대로 훈련되지 못하며 그

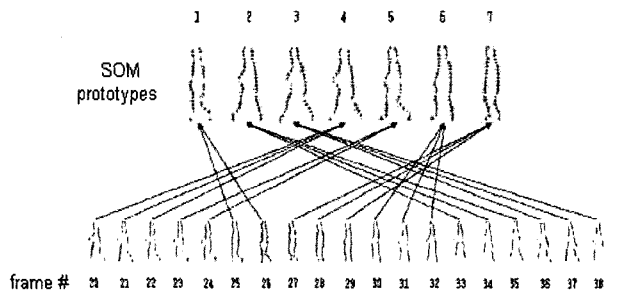


그림 8 걸음걸이 반주기에 대한 벡터 양자화 예

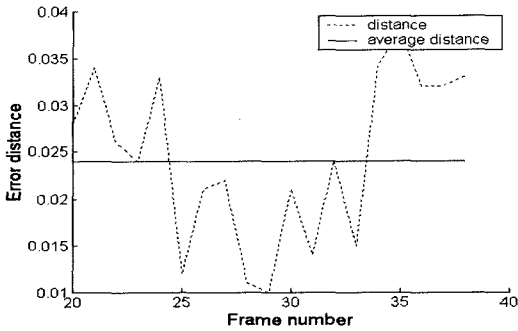


그림 9 그림 6의 양자화 과정에 나타난 오차

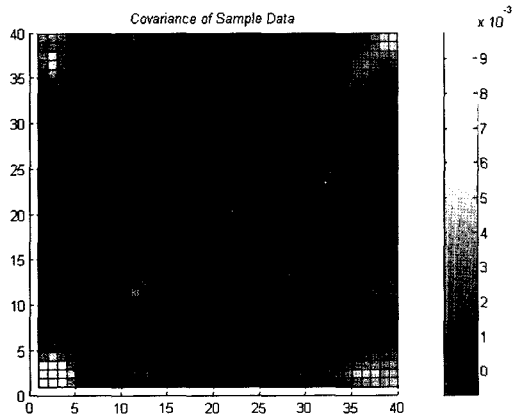


그림 10 공분산 행렬의 시각화. 수직축, 수평축은 모두 프로파일 벡터(그림 6)의 번호

에 따라 신뢰성이 크게 떨어진다. 따라서, CHMM의 입력 벡터의 차원을 낮출 필요가 있었으며 이를 위해서 주성분 분석을 적용하였다.

프로파일 벡터에 PCA를 적용하기 전에 PCA를 적용하는 것이 타당한지를 알아보기 위해 우선, 프로파일 벡터의 공분산을 살펴보았다. 공분산 행렬을 시각화한 그림 10을 보면, 일부 대각 성분의 값들이 비교적 밝게 나타나는데 이는 분산이 큰 곳, 즉 보행시에 신체를 많이 움직이는 부분을 나타낸다. 특히, 사람의 발/다리 부분(1-5, 35-40)과 손/팔 부분(11-15)에 해당하는 곳에서 분산이 높게 나타났으며, 몸통(16-27)에 해당하는 부분에서는 분산이 상대적으로 낮게 나타난 것을 볼 수 있다. 이 그림은 모든 프로파일 집합의 공분산 행렬이고, 동적인 특성을 정확히 표현하지는 않는다. 따라서, 분산이 동적 특성을 제대로 표현한다고는 할 수 없다. 하지만 이와 같은 부분에 시간적인 순서에 따른 변화 패턴의 식별 정보가 많이 숨어 있다고 예상된다. 어쨌든 분산 및 상관관계가 낮은 부분은 인식을 함에 있어서 큰 역할을 하지 않을 것이므로 이들의 정보는 상대적으로

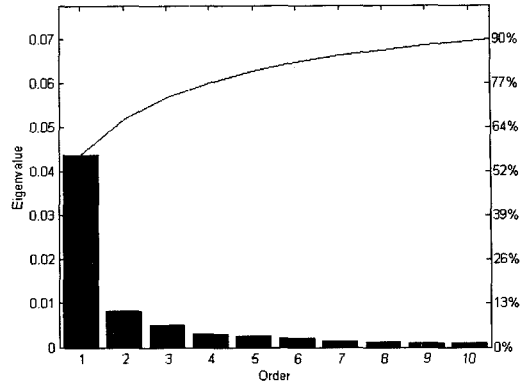


그림 11 공분산에 대한 고유치 기여도

덜 중요하다고 추정된다. 따라서, 40차원의 벡터를 PCA를 적용하여 차원을 줄이더라도 원래의 특징 대부분 나타낼 수 있게 된다.

프로파일 벡터에 PCA를 적용하여 차원을 낮추되 차원을 어느 정도로 낮추어야 하는 것인가는 또 다른 문제이다. 이를 결정하기 위해 고유치의 분포를 그래프로 나타내었다(그림 11). 벡터의 차원을 줄였을 때 선택된 주성분이 총 분산에 대해 최소한 d%를 나타내야 한다고 하면 이는 70%에서 90%사이의 값을 사용했을 때 정보의 대부분으로 유지한다고 알려져 있으며 이는 식 (7)과 같이 표현된다[14]. 식 (7)에서 λ_i 는 i번째 고유치를, K는 선택되는 고유치의 개수를 의미한다.

$$\sum_{i=1}^K \lambda_i \geq d \sum_{i=1}^P \lambda_i \geq \sum_{i=1}^{K-1} \lambda_i \quad (7)$$

그림 11을 보면 40 차원의 벡터를 5차원으로 줄여도 정보의 80%이상을 유지하고 있다는 것을 알 수 있다.

6.5 인식

실험 데이터가 충분하지 못하여 인식을 실험은 leave-one-out 교차 검증으로 행하였다. 인식 실험은 양자화 코드열과 주성분 분석에 의한 연속 벡터열을 입력 데이터로 하여 각각에 대해 DHMM과 CHMM을 인식기로 사용하였다.

6.5.1 DHMM 기반 인식

코드북의 크기(M)와 HMM의 상태수(N)를 동일하게 두고 실험 한 결과가 표 1에 나와 있다. 이 실험에서 M과 N이 모두 7일 때 높은 인식률을 보여, N을 7로 고정시키고, M을 변화시키며 또 다른 실험을 했고(표 2), 반대로 M을 7로 고정시키고 N을 변화시키며 실험을 했다(표 3). 이들 실험에서 N을 7로 하였을 때 가장 인식률이 높았다. 그림 12, 13, 14는 각각의 실험에서 인식기의 후보 개수를 고려하였을 때의 결과를 나타낸 것이다.

표 1 DHMM 인식을 결과(코드북의 크기 M =상태수 N)

DHMM test					
상태수(N)	5	6	7	8	9
Hits (%)	40.48	47.30	88.10	69.05	0.0

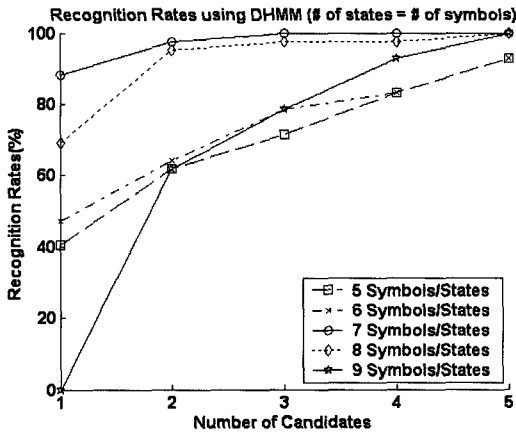


그림 12 후보수를 고려한 DHMM 인식을 변화($M=N$)

표 2 DHMM 인식을 결과(상태수 $N=7$)

DHMM test					
코드북의 크기(M)	5	6	7	8	9
Hits (%)	54.76	64.29	88.10	61.90	0.0

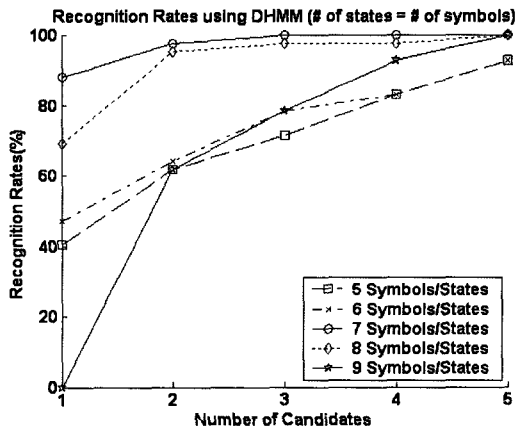


그림 13 후보수를 고려한 DHMM 인식을 변화($N=7$)

표 3 DHMM 인식을 결과(코드북의 크기 $M=7$)

DHMM test					
상태수(N)	5	6	7	8	9
Hits (%)	69.05	73.81	88.10	78.57	83.33

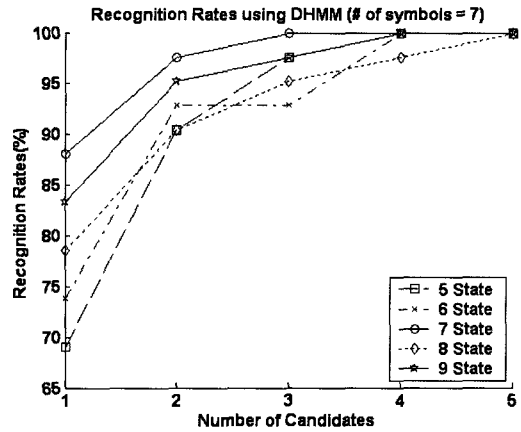


그림 14 후보수를 고려한 DHMM 인식을 변화($M=7$)

6.5.2 CHMM 기반 인식

* 벡터 차원 변화에 따른 성능 변화

DHMM의 결과를 바탕으로 상태수 N 을 7로 고정하였으며 PCA를 이용해서 차원을 바꿔가며 CHMM을 실험하였다. 그 결과 프로파일 벡터를 8차원으로 줄였을 때 88.10%로 가장 높은 인식률을 나타내었다. 차원의 변화에 따른 인식률 결과는 표 4에 나타내었으며, 각 차원에 대해 CHMM의 인식기의 후보 개수를 고려하였을 때의 결과는 그림 15에 나타내었다.

* 상태수 N 의 변화에 따른 성능 변화

앞의 실험에서 CHMM의 N 을 DHMM의 결과를 따

표 4 CHMM 인식을 결과(상태수 $N=7$)

CHMM test					
차원	5	6	7	8	9
Hits (%)	76.19	73.81	69.05	88.10	76.19

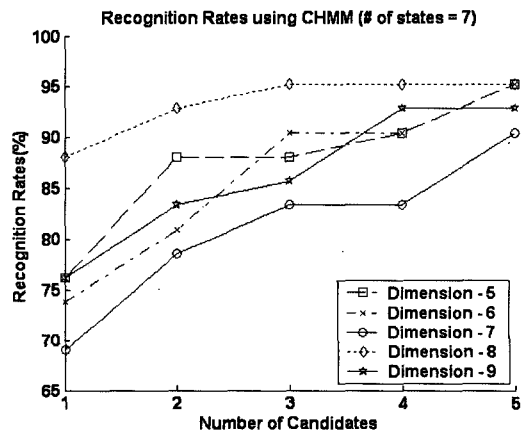


그림 15 후보수의 변화에 따른 CHMM 인식을 변화($N=7$)

표 5 CHMM 인식률 결과(차원=8)

CHMM test					
상태수(N)	5	6	7	8	9
Hits (%)	69.05	69.05	88.10	78.57	69.05

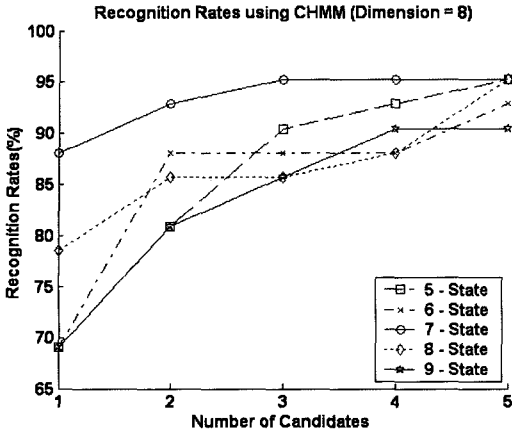


그림 16 후보수의 변화에 따른 CHMM 인식률 변화(차원=8)

라 7로 정하였는데 그것이 옳은 것인지를 확인하기 위해서 가장 높은 인식률을 보였던 8차원 데이터에 대해 차원을 고정시키고, N을 변화시키면서 다른 실험을 하였다. 그 결과가 표 5에 나와 있으며, DHMM의 경우와 같이 N을 7로 했을 때 가장 높은 인식률을 보였다. 그림 16은 인식기의 후보수를 고려했을 때의 인식률을 그래프로 나타낸 것이다.

6.6 성능 비교

본 논문에서 제시한 방법의 적합성을 알아보기 위해 본 논문에 사용된 데이터와 동일한 비디오 데이터를 이용하여 J.J. Little 등[9]과 Kale 등[8]의 방법을 비교해 보았다. J.J. Little 등의 연구에서는 본 논문에서 사용한 데이터와 동일 데이터이지만, 해상도와 초당 프레임수가 각각 2배인 데이터를 사용하였다. 광류(optical flow)에 의한 방식으로 움직임 및 주파수와 위상의 특징을 찾아 벡터를 생성하고, 이를 이용하여 90.5%의 인식률을 보였다. 본 논문의 인식률(88.10%)과 비교하면 조금 높은 수치를 보이는데 이는 사용된 데이터의 해상도와 프레임수가 성능에 영향을 미친 것으로 생각된다.

한편 각 프레임의 실루엣과 고정된 5개의 자세와의 차이를 값으로 하는 벡터를 특징 벡터로 하고, CHMM을 어고딕(ergodic) 구조로 하여 인식을 수행한 Kale 등의 방법의 결과에서는 같은 데이터에 대해 64.3%가 나왔다. 그림 17은 Kale 등에 의해 제안된 방법과 본 논문에서 제안한 방법의 결과를 비교한 것이다. CHMM

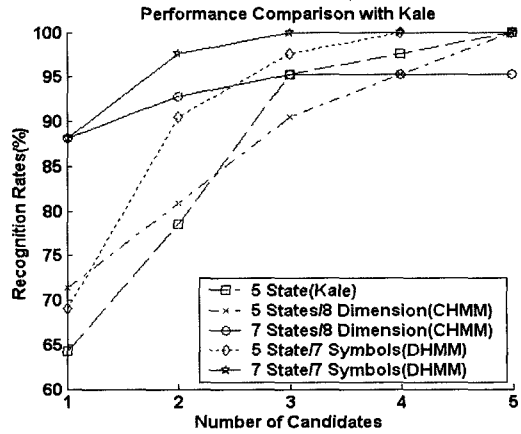


그림 17 Kale 등[8]이 제안한 방법과의 성능 비교

의 상태를 5개로 했을 때(71.41%)는, Kale 등의 결과보다 조금 높은 성능을 보였고, 상태수를 7개로 했을 때(88.10%)는 큰 차이를 보였다. 인식률의 차이를 보인 것은 고차원의 데이터를 5차원의 벡터로 변환하는 과정에서 발생한 오차에 의한 것으로 생각된다. 그리고, DHMM을 인식기로 하여 실험을 한 결과에서도 Kale 등의 결과보다 더 좋은 성능을 보였다.

7. 결론

본 연구에서는 비디오에 나타난 사람의 걸음걸이 특성을 이용하여 인식을 하는 방법을 제안하였다. 이미지 실루엣으로부터 프로파일 벡터를 생성하고, DHMM 훈련을 위해서는 SOM에 의한 벡터 양자화를 수행하였고, CHMM의 파라미터 훈련을 위해서는 주성분 분석을 적용하여 벡터의 차원을 줄였다. DHMM에서는 양자화 코드의 크기를 변경하면서, CHMM에서는 차원의 수를 변경하면서 성능의 변화를 알아보았고, 상태수의 변화에 대해서도 성능 실험을 하였다. 여러 가지의 실험을 종합한 결과 DHMM과 CHMM 모두 상태수를 7개로 했을 때 가장 높은 성능을 보였다.

본 연구에서의 방법은 이미지에서 추출된 실루엣을 기반으로 특징 벡터를 생성하였지만 보행자가 입고 있는 옷이나, 빛 조건 및 보행자와 카메라간의 각도 변화에 따른 실루엣의 큰 변화에 대해서는 고려하지 않았다. 따라서, 향후 이에 대한 연구를 해야 할 것이다.

참고문헌

[1] G. Johansson "Visual perception of biological motion and a model for its analysis," *Perception and Psychophysics*, vol. 14, no. 2, pp. 201-211, 1973.

- [2] J. Aggarwal and Q. Cai. "Human motion analysis -a review," *Computer Vision and Image Understanding*, vol. 73, no. 3, pp. 428-440, March 1999.
- [3] D.M. Gavrila. "The visual analysis of human movement-a survey," *Computer Vision and Image Understanding*, vol. 73, pp. 82-98, 1999.
- [4] L.R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no.2, pp.257-285, February 1989.
- [5] A. Kale, A. N. Rajagopalan, A. Sundaresan, N. Cuntoor, A. RoyChowdhury, V. Krueger, R. Chellappa, "Identification of humans using gait," *IEEE Transactions on Image Processing*, September, 2004.
- [6] C. Cedras and M. Shah, "Motion-based recognition -a survey," *Image and Vision Computing*, vol. 13, no. 2, pp. 129-155, 1995.
- [7] J. Yamato, J. Ohya, and L. Ishii. "Recognizing human action in time-sequential images using hidden Markov model," *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 624-630, 1995.
- [8] A. Kale, A. Rajagopalan, N. Cuntoor, and V. Kruger. "Gait-based recognition of humans using continuous HMMs," *Proceedings of the IEEE Conference on Automatic Face and Gesture Recognition*, pp. 321-326, 2002.
- [9] J.J. Little and J.E. Boyd. "Recognizing people by their gait: the shape of motion," *Videre*, vol. 1, no. 2, pp. 1-32, 1998.
- [10] R. Collins, R. Gross, and J. Shi. "Silhouette-based human identification from body shape and gait," *IEEE Conf Automatic Face and Gesture Recognition*, pp. 351-356, 2002.
- [11] C. BenAbdelkader, R.Cutler, and L.Davis. "Motion-based recognition of people in eigengait space." *IEEE Conf Automatic Face and Gesture Recognition*, pp. 254-259, 2002.
- [12] S. Jabri, Z. Duric, H. Wechsler, and A. Rosenfeld, "Detection and location of people in video images using adaptive fusion of color and edge information," *Proceedings of International Conference on Pattern Recognition*, pp. 627-630, 2000.
- [13] <http://pages.cpsc.ucalgary.ca/~boyd/gait/experiment.html>
- [14] Andrew R. Webb, *Statistical Pattern Recognition Second Edition*, John Wiley and Sons, 2002.



신 봉 기

1985년 서울대학교 자원공학 학사. 1987년 한국과학기술원 전산학 석사. 1995년 한국과학기술원 전산학 박사. 1987년~1999년 한국통신 멀티미디어연구소. 1999년~현재 부경대학교 전자컴퓨터정보통신공학부 부교수. 관심분야는 인공지능,

패턴인식, 컴퓨터시각, 기계학습 등



석 홍 일

2004년 부경대학교 멀티미디어공학과 졸업(학사). 2004년~현재 부경대학교 컴퓨터공학과 석사과정. 관심분야는 컴퓨터비전, 패턴 인식, 인공지능 등