

움직임 실루엣 영상의 일반적인 표현 방식에 대한 연구

A General Representation of Motion Silhouette Image: Generic Motion Silhouette Image(GMSI)

홍성준, 이희성, 김은태*
(Sungjun Hong, Heesung Lee, and Euntai Kim)

Abstract : In this paper, a generalized version of the Motion Silhouette Image(MSI) called the Generic Motion Silhouette Image(GMSI) is proposed for gait recognition. The GMSI is a gray-level image and involves the spatiotemporal information of individual motion. The GMSI not only generalizes the MSI but also reflects a flexible feature of a gait sequence. Along with the GMSI, we use the Principal Component Analysis(PCA) to reduce the dimensionality of the GMSI and the Nearest Neighbor(NN) for classification. We apply the proposed feature to NLPR database and compare it with the conventional MSI. Experimental results show the effectiveness of the GMSI.

Keywords : biometrics, gait recognition, human motion analysis, Motion Silhouette Image(MSI)

I. 서론

걸음걸이는 비강압적이고 인간 친화적이기 때문에 사용자 식별 시스템을 위한 매력적인 생체 정보다. 또한 걸음걸이는 저해상도 환경에서도 멀리서 취득 가능하다. 따라서 최근 걸음걸이 인식 연구는 생체인식 연구자들에게 많은 관심을 끌어들였다.

걸음걸이 인식 방법은 크게 두 가지 방식으로 구분할 수 있다. 모델 기반 인식 방법(model-based approach), 외형 기반 인식 방법(appearance-based approach). 첫째로, 모델 기반 인식 방법은 사람의 몸이나 동작을 모델링하고, 보폭의 동역학적인 특징이나 관절의 운동학적인 특징을 모델을 기반으로 추정하여 특징점(feature)을 획득한다[1,2]. 예를 들어, Lee와 Grimson은 카메라 방향에 수직으로 걷는 사람의 실루엣으로부터 모델템에 기반한 걸음걸이 정보를 획득하는 방법을 제안하였다[1]. Zhang은 걸음걸이 인식을 위해 5-연결 이족 운동 모델을 만들었다. 모델 기반 인식의 경우 우수한 인식 능력을 보여주나, 동작 모델링을 위한 비전 알고리즘이 불완전해 현실적으로 좋은 성능을 보이는 시스템을 구현하기 어렵고, 연산적인 비용이 상대적으로 높다는 단점이 있다.

한편, 외형 기반 인식 방법은 기초적인 구조에 상관없이 간결한 표현 방식을 통해 사람 몸의 전체 동작 패턴을 특성화한다. 이러한 표현 방식들은 몸 형태의 정적인 특성과 동적인 성질을 동시에 획득한다[3-5]. 예를 들어, Nixon은 데이터의 차원을 줄임과 동시에 다른 걸음걸이 시퀀스의 클래스 분리성(class separability)을 최적화 하는 방법을 제안하였다[5]. Lee는 동작 히스토리 영상(Motion History Image: MHI)에서 영감을 얻어 동작 실루엣 영상(Motion Silhouette Image: MSI)라는

외형에 기반한 걸음걸이 표현 방식을 제안하였다. 외형 기반 인식 방법은 모델 기반 인식 방법에 비해 표현과 구현이 간단하고, 연산이 빠르다는 장점이 있다.

본 연구에서는 걸음걸이 인식을 위해 MSI를 일반화시킨 걸음걸이 표현 방식인 일반적 동작 실루엣 영상(Generic Motion Silhouette Image: GMSI)을 제안한다. GMSI는 MSI와 유사하나 MSI보다 유연하며 높은 식별 능력을 가지고 있는 걸음걸이 표현 방식이다. [3]에서와 마찬가지로 특징점 차원을 줄이기 위해 주 성분 분석(Principal Component Analysis: PCA)를 적용하였고, 분류를 위해 최인접물 분류기(Nearest Neighbor classifier: NN)를 이용하였다. 그리고 MSI와 인식 성능을 비교하기 위해 제안한 시스템을 NLPR 걸음걸이 데이터베이스를 실험하였다.

본 논문은 다음과 같이 구성되어 있다. 2장에서는 실루엣 추출에서부터 제안한 걸음걸이 표현 방식을 이용한 특징점 추출 과정과 분류 전략을 소개한다. 3장에서는 실험 방법을 기술하고 실험 결과를 고찰하며, 끝으로 4장에서는 결론을 정리한다.

II. 제안한 걸음걸이 인식 시스템

제안한 걸음걸이 인식 시스템은 전처리 단계, 특징점 추출 단계와 인식 단계로 구성되어 있다. 전처리 단계에서 배경 제거 알고리즘을 이용하여 컬러 비디오 시퀀스로부터 실루엣 영상을 추출하고, 유사성 측정 성능을 높이기 위해 추출한 실루엣 영상의 크기를 정규화한다. 다음으로 정규화된 실루엣 영상으로부터 걸음걸이 특성을 효율적으로 획득하는 걸음걸이 특징점을 추출하고, 특징점으로부터 주 성분을 추출하기 위해 PCA를 적용한다. 끝으로 고유공간(eigenspace)으로 투영한 특징점들 간의 유사성을 측정하기 위해 유클리디안 거리(Euclidean distance)를 사용하며 분류 문제를 풀기 위해 NN 분류기를 이용하였다. 각 단계에 대한 자세한 설명은 아래에서 다룬다.

1. 전처리

배경 모델링을 이용한 배경 제거 알고리즘을 이용하여 동

* 책임저자(Corresponding Author)

논문접수 : 2007. 5. 16., 채택확정 : 2007. 6. 25.

홍성준, 이희성, 김은태 : 연세대학교 전기전자공학과, 생체인식연구센터(BERC)

(imjune@yonsei.ac.kr/4u2u@yonsei.ac.kr/etkim@yonsei.ac.kr)

※ The work was supported by the Korea Science and Engineering Foundation (KOSEF) through the Biometrics Engineering Research Center (BERC) at Yonsei University.



그림 1. 카메라 방향과 45도 각도로 걷는 사람의 비정규화된 실루엣 영상과 정규화된 실루엣 영상의 비교.

Fig. 1. Comparison between (a) original non-normalized and (b) normalized silhouette images of subject walking in the oblique view (45°).

작 영역을 분리할 수 있는데, 본 연구에서는 [6]에서 제안한 통계적인 배경 모델링 기법을 이용한 배경 제거 알고리즘을 적용하였다. 그리고 배경 제거를 통해 얻은 전경 실루엣 영상에 중간값 필터를 적용해 잡음 화소를 줄인다. 그리고 스케일링의 영향을 감소하기 위해 움직임에 따라 가변적인 실루엣 영상의 크기를 고정된 크기(148×88)로 조정한다. 실루엣 영상의 크기는 실루엣의 높이에 대해서 정규화하고, 같은 비율로 실루엣의 너비를 정규화 한다. 그림 1은 비정규화된 실루엣 영상과 정규화된 실루엣 영상의 예를 보여준다.

2. 걸음걸이의 표현

MSI는 실루엣의 정규화된 화소 정보를 갖는 그레이-레벨 영상이다[6]. MSI는 걸음걸이의 중요한 시공간적인 정보를 가지고 있고 다음과 같이 식으로 나타낼 수 있다.

$$MSI(u, v, t) = \begin{cases} 255 & \text{if } I(u, v, t) = 1 \\ \max(0, MSI(u, v, t-1) - 1) & \text{otherwise} \end{cases} \quad (1)$$

여기서 $I(u, v, t)$ 는 t 번째 이진화된 실루엣 영상의 좌표 (u, v) 의 밝기 값을 나타낸다. 그림 2는 MSI 이미지의 예를 보여준다.

본 논문에서는 MSI의 일반화된 표현 방법인 일반적인 동작 실루엣 영상(Generic Motion Silhouette Image: GMSI)를 제안한다. GMSI는 MSI 보다 유연한 특징을 가지며 다음과 같이 정의 된다.

$$GMSI(u, v, t) = \begin{cases} 255 & \text{if } I(u, v, t) = 1 \\ \max(0, GMSI(u, v, t-1) - d) & \text{otherwise} \end{cases} \quad (2)$$

여기서 d 는 새로운 실루엣 영상이 입력되어 GMSI를 갱신할 때, 기존에 생성된 GMSI의 이미지에 시간의 흐름을 반영하고 왜곡하는 정도를 나타내는 상실 인자(forgetting factor)를 나타내고, 다른 매개변수들은 MSI와 동일하다. 이때, MSI는 GMSI에서 $d=1$ 인 특수한 경우에 해당한다. 그림 3은 5가지 상실 인자($d=1, 5, 10, 20, \text{ and } 30$)를 갖는 GMSI의 예를 보여준다.

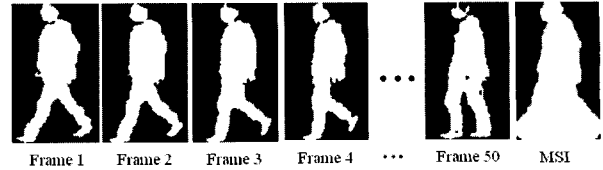


그림 2. MSI 이미지의 예.

Fig. 2. Examples of motion silhouette image.

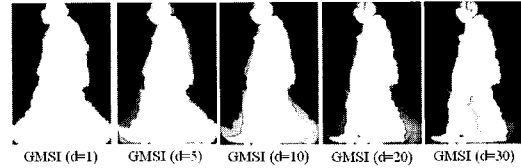


그림 3. GMSI 이미지의 예($d=1, 5, 10, 20, \text{ and } 30$).

Fig. 3. Examples of Generic Motion Silhouette Image ($d=1, 5, 10, 20, \text{ and } 30$).

3. 특징점 추출

주 성분 분석(Principal Component Analysis: PCA)는 데이터의 차원의 크기를 줄이고 상호 연관된 변수들 간의 관계를 분석하는데 이용되는 기법이다. PCA를 이용하면 원본 데이터를 고차원 공간에서 저차원 공간으로 투영시키는 몇 개의 주 성분 벡터를 획득할 수 있다. 이렇게 찾아낸 주 성분 벡터인 주 성분 벡터는 원본 데이터 집합의 변화 순으로 새로운 좌표 체계를 이룬다. 학습을 위한 표본 집합이 주어졌을 때, $x_i^j \in R^M$ 를 클래스 i 에 속하는 j 번째 GMSI라고 하고, N_i 를 클래스 i 에 속하는 GMSI의 수라고 하자. 그럼, 학습에 사용되는 GMSI의 전체 개수는 $N = \sum_{i=1}^C N_i$ 이고, 전체 학습 집합은 $[x_1^1, \dots, x_1^{N_1}, x_2^1, \dots, x_2^{N_2}, \dots, x_C^1, \dots, x_C^{N_C}] \in R^{M \times N}$ 로 나타낼 수 있다. 이 때, 학습 집합에 대한 평균값 μ_x 와 공분산 행렬 Σ_x 은 다음과 같이 구한다.

$$\Sigma_x = \frac{1}{N} \sum_{i=1}^C \sum_{j=1}^{N_i} (x_i^j - \mu_x)(x_i^j - \mu_x)^T \quad (3)$$

$$\mu_x = \frac{1}{N} \sum_{i=1}^C \sum_{j=1}^{N_i} x_i^j \quad (4)$$

변환 행렬 $\Phi_{PCA} = [\Phi_1, \Phi_2, \dots, \Phi_p]$ 은 공분산 행렬 Σ_x 에 대해 특이 값 분해(singular value decomposition: SVD)를 통해 얻을 수 있고, $\Phi_1, \Phi_2, \dots, \Phi_p$ 는 $p (< M)$ 개의 고유값에 해당하는 공분산 행렬 Σ_x 의 고유벡터들이다. 이 고유 벡터들은 서로 직교하는 단위 벡터들이며, 고유공간(eigenspace)을 생성한다. PCA 과정의 마지막 단계는 원본 GMSI를 고유공간에 투영하는 과정이다. 벡터화된 원본 GMSI x_i^j 가 주어졌을 때, 아래와 같이 원본 GMSI에 변환 행렬 Φ_{PCA} 을 곱함으로써 고유공간으로 투영할 수 있다.

$$y_i^j = \Phi_{PCA}^T x_i^j \in R^p \quad (5)$$

PCA를 통해 입력 샘플들의 차원을 줄일 수 있으며, 변환된 고유공간상의 벡터 $y_i^j \in R^p$ 는 원본 GMSI x_i^j 의 중요한 정보들을 간직한다.

3. 인식

고유공간에 투영된 학습 데이터와 실험 데이터 간의 유사성(similarity) 측정을 위해 유클리디안 거리 (Euclidean distance)를 이용하였다. 고유공간 상의 학습 데이터와 실험 데이터 간의 유클리디안 거리 E 는 다음과 같이 계산할 수 있다.

$$E(y, y_i) = \sqrt{(y - y_i)^T (y - y_i)} \quad (6)$$

여기서 y 는 고유공간에 투영된 실험 데이터이며, y_i 는 고유공간에 투영된 클래스 i 에 해당하는 학습 데이터이다.

인식 과정은 간단한 분류 방법인 최근접물 분류기(Nearest Neighbor classifier: NN)를 사용하였다. NN 분류기는 입력된 실험 데이터를 데이터베이스에 저장된 가장 가까운 학습 데이터의 클래스로 분류한다. 다시 말해 실험 데이터는 고유공간 상의 학습 데이터의 특징점 벡터와 고유공간 상의 실험 데이터의 특징점 벡터간의 거리를 가장 작게 하는 클래스 c 로 다음과 같이 분류할 수 있다.

$$c = \underset{i}{\operatorname{arg\,min}} E(y, y_i^j) \quad \text{for } \forall j = 1, 2, \dots, N_i, \forall i = 1, 2, \dots, C \quad (7)$$

여기서 N_i 는 클래스 i 에 해당하는 학습 데이터의 수를 말한다.

III. 실험 및 고찰

실험을 위해 NLPR 걸음걸이 데이터베이스를 이용하였다. NLPR 걸음걸이 데이터베이스는 삼각대로 고정된 디지털 카메라로 실외 환경에서 이들에 걸쳐 촬영된 걸음걸이 영상 시퀀스로 구성되어 있다. 데이터베이스는 모두 20명의 사람을 포함하며, 카메라 시점에 대해 평행 방향(0도), 대각선 방향(45도), 그리고 수직 방향(90도)으로 견도록 하여, 각 방향에 대해 4개의 시퀀스를 촬영해, 총 240(=20×3×4) 걸음걸이 시퀀스로 이루어져 있다. 촬영된 걸음걸이 영상의 해상도는 352×240이다[7]. 그림 4는 NLPR 걸음걸이 데이터베이스의 세 가지 방향으로 걷는 걸음걸이 영상의 예를 보여준다.

1. 인식 성능

먼저 GMSI의 걸음걸이 인식 성능을 알아보기 위해 상실 인자의 변화에 따른 인식률을 평가하였다. 실험에서 상실 인자는 1에서부터 20까지 변화를 주었다. GMSI를 생성한 후에 PCA 과정을 거쳐 20명의 이미지로부터 만들어낸 고유공간



그림 4. 세 가지 방향으로 걷는 NLPR 데이터베이스 걸음걸이 영상의 예.

Fig. 4. Some examples in NLPR database at 3 different viewing angles.

변환 행렬을 이용해 원본 GMSI를 고유공간 상의 특징점 벡터로 투영 시켰다. 본 실험에서는 고유값의 누계가 95%가 되는 고유벡터의 집합을 고유공간 변환 행렬로 사용했다.

표 1. GMSI의 상실 인자에 따른 걸음걸이 인식률.

Table 1. Correct classification rates for different forgetting factors of GMSI.

d	0°	45°	90°
MSI	68.75%	80%	72.5%
GMSI (d=2)	72.5%	83.75%	73.75%
GMSI (d=3)	75%	83.75%	73.75%
GMSI (d=4)	73.75%	77.5%	72.5%
GMSI (d=5)	75%	72.5%	72.5%

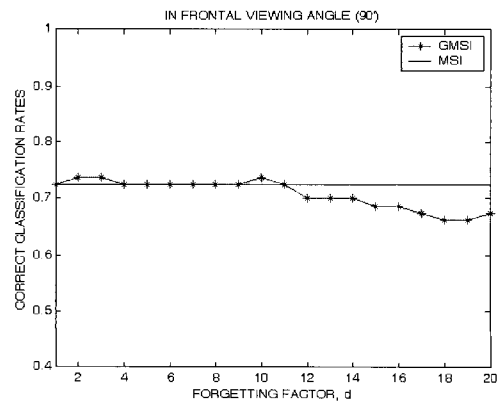
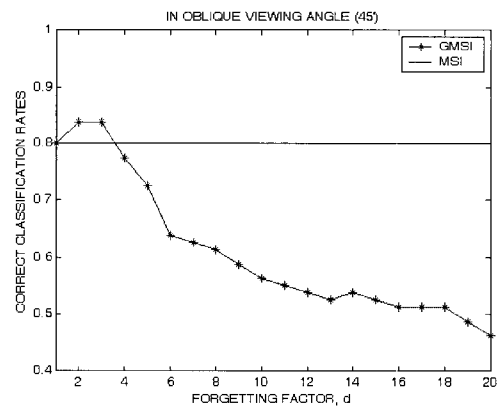
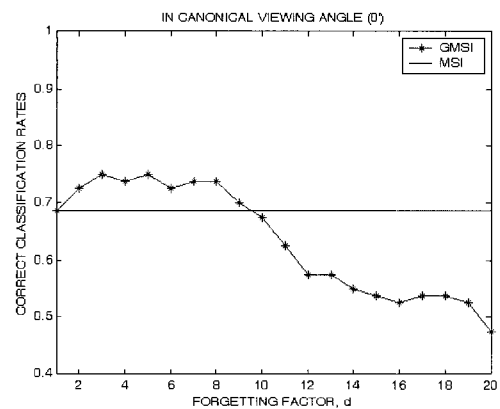


그림 5. GMSI의 상실 인자에 따른 걸음걸이 인식률.

Fig. 5. Correct classification rates for different forgetting factors of GMSI.

한편, NLPR 걸음걸이 데이터베이스의 세 방향으로 걷는 걸음걸이 시퀀스에 대해 각각 성능 평가를 했으며, 제안한 알고리즘의 성능 평가를 위해 leave-one-out 교차 인식(leave-one-out cross validation)을 이용했다. Leave-one-out 교차 인식은 한 개의 샘플을 실험 데이터로 이용하고, 나머지 샘플을 학습 데이터로 사용하여 인식 여부를 확인하고, 같은 실험을 실험 데이터에 해당하는 샘플을 바꿔가며 반복하며 인식 성능을 평가하는 방식을 말한다.

표 1과 그림 5는 실험 결과를 보여준다. 카메라 시점에 90도로 걷는 걸음걸이 데이터베이스에서는 GMSI의 상실 인자가 3과 5일 때 가장 좋은 인식률인 75%를 보여주었다. 카메라 시점에 45도로 걷는 걸음걸이 데이터베이스에 대해서는 GMSI의 상실 인자가 2와 3일 때, 가장 좋은 인식률 83.75%를 보여주었다. 한편, 카메라 시점에 정면으로 걷는 걸음걸이 데이터베이스에 대해서는 GMSI의 상실 인자로 2, 3, 10을 사용하였을 때 가장 좋은 인식률인 73.75%를 보여주었다. 따라서 2, 3, 또는 5를 상실 인자로 취한 GMSI를 특징점으로 이용할 때 기존의 MSI를 특징점으로 사용할 때보다 더 좋은 성능을 보여주고 있음을 보여준다. 하지만 큰 상실 인자 ($d > 6$)를 이용하는 경우 인식률이 감소하고 있음을 볼 수 있다. 따라서 이 실험 결과는 작은 상실 인자를 갖는 GMSI는 시간적인 정보를 효율적으로 사용하여 좋은 성능을 나타내며, 큰 상실 인자를 갖는 GMSI는 GMSI가 새로운 실루엣에 의해 갱신되는 동안 의미 있는 공간 정보를 잃어버리면서 좋지 않은 성능을 보여줌을 확인할 수 있다.

2. 고유값에 따르는 인식 성능의 비교

본 실험에서는 원본 GMSI를 고유공간 상의 특징점 벡터로 투영시키는 고유공간 변환 행렬을 이루는 학습 데이터의

표 1. 고유값에 따른 걸음걸이 인식률의 비교

Table 1. Correct classification rates with different percentages of accumulated values of eigenvalues.

d	0°		
	90%	95%	99%
1	71.25%	68.75%	75%
2	72.5%	72.5%	80%
3	71.25%	75%	80%
4	75%	73.75%	76.25%
5	75%	75%	76.25%
d	45°		
	90%	95%	99%
1	81.25%	80%	85%
2	83.75%	83.75%	87.5%
3	82.5%	83.75%	85%
4	78.75%	77.5%	85%
5	72.5%	72.5%	77.5%
d	90°		
	90%	95%	99%
1	73.75%	72.5%	75%
2	75%	73.75%	75%
3	73.75%	73.75%	73.75%
4	73.75%	72.5%	75%
5	72.5%	72.5%	75%

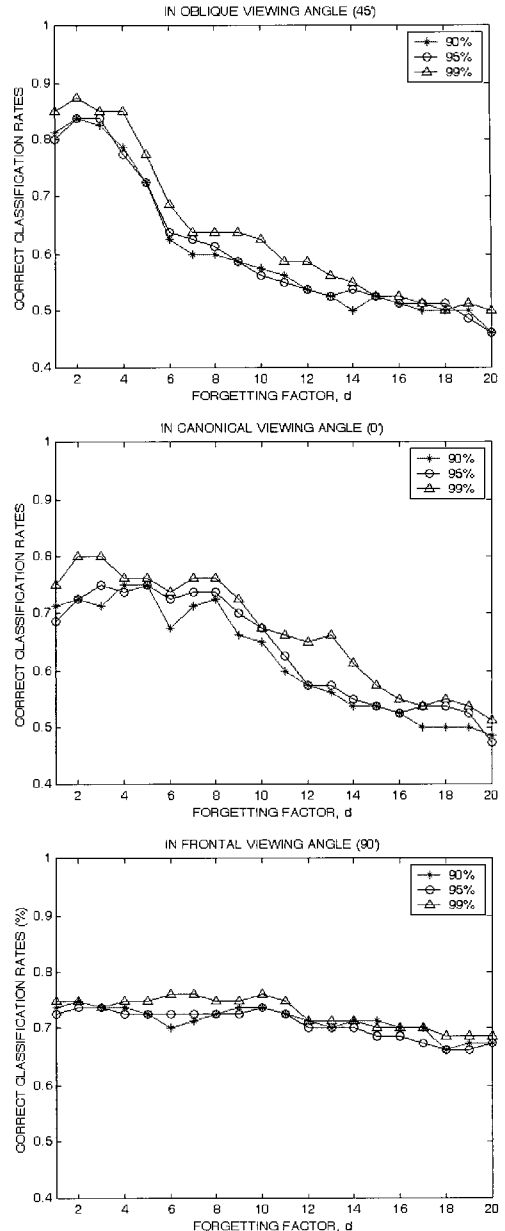


그림 6. 고유값에 따른 걸음걸이 인식률의 비교

Fig. 6. Correct classification rates with different percentages of accumulated values of eigenvalues.

공분산 행렬의 고유벡터의 수에 따른 걸음걸이 인식 성능을 평가한다. 고유값의 누계가 90%, 95%, 99%에 해당하는 고유벡터로 이루어진 변환 행렬을 가지고 실험을 하였고, 표 2와 그림 6은 그 결과를 보여준다. 실험 결과를 보면 걸음걸이 인식률은 고유벡터의 수에 비례적인 관계가 있음을 알 수 있다. 또한 고유공간 변환 행렬을 이루는 고유벡터의 수가 변하더라도, 작은 상실 인자를 갖는 GMSI를 특징점으로 이용한 시스템이 MSI를 특징점으로 이용한 시스템의 인식률을 증가하는 것을 확인할 수 있다.

IV. 결론

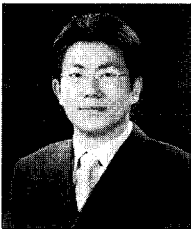
본 연구에서는 걸음걸이 인식을 위해 MSI를 일반화시킨 걸음걸이 표현 방식인 일반적 동작 실루엣 영상(Generic

Motion Silhouette Image: GMSI)을 제안했다. GMSI는 MSI와 유사하나 MSI보다 유연하며 높은 식별 능력을 가지고 있는 걸음걸이 표현 방식이다. MSI는 상실 인자를 1로 갖는 GMSI다. NLPR 걸음걸이 데이터베이스를 이용한 실험 결과는 작은 상실 인자 ($d < 6$)를 갖는 GMSI를 특징점으로 사용한 시스템이 기존의 MSI를 특징점으로 사용한 시스템을 능가하는 성능을 보여줬다.

그러나 실제 시스템을 구현함에 있어 상실 인자의 변화가 걸음걸이 인식률에 영향을 미치기 때문에 GMSI의 상실 인자로서 어떤 값을 취해야 하는가에 대한 일반적인 방법을 제시하기 어렵다. 이 문제를 해결하기 위한 최선의 방법은 학습 데이터 집합을 학습 데이터 집합(training set)과 유효성 데이터 집합(validation set)으로 나누는 것이다. 여기서 유효성 데이터 집합은 실제 걸음걸이 인식 시스템의 GMSI의 상실 인자를 결정하는데 사용하는데, 학습 데이터 집합에 대해 유효성 데이터 집합을 실험하여 가장 좋은 성능을 보이는 GMSI의 상실 인자를 실제 걸음걸이 인식 시스템의 GMSI의 상실 인자로 선택함으로써 좋은 성능을 갖는 GMSI의 상실 인자 선택 문제를 해결할 수 있다.

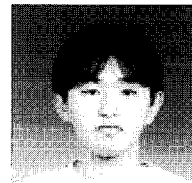
참고문헌

- [1] L. Lee and W. Grimson, "Gait analysis for recognition and classification," *Proc. IEEE Conf. Face and Gesture Recognition*, pp. 55-62, 2002.
- [2] R. Zhang, B. Christian, and M. Dimitris, "Human gait recognition," *Proceeding of the 2004 IEEE Computer Vision and Pattern Recognition Workshops*, pp. 18-28, 2004.
- [3] T. Lam and R. Lee, "A new representation for human gait recognition: Motion Silhouette Image(MSI)," *Lecture Notes in Computer Science*, vol. 3832, pp. 612-618, 2006.
- [4] H. Murase and R. Sakai, "Moving object recognition in eigenspace representation: Gait analysis and lip reading," *Pattern Recognition Letter*, vol. 17, pp. 155-162, 1996.
- [5] P. Huang, C. Harris, and M. Nixon, "Human gait recognition in canonical space using temporal templates," *Vision Image and Signal Processing*, vol. 146, 93-100, 1999.
- [6] T. Horprasert, D. Harwood, and L. S. Davis, "A statistical approach for real-time robust background subtraction and shadow detection," *Proc. Int. Conf. Computer Vision*, 1999.
- [7] L. Wang, T. Tan, H. Ning, and W. Hu, "Silhouette analysis-based gait recognition for human identification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, pp 1505-1518, 2003.



홍성준

2005년 연세대학교 전기전자공학부(공학사). 2005년~현재 동 대학원 전기전자공학과 석·박사 통합과정 재학중. 관심분야는 Multimodal Biometrics, Machine Learning, Optimization Problem.



이희성

2003년 연세대학교 기계전자공학부(공학사). 2005년 동 대학원 전기전자공학과(공학석사). 2005년~현재 동 대학원 전기전자공학과 박사과정 재학중. 관심분야는 Computational Intelligence, Computer Vision, Pattern Recognition.



김은태

1992년 연세대학교 전자공학과(공학사). 1994년 연세대학교 전자공학과(공학석사). 1999년 연세대학교 전자공학과(공학박사). 1999년~2002년 국립한경대학교 제어계측공학과 조교수. 2002년~현재 연세대학교 전기전자공학부 부교수. 2003

년 University of Alberta, visiting researcher. 2003년 대한 전자공학회 해동상 수상. 관심분야는 Computational Intelligence.