

# 3차 텐서기반 MPCA 방법을 이용한 심전도신호의 개인식별

변영현\*, 이재진\*\*, 정하영\*\*, 한하영\*\*, 곽근창\*\*\*

\*조선대학교, 제어계측공학과

\*\*조선대학교, 전자공학부

## Individual Identification on Electrocardiogram using Third-Order Tensor-Based MPCA

Yeong-Hyeon Byeon\*, Jae-Jin Lee\*\*, Ha-Young Jeong\*\*,

Ha-Young Han\*\*, Keun-Chang Kwak\*\*\*†

\*Department of Control and Instrumentation Engineering, Chosun University, Gwangju, Korea

\*\*Department of Electronics Engineering, Chosun University, Gwangju, Korea

(Received : May. 18, 2018, Revised : Jun. 18, 2018, Accepted : Jun. 25, 2018)

**Abstract** : In this paper, performance of individual identification on electrocardiogram using third-order tensor-based MPCA(Multilinear Principal Component Analysis) is performed. This method preserves the data structure by extracting features directly from the tensor representation without structural transformation of the data due to the vectorization process in order to reduce the dimension. It is also less susceptible to small data problems because it can learn more compact and potentially useful representation, and it can efficiently handle large tensors. Here, the third-order tensor is formed by reordering the one-dimensional electrocardiogram signal into a two-dimensional matrix and then taking the time frame into account. Physionet's PTB(Physikalisch-Technische Bundesanstalt) diagnostic database for performance evaluation is used, and MPCA showed 91.85% accuracy.

**Keyword** : Electrocardiogram, Multilinear principal component analysis, Individual identification, Tensor

### 1. 서론

개인식별 기술은 개인의 유일한 특성인 행동적 또는 신체적 특징을 이용해 사용자를 식별하는 기술이다. 현재 보안, 은행, 출입관리, 의료, 오락 등 광범위하게 서비스 범위가 확대되고 있다. 얼굴과 지문 같은 생체 정보를 이용한 개인식별은 위조, 변장 할 염려가 있고 개인식별을 위해서 시스템이 설치된 특별한 장소로 이동해야하는 취약점이 있어서, 최근 신체 내부의 생체 신호를 이용한 개인식별이 연구되고 있다[1]. 건강하고자 하는 인간의 욕구가 기술이 발달함에 따라 간편하게 몸 상태를 측정하여 대비할 수 있는 방안들을 요

구하고 있다. 그래서 여느 회사들은 작고 가볍고 편리하며 다양한 기능들이 포함된 웨어러블 장치들을 개발하고 출시해왔다. 특히 건강 체크에 중요한 다양한 생체 신호들의 수집기능은 필수로 되어가고 있다. 심전도(electrocardiogram)는 심장주기동안 전위의 변화를 측정하는 것으로 그 측정 절차가 간소화되고 편리하게 변화되었고 개인식별에서도 좋은 성능을 보인 연구들이 보고되고 있다[2-5]. 심전도 신호로 개인식별을 하는 것은 심전도 신호가 잠음에 민감하여 동일한 사람에게서 언더라도 편안한 상태, 흥분한 상태, 화난 상태, 운동상태 등에 따라 성능에 큰 차이를 보이기도 한다[1]. 이러한 환경에서도 안정적인 식별성능을 얻기 위한 연구들이 진행되고 있다.

센서와 메모리 그리고 네트워크 기술이 발전하면서, 수 많은 데이터들이 다양한 분야에서 하루가 멀다하게 생성되고 있다. 이러한 빅데이터의 분야에서 성공하려면 세부 특징을 효과적으로 학습하는 게 중요하다. 대부분 빅데이터들은 다차원이고 그것들은 보통 다차원 배열로 표현될 수 있다. 그래서 모바일 인터넷, 클라우드 컴퓨팅, 빅데이터 등에서 텐서 기반 컴퓨팅이 부

† Corresponding Author

성명 : 곽근창

소속 : 조선대학교 전자공학부

주소 : 광주 서구 필문대로 309 조선대학교

전화 : 062-230-6086

E-mail : kwak@chosun.ac.kr

상하고 있다. 다차원 데이터로부터 세부 특징들을 텐서기반으로 학습하기 위해 차원축소(dimensionality reduction) 방법인 다선형 부분공간 학습(multilinear subspace learning)이 개발되었다. 다선형 부분공간 학습의 목적은 고차원 텐서로부터 저차원 텐서로 직접 사상(direct mapping)시키는 것을 학습하는 것이다. 여기서 텐서는 벡터와 행렬의 일반화된 의미를 가진다. 스칼라는 0차 텐서, 벡터는 1차 텐서, 행렬은 2차 텐서, 그리고 3차 이상은 고차텐서라고 부른다.

현실에서 텐서 데이터는 보통은 고차공간에서 명시된다. 고차공간에서 데이터의 직접적인 연산은 두 가지 차원의 저주를 야기한다. 첫째는 고차원 데이터를 다루는 것은 프로세싱 부하와 자원을 많은 사용하게 되어 계산비용이 커지고, 둘째는 사용가능한 데이터의 개수가 그 데이터의 차원에 비해 적을 때 기존 수단들은 적절치 못해 많은 문제들에 타당성을 잃게 한다.

차원축소는 대부분의 물리적 또는 구조적 정보를 유지한 채 고차원 데이터를 저차원으로 표현하는 변환 방법이고 저차원 표현은 특징 추출 과정으로서도 인식할 수 있다. 전통적인 부분공간 학습 알고리즘들은 PCA(Principal Component Analysis), ICA(Independent Component Analysis), LDA(Linear Discriminant Analysis), CCA(Canonical Correlation Analysis), PLS(Partial Least Squares) 같이 벡터에서 연산되는 선형 알고리즘들이었다. 1차 이상의 텐서 데이터에 대해 선형 부분공간 학습(linear subspace learning) 방법들을 적용하기 위해서는 우선 텐서를 벡터로 재정렬(reshape)해야 했다. 이런 벡터화(vectorization)는 2가지 근본적인 제한을 야기한다. 첫째로 벡터화는 원래 데이터에서 보통의 구조와 상관관계를 파괴하고 원래 데이터에서 내포된 고차 의존성과 가외성을 축소시킨다. 그리고 잠재적으로 세부 또는 유용한 특징을 잃게 된다. 둘째로 고차 텐서 데이터로부터 벡터화된 데이터는 매우 고차원이 되어 계산과 메모리 부하가 크고 데이터 수가 차원수보다 적은 문제를 발생시킨다.

텐서기반 컴퓨팅 및 모델링은 매 20년마다 발돋움을 하였다. 스칼라 수준부터 행렬 수준까지 진행되고 그 다음 블록 행렬 수준까지 진행되었다. 이제 행렬기반 컴퓨팅에서 텐서기반으로 이행할 시기에 왔다. 기존의 선형 사상(linear mapping)과는 반대로 다선형 사상(multilinear mapping)은 텐서를 벡터로 재정렬시키지 않는다. 다선형 사상으로 학습된 특징들은 그리드같은(grid-like) 구조와 규칙들로 축소된다. 이것은 3가지 주요 이점을 가진다. 첫째로 데이터의 구조를 보존하고 둘째로 더 세부적이고 유용한 특징을 학습할 수 있고 적은 데이터 수의 문제에 덜 심각해진다. 셋째로 큰 텐서 데이터를 저차원에서 효과적으로 계산할 수 있다(6).

본 논문에서는 3차 텐서기반 MPCA 방법을 이용한 심전도신호의 개인식별에 대한 성능을 비교한다. 2절에서는 특징추출 및 분류 방법에 대해 기술하고, 3절에서는 실험결과 및 분석을 설명하고, 4절에서 결론짓는다.

## 2. 기존의 심전도기반 개인식별 기법

기존의 심전도기반 개인식별을 위해 벡터 기반 기술로는 주성분 분석 기법(PCA: Principal Component Analysis)[7-8]과 매트릭스 기반의 기술로는 2차원 주성분 분석 기법 (2D PCA: two-Dimensional Principal Component Analysis)[9] 있다.

### 2.1 주성분 분석 기법(PCA)

주성분 분석 기법은 분산까지의 통계적 성질을 이용한 2차 통계적 기법으로, 고차원의 입력 데이터의 차원을 효율적으로 축소하는 데에 주로 사용된다. 주성분 분석 기법을 요약 하자면, 전체 데이터를 가지고 그것들의 분산이 큰 몇 개의 고유 방향에 대한 축으로 선형 투영 시켜 차원을 줄이는 방법을 말한다(7-8). 주성분 분석 기법을 이용한 학습 방법은 다음과 같다.

2차원의  $(n \times m)$  데이터를 1차원  $N(n \times m)$  벡터로 변환한다.  $P$ 개의 학습 벡터를  $X = [x^1|x^2 \dots |x^P]$ 와 같이 행렬로 나타낼 수 있으며, 각 벡터와 평균 벡터의 차는 다음과 같다.

$$\bar{x}^i = x^i - mean, \quad mean = \frac{1}{P} \sum_{i=1}^P x^i \quad (1)$$

식(1)에서 구해진  $P$ 개의  $\bar{x}^i$  벡터를 이용해서  $\bar{X} = [\bar{x}^1|\bar{x}^2 \dots |\bar{x}^P]$ 를 구할 수 있으며, 이를 이용한 공분산 행렬(covariance matrix)  $\Omega$ 은 식(2)과 같다.

$$\Omega = \bar{X} \bar{X}^T \quad (2)$$

여기서 공분산 행렬에 의해 얻어지는 고유치(eigenvalue)는 분산을 최대로 하는 방향을 나타내고 이 고유치에 대응하는 고유 벡터(eigenvector)는 특정 방향의 변동성을 나타낸다. 특징벡터는 학습 벡터를 고유벡터  $\bar{V} = [v^1|v^2 \dots |v^P]$ 에 투영 시켜 다음과 같이 구할 수 있다.

$$\tilde{x}^i = V^T \bar{X}^i \quad (3)$$

### 2.2 2차원 주성분 분석 기법(2DPCA)

2차원 주성분 분석 기법은 2차원 데이터를 1차 벡터로 변화하지 않고 직접 공분산을 구하는 방식을 취하는 방법이다(9). 2차원 주성분 분석 기법의 학습 방법을 간단히 살펴보면 다음과 같다.

학습 데이터를  $A_k (k=1, 2, \dots, P)$ 라 하고 학습 데이터 크기는  $R \times C$ 로 가정 한다. 전체 데이터들의 평균 데이터  $mean$ 은 식(4)와 같이 나타낼 수 있다.

$$mean = \frac{1}{P} \sum_k A_k \quad (4)$$

공분산 행렬  $G$ 는 식(5)과 같이 구할 수 있다.

$$G = \frac{1}{P} \sum_{k=1}^P (A_k - \text{mean})^T (A_k - \text{mean}) \quad (5)$$

여기서 2차원 주성분 분석 기법은 데이터를 2차원 형태로 직접 공분산 행렬  $G$ 을 계산하게 되어 수평방향의 공분산 행렬,  $R \times R$ 크기의 공분산 행렬을 얻는다. 즉, 2차원 주성분 분석 기법에서는 수평 방향으로 축소한다는 것을 나타낸다. 고유 벡터  $\Psi$ 와 고유치  $\Lambda$ 는 식(6)에서 구해진다.

$$G\Psi = \Psi\Lambda \quad (6)$$

여기서, 공분산 행렬  $G$ 의 가장 큰  $d$ 개의 고유치에 상응하는 고유벡터를  $V_1, V_2, V_3, \dots, V_d$ 라 한다. 특징 행렬  $C_k$ 는 학습 얼굴 영상 행렬  $A_k$ 를 투영 행렬  $\Psi = [V_1, V_2, V_3, \dots, V_d]$ 에 투영시켜 다음과 같이 구할 수 있다.

$$C_k = A_k\Psi \quad (7)$$

2차원 주성분 분석 기법은 기존의 주성분 분석 기법을 이용할 때보다 매우 작은 공분산 행렬을 사용하기 때문에 계산 속도가 빠르고, 인식률 또한 더 좋다는 장점을 가지고 있고, 단점으로는 특징 행렬의 크기가 크다는 것이다. 이는 많은 메모리를 필요로 하며 매칭 부분의 유사도 측정 시간이 매우 길게 소요되게 하는 요인이 된다[10-12].

### 3. 다선형 주성분 분석

#### 3.1 전처리

센서에서 받아온 데이터는 여러 잡음들이 섞여있기 때문에 정규화와 스파이크를 제거하고 R피크검출을 한다. 검출된 R피크를 가지고 전후 400프레임을 입력으로 사용한다.

#### 3.2 특징추출

다선형 주성분 분석 기법(MPCA; Multilinear Principal Component Analysis)은 주성분 분석 기법의 다선형 확장이며, 최근에 Lu[13]에 의해 결음새 인식에 성공적으로 적용되어졌다. 다선형 주성분 분석 기법은 2차원 주성분 분석법과 같이 얼굴 이미지를 1차 벡터로 변환하지 않고 직접 공분산을 구하는 방식을 취하는 방법이다. 다선형 주성분 분석 기법과 2차원 주성분 분석법과 차이가 있다면, 2차원 주성분 분석법의 경우 입력데이터의 크기가  $w \times h$ 인 경우  $w \times w$  또는  $h \times h$ 의 크기를 가지는 고유벡터 중 하나만을 구하여 사용하는데 다중선형 주성분 분석법의 경우는 두 가지 형태의 고유벡터를 모두 구하여 이용한다.

다선형 주성분 분석 기법을 간단히 설명하면 다음과 같다. 텐서들은 전형적으로  $A$ 로 표시하고, 텐서의 요소는  $N$ 에 의해 표시된다.  $N$ 은 텐서 대상의 차수를 나타내며 각 지수는 하나의 모드를 정의한다. 즉, 벡터들은 1차 텐서이고 행렬들은 2차 텐서이다.  $N > 2$ 를 가진 텐서들은 고차의 벡터와 행렬의 일반화된 형태로써 보일 수 있다. 텐서의 요소는 괄호에서 지수를 가지고 표시되며,  $N$ 차 텐서는  $A \in R^{I_1 \times I_2 \times \dots \times I_N}$ 이다. 그것은  $N$ 개의 지수들  $i_n, n = 1, 2, \dots, N$ 에 의해 표시되고, 각  $i_n$ 은  $A$ 의  $n$ -모드를 할당한다.  $A \times_n U$ 에서 행렬  $U \in R^{I_n \times I_n}$ 에 의한 텐서  $A$ 의  $n$ -모드 곱은 다음과 같은 텐서로 표현되어진다.

$$(A \times_n U)(i_1, \dots, i_{n-1}, j_n, j_{n+1}, i_{n+1}, \dots, i_N) = \sum_{i_n} A(i_1, \dots, i_N) \cdot U(j_n, i_n) \quad (8)$$

두 텐서의  $A, B \in R^{I_1 \times I_2 \times \dots \times I_N}$ 의 스칼라 곱은 다음과 같이 정의된다.

$$\langle A, B \rangle = \sum_{i_1} \sum_{i_2} \dots \sum_{i_n} A(i_1, i_2, \dots, i_n) \cdot B(i_1, i_2, \dots, i_n) \quad (9)$$

텐서  $A$ 의 프로베니우스 놈 (Frobenius norm)은  $\|A\|_F = \sqrt{\langle A, A \rangle}$  과 같이 정의 된다. 표준 다선형 대수식에 따라 텐서  $A$ 는 다음의 곱 형태로 나타난다.

$$A = S \times_1 U^{(1)} \times_2 U^{(2)} \times \dots \times_N U^{(N)} \quad (10)$$

여기서,  $S$ 는 다음과 같다.

$$S = A \times_1 U^{(1)T} \times_2 U^{(2)T} \times \dots \times_N U^{(N)T} \quad (11)$$

$U^{(N)} = (u_1^n u_2^n \dots u_n^n)$ 로  $I_n \times I_n$ 의 직교행렬이다.  $R^{I_1} \otimes R^{I_2} \dots \otimes R^{I_N}$ 공간에서  $M$ 텐서의 집합이  $\{Y_m, m = 1, \dots, M\}$  일 때 총 분산은 식(12)과 같다.

$$\Psi_Y = \sum_{m=1}^M \|Y_m - \bar{Y}\|_F^2 \quad (12)$$

$$\bar{Y} = \frac{1}{M} \sum_{m=1}^M Y_m \quad (13)$$

여기서,  $\bar{Y}$ 는 평균 텐서,  $\Psi_Y$ 는 텐서 총 분산을 나타낸다. 여기서  $n$ -모드 총 분산 행렬은 다음과 같다.

$$C_{T_Y} = \sum_{m=1}^M (A_m^{(n)} - \bar{A}^{(n)})(A_m^{(n)} - \bar{A}^{(n)})^T \quad (14)$$

여기서  $A_m^{(n)}$ 은  $A_m$ 의 n-모드 전개 행렬(unfolded matrix)이다. 다선형 주성분 분석 기법은 텐서기반 스캐터 척도를 최대화하도록 한다. 이 문제는  $N$ 개의 선형 최적화 문제로 분해되어져서 해결되어질 수 있다. 모든 다른 투영행렬  $\tilde{U}^{(1)}, \dots, \tilde{U}^{(n-1)}, \tilde{U}^{(n+1)}, \dots, \tilde{U}^{(N)}$ 이 주어졌을 때, 행렬  $\{\tilde{U}^{(n)} \in R^{I_n \times P_n}, n=1, \dots, N\}$ 의 가장 큰  $P_n$ 개의 고유치에 대응하는 고유벡터들로 구성하는  $\Psi_Y$ 를 식 (15)과 같이 최대화 한다. 고유 벡터는 식(16)과 같이 구해 질 수 있다.

$$\{\tilde{U}^{(n)}, n=1, 2, \dots, N\} = \arg \max_{\tilde{U}^{(1)}, \tilde{U}^{(2)}, \dots, \tilde{U}^{(N)}} \Psi_Y \quad (15)$$

$$\Phi^{(n)} = \sum_{m=1}^M (X_{m(n)} - \bar{X}_{(n)}) \cdot \tilde{U}_{\Phi^{(n)}} \cdot \tilde{U}_{\Phi^{(n)}}^T \cdot (X_{m(n)} - \bar{X}_{(n)})^T \quad (16)$$

여기서,  $\tilde{U}_{\Phi^{(n)}} = (\tilde{U}^{(n+1)} \otimes \tilde{U}^{(n+1)} \otimes \dots \otimes \tilde{U}^{(N)} \otimes \tilde{U}^{(1)} \otimes \tilde{U}^{(2)} \otimes \dots \otimes \tilde{U}^{(n-1)})$  이다.

얻어진 투영행렬  $\tilde{U}^{(n)}$ 은  $\prod_{n=1}^N P_n$ 로서 보일 수 있

다. 그러나 그것들의 모두가 인식을 위해 유용하지는 않기 때문에 클래스 분별력을 통해 선택되어질 수 있다. 클래스 분별력에 의해 순서화된 특징벡터는 다음과 같이 계산되어진다[14].

### 3.3 특징분류

분류에서 특징벡터간의 거리는 분류 모듈의 성능을 결정하기 때문에 매우 중요하다. 여기에서는 맨하탄 거리(L1), 유클리디안 거리(L2), 각도거리(AD: Angle Distance), 마할라노비스 거리(MAH: Mahalanobis), 수정된 각도 거리(MAD: modified angle distance)를 사용한다. 아래 수식들은 각 거리 공식을 나타낸다.

$$L1 = \sum_{h=1}^H |a(h) - b(h)| \quad (17)$$

$$L2 = \sqrt{\sum_{h=1}^H [a(h) - b(h)]^2} \quad (18)$$

$$MAH = - \sum_{h=1}^H a(h) \cdot b(h) \quad (19)$$

아래는 각도 거리(AD:angle distance)를 계산하는 공식을 나타낸다.

$$AD = - \frac{\sum_{h=1}^H a(h) \cdot b(h)}{\sqrt{\sum_{h=1}^H a(h)^2 \sum_{h=1}^H b(h)^2}} \quad (20)$$

수정된 각도 거리를 계산하는 공식은 다음과 같다.

$$MAD = - \frac{\sum_{h=1}^H a(h) \cdot b(h)/g(h)}{\sqrt{\sum_{h=1}^H a(h)^2 \sum_{h=1}^H b(h)^2}} \quad (21)$$

여기에서  $g(h)$ 는 가중치벡터를 나타낸다. 검증데이터의 유사도를 계산하기 위해 최근접 분류기(nearest neighbor classifier)을 적용한다[13].

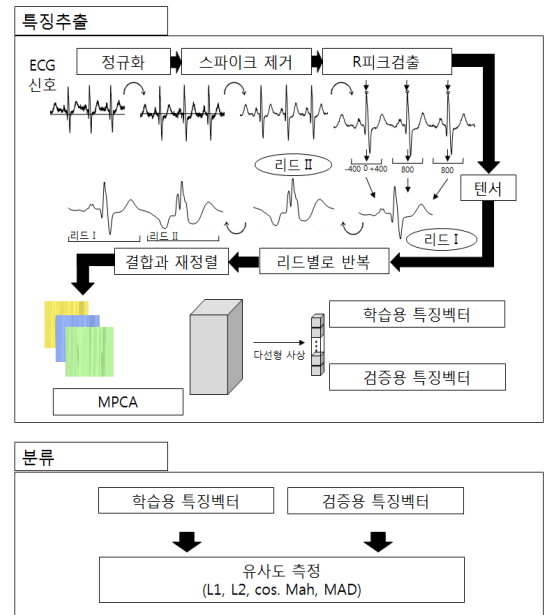


Figure 1. Diagram of proposed method.

## 4. 실험 및 결과 분석

본 실험에서 사용한 컴퓨터 사양은 Intel I7-7700, NVIDIA Geforce GTX 1060이고, ECG데이터는 PhysioNet의 PTB 진단 데이터베이스이다.

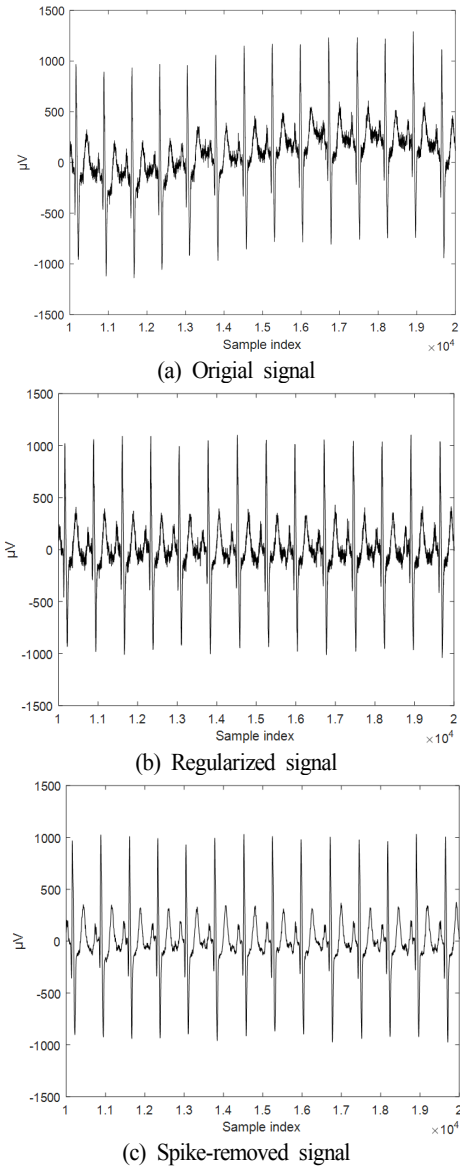


Figure 2. Pre-processing.

PTB데이터는 15개의 리드에서 기록되었으며, 다양한 연령과 다양한 질환자 및 건강한 남녀의 290명으로부터 획득되었고 매번의 기록은 6개월가량의 시간차를 가지고 측정되었다[15]. 사람마다 레코드의 개수와 길이가 달라서 4개이상의 레코드를 가진 사람들로 제한하였고 그때 사람 수는 최대 54명이었다. 사람마다 4개의 레코드 중에서 2개는 학습데이터를 추출하는데 사용하고, 나머지 2개는 검증데이터를 추출하는데 사용하였다. 각 레코드에서는 순차적으로 3초씩 분할하여 10개까지만 데이터를 추출하였다. 사람마다 4개의 레코드가 있고, 레코드마다 10개의 데이터를 추출하여

한 사람은 총 40개의 데이터를 갖게 된다. 이 중에서 20개는 학습으로 사용되고 나머지는 검증으로 사용된다. 전체 사람 수가 54명으로 사람당 20개의 데이터를 가지고 있어서 전체 학습과 전체 검증 데이터의 수는 각각 1080개이다. 개별 ECG 신호는 전처리과정을 통해 정규화가 수행되고 스파이크가 제거되고 R피크점들이 검출된다. 그림 2a는 원신호를 보여주고, 그림 2b는 정규화된 신호를 보여주고 그림 2c은 스파이크 제거된 신호를 보여준다. 그림 3은 검출된 R피크점을 보여준다.

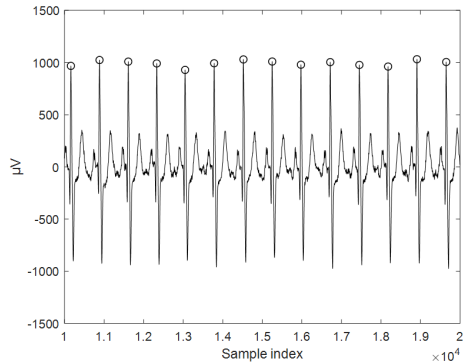


Figure 3. Detected R peaks.

한 데이터는 3초간의 측정값으로 구성되어있고, 이 데이터는 전처리를 거치게 되어 R피크점들의 위치를 알 수 있게 된다. 3초 동안 보통 2~5개의 R피크점이 존재하며 이들 R피크점들을 기준으로 앞뒤 400프레임을 모아서 평균을 취한 후, 다선형 주성분분석을 통해 개인식별을 수행한다. 유사도 측정 방법으로는 L1, L2, AD, MAH, MAD를 이용하였다. 한 리드에서 R 피크점을 기준으로 앞뒤 400프레임을 취하면 중앙까지 포함하여 801프레임을 얻는다. 리드 I와 리드 II에 대해서 2번 수행한 후에 연결하기 때문에 1602프레임이 얻어진다. 학습데이터 수가 1080개이기 때문에 학습용 데이터 크기는 1602x1080이 된다. 마찬가지로 검증용 데이터 크기도 1602x1080이다.

다선형 주성분 분석에 입력하기 위해 1차원의 벡터를 3차원의 텐서로 변경한다. 데이터 4개씩 묶어서 시퀀스 데이터를 생성하면, 한 개의 데이터 크기는 1602x4이 된다. 그리고 1차원 벡터의 맨처음과 맨끝의 1프레임을 버리고 재정렬(reshape)를 하면 텐서의 크기를 40x40x4으로 구성할 수 있다. 학습데이터 1080개에서 4개씩 묶어 3차 텐서를 구성했기 때문에 학습데이터의 수는 270개가 되어 텐서의 크기는 40x40x4x270이다. 마찬가지로 검증용 데이터 크기도 40x40x4x270이다. 심전도의 경우 측정자의 상태에 따라 신호가 민감하게 반응한다. 이럴 경우 다양한 상태에 대한 식별정보를 가지고 있어 민감하게 반응하는 심전도에 어느정도 대응이 가능하다. 그러나 지나치게 데이터를 다양하게 구성할 경우, 오히려 다른 클

래스의 심전도와의 거리를 더 가깝게 하는 경우의 수를 만들 가능성이 있어 인식률이 떨어질 수 있다. 그래서 어느정도 잡음을 허용하면서, 허용한 잡음이 다른 클래스와의 거리를 좁히지 않도록 적당한 데이터를 구성해야한다. 다선형 주성분분석에서 Q는 보존 에너지로서 일반적으로 사용되는 값으로 고정하였고, 반복수는 여러번 반복해도 정확도에 큰 변화가 없어서 1로 사용하였다. 표 1은 다선형 주성분 분석으로 텐서기반 개인식별한 정확도를 보여준다. 특징차원은 10부터 170까지 10간격으로 실험을 수행하였고, 유사도별로 가장 높은 인식률들을 보여준다. 그림 4는 특징차원을 10부터 170까지 10간격으로 실험을 수행하였을 때, 유사도별로 인식률을 보여준다. 그림 5는 심전도 신호를 다선형 주성분 분석으로 추출한 특징을 보여준다. 기존에는 심전도 신호에서 하나의 주기에 대해 1차원적 또는 2차원적 실험을 하였으나 본 논문은 연속된 여러 주기의 신호들을 하나하나 본래 정보를 유지한 채 다차원 배열로 구성하여 기존에 시도되지 않은 방법을 사용하였다. 또한 MPCA의 특성상 추정 파라미터 수가 적어서 메모리, 컴퓨팅 부하가 적게 드는 장점을 가진다.

Table 1. Maximum Accuracies of MPCA

특징추출 방법	특징추출		특징차원	유사도	정확도
	Q	반복			
MPCA	97	1	60	L1	91.85 %
			40	L2	90.00 %
			60	AD	90.37 %
			30	Mah	88.15 %
			60	MAD	91.48 %

### 5. 결 론

본 논문에서는 3차 텐서기반 MPCA 방법을 이용한 심전도신호의 개인식별에 대한 실험을 수행하였다. MPCA는 고차원 데이터를 효율적으로 특징추출을 하여 데이터의 추정파라미터 수가 매우 적어지는 효과를 가진다. 그래서 실제 고차 데이터의 거대한 크기의 데이터를 다룰 때 효율적이다. 심전도 신호는 PTB 데이터베이스를 사용하였고, 전처리와 R피크 검출을 통해서 정규화를 수행하였다. 그 다음 재정렬을 통해 1차원 벡터를 2차원 행렬로 만들고 그 행렬들을 겹쳐 3차원 텐서로 구성하였다. MPCA를 이용하여 특징을 추출하였고 5가지 유사도 계산을 통해 인식률을 살펴 보았다. L1 유사도 방법이 특징차원 60일 때 91.85%로 가장 좋은 정확도를 가졌다. 반면 MAH 유사도 방법은 특징 차원이 커질수록 인식률이 급격하게 감소하였다. 향후 PTB 데이터베이스 뿐만 아니라 직접 데이터를 더 다양하고 더 많은 사람들로 부터 획득

특하여 알고리즘에 대한 성능을 분석하고, 인식률 향상을 위해 연구할 계획이다.

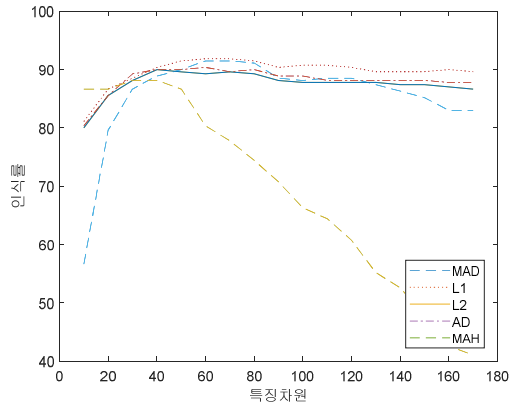


Figure 4. Accuracies of MPCA

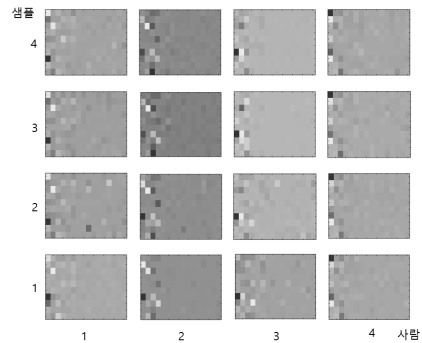


Figure 5. Feature maps of MPCA from electrocardiogram

### 감 사

이 논문은 조선대학교 2017학년도 학부생 새싹형 연구인재 지원프로그램(학부 장학금 사업)의 지원을 받아 연구되었음

### 참고문헌

1. Choi, K. H., Moon, H. M., Pan, S. B., Biometrics system technology trends based on biosignal, Journal of Digital Convergence, 15(1), 381-391, (2017).
2. Yoon, T. H., Human identification by electrocardiogram, Konkuk University Graduate School, 1-55, (2005).
3. Kim, J. J., Lee, S. M., Ryu, G. S., Lee, J. H., Park, K. H., "Hierarchical Authentication Algorithm Using Curvature Based Fiducial Point Extraction of ECG Signal, Journal of Korea Multimedia Society, 20(3), 465-473, (2017).

4. Kim, J. K., Lee, K. B., Hong, S. G., CG-based Biometric Authentication Using Random Forest, *Journal of The Institute of Electronics and Information Engineers*, Vol. 54-6, pp.100-105, (2017).
5. Lim, C. S., A Study on the Analysis of Technology and Service Issues for Wearable Devices and Future Development Direction, *Journal of Korean Institute of Next Generation Computing*, 13(4), 81-89, (2017).
6. Lu, H., Plataniotis, K. N., Venetsanopoulos, A. N., A survey of multilinear subspace learning for tensor data, *Pattern Recognition*, 44, 1540-1551, (2011).
7. Turk, M., Pentland, A., Face recognition using eigenfaces, in *Proc. IEEE Conference Computer Vision Pattern Recognition*, 586-591, (1991).
8. Murase, H. and Nayar, S. K., Visual learning and recognition of 3-d objects from appearance, *Journal of Computer Vision*, 14(1), 5-24, (1995).
9. Yang, J., Zhang, D., Frangi, A. F., Yang, J., Two-dimensional PCA: a new approach to appearance-based face representation and recognition, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 26(1), 131-137, (2004).
10. Kim, D. H., A Vision-based User Recognition in Robot Environments by Using Semi-biometrics, *Busan National University*, (2009).
11. Zhang, D., Zhou, Z. H., Chen, S., Diagonal principal component analysis for face recognition, *Pattern Recognition*, 39(1), 140-142, (2006).
12. Zhang, D. and Zhou, Z. H., (2D)2PCA: Two-directional two-dimensional PCA for efficient face representation and recognition, *Neurocomputing*, 69(1-3), 224-231, (2005).
13. Lu, H., Plataniotis, K. N., and Venetsanopoulos, A. N., MPCA: Multilinear principal component analysis of tensor objects, *IEEE Trans. on Neural Networks*, 19(1), 18-39, (2008).
14. Han, Y. H., Kwak, K. C., Face representation and recognition using third-order tensor-based MPCA method, *Jornal of Advanced Information Technology and Congvergence*, 9(6), 147-154, (2011).
15. Goldberger, A. L., Amaral, L. N., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., Mietus, J. E., Moody, G. B., Peng, C. K., Stanley, H. E., PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals, *Circulation*, 101(23), e215-e220, (2000).