

주성분 분석기법을 이용한 심전도 기반 개인인증

ECG based Personal Authentication using Principal Component Analysis

조 주 희* · 조 병 준* · 이 대 중* · 전 명 근[†]

(Ju-Hee Cho · Byeong-Jun Cho · Dae-Jong Lee · Myung-Geun Chun)

Abstract - The PCA(Principal Component Analysis) algorithm is widely used as a technique of expressing the eigenvectors of the covariance matrix that best represents the characteristics of the data and reducing the high dimensional vector to a low dimensional vector. In this paper, we have developed a personal authentication method based on ECG using principal component analysis. The proposed method showed excellent recognition performance of 98.2 [%] when it was experimented using electrocardiogram data obtained at weekly intervals. Therefore, it can be seen that it is useful for personal authentication by reducing the dimension without changing the information on the variability and the correlation set variable existing in the electrocardiogram data by using the principal component analysis technique.

Key Words : ECG, PCA, Personal authentication, k-NN

1. 서 론

얼굴, 홍채 등의 생체정보는 기존의 패스워드에 의한 방식과 달리 사람마다 고유하고, 변하지 않는 특성으로 인하여 고도의 개인인증 시스템을 구현하는데 널리 사용된다. 최근에는 생체정보 취득과 개인 인증 알고리즘의 성능이 우수하여 다양한 분야에 상용화가 이루어졌다. 그러나 생체정보를 이용한 개인인증 시스템은 생체를 감지할 수 없다는 단점을 지니고 있다. 즉, 생체정보를 이용한 개인 인증 시스템이 사람을 정확하게 식별한다고 해도 시스템은 사람이 실제로 있는지를 보장 할 수 없다. 최근 아이폰과 갤럭시폰이 목공용 접촉제와 경화성 실리콘으로 1시간 만에 보안이 해제된 사례와 2014년 독일 해커단체 ccc가 고화질 사진과 3D 프린터를 이용해 푸틴 러시아 대통령의 홍채를 복제해 공개한 사례를 보아 생체정보는 위·변조에 취약하다는 것을 알 수 있다.

최근에 가짜지문 등 기존의 신체적 특징을 이용한 바이오 인식기술의 위변조 위협에 대한 우려가 존재함에 따라 뇌파·심전도·근전도·맥박 등 살아있는 사람의 생체신호를 이용한 개인인증 기술 개발이 연구되고 있다. 심장의 전기적 활동을 그래프로 보여주는 심전도(ECG:Electrocardiogram) 신호는 심장 질환의 모니터링 및 진단에 널리 사용되고 있다. 그 뿐만 아니라, 심전도 신호는 개인 독특한 특징을 가지고 있어 심전도 신호를 이용하여 개인인증이 가능한 장점이 있다[1]. ECG 기반 생체 인식 인증 시스템의 가장 중요

한 장점은 생체 감지기로 활용할 수 있다는 것이다. 즉, ECG 기반 생체 인증 시스템은 식별 대상자가 물리적으로 존재한다는 것을 보장한다[2].

반면에, ECG 기반 생체 인증 시스템은 ECG 신호의 데이터 획득 프로세스가 매우 비용이 들고 생체 인증 시스템에 비실용적이라는 중요한 단점을 가지고 있다. ECG는 기본적으로 심장 질환의 탐지에 사용되었지만 지난 몇 년 동안 연구자들은 ECG를 개인 식별에 사용하는데 주력했다. 2001년 생체인식을 위한 ECG 사용은 Biel 등에 의해 처음 제안되었으며, 12리드 ECG 기록을 사용하여 심전도 신호에 기초한 개인 식별이 가능함을 제시하였다[3]. 이후에는 Shen과 Tompkins에 의해 3개의 전극(2개의 능동 전극과 접지 전극)으로 구성된 단일 리드 ECG를 사용하여 개인 식별이 가능함을 제시하였다[4]. Kyoso와 Uchiyama는 P파의 지속시간, PQ 간격, QRS 간격 및 QT 간격의 4 가지 매개 변수가 사람을 식별하기에 충분함을 제시하였다[5].

개인마다 ECG 신호의 특징을 추출하기 위하여 ECG 신호의 시간 간격과 진폭을 정확하게 계산해야 하는 단점이 있다. 이러한 문제를 해결하기 위해 DCT[6]와 이산 웨이블릿 변환[7]을 이용하는 방법이 제안되었다. 2010년에는 시간 정보와 스펙트럼 정보를 기반으로 한 ECG기반의 생체인식 알고리즘이 제안되었다[8]. 2011년에는 ECG 생체 인식을 위한 펄스 활성비율로 알려진 새로운 ECG 특징 추출 알고리즘이 소개되었다[9].

본 논문에서는 주성분 분석기법을 이용한 심전도 기반 개인인증 방법을 제안한다. 주성분분석기법은 데이터에 존재하는 가변성과 상관집합변수의 정보를 유지하면서 몇 가지의 변수로 축소하는 통계기법으로 다양한 분야에서 성공적으로 적용되고 있다. 제안된 방법은 일주일 간격으로 취득한 ECG 데이터를 이용하여 타당성을 검증하였다.

[†] Corresponding Author : School of Electronics Engineering, Chungbuk National University, Korea

E-mail : mgchun@cbnu.ac.kr

* School of Electronics Engineering, Chungbuk National University, Korea

접수일자 : 2017년 11월 6일

최종완료 : 2017년 11월 24일

2. 주성분 분석기법을 이용한 개인인증 알고리즘

2.1 알고리즘 구성도

그림 1에서는 본 논문에서 제안한 ECG 신호를 이용한 개인인증 알고리즘의 구성도를 나타냈다. 그림 1에서 보는 바와 같이 ECG 신호를 DAQ보드를 이용하여 취득한 후, 취득된 데이터의 고주파 제거를 위해 필터 처리하는 전처리 과정을 수행한다. 필터 처리된 데이터에 대해 주성분 분석기법을 이용하여 ECG 신호에 대한 특징을 추출하고, 추출된 PCA 계수와 k-NN 분류기를 이용하여 개인인증이 이루어진다.

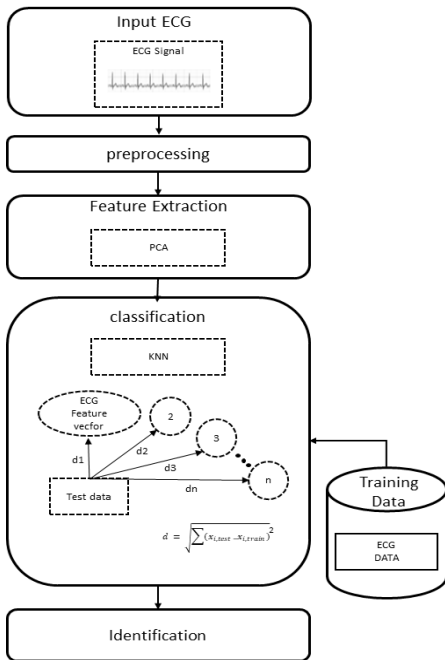


그림 1 제안된 알고리즘 구성도
Fig. 1 Configuration of a proposed algorithm

2.2 주성분 분석기법을 이용한 ECG 신호의 특징추출

주성분분석(PCA) 알고리즘은 데이터의 특징을 가장 잘 표현하는 공분산 행렬의 고유벡터를 구하여 고차원의 벡터를 줄여서 저차원의 벡터로 표현하는 기법으로 널리 사용되고 있다. 시계열의 ECG 생체신호는 공분산 행렬에 대한 고유벡터를 구하고, 고유값 크기 순으로 나열하여 저차원의 특징벡터로 변환을 하여 처리하게 된다. 이에 대한 자세한 알고리즘을 알아보면 다음과 같다.

[단계 1] 식 (1)에 나타난 바와 같이 측정된 p 개의 학습용 ECG 생체 신호에 대한 평균을 구한다.

$$m = \frac{1}{P} \sum_{i=1}^P x_i, \quad x_i \in R^{1 \times n} \quad (1)$$

[단계 2] 식 (2)에 나타난 바와 같이 학습용 ECG 생체 신호에 대한 공분산행렬을 구한다.

$$\Omega = \overline{XX^T} \quad (2)$$

$$\overline{X} = [\overline{x}_1, \overline{x}_2, \dots, \overline{x}_p], \quad \overline{x}_i = x_i - m$$

[단계 3] 식 (2)에 의해 계산된 공분산에 대한 n 개의 고유값과 n 개의 고유벡터를 구한다.

$$\Omega v_i = \lambda_i v_i \quad (3)$$

[단계 4] 고유벡터 v_j 를 고려하고자 하는 차원수(k) 만큼 고유값의 크기 순으로 재정렬한 고유벡터 $V \in R^{n \times k}$ 를 구한 후에 학습 생체 신호 x_i 에 대한 특징벡터 $z_i \in R^{k \times 1}$ 를 식 (4)에 의해 산출한다. 식 (4)에서 알 수 있는 바와 같이 입력 생체 신호에 대한 특징벡터는 공분산에 대한 고유벡터를 이용하여 PCA변환된 공간으로 투영함으로써 얻을 수 있다.

$$z_i = V^T(x_i - m)^T, \quad i = 1, 2, \dots, P \quad (4)$$

[단계 5] PCA기법에 의해 i 번째 검증용 생체신호 x'_i 에 대한 특징벡터 z'_i 를 산출한다.

$$z'_i = V^T(x'_i - m)^T \quad (5)$$

2.3 k-NN 분류기를 이용한 개인 인증

k -NN(k -Nearest Neighbor) 분류방법은 간단성과 직관성 때문에 여러 분야에서 널리 이용되고 있는 예제 기반 학습 방법이다. k -NN과 같이 예제 기반 학습 방법에서는 기본적으로 벡터공간에 학습데이터를 구축한 후, 입력된 데이터에 가장 근접한 k 개의 추출하여 입력 데이터를 분류한다. 일반적으로, 훈련 데이터와 검증 데이터에 대한 특징을 추출한 후 미리 저장해 놓은 기준 패턴과 입력패턴과의 거리를 계산하여 k 개의 가까운 기준패턴의 클래스를 선택하여 미지 패턴의 클래스로 결정하는 분류방법이다. 즉, k -NN 분류규칙은 패턴 x 의 k -최근접 이웃의 각 요소가 어느 클래스에 가장 많이 속하는가를 조사하여 가장 많이 속하는 수의 클래스를 x 의 클래스로 결정하는 방법이다.

k -NN에서 가장 중요한 문제는 k 개를 설정과 훈련데이터와 검증 데이터간의 근접한 거리를 어떻게 결정할 것인가 하는 문제이다. 보통 k 를 결정하는 것은 실험 또는 문제를 잘 알고 있는 전문가를 통하여 휴리스틱하게 결정한다. 또한 기준패턴과 입력패턴과의 거리를 비교하기 위하여 다음 식과 같이 유클리디안 거리기반에 의한 방법을 주로 사용한다.

$$d_2(x) = \sqrt{\sum_i |x_i|^2} \quad (6)$$

그림 2에서는 k -NN($k=5$) 분류규칙에 의한 패턴분류 결과를 나타냈다. 그림 2에서 알 수 있는 바와 같이 미지의 입력 데이터와 미리 저장된 기준패턴과의 거리 비교를 통하여

거리가 가까운 세 개의 데이터를 기준패턴으로부터 결정한다. 즉, 5개의 가장 근접한 이웃들 중에, 4개은 ω_1 에 속해있고, 1개는 ω_3 에 속해있으므로, 최종 분류단계에서는 선택된 기준패턴 중 가장 많은 클래스 값을 갖는 클래스 ω_1 를 선택한다.

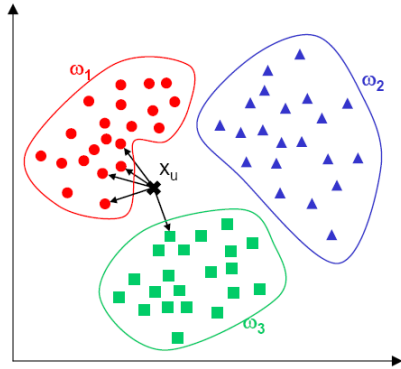


그림 2 k-NN 분류기에 의한 분류
Fig. 2 Classification example by k-NN

클래스의 집합을 $\{\omega_1, \omega_2, \dots, \omega_n\}$ 이라고 할 때 k-NN 분류규칙을 단계별로 설명하면 다음과 같다.

[단계 1] 먼저 주어진 데이터의 k개의 가장 가까운 패턴 $X = \{x_1, x_2, \dots, x_k\}$ 을 기준패턴으로부터 찾는다.

[단계 2] k-최근접 이웃 가운데 voting 방법에 의해 주어진 데이터의 클래스를 결정한다.

$$N(i) = \arg(X \in \omega_i), \quad i = 1, 2, \dots, n, \quad (7)$$

$$i_o = \operatorname{argmax}_i(N(i))$$

3. 실험 결과

3.1 데이터 취득

그림 3에서는 ECG 신호를 취득하기 위한 실험장치를 나타냈다. 그림 3에서 보는 바와 같이 (+)전극을 왼팔에, (-)전극을 오른팔에, 그리고 접지전극을 오른다리에 부착한다. 이와 같은 전극유도법을 표준사지유도 중에서 Lead-I라 한다. 전극 부착 시 팔에 부착하는 위치는 팔꿈치와 손목 사이의 근육부위가 적당하다. 기록 중에는 안정된 상태로 팔을 편안히 내려놓아야 근전도에 의한 잡음을 방지할 수 있다.

측정 장비는 Physiolab의 iDAQ-400(AD Converter & PC interface Module)과 ECG-Amp를 사용한다. iDAQ-400은 생체신호를 디지털화하고 PC로 모니터링하기 위한 베이스 모듈이다. 최대 4채널의 생체신호를 간편하게 모니터링할 수 있으며 최대 2kHz의 속도로 샘플링 할 수 있다. ECG-Amp는 심장의 전기적 신호를 증폭할 수 있는 모듈이다. 심전도 파형 이외에 심전도 분석에 많이 활용되어지는 Beat, BPM도 선택적으로 측정할 수 있다. 수 mV 수준의 미약한 심전도 신호를 증폭하고, 혼입된 잡음을 필터링하

며, 디지털 신호로 변환하기 위해 iDAQ-400으로 전달한다.

그림 4에서는 20초 동안에 500[Hz] 취득한 ECG 신호의 파형을 나타냈다. 그림 4에서 보는 바와 같이 여러 주기의 ECG 신호가 연속적으로 취득됨을 알 수 있다. ECG 생체신호를 이용하여 개인인증을 하기 위해서는 한 주기의 ECG 생체신호를 검출해야 한다. 본 논문에서는 그림 5에서 보는 바와 같이 ECG 생체 신호의 최대값을 의미하는 R파를 기준으로 좌우 200개의 샘플을 취득하여 사람당 총 401개의 샘플로 구성된 ECG 한주기 생체신호를 검출하였다.

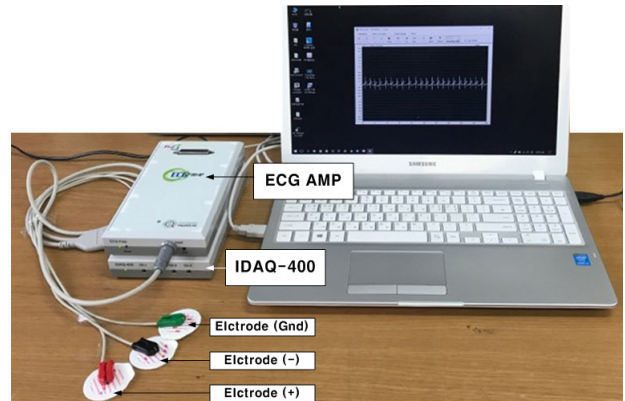


그림 3 실험 장치 구성
Fig. 3 Configuration of experimental apparatus

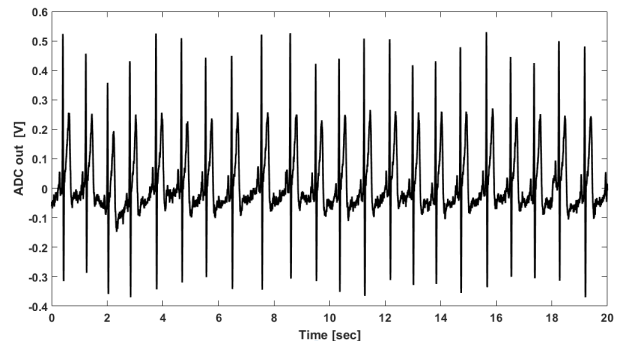


그림 4 ECG 생체 신호
Fig. 4 ECG bio-signal

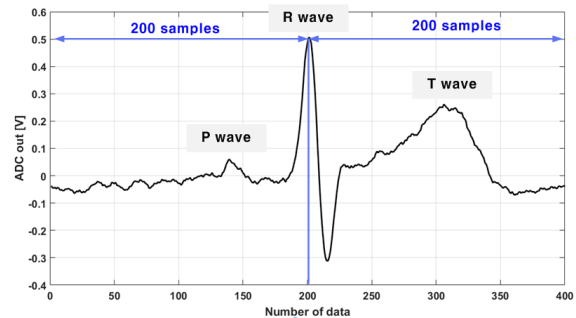


그림 5 R파를 기준으로 취득한 ECG 생체신호
Fig. 5 The ECG bio-signal acquired based on the R wave

3.2 실험결과

데이터 취득은 20대 남성 20명을 대상으로 ECG 신호를 취득했다. 일주일씩 간격을 두고 3번에 나누어 샘플링 주파수 500Hz로 40초간 측정했다. 첫째 날 3번, 둘째 날 3번, 셋째 날 4번으로 한명 당 총 10번을 취득했다. 각 1번의 신호에서 R피크 특징을 이용해 10개의 주기를 취득하여 한명 당 총 100개(10회×10samples/회)로 구성된 ECG 데이터를 구축하였다. 100회의 데이터 중에서 홀수 번에 해당하는 데이터를 학습데이터로 사용하였고, 짝수 번에 해당하는 데이터를 검증데이터로 사용하여 제안된 방법의 성능을 평가하였다.

표 1에서는 k 값(1, 3, 5, 7, 9)에 따른 인식률을 나타냈다. 표 1에서 보는 바와 같이 k가 1일 때 (즉, 가장 최솟값만 선택하는 경우) 인식률이 98.2%로 가장 높게 나타났다. 그림 6에서는 k가 1로 설정하고, 주성분 분석 특징의 수를 변경하면서 실험한 결과를 나타냈다. 그림 6에서 보는 바와 같이 특징벡터의 수를 10개로 하였을 때 인식률이 98.2[%]를 보임을 알 수 있다.

표 2에서는 개인별 인식결과를 나타냈다. 표 2에서 보는 바와 같이 가장 인식률이 낮은 사람은 P20번으로 47개의 생체신호는 본인으로 인식하였으나 3개의 생체신호는 P1번으로 오인식 되었다. 이와 같이 P20번이 P1번으로 인식되는 이유를 분석하여 위하여 P20번에 해당하는 생체신호와 P1번에 해당하는 생체신호를 그림 7에 나타냈다. 그림 7에서 보는 바와 같이 P20번의 ECG 파형과 P1번의 ECG 파형이 유사하여 오인식 된 것으로 분석된다.

표 1 k값에 따른 인식률

Table 1 Recognition rate according to number of k

k=1	k=3	k=5	k=7	k=9
98.2%	97.4%	97.1%	97.0%	96.9%

표 2 개인별 인식 결과

Table 2 Individual recognition result

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20
P1	50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P2	0	50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P3	0	0	48	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0
P4	0	0	0	48	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P5	0	0	0	0	50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P6	0	0	0	0	0	50	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P7	0	0	0	0	0	0	50	0	0	0	0	0	0	0	0	0	0	0	0	0
P8	0	0	0	0	0	0	0	50	0	0	0	0	0	0	0	0	0	0	0	0
P9	0	1	0	0	0	0	0	0	48	0	0	0	0	0	0	1	0	0	0	0
P10	0	0	0	0	0	0	0	0	0	50	0	0	0	0	0	0	0	0	0	0
P11	0	0	0	0	0	0	0	0	0	0	49	0	0	1	0	0	0	0	0	0
P12	0	0	0	0	0	0	0	0	0	0	0	50	0	0	0	0	0	0	0	0
P13	0	0	0	0	0	0	0	0	0	0	0	0	50	0	0	0	0	0	0	0
P14	0	0	0	1	0	0	0	0	0	0	2	0	0	47	0	0	0	0	0	0
P15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	49	0	0	0	1	0
P16	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	49	0	0	0	0
P17	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	48	0	0	0
P18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	50	0	0
P19	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	49	0
P20	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	47

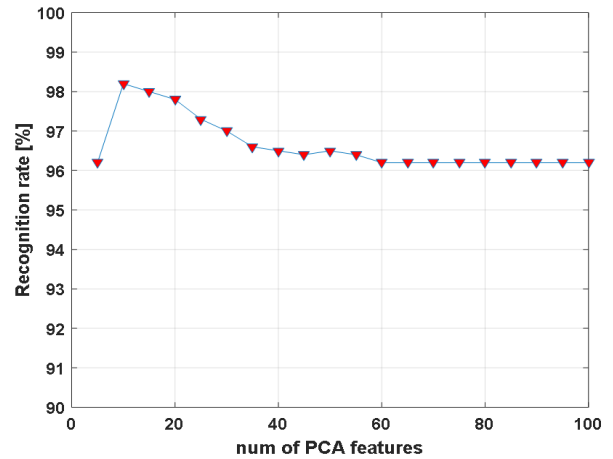


그림 6 PCA 특징수에 따른 인식결과 (k=1)

Fig. 6 Recognition rate according to number of PCA feature(k=1)

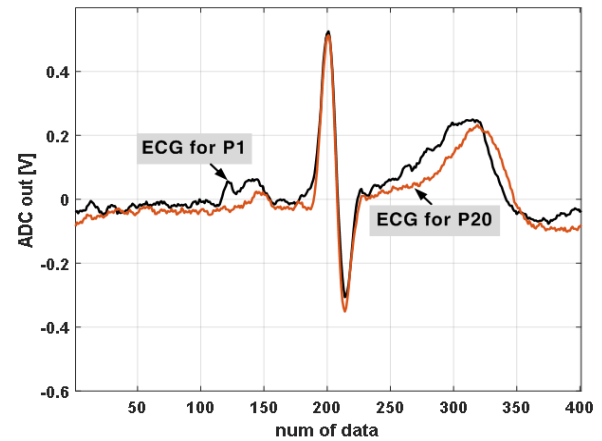


그림 7 P1과 P20에 대한 ECG 파형 비교

Fig. 7 ECG waveform comparison for P1 and P20

4. 결 론

주성분분석 알고리즘은 공분산 행렬의 고유벡터를 구하여 고차원의 벡터를 줄여서 저차원의 벡터로 표현하는 기법으로 널리 사용되고 있다. 본 논문에서는 주성분 분석기법을 이용한 심전도 기반 개인인증 방법을 개발하였다. 제안된 방법은 일주일 간격으로 취득한 심전도 데이터를 이용하여 실험한 결과 98.2[%]의 우수한 인식성능을 나타냈다. 따라서 주성분 분석기법을 이용하여 심전도 데이터에 존재하는 가변성과 상관집합변수의 정보를 변화시키지 않으면서 차원축소를 함으로써 개인인증에 유용함을 알 수 있다.

감사의 글

본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 대학 ICT연구센터육성지원사업의 연구결과로 수행되었음(IITP-2017-2013-0-00881).

References

- [1] A. Kaul, A. S. Arora and S. Chauhan, "ECG based human authentication using synthetic ECG template," *In Signal Processing, Computing and Control (ISPCC), 2012 IEEE International Conference on*, pp. 1-4, 2012.
- [2] Fang, Shih-Chin, and Hsiao-Lung Chan, "Human identification by quantifying similarity and dissimilarity in electrocardiogram phase space," *Pattern Recognition*, vol. 42, no. 9, pp. 1824-1831, 2009.
- [3] Biel, Lena, et al., "ECG analysis: a new approach in human identification," *IEEE Transactions on Instrumentation and Measurement*, vol. 50, no. 3, pp. 808-812, 2001.
- [4] T. W. Shen, W. J. Tompkins, Y. H. Hu, "One-lead ECG for Identity Verification," in *Proceedings of the 2nd Joint Engineering in Medicine and Biology, 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society (EMBS/BMES)*, vol. 1, pp. 62-63, 2002.
- [5] M. Kyoso, and A. Uchiyama, "Development of an ECG identification system," in *Proceedings of the 23rd Annual International Conference of the IEEE on Engineering in Medicine and Biology Society*, vol. 4, pp. 3721-3723, 2001.
- [6] Plataniotis, Konstantinos N., Dimitrios Hatzinakos, and Jimmy KM Lee, "ECG biometric recognition without fiducial detection," *Biometric Consortium Conference, 2006 Biometrics Symposium: Special Session on Research at the IEEE*, pp. 1-6, 2006.
- [7] Chiu, Chuang-Chien, Chou-Min Chuang, and Chih-Yu Hsu, "A novel personal identity verification approach using a discrete wavelet transform of the ECG signal," *Multimedia and Ubiquitous Engineering*,

2008. *MUE 2008. International Conference on IEEE*, pp. 201-206, 2008.

- [8] Li, Ming, and Shrikanth Narayanan, "Robust ECG biometrics by fusing temporal and cepstral information," *Pattern recognition (icpr), 2010 20th international conference on, IEEE*, pp. 1326-2329, 2010.
- [9] S. I. Safie, J. J. Soraghan, and L. Petropoulakis, "Electrocardiogram biometric authentication using pulse active ratio (PAR)," *IEEE Trans. on Inf. Forensics and Security*, vol. 6, no. 4, pp. 1315-1322, 2011.

저 자 소 개



조 주 희(Ju-Hee Cho)

2016년 충북대학교 정보통신공학과 졸업.
2017년~현재 충북대학교 대학원 정보보호경영학과 석사과정.
E-mail : bbqds123@naver.com



조 병 준(Byeong-Jun Cho)

2017년 충북대학교 전자공학과 졸업.
2017년~현재 충북대학교 대학원 제어로봇공학과 석사과정.
E-mail : jjgk91@cbnu.ac.kr



이 대 종(Dae-Jong Lee)

1995년 충북대학교 전기공학과 졸업.
1997년 동 대학원 졸업(공학석사).
2002년 동 대학원 졸업(공학박사).
2006년~2008년 충북대학교 BK21충북정보기술 사업단 초빙교수.
E-mail : leebigbell@gmail.com



전 명 근(Myung-Geun Chun)

1987년 부산대학교 전자공학과 졸업.
1989년 한국과학기술원 전기및전자공학과 대학원 졸업(공학석사).
1993년 한국과학기술원 전기및전자공학과 대학원 졸업(공학박사).
1993년~1996년 삼성전자 자동차 연구소

전임연구원.

2000년~2001년 University of Alberta 방문 교수.

2010년~2011년 Temple University 방문 교수.

1996년~현재 충북대학교 전자공학부 교수.

2008년~현재 TTA PG505 표준위원회 부의장.

E-mail : mgchun@cbnu.ac.kr