

## 스마트 장치에서 비접촉식 전위계차 센서 신호를 이용한 동작 인식 기법

(Gestures Recognition for Smart Device using Contactless Electronic Potential Sensor)

오강한\*, 김수형\*\*, 나인섭, 김영철, 문창협

(KangHan Oh, Soohyung Kim, Inseop Na, Young Chul Kim, Changhub Moon )

### 요약

본 논문에서는 스마트 TV, 스마트폰으로 대표되는 스마트 장치에서 비접촉식 전위계차 센서(CEPS)로부터 추출된 동작 신호를 k-NN 과 DTW 알고리즘을 이용하여 인식하는 방법을 제안한다. 먼저 CEPS 신호는 칼만 필터를 이용해서 잡음을 제거해주고 정규화를 시켜준다. 다음 인식 속도를 항상 시키고 분류에 방해되는 성분제거 하기 위해 PCA 알고리즘을 사용해서 신호의 차원을 축소시켰다. 그리고 k-NN 과 DTW 알고리즘을 사용하여 인식 작업을 수행하였다. 실험 결과에서는 앞서 언급된 2개의 스마트 장치 환경 설정에 대해서 설명하고 각각의 환경에서 추출된 신호를 제안된 알고리즘에 의해서 인식을 하였다. 기존 인식 알고리즘의 결합과 분해를 통해 다양한 결과를 비교 분석함하고 90% 이상의 인식률을 달성함으로써 제안된 방법의 우수성을 증명하였다.

■ 중심어 : 동작인식; CEPS; PCA ; DTW; k-NN;

### Abstract

This paper presents a novel approach to recognize human gestures using k-NN and DTW based on Contactless Electronic Potential Sensor(CEPS) in the smart devices such as smart TV and smart-phone in the proposed method, we used a Kalman filter to remove noise on gesture signal from CEPS and a PCA algorithm is utilized for reducing the dimensionality of gesture signal without data losses. And then in order to categorize gesture signals, k-NN classifier with DTW distance measure is considered. In the experimental result, we evaluate recognition performance with CEPS gesutres signal form the above two types of smart devices, and we can successfully identify five different gestures with more than 90% of recognition accuracy.

■ keywords : Gesture Recognition; CEPS; PCA ; DTW; k-NN;

### I. 서 론

명령어 줄 인터페이스(Command Line Interface(CLI)), 그래픽 사용자 인터페이스(Graphic User Interface(GUI)) 등 사용자 인터페이스(User Interface(UI)) 기술은 날로 발전해왔다. 최근 조명을 받고 있는 감성 유저 인터페이스(Natural User Interface(NUI))는 앞서 언급한 CLI, GUI 보다 더 사람들이 손쉽게 사용할 수 있는 직관적인 UI이다. NUI 기술들을 좀 더 살펴보면 키보드 와 마우스 같은 입력 장치 없이 사람들의 동작, 음성, 사람의 상태 등을 센서가 분석하여 이를 인식하

는 공통점이 있다. NUI 연구는 다양한 센서를 사용하기 때문에 센서 종류에 따라서 다양한 방법론 존재하고 그 중 카메라로부터 영상정보를 입력받아 이것을 분석하는 연구와 마이크로부터 음성 신호를 입력받아 이것을 인식하는 연구는 지금까지 많이 진행 되어져오고 있었고 많은 성과가 있었다[1-3]. 본 연구의 기술개발 목적은 차세대 스마트 TV에 장착된 비접촉식 전위 계차 센서 (Contactless Electronic Potential Sensor (CEPS)) 신호를 이용하여 새로운 NUI를 구현하는 것이 목적이다. 주로 스마트 TV의 NUI 환경은 카메라로 사용자의 동작을 인식한다. 이러한 연구는 많이 진행되어져 오고 있고 실제로 상용화도 되어있지만 조명이 민감한 환경에서 문제가 발생

\* 학생회원, 전남대학교 전자컴퓨터공학과

\*\* 종신회원, 전남대학교 전자컴퓨터공학부

이 논문은 2013년도 정부(교육과학기술부)의 재원으로 한국연구재단의 기초연구사업 지원을 받아 수행된 것임(2013-022495).

접수일자 : 2014년 3월 4일

수정일자 : 1차 2014년 06월 20일

제재확정일 : 2014년 06월 20일

교신저자 : 김수형, e-mail : shkim@jnu.ac.kr

할 수 있고 CEPS 센서와 비교하여 매우 비싼 가격도 단점으로 작용한다. 최근 영국 Plessey Semiconductors사는 EPIC 센서라고 불리는 비접촉식 전위계차 바이오센서를 발표 했다. 비접촉식 전위계 센서는 매우 민감한 디지털 전압계 형태로써 신경과 근육 주변의 전기장의 작은 변화도 측정 할 수 있다. 특히 움직임을 추적하는 데에 전문화된 이번 버전은 안정성, 스위칭 기술, 게임 어플리케이션에 최적화된 형태로 개발되었다. 또한, 심장박동 이외에도 폐, 암구 움직임들을 추적 할 수 있다[4]. 제안된 방법은 CEPS 센서중 하나인 EPIC 센서를 사용하였다. 실험을 위해서 스마트 TV 와 스마트폰 으로 대표되는 2가지 장치 에서 환경을 설정 하고 신호를 추출 한 후 제안된 방법을 사용해서 다양한 인식 결과를 도출 하였다. 논문의 구성은 2장 제안된 방법에서는 “스마트 장치 셋팅”, “신호 전처리”, “특징 추출”, “인식 기법” 등으로 구성되어져있으며, 3장에서는 제안된 방법의 우수성을 증명할 수 있는 다양한 실험결과를 기술 하였고, 4장 결론에서는 향후 집중적으로 개선 해야할 방향에 대해서 설명한다.

## II. 제안된 방법

### 1. 스마트 장치

제안된 방법은 스마트 TV , 스마트폰 2가지 환경을 만들어서 동작 신호를 추출하고 인식 실험을 진행 하였다. 먼저 스마트 TV 환경 셋팅은 그림 1과 같다. 스마트 TV 모서리에 센서를 장착하고 사용자는 스마트 TV 로부터 2 m 떨어져서 5가지 동작을 실행한다. 4개의 센서로부터 추출되는 4개의 신호는 A-B , C-D 신호로 바꾸어서 입력 신호로 사용하였다. 이렇게 추출된 동작 신호는 제안된 방법을 통해서 인식되고 궁극적으로 스마트 TV 를 제어하게 된다.

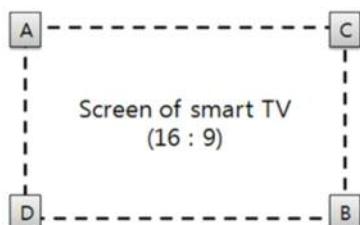


그림 1. 스마트 TV 기반 신호 추출

2번째 환경은 그림 2와 같이 스마트폰 에 센서를 장착해서 신호를 추출하고 동작을 인식한다. 그리고 인식된 결과는 블루투스를 사용하여 스마트 TV로 인식결과를 제공하고 TV를 제어 하게 된다.

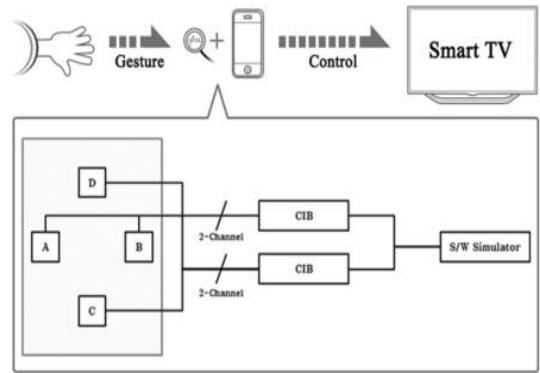


그림 2. 스마트폰 기반 신호 추출  
센서의 배치는 삼자 모양이며 4개의 센서로부터 추출되는 4개의 신호는 A-B , C-D 신호로 바꾸어서 입력 신호로 사용한다.

### 2. 전처리 시스템 설계

이번 장에서는 조정재 등[10] 제안한 전처리 시스템 설계에 관해서 설명한다. 인식 성능 향상을 위해 입력된 신호는 최대 폭의 크기를 2로 만드는 데이터 정규화를 통해 -1부터 +1까지의 절대적 범위 내에서 출력하도록 한다.

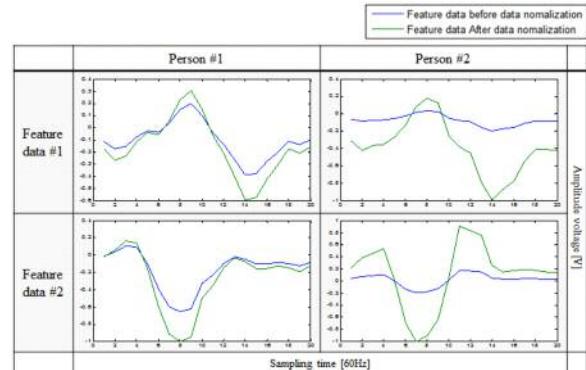


그림 3. 특정 제스처에 대한 데이터 정규화 전후 비교

데이터 정규화 과정 이후 선형보간법을 이용한 데이터 리샘플링을 통해 데이터의 크기와 정밀도를 바꾸어 처리 속도를 조절한다. 리샘플링은 데이터의 크기를 줄여 처리 속도를 향상시키는데 적용되며, 센서의 신뢰도가 낮을 때 여러 번 읽은 데이터를 한 개의 데이터로 다시 구성하는 용도로 사용되기도 한다. 리샘플링은 일정 구간의 입력 데이터를 사용해 새로운 데이터를 발생하거나 기존 데이터의 길이를 줄여서 만든다[9].

### 3. PCA를 이용한 특징 추출

PCA [6] 알고리즘은 고차원 입력 벡터를 저차원의 벡터로 표현하여 몇 개의 주성분 값으로 정보 손실을 최소화 하면서 차원을 축소하는 알고리즘이다. 더불어 PCA를 통해 삭제

되는 벡터들은 인식에 중요하지 않는 성분일 가능성이 높기 때문에 이 장에서는 최대한 벡터의 차원을 감소시키면서 가장 높은 성능을 지니는 최적의 축소된 벡터 개수를 연구한다. 먼저 (A-B) 와 (C-D) 두 신호를 연결 시켜 하나의 신호로 만들 어준다. 다음 신호들로부터 공분산을 계산하고 고유치 분해를 통해서 고유벡터와 고유 값을 추출한다. 마지막으로 축소하고 싶은 차원 개수만큼의 높은 순서대로 고유 값을 저장하고 이에 대응하는 고유 벡터를 사용하여 입력 신호를 저차원으로 사영 시킨다. 표 1은 차원 축소에 따른 성능 변화를 보여 준다. 여기서 -90%는 총 데이터 개수를 축소시킨 것을 의미 하며 0%는 PCA를 적용시키지 않았을 때 결과이다. 그리고 성능은 최대 성능을 기준으로 정규화 시켰다. 실험결과에 따라서 정규화된 데이터는 PCA를 사용하여 10% 만큼의 데이터를 축소시키고 사용한다.

표 1. PCA를 이용한 차원축소에 따른 성능 비교

차원축소	-90%	-70%	-50%	-30%	-10%	0%
성 능	0.69	0.88	0.91	0.94	1	0.83

#### 4. 인식 기법

이번 장에서는 제안된 방법뿐만 아니라 부분적으로 사용된 알고리즘에 대해서 간략히 설명한다. 먼저 비교를 위해서 사용된 “베이지안 분류기” 그리고 제안된 알고리즘에서 사용되는 DTW, k-NN 그리고 제안된 방법인 DTW 거리척도 기반 k-NN 알고리즘 구현에 대해서 기술한다.

##### 가. 베이지안 분류기

베이지안 분류기는 추정된 확률밀도를 이용해서 입력된 패턴을 분류하는 알고리즘이다. 간단히 말해서 훈련 데이터들로부터 각각의 부류의 확률밀도 파라미터를 (공분산행렬  $\Sigma$ , 평균  $\mu$ ) 계산하고 부류의 개수 N 개만큼의 분류기를 생성한다. 이때 각각의 분류기에는 앞서 계산된 확률밀도 파라미터들을 대입 시켜준다. 이 계산식에 입력 패턴  $x$ 를 대입하고 가장 높은 확률이 나오는 부류로 분류하는 원리이다. 이 과정은 수식 (1)(2)에서 표현하고 있다.

$$\text{class} = \arg\max\{p(x|w_i)p(w_i)\} \quad i=1..N \quad (1)$$

이때  $p(w_i)$  부류  $w_i$  발생하는 사전확률이고,  $p(x|w_i)$  우도 함수이다. 수식 (2)는 우도를 표현한다.

$$p(x|w_i) = \frac{1}{(2\pi)^{d/2} |\Sigma_i|^{1/2}} \exp(-\frac{1}{2}(x - \mu_i)^T \Sigma_i^{-1} (x - \mu_i)) \quad (2)$$

수식 (2)를 보면 d는 훈련 벡터의 차원을 표현하고  $\Sigma$ 는  $d \times d$  크기의 공분산 행렬을 의미한다. 그리고  $\mu$ 는 평균 값을 나타낸다. 즉 우도함수가 각각의 분류기를 의미하며 여기서 쓰인  $\Sigma, \mu$  가 훈련데이터라고 볼 수 있다.

#### 나. Data Time Warping(DTW)

DTW [7] 알고리즘은 주로 시계열 신호인 음성 데이터 인식 분야에서 자주 사용되는 알고리즘이다. 그 이유는 음성 신호는 사람에 따라 같은 음성이라도 빠르고 느림이 존재하기 때문에 유클리디언 거리와 같은 거리분류기로 분류할 경우 오차가 많이 발생하기 때문이다. 동작 신호 역시 음성 신호와 유사한 특징을 볼 수 있다. DTW 알고리즘의 2개의 시계열 입력벡터의 유사성을 측정하는 방법은 수식 (3)로 설명할 수 있다.

$$D(i, j) = d(x_i, y_j) + \min \begin{cases} D(i, j-1) \\ D(i-1, j) \\ D(i-1, j-1) \end{cases} \quad (3)$$

where  $X = x_1, \dots, x_n$  and  $Y = y_1, \dots, y_m$

$$d(x_i, y_j) = (x_i - y_j)^2$$

먼저 그림 4와 같이 두 개의 시계열 신호를 이용하여 DTW Matrix를 만든다[7]. 그 다음 수식 (2)를 이용해서 오른쪽 끝 영역부터 역으로 탐색을 시작하는데 수식 (3)에서  $D(i, j-1), D(i-1, j), D(i-1, j-1)$ 은 입력된 지점으로부터 주변 값들을 탐색하고 그중 최솟값을 지니는 경로로 이동하게 된다. 이렇게 만들어진 경로를 기반으로 거리를 계산함으로 시간에 따른 왜곡을 어느 정도 보정 할 수 있다.

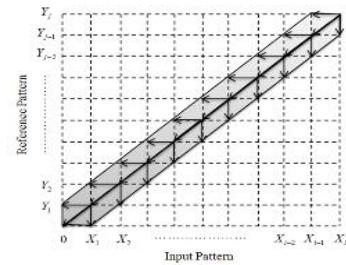


그림 4. DTW Matrix

그림 5는 유클리디언 거리와 DTW 알고리즘의 2개의 신호 거리를 측정하는 방법의 차이를 보여준다.

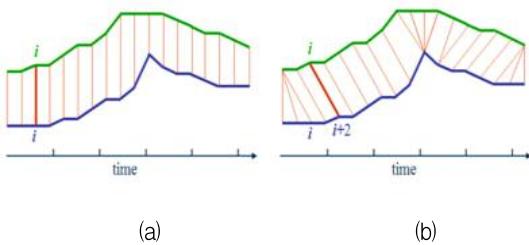


그림 5. DTW 와 유클리디언 거리 차이점 (a) 유클리디언 거리 (b) DTW

#### 다. k-NN 분류기

k-NN [8] 알고리즘은 입력된 벡터를 중심으로 가장 가까운 K개의 훈련벡터를 검출하고 그 중 가장 많은 빈도를 보이는 부류로 분류하는 알고리즘이다. K 근접 이웃 분류기는 k-NN 최근접 알고리즘과 흡사하게 동작하는 분류기이다. 훈련 집합  $X=\{(x_1, t_1), \dots, (x_N, t_N)\}$ 이 주어져 있다고 가정하자. 입력된 샘플  $x$ 를 분류해야 하였다. 다음 특징공간 안에서 입력된  $x$ 를 중심으로  $k$ 개의 특징들이 들어올 때까지 창을 확장시킨다.  $k$ 개가 들어오는 순간의 창의 크기를  $h_x^d$ 라 하면 창의 부피는  $h_x^d$ 가 된다. 창 안에 들어온 특징들 중 클래스  $w_i$ 에 속하는 특징 숫자를  $k_i$ 라 하였다. 수식 (4)는 k-NN 확률분포 추정을 설명한다.

$$P(x|w_i) = \frac{1}{h_x^d} \frac{k_i}{N_i} \quad (4)$$

수식 (5)는 수식 (4)의 결과를 이용해서 분류기를 만든 것이다. 간단히 표현해서 입력된  $x$ 에 가장 가까운 특징  $k$ 개를 구하고 가장 많은 빈도를 보이는 부류로 분류하라는 것이다.

$$\begin{cases} k-nn\text{분류기} : x \text{를 } w_q \text{로 분류하라.} \\ \text{이때 } q = \arg \max k_i \end{cases} \quad (5)$$

그림 6는 2차원 좌표에서 k-NN 분류기의 동작 원리를 보여준다.  $K=5$  일 경우 증가하는 창은 멈추고 빈도수가 많은 검은 부류로 입력 벡터  $x$  분류한다.

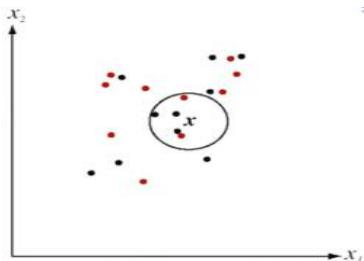


그림 6. k-NN 분류기 ( $K=5$ )

#### 라. DTW 거리척도 기반 k-NN 분류기

보통 K개의 훈련벡터를 검출하기위해 k-NN 알고리즘은 유클리디언 거리 기법을 사용한다. 하지만 앞에서 언급했듯이 동작신호는 시계열 신호이기 때문에 유클리디언 거리 기법은 적합하지 않다. 그래서 제안한 방법은 k-NN 알고리즘을 수정하여 유클리디언 거리를 계산하는 함수 대신 DTW 함수를 적용하였다. 하지만 DTW 알고리즘은 계산속도가 유클리디언 거리에 비해 느리기 때문에 N 개의 가까운 훈련 벡터를 검출할 때 많은 계산시간이 필요하다. 그래서 각각의 훈련 벡터를 K-means 알고리즘을 이용해서 대표 신호를 각각 5개를 추출한 다음 이것을 k-NN 알고리즘의 훈련 벡터로 사용함으로써 계산 속도를 향상 시켰다. 마지막으로 k-NN 파라미터인 K의 개수는 5개를 사용하였다. 그 이유는 4장에서 K의 개수가 5개 이상일 때 더 이상의 성능증가가 없는 것을 확인하였기 때문이다.

### III. 실험결과

실험환경은 Intel(R) Core(TM)2 Quad CPU Q9550이며 MATLAB 2012를 사용하여 CEPS 동작 신호를 분석 하였다. 먼저 스마트 TV기반의 실험에서는 총 동작 표 2 숫자 1부터 5까지의 동작 신호는 5개이며 각각 200개씩 1000개이다. 10명의 실험자를 통해 측정된 동작신호 센서로 부터 3m 거리를 기준으로 측정된 신호들이다. 1000개의 데이터 중 50%를 훈련 세트으로 사용하고 남은 50%를 테스트 세트으로 사용한다. 표 3에서 특정 칸 “원본”은 PCA를 사용하지 않고 원본 특징을 그대로 사용하는 걸 나타낸다. 그리고 인식기에는 베이지안 분류기, k-NN, 그리고 제안된 방법인 k-NN 과 DTW 알고리즘을 결합한 결과를 보여준다.

표 2. 실험에 사용된 동작

Number	Gesture	Image	Number	Gesture	Image
1	Left to right	→	6	Top to right	↗
2	Right to left	←	7	Letter 'V'	↙
3	Up to Down	↑	8	Reverse Letter 'V'	↖
4	Down to Up	↓	9	Circle	○
5	Top to left	↖	10	Reverse Circle	○

표 3. 성능 비교 분석 (스마트 TV)

특징	인식기	성능
원본	베이지안 분류기	64 %
PCA (-10%)	베이지안 분류기	72 %
원본	k-NN	85 %
PCA (-10%)	k-NN	88 %
원본	DTW+KNN	90 %
PCA (-10%)	DTW+KNN (제안된 방법)	92 %

실험 결과를 살펴보면 독립적으로 분류알고리즘 과정을 적용시킨 것 보다 제안한 방법처럼 알고리즘을 결합 시킬 때 좋은 성능을 보인다. 표 4은 스마트폰 기반의 실험결과를 보여준다. 스마트 TV 기반 실험과 다른점은 총 3000개의 동작신호를 3사람으로부터 각각 10개의 동작 신호 [표 2 (1-10) 동작]를 추출하여 실험하였다. 그리고 센서가 장착된 스마트폰 바로위에서 동작을 하고 신호를 추출하였다. 앞의 실험과 마찬가지로 50% 훈련 50%를 테스트 하였다.

표 4. 성능 비교 분석 (스마트폰)

특징	인식기	성능
원본	베이지안 분류기	67 %
PCA (-10%)	베이지안 분류기	71 %
원본	DTW	84 %
PCA (-10%)	DTW	96 %
원본	k-NN	88 %
PCA (-10%)	k-NN	91 %
원본	DTW+KNN	89 %
PCA (-10%)	DTW+KNN (제안된 방법)	92 %

표 5은 k-NN 분류기를 사용할 때 사용할 K의 개수에 따른 성능 변화를 나타낸다. 이론적으로 k가 증가하면 속도는 느려지고 성능은 증가해야 하지만 제안된 방법에서는 K-mean 알고리즘을 사용하여 각각의 동작에서 5개의 대푯값을 추출하였기 때문에 k의 개수가 증가하면 성능이 떨어진다.

표 5. k에 따른 k-NN 인식 결과 비교 분석

K	1	3	5	7	9	12
성능	0.85	0.88	0.92	0.92	0.89	0.89

#### IV. 결 론

이 논문은 CEPS 센서로부터 측정된 동작신호를 인식하는 방법에 대해 기술하였다. 인식 성능 92%는 스마트 TV에 적용하여 상용화하기엔 높지 않는 성능이므로 앞으로 더 많은 개선이 필요하다. 그리고 동작 신호를 추출할 때 비참가자를 인식대상으로 하는 실험이 필요하며 높은 성능을 위해서는 잡음이 제거되고 더 깨끗한 동작 신호들이 필요하다. 또 한 가지 반드시 해결해야 할 과제는 본 논문에서는 신호를 측정하고 동작 신호 자체를 스스로 추출하여 인식률만 실험하였기 때문에 목표하는 과정을 자동화 시키려면 강력한 동작 검출 알고리즘이 반드시 개발 되어야한다. 향후 이러한 문제점을 인식하고 연구를 앞으로 진행을 할 예정이다.

#### 참 고 문 헌

- [1] Kun-Woo Kim, Won-Joo Lee, Chang-Ho Jeon "A Hand Gesture Recognition Scheme using WebCAM", 2008년도 대한전자공학회 하계종합학술대회 제31권 제1호, 2008.
- [2] KwangHyun Lee, "Comparison of SpeechIntelligibility & Performance of Speech Recognitionin Real Driving Environments", The Journal of Korean Society of Speech Sciences pp.99-110 ISSN2005-8063, 2004.
- [3] Y.S. Kang, E.K. Lee, and Y.S. Ho, "Multi-Depth Camera System for 3D Video Generation," *Proc. of International Workshop on Advanced Image Technology (IWAIT)*, pp. 44(1-6), Jan. 2010.
- [4] Madhumita Venkataraman, "Biosensor can monitor your heartbeat from a distance", <http://www.newscientist.com-/blogs/onepercent/2011/11/sensor-monitors-your-heartbeat.html?DCMP=OTC-rss&nsref=online-news>, 2011.
- [5] Greg Welch and Gary Bishop, "An Introduction to the Kalman Filter", UNC-Chapel Hill, TR 95-041, 2006.
- [6] L. I. Smith A Tutorial on Principal Components Analysis, 2003[online] Available: <http://www.snl.salk.edu/~shlens/pub/notes/pca.pdf>
- [7] Eamonn Keogh, "Exact Indexing of Dynamic Time Warping", Proceedings of the 28th international conference on Very Large Data Bases, pp. 406-417, 2002.
- [8] P. Cunningham and S. J. Delany, "k-Nearest Neighbour Classifiers," Technical Report, University College of Dublin, March

2007.

- [9] 김형일. “모바일 단말에서의 가속도 센서 기반 제스처 인식 기술 개발”, 『경희대학교 대학원, 컴퓨터공학과』 학위논문(박사), 2012. pp.1-99.
- [10] 조정재. “비접촉식 전위계 센서를 이용한 적응적 DTW 알고리즘 구현”, 『전남대학교 대학원, 컴퓨터공학과』 학위논문(석사), 2013.

---

저자 소개

---



조정재

2010년 호남대학교 컴퓨터공학과 학사.  
2013년 전남대학교 전자컴퓨터공학과 석사.  
2014년~현재 전남대학교 전자컴퓨터공학과 박사 과정.  
<주관심분야 : 영상패턴인식, 객체검출, 신호처리>



김영철(정회원)

1981년 한양대학교 전자공학과 학사.  
1987년 Univ. of Detroit 전자공학과 석사.  
1993년 Michigan state Univ. 전자공학과 박사.  
1993년~현재 전남대학교 전자컴퓨터공학부 교수

<주관심분야 : 임베디드 SoC 설계, 스마트 단말기 및 TV 인터페이스 및 저전력 설계 기술>



문창협

2013년 전남대학교 전자공학과 학사.  
2014년~현재 전남대학교 전자컴퓨터공학과 석사과정.

<주관심분야 : 정보처리 시스템, 사용자 인터페이스 시스템, 신호처리>



김수형(종신회원)

1986년 서울대학교 컴퓨터공학과 학사.  
1988년 한국과학기술원 전산학과 석사.  
1993년 한국과학기술원 전산학과 박사.  
1997년~현재 전남대학교 전자컴퓨터공학부 교수

<주관심분야 : 영상패턴인식, 문서인식, 인공지능>



나인섭(정회원)

1997년 전남대학교 전산과 학사.  
1999년 전남대학교 전산통계산학과 석사.  
2008년 전남대학교 전산과 박사.  
2012년~현재 전남대학교 전자컴퓨터공학부 교수

<주관심분야 : 패턴인식, 영상처리, 전자도서관>