

3축 가속도 센서 기반 인간 행동 인식을 위한 기계학습 분석

이송미, 조희련, 윤상민
국민대학교

요약

최근 스마트폰의 이용 사례가 증가함에 따라, 스마트폰에 내장되어 있는 다양한 센서를 이용하여 인간의 행동을 인식하기 위한 연구가 많은 각광을 받고 있다. 본고에서는 인간의 기본적인 행동 중에 앉기, 걷기, 달리기 등의 행동 특성을 스마트폰에 내장되어 있는 3축 가속도 센서를 통하여 분석하고 인간의 기본적인 행동을 자동으로 인식하기 위한 방법에 대하여 비교 분석하는 것을 목적으로 한다. 구체적으로는 스마트폰에 내장되어 있는 3차원 가속도 센서로부터 추출된 데이터를 시간축에서 샘플링하여 인간의 행동을 인식하기 위한 기댓값 최대화 알고리즘, 랜덤 포레스트, 딥러닝 기반의 기계학습 방법을 비교하여 각 기계학습 알고리즘의 장단점을 분석한다.

I. 서론

사용자 중심의 사물인터넷 시대에서 인간의 의도 및 감정을 이해하기 위하여 인간의 다양한 행동 패턴을 분석하고 인식하기 위한 연구가 최근 들어 각광을 받고 있다[1]. 인간의 행동 패턴을 분석하고 이해하기 위해서는 먼저 인간의 물리적인 행동들 중 가장 대표적이고 기초가 되는 정적인 행동과 동적인 행동을 구분할 필요가 있다. 기본적인 동적 행동으로는 걷기와 달리기, 차량으로 이동하는 방법으로 구분되고, 정적 행동으로는 앉기 행동으로 구분할 수 있다. 본고에서의 인간의 행동에 대한 범위를 앉기, 걷기, 달리기의 세 가지로 정의하고, 이들 행동을 구분하여 자동으로 인식하기 위한 기술에 대하여 알아보도록 한다.

인간의 정적, 동적 행동을 분석하기 위한 방법은 센서의 부착 여부에 따라 크게 두 가지로 나뉜다. 그 중 한 방법은 사람의 신체에 직접 센서를 부착하여 사람의 행동을 인식하는 방법이고 [3], 다른 하나는 사람이 소지하고 있는 다양한 센서로부터 간접적으로 데이터를 수집, 분석하여 사람의 행동을 인식하는 방

법이다[2]. 센서를 직접 인간의 몸에 부착하여 행동을 인식하는 방식에서는 인간의 뇌파, 맥박, 혈압, 심전도 등을 측정할 수 있는 센서를 부착하여 센서로부터 수집된 정보를 종합적으로 분석하여 인간의 행동을 인식한다. 이와 같이 센서를 직접 인간의 몸에 부착하는 방식은 센서로부터 정확한 데이터를 추출할 수 있다는 장점이 있으나, 일반 사용자의 행동에 제약을 주고, 다양한 센서로부터 수집된 데이터를 종합적으로 분석하기 위한 별도의 디바이스가 필요하다는 단점이 있다. 따라서, 최근에는 사용자의 행동을 방해하지 않으면서 자연스럽게 행동 패턴을 이해하고 분석하기 위하여 스마트폰 등 다양한 휴대용 디바이스에 내장된 센서들로부터 사용자 정보를 수집하여 간접적으로 사용자의 행동을 인식하는 방법이 각광을 받고 있다. 또 스마트폰이 보편화됨에 따라 스마트폰에 내장된 GPS, 3축 가속도 센서, 자이로 센서, 지자기 센서 등을 활용하여 인간의 행동을 인식하는 기술에 대한 관심이 높아지고 있다[5][6]. 특히, 스마트폰에 내장되어 있는 3축 가속도 센서 기반 실시간 행동 인식 기술은 의료, 운동, 스마트 의류 분야 등에 널리 활용될 수 있다. 한편, 사물인터넷을 기반으로 하는 초연결시대에서의 사용자의 행동 인식 및 의도 분석은 다양한 활용분야에서 기본 기술로 요구되고 있다.

본고에서는 3축 가속도 센서를 활용하여 인간의 행동을 인식하기 위한 시스템을 <그림 1>에서와 같이 크게 두 단계로 구성하여 실험을 진행하였다. 첫 번째 단계로, 가속도 센서로부터 회전 성분을 제외한 데이터를 획득, 정형화하는 과정과 두 번째 단계로 기댓값 최대화, 랜덤 포레스트[7], 딥러닝 기반의 기계학습 방법에 따른 인간 행동 성능의 비교 분석하는 것을 목표로 한다.

본고의 구성은 다음과 같다. 2장에서는 기존의 가장 널리 활용되는 구글 행동인식 API 기술을 이용한 행동 인식 기술에 대하여 살펴보도록 한다. 3장에서는 본고에서 제안하는 다양한 기계학습별 행동인식 과정에 대하여 기술적으로 구체적으로 설명하도록 한다. 4장 및 5장에서는 직접 수집한 사용자 행동 데이터로부터 각 기계학습별 데이터에 대한 성능을 측정하고 분석하도록 한다.

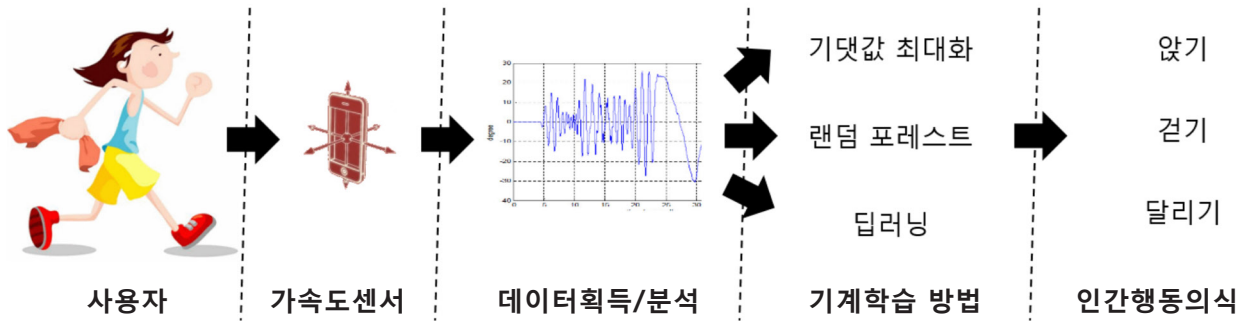


그림 1. 제안하는 인간 행동 인식을 위한 기계학습별 비교 분석 시스템 구성도

표 1. 구글 행동 인식 API를 이용한 행동인식 성능 분석

		On Foot	Still	Vehicle	Unknown
구글 행동 인식	걷기	28.34 %	-	22.40 %	31.15 %
	달리기	27.21 %	-	0.58 %	54.48 %
	앉기	0.78 %	3.50 %	-	95.70 %

II. 관련연구

스마트폰을 이용한 인간 행동 분석에 가장 많이 이용되는 기술은 구글에서 제공하는 구글 행동 인식 API를 이용한 행동 분석 시스템으로, 실시간으로 인간의 행동을 인식할 수 있다[4][8]. 구글 행동 인식 API는 Fused Location Provider (FLP)를 이용하여 사용자의 위치를 측정하고 가속도 센서, 자이로 센서, 기압 센서, 지자기 센서 등 다양한 센서로부터 수집한 데이터를 종합적으로 기계학습 분류기에 넣어 특정 행동 패턴과 가장 유사도가 높은 행동 패턴을 정답으로 판단하는 방법을 취하고 있다. <표 1>에서와 같이, 사용자의 움직임에 대하여 ‘On Foot’, ‘Vehicle’, ‘Unknown’ 세 가지로 구분하여 움직임 패턴을 구분할 수 있다. 특히, 사용자가 차량을 이용하여 이동하지 않았음에도 불구하고 ‘Vehicle’로 인식하는 경우가 매우 많이 있다. 또한, 유사도 값이 일정 임계값 이하로 나오는 경우에는 결과 판단을 ‘알수없음’으로 판단하는 단점이 있다. 특히, 사용자가 앉아있는 경우에 대부분 ‘unknown’으로 인식하였다. ‘Unknown’으로 인식하는 경우는 센서로부터 데이터의 변화가 거의 없거나, 급격하게 사용자의 움직임이 변화하여 센서의 값의 변화 폭이 많은 경우에 각 센서별 데이터의 임계값이 일정 값을 넘지 못하는 경우에 주로 ‘알수없음’으로 인식하는 경우가 매우 많이 발생할 수 있음을 알 수 있다. 따라서, 본고에서는 다양한 기계학습 방법을 이용하여 사용자의 움직임이 급격하게

변화하거나, 앉기와 같이 센서의 데이터의 변화량이 거의 없는 경우에도 행동을 강인하게 인식할 수 있는 방법에 대하여 비교, 분석하도록 한다.

III. 본론

본 절에서는 3축 가속도 센서로 들어온 데이터로부터 잡음에 강인한 데이터를 추출하는 방법에 대해 설명하고, 추출된 데이터에 기댓값 최대화 기법(Expectation Maximization, EM), 랜덤 포레스트(Random Forest), 그리고 1차원 Convolutional Neural Network (CNN) 기반의 딥러닝을 적용하여 기계학습별 행동 인식 성능이 어떻게 다르게 나오는지 비교한다.

스마트폰에 내장되어 있는 3축 가속도 센서는 지구의 중력 가속도를 이용하여 사용자 스마트폰의 X, Y, Z 축 방향으로의 힘의 크기를 측정하는 센서이다. 따라서 가속도 데이터의 각 축의 좌표 값은 사용자의 행동에 따라 바뀌게 되며, 이러한 변화 정보를 이용하여 사용자의 움직임을 인식할 수 있다. 그러나 이때 데이터의 회전 성분은 정확한 행동 인식에 방해가 될 수 있기 때문에, 이러한 회전 성분의 영향을 최소화함과 동시에 움직임을 정확하게 인식하기 위하여 X, Y, Z 축 방향으로의 움직임 값을 다음 식을 이용하여 구하도록 하였다.

$$V = \sqrt{X^2 + Y^2 + Z^2} \quad (1)$$

식 (1)을 이용하면 3축 방향으로의 가속도 값을 대표성분으로 변환시킴으로써 다양한 회전에 의한 오차를 줄일 수 있다. <그림 2>는 식 (1)을 이용하여 변환한 앉기, 걷기, 달리기 데이터의 특징점을 시간 축 위에 그린 결과이다. <그림 2>에서 보이는 바와 같이 앉기, 걷기, 달리기 등의 행동 패턴에 따라 V의 값이 매우 차이가 나는 것을 확인할 수 있다. 이는 사람의 움직임이 빨라질수록 가속도 센서로부터 추출된 V값의 변화 폭이 커짐을

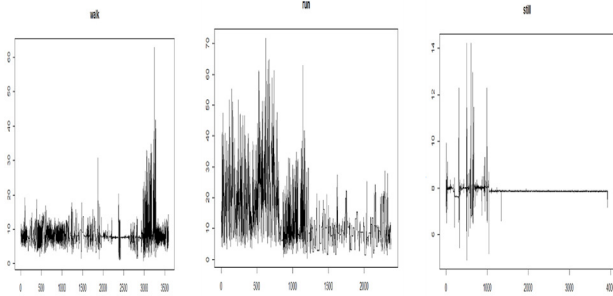


그림 2. 가속도 센서를 통하여 획득된 데이터의 예

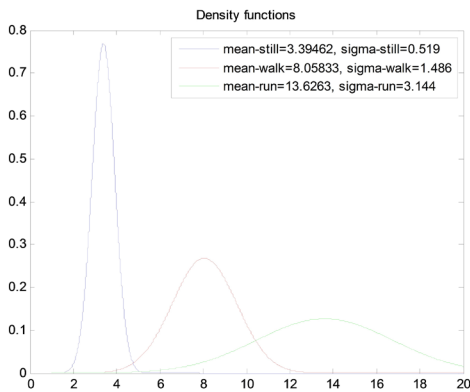


그림 3. 기댓값 최대화 알고리즘을 활용하기 위하여 앉기, 걷기, 달리기 행동 특성에 따른 확률 밀도 함수

의미한다. 비록 개인별 움직임의 차이나 스마트폰의 위치에 따른 차이로 인해 표준편차가 달라지기는 하지만 걷기, 뛰기, 앉기의 각각의 행동 패턴에 따른 차이가 존재하는 것을 <그림 2>를 통해 알 수 있다. 이후에서는 V 값을 바탕으로 기댓값 최대화 알고리즘, 랜덤 포레스트, 및 딥러닝을 이용한 기계학습 패턴인식 기술에 대하여 설명하도록 한다.

기댓값 최대화 알고리즘을 이용하여 행동 패턴을 분류하기로 하고 앉기, 걷기, 달리기의 V 값이 평균 $\mu_i (i=1,2,3)$ 과 분산 $\sigma_i^2 (i=1,2,3)$ 을 갖는 정규분포를 따른다고 가정했을 때, 입력되는 V 값이 앉기, 걷기, 달리기의 각 분포에 속할 가능성을 확률로 표현하여 행동 패턴을 구분하였다. 앉기, 걷기, 달리기 각각에 해당하는 행동 패턴에 대한 확률 밀도 함수는 <그림 3>과 같다. <그림 3>에서 보이는 바와 같이 앉기 행동에 대한 확률 밀도 함수의 분산 값은 매우 작아서 걷기 및 달리기와 구분이 잘되는 반면, 걷기와 달리기는 확률 밀도 함수가 많이 겹치는 것을 확인할 수 있다. 특히, 달리기의 경우 개인의 달리기 속도와 스마트폰의 위치 등으로 인하여 매우 큰 분산 값을 갖는 것을 알 수 있다. 각각의 개인별, 센서별 특성에 따라 많은 편차가 있는 인간의 행동에 대하여 인식하기 위하여 기댓값 최대화 알고리즘을 통하여 분류하도록 한다. 기댓값 최대화 알고리즘을 수행하기 위해서는 먼저 전체 행동 가속도 데이터 군집의 개

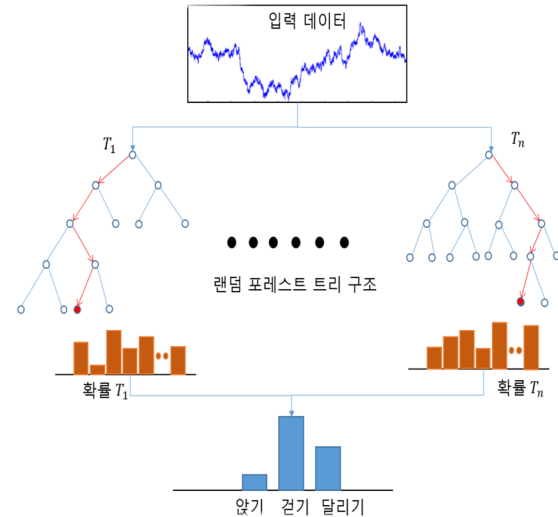


그림 4. 랜덤 포레스트를 이용한 인간 행동 인식 기술

수와 종료 조건을 지정하도록 한다. 그 후에 기댓값 최대화 알고리즘은 E-step 과 M-step으로 나뉘어 실행된다. E-step에서는 각 행동 패턴의 확률 분포에 대하여 입력된 가속도 센서의 V 값들이 속할 확률값을 계산하여 가중치로 변환하여 각 확률 모델에 배정한다. M-step에서는 새로 배정된 입력 값의 평균과 표준편차를 변경하며, 이때 입력 값의 가중치를 고려하여 likelihood 값에 따라 일정 임계값을 넘을 때까지 위의 E-step 및 M-step을 반복하여 최적화하여 행동을 인식하도록 한다. 그러나 위의 <그림 3>에서 보는 바와 같이, 달리기의 경우 분산 값이 매우 커서 다른 행동 패턴과 유사한 경우가 많이 발생하여 인식 성능이 떨어지는 단점이 있다.

랜덤 포레스트 기계학습 기법을 이용하여 각각의 행동 패턴 변동 폭이 큰 문제를 배깅(Bagging) 기법을 이용하여 결정 트리가 가지는 단점을 보완하여 성능을 향상하는 것을 목표로 하였다. 변동 폭이 큰 앉기, 걷기, 달리기의 패턴 분류를 위하여 임의성에 의하여 다른 특성을 갖는 다양한 트리로 구성하도록 하였다. 각각의 트리의 예측이 각 트리간 상관관계를 최소화하여 일반적으로 성능을 향상할 수 있는 것을 기본 원칙으로 한다. 랜덤 포레스트를 이용한 기계학습 방법은 특징 벡터의 크기에 덜 민감하고, 분류 속도가 빠르다는 장점이 있다. 랜덤 포레스트 분류기는 여러 개의 이진 트리를 앙상블 형태로 결합한 방법이다. 각각의 이진 트리의 구조는 랜덤하게 성장하는 방식을 따른다. <그림 4>에서 보이는 바와 같이 n 개의 이진 트리 구조에서의 확률 값을 종합적으로 분석하여 행동을 인식하기 위한 분류기를 형성하도록 하였다. 랜덤 포레스트를 이용한 인간 행동 분석 알고리즘은 비교적 빠른 시간에 강인한 성능을 보이는 장점이 있다.

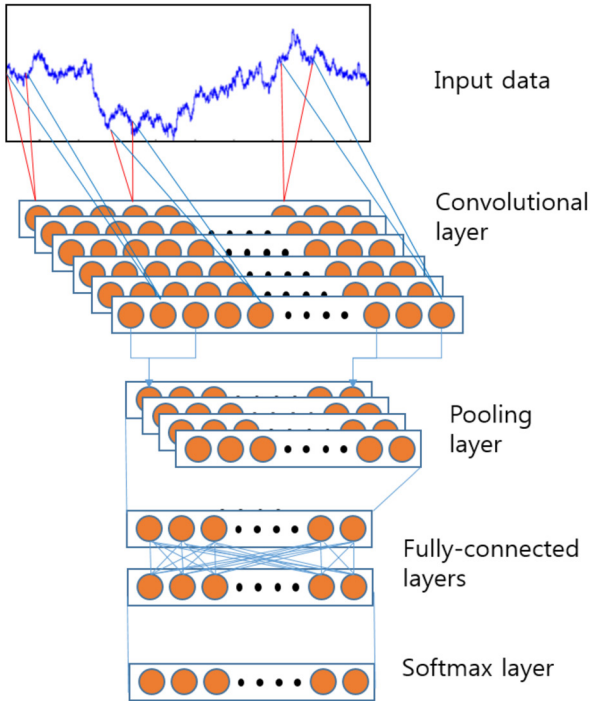


그림 5. CNN기반 딥러닝을 이용한 인간 행동 인식 기술

최근 패턴인식 분야에서는 딥러닝 기법을 이용하여 좋은 인식 성능을 보이고 있다. 특히, 본고에서는 CNN을 기반으로 행동을 인식할 수 있도록 시스템을 구성하였다. <그림 5>에서 보이는 바와 같이, 입력된 가속도 센서 값을 기반으로 패턴에 대한 학습을 위하여 CNN 기반 딥러닝 기술을 이용하였다. 구체적으로 CNN을 이용한 deep neural network 방법은 크게 convolution layer, pooling layer, fully connected layer, 그리고 softmax layer로 구성된다. Convolution layer를 통하여 입력된 가속도 특징 벡터 데이터로부터 필터와의 convolution을 통하여 부분 최댓값을 구하는 과정으로 특징 벡터의 부분 최댓값을 추출하도록 한다. Pooling을 통하여 부분 최댓값을 추출한 후에, weight값들의 여러 층들의 조합을 통하여 학습한 결과를 최종적으로 softmax 알고리즘을 이용하여 앉기, 걷기, 달리기 패턴으로 구분하는 과정을 거치는 과정으로 구성된다. 훈련 데이터를 이용하여 가중치 값을 자동으로 업데이트함으로써 인간의 행동을 인식하였다.

IV. 실험 결과

스마트폰에 내장된 3축 가속도 센서를 이용하여 인간의 물리적 행동을 인식하기 위하여 안드로이드 운영체제를 갖는 스마트폰을 이용하였으며, 스마트폰의 위치를 사용자의 가방, 손,

표 2. 기댓값 최대화, 랜덤 포레스트, CNN 기반 인간 행동 인식 성능 비교를 위한 confusion 행렬

		걷기	달리기	앉기
기댓값 최대화	걷기	74.85%	16.52%	8.63%
	달리기	14.64%	83.68%	1.68%
	앉기	11.18%	1.09%	87.73%
랜덤 포레스트	걷기	83.99 %	5.19 %	10.80 %
	달리기	5.40 %	94.37 %	0.2 %
	앉기	7.75 %	0 %	92.24 %
CNN	걷기	76.84 %	5.60 %	17.51 %
	달리기	0.31 %	99.29 %	0.38 %
	앉기	0.20 %	0 %	99.79 %

주머니 등으로 다양화하여 스마트폰에서 발생할 수 있는 상황을 재연하여 실험하였다. 본고에서 인간의 앉기, 걷기, 달리기의 행동을 인식하기 위하여 1초 단위로 스마트폰으로부터 수집한 3축 가속도 센서의 값을 검출하였으며, 10초 단위로 검출한 V값을 특징벡터로 인간의 행동 인식 실험을 진행하였다. 실험에 사용된 훈련데이터의 양은 총 497,512 개의 특징 벡터를 사용하였으며, 248,765개의 특징 벡터를 테스트 벡터로 사용하여 실험하였다.

<표 2>에서 각각의 기댓값 최대화, 랜덤 포레스트, CNN 기반의 딥러닝을 이용한 인간 행동 인식 성능에 대하여 confusion 행렬로 인식된 행동에 대하여 분석하였다. 세 가지 패턴 인식 알고리즘 중에서 CNN을 이용한 딥러닝 알고리즘의 성능이 기댓값 최대화 및 랜덤 포레스트 알고리즘에 비하여 성능이 우수하다는 것을 알 수 있다. 전반적으로 앉기 행동이 다른 행동에 비하여 특징 벡터의 특성이 매우 다르기 때문에 인식 성능이 높다는 것을 알 수 있다. 하지만, 걷기와 달리기에 대하여 많이 혼동되는 것을 알 수 있다. 이는, 실험 사용자의 행동 특성 및 휴대폰의 위치에 따라 걷기와 달리기에 대한 행동 인식에 많은 영향을 끼치는 것을 알 수 있다. CNN을 이용한 행동인식의 경우, 앉기 및 달리기 행동에 대한 인식 성능이 모두 99%를 넘는 반면에 걷기 행동에 대한 인식 성능은 상대적으로 낮다는 것을 알 수 있다. 식 (1)로부터 추출한 가속도 센서의 특징 벡터의 절대값이 걷기의 경우 달리기와 앉기 사이에 위치하기 때문에 걷기에 패턴이 달리기 및 앉기 행동으로 오인되는 경우가 많다는 것을 확인할 수 있었다. 또한, 앉기와 달리기 행동에 대한 특징 벡터의 값이 매우 큰 차이가 있기 때문에 <표 2>에서 보이는 바와 같이 잘못 인식하는 케이스가 적다는 것을 확인할 수 있었다.

V. 결론

본고에서는 3축 가속도 센서로부터 입력된 데이터를 기반으로 인간의 가장 기본적인 물리적 행동인 앉기, 걷기, 달리기 행동에 대하여 기댓값 최대화, 랜덤 포레스트, 및 CNN 기반의 딥러닝 기법을 이용한 기계학습에 따른 인식 성능을 비교 분석하였다. 특히, 딥러닝을 이용한 인간 행동 인식 시스템이 전통적인 기계학습 방법에 비하여 매우 우수하다는 것을 실험을 통하여 증명하였다. 실험을 통하여 분석한 바와 같이 가속도 센서로부터 회전 특성을 제외한 크기 값을 시간에 따라 분석함에 따라 앉기, 걷기, 달리기 특징에 대하여 사람의 특성, 휴대폰의 위치에 따라 혼동되는 경우가 많기 때문에 이에 대한 고려가 필요하다는 것을 알 수 있다. 이를 위하여, 시간 도메인이 아닌 주파수 도메인에서 각 행동에 대한 가중치를 부여하여 다른 행동 패턴 간 차이를 최대화할 수 있는 기술을 포함하면 행동 인식 성능이 향상될 수 있을 것으로 기대된다. 본 연구에서 분석된 기술을 바탕으로 다른 센서와의 조합을 통하여 차원을 줄임과 동시에 인식 성능을 향상할 수 있는 방법에 대하여 연구하도록 할 계획이다.

본고를 통하여 제안된 3축 가속도 센서를 이용한 행동 인식 기술은 의료, 운동, 스마트 의료 시스템에 적용할 수 있는 기반 기술로 활용이 가능하다. 특히, 본 연구를 통하여 개발된 인간의 기본적인 행동에 대한 이해를 바탕으로 복잡한 인간의 의도 및 감정을 이해할 수 있는 연구로 확장될 수 있다. 인간에 대한 의도에 대한 이해는 사용자 중심의 사물 인터넷 환경에서 상호작용에 활용될 수 있을 것으로 기대된다.

Acknowledgement

이 논문은 2016년도 정부(미래창조과학부)의 재원으로 정보통신기술진흥센터의 지원(R0115-16-1009, 현실체감형 외국어 학습을 위한 스마트러닝 인터랙션 콘텐츠 개발)과 2015년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(NRF-2015R1A5A7037615).

참고 문헌

[1] D. Roggen, G. Troester, P. Lukowicz, L. Ferscha,

- J. Millan, R. Chavarriaga, "Opportunistic Human Activity and Context Recognition," *IEEE Computer*, Vol. 46, No. 2, pp. 36-45, 2013.
- [2] S. Dernbach, B. Das, N.C. Krishnan, B.L. Thomas, D.J. Cook, "Simple and Complex Activity Recognition through Smart Phones," *Proceedings of Int'l Conf. on Intelligent Environments*, pp. 214-221, 2012.
- [3] Adil Mehmood Khan, Young-Koo Lee, Sungyoung Y. Lee, and Tae-Seong Kim, and Hwee-Pink Tan "A Triaxial Accelerometer-Based Physical-Activity Recognition via Augmented-Signal Features and a Hierarchical Recognizer", *IEEE Transactions on Information Technology in Biomedicine*, Vol. 14, 2010
- [4] Google, "Activity Recognition," <https://developer.android.com/training/location/index.html>, (2016. 08.22)
- [5] Mohammad Abu Alsheikh, Ahmed Selimz, Dusit Niyato, Linda Doylez, Shaowei Liny, and Hwee-Pink Tan, "Deep Activity Recognition Models with Triaxial Accelerometers," *AAAI Workshop*, 2016
- [6] Wenchao Jiang, Zhaozheng Yin, "Human Activity Recognition using Wearable Sensors by Deep Convolutional Neural Networks", *Proceedings of the 23rd ACM Int'l Conf. on Multimedia*, pp. 1307-1310, 2015
- [7] Leo Breiman, "Random Forest", *Machine Learning*, Vol. 45, pp. 5-32, 2001
- [8] 백승은, 박상원, "사용자 행위인식을 위한 Google Activity Recognition의 정확성 개선 방안," *한국정보과학회 학술 발표논문집*, pp.1770-1772, 2014

약 력



이 송 미

2016년 국민대학교 학사
2016년~현재 국민대학교 석사 재학
관심분야: 머신 러닝 기반 데이터 분석



조 희 련

1995년 연세대학교 학사
2005년 일본 교토대학 석사
2009년 일본 교토대학 박사
2011년~2014년 연세대학교 글로벌융합기술원
책임연구원
2015년~현재 국민대학교 연구교수
관심분야: 텍스트 마이닝



윤 상 민

2000년 고려대학교 공학사
2002년 고려대학교 공학석사
2010년 독일 TU Darmstadt 공학박사
2002년~2005년 삼성종합기술원 연구원
2010년~2011년 일본 AIST 박사후 연구원
2011년~2012년 연세대학교 융합기술연구소
조교수
2012년~2016년 국민대학교 컴퓨터공학부
조교수
2012년~현재 국민대학교 컴퓨터공학부 부교수
관심분야: 컴퓨터비전, HCI, 머신러닝