

Fuzzy ARTMAP 을 이용한 Dynamic Hand Gesture 실시간 학습 및 인식

Real-time Dynamic Hand Gesture Learning and Classification using Fuzzy ARTMAP

김 기 훈¹, 이 원 형^{2*}

¹⁾ 한동대학교 전산전자공학부 (TEL: 010-9551-1993; E-mail: 21500049@handong.edu)

²⁾ 한동대학교 전산전자공학부 (TEL: 054-260-1388; E-mail: whlee@handong.edu)

Abstract Previously proposed models for time series data such as RNN and LSTM have limitations in learning new categories in real time. This paper uses Fuzzy ARTMAP, which enables a model to quickly learn new categories with a small amount of computation time. In addition, This paper proposes a method for converting static hand joint data into dynamic information for learning and classifying dynamic hand gestures.

Keywords Fuzzy ARTMAP, Real Time Learning, Hand Gesture, Classification

1. 서 론

머신 러닝과 컴퓨터 비전의 기술의 발전으로 실시간으로 영상 정보를 처리하거나 제스처 인식에 대한 연구가 활발하다. 하지만 대부분의 기술들이 미리 학습된 모델을 통한 처리 기술을 적용하고 있어, 새로운 상황이나 학습되지 않은 범주에 대한 학습을 위해서는 데이터 수집과 학습을 위해 시간을 많이 소요하는 한계점을 가지고 있다.

이에 본 연구는 적은 연산으로 새로운 범주를 추가적으로 학습하는데 장점을 가진 Fuzzy ARTMAP[3]의 특성을 이용하여 실시간으로 새로운 범주의 학습이 가능한 동적 손 제스처 학습 및 인식모델을 제안한다.

2. 관련 연구

Adaptive Resonance Theory(ART, 적응공명이론)[1]은 1976년 Gail Carpenter와 Stephen Grossberg가 제안한 방식으로 기존의 Neural Network 학습 시 나타나는 Stability-Plasticity dilemma를 보완하는 알고리즘이다[2]. ART 신경 회로망은 기존에 학습되었던 값들이 변화되지 않고 새로운 내용을 학습시킬 수 있어, 시스템이 빠른 시간안에 안정될 수 있다.

ARTMAP[3]은 ART이론을 적용한 supervised learning 방식으로, Label을 가진 데이터에 대해 새로운 범주를 학습하거나 기존의 범주에 대한 추가적인 학습이 이루어지는 알고리즘이다. 또한, Fuzzy ART[4]는 ART 알고리즘

에 Fuzzy Theory를 적용시켜 연산 속도와 안정성을 높였다. 연산 과정 중에 제곱이나 루트 같은 연산을 사용하지 않고, AND, OR같은 연산을 최소, 최대 연산으로 대체하여 이전보다 적은 계산을 통해 모델이 학습할 수 있어 실시간 동작에서 장점을 가진다.

다만, 앞의 방법은 정적인 공간 정보를 학습하는 것에 최적화되어 있지만, 손 움직임과 같은 동적인 움직임을 학습하기에는 다소 한계가 있다. 또한, 기존 시계열 데이터에 대한 학습으로 제한된 RNN, LSTM은 학습단계에서의 데이터에 대해 모델 전체를 다시 학습시켜야 하는 단점이 있어 실시간 학습에 적합하지 않다. 이에 본 연구에서는 정적 정보에서 동적 정보를 실시간으로 얻어낼 수 있는 효과적인 방법을 제안하였다.

3. 제안하는 방법

3.1 학습 단계 및 인식 단계

학습 단계에서는 센서를 통해 얻어지는 손 관절 정보를 Fuzzy ARTMAP 모델의 입력으로 사용하고, 사용자가 입력한 Label로 실시간학습을 진행한다. 인식 단계에서는 학습된 Fuzzy ARTMAP의 Template matching을 통해, 가장 높은 값을 가지는 Category의 Label을 반환한다.

3.2 동적 손제스처의 Directional Momentum Representation

본 연구는 각 프레임 간 변화의 누적합과 각 좌표축의 방향 분리를 통한 동적 손 제스처의 Directional Momentum Representation을 제안한다.

* 이 논문의 연구는 과학 기술 정보 통신부와 정보 통신 기술 진흥 센터(IITP)의 소프트웨어중심대학 지원사업 (2017-0-00130)의 지원을 받아 수행하였음

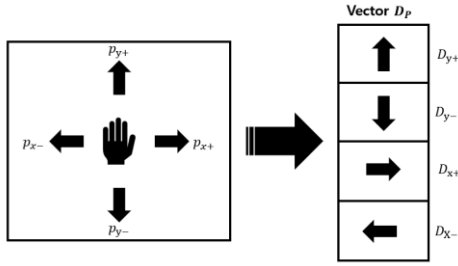


그림 1. Directional Momentum Representation for Hand position

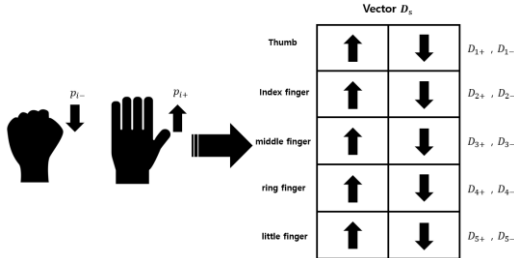


그림 2. Directional Momentum Representation for Hand shape

인식한 손 관절 정보에서 [그림 1]과 같이 손 중앙의 좌표 변화를 각 X, Y축의 + 방향과 - 방향으로 분리 ($p_{x+}, p_{x-}, p_{y+}, p_{y-}$) 하여 Hand position의 변화에 대한 Dynamic Gesture 인식할 수 있는 입력 벡터 $\vec{D_p} = (D_{x+}, D_{x-}, D_{y+}, D_{y-})$ 로 활용하였다.

또한, [그림 2]와 같이, 각 손가락과 손 중앙 사이의 거리 변화를 증가와 감소 여부로 분리 ($p_{1+}, p_{1-}, p_{2+}, p_{2-}, \dots, p_{5+}, p_{5-}$) 하여 Hand shape에 따른 Dynamic Gesture 인식을 위한 입력벡터 $\vec{D_s} = (D_{1+}, D_{1-}, D_{2+}, D_{2-}, \dots, D_{5+}, D_{5-})$ 로 활용하였다.

위와 같이 구성한 입력벡터의 t 시간에서의 각 성분 $D_i(t)$ 을 식(1)과 식(2)를 통해 Directional Momentum을 포함하도록 하였다.

$$D_i(t) = \gamma \cdot D_i(t-1) + \Delta_i/N \quad (1)$$

$$\Delta_i = \max(p_i(t) - p_i(t-1), 0) \quad (2)$$

$$i = \{X+, X-, Y+, Y-, 1+, 1-, 2+, 2-, \dots, 5+, 5-\}$$

위 식을 통해 얻어진 값은 이전과 현재 정보의 누적 합을 나타낸다. 기존 D_i 값에 $\gamma \in [0, 1)$ (forgetting factor)을 곱해주어 매 프레임 마다 감쇄 되도록 하였다. N 는 입력의 범위를 맞춰 주기 위한 정규화 상수이다.

이 방식을 통해 방향성이 다른 두 정보가 서로 상쇄되는 문제점을 해결하였고, 손의 모양 또는 손의 위치에 대한 동적 정보를 함께 측정할 수 있도록 하여 Dynamic Gesture 인식을 위한 적합한 표현방법을 제안하였다.

4. 구현 결과

본 연구에서는 Intel Real-Sense SR305 센서와 C++ 프로그램으로 제공되는 Hand-Viewer 프로그램을 통해 손을

인식하고, 나타난 손에 대한 관절 정보를 Socket 통신을 통해 Python 기반의 프로그램으로 전송하여 학습을 진행하였다. (https://github.com/rigswk/project_316)

앞서 설명한 3.2의 방식을 활용하여 입력 벡터를 구성하여 시스템을 구현한 결과는 [그림 3]과 [그림 4] 아래의 링크에서 확인할 수 있다. 사용자가 학습 시 취했던 Gesture에 대하여 해당 Label로 결과로 출력되는 것을 확인할 수 있다.



그림 3. Hand position change Gesture recognition result:
stay-left-stay (<https://youtu.be/nz32D1FPceo>)



그림 4. Hand shape change Gesture recognition result:
stay-grab-stay (<https://youtu.be/H2nVC2lFWjI>)

5. 결론

본 논문에서는 새로운 범주의 빠른 학습이 가능한 Fuzzy ARTMAP을 이용하고, Directional Momentum 표현 방법의 제안을 통해, 타 학습 방식에 요구되는 추가적인 학습 시간 없이 동적 손 제스처에 대한 실시간 학습과 인식이 가능한 방법을 제안하고 구현 결과를 보였다.

본 논문에서 제안된 방법은 앞으로의 HRI(Human-Robot Interaction) 상황에서 사용자 맞춤형 제스처로 여러 시스템 및 로봇을 제어하는 기술로써 유용하게 활용될 수 있을 것으로 기대한다.

참고문헌

- [1] S. Grossberg, "Adaptive pattern classification and universal recoding. II. Feedback, expectation, olfaction, and illusions", Biol. Cybernet. 23, pp. 187-202., 1976.
- [2] Carpenter, Gail, Grossberg, Stephen "A Massively Parallel Architecture for a Self Organizing Neural Pattern Recognition Machine", Computer Vision, Graphics, and Image Processing, vol 37, pp.54-115, 1987a
- [3] G.Carpenter, S.Grossberg, "ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network," in IEEE Transaction on Neural Networks, vol. 4, pp. 568-588, 1991.
- [4] G. Carpenter, S. Grossberg, "Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system." in IEEE Transaction on Neural Networks, vol. 4, pp. 759-771, 1991.