

동적 하이퍼네트워크 알고리즘을 이용한 사람 동작의 점진적 학습

장수용^o 김지섭 장병탁

서울대학교 컴퓨터공학부

syang@bi.snu.ac.kr, jkim@bi.snu.ac.kr, btzhang@bi.snu.ac.kr

Incremental Learning of Human Motions with Dynamic Hypernetwork Algorithm

Sooyong Jang^o Jiseob Kim Byoung-Tak Zhang

School of Computer Science and Engineering, Seoul National University

요 약

시계열 데이터 학습을 하는데 있어 기존 연구의 주된 접근 방식은 일괄 학습(Batch Learning)이었다. 이 방법은 모델링이 쉽고, 학습효율이 좋다는 이점이 있지만, 학습을 위한 모든 데이터가 준비되어 있지 않고 지속적인 학습을 요구하는 실제 환경에서는 적절하지 않다. 이에 본 논문에서는 점진적 학습(Incremental Learning) 방법에 근간하여 시계열 데이터 패턴을 학습할 수 있는 동적 하이퍼네트워크(Dynamic Hypernetwork) 알고리즘을 소개한다. 알고리즘은 학습 데이터 내의 각 자질 사이의 관계(association)를 부분적으로 저장하고 자기조직화를 통해 추론하는 구조를 지님으로써, 간단한 샘플링만으로 학습이 가능하다. 따라서 샘플을 기존 모델에 추가하는 과정을 통해 새로 주어진 데이터 또한 학습이 가능하므로 점진적 학습에 효과적이다. 실험에서는 Microsoft Kinect 모션 캡처 장비를 이용하여 사람의 다양한 동작을 측정하고 동적 하이퍼네트워크 알고리즘으로 순차적 학습을 진행하여, 알고리즘의 점진적 학습 성능을 확인하였다.

1. 서 론

점진적 학습은 일괄 학습 방식과 비교하여, 많은 양의 데이터를 학습하는데 유리하다[1]. 또한, 더 중요하게는 끊임없이 새로운 데이터가 유입되는 실제 환경에서 지속적인 학습, 혹은 평생학습을 하는 것이 가능하다[2]. 점진적 학습은 모델을 설계하는 측면에서 어려움이 많고 학습능력이 종종 떨어진다는 단점이 있지만, 로봇 에이전트와 같이 지속적인 학습이 필요한 경우 기 언급한 장점으로 인해 유용하게 사용된다.

하지만 다른 많은 수의 학습 알고리즘들과 마찬가지로 기존의 시계열 데이터를 모델링하는 알고리즘들은 주로 일괄 학습 방식을 취하여 연구되어왔다. 대표적인 시계열 학습 알고리즘인 은닉 마르코프 모델(HMM)을 비롯한 대부분의 동적 베이저안망(Dynamic Bayesian Network) 알고리즘들이 일괄 학습 방식을 사용한다[3]. Hinton 등에 의해서 소개된 incremental EM[4], Ishii 등의 stepwise EM[5]과 같은 점진적 학습 방법들도 연구되고 있으나, EM 알고리즘의 특성상 자질(feature)들 간의 관계가 복잡하고 차원이 높은 데이터의 경우, 학습 효율이 떨어져 사용할 수 없게 된다. 실생활에서 접하는 음악데이터나, 사람의 모션데이터 등은 이처럼 주로 고차원인 경우가 많고 적용이 어렵다.

하지만 확률 그래프 모델의 틀을 벗어나면 이러한

제약에 자유롭다. Zhang이 제시한 하이퍼네트워크(Hypernetwork) 알고리즘은 학습데이터 내의 각 자질 사이의 관계(association)를 부분적으로 저장하고 자기조직화를 통해 추론하는 구조를 지님으로써, 간단한 샘플링만으로 학습이 가능하다[6]. 따라서 샘플을 기존 모델에 추가하는 과정을 통해 새로 주어진 데이터 또한 학습이 가능하므로 점진적 학습에 효과적이다. 실제로 하이퍼네트워크 모델을 이미지, 텍스트 등의 정적인 데이터에 적용시켜, 점진적 학습 성능을 살펴본 연구도 진행된 바 있다[7, 8].

이에 본 논문에서는 하이퍼네트워크 알고리즘을 시계열 데이터에 적용 가능하도록 확장시킨 동적 하이퍼네트워크(Dynamic Hypernetwork) 모델을 소개한다. 기존 알고리즘이 주어진 데이터내의 자질을 공간적으로 샘플링 하였다면, 제시하는 모델에서는 샘플링 영역을 시간 축으로도 확장시켜, 시간에 따라 변하는 서로 다른 자질들간의 관계를 학습할 수 있게 하였다.

2. 제안 방법

이 논문에서 제안하는 알고리즘은 다음과 같은 방식으로 진행된다.

2.1. 학습 방법

학습 데이터에서 임의로 시간 t 의 데이터와 함께 타임 스텝만큼의 값을 고른다. 이렇게 선택된 값에서 일부를

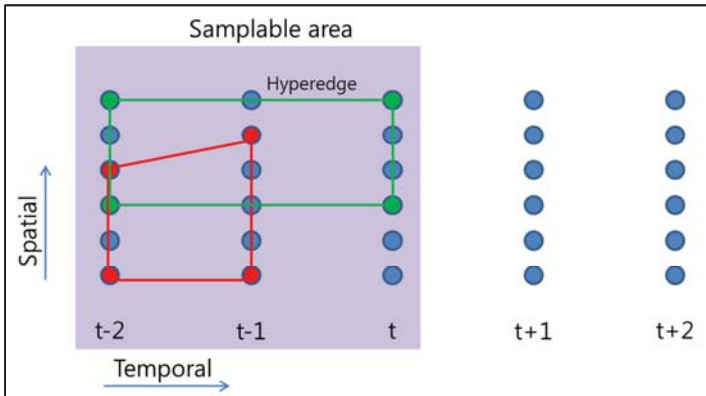


그림 2 하이퍼에지(Hyperedge) 샘플링

샘플링하여 하나의 하이퍼에지(hyperedge)로 만든다(그림 1). 위의 작업을 반복하여서 미리 정해진 개수만큼의 하이퍼에지를 얻고, 이런 하이퍼에지들을 하이퍼네트워크(Hypernetwork)에 추가한다.

2.2. 추론 방법

이렇게 만들어진 하이퍼네트워크를 바탕으로 주어진 시드로부터 데이터를 재생성한다. 다음 스텝의 데이터를 생성하기 위해서, 이전 스텝까지의 데이터와 하이퍼네트워크의 모든 하이퍼에지들과 매칭을 해본다. 모든 하이퍼에지들과 비교를 하면서 둘 사이의 값의 차이를 계산한다, 그리고 이 차이 값을 기준으로 미리 정해진 개수만큼 매칭이 잘 되는 하이퍼에지를 선별한다. 이렇게 선별된 하이퍼에지들의 값들의 가중평균을 다음 스텝의 값으로 정한다. 한 스텝씩 이동하면서 위의 작업을 반복하여 원하는 스텝만큼의 데이터를 생성한다.

3. 실험 및 결과

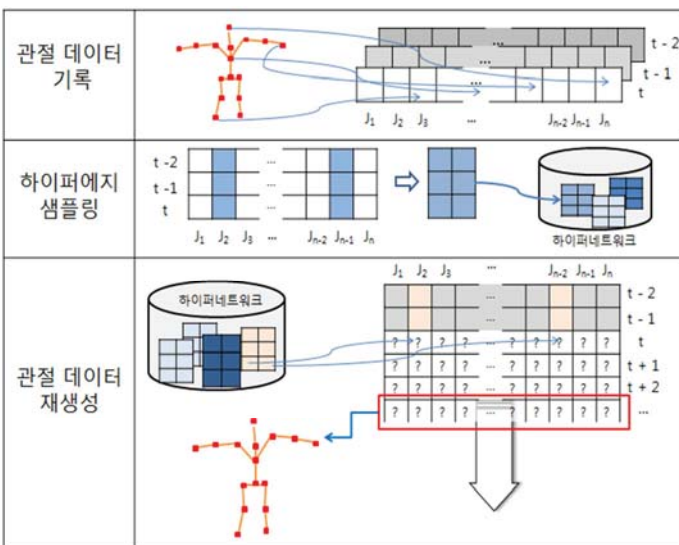


그림 1 동적 하이퍼네트워크를 이용한 시계열 데이터 학습 실험

3.1 데이터

이 논문에서 실험한 행동은 총 5가지로 정의하였다. 각 행동은 박수 치기, 손 흔들기, 악수하기, 가라는 수신호, 오라는 수신호이다.

이렇게 정의한 사람의 행동 데이터를 Brekel Kinect 프로그램을 이용하여 BVH 파일 형식으로 기록하였다. 이 형식으로 기록된 데이터에서는 관절은 총 18개로 팔에 4개씩, 다리에 3개씩, 머리, 목, 엉덩이, 척추에 한 개씩 있다. 그리고 각 관절마다 기본 자세에 대해 ZXY축에 대한 회전각을 기록하였다.

30 FPS로 행동을 다섯 번씩 기록하였으며, 행동에 따라 200 ~ 400 프레임이다. 이렇게 BVH형식으로 수집한 데이터에서 각 관절의 회전각 값을 추출한 후에 다운샘플링을 하여 학습데이터를 만들었다. 이렇게 만들어진 학습데이터는 관절마다 3개의 값을 가져서, 총 54 차원의 데이터 파일을 만들었다.

3.2 실험방법 및 절차

실험은 크게 두 가지로 나누어 진행하였다. 첫 번째 실험은 제안한 알고리즘의 기본적인 학습 성능을 검증하기 위하여, 하나의 행동의 데이터에 대해서만 학습을 하고 동일한 행동의 다른 데이터를 시드(seed)로 하여 실험하였고, 두 번째 실험은 점진적 학습 성능을 검증하기 위해서, 동일한 행동에 대해서 데이터 수를 늘려가면서 하고, 다른 행동을 섞어서도 실험을 하였다.

실험은 그림 2와 같은 과정으로 진행하였다. 먼저 Microsoft Kinect 모션 캡처 센서를 이용하여 행동 데이터를 수집하고, 수집된 데이터를 이용하여 제안된 알고리즘을 적용하였다.

각 실험은 동일하게 세 번씩 수행하여 얻은 결과를 평균을 내어서 실험결과로 활용하였다.

3.2.1 기본적인 학습 성능 검증

각 i 번째 행동의 j 번째 데이터를 a_i^j 라고 정의한다. 즉, a_1^1 은 박수치기의 첫 번째 데이터를 의미한다. 기본적인 학습 성능을 검증하기 위하여, a_1^1 만 학습한 후에, a_1^1, a_1^2, a_1^3 의 세 데이터의 처음 두 프레임을 각각 시드(seed)로 하여 재생성을 하였다. 실험의 결과는 그림 3과 같다. 위의 그래프는 a_1^1 의 처음 두 프레임으로 실험한 결과고, 아래 그래프는 왼쪽은 a_1^2 의 처음 두 프레임으로, 오른쪽은 a_1^3 의 처음 두 프레임으로 재생성하였다.

3.2.2 점진적 학습 성능 검증

점진적 학습 성능을 검증하기 위해서 데이터 수를 늘려가면서 재생성 실험을 하였다. 실험을 더 세분화하여, 동일한 행동 내에서 데이터 수를 늘려가면서 실험을 하고, 다른 행동 데이터를 섞어가면서 실험을 하였다.

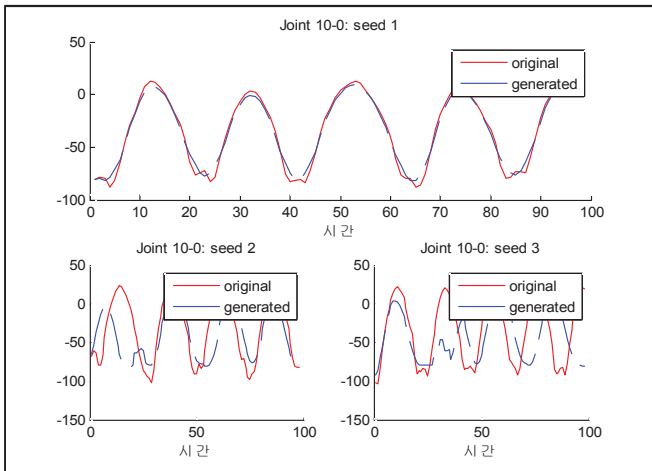


그림 3 한가지 학습 데이터로 학습하여 세가지 다른 시드(seed)로 재생성한 결과

3.2.2.1 동일한 행동만을 사용한 실험

처음에는 (a_i^1) 만을 사용하여 실험을 하고, 다음으로 (a_i^1, a_i^2) 를 사용하여 학습한 후에, a_i^1 의 처음 두 번째 프레임을 시드로 재생성하고, a_i^2 의 처음 두 번째 프레임을 시드로 재생성하였다. 이런 방식으로 (a_i^1, a_i^2, a_i^3) , $(a_i^1, a_i^2, a_i^3, a_i^4)$, $(a_i^1, a_i^2, a_i^3, a_i^4, a_i^5)$ 에 대하여 각각 실험하였다. 실험 결과는 그림 4의 위 그래프에 나타나있다. 성능의 평가척도로 관절값에 대하여 DTW값을 구하고, 이를 생성 프레임 길이로 나눈 후에 자연로그를 취하여서 사용하였다. 그리고 각 학습데이터 수 별로 평균을 취하여 이를 그래프로 그렸다.

3.2.2.2 다른 행동 데이터를 사용한 실험

이 실험에서는 각 행동의 첫 번째 데이터만을 사용하여서 사용하였다. 처음에는 (a_1^1) 만을 사용하였고, 다음에는 (a_1^1, a_1^2) 을 사용하여 실험하였고, 그 다음으로 (a_1^1, a_1^2, a_1^3) , $(a_1^1, a_1^2, a_1^3, a_1^4)$, $(a_1^1, a_1^2, a_1^3, a_1^4, a_1^5)$ 에 대하여 각각 실험을 하였다. 그림 4의 아래 그래프에 실험결과가 나타나있다. 실험 3.2.1과 같은 방식으로 결과를 표시하였다.

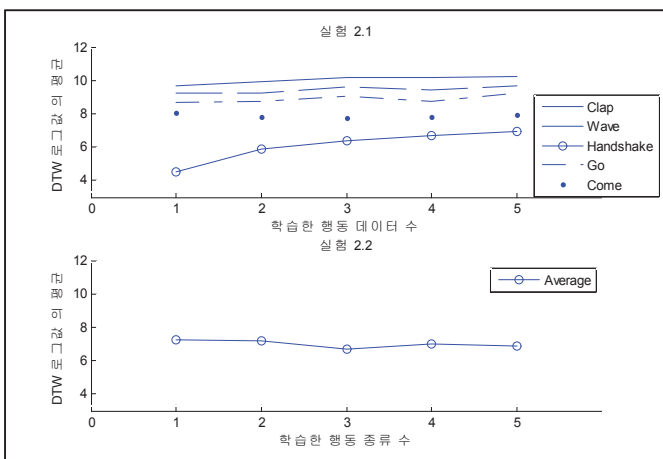


그림 4 학습 데이터 수를 늘려가며 실험한 결과

4. 결론 및 논의

본 논문에서는 시계열 데이터의 점진적 학습을 위한 알고리즘을 제안하고 그 성능을 분석해보았다. 제안된 알고리즘은 새로운 학습데이터의 부분을 기존 모델에 추가하는 방식으로 점진적 학습이 가능한 모델이다.

실험 결과 알고리즘의 기본적인 학습 성능을 확인할 수 있었고, 점진적 학습이 가능하다는 것도 알 수 있었다. 하지만, 점진적 학습의 성과의 평가척도로 사용하였던 DTW의 로그 값은 데이터 수가 증가할수록 오히려 값이 증가하기도 하였다. 이는 제안한 알고리즘이 파라미터 설정에 따라 영향을 많이 받는 특성에서 기인한 것으로 보인다. 예를 들어, 하이퍼에지의 개수를 높게 잡았을 경우, 특정 학습데이터에만 의존적인 하이퍼에지까지 샘플링되어 성능이 감소할 수 있다. 이런 점을 고려하더라도 제안한 알고리즘의 점진적 학습이 가능하다는 점은 의의가 있다고 볼 수 있다.

감사의 글

이 논문은 2012년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구이며(No. 2012-0005643, Videome), 정부(지식경제부)의 재원으로 한국산업기술평가관리원의 지원(10035348, mLife) 및 교육과학기술부의 BK21-IT 프로그램에서 일부 지원되었음.

참고문헌

- [1] C. M. Bishop, "Pattern Recognition and Machine Learning", Springer, 2006, pp. 143-144.
- [2] 장병탁, 차세대 기계학습 기술, 정보과학회지, 2007, 25(3).
- [3] K. Murphy, Dynamic Bayesian Networks: Representation, Inference and Learning, Thesis, 2002.
- [4] G. E. Hinton, R. M. Neal, A view of the EM algorithm that justifies incremental, sparse, and other variants, In Learning in Graphical Model, 1998.
- [5] M. Sato and S. Ishii, On-line EM algorithm for the normalized Gaussian network, Neural Computation, 12:407-432, 2000.
- [6] B.-T. Zhang, Hypernetworks: A molecular evolutionary architecture for cognitive learning and memory, IEEE Computational Intelligence Magazine, 3(3):49-63, 2008.
- [7] B.-T. Zhang, J.-W. Ha, and M. Kang, Sparse population code models of word learning in concept drift, In Proceedings of Annual Meeting of the Cognitive Science Society (CogSci 2012), pp. 1221-1226, 2012.
- [8] 이은석, 이지훈, 장병탁, A Computational Model of Language Learning Driven by Training, 한국인지과학회 춘계학술대회 발표논문집, pp.59-64, 2010.05.