

PHHMM(Product Hierarchical Hidden Markov Model)을 이용한 축구 비디오 분석¹

김무성, 강행봉
가톨릭대학교 컴퓨터정보공학부

A Soccer Video Analysis Using Product Hierarchical Hidden Markov Model

MooSung Kim, Hang-Bong Kang
Department of Computer Engineering
Catholic University of Korea
E-mail : { moostar, hbkang }@catholic.ac.kr

Abstract

일반적으로 축구 비디오 데이터는 멀티모달과 멀티레이어 속성을 지닌다. 이러한 데이터를 다루기 적합한 모델은 동적 베이저안 네트워크(Dynamic Bayesian Network: DBN) 형태의 위계적 은닉 마르코프 모델(Hierarchical Hidden Markov Model: HHMM)이다. 이러한 HHMM 중 다중속성의 특징들이 서로 상호작용하는 PHHMM(Product Hierarchical Hidden Markov Model)이 있다. 본 논문에서는 PHHMM 을 축구 경기의 Play/Break 이벤트 검색 및 분석에 적용하였고 바람직한 결과를 얻었다.

방식[3]을 적용하였다.

I. 서론

축구 비디오 분석 및 검색을 위해서는 스포츠 비디오 특유의 멀티모달 멀티레이어 속성을 고려해야 한다[1]. 이런 멀티모달 멀티레이어 속성을 다루기 위한 기존 모델 중 F.Wang(2005)[2]이 제안한 HHMM기반의 통합모델은 특징융합, 시퀀스 데이터 분석, 학습의 용이성 등의 장점을 가지고 있다. 그들은 특징융합을 위해 여러가지 변형된 HHMM을 제안하였고 이 중 PHHMM이 축구 비디오 데이터에 매우 바람직한 성능을 가짐을 보였다.

그러므로 본 논문에서는 축구 비디오 이벤트의 가장 기본이 되는 Play/Break 분석을 위해 DBN형태의 PHHMM을 적용하였다. 하지만 계산복잡도 감소와 모델의 확장용이성을 위해, 추론 알고리즘은 HHMM을 HMM으로 변환하는 기존의 방식이 아니라, DBN의 1.5Junction Tree

II. PHHMM

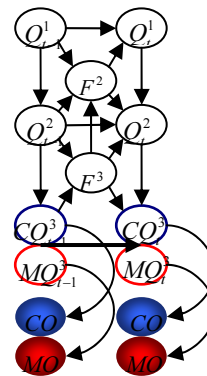


그림 1. 3-level PHHMM

DBN 형태의 HHMM은 최상층, 중간층, 최하층에 따라 다른 확률관계를 가진다[4]. 마르코프 가정에 의해 현재의 층은 바로 인접한 상하층의 상태에만 영향을 받는다. 그림 1은 3 계층의 PHHMM이다. 그림에서 F 노드는 하위층의 상태를 파악하는 지시노드로서, 바로 아래층의 처리가 끝나 종료상태가 되어야만 F 노드가 점화되고, 그에 따라 상위층의 상태는 초기 상태에서 천이상태로 들어갈 수 있다. HHMM과 PHHMM의 차이점은 최하위 은닉층 Q에서 상태간의 상호작용이다. 관찰노드 O의 은닉노드인 최하위층의 노드들간의 상호작용은 PHHMM에서

.¹ 본 연구는 문화관광부 및 한국문화콘텐츠진흥원의 지역문화산업연구센터(CRC)지원사업의 연구결과로 수행되었음

는 동일한 노드가 가지는 일련의 상태벡터로서 구현된다. 이는 기존의 HHMM과 동일한 것 같으나 축구 경기에서 특징들의 상호작용과 동기화를 다루기에 매우 적합한 특징융합 방식이다.

III. Play/Break 이벤트 탐지

축구 경기 비디오의 Play/Break 이벤트는 축구 경기 데이터의 이벤트 계층에서 가장 기본이 된다. Play/Break를 위한 칼라와 모션 특징들은 다음과 같이 계산된다. $M \times N$ 픽셀의 이미지 프레임에 대해, 칼라 특징의 경우

$$1) \text{ geometric moment: } m_{pq} = \sum_{y=0}^{N-1} \sum_{x=0}^{M-1} x^p y^q f(x, y)$$

$$2) \text{ Field Area: } FC = \frac{m_{00}}{MN}$$

3) Horizontal & Vertical Variance :

$$\sigma_x = \frac{1}{M} \sqrt{m_{20} / m_{00} - m_{10}^2 / m_{00}^2}$$

$$\sigma_y = \frac{1}{N} \sqrt{m_{02} / m_{00} - m_{01}^2 / m_{00}^2}$$

모션 특징의 경우

$$1) \text{ Motion Intensity: } m_i = \frac{1}{K} \sqrt{v_x^2 + v_y^2}$$

$$2) \text{ Motion Entropy: } m_e = - \sum_{i=0}^8 \frac{h(i)}{K} \log \frac{h(i)}{K}$$

이때 K는 모션 벡터의 총 수, h(i)는 8방향 중 i번째 방향에 속하는 모션 벡터의 수이다.

IV. 실험결과

실험에 사용된 동영상은 “ MANCHESTER UTD vs Bolton 후반전”, “ Clyde FC vs MANCHESTER UTD 전반전”, “ 스웨덴 vs 불가리아 전반전” 경기에서 추출되었다. 제안된 시스템으로서 PHHMM(Product HHMM)을 구축하였고, 칼라특징만 사용된 HHMM 과 모션특징만 사용된 HHMM 이 성능측정을 위한 기저모델로 사용되었다. 프레임 레벨(frame level)에서 이벤트를 분류한 결과의 정확도는 Play/Break 구분의 주요한 특징이 칼라이므로 역시 칼라특징만 사용된

HHMM 이 좋은 성능을 보였고, PHHMM 도 그와 비슷한 결과를 보였다(표 1 참조).

표 1. 프레임 수준에서의 추론 결과

모델유형	정확도
칼라 HHMM	78.83
모션 HHMM	63.86
제안된 PHHMM	80.05

멀티레이어 속성을 제대로 표현하고 있는지에 대한 척도로서 세그먼트 수준(segment level)의 정확도(Precision)와 회상률(Recall)에서는 PHHMM 이 기존 HHMM 보다 좀더 나은 결과를 보였다(표 2 참조).

표 2. 세그먼트 수준에서의 추론 결과

모델유형	정확도	회상률
칼라 HHMM	74.38	76.11
모션 HHMM	63.21	63.58
제안된 PHHMM	79.16	78.46

V. 결론 및 향후 연구 방향

본 논문에서는 축구 비디오 데이터의 멀티레이어 멀티모달 속성을 다룰 수 있는 통합모델로 DBN 형태의 PHHMM 을 택하여 이를 Play/Break 이벤트 분석에 적용하였고 바람직한 결과를 얻었다. 이후의 연구에서 HHMM 의 확장의 용이성에 의해 다른 이벤트를 추가하여 계층을 더 늘려갈 수 있을 것이다. 하지만 이를 위해서는 은닉 상태들에 대한 좀더 정교하고 명확한 정의와 분류가 필요하다.

참고문헌

[1] Comlumbia University, <http://www.ctr.columbia.edu/~xlx/research/soccer/index.html>

[2] F. Wang, Y.F Ma, H.J. Zhang, J.T. Li, “ A Generic Framework for Semantic Sports Video Analysis Using Dynamic Bayesian Networks” , IEEE, 2005.

[3] K. Murphy, “ Dynamic Bayesian Network: Representation, Inference and Learning,” Ph.D. Thesis, University of California, Berkeley, 2002.

[4] K. Murhpy, M. Pashin, “ Linear Time Inference in Hierarchical HMMs,” NIPS’ 01, Neural Info. Proc. Systems, 2002.