

# 은닉 마코프 모델을 이용한 골프 비디오의 시멘틱 이벤트 검출

김천석<sup>†</sup>, 추진호<sup>\*\*</sup>, 배태면<sup>\*\*\*</sup>, 진성호<sup>\*\*\*\*</sup>, 노용만<sup>\*\*\*\*\*</sup>

## 요 약

본 논문에서는 은닉 마코프 모델을 이용하여 골프 비디오의 시멘틱한 이벤트들을 검출하는 알고리즘을 제안한다. 본 논문의 목적은 하이라이트에 기반한 비디오의 색인 및 요약에 용이하도록 이벤트들을 식별하고 분류하는 것이다. 제안된 알고리즘은 먼저 골프 비디오의 분석을 통하여 4개의 이벤트를 정의하고, 각 이벤트를 구성하는 상태를 이용하여 HMM 모델을 설계한다. 또한 각 이벤트의 HMM을 구성하는 파라메타를 구하기 위해 MPEG-7 시각 기술자에 기반한 10개의 시각 정보 특징들을 이용한다. 실험 결과 제안된 방법은 다양한 골프 이벤트들을 식별하는데 있어 양호한 성능의 검출 결과를 보여 주고 있다.

## Semantic Event Detection in Golf Video Using Hidden Markov Model

Cheon Seog kim<sup>†</sup>, Jin Ho Choo<sup>\*\*</sup>, Tae Meon Bae<sup>\*\*\*</sup>  
Sung Ho Jin<sup>\*\*\*\*</sup>, Yong Man Ro<sup>\*\*\*\*\*</sup>

## ABSTRACT

In this paper, we propose an algorithm to detect semantic events in golf video using Hidden Markov Model. The purpose of this paper is to identify and classify the golf events to facilitate highlight-based video indexing and summarization. In this paper we first define 4 semantic events, and then design HMM model with states made up of each event. We also use 10 multiple visual features based on MPEG-7 visual descriptors to acquire parameters of HMM for each event. Experimental results showed that the proposed algorithm provided reasonable detection performance for identifying a variety of golf events.

**Key words:** Sport Video Analysis (스포츠 비디오 분석), Semantic Event Detection(시멘틱 이벤트 검출), Hidden Markov Model(은닉 마코프 모델)

## 1. 서 론

디지털 기술과 방송 통신 기술의 발전으로 멀티미디어 콘텐츠가 늘어나면서, 비디오 색인, 요약을 위

한 시멘틱한 이벤트 검출에 관한 연구가 진행되고 있다[1-3]. 특히 스포츠 비디오는 전체 내용에 비해 반복적인 구조를 가지는 몇 개의 이벤트로 요약 할 수 있기 때문에 다양한 스포츠 장르에서 시멘틱한

※ 교신저자(Corresponding Author): 김천석, 주소: 대전광역시 유성구 문지로 119번지(문지동 103-6)(305-714), 전화: 042)866-6279, FAX: 042)866-6245, E-mail: cheonseog@icu.ac.kr

접수일: 2004년 2월 20일, 완료일: 2004년 6월 11일

<sup>†</sup> 준회원, 한국정보통신대학교 대학원 공학부 박사과정

<sup>\*\*</sup> 준회원, 삼성전자 디지털 미디어 총괄 근무

(E-mail: wh966@icu.ac.kr)

<sup>\*\*\*</sup> 서울대학교 컴퓨터 연구소 연구원

(E-mail: heartles@icu.ac.kr)

<sup>\*\*\*\*</sup> 준회원, 한국정보통신대학교 대학원 공학부 박사과정

(E-mail: sselmo2000@icu.ac.kr)

<sup>\*\*\*\*\*</sup> 종신회원, 한국정보통신대학교 공학부 부교수

(E-mail: yro@icu.ac.kr)

※ 본 연구는 한국과학재단의 특정기초연구 (승인 번호 2000-1-303-004-3)의 일환으로 수행된 연구 결과입니다.

이벤트 검출에 관한 연구가 활발히 진행되고 있다. 관련 연구로는 최대 엔트로피 방법에 의한 야구 경기 장면 분류, 자막 정보 추출에 의한 미식축구 분석, 오디오와 모션강도 특징을 이용한 축구경기의 골 장면 검출, 압축 영역에서 테니스 경기의 구조적인 분석을 통한 서브 이벤트 검출 연구가 있다[4-7].

스포츠 장르의 하나인 골프는 대중적인 관심의 증가에 따라 다양한 방송 매체에 의한 골프 경기 비디오 콘텐츠는 급속히 증가하고 있다. 그러나 장시간의 경기 내용에 비해 중요한 시멘틱한 장면들이 있는 시간은 상대적으로 짧다. 또한 디지털 방송 기술의 발달은 방송을 시청자 기반의 지능적인 시스템으로 진화시키고 있다. 이와 같은 환경은 다른 스포츠 장르와 마찬가지로 골프 콘텐츠도 내용기반 색인, 요약 정보가 필요하지만, 경기에 대한 관심과 양에 비하여 관련 연구는 미미한 실정이다.

스포츠 비디오 콘텐츠 분석 연구 중 골프 관련 연구로는 Peker[8]의 연구가 있다. Peker는 MPEG-7 움직임 강도 서술자(motion activity descriptor)에 의해 P프레임 마다 추출된 모션 강도 값들을 미디언 필터로 처리하고, 이를 10개 단위로 해서 얻은 평균 값에 대해 양자화 한 후, 설정된 임계값에 의해 드라이브 샷과 퍼팅 샷에 대한 이벤트를 검출하였다. 그러나 이 방법은 실험적인 방법에 의해 임계값을 구하여야 하는 어려움과 이벤트 검출 결과의 정확도가 낮은 문제점이 있다.

본 논문에서는 이 문제를 해결하기 위해 확률적 모델인 은닉 마코프 모델(hidden markov model, HMM)을 이용한 범용적인 골프 비디오의 시멘틱한 이벤트를 검출하는 알고리즘을 제시한다. HMM은 확률적인 모델에 기반 한 글로벌 프레임 워크로 축구, 야구, 테니스 등에서 적용한 사례가 있다[9-14]. 시멘틱한 이벤트 검출은 하이 레벨의 검출로서, 정의에 따라 매우 다양하기 때문에 모든 장르에 적합한 범용적인 검출방법은 아직 없다. 오히려 각 장르의 고유한 규칙에 기반 한 검출 방법이 더 효율적이다. 따라서 본 논문에서도 골프 비디오의 시멘틱한 이벤트 검출 모델을 위해 골프 비디오만 갖는 고유한 규칙을 이용한 룰 기반 방법(rule based on method)을 적용한다.

논문의 구성은 다음과 같다. 2장에서는 이벤트 정의와 전체 시스템 구조에 대해 설명하고, 3장과 4장

에서는 본 논문에서 제안하고 있는 HMM의 구성 방식과 특징 벡터 및 추출에 대해서 설명한다. 그리고 5장과 6장에서 실험에 대한 결과와 결론 및 향후 계획에 대해 기술한다.

## 2. 시스템 개요

골프는 코스 위에 정지하여 있는 볼을 클럽으로 쳐서 정해진 홀에 넣어 그때까지 소요된 타수로 우열을 겨루는 경기이다. 코스는 티 그라운드(tee ground), 페어웨이(fairway), 벙커(bunker), 워터 해저드(water hazard), 퍼팅 그린(putting green), 홀(hole) 등으로 이루어져 있다. 이 가운데 페어웨이와 러프를 총칭하여 스루더 그린이라 하고, 벙커와 워터 해저드를 총칭하여 해저드라고 한다.

골프의 인 플레이(in play)는 한 홀마다 선수가 티 그라운드에서 스트로크를 하는 티 샷부터 시작하여, 어프로치 샷, 퍼팅 샷의 순서로 구성된다. 보통 18홀을 1라운드로 하여 3라운드 또는 4라운드의 경기를 하는데 이 과정을 반복한다. 플레이를 구성하는 각 과정은 일반적으로 시청자나 갤러리들이 관심 있는 이벤트들로 하이라이트나 요약은 이 이벤트들을 이용하여 구성한다.

본 논문에서는 관심 있는 공통 이벤트로서 드라이브 샷, 퍼팅 샷, 그리고 해저드 탈출과 관련된 벙커 샷의 검출을 목표로 한다. 본 논문에서는 비디오의 구조 분석 시 사용되는 용어인 샷의 의미와 혼동을 피하기 위해 이벤트 명의 뒤에 붙은 샷을 뺀 드라이브, 벙커, 그리고 퍼팅으로 정의한다.

골프 비디오에는 플레이와 상관없는 이벤트들도 다수 있다. 즉, 선수가 다른 지역으로 이동하는 장면, 관중 장면, 골프 경기장 코스 소개 장면 등과 같이 다양한 형태로 나타난다. 이런 장면들은 불필요한 이벤트로 검출할 필요가 없으나, 위에서 언급한 3개의 관심 있는 이벤트의 검출 효율성을 높이기 위해서 별도의 이벤트로 정의 하였다. 불필요한 이벤트는 다양한 종류로 발생할 수 있기 때문에 이 이벤트들을 모두 정의 하는 것 보다는, 발생빈도가 가장 높은 선수의 이동 장면인 워킹 이벤트를 대표 모델로 정의하여 분류한다. 표 1은 각 이벤트에 대한 정의이다.

이벤트 검출 과정은 그림 1과 같다. 먼저 입력된 MPEG 압축 비디오로부터 DC 이미지를 추출하고

표 1. 사용된 이벤트 정의

이벤트 정의	
드라이브 (drive)	선수가 티 그라운드, 혹은 페어 웨이에서 퍼팅 그린에 접근하기위한 행위로 티샷(tee shot), 세컨샷(second shot), 어프로치 샷(approach shot)들을 모두 포함한다.
벙커 (bunker)	선수가 모래로 구성된 벙커 해저드(bunker hazard)에서 탈출 하는 행위
퍼팅 (putting)	선수가 그린 주위에서 퍼터를 사용하여 공을 홀에 넣으려고 하는 행위로 퍼팅과 칩 샷(chip shot)을 포함한다.
워킹 (walking)	선수가 다음 플레이를 위해 코스를 건너 나 이동하는 행위로 불필요한이벤트의대 표적인 모델 이다.

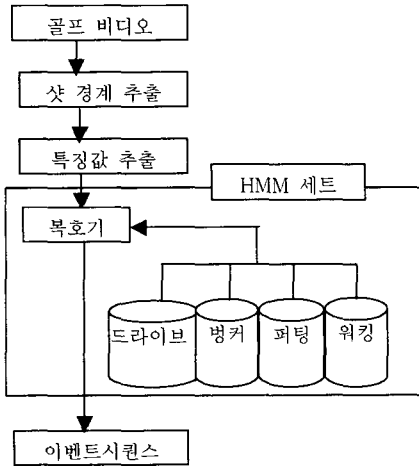


그림 1. 시스템 구조

추출된 이미지 칼라 값을 이용하여 샷 경계 구간을 검출한다[15]. 그리고 각 샷 구간에 대해 이미지 및 동영상 특징 값들을 추출한다[16-18]. 이미지는 선택된 샷 구간내의 모든 프레임에 대해 추출된 특징 값의 평균을 이용한다. 추출된 특징 값들은 복호기로 입력되어 미리 훈련된 HMM 세트에 의해 Viterbi 알고리즘 방법에 의해 최적의 상태 열을 찾아 출력한다[19].

### 3. 각 이벤트의 HMM 모델

#### 3.1 드라이브 이벤트 HMM

드라이브 이벤트의 HMM은 7개(N=7)의 상태로

구성하며, 각 상태의 정의는 표 2와 같다. 그림 2는 드라이브 이벤트에 대한 HMM 모델이다. 드라이브 이벤트는 항상 스윙 상태에서 시작해서 페어웨이나 온 그린 상태에서 끝난다. 그리고 얼굴 상태와 볼 상승 상태 간, 스윙 상태와 페어웨이 상태 간, 그리고 페어웨이와 그린 상태 간에는 천이가 발생하지 않는다.

#### 3.2 벙커 HMM

벙커가 대부분 퍼팅 그린 주변에 위치해 있기 때문에 온 그린(on green)을 위한 어프러치 샷이 대부분이어서, 스윙 이후에 그린을 향해 굴러가는 장면을 보여주는 것이 일반적이다. 따라서 표 3과 같이 2개의 상태(N=2)로 구성한다. 그림3은 HMM모델이다.

#### 3.3 퍼팅 HMM

퍼팅 이벤트는 스윙 상태, 그린 상태 그리고 홀 아웃 상태로 구성된다(N=3). 표 4와 그림 4는 상태

표 2. 드라이브 이벤트에서 상태의 정의

상태	정의
스윙	티 그라운드나 페어웨이에서 스윙 동작
얼굴	스윙 동작 후 카메라가 선수를 클로즈업
볼 상승	스윙 후 볼이 하늘로 상승하는 상태
볼 플라잉	볼이 하늘에서 이동 하는 상태
볼 하강	볼이 하늘에서 페어웨이나 퍼팅 그린을 향해 하강하는 상태
페어웨이	볼이 페어웨이에서 굴러가는 상태
그린	볼이 퍼팅 그린에 있는 상태

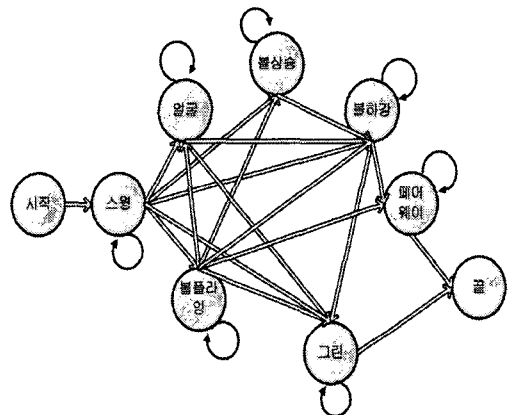


그림 2. 드라이브 이벤트의 HMM.

표 3. 벙커 이벤트에서 상태 정의

상 태	정 의
스 윙	벙커 안에서의 선수의 스윙 동작
그 린	볼이 퍼팅 그린으로 굴러가는 상태

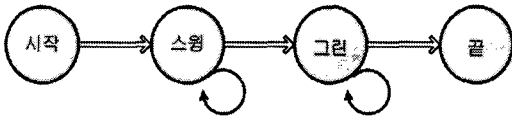


그림 3. 벙커 이벤트의 HMM.

정의와 HMM 모델이다.

워킹 이벤트는 드라이브, 벙커, 퍼팅 이벤트 이외의 관심 없는 불필요한 장면들을 필터링 하기 위해 정의된 것이다. 관심 없는 이벤트들은 각 이벤트간의 연관성이 없으므로, 본 논문에서는 간략하게 하나의 상태만을 가지는 간단한 HMM으로 정의한다(N=1). 표 5와 그림 5는 상태 정의와 HMM 모델이다.

3.5. HMM 파라 미터 추정

이벤트 검출에 은닉 마코프 모델을 적용하기 위해서는 다음과 같은 파라미터들이 필요하다[16].

- (1) N : 상태의 개수를 나타내며 각 개별 상태는  $S = \{S_1, S_2, \dots, S_N\}$  과 같다.

표 4. 퍼팅 이벤트의 상태 정의

상 태	정 의
퍼 텅	퍼팅 그린에서 볼의 퍼팅 동작
그 린	볼이 퍼팅 그린으로 홀 컵을 향하여 굴러가는 상태
홀 아웃	선수가 홀 컵 안이나 주위에 있는 공을 주우러 가는 상태

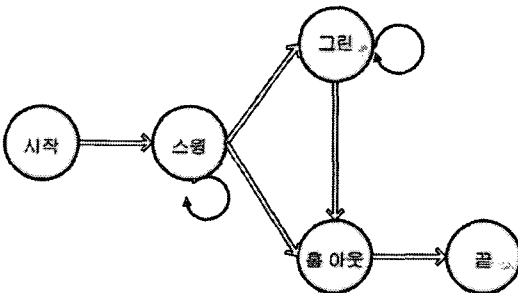


그림 4. 퍼팅 이벤트의 HMM.

표 5. 워킹 이벤트의 상태

상 태	정 의
워 킹	다음 플레이 장소를 위해 선수가 이동하는 동작



그림 5. 워킹 이벤트의 HMM 모델.

- (2) M : 각 상태의 관측 심벌 개수를 나타내며 개별 심벌은  $V = \{V_1, V_2, \dots, V_M\}$  과 같다.

- (3)  $a_{ij}$  :  $i^{th}$  상태에서  $j^{th}$  상태로의 천이 확률 분포.

$$a_{ij} = P(S_j | S_i) \tag{1}$$

여기서,  $1 \leq i, j \leq N$  이다.

- (4)  $b_j(k)$  : j 상태에서 관측 심벌 확률 분포

$$b_j(k) = P(V_k | S_j, at t) \tag{2}$$

여기서,  $1 \leq j \leq N, 1 \leq k \leq M$  이다.

- (5)  $\pi = \{\pi_i\}$  : 초기 상태 분포

$$\pi_i = P(S_i), \quad 1 \leq i \leq N \tag{3}$$

N은 각 이벤트마다 다르며, 초기 상태 분포는 각 상태에서 평균값 0, 분산 1을 가진다. M은 특징벡터 세트의 총 개수로서 구성되는 특징 요소들에 대한 것은 4 장에서 설명한다. 각 이벤트에 대한 HMM 모델을 기반으로 하여, 각 HMM 모델의 관측 심벌 확률 분포와 천이 확률 분포를 직접 구한다. 먼저 관측 심벌 확률 분포를 구하기 위해서 최대 가능도 추정량(maximum likelihood estimator)을 이용한다. 이때 모든 데이터는 식 (4), (5)의 평균 0와 분산 1을 가지는 독립적인 가우시안 랜덤 변수(gaussian random variable)들이다.

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i \tag{4}$$

$$v = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2 \tag{5}$$

다음으로 i번째 상태에서 j번째 상태로의 천이 확률( $a_{ij}$ )은 식 (6)과 같이 계산한다.

$$a_{ij} = \frac{\#xt_j}{N_i} \quad (6)$$

여기서,  $N_{next}$ 는  $i^{th}$  상태에서  $j^{th}$  상태가 나타나는 빈도수를 말하며,  $N_i$ 는 모든 데이터들 중에서  $i^{th}$  상태의 발생 빈도수를 나타낸다.

#### 4. 특징 벡터 세트 결정 및 추출

본 논문에서는 파라메타를 얻기 위해 관측 심벌로서 MPEG-7 비주얼 서술자를 기반 한 카메라 움직임, 모션, 칼라, 에지 성분들로 구성된 10개의 특징들을 사용한다[17-19]. 따라서 특징 벡터 세트와 구성되는 특징 요소는 식 (7)과 같다.

$$feature\_set = \{F_1, F_2, \dots, F_{10}\} \quad (7)$$

여기서,

- $F_1$  은 카메라 움직임 고정 성분 ( $CM_{fixed}$ ),
- $F_2$  는 에지 히스토그램의 수직 성분 ( $EH_{vertical}$ ),
- $F_3 \sim F_9$ 는 칼라 히스토그램 ( $CH_1 \sim CH_7$ ),
- $F_{10}$ 은 모션의 크기( $MM$ )이다.

##### 4.1. 카메라 움직임 (camera motion)

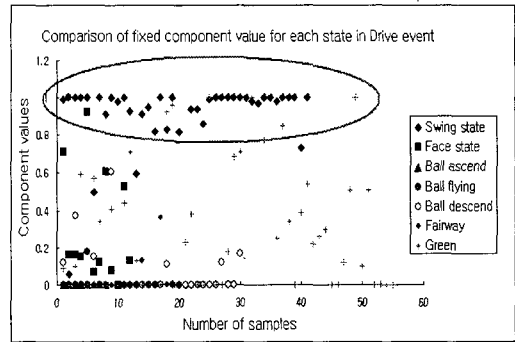
이벤트들의 스윙 상태는 그림 6과 같이 카메라 움직임 고정 성분 값이 다른 상태의 값들보다 높게 나타난다. 따라서 식 (8)과 같이 카메라 움직임의 고정 성분을 하나의 특징으로 이용할 수 있다.

$$CM_{fixed} = \frac{duration_{fixed}}{total\_duration} \quad (8)$$

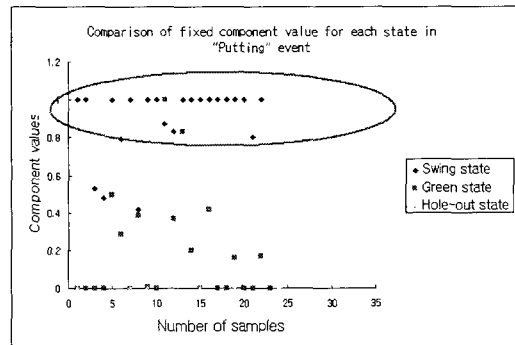
여기서  $total\_duration$ 은 주어진 비디오 클립의 총 시간만을,  $duration_{fixed}$ 는 카메라 움직임 성분 중의 고정 성분이 존재하는 시간을 나타낸다.

##### 4.2. 에지 (edge)

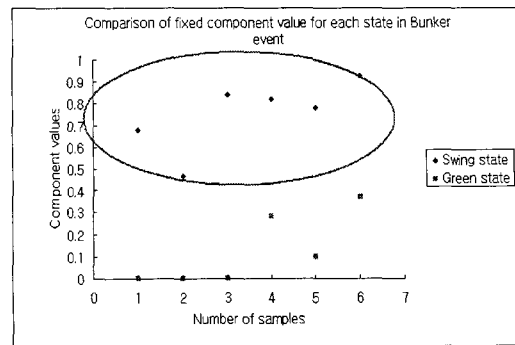
그림 7은 드라이브 이벤트에서 각 상태에 따른 수직 에지 성분값을 비교한 것으로 스윙 상태나 선수 얼굴의 수직 에지 성분이 다른 상태보다 높게 나타난다.



(a)



(b)



(c)

그림 6. 각 이벤트 상태의 카메라 움직임 고정 성분 비교: (a) 드라이브 이벤트, (b) bunker 이벤트, (c) 퍼팅 이벤트.

다. 본 논문에서는 식 (9)의 MPEG-7 에지 히스토그램 기술자의 글로벌 에지 수직성분을 이용한다.

$$Global\_Edge_{vertical} = \sum_{j=0}^{j=15} Local\_Edge_{vertical} \quad (9)$$

##### 4.3. 칼라 (color)

골프 비디오를 분석해 보면 독특한 칼라 특징들이

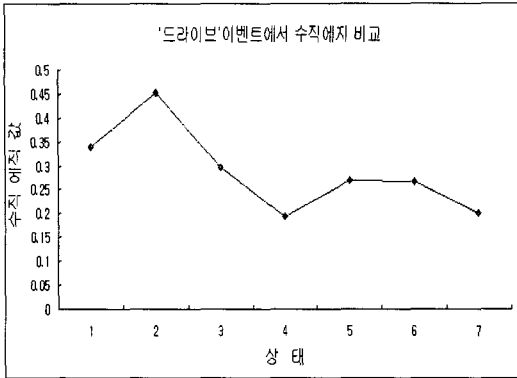


그림 7. 드라이브 이벤트의 글로벌 수직 예지 비교.

있다. 이 칼라 특징 값들을 MPEG-7 스케일러블 칼라 서술자(scalable color descriptor)의 HSV 칼라 공간에서 특정 범위의 색조(hue)와 채도(saturation)를 이용한다. 그림 8은 골프 비디오의 주요 장면들의 색조 히스토그램으로 드라이브 이벤트에서 선수의 얼굴은 range1 (1~39, 209~255), 그린 상태의 경우는 range2 (40~104), 볼 플라잉 상태는 range3 (105~209) 범위에서 독특한 특징이 있다. 따라서 식 (10)과 같은 특징 값을 사용한다.

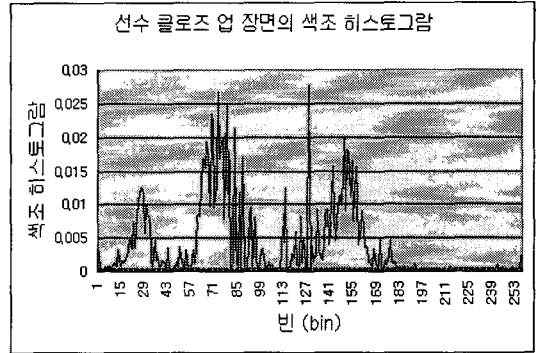
$$CH_k = \frac{1}{N} \sum_{j=1}^N \sum_{i=1}^N (h_j[i]/image\_size) \quad k=1,2,3 \quad (10)$$

여기서 N은 각 상태구간에서 전체 프레임 수이고, image\_size는 프레임 사이즈이다.  $h_j(i)$ 은 j번째 프레임에서 i번째 빈(bin)의 색조 히스토그램 값이다. k=1일 때  $i \in range1$ , k=2일 때  $i \in range2$ , k=3일 때  $i \in range3$  이다.

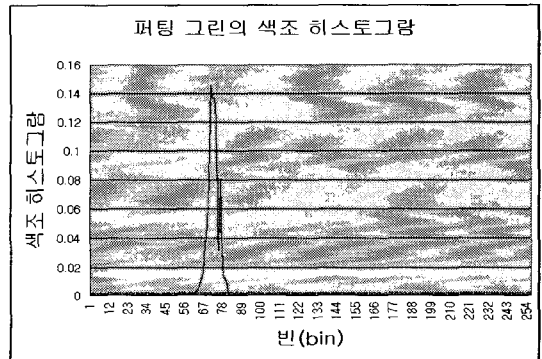
벙커 이벤트에서 나타나는 모래나, 퍼팅 이벤트에서 물과 같은 해저드는 그림 9와 같이 이미지 하단 혹은 상단의 일정한 부분에 존재한다. 또한 모래는 페어웨이나 그린과 유사한 색조 값을 가진다. 그래서 이들 특징 값을 추출하기 위해 채도 성분을 추가하고, 이미지를 상, 중, 하로 균등 3 분할하여 모래는 하단, 물은 상단 부분의 색조와 채도 비율을 모래는 식 (11), (12)에 의해, 물은 식 (13), (14)에 의해 구한다.

$$CH_4 = \frac{1}{N} \sum_{j=1}^N \sum_{i=1}^N (h_j[i]/segment\_size) \quad 40 \leq i \leq 59 \quad (11)$$

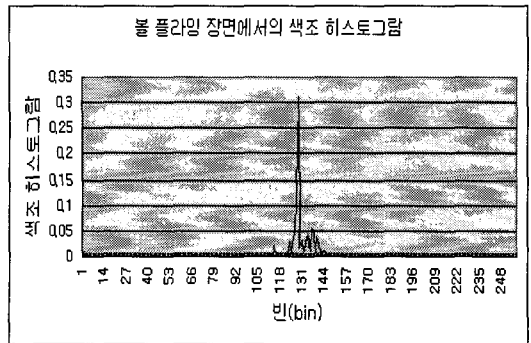
$$CH_5 = \frac{1}{N} \sum_{j=1}^N \sum_{i=1}^N (S_j[i]/segment\_size) \quad 0 \leq i \leq 25 \quad (12)$$



(a)



(b)



(c)

그림 8. 드라이브 이벤트의 주요 상태 장면의 칼라 색조 히스토그램: (a) 얼굴 클로즈업 상태, (b) 퍼팅 그린 상태, (c) 하늘로 상승하는 볼 상태.

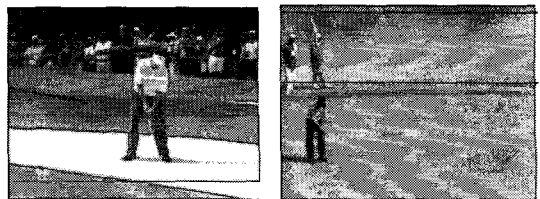


그림 9. 벙커와 퍼팅 이벤트에서 모래와 물의 위치.

$$CH_6 = \frac{1}{N} \sum_{j=1}^N \sum_i (h_j[i]/segment\_size) \quad 110 \leq i \leq 140 \quad (13)$$

$$CH_7 = \frac{1}{N} \sum_{j=1}^N \sum_i (S_j[i]/segment\_size) \quad 35 \leq i \leq 60 \quad (14)$$

여기서  $S_j(i)$  는  $j$  번째 프레임에서  $i$  빈(bin)의 채도 히스토그램 값이다.  $h_j(i)$  은  $j$  번째 프레임에서  $i$  번째 빈(bin)의 색조 히스토그램 값이다. segment\_size는 추출하고자 하는 영역의 사이즈로 벙커의 경우는 하단, 물 해저드는 상단 영역으로 각각 전체 이미지 크기의 1/3 이다.

4.4. 모션 (motion)

그림 10은 각 이벤트의 모션의 크기 변동으로 스윙 상태에서는 거의 존재하지 않다가 그린 상태에서 움직임이 커지게 된다. 이것은 카메라가 볼의 움직임을 추적하기 때문에 나타나는 현상이다. 이때, 모션의 크기는 비디오 스트림의 P 프레임의 모션 벡터 정보를 이용하여 다음과 같이 구한다.

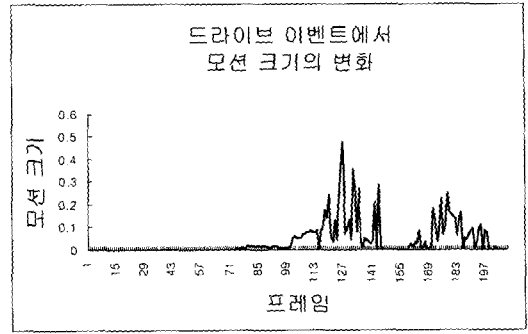
$$MM = \frac{1}{M} \sum_{i=1}^M (MV_i / image\_size) \quad (15)$$

$$MV_i = \sum_{j=1}^{330} \sqrt{mv_x(j)^2 + mv_y(j)^2} \quad (16)$$

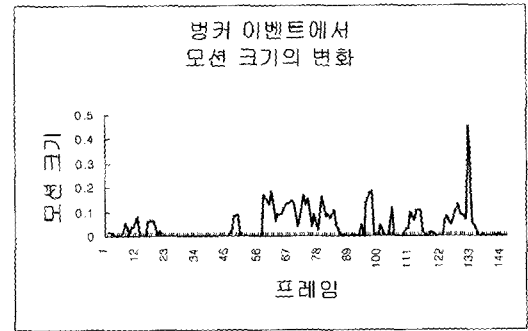
여기서 MM은 하나의 샷 구간에 존재하는 모든 P프레임의 모션 크기의 평균이고,  $MV_i$  는  $i$  번째 P 프레임에서 각 매크로 블록의 모션 벡터 크기의 합이다. 이때, M은 각 상태에서 P프레임의 개수이고,  $mv_x$ ,  $mv_y$  는 매크로 블록의 x, y 방향의 모션 벡터이다. 또한 j는 프레임에서 매크로 블록을 나타낸다.

5. 실험 결과

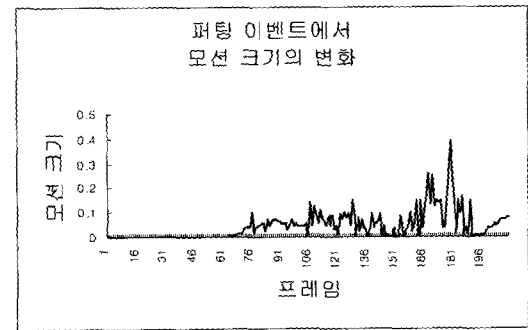
실험은 PGA(professional golfer's association)와 KPGA(korea PGA) 경기에서 추출된 7개의 비디오에 대해 실험하였다. 비디오 총 길이는 180분이며, 드라이브 이벤트 101개, 벙커 이벤트 9개, 퍼팅 이벤트 132개로 구성되어 있다. 이 중에서 3개의 비디오를 이용하여 훈련시키고, 나머지 4개의 비디오를 평가하기 위해 사용 하였다. 훈련에 사용한 데이터는 드라이브 이벤트 58개, 벙커 이벤트 6개, 그리고 퍼팅 이벤트 75개 등 총 90분 분량이다.



(a)



(b)



(c)

그림 10. 각 이벤트에서 모션 크기 변화: (a)드라이브 이벤트, (b)벙커 이벤트, (c)퍼팅 이벤트.

사용한 테스트 데이터베이스는 실험의 강인성을 위해서 서로 다른 경기 장소(미국, 한국)와 상이한 편집 스타일을 가진 방송국에서 제작된 것으로 구성 하였다. 각 이벤트에 대해 식 (17), (18)에 의해 검색율(recall)과 정확도(precision)를 구하였으며, 그 결과는 표 6과 같다.

$$recall = \frac{correct}{correct + miss} * 100 \quad (17)$$

표 6. 이벤트별 검출 결과

Event	Correct (개)	Miss (개)	False (개)	Accuracy (%)	
				Recall	Precision
드라이브	35	8	10	83.7	77.3
벙커	2	1	5	67.0	29.0
퍼팅	47	10	21	82.4	69.0
워킹	87	11	150	88.8	37.8
계	171	30	186	85.0	48.0

$$precision = \frac{correct}{correct + false} * 100 \quad (18)$$

각 이벤트에 대한 검색율은 67%에서 89%로 전체 평균 85%의 양호한 결과를 보여주는 반면에 정확도는 평균 48%이다. 정확도의 저하는 워킹 이벤트의 에러가 주요 원인이다. 그러나 워킹 이벤트는 관심 있는 이벤트 외의 불필요한 장면들을 필터링 하는 개념으로 사용한 것으로 드라이브, 벙커, 퍼팅 이벤트의 결과만 취합하면 평균 70%이다.

표 7은 본 논문에서 제시한 의미 없는 이벤트의 대표 모델로 사용한 워킹 이벤트의 사용 결과이다. 간단한 워킹 이벤트라는 대표 모델을 적용하였지만 드라이브와 퍼팅에서 정확도가 평균 50%, 퍼팅 검색율은 15% 향상 되었다. 반면에 벙커 이벤트의 결과에는 거의 영향이 없다. 이 결과는 워킹 이벤트 모델의 보완에 의해 검출 효율을 증대 할 수 있음과 동시에 벙커 이벤트는 워킹 이벤트 모델의 적용과는 무관함을 의미한다.

다음에 본 논문에서 제안한 방법을 Peker의 결과와 비교하였다. Peker는 움직임 강도(motion activity)의 특징 만을 이용하여 드라이브와 퍼팅 이벤트에 대해 검출 하였다. 표 8은 Peker의 검출 결과와 제안한 방법의 검출 결과를 평균하여 비교한 것으로

표 7. 워킹 이벤트의 효과 (단위 : %)

이벤트	워킹이벤트		
	없음	있음	
드라이브	Recall	88	83
	Precision	51	77
벙커	Recall	67	67
	Precision	29	29
퍼팅	Recall	71	82
	Precision	46	69

표 8. Peker의 결과와 비교. (단위 : %)

Accuracy	Peker 방법 (a)	제안 방법 (b)	상승율  1-(b/a) *100
Recall	59	82	39
Precision	44	73	66

Peker의 방법에 비해 제안한 방법은 임계값을 정하지 않아도 검색율은 39%, 정확도는 66% 향상 되었음을 보여주고 있다.

표 9는 검출된 이벤트의 구간을 실제 구간과 비교한 결과이다. 평균 오차는 입력된 실제 이벤트 구간의 프레임 수와 검출된 이벤트의 구간의 프레임 수의 차에 대한 평균값이다.

$$평균오차 = \frac{1}{N} \sum_{i=1}^N |L(i) - L'(i)| \quad (19)$$

여기서 N은 검출된 이벤트 개수, L'(i)는 검출된 이벤트 구간의 프레임 수, L(i)는 검출된 이벤트에 해당하는 입력된 이벤트의 실제 프레임 수이다. 평균 오차는 약 60 프레임 미만이다. 오차가 발생하는 구간은 준비동작이나 이동하는 장면 등과 같이 이벤트 시작이나, 끝나는 부위에서 발생 된 것으로 전체 의미에는 실제 사용상 큰 문제가 없다. 이벤트 시작에서 종료까지의 구간을 구하기 위해 Peker는 일괄적으로 10초라는 시간 제약조건을 부여하는 것에 반해, 제안 방법은 비교적 정확하게 자동으로 구간을 분류하는 장점이 있다.

(주) 정확도(%) =  $\frac{a}{a+b} * 100$

각 이벤트에서 발생하는 대표적인 에러의 유형은 다음과 같다. 드라이브 이벤트는 그린 근접한 곳에서의 어프로치 샷이 퍼팅 이벤트나 워킹 이벤트로 오인되거나, 관심 없는 이벤트들의 조합이 드라이브 이벤트로 잘못 인식되는 경우이다. 퍼팅 이벤트의 에러

표 9. 검출된 이벤트의 검출 구간 비교

	드라이브	벙커	퍼팅	전체 평균
평균 프레임 수(a)	550	452	626	542.67
평균 오차(b)	59.7	43.5	47.3	50.16
정확도(%)	90	90	92.5	90.8



는, 보통 관심 없는 이벤트들 중에서 퍼팅 이벤트와 유사한 특성을 가지는 이벤트들로 인해 발생된다. 예를 들어 골프 코스나 골프코스를 배경으로 스코어보드가 화면에 나타나는 경우이다. 벙커 이벤트의 경우는 다른 이벤트들의 상태 중에 모래를 포함하고 있는 상태가 있는 경우 오인될 확률이 높다. 예를 들어서 드라이브 이벤트에서 불이 벙커 주변에서 구르거나 불이 벙커 옆에서 멈추게 되어 화면에 벙커가 잡히는 경우 벙커 이벤트로 오인될 수 있다. 이는 벙커의 칼라 특징 외에 다른 특징들이 필요함을 의미한다.

지금까지 나타난 에러의 원인을 종합하면 본 논문에서 정한 각 이벤트의 모델과 특징 벡터 세트에서 약간의 모호성이 존재하고 있음을 보여주고 있다. 특히 불필요한 장면을 검출하는 대표 모델로 빈도가 많은 선수의 이동 장면만 모델링 하였기 때문에 이외의 경우에 대한 좀더 정교한 모델이 필요하다. 따라서 좀더 최적화된 모델과 특징들을 찾으면 더 나은 향상된 결과가 예상된다.

## 6. 결 론

본 논문에서는 골프 비디오의 시멘틱한 이벤트의 검출을 위해 은닉 마코프 모델을 이용하는 알고리즘을 제안하였다. 골프 비디오의 분석을 통해 각 이벤트들에 대한 HMM 모델을 설계하고, 칼라, 모션, 에지, 카메라 움직임 등으로 구성된 10개의 특징들을 이용하여 HMM의 파라미터들을 구하였다.

실험 결과 제안된 알고리즘은 기존의 골프 비디오의 이벤트 검출 방식과 비교해 볼 때, 1) 비교적 적은 특징 벡터 세트들로도 양호한 결과를 얻을 수 있으며, 2) 임계값 기반이 아닌 확률적인 모델 기반이기 때문에 다양한 골프 콘텐츠에 범용적인 적용이 가능하고, 3) 별도의 과정 없이 정확한 이벤트 구간을 검출할 수 있는 장점이 있다. 본 논문에서 제시된 알고리즘은 골프 비디오의 요약 및 하이라이트를 위한 저작 툴로 적용이 가능하다. 향후 좀 더 나은 성능 향상을 위해 HMM 모델의 개선과 멀티 모달리티(multi modality)의 적용에 대한 연구를 계속 수행할 것이다.

## 참 고 문 헌

[1] H. Zhang, "Content-based video indexing and retrieval," IEEE Multimedia, Vol. 1, No.2, pp.

- 62-72, 1994.
- [2] P. Aigrain, H. J. Zhang, and D. Petkovic, "Content-based representation and retrieval of visual media: A state-of-the-art review," Multimedia Tools Application., Vol. 3, pp. 179-202, 1997.
- [3] W.A. Khatib, Y. F. Day and A. Ghafoor, "Semantic Modeling and Knowledge Representation in Multimedia Databases," IEEE Transactions On Knowledge And Data Engineering, Vol. 11, No. 1, 1999
- [4] Wei Hua, Mei Han, Yihong Gong, "Baseball Scene Classification using Multimedia Features," ICME 02, Vol. 1, pp. 821-824, 2002.
- [5] Noboru Babaguchi, Yoshihiko Kawai and Tadahiro Kitahashi, "Event based Indexing of Broadcasted Sports Video by Intermodal Collaboration," IEEE Transactions on Multimedia, Vol. 4, Issue: 1, pp. 68-75, 2002.
- [6] Ziyou Xiong, Radhakrishnan. R, Divakaran. A, "Generation of sports highlights using motion activity in combination with a common audio feature extraction framework," Image Processing 2003, Vol. 1, pp. 5-8, 2003.
- [7] Di Zhong and Shih-Fu Chang, "Structure Analysis of Sports Video using Domain," ICME 2001, pp. 713-716, 2001.
- [8] K.A. Peker, R. Cabassen, and A. Divakaran, "Rapid Generation of Sport Video Highlights using the MPEG-7 Motion Activity Descriptor," Proc. SPIE, Vol. 4676, pp. 318-323, 2002.
- [9] S. Lefevre, B. Maillard, N. Vincent, "3 Classes Segmentation for Analysis of Football Audio Sequences," Digital Signal Processing 2002, Vol. 2, pp.975-978, 2002.
- [10] J. Assfalg, M. Bertini, A. Del Bimbo, W. Nunziati, P. Pala, "Soccer Highlights Detection And Recognition using HMMs," ICME 02, Vol. 1, pp. 825 -828, 2002.
- [11] NguyenNgoc Thanh, Truong Cong Thang, Tae Meon Bae, Yong Man Ro, "Soccer Video Summarization System Based on Hidden

Markov Model with Multiple MPEG-7 Descriptors," CISST, pp. 673-678, 2003.

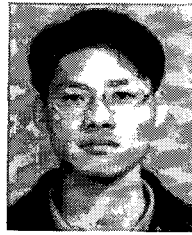
- [12] Chung-Lin Huang and Chih-Yu Chang, "Video Summarization using Hidden Markov Model," Information Technology: Coding and Computing 2001, pp. 473-477, 2001.
- [13] Peng Chang, Mei Han and Yihong Gong, "Extraction Highlights From Baseball Game Video with Hidden Markov Models," Image Processing. 2002, Vol. 1, pp. 609-612, 2002.
- [14] E. Kijak, G. Gravier, P. Gros, L. Oisel and F. Bimbot, "HMM Based Structuring of Tennis Videos using Visual And Audio Cues," ICME 03, Vol. 3, pp. 309-312, 2003.
- [15] Sang Heun Shim, Seung Ji Yang, Jeong Hyun Yoon, Ki Hyun Kim, Yong Man Ro, "Real-time Shot boundary detection for digital video camera using the MPEG-7 Descriptor," Real-Time Imaging VI, Proceeding on SPIE Electronic Imaging, Vol. 4666, pp. 161-171, 2002.
- [16] Manjunath, Philippe Salembier and Thomas Sikor, Introduction to MPEG-7 Multimedia Content Description Interface, John Willey & Sons, Baffins Lane, England, 2002.
- [17] Text of ISO/IEC 15938-3/FDIS Information technology Multimedia content description interface-Part3 Visual
- [18] Text of ISO/IEC 15938-8 DTR (Extraction and Use of MPEG-7 Description)
- [19] Lawrence R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," Proceeding of IEEE, Vol 77, no. 2, pp 257-286, 1989.



**김 천 석**

1981년 홍익대학교 전자공학과 학사  
 1983년 고려대학교 전자공학과 석사  
 2002년~현재 ㈜ 인터정보 연구소장  
 2001년~현재 한국정보통신대학교 공학부 박사과정

관심 분야: 영상 처리, MPEG-7, MPEG-21



**추 진 호**

2002년 홍익대학교 전자전기제어공학부 학사  
 2002년~2003년 한국전자통신연구원 위촉연구원  
 2004년 한국정보통신대학교 석사  
 2004년~현재 삼성전자 디지털미디어 총괄 근무

관심분야 : MPEG-7/21, 비디오 인덱싱, 데이터 방송



**배 태 면**

1996년 경북대학교 전자공학과 졸업, 학사  
 1998년 경북대학교 전자공학과 대학원 졸업, 석사  
 2000년 경북대학교 전자공학과 박사과정 수료  
 2004년 한국정보통신대학교 박사과정 수료

2001년 2월~현재 서울대학교 컴퓨터 연구소 연구원



**진 성 호**

2000년 중앙대학교 전자공학과, 학사  
 2002년 한국정보통신대학교 공학부, 석사  
 2002년~현재 한국정보통신대학교 공학부 박사과정

관심 분야: SmarTV, MPEG-7/21



**노 용 만**

1985년 연세대학교 전자공학과, 학사  
 1987년 한국과학기술원 전자공학부, 석사  
 1992년 한국과학기술원 전자공학부, 박사  
 1987년 컬럼비아대학 연구원

1992년 UC 어바인대학 초빙 연구원  
 1996년 UC 버클리대학 연구원  
 1992년~1997년 대전대학교 조교수  
 1997년~현재 한국정보통신대학교 공학부 부교수  
 관심 분야: 이미지/비디오처리 및 분석, MPEG-7, MPEG-21, 특정 인식