

일반논문 (Regular Paper)

방송공학회논문지 제25권 제4호, 2020년 7월 (JBE Vol. 25, No. 4, July 2020)

<https://doi.org/10.5909/JBE.2020.25.4.576>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

## 중계 영상을 활용한 야구 경기 분석 방법

손 종 웅<sup>a)</sup>, 이 명 진<sup>a)†</sup>

## Baseball Game Analysis Method Using Broadcast Video

Jong-Woong Son<sup>a)</sup> and Myeong-jin Lee<sup>a)†</sup>

### 요 약

레이더나 라이더 센서를 활용한 야구 경기 분석은 많은 비용이 요구된다. 본 논문에서는 중계 비디오에서 피치 샷과 타구 샷을 검출하고, 카메라의 움직임 기반 타구 궤적 생성 알고리즘을 제안한다. 제안하는 알고리즘은 객체 검출과 옵티컬 플로우 기반 피치 샷과 타구 샷 검출 이후, 프레임 간 변환 관계를 통해 프레임 내 타구 위치와 타구 궤적을 계산한다. 제안 방법은 KBO 중계 영상 시퀀스 3개에 대해 성능을 평가하였고 피치 샷과 타구 샷 검출 정확도와 검출률은 89-95[%] 이내의 성능을 보였으며, 평균 타구 위치 거리 차이는 13.6[m], 방향 차이 7.5°, 파울 분류 정확도 98.6%의 성능을 보였다.

### Abstract

Analyzing baseball games using sensors such as radars or riders is expensive. In this paper, we propose an algorithm to detect pitch shots and hit shots using baseball video and to generate ball trajectories within hit shots using camera movement. After the pitch shot and the hit shot detection using object detection and optical flow, we generate the transformation relationship between frames and ball locations in the frame, and calculates the ball trajectory. The performance of the proposed method is evaluated for three KBO baseball video sequences, and the detection accuracy and detection rate of pitch shot and hit shot were within 89-95 [%], and the average error for shot range was 13.6[m], The direction error was 7.5° and foul classification accuracy was 98.6%.

Keyword : Baseball Game Analysis, Ball tracking, Shot detection

a) 한국항공대학교 항공전자정보공학과(Dept. of Electronics and Information Engineering, Korea Aerospace University)

† Corresponding Author : 이명진(Myong-jin Lee)

E-mail: artistic@kau.ac.kr

Tel: +82-2-300-0421

ORCID: <https://orcid.org/0000-0002-3136-2819>

※ 이 논문은 2018년도 정부(교육부)의 재원으로 한국연구재단의 기초연구사업(2018R1D1A1B07050603)과 경기도 재원의 경기도 지역협력연구센터 사업(GRRC-항공2020-B01)의 일환으로 수행된 연구결과임. (This work was partly supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT)(No. NRF-2018R1D1A1B07050603), and by the GRRC program of Gyeonggi province (No. GRRC-KAU2020-B01)).

· Manuscript received April 2, 2020; Revised May 25, 2020; Accepted May 25, 2020.

## I. 서론

SPOTV, KBO(Korea Baseball Organization)에서 제공되는 야구 경기 하이라이트 영상은 주로 득점 장면을 중심으로 구성되어 있다. 그러나, 시청자들은 특정 선수 혹은 수비 포지션 중심의 분석, 수비 하이라이트, 자동화된 기록과 경기 통계에 대한 수요가 존재한다. 자동화된 경기 분석을 위해 그림 1과 같이 여러 센서를 활용한 Pitch f/x, 스탯캐스트(Statcast)<sup>[1]</sup>와 같은 시스템이 상용화되어 있다. 이들은 광학 카메라, 레이더와 같은 고가의 장비를 필요로 하며 스탯캐스트의 경우 한 경기당 약 7TB 이상의 데이터를 이용하여 경기를 분석하기 때문에 경기장 별 분석 시스템을 도입하기 위한 비용 부담이 매우 크다.

샷이란 하나의 카메라가 촬영을 시작해서 정지할 때까지 혹은 다른 카메라의 시점으로 이동할 때까지 얻어진 일련의 프레임들을 의미한다. 비디오 콘텐츠 분석을 위해 샷(Shot) 단위로 비디오를 분할하여 해석하고자 하는 연구들이 진행되어 왔다<sup>[5][8][9]</sup>. 이러한 연구는 다양한 분야에서 적용될 수 있으며 영화나 드라마에서 의미론적인 장면 생성하기 위한 방법<sup>[6][7]</sup>, 야구 경기 이벤트 검출 및 분류를 위해서도 활용된다<sup>[3][4][5][10][11]</sup>.

경기에서 발생하는 득점, 수비 등과 같은 이벤트는 주로 공의 이동에 의해 발생하며 타자가 공을 친 경우에 중계 카메라는 공을 추적하며 이동한다. 따라서, 중계 영상에서 타자가 친 공이 날아가는 타구 샷(Hit shot) 내에서 카메라의 움직임을 활용하면 경기장 내 타구의 이동 궤적을 파악할 수 있다.

본 논문에서는 특정 선수나 포지션 중심의 분석 또는 요약 서비스 등에 활용할 수 있는 야구 중계 비디오를 활용한 타구 궤적 생성 알고리즘을 제안한다. 제안 방법은 피치 샷 검출, 타구 샷 검출, 타구 궤적 생성 과정으로 생성되며, 타구 샷 내 프레임들의 타구 위치를 연결하여 타구의 궤적을 생성한다.

본 논문의 구성은 다음과 같다. 제2장에서는 스포츠 경기 분석, 야구 경기 분석에 관련된 기존 연구들을 설명한다. 제3장에서는 객체 검출과 옵티컬 플로어를 활용한 피치 샷과 타구 샷 검출 이후, 타구 샷 프레임 간의 호모그래피 변환 관계를 활용하여 타구의 위치를 계산하는 알고리즘을

제안한다. 제4장에서는 제안한 방법의 실험 결과를 제시하고, 마지막으로 5장에서는 결론을 맺는다.



그림 1. 스탯캐스트 시스템  
Fig. 1. Statcast system

## II. 관련 연구

### 1. 스포츠 경기 분석

Shih은 계층적인 스포츠 비디오 분석 모델을 제안하였다<sup>[2]</sup>. 그림 2와 같이 콘텐츠 분석 모델을 제시하고 기존 스포츠 비디오 분석 연구들을 제안된 모델 구조에 계층화하여 단계별로 분류하였다. 콘텐츠 분석 계층 모델은 객체, 이벤트, 상황 및 결과 단계로 구성되며 객체 단계에서 객체 검출, 추적 등을 정의한다. 이벤트 단계에서는 이벤트, 하이라이트 검출 단계로 장면 변화 검출, 대표 프레임 설정 등 객체 단위 발생한 이벤트를 정의한다. 상황 및 결과 단계는 상황적 추론 및 의미를 분석하는 단계로 이전 획득 정보를 응용하여 비디오 요약, 분류 등을 수행한다.

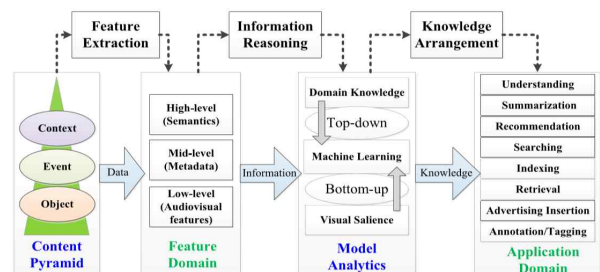


그림 2. 계층적 콘텐츠 분석 모델에 따른 비디오 분석 방법<sup>[2]</sup>  
Fig. 2. Video analysis method according to hierarchical content analysis model<sup>[2]</sup>

## 2. 야구 경기 분석

Liang은 피치 샷(Pitch shot)을 기준으로 특정 이벤트가 발생하기 때문에 각 피치 샷 내 프레임 위에 오버레이되는 캡션(Caption), 스코어보드(Scoreboard)를 인식하여 경기 상황 업데이트 시스템을 제안하였다<sup>[3]</sup>. 제안한 시스템은 그림 3과 같이 PSB(Pitch Scene Block)을 생성하며 이벤트를 탐색한다. 피치 샷 이후에 타구가 발생한다는 점은 대부분의 야구 중계 비디오에서 동일하지만, 캡션이나 스코어

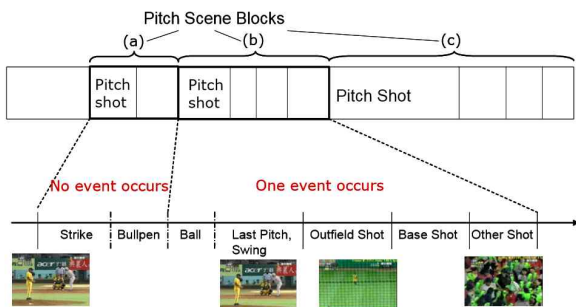


그림 3. PSB를 활용한 야구 이벤트 분류 방법<sup>[3]</sup>

Fig. 3. Baseball events classification method using PSB<sup>[3]</sup>

보드는 영상 제작자마다 달라질 수 있어서 일반화된 분석 방법이 될 수 없다.

Huang은 컬러와 카메라 움직임에 활용하여 야구 경기의 이벤트를 분류하는 시스템을 제안하였다<sup>[4]</sup>. 중계 카메라의 움직임 정보를 얻기 위해 샷 내에서 영상의 가장자리 영역의 움직임 벡터를 활용하여 카메라의 팬(Pan), 틸트(Tilt), 줌(Zoom)을 분류하였다. 야구 경기장의 잔디와 흙 영역의 색상 정보와 카메라 이동 정보를 통해 샷들을 피칭, 내야, 외야, 선수 클로즈업 등 8개의 유형으로 분류하였다. 타구 이벤트가 발생한 경우 샷에 대해 내야와 외야에 대한 위치를 분류하지만, 타구 위치에 대한 상세한 정보는 제공하지 않았다.

## III. 제안하는 중계 비디오를 활용한 타구 궤적 생성 알고리즘

제안하는 타구 궤적 생성 알고리즘은 그림 4와 같이 샷 단위로 분할된 중계 비디오로부터 피치 샷 검출, 타구 샷 검출, 타구 궤적 생성 단계로 구성된다.

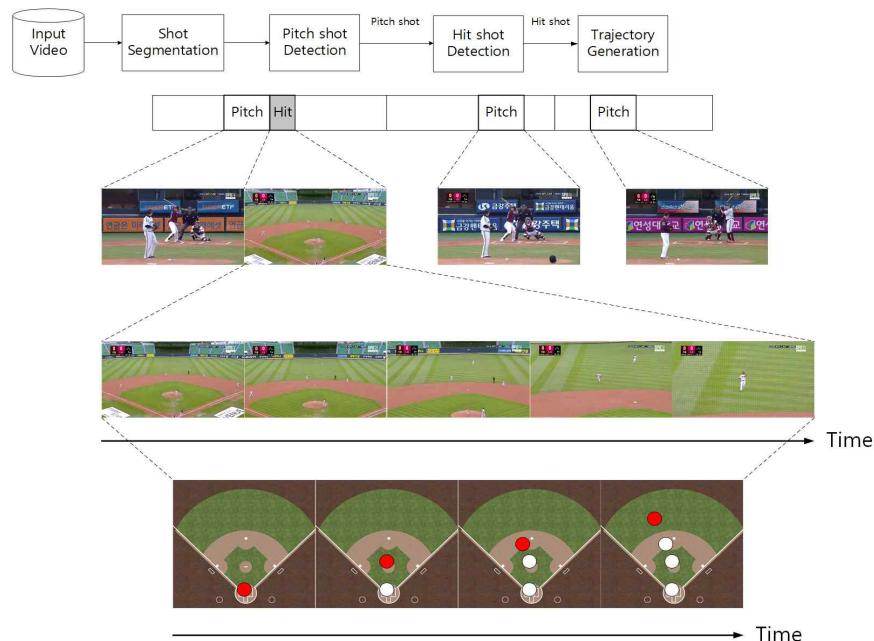


그림 4. 제안하는 중계 비디오 타구 궤적 생성 알고리즘

Fig. 4. Proposed trajectory generation algorithm

## 1. 피치 샷 검출

피치 샷이란 타자에게 투수가 공을 던지는 샷이다. 피치 샷은 다음 세 가지 특성을 갖는다. 첫째, 투수, 타자, 포수, 심판이 특정 위치에 화면에 함께 나타난다. 둘째, 리플레이를 제외한 피치 샷 내 카메라 이동은 없다. 셋째, 투수가 공을 한 번 던질 때 피치 샷의 길이는 15초 이내이다.

피치 샷 검출을 위해 YOLOv3<sup>[12]</sup> 검출기를 사용하며, 사람, 배트, 글러브, 야구공 클래스의 객체는 사용하고, 나머지 검출된 클래스의 객체는 무시한다.

피치 샷 검출 조건은 다음과 같다. 첫째, 하나의 프레임에서 일정 크기의 객체가 동시에 최소 3개 이상 출현해야 한다. 객체 크기는 그림 5와 같이 검출된 사람 객체 바운딩박스의 크기이다. 다음은  $p$ 번째 바운딩박스에 대한 크기 비율 범위이다.

$$th_1 < \frac{w_p h_p}{WH} < th_2, \quad (1)$$

여기에서  $w_p$ 와  $h_p$ 는  $p$ 번째 객체 바운딩박스의 넓이와 높이이며,  $th_1$ 과  $th_2$ 는 객체 바운딩박스의 최소와 최대 크기 비율 문턱치,  $W$ 와  $H$ 는 이미지의 가로와 세로 해상도이다.

둘째, 피치 샷 내 식 (1)을 만족하는 객체들은 프레임 수직 중심선을 기준으로 일정 거리 이내에 각각  $N$ 회 이상 출현해야 한다. 피치 샷에서 타자의 위치 변경을 제외하고 모든 바운딩박스는 특정 위치에서 출현한다.  $p$ 번째 바운딩박스 중심점에서 프레임 수직 중심선까지 거리가  $d_p$ 일 때, 조건은 다음과 같다.

$$d_p < (W/4). \quad (2)$$

셋째, 샷 내에서 배트, 글러브 또는 야구공 객체 클래스가 하나 이상 존재해야 한다. 관중석과 같이 여러 사람이 출현하는 샷은 식 (1), (2)를 모두 만족할 수 있다. 피치 샷은 투수와 타자가 나타나기 때문에 배트, 글러브, 야구공이 적어도 하나 이상 검출되어야 한다.

넷째, 샷 내 카메라 이동이 없어야 한다. 피치 샷을 다시 보여주는 리플레이 샷은 검출 대상에서 제외한다. 리플레이 샷은 피치 샷과 유사하지만, 카메라 이동이 존재한다. 리플레이 샷 구분을 위해 프레임 좌측 하단 200x200 크기 영역의 옵티컬 플로우 크기가 다음 조건을 만족해야 한다.

$$\sum_{i=0}^L \frac{|\overrightarrow{mv_i^{roi}}|}{L} < th_3, \quad (3)$$

여기에서  $\overrightarrow{mv_i^{roi}}$ 는 샷 내  $i$ 번째 프레임에서 설정 영역 중심점의 옵티컬 플로우 크기이며,  $th_3$ 은 옵티컬 플로우 최대 크기 문턱치,  $L$ 은 샷의 총 프레임 수이다.

다섯째, 샷 최대 길이가  $M$ 이하여야 한다. 샷 경계는 영상의 신호 변화에 따라 점진적인 경계와 갑작스러운 경계로 분류된다. 본 논문에서 샷 분할은 컬러 히스토그램 차이와 SSIM(Structural similarity)을 활용하여 검출하였고, 점진적인 변화는 고려하지 않았다. 점진적인 변화를 포함한 샷의 길이가 길어지는 경우를 제외하기 위해 샷 최대 길이는 500프레임으로 설정하였다.

위 다섯 조건들을 모두 만족하는 샷을 피치 샷으로 설정하며, 영상 시퀀스 3개에 대한 파라미터  $th_1$ ,  $th_2$ ,  $N$ ,  $th_3$ 는

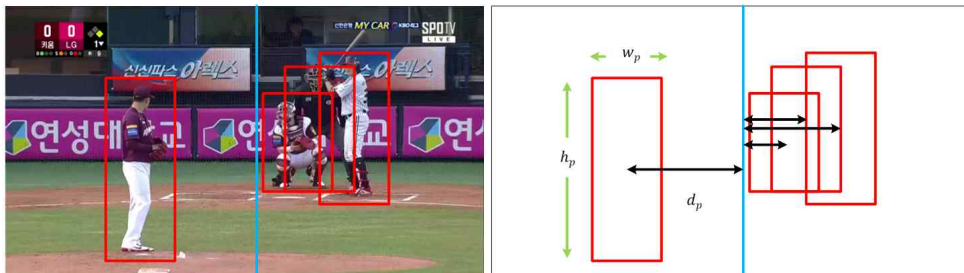


그림 5. 피치 샷 사람 객체 검출  
Fig. 5. Object detection for person class in pitch shot

각각 0.03, 0.15, 30, 15로 설정하였다. 제안 방법은 경기장과 팀이 중복되지 않는 KBO 영상 시퀀스에 대하여 모든 피치 샷과 타구 샷이 포함될 수 있도록 파라미터들을 설정하였다.

## 2. 타구 샷 검출

타구 샷은 고정 위치의 카메라가 공을 추적하는 샷이다. 타구 샷은 다음 두 가지 특성을 갖는다. 첫째, 타구 샷은 피치 샷 직후 나타난다. 둘째, 타구 샷 초기에는 카메라는 줌이 작고 빠르게 회전이동하며, 이후에는 줌이 크고 느리게 회전이동한다.

타구 샷 검출 조건은 다음과 같다. 첫째, 타구 샷 구간  $T$ 에서 크기가 가장 큰 객체 크기는 일정 값보다 작아야 한다.  $T$ 는 샷의 시작 프레임에서 50번째 프레임까지 설정하였으며, 구간 내 크기가 가장 큰 사람 객체는 다음 조건을 만족해야 한다.

$$\frac{w_m h_m}{WH} < th_4, \quad (4)$$

여기에서  $w_m$ 과  $h_m$ 은 샷 구간  $T$ 동안 크기가 가장 큰 객체 바운딩박스의 넓이와 높이이며,  $th_4$ 는 크기가 가장 큰 객체의 최대 크기 문턱치이다.

둘째, 구간  $T$ 에서 샷 옵티컬 플로우 크기가 일정 크기 이상이어야 한다. 샷 옵티컬 플로우를 프레임 옵티컬 플로우와 격자 옵티컬 플로우로 계산한다. 격자 옵티컬 플로우는 그림 6(a)와 같이 직사각형 10x10개의 그리드 각 격자점

에서 옵티컬 플로우를 계산한다<sup>[13]</sup>. 프레임 옵티컬 플로우 크기는 그림 6(b)와 같이 크기 상위  $A$ 개의 격자 옵티컬 플로우 크기 평균으로 정의한다. 객체 움직임을 제외하기 위해 프레임 옵티컬 플로우는 객체 바운딩박스 바깥 격자만 고려한다. 격자 개수를  $C$ ,  $n$ 번째 크기의 옵티컬 플로우를  $\vec{mv}_n$ 라고 할 때, 타구 샷 내  $i$ 번째 프레임에서 내림차순으로 정렬한 옵티컬 플로우 집합  $OF_n$ 는 다음과 같다.

$$OF_n = \{\vec{mv}_0, \vec{mv}_1, \dots, \vec{mv}_C\}. \quad (5)$$

$i$ 번째 프레임의 프레임 옵티컬 플로우는 다음과 같다.

$$FOF_i = \sum_{n=0}^{A-1} \frac{|\vec{mv}_n|}{A}. \quad (6)$$

샷 옵티컬 플로우 크기는 샷 내 구간  $T$ 에 대한 프레임 옵티컬 플로우 크기 평균으로 정의한다.  $k$ 번째 샷의 샷 옵티컬 플로우는 다음과 같다.

$$SOF_k = \sum_{i=0}^{T-1} \frac{FOF_i}{T}. \quad (7)$$

타구 샷은 다음 조건을 만족해야 한다.

$$SOF_k > th_5, \quad (8)$$

여기에서  $th_5$ 는 샷 옵티컬 플로우 최소 크기 문턱치이다. 위 두 조건을 모두 만족하는 피치 샷 직후 샷을 타구 샷으

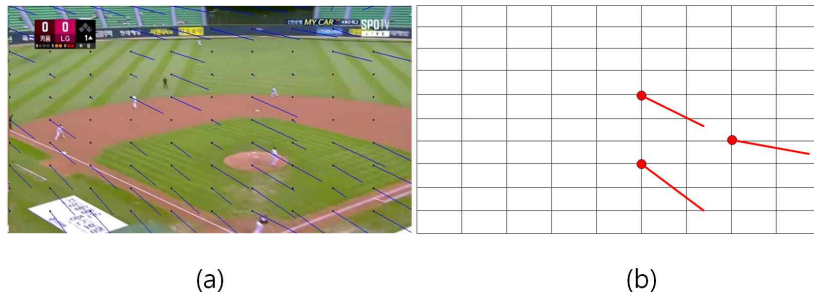


그림 6. 샷 옵티컬 플로우 (a) 격자 옵티컬 플로우, (b) 옵티컬 플로우 크기 상위 세 점(빨간색)  
Fig. 6. Shot optical flow (a) Grid optical flow, (b) Optical flow magnitude top three points(red)



로 설정하며, 영상 시퀀스 3개에 대한 파라미터  $th_4$ ,  $A$ ,  $th_5$ 는 각각 0.1, 3, 0.5로 설정하였다.

### 3. 타구 위치와 궤적 생성

타구 샷 프레임에서 공은 크기가 작고 프레임 밖에 존재할 수 있기 때문에 공의 정확한 위치를 찾기 어렵다. 따라서, 본 논문에서는 공이 타구 샷 각 프레임의 중심점에 위치한다고 가정한다. 타구가 뜬 공일 경우 그림 7(a)와 같이 프레임의 중심과 공의 위치는 다르지만, 그림 7(b)와 같이

샷 끝에는 공이 프레임의 중심과 가까워진다.

타구 위치는 그림 8과 같이 타구 샷 내 프레임들의 중심점을 상공 이미지에 표시하여 생성한다. 상공 이미지는 실측 거리를 계산하기 위해 설정하며, 경기장 규격을 활용하여 생성된다. 기준 프레임은 상공 이미지를 생성하기 위해 경기장 형태가 잘 드러나는 프레임으로 설정한다. 타구 샷 프레임들의 중심점은 타구 샷의 첫 번째 프레임에 표시한다.

상공 이미지 생성은 경기장 규격을 활용한다. 경기장의 본루와 1, 2, 3루 네 점을 꼭지점으로 하는 사각형은 각 변



그림 7. 시간에 따른 중심 영역(빨간색)과 타구 위치 차이, (a) 프레임 밖의 공, (b) 마지막 프레임  
Fig. 7. Difference between center region(red) and ball location over time, (a) Out of frame, (b) last frame

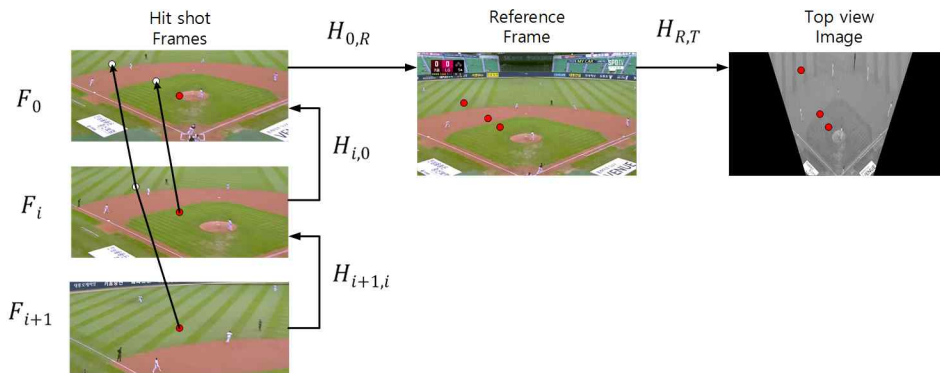


그림 8. 상공 이미지 타구 위치 생성 과정  
Fig. 8. Process for ball location of top view image

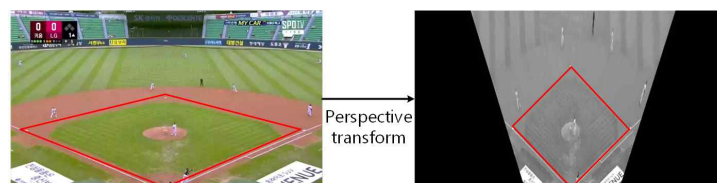


그림 9. 사각형 변환 관계를 활용한 상공 이미지 생성  
Fig. 9. Top view Image generation using rectangle form relationship

이 27.432[m]인 정사각형이다. 타구를 표시하기 위한 상공 이미지는 500x500 해상도와 정사각형 한 변의 길이를 85픽셀로 설정하였다. 기준 프레임은 그림 9와 같이 사각형 간의 변환 관계를 구하여 상공 이미지로 변환될 수 있다.

타구 샷 내  $i$ 번째 프레임을  $F_i$ 라 할 때,  $F_i$ 와  $F_{i+1}$ 은 프레임 내 많은 영역이 겹치며 두 프레임 사이에 특징점 매칭을 통한 호모그래피 변환이 계산될 수 있다.  $i+1$ 번째 프레임의 타구 위치는 프레임의 중심점으로  $(\frac{W}{2}, \frac{H}{2})$ 이며, 이를  $i$ 번째 프레임 상의 좌표로 계산하는 방법은 다음과 같다.  $F_{i+1}$ 에서  $F_i$ 로의 호모그래피 변환 행렬을  $H_{i+1,i}$ ,  $k$ 번째 타구 샷의 길이를  $L_k$ ,  $i$ 번째 프레임 상의  $(i+1)$ 번째 프레임 중심점 위치  $C_i^{i+1}$ 는 다음과 같이 계산한다.

$$C_i^{i+1} = H_{i+1,i} C_{i+1}^{i+1}, i = 0, 1, \dots, (L_k - 1). \quad (9)$$

기준 프레임은 타구 샷의 첫 번째 프레임과 가장 넓은 영역이 겹친다. 따라서, 타구 위치는  $(i+1)$ 번째 프레임부터 타구 샷 첫 번째 프레임까지의 위치 변환을 연속적으로 적용하여 기준 프레임 상의 위치를 계산한다. 기준 프레임을  $F_R$ 이라 할 때, 기준 프레임 상의  $(i+1)$ 번째 프레임 중심점 위치는 다음과 같다.

$$C_R^{i+1} = H_{0,R} H_{1,0} H_{2,1} \dots H_{i+1,i} C_{i+1}^{i+1}. \quad (10)$$

상공 이미지 상의 위치는 상공 이미지 생성 과정에서 계산한 호모그래피 행렬을 활용한다. 상공 이미지를  $F_T$ 라 할 때, 상공 이미지 상의  $(i+1)$ 번째 프레임 중심점 위치는 다음과 같다.

$$C_T^{i+1} = H_{R,T} C_R^{i+1}. \quad (11)$$

식 (10)과 (11)을 활용하여 계산된 타구 샷 프레임들의 중심점들은 하나의 집합으로 다음과 같이 표현한다.

$$P_k = \{C_T^0, C_T^1, \dots, C_T^{L_k-1}\}, \quad (12)$$

본 논문에서는 본루와  $C_T^{L_k-1}$  위치까지 연결한 직선을 타구 궤적으로 설정한다.

본 논문에서는 궤적의 픽셀 단위 거리와 방향을 경기장 규격과 상공 이미지의 비례 관계를 통해 실측 거리를 계산한다. 이미지 상의 정사각형 한 변의 길이가  $B$ 픽셀, 타구 궤적의 직선 길이가  $l$ 일 때, 실측 거리는 다음과 같다.

$$M = l \times \frac{27.432}{B}. \quad (13)$$

## IV. 실험 결과 및 분석

### 1. 실험 환경

본 논문에서는 제안 방법 성능 평가를 위해 1280x720 해상도의 KBO 중계 영상 시퀀스 3개에 대한 실험을 진행하였다. 일반적인 성능 평가를 위해 영상 시퀀스의 경기장 및 팀이 중복되지 않도록 선별하였으며, 각 영상 시퀀스 정보는 표 1과 같다.

### 2. 제안 방식의 PS와 HS의 검출 성능

샷의 경계 변화가 점진적인 샷은 G(Gradual), 갑작스런

표 1. KBO 중계 영상 시퀀스 정보

Table 1. KBO video sequences information

|      | Length  | Game Score | Stadium | Team (vs.) | Game Date | No. of Shots | No. of Pitch Shots | No. of Hit Shots |
|------|---------|------------|---------|------------|-----------|--------------|--------------------|------------------|
| Seq1 | 3:59:58 | 8:7        | Jamsil  | Kiwoom LG  | 2019.8.13 | 3,396        | 338                | 90               |
| Seq2 | 3:04:57 | 6:1        | Sajik   | NC Lotte   | 2019.9.23 | 2,897        | 308                | 87               |
| Seq3 | 3:00:14 | 4:5        | Suwon   | Hanwha KT  | 2019.8.10 | 2,353        | 281                | 69               |

표 2. 제안 방식의 피치 샷과 타구 샷 검출 성능  
Table 2. PS and HS detection performance of proposed method

|            |       | GT  |    | TP  | FP | FN | Precision | Recall |
|------------|-------|-----|----|-----|----|----|-----------|--------|
|            |       | C   | G  |     |    |    |           |        |
| Pitch Shot | Seq1  | 316 | 22 | 332 | 10 | 6  | 97.0      | 98.2   |
|            | Seq2  | 264 | 44 | 260 | 9  | 48 | 96.7      | 84.4   |
|            | Seq3  | 264 | 17 | 257 | 20 | 24 | 92.8      | 91.5   |
|            | Total | 844 | 83 | 849 | 39 | 78 | 95.6      | 91.6   |
| Hit Shot   | Seq1  | 90  | 0  | 85  | 6  | 5  | 93.4      | 93.4   |
|            | Seq2  | 87  | 0  | 77  | 2  | 10 | 97.4      | 88.5   |
|            | Seq3  | 69  | 0  | 60  | 6  | 9  | 90.9      | 87.0   |
|            | Total | 246 | 0  | 222 | 14 | 24 | 94.1      | 90.2   |

샷은 C(Cut)으로 표기한다. 제안 방식의 검출 성능 평가는 GT(Ground truth)와 검출 결과를 비교하여 TP(True Positive), FP(False Positive), FN(False Negative)에 따른 정확도와 검출률을 계산하였으며 표 2와 같다.

샷 분할 과정에서 점진적인 샷 경계를 고려하지 않았기 때문에 점진적인 샷 경계 분포 비율이 가장 높은 Seq2가 가장 많은 FN이 발생하였다. Seq3은 점진적인 샷 경계 분포 비율은 가장 낮지만, 길이가 긴 피치 샷이 존재한다. 제안 방법은 피치 샷 검출 조건에서 샷 길이를 제한하였기 때문에 샷 길이가 긴 피치 샷은 검출하지 못한다.

타구 샷 검출은 피치 샷 직후 샷에 대해서만 고려하기 때문에 피치 샷 검출 결과에 영향을 받는다. 피치 샷 검출에서 발생한 FN이 많을수록 타구 샷 검출에서 FN이 많이 발생하였다.

### 3. 제안 방식의 타구 궤적 생성 성능

타구 궤적 생성 성능 평가는 그림 10과 같이 네이버 스포

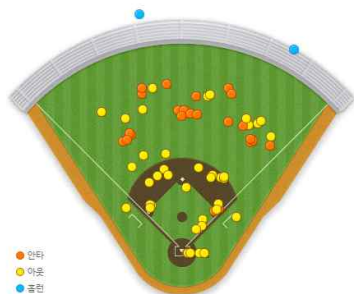


그림 10. 네이버 스포츠 타구 위치  
Fig. 10. Hit ball location on Naver Sports

츠에서 제공하는 타구 기록을 GT로 활용하였다. 타구 기록은 안타, 아웃, 홈런의 위치를 제공한다. 그러나, 제안 방법은 파울을 포함하고 있다. 제안 방법과 타구 기록의 위치를 비교하기 위해 먼저 타구의 유형을 정의하고, 유형에 따라 타구 위치 정확도와 파울 분류 성능을 평가한다.

#### 3.1. 타구 유형 정의

타구 유형은 크게 공의 최종 위치를 기준으로 파울(Foul)과 파울이 아닌 타구(Others) 두 가지로 분류한다. 경기 기록은 파울 위치에서 수비수 실책이나 플라이 아웃이 발생한 위치를 포함한다. 파울은 수비수가 공을 잡으려 하는 FTC (Foul-Try to Catch)와 그렇지 않은 FNTC(Foul-Not Try to Catch)로 분류한다. 파울이 아닌 타구는 송구 유무에 따라 T(Throw)와 NT(Not Throw)로 구분하며, 각 영상 시퀀스의 타구 유형 분포는 표 3과 같다.

표 3. 타구 유형 분포  
Table 3. Ball type distribution

|       | No. of Hit shots | Others |     | Foul |      |
|-------|------------------|--------|-----|------|------|
|       |                  | T      | NT  | FTC  | FNTC |
| Seq1  | 85               | 10     | 48  | 2    | 25   |
| Seq2  | 77               | 22     | 25  | 0    | 30   |
| Seq3  | 60               | 8      | 34  | 4    | 14   |
| Total | 222              | 40     | 107 | 6    | 69   |

#### 3.2. 제안 방식의 타구 위치 정확도

타구 위치 정확도는 송구가 없는 NTG(Not Throw at GT)에 대해서만 고려하며, NTG는 표 3의 NT와 FNTC를 모두 포함한다. 본 논문에서는 타구 기록과 제안 방식의 타구 위치를 비교하기 위해 두 위치의 오차를 측정한다. 오차는 상



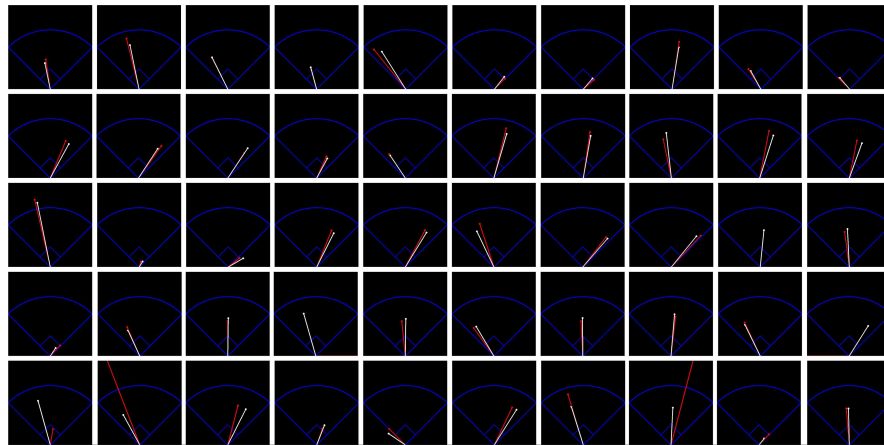


그림 11. NTG 타구 유형 위치 정확도(Seq1), GT(하얀색), 제안 방법(빨간색)

Fig. 11. NTG ball type location accuracy(Seq1), GT(white), proposed method(red)

공 이미지상에서 타구 궤적 직선의 거리 차이와 두 직선 사이의 각도를 계산한다.

그림 11은 Seq1의 NTG 타구 유형에 대한 제안 방법의 위치(빨간색)와 GT 위치(흰색)를 함께 표시한 결과이다. Seq1의 타구 샷 50개를 대상으로 위치를 비교한 결과 오차가 큰 타구 샷이 5개 존재하였다.

오차가 크게 발생하는 원인은 호모그래피 추정 과정에서 오차가 발생했기 때문이다. 프레임 내에 있는 자막이나 CG(Computer graphic)는 카메라의 움직임에 따라 변화하지 않고 고정되어 있으므로 해당 영역에서 매칭된 특징점이 호모그래피 계산에 사용될 경우 카메라 움직임을 파악할 수 없다.

### 3.3. 제안 방식의 파울 분류 성능과 단계별 성능

파울 방향 분류는 FNTC 타구 유형만을 고려한다. 파울은 타구 위치 정확도가 중요하지 않기 때문에 좌우 방향성만을 고려한다.

제안 방법은 식 (13)을 활용해 실측 거리 오차를 계산할 수 있다. 표 4는 제안 방법의 단계별 성능을 나타낸 것이며, 피치, 타구 샷 검출, 타구 위치 정확도, 실측 거리 오차, 파울 방향 분류 각각의 성능을 정리한 표이다. PS(Pitch shot)과 HS(Hit shot) 검출 성능에서 P\_P(Pitch shon precision)는 피치 샷 검출 정확도, P\_R(Pitch shot recall)는 피치 샷 검출률, H\_P(Hit shot precision)는 타구 샷 검출 정확도, H\_R(Hit shot recall)는 타구 샷 검출률을 의미한다.

피치 샷과 타구 샷 검출의 가장 큰 오검출 원인은 점진적인 샷 경계를 고려하지 않는다는 점이다. 따라서, 점진적인 샷 경계를 고려할 수 있는 샷 분할 방법을 사용한다면 더 높은 검출률을 보일 수 있다.

제안 방법은 호모그래피를 연속하여 적용하기 때문에 호모그래피 계산 과정에서 발생한 오차 또한 누적된다. 따라서, 매칭된 특징점 중 카메라 움직임을 반영할 수 있는 특징점을 선택적으로 사용한다면 오차를 줄일 수 있다.

타구 위치의 각도 오차가 낮고 파울 방향 분류 정확도가

표 4. 제안 방법 단계별 성능

Table 4. Proposed method step by step performance

|       | PS & HS detection |      |      |      | Location accuracy |       | Foul direction classification accuracy |
|-------|-------------------|------|------|------|-------------------|-------|----------------------------------------|
|       | P_P               | P_R  | H_P  | H_R  | Distance [m]      | Angle |                                        |
| Seq1  | 97.0              | 98.2 | 93.4 | 93.4 | 12.2              | 8.9   | 100.0                                  |
| Seq2  | 96.7              | 84.4 | 97.4 | 88.5 | 21.2              | 3.6   | 100.0                                  |
| Seq3  | 92.8              | 91.5 | 90.9 | 87.0 | 10.3              | 8.2   | 92.6                                   |
| Total | 95.6              | 91.6 | 94.1 | 90.2 | 13.6              | 7.5   | 98.6                                   |

높은 것을 통해 타구가 이동한 방향에 대한 정확도는 높다는 것을 알 수 있다.

경기장 본루에서 중앙 펜스까지 평균 거리는 약 120[m]이며, 좌우 펜스까지 평균 거리는 약 100[m]이다. 제안 방법의 타구 위치 생성은 약 13.6[m]의 거리 오차를 가지기 때문에 정확한 타구 위치를 생성하는 것은 아니다. 그러나, 제안 방법은 중계 비디오만을 활용하기 때문에 비용이 적고 자동화된 방법으로 경기 분석에 활용할 수 있다.

## V. 결론 및 향후 연구

본 논문에서는 야구 중계 비디오에서 피치, 타구 샷을 검출하고, 타구 궤적을 생성하는 알고리즘을 제안하였다. 제안 방법은 샷 단위 객체 검출과 옵티컬 플로우를 활용하여 피치 샷과 타구 샷을 검출하였다. 타구 샷 프레임의 중심점에 공이 위치한다고 가정하고, 각 프레임의 중심점을 상공 이미지에 표시하여 타구 궤적을 생성하였다. 제안 방법은 피치 샷과 타구 샷 검출, 타구 위치 정확도, 파울 분류에 대해서 평가를 수행하였다. 평균 거리 오차는 10-15%의 오차율로 상용 시스템에 비해 높지만, 각도 오차가 작고 파울 방향 분류 정확도가 높은 것으로 미루어 보아 타구의 방향은 비교적 정확하게 계산할 수 있다.

제안 방법은 호모그래피 생성시 발생 오차가 누적되어 타구 궤적 정확도에 영향을 끼친다. 이를 해결하기 위해 비디오 프레임 내 카메라 움직임을 반영할 수 있는 특징점을 선택적으로 사용하는 호모그래피 생성 방법에 대한 추가 연구가 필요하다.

## 참 고 문 헌 (References)

- [1] Major league baseball statcast leaderboard, <http://m.mlb.com/statcast/leaderboard#exit-velocity>, 2019 (accessed Feb. 11, 2020).
- [2] H.-C. Shih, "A survey of content-aware video analysis for sports," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.28, No.5, pp. 1212-1231, 2017.
- [3] C.-H. Liang, et al., "Baseball event detection using game-specific feature sets and rules," *IEEE International Symposium on Circuits and Systems*, pp. 3829-3832, 2005.
- [4] Y.-F. Huang and L.-H. Tung, "Semantic scene detection system for baseball videos," *International Journal of Digital Content Technology and Its Applications (JDCTA)*, Vol.5, No.9, 2011.
- [5] S. Tappaya, et al., "Multi-modal visual features-based video shot boundary detection," *IEEE Access*, Vol.5, pp. 12563-12575, 2017.
- [6] Z. Rasheed, and M. Shah, "Scene detection in Hollywood movies and TV shows," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol.2, pp. 11-343, 2003.
- [7] I. U. Haq, et al., "Movie scene segmentation using object detection and set theory," *International Journal of Distributed Sensor Networks*, Vol.15, No.6, 2019.
- [8] Z.-M. Lu and Y. Shi, "Fast video shot boundary detection based on SVD pattern matching," *IEEE Transactions on Image processing*, Vol.22, No.12, pp. 5136-5145, 2013.
- [9] J. Sun and Y. Wan, "A novel metric for efficient video shot boundary detection," *IEEE Visual Communications and Image Processing Conference*, pp. 45-48, 2014.
- [10] M.-H. Hung and C.-H. Hsieh, "Event detection of broadcast baseball videos," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.18, No.12, pp. 1713-1726, 2008.
- [11] P. Chang, M. Han and Y. Gong, "Extract highlights from baseball game video with hidden Markov models," *Proceedings. International Conference on Image Processing*, Rochester, NY, USA, 2002.
- [12] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," arXiv preprint arXiv:1804.02767, 2018.
- [13] G. Farneback, "Two-frame motion estimation based on polynomial expansion," *Scandinavian conference on Image analysis*, pp. 363-370, 2003.

---

저 자 소 개

---



**손 종 응**

- 2018년 2월 : 한국항공대학교 항공전자정보공학부 학사
- 2020년 2월 : 한국항공대학교 항공전자정보공학과 석사
- ORCID : <https://orcid.org/0000-0001-5875-4145>
- 주관심분야 : 영상처리, 임베디드시스템



**이 명 진**

- 2001년 8월 : KAIST 전자전산학부 박사
- 2001년 3월 ~ 2004년 2월 : 삼성전자 System LSI 사업부 책임
- 2004년 3월 ~ 2007년 2월 : 경성대학교 전기전자공학전공 조교수
- 2007년 3월 ~ 현재 : 한국항공대학교 항공전자정보공학부 교수
- ORCID : <https://orcid.org/0000-0002-3136-2819>
- 주관심분야 : 영상통신, 영상처리, 임베디드시스템