

# F-058 野球ゲームにおける動画タギングを用いたハイライト生成 Highlight Generation Using Video Tagging of Template Matching and OCR in Baseball TV Game

赤木 信也\*

Shinya Akagi

\*Email: shinyaresearchresource@gmail.com

## 1. はじめに

ハイライト生成は、長時間のスポーツ動画などから盛り上がった箇所を抽出し、視聴者が短い時間で要点を理解できるようにする動画要約手法であり、ニュースなどにおいては人手で実施されている動画要約の自動化が検討されていたり、ハードディスクレコーダーやアプリにおいては自動要約機能を持つ製品が提供されたりしている[1]。しかしながら、多様な動画から多様な盛り上がり箇所を精緻に抽出できる技術は確立しておらず、スポーツなら各スポーツごとに最適化されたハイライト生成手法を研究することが主流となっている。

ハイライト生成の過去研究としては、音声や画像を利用したり[1][2]、スローモーションリプレイを利用したり[3]、視聴者のコメントを利用したりする手法[1][4][5]が提案されている。

著者の過去研究では、ハイライト生成の前段階として、盛り上がり検出に取り組んだ。最初の研究では、YouTubeの切り抜き動画を対象とした盛り上がり箇所検出実験を実施し、音声振幅やコメントを用いた方法では、盛り上がり箇所をピンポイントで推定できた結果を上位10箇所を含む場合が10件中1件程度しかなく、検出精度の低さが課題として挙げられた。特にプロ野球の練習試合動画については、音声振幅を用いた方法では得点シーンやホームラン箇所の推定に失敗しており、個別に最適化された方法を試した方が良いという結果が得られた[6]。その次の研究では、プロ野球ゲームを対象として画像解析によって盛り上がり箇所(得点シーン、ホームラン)を検出できないか分析し、テンプレートマッチングを利用することで高い精度で盛り上がり箇所を検出できることが示唆された[7]。その次の研究では、盛り上がり箇所(ピンチ、チャンス)の検出にスコア情報を活用するため、光学文字認識(OCR)の精度を分析し、OCR製品の認識精度が83%~95%となり、精度的な課題は残るものの実用可能であることが示唆された[8]。その次の研究では、数字2文字のOCRについてCNNベースの手法で精度向上ができないか検討し、認識精度89%~100%を達成できた[9]。

本研究では、著者の過去研究を発展させ、プロ野球ゲームを対象としてテンプレートマッチングとOCRを用いた動画タギング(ボール、ストライク、アウト、ランナー、スコア情報、ホームラン)を行い、そのタグに基づいてハイライト生成(打球間隔などの場面の切り分けと盛り上がり箇所の結合)を行った結果について報告する。

## 2. 研究の技術領域

本研究の最終的な目的は、ハイライト動画の作成を支援または自動化することにある。本研究の技術領域は下記の通りいくつか分かちされており、今回の研究は場面の切り分け・結合の技術領域である。

- (1) 盛り上がり箇所の検出
- (2) 場面の切り分け・結合
- (3) タイトル・説明文の作成

## 3. 動画タギングの自動化と想定利用者

テンプレートマッチングやOCRを用いた動画タギングでは、特徴画像の位置特定をする必要がある。位置特定については自動抽出手法が過去に研究されているが[10]、精度が90%前後であり、99%以上なら採用の利点はあると考えるが、本研究では、決め打ちやリハーサルによって事前に位置特定できるものとし、位置特定の精度も100%になることを前提とする。ただし、その分、決め打ちやリハーサル時の作業効率化の課題が発生するため、最終的には位置特定も自動化したいと考えている。

動画タギングの自動化については、実際のプロ野球などではスコア表示時などに人が介入する必要があるため、その時点で連動してタグを取得すれば動画タギングの自動化を実施する必要がないという指摘が可能である。しかし、本研究で想定している利用者は、動画の視聴者やe-sportsの開催者のようなプロ野球ゲームにおいて柔軟性のあるハイライト生成を実現したい人であり、スコア表示時などに連動してタグを取得できないことが前提となっている。

動画タギングの自動化については、ゲーム内ハイライトを利用すれば良いという指摘も可能だが、実況映像と実況音声を含めたハイライトを生成したい場合はゲーム内ハイライトでは対応できず、柔軟性のあるハイライト生成も実施できないため、本手法を採用する意義はあると考える。

## 4. ハイライト生成の解析手法

ハイライト生成の解析手法は、元動画が再生中の段階で解析する「リアルタイム処理」と元動画が作成済みの段階で解析する「バッチ処理」に分けられる。本研究で採用しているテンプレートマッチング

手法は、バッチ処理では時間がかかりすぎるため[7][8]、実用性を意識する場合は時系列データの分析を極力減らしつつ、リアルタイム処理を実施できるようにする必要がある。

## 5. プロ野球ゲームの特徴

研究対象としては「実況パワフルプロ野球2022」を使用する。プロ野球ゲームの特徴としては、打球・打撃時にストライク、ボール、アウトのカウント、ランナーの有無、得点情報(以下、通常時スコア)が固定の位置に表示されたり、得点時に固定の位置に得点情報(以下、得点時スコア)が表示されたり、ホームラン時やゲームセット時に画面中心に得点情報(以下、ホームランとゲームセット時スコア)が表示されたりする。これらの特徴を利用して、テンプレートマッチングとOCRを用いてピンチやチャンス、得点シーン、ホームラン、ゲームセットを検出する。

また、打球後のインプレー時には通常時スコアが表示されなくなり、インプレー切り替わり後に再度通常スコアが表示されるようになるという特徴があり、かつ、インプレー切り替わり後の打球開始はある程度時間が決まっているため、得点時スコアなどによってインプレーの終了時間さえ分かれば、打球開始からインプレー終了までの場面切り分けが実現できるようになっている。

## 6. 動画タギングの方法

動画タギングでは、ストライク、ボール、アウト、ランナーの状態と、得点、ホームラン、ゲームセットのシーンをテンプレートマッチングで取得し、通常時スコアと得点時スコアをOCRで取得した。テンプレートマッチングとOCRの位置特定は、打球・打撃シーン、得点シーン、ゲームセットシーンの3つの場面の時間を予め特定しておき、それぞれの位置をドラッグ&ドロップで選択できるようにした。3つの場面の時間を予め特定することは手間がかかるが、ドラッグ&ドロップで選択できるようにすることで、位置情報を数値として認識する必要がなく、位置特定の作業効率を向上させることができている。

## 7. 場面の切り分けの方法

場面の切り分けは、ストライク、ボール、アウトのカウントに変化があった時、三振アウトでチェンジになった時、ストライクなどが未表示の画面からストライクなどが表示されている画面に切り替わった時に、場面が切り替わったとして場面の切り分け(シーン分割)を行った。切り替わり直後は、カウントが増加するタイミングであり、前のシーンに関係があるとともに、次の打球開始タイミングを調整するために、切り替わり直後の1秒分を前のシーンに追加している。ヒットやチェンジなどの場合は、打球後から全て同じ場面として切り分けられるため、元々冗長な切り分けとなるのに加えて、更に余分な1秒が追加されることになり、場面結合時に終了タイミングを上手く調整する必要がある。なお、機械学習等を用いてインフィールド、アウトフィールド、ゾーンなどのシーン分類を実施できるようにすれば、ヒットなどの終了タイミングを自動で調整できるようにするが、今回は、盛り上がり箇所として想定される得点シーンやホームランの終了タイミングをテンプレートマッチングでピンポイントで取得できるため、シーン分類は利用していない。

## 8. 盛り上がり箇所の重み付けの方法

盛り上がり箇所として、勝ち越し・逆転シーン、同点シーン、ホームランシーン、接戦時のピンチチャンスシーン(ランナー:満塁、ランナー:二・三塁)に対して重み付けを実施した。得点シーンとホームランシーンの重みは1.0とし、同点の場合は+1.0、勝ち越し・逆転の場合は+2.0を付与した。先制点シーンは勝ち越しに含まれるため、+2.0になる。接戦時のピンチチャンスシーンの重みは1.5とした。

本重み付けは暫定的なものであり、更に細かい条件の重み付けを行えば、より柔軟なハイライト生成が実現できるようになる。

## 9. 結合の方法

重み付けを行ったシーンについて、利用者が指定した時間に収まるように動画を結合する。結合処理にはmoviepyを用いた。シーンごとに重みが付与されており、重みが大きい順に利用者が指定した動画時間を超えるまでシーンを選択していき、選択されたシーン全てを時間順で結合する。

## 10. 開発システム

開発したシステムの全体像を図1に示す。

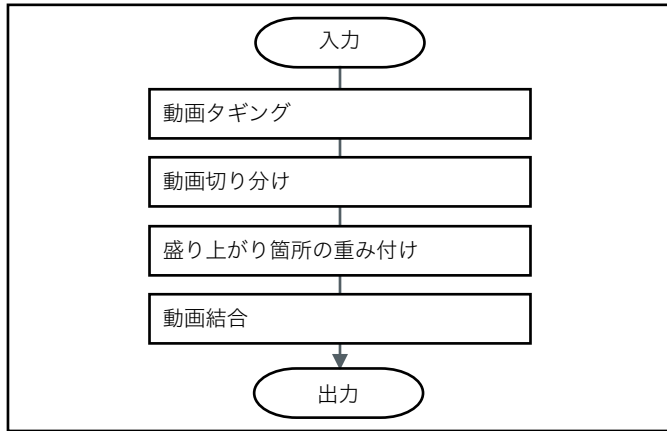


図1 ハイライト生成システム

## 11. 実験方法

実況パワフルプロ野球2022の対戦動画(1時間36秒)において、動画タギングを実施後、3分を指定してハイライト動画を自動生成した。

## 12. 実験結果

2分56秒のハイライト動画が生成され、処理時間は3分18秒かかった。生成された動画に対する評価を著者1名の主観評価によって1~5の5段階で評価を行なった。結果を以下に示す。

| 評価項目                | 評価 |
|---------------------|----|
| 動画タギングの準備が簡便か       | 2  |
| 動画タギングの処理時間が長くないか   | 2  |
| 動画の切り分け・結合に不自然さはないか | 4  |
| 盛り上がり箇所が適切に抽出されているか | 5  |
| 動画結合の処理時間が長くないか     | 5  |

動画タギングの準備が簡便かについては、テンプレートマッチングとOCRの位置特定をする必要があり、ドラッグ&ドロップで選択できる機能を開発しているものの作業の負担は大きく、自動位置特定ができるようになれば良いと考える。

動画タギングの処理時間が長くないかについては、過去研究において並列処理を使用しても相当数の時間がかかることが判明しており、リアルタイム処理が実装できれば改善の目処が立つと考える。

動画の切り分け・結合に不自然さはないかについては、急に始まったり、急に終わったりすることが完全にはない訳ではないが、不自然さは少なかったと考える。なお、ゲームセット場面は取得できているが、ゲーム開始場面も取得できていると良かったと考える。

盛り上がり箇所が適切に抽出されているかについては、同点シーンや勝ち越しシーンに加えて、接戦時のピンチチャンスシーンも取得できており、両チームの盛り上がりを満遍なく取得できていたと考える。動画結合の処理時間が長くないかについては、動画タギングさえ完了していれば、ハイライト生成自体にはそれほど時間がかからなかったと考える。

## 13. 考察

本野球ゲームでは、投球後のインプレー時には通常時スコアが表示されなくなり、インプレー切り替わり後に再度通常スコアが表示されるようになるという特徴があり、かつ、インプレー切り替わり後の投球開始はある程度時間が決まっているため、盛り上がり箇所の切り分けにおける投球開始タイミングを正確に取得できていた。また、得点シーンやホームランといった盛り上がり箇所では、スコア情報が表示されるという特徴があり、テンプレートマッチングでピンポイントでタグ付けできていたことにより、盛り上がり箇所の切り分けにおける終了

タイミングを正確に取得できていた。この条件は実際のプロ野球動画では当てはまらないため、実際のプロ野球動画に応用する場合は、シーン分類等を用いて投球開始タイミングとプレー終了タイミングを取得する必要がある。ただし、実際のプロ野球動画において表示デザインを本野球ゲームに合わせることであれば、本提案手法をそのまま応用することができるようになるため、実際のプロ野球動画においても機械判別可能な表示デザインの採用が進むと良いと考える。しかしながら、実際のプロ野球動画におけるハイライト生成では、スコア表示時などに人が介入する必要があり、その時点で連動してタグを取得すれば機械的な動画タギングの自動化を実施する必要がないため、視聴者が動画要約を実施する場合を除いて、本提案手法を利用する機会は少ないと考える。

## 14. 今後の課題

今回は乱打線だったため、盛り上がり箇所が多く抽出できていたが、投手戦の場合は得点シーンがほとんどなく、ピンチチャンスも少なからず、盛り上がり箇所がほとんど抽出できず分量が足りなくなる恐れがある。また、ピンチチャンスにおいて三振した場合を抽出できなかったり、牽制や盗塁を適切に抽出できなかったり、いくつか未対応の事象が存在している。三振のタグを取得して奪三振ショーに対応したり、盗塁のタグを取得して走塁面の活躍も取得できるようにしたりする必要があると考える。

## 15. まとめ

プロ野球ゲームにおいて、ストライク、ボール、アウト、ランナーの状態、得点、ホームラン、ゲームセットシーンのテンプレートマッチングとスコア情報のOCRを用いて動画タギングを実施することで、動画の切り分け・結合が自然、かつ、盛り上がり箇所を適切に抽出したハイライトを自動生成できることが示された。検出可能な盛り上がり箇所の拡充と柔軟な重み付けの実施、動画タギングの高速化またはリアルタイム処理対応、テンプレートマッチングやOCRの位置特定の自動化または更なる効率化、OCRの精度向上が今後の課題だと考える。

## 16. おわりに

本研究で利用したプログラムなどはGithub[11]およびGoogleドライブ[12][13]に公開している。

## 参考文献

- [1] 望月貴裕:映像自動要約技術の最新動向, NHK技研R&D2020年夏号, 4-15, 2020-08-15
- [2] Y Rui, A Gupta, A Acero: Automatically extracting highlights for TV baseball programs, Proceedings of the eighth ACM international conference on Multimedia, pp. 105-115, 2000
- [3] Po-Chyi Su, Chi-Heng Lan, Chin-Song Wu, Zi-Xin Zeng, Wei-Yu Chen: Transition effect detection for extracting highlights in baseball videos, EURASIP Journal on Image and Video Processing volume 2013, 27, 2013-05-04
- [4] 富田大志, 道満恵介, 井手一郎, 出口大輔, 村瀬洋: スポーツハイライト映像作成のためのTwitter熱狂度に基づいたイベント検出, 電子情報通信学会技術研究報告. MVE, マルチメディア・仮想環境基礎, 112(474), pp. 233-238, 2013-03
- [5] tf0101: YouTube Liveの生放送から盛り上がった箇所を自動抽出するCLIを作った話, <https://qiita.com/tf0101/items/efb2484a0b5b1cdc8291>, 2019-07-08
- [6] 赤木信也: 盛り上がり検出のための音声解析の一考察, 情報科学技術フォーラム講演論文集(FIT), 22巻, 2号, pp. 423-424, 2023-08-23
- [7] 赤木信也: 野球ゲームにおける盛り上がり箇所の自動検出—画像解析の一考察—, 情報処理学会全国大会講演論文集, 86巻, 2号, pp. 59-60, 2024-03-01
- [8] 赤木信也: 野球ゲームにおける盛り上がり検出の自動検出—光学文字認識の一考察—, 情報科学技術フォーラム講演論文集(FIT), 2024-09-05
- [9] 赤木信也: 野球ゲームにおける数字2文字の光学文字認識, 情報処理学会全国大会講演論文集, 2025-03
- [10] K Jung, KI Kim, AK Jain: Text information extraction in images and video: a survey, Pattern Recognition, volume 37, Issue 5, pp. 977-997, May 2004
- [11] Github: create\_baseball\_highlightリポジトリ, [https://github.com/ShinyaAkagiI/create\\_baseball\\_highlight](https://github.com/ShinyaAkagiI/create_baseball_highlight)
- [12] Googleドライブ: output.mp4, <https://drive.google.com/file/d/1sJ0ygXSM0hBk9mCsJXWaqBgZ7Zq-A-jj/view?usp=sharing>
- [13] Googleドライブ: output\_highlight.mp4, [https://drive.google.com/file/d/1xgm0L2CbTRufv-xI-PhJ5EgMwbWEF7gi/view?usp=drive\\_link](https://drive.google.com/file/d/1xgm0L2CbTRufv-xI-PhJ5EgMwbWEF7gi/view?usp=drive_link)