

スポーツニュースの言語情報を利用した プロ野球映像推薦システムの試作

Scene Recommendation System for Professional Baseball using Linguistic Information of Sports News

宮崎 勝 † 住吉 英樹 † 後藤 淳 † 藤井 真人 † 柴田 正啓 †
Masaru Miyazaki Hideki Sumiyoshi Jun Goto Mahito Fujii Masahiro Shibata

1 はじめに

スポーツニュースでは、その日に行われた野球の試合をダイジェストで紹介するのが一般的だが、ダイジェスト映像に含まれる各ハイライトシーンは極めて短時間であるため、そのシーンの前後を含め詳しく見たいという要望が考えられる。そこで、スポーツニュース中のダイジェスト映像のアナウンスコメントを解析して野球のスコアデータと比較することにより、ダイジェストに含まれるハイライトシーンを、アーカイブされた同一試合の中継映像から自動抽出する技術を開発し、検索型コンテンツ視聴システムのプロ野球シーン検索機能として実装した。その際、ハイライトシーンの特徴を制約として利用することで、不完全な情報しか含まれないアナウンスコメントから試合の状況を推定し、精度良くシーンを抽出できることを確認した。

2 CurioView システム

筆者らは視聴者の興味を自然に広げてくれる検索型コンテンツ視聴スタイル CurioView[1] を提案している。テレビというインターフェースの特徴を最大限に活かし、視聴者が現在見ている番組に関連した他の番組や映像をさりげなく検索し、推薦してくれる機能を持つ。CurioView システムは大きく分けて、メタデータの付与・管理を行うメタデータサーバ、関連コンテンツを検索する関連性検索エンジン、関連コンテンツを表示する表示システムから構成されており、検索対象コンテンツや視聴形態に応じて関連性検索エンジンや表示システムを切り替えて運用できる。今回、スポーツニュースのダイジェスト映像のアナウンスコメントと実際の試合状況を記述したスコアデータとの関連性検索を行うエンジンを試作し、Web ブラウザ上に表示する CurioView システムにプロ野球映像推薦機能として実装した。CurioView システムの構成を図 1 に示す。

3 アナウンスコメントを利用したシーン検索

アナウンスコメントの言語情報および試合のスコアデータを解析することにより、ダイジェスト映像に含まれるハイライトシーンを中継映像から検索する手法について述べる。

3.1 試合メタデータ

野球の試合状況を詳細に記述したスコアデータはスポーツデータ配信会社によってリアルタイムに作成され、放送局における番組制作（野球中継時のスコア情報表示）やインターネットのポータルサイトにおける情報提供などに利用されている。今回、このスコアデータを

独自形式の試合メタデータに変換し、1 試合 1 ファイルの XML ファイル形式で蓄積した。この試合メタデータはピッチャーの 1 投球をシーンの最小単位 (sceneUnit) としており、sceneUnit 単位で得点、SBO カウント、進塁状況、投手、打者名、映像ファイル名、および開始点のタイムコード情報などが記述されている。

3.2 インニング番号および選手名によるシーン同定

ダイジェスト映像におけるアナウンスコメントには、インニング番号および選手名が含まれることが多い。そこで、あらかじめスコアデータから選手名のリストを抽出し、アナウンスコメント中にインニング番号および選手名が出現した場合にシーン検索処理を行った。アナウンスコメントは 1 文単位で処理され、各文にインニング番号を示す「回」という表現、および選手名が出現したときに検索処理を行う。どちらか一方だけが出現した場合には前の情報を引き継いで検索処理を行った。インニング番号と選手名を利用したシーン検索処理を図 2 に示す。

この関連性検索エンジンは、検索キーに対応した

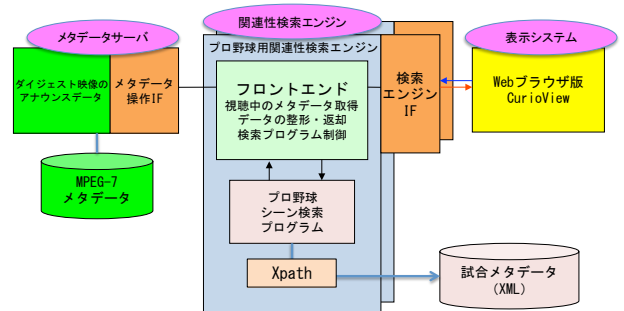


図 1 CurioView システムの構成

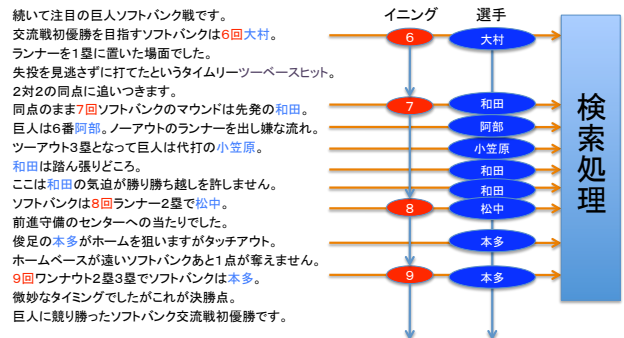


図 2 インニング番号および選手名をキーとしたシーン検索

† NHK 放送技術研究所 (人間・情報科学)

Xpath を自動生成し、試合メタデータに適用することによって該当 sceneUnit を抽出する。6 回の大村選手の打撃を検索する Xpath の例を図 3 に示す。

```

/game/sceneUnit[substring(./@id,1,2)=" 6" ]
[situation/batter/name=" 大村 "]/@id

```

図 3 インニング番号および選手名をキーとした検索 Xpath の例

4 ハイライトシーンに関する制約の利用

ハイライトシーンの特徴を制約として結果をフィルタリングすることにより、シーン検索の精度を高めることができる。以下にその制約の種類とその内容を示す。

- 最終投球制約
ハイライトシーンとして選ばれるのは、必ず打席の最後の投球になっている。そこで、打席の最後の投球のみを検索するよう、生成される Xpath に制約を加える。
- 投手除外制約
通常ある打席では同じ投手が投げ続けることが多いため、投手名での検索ではシーンの特定が困難である。そこで、アナウンスコメント中に現れる選手名が投手名であった場合には検索結果を除外する。
- 打順制約
「俊足の本多がホームを狙いますがタッチアウト」といったランナーの名前が含まれるコメントの場合、そのランナーが打席に立っていたときの打撃が検索されてしまう。ある打撃が検索された次にその前打者の打撃が検索された場合には、打順を考慮して前打者の打撃を除外する。

5 ハイライトシーン検索実験

実際の試合データからスポーツニュースで取り上げられたハイライトシーンを抽出する実験を行った。日本のプロ野球の試合 6 試合を対象として提案手法によるハイライトシーン検索を行った結果を表 1 に示す。

表 1 ハイライトシーン検索実験結果 (6 試合平均)

ダイジェストに含まれる ハイライトシーン数	適合率	再現率
3.83	1.00	0.69

すべての試合で良好な適合率を得ることができたが、特定のアナウンスコメントのパターンでは正しい該当ハイライトシーンを抽出することができず、適合率が下がっている。それらはすべて投手の好投シーンであり、「この後は安定感のあるリリーフ神がしっかりと仕事を果たしました」「得点を与えずしっかり役割を果たします」といったように、シーンの情報がまったくアナウンスコメントに現れないものであった。このようなシーンは言語情報のみから特定することが不可能であり、検索には映像の類似度などを用いた手法 [2] などを併用する必要がある。

6 CurioView システムへの実装

開発したシーン検索エンジンの CurioView システムへの実装を行った。検索結果表示画面を図 4 に示す。

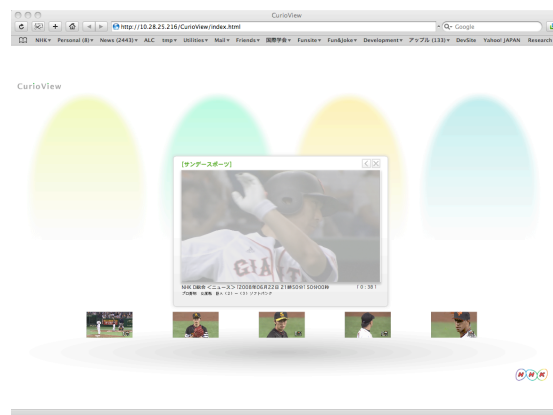


図 4 CurioView システムにおけるプロ野球映像推薦画面

中央の大きな映像が現在視聴しているスポーツニュースであり、ある特定試合のダイジェスト映像に入ったところで検索処理が行われ、得られた sceneUnit に対応した試合映像がサムネイル表示される。視聴者はその興味に応じてサムネイル画像をクリックすることにより、アーカイブされた中継映像の該当ハイライトシーンを視聴することが可能である。

7 おわりに

視聴しているスポーツニュースで紹介されたハイライトシーンを中継映像から自動的に検索し、推薦するシステムの試作について述べた。今回はダイジェスト映像に含まれるシーンと同じシーンを検索する技術を提案したが、ハイライトシーンの前後のシーンや、ハイライトシーンの状況を作るきっかけとなったシーンなどを検索することでよりユーザの興味や理解を拡大するシステムが実現できる。筆者らは野球オンтоロジーを用いたハイライトシーン抽出技術を提案しており [3]、今後はその技術を応用し、関連した他のハイライトシーンの抽出・提示といったより高度なコンテンツ検索・推薦技術の検討を進めていく。

参考文献

- [1] 柴田ほか. 検索技術を使う新しいテレビ視聴スタイル curioview. FIT2008, 第 7 回情報科学技術フォーラム, Vol. 3, No. H-007, pp. 77-78, 2008.
- [2] 望月ほか. スポーツニュース映像をクエリーとした野球放送映像からの同一シーン探索. FIT2009, 第 8 回情報科学技術フォーラム, No. 4J-007, 2009.
- [3] Masaru Miyazaki, Masahiro Shibata, and Nobuyuki Yagi. Baseball digest production system using inductive logic programming. In *ISM '08: Proceedings of the 2008 Tenth IEEE International Symposium on Multimedia*, pp. 503-508, Washington, DC, USA, 2008. IEEE Computer Society.