

# スポーツ映像におけるシーン重要度算出アルゴリズムとその評価

橋本 隆子, 加登岡 隆, 飯沢 篤志

(株)リコー ソフトウェア研究開発本部 {takako, katooka, izw}@src.rioh.co.jp

要約: 放送のデジタル化やモバイル端末の普及により, 多様なコンテンツを低コストで作成する仕組みが必要となっている。我々はこれまで, 映像メタデータと利用者の嗜好情報を利用したパーソナルダイジェスト作成方式(Personal Digest Making Scheme)の提案を行ってきた。PDMS は, スポーツ映像を対象とし, 映像メタデータから発生事象の重要度を自動的に算出し, 複数の重要場面(ダイジェストシーン)を選択するものであり, 従来の手動によるコンテンツ作成に比べ低コストでダイジェストを生成することができる。我々はPDMSを利用して, 携帯端末向けにプロ野球ダイジェスト映像を配信するシステムを構築し, 2002年8月より商用サービスを開始している。開始以来, 既に170余の試合に関する野球ダイジェストの配信を行ってきた。本稿では, この配信結果に基づき, 野球ダイジェスト配信システムの重要度算出アルゴリズムの評価を行う。評価は番組制作ディレクターが手動で選択したダイジェストシーンを正解集合として, 我々のシステムが自動的に選択したダイジェストシーンと比較し, 適合率を求め, 我々の方式が正しいダイジェストシーンを選択しているかどうかを判定することにより行う。評価結果を基に, スポーツ映像における重要事象の解析方法について再検討し, シーン重要度算出アルゴリズムの改良を行う。改良したアルゴリズムに対しても, 適合率を再度求め, 改良の効果を評価する。

## 1. はじめに

ブロードバンド時代に入り, 高機能携帯端末向け動画配信サービスが普及してきた。また2003年12月には, 地上波放送のデジタル化も予定されており, テレビ端末のみならず, 自動車や携帯端末などの移動体向け放送も可能となる。このような状況においては, 多様なコンテンツを低コストで生成配信する仕組みが必要となってくる。

我々はこれまでパーソナルダイジェスト作成方式(Personal Digest Making Scheme, PDMS)<sup>1)2)3)4)</sup>の提案およびシステムの試作を行ってきた。PDMSはスポーツ映像を対象とし, 映像メタデータと利用者の嗜好情報から, 映像上で発生した事象の重要度をルール記述に従って算出し, 重要度の高い順にシーンを選択し, ダイジェストを作成する方式である。我々は, PDMSを利用して, 携帯端末向けにプロ野球ダイジェスト映像を配信するシステム(野球ダイジェスト配信システム)<sup>5)6)</sup>を構築し, これまでにプロ野球パ・リーグの170試合余に関して, ダイジェスト映像を配信してきた。本稿では, この配信結果を基に, 我々が構築した野球ダイジェスト配信システムの重要度算出アルゴリズムの評価を行う。評価結果を元に, スポーツ映像における重要事象の解析方法について再検討し, シーン重要度算出アルゴリズムの改良を提案する。改良アルゴリズムの適合率を再度求め, 改良の効果も評価する。

本稿は以下の構成になっている。第2章でこれまで我々が提案してきたPDMSの概要について述べる。第3章で, PDMSを利用した野球ダイジェスト配信システムの重要度算出アルゴリズムについて述べ, アルゴリズムの評価を行う。なおPDMSの詳細については文献<sup>3)</sup>に, 野球ダイジェスト配信システムの

概要については文献<sup>3)</sup>に述べているので, 詳しくはそちらを参照していただきたい。第4章で, スポーツ映像向け重要度算出アルゴリズムの改良について検討し, 改良後の評価結果について述べる。

## 2. PDMS の概要

本章では, 我々が従来提案してきたパーソナルダイジェスト作成方式(PDMS)の概要について述べる。

### 2.1 PDMS の入力データ

PDMSの入力データは, ビデオ映像, 映像メタデータ(以下メタデータ), 利用者の嗜好情報の3種類である。メタデータは, ビデオ映像上で発生した事象情報を表現する。利用者の嗜好情報は, 利用者の好みのチームや選手などの利用者の嗜好項目とその度合いである。

### 2.2 シーン抽出, アノテーション生成と重要度の算出

PDMSは, メタデータをもとに, 意味的なまとまりとなるシーンを抽出する。次にアノテーションを生成し, これをメタデータに付加する。そしてメタデータの意味的重要度を算出する。このシーン抽出, アノテーション生成及び重要度算出はルール記述(これを「ルールスクリプト」と呼ぶ)に従い, メタデータを複合的に解析することにより実行される。重要度は「重要度パラメタ」の算出ルールに従って算出される。「重要度パラメタ」はメタデータの属性の一つであり, 各種の意味的な特徴を表現する指標である。

PDMSは、重要度パラメタ値を算出するとともに、利用者の嗜好の度合いを表現する嗜好レベルを算出する。嗜好レベルの算出式もルールスクリプトに記述する。

### 3. プロ野球ダイジェスト配信システムにおけるシーン重要度算出アルゴリズムと評価

携帯端末向けプロ野球ダイジェスト配信システムは、球場から直接引き込んだ野球映像をソースとし、野球メタデータをほぼリアルタイムに生成して、あらかじめ登録された利用者の嗜好情報に基づいたダイジェストを配信するシステムである。嗜好情報としては、好きなチーム1つ、好きな選手3名が登録でき、本システムは利用者ごとに「お好みチームダイジェスト」、「お好み選手ダイジェスト」の配信を行う。ダイジェスト生成に当たっては、PDMSの重要度算出方式に基づく重要度算出アルゴリズムを開発している。本章では、我々が開発した野球映像向け重要度算出アルゴリズムの評価を行う。

#### 3.1 アルゴリズムの概要

##### 3.1.1 アノテーションの生成と重要度の算出

本システムでは、表1に示すようなアノテーションを生成している。重要度パラメタとしては、攻撃レベル、投手レベル、守備レベル、興奮レベルの4つのパラメタを定義している。

##### 3.1.2 重要シーンの抽出

プロ野球ダイジェスト配信システムでは、配信するダイジェストシーンのクオリティを実用に耐えるものとするため、あらかじめ1プレイのシーンごとに手動で編集切り出しを行っている。配信の単位もこの1プレイのシーンとし、映像シーンの最終的重要度もこの1プレイのシーンごとに求めている。本システムでは、最終的重要度の高い順に1プレイのシーン(6個~10個)を選択し、ダイジェストとして配信を行う。

#### 3.2 アルゴリズムの評価

本節では、プロ野球ダイジェスト配信システムが正しいダイジ

表1 アノテーション一覧

種別	アノテーション
攻撃的事象を表現	犠飛, 犠打, n点タイムリー, n点本塁打, 満塁本塁打, サヨナラ, 逆転, 先制, 同点, 勝越
投手に関連する事象を表現	3者連続三振
状況を表現	満塁, クリナップの打席

\* 各打席の最後の1球から、進塁やアウトなどで結果が確定するまでの連続したシークエンスのシーン。

表2 シーンリスト例

8月25日	オリックスー西武		
ダイジェストID	イニング	点数	ダイジェスト文
28695	1表	0-0	先頭 松井 右前打 無死一塁
28696	1表	0-0	無死一塁 松井 牽制アウト 1死
28698	1表	0-0	1死 小関 右飛 2死
28701	1表	0-0	2死 貝塚 中飛 チェンジ
28702	1裏	0-0	先頭 早川 ニゴロ 1死
28705	1裏	0-0	1死 大島 左飛 2死
28707	1裏	0-0	2死 谷 遊ゴロ 3者凡退チェンジ
28708	2表	0-0	先頭 カブレラ 右前打 無死一塁
28711	2表	0-0	無死一塁 和田 右飛 1死一塁
28713	2表	0-2	1死一塁 宮地 右越2点本塁打 先制 1死
:	2表	0-2	:

ェストを生成しているかどうかを評価する。評価方法は以下のとおりとした。

- 野球の試合を乱打戦(両チームの得点5点以上)投手戦(両チームの得点2点以下)、ワンサイドゲーム(両チームの得点差5点以上)に分類し、それぞれ4試合ずつ計12試合をサンプルとして選択する。
- 各試合に対して、表2に示すようなシーンリストを作成する。このシーンリストは、各試合の1プレイのシーンの集合であり、本システムが生成した説明文字列(ダイジェスト文)が表示されている。
- 各試合に対して、(株)文化工房<sup>†</sup>・大重ディレクターに以下の3種類のダイジェスト作成を依頼する。
  - A)中立ダイジェスト
  - B)ホームチームダイジェスト
  - C)アウェイチームダイジェスト
 中立ダイジェストは、どちらのチームのファンでもない中立の立場のダイジェストである。ホームチーム/アウェイチームダイジェストは、各チームのファン向けのダイジェストである。各ダイジェストは、それぞれ10個のシーンから構成されるとした。10個のシーンにはプライオリティをつけ、重要な順に並べてもらうこととした。これを正解集合とする。
- サンプル試合それぞれに対して、我々のシステムを利用して、重要度の高い順に上位10シーンから構成されるA)、B)、C)の3つのダイジェストを生成する。
- システムが自動生成したダイジェストと、手動による正

<sup>†</sup> (株)文化工房はスポーツ番組制作会社であり、本システムを共同で構築したアライアンスメンバーである。

解集合を比較する。比較にあたっては、システムがダイジェストとして抽出したシーン10個の中に手動で選択したシーンが何個含まれていたかの比率（適合率）を求める。これにより我々のシステムのダイジェスト生成における精度を判定する。

$$\text{適合率} = \frac{\text{正解シーン数}}{\text{システムによる抽出シーン数(10シーン)}} \quad (1)$$

サンプル試合に関して適合率を算出した結果が表3である。網掛けがかかっているのは、適合率70%以上を達成したダイジェストである。以下に表3の結果に基づく考察を行う。

● 乱打線

表3 適合率の結果

乱打戦

試合ID	対戦カード	結果	適合率		
			中立	Hチーム	Aチーム
322	近鉄 VS 西武	第16戦 6-5	70%	70%	60%
350	西武 VS 近鉄	第19戦 12-10	100%	70%	70%
458	オリックス VS 西武	第24戦 6-4	80%	80%	80%
465	ダイエー VS 西武	第25戦 6-5	80%	50%	60%
平均適合率			82.50%	67.50%	67.50%

ワンサイドゲーム

試合ID	対戦カード	結果	適合率		
			中立	Hチーム	Aチーム
394	西武 VS 近鉄	第24戦 7-2	60%	60%	50%
419	西武 VS ダイエー	第24戦 3-11	90%	40%	80%
360	西武 VS ダイエー	第19戦 10-5	80%	80%	60%
338	近鉄 VS ダイエー	第20戦 3-11	80%	70%	60%
平均適合率			77.50%	62.50%	62.50%

投手戦

試合ID	対戦カード	結果	適合率		
			中立	Hチーム	Aチーム
449	日ハム VS ダイエー	第27戦 1-0	40%	20%	20%
384	近鉄 VS オリックス	第23戦 2-1	40%	50%	40%
349	ダイエー VS 日ハム	第21戦 1-0	30%	50%	20%
330	ダイエー VS 西武	第17戦 1-2	30%	40%	50%
平均適合率			35.00%	40.00%	32.50%

乱打線の適合率は、ワンサイドゲームや投手戦に比較すると良い。適合率が低いダイジェストでは、主要選手のシーンを一部抽出しそこなっている。我々のアルゴリズムでは、クリンナップの打席の重要度は考慮しているが、クリンナップ以外の主要選手の重要度は考慮していない。また代打でヒットを打ったシーンなども抽出しそこなっており、代打の打席とその結果を重視したシーン選択機能が必要であると考え。さらに10シーン程度で構成されるダイジェストの場合は、試合終了シーンをダイジェストに含めるべきであるが、我々のアルゴリズムでは試合終了のシーンの重要度を算出していない。

● ワンサイドゲーム

ワンサイドゲームの場合、中立ダイジェストと勝ったチーム（点を多く入れたチーム）のチームダイジェストの適合率は比較的良いが、負けたチームのダイジェストの適合率が低い。これは現状のアルゴリズムが負けたチームのファン向けダイジェストを効果的に生成できないことを示している。負けたチームのファン向けには、投手の好投場面やファインプレイの場面を抽出できるようにすべきである。また結果として失敗に終わったとしても、チャンスであった場面などを抽出できるようにすべきである。大重氏は、最初負けていて途中で大逆転した、といったシナリオに基づいたシーン抽出をしている。逆転のきっかけとなったヒットの重要度を増したり、三振の数や総得点数などに応じたシーン重要度算出を行ったりするなど、全体の流れに基づいたシーン抽出機能が必要である。シナリオに基づいたシーン抽出方法として、鎌原らによる研究<sup>7)</sup>があるが、ここで提案されている手法などを取り入れ、試合の流れを考慮したシーン抽出を検討すべきであることがわかった。

● 投手戦

投手戦の適合率は非常に低い。我々のシーン重要度算出アルゴリズムが守備側事象の重要度算出に弱い、ということが顕著に表れている。ワンサイドゲームでも言えたことであるが、得点が高い場合、投手の好投シーンやファインプレイなどの守備的な良いシーンをより多く抽出するように工夫すべきである。

上記の考察により、重要度算出アルゴリズムの改良ポイントを以下のようにまとめた。

- 逆転のきっかけとなったヒットの重要度など、試合の流れに従ったシーン抽出機能を検討する
- 試合終了時の重要度を算出する（特に完投、完封における最終打者打ち取りシーンなど）
- 主要選手（クリンナップに限らず）の事象の重要度を上げる
- 守備側の事象の重要度を判定する仕組みを改良する（好投シーン、ダブルプレイ、盗塁死など）
- 攻撃であろうと守備であろうと、そのシーンが良い場面であるかどうかを示す指標が必要である

- f) 得点が入らず負けた試合においても、そのチームにとって良かった場面を抽出する仕組みを検討する

#### 4. シーン重要度算出アルゴリズムの改良と評価

前章における考察により、シーン重要度算出アルゴリズムの改良を行った。本章では、前章で述べた改良ポイントに対して、どのようにアルゴリズムを改良したかについて述べ、その評価結果を報告する。

##### 4.1 アルゴリズムの改良

###### 4.1.1 アノテーションの追加

改良ポイントa, b, c, eに対応するために、アノテーションの追加を行った。追加すべきアノテーションは、試合の流れを表現するアノテーション、試合終了に関連する事象を表現するアノテーション、状況を表現するアノテーションである。表4は、今回の改良において追加したアノテーションの一覧である。本システムの付加するアノテーションは、メタデータを複合的に解析して生成されるものである。例えば、表4における「得点まで繋がったヒット」のアノテーションは、図1のような処理により生成される。ヒットがあり、その後もう一度ヒットが発生し、得点が入って逆転したイニングがあったとする。最初のメタデータ「ヒット」(便宜的に「ヒット(1)」とする)を含む

表4 追加アノテーション一覧

種別	アノテーション
試合の流れを表現	n打席連続本塁打, n者連続本塁打, 得点まで繋がったヒット,
試合に終了に関連する事象を表現	完全試合, ノーヒットノーラン, 完封勝利, 完投勝利, 最終打者打取
状況を表現	得点圏に走者あり, 主要選手の打席, 得点差少ない, 残りイニングわずか

シーンAのオリジナルの重要度は、予めルールスクリプト上で定義されているメタデータ「ヒット」の重要度となる。続いてシーンBにおいて、メタデータ「ヒット」(同じく便宜的に「ヒット(2)」とする)と「加点」が発生し、その加点によって攻撃チームが逆転したとする。するとアノテーション「逆転」が生成される。シーンBの重要度は、メタデータ「ヒット」、「加点」の重要度にアノテーション「逆転」の重要度を加算したものとなる。さらにシーンを遡り、きっかけとなった「ヒット(1)」が検索される。きっかけとなった「ヒット(1)」に対してアノテーション「得点まで繋がったヒット」が生成付加される。「得点まで繋がったヒット」の重要度は、シーンBの重要度に依存して算出され(ここではシーンBの重要度の倍)、シーンAの重要度に付加される(シーンAの重要度再計算処理)。このアノテーションの追加と再計算処理は従来のルールスクリプトに加えて新たに記述され、自動的に実行される。このように発生した事象の重要度をきっかけとなった事象の重要度に反映させ、試合の流れを考慮した重要度計算も可能となる。この他にも表

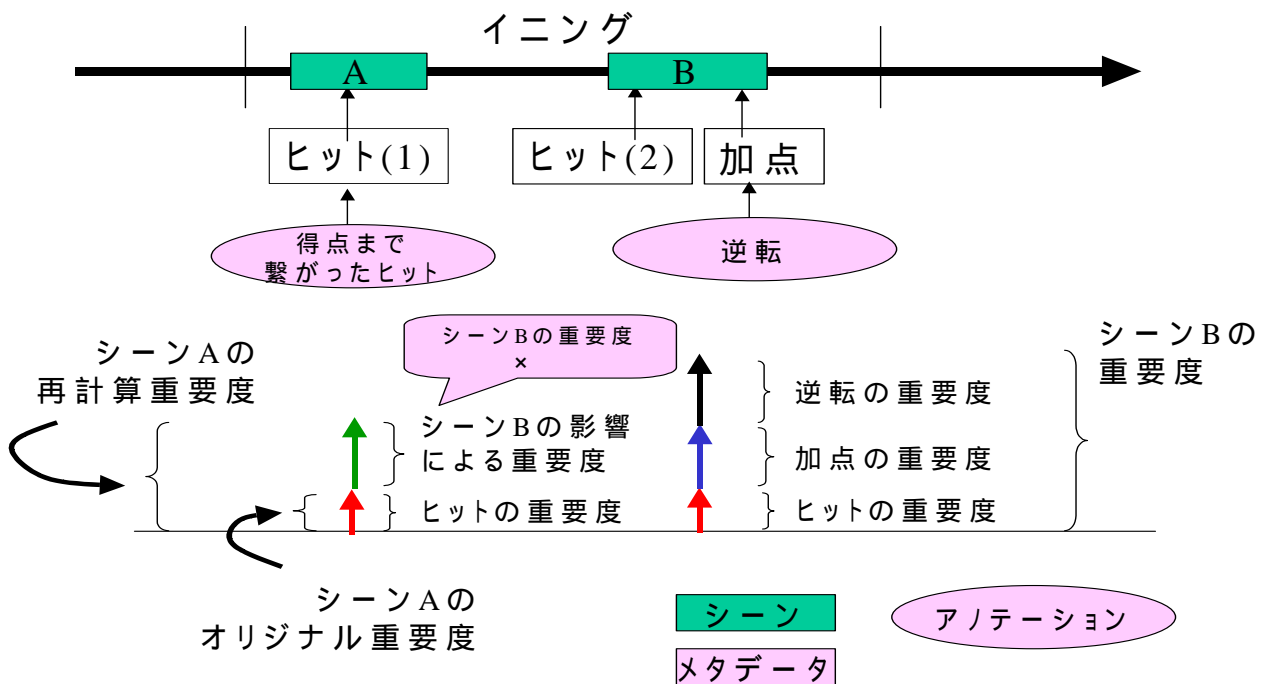


図1 アノテーションの追加と再計算処理

4に示すようなアノテーションを生成付加し、重要度を算出するルール記述を追加することにより、前章の考察における改良ポイントa, b, c, eの要求に対応することができる。

#### 4.1.2 重要度パラメタの整理

改良ポイントdへの対応として、守備的な事象の解析機能を高める必要がある。また改良ポイントeへの対応として、状況を表現できるパラメタが必要である。そこで重要度パラメタの再検討を行った。従来のアルゴリズムでは野球ダイジェストの重要度パラメタは、攻撃レベル、投手レベル、守備レベル、興奮レベルの4つであったが、投手レベルと守備レベルを分割したことにより、守備的なシーンの検索が複雑になっていた。興奮レベルの定義もあいまいであった。そこでパラメタの汎用性を高め、よりわかりやすくするために、以下の3つの重要度パラメタを再定義した。

- 攻撃レベル  
攻撃的重要度を表現するパラメタ。
- 守備レベル  
守備的重要度を表現するパラメタ。従来の投手レベルは、この守備レベルに吸収し、汎用性と検索性を高める。投手レベルが表現していた投手の調子については、次項で述べるチームレベルを利用して表現することとする。
- 状況レベル  
各映像シーンの最初の状況の良し悪しを示す。シーンの最初においてのみ算出される。従来の興奮レベルの定義をより明確にしたもの。

攻撃レベルと守備レベルは、従来通り映像シーン内の各メタデータの属性となるが、状況レベルは、その映像シーンの属性となる。表5は、3つのパラメタの値を制御するメタデータ及びアノテーションのリストである。

この3つの重要度パラメタにより、我々はスポーツ映像の映像シーンの重要度を、式(3)のように表現できると考える。

$$\begin{aligned} \text{シーン重要度} = & \omega_i \times \text{状況レベル} + \omega_j \times \sum_{\text{シーン内のメタデータ}} \text{攻撃レベル値} \\ & + \omega_k \times \sum_{\text{シーン内のメタデータ}} \text{守備レベル値} \end{aligned} \quad (2)$$

#### 4.1.3 チーム別重要度パラメタの導入

改良ポイントfに対応するためには、ある映像シーンがそのチームにとって良いシーンであったか、悪いシーンであったかを示す指標値が必要となる。そこで以下のようなチーム別の重要度

パラメタを導入し、各チームの「チームレベル」を算出することとした。

- チーム別重要度パラメタ  
攻撃レベル、守備レベルをチーム別に求める。該当チームが攻撃時は、攻撃レベルがそのチームのチーム別攻撃レベルとなり、守備の時は、攻撃レベルに(-1)を乗算した値が、そのチー

表5 パラメタ一覧

名前	値をプラスにするメタデータとアノテーション	値をマイナスにするメタデータとアノテーション
攻撃レベル	n打席連続本塁打, n者連続本塁打, 満塁本塁打n点本塁打, 三塁打, 二塁打, 逆転, 内野安打, n塁盗塁, 安打, タイムリー, 先制, 同点, 勝越加点, 振逃げ, 犠牲フライ, 犠牲バント, 得点まで繋がったヒット	オーバーラン, 守備妨害
守備レベル	完全試合, ノーヒットノーラン, 完封勝利, 完投勝利, トリプルプレー, 最終打者打取, ダブルプレイ, 三振, クリナップのアウト, 3者連続三, 好捕, 好送球, チェンジ, 3者凡退, アウト, 凡打, 牽制死, 盗塁死	フォアボール, デッドボール, 失策
状況レベル	代打, 2アウト, 得点圏に走者あり, n塁に走者あり	

ムの攻撃レベルとなる。守備レベルに関しても同様である。以下にチーム別攻撃レベルの算出式を示す。

$$\begin{aligned} \text{攻撃レベル(Hチーム)} = & \{ \text{攻撃レベル}[\text{攻撃チーム} = \text{Hチーム}] \\ & | (-1) \times \text{攻撃レベル}[\text{攻撃チーム} = \text{Aチーム}] \} \end{aligned} \quad (3)$$

$$\begin{aligned} \text{守備レベル(Hチーム)} = & \{ \text{守備レベル}[\text{守備チーム} = \text{Hチーム}] \\ & | (-1) \times \text{守備レベル}[\text{守備チーム} = \text{Aチーム}] \} \end{aligned} \quad (4)$$

- チームレベル  
上記のチーム別重要度パラメタを基に、次のような映像シーンごとのチームレベルを算出する。

$$\begin{aligned} \text{チームレベル} = & \omega_i \times \text{チーム別状況レベル} \\ & + \omega_j \times \sum_{\text{シーン内のメタデータ}} \text{チーム別攻撃レベル値} \\ & + \omega_k \times \sum_{\text{シーン内のメタデータ}} \text{チーム別守備レベル値} \end{aligned} \quad (5)$$

チームレベルが高ければ、攻撃、守備にかかわらずそのシーンは該当チームにとって良い場面となり、低ければ悪い場面となる。また得点数が少ないチームの場合、守備レベルに対する重みを高くすれば、守備的なシーンの重要度が上がり、負けていても良い場面を抽出してくるようになる。これにより改良ポイントeに対応することが可能となる。

また投手の活躍シーンは、その投手が投球しているシーンの中でチームレベルが高いものを検索すれば得ることができる。同様に活躍していない場面は、チームレベルの低い投球シーンとなる。これにより投球レベルで表現していた投手の調子を表現することが可能となる。

## 4.2. 新アルゴリズムの評価

前節で述べた改良に基づき、新アルゴリズムを実装し、ダイジ表7 大重ディレクターの正解集合に対する改良後の適合率

### 乱打戦

試合ID	対戦カード	結果	適合率		
			中立	Hチーム	Aチーム
322	近鉄 VS 西武	第16戦 6-5	80%	100%	90%
350	西武 VS 近鉄	第19戦 12-10	90%	90%	80%
458	オリックス VS 西武	第24戦 6-4	90%	80%	90%
465	ダイエー VS 西武	第25戦 6-5	80%	90%	70%

### ワンサイドゲーム

試合ID	対戦カード	結果	適合率		
			中立	Hチーム	Aチーム
394	西武 VS 近鉄	第24戦 7-2	80%	80%	70%
419	西武 VS ダイエー	第24戦 3-11	80%	70%	100%
360	西武 VS ダイエー	第19戦 10-5	70%	90%	70%
338	近鉄 VS ダイエー	第20戦 3-11	80%	80%	90%

### 投手戦

試合ID	対戦カード	結果	適合率		
			中立	Hチーム	Aチーム
449	日ハム VS ダイエー	第27戦 1-0	60%	100%	50%
384	近鉄 VS オリックス	第23戦 2-1	60%	90%	70%
349	ダイエー VS 日ハム	第21戦 1-0	80%	80%	60%
330	ダイエー VS 西武	第17戦 1-2	80%	70%	100%

エストの生成を行った。より正確な評価を行うため、サンプル試合を合計90試合(乱打戦 33試合、ワンサイドゲーム 29試合、投手戦 28試合)まで増やし、大重ディレクターだけでなく、一般野球ファンのA氏、B氏にも正解集合の作成を依頼した。A氏、B氏に対しても、大重ディレクターと同様、「ディレクターになったつもりで、中立のダイジェスト、ホームチームファン向けダイジェスト、アウェイチームファン向けダイジェストを作るためのシーンを選んで下さい」と依頼をした。3.2節の評価と同様、それぞれのダイジェストは、10個のシーンから構成されるとした。

表6に大重ディレクター、A氏、B氏のサンプル試合(90試合)における平均適合率を示す。平均適合率が70%を越えたものには網掛けをしている。また表7は、3.2節で評価した12試合に特に注目し、大重ディレクターの正解集合に対して、改良アルゴ

表8 大重ディレクターの結果とシステムの結果の比較

2002/9/22 日本ハム-ダイエー

投手戦・中立ダイジェスト

大重ディレクターによる手動選択

シーンID	イニング	点数	ダイジェスト文
34123	1裏	1-0	2死 小笠原 右中間本塁打 先制
34316	9表	1-0	2死二塁 松中 右前打 代走田中瑞
34294	8表	1-0	1死 辻 右中間二塁打 1死二塁
34295	8表	1-0	1死二塁 出口 右飛 辻二塁戻れず アウト ダブルプレー チェンジ
34204	5表	1-0	無死一塁 城島 遊ゴロ 松中二塁ア ウト ダブルプレー 2死
34278	7裏	1-0	2死 小笠原 遊前内野安打 2死一
34240	6表	1-0	1死一、二塁 代打柴原 三振 2死一、二塁
34254	6裏	1-0	塁
34164	3表	1-0	2死一塁 辻 二塁盗塁死 チェンジ
34188	4表	1-0	2死 小久保 三振 3者凡退チェンジ

システムによる自動選択

シーンID	イニング	点数	ダイジェスト文
34316	9表	1-0	2死二塁 松中 右前打 代走田中瑞 本塁アウト 試合終了
34123	1裏	1-0	2死 小笠原 右中間本塁打 先制
34315	9表	1-0	1死二塁 小久保 三振 2死二塁
34294	8表	1-0	1死 辻 右中間二塁打 1死二塁
34295	8表	1-0	1死二塁 出口 右飛 辻二塁戻れず アウト ダブルプレー チェンジ
34284	7裏	1-0	2死一塁 代走古城 二塁盗塁死 チェンジ
34306	8裏	1-0	1死一、二塁 木元 遊ゴロ 林二塁ア ウト ダブルプレー チェンジ
34240	6表	1-0	1死一、二塁 代打柴原 三振 2死一、二塁
34164	3表	1-0	2死一塁 辻 二塁盗塁死 チェンジ
34298	8裏	1-0	先頭 田中幸 三振 1死
34278	7裏	1-0	塁
34188	4表	1-0	2死 小久保 三振 3者凡退チェンジ
34204	5表	1-0	無死一塁 城島 遊ゴロ 松中二塁ア ウト ダブルプレー 2死
34308	9表	1-0	先頭 柴原 中前打 無死一塁
34254	6裏	1-0	1死一、二塁 森本 三振 2死一、二

リズムがどのような適合率を出したかを示している。表6及び7に基づいて評価を行う。

表6, 7の結果から、以下のことがわかる。

- 大重ディレクターの正解集合にたいする平均適合率は、投手戦の中立ダイジェストとホームチームダイジェストを除き、70%以上を達成している。
- 一方、A氏、B氏の正解集合に対する平均適合率は非常に低い。中でもホームチームダイジェストとアウェイチームダイジェストにおける適合率は10~20%のものが多く、中立ダイジェストは、ホームチームダイジェスト、アウェイチームダイジェストと比較すると、適合率はやや高い(50%前後)。

上記のような結果になった理由を考察する。

### 1. 大重ディレクターの投手戦に対する適合率が低いことに関する考察

投手戦の場合、目立った攻撃場面がないため、手動であったとしてもダイジェストシーンを選択することが難しい。大重ディレクターからも投手戦のシーン選択に迷った試合が多かったとの報告を受けている。今回の評価作業終了後、我々のシステムが自動抽出したシーンと、大重ディレクターが手動で選んだシーンを比較するレビューを大重ディレクターと共に行い、特に適合率の低い投手戦について重点的にチェックをした。表8は2002年9月22日の日本ハム対ダイエー戦(投手戦)に対して、大重ディレクターが選択した10シーンと、システムが自動選択したシーンの比較である。なおシステムの自動選択の表では上位15シーンまでを表示している。適合率はシステムが自動抽出した上位10シーンに対して計算を行っている。上位10シーンのうち網掛けのかかった6シーンを両者が共通で選んでいるため、適合率は60%となる。この結果に対して、大重ディレクターからは、「システムが自動で選んだシーンも決して悪いシーンではなく、このシーンでも構わない」という感想をいただいている。

今回の評価作業では、ダイジェストを構成するシーンを10シーンと限定してしまったため、投手戦のようにシーンを選択しにくい試合(重要シーンが10シーンに満たないかもしれない試合)でも手動で必ず10シーンを選択しなければならなくなり、どうしても必要なシーンとどちらでもいいシーンの区別がつくにくくなってしまったと考える。本来ならば正解数を限定せず、ダイジェストを構成するのにふさわしいシーンを選んでもらう、というようにするべきであったかもしれない。システムの結果も手動選択に合わせて10シーンとせず、ある一定の閾値を超えたシーンを自動抽出結果とする、といった方式にすべきであったと考える。

また大重氏が「システムが選んだシーンでも構わない」といったことに対しては、シーンID34188と34315のシーンが両方

とも小久保の三振シーンであること、シーンID34284と34278が共に7回裏のアウトシーンであり似ていること、大重氏だけが選択した4つのシーンはシステムの自動抽出の上位15シーンには含まれており、高い重要度をもっていることなどから、類似度が高いシーンならば、別のシーンを挿入してもダイジェストとしての満足度は変わらないのではないかと、との仮定が成り立つ。その場合、どのような特徴に関して類似度を算出するか、単に重要度が高ければ取り替えてもいいのか、という点が今後問題となってくる。

### 2. A氏、B氏の適合率が低いことに関する考察

次に大重氏に対する適合率に比べ、A氏、B氏の適合率が格段に低いことに関する考察を行う。A氏、B氏の適合率が低い理由として、今回のアルゴリズムの改良が結果として大重氏の嗜好に合わせるためのチューニングになってしまったのではないかと考えられる。我々は3.2節の結果を基に、大重氏が選択したシーンを自動で選択できるようにアルゴリズムを改良した。その際、大重氏が選択したシーンを表現できるようなアノテーションの生成を検討し、アノテーションが付加されているシーンの重要度が上がるように変更したため、大重氏の嗜好にあったシーン抽出が実現できた。このアノテーション記述が大重氏に対するパーソナライゼーションとなったと考える。大重氏はスポーツ番組ディレクターであり、一般ユーザ向けのダイジェストにどのようなシーンが必要であるかという視点でシーンを選択している。一方、A氏、B氏は一般の野球ファンであり、見たいシーンが番組ディレクターの視点と異なる可能性が考えられる。これが大重ディレクターとA氏、B氏の適合率に大きな差が出た原因の一つであると考えられる。このことは、A氏、B氏の結果の中で、中立ダイジェストの適合率が比較的良かったことから推察される。嗜好によらない中立的なダイジェストの場合、A氏、B氏は客観的に重要なシーンを選択しようとしたはずであり、その点で大重ディレクター用にチューニングしたシステムの選択シーンと重なりがあったのではないかと。A氏、B氏に対しても、大重ディレクターに対して行ったようなアノテーション記述ルールの追加(パーソナライゼーション)をすれば、適合率は上がってくる可能性がある。このことからダイジェストにおける個人化は、アノテーション記述のレベルで行うべきことなのかもしれないという仮定が成り立つ。

### 5. まとめ

現在我々が運用している携帯端末向けプロ野球ダイジェスト配信システムの重要度算出アルゴリズムに対し、重要シーンを正しく抽出できているかの評価を行い、アルゴリズムの改良を

行った。改良アルゴリズムについても評価したところ、番組制作者である大重ディレクターのシーン抽出結果に対する適合率は比較的良かったが、投手戦の適合率が良くない、という結果が得られた。また大重ディレクター以外の2名の一般人(A氏, B氏)に対する適合率は悪い、という結果も得られた。この結果にいたった理由を以下のように考えた。

- ダイジェストの場合、似たようなシーンならば、どちらのシーンが入っていてもかまわないということがある。特に投手戦のようにシーンを選択しにくいダイジェストの場合はその傾向が顕著である。
- アルゴリズムの改良においてアノテーションの追加を行ったが、アノテーションの追加及び重要度の再計算ルールの記述が、実は大重ディレクター向けのパーソナライゼーションであったと考えられる。他の2名についても大重ディレクターと同様にアノテーション追加や重要度の再計算ルールを検討すれば適合率が向上すると考えられる。
- ダイジェストの場合、従来の嗜好項目によるパーソナライゼーションだけでは不十分であり、流れを考慮したパーソナライズが必要となる可能性がある。

上記の考察を踏まえ、今後は、ダイジェストの評価方法の再検討を行い、ダイジェストを評価するためにはどのような評価尺度が必要であるかについて提案を行い、より精度の高い評価を行っていく予定である。

## 参考文献

- 1) 橋本隆子, 白田由香利, 真野博子, 飯沢篤志, "TV受信端末におけるダイジェスト視聴システム", 情報処理学会論文誌:データベース, vol.41No.SIG 3, pp.71-84, May 2000. Vol.2000, No.10, 2000-DBS-120-15,pp107-112(2000)
- 2) T. Hashimoto, Y. Shirota, Atsushi Iizawa, and H. S. Kunii: "Personalized Digests of Sports Programs Using Intuitive Retrieval and Semantic Analysis", Alberto H. F. Laender, Stephen W. Liddle, and Veda C. Storey (Eds.): *Conceptual Modeling - ER 2000, Proc. of 19th International Conference on Conceptual Modeling*, Salt Lake City, Utah, USA, Oct. 9-12, 2000, *Lecture Notes in Computer Science*, Vol. 1920, Springer, 2000, pp. 584-585.
- 3) T. Hashimoto, Y. Shirota, A. Iizawa, and H. Kitagawa: "A Rule-based Scheme to Make Personal Digests from Video Program Meta Data", *Proc. of 12th International Conference DEXA2001*, Sep. 3-7, 2001, *Lecture Notes in Computer Science*, Vol. 2113, Springer 2001, pp.243-253.
- 4) 橋本隆子, 白田由香利, 飯沢篤志, 北川博之: "ターニングポイントの解析に基づくダイジェスト作成方式", 情報処理学会論文誌:データベース, Vol.43 No.SIG5, pp1-11 (2002)
- 5) 橋本隆子, 加登岡隆, 飯沢篤志: "モバイル環境におけるプロ野球パーソナルダイジェスト配信システム", DBWeb2002予稿集, pp331-338 (2002年12月)
- 6) T. Hashimoto, T. Katooka, and A. Iizawa: "personal Digest System for Professional Baseball Programs in Mobile Environment", *Proc. Of MDM2003*, Jan.21-24, 2003 (to appear)

- 7) J.Kamahara, T.Kaneda, M.Ikezawa, S.Shimojo, S.Nishio, and H.Miyahara: "Scenario Language for automatic News Recomposition on The News-on Demand", *Technical Report of IEICE DE95-50*, Vol.95, No.287, pp1-8, 1995 (in Japanese)