

# MPEG-1 動画に含まれる動き補償情報に基づく野球プレー推定

## Play Estimation in Base-ball Games based on Motion Compensation Information in MPEG-1 Video

浅井 航平<sup>†</sup>  
Kohei Asai

青木 恭太<sup>†</sup>  
Kyota Aoki

### 1. はじめに

近年、動画サイトが普及し、映像に触れる機会が増えている。これらの映像から利用者の望む映像部分を検索するシステムの必要性が高まっている。とくにスポーツ映像では、プレーに基づく検索が望まれる。

スポーツのプレーを検索する方式として動きに基づく類似映像検索方式がある[1]。この方式は全てのフレーム、全ての画素に対する動き情報の計算が必要であり、類似映像検索のための動き推定の負荷が大きく、対象映像時間の10倍ほどの処理時間を要している。そこで本論文はMPEG-1符号化された映像が保持している動き補償情報を使用する。この動き補償情報を利用することで、動き推定の負荷を削減し、処理の高速化を図る。本論文と類似した動きに注目した方式としては、ホームランなどの大きなカメラワークに着目し、MPEG符号化情報を利用した方式がある[2]。またMPEG符号化情報を利用し、サッカーの映像を自動で分類させる研究[3]やショットチェンジを検出する研究もある[4]。

本論文ではMPEG-1符号化された映像に含まれる動き補償情報を用いて、スポーツ映像において類似プレー検索を行う手法を提案する。また野球映像において投手の打者に対する投球の検出実験を行い提案方式の能力を確認する。

### 2. MPEG 構造体

MPEG-1形式は下記に示すIフレーム、Bフレーム、Pフレームの3種類のフレームを保持している。

#### ・Iフレーム(Intra-coded Frame)

1枚の完全な状態の画像。このフレーム自体は動き補償情報を持たない。本論文では直前のフレームの直後動き情報を反転し、このIフレームの直前動き情報としている。

#### ・Pフレーム(Predicted Frame)

一つ前のIフレームとの差分画像であり、動いた部分だけの情報を持つ。本論文では直前のBフレームの直後動き情報を反転して、Pフレームの直前動き情報としている。

#### ・Bフレーム(Bi-directional Predicted Frame)

Bフレームは直前動き情報、直後動き情報の両方を持つ。本論文では直前動き情報を動き補償情報として利用する。

動き情報は、マクロブロック単位に保持される。マクロブロックは、 $16 \times 16$ 画素の領域であり、動き補償情報を持つ。この動き補償情報を利用してMPEG-1符号化方式では情報量を削減している。

本論文ではBフレームに保持されている動き補償情報のみを利用する。これにより、提案方式は、背景やプレーヤのユニホームの色や模様の変動に対して頑健となる。

### 3. 時空画像照合による類似映像検索

#### 3.1 提案方式概要

本論文ではMPEG-1符号化によるI,B,Pの3種のフレームの内Bフレームから動き補償情報を取り出し、その動き補償情報を利用して動きに基づくスポーツのプレーの検索方法を提案する。図1に提案手法の流れを示す。MPEG-1形式の映像から動き補償情報を画像として取り出し、取り出した動き補償情報画像に対して1次元投影を行い、情報量を削減する。1次元投影を行った後にできた画像を時系列順に並べることで、動きの特徴を保持した時空画像を作成する。プレー推定は、推定したいプレーの例示時空画像と対象映像の時空画像との類似画像検索により行う。

#### 3.2 時空画像構成方式

##### 3.2.1 動き補償情報抽出

本論文では動きに基づく類似映像検索を行う上で必要な動き補償情報をMPEG-1符号化された映像から抽出する。本論文では音声信号を利用しないので、無音のMPEG-1符号化映像を利用する。映像を復号し、そこから動き補償情報を抽出し、マクロブロック単位で動きを抽出する。動き補償情報にマクロブロック単位での画素の移動量が記録されているので、移動量をカラー画像として取り出すことで $16 \times 16$ ピクセル単位で直後のフレームとの動きを得ることができる。

動き補償情報はカラー画像として抽出する。マクロブロックが現フレームから右方向に移動した場合を正、左方向に移動した場合を負としている。Y方向に関しては上方向を正、下方向を負としている。各画素値を $128 \pm 10 \times$ (移動量)として記録する。動きがない場合は(128,128,255)としている。赤値にX方向に対する情報を、緑値にY方向に対する情報を記録している。また、青値は0としている。重みを10としたのは、1フレーム間での移動量を考慮した結果であり、映像の動きの激しさによっては調整する必要がある。本論文で扱った野球の映像では移動先としては最大10という結果が出たので重みを10とした。

##### 3.2.2 1次元投影

前述で取り出した動き補償情報画像をそのまま利用し、各種類似映像検索方式で類似判定を行い、その結果として動きに基づく類似映像検索を実現することも可能である。しかし動き補償情報画像系列をそのまま用いた場合、処理量は大きい。そこでマクロブロックごとに得られた動き補償情報を保持する動き情報画像系列の各画像のXY各方向への投影を作成することで情報量を削減する。

画素の行と列それぞれに対して輝度平均を計算し、投影分布としている。X軸への投影分布及びY軸への投影分布

<sup>†</sup>宇都宮大学 Utsunomiya University

の転置分布を連結する(図2右上の操作)ことによりX方向1次元に空間方向1次元の投影分布を得る。さらにそれを時間方向に連結することにより空間方向1次元, 時間方向1次元で, 各画素にX方向の動き補償情報, Y方向の動き補償情報をもつ時空画像を得る。本論文では空間方向を時空画像のX方向, 時間方向を時空画像のY方向としている。図2に動き補償情報画像から時空画像への生成過程, 図3に時空画像の例を示す。

### 3.3 例示時空画像と検索画像の相関

プレー推定は, 例示時空画像と検索対象映像の時空画像の類似映像検索により実現する。類似映像検索は, X方向動き, Y方向動きの相関を求め行う。相関を求める際, 例示時空画像と検索対象映像時空画像をそれぞれ赤, 緑, 青に単色分離し, グレースケール化している。赤値にX方向, 緑値にY方向の情報が記録されているのでそれぞれ相関を求める。得られた相関値を探索範囲で正規化し類似度ベクトルとしている。閾値を設定し, X方向, Y方向の類似度が閾値を共に超えている位置を例示プレーと推定する。

## 4. プレー認識実験

### 4.1 実験1

実験では提案方式を用いて, 高校野球の映像から投手の打者に対する投球動作を検出した。実験1では, 実験対象にテレビのアナログ放送の高校野球中継を350×240で録画したものを320×240のMPEG-1形式, 無音, 30fpsに符号化したものを用いた。検索対象映像は約70分の映像(132503フレーム)をTEST1, 約90分の映像(152275フレーム)をTEST2とした。投球数は映像TEST1が右投げ1名, 左投げ1名の2投手で168, 映像TEST2が右投げ2名の2投手で180である。投手の打者に対する投球動作は, 投手の背後からバッターを見るような視点で撮影され, 投球動作を撮影している部分のカメラワークはすべて同一である。

実験は, 基準となる例示投球部分時空画像を用いて類似映像検索を行い, 投手の打者に対する投球動作を認識した。

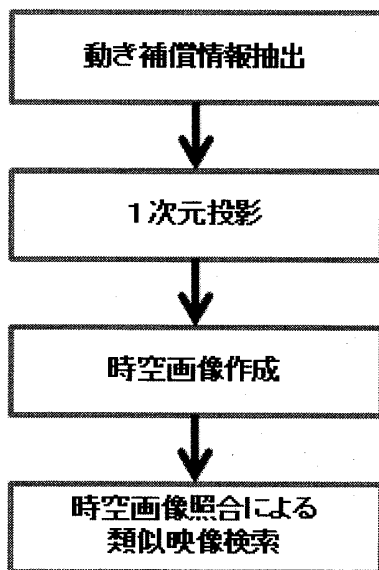


図1. 提案手法の流れ

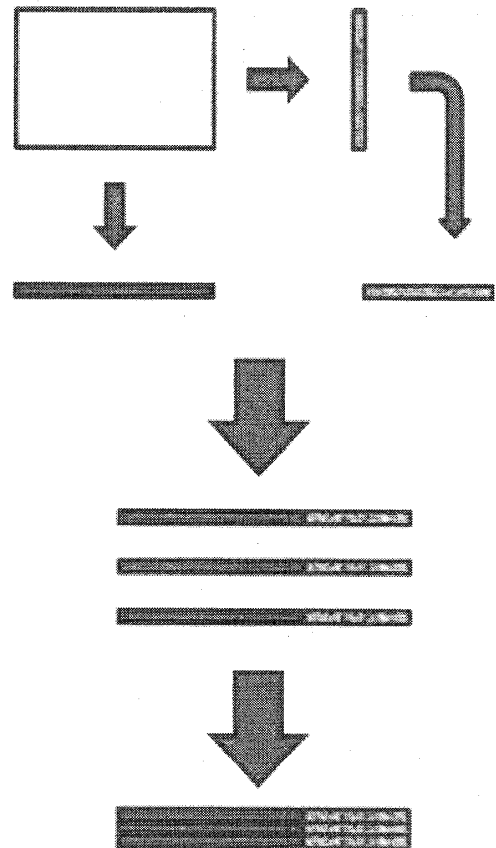


図2. 動き補償情報画像から時空画像への生成過程

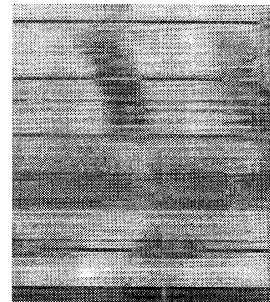


図3. 時空画像の例

基準となる投球動作例は映像TEST2の83004フレーム目から始まる20フレームの部分とした。選択した投球動作例は, 右投手で, 振り上げた左足が下がり始めた時を投球動作の始まりとし, 左足が地面に着いた時を投球動作の終わりとしている。今回の実験では, 見逃し率と誤検出率がほぼ同じ値になるように閾値を設定した。見逃し率は(発見できなかった数)/(すべての投球回数), 誤検出率は(投球位置以外での検出数)/(全発見回数)である。また本実験に使用した計算機のスペックはPentium4,

表1. 実験1映像 TEST1 類似映像検索結果

映像名	フレーム	投球数	投球発見数	見逃し数
TEST1	1-10000	10	10	0
	10001-20000	11	7	4
	20001-30000	13	13	0
	30001-40000	14	12	2
	40001-50000	13	10	3
	50001-60000	15	15	0
	60001-70000	12	8	4
	70001-80000	11	10	1
	80001-90000	14	8	6
	90001-100000	18	18	0
	100001-110000	16	8	8
	110001-120000	10	8	2
	120001-130000	10	7	3
	130001-132503	1	0	1
	計	168	134	34

表2. 実験1映像 TEST2 類似映像検索結果

映像名	フレーム	投球数	投球発見数	見逃し数
TEST2	1-10000	13	11	2
	10001-20000	12	8	4
	20001-30000	13	7	6
	30001-40000	10	8	2
	40001-50000	12	8	4
	50001-60000	11	11	0
	60001-70000	12	9	3
	70001-80000	16	16	0
	80001-90000	13	13	0
	90001-100000	12	11	1
	100001-110000	10	9	1
	110001-120000	10	9	1
	120001-130000	16	10	6
	130001-140000	9	8	1
	140001-150000	9	8	1
	150000-152275	2	2	0
	計	180	148	32

2.59GHz, 980MB の RAM の Windows XP Professional の計算機である。90 分の映像の動き補償情報の抽出におよそ 2 時間半、分析処理に 30 分程度の計 3 時間程度でプレー認識が完了した。高信頼動き推定を行う従来の方式では、映像時間の約 10 倍の処理時間を要しているため、従来の 5 分の 1 程度の時間で類似映像検索が行えている。

表 1, 2 に実験 1 の結果のまとめを示す。映像 TEST1 での発見率は 79.8% で見逃し率と誤検出率は 20.2% である。映像 TEST2 での発見率は 82.2% で見逃し率と誤検出率は

17.8% である。実験全体で見ると発見率は 81.0%, 見逃し率と誤検出率は 19.0% であり約 80% で投球動作が検出されていることがわかる。投手によって細かい動きの異なりはあるが、体幹の動きは総じて同じであり、それにより投球動作を検出できている。図 4 に実験 1 の映像 TEST1 の 5000-10000 フレームの類似度の変動を示している。X 方向の類似度はどのフレームでも 0 付近であるのに対し、Y 方向では投球動作部分で高い類似度を示している。主に Y 方向の動きで投球動作を検出していることがわかる。従来の方式では発見率が 100%, 適合率が 99% であるので提案手法は動き推定を行う方式より精度は落ちている。適合率は検出数を検出数と誤検出数の和で割ったものである。

図 5 に誤検出の例を示す。誤検出は投球動作例においての投球位置と同じような位置で上から下への動きがあった場合に生起している。

#### 4.2 実験 2

実験 2 では実験 1 で使用した、テレビのアナログ放送の高校野球中継を 350×240 で録画したものを 640×480 サイズに MPEG-1 符号化して実験 1 と同様の処理を行う。投球動作例も同様に 640×480 サイズにしている。

映像 TEST1 の発見率は 82.7%, 見逃し率と誤検出率は 17.3% である。320×240 サイズで行った実験 1 と比較すると 2.9 ポイントほど精度が上がっていることがわかる。映像 TEST2 の発見率は 80.6%, 見逃し率と誤検出率は 19.4% である。実験 1 と比較すると映像 TEST2 は発見率が 1.6 ポイント下がっていることがわかる。また 1-10000 フレームの見逃し数、誤検出数が多くなっているのが確認できた。その原因としては投手以外の動きの影響だと考えられる。320×240 サイズでは投手以外の人間の動きや観客席の動きやランナーの動きであるノイズが多少あっても認識できていた。サイズを大きくするとノイズの影響が大きくなることがわかる。実験全体で見ると発見率は 81.7% で実験 1 とほぼ変わらない割合で検出している。

#### 4.3 実験 3

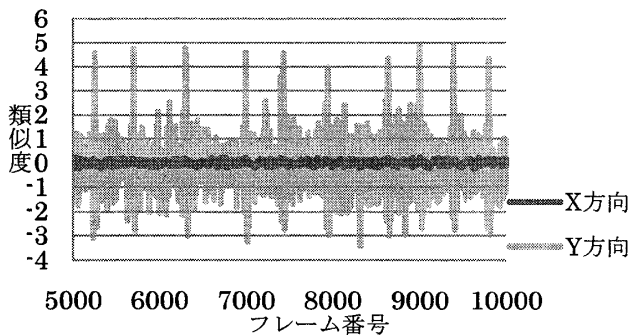
実験 1 では投球動作例を 20 フレームとしたが、実験 3 ではそれより長い 61 フレームの投球動作例を利用する。使用する映像は、

実験 1 で使用した 320×240 サイズの映像 TEST1 である。61 フレームの投球動作例は映像 TEST1 の 5226 フレームから始まる右投手の投球動作を選択した。投球動作例は、投手がウィンドアップを終え、左の足が地面から離れた時を始まりとし、投手が投げたボールがキャッチャーのミットに収まるときを終わりとしている。

発見率は 31.0%, 見逃し率と誤検出率は 69.0% である。発見した投球動作は全体の 3 割と 20 フレームを利用した

表3. 実験結果まとめ

	映像名	X方向 閾値	Y方向 閾値	投球 発見率	見逃し 率	誤検出 率
実験 1	TEST1	-0.42	3.05	79.8%	20.2%	20.2%
	TEST2	0.15	2.81	82.2%	17.8%	17.8%
実験 2	TEST1	-0.41	2.97	82.7%	17.3%	17.3%
	TEST2	0.16	2.43	81.0%	19.0%	19.0%
実験 3	TEST1	2.00	1.50	31.0%	69.0%	69.0%

図4. 実験1映像TEST1の5000-10000フレーム  
における類似度の変動

実験1と比較して低下している。61フレームの投球動作例はボールがキャッチャーミットに収まるシーンまで含まれているので、バッターのスイング動作に影響をうける。また、セットポジションなどの場合もあり、投球タイミングが変わる場合がある。長い投球動作例は、同様の投げ方および打者のプレーを含めて認識することに適していると考えられる。表3に実験結果全体のまとめを示す。

## 5. むすび

本論文ではMPEG-1符号化された映像が持つ動き補償情報を利用した動きに基づくプレー認識方式を提案した。

実画像実験では、高校野球映像から約8割の精度で投球動作を検出し、動き推定を行う方式よりも高速に類似映像検索を行えることを確認した。

提案方式は動き推定を行う方式に比べて精度は低下しているので、精度の向上が今後の課題である。また速度の点においては、本論文では動き補償情報をカラー画像として書き出しているため、この点が速度低下の要因となっている。この過程を省き直接時空画像を作成することで処理時間はさらに短縮できると考えられる。

## 謝辞

本研究に際して、共同研究者としてお力添えをいただいた小熊宇宙氏に深く感謝いたします。

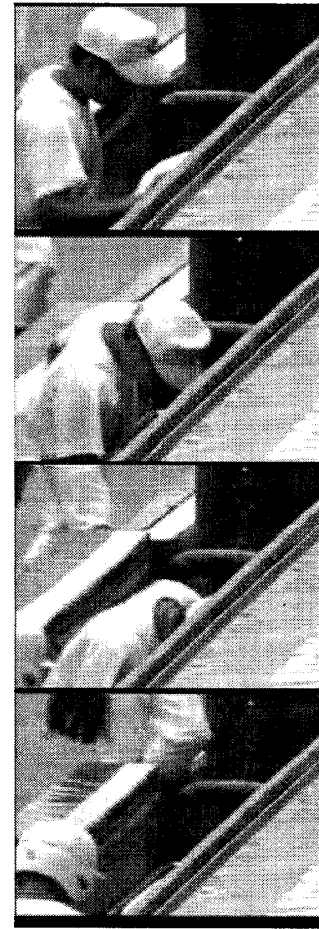


図5. 誤検出例

## 参考文献

- [1] 青木 恭太: "動きに基づく類似映像検索," 信学論, vol.J92-D, no.10, pp.1743-1745, October, (2009).
- [2] 片岡 良治, 遠藤 斉: "MPEG符号化情報に基づく類似シーン検出方法," 情報処理学会論文誌, NoSIG3, 41, pp37-45, (2000).
- [3] 玄蕃 一輝, 金子 正秀, 樽松 明: "MPEG符号化されたサッカー映像に対するシーンの自動分類のための情報抽出," 映像情報メディア学会誌, vol.55, no.3, pp.417-421, (2001).
- [4] 新倉 康巨, 阿久津 明人, 谷口 行信, 佐藤 隆, 外村 佳伸, 浜田 洋: "MPEG符号化映像ショットチェンジ検出のための動き補償解析ハイブリッド法の提案," 信学論, vol.J81-D-II, no.8, pp.1838-1846, August, (1998).
- [5] 浅井 航平, 青木 恭太: "MPEG-1動画に含まれる動き補償情報に基づく野球プレー推定," 信学技報, vol.110, no.22, pp.1-6, (2010).