Receding Horizon 戦略に基づく固定カメラを用いた競技動画の再構成

岡 島 寛*・橋 爪 壮一郎*
大 石 凌 史*・宮 原 智 新**

Reconstruction of Sports Competition Video Using Fixed Camera Based on Receding Horizon Strategy

Hiroshi OKAJIMA*, Soichiro HASHITSUME*,
Ryoji OISHI* and Tomoyoshi MIYAHARA**

At the sports and games for elementary and junior high school students, their parents often become personal photographer for the game. They see their child over the frame despite being in the playing field. In addition, problems such as trying to take a good place to compete may occur. We can solve these problems by the conference management side distributing the sports competition videos. So in this paper, we show a method to automatically reconstruct movies using competition movies taken by fixed point cameras and to present high quality movies to viewers.

Key Words: sports competition, video reconstruction, receding horizon control, optimization

1. はじめに

東京オリンピックが開催される 2020 年に向けて、日本のスポーツ界全体が盛り上がりを見せている。さまざまな競技において活性化の機運があり、スポーツ競技映像をテレビなどで観戦する機会は多くなっている。一般にスポーツ競技映像は、プロ競技のように大規模な競技大会においては専用の予算を充てることができるため、専門のカメラマンによってその競技が撮影され、プロの編集者によって編集されたものを観戦することが可能である。その一方で、小中学生などの競技大会のように小規模の大会においてはコストがかからないため専門のカメラマンなどを配備し、観客に編集された高品質の競技映像を提供することはできない。

そのため、小規模大会においては競技を映像として記録するためには競技者の保護者や関係者などの観客が独自に該当の競技者の動画撮影を行なうことが多い。しかし、撮影者は競技場にて競技者を目前にしているにもかかわらずビデオカメラのフレーム越しに競技を観戦することになるというデメリットがある。このとき、競技者の保護者の立場であれば、

フレーム越しではなく直接観戦したいというのが本音であろう。また、撮影者はより良い映像を撮影する目的で、場所取りトラブルなどの観客間トラブルに発展するケースも存在する。これらの問題に対しては、大会の運営者が低コストで高品質な動画を撮影・配信することは強力な解決策となる。

このとき、低コストで質の高い動画を配信するためには、カメラマンや編集者を必要としない動画の撮影手法が必要不可欠である。定点固定カメラによる撮像を行えばカメラマンは不要であり、運営者が競技会場に設置するだけでコストを掛けずに動画を入手することができる。しかしながら、動画としての質を考えた場合、定点固定カメラの映像自体では満足できるものにならない。そこで本研究では、定点固定カメラ映像から仮想カメラワーク処理により動画を再構成し、良質な動画を制作する方法について考える。

競技の映像をカメラ姿勢やガイドレール上の位置を、装置を用いて自動制御することで撮像し、カメラワークを行なう方法は従来より研究されている^{1)~3)}。また、撮影した映像のシーン解析^{4)~7)}や、撮影した映像からトリミングによる仮想的なパンニング・ズームを行なう方法^{8),9)}は従来より検討されている。講義について、黒板と教員をうまく撮影する自動撮影システム¹⁰⁾の研究も行なわれている。

これに対して本論文では、動きが速いスポーツ競技を対象として、定点固定カメラにより撮影された動画を自動で再構成し、視聴者にとって良質な映像を提示することを目的とする。本論文では、定点固定カメラによる動画に対してパン動作のようなドラスティックな画角の切り替えや、ズーム

* 熊本大学大学院自然科学研究科 熊本市中央区黒髪 2-39-1

** (株) diffeasy 福岡市博多区祇園町 8-13

* Graduate School of Science and Technology, Kumamoto University, 2-39-1 Kurokami, Chuo-ku, Kumamoto

** diffeasy Corporation, 8-13 Gion-machi, Hakata-ku, Fukuoka

(Received July 31, 2018)

(Revised October 24, 2018)

インやトラック動作など仮想カメラワークを自動化し、良質で視聴価値の高い動画を再構成することを考える。従来研究とは異なり、定点カメラで得たオフラインの映像を基に動画を再構成することから、素材動画がすでに得られているものと仮定することができる。このとき、通常のカメラマンによるカメラワークのように現在時刻の状況（もしくは過去のフレーム履歴）に基づいて現在時刻でのフレームを決定する必要がなく、未来の情報を用いることができる点に本研究特有の特徴がある。

画角の切り替えを行なう場合は、通常は個々のフレーム内での画角の決定がなされる。フレームごとでの最適化を施す場合には、各フレームごとに評価関数値が最適となるポイントが大きく変わり得る。このことから、再構成された動画像にチャタリングが生じるなどの不具合が起こる可能性が高い。特に、本システム構成ではスポーツ競技を対象とするため、自動での対象物や対象者の検出も必要となってくることから、各フレームで対象物の認識領域が大幅に異なるようなケースが頻繁に生じるため画角の決定自体が困難になりやすい。また、将来的には複数台の固定カメラで同時に撮影し、これらの切り替えも含めたカメラワークを想定した仮想カメラワークシステムを考えることも必要になってくるが、従来行なわれている研究の拡張では複数台カメラの問題設定への対応は難しい。

本論文では、スポーツ競技の動画像再構成における有力な設計方針として Receding Horizon (RH) 戦略に基づいた動画像の再構成システムを提案する。RH 制御は制御工学分野で広く利用されている^{11)~16)}ものであり、現在時刻から有限時間区間での最適化を逐次的に解く制御手法である。具体的な処理としては、高解像度カメラから得た動画像に基づいて、注目領域のトリミングを行なうことで仮想的にパンニング、ズーム動作などのカメラワークを実行する。RH 戦略により、逐次的にフレームの注目領域をトリミングする手法を提案する。

まず、第2章では、本論文で対象とする空手競技の動画像撮影におけるポイントについて述べる。つぎに、第3章では、固定カメラ映像に基づいたフレームからの情報取得の方法について述べる。本論文では、OpenCV¹⁷⁾を用いることで動画像処理の実装を行なう。このとき、色領域に基づいた選手領域の情報取得とフレーム間差分を用いた選手領域の取得の方法について述べ、さらにそれらを組み合わせる方法について述べる。第4章では、RH 戦略に基づいたフレーム決定法について述べる。第5章では、フレームごとの対象情報が大きくずれない動画像の例として「ビタゴラ装置」を用い、動画像再構成における RH 戦略の有効性についての検討を行なう。最後に、第6章でスポーツ競技動画への提案手法の適用事例とその考察を行なう。

2. 競技における動作と動画像撮影のポイント

固定カメラによる動画像を用いて視聴価値を向上し良質な

Fig. 1 Sports competition video by fixed cameras

動画像を作成する上で、競技特有の撮影上のポイントについて考えることが必要になる。本章では、空手競技中において生じ得る動作と、動画像撮影時のトリミング方針との関係について考える。本論文において競技映像撮影時には、Fig. 1のように、定点固定カメラを用いて競技場を中心として一方もしくは多方向から競技者を撮影する場合を想定する。これらの動画像の組み合わせや、仮想ズームなどを施すことで良質な動画像の構成を実現する。

本論文では、スポーツ競技の1例として空手を用いるが、1対1競技であれば類似の拡張で手法を適用できることから、剣道などの他競技への応用は難しくないことに注意する。

2.1 空手競技開始前

空手競技では1対1での対戦がなされ、選手2名および審判1名が競技場内に存在することになる。競技開始前は両選手ともに場外から入場し開始線で静止する。その後、審判の合図で競技が開始されることになる。競技開始までの時間はズームアウトした領域での撮影に固定することで初期フレームを固定とすることが妥当であるといえる。

2.2 空手競技の競技中

競技中は両選手ともに主に動き回ることとなる。通常のカメラマンによる撮影では被写体を中央に捉え続けるようにトラック動作をするだけでなく、被写体の進行方向側にスペースを広く取ることが良いと知られている¹⁸⁾。したがって、両選手が共に同一の方向に移動するような場合は、両選手を中央に捉えつつ進行方向側にスペースを開けるカメラワークが有効であると考えられる。

2.3 ポイント取得による競技中断

ポイント取得時には競技が一時中断する。中断した競技は審判の合図で再開されることになるが、再開までの時間は競技開始前と同様にズームアウトした領域として固定するほうが視聴者の立場から妥当であると考えられる。

2.4 フレームアウトの頻度と動画像の質

試合時間中の一部の時間であれば、選手の片方にフォーカスし、1名しかフレームに映らない状況も許容されると考えられる。その一方、両名ともにフレームアウトした時間が一定時間続くような場合には視聴者にとって動画像の質は悪いものと判断される。

Fig. 2 Example of sports competition videos

3. 競技者の検出方法

Fig. 2 に競技中のフレーム画像の一例を示す。Fig. 2 のように、動画中には中央で対戦している 2 名の選手以外にも、審判や観客などが同時に映り込んでいる。このような動画画像からフレームの再構成を行なう際には、競技を行なっている選手の領域のみを抽出することが必要になる。

本論文においては、競技者について、選手領域の重心位置抽出と、選手領域面積の抽出の 2 つの値の抽出を考える。重心位置を特定することは、注目ポイントをどこに配置するかを決定する上で重要である。また、選手領域面積はカメラから選手までの近さとの相関が高いものと考えられることからズーミングを行なう上で重要な情報になる。

以上を踏まえた上で本論文では色領域の抽出とフレーム間差分の 2 通りの方法について説明し、それらを組み合わせた手法を示す。色領域抽出とフレーム間差分の手法は一般によく用いられる検出手法であり、OpenCV の標準関数の中に組み込まれていることから容易に実装が可能である。

3.1 色領域の抽出

競技者の領域を抽出する方法の 1 つとして色を用いた抽出について説明する。道着の色はあらかじめ既知であるものと仮定できるため、その色領域を抽出することで選手の位置の推定が可能となる。ここで、元画像は通常、赤・緑・青を組み合わせで色を表現する RGB 色空間を用いて各色を表現している。RGB 色空間の場合、直感的に何色かを判定することが難しい。このことから色空間を「色相」、「彩度」、「明度」で表現する HSV 色空間へと変換した上での処理を行なう。まずは取得した RGB 情報から HSV 色空間への変換の後、H, S, V それぞれの要素について閾値処理を行ない選手領域の抽出を行なう。

例として用いる Fig. 2 では道着の色は白色である。白色は、HSV 色空間では任意の色相で且つ彩度が低く明度が高い色として定義されている。したがって、白色の場合は S と V のみで閾値処理を行えばよい。Fig. 2 に対して、白色領域のみを抽出した画像を Fig. 3 に示す。まず、競技者 2 名ともに白色領域として抽出できていることが確認できる。一方、カメラの全領域内には試合をしていない選手や、道着と同色の服を着ている人物が存在する。そのため色抽出だけでは競技者領域の認識には不十分であり、その抽出領域を適切に選定す

Fig. 3 White color extraction result from competition video

Fig. 4 Inter-frame difference: comparison of two images

ることが必要になる。

3.2 フレーム間差分

第 3.1 節とは別の領域の抽出方法として、数フレーム前と現在のフレームの差分を取るフレーム間差分が知られている。(1) 式に n フレーム前画像を用いたフレーム間差分の原理を示す。

$$I_{diff}(x, y, t) = I(x, y, t) - I(x, y, t - n) \quad (1)$$

ここで、 x, y は各画像の画素の座標、 I_{diff} は差分画像、 $I(x, y, t)$ は時刻 t での画像、 $I(x, y, t - n)$ は n フレーム前の画像を表わす。(1) 式は差分画像が現時刻の画像と n フレーム前の画像との差であることを示している。この差分画像から選手領域を抽出するために、(2) 式で表わす閾値処理を行ない 2 値化する。

$$I_{binary}(x, y, t) = \begin{cases} 255 & (I_{diff}(x, y, t) > T) \\ 0 & (I_{diff}(x, y, t) \leq T) \end{cases} \quad (2)$$

ここで、 I_{binary} は 2 値化された画像であり、 T は閾値を表わす。

Fig. 2 とその 3 フレーム前の画像を Fig. 4 に示す。また、その 2 つの画像の差分を取り、閾値処理した画像を Fig. 5 に示す。

Fig. 5 より、2 つの画像の差分を取ることで選手領域が抽出できていることが確認できる。しかし、色領域からの抽出の場合と同様に、カメラの全領域内には試合をしている選手以外にも動いている人物や物体が存在することが確認される。たとえば Fig. 5 の左側に大きく見える白色部分は審判の差分領域である。このように、この方法においても選手領域のみを完全に検出することはできないことがわかる。選手領域のみを抽出するためには差分を取る領域を適切に選定するだけでなくほかの工夫も必要であることがわかる。

Fig. 5 Inter-frame difference

Fig. 6 Color extraction images

Fig. 7 Combination of color extraction and interframe difference

3.3 色領域抽出とフレーム間差分の併用

第 3.1 節, 第 3.2 節で述べた色抽出とフレーム間差分を組み合わせることにより, 選手領域のロバストな抽出を試みる. 選手領域の抽出手順を以下に示す. 現在のフレームと 3 フレーム前の画像から道着の色を抽出する. その後, その 2 枚の画像の差分を取ることで道着と同色でかつ動いているものを抽出し, 選手領域として定義する. (3) 式に色抽出とフレーム間差分の原理を示す.

$$I_{bd}(x, y, t) = I_b(x, y, t) - I_b(x, y, t - n) \quad (3)$$

ここで, I_b は道着の色を抽出した画像であり, I_{bd} は道着の色を抽出した後フレーム間差分によって得られる画像を表わす.

Fig. 6 に Fig. 4 の画像から白色を抽出した画像を示し, Fig. 6 の二枚の画像の差分を取った画像を Fig. 7 に示す.

Fig. 3 や Fig. 5 と比較して Fig. 7 では選手領域以外の部分が少なくなっている. しかし, 選手以外の白色でかつ動いているもの, 今回は審判のシャツ領域がまだ残っている. よってこの場合においても, ポイント取得の判定領域を適切に選定する必要がある. しかしながら, 競技上領域に限れば取得できている領域は競技者の領域のみであることがわかり, 選手の位置情報の取得が可能であることが Fig. 7 から確認できる.

Fig. 8 Overview of RH control

4. Receding Horizon 戦略に基づいた仮想カメラワークの設計

定点カメラによって撮影された動画をトリミングし, 仮想的なカメラワークを行なう. 本章では, RH 戦略に基づいた設計法について考える. 撮影された動画を使用することから, 現時刻の情報のみでなく過去・未来の情報に基づいて仮想的なカメラワークを行なうことができる点が RH 戦略によって生かされることになる.

4.1 RH 制御の概要

RH 制御の方法に基づいて動画の選択およびトリミング領域の選択を行なうことで仮想的なカメラワークを行なうことを考える. ここで RH 制御とは, 数フレーム先までの区間情報を用いた評価関数を考え, その評価関数に対する最適化問題を解き, その解を利用して現在の時刻で使用する入力解を決める手法である. 逐次的に解を更新することで実用的な解を実用的な計算リソースの下で求めることができ, 制御工学分野で広く用いられている.

本論文で利用する RH 制御の概要を Fig. 8 に示す. Fig. 8 のように現時刻から過去・未来の数フレームを予測区間とし, その情報を元に有限区間での最適化問題を解き, 現時刻での仮想的なカメラワークを最適化する. つぎの時刻では区間をずらし, 再び有限区間での最適化問題を解く. これを繰り返していくことで, 仮想的なカメラワークのすべての時刻における最適化を実行することができる.

各時刻におけるフレームは中心座標 (x, y) と拡大率 l_t により決定することができる. フレームの画面比率は横 × 縦が 16×9 であり, 素材映像である俯瞰視点映像は 3840×2160 画素で表わされる (Fig. 9). 最適化で求める設計変数のフレームへの適用方法を示している. また, 処理を行なう過程で素材映像をリサイズしているため, 実際は 1536×864 画素のフレームをフルショットとして設計変数を決定している.

Fig. 9 Application of design variable to frame

5. ビタゴラ装置動画に対する仮想カメラワーク

5.1 ビタゴラ装置の概要

本章では、第4.1節のRH戦略を用いた仮想カメラワーク手法の有効性について検証する目的で、動体領域の抽出が比較的容易に行なえ、外乱などの少ない映像例としてビタゴラ装置の動作を撮影した素材映像を用いて仮想カメラワークの検証実験を行なう。特に本章ではRH戦略を用いた場合について、トリミングの時間的な変動に着目して検証を進める。

ビタゴラ装置はループ・ゴールドバーク・マシンとも呼ばれ、アメリカ合衆国のRube Goldbergが発案した表現手法で、たとえば転がってきた物体がつぎの物体に衝突して運動が引き継がれるような、最初にきっかけを与えた以降は連鎖的な運動のみで進行するからくり装置のことを指す。本章で用いる素材映像は、NHK Eテレの番組「ビタゴラスイッチ」を俯瞰視点の固定カメラで撮影した映像と考えると容易に想像できる。テレビ番組の「ビタゴラスイッチ」での動画は、多視点の非固定カメラで撮影され、人の手によって編集されている。これに対して、本論文では固定カメラ動画からの再構成を行なう。具体的には、装置の一連の運動を初めから終わりまで俯瞰視点の固定カメラで撮影して、その映像に提案手法を適用することによってビタゴラスイッチと同様に動体を追跡し、さもカメラワークや編集がされているかのような動画を再構成することがここでの目的である。

仮想カメラワークを実装するにあたって、素材となる映像を高解像度の固定カメラで撮影する。ビタゴラ装置全体がフレーム内に入り、動体はフレーム外には出ないことを前提として撮影を行なう。この条件下で撮影した動画を用いることで、注目領域の決定を容易に実現可能にしている。

Fig. 10に本論文で用いるループ・ゴールドバーク・マシンの外観を示す。まず、ドミノを倒すところからスタートし、複数のビー玉の連鎖を介してゴールにたどり着くものである。動体は「ドミノ」、「スティックのり」、複数の「ビー玉」の3種類であり、大きさがほとんど一定している。ビー玉が動く場所において動体の大きさはある程度固定されており、検出された動体領域のパラッキの検証が容易な装置構造である。

5.2 評価関数の設定

本章において、評価関数 E は次式のように設定している。

Fig. 10 Entire frame of Rube Goldberg Machine

Fig. 11 The trajectory of COG of moving point

$$E = w_1 \sum_{i=0}^n |\alpha s_{t+i} - l_{t+i}| + w_2 \sum_{i=1}^n (l_{t+i} - l_{t+i-1})^2 \quad (4)$$

ここで、評価区間として $n = 27$ を用いている。 w_1 および w_2 は重みであり、 α は1より大きな定数である。 s_{t+i} は時刻 $t+i$ における動体領域の大きさを表わしている。評価関数 E における第1項は、トリミング領域における動体領域の面積比率を α となった場合に値が小さくなる項であり、第2項はトリミング領域の大きさの時間変動が少ないと評価値が小さくなるような項である。第2項は、拡大率が急激に変化すると非常に見にくい動画になってしまう¹⁹⁾ため導入している。この評価値 E が小さくなるように $l_{t+i}, i = 0, \dots, n$ を最適化することになる。

なお、本章における固定カメラ動画では、動体領域の時間的な不連続性がなかったことから、最適化のための変数に画像の中心座標 (x, y) は含めていない。

5.3 検証実験と考察

Fig. 11に重心座標の軌跡を示す。Fig. 11から、第3章の手法によって座標の小刻みな振動や急激な変化はほとんどない軌跡が得られていることがわかる。

Fig. 12は、5つの時点におけるフレーム例を示している。つぎに、Fig. 13には、各フレームごとに l_t の最適化を行なった手法と、約1秒 ($n = 27$) の評価区間を設定してRH戦略を施した提案手法のそれぞれの拡大係数 l_t の推移を示している。なおフルショット時の l_t の値は $l_t = 100$ と設定している。

両者を比較すると、フレームごとの最適化では過剰な拡大縮小や小刻みな拡大係数の変化が生じているが、提案手法ではこれらの変動が抑制されていることが確認できる。すなわち、区間情報を用いて最適なフレーム決定を行なうRH戦略によってフレーム決定を行なうことで、第3.3節の抽出手法

Fig. 12 Re-construction result by proposed method

Fig. 13 Time series of l_t

では吸収できなかったフレーム検出能力のバラツキの影響を抑えていることがわかる。

本研究では、競技映像の再構成に関する研究の端緒として動体領域の抽出が容易であるという前提条件の下で RH 制御を応用し、フレーム決定を評価関数によって行なう手法を提案した。俯瞰視点の固定カメラで撮影した映像からフレーム間差分法により動体を抽出し、現在時刻から未来（過去）の動体に関する情報を用いて評価関数を設計し、現在時刻のフレームを決定することで違和感の少ない映像を生成することを目的として研究を行なった。

本研究を行なったことで、RH 制御を応用し評価関数によって動画の再構成をすることが可能であることが確認できた。特に、Receding Horizon Policy に基づいて評価関数を設計することで、1 フレームごとの最適化では実現できない滑らかなカメラワークを実現できた。

なお、本章の実験結果は URL ²⁰⁾ の動画で確認することができる。

6. 空手競技動画に対する仮想カメラワーク

6.1 動画の撮影方針

ここでは、第2章で示した空手競技動画に対する仮想フレームワークにおけるトリミングの方針を考慮した評価関数を設定し、提案手法を用いることでポイントを押さえた動画の再構成を行なう。本節では、空手競技動画の撮影の方向性について述べる。本実験においては、三脚を用いて固定された撮影環境での小学生の空手競技動画を用いた。この動画

Fig. 14 Comparison between games and Point acquisition

像においては審判の映り込みはあるが、第3章の方針により選手と審判の識別が可能な動画である。

通常のカメラワークやスポーツの国際映像などにおいて、同一フレームを用いた映像表示時間は数秒から数十秒程度であり、均一的な値とは限らない。しかしながら、たとえば1秒ごとにカメラを切り替えるなどした場合、映っている映像の良し悪しにかかわらず視聴しづらいものとなり、質の悪い動画になる。よって、競技動画ではサンプリング周期は1秒程度で十分であり、すべての時間フレームを最適解の導出に使用する必要はない。

6.2 ポイント取得による競技の中断判定

第2.3節で述べたとおり、ポイント取得によって競技が中断された場合、ズームアウトして競技場全体を映すのが良いと考えられる。よって競技中にポイント取得によって競技が中断されたかどうかの判定が必要になる。これは、第3.3節で述べた方法を用いて実現している。ポイント取得によって競技が中断されている場合、両選手はあまり動かず所定の位置で待機することになる。そこで、Fig. 14 に示すように、競技場内に白色でかつ動いているものがあれば試合中、そうでない場合はポイント取得によって競技が中断されていると判定している。

6.3 評価関数の設定および提案手法の実装

第4章で述べた RH 制御に基づいた仮想カメラワークの最適化を行なう。仮想カメラワークはフレームの中心の x 座標と拡大率 l_t により決定する。なお、空手動画において、競技場の横からの撮影を行っており上下の動きに比べて左右の動きが大きいことから、ここでは簡単のためにフレームの中心の y 座標については一定とする。

第6.1節で述べた設計指針に基づいて評価関数を設定する。このとき、仮想カメラワークでの動画再構成の目的は視聴しやすい付加価値の高い動画を作成することである。

評価関数 E は以下の (5) 式に設定している。

$$\begin{aligned}
 E = & w_1 \sum_{i=n_1}^{n_2} (G_x(t+i) - x(t+i))^2 \\
 & + w_2 \sum_{i=n_1+1}^{n_2} (l(t+i) - l(t+i-1))^2 \\
 & + w_3 \sum_{i=n_1+1}^{n_2} (f(x(t+i), l(t+i)) \\
 & - f(x(t+i-1), l(t+i-1))) \\
 & + F(t)
 \end{aligned} \tag{5}$$

本節の空手競技動画の再構成において、評価関数の評価区間として $n_1 = -20$, $n_2 = 40$ を用いる。ここで、 $G_x(i)$ は現在処理しているフレームの過去 20 フレームから未来 40 フレームのフレーム i における選手領域の重心の x 座標である。 x は表示する領域の中心の x 座標、 l はフレームの拡大率、 $f(x, l)$ は動画中の動体割合を求める関数である。また、 $w_1 \sim w_3$ は重み係数である。第 1 項目は素材となる動画のどこに注目するかを決定する項である。過去 20 フレームから未来 40 フレームの自乗誤差平均を用いることで、選手がこの先どう移動するかを踏まえた仮想カメラワークによるトラックが可能となる。また、この評価関数設定によって第 2.2 節で述べた前開きのカメラワークに近づけることができる。第 2 項目は、前フレームとの拡大率の変化を抑える項である。拡大率が急激に変化すると非常に見にくい動画画像になってしまうため設定した評価項目である。第 3 項目は、画面に入る動体領域の割合を一定値に近づけるようにしている。画面に入る動体領域割合を関数 f によって特徴付けており、これらの前後フレームでの差が小さくなると第 3 項が小さくなる。これによって、選手の動きがゆっくりの場合は動体領域面積が小さくなることから拡大となるように設定され、選手の動きが速い場合は動体領域が大きくなることから少し引き気味に映すように構成することを可能にしている。

また、詳細は記載しないが抽出面積の関数 F を定義し max 関数に組み込むことで、競技領域内の面積が一定量を下回った場合には中断したものとして $l = 100$ がほかの項目に比べて大幅に小さな値となるよう設定している。この項によって、第 2.3 節の中断に対応する $l = 100$ を RH 戦略による動画の再構成において取り扱うことを可能にしている。

6.4 競技動画をういた検証実験

作成中のシステムを使用して動画の再構成を行なう。素材動画の解像度は 1920×1080 であり、フレームレートは 30 fps である。素材動画を 15 フレーム (0.5 秒) ごとにサンプリングし、その情報を元に RH 制御に基づいた仮想カメラワークの最適化を行なう。本論文では、差分進化法 (Differential Evolution: DE)²¹⁾ を用いて最適化問題を解く。DE は進化的戦略の 1 つであり、非線形問題、微分不可能な問題、非凸問題、多峰性問題などのさまざまな最適化問題に適用されてきており、これらの問題に対して高速で頑健なアルゴリズムであることが示されてきている。なお、本論文はオフラインの動画再構成を行なうため実時間性は必要ないことに注意する。

Fig. 15 に競技開始前の動画を示す。上段が素材動画 (フルショット)、下段が提案手法により作成した動画である。また、上下段共に左から右の画像へと進んでいく。Fig. 15 は、両選手が競技開始のために左右から入場し、審判の合図で礼をする場面である。第 2.1 節で述べたように競技開始前は提案手法の動画において素材動画と同じくフルショットになっていることが確認できる。

Fig. 16 に競技中の動画を示す。上段が素材動画 (フ

(a) Original video

(b) Created video

Fig. 15 At the start of the game

(a) Original video

(b) Created video

Fig. 16 During the game

(a) Original video

(b) Created video

Fig. 17 Point acquisition

ルショット)、下段が提案手法により作成した動画である。また、上下段共に左から右の画像へと進んでいく。Fig. 16 は、右の選手の蹴りをかわすために左の選手が下がる場面である。このとき、下段では第 2.2 節で述べた前開きを模擬し、選手の進行方向側にスペースを開けているようすがわかる。また、すべての場合で両選手が画角に収まる範囲内で拡大されており、トラッキングもできていることが確認できる。

Fig. 17 にポイント取得時の動画中の 1 フレームを示す。上段が素材動画 (フルショット)、下段が作成した動画である。また、上下段共に左から右の画像へと進んでいく。

Fig. 17 は右側の選手の蹴りによってポイント取得した場面である。ポイント取得と審判が判断し、試合を中断させた直後には提案手法ではフルショットになっていることがわかる。

以上より、第2章で述べたカメラワークの方針が実装できていることが確認できる。また、動画像全体を通して不自然な拡大縮小や、選手領域以外に注目していないことを目視によって確認できた。しかし、試合中の動画像の拡大率がほぼ一定であり、第6.3節で述べたような選手の動きによって拡大縮小が行なえていない。また、本論文では注目領域の y 座標は一定としたため、Fig. 17 のように選手がカメラ側によってくる場合、足元が映らなくなってしまうといった問題がある。よって、今後は評価関数の形を工夫することや、注目領域の y 座標を組み込むなどの改良が必要である。

本研究で作成したシステムは RH 制御を用いてカメラワークの決定を行なっている。そのため、多方向からのカメラによる競技動画に対応させることも容易である。その場合、複数の動画像の中から最も評価関数値が低くなる組み合わせを RH 制御を用いて求めることで実現可能である。このとき、評価関数に動画像の切り替え回数に関する項を付け加えることで、細かな動画像の切り替えを防ぎ、見やすい動画像が作成できる。

なお、本章の実験における元動画および仮想カメラワークの結果はそれぞれ URL²²⁾ および URL²³⁾ で確認することができる。

7. おわりに

本論文では、固定カメラで撮影された競技動画に対するオフラインでの動画像の再構成手法を提案した。映像中から色抽出とフレーム間差分を併用することで選手領域の重心と面積を抽出した。また、RH 制御に基づいたカメラワークの最適化手法を提案し、ピタゴラ装置を用いた動画像と空手競技の動画像に適用することで検証を行なった。RH 戦略に基づいた提案手法により、複数のフレーム情報を利用することで良質な動画像を再構成できることを示した。

冒頭でも述べたように、領域設定や特徴設定などを適切に行なえば本手法は空手競技のみならず剣道やバドミントンなどのさまざまなスポーツ競技に適用することができる。また、スポーツ競技のみならず、運動会や学芸会、バンド映像などさまざまな分野への適用が可能である点でも本研究は意義深いものである。評価関数の形をより視聴者の感覚に近いものとして設定し、各競技に適した評価関数構造を設計することが今後の課題である。空手動画像において、 y 軸を固定として RH 最適化を行なったが、選手の表情なども捉えるような枠組みを構築する場合には y 軸方向も設計変数として考慮する必要がある。また、現在は一台の固定カメラ映像での評価および検証を行なっていることから、複数カメラを用いた場合には、トラック動作やパン動作、ズーム動作以外にもワイプやマルチ画面などもカメラワークに必要なってくるため問題はより複雑になる。今後は、複数カメラの映像に RH 戦

略に基づいた本論文の手法を適用することを考えている。

参考文献

- 井口, 土居, 眞鍋, 千原: アクティブカメラの協調によるスポーツ映像撮影システム, 信学技報, IE 100-179, 37/44 (2000)
- 井口, 土居, 眞鍋, 千原: スポーツ映像放送のための実時間映像解析によるマルチカメラの自動制御と自動スイッチング, 映像情報メディア学会誌, **56-2**, 271/279 (2002)
- 醍醐, 小沢: 顔方向追跡に基づいた球技放映カメラのデジタルパン制御～バスケットボールを例として～, 映像情報メディア学会誌, **58-9**, 1299/1306 (2004)
- 宮森, 越後, 飯作: 短時間動作記述を用いた映像シーン表現と検索方式の検討, 信学技報, PRMU98-190, 107/114 (1999)
- 吉田, 北, 小沢: サッカー中継のシーン解析, 電子情報通信学会総合大会, 398 (1996)
- 瀧, 松本, 長谷川, 福村: チームプレー分析を目的としたサッカー動画処理システムの開発, 信学技報, PRU95-27, 49/55 (1995)
- 北, 小沢: バレーボール中継におけるシーン解析, 信学技報, PRU95-242, 47/52 (1996)
- 加藤大一郎: 長野オリンピックにおける新しい映像表現技術「仮想対決システム」, 計測と制御, **38-4**, 281/284 (1999)
- 熊野, 有木, 塚田: ボールと選手に着目したデジタルカメラワークの実現法—デジタルシューティングによるサッカー解説映像生成システムに向けて—, 映像情報メディア学会誌, **59-2**, 271/278 (2005)
- 桐井, 篠木, 藤吉: HD 映像からの視聴者の注目点を考慮した講義映像の自動生成, デジタルコンテンツシンポジウム講演予稿集, **2-3**, 1/6 (2007)
- 大塚敏之: 非線形最適フィードバック制御のための実時間最適化手法, 計測と制御, **36-11**, 776/783 (2001)
- 児島, 大塚: モデル予測制御の考え方, 計測と制御, **42-4**, 310/312 (2003)
- 向井, 畑中, 藤田: ハイブリッドシステムの制御-IV: モデル予測制御, システム/制御/情報, **51-11**, 512/519 (2007)
- E.F. Camacho and C. Bordons: *Model Predictive Control*, Springer-Verlag (1999)
- 端倉, 小浦, 梅田, 児島: 太陽光発電大量導入時の蓄電レベル制御—周期的ロバストモデル予測制御の提案—, 計測自動制御学会論文集, **51-9**, 614/626 (2015)
- 郭, 向井, 川邊, 地形情報を用いたモデル予測型隊列走行制御系の省燃費効果に関する一考察, 計測自動制御学会論文集, **49-7**, 678/687 (2013)
- 永田雅人: 動画像処理ライブラリ OpenCV, 映像情報メディア学会誌, **61-11**, 1602/1605 (2007)
- 加藤, 山田, 阿部, 石川, 尾幡: 被写体を追尾撮影時の放送カメラマンのカメラワーク特性分析, テレビジョン学会誌, **50-12**, 1941/1948 (1996)
- 氏家, 鶴飼, 斎田: 映像酔いに対する運動ボタンと映像コンテンツの影響, 日本バーチャリアリティ学会論文誌, **9-4**, 377/386 (2004)
- 論文用動画 1: RH 戦略に基づく固定カメラを用いた競技動画像の再構成, <https://www.youtube.com/watch?v=XDAE5aqXELE>
- R. Storn and K. Price: Differential Evolution — A simple and efficient adaptive scheme for global optimization over continuous spaces, Technical Report TR-95-012, International Computer Science Institute, Berkeley (1995)
- 論文用動画 2: RH 戦略に基づく固定カメラを用いた競技動画像の再構成, <https://www.youtube.com/watch?v=nApA2eAzHdQ>
- 論文用動画 3: RH 戦略に基づく固定カメラを用いた競技動画像の再構成, <https://www.youtube.com/watch?v=pWO-XXSxaF8>

[著 者 紹 介]

岡 島 寛 (正会員)

2004 年大阪大学大学院工学研究科電子制御機械工学専攻博士前期課程修了。2007 年同博士後期課程修了。同年熊本大学自然科学研究科助教, 2015 年同准教授となり現在に至る。車両の運動制御, 通信系の制御などの研究に従事。博士 (工学)。IEEE, システム制御情報学会の会員。

橋 爪 壮一郎

2017 年熊本大学工学部情報電気電子工学科卒業。現在, 熊本大学大学院自然科学研究科情報電気電子工学専攻に在学中。

大 石 凌 史

2018 年熊本大学工学部情報電気電子工学科卒業。現在, 熊本大学大学院自然科学教育部情報電気工学専攻に在学中。

宮 原 智 新

2009 年熊本大学理学部理学科卒業, 同年, 日本アイビーエム共同ソリューションサービス入社, 現在, (株) diffeasy に勤務。