

スポーツ選手の動きを比較するための動画合成手法

奥山 龍也[†] 石川 知一[†] 柿本 正憲[†]

[†]東京工科大学 メディア学部メディア学科

1. はじめに

近年、CG の分野において画像合成や動画合成は広く研究されており、興味深い映像を提供することができるようになった。我々は、スポーツ選手の動きの違いをわかりやすく示すことで、動きの良い選手とその他の選手でどのような動きの違いがあるかを見つけることができ、スポーツごとに最適化を行うことが容易になると考えた。従来方法では手動で1枚1枚の画像を加工し合成していた。しかしながら、動画の加工は膨大な時間的コストを必要とするため、容易に合成を行うことはできなかった。また、全自動で合成を行うためにはクロマキー合成のように、カメラ撮影時における物理的な制限や工夫が必要だが、スポーツ大会などでは、選手の障害になってしまうため、行うことができない。本研究ではより良い動きを知りたい選手や、スポーツ観戦でより興味深い映像を得たい人のために、実況中継で使用できるよう、カメラ撮影時の物理的制限をつけずに、全自動での動画合成を実現する。

1つ目の制限を設けたのは、比較対象としてまったく違う角度から撮影されたものを合成しても、動作の違いがわからないからである。2つ目の条件は、合成開始フレームを決定するために必要な制限である。3つ目の制限では、本研究は同一物体(図1)を使用し、位置合わせを行うために必要なものである。本研究では、これら3つの制限を満たすアルペンスキーの動画を対象とし、研究を行った。また、アルペンスキーは選手の動作が非常に速く、カメラのパンする速さも比例するため、同一物体を検知することが難しい。トラック競技などは固定カメラや、パンの速さが遅いものが多い。そのため、アルペンスキー動画が合成できれば、トラック競技などの種目にも本研究を応用できると考えた。なお、本稿ではYouTubeに公開されているバンクーバーオリンピック(olympicvancouver2010)の動画[2]を実験素材として利用した。

2. 関連研究

スポーツ動画に限らず、動画の合成はさかんに研究されている分野である。その中でも Rüegg ら[1]の提案する手法では2つの動画を違和感がない境界(シーム)で切り抜き、合成する研究であり、対応できる動画が多く、クロマキー合成のように動画撮影のために必要な条件はない。しかしながら全自動ではなく、あらかじめ同一物体(2つの動画のどちらにも映っているもの)にブラシストロークでマークをつけることで位置を合わせる。本研究では後述する3つの制限を設けるが全自動で合成を行う。



図1 同一物体の例

3. 提案手法

3.1 対象動画の制限

本研究ではカメラ撮影時の制約をつけない代わりに、以下の動画に対する3つの制限を設ける。

- (a) ほぼ同じ角度から撮影されている
- (b) 合成するフレーム指標(タイムなど)がある
- (c) 比較する動画に同一物体が映っている

3.2 全体の流れ

本研究の処理の全体の流れを(図2)に示す。

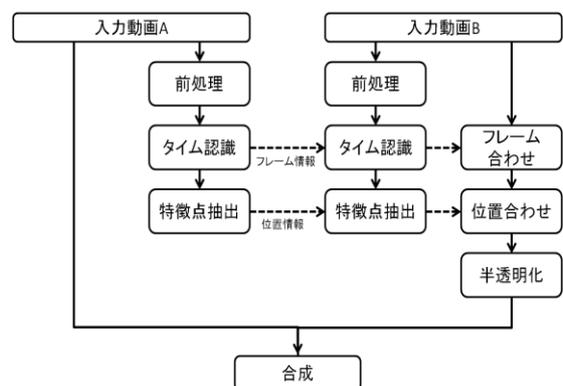


図2 本手法の流れ

“A video composition method for comparing the movements of athletes”

Tatsuya OKUYAMA, Tomokazu ISHIKAWA,

Masanori KAKIMTO

School of Media Science, Tokyo University of Technology,
1404-1 Katakura-machi, Hachioji-shi, Tokyo 192-0982 Japan

3.3 動画の前処理

処理を行う前に、スポーツ動画はサイズが大きいことを想定し、後の処理を高速に行うためにリサイズを行う。その後カメラのノイズや、コース外の物体を除去するためにメディアンフィルタをかけ、2値化処理を施しマスク(図3)を生成する。



図3 マスク画像

3.4 動画のフレーム合わせ

本研究で使用するアルペンスキーの動画はタイムアタック方式の競技のため、必ず画面上に経過タイムが表示されている。このタイムをテンプレートマッチングによって認識し、2つの動画のフレームを合わせる。あらかじめ0から9までの数字をテンプレートとして用意しておき、動画A、Bそれぞれに経過タイムの表示されている右下の領域(図4)にテンプレートマッチングを行う。



図4 タイム表示の例

テンプレートマッチングによって取得したタイムが同じフレームを合成する。

3.5 動画の位置合わせ

本研究での位置合わせは、2つの動画中の同一物体を抽出することにより行う。特徴点抽出には、アルペンスキーの動画にコースとして立っているフラッグやポールを利用する。アルペンスキーの大回転に使用されるフラッグは、赤と青の2種類があり、回転に使用されるポールも、同様に赤と青がある。フラッグを検出する際には、テンプレートマッチングを行い、ポールを検出する際には、直線を検出するハフ変換を行う。また、ノイズを除去するために、色情報を使用する。色情報を使用する際の問題点として、コース外にはスポーツ観戦をしている人などがおり、服などにマッチングしてしまう可能性がある。そこで3.3で述べたマスクを利用し、マッ

グ範囲をしぼる。その後、2つの動画のマスクから重心を割り出す。マスク重心から大まかに位置合わせを行い、抽出したそれぞれの特徴点を、2つの動画間で比較し、最も距離の近かった特徴点を対応付ける。対応付けしたすべての特徴点の距離が最小になるように、合成する動画の移動距離を求める。

4. 結果

提案手法はC++, OpenCV, Lilyを使用し、環境はIntel Quad CPU Q8400 2.66GHz, 4.0 GB RAMのPCで開発をした。使用した動画は2つともフレーム幅1280、フレーム高720、フレームレート30、長さは10秒である。合成処理は3分かかり、中でも特徴点抽出が8割を占める。合成の結果は(図5)に示す。(a)、(b)、(c)は連続したフレームである。

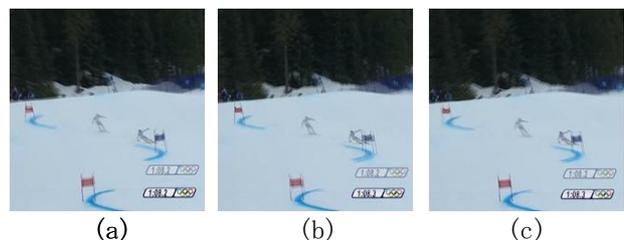


図5 合成結果

5. おわりに

本研究ではスポーツ動画の特徴を利用し、テンプレートマッチングやハフ変換により全自動での動画合成を実現した。違う選手のスポーツ動画を合成することで、動きの比較することができ、より良い動きの提示、スポーツ観戦している人へのサービスとして、興味深い映像を示すことができる。今後の課題は実況中継で実際に使えるような高速性、マッチングの精度、特にハフ変換は精度によってはノイズになってしまう場合があったため、手法の組み合わせや、新たなパラメータを加え、精度を高めなければならないことが挙げられる。また、比較するフレームでズーム倍率が違う場合は、特徴点の位置がずれてしまうため、歪みを補正することを検討している。

参考文献

- [1]Rüegg, Jan, et al. "DuctTake: Spatiotemporal Video Compositing." Computer Graphics Forum. Vol. 32. No. 2pt1. Eurographics, 2013.
- [2]<http://www.youtube.com/user/olympicvancouver2>
010