

# スポーツ連続写真の等身大可視化

武藤駿嗣<sup>1</sup> 新野大輔<sup>1</sup> 井尻敬<sup>1</sup>

**概要.** 連続写真とは、スポーツ等における動作を一目で観察できるよう、撮影された動画から対象のみを切り出して一枚の静止画に合成したものである。本研究では、Mixed Reality (MR) 技術を利用することで、スポーツ動作の連続写真を現実世界に等身大で可視化する手法を提案する。ユーザは、ビデオスルー型のヘッドマウントディスプレイ (HMD) を装着し、HMD に付属したカメラを利用してスポーツ競技者の動作を撮影する。撮影が終わると、得られた動画から競技者の領域が切り出され、MR 空間に等身大のビルボードとして配置される。モニタ上で可視化される通常の連続写真と異なり、ユーザは等身大のビルボードを観察できるため、歩幅・ジャンプの高さなどの競技のスケール感を確認できる。また、我々は、単純な連続写真表示に加え、動画表示、残像付き動画表示という3つの表示法を提供する。提案システムの有用性を示すため、ランニング・立ち幅跳び・テニス・バドミントンといった異なる競技に対して提案法を用いた可視化を行い、等身大で競技者の動作を可視化できることを確認した。

## 1. はじめに

連続写真とは、動きのある被写体を一定の間隔で複数回撮影し、被写体を切り出した下で一枚の写真に合成する表現方法である。例えば、ゴルフスイングの軌跡可視化や走り高跳びの空中姿勢の可視化など、連続写真はスポーツ動作の可視化に用いられることが多い。連続写真には、一枚の静止画で一連の動作を可視化できるという利点があり、スマートフォンにより撮影された写真群から一枚の連続写真を合成する手法[1]なども研究されている。しかし、通常の連続写真は紙やモニタを用いた可視化がされることが多く、実際の動作の現実世界におけるスケール感が実感しにくいという課題がある。

練習支援や新たな観戦法確立のため、Mixed Reality (MR) 技術を用いたスポーツの可視化に関する研究がなされている。Matsui ら[2]や Han ら[3]は、スポーツ動画からカメラパラメータを推定し、異なる視点の映像を合成する手法を提案した。また Inamoto ら[4]や Rematas ら[5]はこのアイデアをさらに進め、サッカーの試合をテーブル上の自由視点映像として表示する手法を提案した。しかし、これらの手法は、テーブル上やモニタにより競技者の動きが可視化されるため、実際の競技動作のスケール感を体感することは難しい。

そこで本研究では、スケール感を実感できるスポーツ動作のリプレイ法の確立を目指し、MR 技術を用いたスポーツ連続写真の等身大可視化法を提案する。図1の通り、ビデオスルー型のヘッドマウントディスプレイ (HMD) を装着した観察者が、HMD に搭載されたカメラにて競技者の動作を撮影する (図 1a)。撮影された一連の画像は、MR 空間内にビルボードとして配置される (図 1b)。本研究では、連続写真表示・動画表示・残像付き動画表示という3種類の表示法を提供する。競技者の動作を可視化するビルボードは、実際に動作が行われた場所に等身大で可視化され、観察者は好きな視点・距離からビルボードを観察で

きる。間近で等身大の競技者のビルボードを観察することで、ジャンプの高さやステップの歩幅等、実際の距離感を実感しながら動作を観察することが可能となる。

提案手法の有用性を評価するため、ランニング、立ち幅跳び、テニスなどの動作を撮影し、提案手法により可視化を行った。結果、競技者の動きをMR空間内にビルボードとして正しく配置し、表示できることを確認した。結果の一例を図1に示す。

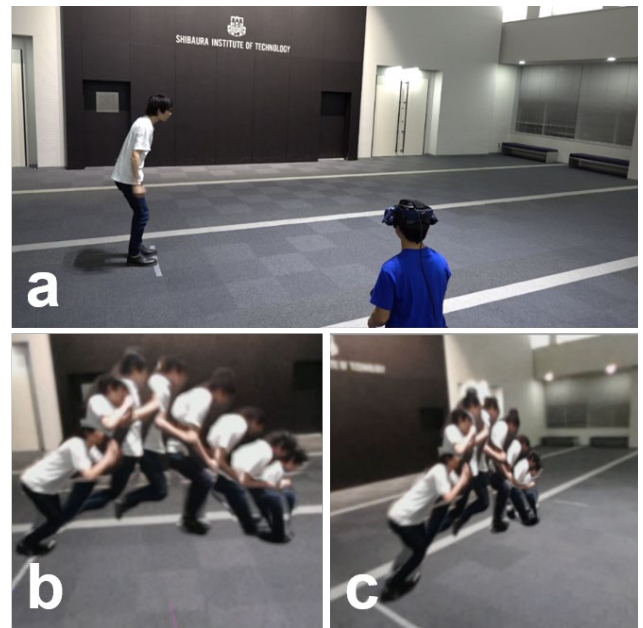


図1. スポーツ連続写真の等身大可視化。観察者はHMDに搭載されたカメラにて競技者の動作を撮影する(a)。競技者はビルボードとしてMR空間に等身大で配置され(b)、好きな視点から観察できる(c)。

## 2. 関連研究

スポーツの練習支援や新たな観戦方法を確立するため、Virtual Reality技術を活用したスポーツ動作の可視化法が研究・開発されている。ここでは、本研究に関連の深い自由視点映像とスマートミラーシステムに焦点を絞り、関連研

<sup>1</sup> 芝浦工業大学

究を紹介する。

**スポーツシーンの自由視点映像化.** 単眼カメラや複数カメラにて撮影された映像から、ユーザが自由に視点を移動できる自由視点映像を構築する技術が研究されている。Matsui ら[2]や Han ら[3]は、単一のスポーツ放送動画からカメラの位置や内部パラメータを求め、仮想的にカメラ位置を変更できる手法を提案した。また、Inamoto ら[4]は、サッカー映像からサッカーの自由視点映像を構築した。加えて Rematas[5]らは、自由視点映像をテーブルトップ上で可視化した。これらの自由視点映像システムは、ユーザが好きな角度からスポーツを観戦できるという利点はあるものの、モニターやテーブルトップ上での可視化を想定しており、競技のスケール感を体感できるものではない。

**スポーツ練習用のスマートミラーシステム.** スマートミラーとは、鏡と計算機を組み合わせることで、鏡の前の人の様子を写すと同時に様々な情報を提示するシステムのことである。鏡とモニターを組み合わせる実装方法や、カメラとモニターを組み合わせる実装方法が存在する。ユーザをほぼ等身大で表示すると同時に、手本となる姿勢やユーザの姿勢情報なども表示できるため、スポーツ練習支援に利用可能である。実際、ダンス [6]、テニス[7]、バレエ[8]などの練習のためのスマートミラーシステムが提案されている。しかし、スマートミラーシステムは、視点方向が限定されており、自由に近づいたり方向を変えたりしながら観察することはできない。

### 3. スポーツ動作の等身大可視化

#### 3.1 提案システムの概要

本研究では、スケール感を実感しやすいスポーツ動作のリプレイ法の確立を目的とし、MR 技術を用いた可視化法を提案する。提案システムにおいて、ユーザはビデオシーンスルー型の HMD を装着し、競技者から 1~3 メートル程度離れた位置から競技者の動作を観察する(図 2)。ユーザは、手に持ったコントローラの録画ボタンを押しこむと、ボタンを押している間、動画が録画される。録画が終了すると、システムは、動画像中から競技者の領域を切り抜き、一連のフレームをそれぞれ平面上の板に張り付けたビルボードとして MR 空間内に配置する。その後ユーザは、MR 空間を移動し好きな視点から、このビルボードを観察できる。また、一度撮影したビルボードは記録され、あとから MR 空間に読みだすことが可能である。

本研究では、上記の MR 環境の実現に HTC Vive Pro を利用した。このデバイスを選択した理由は、現在利用可能な光学シーンスルー型の HMD よりも画角が広く、また、計算能力の高い計算機と併用できるためである。このデバイスは、前面搭載された 2 個のカメラから、それぞれ解像度 1150×750 の画像を 30fps で録画し、これを解像度 2880×1600 のディスプレイに直接表示することでビデオシー

スルー型の HMD として機能する。提案システムでは、競技者の撮影もこの前面カメラを利用して行う。そのため、見たままのものを撮影できる利点があるものの、撮影される画像の解像度が低いという課題がある。解像度の高いカメラを追加で搭載し、連続写真の高解像度化を行うことは将来課題である。



図 2. 撮影環境.

#### 3.2 競技者可視化のためのビルボード生成

提案システムの実現には、撮影された一連の画像から人物領域を分割しビルボードとして 3 次元空間に配置する必要がある。本研究では、Mask R-CNN[9]を利用し、人物領域を分割する(図 3a)。さらに、分割された前景領域の周囲 5 画素の範囲で透明度(アルファ値)を線形に減衰させることでスムーズな人物画像を生成する。

次に、領域分割した画像をビルボードとして MR 空間に配置する手法を解説する。観察者と競技者は同一の床面  $y = 0$  に立っており、HMD のカメラは、ほぼ水平方向を向いていると仮定する。撮影されたある画像に対し、MR 空間における撮影時のカメラ位置を  $\mathbf{x}_c$ 、カメラの姿勢を表す回転行列  $\mathbf{R}_c$  とする。画像内の人物領域の最も下部の画素位置(足の位置)を  $(u_0, v_0)$  とする。利用したカメラがゆがみのないピンホールカメラであると仮定すると、画素  $(u_0, v_0)$  の MR 空間における位置は、 $\mathbf{x}_0 = \mathbf{R}_c(u_0 p_x f, v_0 p_y f, f)^T + \mathbf{x}_c$  と表せる。ただし、 $p_x$  と  $p_y$  は、それぞれ、水平と垂直方向の画角に関する係数であり、 $f$  は焦点距離である。我々は、カメラ位置  $\mathbf{x}_c$  から足の位置  $\mathbf{x}_0$  の方向へ直線を生成し、その直線と床面  $y = 0$  との交点を求める。最後に、求めた交点位置に撮影時のカメラ方向を向くようなビルボードを配置し、これに画像を投影する(図 3b)。

ただし、すべてのフレームに対して足位置に対するレイキャストを行うと、ジャンプ等により競技者が空中にいる場合に正しい位置を取得できない。そこで、本研究では、撮影の開始・終了時点は、競技者は床面に接していると仮定し、開始・終了フレームのみ前述の方法でビルボード位置を決定する。中間のフレームについては、開始・終了フレームにおける深度値を線形補間することでビルボード位置を決定する。

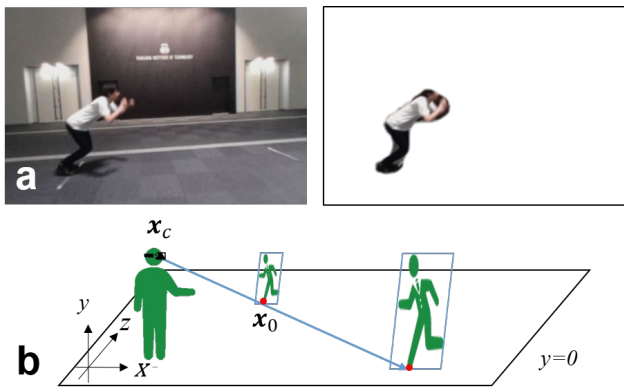


図 3. ビルボードの生成法. 動画の各フレームに対して Mask R-CNN[9]により領域分割を施す(a). 動画撮影時に MR 空間におけるカメラの位置・姿勢を記録しておく, これを利用して床面に対してレイキャストを行うことでビルボードの位置を決定する(b).

### 3.3 ビルボードの表示

提案システムでは, ビルボードの表示方法として連続写真表示・動画表示・残像付き動画表示を提供する. 連続写真表示では, 生成した時間的に密なビルボード群(図 4a)から一定時間間隔で選出されたビルボードが表示される. これにより, 既存の連続写真と同様のものが MR 空間に提示される(図 4b). 動画表示は, 一連のビルボードを順番に表示するもので, MR 空間における動画のような効果が得られる(図 4c). 最後に, 残像付き動画表示は, 動画表示と同様にビルボードを順に表示するが, 一度表示したビルボ

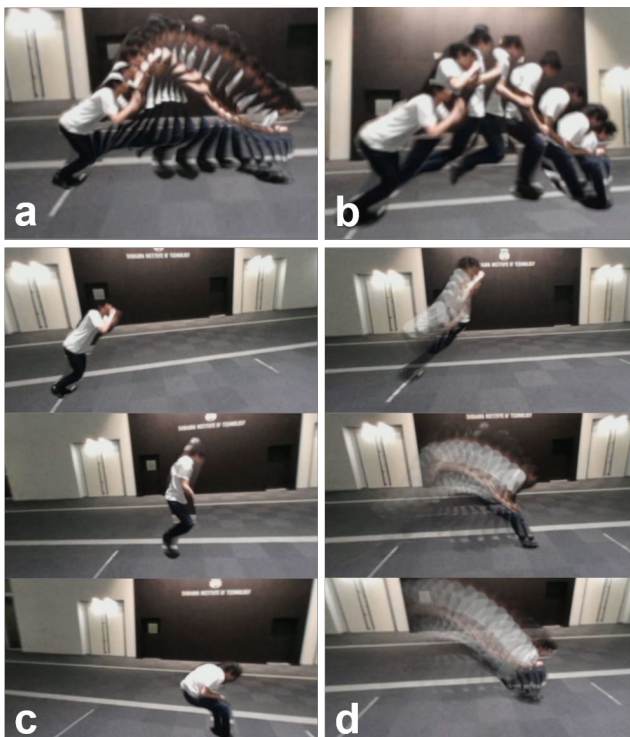


図 4.3 種類の表示法. すべてのビルボード(a), 連続写真表示(b), 動画表示(c), 及び, 残像付き動画表示(d)の表示例.

ードに対しその透明度が徐々に減衰する効果を加える. これにより, 移動の軌跡が残像として残るような効果が生まれる(図 4d). ユーザは手に持ったコントローラにより, この表示モードを切り替えることが可能である.

## 4. 結果と考察

提案手法の有用性を確認するため, ランニング, 立ち幅跳び, 軟式テニス, バドミントンについて, 提案法による可視化を行った. 結果を図 5, 及び, 添付する動画に示す. すべての例において, 実際に動作を行った場所とほぼ同じ位置に, ビルボードが生成されることを確認した.

図 5b は, 立ち幅跳びの例である. この例のようにジャンプを含むものでは, ビルボードに近寄って自身の体と比較することで, ジャンプの高さや距離を確認できる. また, 図 5c は, バドミントンにおいて自身の前方のシャトルを拾う動作である. この例では, ラケットを振る直前のステップが大きくとられていることがわかる. この連続写真が MR 空間に表示されるため, 同じ位置を歩いてみることで実際のステップの距離を確認できる.

提案システムの主な課題は, 計算時間と画像の質である. 現在, Unity 上で動作する比較的低速な Mask R-CNN を利用して人物領域の分割を行っているため, 100 枚程度の画像の領域分割に 5 分ほど時間がかかる. この部分を高速化し, 撮影後すぐにビルボードが表示されるように改良する予定である. また, 現在の提案システムは, HTC Vive Pro に搭載されたカメラにより画像を取得する. このカメラは広角であるため MR 環境実現には適しているが, 人物領域は画像全体からすると比較的小さいため, 高画質な連続写真は得られない. 今後, 人物部分を撮影するためのカメラを付加し, 連続写真の高品質化を行いたい.

## 5. まとめ

本研究では, スポーツ動作のスケール感を実感しやすい可視化法の実現を目指し, スポーツ動作を等身大の連続写真として表示する可視化法を提案した. 提案システムは, ビデオスルー型の HMD を用いて実装されており, 動画像から人物領域を切り出し, これをビルボードとして MR 空間に配置することで, 等身大の連続写真可視化が実現される. 多様な観察方法を適用するため, 本研究では, 連続写真表示・動画表示・残像付き動画表示といった 3 種の表示法を提供した. また, 複数のスポーツ動作について提案法により等身大可視化を行い, その有用性を確認した. 提案システムは, 特に計算時間と表示画像の質において改良の余地があり, 高速化・高品質化が重要な課題である.



図 5. 提案手法による可視化例. ランニング(a)・立ち幅跳び(b)・バドミントン(c)の連続写真表示をそれぞれ別視点から見た結果, テニス(d)の残像付き動画表示の結果.

**謝辞** 本研究のユーザスタディに参加していただいた皆様に感謝の意を表す. 本研究は, 日本学術振興会科学研究費助成金基盤 C(18K11606)の支援を受けて行われたものである.

## 参考文献

- 1) K. Hasegawa and H. Saito, "Synthesis of a stroboscopic image from a hand-held camera sequence for a sports analysis," *Computational Visual Media*, vol. 2, no. 3, pp. 277–289, 2016.
- 2) K. Matsui, M. Iwase, M. Agata, T. Tanaka, and N. Ohnishi, "Soccer Image sequence Computed by a Virtual Camera," In *Proc. Computer Vision and Pattern Recognition*, pp. 860-865, 1998.
- 3) J. Han, D. Farin, and P. H. N. de With, "A real-time augmented reality system for sports broadcast video enhancement," In *Proc. ACM Multimedia*, 2007, pp. 337–340.
- 4) N. Inamoto and H. Saito, "Free viewpoint video synthesis and presentation of sporting events for mixed reality entertainment," In *Proc. 2004 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology ACE '04*, 2004, pp. 42–50.
- 5) K. Rematas, I. Kemelmacher-Shlizerman, B. Curless, and S. Seitz, "Soccer on your tabletop," In *Proc. Computer Vision and Pattern Recognition*, pp. 4738–4747, 2018.
- 6) J.C.P. Chan, H. Leung, J.K.T. Tang, and T. Komura, "A Virtual Reality Dance Training System Using Motion Capture Technology," *IEEE Trans. Learning Technologies*, vol. 4, no. 2, pp. 187- 195, 2011.
- 7) M. Oshita, T. Inao, T. Mukai, and S. Kuriyama, "Self-training system for tennis shots with motion feature assessment and visualization," *International Conference on Cyberworlds*, pp. 82–89, 2018.
- 8) F. Anderson, T. Grossman, J. Matejka, and F. George, "YouMove:

enhancing movement training with an augmented reality mirror," In *Proc. User Interface Software and Technology*, pp. 311–320, 2013.

9) K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask R-CNN," In *Proc. International Conference on Computer Vision*, 2017.