

研究論文

動作学習のための遅延同期ビデオフィードバックシステム

三上 弾^{1,a)} 松本 鮎美¹ 門田 浩二² 川村 春美¹ 小島 明¹

受付日 2013年8月22日, 採録日 2014年1月25日

概要: 本稿では, スポーツをはじめ様々なシーンで行われる動作の学習を支援するシステムとして, 遅延同期ビデオフィードバックシステムを提案する. 動作の学習において, 練習者が試行直後に, 手本となる動作を参照しながら視覚的なフィードバックを受けることは学習効率向上のためにきわめて有用である. これまでも, 練習者の動作と対応した手本動作を提示するシステムが提案されてきているが, 動きの認識にモーションキャプチャを利用するなど, スポーツ愛好者の日常利用は困難なものであった. 本稿では, 撮影条件に厳しい制限を加えることなくカメラ 1 台の入力のみで利用可能, 幅広い動作を対象とすることが可能, さらにタブレット端末でリアルタイム処理可能な, 遅延同期ビデオフィードバックシステムを提案する. これらの特長を実現するため, 提案システムでは, 手本動作と練習者の動作との同期に, Motion History Image (MHI) を特徴量として用いる. 実験により, 提案手法では, 投球, バスケットボールのフリースローおよびレイアップシュート, 側転など幅広い動作について, 動作者の変化, 服装の変化, 照明環境の変化に対してロバストに精度良く同期ビデオフィードバック可能であることが確認された.

キーワード: 動作学習, 遅延同期再生, Motion History Image, MHI

A Video Feedback System Providing Motions Synchronized with Reference Examples for Motor Learning

DAN MIKAMI^{1,a)} AYUMI MATSUMOTO¹ KOJI KADOTA² HARUMI KAWAMURA¹ AKIRA KOJIMA¹

Received: August 22, 2013, Accepted: January 25, 2014

Abstract: This paper proposes a video feedback system for motor learning which provides a video feedback synchronized with reference example with a delay of a few second. For motor learning, it is quite effective that a trainee checks the self-motion by comparing to the reference motion which is synchronized with the self-motion. So far, some systems realized a synchronized video feedback system by using a marker-based motion capture system for motion synchronization. However, it is difficult to be popularized due to the severe limitation to the capturing environment and the high cost operation of motion capture system. The proposed method provides a video feedback system that works with no strict limitation to the video capturing environment, with a input from one camera, can be applied to wide variety of motions, and works with tablet devices. To achieve these traits, it uses motion history image, or MHI, as a feature for motion synchronization. Experiments verified that the proposed method achieved precise synchronization of motions of different actors, different clothes, and under different illumination conditions.

Keywords: motor learning, synchronized playback with delay, motion history image, MHI

1. はじめに

スポーツ・ダンスをはじめとして, 我々人間は, 様々なシーンで, 様々な動作を学習する. 赤ちゃんが歩行を学習したり, 子どもたちが体育の授業で新しい運動を覚えたり, さらには, リハビリなどでかつてはできた動作を再学習したりと, 動作の学習は人生の様々なステージで起こりうる

¹ NTT メディアインテリジェンス研究所
NTT Media Intelligence Laboratories, Yokosuka, Kanagawa
239-0847, Japan

² 大阪大学大学院医学系研究所
Graduate School of Medicine, Osaka University, Toyonaka,
Osaka 560-0043, Japan

^{a)} mikami.dan@lab.ntt.co.jp

課題といえる。我々は、このような動作の学習を効率化することはきわめて意義深いものと考え、誰もが日常的に利用可能かつ効果的な動作学習促進システムを目指した研究を行っている。

スポーツ科学の分野では、適切な視覚的フィードバックにより動作学習が効果的になることが知られている。最も単純な視覚的フィードバック方法として鏡の利用があげられる。実際、鏡の前での反復練習は古くから行われている。しかしながら、鏡を通じた視覚フィードバックには2つの問題がある。第1は確認可能な視点が限られる点である。第2は、動作中に確認を行わなければならないために動作が影響を受ける可能性があるという点である。

このような問題を解決するために、ビデオフィードバックが活用されており、Golf [1], 体操競技 [2], テニス [3] などへの適用で効果が検証されている。本稿では、ビデオフィードバックの中でも特に、動作学習に特に適すると考えられる、短時間の遅延を加えて再生するシステム（たとえば、siliconcoach社 TimeWarp [4] あるいはダートフィッシュ・ソフトウェア [5] など）に注目する。これらのソフトウェアは、練習者がカメラの前で行った動作を、短時間の遅延を加えて再生する機能を有する。これにより、鏡のように視点の制約を受けることなく、動作完了後すぐに練習者自身の動作を確認できるという点でスポーツの動作学習にきわめて適したシステムとされ、鉄棒 [6], フライングディスク [7] などで良好な指導実績が報告されている。

さらに、視覚的フィードバックにおいて、手本動作と練習者の動作が同期再生され比較可能なことは、反復練習の段階において重要である。そのために、Chuaらは武道の一種であるTaiChiを対象とし、練習中の身体姿勢をモーションキャプチャによりリアルタイムに取得し、あらかじめモーションキャプチャにより取得したエキスパートによる動作（手本動作）とを時間的に同期して表示するシステムを提案している [8]。またChoiらは、居合道を対象として、モーションキャプチャにより取得した練習者の姿勢時系列と手本動作の姿勢時系列との比較により練習者の習熟度を推定した [9]。これらのシステムは、モーションキャプチャによる姿勢推定により、手本動作と練習者動作との比較によるフィードバックを可能としている。しかしながら、モーションキャプチャを利用しての姿勢推定には大きな問題がある。1つは設置環境およびコストである。モーションキャプチャは高価であり、誰もが日常的に利用可能な場所に常設することは難しい。そしてキャリブレーションの必要性から設置自体も高コストである。もう1つの問題はモーションキャプチャにより得られる姿勢がきわめて高次元であるという点である。モーションキャプチャでは、用途により異なるが一般に20から50点程度のマーカをつけて三次元位置座標を取得するため、姿勢は1時刻あたり60から150次元の情報となる。前述の文献 [8], [9] の

ように特定の動作を対象とした場合には、モーションキャプチャにより得られた高次元の姿勢情報を、動作を特徴づける低次元の特徴量へと変換することが可能である。しかしながら、提案手法のように汎用的な動作を対象とした同期ビデオフィードバックを目指す場合には、動作依存の特徴抽出は困難である。特に、練習段階の動作は手本動作との姿勢に差異があることが想定され、高次元姿勢情報をそのまま利用した動作間の同期は困難となる。

我々は、誰もが日常的に利用可能な動作学習のためのビデオフィードバックシステムとして、1台の単眼カメラ入力による遅延同期ビデオフィードバックシステムを提案する。提案システムは、手本動作、練習者の動作ともに1台の単眼カメラにより取得するというきわめて簡易な構成であり、練習者の動作と手本動作とを時間的に同期させ、短時間の遅延を加えて再生する。提案システムには3つの特長がある。第1は、照明条件・背景の変化、服装の変化など種々の撮影条件の変化に対して頑健に同期再生が可能であるという点である。第2は、学習対象の動作を選ばないという点である。第3は、タブレット端末によるリアルタイムフィードバックが可能であり、電源を確保できない実際の練習フィールドでも利用可能という計算コストの小ささである。これらの特長を生み出すために提案システムでは、手本動作と練習者動作の同期のための動作特徴量としてMotion History Image (MHI) [11]を用いる。評価実験を通じて提案手法が、様々な動作・様々な環境に対して遅延同期再生が可能であることを確認した。

本稿の構成は以下のとおりである。2章で従来研究について述べる。3章で提案手法の詳細を記し、4章で評価実験およびその結果を示す。5章で提案システムの今後の発展について述べた後、6章でまとめる。

2. 従来研究

提案手法と類似のコンセプトを持つシステムに、MVP2000 [10] と呼ばれる商用システムがある。MVP2000は提案手法と同様に、誰もが利用可能な動作学習支援を目標としたシステムであり、その最も重要な機能の1つとして同期ビデオフィードバックを実現している。MVP2000では、音に基づいた同期を行う。すなわち、動作に付随してインパクト音が発生する動作については同期ビデオフィードバックが可能である。たとえば、実際にボールを打つゴルフスイング、実際にボールを打つテニスサーブについては、良好な同期再生が可能とされている。

しかしながら、1章で示したとおり、動作学習は様々なステージにおいて様々な目的で生じるものである。提案手法ではインパクト音の発生する限られた動作のみでなく、様々な動作の学習に対して、誰もが利用可能なシステムを目指す。こうした観点から、利用の際の制限（服装・照明環境など）を極力少なく日常利用を可能とし、音などの動

作に付随する情報でなく動作そのものを利用することで幅広い動作に対して汎用的に利用可能とする，加えて計算コストの小さい動き特徴 (MHI) を利用することでタブレット型端末でのリアルタイム動作を可能とした点が提案手法の大きな貢献である。

3. 提案手法：遅延同期ビデオフィードバックシステム

本稿では，効果的に動作を学習することを目標に，手本動作と練習者の動作を同期し，わずかな遅延を加えて再生する，同期遅延ビデオフィードバックシステムを提案する。図 1 に提案手法の利用イメージを示す。提案システムでは練習者は，PC (タブレット型端末を含む) に接続されたカメラの前で練習を行う。練習者が練習対象動作を行うと，システムが動作を認識し，手本動作と練習者の動作とを同期させ，短時間の遅延を加えてディスプレイに表示する。これにより練習者は，練習動作直後で動作の記憶が鮮明なうちに，自分の動作を，手本動作を参照しながら確認することが可能となる。

提案システムは手本動作登録ステップおよび遅延同期再生ステップの，2つのステップからなる (図 2 参照)。以降，各々のステップについて詳しく説明する。



図 1 提案手法の利用例

Fig. 1 Typical example of the proposed system.

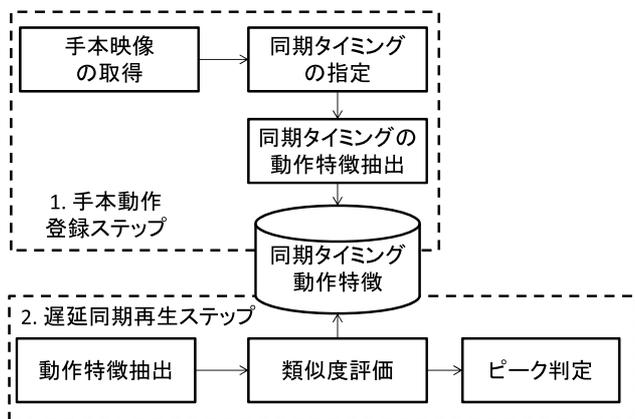


図 2 提案手法の流れ：手本動作登録ステップ (上部) と遅延同期再生ステップ (下部) からなる

Fig. 2 Flowchart of the proposed method: The proposed method consists of two steps, registration of reference motion (upper part) and synchronization video playback with delay (lower part).

3.1 手本動作登録ステップ

提案手法は，背景・照明などの撮影環境変化に対してロバストな手法であり，基本的にどのような映像であっても手本動作として利用可能である。つまり，外部データベースに存在する手本とした動画や，その場で撮影したコーチの動画などを読み込んで手本動作として登録可能である。

手本動作登録ステップでは，まず手本動作を読み込み，続いて同期させるタイミング T_{sync} を設定する。同期タイミング T_{sync} の設定は，動作が本質的に時間的連続性を持つ情報であることから必要となり，同期タイミング T_{sync} で同期した手本映像と練習者映像をフィードバックする。すなわち，手本動作と練習動作の時間発展の速度が異なる場合には，同期タイミング以外の部分において2つの動作は同期しない。実際の練習においては，同期タイミング T_{sync} は練習の目標により変化することが想定される。そのため，提案手法では，目標に合わせた同期タイミングをユーザが手動で設定するものとする。たとえば投球動作において，動作開始からテイクバック部分を主に練習したい場合には，テイクバックを同期タイミングとして設定することができる。一方，テイクバックからボールリリースまでの動作に特に注目したい場合には，リリースを同期タイミングとして設定することが可能である。

次に同期タイミング T_{sync} での動作特徴の抽出と登録を行う。このために，まず同期タイミング T_{sync} において動作領域を設定する。その例を図 3 に示す。そして，動作領域における動作特徴を，手本動作特徴とする。動作領域の設定ではシステム利用者が，動作が生じた領域をマウスクリックにより矩形で指定する。これには，2つの効果がある。第1は動作検出精度の向上である。動作領域を指定することで，仮に背景部分で動きが生じた場合であっても手本以外の動きを含まない手本動作特徴を得ることができる。第2は処理の高速化である。提案手法では動作領域における MHI を手本動作特徴，すなわちテンプレートとし



(a) 動作領域の設定

(b) 手本動作特徴

図 3 (a) 動作領域の設定 (赤矩形) と，(b) 対応する手本動作特徴
Fig. 3 Setting of motion area and corresponding extracted motion feature.

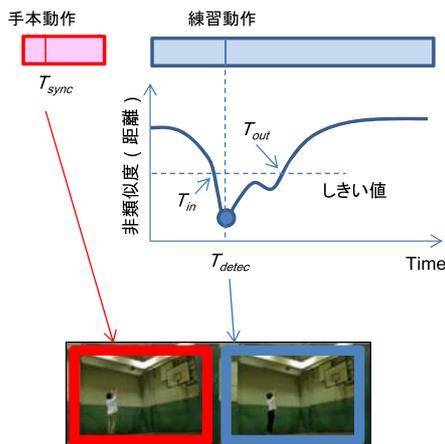


図 4 遅延同期再生ステップ
Fig. 4 Synchronization step.

た距離計算を行う。そのため動作領域の指定によりテンプレートが小さくなることは処理の高速化につながる。これは手本動作作成時に1度だけすればよい作業であり、また、動作領域は手本動作を含む部分をおおよそ含むように設定すればよく、大きな負担にならないと考えている。

3.2 遅延同期再生ステップ

遅延同期再生ステップでは、各フレームにおいて動き特徴を抽出し、登録された手本動作特徴との比較により非類似度(距離)の算出を行う。提案システムでは、非類似度(距離)がしきい値以下となった時刻(T_{in})から、非類似度(距離)がしきい値以上となる時刻(T_{out})のうち非類似度(距離)が最小となったフレームの時刻を同期タイミング T_{detec} と判定し、手本動作の設定された同期タイミング T_{sync} と、練習者動作の検出された同期タイミング T_{detec} とが同時に再生されるようフィードバックを行う(図4参照)。

3.3 遅延同期ビデオフィードバックシステムへの要件と特徴量

ここで遅延同期ビデオフィードバックシステムへの要件を整理し、手本動作と練習者の動作の同期に用いる特徴量および非類似度(距離)算出方法について述べる。提案手法は次の3つの利用条件を想定にしている。

- (1) 服装・照明環境に対してロバストであること、
- (2) 様々な動作に適用可能であること、
- (3) 計算コストが小さく、タブレット端末でリアルタイム動作すること。

以上の要件を満たす特徴量として Motion History Image (MHI) [11] を利用する。MHI の例を図 5 に示す。MHI は、残像のような動き表現および特徴量であり、

$$H(x, y, t) = \begin{cases} 255 & \text{if } d(x, y, t) = 1 \\ \max(H(x, y, t - 1) - \rho, 0) & \text{otherwise,} \end{cases} \quad (1)$$

オリジナル画像列

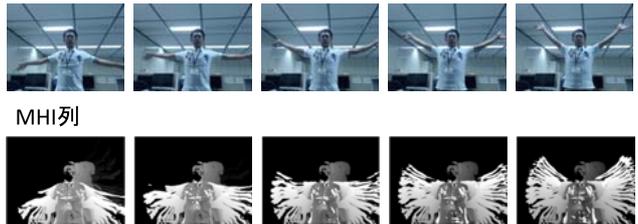


図 5 MHI の例：人物が両手を上に挙げていく動作
Fig. 5 An example MHI for a human raises his arms.

により求める。ここで $H(x, y, t)$ は、画素 (x, y) 、時刻 t でのMHIのピクセル値を表す。また $d(x, y, t)$ は動き検出関数であり、 $d(x, y, t) = 1$ は画素 (x, y) 、時刻 t において、動きが検出されたことを意味する。動き検出関数としては、フレーム間差分などが一般に用いられる(提案手法でもフレーム間差分を用いる)。 ρ は減衰量を表し、小さいほど長く残像が残る。 $\max(a, b)$ は、 a, b のうち大きい方を返す関数である。

式(1)のとおりMHIは、減算とif文のみで構成されきわめて計算コストが小さい。また図5の例からも分かる通り、動き検出関数としてフレーム間差分を用いるため、服装・照明環境にロバストな動き特徴が検出できている。

提案手法では、時間情報を持つ動き特徴(手本動作特徴と練習者動作特徴)がMHI、すなわち1枚の画像により表現される。そこで動き特徴間の非類似度(距離)計算を、テンプレート照合の枠組みにより行う。本稿の範囲では、非類似度(距離)としてユークリッド距離を用いるものとする。時刻 t の距離 $S(t)$ は次の式で求められ、小さいほど類似していることを表す。

$$S(t) = \min s(x, y, t), \quad (2)$$

$$s(x, y, t) = \frac{\sum_{y'=0}^{h-1} \sum_{x'=0}^{w-1} (T(x', y') - I(x+x', y+y', t))^2}{\sqrt{\sum_{y'=0}^{h-1} \sum_{x'=0}^{w-1} T(x', y')^2 \sum_{y'=0}^{h-1} \sum_{x'=0}^{w-1} I(x+x', y+y', t)^2}} \quad (3)$$

ただし、 $s(x, y, t)$ は時刻 t 、探索位置 (x, y) の非類似度(距離)を表す。つまり、各時刻 t で、最も類似した領域との非類似度(距離)を求め、非類似度(距離)とする。 $T(x, y)$ 、 $I(x, y, t)$ はそれぞれテンプレート画像(=手本動作特徴)の座標 (x, y) での画素値、時刻 t における練習者動作特徴の座標 (x, y) での画素値を表す。また、 w, h はそれぞれ手本動作特徴の幅と高さである。

4. 実験

提案手法の有効性を評価するために評価実験を行った。まず提案システムの典型的な動作例を示し有効性について確認を行う。その後、定量評価の結果を示す。

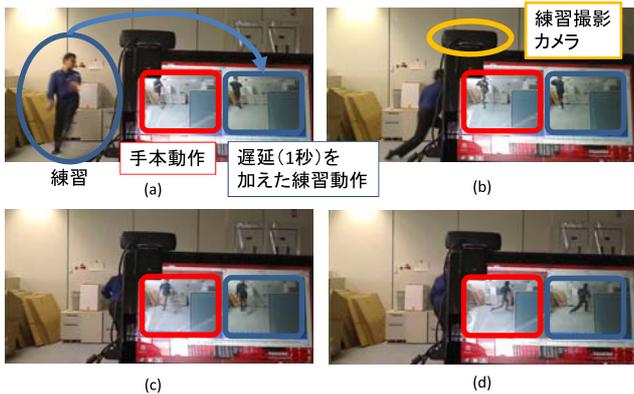


図 6 タブレット端末でのリアルタイム動作：(a) 練習動作（投球動作）開始直後。(b) 練習動作はリリース。遅延動作・手本動作は、左足を挙げた状態。(c), (d) 正しい同期が確認できる

Fig. 6 Real time processing of proposed method; (a) Beginning of pitching motion. (b) The timing of ball release. Delayed trainee's motion and the reference motion are at the timing of the left foot raised. (c), (d) Appropriate synchronization were verified.

4.1 典型的な動作

ここでは提案手法の典型的な動作例に基づいて提案手法の特性の説明を行う。

図 6 に提案システムの動作の様子を示す。図 6 では、画面左側で練習者が動作している。画面中右下側には、提案システムが動作しているタブレット端末が写っており、タブレット端末内左側に手本動作、右側に遅延（ここでは 1 秒の）を加えた練習動作が表示されている。図 6(a) は、練習動作（投球動作）開始直後である。図 6(b) で、練習者の動作はリリース状態である。これに対し遅延を加えた動作では、左足を挙げた状態である。そして手本動作も左足を挙げた状態であり、正しく同期されていることが分かる。図 6(c), (d) でも同様に、正しい同期がなされていることが確認できる。なおタブレット端末は、CPU インテル Core-i5, 1.8GHz, メモリ 8GB の市販のタブレット型 Windows 端末であり、処理には Web カメラによる映像のキャプチャ、映像特徴抽出 (=MHI 作成), 同期タイミング検出, 同期ビデオフィードバックに加えて、Web カメラにより撮影した映像の圧縮と保存が含まれる。このようにバッテリー駆動のタブレット端末によりリアルタイム動作することは、提案手法の大きな特長である。

図 7 に、体育館でのバスケットボールのレイアップシュート動作の同期フィードバック例を示す。この例では図 7(d) の時刻を同期タイミングとして指定した。手本動作と練習動作は、異なる服装の異なる人物により行われている。さらに、レイアップシュート動作では、比較的大きな位置変化をとともう動作である。このような動作に対して提案手法では、図 7 に示すとおり、正しく同期できていることが見て取れる。

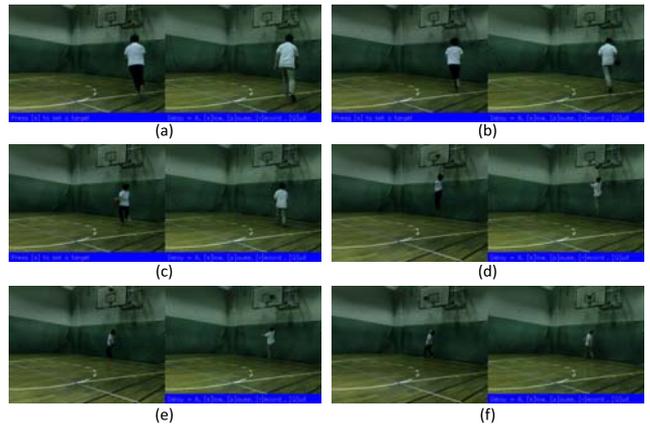


図 7 屋内におけるバスケットボールのレイアップシュートの例：手本動作と練習動作は、異なる人物・異なる服装により行われた。この例のように動作中の場所移動が大きい場合にも適応可能である

Fig. 7 Snapshots for basketball layups in gym; actors of reference and training are different and wear different clothes. As shown in figures, the proposed method can be applied to motions with large changes in position.



図 8 側転の同期例：手本動作は屋内、練習動作は屋外にて、異なる服装により行われた。このように照明環境および背景がまったく異なり、服装も異なる場合でも同期フィードバック可能である

Fig. 8 Snapshots for cartwheel; the actors wear different clothes. The reference motion was carried out in gym and the training was carried out in outdoor. As shown in figures, the proposed method synchronized well under different illumination conditions and different backgrounds.

図 8 に側転の同期例を示す。この例では (c) の時刻を同期タイミングとして指定した。手本動作は体育館で撮影され、練習は直射日光下の屋外で異なる服装で行われた。図 8 は、手本動作の方が動作の速度が遅く、練習動作の方が速度が速い。そのため、図 8(a) では、手本動作はすでに倒立に近い状態となっているが、練習動作はまだ手をついていない。図 8(b) の時刻になると、両動作間の状態はだいぶ近づき、図 8(c) ではほぼ同じタイミングとなっている。そして、図 8(d) では練習動作の方がより傾いた状態へ速く進んでいることが確認できる。このように服装に加えて、直射日光下を含めた照明環境、人物の大きさ、動

作の速度の変化に対して頑健であることは提案手法の重要な特長である。

図 9 に不完全動作との同期の例を示す。提案手法は動作の学習を目指したものであり、上達途上にある動作と手本動作との同期は重要である。図 9 で、練習者の側転は足が挙がらない(図 9(c) 参照)という練習段階の側転にしばしば見られる状態である。提案手法ではこのような不完全な動作に対しても適切な同期フィードバックが可能である。

図 10, 図 11 に投球動作中の異なるタイミングを同期



図 9 不完全動作との同期例：動作学習過程においては、上達途上にある不完全な動作との同期が必要となる可能性がある。提案手法では、側転で足が挙がらないような動作とも適切な同期が可能である。これにより、上達途上の練習者へのフィードバックが可能となる

Fig. 9 An example of synchronization with incompleted motion: Because the proposed method aims at motor learning, it is required to synchronize the reference motion with trainee's incompleted one. This makes the proposed method possible to give feedbacks to trainees of various skill levels.

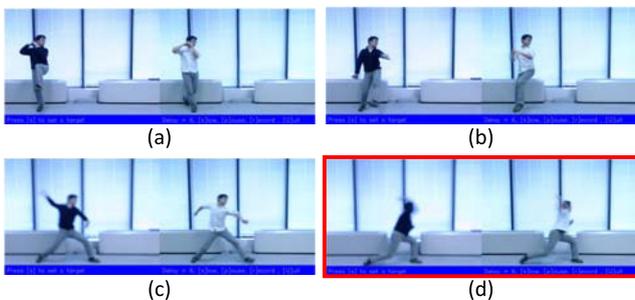


図 10 リリースタイミングでの同期

Fig. 10 Synchronization at ball release timing.

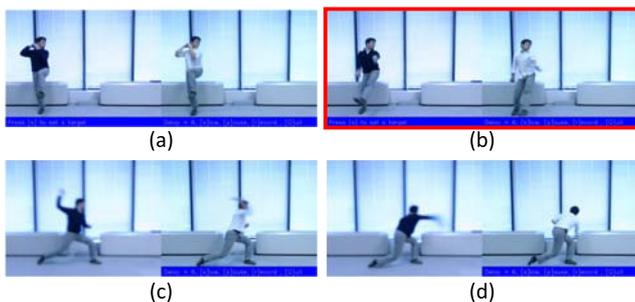


図 11 テイクバックでの同期

Fig. 11 Synchronization at taking back.

タイミングとして設定した例を示す。図 10, 図 11 において、赤枠で囲まれたフレームが同期タイミング (T_{sync}) であり、図 10 ではリリース, 図 11 ではテイクバックを同期タイミングとして設定した。なお、違いを分かりやすくするために、練習動作(白いシャツでの動作)時には、手本動作よりも高速に動いた。

図 10 ではリリース時(d)での動作がほぼ同期しているのに対し、それ以前のタイミングでは動作の進行度合が大きく異なっている。一方、図 11 では、テイクバック時(b)での動作がほぼ同期しているのに対し、その後のリリースに向けた状態(c), (d)では大きくずれている。提案手法では、このようにユーザが同期させたいタイミングを、同期させるタイミング (T_{sync}) として指定し、その時刻で動作の進行度合いが等しくなるように同期再生を行うことが可能である。

提案手法は同期ビデオフィードバックシステムであり、仮に過剰検出が発生しても手本動作が再生されるのみである。そのため過剰検出による実質的な問題は多くない。しかし提案システムの性能確認のため、図 3 の手本動作を用いた場合の非類似度(距離)推移と実際の投球タイミングを図 12 に示す。図 12 より、若干の過剰検出は見られるものの、高い手本動作検出性能が確認できる。

さらに、動作領域の大きさを変化させた場合の非類似度(距離)の時間変化を図 13 に示す。図 13 で、T1 から T4 は設定した動作領域であり、各々の動作領域に対応する非類似度(誤差)の時間変化が表示されている。T1 から T3 のように、動作領域の大きさを変化させた場合であっても大きな精度の変化はなく、さらには、T4 のように脚部分

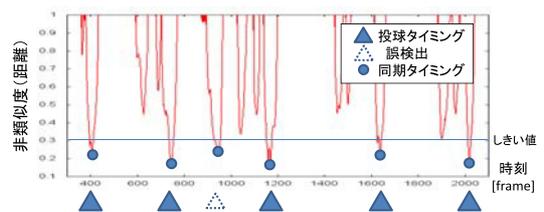


図 12 非類似度(距離)の時間推移

Fig. 12 Temporal transition of dissimilarity (distance).

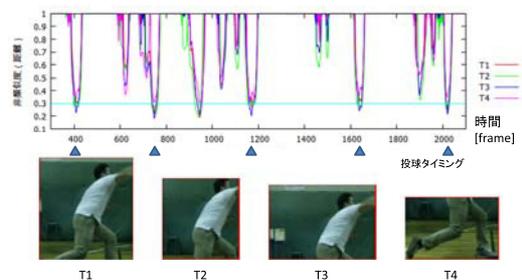


図 13 動作領域を変化させた際の非類似度(距離)の時間変化

Fig. 13 Temporal transitions of dissimilarity (distance) against changes in motion area settings.

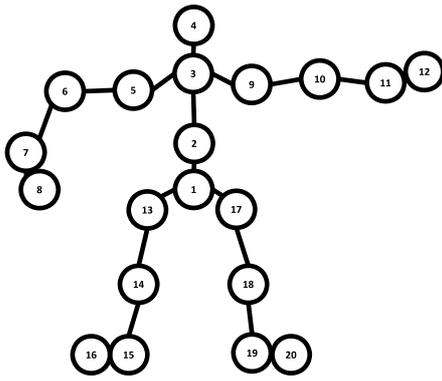


図 14 Kinect for Windows SDK により得られる姿勢情報
 Fig. 14 Pose information obtained from Kinect for Windows SDK.

のみを動作領域として指定した場合であっても極端な精度低下は見られなかった。

以上より、提案手法が服装の変化、照明環境の変化、フォームの変化に対して頑健に、そして様々な動作に対して同期ビデオフィードバック可能であることが示唆される。

4.2 定量評価

提案手法での様々な動作に対する同期性能を評価するために定量評価実験を行った。実験では、フリースローシュート (バスケットボール)、投球動作 (野球)、レイアップシュート (バスケットボール)、側転の 4 つの動作について同期性能の評価を行った。評価にあたっては、1 名の被験者の動作から手本動作を作成し、本人および他者 2 名の 3 名が各 10 回ずつ試行を行い正しく同期される割合を検証した^{*1}。

比較対象として、服装・人物の変化に対して頑健な姿勢推定が可能となりうる Microsoft Kinect (入力デバイス) および Kinect for Windows SDK (解析ソフトウェア) による姿勢抽出による同期を行った。Kinect による人物姿勢を図 14 に示す。Kinect では、頭部・肩・肘・手・膝など 20 点の 3 次元情報が取得可能である。ここでは、図 14 中、1 からの変位、つまり 19 点 × 3 次元 = 57 次元ベクトルにより姿勢を表現することとした。提案手法と同様、同期させるタイミングを実験者により指定し、そのときの姿勢との非類似度 (距離) により同期タイミング検出を行った。

なお、提案手法、従来手法とも、フリースローシュートおよび投球動作に関しては、ボールリリースの瞬間を同期タイミング T_{sync} として設定し、検出された同期タイミング (T_{detec}) と、練習動作でのボールリリースタイミングとの誤差が 2 フレーム以下の場合に正しく同期されてものとした。また、レイアップシュートでは、体が最も高く上がったタイミング、側転では足が垂直に上がったタイミングを同期タイミングとして設定し、同様の基準で精度評価

*1 他者は、2 名による試行の平均。

表 1 フリースローシュート同期性能評価：10 回試行中正しく同期できた回数

Table 1 Number of successful video synchronization for ten trials: Free throw shoot.

	提案手法	Kinect
手本動作を行った被験者	9	7
他者	9.5	2.0

表 2 投球動作での同期性能評価：10 回試行中正しく同期できた回数

Table 2 Number of successful video synchronization for ten trials: Pitching.

	提案手法	Kinect
手本動作を行った被験者	10	2
他者	9.0	0.5

を行った。

4.2.1 フリースローシュート

まずフリースローシュートでの同期再生性能を表 1 に示す。提案手法では、手本動作を行った人物でも他の被験者であっても高い精度で同期ビデオフィードバックがなされていることが分かる。一方で、Kinect の場合、手本動作を行った被験者本人の動作については比較的高い精度で同期フィードバックされたものの、他の被験者については精度がきわめて低い。

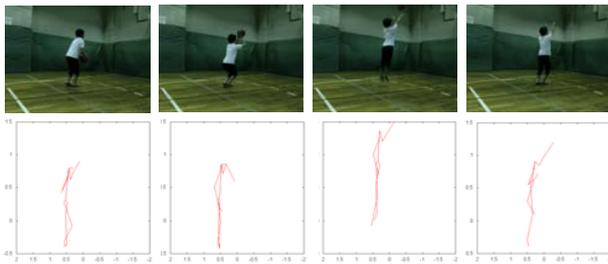
Kinect により推定された姿勢は 57 次元の高次元ベクトルでありフリースローシュート動作の特徴を抽出したものではない。そのため、姿勢推定誤差および個人間のフォームの差の影響を強く受けているものと考えられる。

4.2.2 投球動作

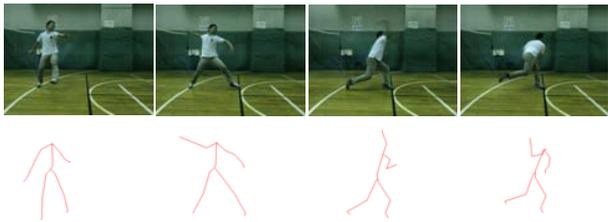
次に、投球動作の同期性能を表 2 に示す。投球動作の場合 Kinect での同期性能はさらに低下した。これは、フリースローシュートと比較して動作のスピードが速く、また動作中に遮蔽が多く発生するため姿勢推定精度が低下したことが原因と推測される。図 15 にフリースローシュートおよび投球動作中の姿勢推定結果を示す。フリースローシュートの場合、動作開始から終了までおおむね正しく姿勢推定が行われていることが分かる。これに対し投球動作では、特にボールリリース以降、動作が高速であること、および胴体による腕の遮蔽が生じることが原因で正しい姿勢推定が行えていないことが分かる。

4.2.3 レイアップシュートおよび側転

レイアップシュートおよび側転の同期フィードバック性能を表 3 に示す。なお、レイアップシュートおよび側転については提案手法の結果のみを示す。これは、比較対象では同期再生が不可能であったためであり、その理由については後述する。表 3 に示すとおり、高い精度で適切な同期ビデオフィードバックができていることが分かる。



(a) フリースローシュート時



(b) 投球動作時

図 15 フリースローシュートおよび投球動作のスナップショットと Kinect による姿勢情報

Fig. 15 Snapshots for free throw shot and Kinect-based postures.

表 3 レイアップおよび側転での同期性能評価：10 回試行中正しく同期できた回数

Table 3 Number of successful video synchronization for ten trials: Layup and Cartwheel.

	レイアップ	側転
手本動作を行った被験者	9	10
他者	8.0	10.0

レイアップシュート

Kinect で奥行情報取得および姿勢推定可能な奥行が約 4m までである。そのためレイアップシュートのようにある程度の移動をとまなう動作においては Kinect による同期ビデオフィードバックは困難であることが分かった。レイアップシュートをはじめ、跳び箱など、移動をとまなう動作は多い。提案手法では、このような移動をとまなう動作においても適切な同期ビデオフィードバックが可能である。

側転

側転動作は移動距離は大きくなく、また、移動速度もそれほど速くない。しかしながら側転中の姿勢の Kinect での推定は困難であった。図 16 に側転中の姿勢情報を示す。図 16(d) で左手が地面に着き倒立状態になるところである。その後、適切な姿勢推定がされないまま、図 16(e) から図 16(g) まで経過し、図 16(h) で正しい姿勢推定が再開される。Kinect SDK の奥行情報からの骨格情報の推定は、大量の奥行情報と骨格情報との組を用いた機械学習によって [12]。側転での頭と脚が上下反転する姿勢は学習さ

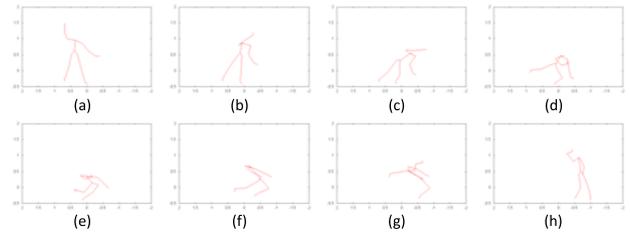


図 16 Kinect により得られる側転中の姿勢情報

Fig. 16 Postures during cartwheel obtained by Kinect.

れておらず、骨格情報推定に失敗しているものと推測される。このように Kinect でも推定対象姿勢の学習が必要であり、適用可能な動作は制限される。

これに対し提案手法では、移動をとまなうレイアップシュート、高速な動きを含む投球、そして倒立状態を含む側転と様々な動きに対して、変更・調整を加えることなく同期フィードバック可能である。この幅広い動作に対する適用可能性はきわめて重要である。

これに加えて、Kinect は直射日光に弱いことが指摘されている。Kinect では赤外線パターンを照射し、それをカメラにより撮影し、パターンの歪みから奥行情報を推定している。直射日光下では赤外線パターンの撮影が困難となり、奥行情報の精度が低下する。提案手法では、直射日光下での同期フィードバックが可能だけでなく、手本動作あるいは練習動作のどちらか一方が直射日光下で照明環境が大きく変わる環境であっても適切な同期フィードバックが可能である。これは、スポーツシーンでの利用を想定した場合、非常に重要である。

5. 動作学習支援に向けた今後の発展

本章では、競技者などへのヒアリングを通じて聞かれたビデオフィードバック方法について検討を行う。

第 1 にオーバーレイ表示があげられる (図 17, 図 18)。手本動作と練習者動作を横に並べた場合には動作者間の距離が離れてしまい、動作の違いの検証が困難な場合がある。そこでより近接した状態で確認できるオーバーレイ表示への要望が多く寄せられた。実際、図 8 と図 18 は同一シーケンスであるが、図 18 の方がより顕著に違いが伝わる。オーバーレイ表示は実装済みである。

第 2 に時間伸縮があげられる。提案手法は、同期タイミングを設定し、同期タイミングでの動作が同時に再生されるようにビデオフィードバックを行った。しかしながら手本動作と練習動作の動作速度が異なる場合に、時間伸縮を施して全体が同期された状態で動作を確認したいという意見が聞かれた。これは、同期タイミングを動作前半と動作後半の 2 点設定し、その間の動作を時間伸縮させることで実現可能である。今後の実装を検討していきたい。

第 3 に、手本動作の更新があげられる。特に高いレベルにある競技者へのヒアリングでは、直前の状態からの変化

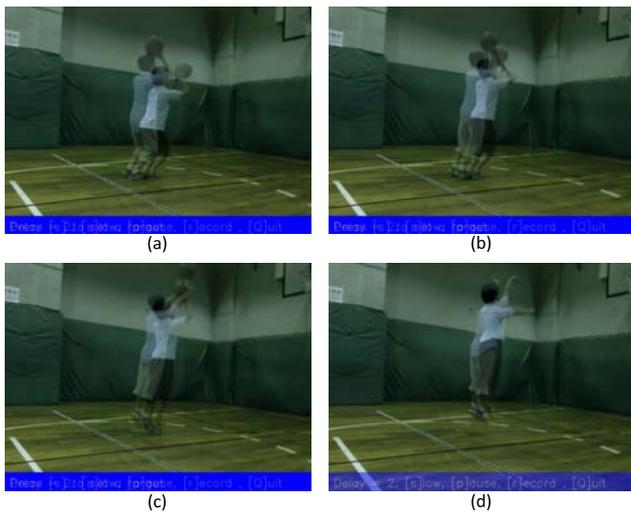


図 17 オーバレイ表示による同期ビデオフィードバック (フリースロー)

Fig. 17 Video feedback with overlay video representation: Free throw.

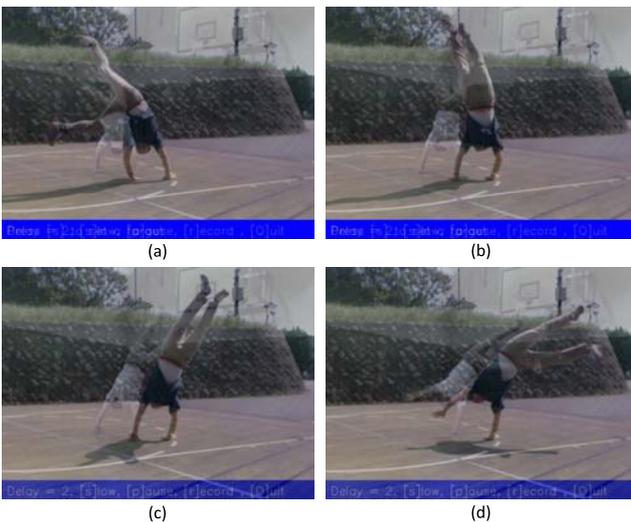


図 18 オーバレイ表示による同期ビデオフィードバック (側転)

Fig. 18 Video feedback with overlay video representation: Cartwheel.

(意識の変化と、意識の変化による動作の変化)を確認できることが重要との意見が聞かれた。現在の実装では、手本動作は起動時に1度設定するのみであるが、簡単な操作あるいは動作が検出されるごとに自動的に手本動作を更新する仕組みが望まれる競技者レベルも存在すると考えられる。これについても今後の実装を検討していきたい。

6. まとめ

本稿では、動作学習のための、同期遅延フィードバックシステムを提案した。提案手法では、様々な動作について、服装・照明環境・動作の違いに対して頑健な同期遅延フィードバックを、バッテリー駆動のタブレット型端末で可能とした。提案手法は、屋外、屋内を問わず様々な場所で

の、種々の反復動作を対象とした遅延同期ビデオフィードバックのために有効な可能性が高いことが分かった。ただし、より激しい動きをとまなう動作、逆に、ほとんど動きをとまなわれないような動作での利用、大きな照明変化をとまなう状況、練習者の背景で大きな動きが存在する状況での利用などを考えた場合、精度の良い遅延同期再生が困難なケースが生じうる。今後、提案手法の限界を明らかにすると同時に、MHIの拡張[13]の利用などの検討をしていきたい。

また、最適なビデオフィードバック方法は、練習者のレベル・特徴、動作の種類などに影響を受ける可能性がある。今後、より良いビデオフィードバック確立に向け、動作パフォーマンスへの効果の検証を行っていきたい。

参考文献

- [1] Guadagnoli, M., Holcomb, W. and Davis, M.: The efficacy of video feedback for learning the golf swing, *Journal of Sports Science*, Vol.20, Issue 8, pp.615-622 (2002).
- [2] Boyer, E., Miltenberger, R., Batshe, C. and Fogel, V.: Video modeling by experts with video feedback to enhance gymnastics skills, *Journal of Applied Behavior Analysis*, Vol.42, No.4, pp.855-860 (2009).
- [3] Van Wieringen, P.C., Emmen, H.H., Bootsma, R.J., Hoogesteger, M. and Whiting, H.T.: The effect of video-feedback on the learning of the tennis service by intermediate players, *Journal of Sports Science*, Vol.7, Issue 2, pp.156-162 (1989).
- [4] SiliconCoach: Timewarp, available from <http://www.siliconcoach.com/Products/Timewarp/>.
- [5] Dartfish: Dartfish, available from <http://www.dartfish.co.jp/soft/>.
- [6] 小澤治夫, 石田 譲, 岡崎勝弘, 西島尚彦: 鉄棒単元におけるスポーツミラーによる運動画像の即時フィードバックの効果, 北海道教育大学釧路校研究紀要, Vol.35, pp.1-6 (2003).
- [7] 村山光義, 村松 憲, 佐々木玲子, 清水静代, 野口和行: 動作映像の即時フィードバックを用いた技術指導の効果: フライングディスク・サイドアームスロー導入時の事例, 慶應義塾大学体育研究所紀要, Vol.46, No.1, pp.1-15 (2007).
- [8] Chua, P.T., Crivella, R., Daly, B., Hu, N., Schaaf, R., Ventura, D., Camil, T., Hodgins, J. and Paush, R.: Training for physical tasks in virtual environments: Tai Chi, *IEEE Virtual Reality*, pp.87-94 (2003).
- [9] Choi, W., Mukaida, S., Sekiguchi, H., and Hachimura, K.: Quantitative analysis of IAIDO proficiency by using motion data, *ICPR 2008*, pp.1-4 (2008).
- [10] 日本ナレッジ: MVP2000, available from <http://www2.know-net.co.jp/mvp2000/>.
- [11] Bobick, A. and Davis, J.: The Representation and Recognition of Action Using Temporal Templates, *IEEE Trans. PAMI*, Vol.23, No.3, pp.257-267 (2001).
- [12] Shotton, J., Fitzgibbon, A., Cook, M. and Blake, A.: Real-time Human Pose Recognition in Parts from Single Depth Images, *CVPR 2011* (2011).
- [13] Ahad, M.A.R., Tan, J.K., Kim, H. and Ishikawa, S.: Motion History Image: Its Variants and Applications, *Machine Vision and Applications*, Vol.23, Issue 2, pp.255-281 (2012).



三上 弾 (正会員)

2000年慶應義塾大学理工学部情報工学科卒業。2002年同大学大学院理工学研究科修士課程修了。2012年筑波大学大学院システム情報工学研究科博士後期課程修了。2002年日本電信電話(株)入社。現在、NTTメディアインテリジェンス研究所画像メディアプロジェクト研究主任。

コンピュータビジョン、映像ハンドリング、および映像インタラクションに関する研究に従事。MIRU2009優秀論文賞、平成22年度電子情報通信学会論文賞および喜安善市賞、ISBS2012 NIA (poster) 2nd placeほか受賞。IEEE、電子情報通信学会各会員。博士(工学)。



川村 春美 (正会員)

1989年東京女子大学文理学部数理学科卒業。同年日本電信電話(株)入社。主に、色知覚モデル、色知覚特性に基づくカラー画像処理に関する研究に従事。現在、NTTメディアインテリジェンス研究所主任研究員。1998年度電子情報通信学会学術奨励賞受賞。2012年日本バーチャルリアリティ学会サイバースペース研究賞受賞。電子情報通信学会、日本心理学会各会員。

2012年日本バーチャルリアリティ学会サイバースペース研究賞受賞。電子情報通信学会、日本心理学会各会員。



松本 鮎美

2007年九州芸術工科大学音響設計学科卒業。2009年九州大学大学院芸術工学府修士課程修了。2009年日本電信電話(株)入社。現在、NTTメディアインテリジェンス研究所画像メディアプロジェクト研究員。主にコンピュータビジョン、人物動作解析に関する研究に従事。2011年日本バーチャルリアリティ学会サイバースペース研究賞受賞。

2011年日本バーチャルリアリティ学会サイバースペース研究賞受賞。



小島 明

1988年東京大学工学部計数工学科卒業。1990年同大学大学院工学系研究科修士課程修了。同年日本電信電話(株)入社。以来、VOD、電子図書館、マルチメディア検索、映像監視、高臨場映像通信の研究・実用化に従事。現在、NTTメディアインテリジェンス研究所主幹研究員。電子情報通信学会、画像電子学会、ACM等各会員。

現在、NTTメディアインテリジェンス研究所主幹研究員。電子情報通信学会、画像電子学会、ACM等各会員。



門田 浩二

1996年京都教育大学教育学部体育学科卒業。1998年同大学大学院教育学研究科修士課程修了。2002年大阪大学大学院人間科学研究科博士後期課程単位修得後退学。2004年大阪体育大学大学院スポーツ科学研究科教務助手。2005年JST-ERATO下條潜在脳機能プロジェクト博士研究員。2011年大阪大学大学院医学系研究科助教。人間の感覚運動制御と学習の神経基盤、およびスポーツスキルに関する研究に従事。第26回NICOGGRAPH最優秀ポスター賞、デサントスポーツ科学振興財団研究助成優秀入選ほか受賞。北米神経科学学会、日本神経科学学会、日本スポーツ心理学会各会員。博士(人間科学)。

2005年JST-ERATO下條潜在脳機能プロジェクト博士研究員。2011年大阪大学大学院医学系研究科助教。人間の感覚運動制御と学習の神経基盤、およびスポーツスキルに関する研究に従事。第26回NICOGGRAPH最優秀ポスター賞、デサントスポーツ科学振興財団研究助成優秀入選ほか受賞。北米神経科学学会、日本神経科学学会、日本スポーツ心理学会各会員。博士(人間科学)。