Vol.2021-HCI-192 No.28
2021/3/16

スポーツ練習支援のための複数高速度カメラを用いた多視点自動撮影システム

向井 稜[1,a)]　武藤 駿嗣[1]　井尻 敬[1]

概要：本研究では，スポーツ練習の効率的な振り返りの支援を目的とし，複数の高速度カメラによって練習の様子を自動的に撮影するシステムを提案する．提案システムにおいて，ユーザは，撮影したい動作（バットスイングやテニスストロークなど）をキーモーションとして事前に登録し練習を行う．練習中，提案システムは，Web カメラによりユーザの姿勢を推定し，登録されたキーモーションと似た動きを検出したら複数の高速度カメラより撮影を行う．提案システムでは，多様な動作に対応できるよう，スタートシャッターとエンドシャッターという二通りの撮影法を提供する．提案システムの有用性を評価するため，バットスイングとテニスストロークを対象にユーザスタディを実施した．その結果，比較的高い精度でこれらの動作を検出できることを確認した．また，実験参加者より提案システムの練習利用に関して肯定的な反応が得られた．

1. はじめに

スポーツの技術向上には練習中の動作の振り返りが重要である．そのため，鏡や動画撮影により自身の姿勢を確認しながら練習が行われることも多い．また，スマートフォンにも搭載されるなど，高速度撮影機能を持つカメラが普及しており，これを利用することで高速な動作の詳細な振り返りも可能になりつつある．しかし，一般的に，汎用高速度カメラには撮影可能時間に制限があり，スポーツ練習に利用するには，動作のタイミングに合わせて録画開始・終了を指定する必要がある．そのため，練習者とは別に撮影者が必要となり，一人で練習と高速度撮影を行うことは難しい．

スポーツやダンスなどの練習支援のため，理想のフォームと実際のフォームとの差を可視化する試みがなされている．例えば，テニス [1]，ダンス [2], [3]，ゴルフ [4]，スラックライン [5] を対象に，練習者の姿勢と理想のフォームとの違いを，プロジェクタにより実装されたスマートミラーシステムなどで可視化する手法が提案されている．他にも，バットスイング [6] やヨガのポーズ [7] について，練習者の姿勢をセンサや画像から推定し，その評価をフィードバックする手法が提案されている．しかし，これらの手法は，特定スポーツ・特定動作に特化してデザインされており，多様なスポーツ動作の練習支援にそのまま利用するのは難しい．また，より一般的なスポーツ動作・武道動作の練習支援のため，武道の高速な動作 [8] や事前登録したス

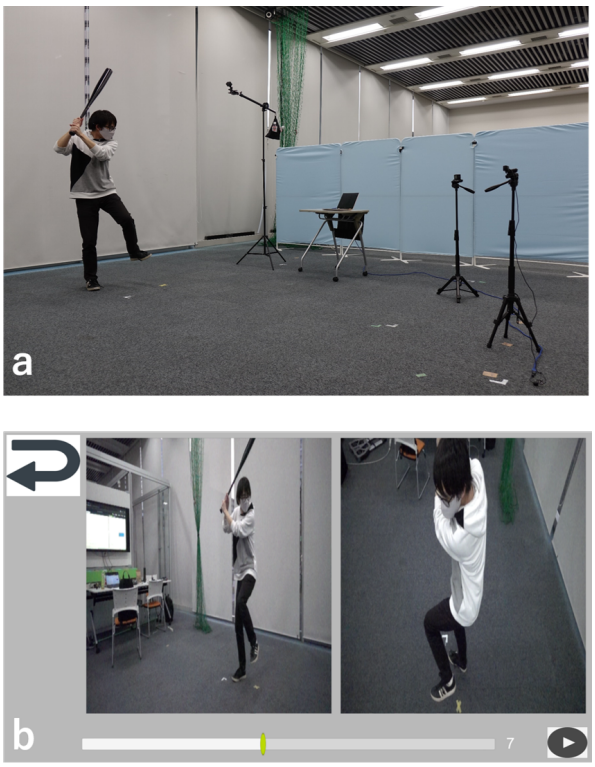

図 1　提案システムを用いた練習の様子．提案システムにおいて，ユーザは好きな位置に複数の高速度カメラを設置し，Web カメラの前で練習を行う (a)．練習中の特定動作は自動で高速度撮影され，後から複数視点のスロー動画として視聴できる (b)

1　芝浦工業大学

a)　ma20093@shibaura-it.ac.jp

ⓒ 2021 Information Processing Society of Japan

ポーツ動作 [9] を自動録画し，スマートミラーにより再生できるシステムが提案されている．特に，新野らの手法 [9] では，高速度カメラによる撮影がなされている．しかし，これらのシステムは，プロジェクタ等の装置を必要とするため，実際のスポーツ練習環境における利用は難しい．また，複数視点からの同時撮影などは議論されていない．

本研究では，スポーツ練習における高速な動作の振り返りの支援を目的とし，特定の動作を複数の高速度カメラにより自動撮影できるシステムを提案する（図 1）．提案システムは，高速度撮影機能付きの小型アクションカメラ（以下，高速度カメラ），タブレット PC，Web カメラにより構成される．ユーザは，Web カメラの前で撮影したい動作を 1 回行い，その動作を登録する．その後，高速度カメラを好きな位置に配置し，Web カメラの前で練習を行う．練習中，提案システムはユーザの動きを追跡し，登録動作と似た動作を検出したら，高速度カメラより撮影を行う．これにより，練習者は一人でも様々な位置から動作を撮影でき，自身の姿勢の詳細な確認が可能となる．また，本研究では，多様な動作を正しく録画するため，動作の始まりを検出するスタートシャッターと動作の終わりを検出するエンドシャッターの二通りの撮影法を用意する．

提案システムの有用性を確認するために，野球のバットスイングおよびテニスのフォアハンドストロークを撮影するユーザスタディを実施した．結果，バットスイングの検出精度は 95.2%，フォアハンドストロークの検出精度は 83.3%であった．また，提案システムによるバットスイング練習を体験した実験参加者に対し，提案システムの使いやすさに関するアンケート調査を行い，自動撮影機能・動画再生機能等について肯定的な反応を確認した．

2. 提案手法

2.1 提案システムの概要

本研究では，スポーツ練習における高速な動作の振り返りの支援を目的とし，練習の様子を複数の高速度カメラから自動撮影し見直せるシステムを提案する．図 2 に提案システムの構成を示す．提案システムは，タブレット PC（図 2b），Web カメラ（図 2c），高速度撮影機能付きアクションカメラ（図 2de）より構成される．Web カメラは，ユーザの動作検出用のためのもので，ユーザの全身が映る位置に三脚等を用いて設置される．また，高速度カメラはユーザの好みの位置に配置され，動作の高速度撮影に利用される．図 2a は，各要素を配置した例である．

提案システムにおいて，ユーザは，練習前に撮影したい動作をキーモーションとして登録する．練習開始後，提案システムは，Web カメラから得られる動画に PoseNet[10] を適用することでユーザの姿勢情報を取得し，キーモーションと似た動きを検出した際に高速度動画を撮影する．高速度動画の自動撮影において，本研究では，撮影開始時点の動作を検出しそこから 2 秒間を記録する『スタートシャッター』と，撮影終了時点の動作を検出しその前 2 秒間を巻き戻して記録する『エンドシャッター』の 2 種類を用意する．

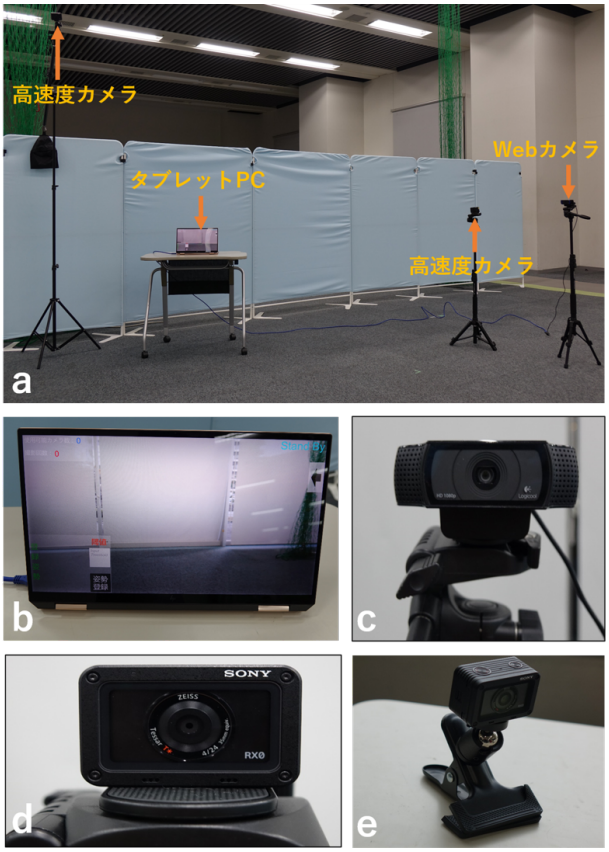

図 2 提案システムの全体像 (a)．提案システムは，タブレット PC(b)，Web カメラ (c)，複数の高速度カメラ (de) により構成される．

本研究では，高速度カメラとして，Wifi 通信により撮影開始・終了を指定可能な Sony RX0 を利用する．この高速度カメラは，クリップによる固定が可能なため（図 2e）様々な場所に手軽に配置可能である．この高速度カメラは，その仕様上，5 台まで同時に接続し，高速度撮影を行うことが可能である．高速度撮影した動画は，約 20 秒程度の時間をかけてカメラ内の SD カードへ保存される．提案システムでは，SD カード保存時に動画をタブレット PC へ転送し，転送後タブレット PC にて再生することも可能である．ただし，高速度カメラからタブレット PC の転送には通信量に限界があり，複数の設定で実験したところ，安定して動画転送を行えるのは 2 台分の動画像であった．

2.2 ユーザインタフェース

提案システムによる撮影方法を説明する．ユーザはすべての操作を図 3a に示すタブレット PC 画面にて実施する．練習は，Web カメラから全身が映る位置にて行われ，Web カメラより撮影された映像は常にタブレット PC に表示される．

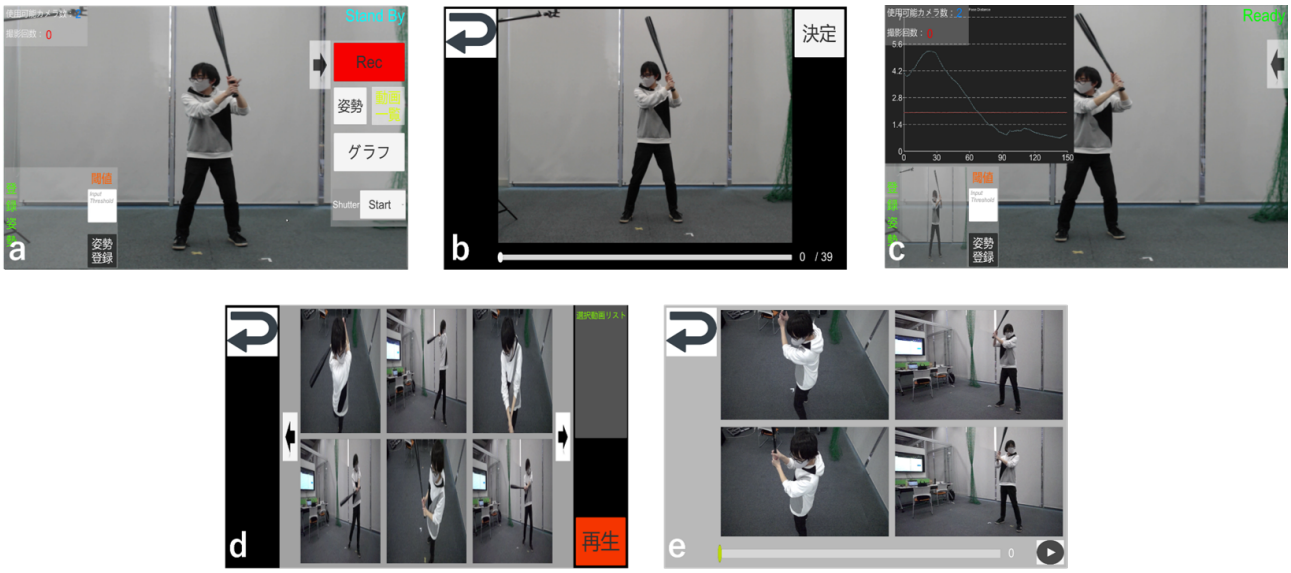

図 3 ユーザインターフェース．ユーザが練習を行う様子は Web カメラにより撮影され，タブレット PC 上に表示される (a)．ユーザは，練習時に撮影したい動作を行うことでキーモーションを登録する (b)．現在の動きとキーモーションとの差分はグラフ (c) により可視化される．撮影モード中にキーモーションと似た動きが検出されると，高速度撮影が行われ，得られた高速度動画はタブレット PC 上で再生できる (de)．

キーモーションの登録． ユーザは，練習を始める前に，撮影したい動作をキーモーションとして登録する．システム起動後，画面左下の『姿勢登録』ボタン（図 3a）を押すと，3 秒のカウントダウンが開始され，その後に 3 秒分（60 フレーム）の動画が Web カメラにより撮影される．ユーザはこの 3 秒間に，練習中に撮影したい動作と同じ動作を行う．撮影された動画はタブレット PC 上に提示され，ユーザは，スライダーバーを操作することで動画中の連続する 20 フレームを抽出し，キーモーションとして登録する（図 3b）．

高速度動画の自動撮影． キーモーションの登録後，『Rec』ボタンを押すと撮影モードが開始される．この状態でユーザがキーモーションと似た動きをすると，システムはそれを自動検出し高速度カメラに信号を出し，高速度撮影を行う．この撮影モードにおいて，図 3c の通り，現在の動きとキーモーションとの差分はグラフとしてリアルタイムに可視化される．ユーザはこのグラフを観察し，撮影開始のための閾値を調整可能である．多様なスポーツ・多様な動作を扱うには，撮影開始の閾値を動作ごとに調整する必要があるため，このような可視化が必要となる．

撮影対象動作の『直前』または『直後』のどちらに，よりシステムが検出しやすい特徴のある動作が現れるかは，対象動作に依存すると考えられる．そこで，本研究では，『スタートシャッター』と『エンドシャッター』の二つを用意する．スタートシャッターは，キーモーションに似た動きを検出した時点から 2 秒間高速度動画を撮影するものである．一方，エンドシャッターはシャッターを切った瞬間から 2 秒間遡り動画を記録する．ユーザは，キーモーション登録時にこの二つを選択することが可能である．

高速度動画の視聴． 練習の後，ユーザは撮影された高速度動画を視聴できる．タブレット PC 画面の右にある『動画一覧』ボタン（図 3a）を押すと，撮影された動画の一覧が表示される（図 3d）．ユーザは，動画を一つまたは複数選択し『再生』ボタンを押すことで選択した動画を視聴できる（図 3e）．この複数再生機能により，異なる角度から撮影された高速度動画を同時に視聴することや，同じ視点から撮影された複数の動画を比較することが可能となる．

3. キーモーション検出法

キーモーションと類似する動作を検出する手法について説明する．撮影モードにおいて，提案システムは，各時点から遡った直近 20 フレーム分の動作と，登録されたキーモーションとの距離を計算し，これが閾値以下であれば，類似動作と判断する．この動作間距離の計算のため，本研究では，Web カメラにより撮影される各フレームに PoseNet[10] を適用し，17 個のランドマーク点群を取得する（図 4）．

本研究では，現時点の動作とキーモーションとの距離 D を以下の通り計算する，

$$D = \sum_{f=1}^{T} \sum_{i=5}^{16} w_i ||\mathbf{p}_i^f - \mathbf{q}_i^f||. \tag{1}$$

ここで，$T = 20$ はフレーム数，$\mathbf{p}_i^f \in R^2$ はキーモーションにおける f 番目のフレーム・i 番目のランドマーク点の 2 次元座標，$\mathbf{q}_i^f \in R^2$ は直近 20 フレームにおける f 番目

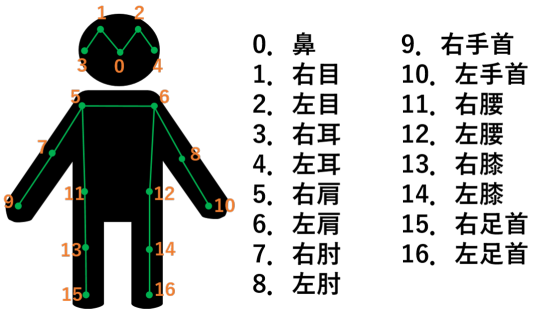

図 4 PoseNet より得られる 17 個のランドマーク点群.

のフレーム・i 番目のランドマーク点の 2 次元座標である．また，w_i は，i 番目のランドマーク点に対応する重みである．大きく動くランドマーク点ほど大きな重みづけがなされるよう，w_i を以下の通り計算する，

$$w_i = \frac{\sum_{f=1}^{T} ||\mathbf{p}_i^{f+1} - \mathbf{p}_i^{f}||}{\sum_{k=5}^{16} \sum_{f=1}^{T} ||\mathbf{p}_k^{f+1} - \mathbf{p}_k^{f}||}. \tag{2}$$

また，事前に行ったバットスイングの実験において，動きの少ない顔のランドマーク点（$i = 0, 1, 2, 3, 4$）は除いた方が検出精度が向上したため，本研究における距離計算では顔のランドマーク点を利用しないこととした．

PoseNet によって得られるランドマーク点は，練習する場所かカメラとの距離に応じて全体の位置や大きさが変化する．そのため，本研究では得られた 17 個のランドマーク点に対しバウンディングボックス計算し，x 座標，y 座標それぞれを [0,1] に正規化した物を，$\mathbf{p}_i^f, \mathbf{q}_i^f$ として利用する．

4. バットスイングに関するユーザスタディ

提案システムの検出精度とユーザビリティを確認するため，ユーザスタディを実施した．ここでは，野球のスイング練習を想定し，複数の野球経験者に提案システムを利用した練習を体験してもらい，提案システムの使いやすさと検出精度を調査した．

4.1 タスク内容

各実験参加者の行うタスクの内容は以下の通りである．まず，実験参加者は本研究の目的・内容，および，提案システムの使い方について説明をうける．続いて，提案システムに慣れるため約 15 分の撮影練習を行う．この撮影練習では，スタートシャッターとエンドシャッターそれぞれによるキーモーション登録，閾値調整，自動撮影を自由に体験する．その後，各参加者は，スタートシャッターを用いてバットスイング 5 回分を撮影し，続いてエンドシャッターを用いてバットスイング 5 回分を撮影する．このとき，対象動作以外を誤まって撮影してしまったものや検出ミスはカウントせず，スイング動作を 5 回撮影できるまで練習を行うこととする．撮影の際に，キーモーション検出がうまく動作しないと感じた場合，キーモーションや閾値の再設定を自由に行える．また，撮影に利用する高速度カメラは，練習者正面と撮影者を上から見下ろす位置に配置する．

上記の撮影の後，各参加者は，撮影された自身のスイング動作を観察し，最後に，提案システムの使いやすさや練習効果に関するアンケートに回答する．アンケートは，図 5 に示す 5 項目に対し，それぞれ 4 段階のリッカート尺度で回答するものである．また，各項目について，フリーディスカッション形式でコメントをもらう．

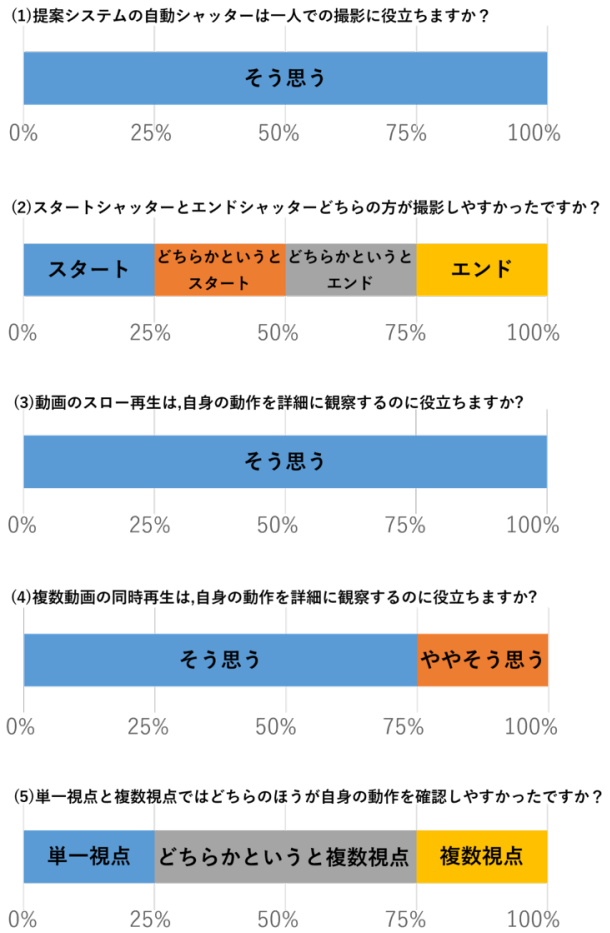

図 5 アンケート項目とその結果.

4.2 ユーザスタディ結果と考察

いずれも野球経験のある 4 名の実験参加者に，前述の通り提案システムを利用した撮影を体験してもった．アンケートへの回答結果を図 5 に示す．また，各項目への代表的なコメントを以下に示す．

(1) 提案システムの自動シャッターは一人での撮影に役立ちますか？この項目について，すべての参加者がそう思うと回答した．また，『事前に登録した動作を検出して自動で撮影してくれるのは便利』『いちいちシャッターを押す必要がなく連続で撮影できるため練習の邪魔にならない』という意見が得られた．この項目への回答から，提案シス

テムの自動撮影機能は一人でのスポーツ動作の撮影に役立つと参加者が肯定的にとらえたことが分かる．

(2) スタートシャッターとエンドシャッターどちらの方が撮影しやすかったですか？この項目への回答は，参加者ごとに異なる結果となった．肯定的なコメントとして，『スタートシャッターは動作の登録も簡単で動作の未検出が少ないため使いやすい』『エンドシャッターは構えてから自分のタイミングでスイングできるため，実践に近い練習ができる』という意見が得られた．一方，『スタートシャッターでは撮影したい動作の初めの動作を登録するため，構えてから自分のタイミングでスイングできないのが使いにくい』『エンドシャッターでは，キーモーションと似た動きをしても差が小さくならず，動作登録を何回もする必要があった』という否定的な意見もあった．この項目への回答から，対象とする練習動作や練習スタイルにより，スタート・エンドシャッターを切り替える必要があると考察できる．また，ユーザが明示的に選ばずとも，システムが自動で検出精度の高いシャッター方式を選択する機能の提供なども重要な将来課題である．

(3) 動画のスロー再生は，自身の動作を詳細に観察するのに役立ちますか? この項目については，参加者全員がそう思うと回答した．また，『自分の視線をじっくりと観察できるため，打つ瞬間までボールを見ているかを確認できるのが良い』『スローでバットの軌道，ミートポイント（腕の返る位置），体の使い方（腰の回し方など）を見れるため，自身の動作をより詳細に観察しやすい』との意見が得られた．この項目への回答から，スポーツ練習において，高速度撮影した動画を見返すことは有効であると多くの参加者が肯定的にとらえたことが分かる．

(4) 複数動画の同時再生は，自身の動作を詳細に観察するのに役立ちますか? この項目について，4 名中 3 名の参加者がそう思うと回答した．また，『同じ視点からの動画を並べて比較できるのがよい』『2 視点の動画を 2 つ同時に見ることでより詳細な比較を行える』『野球の試合で同じ視点から，全ての打席を撮影することで各打席の分析にも役立つ』という意見が得られた．一方，『各動画の開始動作のタイミングがずれているため比較しにくい』という否定的なコメントもあった．この項目への回答から，複数動画の再生機能は肯定的に受け入れられたことが分かる．より動作比較を行いやすくするために，複数動画内の動作タイミングを揃えて再生する機能の提供が今後の課題である．

(5) 単一視点と複数視点ではどちらのほうが自身の動作を確認しやすかったですか？この項目への回答は参加者ごとに異なる結果となった．また，『複数視点の動画は，単一視点からは見られない体の部位を確認できるのがよかった』『今まで複数視点からの動画を同時に視聴したことがないため，複数の単一視点の動画を並べてみたほうが動作を確認しやすい』『カメラの置く位置を自分で決めたい，視点数がもっとあるとよい』等の意見が得られた．この項目への回答から，ユーザに依存するものの，複数視点からの高速度動画撮影は，動作の詳細な確認の助けになる可能性が示唆された．今後，ユーザがカメラ位置を決められるより自由度の高いユーザスタディやスポーツ指導者を交えたユーザスタディを行うことで，複数視点撮影の有効性をより深く検証したい．

4.3 キーモーションの検出精度

このユーザスタディでは，4 名の実験参加者にスタートシャッターとエンドシャッターを用いて，それぞれ 5 回ずつのスイング動作を撮影したもらった．このタスク中に，全参加者が実際に行ったスイング回数，検出できなかったスイング数（未検出数），スイング以外をスイングとして検出した数（誤検出数）を表 1 に示す．スタートシャッターを用いた場合，実験参加者がスイングを 20 回撮影するまでに起こった検出ミスは，1 回のみであった（検出率 95.2%）．また，エンドシャッターを用いた場合，実験参加者のスイングを 20 回撮影するまでに起きた検出ミスは，8 回であった（検出率 71.4%）．また，誤検出は一度も発生しなかった．

この通り，検出率が高く誤検出が起きなかったことから，提案手法はバットスイング練習に十分利用できると考えられる．また，スタートシャッターのほうがエンドシャッターより検出精度が高かった理由は，スイングにおける振り終わりの姿勢は，振り方や意識するコースにより異なるためだと考えられる．

表 1 キーモーションの検出精度評価結果.

	シャッター	スイング回数	未検出数	誤検出数
実験参加者 1	スタート	5 回	0 回	0 回
	エンド	7 回	2 回	0 回
実験参加者 2	スタート	5 回	0 回	0 回
	エンド	6 回	1 回	0 回
実験参加者 3	スタート	6 回	1 回	0 回
	エンド	9 回	4 回	0 回
実験参加者 4	スタート	5 回	0 回	0 回
	エンド	6 回	1 回	0 回

5. テニスストロークへの適用例

バットスイング以外の動作にも提案システムが適用可能であることを確認するため，軟式テニスストロークの撮影を実施した．この例では，練習者正面に Web カメラとタブレット PC を設置し，ボールを手出しして，フォアハンドにてストロークする様子を撮影した．動作検出には，エンドシャッターを利用した．また，高速度カメラ 4 台を練習者の，正面・上・左前方・後ろに配置し，4 台同時に撮影を行った．図 6 に練習シーンを示す．

図 6 提案システムを用いたテニスのフォアハンドストローク練習.

図 7 提案システムにより異なる 4 個の異なる高速度カメラから自動撮影されたストロークを同時再生した例.

図 7 に，撮影した動画を示す．提案システムを利用すると，練習者は撮影者の協力なしに，同じストローク動作を多視点から撮影した動画を取得できる．また，この例では，24 回のストロークを行い，そのうち 20 回を正しく検出できた（未検出 4 回，検出率 83.3%）.

6. まとめと展望

本研究では，スポーツ練習の効率的な振り返りの支援を目的とし，練習中の特定動作を自動的に複数の高速度カメラから撮影し見直すことができるシステムを提案した．提案手法の有用性を確認するために，バットスイングおよびテニスストロークに着目したユーザスタディ実施した．結果，バットスイングは 95.2%，テニスストロークは 83.3% という精度で正しく検出できることを確認した．この結果は，我々の目的とする練習応用に対して十分なものだと考えている．また，ユーザスタディにおいて参加者へアンケートを実施した結果，提案システムの提供する自動録画機能・動画再生機能等について肯定的な意見が得られた．

現在の実装では，複数の高速度カメラより同時に撮影を行った場合に，ネットワーク遅延等の影響により，撮影開始のタイミングがずれるという問題がある．このずれは，複数視点動画を比較する際に動作のずれを引き起こす．そのため，複数動画内の動作タイミングをそろえて再生する機能を実現する必要がある．また，本研究ではバットスイング・テニスストローク練習を扱ったが，より多様なスポーツへの適用や，試合中の動きへの適用などが重要な将来課題である．

謝辞 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る．本研究は，日本学術振興会科学研究費 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] Oshita, M., Inao, T., Mukai, T. and Kuriyama, S.: Self-Training System for Tennis Shots with Motion Feature Assessment and Visualization, *2018 International Conference on Cyberworlds (CW)*, IEEE, pp. 82–89 (2018).

[2] Anderson, F., Grossman, T., Matejka, J. and Fitzmaurice, G.: YouMove: enhancing movement training with an augmented reality mirror, *Proceedings of the 26th annual ACM symposium on User interface software and technology*, pp. 311–320 (2013).

[3] Chan, J. C., Leung, H., Tang, J. K. and Komura, T.: A Virtual Reality Dance Training System Using Motion Capture Technology, *IEEE transactions on learning technologies*, Vol. 4, No. 2, pp. 187–195 (2010).

[4] Ikeda, A., Hwang, D.-H. and Koike, H.: Real-time Visual Feedback for Golf Training Using Virtual Shadow, *Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces*, pp. 445–448 (2018).

[5] Kosmalla, F., Murlowski, C., Daiber, F. and Krüger, A.: Slackliner - An Interactive Slackline Training Assistant, *Proceedings of the 26th ACM international conference on Multimedia*, pp. 154–162 (2018).

[6] Ghasemzadeh, H. and Jafari, R.: Body sensor networks for baseball swing training: Coordination analysis of human movements using motion transcripts, *2010 8th IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*, IEEE, pp. 792–795 (2010).

[7] Chen, H.-T., He, Y.-Z. and Hsu, C.-C.: Computer-assisted yoga training system, *Multimedia Tools and Applications*, Vol. 77, No. 18, pp. 23969–23991 (2018).

[8] Hämäläinen, P.: Interactive Video Mirrors for Sports training, *Proceedings of the third Nordic conference on Human-computer interaction*, pp. 199–202 (2004).

[9] 新野大輔， 井尻敬：姿勢検出法を活用したスポーツ反復練習のためのスマートミラーシステム，エンタテインメントコンピューティングシンポジウム 2019 論文集，Vol. 2019, pp. 138–144 (2019).

[10] Dan, O. and Irene, A.: Real-time Human Pose Estimation in the Browser with TensorFlow.js, Google Creative Lab (online), available from ⟨https://medium.com/tensorflow/real-time-human-pose-estimation-in-the-browser-with-tensorflow-js-7dd0bc881cd5⟩ (accessed 2019-07-22).