

動作学習のための映像情報を用いた オンサイト視覚フィードバック方式

永井隆昌^{†1} 宍戸英彦^{†2} 亀田能成^{†2} 北原格^{†2}

概要: 本稿では、多視点映像を取得してからバレットタイム映像を生成・提示するまでの全ての処理をリアルタイムで実行するオンサイト視覚フィードバック方式について述べる。スポーツシーンのような動的な状況下においてオンサイト視覚フィードバックを実現するためには、バレットタイム映像を観察する際の注視点を自動的に設定する必要がある。本研究では、RGB-Dカメラによる被写体の3次元位置検出とリアルタイムバレットタイム映像生成法を組み合わせて、測定した3次元位置に基づいた注視点の自動設定を実現する。さらに、対象とする動作を学習するために重要なフレームを検出し、そのフレームを観察するのに最も適した視点に多視点映像を自動的に切替える機能を組み込むことにより、閲覧操作の単純化を試みる。

キーワード: 多視点映像, 視覚フィードバック, バレットタイム映像, 動作学習, リアルタイム処理

On-Site Visual Feedback Method Using Video Information for Motor Learning

TAKASUKE NAGAI^{†1} HIDEHIKO SHISHIDO^{†2}
YOSHINARI KAMEDA^{†2} ITARU KITAHARA^{†2}

Abstract: This paper proposes an on-site visual feedback method for using multiple viewpoint videos in motor learning. There are two requirements to be satisfied in order to use multiple viewpoint videos in motor learning. The first requirement is to execute all processes from acquisition of multiple viewpoint videos to presentation in real-time, and to realize both real-time visual feedback and delayed visual feedback. In this research, real-time processing is realized by controlling all cameras with one computer and losing offline work. The second requirement is to minimize the operation of the viewer so that the viewer can concentrate on the observation of the motion. This research automates two operations, "target point setting" and "viewpoint switching", which are operations by viewer. We introduce the RGB-D camera that detects the position of the subject in real-time. By automatically setting the target point on the subject based on the 3D position of the subject measured by the RGB-D video, the "target point setting" operation is automated. Furthermore, the posture of the subject is detected using the RGB-D video and machine learning. By detecting the moment when the viewer wants to observe and automatically switching the multiple viewpoint videos in the direction in which the viewer wants to observe the posture, the "viewpoint switching" operation is automated. We verify that the proposed method works correctly by demonstration experiment.

Keywords: Multiple viewpoint videos, Visual feedback, Bullet-time video, Motor learning, Real-time processing

1. はじめに

視覚フィードバックの導入により、運動トレーニングや技能習得の効率化を目的とした研究開発に注目が集まっている。視覚フィードバックにより、被写体の身体や動作の客観的な観察や状況把握が可能となり、パフォーマンス向上に効果があることが報告されている¹⁾⁹⁾。本稿では、パフォーマンス中に行われるリアルタイム視覚フィードバックと、パフォーマンス直後に行われる遅延視覚フィードバックの二つの視覚フィードバックに着目し、それらを総称してオンサイト視覚フィードバックと呼ぶ。

従来の視覚フィードバック方式では、鏡や少数台のカメラで撮影した映像をモニタに提示しているため、多様な角度で被写体を観察することは困難である。3次元形状復元

処理やモーションキャプチャ装置を用いて被写体のCG (Computer Graphics) モデルを生成し、それをHMD (Head Mounted Display)などで観測⁷⁾¹¹⁾することにより、多様な角度からの観察が可能となるが、計算機で合成された人工的な環境や自己像に起因する違和感が観察の妨げになることが懸念される。

我々は、Figure 1のように、多数のカメラで同一シーンを様々な角度から撮影した多視点映像を用いることにより、上述した問題を解決する視覚フィードバック方式の実現を目的とした研究に取り組んでいる。多視点映像を効果的に提示する技術の一つとしてバレットタイム映像¹²⁾¹³⁾が知られている。バレットタイム映像は、撮影時のカメラ配置に応じて多視点映像を順番に切替えながら提示することにより、被写体の周りを飛び回るような視覚効果を実現する映

^{†1} 筑波大学 システム情報工学研究科
University of Tsukuba. Graduate school of system and information engineering.

^{†2} 筑波大学 計算科学研究センター

University of Tsukuba. Center for computational sciences.

像技術である。自由視点映像と異なり、撮影シーンの3次元形状を推定する必要がないため、高品質な映像を高速に生成・提示することが可能であり、我々が実現を目指すオンサイト視覚フィードバック方式との親和性が高い。

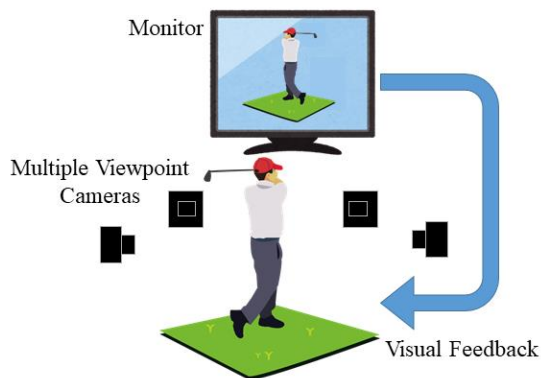


Figure 1: Visual feedback of multiple viewpoint videos: The subject can practice while viewing videos from various directions reflected on the monitor.

バレットタイム映像をオンサイト視覚フィードバックに導入するためには、以下二つの要件を満たす必要があると考えられる。

一つ目は、“バレットタイム映像のリアルタイム生成”である。リアルタイム視覚フィードバック、遅延視覚フィードバックの両方が動作学習に効果的であることが報告されているが、遅延視覚フィードバックは閲覧者が見たいと思った瞬間に見たい映像を提示可能なシステムが望ましい。しかし、多数のカメラで撮影された膨大な多視点映像データの計算機への取り込み処理、および、バレットタイム映像生成・提示に要する計算コストの問題により、多視点映像をオンサイト視覚フィードバックに用いることは困難である。本研究では、全ての多視点映像を1台の計算機に集約することで映像伝送時間などのバレットタイム映像の生成に要する時間を短くし、リアルタイムかつ遅延視覚フィードバックが可能なシステムを実現する。

二つ目の要件は、“閲覧操作の簡易化”である。被写体の観察に閲覧者が集中するために、閲覧者の操作を極力簡易化する必要がある。そのために我々は、(a)注視点再設定の自動化、(b)多視点映像の切替えの自動化、の二つの処理をシステムに導入した。

(a)注視点再設定の自動化：バレットタイム映像では、滑らかな多視点映像の切替えを実現するために、全てのカメラ（以下、多視点カメラ）の光軸が撮影空間中の一点（注視点）で交わるようにカメラの姿勢を調整する必要がある。本稿ではこの処理を注視点再設定処理と呼ぶ。富山ら¹⁴⁾や明智ら¹⁵⁾は、閲覧者が自由に注視点の再設定が可能なバレットタイム映像生成手法を提案している。しかしそれらの手法では、被写体が移動する度に注視点情報を入力する必

要があるため、スポーツのように被写体が絶えず移動する動的なシーンへの適用は現実的とは言い難い。本研究では、リアルタイム3次元センサとリアルタイムバレットタイム映像生成方式を組み合わせることにより、注視点再設定の自動化を実現する。

(b)多視点映像の切替えの自動化：動作学習において、ある瞬間（フレーム）の被写体の姿勢が、その動作の中でとくに重要である場合、その瞬間をキーフレームと呼ぶことにする。多視点映像を用いることにより、キーフレームを様々な角度から観察可能となる。しかし、視点数の増加に伴い、各キーフレームにおいて、観察に適した方向から撮影された多視点映像への切替え作業に必要な時間が増加する。本研究では、機械学習によりキーフレームを学習し、動作中のキーフレームを検出することで、各キーフレームを観察するのに適した方向から撮影された多視点映像への切替え処理の自動化を実現する。

2. 多視点映像によるオンサイト視覚フィードバック方式

リアルタイムバレットタイム映像提示方式に、注視点再設定の自動化機能と多視点映像の切替えの自動化機能を付加したオンサイト視覚フィードバック方式を Figure 2 に示す。被写体の周囲に多視点カメラと RGB-D カメラを設置する。多視点カメラはバレットタイム映像を生成するための多視点映像を撮影する。RGB-D カメラで撮影した RGB-D 映像に人物姿勢検出処理を適用し、被写体の3次元位置やボーン情報を推定する。多視点映像と RGB-D 映像に SfM (Structure from Motion)¹⁶⁾を適用し、カメラパラメータを推定する。それらのカメラパラメータと RGB-D 映像から推定した被写体の3次元位置を用いて、注視点再設定の自動化を実現する。RGB-D 映像により被写体のジェスチャー検出を行い、キーフレームを観察したいカメラ位置への多視点映像の切替えの自動化を実現する。バレットタイム映像生成、注視点再設定の自動化、多視点映像の切替えの自動化の三つの処理をリアルタイムで実行することにより、バレットタイム映像を用いたオンサイト視覚フィードバックを実現する。

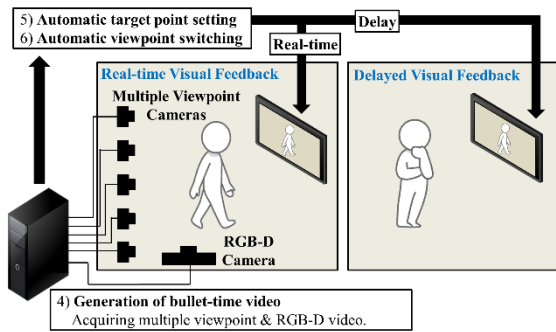


Figure 2: System overview: Multiple viewpoint videos and RGB-D video are acquired with one computer. Estimate the 3D position of the subject from the RGB-D video and realize automatic target point setting. We perform gesture detection of the subject using RGB-D video and realize automatic viewpoint switching.

3. 視覚フィードバックのためのパレットタイム映像の生成

3.1 多視点映像の撮影

多視点映像を計算機にリアルタイムで取り込む処理について述べる。Figure 3 に示すように、被写体を取り囲むように多視点カメラを設置する。カメラの設置位置はパレットタイム映像の切替りの滑らかさに影響するため、可能な限り同一平面上に配置する。一方、カメラの姿勢は後述する注視点再設定処理によって調整可能であるため、おおよそ同一領域を撮影するよう設定する。Figure 4 に示すように、多視点カメラを外部同期ケーブルで同期信号発生器に結線することで、多視点カメラを同期する。また、多視点カメラを映像ケーブルでパレットタイム映像生成用計算機に結線し、同期多視点映像を計算機に取り込む。撮影した映像を計算機に直接取り込むことにより、記録メディアを介したオフライン映像データのハンドリングの作業をなくしリアルタイム撮影を実現する。

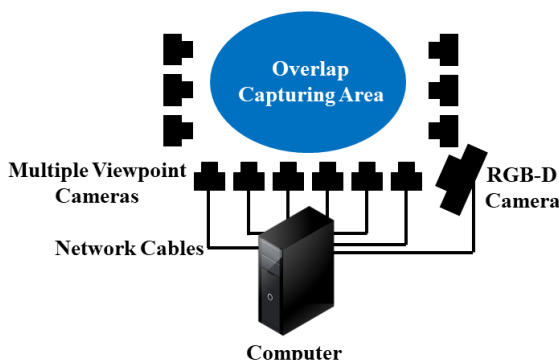


Figure 3: Bullet-time video generation method: Multiple viewpoint cameras and RGB-D camera are arranged around the subject such that the overlap capturing area becomes as large as possible.

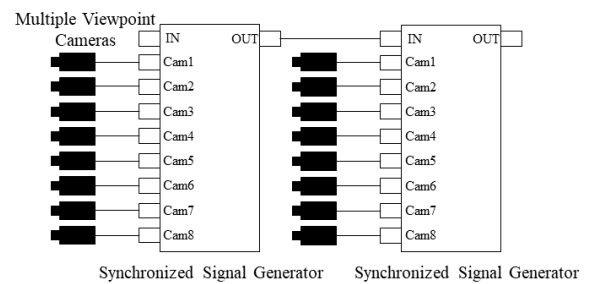


Figure 4: Synchronization of multiple viewpoint cameras: Multiple viewpoint cameras can be synchronized by connecting cameras to synchronized signal generator. It is also possible to synchronize the multiple viewpoint cameras between different generators by connecting the generators with the extension cable.

3.2 多視点カメラのキャリブレーション

多視点カメラの内部・外部カメラパラメータを推定するために、SfM を用いてカメラキャリブレーションを行う。SfM では、まず画像特徴量¹⁷⁾を用いて画像から特徴点を検出し、多視点画像間での対応点を求める。その対応点情報からカメラのモーション（相対的な位置姿勢情報）を推定し、同時に、ステレオ視により対応点の3次元位置を推定する。これらの処理を全ての画像対について行った後、バンドル調整¹⁸⁾を適用することで対応点の3次元位置とカメラパラメータを推定する。

N 個の多視点カメラについてキャリブレーションを行う際、カメラ $C_m (m = 1, \dots, N)$ の内部パラメータ行列 A_{cm} は式 (1) のように表される。

$$A_{cm} = \begin{bmatrix} f_m & 0 & c_{x,m} \\ 0 & f_m & c_{y,m} \\ 0 & 0 & 1 \end{bmatrix} \quad (1)$$

ここで、SfM によって推定したカメラ C_m の焦点距離を f_m 、カメラの光軸と画像平面の交点の画像座標を $(c_{x,m}, c_{y,m})$ とする。Figure 5 に示すように、世界座標系に対するカメラ C_m の位置と姿勢をあらわす回転行列 R_{cm} と平行移動ベクトル t_{cm} も SfM によって推定することができる。以上の処理は、多視点カメラ設置後に一度だけ行い、実行時には推定したパラメータを利用する。

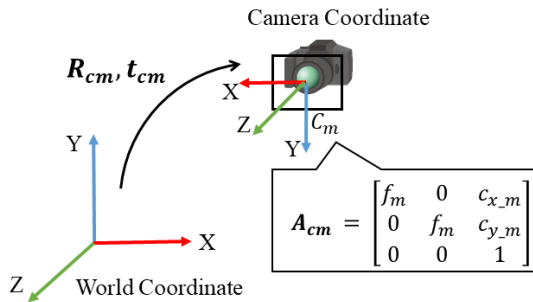


Figure 5: Camera parameters estimation using SfM.

3.3 バレットタイム映像の生成

バレットタイム映像を生成する過程を Figure 6 に示す。システムを起動すると、計算機は多視点カメラを起動し、多視点映像の取り込みを開始する。同期信号発生器は一定時間毎に映像更新の信号をカメラに送り、計算機は全てのカメラからフレームを取得する。カメラ配置に従って多視点映像を切替えることによりバレットタイム映像を生成する。その際、多視点映像の切替えは、隣り合うカメラ間で行われる。ある時間において表示される映像は、閲覧者が選択したカメラの映像のみである（選択されたカメラが撮影した映像を画面に提示する）。閲覧者が観察画面（観察視点）を変更すると、同様に、視点位置に対応するカメラを選択し、そのカメラが撮影した映像を画面に出力する。一連の操作を繰り返すことによりバレットタイム映像が生成される。

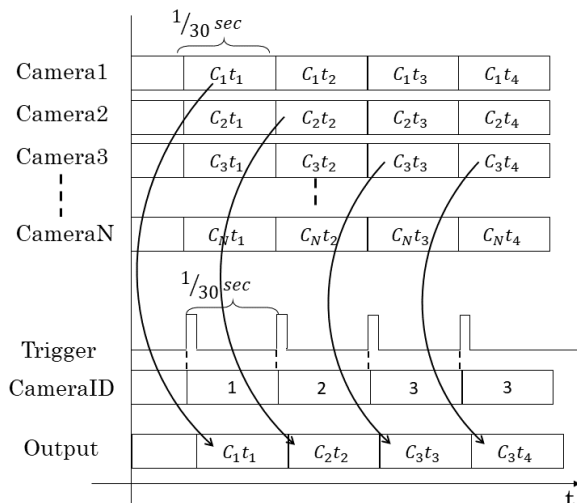


Figure 6: Generation of bullet-time video: The synchronized signal generator sends a trigger to all cameras at regular time intervals, and the computer acquires frames from all cameras ($C_m(m = 1, \dots, N)$). The viewer selects the camera (CameraID), and system outputs the video captured by selected camera to the screen.

3.4 バレットタイム映像によるオンサイト視覚フィードバック

バレットタイム映像をオンサイト視覚フィードバックする処理について述べる。本研究では、オンサイト視覚フィードバックとして、Figure 7 のように被写体のパフォーマンスと同時に映像を提示するリアルタイム視覚フィードバックと、Figure 8 のようにパフォーマンス直後に被写体を撮影した映像を提示する遅延視覚フィードバックの両方を実現し、練習の目的に応じて切替えて使用可能とする。リアルタイム視覚フィードバックと遅延視覚フィードバックは両方ともに動作学習の習熟度に対して効果的であると報告されている。練習現場で即座に映像を確認可能であるため、練習の効率化も期待できる。

リアルタイム視覚フィードバックは、多視点カメラから取得した多視点映像を即座に画面に出力することで実現される。リアルタイム視覚フィードバックと同時に、取得した多視点映像はデータベースに記録する。遅延視覚フィードバックは、被写体がパフォーマンスを終えた直後に、データベースに記録された多視点映像を画面に出力することで実現する。リアルタイム視覚フィードバックにより、被写体はリアルタイムで、様々な方向から撮影した自己像を確認しながら練習可能となる。

一方で、リアルタイム視覚フィードバックでは、ユーザは常に映像が提示される画面を見続ける必要があるため、顔の向きが固定されるなどパフォーマンスへの悪影響が懸念される。本システムでは、遅延視覚フィードバックによってこの問題を解決する。遅延視覚フィードバックでは、パフォーマンス中は動作に集中し、パフォーマンス後に自己の動作を様々な方向から確認することが可能となる。



Figure 7: Real-time visual feedback: The system presents videos at the same time as subject performance.

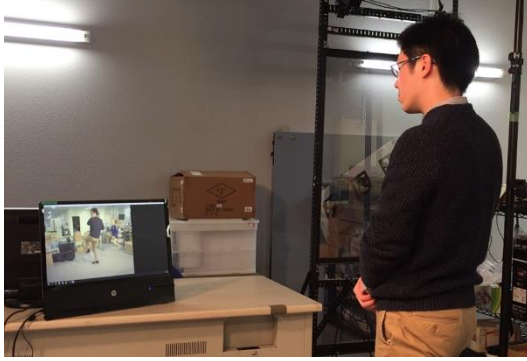


Figure 8: Delayed visual feedback: The system presents videos of the subject immediately after the performance.

4. 注視点再設定の自動化

4.1 注視点再設定処理

4.1.1 注視点の3次元位置推定

バレットタイム映像の観察者が画面のポインティング操作などにより、ある多視点画像 $I_m (m = 1, \dots, N)$ において新たな注視点を与えた際の、注視点の3次元位置推定処理について述べる。まず、3.2節で推定したカメラパラメータ情報を用いて、注視点が入力された画像 I_m と、画像 $I_n (n \neq m, n = 1, \dots, N)$ 間の基礎行列 F を算出する。基礎行列 F により、画像 I_m に与えられた注視点 $g_m(u_m, v_m)$ に対応する画像 I_n 上のエピポーラ線を求める。エピポーラ線上で、与えられた注視点の対応点を探索することにより、注視点の対応点 $g_n(u_n, v_n)$ を推定する。2枚の画像間での対応点座標にステレオ視を適用することで、注視点の3次元元座標 $M(X_t, Y_t, Z_t)$ を算出する。

4.1.2 2次元射影変換行列の算出

前節で与えた新たな注視点が同一座標で観測されるように全ての多視点画像に対して2次元射影変換を施す処理をFigure 9に示す。画像 I_m の光学中心から注視点に向かう単位ベクトル e_{mz} を、画像 I_m のカメラ座標系における新たなz軸とする。 e_{mz} と世界座標系のY軸負方向の二つのベクトルを外積した単位ベクトル e_{mx} を新たなx軸とする。 e_{mx} と e_{mz} の外積 e_{my} を新たなy軸とする。 e_{mx} , e_{my} , e_{mz} を用いて、回転行列 R'_{cm} を式(2)により求める。元画像に R'_{cm} を施すことにより注視点が画像中央で観察される。

$$R'_{cm} = \begin{bmatrix} e_{mx} \\ e_{my} \\ e_{mz} \end{bmatrix} \quad (2)$$

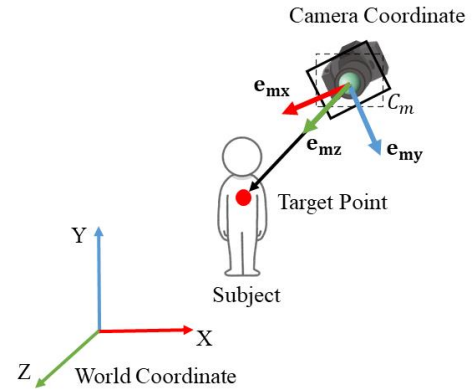


Figure 9: Conversion of a camera coordinate system.

4.1.3 内部パラメータ行列の算出

多視点映像の切替わり時に注目物体の見た目の大きさが変化しないように内部パラメータを調整する。各カメラの焦点距離の平均 f_{ave} 、注視点を与えられた画像を撮影するカメラからシーン中の注視点までの距離 d_i 、各多視点カメラから注視点までの距離 d_m を求め、式(3)により新たな内部パラメータ行列 A'_{cm} を算出する。ここで、 f'_m は式(4)により求めた値である。元画像に A'_{cm} を施すことにより、多視点映像の切替え時に注目物体の大きさが一定に保たれる。

$$A'_{cm} = \begin{bmatrix} f'_m & 0 & c_{x,m} \\ 0 & f'_m & c_{y,m} \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

$$f'_m = f_{ave} \frac{d_i}{d_m} \quad (4)$$

注視点へのズームイン・アウト処理を再現するスケール変換行列 S を加えて、式(5)により2次元射影変換行列 H_m を算出する。各画像に求めた2次元射影変換行列を施すことにより、全ての画像で注視点が同じ座標に投影され、かつ多視点映像の切替え時に注目物体の大きさが一定に保持されるバレットタイム映像が生成される。

$$H_m = SA'_{cm}R'_{cm}R_{cm}^{-1}A_{cm}^{-1} \quad (5)$$

4.2 注視点再設定の自動化

被写体に取り付けたマーカを手がかりに3次元位置姿勢を測定する手法では、被写体の動きを制限しかねず、スポーツシーンへの適用が困難である。2台以上のカメラで撮影した画像にボーン推定処理¹⁹⁾などを適用し、人物ボーン推定結果にステレオ視を適用することによって被写体の位置姿勢を求めることも可能となりつつある。しかし、スポーツ現場に持込可能な計算機資源で多視点映像を取得しつつ2台以上のカメラ映像に対しリアルタイムでボーン推定処理を行うことは現実的ではない。これらの状況を踏ま

え、本稿では、被写体の位置姿勢のリアルタイム計測が可能な RGB-D カメラを導入する。バレットタイム生成処理のリアルタイム性を損なわずに、被写体上にリアルタイムで注視点の自動再設定を行う方法について述べる。

RGB-D 映像に人物姿勢検出処理を適用し、被写体の 3 次元位置をリアルタイムで取得する。この時、被写体の推定位置は、RGB-D カメラの赤外線カメラ (Depth カメラ) を原点とした 3 次元座標である (RGB-D カメラ座標系)。3.2 節で述べたカメラキャリブレーション処理を RGB-D カメラにも適用し、外部パラメータ R_k , t_k を推定する。それらの値を用いて Figure 10 に示すように、RGB-D カメラ座標系の 3 次元座標 $M_k(X_k, Y_k, Z_k)$ を世界座標系の 3 次元座標 $M(X_t, Y_t, Z_t)$ に変換する。

RGB-D 映像から推定した被写体の 3 次元位置を新たな注視点として与えることで、4.1.1 節で述べたポインティング操作を行うことなく、注視点を自動で再設定することが可能となる。3 次元位置推定と注視点再設定処理をフレーム毎に繰り返すことにより、自動で被写体上に注視点が再設定される被写体追跡型バレットタイム映像が生成される。

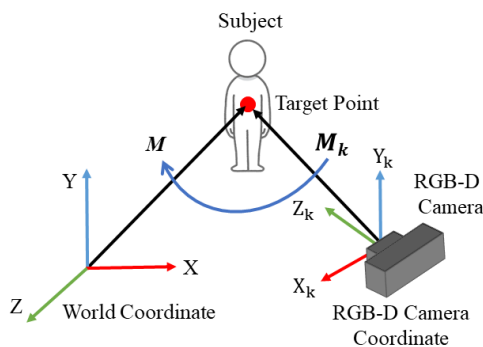


Figure 10: Conversion of RGB-D camera coordinate system to the world coordinate system.

5. 多視点映像の切替えの自動化

対象とする動作を理解するために注意深い観察を要する瞬間をキーフレームと呼ぶ。Figure 11 は、対象とする動作中にキーフレームが二つ (Posture1, Posture2) 存在し、被写体を前方 (Cam1) と後方 (Cam2) から撮影している例である。Posture1 では Cam1 の方が Cam2 より観察に適しており、Posture2 では Cam2 の方が Cam1 より観察に適しているとする。本処理では、Figure 11 のように対象とする動作中にキーフレームが複数存在し、観察に適した方向がキーフレーム毎に異なる場合を対象とする。本節では、Posture1 においては Cam1 に、Posture2 においては Cam2 に、自動で多視点映像を切替える処理について説明する。

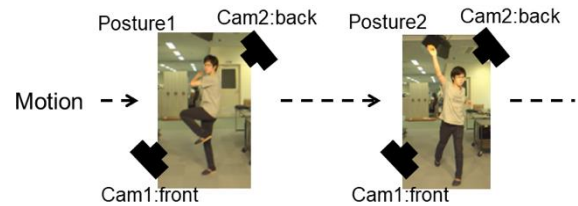


Figure 11: There are two keyframes (Posture1 and Posture2) during the target motion. Multiple viewpoint cameras capture the subject from the front (Cam1) and back (Cam2). For Posture1, Cam1 is better suited for observation than Cam2. For Posture2, Cam2 is better suited for observation than Cam1.

キーフレームの検出は、機械学習によりキーフレームを学習し、動作のジェスチャー判定を行うことで実現する。機械学習では、対象とする動作を撮影した RGB-D 映像と、その映像中の各フレームがキーフレームであるか否かを目視で判定した正解ラベルの、二つのデータを入力し学習を行う。学習結果に対して被写体の動作を撮影した RGB-D 映像を入力することで、被写体の動作の各瞬間がキーフレームである確率を出力し、確率が一定の値以上である場合に、その瞬間をキーフレームとして扱う。

本処理では、Figure 11 のように、各キーフレームと、各キーフレームにおいて観察に適した方向の関係が事前にわかっていることを前提とし、この関係を記した対応表を作成する。対応表の例を Table 1 に示す。キーフレームを検出すると、対応表を参照し、検出したキーフレームと対応するカメラ番号を読み取る。多視点映像を対応表に記されたカメラ番号まで切替えることで、多視点映像の切替えの自動化を実現する。

Table 1: Example of correspondence table of keyframes and the viewpoint

keyframe	viewpoint
Posture1	Cam1
Posture2	Cam2

6. 実証実験

6.1 実験環境

本実験では、Figure 12 に示す 20 台の多視点カメラを U 字型 (約 110 度の角度でカメラが被写体を取り囲む状態) に配置し、被写体が撮影空間内を自由に移動する様子を撮影した。約 120cm の高さに固定した棒に、クランプを用いてカメラを取り付ける。3 次元位置センサとして Microsoft 社の Kinect v2 を使用し、被写体の 3 次元位置・姿勢をリアルタイムに取得する。



Figure 12: An overview of camera setting. 20 multiple viewpoint cameras are placed in a U-shape.

多視点映像の取り込みからバレットタイム映像提示までの処理は、Intel(R) Xeon(R) CPU E5-1620 v4 3.5GHz, GPU: NVIDIA GeForce GT710 を搭載したデスクトップ PC で行う。また、本処理中の2次元射影変換処理はピクセル毎に独立して計算可能なため、GLSL を用いて GPU による並列計算を行う。撮影には Basler 社のネットワークカメラ acA1300-30gc を 20 台使用し、全てのカメラを 1 台の PC とネットワークケーブルで有線接続する。同期信号発生器として ViewPlus 社の VPTRIG-8-02 を使用する。1 台当たりカメラ 8 台までの同期が可能であり、本実験では 3 台の同期信号発生器を使用し 20 台のカメラを同期する。画像の解像度は SXGA (1280 画素×1024 画素)、撮影速度は 30fps である。Kinect v2 は多視点カメラと同様に PC に USB ケーブルで接続し、30fps で被写体追跡を行う。5 節で述べたキーフレームの学習と検出処理は、Kinect v2 付属の Gesture Builder を用いて実行する。3.2 節で述べた SfM の処理は VisualSFM²⁰⁾を用いて実行する。

6.2 適用結果と考察

提案システムを用いて、バレットタイム映像を生成した結果を紹介する。Table 2 にシステム中で行う処理と各処理に要する平均時間とその標準偏差を示す。フレームの抽出から映像出力までのシステム内部で行う処理は、全体で約 32ms 程度で完了する。一方、映像の遅延時間は約 133ms であり、これはカメラから計算機までデータを伝達する時間が大部分を占めている。約 133ms の遅延は時間的ずれを知覚させないため²¹⁾、リアルタイム視覚フィードバックとして短い時間であると考えられる。

遅延視覚フィードバックに要する時間は約 5msec であり、ほぼタイムラグなしでの遅延視覚フィードバックが可能である。多視点カメラの映像を全て指定時間分保持しておくため、遅延視覚フィードバック時も、自由に多視点映像を切替え、被写体の姿を様々な角度から見返すことが可能となっている。

注視点の自動再設定処理の結果について述べる。被写体は Figure 13 に示す経路で移動する。Figure 14 は Figure 13 の経路上の各位置①から⑥における提示映像である。移動

する被写体上の適切な位置に注視点が自動再設定されることがわかる。

Table 2: Processing time of the system

processing	average time (s)	standard deviation
frame acquisition	3.14×10^{-2}	1.62×10^{-3}
3D position estimation	4.30×10^{-5}	2.03×10^{-4}
detect gesture	1.00×10^{-6}	3.16×10^{-5}
compute H mat	3.30×10^{-5}	1.79×10^{-4}
video output	1.00×10^{-6}	3.16×10^{-5}
total	3.16×10^{-2}	

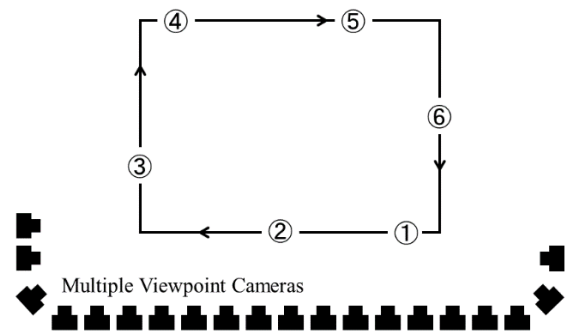


Figure 13: Movement route of the subject. The subject moves from circle 1 to circle 6.

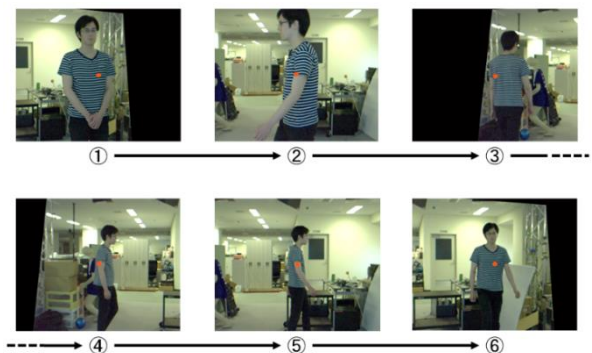


Figure 14: Subject tracking (red dots are the target points).

多視点映像の切替えの自動化について結果を述べる。本実験では、対象とする動作をピッチングとする。Figure 15 に示すように、“足の振り上げ”・“加速”・“リリース”の三つの姿勢をキーフレームとして設定する。Table 3 に、キーフレームとキーフレームを観察するカメラ番号の対応表を示す。本処理をシステムに適用した結果を Figure 16 と Figure 17 に示す。Figure 16 は、ピッチング動作の各タイミングにおいて、どのカメラを提示しているかを示している。Figure 17 は、Figure 16 のタイミング(a)(b)(c)における提示映像である。キーフレームを検出すると、多視点映像が指定したカメラ番号へ自動で切替わることがわかる。

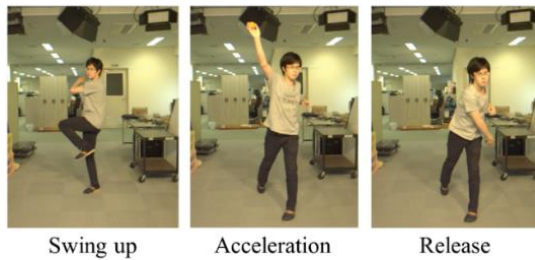


Figure 15: Postures to detect as keyframes.

Table 3: Correspondence table of keyframes and the viewpoint

keyframe	viewpoint
Swing up	1
Acceleration	20
Release	11

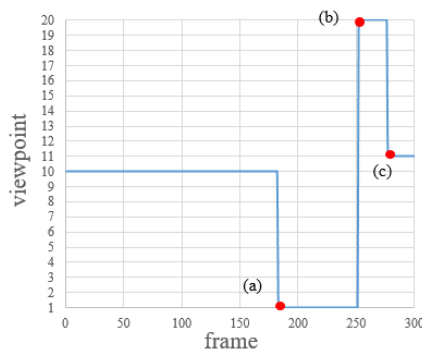


Figure 16: Viewpoint in each frame of pitching motion.

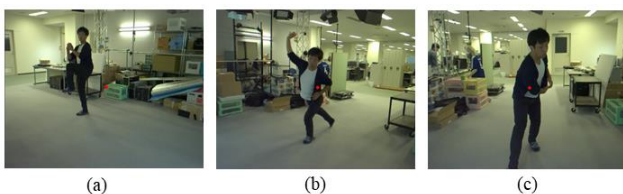


Figure 17: Output images in (a), (b) and (c) of Figure 16.

7. おわりに

本研究では、注視点再設定と多視点映像の切替を自動化したバレットタイム映像を生成し、リアルタイム視覚フィードバック・遅延視覚フィードバックの両方が可能なシステムを提案した。実証実験の結果から、リアルタイムでバレットタイム映像を生成可能であり、ほぼタイムラグなしで遅延視覚フィードバックも可能であることを示した。また、手動の操作を必要とせずに、適切な注視点の自動再設定と多視点映像の切替えが可能であることを示した。

参考文献

[1] 村山光義, 村松憲 (2007) “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-”, 慶応義塾大学体育研究所紀要, 1-15
[2] 小澤治夫, 石田讓, 岡崎勝博, 西嶋尚彦 (2003) “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”, 釧路論集, 北海道教育大学釧路校研究紀要, 第35号, 1-6
[3] 佐々木直基 (2011) “視覚的フィードバックが運動スキル獲得に与える影響”, びわこ成蹊スポーツ大学研究紀要, 121-127
[4] M.W.Kernodle and L.G.Carlton (1992) “Information Feedback and the Learning of Multiple-Degree-of-Freedom Activities.” Journal of Motor Behavior, 24(2), 187-195
[5] M.Guadagnoli, W.Holcomb and M.Davis (2002) “The Efficacy of Video Feedback for Learning the Golf Swing.” Journal of Sports Sciences, 20(8), 615-622
[6] E.Boyer, R.G.Miltenberger, C.batsche and V.Fogel (2009) “Video Modelling by Experts with Video Feedback to Enhance Gymnastics Skills.” Journal of Applied Behavior Analysis, 42(4), 855-860
[7] 本莊直樹, 伊坂忠夫, 満田隆, 川村貞夫 (2005) “HMDを用いたスポーツスキルの学習方法の提案”, 日本バーチャルリアリティ学会論文誌, Vol.10, 63-69
[8] 柴田傑, 玉本英夫, 松本奈緒, 三浦武, 横山洋之 (2014) “学習者中心のインタラクティブ舞踊学習支援システムの開発”, 電子情報通信学会論文誌, 1014-1023
[9] 中村裕基, 柳井香史郎, 中川純希, 温文, 山川博司, 山下淳, 浅間一 (2015) “重畳映像を用いた動作学習支援システムにおける映像提示視点の自動決定”, サービス学会第3回国内学会講演論文集, 236-240
[10] U.Yang and G.J.Kim (2002) “Implementation and Evaluation of “Just Follow Me”: An Immersive, VR-Based, Motion-Training System.” Presence, 11(3), 304-323
[11] J.C.P.Chan, H.Leung, J.K.T.Tang and T.Komura (2010) “A virtual reality dance training system using motion capture technology.” IEEE Transactions on Learning Technologies, 4(2), 187-195
[12] J.G.Lou, H.Cai and J.Li (2005) “A Real-Time Interactive Multi-View Video System”, MULTIMEDIA '05 Proceedings of the 13th annual ACM international conference on Multimedia, 161-170
[13] K.Ikeya and Y.Iwadate (2016) “Multi-Viewpoint Robotic Cameras and their Applications.” ITE Transactions on Media Technology and Applications, 4(4), 349-362
[14] 富山仁博, 岩館祐一 (2006) “多視点ハイビジョン映像生成システムの試作 -全日本体操選手権での中継番組利用-”, 電子情報通信学会技術研究報告, PMRU, パターン認識・メディア理解, 43-48 (2006)
[15] N.Akechi, I.Kitahara, R.Sakamoto and Y.Ohta (2014) “Multi-resolution bullet-time effect”, SA' 14 SIGGRAPH Asia 2014 Posters Article No.30
[16] N.Snavely, S.M.Seitz and R.Szeliski (2006) “Photo Tourism: Exploring Photo Collections in 3D”, ACM Transactions on Graphics, Vol25, 835-846
[17] D.G.Lowe (2004) “Distinctive Image Features from Scale-Invariant Keypoints”, International Journal of Computer Vision, 91-110
[18] C.Wu, S.Agarwal, B.Curless and S.M.Seitz (2011) “Multicore Bundle Adjustment”, IEEE Conference on Computer Vision and Pattern Recognition, 3057-3064
[19] Z.Cao, T.Simon, S.E.Wei and Y.Sheikh (2017) “Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields”, CVPR2017, 7291-7299
[20] C.Wu, “VisualSFM : A Visual Structure from Motion System”, <http://ccwu.me/vsfm>
[21] S.J.Blakemore, D.M.Wolpert and C.D.Frith (1998) “Central cancellation of self-produced tickle sensation”, Nature Neuroscience, 635-640