

練習動作と参照動作との3次元重畳フィードバック による動作トレーニング

三上 弾¹ 草地 良規¹

概要：動作の獲得を支援する手法として 3D 遅延同期ビデオフィードバックを提案する。これまで遅延同期ビデオフィードバックとして、練習者の練習動作と、対応する参照動作とを自動的に対応付けリアルタイムに練習者にフィードバックするシステムを検討してきた。遅延同期ビデオフィードバックには、リアルタイムのフィードバックが可能、可搬性がありどこでも使えるなどの長所がある。一方で、単眼カメラ1台で動作するためカメラの奥行き方向など確認しにくい部分が存在してしまう。この問題を解決するために本稿では、3D ディスプレイとステレオカメラへと拡張した3D 遅延同期ビデオフィードバックを提案する。本稿ではまず前提として、3D ディスプレイを用いて、2つの立体映像を重畳表示し人間が正しく知覚可能であるかを検証した。さらにその特性に基づいて、今後の検討の方向性を明らかにした。

Motor Learning with Use of 3D Overlay of Reference Motion and Practice Motion

DAN MIKAMI¹ YOSHINORI KUSACHI¹

1. はじめに

人間の生活において新しい動作を習得し、獲得することはライフステージを通して行われる。例えば、赤ちゃんがはいはいや歩行を獲得し、成長に伴い、走ったり、またボールを投げたり、打ったりなどの複雑な連携を伴う動作を獲得していくようになる。また、怪我や病気のリハビリのシーンなどでは、既に獲得した動作についても、新しい環境下での再学習が必要になることも往々にして存在する。このように人生の多くの場面で必要とされる動作の獲得を効果的なものにするには非常に重要なことであると考えている。

人間の動作獲得は、ニューロサイエンス(脳科学)、運動科学・バイオメカニクス、現場(例えば運動療法士やスポーツチームのコーチ)など様々な研究分野が異なる立場で検討が進められている。それに加えて近年では、VR/MR/ARや画像処理研究者らが情報処理技術のアプリケーションといった立場から動作獲得を支援するシステムを提案し

ている。Chuaらは太極拳のトレーニングシステムを開発した[1]。これはリアルタイムのモーションキャプチャと、その3Dレンダリングとの組み合わせにより、練習者の動作と対応するお手本の動作を練習者が装着しているヘッドマウントディスプレイに提示するものである。これは、お手本および練習者の全身の姿勢をモーションキャプチャにより取得していることから、練習者とお手本とを仮想空間において任意の位置関係で配置し、任意の視点からそれを観察し練習することを実現している。この視点に関する高い自由度はフォームの確認のために非常に有益であるものの、システムが非常に大がかりであり実験室環境でのみ実施が可能である。そのため、一般的な練習者が日々練習道具として使い続けられるものとは言い難い。太極拳については[2]など動作の練習に関する数多くの取り組みがなされているが、概して日常的な利用を想定したものではない。

我々も類似の課題設定から遅延同期ビデオフィードバックシステムと呼ばれる動作獲得支援システムを研究開発してきた[3]。遅延同期ビデオフィードバックシステムでは、練習者の動作と参照対象の動作とをリアルタイムかつ自動的に同期し、数秒の遅延を加えて練習者にフィード

¹ 日本電信電話株式会社
NTTメディアインテリジェンス研究所

バックする。これは、前述の Chua ら [1] に比較して、シンプルな構成でありノート PC1 台と Web カム 1 台で動作し、電源のない練習場などで日々活用することが可能なシステムとなっている。

一方で、モーションキャプチャなどを利用することなく単眼カメラ 1 台のみに基づいた提示を行うため、特にカメラの奥行き方向など情報が欠落し、フォーム全体として考えた場合、十分なフィードバックが出来ていない可能性もある。特に、動作中の人間の姿勢のように 3 次元的に複雑な動きの場合にはその影響が大きい可能性がある。

そこで、本研究では、遅延同期ビデオフィードバックのシンプルで可搬な構成、リアルタイムフィードバック可能という特長を生かしたまま、3 次元的な提示能力を高めることを目標として、3D 遅延同期ビデオフィードバックを提案する。これは、遅延同期ビデオフィードバックから極めてシンプルな拡張であり、カメラをステレオカメラへと置き換え、またディスプレイとして 3D 提示可能なものへと変更する。

本稿では、3D 遅延同期ビデオフィードバックによる動作の獲得に向けた第一段階として、以下の 2 点について検証した後、検証結果に基づいた議論を行う。

- (1) 3D 遅延同期ビデオフィードバックシステムが計算コストと観点で動作するかの検証
- (2) 3D ディスプレイを用いて、2つの 3D 映像を重畳表示した時に、人間が2つの立体像が重畳された表示として知覚できか否かの検証

2. 遅延同期ビデオフィードバック

本節では提案システムで前提とする遅延同期ビデオフィードバックシステムについて述べる。図 1 に遅延同期ビデオフィードバックの典型的なセットアップを示す。システムは PC、カメラ、ディスプレイにより構成される。PC にカメラが接続され、そのカメラで練習者を撮影する。練習者の動作は、数秒の遅延を伴い、ディスプレイ上に提示される。さらに、練習者による練習動作が検出されると、練習者による練習動作とあらかじめ登録された参照動作とを自動的に同期して提示する。

システムの処理の流れを図 2 に示す。処理は 2 つのパートに大別される。まず、リファレンス映像登録ステップにおいてリファレンスとなる映像を登録する。リファレンス映像登録ステップでは、動作中で同期させるタイミングを指定し、そのタイミングにおける動作テンプレートを作成する。実際に練習する際には、まず動作テンプレートを読み込む。システムの動作が開始すると取得された映像は、数秒の遅延を伴ってディスプレイに提示される。それと同時に時々刻々リファレンス動作の検出が行われる。リファレンス動作が検出された場合、練習者の行ったリファレンス動作と、リファレンス動画とが同期して再生される。再

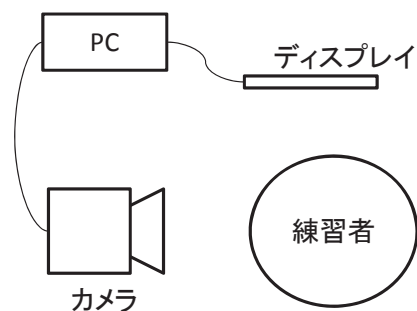


図 1 遅延同期ビデオフィードバックの典型的なセットアップ

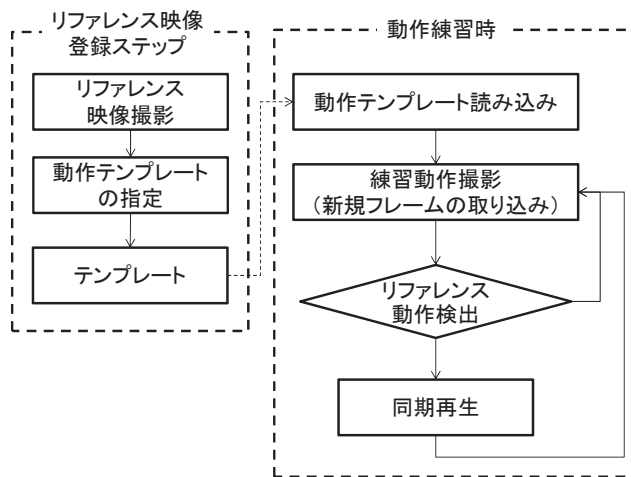


図 2 遅延同期ビデオフィードバックの処理の流れ

生にあたっては、練習者の動きと、リファレンス動画とをサイドバイサイドに並べる (図 3(a) 参照)、あるいは、それらを重畳する (図 3(b) 参照) などが考えられる。

遅延同期ビデオフィードバックシステムには 3 点の特長がある。第 1 に、1 台のノート PC (図 1 における PC とディスプレイに相当) と Web カムのみで動作させることができ、スポーツの練習場など電源の確保が難しい環境など極めて可搬性が高く、いつでも利用することができる。第 2 に、数秒という短時間の遅延で動作を提示することで、練習者が動作した時の記憶が新鮮なうちに自身の動作を確認することができる。なお、鏡のように動作しながらフォームを確認する必要がある状況については、フォームの確認のために却ってフォームが崩れるなどの指摘もなされている。第 3 に参照動作と練習動作とが同期して提示されるため、差異の観察が容易になっている。

しかしながら、人間の動作は 3 次元空間で行われるものである。そのため、ひとつの方向のみからの観測では情報が欠落し、3 次元空間における姿勢 (フォーム) の違いを正しく認識できない可能性がある。もちろん多数のカメラを設置することであらゆる角度からの観測を実現すれば解決する可能性があるが、練習環境に大量のカメラを持ち込みそれらを設置することは非現実的である。



(a) サイド・バイ・サイド表示



(b) 重畳表示

図 3 遅延同期ビデオフィードバックの提示

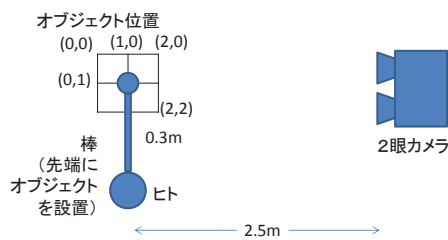


図 4 3D 遅延同期ビデオフィードバックのセットアップ



図 5 3D 遅延同期ビデオフィードバックのアウトプット例；左側に左目用映像，右側に右目用映像が提示される．左右両映像ともに，練習者の動作と参照の動作とが重畳されている．

3. 3D 遅延同期ビデオフィードバック

本稿の提案の目標は，2節の遅延同期ビデオフィードバックの課題，すなわち，動作の獲得において，動作が3次元空間で実施されるにも関わらず単眼カメラを利用することにより欠落する情報を補うことである．そのために，提案システムでは，3D ディスプレイとステレオカメラによる遅延同期ビデオフィードバックを行う．以降，第2節記載の遅延同期ビデオフィードバックと区別するために，3D 遅延同期ビデオフィードバックと呼ぶこととする．

そのセットアップを図4に示す．構成は遅延同期ビデオフィードバックと変わらない．カメラがステレオカメラとなり，ディスプレイが3D ディスプレイとなる．処理の流れも基本的に図2記載の通りである．主な違いは，同期した練習動作とリファレンス動作を提示する方法である．

従来の遅延同期ビデオフィードバックを3D表示に拡張するために，サイドバイサイドの表示を左目用，右目用と2つの視点に割り当て，練習者およびリファレンスのふたつの動作の提示に重畳表示を用いる．その例を図5に示す．実際に練習者に提示される際には，3D ディスプレイを通して重畳して提示されるため，4枚の画像が重畳されたものとなる．

4. 実装と実験

3D 遅延同期ビデオフィードバックの実現可能性を評価するために実装および実験を行った．以下にその結果について示す．

4.1 システムとしての動作

第1に3D 遅延同期ビデオフィードバックシステムが実装可能であるかを検証した．動作を同期する処理についてはアルゴリズムが従来の遅延同期ビデオフィードバックと全く同じであるため，同期の精度に関する評価は行わない．しかしながら，3D 遅延同期ビデオフィードバックでは，扱う映像が増え，また提示にあたって重畳する処理などが増える．すなわち，ここでの検証対象はマシン負荷の観点である．

実装に用いたシステムは以下の通りである．

PC Windows10 64bit, Intel Core i5-6500 3.2GHz メモリ 16.0GB

3D プロジェクター EPSON EH-TW6600(シャッター式メガネを利用)

ステレオカメラ Web カム (Logicool C920) を2台，間隔 7.5 でセット．各カメラ 640x480 ピクセルの解像度，30fps にて映像を取得．

動作中のスナップショットを図6に示す．図6に示す通り，正しく動作することが確認できた．

4.2 2つの立体画像の重畳 (1)

3次元遅延同期ビデオフィードバックでは2つの3D映像を重畳表示することで，単眼カメラでの提示では欠落する情報を補うことを狙っている．著者らの知る限り，3D ディスプレイを利用して2つの立体映像を重畳させた時に



(a) 同期前；練習映像のみが提示されている



(b) 同期再生直後；同期再生が開始され、練習動作と参照動作が重畳して提示されている。ただし、時間発展の速度が異なるためフォームの状態にはズレが生じている



(c) 同期タイミング；指定されている同期タイミング、動作タイミングが揃っている様子が分かる

図 6 3D 遅延同期ビデオフィードバックの動作例

どのように知覚されるかの調査はされてきていない。そこで、単眼カメラで最も知覚が困難となることが予想される奥行きを中心に評価実験を行った。

タスク

2つのオブジェクトを被験者に対して提示し、それのうちより手前にある方を回答させた。なお、奥行きが同じ場合には同じと回答するよう指示した。

手順

実験での配置を図 7 に示す。ステレオカメラから約 2.5m 離れた位置に立った人物が保持する棒の先端にオブジェクトを付け (図 8 参照), オブジェクト位置を図中 (0,0), (0,1) …, (2,2) の 9 点で保持し撮影を行った。各オブジェクト位置間は約 0.3m とした。撮影には、ステレオカメラを用いた。

オブジェクトは 2 種類用意した。つまり 2 種類のオブ

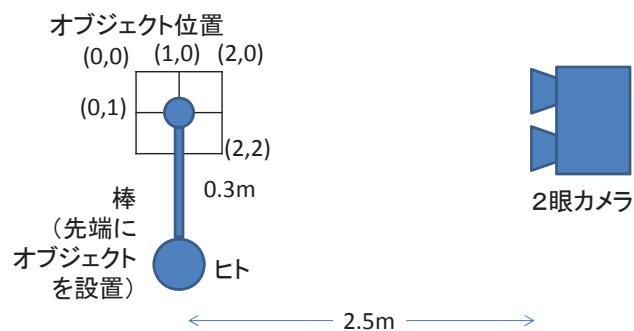


図 7 評価実験のセッティング



図 8 提示したオブジェクトと撮影時の状況

ジェクトを 9 地点で撮影した。撮影時の状況は図 8 の通りである。撮影した画像群の中から、各オブジェクトについて 1 つずつを選択し、それらを重畳させて提示し、被験者に回答させた。課題において提示された画像の例を図 9 に示す。なお被験者は 30 歳台から 50 歳台の男女 8 名とした。比較対象として、ステレオカメラ画像の 1 眼分の画像を用いて、それらを重畳提示させたものを提示し回答させた (図 9(b) 参照)。

実験では順序効果を考慮して、以下の通り行った。30 問からなる課題を 2 セット用意し、タスク 1, 2 とした。同時に被験者を 2 グループ用意し、グループ A および B とした。そして、グループ A はタスク 1 を 3D 重畳提示、タスク 2 を 2D 重畳提示で回答させた。一方、グループ B はタスク 1 を 2D 重畳提示で回答させた後、タスク 2 を 3D 重畳提示で回答させた。

結果

表 1 に結果を示す。表 1 は上から全体を通しての正解率、さらに、詳細としてオブジェクト位置の奥行き方向のズレ幅毎の正解率を示している。全体を通しての正解率を確認すると、2D 重畳提示では 46.3%であったものが、3D 提示では 70.4%と大きく向上している。ここから、2D 提示では奥行き情報が正しく伝わらなくなってしまったもの



(a) 3D での提示例

(b) 2D での提示例

図 9 被験者実験での提示例

表 1 結果

正解率	3D	2D
全体	70.4.8 %	46.2 %
奥行きズレ (2)	100 %	77.3 %
奥行きズレ (1)	81.3 %	54.2 %
奥行きズレ (0)	47.0 %	25.0 %

が、3D 提示により奥行き情報をより正確に把握できるようになっていることが分かる。すなわち、静止画での提示では、2つの立体画像の重畳提示を人間が正しく知覚できたことが示唆される。

正解率の変化をより詳しくしらべるために、2つのオブジェクトの奥行きの違いによる正解率の違いを調べた。奥行きの違いが2の場合、3D 提示での正解率は100%となり、間違ふことなく回答できたことが分かる。一方で、2D 提示の場合には、奥行きの違いが2と大きくある場合でも誤回答が見受けられた（正解率は77.3%）。奥行きの違いが1の場合にも、3D 提示と2D 提示を比較すると3D が圧倒的に高い精度で奥行き情報を知覚できたことが分かる。

なお、奥行きの違いがある状態で手前に存在するものを回答する課題と比較して、2つのオブジェクトが同じ奥行きにあり、奥行きが同じと回答することは非常に難しい課題である。この詳細については後述したい。

さらに、提示される2つのオブジェクトの水平方向の位置のずれの知覚への影響を調べた。それを図10に示す。図10から、2Dにおいて位置の違いが大きくなった時（1から2への変化時）に急激に正解率が低下しているのが見て取れる。3D 提示についても精度は低下するもののその度合は緩やかであり、3D 提示により、画面内の広い範囲で位置情報が正しく伝わっていることが示唆される。これはフォームの確認においては非常に重要な特性である。

被験者実験後の被験者との議論でも、すべての被験者において、3D 提示により2つの立体画像が重畳されたように知覚できたことが確認できており、奥行きを知覚が行えたことを確認することができた。

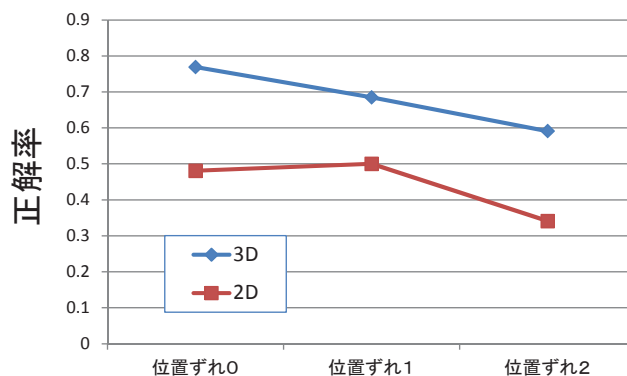


図 10 提示位置のズレに起因する正解率の変化

4.3 2つの立体画像の重畳 (2)

ここまでの被験者実験は全く異なるオブジェクトを利用したものであり、非常にシンプルな課題となっていた。そのため、我々の想定する動作の学習での利用にとって、より現実に近い環境を想定し、人物の姿勢について3D 立体重畳表示における知覚性能を評価した。本実験では実験者が、2種類の服装で類似した2つの姿勢を取った。2種類の姿勢において服装を黒と白と変化させた。両方の場合で足が地面と接する位置に違いはないものの、白い服を着たケースでよりカメラ側に上体を傾けた状態をとった。

被験者に対してまず、服装（黒）および服装（白）をひとつずつ提示し、どちらが手前にいるように感じるか確認した。提示した画像を図11に示す。その結果、すべての被験者が白が手前に感じるかと回答した。一方で、それらを重畳して提示し（図11(c)参照）同様の質問をしたところ、1名が白、1名が不明と回答したものの、残り全員が黒が手前にいるように感じるかと回答であった。

こちらは被験者実験としてのコントロールが不十分なものの、表1の結果と合わせて考えると、以下の考察をし得ると考えている。

- (1) 奥行きに一定以上の差がある場合、3D 重畳提示により高い精度で奥行き情報を利用者に伝えることが可能
- (2) その差が小さくなると、被写体の色など視認性が影響



(a) 服装（黒）の提示



(b) 服装（白）の提示



(c) 重畳の提示

図 11 被験者実験での提示例

し、濃い色の方が手前に定位して感じられる可能性がある。

ただし、現状では断定できるものではないため今後検討を続けていきたい。

4.4 2つの立体映像の重畳

ここまでは静止画を対象として評価を実施した。次に本実験では被験者に対して、2つの3D映像を、重畳表示した映像を提示し、知覚の状況を確認した。その結果、ふたつの重畳された立体像として知覚することが困難であることが分かった。原因は今後詳しく考察して行きたいが、ここで著者らが考える可能性を示しておきたい。

第1にモーションブレイカーによる影響が考えられる。今回カメラが30fpsで動作するものを用いた。そのため、特に運動動作においては激しいモーションブレイカーが発生している。そのため、立体像としての知覚が困難となっていることが予想される。

第2が背景と重畳方法の影響である。本実験における重畳提示にあたっては単純に画像全体の画素値を0.5倍し、それらを足し合わせる処理を行った。また、背景を同一の

環境で行ったため、重畳提示の際に前景部分のコントラストが低下し見にくい傾向がある。前景抽出を行うことも考えられるが、本稿の範囲では計算コストの観点などから実施していない。こちらも今後あわせて検討していきたい。

5. まとめと今後の計画

本稿では、動作の獲得に主眼を置いたシステムである遅延同期ビデオフィードバックシステムについて、単眼カメラに基づくために欠損してしまう情報、特に奥行き情報について提示可能とするために、3Dへと拡張を行い、システムとしての実現可能性を確認した。さらに、3D映像を重畳した時の人間の知覚可能性が不明なため、実験により明らかにした。

本稿での議論に基づき今後以下を計画している。

第1に遅延同期ビデオフィードバックにおける2D/3Dの適切な組み合わせの検討を行いたい。前述の通り、映像においては、立体の重畳表示としての知覚が困難なことが分かっている。このことから、映像としてフィードバックしている最中は2D重畳提示を行い、一時停止などの利用者の操作をきっかけとして姿勢をより詳しく観測するモードとして3D重畳提示を実施するなどが考えられる。

第2に高解像度化、高フレームレート化が挙げられる。現在の30fpsという低いフレームレートは動作の確認において不十分な可能性があり、加えて3D提示においてモーションブレイカーが影響して立体としての知覚を妨げている可能性もある。このようなことから、高解像度化および高フレームレート化を検討したい。しかしながら、遅延同期ビデオフィードバックのその場で簡易に実現可能という特長を維持するためには、リアルタイムでの動作が必須であり更なる処理の高速化などとセットで検討する必要があると考えている。

第3にHidden Stereo技術との連携を考えたい。Hidden Stereoとは、Fukiageらにより提案された手法であり、3Dメガネをかけた視聴者には3Dで、裸眼の視聴者には2Dで提示される映像提示手法である。3D遅延同期ビデオフィードバックの用途を考えると、選手あるいはコーチのどちらかのみが裸眼であったり、動作を行っている練習者以外に、多くの選手やコーチが指導中の映像を視聴し、その中に裸眼での視聴者が含まれる可能性が考えられる。しかしながら、3D遅延同期ビデオフィードバックの提示を裸眼で視聴すると、4つの画像が重畳提示されているものであり非常に分かりにくい。これを解決する手段のひとつとしてHidden Stereo技術が利用可能なのではないかと考えており今後検討していきたい。

参考文献

- [1] Philo Tan Chua et al., "Training for physical tasks in virtual environments: Tai Chi," Proc. IEEE Virtual Re-

- ality, pp. 87-94, 2003.
- [2] T. He et al., “Immersive and collaborative Taichi motion learning in various VR environments,” Proc. IEEE Virtual Reality (VR), pp. 307-308, 2017.
 - [3] 三上弾, 松本鮎美, 門田浩二, 川村春美, 小島明 “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,” 情報処理学会トランザクション: コンシューマ・デバイス & システム, Vol.4, No.1, pp.1-10, 2014.
 - [4] T.Fukiage, T.Kawabe, and S. Nishida, “Hiding of Phase-Based Stereo Disparity for Ghost-Free Viewing Without Glasses,” ACM Transactions on Graphics (Proc. of SIGGRAPH 2017), Vol.36, No.4, 2017.