

発達教育支援に着目した多視点映像閲覧インタフェースの開発

北村誠之^{†1} 宍戸英彦^{†2} 榎本拓哉^{†3}
亀田能成^{†2} 山本淳一^{†3} 北原格^{†2}

概要：保育士・セラピスト、児童による介入作業をビデオカメラで撮影した映像情報を用いて、発達教育従事者の技能教育支援を目的とした取り組みが行われている。本稿では、そのような作業現場での利用に適した多視点映像閲覧インタフェースを紹介する。その際、ユーザとなる発達教育従事者にとって操作性の高いインタフェースを実現することが重要である。我々は、発達支援現場における評価アンケートと要望調査を実施し、その結果に基づいたインタフェースを提案する。パイロットシステムを用いた比較実験により、提案インタフェースの操作性向上について検証する。

キーワード：多視点映像閲覧，パレットタイム映像，ユーザインタフェース，マルチタッチ操作，発達教育支援

1. はじめに

複数台のカメラを用いることで被写体を多面的に観測可能な多視点映像は、スポーツ中継や映画などの映像メディア[1]やスマートフォンアプリ[2]など多くの場面で活用されている。また、技能教育やスポーツトレーニング、遠隔作業指示のような「相手の視線、手の動きなどを色々な方向から見て学ぶ」ことが重要な協調作業においても多視点映像の利活用が進んでいる [3]。

一般的な多視点映像の閲覧法は、(1)撮影した映像を一覧表示する「ディスプレイマトリックス提示方式」、(2)コンピュータビジョン(CV)技術を用いて撮影映像を計算機内部で統合し、コンピュータグラフィックス(CG)技術を用いて任意視点からの映像を生成する「自由視点映像提示方式」[4]がある。(1)は、撮影映像をそのまま提示するため映像品質は高いが、人間が同時に観察可能な視覚情報チャンネルは1系統であるため、視点数が増加するにつれて観察が困難になるという問題が存在する。(2)は、観察者が自由に視点を操作しながら必要な視覚情報を選択的に獲得することができるが、3次元映像処理過程で生じる誤差によって提示映像の品質が損なわれる問題が存在する [5]。

Bullet-Time 映像は、(1)と(2)の長所を合わせ持つ映像提示法である。被写体の周囲に複数のカメラを、各カメラの光軸が3次元空間内の一点で交わる状態で配置し、撮影映像をカメラ配置に従って順番に切り替えることにより視点の移動感を再現する。撮影した映像をほぼそのままの状態提示するため画質が高い。さらに、我々が開発した拡張Bullet-Time 映像方式[6]では、CV 技術を部分的に導入することで、撮影空間中の任意位置を注視し続けながら視点移動することが可能である。

本研究では、Bullet-Time 映像導入現場の一つとして発達教育現場での利用に着目する。発達教育現場では保育士

やセラピストの協調作業の様子を撮影した映像を用いて、セラピスト養成訓練や子供の行動分析を目的とした取り組みが実施されている[7]。このような現場では、子供や保育士といった複数の被写体が自由に動き回るため、カメラ1台を用いた撮影では、被写体が撮影範囲外へ出てしまったり、被写体同士によるオクルージョンが発生したり、などの問題が発生する。我々は、多視点映像を用いることにより、上述した問題の解決を試みている [8]。

発達教育の現場では、子供の視線方向や手の動きといった情報が求められるため、被写体を注視し続けながら多面的な観察が可能な Bullet-Time 映像が有効である。しかし、作業現場で映像を閲覧する保育士やセラピストにとって、3次元画像処理や計算機操作に関する知識を必要とする閲覧インタフェースは使いやすいものとは言えない。そこで我々は、現場のユーザが操作しやすい映像閲覧インタフェースとして、図1に示すマルチタッチ入力による閲覧インタフェースに関する研究に取り組んでいる[9]。タッチ操作は、スマートフォンやタブレット端末の急速な普及を背景に、多くの人々にとって馴染みやすい操作インタフェースとなっている。また、提示ディスプレイに直接ポインティングするため、シンプルかつ直感的な操作が可能である。

映像閲覧に求められる操作方法は、利用現場によって異なるのが一般的である。我々は、現場の要望に応えるユーザインタフェースの考案 (Plan)、インタフェースの実装

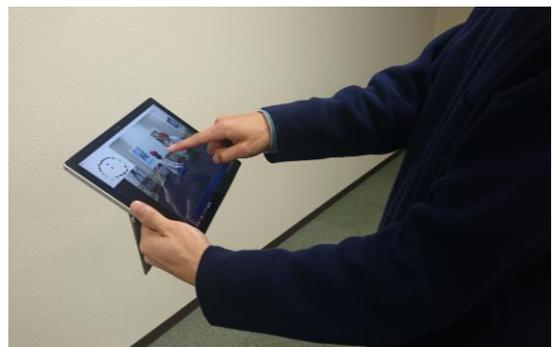


図1 タッチ入力による多視点映像閲覧

†1 筑波大学システム情報工学研究科
†2 筑波大学 計算科学研究センター
†3 慶應義塾大学文学部人文社会学科

(Do), 現場での実利用を通じたアンケート調査 (Check), その結果得られた問題点に対するシステム改善 (Act)の, PDCA (plan-do-check-act)サイクルを繰り返し実施することが, 使い勝手の良いユーザインタフェースを実現するために重要であると考えている.

本稿では, 発達教育支援に着目した **Bullet-Time** 映像閲覧インタフェースを提案する. 発達教育現場における提案インタフェースの実利用とアンケート調査を通じてインタフェースの問題点やユーザの要望を調査し, その結果をもとにインタフェースを変更・改善する. 改善前後のインタフェースを用いた評価実験により提案インタフェースの使いやすさ向上について検証する.

2. 関連研究

2.1 多視点映像の利用

Bullet-Time 映像は, その生成画質の高さから映画やスポーツ中継といった映像メディア分野での利用が進んでいる. **Eye Vision**[1]は, グランドを取り囲むように複数台のカメラを設置し, ロボットアームを用いて各カメラを連動して制御するシステムであり, アメリカンフットボールの試合中継などで実用化されている. 一方で, 視点位置や注視点は映像制作者が設定するため, 閲覧者には選択権がない. **Bullet-Time** 映像の注視点(カメラ移動時に中心となる地点)は, 通常カメラ設置時に決められた箇所に固定されるが, 動的なシーンを撮影する場合, 被写体が常にカメラ設置時の注視点に存在するとは限らない. **Akechi** ら[6]は, 多視点カメラのキャリブレーションと多視点映像から被写体の3次元位置推定処理により, 3次元空間中の任意の位置に **Bullet-Time** 映像の注視点を再設定する手法を提案している.

2.2 多視点映像の閲覧法

Mase ら[10]は, 多視点映像を閲覧する際に多数の視聴者の視聴履歴を用いることで, 閲覧視点を推薦するアルゴリズムを提案している. 視聴履歴からある時刻において各視点の選択頻度を表した **Viewgram** を作成し, それを用いて最適な視点移動を自動で生成するアルゴリズムである. 視聴者が自分で操作しなくてもよいという利点があるが, 発達教育現場のような被写体の詳細な観察が必要となるシーンへの適用は想定されていない.

Apostu ら[11]は, 多視点映像の視点選択を直感的に可能な4種類のユーザインタフェースを作成している. それぞれのユーザインタフェースは, シーン内の被写体やカメラの位置関係の把握に伴う認知努力を低減することを目的とし, 快適な視野及び視点切り替えを可能にすることを目指している. 被験者実験でそれぞれのユーザインタフェースを評価し, 複数のシーンに対してどのユーザインタフェースが最適であるかを議論している. 評価実験の結果, 提示映像内に他の視点からの方向を矢印で重畳し, その矢印を

選択することで視点を切り替えるユーザインタフェースが最も使いやすいという結論を得ているが, **Bullet-Time** 映像撮影のように膨大な台数のカメラによる撮影の場合, 画面に重畳する矢印の本数が多くなり, 映像閲覧の妨げとなる可能性がある. 本研究でも用いるストローク操作で視点を切り替えるようなユーザインタフェースは, シンプルで使いやすいという結果が出ており, 多くの一般ユーザでも抵抗なく使用できると考える.

2.3 タッチ入力を用いた視点操作

ユーザインタフェースを設計する際, タッチ操作はシンプルで直感的な操作が可能であることから広く用いられている. **Kashiwakuma** ら[12]は, 自由視点映像閲覧時における仮想カメラ操作をマルチタッチジェスチャで行うユーザインタフェースを提案している. 仮想カメラの位置姿勢の操作と数多くのタッチジェスチャの組み合わせの中から, 実験を通して最適な対応関係を定義した.

Ortega ら[13]は, 仮想空間(宇宙)内を動き回る飛行船を操作する際に求められるユーザインタラクションについて議論している. ディスプレイに対してのマルチタッチジェスチャと空中で行うジェスチャをそれぞれ提案しているが, ユーザ調査の結果, 多くのユーザがマルチタッチで行う簡単なジェスチャを好むという結果を示している. 特にタッチジェスチャの中でも, 片手の指1,2本で行うシンプルなストロークジェスチャが好まれる.

3. 発達教育支援に着目した多視点映像閲覧インタフェース

本研究では, 多視点映像の利用現場の一つとして発達教育現場に着目する. 子供とセラピスト・保護者の協同作業シーンの多視点映像を撮影し, 直感的に操作可能なインタフェースを用いた映像閲覧を目的とする(Plan). 介入作業現場では, 動的な被写体の目線や手の動きといった情報が求められる. 目線は特に該当部分を拡大し, 複数方向から観察しないと判断できないため, 目元を拡大した状態で視点を連続的に切り替え可能な閲覧システムが有効である. また閲覧時には, ユーザであるセラピストや保護者が自分の思い通りに注目領域を設定し, 視点切り替えが可能な操作が求められる. そのような操作を実現するために, 多くのユーザに馴染みあるタッチ操作が有効であると考えられる.

多視点映像の生成提示には, **Akechi** らの拡張 **Bullet-Time** 方式[6]を用いる. また, **Bullet-Time** 映像を閲覧するインタフェースは, マルチタッチ入力による閲覧インタフェースを用いて実装する(Do).

現場での実利用を通じたアンケート調査の結果, 主に三つの問題点及び要望が挙げられた(Check). 一つ目は, 入力ジェスチャの曖昧さのためにジェスチャが誤認識される問題である. 各々のジェスチャの独立性を高め, 閲覧者の思

い通りに操作できるようにするインタフェースを考案することで対応する。二つ目は、注視点位置の再設定によって映像閲覧中に生じる余白領域が映像理解の妨げとなる問題である。これに対しては、提示映像を撮影する仮想カメラのズーム値を調整することにより影響の軽減を図る。三つ目は、注視点の3次元位置が誤って推定されることにより、提示映像が乱れる問題である。これに対しては多視点映像情報を十分に活用した注視点設定処理を実装することで、推定精度を向上し対応する。

上述した発達教育現場のユーザによるユーザビリティテストを通じて、閲覧インタフェースの完成度を高める(Act)。

4. Bullet-Time 映像の生成閲覧法

本節ではインタフェースの実装 (Do)について述べる。

4.1 Bullet-Time 映像生成

被写体の周囲に等間隔で配置したカメラで多視点映像を撮影し、その映像に対して Structure from Motion (SfM)を適用することで各カメラの内部・外部パラメータを推定する[14]。Bullet-Time 映像の注視点を設定する場合は、ユーザが閲覧中の画面上で注視点の位置(2次元座標)を入力する。その注視点が他の視点画像で観察される位置(対応情報)に基づいたステレオ処理により注視点の3次元位置を算出する。注視点の3次元情報と各多視点カメラのカメラパラメータを用いて、新たに指定された注視点が多視点画像の中央で観察されるよう各撮影カメラを仮想的に回転させる(ホモグラフィ変換行列を算出する)。

4.2 マルチタッチ入力を用いた仮想カメラ操作

Bullet-Time 映像を閲覧する際に求められる仮想カメラ操作は、主に注視点の設定と視点切り替え、ズームイン・アウトである。そこで我々は、先行研究[11][12][13]やタブレット端末上で動作する一般的なアプリケーションの操作方法を参考にマルチタッチ入力による操作インタフェースを提案する。表1に仮想カメラ操作とタッチジェスチャの対応を示す。

注視点の設定方法は、注目箇所への正確な入力が必要になると考え、図2(a)に示す指1本でのタップ操作を採用する。タップ操作を認識すると、4.1節で述べた注視点再設

定処理が行われる。視点の切り替え方法は、図2(b)に示す指1本でのストローク操作を採用する。カメラは被写体の周りを囲むように並んでいるため、指をディスプレイ上で左右に移動させることで、その方向に対応した視点に切り替わる。また図2(c)に示す2本指でのストローク操作で、映像閲覧中に提示画像をウィンドウ内で平行移動させる。指の動きにあわせて画像を上下左右に移動させることが可能である。注視点へのズームイン・アウトは、図2(d)に示す2本指でのピンチイン・アウトを採用する。注視点付近を拡大して観察したい場合はピンチイン、引いて観察したい場合はピンチアウトを用いる。

表1 タッチジェスチャと仮想カメラ操作の対応

タッチジェスチャ (指本数)	仮想カメラ操作
タップ (1本)	注視点設定
ストローク (1本)	視点切り替え
ストローク (2本)	画像の平行移動
ピンチイン・アウト (2本)	ズームイン・アウト

5. 発達教育現場でのアンケート調査

4節で述べたタッチインタフェースは、実装後順次、発達教育現場での導入を開始し、子供とセラピストの協調作業様子を撮影した映像を用いた子供の行動分析や子供の学習用教材として利用が進んでいる。本節では、考案インタフェースが現場の要求に適切に役立っているか確認するために、本インタフェースを使用しているセラピスト・保護者を対象としたアンケート調査 (Check)を実施する。その結果により、考案・開発時には気がつかなかった問題点や機能について検討する (Act)。

5.1 アンケート調査の方法

アンケート実験参加者は、本インタフェースを利用して発達教育における介入作業時に共同注意(子供とセラピストの注意が一致している状態)発生の有無を確認した後、アンケートに回答する。アンケートでは、「観察したいものを見ることができましたか」「自分の望むような操作で映像閲覧が行えましたか」などを質問し、できなかったと答え



図2 タッチジェスチャ ((a)タップ, (b)ストローク1本, (c)ストローク2本, (d)ピンチイン・アウト)

た部分ではできなかった理由を具体的に記述するよう指示した。

5.2 アンケート調査の結果

セラピスト・保護者の計6名に対してアンケート調査を実施した。ユーザが使いにくかった、思い通りに操作できなかったといった問題点について考察する。

問題点の一つ目は、「ジェスチャが誤って認識される」といった意見である。例えば、ストローク操作で視点を切り替えようとしたら、誤ってすぐにディスプレイから指を話してしまいタップ操作と認識されてしまう場合である。

二つ目の問題点として、「視点を切り替えていると黒い領域（余白）が出てきて見にくい」といった意見が多く得られた。この余白領域は、拡張 Bullet-Time 映像生成処理において、注視点を再設定した際、多視点画像を仮想的に回転したことで撮影範囲外の領域が画面中に入ってしまうため生じる。

三つ目の問題点は、「望んだ箇所へ注視点の設定ができない」といった意見が多くの被験者から得られた。これは注視点再設定を行う際に、ステレオ処理によって誤った3次元位置が推定されることが原因である。誤推定された3次元位置に基づいた視点切り替えの結果、提示映像の見え目が大きく損なわれる。

被験者から得られた要望に「事前に注視点の位置を記録しておき、それを切り替えるだけで注視点を設定したい」というものがあつた。初期インタフェースでは、画面上の注目領域にタッチして注視点設定を行うが、上述したように誤った3次元位置が推定される場合がある。事前に注視点を設定しそうな地点は正確な3次元位置を記録しておき、それらを切り替えることで、注視点設定をしたいというものである。

5.3 アンケート結果に基づいた追加機能の検討

アンケート結果より、以下の追加機能の検討を行った (Act)。

- (1) タッチジェスチャの精度向上及び追加
- (2) 余白領域の調整
- (3) ラフな3次元モデルを用いた注視点推定精度の向上
- (4) 事前保存による固定注視点設定処理の追加

(1)のタッチジェスチャの精度向上では、タッチジェスチャ同士の誤認識をなくすためにジェスチャを変更する。また映像が乱れてしまった場合にプログラム起動時に初期化するジェスチャも追加する。(2)の余白領域の調整は、余白ができるだけでないような画像処理を追加する。(3)の注視点推定精度向上では、注視点の設定処理を変更し、誤推定を減らす。(4)の固定注視点設定処理は、ユーザが注目領域にタッチして注視点を設定するのではなく、事前に保存しておいた被写体などの3次元位置を用いて注視点の設定を行う。

6. 閲覧インタフェースの改善

本節では、1巡目のシステム改善 (Act)を受けたインタフェースの考案 (2巡目の Plan) と実装 (2巡目の Do) について述べる。

6.1 タッチジェスチャの精度向上及び追加

ユーザがインタフェースを操作する際、ユーザが自分の思い通りに映像閲覧できるようにタッチジェスチャを変更する。ジェスチャの誤認識として最も多く指摘されたのは、注視点設定を行う際のタップ操作である。他の操作をするときや思いがけず画面に触れてしまうと、その部分にタップしたと認識され注視点設定が行われてしまう。そこで誤って操作する心配がないようなジェスチャへの変更が必要である。また、操作ミスなどにより提示映像が乱れてしまった場合には、他のジェスチャを駆使して映像閲覧を続けるよりも、最初の状態から閲覧しなおすことが好ましいと考える。そこで初期化するための新たなジェスチャを導入する。

より独立性の高い注視点設定操作として、注目領域へのタップ操作から長押し操作に変更する。短時間注目領域に対してタッチし続けたらその領域に対して注視点設定処理を行うようにすることで、誤って設定するような場面を防ぐ。またもし映像が乱れてしまい閲覧の継続が困難となった場合に初期化できるよう、図3に示す5本指でのシェイク操作を追加する。5本の指を全てディスプレイにタッチさせ、手を振るような操作を行うことでプログラムを起動時に戻す。



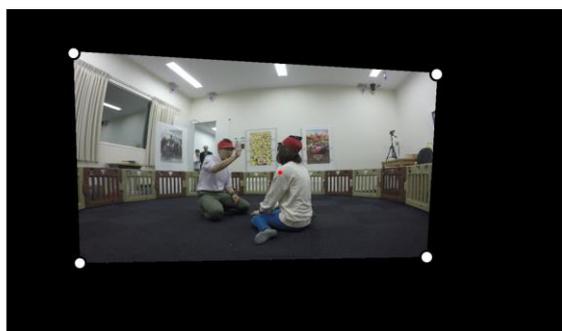
図3 シェイクジェスチャによる初期化

6.2 余白領域の調整

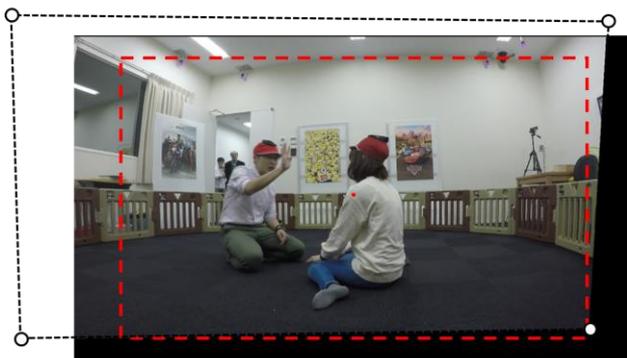
Bullet-Time 映像を生成時の画像処理の結果、提示画像内に余白領域が生じる。余白領域が存在すると、映像が見にくいだけでなく、システムが正しく機能していないのではないかとユーザに不安感を与える恐れもある。そこでこの余白領域を減らすことで Bullet-Time 映像の閲覧性を向上させることを目指す。具体的には、提示画像の平行移動処理と仮想カメラのズーム値調整により余白領域を小さくする。



図 5 余白領域が異なる Bullet-Time ((上) 元の Bullet-Time 映像 (下) 余白領域を調整した Bullet-Time 映像)



(a)



(b)

図 4 余白領域の調整処理

((a)余白が存在する画像, (b)ズーム比率調整後の画像)

まず,再設定された注視点に対して平行移動処理を行う.ユーザが注視点の3次元位置が推定すると,各カメラ画像における注視点の画像座標を透視投影行列により算出する.この画像座標群において座標の重心を求め,その重心位置に注視点が来るように画面全体を平行移動する.続いて,平行移動後の画像の注視点付近を拡大して余白を小さくする.

図 4(a)に余白領域が大きい Bullet-Time 映像を示す.この余白領域を軽減するために,図中のホモグラフィ変換後の画像のコーナー4点(白○)の座標を利用する.図 4(b)に示すような矩形領域(赤枠)を設定し,コーナー4点すべ



図 6 3次元モデルを用いた注視点設定

(上) 撮影空間の3次元モデル (下) 3次元モデルの点群

てがこの領域外に出るような仮想カメラのズーム値を算出する.この赤枠の大きさは事前に定義する必要がある.

以上の処理によって得られる Bullet-Time 映像を図5に示す.元の Bullet-Time 映像と比べると小さくなっていることが確認できる.余白領域が小さくなる一方で,注目物体の見かけの大きさが拡大し,シーン全体の状況把握への支障が懸念される.そこで,どの程度の余白ならば映像閲覧の妨げにならないかに関する予備調査を実施した.

6.3 3次元モデルを用いた注視点推定精度向上

5.2 節でも述べたように,ステレオ処理における対応点探索の失敗により誤った3次元位置が推定されることで,注視点設定に失敗する場合がある.そこで,多視点映像情報を用いて観察空間のラフな3次元形状をあらかじめ推定し,注視点再設定時にはその3次元情報を用いる処理を追

加する。

3次元モデルは、SfMとMulti View Stereo法(MVS)を組み合わせて生成する[15]。SfMで多視点画像群の対応点を推定し、撮影カメラの位置姿勢及び撮影空間の疎な3次元点群を推定する。この推定された情報に対してMVSを適用し、密な3次元点群の推定及びメッシュの作成を行う。図6に復元された3次元モデルと注視点として利用する点群の一例を示す。SfMで推定した透視投影行列を用いて、復元された点群を各カメラ画像に投影することで、2次元座標(画像座標)と3次元情報(復元モデル)を対応付ける。一つの画像座標上に複数の3次元点が投影された場合には、その中で最もデプスが小さい点を注視点に設定する。

6.4 仮想空間生成による固定注視点設定

注視点を設定する際、セラピストや保護者が設定する領域は限られている。介入作業シーンで求められる情報は被写体の視線や手の動きであることから、それらを観察できる部分に注視点を設定する可能性が高い。そこで注視点設定法の改善策として、撮影空間内で注視点として設定する可能性が高い地点の3次元位置を事前に記録しておき、映像閲覧の際にはその地点を切り替えることで注視点設定を行う方法を追加する。以降、これを固定注視点と呼ぶ。この設定方法では映像閲覧時に3次元位置を推定する必要がないため、注視点設定が失敗するという事とはなく、スムーズな閲覧が可能になると考える。セラピストに介入作業においてどの部分に注視点を設定することが多いか調査した結果、被写体の体や顔、おもちゃに固定注視点を設定するのが適切であると考えられる。

固定注視点の選択方法は、図7に示すような固定注視点とカメラアイコンを配置した仮想空間を用いる。仮想空間は、SfMで得られたカメラパラメータと事前に推定した固定注視点の3次元位置情報を描画することで生成し、仮想空間内の青丸にタッチすることで、その点に対応した注視点が設定される。仮想空間内のオブジェクト座標とサブウィンドウの画像座標の対応は、モデルビュー変換、透視投影変換、ビューポート変換を用いた座標変換により求まる。

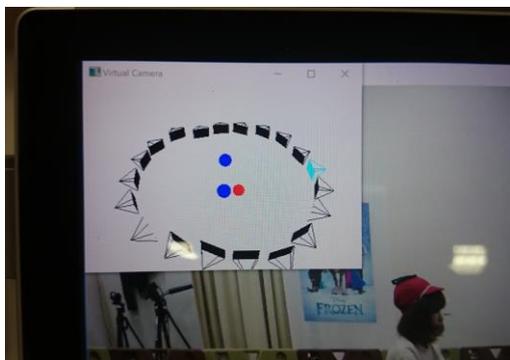


図7 サブウィンドウ上の仮想空間
(青丸：固定注視点，赤丸：現在の注視点)

7. 改善項目の評価実験

本節では、協調作業シーンを撮影した多視点映像を用いて、余白領域が閲覧者にどの程度影響を与えるのかを調査する(2巡目のCheck)。また現場でのアンケートをもとにして改良した提案インタフェースの操作性が向上したかを検証する実験計画を提案する。

7.1 撮影実験

慶應義塾大学こども支援ラボにおいて撮影実験を行った。図8に示すように、数人で協調作業を行う空間を取り囲むようにパネルを20枚配置し、パネル毎に1台のカメラを取り付けて撮影した。撮影に使用したカメラはGoPro社HERO4ブラックエディションである。3840画素×2160画素の4K映像を毎秒30枚撮影した。カメラキャリブレーションや3次元モデルの復元には、VisualSFM[16]とOpenMVSを用いる。

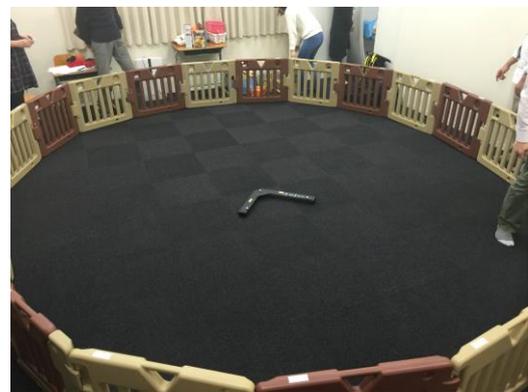


図8 撮影環境

7.2 余白領域の予備調査

予備調査では、閲覧者にとって映像閲覧の邪魔にならない余白領域の色・大きさを調査する。この調査は、理系学生(画像処理を専攻)、文系学生、セラピスト・保護者の3グループに対して実施する。現場で使うセラピストや保護者をはじめ、普段からBullet-Timeのような映像を見慣れている学生に対しても行い、調査結果に違いがでるかどうかを検証する。

色の調査では、図9に示す余白領域の色が黒、白、グレーの計3色のBullet-Time映像を用意し、一対比較法を用いて評価する。被験者に対して色の異なるBullet-Time映像を2つずつ組み合わせたビデオを計3つ見せ、余白の存在が気にならなかった方を選択させる。

大きさの調査では、余白領域の大きさが異なる三つのBullet-Time映像を用いて実施する。余白の調整は6.2節で述べた赤枠の大きさを適宜変更することで生成する。色の調査と同様に、大きさの異なるBullet-Time映像を2つずつ組み合わせたビデオを計3通り見せ、「見やすい」映像、つまり余白の大きさや映像の連続性などを総合的に考慮した



図 9 余白の色が異なる Bullet-Time 画像



図 10 評価タスクに用いる多視点映像 ((上) 成功場面, (下) 失敗場面)

上で映像理解が容易である方を選択させ、一対比較法で評価する。

7.3 改良前後のインタフェースの評価

本研究で提案した新たなインタフェースにより、映像閲覧が容易になったかどうかを検証する。実験では4章で述べた改良前のインタフェース (インタフェース A) と6章で述べた機能を追加した改良後のインタフェース (インタフェース B) の二つを用いた比較実験により改善効果を調査する。インタフェース A/B の実装には、Microsoft 社 Surface Pro 4, CPU : Intel Core i5 2.40GHz, メモリ: 4.00GB を用いた。

7.3.1 評価タスク

被験者はそれぞれのインタフェースを用いて子供の視線を判定する評価タスクを行う。図 10 にその例を示す。子供 (女) と大人 (男) が協調作業をしているシーンにおいて、子供が大人のアクションに対して反応しているか (視線を向けているか) を判定する。図 10 の場合、上のシーンは子供が大人の手を挙げたアクションに対して反応しており (成功場面)、下のシーンでは大人の顔を見させるアクションに対して違う方向に視線を向けている (失敗場面)。成功・失敗の判定の正答率と被験者が判定までにかかった時間で評価する。

実験の注意事項として、インタフェース A/B の順番は被験者によってランダムとする。実験前に一定時間インタフ

ェースの操作方法を学習するための事前練習を設ける。

7.3.2 NASA-TLX

インタフェースの主観的メンタルワークロードの測定として、NASA-TLX(NASA Task Load Index)[17]を用いる。評価タスクを実施した後、PC 版の NASA-TLX に答える。インタフェースの事前練習を行った後に NASA-TLX の説明を行い、各尺度について理解させる。また各尺度の重みづけも同時に行う。各尺度の数値及びワークロード評価 (WWL 得点) を用いて評価する。

8. おわりに

本研究では、発達教育支援に着目し、PDCA サイクルを用いて現場での利活用に資する Bullet-Time 映像閲覧インタフェースを実現するアプローチを提案した。考案したインタフェースを現場のセラピストや保護者に使用してもらい、アンケート調査を実施した。その結果判明した問題点や要望に基づいて新たな機能を検討し、インタフェースの改善を行った。注視点の推定精度向上や余白領域を減らすことで、より映像閲覧がスムーズにできるインタフェースを開発した。使いやすさの向上を検証するために、改善前後のインタフェースを用いた評価実験について検討した。

謝辞 本研究は、JST CREST(JPMJCR14E2)および、科研費 (17H01772) の助成を受けたものである。

参考文献

- [1] T. Kanade et al, “Eye Vision”, <http://www.ri.cmu.edu/events/sb35/tksuperbowl.html>, 2001
- [2] Fyusion Inc, “Fyuse”, <https://fyu.se>, access 2017
- [3] H. Joo, H. Liu, L. Tan, L. Gui, B. Nabbe, I. Matthews, T. Kanade, S. Nobuhara, Y. Sheikh, “Panoptic Studio: A Massively Multiview System for Social Motion Capture”, International Conference on Computer Vision, 2015
- [4] T. Koyama, I. Kitahara, and Y. Ohta, “Live Mixed-Reality 3D Video in Soccer Stadium”, Proceeding of the Second IEEE/ACM International Symposium on Mixed and Augmented Reality, 2003
- [5] T. Shin, N. Kasuya, I. Kitahara, Y. Kameda, Y. Ohta, “A Comparison between Two 3D Freeviewpoint Generation Methods: Player-Billboard and 3D Reconstruction”, 3DTV-Conference, 2010
- [6] N. Akechi, I. Kitahara, R. Sakamoto, Y. Ohta, “Multiple Resolution Bullet-Time Effect”, ACM SIGGRAPH-Asia, 2014
- [7] 山本淳一, 大森貴秀, 皆川泰代, “自閉性障害幼児の家庭訪問型発達支援モデルの構築と包括的評価”, 科学研究費補助金研究成果報告書 (2013), 2014
- [8] 志田全弘, 明智奈央, 北原格, 大田友一, 松田壮一郎, 山本淳一, 鈴木健嗣, “簡単に組み立て可能な多視点映像の撮影および閲覧方式”, 電子情報通信学会 CNR/PMRU 研究会, 信学技報, vol.114, no.455, CNR2014-57, pp.143-148, 2015
- [9] 北村誠之, 亀田能成, 北原格, “マルチタッチ入力を用いた多視点映像切り替えインタフェース”, 日本バーチャルリアリティ学会大会, 2016
- [10] K. Mase, K. Niwa, T. Marutani, “Socially assisted Multi-view Video Viewer”, ACM International Conference on Multimodal Interfaces, pp.319-322, 2011
- [11] S. Apostu, A. Al-Nuaimi, E. Steinbach, M. Fahrmaier, X. Song, A. Moller, “Toward the Design of an Intuitive Multi-View Video Navigation Interface Based on Spatial Information”, Proceedings of the 15th International Conference on Human-Computer Interaction with Mobile Devices and Services, pp.103-112, 2013
- [12] J. Kashiwakuma, I. Kitahara, Y. Kameda, Y. Ohta, “A Virtual Camera Controlling Method Using Multi-Touch Gestures for Capturing Free-viewpoint Video”, The 11th edition of the European Interactive TV Conference (EuroITV), pp.67-74, 2013
- [13] F. Ortega, A. Galvan, K. Tarre, A. Barreto, N. Rische, “Gesture Elicitation for 3D Travel via Multi-Touch and Mid-Air Systems for Procedurally Generated Pseudo-Universe”, IEEE Symposium on 3D User Interfaces, pp.144-153, 2017
- [14] C. Wu, “Towards Linear-Time Incremental Structure from Motion”, Proceedings of the 2013 International Conference on 3D Vision, pp.127-134, 2013
- [15] Y. Furukawa, J. Ponce, “Accurate, Dense, and Robust Multi-View Stereopsis”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.32, No.8, 2010
- [16] C. Wu, “VisualSFM: A Visual Structure from Motion System”, <http://ccwu.me/vsfm>, 2011
- [17] 芳賀繁, 水上直樹, “日本語版 NASA-TLX によるメンタルワークロード測定”, 人間工学, pp.71-79, 1996