

多視点映像を用いたコミュニケーションの記録と観察

北原 格^{†*} 志田 全弘[†] 亀田 能成^{†*}

大田 友一[#] 松田 壮一郎^{#*} 山本 淳一^{‡*} 鈴木 健嗣^{‡*}

†筑波大学大学院システム情報工学研究科 〒305-8577 茨城県つくば市天王台 1-1-1

+筑波大学計算科学研究センター 〒305-8577 茨城県つくば市天王台 1-1-1

#筑波大学 〒305-8577 茨城県つくば市天王台 1-1-1

‡慶應義塾大学文学部人文社会学科 〒108-8345 東京都港区三田 2-15-45

*独立行政法人科学技術振興機構, CREST 〒332-0012 埼玉県川口市本町 4-1-8

E-mail: †s1520786@u.tsukuba.ac.jp, {kameda,kitahara}@iit.tsukuba.ac.jp

あらまし 複数人によるコミュニケーションの記録・観察を目的とした、多視点映像の記録閲覧方式を提案する。具体的には、多視点カメラ群を簡易かつ安定に設置・撮影する多視点映像撮影方式と、それらで撮影した映像の直感的な観察を可能とする閲覧方式について述べる。本方式を児童発達介入や幼児保育の現場に導入することにより、セラピストや保育者の技能確認・共有が実現され、エキスパート養成への一助となることが期待される。本方式では、不特定多数の子供が移動する空間での撮影を考慮している。カメラを埋め込んだ幼児用ゲージを複数枚用意し、それらを繋ぎ合わせることで、撮影スペースを確保すると同時にカメラ設置を行う。カメラ操作（録画・停止）やモニタリング作業を無線通信経由で行うことにより、ワイヤリング作業が不要となり、簡単に設置可能である。多視点映像の閲覧方式には、高解像度で撮影された多視点カメラを連続的に切替える **Bullet-Time** を用いる。閲覧者は、カメラ間の位置関係（空間的連続性）を把握しながら、直感的に映像の切り替えが可能である。さらに、**Bullet-Time** の注視点を任意に設定し、そこへのデジタルズーム処理を可能とすることにより、注目箇所の詳細な観察を可能とする。最後に共同注視の確認を例題に本手法の有効性を検証する。

キーワード 多視点映像, **Bullet-Time** 映像, 発達支援, セラピスト, 技能確認

1. はじめに

情報コミュニケーション技術 (ICT) の発展に伴い、児童発達支援や幼児保育の現場への導入が進んでいる。特に、88名に1名という比較的高い有病率 [1] の自閉症児の支援には高い社会的要請がある。宮本ら [2] は、自閉症児療育のエキスパート養成支援技術では、自閉症児のコミュニケーション促進を目指し、保育士、教員、臨床心理士、言語聴覚士（セラピスト）といった支援実績のあるエキスパートや自閉症児の保護者の療育支援の機会を、ICT を用いて増加させることで支援技術の熟達度を高めている。

本研究は、セラピストの介入作業の様子を映像記録し、それを確認することにより作業の熟達度向上を実現するアプローチに基づいている。我々は、作業現場で映像を記録することには、二つの利点があると考えている。一つ目に、「見て覚える」という言葉が示すように、熟練セラピストの育児現場への介入の様子を観察することは、非熟練セラピストの育成にとって有用である。二つ目に、山本らが策定した **KEIP** 指標 [3] との親和性が高いことである。**KEIP** 指標とは、あらかじめ設定された評価項目を用いて、セラピストの熟達度を客観的に評価するものであり、チェックリストは **KEIP** フィデリティスケールと呼ばれる。その項目の中

には「おもちゃ（教材）を置く場所、使う位置は適切である」というものや「子どもの注意をひくための、表情、ジェスチャー（視覚刺激）、身体接触（触覚刺激、内受容感覚刺激）などの動作の強さ、速さ、リズム、タイミングが適切である」というものがある。これらの項目を評価する際、おもちゃが子どもや支援者の陰にいたり、子どもや支援者がカメラに背中を向けていたりすると適切な評価が困難である。カメラを手で持って環境中を移動することで、適切な角度からの観察が可能となるが、介入作業現場では、子供、セラピスト、教材が、時々刻々不規則に動き回るため、一台のカメラでくまなく観察することは困難であろう。

複数の視点から対象空間を撮影することで、上述したような動的に変化する環境においても、適切な角度からの観察が実現可能である。本研究でも多視点カメラで映像記録を行う。児童発達支援や幼児保育のような煩雑な場所での撮影を考えると、撮影装置を設置する時間はできるだけ短く、かつ、設置機材は子どもがぶつかっても怪我をする危険がないものを選定する必要がある。しかし、一般的な多視点映像撮影方式では、多数のカメラの設置やカメラ間のワイヤリング作業などに長い準備時間を要し [4]、また、三脚などでカメラを固定した場合、子供やセラピストが接触した場合、

転倒などの怪我や機材破損の恐れがある。我々は、幼児の遊具範囲を指定する幼児用ゲージにカメラを埋め込み、それらを繋ぎ合わせることで、撮影スペースを確保すると同時にカメラ設置を行う撮影装置“メディアサークル”を開発した。また、撮影操作やモニタリングを無線通信で実現することにより、ワイヤリング作業が不要となり、設置の大幅な簡易化を実現する。

多視点映像を閲覧する際、カメラ台数分のモニタ各視点からの見え方を提示すると、視点数の増加に伴い、目の前に膨大なモニタが並ぶことになり、視認性の低下が懸念される。少数台のモニタに閲覧映像を切り替えて提示すれば視認性の確保は可能であるが、撮影空間の状況やカメラ同士の位置関係が不足していると、適切な切り替え処理は困難である。多視点映像から撮影空間の3次元情報を復元し、それらを用いて任意視点からの見え方を再現する自由視点映像技術[5]を用いることにより、空間的・時間的連続性を用意に把握しながらの観察を可能にする試みが行われている[6]。この場合、3次元情報を復元するために多視点カメラをキャリブレーションする必要となるが、短時間での設置作業の妨げとなる。また、子どもやセラピストがぶつかり撮影機材が移動した場合、再度カメラキャリブレーションが必要となる。また、3次元形状推定誤差により、提示映像の品質が劣化するという問題も存在する。

Bullet-Time方式は、光軸が被写体上の1点(注視点)で交わる状態で一定間隔に並べたカメラで撮影した画像を、カメラ配置に応じて順番に切り替えるカメラワークであり、視点移動感を観察者に与えることができる。撮影した映像をほぼそのまま提示するため、映像品質が維持されること、カメラ同士の位置関係を把握しながら映像切り替えるため、適切な視点移動が可能であること、映像生成に要する計算コストが低いことなどの利点が存在する。また本研究では、子どもの表情や視線などの詳細な情報の視認が求められるため、高品質な映像生成提示が可能なBullet-Time方式との親和性が高い。一方で、従来のBullet-Time方式では、注視点は、カメラ設置時に指定した位置に固定されているため、注目対象が注視点から移動してしまった場合、適切なBullet-Time映像の生成が困難である。我々は、多視点映像にStructure from Motion (SfM) [7][8]を適用することにより、注視点の3次元的な移動が可能なBullet-Time処理を実現している[9]。本稿で紹介する提示方式もそれに基づいている。

2. 多視点映像撮影装置

本節では、簡易かつ安定に設置・撮影可能な多視点映像撮影方式“メディアサークル”について述べる。

メディアサークルは、被写体となる子どもの活動の妨げにならないよう、育児現場で一般的に用いられている幼児用ゲージによって撮影範囲を設定し、図1上に示すゲージのパネルn枚とカメラn台で構成される。幼児用ゲージを円形に組み上げることにより、図1下に示すように自立するため、安定な設置が実現される。カメラはマジックテープを用いてパネルに取り付ける。迅速かつ容易に取り外しができる反面、多少の揺れでもカメラがずれることはない。坂本ら[10]は、光軸方向が不揃いな多視点映像に対してBullet-Timeを適用すると、滑らかな視点移動感を与えることが困難なることを指摘している。本方式で提案する撮影装置では、高さが一定の幼児用ゲージを用いるため、パネルを繋ぎ合わせるだけの簡単な組み立て作業ながら、光軸方向が安定した多視点映像の撮影が可能となる。



図 1: メディアサークル

3. 注視点の動的設定が可能な Bullet-Time

3.1. SfM

多視点画像にSfMを適用し、各カメラの内部及び外部パラメータを求める。内部パラメータ行列は、カメラの焦点距離 f 、画像の中心座標 (C_u, C_v) 、外部パラメータは世界座標に対するカメラの位置と姿勢からなる。これらのパラメータから3次元空間と2次元画像空間の射影関係を表す射影変換行列 P を求める。

3.2. 注視点の動的設定

Bullet-Time映像の注視点を動的に設定するためには、まず注視点の3次元座標を推定する。閲覧者は、 i 番目の画像を閲覧しながら、新たに注視点とする画像座標 Z_i を入力する。画像 j ($j \neq i$)と画像 i のカメラパラメータから基礎行列 F を求め、注視点のエピポーラ線を求める。このエピポーラ線上で、SSD (Sum of Squared Difference) などに基づいた対応点探索を行い、

画像 j における対応点 Z_j を推定する．2 台のカメラにおける注目点の画像座標 Z_i, Z_j と射影変換行列から，ステレオ視により，注目点の 3 次元座標 (X_i, Y_i, Z_i) を算出する．

次に，得られた注視点が全ての多視点画像上で同一座標に投影されるよう，各画像に 2 次元射影変換（レンズ中心周りの回転変換）を施す．各カメラの光学中心から注目点に向かう単位ベクトル e_{mz} を新しい z 軸とし， e_{mz} と世界座標系の下方向の両方に直交する単位ベクトル e_{mx} を新しい x 軸とする．新しい y 軸は， e_{mz} と e_{mx} の外積から求まる．求めた e_{mx}, e_{my}, e_{mz} を用いると，式(1)のように注目点を画像中心に合わせたときの回転行列 R'_m が得られる．

$$R'_m = \begin{bmatrix} e_{mx} \\ e_{my} \\ e_{mz} \end{bmatrix} \quad (1)$$

上述した手法により注視点を一致させただけでは，注視物体の画面上での観測サイズが不揃いになるため，映像の連続性が損なわれてしまいます．ここでは，注目物体の観測サイズを一定に保つ処理について述べる．まず，提示画像を撮影する仮想カメラの内部パラメータ行列 A'_m を式(2)のように設定する．ここで， f'_m は式(3)によって算出される値である．各カメラの焦点距離の平均値を f_{ave} ，ズームイン操作が行われたカメラにおける注視点の奥行きを d_i ，各カメラにおける奥行きを d_m とする． A'_m を多視点画像に適用することで，観測サイズを補正する．

$$A'_m = \begin{bmatrix} f'_m & 0 & u_0 \\ 0 & f'_m & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

$$f'_m = f_{ave} \frac{d_i}{d_m} \quad (3)$$

4. 撮影実験

慶應義塾大学子ども支援ラボにおいて撮影実験を実施した．GoPro 社の GoPro HERO4 Black Edition を 20 台用いて，3840[画素]×2160[画素]（4K）の解像度の映像を毎秒 30 枚のフレームレートで撮影した．約 1 時間の撮影実験で要したデータ量はカメラ 1 台あたり約 8GB であった．4K 撮影時の水平画角は 122.6° で垂直画角は 94.4°，焦点距離は 14mm である．Bullet-Time 方式での閲覧を行うための計算機には，CPU：Intel Core i5 1.70GHz，GPU：NVIDIA GeForce GT620M，メモリ：10.0GB RAM を搭載したノート PC を用いた．

多視点画像のカメラキャリブレーションは，SfM の公開ライブラリである VisualSfM [11] を用いた．カメラを固定するメディアサークルのパネルには，ベビーサークルのミュージカルキッズランド EX マロン S を使用する．大きさは幅 71[cm]×厚み 5[cm]×高さ 56[cm] で，今回はこれを 20 枚利用した．図 2 に撮影した画像の一例を示す．



図 2：撮影画像の一例

5. 共同注視の確認作業への適用

KEIP フィデリティスケールの項目の一つに「共同注視」がある．共同注視とは，図 3 に示すように，介入作業参加者（通常はセラピストと子ども）が注視対象を共有している状態である．自閉症児は一般的に周辺環境への関心が低いため，図 4 に示すように，セラピストと注視物体を共有できていないケースが発生し，結果として，コミュニケーション能力の開発に支障をきたす場合がある．

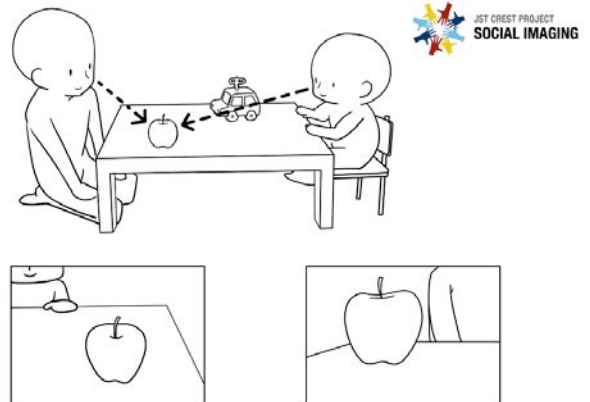


図 3：共同注視ができている例

支援技術の熟達度が高いセラピストは，共同注視を引き出しながら介入作業を進めることができる．一方で，セラピストが，実際に共同注視が起きていることを確認しようと子ども顔に目を向けると，視線が対象物体から逸れるため，子どもの注意も物体から逸れてしまい，共同注視ができていたかの確認が難しくなる，といった問題が存在する．

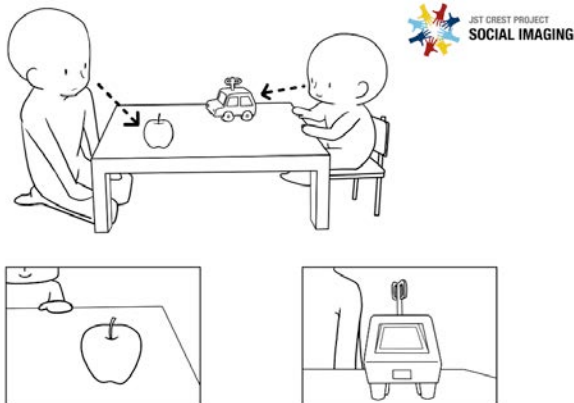


図 4：共同注視ができていない例

我々が開発した多視点映像閲覧方式を用いることにより、共同注視の有無を確認することができれば、セラピストは介入作業に集中することができ、支援技術の熟達度のより正確な計測・分析が可能となる。

4節で紹介した撮影実験において、セラピストが注視しながら指差した卓上の物体（りんご）に対して子どもが共同注視しているシーン（図5）と共同注視していないシーン（図6）の多視点映像を撮影した。

セラピストによって指差されたりんごをBullet-Timeの注視点として注視点の再設定およびズームイン処理を行いながら、図7、図8に示すように、様々な方向からセラピストと子どもの様子を観察することにより、共同注視の有無を容易に確認することができる。



図 5：共同注視が起きているシーン



図 6：共同注視ができていないシーン



図 7：共同注視が起きているシーンの観察

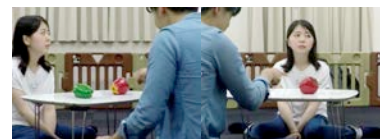


図 8：共同注視ができていないシーンの観察

6. おわりに

複数人によるコミュニケーションの記録・観察を目的とした、多視点映像の記録閲覧方式を提案した。多視点カメラ群を簡易かつ安定に設置・撮影する多視点映像撮影方式“メディアサークル”と、それらで撮影した映像の直感的な観察を可能とする閲覧方式（拡張Bullet-Time）について述べた。実際の支援現場で実施した撮影実験を紹介し、共同注視の確認を例題に本手法の有効性を検証した。

本研究は、科学技術振興機構(JST)の戦略的創造研究推進事業(CREST)における研究領域「人間と調和した創造的協働を実現する知的情報処理システムの構築」の研究課題「ソーシャル・イメージング：創造的活動促進と社会性形成支援」により行ったものである。

文 献

- [1] K.Supekar, L.Q.Uddin, A.Khouzam, J.Phillips, W.D.Gaillard, L.E.Kenworthy, B.E.Yerys, C.J.Vaidya, V.Menon, “Brain hyperconnectivity in children with autism and its links to social deficits,” Cell Reports, Volume5, Issue3, pp.738-747, November.2013
- [2] R.Wakimizu, H.Fujioka, A.Yoneyama, A.Iejima, S.Miyamoto, “Factors associated with the empowerment of Japanese families raising a child with developmental disorders,” RESEARCH IN DEVELOPMENTAL DISABILITIES, 32(3), pp.1030-1037, May.2011
- [3] 山本 淳一, 松崎 敦子, “応用行動分析学による包括的コミュニケーション発達支援プログラム：慶應早期発達支援プログラム(KEIP)の開発・適用・普及,” 子どもの健康科学, 14(1), pp.23-29, March.2014
- [4] 北原 格, 大田 友一, “大規模空間を対象とした自由視点映像生成のための3次元形状表現手法,” 日本バーチャルリアリティ学会論文誌, Vol.7, No.2, pp.177-184, 2002
- [5] R. Szeliski, “Computer Vision: Algorithms and Applications”, Springer-Verlag New York Inc., 2010
- [6] SatCam
- [7] C.Wu, S.Agarwal, S.M.Seitz, “Multicore Bundle Adjustment,” IEEE Conference on Computer Vision and Pattern Recognition, pp.3057-3064, June.2011
- [8] N.Snavely, S.M. Seitz, R.Szeliski, “Photo tourism: Exploring photo collections in 3D,” ACM Transactions on Graphics (SIGGRAPH Proceedings), 25(3), 2006
- [9] N. Akechi, I. Kitahara, R. Sakamoto, Y. Ohta, “Multi-Resolution Bullet-Time Effect”, ACM SIGGRAPH-Asia, December.2014
- [10] 坂本 竜基, 陳 鼎, “解像度を維持しつつ自然な Bullet Time を実現する射影変換の検討,” 情報処理学会 インタラクション 2014, pp.65-70, 2014
- [11] C.Wu, “VisualSFM : A Visual Structure from