

# 複数一人称視点映像閲覧における行動空間とカメラ位置姿勢の3次元可視化による効果

杉田 祐樹<sup>1,a)</sup> 樋口 啓太<sup>1,b)</sup> 米谷 竜<sup>1,c)</sup> 佐藤 洋一<sup>1,d)</sup>

**概要：**本研究では、頭部に装着したウェアラブルカメラで協調作業の様子を記録した複数一人称視点映像の閲覧において、撮影者の位置情報の把握を支援するためのインターフェースを提案する。複数の一人称視点映像に加えて、structure-from-motion で3次元復元した撮影者の行動空間とウェアラブルカメラの位置姿勢、および移動軌跡を可視化したビューを提示することによって、撮影者の絶対位置や相対的位置関係の把握を支援する。複数のタスクによるユーザ評価実験の結果、提案インターフェースを閲覧することによって撮影者の絶対位置・相対的位置関係のより正確な把握が支援されることが示され、また複数の一人称視点映像を見比べる場合や、特定の一人称視点映像に集中している割合が高くなるようなタスクにおいても、提案ビューが撮影者位置の把握に役立ち、同時に提案ビューの参照が位置情報以外の把握に悪影響を及ぼさないことも確認された。

## Effects of 3D Visualization of Workspace and Camera Poses in Browsing Multiple First Person Videos

SUGITA YUKI<sup>1,a)</sup> HIGUCHI KEITA<sup>1,b)</sup> YONETANI RYO<sup>1,c)</sup> SATO YOICHI<sup>1,d)</sup>

**Abstract:** This work introduces a new interface for browsing multiple videos captured with head-mounted cameras worn by multiple persons who are performing a collaborative work. The proposed interface is designed to assist viewers to grasp geometrical relationships of the workers such as where the worker of focus is located in workspace or how multiple workers are positioned while carrying out a certain task of interest. In addition to the multiple first person videos being displayed, the position and orientation of each camera is visualized in the 3D structure of the workspace obtained by structure-from-motion. Our user study confirmed that the proposed interface helps a viewer to grasp geometrical relationships of workers accurately. In particular, the interface provides a meaningful positional reference when a viewer examines a single first-person video or compares multiple first-person videos quickly. In addition, the interface does not prevent a viewer to understand video contents in other than the geometrical aspect.

### 1. はじめに

近年急速に普及が進んでいるウェアラブルカメラでは、カメラの設置場所を考慮する必要がなく、撮影者が両手を自由に使用できるといった利点によって空間的制約にとらわれない自然な行動の記録が可能である。頭部にカメラを

装着して撮影した一人称視点映像では作業空間に接近した詳細な手元の映像が得られ、その画面の動きは撮影者の頭の動きを示唆し、映像の中心付近にある物体は撮影者の注目を示唆するといった特色を持つ。このような特色に着目して、コンピュータビジョンの技術を用いて撮影者が何をしているかを認識する研究 [1] やライフログ等に応用する研究 [2], [3] が行われている。さらに、一人称映像を複数人数による協調作業の記録と解析に用いようとする研究も行われ始めており、複数の映像に映る重要物体といった集合視に着目して解析を行うアプローチ [4] や、HCIにおける遠隔協調システムへの応用 [5] 等が考案されている。

しかし一方で、一人称視点映像を人間が閲覧しようとする

<sup>1</sup> 東京大学生産技術研究所  
Institute of Industrial Science, the University of Tokyo, 4-6-1  
Komaba, Meguro-ku, Tokyo 153-8505, Japan  
a) yusugita@iis.u-tokyo.ac.jp  
b) khiguchi@iis.u-tokyo.ac.jp  
c) yonetani@iis.u-tokyo.ac.jp  
d) ysato@iis.u-tokyo.ac.jp

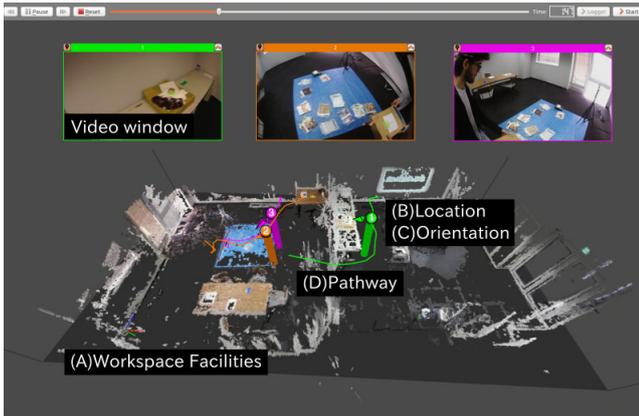


図 1: ワークスペースビューによる可視化 – (A)3次元復元された作業空間, (B) カメラの位置, (C) 向き, (D) 移動軌跡が表示されている。  
**Fig. 1** Proposed workspace-view visualization: (A)3D workspace structure, (B)camera positions, (C)orientations and (D)pathways are visualized.

際には多大な困難が伴う。一人称視点映像に特有な断続的な頭の動きは画面の動きとして顕在化する。また、撮影者の移動に伴う視点移動や、視野の断続的な変化は撮影者の位置情報の把握を困難にする。こうした閲覧の難しさは複数映像を同時に閲覧しようとする場合により顕著になる。例えば、複数の撮影者が広い作業空間内を移動しながら行なう引っ越しや部屋の模様替えのような協調作業の場合、(1) 撮影者がどのような経路を通過して作業空間内のどこからどこへと移動するか、(2) 各撮影者の両隣や向かいにいる人物は誰か、そして(3) 誰と誰がグループとなって一緒に物を運搬しているかを追いかけてその変化を逐一把握することが困難となる。

本研究では、協調作業における上記のような撮影者の絶対位置や相対的位置関係を把握する難しさに着目し、その支援を行なうために図 1 のように (A) 作業が行われている空間を 3 次元復元し、(B) 各カメラの位置と (C) 向き、そして (D) 移動軌跡をその上に表示する可視化手法 (ワークスペースビュー) を提案する。これは複数一人称視点映像の撮影者位置把握の支援という点において我々の知る限りでは初めての試みである。我々は 8 つの協調作業データセットを構築してユーザ評価実験を行ない、提案手法が協調作業における撮影者位置情報のより正しくかつ容易な把握を支援することを示し、かつワークスペースビューの参照によって位置情報以外の一人称視点映像の内容理解が妨げられないことを確認した。

## 2. 関連研究

### 2.1 一人称視点映像自体の編集

一人称視点映像には撮影者の断続的な頭の動きによって多大な画面の動きやブレが表れている。この頭の動きを滑らかにするアプローチが [6] らによって考案されている。これらの手法は一人称視点映像を早送り再生しても觀賞に耐えることを可能にする。一方、複数一人称視点映像から

重要な要素だけを抽出して映像自体を要約して短縮するアプローチでは、要約映像に用いるショットの選択に重要な人物や物体 [7] や、集合視 [8] などが用いられる。本研究ではこれらとは異なり、複数一人称視点映像を編集せずに閲覧する場合において、映像のみからでは把握の難しい撮影者位置情報を可視化するという立場を取る。

### 2.2 映像閲覧インターフェース

複数の映像を同時再生するユーザインターフェースとしては、固定監視カメラを想定したものが古くから研究され実用化も活発に行われている。並列提示された映像に加えて、作業空間のマップを提示する手法も数多く開発され、3D モデルを利用したマップ [9]、複数の魚眼レンズを元に構築したマップ [10] 等が利用されている。またマップにシーン解析の結果を可視化する手法 [11] も考案されている。本研究では、撮影者と共に移動するウェアラブルカメラで記録した映像の閲覧支援のために作業空間を俯瞰する視点からの 3 次元マップを提示し、それが積極的に利用されるかどうかを検証する。複数一人称視点映像の閲覧に関して、[12] は複数一人称視点映像を HMD 上に並べて表示し、共同作業者の視線も提示してリアルタイムでの没入型協調作業支援システムを開発した。また [13] では複数人のカメラ位置と姿勢を復元し、共同注視領域を算出して可視化を行なった。本研究ではこれらの手法で用いられているような可視化手法 (映像の並列提示や位置姿勢の提示) が複数一人称視点映像の撮影者位置を把握する上でどのように役立つかを綿密なユーザ評価実験に基づいて検証する。

## 3. 閲覧システムの実装

本研究では、広い空間内で行われ撮影者の移動が頻繁に発生するような協調作業の映像記録を、第 3 者があとで見返すような閲覧スタイルを想定したシステムを構築する。

### 3.1 作業空間とカメラ位置姿勢の復元

作業空間の復元には事前に撮影した作業空間の映像を用いて VisualSFM<sup>\*1</sup> による 3 次元復元のフレームワーク (Structure from Motion と Bundle Adjustment による粗い形状復元および Patch-based Multiview Stereo による密な形状復元) を適用する。復元された作業空間の点群には Point Cloud Library<sup>\*2</sup> を用いたノイズ除去とデータサイズの圧縮を適用する。カメラ位置姿勢の復元では、まず作業空間の撮影に用いた映像フレームや他のカメラの映像フレームとの特徴点マッチングを行ない、次に VisualSFM のフレームワークを適用して計算する。

### 3.2 閲覧システムの可視化

閲覧システムの可視化は図 1 に示したように行なう。3

\*1 <http://ccwu.me/vsfm/>

\*2 <http://pointclouds.org/>

次元復元した作業空間の点群は真上から約 45 度の俯瞰視点から見下ろしたものを表示する。このような斜め上からの視点を用いる理由は、作業空間の情報を 3 次元で提供するとともに撮影者の頭部という低い視点で撮影された一人称視点映像とのスムーズな視点の対応を考慮してのものである。本研究では画面に表示された作業空間の点群をワークスペースビューと呼ぶこととする。ワークスペースビューには (A) 作業空間の点群のほか、(B) カメラ位置を丸印で、カメラの作業空間床面からの高さを円柱でそれぞれ表示し、(C) カメラの向きを矢印で、そして (D) 移動軌跡を線で表示する。移動軌跡は現在位置の前後約 5 秒分の長さを表示する。これらの各カメラの可視化と各撮影者の一人称視点映像との対応は色と撮影者 ID が記されたラベルによって行なう。

#### 4. ユーザ評価実験

提案手法が協調作業を記録した複数一人称視点映像の閲覧において、撮影者位置情報の効果的な把握を支援することを示すために 2 種類の実験を実施する。それぞれの実験に対してタスク群 (実験 1: タスク 1-3, 実験 2: タスク 4-5) を設定する。実験 1 では、ワークスペースビューを閲覧することによって、一人称視点映像をタイル状に並べたベースラインよりも撮影者位置情報を正しくかつ容易に把握できることを確認する。一方、実験 2 ではワークスペースビューを補助的に閲覧するような場面を想定して、提示する一人称視点映像数を変化させたり (タスク 4)、様々な役割推定に取り組んだりしてもらい (タスク 5)、そのような場合でも撮影者位置情報をより正しくかつ容易に把握することができると同時に、ワークスペースビューの参照によって新たに発生する一人称視点映像とワークスペースビューとの頻繁な見比べ行動が位置情報以外の一人称視点映像の内容把握に悪影響を及ぼさないことを確認する。ユーザ評価実験ではアンケートとインタビューも実施し、閲覧行動の変化や提案手法へのフィードバック等の知見を得る。

実験 1: ワークスペースビューの閲覧によって、撮影者の位置情報をより正しくかつ容易に把握できることを確認する  
タスク 1: 撮影者の絶対位置変化の把握

1 名の撮影者について、実験用映像の再生時間内における移動経路を白地図に記入してもらう。実験用映像は 12 秒程度の長さで 1 名分の一人称視点映像だけが提示される。  
タスク 2: 撮影者の相対的位置関係を把握

実験用映像の再生終了時における 3 名の撮影者の相対的位置関係を白地図に記入してもらう (1 名分が既に記入済)。実験用映像は 30 秒程度の長さで 3 名分の一人称視点映像が提示される。

タスク 3: 同一グループに属する撮影者の把握

5 名の撮影者が複数のグループに分かれて紙への描画や箱の運搬を行なっている実験用映像を閲覧し、その中の 1

つのグループの全メンバーを白地図に記入してもらう。実験用映像は 30 秒程度の長さで 5 名分の一人称視点映像が提示される。

実験 2: ワークスペースビューを補助的に閲覧する場合において、一人称視点映像の閲覧を妨げないことを確認する  
タスク 4: 一人称視点映像数を変化させた場合の移動経路の把握

3 名または 5 名の撮影者が協力して物体を運搬している実験用映像を閲覧し、その中で 2 回目に発生した運搬の経路を白地図に記入してもらう。提示する一人称視点映像数は  $1 \cdot 3 \cdot 5$  と変化させる (FPV1, 3, 5)。FPV1 の場合は 1 名の撮影者による運搬を把握し、FPV3, FPV5 の場合は 2 名の撮影者による協調的運搬を把握する。実験用映像はそれぞれ 40 秒程度の長さである。

タスク 5: 協調作業における撮影者の役割把握

3 名または 5 名の撮影者が協調作業を行なっている実験用映像を閲覧し、特定の役割を担っている撮影者を把握して白地図に記入してもらう。実験用映像はそれぞれ 50 秒程度の長さである。本人の一人称視点映像のみから役割推定ができる場合 (ROLE1: 雑誌整理におけるタグ配布者)、手が映らず本人以外の映像から役割推定を行なう必要がある場合 (ROLE2: 掃除中にゴミ袋を縛っている人物)、そして同時に複数の撮影者の役割推定を行なう場合 (ROLE3: ブロック組み立て作業における 2 名の「何もしない」作業者) の 3 種類を行なう。ROLE2 では同時に操作対象物体の位置も記入してもらう。提示する一人称視点映像数はそれぞれ 3 名・5 名・5 名分である。本タスクでは撮影者の移動が比較的少なく相対的位置関係があまり変化しないようなシーンをを用いる。

##### 4.1 データセットの構築

本研究では 3 名または 5 名で行われる協調作業を頭に装着したウェアラブルカメラで記録して 8 種類のデータセット (1-8) を構築した。各データセットで行われる協調作業は (1) 箱の梱包と運搬、(2) ポスターの見回り、(3) 雑誌と箱の集配と整理、(4) 机を囲んだ乾杯、(5) 2 人 1 組での描画、(6) 箱の運搬、(7) 部屋の掃除、(8) ブロックの組み立てである。各データセットはそれぞれ 30-630 秒程度の長さがあり、異なる 5 つの場所のうち 1 ヶ所以上で収録された。ユーザ評価実験における各タスクでは、これら 8 種類のデータセットから互いに重複しないように多数の部分データを切り出して実験用映像として使用する。作業空間と撮影者のカメラ位置および姿勢の復元には GPU が搭載されたマルチコア環境のデスクトップ PC を用いた。作業空間の 3 次元復元には 165-289 分、カメラ位置姿勢の復元には 88-2576 分の時間を要した。複数の撮影者の映像はあらかじめ手動によるフレームの時間同期が行われている。

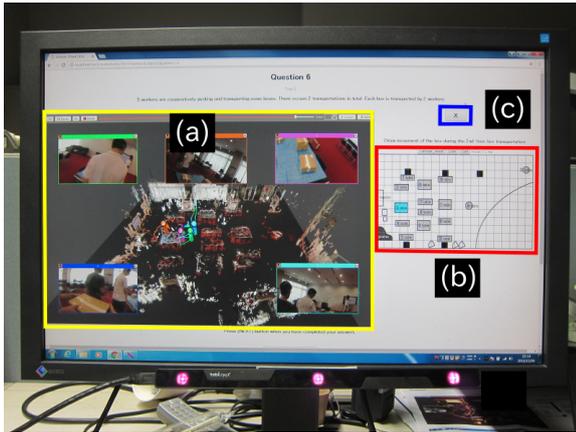


図 2: 実験用インターフェース – 24 インチモニター上に (a) 実験用映像と (b) 解答入力用キャンパスが表示されている。被験者は (c) ボタンで映像を一度だけ再生できる。

Fig. 2 Experimental user interface: (a)video for experimental use, (b)canvas for drawing answers and (c)play-button are placed on 24" screen.

## 4.2 実験の詳細

### 4.2.1 実験手順

評価実験では、被験者に実験 1(タスク 1-3) または実験 2(タスク 4-5) のどちらか一方に取り組んでもらい、各タスクではベースライン(ワークスペースビューなし)と提案手法(ワークスペースビューあり)の両方を閲覧してもらった。各タスクの開始前には本番タスクとは異なる実験用映像を使用した練習問題に取り組んでももらった。各タスクの解答終了後には難易度に関してアンケートに回答してもらった。全タスクの終了後には実験全体を振り返ってワークスペースビューによる可視化が理解しやすかったかどうかについてアンケートを実施し、その後インタビューを実施した。

### 4.2.2 実験条件

ユーザ評価実験はコンピュータビジョン分野の大学院生や博士研究者 14 名を集めて実施した。被験者のうち 8 名が実験 1 に取り組み、8 名が実験 2 に取り組んだ。各タスクではベースラインまたは提案手法に取り組む順序と提示する実験用映像の種類の組み合わせに対してカウンタバランスを取った。実験は図 2 に示すような環境下で実施した。実験用映像は 24 インチモニター上に提示され、再生はモニター上のボタンによって行なわれる。被験者には再生ボタンを押す前に解答すべき内容の把握と、白地図による作業空間の大まかな把握を行なってもらった。各映像の再生は 1 度のみとし、巻き戻し・一時停止等の操作は一切できない。各映像の再生開始前には 3 秒のカウントダウンが行われ、映像は再生終了後には速やかに非表示となる。各タスクに関する指示や解答入力も同一のモニター上で行なう。解答はモニター上のキャンパスに対してマウスで入力する。キャンパスには罫線が引かれ、机やドア、本棚等のラベルを付与した白地図が表示されている。また、実験終了後のインタビューではその様子を録音した。

## 4.3 評価方法

客観評価 被験者がキャンパス上に書き込んだ解答を手動で採点して正答率を算出する。

主観評価 難しいを 1、易しいを 7 とした 7 段階でタスクの主観難易度を回答してもらった。

## 5. 実験結果

### タスク 1-3

タスク 1-3 の正答率と主観難易度をそれぞれ図 3(A, D) に示す。Wilcoxon の符号順位検定の結果、タスク 1・2 において提案手法では有意な正答率の上昇が確認された(それぞれ  $p < .05$ ,  $p < .001$ )。主観難易度では、全てのタスクにおいて提案手法での有意な易化が確認された(それぞれ  $p < .01$ ,  $p < .001$ ,  $p < .01$ )。

### タスク 4

FPV1, FPV3, FPV5 における正答率および主観難易度をそれぞれ図 3(B, E) に示す。タスク正答率では有意差は確認されず、また FPV1 での提案手法とベースラインの双方、FPV3 での提案手法において全被験者が満点となった。主観難易度では、FPV5 において提案手法での有意な易化が確認された ( $p < .05$ )。

### タスク 5

ROLE1-3 におけるタスク正答率と主観難易度を図 3(C-1, F) に示す。正答率では、ROLE3 において提案手法での有意な上昇が確認されたが ( $p < .1$ )、その他では確認されなかった。また ROLE1-3 のそれぞれについて役割担当者を特定する項目のみの正答率を図 3(C-2) に示す。すべての場合において有意差は確認されなかった。またタスク難易度では ROLE3 において有意に易化した ( $p < .05$ )、その他では有意差は確認されなかった。

## 5.1 実験後アンケートの結果

実験終了後アンケートでは、(Q1) ワークスペースビューによる可視化を容易に理解して利用できたかどうか、(Q2) ワークスペースビューにおいて作業空間の 3D モデルを閲覧する視点は適切だったかを最も不適切を 1 とした 1-7 の 7 段階で尋ねた。Q1 では  $6.3 \pm 0.8$ 、Q2 では  $6.4 \pm 0.7$  という結果を得た。

## 5.2 インタビュー結果の集約

実験終了後のインタビューでは、被験者 14 名に次の 5 つの質問(質問 2 と 4 は 8 名)に回答してもらった。

1 ワークスペースビューにおける可視化 3 要素はどのような場面で役に立ったか?

カメラ位置: 誰がどこにいるのかの把握(7 名)、移動軌跡の把握(3 名)、空間の把握(人の位置を起点にして作業空間の背景や物体の位置が分かった)(2 名)、一緒に行動している撮影者の把握(2 名)。

カメラの向き: 意識的には使用しなかった(6 名)、全く

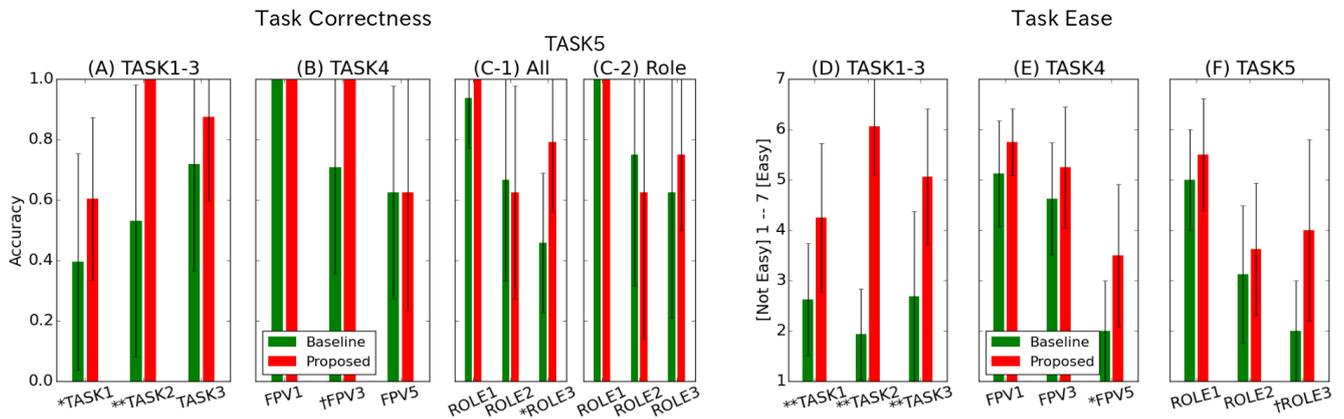


図 3: (A-C) タスク正答率と (D-F) タスク主観難易度 - (†, \*, \*\*) はそれぞれ有意水準 ( $p < .1, .05, .01$ ) を示す。

Fig. 3 (A-C)Task Correctness and (D-F)Task Ease. (†, \*, \*\*) represents each significance level ( $p < .1, .05, .01$ ).

使用しなかった (3 名)、撮影者の移動方向の把握 (3 名)、一人称視点映像との視点の対応 (映像でこちらを向いている人物が誰なのかを向きから判断)(3 名)、作業場所の把握 (2 名)、複数人による協調の有無の判断 (1 名)、位置ズレの補正 (1 名)。使用すればよかった (2 名) や向き表示がなければ難易度が上昇しただろう (1 名) と回答した被験者も見られた。

カメラの移動軌跡: 撮影者や物体の移動経路を把握 (5 名)、先読み (撮影者位置やペア形成など)(4 名)、過去位置の確認 (1 名)、位置が時々ズれるので使用しなかった (1 名)、位置だけで十分だったので使用しなかった (1 名)。2 役割推定ではどのようにワークスペースビューを利用したか?

複数の一人称視点映像を見比べて役割担当者の外見を決定した後、ワークスペースビューを用いてそれが誰であるか (ID) を位置とともに決定した (3 名)。上記の回答をした被験者のうち 2 名は位置の解決が終わらないうちに映像が終了してしまい時間が足りなかった、と述べた。映像がどの向きから撮影されたかを把握して映像に映っている人物が誰であるかを把握しながら役割担当者を決定した (2 名)。映像に映っている人物が誰でどこにいるのかわかりづらかった (2 名)、撮影者がもっと動いていれば軌跡を利用した (1 名) という意見もあった。

3 ワークスペースビューによる可視化はどのような場面で役に立たなかったか / あることによって却って混乱したか?

役に立たなかった場面: 撮影者が動いていない (位置のみ役に立った)(4 名)、手元での作業内容を把握する (1 名)、手元が見えない (位置のみ役に立った)(1 名)、運搬 (1 名)。また、移動軌跡のうち後方の軌跡が不要 (1 名) あるいは前方の軌跡が不要 (1 名) という意見もあった。

却って混乱を招いた場面: 特になし (6 名)、撮影者位置推定の誤り (3 名)、向き推定の誤り (1 名)、撮影者が動かない (1 名)、関係のない人物が映り込んでいる (1 名)。

4 提示する一人称視点映像数が増加した場合に、どのような困難がどのように変化したか?

ワークスペースビューなし: 見るべき映像数が増加して処理量が増大 (3 名)、FPV3 がちょうど良い (2 名)、心理的の圧迫感が増大 (1 名)、FPV1 は情報量が少なく逆に難しい (1 名)。

ワークスペースビューあり: FPV5 では何を見るべきかわからず混乱した (3 名)、見るべき映像数が増加して処理量が増大 (1 名)、物体の移動を追う際に別の場所で行われている作業を同時に把握するのが困難だった (1 名)、ワークスペースビューからは一人称視点映像に映る人物が誰であるかを把握できなかった (1 名)、視点对応の困難が増大した (1 名)、一人称視点映像とワークスペースビュー間の視線移動が大変だった (1 名)、撮影者位置表示に対応する一人称視点映像を探す困難が増大した (1 名)。

#### 5 提案手法への改善要望点

位置・向き・移動軌跡の推定精度向上 (4 名)、撮影者位置表示と一人称視点映像を近づけてほしい (4 名)、グループになっている撮影者の映像は互いに近づけてほしい (1 名)、撮影者位置表示から一人称視点映像へのリードが必要 (1 名)、3D モデルの粗さの改善 (1 名)、矢印を意識的に見なくても体の向きが分かる可視化 (1 名)、映像が多い際に何を見るべきかの示唆 (1 名)、映像が多い際の心理的の圧迫感の解消 (1 名)、一人称視点映像のラベルが見づらい (1 名)、映像内での人や物体のローカルな位置把握とワークスペースビュー上でのグローバルな位置把握が二度手間 (1 名)、オブジェクトを表示してほしい (1 名)。

### 6. 実験結果に対する考察

実験 1: ワークスペースビューの閲覧によって、撮影者位置情報をより正しくかつ容易に把握できることが確認された

タスク 1-3 の全てにおいてタスク正答率が上昇し、主観難易度が有意に易化した。特に相対位置の把握で大幅な改善が見られた。

実験2: ワークスペースビューを補助的に閲覧する場合でも、一人称視点映像の閲覧を妨げないことが確認された

タスク4-5ではタスク5のROLE2以外の全てにおいて正答率の悪化傾向は見られなかった。またタスク5で役割担当者決定の項目のみに着目した場合でもROLE2以外では正答率の悪化傾向は見られなかった。ROLE2の役割担当者決定の正答率では有意でない悪化傾向が見られたが、このタスクでは役割担当者本人以外の映像を長時間閲覧する必要がある。カメラの向き表示を使用しなかった、または使用しても映像中の人物との対応付けに役に立たなかった場合、一人称視点映像のみを見比べて役割担当者の外見を決定してからワークスペースビューを参照して、その人物のIDと場所を決定する、という方針を取った被験者では「誰」であるかを把握する作業が二度手間となって処理時間が不足してしまったと考えられる。一方、主観難易度では全てのタスクにおいて易化傾向が観察されている。

#### ワークスペースビューの使用方法に関する知見

カメラ位置表示は撮影者位置の決定のほか撮影者位置を起点とした空間の把握等で役に立った。カメラ向き表示は作業場所、協調作業の有無、移動方向、視点の対応等の様々な用途に用いられた一方で、使用方法を見出せず意識的に使用しなかった被験者が多数であった。カメラの移動軌跡は経路把握だけでなくペア形成や位置の先読み、動きを用いた人物との紐付け等の使用方法が見られた。

#### 可視化手法の改善すべき点

最も多い改善要望はカメラ位置姿勢の復元精度と一人称視点映像の提示位置に関するものであった。現在のカメラ位置姿勢復元では多くのエラーフレームを含み、復元に成功した前後のフレームによる単純線形補間を行なっている。[8]で用いられているようなエピソード幾何による補間やモーションモデル等を取り入れて、より厳密に補間していく必要があると考えられる。映像の提示位置に関しては、カメラ位置表示との距離やグループになっている撮影者の映像同士の距離を近づけて視線移動を軽減する必要がある、ワークスペースビュー内に映像を配置するほか、リードの付与や色対応の改善等の対策を検討していく必要がある。またワークスペースビューと一人称視点映像との視点の対応がスムーズでなかった点も挙げられる。このことはカメラ向き表示の利用度が低かった点と密接に関連していると考えられ、向きの意識的利用の拡大を促す可視化や、あるいは一人称視点映像を閲覧している間にワークスペースビューを回転させて映像との視点を一致させる、といった無意識的な視点対応をサポートする対策が効果的に働く可能性がある。また本研究では映像の事後閲覧を想定しているものの、反復・一時停止を認めない条件で実験を実施して位置把握への効果を確認した。SLAMやセンサフュージョン等を活用した高速な位置姿勢復元手法を用いること

でリアルタイムシステムへの応用にも検討の余地がある。

## 7. おわりに

本研究では、協調作業を記録した複数一人称視点映像の閲覧において、撮影者の絶対位置や相対的位置関係の把握を支援するために、撮影者の行動空間とウェアラブルカメラの位置姿勢および移動軌跡を可視化したワークスペースビューを提案した。ユーザ評価実験の結果、提案ビューの閲覧が効率的な撮影者位置の把握を支援し、かつ位置情報以外の一人称視点映像の内容理解を妨げないことを確認した。今後は一人称視点映像とワークスペースビューとの配置関係や視点の対応に関する可視化の改善、ユーザインターフェースとして実用的な場面でのテスト等を検討していきたい。

謝辞 本研究はJST CRESTの助成を受けて行われた。

#### 参考文献

- [1] Pirsiavash, H. and Ramanan, D.: Detecting Activities of Daily Living in First-person Camera Views, *CVPR* (2012).
- [2] Ishiguro, Y., Mujibiya, A., Miyaki, T. and Rekimoto, J.: Aided Eyes: Eye Activity Sensing for Daily Life, *AH* (2010).
- [3] Leung, T.-S. and Medioni, G.: Visual Navigation aid for the blind in dynamic environments, *CVPRW* (2014).
- [4] Kera, H., Yonetani, R., Higuchi, K. and Sato, Y.: Discovering Objects of Joint Attention via First-Person Sensing, *CVPRW* (2016).
- [5] Kasahara, S. and Rekimoto, J.: Jackin: Integrating first-person view with out-of-body vision generation for human-human augmentation, *AH* (2014).
- [6] Liu, S., Yuan, L., Tan, P. and Sun, J.: Steadyflow: Spatially smooth optical flow for video stabilization, *CVPR* (2014).
- [7] Ghosh, J., Lee, Y. J. and Grauman, K.: Discovering important people and objects for egocentric video summarization, *CVPR* (2012).
- [8] Arev, I., Park, H. S., Sheikh, Y., Hodgins, J. and Shamir, A.: Automatic editing of footage from multiple social cameras, *ACM TOG*, Vol. 33, No. 4, p. 81 (2014).
- [9] Rieffel, E. G., Girgensohn, A., Kimber, D., Chen, T. and Liu, Q.: Geometric tools for multicamera surveillance systems, *ICDSC* (2007).
- [10] DeCamp, P., Shaw, G., Kubat, R. and Roy, D.: An immersive system for browsing and visualizing surveillance video, *ACM Multimedia* (2010).
- [11] Roth, P. M., Settgest, V., Widhalm, P., Lancelle, M., Birchbauer, J., Brandle, N., Havemann, S. and Bischof, H.: Next-generation 3D visualization for visual surveillance, *AVSS* (2011).
- [12] Kasahara, S., Ando, M., Suganuma, K. and Rekimoto, J.: Parallel Eyes: Exploring Human Capability and Behaviors with Paralleled First Person View Sharing, *CHI* (2016).
- [13] Park, H. S., Jain, E. and Sheikh, Y.: 3d social saliency from head-mounted cameras, *NIPS* (2012).