

音環境比較による会話場検出を利用した 共有体験データのインデキシング

遠山 魁¹ 角 康之¹

概要: 本研究では音環境比較手法で得られる会話場情報をインデクスとした体験映像の閲覧支援への応用を提案する。本研究が指す体験映像とは、会話相手などが映りこんだ個人の体験を記録した一人称映像のことを指す。また、音環境比較による会話場検出とは、音の特徴量から会話グループを検出する手法のことを指す。体験映像の閲覧支援を試みる理由としては、カメラなどの性能の向上によって記録が容易になり、長時間の映像を得ることが可能になった一方で、閲覧したいシーンを振り返るのに時間がかかってしまうからである。そして音環境比較による手法で得られる会話相手と会話を行っている場所の情報は、映像閲覧のインデクスとして有用であり応用が可能であると考えられる。そこで、本研究ではこの音環境比較による会話場検出手法から得られた情報を映像閲覧へ応用することを目的とする。具体的には会話場情報をもとに映像のシークバーにインデクスを振ることによって、追体験と体験共有の支援を行う。1時間ほどのポスター発表の場で体験映像の記録と会話場の検出を行い、映像に会話場情報のインデキシングした。そして、本システムの評価実験を行った結果、会話場情報によるインデキシングが閲覧の支援に役立つことを示した。

Indexing of Shared Experience Data Depending on Conversation Groups Based on Auditory Similarity

KAI TOYAMA¹ YASUYUKI SUMI¹

1. はじめに

本研究は音環境比較による会話場検出手法の応用例として体験データの閲覧支援システムの提案を行う。本研究が指す体験データとは、会話相手などが映りこんだ一人称視点映像を指す。また、共有体験データと表現する理由としては、アクションカメラなどによって記録された一人称視点の体験映像が個人のために限らず他者と共有するための役割を持っていると考えられるからである。そして、体験データの閲覧支援を試みる理由としては以下のとおりである。ウェアラブルカメラなどの登場により、日常の体験を常時記録可能となった一方で、長時間にわたる映像の中から特定のイベントを見つけ出すことが難しいためである。そこで、本研究は主に会話場という抽象的な空間に着目し、会話をしていたり、同じ体験を共有しているグループ

と場所の情報を基にインデキシングすることで映像の閲覧支援を行う。誰と、どこで会話をしているという情報を会話場情報と呼ぶ。そして、会話場情報を得る手法として音環境比較による会話場検出手法を用いる。音環境比較による会話場検出手法を用いる理由としては2点挙げられる。まず1点目は、図1に示すような部屋のサイズや人の密度が異なる状況においても検出が可能である。次に、図2のような物理的空間が離れた会話グループも柔軟に検出できる点である。この手法においては似たような音を聴いている人をグルーピングする手法であるため、会話というインタラクションに限らず傾聴であったり、1対多のスピーチのような状況下でも会話グループとみなす。ただし、同じ体験を共有していると考えられるため、会話というインタラクションに限らず傾聴をしているようなグループも本研究では会話グループに含めインデキシングを行っていく。

本研究は、会話というインタラクションに着目する。そして、会話場の情報は音環境比較による会話場検出手法

¹ 公立はこだて未来大学
Future University Hakodate



図 1 会話者間の距離が近いグループ



図 2 空間的に離れていて物理的サイズの大きい会話グループ

によって得る。具体的には、”どこで誰と会話している”といった情報が得られ、この情報をインデクスに利用する。

2. 関連研究

2.1 音環境比較による会話場検出

音環境比較による会話場検出とは、中蔵らが提唱した手法である [1]。中蔵らは、同じ体験を共有している人は似た音にさらされているという仮定の下、音の類似性から会話グループの検出を行う手法を提案した。参加者それぞれに小型の PC 端末を持たせ、端末からお互いの会話音声や周囲の音を計測する。そして、計測された音声の有音区間であればフーリエ変換 (FFT) し、得られた音声の 50 から 1600Hz の周波数特性を抽出する。また、抽出した周波数特性をベクトルとみなし、端末の組み合わせ分比較しあい類似度を算出し、閾値を超えた組み合わせを会話グループであると判定する。この手法の強みとしては、距離によらない検出手法である点と、拘束性の低さが挙げられる。会話グループや近隣推定を行う研究の多くは、距離に着目した検出を行っている。従来研究の多くが会話者間の距離に着目している理由は、人類文学者の Edward Hall が定義した対人距離 [2] を踏襲しているためである。それらの具体的な手法としては、赤外線 [3, 4, 5] や Bluetooth [6]、WiFi [7, 8] などといった電波、または超音波 [9] などの様々なセンサから推定を行っている。しかし、図 1 に挙げられるような混雑し、相手との距離が近いような状況下であったり、図 2 のような複数人が同じスピーチを聴いているような距離が離れたような状況下において、正しい結果を得ることは難しいと考えられる。また、精度を上げるためにセンサを複合的に用いたり、モーションキャプチャなどを用いる場合は拘束性が高くなり実環境での計測が難しい。また、Neary の検出精度については適合率が 96.6%、再現

率が 67.9% であり、実環境においても十分な精度で検出することができる。そして、Neary のような同時録音型の手法を利用する理由としては、シグネチャ型 [10, 11] のようなあらかじめ録音した場所ごとの環境音と比較するだけでは、会話であったり、物音などといった非定常音によるグルーピングができないためである。本研究で得たい会話場情報は、具体的には会話や傾聴であったり、抽象的には同じ体験を共有しているグループを検出したいため、同時録音型による手法が好ましいと考える。以上により、本研究ではこの中蔵らのアルゴリズムを参考にし、会話場検出を行う。しかし、中蔵らの手法では会話している人の組み合わせの情報のみに着目しており、会話グループが存在する場所については考慮していない。そこで本研究では計測端末を場所に置くことで、会話グループの大まかな場所の検出も行う。

2.2 インタラクシオン情報をライフログに応用した研究

会話などといったインタラクシオン情報をライフログへ活用した研究としては角らのコミックダイアリ [12] や小関らのぱらぱらアニメによる体験データの要約・編集支援システム [13] が挙げられる。角らのコミックダイアリは博物館などでのコミュニティウェアで得られたライフログ情報を基に漫画の生成を行う研究である。参加者は PalmGuide と呼ばれる専用の端末を持ち歩き、博物館内の展示物の見学情報や、他の参加者との名刺交換を行うことで場所とインタラクシオンの情報を蓄積することができる。そして、得られた情報を基に漫画を生成することで体験の気づきや他の参加者との体験の共有を促すことを目的としている。次に小関らのぱらぱらアニメによる体験データの要約・編集支援システムは、ポスターセッションでのインタラクシオン情報を基に映像の閲覧と編集を行うシステムの提案をしている研究である。この研究も同様に赤外線タグを用いて誰がどのポスター発表を聴いているか、議論しているかといった情報の蓄積を行っている。また、会話している人数に応じてシーンの切り出しや強調表示をしている。これらの研究は赤外線を用いており、赤外線によって得られる近接性からインタラクシオンの検出を行っている。距離からグループの推定を行う手法の問題点としては、部屋のサイズや人の密度が異なる状況においては会話者間の距離が一定とは限らないことが挙げられる。そこで本研究では、音環境比較による手法を用いて会話というインタラクシオンの検出を行い、インデキシングを行う。

本研究と同様に一人称視点映像にインデキシングを試みる研究としては、Higuchi らの EgoScanning [14] が挙げられる。Higuchi らは、一人称手掛かりと呼ばれる情報をインデクスとすることで、日常の体験を記録した長時間の一人称映像の高速閲覧支援を目的としている。一人称手掛かりとは、静止と動作の情報、手の情報、人の情報の 3 つの情

報のことを指し、一人称映像中の手掛かりになる情報であると定義している。静止と動作の情報は記録された体験映像が止まったり、動いたりする映像の変化を指す。つまりは、撮影者が急に止まったり動きだしたりする様子のことである。また、手の情報は撮影者が物を持ったりする際の手の映り込みの情報を指す。この情報から撮影者の興味や動作の対象を得ている。そして人の情報とは撮影者と対面している相手の情報を指す。この情報は本研究と同様に会話などといったインタラクションを行う相手の情報を指している。これらの情報を画像認識によって検出し、これらの情報が含まれる部分以外の映像を高速再生することで閲覧の支援を行っている。この研究の強みとしては、人とのインタラクションに限らず撮影者の行動も含めた複数の指標を用いている点が挙げられる。また、それらの複数の指標を画像処理のみを用いている点も挙げられる。一方で、この画像処理による限界としては、照明の状況が悪い場合には手と人の検出がうまくできないことが挙げられる。また、会話相手ではないが近くにいるような状況下においても誤検出することが考えられる。また、場所や人数、相手が誰であるかといった情報は加味していない。そこで、本研究はインデックスの情報は会話というインタラクションにのみ着目し、音環境比較による会話場検出を行う。また、環境側にも計測端末を設置することで会話グループの位置の検出を試みる。

3. 音環境比較による会話場検出を利用した共有体験データのインデキシング

本研究は音環境比較によって得られる「誰と誰が、どこで会話をしている」といった会話場の情報をインデックスにすることで、一人称映像の閲覧支援への応用を目的とする。この章では音環境比較による会話場検出を利用した共有体験データのインデキシングの実装方法について述べる。最後に本システムを利用した実験と評価について述べる。

3.1 音環境比較による会話場検出

本節では、音環境比較による会話場の検出手法について述べる。音環境の計測方法については人ノード、環境ノードともに Nexus5 を用いた。そして、Nexus5 に内蔵されたマイクを使って音の計測を行う。また、同一の端末に揃えた理由としてはマイクの周波数特性を考慮したためである。端末の装着や設置については図 3、4 に示すとおりである。人ノードは内臓マイクを上向きにして胸ポケットなどに装着する(図 3)。また、会話グループの場所を検出するための環境ノードは、任意の地点や物に設置をする(図 4)。このように特定の場所、物に設置をする理由としては、人ノード同士のグルーピングと同様に、特定の環境に固定された端末とグループであると判定されることで、グルーピングされた人物が特定の環境付近で会話をしているとみ



図 3 端末を装着する様子



図 4 環境側に端末を設置した様子

なすことができるからである。

3.1.1 会話場検出のアルゴリズム

本節では音環境比較による会話場検出のアルゴリズムの詳細について述べる。本システムで用いるアルゴリズムは主に Neary の手法を参考としている。また、会話場検出に使う音声はあらかじめ計測し終えてから処理を行う。リアルタイムによる判定も可能ではあるが、多少の時間ズレが生じる問題があり、ライフログへの活用にあたって精度の粗い判定結果を用いることは不適切であると考えたためである。図 5、6 に会話場検出のアルゴリズムのブロック図を示す。各ノードから計測された音声の周波数特性の抽出、各ノードの周波数特性の比較、1 秒ごとの会話場の判定、得られた判定結果のスムージングの 4 つの段階にわけて説明する。

まずは音声を 6 秒ごとのバッファに切り分ける(図 5 中 1)。次に 6 秒分のバッファに音声をためる(図 5 中 2)、6 秒分の音声を FFT し(図 5 中 3)、分解能 1Hz の 50~1600Hz、つまり 1551 個の周波数特性を抽出する(図 5 中 4)。音声をすべて FFT するまでこれらの一連の処理を繰り返す。このようにして、計測時刻と 6 秒分の周波数特性の組み合わせを CSV へ出力していく。全てのノードから得られた音声を処理し、最終的には各ノード毎の 6 秒ごとの周波数特性が得られる。

次に、得られた各ノードの 6 秒ごとの周波数特性を比較

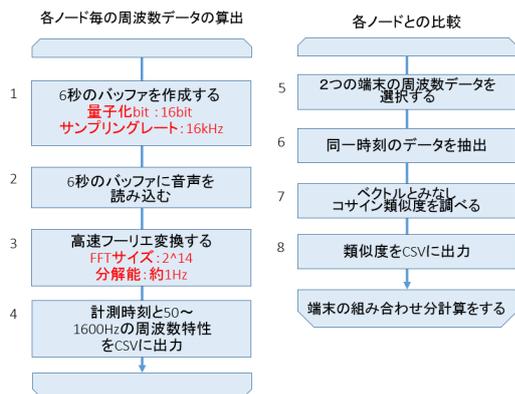


図5 会議場検出アルゴリズム：類似度の算出

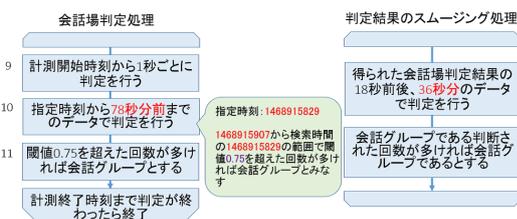


図6 会議場検出アルゴリズム：会議場の判定



図7 各ノードの周波数特性の比較

していく(図7)。比較には同一時刻の6秒分の周波数特性50~1600Hzを用いる。また、50~1600Hzの周波数特性を1551次元のベクトルとみなしてコサイン類似度を求める。比較した端末のペアと計測時刻、類似度をCSVに記録する。以上の処理を全てのノードの組み合わせ分で行う。

次に、会議場の判定方法について説明する(図6中左)。会議場の判定方法は1秒ごとに時間窓をスライドさせ、閾値0.75を超えた回数が多ければ会議グループ、少なければ会議グループではないという多数決判定を行う。その判定を計測開始時刻から終了時刻までの1秒ごとに判定を行う。また閾値に関してはNearyを踏襲したものである。一方で時間窓を78秒としたのは判定結果の再現率や適合率を加味したうえで適切だと判断したためである。しかしこれらの値はあくまでデータごとのチューニングであり、一般化されたものではない。

最後に、得られた結果のスミージング判定について説明する(図6中右)。得られた1秒ごとの判定結果からさらにスミージング処理を行う。具体的には判定時刻から前後18秒を抽出し、計36秒の判定結果の内、各ペアが会議グループであるという判定が多ければ会議グループ、そうでな

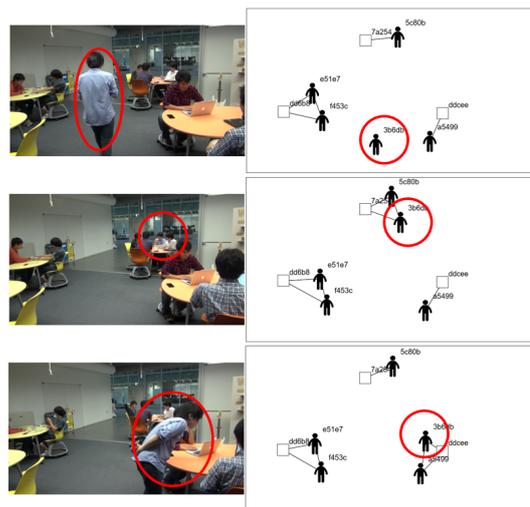


図8 会議場の検出例：人の移動の変化

れば会議グループではないという多数決判定を再度行う。また、この時間窓に関してもデータごとに得られたものであり、一般化されたものではない。

3.1.2 音環境比較による会議場検出の実例

本節では、以上で述べた音環境比較手法を用いた会議場検出の実例を示す。

図8は会議場間を移動する様子を示したものである。図8中の左側は実際の映像、右側は音環境比較による手法で得られた判定結果を元に視覚化したノードグラフである。人型のアイコンが人ノード、四角が環境ノードを示している。図8の例では会議グループが3つ存在しており、右側、左側、奥にそれぞれ2、3人が会話している。また、環境ノードがそれぞれの机の上に設置されており、人ノードは椅子に座っている人の一部が装着している。そして、赤丸で囲われた男性が会議場間を移動している参加者である。この赤丸で囲まれた男性は3b6dbというIDの端末を持っており、このIDと対応している。上段右の判定結果によるとこの段階ではどの会議グループにも属していない。次に、中段では奥の会議グループに参加しており、判定結果も環境ノード(7a254)とその付近で会話をしている参加者(5c80b)とグルーピングされている。下段では右手前の会議グループへ移動し、会話をしており、判定結果も右手前テーブルの環境ノード(ddcee)とグルーピングされていることがわかる。以上からこのような移動の連続的な流れを音環境比較によって検出することができる。また、環境側に設置されたノードとグルーピングされることでどこで会話をしているかといった場所の情報も得ることができる。

次に、図9は会議場が広がった例である。図9の例では、奥と手前に会議グループが計2つ存在している。また、それぞれの机の上には環境ノードが設置されており、人ノードは装着している人としていない人が混じっている。上段は2つの会議グループが存在しており、判定結果も2つのグループが存在していることを示している。次に、下段は

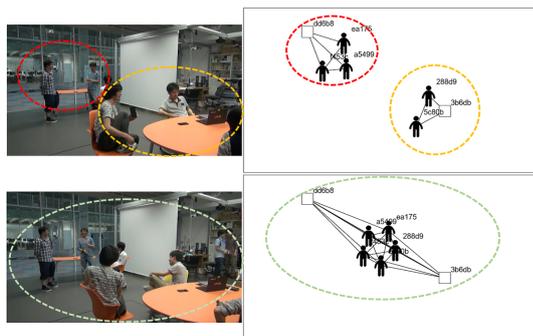


図 9 会話場の検出例：物理的空間の広がり

2つの会話グループ間で会話をはじめている様子で、判定結果も環境ノードを含めた大きな会話場であるとみなしている。以上のことから、会話場が統合するような状況も検出することができる。また、物理的に離れている環境ノード同士がグルーピングされることで大まかな会話場のサイズを認識できると考える。

以上二つの例より、音環境比較による会話場検出で得られる情報としては、誰と誰がどこで会話をしているという会話場の情報と、そのダイナミクス、規模感を得ることができる。本研究では、この会話場情報を基にインデクスを振ることで体験データの閲覧支援への応用を試みる。

3.2 実環境における会話場検出

評価に用いるデータセットも兼ねて、実環境での会話場計測を行った。他大学の学生とのポスター発表の場と、学会のポスターセッションで計測を行った。シチュエーションとしてポスターセッションを選択した理由としては、会話場のダイナミクスが得やすいという点と、記録の対象として優れていると考えたためである。前者については、平行にポスター発表が行われているため人が行き来しやすい場だからである。また、後者については、ポスターセッションでの議論は記録や振り返りを行う対象になりうると考えたからである。

計測方法についてはどちらも同様に、GoProを胸部に装着し一人称映像の記録を行った。また、環境ノードは机の上やポールに装着し、発表者の近くに設置した。また、予備検討的に得られた会話場情報を基に1人称映像を確認した結果、お互いの映像中に会話しているお互いの姿が確認できた(図10)。

3.3 会話場情報を利用した閲覧支援システム

提案する閲覧支援システムは図11に示すとおりである。図中左側に定点カメラ。右側には得られた会話場情報を基に可視化したノードグラフを表示している。また、各参加者が計測した一人称視点映像はこれらのノードグラフ上の人型アイコンをダブルクリックすることで表示できる。システムの上部にあるシークバーを操作し閲覧を行う。ま

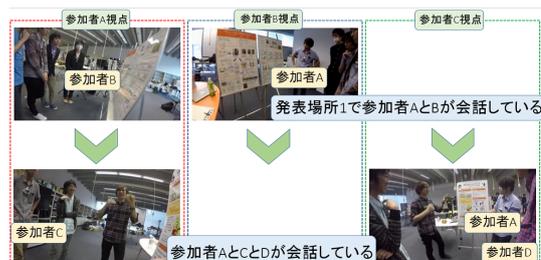


図 10 得られた会話場情報と一人称映像のシーンの比較

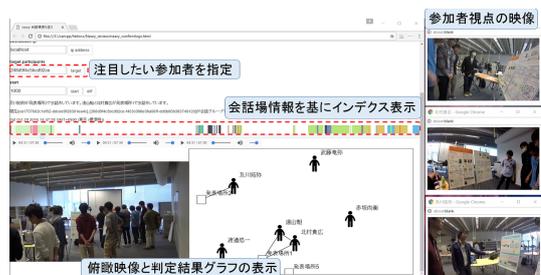


図 11 会話場情報を利用した閲覧支援システム

た、シークバー上部に会話場情報のインデクスが表示されており、会話グループと場所の情報ごとに色分けをして表示した。また、参加者の名前を指定することで、注目したい参加者が含まれるインデクス情報を表示することもできるようにした。このように参加者ごとの情報を表示できるようにした試みとしては、本研究が振り返りに用いる映像が自身だけではなく、他の参加者の視点も振り返りに有用であると考えたためである。例えば自分が聴講できなかったポスター発表を他人と視点を通して閲覧したり、複数人が集まって聴講している発表の様子をきつままで閲覧することが可能になると考えられる。このような自身が経験していない他者の経験も閲覧が容易になり、その他者の視点から追体験が可能になると考えられる。

4. 提案システムを用いた閲覧支援の評価と考察

4.1 閲覧支援システムの評価実験概要

「誰と、どこで会話をしていた」といった会話場情報を一人称映像の閲覧支援に活用することが、体験映像の振り返りに有用であるかどうかの検証を行った。方法としては、与えられた特定のイベントに到達するまでの時間を、閲覧支援がある場合とない場合とで比較を行った。また、自身の映像の振り返りだけではなく、他者の映像の追体験においても有用であるかどうかの確認を行った。

4.1.1 データセットと評価方法

ポスターセッションでの体験映像と会話場情報を実験の評価に用いた。他大学とのポスターセッションでの計測データセットをセッションA、学会のポスターセッションでの計測データセットをセッションBとする。データはいずれも1時間ほどの長さとした。また、被験者はいずれか

のポスターセッションの参加者3人を選択した(以下、被験者 a、b、c とする)。また、セッション A の参加者は被験者 a のみで、セッション B の参加者は被験者 b、c である。

与えるタスクは各セッション中において検出された会話場の情報を基に作成した。「被験者 a が特定の参加者とともにもポスター発表を見ているシーンはどこか」などといった情報を与えた。各セッションごとに4つのタスクを設定した。セッション A の4つのタスクにはそのセッションに参加している a の体験を基に作成した。またセッション B のタスクは、セッション B の参加者である被験者 b と c の体験を基に作成した。

実験に際して、本人の体験に基づくデータセットからなる4つのタスクと、本人の体験が含まれないデータセットからなる4つのタスクの計8つのタスクを与えた。またそれぞれの4つのタスクの内2つのタスクはシステムによる支援を行い、残る2つは通常の映像再生ソフトウェアを使ってタスクの発見を行ってもらった。

本システムを使うことで、本人の体験が記録された映像、本人の体験ではない映像、そしてその両方において特定のイベントを振り返るまでの時間が短縮されるかを計測した。

4.2 閲覧支援システムの結果

タスク A の遂行にかかった時間を表 1、タスク B にかかった時間(秒)を表 2 に示す。A-1~A-4 は被験者 a が参加したポスターセッションであるセッション A のデータセットから作成されたタスクを示している。また B-1~B-4 は被験者 b、c が参加したセッション B から作成したタスクを示している。赤字は被験者本人の一人称映像が含まれており、本人の体験に基づいたタスクの発見にかかった時間を示している。また、緑色で塗りつぶされた箇所は、本システムによる閲覧支援を受けている状態でのタスク遂行時間を示している。また、表 3 は、自分の体験に基づく映像の振り返りと他人の体験映像の振り返りにかかった時間の平均を示している。この表から、会話場情報による閲覧支援を行うことで支援がない場合よりも時間が短縮された。他人の体験を閲覧する際には大きな差は見られなかったが、自分の体験に基づく映像を閲覧する際には会話場情報に基づいた支援によって大きく短縮されるのが確認できた。また、t 検定を行ったところ、自分の体験の振り返りにおいて支援がある場合と支援がない場合において 0.042 であったため ($p < 0.05$) より有意差であることを確認できた。一方で、他人の体験を閲覧する場合においては有意差は見られなかった。

4.3 閲覧支援実験の考察

本研究では会話場情報を基にインデキシングをすることが映像の閲覧支援になることを示すため、閲覧支援実験による評価を行った。人が自由に行き来するオープンスペー

	A-1	A-2	A-3	A-4
被験者A	68	83	18	31
被験者B	261	128	13	90
被験者C	58	352	59	71

表 1 タスク A のイベント発見にかかった時間(秒)。赤字は被験者本人の体験映像に基づいたタスクを指す。また、緑色で塗りつぶされている箇所は本システムを利用したタスクを指す。

	B-1	B-2	B-3	B-4
被験者A	147	42	80	67
被験者B	43	179	162	216
被験者C	52	276	30	54

表 2 タスク B のイベント発見にかかった時間(秒)。赤字は被験者本人の体験映像に基づいたタスクを指す。また、緑色で塗りつぶされている箇所は本システムを利用したタスクを指す。

タスク		平均秒数
自分	支援なし	142.83
自分	支援あり	59.17
他人	支援なし	117
他人	支援あり	111
合計	支援なし	129.91
合計	支援あり	85.08

表 3 自分の体験に基づいた映像の振り返りと他人の体験の振り返りにかかった秒数の平均。

スで、1時間ほどのポスターセッションで計測したデータセットをあらかじめ二つ用意し、3人の被験者に一人称映像の振り返りと追体験を行ってもらった。結果としては、自分の体験が記録された映像においては閲覧支援になりうることを示された。よって、自分の記憶を頼りに映像を閲覧するよりも「誰と、どこで会話をしていた」という会話場情報による支援によって特定のイベントの発見にかかる時間を短縮することができた。一方で、自身の体験ではない他者の体験の映像から、他者のイベントを発見する時間短縮を行えることは示すことができなかった。また、被験者からは、「インデキシングされた情報がその日の自身の記憶に近いものであった。」という意見や、「会話内容について詳細は覚えていないが、会話をしていた相手や一緒に発表を聞いていた人は記憶に残っている。」といった意見が得られた。よって会話相手と場所の情報は十分に検索クエリになりうると思われる。

5. まとめ

本研究では、音環境の近さから「誰と、どこで会話をしていた」という会話場情報を検出し、映像にインデキシングすることで閲覧支援への活用を提案した。自身の体験に基づくタスクと、他人のタスクの両方を用意し、それぞれにおいて本システムによる閲覧支援がある場合とない場合とでタスク発見にかかる時間の比較を行った。結果としては閲覧支援がある場合の平均秒数はない場合と比べて短縮

された。また両側 t 検定を行ったところ、自分の体験映像の振り返りにおいて有意差が得られた。よって、自分の体験を振り返る際には「誰と、どこで会話をしていた」という情報は映像中のイベント発見に役立つことが示された。

6. 今後の展望

本研究では、音環境比較による会話場検出を体験データのインデキシングに利用した閲覧システムを提案した。音環境比較による検出手法が従来手法よりも有用であることを示すための応用例の一つとして閲覧支援の提案を行った。また、インデックスの指標として、会話場に注目した。会話場を得る方法として近い音にさらされているペアを同じコンテキストを共有しているグループとみなし、音環境比較による手法を用いた。自分の体験に基づく映像の閲覧においては、記憶に頼るよりも会話場情報を基に閲覧をする方が時間が短縮されることが示された。今後の展望としては、音環境比較による手法で得られる会話場情報は、傾聴しているような状況やその場に居合わせているような状況も含まれている。ただし、会話というインタラクションだけではなく、その場に居合わせたりお互いのテリトリに共存しているような状況からも気づきを促せると考えている。いわば潜在的なグループ内で、その場で得られた知識や体験の共有につながると考えられる。自身の体験を振り返るライフログとしての活用だけではなく、他者との知識、体験共有のためのコミュニティウェアやグループウェアなどの活用も行えらる。そこで、会話というインタラクションを行っている状況と、お互いがその場に共存しているような状況とをより詳細に分別することで、潜在的なグループ間での知識体験共有への活用も目指していきたい。具体的には、音環境比較による手法だけではなく体験データから得られる顔の距離や数といった情報をあわせることで、より具体的なコンテキスト情報の取得をし、インデキシングすることを試みたいと考えている。また、他者の体験を振り返るための閲覧支援が行えるようなデザイン設計にも取り組んでいきたい。

参考文献

- [1] T. Nakakura, Y. Sumi, and T. Nishida, "Neary: conversation field detection based on similarity of auditory situation," Special Section on emerging Technologies of Ubiquitous Computing Systems, No.6 June 2011, pp.1164-1172.
- [2] E. T. Hall. The Hidden Dimension. Doubleday & Company, Inc., 1966.
- [3] R. Borovoy, F. Martin, S. Vemuri, M. Resnick, B. Silverman, and C. Hancock, "Meme tags and community mirrors: moving from conferences to collaboration," in Proceedings of the 1998 ACM conference on Computer supported cooperative work, 1998, pp. 159168.
- [4] T. Choudhury. Sensing and Modeling Human Networks. Doctoral thesis, Massachusetts Institute of Technology,

- September 2003.
- [5] 森脇 紀彦, 佐藤 信夫, 脇坂 義博, 辻 聡美, 大久保 教父, 矢野 和男. 組織活動可視化システム「ビジネス顕微鏡」. The institute of electronics 2007-HCS-44(39), 2007-09, 一般社団法人電子情報処理学会.
 - [6] T. M. T. Do and D. Gatica-Perez, "Contextual grouping: discovering real-life interaction types from longitudinal bluetooth data," in 2011 IEEE 12th International Conference on Mobile Data Management, 2011, vol. 1, pp. 256265.
 - [7] J. Rekimoto, T. Miyaki, and T. Ishizawa. LifeTag: WiFi-based Continuous Location Logging for Life Pattern Analysis. 3rd International Symposium on Location- and Context-Awareness (LOCA2007), pages pp.3549, 2007.
 - [8] Y. Sumi, J. Ito, and T. Nishida. PhotoChat: communication support system based on sharing photos and notes. CHI 2008 Extended Abstracts, pages pp.32373242, April 2008.
 - [9] B. Thiel, K. Kloch, and P. Lukowicz, "Sound-based proximity detection with mobile phones," 2012, pp. 14.
 - [10] P. G. Kannan, S. P. Venkatagiri, M. C. Chan, A. L. Ananda, and L.-S. Peh, "Low cost crowd counting using audio tones," in Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems, 2012, pp. 155168.
 - [11] M. Azizyan, I. Constandache, and R. R. Choudhury, "SurroundSense: Localizing Mobile Phones Using Ambient Light, Sound, Color, and Motion."
 - [12] 小関 悠, 角 康之, 西田 豊明, 間瀬 健二. ぱらぱらアニメによる体験データの要約・編集支援システム. コンピュータソフトウェア, Vol.24, No.3, pp.41-50. 2007.
 - [13] 坂本 竜基, 角 康之, 中尾 恵子, 間瀬 健二, 國藤 進. コミックダイアリ: 漫画表現を利用した経験や興味の伝達支援. 情報処理学会論文誌, Vol.43, No.12, pp.3582-3595, 12, 2002.
 - [14] K. Higuch, R. Yonetani, and Y. Sato, "EgoScanning: Quickly Scanning First-Person Videos with Egocentric Elastic Timelines," 2016.