

発話を空間コンテキストに 埋め込むことによる体験共有システム

藤倉 稜^{1,a)} 角 康之^{1,b)}

概要：本稿では、発話を空間コンテキストに埋め込むことによる体験共有システムの提案をする。日常生活の中で、人は独り言や他人との会話など発話を常に行っている。その中で記録される発話もあれば、何事もなくその場だけで終わる発話も存在する。そこで本研究では、人間が移動時に行う発話に着目し、それを空間コンテキストに埋め込んで記録することでその場所にきた他人に感動や気づき、知識などを共有するシステムを提案する。これにより、今まで共有されることのなかった体験の共有と新たな知識の流通を目指す。

Experience Sharing System by Embedding Conversations into Spatial Context

RYO FUJIKURA^{1,a)} YASUYUUKI SUMI^{1,b)}

Abstract: This paper presents a system that allows sharing of experiences by embedding conversations into spatial context. Most people usually engage in conversations on a daily basis. In those conversations, some are recorded while others remain private with conversation partners. In this paper, we record conversations between people as they move in a space and map them with their spatial context. Ultimately, it is our goal to present new and unshared knowledge, and to facilitate a fuller experience between conversation participants in a shared space.

1. はじめに

人間が他人と何かの体験を共有する時、一般的に体験時の感動や知識を言語化して伝え合う傾向がある。この場合、実際の感動が十分に伝わらなかったり、知識が身に付かないという限界がある。つまり、言語情報だけで実際の体験と同等の情報を伝えるのは難しい。近年スマートフォンなどの普段から携帯する撮影機器が急速に普及し、至る場所で様々な状況を対象とした写真が撮影されている。そしてそれらは、Twitter や LINE などの SNS で他人と共有されることがしばしばある。体験の共有に言語情報だけでなく写真を用いることで、格段に実際の感動に近いものが得られると考えられる。しかし当然、共有した写真を見た

ときに受け手が得られる気づきや感動は、受け手がそこにいたわけではないので実際に写真を撮影した人の感動より薄いと考えられる。一方、体験というのはイベントなどの体験と何気ない行動から発生する日常的な体験に分けられると筆者は考える。前者を他人に共有する場合、体験の様子を撮影したりメモを取ったり情報を記録しようとする。しかし、後者の体験はふいに訪れる為、他人に共有されない場合がしばしばある。例えば、「ここのパン屋、今日から1ヶ月間オープンセールでパンが安いのか〜」などと誰かが聞いたときに面白いだとか、有用だと思われる情報があったとしても、体験者がこれを何らかの形で残そうという考えに及ばない場合は共有されることがないままその発言は埋もれてしまう。折角の自分の貴重な時間で体験した事を他人に共有しないのは、勿体ないことであると筆者は考える。また、今の例は一人での体験の例だが他人との会話の中でそういった知識が深まることもある。このような

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

^{a)} r-fujikura@sumilab.org

^{b)} sumi@acm.org



図 1 会話の移り変わり

知識は現在だと SNS などに投稿され共有されることがよくある。しかし、SNS での共有ではその場所で起きたことがそれを有用だと思う人や興味を持つ人に届くかどうか定かではない。そこで、本研究ではそのような瞬間的な体験の中で生まれる知識を流通させるシステムを提案する。また、図 1 のように現在いる地点や向いている方向によって人は発話の内容を変えることから、空間コンテキストに発話を埋め込むことでシステムを実現する。ここでいう空間コンテキストとは、どこにいるかといった位置情報とは異なり発話する人がどこにいて、何に注目してどこを向いていたかといった空間的な状況のことである。それに発話を関連付けることで、その場で過去に行われた発話を空間に紐付けて記録できると考える。具体的な手法として、図 2 に示した概要のように、1) 独り言や会話を空間コンテキストに埋め込む、2) 埋め込まれた独り言や会話を近くを通る人に提示する、といった流れで知識の流通を目指す。図中では、空間コンテキストに埋め込まれた会話が通りかかった人に提示されている。また、提示された発話により新たに発生した発話も場所に紐付けられ、空間に新たな知識が蓄積している。これらの流れのループにより知識の流通を目指す。本論文では、発話を空間コンテキストに埋め込むことによる体験共有システムの提案と実装における予備検討の結果を報告する。

2. 関連研究

感動や知識の源泉である体験を通した協調や知識創造を増幅するメディア技術は体験メディアと呼ばれるが、角らはその一つとして複数の視点による映像・音声を協調的に記録するシステムを提案している [1]。このシステムでは、各体験者の一人称視点、会話や共同作業の相手による二人称視点、環境側から体験者同士のインタラクションをとらえる三人称視点の映像を同時記録する為に、各体験者の頭部に装着するカメラと、体験現場を捉える環境カメラを併用して、体験の記録が出来る。角らは、このシステムを用いて得られたデータをもとに自らが参加した体験シーンを後で時間をかけて追体験することを可能にし、また、二人

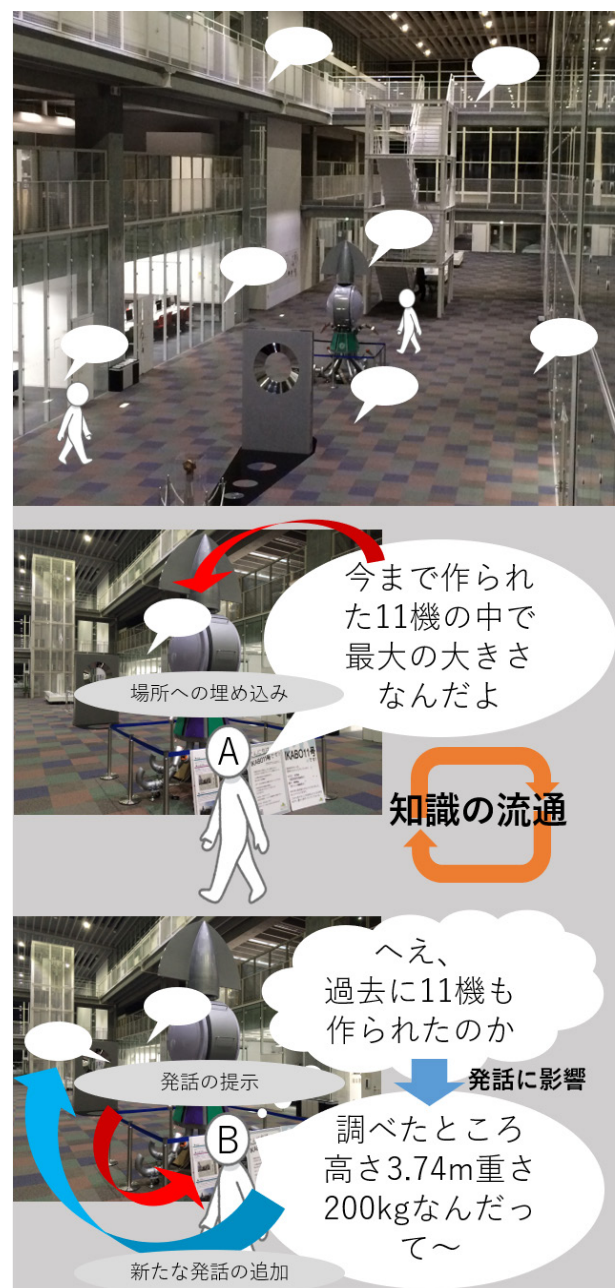


図 2 提案システムの概要

称、三人称の映像を見ることで、新たな視点への気づきを促すことが期待されたと考えた。それを検証するメタ認知実験では、体験記録システムが体験に基づいた知識の創造プロセスに大きな影響を与えることが明らかになった。本研究では、会話情報を空間コンテキストに埋め込むことで体験の記録を行う。これにより実空間とリンクしたすぐに利用可能な知識の流通が行われると考える。

岡村らは車内会話に注目し、GPS ロガーとビデオカメラを用いて位置情報が紐付いた車載映像の記録を行った。また、自動車の運転シミュレータ上で、記録した映像データを再生するシステムを構築し、場所に紐付いた過去の会話から新たな会話が想起されることを確認した [2]。本研究でも、場所に紐付いた過去の会話から新たな会話が想起されたり、知識の創造を促すことが期待されと考えられる。

西村らはユーザの位置情報に基づいた情報支援がリアルタイムでされることを重要視し、それを実現するシステムを提案した [3]。このシステムでは、無電源小型情報端末 (Compact Battery-less Information Terminal: CoBIT) を用いて適切な位置で適切な方向に端末を向けるだけでインタラクティブに音声情報を取得することを可能にする。本研究では、通信を行うことなく他人が空間テキストに埋め込んだ音声情報を提示するという違ったアプローチから情報支援を行う。

角らは、体験から得られる知識共有や体験の現場におけるさらなる体験創造を支援するため PhotoChat というシステムを開発した [4]。PhotoChat はカメラを内蔵した携帯型のパソコン上で動作することを想定したソフトウェアで、写真撮影が出来る。また、ペンインターフェースを用いて、写真の上書き込みを行うことができ、実時間で他のユーザと共有されるため互いの興味への気づきを促し、その上で自然に会話を行うことを可能にする。このシステムでは、体験時に湧き起った感情をすぐに他人と共有することができる為互いの新たな気づきを促すと考えられる。PhotoChat ではテキスト情報を用いた体験共有を支援しているが、本研究では、音情報に着目しそれらをリアルタイムで空間コンテキストに埋め込むことで場所に知識を紐付ける手法をとる。

松村らは人と街の間に埋め込まれた会話的知識を流通させるために、自動車に乗って移動する際の車内会話を記録・提示する手法を提案した [5]。松村らは移動中にあるエンターテインメント性に注目し、自動車を利用して移動する際に生まれる会話の中で深まる知識を会話的知識 (Conversational Knowledge) と呼んだ。この会話的知識、特に車内会話から生み出される知識を、場所に車内会話を紐付けるという手法で流通させるシステムを提案した。本研究では、車内という限られた状況ではなく人が移動する場面全てを対象として、空間コンテキストに発話を紐付ける事でこの会話的知識の流通を目指す。

Sawhney らは歩行中に仮想空間上の複数の情報とインタラクションするために音声情報の提示をするシステムを開発した [6]。本研究での試みは、仮想空間ではなく実空間での行動誘発・誘導に音場展示を利用する、という位置づけである。

Mueller らは特定の場所に立ち止まるとその場で過去に行われた会話が聞こえてきたり、さらに詳しくポスターや展示物などの特定の場所に注目すると、そういったコンテキストを共有した会話に絞り込まれる Ambient Sound Shower というシステムを開発した [7]。本研究はこの研究と目的が同一であるが、システムの実現をするのに別の手法を用いて再実装しているという位置づけである。

3. 研究概要

本研究では、発話を空間コンテキストに埋め込むことによる体験共有システムの提案をする。図 2 にて提案システムの概要を示したが、本研究の目的は体験時に湧き起こる感動や気づき、また会話から生まれる知識を流通させることである。それを実現する手法として、独り言や会話といった発話を空間コンテキストに埋め込み、埋め込まれた発話を近くを通る人に提示するシステムを開発する。以下では、システムを実現するための要件について述べる。

3.1 発話の埋め込み

独り言や会話といった発話を空間コンテキストに埋め込むには、発話の音声情報を記録すると同時に位置情報や頭部方向、またどこに埋め込むのかといった詳細な位置を記録する必要がある。本システムではこの条件を満たすデバイスとして Microsoft 社が提供する HoloLens を使用する。HoloLens とは、目の前の現実世界と、3D ホログラムの世界を融合させた Mixed Reality (複合現実) を実現させる HMD 方式の拡張現実ウェアラブルコンピュータである。HoloLens には深度カメラによる空間認識機能が搭載されており、様々なオブジェクトを現実空間に合わせた位置で設置することができる。またその位置情報を記憶することが可能である。さらにマイク機能も搭載しているため、人の発話を録音しそれと位置情報を紐付けて空間に残すことが可能である。この機能を利用することで体験における独り言や会話などの発話と空間コンテキストを紐付けるシステムを実現できる。さらに 3.3 節でも述べるが、HoloLens にはスピーカも搭載されており、他人が追体験する場合 HoloLens をそのまま用いることで一つのデバイスで記録と体験を担うことができる。

3.2 発話の切り出しと分類

発話を空間コンテキストに埋め込んだとして、その発話を聞く人は記録時と同じ経路でかつ同じ速度で移動しなければ連続した意味の伝わる発話を聞くことはできない。

よって、ある程度内容が固まっている発話をひとまとまりとして切り出すことが必要になる。また、全ての発話が有用だと限らないため発話の分類も必要である。発話の分類手法の先行研究として岡村らが提案した指差しを手掛かりにしたシステムがある [2]。これは、シミュレータ上での位置推定であったため利用できたが、本研究では現実世界の場所に紐付けたうえで分類するため活用できない。しかし、HoloLens には片手の指先 1 本をクイックと手前に引くような操作を AirTap と呼びそれを検知する機能があらかじめ備わっている。これをトリガーとすることで有用な発話だけを記録できると考える。また、このようにして切り出された発話は埋め込まれた場所のうち適切な点に関連付けられる必要がある。

3.3 発話情報の提示

前節の手法で切り出された発話を別の、つまりユーザーに提示する方法はいくつか考えられる。ここでは、体験を記録するために用いた HoloLens を使用する方法について述べる。まず、記録された発話は発話が行われた場所に紐付けられて実際の体験時と同じ時間幅で再生されるように設定する。その状態でユーザーに HoloLens を着用してもらう。体験時には移動を行うと想定しているため発話も時間とともに位置が変わっていく必要がある。これにより、体験時の発話にユーザーが移動を促されるような形で追体験が可能となる。また、HoloLens ではステレオサウンドを実現できるため、どの方向から声が聞こえるといったことも知覚できると考えられる。さらに、発話の同時提示についても考える必要がある。一人の体験時の発話のみを提示するのか、複数人の体験時の発話を提示するのかということである。本研究の目的である、体験時に湧き起こる感動や気づき、また会話から生まれる知識を流通させることを達成するには、情報がいくつかある方が適していると考えられる。複数の発話が混在した場合、必要な情報が聞き取れない可能性があるという懸念は、複数発話の聞き分けにより払拭されることが考えられる。なぜなら人間は音声について選択的聴取が可能なので、ユーザーは気になる発話を選択的に聴取できるからである。発話をいくつ同時提示するかは別に検討する必要がある。

4. 予備検討

4.1 HoloLens の機能確認

3 章で触れたが、本研究では HoloLens を用いることで体験時の発話を記録し提示する。本節ではシステムにて実装したい機能を HoloLens にて実現可能か検証したのでそれを報告する。まず、位置情報の記憶についてだが、あらかじめ用意されているホログラムを HoloLens 内で指定した位置に設置した。ここでいうホログラムとは立体的に空間上に設置された、HoloLens を通してでしか視認できな

い映像のことを指す。次に HoloLens の電源を切り別の位置に移動したうえでもう一度起動し、ホログラムを設置した場所に行ったところ cm 単位の位置のずれはあったが、同じホログラムが残っていることが確認できた。このことから、発話を空間コンテキストに埋め込んでおくという機能は実装できると考えられる。

次に 3.3 節で触れたステレオサウンドについて、検証を行った。HoloLens での開発には unity を用いることが一般的である。unity では音声情報に 3 次元的な位置情報を付与することができる。つまり unity 上でのスピーカーと音声情報の相対位置によって聞こえてくる音の方向や音量を変化させることができる。実際に HoloLens を着用して検証したところ方向、音量から音源の位置を特定できた。よって、HoloLens でステレオサウンドが実現できると考えられる。

4.2 複数発話の聞き分けと場所特定の可否の検討

3.3 節で触れたように人間は音声について選択的聴取が可能であるということが知られている。よくある例だと、人混みの中で騒がしい状況にいても自分の名前を呼ばれると反応できるというものがわかりやすい。しかし、それが HoloLens を用いた発話の聴取でも同じことが可能かは未検証であったため、それについて検証した。まず、過去に録音した全く異なる発話データを 4 つ用意した。それらを図 3 のように unity 上で現実世界で 5m 間隔になるよう配置し、HoloLens を装着して音声の選択的聴取が可能か確認した。結果、移動を伴いながら 4 つの音源全てを聞き分けることができた。これにより、HoloLens を用いて発話の同時提示をした場合、ユーザー選択的聴取が可能であることが示された。

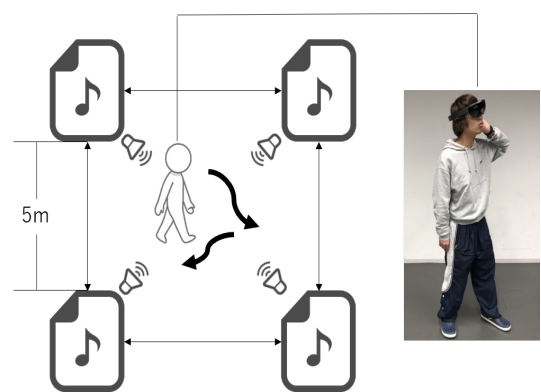


図 3 予備検討の様子

5. おわりに

本稿では、発話を空間コンテキストに埋め込むことによる体験共有システムについて提案した。発話を記録、再生するデバイスとして、発話を空間コンテキストに従って記

録可能である HoloLens を使用した。また、HoloLens を使用してシステムを開発するうえで必要となる機能を検証するために、予備検討を行った。本研究により、日常生活の中に埋もれていた知識が流通し、新たな感動や気づきが他人に共有されることが期待される。現状の課題として、記録した発話と空間コンテキストを紐付けるシステムの開発が挙げられる。それを解決する為に、HoloLens の自己位置推定に使用されているアンカーを利用する予定である。アンカーとはホログラムが確実に所定の位置に正しく留まるために、他の空間アンカーや静止座標系に対して必要に応じて調整される座標系を有している。またアンカーをセーブしたりロードしたりする事で、アプリはセッションをまたいで現実世界の同じ座標を利用する事ができる。

今後は、システムの開発に着手し、記録した発話を空間コンテキストに埋め込みそれを体験できるシステムを作り上げ、我々の体験や知識にどう影響を及ぼすか確かめていく。

参考文献

- [1] 角康之, 伊藤禎宣, 松口哲也, シドニーフェルス, 間瀬健二. 協調的なインタラクションの記録と解釈. 情報処理学会論文誌, Vol. 44, No. 11, pp. 2628–2637, nov 2003.
- [2] 岡村剛, 久保田秀和, 角康之, 西田豊明, 塚原裕史, 岩崎弘利. 車内会話の量子化と再利用. 情報処理学会論文誌, Vol. 48, No. 12, pp. 3893–3906, dec 2007.
- [3] 西村拓一, 中村嘉志, 伊藤日出男, 山本吉伸, 中島秀之. 無電源小型情報端末を用いたインタラクティブ情報支援. 人工知能学会全国大会論文集, Vol. JSAI03, pp. 106–106, 2003.
- [4] 角康之, 伊藤惇, 西田豊明. Photochat : 写真と書き込みの共有によるコミュニケーション支援システム. 情報処理学会論文誌, Vol. 49, No. 6, pp. 1993–2003, jun 2008.
- [5] 松村耕平, 角康之. 自動車内における会話と場所の関連性の分析 : タイムリーな情報の流通に向けて. 情報処理学会論文誌, Vol. 56, No. 4, pp. 1258–1268, 4 2015.
- [6] Nitin Sawhney and Chris Schmandt. Nomadic radio: Speech and audio interaction for contextual messaging in nomadic environments. *ACM Transactions on Computer-Human Interaction*, Vol. 7, No. 3, pp. 353–383, September 2000.
- [7] Christof Müller, Yasuyuki Sumi, Kenji Mase, and Megumu Tsuchikawa. Experience sharing by retrieving captured conversations using non-verbal features. In *The First ACM Workshop on Continuous Archival and Retrieval of Personal Experiences (CARPE 2004)*, pp. 93–98, 2004.