

人間の目と耳を集合的に共有するための 分散型デジタルプラットフォーム

木村 梨沙¹ 中島 達夫¹

概要：本論文では、人間の目と耳を集合的に共有する CollectiveEyes と CollectiveEars の2つの分散型デジタルプラットフォームを紹介する。最近、物理的な持ち物の共有を促進する共有経済分散プラットフォームが普及してきている。しかし、他の種類の物理的資源、特に人体の一部の共有を考慮することは、将来のサービスを開発するための興味深い新規の方向性を提供する可能性がある。CollectiveEars と CollectiveEyes の概要を解説した後、それらのデジタルプラットフォームに関する開発の経験に関して説明する。

キーワード：共有経済、デジタルプラットフォーム

Distributed Digital Platforms for Sharing Human Seeing and Hearing Collectively

RISA KIMURA¹ TATSUO NAKAJIMA¹

Abstract: This paper presents CollectiveEars and CollectiveEyes that collectively sharing human eyes and ears. Recently, sharing economy distributed platforms to encourage sharing personal physical belongings become popular. Considering other types of physical resources, in particular, parts of human physical bodies offers an interesting opportunity to develop novel services. We present overviews of these digital platforms and show some experiences with them.

Keywords: Sharing Economy, Digital Platforms

1. はじめに

分散型デジタルプラットフォームの最近の進歩により、サービスプロバイダは多様なサービスを提供できるようになってきた [5, 6, 8, 10, 11]。また、バーチャルリアリティ技術のような新興技術は、仮想空間内の実際の施設、サービス、店舗などを表すための新しい開発を促進してきた [10, 13]。これらは、社会インフラとしての仮想空間を含み、私たちの日常の活動を拡張する場として考えることが可能となってきた。

様々なリソースへのアクセスをデジタルテクノロジーと共有することでピアツーピアベースで共有することを可能とする Uber や Airbnb などの共有経済分散デジタルプラットフォームは、最近大きな注目を集めている [1]。従来のシェアリングエコノミーでは、一般に、車などの人の持ち物、ま

たは人の空き時間などの物理的なリソースが含まれている。そこで、他のタイプの物理的リソース、特に人体を共有することの実現可能性を調査することは、共有経済デジタルプラットフォームの現在の範囲を拡大する興味深い方向性を再考するための有望な機会を提供する可能性がある。

この論文では、人間の目と耳を集合的に共有する CollectiveEars と CollectiveEyes の2つの分散型デジタルプラットフォームを紹介する。この研究の貢献は、人間の目と耳を共有するためのデジタルプラットフォームの実現可能性を明らかにすることであり、基本的なアイデアを実証するための実用的なプロトタイプを作成し、それらの経験について述べる。

2. 関連研究

Cities and Memory [14]は、世界中のさまざまな都市から集

¹ 早稲田大学 基幹理工学部 情報理工学科
Waseda University, 3-4-1 Okubo Shinjuku, Tokyo 169-8555, Japan
{r.kimura, tatsuo}@dcl.cs.waseda.ac.jp

められたサウンドレコーディングに貢献するフィールドレコーダ、サウンドアーティスト、ミュージシャン、サウンド愛好家のための世界規模の共同プロジェクトである。現在、500 人を超える人々が寄稿しており、世界 80 以上からの 2,000 を超えるオーディオサウンドを Web 上に録音している。対照的に、CollectiveEars は、同時に使用する人々から収集された世界中の複数の音を提示し、ユーザは個人のリスニングスペースで音をナビゲートして、世界中の様々な音を聞くことを可能とする。

Jack-In は、人間の能力と人間の存在を増強するためのコンセプトである[3]。このコンセプトにより、人間と他のアーティファクトの間、または人間の間での没入型接続が可能になる。私たちのアプローチで使用されている他の人の視力をハイジャックする手法は、Jack-In の概念に似ている。KinecDrone は、空を飛ぶという体の感覚を高める[2]。ドローンによってキャプチャされたビデオは、ユーザのスマートグラスに送信される。ユーザは空を飛んでいるように振る舞いながら、部屋にいる間にドローンを飛ばして撮影したシーンを見ることが可能となる。したがって、ユーザは自分が本当に空を飛んでいるように感じる事が可能となる。対照的に、CollectiveEyes を使用すると、ユーザは個々の人間/機械の目ではなく、人間の目を集合的に共有することが可能となる。

3. CollectiveEyes による集合的な人間の目の共有

CollectiveEyes は、人間の集合的な視覚的視点をキャプチャするために使用されるデジタルプラットフォームである。CollectiveEyes では、各人がカメラとマイクを含むウェアラブルデバイス（通常はウェアラブルグラス）を装備していると想定している。CollectiveEyes の現在のバージョンは、ヘッドマウントディスプレイ（HMD）を使用し、HMD の前にカメラとマイクを配置している。HMD は、カメラによってキャプチャされたビューを投影する。ただし、CollectiveEyes の将来のバージョンでは、軽量のスマートグラスを使用することで、日常の使いやすさを向上することが可能となると考えている。また、視点の表示に関しては Unity を使用している。

3.1 テーマチャンネルによる視点の選択

ユーザはテーマチャンネルを指定して、データベースに保存されているキャプチャされた視点から、TV チャンネルセレクトのように、ユーザの周囲に表示される複数の視点を選択することを可能とする。



図 1. ユーザが見たい視点の検索

Figure 1. Searching a View that a User Likes to See

各人の視点は、HMD に取り付けられたカメラを介してキャプチャされるが、その人は、CollectiveEyes のハッシュタグとして自分の視点に表示できる適切なテーマタグを定期的に送信する。現在のプロトタイププラットフォームでは、視点がキャプチャされた人がテーマタグをプラットフォームに手動で登録する。HMD を着用している人から抽出された視点は、テーマタグによってクラスタ化され、一部の視点が仮想空間にアイコンとして表示される。アイコンは、それぞれのアイコンクラウドとしてテーマタグにクラスター化され、図 1 に示すように表示されます。ユーザは、HMD を介してアイコンクラウドの 1 つを選択する。

3.2 視線ジェスチャによる視点のナビゲーション

CollectiveEyes の最も重要な特徴の 1 つは、すべてのコントロールに視線ベースのジェスチャを使用することである。CollectiveEyes では、次の 5 つの基本的なコマンドを定義した。最初のコマンドは選択コマンドである。このコマンドは、ユーザの視線を上から下に移動して対象者を選択するために使用される。2 番目のコマンドは非選択コマンドである。このコマンドは、ユーザが視線を下から上に移動して前のビューに戻るために使用する。3 番目のコマンドはハイジャックコマンドである。このコマンドは、ユーザが選択した視界内で人物を見て、その人物の周りでユーザの視線を上から下に動かすことにより、ユーザがハイジャックしたい別の人物を選択するために使用される。4 番目のコマンドは変更コマンドである。このコマンドは、ユーザが現在の視点で視線を上から下に移動することにより、現在のビューをユーザの近くにあるランダムに選択された別の人のビューに変更するために使用される。5 番目の最後のコマンドは入れ替えコマンドである。このコマンドは、ユーザの視線を視点の下から上に移動して、ユーザが置き換えたいビューを削除するために使用される。

3.3 複数の視点の閲覧

複数のビューを表示する場合、図2に示すように、複数の視点は仮想空間に表示される。現状のCollectiveEyesでは、4つの視点が自動的に選択され、仮想空間に表示される。視点の1つがユーザの興味を引くものでない場合、削除されたビューの代わりに別のビューが表示される。さらに、変更コマンドを用いることにより、すべての視点を同時に置き換えることを可能とする。



図2. 複数の視点の閲覧

Figure 2. Watching Multiple Eye Views

3.4 CollectiveEyesのソフトウェア構造

表1は、CollectiveEyes分散プラットフォームの各ノードのソフトウェア構造を示している。現在のプラットフォームでは、HMDまたはパブリックディスプレイを使用して他の人の視界を見ることができ、Fove0iまたはTobiiEye Tracker4Ciiを使用して人の視線を追跡することができる。現在、標準のWebカメラとマイクをHMDまたはパブリックディスプレイに接続している。現在のプロトタイプは、表1に示す4つのモジュールから構成されている。

表1. CollectiveEyesのソフトウェア構造
Table 1. Software Structure for CollectiveEyes

ビュー生成モジュール	このモジュールは、ユーザがUnityを使用して視聴するビューを生成する。ビューは、次の3つのモードを切り替えることができる。最初のモードはマップを表示し、2番目のモードは複数のアイビューを表示し、3番目のモードは選択された視線ビューを表示します。ビューは、視線ベースのジェスチャによってモー
------------	---

	ドを切り替えることができる。
ビューナビゲーションモジュール	このモジュールは、Unityを使用してカメラでキャプチャされたシーンに、移動または点滅するオブジェクトを重ね合わせる。また、3Dサウンドを使用して、人の頭の向きを指示することも可能とする。
視線ジェスチャ管理モジュール	このモジュールは、視線追跡デバイスを介して人の視線を抽出する。目の動きは、それぞれのジェスチャコマンドに変換される。
キーワードと視点ビューマッピングデータベースモジュール	このモジュールは、指定された地域にいる人々を見つけるために使用される分散ストレージである。

4. CollectiveEarsによる集合的な耳の共有

多様な人々の耳を共有するというコンセプトの実現可能性を実証するためにCollectiveEarsを開発した。Unityを採用して、ユーザの3Dリスニングスペースで複数のサウンドを提示する。CollectiveEarsのエンドユーザは、ユーザの頭部のジェスチャを検出するための加速度センサを備えたヘッドホンを装着する。

4.1 複数サウンドの提示

CollectiveEarsは、複数のサウンドを表示するための空間表示方法を提供する。空間提示法では、音源をユーザの周囲に配置し、複数の音を同時に提示する。ここでは、Unityの3Dサウンド機能を使用し、ユーザのリスニングポジションとして1つのリスナーオブジェクトを準備し、各音源として複数のオーディオソースオブジェクトを用意する。すべての音源がリスナーオブジェクトからほぼ同じ音量で聞こえるように、すべてのオーディオソースオブジェクトはリスナーオブジェクトから等距離に配置されている。図3は、プラットフォームのユーザーインターフェイスにおける8つの音源の配置を示している。この図では、リスニング位置にあるリスナーオブジェクトは正方形で表され、音源にあるオーディオソースオブジェクトは円で表されている。

i <https://www.getfove.com/>

ii <https://www.tobii.com/>

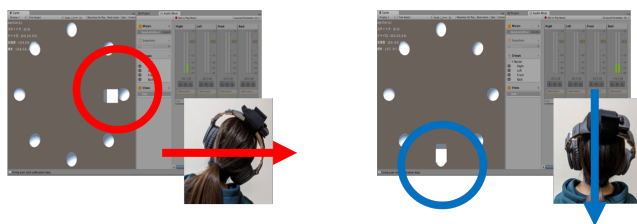


図 3. 頭部ジェスチャを用いた音ナビゲーション

Figure 3. Head Gesture based Sound Navigation

4.2 テーマチャンネルによる音の検索

ユーザはテーマチャンネルを指定して、データベースに保存されているサウンドから、TV チャンネルセレクトのように、ユーザの周囲に表示される複数のサウンドを選択する。現在、異なる場所での類似した音、同じ場所での異なる音、通りで話す音、ランダムな音の4つのテーマチャンネルが提供されている。異なる場所での類似した音チャンネルは、鳥のさえずりや川の音など、さまざまな場所で複数の同様の音を提供する。このチャンネルにより、ユーザは同様の音の多様性を認識することが可能となる。同じ場所での異なる音チャンネルは、ユーザが指定した場所でさまざまな異なるサウンドを提供する。このチャンネルにより、ユーザはターゲットの場所にさまざまな音があることに気付くことが可能となる。通りで話す音チャンネルは、さまざまな都市の通りで複数の人間が話す音を提供する。このチャンネルを使用すると、ユーザはさまざまな都市で現在アクティブになっているいくつかの会話を聞くことが可能となる。ランダムな音チャンネルは最後のチャンネルで、ランダムに多様な複数のサウンドを提供する。ヘッドジェスチャは、指定されたチャンネル内で現在提示されている音を置き換えることができるため、ユーザは提示された音を継続的に置き換えることにより、世界中から多数の音を聞くことが可能となる。

4.3 頭部ジェスチャによりサウンドナビゲーション

空間提示方式では、頭を傾ける方向の音を大音量で聞くことができるサウンドフォーカシング機能を実装した。たとえば、ユーザーが頭を右に傾けると、Unity の3D 空間のリスナーオブジェクトの右側にあるオーディオソースオブジェクトの音が大きく聞こえるようになる。図3(左)は、頭を傾けていないときの各オーディオソースオブジェクトの配置と音の大きさを示している。オーディオミキサーの各ボリュームバーは、すべての方向の音がほぼ同じ音の大きさであることを示している。図3(右)は、各オーディオソースオブジェクトの配置と、頭を後ろに傾けたときの音の大きさを示している。頭を後ろに傾けると、リスナーオブジェクトが後ろに移動し、後ろに配置された音が大きくなる。

4.4 CollectiveEars のソフトウェア構造

現在の CollectiveEars のプロトタイププラットフォームを使用する場合、ユーザは Nintendo JoyCon デバイスを搭載したヘッドフォンを着用する。現在の CollectiveEars は、人々が eSense デバイス[4]などのマイクを装着して聴覚を録音することを前提としている。現在のプロトタイプは、表2に示す4つのモジュールから構成されている。

表 2. CollectiveEars のソフトウェア構造

Table 2. Software Stricture of CollectiveEars

音生成モジュール	このモジュールは、ユーザが Unity を使用して聞くサウンドを生成する。複数のサウンドは、ユーザの3D サウンドスペースに配置される。音は、頭ベースのジェスチャを介してナビゲートすることが可能である。
音取得モジュール	このモジュールは、ウェアラブルマイクデバイスを使用して各ユーザのサウンドをキャプチャする。
頭部ジェスチャ管理モジュール	このモジュールは、加速度センサーデバイスを介して人の頭の動きを抽出する。頭の動きは、それぞれのジェスチャコマンドに変換される。
キーワードサウンドマッピングデータベースモジュール	このモジュールは、特定のキーワードに関連するサウンドを見つけるために使用される分散ストレージである。

5. 現在のプロトタイプからの経験と現在の研究の限界

このセクションでは、CollectiveEars と CollectiveEyes に関するいくつかの経験と、考えられる将来の方向性について説明する。

5.1 CollectiveEars and CollectiveEyes の経験

人々は視覚的なシーンからさまざまな具体的な情報を取得する。視界内のさまざまなオブジェクトを簡単に認識し、複数の視界の違いを理解する。CollectiveEyes を使用すると、ユーザはさまざまな視界を比較して、それらの違いを見つけることが可能となる。ただし、人間の視力の視野範囲は

上下 120 度, 左右 180 度と限られている. ユーザが範囲外を見たい場合, CollectiveEyes には, ユーザがアクセスしている別の人の現在の視点を変更するための追加のメカニズムが必要となる. もちろん, 360 度カメラを装着することで制限を克服できるかもしれないが, 特に何かを見つめているときに, 人々の視点に現在興味のあるものが含まれている可能性があるため, 制限は別の観点からもメリットになる可能性もある.

一方, 音は通常 360 度からキャプチャされる. ユーザは, ユーザが聴覚にアクセスする別の人の 360 度の音を聞くことが可能である. ユーザは通常, 現在の環境で選択的に何かを聞くが, CollectiveEars は, その人が現在注意深く聞いていることを知るのは困難である. また, 通常, 音だけで色々なものを認識することは容易ではない. たとえば, 他の人が現在聞いている音を聞いた場合, 一般的な音が認識できないと, その人がどのような状況にあるのかわからなくなる可能性がある. したがって, 多くの人の視点を比較することは効果的だが, 多くの人の聴覚を比較することは効率指向の活動には役立たないかもしれない. しかしながら, 多様な音を同時に体験することは多くのユーザにとって楽しい経験を提供できるかもしれない. したがって, CollectiveEars の有望な方向性の 1 つは, 新しいタイプのサウンドベースのエンターテインメントを探すことにもあると思われる.

聴覚のもう 1 つの本質的な特徴は, キャプチャされた音が人間の能力に強く依存しないため, 人間の聴覚と機械の聴覚の間に大きな違いがないことである. たとえば, 視覚はアクセス先の人物の視線を目標の出来事に向ける必要があるが, 聴覚は意識する必要はない. したがって, 人間の聴覚を機械の聴覚に置き換えるのは簡単であり, CollectiveEars は Cities and Memory のようなキャプチャサウンドサービスと大きな違いはないかもしれない. また, 類似の複数の音を明確に区別することは困難である. 一方, 人間は, 同様の複数の視界は, わずかな違いでも簡単に認識できる. したがって, 複数の視界が人間の思考を促進するための情報源になる. CollectiveEars に強力な実用的ユースケースがないという事実も, 1 つの本質的な理由となる.

5.2 将来課題

分散型デジタルプラットフォームは, 現在, 日常の活動を支援するための日常のインフラストラクチャになっている. 世界モデルを組み込むことで, 将来のデジタルプラットフォームがさまざまな人々の状況や文化をカスタマイズできるようにする新しい機会を提供する. このセクションに示されているように, より良い視覚と聴覚の提示方式は, それぞれの人によって異なる. ただし, これらの違いを認識するために, デジタルプラットフォームは私たちの現実の

世界についてもっと知る必要がある. 世界モデルは, さまざまなサービスからアクセスできる実世界に関するさまざまな現在および過去の情報を提供する [7]. 次のステップでは, CollectiveEars と CollectiveEyes に適した世界モデルを調査し, そのモデルをデジタルプラットフォームに組み込むことを検討する.

異化とは, なじみのあるものの知覚を高めるために, なじみのない, または奇妙な方法で一般的なものを聴衆に提示する芸術的手法であり, この手法は, 人々に隠されているが本質的な意味を認識させるのに効果的である [12]. この概念は, 次のステップにおいて CollectiveEyes と CollectiveEars を活用するために非常に有効であると考えられる. また, 異化技術は, ユーザが多様な複数の同じ機会を見ることによって, より多様な方法で考えるのに役立つ可能性がある. これらは, CollectiveEyes と CollectiveEars の最も有望な方向性であると考えている. なぜなら, 人間が深く反省的に考えるのを助けることは, 人間の社会を進化させる最も重要な能力の 1 つだからである. 異化を促進するためのデジタルプラットフォームは, CollectiveEars と CollectiveEyes の有望な方向性と考えられる.

6. 結論

この論文では, ユーザが人間の目と耳を集散的に共有できるようにする CollectiveEars と CollectiveEyes という名前の 2 つの分散デジタルプラットフォームを紹介した. これらのプラットフォームは, 革新的な将来のサービスを開発するための新しい機会を提供する. 次のステップでは, 実際のフィールドにプラットフォームを展開し, フィールドスタディを実施して, CollectiveEars と CollectiveEyes のより多くの機会と落とし穴を抽出する予定である.

参考文献

- [1] Juho Hamari, Mimmi Sjöklint, Antti Ukkonen, A., The Sharing Economy: Why People Participate in Collaborative Consumption. Journal of the Association for Information Science and Technology, 2015, Vol.67, No.9, pp.2047-2059.
- [2] Kohki Ikeuchi, Tomoaki Otsuka, Akihito Yoshii, Mizuki Sakamoto, and Tatsuo Nakajima, KinecDrone: enhancing somatic sensation to fly in the sky with Kinect and AR.Drone. In Proceedings of the 5th Augmented Human International Conference (AH'14). Article 53, 2 pages. 2014, DOI: <http://dx.doi.org/10.1145/2582051.2582104>
- [3] Shunichi Kasahara and Jun Rekimoto, JackIn: integrating first-person view with out-of-body vision generation for human-human augmentation. In Proceedings of the 5th Augmented Human International Conference (AH '14). Article 46, 8 pages. 2014, DOI: <http://dx.doi.org/10.1145/2582051.2582097>
- [4] Fahim Kawsar, Chulhong Min, Akhil Mathur, Alessandro Montanari, Earables for Personal-Scale Behavior Analytics. IEEE Pervasive Computing, 2018, Vol. 17, No. 3.
- [5] Risa Kimura, Keren Jiang, Di Zhang, Tatsuo Nakajima., Society

- of “*Citizen Science through Dancing*” In Proceedings of the 11th International Symposium on Ambient Intelligence. 2020.
- [6] Risa Kimura and Tatsuo Nakajima, Gathering People’s Happy Moments from Collective Human Eyes and Ears for a Wellbeing and Mindful Society. In Proceedings of the 14th International Conference on Augmented Cognition. 2020.
- [7] Risa Kimura. and Tatsuo Nakajima, Digitally Enhancing Society through Structuralism: Virtualizing Collective Human Eyesight and Hearing Capabilities as a Case Study. In Proceedings of the 8th International Conference on Distributed, Ambient and Pervasive Interactions. 2020.
- [8] Yefeng Liu, Todorka Alexandrova, and Tatsuo Nakajima, Using stranger as sensors: temporal and geo-sensitive question answering via social media. In Proceedings of the 22nd international conference on World Wide Web (WWW '13). ACM, New York, NY, USA, 803-814. 2013. DOI: <https://doi.org/10.1145/2488388.2488458>
- [9] Tatsuo Nakajima and Ichiro Satoh. A software infrastructure for supporting spontaneous and per-sonalized interaction in home computing environments. *Personal and Ubiquitous Computing* Vol.10, p.379–391, 2006. <https://doi.org/10.1007/s00779-005-0056-1>
- [10] Kodai Oiwake, Kosuke Komiya, Hina Akasaki, Tatsuo Nakajima, VR Classroom: Enhancing Learning Experience with Virtual Class Rooms, In Proceedings of the 11th International Conference on Mobile Computing and Ubiquitous Network, pp.1-6, 2018
- [11] Mizuki Sakamoto, Tatsuo Nakajima and Sayaka Akioka, Gamifying collective human behavior with gameful digital rhetoric. *Multimedia Tools and Applications*, Vol. 76, No.10, pp.12539-12581. 2017, DOI: <https://doi.org/10.1007/s11042-016-3665-y>
- [12] Viktor Shklovsky, Art as technique. In *Contemporary Literary Criticism. Modernism Through Poststructuralism*, R. Con Davis, Ed. 1917. Longman Press, New York and London
- [13] Hiroki Tsuboi, Shuma Toyama, Tatsuo Nakajima, T., Enhancing Bicycle Safety Through Immersive Experiences Using Virtual Reality Technologies”. In Proceedings of the 12th International Conference on Augmented Cognition, pp.444-456, 2018.
- [14] Cities and Memory, <https://citiesandmemory.com/> (Accessed October 1 2021)