

多重視点を表示するための分散基盤システム

木村 梨沙¹ 中島 達夫¹

概要：本論文では、複数の視点を表示する分散プラットフォームの設計と実装に関して述べる。提案するプラットフォームでは、多様な視点を現在の視点に拡張現実感技術を利用することにより表示をおこなう。また、ハンドジェスチャとヘッドジェスチャを用いることにより、これらの視点の操作を可能とする。また、ユースケースシナリオを用いて提案するプラットフォームの可能性と問題点を議論する。

キーワード：人間の視点、多重視点の表示、拡張現実感、ハンド/ヘッドジェスチャ

A Distributed Software Platform to Project Multiple Viewpoints

RISA KIMURA¹ TATSUO NAKAJIMA¹

Abstract: The paper presents the design and implementation of a distributed platform to present multiple viewpoints. In the proposed platform, multiple viewpoints are superimposed on the current viewpoint by using augmented reality technologies. The hand/head gesture is used to control the viewpoints. Finally, we discuss some opportunities and pitfalls of the proposed platform.

Keywords: Human Visual Capabilities, Presenting Multiple Viewpoints, Augmented Reality, Hand/Head Gesture

1. はじめに

人間にとって見ることは生活において外界の情報を取得し、状況に応じた行動をとるために必要不可欠である。しかし、見ることはあまりに当たり前であるため、多くの場合、世界を無意識に見ることも多く、現代の日々忙しい日常の中で、見ることの重要性を深く考える機会は失われている。見ることは単に情報を収集するための手段だけではなく、新しいことに気がつくとか、深く物事を考える機会を提供するなどの人間がより良い人生を過ごすためにも必要不可欠なものである。例えば、[14]では、プロダクトデザインやサービスデザインにおいて物事をよくみて反省的に考えることがより良い製品を作る上で重要であると主張している。

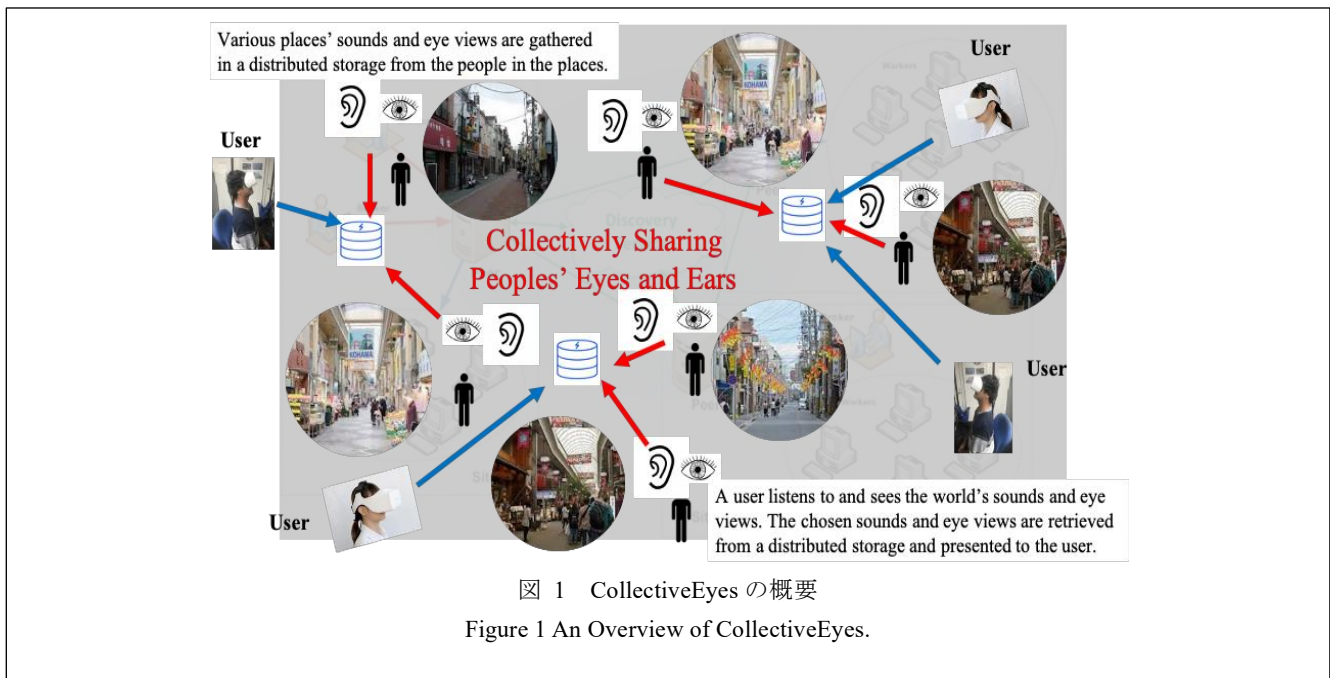
芸術は日常に対して通常とは異なる視点を提供することにより、鑑賞者に対して現実を深く理解させようとしている。ヴィクトル・シクロフスキーは、芸術の目的は、現実を異化することであると主張し [15]、ジェネビーブ・ベルは、日常生活を異化することは、日常品のデザインを理解する上で重要であると主張している [3]。パブロ・ピカソとジョルジュ・ブラックは、20世紀初頭にキュビズムを創始した。それまでの具象絵画が1つの視点に基づいて描かれていたのに対し、複数の視点による対象の把握と画面上の再構成により絵画を描いている [13]。具体的には、ルネサンス以来の単一焦点による遠近法の放棄と形態上の極端な解体・単純化・抽象化を特徴としている。フォトモンタージュは、2枚以上の写真を切り取り、接着し、再配置し、重

ね合わせて新しい画像にすることにより、合成写真を作成することにより作成された芸術作品である。デイヴィッド・ホックニーは、まるでキュビズムのように写真を多面的な角度から撮影し、同時に同じ画面に張り付けることで作品を制作することや、広く一般的に用いられている透視図法とは逆の視点で捉えた逆遠近法を用いることにより人間の眼に新たな視覚を提供しようとしている [11]。モンタージュとは、映画技法で、複数の映像の断片を組み合わせでひとつの連続したシーンを作る方法である。特に、映画作家セルゲイ・エイゼンシュテインのモンタージュは、より作家性や左翼的なイデオロギーと結びついており、むしろ観客の安定した日常を揺さぶり、それを積極的に問いかけるものとして提示された。エイゼンシュテインのモンタージュ論は、単なる映像形式的な方法論を超えて、いわば映像というメディアそのものの本質を思想として捉え、開発していこうとする試みであった [4]。

情報技術の発展は、上述のような芸術により様々な探索が行われてきた、人間の見るということを大きく変革する可能性を提供と思われる。特に、大量のビデオ映像を容易に撮影して蓄積できることは、芸術のように美術館において鑑賞することを前提とせず、日常生活の中で今見ているものに対して異化作用を生じさせることで、世界に対する意識を変革できる可能性があるのではないかと考える。

本論文では、複数の視点を表示するプラットフォームの設計と実装に関して述べる。提案するプラットフォームでは、以上に述べたような芸術手法の考え方を利用して、多

¹ 早稲田大学
Waseda University



様な視点を現在の視点に拡張現実感技術を利用することにより表示をおこなう。また、ハンドジェスチャとヘッドジェスチャを用いることにより、これらの視点の操作を可能とする。また、ユースケースシナリオを用いて手案手法の可能性や問題点の検討をおこなう。

2. 関連研究

WindowSwap [16]を使用することで、世界中の人々がビデオショットウィンドウから風景を見ることができるようになる。人々が風景を目にするのは珍しいことではないが、知らない人がこのように窓の外を見ているという一体感を味わうことを可能とする。人々はグーグルマップのストリートビューを用いても世界の風景を見ることができるが、WindowSwap の特徴は人々が誰かの生活により密接に関連した視点を獲得することができるということである。表示される映像には、各動画の左上に撮影者の名前、右上に地名が表示される。部屋からの眺めなので、トースター、植物、マグカップなどのそこに住んでいる人々の持ち物も映る。

Jacqui Kenny は非常にユニーク旅行写真家である [17]。彼女の写真は、アラブ首長国連邦のラクダレース、ペルーの街のプール、テキサスの乾燥した風景の中のパステルカラーの家、またはチリの山々を背景に子供たちが遊んでいる場所にユーザーを連れて行くことを可能とする。それらの写真は、非常に印象的で、独特の視点と世界観を備えており、それらはすべて Google ストリートビューの画面をつかむ静止画によって撮影されている。彼女は、「私たちがお互いにどれほど似ているかを実感しました。家から旅行したにもかかわらず、不思議なことに、これまで以上に世界とのつながりを感じることができました。」と語っている。デジタルフォトモンタージュ [1] は写真の集合の一部を

単一の合成画像に結合するためのインタラクティブなコンピューター支援フレームワークである。このフレームワークでは、主に2つの手法を利用している。グラフィットの最適化では、構成画像内の適切な継ぎ目を選択して、可能な限りシームレスに複数の写真を組み合わせることを可能とする。また、ポアソン方程式に基づくプロセスである勾配領域融合により、合成写真に残っている目に見える人工性をさらに削減している。

Jack-In Head は、人間の能力と人間の存在を増強するための概念である [6]。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。このアプローチで使用されている他の人の視力にアクセスする手法は、ジャックインの概念に似ている。Procyk は、ヘッドマウントカメラとモバイルスクリーンを備えたペアの一人称ビデオ交換を調査した [11]。この作品は、ペアの相互ビデオストリーミングが公共の場で並行体験を達成したことを示した。KinectDrone は、空を飛んでいるという私たちの体の感覚を高めるシステムである [5]。ドローンによってキャプチャされたビデオは、ユーザーのスマートグラスに送信される。ユーザーは空を飛んでいるかのように振る舞い、部屋にとどまりながら飛行ドローンが捉えたシーンを見るようになる。

3. 背景:目の共有システム

図 1 に示す CollectiveEyes は、人間の視力と聴覚を共有する分散プラットフォームであり、ユーザーが他の人の視覚と聴覚の機能を採用できる新しいサービスの開発を可能とする [7]。デジタル共有プラットフォームは、さまざまな魅力的な革新的なサービスを可能にするためのアプローチとして近年人気が高まっている [2]。これらのデジタルプラットフォームでは、私物や人々の暇な時間が他の人の

目的のために提供される。私たちのアプローチでは、人間の目と耳を他の人に貸すことを検討することに関心がある。

CollectiveEyes プラットフォームは、ユーザーの視線ベースのジェスチャを採用して、ユーザーが見たい目のビューを選択する。私たちは、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メガネ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、カメラとマイクを HMD の前に配置する。HMD は、カメラによってキャプチャされたビューを仮想空間に投影する。

複数の視点を表示する場合、それらの視点は仮想空間に表示される。CollectiveEyes は、空間ビューモードを提供する。空間ビューモードを使用すると、それぞれの視点を提示するための複数のビュー (図では4つのビュー) が自動的に選択され、仮想空間に表示される。表示された視点の1つがユーザーの興味を引かない場合は、削除されたビューの代わりに別のビューが表示される。

多数の視点から目的とする視点を探すため、CollectiveEyes はチャンネルという抽象化を提供している。チャンネルはユーザの興味を指定するための抽象化で、CollectiveEyes は複数の種類のチャンネルを提供している。キーワードベースチャンネルは、ユーザがしているキーワードを元に視点の選択をおこなう。位置ベースチャンネルは、視点が提供された位置に基づいて視点の選択をおこなう。最後のテーマベースチャンネルは CollectiveEyes が提供するいくつかのテーマに従って視点の選択をおこなう。

CollectiveEyes は、様々な目的のために拡張されてきている。[8]では、日常生活におけるポジティブ感情を増幅するプラットフォームとして拡張され、[8]では、タンパク質のドッキングを探索するシズンサイエンスゲームのための社会視聴システムとして拡張されている。また、耳の共有を可能とする CollectiveEars も提案されている [10]。CollectiveEyes は人の視点を共有することを目的として提案されたが、非人間の視点を導入することによりより多様な視点の共有が可能となる。本論文が提案するシステムは、CollectiveEyes を非人間の視点も導入する際のインフラストラクチャと考えることができる。

4. プロトタイプシステムの設計と実装

本節では、現在実装中のプロトタイプシステムの概要に関して紹介する。本プロトタイプシステムは前節で紹介した CollectiveEyes を拡張したものとなっている。4.1 節では多数の視点を表示するための抽象化として提案するアンビエントフォトモンタージュに関して説明する。4.2 節では、本システムを使ってユーザはどうやって多数の視点にアクセスするかを示すシナリオを示す。最後に、4.3 節では現状の実装の概要に関して紹介する。

4.1 アンビエントフォトモンタージュ

キュビズムやフォトモンタージュは、同一の空間について一人の撮影者が複数の視点から多重化して一つのものとして表現している。これらのアートにおける考えを CollectiveEyes を組み合わせることにより、1つの空間に複数の人物の視点を多重化することが可能になる。これが、アンビエントフォトモンタージュという我々が提案するアプローチの基本的な考え方である。

アンビエントフォトモンタージュでは、従来のキュビズムやフォトモンタージュのように、1個人の様々な視点から見た空間だけでなく、様々な人々の視点をユーザの現在の視点上で見ることを可能とする。

アンビエントフォトモンタージュの主な機能を以下にまとめる。

・視点の共有

CollectiveEyes により、様々な人々の視点を共有し、それらの視点を見るのが可能となる。様々な人々の視点を見ることにより、世界中の人々がどのようなものを見ているのか気が付く事ができ、また、同じ空間の様々な視点を見ることにより新たな発見が期待できる。



図 2 空間的表示方法

Figure 2 Spatially Presentation Method



図 3 時間的表示方法

Figure 3 Temporally Presentation Method

・複数視点の多重表示

キュビズムやフォトモンタージュのように、複数の視点を現在の視界上に表示する。現状のアンビエントフォトモンタージュは空間的表示方法と時間的表示方法の2つの方式を提供している。図2は空間的表示方式、図3は時間的表示方式の基本的考え方を示している。

・360度動画の空間的複数表示

ユーザの視点によりキャプチャーした360度動画を多重化して複数表示するために、複数の動画を角度ごとに配置し、ユーザーが見渡すと動画が切り替わる。表示する動画の数として2~8を選ぶことが可能である。

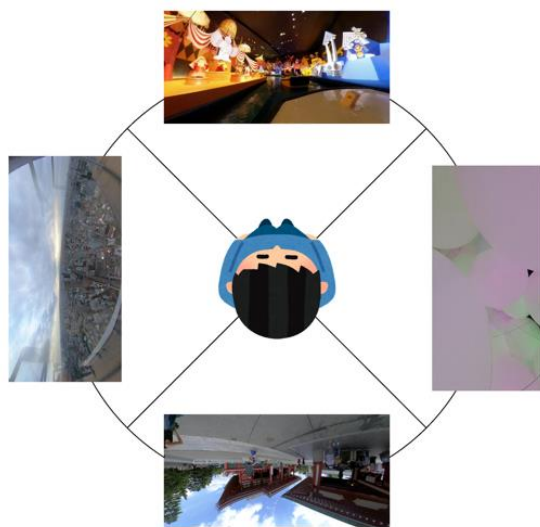


図 4 360度動画の空間的複数表示

Figure 4 Spatial Presentation of 360 Degree Videos

・非人間視点の表示

人々の視点だけではなく、ペットなどの動物からの視点、ドローンによる鳥瞰的視点、街中や自動車等に埋め込まれた様々なカメラからの視点を導入することにより、人の視点のみからでは得られない情報も提供する。

・視点へのエージェンシの追加

ユーザの視点にその視点を提供している人々の情報を追加することにより、視点を提供している人々のエージェンシが感じられるようにする。現状では、人々のその時の感情などプライバシーの公開にならない情報を提供することを検討している。

4.2 シナリオ

本節では、アンビエントフォトモンタージュ・プロトタイプシステムの利用場面を想定し、ユースケースシナリオを示す。

「ゆきさんは東京に住む高校生です。コロナウイルス感染

対策のため、外出をすることを極力避けてきたゆきさんは、外にいる気分を味わいたくて CollectiveEyes を使って渋谷の様子を見てみることにしました。ゆきさんは渋谷の様々な映像を同時に見たかったので、アンビエントフォトモンタージュ機能を利用しました。まず、360度映像の空間的表示で頭を動かしながら9つの視点映像を同時に見比べました。」(図5)



図 5 多視点の表示

Figure 5 Presenting Multiple Views

「その後、9つの視点映像を選択し、360度視点映像を見ました。その動画では他のユーザである人間の視点を動画として提供します。映像には提供したユーザの感情が表示されており、ゆきさんは同じく楽しい気持ちになることができました。視点映像には他の視点映像もフォトモンタージュ的に複数表示されており、ゆきさんはハンドジェスチャーを使うことでそれらの視点の透明度を上げたり下げたりしながら様々な視点の動画を同時見てそれぞれのユーザの感情や視点の意味を考えました。」(図6)

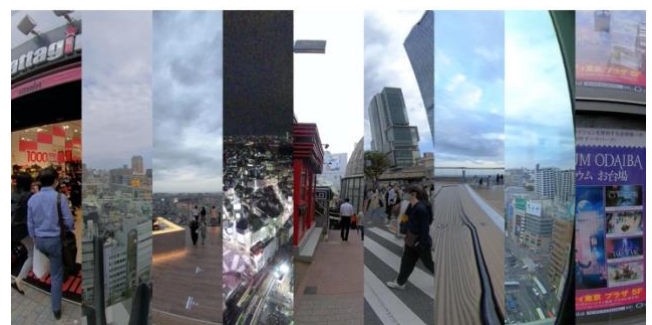


図 6 360度動画の多視点

Figure 6 Multiple Views from a 360 Degree Video

「ゆきさんは複数の視点を見て回っていると、モノの視点の動画を見つけました。非人間の視点では提供したユーザの感情というものは無いので、ゆきさんは自分がこのモノの目線に立ったらどのように感じるか考えながら視点を鑑賞しました。」(図7)



図 7 非人間の視点
Figure 7 Non-Human Views

4.3 実装

図 8 は、CollectiveEyes プロトタイププラットフォームの各ノード上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、頭につける加速度センサを利用して頭の動きを使って 360 度映像をナビゲーションすることができる。また、複数の視点の選択のために、Leap Motion を用いたハンドジェスチャを用いる。視点のナビゲーションのためには、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。

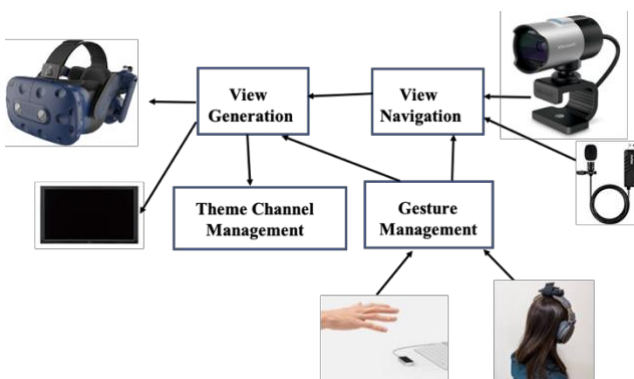


図 8 プロトタイプシステムの概要
Figure 8 An Overview of a Prototype Implementation

現在のプロトタイプは、以下の 4 つのモジュールから構成されている。

ビューの生成モジュール：このモジュールは、ユーザーが 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。最初のモードはマップを表示、2 番目のモードは複数の視点を表示、3 番目のモードは選択された視点を表示する。ビューは、ハンドジェスチャによってモードを切り替えることが可能である。

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できる。

ジェスチャ管理モジュール：このモジュールは、頭の動きをモニタリングするための頭につけた加速度センサーとハンドジェスチャを認識するための Leap Motion からの入力を管理する。

テーマチャンネルマネジメントモジュール：このモジュールは、テーマチャンネルを管理するためのものである。



図 9 プロトタイプシステムの複数視点の表示 (空間的表示方式)

Figure 9 Multiple Views in a Prototype System (Spatial Presentation Method)



図 10 プロトタイプシステムの複数視点の表示 (時間的表示方式)

Figure 10 Multiple Views in a Prototype System (Temporal Presentation Method)

図 9 と図 10 は、現状のプロトタイプシステムのスクリーンショットを表示している。空間的表示方法の場合、手を上下に動かすことにより他の視点の透明度を変更が可能である。時間的表示方法の場合、手を上下に動かすことにより視点をスキップすることが可能である。

5. 議論と今後の課題

本節では、第 4 節で十分に議論しきれなかったいくつかの課題に関して議論をおこなう。第 1 の議論点は、ユーザーからの視点の選択法である。3 節で説明したように、CollectiveEyes はチャンネルと呼ぶ抽象化を提供している。チャンネルはユーザーの興味を指定するための非常に強力な抽象

化である。チャンネルの性質を様々な定義することで視点の多様な選択を可能とする。しかし、どのようなチャンネルがどのような場合に有効であるかは必ずしも現状で明らかでない。また、視点をどのようにチャンネルにカテゴリ化するかも今後の検討を要する。現状では、視点の提供者が視点がどのチャンネルに含まれるかを各自にマニュアルに定義するようにしているが、利便性や動機付けを考慮して再検討する必要がある。

第2の議論点は、集合的に見る機能 (Collective Seeing)の実現に関するものである。CollectiveEyes は1つの視点を複数のユーザが同時に見ることを可能とする。これは、複数の人が同一の人となった感覚を提供できる可能性を示唆する。異なる場所にいる異なる人々が同一の場所で同一の人として同じ視点を共有することができるようになる。しかし、視点を共有していると感じるためには、共有感覚を提供するための仕組みが必要となる。特に、それぞれはそれぞれが意識することに応じて異なる方向を見ようとする。その自律性が提供されないとユーザは自分自身であるというエージェンシを感じにくい。つまり、自分自身であり、他人でもあるという共感感覚の提供が必要になる。

第3の議論点は、アンビエントフォトモンタージュを使い複数の視点をスーパインポーズするとき、どのような観点で表示する視点を選ぶのか、表示する視点をどのように変更するかである。同種の視点のみの提供は新たな発見に結びつきにくい。しかし、興味がなくさんの視点が表示されることは、ユーザの興味を失わせる原因となる。そのため、新たな発見に導く可能性を残す程度に多様性を提供することが望ましい。この場合も、興味がなくというユーザの意志を反映させることで、意味がない多様性を提供することを低減することが望ましい。我々は、Amazon Discover [18] が利用している方針を採用する。つまり、ユーザが表示された視点に対して、興味があるかないかを指定することで、興味がなくカテゴリーの視点が表示される可能性を低下させる。これにより、ユーザが指定したチャンネルに属しない視点を若干表示させることで、ユーザが異なるチャンネルに属する視点をアクセスする機会を提供することを可能とする。

最後の第4点は、ユーザが現在見ている視点にどう意識させるかということである。人は多くの場合、自分が現在見ているものを必ずしも常に意識しているわけではない。自分の興味あることや、現在のゴールを達成するのに必要な視覚情報を得る必要が生じて初めて視点が意識される。アンビエントフォトモンタージュは、現在ユーザが見ている視点に様々な視点をスーパインポーズすることにより、異化作用が生じて現在の視点内の気がつかなくなったことに気がつく可能性を提供する。しかし、確かにスーパインポーズした視点の存在は違和感を生じるので異化作用を生じるかもしれないが、スーパインポーズされた複数の視点感

や、それらと現在の視点の間に意味ある差異が存在しないと効果的な異化作用を生じることが不可能である。映画のモンタージュやホックニーのフォトモンタージュは、複数の写真や映像間の意味ある差異が絵画の芸術的效果や映画の劇的效果を生じる。アンビエントフォトモンタージュでも、如何にして複数の視点が意味ある差異を提供できるようにするか今後の課題となる。

6. おわりに

本論文では、複数の視点を表示するプラットフォームの設計と実装に関して提案した。現状では、本格的な実装に向けた大枠の方針を決定するためのペーパープロトタイプを持ちた実験と初期的な実機プロトタイプの実装をおこなってきた。今後は、5節で述べたような課題を検討しつつ、実装を進め、ユーザ評価をおこなう環境を整えていく予定である。

参考文献

- [1] Aseem Agarwala, Mira Dontcheva, Maneesh Agrawala, Steven Drucker, Alex Colburn, Brian Curless, David Salesin, and Michael Cohen. 2004. Interactive digital photomontage. *ACM Transaction on Graphics*. 23, 3, 294–302. 2004.
- [2] Juho Hamari, Mimmi Sjöklint, and Antti Ukkonen. The Sharing Economy: Why People Participate in Collaborative Consumption. *Journal of the Association for Information Science and Technology*, Vol.67, No.9, pp.2047-2059. 2015.
- [3] Genevieve Bell, Mark Blythe, and Phoebe Sengers. Making by making strange: Defamiliarization and the design of domestic technologies. *ACM Trans. Comput.-Hum. Interact.* 12, 2 (June 2005), 149–173. 2005. DOI:<https://doi.org/10.1145/1067860.1067862>
- [4] Sergei M. Eisenstein. *Sergei Eisenstein Selected Works: Towards a Theory of Montage*. Tauris Academic Studies. 2010.
- [5] Kohki Ikeuchi, Tomoaki Otsuka, Akihito Yoshii, Mizuki Sakamoto, and Tatsuo Nakajima. KinectDrone: enhancing somatic sensation to fly in the sky with Kinect and AR.Drone. In *Proceedings of the 5th Augmented Human International Conference (AH'14)*. 2014.
- [6] 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, 2014*, Article 46, 8 pages.
- [7] Risa Kimura and Tatsuo Nakajima. Collectively Sharing People's Visual and Auditory Capabilities: Exploring Opportunities and Pitfalls. *SN Computer Science*, Vol. 1, No. 5, Article 298, 2020.
- [8] Risa Kimura, Keren Jiang, Di Zhang, Tatsuo Nakajima. Society of "Citizen Science through Dancing". In: Novais P., Vercelli G., Larriba-Pey J.L., Herrera F., Chamoso P. (eds) *Ambient Intelligence – Software and Applications*. ISAmI 2020. *Advances in Intelligent Systems and Computing*, vol 1239. Springer, Cham. 2020. https://doi.org/10.1007/978-3-030-58356-9_2
- [9] Risa Kimura, Tatsuo Nakajima. Gathering People's Happy Moments from Collective Human Eyes and Ears for a Wellbeing and Mindful Society. In: Schmorow D., Fidopiastis C. (eds) *Augmented Cognition*. *Human Cognition and Behavior*. HCII 2020. *Lecture Notes in Computer Science*, vol 12197. Springer, Cham. 2020. https://doi.org/10.1007/978-3-030-50439-7_14
- [10] Risa Kimura, Tatsuo Nakajima. Opportunities to Share Collective Human Hearing. *Hearing*. In *Proceedings of the 11th Nordic*

Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society (NordiCHI '20), 2020, 4 pages.
<https://doi.org/10.1145/3419249.3421244>

- [11] Marco Livingstone, David Hockney, Thames and Hudson Ltd, 2017.
- [12] Jason Procyk, Carman Neustaedter, Carolyn Pang, Anthony Tang, and Tejinder K. Judge. Exploring video streaming in public settings: shared geocaching over distance using mobile video chat. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14). 2014
- [13] John Pultz, Catherine B. Scallen, Cubism and American photography 1910-1930, The Sterling and Francine Clark Art Institute, 1984.
- [14] Donald Schön. 1983. The Reflective Practitioner: How Professionals Think in Action. Basic Books, New York.
- [15] Viktor Shklovsky. Art as technique. In Contemporary Literary Criticism. Modernism Through Poststructuralism, R. Con Davis, Ed. Longman Press, New York and London. 1917.
- [16] WindowSwap, <https://www.window-swap.com/>(Accessed August 1 2020)
- [17] The Agoraphobic Traveller, <https://www.pinterest.jp/ashleyschleper/jacqui-kenny/> (Accessed November 15 2020)
- [18] Amazon Discover, <https://www.amazon.co.jp/discover/> (Accessed November 15 2020)