

Software Defined Media:

視聴空間サービスのソフトウェア制御

塚田 学 小川 景子 池田 雅弘 曾根 卓朗 丹羽 健太

齊藤 翔一郎 粕谷 貴司 砂原 秀樹 江崎 浩

インターネットを前提とした視聴サービスが登場し、中でも空間に存在する視聴対象を三次元的に解釈するオブジェクトベースの視聴サービスの重要性が増している。そこで著者らは2014年からSoftware Defined Media (SDM) コンソーシアムを設立し、オブジェクトベースのデジタルメディアと、インターネットを前提とした視聴空間の研究を続けてきた。その中で、Internet of Things (IoT) を利用したスマートビルディングと相互運用を目指して、視聴サービスのためのビル設備を仮想化するSDMアーキテクチャを提案している。SDMアーキテクチャに基づき、試作システムを構築し、デモンストレーション展示ハッカソン開催を行った。さらに、試作システムを利用したアプリケーション開発と収録データのデータベース化を行った。プラットフォーム上で動作するアプリケーションの遅延を測定し、多くのインタラクティブなアプリケーションを支援しうるリアルタイム性を確認した。また、試作システムを利用したインタラクティブなアプリケーションを被験者に体験してもらい、主観実験によって開発者の演出意図を高水準で満たせたことを示した。

Internet users come to enjoy various audio-visual services these days. Among these services, object-based audio-visual services are becoming important. We started the Software Defined Media (SDM) consortium to investigate object-based digital media and Internet-by-design audio-visual service. We design the SDM architecture that virtualizes the underlying audio-visual infrastructure installed in the place. We implemented a prototype system based on the architecture and presented in Interop 2015 Tokyo. Moreover, we organised hackathons to investigate the innovative application of the platform. We also study a tangible interface using the SDM prototype. To make a database of object-based audio-visual data, a baroque orchestra was recorded by various methods with 82 channels of microphones in total. We measured the delay of the application on the prototype platform and the result shows that the system has enough real-timeless to support major applications. The experimental evaluation of an interactive SDM application shows that the prototype system allows reflecting the intention of the application developer well.

Software Defined Media: Flexible Services Platform for Audio-Visual Applications.

Manabu Tsukada, Hiroshi Esaki, 東京大学情報理工学系研究科, Graduate School of Information Science and Technology, The University of Tokyo.

Keiko Ogawa, Hideki Sunahara, 慶應義塾大学メディアデザイン研究科, Graduate School of Media Design, Keio University.

Masahiro Ikeda, Takuro Sone, ヤマハ株式会社, Yamaha Corporation.

Kenta Niwa, Shoichiro Saito, 日本電信電話株式会社, NTT Media Intelligence Laboratories.

Takashi Kasuya, 株式会社 竹中工務店, Takenaka Corporation.

コンピュータソフトウェア, Vol.34, No.3 (2017), pp.37-58. [研究論文] 2016年8月22日受付.

1 はじめに

2020年の東京オリンピック・パラリンピックの開催に向けて、世界からの訪問者への高度なおもてなしを目指した社会インフラのスマート化が急がれている。中でも、オリンピックの理念である「スポーツを通して心身を向上させ、さらには文化・国籍など様々な差異を超え、友情、連帯感、フェアプレーの精神をもって理解し合うことで、平和でよりよい世界の実現に貢献する」という目標を支援するため、地球規模での一体感を生み出すような視聴メディアに関する研究開発の重要性が高まっている [1]。そのためには、グ

ローバルなインターネットを前提とした映像・音響の視聴空間と視聴メディアの設計が必要である。

近年は、多くのスマートフォンに標準的に内蔵される収録機器を常時持ち歩くことが増え、収録された映像音声の情報はインターネット上で瞬時に伝達・共有・加工される状況が出来上がった。さらには、収録対象から映像素子に入力されたビットマップ情報と、ステレオマイクに入力された2チャンネルの音声情報としてそのまま伝送し、受信側でそのまま再生するだけでなく、空間に存在する収録対象を三次元モデルとして解釈し複数の視聴オブジェクトに分解して伝送し、受信側ではこれらのオブジェクトを用いて空間を再合成するオブジェクト指向の方式が注目を集めている。これにより、ヘッドマウントディスプレイ (HMD)、3D テレビ、立体音響装置などの受信側のシステムの構成に合わせた柔軟な三次元表現が可能となるだけでなく、他のコンテンツの視聴オブジェクトを別途受信し組み合わせることで、今までにない表現が可能になる。

例えば、音響においては、Dolby Atmos [2] や AuroMax [3] などの映画館やホームシアター、さらには個人向けモバイル機器を対象に、音のオブジェクトから三次元の音場を生成する立体音響システムが登場している。また、映像においては、複数の地点・角度から撮影された映像・動画から、撮影した空間に存在する三次元オブジェクトの抽出が可能であり、抽出した三次元オブジェクト情報を用いて、任意の視点(自由視点)からの映像の作成・再構築が可能となりつつある [4] [5]。こうした流れに目を向けると、今後はインターネットで収録環境と再生環境を双方向で接続し、視聴オブジェクトを交換しながら、三次元表現を持つ情報空間をエッジヘビーコンピューティングまたはクラウドで計算処理することによって、映像音声を作り出されていくことになるであろう。

映像と音声のオブジェクト化が融合することで、従来の配布型コンテンツビジネスを超えた、新しい次元のインタラクティブなオリンピック・パラリンピックの視聴形態などこれまでにないビジネス領域や、これまでにないデジタルメディアを用いた表現方法などを開拓・開花させることが期待される。例えば、音

と映像をオブジェクト化することで、スポーツ観戦者の要求に応じてリアルタイムに応援スタンド側からの視聴や、選手の一人称視点からの視聴を切り替えることが可能になるだろう。さらに、コンサートの視聴においては、複数の歌手の中の1人や、演奏者のうちの1人をズームアップすると映像に追従して音もズームアップするなどのインタラクティブなコンテンツ制作が可能になる。

さらに映像・音響を作り出す空間に目を向ければ、例えば、劇場、ホームシアター、テレビ会議システム、講義室など、様々な用途に対して専用に設計されたシステムが存在している。これらの映像・音響システムには、收音システム、拡声システム、制作システム、BGMや館内放送のためのシステムなどのサブシステムから構成されている。こうしたサブシステム間で信号の伝送を行う場合は高精度同期と低遅延を求められるため、従来はアナログ線でシステム間を結ぶ必要があった。例えば、劇場では、演劇、軽音楽、講演といった様々な演目に対応する必要があり、各演目によって要求される演出が異なるため音声信号の伝送路も必要に応じて組み替えるが、アナログパッチの差し替えや大規模なマトリクスなどによって組み換えを実現しているため、処理が複雑になり、また事前の設定通りに組み替えるため、柔軟性に欠けていた。

しかし現在では、IPを利用した1Gbps環境等の大容量通信環境の一般化により、各サブシステム間をIPによりネットワーク化することが可能となった。これにより、映像音声の入力から制作、出力までの伝送路の組み換えを一貫してソフトウェアによって制御するシステムの実現が期待される。ソフトウェアによって定義された演出を、目的に応じて柔軟に変更することが可能になれば、1つのシステムを講演会、遠隔会議、音楽イベント、映画や放送の上映など様々なアプリケーション向けに最適化し活用することも容易になる。例えば、アプリケーションを載せ換えることで、映画館を遠隔音楽教室として利用するなどのユースケースが考えられる。演出者の意図を汲み、柔軟に視聴空間をソフトウェア制御するためには、映像・音声の設備機能を抽象化して管理し、演出に関する意思決定を行うアプリケーションと実際の映像・音響の

入出力を行うデバイスを引き離して (decoupling) 扱うべきである。その際、実際の入出力デバイスを仮想化し、複数のアプリケーションから同時並列に利用可能にする。また、視聴空間のソフトウェア制御は、同じく IP ネットワークの一般化により生まれた照明や空調などをソフトウェア制御するスマートビルディングの技術との親和性を持ったシステムであることが望ましい。

オブジェクト指向のデジタルメディアと、ネイティブ・デジタルなインターネット環境が前提の映像・音響空間を用いたビジネス創造を目指し、著者らは 2014 年 1 月に Software Defined Media (SDM) コンソーシアム^{†1} を設立した。SDM とは、映像・音響システムの IP ネットワーク化を背景に、これらの設備の機能に対して抽象化・仮想化を行い、サービスとしての映像・音響を提供するための基盤的なアプローチである。

本論文では、SDM の実現のためにコンソーシアムで行ってきた研究について報告する。2 章では関連研究を概観し、3 章で本論文の研究動機と SDM の目的について述べる。4 章では、SDM を実現するためのアーキテクチャについて説明し、5 章でそのアーキテクチャに基づいて構築した試作システムについて報告を行う。6 章では、試作システムをデモンストレーション展示した事例や、試作システムをプラットフォームとして利用し、アプリケーションの作成を行ったハッカソン開催、さらに試作システムを利用したタンジブルインターフェースの研究について報告する。さらに、アプリケーション非依存な視聴オブジェクトの管理を行うため、最大 24 名によるコンサートを、さまざまな方式のマイクとカメラで収録し、収録データをデータベース化した方式について述べる。7 章では、SDM 試作システムの性能評価を行う。8 章にてまとめを行い、今後の課題について述べる。

2 関連研究

空間表現を持った録音再生はステレオサウンド (2 チャンネル) から始まりサラウンドサウンド (多チャ

ネル) へと発展してきた。また、2016 年 8 月より試験放送の始まったスーパーハイビジョン [6] は、高さ方向への表現を持った 22.2 マルチチャンネルの音響システム [7] と、8K 映像を採用している。一方、チャンネルベースの再生方法は各チャンネルに割り当てられた情報から音を再生するのに対し、オブジェクトベースの再生方法として音源の三次元位置と音色、再生環境のスピーカの位置からソフトウェア制御によって音場をレンダリングする方式が登場した。例えば映画館での採用の進む Dolby Atmos [2] や AuroMax [3] の他に、オブジェクトベースの再生方式は国際標準化機構 (ISO) と国際電気標準会議 (IEC) の Moving Picture Experts Group (MPEG) において MPEG-H [8][9] として標準化が進んでいる。

録音技術においても、マイクロフォンアレイからの入力を信号処理することによって、遠くにある狙った音だけをクリアに集音する技術 [10] [11] の研究開発が進んでいる。こうした研究開発は、再生技術で進むチャンネルベースの技術から、音源の三次元位置情報を記録するオブジェクトベースの技術への発展と通じる技術である。

一方、映像収録技術においても、4K や 8K などの高精細映像化を進める方向と平行して、収録対象の三次元解釈を取り入れる方式が登場している。例えば、視聴者が任意に選択する視点に合わせて三次元の空間を表示することが可能な自由視点映像の合成技術が注目されている [12] [13]。特に、サッカーのような屋外大空間で撮影される映像を対象とした合成方式として、簡単化三次元モデルを用いる手法が提案されている [14] [15]。

スピーカ、マイク、カメラやディスプレイなどの視聴空間の設備は、建築設備システムの一部と捉えることができる。建築設備システムでは、空調や照明を制御することで、電力を削減することなどを目的とする Building Energy Management System (BEMS) の導入が進んでいる。ビル設備を IP ネットワークで繋ぎ高度な制御を行うためのオープンプロトコルとして、IEEE1888 が制定された [16]。また、ビル設備の IP ネットワーク化を背景に、Machine-to-Machine (M2M) での利用が想定されている MQ

^{†1} <http://sdm.wide.ad.jp/>

Telemetry Transport(MQTT) [17] によるビル設備制御の研究開発も進んでいる [18].

映像音声の遠隔配信は IP ネットワークの高速化に伴い、4K60p 映像を複数同時に配信する実証実験が開始されている。高精度な映像音声の双方向での遠隔配信を行うことで、遠隔コラボレーションシステムを構築すると、立体感や情感など、臨場感を高めることができる [19]。さらに、同様の技術を利用してデジタルシネマ劇場へライブ配信することで、劇場をパブリックビューイングの会場として転用することが可能になる。これにより、同じく高精度化するホームシアターから劇場を差別化し、劇場施設の提供できる価値を大きく向上させると考えられ、ビジネス的な展開が期待されており、商用実証実験が行われた [20]。

遠隔のユーザ同士が快適に会話し、効率的に協働作業するためのコミュニケーションシステムの開発が進んでいる。NTT では、同じ部屋にいる感覚を同室感と名付け、同室感コミュニケーションシステム「t-Room」を開発した [21] [22]。t-Room では、ユーザ全員が周囲の音や映像に関して同じ認識や知覚を対称的に共有することを目指し、同じ大きさの部屋に複数の背面スクリーンを囲い込んで設置している。

バーチャルリアリティの視覚ディスプレイとしては CAVE [23] や多面型全天周ディスプレイ (CABIN) [24] を始めとする没入型多面ディスプレイの研究が行われて来た。こうした 4 面の壁と地面を含めた没入多面ディスプレイは HMD とは違い、その場にいる多人数のユーザに同時に仮想世界の体験を提供できる。したがって、こうしたディスプレイに追従する音響の提供はヘッドホンではなく、複数のスピーカによるインタラクティブ高臨場感音場提示手法が検討されて来た [25]。また、NHK 放送技術研究所は、2 眼立体の三次元コンピュータグラフィックス再生と、映像に追従してインタラクティブに三次元の音場をスピーカアレイを用いて再生するシステム [26] を組み合わせインタラクティブ 3D 映像音響再生システムを開発した [27]。

一方で、音声を扱うタンジブルユーザインターフェイスについての研究が進んでいる。ミュージックボトル (musicBottles) はユーザが小瓶の蓋を開けると

音楽が流れ、閉じると音楽が止まるタンジブルユーザインターフェイスであり、小瓶の蓋を開け閉めすることでデジタルコンテンツへのアクセスを実現する [28] [29]。Tangible Sound は、水流と音楽の流れを関連付けたインタラクティブシステムを試作し、音楽表現を拡張した [30] [31]。

オープンデータのメタデータ記述手法として Resource Description Framework(RDF) が広く利用され、それらのデータが URI で相互参照され構築されるデータのネットワークは Linked Open Data(LOD) と呼ばれる。RDF や LOD は、現在例えば、人文社会科学系大規模データベース [32] や、ノックアウトマウスの表現型のデータベース [33] に使われ、データの横断的な利用を促進している。さらに、データ間の関係を示す語彙やデータの種類(クラス)を表す語彙は RDF Vocabulary と呼び、音楽データを記述する RDF Vocabulary は Music ontology [34] として定義されて、英国放送協会 (BBC) などで広く利用されている。

以上、三次元の映像と音声に関する研究、遠隔映像配信システム、遠隔コミュニケーションシステム、バーチャルリアリティやタンジブルインターフェイスの研究、メディアの管理手法を概観した。しかし、映像音声の設備機能をスマートビルディングの一部として捉え直し、それらの機能の仮想化と抽象化を通じて、アプリケーション開発者に様々な表現を可能にするプラットフォームを目指し取り組んだ研究はこれまでなかった。

3 研究動機と SDM の目的

SDM では、映像・音響の入出力設備機能を抽象化して管理することで、全体の意思決定を行うアプリケーションと、実際の映像・音響の入出力を行うデバイスを引き離して (decoupling) 扱うことが可能となる。これにより、演奏・上映される演目や、その演出上の目的のために、柔軟に映像・音響のシステムを構成することができる。また、演出に関する意思決定をアプリケーションとして一元化することで、映像音響システム全体の一貫性を持たせることを容易にする。これにより次の機能を実現する。

3.1 三次元の映像・音声の演出をソフトウェアで制御

今日、会議室、劇場などの用途別に特化した映像音声の設備が存在し、機能ごとにハードウェアベースのシステムが構築されているため、視聴空間の目的外利用は難しい。SDMでは、ソフトウェアの設定や処理によって、視聴空間の多目的な利用が可能となる汎用の設備による視聴空間を実現する。収録対象から映像素子に入力されたビットマップ情報と、ステレオマイクに入力された2チャンネルの音声情報としてそのまま伝送する従来方式では、映像音声の演出を制御することは難しいため、SDMでは、視聴オブジェクトを三次元表現を持った情報空間上で管理しながら、ソフトウェアの演出によって再生環境に適した映像と音響を作り出す。

3.2 映像・音声の複数ソースのミキシング可能性

出来合いのコンテンツだけでなく、インターネットと放送システムの伝送メディアと、リアルタイムとオンラインの複数のコンテンツから、視聴オブジェクトを受信して、受信側でのミキシングを可能とする。これによって、各拠点の視聴者の要求に合わせた柔軟な映像音声を作り出す。

現在、デジタルコンテンツの配布形態は、インターネットによるものや、TV放送やDVDパッケージなど、様々な形態が存在する。これらのコンテンツを大別するとリアルタイムなもの、アーカイブされたコンテンツが存在し、汎用的な再生環境においては、両方のコンテンツが混在することになる。現在のコンテンツ配信においては、リアルタイムなもの、アーカイブされたものに関わらず、視聴者は放送側が作成した映像音声のコンテンツを単一のソースから受信する形態が一般的である。

SDMでは、放送側が作成した映像音声のコンテンツを単一のソースからの受信だけでなく、複数のソースからコンテンツ素材の受信を可能にし、ソフトウェアによる受信側でのミキシングを実現する。さらに、複数ソースのコンテンツをミキシングする主体は、放送側と受信側の中間に位置するクラウドでの共同でのミキシングを行うことも想定した仕組みを構築する。

これにより、創造性豊かなクリエイターの手によるコンテンツが、第三者クリエイターらの共同作業によって進化し、視聴者の好みによって選択される視聴環境を作る。

3.3 ソフトウェアレンダリングによる拡張演出

SDMで交換される視聴オブジェクトは実空間由来のものである必要はなく、コンピュータを使って設計された映像音声による効果を組み合わせることによって、視聴者による満足度や臨場感を高める。

例えば、スポーツ中継などの臨場感を伝える場合、収録環境において収録した映像と音声を、再生環境において忠実に再現するだけでなく、様々な拡張演出を施したほうが視聴者による満足度や臨場感が高まる可能性がある。SDMでは、コンピュータを使って設計された映像音声による効果を組み合わせたソフトウェアレンダリングによる拡張演出を行う。例えば、応援するチームの応援を増加し、相手チームのブーイングを低減させることにより、応援するチーム側のスタンドで観戦しているような臨場感が得られるようなユースケースが考えられる。また、スタジアム会場での観戦人数、観客の興奮度、競技の内容などを計測し定量化し、再生環境にフィードバックすることにより、リアルタイムの拡張演出を可能にする。この拡張演出によって、スタジアムで収録された映像音声をそのまま再生するよりも、臨場感のある空間を作り出すことを目指す。

3.4 ユーザ・インタラクシオン性

SDMでは、視聴者の関心事を再生側のソフトウェアシステムにフィードバックし、演出の制御や、視聴者の望む映像音声のソースを選択可能にする。現在のコンテンツ視聴においては、放送局において制作されたコンテンツを受け取り再生するため、視聴者が特定の関心事に対して、能動的にズームアップしたり、違うアングルから視聴を行うことはできない。SDMでは、視聴者の関心事を再生側のソフトウェアシステムにフィードバックし、演出の制御や、視聴者の望む映像音声のソースを選択可能にする。ユースケースとしては、ファンのプレイヤーの一人称視点での映像音

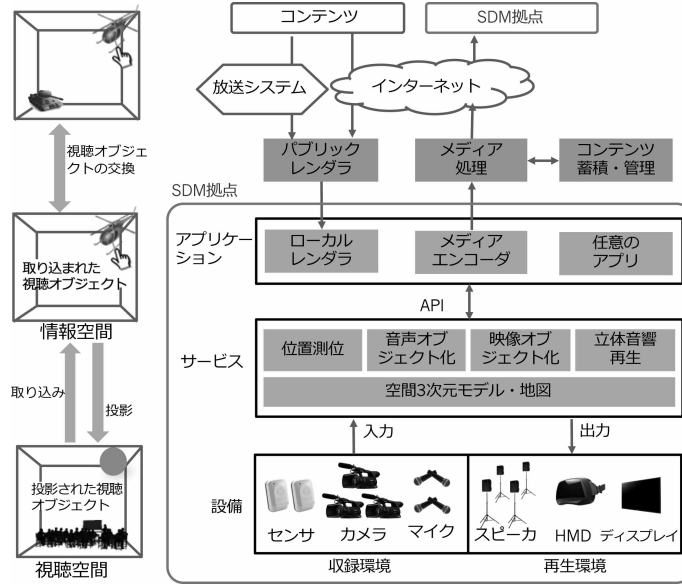


図1 SDM アーキテクチャ

声の視聴などが考えられる。

また、視聴者間での相互作用は、時間と空間の制約を受けるため、あるコンテンツに対して同一の関心事を持つ視聴者間においても相互作用は制限されている。SDMでは、同一の関心事を持つ視聴者間で、時間と空間の制約を超えた視聴者間の相互作用を実現する。例えば、スポーツ観戦において、テレビやパッケージメディアの視聴では、視聴する時間や場所を同じにする者以外の相互作用は限られているが、SDMでは同じチームのサポータ同士の声援が相互の視聴空間に再生するなどのアプリケーションが考えられる。

4 SDM アーキテクチャ

前章に述べたSDMの機能を実現するSDMアーキテクチャを図1に示す。

SDMは図1左に示すように、**視聴空間**(実空間)に存在する収録対象を三次元モデルとして解釈し、複数の視聴オブジェクトに分解して**情報空間**に取り込み、視聴オブジェクトの交換など様々な計算処理を行って、その結果を自在に視聴空間に**投影**することで視聴空間の再合成を行う仕組みである。上記の視聴空間と情報空間をつなぐ処理を行うSDMの基本単位を**SDM拠点**と呼ぶ。また、SDM拠点のシステム

は、図1右に示す通り、設備層、サービス層、アプリケーション層に分けられる。

設備層には、収録する空間に存在する映像音声の収録対象を三次元的に解釈しオブジェクト化するためのセンサ、カメラ、マイクなどの設備と、再生空間において視聴オブジェクトから三次元の映像音声を作り出すためのスピーカ、HMD、ディスプレイなどの設備がある。設備例はこれに限らず、これから登場する革新的デバイスの機能も取り込めるよう設備層の差異を吸収できる形で、上位のサービス層を設計する。また、SDM拠点は収録専用・再生専用の空間だけでなく、収録環境と再生環境は同一の実空間上に配置される可能性も想定して設計する。

サービス層は、設備層の設備機能を抽象化しアプリケーションに対してApplication Programming Interface(API)を通じてサービスとして提供するための層である。視聴オブジェクトと管理するための基礎となる空間三次元モデルと地図を定義し、位置測位、音声オブジェクト化、映像オブジェクト化、立体映像再生などのサービスを提供する。APIは開発者が創造的なアプリケーションを実現できるよう柔軟かつ拡張性を持つ。

アプリケーション層は、APIを通じてSDMサー

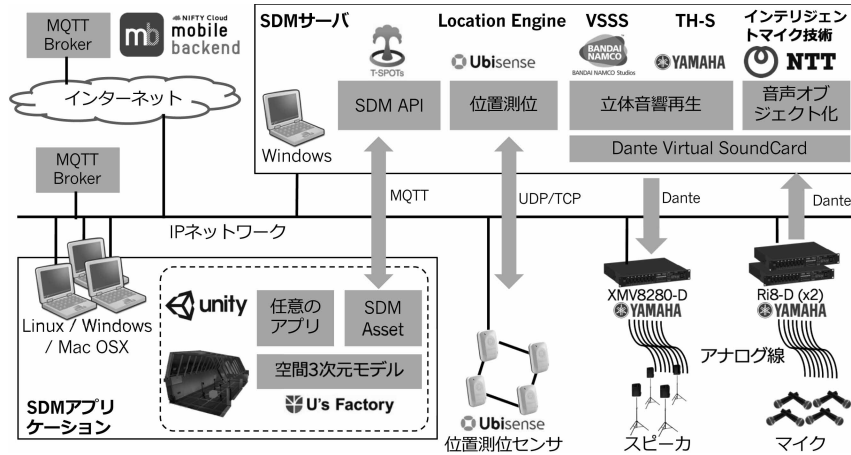


図 2 SDM 試作システムの概要

ビスを利用する任意のアプリケーションが動作する層である。この層にローカルレンダラと呼ぶアプリケーションがあり、インターネット上に存在するリアルタイムとアーカイブのコンテンツを取得してSDMサービスと組み合わせてSDM拠点の視聴空間に対して視聴者の要望に合わせた映像音声を作り出す。レンダリングの機能は、ローカルだけではなく既存の放送システムやインターネット経由で取得したコンテンツを共同編集するなどの機能を持つパブリックレンダラにより前処理される。また、アプリケーション層に定義されるメディアエンコーダによりSDM拠点を、他のSDM拠点からは放送局として機能させる。SDM拠点からの放送はリアルタイムに他のSDM拠点に転送されるだけでなく、視聴オブジェクトの蓄積・管理などのコンテンツ管理を行う。SDMでは、放送局から送信されるリアルタイムとアーカイブのコンテンツに対し、標準化されたメタデータや、タグを付与することで、ソフトウェアによって柔軟なコンテンツ管理や、状況に応じたコンテンツ選択を制御する。メディアデータとコンテンツ情報、位置情報などのメタデータを持つ視聴オブジェクトを基本単位として扱う。

5 SDM 試作システムの概要

SDM コンソーシアムでは東京大学 I-REF 棟 6 階の 65 名収容可能な大教室であるヒロビーにて、SDM アーキテクチャに基づく SDM 拠点の試作システムを

構築している。本試作システムは建物の入居者や管理者に空調・照明の遠隔制御やエネルギーの見える化など、付加価値の高いサービス提供を行うスマートビルディングのためのシステム [18] との相互運用を目指して開発している。したがって、アプリケーション環境や利用する通信プロトコルなどを共用するが、本章では映像音声に関する SDM 試作システムを抜き出し、その概要を図 2 に示す。

SDM 試作システムでは、I-REF 棟の LAN に SDM サーバ、SDM アプリケーションを動作する PC、位置測位センサ、XMV8280-D、Ri8-D、MQTT ブローカを接続するネットワーク構成を取っている。MQTT ブローカは LAN に接続したものの他、クラウド上に構築した構成も利用する。SDM サーバは Windows で動作しており、位置測位サービス、立体音響再生サービス、音声オブジェクト化サービスを提供する。試作システムは、SDM アプリケーションの動作するクライアントからの SDM API を通じた要求により、サービスを提供する。これにより複数のアプリケーションからの映像・音声の入出力要求に対し柔軟に対応できる設備機能の仮想化環境が提供できる。現在 SDM アプリケーションから SDM API を通じて利用可能な状態で試作システムへ統合されているサービスは、立体音響再生サービスと、室内位置測位サービスであり、音声オブジェクト化サービスの SDM API を通じた利用は統合の作業中である。また、現在の試



図3 Unity 開発画面とヒロビーの空間三次元モデル

作システムでは、映像オブジェクト化サービスは提供していない。

試作システムでは、各サービスのソフトウェアは、単一のサーバにインストールされる構成であるが、個別のサーバを運用することも可能である。

以下、アプリケーション環境と SDM API については 5.2 節で解説する。続けて、5.3 節 5.4 節、5.5 節では、それぞれ室内位置測位、立体音響再生、音声オブジェクト化の SDM サーバで提供されるサービスについて解説する。

5.1 アプリケーション環境

SDM アプリケーションは、Windows / MAC OSX / Linux で動作する Unity^{†2} を用いて開発する。SDM コンソーシアムでは、SDM API を用いた Unity アプリケーションを開発しやすくするため、API に接続する部分をサンプルコードとして、Unity Asset として提供している。また、I-REF 棟ヒロビーの空間三次元モデルは、U's Factory 社の Info360^{†3} により作成した。図 3 にヒロビーの三次元モデルを取り込んだ Unity 開発画面を示す。

アプリケーションは MQTT の通信を用いて SDM API にアクセスする(図 4)。これにより、サーバとアプリケーションは各々の対象とするサービスに関するメッセージを双方向に 1 対多で送信できる。現在、Unity から利用可能な SDM API で利用可能な

†2 <https://unity3d.com/>

†3 <http://us-factory.jp/robot/>

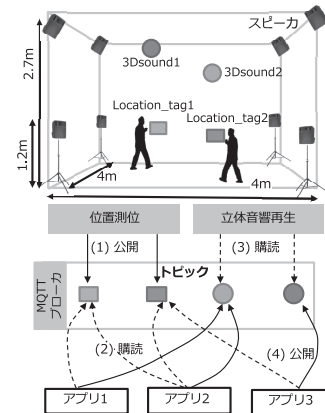


図4 ハードウェア構成と SDM サービスへのアクセス

表1 MQTT トピック命名規則

属性	階層	説明
Who	1	サービス事業者名
	2	敷地名・サイト名
	3	棟・建物名
	4	階数
	5	分割エリア・フロア名・教室名
Where	6	サービスカテゴリ (SDM や空調、照明 など)
	7	種別 (立体音響や位置測位など)
	8	型式・機器メーカー
	9	トピック内で一意の識別名称
	10	パラメータ名 (音源操作、温度、湿度など)
入出力	11	(R または W)

サービスは、Ubisense を利用した位置測位サービスと、バンダイナムコスタジオの VSSS またはヤマハの TH-S を利用した立体音響再生サービスである。

5.2 SDM API へのアクセス

情報空間で管理される視聴オブジェクトは MQTT トピックに関連づけられる。MQTT トピックは表 1 に示す固定の 11 階層の構造を採用した。第 1 階層は、誰が (Who) を示す階層としサービス事業者名を指定する。次の第 2 から第 5 階層までを、どこで (Where) を示す広域から狭域へと絞り込む構造とした。第 6 から第 10 階層までを入出力対象 (What) を示すものとし、最後尾の第 11 階層目で、入出力を示す R または W のパラメータを付与する。命名が困難な階層についてはハイフン「-」を用いる。

位置測位サービスで管理される無線タグの三次元位置を取得したいアプリケーションは、無

線タグの位置に関連付けられた MQTT トピック [3dbcs.biz/UTokyo/iREF/6F/Hilobby/SDM/Ubisense/id/Person01/Position/R] を購読することで、リアルタイムにタグの位置の配信を受けることができる。また、三次元空間の特定の位置に特定の音を再生したいアプリケーションは、その音に関連付けられたトピック [3dbcs.biz/UTokyo/iREF/6F/Hilobby/SDM/midi/sc/VSSS01/controlChange/W] に三次元空間などの情報を公開することにより音の制御を行う。これらの MQTT トピックを購読するアプリケーションは、1 対多でメッセージを受信できるため、他のアプリケーションとの連携が容易である。

MQTT トピックに公開するメッセージの詳細については、次に 5.3 節と 5.4 節でそれぞれのサービスとともに解説する。

5.3 室内位置測位サービス

位置測位サービスでは、Ubisense 社^{†4}の UWB(超広帯域)無線を利用したリアルタイム位置情報システムを利用する。このシステムは、8.5GHz ~ 9.5GHz の周波数 (1GHz 幅) を用いた無線タグとヒロビーの高所 4 カ所に設置したセンサを使って、タグの正確な位置をリアルタイムで測定する。センサ間は高精度で時間同期がなされており、無線タグからの電波を最低 2 台のセンサーが受信するとそれぞれのセンサーでの電波の受信角度を計算し、受信時間の差を割り出すことで、タグの位置を 15 ~ 30cm の精度で測定する。図 2 に示す通り、Ubisense 社の位置測位サービスと位置測位センサとの通信は、同社の提供する Location Engine による UDP/TCP で行われる。

位置測位サービスでは、ヒロビーに存在する 5 つの Ubisense の無線タグ位置を取得し、その三次元位置を SDM サーバから MQTT のタグごとのトピックに公開する (図 4 の (1))。この時の MQTT メッセージのペイロードは、部屋の基準点からの絶対座標を縦 (x)、横 (y)、高さ (z) をメートル単位で記述し、最後に時間 (time) を含める。例えば、

[x=2.8184&y=6.1839&z=0.564161&time=10/18/2016 2:14:95 PM] というメッセージを発行する。アプリケーションでは、取得したい無線タグの位置に関連付けられた MQTT トピックを購読することで、必要なタグの三次元位置をリアルタイムに取得できる (図 4 の (2))。この際、無線タグの位置の他に、タグに付属するボタンが押されたイベントなどを取得可能である。SDM サーバはボタン押下イベントを、対応する MQTT トピックに公開することで、そのトピックを購読するアプリケーションは、そのタグのボタン押下イベントを利用できる。

5.4 立体音響再生サービス

立体音響再生サービスにはバンダイナムコスタジオの開発する Virtual SoundScape System (VSSS) とヤマハの開発する Theater Surround (TH-S) を利用する。これらのサービスは、サーバにあらかじめ登録された複数の音源の三次元位置をアプリケーションから操作することが可能である。サービス起動時には登録された音源に関連付けられた MQTT トピックを購読し、アプリケーションからの操作を受け付ける (図 4 の (3))。この時、音源ごとに MQTT トピックが設定され、アプリケーションはそれぞれのトピックにメッセージを公開することにより、音源の音量、三次元座標指定、ピッチ指定、再生開始・終了を操作することができる (図 4 の (4))。VSSS のある音源を左右に動かす場合は、設置されたスピーカの中心を基準に左右 10 メートルを 128 等分した 0 から 127 の値を value 指定し、[value=127&channel=1&control=11&datetime=10/18/2016 2:14:95 PM] のようなペイロードを付与して関連づけられた MQTT トピックに公開する。channel は音源番号、control は音源移動の方向を表す。音源種類ごとに MQTT トピックが分かれていることにより、どのアプリケーションからでも関心のある音源種類に関連付けられた MQTT トピックを購読することで、その音の状態を取得することができる。

VSSS と TH-S は、登録された音源の操作 (音源の音量、三次元座標指定、ピッチ指定、再生開始・終了) を MQTT メッセージで受け、その操作に基づいて各スピーカをどのように駆動させるか計算する。計算

^{†4} <http://ubisense.net/>

結果の音声情報は、Dante™規格で XMV8280-D へ送信され、アナログ入力に変換されたものが各スピーカの入力となり音声再生される。Dante 規格の送受信のため、SDM サーバには Audinate 社の Dante Virtual SoundCard をインストールしている (図 2)。現在、図 4 上部に示す通り、4 メートル四方のエリアの 4 隅に高さ 1.2 メートルと 2.7 メートルのスピーカを配置した構成を I-REF 棟に構築している。

以下に、VSSS と TH-S の概要を示す。

- **Virtual SoundScape System(VSSS)**

VSSS とは、バンダイナムコスタジオの開発する、最先端の音響演出を再現・体感できるイベント及び施設向けの音響制御システムである。本システムは、バンダイナムコグループのゲーム製品で実際に使用されている音響制御のライブラリを応用した Windows 上で動作するアプリケーションである。ゲームに利用されるため、プレイヤーの操作に応じて即時に反応するためのインタラクティブな技術を提供する。

- **Theater Surround(TH-S)**

TH-S とは、ヤマハの開発する任意音源の Panning アルゴリズムを組み込んだ、立体音響システムである。本システムは、ヤマハ株式会社が開発された劇場向け音像移動処理を核とした Windows 上で動作するアプリケーションである。演劇やミュージカルの演出における使用を主な想定アプリケーションとして開発されているため、任意の入力音源をデザイナーの操作に応じて 3D 面内の任意の位置に定位させる技術を提供する。

5.5 音声オブジェクト化サービス

音源オブジェクト化サービスには、NTT メディアインテリジェンス研究所が開発した遠くにある狙った音だけをクリアに集音する技術であるインテリジェントマイク技術を利用する。

まず、前節の立体音響再生サービスとは逆方向に、複数のマイクからのアナログ入力はヤマハ Ri8-D にて Dante 規格へと変換され SDM サーバに送られる。インテリジェントマイク技術では、音源を分離して収録するために、複数のマイクを用いて集音すべき

音の性質を明らかにし [11]、その性質を満たすような受音系を構築する [35]。多マイクロホンで観測した信号に対して、ビームフォーミングやウィナーフィルタリングといった信号処理をすることで [36]、ユーザが狙った位置にある音をクリアに集音することを実現している。インテリジェントマイク技術により分離されて収録された音は、Dante 規格でヤマハ XMV8280-D へ送られ、アナログ信号に変換されて視聴者のヘッドセットなどで再生を行う。

音声オブジェクト化サービスは、現在 MQTT で SDM API へアクセスすることで試作システムで利用可能になるようには統合されておらず、代わりに利用者は SDM サーバの GUI を直接操作することで集音対象のエリアを指定する。今後、SDM API への統合を目指し、情報空間で (試作システムにおいては Unity 上で) 仮想のマイクを動かすことによって、現実空間に存在するその場所の音を集音可能にする予定である。その際、仮想マイクの位置を制御するための MQTT トピックを割り当て、複数のアプリケーションで仮想マイクの三次元位置を共有可能にする。仮想マイク位置で集音した音声をあらかじめ決めておいたヘッドホンやスピーカで再生する場合は、Dante で送信先の機器 (ヤマハ XMV8280-D や Dante Virtual SoundCard をインストール済みの PC) を指定すれば実現できる。一方、その仮想マイクが集音した音声を複数の MQTT トピック購読者に配信する場合は、MQTT トピック購読者を Dante の宛先に一致させる仕組みが必要となり、今後の課題である。

6 SDM アプリケーション

SDM コンソーシアムでは Interop 2015 Tokyo にて SDM 試作システムを用いてデモンストレーション展示を行い、試作システムを提供したハッカソンを複数開催した。また、試作システムをタンジブルインターフェースに利用する研究を行った。さらに、アプリケーション非依存な視聴オブジェクトの管理の研究のために、オーケストラの収録を行いデータベースを構築した。以下に試作システムを利用した SDM アプリケーションとアプリケーション非依存な視聴オブジェクトの管理に関して解説する。

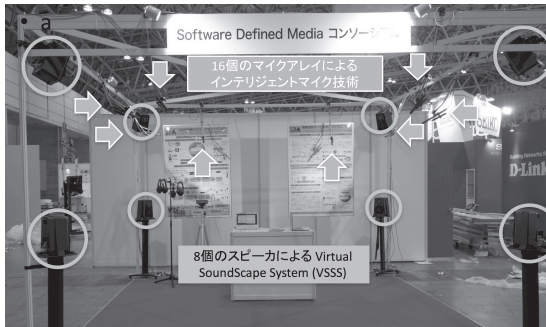


図5 Interop 2015 での展示

6.1 試作システムのデモンストレーション展示

2015年6月10日から12日にかけて幕張メッセで開催された Interop 2015 Tokyo にて、Software Defined Media コンソーシアムとして出展した。Interop 2015 への参加者は、主催者によると延べ136,341人であり、参加者は情報システム、ネットワークエンジニアリング、セールスと研究を職業とする者の割合が多かった。

出展ブースの様子を図5に示す。任意の領域で発生する音をソフトウェアで指定して収録できるインテリジェントマイク技術と、オブジェクトベースの立体音響をインタラクティブに制御する VSSS 技術を展示した。Interop 2015 で展示したデモシステムは5章で述べた試作システムのサブセットであり、MQTT による SDM API アクセスの代わりに、来場者が直接 SDM サーバの GUI を操作する。

インテリジェントマイク技術のデモンストレーションでは、図5に示すブースに、およそ2.5メートル上空のブースフレームに計16チャンネルのマイクアレイを設置した。その16のマイクからのアナログ入力は2台の Ri8-D により Dante 規格へ変換され IP ネットワークを通じて、インテリジェントマイク技術がインストールされた SDM サーバに送られる。訪問した見学者は、この SDM サーバの GUI を操作し、希望する区画を選択すると、インテリジェントマイク技術によりその区画の音だけが分離されて収録される。収録された音は、再び IP ネットワークを経由して Dante 規格でヤマハ XMV8280-D へ送られ、アナログ信号に変換されて視聴者のヘッドセットで再生される。こ

れにより来場者は指定した区画の音だけを聞くことができる。

VSSS 技術の展示では、図5に示すブースの4隅には床からそれぞれ1.2メートル地点と2.7メートル地点に計8個のスピーカーが設置されている。ブースへの訪問者は SDM サーバに MIDI ケーブルで繋がる iPad を操作して、音源の選択や、その音源の三次元座標を入力すると、SDM サーバで8つのスピーカーから鳴るべき音を計算し、その結果は IP ネットワークを経由して Dante 規格でヤマハ XMV8280-D へ送られ、アナログ信号に変換されて実際のスピーカーから、インタラクティブに立体音響が再生されるデモンストレーションを展示した。

展示した両技術は、入力機器からのアナログの入力チャンネルを一旦デジタルに変換し IP ネットワークを通じて別の機器に送信し、ソフトウェア処理を施し、計算結果を再び IP ネットワークを通じて送信し終端でアナログに変換する構成になっている。これは、来場者に3.1節で示した三次元の映像・音声の演出をソフトウェアで制御するという SDM のコンセプトの理解に役に立った。また、VSSS 技術の展示では、来場者はコンピュータで設計された臨場感のある立体音響を聴取することで、3.3節で示した実際の映像音声を録音したコンテンツに対して、拡張的な演出を組み合わせる可能性について体感することができた。

6.2 試作システムを利用したハッカソンの開催

2015年7月18日、19日の週末と2015年12月12日、13日の週末に、東京大学 I-REF 棟のヒロビーにおいて、それぞれ第2回、第3回のスマートライフハッカソンを開催し、5章で述べた SDM 試作システムを公開し、参加者に SDM API を利用したアプリケーションの開発をしてもらった。これにより、SDM プラットフォームへの要望を洗い出す一方、従来になかったような映像音声の表現や、画期的な SDM の利用方法を開拓することを目指す。スマートライフ・ハッカソンでは他に、空調や照明、ウェブカメラなどの東京大学 I-REF 棟のスマートビルディング設備を、SDM 試作システムとアプリケーションと通信プロトコルを共用し利用可能な状態で提供した。これにより



図6 SDM 試作システムを利用したハッカソンの様子

参加者はスマートビルディング設備と同様に、音声オブジェクトを制御することで、仮想化された音響機器を利用することができた。

第2回ハッカソンには28人の参加があり、13チームに分かれてアプリケーションの開発を行った。この時提供したサービスは5章で述べたSDM試作システムのうち、VSSSを利用した立体音響再生サービスのみであったが、ハッカソン当日にUnity上のSDM Assetを開発し、githubで公開し、他の参加者が利用できるようにする開発者が現れるなど、Unity上でのSDM利用が容易になった。また、ハッカソン後に行ったアンケート(回答者10名)では、全員が「大変満足」「まあまあ満足」と言う結果であった。また、ハッカソンで興味を持った技術の一つ挙げる質問では6割がSDMを挙げるなど、SDMに注目が集まった。

第3回スマートライフハッカソンの様子を図6に示す。第3回ハッカソンには19人の参加があり、9チームに分かれてアプリケーションの開発を行った。第3回では、前回のVSSSを利用した立体音響再生サービスに加え、TH-Sを利用した立体音響再生サービス、Ubisenseを利用した位置測位サービスなどを提供した。特に、今回初めて用意したUbisenseを利用した位置測位サービスが人気を集め、ほぼ全てのチームが作品に取り入れている。

ハッカソン会場には開催期間の2日間、参加者が起動するアプリケーションが同時に動作している状態が続き、各々が任意の時間に任意の音声オブジェク

トを制御していた。これにより、3.1節で示した、仮想化された音響機能を音声オブジェクトをソフトウェアで制御することで利用し、実際の音響設備から音声生成される実例を体験できた。また、これにより複数の参加者が同時に音声オブジェクトを制御することで、立体音響サービスの提供するロボットの音、犬の音、花火の音などが同時に鳴り響く時間もあり、3.2節で示した、ある視聴空間に複数のソースからの音声をミキシングする体験を提供できた。

6.3 オブジェクトベースの音響システムのためのタンジブルインターフェース

SDMコンソーシアムでは、SDM試作システムを利用したオブジェクトベースの音響システムのためのタンジブルインターフェースの研究(Tangible Sound Object) [37] [38]を行っている。

Tangible Sound Objectは一定空間内で再生されている複数の音源をタッチパネル・スクリーンや光等の視覚的なインターフェースを使用せずにユーザー自らの手で物体として移動、配置が可能なタンジブルインターフェースである。Tangible Sound Objectを使用することでユーザーは音源(音のオブジェクト)の形、サイズ、位置、距離をイメージできるようになる。Tangible Sound Objectは、SDM試作システムで提供されるオブジェクトベースの立体音響サービスと三次元での位置情報の取得が可能な室内位置測位サービスを組み合わせることにより、本来は物理的には触れることの出来ない音を物体として扱うことを可能とする。

本システムは5章で解説した試作システムのうち、5.3節の室内位置測位サービスと5.4節の立体音響再生サービスを組み合わせて、設計した。図7は、室内位置測位サービスを使って視聴空間に存在する無線タグの三次元位置が情報空間に取り込まれている様子を示す。

ユーザーは自らの手で位置情報タグ、または位置情報タグが内蔵された物理的なオブジェクトを掴んで動かし、配置することにより音のオブジェクトを三次元に配置することができる。図8はユーザーが実際に空間内の任意の位置に複数の音のオブジェクトを配置して



図 7 視聴空間の無線タグの三次元位置が情報空間に取り入れられる



図 8 再生されている音をスペース内の任意の位置に移動させる

いる様子である。各風船には位置情報タグが内蔵されている。

Tangible Sound Object の研究は、チャンネルごとの音声データを記録しておくのではなく、3.1 節で述べた通り、音源の三次元位置情報をソフトウェアで扱う SDM アーキテクチャを利用することで、実現することが容易になったアプリケーションの好例である。さらに、3.4 節で述べた、ユーザインタラクションによって聴取する音声は動的に変化するコンテンツを提供することができた。

6.4 アプリケーション非依存な 視聴オブジェクトの管理

SDM コンソーシアムでは、試作システムとそのシステムを用いたアプリケーションを開発する一方、アプリケーション非依存な視聴オブジェクトの管理の研究のために、オーケストラの収録を行いデータベース

を構築した [39]。

現在、オーケストラ収録を後に視聴するためのパッケージメディアでは通常、カメラとマイクの置かれた位置と角度によって、視聴するコンテンツが大幅な制約を受け、視聴者の意思による自由な場所と角度、距離での視聴は困難である。こうした課題を解決するため、収録されたメディアはアプリケーションに非依存に、複数の視聴オブジェクトなどに分解して解釈し、管理する機構が必要であると考えられる。

こうした機構により自由度の高いシステムが構築可能となり、汎用的な視聴オブジェクト管理によって様々なアプリケーション／コンテンツを様々な場所で提供することが可能となる。この結果、自由度の高いシステムによってより高い忠実度で音の場を収録したり、再生したり、共有したりすることが可能になるということが期待される。例えば、ある映像が、ある奏者にズームアップするとその奏者が奏する音も連動してズームアップするなどのインタラクティブなコンテンツを簡単に作成することが可能になる。

そこで、上記のような SDM を実現するデータベースを作成するため、2016 年 01 月 10 日、慶應義塾大学 日吉キャンパス内に建つ 509 席の席数を有する藤原記念ホールにて開催された慶應義塾大学 コレギウム ムジクム古楽アカデミーのコンサートを収録した。収録対象は、17 世紀のドイツ宮廷音楽であるヨハン・フリードリヒ・ファッシュの「管弦楽組曲ト長調 FaWV K:G2」など、最大 24 名の演奏家による全てアコースティックの演奏である。楽器には、テオル

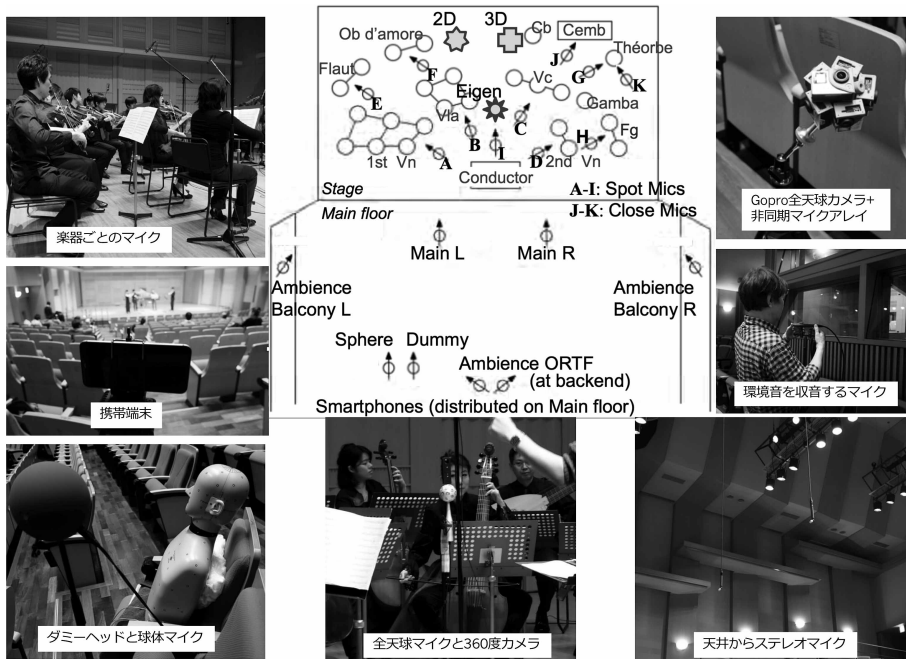


図9 オーケストラ収録のカメラとマイクの配置

べ、チェンバロ、バロック・ヴァイオリン、バロック・オーボエ、ヴィオラ・ダ・ガンバなどの現代では珍しい楽器が含まれており、楽器ごとの演奏と各楽器の合わせた音を分離して収録した。

図9にカメラとマイクの配置を示す。メインマイクロホンとして単一指向性の Schoeps MK4+CMC6 をプロセニウム (客席からみて舞台を額縁のように区切る構造物) 付近のステージバトンに左右それぞれ舞台中央から2メートルの位置に吊るした。アンビエントマイクとしては、無指向性の DPA4006 を左右のサイドバルコニーにそれぞれ1つずつ設置するとともに、主階席の最後席に ORTF マイクを設置した。これらは全て録音セッションを指示したプロの録音エンジニアの判断のもと、標準的な室内楽オーケストラ録音用のマイクとして、最適な位置を選んで配置された。舞台内にはセクションごとと同じく録音エンジニアの判断のもと標準的な個別録音として最適な位置を選び設置したが、それぞれのマイクロホンは単一指向性と双指向性の組合せとなっている。この組合せは従来から使われている M/S 型のマイク配置であるため、マイクロホンからの出力に既存のマトリクス処理

を施せばその係数を変えることでステレオ集音の拡がり感を変化させることができる。またさらに高度な処理を使うことで、側方成分を利用した主方向成分の効果的な強調といったことも期待できる。

加えて音響的な出力レベルは弱いユニークな楽器である テオルベ、チェンバロには個別集音のための近接マイクが単一指向性マイクロホンを使って設置されている。ポイントアレイマイクロホンとしては、それぞれ6個と8個の単一指向性マイクロホンを利用した、3D XYZ および 2D 360度アレイを設置すると共に、Eigenmike™を設置した。3D および 2D マイクは、それぞれ GoPro™ を6台組み合わせた三次元動画収録、および Ricoh Theta™ S を使った二次元動画収録との組合せで収録された。また客席エリア内の代表点にはダミーヘッドマイクと球体マイクも設置されている。さらに主階席4点、バルコニー2点の合計6点にはスタンド上に取り付けたスマートフォンを設置し、これにより録画と録音を行った。

収録した映像音声データはSDMによる魅力的なコンテンツ制作とアプリケーション開発に利用するため、視聴オブジェクトの概念を基礎とするデータベー

スを構築した。メディアデータとコンテンツ情報、位置情報などのメタデータを持つ視聴オブジェクトを基本単位として扱う。オブジェクト型の概念を使うことで、マイクで集音した複数の音から、音声オブジェクト化サービスによってより上位の抽出された音を生成でき、それらに同じオブジェクトの構造を使うことが可能になる。次にこれらのオブジェクトに意味情報を追加すれば、さらに上位層のオブジェクトを想定することができる。このようにこのデータベースには、下位層から意味情報を付加することで1つ上の層を生成できる再起的な構造を持たせた。さらに例えば従来型録音で典型的に使われるマイク群をひとまとめにしたさらに上位層のオブジェクトなども作成している。このデータベースはこのように階層的にオブジェクトを構成することでその意味的 (Semantic) な価値を向上させることを意図している。今回作成したデータベースでは、GoPro™, Ricoh Theta™ S, 4K ビデオカメラ、スマートフォンと4種類の映像データを位置情報とコンテンツ情報を付与して管理しているため、コンテンツやアプリケーションを作成する際、上記の音声データと組み合わせて利用することが容易である。音声データと同様、映像でも映像オブジェクト化サービスを利用し、収録対象を解釈したより上位のデータに変換し階層的にデータベースに格納することを想定するが、上位のデータの作成は今後の課題である。本データベースは、実験素材としてコンソーシアムのメンバーに提供される。

7 SDM 試作システムの性能評価

本章では SDM 試作システムの実験評価を示す。7.1 節では SDM アプリケーションの遅延を測定し、アプリケーションのリアルタイム性を検証する。また、7.2 節では、SDM 試作システムを利用したアプリケーションである Tangible Sound Object (詳細は 6.3 参照) を評価し、SDM 試作システムの有効性を考察する。

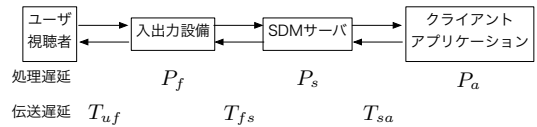


図 10 SDM 試作システムにおける遅延

7.1 リアルタイム性の試験

7.1.1 実験構成

立体音響再生サービスや、位置測位サービスなどの SDM サービスは、リアルタイム性や、インタラクティブ性を重視するアプリケーションに利用されることが想定される。SDM 試作システムの総遅延 D は図 10 に示すとおり、ユーザ、入出力設備、SDM サーバ、アプリケーションを伝わる情報の伝送遅延 T_{uf} , T_{fs} , T_{sa} と、処理遅延 P_f , P_s , P_a の総和となる (式 1)。

$$D = T_{uf} + T_{fs} + T_{sa} + P_f + P_s + P_a \quad (1)$$

T_{uf} は入出力設備とユーザ間の遅延で、音声の伝わる遅延やセンサ更新頻度などが含まれる。次に T_{fs} は入出力設備と SDM サーバ間の伝送遅延で、この間の IP ネットワークで送信されたパケットが宛先に到達する時間が含まれる。 T_{sa} は SDM サーバとアプリケーション間の遅延であり、試作システムにおいては MQTT による SDM API のアクセスにかかる時間となる。また、 P_f は入出力設備の処理遅延であり、AD/DA 変換の処理遅延が含まれる。 P_s は SDM サーバでの処理遅延であり、入出力設備による入出力を SDM の提供する抽象化されたサービスとの変換して API として提供するための処理遅延である。 P_a はクライアントアプリケーションの処理遅延である。本節では、提供する SDM サービスに依存せず、共通で発生する遅延である T_{sa} を計測するため、アプリケーションから SDM サーバまでを MQTT を利用したメッセージ配信の遅延を評価し、総遅延 D を含めた評価は次節の主観評価で行う。

一般に MQTT によるアクセスは利用するブローカのインターネット上の位置に依存する。視聴オブジェクトの送受信を行うクライアントがローカルで完結するアプリケーションでは、ブローカをローカルに置

表 2 検証システムの構成

	OS	CPU	メモリ	ソフトウェア
ローカルブローカ	Ubuntu 14.04 LTS	Intel®Core™ 2-CPU6600 2.40GHz	5.8GB	Mosquitto (ver 1.4.2 MQTT v3.1 broker)
クラウドブローカ	Ubuntu 14.04 64bit small	1vCPU	1GB	Mosquitto (ver 1.4.4 MQTT v3.1 broker)
PC	Windows 7 Enterprise	Intel®Core™ i7-3630QM CPU 2.40GHz	16.0 GB	Unity 5.3.0

く構成がもっとも遅延が小さくなる。一方、視聴オブジェクトを遠隔複数拠点に配信する場合は、拠点間の中心にブローカを置く構成が拠点間の遅延の差異がもっとも小さくなる構成となる。その中間で、あるビル内で視聴オブジェクトを送受信するアプリケーションでは、ビル設備のデータセンターにブローカを置く構成が合理的となると考えられる。したがって最適なブローカの構成はアプリケーションによって異なる。そこで、本章の評価では、遅延が最小となるブローカをローカルにおいた構成と、クラウド上に置いた構成で、往復遅延を計測したものを示す。ローカルネットワーク上のブローカ、クラウド上のブローカ、Unityを起動するPCの性能は表2のとおりである。

SDMアプリケーションとSDMサーバが同一のネットワークに存在する試作システムの構成では、アプリケーションとサーバが同一のトピックを購読している時、あるアプリケーションが公開したMQTTメッセージがブローカを経由してサーバに届く時間と、アプリケーションに戻って来る時間はほぼ等しい。これより、アプリケーションとブローカ間のメッセージ配信の往復時間を計ることで、アプリケーションからSDM APIアクセスの遅延を測ることができる。

7.1.2 実験結果

図11は、ブローカをローカルに置いた構成で、ペイロードサイズを20bytesから1420bytesまで100bytesずつ変化させて、往復遅延を計測したものである。MQTTのペイロードサイズはプロトコル仕様では256MBまで許容されるが、視聴オブジェクトの三次元位置などの軽量のメッセージを交換するSDMの特性を考慮し、MTUでフレーム分割されないサイズでの計測を行った。この時、各サイズのテストでメッセージを100回送信することで、メッセージ間隔は17msを取った。これは、Unityの平均的な描画は60フレーム/秒で行われるため、1フレー

ムに1度のSDM APIアクセスを行うアプリケーションの場合、平均17ms間隔でメッセージ送信が行われるためである。また、クライアントとブローカの一時的な計算機資源の揺らぎの影響を避けるため、MQTTメッセージは20bytesから1420bytesまでサイズのメッセージを順番に送る手順を100回繰り返した。メッセージサイズを20bytesから1420bytesまで100bytesずつ変化させたところ、多くのメッセージが0.4ミリ秒から0.75ミリ秒の範囲に収まり、メッセージサイズが大きくなるに従い、わずかに往復遅延が増えた。

図11と図12および図13の箱ひげ図はそれぞれ、最大値、第1四分位点、第3四分位点、最小値を表す。

図12はメッセージサイズを60bytesに固定して、送信間隔を1ミリ秒から17ミリ秒まで変化させた場合の往復遅延を示す。メッセージ間隔が1ミリ秒から17ミリ秒の時でも、往復遅延は0.5ミリ秒程度と変わらなかった。図12から、こうした1フレームに1度のSDM APIアクセスを行うアプリケーションが17ほど存在しても往復遅延は0.5ミリ秒で収まることわかる。また、メッセージ間隔を空けずに送信した場合を横軸の間隔0ミリ秒に表したが、平均して2ミリ秒での往復遅延に収まった。

一方、図13はブローカをクラウドに置いた構成で、メッセージサイズを変化させた時のメッセージ配信の往復時間の割合を示したものである。RTTはメッセージサイズに依存せず、平均17.5ミリ秒であった。ほとんどのメッセージは50ミリ秒以内に返答され、おおよそ80%のメッセージは30ms以内の往復遅延となったことが示された。これは、ブローカをクラウドに置いた構成では、SDM APIへのアクセスが2また3フレームのリフレッシュ時間で完了することを示す。また、各メッセージサイズによって往復遅延の違いは見られなかった。図11において観測されたミリ

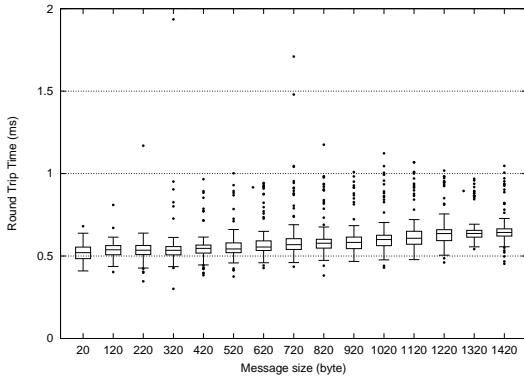


図 11 ブローカをローカルに置く構成でメッセージサイズを変化させた時の往復遅延

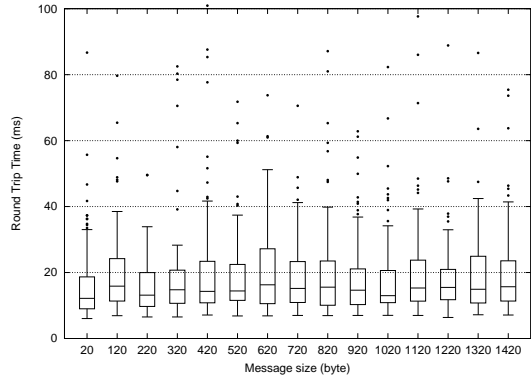


図 13 ブローカをクラウドに置いた構成でメッセージサイズを変化させた時の往復遅延

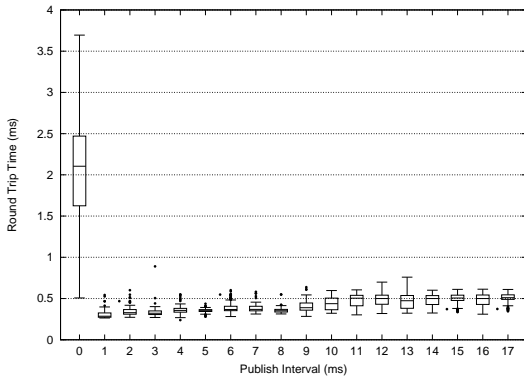


図 12 ブローカをローカルに置く構成でメッセージ間隔を変化させた時の往復遅延

秒以下の違いは誤差に埋もれて観測できなかった。また、ブローカをローカルに置いた構成と比べ、極端な外れ値が観測された。全 1500 個の MQTT メッセージのうち、37 のメッセージが 100 ミリ秒以内に往復が完了せず、2 つのメッセージは 500 ミリ秒以内の往復遅延に収まらなかった。この極端な外れ値の原因については、クラウドオペレータとの検証が必要であり、今後の課題である。

遅延のある演奏系を利用した音楽セッションにて遅延の認識に関して行われた実験によると 30 ミリ秒以内の遅延は認識されず、30 ミリ秒から 50 ミリ秒の遅延がある環境では遅延を認識するが演奏は可能、50 ミリ秒以上だと演奏に困難になることがわかっている [40]。このことから、SDM を利用したアプリケー

ションが音楽セッションのようなユーザの協調作業に利用される場合は、システム部分の遅延を 50 ミリ秒以内に抑えることが望ましい。ブローカをローカルにおいた構成では 0.5 ミリ秒で SDM API アクセスを完了したことから、この構成であれば、SDM 試作システムは多くのアプリケーションの遅延要求に応じた構成を提供できるインタラクティブ性を持っていることを確認できた。また、ブローカをクラウドに置いた場合は遅延が 30 ミリ秒であったことから、ユーザに認識可能な遅延が生まれることが予想されるため、利用用途によってブローカの位置を十分に検討する必要がある。さらにクラウドのブローカを利用した実験で確認されたように遅延のばらつきが大きい場合は、ユーザの協調作業には困難が生じることが知られており、そうした利用用途では、同期した端末間でタイムスタンプを参照することでユーザへの遅延を一定にする機構が有効であると考えられる。

7.2 SDM アプリケーション (Tangible Sound Object) の主観評価

7.2.1 実験構成

SDM アプリケーションの検証では、6.3 節に示した SDM 試作システムを利用した Tangible Sound Object の評価を行った。被験者には無線タグに関連づけられた複数のサンプル音源を再生して聴いてもらい、音声オブジェクトの音質、位置、距離について評価してもらった。また、Tangible Sound Object

表 3 音源の種類

No.	音声	三次元音響表現
V1	蚊	○
V2	犬	地面のみ
V3	ロボット	地面のみ
V4	花火	○
T5	歌声	○
T6	ピアノ	○

の目的である音声に触る（設置したり動かしたりする）という体験をどの程度達成できたのか、調査を行った。

本実験ではブローカをクラウドに置く構成とした。実験で使用した立体音響再生サービスの音源は表3の通りである。V1からV4はVSSSによる立体音響再生であり、T5とT6はTH-Sによるものである。V2とV3の音源以外は、高さを含む立体音響再生であり、V2とV3(犬とロボットの足音)は、それらが地面を歩きまわる音が聞こえる仕様になっているため、高さ方向の表現が無効化されている。

被験者は一定のスピーカーに囲まれたスペースの中心の椅子に座り、音響システムから再生された音を聴いた。全ての音源は二度以上再生され、一度目は一定の位置で、二度目はある位置からある位置へ動く音とした。音はランダムな順番・位置で再生し、音源の選択・配置・移動にはTangible Sound Objectを使用した。その後、無線タグを持って歩いたり、任意の場所に置いてもらったりして、音に触る体験をもらった。

Tangible Sound Objectを体験してもらった後、聞き取りを行った。評価項目は以下のとおりである。評価は各項目に対し10段階(1=悪い・全くわからなかった, 10=良い・はっきりと分かった)で答えてもらった。

1. 音声の聞こえ方はどうだったか？(音質, ノイズ, 途切れなど)
2. 音声オブジェクトを動かした際、遅延はあったか？
3. 位置の精度はどうだったか？
4. 音声の動きを追えたか？
5. 音声を操作する際、音声オブジェクトとして捉

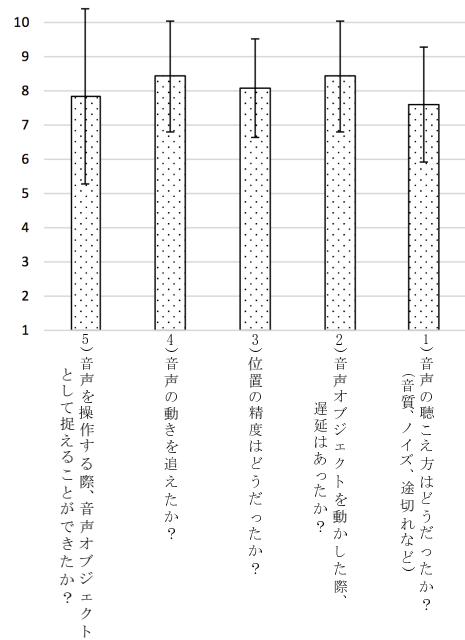


図 14 実験評価の結果(棒グラフは回答の平均値, 線は標準偏差を示す)

えることができたか？

7.2.2 評価場所と被験者

評価実験は東京大学のI-REF棟6階のヒロビーで行われ、男子6名、女子6名の10代から50代の12名の被験者が参加した。被験者全員が本システムに初めて触れる非エキスパートであった。

7.2.3 実験結果

結果は図14のとおりである。棒グラフは回答の平均値を示し、線は標準偏差を示す。

音質、ノイズ、途切れなどの音声の聞こえ方に関する評価項目1)では、最小値は4であったが、平均値7.6となり良好な結果が得られた。また、オブジェクト操作の遅延に関する評価項目2)では全員が5以上の評価で、平均値は8.4と高い評価を得られた。Tangible Sound Objectのユーザが感じる遅延は、ユーザが無線タグを操作しUnityアプリケーションで室内位置測位の結果を取得するのにかかる式1で求められる総遅延 D_{loc} と、Unityアプリケーションが対象の場所に立体音響再生をさせるSDM APIにアクセスし実際に音が視聴者に届くまでの D_{aud} を

加算したものである。クラウドのプロセッサを利用した際の SDM API アクセスにかかる遅延は平均 17.5 ミリ秒 (図 13) であったため、今回の評価は少なくとも平均 35 ミリ秒の遅延が存在した場合のものであり、ローカルプロセッサを使うことにより改善の余地があるものと思われる。

音源の位置に関する評価項目 3) では、全員が 6 以上で平均値 8.0 であった。音源は特に止まっている時ははっきりとどこに位置しているか判断できた。ほとんどの被験者は音源が前、後ろ、右、左、斜め前、斜め後ろにある時は判断できた。上下の表現は比較対象があれば判断できた。例えば、音が上にある場合を聞けば、音が下に降りてきた場合を認識できた。花火の音源は実際の花火も下から上に移動する高さの表現がある為、音の高さを表現するのに最も適していた。一方、再生している音声オブジェクトが視聴者の近傍に存在するとき、被験者は音声オブジェクトの相対位置を認識することが困難であった。

音源の動きに関する評価項目 4) では、最小値を 5 とする平均 8.4 の評価を得た。犬 (V2) の音源は犬の鳴き声が連続していない音源であり、音源が動いた際には追いきることが分かった。また、花火の音は連続してはいなかったが、音が下から上に長い距離を 4 秒かけて移動しており、その移動は最もはっきりと聞こえた。歌声 (T5) の音は動きを追いやすかったが、歌のフレーズが途切れるところでは、音の動きを追うことが難しかった。その他、音源で高さの表現を持つ歌声 (T5)、ピアノ (T6) は日常では高さの表現を持たないため、被験者から興味深いとのコメントがあった。

評価項目 5) は、音声を触る (設置したり動かしたりする) という体験を目指した Tangible Sound Object の目的が達成できたかを評価する項目である。本項目では、最小値 2、最大値 10、平均値 7.8 の評価を得て、被験者によって結果にばらつきがあった。満点の 10 をつけた被験者は 4 人存在し、音声を設置したり動かすという行動により、見えないものに触れる体験ができたというコメントがあった。一方、目に見えない何かを触った経験がないため、音を触る感覚というのが体験できないという意見もあった。

Tangible Sound Object のアプリケーション部分は約 200 行の Unity 上のコードであり、SDM 試作システムは、比較的容易に視聴空間をソフトウェア制御するため機能を提供できるプラットフォームであることを示している。本節の主観評価は、そうして開発されたソフトウェアは開発者の演出意図を高水準で満たせることを示した。

8 まとめ

本論文では、2014 年より開始した SDM コンソーシアムで進めてきた、視聴空間サービスのソフトウェア制御に関する研究について報告した。その中で、視聴サービスのための設備を仮想化する SDM アーキテクチャを提案し、そのアーキテクチャに基づく試作システムを実装した。また、試作システムのデモンストレーション展示、システムを利用したハッカソン開催、システムを利用したタンジブルインターフェイスの研究を進めてきた。さらに、アプリケーション非依存な視聴オブジェクトの管理を行うため、最大 24 名によるコンサートを、さまざまな方式のマイクとカメラで収録し、収録データのデータベース化を行った。試作システムを利用した実験評価では、MQTT を利用した SDM API アクセスの遅延を測定した。測定結果により、ネットワーク構成によって遅延が異なることがわかったが、多くのインタラクティブなアプリケーションに耐えうるリアルタイム性を確保することが可能であることがわかった。また、試作システムの機能を利用して開発された Tangible Sound Object を主観評価することで、SDM 試作システムが開発者の演出意図を高水準で満たせる開発プラットフォームであることが示された。

今後の課題は、まず SDM 拠点の試作システムの拡張である。5.5 節で述べた、音声オブジェクト化サービスを MQTT による SDM API アクセスへの統合するにあたっての課題を解決する。映像オブジェクト化サービスとして、KDDI 研究所の開発する自由視点映像技術の統合も検討する。次に、SDM 拠点をつなぐ SDM アーキテクチャの機能 (パブリックレンドラ、メディア処理、コンテンツ蓄積・管理) の研究開発を進め、効率的に多拠点を接続する。コンテンツを

蓄積する機能については、アプリケーション非依存的な視聴オブジェクトの管理 [39] の概念を拡張する。

謝辞 本論文の議論に参加していただいた SDM コンソーシアム・メンバーに感謝します。

参考文献

- [1] 文部科学省: 特集 2 東京オリンピック・パラリンピック競技大会が開催される 2020 年为目标とした科学技術に関する展開, 平成 26 年版科学技術白書-DATA GO JP, 2014.
- [2] Dolby Atmos© Specifications, Technical Report Issue 3, Dolby Laboratories, 2015.
- [3] AUROMAX© Next generation Immersive Sound system, 2015.
- [4] Smolic, A., Mueller, K., Merkle, P., Fehn, C., Kauff, P., Eisert, P. and Wiegand, T.: 3D Video and Free Viewpoint Video—Technologies, Applications and MPEG Standards, in *2006 IEEE International Conference on Multimedia and Expo*, July 2006, pp. 2161–2164.
- [5] Sankoh, H., Sugano, M. and Naito, S.: Dynamic camera calibration method for free-viewpoint experience in sport videos, in *Proceedings of the 20th ACM international conference on Multimedia*, ACM, 2012, pp. 1125–1128.
- [6] 金澤勝: 超高精細映像・音響システムスーパーハイビジョン, 電子情報通信学会誌, Vol. 92, No. 12(2009), pp. 1046–1049.
- [7] 濱崎公男, 火山浩一郎: 22.2 マルチチャンネル音響システム, 平成 17 年電気学会電子・情報・システム部門大会講演論文集, 2005, pp. 3–7.
- [8] Herre, J., Hilpert, J., Kuntz, A. and Plogsties, J.: MPEG-H 3D Audio —The New Standard for Coding of Immersive Spatial Audio, Vol. 9, No. 5(2015), pp. 770–779.
- [9] Bleidt, R., Borsum, A., Fuchs, H. and Weiss, S. M.: Object-Based Audio: Opportunities for Improved Listening Experience and Increased Listener Involvement, *SMPTE Motion Imaging Journal*, Vol. 124, No. 5(2015), pp. 1–13.
- [10] Johnson, D. H. and Dudgeon, D. E.: *Array signal processing: concepts and techniques*, Simon & Schuster, 1992.
- [11] Niwa, K., Hioka, Y., Furuya, K. and Haneda, Y.: Diffused Sensing for Sharp Directive Beamforming, *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 21, No. 11(2013), pp. 2346–2355.
- [12] Kanade, T., Rander, P. and Narayanan, P. J.: Virtualized Reality: Constructing Virtual Worlds from Real Scenes, *IEEE Multimedia, Immersive Telepresence*, Vol. 4, No. 1(1997), pp. 34–47.
- [13] 三功浩嗣, 石川彰夫, 内藤整, 酒澤茂之: 被写体 3 次元形状モデル投影型背景分離方式, 映像情報メディア学会誌, Vol. 64, No. 11(2010), pp. 1685–1697.
- [14] Koyama, T., Kitahara, I. and Ohta, Y.: Live mixed-reality 3D video in soccer stadium, in *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality*, 2003, pp. 178–186.
- [15] Iwase, S. and Saito, H.: Parallel tracking of all soccer players by integrating detected positions in multiple view images, in *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, Vol. 4, 2004, pp. 751–754.
- [16] ISO/IEC/IEEE International Standard – Information technology – Ubiquitous green community control network – Security, *ISO/IEC/IEEE 18883 First Edition 2016-04-15*, 2016, pp. 1–35.
- [17] Information technology – Message Queuing Telemetry Transport (MQTT) v3.1.1, 2016.
- [18] 粕谷貴司, 近藤正芳, 茂手木直也, 松岡康友, 矢野雅, 秋山貴紀, 境野哲, 貞田洋明, 堀越崇, 畠山英之: スマートシティのための MQTT プラットフォームの検証, 電子情報通信学会技術研究報告. ASN, 知的環境とセンサネットワーク, Vol. 114, No. 65(2014), pp. 45–46.
- [19] 金順暎, 仲地孝之, 江村暁, 藤井竜也, 羽田陽一: 4K マルチ映像と 6 チャンネルエコーキャンセラを用いた超高臨場遠隔コラボレーションシステム, 電子情報通信学会技術研究報告. CQ, コミュニケーションクオリティ, Vol. 112, No. 10(2012), pp. 87–92.
- [20] 藤井竜也, 藤井哲郎, 小野定康, 白川千洋, 白井大介: デジタルシネマ劇場へのライブ配信 (ODS) 技術, 電子情報通信学会基礎・境界ソサイエティ *Fundamentals Review*, Vol. 5, No. 1(2011), pp. 80–89.
- [21] Hirata, K., Harada, Y., Takada, T., Aoyagi, S., Shirai, Y., Yamashita, N., Kaji, K., Yamato, J. and Nakazawa, K.: t-Room: Next Generation Video Communication System, in *IEEE GLOBECOM 2008–2008 IEEE Global Telecommunications Conference*, 2008, pp. 1–4.
- [22] 平田圭二: 未来の電話を考える—遠隔コミュニケーションシステム t-Room (特集コミュニケーション環境の未来に向けた研究最前線), *NTT 技術ジャーナル*, Vol. 19, No. 6(2007), pp. 10–12.
- [23] Cruz-Neira, C., Sandin, D. J. and DeFanti, T. A.: Surround-screen Projection-based Virtual Reality: The Design and Implementation of the CAVE, in *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '93, New York, NY, USA, ACM, 1993, pp. 135–142.
- [24] 廣瀬通孝, 小木哲朗, 石綿昌平, 山田俊郎: 多面型全天周ディスプレイ (CABIN) の開発とその特性評価, 電子情報通信学会論文誌. D-II, 情報・システム, II-情報処理, Vol. 81, No. 5(1998), pp. 888–896.
- [25] 小木哲朗, 茅原拓朗, 加藤允文, 浅山宏, 廣瀬通孝: 没入型多面ディスプレイのためのインタラクティブ高臨場感音場提示手法, 日本バーチャルリアリティ学会論文誌, Vol. 8, No. 1(2003), pp. 75–83.
- [26] 大久保洋幸, 大谷真道, 小野一穂, 正岡顕一郎, 池沢龍, 小宮山撰, 浅山宏, 湯山一郎: CG 同期したインタラ

- クティブ音場再生システムについて, 日本バーチャルリアリティ学会論文誌, Vol. 5, No. 3(2000), pp. 965–973.
- [27] 大久保洋幸, 中山靖茂, 池永敏和, 小宮山撰: インタラクティブ 3D 映像音響再生システム, NHK 技研 R&D, Vol. '04 [NHK] 技研公開 講演・研究発表 特集号 (1)), No. 86(2004), pp. 72–79.
- [28] Ishii, H., Fletcher, H. R., Lee, J., Choo, S., Berzowska, J., Wisneski, C., Cano, C., Hernandez, A. and Bulthaup, C.: musicBottles, in *ACM SIGGRAPH 99 Conference Abstracts and Applications*, SIGGRAPH '99, New York, NY, USA, ACM, 1999, p. 174.
- [29] Ishii, H., Mazalek, A. and Lee, J.: Bottles As a Minimal Interface to Access Digital Information, in *CHI '01 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '01, New York, NY, USA, ACM, 2001, pp. 187–188.
- [30] 米澤朋子, 安村通晃, 間瀬健二: Tangible sound: 流体を用いたインタラクションによる音表現とその拡張, 第 16 回 NICOGRAPH MULTIMEDIA 論文, 2000.
- [31] Yonezawa, T. and Mase, K.: Tangible Sound: Musical Instrument Using Fluid Media, in *International Computer Music Conference (ICMC)*, 2000, pp. 551–554.
- [32] 後藤真: 人文社会系大規模データベースへの Linked Data の適用—推論による知識処理, 情報知識学会誌, Vol. 25, No. 4(2015), pp. 291–298.
- [33] Dickinson, M. E., Flenikien, A. M., Ji, X., Teboul, L., Wong, M. D., White, J. K., Meehan, T. F., Weninger, W. J., Westerberg, H., Adissu, H., et al.: High-throughput discovery of novel developmental phenotypes, *Nature*, (2016).
- [34] Raimond, Y., Abdallah, S. A., Sandler, M. B. and Giasson, F.: The Music Ontology, in *Proceedings of the International Conference on Music Information Retrieval*, 2007, pp. pp. 417–422.
- [35] 丹羽健太, 加古達也, 小林和則: 凹型反射板を用いた相互情報量最大化型マイクロホンアレイ, 日本音響学会秋季講演論文集, 2014, pp. 637–638.
- [36] Niwa, K., Hioka, Y. and Kobayashi, K.: Post-filter design for speech enhancement in various noisy environments, in *14th International Workshop on Acoustic Signal Enhancement (IWAENC)*, 2014, pp. 35–39.
- [37] Ogawa, K.: Tangible sound : a tangible interface for object-based sound systems, Master's thesis, KMD:Graduate School of Media Design, Keio University, 2016.
- [38] 小川景子, 塚田学, 江崎浩, 砂原秀樹: Tangible Sound Object: オブジェクトベースの音響システムのためのタンジブルインターフェース, マルチメディア, 分散, 協調とモバイル (DICOMO2016) シンポジウム, 2016.
- [39] Ikeda, M., Sone, T., Niwa, K., Saito, S., Tsukada, M. and Esaki, H.: New Recording Application for Software Defined Media, in *Audio Engineering Society Convention Paper, 141st AES Convention*, Los Angeles, USA, 2016.
- [40] 西堀佑, 多田幸生, 曽根卓朗: 遅延のある演奏系での遅延の認知に関する実験とその考察, 情報処理学会研究報告音楽情報科学 (MUS), Vol. 2003, No. 127(2003), pp. 37–42.



塚田 学

2005 年慶應義塾大学環境情報学部卒業。2007 年慶應義塾大学政策・メディア研究科修士取得。2007 年よりフランス・パリ国立高等鉱業学校 (Mines

ParisTech) ロボット工学センター博士課程在籍および、フランス国立情報学自動制御研究所 (INRIA) の IMARA チームにて研究員として勤務。2011 年博士号取得。現在は、東京大学大学院情報理工学系研究科の特任助教。2014 年より WIDE プロジェクトのボードメンバーおよび SDM コンソーシアム・チェア。自動車の情報化など、次世代インターネット IPv6 における移動体通信, SDM に取り組む。



小川 景子

2013 年慶應義塾大学経済学部卒業。2016 年同大学大学院メディアデザイン研究科修士課程修了。修士論文で研究科委員長表彰。グローバルイノ

ベーションデザイン (GID) プログラムにて、ロンドンのロイヤル・カレッジ・オブ・アート (RCA) とインペリアル・カレッジ・ロンドン (Imperial), そしてニューヨークのプラット・インスティテュート (Pratt) へ留学。現在は、慶應義塾大学メディアデザイン研究科研究員。



池田 雅弘

1993 年京都大学工学部建築学科卒, 95 年同大学大学院博士前期課程工学研究科建築学専攻修了, 同年ヤマハ株式会社入社。以降, 同社の研究開発部門・事業部門にて建築音響・空間音響・プラットフォーム開発関連の研究開発に従事。日本音響学会, AES, 米国音響学会, 各会員。



曾根卓朗

1982年北海道大学工学部応用物理学科卒。1982年日本楽器製造株式会社(現ヤマハ株式会社)入社。主に組込みシステム向けソフトウェア開発と研究開発業務に従事。業務用通信カラオケシステム、携帯着メロソフトウェアのシステム設計・開発等を担当。楽器・音響開発本部技術開発部所属。



丹羽健太

2006年名大・工・電気電子情報工卒、2008年同大学院修士課程了、2014年同大学院博士課程了、博士号(情報科学)取得。2008年日本電信電話(株)入社。マイクロホンアレイを用いた收音処理の研究に従事。2010年日本音響学会より栗屋潔学術奨励賞受賞。日本音響学会、電子情報通信学会、IEEE各会員。



齊藤翔一郎

2005年東京大学工学部計数工学科卒業。2007年同大学院情報理工学系研究科修士課程修了。同年NTTサイバースペース研究所(現NTTメディアインテリジェンス研究所)入社。エコーキャンセラ・マイクロホンアレイを中心とした音響技術の研究開発に従事。日本音響学会、電子情報通信学会、IEEE各会員。



粕谷貴司

2005年東京農工大学情報コミュニケーション工学科卒業。2007年慶應義塾大学大学院政策・メディア研究科修了。2008年より竹中工務店入社。ワークプレイスプロデュース本部を経て、2015年より情報エンジニアリング本部。建物設備に対する情報エンジニアリングの業務や研究に従事。



砂原秀樹

慶應義塾大学大学院メディアデザイン研究科教授。1960年兵庫県生まれ。88年慶應義塾大学理工学部博士課程修了。電気通信大学情報工学科助手。94年奈良先端科学技術大学院大学情報科学センター助教授を経て、2001年から教授。2005年情報科学研究科教授。2008年4月より現職。1984年からJUNET、1988年からWIDEプロジェクトを通じて、日本におけるインターネットの構築とその研究に従事。



江崎 浩

1987年九州大学大学院・工・電子修士課程了。同年(株)東芝入社。1990年米国ニュージャージー州ベルコア社。1994年コロンビア大学・客員研究員。1998年東京大学大型計算機センター・助教授。2001年同大学院・情報理工学系研究科・助教授。2015年同大学院・同研究科・教授。現在に至る。博士(工学、東京大学)。MPLS-JAPAN代表、IPv6普及・高度化推進協議会専務理事、WIDEプロジェクト代表、JPNIC 副理事長。