SDM360^2:音楽イベントのための自由視聴点映像音声のインタラクティブ再生

塚田 学[1]　菰原 裕[1]　新居 英明[2]　粕谷 貴司[3]　高坂 茂樹[4]　小川 景子[5]　江崎 浩[1]

概要：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解釈し，コンテンツとして活用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．2014 年より，Software Defined Media（SDM）コンソーシアムでは，オブジェクトベースの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行っている．現在，音楽イベントの DVD などのパッケージメディアは，マイクやカメラなどの収録機材の位置によって大きく制約を受けるコンテンツである．本研究では，クラシックコンサートとジャスセッションのイベントを収録し，インタラクティブに自由視聴点での三次元映像音声を再生するアプリケーション「SDM360^2」を設計，実装し，評価を行なった．さらにビルボードジャパンが開催した Live Music Hackasong にて SDM3602 のデモンストレーションを行い，審査員と一般の来場者の投票により，優秀賞を受賞した．

SDM360^2: An interactive audio-visual service for music event

Manabu Tsukada[1]　Yu Komohara[1]　Hideaki Nii[2]　Takashi Kasuya[3]　Shigeki Takasaka[4]　Keiko Ogawa[5]
Hiroshi Esaki[1]

1. はじめに

近年，スマートフォンが普及し，標準的に内蔵される映像音声の収録機能によって様々なイベントが収録される場面が増えている．それに伴い，収録されたコンテンツが，Youtube やニコニコ動画などの動画共有サービスで共有されるという，インターネットを利用して，伝達・共有・加工が行わ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をそのまま転送するだけでなく，空間に存在する収録対象を 3 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設備に合わせた柔軟な 3 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て，組み合わせることで，今までにない表現への可能性が広がる．例えば，音声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応援スタンド側からの視聴や，選手 1 人称視点からの視聴を切り替えたり，アイドルグループの中のお気に入りの人物を近くで鑑賞したりというインタラクティブなコンテンツの製作が可能になる．

このように，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加する中，著者らは 2014 年から Software Defined Media (SDM) コンソーシアム[*1]を設立し，オブジェクトベースのディジタルメディアと，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 [1], [2].

[1] 東京大学
The University of Tokyo
[2] 株式会社 IIJ イノベーションインスティテュート
IIJ Innovation Institute Inc.
[3] 株式会社 竹中工務店
Takenaka Cooperation
[4] エスイーディー 株式会社
Sophisticated Engineers Department (SED)
[5] 慶應義塾大学
Keio University

[*1] http://sdm.wide.ad.jp/

現在，オーケストラやライブなどのイベント収録を，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，本研究では，クラシックコンサートとジャスセッションのイベントを収録し，インタラクティブに自由視聴点での三次元映像音声を再生するアプリケーション SDM360^2 を開発した．

本論文では，まず第 2 章で，本研究の目的を述べる．次に，第 3 章で，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収録と，Musilogue Band ジャズセッションの収録に関する構成を説明する．第 4 章では，収録されたデータを元に開発した，インタラクティブに自由視聴点での三次元映像音声を再生するアプリケーションである SDM360^2 の設計と実装を述べる．第 5 章では，SDM360^2 の性能評価を行い，第 6 章にて，「ライブを拡張する」というテーマでお行われたビルボードハッカソンで出展したデモンストレーションの報告を行う．第 7 章で，関連研究の調査を行う．第 8 章で，本論文の結論と今後の課題について述べる．

2. 本研究の目的

本研究の目的は，音楽イベントの収録データを利用して，自由視点映像音声のインタラクティブな再生を行うことである．また，収録データを利用した様々な演出を可能にするアプリケーション開発環境を要求事項を抽出するため，そのために SDM360^2 を試作する．こうしたアプリケーション開発環境により，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で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．

図 1 は，プロ用の機器の他に，観客が持ち込む商用デバイスによって収録されたデータから三次元映像音声コンテンツを協調制作するフレームワークを模式的に示したものである．SDM360^2 と目指すアプリケーション開発環境は以下の要件を想定する．

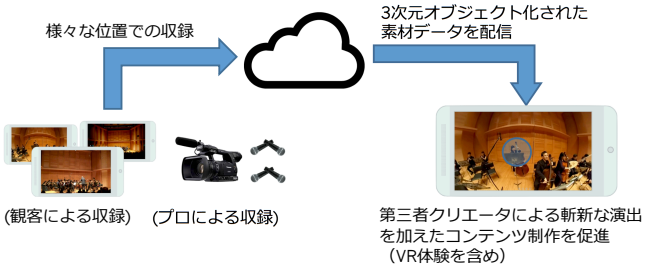


図 1 三次元映像音声コンテンツの協調制作フレームワーク

三次元の映像・音声の演出をソフトウェアで制御： 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．

ユーザ・インタラクション性： 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，ファンの歌手の一人称視点での映像音声の視聴などが考えられる．

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も可能であり，それによって視聴者の臨場感や満足度を高めることが期待される．これ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ソフトウェアレンダリングによって組み合わせた拡張演出と考えられる．例えば実際のイベント会場では，見たり聞こえたりしないような映像音声の効果を追加することで，現実拡張を行うといったことが想定できる．

映像・音声の複数ソースのミキシング可能性： 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ーターの手によるコンテンツが，第三者クリエー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．

再利用性の高いデータ公開方式： 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．本研究では SDM360^2 を開発することで，三次元の映像・音声を利用したアプリケーション開発環境の要求に応えるデータベースの設計に関する知見を得る．

3. 素材データの収録

SDM360^2 の開発や，その他の 3 次元映像音声のアプリケーション開発に用いるため，収録データ素材として，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と，Musilogue Band ジャズセッションを収録した．

3.1 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収録

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，テオルベ，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さった音を分離して収録した [3].

図 2 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 をプロセニアム（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に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が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^{TM} を設置した．3D および 2D マイクは，それぞれ GoPro^{TM} を 6 台組み合わせた三次元動画収録，および Ricoh Theta^{TM} 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．2D と 3D アレイマイク，Eigen マイク，スマートフォンはローカルのストレージに記録し，その他の全てのマ イク出力は Dante 付の Head Amp Unit を介し IP ネットワークで録音ブースに設置された PC 上の DAW に伝送された．また主階席後方に設置したカメラで 4k Video も同時に撮影している．

3.2 Musilogue Band ジャズセッションの収録

2017 年 1 月 26 日に六本木ミッドタウン内にある Billboard Live Tokyo で Musilogue Band のコンサートを収録した．リハーサル中に録音した映像，音声のコンテンツを格納し SDM360^2 のコンテンツを作成し Live Music Hackasong で発表を行った．

編成は ElectricPiano(Fender Rhodes)，Electric Bass，Drums の instrumental である．会場のリスナーへの表現手法としては全ての楽器に Mic 或いは電気信号を増幅する機材に接続し，SR(Sound Reinforcement) 用 MixingConsole で調整されたのちに大規模な Speaker で再生される．

音圧レベルは場所にもよるがおおよそ 100dBspl となる．

映像データを収録するにあたりリコー社製品の全天球カメラ「RICOH THETA S」を 8 台配置した．

音声は全て SR(Sound Reinforcement) 用 MixingConsole に纏められていて，Console の HA 分岐をデジタルで DAW に転送するシステムとなっている．録音される音量レベルは全て MixingConsole の Setting に依存する為，後に整音作業が必須となる．

回線は図 3 の通りである．

舞台内の Microphone Setting については楽器単体の他に Audience マイクを 4 本設置する．音源より近いポジションとしてステージ袖の上手，下手に 1 本ずつ，もう 2 本は天井から無指向性の DPA4090 を 2 本つりさげての Setting とした．

Near の Mic はステージの角の位置，高さ 1m くらいの所に setting する．Far についてはステージから奥に 5m，高さ 5m のところにセンターを中心に約 1m の幅で 2 本設置した．

録音したデータを編集用の DAW にて DrmsMix，BassMix，Piano(Rhose)Mix，Ambience near L，R　Ambience far L，R，それぞれ楽器ごとの単体で纏め上げた音源に編集し，アプリケーションに組み込む作業を行う．

4. SDM360^2：インタラクティブな自由視聴点での三次元映像音声の再生

SDM コンソーシアムでは 2 章で述べた本研究の目的と



図 2　オーケストラ収録のカメラとマイクの配置

INPUT PLAN				
No.	INST	MIC	STAND	Remarks
1	O/H L	C-414	ST-210	
2	O/H R	C-414	ST-210	
3	F TOM	e-904	-	
4	M TOM	e-904	-	
5	H TOM	e-904	-	
6	HH	PG81	ST-259	
7	SNr Top	SM-57	ST-259	
8	SNr Bottom	SM-57	ST-259	
9	Kick OFF	BETA52A	ST-259	
10	Kick ON	ATM-25	ST-259	
11	E.Bass Line	SM-58	ST-259	
12	E.Bass Mic	Type-85	-	
13	Rhodes L	Type-85	-	
14	Rhodes R	Type-85	-	
15	Ambience near L	4090	ST-259	
16	Ambience near R	4090	ST-259	
17	Ambience far L	4090	-	
18	Ambience far R	4090	-	

図 3　マイク回線のリスト

要求事項を基に，インタラクティブに自由視聴点での三次元映像音声を再生するアプリケーション SDM360^2 を開発した．SDM360^2 とは，映像と音声を全周 360 度で視聴できることから命名した．

現在，SDM360^2 では，収録した慶應義塾大学のクラシックコンサートと，ビルボードライブ・ジャパンで収録したジャズのセッションを視聴できる．

ユーザは，収録された演奏が行われている三次元空間をステージと観客席を含めインタラクティブに自由に移動でき，SDM360^2 はその地点で見える映像と聞こえる音声を再生する．また，演奏されている楽器や演奏パートを音声にズームインしたり，それらを自由に有効化・無効化することができる．

4.1　設計

SDM360^2 の全体の設計概要を図 4 に示し，映像音声の処理とユーザインターフェイスに分けて解説する．

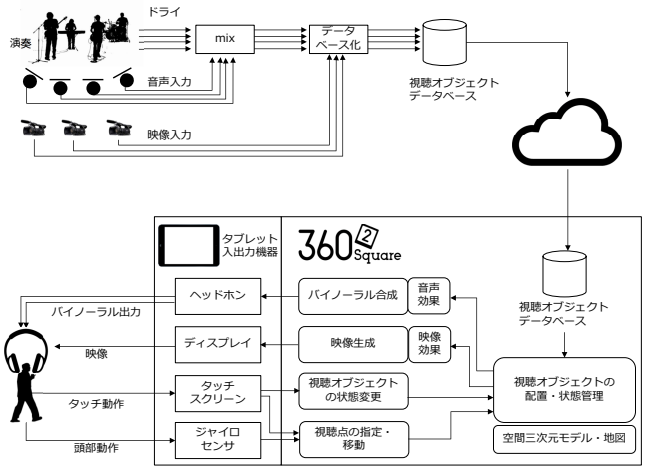


図 4　SDM360^2 の設計

4.1.1　インタラクティブ映像音声の処理

視聴オブジェクトデータベースから取り出した映像と音声のメディアは，それに付随するメタデータを元にあらか

じめ作成しておいた空間三次元モデルの上に配置される．配置された視聴オブジェクトは状態を持ち，自由に有効化・無効化できるように管理される．

また，その 3 次元モデルの中に単一の視聴点を定義して，その場所で聞こえる音をバイノーラル合成する．その際，音の減衰率やリバーブ効果をつけるなどの処理を追加可能である．また，移動速度によって音声のドップラー効果を付与することも想定する．同様に，その場所で見える映像を生成しディスプレイで再生する．映像にも移動や実際には収録されてない物などを仮想的に追加するなどの処理を追加できる．

ユーザのタッチ動作のインタラクションより，視聴オブジェクトの状態を有効化・無効化でき，さらに視聴点を自由に移動したりできる．また，頭部動作により向いた方向に視聴方向を移動できる．

4.1.2 ユーザインターフェイス

SDM360^2 は，HMD で鑑賞に応用できるが，現在は簡易なインタラクティブ性を重視してタブレットの入出力機器を想定して開発する．

ユーザは，ヘッドホンからバイノーラル出力を受け両耳で音を聴取し，ディスプレイにより映像を視認する．

一方，ユーザからのインタラクションは，図 1 に示す，スクリーンのタッチ動作で行う．視聴点自体を移動させる通常モードと，位置は固定し視線を移動させる位置固定モードによって，タッチ動作による結果が異なるが，それぞれ 1 指ドラッグ，2 指回転，2 指ピンチをサポートする．

また，視聴点の移動に関しては，画面に常に表示されるボタン UI を設置し，プリセットされた位置と視線に移動できるように設計する．

表 1 タッチ動作

動作	通常モード	位置固定モード
1 指ドラッグ	ドラッグ方向へ位置を移動	ドラッグ方向へ視線を移動
2 指回転	視線を水平方法に回転	視線を中心に回転
2 指ピンチ	ズームイン・ズームアウト	

さらに視線移動に関しては，タブレットのジャイロセンサーを使い，タブレットの向いている方向を取得して，視点移動できるモードを用意する．

4.2 実装

SDM360^2 は，Unity Version 5.5.0f3 を用いて開発した．

本アプリケーションは起動すると SDM360^2 ロゴが現れ，コンテンツ選択画面へ遷移する．現在選択可能なコンテンツは，慶應義塾大学のクラシックコンサートと，ビルボードライブ・ジャパンでのジャズのセッションである．

クラシックコンサートを選択すると図 5 に示すクリーンショットのようなヘルプ画面に遷移して，使い方を提示する．ユーザは使い方を理解したのち，「Start」ボタンを押下することにより，コンテンツ視聴が開始する．



図 5 SDM360^2 のヘルプ画面

コンテンツ視聴は，図 6 に示すクリーンショットのような画面で行われる．図 6 は図中で音符を描画した立方体で示される音声オブジェクトの配置が見やすい視聴点に移動した時の映像である．実際に複数の場所で収録された音声は，3 次元モデル上に配置され，Unity の Audio Listerner と呼ばれるコンポーネントを利用して，その聴取点で聞こえるようバイノーラル合成されている．図中 (1) で示されるように，実際には見えない音を音量に大きさの変化する白い輪によって可視化する映像効果を追加した．また，図中 (2) で示される通り音声オブジェクトにタッチすることで，その音声のパートの有効化・無効化することができ，無効化された音声オブジェクトは回転が止まり，音声可視化の白い輪が消える．図中 (3) に示す通り，各音声オブジェクトにはメタデータが付属しており，タッチによってそのメタデータが詳細表示に切り替わる．また，図中 (4) の映像オブジェクトをタッチすることで，視聴点がその映像オブジェクトの位置へ移動し，位置固定モードに切り替わる．



図 6 通常モードのスクリーンショット

位置固定モードのスクリーンショットを図 7 に示す．位置固定モードでは，Ricoh Theta S で撮影された 360 度動画を，球体内部に投影して 360 度ビューを実現している．これは，Easy Movie Texture (v3.56) が提供する MediaPlayerCtrl スクリプトによって実現している．位置固定モードにおいても，同様にタッチ動作により音声オブジェクトの有効化・無効化が可能である（図中 (5)）．また，図中 (6) に示す通り，ドラッグによって視線を全周に移動させることができる．図中 (7)〜(10) は，通常モードと位置固定モードの両方に常に表示されるボタン UI の説明である．図中 (7) のボタン群は，タッチ動作によりステージや客席などのプリセット位置へと移動するためのボタンである．図中 (8) は，音声オブジェクトを操作しやすいステージ上の視聴点へ移動するボタンであり，スクリーンショットを図 6 に示した．図中 (9)，(10) のボタンにより，ユーザはいつでも自由にヘルプ画面を参照し，コンテンツ選択画面へ遷移することができる．



図 7 位置固定モードのスクリーンショット

5. 性能評価

本節では前節で解説した実装を用いて実験による性能評価を行う．実験で用いた機器は macOS Sierra (Version 10.12.3) がインストールされた MacBook Pro (15-inch, 2016) である．プロセッサは 2.9 GHz Intel Core i7，メモリは 16 GB 2133 MHz LPDDR3 であり，グラッフィックチップは Radeon Pro 460 4096 MB と，Intel HD Graphics 530 1536 MB を搭載している．

実験は，Unity Version 5.5.0f3 のエディタ上で SDM360[2] を実行し，マウスで操作し，各イベントのタイムスタンプを記録することでイベント発生時の Unity のプロファイラの分析結果を確認することで行った．Unity のプロファイラは CPU 使用率，レンダリング，メモリ，オーディオの 4 種類の結果を報告する．大きく影響すると思われる GPU については，該当機器では情報を取得できなかったため，GPU 使用率の分析は今後の課題である．

図 8 にビルボードでのジャズセッションのコンテンツ（8 個の 360 度動画と，7 個のオーディオ音源が含まる）を表示した際の CPU 使用率，レンダリング，メモリ，オーディオの変動を示す．X 軸は，コンテンツ開始からのフレームの番号である．変動は，全ての操作イベントの発生ではなく，視聴点が移動するイベントの影響が多かったため，図 8 では，影響の大きかったイベントのみを表示した．

全体として測定した値は，それぞれ図 5〜7 で示した，ヘルプ画面，通常モード，位置固定モードの 3 つの傾向のカテゴリに別れた．図 8 のグラフでは，それぞれの代表的な値を示す位置に Y 軸の値を示している．

CPU 利用率では，全体として概ね 30 フレーム/秒を維持していることがわかる．また，ヘルプ画面表示とイベント発生時は，最も処理が遅くなっているが，それ以外の時間には 60 フレーム/秒の性能を維持できている．全体として最も CPU 時間を消費しているプロセスは，360 度動画を球体内面に貼り付ける MediaPlayerCtrl スクリプトであり，8 つの 360 度動画の貼り付けを行うための CPU 時間が全体の 66.0%を占めている．図 8 では，MediaPlayerCtrl スクリプトの処理が大半を占めるスクリプト処理が，ヘルプ画面，通常モード，位置固定モードの代表的な値それぞれで，11.4ms，11ms，9.6ms の時間がかかっている．ヘルプ画面では，その他，ステージの全景を描画するための時間が，5.9ms かかっており，フレームレートが落ちる原因となっている．

レンダリングでは，最終的な DrawCall 数を意味する Batches 数，マテリアルへの DrawCall 数を意味する SetPass calls 数，三角メッシュ数，頂点数を示している．表 2 に示すとおり，全ての値が，ヘルプ画面，通常モード，位置固定モードの順に減少していることがわかる．

表 2 各モードにおけるレンダリングの数

	ヘルプ画面	通常モード	位置固定モード
Batches 数	594	294	110
SetPass calls 数	421	205	80
三角メッシュ数	348.2k	174.3k	29.0k
頂点数	219.5k	110.2k	27.7k

メモリ使用量は，全体を通して小刻みな増減を繰り返すものの，表示しているモードによる違いは見られなかった．メモリ使用量は，実験を通して全体で 2.5GB〜2.7GB ほどであった．内訳としては，ネイティブコードへのメモリ割り当て量（Unity）が 420MB， 全体のヒープサイズとマネージドコードが使用しているヒープサイズ（Mono）が 1.94GB ほどであった．また，ドライバが Texture，レンダリングのターゲット，Shader，Mesh データに使用している推定メモリ量（GfxDriver）が，約 100MB，プロファ

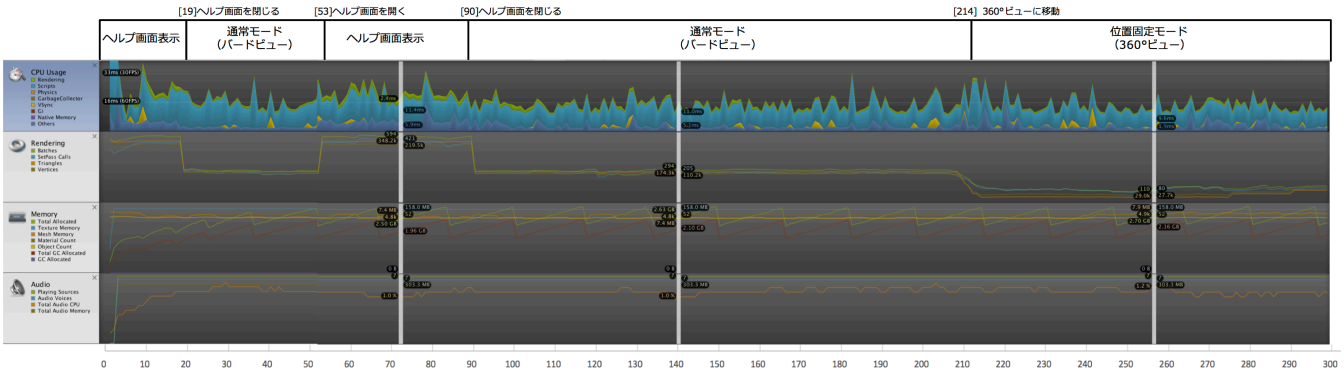


図 8 SDM360^2 の性能評価

イラーデータに使用されているメモリ量（Profiler）が約 20MB であった．

オーディオは，実験を通してほぼ変動しなかった．実験を通じて 7 つの音源が再生されていた．これは，ユーザ操作によって無効化された音源も，有効化した時に演奏のずれがないように停止することなくミュートしていたためである．オーディオドライバが使用している推定メモリ量（FMOD）が約 300MB，CPU 利用率は約 1%であった．

本節では，SDM360^2 を動作させた Macbook Pro での Unity の CPU 使用量，メモリ使用量，レンダリング，オーディオの統計を調べた．360 度動画の映像 8 つ，オーディオ 7 つのコンテンツをインタラクティブに自由視聴点において，問題なく再生を行うことができた．

6. SDM360^2 のデモンストレーション

2017 年 1 月 26 日にビルボードライブ東京で「ライブ体験の拡張」をテーマとして行われたハッカソン【LIVE MUSIC HACKASONG】にて SDM360^2 のデモンストレーションを行った．審査員と約 100 名の一般の来場者は，Musilogue Band の実際のジャズ演奏を聴いてから，事前にリハーサル収録してあったデータから，SDM360^2 で視聴できるというデモンストレーションを行った．また，本イベントは，イベント運営に当たったビルボードジャパンによって LINE live を利用して，スマートフォン及び PC でライブ視聴できるように放送されており，のべ 3000 人を超える視聴があった．また，審査員には iPad Pro 及び iPad air2 を渡して手元で鑑賞できるようにしたほか，一般の来場者には Unity から出力されるステレオ音声出力を舞台の左右に設置されているスピーカから放送し，スクリーンには Unity の Play 画面を表示することで鑑賞できるようにした．

また，最終審査では SDM360^2 が自由な視聴点から聞きたい音にフォーカスできる点が評価され，審査員と一般の来場者からなる投票で，10 組の中から優秀賞をを受賞した．デモンストレーション，及び審査・受賞の様子は，様々なメディア[*2] [*3] [*4] で掲載された．

7. 関連研究

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 [4]．さらに，同様の技術を利用してディジタルシネマ劇場への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 [5]．

遠隔のユーザ同士が快適に会話し，効率的に協調作業するためのコミュニュ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」のを開発した [6][7]．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[8] や多面型全天周ディスプレイ (CABIN)[9] 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 [10]．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

*2 http://www.musicman-net.com/business/64896.html
*3 http://www.billboard-japan.com/d_news/detail/46826/2
*4 http://www.musicman-net.com/business/64896.html

レイを用いて再生するシステム [11] 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 [12].

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，人文社会系大規模データベース [13] や，ノックアウトマウスの表現型のデータベース [14] 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[15] 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．

8. まとめと今後の課題

本論文では，2014 年より，SDM コンソーシアムで取り組んで来た視聴空間サービスのソフトウェア制御に関する研究の中でも，音楽イベントのための自由視聴点映像音声のインタラクティブな再生を行う SDM360^2 に関して報告した．まず，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と Musilogue Band ジャズセッションを収録し，その収録データを利用した 3 次元映像音声を利用した様々な演出を可能とするアプリケーション開発環境の要求事項を考察した．要求事項に基づき SDM360^2 を設計，実装し，性能を評価した．さらにビルボードジャパンが開催した Live Music Hackasong にて SDM3602 のデモンストレーションを行い，審査員と一般の来場者の投票により，優秀賞を受賞した．

今後の課題として，本論文ではタブレットを対象デバイスとして選択したが，HMD などの没入感を高めた視聴サービスを検討する．この場合には，適したユーザインターフェイスを開発する必要がある．また，3 次元映像音声を利用した様々な演出を可能とするアプリケーション開発環境として，収録データを高い再利用性を持った方式で公開・共有する必要がある．現在，オープンデータのメタデータ記述手法 RDF を利用した LOD として収録データを公開，共有することを検討している [16].

参考文献

[1] 塚田学, 小川景子, 池田雅弘, 曽根卓朗, 丹羽健太, 齊藤翔一郎, 粕谷貴司, 砂原秀樹, 江崎浩. Software Defined Media: 視聴空間サービスのソフトウェア制御. 日本ソフトウェア科学会学会誌『コンピュータソフトウェア』「ネットワーク技術」特集, August 2017.

[2] Manabu Tsukada, Keiko Ogawa, Masahiro Ikeda, Takuro Sone, Kenta Niwa, Shoichiro Saito, Takashi Kasuya, Hideki Sunahara, and Hiroshi Esaki. Software Defined Media: Virtualization of Audio-Visual Services. *IEEE International Conference on Communications (ICC2017)*, May 2017. Paris, France.

[3] Masahiro Ikeda, Takuro Sone, Kenta Niwa, Shoichiro Saito, Manabu Tsukada, and Hiroshi Esaki. New recording application for software defined media. In *Audio Engineering Society Convention Paper, 141st AES Convention*, Los Angeles, USA, September 2016.

[4] 金順映, 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, pp. 87–92, apr 2012.

[5] 藤井竜也, 藤井哲郎, 小野定康, 白川千洋, 白井大介. ディ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, pp. 80–89, 2011.

[6] K. Hirata, Y. Harada, T. Takada, S. Aoyagi, Y. Shirai, N. Yamashita, K. Kaji, J. Yamato, and K. Nakazawa. t-room: Next generation video communication system. In *IEEE GLOBECOM 2008 - 2008 IEEE Global Telecommunications Conference*, pp. 1–4, November 2008.

[7] 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-ナル, Vol. 19, No. 6, pp. 10–12, 2007.

[8] Carolina Cruz-Neira, Daniel J. Sandin, and Thomas A. DeFanti. 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, pp. 135–142, New York, NY, USA, 1993. ACM.

[9] 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, pp. 888–896, may 1998.

[10] 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, pp. 75–83, March 2003.

[11] 大久保洋幸, 大谷眞道, 小野一穂, 正岡顕一郎, 池沢龍, 小宮山摂, 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, pp. 965–973, sep 2000.

[12] 大久保洋幸, 中山靖茂, 池永敏和, 小宮山摂. インタラクティブ 3D 映像音響再生システム. NHK 技研 R&D, ('04〔NHK〕技研公開 講演・研究発表 特集号 (1)), No. 86, pp. 72–79, jul 2004.

[13] 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, pp. 291–298, 2015.

[14] Mary E Dickinson, Ann M Flenniken, Xiao Ji, Lydia Teboul, Michael D Wong, Jacqueline K White, Terrence F Meehan, Wolfgang J Weninger, Henrik Westerberg, Hibret Adissu, et al. High-throughput discovery of novel developmental phenotypes. *Nature*, 2016.

[15] Yves Raimond, Samer A Abdallah, Mark B Sandler, and Frederick Giasson. The Music Ontology. *Proceedings of the International Conference on Music Information Retrieval*, pp. pp. 417–422, 2007.

[16] 菰原裕, 塚田学, 江崎浩, 曽根卓朗, 池田雅弘, 高坂茂樹, 新麗, 新善文. SDM Ontology: Software Defined Media のメタデータ管理のための Ontology. マルチメディア，分散，協調とモバイル (DICOMO2017) シンポジウム, June 2017.