Web360^2: インタラクティブな3D視聴体験を提供するWebアプリケーション

加藤 慎[1,a)]　池田 友洋　川守田 光昭[2]　塚田 学[1]　江崎 浩[1]

概要：インターネット上の動画配信サービスが拡大し続けており，全世界のダウンストリームトラフィックシェアのうち動画ストリーミングのトラフィックシェアが過半数を占めている．VR サービスの普及や民生用の 360° カメラの登場などにより 3D コンテンツの視聴体験への敷居が下がり，3D コンテンツサービスが拡がり始めている．近年，3D コンテンツサービスを Web ブラウザ上で体験するための技術が進み，専用のアプリケーションをインストールする必要もなく，Web 上で手軽にサービスを楽しむことができるようになってきた．本研究では，Web 上で 360° 動画およびオブジェクトベースオーディオによる立体音響を再生し，視聴者にインタラクティブな 3D コンテンツの視聴体験を提供する Web アプリケーション「Web360^2」を設計，実装し，さらに評価を行った．

キーワード：3D コンテンツ，A-Frame，WebVR，Web Audio，オブジェクトベースオーディオ，Software Defined Media

Web360^2: An Interactive 3D Audio-visual Service with Web

1. はじめに

近年，インターネット上の動画配信サービスが急激に拡大している．Sandvine が公開した 2018 年における全世界のインターネットトラフィックの統計レポート [1] によると，動画ストリーミングトラフィックがダウンストリームトラフィックの 57.69%を占めている．同社の 2019 年におけるモバイル端末インターネットトラフィックの統計レポート [2] では，ダウンストリームトラフィックのうち YouTube が 37.04%を占めており，アプリケーションのトラフィックシェアにおいて他を大きく引き離して 1 位を獲得している．そのほか動画ストリーミングトラフィック関連のアプリケーションでは，Facebook Video (2.53%)，Netflix (2.44%) がトラフィックシェア上位 10 のアプリケーションに含まれている．また，VR の登場にともなって 360° 動画配信サービスも普及し始めている．上に挙げた YouTube や Facebook，Netflix でも 360° 動画に対応している．

多くの動画配信サービスには視聴に関して専用のアプリケーションのほかに Web ブラウザ版が用意されている．Web は，主流の Web ブラウザを使用していれば誰でもサービスを利用できるという点で非常に手軽な環境であり，技術の進化により Web 上での 360° 動画の視聴体験や立体音響体験が可能となっている．一方で，これらの技術を用いた 3D コンテンツのインタラクティブな再生を可能とする Web アプリケーションは少ない．本研究では，クラシックコンサートとジャズセッションのイベントの収録データをもとに，インタラクティブに 360° 動画および立体音響を視聴体験する Web アプリケーション Web360^2 を開発した．

本論文では，2 節で立体音響技術の手法やインタラクティブな 3D 視聴体験についての先行システムを紹介する．3 節で本研究の目的を述べ，4 節で本研究の目的を実現するために開発した，インタラクティブな 3D 視聴体験を提供する Web アプリケーションである Web360^2 の設計や実装について述べる．5 節では Web360^2 について行われた 93 人の被験者による主観評価の結果を報告する．最後に，6 節で本論文のまとめと今後の課題を述べる．

[1] 東京大学
The University of Tokyo, Bunkyo, Tokyo 113–8654, Japan
[2] Bitlet 合同会社
Bitlet LLC., Chuo, Tokyo 103–0021, Japan
[a)] shin@hongo.wide.ad.jp

© 2019 Information Processing Society of Japan

2. 関連研究

立体音響技術には，様々な方法が存在する．現在，国際標準化機構 (ISO) および国際電気標準会議 (IEC) のワーキンググループである Moving Picture Experts Group (MPEG) が標準化を進めている MPEG-H において，MPEG-H Part 3 として 3D Audio（立体音響）に関する項目がある [3]．MPEG-H 3D Audio が許容する立体音響の手法は，チャンネルベース，オブジェクトベース，HOA（Higher-order Ambisonics，高次アンビソニックス）の 3 つである [4]．

チャンネルベースオーディオは，2.0 チャンネルのステレオサウンドから 5.1 チャンネルを代表とする多チャンネルのサラウンドサウンドやイマーシブサウンドへと発展している．チャンネルベースオーディオはスピーカーの数や視聴者に対するスピーカーの位置関係の情報をもとに，再生環境に紐付けて音響データを記録・編集する．したがって，再生環境の変更に対しては環境に紐付けられた別の音響データが必要になる．あるいは，もとの音響データを改変する必要がある [5]．また，視聴者の動きに追従して音響に動きを加えることも難しく，柔軟性やインタラクティブ性に欠ける．

オブジェクトベースオーディオでは，音響データとともに 3 次元の位置情報がメタデータとして記録される．再生環境を事前に想定する必要はなく，スピーカーの数や位置に適した音響を柔軟にレンダリングしながら再生することができる．また，音響データがもつ 3 次元位置情報の時間変化に対応しながらレンダリングすることが可能であるため音響オブジェクトの動きを表現できるほか，視聴者の位置移動や頭部の回転動作に音響を追従させることもできる．コンテンツ制作において，コンテンツに含まれる楽器の演奏や音声，効果音にはパンニングツールによって 3 次元位置情報が割り当てられており，今日のスタジオには，すでに膨大な数のオブジェクトベースコンテンツが存在している．大部分は配信される際にオブジェクトベースオーディオとしてではなく，従来のチャンネルベースオーディオとしてミキシングされてしまう [5] が，主に映画館やホームシアターへの導入が進む Dolby Atmos [6] [7] や DTS:X [8]，AuroMax [9] ではオブジェクトベースオーディオを採用，あるいは，従来のチャンネルベースオーディオとともに併用している．

HOA は音響を，1 点を中心に広がる球状の音場として記録する手法である．数学的には音場を球調和関数展開を利用して表現する方法であり，従来より研究されてきたアンビソニックスでは球調和関数の 0 次および 1 次係数にのみ着目するが，HOA では 2 次以上の係数も考慮することで，より高品質な音場を表現できるようになった [10]．HOA/アンビソニックスのもつ空間的な情報はオブジェクトベースオーディオのような明示的な幾何学的メタデータではないため，音場内のそれぞれの音響オブジェクトへアクセスすることは難しい [5]．HOA/アンビソニックスの収録には，360° 全方位を収音するためにアンビソニックマイクと呼ばれる特殊なマイクが必要となる．収録されたデータはマイクを中心とする球状の音場を再現するため，視聴者の頭部回転動作に対しては音場の回転によって容易に追従することができる．一方で，視聴者は常に音場の球の中心に位置することが求められるため，位置移動への追従は難しい．全天球カメラによって撮影された 360° 動画も同様であり，映像の回転は容易だが中心点から移動することはできない．このように，HOA/アンビソニックスと全天球カメラは同じ特性をもち，相性がよいと言える．最近では，RICOH THETA Z1, THETA V [11] などの民生用全天球カメラにもアンビソニックマイクが搭載されるようになり，YouTube や Facebook の 360° 動画サービスでも HOA/アンビソニックスに対応している．

著者らが所属する SDM コンソーシアム [12] [13][*1]の先行研究 [14] では，タブレット端末を前提とした収録済みコンテンツの再生アプリケーション SDM360^2 を開発し，[15] では，HMD（Head Mounted Display，ヘッドマウントディスプレイ）を前提としたライブ配信コンテンツ対応のアプリケーション LiVRation を開発した．360° 動画を投影し，オブジェクトベースオーディオの手法をもとに音響オブジェクトをマッピングした仮想空間内で，視聴者の頭部動作に応じた映像・音響の追従を行う．また，視聴者の操作によって仮想空間内を移動することができる．さらに，各音源がオブジェクトとして独立していることを利用して，音源ごとに ON/OFF を切り替えたり，音量調節を行うことができるなどインタラクティブな 3D コンテンツの再生を実現している．どちらも Unity を用いて開発されており，開発・実行には Unity が必要である．

Google Creative Lab と Song Exploder が 2017 年に公開した実験サイト Inside Music [16] は，WebVR と Web Audio の技術を用いており，Web 上で仮想空間におけるオブジェクトベースオーディオによる立体音響を体験できる．仮想空間内には 360° 動画の投影はされず，音響オブジェクトなどのコンポーネントのみが投影される．空間内は自在に水平移動することができ，視聴位置や視聴角度の変化に音響が追従する．また，各音源オブジェクトをクリックすることで ON/OFF の切り替えができる．専用のアプリケーションのインストールは不要で，Web ブラウザを利用することができれば，HMD や PC，モバイル端末から非常に手軽に体験できる．ソースコードは GitHub[*2]に公開されており，Apache License, Version 2.0 [17] に則り，コードの利用を提示することで Inside Music のシステムを利用

[*1] https://sdm.wide.ad.jp/
[*2] https://github.com/googlecreativelab/inside-music

© 2019 Information Processing Society of Japan

したインタラクティブな VR 体験アプリケーションの制作が可能になる．

3. 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対象として，Web 上で自由視点 3D コンテンツのインタラクティブな再生を行うことである．具体的には Web 上で，収録された 360° 動画を用いた自由視点視聴，および，視聴者の動作に追従し，個別の音源をインタラクティブに制御することができる立体音響の提示を目指す．そのため立体音響の手法として，従来のチャンネルベースや 360° 動画との相性がよいとされる HOA/アンビソニックスではなく，個別の音源の制御が容易なオブジェクトベース方式を採用する．本研究では，以上の目的を実現するために Web360^2 を試作し，実験を通して評価を行った．2 節で紹介した関連システムと Web360^2 の比較概要を**表 1** にまとめる．また，Web360^2 のシステム要件として以下を想定した．

Web アプリケーション：
専用のアプリケーションをインストールすることなく，Web ブラウザから直接利用可能な手軽さを目指す．

自由な映像音声の視聴：
視聴者が自由に視聴位置や視聴角度を決めることができ，定まった位置および角度と，視聴オブジェクトがもつ 3 次元位置情報をもとに自動で適切な映像・音響を合成し提示できる．

インタラクティブ性：
視聴者の操作によって，各音源のオブジェクトに個別にアクセスし ON/OFF の切り替えができる．これにより，視聴者の興味関心に基づいた自由度の高い視聴体験が可能となる．

ストリーミング配信：
360° 動画を HTTP Live Streaming (HLS) でストリーミング配信する．ストリーミング配信に対応することで，再生開始までの待機時間の短縮や余計な通信を省くことが可能となる．

4. Web360^2

3 節で述べた本研究の目的やシステム要件をみたすシステムとして Web360^2 を開発し，Github Pages[*3]にて公開している．ただし，現行バージョンでは各コンテンツにつき 1 つのみ視聴位置を実装しており，視聴位置変更の操作には対応していない．本節では Web360^2 のシステム設計や実装について述べる．

4.1 設計

Web360^2 のシステム設計概要を**図 1** に示す．視聴者の頭部動作や音響オブジェクトへのタッチ操作を入力として受け付け，操作に応じた映像や音響をシステム内でリアルタイムにレンダリングして出力することでインタラクティブ性を実現する．映像・音響のメディアファイルはそれぞれ独立しており，システム内で同期をとりながら再生する．映像ファイルには無音の 360° 動画，音響ファイルには AudioSprite[*4]によって各音源ファイルを 1 つにまとめた音源群が収録されている．AudioSprite は複数の音声ファイルを，時間間隔をとりながら並べて，1 つの音声ファイルにまとめる技術であり，再生の際のリクエスト数を減らすことができるため読み込み時間を短縮することができる．また，すべての音声が 1 つのファイルにまとまっていることで，音声ごとの読み込み遅延を気にする必要がなくなり，音声間の同期がとりやすいことも利点として挙げることができる．

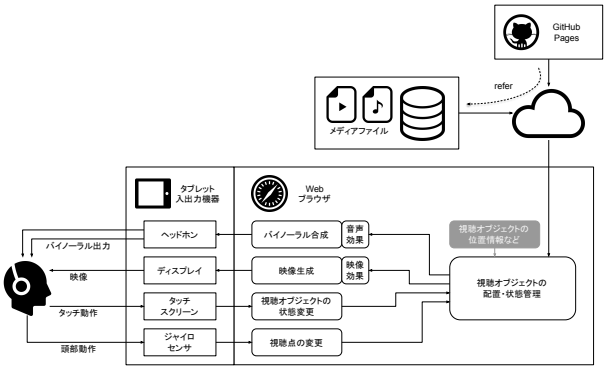

図 1 Web360^2 のシステム設計．
Fig. 1 System design of Web360^2.

360° 動画は HLS でストリーミング配信され，視聴者を中心とする仮想的な球面内側に投影して再生される．また，音響オブジェクトのもつ 3 次元位置情報からこの仮想的な球内部の各地点に音響可視化オブジェクトを描画し，各音響の強弱を表現したり，タッチ操作によって ON/OFF を切り替えることができる．視聴者の頭部動作は端末に搭載されたジャイロセンサを通してシステムに入力され視聴角度が変更される．視聴角度と ON 状態にある各音響オブジェクトの位置関係から聴こえるべき音響を生成し出力する．

4.2 素材データ

Web360^2 を開発に用いた素材データは，2016 年 1 月 10 日に慶應義塾大学日吉キャンパス内に建つ藤原洋記念ホールで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の収録データ，および，2017 年 1 月 26 日に六本木ミッドタウン内にある Billboard Live Tokyo で開催された Musilogue Band のコンサートの収録データ

[*3] https://sdm-wg.github.io/web360square/

[*4] https://github.com/tonistiigi/audiosprite

© 2019 Information Processing Society of Japan

表 1　関連システムとの比較.
Table 1　A comparison with related systems.

システム	環境	360° 動画	視聴位置操作	視聴角度操作	各音源操作	ライブ配信
SDM360^2	Unity	✓	✓	✓	✓	×
LiVRation	Unity + HMD	✓	✓	✓	✓	✓
Inside Music	Web	×	✓	✓	✓	×
Web360^2	Web	✓	✓（実装計画中）	✓	✓	×

である．[14] や [18] に，それぞれの収録環境や収録方法が記述されている．

4.3　実装

Web360^2 は，主に WebVR 用のフレームワークである A-Frame と Web Audio API を用いて実装した．実装概要を図 **2** に示し，視聴体験中のスクリーンショットを図 **3** に示す．

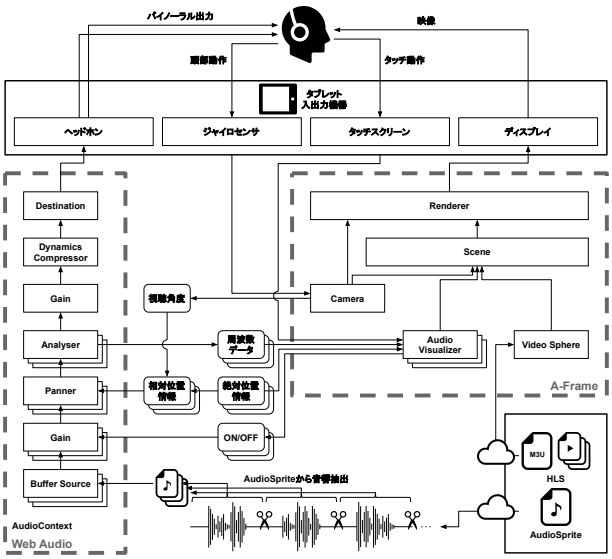

図 **2**　Web360^2 の実装概要.
Fig. 2　Implementation of Web360^2.



図 **3**　Web360^2 のスクリーンショット.
Fig. 3　Screenshot of Web360^2.

4.3.1　A-Frame を用いた処理

A-Frame を用いて，360° 動画の投影や音響可視化オブジェクトの描画を行う．360° 動画は HLS でストリーミング配信され，A-Frame の `a-videosphere` オブジェクトを用いて，カメラを原点に広がる球面内部に映像を投影する．HLS は Safari や Edge，一部のモバイル端末用ブラウザではネイティブサポートされているが，Chrome や Firefox，IE などサポートしていないブラウザも多く存在するため，`hls.js`[*5]を用いて対応した．音響の可視化は Web Audio の Analyzer ノードから提供される各音響の周波数領域データを用いて表現している．

また，視聴者のインタラクティブな操作は主に A-Frame が担う処理を介して入力される．視聴者の頭部動作は端末に搭載されたジャイロセンサに検知され，A-Frame のカメラオブジェクトに入力されることで視聴角度が変更される．タッチ操作は A-Frame でレンダリングしたオブジェクトにタッチすることで入力として認識され，各音響の ON/OFF 状態を変更することができる．

4.3.2　Web Audio API を用いた処理

Web Audio の AudioContext 上で各ノードを鎖状に繋いでいき，視聴角度と ON 状態の各音響オブジェクトから聴こえるべき音響をリアルタイムレンダリングする．はじめに各音源が 1 つにまとめられた AudioSprite のファイルデータから各音源を抽出し，Buffer Source ノードに入力する．次の Gain ノードでは，OFF 状態の音響のゲインを 0 にすることで音響の ON/OFF 状態を表現する．映像は視聴角度と各オブジェクトの絶対座標に応じて，A-Frame が自動的に適切な相対位置を計算しながらレンダリングするが，Web Audio 上で音響オブジェクトは視聴角度の変動に自動追従できないため，カメラの視聴角度と絶対座標から相対座標を随時計算して Panner ノードに受け渡すことで音響の追従性を実現する．例えば，視聴角度に対して音響オブジェクトが右側に存在すれば右側で音が鳴っているような音響を生成することができる．また，各音響の周波数領域データは Analyzer ノードで計算され，A-Frame の音響の可視化処理に用いる．以上の処理までは各音源ごとにされ，最後にマスターゲインとなる Gain ノードや Dynamics Compressor ノードでまとめられバイノーラル出力される．

[*5] `https://github.com/video-dev/hls.js/`

© 2019 Information Processing Society of Japan

4.3.3 音響の可視化

Web Audio の Analyser ノードでは，時間領域データや FFT（Fast Fourier Transform，高速フーリエ変換）による周波数領域データを取得することができる．Web360^2 の音響可視化では周波数領域データのみを用いる．音響可視化オブジェクトを構成する A-Frame コンポーネントの `tick` ハンドラを用いて毎秒 60 回から 120 回程度のオーダーで，周波数データを取得する `AnalyserNode.getByteFrequencyData()` メソッドを呼び出す．取得した周波数データの処理行程の概要を**図 4** に示す．`AnalyserNode.getByteFrequencyData()` メソッドによってナイキスト周波数以下の周波数データが得られるが，極端な高周波数帯・低周波数帯にはあまりデータが存在せず，有効な値が含まれる周波数の範囲は限られていることが多い．そこで，一度でも有効な値が現れたことがある周波数の上限・下限をとり，この範囲を有効周波数帯とする．有効周波数帯は再生開始から周波数データが取得される度に更新される．有効周波数帯の中に含まれるデータを周波数の高さの順に可能な限り均等な 32 グループに分けて，グループ内で平均値を計算する．得られた 32 グループそれぞれの平均値は音響可視化オブジェクトの 32 本の棘の長さと色に変換されて描画される．周波数の高さを音響可視化オブジェクトの棘の長さに対応させて描画してみたところ，近接する周波数グループ同士に相関があるために部分的に棘が盛り上がるような不格好な形状となったため，周波数の 32 グループと音響可視化オブジェクトの 32 本の棘の対応関係はランダムに定まるようにした．また，棘の色は OFF 状態ではグレーであり，ON 状態では平均値が大きくなるにつれて青，緑，黄，オレンジ，赤と変化していく．

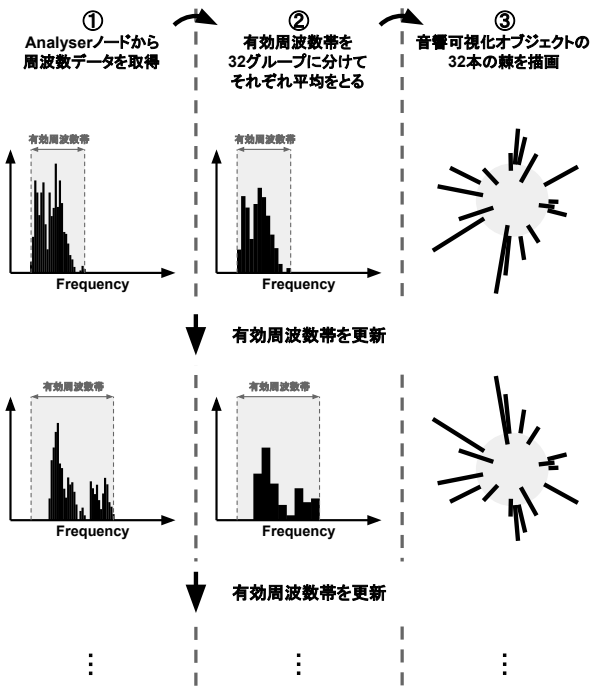

図 4　音響を可視化する処理．
Fig. 4　Process of audio visualiser.

4.3.4 映像と音響の同期処理

映像・音響のメディアファイルはそれぞれ独立しているため同期をとりながら再生する必要がある．映像の再生時間は JavaScript の `HTMLMediaElement.currentTime` プロパティによって取得・設定することができる．一方，音響については Web Audio API に読み取り専用の `AudioContext.currentTime` プロパティが用意されているが，これは音響の再生時間ではなく AudioContext が生成されてからの経過時間を示す．音響が停止中であっても `AudioContext.currentTime` プロパティの値は単調に増加し続けるため，システム内では音響の再生・停止を管理する Web Audio API の `AudioBufferSourceNode.start()` メソッドおよび `AudioBufferSourceNode.stop()` メソッドが呼び出される際に `AudioContext.currentTime` プロパティの値を記録するなどして音響の再生時間を算出している．

再生開始時やループ再生により先頭に戻った際には映像の再生時間を算出された音響の再生時間に強制的に同期させる．また，音響可視化オブジェクトの A-Frame コンポーネントがもつ `tick` ハンドラによって毎秒 60 回から 120 回程度のオーダーで，算出された音響の再生時間と映像の再生時間の差分を管理する．差分値に対して閾値を設定し，閾値を超えたときに映像の再生時間を音響の再生時間に強制的に同期させる対応は可能ではあるが，同期処理が入るたびに映像再生の滑らかさが失われ，体感品質を損なってしまう．そこで，現行バージョンでは差分の閾値を 0.1 秒として，閾値よりも映像が遅れたら映像の再生速度を 2 倍にし，閾値よりも映像が先行したら映像の再生速度を 0.5 倍にすることで，映像と音響の再生時間のずれが徐々に閾値以内に収束するような比較的制約のゆるい同期をとっている．再生速度の変調は多少の違和感を生じてしまうが，映像再生の滑らかさを保つことができる．映像の再生速度は JavaScript の `HTMLMediaElement.playbackRate` プロパティによって設定する．

5. Web360^2 の評価

Web360^2 について，アンケート調査による主観評価を行った．

5.1 実験方法

6 月 12 日から 6 月 14 日までの 3 日間にわたって開催された Interop Tokyo 2019 の SDM コンソーシアムブースでデモンストレーション展示を行った際に，Web360^2 を体験した来場者にアンケートを回答してもらい，調査を実施した．体験では，2 台のヘッドホン (Sony WH-1000XM2) と

© 2019 Information Processing Society of Japan

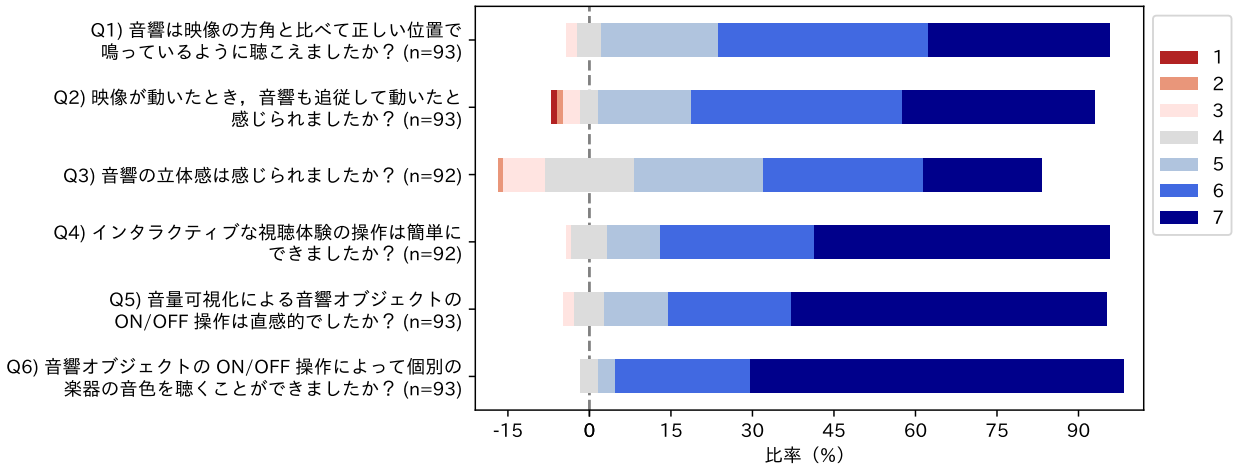

図 5　アンケートの回答比率.
Fig. 5　Response rates for questionnaires.

2 台のタブレット (Apple iPad Pro 12.9-inch (2018), Apple iPad Pro 11-inch (2019)) をそれぞれ有線接続して，Wi-Fi で Web360^2 にアクセスした．

はじめにアプリケーション概要と操作方法を実際に操作する様子を見せながら説明し，その後，体験者に自由に操作してもらった．体験終了後，アンケートの回答をお願いしてデータを取得した．

5.2　被験者

被験者は Web360^2 を体験した Interop Tokyo 2019 の来場者であり，男性 81 人，女性 9 人，無回答 3 人の合計 93 人で行った．被験者の年齢構成は 10 代 1 人，20 代 24 人，30 代 26 人，40 代 18 人，50 代 17 人，60 代 6 人，無回答 1 人である．また，社会人 81 人，学生 3 人，教員 5 人，無回答 4 人であった．

5.3　質問項目

アンケートの質問項目は，以下に示す設問 Q1 から Q6 までの 6 つであり，それぞれ 1 から 7 までの 7 段階のリッカート尺度を用いて回答してもらった．それぞれの回答について，最低評価を 1，中間を 4，最高評価を 7 とする目安を記載した．

Q1 音響は映像の方角と比べて正しい位置で鳴っているように聴こえましたか？
Q2 映像が動いたとき，音響も追従して動いたと感じられましたか？
Q3 音響の立体感は感じられましたか？
Q4 インタラクティブな視聴体験の操作は簡単にできましたか？
Q5 音量可視化による音響オブジェクトの ON/OFF 操作は直感的でしたか？
Q6 音響オブジェクトの ON/OFF 操作によって個別の楽器の音色を聴くことができましたか？

設問 Q1 および Q2 は映像と音響の組み合わせが違和感なく一体として知覚できるかを問う設問である．設問 Q1 は映像上の音響オブジェクトの位置と出力される音響の鳴っている方角・位置の知覚に差異がないかを問い，設問 Q2 は頭部動作による視聴角度の動きに対する音響の追従性を問う．設問 Q3 はシステム内で合成された音響に立体感を感じるかを問う設問である．また，設問 Q4 および Q5 はタブレットをもった状態での頭部動作やタッチ動作によるインタラクティブな視聴体験についての設問であり，設問 Q4 は全般的な操作の容易さを問い，設問 Q5 は音響オブジェクトの可視化と ON/OFF 操作について問う．設問 Q6 は各音響の ON/OFF 操作を通して，全体の音響を分解し個別の楽器の音色がを聴くことができたかを問う．

アンケートの末尾に「感想・要望・その他」という自由記述欄を用意して，Web360^2 の体験に関するコメントを得た．

5.4　評価結果

アンケートの調査結果を図 **5** に示す．グラフの横軸は 1 から 7 までの 7 段階の回答の比率を表し，それぞれ合計すると 100%になる．中間評価を示す 4 の回答比率の中間を横軸の原点 0%に置き，正の方向に高評価を示す 5, 6, 7 の回答比率を，負の方向に低評価を示す 3, 2, 1 の回答比率を伸ばす．つまり，高評価 (5, 6, 7) の回答が多いほどグラフは正の方向に偏り，逆に低評価 (3, 2, 1) の回答が多いほどグラフは負の方向に偏っていく．グラフ縦軸は前項 5.3 で述べた各設問 (Q1–Q6) であり，それぞれ末尾の括弧内に有効回答数を示す．

すべての設問に対して概ね高評価を得ることができ，い

© 2019 Information Processing Society of Japan

ずれも最高評価の 7 と次点評価の 6 の回答割合のみで半数を上回った．特に後半の 3 つの設問 (Q4–Q6) については，最高評価の 7 のみで半数を超えている．一方で，音響の立体感を問う設問 Q3 については，評価が分散しており，あまり実感できなかったという評価（低評価の 3, 2, 1）も 8.7%とほかの設問に比べて多い結果となった．音響の立体感を構成する要素は音像の方角や奥行き，音源の臨場感，空間の臨場感など複数あると考えられるが，設問 Q3 は表現が曖昧であり，質問がどの要素を問うているのか，被験者がどの要素について応えているのか，明確にわからないという反省点が挙げられる．音響の立体感は高品質な 3D 視聴体験を実現する上で重要な要素であるので，明確な評価結果が得られるために，また，それによって品質向上をめざすために，次回以降の主観評価実験における質問項目の設定では十分に注意したい．

「感想・要望・その他」の自由記述欄でも，直感的な操作性や個別の音響を消す機能，音響の追従によりその場にいるような没入感が得られたことへの好感触な感想があった一方で，360° で音響が鳴っている様子があまり実感できないという意見があり，各設問内容への実感がコメントに現れていた．また，仮想空間内の移動への要望があったが，別の位置で収録された 360° 動画と組み合わせて，視聴位置の移動操作として実装する予定である．そのほか，各音源のボリューム調整やエフェクトの付与への要望などがあり，今後の開発において検討していきたい．

6. まとめと今後の課題

本稿では，インタラクティブな 3D 視聴体験を提供する Web アプリケーション Web360^2 について述べた．システム要件として，Web アプリケーションであること，自由な映像音声の視聴が可能であること，インタラクティブ性があること，ストリーミング配信に対応することを設定し，これに基づき Web360^2 の設計，実装を行い，視聴位置の変更以外の最低限の要件はみたされた．また，Interop Tokyo 2019 の来場者 93 人に対して，Web360^2 の体験とアンケートの回答を協力してもらうことで主観評価をとり，Web 上でのインタラクティブな 3D 視聴体験について高評価を得た．一方で，主観評価項目の一部に表現が明確でなく意図が伝わりづらい点が確認された．主観評価を有意義なものとするため，評価項目を検討し直し洗練させる必要があるだろう．

今後の課題としては，異なる視聴位置へ移動できる機能を実装して自由視点の要件をみたすことが挙げられる．また，SDM コンソーシアムでは，映像・音響データだけでなく，位置情報や楽器の向き，演奏された曲目，会場の情報，収録のプロセスなど詳細な収録環境のメタデータを記録する機構 SDM Ontology [19] [20] の提案・開発を進めている．現行の Web360^2 では事前に決められたコンテンツのみ視聴体験可能であるが，SDM Ontology と連携することで，SDM Ontology 上に格納されたコンテンツの情報を問い合わせ，得られた映像・音響データ及び位置情報によって動的にインタラクティブな視聴空間を構築できるようになると考えられる．これは Web360^2 が SDM Ontology を利用することによる利点であるが，逆に，SDM Ontology に格納されたデータは Web360^2 を通して視聴空間に変換可能であるため，SDM Ontology に格納されたデータの正当性を視覚的に確認することも可能となる．このように相互に活用しながら開発を続けていき，より成熟されたアプリケーションへと発展させていきたい．

参考文献

[1] Sandvine: Global internet phenomena report, Technical report, Sandvine (2018).
[2] Sandvine: Mobile internet phenomena report, Technical report, Sandvine (2019).
[3] ISO/IEC: 23008-3:2019, Information technology – High efficiency coding and media delivery in heterogeneous environments – Part3: 3D audio, Standard, ISO/IEC (2019).
[4] Beack, S., Sung, J., Seo, J. and Lee, T.: MPEG Surround Extension Technique for MPEG-H 3D Audio, *ETRI Journal*, Vol. 38, No. 5, pp. 829–837 (2016).
[5] Herre, J., Hilpert, J., Kuntz, A. and Plogsties, J.: MPEG-H 3D audio—The new standard for coding of immersive spatial audio, *IEEE Journal of selected topics in signal processing*, Vol. 9, No. 5, pp. 770–779 (2015).
[6] Dolby Laboratories: Dolby Atmos®Specifications, Isuue 3, Dolby Laboratories (2015).
[7] Dolby Laboratories: Dolby Atmos®Home Theater Installation Guidelines, Technical report, Dolby Laboratories (2018).
[8] DTS, Inc.: Home Theater Sound Gets Real, DTS, Inc. (online), available from ⟨`https://dts.com/dtsx`⟩ (accessed 2019-09-05).
[9] Auro Technologies: AUROMAX®Next generation Immersive Sound system, Technical report, Auro Technologies (2015).
[10] Frank, M., Zotter, F. and Sontacchi, A.: Producing 3D audio in ambisonics, *Audio Engineering Society Conference: 57th International Conference: The Future of Audio Entertainment Technology–Cinema, Television and the Internet*, Audio Engineering Society (2015).
[11] Ricoh Company, Ltd.: 360-degree camera RICOH THETA, Ricoh Company, Ltd. (online), available from ⟨`https://theta360.com/`⟩ (accessed 2019-09-05).
[12] 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，コンピュータソフトウェア，Vol. 34, No. 3, pp. 3_37–3_58 (2017).
[13] Tsukada, M., Ogawa, K., Ikeda, M., Sone, T., Niwa, K., Saito, S., Kasuya, T., Sunahara, H. and Esaki, H.: Software defined media: Virtualization of audio-visual services, *2017 IEEE International Conference on Communications (ICC)*, IEEE, pp. 1–7 (2017).
[14] 塚田 学，菰原 裕，粕谷貴司，新居英明，高坂茂樹，小川景子，江崎 浩：SDM360^2: インタラクティブ 3D コンテンツの自由視聴点再生，情報処理学会論文誌デジタ

© 2019 Information Processing Society of Japan

ルコンテンツ (DCON)，Vol. 6, No. 2, pp. 10–23 (2018).
[15] 粕谷貴司，塚田　学，菰原　裕，高坂茂樹，水野拓宏，野村譲誉，上田雄太，江崎　浩：LiVRation: VR による自由視聴点映像音声のインタラクティブ再生，マルチメディア, 分散協調とモバイルシンポジウム 2018 論文集，Vol. 2018, pp. 514–522 (2018).
[16] Google Inc. and Song Exploder: Song Exploder Presents: Inside Music, Google Inc. (online), available from ⟨https://experiments.withgoogle.com/webvr/inside-music/view/⟩ (accessed 2019-09-06).
[17] Apache Software Foundation: Apache License, Version 2.0, Apache Software Foundation (online), available from ⟨https://www.apache.org/licenses/LICENSE-2.0⟩ (accessed 2019-10-14).
[18] Ikeda, M., Sone, T., Niwa, K., Saito, S., Tsukada, M. and Esaki, H.: New recording application for software defined media, *Audio Engineering Society Convention 141*, Audio Engineering Society (2016).
[19] 菰原裕，塚田学，江崎浩，曽根卓朗，池田雅弘，高坂茂樹，新麗，新善文ほか：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シンポジウム 2017 論文集，Vol. 2017, pp. 110–115 (2017).
[20] Atarashi, R., Sone, T., Komohara, Y., Tsukada, M., Kasuya, T., Okumura, H., Ikeda, M. and Esaki, H.: The software defined media ontology for music events, *Proceedings of the 1st International Workshop on Semantic Applications for Audio and Music*, ACM, pp. 15–23 (2018).

© 2019 Information Processing Society of Japan