



6DoF 音響サービスに用いる音響メタデータの要求条件

Requirements for audio metadata employed in 6DoF audio services

杉本岳大¹⁾, 長谷川知美¹⁾, 大出訓史¹⁾

Takehiro SUGIMOTO, Tomomi HASEGAWA, and Satoshi OODE

1) 日本放送協会 (〒 157-8510 東京都世田谷区砧 1-10-11, sugimoto.t-fg@nhk.or.jp)

概要: イマーシブメディア用の音響技術実現に向けて, リスナの six-degrees-of-freedom (6DoF) に対応した音響メタデータ (6DoF 音響メタデータ) の開発を進めている. 6DoF 音響メタデータの記述法を確立するためには, 6DoF 音響サービスの実現に資する音響メタデータの要求条件を明確化する必要がある. 本研究では, ①コンテンツ制作者への意向調査, ②ゲームにおける音響制作のサーベイ, ③6DoF 音響制作ワークフローの検討, の各調査に基づいて, 6DoF 音響メタデータの要求条件を抽出した.

キーワード: 6DoF 音響, 音響メタデータ, イマーシブメディア, オブジェクトベース音響

1. 緒言

近年, イマーシブなコンテンツに対応したメディア (以下, イマーシブメディア) の開発が, 様々なジャンルで進んでいる. 中でもゲーム業界は, 他業界と比してコンテンツのイマーシブ化が先行しており, 技術革新が目覚ましい [1]. 放送業界においても, AR (augmented reality) や VR (virtual reality) を応用したコンテンツのイマーシブ化の機運は着実に高まっている [2, 3]. さらに, イマーシブ化は映像だけでなく, 音響技術においても顕著に活発化しつつある [4, 5, 6]. 22.2 マルチチャンネル音響を始めとする従来の 3 次元音響が 3DoF (three-degrees-of-freedom, 直交座標系の 3 座標軸回りの回転) 対応であることに對し, イマーシブメディア用の音響技術は 6DoF, つまり 3DoF に加えてリスナの移動にも対応できる技術を目指している [4] (図 1).

ところで, 6DoF 音響コンテンツの制作プラットフォームとしては, リスナと発音源との位置関係の変化を再生音に反映しやすいオブジェクトベース音響 [7] が適している. オブジェクトベース音響のコンテンツは, 音声オブジェクトを構成する音声信号と, 音声信号の位置や機能などを記述した音響メタデータ [8] を組み合わせて構成する. そこで, 6DoF 音響コンテンツの制作に向けて, 6DoF に対応した音響メタデータ (以下, 6DoF 音響メタデータ) の開発に着手した.

6DoF 音響メタデータの記述法を確立するためには, まず 6DoF 音響サービスの実現に資する音響メタデータの要求条件を明確化する必要がある. そこで本研究では, 6DoF 音響コンテンツの制作に関する調査を実施し, その結果に基づいて 6DoF 音響メタデータの要求条件を抽出する.

2. 6DoF 音響メタデータの要求条件の抽出方法

オブジェクトベース音響コンテンツの記述で用いる音響メタデータとしては, International Telecommunication Union-Radiocommunication Sector (ITU-R: 国際電気通信連合の無線通信部門) が勧告した Audio Definition Model (ADM) [8] が広く用いられている. ADM では, 音声オブジェクトを構成する音声信号の機能や性質を, 音響メタデータで詳細に記述して規定することになっており, 本研究の 6DoF 音響メタデータも, ADM に準じた書式での記述を目指している.

本研究では, 先行して実施した 6DoF 音響コンテンツの制作に関する①コンテンツ制作者への意向調査, ②ゲームにおける音響制作のサーベイ, ③6DoF 音響制作ワークフローの検討, に基づき [9], 希望点列挙法を用いて, 6DoF 音響サービスの実現に資する 6DoF 音響メタデータの要求条件を, 6DoF 音響コンテンツを構成する要素項目ごとに抽出した. また制作における主たる用途に合わせて, 得られた 6DoF 音響メタデータの要求条件を音響空間デザイン用と演出用に大別した.

3. コンテンツ制作者への意向調査から抽出した 6DoF 音響メタデータの要求条件

意向調査では, コンテンツ制作者 64 名に対して, 6DoF コンテンツ制作におけるコンセプトおよび 6DoF コンテンツにおける音響制作に関して回答を求めた [9]. 表 1 に, 意向調査の結果から抽出した 6DoF 音響メタデータの要求条件を示す. 表 1 中の各要求条件について, 以下に説明する.

- 位置/向き (発音源): 6DoF 音響空間における発音源の位置および向き
- 放射パターン: 発音源から放射される音波の放射パターン

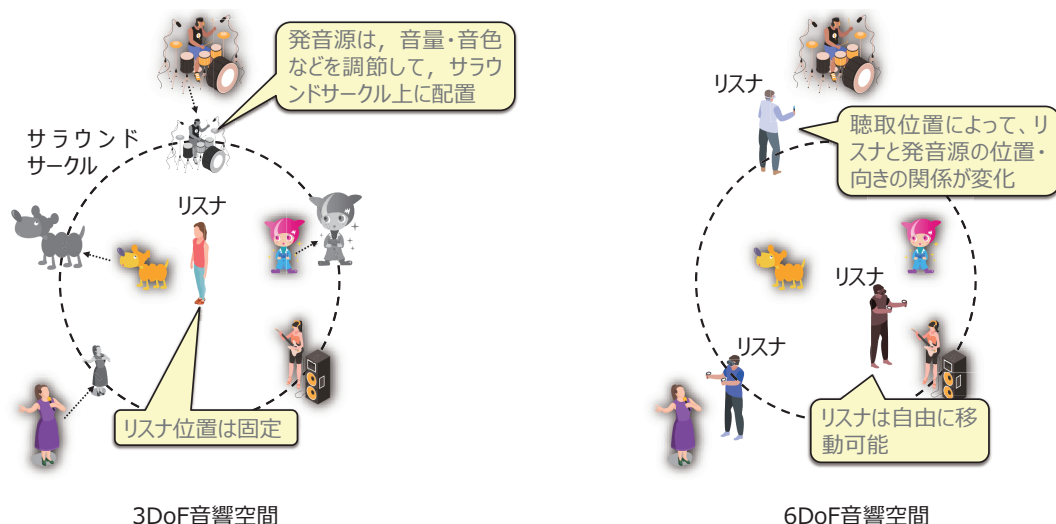


図 1: 3DoF 音響空間と 6DoF 音響空間の比較. 6DoF では、聴取位置によってリスナと発音源の位置・向きの変化する。

- インタラクションフリーフラグ: 背景音の音量や定位を、リスナの動作に依存させないためのフラグ
- 音響特徴量: 壁面や地面などの映像で提示された物体の、音響的な反射率や透過率などのパラメータ
- 距離減衰特性 (音響則): 6DoF 音響空間内での音波伝播に伴う減衰特性
- 位置/向き (リスナ): 6DoF 音響空間におけるリスナの位置および向き
- 視点誘導用定位シフトフラグ: リスナの視点誘導のための定位変更を許可するフラグ
- インタラクティブ再生フラグ: リスナのアクションをトリガとする、イベントベースの再生音を示すフラグ
- 必須オブジェクトフラグ: レンダリング時に聴きとりやすさを改善する必要のある音を示すフラグ

4. ゲームの音響制作のサーベイから抽出した 6DoF 音響メタデータの要求条件

ゲームの音響制作については、CEDEC2022 [1] をサーベイした [9]. 表 2 に、全講演のサーベイ結果から抽出した 6DoF 音響メタデータの要求条件を示す. 表 2 中の要求条件のうち、初出のものについて以下に説明する。

- 室内音響パラメータ: 残響時間, C80 (音楽の明瞭度), D50 (スピーチの明瞭度), インパルス応答など. 採用するパラメータの種類や数によって、コンプレキシティを制御することも可能
- 距離減衰特性 (音響空間モディファイ): 音響則として既定した距離減衰特性の特徴や適用範囲を調節するためのパラメータ
- 定位縮退フラグ: 6DoF 音響空間における発音源の定位を、特定の方向に偏在させるためのフラグ

表 1: コンテンツ制作者への意向調査から抽出した 6DoF 音響メタデータの要求条件

	項目	内容
デザイン用	発音源	<input type="checkbox"/> 位置/向き <input type="checkbox"/> 放射パターン <input type="checkbox"/> インタラクションフリーフラグ
	非発音源	<input type="checkbox"/> 音響特徴量
	音響則	<input type="checkbox"/> 距離減衰特性
	リスナ	<input type="checkbox"/> 位置/向き
演出用	音響空間モディファイ	<input type="checkbox"/> 視点誘導用定位シフトフラグ
	時間軸	<input type="checkbox"/> インタラクティブ再生フラグ <input type="checkbox"/> インタラクティブミュージックフラグ
	人声	<input type="checkbox"/> 必須オブジェクトフラグ

表 2: ゲームの音響制作のサーベイから抽出した 6DoF 音響メタデータの要求条件

	項目	内容
デザイン用	発音源	<input type="checkbox"/> 位置/向き <input type="checkbox"/> 放射パターン
	非発音源	<input type="checkbox"/> 室内音響パラメータ
	音響則	<input type="checkbox"/> 距離減衰特性
	リスナ	<input type="checkbox"/> 位置/向き
演出用	音響空間モディファイ	<input type="checkbox"/> 距離減衰特性 <input type="checkbox"/> 定位縮退フラグ <input type="checkbox"/> 聴取点シフトフラグ
	時間軸	<input type="checkbox"/> インタラクティブ再生フラグ <input type="checkbox"/> インタラクティブミュージックフラグ
	人声	<input type="checkbox"/> ダッキングメタデータ

表 3: 6DoF 音響制作ワークフローの検討から抽出した 6DoF 音響メタデータの要求条件

	項目	内容
デザイン用	発音源	<input type="checkbox"/> 位置/向き <input type="checkbox"/> 放射パターン <input type="checkbox"/> インタラクティブフリーフラグ
	非発音源	
	音響則	<input type="checkbox"/> 距離減衰特性
	リスナ	<input type="checkbox"/> 位置/向き
演出用	音響空間 モディファイ	<input type="checkbox"/> 距離減衰特性 <input type="checkbox"/> 基準位置
	時間軸	<input type="checkbox"/> インタラクティブ再生フラグ
	人声	<input type="checkbox"/> 必須オブジェクトフラグ <input type="checkbox"/> 音響則フリーフラグ

- 聴取点シフトフラグ: 再生音の総合的な音質を考慮して、リスナ位置とは異なる位置でレンダリングするためのフラグ
- インタラクティブミュージックフラグ: リスナ主導のシーンチェンジに合わせて、異なる音楽素材間をシームレスに移行するためのフラグ

5. 6DoF 音響制作ワークフローの検討から抽出した 6DoF 音響メタデータの要求条件

3次元音響やVR音響の制作実績を有するバランスエンジニアに聞き取り調査をおこない、6DoF音響制作ワークフローを検討した[9]。表3に、検討した6DoF音響制作ワークフローから抽出した6DoF音響メタデータの要求条件を示す。表3中の要求条件のうち、初出のものについて以下に説明する。

- 基準位置: バランスエンジニアの手で再生音のバランスが調節された位置、またはリスナに対する聴取推奨位置
- 音響則フリーフラグ: 距離減衰や回折、遮蔽などの音響則を無視してレンダリングするためのフラグ

6. 考察

異なる組織間での6DoF音響コンテンツの交換や、制作システムの相互接続性向上を考慮すると、音響メタデータは汎用性が高いほど有用であるため、第3章から第5章で抽出した各要求条件を統合して、6DoF音響メタデータの最終的な要求条件と規定した。

表4に、統合した6DoF音響メタデータの要求条件を示す。なお複数の調査で挙げられた要求条件は、赤字で示した。表4から、6DoF音響メタデータには、6DoF音響空間をデザインする機能だけでなく、コンテンツ制作者の演出意図を反映するための機能も必要であることがわかる。

表 4: 統合した 6DoF 音響メタデータの要求条件。複数の調査で挙げられた要求条件は、赤字で示した。

	項目	内容
デザイン用	発音源	<input type="checkbox"/> 位置/向き <input type="checkbox"/> 放射パターン <input type="checkbox"/> インタラクティブフリーフラグ
	非発音源	<input type="checkbox"/> 音響特徴量 <input type="checkbox"/> 室内音響パラメータ
	音響則	<input type="checkbox"/> 距離減衰特性
	リスナ	<input type="checkbox"/> 位置/向き
演出用	音響空間 モディファイ	<input type="checkbox"/> 視点誘導用定位シフトフラグ <input type="checkbox"/> 距離減衰特性 <input type="checkbox"/> 定位縮退フラグ <input type="checkbox"/> 聴取点シフトフラグ <input type="checkbox"/> 基準位置
	時間軸	<input type="checkbox"/> インタラクティブ再生フラグ <input type="checkbox"/> インタラクティブミュージックフラグ
	人声	<input type="checkbox"/> ダッキングメタデータ <input type="checkbox"/> 必須オブジェクトフラグ <input type="checkbox"/> 音響則フリーフラグ

7. 結言

本稿では、①コンテンツ制作者への意向調査、②ゲームにおける音響制作のサーベイ、③6DoF音響制作ワークフローの検討、の各調査に基づき、6DoF音響メタデータの要求条件を抽出した。

今後は、本検討を受けて6DoF音響メタデータの具体的な記述法を開発し[10]、本研究で抽出した6DoF音響メタデータの要求条件が幅広いジャンルのコンテンツの記述に対応できるかについて、実際にコンテンツを制作しながら検証していく。

参考文献

- [1] CEDEC2022, <https://cedec.cesa.or.jp/2022/>
- [2] Rec. ITU-R BT.2420-5: Collection of usage scenarios of advanced immersive sensory media systems, 2022.
- [3] NHK 放送技術研究所: Future Vision 2030-2040, https://www.nhk.or.jp/strl/future_vision/index.html
- [4] 杉本岳大, 堀内俊治: イマーシブメディア実現に向けた音響技術, 映像情報メディア学誌, Vol. 77, pp. 17-21, 2023.
- [5] ISO/IEC 23090-4: Immersive Audio (規格化中)
- [6] J. Herre and S. Disch: MPEG-I Immersive Audio – Reference model for the virtual/augmented reality audio standard, Vol. 71, pp. 229-240, 2023.
- [7] Rec. ITU-R BS.2051-3: Advanced sound system for programme production, 2022.
- [8] Rec. ITU-R BS.2076-2: Audio Definition Model, 2019.
- [9] 杉本岳大, 長谷川知美: 6DoF音響コンテンツの制作要件, 映像情報メディア学会年次大会, 32B-4, 2023.

- [10] 長谷川知美, 杉本岳大, 久保弘樹, 大出訓史, 岩崎泰士, 大久保洋幸: 6DoF 音響サービスのための発音源メタデータの記述法, 日本バーチャルリアリティ学会大会論文集, 2023.