

地上デジタル放送方式高度化の 音声符号化方式に関する中間報告

概要

2022年10月11日

一般社団法人 電波産業会
デジタル放送システム開発部会
音声符号化方式作業班

中間報告の内容／検討状況

- 符号化音声入力フォーマット
 - 検討済（残課題なし）
 - 48kHz, 16bit以上, 最大56ch
- 音声符号化方式
 - 検討済（残課題なし）
 - 2方式（MPEG-H 3D Audio, AC-4）

提案方式の概要

- 符号化音声入力フォーマット

項目	値
標本化周波数	48kHz
入力量子化ビット数	16bit以上
対応する音声信号	オブジェクトベース チャンネルベース
最大入力チャンネル数	56ch ^{注1}

注1

同時デコード数の最大は28ch、
チャンネルベースの音声信号の1グループの構成は最大22ch、低域を強調する2chとする

提案方式の特徴と選定理由

- 符号化音声入力フォーマット
 - 高度広帯域衛星デジタル放送のフォーマットを考慮
- 標本化周波数：現状の実運用動向を鑑み、**48kHz**のみとした
 - **48kHz**以外の標本化周波数については、実運用動向から放送局設備へのインパクトが大きいこと、当面サービスが想定されていないことを鑑み、規定しないこととした
- 入力量子化ビット数：現状の実運用動向を鑑み、**16 bit**以上とした
- 対応する音声信号：チャンネルベースに加えて、多様な音声サービスを高効率に実現可能なオブジェクトベースを採用した

提案方式の特徴と選定理由

- 符号化音声入力フォーマット

- 最大入力音声チャンネル数：22.2 ch音響に対応し、かつ、オブジェクトベース音声を用いた差し替え音声によるサービスを考慮して、MPEG-H 3D Audioで規定されているレベル4^{注1}で規定された最大入力音声チャンネル数^{注2}とした

注1 レベル：MPEGで規定されているデコーダの回路規模を決めるためのパラメータ。標本化周波数、最大入力信号数、最大同時デコード数、最大出力信号数が規定されている

注2 最大入力音声チャンネル数：これまでのチャンネルベースの放送サービスにおいて、二カ国語放送などの副音声は別のストリームで実施されてきたが、オブジェクトベースでは副音声の音声信号も一つのストリームとして伝送することを想定するために、最大入力音声チャンネル数は実際に再生される音声信号数よりも多く見積もられている。

提案方式の概要

- 音声符号化方式

	MPEG-H 3D Audio	AC-4
準拠規格	ISO/IEC 23008-3	ETSI TS 103 190-2
プロファイル	Baseline	規定なし
ビットストリーム形式	MHAS形式	raw_ac4_frame 形式

提案方式の概要と選定理由

- 音声符号化方式（**MPEG-H 3D Audio**）
 - 準拠規格： **ISO/IEC 23008-3**
 - オブジェクトベースによる音声サービスが実現可能な符号化方式
 - **MPEG-2/4 AAC**に比べて高効率な音声符号化方式であることを主観評価実験により確認
 - プロファイル： **Baseline**
 - オブジェクトベースによる音声サービスを実施可能な最も回路規模が小さいプロファイル
- 音声符号化方式（**AC-4**）
 - 準拠規格： **ETSI TS 103 190-2**
 - オブジェクトベースによる音声サービスが実現可能な符号化方式
 - **MPEG-2/4 AAC**に比べて高効率な音声符号化方式であることを主観評価実験により確認

要求条件への適合性

・システム

要求条件	提案方式による適合性 (MPEG-H 3DA)	提案方式による適合性 (AC-4)
HDTVを超える高画質・高音質・高臨場感サービスを基本として、多様な画質のサービス等を提供できること。	音の位置を自由に座標でも配置できる高臨場感な3次元立体音声サービスに対応。 オブジェクトベースに対応した多様なサービスが可能。	同左。
多様で柔軟な高機能サービスを提供できること。	オブジェクトベースに対応した多様なサービスが可能。	同左。
高齢者、障害者等様々な視聴者向けの放送サービスについても考慮すること。	オブジェクトベースに対応した多言語サービス、視覚及び視聴覚障害者用副音声サービス、高齢者用音声サービスを高効率に実現可能。	同左。
視聴者に違和感を与えない程度の映像・音声の遅延差であること。	システムレイヤ構成により実現可能。	同左。

要求条件への適合性

・システム

要求条件	提案方式による適合性 (MPEG-H 3DA)	提案方式による適合性 (AC-4)
高い実時間性を実現するため、できるだけ遅延時間を短くすること。	提案した符号化方式は短いブロック長を持つ。	同左。
放送の要件に応じて伝送パラメータの選択や組み合わせの指定を行うことができ、またそれに合わせて受信機制御ができる方式とすること。	音響メタデータにより受信機の制御が可能。	同左。
送出する映像、音声、データのフォーマットやビットレート、チャンネル数等を任意に選択、変更できること。	多くの音声モードやビットレートに対応。	同左。
諸外国も容易に導入できるシステムとなるよう考慮すること。	提案方式が他国で導入済。 (提案した符号化音声入力フォーマットとは異なる仕様が主流)	同左。

要求条件への適合性

• 放送品質

要求条件	提案方式による適合性 (MPEG-H 3DA)	提案方式による適合性 (AC-4)
放送サービスに応じて音声のフォーマットやビットレートを変更できること。	各種音声フォーマット・ビットレートに対応。	同左注。
UHDTVサービスに対応した、高音質・高臨場感な音声サービスに適した音質が望まれることを考慮し、できるだけ高い音質を保つこと。	最大22.2chの音声モードにより高臨場感音声サービスに対応。 主観評価実験によりMPEG-2/4 AACよりも高効率に符号化可能であることを確認。	

注：22.2chへの対応についてはAC-4の要求条件の適合性に関する補足を参照

要求条件への適合性

• 技術方式

要求条件	提案方式による適合性 (MPEG-H 3DA)	提案方式による適合性 (AC-4)
国際標準と整合した方式を用いること。	ISO/IEC規格準拠。	ETSI規格準拠。
高効率かつ高音質な符号化方式であること。	MPEG-2/4 AACよりも高効率に符号化可能。	同左。
多チャンネル音声放送をはじめとした、さまざまなサービス要件に柔軟に対応できる符号化方式であること。	最大22.2chに対応。 オブジェクトベースに対応した多様なサービスが可能。	最大22.2chに対応 ^注 。 オブジェクトベースに対応した多様なサービスが可能。
現行設備や受信機への負担等を考慮して選定される種々の音声入力フォーマットに対応できること。	多くの入力フォーマットに対応。	同左。

注：AC-4の要求条件の適合性に関する補足を参照

AC-4の要求条件の適合性に関する補足

- 欧米各国で採用されているAC-4 Level 3で22.2chを符号化するためには、聴感的に重要度の低いチャンネルを適応的にまとめて、最大17.1個の座標付きモノラル信号として符号化する。復号側では、これを22.2chにレンダリングして出力を得る。
- 22.2chをそのまま符号化する必要がある場合、今後の議論で具体化される同時最大デコードチャンネル数などに適合する形で、現在のAC-4規格上で"Reserved"となっているLevel 4として規定する必要がある。