

## テレビの音量レベル差と放送規格 第 2 回 放送規格と音量レベル差

JPPA 技術委員会

オーディオ部会委員 丸谷正利

### はじめに

今回は「サラウンドの音はなぜ低いのか」をテーマに、サラウンド音声の再生環境の違い、ダウンミックスの問題、ミキシングレベルの問題、などについて述べました。今回はその続きと、アナログ放送、デジタル放送の音量差、欧米のデジタル放送におけるサラウンド音声などについて述べます。

### 6 ポストプロダクション以外での対応

#### 6-1 二音声放送による対応

二音声放送（二音声エレメンタリストリーム=2ES）とは、サラウンド音声とステレオ音声を同時に送出する方法です。二ヶ国語（デュアルモノ）放送とは異なり音声ストリーム（ADTS）を二つ使用し、テレビでは「音声1」「音声2」のように表示されます。視聴者はどちらかの音声を選択できますが、デフォルトでは音声ストリーム1が選択されるようになっています。現在の放送局の運用ルールでは、二音声放送の場合は音声ストリーム1がサラウンド、音声ストリーム2がステレオとなっています。

この方法で、適正レベルで制作したサラウンド音声と“ステレオミックスした”ステレオ音声を使用し、視聴者に自分の再生環境に合った音声ストリームを選択させることで、音量差のない番組提供が可能となります。サラウンド番組やサラウンド CM は、搬入規定により、1ch から 6ch にサラウンド音声、7-8ch はステレオ音声となっており、ポストプロダクション時にステレオ音声も制作しているので局側の対応で実現可能な方法です。

実際にサラウンドの二音声番組を聴取され

た方もおられると思いますが、残念ながらこの方法を使用した番組でもサラウンド音声のレベルが大きく制作されているように感じます。その理由を考えてみると、音声1がサラウンドになっているためではないか、と思います。現在の二音声放送運用ルールでは、一般視聴者が意識して音声2を選択しない限り、サラウンド番組はダウンミックス音声で聴取することになります。したがって「サラウンドの音が低い」と言うクレームになることを避けるために、サラウンド音声のレベルを大きくしているのではないかと考えられます。

二音声放送では、通常図6-1の音声ストリーム構成が使用されています。この場合、音声ストリーム1は本編がサラウンド音声、CMはステレオ音声で構成され、音声ストリーム2は本編、CMともにステレオ音声で構成されています。番組内で一度音声ストリーム2を選択すれば、番組の終了までステレオ音声での聴取が可能となります。しかし、途中でチャンネルを切り替え、再びもとのチャンネルに戻るとテレビのデフォルト設定である音声ストリーム1が選択され、音声ストリーム2へ戻ることが出来ません。あらためて2を選択し直す必要があります。

映像ストリーム	本編 映像1	CM 映像	本編 映像2
音声ストリーム1	本編 サラウンド音声1	CM ステレオ音声	本編 サラウンド音声2
音声ストリーム2	本編 ステレオ音声1	CM ステレオ音声	本編 ステレオ音声2

図 6-1 番組全体を 2ADTS 構成にする

このように、二つの音声ストリームを使用した場合の問題点は、チャンネル切り替えなどで一度音声ストリーム1になると、その後

音声ストリーム 2 を選択したチャンネルに戻しても音声ストリーム 1 が選択されることで、す。

これを解決するには次のような方法が考えられます。

**(a) 音声ストリーム 1 でステレオ音声を送出する**

これは、聴取者のほとんどがステレオ環境であることと、番組もステレオ及びモノラルが中心であることを考慮した考え方です。音声ストリーム 1 でステレオ、2 でサラウンドを送出すれば、ほとんどの聴取者は音声ストリームを意識することなく、今までどおりの視聴方法で音量差のない番組聴取が可能となります。逆に、サラウンド聴取者は音声ストリームを選択する手間が必要になりますが、適正音量で聴取することが出来ます。この場合、「音声ストリーム 1 でサラウンド」という、放送局の運用ルール変更が必要となります。

**(b) 選択した音声ストリームを保持する**

もうひとつは、テレビに聴取者の選択した「チャンネル+音声ストリーム」を一時記憶（電源 OFF まで）させる方法です。チャンネルを切り替えても一時記憶した設定値で元の聴取環境に戻すことが出来ます。この方法はテレビセットメーカーの対応が必要となります。

**6-2 ARIB 規格の改定**

最後に ARIB STD-B21 規格を変更するという方法があります。つまり、係数 a を a=1 または削除することで、聴取時の音量レベル差、制作時のミキシングレベルの問題などを一挙に解決することが出来ます。ミキサーの立場からはこの方法が一番好ましいのですが、「今まで販売したテレビとの互換性がなくなる」という意見もあり、簡単には実現しないようです（ちなみに、米国や欧州のデジタル放送のダウンミックスには、係数 a に相当す

る部分がないので、このような音量レベル差の問題は発生していないようです）。

**7 本編と CM の音量レベル差を考える**

ここまではサラウンド番組単体の音量レベルについて考えてきましたが、次に実際の放送で発生するサラウンド番組（本編）と CM の組み合わせについて考えて見ます。

現状では「番組はサラウンドでも、番組の中の CM や番宣はステレオ」というのが普通です。このような状況で本編と CM の音量レベルの関係はどのようになるのでしょうか。

ここでは、ダウンミックス音声とステレオ音声を同等音量になるように制作したサラウンド番組（+3dB 制作）を例にしています。

サラウンド番組本編と CM の音声レベル・音声モードは以下の四つのパターンが考えられます。

パターン 1（図 7-1）は、現時点のサラウンド放送で一番運用されている方法です。この場合、ダウンミックス聴取者は本編・CM と同じような音量で聴取できますが、サラウンド聴取者は CM になると音量が小さくなります。

映像	本編映像1	CM映像など	本編映像2
音声レベル	本編サラウンド音声1 (+3dB)	CMステレオ音声 (0dB)	本編サラウンド音声2 (+3dB)
音声モード	サラウンド	ステレオ	サラウンド

図 7-1 音声モードパターン 1

パターン 2（図 7-2）は、1 と同様に音声モードを切り替えています。CM の音声レベルをサラウンドに合わせて大きくした場合です。この場合、サラウンド聴取者は本編・CM と同じような音量で聴取できますが、ダウンミックス聴取者は CM の音量が大きくなります。この方法では CM バンクで同じ CM をサラウンド用（+3dB）とステレオ用で別管理する必要があり、非現実的といえます。

映像	本編 映像1	CM 映像など	本編 映像2
音声レベル	本編 サラウンド音声1 (+3dB)	CM ステレオ音声 (+3dB)	本編 サラウンド音声2 (+3dB)
音声モード	サラウンド	ステレオ	サラウンド

図 7-2 音声モードパターン 2

パターン 3 (図 7-3) は、音声モードをサラウンドに統一した場合です。ただし、CM の音声は L/R のみ (ステレオ) でレベルは +3dB にします。この場合、CM の音声は L/R にしかありませんが (C/Ls/Rs/LFE は無音)、モードがサラウンドなのでダウンミックス機能が有効となります。このパターンではダウンミックス聴取者もサラウンド聴取者も、本編・CM と同じ音量で聴取できます。これもパターン 2 同様、同じ CM を CM バンクで別管理する必要があります。

映像	本編 映像1	CM 映像など	本編 映像2
音声レベル	本編 サラウンド音声1 (+3dB)	CM ステレオ音声 (+3dB)	本編 サラウンド音声2 (+3dB)
音声モード	サラウンド	サラウンド	サラウンド

図 7-3 音声モードパターン 3

パターン 4 (図 7-4) は現実的な方法ではありませんが、ステレオ CM の音声を擬似サラウンド化して使用した場合です。CM の音声レベルは本編と同じ音量になるように調整します (+3dB 化)。この方法もパターン 3 と同様、ダウンミックスもサラウンドも同じ音量で聴取することが出来ますが、当然 CM を別管理しなければなりません。また、擬似サラウンド化した CM 音声は、ダウンミックスが適正に行われるか (ステレオと同じように聴こえるか) 検聴する必要もあります。

映像	本編 映像1	CM 映像など	本編 映像2
音声レベル	本編 サラウンド音声1 (+3dB)	CM疑似 サラウンド音声 (+3dB)	本編 サラウンド音声2 (+3dB)
音声モード	サラウンド	サラウンド	サラウンド

図 7-4 音声モードパターン 4

以上、現実的なもの非現実的なものも含めて四つのパターンを、音声モードと音声レベルの関係から述べてみました。これらの問題も、先に述べたダウンミックス係数を 1 (a=1) にすることで解決することが出来ます。

## 8 放送方式による音量レベル差の問題

地デジを聴取していて気付かれると思いますが、NHK と民放では音量レベル差を感じる事が多々あります (NHK が大きい)。また、アナログ放送と地デジでも明らかに音量レベル差があります (アナログが大きい)。

このような (受信機も含めた) 放送規格の問題で発生する音量レベル差について整理してみました。これらは規格上発生する (であろう) 理論上のレベル差で、必ずしもこのようなレベル差が、実際の放送で発生するという意味ではありません。番組によってはそれほどレベル差を感じないこともあります。また、発生するレベル差も放送局により異なります (後述)。

### 8-1 アナログ放送とデジタル放送の音量レベル差について

地デジを受信している視聴者がわざわざアナログ放送を視聴することは少ないと思いますが、現在のサイマル放送は“こうなっている”ということを知ってもらうために、アナログ放送と地デジのレベル差について整理してみました。

民生用 AV 機器には相互接続のための規格があります (JEITA CP-1203 AV 機器の電氣的接続要件)。テレビのチューナ部基準出力レベルもこの規格の中で決められています。この規格では、アナログ放送の場合、

- ・100%変調信号を受信したとき 500mV の出力電圧

となっています。一般には基準信号 (0VU) の変調度を 63%としている放送局が多いので、これに換算すると 315mV となります。

一方、地デジなどデジタル音声の基準出力レベルは、

- ・ -12dBFS の信号を受信したとき 500mV の出力電圧

となっており、これを地デジの基準レベル -18dBFS に換算すると 250mV となります。その差は 2dB となり、テレビで同じ基準レベルの信号（音声）を受信しても、アナログ放送よりも地デジの方が 2dB 小さな基準出力レベルとなります。

また、民放の地デジ放送は -20dBFS を基準レベルとしているので、受信機の出力信号はさらに低く 200mV となり、アナログと比べ 4dB 小さくなります。

### 8-2 デジタル放送の音量レベル差

地デジなど、デジタル放送の基準レベルは ARIB 規格で -18dBFS となっています。これは NHK の基準レベルと同じですが、民放の場合は -20dBFS を基準レベルとして放送しているので、NHK と民放との間に 2dB のレベル差が発生することになります。したがって、民放よりも NHK の方が少し大きな音量になるのは当然と言えます。

以上を整理すると、アナログ放送の受信機基準出力レベルを 0dB とした場合、NHK のデジタル放送は -2dB、民放のデジタル放送は -4dB となります。ただし、アナログ放送では放送局によりレベル差がかなりあります。これは放送局により変調度に違いがあること、送出時に使用するリミッタ/コンプレッサなどにより変調度の動特性に違いが生じること、などが原因と考えられます。ここでは標準的な場合を例として上げてみました。

以上のような“基準レベルによる差”に、ダウンミックスによるレベル低下を加えると、アナログ放送の音声と民放のダウンミックス音声の差は 7dB にも達する場合があります。同じ地デジでも NHK のステレオ音声と民放の

ダウンミックス音声では 5dB の差が発生することになり、無視できない値となります。

## 9 米国と欧州のデジタル TV 放送

### 9-1 米国のサラウンド音声

米国のデジタル TV 放送は、音声規格として Dolby AC-3 が採用されています。AC-3 のステレオ用ダウンミックス式は、

$$L_o = 1.0 * L + clev * C + slev * L_s$$

$$R_o = 1.0 * R + clev * C + slev * R_s$$

clev : C のダウンミックス係数

slev : Ls/Rs のダウンミックス係数

となっており、日本の「係数 a」に相当する部分がありません。式から分かるように C チャンネルもダウンミックス係数が選択可能となっています。clev のダウンミックス係数を表 9-1、slev のダウンミックス係数を表 9-2 に示します。

cmixlev	clev
'00'	0.707 (-3.0dB)
'01'	0.595 (-4.5dB)
'10'	0.500 (-6.0dB)
'11'	reserved

表 9-1 clev 係数 (ATSC A52/B より)

surmixlev	slev
'00'	0.707 (-3.0dB)
'01'	0.500 (-6.0dB)
'10'	0
'11'	reserved

表 9-2 slev 係数 (ATSC A52/B より)

また、2005 年の規格改定で LFE チャンネルもダウンミックスに加えることが可能となりました。この場合の LFE ミックスは下式のようになっています。

mix LFE into left with (LFE mix level - 4.5) dB gain

mix LFE into right with (LFE mix level - 4.5) dB gain

ここで「LFE mix level」は、

$$\text{LFE mix level (dB)} = \text{LFE mix level code} + 10$$

で計算された値となります。

「LFE mix level code」は 5 ビット (0~31) で構成され、0 から 31 の値を持ちます。したがって「LFE mix level(dB)」の有効値は+10dB から -21dB となります。LFE をダウンミックスに加える場合は「mix LFE into～」式が固定値となっているので、LFE mix level で LFE チャンネルのレベル調整を行います。

### 9-2 欧州のサラウンド音声

一方、欧州の地上デジタル TV 放送 (DVB-T) の音声規格は MPEG-2 BC を採用しています (最新規格では AC-3, DTS, MPEG-4 AAC など使用可能)。MPEG-2 BC 使用時のダウンミックスは以下のようになっています。

```
downmixing_levels_MPEG2 ( )
{
  center_mix_level_on
  center_mix_level_value
  surround_mix_level_on
  surround_mix_level_value
}
```

center\_mix\_level\_on の値が “1” ならば、C チャンネルの音声ダウンミックスの対象となります。同様に surround\_mix\_level\_on の値が “1” ならば、Ls と Rs チャンネルが対象となります。

mix_level_value	Multiplication factor
'000'	1.000 (0.0dB)
'001'	0.841 (-1.5dB)
'010'	0.707 (-3.0dB)
'011'	0.596 (-4.5dB)
'100'	0.500 (-6.0dB)
'101'	0.422 (-7.5dB)
'110'	0.355 (-9.0dB)
'111'	0.000 (-∞ dB)

表 9-3 mix level value (TS-101-154)

mix\_level\_value は 3 ビットで構成され、表 9-3 のような dB 値を持っています。ダウンミックスの対象となった C チャンネル、Ls/Rs

チャンネルにこの係数を掛けて、L チャンネル / R チャンネルに加算し、ダウンミックスステレオ音声を作ることになります。

MPEG-2 BC の場合も「係数 a」に相当する部分は存在せず、原則としてダウンミックスによる音声レベルの低下は発生しません。また、mix\_level\_value の選択についてはオーバーロードを避ける値にする、との記述があり、ダウンミックスのレベル管理は制作者側で行うこととなります。

### 10 ドルビーボリュームについて

最後にテレビに実装する、新しい音量レベル差調整技術を紹介しておきます。この技術が普及するとテレビの音量レベル差問題もある程度解決できるかもしれません。また、テレビ側で自動的に音量制御が行われるとすれば、CM の音量競争の抑制につながる (音量を大きくしても効果がない) 可能性も考えられます。

#### 10-1 Dolby Volume とは

2007 年に発表された Dolby Volume は、テレビのチャンネル間音量レベル差、番組間音量レベル差や AV 機器の入力間音量レベル差などを、聴覚心理モデルを用いて DSP 処理し、音量レベルが一定になるようリアルタイムに制御するものです。Dolby Volume にはボリュームレベラーとボリュームモデラーという二つの機能が搭載されており、この二つの機能を組み合わせて動作します。

セットメーカーはライセンス生産された Dolby Volume DSP を実装してこの機能を利用します。2008 年 4 月に東芝から、世界初の Dolby Volume 搭載テレビが発売され、米国では AV アンプも何機種か販売されています。

#### 10-2 Dolby Volume の動作原理

音量レベル差の基本的な検出方法は、テレビなどの機器に設定する、デフォルト音量レ

ベルと入力信号レベルを比較することで行われます。

Dolby Volume は、音声帯域を複数の周波数バンドに分割して処理が行われており、バンド数が多いほど高品質な処理が可能となりますが、反面 DSP 処理の時間も長くなり、音声ディレイも大きくなります。DSP メーカーにより処理できるバンド数が異なるようですが（最新の DSP では 20 バンド処理が可能なものもある）、例えば 5 バンド処理まで可能な DSP の場合、3 バンド処理を選択しても一般視聴者には十分とのことで、この場合は気になるほどの音声ディレイも生じないようです。音声ディレイは DSP の性能アップと共に少なくなるので、よりいっそうの性能アップも期待できます。

### 10-3 ボリュームレベラー機能

この機能はチャンネル間、番組間、入力間の音量レベル差を解消し、これらを切り替えても一定音量で聴取できるように制御する機能です。ドルビーのホワイトペーパーには「音声信号の聴感上の音量を常時測定し、知覚された音声の音量レベルが一定になるよう、知覚型処理エンジンを使用してマルチバンドのゲイン補正をダイナミックに行う、強力かつユニークなソリューションです」とあります。

### 10-4 ボリュームモデラー機能

二つ目はボリュームモデラー機能です。人は聴取音量を下げてゆくと、知覚感度以下になる周波数帯を生じますが、これを知覚感度以下にならないようにレベル制御する機能です。つまり、等ラウドネス曲線（古くはフレッチャー／マンソン、現在はロビンソン／ダッドソンの曲線）を利用したレベル制御です。人間の聴覚は中域と比べ低域、高域の感度が低くなります。聴取レベルを下げてゆくと、中域の音は聴こえていても、低域や高域の音

がだんだん聴こえなくなります。これを適正に補正し、小音量でもバランスの良い音になるような制御を行っています。

以上述べた二つの機能は Dolby Digital や DD Plus のドルビーメタデータとも関連して動作します。ドルビーメタデータがあればそれを優先して動作するようになっています。

### おわりに

サラウンド番組を中心にテレビの音量レベル差問題を二回に亘って述べてきました。サラウンド番組についてはテレビ局の対策もあり、現在では一般聴取者からの「サラウンド番組の音が小さい」というクレームは少なくなっているものと思われます。しかし、現在の対応は根本的な解決策とは言えません。制作側の負担も少なく、視聴者も満足できる解決策が求められます。

二回連載となった本稿は、2008 年 11 月の技術委員会オーディオ部会で配布した技術資料を加筆修正したものです。

### 参考文献

1. ARIB STD-B21v4.7
2. ARIB STD-B31v1.7
3. ARIB STD-B32v2.1
4. Dolby Digital Professional Encoding Guidelines
5. JEITA CP-1203
6. ATSC A/52B
7. ETSI TS 101 154 v1.8.1
8. Dolby Volume White Paper J07/18114