

最新イマーシブ・オーディオ・フォーマット「DTS:X」

dts Japan 株式会社

山口 博紀

本稿では弊社の最新の3DサラウンドフォーマットであるDTS:Xに関して紹介させていただきたいと思います。DTS Digital Surround というフォーマットで弊社のサラウンドはスタートしました。6.1chをサポートしたDTS-ES、7.1chおよびロスレスを実現したDTS-HD Master Audioなどを経て現在のDTS:Xに至ります。

DTS:Xの説明の前に、そもそもどうしてマルチチャンネルなのかを少しだけ説明したいと思います。多くのコンテンツはステレオ2chで作られています。例えば一般的なCDなどの音楽ですと、ヴォーカルが左右の真ん中にあり、さまざまな楽器が演奏しているような配置で左右のスピーカーの間に配置されているような形です(図1)。これはもちろん素晴らしく完成された世界なのですが、現実の環境では壁や天井からの反射の音や、ライブでしたら周りの観客の声や拍手は横や後ろからも聞こえます。つまり、より現実に近い音の世界を再現するためにマルチチャンネルが使われるのです。



図1 ステレオの音像イメージ例

さらに、「生演奏を聴く」場合と「録音されたものを聴く」場合の理想的な部屋の違いも考慮する必要があります。少しおおざっぱな言い方になりますが、音を扱う部屋として生演奏を聴く部屋は適度な響きを持った部屋が適し、録音された音を聴く部屋は響きの少ない部屋が適しています。なぜかといいますと演奏時は楽器の音が反射することでよりよい響きを生むのに対し、録音された音には基本的にはこの響きがすでに入っているため余計な響きがない状態で聴いたほうがよいからです。このように考えますと、録音された音を聴く場合には、反射音や環境音などは後ろや横にスピーカーを置いてそこからその成分を鳴らすのが形としてはより理想的といえます。映画の音響制作においては、自分が映画の世界にいるような体験をしてほしいため、場面によつ

て後ろから声がしたり、音が鳴ったり、はたまたヘリコプターが頭上に来たりと自分の周り 360 度にわたって様々な音が聞こえるように演出します。

また、マルチチャンネルの効果には別の側面もあります。皆さん、理想的には生の演奏を聴きたいですね。例えばあるバンドが演奏をしています。ヴォーカル、ギター、ドラム、ベース、キーボードによる組み合わせだとすると、それぞれの楽器が鳴っています。これをスピーカーに置きかえると、最低でも 5 つのスピーカーを使って音を鳴らしたほうがより本物に近くなるはずですが。ドラムなどは実際にはもっと細かく分けたほうがよいのかもしれませんが。ステレオの場合はこれをある意味では無理やり 2ch にミックスダウンしています。複数の音を少ないスピーカーに押し込むため、生演奏でそれぞれの楽器で別々に鳴っている場合に比べると、どうしても聞こえにくい音が出てきます。世の中に出てくる作品はミキサーの方々の努力と技術でこの問題をうまく調整して、2ch でもバランスよく音が聞こえるようになっています。ですが、スピーカーの数を増やすことで、それでもどうしても聞こえにくくなる部分やつぶれてしまう部分を分散させてよりよく聞こえるようにできます。もちろん単純に自分の横や後ろでギターやベースを鳴らすという意味ではなく、(演出としてはあえてそうする場合もありますが) 例えば楽器の響きの部分をサラウンドに持ってきたり、サラウンドチャンネルと組み合わせて LR より少し外側にさらに音を広げたりなど、様々な可能性が考えられます(図 2)。

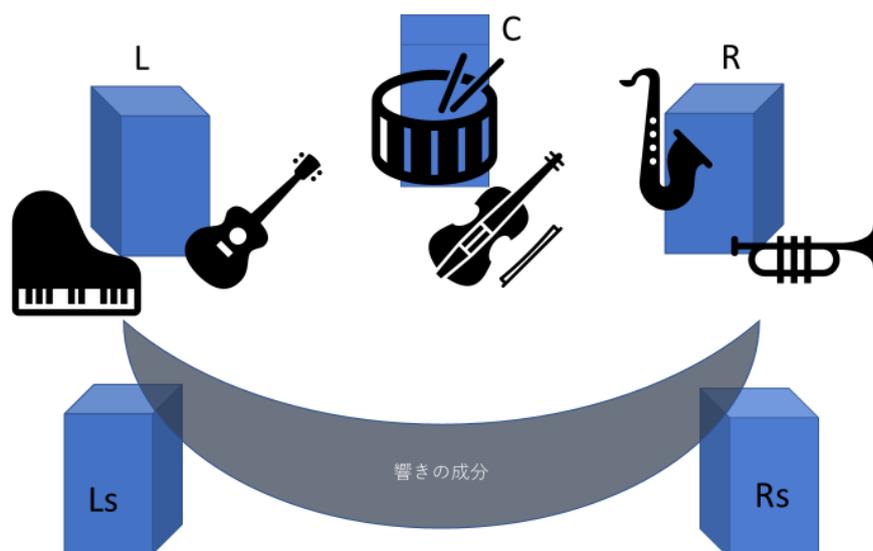


図 2 サラウンドの音像イメージ例

昨今のトレンドは、サラウンドというよりはハイレゾでしょうか。CD は 44.1kHz/16bit というデータ量ですが、これが 96kHz/24bit のハイレゾになると情報量としてはおよそ 2.9 倍になります。2ch という手軽で使いやすい環境の中でハイレゾはチャンネル数を 3 変えずにデータ量を増やすことで精度を高くし、さらに対応した再生機器の精度・表現力を上げることで 2ch のままでもより高品位な音が聴けるようにするというアプローチです。

それに対してサラウンドはスピーカーの数を増やします。同じようにデータ量で考えると 2ch が 5.1ch(実質は 6ch 相当)になるため、データとしては 3 倍になります。チャンネルを増やしてデータ量を増やす、つまり、音の聞こえる方向を増やすことに使うことでマスキングを減らし、

かつ、空間的な定位の範囲を広げることで高品位な音を届けるというアプローチになります。もちろん、ハイレゾかつサラウンドというコンテンツもあります。これらがコンテンツの特性にあわせて使い分けられているというのが現在で、とりわけ、映画の世界ではサラウンドのアプローチが使われていると言えます。

DTS:X では横や後方だけでなく、高さ方向にもスピーカーを増やしてリスナーを半球状に取り囲むことで今までにない臨場感を出すことができます。そのため従来の「サラウンド」に対して「イマーシブ」という言葉をつかっています(図 3)。イマーシブは日本語では「没入」と訳されることが多いですが、実際は若干ニュアンスが違います。英英辞典などを見るとわかりやすいのですが、単純に物理的にまたはコンピュータシステムの的に 360 度にわたって音や映像が隙間なく



図 3 イマーシブなスピーカーの配置例

埋められている状態をいいます。例えばプラネタリウムや VR ゴーグルのような形です。没入というと海中のような、より閉鎖的な空間だったり、精神的にも深く入り込む・沈み込むようなニュアンスが含まれます。素晴らしいコンテンツでユーザーが結果的にそのようになることはもちろんありますが、「イマーシブ」という言葉そのものにはそこまでの強いニュアンスはないです。

イマーシブの利点は単にスピーカーが増えた、上方から音が出せるというだけではありません。前述の通り、さまざまな音のバランスがとりやすくなるということも大きなポイントになります。例えば、映画を作る際に必要な大きな 3 つの要素 DME(Dialog, Music, Effect)がありますが、どうしてもこれらはお互いにぶつかってしまう場合もあり、すべてを効果的に聴かせるのに苦労する場合があります。しかしながらイマーシブなフォーマットを使うことで、より様々な要素を分散させることができるため、それぞれを引き立たせるようなミキシングがより容易にかつ効果的になります。

DTS:X は「オブジェクトオーディオ」にも対応しています。従来のフォーマットの考え方は「どのチャンネルからどのような音を出すか」というのをあらかじめ決められたチャンネルの配置に従って音を作るという形でした。それに対してオブジェクトオーディオでは「どの音をどのスピーカーから出すか」を再生時に決めます。再生時に決めるので例えばスピーカーの数や配置が少しずつ違って、一つのフォーマットで可能な限り狙った位置から音を出したり動かしたりができます。映像に置き換えると映画と 3D のゲームのような違いでしょうか。映画ですと基本的に

はあるシーンを上や反対側から見ることはできません。それに対して 3D のゲームは絵をリアルタイムにレンダリングしているので、ユーザーの見た方向の絵を瞬時にレンダリングできます。音に関してそれに近いリアルタイムのレンダリングをするのがオブジェクトオーディオの特徴です。

オブジェクトのレンダリングの基本的な考え方は VBAP(Vector Base Amplitude Panning)に基づいています。聞きなれない言葉なのですが、実際は皆さんが日常的に体験している話です。例えばステレオスピーカーで音楽を聴いて、スピーカーの左右の真ん中あたりで音を聞きます。(理想的には L スピーカー, R スピーカー, リスナーの頭の 3 点を結んだときに正三角形になるような位置関係) この時音は実際には左右のスピーカーからしか出ていないのですが、ヴォーカルは左右のスピーカーのちょうど真ん中から聞こえます。これが VBAP の基本的な考え方で、ある

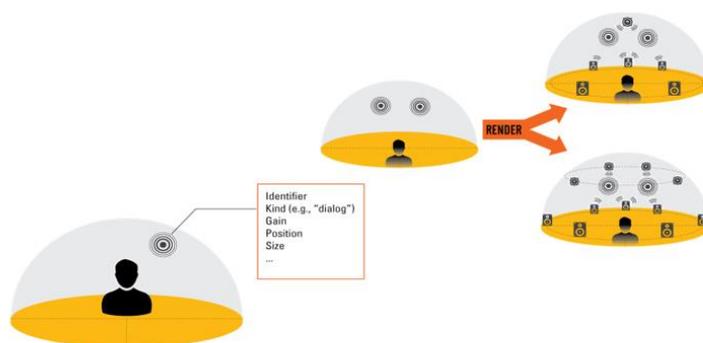


図 4 オブジェクトオーディオ

音を左右にどう振り分けるかを調整することでスピーカーの間であればどの位置にでも音を置くことができます。DTS:X ではリスナーを取り囲むように 360 度の様々な位置にスピーカーがあるので 3 つのスピーカーで作られる 3 角形をいくつも作ることが可能です。その中の好きな位置に音を定位させたり動かしたりしてオブジェクトオーディオを実現します(図 4)。

オブジェクトオーディオの別の使い方として「インタラクティブ」な使い方が考えられます。その一つが「ダイアログコントロール」という機能です。これはコンテンツを作る方がそのように作った場合にのみ有効なのですが、セリフのオブジェクトのみを最大+6dB 持ち上げることができます。夜など小さい音量で映画を見る時にセリフが聞こえにくい場合があります。今までの AV アンプですとセンターチャンネルの音量を大きくするような機能がありましたが、センターにはセリフ以外の音も入っているのでそれらもすべて音量が上がります。しかし、ダイアログコントロールはセリフというオブジェクトのみの音量を上げるため、よりクリアにセリフを聞きやすくすることができます(図 5)。

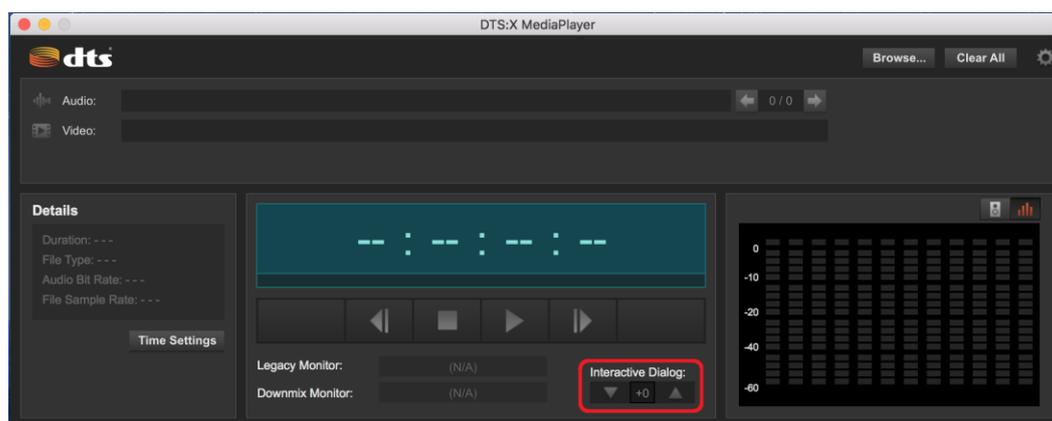
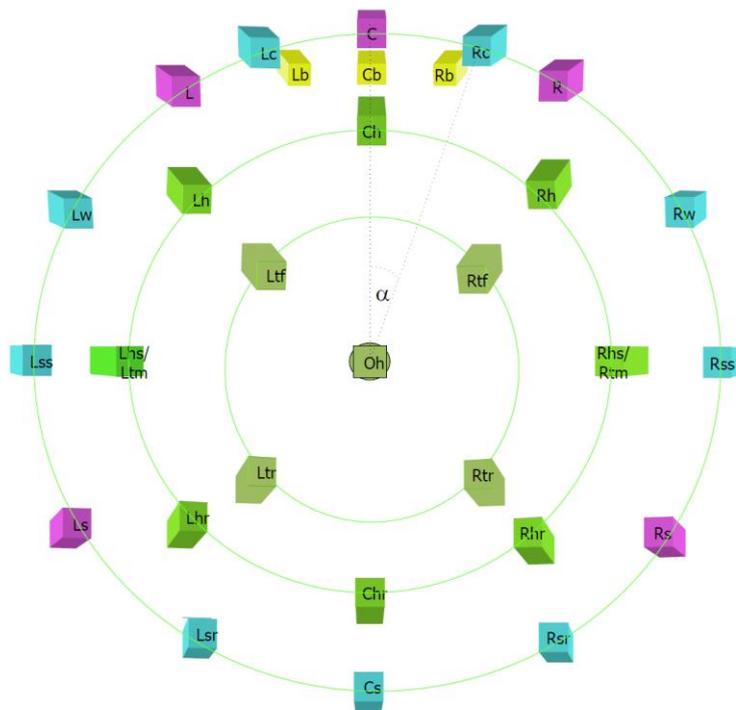


図 5 ダイアログコントロール

DTS:Xの利点としてスピーカーの配置の自由度が高いという点があります。イマーシブに対応するためハイトやトップのスピーカーがあることは推奨ですが、ない場合でもオブジェクトは動作しています。極端なことを言いますと、2chの環境でもちゃんとダイアログコントロール(つまり、オブジェクト)は動きます。また高さ方向のスピーカーの配置もハイト、トップまたはそれらの組み合わせでも問題なく動作しします。自由度の高さはDTS:Xの魅力といえます(図6)。



- 2.0ch : L、 R
- 5.1ch : L、 R、 C、 Ls、 Rs
- 7.1ch : L、 R、 C、 Lss、 Rss、 Lsr、 Rsr
- 7.1.4ch : L、 R、 C、 Lss、 Rss、 Lsr、 Rsr、 Lh、 Rh、 Lhr、 Rhr

図 6 DTS:X が想定するスピーカー配置

DTS:Xは過去のDTS Digital Surround、DTS-HD Master オーディオの延長線上に存在します。つまり、これらのフォーマットと互換性を持っています。例えばDTS:X収録のディスクをDTS Digital Surroundしか対応していないサウンドバーに接続しても、DTS Digital Surround互換の部分が動作をして音が鳴ります。オブジェクトや、高さ方向の成分が失われることもありません。同様にDTS-HD Master Audio対応のAVRに接続しても、ロスレスのDTS-HD Master Audioとして再生されます。つまり、DTS:Xとしてストリームを作ってしまうと過去のDTSのあらゆるデバイスで互換性を保ったまま問題なく再生ができるということです(図7)。

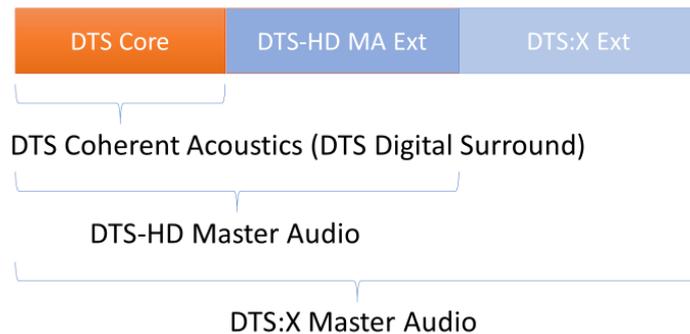


図7 DTS:Xビットストリームの構成

DTS:Xはもちろんロスレスにも対応しております。ディスクに収録する際にデータ量を抑えたい場合はロッキーの、高品質な音を収録したい場合はロスレスというような選択ができます。どちらの場合でもオブジェクトの動作等に制限はありません。

さて、DTS:Xは非常に強力なフォーマットですが、新しいフォーマットでもあるためどうしてもコンテンツの普及には時間がかかります。また、DTS Digital SurroundやDTS-HD Master Audioで作られた様々な名作をイマーシブな環境で聴きたいという方も多いと思います。そのような要求にこたえるためDTS:Xの中にはNeural:Xと呼ばれるアップミキサーが入っています。簡単に言いますと、この機能は2.0/5.1/7.1chといったコンテンツを解析し、上方(2.0から作る場合はそれにプラスして後方)に持ってくるほうがより効果的な音を見つけて分離し7.1.4chに拡張する技術です。DTS:Xの中に含まれている裏方のような存在のため、あまり表に出て話題になることは少ないのですが、非常に強力なアップミキサーです。例えば2chの音楽をアップミックスします。弊社の過去のものも含めて、ほとんどのアップミキサーは音が不自然に広がってしまい音場が変わってしまうように聞こえますが、Neural:Xの場合は、ヴォーカル回りなどの定位がほとんど変わりません。ある意味ではON/OFFの違いが一瞬では分かりにくいのですが、よく聴くと前方中心部分の定位はほとんど変化がないのに音全体としては後方や上方に気持ちよく広がり、響いている状態に変わっています。

また5.1chや7.1chの映画をNeural:Xで実際にハイトのどのチャンネルからどのタイミングで音が出るか聴いてみてください。非常に面白い動作をします。ここでは詳しく書きませんが、実はアップミックスだからと言って常に上方から音が出るわけではありません。是非、さまざまなコンテンツを聴いてみてください。

このように DTS:X は次世代フォーマットとして様々な特徴を持ったフォーマットです。対応する再生機器やディスクも増えており、国内でも対応した映画館が少しずつ増えてきています。DTS:X にますます注目をしていただければと思います。



執筆者プロフィール

山口 博紀 (やまぐち ひろき)

北海道大学大学院システム情報工学専攻卒業

2001年より外資系半導体企業にて R&D,

FAE, セールスを担当

2014年より dts Japan(株)にて

フィールド・アプリケーション・エンジニア担当

DSP/SoC への信号処理の実装や規格関連のシステムおよび

アルゴリズムの開発、実装、技術サポートに従事

Audio Engineering Society 会員