

CELLレグザ™の高音質化技術

High-Performance Audio Technology of CELL REGZA™

桑原 光孝

竹内 広和

■ KUWABARA Mitsutaka

■ TAKEUCHI Hirokazu

デジタルハイビジョン液晶テレビ (TV) “CELLレグザ” は、単に高機能だけでなく、画質、デザイン、機能などあらゆる観点から、最高の質の実現を目指している。しかし従来の薄型TVの開発では、形状面の制約から高音質を追求することが困難であった。

そこで東芝は、音質最優先という観点で、これまでのTVとはまったく異なるオーディオシステムを開発した。新開発のスピーカシステムとマルチチャンネルデジタルアンプにより本格的なオーディオシステム並みの高音質を実現するとともに、CELLプラットフォーム™を活用したコンテンツ適応音質制御により再生するシーンに最適な音声処理をリアルタイムに行うことができる。

As the flagship high-definition (HD) TV of the REGZA™ lineup, the CELL REGZA is aimed at realizing not only high performance but also the highest quality from every point of view, including image quality, design, and functionality. However, it was difficult to achieve high-quality sound due to the dimensional restrictions in the development of flat-panel HDTVs.

To rectify this situation, Toshiba has developed an audio system with a new innovative concept giving highest priority to the sound quality. Almost the same level of sound quality as in a high-grade audio system is realized by a newly developed speaker system and multichannel digital amplifier. Moreover, the CELL Platform™ at the heart of the CELL REGZA system incorporates a content-based adaptive audio enhancement control function that provides real-time sound processing suitable for movie scenes.

1 まえがき

薄型TVの開発では、従来、デザインが重視されて機構的な制約を多く受けることから、高音質という観点でのアプローチが十分できない状況であった。これは市場の薄型TV全般に言えることであり、AV評論家の薄型TVの音に対する評価はどれも厳しい。

この課題を解決するため、東芝は、CELLレグザの開発にあたり、TV用スピーカというアプローチではなく音質最優先という観点で、従来とはまったく異なる規模、性能、及び音質を備えた、ハードウェアとしてのオーディオシステムを開発した。

また、TV番組には、音楽やスポーツのライブ放送、ニュースなど様々なコンテンツが存在する。ユーザーの利便性を向上させるため、これらのコンテンツをリアルタイムで解析して再生するシーンに応じて適切な音質に自動制御する、コンテンツ適応音質制御技術も開発した。

ここでは、音質を改善するハードウェアとしてのアプローチと音質決定のプロセス、及びCELLプラットフォームを活用した音質制御について述べる。

2 オーディオシステムの開発方針

CELLレグザ オーディオシステムの開発方針を以下に示す。

- (1) 要素部品の吟味 スピーカユニットや電子回路部品など要素部品の性能が良くなければ、結果の到達点は低くなる。CELLレグザでは、スピーカユニットだけでなく、電子回路部品に対しても音質的に優れた素材を投入する。
- (2) 優れた基本性能の実現 周波数特性に代表される基本性能を高めることで、要素部品の性能や構造の不備をデジタルイコライザなど電子回路の補正なしでも優れた音質が得られる設計とする。
- (3) 求める音に対する明確な指針 優れた音質の実現のためには、音作り指針の明確化がもっとも重要である。CELLレグザはTVであり、映像も音声もシステム内で完結する。AV再生というのは単に映像と音声が良いだけでなく、両者がそれぞれを引き立てるようにバランスさせることが重要である。画質に対する“感動を与える王道の画質”という明確な指針は、そのまま音にも当てはめることができる。CELLレグザの圧倒的な映像ダイナミックレンジを引き立てるには、オーディオもダイナミックレンジの再現に優れ、ハイスピードで鮮烈であることが重要であり、これを音質の指針とした。
- (4) CELLレグザだけのかつてない機能 TVの音が軽視されてきた背景に、映像だけをTVで再生し音声は外部の優れたオーディオシステムに任せる、というAVシステム構築の傾向がある。CELLレグザのオーディオシステム

がどのように優れていても、数百万円のオーディオ機器の音質レベルを期待するユーザーには使用されない状況になる可能性がある。CELLレグザ オーディオシステムでは、そういう高級オーディオ機器との親和性を高める機能を搭載することを開発方針に加えた。

3 開発アイテム

3.1 スピーカユニット

スピーカの構成は、ある程度の小型化と高音質化を両立させるためウーファ（低音域用）とツイータ（高音域用）から成る2ウェイシステムとした。ウーファとツイータは専門メーカーとの共同開発の形をとった。

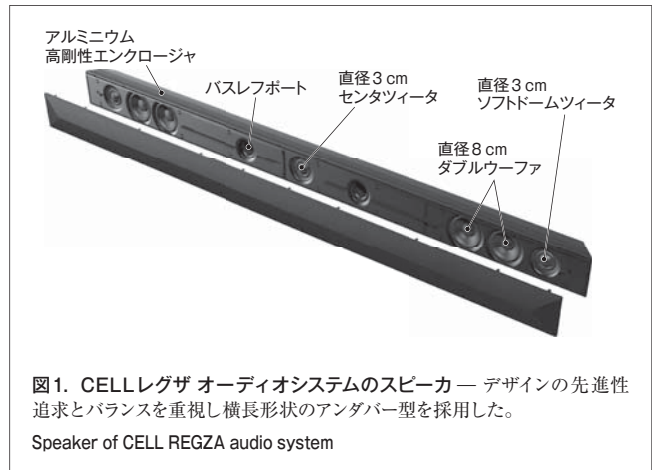
ウーファとしては、TV用ではなくハイファイ（HIFI）用として使える優れた小口径ウーファの実現を目指した。フレームは開発費を抑えるため既存の製品を流用し、この段階で口径を8 cmとした。CELLレグザの画質に見合った音質を実現するため、磁気回路や、コーン紙、エッジ、ダンパなど、スピーカを構成する要素を吟味した。その結果、2種類のコーン紙を張り合わせた多重抄紙コーン、ネオジウムマグネットを使った反発型磁気回路、耐久性に優れた大振幅に耐えられるゴム製エッジ、音質に固有の癖の少ない綿ダンパなどを投入することでウーファを完成させた。

ツイータもHIFI用ユニット開発という目的から、織物による布製の振動板を搭載するソフトドームツイータを採用した。金属製の振動板によるハードドーム型では高域特性が伸びすぎ、また、可聴帯域内に癖の出ることが多い。あらゆる音に柔軟に対応するためには、低い周波数から使用可能な優れた大口径ソフトドームツイータが最適と判断した。マグネットがフェライト又はネオジウムの2種類の試作品を評価し、ウーファとの質感統一に優れたネオジウム製を採用した。クロスオーバー周波数は、ツイータのひずみ特性を配慮して3 kHzに設定した。

3.2 エンクロージャ

スピーカの基本構造は、音質の追求という観点からTV本体と完全に分離する独立型にした。スピーカユニットを搭載するきょう体（エンクロージャ）を独立型にすることで、スピーカ設計だけで音質を管理できる。デザインの先進性追求とバランスを重視し、エンクロージャにはアンダバー型を採用したため、8（高さ）×8（奥行き）×130（幅）cmの横長形状になった。材質は、強度の確保と加工の容易さ、特別な金型が不要な点を考慮し、厚さ3 mmのアルミニウム引き抜き材を主材とし、高耐入力と低音域の再生能力を確保するためダブルウーファとした。

また、エンクロージャ形式には癖が少なく、かつ優れた低域特性が得られるバスレフ方式を採用し、片チャンネル当たり2Lの容量を確保して60 Hz（-10 dB）の低音域再生能力を実



現した。このアンダバー型のアルミニウム高剛性エンクロージャには、片チャンネル2機の直径8 cmのダブルウーファと1機の直径3 cmのソフトドームツイータ、更に3.3.4項で述べるセンタスピーカモードで使用する直径3 cmのセンタツイータを含めて、合計7機のスピーカユニットを搭載した（図1）。

3.3 エレクトロニクス部

3.3.1 アンプ構成

CELLレグザ オーディオシステムでは、限られたアンプパワーでの大出力確保と、コイル（L）とコンデンサ（C）で構成されるLCネットワークを廃することで高音質化を図るため、当社で初めてTV用帯域分割型マルチアンプ方式を採用した。ウーファ用に20 W+20 W、ツイータ用に10 W+10 W、そしてセンタツイータ用に20 Wモノラルとしてデジタルアンプを3機搭載した。

マルチアンプ方式にすることで、実際の音楽再生では30 Wのアンプと等価の再生が可能であり、一方で各スピーカの耐入力は20 W以下ですむ。ツイータ用のアンプパワーを10 Wとしたのは、ウーファとの能率の違いからウーファ用の20 Wと等価の音圧が確保できるからである。また、センタツイータには20 Wのパワーを供給し、センタモード時に能率が上がるウーファに対応した。このパートタイムパワーの配慮により電源の総合容量を60 Wに抑えることができる。

3.3.2 チャンネルデバイダ

マルチアンプを構成するうえで音質の要となるチャンネルデバイダは、可能な限り急しゅんなフィルタを用いて、ウーファとツイータを理想に近い形で駆動することを重視した。検討の結果、優れた特性を持つ有限インパルス応答（FIR）フィルタを使ったリニアフェーズ型を採用した。このフィルタでは、クロスオーバーポイント（2.96 kHz）の位相ずれが発生せず、ハイパスフィルタ（HPF）は-200 dB/oct以上の極めて急しゅんな遮断特性を持っている（図2）。なお、リニアフェーズフィルタはDSP（Digital Signal Processor）内にソフトウェアとして実装した。

3.3.3 ジッターダクシオン・アップサンプリング回路

HIFIオーディオでは音質を左右する要因としてマスタクロック

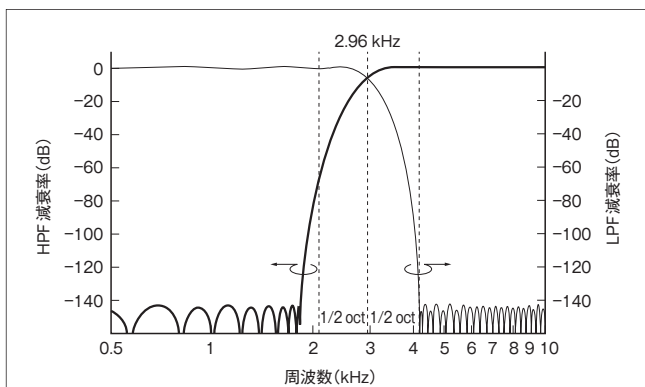


図2. リニアフェーズフィルタの特性 — FIRフィルタを用いたリニアフェーズフィルタは、クロスオーバーポイントの位相ずれが発生せず、急しゅんな遮断特性を備えている。

Characteristics of linear phase filter

クジッタに着目し、それに対応した位相ノイズの少ない水晶発振器や、TCXO (温度補償型水晶発振器)、OCXO (恒温槽付き水晶発振器)などを応用した製品も存在する。CELLレグザでは、新たにジッタリダクションICを投入してシステムクロックのクリーンアップを実施し、デジタルアンプ及びDSPにクロックを供給している。

また、DSP内で入力されるすべてのリニアPCM (パルス符号変調)フォーマットをサンプリング周波数192 kHz/量子化ビット数24ビットにアップサンプリングしてデジタルアンプに伝送し、各デバイスの持つ能力を最大限に発揮するよう配慮した。

3.3.4 センタスピーカモード アンダバー形式のスピーカ構造に期待できる最大の効果は、TV画面との親和性を向上させたセンタスピーカモードの搭載である。マルチアンプ構成を採用したCELLレグザオーディオシステムではこの対応が容易に実現でき、また外部に高級オーディオ機器を配置したユーザーにとってもこのシステムがマルチチャンネル用のセンタスピーカとして活用可能である。

なお、センタチャンネルモードでは4機のウーファとセンタツイータがモノラルスピーカとして動作し、左右のツイータは動作しない。HDMI[®](注1)によりアンプの起動を感知し、自動的にセンタモードとステレオモードを切り替える機能も搭載した。

3.3.5 音質チューニング 音質のチューニングについては、電子回路補正なしで基本音質を確保し、その後、デジタルイコライザによる電子回路補正を加えて最終的な音質を実現した。

デジタルイコライザではピークディップの補正を実施したが、過度の補正は著しい音質の劣化につながる事がわかり、補正量は最大3 dBに抑えた。DSP内蔵のバイカッドフィルタによりピンポイントの調整を加えて最終的な状態とした。

(注1) HDMI及びHigh-Definition Multimedia Interfaceは、HDMI Licensing, LLCの商標又は登録商標。

4 コンテンツ適応音質制御

この制御技術では、受信中のオーディオ信号をリアルタイムに解析することで、様々なコンテンツに対して現在のシーン状態を検出し、その結果に基づいて音質を制御する。例えば音楽シーンの場合、サラウンドや低域強調 (Bass-Boost) といった音響効果を生かして広がりのある臨場感豊かな音質設定にする。またトークシーンの場合には、イコライザなどにより音声明瞭 (めいりょう)度を向上させる。このように、それぞれのシーンに応じて適切な音質設定に自動制御することで、コンテンツに合った効果的な演出を実現できる。

一方、コマーシャルメッセージ (CM) では、番組本編と比較してあらかじめ音響効果を強めたりする場合があります。CMに対して番組本編と同様な音質制御を適用すると過度な音響効果になることが懸念される。過度の音響効果を緩和するため、この技術では、先のシーン検出に加えてリアルタイムにCMを検出する処理も備え、番組本編だけに効果的な音質制御を行うとともに、CM時には演出効果を抑えた聞きやすい音質制御が可能になっている (図3)。

それぞれの検出技術の特長を、次に述べる。

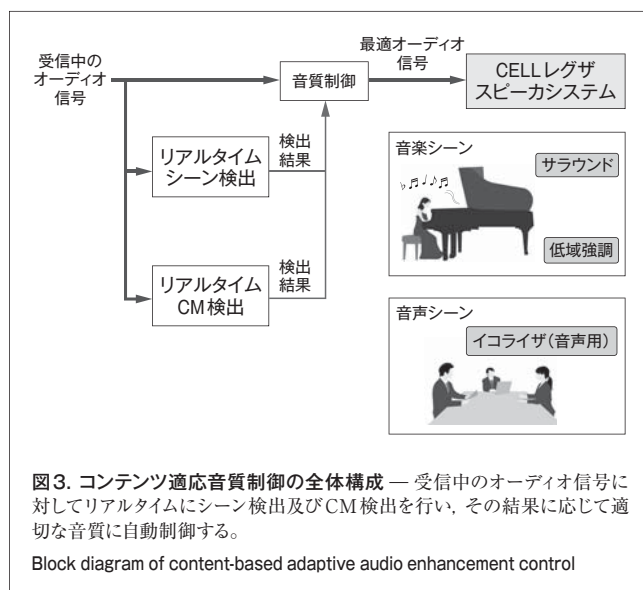


図3. コンテンツ適応音質制御の全体構成 — 受信中のオーディオ信号に対してリアルタイムにシーン検出及びCM検出を行い、その結果に応じて適切な音質に自動制御する。

Block diagram of content-based adaptive audio enhancement control

4.1 リアルタイムシーン検出

シーン検出処理は、受信中のオーディオ信号の音響特性をリアルタイムに解析し、音楽と音声との類似度をそれぞれ“音楽レベル”と“音声レベル”として検出する。検出には、音楽と音声の間で違いが出やすいステレオ信号での左右チャンネル間のパワー比などの時間域特性や、スペクトル変動などの周波数域特性といった音響特性を統計的に分析して得られる特徴パラメータに基づいて、類似度を算出し、シーン判定を行う。

純粋な音楽シーンや音声シーンは、この方法で比較的精度

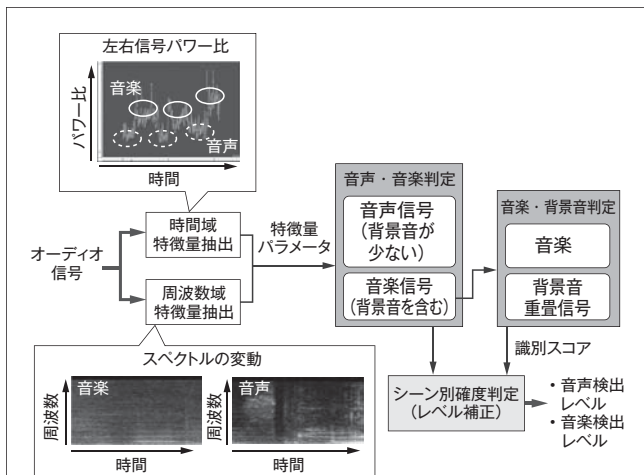


図4. リアルタイムシーン検出の仕組み — 音楽と音声との音響的類似度及び背景音の影響を考慮した2段階の判定に基づいて音種別(シーン)ごとの検出レベルを算出する。

Mechanism of real-time scene detection

よく検出できる。しかし、実際のTV放送信号では拍手や、歓声、BGM(Back Ground Music)といった背景音が重畳されていることが多く、これらの影響で誤判定や判定のばたつきが発生するため、その判定結果を音質制御に反映させようと聴感上で違和感が生じるという問題がある。

そこで、背景音の影響を緩和するため、音種別の判定を2段階に分けた(図4)。最初の判定では、背景音の少ない音声信号と背景音を含む音楽信号に大きく分類する。2段階目の判定では、背景音を含む音楽信号を更に音楽信号と背景音が比較的大きく重畳されている背景音重畳信号に細分化する。特に2段階目の判定では、ベースとなる楽器の低周波成分が最近の音楽では特定の帯域に集中しやすいことを考慮して背景音と音楽との識別精度向上を図り、背景音に対するロバスト性を高めている⁽¹⁾。

また、シーン判定結果のばたつきを軽減するため、音種別判定で得られる識別スコア(類似度)に基づいて音種別ごとの検出レベルを多値化し、識別スコアの継続性や背景音の強さを考慮してレベル補正することで、検出結果の安定化を実現している。

4.2 リアルタイムCM検出

国内のTV放送のCMでは、一般的にCM区間が15秒単位であることが多く、またCM境界の前後には無音が入挿されている。そこで、まず受信中のオーディオ信号の無音を監視し、無音検出後に受信した音声信号の音響特徴量とデータベースに蓄積されたCM冒頭の音響特徴量(約4,000件)を比較することで、CMと番組本編を瞬時(数百ms以内)に識別している。ただし、未知のCMに対しては、CM終了後に冒頭の音響特徴量をデータベースとして保存し、以降の識別に利用している(図5)。

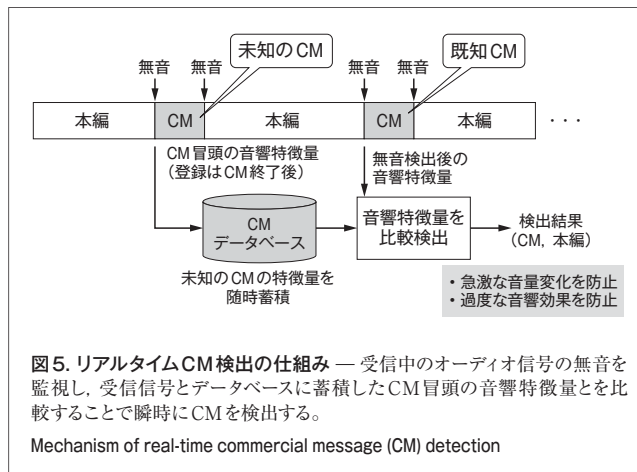


図5. リアルタイムCM検出の仕組み — 受信中のオーディオ信号の無音を監視し、受信信号とデータベースに蓄積したCM冒頭の音響特徴量とを比較することで瞬時にCMを検出する。

Mechanism of real-time commercial message (CM) detection

5 あとがき

CELLレグザ オーディオシステムは、画期的な回路や、DSPの開発、優れたスピーカシステムなどを導入することで、非常に優れた音質を実現できた。一体型のTVで優れた音質が実現できることは、高級オーディオシステムとの組合せとは異なる感動につながる。また、コンテンツ適応音質制御により、従来はユーザーがマニュアルで操作していた音質設定を、コンテンツに応じてきめ細かく自動的に行うことが可能となり、利便性と音質向上の両立を図ることができた。

今後は、より多様なシーンに対応できるよういっそうの性能向上に取り組み、更なる高音質化と他機種への展開を図っていく。

文 献

- (1) 米久保裕, ほか, “背景音にロバストな音声・音楽信号の識別方式の検討”, 第7回情報科学技術フォーラム (FIT2008), 神奈川, 2008-09, 電子情報通信学会, 情報処理学会, 2008, p.185-186.



桑原 光孝 KUWABARA Mitsutaka

東芝デジタルメディアエンジニアリング(株) デジタルメディアグループ デジタル機器開発技術担当チームマネージャー。DVD及びTV機器の設計・開発に従事。Toshiba Digital Media Engineering Corp.



竹内 広和 TAKEUCHI Hirokazu

ビジュアルプロダクツ社 コアテクノロジーセンター エンベディッドシステム技術開発部主務。オーディオ信号処理技術の開発に従事。Core Technology Center