

音のデザイン

木村 勝*
堀田 厚*

Designs of Sound Quality

Masaru Kimura, Atsushi Hotta

要 旨

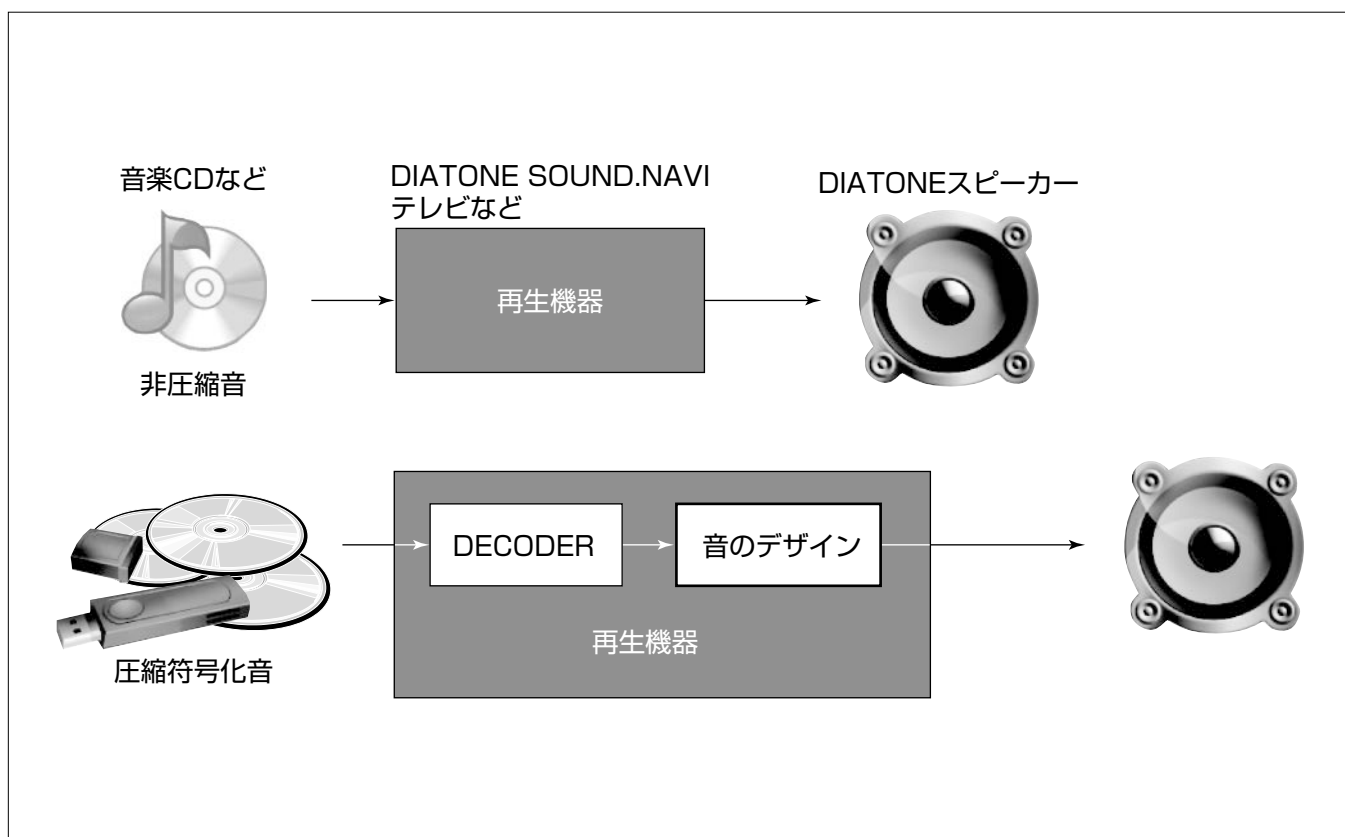
音楽の圧縮符号化技術は、地上デジタル放送や音楽配信サービス、DVD(Digital Versatile Disk)のようなパッケージ等、広範囲に使用されるようになってきた。

三菱電機では、“DIATONE”スピーカーに代表されるように、再生音を忠実に再現する技術について研究開発を続けてきた。しかし、圧縮符号化された音は、CD(Compact Disk)などの非圧縮音と比較すると一般に音質が劣化しているため、高忠実度再生技術だけでは高音質を実現することが困難になってきている。このため、音質・特性を積極的に変更する“音のデザイン”についても研究開発を行っている。本稿では、当社の音のデザインを支えている主要技

術の一つである圧縮符号化音の高音質化技術について技術的な内容と特長について述べる。

圧縮符号化音の高音質化技術は、圧縮符号化の音質劣化傾向に対応した2つの処理から構成されている。1つ目はステレオ回復処理であり、ステレオ左右信号の相関を利用し、豊かで立体感のあるステレオ音像を再現する。2つ目はBWE(Band Width Extension)であり、符号化で失われた高域成分を推定・再現することで、歯切れの良い素直な音を実現する。

主観評価の結果、圧縮符号化音を大幅に上回り、CD原音同等、又はそれ以上の音質が再現できることを確認した。



オーディオ技術のイメージ

当社が取り組んでいるオーディオ関連技術のイメージを示している。音楽CDなどの再現では、忠実な再現が必要であり、DIATONEスピーカーに代表される忠実再生技術に取り組んでいる。一方、音質劣化がある圧縮符号化音に対しては、音質・特性を積極的に加工し、あるべき姿にデザインすることで、CD以上の高音質化、又はユーザーの嗜好(しこう)にあった音の再生の実現を目指している。

1. ま え が き

音楽の圧縮符号化技術は、地上デジタル放送や音楽配信サービス、DVDのようなパッケージ等、広範囲に使用されるようになってきた。当社では、再生音を忠実に再現する技術だけでなく、音質が劣化する圧縮符号化音に対しては、再生音の音質・特性を積極的に加工し、あるべき音にデザインする技術について研究を行っている¹⁾。

本稿では、音のデザインを支えている主要な技術の1つとして、圧縮符号化音の高音質化技術について述べる。

2. 圧縮符号化の特性

2.1 音質劣化の傾向

圧縮符号化は、聴覚特性を利用した圧縮技術によって、大幅な情報圧縮を行っても、大きな音質劣化がないことが知られている。しかし、劣化が全くないわけではなく、注意深く聞くと原音とは異なる音質であったり、ステレオ音像の広がり減少したように聞こえたりする等の音質劣化が知覚されることが多い。圧縮符号化を行ったオーディオ信号には、次の特徴が現れる。

- (1) 高域成分の欠落
- (2) 左右差信号の欠落

これらの特徴が音質劣化の要因となり得ることを述べる。

2.2 高域成分の欠落

図1、図2は、圧縮符号化方式MP3(MPEG(Moving Picture Experts Group) AUDIO Layer 3、本稿ではLAME Ver.3.98.2で、128kbpsでエンコードしたものを使用)、AAC(Advanced Audio Codec、本稿ではiTunes Ver.8.0.2.20で、128kbpsでエンコードしたものを使用)で符号化した信号のスペクトログラム(時間・周波数特性)を示している。図で、上段が圧縮符号化前(原音)の特性、下段が圧縮符号化後の特性である。

図から、MP3、AACともに圧縮符号化によって、およそ16kHz以上の高域成分が欠落していることが分かる。ビットレートが下がるとこの傾向は更に顕著になり、64kbpsではおよそ12kHz以上の全ての高域成分が失われる場合もある。

高域成分は、音の華やかさや立ち上がりの鋭さ、楽器の生々しさ等に関係するため、高域成分が欠落することは、音質劣化の要因となり得る。

2.3 左右差信号の欠落

再生するオーディオ信号に空間的な情報が多く含まれるほど、ステレオ音像の広がり感や奥行き感が得られる。また、空間的な情報は左右信号の差が大きいほど豊かになる。

図3、図4はMP3、AACで符号化したオーディオ信号

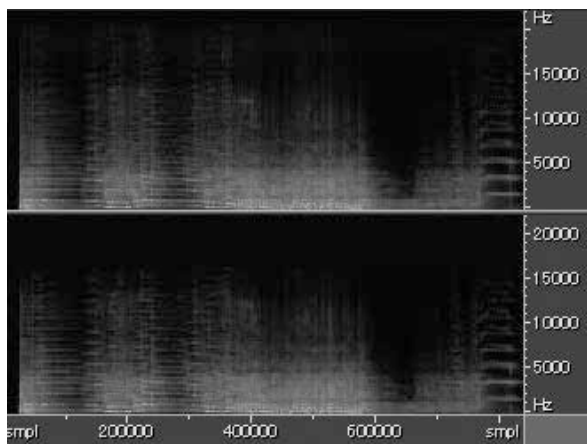


図1. 圧縮符号化のスペクトログラム(MP3)

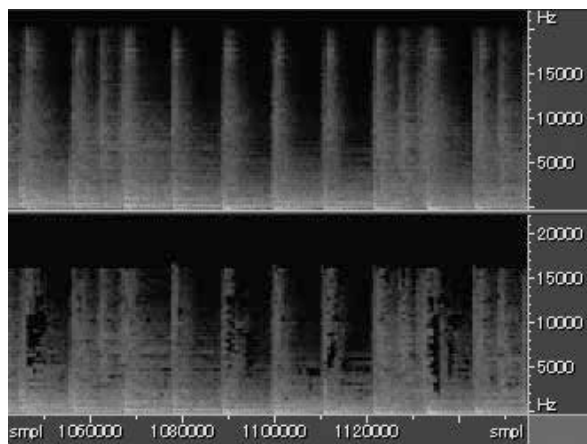


図3. 左右差信号の特性(MP3)

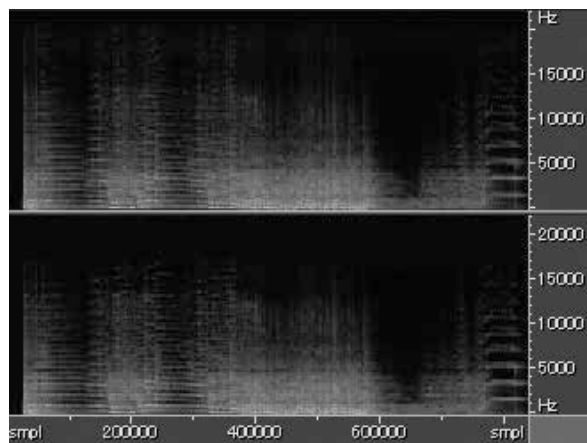


図2. 圧縮符号化のスペクトログラム(AAC)

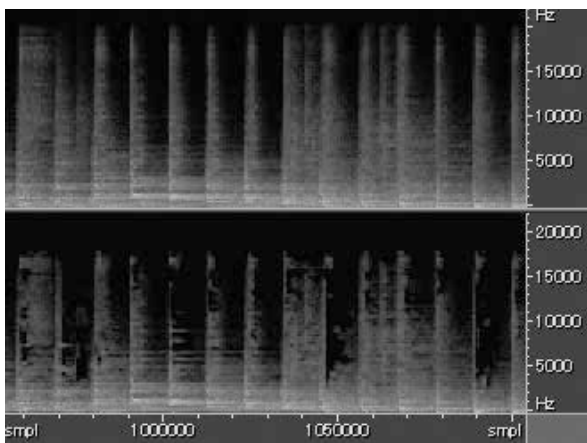


図4. 左右差信号の特性(AAC)

の左右差信号のスペクトログラムである。図で、上段が圧縮符号化前(原音)の特性、下段が圧縮符号化後の特性である。

図から、左信号と右信号の差分の中高域のスペクトルが部分的に欠落し、いわゆる“歯抜け”のようにになっていることが分かる。スペクトル上の欠落は、その帯域の音信号が左右同じ音になっていることを示すものであり、ステレオ音像の広がり感、奥行き感が劣化する要因になり得る。

ビットレートが下がるとこの傾向は更に顕著になり、64kbpsではかなり広範囲の周波数成分が失われることもある。

3. 高音質化技術

3.1 処理構成

図5は圧縮符号化音の高音質化技術の処理構成である。この技術は、圧縮符号化による2つの音質劣化要因に対応し、失われた左右差信号成分を推定補完するステレオ回復器と、失われた高域成分を推定して帯域を拡張するBWEから構成される。

3.2 ステレオ回復器

ステレオ回復器では、左信号と右信号の無相関成分を抽出し、これを左右差信号の推定信号として代用する。これによって、左右差信号を強調し、ステレオ音像の広がり感を回復している。

一般的に、ステレオ音声の左信号と右信号のスペクトル強度がほぼ等しいとき、左右信号の和信号は相関成分、差信号が無相関成分を表すと解釈できる。この方法は、その特徴を利用し、無相関成分を算出し、差信号を再現し強調する方法である。

3.3 BWE

BWEでは、入力信号を、調波構造を持つ周期成分と、立ち上がり成分などの非周期成分とに分割し、周期成分には調波構造を保持するように帯域拡張を行い、非周期成分には立ち上がりを強調する処理を行っている。

一般的に、高域成分が欠落すると、人の声や楽器等の周期性のある音では高調波成分がなくなることで、華やかさ、伸びやかさが欠如し、打楽器などの非周期的な音では、立ち上がりが鈍くなる。これに着目し、周期的な音と非周期的な音で異なる高域の再現を行い、品質改善を図る方法である。

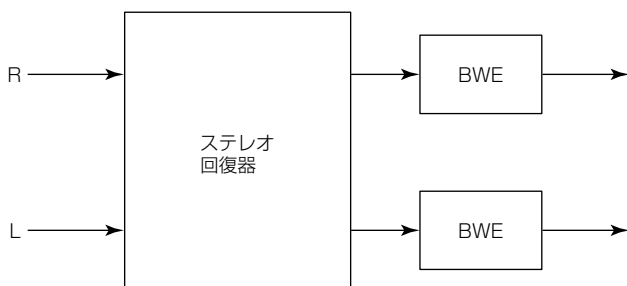


図5. 圧縮符号化音の高音質化処理

4. 評価

高音質化技術の効果を評価するため、音楽信号を用いて周波数特性の確認と、主観評価を行った。

4.1 処理後の特性

図6、図7は、MP3、AACで符号化したオーディオ信号(上段)と、高音質化技術で処理した信号(下段)のスペクトログラムを示している。なお、評価曲として欧州で標準的に音質評価に用いられているSQAM(Sound Quality Assessment Material)のtrack 59(Violin)を用いた。

図から、MP3、AACどちらの方式でも処理前は16kHz以上の高域成分がないが、この処理によって20kHz程度まで高域成分が伸びていることが確認できる。これらの図は一例であるが、SQAMに収録された70曲で同じ傾向となることを確認している。

また、図8、図9はMP3、AACで符号化したオーディオ信号の左右差信号(上段)と、高音質化技術で処理した信号の左右差信号(下段)のスペクトログラムを示している。なお、評価曲としてSQAMのtrack 69(ABBA)を用いた。

図から、この処理によって中高域成分の“歯抜け”が改善

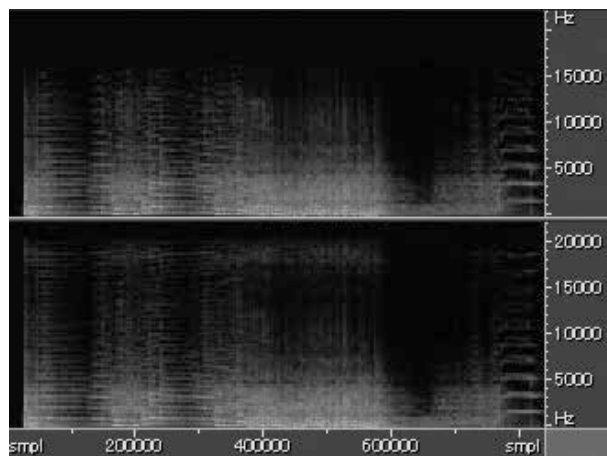


図6. 圧縮符号化の高音質化処理(MP3)

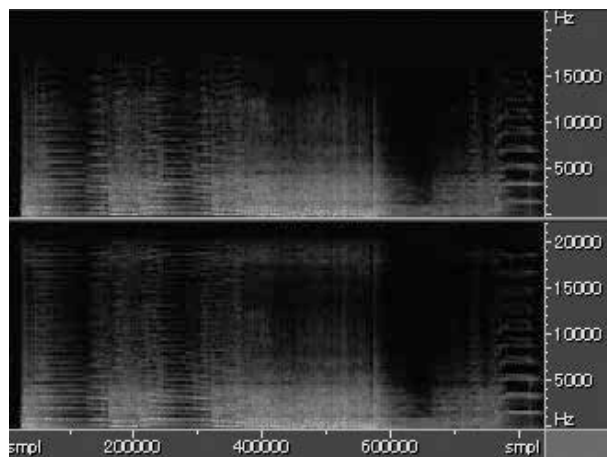


図7. 圧縮符号化の高音質化処理(AAC)

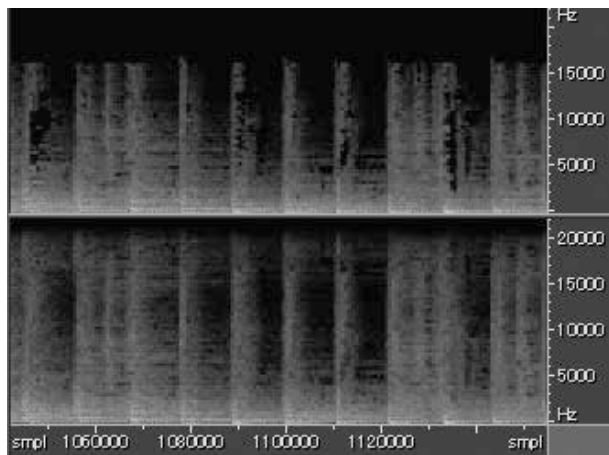


図 8. 圧縮符号化の高音質化処理(MP3)

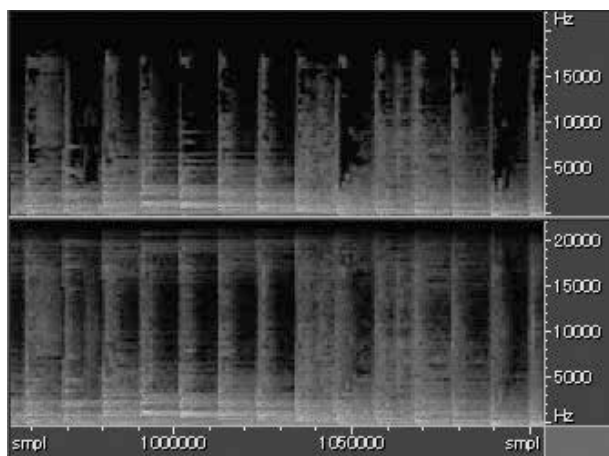


図 9. 圧縮符号化の高音質化処理(AAC)

されていることが確認できる。なお、SQAM収録曲70曲で同様の傾向となることを確認している。

4.2 主観評価

オーディオ信号の主観評価方法を規定したITU-R (International Telecommunications Union-Radiocommunication sector) 勧告BS.1534⁽²⁾で用いられている“隠れ基準及び隠れアンカー付き多重刺激二重盲検法”を参考に主観評価を行った。すなわち、被験者は、評価音が基準音よりも良ければ50点以上の評点を、悪ければ50点以下の評点を0~100の範囲で採点し、その際、安定した評価ができるよう、明らかに品質の低いサンプルを含めて評価する。このサンプルを隠れアンカーと呼ぶ。評価では、基準音と同程度であれば50点と採点する。隠れ基準はCD音声、隠れアンカーは、6 kHz以上の高域成分を緩やかにカットした曲とした。

また、評価曲として、SQAMの中から(1)track 48 (Vocal), (2)track 59 (Violin), (3)track 66 (Orchestra), (4)track 70 (Eddie Rabbitt)の4曲を選定した。

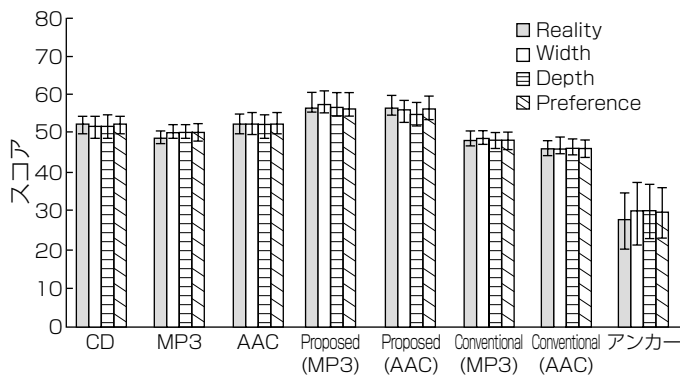


図10. 主観評価結果

評価項目は、“楽器・声のリアルさ (Reality)” “音像の広がり (Width)” “音像の奥行き (Depth)” “好き／嫌い (Preference)” の4項目である。被験者はオーディオエンジニア7名、また比較のため、従来技術による高域再現手法⁽³⁾を含めた8種類の比較を行った。

評価結果を図10に記す。図では、隠れ基準音、アンカーを含めた各評価音のスコアの平均点を示している。また、t検定に基づく95%信頼性区間も示している。

図10から、高音質化技術(図中Proposed)によって元の圧縮符号化音に比べ、8ポイント近くの大きな改善効果があり、従来方法(図中Conventional)からも優位に改善されていることが確認できた。

さらに、提案方式は、全て、CD原音と同等以上のスコアを得ており、被験者が全般的に良い評価をしたと言える。

5. む す び

当社の音のデザイン技術を支える圧縮符号化音の高音質化方式について述べた。この方式は、圧縮符号化によって生じる2つの音質劣化傾向に対応してステレオ回復器とBWEとから構成されており、主観評価によって有効性が確認できた。

今後、圧縮音声に限定せず、より高い品質の提供、ユーザーの嗜好や曲にあわせた音質の制御等、オーディオ全般について音のデザインの研究開発を継続していく。

参考文献

- (1) Kimura, M., et al.: Improvement in Stereophonic Sound Images of Lossy Compression Audio, IEEE 2nd GCCE, 90~91 (2013)
- (2) Recommendation ITU-R BS.1534: Method for the subjective assessment of intermediate quality levels of coding systems (2003)
- (3) 佐藤 蜜: 圧縮で失われた音の復元技術, 電気学会誌, 127, No.7, 411~412 (2007)