

スペクトル制御エンベロープによる混合音中の 周期および非周期成分の選択的イコライザ

亀岡 弘和[†] 後藤 真孝^{††} 嵯峨山 茂樹[†]

[†] 東京大学大学院情報理工学系研究科 ^{††} 産業技術総合研究所

E-mail: [†] {kameoka,sagayama}@hil.t.u-tokyo.ac.jp, ^{††} m.goto@aist.go.jp

あらまし 本報告では、多重音中に含まれる周期的および非周期的な音源成分を、2つのスペクトル制御エンベロープを用いた振幅スペクトル変形操作を介し、信号分離することなく選択的にイコライズする方法を提案する。本方法の目指す応用事例は、音楽音響信号の中のドラム音とそれ以外の楽器音のレベル調整や音質補正を個別に行えるインタフェースや、白色雑音などにより劣化した楽曲信号の音質修復などである。提案方法の動作確認のために実装した GUI を操作しながら得たいくつかの効果例を紹介する。

Selective Amplifier of Periodic and Non-periodic Components in Concurrent Audio Signals with Spectral Control Envelopes

Hirokazu Kameoka[†] Masataka Goto^{††} Shigeki Sagayama[†]

[†]Graduate School of Information Science and Technology, The University of Tokyo

^{††}National Institute of Advanced Industrial Science and Technology (AIST)

Abstract We propose a new equalizing technique that enables a separate control of periodic and non-periodic source components in music signals through a spectral mapping using the two spectral control envelopes. The applications we have in mind are for example a selective amplifier of drum sounds and other instruments in a polyphonic music signal and a recuperation of a degraded music signal with an additive white noise.

1 まえがき

今日のデジタルオーディオプレイヤーソフト (例えば、iTunes や WinAmp 等) の多くには、ユーザ設定やプリセット選択などにより特定周波数帯域のフィルタゲインを調節し、音質を加工できるイコライザ機能がほぼ標準的に搭載されている [1]。これは、必ずしもリスナは既存の楽曲を与えられたまま聴きたいだけでなく、音質を自由に加工して聴きたいという潜在的な欲求も持っていることを示唆すると考えられ、もしさらに自由度の高い (例えば、混合音中の特定の成分のみをイコライズできるような) 音質加工技術が提供できれば、リスナはより能動的に音楽を楽しむようになる可能性がある。そこで、本論文では、次世代音楽鑑賞用インタフェースに向け、新しいイコライザ機能を実現するための技術を提案する。

実際既に、新しい音質加工技術の一例として、INTER:D [7] や Drumix [8] と呼ぶ、楽曲音響信号の中のドラム音だけを選択的にイコライズできたり、他の音に置換できるインタフェースが吉井らによって提案されている。彼らがとったのは、ドラム音とそれ以外の構成音のスペクトログラムを精度良く分離さえできれば良いであろうという立場であったが、そのためにはドラム音の発音時刻の特定とスペクトログラムの高精度な分離推定のための高度なパターン認識や信号処理技術を要し、精度向上のための試みも継続的に取り組まれている [9]。

一方で我々は、特定の楽器に対する固有の知識に基づいたスペクトルの分離推定過程を経るのではなく、対象楽器を限定しない汎用的な技術の実現を目指し、個別にイコライジングする対象をより抽象化する。具体的には、ラインスペクトル状の成分と平坦なスペクトル成分を、分離することなく個別に変形 (イコライジング) するために必要な方法論としては何が考えられるか、という問題設定から出発することにした。良く知られる通り、音声の母音や音高を持つ楽器音の定常部などにおける周期的な信号は、基本周波数の整数倍の周波数にラインスペクトル状の成分を持つ調波構造と呼ばれる構造を持つが、一方、音声の子音や楽器音のアタック時、ドラム音などの多くに含まれる非周期的な信号は、広い帯域に広がる平坦なスペクトル成分を持つ傾向がある。従って、もしこのようなイコライザが実現されれば、例えば、ドラム音のレベル調整、古いレコードの劣化音に含まれる白色・ピンクノイズの抑制や、ライブ音源における聴衆の歓声等の背景音のレベル調整等、音質補正、音圧補正、ノイズリダクションを含む広範囲な応用可能性が期待される。さらに、ある楽曲のドラム音などのミキシングレベルを別の楽曲にリスナが手軽に転写する用途にも適用できる。

以上の応用可能性を踏まえ、我々は、図 1 に示すように、振幅スペクトルの山頂と谷底周辺を通るような 2 本のエンベロープをそれぞれ上下に操作することで帯域ごとのスペクトルピークとディップのレベル差を調節し、ラインスペクトル状と平坦な成分を個別にイコライズできる音質加工手法を

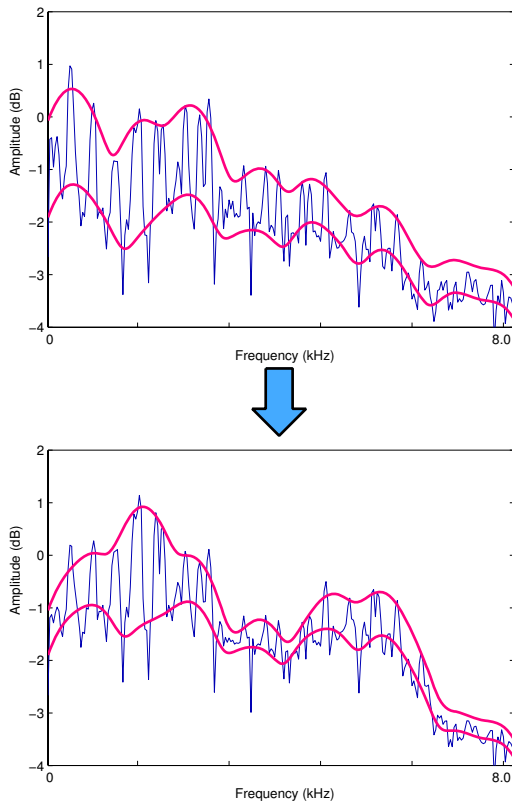


図1 スペクトル制御エンベロープを介した振幅スペクトル中のラインスペクトル状および平坦な成分の変形操作

提案する。本手法は、振幅スペクトルの山頂周辺を通る包絡（**トップエンベロープ**と呼ぶ）と谷底周辺を通る包絡（**ボトムエンベロープ**と呼ぶ）を得て、それらを制御点とした各周波数における振幅スペクトルの非線形変換を介し、音量や音質の調節を可能にする。その後、加工した振幅スペクトルと元の位相スペクトルから、音質加工後の音響信号を生成する。以上のようにトップエンベロープとボトムエンベロープを介して振幅スペクトルを間接的に変形する目的のため、これらを併せて**スペクトル制御エンベロープ**と呼ぶことにする。

以上のような振幅スペクトルの非線形変換に基づく音質加工手法は、従来のように帯域ごとに異なるゲインを決めるイコライザ [2, 3, 4, 5, 6] などでは実現が困難であり、対象が多重音であってもその各成分を同定・分離する必要がない利点と、高い自由度を持ちつつ直観的な変形を可能にする利点を持つ。

以上のように具体的に我々が目指すのは、制御エンベロープの操作を介した直観的な音質加工が可能なインタフェースであるが、その実現のために

- 制御エンベロープの適切な関数クラスの選定
- 制御エンベロープの取得方法
- 制御エンベロープの変換に伴う振幅スペクトル変形法

について以下で論じていく。

2 エンベロープの関数クラスの定義とその性質

制御エンベロープの操作性の良さは、エンベロープの関数クラスの選び方に大きく関係する。例えば、スペクトルの山頂付近を通るような包絡曲線については有名な線形予測分析

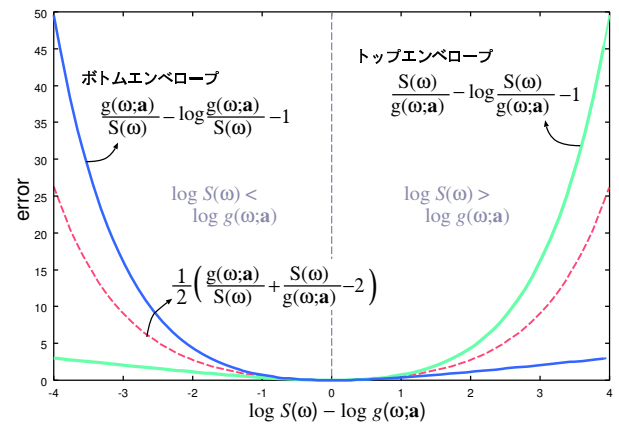


図2 スペクトル歪み尺度の被積分関数を $\log S(\omega) - \log g(\omega; \mathbf{a})$ の関数と見立てた場合のグラフ。トップエンベロープ用の目的関数においては、 $\log S(\omega)$ が $\log g(\omega; \mathbf{a})$ より大きい場合には小さいときに比べて過大なペナルティとして課す。一方、ボトムエンベロープ用の目的関数では、 $\log S(\omega)$ が $\log g(\omega; \mathbf{a})$ より小さい場合には、大きいときに比べて過大なペナルティとして課す。

(Linear Predictive Coding) で得ることはできるが、LPC エンベロープの場合、パラメータが LPC 多項式の各項の係数であり、エンベロープのピークがその多項式の根に対応しているため、パラメータの動かし方がエンベロープの操作に直感的に対応しやすいとは必ずしも言い難い。その意味で、エンベロープの関数クラスの定義は重要な要件の一つである。

直感的操作のしやすさに留意し、エンベロープ曲線を次のような正値係数 $\forall i: a_i > 0$ による核関数の線形和：

$$g(\omega; \mathbf{a}) = \sum_{i=1}^I a_i \Psi_i(\omega) \quad (1)$$

を用いることにする。パラメータは係数 $\mathbf{a} = (a_1, \dots, a_I)^T$ である。ただし、核関数 $\Psi_i(\omega)$ はすべての $\omega \in \mathbb{R}$ において

$$0 < \Psi_i(\omega) < \infty \quad (2)$$

を満たす任意の有界関数とするが、以後 $\Psi_i(\omega)$ は、直感的なイメージを与えるために Gauss 分布を想定することにする。 $\Psi_i(\omega)$ を、ナイキスト周波数範囲内に一定間隔で並べる (なお、ある程度の滑らかさを仮定するためには、例えば各 Gauss 分布の標準偏差をその間隔と等しく定めるなどすれば良い。) ことにすると、係数 $\mathbf{a} = (a_1, \dots, a_I)^T$ は、各帯域でのエンベロープの高さの調節用のパラメータに相当するため、パラメータを介したエンベロープの操作性は極めて直感的である。

3 制御エンベロープの反復推定アルゴリズム

3.1 目的関数の定義

スペクトルの山頂近くを通るトップエンベロープを得るための一つの手段としては、スペクトルの山頂近くに張りついていての方がよりペナルティが低く抑えられるような歪み尺度を導入し、それを最小とするような \mathbf{a} を求めれば良い。従ってまず、振幅スペクトルのエンベロープ関数からの正値の誤差には過大なペナルティを、負値の誤差には過小なペナルティを与えるような何らかの非対称な歪み尺度を考えたい。そのような歪み尺度の例として、LPC の枠組で導出された板倉斎藤距離 [10] と呼ぶ尺度が広く知られており、上記の目的

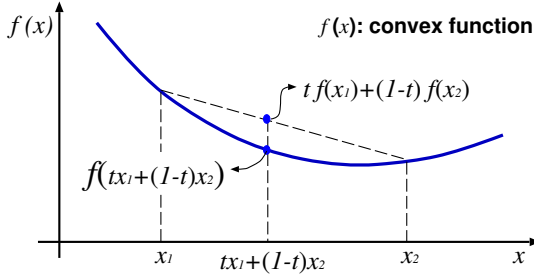


図3 凸不等式の概念図。

を達成できる目的関数の1つである。ただしここでは、文献[10]においてなされているような上記尺度の物理的意味に関する議論は行わない。以上より、トップエンベロープを求めるための最小化すべき目的関数は

$$J_{\top} = \int \left(\frac{S(\omega)}{g(\omega; \mathbf{a})} - \log \frac{S(\omega)}{g(\omega; \mathbf{a})} - 1 \right) d\omega \quad (3)$$

である。一方、スペクトルの谷底近辺を通るボトムエンベロープを得るためには、逆に振幅スペクトルのエンベロープ関数からの負値の誤差には過大なペナルティを、正值の誤差には過小なペナルティを与えるような非対称な歪み尺度を導入すれば良いが、板倉齊藤距離の要素を逆転した歪み尺度がまさにこのような尺度の1つとして適用できる。従って、ボトムエンベロープを求めるための最小化すべき目的関数としては

$$J_{\perp} = \int \left(\frac{g(\omega; \mathbf{a})}{S(\omega)} - \log \frac{g(\omega; \mathbf{a})}{S(\omega)} - 1 \right) d\omega \quad (4)$$

が考えられる。両目的関数の被積分関数を $\log S(\omega) - \log g(\omega; \mathbf{a})$ の関数と見立てたグラフ(図2)を見てみると、両目的関数がトップエンベロープおよびボトムエンベロープを推定するのに適切な尺度であることが良く分かる。実際には、上記以外の歪み尺度に基づく目的関数を考えることもできるが、式(3)、(4)を採用することで次節以降に示すエレガントな反復最適化アルゴリズムを導くことができる。

3.2 ボトムエンベロープの推定法

まず、はじめにボトムエンベロープのパラメータ \mathbf{a}_{\perp} の推定方法について述べる。 \mathbf{a}_{\perp} は、

$$J_{\perp} = \int \left(\sum_{i=1}^I \frac{a_i \Psi_i(\omega)}{S(\omega)} - \log \sum_{i=1}^I a_i \Psi_i(\omega) + \log S(\omega) - 1 \right) d\omega \quad (5)$$

を最小化する \mathbf{a} を求めることで得られるが、第二項の対数関数の非線形性により、解析解の求解は困難であることが分かる。しかしながら、この最適化問題は、勾配法やサンプリングアルゴリズムに頼らずとも、EMアルゴリズムのアイデアをヒントにした次に示すような工夫により、極めて効率的な反復法が導けることを示す。そのためには、 \mathbf{a} に関する最小化が解析的に行え、かつ目的関数の上限となっているような補助関数を設計することが目標となる。具体的な方策を以下で提案する。

負の対数関数が凸関数であることに注意すると、

$$\forall \omega : \sum_{i=1}^I m_i^{\perp}(\omega) = 1, \quad \forall i, \omega : 0 < m_i^{\perp}(\omega) < 1 \quad (6)$$

を満たす任意の重み関数を用い、凸不等式：

$$\begin{aligned} \forall \omega : -\log \left\langle \frac{a_i \Psi_i(\omega)}{m_i^{\perp}(\omega) S(\omega)} \right\rangle_{m_i^{\perp}(\omega)} &= -\log \sum_{i=1}^I \frac{a_i \Psi_i(\omega)}{S(\omega)} \\ &\leq -\left\langle \log \frac{a_i \Psi_i(\omega)}{m_i^{\perp}(\omega) S(\omega)} \right\rangle_{m_i^{\perp}(\omega)} \end{aligned}$$

が立てられる。ただし、 $\langle x_j \rangle_{w_j}$ は、 x_j の w_j による重みつき和の演算子： $\sum_j w_j x_j$ を表す。凸不等式の概念図を図3に示す。以上より、

$$J_{\perp} \leq \int \left(\log S(\omega) - 1 + \sum_{i=1}^I \left(\frac{a_i \Psi_i(\omega)}{S(\omega)} - m_i^{\perp}(\omega) \log \frac{a_i \Psi_i(\omega)}{m_i^{\perp}(\omega) S(\omega)} \right) \right) d\omega \quad (7)$$

なる不等式が得られ、上式右辺は目的関数の上限値であることが分かる。この右辺は、変関数 m^{\perp} の汎関数であり、これを補助関数 $J_{\perp}^+[m^{\perp}]$ と呼ぶ。ここで注目すべき点は、もし $m_i^{\perp}(\omega)$ を固定すると、 $J_{\perp}^+[m^{\perp}]$ を最小化する a_i の解析的な解は陽に与えられる(後述)ことである。従って、目的関数から直接 a_i の最適化を行う代わりに、補助関数を用いて効果的に a_i の最適化を行っていくことができる。まず、 a_i を固定したまま、 $m_i^{\perp}(\omega)$ に関して補助関数を最小化すると、式(7)は等号が成り立つ。その等号成立状態から、両関数に含まれる変数 a_i に関して補助関数を最小化すると、自動的に目的関数も減少する。なぜなら、補助関数は目的関数の上限値を表し、不等式(7)より最小化した補助関数よりさらに目的関数は小さくなっていることが保証されるからである。板倉齊藤距離は下に有界なので、 $J_{\perp}^+[m^{\perp}]$ を $m_i^{\perp}(\omega)$ と a_i に関して一方を固定しながら交互に最小化すれば、目的関数は停留点に収束することが示される。この反復最適化の概念図を図4に示す。

さて、まず、条件(6)を満たしながら $J_{\perp}^+[m^{\perp}]$ を $m_i^{\perp}(\omega)$ に関して最小化するには、

$$\sum_{i=1}^I \int \left(\frac{a_i \Psi_i(\omega)}{S(\omega)} - m_i^{\perp}(\omega) \log \frac{a_i \Psi_i(\omega)}{m_i^{\perp}(\omega) S(\omega)} - \lambda(\omega) \left(\sum_{i=1}^I m_i^{\perp}(\omega) - 1 \right) \right) d\omega \quad (8)$$

の $m_i^{\perp}(\omega)$ に関する括弧内の偏微分

$$\log \frac{S(\omega)}{a_i \Psi_i(\omega)} + \log m_i^{\perp}(\omega) + 1 - \lambda(\omega) \quad (9)$$

を0とする $m_i^{\perp}(\omega)$ を求めれば良い。ただし $\lambda(\omega)$ は Lagrange 乗数である。式(9)を0にする $m_i^{\perp}(\omega)$ は

$$\hat{m}_i^{\perp}(\omega) = \frac{a_i \Psi_i(\omega)}{S(\omega)} e^{\lambda(\omega) - 1} \quad (10)$$

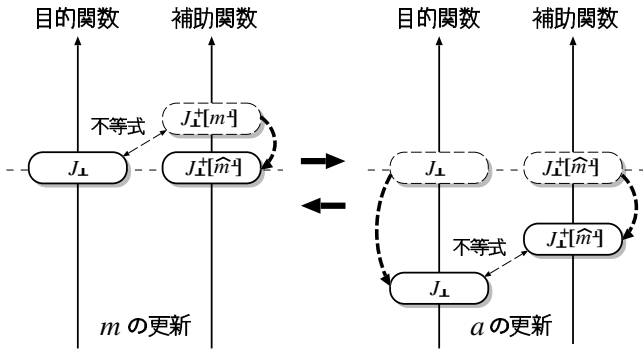


図4 任意の a_i と $m_i^\perp(\omega)$ の組合せのときは補助関数は目的関数より大きく、補助関数にのみ含まれる変関数 $m_i^\perp(\omega)$ に関して補助関数を最小化すると両者は等号で結ばれる (左図)。その状態で両者の共通の変数 a_i に関して補助関数を最小化すると、当然補助関数自体は減少するが、その値よりも必ず目的関数値は小さくなっている (右図)。

で与えられ、条件 (6) より最終的に

$$\hat{m}_i^\perp(\omega) = \frac{a_i \Psi_i(\omega)}{\sum_{\forall i} a_i \Psi_i(\omega)} \quad (11)$$

を得る。 $\forall i: a_i > 0, \omega: \Psi(\omega) > 0$ であるから $\forall i, \omega: \hat{m}_i^\perp(\omega) > 0$ である。一方、 $m_i^\perp(\omega)$ 固定で、補助関数の a_i に関する偏微分

$$\int \left(\frac{\Psi_i(\omega)}{S(\omega)} - \frac{m_i^\perp(\omega)}{a_i} \right) d\omega \quad (12)$$

を 0 とする a_i は

$$\hat{a}_i = \frac{\int m_i^\perp(\omega) d\omega}{\int \frac{\Psi_i(\omega)}{S(\omega)} d\omega} \quad (13)$$

で与えられる。 $\forall i, \omega: \hat{m}_i^\perp(\omega) > 0$ であるから明らかに $\forall i: \hat{a}_i > 0$ である。よって、 $\forall i: a_i$ の初期値を正值にすれば必ず収束値も正值となる。式 (11) と式 (13) の反復計算における \mathbf{a} の収束先をボトムエンベロープのパラメータ \mathbf{a}_\perp とする。以下では、 \mathbf{a}_\perp が上記のアルゴリズムで既に求まっていることを前提にする。

3.3 トップエンベロープの推定法

トップエンベロープとボトムエンベロープとをある程度以上は近接させないという付加条件を与えるために、上記で求めたボトムエンベロープパラメータ \mathbf{a}_\perp を用いて次のようなペナルティ関数項：

$$P(\mathbf{a}) \triangleq \sum_{i=1}^I \frac{\eta_i a_i^\perp}{a_i} \quad (14)$$

を導入して、目的関数に付加する。上記項は、 a_i が a_i^\perp と比べて大きいほど、 $P(\mathbf{a}) \rightarrow 0$ なのでペナルティの効力を小さくするが、 a_i^\perp に近い (または小さい) ほどペナルティの効力を大きくする働きがある。また、 $\forall i: \eta_i$ は、ペナルティの影響力の大きさを表すものであり、実験的に定める正值のパラメータとする。

以上より、トップエンベロープのパラメータ \mathbf{a}_\top は、

$$J_\top = \int \left(\frac{S(\omega)}{\sum_{\forall i} a_i \Psi_i(\omega)} + \log \sum_{i=1}^I a_i \Psi_i(\omega) - \log S(\omega) - 1 \right) d\omega + P(\mathbf{a}) \quad (15)$$

を最小化する \mathbf{a} を求めることで得られるが、第一項と第二項の逆数関数と対数関数の非線形性により、これを \mathbf{a} に関する偏微分を 0 とする \mathbf{a} を求めるには非線形な連立方程式を解く必要があり、解析解の求解は困難であることが分かる。 \mathbf{a} に関しての最小化が解析的に行え、かつ目的関数の上限値となっているような補助関数が作れるかどうか、前節と同様の効果的な反復最適化のための要件となる。さて、まず第一項に注目し、正領域において逆数関数が凸関数であることに注意すると、

$$\forall \omega: \sum_{i=1}^I m_i^\top(\omega) = 1, \quad \forall i, \omega: 0 < m_i^\top(\omega) < 1 \quad (16)$$

を満たす任意の重み関数を用いて凸不等式：

$$\begin{aligned} \forall \omega: 1 &\left\langle \frac{a_i \Psi_i(\omega)}{m_i^\top(\omega) S(\omega)} \right\rangle_{m_i^\top(\omega)} = \frac{S(\omega)}{\sum_{\forall i} a_i \Psi_i(\omega)} \\ &\leq \left\langle 1 / \frac{a_i \Psi_i(\omega)}{m_i^\top(\omega) S(\omega)} \right\rangle_{m_i^\top(\omega)} = \sum_{i=1}^I \frac{m_i^\top(\omega)^2 S(\omega)}{a_i \Psi_i(\omega)} \end{aligned}$$

が成り立つので、

$$J_\top \leq \int \left(\sum_{i=1}^I m_i^\top(\omega)^2 \frac{S(\omega)}{a_i \Psi_i(\omega)} - \log S(\omega) + \log \sum_{i=1}^I a_i \Psi_i(\omega) - 1 \right) d\omega + P(\mathbf{a}) \quad (17)$$

なる不等式を得る。これでもまず、各要素 $a_i \Psi_i(\omega)$ の和の逆数を各要素の逆数の総和の形にした目的関数の上限関数を作ることができた。上式右辺を $J_\top^+[\mathbf{m}^\top]$ とする。

次に、第二項の上限値を設計することが目標であるが、前節と異なり、正の対数関数は凹関数なので、凸不等式では上限関数が作れない。しかしながら、任意の ω における

$$\log \sum_{i=1}^I a_i \Psi_i(\omega) \quad (18)$$

に対し、任意の点 $t(\omega)$ の周りの 1 次テイラー展開：

$$\log t(\omega) + \frac{1}{t(\omega)} \left(\sum_{i=1}^I a_i \Psi_i(\omega) - t(\omega) \right) \quad (19)$$

が上限となっていることを利用すれば、式 (17) の右辺のさらなる上限関数が

$$\begin{aligned} \int \left(\sum_{i=1}^I m_i^\top(\omega)^2 \frac{S(\omega)}{a_i \Psi_i(\omega)} + \log t(\omega) - \log S(\omega) \right. \\ \left. + \frac{1}{t(\omega)} \left(\sum_{i=1}^I a_i \Psi_i(\omega) - t(\omega) \right) - 1 \right) d\omega + P(\mathbf{a}) \quad (20) \end{aligned}$$

と書ける。上式を $J_T^+[m^\top, t]$ とすると、すなわち、

$$J_T \leq J_T^+[m^\top] \leq J_T^+[m^\top, t] \quad (21)$$

が成り立つ。 $J_T^+[m^\top, t]$ は、変関数 $m_i^\top(\omega)$ と $t(\omega)$ が固定のとき、パラメータ \mathbf{a} に関する最適解が解析的に求まる(後述)ことから目的関数最小化のための補助関数として扱える。これを踏まえ、この補助関数を目的関数と等しくする $m_i^\top(\omega)$ と $t(\omega)$ を以下で具体的に求めることにする。

前節と同様に、まず $m_i^\top(\omega)$ に関して式 (20) を最小化するには

$$\int \left(\sum_{i=1}^I m_i^\top(\omega)^2 \frac{S(\omega)}{a_i \Psi_i(\omega)} - \gamma(\omega) \left(\sum_{i=1}^I m_i^\top(\omega) - 1 \right) \right) d\omega$$

の括弧内の $m_i^\top(\omega)$ に関する偏微分：

$$\frac{2m_i^\top(\omega)S(\omega)}{a_i \Psi_i(\omega)} - \gamma(\omega) \quad (22)$$

を 0 とする $m_i^\top(\omega)$ を求めれば良い。ただし、 $\gamma(\omega)$ は Lagrange 乗数である。式 (22) を 0 とする $m_i^\top(\omega)$ は

$$\hat{m}_i^\top(\omega) = \frac{\gamma(\omega)a_i \Psi_i(\omega)}{2S(\omega)} \quad (23)$$

で与えられ、条件 (16) より最終的に

$$\hat{m}_i^\top(\omega) = \frac{a_i \Psi_i(\omega)}{\sum_{\forall i} a_i \Psi_i(\omega)} \quad (24)$$

を得る。 $\forall i: a_i > 0, \omega: \Psi(\omega) > 0$ であるから $\forall i, \omega: \hat{m}_i^\top(\omega) > 0$ である。 $m_i^\top(\omega)$ が上式で与えられるとき、不等式 (17) の等号、 $J_T = J_T^+[m^\top]$ が成り立つ。一方、明らかに

$$\hat{t}(\omega) = \sum_{i=1}^I a_i \Psi_i(\omega) \quad (25)$$

のときに $J_T^+[m^\top] = J_T^+[m^\top, t]$ となる。従って、 $m_i^\top(\omega)$ と $t(\omega)$ が式 (24) と式 (25) で与えられるとき $J_T = J_T^+[m^\top, t]$ である。

以上で、補助関数を目的関数に等しくするステップが具現化できたので、次に $m_i^\top(\omega)$ と $t(\omega)$ が固定の下で補助関数を最小化する \mathbf{a} を求められることを確認する。式 (20) の a_i に関する偏微分：

$$\int \left(\frac{\Psi_i(\omega)}{t(\omega)} - \frac{m_i^\top(\omega)^2 S(\omega)}{\Psi_i(\omega)} \frac{1}{a_i^2} \right) d\omega - \frac{\eta_i a_i^\perp}{a_i^2} \quad (26)$$

を 0 と置くと、 $\forall \omega: 0 < t(\omega), \Psi_i(\omega) < \infty$ であるから、

$$\tilde{a}_i = \left(\frac{\int \frac{m_i^\top(\omega)^2 S(\omega)}{\Psi_i(\omega)} d\omega + \eta_i a_i^\perp}{\int \frac{\Psi_i(\omega)}{t(\omega)} d\omega} \right)^{1/2} \quad (27)$$

が得られる。ただし、 $\tilde{\mathbf{a}}$ が \mathbf{a}_\perp を下回らないようにペナルティ項を付加したとは言え、 a_i が a_i^\perp より常に大きいことが保証されるわけではないため、 a_i は

$$\forall i: \hat{a}_i = \begin{cases} \tilde{a}_i & \text{if } \tilde{a}_i \geq a_i^\perp \\ a_i^\perp & \text{otherwise} \end{cases} \quad (28)$$

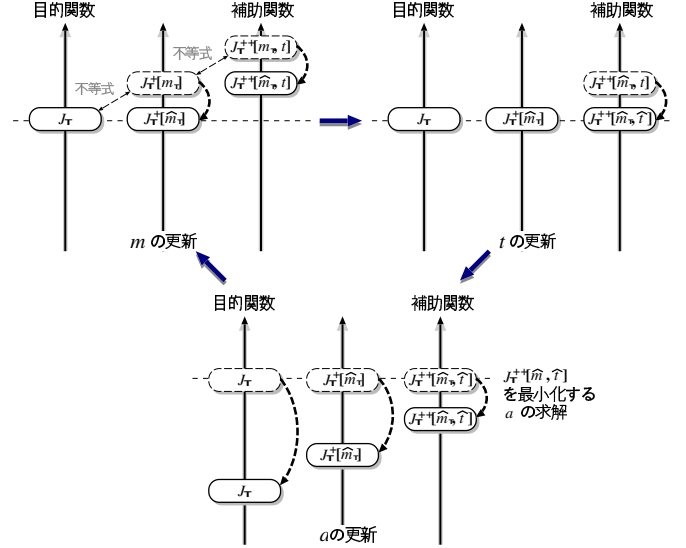


図5 任意の a_i と $m_i^\top(\omega)$ と $t(\omega)$ の組合せのときは補助関数は目的関数より大きく、補助関数にのみ含まれる変関数 $m_i^\top(\omega)$ と $t(\omega)$ に関して補助関数を最小化すると両者は等号で結ばれる(右上図)。一方、その状態で両者の共通のパラメータである a_i に関して補助関数を最小化すると、当然補助関数自体は減少するが、その値よりも必ず目的関数値は小さくなっている(下図)。

のように更新することにする。なお、 $\forall i, \omega: \hat{m}_i^\top(\omega) > 0$ であるから明らかに $\forall i: \hat{a}_i > 0$ である。よって、 $\forall i: a_i$ の初期値を正值にすれば必ず収束値も正值となる。以上より、式 (24) と式 (25) と式 (27) の反復計算により、前節と同様なアイデアにより目的関数を効果的に小さくしていくことができ(図5参照)、適当な収束判定により得られる最終的な \mathbf{a} の値をトップエンベロープのパラメータ \mathbf{a}_T とする。

4 振幅スペクトルの変換則

上記のようにして、トップエンベロープとボトムエンベロープが効果的に得られることが分かった。ここでは、トップエンベロープ付近の成分とボトムエンベロープ付近の成分がそれぞれうまくラインスペクトル状と平坦な成分に対応するという仮定の下で、これらの2本の制御エンベロープを調節または他の楽曲から転写したときに、それに合わせていかにして振幅スペクトルを変換するかについて述べる。

ここで我々が決めておきたいのは、トップエンベロープを上下したときには平坦な成分はできるだけ変化させずラインスペクトル状の成分だけを变化させられ、一方で、ボトムエンベロープを上下したときにはラインスペクトル状の成分はできるだけ変化させず平坦な成分だけを变化させられるような何らかの変換則である。さて、振幅 $S(\omega)$ を $\varepsilon (> 0)$ 倍にするということは、対数振幅 $\log S(\omega)$ の写像先が $\log S(\omega) + \log \varepsilon$ である、すなわち写像関数が傾き 1 の恒等写像関数を上方向に $\log \varepsilon$ だけずらしたもので与えられるということである。従って、対数振幅 $\log S(\omega)$ の写像関数が傾き 1 の直線で与えられる場合、振幅が定数倍されることを意味する。これを踏まえると、ある ω における対数振幅 $\log S(\omega)$ がトップエンベロープ値付近あるいは以上であれば、元のトップエンベロープ値から目標先のトップエンベロープ値までの変化分と同等の大きさの定数を $\log S(\omega)$ に加え、ボトムエンベロープ値付近あるいは以下であれば、元のボトムエンベロープ値か

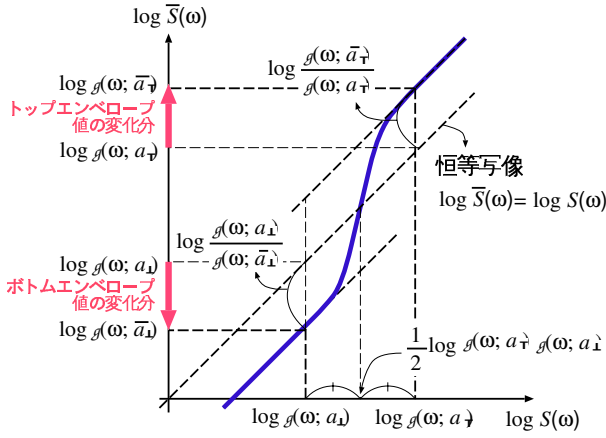


図6 変換元と変換先の制御エンベロープによって決定される、対数振幅 $\log S(\omega)$ の写像関数。例えば、 $\log S(\omega)$ が変換元のトップエンベロープ周辺かそれ以上に大きい場合は、変換先のトップエンベロープまでの変化分とほぼ等しい大きさの定数が $\log S(\omega)$ に加えられる。

ら目標先のボトムエンベロープ値までの変化分と同等の大きさの定数を $\log S(\omega)$ に加えるような写像関数を構成するのが我々の目指すところである。このような操作には、閾値処理などで $\log S(\omega)$ がトップエンベロープとボトムエンベロープのどちらに近いかを決定し、エンベロープの変化分の定数を $\log S(\omega)$ に加える（不連続な写像関数を構成する）方法も考えられるが、中間的な成分の整合性を大きく損なわないようにするには図6のような滑らかな写像関数を構成すれば良い。そこで、その実現方法の一例として我々は、原エンベロープ $g(\omega; \mathbf{a}_\top)$ と $g(\omega; \mathbf{a}_\perp)$ を、目標エンベロープ $g(\omega; \bar{\mathbf{a}}_\top)$ と $g(\omega; \bar{\mathbf{a}}_\perp)$ に変換するときの、 $\log S(\omega)$ から $\log \bar{S}(\omega)$ への変換則を次のような写像関数：

$$\log \bar{S}(\omega) = \log \frac{S(\omega)g(\omega; \bar{\mathbf{a}}_\perp)}{g(\omega; \mathbf{a}_\perp)} + \log \frac{g(\omega; \bar{\mathbf{a}}_\top)g(\omega; \mathbf{a}_\perp)}{g(\omega; \mathbf{a}_\top)g(\omega; \bar{\mathbf{a}}_\perp)} \times f\left(\frac{\log S(\omega) - \frac{1}{2} \log g(\omega; \mathbf{a}_\top)g(\omega; \mathbf{a}_\perp)}{\alpha (\log g(\omega; \mathbf{a}_\top) - \log g(\omega; \mathbf{a}_\perp))}\right) \quad (29)$$

で表現することにする。ただし、 $f(\cdot)$ は sigmoid 関数：

$$f(x) = \frac{1}{1 + e^{-x}} = \begin{cases} 1 & (x \rightarrow \infty) \\ 0 & (x \rightarrow -\infty) \end{cases} \quad (30)$$

である。特に、上式で $\alpha \rightarrow 0$ のとき、

$$\log \bar{S}(\omega) = \begin{cases} \log S(\omega) + \log \frac{g(\omega; \bar{\mathbf{a}}_\perp)}{g(\omega; \mathbf{a}_\perp)} \\ \text{if } \log S(\omega) \geq \frac{1}{2} \log g(\omega; \mathbf{a}_\top)g(\omega; \mathbf{a}_\perp) \\ \log S(\omega) + \log \frac{g(\omega; \bar{\mathbf{a}}_\perp)}{g(\omega; \mathbf{a}_\perp)} \\ \text{if } \log S(\omega) < \frac{1}{2} \log g(\omega; \mathbf{a}_\top)g(\omega; \mathbf{a}_\perp) \end{cases} \quad (31)$$

のように、先に述べた閾値に基づく変換則と等価になる。

5 実装

以上の各要素技術を用い、スペクトル制御エンベロープを介した周期成分と非周期成分の選択的イコライザを実装した。

スペクトル制御エンベロープは実在する物理量であるわけではないため、3章で提案した反復推定アルゴリズム自体の定量的な性能評価は行うことはできない代わりに、制御エン

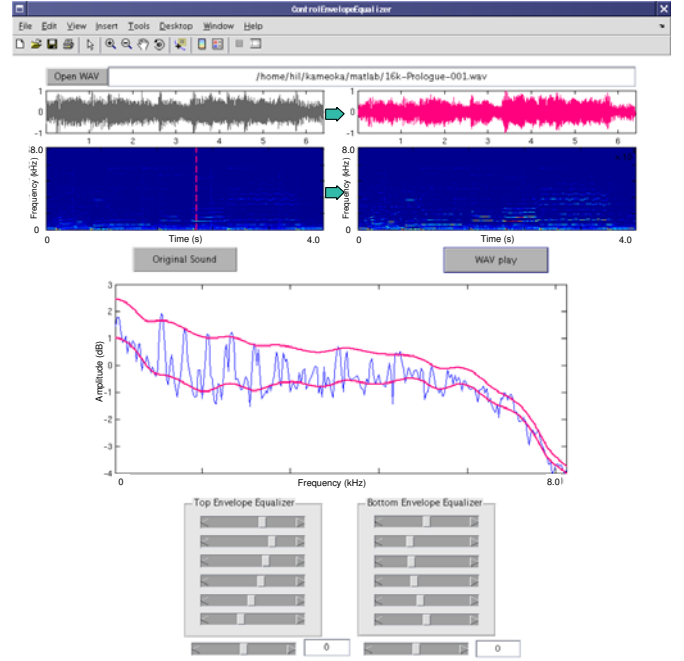


図7 スペクトル制御エンベロープイコライザのグラフィカルユーザインタフェース。左上に表示されるスペクトログラム上でクリックした場所（赤色の破線）での切ロスpekトルが制御エンベロープとともに GUI 中央のグラフに表示され、GUI 下部のスライダを通した制御エンベロープの操作に伴った変換の様子が観察できる。

ベロープを介したイコライザの効果を示す実験結果例をいくつか挙げることにする。実験条件は以下のとおりとした。サンプリング周波数 16kHz のモノラルのデジタル信号を入力とし、16ms 長の Hamming 窓を 8ms ずつずらしながら短時間 Fourier 変換を行いスペクトル系列を算出した。3章の反復推定アルゴリズムでは、核関数にはガウス分布関数を選び、ガウス分布をナイキスト周波数範囲内に一樣間隔で 20 個配置したものをエンベロープ関数とした。すなわち、 i 番目のガウス分布関数の中心は $400 \times i$ Hz である。また、各ガウス分布の標準偏差パラメータは 240Hz とした。 a_1, \dots, a_{20} の初期値は $\frac{1}{20} \int S(\omega) d\omega$ とした。反復アルゴリズムの収束性は実験的に確認され、実装では反復回数をいずれも 30 回とした。また、3.3 章および 4 章において導入した定数パラメータ η 、 α はそれぞれ $\eta_i = 100.0/i$ ($i = 1, \dots, 20$)、 $\alpha = 0.2$ とした。

MATLAB で以上の条件のもとで実装したグラフィカルユーザインタフェースの概観図を図7に示す。GUI 下部に配置されているトップエンベロープおよびボトムエンベロープ用に 6 個ずつ用意されたスライダはそれぞれ、エンベロープ関数中の $i = 1, 2, i = 3, 4, 5, i = 6, 7, 8, i = 9, 10, 11, 12, i = 13, 14, 15, 16, i = 17, 18, 19, 20$ 番目の核関数の重み係数に、すべてのフレームに渡って一様に乗じる正值のゲインを調節するためのものである。このスライダを調節することにより制御エンベロープに対し帯域ごとの強調や抑圧が行え、そのたびに 4 章のマッピングを介して振幅スペクトルが変形される。また、特定時刻の短時間振幅スペクトルや全体のスペクトログラムが上記操作に伴ってどのように変形されるかが GUI 上部で確認できるようになっている。

図8に、RWC 研究用音楽データベース [11] 収録のポピュラーとジャズ音楽の各 1 曲ずつに対し、原スペクトログラムとともに上記 GUI を用いて非周期成分および周期成分を強調した変換信号のスペクトログラムを示す。各実験データの

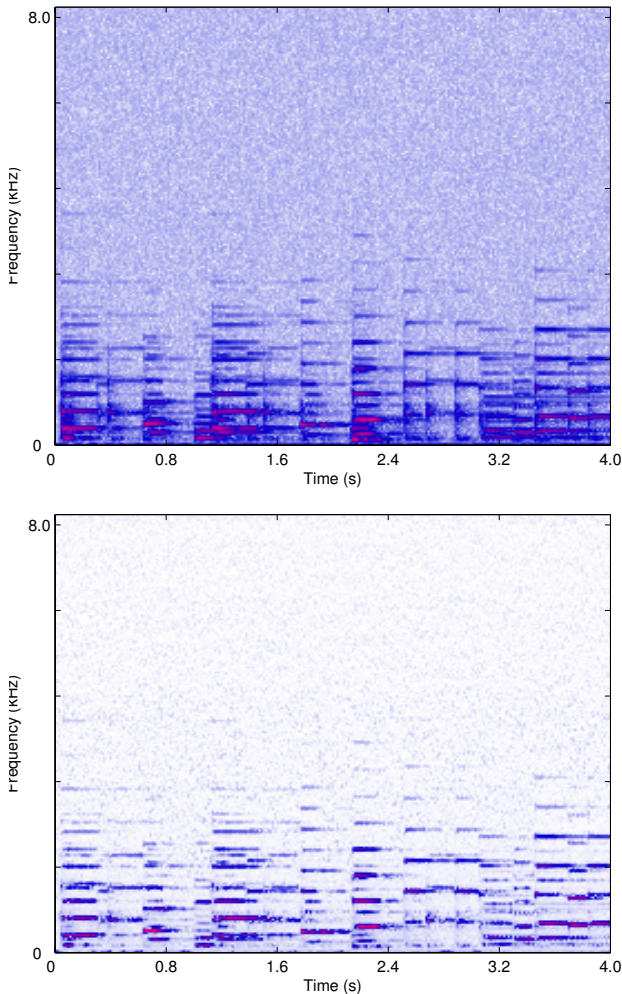


図9 RWC 研究用音楽データベース [11] 収録の “Jive (Piano Solo)” (中村真) に白色雑音を加えた信号のスペクトログラム (上) と白色雑音を抑圧した変換信号のスペクトログラム (下)。

タイトル、アーティスト名、カタログ番号はそれぞれ以下の通りである。

- PROLOGUE, 緒方智美, RWC-MDB-P-2001 No.7
- Lounge Away, YOU Band, RWC-MDB-J-2001 No.28

左中図と左下図を見てみると、左上図の原スペクトログラムの中で垂直方向に及ぶドラム音と水平方向に及ぶボーカル音のスペクトル成分が選択的に強調および抑圧されていることが分かる。また、聴感上もほとんど劣化なくドラム音やボーカル音の音量や音質を調節できることが確認できた。右中図では、右上図の通り極めて小さい音量で収録されているドラム音をサックスなどの音源の音量を抑えながら強調することができている。

また、白色雑音を付加した楽曲信号に対する雑音抑圧効果を確認するため、RWC 研究用音楽データベース [11] 収録の

- Jive (Piano Solo), 中村真, RWC-MDB-J-2001 No.1

に白色雑音を加えた信号を実験データとして非周期成分の抑圧を行った。図9に結果例を示す。原スペクトルの時間周波数領域に渡ってほぼ一様に分布する雑音成分が、ボトムエンベロープ操作を介した振幅スペクトル変換により低減されていることが分かる。

6 まとめ

本稿では、混合音中の周期成分および非周期成分を選択的にレベル補正や加工を施すことを目的とし、混合音スペクトル中のラインスペクトル状と平坦な成分を個別操作する方法としてスペクトル制御エンベロープを介したスペクトルの変形方法を提案し、GUIを実装した。本GUIを用い、ポピュラーやジャズ音楽などに含まれるドラム音の強調および抑圧効果や、白色雑音が混入された音楽音響信号に対する雑音抑圧効果の例を示し、提案方法の基本性能を確認した。

謝辞

本研究のインタフェース実装に関してご助言とご協力頂いた中野倫靖氏 (筑波大) に深く感謝する。

参考文献

- [1] Rose, J., “Audio Postproduction for Digital Video (DV Expert Series),” *CMP Books*, 2002.
- [2] M. N. S. Swamy and K. S. Thyagarajan, “Digital Bandpass and Bandstop Filters with Variable Center Frequency and Bandwidth,” *Proc. IEEE*, Vol. 64, pp. 1632–1634, 1976.
- [3] S. Erfani and B. Peikari, “Variable Cutt-off Digital Ladder Filters,” *Int. J. Electron.*, Vol. 44, pp. 535–549, 1978.
- [4] E. C. Tan, “Variable Lowpass Wave-digital Filters,” *Electron. Lett.*, Vol. 18, pp. 324–326, 1982.
- [5] P. A. Regalia and S. K. Mitra, “Tunable Digital Frequency Response Equalization Filters,” *IEEE Trans. Acoust., Speech, Signal Process.*, ASSP-35, No. 1, pp. 118–120, 1987.
- [6] S. J. Orfanidis, “Digital Parametric Equalizer Design with Prescribed Nyquist-frequency Gain,” *Journal of the Audio Engineering Society*, Vol. 45, No. 6, pp. 444–455, 1997.
- [7] Yoshii, K., Goto, M. and Okuno, H. G., “INTER:D: A Drum Sound Equalizer for Controlling Volume and Timbre of Drums,” In *Proc. 2nd European Workshop on the Integration of Knowledge, Semantic and Digital Media Technologies (EWIMT2005)*, 2005.
- [8] 吉井和佳, 後藤真孝, 駒谷和範, 尾形哲也, 奥乃博, “Drum: ドラムパートのリアルタイム編集機能付きオーディオプレイヤー,” インタラクシオン 2006 論文集, 2006.
- [9] Yoshii, K., Goto, M., Komatani, K., Ogata, T., Okuno, H. G., “An Error Correction Framework based on Drum Pattern Periodicity for Improving Drum Sound Detection,” In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP2006)*, 2006.
- [10] 板倉文忠, “統計的手法による音声分析合成系に関する研究,” 博士論文, 名古屋大学大学院工学研究科, 1972.
- [11] 後藤真孝, 橋口博樹, 西村拓一, 岡隆一, “RWC 研究用音楽データベース: 研究目的で利用可能な著作権処理済み楽曲・楽器音データベース,” *情報処理学会論文誌*, Vol.45, No.3, pp.728–738, 2004.

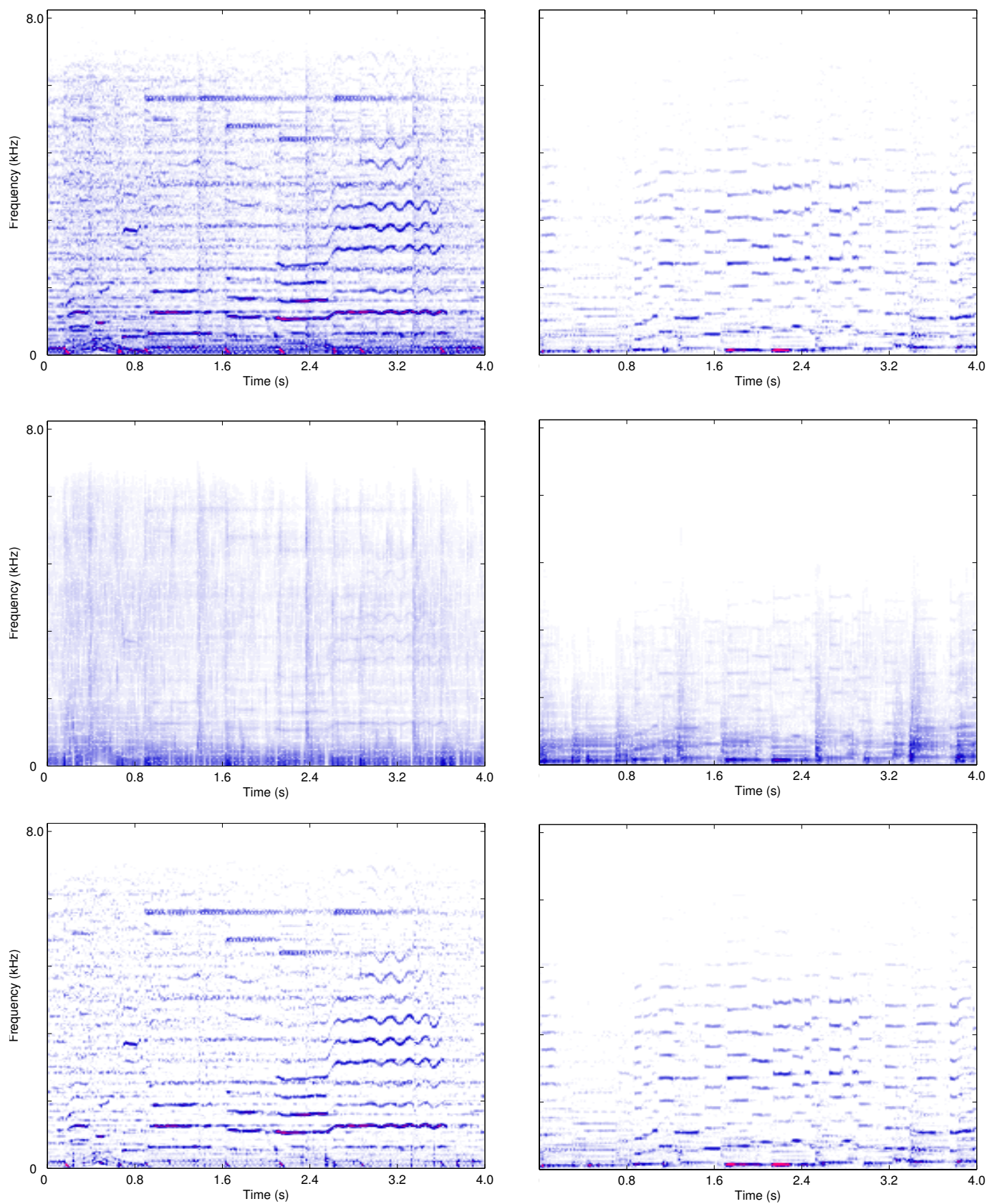


図 8 元のスペクトログラム (上) と非周期成分および周期成分を強調した変換信号のスペクトログラム (中、下)。上図は、RWC 研究用音楽データベース [11] 中の、緒方智美: “PROLOGUE” (RWC-MDB-P-2001 No. 7)(左) と YOU Band: “Lounge Away (Octet)” (RWC-MDB-J-2001 No. 28)(右) に対する解析および変換結果例である。