

# 楽音のスペクトル構造に基づいた音信号混合法

池田 友和<sup>1,a)</sup> 大脇 渉<sup>1,b)</sup> 高橋 弘太<sup>1,c)</sup>

**概要:** スマートミキサーは、複数の入力信号を時間周波数平面に展開し、平面上の局所領域において非線形な混合処理を行うことで、使用者の優先させたい音を目立たせた混合出力を得ることのできる新しい音信号混合法である。従来のスマートミキサーは、音声とBGMの混合を対象とした局所的な処理であり、楽音同士の混合については検討されていなかった。本研究では、楽音同士の混合を対象として、スペクトルの全体構造に基づいた処理を行い、協和性理論を規範とし、調和の取れた混合出力を得るミキシング手法を提案する。提案手法でのミキシング結果を聴取実験と不協和度によって評価し、楽音同士の混合における、スペクトルの全体構造に基づいた処理の不協和感の緩和への有効性を確認した。その結果、440 Hz以下の半音音程の不協和な定常音について、不協和感の緩和効果があることが確認できた。

TOMOKAZU IKEDA<sup>1,a)</sup> WATARU OWAKI<sup>1,b)</sup> KOTA TAKAHASHI<sup>1,c)</sup>

## 1. はじめに

スマートミキサー [1] は、我々が提案する新しいミキシング技術である。これまで研究されてきたスマートミキサーでは、音声とBGMが同時に再生される環境において、局所的な処理により音声を目立たせていた。本研究では、楽音同士が同時に再生される環境において、スペクトル構造に基づいた処理により調和の取れた音を作ることを目的としたスマートミキサーの新たなアルゴリズムを提案する。

本研究の最終目標は、スマートミキサーによって、音楽理論や作曲理論における拘束を緩和し、より自由な作曲を可能にすることである。作曲を行う際、音楽理論や作曲理論を踏まえた結果、諦めなくてはならない音色や音程、フレーズが生じる。そのような本来諦めなくてはならない音を、スマートミキサーによって、使えるようにしたいと著者は考えている。

ところで、どのような音が調和の取れた音であるのだろうか。協和な音程だけで構成されている音、主役と脇役がはっきりわかる音など、人によって解釈は異なるであろう。音の調和は主観的なものであり、基準が必要である。

そこで、今回は、調和の取れた音として、心理音響的な不協和 [2] を抑制した音とした。心理音響的な不協和は、うなりにより生じる音の粗さを由来とする協和と不協和の概念であり、不協和度によってモデル化されている。

本稿では、うなりに起因する倍音を抑制し、心理音響的な不協和を緩和するミキシング手法の有効性の評価を示す。

## 2. スマートミキサー

スマートミキサーの基本原理は、入力信号を短時間フーリエ変換によって時間周波数平面に展開し、時間周波数平面上で非線形な重ね合わせを行うことである。

従来の音信号のミキサーは、単純な加算器である。入力信号をそのまま加算すると、加算後のスペクトル上で各入力信号のスペクトルの特徴が崩れうる。このとき、混合音において聞き取りにくい部分が生じる。そこで、プロのミキシングエンジニアは、聴感上好ましい出力信号に仕上げるために、加算の前後に線形・非線形な様々な処理を追加する。これらの追加処理は、豊富な知識や経験に基づいた高度な技術が必要であり、素人には難しい。

そこで、ミキサー自体を高度化して、入力信号を解析し処理を行う機構を持たせたスマートミキサーが考案された。

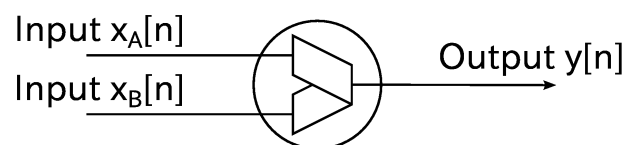


図 1 スマートミキサーを表す記号

<sup>1</sup> 電気通信大学情報理工学研究所  
Graduate School of Informatics and Engineering,  
The University of Electro-Communications,  
Chofugaoka 1-5-1, Chofu-shi, Tokyo, 182-8585, Japan  
a) ikeda-tomokazu@uec.ac.jp  
b) wataru@uec.ac.jp  
c) kota.t@uec.ac.jp

これまでのスマートミキサーの研究として、我々は、音量関係を利用した手法 [1] や調波構造を利用した手法 [3], フォルマントを規範とした手法 [4], 調波構造とフォルマントを組み合わせた手法 [5], 騒音下に適応させた手法 [6] などを提案してきた。いずれも、音声と音声以外の音（以下、BGM と呼ぶ）の混合を前提とし、音声と BGM の間に優先関係を設け、音声を目立たせる処理を行っている。一例として、調波構造を利用したスマートミキサーで音声と BGM の混合処理を行った結果の時間周波数平面を図 2 に示す。音声には SRV-DB[7] の「発話のプロフェッショナルによる編集手帳（読売新聞）の読み上げ」のデータを、BGM には「RWC 研究用音楽データベース：ポピュラー音楽 [8]」のデータを用いている。

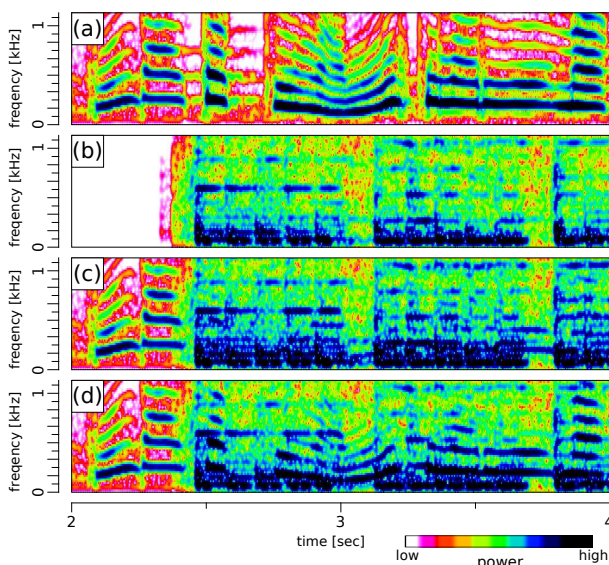


図 2 スマートミキサー [3] の処理結果の時間周波数平面。  
 (a) 音声, (b)BGM, (c) 単純加算, (d) スマートミキサー

図 2 において、単純加算 (c) では、音声のスペクトルが BGM のスペクトルに埋もれてしまっているが、スマートミキサー (d) では、音声のスペクトルが浮き出ていることがわかる。また、聴感上も音声聞き取りやすくなっていることが聴取実験により明らかになっている。

以上のように、これまでは聞き手に焦点を合わせ、優先関係を設けることで、音声の聞き取りを良くする研究が行われてきた。しかし、優先関係は設けず、音自体を磨き、理想の音色やハーモニーにするという研究はされていなかった。そこで、本研究では、楽音同士の混合を前提として、調和の取れた音を作り出す新たなスマートミキサーのアルゴリズムを提案する。調和の取れた音を自由に作り出すことができれば、音程関係に依らずに、協和・不協和感を操ることができるであろう。そして、将来的には、楽音同士のスマートミキサーでミキシングすることにより、作曲の原則や禁則を緩和した、より自由な、全く新しい作曲が可能になるであろう。

今回は、前提としてヴァイオリンのような倍音列がおおよそ整数倍になる単旋律同士の混合を考える。また、調和の取れた音の規範として、今回は協和性理論を用いる。協和性理論とは、うなりにより生じる心理音響的な不協和を不協和度で定量化した理論である [9]。

### 3. 不協和度

音楽的な不協和音と心理音響的な不協和音は異なる [2]。音楽的な不協和音は、西洋調性音楽の構造に基づいた音程関係（度数）やコードによって表される。不協和音は緊張感を生み出し、協和音へ戻ろうとする力を持つ。そして、西洋調性音楽の構造に基づいた不協和音の定義は音楽のスタイルや、時代に依存する。例えば、10 世紀においては、V 度の音程関係は協和音とされていなかったが、16 世紀以降では、オクターブと同等の協和音とされている。一方、心理音響的な不協和音は同時に鳴らされた特定の音のラフネス（粗さ）に由来するものである。Helmholz は、聴覚システムについて「重なっているハーモニーにおいてうなりと荒々しさを好まない」と予測し、心理音響的な不協和音を提唱した [10]。この、心理音響的な不協和音は音楽のスタイルや時代の違いで変化しないとされている。

その後、Plomp と Levelt の心理実験 [11] により、2 純音の周波数差と不協和の度合いを表す不協和度曲線（図 3）が提案された。

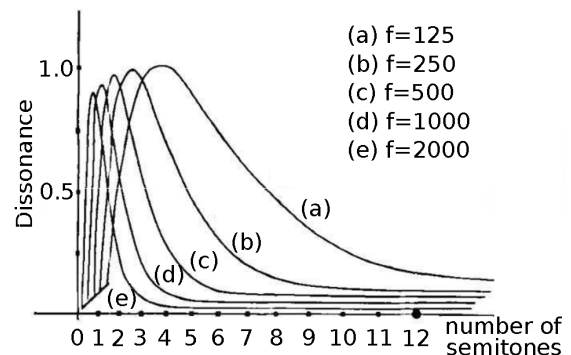


図 3 不協和度曲線 [12]

また、Sethares は不協和度曲線と臨界帯域幅の関係を基に、倍音まで含めた 2 音の不協和度を  $D$  として以下のように定義した [12]。

$$D = \frac{1}{2} \sum_{p=1}^P \sum_{q=1}^Q v_1[p]v_2[q] \quad (1)$$

$$\times \left( e^{-as(f_2[q]-f_1[p])} - e^{-bs(f_2[q]-f_1[p])} \right) \quad (2)$$

$$s = \frac{d^*}{s_1 f_1 + s_2}$$

ここで、第 1 の音の第  $p$  倍音 ( $1 \leq p \leq P$ ) の周波数を  $f_1[p]$ , 第 2 の音の第  $q$  倍音 ( $1 \leq q \leq Q$ ) の周波数を  $f_2[q]$ , 第 1 の音の第  $p$  倍音と第 1 倍音のパワー比を  $v_1[p]$ , 第 2 の音の第  $q$  倍音と第 1 倍音のパワー比を  $v_2[q]$  とした。ま

た, Plomp と Levelt の心理実験との比較によって  $a = 3.5$ ,  $b = 5.75$  としている.  $s$  は音域で異なる臨界帯域幅に対応するための関数であり,  $s_1 = 0.021$ ,  $s_2 = 19.0$ , 最大不協和点  $d^* = 0.24$  としている. 以下の図は,  $v_1[p]$  と  $v_2[q]$  を変えた 3 パターンについて, 音程の周波数差と不協和度の関係を算出した図である.

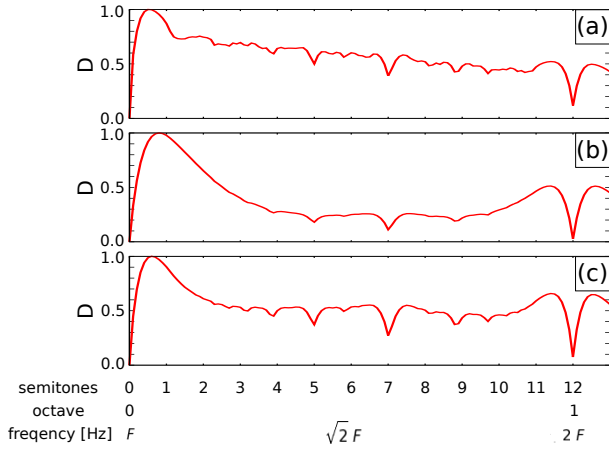


図 4 各度数の音程と不協和度の関係. (a)  $v_1[p] = v_2[q] = 1.0$ , (b)  $v_1[p] = \frac{1}{p}, v_2[q] = \frac{1}{q}$ , (c)  $v_1[p] = 0.88^{p-1}, v_2[q] = 0.88^{q-1}$

## 4. 提案手法

心理音響的な不協和音はうなりが原因である. そこで, 本稿では倍音同士のうなりを抑制する新しいミキシング手法を提案する. 提案手法のブロック図を図 5 に示す.

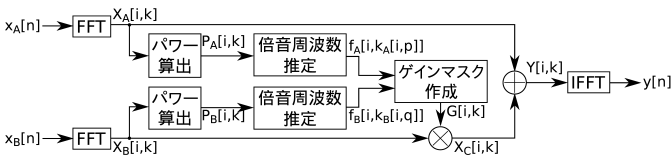


図 5 提案手法のブロック図

入力信号を  $x_A[n]$ ,  $x_B[n]$  とする. ここで  $n$  は時間サンプル番号を表す. まず, それぞれの入力信号を短時間フーリエ変換により, 複素時間周波数表現  $X_A[i, k]$ ,  $X_B[i, k]$  に展開する. ここで  $i$  は時間フレーム番号,  $k$  は周波数ビン番号である. 次に,  $i$  フレーム,  $k$  ビンにおけるパワー  $P_A[i, k]$ ,  $P_B[i, k]$  を以下のように算出する.

$$P_A[i, k] = X_{A(\text{Re})}^2[i, k] + X_{A(\text{Im})}^2[i, k] \quad (3)$$

$$P_B[i, k] = X_{B(\text{Re})}^2[i, k] + X_{B(\text{Im})}^2[i, k] \quad (4)$$

次に, パワーを基に倍音周波数推定を行い, 各倍音の推定周波数  $f_A[i, k_A[i, p]]$ ,  $f_B[i, k_B[i, q]]$  を求める. ここで  $k_A[i, p]$  は入力音  $A$  の  $i$  フレームにおける第  $p$  次倍音に該当するビン番号を表し,  $k_B[i, q]$  は入力音  $B$  の  $i$  フレームにおける第  $q$  次倍音に該当するビン番号を表す. そして, 各入力音の推定倍音周波数から倍音同士のうなりを抑制するための

ゲインマスク  $G[i, k]$  を作成する. 作成したゲインマスク  $G[i, k]$  を  $X_B[i, k]$  に乗算することによって, 加算したときうなりが生じないようにうなりに起因する入力音  $B$  の倍音を抑制した複素時間周波数表現  $X_C[i, k]$  を得る.

$$X_C[i, k] = G[i, k]X_B[i, k] \quad (5)$$

その後,  $X_A[i, k]$  と  $X_C[i, k]$  を時間周波数平面上で加算した  $Y[i, k]$  を算出する.

$$Y[i, k] = X_A[i, k] + X_C[i, k] \quad (6)$$

$Y[i, k]$  を短時間逆フーリエ変換し, 倍音同士のうなりが抑制された出力信号  $y[n]$  を得る.

### 4.1 倍音周波数推定

周波数推定法は, ゼロクロス法や自己相関数法, ケプストラム法, 線型予測法などの様々な方法 [13] が存在するが, 本稿では, 高速化を図るため, 時間周波数平面上で周波数を推定する独自のアルゴリズムを考案した. 倍音周波数推定の内部の処理は, 基音ビン推定, 倍音ビン推定, 周波数推定の順に行う.

#### 4.1.1 基音ビン推定

求める基音ビンを  $k_f[i]$  とする. 倍音列は基音のおおよそ整数倍上に存在するという性質を利用して基音ビン推定を行う. まず, 算出したパワーが, 周波数ビン方向に極大となる周波数ビンを求める. それらのビンの内, パワーが閾値  $\alpha[i] = aP_{\text{ave}}[i]$  以上となるビンを, 周波数の昇順に並べたときの  $m$  番目のビン番号を  $k_p[i, m]$  とする. ここで,  $P_{\text{ave}}[i]$  は  $i$  フレームにおけるパワーの平均値を表す. また,  $a$  は重み付けのための係数であり, 今回は  $a = 0.1$  とした.

次に, 極大ビン  $k_p[i, m]$  から, 隣接極大ビン差  $k_d[i, m]$  を求める.

$$k_d[i, m] = k_p[i, m + 1] - k_p[i, m] \quad (7)$$

隣接極大ビン差  $k_d[i, m]$  の中央値を基音ビン  $k_f[i]$  とする.

#### 4.1.2 倍音ビン推定

第  $l$  次の倍音に該当するビンを  $k_h[i, l]$  とする. 理想的な倍音列であれば, 第  $l$  次倍音ビンは  $k_h[i, l] = lk_f[i]$  で算出できるが, 周波数分解能によるずれや, 入力音源の倍音列が完全な整数倍になっていない可能性を考慮し,  $lk_f[i]$  から上下  $\beta[k_h[i, l]]$  の領域内におけるパワーが最大のビンを第  $l$  次倍音ビン  $k_h[i, l]$  とする. ここで  $\beta[k_h[i, l]]$  は倍音次数ごとに計算領域決定のためのパラメータであり, MPEG オーディオにおける国際規格に用いられている聴覚心理モデルの純音判定 [14] に用いる帯域幅を規範とした.

$$\beta[k_h[i, l]] = \begin{cases} 2 & (2 < k_h[i, l] < 63) \\ 3 & (64 \leq k_h[i, l] < 127) \\ 6 & (128 \leq k_h[i, l] < 255) \\ 12 & (256 \leq k_h[i, l]) \end{cases} \quad (8)$$

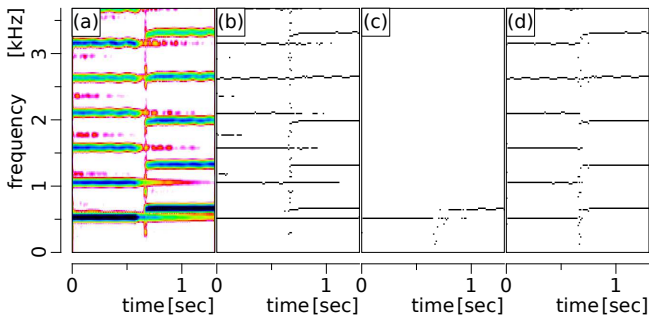


図 6 パワー算出から倍音ビン推定までの各時間周波数平面。  
(a) パワー, (b) 極大値, (c) 基音ビン, (d) 倍音ビン

#### 4.1.3 周波数推定

第  $\ell$  次倍音の推定周波数を  $f_h[i, k_h[i, \ell]]$  とする。今回は、4 次のラグランジュ補間で補間関数を作成し、補間関数上のパワーが最大となる周波数ビンを周波数に変換する方法で倍音周波数  $f_h[i, k_h[i, \ell]]$  の推定を行った。

#### 4.2 ゲインマスク作成

まず、4.1 節の手法で入力音  $A, B$  それぞれの倍音周波数  $f_A[i, k_A[i, p]], f_B[i, k_B[i, q]]$  の周波数差  $f_d[i, p, q]$  を算出する。

$$f_d[i, p, q] = |f_A[i, k_A[i, p]] - f_B[i, k_B[i, q]]| \quad (9)$$

うなりと見做す上限周波数を  $f_u$  とする。算出した  $f_d[i, p, q]$  の値が  $f_u$  以下であるとき、 $k_B[i, q]$  を入力音  $A$  に対する入力音  $B$  のうなりに起因する倍音ビンとして検出し、そのビンを中心に幅  $W$  の領域は  $G[i, k] = 0$ 、領域外は  $G[i, k] = 1$  であるゲインマスク  $G[i, k]$  を作成する。今回は  $W = 5$  と設定した。 $f_u$  の値については 5.1 節に記載する。

ゲインマスク  $G[i, k]$  を  $X_B[i, k]$  に乗算することでうなりに起因する入力音  $B$  の倍音を抑制した  $X_C[i, k]$  を得る。

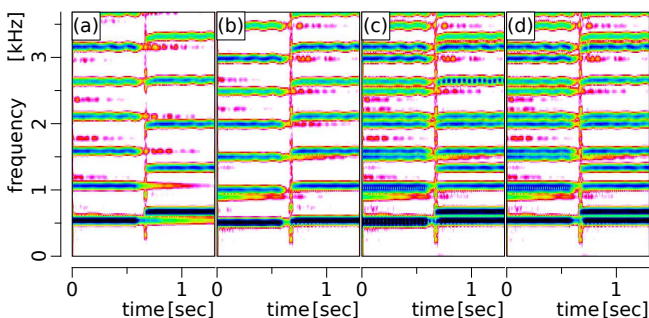


図 7 提案手法による時間周波数平面の比較。  
(a) 入力音 A, (b) 入力音 B, (c) 単純加算, (d) 提案手法

## 5. 実験

実信号を用いてミキシングを行い、単純加算との比較を聴取実験と不協和度によって評価した。

### 5.1 実験諸元

実験諸元を以下に示す。

表 1 実験諸元

パラメータ	記号	値
サンプリング周波数	$F_s$	44100 Hz
量子化ビット数	$N_{bit}$	16 bit
FFT 点数	$N$	2048 点
解析時窓関数	$N_F$	ハニング窓形 1024 点
合成時窓関数	$N_I$	ハニング窓形 512 点
フレームシフト幅	$N_S$	512 点
極大値決定係数	$a$	10
ゲインマスク抑制幅	$W$	5 ビン (107.67 Hz)
うなりと見做す上限周波数	$f_u$	提案手法 1 : 20 Hz 提案手法 2 : $\gamma f_A[i, k_A[i, 1]]$ [Hz] ( $\gamma \approx 0.0905$ )

$f_u$  について、固定値と周波数依存値の 2 種類を候補とした。まず、固定値として、 $f_u = 20$  を提案手法 1 とした。次に、周波数依存値として、 $f_u$  を基音の周波数に依存させた  $f_u = \gamma f_A[i, k_A[i, 1]]$  を提案手法 2 とした。なお、 $\gamma$  は主観実験により、平均律音程における、入力音  $A$  の基音の周波数と 1.5 半音上の周波数との差になるように  $\gamma = 2^{\frac{1.5}{12}} - 1 \approx 0.0905$  とした。

入力信号には以下に示す 6 個のセットを使用した。

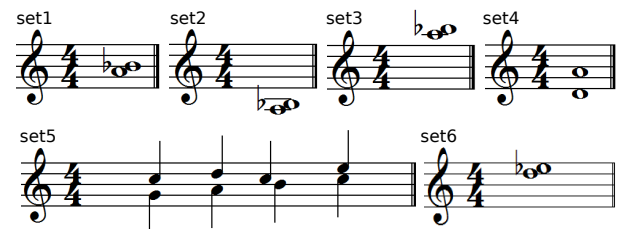


図 8 実験に使用した音源セット

set1 から set4 までの入力音は、基本周波数の正弦波に、第 50 次までの整数倍音の正弦波を重ね合わせた混合波を作成し使用した。なお、第  $n$  次倍音の振幅を  $\frac{1}{n}$  とし、各倍音の正弦波の初期位相をランダムとした。set1, set2, set3 は、それぞれ別の周波数帯域における、半音で重なる不協和音の混合である。set4 は set1 と同じ周波数帯域における、完全 V 度で重なる協和音の混合である。set2, set3 は set1 との比較により、同じ不協和な音程関係で周波数帯域が異なる場合での効果の程を確認する目的で作成した。また、set4 は set1 との比較により、同じ周波数帯域での協和な音程と不協和な音程での効果の程を確認する目的で作成した。set5 は、シンセサイザで作成したヴァイオリンの音色で、0.75 秒毎に別の音程に推移するフレーズ同士の混合である。set6 の入力音は、1.0 秒のヴァイオリン演奏音 (ロングトーン) で、半音で重なる不協和な混合である。

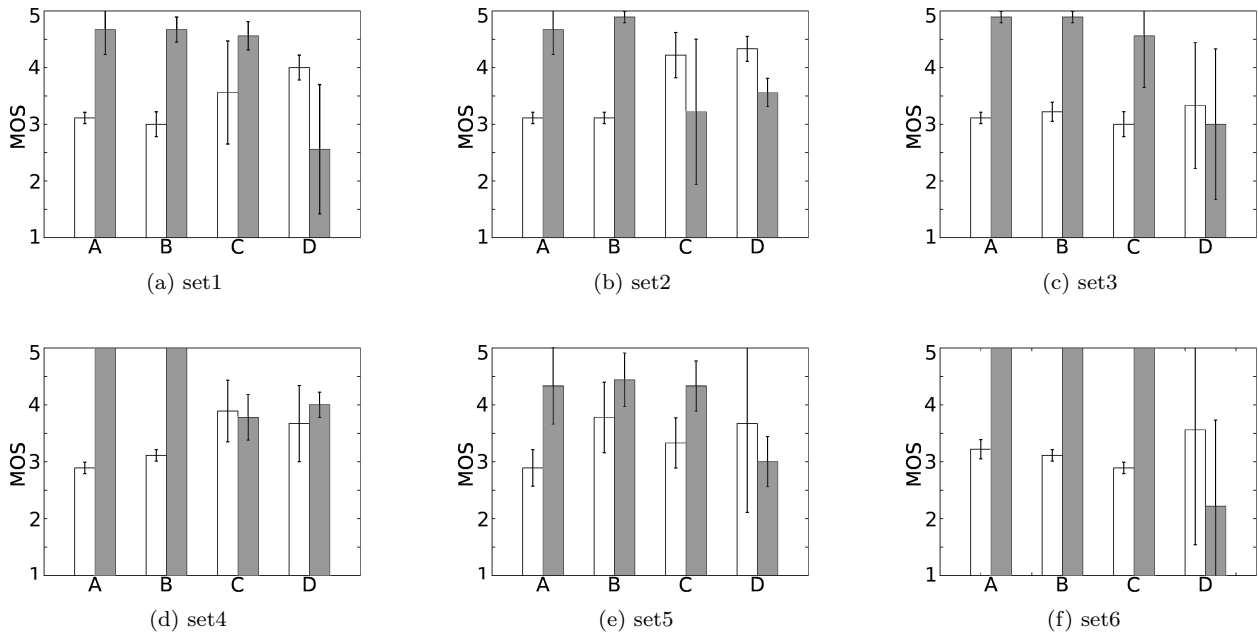


図9 聴取実験結果。(白色：不快度の緩和, 灰色：音色変化)

## 5.2 聴取実験

各音源 set について表 2 に示す 4 種類の処理を行った。各処理に対し、単純加算と比較した「不快度の緩和」と「音色変化」の 2 つの評価項目について、5 段階の MOS 値 (1: Bad, 2: Poor, 3: Fair, 4: Good, 5: Excellent) の評価を行った。被験者は成人男性 9 名であり、全員ヘッドフォンで聴取を行った。結果を図 9 に示す。

表 2 各処理内容の説明

記号	内容
A	単純加算 (各信号に対し一切の処理を行わずに加算)
B	従来スマートミキサー (調波構造に基づいた処理 [3])
C	提案手法 1 (20 Hz 以下のうなりを抑制)
D	提案手法 2 ( $\gamma f_A[i, k_A[1]]$ [Hz] 以下のうなりを抑制)

## 5.3 不協和度による評価

処理後の音源について、式 (1) より各時刻における不協和度を算出した。算出された不協和度について、単純加算と比較した減少値の時間変化を図 10 に示す。

## 6. 考察

提案手法におけるミキシング結果の不快度の緩和と音色変化について考察する。

まず、不快度の緩和について図 9 を用いて考察する。set1, set2, set3 において、提案手法 1 および提案手法 2 を単純加算と比較すると、set1 と set2 ではいずれの提案手法も不快度を緩和する効果が見られるが、set3 では見られない。うなりに起因する倍音の周波数は、set1 では約 4 kHz, set2 では約 8 kHz であったのに対し、set3 では約 16 kHz という高周波数帯域であった。そのため、そもそも単純加算時に生じているうなりに気づいておらず、抑制しても抑制前との変化に気が付かなかったと考えられる。このことから、提案手法は基音の周波数が低いほど効果を発揮

すると思われる。また、set1 と set4 より、提案手法は、協和音程でも不協和音程でも同等の効果を発揮することがわかる。平均律音程において整数倍の倍音列で構成された 2 音が完全 V 度の音程である時、1 度音の第 3(n+1) 次倍音と V 度音の第 2(n+1) 次倍音 (n は自然数) は低いうなり周波数 (約 1.955 cent) でうなる。完全 V 度の音程でこのうなりが発生しないように調整した音律が純正律音程であるので、この聴取実験から、純正律の完全 V 度の方が平均律の完全 V 度よりも心地よい響きであるということも示唆される。set5, set6 を見ると、提案手法 1 よりも提案手法 2 の方が点数が高く、提案手法 2 の分散が大きくなっている。このことから、入力音がフレーズの場合や、演奏音の場合は、不快度の緩和効果は多少あるものの、個人差が大きくなると考えられる。

次に、音色変化について図 9 と図 10 を用いて考察する。set1 について、図 10 では提案手法 2 のみ不協和度の減少値が高く、図 9 では提案手法 2 のみ音色変化の得点が低い。この傾向はいずれのセットでも見られる。また、set6 は他のセットに比べて不協和度の減少値が大きいのに対し、音色変化の得点が著しく低い。このことから、不協和度と音色変化には相関関係があることがわかる。

以上より、うなりの抑制は基音が 440 Hz 以下の半音音程の不協和な定常音の不快度を緩和させる効果があるが、同時に音色変化も知覚させてしまうことがわかった。

また、図 10 より、提案手法 1 では不協和度がほぼ減少していないのに対し、提案手法 2 では、いずれの音源でも不協和度が減少し、特に不協和な関係の混合時に減少している。このことから、うなりと見做す上限周波数  $f_u$  は基音の周波数に依存させるべきであり、提案手法は不協和度の減少に効果があるということがわかる。

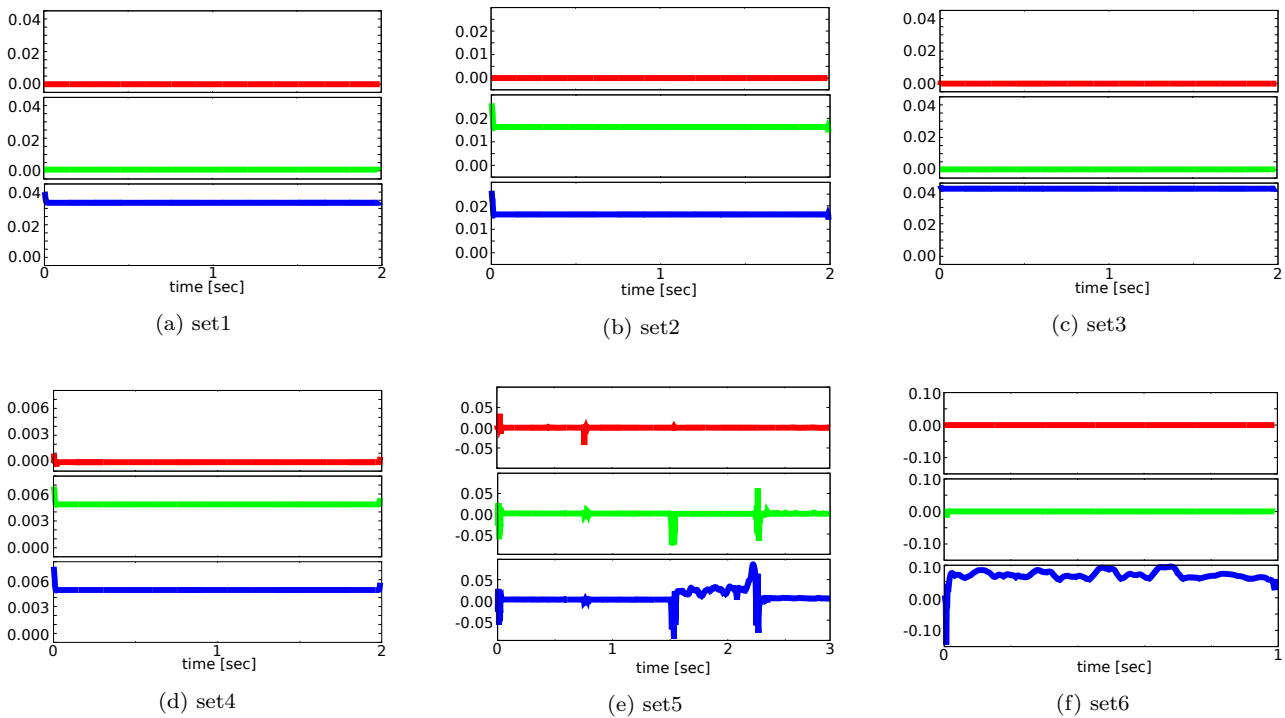


図 10 不協和度の減少値. (上から, 従来スマートミキサー, 提案手法 1, 提案手法 2)

## 7. 結論

本稿では, 楽音を対象とし, 協和性理論を規範とした新しいスマートミキサーのアルゴリズムを提案した. はじめに, 倍音同士のうなりを抑制する手法を考案し, プログラム上での実装を行った. 次に, 実音源によるミキシング結果を聴取実験と不協和度によって評価した. その結果, 第一に, 440 Hz 以下の半音音程でぶつかる定常音について, 倍音同士のうなりの抑制は不協和感の緩和効果があることがわかった. 第二に, 定常的でない音や演奏音の混合に対しては, 不協和感の緩和効果は個人差が大きく, うなりの抑制だけでは不十分だとわかった. 第三に, 不協和度の変化と音色変化には相関があることがわかった. 第四に, うなりと見做す上限周波数は固定ではなく, 基音の周波数に依存させるべきであることがわかった.

## 8. 今後の課題

今後は, うなりのパワー, 発生周波数帯域, うなり周波数を基に抑制量を最適化するアルゴリズムを導入する. また, スペクトル重心を用いた, 音色変化・音質劣化の知覚を抑えるアルゴリズムや前後の音程や調情報などの音楽情報を取り入れたアルゴリズムの導入を行っていく.

### 謝辞

本研究は JSPS 科研費 20500111, 23500147, 26330188 の助成を受けたものです.

### 参考文献

- [1] 大脇渉, 高橋弘太 “スマートミキサー ~新しい音信号混合合法~”, 信学技報, vol.102, pp.37-42, (2011).
- [2] Albert S, Bregman: *Auditory Scene Analysis*, The MIT Press, pp.502-511, (1990).
- [3] Owaki, W, Takahashi, K “Novel sound mixing method for voice and background music”, Proceedings of IEEE, ICCASP, pp.290-294, (2015).
- [4] 大脇渉, 旭岡舜介, 高橋弘太 “フォルマント構造維持を規範とした音声信号混合合法”, 研究報告音楽情報科学 (MUS), vol.107, no.66, pp.1-4, (2015).
- [5] 大脇渉, 高橋弘太 “歯擦音と母音の識別性を重視する音声信号混合合法”, 信学技報, vol.115, no.126, pp.7-10, (2015).
- [6] 長谷川政良, 大脇渉, 高橋弘太 “騒音下における音声再生の識別性を重視したスマートミキサー”, 騒音・振動研究会, N-2015-40, (2015).
- [7] “話速バリエーション型音声データベース:SRV-DB”, 入手先 (<http://www.it.ice.uec.ac.jp/SRV-DB/>) (2015.10.09).
- [8] 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:ポピュラー音楽データベースと著作権切れ音楽データベース”, 研究報告音楽情報科学 (MUS), vol.42, no.6, pp.35-42, (2001).
- [9] 内田遼, 矢向正人, “協和性理論を用いた歌唱時音声の分析手法に関する研究”, 研究報告音楽情報科学 (MUS), vol.97, no.14, (2012).
- [10] ホアン・G・ローダラー, 高野光司, 安藤四一: 新版 音楽の科学 音楽の物理学, 精神物理学入門, 音楽之友社 (2014).
- [11] R. Plomp and W. J. M. Levelt, “Tonal Consonance and Critical Bandwidth”, J. Acoust. Soc. Am.38, pp.548-560, (1965).
- [12] William A, Sethares, “Local consonance and the relationship between timbre and scale”, Journal of the Acoustical Society of America, vol.94, no.3, pp.1218-1228, (1993).
- [13] 古井貞熙: 新音響・音声工学, 近代科学社 (2006).
- [14] ISO/IEC 11172-3. Information technology - Coding of moving pictures and associated audio for digital stage media at up to about 1.5Mbit/s - Part3: Audio, (1993).