

音楽を探る —オペレーションズ・リサーチ的手法を使って—

Julián Villegas, Michael Cohen

オペレーションズ・リサーチの音楽への適用例を2つ紹介する。はじめに、音階を多次元多様体へマッピングする手法とその利点を述べる。その応用として、らせん鍵盤および3次元音階をインタラクティブに聴覚的・視覚的に表現する方法について解説する。次に、ローカルミニマム探索手法により音響ストリームの不協和音をリアルタイムに軽減する方法について述べる。

キーワード：Mathematical modeling, OR practice, discrete optimization, musical scales, automatic reintonation

1. イントロダクション

オペレーションズ・リサーチにおいては、問題の理解・解決に数理モデルが重要な役割を担っている。特に、生産システムなどの問題解決では、顕著な成功を収めている。また、その方法論の背後にあるアイデアの普遍性により、本質的に異なる様々な分野へも応用されている。音楽理論では、例えば、Chew[4]が、音楽計算と音楽認知理論の研究室を作り、音色のリアルタイム解析、可視化、情報取得、音楽生成などの音楽問題に取り組んでいる。

音楽理論が数学の力を借りているのと同様に、数学も、Tymoczko[12]が行っているように、高次元空間の解析に音楽の力を借りている。Tymoczkoによれば、作曲家ショパンは、数学者が取り組むはるか以前に、4次元構造をシステムティックに使っていたようである。音楽と数学の関係の成り立ちはともかく、その強い結びつきは否定できず、その歴史は複雑に絡み合っていることは、事実である。

本論では、ORを使って二つの音楽問題にアプローチする。どちらの問題も長く研究されてきたが、完全な解答は、現在までのところ得られていない。はじめに、長調 (major) と短調 (minor) の和音の違いの本質を、数学的モデルで考察する。次に、ライブ演奏中に、完全に協和された音色を得る方法について述べる。これらの方法について述べることにより、ORの

方法の有用性を示し、さらに、新たな応用先を探るきっかけになることを目的とする。

2. 音階とトポロジー

音楽の起源は、音である。その現象のはかなさから、古来より、ほとんどの文明で音楽のコンテンツを記録・保存するようなメカニズムの考案が追求されてきた。中でも、11世紀ごろに発明された Arezzo の記譜法 (五線譜) は、その普遍性と一般への普及により最も成功したものといえる。五線譜では、低い周波数が下の位置の音符に、高い周波数が上の位置の音符によって表される。最もよく使われる、7種類の音 (ドレミファソラシド) と変化記号 (シャープ、フラット) でオクターブを分解する (1オクターブに12の音階) という方法も Arezzo が提案したものである。

2.1 音階表現と音の再生

音階のシンボリックな表現と実際の音には、本質的な差がある。音階は、音のピッチ (高低) の離散的な記録であるのに対して、実際の音は、音程 (interval)、チューニング (tuning)、イントネーション (intonation) によって様々なピッチが異なる。ここでは、音符に対応する周波数 (チューニング) と楽器が奏でる実際の音 (イントネーション) を区別する。音色と音のピッチの関係の問題は、量子力学の不確定原理と似ている。量子力学では、観測を含む全体が、個別の量子的なイベントに決定的な役割を果たし、実在に影響を与える[8]。音楽では、音楽的コンテンツとそれを聴く聴衆の全体が、個別の音色を決定する。Wild & Schubert[32, p. 129]に述べられているように、A[♯]

and G[#] のような異名同音（二音間の周波数比がほぼ等しいので、音程としては同じ音）のピッチの違いは、他の音との関係（音の経路）に依存するのである。

和音 (harmony) の音程 (interval: 音の周波数の差) は、できるだけ簡単な整数倍 (pure interval: 純正律) になっているのが好ましい。例えば、基本周波数比が 3:2 に分かれる二つの音程は、完全 5 度、4:3 なら完全 4 度、5:4 なら長 3 度などと呼ばれ、これらが、音の協和を最大限に引き出すことが知られている。当然、純粋な音程だけで構成された 12 音階が望ましいが、1 オクターブの中でそのような音階を構成することは不可能であることも知られている [19]。異名同音においては、「コンマ」と呼ばれる非常に小さいピッチの差が生じる。コンマは、実際に聴き分けることが可能である。この現象を避けるために、コンマを音階上に微妙に分散させる、さまざまな調律法が提案されている。中でも、12 の音程上に均等にコンマを分散させる平均律がもっともよく使われている。

2.2 音階の幾何学的表現

五線譜は、音階（認識される音の高低と時間認識）の線形表現であるが、その他の性質については、十分表現することができない。例えば、整数倍のオクターブだけ離れた音程は、ユニゾン（同じ音程の音）以外のその他の音程に比べると近いと判定され、chroma と呼ばれる（すなわち、C とよばれる音は、すべて同じ chroma を持ち、C[#] や D などについても同様である。伝統的な西洋音楽において chroma によって可能な 12 の値は、ピッチクラスを構成する）。

完全 5 度の音程は、和音として、他の音階よりも近いと認識される。これは、5 度圏 (circle of fifth) という構造体として知られている。Shepard [20] は、5 度圏、chroma 周期、音のピッチをすべて包含した音階表現を提案している。これら 3 つの音階上の性質は、5 次元に表現され、図 1 にあるような幾何学的構造体上にマップされる。

Shepard 表現は、完全であるが実用的でない場合が多い。したがって、そのモデルをより低次元空間へ射影したものが、研究者に広く使われている。Chew & Chen [3] や Herder & Cohen [9] など、多数の研究者が、低次元モデルを使って、音色をリアルタイムで追跡、可視化する方法を提案している。これらの研究に使われているモデルをすべて述べるのは、この論文のスコープの範囲外ではあるが、以下のセクションでの議論に役立つので、Euler の tonnetz を紹介

する。

tonnetz (ドイツ語で「音色の網」) は、ピッチのクラスの格子表現であり、Euler [6] が提案し、後に Riemann が音楽分析のために使用した。この構造体は、図 1 のドーナツ状の断面に相当する。tonnetz では、図 2 のように、完全 5 度を水平線上に配置し、同レベルの連続する 2 音符と 1 レベル上の 1 音符で、長調の和音を表す。対称的に、同レベルの連続 2 音符と 1 レベル下の 1 音符で、短調の和音となる。左下から右上への斜めの変化（またはその逆）は、長 3 度 (major third: M3) であると呼ばれ、左上から右下への斜めの変化は、短 3 度 (minor third: m3) と呼ばれる。音程が完全で、tonnetz 音符が同じ名前で、別のレベルに位置しているときには、イントネーションが異なる。例えば、図 2 上の C と C^{''} では、その周波数比が、81:80 であり、syntonic comma と呼ばれる。Euler の tonnetz のこのような性質は、平均律以

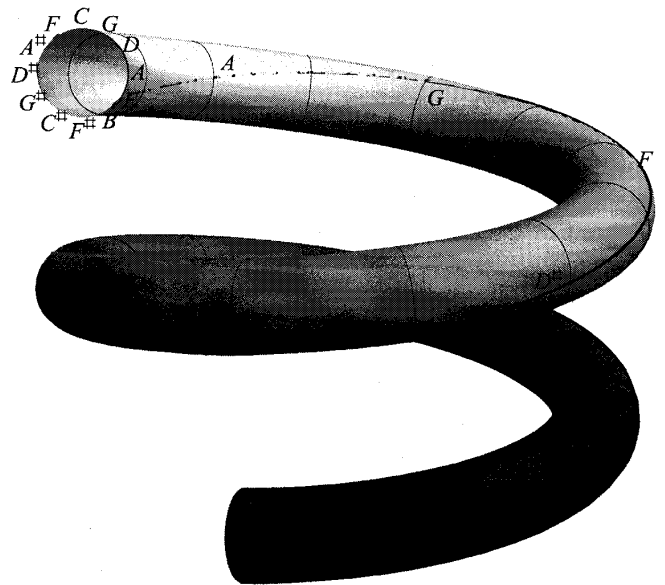


図 1 Shepard が提案する 5 次元モデル。5 度圏が断面の副軸に、chroma が主軸に相当する。ピッチは、この構造体の高さに相当する。

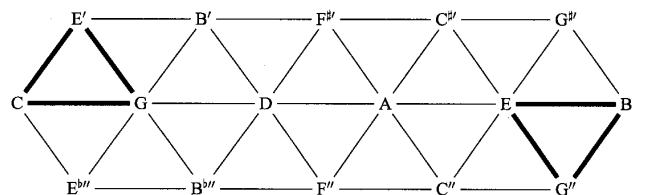


図 2 Euler の tonnetz. ハイライトされた上向きの三角形は、長調の和音の一つである (C) を表す。短調の和音は、下向きの三角形として表される (ハイライト: Em)。

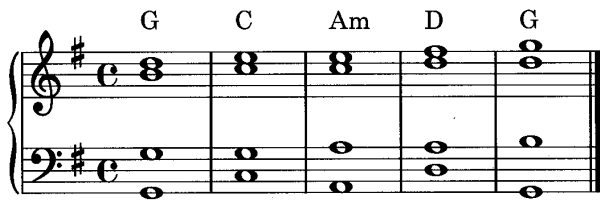


図3 はじまりと終わりにピッチドリフトのある簡単なコード進行 (I IV ii V I).

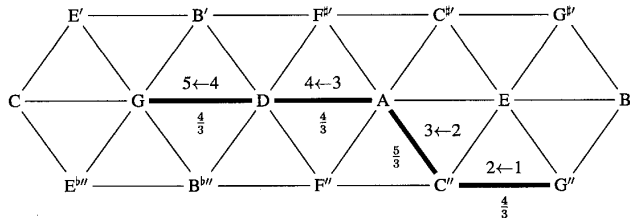


図4 図3のコード進行にある低音部のピッチドリフト解析. 太線は, ベース進行に従った経路(右から左へ)を表す. 経路の上下の数字は, それぞれ, 音符の変化と完全な音程比を表す.

外の音階を使って演奏された, 楽曲の解析に役立つ. 簡単だが, よく使われる図3のようなコード進行を考えよう. 解析を簡単にするために, 同じ名前だが異なる小節に位置する音符のイントネーションが変わらないとしよう. 図4上で, ベース(低音部)進行は, GからC, そしてA, Dを經由し, 最後にGへと移る. もし, ベース進行の音程が完全にトレースできたなら, 最後のGのピッチ(音の高さ)は, 次のように計算できる.

$$G = \frac{4}{3} \times \frac{5}{3} \times \frac{4}{3} \times \frac{4}{3} G'' = \frac{320}{81} G'' \quad (2.1)$$

式2.1をオクターブ表現の剰余類で考えれば, 同一オクターブ内でのオリジナルなG''の音程は,

$$G = \frac{320}{81} \frac{1}{2^2} G'' = \frac{80}{81} G'' \quad (2.2)$$

となる. 式2.2により, Gは, syntonic commaに相当する分だけ低いピッチであることがわかる. 数学的にピッチドリフトを評価することは, いつでも可能だが, 長い曲では, かなり苦勞することになる. しかし, tonnetzによる視覚的なチェックにより, 楽曲の音色の解析は, かなり容易になる.

2.3 ストレッチ・チューニング

既に述べたように, オクターブの分割では, 平均律がもっとも簡単な方法である. このような方法で構成された12平均律は, 対数的に一律な音階比

$$\rho = 2^{x/12} \quad (2.3)$$

で構成される. ここで, $0 \leq x \leq 11$ は, 音程 f_0 を基準とした場合の半音階を表す.

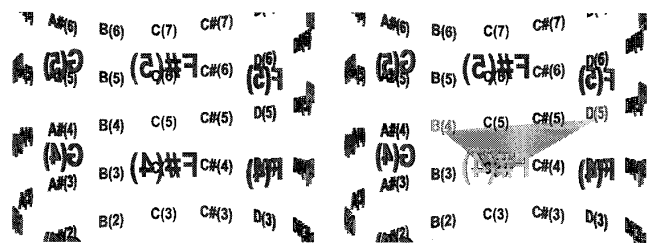
「ストレッチ・チューニング」という用語は, Mathews & Pierce[13]によって導入され, ρ 以外の音階比を使った平均律の構成方法を指す. Mathewsらは, 3種類のストレッチ・チューニングを紹介している. 「melodic」は, 音階上の音符間の音程のみが引き延ばされる. 「harmonic」では, 倍音が stretching factor 分だけあるべき位置から動かされる. 最後に, これら二つの方法のコンビネーションがあげられる. Houtsma et al.[10]とPierce[15]は, 部分音(partial)と音階が, 同じファクターで伸縮されたときには, 心地よい旋律になることを示した.

2.4 らせん鍵盤

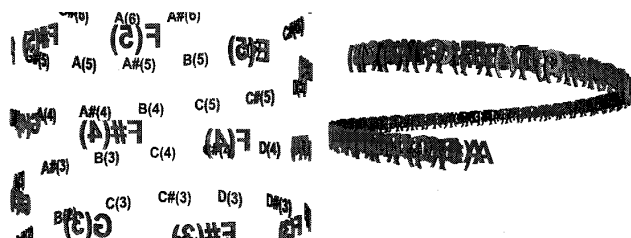
らせん鍵盤[9]は, 12-TET (12平均律)を視覚的, 聴覚的に表すことができるJavaアプリケーションである. このモデルでは, キーボードは, 左巻きのらせんで表され, もっとも低い音が, 底の部分に配置される. 図5(a)のように, 通常の(ストレッチされていない)らせんでは, 1回転が1オクターブに相当する.

らせん鍵盤によって, さまざまな座標空間表現・音表現が可能である. 図5(b)にあるように, コードのリアルタイム進行の認識は, 「chord kites」として表現される. らせん鍵盤は, デュアル偏光プロジェクター, シルバースクリーン, それにパッシブアイウェアによって, 3次元立体グラフィック化が可能である. また, chromastereoptic (色立体視)技術で3D化されたキーボードへの指示を構成することができる[23]. この技術は, 偏光レンズに比較すると簡易だが効果的であり, 研究室外でのデモ等に使用されている. オーディオは, 半球型のラウドスピーカアレイにより, そのchromaの方位に従い, 空間的に表現される[1][33]. より簡易な実装方式では, 音の空間表現は, らせんの中心のリスナーの論理位置の左右軸への空間方位の射影を使い, ヘッドホンによって表現される.

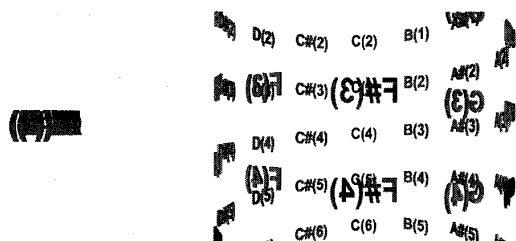
らせん鍵盤は, さまざまな機器と(3D楽譜を含む)MIDI情報を送受信することができる. ユーザーは, ジョイスティックやコンピュータキーボードを使って, 移動・ズームなどの様々な空間動作をすることにより, らせんを動き回り, 回転したり, 変形したりすることができる. このモデルは, 多様な特性表現と実時間でストレッチ・チューニングが可能であることに特徴がある[30][31].



(a) Normal (octave=2:1) (b) A chord kite for Bm chord in second inversion



(c) Stretched (pseudo-octave=2.1:1) (d) Compressed (pseudo-octave=1.1:1)



(e) The most degenerated scale (pseudo-octave=1:1) (f) Inverted (pseudo-octave=1:2)

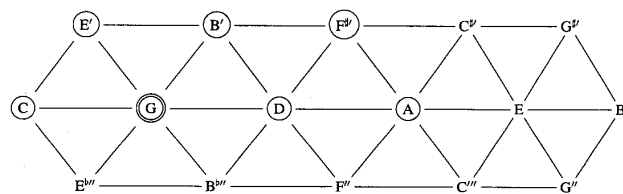
図5 異なる収縮率を持つらせん鍵盤. 残念ながら, 白黒印刷では, chromastereoptic (色立体視) 技術による立体視は再現できない.

2.5 らせん鍵盤におけるストレッチ・チューニング

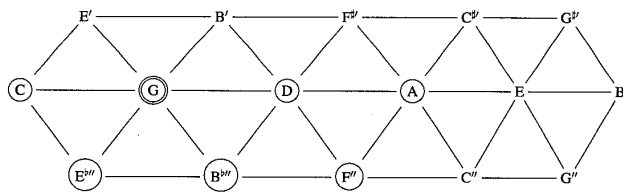
ユーザーが音ストリームを空間的にレンダリングする際に, 自在にストレッチを行うことは, とても面白い音楽的体験となる. 収縮操作を上手に使えば, 図5(e)のように音階全体を一つの音色 (部分波形のない, 単一の sin 波形) に収縮させたり, 逆に, 図5(f)のように, ストレッチ率を1以下にし, さまざまな平均律を作り出すことや, 単位周辺のハーモニーに収縮することにより, 複雑なビートを作り出すことができる. 図5は, そのような例を示している.

2.6 短三和音 (Minor Triad) 問題

短調のコード (minor chord) は, 短3度 (minor third) の上に長3度 (major third) を重ねることで作ることができる. 短調のコードは, 調和性音楽に重要な役割を占めるが, 長調のコード (major chord)



(a) G major scale



(b) G natural minor scale

図6 GメジャーとGマイナーの tonnetz 表現. 丸い音符は音階上の音符に相当する. 二つのモードは, 主音 (2重丸がついている G) を含む水平軸で対称となっている.

に比べると, その起源については, よくわかっていない. 長調の3和音 (major triad: 長3度の上に短3度を重ねたもの) は, 基本周波数と最初の6つの倍音 (harmonics) からなる. その倍音は, 基本周波数と同じオクターブに射影することができ, 長3度に属し, 音の音波エネルギーの大部分を持つ周波数である. このように, 音楽コードと倍音の関係が, 短調のコードの起源にもなるとされてきた[25]. 長調と同様に, 短調のコードの周波数は, 倍音のスペクトラム上に存在する. しかし, 短調コードのこれらのコンポーネントは, 通常, 音波エネルギーが小さいので, 聴き分けるのは, 困難である. したがって, 倍音だけで短調のコードの起源を説明するには, 少し無理がある.

Terhardt[24]は, 短調コードの起源に別の説明があることを示した. Terhardtは, 基音 (root position) にある短調のコードは (同様に長調のコードも), その基音の低音部に相当する仮想的なピッチを引き起こすということを示した. 低倍音 (subharmonic) の感覚的テンプレートという Terhardt のアイデアは, Simpson[21]に引き継がれ, 長調と短調のコードが, そのコードの一つ下のオクターブの基音に相当する神経細胞の活性パターンを引き起こすことがわかった.

数学モデルによって, 完全に短三和音の起源を説明することはできなくても, 短三和音を正確に記述することができ, その性質を明らかにすることは, 可能である. Mazzola[14]によると, 短調のコードとその音階は, 長調のコードと音階と対称になっている. 図2

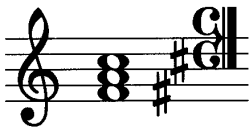


図7 Fメジャーの和音の対称変換。A₄を原点としたとき、Fメジャー (F A C) は、F[#]マイナー (F[#] A C[#]) に変換される。すなわち、元の長3度 (F A) は、(A C[#]) に逆変換され、短3度 (A C) は、(F[#] A) に変換される。

のように、長調と短調のコードは、tonnetz 内の上向きと下向きの三角形で表すことができ、図6は、tonnetz における完全に対称な音階を表している。

短調の音階と3和音の心理音楽的な説明の幾何学的表現形式をみてみよう。1/2の比率のストレッチ・チューニング (melodic および harmonic) を使うと、Terhardtの低倍音のパターンと対称な音階を視覚および聴覚的にらせん鍵盤で表すことができる。より正確には、melodicな音階の逆転は、視覚化および聴覚化が共に可能だが、低倍音のパターンは、聴覚化のみ可能である。図5(f)は、チューニング圧縮によるらせん鍵盤の反転を表す。

図7のように、例えば、A₄を反転の原点とすると、Fメジャーの和音は、F[#]という短調の和音に変換される。別の原点をとっても、別の基音になってしまうが、やはり短調へと変換される。このような対称性は、図6のように、音階の主音 (tonic) に依存する。チューニングの圧縮による変換では、その絶対的な音階の「距離」は、保存される。この例では、FとAの長3度 (M3) という距離は、(AとC[#]でも) 保存されているが、相対的な音階上の位置は、逆転する。この例では、M3が原点の下にあったのが、上に変化している。音階上での配置が逆転しているため、どんな長調のコードでも、基音の異なる短調のコードに変換される。また、その逆も可能である。

半分の圧縮率をもつチューニングを行うと、通常では、難しい楽譜における運指 (指の使い方) を簡単にすることができる。付け足したが、人間のほとんどが左利きだったら、キーボードのキーと音のアレンジが今と同じだったであろうかと著者らは、想像してしまう。

3. 音楽的協和 (consonance) のリアルタイム最大化

ここまで、純粋な音程の音階は、望ましいが、実現不可能であり、12-TETがデファクトのオクターブ

配分となっていることを明らかにした。12-TETは、commaを音階上に一様に分散させ、ピッチドリフトのない音楽転調 (musical modulation) を実現するひとつの最適解となっている。しかし、この可用性は、逆に協和を犠牲にしている。

12-TETが提供する可用性を、和音を壊さずにどのように実現するかは、音楽の歴史を通じてたくさんの研究者が取り組んできた課題である。一般的に、同音異名 (enharmonic) な音符に対してピッチを増やすこと [28] [2], 旋律内で音色のピッチを調整すること [5] [18] が研究されてきた。キーの種類を増やすことは、楽器の演奏性を損なうし、旋律を操作しても、いつでも楽器で完全に表現できるわけではない。これらの問題を解決するために、12-TETにおいて、同じ周波数帯に同時に存在する周波数コンポーネントをダイナミックに最小化する方法を我々は、提案している [34]。

周波数の近い2つの音色を重ね合わせると、相互干渉パターンが生じる。1に近い圧縮率を持つ (harmonic な) チューニングにおけるこの現象について考える。この問題は、音楽的協和と非協和 (dissonance) を理解する上で大変重要である。

周波数が近い2つの音は、人間の聴覚センスでは、独立に聞き分けることができない。重ね合わせた音の周波数がある一定のバンド幅によって分けることができれば、そのコンビネーションは、二つの異なる音として認識される。ひとつに認識されていた音が、2つの音に認識されるような遷移が起こるとき、音響心理学的に、ラフネス (roughness) と呼ばれる現象が生じる [7]。定量的には、ラフネスは、音楽的非協和と関連している [26]。純正律は、ラフネスを最小化することも知られている [16]。

Vassilakis [27] によると、周波数 f_i , 振幅 a_i の sin 波形のペア、 $\langle f_1, a_1 \rangle, \langle f_2, a_2 \rangle$ のラフネス r は、次のように計算できる。

$$r = \frac{(a_1 a_2)^{0.1}}{2} \left(\frac{2 \min(a_1, a_2)}{a_1 + a_2} \right)^{3.11} (e^{-3.5F} - e^{-5.75F}) \quad (3.1)$$

ここで、

$$F = S(\min(f_1, f_2)) |f_1 - f_2| \quad (3.2)$$

および

$$S(f) = \frac{0.24}{0.0207f + 18.96} \quad (3.3)$$

である。

Plomp & Levelt [16] と Sethares [17] のアプローチ

に従い、Vassilakis は、複雑な音色のラフネス R をそれぞれの周波数ペアのラフネスの集積と考え、

$$R = \sum_{h=1}^n \sum_{i=h+1}^n \sum_{j=1}^p \sum_{k=1}^p r(\langle a_{hj}, f_{hj} \rangle, \langle a_{ik}, f_{ik} \rangle) \quad (3.4)$$

と表現した。ここで、 n は、オーディオストリームの数、 p は、計算に寄与すると考えられる倍音の数を表す。 x_{uv} は、 u 番目のオーディオストリーム内の v 番目の振幅 (a) または周波数 (f) を表し、 r は式 3.1 で定義されたラフネスを表す。

3.0.1 Vassilakis モデルの改良

ここでは、異なる音源から発生する音色間の非協和現象をモデル化するために、内的ラフネスと外的ラフネスの2つを区別する。内的ラフネスは、単一音源内で生じ、その音質 (timber) に強く関係している。それに対して、外的ラフネスは、異なる音源間から生じるサウンドのコンビネーションによって生じる。

式 3.5 で、外的ラフネス R_E は、計算できる。各項で、異なる音源のみ (h と i) をとりあげて、その中の倍音についてラフネスを計算し、累積している。

$$R_E = \sum_{h=1}^{n-1} \sum_{i=h+1}^n \sum_{j=1}^p \sum_{k=1}^p r(\langle a_{hj}, f_{hj} \rangle, \langle a_{ik}, f_{ik} \rangle) \quad (3.5)$$

3.1 近傍

ラフネスを最小化するような別のイントネーションを探す場合に、その探索範囲を限定することが必要となる。限定しなければ、すべての音色は、もとのピッチより数オクターブ高い単一の音色 (ユニゾン: unison) に最適化されてしまうであろう。探索範囲を限定するには、所与のコードの「性質」をできるだけ保存したいという理由もある。短調の3和音は、対応する長調のバージョンに比べ、通常ラフネスが高いが、これらのコードは、聞き覚えが高いので、その認識上の特性量が変わらない場合以外は、コードの変更はしない方がよい。

人間は、その聴覚可能なスペクトラム上で最も能力の高い領域では、8.3 cents のピッチ差を聴き分けることができる。12-TET と純粋音階の最大誤差は、16 cents であり、これは、短3度とその逆の長6度において生じる。ある音色の ± 8 cents 近傍の音色は、認識上の変化を伴わずに、12-TET 音階を「浄化」できる可能性がある。

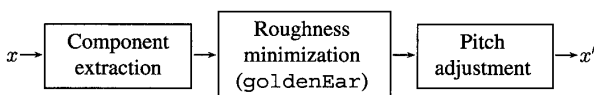


図8 機能ブロックのダイアグラム

3.2 プロトタイプの概要

楽器がもつ物理的な制約とは別に、異なるチューニングシステム上で、イントネーションの違いをうまく補正することは難しい。Kopiez[11]によると、プロの音楽家できえ、異なるチューニングシステム上で、同じイントネーションで演奏してしまっている。

筆者らは、Vassilakis のモデルに基づき、音楽家が演奏中により高い協和レベルを得るようにサポートするコンピュータプログラムを開発した。GoldenEar と名付けられたリアルタイムのオーディオフィルタリングシステムのプロトタイプは、概念的には、図8のように3つのステージに分けられる。第2ステージは、本研究専用開発した。GoldenEar のソースコードとそのライブラリーは、Mac OS と MS Windows で使用でき、ドキュメントとサンプルは、文献[29]より取得できる。

3.3 要素抽出

ピッチ検出と要素抽出のプロセスでは、ピッチとそれを構成する周波数と振幅が記録される。新しい音が認識されると解析が行われる。サンプリング周波数 $f_s = 48$ kHz と $n = 2048$ サンプル (512 ごとにオーバーラップ) のサンプルウィンドウを使い、 $f > 2f_s/n$ に相当するピッチが検出され、解析される。これは、 G_2 よりも少し高い周波数に相当する。

3.4 ラフネスの最小化 (GoldenEar)

GoldenEar は、オリジナル音源のピッチからの cent を調整し、協和を最大にするチューニングを見つけることができる。振幅感度も高いので、生演奏や質の悪いレコーディングなどのバックグラウンド雑音にも対応できる。GoldenEar は、3つまでのパラメータを設定することができる。最初のパラメータは、音源の数、2番目は、近傍のサイズ (8 は、 ± 8 cent の大きさの近傍を表す)、3番目は、使われる要素の数を指定する。指定しない場合には、デフォルト値 (1音源、サイズ ± 8 cent、20 要素) が解析に使われる。

オリジナルの周波数から、1 cent ごとに近傍の端まで最小のラフネスを探索する。GoldenEar の探索は、直近の音について行われるので、新しい音の到着順序が重要である。

このアプリケーションの処理時間 (約 12 ms) は、たいていのリアルタイムシナリオには有効である。しかし、まだ改良の余地はある。例えば、ラフネスがある一定値より小さいときには、計算手順を省略して、より早く収束値を探索することができる。最適化ルー

チン自体にも改良の余地がある。ここでは、近傍の可能な周波数を、1 cent 刻みに、しらみつぶしに探索している。代わりに、Simulated annealing や Simultaneous Perturbation Stochastic Approximation (SPSA) [22] などのヒューリスティックな手法も使うことができる。SPSA は、Adaptun の Sethares [18] がその使用について検討をしている。我々も初期プロトタイプで、SPSA を使った。SPSA は、各反復で、摂動範囲が少しずつ小さくなるという点で、Simulated annealing と似ているが、Simulated annealing が大域的な最小値を探索するのに対して、SPSA は、局所的な最小値を探索する点異なる。SPSA は、各反復で、二つの関数だけで評価することができ、ノイズの大きな場合にも適用でき、gradient の情報も不要である。いずれにせよ、たいいていの場合には、近傍は、十分小さく、最小化手法の違いが、演奏に大きな影響を与えることは少ない。ラフネスの計算を両端、センターの3回行うことで、演奏の性能をあげることができる。3つのポイントは、二つの線分で結ぶことができ、これを使って、再イントネーションの周波数を少ない誤差で評価できる。その他にも、最小化アルゴリズムを並列処理することでも、同じような効果が得られる。現在、我々は General-Purpose Computation on Graphics Hardware (GPCGH) を使って、インプリメンテーションを評価している。

3.4.1 既存のサウンドへの再チューニング

到着したピッチが近傍のサイズよりも大きい場合に、ピッチ探索のメカニズムの最小探索が開始される。すなわち、

$$\frac{\max(p_{new}, p_{last})}{\min(p_{new}, p_{last})} > v \quad (3.6)$$

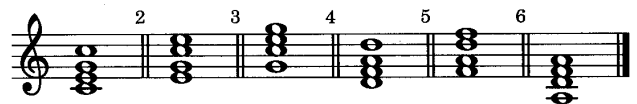
の場合である。ここで、 p_{new} は、新規に到着したピッチで、 p_{last} は、オーディオストリームの最後に記録されたピッチであり、 v は、近傍のサイズを表す。

3.5 ピッチ補正

ピッチの補正には、オリジナル音源にシングル側波帯変調 (single-sideband modulation: SSB) が用いられる。SSB は、2つの側波帯を消去する AM の改良方式である。この方法は、計算時間が短く、ほとんどの場合、その変調はごく小さいので、その違和感は、人間には認識ができない。

3.5.1 結果

異なる音楽コードに、我々の方法を適用した結果を図9に示す。音源には、参照しやすいように(合唱パ



(a) Isolated chords for testing the prototype. Each bar was played independently.

	1	2	3	4	5	6
S	0	0	0	0	0	0
A	2	8	-8	2	-8	8
T	-8	8	-2	8	-7	-2
B	0	0	0	0	0	0

(b) Deviation in cents for the isolated chords.

図9 再チューニングの音楽コード例

ート形式で) B , T , A , S とラベルをつけた。第1節と第3節は、純正律に近い。

図9(b)が、到着順序の重要性を示している。コードは、 B から順に生成される。 T が到着すると、その近傍で一番良いイントネーションへと再チューニングする。第2節では、 T が純短3度 ($m3$) へ近づいたために、近傍内の最高値へ再チューニングしている。第5節と第6節は、このプロトタイプの欠点を表している。コード内の音符は、音程順に送られ、低い音から高い音にソートされるが、第5節では、 A が T よりも早く再チューニングされている。各音は、順番に処理されていくが、これはあくまでもロジック上の仮想的なもので、実際には、この順番に音が本当に発生しているわけではない。このため、ピッチの検出メカニズムは、 A ピッチの方が早いと判定してしまう。

ライブ演奏でこのツールを使う場合には、一般の聴衆に届くオーディオと演出家に(モニターするために)フィードバックされるオーディオを分離する必要がある。演出家向けのオーディオには、ピッチ変更処理を行わない生の音のコピーを届け、処理済みのオーディオとミックスされないように注意する必要がある。このためには、例えば、アコースティック分離型ヘッドホンや指向性の強いマイクが必要になる。しかし、このソフトウェアの最適な応用先は、サウンドスタジオでのプロダクションとポストプロダクション処理で

あろう。

このシステムは、調和スペクトラムで純正律に近く、ラフネスの少ないコードを、近傍の範囲内で生成する。既存の再イントネーションシステムに似ているが、このシステムでは、チューニングに必要な情報を前もってに設定する必要がなく、近傍メカニズムが自動的にピッチドリフトのコントロールを行う。

4. 結論と今後の計画

現在、筆者らは、らせん鍵盤とその拡張が、どのように音楽構造の理解を助け、どのようにアーティストの演奏を高めるかということに興味がある。また、筆者らの研究室では、プラットフォームの可用性を高めるために、らせん鍵盤に触覚インターフェースを統合することも計画している。

GoldenEar は、希望した値（純正律）とそれほど遠くない値に、12-TET のコード進行に再チューニングできた。非ハーモニー音楽とハーモニー音楽との混成、例えば、ベルや鉄琴（メタロホン）と既存のシンフォニー楽器の混成などに対しても、このシステムを使用することを計画している。

また、リアルタイムで与えられた音の重要性の情報がないため、我々のアルゴリズムがうまく動かない場合（例えば、他の音色をチューニングするために主音が必要なとき）があることがわかった。通常、このような情報は、演出家が、楽譜から読み取ることによって得られる。しかし、このソフトウェアのリアルタイムでの即応性を考えると、再チューニングする前に、コードを構成しておくことは、不可能である。最適化処理による演奏と並行またはその直前に音色の特性を抽出するシステムも考えられる。GoldenEar は、オリジナル演奏のラフネスとその最適解のラフネスを比較しているので、GoldenEar からの出力は、少なくとも生演奏並であるといえる。

（訳：豊泉 洋 早稲田大学）

参考文献

- [1] K. Amano, F. Matsushita, H. Yanagawa, M. Cohen, J. Herder, Y. Koba and M. Tohyama. The Pioneer Sound Field Control System at the University of Aizu Multimedia Center. In *5th IEEE Int. Workshop on Robot and Human Communication*, pp. 495-499, Tsukuba, Japan, 1996.
- [2] W. Carlos. Tuning: At the crossroads. *Computer Music Journal*, 11 (1): 29-43, 1987.
- [3] E. Chew and Y.-C. Chen. Mapping MIDI to the spiral array: Disambiguating pitch spelling. In *Computational Modeling and Problem Solving in the Networked World-Proc. of the 8th INFORMS Computer Soc. Conf.*, pp. 259-275. Computational Modeling and Problem Solving in the Networked World, Kluwer, 2003.
- [4] E. Chew. Math & Music—The Perfect Match. *OR/MS Today*, pp. 26-31, 2008.
- [5] J. deLaubenfels. Studio j: John deLaubenfels' basement project, 2002. [Online; accessed 9-September-2008]: <http://personalpages.bellsouth.net/j/d/jdelaub/jstudio.htm>.
- [6] L. Euler. *Tentamen novae theoriae musicae ex certisismis harmoniae principiis dilucide expositae*. Saint Petersburg Academy, 1739. Online; Available from www.math.dartmouth.edu/~euler/pages/E033.html.
- [7] H. Fastl and E. Zwicker. *Psychoacoustics: Facts and Models*. Springer series in information sciences. Springer, Berlin, 3rd edition, 2007.
- [8] W. Heisenberg. *Physics and Philosophy*. Penguin classics, 2000. ISBN 0-14-118215-6.
- [9] J. Herder and M. Cohen. The Helical Keyboard: Perspectives for Spatial Auditory Displays and Visual Music. *J. of New Music Research*, 31 (3): 269-281, 2002.
- [10] A. J. M. Houtsma, T.D. Rossing and W.M. Wagenaars. Auditory demonstrations, 1987. Philips compact disc No. 1126-061.
- [11] R. Kopiez. Intonation of Harmonic Intervals: Adaptability of Expert Musicians to Equal Temperament and Just Intonation. *Music Perception*, 20: 383-410, 2003.
- [12] M. D. Lemonick. Time Magazine: The Geometry of Music, Jan 2007. [Online; accessed 20-February-2009]: <http://www.time.com/time/magazine/article/0,9171,1582330,00.html>.
- [13] M. V. Mathews and J.R. Pierce. Harmony and Nonharmonic Partial. *J. Acoust. Soc. America*, 68: 1252-1257, 1980. ISBN 3-540-76173-X.
- [14] G. Mazzola, S. Göller and S. Müller. *The topos of music: geometric logic of concepts, theory, and performance*. Birkhauser Verlag, Boston, MA, 2002.
- [15] J. Pierce. *Music, Cognition, and Computerized Sound*, chapter Consonance and Scales. MIT Press, Cambridge, Massachusetts, 2001. ISBN 0-262-53190-9.
- [16] R. Plomp and W. Levelt. Tonal consonance and

- critical bandwidth. *J. Acoust. Soc. America*, 38(4): 548-560, 1965.
- [17] W. Sethares. Adaptive tuning for musical scales. *J. Acoust. Soc. America*, 96(1): 10-18, 1994.
- [18] W. Sethares. Real-time adaptive tunings using Max. *J. New Music Research* Vol. 31, No. 1, pp.1-7, 2002.
- [19] W. A. Sethares. *Tuning, timbre, spectrum, scale*. Springer, London, 2nd edition, 2005.
- [20] R. Shepard. *The Psychology of Music*, chapter Structural Representation of Musical Pitch. Academic-Press, New York, NY, USA, 1982.
- [21] J. Simpson. Cochlear modeling of sensory dissonance and chord roots. Master's thesis, University of Waterloo, 1994.
- [22] J. C. Spall. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Transactions on Automatic Control*, 37: 332-341, 1992.
- [23] R. A. Steenblik. Chromastereoscopy. In D. F. McAllister, editor, *Stereo Computer Graphics and Other True 3D Technologies*, pp. 183-195. Princeton University Press, 1993. ISBN 0-691-08741-5.
- [24] E. Terhardt. Pitch, consonance and harmony. *J. Acoust. Soc. America*, 55(5): 1061-1069, 1974.
- [25] H. v. Helmholtz. *On the Sensations of Tone as a Physiological Basis for the Theory of Music*. Dover Publications, 31 East St., Mineola, N. Y. 11501, 1954.
- [26] H. v. Helmholtz. *On the Sensations of Tone as a Physiological Basis for the Theory of Music*, chapter VIII, On the beats of simple tones. Dover Publications, II english edition, 1954.
- [27] P. N. Vassilakis. *Perceptual and Physical Properties of Amplitude Fluctuation and their Musical Significance*. PhD thesis, University of California—Los Angeles, 2001.
- [28] N. Vicentino. L'antica musica ridotta alla moderna prattica, 1555. [Online ; accessed 9-September-2008] : <http://visualiseur.bnf.fr/CadresFenetre?O=NUMM-58223&M=telecharger&Y=Image>.
- [29] J. Villegas. Julián Villegas Website. [Software]. Retrieved July 9, 2008. Available from www.julovi.net, 2008.
- [30] J. Villegas and M. Cohen. Melodic stretching with the helical keyboard. In *Proc. 2nd Int. Conf. on Enactive Interfaces 2005*, Genoa, Italy, November 2005. IST Network of Excellence ENACTIVE.
- [31] J. Villegas, Y. Kawano and M. Cohen. Harmonic stretching with the helical keyboard. *J. of Three-Dimensional Images*, 20(1): 29-34, 2006.
- [32] J. Wild and P. Schubert. Historically informed retuning of polyphonic vocal performance. *J. of Interdisciplinary Music Studies*, 2(1-2): 121-139, 2008.
- [33] Y. Yokomatsu. Primassa: Polyphonic Spatial Audio System with Matrix Mixer and Speaker Array Integrated with CVE. Master's thesis, University of Aizu, 2007.
- [34] E. Zwicker, G. Flottorp and S. Stevens. Critical bandwidth in loudness summation. *J. Acoust. Soc. America*, 29(5): 548-557, 1957.