

## かな書道画像からの音楽生成の試み Music generation from calligraphy of kana style

梅村 祥之<sup>1)</sup>  
Yoshiyuki Umemura

### 1 はじめに

音楽の自動生成の研究が古くから行われている [1]. その中に、フラクタル図形の形に音符情報をマッピングして音楽生成する研究がある [2]. 我々の研究室でも、自然界に存在する様々な美しいパターンや数学的に生成した美しいパターンを基に音楽生成する次のような研究を行ってきた. 山岳の標高データから、山のシルエット像を旋律の形に見立ててメロディ生成する研究 [3][4], 樹木の CG に基づいて、枝の先端にある葉の位置を順番にたどりながらメロディライン生成する研究 [5], 樹木が生長する様子の CG 動画の動きに連動したリズムで音楽を生成する研究 [6], スピログラフという回転対称の幾何学図形の形状からメロディを生成する研究 [7], 二重振り子の複雑な運動に基づいてメロディラインを生成する研究 [8] を行ってきた.

この研究の流れの中で、本研究では芸術作品である書を対象とした音楽生成を行う. かなを主体に続け字で書かれた書作品は、流れるような筆使いからリズム感を感じることができ、音楽との親和性が良いと思われる. さて、書道は日本の古典芸術である. 平安時代、かなが出現し、漢字と融合した和様書道が完成した. 書家が制作した作品が書道展で鑑賞されるとともに、掛軸、屏風、額などの室内装飾としても用いられる. 一方、書の作成過程を見て楽しむ席上揮毫がある. しかし、静かな環境で行うのが一般であり、BGM が演奏されることはない. また、高校などのクラブ活動で書道パフォーマンスが行われ、BGM として既存のポップス曲が使われることが多い. それらに音楽情報処理が活用されてはいない.

本研究の狙いは、書とテクノロジーを融合した新たなメディアアートの創出を通じて、書の鑑賞をより楽しくすることである. 利用場面として、室内インテリアとしての壁掛けディスプレイでの展示やホテルのロビーにおけるデジタルサイネージでの展示なども考えられる.

## 2 書画像から運筆の推定

### 2.1 書画像の取得

本研究で音楽生成の対象として用いる素材として、和歌の書を選択する. その中でもかなを主体として漢字も含み、草書体の続け文字として書かれた書を選択する. 続け文字で表現されたかな文字は流れるような曲線美を有し、2次元画像でありながら、動きやリズムを感じるができると考えられ、本研究の要求を満たす素材と考えられるからである.

今回、慶應大学に蓄積された文化財を公開する Web サイト Keio Object Hub [9] から公開されている和歌を収録した江戸後期の写本である歌合集 [10] の 4 頁に掲載された 2 首の和歌を処理対象の原画像とする. 図 1 に画像を示す.

以下、書画像から運筆データを抽出するにあたって、次の 3 ステップを行う. 第 1 ステップはカラー画像であ

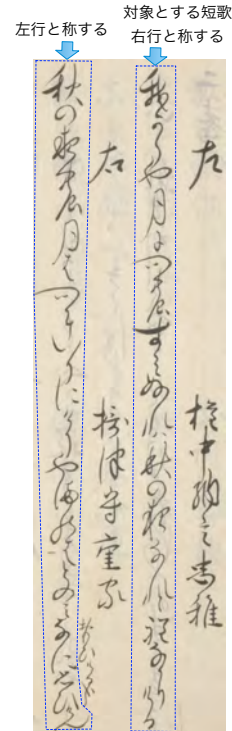


図 1 音楽生成の対象とする書画像. 江戸後期の写本歌合集収録の和歌.

る原画像から 2 値画像に変換する処理である. 第 2 ステップは 2 値画像を細線化して線画像にする処理である. 第 3 ステップは線画像から、筆順と筆の速度の情報を備えたデータに変換する処理である.

### 2.2 運筆データ生成の前処理

本節では第 2 ステップを説明する. 2 値画像処理の分野で、古くから、細線化、ないしスケルトン抽出という処理があり、様々なツールが提供されている. 本研究では R Image Thinning Library [11] を用いる. このツールを用いると書画像に対してある程度うまく動作するものの、太い線が二重線になったり、縦線と縦線が狭い角度で交差する場合に誤検出する場合がある. 本研究は音楽情報処理に関する研究であり、そこに至る前処理である画像処理に重点を置かないため、画像の誤検出を手動で修正する. 線データの形式として 4 連結と 8 連結がある中で、ここでは後の処理が容易となる 8 連結を用いる. 手修正を含めた細線化の結果を単独の結果としては図示せず、データの構造化の説明と合わせて図 2 の中で示す. 同図下に、階層図の形式で、細線化で得られた線が複数集まってグループ化されること、1 本の線が細分化されることとの関係を示す. 本論文では、連続した線をセグメントと称する. 複数のセグメントが集まって 1 文字となる場合、そのまとまりをグループと称する. 続け字のため、続いた複数文字で 1 グループとなる場合もある. 実験者が手作業でグループ化を行う.

1) 広島工業大学 Hiroshima Institute of Technology

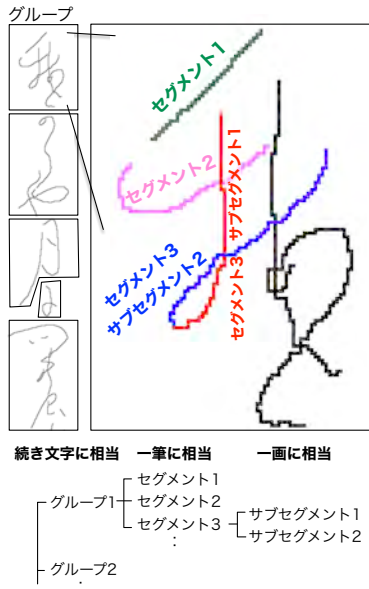


図2 細線化の結果および線の構造化

グループ化の逆で1セグメントが複数の部分に細分化されるとき、それぞれをサブセグメントと称する。同図ではセグメント3が2個のサブセグメントに分割されている。この分割も実験者の主観で行う。以上の方法で、今回の対象データに対して、セグメントのグループ化と、サブセグメントへの細分化を行った結果を同図に示す。

### 2.3 運筆速度の推定

本節では第3ステップである筆順と運筆速度の推定について説明する。筆順に関しては、実験者が手作業でデータをインプットし、運筆速度に関しては、次の簡単なアルゴリズムを用いて推定する。

図3を参照しながら推定アルゴリズムを説明する。図は文字「あ」を用いた説明用イラストである。速度が速い箇所の代表として、赤色の縦線を取り上げ、速度が遅い箇所の代表として、茶色の湾曲した線を取り上げる。図のように、直線に近い場合には速度を増加させ、蛇行した場合には速度を減少させる。具体的にはある点の速度を推定するのに、前後、何点かの定まった点数での縦幅と横幅のうち、大きい方の幅に係数をかけて速度とする。同図の場合、赤色の線、茶色の線、双方とも、縦幅 $l_1$ が採用される。なお、セグメントの始点と終点は速度を0とする。以上のアルゴリズムで求めた結果をストロボ撮影のような表示で可視化して図4に示すとともに、後述するYouTubeビデオでも動きを確認できる。

## 3 運筆から音楽情報への変換

### 3.1 音価と音高の仮決め

図5に対象画像の最初の1文字の最初の部分を例として音高決定のアルゴリズムを説明する。サブセグメントごとに1音符を割り当てる。サブセグメントは1文字の中の1画分であり、1画ごとに音符が奏でられることになる。音価は、基本的には、筆が1画分移動する時間によって決定される。その後、小節区切りと同期させる処理によって、オンセット時刻ないしオフセット時刻が変更される場合と和音との関係でオフセット時刻が引き延ばされる場合があり、それらの詳細は後述する。なお、

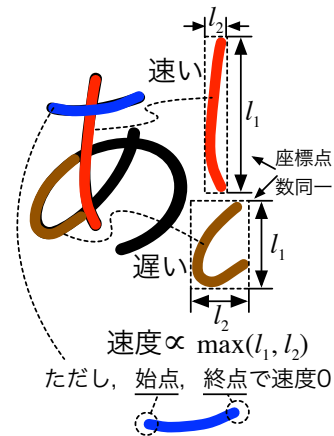


図3 速度推定法を説明するイラスト図

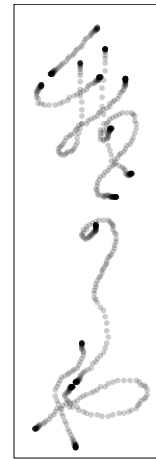


図4 右行の冒頭部分に関する速度推定結果

図は小節区切りによる同期処理は行われており、和音による伸長処理は行われていない。

音高を求める際にグループの情報を用いる。グループは1文字ないしは続き文字のまとまりに相当する単位である。グループの中心座標を画素の重心として求める。通常は描画範囲の中心とほぼ等しい。図に中抜き白色の点でグループの中心位置を示す。図に示すように、運筆順に従って中心から筆の位置まで赤色の直線を描画している。水色と黄緑色のスポットライトは筆の位置を中心としたスポットライトである。サブセグメント、すなわち音符ごとに色を交互に変えて、音符の切り替わりを知覚しやすくしている。音高を中心から筆までの長さに応じて決定する。長さが長いほど、すなわちグループの辺縁をたどるほど音高を低くし、長さが短いほど、すなわち中心をたどるほど音高を高くする。図に示す7音符の音高がこの対応関係で定まることが見て取れる。長さを音高に対応させる際、長さを線形変換して音階単位の音高に変換する。音階は全音階を使用する。線形変換において切片と傾きに関するパラメータが存在する。これらを変更すると、音域が上昇したり下降したりし、また、音域の幅が広がりたり狭まったりする。本論文の音楽生成法において、2行を同時にたどることによって2重奏を奏でる。両者の音域を離れさせることによって、それぞれの旋律を聴き取りやすくする。

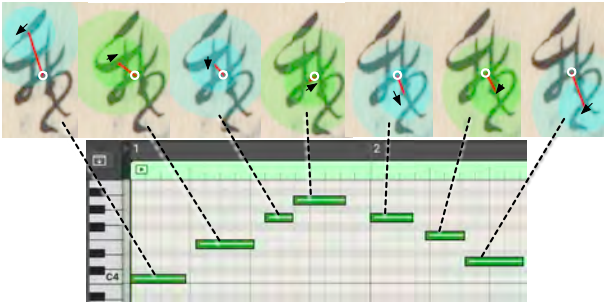


図 5 グループ中心から筆までの距離に基づく音高決定方法

### 3.2 小節との同期

元々、運筆における 1 画の始めと終わりのタイミングは小節区切りと同期していない。そのようなバラバラのタイミングで音符を生成すると音楽性が損なわれる。そのため、小節線近くにオンセットやオフセットが来る音符のオンセット、オフセット時刻を移動させることにより、小節に同期させる。具体的なアルゴリズムに関しては、先に開発した二重振り子による音楽生成で用いた手法と同様の処理であり、文献 [8] で詳細を説明しているため、説明を省略する。

### 3.3 和音に伴う処理

和音を使った伴奏を付加する。和音は半小節あたり 1 和音を割り当てることとする。和音系列の自動生成は行わず、制作者が与えるものとする。いくつもある著名な和音進行から選択して与えれば良い。生成例で使用する和音進行は J-POP でよく使われる小室進行 (Am → F → G → C) である。演奏法にはアルペジオ、バッキングなど様々あり、適切に選択することによって、クラシック調、POPS 調、バラード調と様々な曲調を生じさせることができる。

主旋律、副旋律の音高が和音の構成音の音高に対して不協和音程になると濁った響きが生じる。古典派までの音楽理論において、それらを非和声音と言ひ、経過音などの補助的な役割に用いることが推奨されていた。ロマン派以降に進むにつれて、音楽の幅を広げるため、不協和音程が積極的に使われるようになってきた。協和音主体の旋律は違和感ない感じを与える反面、刺激のない平凡な感じを与える。本研究では、協和音程に主体を置きながら、書に連動した音楽という新たなジャンルであるため、所々、不協和音程を盛り込んで緊張感を高めるようにする。

図 6 を参照しながらアルゴリズムを説明する。図の上段「変更前」を参照されたい。対象とする書の右行に対する旋律の冒頭 2 小節のピアノロールを示す。半小節ごとに使用する和音を記号 Am 等で示す。ピンクの帯をつけた音高は半小節ごとの和音の構成音の音高を表す。帯に重なる音高は協和音程となり、重ならない音符は 2 度の不協和音程となる。ピアノロールに登場する 9 音符のうち、6 音符が協和音程、3 音符が不協和音程となる。

協和音と不協和音が混在する中で協和音を強調させるため、協和音であれば、オフセット時刻を小節の最後まで延長させる。処理結果を図の下段「変更後」に示す。6 音符が延長される。ここで、第 7 音符は音符の前半が協和音程、後半が不協和音程である。音符の継続時間中、不協和音程となる時間の方が長いので、不協和音と



図 6 和音に伴う処理。ピンクで示す和音構成音と重なる音符を小節の終わりまで延長する。

認定する。

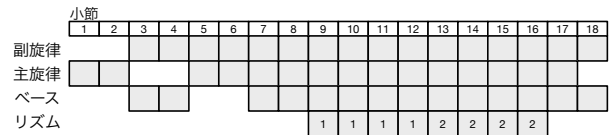


図 7 パートの編成

## 4 生成例

これまでの説明に用いた書画像からの映像を伴う音楽生成を行う。本研究では、生成した音楽データを MIDI ファイルに出力し、Apple 社製の音楽編集ソフト Logic Pro 10.8 で読み込んで楽譜の作成とオーディオファイルの生成を行う。このソフトに備わった楽器で各パートを演奏する。まず、これまで説明しなかったパートの編成を述べる。全体を図 7 に示す。第 1 パートは副旋律で、高音域を Chinese Di Zi Flute で演奏する。第 2 パートは主旋律で、低音域を Trumpet で演奏する。第 3 パートは和音の根音を Cello で演奏する。第 4 パートはリズム伴奏で、Timpani で演奏する。生成曲の楽譜を図 8 に示す。

曲の始めの 2 小節の間、主旋律をソロで演奏する。次の第 3-4 小節は主旋律の演奏を停止し、副旋律と根音を演奏する。第 5-6 小節は、根音を停止し、主旋律と副旋律の二重奏とする。第 7-8 小節は二重奏に根音を加える。第 9-16 小節にかけて、さらにリズム伴奏を加える。前半 4 小節はゆったりとしたリズムパターンを用い、後半 4 小節は小刻みなリズムパターンを用いる。第 17-18 小節はリズム伴奏を停止し、最後の第 18 小節は主旋律も終了し、副旋律と根音の演奏となる。このように、単純なパターンから徐々にパートを増やしてゆき、エンディングで静かに締めくくる。狙いは単純な音から開始することによって、筆の動きに関する映像と音の対応関係を知覚しやすくすることと、ソロ演奏から徐々にパートを増やして盛り上げ、最後に、静かなエンディングにして緩急をつけることである。生成曲は曲単独ではなく、書の運筆映像とともに楽しむものである。映像中の静止画 4 シーンを先の楽譜に添えている。

生成した作品は YouTube の次の URL で公開している。

<https://youtu.be/6vJWVQcRKog>

## 5 音楽特徴の分析

世界の民謡を集めた The Essen folksong collection(以下, Essen コーパスと記載する) [12] の中に収録されているヨーロッパ曲 6202 曲を用いて, 音楽としての性質を分析する。

### 5.1 ピッチクラスの頻度分布

全音階を構成する第 1 音から第 7 音には主音, 属音等の役割があり, それらの出現頻度分布が音楽の特徴量と考えられることから, Essen コーパスと本生成曲で出現頻度分布を比較する。ところで, 長調の曲と短調の曲ではピッチクラス分布の形状が多少異なる。文献 [13] において, Essen コーパスの長調と短調に関する分布がそれぞれ Figure 9.1, Figure 9.2 に示されている。Essen コーパス中の長調 5440 曲, 短調 761 曲について, 主音を C に移調して集計した結果をそれぞれ図 9(a),(b) に示す。本生成法で生成した 1 曲はハ長調として生成しているため, 移調を行わない。集計した頻度分布を同図 (c) に示す。Essen コーパスの長調とも単調とも異なり, 下属音, 導音の多い独特の分布である。

### 5.2 水平音程の頻度分布

水平音程は隣接音符間の音程であり, 同音進行, 順次進行, 跳躍進行の 3 種類に分類される。一般に, 順次進行は旋律を滑らかにする作用があるため旋律中に多く出現する。一方, 跳躍進行は順次進行の合間に時々挿入され, 躍動感につながる。このように水平音程の頻度分布は旋律の性質を表す重要な指標と考えられる。分析にあたり, 文献 [14] を参考に, 音程に対する数値の割り当て方として, 同音進行を 0, 半音下降を -1, 半音上昇を 1 のように割り当てる。結果を図 10 に示す。本生成曲は Essen コーパスに比べて跳躍進行が多いものの, ある程度類似している。

### 5.3 音高輪郭

本論文のアルゴリズムで音高を定めると, 筆位置はグループの周辺である左上から書き始められ, 途中で中心を通り, 最後に右下で書き終わることが多いため, 低い音高から始まり高い音高を経て低い音高に戻る山形の旋律を描く。先の図 5 はそのようなカーブになっていた。主旋律に該当する対象画像の右行全体について図 11 に示す。11 個のグループのうちの多くにそのような傾向が見られる。これは音高輪郭に関する一般的な性質とも合致する。実際, Essen コーパスを使って音高輪郭の平均の形状が求められ, 山型になることが調べられている [13]。文献と同様の処理を行って求めた平均音高輪郭を先の図 11 に添える。

## 6 まとめ

かなを主体に続け字で書かれた書作品から運筆を推定して, 筆の動きを映像化し, 動きに合わせた多旋律音楽を生成して映像音楽作品を生成する技術を開発した。提案した旋律生成アルゴリズムによって生成される旋律が一般の音楽の特性に類似する点を多く含み, 違和感のない音楽として感じられるとともに, クラシックやポップスとは異なる和風の雰囲気も備えるものとした。現在, 生成法の基本技術を開発した段階のため, 今後, 生成曲の多様性を増すための技術開発を行い, 多くの作品を制作して, 作品の質に関する主観評価を行う予定である。

## 参考文献

- [1] 松原正樹, 深山寛, 奥村健太, 寺村佳子, 大村英史, 橋田光代, 北原鉄朗, “創作過程の分類に基づく自動音楽生成研究のサーベイ”, コンピュータソフトウェア, Vol. 30, No. 1, pp. 101-118(2013).
- [2] Nierhaus, G., “Chaos and Self-Similarity”, In “Algorithmic Composition”, Springer(2009).
- [3] 梅村祥之, 伊達彩斗, “地図標高データを用いたメロディ生成の試み”, 情報処理学会, 第 115 回音楽情報科学研究会, Vol.2017-MUS-115, No.39(2017).
- [4] 梅村祥之, 富士直斗, “地図標高データを用いたメロディ生成法”, 情報処理学会, 第 123 回音楽情報科学研究会, Vol.2019-MUS-123, No.14(2019).
- [5] 岡田康汰, 梅村祥之, “樹木の生長と連動した自動作曲”, 広島工業大学紀要, 研究編, Vol.57, pp.7-13(2023).
- [6] 久下大介, 梅村祥之, “樹木の葉をたどりながら奏でる自動作曲システム”, 広島工業大学紀要, 研究編, Vol.57, pp.15-22(2023).
- [7] 梅村祥之, 富士直斗, “スピログラフ図形に基づく音楽生成”, 情報処理学会, 第 139 回音楽情報科学研究会, Vol.2024-MUS-139, No.7(2024).
- [8] 梅村祥之, “二重振り子の動きを利用した音楽生成の試み”, 情報処理学会, 第 139 回音楽情報科学研究会, Vol.2024-MUS-139, No.8(2024).
- [9] 慶應義塾ミュージアム, “Keio Object Hub”, 入手先 <https://objecthub.keio.ac.jp/ja/guide>, (参照 2024-6-11).
- [10] 小沢蘆庵, “歌合集”, 入手先 <https://objecthub.keio.ac.jp/ja/object/10191>, (参照 2024-6-11).
- [11] Jake Drew, Dr. Michael Hahsler, “R Image Thinning Library”, 入手先 <http://opensource.org/licenses/gpl-3.0.html>, (参照 2024-6-12).
- [12] Center for Computer Assisted Research in the Humanities: Themefinder, <http://essen.themefinder.org>, (参照 2024-6-13).
- [13] Huron, D., Chapter 5 Statistical Properties of Music, 5 Melodic Arch. In “Sweet anticipation”, pp. 85-88, The MIT Press(2006).
- [14] Temperley, D., “Music and Probability”, The MIT Press(2007).

♩ = 30

Chinese Di Zi Flute

Trumpet

Cello

Timpani

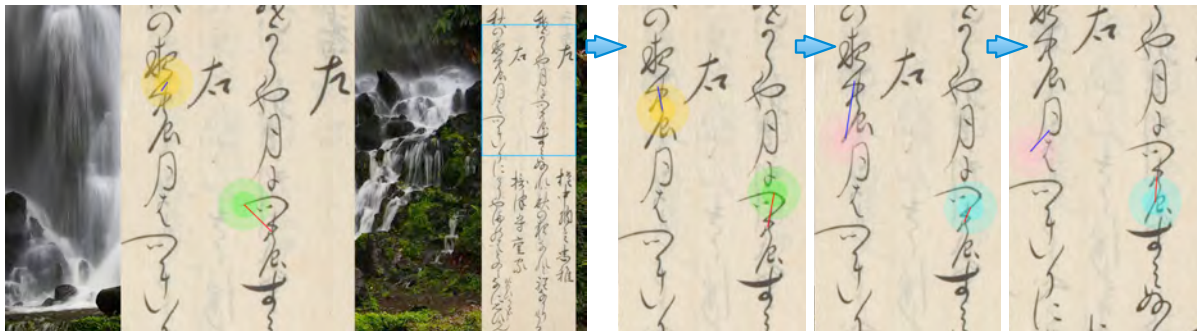


図 8 生成曲の楽譜 <https://youtu.be/6vJWVQcRkKog>

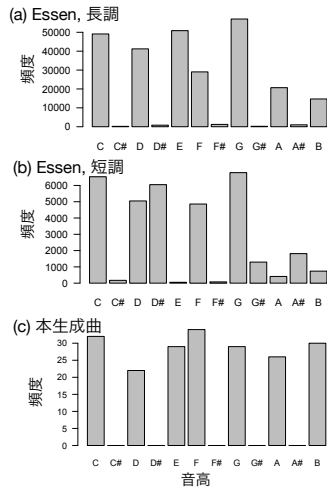


図9 ピッチクラスの頻度分布

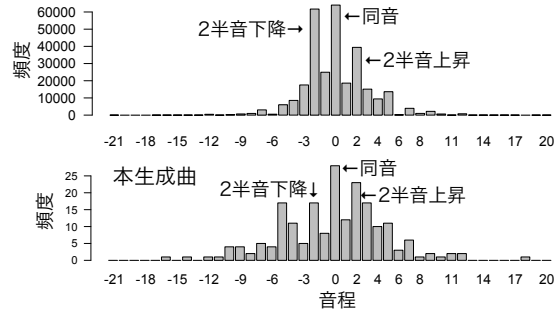


図10 水平音程の分布に関する Essen コーパスと本生成曲の比較

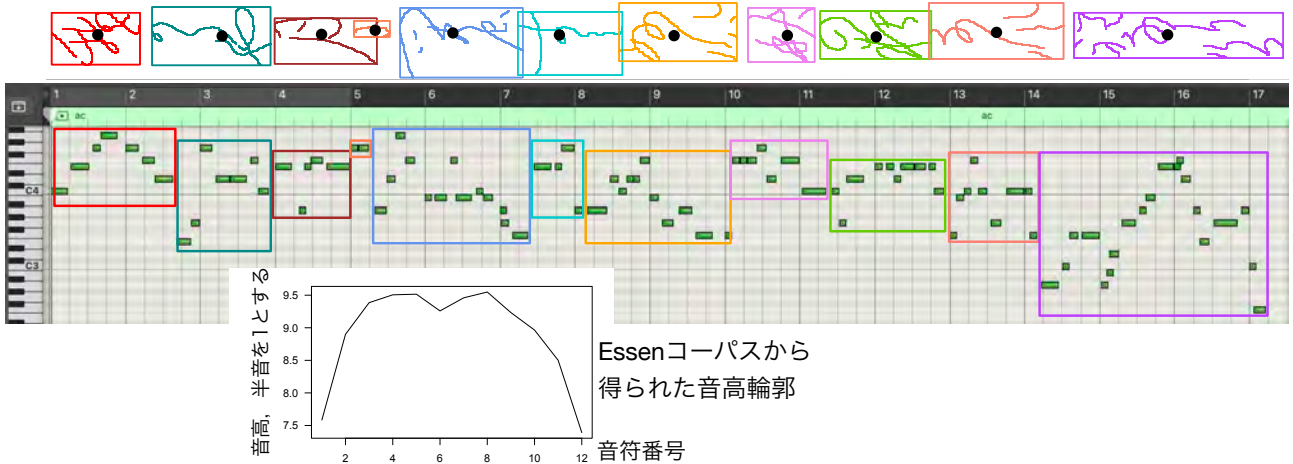


図11 主旋律に該当する右行のグループごとの音高輪郭