

抽象絵画における鑑賞者の視線行動を利用した 音楽作曲の検討

小楠 竜也† 大谷 淳† 榎沢 順‡

本検討は抽象絵画と音楽における互いの表現要素を工学的アプローチによって結びつけることで両者を融合した新しい表現を提案する。そのアプローチの方法として絵画鑑賞者の視線行動による画面上の注視要素の時系列的な変化を音楽の時間情報として利用する方法を考えた。今回は抽象絵画を対象に鑑賞者の視線行動を計測し、視線の停留時間は平均と異常値の両極端に別れることや、要素ごとの停留回数には要素の大きさを中心からの距離が大きく影響しているといった傾向が見られた。また、色と音の印象や共感覚保持者の感覚を参考に色と音のマッピングを検討することで注視要素を音に変換し、音群の音楽を踏襲する作曲方法の可能性を示した。

Study of Music Composition Utilizing Gaze Behaviors of Persons Viewing Abstract Paintings

Tatsuya Ogusu† Jun Ohya† Jun Kurumisawa‡

This paper proposes a method that allows music composition based on abstract paintings. Our approach utilizes gaze behaviors of viewers who see abstract paintings, where temporal changes in gaze positions are utilized as temporal information for music. In this paper, we measured the viewers' gaze behaviors when they see different abstract paintings, and found the following things: (1) durations in which the fix one's gaze on some position are classified into "average length" group and "extremely long" group, and (2) the number of gaze's stay in each element depends on the element's size and distance from the center of the painting. By analyzing mapping between color and sound based on synesthesia, the element at which the gaze stays is converted to sound so that a new music composition method for generating tone clusters is shown.

1. はじめに

1.1 背景と目的

現代美術において抽象絵画が描かれるようになった根源のひとつに《音楽》がある。ポール・ゴーギャンは「絵画の理想は音楽のように感覚を通して魂に間接的に働きかけるものだ」と語っている[1]。また、初期の抽象画家であるワシリー・カンディンスキーは、自身の著書[2]において音響と色彩および音楽と絵画の親近性について説いている。

現代ではコンピュータやセンサーを用いて様々な視覚メディアや音響を融合させたマルチメディアな表現が頻繁に行われるようになったが、前述したような抽象絵画の音楽性を表現するといった試みは少なく、具体的な方法も確立されていないといえる。本検討の目的は、抽象絵画と音楽における互いの表現要素を工学的に結びつけることによって、抽象絵画から音楽を作曲する新しい表現の方法を提案する。

1.2 アプローチ

1.2.1 絵画鑑賞者の注視行動による時間的変化の抽出

音楽は時間的芸術と呼ばれるように、表現には時間を必要とする。しかし、静止画である絵画は映像のように時間情報を持たない。そこで、本検討では絵画鑑賞者の視線行動に着目した。

通常、絵画を鑑賞する際、鑑賞者の視線は画面上の様々な位置へ移動する事が知られている[3]。従って、絵画鑑賞者の視線を計測することによって得られる視線の移動軌跡(図1, ①)から注視した要素を順番に取得し、時系列的に配置(図1, ②)することで音楽の時間と結びつける方法を考えた。

静止画から時間情報を抽出して音楽作曲を行う方法として、これまでも次のような例がある。1970年代に現代音楽の作曲家ヤニス・クセナキスによって開発されたUPICは、ユーザーがペンとタブレットを用いて点や線などの図形を描き、それをコンピュータが読み取ることで電子音が発せられる作曲ツールである。しかし、UPICは図形の上下の座標を音高に割り当て、左から右にスキャンしていく五線譜の記法と同じ

†早稲田大学大学院 国際情報通信研究科

‡千葉商科大学 政策情報学部

†Waseda University Graduate School of Global Information and Telecommunication Studies

‡Chiba University of Commerce Faculty of Policy Informatics

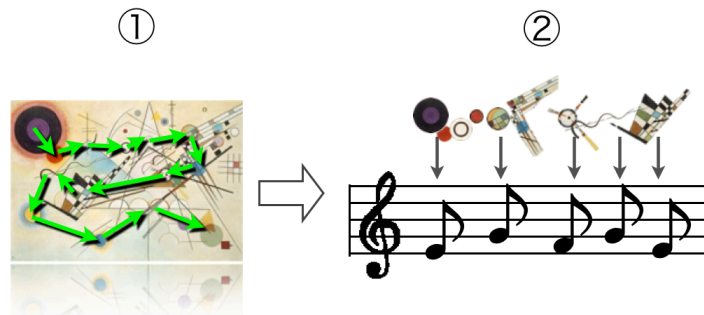


図1 基本的な考え方

であり、描かれた図形はすべて楽譜上の音符と同じ位置関係で扱われる。従って、構図などの絵画的な表現性を反映できない。UPICの後継機もこれまでにいくつか検討[4]されているが、基本的な考え方はすべて楽譜の拡張であり、絵画的な表現性を取り入れようとした試みは少ない。

本論文では実際に抽象絵画による視線行動を計測し、統計的な解析によって見出された特徴や傾向について2・3節で述べる。

1.2.2 色と音のマッピングによる音群の生成

抽象絵画は具象的な対象物がないため、平面上に構成された色彩そのものが表現である。このような極論的な概念は、音楽においては音群的音楽の概念に類似する。音群的音楽は現代音楽のひとつであり、音楽を音の集合体（音群、音響テクスチャー）として捉え、作曲家独自の方法によって生成された音群を時系列に配置することが基本的な作曲法である[5]。本検討では音群的音楽の作曲法に習い、注視要素から得られる色彩情報によって音群を生成する方法を提案する。

音群の生成方法としては様々な方法が作曲家によって恣意的に行われてきた。しかし、今日では色と音の心理的効果や、共感覚の研究により色と音の対応関係を解明する研究が行われている。そこで、今回は過去に行われてきた音と色の印象評価実験や共感覚の研究などから参考データを引用し、両者を印象や共感覚保持者の感覚をもとにデータを掛け合わせる方法を検討した。具体的な方法については4節で述べる。

2. 抽象絵画における視線行動の計測

2.1 視線計測の目的と方法

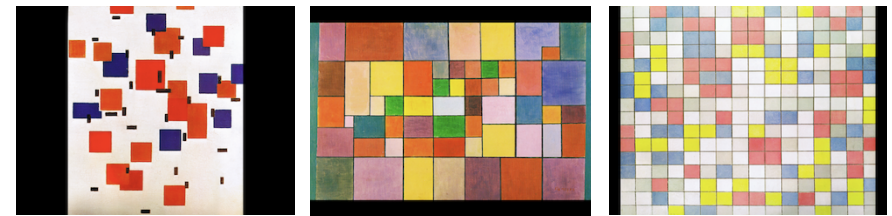


図2※ 提示画像1

図3※ 提示画像2

図4※ 提示画像3

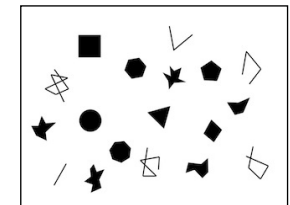


図5 提示画像4

これまでも視線行動の計測と解析は様々な研究で行われてきた[3]。絵画を刺激対象にした視線計測においては、具象絵画を対象とした視線計測は行われてきたが、抽象画に対してはあまり行われて来なかった。しかし、実際には色や形状など抽象的な要素においても、それぞれ誘目性が異なることが報告されている[6]。従って、今回は抽象絵画においても、色彩や要素の配置によって視線行動に影響があると仮定し、どのような傾向があるかを解析するために抽象絵画における視線行動の計測を計画した。

今回、視線の計測には株式会社ナックイメージテクノロジーのアイマークレコーダー「VOXER」を使用した。VOXERは角膜反射法を用いた非接触型の視線計測装置であり、装置から赤外線を投影し、被験者の角膜に反射した像点の座標位置を測定する方法によって計測される。

刺激提示用にアスペクト比4：3、サイズが22インチのPCディスプレイを使用し、ディスプレイと被験者の眼との距離は約75センチで固定した。

実験に使用した画像は既存の抽象絵画12枚と筆者が作成した単純図形画像6枚である。このうち単純図形画像のうち3枚については明確なデータが得られなかったため、今回の解析からは除外した。刺激対象画像の一部を図2～5に示す。

2.2 視線計測の手順

今回の視線計測では正常な視力を持つ大学生および大学院生20人の男女を被験者とした。

※図2: ピエト・モンドリアン『白地の上の純粋色面によるコンポジションA』

図3: パウル・クレイ『北方のハーモニー』

図4: ピエト・モンドリアン『明るい色による市松模様のコンポジション』

計測の手順は以下の通りである。

1. まず、被験者には実験で行なってもらおうことの内容と、実験中の注意事項を口頭または用紙で説明する。ここで、実験の注意事項として「なるべくあごを引く」「ディスプレイの外を見ない」など、正確な視線情報を取得するために必要なことを告げた。

2. 次に、被験者は視線検出装置の前に座り、目の前のディスプレイに表示された刺激画像を順番に鑑賞する。刺激提示の流れは以下の①②③の繰り返しで行った。

(カッコ内に時間を示す)

- ①待機画像 (10秒)
- ②刺激提示 (30秒)
- ③簡易アンケート記入 (10~20秒)

このとき、③においては直感で回答してもらうため、20秒を制限時間として設けた。

3. 最後に総合アンケートと被験者プロフィールシートを記入してもらう。

3. 視線行動の解析と考察

3.1 停留の定義

視線データの解析にあたって、視線の停留を定義しておく必要がある。今回は視線座標が半径 25pixel (停留候補領域以内) で移動した場合を停留候補としてカウントし、連続して200ms以上の停留候補がカウントされた場合を1回の停留としてとらえることにした。また、停留候補の範囲の決め方は停留開始点を含むすべての停留候補の重心から 25pixel 以内とした。

3.2 停留時間の異常値

すべての観測された停留点から、絵画と被検者ごとに停留点座標と停留時間を円と円のサイズで表現したグラフを作成した。一部の結果を図6と図7に示す。

多くの被験者の場合、図6のように一定の停留時間内において停留点が分布していた。しかし、なかには図7の大きな円に表されている通り、極端に永い停留時間が観測されるケースが見られた(図7では1つの停留において約3.5秒間の永い停留が観測された)。このことから、停留時間は平均と異常値の両極端に分けられると考えられる。また、停留時間が異常値か否かを判断する異常値検定を以下の式1を用いて行なった。なお、式1において[$T_i > 3$] となった場合を異常値と見なした。

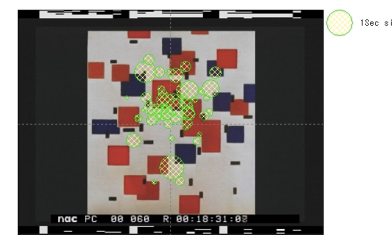


図6 被験者Aの停留点分布

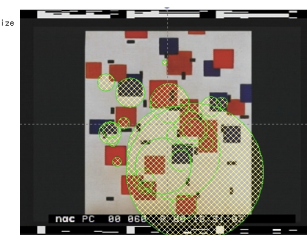


図7 被験者Bの停留点分布

$$T_i = \frac{x_i - \mu}{\sigma} \quad (式1)$$

(x_i = 停留点 i における停留時間、 μ = 平均、 σ = 不偏標準偏差)

各絵画と被験者毎に検出した異常値検出数は0~2回であり、稀に3回検出する場合もあった。また、被験者別と絵画別に、それぞれの異常値検出数の合計を停留回数合計で割った数を百分率にした異常値検出確率を割り出した結果、被験者毎の個人差や絵画毎の差は見られたが、全体でおよそ1.5%~3%の間に収まり、大きな差は無いといえる。

3.3 停留時間と停留回数における個人差

観測された全ての停留時間と停留回数において被験者間に個人差があるか否かを調べるため、全12枚の絵画に対して被験者別に停留回数の合計と停留時間の中央値を算出した。結果を図8、図9に示す。

図8を見るかぎり、ばらつきはあるものの被験者毎にある程度のもどまりが見取れる。図9は値が 400ms を中心にまとまっている被験者と、広範囲にばらつきがある被験者の2種類に分かれた。また、両者の結果を被験者別に平均値をとって相関係数を算出したところ[-0.702]となり、やや強めの負の相関があるといえる。従って、すべての絵画における停留回数・停留時間のそれぞれの値には被験者毎に個人差があり、停留回数が多い被験者は停留時間が短く、反対に回数が少ない被験者は停留時間が平均的に永くなり、絵画毎のばらつきが大きくなる傾向があることがわかった。

また、図8と図9の結果を、3.2節の被験者別の異常値出現確率との相関係数を調べたところ、停留回数との相関係数は 0.0256、停留時間との相関係数は -0.261 となり、相関はほとんど無いといえる結果となった。

3.4 要素毎の停留回数と面積および画面中心からの距離との相関

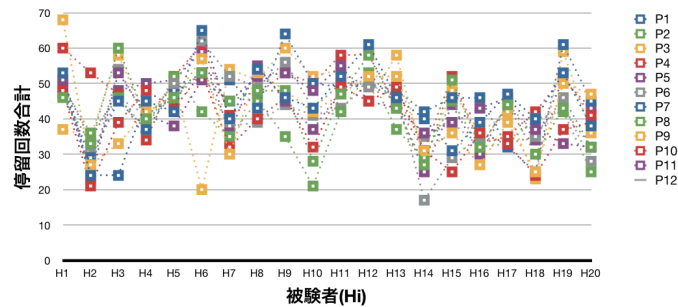


図8 絵画毎の被験者別停留回数合計

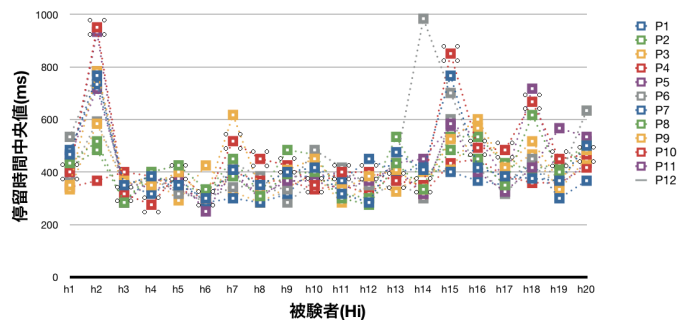


図9 絵画毎の被験者別停留時間中央値

先行研究[6]によると、大きな要素ほど誘目性が高く、また、画面の中心に近い要素の方が誘目性が高いと言われている。従って、一部の提示画像を対象に要素毎の占有画素領域におけるそれぞれの停留回数を算出し、各要素の物理的特徴量として《面積》と《画面中心からの距離》との比較を行なった。また、ここでは各要素における停留回数は観測された停留点座標がそれぞれの要素が占める画素の領域に入っているか否かによってカウントした。なお、停留点座標は固視微動や計測器の誤検出などにより実際の注視点とのズレが頻繁に発生する為、停留点座標を中心とした縦横50pixelを許容範囲とし、要素が占める画素領域がこの範囲内にあった場合をその要素の停留回数としてカウントすることにした。

各要素毎の面積を、その要素が含まれる領域の全画素数で計測し、要素が含まれる領域の x, y 座標の平均値(重心)から、画面の中心座標とのユークリッド距離を要素の画面中心からの距離と定義した。各要素の停留回数の平均を求め、面積および中心からの距離との散布図を作成した。結果の一部を図10と図11に示す。

図10と図11を見るかぎり、極端に平均停留回数が多い要素が1~2個存在するというのがわかる。また、基本的に平均停留回数は面積の大きさと正の相関、画面中心からの距離に対しては負の相関という傾向にあることがわかる。しかし、図3、図4、図5の提示画像

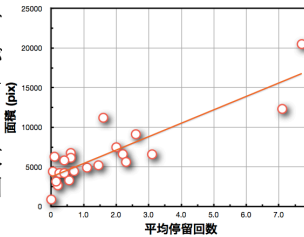


図10 各要素の停留回数と面積の散布図

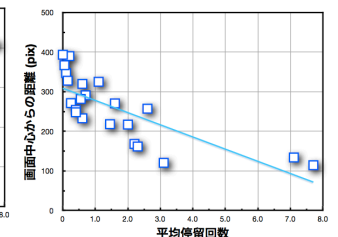


図11 各要素の停留回数と画面中心からの距離の散布図

に限っては、面積との相関が見られなかった。その原因は定かではないが、図3の絵画は画面中心から外側に向かって要素の面積が大きくなり、図4は全ての要素の面積にあまり差が無いなどの特徴がある。また、図5においては占有画素数の合計と領域の広さに違いがあり、大きさを面積で測る以外の方法が必要と思われる。これらの特徴が面積との相関係数を低くする要因ではないかと考えられる。

4. 色と音のマッピング

4.1 色と音の印象評価を用いた色による音色操作

色には明るさによって大きく見えたり、感情や情緒に影響を与えたりといった固有の性質があるといわれている。一方で、音楽や環境音に対しても様々な感情効果や心理的反応があることがわかっている。これらの印象を調べる方法として、「大きい—小さい」「重い—軽い」などの形容詞対により評価するSD法が知られている。そこで、今回は色と音に対してSD法による印象評価実験を行った様々な研究を参考に、両者の形容詞対毎の評価値を参照して互いを結びつける方法を以下のように検討した。

(A) 色に関する印象評価として柳瀬徹夫氏の研究[7]を参照する。この実験では、100名の学生を対象に[赤・黄・緑・紫・青・白・黒]の7色を20種類の形容詞対で評定し、各色の平均評価値を1~7段階に順位付けした結果が得られている。

(B) 音における印象評価として、工業ノイズや報知音、電子音などを対象に印象評価を行なった研究結果[8][9][10]をもとに[音高・大きさ・高周波成分・音の立ち上がり速度・ゆらぎ(ビブラート、トレモロなど)]の5つの音色要素に対して39種類の形容詞対による印象評価の結果を参照する。なお、具体的な評価値まで考慮すると対応付けが困難になるため、ここではすべての形容詞対に対する印象傾向を2段階に単純化した。

(A) (B) で参照したすべての形容詞対から、それぞれ同じ意味となる形容詞対6種類を抽出した。色から得られた形容詞対の評価値を表1、音色から得られた形容詞対の評価値を表2に示す。なお、表1の数値は形容詞対の左=1、右=7の7段階の順位を示している。(例：明るい～暗い=1～7)

以上の色と音に共通する6種類の形容詞対評価値を用いて、それぞれをプログラム上で実装するための計算方法を以下に述べる。

- ①入力色RGB (視線による注視要素から得られた平均色) をHLS色空間に変換する。
- ②色相H ($0 < H < 360$) の値から表1の赤・黄・緑・青・紫のどの色相に近似するかを決定する。このとき、各色のH値は赤[0]、黄[60]、緑[120]、青[240]、紫[300]とする。また、決定された色に対する6種類の形容詞対評価値をXとする。
- ③輝度L ($0 < L < 1$) の値から表1の白と黒のどちらに相当するかを決定する。このとき、L=1は白、L=0は黒とする。また、決定された白と黒のいずれかに対する6種類の形容詞対評価値をYとする。
- ④XとYのどちらの評価値が適当であるかを彩度S ($0 < S < 1$) の値によって判断する。もし、Sが1であればXを、Sが0であればYの形容詞対評価値を採用する。
- ⑤採用された6種類の形容詞対評価値を、それぞれ対応する音の形容詞対に当てはめる。この時、音の形容詞対評価値は色と同じ1～7にスケールアップし、対応する形容詞対評価値を代入する。また、対応する形容詞対が服すある場合は、それぞれの評価値の平均値を代入する。例えば赤の場合、[明るい - 暗い] [軽い - 重い] [澄んだ - 濁った] はそれぞれ3, 4, 5, となり、それらが対応する音の形容詞対 (上記3つの場合、音高) に、評価値の平均を代入する。

以上のように色と音の印象評価実験において互に対応する形容詞対を用いれば、注視要素の色と印象がマッチした音色操作が可能であると考えられる。

4.2 共感覚保持者の感覚による色の調性変化

共感覚とは、ある知覚によって全く別の感覚が刺激される特殊な感覚のことである。この感覚がある人々を共感覚保持者と呼び、彼らは文字に色が見えたり、音楽を聞くと様々な色が見えるという[11]。今日では色と音の適切なマッピング方法の検討として共感覚保持者の共通感覚を見出す研究が行われている。

長田典子氏らが行った研究[12]では、4人の共感覚保持者において様々な調の音楽を聞かせ、同時に知覚する色を答えてもらったところ、ある程度共通性が見られたという。本検討ではこの研究データをもとに色から調性を決定する方法を考えた。以下にその方法を述べる。

- ①4人の共感覚保持者がそれぞれの調に対して知覚した色を、すべてRGB色空間上にプロットする。

表2 音の印象評価値

音高		大きさ (音圧)		高周波成分		立上り時間 (アタック)		ゆらぎ	
高い	低い	大きい	小さい	多い	少ない	早い	遅い	大きい	少ない
明るい	暗い	強い	弱い	明るい	暗い	硬い	軟らかい	軟らかい	硬い
軽い	重い	明るい	暗い	はっきり	ぼんやり				
澄んだ	濁った								

表1 色の印象評価値

形容詞対	赤	黄	緑	青	紫	白	黒
明るい - 暗い	3	1	4	5	6	2	7
軽い - 重い	4	1	3	5	6	2	7
澄んだ - 濁った	5	2	3	4	7	1	6
強い - 弱い	3	1	5	4	6	7	2
はっきり - ぼんやり	4	1	5	6	7	3	2
軟らかい - 硬い	4	1	3	6	5	2	7

- ②入力色 (視線による注視要素から得られた平均色) を上記RGB色空間上にプロットし、すでに①で入力された色のうち最も距離が近い色を判定し、その色に対応する調を判定結果として出力する。このとき、判定された色における調が複数ある場合は、それらのうちランダムで選択する。

共感覚保持者が音から感じる色には水色や灰色など、色相や彩度だけでは決定できない色が多いため、今回は上記の通りRGBの複数のパラメーターを考慮したマッピング方法を考えた。

5. 音群の生成と作曲

視線行動の計測により得られた視線データと対象画像を使用し、4節で提案する色と音のマッピング方法を取り入れて音響化とパラメーターの抽出を試みた。ここでは図3の絵画を対象に一人の被検者による視線データを用いて 0ms ~ 6039ms までの注視要素と抽出されたパラメーターをそれぞれ図12、図13、図14に示す。なお、図13と図14の(1)~(8)は図12の注視要素の順番を表す。

図12は観測された注視要素を時系列に並べたものである。図3の絵画は四角形の集合なので、色と大きさのみ異なる四角形の要素が抽出されているのがわかる。key 行のアルファベットは注視要素の色から選択された調名である。これら選択された調から和音やモチーフを生成させることで色彩と協調する楽節を生成させることができると考えられるが、適切な和音およびモチーフ生成方法の検討は今後の課題である。

図13は注視要素から 4.1節の計算方法によって算出された6種類の形容詞対評価値である。それぞれの値は色の印象によって変化している。また、これらの印象にあわせて音色パラメーターを操作し、ノイズ音の音色操作を行った結果を図14に波形で示す。なお、ここでは音色の変化のみを見るためにホワイトノイズを音源とした。

図14では音圧や音の立ち上がりに変化が見られるが、ゆらぎ (ここではトレモロ効果を割り当てた) や高周波成分の違いはあまり見られない。これらは論理的な値と人

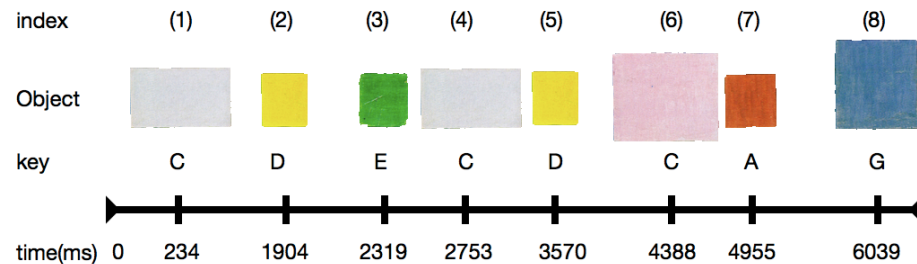


図12 注視要素の時系列配置と選択された調名

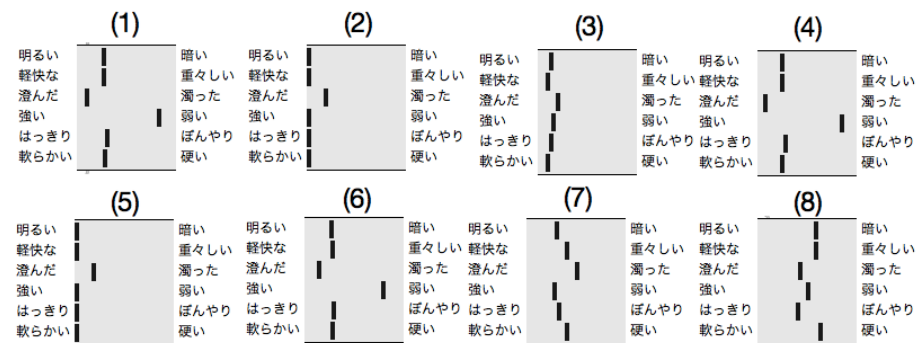


図13 注視要素の色から得られた印象評価値

間の感覚が異なることも考慮し、必要な重み付けを行うなどの調整が必要と思われる。

このように注視要素から抽出されたパラメータはいくつかの調整の余地があるものの、複数の値が様々に変化することで表現力に富んだ音群の生成が可能であるという見通しが得られた。

6. まとめと今後の展望

抽象絵画における視線行動を利用することで絵画から時系列的情報を抽出し、音楽の時間的構造に対応付けることにより、絵画から音楽を作曲する表現方法の検討を行った。本論文では抽象絵画における視線行動の計測を行い、統計的な解析によっていくつかの特徴や傾向が解明できた。また、色と音の印象や共感覚保持者の感覚を元にマッピングする方法を検討し、注視要素の色彩から音色や調性を操作する方法を提

案した。これら視線行動と音と色のマッピング方法を組み合わせることにより、抽象絵画から時系列的な音の変化を発生させることで音群的音楽の作法を踏襲する新しい作曲方法の可能性を提示できた。

現状ではまだ視線データの解析結果を取入れた検討はされていないが、今後は、これらの視線行動特徴をモデリングすることで、様々な抽象絵画やユーザーが描いた抽象絵画などに適応することが考えられる。

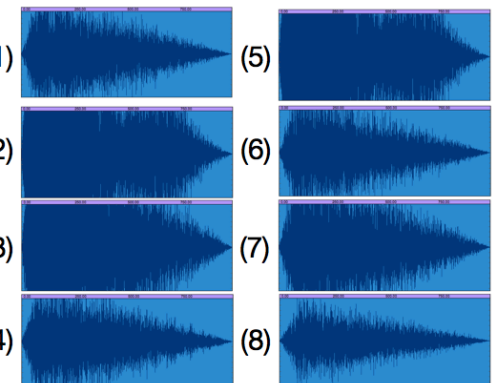


図14 色の印象から生成した音響の波形

参考文献

- [1]視覚デザイン研究所・編集室：“体験！現代美術”，視覚デザイン研究所(1991)
- [2]Wasily Kandinsky,西田秀穂(訳),西村規矩夫(訳)：カンディンスキー著作集，美術出版社(2000)
- [3]池田光男：“眼はなにを見ているか 視覚系の情報処理”，平凡社(1988)
- [4]長嶋洋一：“PGS (Polyagogic Graphic Synthesizer) の検討” 情報処理学会研究報告.[音楽情報科学] 2005(14), pp.39-42(2005)
- [5]中村滋延：“現代音楽×メディアアート 音響と映像のシンセシス”，九州大学出版会(2008)
- [6]田中昭二，井口征士，岩館祐一：“画像領域の物理的特徴に基づく誘目度評価モデルの検討”，電子情報通信学会論文誌, J83-A(5), pp.576-588 (2000)
- [7]柳瀬徹夫：“色彩計画，工業デザイン全集，第4巻所収”，日本出版サービス(1982)
- [8]上田隆司，佐久間邦郎：“ホーニング音の研究：SD法による音色評価”，精密工学会誌 63(7), pp.1039-1043 (1997)
- [9]上野山努，櫻村雅章，小沢慎治：“ドラム音の音色における感知情報と工学的パラメータとの対応付け”，日本音響学会誌 49(10), pp.671-681 (1993)
- [10]崔鍾大 ほか：“自動車内のリバーズ報知音にとって望ましい音響特性”，日本音響学会誌 61(3), pp.118-125 (2005)
- [11]リチャード・E・シトーウィック：“共感覚者の驚くべき日常一形を味わう人、色を聴く人”，草思社(2002)
- [12]長田 ほか：“音と色のノンバーバルマッピング—色聴保持者のマッピングルール抽出とその応用—，電子情報通信学会論文誌A, J86-A(11), 1219-1230(2003)