

特微量抽出に基づく画像と音楽を融合した 芸術作品の鑑賞コンテンツの生成システムの提案

宮本 拓実[†] 王 元元[†]

† 山口大学工学部知能情報工学科 〒755-8611 山口県宇部市常盤台 2-16-1

E-mail: †i081ff@yamaguchi.ac.jp, ††y.wang@yamaguchi-u.ac.jp

あらまし 昨今では、オンライン上でも芸術作品が鑑賞できるようになった。しかしながら、芸術作品を自発的に触れる人が少ないため、興味・関心を持ちやすくするための新たなアプローチが必要である。そこで本研究では、絵画作品の画像と音楽の特徴量抽出に基づき、異なる絵画作品の感情や雰囲気に対応させた音楽を融合した鑑賞コンテンツを生成するシステムを提案する。本システムでは、ユーザが鑑賞したい絵画作品の画像を複数枚選択し、その画像群の特徴量と音楽ごとの特徴量との類似度算出に基づき、画像群と同じ印象を持つ音楽を選定する。次に、楽曲の展開に応じて絵画作品の画像を提示するために、選定された1つの楽曲を複数のセクションに分割し、それぞれのセクションと画像ごとの特徴量との類似度を算出することで、画像と音楽を融合した鑑賞コンテンツを生成する。これにより、絵画作品から感じる印象と、音楽から感じる印象の相乗効果により、没入感が増加や作品の魅力を感じられるオンライン鑑賞になることが期待できる。

キーワード 芸術鑑賞支援, オンライン鑑賞, 鑑賞コンテンツ生成, 特徴量抽出, 画像と音楽の融合

1 はじめに

昨今では新型コロナウイルスの影響を受け、芸術作品を鑑賞する機会は減少している。文化庁の世論調査[1]によると、この1年間に文化芸術イベントを直接鑑賞したものがないと答えた割合は60.3%であり、コロナ前の29.8%と比べて激増している(図1)。その背景として鑑賞しなかった理由には、「新型コロナウイルスの影響により、公演や展覧会が中止になった、又は外出を控えたから」と回答した人の割合が37.6%、「関心がない」と回答した割合が22.8%であった。このように文化芸術に対する直接鑑賞はコロナ禍により大きく影響を受けた。

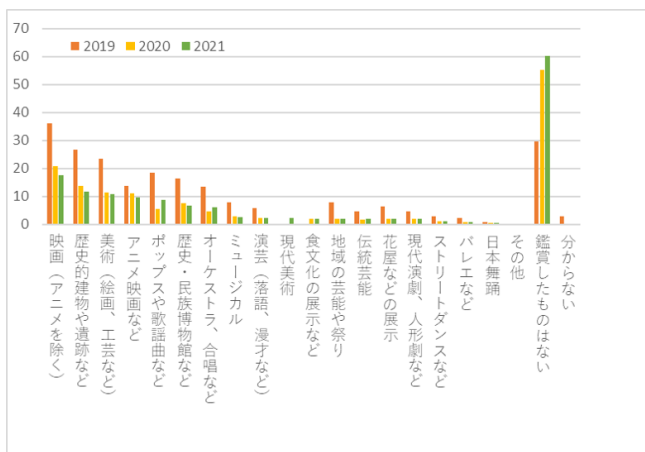


図1 直接鑑賞した文化芸術のジャンル

その影響下で、Google Arts & Culture¹のストリートビューや

世界の美術館が独自に提供しているオンラインビューイングやバーチャルツアーによるオンライン鑑賞が展開されている。時間と空間の制約を介さず、家の中でも世界中の名画を好きなだけ鑑賞できるという利点がある。しかし、オンライン鑑賞の問題点として「直接鑑賞との体験価値の差」があげられる。実際の美術館での鑑賞では、作品を五感でより高い解像度で鑑賞する事ができる。オンライン鑑賞では体験できない体験価値を補うことが昨今の芸術鑑賞に必要であると考えられる。

本研究ではオンライン鑑賞の体験価値に深みをつけ、芸術鑑賞に対する興味関心を深めることを目的とする。特に鑑賞の体験価値として聴覚を補うため、音楽との融合による芸術作品の鑑賞システムを提案する。提案手法では、ユーザが選択した絵画作品の画像に対し、それらの画像群の特徴量と音楽ごとの特徴量を類似度算出し、画像群に適する音楽を選択する。次に、その楽曲の展開に対して、それぞれに適した雰囲気絵画作品を提示するために、楽曲を複数のセクションに分割し、それぞれのセクションと画像の特徴量を類似度算出し、絵画作品と音楽の融合した、鑑賞コンテンツを生成する。絵画と音楽の融合から、より作品への没入感や迫力を向上するための相互作用となることを期待する。

本論文の構成は以下のとおりである。2章では、芸術鑑賞における体験価値の向上に関する研究と画像と音楽の融合に関する研究について紹介する。3章では、画像と音楽を融合した鑑賞コンテンツの生成について説明する。4章では、提案手法によって生成された鑑賞コンテンツの有用性を図るための評価実験について述べる。最後に、5章では、まとめと今後の課題について述べる。

1 : <https://artsandculture.google.com>

2 関連研究

芸術鑑賞における体験価値の向上や画像と音楽の融合に関する関連研究と本研究の関連について述べる。

2.1 芸術鑑賞における体験価値の向上

Usui ら [2] は、芸術鑑賞教育を目的とした、絵画のインタラクティブな表示方法である D-FLIP Paintings を提案している。ユーザ間の多様なインタラクションに応じて、数百枚の絵画をリアルタイムで柔軟にレイアウトすることで、コミュニケーション型美術鑑賞のための主体的な学習や生徒間の対話による新たな気づき、興味関心の向上が期待される。Spence ら [3] は、美術館の作品を親しい人にギフトという形で写真を送るアプリケーションを提案している。これにより、新たな視点での作品鑑賞や人との繋がりが生まれる。Muntean ら [4] は無形の民族文化を学ぶために、それらを体験する展示設計とシステムを提案した。民族文化の展示物を現代との比較を通して、鑑賞者に探索させる体験を付与している。

このようにさまざまな現代技術を用いたアプローチによって鑑賞者自身が芸術作品に興味を持ち、さまざまな視点から鑑賞することで新たな発見や学習、創作意欲の向上など鑑賞体験価値が深まる。しかしながら、コロナ禍の影響で、オンラインでも芸術鑑賞が可能となった今、オンライン鑑賞で同様の学びや鑑賞体験が十分に提供されていない。そこで、本研究ではオンラインでの絵画鑑賞に新たな学びや創造性を育むためのアプローチとして音楽との融合を図る。

2.2 画像と音楽の融合

音楽と写真を同期させる最初の試みは Sakai [5] らによって 2013 年に報告されている。音楽と写真から Arousal (覚醒度) と Valence (感情価) を抽出し、写真の HSV と音楽の 19 の音響特徴に基づいて Arousal と Valence の値を算出し、感情的な音楽推薦システムを提案した。福本ら [6] は印象派絵画の色彩が Zipf の法則に従うことから統計的特徴を利用し、絵画と文章、絵画と音楽の融合による表現・鑑賞手法の実現を提案した。Lee ら [7] は Russell の Arousal-Valence モデルを用いて絵画と音楽の融合システムを提案した。絵画から色の組み合わせを抽出し、事前に定義された色情動モデルを用いて、統計的学習過程を経ずに Arousal と Valence の値を推定している。

これらの研究では、音楽と画像の対象が 1 対 1 で対応付けられていたが、本研究では、鑑賞者が複数の絵画作品を鑑賞する際の音楽との融合を目指す。音楽には 1 つの曲の中にさまざまな展開や雰囲気があり、それらを鑑賞する作品ごとに融合させることで、鑑賞体験の向上を目的とする。また絵画作品には作品画像だけでなく、説明文にも作品の特徴や雰囲気の要素を捉えることができるため、本研究ではテキスト情報からも特徴量を抽出する。

3 画像と音楽を融合した鑑賞コンテンツの生成

本研究では、まず鑑賞者が鑑賞したい絵画作品の画像を複数枚選択し、その画像群全体の特徴量と音楽データベース内の各楽曲の特徴量との類似度算出に基づき、画像群と同じ印象を持つ音楽曲を選定する。次に、楽曲の展開に応じて絵画作品の画像を提示するために、選定された 1 つの楽曲を複数のセクションに分割し、それぞれのセクションと画像ごとの特徴量との類似度を算出することで、画像と音楽を融合した鑑賞コンテンツを生成する手法を提案する。図 2 に生成システムの概要図を示す。



図2 提案する鑑賞コンテンツ生成システムの概要図

3.1 画像の特徴量抽出

本研究では、メトロポリタン美術館²のコレクションのパブリックメインの作品のうち、41 作品を画像データとして使用し、これらの作品の解説文の感情値、及び画像の彩度と明度を抽出する。作品の解説文の感情分析には感情分析ツールの VADER [8] を用いる。VADER における総合的な感情の評価値として定義されている compound 値を求め、その値を感情値とする。画像の彩度と明度においては標準化して算出する。これら 3 つの特徴量からなる絵画作品ベクトルをそれぞれの作品ごとに生成する。その後それらの平均値をとり、作品群全体の絵画作品ベクトルを生成する。

本論文は、鑑賞コンテンツを 3 分程度の動画コンテンツにするために、鑑賞者には選択する絵画の画像データを 5~7 枚の数で指定し、それに対応した楽曲が選出される。

3.2 楽曲の特徴量抽出

本研究で使用した楽曲は、「毎日 3 分！親子でクラシック★ベスト 50」³から 41 曲を対象とした。歌詞が入ると鑑賞時の妨げになってしまうことや、ジャンルによって嗜好性の差や印象の違いが発生することを懸念し、広く聞かれる機会の多いクラシック音楽を選定した。また、楽曲は 3 分以下のものとなっている。

ここでは特徴量抽出に算出するために、2 種類のツールを用いる。Spotify⁴の提供する audio-features と Python の MIR ライブラリである Librosa⁵を用いた。この audio-features は Spotify

2: The Metropolitan Museum of Art: <https://www.metmuseum.org/visit/plan-your-visit/met-fifth-avenue>

3: 日本コロムビア株式会社: 毎日 3 分！親子でクラシック★ベスト 50

4: <https://www.spotify.com>

5: <https://librosa.org/doc/latest/index.html>

のデータベースにある、楽曲全体の音響特徴量しか算出することができないため、分割されたセクションごとの特徴量を抽出するために Librosa を用いた。2 種類のツールを用いることで算出する特徴量間に差異が発生しきまらないよう、画像と音楽のそれぞれの特徴量の対応付けを吟味する必要がある。以下に選出した音響特徴量と画像特徴量との対応付けを記述する。

3.2.1 楽曲全体の特徴量

楽曲全体に対する特徴量は、Spotify が提供する定量的評価値の Valance, Energy, Tempo を用いる。Valance は 0.0 から 1.0 までの値で、楽曲が伝える音楽的なポジティブさを表現する。Energy は 0.0 から 1.0 までの値で、強度と活性の知覚的な尺度を表す。各 Tempo は曲の全体的な推定テンポを 1 分あたりの拍数 (BPM) で表す。ここで特徴量の選出においては、中野ら [9] が、各特徴量のデータ分布、分散などについて全曲を調べた結果、Energy, Popularity, Tempo, Valance は値が分散しており、雰囲気異なる音楽に対しても対応できる可能性が高いと考えられているため、上記の Valance, Energy, Tempo の 3 つの特徴量を使用する。これらを標準化し、それぞれの 3 つの特徴量からなる楽曲ベクトルを生成する。図 3 に音楽選定のための特徴量抽出する評価値と画像特徴量との対応付けを記述する。

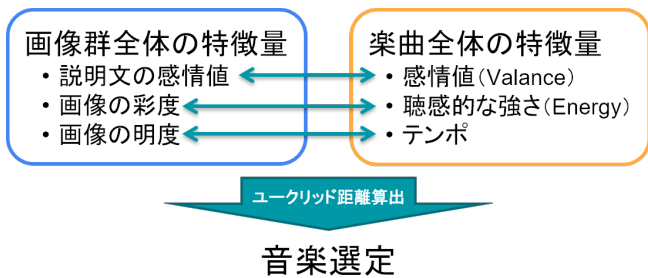


図3 音楽選定のための特徴量抽出

ここで対応付けについて、説明文の感情値は同じ感情値である Valance を用いた。彩度と明度に関しては、安達ら [10] が、彩度の高低と楽曲の力強さの度合いとが対応し、明度が高くなるにつれ、曲の印象も明るくなる傾向にあることを結論付けた。また、星野ら [11] は、テンポの遅速は明度段階にほぼ対応すると述べた。ここから彩度と Energy, 明度と Tempo を対応付けた。

3.2.2 楽曲の各セクションの特徴量

選出した背景音楽に対して分割されたセクションごとの特徴量は Librosa の spectral contrast, rms energy, tempo を用いる。spectral contrast はスペクトル重心で周波数の明瞭さ、rms energy はエネルギーの二乗平均平方根、tempo は楽曲の速度を表す。これらを標準化し、3 つの特徴量からなる楽曲セクションベクトルを生成する。図 4 に音楽と画像の融合のための特徴量抽出する評価値と画像特徴量との対応付けを記述する。

ここで、スペクトル重心を用いて音楽の明るさの定量尺度化を試みたため、これを説明文の感情値と対応付けた。rms energy と tempo は上記の記述より、彩度と明度を対応付けた。

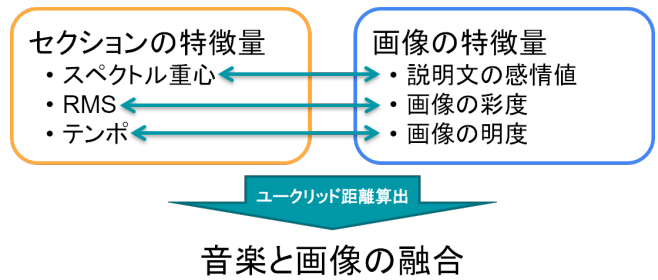


図4 音楽と画像融合のための特徴量抽出

3.3 画像と音楽の対応付け

3.1 節の方法で絵画作品ベクトルを算出した絵画作品群 X と 3.2 節の方法で音楽データベース内の n 番目の楽曲ベクトルを算出した Y_n とのユークリッド距離を下記の式 (1) により算出し、最も類似度の大きい組み合わせを決定する。

$$d(X, Y_n) = \sqrt{(X - Y_n)^2} \quad (1)$$

次に選定された音楽に対して、Spotify の audio-analysis を用いて楽曲をセクションごとに分割する。分割されたセクション数を j とし、Librosa を用いて音響特徴量を抽出し、 j 番目のセクションベクトル y_j とする。その後、絵画作品群の各作品ごとの i 番目の絵画作品ベクトル x_i と先ほど生成した y_j のユークリッド距離を下記の式 (2) により算出し、最も類似している順に作品とセクションを決め、それぞれの組み合わせを決定する。

$$d(x_i, y_j) = \sqrt{(x_i - y_j)^2} \quad (2)$$

4 評価実験

4.1 プロトタイプシステム

本節では、提案手法に基づいて融合された絵画作品と背景音楽の例を示し、その特徴について述べる。今回使用した絵画作品を図 5 に示す。

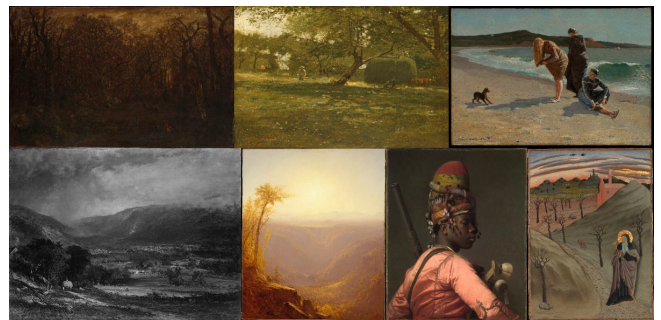


図5 使用した絵画の画像群

表 1 にこれらの絵画作品の特徴量を標準化して算出した数値を示す。この絵画群の平均特徴量に対して音楽データベース内の楽曲特徴量とのユークリッド距離を算出し、最も類似度の高い楽曲を選定する。今回の場合は「ジュール・マスネ: タイスの瞑想曲」が最もふさわしい音楽として選択された。次に選定されたこの曲を 7 つのセクションに分割し、それぞれのセクショ

表1 絵画作品の特徴量

画像	説明文の感情値	彩度	明度
1	-1.684	1.387	-1.625
2	-1.249	1.824	-0.090
3	-2.312	-2.424	-0.888
4	-1.837	0.085	-0.945
5	-2.831	-0.160	-0.169
6	-2.453	-0.452	-0.189
7	-0.973	0.217	1.680
平均値	-1.906	0.068	-0.318

表2 タイスの瞑想曲の特徴量

セクション	spectral contrast	rms	tempo
1	-0.675	-0.4670	0.477
2	-1.109	-1.229	0.477
3	-0.229	0.104	-0.952
4	-0.844	0.239	1.085
5	0.389	-0.081	0.867
6	2.057	2.161	-1.924
7	0.410	-0.724	-0.033
楽曲全体	-1.076	-0.064	-0.840

ンの音響特徴量を標準化して算出した数値を表2に示す。

ここからセクションごとに各画像とのユークリッド距離を算出し、それから最も適した画像とセクションの組み合わせを選定する。表3に各ユークリッド距離の値を示す。またこれにより決まった楽曲に対する画像の対応順を図6に示す。

表3 各画像と各セクションとのユークリッド距離

画像	sec1	sec2	sec3	sec4	sec5	sec6	sec7
1	2.980	3.404	2.053	3.061	3.560	3.832	3.373
2	2.432	3.108	2.178	2.015	2.689	3.796	3.041
3	2.006	1.809	2.347	2.331	3.156	5.416	2.906
4	2.892	2.177	3.277	3.625	3.985	6.418	3.321
5	1.417	1.886	2.738	0.609	1.613	5.094	2.395
6	1.918	2.068	1.608	2.266	2.877	4.520	2.556
7	2.272	2.127	2.731	2.384	3.384	5.689	3.293



図6 タイスの瞑想曲に対する画像の対応順

4.2 絵画作品と音楽との融合の有用性評価

本提案手法による芸術作品画像と音楽の融合の有用性を評価するため、被検者に対して生成した鑑賞コンテンツを鑑賞した後に、アンケート調査を実施した。まず、事前に大学生10人を被検者とし、鑑賞コンテンツによるデモ動画を5つ鑑賞してもらい、「音楽と芸術作品を鑑賞して適している（違和感がない）」と思いましたがか?という問いに「はい」、「いいえ」、「わからない」で答えてもらった。この中から「はい」を含む割合が80%以上であるコンテンツ2つとそれ以外から1つを選択し、アンケート評価の対象コンテンツとする。

4.2.1 実験概要

本実験では、被検者として大学生の男子9名、女子11名の計20名を対象とした。アンケート内容は以下の通りである。

以下の質問は各デモ動画を鑑賞した後にお答えください。

- Q1: このデモ動画に使われている音楽と芸術作品を鑑賞して適している（違和感がない）と思いましたがか?
- Q2: デモ動画を鑑賞して、楽曲の明るさと芸術作品の雰囲気は適していると思いましたがか?
- Q3: デモ動画を鑑賞して、楽曲のテンポと芸術作品の雰囲気は適していると思いましたがか?
- Q4: デモ動画を鑑賞して、楽曲の強弱と芸術作品の雰囲気は適していると思いましたがか?
- Q5: もう一度音楽だけを鑑賞いただいた後、自身がその楽曲において適切だと感じる順番に作品群を並び替えてください。

以下の質問はすべてのデモ動画を鑑賞した後にお答えください。

- Q6: デモ動画を鑑賞して、芸術作品だけを鑑賞するより、音楽とともに芸術作品を鑑賞する方がよいと思いましたがか?理由があればお答えください。
- Q7: デモ動画を鑑賞して、このようにオンラインで芸術作品を鑑賞したいと思いましたがか?理由があればお答えください。
- Q8: 最後に改善すべき点、感想等あれば教えてください。

鑑賞する前に事前調査として被験者の属性質問を行った。ここでは性別や年齢による感じ方の違いや、普段芸術や音楽に触れる機会の頻度、オンラインでの鑑賞経験や興味によって本提案システムを利用した際の体験の変化に影響を及ぼす可能性を発見するためである。Q1, Q2, Q3, Q4では、事前調査によって選出された3つの鑑賞コンテンツを鑑賞してもらい、音楽と芸術作品の融合が適切か（違和感がないかどうか）を判断してもらい、その後に楽曲の明るさ、テンポ、強弱の観点から芸術作品との雰囲気が適切かどうかを5段階評価で回答してもらった。これは鑑賞コンテンツが違和感なく音楽と絵画作品が融合し鑑賞できるか、楽曲の明るさ、テンポ、強弱といった音響特徴量が鑑賞コンテンツの有用性の要素としてどれくらい重要かを発見するためである。また、Q5では、鑑賞コンテンツの鑑賞後にもう一度音楽だけを鑑賞した後、被験者がその楽曲の展開や強弱に対して、適切だと感じる順番に絵画作品を並び替えてもらう。これにより、分割したセクションごとの音楽と各絵画の組み合わせの適切さを計ることができる。Q6, Q7では、すべての鑑賞コンテンツを鑑賞した後にシステム全体の満足度を5段階評価で回答してもらい、その理由を記述式で回答してもらった。Q8では、改善すべき点や感想などを記述式で回答してもらった。

4.2.2 実験結果

属性調査の結果によると音楽を多く聞く機会がある人は85%（17名）が多いと答え、芸術鑑賞をに行く機会は95%（19名）が少ないと答えた。オンライン鑑賞の経験の有無は75%（15名）がないと答えたが、オンライン鑑賞に前向きに興味を示し

たのは 75%（15 名）だった。

Q1 の選択結果を図 7 に示す。音楽と絵画作品の融合が適していると答えたのはデモ動画 1 では 85%（17 名）、デモ動画 2 では 65%（13 名）デモ動画 3 では 45%（9 名）であった。これらを元に計算した適合率は 0.77 であった。Q2, Q3, Q4 の 5 段階評価による評価値は、デモ動画 1 の平均値は 4 で、デモ動画 2 の平均値は 3.75、デモ動画 3 の平均値は 3.3 であった。その中でも Q3 のテンポに対する質問の評価値はデモ動画 1, 2 では最も高く、デモ動画 3 では最も低かった。

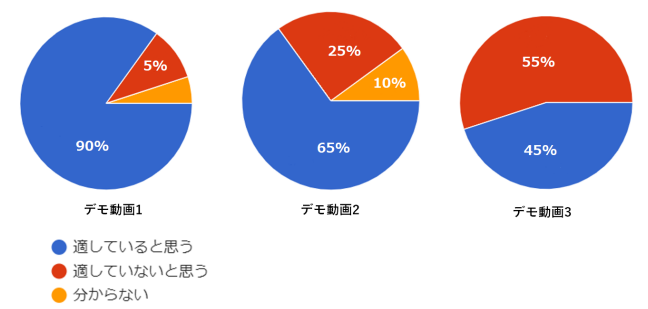


図 7 Q1 の選択結果

また、Q5 の結果をスピアマンの順位相関係数によって算出し、その結果を表 4 に示す。平均値を見ると、特にデモ動画 2 とデモ動画 3 に強い正の相関があることが分かった。Q6 と Q7 の結果を図 8 に示す。Q6 の 5 段階評価の平均値は 4.4 で、その理由は「聴覚からの情報が加わることで作品に奥行きが出るように感じたから」、「芸術作品の雰囲気と音楽がマッチしていると鑑賞が楽しくなるから」という回答が見られた。Q7 の 5 段階評価の平均値は 3.55 で、その理由は「気軽に鑑賞できるから」「わざわざ足を運ばなくても見れるから」という肯定的な意見がある一方、「絵が描かれた背景や作品名作者などの情報が知りたいから」、「音楽とともに芸術作品を鑑賞しても、しなくても変わらないと思ったから」という回答が見られた。

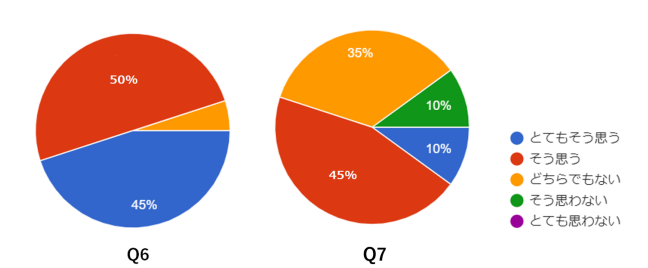


図 8 Q6 と Q7 の結果

4.3 考 察

まず、全体で約 7 割が音楽と芸術作品を鑑賞して適していると答えた結果から、システム全体としてはやや精度を高くすることができたと思われる。またデモ動画とアンケート結果の比較から、全体的に明度が低い作品と、テンポがゆったりとした曲との組み合わせが最も適することが分かった。スピアマンの順位相関係数の結果より、選択画像が多いほど相関係数が低く

表 4 スピアマンの順位相関係数

被験者	デモ動画 1	デモ動画 2	デモ動画 3
A	0.503	0.988	0.721
B	0.721	0.903	0.879
C	0.636	0.794	0.855
D	0.564	0.976	0.733
E	0.358	0.939	回答なし
F	0.503	0.939	0.867
G	0.564	0.770	0.879
H	0.491	0.964	0.842
I	0.418	1.000	0.770
J	0.588	0.952	0.770
K	0.685	0.939	0.867
L	0.661	0.867	0.733
M	0.636	0.867	0.733
N	0.661	0.952	0.915
O	0.564	0.770	0.964
P	0.661	0.794	0.879
Q	0.661	0.842	0.648
R	0.685	0.794	0.867
S	0.661	0.794	0.636
T	0.661	0.806	0.624
平均値	0.594	0.878	0.799

なることが明らかになった。さらに楽曲展開に適した組み合わせを実現するためには、楽曲の構成や展開に応じた選択画像数に調整したり、色情報や音響特徴量をさらに増やし、適正調査を重ねる必要がある。

最終的に芸術作品だけを鑑賞するより、音楽とともに芸術作品を鑑賞の方がよいと答えたのは 95%（19 名）であり、前向きな結果が得られた。しかしながら、このようにオンラインで芸術作品を鑑賞したいかと答えたのは 55%（11 名）であった。この結果より、ただ音楽と芸術作品を共に鑑賞するだけではまだまだ体験価値や興味関心が低いということが分かった。これを踏まえ、鑑賞体験を向上させるための機能や、さらなるアプローチが必要であると感じた。

5 おわりに

本研究では、絵画作品の画像と音楽の特徴量抽出に基づき、異なる絵画作品の感情や雰囲気に対応させた音楽を融合した鑑賞コンテンツを生成するシステムを提案した。その結果、従来のオンライン鑑賞より、音楽と融合した鑑賞コンテンツの方が良いという結果が得られた。

今後の課題としては、今回の提案手法ではまだ特徴量が不足や吟味の必要性があったり、オンライン鑑賞の体験価値を向上するにはまた違ったアプローチを加える必要があることがわかった。また、今回は絵画作品に焦点を当てたシステムであったが、彫刻や工芸品、書道など多様な芸術作品に対してもオンライン鑑賞の体験価値を深められるものにしていきたい。さらに、今回の研究を踏まえて、リアルでもオンラインでも鑑賞体験や学習を深められるよう、知見を深め改善をしていきたい。

文 献

- [1] 文化庁: 文化に関する世論調査(令和3年度調査), 2022年3月.
- [2] Shoko Usui, Chen He, Katsumi Sato, and Yoshifumi Kitamura: A Dynamic, Flexible & Interactive Display Method of Paintings for Communicative Art Appreciation among Students, in Proceedings of the SIGGRAPH Asia 2015 Symposium on Education (SA '15), Article No.: 3, pp. 1-8, 2015.
- [3] Jocelyn Spence, Benjamin Bedwell, Michelle Coleman, Steve Benford, Boriana N. Koleva, Matt Adams, and Ju Row Farr: Seeing with New Eyes: Designing for In-the-Wild Museum Gifting, in Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19), pp. 1-13, 2019.
- [4] Reese Muntean, Alissa N. Antle, Brendan Matkin, Kate Hennessy, and Jordan Wilson: Designing Cultural Values into Interaction, in Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17), pp. 6062-6074, 2017.
- [5] Shoto Sasaki, Tatsunori Hirai, Hayato Ohya, and Shigeo Morishima: Affective Music Recommendation System Reflecting the Mood of Input Image, in Proceedings of the 2013 International Conference on Culture and Computing, pp. 153-154, 2013.
- [6] 福本麻子, 塚田浩二, 蔡東生, 安村通晃: Ziplayer: 絵画, 文章, 音楽を組み合わせたメディアプレイヤーの提案, 情報処理学会シンポジウム論文集(情報処理学会ワークショップ論文集), Vol. 2006, No. 4, pp. 79-80, 2006年3月.
- [7] Taemin Lee, Hyunki Lim, Dae-Won Kim, Sunkyu Hwang, and Kyunghyun Yoon: System for matching paintings with music based on emotions, in Proceedings of the SIGGRAPH ASIA 2016 Technical Briefs (SA '16), Article No.: 31, pp. 1-4, 2016.
- [8] C. Hutto, and Eric Gilbert: VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text, in Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media (ICWSM '14), Vol. 8, No. 1, pp. 216-225, 2014.
- [9] 中野美由紀, 石山悟志, 君野史明, 小久保勇気, 斉藤大介, 鈴木智尚, 宮崎晃一, 吉村慎祐: 利用者の印象に基づく音楽レコメンドサービス: 画像を利用した利用者指向の音楽レコメンド手法の考察, 情報処理学会第81回全国大会講演論文集, Vol. 2019, No. 1, pp. 383-384, 2019年2月.
- [10] 安達太郎, 岩宮真一郎: 色彩と音楽とが互いに及ぼす影響-ショパンのエチュードを手がかりに, 第5回学生のための研究発表会講演論文集(日本音響学会九州支部), pp. 13-16, 2003年.
- [11] 星野喜久三: 色とテンポの関係に関する研究, 北海道学藝大學紀要, 第一部, Vol. 10, No. 1, pp. 233-239, 1959年7月.