

歌詞や曲調の特徴に基づいた ステージ照明演出自動生成システムの実現

神野満里奈^{†1} 福原義久^{†2}

本研究では、照明色とその配置、および照明の動きを制御し、ステージの感動をより効果的に演出するための照明演出システムを実現した。具体的には1) Attentionとカラーイメージスケールを用いた照明色の選択、2) 曲調に基づく照明色の配置と明るさの決定、3) 音量の変化に基づく照明の動きの制御の3つの部分から構成される。提案システムは我々のこれまでの研究から、歌詞の解釈方法の改善や照明の動きの追加などのアップデートをおこない、よりステージと聴衆の一体感の向上を目指した。

Realization of an automatic stage lighting production generation system based on lyrics and music tune.

MARINA KANNO^{†1} YOSHIHISA FUKUHARA^{†2}

This research realized a lighting production system that controls lighting colors, their placement, and lighting movements to produce more effective stage impressions. Specifically, the system consists of three parts: 1) selection of lighting colors using the Attention and Color Image Scale, 2) determination of lighting color placement and brightness based on the music tone, and 3) control of lighting movement based on changes in sound volume. Based on our previous research, we have updated the proposed system by improving the interpretation method of lyrics and adding lighting movements, aiming to further improve the sense of unity between the stage and the audience.

1. はじめに

舞台照明とは演劇、ダンス、オペラ、その他のパフォーマンスアートの制作に適用される照明の技術のことであり、初期のギリシャの劇場から取り入れられ始めた。その光源は、日光、キャンドル、石油ランプ、ガスと移り変わり、1800年代後半には白熱電球が生み出されたことによって多くの照明が電気照明に変わった。しかし、この光源は画期的ではあったが、新しい照明方法などは開発されず、従来の脚光、線形照明、ボーダーライトなどが電気照明に置き換えられただけであった。その後、より明るく大きな照明が開発され、照明デザイナーは、さまざまな色や焦点を試すことで、複雑な照明計画を立てられるようになった[1]。また、照明コンソールやDMXプロトコルが開発されたことにより、照明演出はオン・オフだけでなく、より複雑なものになっていった。例えば、舞台進行に伴う照明情報の保存が可能になったことで、数多くの照明を同時に操作することが可能となった。そして、2000年代にはLEDライトの普及により、これまで使用していた手動で切り替える板状のカラーフィルターを使わずに光の色を瞬時に変更することが可能になった[2]。近年ではPCからの制御も容易になり、舞台だけではなく、家庭用としても照明制御の応用は裾野を広げている。

例えば、ゲームの内容や状況に応じて、室内の照明を動

的に変化させるといった技術も普及している。これらの技術は、新しい体験をもたらしてはいるものの、照明がもたらす印象については事前に人間が定義した簡単なルールに従っている。例えば、被弾すると赤い照明が点灯する、といった具合である。しかし楽曲のような、より複雑なニュアンスを伝え、感動や一体感を高めていくような目的のためには、楽曲の内容そのものを解釈し照明演出を決定していくような仕組みが必要である。

一方、実際のライブやコンサートなどの舞台照明の設定では、専門の技術者や照明アーティストが照明演出を決定しているわけであるが、今やVR空間上でもステージを体験でき、安価な照明機材の普及により家庭でも制御さえ可能であれば、それなりの照明演出を実現することができるようになってきている。そのような現状を鑑み、我々はプロのアーティストだけではなく、アマチュアや家庭、あるいはVtuberといったさまざまな利用者が誰でも簡単に照明演出を体験するには、ある程度の自動化が必要であると考えた。

では、どのように楽曲に合わせて照明演出を制御するかであるが、Oksanenは視覚と聴覚の情報を組み合わせて観客に語りかけることで、共感的な結果をもたらすことができる[3]と述べており、この考えに基づく楽曲の持つ印象をいかに正しく照明演出に反映させるかが課題である

^{†1} 武蔵野大学
Musashino University,
^{†2} 武蔵野大学、アジア AI 研究所

Asia AI Institute, Musashino University

ことがわかる。また、聴覚情報である楽曲と視覚情報である色に関する研究として、言語のイメージと配色の関連性を分析し、その結果を利用した配色支援システムを提案する研究[4]や、楽曲に相応しい色彩の組み合わせについての研究[5][6]、などが行われてきた。演劇作品においては俳優の発話感情で照明色を変化させる研究[7]も存在する。このように視覚と聴覚の情報を組み合わせて、共感的な結果をもたらす研究は多く行われてきたが、楽曲と色に関する研究において、これらの成果を舞台照明に応用することは行われていなかった。

そこで我々は、曲調や歌詞がもたらす印象をできる限り忠実に照明演出に反映させ、聴覚と視覚どちらからも同じような印象をオーディエンスに伝え共感に働きかける手法について研究している。

我々のこれまでの研究では歌詞中の単語全てで調色を行っていたため、歌詞中の他の単語との相対評価を行っていなかったが、本論文では **Attention** を用いることで歌詞の中での重要ワードを取り出し、取り出された重要ワードで調色を行ったためより歌詞の印象を忠実に照明色に反映することができた。また、照明自体に動きを動かすことで、曲に合わせた静的・動的な印象を提供することができるようになった。

具体的には、照明の調光を歌詞、調色および配置・動きを曲調の特徴を用いて変化させることで、曲のストーリーを考慮した自然な照明の移り変わりを実現した。

照明器具の高機能化と低価格化、VR コンテンツの普及などを経てリアル / バーチャルを問わず個人でも本格的な照明演出を楽しめるようになった現在、フルカラーの調光・調色装置を用いた自宅でのカラオケや音楽鑑賞、あるいは VR 空間内などにおけるライブやコンサートが提案手法によってより本格的に楽しめるようになるだろう。

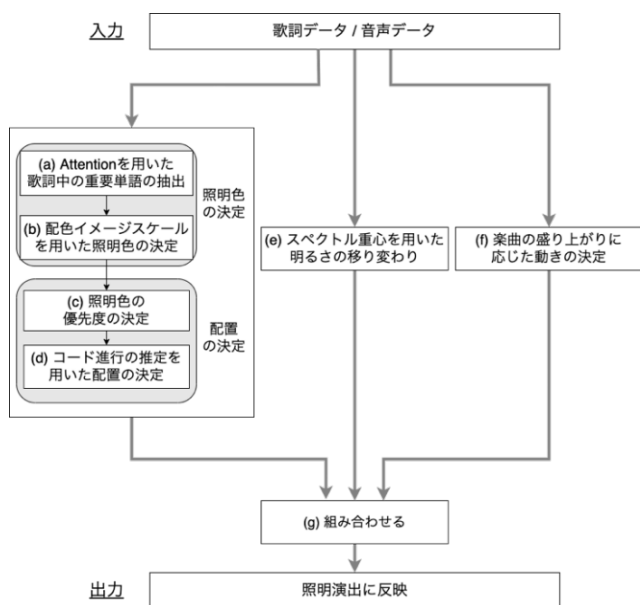


図1 照明制御の全体像

2. 曲に合わせた調光・調色・配置と動きの決定

本研究では照明演出を構成する要素の中で、照明自体の明るさを決定する「調光」、照明色を決定する「調色」、照明色の並びを決める「配置」、照明自体を動かす「動き」の4要素を楽曲の特徴を用いて決定する。

2.1 提案手法の概要

提案手法は、照明色とその配置の決定、明るさ、動きの設定の3つのプロセスを組み合わせることで実現している(図1)。本研究では楽曲を構成する要素である歌詞と曲調をどちらも照明演出の決定に組み込むことで、楽曲の持つ印象をより忠実に再現することを目指す。具体的には、楽曲の歌詞を用いて、カラーイメージスケールに示された色から適切な照明色を選択し、曲調を用いて照明色の配置・明るさ・動きを決定しライトに出力する。次にそれぞれのプロセスについて詳しく述べる。

2.2 照明色の選択

楽曲と色彩を結びつける研究はこれまでも多く行われてきた[5][6][8][9]。色彩の決定方法は楽曲の曲調を用いる場合、歌詞の特徴を用いる場合とあるが、本研究では歌詞の持つストーリー性や印象を照明演出に反映させるために、歌詞の特徴とカラーイメージスケールを用いた調色を行う。このとき、歌詞の文脈の繋がりを考慮するために、歌詞に出てくる単語全てでは調色を行わず、歌詞のなかでも重要な要素となっている単語のみを **Attention** を用いて抽出し、抽出された単語を使用して調色を行う。(図1の(a,b))

2.2.1 Attention を用いた歌詞中の重要単語の抽出

人間は文章を見聞きするとき、その文の中で重要な要素となる部分が無意識に抽出している。これを深層学習で数値化したものが、**Attention** と呼ばれる入力されたデータのどこに注目すべきかを、動的に特定する仕組みである。2017年に発表された Vaswani らによる "Attention Is All You Need" [10] の中で初めて登場した深層学習モデルの Transformer は、**Attention** のみを使用するニューラル翻訳モデルであり、わずかなトレーニングで LSTM よりも優れた結果を得ている。**Attention** とは、系列データを扱う際に重要な点に注意(**Attention**)を向ける手法であり、加法的注意、内積注意、自己注意がある。その中でも Self-Attention は汎用性が高く効果的であり、言語処理のすべてのタスクで高い性能を発揮することが知られている。

本研究では、Transformer を1つのユニットとして使用している BERT モデルを用いて **Attention** の値を算出

し、人間がその楽曲の歌詞の中で重要視している箇所を抽出。それに沿った調色を行う。

Attention をもちいて重要なワードを抽出するには、歌詞をいくつかのジャンルごとに分類しておく必要がある。我々はアーティストが作詞を行う際その楽曲のテーマを決定することにならない、今回は作詞の中でテーマとして設定されることが多い「恋愛」「青春」「応援」「悲しみ」「ストーリー」の5つを歌詞ジャンルとして設定し、分類を行った。

以上の流れをまとめると提案手法は以下ようになる

- i) 既存の楽曲の歌詞を 100 曲選び、「恋愛」「青春」「応援」「悲しい」「ストーリー」の5つに分類にアノテーションする。アノテーション済みの楽曲データを日本語事前学習済みBERTモデル[11]に学習させる。
- ii) 調色を行う楽曲の歌詞を、歌詞カードを参考に文節ごとに分割する。
- iii) 分割した歌詞の歌詞ジャンルと、歌詞ジャンルの決定において判断根拠となった各歌詞単語の Attention を求め、Attention の値が高い上位 50%の単語を取り出す。
- iv) 取り出された単語の中で、文の構成において重要となる要素(名詞, 動詞, 形容詞)のみを抽出し、楽曲の重要ワードとする(図2)。

なお、Attention の算出にはBERT[12]、重要ワードの抽出にはNLTK[13]を用いた。

表 1 配色イメージスケールの例

配色イメージスケール内の単語	定義されている5色
まぶしい	
さわやかな	
華麗な	
健康的な	
くつろいだ	

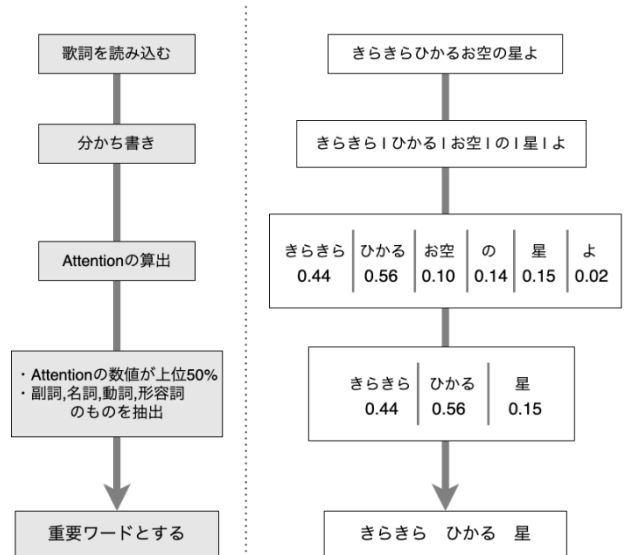


図 2 Attention を用いた重要度の高い歌詞単語の抽出

2.2.2 配色イメージスケールを用いた照明色の決定

次に、重要ワードから適切な色を求めるため、本研究では小林らによる5色配色イメージ・スケール [14][15]を用いた。配色イメージ・スケールとは、イメージの共通感覚を心理学的研究の蓄積に基づき360のイメージ語を5色の配色で表現したものである。

ただし文化の違いやアーティストの感性を重視する場合は、この配色イメージスケールは適切ではないかもしれない。その場合は、言語・文化に基づいた配色イメージスケールやアーティストの定義した配色イメージスケールを事前に作成しておけばよいだろう。

提案手法では、以下の手順で歌詞に適合する配色を取得した。

- i) 2.2.1 で得られた重要ワードごとに分散表現ベクトルを取得する。
- ii) 配色イメージ・スケールの中から、歌詞単語のベクトルとコサイン類似度がもっとも近いものを選択する。配色イメージ・スケールでは、1つの単語につき5つの色を使った配色がされているため表1に示されるような5つの照明色の候補が得られる。例えば「笑顔」という単語は、配色イメージ・スケール内で最も類似度が近い「さわやかな」(浅青緑, 白, サルビア, 白群, 空色)に分類される。

なお、分散表現ベクトルの取得にはword2vec[16]で学習済みの「日本語 Wikipedia エンティティベクトル」[17]を使用した。また配色イメージ・スケールで用いられる単語の中で、学習済みモデルに存在しない単語は除外した。

2.3 照明色の配置

舞台照明では、複数の照明に対して、さまざまな色を割り当てる必要がある。その際、すべての照明は同じ色かもしれないし、ばらばらの色の場合もある。また、グラデーションに配色されている場合もあれば、そうでない場合もあるだろう。このように複数照明がある場合に、個々の照明色を決定するには、使用する色とその配置について考慮しなければならない。本研究では2.2で決定した5色の配色の中から、曲調を基に照明色の配置を決定する(図1(c,d)).

2.3.1 照明色の優先度の決定

曲調は楽曲の印象を決定するための重要な要素の一つである。楽曲に相応しい色彩の組み合わせについて、これまで多くの研究が行われてきた[5][6]。本研究では先行研究に倣い照明色の優先度の決定に BPM(Beat Per Minutes)を利用した。BPMとは楽曲の1分間の拍数を表す数値である。BPMの値が大きいほど、テンポが速い楽曲、BPMの値が小さいほど、テンポが遅い楽曲であると言える。先行研究より、暖色系と速いテンポ、寒色系及び中性色系と遅いテンポ、紫と最も遅いテンポは共通の印象を持つことがわかっている。そのため本研究では、各楽曲のBPMを抽出しそれらに相応しい色を定義した(表2)。

算出したBPMを元に前のステップで選出した5色の配色の中から、表2に定義した相応しい色に類似している色を、CIE DE2000[18]を用いて色差の距離を求めることで決定し、距離が最も近いものから優先度が最も高い照明色とし、中心に配置するものとする

2.3.2 照明色の配置の決定

色の配置の中には、いきいきとした印象を与える色相差のコントラスト配色と、落ち着いた印象を与える色相差のグラデーション配色(表3)がある。舞台照明では曲の印象に合わせてコントラストまたはグラデーションになるように、照明を舞台の中心からシンメトリーに配置する事が多い[19]。本研究では、楽曲のコード進行の特徴から、その楽曲の印象に合致する照明色の配置を決定する。

楽曲におけるコード進行とは、コードすなわち和音が楽曲の中でどのように進んでいくかを示したものである。コードには大きく分けて「メジャーコード」「マイナーコード」の2種類に分類することができる。それぞれのコードには特徴があり、「メジャーコード」は聴講者に明るい印象を与え、反対に「マイナーコード」は暗い印象を与えられている[20]。そのためコード進行において、どちらのコードが主に使われているかがその楽曲の印象を決める重要な要素になる。本研究では、コード進行を照明色の配置を決定するための要素とすること

で、オーディエンスが楽曲の印象と照明の印象が一致していると感じる演出を実現する。

提案手法では、以下の手順で配置を決定した。

- i) 楽曲の音声データを読み込み、その楽曲のクロマグラムを算出する。
- ii) クロマグラムのピッチクラスに対応するメジャーコード・マイナーコードそれぞれの辞書を作成する。
- iii) 毎秒ごとにクロマ特徴を求め降順にソートし、値の大きい3つのクロマ特徴に対応するインデックスを求める
- iv) 値の大きい3つのクロマ特徴に対応する音名の集合を求め、それがメジャーコード・マイナーコードどちらかに該当する場合それを表記し、楽曲のコード進行を推定する。
- v) 楽曲内のコード進行が主にメジャーコードでされている場合は2.3.1で決定した最も優先度の高い色を中心に2.2で決定した照明色をコントラスト配色で配置する。反対に、コード進行が主にマイナーコードでされている場合は2.3.1で決定した最も優先度の高い色を中心に2.2で決定した照明色をグラデーション配色に配置する。

上記の手順を「きらきら星」でおこなった場合の例を示す(表4)。きらきら星はBPM76であるので、相応しい色は青であることがわかる。また、コード進行はメジャーコードが多く使われているため、コントラスト配色に照明色を配置する。

表2 各音楽の特徴に相応しい色

BPM	相応しい色
~60	紫
61~80	青
81~100	青緑
101~120	緑
121~140	橙
141~	赤

表3 配置の例









配色名	実際の配色
色相差の コントラスト配色	
色相差の グラデーション配色	

表4 「きらきら星」の照明

歌 詞	照明の色
きらきら	
ひかる お空の	
星よ	
瞬きしては	
みんなを	
みてる	

2.4 スペクトル重心を用いた明るさの移り変わり

照明の調光・調色に歌詞だけでなく曲調も考慮するために、本研究では楽曲のスペクトル重心を利用した。スペクトル重心とは音色の明るさを表したものであり、重心の値が大きければ、”明るい”音色をもつことに対応する[21]。本研究では、使用するライトの明るさ50%を基準とし、スペクトル重心が0.1変わるごとに明るさを10%変えることで、楽曲の音色に対応した照明の明るさを実現した（図1の(e)）。

スペクトル重心の抽出には LibROSA[22]を用いた。

実際に「きらきら星」でスペクトル重心を抽出したものを図3に示す。

2.5 楽曲の盛り上がりに応じた動きの決定

照明演出ではライト自体に動きを持たせることが多い。本研究ではこうしたライトの動きを楽曲の盛り上がりに応じて決定する(図1の(f))。

楽曲の盛り上がりを決定するための指標は多くあるが、その中でも音量の変化を時系列で表す、音量の実効値である音量RMSを用いるのは有効的な手段とされる[23]。本研究では音量RMSの最小値を0、最大値を1とし、音量RMSが0.5未満の時はライトは静止、音量RMSが0.5以上の時はライトは135°動くものとする。複数のライトを使用する場合、左右に振り動かすようにする（図4,5）。

音量RMSの抽出には LibROSA[22]を用いた。

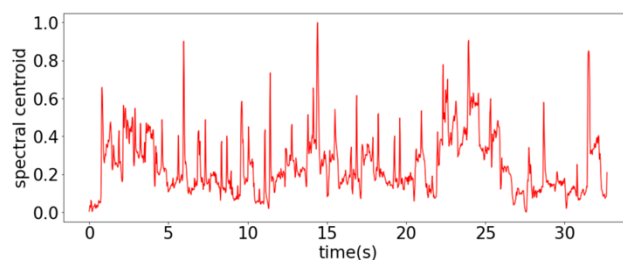


図3 「きらきら星」のスペクトル重心

2.6 歌詞と曲調による調光・調色・動きの決定

上記の5つの結果より、単語ごとの5色の照明色、曲調を用いた色の優先度と配置に従い、単語ごとの5色の照明色をシンメトリーに並び、スペクトル重心による明るさの決定、音量RMSによる動きの決定をライトに出力する。

3. 実験結果

上記の実験方法に従って舞台照明のシミュレーションした。今回は舞台照明シミュレーションソフトを使用した(図6)。例として童謡「きらきら星」での実験結果を表6に示す。シミュレーションには Capture Visualisation AB社が開発したライティングデザイン、ビジュアルライゼーション、ドキュメンテーションソフトウェアである「Capture 2022」を利用した。

実験結果の動画はウェブサイト[24]に掲載。



図4 ライト静止時



図5 ライトを振り動かした時









図6 ステージを模したもの

4. 評価方法

提案システムを実際に舞台照明として使用した際の印象を評価するために、10代から50代の男女計26人の協力を得て、楽曲と舞台照明の印象の一致度についてアンケートを取った。今回は提案システムと楽曲の一致度を図るため、BPM186とアップテンポな楽曲Aと、BPM60とスローテンポな楽曲Bを用意し、提案手法による照明効果を適用した動画を視聴してもらった。評価方法は予め設定した4つの設問に10段階での評価を行うものである。設問内容はQ1「楽曲の印象と照明の色は一致していたか」、Q2「楽曲の印象と照明色の配置（色の並び）は一致していたか」、Q3「楽曲の印象とライトの動きは一致していたか」、Q4「楽曲の印象と全体的な照明演出は一致していたか」とした。また、照明演出と楽曲の「ここが特に楽曲と一致していた」と感じた箇所、「違和感がある」と感じた箇所があった場合はそれらを記述する形で評価を行った。

表6 「きらきら星」の照明

歌 詞	照明の色
きらきら	
ひかる お空の	
星よ	
瞬きしては	
みんなを	
みてる	

実験結果の動画は下記リンク先

・きらきら星【<https://youtu.be/95DnNWLOGfQ>】

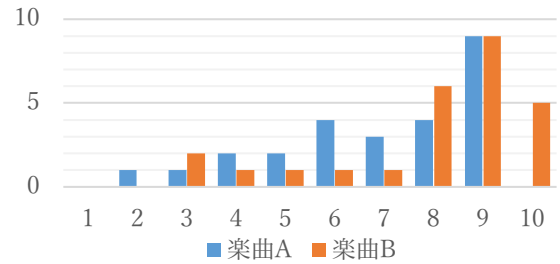


図7 Q1

「楽曲の印象と照明の色は一致していたか」

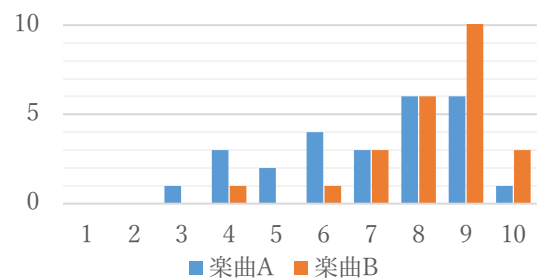


図8 Q2

「楽曲の印象と照明色の配置（色の並び）は一致していたか」

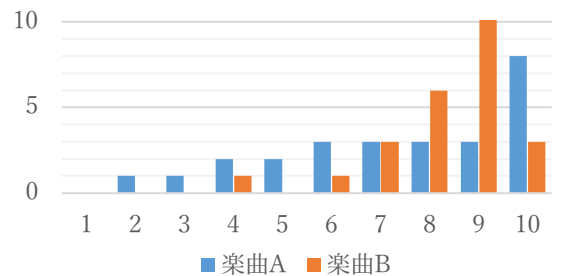


図9 Q3

「楽曲の印象とライトの動きは一致していたか」

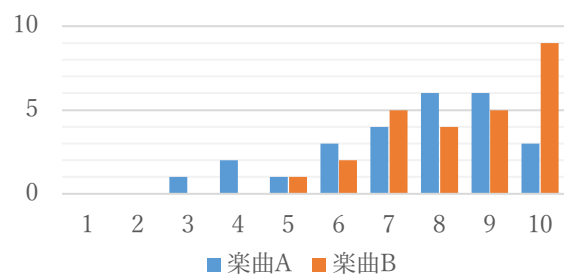


図10 Q4

「楽曲の印象と全体的な照明演出は一致していたか」

5. 評価結果

上記の手順通りに評価を行った結果は図 7,8,9,10 のようになる。

楽曲 A, 楽曲 B どちらも本手法通りに決定した照明色, 配置, 動き, 全体的な照明演出は, 楽曲の印象に一致するという意見が多かった。このことから本研究での手法で提案された照明演出は, 多くの人に楽曲と同じような印象を与えることがわかった。

一方で楽曲と照明演出が一致しないと答える人も一定数いた。楽曲 A における Q1, Q2 の回答は特に意見にバラつきがあった。照明演出と楽曲の「違和感がある」と感じた箇所の記述の中にも, 照明色や配置に関するものがいくつ書かれていた。理由としては, 楽曲 A では提案される照明色は寒色系が多かったが, BPM 値が 186 と高かったため, 曲調から定義した優先色(表 2)は赤であった。この歌詞から提案される色と, 曲調から提案される色の差が違和感に繋がったのではないかと考えられる。より多くのオーディエンスが, 楽曲と舞台照明の印象の一致を感じられるよう改善が必要と考えられる。

6. まとめ

実験結果より, 楽曲の世界観に没入することができるような新しい舞台照明演出を, 自然言語処理や曲調抽出により実現することができた。

これにより, 近年安価に入手可能となったフルカラーの調光・調色装置を用いて自宅でのカラオケや音楽鑑賞, あるいは VR 空間内のライブやコンサートなどにおいて, 誰でもが本格的な照明演出を実現することができるようになることが期待される。

7. 展望

本研究ではあらかじめ読み込ませた楽曲に合わせて照明演出を決定するにとどまっている。今後は舞台上の演者の動きに合わせてインタラクティブに動かすといったことに挑戦したい。これにより, 曲の持つ感動をより深く, 効果的に演出することができるようになるだろう。

今回は評価におけるアンケート対象を, オーディエンスのみとしたが, 今後は実際のライブやコンサートでの有用性を図るため, プロの照明技術者や音楽アーティストにも参加を呼びかけ, 様々な角度からの評価・検証をおこないたい。

また, ある会場でのパフォーマンスを VR 上や別の部屋などに, 照明演出を含めて同時に提供することで, リアルやバーチャルを超えて同じ体験・感動を共有するといったことに挑戦したい。

参考文献

1. Susan Brady, Nena Couch, “A Primer for the History of Lighting Design”, Documenting: Lighting Design (pp.1-17)
2. Saki Ogawa, Masafumi Hagiwara, “Color Combination Support System using Image Words”, International Journal of Affective Engineering, Vol.15 No.2, pp.287-296, 2016.
3. Turk, M. “Multimodal interaction: A review.” In Pattern Recognition Letters 36. 189-195. (2014)
4. Julle Oksanen, “DESIGN CONCEPTS IN ARCHITECTURAL OUTDOOR LIGHTING DESIGN Based on Metaphors as a Heuristic Tool” , 2017
5. Kikuzo Hoshino, “A Study on the Relationship between Color and Tempo”, 1959
6. Aya Yajima , You Kaku , Miho Saito ,” Psychological effects of combination of music and color
7. N. Vryzas, A. Liatsou, R. Kotsakis, C. Dimoulas, and G. Kalliris. 2017. “Augmenting Drama: A Speech Emotion - Controlled Stage Lighting Framework.” ,2017
8. T. Kitagawa, T. Nakanishi, Y. Kiyoki, “An Implementation Method of Automatic Metadata Extraction Method for Image Data and Its Application to a Semantic Associative Search”
9. S. Murata, R. Okada, F. Sasaki, and T. Nakanishi, “Time Series Sensory Color Representation Method for Songs Focusing on Words in Lyrics”, 13th Forum on Data Engineering and Information Management, DEIM Forum 2021 D25-4, 2021.
10. Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, Illia Polosukhin “Attention is all you need.” arXiv preprint arXiv:1706.03762.
11. 東北大学 乾研究所, “日本語 BERT 訓練済みモデル” <https://github.com/cl-tohoku/bert-japanese> , 2019
12. Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova, “Pre-training of Deep Bidirectional Transformers for Language Understanding”, 2018
13. NLTK <http://www.nltk.org/>
14. Shigenobu Kobayashi・NIPPON COLOR & DESIGN RESEARCH INSTITUTE INC., 『Art of Color Combinations』, Kodansha,1995.
15. Shigenobu Kobayashi, COLOR IMAGE SCALE, Oxford University Press, 1992
16. Mikolov, Tomas *et al.* “Efficient Estimation of Word Representations in Vector Space”, 2013, ICLR
17. 日本語 Wikipedia エンティティベクトル. http://www.cl.ecei.tohoku.ac.jp/~m-suzuki/jawiki_vector/
18. Luo, M.R., G. Cui, and B. Rigg, “The development of the CIE2000 color-difference formula CIEDE2000,” Color Research & Application, 26, pp. 340-350 (2001).
19. Naoki Fuzii 『Introduction to Stage and Stage Lighting: From Stage Basics to DMX and Moving』, 2016

20. Norman D. Cook "A Psychophysical Explanation for Why Major Chords are 'Bright' and Minor Chords are 'Dark'"
21. Masashi Yamada "Human Communication through the Music"
22. LibROSA <https://librosa.org/doc/latest/index.html>
23. Kenichi Koyama, Hiromi Ishizaki, Keiichiro Hozoku, Chihiro Ono, Jiro Katoh, "A Study on Feature Extraction for Highlights Detection from Musical Performance Videos",2011
24. 神野満里奈の YouTube,
<https://www.youtube.com/channel/UCLsECF3H9iuIGexUSoYZhtQ>