

機械学習による手書きシンボルの判別を利用した 照明制御信号操作デバイスの提案と実装

宮川一歩[†] 赤池英夫[†]

概要: 本研究では、照明機材の操作などにおけるリアルタイム制御の補助を目的として、機械学習による手書きシンボルの判別を制御パターンの判別手法として採用した、小型制御用デバイスについて提案を行う。本研究の目的は、照明機材などの大規模な機材を操作する場合などを主として、視覚的自由度の少ない状況にあるユーザが、直感的に制御パターンの呼び出しを行えるようなデバイスおよびそれに適したシンボルを実装し、リアルタイム性を確保しつつ、ユーザが照明等の操作を行いやすい環境を提供することである。

本稿では、研究内容及び今後の課題について報告し、機械学習による画像分類をハードウェアの入力手法として取り入れることの有用性を検討する。

1. はじめに

近年、インターネット上を主として活動する、仮想的なアバターを利用した活動が広まっており、それに伴った、3D アバターのリアルタイム投影及び音源・機材の同期を行う形でのライブ演目が増加している。

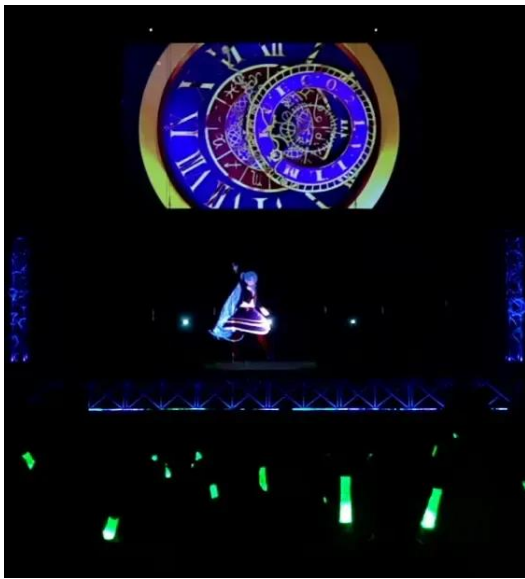


図 1 ヴァーチャルシンガーによる
ライブ演目の一例[1].

このような需要の増加に伴い、会場の動向に応じた、即時的な機材操作の必要性が徐々に生じてきている。しかし、ライブ会場のような大規模な機材を同時に操作する現場では、各機材に対応したスライダによりパラメータを逐一操作し、特定のパターンを再現しなければならないことから、機材と会場の双方の同期を取るような演目には向いていない。

zero88



図 2 照明機材卓の例

(<https://rizing.co.jp/?pid=166825351> より引用)

リアルタイムでの機材制御手法の一例として、入力されたユーザのジェスチャーに応じて照明パターンを切り替える手法などが検討された例がある[2]。この手法では、加速度センサーにより人間の手の動きをジェスチャーとして入力し、それに伴って照明を制御するといった手法となっている。しかし、これらの手法の課題点として、人間の動作を入力としていることから、環境により精度が大きく左右される問題や、加速度センサー等の設置による場所の占有、ジェスチャーを行ってから制御が行われるまでにある程度のインターバルが発生する問題点などが存在している。特に今回の目的の一つである、ライブ会場等における照明機材の操作では、狭い領域での機材操作を求められることから、体の動作を利用した入力は難しく、より小さな領域内におけるジェスチャーに近い形でのパターン判別が必要となる。

近年機械学習技術の向上に伴い、小型デバイス上における機械学習を用いたプログラムの実行、および画像の判別が可能となりつつあり、この手法をシステムにおけるパターンの判別手法として取り入れることで、手元領域のみで、ジェスチャー操作に近い手法による機材の操作を実現でき

ると考えられる。

そこで本研究では、デバイス上の操作位置に依存しない、機械学習による手書きシンボルの判別を入力手法として取り入れた、デバイス上で入力されたシンボルに基づいて**照明制御信号**の操作を行うデバイスを提案し、ユーザ操作に対する本デバイスの有用性を検討する。

2. 関連研究

本研究では、機械学習を用いた画像判別による照明制御信号の操作デバイスを提案する。

本節では、関連研究として、従来の手書き判別手法に関する研究及び、特定の操作を外部入力により制御する手法に関する研究を紹介し、本研究との関連性について述べる。

2.1 位置に依存しない手書き点字入力方式

Sana Shokat ら[3]は、タッチスクリーンベースでの、入力位置に依存しない点字方式での文字入力手法を提案し、深層学習 (DL)ベースによるパターン認識の有用性について言及した。

実験の結果から、入力シンボルの組み合わせについて重複がないように適切に指定し、DL 方式による判別を行うことで、比較的高精度に入力内容を判別することができ、また、入力位置に依存しない入力方式は、視覚的に不自由な状況にあるユーザに対して、有用な方式であることが示唆された。

2.2 無線 LAN の信号による手書き文字の判別

Zhengxin Guo ら[4]は、無線 LAN アクセスポイント間の人間の動作を無線 LAN の信号の変化として受け取り、DL を利用したデータ分類により、人間のジェスチャーに対応した手書き文字の判別を行う手法を提案した。

研究結果から、DL における LSTM などの手法を組み合わせることにより、実験環境として示された限定的な状況においては、高精度に認識可能であることが示された。

また、この手法における問題点として、無線 LAN 信号が周囲環境の影響を受けることによる判別精度の低下、形状が似ている文字同士の混同などの課題点が示された。

この課題点に関しては、シンボル形状の調整や、デバイスに直接書き込みを行う方式を取ることで、回避可能であると予想される。

2.3 マイコン上における DL による画像の分類実行

Mitulgiri ら[5]は Raspberry Pi 上において、画像分類システムを実現し、マイコンボード上でも画像分類システムが実行可能であることが明らかになった。

この研究では、Raspberry Pi 上のシステムと、クラウド上に存在している機械学習サービスをリンクすることにより

画像分類を実現しており、デバイス上に接続されたカメラを通して入力された画像を、クラウド上で分類することにより、マイコンボード上から分類結果を利用可能であることが示唆された。

2.4 本研究との関連性

2.1 節より、タッチスクリーン上において、入力位置に依存しない入力方式を適用することにより、視覚的に不自由な状態のユーザが、利用しやすい入力システムを実現することが可能であることが示唆されている。これは、本研究の目的の一つである、ごく限られた空間内で機材の操作を行っているような場合にも適用することができ、操作時に入力位置を気にすることなくシームレスに機材の制御を行うことができると考えられる。

2.2 節より、手書きのシンボルを入力として利用することの利点と課題点が示された。この内容から、人間の動きをそのまま入力として利用することにより、遠隔で機材を操作する際に、ユーザの望んでいる動作をそのまま機材に伝達することができるという利点が存在していることが示された。このことから、今回のような、制御用のスペースから全体の機材を制御する場合においては、有効な手段であると考えられる。

また、Zhengxin Guo らの研究より、体の動き全体をジェスチャーとして利用する場合、可動スペースの確保や、周囲の人の動きによる影響などを受けやすいといった問題点も示された。このことから、体全体の動きから入力内容を推測するのではなく、手元などの限られた部位の動きのみをジェスチャーとして利用することにより、安定した分類精度を得られることが期待できる。

また、2.3 節において、マイコンボード上において機械学習による画像分類が実行可能であることが示され、これを入力判別方法として採用することにより、入力位置に依存しない入力手法の実現が期待できる。

3. 研究内容

本研究では、既存の研究では主に、入力内容の反映などのために用いられてきた、DL による手書き文字・記号の判別を、機材の制御パターンを選択に用いることで、より直感性の高いシステムを実現することを目指している。

本節では、実際の研究内容について示し、本システムの利点と課題点を明らかにする。

3.1 本研究の目的

本研究の目的は、リアルタイムでの操作が必要とされる機材の操作に対して、機械学習を用いた手書きシンボルの判別による制御デバイスを提案し、制御信号の選択に手書き判別の処理を適用することによる、ユーザに与える操作

感の影響を調査し、その有用性を明らかにすることである。

また、手書きシンボルの判別処理においても、誤検出が少なくなるようなシンボルの形状について検証し、シンプルかつ他のパターンと衝突しないようなシステム設計を目指す。

3.2 設計内容

本研究では、手書きシンボルによる制御デバイスとして、Raspberry Pi Zero 2 W (以降 Zero 2) を中核とした、以下のよう実装を提案する。

本研究におけるデバイスは、図 3 に示す 5 inch タッチディスプレイおよび Zero 2 により構成し、ディスプレイ側面に固定する形でマイコンを接続し、デバイス全体としては、マウス、あるいはスマートフォンのような感覚で持ち運び、運用ができるようにすることを目標としている。



図 3 使用するディスプレイおよび GUI 表示の例

デバイス上では、簡素なディスプレイ表示による操作の受付が行われ、スクリーン上部の欄には手書きシンボルを入力してもらい、キャンバス領域の下部に存在しているスライダーを右にスライドすることで、入力内容が受理され、ダブルタップすることにより、入力内容を消去することができる。アプリケーション面の実装内容に関しては、次節で触れており、判別されたシンボルに対して、事前を選択された規格に基づいて照明制御信号の送信を行う。

3.3 手書きシンボルの判別

本研究におけるデバイスは、Zero 2 により構成されることから、独立した状態で画像認識をする必要がある。このことから、Zero 2 上に Raspbian 環境を構築したうえで、Raspbian 上で Python プログラムを実行することにより、手書きシンボルの判別を行う。

タッチスクリーン上に表示される入力受付用 GUI は、Kivy [6]により構成され、図 4 のようになっている。GUI 上に書き込まれた手書きシンボルは、いったん画像ファイルとして保存されたのち、OpenCV を用いたプログラムによりシンボルの判別が行われ、判別結果に基づく照明制御信号がデバイスから発信されるといった流れになっている。

3.4 デバイス全体の構造

本研究において使用するデバイスは、5 inch ディスプレイに Zero 2 を接続した形に設計し、コンピュータと接続する USB ポートに加え、MIDI 信号出力用の DIN ソケットが利用可能な形とする。

本体ディスプレイ上には、前節で述べた、手書きシンボルを受け付けるための GUI のみが表示されるようにし、ソフトウェア側におけるシンボルの分類結果を基に、本体 DIN ソケットより、照明機材の操作に使用する MIDI 信号を出力する形とし、照明卓を通して MIDI 信号に応じて DMX 規格の信号が送信される。

本デバイスのディスプレイに 5 inch のディスプレイを採用した理由として、本デバイスを小型な設計とすることで、本デバイスの運用を容易にし、今回の目的であるライブ会場等における用途だけでなく、自宅等の個人環境における、IoT デバイスの制御等にも利用可能な形とするためである。

また、本研究において Zero 2 による新規デバイスを提案したことの理由として、ライブ会場等における大規模な照明システムの場合、照明卓の操作を行う信号としては MIDI 信号などが採択されている場合が多く、最終的には、利用目的に合わせて出力信号を切り替えられるようなシステムを構築したいと考えていることから、出力信号をカスタマイズできるマイコンを採用している。

3.5 デバイス上におけるシンボルパターン

本デバイスにおける懸念点の一つとして、画像分類の認識結果によるパターンの選択を行っていることから、誤分類により、ユーザの意図していない動作が出力されてしまう恐れがある。特に今回意図されている用途においては、誤った信号の送信は大規模なトラブルを起こす原因となりかねない。

そのため、本研究では、単純にローマ字や数字などの既存のシンボルセットではなく、できる限りシンボル同士の類似形状を持たないシンボルセットを提案し、パターン選択用のシンボルとして採用することで、画像分類の誤判別による制御用信号の誤送信を防ぐことを意図している。

現時点において端末内に記憶しているシンボルは、図 4 に示す通りであり、できる限り類似形状が生じることを防ぎ、類似点を持つシンボルに関しては、シンボル間にわかりやすい差異を作ることにより、対処している。

実験では、これらのシンボルの入力速度だけでなく、誤分類の割合についても検証を行い、誤分類しやすいシンボ

ルの組み合わせに関しては、再度形状の見直しを行うことにより、分類精度の改善を図る。

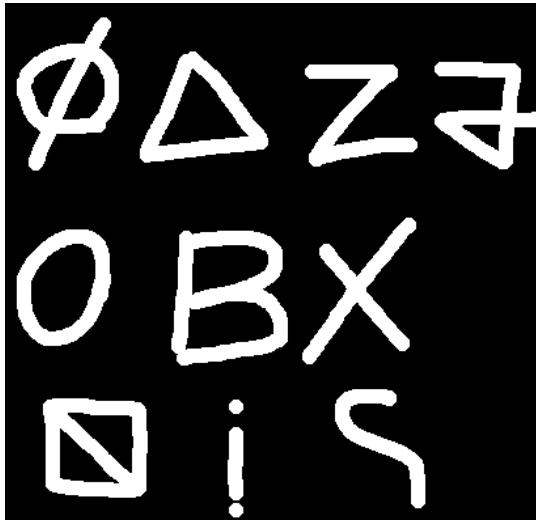


図 4 使用されるシンボルの例

3.6 評価実験

実験では、コンピュータ上における照明機材のシミュレータ(今回は、Daslight 5 [7] を使用)と本デバイスを接続し、被験者に呼び出してほしいパターンを連続的に提示し、それに伴い本デバイスに対してシンボルを入力してもらい、パターンが提示されてから実際に動作するまでの時間の測定を行うとともに、シンボル入力においてどの程度誤入力があったかを数値化し、本手法においてどの程度の信頼性が担保されているかを評価する。

実験で与えられるパターンは、現在提案している 10 種類程度のパターンをランダムに表示するようにし、ユーザはパターン提示を行うスクリーンを注視しつつ、手元のデバイスを用いてシミュレータ上の照明機材を操作してもらう。対照実験としては、本デバイスではなくキーボードをシミュレータ上に接続し、同じくパターンが提示されてからの時間および提示されたパターンと実際に実行されたパターンの合致率を検証、比較する。

4. 今後の課題

今回の実験では、本研究における目的の一つである、照明機材のパターン制御に重点を置いて、検証を行っている。本研究において提示した、手書きシンボルにより機材を操作する手法には、画面上の入力位置に依存せず、覚えやすい単純なシンボルを採用することにより、照明機材の操作と手先の動作が結び付けやすいという利点がある。このようなデバイスの特性は、より一般的な機器の操作や、視覚的に不自由な者が利用しやすいようなシステムの構築に、有効的であると考えられる。

このことから、今後の展望としては、本デバイスのシン

ボル識別精度をより向上させていくと同時に、利用用途に合わせて出力信号を切り替え、同時に遠隔接続も考慮したシステム設計にすることで、より一般化した手法として運用できるのかといった点について検証していきたい。

また、より細かなパラメータの操作に対して、シンボル判別後に操作の分岐を可能にし、特定のシンボルが検出された際に各パラメータを制御するモードへと切り替えられるような機能を加えることなどにより、操作の拡張性をより高め、操作の直感性が失われない範囲で、機能の追加を行っていく。

5. おわりに

本研究では、リアルタイムでの機材のパターン操作に対して、手書きシンボルの判別による照明制御用信号の操作を行うデバイスを提案した。ユーザは、デバイスのタッチスクリーン上に、スクリーン上の位置によらず各制御信号のパターンに対応したシンボルを入力することにより、対応した操作を呼び出すことができる。入力位置によらず、シンプルかつ重複性も少ないシンボルを入力として採用することにより、ユーザのごとの身体的な差異等によるシンボルの誤認識を防ぎ、かつ求められる状況に合わせて即時に受付できることで、リアルタイム性も確保することができると考えた。

現時点における本研究の懸念点として、タッチスクリーンが小型なことによるシンボルの誤認識や、意図していない信号が送信されることによる、機材への影響などが挙げられる。これに対して、デバイス形状の見直し、より適したシンボル形状の提案などにより、これらの問題に対して対処可能であると考えられる。

参考文献

- [1] “Virtual Live Lab”, <https://miku.ec.com/> (参照 2024-12-24)
- [2] Bojan Mrazovac, Milan Z. Bjelica, Djordje Simic, Srdjan Tikvic.: Gesture Based Hardware Interface for RF Lightning. *Control, 9th IEEE International Symposium on Intelligent Systems and Informatics (SISY 2011)*
- [3] Sana Shokat, Rabia Riaz, Sanam Shahla Rizvi, Abdul Majid Abbasi, Adeel Ahmed Abbasi, and Se Jin Kwon.: Deep learning scheme for character prediction with position-free touch screen-based Braille input method. *Human-centric Computing and Information Science*, article number: 41 (2020)
- [4] Zhengxin Guo, Fu Xiao, Biyun Sheng, Huan Fei, Shui Yu.: WiReader: Adaptive Air Handwriting Recognition Based on Commercial Wi-Fi Signal. *IEEE Internet of Things Journal*. (2020)
- [5] Mitulgiri H. Gauswami, Kiran R. Trivedi.: Implementation of machine learning for gender detection using CNN on raspberry Pi platform. *2nd International Conference on Inventive Systems and Control*. (2018)
- [6] Kivy : <https://kivy.org/> (参照 2024-12-24)
- [7] Daslight 5 <https://www.daslight.com/ja/daslight5> (参照 2024-12-24)