

魚眼カメラと AI 技術を活用した屋外照明制御システムの開発

松本陽斗、水野温*

Development of Outdoor Lighting Control System Using Fisheye Camera and AI Technology

Haruto Matsumoto, Motomu Mizuno*

要 旨

国内西洋りんごの発祥の地である七飯町の道の駅「なないろ・ななえ」には、りんごをモチーフとしたモニュメントが設置されている。令和2年には、施設の集客力向上を目的として、AIを活用してモニュメントの照明を動的に制御するシステムの開発・実装を行った。コロナ禍の影響もあり集客効果を定量評価するには至っていないが、施設管理者からの評判も良く、2年間に渡り安定して継続運用された。鑑賞者等へのヒアリング調査等により、鑑賞者の対話的動作で変容・完成する芸術作品が求められていることがわかった。

そこで、本研究では施設のさらなる集客力向上を目的として、鑑賞者が水晶占いのように手をかざすことで操作可能な対話型照明制御の実現に向けて、実証システムの開発を行った。手掌部検出には魚眼カメラと AI を用い、魚眼画像の歪み及び手掌部の多様性に対する汎化性能を検証した。その結果、夜間に照明明滅する環境において高精度に動作することを確認し、本システムの実用性を認めた。

1. はじめに

商業施設等におけるオブジェやモニュメント等の芸術作品には、施設の魅力を高め、鑑賞者の滞在時間を延ばす等の効果が認められており、空間演出や経済活動の活性化にも重要な役割を担っている^{1,2)}。また、ライトアップやイルミネーション等の照明演出は、誘目性や注目度が高く、さらなる集客効果の向上が期待できる³⁾。そこで、令和2年に、道の駅「なないろ・ななえ」の集客効果向上を目的として、既設モニュメント(図1)の照明に AI 技術を取り入れ、鑑賞者の距離と人数に応じて点灯パターンが変化する動的制御システムを開発した⁴⁾。実装後の施設管理者や鑑賞者等へのヒアリング調査等により、鑑賞者の操作に応じて照明が対話的に変容する芸術作品(インタ

ラクティブアート)が求められていることがわかった。芸術と情報技術の融合表現手法であるインタラクティブアートは、鑑賞者自身が作品の一部となるため、充実感や満足感を得やすい。特に光や動きを伴う作品は、ソーシャルネットワークキン



図1 モニュメント外観

*合同会社ヒルズ社

責任著者連絡先 (Haruto Matsumoto) : matsumoto@techakodate.or.jp

グサービスに投稿する写真や動画等として人気が高く、設置後に施設への来場者数が増加した事例も多い⁹⁾。そこで、当該モニュメントのインタラクティブアート化を実現する屋外照明制御システムの開発を行った。

コロナ禍の影響や経験を考慮しつつ、娯楽性が高い対話制御の入力方法として、水晶占いのように手をかざす動作による非接触入力式に着目した。そして、1台で全周撮影可能な魚眼カメラとAIを用い、検出した手掌部の状況に応じた照明演出を実現する基本システムを試作して、実用性を検証した。

2. システム設計及び試作

2.1 ハードウェア設計及び試作

照明器具における現在の主力光源はLEDであり⁶⁾、制御対象の既設照明も、複数のLEDを光源とした多彩な点灯色を表現可能なパーライトである。既設照明の通信制御方式には舞台照明等の制御で用いられるDMX512(Digital Multiplex with 512 pieces of Information)⁷⁾を採用していることから、演出性の高い照明制御を実現するため、PCを用いるシステム(図2)を構成した。

手掌部検出や照明制御には高負荷な処理が求められるため、高性能なPCを利用して各種処理を実行する必要があると考えられた。そこで、制御用PCには、一般的に入手可能な小型PCの中から、クロック周波数が1.7GHzで10コア/12スレッドのCPUを搭載した高性能な機種を選定した。また、DMX512を用いた照明制御では各機材同士を有線接続する必要がある。そのため、制御

機器の設置場所はモニュメントの近傍となることから、制御機器の大型化による視覚的な印象や通行性の低下を避けなければならない。そのため、各制御機器はメンテナンス性や対候性等も考慮し、IPX3相当の防水性能を有する小型のボックスに格納した。図3に照明及びモニュメントの近傍に設置した屋外仕様の試作制御ボックスを示す。

2.2 ソフトウェア設計

手の動きに応じて対話的に照明制御を行うためには、画像上の手掌部の位置を検出する必要がある。画像から手掌部等の複雑な対象を検出する場合、検出精度の高さ等からAIの中でも深層学習による手法が多く用いられている。主な画像認識手法には、画像分類と物体検出、セマンティックセグメンテーション⁸⁾があり、本件では画像上における検出対象の位置が得られる手法の中で、一般的に最も高速な処理が可能な物体検出を選定した。物体検出の手法には、Faster R-CNN⁹⁾やSSD¹⁰⁾、YOLO¹¹⁾等があるが、本件では、このうち高速かつ高品質な検出が可能で、長期使用の実績が多いYOLOv5¹²⁾を用いた。YOLOv5には、ニューラルネットワークの中間層の数等が異なる5種類のモデルがある¹³⁾。各モデルは検出精度や処理速度が異なり、それぞれはトレードオフの関係にある。インタラクティブアートでは、自身の操作が照明演出の変容を引き起こしているという感覚(運動主体感)¹⁴⁾が重要である。しかし、運動主体感は、遅延が200ms以上になると顕著に低下し始め、400ms以上ではほ

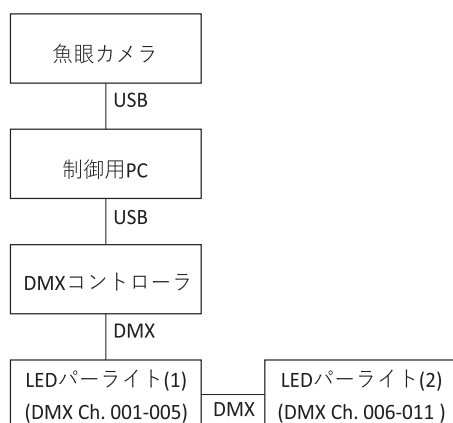


図2 照明制御システム構成



図3 PC仕様及び制御ボックス外観

とんど感じられなくなる等、遅延が感覚に影響を与えることがわかっている^{15),16)}。そこで、各モデルについて、任意のタイミングにおける100回分の処理速度を計測した結果、図4に示す結果が得られた。そこで、200ms以下での処理が可能なモデルの内、最も高い検出精度が期待できるYOLOv5sをモデルに用いることとし、高精度な手掌部検出の実現に向けて、画像の歪みと汎化性能(未知の色情報等の検出性能)について実験・検証を行った。

3. 画像の歪みに関する実験

3.1 実験方法

一般的な手掌部検出の学習データには、収集性等の点から標準的なレンズで撮影した歪みの小さい画像を利用する。しかし、魚眼カメラで撮影した魚眼画像は、図5に示すように画像端に近づくほど大きく歪む特徴がある。そのため、通常画像を学習データとした場合、被写体形状に大きな差異が生じることから検出精度の低下が予測された。

魚眼画像から手掌部を検出するには、画像を未処理のまま利用する方法と、キャリブレーション

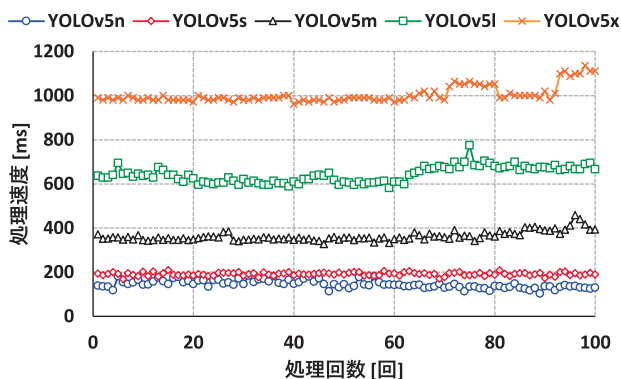


図4 照明制御システム構成



図5 魚眼画像の歪み

を行い魚眼画像の歪みを補正して利用する方法がある。そこで、本実験では手掌部検出用の学習済みモデルの作成方法として、(1)学習データに通常画像を用い魚眼画像から検出を行う方法、(2)学習データに魚眼画像を用い魚眼画像から検出を行う方法、(3)学習データに通常画像を用い歪みの補正を行った魚眼画像から検出を行う方法の3通りの組み合わせを定めた。

手掌部が映る魚眼画像は、図6に示す魚眼カメラを用いて撮影した魚眼画像60枚(学習用50枚、検証用10枚)、一般的なカメラを用いて撮影した通常画像50枚(学習用)を使用し、3種類の学習済みモデルを作成した。精度検証の指標には、モデルの予測結果と実際の値の一致度を示す平均適合率(Average Precision)に着目し、1.00に近いほど検出精度が高いことを示すmAP50-95(mean Average Precision)を用いた。

3.2 実験結果及び考察

モデル(1)～(3)のmAP50-95を算出した結果を図7に示す。モデルの複雑さ等によりmAP50-95は変動するが、一般的に0.50以上であれば高精度であるとされている。本実験では全結果が0.73以上を示したが、特に学習データに魚眼画像を用いたモデル(2)において約0.92と最も高くなった。検出結果を検証したところ、モデル(1)では歪みが大きい画像端付近で検出不可や誤検出が確認され、モデル(3)では補正により画像端の一部が切り取られ検出に利用できないことがわかった。これらのことから、本件においては学習データに魚眼画像を用いたモデル(2)が最適であると考えられた。



モデル	ELP-USB16MP01-BL180
画角	180度
センサー	CMOS IMX298
画素数	1600万画素

図6 魚眼カメラ仕様

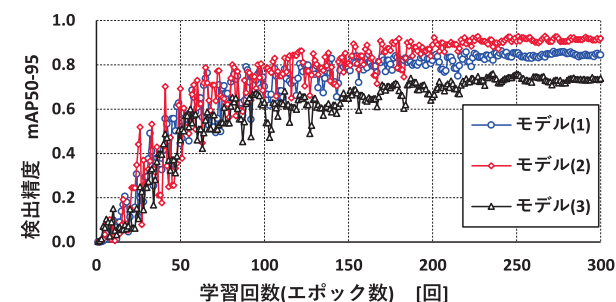


図 7 魚眼画像の歪みによる検出精度の比較

4. 汎化性能に関する実験

4.1 実験方法

手掌部の検出では、季節や嗜好により多様な色の手袋の装着等が予想される。汎化性能の向上を図る上で、色に依存しない特徴を学習させて検出することが有効であると考え、暗所で照明が明滅する環境で、手掌部が映る魚眼画像を撮影し、検出精度に関する実験を行った。本実験では、撮影した魚眼画像計 988 枚(学習用 889 枚、検証用 99 枚)に対して図 8 に示した色情報の変更を行い、(A)未処理画像、(B)グレースケール化画像、(C)未処理画像及び同数の色相を変更した画像とした 3 種類の学習済みモデルをそれぞれ作成し、歪みに関する実験と同様に mAP50-95 による検証を行った。

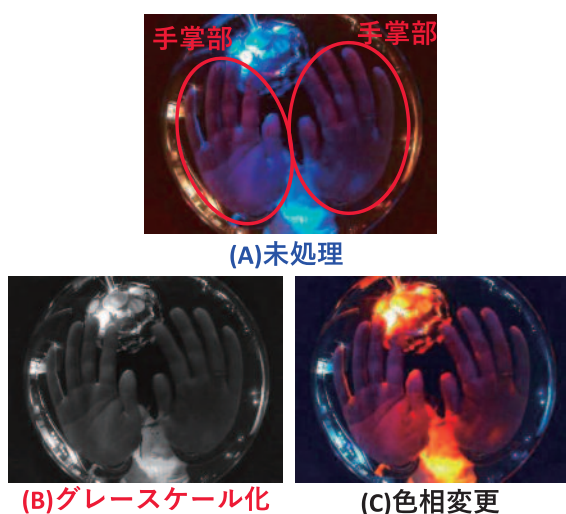


図 8 画像の色情報の変更例

4.2 実験結果及び考察

モデル(A)～(C)の mAP50-95 を算出した結果を図 9 に示す。モデル(A)と(B)の mAP50-95 は約 0.78 と高い値を示したが、モデル(C)では 0.46 未満の結果となった。これは色相のみ変更した画像を追加したことによって、検出対象外の背景等の外乱にモデル(C)が過剰に適合し、過学習が発生したためと考えられる。モデル(B)はグレースケール化による色情報の減少等により、モデル(A)に比べると手掌部付近が暗い場合に時折検出不能となることを確認した。本条件のように手掌部自体の色情報に比べて照明の点灯色の影響が支配的な場合では、未処理画像を用いた検出が最も高精度であることがわかった。

5. 検出精度の評価・動作試験

開発したシステムの検出精度の評価として、顔や手、全身姿勢等の検出が可能なオープンソースの機械学習ライブラリである MediaPipe (Google 社)¹⁶⁾との比較を行った。MediaPipe における手掌部の検出は、開発システムの物体検出とは異なり、指や掌等の合計 21 か所の位置を検出し、手掌部の姿勢を推定することで行われることから、単純な mAP50-95 の比較による精度評価は困難である。そこで、式(ア)で表される検出の正解率と、式(イ)の適合率(誤検出しない確率)、式(ウ)の再現率(検出対象を見逃さない確率)を用いて比較を行った。なお、評価には本環境にて撮影した、手掌部が映る魚眼画像 71 枚

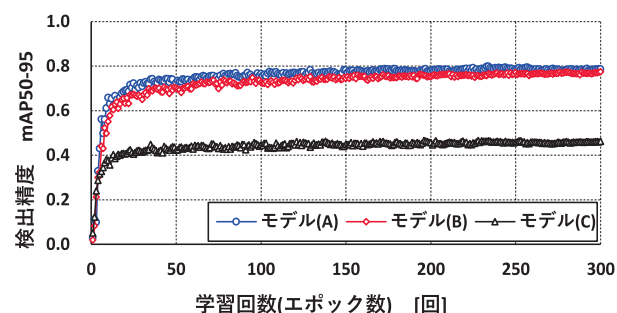


図 9 色情報の変更による検出精度の比較

と、手掌部が映らない魚眼画像 29 枚の合計 100 枚を用いた。また、検出成否の基準には物体検出の正確さを 0.0 から 1.0 の範囲で示す信頼度スコアを用い、閾値を 0.5 とした。

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \dots (\text{ア})$$

$$Precision = \frac{TP}{TP + FP} \dots (\text{イ})$$

$$Recall = \frac{TP}{TP + FN} \dots (\text{ウ})$$

TP: 正しく物体を検出した数

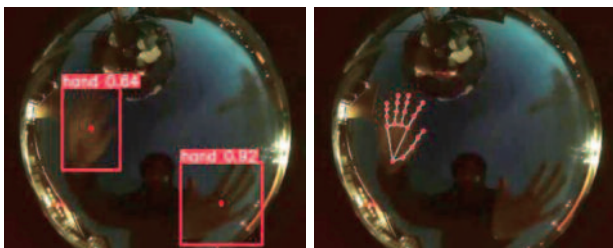
FP: 誤って物体を検出した数

TN: 正しく物体がないと判断した数

FN: 物体を検出できなかった数

図 10 に示す結果から、双方において適合率は最大値である 1.00 に極めて近い値が得られた。しかし、正解率と再現率については、開発したシステムでは、それぞれが約 0.69、約 0.64 と高い値を示したのに対し、MediaPipe では、それぞれが約 0.24、約 0.08 と極めて低くなることがわかった。暗所で照明が明滅する本環境において、魚眼画像から手掌部検出を行う場合、汎用利用を主目的とした MediaPipe の標準的な学習済みモデルでは、高精度な検出を行うことは困難であることがわかった。その理由として、暗所で

	開発システム	Mediapipe
TP (正しく物体を検出した数)	90	11
FP (誤って物体を検出した数)	3	0
TN (正しく物体がないと判断した数)	29	29
FN (物体を検出できなかった数)	51	133
正解率	0.69	0.23
適合率 (誤検出しないう率)	0.97	1.00
再現率 (検出対象を見逃さない確率)	0.64	0.08



開発システム 検出例

MediaPipe 検出例

図 10 開発したシステムと MediaPipe の検出精度比較

は手掌部と背景の境界が曖昧になることや、画像に写る手掌部の色も明所環境下とは大きく異なること、魚眼画像に写る手掌部の歪みが通常画像に比べて大きいこと等が考えられる。そのため、環境に合わせた独自の学習済みモデルを用いる必要性を確認した。

最後に、開発したシステムを道の駅「なないろ・ななえ」のモニュメント照明に実装し、動作試験を行った。実験風景と検出結果の例を図 11 に示す。目視確認の結果、魚眼画像から手掌部を高精度に検出でき、検出結果に応じた対話制御が正常に動作することを確認した。なお、本システムは動作試験から 5 カ月以上継続して稼働しており、防雨性等の確認が行えたことから、屋外での実用性能を十分に有しているものと判断する。

6. まとめ

暗所で照明が明滅する環境において、魚眼カメラに映る手掌部を高精度に検出し、検出結果に応じて非接触で対話的に照明を制御する屋外照明制御システムを開発し、モニュメントのインタラクティブアート化を実現した。今後は、追加学習により指の状態等を検出可能として制御に活用することで、より多様な照明演出が可能な制御システムの実現を図る。



図 11 実験風景と手掌部の検出例

謝辞

機器設置に伴うネットワーク環境設定では、エフエスシー(株) 若山みつる様、動作試験等では、(一社)七飯町振興公社の職員の皆様にご協力いただいた。各位に深く感謝する。

参考文献

- 1) 園田聡：日本の公共的空間の整備・活用におけるプレイスメイキングの展開に関する研究、(2015)、PhD Thesis. 工学院大学
- 2) 土田和香、大川碧望、佐藤慎也：日本建築学会大会学術講演梗概集、2023、(2023)、pp.649-650
- 3) Tavares P, Ingi D, Araújo L, Pinho P, Bhusal P: Sustainability 13, no.22:12657, (2021), <https://doi.org/10.3390/su132212657>
- 4) 松本陽斗、村田政隆、松村一弘、水野温：北海道立工業技術センター研究報告、第 17 号、(2022)、pp.49-54
- 5) 例えば、アートと経済社会について考える研究会の報告書：経済産業省(2023)、P260
- 6) 市川重範、鳥居龍太郎：電気設備学会誌、39、1(2019)、pp.18-21
- 7) 委員会報告：電気設備学会誌、37、9(2017)、pp.669-675
- 8) 日本ロボット学会誌、Vol.35、No.3(2017)、pp.180-185
- 9) Shaoqing Ren, Kaiming He., Ross Girshick, Jian Sun: Advances in neural information processing systems (2015), pp.1-9
- 10) Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, Alexander C. Berg.: European conference on computer vision (2016), pp.21-37
- 11) Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi: Proceedings of the IEEE conference on computer vision and pattern recognition, pp.779-788
- 12) Glenn Jocher, ultralytics/yolov5, <https://github.com/ultralytics/yolov5>
- 13) Shaun G.: Trends in Cognitive Sciences, Vol.4 (No.1), (2000), pp.14-21
- 14) Takaki Maeda, Keisuke Takahata, Taro

Muramatsu, Tsukasa Okimura, Akihiro Koreki, Satoru Iwashita, Masaru Mimura, Motoichiro Kato: Psychiatry Research, 209(3), (2013), pp.386-392

15) Michihiro Osumi, Satoshi Nobusako, Takuro Zama, Naho Yokotani, Sotaro Shimada, Takaki Maeda, Shu Morioka: PloS one 14.7 (2019), e0219222.

16) MediaPipe, <https://google.github.io/mediapipe/>