

Visual Computing, Devices & Communications

# 画像電子学会誌

*The Journal of the Institute of  
Image Electronics Engineers of*

*The Journal of the Institute of  
Image Electronics Engineers of  
Japan*

*The Journal of the Institute of  
Image Electronics Engineers of  
Japan*

## AI 技術と画像電子関連技術論文特集

◆【特集論文】大規模言語モデルの強化学習と画像クラスタリングに基づく卒業アルバムのレイアウト生成

◆【特集論文】髪型の特徴が国籍分類に与える影響：深層学習による東アジア顔画像の分析

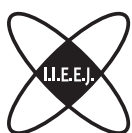
### 一般の部

◆【論文】指向性をもつ新型ノズルを用いたホログラフィック空間像投影用のミストスクリーンの安定性の向上

◆【実践論文】特別史跡百済寺跡における Indirect AR・VR 統合体験システムの開発と社会実装

◆【システム開発論文】乗客行動の認識に基づくバス発車判断の自律化システム

◆【システム開発ショートペーパー】可動カメラを用いた高速道路上の落下物検出



一般社団法人

画像電子学会

Vol.55 No.2

2026

池上通信機株式会社  
 科学技術振興機構  
 公益財団法人画像情報教育振興協会  
 桂川電機株式会社  
 株式会社ゲネシスコンマース  
 コニカミノルタ株式会社

Z A Z A 株式会社  
 大日本印刷株式会社  
 日本テレビ放送網株式会社  
 日本電信電話株式会社  
 日本放送協会  
 株式会社日立製作所

富士フイルムビジネスソリューション株式会社  
 古野電気株式会社  
 三菱電機株式会社  
 株式会社リコー

編集委員会

- (委員長) 内田 理
- (副委員長) 小林直樹, 竹島由里子, 石川雅浩.
- (編集顧問) 安田靖彦, 富永英義, 小宮一三, 小野文孝, 羽鳥好律, 松本充司, 加藤茂夫, 田中 清, 児玉 明.
- (編集理事) 石川雅浩, 新田高庸, 向井智彦, 今給黎 隆
- (編集幹事) 河村尚登, 森谷友昭, 山田雄一郎.
- (編集・査読委員) 荒井良徳, 池田充郎, 井尻 敬, 石川知一, 五十嵐悠紀, 上平員丈, 大井 翔, 大木眞琴, 木村俊一, 久下哲郎, 倉掛正治, 小池崇文, 櫻井快勢, 佐藤甲斐, 佐藤周平, 白川真一, 新谷幹夫, 関野雅則, 田中賢一, 田中 清, N.P.チャンドラシリ, プレマチャンドラ・チンタカ, 豊浦正広, 長谷川まどか, 濱本和彦, 藤澤 誠, 藤代一成, 牧田孝嗣, 吉田典正, 吉川 浩, Chee Seng CHAN, Paramesran RAVEENDRAN, KokSheik WONG.
- (査読委員) 荒川賢一, 荒木昭一, 有川智彦, 伊藤貴之, 五十嵐悠紀, 岩切宗利, 岩橋政宏, 大澤秀史, 尾上孝雄, 金井 崇, 金子俊一, 金森由博, 金子 格, 金田和文, 北郷正輝, 勝間ひでとし, 栗原恒弥, 黒沢俊晴, 洪 博哲, 小館亮之, 駒形英樹, 小町祐史, 今間俊博, 斎藤隆文, 齋藤 豪, 齊藤文彦, 佐藤真知子, 篠原克幸, 島村 潤, 下馬場 朋禄, 白井啓一郎, 杉崎栄嗣, 瀬崎 薫, 瀬政孝義, 高島洋一, 高野邦彦, 田中芳樹, 高橋時市郎, 谷口行信, 田村 徹, 辻 宏行, 鉄谷信二, 中村康弘, 納富一宏, 包 躍, 林 正樹, 福江潔也, 堀田裕弘, 本宮隆広, 茅 暁陽, 松木 眞, 松田浩一, 三田雄志, 三ツ峰秀樹, 六浦光一, 森島繁生, 柳原政弘, 藪下浩子, 山口隆二, 山崎龍次, Hernan AGUIRRE, Yoong Choon CHANG, Robin Bing-Yu CHEN, Mochamad HARIADI, Pizzanu KANONGCHAIYOS, Teck Chaw LING, Keat Keong PHANG, Nordin BIN RAMLI.
- (事務担当) 大嶽英宗, 浮ヶ谷 修, 福島理恵子, 本田京子.

入会のご案内

入会ご希望の方は下記ご参照の上、学会ホームページよりお申込頂くか、事務局にその旨ご連絡ください。

○会員の種別

- 正会員：本会の目的に賛同する個人
- 学生会員：本会の目的に賛同する学生
- 賛助会員：本学会を援助する個人または法人
- 特殊会員：本学会の目的に賛同する個人以外の図書室、研究室など

○入会金および年会費

- 入会金：正 会 員 1,000 円 学生会員 500 円
- 年会費：正 会 員 10,000 円
- 学生会員 3,000 円
- 賛助会員 50,000 円 (1 口)
- 特殊会員 12,500 円

○ご連絡先

- 〒116-0002 東京都荒川区荒川 3-35-4
- ライオンズマンション三河島第二 101 号
- TEL (03)5615-2893 FAX (03) 5615-2894
- E-mail: hensyu@iieej.org (編集)
- kikaku@iieej.org (研究会・会員情報)
- hyoujun@iieej.org (テストチャート)
- <http://www.iieej.org/>
- <http://www.facebook.com/IIEEJ>

# 画像電子学会誌

第55巻 第2号 通巻276号 (2026年4月)

## 目次

### 随想

- 95 生成AIが発展する時代の論文執筆と学会の在り方 今給黎 隆

### AI技術と画像電子関連技術論文特集

- 96 AI技術と画像電子関連技術論文特集に寄せて 編集委員会

### 論文

- 97 大規模言語モデルの強化学習と画像クラスタリングに基づく卒業アルバムのレイアウト生成 松田 展, 佐野睦夫, 葉山恒生, 大賀将弘

- 106 髪型の特徴が国籍分類に与える影響: 深層学習による東アジア顔画像の分析 永岡正悟, 張 珏, ナイワラ P. チャンドラシリ

### 一般の部

### 論文

- 114 指向性をもつ新型ノズルを用いたホログラフィック空間像投影用のミストスクリーン安定性の向上 高野邦彦, 須田諒真, 坂田 凜, 原田博貴, 前川海斗, 篠崎大和, 戸塚雄大, 佐藤甲斐, 浅井紀久夫

### 実践論文

- 122 特別史跡百濟寺跡におけるIndirect AR・VR統合体験システムの開発と社会実装 岡田 錬, 濱田悠太, 大野倫敬, 川原佑斗, 神田啓冴, 和田築来, 菅沼祐耶, 中山堪太, 河合紀彦

### システム開発論文

- 134 乗客行動の認識に基づくバス発車判断の自律化システム 山崎将幸, 滝沢徳高

### システム開発ショートペーパー

- 144 可動カメラを用いた高速道路上の落下物検出 軒田佳林, 佐治 斉

### グループ紹介

- 150 神奈川工科大学 鈴木研究室 鈴木 浩

### 会告・ニュース

- 152 理事会だより 161 2027年4月号(英文誌6月号)論文特集号論文募集  
153 画像電子学会研究会等予定 一 社会貢献と画像電子関連技術特集一  
154 2026年度 第54回 画像電子学会 年次大会Call for Papers 162 論文投稿の手引き  
156 画像電子学会 第317 回研究会 in 東京 166 日本画像学会誌・日本写真学会誌・日本印刷学会誌目次  
158 Advanced Image Seminar 2026 (AIS2026) —画像×AIの産業応用の最新動向— 171 会報  
160 Call for Papers Special Issue on “Extended Papers Presented in IEVC2026” 171 編集後記

**The Journal of  
the Institute of Image Electronics Engineers of Japan**  
Vo1.55 No.2 April 2026  
CONTENTS

---

**Foreword**

- 95 The Meaning of Paper Writing and the Role of Academic Societies in the Era of Generative AI Takashi IMAGIRE

**Special Issue on Image Electronics Technologies Related to AI**

- 96 Upon the Special Issue on Image Electronics Technologies Related to AI Editorial Committee

**Contributed Papers**

- 97 Layout Generation for Graduation Albums Based on Reinforcement Learning with Large Language Model and Image Clustering Hiraku MATSUDA, Mutsuo SANNO, Nobutaka HAYAMA, Masahiro OGA
- 106 Hairstyle Matters: Deep Learning-Based Analysis of Nationality Classification in East Asian Facial Images Shogo NAGAOKA, Jue ZHANG, Naiwala P. CHANDRASIRI

**Regular Section**

**Contributed Paper**

- 114 Improvement of the Stabilization of the Mist Screen for the Spatial Holographic Projection Employing the Newer Structured Nozzle to Obtain the Mist-Flow with Enhanced Directionality Kunihiko TAKANO, Ryoma SUDA, Rin SAKATA, Hirotaka HARADA, Kaito MAEKAWA, Yamato SHINOZAKI, Yudai TOTSUKA, Koki SATO, Kikuo ASAI

**Practice Oriented Paper**

- 122 Development and Social Implementation of an Integrated Indirect AR/VR Experience System at Kudaradera Temple Ruins, a Special Historic Site Ren OKADA, Yuta HAMADA, Michitaka ONO, Yuto KAWAHARA, Keigo KANDA, Kizuku WADA, Yuya SUGANUMA,

**System Development Paper**

- 134 Autonomous Decision System for Bus Departure Based on Passenger Behavior Recognition Masayuki YAMAZAKI, Hotaka TAKIZAWA

**System Development Short Paper**

- 144 Fallen Object Detection on Expressway Using a Pan-Tilt Camera Karin NOKITA, Hitoshi SAJI

**Research Group Introduction**

- 150 Suzuki Lab., Kanagawa Institute of Technology Hiroshi SUZUKI

## 生成 AI が発展する時代の論文執筆と学会の在り方

今給黎 隆 (編集理事, 東京工芸大学)

## The Meaning of Paper Writing and the Role of Academic Societies in the Era of Generative AI

Takashi IMAGIRE (Editorial Director of IEEEJ, Tokyo Polytechnic University)



本号 (画像電子学会誌 Vol.55, No.2) の論文特集テーマは、「AI 技術と画像電子関連技術」となっております。近年、AI 技術は急速に発展し、生成 AI や大規模言語モデル (LLM) などの大きな話題が出現する間隔が日に日に短くなってきております。画像生成や画像認識の分野が AI の登場により新たな局面を迎えているのと同様に、私たちの研究活動スタイルそのものも急激な変化の中にあります。皆さんも日々の研究において、生成 AI を積極的に活用されていることでしょう。私自身、英語に対して苦手意識を持っておりましたが、近年の AI 技術により英作文のハードルは劇的に下がりました。また、学内外の公務や委員会活動に追われ、研究時間の捻出に苦労されている皆様にとって、論文執筆の補助ツールとしての AI は、もはや欠かせないパートナーとなっているのではないのでしょうか。しかし、論文が「書きやすくなった」ことは、必ずしも良い面ばかりをもたらすわけではありません。投稿数の増加は学会の財政面で喜ばしい一方、論文を取りまとめる編集委員会の立場から俯瞰すると、それは運営側の負担増をダイレクトに意味します。

現状、本学会においては投稿の増加を歓迎する姿勢ですが、他分野の論文誌では深刻な問題が顕在化しつつあることをよく耳にします。例えば、不採録判定を受けた直後に、生成 AI を用いて機械的に修正されたと思われる論文が即座に再投稿されるケースがあります。従来においても、著者の推敲が不十分のため査読段階が「校正作業」の役目を担ってしまう例はありましたが、生成 AI の普及は、文体だけではなく内容としても、この傾向を加速させます。容易に「もっともらしい体裁」を整えられるようになったことで、論理的飛躍やデータの検証が不十分のまま投稿される頻度が高まりつつあります。こうした状況は、ボランティアベースで成り立つ学会運営の持続可能性を揺るがしかねません。査読件数の爆発的な増加は査読者の消耗を招き、「査読に時間を取られすぎて自身の研究に支障が出る」という理由で、査読依頼を辞退されるケースが増えることが危惧されます。結果として、編集者が内容を深く吟味する前に形式的に却下する「デスクリジェクション (Desk rejection)」を増やさざるを得ない未来が現実味を帯びてきそうです。さらに深刻なのは、AI による「高度な欺瞞」への対応です。AI は、都合の良い結果だけを抽出したり、人間が一瞥しただけでは真偽を判断できない微細なデータ調整が可能です。査読者がそのような可能性の検証作業に忙殺されることは、学術の進歩という観点から見れば極めて不毛な現象と言わざるを得ません。

研究者が生成 AI で論文を書き、査読者もまた効率化のために生成 AI で評価を行う。そのような「AI 対 AI」の構図が完成してしまったとき、果たして「学会」というプラットフォームは必要なのでしょうか。単なる情報の掲載であれば、人間を介在させる必要はなくなります。むしろ人間がいない方が効率的に処理できるでしょう。しかし、こうした悲観的な予測が立ち込める今だからこそ、学会には本質的な変化が求められます。情報の洪水の中から、我々にとって本当に重要な価値を見極める「審美眼」こそが、これからの学会員の専門知識の見せ所となるはずですが、それは単なる既知の技術や知識の集積ではありません。AI の推論によりある程度の価値判断は可能ですが、長年、特定の分野で深い思考を積み重ね、苦難を乗り越えてきた研究者だからこそ持ち得る「直感」こそが、AI の安直な利用では到達し得ない領域を拓くこととなるでしょう。各分野に特化し、現場の空気感を知る皆様の意見には、AI の及ばない重みがあります。生成 AI は「情報を編む速度」を劇的に変えましたが、私たちの専門分野における「真理」とは、そのような効率や速度の先にあるものではないはずです。

論文は、単なる記号の羅列やデータの提示ではありません。それは研究者が現場で感じた小さな違和感、実験室での重苦しい沈黙、そして幾多の失敗の末によりやく掴み取った「手触りのある知」の結晶です。私たちは、現実と仮想の境界が曖昧になる時代の中で、より主観的な「心の体験」や「社会の本質」に最適化された世界を希求しています。今後も AI と協力して論文を執筆することはもはや避けられませんが、AI が生成する「もっともらしい言葉」の奔流に呑み込まれそうになったときこそ、立ち止まりましょう。私たちの役割は、単に情報の流通量を増やすことではありません。この分野にしかない「純粋な真理」を研ぎ澄まし、次世代へと繋いでいくことです。10年、20年先を見据え、効率化の波に洗われてもなお残る「本物」を追求する姿勢—それこそが、これからの研究者に求められる最も尊い資質であると信じています。微力ながら、私も編集理事として、皆様がこうした深い探求に没頭できるような学会の環境作りに尽力してまいります。どうぞよろしくお願ひ申し上げます。

## AI 技術と画像電子関連技術論文特集に寄せて

編集委員会

近年の AI (人工知能) 技術の急速な進歩は、画像や映像の処理・認識・生成をはじめとする、さまざまな画像電子関連技術に飛躍的な進展と向上をもたらしました。応用分野も、自動運転、医療画像診断、顔認証システム、工場等での異常検知、監視カメラ/ドローンなどと枚挙にいとまがありません。一方で、フェイク画像の作成など、技術の使い方によって生じる社会的問題への対策も大きな課題となることが予想されます。いずれにせよ、今後、画像電子関連分野において AI 技術がますますその重要性を増していくことは間違いないでしょう。

本特集では、これらの AI 技術の進歩と、それに伴い進化する画像電子関連技術の研究成果や、それらを実際に応用した結果の評価に関する論文、システム開発論文、実践論文、さらにサーベイ論文等を広く募集しました。

その結果、計 6 件の投稿をいただき、和文投稿 3 件のうちフルペーパー 2 件が今号で採録されました。採録となった論文の 1 件目は、「大規模言語モデルの強化学習と画像クラスタリングに基づく卒業アルバムのレイアウト生成に関する研究」です。この論文では、大量の写真を配置するために大きな労力と知識・経験が必要とされる卒業アルバムの写真レイアウトの自動化に向け、入力画像から抽出された属性情報やクラスタリング結果を LLM のプロンプトテキストとして入力するアプローチが提案されています。また、2 件目は、「髪型の特徴が国籍分類に与える影響：深層学習による東アジア顔画像の分析」です。この論文では、顔画像を国籍ごとに分類する際、顔全体だけでなく髪型にも国籍ごとの特徴が含まれることが、東アジア人の顔画像を対象とした画像認識 AI モデルを用いた検証を通じて報告されています。

なお、英文投稿論文 3 件については、採択論文が本年 6 月発行の英文論文誌に特集として掲載される予定です。また、特集号の発行スケジュールに審査が間に合わなかった論文が、その後採録された場合は、和文・英文とも一般論文として掲載の予定です。

AI 技術は、画像処理やビジュアルコンピューティングを含むあらゆる分野に革新的な発展をもたらしており、今後もさらなる深化および広範な展開が期待されます。本特集号を通じて、その研究の一端を皆様にご紹介できることを嬉しく思います。また、本誌では、今後も一般論文として AI 関連の論文、システム開発論文、実践論文等を受け付けておりますので、最新の成果を随時ご投稿いただければ幸いです。

末筆ながら、ご多忙の中、論文の査読をしていただいた方々、ならびに編集・出版にご協力いただいた皆様に、心よりお礼申し上げます。

\* 特集号編集理事

新田 高庸 (会津大学), 向井 智彦 (東京都立大学)

(向井 智彦 記)

## 大規模言語モデルの強化学習と画像クラスタリングに基づく卒業アルバムのレイアウト生成

松田 展<sup>†</sup> (学生会員)      佐野 睦 夫<sup>††</sup> (正会員)      葉山 恒 生<sup>†††</sup>      大賀 将 弘<sup>†††</sup>

<sup>†</sup>大阪工業大学大学院 情報科学研究科    <sup>††</sup>大阪工業大学 情報科学部    <sup>†††</sup>ダイコロ株式会社

### Layout Generation for Graduation Albums Based on Reinforcement Learning with Large Language Model and Image Clustering

Hiraku MATSUDA<sup>†</sup>(Student Member), Mutsuo SANNO<sup>††</sup>(Member), Nobutaka HAYAMA<sup>†††</sup>, Masahiro OGA<sup>†††</sup>

<sup>†</sup> Graduate School of Information Science and Technology, Osaka Institute of Technology,

<sup>††</sup> Faculty of Information Science and Technology, Osaka Institute of Technology,

<sup>†††</sup> DAICOLO CO.,LTD.

〈あらまし〉卒業アルバムの製作には、大量の写真を配置するための大きな労力が伴う。しかし、この作業にはレイアウトに対する経験や知識が必要となるため、自動化を行うことは難しい。また、卒業アルバムのレイアウトを作成する際には、被写体の人数や内容に基づく画像の大きさの調整などの、卒業アルバムに特有のルールに基づいた画像の配置を行う必要があるため、これらの要件を満たすレイアウトを生成する必要がある。しかし、これらの課題に対して取り組んだ研究はこれまでに行われていない。そこで、本研究では大規模言語モデル (LLM) の強化学習や画像のクラスタリングを用いることで、レイアウトを生成する手法を提案する。実験では、実際に使用された卒業アルバムのレイアウトを使用し、適切なレイアウトを生成できることを確認した。

キーワード：レイアウト生成, 大規模言語モデル, クラスタリング, GRPO, 教師ありファインチューニング

<Summary> Creating graduation albums involves significant effort to arrange large numbers of photos. However, this task requires experience and knowledge of layout design, making automation difficult. Furthermore, when designing the layout for a graduation albums, images must be arranged according to rules specific to graduation albums, such as adjusting image sizes based on the number of subjects and the content, and a layout that meets these requirements must be generated. However, no research has been conducted to address these challenges to date. Therefore, this study proposes a method for generating layouts using reinforcement learning with Large Language Model (LLM) and image clustering. Experiments using actual graduation album layouts confirmed that the method can generate appropriate layouts.

**Keywords:** layout generation, large language model, clustering, grpo, supervised fine-tuning

#### 1. はじめに

卒業アルバム制作において、レイアウトの作成はアルバムの外観を左右する重要な作業工程である。この作業を担当するレイアウトターは、経験や知識を基に適切なレイアウトを作成する。しかし、この作業では多くの画像を対象に配置を行う必要があるため、レイアウトターにとって負担が大きい。また、在校生の転校による修正や、発注側の要望による内容の変更にも、限られた納期のもと対応しなければならない。後続の作業には校正や印刷があるため、制作の遅延を防ぐためにも迅速な作業が求められる。この問題を解決するためには、レイアウトの一部作業の自動化を行う必要があると考えられ

る。ただし、これまで卒業アルバムのレイアウト生成に重点を置いた研究は少ない。卒業アルバムは学校全体の生徒を被写体としていることから、家族アルバムなどと比較し、画像の枚数が非常に多いという特徴がある。また集合写真を大きく配置することや、類似する画像を近くに配置するなどの制約条件があり、これらの条件を考慮する必要がある。

そこで、本研究ではレイアウトターの業務改善に焦点を絞り、大規模言語モデル (LLM) 及びクラスタリングを用いた卒業アルバムのレイアウトを自動生成するシステムの実装を行う。提案手法では、前処理を行った画像を LLM に入力し、適切なレイアウトを生成する。また、レイアウトターの意図を反映するため、画像の領域及び配置は手動で作成することも可能で

ある。実験では、アンケート形式の主観的な評価実験及び精度評価を行い、システムの実用性を確認する。そして、評価実験の結果を基に考察及び今後の展望を述べる。

### 2. 従来手法

近年、レイアウトの自動生成に関する取り組みとして、LLM や大規模マルチモーダルモデル (LMM) を用いた研究が盛んにおこなわれている。Tang<sup>1)</sup>らは、html 形式のレイアウトのテンプレートに数値を埋め込むことで、補完形式の生成を行う LayoutNUWA を提案している。Tian<sup>2)</sup>らは、再帰的なレイアウトの構造化と Chain-of-Thought (CoT)<sup>3)</sup>を用いて、LMM でのレイアウト生成を行う ReLayout を提案した。Li<sup>4)</sup>らは、LLM に特化した強化学習手法である Group Relative Policy Optimization (GRPO)<sup>5)</sup>を用いて、レイアウトの最適化を行う LaySPA を提案した。フォトコラージュの自動生成に関する研究として、Yu<sup>6)</sup>らにより提案された SoftCollage が挙げられる。この手法では、誤差逆伝播を用いて、画像の縦横比を保持したまま余白を生じないレイアウトを生成する。また、Zhang<sup>7)</sup>らは、深層強化学習を用いて画像の配置や順序の調整を行うフォトコラージュの手法を提案した。

LLM を用いたレイアウト生成に関する研究では、主に雑誌やポスターのレイアウトを対象にしているが、卒業アルバムのような多くの画像を扱う研究は行われていない。フォトコラージュの自動生成に関する研究では、複数の枚数を用いたレイアウトを生成できるものの、被写体の人数などの詳細な属性は考慮されていない。また、卒業アルバムで画像の間隔をあける必要があるが、前述の研究では余白を作らないように手法を設計している。このように、従来の手法では卒業アルバム制作の要件に対処することはできていない。

そのため、本研究では GRPO や CoT 及び、レイアウトの構造化などの手法や、画像に対する特徴抽出及び分析を用いて、前述の課題の解決を図る。そして、実際の卒業アルバムのレイアウトを学習したモデルによる、卒業アルバムの効率的なレイアウト生成に取り組む。

### 3. 提案手法

本研究では、強化学習やクラスタリングなどの技術を用いて、卒業アルバムのレイアウト生成に特化した LLM の学習を行う。提案手法では、レイアウトの作業の一部をシステムが行う形式で卒業アルバムのレイアウト制作を支援する。まず、システムに見開き単位で画像群を入力した後、属性の抽出及びクラスタリングなどの前処理を行い、プロンプトを作成する。そして、プロンプトを LLM に入力することでレイアウトを生成する。この生成段階では、画像がクラスタリングされた領域、その領域内における画像配置の順で生成を行う。これらのうちどちらかをレイアウトが行うか、どちらもシステムが行うかについて任意で選択することができる。前者の場合では、よりレイアウトの意図を反映したレイ

アウトを作成することができる。また、後者の場合は全ての作業をシステムが行うが、生成したレイアウトをレイアウトの経験や知識に従い微調整することもできる。図 1 に本手法の全体的な流れを示す。また、実際の卒業アルバムの特性を調査するため、ダイコロ株式会社<sup>8)</sup>に所属するレイアウトターに対しヒアリング調査を実施した。以降に述べる卒業アルバムの特性は、前述の調査に基づいたものである。

卒業アルバムには多くの画像が含まれているが、それらは類似したもの同士が近くに配置される傾向にある。そのため、ページごとに画像をクラスタリングし、ページ全体をクラスタ領域に分割することで、レイアウトの構造化を行う。前処理では画像のクラスタリング及び、顔検出や画像評価モデルによる重要度の算出を行っている。これらの特徴は、レイアウトターに対するヒアリング調査に基づき選定している。これらをプロンプトに埋め込むことにより、視覚的な情報を LLM に与えている。また、複数枚の画像から抽出した属性情報をテキストに変換した後に LLM に入力し、レイアウトの生成を行う。レイアウトの生成はクラスタ領域と画像配置に分けて行い、これらの処理は同一のモデルを用いている。

クラスタ領域を生成するタスクでは、レイアウトから抽出したクラスタの範囲を示す矩形を学習データとしている。そして、画像の配置を行うタスクでは、それぞれのクラスタ領域ごとに画像の配置を切り出し、学習データとしている。LLM の学習には教師ありファインチューニング及び、LLM に特化した強化学習手法の GRPO を用いることで、適切なレイアウトを作成する。図 2 に学習の流れを示す。なお、前述のように、レイアウトターの意図を反映したレイアウトを作成するため、任意でクラスタ領域、若しくは画像配置を手動で作成可能な形式としてシステムを設計している。その例として、レイアウトターがクラスタ領域を指定した後、クラスタ内の画像を LLM が配置することにより、レイアウトターの意図を反映しつつレイアウト作成の作業を効率化することができる。

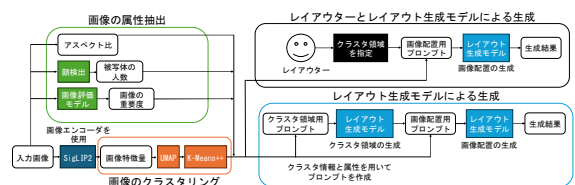


図 1 本手法の流れ

Fig.1 Flow of the method

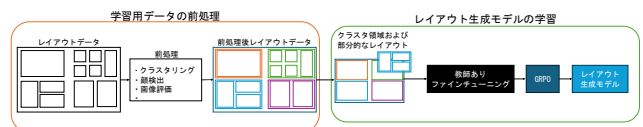


図 2 学習の流れ

Fig.2 Learning process

また、本稿では LLM によるクラスタ領域生成に焦点を絞る。そして、前処理や学習方法の詳細に関しては、以降の節で述べていく。なお、卒業アルバム全体のページ構成は学校の要望などにより事前に決められているため、本手法ではページ構成ではなく見開きでのレイアウトの生成に焦点を絞っている。

### 3.1 レイアウトの形式

レイアウト生成では、複数の要素を取り扱うため、LLM に解釈が容易な形式でプロンプトを作成する必要がある。従来研究<sup>14)</sup>では、レイアウトの表現に html や json が使用されていたが、本手法では xml を使用する。xml の特徴として、構造化データの表現が容易である点が挙げられる。xml では入れ子構造で要素を表現することが容易である。また、属性を自由に設定できることから、座標値やスケールなど多くの数値を扱うことに適している。図3にxml形式のレイアウトの概要を示す。

また、卒業アルバムのレイアウトでは、画像の美観を損なうためアスペクト比を変えないという制約がある。そこで、本手法では画像をスケールとアスペクト比で表現する。これにより、アスペクト比を保持したまま、スケールを変化させることにより画像の大きさを自由に変更することができる。

そして、クラスタ領域はアスペクト比を自由に変更することができるため、x 座標と y 座標で位置を表現し、縦幅と横幅で領域の面積を表す。これらの数値は、図3のようにxmlの属性に埋め込まれる。このように、本手法では数値を多く取り扱うため、LLM の数学的能力を向上させる代表的な手法として知られる Chain-of-Thought (CoT)<sup>3)</sup>をプロンプトに用いる。この手法では、LLM に出力に至るまでの思考課程を生成させることにより、出力の精度を向上させることができる。

本手法ではクラスタごとに配置領域の横幅及び縦幅が変化するため、領域外に画像が配置されてしまう可能性が懸念される。それを防ぐために、画像の座標がレイアウトの範囲外に配置されないよう CoT を用いて制限する。具体的には、画像の x 座標及び y 座標がクラスタ領域の横幅及び縦幅より小さく、0 より大きいことを画像ごとに確認する。

```

xml形式で表現したクラスタ領域のプロンプト
・属性…layout: レイアウト全体を表す, cluster: 各クラスタ領域を表す

<?xml version="1.0" encoding="UTF-8"?>
<layout task="タスクの説明" layout_width="ページの横幅" layout_height="ページの縦幅">
  <cluster sum_images="画像枚数" sum_scores="重要度の合計" x="中心x座標" y="中心y座標">
    <cluster_width="クラスタの横幅" cluster_height="クラスタの縦幅"/>
  </cluster>
</layout>

xml形式で表現した画像配置のプロンプト
・属性…layout: レイアウト全体を表す, image: 各クラスタ領域に属する画像を表す

<?xml version="1.0" encoding="UTF-8"?>
<cluster task="タスクの説明" cluster_width="クラスタの横幅" cluster_height="クラスタの縦幅">
  <image num_person="被写体の人数" x="中心x座標" y="中心y座標" scale="画像の大きさ">
    aspect_ratio="アスペクト比" score="画像の重要度"/>
  </image>
</cluster>
    
```

図3 xml形式で示したレイアウトの概要  
Fig.3 Layout overview in xml format

### 3.2 顔検出による人数カウント

卒業アルバムでは、集合写真は大きく配置される傾向にある。その理由として、顔を判別できるようにある程度の大きさを確保する必要がある点が挙げられる。そのため、顔検出を用いて被写体の人数カウントを行う。また、卒業アルバムにおいて、レイアウトに使用する写真やページ構成は事前に選定されているため、本研究では被写体の分類や人物の重複は考慮しないものとする。

本手法では、処理速度と精度に優れる YuNet<sup>9)</sup>を用いて顔検出を実装した。検出した人数は xml 形式のレイアウト中に、num\_person 属性として埋め込む。また、人物が検出されなかった場合は人数を 0 とする。

### 3.3 画像評価モデル

卒業アルバムでは、視覚的に重要度の高い画像が大きく配置される。そのため、画像の重要度を評価し、LLM にプロンプトとして与えることで画像の重要度を考慮したレイアウトを作成することができる。ただし、従来の画像評価モデルは、景観や構図などの美的評価を基準とした評価を行うため、人物の躍動感などを重要度の基準とする卒業アルバムには適さない。そのため、卒業アルバムに特化したデータセットを用いて学習する必要がある。

画像評価用のデータセットを作成するには、大量の画像に対し手作業で評価を与えることが必要となる。しかし、レイアウトが日々の業務と並行しこの作業を行うことは困難である。そこで、本手法ではレイアウト中の画像の相対的な面積を画像の評価とすることで、自動的に画像評価用のデータセットを作成する。図4に画像評価モデルの学習の流れを示す。

本手法では、レイアウト中の画像の大きさを基に、画像の重要度を算出する。ページに含まれる任意の画像の面積 $A_i$ とページに含まれる画像の平均面積 $A_{mean}$ の比を求めることで重要度 $I_i$ を算出する。また、画像の重要度は0.1-10の範囲に制限している。式(1)にスコアの算出式を示す。

$$I_i = \frac{A_i}{A_{mean}} \quad (1)$$

式(1)からわかるように、各画像の重要度 $I_i$ はページごとの画像の平均面積 $A_{mean}$ に依存する。対象の画像の面積が平均値よりも大きければ、その画像の重要度は高くなる。逆に、平均値よりも小さければ重要度は低くなる。

学習段階では式(1)から算出したスコアを使用する。しかし、レイアウトにおける未知画像の面積は不明であるため、先ほどのスコアを算出することができない。そこで、レイアウトのデータを学習した画像評価モデルを用いて、未知画像の重要度を算出する。

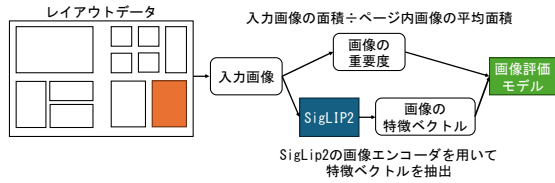


図4 画像評価モデルにおける学習の流れ

Fig.4 Learning process in image evaluation model

本手法では、画像評価モデルに Aesthetic Predictor v2<sup>10)</sup> を使用する。このモデルの選定理由として、画像からの特徴抽出に画像言語モデルのエンコーダを用いていることが挙げられる。3.4 節に述べる画像のクラスタリングでは、画像言語モデルから抽出した特徴量を入力として用いるため、特徴量を流用することにより処理の効率化を図ることができる。また、4 層の線形層と 3 層のドロップアウト層で構成されていることから、処理が高速であるという特徴がある。これらの点は、多くの画像を扱う卒業アルバム制作において適している。

学習データとして、レイアウトから抽出した画像及び式(1)から算出したスコアのペアを用いることで、未知データからスコアを推定するモデルを作成する。また、最先端の画像言語モデルである SigLIP2<sup>11)</sup> を用いて画像からの特徴抽出を行う。そして、算出したスコアは xml 形式のレイアウト内に score 属性として埋め込む。

### 3.4 画像のクラスタリング

卒業アルバムの特性として、被写体の人数に比例し、画像枚数が多くなるという点が挙げられる。場合によっては、1 つのレイアウトに画像が 60 枚ほど含まれることもある。そのため、画像間の重なりを考慮しつつ、全ての画像の位置や大きさを一度に予測することは難しいと考えられる。

この点を踏まえ、画像のクラスタリングを行い、類似する画像同士をグループ化することでレイアウト全体を構造化する。そして、クラスタ単位で領域を予測した後、それぞれの領域内に画像を配置するというアプローチをとる。これにより、全体ではなくクラスタ単位で画像の配置を学習できる。

本手法では、画像言語モデルを用いて抽出した特徴量に対し、BERTopic<sup>12)</sup> と同様の手法を用いてクラスタリングを行う。まず、次元削減手法の UMAP<sup>13)</sup> で特徴量の次元を削減した後、HDBSCAN<sup>14)</sup> でクラスタリングを行う。HDBSCAN の特徴として、事前にクラスタ数を指定する必要がなく、高精度である点が挙げられる。そのため、レイアウトターがクラスタ数を指定せずとも、類似する画像を高精度でクラスタリングできるという利点がある。モデルの入力には、画像評価モデルに使用した画像特徴量を使用し、処理の効率化を図っている。

### 3.5 教師ありファインチューニング

LLM に GRPO を適用させる前に、学習データを使用し教師ありファインチューニング (SFT) を行う。GRPO では、3.6

節のフォーマット報酬の解説で述べるように、既定のフォーマットに合致するよう学習が行われる。そのため、事前に SFT を行うことにより、あらかじめフォーマットに沿った出力を行うモデルを作成できる。そのモデルに対し GRPO を適用することにより、学習効率の改善を図ることができる。

### 3.6 GRPO によるレイアウトの最適化

GRPO は、DeepSeekMath<sup>5)</sup> で提案された、LLM に特化した強化学習手法である。この手法の特徴として、状態価値を求めるネットワークを不要としたことにより、効率的な学習を行える点が挙げられる。

卒業アルバムのレイアウトには「画像をできるだけ重ねない」などのルールが存在する。GRPO では、任意の形式で報酬関数を設定することができるため、それらのルールに従うレイアウトに高い報酬を与えることができる。このように、適切な報酬関数を設定することにより、目的に則したレイアウトを作成できると考えられる。レイアウト生成において、GRPO は LaySPA<sup>4)</sup> の学習に用いられている。

本手法では、クラスタ領域と画像配置ともに共通の報酬関数を設定する。クラスタ領域と画像生成のタスクで大きく異なる報酬関数を用いた場合、それぞれのタスクごとに個別のモデルを学習する必要がある。その場合、モデルの切り替えが必要になるため、処理の非効率化が懸念される。そのため、本手法では共通の報酬関数を用いることで、1 つのモデルに対しクラスタ領域と画像配置の学習を行っている。

また、xml のフォーマットが誤っていた場合には、一律に報酬を 0 にする。なお、それぞれの報酬の最大値は 1 に制限している。矩形の重なりや面積などの一項目への過剰な適合を防ぐため、各報酬は重み係数 1.0 の等重みに設定している。フォーマット報酬は、Unsloth<sup>15)</sup> が公開している実装形式に基づく。以下に報酬関数の詳細を示す。

#### 形式に基づくフォーマット報酬

事前に設定したフォーマットに従う出力に報酬を与える。今回は、`<start_working_out>` と `<end_working_out>` の間に思考過程、`<solution>` と `</solution>` の間に xml 形式のレイアウトが生成された場合に報酬  $R_{format}$  を 1 与え、そうでない場合は報酬  $R_{format}$  が 0 となる。

#### 要素の出現回数に基づくフォーマット報酬

出力に `<start_working_out>`、`<end_working_out>`、`<solution>`、`</solution>` のいずれか 1 つが含まれるごとに報酬  $R_{format-cnt}$  を 0.25 与える。これにより、フォーマットに従う出力を促す。

#### 矩形の重なり面積に基づく報酬

画像の過剰な重なりが生じると、被写体の顔などの重要な情報が隠れてしまう可能性が懸念される。このような状況を防ぐため、卒業アルバムでは画像の重なりを避けるように配置を行う必要がある。この点を踏まえ、予測された矩形の面

積 $A_{pred}$ と、正解の矩形の面積 $A_{true}$ の間で生じた重なりを基に、報酬 $R_{overlap}$ を算出する。式(2)に報酬 $R_{overlap}$ の算出式を示す。

$$R_{overlap} = 1 - \frac{A_{max} \cap A_{min}}{A_{min}} \quad (2)$$

ここで、式(2)に示した $A_{min}$ は $\min(A_{pred}, A_{true})$ から、 $A_{max}$ は $\max(A_{pred}, A_{true})$ から求められる。 $A_{min}$ と $A_{max}$ が重なり合った面積と $A_{min}$ の比を求めることにより、重なりが矩形のどの程度を占めるかについて判定できる。この値を報酬 $R_{overlap}$ とすることにより、画像やクラスタの重なりが少ない配置の学習を促進する。図5に報酬 $R_{overlap}$ の算出課程を示す。

#### 矩形の面積に基づく報酬

本手法では、画像から求めた重要度や、被写体の人数をもとに画像の大きさを予測する。そして、予測された矩形の面積 $A_{pred}$ と、正解の矩形の面積 $A_{true}$ を用いて、どの程度予測結果の面積が正解に近いかを評価する。この比を報酬 $R_{ratio}$ とすることで、画像やクラスタ領域の正確な面積を予測させる。式(3)に報酬 $R_{ratio}$ の算出式を示す。

$$R_{ratio} = \frac{\min(A_{pred}, A_{true})}{\max(A_{pred}, A_{true})} \quad (3)$$

#### 領域外へのはみ出しの有無に基づく報酬

卒業アルバムのレイアウトにおいて、画像のはみだしは美観を損ねるため、レイアウトの領域外に画像がはみ出さないように配置を行う必要がある。予測した矩形の面積 $A_{pred}$ と、レイアウトの領域からはみ出た矩形の面積 $A_{overflow}$ から、領域外へのはみ出しに基づく報酬 $R_{overflow}$ を算出する。このように、領域内に配置された要素へ報酬を与えることにより、レイアウトの範囲内への配置を促進する。式(4)に報酬 $R_{overflow}$ の算出式を示す。

$$R_{overflow} = 1 - \frac{A_{overflow}}{A_{pred}} \quad (4)$$

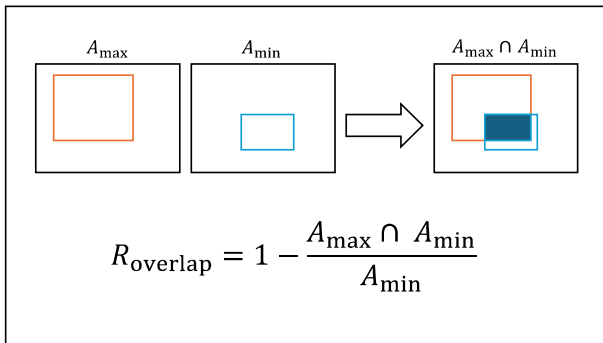


図5 重なりに基づく報酬の算出

Fig.5 Calculation of rewards based on overlap

## 4. 評価実験

本手法の精度検証を行うため、卒業アルバムのレイアウトから構成されたデータセットを使用し実験を行う。本実験で使用するレイアウトは、ダイコロ株式会社<sup>8)</sup>の協力のもとで収集した。学習データに対しては、手動でアノテーションを行い、推論時にはクラスタリングを用いて自動でクラスタを推定する。

また、本実験では LLM のモデルとして、Qwen3-4B<sup>16)</sup>を選定した。このモデルは、数学やコーディングのタスクにおいて優れており、6 億から 320 億までのパラメータを持つモデルが公開されている。本システムでは個人情報を扱うため、ローカル環境で実行できる小型のモデルが好ましいが、パラメータ数が多いほど性能が向上する傾向にある。

これらの点を考慮し、本システムでは 40 億パラメータを持つ Qwen3-4B を選定した。このモデルは、同規模のパラメータを持つ Gemma-3-4B<sup>17)</sup>や Qwen2.5-7B<sup>18)</sup>と比較し、数学やコーディングのタスクにおいて高い精度を達成している。また、このモデルは商用利用が可能な Apache 2.0 ライセンスである。

実験内容として、Qwen3-4B に教師ありファインチューニング (SFT) 及び GRPO による学習を行い、レイアウト領域生成と画像配置の結果から、正解レイアウトとの面積比 (ratio) と重なり (overlap) 及びレイアウトの範囲内への配置 (inside) を求めることで精度を測る。ratio と inside の算出は、3.6 章に示した式(3)と式(4)から、overlap は式(5)から算出する。ここで、 $A_{sum-rect}$ はクラスタ若しくは画像の重なりを含まない面積の合計、 $A_{layout-rect}$ は重なりを含む面積の合計を示す。

$$overlap = 1 - \frac{A_{layout-rect}}{A_{sum-rect}} \quad (5)$$

そして、各報酬関数を取り除いたモデルを学習し、それぞれの表指標に対して比較を行う。このアブレーション実験を通して、各報酬関数がどのような働きを行っているかについて確認する。

また、従来手法との比較実験では、最先端のフォトコラージュ生成手法である SoftCollage<sup>6)</sup>との比較を行い、生成されたレイアウトについて考察を述べる。比較検討用の画像データには、フォトコラージュ用に特化した AIC データセット<sup>6)</sup>を用いる。

最後に、本システムの現段階における実用性を確認するため、ダイコロ株式会社に所属する 4 名のレイアウトターの協力のもと、5 段階の尺度に基づいた主観的な評価実験を行う。そして、得られた結果を基に、本システムの性能に関する考察を行い、改善すべき点について述べていく。

#### 4.1 実験設定

モデルのトレーニングには LLM の学習用ライブラリである Unsloth<sup>15)</sup>を用い、SFTによる事前学習は計2エポック実行する。GRPOによる学習は計3エポック実行する。学習及び精度評価に使用するレイアウトは、あらかじめクラスタリングや重要度の算出などの前処理を行っている。

クラスタ領域の学習には計1112枚、精度評価には計145枚のレイアウトを使用した。そして、画像配置の学習及び精度評価には、レイアウトのクラスタから切り出した部分レイアウトを用いる。また、画像評価モデルの学習には、学習用の画像データのみを用いている。

レイアウトの主観的な評価実験では、5段階の評価尺度のもと、4名のレイアウトターが主観的な評価を行う。評価項目は、「類似する画像が適切に配置されていますか?」、「写真に写った人物の大きさは適切ですか?」の2項目である。評価対象のレイアウトは、ダイコロ株式会社の協力のもとで収集した合計349枚の卒業アルバムのサンプル画像を使用して生成する。また、卒業アルバムではレイアウトごとに使用する画像の枚数が変化する。画像枚数の変化による評価の違いを検証するため、この画像からランダムに30枚、40枚、50枚、60枚の復元抽出を3回ずつ行い、計12枚のレイアウトを生成する。これらの画像枚数の変化に対する評価の違いについても比較を行う。

#### 4.2 実験結果

表1にクラスタ領域生成の実験結果を示す。これらの評価指標は、overlapのみ0に近いほど評価が高くなる。また、最も良好な結果を太字で示している。

表1から、各評価項目に対応した報酬関数を除いたモデルでは、精度が低下することが確認できる。その例として、ratioに関連する報酬関数を除いたモデルでは、ratioの値が他のモデルと比較し1/2ほどとなっている。

表1 クラスタ領域生成の実験結果

Table 1 Experimental results on cluster region generation

報酬関数	ratio	overlap	inside
ratio+overlap+inside	<b>0.906</b>	0.014	0.999
ratio+overlap	0.898	0.005	0.999
ratio+ inside	0.901	0.019	<b>1.0</b>
overlap+inside	0.450	<b>0.0</b>	<b>1.0</b>

表2 画像配置の実験結果

Table 2 Experimental results of the image placement

報酬関数	ratio	overlap	inside
ratio+overlap+inside	0.767	0.085	0.983
ratio+overlap	<b>0.768</b>	0.087	0.971
ratio+ inside	<b>0.768</b>	0.236	0.989
overlap+inside	0.043	<b>0.007</b>	<b>0.991</b>

また、全ての報酬関数を用いて学習したモデルでは、最高値と最低値が0.001と極めて差が少ないinsideを除く評価項目において最も低い精度を記録していないことから、全体的に良好な精度を達成していることが確認できる。表2に画像配置の実験結果を示す。表1と同様に、最も良好な結果を太字で示している。

表2からも表1と同様に、全ての報酬関数を用いたモデルが良好な精度を達成しており、各報酬関数を除いたモデルでは、対応する評価項目が最も低い精度となっていることが確認できる。また、タスクごとの生成結果を図6に示す。また、本手法で生成したレイアウトとSoftCollageを用いて生成したレイアウトを図7に示す。本実験では、含まれる画像枚数を変化させつつレイアウトを生成している。そして、レイアウトターによる主観的な評価の結果を示す。全体の評価結果を図8に、枚数ごとの評価結果を図9に示している。

画像全体の評価実験の結果、「類似する画像が適切に配置されていますか?」の平均値が2.46、「写真に写った人物の大きさは適切ですか?」の平均値が2.02であることが確認できた。

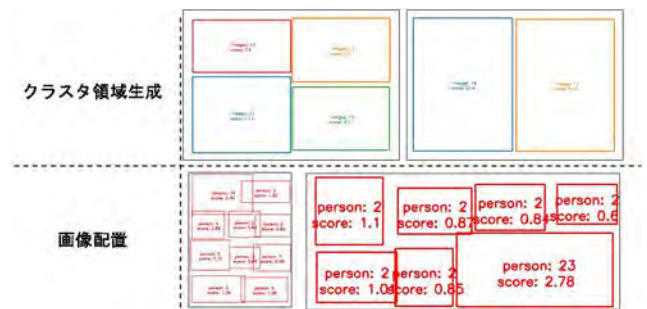


図6 タスクごとの生成結果

Fig.6 Generation results for each task



図7 生成結果の比較

Fig.7 Comparison of generation results

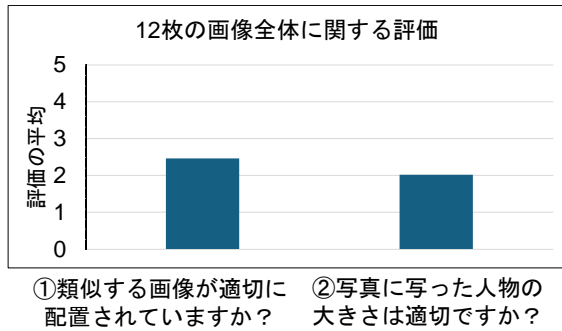


図8 評価項目ごとの評価結果の平均値

Fig.8 Average score for each evaluation criterion

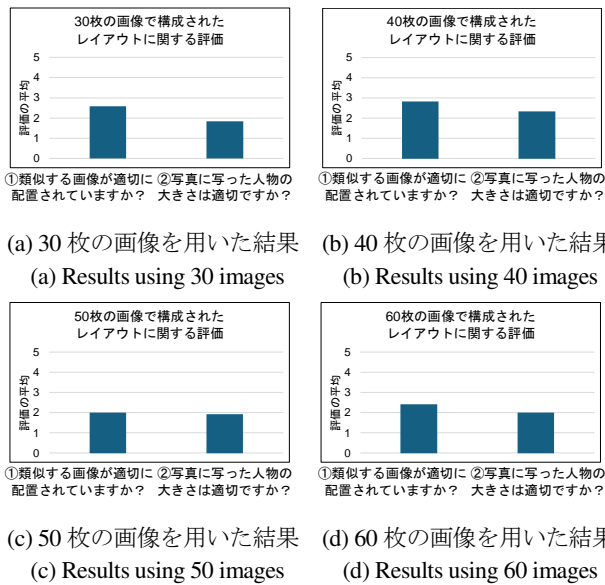


図9 画像枚数ごとの評価結果の平均値

Fig.9 Average evaluation results by different number of images

また、画像枚数ごとの評価結果では、「類似する画像が適切に配置されていますか?」の評価が最も高かったものは、40枚の画像を使用したレイアウトでその平均値は2.83、最も低かったものは50枚の画像を使用したレイアウトでその平均値は2.00であり、両者の差は0.83であった。そして、「類似する画像が適切に配置されていますか?」の評価が最も高かったものは、40枚の画像を使用したレイアウトでその平均値は2.33、最も低かったものは30枚の画像を使用したレイアウトで、その平均値は1.83であり、両者の差は0.50であった。

### 4.3 考察

表1から、クラスタ領域生成では、ratioの値は0.906と、正解のクラスタの面積から大きな乖離はみられなかった。また、overlapは0.014と、重なりがほぼ生じていないことが確認できる。画像配置ではratioが0.768となり、また、overlapは0.085に抑えられている。そして、双方ともにinsideが0.98以上と、ほとんどの要素がレイアウトの範囲内に配置されて

いることがわかる。また、アブレーション実験の結果からは、各報酬関数が対応する評価項目に対して性能を向上させていることが確認できた。そのため、今後レイアウトの要望により制約条件が追加されたとしても、報酬関数を追加することで対応可能だと考えられる。

また、図6から、画像配置とクラスタ領域生成ともに、要素間の重なりを抑えつつ、レイアウトの範囲内に配置を行っていることがわかる。これらの点は、GRPOの報酬関数による学習が有効に作用したと推察される。また、GRPOで学習したモデルにより生成されたレイアウトは、含まれるスコアが高い要素ほど大きく配置される傾向にあることがわかる。これは、スコアが高い画像ほどレイアウトにおける重要度が高いという特性を反映しているといえる。

図7を確認すると、本手法で生成したレイアウトは、類似した画像同士を1つのクラスタにまとめていることがわかる。また、算出した重要度に比例して、画像が大きく配置されていることが確認できる。それに対し、SoftCollageで生成したレイアウトでは画像同士が隙間なく配置されていることがわかる。その理由としては、SoftCollageは二分木ベースでページを分割し、その領域に画像を当てはめる方式でレイアウトを生成している点が挙げられる。ただし、卒業アルバムでは、画像を引き立たせるために意図的に余白を持たせるため、画像間の空白を確保することは必要である。また、双方とも画像枚数に関わらず破綻のないレイアウトを作成できている。

これらの点から、本手法では画像間の関連性を捉えつつ、画像の重要度も取り入れたレイアウトを作成できることを確認できた。また、本実験では卒業アルバムの画像のみを学習し、重要度を算出するモデルを学習したため、さらなる画像データを学習することにより、精度が改善する可能性がある。

次に、レイアウトによる主観的な評価についての考察を記す。評価実験の結果、「類似する画像が適切に配置されていますか?」の項目は全体的に2~3程度の評価が多いことが確認できた。そのため、本手法による画像配置はレイアウトにより一定の評価を得られたものの、実用化に向けてさらなる改善が必要であると考えられる。

それに対し、人物の大きさの評価項目は画像の配置の評価項目より評価がやや低い傾向にあった。その原因としては、本手法では面積ではなく人数のみを考慮している点が考えられる。卒業アルバムでは被写体の顔の大きさをそろえることが好ましいため、画像に占める人物の面積に関する報酬関数を追加し、精度を向上させることが必要である。

画像枚数による評価に関しては、双方の評価項目ともに画像枚数ごとの評価の差が1以下に収まっていたことからわかるように、画像枚数によって評価に大きな差が生じないことが確認できた。卒業アルバムのレイアウトでは、生徒の人数などの要因により画像の枚数が大きく変化する。そのため、画像枚数に評価が大きく左右されることのない結果は好ましいものであると言える。

また、「写真と写真の隙間が不ぞろい」であるというコメントが寄せられたが、間隔を統一する報酬関数を追加することで対応可能である。これらの結果から、画像のクラスタリングにおいて一定の評価を得られたものの、実際の制作現場で運用するためには、今回の評価実験で得られた結果を基に、適切な報酬関数を設定することが必要であると考えられる。

以上から、卒業アルバムに向けた報酬関数を用いて、GRPOによる学習を行うことで適切なレイアウトを生成できることが確認できた。また、本実験では、HDBSCANを用いて自動的にクラスタ数を決定したが、本システムがレイアウトターと共同で卒業アルバムを制作するためのものであることを考えると、レイアウトターがクラスタ数を指定する機能を追加することが好ましいと推察される。K-means++<sup>19)</sup>などの事前にクラスタ数を指定するクラスタリング手法を使用することで、レイアウトターの意図を反映することができる。このように、複数のクラスタリング手法を選択する機能を追加することを検討している。また、本実験では行わなかったものの、クラスタ領域生成、若しくは画像配置のどちらかをレイアウトターが手動で行うことも可能である。このように、本手法では高い自由度でレイアウトを作成することができる。

## 5. おわりに

本研究では、GRPOやクラスタリングなどの技術を用いて、卒業アルバムのレイアウト生成を行う手法を提案した。本手法では、入力した画像に対して属性を抽出した後、クラスタリングを行うことで画像間の関係性を構造化している。これに基づき、GRPOにより学習を行ったLLMを用いて、適切なレイアウトを作成できるようにしている。

実験では、実際の卒業アルバムのレイアウト及びAICデータセットを用いて、本手法の精度検証と生成結果の視覚的な評価を行った。その結果、GRPOによる卒業アルバムの性質の学習及びクラスタリングの効果により、画像同士の重なりを約8%に抑えつつ、適切なクラスタ領域内への画像の配置を行えることが分かった。クラスタ領域の生成に関しては、クラスタ間での重なりを約1%に抑えられることが分かった。また、従来手法と比較したところ、クラスタの特徴を捉えつつ、重なりが少ないレイアウトを作成できることを確認できた。また、主観的な評価実験の結果を基に、システムの実用化に向けた改善を予定している。そして、本研究では1社のみを対象にした評価実験を行っているため、複数企業の協力のもと、更なる評価検証を行うことを検討している。

## 謝辞

本研究では、ダイコロ株式会社よりレイアウトデータの提供や、卒業アルバム制作に関する助言を頂きました。皆様のご支援に心より感謝申し上げます。

## 参考文献

- 1) Z. Tang, C. Wu, J. Li, N. Duan: "LayoutNUWA: Revealing the Hidden Layout Expertise of Large Language Models", The Twelfth International Conference on Learning Representations (ICLR 2024), pp. 38369-38388 (2024).
- 2) J. Tian, X. Yu, Y. Wang, P. Wang, G. Guo, S. Gao: "ReLayout: Integrating Relation Reasoning for Content-aware Layout Generation with Multi-modal Large Language Models", arXiv preprint, arXiv:2507.05568 (2025).
- 3) J. Wei, X. Wang, D. Schuurmans, M. Bosma, B. Ichter, F. Xia, E. Chi, Q. Le, D. Zhou: "Chain-of-Thought Prompting Elicits Reasoning in Large Language Models", Proc. of the 36th International Conference on Neural Information Processing Systems (NIPS 22), pp. 24824-24837 (2022).
- 4) S. Li: "LLMs as Layout Designers: A Spatial Reasoning Perspective", arXiv preprint, arXiv:2509.16891 (2025).
- 5) Z. Shao, P. Wang, Q. Zhu, R. Xu, J. Song, X. Bi, H. Zhang, M. Zhang, Y.-K. Li, Y. Wu, D. Guo: "DeepSeekMath: Pushing the Limits of Mathematical Reasoning in Open Language Models", arXiv preprint, arXiv:2402.03300 (2024).
- 6) J. Yu, L. Chen, M. Zhang, M. Li: "SoftCollage: A Differentiable Probabilistic Tree Generator for Image Collage", Proc of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2022), pp.3729-3738 (2022).
- 7) M. Zhang, M. Li, L. Chen, J. Yu: "Aesthetic Photo Collage with Deep Reinforcement Learning", arXiv preprint, arXiv:2110.09775 (2021).
- 8) ダイコロ株式会社, ダイコロ株式会社【卒業アルバム スクールアルバム】, <https://www.daicolo.co.jp/> (2025).
- 9) W. Wu, H. Peng, S. Yu: "Yunet: A tiny millisecond-level face detector", Machine Intelligence Research, Vol. 20, No. 5, pp. 656-665 (2023).
- 10) LAION, LAION-AESTHETICS, <https://laion.ai/blog/laion-aesthetics/> (2022).
- 11) M. Tschannen, A. Gritsenko, X. Wang, M. F. Naeem, I. Alabdulmohsin, N. Parthasarathy, T. Evans, L. Beyer, Y. Xia, B. Mustafa, O. Hénaff, J. Harmsen, A. Steiner, X. Zhai: "SigLIP 2: Multilingual Vision-Language Encoders with Improved Semantic Understanding, Localization, and Dense Features", arXiv preprint, arXiv:2502.14786 (2025).
- 12) M. Grootendorst: "BERTopic: Neural topic modeling with a class-based TF-IDF procedure", arXiv preprint, arXiv:2203.05794 (2022).
- 13) L. McInnes, J. Healy, J. Melville: "UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction", arxiv preprint, arXiv:1802.03426 (2018).
- 14) R. J. G. B. Campello, D. Moulavi, J. Sander: "Density-Based Clustering Based on Hierarchical Density Estimates", Pacific-Asia conference on knowledge discovery and data mining (PAKDD 2013), pp. 160-172 (2013).
- 15) GitHub, Unsloth, <http://github.com/unslothai/unsloth> (2023).
- 16) Qwen Team: "Qwen3 Technical Report", arXiv preprint,

arXiv:2505.09388 (2025).

17) Gemma Team: "Gemma 3 Technical Report", arXiv preprint, arXiv:2503.19786 (2025).

18) A. Yang, B. Yang, B. Zhang, B. Hui, B. Zheng, B. Yu, C. Li, D. Liu, F. Huang, H. Wei, et al.: "Qwen2.5 Technical Report", arXiv preprint, arXiv:2412.15115 (2024).

19) D. Arthur, S. Vassilvitskii: "k-means++: the advantages of careful seeding", Proc. of the eighteenth annual ACM-SIAM symposium on Discrete algorithms (SODA '07), pp.1027-1035 (2007).

(2025年12月2日 受付)  
(2026年4月1日 再受付)



松田 展 (学生会員)

2022年 大阪工業大学情報メディア学科卒業。現在 大阪工業大学大学院情報科学研究科 博士後期課程在籍。画像処理及び自然言語処理に関する研究に従事。



佐野 睦夫 (正会員)

1983年 京都大学大学院工学研究科修士課程修了。同年4月 電電公社(現NTT)入社。パターン認識, ロボットビジョンやコンテンツ流通の研究開発に従事。2002年 大阪工業大学情報科学部教授。コミュニケーション科学, HCI, 行動科学等の研究に従事。現在に至る。画像電子学会, 電子情報通信学会, 情報処理学会, IEEE など正会員。博士(工学)



葉山 恒生

2011年 ITベンチャーキャピタル会社を経てダイコロ株式会社に入社。現在 IT事業本部所属。同社にて執行役員本部長を務め思い出関連サービス(「思い出メーカー」)のIT事業責任者として産官学連携による共同開発を推進。



大賀 将弘

ダイコロ株式会社 IT事業本部所属。アルバム制作および写真販売領域において顔認証およびAI技術を活用したサービス開発に従事。

## 髪型の特徴が国籍分類に与える影響：深層学習による東アジア顔画像の分析

永岡 正梧<sup>†</sup> 張 珏<sup>†</sup> ナイワラ P. チャンドラシリ<sup>†</sup>(正会員)<sup>†</sup>工学院大学 工学研究科 情報学専攻

## Hairstyle Matters: Deep Learning-Based Analysis of Nationality Classification in East Asian Facial Images

Shogo NAGAOKA<sup>†</sup>, Jue ZHANG<sup>†</sup>, Naiwala P. CHANDRASIRI<sup>†</sup>(Member)<sup>†</sup>Infomatics Program, Graduate School of Engineering, Kogakuin University

〈あらまし〉本研究では、日本・中国・韓国の三カ国の顔画像を対象に、年代や性別といった属性情報を考慮したうえで、ResNet-101, AlexNet, VGG-19 の 3 種類の事前学習済みモデルに転移学習を適用し、国籍分類を行った。その結果、ResNet-101 では 86.9% という高い分類精度を達成した。さらに、本研究では従来あまり注目されてこなかった「髪型」に国籍ごとの特徴が存在するという仮説を立て、顔全体だけでなくパーツごとに分割した画像で学習を実施した。また、モデルが着目している領域を可視化するために Grad-CAM を用い、さらに主成分分析 (PCA) による特徴解析も行った。その結果、髪型のみを用いた分類精度が目や口といった他のパーツとほぼ同等であることが示され、国籍分類において髪型が顔の他部位と同様に重要な情報源となる可能性が示唆された。

キーワード：国籍分類, 転移学習, 領域抽出, 髪型特徴

<Summary> This study investigates nationality classification based on facial images of individuals from Japan, China, and Korea. We applied transfer learning to three pre-trained convolutional neural networks—ResNet-101, AlexNet, and VGG-19—while considering age and gender as additional factors. Among these models, ResNet-101 achieved the highest classification accuracy of 86.9%. Furthermore, assuming that hairstyles possess nationality-specific characteristics, which have been largely overlooked in prior research, we conducted additional experiments focusing on individual facial regions. Using Grad-CAM visualization and Principal Component Analysis (PCA), we analyzed the distinctive features contributing to classification. The results indicate that the accuracy of nationality classification based solely on hairstyle was comparable to that based on the eyes and mouth, suggesting that hairstyle plays a role as important as other facial features in identifying nationality.

**Keywords:** classification of nationality, transfer learning, region extraction, hairstyle features

## 1. はじめに

近年、日本を訪れる外国人観光客の数は増加傾向にある。なかでも、日本、中国、韓国の三カ国は地理的にも文化的にも関係が深く、これらの地域からの渡航者や就労者の往来は特に盛んである。一方で、この三国籍の人々の顔立ちは互いに類似していることが多く、国籍による分類は容易ではないと指摘されている<sup>1)</sup>。顔認識技術は、目・鼻・口といった顔の主要パーツや輪郭情報をもとに分類を行うことが一般的である。しかし近年、韓国や中国では、美容整形が社会的に広く普及しており、顔の形態的特徴が均質化する傾向があるとされている。このような状況では、国籍分類において従来重視されてきた顔パーツそのものの差異が、分類の有効な手がかりになりにくい可能性がある。一方、

髪型スタイルのトレンドは、各地域の文化的背景や流行の影響を受けやすいため、結果として集団的な視覚的特徴となる。そのため髪型は国籍を分類する際の重要な補助的特徴になると考えられる。近年の顔認識分野の研究では、髪型に含まれる高周波成分が分類モデルの判断要素として重要な役割を果たす可能性が示唆されている<sup>2)</sup>。さらに、深層学習技術の進歩によって、画像認識の精度と応用範囲は飛躍的に拡大している<sup>3-5)</sup>。顔認識技術はその代表例であり、監視カメラやスマートフォンの顔認証機能など、私たちの日常生活のさまざまな場面に組み込まれている。

この技術を活用することで、性別や年齢といった個人属性を推定するシステムがすでに実用化されており、人の属性情報を取得する一つの有力な手段となっている<sup>6)</sup>。さらに、これらのシステムに国籍分類機能を組み込むことで、デジタル展示の内容を利用者の国籍

に応じて動的に変化させるといったサービスの発展が可能となる。加えて、空港や駅における入国審査の効率化、防犯カメラによる犯罪捜査の支援、多言語対応システムの自動化など、多様な分野への応用が期待される。そこで本研究では従来あまり注目されてこなかった「髪型」に国籍ごとの特徴が存在するという仮説を立て、顔全体だけでなくパーツごとに分割した画像で学習を実施した。また、モデルが着目している領域を可視化するために Grad-CAM を用い、さらに主成分分析(PCA)による特徴解析も行った。その結果、髪型のみを用いた分類精度が目や口といった他のパーツとほぼ同等であることが示され、国籍分類において髪型が顔の他部位と同様に重要な情報源となる可能性が示唆された。

## 2. 関連研究

Yu Wang ら<sup>1)</sup>は、事前学習済みのニューラルネットワークを活用し、X (旧 Twitter) 上から収集した日本・中国・韓国の顔画像を対象として、各種顔パーツに着目した国籍分類を行った。データ収集においては、まず対象とする国籍の著名人のフォロワー情報を取得し、ユーザーの投稿言語から母語を推定することで高精度なラベル付けを実現した。その後、OpenCV による顔検出処理と 12KB 以上の画像サイズによるフィルタリングを適用し、ノイズを除去した上で、最終的に約 4 万枚のラベル付き画像データセットを構築している。これらのデータを用いて VGG 及び ResNet ベースの CNN をファインチューニングし、ネットワークの深さと分類精度の関係を検証した結果、特に ResNet-50 を用いたモデルでは 3 クラス分類において 75.03% のテスト精度が得られ、これは人間による平均分類精度 (38.89%) を大きく上回る性能であることが示された。

Huber ら<sup>2)</sup>は、顔認識モデルにおける人口統計的バイアスの要因を明らかにするため、画像の周波数成分に着目した分析を行った。顔画像をフーリエ変換し、低・中・高周波帯域をマスクしてスコア変化を評価した結果、人種グループによってモデルが依存する周波数帯域が異なることを示した。特に、高周波成分には髪型や髪質などの特徴が含まれ、分類に大きく寄与していることが確認された。

永岡ら<sup>3)</sup>は、日本・中国・韓国の顔画像を用いて国籍分類を実施したが、髪型に関しては十分な考察が行われていなかった。本研究では、追加の評価実験を含む分析を通じて、国籍分類において髪型が顔の他部位と同様に重要な情報源として寄与する可能性が示唆された。

## 3. 提案手法

### 3.1 画像データセットの作成

本研究では、日本、中国、韓国の三カ国の顔画像を対象として解析を行った。しかし、年齢や性別などの属性情報が明確に付与された既存の公開データセットは入手が困難であったため、年齢及び国籍情報の信頼性を確保できる点と髪型のトレンドの最先端が芸能人であるという点から、対象となる

画像はすべて芸能人のものに限定し、出典が明確な 4 つの Web サイトから取得し<sup>8)~11)</sup>、独自にデータセットを構築した。収集した顔画像は各国 400 枚、合計 1,200 枚であり、国籍のみならず性別及び年代の分布が偏らないように調整を行った<sup>12)</sup>。本研究における「国籍」とは、法的な国籍ではなく、文化的背景により形成された外見特徴がどの国に近いかという観点から定義している。

年代別の内訳は、少年期 (5~14 歳) 20 枚、青年期 (15~24 歳) 60 枚、壮年期 (25~44 歳) 60 枚、中年期 (45~64 歳) 40 枚、高年期 (65 歳以上) 20 枚とし、男女それぞれ 200 枚ずつを収集した。また、顔の大きさを統一し、ランドマーク検出に基づくアライメント処理を実施し、画像データの正規化を行った。

### 3.2 転移学習による学習

本研究では、深層学習の手法の 1 つである転移学習を用いて国籍分類を実施した。転移学習とは、大規模データセットであらかじめ学習されたモデルを別のタスクへ応用する深層学習技術の一種であり、限られたラベル付きデータであっても、短時間かつ低い計算コストで高精度なモデルを得ることが可能である<sup>13)</sup>。

まず、顔画像による国籍分類に適した事前学習済みネットワークを特定するため、複数モデルの比較を行った。学習対象となる顔画像は MATLAB の顔検出モデルを用いて抽出し、データは学習用と検証用に 3:1 の割合でランダム分割した上で、4 分割交差検証を適用した。

転移学習の条件は、エポック数を 100、学習率を 0.0003 に統一して実験を行った。比較対象とする事前学習済みモデルには、ResNet-101<sup>14)</sup>、AlexNet<sup>15)</sup>、VGG-19<sup>16)</sup> の 3 種類を採用した。ResNet-101 は先行研究でも使用されており、さらに学習時間や精度特性が異なる AlexNet 及び VGG-19 を比較対象として加えることで、モデル間の性能差を検証できる構成とした。

### 3.3 Grad-CAM の適用

髪型が国籍判別において有効かを判断するために、他の顔パーツと比較をする必要があった。そのため、比較対象とする顔パーツを決定するために、機械学習モデルがどの顔の部分に着目しているかを明らかにする Grad-CAM を使用した。本研究では 3.2 で得られた ResNet-101 の分類結果画像に Grad-CAM<sup>17),18)</sup>を適用した。

### 3.4 顔パーツ別分類実験

3.3 で得られた結果を踏まえ、モデルが分類時に注目している顔の部位と、本研究で着目する髪型領域との比較を行うため、まず MATLAB の Computer Vision Toolbox に含まれる特徴検出機能を用いて、各顔パーツを抽出し、それらを転移学習の入力データとして使用した<sup>19)</sup>。この手法による髪型領域の抽出は一定の精度で可能であったが、さらに抽出精度を高めるために、Python を用いセマンティックセグメンテーションに特化した高精度・高速なニューラルネットワークである BiSeNet を導入し、髪型領域の抽出処理を実施した<sup>20),21)</sup>。

学習に使用するネットワークは、3.2 で最も高い精度を示した事

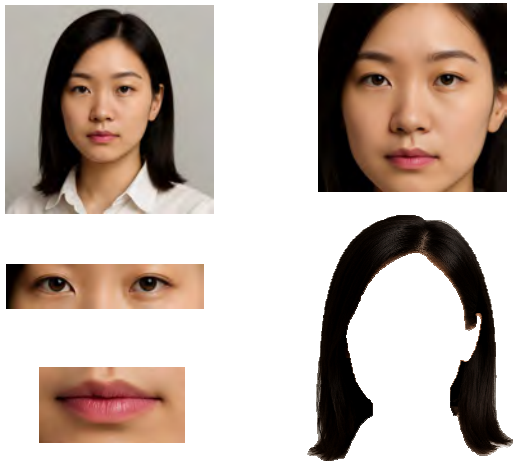


図1 各顔パーツの抽出画像例 (GPT-4o によって生成)

Fig.1 Extracted image examples of facial features (Generated by GPT-4o)

前学習済みモデルを採用した。画像データは3.2と同様に、学習データと検証データを3:1の比率でランダム分割し、4分割交差検証によって評価を行った。

図1に、各領域の抽出結果の一例を示す。これらはGPT-4oによって生成した画像を用いて処理したものである。本研究では髪型領域について、目や口の顔パーツと比較して、国籍分類における精度を分析した。各パーツの比較においては、パーツの位置関係は考慮せず、パーツ自体の形状とテクスチャーを特徴量として用い、同一条件で精度の比較を行った。

### 3.5 PCA (主成分分析) による特徴解析

本研究では、各国籍が持っている固有の髪型特徴を明示することを目的として、画素値による主成分分析 (PCA) を行った。本手法は、各国籍の髪型が具体的にどのような特徴を持つかを可視化・分析するために用いた。3.2, 3.3 で用いたデータセットを全て直接解析した場合、データ量が多く、国籍と主成分の対応関係が読み取りにくい。まず男女別 (男・女) と男女混合の三条件で髪型領域の平均顔画像を各国ごとに作成し、計9通りの平均顔画像を解析対象とした。この結果から三次元 PCA 空間上の分布を可視化し、各国の特徴量を考察した。第1主成分 (PC1)、第2主成分 (PC2)、第3主成分 (PC3) が具体的にどの髪型特徴量を表しているのかわかるようにするために、平均顔ベクトル  $y$ 、主成分ベクトル  $v$ 、各成分の標準偏差  $\sigma$ 、 $\pm\sigma$  した固有ベクトル  $I$  とし、これらを用いた式(1)を使用し、各主成分が捉える形態変動を  $\pm\sigma$  の変動として可視化した<sup>2)</sup>。

$$I = y \pm \sigma v \quad (1)$$

PCA で使用するために作成した髪型領域の平均顔画像を図2に示す。三国籍の表記としてそれぞれ日本を JPN、中国を CHN、韓

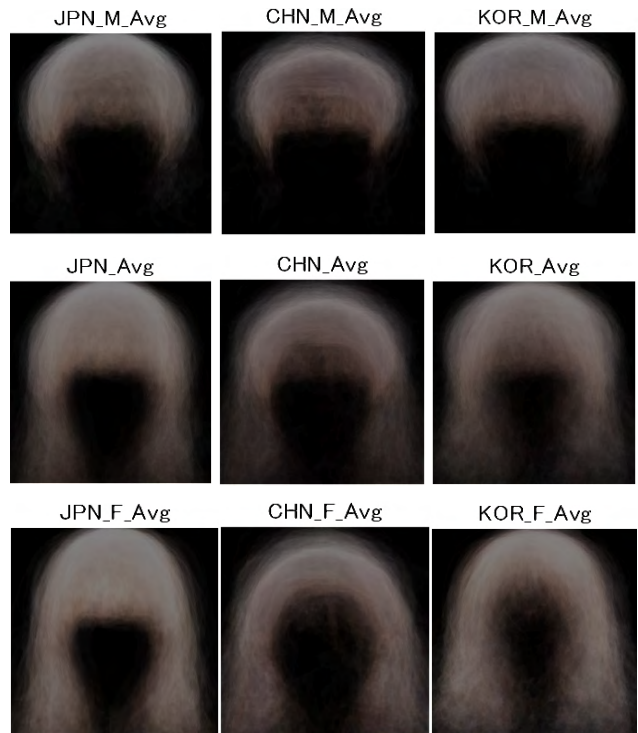


図2 髪型領域の国別(\_Ave)、国別男女別(M/F\_Ave)平均顔画像 9 枚

Fig.2 Nine average face images of the hairstyle region by country (\_Ave) and by gender within each country (M/F\_Ave)

国を KOR としており、M は男性のみで作成した平均顔画像、F は女性のみで作成した平均顔画像に対応している。

### 3.6 主観評価実験

転移学習による機械的分類に加え、人間による主観的判断において髪型が国籍分類に与える影響を検証するため、google フォームを用いて主観評価実験を実施した。本実験では、被験者に対して髪型領域を含む画像を提示し、その形態的特徴に対する印象評価と国籍推定を同時に行うことで、髪型形態と国籍イメージの関連性、モデル分類結果と人間の分類精度の比較、髪型のどの特徴が国籍分類に寄与するのかを明らかにすることを目的とした。また個人の文化的背景や経験による影響があるかも考察することを目的とした。

被験者には、髪型領域を抽出した画像をランダムな順序で提示し、国籍の推定と髪型特徴に関する4項目について7段階リッカート尺度で回答を求めた。項目は、前髪の有無、耳上から頭頂部のボリューム、耳から下の髪型のボリューム、全体的な髪型のすっきりさである。これらは3.5で考察された形態変動の結果から項目を設定した。回答は1が「とても弱い」、4が「どちらでもない」、7が「とても強い」を示す。また、回答前に被験者の属性 (海外生活の有無、他国籍メディア接触頻度、他国籍の人との接する機会) も取得し、個人属性による傾向差も分析可能とした。これらに加えアンケートの最後にどのような髪型特徴がそれぞれの国籍らしいかの回答も併せて求めた。被験者は現在日本国籍である20歳~24歳の21名

表 1 三カ国の分類精度

Table 1 Classification accuracy of three countries

	中国人	日本人	韓国人
AlexNet	85.0%	76.0%	67.0%
VGG-19	85.1%	70.0%	77.0%
ResNet-101	93.1%	88.0%	79.0%

であり、男性 18 名、女性 3 名である。海外生活経験有りとなされた被験者は 3 名おり、実際に住んでいた国は中国、アメリカ合衆国、インドとオーストラリアである。

#### 4. 実験結果

##### 4.1 転移学習による事前学習済みネットワークの比較結果

ResNet-101, AlexNet, VGG-19 の 3 種類の事前学習済みネットワークを用いた、三カ国別の分類精度をまとめたものを表 1 に示す。

ResNet-101 を用いた国籍分類の結果、平均精度は 86.9% となり、先行研究を上回る高い精度が得られた。国別の分類精度では、中国、日本、韓国の順に高く、韓国人のみ 79% と他の 2 国と比較してやや低い値を示したものの、全体としては既往研究よりも良好な結果となった。

一方、AlexNet を使用した場合、平均精度は 76.0% であり、先行研究とほぼ同等の水準であった。国別では中国、日本、韓国の順に精度が高く、日本人が 76.0%、韓国人が 67.0% と、いずれも高精度とはいえない結果であった。

また、VGG-19 を用いた結果では、平均精度は 77.6% となり、既存の報告よりわずかに高い値が得られた。国別の精度は中国、韓国、日本の順に高く、日本人は 70.0%、韓国人は 77.0% であり、他の 2 種類のモデルでは最も低かった韓国人の精度が VGG-19 では日本人を上回った点が特徴的である。

これら 3 種類のネットワークを比較した結果、平均精度・国別精度ともに ResNet-101 が最も優れた性能を示した。

##### 4.2 Grad-CAM 適用結果

4.1 で行った ResNet-101 を用いた分類での正解画像に Grad-CAM 適用した結果、学習において、目と口に特に注目をしていることが推察されたため、髪型との比較に使用するパーツは目と口を採用した。

##### 4.3 顔パーツのみを用いた学習

4.1 の結果から、ResNet-101 が最も高い分類精度を示したため、髪型を中心とした顔パーツ別の国籍分類を ResNet-101 を用いて実施した。髪型領域を抽出した画像を用いた分類結果と、目・口など他パーツを用いた結果を比較することで、髪型情報が国籍分類において持つ分類能力を定量的に評価した。各パーツでの分類精度を表 2 に示す。

元画像を用いて転移学習による国籍分類を行った結果、平均精度は 88.4% となり、先行研究を上回る性能が得られた。国別の分類精度では、日本、中国、韓国の順に高く、韓国人のみ 81.2% と他の 2

表 2 各パーツを用いた分類精度

Table 2 Classification accuracy by using each part

	中国人	日本人	韓国人
元画像	91.0%	92.0%	81.2%
髪型	89.0%	85.0%	75.0%
目	89.0%	95.0%	75.0%
口	86.1%	95.0%	70.0%

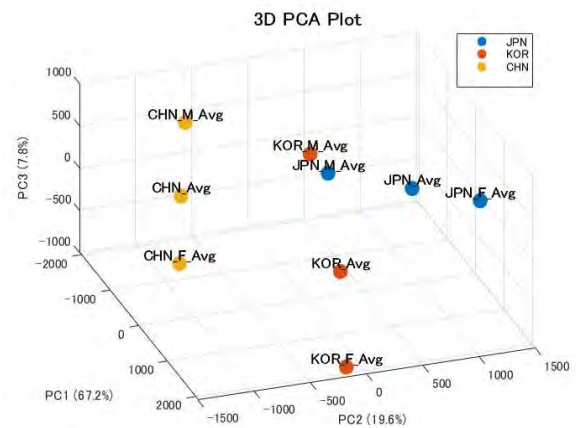


図 3 平均顔画像（髪型領域）での三次元 PCA 空間

Fig.3 Three-dimensional PCA space in average face images (hairstyle region)

国に比べやや低い結果であったが、全体として高い分類性能を示した。また、顔領域のみを用いた場合と比較して、日本人及び韓国人の分類精度が向上する傾向が見られた。

髪型領域のみを用いて同様の実験を行ったところ、平均精度は 82.8% となり、顔全体を使用した場合に比べてやや低い値となった。国別では中国、日本、韓国の順に精度が高く、韓国人は 75.0% と他の 2 国よりも低い結果を示した。

さらに、目領域のみを用いた場合、平均精度は 86.1% となり、顔全体での結果とほぼ同等の精度が得られた。特に日本人の精度は 95.0% と、全実験条件の中で最も高い値を示した。一方、韓国人の精度は 75.0% であり、髪型領域を用いた場合と同程度であった。

口領域のみを用いた場合の平均精度は 83.8% となり、顔全体よりもやや低いものの髪型領域と同程度の精度を示した。日本人の精度は目領域と同様に 95.0% と非常に高く、韓国人は 70.0% と最も低い結果となった。

##### 4.4 PCA による特徴分析結果

髪型領域画像の平均顔画像に PCA を行った結果、得られた三次元 PCA 空間を図 3 に示す。

髪型領域に基づく PCA の結果から、日本国籍は PC2 及び PC3 の両軸でプラス方向に大きく分布する傾向が見られた。一方、中国国

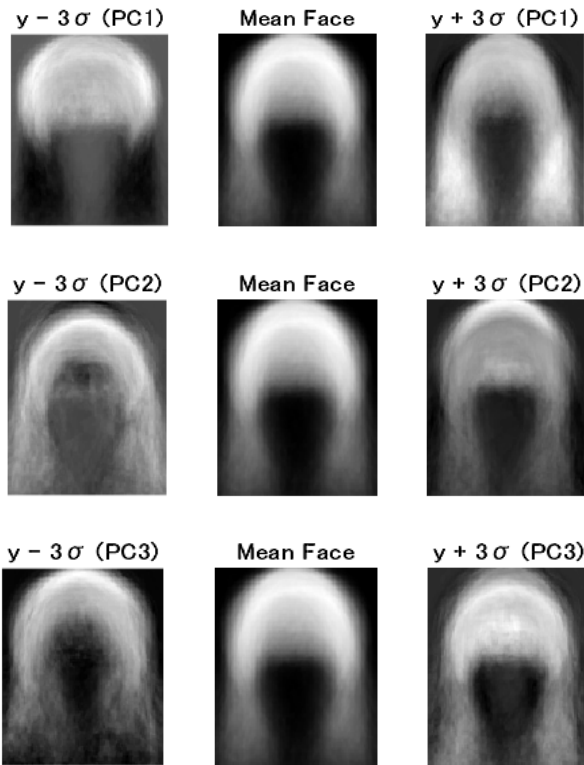


図4 髪型領域における固有ベクトル比較画像

Fig.4 Eigenvector comparison images in the hairstyle domain

籍ではPC2がマイナス方向, PC3がプラス方向に強く偏っており, 韓国国籍ではPC3が他の2国と比較してマイナス方向に大きく, PC2はほぼ中間的な位置に分布している. また, PC1に関しては, 顔領域を用いた場合とはプラス・マイナスの分布方向が逆転する傾向が確認された. 式(1)を用いて得られた固有ベクトルの可視化結果を図4に示す. 髪型領域における固有ベクトル画像の解析結果から, PC1ではプラス方向で髪の長さが増し, ロングヘアのような外観に変化している一方, マイナス方向では短髪化する傾向が確認された. PC2では, プラス方向において頭頂部付近のボリュームが大きく, 丸みを帯びたシルエットが現れるのに対し, マイナス方向では膨らみが少なく, 全体としてコンパクトな印象となっている. この軸では, 前髪の有無に関する変化も顕著に表れていると考えられる. さらに, PC3では, プラス方向で耳の位置から下の髪がすっきりと収まる形状となる一方, マイナス方向ではボリュームが増し, 顔まわりの髪型に大きな変化が生じていることが確認された.

4.5 主観評価実験

アンケート調査から得られた, 属性ごとの国籍分類の正解率をまとめたものを表3に示す.

この結果から, 全体的に日本人の分類精度が最も高く, 中国人及び韓国人は類似しており分類が難しい傾向が確認された. 特に海外生活経験の有無に注目すると, 海外生活経験を有する被験者では, 中国人の正答率がやや高い一方で, 日本

表3 属性別正解率

Table 3 Percentage of correct answers by attribute

	中国	日本	韓国
全体	36.0%	57.4%	36.4%
海外生活 経験無し	34.4%	58.9%	36.7%
海外生活 経験あり	45.0%	51.7%	35.0%



図5 各国籍におけるアンケートの正答率が高い髪型画像

Fig.5 Hairstyle images with high accuracy in nationality identification surveys

人の正答率が低下する傾向が見られた. 一方, 海外生活経験のない被験者群では, 日本人の正答率が高いという結果が得られた.

各国籍で正答率が高かった男女の髪型画像を図5に示す. 1行目が日本, 2行目が中国, 3行目が韓国となっており, Mが男性, Fが女性を表している.

日本男性の画像では「前髪あり」「頭頂部から耳上にかけて丸み強い」とする回答が多く, 日本女性ではそれに加え

て「全体的なボリュームが多い」との回答が多く見られた。中国男性の画像では「前髪なし」「頭頂部から耳上にかけて丸みが弱い」「耳下のボリュームが少ない」「全体的なボリュームが多い」といった回答が多く、中国女性の画像では「前髪あり」「頭頂部から耳上までの丸さが強い」「耳下のボリュームが多い」という傾向が確認された。

韓国男性の画像では「前髪あり」「頭頂部から耳上にかけて丸みが強い」「耳下のボリュームが少ない」との回答が多く、韓国女性の画像では「前髪なし」「耳下のボリュームが多い」とする回答が多かった。

## 5. 考察

### 5.1 転移学習による事前学習済みネットワークの比較からの考察

3種類のネットワークを用いて転移学習を実施した結果、いずれの国籍においても ResNet-101 が最も高い分類精度を示した。このことから、AlexNet 及び VGG-19 と比較して、この3種類のネットワークの中では ResNet-101 が顔画像を用いた国籍分類に最も適した事前学習済みモデルであると判断できる。また、先行研究と比較して ResNet-101 の精度が向上した背景には、本研究において年齢・性別といった属性情報を考慮し、データセットを構築したことが影響していると考えられる<sup>14)</sup>。さらに、データ数を増やす際にも同様に属性バランスを保つことで、さらなる精度向上が期待できる。一方、いずれのモデルにおいても韓国人の分類精度が他の国と比べて低い傾向が見られた。この結果は、韓国人の顔特徴が他国と比較して、分類に寄与する明確な形態的差異として現れにくい可能性を示唆している。

### 5.2 顔パーツのみを用いた学習からの考察

顔パーツごとの転移学習結果から、いずれのパーツにおいても韓国人の分類精度が他の2国に比べて低く、韓国人の顔的特徴が特定の部位に強く現れにくいことが示唆される。一方、日本人では目及び口領域において非常に高い精度が得られ、これらのパーツに日本人特有の形態的特徴が存在する可能性がある。注目すべきは、髪型領域を用いた分類結果である。日本人のみ精度が他パーツよりも低下したが、中国人及び韓国人では比較的高い精度が得られており、髪型にも国籍ごとの特徴が反映されていることが確認できた。特に、髪型による分類精度が目や口と同程度であったことから、髪型は顔認識における補助的要素ではなく、国籍分類に寄与する重要な分類要素の一つであると考えられる。さらに、元画像を用いた分類における混同行列の分析では、真のラベルが中国人の画像を韓国人と誤分類する例や、その逆のパターンが多く見られたことから、中国人と韓国人は髪型や顔形状において類似した傾向を持つと考えられる<sup>29)</sup>。一方で、目領域を用いた分類では、真のラベルが韓国人である画像が日本人と誤分類される例が多く、韓国人は日本人に近い目の形態的特徴を有している可能性が示唆される。

これらの結果から、髪型は他のパーツと同様に国籍分類に寄与する有効な特徴であり、従来の顔パーツ中心の分析に髪型情報を組み

合わせることで、より多面的な国籍分類モデルの構築が可能であると考えられる。

### 5.3 PCAによる特徴分析の考察

髪型領域における固有ベクトルの比較では、PC1が髪の長短を表す軸となっており、性差に関連する特徴を反映していると考えられる<sup>24)</sup>。PC2は頭頂部付近の髪のボリュームや丸みの変化を示しており、髪型の立体的形状を捉えている軸とみなせる。また、前髪の有無もこの主成分に影響していると推測される。PC3では耳の位置から下の髪のボリューム変化が顕著であり、頭頂部以外の下部髪型の特徴量を主に表していると考えられる。

これらの主成分が示す特徴量と三次元 PCA 空間の分布関係から、各国籍の傾向を以下のように整理できる。日本国籍は、頭頂部に丸みがあり、全体的にすっきりとした髪型が多く、前髪を下ろす傾向が強い。また、男女で前髪の有無に差が見られることが特徴である。中国国籍は、全体的にコンパクトで整った印象があり、前髪を上げるスタイルが多く見られる。韓国国籍では、耳下のボリュームが多い傾向を示すものの、前髪や頭頂部の形状に関しては明確な共通傾向が見られなかった。

これらの結果から、髪型は国籍ごとの特徴を反映する重要な要素の一つであり、特に頭頂部の形状やボリューム、前髪の有無等が文化的・形態的差異として表れている可能性が示唆される。これらの結果から得られた各国籍が持っている固有の髪型特徴量を活用し、国籍分類モデルを作成することで更なる分類精度の向上につながると考えられる。

### 5.4 主観評価実験に関する考察

属性別の正解率の結果から、日本の正答率が三カ国の中では約57%と高く、今回の被験者は全員日本国籍であったことから、自国籍の人々を見る機会が多く髪型の傾向をよく理解しているため正解率が他国よりも高くなったと考えられる。また、海外生活経験の有無による正解率では、海外生活経験有りと答えた被験者の中に中国での生活経験がある被験者がいたため、中国の正答率が高くなったと考えられる。これらのことから髪型に文化的な傾向があることが示唆され、接触経験が国籍分類の基準に影響を与えていると考えられる。また、髪型は単なる外見的特徴にとどまらず、文化的背景や社会的流行を反映する集団的特徴量として国籍分類に寄与しているといえる。学習モデルと精度を比較すると、学習モデルの方がはるかに高い分類精度を示し、人間は個人の経験や文化的要素を用い判断するため認知バイアスの影響を受けやすいと考えられる。そのためこのような髪型だけの国籍分類においては深層学習モデルが非常に有効であることが示唆される。

また、正答率が高い画像から、男女いずれにおいても「髪型の形状(丸みやボリューム)」や「前髪の有無」が国籍分類における判断材料として被験者に強く意識されていることが示唆された。特に、日本人は前髪の有無や頭頂部あたりに丸みの大小、中国人は全体的なボリュームの大小、韓国人は耳下部分のボリュームの大小がそれぞれ特徴的に認識されていると考えられる。

5.3 より、機械学習モデルと人間どちらにおいても国籍分類における判断材料として、「髪型の形状とボリューム」や「前髪の有無」が強く影響されていることが示唆されたため、これらの特徴量に焦点を置いた国籍分類モデルを作成することで分類精度の向上につながると思われる。

## 6. 終わりに

本研究では、日本・中国・韓国の三カ国の顔画像を対象に、ResNet-101, AlexNet, VGG-19 の3種類の事前学習済みネットワークを用いて転移学習を行い、国籍分類を実施した。その結果、ResNet-101 が最も高い分類精度 (86.9%) を示し、国籍分類タスクに適したモデルであることが確認された。

さらに、本研究の主題である髪型領域の重要性について検証を行ったところ、髪型のみを入力とした場合でも平均82.8%の高い分類精度を達成し、目や口といった従来の主要パーツと同程度の分類性能を有することが明らかとなった。また、顔領域及び髪型領域に対して主成分分析 (PCA) を適用し、固有ベクトル画像を可視化した結果、日本国籍は目が明瞭で前髪を下ろした丸みのある髪型、中国国籍は口元が水平で前髪を上げたすっきりとした髪型、韓国国籍は切れ長の目と耳下にボリュームを持つ髪型という傾向が確認された。評価実験においてもこれらの特徴は国籍分類において非常に重要な特徴であるということが明らかとなった。

これらの結果から、髪型が国籍ごとの文化的・社会的特徴を反映する分類要素であることが本研究により実証された。美容整形などによる従来の顔パーツ中心の国籍分類では分類が困難なケースにおいても、髪型情報を加味することでより高精度かつ文化的背景を踏まえた国籍推定が可能であることを本研究は示している。

今後の課題として、本研究を通して得られた各国籍の特徴を用いた精度向上のための新しいシステムを構築することが必要であると考える。また、顔認識に特化して事前学習をされたモデルや視覚言語モデルなどを用いて、本研究と同様の検証を行い、分類精度にどのような影響を与えるか考察する必要がある。さらに、データセットの規模拡大やそれぞれの年代ごとに分けた検証を行い、そこから考えられる特徴の傾向などの分析も今後の重要な課題である。

## 参考文献

- 1) Y.- Wang, Y.- Feng, H.-Liao, J.- Luo: "Do They All Look the Same? Deciphering Chinese, Japanese and Koreans by Fine-Grained Deep Learning", IEEE Conference on Multimedia Information Processing and Retrieval (2018).
- 2) M. Huber, F. Boutros, N. Damer: "Frequency Matters: Explaining Biases of Face Recognition in the Frequency Domain", arXiv preprint arXiv:2501.16896 (2025).
- 3) A. Kalkatawi, U. Saeed: "Ethnicity Classification Based on Facial Images Using Deep Learning Approach (MaxViT)", International Journal of Advanced Computer Science and Applications (IJACSA) (2024).
- 4) K.- Wu, C.- Lin, J.- Zhang, Z.- Wang, K.-Xu: "MonoHair: High-Fidelity Hair Modeling from a Monocular Video", Proceedings of Computer Vision and Pattern Recognition (CVPR) (2024).
- 5) M. Obayya: "Optimal Deep Transfer Learning Based Ethnicity Recognition", Image and Vision Computing (2022).
- 6) NEC ソリューションイノベータ, "FieldAnalyst(フィールドアナリスト)", <https://www.nec-solutioninnovators.co.jp/sl/fieldanalyst/lp/> (2025).
- 7) 永岡正梧, ナイワラ .P チャンドラシリ: "顔画像によるアジア三カ国の国籍判定", 電子情報通信学会 研究会技術報告, Vol. 125, No. 105, AI2025-7, pp. 39-44 (2025).
- 8) タレントデータバンク, [https://www.talent-databank.co.jp/\(2025\)](https://www.talent-databank.co.jp/(2025)).
- 9) 美韓, [https://www.bihann.com/\(2025\)](https://www.bihann.com/(2025)).
- 10) Mydramalist, [https://mydramalist.com/list/3DgPgApL\(2025\)](https://mydramalist.com/list/3DgPgApL(2025)).
- 11) douban, [https://movie.douban.com/\(2025\)](https://movie.douban.com/(2025)).
- 12) N. AlDahoul, M. J. Toledo, H. R. Kasireddy, Y. Zaki: "Exploring Vision Language Models for Facial Attribute Recognition: Emotion, Race, Gender, and Age", arXiv preprint arXiv:2410.24148 (2024).
- 13) MathWorks, "転移学習とは", <https://jp.mathworks.com/discovery/transfer-learning.html> (2024) .
- 14) K.-He, X.-Zhang, S.-Ren and J.-Sun: "Deep Residual Learning for Image Recognition", 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.770-778 (2016).
- 15) A. Krizhevsky, I. Sutskever, G. E. Hinton: "Image Classification with Deep Convolutional Neural Networks", Neural Information Processing Systems (NeurIPS), pp. 1097-1105 (2012).
- 16) K.Simonyan, A.Zisserman: "Very Deep Convolutional Networks for Large-scale Image Recognition", International Conference on Learning Representations (ICLR) (2015).
- 17) R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra: "Grad-CAM: Visual Explanations from Deep Networks Via Gradient-based Localization" IEEE International Conference on Computer Vision (2017).
- 18) S.-Yao: "Layer-wise Interpretability Investigation of Facial Expression Recognition Models Based on Grad-CAM", Proc. of 1st International Conference on Innovations and Applications of Artificial Intelligence (ICIAAI) (2024).
- 19) MathWorks: "Computer Vision Toolbox", [https://jp.mathworks.com/help/vision/index.html?utm\\_source=chatgpt.com](https://jp.mathworks.com/help/vision/index.html?utm_source=chatgpt.com), (2024).
- 20) K. Khan, J. Ali, I. Uddin, S. Khan, B. Roh "A Facial Feature Discovery Framework for Race Classification Using Deep Learning", arXiv preprint arXiv:2104.02471 (2021).
- 21) C.-Yu, J.-Wang, C.-Peng, C.-Gao, G.-Yu, N.-Sang: "BiSeNet: Bilateral Segmentation Network for Real-time Semantic Segmentation", Proc. of European Conference on Computer Vision (ECCV) (2018).
- 22) M. Turk, A. Pentland: "Eigenfaces for Recognition", Journal of Cognitive Neuroscience, Vol. 3, No. 1, pp. 71-86 (1991).
- 23) S. Adapa, E. Vamsidhar: "Deep learning based face shape classification

system with binary feature selection model”, Expert Systems with Applications, Vol. 296, Article 129115 (2025).

24) S. Makinist, G. Aydin: “Gender Classification Using Face Vectors: A Deep Learning Approach Without Classical Models”, Information, Vol. 16, No. 7, 531 (2025)

(2025年10月31日 受付)

(2026年4月3日 再受付)



永岡 正梧

2025年 工学院大学情報学部情報デザイン学科卒業。現在 同大学大学院工学研究科情報学専攻修士課程在学中。顔画像による国籍分類に関する研究に従事。



張 珏

2007年 筑波大学大学院人間総合科学研究科感性認知脳科学博士課程修了。工学院大学情報学部を経て2016年 同情報学部准教授、現在に至る。主に感性情報学、デザイン学、コンテンツ設計に関する研究に従事。日本感性工学会、日本デザイン学会、ヒューマンインタフェース学会、IEEE等の各会員。博士(感性科学)。



ナイワラ P. チャンドラシリ (正会員)

1996年 豊橋技術科学大学情報工学科卒業、2001年 東京大学大学院工学研究科電子情報工学専攻 博士後期課程修了(工学博士)。現在 工学院大学 情報学部教授。コンピュータビジョン、機械学習、ヒューマンコンピュータインタラクション研究に従事。IEEE、画像電子学会等の会員。

## 指向性をもつ新型ノズルを用いたホログラフィック 空間像投影用のミストスクリーンの安定性の向上

高野 邦彦<sup>†</sup>(正会員) 須田 諒真<sup>†</sup> 坂田 凜<sup>†</sup> 原田 博貴<sup>††</sup>  
前川 海斗<sup>†††</sup> 篠崎 大和<sup>††</sup> 戸塚 雄大<sup>†</sup> 佐藤 甲斐<sup>††††</sup>(フェロー)  
浅井 紀久夫<sup>††††</sup>

<sup>†</sup>東京都立産業技術高等専門学校, <sup>††</sup>長岡技術科学大学, <sup>†††</sup>東京都立大学, <sup>††††</sup>湘南工科大学, <sup>†††††</sup>放送大学

### Improvement of the Stabilization of the Mist Screen for the Spatial Holographic Projection Employing the Newer Structured Nozzle to Obtain the Mist-Flow with Enhanced Directionality

Kunihiko TAKANO<sup>†</sup>(Member), Ryoma SUDA<sup>†</sup>, Rin SAKATA<sup>†</sup>, Hirotaka HARADA<sup>††</sup>,  
Kaito MAEKAWA<sup>†††</sup>, Yamato SHINOZAKI<sup>††</sup>, Yudai TOTSUKA<sup>†</sup>, Koki SATO<sup>††††</sup>(Fellow),  
Kikuo ASAI<sup>††††</sup>

<sup>†</sup> Tokyo Metropolitan College of Industrial Technology, <sup>††</sup> Nagaoka University of Technology,  
<sup>†††</sup> Tokyo Metropolitan University, <sup>††††</sup> Shonan Institute of Technology,  
<sup>†††††</sup> The Open University of Japan

〈あらまし〉 本稿ではミストスクリーンを利用したホログラフィ投影空間法において、スクリーンを構成するミスト流の安定化を図るために気流制御用の新型ノズルを用いたスクリーン装置を検討するとともに、その再生像の安定度や輝度の向上を図るための基礎検討を行った。提案手法では、奥行き方向にボリューム幅を持たせたノズルを構成してスクリーン内のミスト流の指向性を高め、エアカーテン状のスクリーンを構成するように改善を図った。これにより、従来システムに比べ投影像の安定性と輝度が改善されることがわかった。

キーワード：動画ホログラフィ, 空間投影法, 霧スクリーン, エアカーテン, 粒子画像流速測定法 (PIV 法), 計算機合成ホログラム, DMD パネル

<Summary> In this paper, we made a basic study for the improvement of the screen system to raise the stability of mist flow and to produce highly bright reproduced images employing the newer structured nozzle. As this result, we succeeded in constructing a reformed screen system adopted a nozzle producing strongly directed mist flow with voluminous form to the depth direction. It enabled us to generate a mist screen of the form like air curtain. We constructed a system for raising the stability of the screen adopted one unit of ultrasonic vibrator, and confirmed the effectiveness of this system. Like this way, we achieved improvement in stability and brightness compared to previous systems.

**Keywords:** electro holography, projection onto the space, mist screen, air curtain, particle image velocimetry, computer generated hologram, DMD panel

### 1. はじめに

近年の街中の商店街では、窓ガラス上に商品紹介など広告映像などがシースルーで提示されることが多くなっている。その手段としてはプロジェクタからの投影像を光線の進行方向を制御可能な回折スクリーン（通称：ホログラムスクリーン）に投影する方法などが採られている<sup>1)</sup>。広告

表示された商品の魅力度をより強く訴求するためには映像コンテンツの一部を立体像で提示する方法が有効になる。その点において空間像提示方式の動画ホログラフィは有効な選択肢の一つと考えられる。一般的にホログラフィを利用した再生法（動画ホログラフィ）では、サブミクロン・オーダーの干渉縞パターンを LCD や DMD などの空間光変調素子 (Spatial light Modulator; 以下 SLM) に表示して、

その回折波面群から空間像を再生している。現在入手可能な SLM の解像力は縦・横ともに、数ミクロンオーダーと可視光線の波長の 10 倍程度の粗い画素ピッチであるため、本来のホログラム像再生に必要な空間帯域幅が不足しており、空間像を拡大観察するための技術開発が必須になると考えられる<sup>2),3)</sup>。その対策としては空間像を空間スクリーンなどに拡大投影する方法が有効であることから、空間スクリーン<sup>4)~13)</sup>、フォトリフラクティブ結晶体を利用した手法<sup>14)</sup>、複数のミラー<sup>15),16)</sup>や HOE<sup>17),18)</sup>を用いた手法が検討されている。ホログラフィ再生像の視域とサイズの両方を拡大する点において、ミストスクリーンを利用した空間投影法<sup>4)~6),11),12)</sup>は有効であることが知られており、スクリーンの大型化が期待されている。この方法では空間内に噴霧したミスト粒子(散乱体)のみから空間スクリーンとして機能するため、将来的にはスクリーンのサイズ設定に対しても自由度が期待される。その一方で、ミスト流の揺らぎがスクリーンとしての形状や大きさの安定性に直結するため、スクリーンのスクリーン形状の安定化にはミスト流の揺らぎ軽減が課題であり、そのための技術開発が重要になることが考えられる。スクリーンの形状を安定化する試みとしては、ファンを利用して強制的に気流を発生させる方法<sup>5),6),12)</sup>や、ミスト粒子を空間内に閉じ込める方法<sup>13)</sup>などが検討されている。さらに、複数スクリーンの配置調整による画質改善<sup>19),20)</sup>や、ノズルを円形配置として円筒形状のスクリーンを構成する方法<sup>21)~26)</sup>などの実施例がある。空間投影方式の利点を活かし、同時に複数人が空間投影像を観察するためにはスクリーンの大型化が必要とされている。そのためには安定した像表示が可能な領域(以下、有効表示面)を拡げる方法の確立、すなわち、ミスト流の形状安定化に向けた検討が必要と考えられる。スクリーンの形状安定化に向けた先行研究としては、独立駆動するファンをスクリーン上下に設置する方法<sup>6),12)</sup>、スクリーン下部に羽根なしファンを設置する方法<sup>5)</sup>、スクリーン上部に複数のファン設置する方法<sup>25),26)</sup>などが知られており、これらの検討により、スクリーン形状の安定化に一定の効果が確認されている。この方法をスクリーンの大型化に適用する場合には、使用するファンのサイズそのものが開口の大きさに直結すると考えられる。そのため、ボリュームを確保しながら、かつ、ファンのサイズよりも大きい開口幅をもつノズルを構成して、ミスト流の指向性を強める必要がある。本稿では、スクリーンを大型化するための基礎検討として、スクリーン内でのミスト流の指向性を強めることによりエアカーテン状<sup>27)~29)</sup>のミスト流を生成し、これをスクリーンとして活用する方法に着目した。まず、スクリーン生成装置の構成を示すとともに、開発した指向性強調型のノズル(以下、改良型)の構造について述べる。次に、粒子画像流速測定法(以下; PIV 法)<sup>30)</sup>に基づく流体

解析の結果から、スクリーン内のミスト流の安定性が向上することを示す。さらに実際にスクリーンに投影した面状の四角形(静止物体)の結果から、瞬間的に気流変動が発生した場合においても、スクリーンの有効表示面積の減少割合が、従来法<sup>5)</sup>に比べ改善されることがわかり、面状などグラデーションのある物体の表示にも可能性が示された。

## 2. 提案する空間像投影システムの構成

### 2.1 システムの全体構成

本稿で改良した改良型の空間像投影システムの全体構成は図 1 に示す通りであり、赤色光のレーザ光源(civil-lasers, 発振波長: 635nm, 以下, LD), TI 社製の DLP LightCrafter6500(表 1 参照)の Digital Micro-mirror device パネル(DMD)と後述する空間スクリーン生成装置から成る。同パネル上には、制御用 PC により光線追跡フレネル型の CGH パターン<sup>5)</sup>を電気的に書き込むことによりホログラム面を形成する。上述のホログラム面に照射されたレーザ光から反射波面群をミストスクリーンへ投影している。なお、本稿では LD から DMD パネル、DMD パネルから空間スクリーン(像表示位置)への直線距離は順に 133cm, 86cm(一例)と設定した。

### 2.2 ミスト流の指向特性を強調する新型構造のノズル

ホログラム再生像を投影するスクリーンには、時間的かつ空間的に高い安定性が求められる。この観点から筆者らは吸気用・排気用の 2 か所にファンを設置する方法や、羽根なしファン 1 台のみを排気部分に設置する方法<sup>5)</sup>により強制的にミスト流を生成してスクリーンの揺らぎの軽減させる方法を検討した。さらにスクリーンの揺らぎを軽減さ

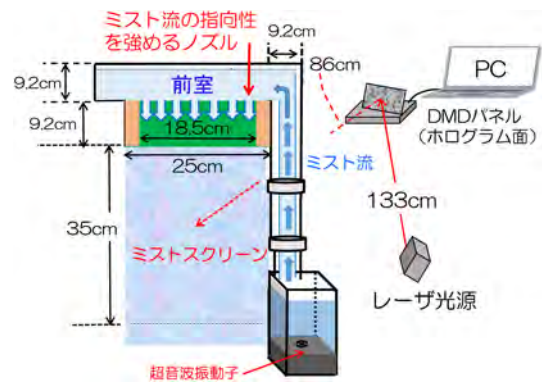


図 1 改良型ミストスクリーン生成装置を含むホログラム投影像表示システムの全体構成

Fig. 1 Total scheme of improved image reconstructing system used in the experiment

表 1 DMD パネルの仕様  
Table 1 Specification of DMD panel

パネルサイズ	0.65in.
リフレッシュレート	250Hz(システム全体)
画素数	2,073,600(V1,920 × H1,080)
画素ピッチ	7.56μm × 7.56μm

せるには、高い指向性のミスト流を作り出すためのノズルを構築することが必要不可欠と考えられる。ここでは、本稿で構築したミストスクリーン生成装置の考え方とノズルの改善方法について述べる。

本稿では次の手順でミストスクリーンを構成している。

- ハイブリッド超音波加湿器 ((株)山善社製の KH-A55 型、最大出力は 400mL/h) に内蔵されている超音波振動子により、直径 2~3 $\mu$ m(公称値) の水粒子を発生させた。
- 後述するノズルから均一なミスト流を作り出すことを目的として、超音波振動子から押し出された水粒子(ミスト流)を、管を通じてノズル部の前室へ移動させた。ノズルから噴出されるミストの量と安定度はトレードオフの関係があるため、本稿では良好な表示像が得られるよう、超音波振動子から生成されるミスト量をもとに、一例として前室のサイズを図 2 のように設定している。なお、超音波振動子から生成されるミスト量を増加させた場合には、前室サイズも大型化する必要があると考えられる。
- 後述するノズルにより、指向性を高めたミスト流を自由落下させた。

ミスト流を噴出させるノズルの開口の大きさは、スクリーン内のミスト分布に影響を与えと考えられる。超音波振動子から供給されたミストを噴出するノズルの構造は図 2(a) のとおりであり、1mm の厚みをもつプラスチック塩化ビニール板の両端(開口部両端の内側)に、8~9mm の厚みをもつベニア板(素材)を接合することによりスリット状としてノズルを構成している。ノズルの開口横幅を 18.5mm を設定したのは、ボリューム確保のために用いた素材(ベニア板)の厚みを考慮しつつ、超音波振動子から供給されるミスト流の勢いをベンチュリ効果<sup>31),32)</sup>により強め、その指向性が高まるようにしたためである。従来型<sup>5)</sup>でのノズル(b)ではスクリーンにボリュームをつける目的で水平・奥行き方向に三層のスリットが設定されていたため、同時に気流制御が複雑化する要因となっていた。本手法の(a)では、超音波振動子から供給したミストを、奥行き方向に 8~9mm 幅のボリュームを付けながら、水平方向のみに 18.5cm 幅のノズルで拡げ、ミスト流の落下方向を整えている。同時にノズル幅を絞ることにより、ベンチュリー効果による部分的な減圧効果を得ている。これにより指向性の高いエアカーテン状のミストスクリーンを実現しており、気流制御機構の単純化を図っている。

時間経過に伴うミスト流の乱れの軽減に対して、本手法で構成したノズルは図 3 に示すように有効な手段の一つと考えられる。なお、表示領域の両端(左側と右側)ではミスト濃度が高い箇所があるが、これはノズル開口から噴出されたミストの一部分が左右にもやや拡がり、結果的にミ

スト量が増加するためと考えられる。この改善には、噴霧器の出力をさらに上げるなど、スクリーンへのミストの供給量を増やすことによりミスト流を指向性を一層強めることが必要になると考えられる。

### 2.3 スクリーン生成装置の駆動音について

本手法では、観察者は従来型<sup>5),6)</sup>と同様にスクリーンから 75cm 程度離れた場所から像観察をするシステムとしている。観察位置での駆動音のレベル(Rion, NL-05 型で計測)は、羽根なしファンを利用したシステムでは、暗雑音の補正後(A 特性)の値が 70~75dB(A) 程度、吸気・排気用として 2 基のファンを用いた方法でも 62~65dB(A) と、米国の Centers for Disease Control and Prevention(CDC) の基準<sup>33)</sup>において、「騒音が気になる、うるさいレベル」となっており装置の騒音レベルの低減が課題となっていた。それに対して本稿でのスクリーン生成装置ではファンレスの構成としたため、装置の駆動音を 49~51dB(A) 程度と大幅に低減でき、CDC 基準で「静かな事務所、家庭用クーラー(室外機)」、すなわち騒音レベルが気にならないレベルになることが確認された。

## 3. 結果及び検討

ここでは、後述する評価実験によりスクリーンの光学特性と再生像の輝度評価を行った結果について述べる。評価範囲は図 4 に示す条件で設定した。なお従来型<sup>5)</sup>との表記の統一性を考慮して、中段の中央の位置を①、その周囲を

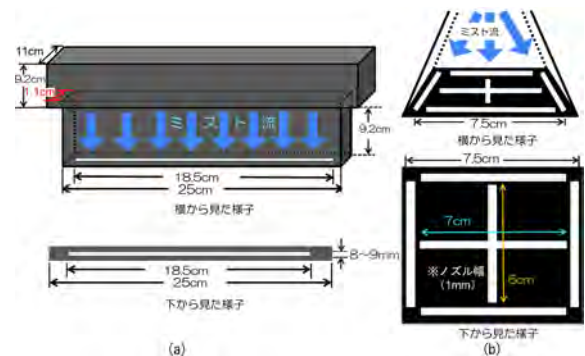


図 2 ノズル構造の比較;(a) 指向性強調型のノズル,(b) 従来型のノズル

Fig. 2 Comparison of structures of the nozzle; (a)Improved nozzle, (b)Previous nozzle



図 3 提案手法での気流の様子

Fig. 3 Stabilization of water flow in the proposed method

①～⑧と表記している。

### 3.1 流体解析によるスクリーンの光学特性の評価

ここではスクリーン上での水粒子の動きを Library 社の流体解析ソフトウェア Flow PIV を利用した解析結果から確認する。本稿では水粒子の平均移動ベクトルのうち、図 5(a) に示す角度  $\theta$  を移動方向と見做し、その標準偏差  $\sigma$  [deg] に着目した。水粒子が同図 5(b) に示すように指向性が高い(多くの水粒子が揃って鉛直下向きに落下する)場合には  $\sigma$  が 0 に近づき、逆に同図 (c) のように、指向性が低い(水粒子の落下方向がばらつく)場合には  $\sigma$  の値が大きくなると考えられる。また、評価範囲を変更しても  $\sigma$  の傾向に差異がなければ、スクリーン全体として表示像の傾向には大きな相違がないことが考えられる。

この考え方のもと、改良型ならびに従来型<sup>5)</sup>における  $\sigma$  の傾向を比較した結果を表 2 に示す。表 2 には、参考実験として、改良型装置のノズルと羽根なしファンを併用した場合について比較した結果を併記している。

表 1 より次の傾向が見て取れる。

- (a) 改良型では評価位置が変わっても傾向差が少ないため、スクリーンの位置による像表示の傾向に顕著な差は発生しない。
- (b) 従来型<sup>5)</sup>の場合は、上・中段では傾向差はないが、ノズルから離れた下段では揺らぎの影響が顕著になる。

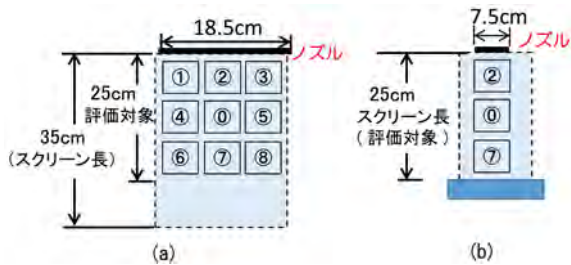


図 4 各装置における評価範囲の設定条件; (a) 改良型の場合 (b) 従来型<sup>5)</sup>の場合

Fig. 4 Evaluated position on the screen; (a) New system, (b) Previous system

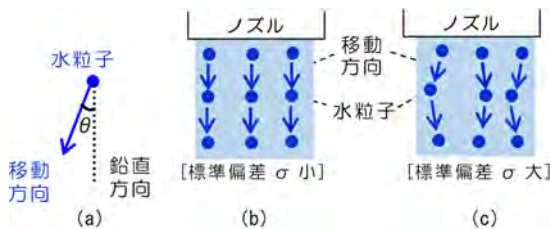


図 5 水粒子の移動平均ベクトルの鉛直方向に対する角度  $\theta$  の標準偏差  $\sigma$  とスクリーンの安定性との関係; (a) 水粒子の移動平均ベクトルの角度  $\theta$ , (b) 指向性が高い場合, (c) 指向性が低い場合

Fig. 5 Relationship between the angular deviation  $\sigma$  of the average velocity  $\theta$  of the water particles and the stability of the screen; (a) The definition of the average velocity  $\theta$ , (b) Case of using the highly-directional mistflow, (c) Case of using the low-directional mistflow

(c) 改良型ノズルと羽根なしファンを併用した場合には、改良型と従来型の何れの場合よりも揺らぎの影響が顕著になる。これは、ノズルによって指向性を強めたミスト流が、羽根なしファンが作る気流により乱れることによるものと考えられる。これより、本手法で構成したノズルを使用する場合には、外部ファンを用いない構造とすることが、スクリーンの安定性を得る上で有効となることがわかる。

以上より、改良型のノズルを採用することが、スクリーンの有効表示範囲が拡大する上で有効であると考えられる。以降ではスクリーン装置としての提案手法と従来型<sup>5)</sup>の比較結果について述べる。

### 3.2 多点数の面状物体を利用したスクリーン上の輝度変動の傾向比較

ここでは、時間経過に対する投影物体の輝度変動の様子を調べることにより、スクリーンの安定性を確認する。スクリーンの設定範囲内に表示された投影像の輝度変動の傾向を調べるため、CGH 生成時の入力物体には面状の正方形(総点数:900)に用いている。表示エリア全体に像表示をする目的で改良型システムにおいて上段中央②の範囲内に表示された投影像を 30fps の時系列の画像(写真)として撮像した上で、取得した写真画像に対して図 6 に示すように輝度分布のサーフェスプロットにより輝度変化を調べた。同図は、本システムにおいて CGH 生成時の入力物体の総点数を限界条件(1,000 点)よりも多めの 1,764 点とした場合の再生結果を示しており、スクリーン上のマーキング位置に、局所的な輝度むらが発生する様子が確認できる。同様に、スクリーン内で瞬間的に気流が乱れた箇所においても、像の一部分の輝度が著しく低下することが知られているため、スクリーンの安定性の確認を目的として同手法を適用している。正方形の総点数を前述のように 900 とした上で、時間経過に対する投影像の輝度分布の一例を図 7 に示す。さらに比較参考のため、羽根なしファンを利用した手法<sup>5)</sup>に対する結果を図 8 に示した。図 7 及び図 8 における各投影像は、140 枚の投影像の時系列写真において、顕著な差異が認められた投影像とその輝度分布を、時間経過にしたがって任意で抽出したものである。同図から次の

表 2 平均速度ベクトルの方向のバラツキの比較  
Table 2 Relationship between the angular deviation  $\sigma$  [deg] of the average velocity of water particles and the irectivity performance of the screen

評価位置	上段			中段		
	①	②	③	④	⑤	⑥
改良型	19.84	17.27	19.41	20.4	18.94	19.08
従来型	—	17.77	—	—	18.85	—
参考実験	20.27	18.12	21.13	23.31	24.61	23.30
評価位置	下段					
	⑦	⑧	⑨			
改良型	20.96	21.49	20.03			
従来型	—	24.16	—			
参考実験	27.87	27.63	27.96			

傾向が見て取れる。

- 高輝度で多点数の面状物体の表示が可能となり、スクリーンを構成するミスト流の安定度が改善される。なお、CGHパターンに記録する入力物体の総点数を1,000以上に設定した場合には、投影物体の輝度むらの影響が顕著になることが確認された。この結果は拡散板面スクリーンに像投影をした場合の結果と同程度であることから、本システムの場合においても、時分割多重再生法を用いない場合に表示可能となる物体の総点数は500~1,000程度となると考えられる。
- 再生図形としての確認はできるが、図7のマーキング部分のように、瞬間的に発生したミスト流の変動などにより、スクリーン内で局所的に輝度低下の発生する。これはミスト流の伝送経路やノズル部内において、瞬間的にミスト流の変動が発生したためと考えられる。
- 図8の羽根なしファンを利用した場合において、瞬間的にファン動作が不安定になり気流乱れが発生すると、ミスト濃度が著しく減少して、同図の左から2番目のように、表示図形の形状が変化する場合がある。それに対して、図7(本手法)の場合には、顕著な輝度低下が発生する範囲が大幅に減少する。
- 図7(本手法)の場合には、瞬間的に筋状のノイズが発生する場合がある。これはファンで強制的にミストを押し出していないため、ノズル部に水滴が溜まった場合に、それが落下する際に筋状のノイズが発生したと考えられる。

以上の結果から、ノズルの指向性を高めてミストスクリーンを構成する本手法は、スクリーンの安定化に対して有効な選択肢の一つになると考えられる。また、参考までに観察者がスクリーン内に手を入れた場合の投影像への影響を確認したところ、従来法<sup>5)</sup>の場合には気流乱れが発生し、図8(左から2番目)と同様の現象が確認された。提案手法の場合にはこの現象は確認されなかった。この点からも提案手法の有効性が伺える。なお、スクリーン上の局所部分で瞬間的に発生する輝度変動(筋状のノイズ)の軽減には、ミスト流の伝送経路の調整やノズル部構造の調整が必要になると考えられるが、これは今後の検討課題としたい。

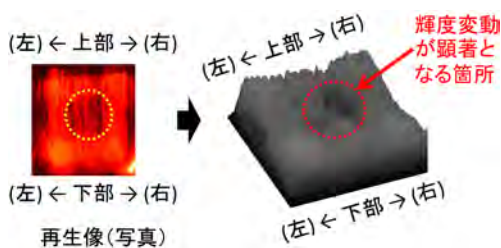


図6 スクリーン上の輝度変動の確認方法

Fig. 6 Confirmation method of the brightness fluctuations on the screen based on the projected images

### 3.3 面状物体の投影像を用いた有効表示面の比較

前節と同じ面状物体による再生結果から、各評価位置において投影物体が高輝度で得られる割合Pを比較する。この値Pの値が100に近いほど、高輝度で物体が投影できる表示面が広がると考えられる。ここでは、物体表示と見做せる範囲を、一般的なディスプレイにおける輝度評価<sup>34)</sup>を参考にして、図9の撮像データ(感度補正をした上で撮像した8bitのグレースケール画像)から最大濃度の半値幅に設定した(選択された画素値をヒートマップとして示している)。このようにして抽出された画素を白、それ以外を黒とした上で、全画素に対する白画素の百分率をPとしており、Pの値は文献5)と同様に算出している。各スクリーン装置に対して求めた「高輝度となる割合P」の算出結果を表3に示す。これを見ると、(a)上段(①,②,③)と中段(④,⑤)の位置では従来型<sup>5)</sup>に比べ、改良型において投影像の輝度が向上することが見て取れる。

### 3.4 ミスト量に対する多点数物体の再生像の変化

前節と同様に、改良型において上段中央②の範囲内に、線と面の要素を含む多点数物体から生成したCGHを用いた上で、時分割多重再生法<sup>14),35),36)</sup>により空間内に多重投影した。初代ミッキーマウス(2024年からパブリックドメ

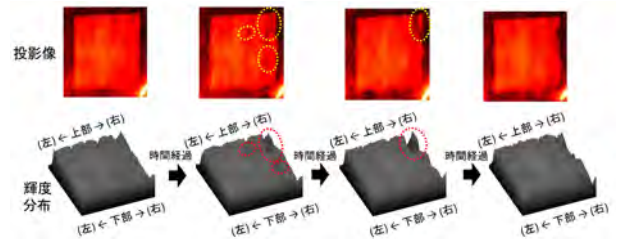


図7 多点数物体の再生結果の一例(本手法)

Fig. 7 Example of the images reconstructed from the CGH of the objects described by multiple points (New system)

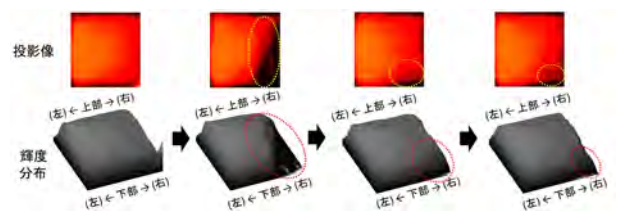


図8 多点数物体の再生結果の一例(従来法)

Fig. 8 Example of the images reconstructed from the CGH of the objects described by multiple points (Previous system)

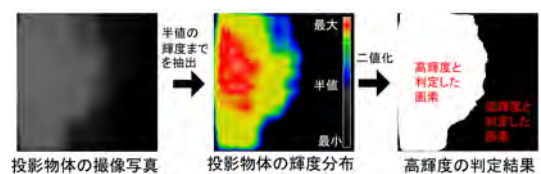


図9 高輝度となる割合Pの考え方

Fig. 9 The ratio P in the projected images

インへ移行)を参考にして、入力物体を構成している。スクリーンの表示領域の拡大効果を確認するため、入力物体のサイズは0.89cm(横)×1.35cm(縦)の縦長の物体に設定している(総点数は9,632)。本節では、座標分割数を10(各要素CGHへの入力物体点数は、順に964[2個]と963[8個]に対応)に設定した上で、スクリーンに供給するミストの量を変化させながら像特性を比較した。各CGHパターンに記録する入力物体の総点数の適正な設定条件は、文献35),36)を参考にして500~1,000前後となるよう設定している。投影像の一例は図10に示す通りであり、横5.2cm×縦8cm前後に拡大投影している。各CGHパターンは、本手法で用いたDMD表示装置の限界値となる0.105ms間隔で同パネル上に構成しており、時分割による多重再生により像投影を行っている。これを見ると以下のことが確認できる。

- 時分割多重投影により、高輝度の多点数物体が表示可能となる。
- ミスト流が極端に少ない場合には、物体の一部分の輝度が低下傾向となる。再生物体は確認できるが、再生物体の箇所によっては、著しい輝度の低下が認められる。
- ミスト量が中程度の場合には、曲線と面状の各要素において、顕著な画質の改善効果が認められる。(図10の黄色及び青色マーキング部分を参照)
- ミスト量が極端に多くなると、背景ノイズが増加傾向となり、また、表示像にボケが多く発生することから、線状と面状の識別が困難になる。(図10の緑色マーキング部分を参照)
- ミスト量の設定条件を適切に調整することにより、スクリーンの下部に表示された「足」の部分も鮮明に表示される。

以上より、本手法において直線や曲線、面状の各要素を含む多点数による点物体の表示が可能になることがわかる。これは、スクリーンを構成するミスト流の安定化が図られ

たことによる効果の一つと考えられる。

### 3.5 提案手法による動画投影像の一例

参考実験として、改良型の装置において、上段中央②の範囲内にブランク(空白)と、7セグメントの数字に変化する5種類の入力図形から生成したCGH(#01~#05)から成る6フレームを、0.12s間隔でDMDパネル上に構成して動画表示を行った。得られた再生結果の一例は図11に示す通りである。ミスト流の瞬間的な乱れなどによって、マーキング位置に不要光がやや発生するが、滑らかな動画像として観察できた。

## 4. 結論

本稿では、ミストスクリーンの安定化を目的として気流制御用の新型ノズルを用いたスクリーン装置を検討した。本稿で構成したノズルを用いることにより、奥行き方向にボリューム幅を持たせ、かつ指向性を高めたミスト流を作り出し、エアカーテン状のスクリーンが構成可能となることがわかり、提案手法がスクリーン生成装置を構成する上で有効となることがわかった。スクリーン内の局所領域で瞬間的に発生する筋状ノイズの軽減に向け、ノズル部を含めた更なる改善策が必要となるが、これは今後の検討課題としたい。

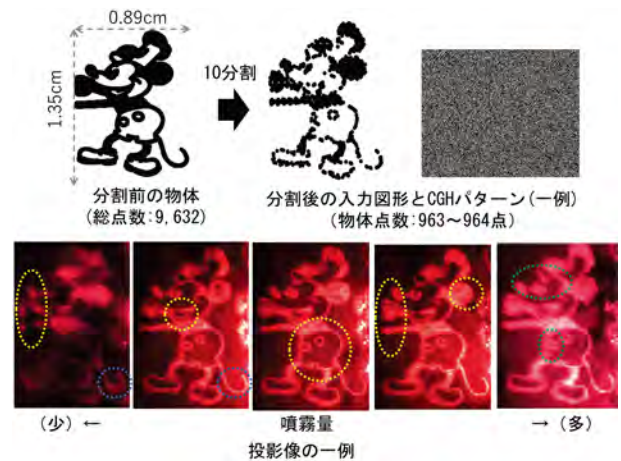


図10 多点数物体の再生結果の一例  
Fig. 10 Example of the images reconstructed from the CGH of the objects described by multiple points using time-shared reconstructing process

表3 輝度平均と高輝度となる割合Pの比較

Table 3 Results of comparing the average brightness of displayed objects and the percentage P of high brightness of the objects

評価位置	投影像の輝度平均 (8bit 濃度の平均 [a.u])		高輝度となる割合 P (濃度が高い割合 [%])	
	従来型	改良型	従来型	改良型
①	—	108.2	—	91.4
②	103.0	110.0	97.9	95.7
③	—	103.9	—	93.0
④	—	116.6	—	96.7
⑤	97.6	118.8	92.0	98.3
⑥	—	105.7	—	83.8
⑦	—	75.9	—	76.6
⑧	78.6	71.9	69.5	73.2
⑨	—	69.2	—	68.2

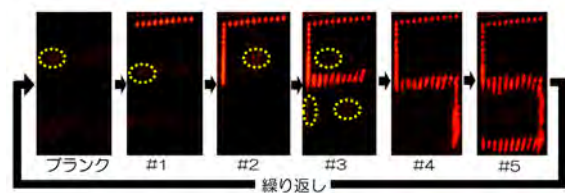


図11 動画投影像の一例  
Fig. 11 An example of the projection of animated moving images

## 謝 辞

本研究は科学研究費補助金(課題番号:23K11352)の一環として実施された。ここに深謝の意を表し厚く御礼を申し上げます。

## 参考文献

- 1) 市橋保之, 涌波光喜, 吉田俊介: “ホログラフィック光学素子と複数の小型プロジェクタを用いた AR-3D ディスプレイシステムの実装”, 映像情報メディア学会誌, Vol.77, No.4, pp.546-551 (2023).
- 2) 吉川浩, 丹治英一郎: “高効率符号化を用いたホログラフィック 3 次元画像の圧縮法”, テレビジョン学会誌, Vol.47, No.12, pp.1678-1680 (1993).
- 3) P.S.Hilaire, S.A.Benton, M.Lucente, M.L.Jepsen, J.Kollin, H.Yoshikawa, J.Umderkoffler: “Electronic Display System for Computational Holography”, Proc. of SPIE,1212, pp.174-182 (1990).
- 4) K.Takano, K.Sato and M.Ohki: “Improved Scattering Screen for a Multiplanar Volumetric Holographic Display”, Opt. Eng.,Vol.50, No.9, 091315 (2011).
- 5) K.Takano, Y.Totsuka, Y.Takano, K.Hara, M.Kotoyori, K.Sato and K.Asai: “On an improvement of the stability of a spatial mist screen employed a fan without blades in the display of a projected holographic image”, Proc. of SPIE, 131640L [WAIT2024] (2024).
- 6) 高野邦彦, 大岩徹志, 福嶋駿, 梶村一太, 小野治親, 高崎和之, 若林良二, 圓道知博, 佐藤甲癸, 浅井紀久夫: “青紫色光源を用いたホログラフィ再生像の空間投影法の検討”, 画像電子学会誌, Vol.52, No.3, pp.405-412 (2024).
- 7) 妹尾孝憲, 大友明, 山本健詞, 大井隆太郎, 市橋保之, 栗田泰一郎: “電子ホログラフィ映像の拡大投影の検討: “像サイズと視域の拡大に向けて”, 映像情報メディア学会技術報告, Vol.36, No.12, pp.29-32 (2012).
- 8) M.L. Huebschman, B.Munjuluri, and H.R. Garner: “Dynamic holographic 3-D image projection”, Optics Express, Vol.11, No.5, pp. 437-445 (2003).
- 9) Y.Ochiai, K.Kumagai, T.Hoshi, J.Rekimoto, S. Hasegawa, and Y.Hayasaki: “Fairy Lights in Femtoseconds: Aerial and Volumetric Graphics Rendered by Focused Femtosecond Laser Combined with Computational Holographic Fields”, Proc.of ACM SIGGRAPH 2015, pp.9-13 (2015).
- 10) K.Takano, K.Miyama, K.Hamano, K.Motojima, K.Sato, M.Ohki and K.Asai: “On some improvement of the waterfall screen for the stable projection of holographic images”, IIEEJ Trans. on Image Electronics & Visual Computing, Vol.2, No.1, pp.48-58 (2014).
- 11) 高野邦彦, 佐藤甲癸, 大木眞琴: “微粒子の散乱を用いたホログラフィ用立体スクリーンの検討”, 映像情報メディア学会誌, Vol.57, No.4, pp.476-482 (2003).
- 12) 高野邦彦, 桑崎博司, 佐藤甲癸, 浅井紀久夫: “ホログラフィック投影法のためのミストスクリーンの安定化に向けた基礎検討”, 画像電子学会誌, Vol.45, No.1, pp.98-104 (2015).
- 13) 高野邦彦, 宮崎真宏, 八重樫研介, 山口真, 大木眞琴, 佐藤甲癸, 浅井紀久夫: “薄膜封入型ミストスクリーンの構成手法とそのホログラフィ立体像への効果”, Vol.45, No.1, pp.20-26 (2016).
- 14) M.Sasaura, T.Horikoshi, Michio Ono, T.Imai, S.Yagi, E.Kubota, A.Tate, H.Kojima and N.Sonehara: “Proposal and demonstration of cached holographic 3D display system using photorefractive crystals”, Journal of Crystal Growth (Elsevier), Vol.229, No.1, pp.199-204 (2001).
- 15) 内田十内, 成島佑華, 小田好洗, 森口嘉軌, 高田直樹: “ホログラフィックプロジェクタを用いたリアルタイム空中描画システムの開発”, FIT2022, I-021, pp.303-304 (2022).
- 16) S.Yoshida: “fVisiOn: 360-degree viewable glasses-free tabletop 3D display composed of conical screen and modular projector arrays”, OPTICS EXPRESS, Vol.24, No.12, 13194-13203 (2016).
- 17) H.Jeon, S.Lim, Y.Jeon, W.Baek, D.Heo, Y.Kim, H.Kim, and J.Hahn: “Holographic printing for generating large-angle freeform holographic optical elements”, Optics Letters, 47, 2, DOI:10.1364/OL.448480 (2022).
- 18) Y.Takatsuka, K.Yoshimoto, and H.Takahashi: “Retinal projection type 3D head-mounted display using an HOE lens array”, IS & T International Symposium on Electronic Imaging 2016, Stereoscopic Displays and Applications XXVII, SDA-452 (2016).
- 19) C. Lee, S. DiVerdi and T. Höllerer: “Depth-fused 3D imagery on an immaterial display”, IEEE Trans. on Visualization and Computer Graphics, Vol.15, No.1, pp.20-33 (2009).
- 20) 橋本直己, 東佑圭: “霧を用いた空中立体像提示システムの提案”, 映像情報メディア学会誌, Vol.65, No.7, pp.1007-1010 (2011).
- 21) I. Rakkolainen, S. DiVerdi, A. Olwal, N. Candussi, T. Hüllerer, M. Laitinen, M. Piirto and K. Palovuori: “The interactive fogscreens”, ACM SIGGRAPH 2005 Emerging Technologies, Article No.8 (2005).
- 22) 山口慶二郎, 遠藤祐二, 住友翔, 三輪敬之: “霧を用いた表現メディア空間の設計と遊び場への活用”, ヒューマンインタフェースシンポジウム 2012, pp.755-758 (2012).
- 23) 八木明日華, 井村誠孝, 黒田嘉宏, 大城理: “多視点観察可能なフォグディスプレイ”, 日本バーチャルリアリティ学論, 17, 4, pp.409-417 (2012).
- 24) 三寄一樹, 池田竜也, 森裕司, 大塚達也, 三輪達也, 三輪敬之: “円筒状霧ディスプレイの開発とその共同描画への適用”, 第 17 回計測自動制御学会システムインテグレーション部門講演会, SI2016 (2016).
- 25) 井村誠孝: “円筒型フォグスクリーンへのプロジェクションによる多視点表示”, 光学, Vol.43, No.10, pp.469-474 (2014).
- 26) 澁澤翔也, 蛭田航平, 国吉光: “ミストスクリーンの投影範囲の拡大”, 関東学生学会第 56 回学生員卒業研究発表講演会論文集 (2017).
- 27) Ran Duan, Jiagen Liu: “Comparative study on pollution-block performance of ceiling-supply air curtain system for commercial building entrance”, Proc. of the 16th IBPSA Conference, pp.650-656 (2019).
- 28) R.A.Eddine, Y.Nezu and Y.Oshima: “Assessing horizontal air curtain effectivity using Particle Image Velocimetry (PIV) analysis for preventing the spread of airborne infection in lecture halls”, Journal of Environment and Safety, Vol.14, pp.1-5 (2023).
- 29) K.Sugimoto and S.Mizuno: “Visualization and Interaction of Light Rays through Aerial Imaging”, Proc. of SPIE, 131641R [WAIT2024] (2024).
- 30) T.Ishima: “Fundamentals of Particle Image Velocimetry (PIV)”, Journal of the Combustion Society of Japan, Vol.61, No.197, pp.224-230 (2019).
- 31) 淡野公一, 川崎悠人, 川崎清人, 田村宏樹, 外山貴子: “ベンチュリー効果を用いたナノミスト噴霧器とその応用”, 平成 24 年度電気関係学会九州支部連合大会, 07-2P-12, p.571 (2012).
- 32) 上澤伸一郎, 堀口直樹, 柴田光彦, 吉田 啓之: “ベンチュリー管内の水-蒸気二相流挙動に関する研究”, 日本機械学会 第 22 回動力・エネルギー技術シンポジウム 講演論文集, No.17-15 (2017).
- 33) N.Natarajan, S.Batts and K.M. Stankovic: “Noise-Induced Hearing Loss”, Journal of Clinical Medicine, 12, 2347 (2023).
- 34) 政倉祐子, 田村徹, 長嶺邦彦, 富岡聡: “LED バックライトの輝度ムラ評価法に関する実験的検討”, 映像情報メディア学会誌 Vol. 63, No. 10, pp. 1423-1428 (2009).

- 35) 高野邦彦, 清野智之, 長尾俊介, 小室重行, 佐藤甲癸, 浅井紀久夫, 大木眞琴: “時分割再生法を用いた多点多数物体のホログラム再生像特性”, 画像電子学会誌, Vol.39, No.6, pp.1067-1073 (2010).
- 36) 高野邦彦, 横田和海, 笠井湧喜, 安藤香月, 川崎悠也, 佐藤甲癸, 浅井紀久夫: “青紫色レーザー光源を用いたホログラフィックカラー再生について”, 画像電子学会誌, Vol.51, No.2, pp.157-163 (2022).

(2025年3月20日受付)

(2025年11月28日再受付)



**高野 邦彦 (正会員)**

2002年 湘南工科大学大学院工学研究科電気工学専攻博士後期課程修了。同年, 東京都立航空工業高等専門学校(2006年 東京都立産業技術高等専門学校に組織変更) 助手, 講師, 准教授を経て, 2024年 同校教授。動画ホログラフィ, 画像処理などの研究に従事。博士(工学)。



**須田 諒真**

2024年 東京都立産業技術高等専門学校ものづくり工学科情報通信工学コース卒業。在学時は動画ホログラフィに関する研究に従事。



**坂田 凜**

2024年 東京都立産業技術高等専門学校ものづくり工学科情報通信工学コース卒業。在学時は動画ホログラフィに関する研究に従事。



**原田 博貴**

2024年 東京都立産業技術高等専門学校ものづくり工学科情報通信工学コース卒業。2026年 長岡技術科学大学電気電子情報工学分野卒業。現在, 同大学大学院工学研究科電気電子情報工学分野修士課程に在学中。LiDARを用いた雪害対策などの研究に従事。



**前川 海斗**

2024年 東京都立産業技術高等専門学校ものづくり工学科情報通信工学コース卒業。2026年 東京都立大学システムデザイン学部電子情報システム工学科卒業, 現在, 同大学大学院システムデザイン研究科電気電子工学域博士前期課程に在学中, THz動作に向けた半導体デバイスの開発などの研究に従事。



**篠崎 大和**

2024年 東京都立産業技術高等専門学校ものづくり工学科情報通信工学コース卒業。2026年 長岡技術科学大学電気電子情報工学分野卒業。現在, 同大学大学院工学研究科電気電子情報工学分野修士課程に在学中。応用波動光学, 光電磁気学など分野の研究に従事。



**戸塚 雄大**

2023年 東京都立産業技術高等専門学校ものづくり工学科情報通信工学コース卒業。2025年 同校専攻科創造工学専攻電気電子工学専攻修了。在学中は動画ホログラフィならびに光計測などの研究に従事。



**佐藤 甲癸 (フェロー)**

1975年 早稲田大学大学院理工学研究科電気工学専攻博士後期課程修了。1975年～1977年 同大学理工学研究所研究員。1977年 相模工業大学(現在校名変更により湘南工科大学)勤務。1995年 同教授。2012年 定年退職。その後湘南工科大学研究員, 早稲田大学理工学研究所招聘研究員, 東海大学非常勤講師。2016年 同大学非常勤講師定年退職。その間に応用物理学会編集委員, 電子情報通信学会動画ホログラフィ時限研究専門委員会委員長, 映像情報メディア学会映像表現研究会主査, 画像電子学会庶務理事, 技術専門委員長, 編集委員, 企画委員, 通信放送機構(TAO)高度立体動画通信プロジェクトサプリーダーなどを歴任。専門分野はホログラフィなどの立体映像技術, 光エレクトロニクス全般。電子情報通信学会, 映像情報メディア学会(フェロー), 画像電子学会(フェロー), SPIE各会員。工学博士。



**浅井 紀久夫**

1996年 名古屋大学大学院工学研究科博士後期課程退学。1996年 大学共同利用機関放送教育開発センター助手, 放送大学准教授を経て, 現在, 同大学教授。この間, 2000年にイリノイ大学シカゴ校(アメリカ), 2002～2003年にアルバータ大学(カナダ), 2009～2010年にカンタベリー大学(ニュージーランド)客員研究員。ヒューマンコンピュータインタラクションの研究, 高臨場感を提示する仮想環境システムの開発などに従事。電子情報通信学会, 人工知能学会, ACM各会員。博士(工学)。

## 特別史跡百済寺跡における Indirect AR・VR 統合体験システムの開発と社会実装

岡田 鍊<sup>†</sup> 濱田 悠太<sup>†</sup> 大野 倫敬<sup>†</sup> 川原 佑斗<sup>†</sup>  
 神田 啓冴<sup>†</sup> 和田 築来<sup>†</sup> 菅 沼 祐 耶<sup>††</sup> 中山 堪太<sup>††</sup>  
 河合 紀彦<sup>†</sup>(正会員)

<sup>†</sup>大阪工業大学 情報科学部, <sup>††</sup>大阪工業大学大学院 情報科学研究科

## Development and Social Implementation of an Integrated Indirect AR/VR Experience System at Kudaradera Temple Ruins, a Special Historic Site

Ren OKADA<sup>†</sup>, Yuta HAMADA<sup>†</sup>, Michitaka ONO<sup>†</sup>, Yuto KAWAHARA<sup>†</sup>,  
 Keigo KANDA<sup>†</sup>, Kizuku WADA<sup>†</sup>, Yuya SUGANUMA<sup>††</sup>, Kanta NAKAYAMA<sup>††</sup>,  
 Norihiko KAWAI<sup>†</sup>(Member)

<sup>†</sup> Faculty of Information Science and Technology, Osaka Institute of Technology,  
<sup>††</sup> Graduate School of Information Science and Technology, Osaka Institute of Technology

〈あらまし〉本研究では、大阪府枚方市にある特別史跡百済寺跡を対象として、制約条件の多い文化財現場における観光体験を支援する Indirect AR・VR 統合体験システムを開発し、社会実装を行った。本システムは、事前に取得した全方位画像を用いる Indirect AR (事前生成型拡張現実) による AR 体験と、視点の高さや天候・季節の切り替えが可能な VR 体験を組み合わせることで、現地環境に適した体験設計を実現している。提案システムは、2021 年から Android アプリとして段階的に改良を重ねながら自治体主催イベントにおいて運用を行い、その後、アプリのインストールを必要としない Web アプリ版へと展開した。さらに、自治体による QR コード設置を通じて、来訪者がいつでも利用可能な形での常設運用が行われている。これらの開発および運用の実践を通じて得られた知見を踏まえ、リアルタイム映像を用いた AR の適用が困難な文化財環境においても、AR/VR 体験を成立させるための設計方針および運用プロセスを、本論文では整理・提示する。本研究で示した知見は、文化遺産や遺跡など制約の強い屋外環境における AR・VR 活用において、有用な実践的指針となることが期待される。

**キーワード**：事前生成型拡張現実, バーチャルリアリティ, 全方位画像, 社会実装, 遺跡, 史跡

<Summary> This study develops and socially implements an integrated Indirect AR/VR experience system to support tourism experiences at the Kudaradera Temple Ruins, a Special Historic Site located in Hirakata City, Osaka, Japan, where strong constraints exist due to cultural heritage preservation. The proposed system combines an AR experience based on Indirect AR using pre-captured omnidirectional images with a VR experience that allows changes in viewpoint height as well as weather and seasonal conditions, thereby realizing an experience design suitable for the on-site environment. The system was first operated as an Android application at municipality-organized events from 2021, during which it was incrementally improved through repeated practical use. Subsequently, the system was extended to a Web-based application that does not require application installation. Furthermore, permanent operation has been achieved through the installation of QR codes by the local government, enabling visitors to access the system at any time. Based on knowledge obtained through these development and operational practices, this paper organizes and presents design principles and operational processes for realizing AR/VR experiences even in cultural heritage environments where the application of real-time camera-based AR is difficult. The findings of this study are expected to serve as practical guidelines for the use of AR and VR technologies in outdoor environments with strong constraints, such as cultural heritage sites and archaeological ruins.

**Keywords**: indirect augmented reality, virtual reality, omnidirectional image, social implementation, ruins

## 1. はじめに

日本には多くの遺跡が存在するが、その多くは建造物が現存せず、復元されていないことも多い。図1に示す大阪府枚方市にある百済寺跡<sup>1)</sup>もその一例であり、基壇、柱跡、築地塀が復元されているが、その上に建っていた建造物は再現されていない。百済寺跡は奈良時代後半に建てられたと考えられている百済寺の跡地であり、2025年7月現在、日本で64件しか指定されていない特別史跡の一つである。それにも関わらず、知名度がそれほど高くないため、知名度および魅力の向上を目指し枚方市では様々な手段で広報が行われている。

こうした課題への対応として、近年では、カメラ映像中の物体やマーカを認識し、現実空間に仮想物体を重ね合わせる技術である拡張現実 (AR: Augmented Reality)<sup>2)</sup>が活用されている。たとえば観光地においては、特定の物体やポスターをマーカとして利用し、それをスマートフォンのカメラで撮影することで、その場所に関する観光情報やキャラクターが表示される。このようなARを遺跡に応用することで、かつて存在した建築物や街並みをあたかも現代に蘇らせたかのような体験が可能となる。しかし、百済寺跡のような場所では、人工的なマーカの設置は景観上不適切であり、また、マーカとして利用可能な特徴的な物体もほとんど存在しない。そのため、マーカを用いた従来型のARは利用が困難である。

この問題を解決する手段として、本研究ではIndirect AR (事前生成型拡張現実)<sup>3)</sup>に着目する。この手法では、事前に取得した全方位画像に仮想物体を合成し、ユーザのモバイル端末の姿勢に応じて適切な方向の画像を切り出して表示する。これにより、提示される画像には周囲の観光客が映り込まず、現実シーンと仮想物体の間には位置ずれは一切生じない。また、リアルタイムのカメラ映像は用いずデバイス内蔵のセンサを用いるため、周囲の環境に影響されない頑健なARを実現できる。ただし、全方位画像を用いたIndirect ARは視点位置を移動できないという制約を持ち、これが一般的なARに比べて不利な点とされる。一方で、百済寺跡のように復元した基壇による段差が多く、歩きながらARを体験することが危険な環境においては、視点固定という制限がむしろ安全性の向上に寄与する。さらに、全方位画像を使用するIndirect ARは、同じく全方位画像を用いるパノラマVRと共通する技術基盤を有しており、両者間には高い親和性がある。そのため、ARからVRへの体験の移行も容易であり、統合的な体験設計を実現しやすい点もこの方式の利点のひとつである。

以上より、本研究では、過去に存在した建物を視覚的に再構成し、タブレットPC/スマートフォンを用いたARおよびVRにより提示する統合型体験システムを構築することで、遺跡観光の魅力の向上を図った。提案システムでのAR体験により、ユーザがその場に立って金堂や塔を目前に見ているかのような感覚を覚える。さらに、空中に浮かび上がったり、



図1 現在の百済寺跡

Fig. 1 Current scene of Kudaradera Temple Ruins

異なる季節や天候の風景へとワープしたりするVR体験へとシームレスに移行できる。提案システムは、最初はAndroidアプリ<sup>4)</sup>として実装し、枚方市主催の現地イベントにて高齢者を含む一般参加者による体験が行われ、好意的な評価を得た。さらにアプリのインストールを必要としないWebアプリ版を開発し、現地に設置されたQRコードを通じて誰でも即時にアクセス可能とし、実際に枚方市による常設運用が行われている。

本研究の貢献は、Indirect ARを用いた遺跡観光システムに関する新規な技術的手法を提案することではなく、制約条件の多い文化財現場において、AR/VR体験を成立させるための設計方針および運用モデルを、実践を通じて整理・提示した点にある。具体的には、リアルタイム映像を用いたARの適用が困難な環境を前提とし、全方位画像を用いたIndirect ARとVRを組み合わせた体験設計、ならびに、複数年にわたるイベント利用から常設運用へと段階的に改良を重ねた運用プロセスを示す。これらの知見は、文化財や遺跡など制約の強い屋外環境においてAR/VR体験を導入する際の実践的指針として有用である。

以下では、まず一般的なARの課題と活用事例および本研究の貢献について述べる。次に、Android版アプリの開発およびそれを用いたイベントの実施について述べる。次に、Web版アプリの開発およびその運用について述べる。最後に、全体の考察および今後の課題について述べる。

## 2. ARの課題と活用事例および本研究の位置づけ

### 2.1 ARの課題

遺跡でのARにおいて、過去に存在した建築物を復元しARによって提示する場合、現実空間と仮想建築物の位置合わせ、つまり幾何学的整合性が重要である。幾何学的整合性を実現するための、リアルタイムのカメラ映像を用いたカメラ位置姿勢推定手法としては、マーカの使用<sup>5),6)</sup>、特徴点ランドマークの利用<sup>7)</sup>、およびSLAM<sup>8)~10)</sup>がある。また、AppleのARKitやGoogleのARCoreといったライブラリでは、アルゴリズムは公開されていないがこれらの手法と端末の内部センサを融合してモバイル端末の位置姿勢が求められている。しかし、前述のように、周囲に多数の人がいる状況では画像の大部分が人などの障害物で覆われてしまい、カメラの

正確な位置姿勢推定が困難となり、仮想物体が適切な位置に描画されないという問題がある。さらに、カメラ画像の大部分が遮られると、ユーザが実際に見たい風景自体が見えなくなり、視認性の著しい低下が起きる。近年では、リアルタイム映像から不要な物体を除去する隠消現実 (DR: Diminished Reality) の研究<sup>11)~14)</sup>も行われているが、カメラが自由に動く環境下で高品質に移動物体を除去する技術はまだ確立されていない。そのため、リアルタイムカメラ映像を用いた AR は依然として頑健性や視認性の面で制約が残る。

## 2.2 遺跡での AR の活用事例

観光地や文化施設において、AR を用いた体験が各地で進められている。特に歴史的景観が失われた遺跡や文化財において、現地に復元した仮想建造物を重ねる AR は、往時の情景を直感的に想起させる技術として注目されている。多数の AR の事例があり、報告書・ウェブサイト<sup>15),16)</sup>などにまとめられている。

その中でも筆者らのシステム開発・イベント実施と同時期となる遺跡での活用事例として、奈良県桜井市による「YAM-ATO 桜井周遊 AR ガイド」<sup>17),18)</sup>が挙げられる。これは、市内の纏向遺跡や山田寺跡などの文化財を対象に、建造物の仮想物体を AR・VR により提示するシステムであり、Android および iOS 用のアプリによりスマートフォン・タブレット PC を用いて当時の寺院や都市空間を体感することができる。論文などでシステムの詳細は公開されていないが、筆者らの体験に基づくと、リアルタイム映像を用いた SLAM により回廊をウォークスルーできる AR 体験および、「CG 復元 VR 体験」という名前で Indirect AR 技術を用いた AR 体験が可能である。また、一部空中に視点移動する機能も実装されている。

上述のシステム以外にも、Indirect AR<sup>3)</sup>はその特性を活かして、国外においてリアルタイムカメラを用いる AR の代わりとして様々な用途に応用されてきた<sup>19)~21)</sup>。本研究と特に近い事例として、Liestøl ら<sup>20)</sup>は文化遺産において Indirect AR を用いて過去の建物を再構築・提示する試みを行っている。特に、視点の変化に関する多様なバリエーションが試行されており、本研究でも同様の機能を導入している。ただし、その詳細なシステム構成については明らかにされていない。

## 2.3 本研究の位置づけ

本研究は、百済寺跡という制約条件の多い文化財現場を対象として、遺跡観光における AR/VR 体験をどのように成立させるかという実践的課題に取り組むものである。前節までに述べたように、リアルタイム映像を用いた従来型 AR は、視認性や頑健性の観点から、遺跡環境において適用が困難な場合が多い。

そこで本研究では、百済寺跡の環境特性を踏まえ、事前に取得した全方位画像を用いる Indirect AR を基盤技術として採用し、さらに VR 体験を組み合わせることで、現地環境に



図 2 全方位カメラと 7.5m の三脚を用いた撮影

Fig. 2 Capturing with an omnidirectional camera and a tripod

適した統合的な体験設計を行った。また、体験設計にとどまらず、2021 年からの Android アプリとしてのイベント運用を起点として Web アプリ版への展開を行い、さらに 2025 年の自治体による QR コード設置を通じた常設運用へと、段階的に改良を重ねた運用プロセスを実践した点に本研究の特徴がある。

## 3. Android アプリ版提案システム

### 3.1 システムの概要

本章では、2021 年、2022 年、2023 年にかけて、段階的に機能追加および改良を行いながら開発・運用した Android アプリ版の提案システムについて述べる。各図では、対応する年次を併せて示す。なお、アプリをインストールせずに利用可能な Web アプリ版については、第 5 章で述べる。

本システムの構築では、まず事前に全方位画像を撮影・編集し、仮想の建築物を合成する。Indirect AR の機能では、合成した画像を球にマッピングし、仮想視点を球の中心に固定し、デバイスの向きに応じて視線を変えることで、全方位画像を切り出して表示する。VR の機能では、より高所からの視点の全方位画像や空を置き換えた全方位画像を利用し、天候・季節によってエフェクトなどを追加することで没入感のある体験を実現する。以下に、システム構築の各工程および主な機能について詳述する。

### 3.2 全方位画像の撮影

百済寺跡内のユーザが体験を行う場所で全方位画像を撮影した。カメラには「RICOH THETA X」を使用し、解像度は最大 11K (11008 × 5504 ピクセル) である。撮影には最大 7.5 メートルまで伸長可能な三脚「LUMICA Bi Rod 6G-7500」を使用した。本システムでは視点の高さを変化させる機能があるため、地上に立った視点の高さと、図 2 に示すように約 7.5 メートルの高さの 2 種類の画像を撮影した。それ



(a) 地上からの撮影画像



(b) 地上 7.5m からの撮影画像

図 3 撮影した全方位画像

Fig. 3 Captured omnidirectional images



図 4 不要物体を消去した全方位画像

Fig. 4 Omnidirectional image with unnecessary objects removed

それぞれの撮影結果例を図 3 に示す。なお、近年ではドローンによる空撮も行われているが、場所によっては飛行許可の取得が難しいため、本研究では三脚を用いた。

### 3.3 撮影画像の編集

撮影された全方位画像は、Indirect AR における現実背景として用いられるが、表示品質および VR 演出のために編集を行う。まず、画像内に含まれる不要物体を除去する。今回のシーンでは、「三脚」「人物」「カラーコーン」などの不要物を対象とした。これらの除去には、任意のインペインティング手法を用いることができるが、本研究では Adobe Photoshop の「コンテンツに応じた塗りつぶし」機能を用いた。図 4 に、図 3(a) の画像から不要物を除去した結果例を示す。

次に、VR 体験における天候の変化を実現するため、晴天画像を曇天に変換する。まず、画像中の空部分をセマンティックセグメンテーションにより抽出し、その領域を曇天の空に差し替える。本研究では同じく Photoshop の「空の置き換え」機能を使用した。最初にこの機能を使って空領域の色を単色の青に変換し、次に図 5 に示すように色温度を設定して曇り空を表現した。なお、これらの編集はすべて手動作業であり、不要物体の除去および空の置き換えを含めて、全方位

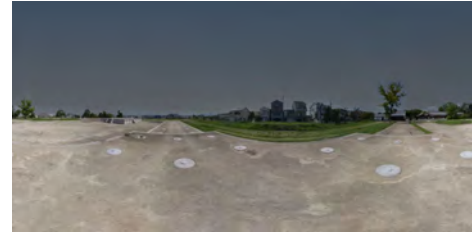


図 5 空を変更した全方位画像

Fig. 5 Omnidirectional image with sky changed

画像 1 枚あたりおおよそ数十分程度の時間を要した。

### 3.4 3D モデルの作成と画像への合成

枚方市から提供されたイラスト資料<sup>22)</sup>に基づき、3D モデリングソフトウェア Blender などを用いて百済寺の 3D モデルを作成した。具体的には、金堂、西塔、東塔、およびそれらを囲む回廊をモデリング対象とした。なお、西塔、東塔について、2021 年および 2022 年では、著者らが自らイラストを参考にモデリングを行ったが、2023 年には、奈良文化財研究所の評価を受けながら、建立の時代や規模が近い奈良県葛城市の當麻寺の西塔などを参考に改良を行った。モデリング作業には、3D モデリングソフトウェアへの習熟度、建築構造の理解・検討、および専門家による評価を受ける場合にはその打合せを含むため、対象モデルや改良内容に応じて、数日から数週間の時間を要した。

次に、編集済みの全方位画像を Blender における環境テクスチャとして設定し、原点を視点としたときに位置ずれが生じないように 3D モデルを手動で設置し、原点視点で全方位画像のレンダリングを行う。2021 年、2022 年の塔のモデルによるレンダリング結果および 2023 年の奈良文化財研究所の評価を受け改良したモデルによるレンダリング結果をそれぞれ図 6 に示す。背景の全方位画像は各年で別途撮影したものをを用いており、撮影条件の違いにより、塔の大きさ、位置、および姿勢に差が生じている。なお、この Blender を用いた合成作業についても、ソフトウェアへの習熟度に大きく依存し、1 視点あたり数十分から数時間の時間を要した。

### 3.5 AR と VR の統合システム

本システムは、Unity (バージョン 2021.3.2f1) を用いて、タブレット PC およびスマートフォン上で動作する Android アプリとして実装した。図 7 に 2022 年のシステムのインタフェースを示す。各機能の実装について以下に詳述する。

#### 3.5.1 見回し機能

3D モデルを合成した全方位画像を仮想空間内に配置した球に貼り付けて表示する。AR 体験時には、端末内蔵の電子コンパスとジャイロセンサを用いてデバイスの姿勢を取得し、球の中心の仮想カメラの回転を端末の回転と同期させることで、ユーザが任意の方向を見回すことが可能となる。図 8 にモバイル端末を回転させた際の画面表示を示す。



(a) 2021年のモデル (Ver.1)



(b) 2022年のモデル (Ver.2)



(c) 奈良文化財研究所の評価を受けた2023年のモデル (Ver.3)

図6 作成した百濟寺を合成した全方位画像

Fig. 6 Omnidirectional image with reconstructed Kudadera temple



図7 2022年のシステムのインターフェース

Fig. 7 System interface of 2022



図8 端末を回転させた場合の表示画面

Fig. 8 Screen when the mobile device is rotated

### 3.5.2 季節および天候の切り替え

季節や天候の切り替えは、Unityのパーティクルシステムを用いて実現した。パーティクルは空中から降下するように



(a) 秋のシーン

(b) 春のシーン



(c) 雪のシーン

(d) 雨のシーン

図9 季節と天候の切り替え

Fig. 9 Change in seasons and weather conditions

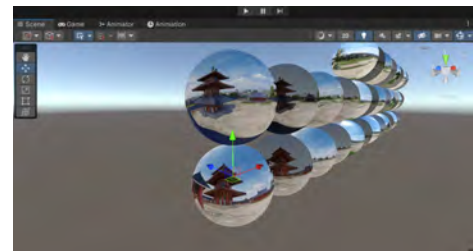


図10 Unityでの視点移動の実装

Fig. 10 Implementing viewpoint movement in Unity

設定されており、秋には紅葉(図9(a))、春には桜の花びら(図9(b))を用いた表現を行った。冬の雪景色では、編集済みの曇天画像とともに雪のパーティクルを用いた(図9(c))。雨の場合は線状のパーティクルと曇天画像を組み合わせることで表示した(図9(d))。これにより、来訪者は現在の気象や季節に限定されず、多様な環境下における寺院の雰囲気を楽しむことが可能となる。

### 3.5.3 地上視点と空中視点間の移動

地上視点と空中視点間の移動は、図10に示すように、地上および上空から撮影された全方位画像に3Dモデルを合成した画像をそれぞれ球にマッピングし、それぞれの球をUnity上で垂直方向に接続し、カメラ位置を上下に推移させることで実現した。なお、異なる季節や天候の全方位画像でも同様の視点移動を実現するため、それぞれの季節・天候での球体は水平方向に並列配置し、季節・天候に応じて各球の中心に仮想視点を移動した。しかし、図11に示すように、接続部付近に仮想カメラを配置すると視野の歪みが大きくなり、下の視点から上の視点へ移行した際には、一度沈んだ建物の下部や地面が再び現れるため、違和感が生じる。この問題に対して、接続部付近でのカメラ移動速度を急加速することで視覚の違和感を軽減している。図12は空中視点における表示画面である。この視点は約7.5メートルの高さから撮影されたものであるため、極端に高所という印象はないが、建物全体を見渡せるだけでなく、東塔および西塔の屋根瓦を詳細に観察できる視点を提供する。これにより、当時の鳥の視点や塔の上からの視点を想像させる体験を実現する。



図 11 視点移動中に見えるテクスチャの歪み

Fig. 11 Texture distortion during viewpoint shift



図 12 空中視点

Fig. 12 Aerial viewpoint



図 13 仮想物体が非表示の画面

Fig. 13 Screen with virtual objects hidden

### 3.5.4 仮想オブジェクトの表示・非表示切り替え

図 4 および図 5 に示すような 3D モデルを合成する前の全方位画像もシステム内に用意しておき、これと 3D モデル合成済みの画像を切り替えることで、仮想物体の有無を画面上で切り替えることが可能である。これにより、図 13 に示すような現実のみの画像も提示することで、たとえ提示画像がリアルタイム映像でなくても、ユーザにとってのリアリティが高まり、AR 体験の品質向上につながる。

### 3.5.5 塔の中への移動

本項から、2023 年に追加・改良した機能について示す。まず、奈良文化財研究所の評価に基づき、塔の外装を改良するとともに、塔の内部までモデリングを行った。図 14 に示すように、2023 年のシステムのインターフェースでは、スクリーン左の「中へ」アイコンを新たに配置し、タップすることで塔の内部へ視点を移動し、塔の内部を観察できる。実際の三重塔や五重塔の内部が公開されることは少なく、貴重な体験となる。

### 3.5.6 当時の人物の再現

2023 年のシステムでは、奈良時代の人物をモデリングし、仮想空間の球の中で移動させることで、図 15 に示すように当時の人物が寺院の中を歩いている様子を再現した。背景と建物には動きがないため、このような移動物体を配置することで、より臨場感高く当時の様子を想像できるようになる。



図 14 塔の中への移動

Fig. 14 Moving inside the tower



図 15 奈良時代の人物の描画

Fig. 15 Display of a human in the Nara period



図 16 ズーム機能

Fig. 16 Zoom function

### 3.5.7 ズーム機能

2023 年のシステムでは、二つの指を用いたピンチ操作により、仮想カメラの画角を変更し、図 16 に示すようにズームイン、ズームアウトの表示を可能にした。これによりモデリングされた塔などのより詳細まで観察できるようになる。

## 4. イベントでの利用

### 4.1 概要

本章では、提案システムをイベントにて実際に活用した事例について報告する。2021 年 11 月 23 日、2022 年 11 月 12 日、2023 年 11 月 24-25 日の計 3 回、枚方市主催の百済寺跡現地で開催されたイベントにおいて、枚方市広報誌などを通じて募集した一般市民を対象に、百済寺跡を紹介する一環として提案システムの体験を提供した。参加者数は、2021 年が 65 名、2022 年が 19 名、2023 年が 65 名であった。なお、毎年システムの改良を重ね、表示するモデルおよびシステムの機能はそれぞれの年で表 1 に示すとおりであった。また、枚方市主催イベントとは別に、枚方まつり 2023 において、現地とは異なる場所でも同システムの体験を提供した。なお、モバイル端末としては、NEC 製タブレット PC (LAVIE

表 1 各年のイベントで用いた提案システムの機能  
**Table 1** Functions of proposed system for events

	表示モデル	実装・機能
2021年	塔 (Ver.1) 金堂 (Ver.1)	見回し
2022年	塔 (Ver.2) 金堂 (Ver.2) 回廊	2021年と同じ機能 (見回し (図 8)) 不要物消去 (図 4) 季節・天候切り替え (図 5, 9) 地上・空中視点の移動 (図 10-12) 表示・非表示切り替え (図 13)
2023年	塔 (Ver.3) 金堂 (Ver.2) 回廊	2022年と同じ機能 塔の中への移動 (図 14) 歩く人物の表示 (図 15) ズーム (図 16)



図 17 百済寺跡での提案システムの使用

**Fig. 17** Use of the proposed system at Kudaradera Temple Ruins

T1295/DAS および T1195/BAS) を使用した。

#### 4.2 システム利用の様子とアンケート調査

図 17 に百済寺跡において実際にタブレット PC 上で動作している 2022 年の提案システムの様子を示す。図から、タブレット画面を通して寺院の風景を見ることができ、方角もおおむね一致していることがわかる。また、一般的に人は自分の目で見る景色と端末のカメラで撮影した画像で、明るさや色合いの違いがあることを経験しているため、この背景画像が事前に撮影されたものにもかかわらず、雲の模様といった大きなテクスチャの変化がない限りは当日の景観と見分けがつきにくい。

図 18 に 2022 年および 2023 年のイベントで参加者が同時に提案システムを体験している様子を示す。図から、複数人が同時に周囲の人の存在に影響されずにシステムを体験できていることが確認できる。また、図 17 に示すように、視点の高さや天候、季節の変更は画面上のアイコンをタップするだけで簡単に切り替えることができるため、高齢者でも事前に簡単な説明を受けるだけで問題なく操作できた。さらに、参加者からは、事前撮影された画像であるにもかかわらず現地の風景との不一致を感じなかったという意見が多く、AR 体験に加えて、天候や季節、視点がシームレスに切り替わる VR 体験も楽しんでいた。参加者からは、「モデルの再現度が高く、まるで実際にそこに存在するかのようだった」との肯定的な意見が寄せられた。



(a) 2022 年のイベント (b) 2023 年のイベント

図 18 イベントでの人々の同時体験の様子

**Fig. 18** Event where people use the system together

表 2 イベント参加者のうちアンケート回答者の年齢分布  
**Table 2** Age distribution of questionnaire respondents among event participants

年齢層	2021 年人数	2022 年人数
20 歳未満	4	0
20 歳代	1	0
30 歳代	2	0
40 歳代	1	0
50 歳代	4	0
60 歳代	20	6
70 歳代以上	29	11
合計	61	17

表 3 体験イベントの満足度

**Table 3** Participant satisfaction in the experience event

満足度	2021 年人数	2022 年人数
満足	58	17
どちらでもない	1	0
不満	1	0
未回答	1	0
合計	61	17

表 4 百済寺跡の再訪問の希望度

**Table 4** Intention to revisit the Kudaradera Temple Ruins

希望度	2021 年人数	2022 年人数
ぜひ訪れたい	41	8
機会があれば訪れたい	18	8
あまり訪れたいと思わない	0	1
訪れたくない	0	0
未回答	2	0
合計	61	17

さらに、2021 年および 2022 年に実施した体験イベントでは、枚方市により参加者を対象としたアンケート調査が実施された。2021 年のイベントでは参加者 65 名中 61 名から、2022 年のイベントでは参加者 19 名中 17 名から回答を得た。年齢分布、満足度、百済寺跡への再訪の希望度に関する結果を表 2, 3, 4 に示す。表より、参加者は 60 歳代以上の高齢者が多くを占めていたが、若い年代の参加者も一定数含まれていることがわかる。また、満足度で「満足」、再訪の希望度で「ぜひ訪れたい」または「機会があれば訪れたい」と回答した参加者がほとんどで、高い満足度が得られている。複数年にわたって同一の参加者を追跡する調査は行っていない

ため、年次ごとの操作性や満足度の直接的な比較はできないが、高齢者を含む幅広い年代の参加者が容易にシステムを利用し、満足度の高い AR 体験であったこと、また百済寺跡の魅力が伝わり再度訪問したくなるような動機の向上も、アンケート結果からも確認できた。

また、現地（百済寺跡）開催ではない枚方まつり 2023 では、端末で提示される映像と体験者周辺の景色が一致しないため、百済寺跡および奈良時代の様子を VR として提示し、200 人以上にシステムを体験してもらった。このような体験は、来場者の関心を喚起し、後に現地へ足を運ぶきっかけとなった可能性があると考えられる。

#### 4.3 考察と課題

本イベントを通して、歴史的資料および奈良文化財研究所の評価に基づいて再現された 3D モデルを含む風景を、複数人が同時に体験できる提案システムの有効性を確認した。また、周囲に他の参加者がいる状況でも頑健に動作することから、提案システムはこのような観光地における同時参加型の活用に適していることが実証された。また、本システムでは地上視点からの AR 体験だけでなく、タップ操作によって日常生活では得られない上空からの視点をういた VR 体験が可能である。加えて、天候や季節の変更が可能であるため、歴史的遺跡の多様なシーンを体験することができる。これにより、来訪者が異なる季節にも再訪したいという動機づけになると考えられる。

ただし、イベントではあらかじめインストール済みのモバイル端末を用意し、多くの参加者に体験してもらったが、より多くの人にその魅力を伝えるためには、誰でもいつでも気軽に現地で体験できることが望ましい。これに対して、このシステムは Android OS のタブレット PC やスマートフォンでのみ動作するアプリとして開発したため、iOS など他の OS を使用する来訪者は体験ができず、使用できるユーザが限定的である。また、観光地での見知らぬアプリのインストールに抵抗を感じる利用者も多く、体験開始の心理的障壁となる。近隣の施設で体験用の端末を貸し出す運用も考えられるが、台数の確保や管理、返却漏れなどの運用コストやリスクが高く、百済寺跡のような開放的な野外環境では適切な方法とは言いがたい。このため、モバイル端末にインストールするアプリではなく、一般的なブラウザで体験できる Web アプリとして開発することとした。次に、Web アプリの開発およびその運用について述べる。

### 5. Web アプリ版提案システムと運用

本研究では、ユーザがより気軽に体験可能な Web アプリ版システムを新たに開発した。スマートフォンやタブレットのブラウザ上で動作するこの Web アプリにより、利用者は専用アプリをインストールすることなく、現地に設置された案内板の中の QR コードを読み取るだけで、すぐに百済寺跡



図 19 百済寺跡での Web アプリの案内

Fig. 19 Guidance of Web app in Kudaradera Temple Ruins

の AR 体験を開始することが可能である。図 19 が 2025 年 3 月 31 日より実際に百済寺跡現地に案内板が設置され運用されている様子である。なお、Web アプリ版では、Indirect AR と VR による閲覧機能だけでなく、リアルタイムのウェブカメラ映像を用いた音声ガイドシステムも開発した。

#### 5.1 システムの技術基盤

Web アプリ版は、Three.js および AR.js を基盤として構築した。Three.js は WebGL ベースの 3D 描画ライブラリであり、軽量でかつ様々な端末での動作実績を有する。本システムでは仮想物体を合成した全方位画像を球にマッピングして、球の中心視点からの視野を画面に提示する。また、AR.js を活用することで、マーカ型 AR の表示にも対応しており、カメラを起動して特定のマーカを読み取ると、その場に 3D モデルが出現するような体験を提供できるよう設計されている。

#### 5.2 機能設計

Web アプリ版の機能は Android アプリ版の機能を基本的には踏襲しているが、インタフェースの変更および一部機能の簡略化を行った。図 20 に Web アプリ版の機能一覧を示す。なお、この図は上記の百済寺跡現地に設置している案内板の一部である。また、Android アプリ版と Web アプリ版の対応関係を表 5 に示す。表中の△については、機能の一部簡易化を示す。各機能について、Android 版と同様の機能・異なる機能に分け、それぞれ以下で述べる。

#### 5.3 Android 版アプリと同様の機能

- **見回し** Android アプリ版と同様に、電子コンパス・ジャイロセンサを用いることで、モバイル端末の見回しに応じて適切な方向の視野を提示する。なお、PC や上記の内部センサがない端末でウェブサイトを閲覧した場合には、マウスのドラッグまたは画面タッチによるスワイプで見回すことができる。
- **天気・季節の変更** 左上のアイコンにより天気・季節の変更ができる。ただし、Android 版では 3 次元的に花びらや雪などを描画したが、Web アプリ版では画面上の 2 次元的な動きとした。また、天候のアイコンについては、



図 20 Web アプリ版の機能一覧 (現地設置の案内板の一部)

Fig. 20 List of functions of web app

表 5 Android アプリ版と Web アプリ版における体験機能の比較

Table 5 Comparison of experience functionalities between the Android app and the Web app

機能	Android 版	Web 版
見直し	○	○
天気・季節の切り替え	○	△
視点高さの切り替え (地上/空中)	○	△
仮想物体の表示・非表示の切り替え	○	○
歩行人物の表示	○	×
ズーム操作	○	×
利用規約・歩きスマホ注意の表示	×	○
複数個所での AR 体験	×	○
音声・テキストガイド	×	○

各天気・季節のアイコンは最初は表示せず、左上の赤いアイコンをタップすることで、各アイコンを表示する折り畳み機能を実装した。これにより、アイコンによる視野の妨げを抑制し、没入感を損なわない工夫を行った。

- **視点の高さの切り替え** 右上のアイコンにより、地上視点と空中視点を切り替えることができ、建造物の俯瞰や屋根の詳細な観察も可能である。ただし、視点の滑らかな移動は省略し、画像の切替とした。
- **仮想物体の表示・非表示の切り替え** 図 21 に示すように、右上の上から 2 つ目のアイコンにより、塔・金堂・回廊の仮想物体の表示と非表示の切り替えができる。

#### 5.4 Android 版アプリと異なる新たな機能

- **利用規約および歩きスマホ注意の表示** イベントにおいてイベントの運営者やアプリの開発者が傍について体験するのは異なり、QR コードを設置して常時使用可能な Web アプリとして運用する場合には、トラブルが起きないよう責任の所在について明確にする必要がある。このため、Web アプリの起動と同時に利用規約を表示する。また、百済寺跡は基壇による段差が大きく、歩きスマホは転落や転倒の危険がある。このため、利用規約の同意後に図 22 に示す歩きスマホの注意喚起画面を表示する。
- **複数個所での AR 体験** Android アプリ版では 1 箇所の



図 21 仮想物体の表示・非表示機能

Fig. 21 Interface of web app and function of switching between display and non-display



図 22 歩きスマホ注意の表示

Fig. 22 Warning sign against walking while using smartphones

みからの AR 体験であったのに対し、Web アプリ版では、3 箇所において地上視点および空中視点で全方位画像を撮影し、それぞれの画像で仮想物体を合成しておくことで、3 箇所で AR 体験ができる機能を実装した。図 23(a) に示すように、右下の百済寺を俯瞰するイラストをクリックするとイラストが拡大され、AR 体験ができる場所がアイコンで表示される。異なる場所で AR 体験をしたい場合には、その場所のアイコンをタップするとともに、実際にその場所に移動することで、図 23(b) に示すように、その場所での風景で AR 体験ができる。この機能により、百済寺跡の構内を散策し、様々な視点での風景を楽しむことを促すことができる。

- **音声・テキストガイド** 左下のアイコンをタップすると、図 24 に示すような音声・テキストガイドシステムが起



(a) AR 体験の場所を示すアイコン (b) イラスト内左上視点での AR 画像

図 23 複数視点での AR 体験

Fig. 23 AR experience at multiple positions



図 24 音声・テキストガイドシステム

Fig. 24 Guide system using voice and text

動する。案内板の中のマークをカメラで写すと簡易的な 3D モデルが表示され、「次へ」ボタンをタップすると、史跡に関する解説がテキストで表示されるとともに音声再生される。これにより風景を楽しむだけでなく、百済寺跡の歴史を学ぶきっかけを与えている。

## 6. 考察と今後の課題

Android アプリ版でのイベントでの参加者の体験や、QR コードの設置による Web アプリ版での常時運用から、遺跡観光において Indirect AR と VR を統合したシステムが現地での風景を楽しむのに現状最適なものであると考え。しかし、Indirect AR の手法の特徴から、より臨場感高く AR を楽しむためにはいくつかの課題がある。

Indirect AR は事前に撮影された画像を用いる方式であるため、今回のイベントの事例では実際の景観と画像との間に大きな乖離はなかったものの、季節や天候が大きく異なる状

況では、ユーザ体験に悪影響を及ぼす可能性がある。これに対して、さまざまな天候や季節で事前に撮影された画像を用意し、利用時に適切なものを選択する手法<sup>23)</sup>が提案されている。また別の方法として、ユーザがその場で撮影したカメラ画像を用いたキャリブレーションにより、背景の色調やテクスチャをより現在の景色に近づける手法<sup>24)</sup>も提案されている。このような仕組みを取り入れることで、その時の風景と提示される画像の風景の差異を抑制することが望まれる。

また、本システムではユーザがその場から見回すことしかできず、視差を再現できないため、建造物の立体感やスケール感が伝わりづらいという問題もある。百済寺跡のような遺跡では、段差や他の来訪者の存在により、歩行中にモバイル端末の画面を見続ける歩きスマホは危険であるが、その場に立ち止まりながら端末を動かせる範囲での視差の提示はスケール感の提示に有効であると考えられる。このため、文献<sup>25),26)</sup>で提案されているようなその場で端末を平行移動したときの視差を再現する機能を追加することも検討する必要がある。

## 7. おわりに

本研究では、大阪府枚方市にある特別史跡百済寺跡を対象として、制約条件の多い文化財現場における観光体験を支援する AR・VR 統合体験システムを構築した。提案システムは、事前に取得した全方位画像を用いた Indirect AR による AR 体験と、視点の高さや天候・季節の切り替えが可能な VR 体験を組み合わせることで、現地環境に適した体験設計を実現している。また、本システムは Android アプリとして実装しただけでなく、アプリのインストールを必要としない Web アプリ版へと展開し、自治体による QR コード設置を通じて常設運用が行われている。これらの実践を通じて得られた知見を基に、リアルタイム映像を用いた AR の適用が困難な文化財環境においても、AR/VR 体験を成立させるための設計方針および運用プロセスを、本論文において整理・提示した。本研究で示した知見は、今後の文化遺産保全および観光振興における AR・VR 活用において、一つのモデルケースとなることが期待される。

## 謝 辞

本研究の実施にあたり、資料やフィードバック、ならびにイベント実施および社会実装の機会をご提供いただいた枚方市文化財課の皆様へ深く感謝する。また、百済寺の塔モデル構築にご協力いただいた奈良文化財研究所および大阪工業大学工学部建築学科の皆様、ならびにコンテンツ作成およびシステム開発にご協力いただいた同大学情報科学部情報メディア学科の皆様へ感謝する。

## 参考文献

- 1) 特別史跡 百済寺跡 | 枚方市ホームページ, <https://www.city.hirakata.osaka.jp/0000049862.html> (2025 年 7 月 7 日閲覧).

- 2) R.T. Azuma: "A Survey of Augmented Reality", Presence: Teleoperators and Virtual Environments, Vol.6, No.4, pp.355-385 (1997).
- 3) J. Wither, Y.-T. Tsai, R. Azuma: "Indirect Augmented Reality", Computers & Graphics, Vol.35, No.4, pp.810-822 (2011).
- 4) Y. Suganuma, M. Oda, K. Nakayama, S. Nishikawa, S. Hata, K. Paul, S. Wada, N. Kawai: "Integrated System of Augmented and Virtual Reality for Ruins Tourism", Proc. of NICOGRAPH International 2023, p.85 (2023).
- 5) H. Kato, M. Billinghurst: "Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System", Proc. of IEEE International Workshop on Augmented Reality, pp.85-94 (1999).
- 6) M. Fiala: "Designing Highly Reliable Fiducial Markers", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.32, No.7, pp.1317-1324 (2009).
- 7) T. Taketomi, T. Sato, N. Yokoya: "Real-time and Accurate Extrinsic Camera Parameter Estimation using Feature Landmark Database for Augmented Reality", Computers & Graphics, Vol.35, No.4, pp.768-777 (2011).
- 8) G. Klein, D. Murray: "Parallel Tracking and Mapping for Small AR Workspaces", Proc. of IEEE International Symposium on Mixed and Augmented Reality, pp.225-234 (2007).
- 9) R.A. Newcombe, S.J. Lovegrove, A.J. Davison: "DTAM: Dense Tracking and Mapping in Real-Time", Proc. of IEEE International Conference on Computer Vision, pp.2320-2327 (2011).
- 10) R. Mur-Artal, J.M.M. Montiel, J.D. Tardos: "ORB-SLAM: a Versatile and Accurate Monocular SLAM System", IEEE Trans. on Robotics, Vol.31, No.4, pp.1147-1163 (2015).
- 11) N. Kawai, T. Sato, N. Yokoya: "Diminished Reality Based on Image Inpainting Considering Background Geometry", IEEE Trans. on Visualization and Computer Graphics, Vol.22, No.3, pp.1236-1247 (2016).
- 12) N. Kawai, T. Sato, Y. Nakashima, N. Yokoya: "Augmented Reality Marker Hiding with Texture Deformation", IEEE Trans. on Visualization and Computer Graphics, Vol.23, No.10, pp.2288-2300 (2017).
- 13) S. Mori, S. Ikeda, H. Saito: "A Survey of Diminished Reality: Techniques for Visually Concealing, Eliminating, and Seeing through Real Objects", IPSJ Trans. on Computer Vision and Applications, Vol.9 (2017).
- 14) M. Kari, T. Grosse-Puppenthal, L.F. Coelho, A.R. Fender, D. Bethge, R. Schütte, C. Holz: "TransforMR: Pose-Aware Object Substitution for Composing Alternate Mixed Realities", Proc. of IEEE International Symposium on Mixed and Augmented Reality (2021).
- 15) 独立行政法人国立文化財機構奈良文化財研究所: "デジタルコンテンツを用いた遺跡の活用: 遺跡整備・活用研究集会報告書", 遺跡整備・活用研究集会報告書 27 (2016).
- 16) WebAR Lab - WebAR の最新情報が分かるメディア -, <https://webar-lab.palnar.com/> (2025年7月7日閲覧).
- 17) 「YAMATO 桜井周遊 AR ガイド」に新コンテンツが追加されました! / 桜井市, [https://www.city.sakurai.lg.jp/sosiki/kyouikuinkaijijimukyoku/bunkazai/YAMATO\\_sakuraiAR.html](https://www.city.sakurai.lg.jp/sosiki/kyouikuinkaijijimukyoku/bunkazai/YAMATO_sakuraiAR.html) (2025年7月7日閲覧).
- 18) 「史跡の魅力と考古学的価値を高める AR アプリ開発!」, <https://www.fujiya-net.co.jp/works/dx-promotion/202404011> (2025年7月7日閲覧).
- 19) D. Chatzopoulos, C. Bermejo, Z. Huang, P. Hui: "Mobile Augmented Reality Survey: From Where We Are to Where We Go", IEEE Access, Vol.5, pp.6917-6950 (2017).
- 20) G. Liestøl, A. Morrison: "Views, Alignment and Incongruity in Indirect Augmented Reality", Proc. of IEEE International Symposium on Mixed and Augmented Reality - Arts, Media, and Humanities (ISMAR-AMH) (2013).
- 21) J. Gimeno, C. Portalés, I. Coma, M. Fernández, B. Martínez: "Combining Traditional and Indirect Augmented Reality for Indoor Crowded Environments. A case study on the Casa Batlló museum", Computers & Graphics, Vol.69, pp.92-103 (2017).
- 22) 【文化財キッズページ】百済寺跡(くだらでらあと)|枚方市ホームページ, <https://www.city.hirakata.osaka.jp/0000035813.html> (2025年7月7日閲覧).
- 23) F. Okura, T. Akaguma, T. Sato, N. Yokoya: "Addressing Temporal Inconsistency in Indirect Augmented Reality", Multimedia Tools and Applications, Vol.76, pp.2671-2695 (2017).
- 24) H. Ikebayashi, N. Kawai: "Intensity and Texture Correction of Omnidirectional Image Using Camera Images for Indirect Augmented Reality", Proc. of International Workshop on Frontiers of Computer Vision (2024).
- 25) F. Okura, Y. Nishizaki, T. Sato, N. Kawai, N. Yokoya: "Motion Parallax Representation for Indirect Augmented Reality", Proc. of IEEE International Symposium on Mixed and Augmented Reality, pp.105-106 (2016).
- 26) T. Ueda, Y. Suganuma, N. Kawai: "Motion Parallax Reproduction in Indirect Augmented Reality Using 3D Gaussian Splatting for Mobile Devices", Proc. of IEEE International Symposium on Mixed and Augmented Reality Adjunct, pp.849-850 (2025).

(2025年7月7日 受付)

(2026年1月27日 再受付)



**岡田 錬**

2026年 大阪工業大学情報科学部情報メディア学科卒業。現在、大阪工業大学情報科学研究科博士前期課程に在学中。



**濱田 悠太**

2026年 大阪工業大学情報科学部情報メディア学科卒業。現在、大阪工業大学情報科学研究科博士前期課程に在学中。



**大野 倫敬**

2026年 大阪工業大学情報科学部情報メディア学科卒業。現在、キャノンマシナリー株式会社に勤務。



**川原 佑斗**

2026年 大阪工業大学情報科学部情報メディア学科卒業。現在、パナソニックシステムデザイン株式会社に勤務。



**神田 啓 冴**

2026年 大阪工業大学情報科学部情報メディア学科卒業。現在、東京コンピュータサービス株式会社に勤務。



**和田 築 来**

2026年 大阪工業大学情報科学部情報メディア学科卒業。現在、株式会社サイバーリンクスに勤務。



**菅 沼 祐 耶**

2024年 大阪工業大学情報科学部情報メディア学科卒業。2026年 大阪工業大学情報科学研究科博士前期課程修了。現在、株式会社日立ソリューションズに勤務。



**中山 堪 太**

2024年 大阪工業大学情報科学部情報メディア学科卒業。2026年 大阪工業大学情報科学研究科博士前期課程修了。現在、理想科学工業株式会社に勤務。



**河合 紀 彦 (正会員)**

2010年 奈良先端科学技術大学院大学情報科学研究科博士後期課程修了。同年 日本学術振興会特別研究員 (PD) 兼、カリフォルニア大学バークレー校博士研究員。2011年 奈良先端科学技術大学院大学情報科学研究科助教。センスタイムジャパンを経て、2020年 大阪工業大学情報科学部情報メディア学科准教授、現在に至る。博士 (工学)。コンピュータビジョン、複合現実感、バーチャルリアリティに関する研究に従事。

## 乗客行動の認識に基づくバス発車判断の自律化システム

山崎 将幸<sup>††</sup> 滝沢 穂高<sup>††††††</sup>(正会員)

<sup>†</sup>トヨタ自動車株式会社, <sup>††</sup>筑波大学情報理工学学位プログラム, <sup>†††</sup>筑波大学システム情報系,  
<sup>††††</sup>筑波大学学術情報メディアセンター

### Autonomous Decision System for Bus Departure Based on Passenger Behavior Recognition

Masayuki YAMAZAKI<sup>††</sup>, Hotaka TAKIZAWA<sup>††††††</sup> (*Member*)

<sup>†</sup>Toyota Motor Corporation, <sup>††</sup>Doctoral Program in Computer Science, University of Tsukuba,  
<sup>†††</sup>Institute of Systems and Information Engineering, University of Tsukuba,  
<sup>††††</sup>Academic Computing and Communications Center, University of Tsukuba

**〈あらまし〉** 本研究では、バス発車時の安全確認における運転手の負担軽減と安全性の向上を目指し、RGB カメラ画像認識を用いた乗客安全確認システムを提案する。提案システムは、カメラ映像から乗客定員超過、駆け込み、車内移動、ふらつき、離席、転倒をリアルタイムに検知する。これらの情報を基に乗客安全の観点で発車可否を判断することで、運転手や自動運転システムによる交通状況の確認結果と合わせて発車判断の自律化を目指す。車載環境での実用性を考慮し、演算負荷を軽減させた高速軽量なシステムとして実装した。実証実験の結果、提案システムは実用的な運行環境下において高い精度を示し、発車判断の自律化に貢献できる可能性が示唆された。今後は、提案システムを用いた車両制御システムとして、運転手や乗客への心理的な影響評価や社会的受容性を検討していく。

**キーワード**：バス乗客認識, 人物行動認識, 車載画像認識, リアルタイム処理, 判断の自律化

**<Summary>** This study proposes a passenger safety confirmation system using RGB camera image recognition to reduce the driver's workload and improve safety during safety confirmation at bus departure. Our proposed system detects in real time from camera images whether the vehicle is over capacity, rushing, moving inside the vehicle, staggering, leaving the seat, or falling over. Based on this information, the proposed system determines whether or not to start the vehicle from the viewpoint of passenger safety, aiming to automate the decision to start the vehicle together with the results of traffic condition checks by the driver and the automated driving system. Considering practicality in an in-vehicle environment, the system was implemented as a high-speed, lightweight system with reduced computational load. Demonstration experiments showed that the proposed system is highly accurate under practical operating conditions, suggesting that it may contribute to the automation of vehicle departure decisions. In the future, we will evaluate the psychological impact of the proposed system on drivers and passengers and study its social acceptability as a vehicle control system.

**Keywords:** bus passenger recognition, person behavior recognition, in-vehicle image recognition, realtime processing, autonomous decision making

### 1. はじめに

従来、バス運転手は乗客の安全を目視で確認し、車両の発車判断を行う必要があった。特に発車時は乗客の転倒事故が発生しやすく、運転手は細心の注意を払う必要がある。しかし、車内の混雑時など視界不良な状況下で

は、乗客の状況を正確に把握することは困難である。その結果、発車判断は運転手の経験や余裕に依存<sup>1)</sup>し、判断にばらつきが生じることがある。運転手が安全確認を十分に行わずに発車判断を行うと事故のリスクが高まり、発車判断が遅れると運行遅延のリスクが高まる。このよ

うに、発車判断は安全性を確保しつつ効率的な運行を実現するための重要な課題である。

一方、日常的にはバス停での乗客の駆け込み、立ち乗り、発車直前に車内を移動する乗客行動が存在する。これらの乗客行動は乗客自身の事故リスクを高めるだけでなく、運転手の発車判断を遅延させる要因ともなる。また、発車判断は依然として運転手の目視に依存しており、人的な運行コストを削減する際の障壁となる。例えば、自動運転バスでは発車判断を簡素化するために着座に制限された運用が多く、発車判断は乗客の輸送効率を改善する際の障壁でもある。乗客安全のための人員を同乗させる運用もあり、運転手無人化による効率化効果は限定される。この課題に対し、自動運転タクシーではシートベルトセンサや座席重量センサによって乗客の着席状態を自動判定し、乗員の着座を発車判断の条件としている。しかしながら、中型以上のバスでは、全座席へのセンサ設置はコストやスペースの観点から困難な場合がある。従来、発車時の乗客認識に焦点を当てた研究事例は少なく、発車判断のための着座検知は4~6人程度の車両<sup>2)</sup>が対象であり、多人数や立ち乗りなど日常的な乗車行動には対応していない。

そこで本研究は、日常的な発車時の乗客認識に焦点を当て、バス発車判断を自律化するための車載画像認識による乗客安全確認システムを提案する。この自律化は、従来の運転手が行っていた認知・判断・操作をシステムで代替するものであり、発車判断を均質化し、事故と運行遅延のリスクを抑制する。さらに、運転手の作業負担を軽減し、省力化として運行効率の改善にも寄与することを目指す。発車判断は乗客安全と交通環境の確認で構成され、前者は本システム、後者は運転手や自動運転システムが担うものとする。本システムは魚眼カメラの広い視野角を活用し、乗客認識に基づいて発車前の乗客安全を確認する。

本システムの目標は、運転手の平均的な発車判断時間(4秒)を基準に、バス車載リアルタイムシステムとして実装することである。その際、演算機の性能要件低減のために、CPU6コアと1スロット型低消費電力GPU(NVIDIA RTX A4000)を対象とし、演算周期6fpsかつ演算時間100msec/フレーム以内の制約下で高速かつ軽量のシステムを構築する。

## 2. 関連研究

### 2.1 RGBカメラを用いた行動認識

RGBカメラを用いた行動認識には、以下の4つの課題が存在する。第一はカメラ画角とドメイン適合である。RGB画像は奥行き情報を持たないため、画像特徴はカメラ画角やセンサ特性に強く依存する。これにより、取



図1 本手法のカメラ構成と乗客行動の例  
Fig. 1 Camera locations and examples of passenger's behavior

得されるデータは姿勢の見えやオクルージョン、照明、画像歪み、背景領域、フレームレートなど、様々な条件が固定される。このため、あるドメインで学習された認識モデルを別のドメインに適用することは困難である。この対策として、骨格推定に基づいた手法が用いられ、RGB画像から直接推論する手法と比較して、オクルージョンや照明、視野角、背景の影響が低減される<sup>3)</sup>。

第二はデータセットの作成である。前述のドメイン適合の課題から、行動認識の対象や適用ドメインに応じて、独自でデータの収集とアノテーションが必要となる場合が多い。特に、教師ラベルの定義はデータセットによって異なり、例えば「歩く」と「走る」の区別は移動速度で定義するなど、見た目ではない基準で正解付けされることがあり、複数のデータセットを単純に統合することは困難である。この対策として、教師なし学習の手法が提案されている。具体的には、ドメイン適合先の教師ラベルなしデータに対してデータ蒸留<sup>4)</sup>や蒸留学習<sup>5),6)</sup>により疑似的な正解ラベル付けを実現する。また、データセットはプライバシーや機密性の課題があり、公開データセットは限られている。

第三は行動認識における時系列処理である。行動は種類によって所要時間が異なり、例えば「起立」や「転倒の瞬間」のように瞬間的な動作から、「歩行」や「転倒状態」のように状態が一定時間継続する動作まで存在する。さらに、行動速度は人物の状況や個人差によって変動する。この課題に対して、時間的な特徴を捉えるためにLSTM, ST-GCN, 3DCNNなどを活用した手法<sup>7)~9)</sup>が提案された。しかしながら、画像特徴と時間的な特徴をバランス良く考慮することには限界があることが指摘されている<sup>10)</sup>。

第四は行動認識の演算負荷である。3D 骨格推定や3DCNNを用いる手法は高精度化に寄与する一方で、演算負荷は高くなる。また、2D 骨格推定やLSTMを用いる手法は演算負荷を比較的抑制できる一方、精度は低下する。骨格推定の手法はBottom-upとTop-downに分類され、骨格推定に基づいた行動認識は骨格推定の精度に依存する。Bottom-up手法は、OpenPose<sup>11)</sup>に代表されるように、RGB画像を入力として骨格情報を直接推定するため、追加学習を行わない場合は天井カメラからの俯瞰視点やオクルージョンに対して認識精度が低下しやすい。Top-down手法は、HRnet<sup>12)</sup>に代表されるように、YOLO<sup>13)</sup>などの物体検出技術を用いて人物の位置を特定した後、その人物を囲む矩形領域に対して骨格推定を演算するため、比較的高い精度が得られる。しかし、演算対象となる人物の数が増加するにつれて演算時間と演算負荷は増大する。このように、システム実装においては認識精度と演算負荷の間にトレードオフの関係が存在するため、実装対象に合わせて様々な高速化手法が研究されている。

## 2.2 バス乗客の行動認識

RGBカメラを用いたバス乗客の行動認識技術は、主に以下の2つのタスクに分類できる。第一は乗客人数カウントである。このタスクは、バス車内の乗客数を推定し、混雑状況や人の流れを把握することを目的とする。多くの研究では、物体検出技術とMOT (Multi-Object Tracking) 技術を組み合わせた手法が提案された<sup>14)~16)</sup>。これらの研究では、事前にバス車内の画像データを収集し、そのデータを用いて物体検出モデルを学習させることで、乗客の位置認識のためのドメイン適合が図られた。しかしながら、組込システムへの実装では、モデル軽量化に伴う性能限界が指摘された<sup>14)</sup>。より効率的な技術構成や、計算資源の制約下での性能向上に関する研究が求められる。

第二は乗客の行動監視である。このタスクは、乗客の姿勢や動作、周囲の状況との関係性を分析し、転倒や喧嘩などの異常行動を検知し<sup>17)~20)</sup>、乗務員や関連スタッフに警告を発することを目的とする。異常行動の検知では、骨格推定とLSTMを用いた時系列的な姿勢解析<sup>17)</sup>、3DCNNを用いた時系列の画像テクスチャ解析<sup>18)</sup>、飲食物や煙草の検出による行動解析<sup>20)</sup>などの手法が提案された。このように、既存研究は監視カメラの応用に主眼を置いたものが多く、発車時などのバス固有のユースケースに対応した研究は少ない状況である。一方、システム事例では、フジクラ<sup>21)</sup>や東海理化<sup>22)</sup>が離席検知を、SBドライブ<sup>23)</sup>はさらに移動検知と転倒検知を発表した。しかし、これらのシステムは運転手への注意喚起を主な目

的としており、人の作業を完全に置き換えるものではないため、一部の行動種別の検知に留まる。このため、立ち乗りや駆け込み乗車など、運転手が特に注意を払うべき乗客行動には対応していない。

## 3. 提案手法

### 3.1 発車判断の基準と検知対象

本研究では、バスの発車時における乗客の安全確認を自律化するために、運転手の判断基準を設計し、システムへ構築する。そのために、RGBカメラ画像認識を用いた複数検知に基づく発車可否判定の手法を提案する。本手法は、2つのカメラを入力として、知覚層と検知層、判定層で構成される。図1に示すように、RGB魚眼カメラをバス車両の車外と車内に各1個を設置する。車外カメラは、地面からの設置高2.5mの車両屋根上に設置し、乗降口を撮影する画角とする。車室内カメラは、乗降口上部の天井に、車内床からの設置高1.9mで下向きに設置する。

図2に、乗客の安全確認のための判断基準を示す。判断基準は、演算量の少ない順に適用され、図2の[3]以降の判断基準は、いずれか一つでも該当した乗客が存在する場合、発車不可と判定する。

以下に、6つの判断基準と、それらに基づく検知対象を示す。

- 1 定員超過: 車内乗客数が乗客定員数を超過している状態
- 2 駆け込み: 車両外側から乗降口に近づく乗客
- 3 車内移動: 車内を移動する乗客
- 4 身体ふらつき: 立ち乗り/離席中などで大きな身体動作を行う乗客
- 5 座席離席: 座席位置で着座していない乗客
- 6 転倒状態: 転倒した状態が3秒以上続く乗客

図3に、提案手法の処理フローを示す。知覚層は骨格推定に基づいた行動認識を用いて、全3ステージで構成される。第1ステージは人物検出、第2ステージは骨格推定、第3ステージは着座認識と転倒認識を行う。車外カメラ画像は第1ステージのみ、車室内カメラ画像は全3ステージまで処理し、検出結果を出力する。検知層は、

表1 検知層の各検知方法と知覚層の各出力の対応表  
Table 1 Mapping between detection methods and perception outputs

検知方法	知覚層の出力			
	人物検出	骨格推定	着座認識	転倒認識
[1] 定員超過	○			
[2] 駆け込み	○			
[3] 車内移動	○			
[4] 身体ふらつき		○		
[5] 座席離席	○		○	
[6] 転倒状態	○			○

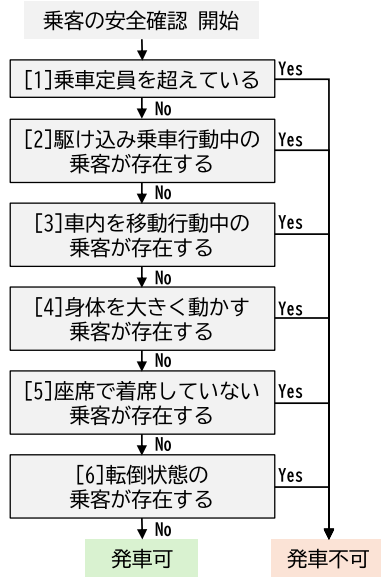


図2 本手法における発車可否の判断基準

Fig. 2 Criteria of departure decision in our proposed method

上記6つの判断基準に対応した検知方法で構成され、各検知方法は各々の検知有無を出力する。表1に、検知層の各検知方法の入力となる知覚層の出力の対応を示す。詳細は後述の各節で述べる。表1は、各検知方法への入力となる知覚層の出力結果をチェックマークで示したものであり、詳細は後述の各節で述べる。判定層では、検知層からの出力に基づいて発車可否判定を行い、「発車可」または「発車不可」を出力する。実運行では、本システムが「発車不可」を出力した場合、運転手は乗客に対し、着席や降車の誘導、転倒者対応などの現場判断を行う。上記6つの検知が全て「検知無」の場合、「発車可」を出力する。いずれか1つでも「検知有」の場合、その時点で判定を終了し、「発車不可」を出力する。以下に、知覚層と検知層の詳細を述べる。

### 3.2 人物検出

人物検出は物体検出技術とMOT技術を用い、全身検出と頭部検出を行う。まず、各カメラ画像を入力として物体検出の推論を実行する。次に、現フレームにおいて検出矩形が存在しない場合は、予測矩形を検出矩形として扱う。最後に、推論された部位に対してMOTを適用し、過去フレームにおける検出矩形から現フレームにおける予測矩形を算出するとともに、ReID(Re-Identification)を行う。全身検出矩形と頭部検出矩形を出力する。

### 3.3 骨格推定

骨格推定では、演算時間と演算負荷の抑制及び俯瞰視点の魚眼歪みに対する頑健性を両立するため、2D骨格推定のTop-down手法を採用する。Top-down手法による骨格推定は、人物検出の検出人数分、全身検出矩形ご

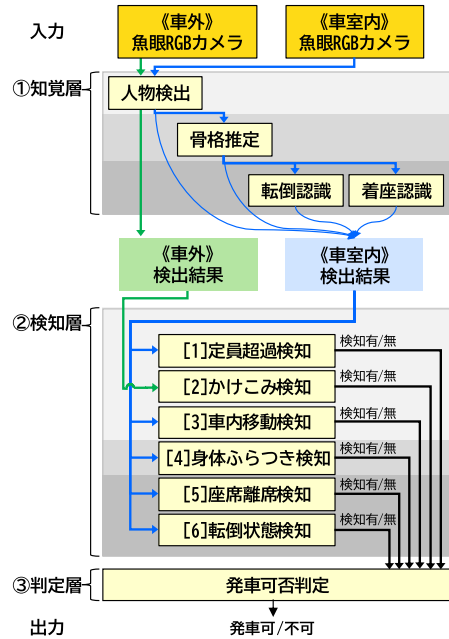


図3 本システムの処理フロー

Fig. 3 Processing flow

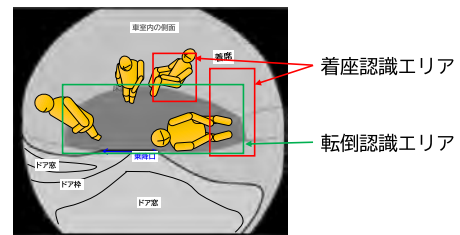


図4 着座認識・転倒認識における演算対象エリア

Fig. 4 Computational areas for sit and fall recognition

とに処理を行うため、HRNet<sup>12)</sup>など単純な逐次処理では人数に応じて演算時間と演算負荷が増大する。本手法では、この課題を回避するため、モデル入力にバッチ処理<sup>24)</sup>を適用し、バッチサイズを想定人数より大きく設定することで、複数人分の矩形を一括で推論する。同時に、誤推定抑制のため、サイズの小さい全身検出矩形を閾値処理により演算対象外とする。

この結果、メモリ転送を含めた演算時間の短縮と負荷軽減を実現し、特にCPUコア数やメモリバス帯域が少ない演算環境で効果が発揮される。

### 3.4 着座認識・転倒認識

着座認識は着座状態と離席状態(非着座状態)を判別する。転倒認識は転倒状態と非転倒状態を判別する。着座認識と転倒認識は、基本的な処理フローは同一であり、判別モデルを採用する。両者の違いは、認識エリアと判別モデルが異なる。認識エリアは、着座認識では車内画像の座席領域、転倒認識では車内画像の床領域に対して、画像座標上に多角形として事前に設計する。図4は、2つの着座認識エリアと1つの転倒認識エリアを設計した例である。

処理の流れとして、まず全身検出矩形の単位で、各々の認識エリアとのIoU(Intersection over Union)マッチングを計算する。次に、IoUが閾値以上の全身検出矩形に対し、その骨格推定結果 $(X, Y)$ に対して式(1)により座標のスケーリング処理を行う。式(1)では $(X', Y')$ の値域を0~1.0へ正規化するために、首から腰までの骨格点を結ぶ直線の長さを $L_x, L_y$ とする。首と腰の骨格点が検出されない場合は、全身検出矩形の縦・横サイズを各々 $L_x, L_y$ とする。最後に、 $(X', Y')$ を入力とし、各々の判別モデルを用いて判別処理を行い、判別結果を出力する。

$$\begin{pmatrix} X'_i \\ Y'_i \end{pmatrix} = \begin{pmatrix} \frac{1}{L_x} & 0 \\ 0 & \frac{1}{L_y} \end{pmatrix} \begin{pmatrix} X_i \\ Y_i \end{pmatrix} = \begin{pmatrix} \frac{X_i}{L_x} \\ \frac{Y_i}{L_y} \end{pmatrix} \quad (1)$$

### 3.5 定員超過検知

定員超過時は乗客への降車を促すため、定員超過状態を検知する。本検知では、乗降時に乗降口を通過した頭部検出矩形に対し、乗車方向と降車方向別にカウントする。乗車方向の個数から降車方向の個数を減算し、その結果が乗客定員を超える場合は「検知有」、超えない場合は「検知無」を出力する。頭部が検出されない乗客は、全身検出矩形をカウントに用いる。

### 3.6 駆け込み検知

発車直前に乗車しようと近づく乗客を検知する。人の反応速度0.2秒より速く、乗客の接近を予測することでブレーキを踏むまでの反応時間を確保する。本検知では、全身検出矩形の時間的な移動速度と、事前に設定した駆け込み乗車検知エリアへの到達予測に基づいて検知する。その検知エリアの例は、図1左上の車外カメラ画像内に赤色の多角形で示す。まず、車外カメラ画像の全身検出矩形に対し、SORT<sup>25)</sup>の方法を用い、矩形位置の予測ステップを2回連続で適用することで2フレーム先の矩形位置を予測する。そして、予測された矩形と検知エリアとの重複を比較し、重複する場合は「検知有」、重複しない場合は「検知無」を出力する。

### 3.7 車内移動検知

車内を移動中の乗客は出発準備が完了していないとみなし、その移動行動を検知する。移動検知の基準は80cm/2秒とする。瞬間的な動作(足の開閉など)を移動と誤検知しないために、2秒間の移動距離で検知する。基準として用いる閾値は、実際に車内各位置で人物が80cm移動した実験を複数回行い、その際の魚眼補正後画像に対して頭部矩形を付与し、その中心座標のフレーム間移動量の最小値から求める。本検知では、車外カメ

ラ画像の頭部検出矩形に対して、その時間的な移動量から検知する。頭部の未検出は1秒未満までを許容とするために、その移動量の算出区間と検知結果の算出区間を区別して2段構成とする。

本処理は以下3つの処理で構成される。第一に、頭部検出矩形を基に、各フレームにおいて頭部矩形の中心座標の移動量を、過去1フレームとの差分として算出する。このとき、画像から乗客の移動距離を近似的に求めるために、OpenCVで用いられる魚眼レンズ補正を適用し、魚眼歪みの影響を低減させる。第二に、6fpsの場合、現フレームから過去6フレームにおける移動量の平均値が閾値以上である場合は移動量フラグを1、閾値未満の場合は移動量フラグを0とする。このとき、頭部矩形が未検出のため移動量が算出できないフレームは、移動量の平均値の算出対象から除外する。第三に、過去6フレームにおいて、移動量フラグ1が継続した場合は「検知有」、移動量フラグ0が継続した場合は「検知無」、それ以外の場合は前回値を出力する。

図5に本処理の例を示す。時刻 $t$ を基点として説明すると、 $t-10$ から $t-5$ は頭部の移動量を算出する区間であり、 $t-5$ 時点の第二処理における閾値判定で移動量フラグが1となる。次フレーム以降から $t$ まで、6フレーム連続で移動量フラグが1であるため、時刻 $t$ における判定結果は「検知有」となる。それ以前の $t-2$ と $t-1$ は6フレーム連続の条件を満たさなため「検知無」となり、 $t-2$ 以前は必要な算出区間を満たしていないため待機状態となる。このように、本検知のレイテンシは合計10フレーム(1.67秒)である。

### 3.8 身体ふらつき検知

立ち乗り時や着座前などの体勢を崩しやすい乗客を想定し、車内の移動を伴わない範囲で大きな身体動作がある場合、その乗客は出発準備が完了していないとみなし、この行動を検知する。判断基準としての大きな身体動作とは、一歩踏み出す程度の動きと定義し、本検知の閾値は実験的なデータ取得に基づいて設定する。

本検知では、骨格点の時間的な移動量に基づいて大きな身体動作を検知する。時系列処理は図5の車内移動検知と同様とし、車内移動検知の違いは本検知が骨格点を特徴量とすることである。

本処理は以下3つの処理で構成される。第一に、骨格検出結果を基に、各骨格点の移動量を過去1フレームとの差分として各フレームで算出し、骨格ごとに設定した加重係数を用いて加重平均を算出する。このとき、未検出の骨格点は、差分算出の対象から除外する。第二に、6fpsの場合、現フレームから過去6フレームにおける第一の移動平均値が閾値以上であれば骨格移動量フラグを

frame	t-12	t-11	t-10	t-9	t-8	t-7	t-6	t-5	t-4	t-3	t-2	t-1	t
検知結果	待機	待機	待機	待機	待機	待機	待機	待機	待機	待機	無	無	有
移動量 フラグ 0: 閾値未満 1: 閾値以上						0	0	1	1	1	1	1	1

図5 車内移動検知における時系列処理 (6fps の場合): 時刻 t において「過去 6 フレームにおける移動量の平均値が閾値以上」の状態が 6 フレーム連続で続いた例

Fig. 5 Time-series processing for in-vehicle movement detection (6 fps): An example where the condition “the average movement over the past 6 frames is above the threshold” continues for 6 consecutive frames at time t

1, 閾値未満であれば骨格移動量フラグを 0 とする。過去 6 フレームにおいて、骨格移動量フラグ 1 が続いた場合は「検知有」、骨格移動量フラグ 0 が続いた場合は「検知無」、それ以外の場合は前回値を出力する。

### 3.9 座席離席検知

座席位置に存在する乗客が着座していない状態を検知する。着座認識の出力尤度が閾値未満の場合、離席フラグを 1 とする。閾値以上の場合、離席フラグを 0 とする。著者の計測によれば、離席から着座までの時間が最短で約 0.5 秒であったことから、着座状態が 0.5 秒以上継続しない場合は離席とみなす。したがって、6fps の場合、過去 6 フレームにおいて 4 フレーム以上が離席フラグ 1 の場合は「検知有」、それ以外の場合は「検知無」を出力する。

### 3.10 転倒状態検知

車内で 3 秒以上転倒している状態の乗客を検知する。着座認識の出力尤度が閾値以上の場合、転倒フラグを 1 とする。閾値未満の場合は転倒フラグを 0 とする。転倒認識は 1 秒ごとに実行し、過去 3 秒間で転倒フラグ 1 が連続した場合は「検知有」、それ以外の場合は「検知無」を出力する。

## 4. 実 装

### 4.1 知覚層と検知層の実装

人物検出は、軽量かつアンカーフリーな手法である YOLOX-m<sup>13)</sup>を用いた。COCO データセット<sup>26)</sup>とオリジナルデータセットを用いて学習を行い、FP16 へ量子化した。学習データセットは合計 29,145 フレーム、教師ラベルは全身矩形 97,240 個、頭部矩形 9,564 個である。教師ラベルの作成では、手作業による負担を軽減するため、Mask R-CNN<sup>27)</sup>によるインスタンスセグメンテー

ションを適用し、その出力結果から得られた人物領域の外接矩形を疑似的な正解矩形とした。必要に応じて、手作業による修正を加えた。検証データセット (577 画像, 2,478 人) を用いた評価の結果、学習前の AP(Average Precision)@50 が 0.756 に対し、学習後は 0.938 となり、0.182 ポイントの改善が確認された。MOT には、オクルージョンへの頑健性と軽量な演算処理を考慮し、矩形ベースの手法である OC-SORT<sup>28)</sup>を用いた。

骨格推定には、軽量かつ高精度な手法である ViTPose-B<sup>29)</sup>を用いた。COCO データセットとオリジナルデータセットを用いて学習を行い、FP16 へ量子化した。学習データセットは 74,600 フレームで構成され、教師ラベルは骨格点 275,219 人分である。教師ラベルの作成では、上記のうちオリジナルデータセットは 10,485 フレーム 12,754 人に対し、高精度な骨格推定器である HRnet<sup>12)</sup>を適用して疑似的な正解を作成し、その結果を必要に応じて、手作業で修正した。学習後の検証データセット (5,000 フレーム, 11,004 人) に対する AP@50 は 0.726 であった。上記の学習により RGB 魚眼カメラの画角へ対応したことで、COCO2017val データセットに対する AP@50 は 0.885 となり、ViTPose-B の論文掲載値 0.905 に対し、0.020 ポイントの低下に留まることを確認した。

着座認識と転倒認識は GPU へ演算負荷を与えないために、CPU 上で演算可能な手法である LightGBM<sup>30)</sup>を各々の 1 クラス判別モデルとして用いた。教師ラベルとして、着座認識には着座状態 3,015 フレームと離席状態 (非着座状態) 2,355 フレーム、転倒認識には転倒状態 8,411 フレームと非転倒状態 8,411 フレームを手作業で付与して、各々の判別モデルは学習させた。

## 4.2 システム構成

本システムは、乗客定員 15 名の小型バス車両に 2 個の RGB 魚眼カメラと車載演算機を搭載する。RGB 魚眼カメラの仕様は、解像度 1280 × 960 画素、FOV170 度、フレームレートは 6fps である。車載演算機の構成は、Intel Core i7-12800H (14 コア)、メモリ 32 GB、1 スロット型低消費電力 GPU (NVIDIA RTX A4000)、UbuntuOS 20.04 である。なお、他のサービスアプリケーションとの共存を考慮し、本システムに割り当てる CPU コア数を 6 コアに制限するとともに、演算周期を 6fps に制限する。

本実験では、バス停車時の発車判断を対象とし、本システムは車両停車中に演算を行う。本システムの出力結果は、発車直前の安全確認時に車内安全確認の結果として車両制御または画面モニタ、乗客向けアナウンスへ出力する。発車直前の安全確認時とは、乗客の乗車完了後にドアが閉まったタイミングから発車までの期間とする。アナウンス内容は検知層の検知結果を参照し、乗客の移動やふらつき、離席が検知された状況では、従来バス同様「着席または手すりをお掴まりください」とした。また、定員超過やかけこみ、転倒が検知された場合は運転手へ通知する。

後述の実証では、発車不可時は 3 回目まで乗客安全確認とアナウンスを再実行する。その後、発車不可が継続する場合は、運転手の現場判断に委ねる。なお、最終的な発車判断は、本システムによる乗客安全確認の結果と、運転手または自動運転システムによる車外の交通環境に対する安全確認が完了した場合に実行される。

## 4.3 演算時間の計測結果

実装したシステムの処理時間を表 2 に示す。CPU6 コア環境下において、本計測のために模擬的に全機能を演算させた状態での処理時間は平均 77msec であった。主な内訳では、知覚層におけるモデル推論処理が大部分を占め、特に骨格推定が最も時間を要した。バッチ処理を導入したことで、骨格推定のモデル推論を最大 24 名分までを 1 回の実行に集約し、本システムの実現に有効なことを確認した。その効果は、骨格推定の 1 回あたりの平均処理時間は 52msec に対し、乗客定員 15 名の場合

表 2 処理時間の計測 (N=1000)

Table 2 Measurement of computation time (N=1000)

		平均値
全体		77 msec
主な内訳	知覚層	
	人物検出 (車外)	12 msec
	人物検出 (車室内)	14 msec
	骨格推定 (車室内)	52 msec
	着座認識 (車室内)	9 msec
	転倒認識 (車室内)	8 msec
検知層+判定層		2 msec

には 14 回分の推論を削減でき、合計 728msec の処理時間短縮効果が得られた。なお、バッチ処理導入による 1 回あたりの処理時間増加は 1msec 未満であった。なお、各処理は並列実行されるため、表 2 は、全体の処理時間が内訳の合計値より少ない点に留意する必要がある。

本システムにおける発車判断時間は目標 4 秒に対し、転倒者が存在する場合は 3 秒、存在しない場合は時系列処理のレイテンシである 1.67 秒であり、その目標の達成を確認した。

## 5. 評価

### 5.1 精度評価

本システムの評価データセットは、バス車両へシステムを搭載し、学習データセットとは異なる場所で、評価のために乗客の動作を模擬して撮影した。判定層は単純な論理演算であり、最終的な判定結果は検知層の性能に依存するため、検知層の評価に焦点を当てた。駆け込みや車内移動など、判断基準ごとにシーンを分類して精度評価を実施した。各判断基準の正解ラベルは手作業で付与した。混合行列では「存在する」を Positive、「存在しない」を Negative として扱い、Precision と Recall、F1score を評価した。以下に、検知層の精度と本システムの限界について、6 つの評価基準ごとに述べる。

定員超過検知の精度は、人物検出の精度に依存するため、帽子やヘルメットの有無、服装 16 種、昼夜の組み合わせで 96 パターンの乗降行動を評価した。乗降口で人物を 100% 検知したため、検出漏れや過剰検出は発生しなかった。人物検出の挙動を検証すると、頭部検出矩形は、黒いヘルメットを着用した乗客に対して乗降口では 3/32 パターンで未検出であったが、その場合は全身検出矩形の結果で補完されたため、結果として 96 パターン想定通り検知され、実用水準であることを確認した。

駆け込み検知は車外の人物が駆け込む方向 5 種と速度 2 種、昼夜の組み合わせで評価した。駆け込み方向は駆け込みの開始位置をバス車外後方から前方にかけて 5 段階とし、速度は早歩きと駆け足を想定して、乗降口へ駆け込む動作とした。いずれのパターンでも想定通り 100% 検知された一方、バス停留所の利用状況に影響を受けることがわかった。駆け込み乗車検知エリアが通路と重なる場合に、乗客ではない通行人は誤検知の要因となった。対策として、バス停留所に依りて本処理を無効化する、またはバス車両の運用で停車位置を通路から離すなどの方法が考えられる。

車内移動検知は、車内の乗客が移動する方向と移動量、体勢の向きや隠れ具合の組み合わせで評価した。移動する方向は、車内 4 か所の位置で、移動行動と非移動行動を実施した。移動行動は往來の 2 パターン、非移動行動

はその場で足を開く、足元の物を拾うの2パターンとした。体勢は画像に対して上下左右4パターン、隠れ具合はカメラに対して手前に別の人物が立ち乗りした状態で混雑時の状況を模擬した。表3に1,084フレームの評価データセットにおける精度を示す。Precisionは0.992, Recallは0.859, F1スコアは0.920であり、良好な精度が得られた。ただし、以下のシーンでは精度低下が確認された。車室内カメラから1m以内の近くに立ち乗りの乗客Aが存在し、乗客Bがオクルージョンとなる場合である。乗客Bは判断基準である80cmの移動が乗客Aによって隠れやすく、5回中3回で未検知となった。

身体ふらつき検知は、乗客が車内の各位置で、物を拾う動作やよろける動作など手足や上半身を動かし、体勢の向きや隠れ具合で評価した。位置及び体勢、隠れ具合は車内移動検知と同様とした。表4に、1,156フレーム

表3 車内を移動する乗客が存在するシーンの検知精度 (N=1,084 フレーム) Precision 0.992, Recall 0.859, F1score 0.920

Table 3 Detection accuracy for scenes with passengers moving in the vehicle (N=1,084 frames) Precision 0.992, Recall 0.859, F1score 0.920

		正解 (当該乗客が~~)	
		存在する	存在しない
検知	存在する	11.25%	0.09%
結果	存在しない	1.85%	86.81%

表4 大きな身体動作をする乗客が存在するシーンの検知精度 (N=1,156 フレーム) Precision 0.975, Recall 0.990, F1score 0.983

Table 4 Detection accuracy for scenes with large moving passengers (N=1,156 frames) Precision 0.975, Recall 0.990, F1score 0.983

		正解 (当該乗客が~~)	
		存在する	存在しない
検知	存在する	21.88%	0.56%
結果	存在しない	0.21%	77.35%

表5 座席で座っていない乗客が存在するシーンの検知精度 (N=4,361 フレーム) Precision 0.854, Recall 0.849, F1score 0.852

Table 5 Detection accuracy for scenes with passengers not seated in seats (N=4,361 frames) Precision 0.854, Recall 0.849, F1score 0.852

		正解 (当該乗客が~~)	
		存在する	存在しない
検知	存在する	6.33%	1.12%
結果	存在しない	1.08%	77.35%

表6 転倒した乗客が存在するシーンの検知精度 (N=5,237 フレーム) Precision 0.977, Recall 0.589, F1score 0.735

Table 6 Detection accuracy for scenes with fallen passengers (N=5,237 frames) Precision 0.977, Recall 0.589, F1score 0.735

		正解 (当該乗客が~~)	
		存在する	存在しない
検知	存在する	16.42%	0.38%
結果	存在しない	11.42%	71.76%

の評価データセットに対する精度を示す。Precisionは0.975, Recallは0.990, F1スコアは0.983であり、良好な精度が得られた。ただし、カメラに対して身体が横向きに映る場合は、骨格点が重なって映るため、奥行き方向への身体の動きは検知が1~2フレーム遅れる傾向があった。

座席離席検知は、乗客が各座席で離席と着席を行い、体勢の向きや隠れ具合の組み合わせで評価した。位置及び体勢、隠れ具合は車内移動検知と同様とした。表5に、4,361フレームの評価データセットにおける精度を示す。Precisionは0.854, Recallは0.849, F1スコアは0.852であるが、後述の実証では実用水準であることを確認した。2つの後部座席に斜めに座り、その体勢がカメラに正対する場合、着座と立ち姿勢の区別が困難となり、精度の低下が確認された。

転倒状態検知は、乗客が車内の各位置で、四つん這いやしゃがみ込み、寝転びなどを行い、体勢の向きや隠れ具合の組み合わせで評価した。位置及び体勢、隠れ具合は車内移動検知と同様とした。表6に、5,237フレームの評価データセットにおける精度を示す。Precisionは0.977, Recallは0.589, F1スコアは0.735であり、一般的なシーンでは検知可能であった。Recallの低さは以下3つの特定シーンで未検知となったためである。第一は、認識エリア内に足のみが映る場合である。第二は、四つん這いで手や膝が隠れる場合である。第三は、黒い服を着てしゃがみ込んだ状態で倒れた場合である。これらは、身体の隠れやコントラストの低さにより骨格点が検出されなかったためである。

## 5.2 実証実験

本手法は運転手が同乗するバスにおいて、1か月間で570件の発車時に実証した。実証条件は8時から17時までの日中、成人の乗客、最大7名乗車で実施した。本実証は前述の小型バスを使用し、実用上の快適性を考慮して、立ち乗り乗客は一人あたり0.8m<sup>2</sup>の床面積を確保できる最大4名を想定する。これに着座乗客3名を加え、最大7名乗車で運行を実施した。乗客数は平均3名であり、夕方の通勤時間帯の1時間は7名乗車であった。実証の結果、表7に示すとおり、判定実行回数は全643回であり、「発車可」をPositive、「発車不可」をNegative

表7 実証実験における発車可否判定の結果 (N=643回) Precision 1.000, Recall 0.943, F1score 0.971

Table 7 Departure decision result in Demonstration experiments (N=643 execution count) Precision 1.000, Recall 0.943, F1score 0.971

		正解	
		発車可	発車不可
判定	発車可	566	0
結果	発車不可	34	43

として集計した。乗客の移動中など、本来発車不可の状況での「発車可」判定結果は0回、Precisionは1.000であり、安全に運行されたことを確認した。

一方、発車可の状況に対し「発車不可」と誤判定された回数は34回、Recallは0.943であった。その内訳は、着席や吊り革を握る際の腕の動きに対して身体ふらつき検知が敏感に検知されたことが22回、前節で述べた2つの座席に乗客が斜めに着座した状況に対して着座検知が敏感に検知されたことが12回であった。前者は一時的な動きであり「発車不可」が3回連続で判定されなかったため、発車判断はシステムで完結された。後者は3回連続で「発車不可」が継続したため、運転手へ現場判断を委ねて対処された。その件数は4件であった。

なお、この場合は運転手のアナウンスにより乗客自身が姿勢を是正され、最終的に、発車可能な状況で「発車不可」が継続し続ける事態は0件であり、運行の遅延は生じなかった。

以上の結果、実証条件下では、本システムが安全性を確保しながら、乗客安全確認570件中566件を自律化し、残り4件を運転手との協調で対処したことで、本システムが運転手の乗客安全確認件数の99.2%を代替可能なことが確認された。これにより、運転手の作業量は削減され、作業負荷軽減や運行の省力化が期待される。

## 6. む す び

本研究では、バス運行における発車判断の自律化という課題に対し、画像認識を用いた複数検知に基づく乗客安全確認システムを提案し、実験と実証を通してその有効性と限界を明らかにした。本システムは、RGB魚眼カメラの画像入力に基づいて、定員超過、駆け込み、車内移動、身体ふらつき、座席離席、転倒状態といった危険要因を画像認識で検知し、発車可否を判断する。実証実験の結果、日中かつ成人乗客、小型バス車両を対象とした条件下では、本システムは実用的な運行に十分対応可能であり、乗客安全の観点で発車判断を自律化し、運転手の負担軽減に貢献できる可能性が示された。一方で、個々の検知機能では、車室内カメラに立ち乗りの乗客が接近した状況では、その後方の乗客が映らない場合が生じ、単一カメラの検知では限界があることが確認された。この対策として、複数カメラの導入や、運用面でカメラ付近の立ち乗りを制限する方法が考えられる。また、カメラに正対する乗客の着座や四つん這いの検知に改善の余地があることも確認された。

今後の課題として、夜間や子供、より大きなバス車両を対象とした実証の実施に加え、さらなる乗客行動データの収集や解析を通じて発車判断の性能改善、及び複数

カメラや3D骨格推定などを用いた検知精度の向上が挙げられる。また、運転手や乗客へのインタラクションを通じた心理的な影響や社会的受容性に関する評価は重要な課題である。

## 参 考 文 献

- 1) 大谷真弘, 多田昌裕, 岡田昌也: “アイカメラを用いたバス運転手の公道上運転行動解析”, 交通工学論文集, Vol.4, No.1, pp.294-301 (2018).
- 2) 川端伸一朗, 加藤晋: “ラストマイル自動走行における車両内外安全に対応した乗降状態確認システムの研究開発”, 電子情報通信学会技術研究報告; 信学技報, Vol.118, No.343, pp.159 - -163, (2018).
- 3) C. Wang, J. Yan: “A Comprehensive Survey of RGB-Based, Skeleton-Based Human Action Recognition”, in IEEE Access, Vol. 11, pp.53880-53898 (2023).
- 4) I. Radosavovic, P. Dollár, R. Girshick, G. Gkioxari, K. He: “Data Distillation: Towards Omni-Supervised Learning”, IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp.4119-4128 (2018).
- 5) Camarena F, Gonzalez-Mendoza M, Chang L.: “Knowledge Distillation in Video-Based Human Action Recognition: An Intuitive Approach to Efficient and Flexible Model Training”, Journal of Imaging, 10, No.4, 85 (2024).
- 6) V. Tran, N. Balasubramanian, M. Hoai: “Progressive Knowledge Distillation For Early Action Recognition,” 2021 IEEE International Conference on Image Processing”, pp.2583-2587 (2021).
- 7) I. Lee, D. Kim, S. Kang, S. Lee: “Ensemble deep learning for skeleton-based action recognition using temporal sliding LSTM networks”, IEEE/CVF International Conference on Computer Vision, pp.1012-1020 (2017).
- 8) S. Yan, Y. Xiong, D. Lin: “Spatial temporal graph convolutional networks for skeleton-based action recognition”, In AAAI Conference on Artificial Intelligence, pp.7444-7452 (2018).
- 9) S. Ji, W. Xu, M. Yang, K. Yu: “3D Convolutional Neural Networks for Human Action Recognition”, in IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.35, No.1, pp.221-231, (2013).
- 10) R. Li, S. Li: “Human behavior recognition based on attention mechanism”, International Conference on Artificial Intelligence and Education, pp.103-107 (2020).
- 11) Cao Z., Hidalgo G., Simon T., Wei S., Sheikh Y.: “OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 43, 172-186 (2018).
- 12) K. Sun, B. Xiao, D. Liu, J. Wang: “Deep High-Resolution Representation Learning for Human Pose Estimation”, IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp.5686-5696 (2019).
- 13) Ge Zheng, Liu Songtao, Wang Feng, Li Zeming, Sun Jian: “YOLOX: Exceeding YOLO series in 2021”, arXiv:2107.08430 (2021).
- 14) Pronello Cristina, Garzon Ruiz, Ximena Rocio: “Evaluating the Performance of Video-Based Automated Passenger Counting Systems in Real-World Conditions: A Comparative Study”, Sensors, Vol.23, 18, 7719 (2023).
- 15) Ahmad Misbah, Ahmed Imran, Ullah Kaleem, Ahmad Maaz: “A Deep Neural Network Approach for Top View Peo-

ple Detection Counting”, IEEE 10th Annual Ubiquitous Computing Electronics & Mobile Communication Conference, pp. 1082–1088 (2019).

- 16) Alessandro Olivo, Giulio Maternini, Benedetto Barabino: “Empirical Study on the Accuracy and Precision of Automatic Passenger Counting in European Bus Services”, The Open Transportation Journal, Vol.13, No.1, pp.250–260 (2019).
- 17) D. Tsiktisiris, N. Dimitriou, A. Lalas, M. Dasygenis, K. Votis, and D. Tzovaras: “Real-Time Abnormal Event Detection for Enhanced Security in Autonomous Shuttles Mobility Infrastructures”, Sensors, Vol.20, No.17, p.4943 (2020).
- 18) H. Y. Lin, C. H. Tseng: “Abnormal Activity Detection and Classification of Bus Passengers With In-Vehicle Image Sensing”, in IEEE Access, Vol.12, pp.23057–23065 (2024).
- 19) C. H. Tseng, H. Y. Lin: “A Vision-Based System for Abnormal Behavior Detection and Recognition of Bus Passengers”, IEEE 25th International Conference on Intelligent Transportation Systems (ITSC), pp.2134–2139 (2022).
- 20) A. Mishra, S. Lee, D. Kim, S. Kim: “In-Cabin Monitoring System for Autonomous Vehicles”, Sensors, Vol.22, No.12, p.4360 (2022).
- 21) 株式会社フジクラ, AI を用いた路線バスの乗客安全監視システムの実証実験を開始 ～将来的な自動運転社会の実現に貢献～, <https://www.fujikura.co.jp/news/pressrelease/20211011ai.html> (2021).
- 22) 株式会社東海理化, バス乗車乗客の安全委寄与する画像 AI と行動デザイン技術, MEET UP CHUBE × 関西・共創の森 DAYS, [https://www.chubu.meti.go.jp/b35innovation/meetupchubu/event/pdf/slide52-1\\_5.pdf](https://www.chubu.meti.go.jp/b35innovation/meetupchubu/event/pdf/slide52-1_5.pdf) (2024).
- 23) SB ドライブ株式会社, DaiLY by dispatcher, [https://www.nishitetsu.co.jp/ja/news/news20190111102972/main/0/link/18\\_144.pdf](https://www.nishitetsu.co.jp/ja/news/news20190111102972/main/0/link/18_144.pdf) (2019).
- 24) 山根大河, 梅田靖, 木下裕介, 増田周弥, 濱田徳亜: “産業応用を目指した外観検査 CNN の軽量化”, 2022 年度精密工学会春季大会, pp.17–18 (2022).
- 25) Alex Bewley, Zongyuan Ge, Lionel Ott, Fabio Ramos, Ben Uprocft: “Simple online and realtime tracking”, 2016 IEEE International Conference on Image Processing (ICIP), 3464–3468 (2016).
- 26) Lin Tsung-Yi, Maire Michael, Belongie Serge, Hays James, Perona Pietro, Ramanan Deva, Dollár Piotr, Zitnick C. Lawrence: “Microsoft COCO: Common Objects in Context”, European Conference on Computer Vision, pp.740–755 (2014).
- 27) K. He, G. Gkioxari, P. Dollár, R. Girshick: “Mask R-CNN”, IEEE International Conference on Computer Vision, pp.2980–2988 (2017).
- 28) Cao Jinkun, Pang Jiangmiao, Weng Xinshuo, Khirodkar Rawal, Kitani Kris: “Observation-centric sort: Rethinking sort for robust multi-object tracking”, IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp.9686–9696 (2023).
- 29) Yufei Xu, Jing Zhang, Qiming Zhang, Dacheng Tao: “ViTPose: Simple Vision Transformer Baselines for Human Pose Estimation”, Advances in Neural Information Processing Systems, Vol.35, pp.38571–38584 (2022).
- 30) Ke Guolin, Meng Qi, Finley Thomas, Wang Taifeng, Chen Wei, Ma Weidong, Ye Qiwei, Liu Tie-Yan: “LightGBM: A Highly Efficient Gradient Boosting Decision Tree”, Ad-

vances in Neural Information Processing Systems, Vol.30, pp.3149–3157 (2017).

(2025 年 6 月 23 日 受付)

(2025 年 12 月 25 日 再受付)



### 山崎 将 幸

2006 年 岩手県立大学大学院ソフトウェア情報学研究科 博士前期課程修了。精密機器企業で認識技術や画質設計の研究開発、製品設計に従事。2017 年～現在 トヨタ自動車株式会社に勤務、認識技術及び車載システムに関する研究開発に従事。筑波大学情報理工学位プログラム博士後期課程に在学中。



### 滝 沢 穂 高 (正会員)

1998 年 大阪大学大学院工学研究科機械工学専攻 博士後期課程修了, 博士 (工学). 1998 年 豊橋技術科学大学知識情報工学系 助手. 2003 年～2004 年 シカゴ大学放射線科 客員研究員. 2005 年 筑波大学大学院システム情報学研究科コンピュータサイエンス専攻 講師. 現在 筑波大学システム情報系 教授, 学術情報メディアセンター長. IEEE, 電子情報通信学会, 情報処理学会, 画像電子学会, 日本生体医工学会, 日本医用画像工学会, 日本コンピュータ外科学会の会員.

## 可動カメラを用いた高速道路上の落下物検出

軒田佳林<sup>†</sup> 佐治 齊<sup>†</sup>(正会員)<sup>†</sup>静岡大学大学院総合科学技術研究科情報学専攻

## Fallen Object Detection on Expressway Using a Pan-Tilt Camera

Karin NOKITA<sup>†</sup>, Hitoshi SAJI<sup>†</sup> (Member)<sup>†</sup>Department of Informatics, Graduate School of Integrated Science and Technology, Shizuoka University

〈あらまし〉 高速道路における落下物処理件数は年間約 30 万件に達し、事故や渋滞を引き起こす可能性のある深刻な社会課題である。現在、高速道路の監視は可動式カメラを用いているが、落下物の発見は依然として目視に依存しており、早期発見が困難である。この課題に対して、本研究では交通監視カメラ映像を用いて自動で落下物を検出するシステムの実現を目的とする。従来の研究では背景差分や物体検出が用いられてきたが、実環境では影や道路外の背景により誤検出が生じやすい。そこで本研究では、画角変動を検出し、道路領域に限定した処理を行うことで、背景による誤検出を低減し、実用性の高い落下物検出手法を提案する。実験の結果、影や背景を含む環境下においても一定の精度で落下物を検出できることを確認した。今後は道路監視カメラ映像を用いて実証実験を行うことで、より実運用に近いシステムの構築を目指す。

キーワード：高速道路、パンチルトカメラ、道路領域、エッジ、落下物検出

<Summary> The number of fallen objects on expressways reaches about 300,000 cases per year, which leads to accidents and traffic congestion. This is a serious social problem. Currently, operators mainly monitor expressways with movable cameras, but finding fallen objects depends on humans looking at the video. This makes early detection difficult. Therefore, this study aims to realize an automatic fallen object detection system using traffic surveillance camera video. Previous studies used background subtraction and object detection methods. However, in real environments, these methods often make false detections because of shadows or background outside the road. In this research, we propose a method that detects camera motion and only processes the road and lane-marking areas. This reduces false detections due to noise and improves practical usability. The experiments showed that fallen objects could be detected with consistent accuracy, even in environments with shadows or background noise. In the future, we plan to test the system with real expressway videos to make it closer to real-world operation.

**Keywords:** expressway, pan-tilt camera, road area, edge, fallen object detection

## 1. はじめに

近年、高速道路上における落下物処理件数は年間約 30 万件<sup>1)</sup>に上り、その数は依然として高止まりの傾向を示している。高速道路では、車両の走行速度が速いため落下物を視認してから回避行動を取るまでの時間的余裕が極めて短く、落下物が直接的な交通事故の原因となる事例もある。そのため、高速道路上の落下物は交通安全上、極めて危険な存在である。また、事故処理や除去作業に伴う車線規制により、渋滞の原因ともなり得ることからその影響は大きい。こうした事故対応や交通管理の観点から、高速道路では道路上におよそ 1km

間隔でカメラを設置し、映像による監視を行っている<sup>2)</sup>。これらのカメラは水平・垂直方向に旋回が可能で、カメラ方向の自動制御も進められている。しかし、現在のところ落下物の発見は通行車両からの通報に頼っており、カメラ映像の確認は通報を受けてから行われるため、早期の発見は難しいという課題がある。このため、交通監視カメラ映像を用いた落下物の自動検出が望ましいと考える。

画像解析技術を活用した落下物検出に関する研究は広く行われている。天本ら<sup>3)</sup>は、背景差分と時間差分を用いて、その差分領域を移動物体、静止物体、照明変動の三種類に分類することで、車両や障害物を検出した。他にも深層学習ベー

スの研究として、Xuら<sup>4)</sup>は、路上放置物データセットを作成し、物体検出、セマンティックセグメンテーション、異常検出といった主要な手法の比較や背景と落下物を分離する手法の提案を行った。また、山本ら<sup>5)</sup>は異常検知モデルであるVAE(Variational Auto-Encoder)を用いて道路を学習させ、落下物がある領域を異常と判定することで落下物を検出した。しかし、これらの手法は画像全域を対象とするため、影や背景の影響による誤検出が起きやすい。セマンティックセグメンテーションによって道路領域を抽出することは可能だが、小さな落下物は道路自体との区別が付きにくい。また、事前に道路領域のマスク画像を用意することも有効だが、可動カメラの場合は画角変動ごとにマスク画像を生成する必要がある。

そこで、本提案は、これらの問題を解決しかつ実環境に適用し得る落下物検出手法の構築を目的とする。具体的には、画角変動を検出し道路領域を自動的に抽出することにより、誤検出の原因となる背景を効果的に除去し、さらに固定的な道路情報である道路領域や白線領域のエッジを除去することにより、種類や大きさが多様な落下物の検出精度を高める。本手法の適用により、高速道路上の落下物の早期検出を実現し、交通安全への寄与を図る。

## 2. 手法概要

本手法の概要図を図1に示す。本手法では、可動カメラで撮影された高速道路映像から落下物を検出するために、画角検知、背景生成、ノイズ除去、道路領域・白線除去を経て落下物を抽出し、矩形で可視化する一連の処理手法を提案する。まず、入力映像に対してオプティカルフロー解析<sup>6)</sup>を行い、フレーム間の画素移動量からカメラの画角変動を判定して連続した画角固定画像を数百フレーム抽出する。画角固定画像を基に背景画像を作成し、亀裂除去とエッジ検出によって背景画像と背景エッジ画像を得る。そして、道路領域と白線領域を求め、背景と白線を除去することで落下物のエッジを抽出する。最後に、落下物のエッジを外接矩形で囲むことで視認性を向上させる。本手法は、可動カメラ映像という背景が変化する環境下においても、落下物の検出が可能である。また、映像を一定フレームごとに区切って処理を繰り返すことで日照変化に対応し、実環境に適用可能なシステムを実現している。

## 3. 手法

### 3.1 画角変動検知

可動カメラで撮影された映像から連続したフレーム間での画素の移動量を解析し、画角変動やカメラ揺れを検知する。最終的に、画角が固定された200フレームを抽出し、背景画像の作成に用いる。

まず、入力画像に対してAKAZE特徴量<sup>7)</sup>を用いて特徴点を抽出する。このとき、画像を細かく分割した各領域から特

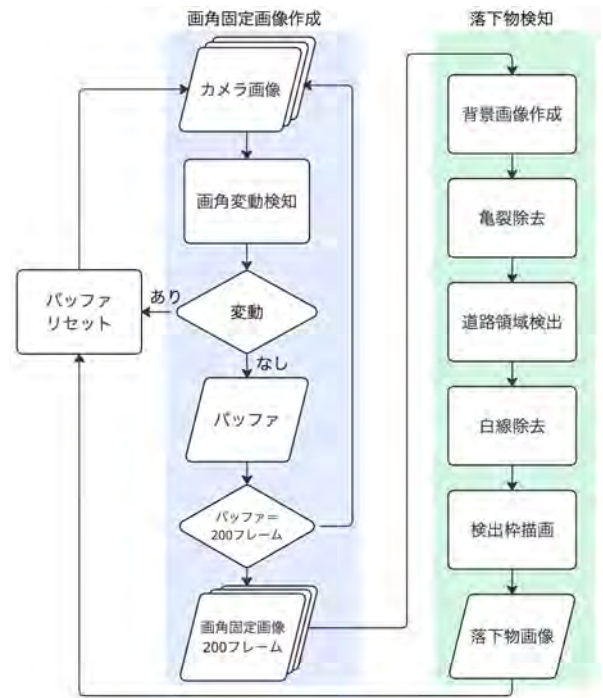


図1 手法概要図

Fig. 1 Overview of the proposed method

徴点を抽出することで、画面全体に均等に特徴点が分布するようにする。前フレームと現フレームで一致した特徴点をもとに、Lucas-Kanade法<sup>8)</sup>を用いてオプティカルフロー<sup>6)</sup>を計算し、各特徴点の移動距離を求める。移動距離が一定以上の特徴点の割合を求め、80%未満の場合はカメラ画角変動なしと判断して現フレームを一時的に保存する。この処理を繰り返して画角変動なしの画像を200枚取得できた時点で、画角固定画像200フレームとして出力する。一方で、移動距離が一定以上の割合が80%を超える場合は、カメラ画角変動ありと判断し、これまで一時的に保存していたフレーム群を破棄して次のフレームから再び画角変動検知を試みる。

### 3.2 背景画像作成

画角固定画像200フレームから5フレーム間隔でサンプリングを行う。これにより計算量を削減し、通過車両を除去する効果を高める。次に、サンプリング後のフレーム群の各画素における画素値の中央値を算出し、背景画像とする。フレーム群を時間軸方向に並べ、画素単位で中央値を取ること共通して存在する静止領域を抽出する。

### 3.3 亀裂除去

背景画像の路面に含まれる亀裂やノイズを抑制後、自動閾値決定法を用いたエッジ処理によってエッジ画像を作成する。まず、背景画像に対し、背景エッジ画像を作成する手法として、画像の勾配を算出し、閾値処理を行うことによりエッジを抽出するCannyのエッジ検出法<sup>9)</sup>を適用する。閾値は自動閾値決定法を用いており、画素値の中央値を0.66倍した値を下限閾値、1.33倍した値を上限閾値とすることで中央値に基づく自動的な閾値設定を行う。得られたエッジ画像をもとに、

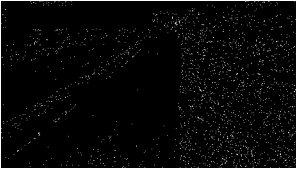


図2 亀裂除去前のエッジ画像  
Fig. 2 Edge image before crack removal



図3 亀裂除去後のエッジ画像  
Fig. 3 Edge image after crack removal

局所領域のエッジ分散を算出し、閾値を超える領域をノイズ領域として抽出する。次に、画像の垂直方向位置に応じて平滑化強度を変化させるための距離係数を導入した。画像の上端を0、下端を1とする正規化距離  $d(y)$  を式(1)によって求め、この距離に比例して平滑化強度  $s(y)$  を式(2)のように設定した。 $y$  は垂直方向位置、 $H$  は画像高さを表し、 $\alpha$  は平滑化強度の増加率を制御する係数として手動で設定する。ノイズ領域に対して、平滑化強度に応じてガウシアンフィルタ、バイラテラルフィルタ、オープニング処理を行うことで、カメラ近くは強く、遠方は弱く亀裂を除去する。亀裂除去前の背景画像にエッジ検出を行った結果を図2に、除去後の背景画像にエッジ検出を行った結果を図3に示す。亀裂除去前はノイズや路面の亀裂が多く、エッジ画像に多くの不要なエッジが含まれている。亀裂除去を行うことでノイズや亀裂が抑制され、主要な輪郭線を抽出することができる。

$$d(y) = \frac{y}{H} \quad (1)$$

$$s(y) = 1 + \alpha d(y) \quad (2)$$

### 3.4 道路領域検出

道路領域検出では、背景画像と背景エッジ画像から車道を抽出して道路マスク画像を作成する。まず、2値の背景エッジ画像に対して膨張処理を行うことで、線の切れ目を補完して道路境界の連続性を向上させる。膨張後の画像に反転処理を行うことで画像内の面部分を白、エッジ部分を黒にする。次に、背景画像をHSV色空間に変換して道路の灰色成分を抽出する。HSV色空間はHue(色相)、Saturation(彩度)、Value(明度)の3要素で構成されており、路上の灰色成分は彩度が低く、明度が中程度であるため、彩度が0~50、明度が100~200の範囲にある画素を灰色領域として抽出する。そして、エッジを反転した画像と灰色成分を抽出した画像との間で論理積を求めることで、エッジに囲まれ、かつ灰色成分を有する領域を抽出する。

続いて、得られたマスクに対してラベリング処理を実行し、各連結成分の面積を計算する。今回は、カメラから一番近い道路を検出対象とするため、最も面積の大きい領域を道路領域とする。そして、道路領域の輪郭を求め、内部を塗りつぶすことで穴のない2値の道路マスク画像を作成する。最後に、収縮処理を施すことで、境界部のノイズを除去する。

### 3.5 白線除去

白線検出では、白線の色と形状特徴を利用して白線を検出し、白線マスク画像を作成する。事前に白線エッジ画像と背景画像に対して道路マスク画像を適用し、背景を除去した2つの画像を白線除去の入力とする。まず、背景画像をHSV色空間に変換して白色領域を抽出する。白色は彩度が低く明度が高いため、彩度が0~30かつ明度が200~255の範囲を白色領域として抽出する。続いて、白線のエッジは全て同一角度であり、線分が長いという特徴を利用して白線のエッジを抽出する。まず、背景エッジ画像に対してハフ変換<sup>10)</sup>を適用して直線を検出する。ハフ変換では、直線を極座標形式である  $\rho = x \cos \theta + y \sin \theta$  で表す。 $\rho$  は原点から直線までの距離、 $\theta$  は直線の角度である。投票により直線を検出し、直線集合式(3)を得る。直線集合  $l_i$  について、角度  $\theta_i$  と長さ  $w_i$  を式(4)と式(5)により算出する。 $\theta_i$  は、 $0 < \theta_i < \pi$  で正規化される。直線の長さ  $w_i$  を重みとしたヒストグラムを式(6)と定義し、式(7)によって、最大化する角度  $\theta^*$  を白線の角度として求める。ここで、式(6)における関数  $\delta(\theta_i = \theta)$  は、 $\theta_i = \theta$  のとき1、それ以外るとき0となるインジケータ関数である。白線の角度と同一角度の直線を抽出し、白色領域と重なる領域を白線領域として白線マスク画像を作成する。最後に、背景を除去したエッジ画像と白線マスク画像の重なった画素を黒画素にすることで白線を除去する。

$$l_i = (x_{1i}, y_{1i}, x_{2i}, y_{2i}), \quad i = 1, \dots, N \quad (3)$$

$$\theta_i = \arctan 2(y_{2i} - y_{1i}, x_{2i} - x_{1i}) \pmod{\pi} \quad (4)$$

$$w_i = \sqrt{(x_{2i} - x_{1i})^2 + (y_{2i} - y_{1i})^2} \quad (5)$$

$$H(\theta) = \sum_{i=1}^N w_i \delta(\theta_i = \theta) \quad (6)$$

$$\theta^* = \arg \max_{\theta} H(\theta) \quad (7)$$

### 3.6 検出枠描画

背景と白線を除去した画像には落下物のエッジが残るため、エッジを矩形で囲うことで視認性を向上させる。まず小領域除去処理を行い、指定した面積未満の領域を削除する。次に、エッジの輪郭を検出し、外接矩形を描画する。そして、検出枠と背景画像を合成することで、落下物の位置を視覚的に明示する。

## 4. 評価

### 4.1 画角変動検知の実験

まず、画角変動検知について、学内で撮影した映像を用いて実験を行った。画角変動は、カメラの向きを手動で水平垂直方向に変動させ、数秒固定した後に背景画像が生成されるかを確認した。入力映像の詳細を表1に、入力映像を図4に

表 1 模擬実験の入力映像の詳細

Table 1 Details of the input video for the simulation experiment

場所	静岡大学 浜松キャンパス
解像度	1920 × 1080 (pixel)
フレームレート	30 (fps)
撮影時間	30 (sec)
カメラの高さ	3 (m)



図 4 入力映像

Fig. 4 Input video



図 5 背景画像

Fig. 5 Background image



(a) 黒色ブロック  
(a) Black block

(b) 灰色ブロック  
(b) Gray block

(c) 白色ブロック  
(c) White block



(d) 布  
(d) Cloth

(e) プラスチック  
(e) Plastic

(f) タイヤ片  
(f) Tire fragment

図 6 落下物

Fig. 6 Fallen objects

示す。処理の結果、画角が変動した場合は背景画像を作成せず、画角が固定された時点から背景画像を作成することができた。また、車両が走行している場合は、背景画像によって車両を除去することができた。作成した背景画像を図 5 に示す。

#### 4.2 落下物検出の実験

続いて、本手法の有効性を確認するため、落下物検出の実験を行った。入力映像には、学内で撮影した模擬実験のデータとパブリックになっている RAOD データセット<sup>4)</sup>の 2 種類のデータを用いた。まず、学内で撮影した模擬実験について述べる。入力映像の詳細は先の実験と同様で表 1 である。落下物は黒色、灰色、白色の 3 種類のブロックを 2 つずつ用意し、各色 20m と 10m の位置に配置した。落下物を図 6 に示す。実験は、各色のブロックに対して 2 回行い、1 回目は道路上、2 回目は白線上にブロックを置いた。精度評価については、落下物の領域を一部でも含む検出枠が存在する場合を検出成功とし、落下物以外の箇所に検出枠が存在する場合を誤検出とした。実験結果を表 2 に示し、結果画像の一例を図 7 に示す。

次に、RAOD データセットの実験について述べる。RAOD データセットは海外の高速道路で撮影された監視カメラ映像を用いた落下物検出のためのデータセットであり、様々な環境条件下での映像が収録されている。実験には、RAOD データセット内から落下物が車道内にある映像を抜粋し、昼の晴天時の映像を 8 本、昼の雨天時の映像を 1 本、夜の晴天時の映像を 1 本利用した。また、映像上部にはカメラ情報が表示されるため、上部を白く塗りつぶした。RAOD データセットの詳細を表 3 に示す。落下物は、タイヤ片、布、プラスチック、ビニール袋の 4 種類であり、落下物の例を図 6 に示す。入力映像に対して実験を行った結果を表 4 に示す。精度評価において、適合率は落下物と検出した中で実際に落下物の領域である割合、再現率は落下物の領域の中で、落下物と検出できた割合を示している。次に、昼の晴天時の結果画像の一例を図 8 に、道路領域画像を図 9 に、白線画像を図 10 に示

表 2 模擬実験の結果

Table 2 Results of the simulation experiment

落下物	検出結果
黒色ブロック	16/16
灰色ブロック	8/16
白色ブロック	8/16



図 7 模擬実験の結果画像

Fig. 7 Result image of the simulation experiment

表 3 RAOD の入力映像の詳細

Table 3 Details of the input video for RAOD

場所	海外の高速道路 <sup>4)</sup>
解像度	1920 × 1080 (pixel)
フレームレート	29.97 (fps)

表 4 RAOD の実験結果

Table 4 Results of RAOD

状況と落下物	検出結果	適合率	再現率
昼の晴 タイヤ片	2/2	0.67	0.83
昼の晴 布	2/2		
昼の晴 プラスチック	4/5		
昼の晴 ビニール袋	0/1		
昼の雨 布	1/1		
夜の晴 タイヤ片	1/1		

す。また、昼の雨天時の結果画像を図 11 に、夜の晴天時の結果画像を図 12 に示す。



図8 昼の晴天時の結果画像  
Fig. 8 Result image under clear daytime conditions



図9 道路領域画像  
Fig. 9 Road area image

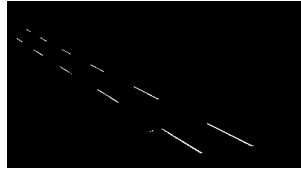


図10 白線画像  
Fig. 10 White line image



図11 雨天時の結果画像

Fig. 11 Result image under rainy conditions



図12 夜間の結果画像

Fig. 12 Result image under night conditions

#### 4.3 考察・システム制限について

実験の結果、画角変動の検知を行い、画角が固定された時点から背景画像を作成することが問題なく実現できた。また、模擬実験では黒色ブロックは16個全てを検出することができた。灰色ブロックと白色ブロックは、道路上では全てを検出できたが、白線上では未検出となった。また、本実験は学校敷地内で実施したためマンホールが誤検出された。実際の高速道路ではマンホールは存在しないことから、誤検出は低減すると考える。

RAOD データセットの実験では、様々な種類の落下物を検出できたことから、本手法の有効性が示されたと考える。また、道路領域、白線検出どちらも安定して検出できた。白線検出においては、一部落下物の白い領域が白線として誤検出されたが、落下物のエッジは残存していたため、落下物の検出には影響しなかった。しかし、白色に近い色の落下物は白線と誤検出されやすく、検出されにくいことが分かった。未検出においては、道路面が日照による影響で白色に近くなる場合、白線と誤検出されて検出範囲が狭くなるのが原因であった。また、画角によっては落下物がない道路を検出する場合があった。

本提案手法の領域の限定により、背景・影・白線の誤検出を低減でき、処理対象を削減することで計算コストの削減に直結する。また、多少の雨や明かりのある夜間においても落下物を検出できることが分かった。今後は、落下物の多様な色や形状への対応を進める必要がある。

続いて、適用可能な環境条件について以下にまとめる。検出物については静止物体を対象とし、検出範囲は白線内の車道である。画角端の4辺のいずれかに白線が交差し、道路面が四角で囲われるカメラ画角を前提とする。カメラからの距離60m以内の落下物は検出が可能であった。60mを超える領域では、路面の輝度変化による道路領域の未検出や白線の未検出が発生した。天候については、特に晴れた日中の環境において高い実用性が期待できる。雨や雪などの悪天候時には、路面の輝度低下やカメラに水滴が付着するといった環境の影響により検出精度が低下する場合がある。また、本手法は直線的な道路に適しており、角度の急な曲線の場合は白線検出が困難となる。今後は他の高速道路映像に対しても実験を行い、システム制限を緩和できるよう改善する。

## 5. むすび

本稿では、高速道路上の落下物検出の課題に対して、可動カメラ映像から画角変動を検知し、背景と白線を除去することで落下物を検出する手法を提案した。実験の結果、画角変動を検知し、背景画像作成や道路領域検出を自動で行うことができた。落下物検出に関しては、83%の再現率を達成した。この結果から、本手法は背景の影響を除去し、落下物検出を一定精度で実現できることが確認された。一方、落下物が白線の上にある場合や白色に近い場合に検出精度が低下するため、今後は、落下物の色や位置に依存しない手法を検討する。また、落下物の大きさを推定する手法についても検討を進めていきたい。

## 謝 辞

本研究の実施にあたり、多大なるご協力をいただいた中日本高速道路株式会社の皆様に深く感謝する。

## 参考文献

- 1) 国土交通省, 高速道路会社の落下物処理件数, [https://www.mlit.go.jp/road/sisaku/ijikanri/pdf/rakkabutu\\_nexco.pdf](https://www.mlit.go.jp/road/sisaku/ijikanri/pdf/rakkabutu_nexco.pdf) (2025).
- 2) NEXCO 東日本, 高速道路上の事故や落下物などの事象を交通監視カメラ映像から自動で検知する技術開発・実証を開始します, [https://www.e-nexco.co.jp/pressroom/head\\_office/2020/1028/00008642.html](https://www.e-nexco.co.jp/pressroom/head_office/2020/1028/00008642.html) (2025).
- 3) 天本直弘, 藤井明宏: “画像処理技術による障害物検出と移動物体追跡方法”, 電子情報通信学会論文誌 A, Vol.J81-A(4), pp.527-535 (1998).
- 4) Y. -J. Xu, H. Hu, X. -Y. Zhu, Y. -B. Nan, K. Wang, Z. -X. Liu, S. -G. Lian: “RAOD: A Benchmark for Road Abandoned Object Detection from Video Surveillance”, IEEE Access, Vol.12, pp.123985-123994 (2024).

- 5) 山本健生, 酒井俊: “フレーム差分画像を用いた VAE と近傍探索による落下物・不審物検知”, 人工知能学会全国大会論文集, Vol.JSAI2022, pp.1F4GS1002-1F4GS1002 (2022).
- 6) B. K. P. Horn, B. G. Schunck: “Determining Optical Flow”, Artificial Intelligence, Vol.17, No.1, pp.185-203 (1981).
- 7) P. F. Alcantarilla, J. Nuevo, A. Bartoli: “Fast Explicit Diffusion for Accelerated Features in Nonlinear Scale Spaces”, Proceedings of the British Machine Vision Conference (2013).
- 8) B. D. Lucas, T. Kanade: “An Iterative Image Registration Technique with an Application to Stereo Vision”, Proceedings of the Imaging Understanding Workshop, pp.121-130 (1981).
- 9) J. Canny: “A Computational Approach to Edge Detection”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.PAMI-8, No.6, pp.679-698 (1986).
- 10) N. Kiryati, Y. Eldar, A. M. Bruckstein: “A Probabilistic Hough Transform”, Pattern Recognition, Vol.24, No.4, pp.303-316 (1991).

(2025年11月13日受付)



軒田 佳林

2024年 静岡大学情報学部情報科学科卒業。現在 静岡大学大学院総合科学技術研究科修士課程在籍。情報科学と画像処理を専門とし、高所監視カメラを用いた交通流解析や、交通監視カメラを用いた落下物検知に関する研究に取り組んでいる。



佐治 斉 (正会員)

東京大学工学部機械工学科卒業，東京大学大学院理科系研究科情報科学専攻博士課程修了。博士(理学)。現在静岡大学情報学部教授。画像処理技術の社会応用を専門とし，航空・衛星画像を用いた，被災地解析や道路交通情報解析の研究に取り組んでいる。

## 神奈川工科大学 鈴木研究室

鈴木 浩 (正会員)

神奈川工科大学

Suzuki Lab., Kanagawa Institute of Technology

Hiroshi SUZUKI (Member)

Kanagawa Institute of Technology

## 1. はじめに

神奈川工科大学は、1963年に幾徳工業高等専門学校として開学し、その後の発展を経て現在に至る大学である。現在は、工学部、情報学部、健康医療科学部の3学部を擁している。当研究室は、2016年に開設された研究室であり、神奈川工科大学 情報学部 情報メディア学科に所属し、厚木キャンパス（神奈川県厚木市）に位置している。2025年度現在で、学部4年生11名、修士課程2名から構成されている。本稿では、当研究室の研究内容を簡単に紹介する。

## 2. 研究内容

鈴木研究室では、人とデジタル技術の関係性に着目し、年齢や経験、身体能力の違いを超えて誰もが参加できるインタラクティブシステムの研究・開発を行っている。研究の中心にあるのは、技術そのものの新規性ではなく、「人がどう関わり、場がどう変化するか」という視点である。

本研究室では、子どもから高齢者までを対象とし、身体的・認知的・社会的な関与を自然に引き出す体験設計を重視している。紙に描く、手で触れる、実物を転がすといったアナログ的な行為と、デジタル処理やネットワーク技術を組み合わせることで、直感的で参加しやすいインタラクションを実現してきた。

研究成果は、論文発表にとどまらず、科学館展示、地域イベント、高齢者施設への導入など、実社会の現場で検証されることを前提としている点の特徴である。現場での運用を通じて得られる課題や気づきを次の設計に反映し、改良を重ねる循環型の研究スタイルをとっている。

近年は、高齢者向けユニバーサルeスポーツと、子どもの創造力を育むインタラクティブシステムという二つの軸を中心に、人と人をつなぎ、世代を越えた関わりを生み出すメディア技術の可能性を探究している。本稿では、当研究室における最近の研究や取り組みについていくつか紹介する。

2.1 ユニバーサルeスポーツについて<sup>1)</sup>

高齢者向けユニバーサルeスポーツの研究として、鈴木研



図1 サイコロ型UIを利用したサンコロビンゴ<sup>1)</sup>

Fig. 1 Sancoro Bingo using a Dice-based User Interface<sup>1)</sup>

究室ではサイコロ型インタラクティブゲーム「サンコロビンゴ<sup>1)</sup>」を中心に開発と実践を行っている。サンコロビンゴは、3つのサイコロ型コントローラを利用してビンゴカードのマスの数字が答えとなる数式をつくることでビンゴカードを埋めていくシンプルなルールを持ちながら、計算、判断、身体動作、他者との会話といった複数の要素を自然に含んでいる点の特徴である（図1参照）。

本研究室では、勝敗や競技性のみを目的とするのではなく、「参加し続けられること」「場が和らぐこと」「観戦や助言といった周辺行為も含めて楽しめること」を重視している。そのため、操作の難易度を極力下げ、初めての参加者でも短時間で理解できる設計を採用している。

近年は、高齢者施設への継続的な導入や常設展示を通じて、単発イベントでは見えにくい継続性や習慣化の効果について検討を進めている。また、遠隔地をネットワークで接続した対戦や観戦の実証も行い、地域や施設を越えた交流の可能性を探っている。

このようにサンコロビンゴは、eスポーツを「競技」ではなく「社会参加のきっかけ」として捉え直す試みであり、高齢者の未病改善や孤立防止に資する新しいeスポーツ像の提示を目指している。

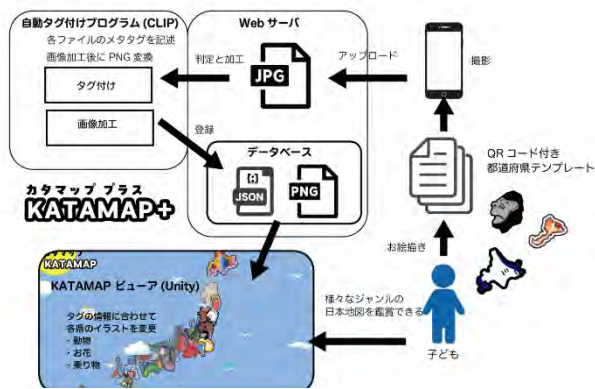


図2 KATAMAPシステムの概要図<sup>2)</sup>  
 Fig. 2 Overview of the KATAMAP System<sup>2)</sup>

## 2.2 子ども向けインタラクティブシステムの活用<sup>2)</sup>

子ども向けの研究として、鈴木研究室では、子どもの創造的な行為を起点としたインタラクティブシステムの開発と活用を行っている。重視しているのは、完成されたデジタルコンテンツを一方的に提示するのではなく、子ども自身の「描く、つくる、動かす」といった行為を出発点として、それがデジタル技術によって広がる体験を生み出すことである。こうした仕組みにより、子どもが主体的に参加し、自分の表現や行動が周囲に影響を与える感覚を得られる場づくりを目指している。その代表的な取り組みの一つが「KATAMAP」である。KATAMAP<sup>2)</sup>は、子どもが描いた都道府県の形の絵をデジタル化し、日本地図上の対応する位置に配置・表示するシステムである(図2参照)。自分の作品が地図の一部として他者の作品とともに共有されることで、単なるお絵描きにとどまらない体験が生まれる。また、他の参加者の表現にも自然に触れることができるため、地理への関心に加え、多様な表現への気づきを促すことも目的としている。個人の創作と共有の場とを結びつけている点が、本システムの特徴である。

また、子どもの創作と組み合わせたプロジェクト「ペーパーヨーヨーシューティング<sup>3)</sup>」は、紙で制作したヨーヨーを用いた身体的な遊びと、デジタル演出を組み合わせたインタラクティブ作品である(図3参照)。子どもたちはまず自分のヨーヨーを制作し、その後「投げる、当てる、動かす」といった行為を通じてシステムと関わる。このように、工作と遊び、さらにデジタル体験が一体となっている点に特徴がある。自分でつくったものを用い、その動きに応じて映像や反応が変化することで、自分の行為がシステムに影響を与えていることを直感的に理解できる構成となっている。

## 3. おわりに

これらのシステムは、科学館や地域イベントにおける展示、体験コーナー、ワークショップなどを通じて活用されており、子どもから大人まで幅広い参加者が関わることのできる実践



図3 PPYYシステムの実施風景<sup>3)</sup>  
 Fig. 3 Scene from the Paper Yo-Yo Shooting System<sup>3)</sup>

例となっている。特に、専門的な知識や高度な操作を必要とせず、まずは、描く、つくる、触る、動かすといった身近な行為から参加できるため、年齢や経験に関わらず体験しやすい点が高く評価されている。また、展示やワークショップの場では、完成した作品を一方的に鑑賞するだけではなく、参加者自身がシステムの一部となりながら体験をつくり出していくことができる。鈴木研究室では、このような実践を通じて、作品の完成度の高さを見せることより、参加者が体験の中で「つくることの楽しさ」や「自分の表現が他者や社会とつながっていく感覚」を得られることを重視している。すなわち、創作と体験の循環を支えるシステムを設計することによって、表現の喜びや参加の意義を実感できる場を生み出すことを目指している。

## 参考文献

- 1) Sakura Mizobuchi, Hiroshi Suzuki, Design and Evaluation of "Sancoro Bingo": an Interactive Game Encouraging Continued Participation Among Older Adults, Games and Learning Alliance. GALA 2025. Lecture Notes in Computer Science, vol 16307. Springer, pp.182-191, January 2026.
- 2) 古金ひな, 鈴木浩: "触れて探す地理学習のためのインタラクティブ地図ビューアー KATAMAP+", 情報処理学会 情報処理学会インタラクティブ2026, 2B27, pp. 642-643 (2026).
- 3) 嶋野辰己, 鈴木浩: "ペーパーヨーヨーを利用したコンテンツにおける複数人検出システムの検討", 研究報告デジタルコンテンツクリエイション (DCC), 2025-DCC-40, 号 9, pp.1-3 (2025).



鈴木 浩 (正会員)

情報科学芸術大学院大学(IAMAS)出身。博士(工学)。先端技術を活用したワークショップやエデュテインメントシステムの開発・制作を行い、全国の科学館や教育関連イベントで展示・発表活動を展開している。近年は、年齢や経験を問わず参加できるユニバーサル e スポーツ「サンコロビンゴ」を開発し、高齢者施設や e スポーツイベントを中心に普及活動を行っている。現在 神奈川工科大学 情報学部 情報メディア学科教授。

## 理事会だより

### 第 354 回 理事会

1. 日時：2026 年 4 月 14 日(火)17:00-18:00
2. 場所：事務局を主たる場所としてオンライン（ZOOM）併催
3. 出席者：理事 23 名

総出席者：26 名

【会長】高村

【編集】今給黎、新田

【財務】阿部、西村、中村

【技術専門】丸山、草野、早瀬

【監事】竹島、金盛

<オブザーバ> 小野元会長

<事務局> 大嶽

【副会長】内田、藤澤、佐藤、山本

【企画】石樽、馬場、河合、澤口、菅野

【総務】鈴木

【地方】藤井、甲斐

<次期会長> 土橋

#### 4. 議題・資料

	担当	付番
前回議事録確認	<内田副会長>	354-000-01
<b>議 題</b>		
(1) 議題 1：第 54 回年次大会進捗状況	<藤澤実行委員長>	354-001-01
(2) 議題 2：AIS2026 進捗状況	<草野技術専門理事>	354-003-01
(3) 議題 3：IEVC2026 速報	<内田副会長>	354-004-01
(4) 議題 4：第 11 回西田賞の選定について	<土橋選定委員長>	354-005-01
(5) 議題 5：第 55 回総会資料について	<事務局>	354-006-01
<b>報 告</b>		
(6) 報告 1：編集委員会報告	<内田副会長>	354-007-01
(7) 報告 2：企画委員会報告	<石樽企画委員長>	354-008-01
(8) 報告 3：フェロー、ベイン賞 推薦状況	<事務局>	354-009-01
(8) 報告 4：2026 年度理事会日程について	<事務局>	354-010-01

5. 議事 司会：内田副会長、議事録：鈴木総務理事

2026年度画像電子学会研究会等予定

研究会等名	開催日	場所	テーマ	締切	記事
Advanced Image Seminar 2026	6月5日(金)	オンライン開催	画像×AIの産業応用の最新動向	申込締切:6/3(水)	
第317回研究会	6月22日(月)-23日(火)	早稲田大学 西早稲田キャンパス 55号館 N棟第二会議室	画像一般	申込締切:4/30(木) 原稿締切:6/1(月)	共催
第54回年次大会Media Computing Conference 2025	8月24日(月)-26日(水)	筑波大学		<b>■ Conference Track</b> 発表申込(要旨提出)期限:6/15(月) 原稿提出期限:7/22(水) <b>■ Journal Track</b> 学会誌用原稿提出期限:5/25(月) 発表申込(要旨提出)期限:6/15(月) 原稿提出期限:7/22(水)	
Visual Computing2026	9月8日(火)-10日(木)	北海道大学		概要申込締切:5/22(金) 原稿締切:5/29(金)	
画像関連学会連合会第13回秋季大会	10月26日(月)-27日(火)(予定)	名古屋大学 野依記念学術交流館(予定)			共催
第318回研究会-高臨場感ディスプレイフォーラム 2025	11月(予定)				共催
ビジュアルコンピューティングワークショップ 2026	11-12月(予定)				
第12回DMH研究会	2月(予定)				
第319回研究会	2月(予定)				
第55回VMA研究会/第21回 視覚・聴覚支援システム研究会	3月(予定)				
第13回DMH研究会	3月(予定)				
第320回研究会-映像表現/芸術科学フォーラム 2027	3月(予定)				共催
ドローン研究会	3月(予定)				

\*研究会の場合、いずれも「画像一般」を含む

\*空欄部は未定

\*上記の予定は変更になる場合があります

問い合わせ先: 〒116-0002 東京都荒川区荒川3-35-4-101

一般社団法人 画像電子学会

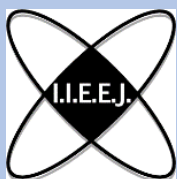
TEL:03-5615-2893 FAX:03-5615-2894

E-mail:kikaku@iieej.org

<https://www.iieej.org/>

<https://www.facebook.com/IIEEJ>

# First Call for Papers



## 2026 年度 第 54 回 画像電子学会 年次大会 Media Computing Conference 2026

筑波大学つくばキャンパス  
(茨城県つくば市)



2026 年度 画像電子学会 第 54 回年次大会(Media Computing Conference 2026)を, 2026 年 8 月 24 日～8 月 26 日に筑波大学にて開催いたします。一般セッション, 学生セッション, 企画セッションのほか, 特別企画(基調講演, 招待講演等)も計画中です。会員の皆様, 奮ってご参加ください。

### 1. 開催日程・場所

開催日程: 2026 年 8 月 24 日(月)～8 月 26 日(水)

場所: 筑波大学つくばキャンパス春日エリア (〒305-8550 茨城県つくば市春日 1-2)

### 2. 講演・参加申込み方法、参加費

Conference Track と Journal Track の二形態で募集します。

#### ◇ Conference Track

従来の年次大会の講演募集と同様で, 「一般セッション」「学生セッション」「企画セッション」で募集します。

- 一般セッション: 画像電子学会の会員が一堂に会し, 画像電子工学全般に関する研究の発表と意見の交換を行うことを目的とします。学生会員の発表も歓迎します。
- 学生セッション: 画像電子工学に関する研究を行っている学生が一堂に会し, 研究の発表と意見の交換を行うことを目的とします。課題提案を含め, 卒論, 修論などで検討された内容も歓迎します。様々な知見を持った専門家との議論によって研究を発展させる良い機会ととらえて積極的な応募をお願いします。なお, 博士課程における研究成果はなるべく一般セッションでの発表をお願いします。
- 企画セッション: 最近のメディアコンピューティング技術のトピックスを本学会研究委員会等が企画し, 発表を通して, 意見交換を行うことを目的としています。企画セッションについては現在検討中です。

講演分野は 5. 講演分野をご覧ください。

#### ◇ Journal Track(査読付き発表)

◇ 画像電子学会誌への投稿論文を併せて提出してもらう形態です。対象セッションは Conference Track と同様「一般セッション」「学生セッション」「企画セッション」とします(各セッションの概要は Conference Track 欄をご覧ください)。

- 掲載号: 2027 年 1 月号(「年次大会ジャーナルトラック特集号」)
- 投稿種別: 通常の画像電子学会誌の投稿種別と同様, 以下の 5 種類とします。  
「論文」「ショートペーパー」「システム開発論文」「資料論文」「実践論文」

原則として採否は年次大会までに決定し, ご連絡いたします。採録が決定した論文は, 査読付発表として取り扱います。

### 3. 論文提出方法・提出期限

#### ◇ 論文提出方法

- Conference Track

EasyChair を利用して要旨, 及び大会予稿集掲載用原稿を提出いただきます。  
5 月下旬に申し込み受付を開始する予定です。

- Journal Track

学会誌への投稿論文は学会誌の投稿システム(Editorial Manager)を利用して提出いただきます。  
なお, 要旨, 及び大会予稿集用原稿を Conference Track と同様, 別途提出いただきます。

#### ◇ 論文提出期限

- Conference Track
  - ・ 発表申込(要旨提出)期限:2026年6月15日(月)
  - ・ 大会予稿集用原稿提出期限:2026年7月22日(水)
- Journal Track
  - ・ 学会誌用原稿提出期限:2026年5月25日(月)
  - ・ 発表申込(要旨提出)期限:2026年6月15日(月)
  - ・ 大会予稿集用原稿提出期限:2026年7月22日(水)

#### 4. 参加申し込み方法・参加費

後日、年次大会 Web サイトに掲載します。参加申し込み受付は5月下旬開始予定です。

#### 3. 講演分野

- [1] 画像符号化(静止画像符号化, 動画画像符号化, 可逆符号化, 変換符号化, 動き補償, エントロピー符号化, 符号化ハードウェア, 符号化ソフトウェア, 国際標準化 等)
- [2] セキュリティ(ステガノグラフィ, データハイディング, 著作権保護, 不正コピー防止, 改ざん防止, フィンガープリンティング, 攻撃技術, 透かし解析技術, ブロックチェーン, コバートチャネル, サプリミナルチャネル, 画像暗号化, 画像認証, バイオメトリクス, イメージフォレンジックス 等)
- [3] 通信・ネットワーク(画像通信, 光ネットワーク, 符号化制御技術, 画質制御技術, 画像検索技術, 通信プロトコル, マルチメディア通信, クラウド, 標準化方式 等)
- [4] 画像処理・認識(画像復元, 画像強調, フィルタ技術, 特徴抽出・追跡, オブジェクト抽出・検出, 動き解析, 領域分割, 画像認識, 文書画像, 顔画像解析, 動画画像解析, 文字認識, 高解像化処理, コンテンツ制作 等)
- [5] 画像入出力・色・画質評価(カラー画像処理, ディスプレイ技術, プリンタ技術, 色評価, 心理評価 等)
- [6] バーチャルリアリティ・CG・コンピュータビジョン(ステレオ・多視点画像解析, 3次元画像技術, 仮想現実感・拡張現実感, 映像生成, コンピュータグラフィックス, モデリング, レンダリング, ビジュアライゼーション, アニメーション, 画像再構成・復元, 照明・反射解析, マルチメディア処理, メディア統合, ホログラフィ, ポイントクラウド, ライトフィールド 等)
- [7] 画像基礎・画像アプリケーション, その他画像一般(人工知能, ニューラルネットワーク, リモートセンシング, 機械学習, 深層学習, 進化計算, 遺伝的アルゴリズム, 画像データベース, 教育支援システム, Web 関連技術, ヒューマンインタフェース, ユーザビリティ, アクセシビリティ, 医用画像, IoT, ビックデータ, クラウドコンピューティング, デジタルサイネージ, 標準化教育, 災害応用, スマートシティ, ドローン, 産業応用, テレワーク, 遠隔教育, ロボット工学, 医療・農業・建設業などとの境界領域 等)

#### 4. 問合せ先

画像電子学会事務局

〒116-0002 東京都荒川区荒川三丁目 35 番 4 ライオンズマンション三河島第二 101 号

TEL: 03-5615-2893

FAX: 03-5615-2894

E-mail: kikaku@iieej.org

**画像電子学会 第 317 回研究会 in 東京 発表募集 (6/22~6/23 開催)**  
**共催：映像情報メディア学会 (映像表現&コンピュータグラフィックス研究会)**

下記により、画像電子学会 第 317 回研究会 (共催：映像情報メディア学会 (映像表現&コンピュータグラフィックス研究会)) の発表論文を募集いたします。

ご関係の各位には奮ってご応募ください。今回は現地とオンラインによるハイブリッド形式で開催します。

画像電子学会の第一種研究会では「コメント・サービス」に加え、優れた研究発表を行った若手研究者に対して「研究会若手奨励賞」を授与する予定です。

—記—

■テーマ：

画像一般 (色再現、立体映像、CG、VR、アニメーション技術、画像処理、ME 関係、画像符号化など幅広く扱います。年次大会投稿に向けた議論の場としても歓迎します。)

■開催日：

2026 年 6 月 22 日 (月) ~6 月 23 日 (火) (予定)

(発表件数により期間を短縮する可能性があります)

■会場・発表形式：

現地とオンラインのハイブリッド開催、質疑応答含み 1 件当たり 30 分前後 (実施方法や発表時間は、今後の状況により変更になる場合があります。)

□現地での発表・聴講形式：

早稲田大学 西早稲田キャンパス 55 号館 N 棟第二会議室 (〒169-8555 東京都新宿区大久保 3-4-1)

<https://www.waseda.jp/top/access/nishiwaseda-campus>

発表者は各自の PC にてビデオ会議システム Zoom の画面共有機能により画面を共有しつつ発表を行っていただきます。発表前日までに発表者に Zoom への接続情報を電子メール等にてお送りします。

(接続トラブルなどを想定し、念のため、USB メモリ等にて PowerPoint ファイル等を別途ご持参下さい。)

□オンラインでの発表・聴講形式：

発表者は各自の PC にてビデオ会議システム Zoom の画面共有機能により画面を共有しつつ発表と質疑応答を行っていただきます。各自マイク、カメラ、スピーカーを備えた PC をご用意ください。発表前日までに発表者と参加者に Zoom への接続情報を電子メール等にてお送りします。

■申込締切日：2026 年 4 月 30 日 (木) → 5 月 7 日 (木) に延長

下記、共催映像情報メディア学会の URL から直接お申込みください。(ページ内の案内に沿ってお申

込み下さい)

<https://www.ite.or.jp/ken/program/index.php?tgid=IEEEJ>

■原稿締切日：2026年6月1日(月)

原稿の目安：A4サイズ 2段組 2～4頁

■参加費(予稿集付)：2,000円(2日開催の場合)、1日開催の場合は1,000円

- ・学生で予稿集(ダウンロード)が不要の場合は参加費無料です。
- ・参加申込、および、参加費の支払い方法は別途ご案内いたします。

■コメント・サービスについて：

研究会(本学会主催/共催の第一種研究会)では、発表内容がその後学会誌論文としての投稿に発展することを促進するため、発表者の希望に基づいて、当学会が指名した研究者が事前に発表者の最終原稿を読んだ上で、当日発表時に質問やコメントをし、さらに後日、コメントを発表者に送付する「コメント・サービス」を実施しています。

このサービスを活用すると、発表に関してより踏み込んだ議論が期待できますので、今回の研究会での発表内容を将来的に学会誌論文として投稿することを検討されている発表者は、申込締切後に画像電子学会より送付されるアンケートにて「コメント・サービス希望」とご回答ください。なお、本サービスの対象発表数には限りがありますので、最終原稿に基づき選抜させていただくことがあります。あらかじめご了承ください。

■研究会若手奨励賞について：

画像電子学会第一種研究会において優れた研究発表を行った若手研究者に対して「研究会若手奨励賞」を毎年若干名授与します。受賞資格は以下のとおりです。研究会奨励賞の対象者(研究会における発表登壇時において35歳以下)は発表申込み時に生年月日をご記入ください。

- 本奨励賞の受賞時において本学会会員(学生会員含む)であること。
- 研究会における発表登壇時において35歳以下であること。
- 過去に同賞を受賞したことがないこと。

■お問い合わせ先：一般社団法人 画像電子学会

〒116-0002 東京都荒川区荒川3-35-4 ライオンズマンション三河島第二 101号室

TEL: 03-5615-2893 FAX: 03-5615-2894 E-mail: [kikaku@iieej.org](mailto:kikaku@iieej.org)

\*詳細については学会ホームページで順次掲載いたします。

画像電子学会ホームページ <https://www.iieej.org/>

■共催：

映像情報メディア学会(映像表現&コンピュータグラフィックス研究会)

■その他：

当学会では研究会予稿を電子化しており、予稿(電子ファイル)の事前配布を可能とするため、予稿の公開日は研究会開催日の1週間前となる予定です。 以上

画像電子学会セミナー  
Advanced Image Seminar 2026 (AIS2026)  
「画像×AIの産業応用の最新動向」

画像認識技術と各種 AI の急速な進展は目覚ましく、産業分野における革新的な応用が実現されています。本セミナーでは、「画像×AI」をキーワードに、製造業、物流、医療、農業、スマートシティなど多様な産業領域における最新の技術動向と実用事例を、第一線で活躍されている講師をお招きし解説していただきます。

主催：画像電子学会

協賛（予定）：映像情報メディア学会、情報処理学会、画像関連学会連合会（日本画像学会、日本写真学会、日本印刷学会）

後援（予定）：電子情報通信学会、公益財団法人 画像情報教育振興協会

開催日：2026年6月5日（金）13:00-16:30

場所：オンライン開催

参加費：会員（賛助会員／協賛学会／後援協会会員を含む）12,000円 / 会員（参加）3,000円 / 学生3,000円 / 非会員18,000円

●振込手数料はご負担願います。

●参加者には 講演後に、講演資料をダウンロード可能とする予定です。

※お申し込み頂いた非会員の方で、入会希望の方は、入会金・初年度会費 が免除となります。

振込先：みずほ銀行 浜松町支店

口座番号：（普）1961408 口座名：一般社団法人画像電子学会

申し込み先：<https://forms.gle/SWCuCN6YRnzLb1DS9> よりお申し込みください。

申込締切：2026年6月3日（水）

全般お問合せ先：画像電子学会 事務局

〒116-0002 東京都荒川区荒川三丁目 35 番 4 ライオンズマンション三河島第二 101 号

TEL: 03-5615-2893, FAX 03-5615-2894, e-mail: kikaku@iieej.org

【プログラム】

(1) 13:00-13:50 NTT 株式会社 人間情報研究所 研究主任 折橋 翔太 氏

タイトル：次世代メディア処理 AI 「MediaGnosis」における画像処理技術の研究と産業応用

(概要) 「MediaGnosis」は、NTT が開発する音声・音響処理、画像処理、自然言語処理、クロスモーダル処理を統合的に扱う次世代メディア処理 AI です。マルチメディア情報をオールインワンで処理することで、効率的な学習と高精度かつ総合的な推論を実現します。本講演では、画像処理技術を中心に、最新の研究動向から産業応用事例までご紹介します。

(2) 13:50-14:40 神奈川工科大学 情報学部 情報システム学科 教授 三枝 亮 氏

タイトル：共生社会を実現するフィジカル AI の研究開発

(概要) 本講演では介護・医療・福祉分野の現場課題を解決する方法として、人と共生するフィジカル AI の事例を紹介します。人支援技術において人と装置の相互作用は本質であり、物理空間で装置を制御するフィジカル AI は、視覚を初めとする多様なモダリティの情報処理に支えられています。人と共生するフィジカル AI を題材として、人支援技術における画像認識や画像生成のニーズと重要性を共に議論できれば幸いです。

14:40-14:50 休憩

(3) 14:50-15:40 明星大学 情報学部 情報学科 准教授 植木一也 氏

タイトル：視覚・言語モデル (VLM) を用いた大規模映像検索

(概要) 近年の視覚・言語モデル (VLM) の発展により、映像検索は単純なキーワードを用いた検索から、マルチモーダル推論に基づく適応的検索へと進化している。本講演では、過去 10 年間にわたる映像検索ベンチマークでの取り組みを踏まえ、検索精度向上のための施策について紹介する。具体的には、推論時の検索クエリの拡張、キャプション生成によるモデルのドメイン適応、VQA を用いた意味的検証による再ランキングという 3 つの手法を紹介し、文脈理解に基づく映像検索の新たな可能性と今後の展望を議論する。

(4) 15:40-16:30 法政大学 情報科学部 デジタルメディア学科 准教授 佐藤周平 氏

タイトル：CG による映像制作支援のための AI 技術の利活用

(概要) コンピュータグラフィックス (CG) による映像制作は、一般に多くの時間と労力を要する。このような課題に対し、近年急速に発展している AI 技術を活用した研究が数多く行われている。本講演では、AI 技術を利用した流体の流れのデザイン、テキストからの樹木のモデリング、山岳地形画像に対する再照明およびセグメンテーションなど、講演者が現在取り組んでいる CG による映像制作を支援するための研究について紹介する。

## *Call for Papers*

### **Special Issue on “Extended Papers Presented in IEVC2026”**

IIEEJ Editorial Committee

The 9th IIEEJ International Conference on Image Electronics and Visual Computing (IEVC2026) was successfully held in Hiroshima, Japan, from March 16 to 19, 2026. The conference brought together researchers, engineers, developers, and students from academia and industry to present and discuss the latest advances in image electronics and visual computing across a wide range of fields.

Following the success of IEVC2026, the IIEEJ Transactions on Image Electronics and Visual Computing is pleased to announce a special issue entitled “Extended Papers Presented in IEVC2026.” This special issue aims to publish high-quality, extended versions of the papers presented at IEVC2026. Authors are invited to submit significantly enhanced manuscripts that expand upon their conference papers to include substantial new contents, such as additional experiments, deeper analysis, improved methodologies, extended discussions, and so on. Accepted papers will be published in either Vol.14, No.2 (December 2026 issue) or Vol.15, No.1 (June 2027 issue), depending on the submission timing, review process, and scheduling issues.

#### 1. Submission Guidelines:

- Submissions must be extended versions of the papers accepted and presented at IEVC2026, for either in the General Papers (oral and poster presentations) category or in Late Breaking Papers (LBP) category. In particular, for papers accepted in the General Papers category, the submitted manuscript must significantly extend the version included in the proceedings to be published in the IEEE Xplore Digital Library (generally at least 30% new material is recommended).
- Manuscripts must follow the official formatting and submission guidelines of the journal.
- Detailed guidelines regarding copyright and how to cite the IEVC2026 proceedings will be prepared before the submission deadline, and provided individually. Therefore, authors are requested to contact the IIEEJ editorial office via the email address listed in Section 5 prior to submission.

#### 2. Review Process:

Submitted manuscripts will undergo the standard double-blind peer review process of the journal.

#### 3. Publication of Special Issue:

IIEEJ Transactions on Image Electronics and Visual Computing  
Vol. 14, No.2 (December 2026 issue) or Vol. 15, No. 1 (June 2027 issue)

#### 4. Submission Deadline:

- For publication in Vol. 14, No. 2 (December 2026 issue): Friday, June 12, 2026
- For publication in Vol. 15, No. 1 (June 2027 issue): Friday, October 30, 2026

#### 5. Contact details for Inquiries: IIEEJ editorial office E-mail: [hensyu@ieej.org](mailto:hensyu@ieej.org)

#### 6. Online Submission URL: <http://www.editorialmanager.com/ieej/>

2027 年 4 月号 ・ 英文誌 6 月号 論文特集号 論文募集  
— 社会貢献と画像電子関連技術特集— (締切 2026 年 10 月 30 日[金])

画像電子学会 編集委員会

近年、地球環境問題、少子高齢化、災害対策、地域格差の是正など、社会が抱える課題は深刻化しさらに多様化・複雑化の一途をたどっています。こうした中で、画像・映像技術、センシング技術、AI、ロボティクス、ドローン利用、データ解析などの進展により、画像電子技術についてもその活用が進められており、防災・減災支援、医療・福祉支援、教育・文化振興、環境保全、地域活性化など、さまざまな分野において課題解決や業務効率化にとどまらず新しい社会的価値の創出も期待されています。一方で、これらの技術を社会に実装する際には、公平性・倫理性・プライバシーなどの観点も考慮する必要があります。技術と社会とのあるべき関係性を再構築する取り組みが求められています。

本特集では、「社会貢献」をキーワードに、画像電子関連技術が社会の持続的発展や人々の生活の質の向上にいかに関与できるかをテーマとした研究成果を広く募集いたします。具体的には、上記の分野への応用研究に加え、社会実装に関する評価や実践的事例、さらに技術の社会的影響に関する調査・提言などを対象とします。

### 1. 対象トピック

以下に対象となるトピック例を挙げますが、これらに限らず、社会貢献と画像電子関連技術に関連するものを広く募集します。

- ・ 地域活性化・地域問題解決・災害予防・災害復興
- ・ 教育・医療・健康・家庭・福祉
- ・ 安心・安全技術、危機管理全般
- ・ ウェアラブル、VR/AR、スマホ利活用
- ・ ユニバーサルデザイン、感性デザイン、ユーザビリティ、UI/UX、人間工学、障害者対応
- ・ AI 活用・自動化・省力化、ロボティクス、ドローン、データ解析、アーカイブ

### 2. 投稿論文種別 (著者に学会員を含む等の掲載件は適用されます)

フルペーパー、ショートペーパー、システム開発論文、実践論文、サーベイ論文

### 3. 論文の取り扱い

投稿様式、査読プロセスとも通常の投稿論文と同様です。採録決定が特集号に間に合わなかった場合には、通常の投稿論文として取り扱います。英文での投稿も受け付けます (英文論文誌 掲載)。

### 4. 特集号発行

画像電子学会誌 (電子版) 2027 年 4 月号, 英文論文誌 2027 年 6 月号

### 5. 論文投稿締切日

2026 年 10 月 30 日 (金)

### 6. 投稿方法および問い合わせ先

画像電子学会 <http://www.iieej.org/>

〒116-0002 東京都荒川区荒川三丁目 35 番 4 ライオンズマンション三河島第二 101 号室

TEL : 03-5615-2893, FAX : 03-5615-2894, E-mail : [hensyu@iieej.org](mailto:hensyu@iieej.org)

論文投稿の手引き

<https://www.iieej.org/call-for-papers/guide-paper-submission/>

2017年1月6日改訂  
2018年7月6日改訂  
2020年10月2日改訂  
2024年11月6日改訂  
2024年12月19日改訂

## 論文投稿の手引き

### 1. 論文投稿

#### (1) 投稿前

- ・ 著者は「論文投稿の手引き」、「スタイルフォーマット」を学会ホームページの「学会誌」、「英文論文誌」欄からダウンロードして、投稿論文の準備をする。
- ・ 「スタイルフォーマット」は TeX 版、Word 版の 2 種類が準備されており、出版コストおよび出版に至る労力削減のために、TeX 版の利用を推奨している。
- ・ 原稿には以下の 6 種類がある。
  - 論文：画像電子技術に関する独創的な研究成果、開発結果、検討結果を学術論文（フルペーパー）としてまとめたもの。新しいアイデアの提案を目的とする通常論文であり、新規性、有用性、信頼性、了解性に基づいて評価する。原則として刷り上がり 8 ページ以内とする。
  - ショートペーパー：研究速報的な性格を持ち、部分的にある程度良好な結果、新しい知見が得られたことを研究報告としてまとめたもの。原則として刷り上がり 4 ページ以内とする。（論文・システム開発論文のいずれとも組み合わせが可能とする）
  - システム開発論文：通常論文の新規性、有用性とは別に、既存技術の組み合わせ、或いはそれ自身に新規性があり、開発結果が従来法より優れているか、或いは他システムへ応用可能であり、新しい知見が示されているもの。原則として刷り上がり 8 ページ以内とする。
  - 資料論文：試験、運用、製造などの新しい成果をまとめたもの。あるいは現場の新技術紹介や提案などで、必ずしも独創的でなくてもよい。原則として刷り上がり 8 ページ以内とする。
  - サーベイ論文：新しい研究の発表を目的とせず、既存の研究・開発をある視点の下に体系付け、比較してその位置づけを明確にし、時代的な変遷を明らかにすること等を目的とする論文。参考文献の網羅、全体像の呈示、客観的評価、特定機関の宣伝に与しないことなどが要求される。また、以下の要件を満たすことが望ましい。その分野に関する状況・課題・及びそれらが、研究者や関係者に与える影響を理解していること。その分野への参入、開発・応用等への進展を促しさらには他分野への展開や新たな研究を触発する可能性を有すること。原則として刷り上がり 8 ページ以内とする。必要に応じて複数パート構成とすること。
  - 実践論文：事例<sup>\*</sup>やフィールドワークに関する発見、工夫、既存技術の新たな組み合わせ等で、事例設計の客観性・相互主観性を有し、個別の事例（フィールドワークも含む）から、他の事例（より一般的な事例も含む）への発展可能性（一般法則化）、利活用・応用可能性などを示せるもの。原則として刷り上がり 6 ページ以内とする。  
※事例とは、画像関連技術の開発・利活用に関する、フィールドワーク、システム開発、システム導入、ソフトウェア実装、アプリケーション開発、表示・提示方法、入出力インタフェース開発、試験・評価方法、サービス、データベース作成等の実例を示す。
- ・ 原稿の投稿および掲載の条件として、論文、ショートペーパー、資料論文、実践論文のいずれも、著者のうち必ず 1 名以上は本学会の正会員または学生会員であることを必要とする。
- ・ 原稿の二重投稿を禁止する。論文、ショートペーパー、システム開発論文、資料論文、実践論文

のいずれも、その採録決定以前に投稿原稿と同一内容のものが、投稿原稿と同一の著者もしくは少なくともその中の1名を含む著者によって他の公開出版物に掲載または投稿中の場合には、原則として採録としない。公開出版物とは、内外の書籍、雑誌および官公庁、学校、会社等の機関誌、紀要などをいう。ただし、本学会または他学会（国外の学会を含む）等の年次大会、研究会、シンポジウム、コンファレンス、講演会などの予稿、特許公開公報、当学会編集委員会で認めたものなどについてはこの限りではない。また、本会誌にショートペーパーとして採録になったものについて、成果を追加し内容を充実させ、論文として再び投稿したものについては二重投稿とみなさない。

## (2) 投稿

- ・ 投稿原稿では、全ての著者情報を削除すること。ただし、参考文献はこの限りではない。
- ・ 著者は以下の論文投稿のページよりユーザ登録を行った上で、ログインして必要事項の記入と「スタイルフォーマット」を使用し作成した論文原稿のアップロードを行う。論文原稿のファイルはpdf形式（TeXのdviを変換したもの）またはWord形式とする。「スタイルフォーマット」を使用するため、原則として図表等は論文原稿内に組み込む。（査読の参考となる別のファイルを同時にアップロードすることは可能）

<https://www.editorialmanager.com/iieej/>

- ・ 投稿に関して質問がある場合には、学会事務局編集担当者に相談のこと。（連絡先：〒116-0002 東京都荒川区荒川 3-35-4 ライオンズマンション三河島第二 101号 TEL: 03-5615-2893 FAX: 03-5615-2894 E-mail: hensyu@iieej.org 画像電子学会事務局編集担当）

## 2. 論文査読と対応

### (1) 査読

- ・ 論文原稿は該当専門分野に詳しい査読者によって審査され、「採録」、「条件付採録」または「返戻」のいずれかに判定される。査読結果は、学会事務局より連絡著者宛にメールで通知される。なお、最終的な採録判定についてはショートペーパーを除き、いずれのカテゴリーの論文であっても2名以上の査読者の採録を必要とする。
- ・ 評価方法
  - 通常論文は、次の基準で評価する。
    - ✓ 新規性：投稿内容に著者の新規性があること。
    - ✓ 有用性：投稿内容が学術や産業の発展に役立つものであること。
    - ✓ 信頼性：投稿内容が読者から見て信用できるものであること。
    - ✓ 了解性：投稿内容が明確に記述されており、読者が誤解なく理解できるものであること。
    - ✓ 文献調査：投稿内容に関連する文献調査が行われており、必要な比較が行われていること。
  - ショートペーパーは、通常論文の新規性、有用性とは別に、研究内容に速報性があり、斬新性の尺度からある程度良好な結果、部分的にでも新しい知見が得られたことを評価する。
  - システム開発論文は、通常論文の新規性、有用性の代わりに次の基準で評価する。
    - ✓ システム開発の新規性：既存技術の統合であっても、組み合わせの新しさ、システムの新しさ、開発したシステムで得られた知見の新しさなどをシステムの新規性として認める。
    - ✓ システム開発の有用性：類似システムと比較し、総合的或いは部分的に優れていること。システムとして先駆的な新しい応用コンセプトが示されていること。組み合わせが実用化のために合目的最適性を有すること。性能限界や実利用システムの性能事例が示されていること。
  - 資料論文は、通常論文における新規性要件については、試験、運用、製造などの新しい成果、

現場の新技术紹介や提案など、必ずしも独創的でなくても優先性があれば新規性として認める。また、有用性要件については新しい成果が既存技術と比較し優れており、学術や産業の発展に役立つものであればそれを評価する。

- サーベイ論文は、通常論文における新規性に代えて網羅性、俯瞰性、客観性について評価する。信頼性、了解性、及び文献調査においては、基本的に通常論文と共通の観点により評価する。
- 実践論文は、通常論文の新規性、有用性に代えて合目的性、独自性、適応性の基準で評価する。また、信頼性、了解性、及び文献調査においては、基本的に通常論文と共通の観点により評価する。(なお、5段階評価において通常論文の評価項目を流用する場合は通常論文の新規性を「合目的性」に、有用性を「独自性」に、信頼性を「適応性及び信頼性の総合」に読み替えて記入するものとする。)
- ✓ 合目的性：投稿内容に合目的性があること。
- ✓ 独自性：投稿内容に著者の独自性や拡張性があること。
- ✓ 適応性：投稿内容が先行事例より優れていること。また、当該事例の合目的性に対する要求条件及び適用範囲において、当該事例が同一条件下で先行事例と比べ、複数の視点で高い客観性・相互主観性を有し、サービスの向上に貢献するものであること。

これらの具体的な例を以下に示す。

- ✓ 合目的性：当該事例への提案内容の導入が一定の目的にかなった方法であることが示されていること。また、分析結果の適用範囲において高い合目的性を有すること。
- ✓ 独自性：当該事例において先駆的な独自応用コンセプトが示されていること。先行事例の組み合わせであっても組み合わせの独自性、導入事例の適用範囲の独自性、事例構成の独自性、他の事例(システム、サービス)への拡張性、現状の画像電子技術に係るサービスの有効性、得られた知見の独自性などを事例の独自性として認める。
- ✓ 適応性：提案事例のもつ適応性を判断するために充分な導入分析結果及び考察が示されていること。提案事例と先行事例との適応性に関する比較結果が、共通条件下で示されていること。提案事例が先行事例と比較し、総合的或いは複数の部分で性能が優れていること、或いは課題、合目的性に対する要求条件及び適用範囲において、より高い客観性・相互主観性を有していること。

## (2) 査読結果受領後の対応

- ・ 「採録」の場合、著者は最終原稿の準備(3.に記載)に進む。
- ・ 「採録(参考意見有り)」の場合、著者は査読者からの参考意見を考慮して論文を修正し、最終原稿の準備(3.に記載)に進む。
- ・ 「条件付採録」の場合、著者は査読者から示された「採録の条件」に基づき、指定された期日まで(60日以内)に論文を修正し、照会内容への対応を明記した回答書とともに再提出し、再度審査を受ける。修正論文には、修正箇所がわかるように、色を付けるか、下線を引くなどすること。また、回答書は必ず添付し、採録条件全てに対し、修正点、主張点、今後の課題などを丁寧に説明すること。
- ・ 「返戻」の場合、著者は以降のプロセスに進むことができない。査読者から示された「返戻の理由」を吟味して論文内容を再検討し、あらためて投稿することができる。

## (3) 修正原稿査読

- ・ 「条件付採録」で再提出する場合は、回答書・修正清書原稿・修正箇所マーク原稿の修正原稿一

式を投稿ページより再提出する。但し、「条件付採録」で指定された期日を過ぎて提出された場合は、新規の初回投稿論文として取り扱われるので注意すること。

- ・ 再提出された修正原稿は原則として同一の査読者によって再審査され、「採録（参考意見付記を含む）」または「返戻」のいずれかに判定される。
- ・ 以降の対応は（2）と同様である。

### 3. 最終原稿提出から出版まで

#### (1) 最終原稿提出

- ・ 「採録」通知受理後は、事務局から最終原稿作成に関するメールが送付されるので、著者はその指示に従い最終原稿一式（電子データ）を準備し、指定の期日までに、事務局へ送付する。
- ・ 最終原稿一式とは、「スタイルフォーマット」を使用して作成した論文原稿のソースファイル（TeX版または Word 版）、その pdf ファイル、すべての図表の eps ファイル（bmp, jpg, png など含む）、著者紹介用写真（縦横比 3:2, 胸より上, 300dpi 以上の eps ファイル, 或いは, jpg ファイル）であり、これらを zip 等の圧縮ファイルに格納して提出する。
- ・ 最終論文原稿には、著者名・組織名、著者紹介、必要により謝辞を記入する。（スタイルファイル内のマクロのフラグを解除する）
- ・ 採録が決定した著者は、当該論文の掲載料を支払わなければならない。なお、別刷りの購入は任意。（別紙 1 「別刷価格表」参照）

#### (2) ゲラ刷りの校正

- ・ 論文掲載号の発行のおよそ 1 ヶ月前に論文原稿のゲラ刷り（ハードコピー）の確認が依頼される。著者は指定の期日まで（約 1 週間）に校正を行い、修正した原稿をスキャンした pdf ファイルを作成し、事務局へメールにて送付する。この際、「別刷り購入票」および「コピーライトフォーム」を記入し、同様にスキャンした pdf ファイルをアップロードする。
- ・ 本会誌に掲載されるすべての記事については、電子的形態による利用も含め、著作権は原則として本会に帰属する。
- ・ 「別刷申込書」および「コピーライトフォーム」は学会ホームページの「学会誌」欄からダウンロードする。（別紙 2 「別刷申込書」、別紙 3 「コピーライトフォーム」参照）

#### (3) 出版

- ・ 最終校正された論文原稿は学会誌、或いは、英文論文誌（電子ファイル形態）としてオンライン出版されるとともに、学会ホームページに掲載される。
- ・ 著者は、グリーンオープンアクセス（セルフアーカイブ）として、本誌からの出版ののちすぐに、自身の著者最終稿を機関リポジトリなどの公的なオンラインリポジトリから公開することができる。

以上

## 目 次

## 巻頭言

年頭にあたり ..... 藤井雅彦... 2(2)

## 論 文

デジタルアート作品のグリーンノイズ拡散型電子透かしによる著作権保護および  
印刷物のオリジナルの判定 ..... 河村尚登, 西村克彦... 3(3)  
Frequency Response and Relaxation of Viscoelastic Inks Based on Jetting Behavior and Droplet Velocity  
..... Naoki MORITA, Akira MIHARA, Shigeki TAKAHASHI, and Shinri SAKAI... 10(10)

## Imaging Today

「ICJ2025 特集—インクジェット, デジタルテキスタイル—」

Introduction ..... 齋藤将史, 朝武 敦, 名越応昇, 宮本栄一... 17(17)  
サステナブルな社会に向けたデジタルテキスタイルの可能性と課題 ..... 藤井雅彦... 18(18)  
コニカミノルタのテキスタイルプリンター技術 NASSINGER SP-1e, O'ROBE ..... 坪谷真吾... 24(24)  
熱転写へ応用するデジタルプリント DTF の最先端技術 ..... 松澤祐至... 29(29)  
拡散方程式に基づく動的表面張力の推定 ..... 田中裕幸, 垣内 徹... 37(37)  
状態方程式を組み込んだインク循環システムの等価回路解析手法 ..... 吉田崇裕, 青木勇太... 44(44)  
用紙波うち低減を実現した RICOH Pro Z75 搭載の紫外線と赤外線を用いた新乾燥技術の開発  
..... 浅田幸輝, 神原一暁, 飛弾竜作, 野沢健二, 秦 輝鮮... 52(52)  
高精細インクジェット印刷装置 ELP04-PCB の開発とそれを用いたフレキシブル基板の量産  
..... 石井洋典, 中農士誠, 安江秀夫, 登口将史, 古山善将,  
中村純也, 田中康大, 織江章裕, 石橋 達, 清水信哉... 59(59)  
連続式インクジェットプリンターの液滴生成過程におけるノズル流量の周期的変動の影響  
..... 栗田雅章, 田川義之... 66(66)  
貧溶媒液中乾燥法を用いたペロブスカイト含有水系高分子微粒子インクの開発  
..... 影山凱紀, 菊地守也, 高橋茂樹, 川口正剛, 榎本航之... 72(72)

## Imaging Highlight

超音波非破壊検査向けの液体塗布が不要な「滑るドライ接触媒質」 ..... 平尾明子... 80(80)

## 教育講座

画像技術を支える人間工学の基礎 (I)-人間工学の紹介 ..... 辛島光彦... 84(84)

## 研究室訪問

兼橋真二研究室 東京農工大学工学部応用化学科 ..... 95(95)

## 会報

..... 97(97)

## 会告

..... 102(102)

## 投稿案内

..... 103(103)

## 日本写真学会誌の目次

..... 104(104)

## 日本印刷学会誌の目次

..... 105(105)

## 画像電子学会誌の目次

..... 106(106)

## Journal of Imaging Science and Technology の目次

..... 107(107)

## 画像閑話

## 令和七年編集委員

編集委員長 宮 本 栄 一 (株式会社キエムトソリューションズ) 編集幹事 石 塚 一 輝 (コニカミノルタ)

編集副委員長 名 越 応 昇 (三菱製紙) 編集副幹事 中 村 一 希 (千葉大学)

山 口 大 地 (リコー)

## 編集委員

新 井 啓 之 (日本工業大学) 次 田 将 大 (東海大学) 名 越 応 昇 (三菱製紙)

石 塚 一 輝 (コニカミノルタ) 常 安 翔 太 (東京工芸大学) 前 田 秀 一 (東海大学)

岩 田 基 (大阪公立大学) 椿 頼 尚 (シャープ) 宮 本 栄 一 (株式会社キエムトソリューションズ)

梅 津 信 二 郎 (早稲田大学) 朝 武 敦 (コニカミノルタ) 村 上 和 歌 子 (リコー)

加 藤 知 之 (東芝テック) 内 藤 裕 義 (大阪公立大学) 山 口 大 地 (リコー)

岸 由 美 子 (リコー) 中 村 一 希 (千葉大学) 山 崎 弘 (元コニカミノルタ)

齋 藤 将 史 (株式会社キエムトソリューションズ) 中 山 信 行 (東京工芸大学) 横 山 優 樹 (富士フイルムビジネスイノベーション)

## 目 次

## 論 文

- エレクトロクロミズムと構造色を融合した新規ハイブリッド光反射デバイス  
 .....白井颯太, 小林範久, 中村一希...110(2)
- ブルーノイズマスク生成におけるエネルギー核の設計と一様性向上 .....石坂敢也...121(13)

## ICJ2025 特集②—インクジェット—

- Introduction .....齋藤将史, 朝武 敦, 名越応昇, 宮本栄一...131(23)
- システム制約を考慮したインク循環供給系の周波数特性計測 .....森 敦司...132(24)
- 透明導電酸化物を用いた不可視インクのインクジェット応用  
 .....船渡結香, 兵藤 哲, 高橋茂樹, 前田秀一...140(32)

## Imaging Today

- 「SDGs など怖くない」  
 Introduction .....椿 頼尚, 朝武 敦, 名越応昇, 山崎 弘...150(42)
- いまの「あたりまえ」を変える～物理劣化・物理再生理論によるマテリアルリサイクル～  
 .....八尾 滋...151(43)
- 資源循環の取組み 再生複合機 .....岡崎 仁...159(51)
- バイオマス素材としてのオレオケミカルの展開 .....森重貴裕...167(59)
- PFAS: 分解技術と資源循環 .....太田響喜, 柴田哲男...171(63)
- 紙製包装を用いた環境対応 .....根本慎二, 脇本篤宏, 池田雅美...181(73)
- トナーやインクジェットに関する環境規制動向 .....山崎 弘...188(80)

## Imaging Highlight

- インク開発と海外法規制 .....齋藤将史...198(90)

## 教育講座

- 人間工学の基礎 (II)—ユーザビリティが高いヒューマンコンピュータインタラクションの設計—  
 .....辛島光彦...204(96)

## 研究室訪問

- 増尾研究室 関西学院大学 生命環境学部 (神戸三田キャンパス) .....214(106)

会報 .....218(110)

会告 .....222(114)

投稿案内 .....223(115)

日本印刷学会誌の目次 .....224(116)

画像電子学会誌の目次 .....225(117)

Journal of Imaging Science and Technology の目次 .....226(118)

## 画像閑話

## 令和八年編集委員

編集委員長	宮 本 栄 一 (社研キエムトソリューションズ)	編集幹事	石 塚 一 輝 (コニカミノルタ)
編集副委員長	名 越 応 昇 (三菱製紙)	編集副幹事	中 村 一 希 (千葉大学)
	山 口 大 地 (リコー)		

## 編集委員

新 井 啓 之 (日本工業大学)	次 田 将 大 (東海大学)	名 越 応 昇 (三菱製紙)
石 塚 一 輝 (コニカミノルタ)	常 安 翔 太 (東京工芸大学)	前 田 秀 一 (東海大学)
岩 田 基 (大阪公立大学)	椿 頼 尚 (シャープ)	宮 本 栄 一 (社研キエムトソリューションズ)
梅 津 信二郎 (早稲田大学)	朝 武 敦 (コニカミノルタ)	村 上 和歌子 (リコー)
加 藤 知 之 (東芝テック)	内 藤 裕 義 (大阪公立大学)	山 口 大 地 (リコー)
岸 由美子 (リコー)	中 村 一 希 (千葉大学)	山 崎 弘 (元コニカミノルタ)
齋 藤 将 史 (社研キエムトソリューションズ)	中 山 信 行 (東京工芸大学)	横 山 優 樹 (社研フイルムビジネスソリューション)

口絵・口絵解説

- 1 「画像からくり」  
第72回 鏡に映る顔：絵画と写真におけるテクニック  
桑山哲郎
- 3 **会告** 2026年度 日本写真学会年次大会発表募集要綱参加募集要綱・日本画像学会誌・日本印刷学会誌・IEEEJ  
Transactions on Image Electronics and Visual Computing

One Point Lecture

- 6 フォトマスター検定過去出題問題の解答と解説 / 水口 淳

**特集** 学会創立 100 周年記念特集：「語り継がれる芸術写真史」

- 13 **解説** 写真史の旅  
田中里実

**特集** 画像保存

- 27 画像保存 -特集にあたって-  
水口 淳

- 28 **解説** ダゲレオタイプあれこれ—国内外の実例を交えて—  
三木麻里

- 34 **解説** 湿板写真を知る：日本の湿板写真の歴史と特徴および制作技法  
安藤千穂子, 田村政実

- 41 **解説** 失われた技法「写真油絵」の復元  
谷 昭佳

- 50 **解説** 写真保存の過去・現在・未来 — 写真保存と写真史における40年から  
付・画像保存セミナーの歩み 2005～2019  
高橋則英

- 62 **解説** 写真の保存額装について  
西丸雅之

- 69 **解説** 人工知能が読み解く行動様式のアークाइブ：体験型映像インスタレーションの試み  
大海悠太, 野口 靖

- 73 **会報** 理事会抄録, 投稿規定等, 執筆の手引き, 著作権の取扱い, 入会のおすすめ

■ 卷頭言		
境界を越え、知をつなぐ——「印刷」の新しい景色を求めて	常安翔太	1
■ 総説 1 特集「印刷の歴史に学ぶ」		
現代のカラー印刷が誕生するまで：西洋色彩版画小史	大森弦史	2
浮世絵版画の歴史と技法	赤木美智	7
型友禅における型紙自動彫刻システムの実用化		
—カッティングプロッタ活用による型紙供給の維持と業界支援—	沖田実嘉子・藪内 快	16
伊勢型紙の歴史と技法	代田美里	20
コロタイプとは	山本 修	27
■ 総説 2 特集「アジアシンポジウム」		
Comparison of CMYK-OGV (Expanded Color Gamut) and XCMYK (eXpanded-gamut CMYK)	Dao Thi Thanh Van and Linlin Liu	32
Quantifying the Correlation between Ink Trapping Efficiency and Color Difference ( $\Delta E_{ab}$ ) in Offset Printing: A Predictive Model for Process Optimization	Nguyen Thanh Phuong, Nguyen Long Giang, Le Cong Danh, Truong The Trung, Che Quoc Long and Nguyen Van Nhat	43
Synthesis and Characterization of ZnS:Mn Quantum Dots for Luminescent Security Ink Applications	Nguyen Thanh Phuong, Nguyen Long Giang and Le Cong Danh	52
Application of XGboost and SVR Models in Improving the Accuracy of Color Prediction Using Optical Methods	Nguyen Van Nhat, Bui Ha Duc and Nguyen Long Giang	59
Overview of Drivers in Implementating Green Supply Chain Management Practices of Green Packaging Suppliers in Vietnam's Packaged Food Sector	Nguyen Thi Kim Thi, Nguyen Anh Tuan and Nguyen Long Giang	66
Optimizing Digital Printed Product Counting Process Using Computer Vision and Deep Learning Technology Based on YOLO Model	Le Thanh Khoa, Nguyen Van Thai and Nguyen Long Giang	73
■ 印象記		
令和7年第21回「紙メディアシンポジウム」印象記		
紙メディアの本質的価値と情報コミュニケーション	大塚真理子・松本 圭	77
2025年度プリメディア研究例会印象記	田中 貞	79
2025年度スクリーン印刷技術研究会セミナー		
「スクリーン印刷の新工法と機能性インキ、製版技術の動向」	江連真市	81
■ Abstract		82
■ 学会だより		86
■ 日本印刷学会賛助会員名簿		92

日本印刷学会誌編集委員

編集委員長	矢口博之 (東京電機大学)	杉山 徹 (大日本印刷 (株))	小泉真一 (共同印刷 (株))
委 員	東 吉彦 (東京工芸大学)	前田満穂 ((独) 国立印刷局)	
	光本知由 (富士フイルム (株))	伊藤 学 (TOPPAN ホールディングス (株))	
	矢島久夫 (東洋インキ (株))	大場久恵 (東京電機大学)	常安翔太 (東京工芸大学)
	次田将大 (東海大学)		
幹 事	小関健一		
表紙デザイン	益田宏樹		

■ 卷頭言	
創立 100 周年の先を目指して . . . . .	東 吉彦 101
■ 総説 特集「デジタルとアナログとの融合」	
デジタル時代におけるアナログ体験価値の再定義	
— instax™ (チェキ) に見る人間中心設計とグローバルブランド構築— . . . . .	高井隆一郎 102
出版事業におけるデジタルとアナログの融合による価値創出 . . . . .	長山知史 108
電子図書館サービスにおける近年の動向 . . . . .	花田一郎 112
■ 論文	
自治体広報誌の文字組版における生成 AI 導入可能性の検証 . . . . .	佐々木喜一郎 119
■ 印象記	
2026 年冬期セミナー印象記 . . . . .	古賀裕之 130
2026 年第 1 回 P&I 研究会シンポジウム印象記 . . . . .	鈴木 学 132
■ Abstract . . . . .	133
■ 学会だより . . . . .	134

日本印刷学会誌編集委員

編集委員長	矢口博之 (東京電機大学)		
委員	東 吉彦 (東京工芸大学)	杉山 徹 (大日本印刷 (株))	小泉真一 (共同印刷 (株))
	光本知由 (富士フイルム (株))	前田満穂 ((独) 国立印刷局)	
	矢島久夫 (東洋インキ (株))	伊藤 学 (TOPPAN ホールディングス (株))	
	次田将大 (東海大学)	大場久恵 (東京電機大学)	常安翔太 (東京工芸大学)
幹 事	小関健一		
表紙デザイン	益田宏樹		

## 会報

### ○新入会員紹介 (2026年3月31日)

正会員

檜垣 徹 (広島大学)

松村 哲哉 (日本大学)

学生会員

HE CHUYI (京都芸術大学)

松田 展 (大阪工業大学)

LI XULINGZHI (名古屋造形大学)

大倉 司 (明治大学)

### ○会員現況

名誉会員	18名
正会員	440名
学生会員	52名
賛助会員	16社 21口
特殊会員	25件

## 編集後記

画像電子学会誌 2026年第2号(4月号)をお届けいたします。本号は、「AI技術と画像電子関連技術論文特集」号で、特集のまえがき〈AI技術と画像電子関連技術論文特集に寄せて〉にもある通り、英文論文を含む合計6編の投稿の中から、採択された和文論文2編を掲載しております。また、英文論文については採録論文が、本年6月号の英文論文誌に掲載の予定です。多岐にわたるAI技術の応用例を、掲載論文を通して知ることは他の分野での研究においても、必ずや何らかの示唆につながるものと信じております。なお、少し先ですが、来年(2027年)の学会誌4月号・英文誌6月号の論文特集テーマは「社会貢献と画像電子関連技術特集」となっており、投稿締め切りは2026年10月30日です。複雑化・深刻化する種々の社会課題の解決に向けた画像電子関連技術の最新研究成果のご投稿を心よりお待ちしております。

本号には、他に、一般論文4編(論文、システム開発論文、システム開発ショートペーパー、実践論文)をはじめ、今給黎編集理事による随想、鈴木総務理事によるグループ紹介など、会員の皆様に興味を持っていただけそうな記事を多数掲載しております。ぜひご一読ください。

さて、2026年3月16日から4日間にわたり、The 9<sup>th</sup> IEEE International Conference on Image Electronics and Visual Computing (IEVC2026)が広島大学霞キャンパスで無事開催されました。今回は2024年に本学会とSister Society契約を結んだIEEE Consumer Technology Society (IEEE CTSoc)の技術的後援を得て行われる最初のIEVCとなり、OpeningではIEEE CTSocの会長であるFabrice Labeau教授にご挨拶いただきました。今回からGeneral Papersカテゴリで採択された発表論文はIEEE Xploreに掲載されることになっており、今後の本学会の国際的認知度の向上と投稿論文の増加に是非つなげてまいりたいと考えています。また、英文誌2026年12月号及び2027年6月号ではIEVC2026のExtended Papersの特集号を予定しております。IEVC2026で発表をして下さった方はこちらにもぜひご投稿ください。

最後に、今後開催予定の学会行事についてご案内申し上げます。まず、今年6月の定例研究会は早稲田大学西早稲田キャンパスでの開催を企画しています。次に「2026年度第54回画像電子学会年次大会」は、2026年8月24日から3日間、筑波大学春日キャンパスにて開催されます。年次大会のJournal Trackの学会誌用原稿提出期限は5月25日、Conference Trackの発表申込期限は6月15日となっております。是非ご投稿・ご参加をご検討ください。

新年度が始まり、会員の皆様は何かとご多忙のことと存じますが、今後とも、画像電子学会へのご支援を賜りますよう、何卒宜しくお願い申し上げます。

(編集理事 会津大学 新田 高庸)

