

モジュール化による視覚障害者の3Dプリンタ操作支援システムの汎用化

日野 正大[†] 田川 直弥[†] 南谷 和範^{††} 岩村雅一^{b)†}

[†] 大阪公立大学大学院情報学研究科, 〒599-8531 堺市中区学園町 1-1

^{††} 大学入試センター 研究開発部, 〒153-8501 東京都目黒区駒場 2-19-23

E-mail: [†]sd24187j@st.omu.ac.jp, ^{†††}minatani@rd.dnc.ac.jp, ^{††††}masa.i@omu.ac.jp

あらまし 視覚障害者は、家電などの機器のディスプレイに表示された視覚情報へのアクセスに困難を抱える。そのため、文字情報を視覚障害者に提供するための読み上げシステムが開発されてきた。既存手法は、ディスプレイに表示された文字の読み上げはできるものの、メニューの操作やスクロールする文字列に対応していないため、視覚障害のある利用者にとって、そのような操作は困難であった。この問題に対して田川らは、メニューの操作やスクロールする文字列に対応した読み上げ手法を提案したが、特定の機器のみを対象としており、他の機種への移植性に課題があった。そこで本研究では、将来的にさまざまな家電製品に応用可能な読み上げシステムの構築を目指し、その第一歩として複数の3Dプリンタに対応可能なシステム設計に取り組む。従来の読み上げシステムが特定デバイスに依存していた問題を解消するため、認識および読み上げ機能をモジュール化し、異なる機種間で切り替え可能な構成とした。

キーワード 視覚障害者, コンピュータービジョン, 3Dプリンタ, スクロールテキスト, 文字認識

Generalization of Display Information Recognition Systems for People with Visual Impairments through Modularization

Shota HINO[†], Tagawa NAOYA[†], Kazunori MINATANI^{††}, and Masakazu IWAMURA[†]

[†] Graduate School of Informatics, Osaka Metropolitan University, 1-1 Gakuencho, Naka, Sakai, Osaka 599-8531, Japan

^{††} Research Division, National Center for University Entrance Examinations, 2-19-23 Komaba, Meguro-ku, Tokyo 153-8501, Japan

E-mail: [†]sd24187j@st.omu.ac.jp, ^{†††}minatani@rd.dnc.ac.jp, ^{††††}masa.i@omu.ac.jp

Abstract People with visual impairments face difficulties in accessing visual information displayed on the screens of household appliances and other devices. To address this issue, reading systems have been developed to provide textual information through speech output. However, existing methods can only read aloud the characters displayed on a screen and do not support operations such as navigating menus or handling scrolling text, making such interactions challenging for users with visual impairments. To tackle this problem, Tagawa et al. proposed a reading method capable of handling menu operations and scrolling text. Nevertheless, their approach was limited to a specific device, raising issues of portability to other devices. In this study, we aim to develop a reading system that can be applied to various household appliances in the future. As a first step, we focus on designing a system compatible with multiple 3D printers. To overcome the device-dependency of conventional systems, we modularized the recognition and speech components, enabling flexible switching between different device models.

Key words People with Visual Impairments, Computer vision, 3D printer, roles of text, character recognition

1. ま え が き

視覚障害者（PVI: People with Visual Impairments）は、視覚情報にアクセスすることが困難であるため、音声や点字による情報提示が必要とされ、今日その支援に関する様々な研究や開発が行われている。その中でも、家電などの機器のディスプレイに表示された情報をPVIに提供するための読み上げシステム

（TTS System: Text-to-Speech System）[1]～[9]の研究が盛んに進められている。TTS Systemは、カメラやOCR（光学文字認識）を用いてディスプレイ上の情報を認識し、スピーカーを通じて音声で出力するものである。例えば、Seeing AI[10]は、スマートフォンのカメラを用いて周囲の情報を認識し、音声で読み上げるアプリケーションである。これは、様々なデバイスのディスプレイに対応可能であるが、単に表示されている文字を

順に読み上げるだけであるので、読み上げられた文字を聞きながらメニューを操作したり、スクロールする文字列を読み上げるのは難しいという問題がある。

この問題に対して、田川らは家庭用 3D プリンタに着目した TTS System の研究 [11] を行い、視覚障害者の 3D プリンタのアクセシビリティの向上を実現している。田川らが開発したシステムは、台に取り付けた Web カメラで 3D プリンタのディスプレイを撮影し、ノート PC で文字認識を行うという仕組みである。このシステムは、ユーザーの選択した部分だけを読み上げる機能が搭載されるなど、操作性に優れている。しかし、画面画像から学習した OCR を用いており、特定の 3D プリンタ固有の文字レイアウトにのみ対応しているため、機種間で画面に表示される文字が異なる場合には対応できない。また、メニューの選択部分の検出も特定の 3D プリンタに特化した処理を行なっているため、機種間でメニューの選択方式が異なる場合には対応できず、汎用性がないという課題がある。

そこで本研究では、これらの既存システムの問題点を踏まえ、将来的に電子レンジやプリンタなどのさまざまな家電製品に応用可能な TTS System の構築を目指し、その第一歩として複数の 3D プリンタに対応可能なシステム設計に取り組む。具体的には、田川らのシステムが適用可能な 3D プリンタ (Prusa i3 MK3) とは表示方式が異なる Prusa mini にも対応可能なシステムを開発した。Prusa i3 MK3 と Prusa mini では、メニュー選択の方式が異なっているためどちらの機種にも適応可能なシステムには、機種ごとの画面構造に応じた選択部分の検出や認識が必要である。そこで、選択部分の検出や認識などの処理をモジュール化し、機種に応じて切り替えられる設計とした。また、処理のマルチスレッド化による処理速度の向上及び既存アルゴリズムの改良によるスクロール文字列の認識精度向上も図った。

2. 関連研究

2.1 家電用文字認識の既存研究

家庭用電化製品のディスプレイ上の文字を視覚障害者のために読み上げる方法については、これまでに広く研究が行われてきた。既存研究における主な関心事は、文字の位置をどのように特定するか、文字の良質な写真をどのように撮影するか、そして文字をどのように適切に読み上げるかという点である。Morris ら [12] による先駆的な研究では、家庭用電化製品の液晶ディスプレイ (LCD) や発光ダイオード (LED) スクリーンに対して、ウェブカメラを用いたアドオン型スクリーンリーダー (SR) を提案している。文字検出および認識は、それぞれマーカーとテンプレートマッチングを用いて実現されている。彼らのアドオン型 SR には、ディスプレイ上のすべての情報を読み上げるモード、重要な部分のみを読み上げるモード、そして画面上に新たに現れた変更箇所を読み上げるモードの 3 つの動作モードがある。彼らは、スクロールする文字列 (彼らの論文では *scrolling display* と呼ばれている) の重要性を指摘しており、それを今後の課題とした。Fusco ら [13], [14] は、*Display Reader* と呼ばれるスマートフォンアプリを提案しており、これはマーカーを用いてディスプレイ上の文字を検出し、PVI が「Be My

Eyes」 [15] を通じて遠隔の晴眼者に文字の内容を読んでもらう支援を行う。Finnegan ら [16] は、血糖値計や血圧計のディスプレイ上に表示される 7 セグメント数字を、画像処理およびコンピュータビジョン技術を用いて読み上げる手法を提案している。Tekin ら [17]、Rasines ら [18]、および Kasar [19] も、同様に画像処理およびコンピュータビジョン技術を用いた手法を提案している。さらに、Moreira ら [20] および Boonag ら [21] は、物体検出技術を用いて医療機器のディスプレイ上の文字を読み上げる手法を提案している。加えて、Lee ら [22] は、遠隔の晴眼者支援の文脈において、ディスプレイ上の文字認識の必要性を指摘している。

Guo ら [23] は、PVI が電子レンジの操作パネルなどのフラットなインタフェースにアクセスすることを支援するインタラクティブツール「*VisLens*」を提案している。VisLens v2 では、ディスプレイ上の文字を読み上げる機能が追加されている。文字画像の前処理を行った後、オープンソースの OCR エンジンである Tesseract OCR [24] を適用して文字を読み上げる。

以上に述べたように、我々の知る限りでは、スクロールする文字列およびメニュー項目をアドオン型 SR で読み上げることを可能にしたのは、田川らの研究 [11] が初めてである。スクロール文字の重要性は [12] において指摘されている。

2.2 3D プリンタ支援システム

ここでは、操作性に優れた TTS System として、田川らの研究 [11] を取り上げる。近年登場した 3D プリンタは、触察の機会を増やすという点と自助具の作成という点で視覚障害者から注目を集めている。しかし、3D プリンタの操作は視覚に依存している部分が多く、視覚障害者にとっては操作が困難であることから、そのアクセシビリティの向上が求められている。田川らは、そこに着目し、視覚障害者の 3D プリンタのアクセシビリティの向上を目指して、3D プリンタのディスプレイに表示される文字を認識し、音声で読み上げるシステムを開発した。具体的には、図 1 のように台に取り付けた Web カメラで 3D プリンタのディスプレイを撮影し、ノート PC で文字認識を行うという仕組みになっている。田川らのシステムは、Prusa i3 MK3 という 3D プリンタ固有の文字レイアウトに対応できるよう設計されており、画面画像から学習した OCR を用いて、3D プリンタのディスプレイに表示される文字を認識することができる。さらに、読み上げる文字は、メニューの選択肢のうち選択されているもののみを読み上げるように設計されており、これにより、読み上げられた文字を聞きながら 3D プリンタを操作することが可能になっている。また、Prusa i3 MK3 は、図 2 に示すように、カーソル (>) を上下に移動させることでメニューの選択を行うため、カーソルが現在選択しているメニューの情報を示すために、カーソルのある行の文字列を読み上げている。具体的には、カーソルのある行の文字列を読み上げるために、画面全体を OCR で認識し、カーソルの位置を特定してから、カーソルのある行の文字列を読み上げるという処理を行っている。

しかし、田川らのシステムには、いくつかの制約がある。まず、3D プリンタの画面に表示される文字は機種によって異なるため、田川らのシステムは他のデバイスに対応することができ

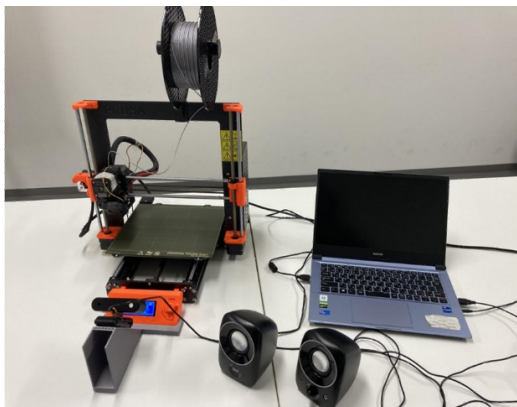


図 1: 田川らのシステム ([11] より引用)

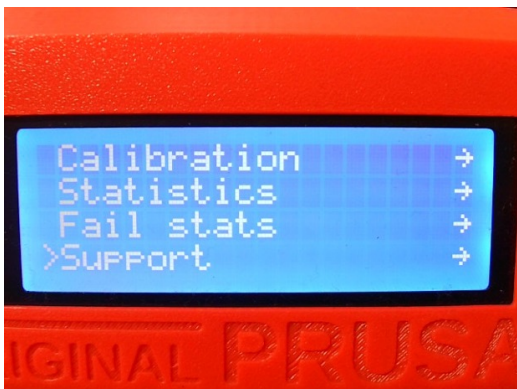


図 2: Prusa i3 MK3 のメニュー選択

ない。さらに、表示されるメニューのうち選択されている部分の検出は、カーソル位置の検出によって行なっているため、メニューの選択方式が異なる他のデバイスに対応することができない。これらの制約により、田川らのシステムは特定の 3D プリンタ (Prusa i3 MK3) に特化したシステムとなっており、汎用性に欠けるという問題がある。

3. 提案システム

本研究では、既存の TTS System の問題点を踏まえ、将来的にさまざまな家電製品に応用可能な TTS System の構築を目指し、その第一歩として複数の 3D プリンタに対応可能なシステム設計に取り組む。具体的には、田川らのシステムが適用可能な 3D プリンタ (Prusa i3 MK3) とは表示方式が異なる Prusa mini にも対応可能なシステムを開発した。先に述べたように田川らのシステムが特定デバイス (Prusa i3 MK3) にのみ最適化され、処理も複雑であったという課題に対し、処理の簡素化とモジュール化を進めることで、複数デバイスへ容易に適用可能な TTS System を構築した。各処理を独立モジュールとして再設計し、OCR モジュールや選択位置判定モジュールなど、デバイス依存性の高い部分のみを差し替えて最適化できるアーキテクチャとした。これにより、Prusa i3 MK3 (カーソル選択) および Prusa mini (白背景によるハイライト) を同一フレームワークでサポートしている。さらに、処理高速化のために施した処理パイプラインのマルチスレッド化によって発生する競合

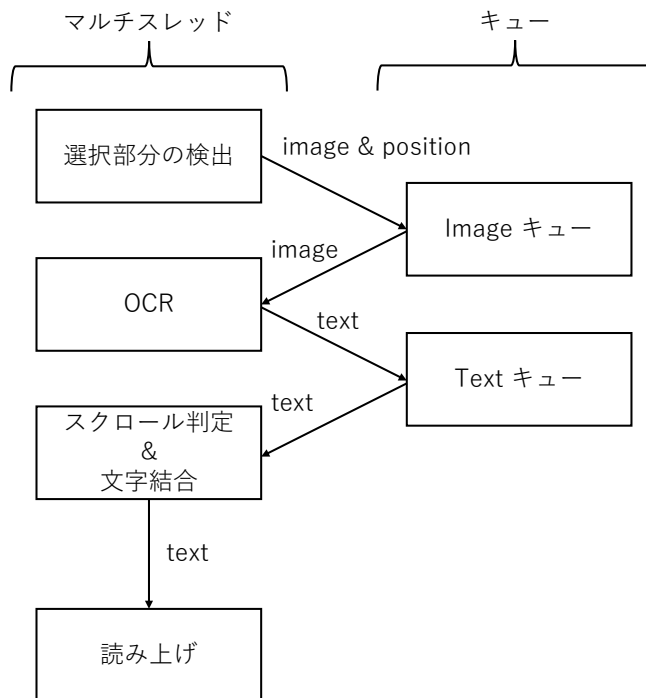


図 3: システム全体フロー

を防ぐため、スクロール判別アルゴリズムを改良した。

3.1 システム全体フロー

本手法の処理は図 3 のようなパイプラインで構成され、各段を独立スレッドとして並列実行することでスループットを向上させた。

- (1) メニュー選択部分の検出
- (2) 文字認識
- (3) スクロール判定 & 文字列結合
- (4) 読み上げ (TTS)

上記の各段はキューで非同期連結し、フレーム入力から音声出力までのレイテンシを低減している。

3.2 モジュール構成と汎用性

本システムは、先に述べたフローごとに以下の主要モジュールから構成される。

- **選択部分検出モジュール**: デバイス固有の選択表現 (カーソル/白背景) に対応。
- **OCR モジュール**: 画面文字の認識を担当。
- **スクロール判定 & 文字結合モジュール**: 連続フレーム間の变化検出と横スクロール文字列の結合を担当
- **TTS モジュール**: 確定したテキストを音声化。

従来のシステムでは、画面様式 (デバイス) ごとにシステム全体を再学習・改修する必要があったが、本手法では、OCR モジュールや選択位置判定モジュールなど、デバイス依存性の高い部分のみを差し替えて最適化できるため、デバイス追加時は当該モジュールのみを最適化すればよい。

3.3 選択部分検出モジュール

先に述べたように Prusa i3 MK3 では、メニューの選択は、カーソル (>) を上下に移動させることで行うため、カーソルが現在選択しているメニューの情報を示すために、カーソルの



図 4: Prusa mini のメニュー選択

ある行の文字列を読み上げる必要がある。しかし、Prusa mini では、図 4 に示すように、選択している行の文字列が白背景で表示されるため、田川らの手法を用いることはできない。そこで、Prusa mini では、選択している行の文字列を読み上げるために、カーソルの位置を特定する処理ではなく、白背景の行の文字列を直接読み上げるようにした。具体的には、図 5 のように画面を検出したのち、そこからディスプレイ部分を切り出し、さらにそこから白背景部分を切り出すというステップで選択部分の検出モジュールを構成した。また、本手法のフローに合わせるために Prusa i3 MK3 の選択部分検出モジュールも田川らの手法をもとに改良した。具体的には、図 6 のように 4 行の選択肢の中からカーソル位置をピクセル値をもとに検出し、その行のみを切り取り、それを OCR モジュールに渡すというように変更した。

3.4 OCR モジュール

田川らのシステムでは、文字認識に Prusa i3 MK3 の画面画像から学習した OCR を用いていた。しかし、Prusa mini の画面は Prusa i3 MK3 とは異なるため、田川らのシステムを用いることはできず、本手法では Prusa mini の画面に適用可能な OCR を追加した。ここでは、田川らのシステムのように Prusa mini の画面画像から学習した OCR を用いるのではなく、汎用性のある OnnxOCR [25] を用いて文字認識を行うことにした。OnnxOCR は、PaddleOCR ベースの検出・認識モデルであり“軽量・高速”な OCR エンジン（ライブラリ）である。多くの言語に対応しており、今後、他の家電製品にも応用可能な汎用性を持たせるためにこれを選択した。

3.5 スクロール文字結合モジュール

パイプラインをマルチスレッド化した結果、フレーム取得・OCR 結果のタイミングずれなどにより、スクロール中の一時的な誤読（文字抜け・重複）やフレームの順序逆転が生じうる。これを抑制するため、田川らのシステムで採用されていたスク

表 1: 操作タスク

Task	メニュー遷移
Task 1	User Interface → Menu Timeout
Task 2	Input Shaper → Y-axis Filter
Task 3	Language & Time → Time Zone Hour Offset
Task 4	NetworkStatus → Gateway

ロール文字の結合アルゴリズムを改良した。田川らのアルゴリズムでは、各文字位置ごとに投票テーブルを作成し、各文字位置で最も投票数が多い文字を選択して結合することで、横に流れるテキストをより正確に生成していた。しかし、田川らのアルゴリズムでは、図 7 のように位置合わせは後ろにのみずらしていくため、新しいフレームの文は前の文の後ろにしか結合できない。これでは、フレームの順序逆転に対応できないため、本手法では、図 8 のように前方向にも挿入位置探索を行い、新しいフレームの文の正しい位置が前の分の先頭方向であったとしても位置合わせ及び文字の結合が正しく行えるようにした。

4. 予備実験

提案手法の有用性を調査するために、実験を行った。今回は、複数人の視覚障害者による実験が実施できなかったため、盲目の著者 1 名を実験参加者とした予備実験で代替した。実験では、田川らのシステムの有用性の評価と同様に、こちらが用意した 3D プリンタの操作タスクを提案手法と普段実験参加者が利用している情報伝達アプリの Lookout でそれぞれで行った。操作タスクは、表 1 のようにメニューを遷移し目的のメニューに到達することとし、4 種のタスクをそれぞれの手法で 1 回ずつ計 8 回行い、それぞれの手法での操作完了までの時間を計測した。同じタスクを別々の手法で行う場合、先に行う手法の方で項目の順番を覚えてしまい、後に行う手法が有利になる可能性がある。そこで、提案手法と従来手法を交互に変えてタスクを行い、それぞれの手法のタスクの順番をシャッフルした。結果は、図 9 の通りであり、全てのタスクにおいて提案手法の操作時間の方が短く、タスクの成功率も上回っている。また、実験参加者から“認識精度は問題なく使えるレベルである”、“レスポンスの速度も許容範囲でありリアルタイム性がある”、“スクロール中にピープ音になっていることは操作する上で有用である”といったフィードバックを得た。以上のことから本手法が有用である可能性が示された。

5. 議論

5.1 応用可能性

モジュール境界が明確になったことで、各デバイスに適した OCR の柔軟な選択が可能になった。とくにマルチスレッド化により、従来はレイテンシの観点から実運用が難しかった高精度・高負荷 OCR も使用できる可能性が広がり、多様な家電ディスプレイへの展開が見込める。また、スクロール文字の結合アルゴリズムの改良により改良した結合アルゴリズムにより、途中から撮影したスクロールテキストでも正しい系列復元が可能



図 5: 白背景の検出

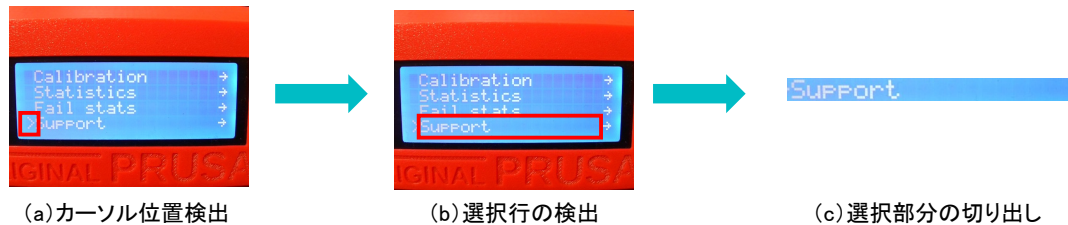


図 6: カーソルの検出

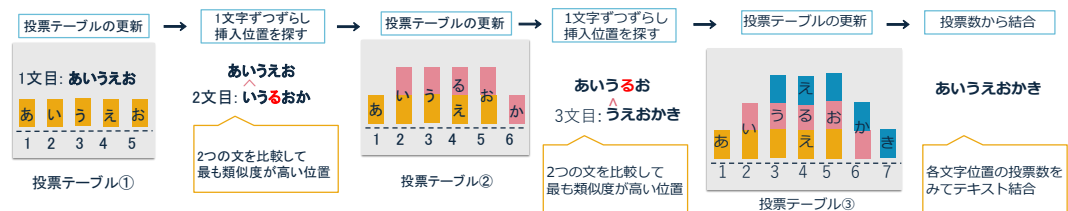


図 7: 田川らのアルゴリズムの概要

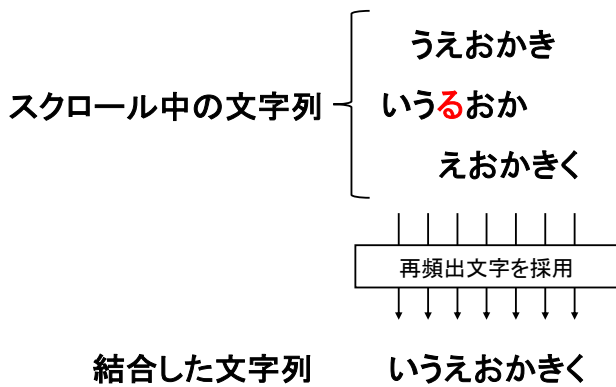


図 8: 本手法のアルゴリズムの概要

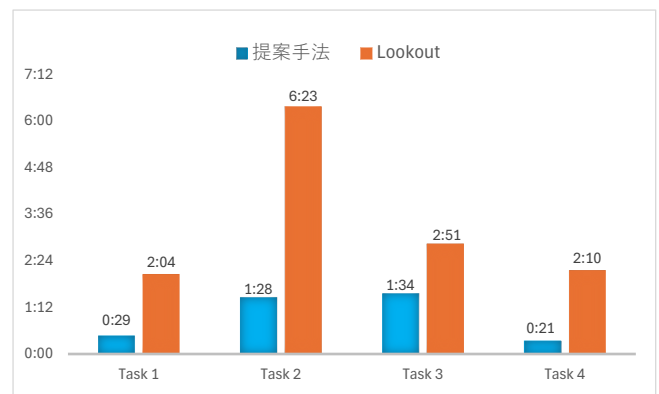


図 9: 実験結果

になった。これは駅・新幹線内の掲示板やニュースティッカー、電光掲示サイネージ等への適用を後押しし、PVI が時変視覚情報を正確に取得できる機会を拡大する。

5.2 課題

現状、Prusa mini における文字認識モジュールは文字列の抽出に特化しており、非文字オブジェクト（スイッチやアイコン）の検出・認識には未対応である。そのため、今後は非文字オブジェクトを既存 OCR へ追加学習することやテンプレートマッチングなどを含めた対応が必要である。また、汎用 OCR (OnnxOCR) では、Prusa i3 MK3 において学習済みデバイス特

化 OCR の精度を上回れない場面が残る。低解像度・発光ノイズ等の条件で誤認識が散見されているためその対応も今後必要である。

6. まとめと今後の展望

本研究では、既存の TTS System が特定デバイスにのみ最適化されており、汎用性がないという課題に対して、処理の簡素化とモジュール化を行い、選択部分検出 / OCR / スクロール判定 & 文字結合 / TTS を独立モジュールとして再設計した。この結果、Prusa (カーソル選択) および Prusa mini (白背景ハイ

ライト)の双方に同一フレームワークで対応可能となった。また、パイプラインをマルチスレッド化して実行し、フレーム入力から音声出力までのレイテンシを低減した。あわせて、マルチスレッド化に伴うタイミング不整合に起因する誤読(文字抜け・重複)やフレームの順序逆転に対処するため、スクロール文字列の結合アルゴリズムを改良した。1人の実験参加者を対象にした予備実験では、4つの操作タスクにおいて、提案手法は操作時間・成功率の両面で従来手法を上回り、実時間操作に必要な応答性と実用的な認識精度を示した。今後は、残された課題(非文字UIの扱い、汎用OCRの精度限界)の解決に向けて、既存OCRと非文字理解の統合を進めることで、PVIに対する実環境での支援の質をさらに高められると期待される。また、複数人の視覚障害者による実験を実施することも今後の課題である。

謝辞 本研究は、JST RISTEX「SDGsの達成に向けた共創的研究開発プログラム(ソリューション創出フェーズ)」JPMJRX21I5の補助を受けて実施した。

文 献

- [1] Chiraphat Boonnag, Piyalitt Ittichaiwong, Wanumaidah Saengmolee, Narongrid Seesawad, Amrest Chinkamol, Saendee Rattanasomrerk, Kanyakorn Veerakanjana, Kamonwan Thanontip, Warissara Limpornchitwilai, and Theerawit Wilaiprasitporn. Pacman: A framework for pulse oximeter digit detection and reading in a low-resource setting. *IEEE Access*, Vol. 10, No. 15, pp. 13196–13204, 2022.
- [2] E Finnegan, M Villarroel, Carmelo Velardo, and L Tarassenko. Automated method for detecting and reading seven-segment digits from images of blood glucose metres and blood pressure monitors. Vol. 43, pp. 1–15, 11 2019.
- [3] Giovanni Fusco, Ender Tekin, Nicholas A. Giudice, and James M. Coughlan. Appliance displays: Accessibility challenges and proposed solutions. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '15, 2015.
- [4] Giovanni Fusco, Ender Tekin, Richard E. Ladner, and James M. Coughlan. Using computer vision to access appliance displays. In *Proceedings of the 16th International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '14, 2014.
- [5] Thotringam Kasar. Recognition of seven-segment displays from images of digital energy meters. In *Advances in Image and Video Processing*, pp. 1–10. Springer Singapore, 2016.
- [6] Lucas P. Moreira. Automated medical device display reading using deep learning object detection, 2022. Preprint.
- [7] Tim Morris, Paul Blenkhorn, Luke Crossey, Quang Ngo, Martin Ross, David Werner, and Christina Wong. Clearspeech: A display reader for the visually handicapped. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 14, No. 4, pp. 492–500, 2006.
- [8] Irati Rasines, Pedro Iriondo, and Ibai Díez. Real-time display recognition system for visually impaired. In *Computers Helping People with Special Needs*, pp. 623–629. Springer Berlin Heidelberg, 2014.
- [9] Ender Tekin, James M. Coughlan, and Huiying Shen. Real-time detection and reading of LED/LCD displays for visually impaired persons. In *2011 IEEE Workshop on Applications of Computer Vision (WACV)*. IEEE, 2011.
- [10] Seeing AI. <https://www.microsoft.com/en-us/garage/wall-of-fame/seeing-ai/>. Accessed: 2025-11-3.
- [11] Naoya Tagawa, Masakazu Iwamura, Kazunori Minatani, and Koichi Kise. Making 3d printer accessible for people with visual impairments by reading scrolling text and menus. In *Proceedings of the 26th International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '24, 2024.
- [12] Tim Morris, Paul Blenkhorn, Luke Crossey, Quang Ngo, Martin Ross, David Werner, and Christina Wong. Clearspeech: A display reader for the visually handicapped. Vol. 14, No. 4, pp. 492–500.
- [13] Giovanni Fusco, Ender Tekin, Richard E. Ladner, and James M. Coughlan. Using computer vision to access appliance displays. In *Proceedings of the 16th international ACM SIGACCESS conference on Computers and Accessibility*, ASSETS '14. ACM Press, 2014.
- [14] Giovanni Fusco, Ender Tekin, Nicholas A. Giudice, and James M. Coughlan. Appliance displays: Accessibility challenges and proposed solutions. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '15. ACM Press, 2015.
- [15] Be My Eyes. <https://www.bemyeyes.com> (accessed 2024-5-24).
- [16] E. Finnegan, M. Villarroel, C. Velardo, and L. Tarassenko. Automated method for detecting and reading seven-segment digits from images of blood glucose metres and blood pressure monitors. Vol. 43, No. 6, pp. 341–355.
- [17] Ender Tekin, James M. Coughlan, and Huiying Shen. Real-time detection and reading of led/lcd displays for visually impaired persons. In *2011 IEEE Workshop on Applications of Computer Vision (WACV)*. IEEE.
- [18] Irati Rasines, Pedro Iriondo, and Ibai Díez. *Real-Time Display Recognition System for Visually Impaired*, pp. 623–629. Springer Berlin Heidelberg.
- [19] Thotringam Kasar. *Recognition of Seven-Segment Displays from Images of Digital Energy Meters*, pp. 1–10. Springer Singapore.
- [20] Lucas P. Moreira. Automated medical device display reading using deep learning object detection, 2022.
- [21] Chiraphat Boonnag, Piyalitt Ittichaiwong, Wanumaidah Saengmolee, Narongrid Seesawad, Amrest Chinkamol, Saendee Rattanasomrerk, Kanyakorn Veerakanjana, Kamonwan Thanontip, Warissara Limpornchitwilai, and Theerawit Wilaiprasitporn. Pacman: A framework for pulse oximeter digit detection and reading in a low-resource setting. Vol. 10, No. 15, pp. 13196–13204.
- [22] Sooyeon Lee, Rui Yu, Jingyi Xie, Syed Masum Billah, and John M. Carroll. Opportunities for human-ai collaboration in remote sighted assistance. In *27th International Conference on Intelligent User Interfaces, IUI '22*. ACM.
- [23] Anhong Guo, Xiang 'Anthony' Chen, Haoran Qi, Samuel White, Suman Ghosh, Chieko Asakawa, and Jeffrey P. Bigham. VizLens: A robust and interactive screen reader for interfaces in the real world. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology, UIST '16*. ACM.
- [24] Tesseract OCR. <https://github.com/tesseract-ocr/tesseract>. Accessed: 2025-07-18.
- [25] OnnxOCR. <https://github.com/jingsongliu/jing/OnnxOCR>. Accessed: 2025-11-3.