

## [研究助成 (A)]

## 視覚障害者支援のための可視情報伝達技術の構築

## Construction of Visual Information Transmission Technology to Assist People with Visual Impairment

2181004



研究代表者

大阪府立大学  
大学院工学研究科

准教授

岩村 雅一

## [研究の目的]

物体認識技術の著しい発展により、人と同程度の認識性能が実現できるようになった。そのため、これらの技術を視覚障害者の「目」として使う試みが増えている。視覚情報は膨大である反面、視覚障害者が情報を受け取ることができるのは触覚と聴覚のみである。そのため、むやみに周囲の情報を伝えることはかえって混乱を招く。このように、認識技術の利用者である視覚障害者に認識結果をどのように伝えるかという問題は実用上非常に重要であるものの、これまでほとんど考慮されてこなかった。本研究では認識結果をいかに視覚障害者に伝えるかという問題に初めて焦点を当て、解決を試みる。具体的には、視覚障害者に視覚情報（認識結果）を伝える前に、情報を取捨選択する方法を考案し、その実現並びに評価を行う。

## [研究の内容, 成果]

## 1. 能動的な情報取得と受動的な情報取得

深層学習に代表される近年の認識技術の飛躍的な精度向上により、物体認識や文字認識などの認識技術を視覚障害者の目の代わりに用いる試みが盛んである。中でも、Seeing AIやTapTapSee, Envision AIなどのアプリは、目の前に何があるかを知るために多くの視覚障害者に使われ、重宝されている。ただし、その暗

黙の前提として、(1) 物体や文字情報に気付く、(2) その位置を把握する、(3) その場所にカメラを向けて利用者が撮影を行うという、利用者の「能動的な行動」が求められる。しかし、そもそも視覚障害者が情報の存在に気づけるのか、そして、その位置を把握できるのかといった疑問も生じる。また、位置が把握できても、認識に適した写真を視覚障害者が撮影することにも困難を伴う [1, 2]。このように、「能動的な情報取得」が視覚障害者には難しいという問題は、重要であるにもかかわらず、これまであまり注目されてこなかった。

そこで本研究では、利用者が存在や位置を知らない情報を取得できる「受動的な情報取得」という枠組みを考える。受動的な情報取得とは、能動的な情報取得が前提としている3要件に依存しない形での情報取得である。このような情報取得を実現する1つの方法として、利用者の周囲を常時撮影して、常時認識することが考えられる。常時撮影すれば、利用者が気づいていない情報もカメラで撮影できるし、撮影した画像を全て認識すれば、利用者が意識していない周囲の視覚情報が得られるからである。その結果、提案する「受動的な情報取得」は、従来の「能動的な情報取得」とは比べものにならない程、膨大な情報（認識技術の出力として得られる認識結果）が得られる。しかし、得られた情報を全て利用者に伝えれば、利用者の混乱を招くため、受動的な情報取得では、情報を要約し

たり、取捨選択したりして、利用者に伝える情報を削減することが重要である。

本研究では、提案する「受動的な情報取得」について、以下の2つを実施した。1つ目は、「受動的な情報取得」という方式を視覚障害者が本当に必要としているのかの検証である。視覚障害者を対象としたオンラインアンケートを実施して、視覚障害者が周囲の視覚情報を必要としているのか、現状の能動的な情報取得に困難な点があるのかなどを調査した。2つ目は、提案する方式を実現する際に、認識技術の出力として得られる情報（認識結果）を要約したり取捨選択することの重要性の検証である。擬似的に買い物を行うタスクを用意して、9名の視覚障害者に参加してもらった実験を行った。その実験の中で、時間的ならびに意味的に情報を要約することの善し悪しについて、客観評価ならびにインタビューによる主観評価を行った。

## 2. 受動的な情報取得の需要の調査

本研究で提案する視覚障害者の受動的な情報取得は、我々が知る限り、これまで着目されておらず、既存研究が存在しない。そのため、受動的な情報取得が視覚障害者に本当に必要とされているのかという需要を調査するため、テキストや色、明るさ等の触っただけでは分からない情報（以下、周囲の視覚情報と呼ぶ）に関するオンラインアンケートを行った。オンラインアンケートは3部構成で23問から成る。内訳は、第1部は回答者の視覚障害の状況について6問、第2部は周囲の視覚情報に対する回答者の経験について7問、第3部は回答者の情報技術の利用に状況について10問である。本アンケートは、視覚障害者が入っているメーリングリストにオンラインアンケートの回答依頼メールを送り、視覚障害者20名（全盲8名、弱視12名）の方から回答を得た。更に3.の実験の参加者にも同じ質問をした結果、全部で視覚障害者29名（全盲14名、弱視15名）から回答を得た。回答者の年齢層は20代、30代、40代

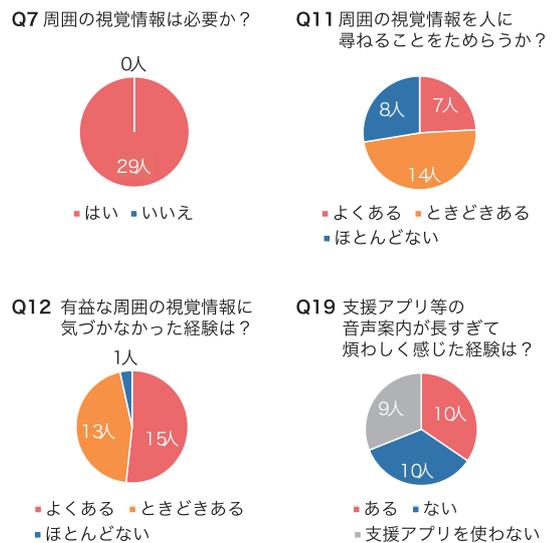


図1 アンケートの結果 (抜粋)

が各4名、50代が6名、60代が11名である。

需要の調査に大きく関わる部分のアンケートの内容を示し、回答割合を図1に示す。

**Q 7 : 周囲の視覚情報を必要と感じていますか？**

先天盲・中途失明に関わらず全員が必要と回答した。文字や色の情報等、生活全般で視覚情報が欲しいとの回答が得られた。

**Q 11: 周囲の視覚情報を得る時に、人に尋ねるのをためらうことがありますか？**

「よくある」「ときどきある」が全体の約70%を占めた。このことから視覚情報の取得を支援するデバイスの需要は十分あると考えられる。

**Q 12: 有益な周囲の視覚情報に気づかず、後からそれを知って有益な情報を逃していたと感じたことはありますか？**

半数が「よくある」と回答し、残りのほとんどの方も「ときどきある」と回答した。トイレの場所や空席に気づけなかった、近くに人がいることに気づかず物事を尋ねることができなかった、同伴者にすべての情報を読み上げてもらえず自身が望んでいるものがあっても関わらず購入できなかった、新商品や特売品が分からなかったなど、多岐に渡る場面の例が挙げられた。このことから、視覚障害者が気づけてい

ない視覚情報を取得できる受動的な情報取得の需要は十分あると考えられる。

**Q 19:** 支援アプリ等の音声案内が長すぎて煩わしく感じた経験がありますか？

支援アプリ等を使っていると回答した20名の内、約半数の10名が「ある」と回答した。自身の求めている情報が読み上げられるまでに時間が掛かったり、結局何が読み上げられたのか分からず時間を無駄にするという意見があった。これはスマートフォンカメラでの撮影の難しさ、検出された情報を要約していないことに起因していると考えられる。

**Q 20:** 支援アプリの不満な点があれば教えてください。

撮影対象をフレーム内に収めることができず、上手く撮影することができないという意見が一定数見受けられた。これは、視覚情報がどこにあるのか正確に把握できていないことや、カメラでの撮影に不慣れであることが原因だと考えられる。

Q 12で回答者のほぼ全員が周囲の視覚情報に気づかなかった経験があると回答していることから、既存の能動的な情報取得だけでは視覚障害者を十分支援できない場面の存在がうかがわれる。そうした場面では、本研究で検討する受動的な情報取得が役に立つ可能性がある。また、Q 19では回答者の約1/3が既存の能動的な情報取得で音声案内が長すぎて煩わしく感じた経験があると回答している。受動的な情報取得では能動的な情報取得より検出される情報の量がより増えるため、情報の要約がより重要になると考えられる。

### 3. 情報の要約方法の検討と評価

本研究では、汎用的な情報要約アプローチである、一度伝えた情報を再度伝えるかという時間的要素、概念的に上位の言葉で情報をまとめて伝える意味的要素を考慮し、その有効性を検証する。図2のように、時間的要素と意味的要素それぞれ有り無しの4つの要約方法を比較す

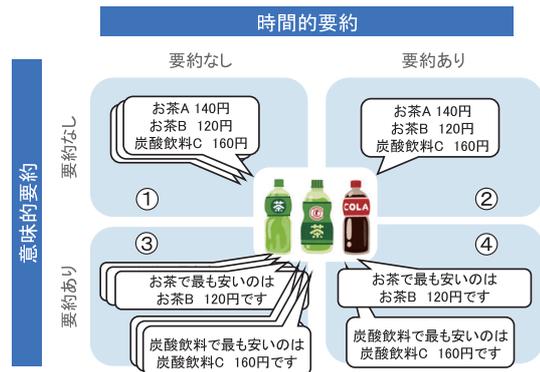


図2 本研究で検証する4種類の要約方法



図3 実験の様子

る。なお、実験では「お茶A」、「お茶B」、「炭酸飲料C」などではなく、市販されている実在の商品名を用いた。

視覚障害者9名(全盲6名、弱視3名)に音声案内システムを使用して疑似的な買い物をして貰った。そして、音声案内システムが読み上げる商品名と価格から最も安い商品を探すというタスクを実験参加者にして貰った。このタスクを課した理由は、情報を正確に受け取る必要があるタスクでの所要時間とスコア(商品選択の正確さを表す)を測り、定量的に評価を行うためである。実験の様子を図3に示す。実際の買い物風景とは大きく異なるものの、実験参加者からは目が見えていないがゆえに普段の買い物と異なる感覚はあまりなく、ウィンドウショッピングをしている感覚に近いかもしれないといった意見が得られた。実験では、図2に記載の要約方法4種類を①から④の番号順に試してもらい、タスクを終えるまでの所要時間



図4 試作した音声案内システム

と商品選択のスコア、及びアンケートを通して評価した。実験では、図4に示す広角レンズを搭載したウェアラブルカメラを用いて音声案内システムを作成した。誤認識による影響を除外するため、商品情報を含むQRコードを撮影し、その商品情報を骨伝導ヘッドセットを通して音声でアナウンスする。

まず、主観的評価の結果を示す。ほとんどの実験参加者が、要約方法③または④が使いや

すかったと回答した。このことから意味的要素での要約に有効性が認められる。理由として、認識した順にランダムに情報をアナウンスされるよりも、上位のカテゴリでまとめられることにより情報が減り整理しやすくなったことが、実験参加者の評価から分かった。

次に、客観評価の結果として、飲み物で行ったタスクとお菓子で行ったタスクの平均所要時間と平均スコアに基づく客観評価の結果を示す。所要時間は、音声案内システムの最初の認識が始まってから実験参加者が回答するために手を挙げるのに掛かった時間である。ただし、所要時間が180秒を超えた場合はそこで実験を打ち切り、その段階で回答を求める。スコアは回答の正確さを表す指標であり、100%が最も良い場合、0%が最も悪い場合を表す。実験参加者ごとの平均所要時間を図5に、要約方法ごとの平均スコアを図6に示す。

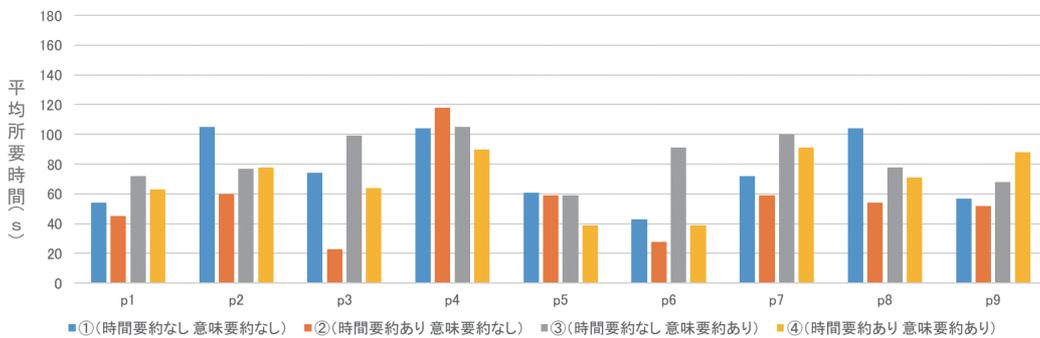


図5 実験参加者ごとのタスクの平均所要時

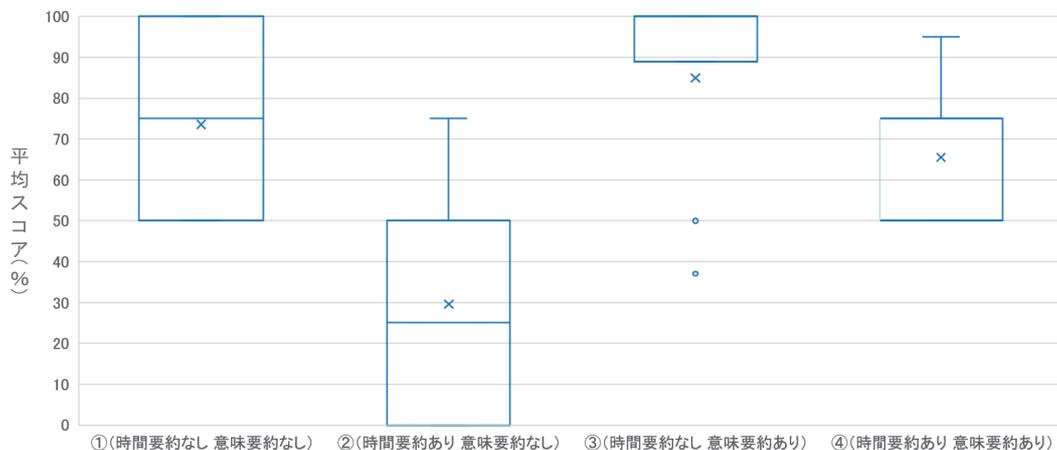


図6 各要約方法におけるタスクの平均スコア

平均所要時間では、要約方法ごとに相関性は見られなかった。実験参加者ごとに見ると、どの要約方法であってもかかる時間は基本的に近い値を取っており、タスクの所要時間は実験参加者に依存していることが分かる。また先天的な失明の実験参加者ほど、音声での情報の取得を苦しめない可能性があるかと予想していたが、そのような相関性は確認できなかった。現在の視力になった時期で注目すると、現在の視力になった時期の遅い p1, p5, p6, p9 の平均所要時間が他の5名と比べ60秒前後と短い。p6の方からは、目が見えていたところに今回の実験でアナウンスされた商品を見たことがあり、商品をイメージしやすく情報を受け取りやすいといった意見があった。このことから、中途失明の方は目が見えていたころの視覚的イメージから情報の受け取り方に異なる特徴がある可能性が考えられる。しかし、p1, p5, p9の方から似たような意見は出ておらず、より詳細な考察を行うにはさらに多くの視覚障害者を対象に実験を行う必要がある。

平均スコアでは、要約方法の時間的要素の有無で異なる結果が得られた。一度だけ読み上げる要約方法②と④のスコアが低く、認識する度に繰り返し読み上げる要約方法①と③のスコアが高かった。これは、聞き逃した場合や商品名を忘れた場合に再確認することができなかったためだと考えられる。また、要約方法①と③、要約方法②と④で比較すると、カテゴリでまとめる効果があることが見て取れる。これらから、時間的要素での要約は聞き逃しの観点から単純に用いることは難しいことがわかった。一方、意味的要約には一定の有効性があることがわかった。

#### [今後の研究の方向, 課題]

本報告の詳細は成果 [1, 2] を参照いただきたい。スペースの関係で、本報告は成果 [1] の抜粋である。成果 [2] では、認識結果の伝達方法として質問応答 (QA) システムを用いて評価した。その結果、QA システムを適切に設定することで、利用者の満足度がより高くなることがわかった。

今後は買い物タスク以外の場面においても調査して、実用的なシステムの実現を目指したい。

#### [参考文献]

- [1] Kacorri, H., Kitani, K.M., Bigham, J.P. & Asakawa, C.: People with Visual Impairment Training Personal Object Recognizers: Feasibility and Challenges, Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (2017).
- [2] Iwamura, M., Hirabayashi, N., Cheng, Z., Minatani, K. & Kise, K.: VisPhoto: Photography for People with Visual Impairment as Post-Production of Omni-Directional Camera Image, ACM CHI Conference on Human Factors in Computing Systems (CHI2020) Extended Abstracts Proceedings (2020).

#### [成果の発表, 論文等]

- [1] 河合 隆哲, 高嶋 慶伍, 岩村 雅一, 南谷 和範, 黄瀬 浩一: 視覚障害者の受動的な情報取得に向けての情報要約に関する調査, 第181回ヒューマンインタフェース学会研究会, 23, 1, pp.5-14 (Mar. 2021)
- [2] 高嶋 慶伍, 河合 隆哲, 岩村 雅一, 南谷 和範, 黄瀬 浩一: 視覚障害者の周囲の視覚情報取得における QA システムの有用性に関する調査, 第181回ヒューマンインタフェース学会研究会, 23, 1, pp.15-24 (Mar. 2021)