

視覚障害者向け周辺情報提示システム

千葉 穂乃美[†]青山学院大学[†]

社会情報学部

宮治 裕[‡]青山学院大学[‡]

社会情報学部

1. はじめに

人間は情報の8割以上を視覚から得ている [1] ため、視覚に障害のある人は得られる情報が限られている。しかし、視覚障害を抱えている人は世界では少なくとも22億人とされていて [2]、その中でも視力がまったくない人や、弱い人は自分の周辺の情報を取得することが難しい。現状の点字や音による視覚障害者への周辺情報の提示は不十分であり、情報を取得できる時間や場所が限られてしまうという問題がある。これらの問題を改善するために、視覚障害者に対して周辺の情報を音声によって、場所を選ばず個別に任意のタイミングで提示するシステムが必要だと考えた。

そこで本研究では、自分の周辺の情報を取得することが難しい視覚障害者に対して文字情報、物体情報とその方位を音声で提示するシステムを提案する。

2. 関連研究

視覚障害者の周辺の情報提供に関して、宮田ら (2016) は、全方位カメラと方位センサを利用し、認識された文字情報を音声で提示するウェアラブルインタフェースを提案しているが、インタフェースが重く長時間装着することが不快だという課題が残った [3]。

本研究では、スマートフォンを使用することでインタフェースを小型化する。

3. 周辺情報提示システム

3.1. システム概要

本システムは、全方位カメラの THETA V と眼鏡型ウェアラブルデバイスの JINS MEME を使用する。JINS MEME ではユーザーの瞬き、顔の角度、向いている方位などを取得できる。

ユーザーは THETA V が固定された帽子を被り、JINS MEME と耳を塞がないイヤホンを装着し、Android を首にかけて本システムを使用する。

システムの操作方法は、以下の通りである。

撮影

1. ユーザーが頷く
2. JINS MEME が頷きを検知する
3. アプリが THETA V に2秒後にシャッター

を切るように指示する

4. ユーザーは2秒以内に顔を前に戻す
5. THETA V のシャッターが切られる
6. 撮影はアプリの画面上のボタンからもおこなうことができる
7. 再度撮影をおこなう時は、上を向きリセットしてから動作を繰り返す必要がある

物体情報 (文字情報) の取得

1. 撮影後、ユーザーは知りたい物体情報 (文字情報) のある方位を向く
2. 顔を右 (左) に傾ける
3. JINS MEME が傾きを検知する
4. アプリからその方位にある物体情報 (文字情報) がイヤホンを通して提示される
5. アプリの画面上でも、文字で物体情報 (文字情報) を確認することができる
6. 物体情報 (文字情報) の取得はアプリの画面上のボタンからもおこなうことができる
7. 再度物体情報 (文字情報) の音声聞く時は、上を向きリセットしてから動作を繰り返す必要がある

3.2. システム構成

本システムは、撮影部、方位更新部、物体認識・文字認識部、音声提示部の4つで構成されている (図1)。

撮影部では、THETA V を用いて全方位画像の撮影をおこなう。撮影された正距円筒図法の画像は前方、後方、左方、右方の4つの画像に分割して使用する。また、分割された画像は、スマートフォンの画面上に表示される。撮影がされるたびに、表示される画像は更新される。

方位更新部では、JINS MEME を用いてユーザーの向いている方位の更新をおこなう。45度以上の変化があった場合、ユーザーの向いている方位が変わったとする。音声で提示される情報は、常にユーザーの前方にある情報であるため、方位が更新されることにより、提示される情報が変更される。また、スマートフォンの画面上の方位表示も変更される。

物体認識・文字認識部では、撮影され分割された4枚の画像にうつっている物体・文字の認識を、Python で構築されたサーバ内でおこなう。本研究では、Microsoft Azure Cognitive Services で画像を分析するために提供されている Computer Vision API を使用して、前方、

Surrounding information presentation system for the visually impaired

[†] Honomi Chiba, Aoyama Gakuin University

[‡] Yutaka Miyaji, Aoyama Gakuin University

後方, 左方, 右方それぞれの認識情報を文字で受け取る。物体情報と文字情報は, スマートフォンの画面上で文字で確認することもできる。撮影がされるたびに, 認識が実行される。

音声提示部では, 物体認識・文字認識部で取得された日本語の情報を音声に変換する。ユーザーが顔を右(左)に傾けた時点で前方にある物体情報(文字情報)を, 安全に考慮した耳を塞がないイヤホンを通して提示する。

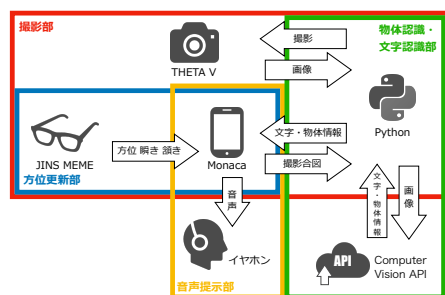


図 1: システム構成

4. システムの評価

4.1. 方法

本実験は, 構築した情報提示システムの精度や有効性を確認することが目的である。

実験1では, 対象物とそこに含まれる文字が正しく認識できるかを確認する動作検証をおこなった。屋外にある横書きの看板と縦書きの看板を11時と17時の2回, カメラの前方, 後方, 左方, 右方の4方向に位置するように撮影し, 対象物と文字が正しく認識できるかを検証した。

実験2では, 実験協力者10人に本システムを使用してもらい, アンケートを実施し, 提案インタフェースの有効性を評価した。

4.2. 結果

実験1では, 横書きの看板も縦書きの看板も全方位で11時と17時ともに, 正しく「看板」と認識した。横書きの看板では「ご」を「こ」と, 縦書きの看板では「犬」を「大」と認識してしまった。

次に実験2の結果を図2に示す。自由記述では, 精度の高さや操作のしやすさに対する意見が多かった。

4.3. 考察

実験1の結果より, 物体認識の精度は高いといえる。また, 看板内の文字は, 正しく認識できなかった文字が一文字ずつであり, 正しく認識できなかった文字は, 濁点や漢字の右肩の点が認識されていないことから, ある程度の精度で文字認識ができていているといえる。横書きの看板内の文字は, 11時にはすべて正しく認識されているのに対して, 17時には正しく認識できなかった文字

項目	平均値	分散値
質問 1 このインタフェースを長時間装着すると不快か	4.50	0.28
質問 2 このインタフェースを装着するには重いか	4.30	0.68
質問 3 このインタフェースを装着して外出するのは恥ずかしいか	3.50	1.83
質問 4 このインタフェースを装着して外出すると危険か	4.40	0.49
質問 5 このインタフェースを装着すると周りの音が聞こえなくなるか	5.00	0.00
質問 6 このインタフェースを装着すると動き辛い	4.00	1.33
質問 7 提示される物体情報は分かりやすいか	4.60	0.27
質問 8 提示される文字情報は分かりやすいか	4.40	0.27
質問 9 周辺の情報が取得できない状態の時に, 提示される物体情報は役に立つか	4.70	0.23
質問 10 周辺の情報が取得できない状態の時に, 提示される文字情報は役に立つか	4.80	0.18
質問 11 視力が低い状態の時に, アプリケーションの画面上の情報は役に立つか	4.90	0.10
質問 12 このシステムの操作は簡単か	4.50	0.28
質問 13 このシステムの操作は思い通りにできるか	4.30	0.46

図 2: 実験 2-アンケート結果

があるため, 撮影時に逆光になってしまったことが理由として考えられる。そのため, カメラにライトを取り付けることによって, どの時間帯でも精度が落ちることのないようにすることが, 改善点として挙げられる。

実験2の結果より, 全体的に良い評価を得ることができているが, 質問3の「このインタフェースを装着して外出するのは恥ずかしいか」の平均値が3.50と低いことが目立った。自由記述でも, 「ファッション感覚で身につけられる見た目になればより良いと思う」という意見があったことから, 本研究で使用した野球帽にカメラを固定したインタフェースが好みではない人がいたことが考えられる。また, 質問6の動きやすさの問いの結果からもカメラを固定する帽子の種類を変えることが, 今後の課題として挙げられる。さらに, 自由記述では, 情報が提示されるまでの時間が気になるという意見があったため, 物体認識後に文字がある場合のみ, 文字認識をおこなうことで, 速度向上を検討する必要がある。

5. おわりに

本研究では, 全方位カメラと眼鏡型ウェアラブルデバイスを用いて, 自分の周辺の情報を取得することが難しい視覚障害者に対して文字情報, 物体情報とその方位を音声で提示するシステムを提案した。

本システムの精度を検証するため屋外にある縦書きの看板と横書きの看板を対象物として選び, 実際に認識できるか実験をおこなった結果, 認識の精度が高いことが確認された。また, 実際に実験協力者に使用してもらったところ, 精度の高さと操作のしやすさに関して良い評価を得ることができ, 有効性を示すことができた。

参考文献

- [1] 教育機器編集委員会編. 産業教育機器システム便覧. 日科技連出版社, 1972.
- [2] World Health Organization. *World report on vision*. October 2009.
- [3] 宮田武嗣, 岩村雅一, 黄瀬浩一. 立体音響で教える全方位単語感知システム. 電子情報通信学会技術研究報告, Vol. 115, No. 516, pp. 179–184, March 2016.