

弱視者の視線解析と視認支援に向けた取り組み（その2）

— 視点誘導システムの提案とピクトグラム認識 —

筑波技術大学保健科学部¹⁾ 日本薬科大学薬学部²⁾ 筑波技術大学³⁾

巽 久行¹⁾ 村井保之²⁾ 関田 巖¹⁾ 宮川正弘³⁾

要旨：公共空間のバリアフリー化が進み、誰もが取得しやすい情報伝達手段として、ピクトグラムに代表される公共サインが設置されている。これらのサインは、案内や誘導、説明や規制等の大切な情報にも関わらず、弱視者に適切に伝達されているとは言い難い。本研究の目標は、弱視者が視認できていないピクトグラムに対して、どのような情報保障や視認支援を行なえるのかを考察することにある。著者らは既に、視線追跡装置を用いてピクトグラムの認識をブースティング法による識別で行なっているが、その識別の現状と弱視者への誘導支援について述べる。

キーワード：弱視, 視認支援, 視線解析, 視点誘導, ピクトグラム認識

1. はじめに

公共空間には誰もが取得しやすい情報伝達手段としてピクトグラム（絵文字、例えば非常口のマーク）に代表される公共サインが設置されている。我々は過去に弱視者の視線を解析することを目的として、視線追跡装置を用いた屋内・屋外歩行時における公共サインの視認具合を調査したが、ほとんど視認されていないという結論を得た。

本研究で考察している弱視者の視認を支援するシステムとは、視界に入っているが視認できていない対象に対して、視点先の画像を手元の端末で拡大表示する、視点の誘導を音声で行う、などといった機能を持つものである [1, 2]。我々はこれまで、視線追跡装置を用いて、このような視界や視認を訓練する映像シミュレータの研究をしてきた [3]。本報告は、この研究課題に対するこれまでの経過と成果をまとめたものである。

2. システムの概要

公共サインとは、直感的な情報伝達を目的とした公共空間に設置する掲示物を指す。ピクトグラムに代表的される視覚記号を、ユニバーサルデザインの基準に沿って設計したものが多く、種類は様々であるが、一般的に目的や設置場所に応じて、定点機能、説明機能、案内機能、誘導機能、規制機能の、5つに分類されることが多い。

視線追跡装置とは、瞳孔の動きを解析して被験者の視点を追跡できる機器である。本研究で使用したものは、図1に示す米国 ASL 社のモバイル・アイ（Mobile Eye-GX、我々の名称では Mark3）という装置で、メガネ型の

光学ユニットと視点データの記録ユニットからなる。

図2に示すように、光学ユニットには2つの小型カメラが装着されており、一つは被験者の目線で風景を撮影する CCD カメラ（シーンカメラと呼ばれる）で、もう一つは被験者の瞳孔の動きを計測する赤外線カメラ（アイカメラと呼ばれる）である。この2つのカメラが連動することにより、被験者の視界中での視点（位置）を特定することができる。計測された映像と視点は、記録ユニットの SD メモリカードに保存できるほか、Wi-Fi 規格でパソコンにも無線転送できる。



図1 視線追跡装置：モバイル・アイ

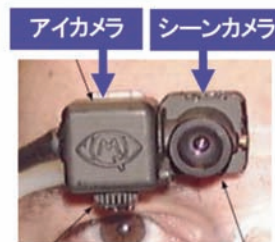


図2 光学ユニット

3. 弱視者向けの視線追跡

我々が目標とする弱視支援機器の要件は、1) 無理なく視点が捉えられること、2) 視力だけでなく視野も補えること、3) 歩行時のランドマーク（例えば、公共サイン等）の発見に効果を発揮すること、などである。残存視力を活かしながら、これらの要件を満たす技術として我々が着目したのが、眼球運動計測にもとづく視線追跡の応用である。

晴眼者に比べて、弱視者の瞳孔を認識させて校正（キャリブレーション）を行うことは難しい。我々は、弱視者に視線追跡装置を用いる場合、図3に示すように、シーンカメラで捉えた映像内に、拡大したアルファベットやランドルト環を表示して、その文字や環の空き部分が正しく読み取れば、文字の中心や環の空き部分の中心に弱視者の視点があったと判断して、これを校正位置としている。

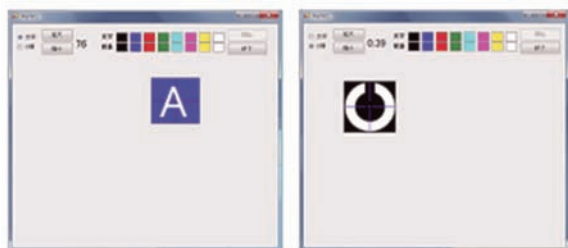


図3 弱視者の校正方法

4. 視点映像の解析と拡大

視線追跡装置に保存された弱視者の視点データは、付属ソフトウェア EyeVision により、1秒間に約30フレームの視界動画ファイルと視点情報ファイルに展開できる。視点情報ファイルは各フレームの静止画像における、視点座標、瞳孔の位置や大きさなどである。

我々は、弱視者の視認を解析するために、視点の軌跡を描画している。これは各フレーム画像に対応する視点情報ファイル内に記録された視点座標を描画するものである。ある開始フレームから一定の時間で描画することにより、最終フレーム画像内に、開始フレームからの視点が折れ線となって残る。その際、開始フレームから最終フレームまでの画像がほぼ同じであれば、被験者の視界があまり動かずに（被験者の頭や体があまり動かずに）視認が行われたことになる。大抵の場合、弱視者は視認に時間がかかるので、視認状態に入ったときは晴眼者と比べて、比較的動作の少ない視認結果が得られた。

我々が目標とする弱視支援機器の機能の一つに、視認できない対象に対して、その画像を“拡大表示する”というものがある。そのためには弱視者が視線を追う対象を検知することが必要である。我々は、視線対象の検知は視点の滞留時間で捕えられると考え、滞留時間を描画するプログラムを作成した。これは視認領域の時間密度をフレ-

ム毎に積算して、視点の滞留具合を描画する。

図4に、図書館内で受付を探しているときの弱視者の視点軌跡と、視点滞留時間の描画を示す（視点が滞留しているほど赤色が濃い）。このような視点密度をもとに、対象画像の切り出しと拡大表示を行なった例を、図5に示す。



図4 視点軌跡と視点滞留時間の描画

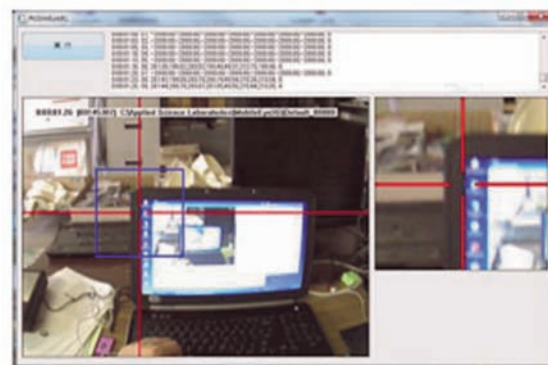


図5 視点先の映像を拡大提示

同図において、左の画像がシーンカメラのリアルタイム画像であり、滞留が認められた最終フレームの視点位置（赤のクロスライン）の辺りを拡大表示している（この例では、青の四角を2倍に拡大した）。滞留時間が長い視認映像が手元のモバイル端末で確認できることで、これまで目で追ったが視認できなかった対象も、弱視者自身の残存視力で、直ちにその場で理解できる。

5. 視点誘導に向けて

視点誘導は、視力障がいよりも視野障がいに有効な支援である。視野障がいのある弱視者は、対象を見つけることが難しいので、代わりに支援機器が検出する。

近年、パターンマッチング法の一つに、検出力の弱い分類器を多く組み合わせることで、全体として強い分類器となるような、ブースティング（Boosting）と呼ばれる分類法がある[4]。本研究では、各分類器の学習に Haar-like 特徴（Haar-like Feature）と呼ばれる、水平・垂直方向にスケー-

リングしたパターン特徴を使用している [5]。最近、OpenCV (Open Source Computer Vision Library) と呼ばれる画像処理ライブラリが公開されており [6]、その中のブースティング・ルーチンを用いて、比較的簡単に顔検出や物体検出のプログラムを作成することができる。

図 6 は、正しい非常口のピクトグラムを正解画像として約 1～2 時間の学習をした後に、体育館内と通路の画像から、非常口のピクトグラムを検出した結果である (図中、検出結果が白丸で表示されている)。学習時間が少ないために非常口でない箇所も検出されるが、これらは適当な抽出 (例えば、エッジ抽出や輪郭抽出など) を使用した真偽判定で除外することができる。



図 6 Boosting 法による検出例

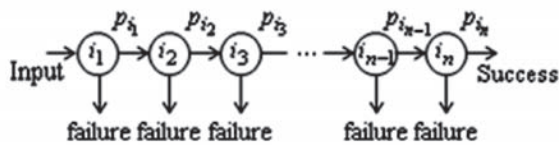


図 7 分類器の典型的な構造

ここで、分類器の最適配置問題を考察する。そのために、図 7 に示すような典型的な構造を用いて議論する。ここで、目標となる対象図形 (例えば、非常口サイン) には、 n 個の特徴 f_i (但し、 $i=1, \dots, n$) を持っているとして仮定し、図中のノード i_j は特徴 f_{ij} の検査を行なうものとする。

入力 (Input) が特徴 f_{ij} を持つならば、処理は次のノードに進み、そうでなければ失敗 (failure) である。従って、入力がすべての特徴を満足するときのみ、分類器は成功 (Success) と判定して、対象と認識する。以下、特徴を検査するノードをフィルタと呼ぶことにする。ここで、各 f_i は肯定枝に対して、平均処理時間 t_i および平均確率 p_i (但し、 $0 \leq p_i \leq 1$) を持つ (ここで、 t_i および p_i は入力データに依存するので、この定義はかなり単純化したものである)。

以下に、この条件のもとに分類器が最少の処理時間となるような、フィルタの最適順序が得られることを示す。すなわち、分類器の予想処理時間 (もしくは、半決定的連鎖のコスト) は、次のようになる。

$$T_a = t_{i_1} + p_{i_1} \cdot (t_{i_2} + p_{i_2} \cdot (t_{i_3} + \dots + (t_{i_{n-1}} + p_{i_{n-1}} \cdot t_{i_n}) \dots)) \quad (1)$$

ここで a は、系列 i_1, \dots, i_n を表すこととする。

今、 a' を、 a のある j (但し、 $1 \leq j \leq n$) において、隣同士の i_j と i_{j+1} とを交換することによって、 a から得られる系列とする。以下、2 つの系列 a と a' を、隣接系列と呼ぶことにする。以上の準備より、次の式が成り立つ。図 7 において、隣接するフィルタを交換する、すなわち、2 つの隣接系列の処理時間に関して、 $T_a \leq T_{a'}$ ならば、そのとき、

$$\frac{t_{i_j}}{1-p_{i_j}} \leq \frac{t_{i_{j+1}}}{1-p_{i_{j+1}}} \quad (2)$$

である。逆に、式 (2) が 2 つの隣接系列の間で成り立つならば、 $T_a \leq T_{a'}$ となる。

従って、 $t_i / (1-p_i)$ を q_i とすると、 q_i の昇順にフィルタを並べ替えることができる。すべての i (但し、 $1 \leq i \leq n$) に対して等価コスト系列の違いを除けば、この並べ替えは一意に決定され、それが最適な処理である。

実際の最適化を考える場合、 t_i および p_i に対する統計データを得ることは困難なので、昇順で t_i を並べ替えて、初めの少数の j に対して、式 (2) を検査することが有効である。但し、この議論では p_{ij} と t_{ij} の条件付振る舞いを完全に無視しているため、厳密には p_{ij} の代わりに条件付確率 $p(i_1, \dots, i_{j-1} | i_j)$ (但し、 $j=2, \dots, n$) を用いなければならない。同じことは t_{ij} についても言える。ここで最適な半決定的連鎖を構成するアルゴリズムは、動的計画法の原理 (Bellman の最適性原理) に従うことが知られている。このため、最適な系列を得るには $f_i=0,1$ に対する 2^n 個の項目を持つ条件付確率表 $P(f_1, \dots, f_n)$ を必要とするが [7]、実際にそのようなデータを得ることは簡単ではない。

上に述べた議論は、一般に分類器の構成方法において、有効に分類できる処理を先にするほど、より最適な手続系列が得られることを示している。

ピクトグラムにおいて、目的地や次の定点方向を示す誘導機能や、禁止・指示・警告などを示す規制機能に、色がつくことが多い。例えば、誘導機能のサインである“非常口”には緑色が、規制機能のサインである“立入禁止”には赤色が、規制機能のサインである警告等には黄色が、一般に多く使われている。そこで、分類器列の始めの方で色調による分類を行なえば、ピクトグラムの機能による類別が行なえて、より高速な認識が可能になると考えられる [8]。

画像上で同じ色の画素数が多い領域は、色調空間内では高密度クラスタとなるので、色分類はクラスタ検出になる。一般に、分離可能性 (認識率) の難易度は、データ数に比例し、特徴空間の次元数に反比例する。本研究では、類似画像のクラスタ検出として、認識率よりも計算時間を優先したため、分類器として最近傍法を、特徴量として単純で良く用いられているカラーヒストグラムを採用した。

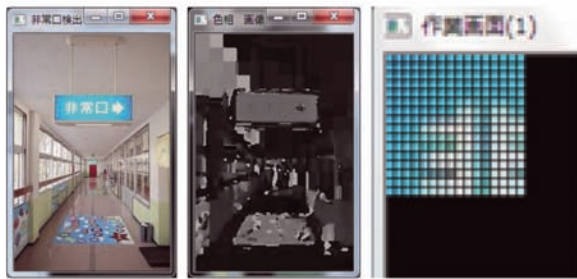


図8 色調による検出

図8に、色調による検出結果を示す。同図において、左は原画像、中央は色調による処理、右は誘導機能のサインに多くある緑色の色調を用いて、非常口画像の一部を切り出して拡大したものである。

また、パターンマッチング法の一つに、特徴点の検出と特徴量の記述を行うシフト法（SIFT: Scale-Invariant Feature Transform）がある。シフト法は、米国オレゴン州立大学でソースコードが公開されているが[9]、実用的な処理速度を得るためにはデータベースを作成する必要がある。本研究では、シフト法とは異なる局所特徴量アルゴリズムで、米国 Evolution Robotics 社の認識エンジンである ViPR（Visual Pattern Recognition）を使用している。この ViPR のアルゴリズムは、対象画像から得られた特徴量とモデルに予め登録してある特徴量との比較を行う際に、変化量をベクトルに変換した特徴点として、最近傍探索を高速に行なうことができる。図9に、ViPR 法による検出例を示す。

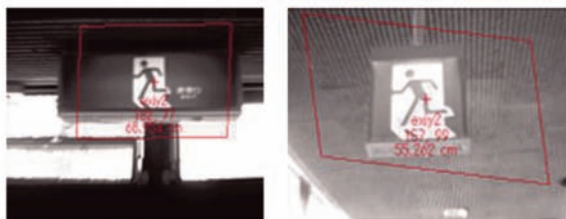


図9 ViPR 法による検出例

6. おわりに

本研究の目的は、弱視者の情報保障を向上させることであり、視界や視認を訓練する映像シミュレータとして標識や信号、各種の公共サインなどを、弱視者自身が行うことができる機器を開発することである。我々が考えている視認支援システムは、視力障がいや原因で視認できない対象に対して、視点先の映像を拡大提示すること、視野障がいや原因で視認できない対象に対して、視点先への視線の誘導を行うこと、などといった機能を持つものである。これらの処理のために、システムでは視界画像内でパターン認識を行うが、これを解決するために、1) ブースティン

グ法やシフト法に基づく汎用性の高い識別アルゴリズムを用いて、2) 学習機能を付加して誤識別を除外することを試みながら、識別能力を徐々に高めることを行なっている。

弱視者の視認を支援する補償機器を開発するには、弱視者の視点を解析することが必要であり、逆に開発した過程から、検出しやすい公共サインの設置や図柄も提言したいと考えている。また、実用的な視認支援システムを構築するには、公共サイン自身が情報を発信するような研究も必要である。

謝辞

本研究は平成23年度科学研究費補助金（基盤研究(B), 21300079: “公共サインを目印とした弱視の歩行訓練映像シミュレータの開発”) の助成を受けて行われた。ここに記して謝意を表する。

参考文献

- [1] H. Tatsumi, Y. Murai, M. Kawahara, I. Sekita, M. Miyakawa : “Eye Tracking Application for Low Vision, --A Proposal of Vision Navigator--” , Proc. 2010 IEEE Int. Conf. on Systems, Man and Cybernetics, No.492, pp.1578-1583, Oct. 2010.
- [2] Y. Murai, M. Kawahara, H. Tatsumi, I. Sekita, M. Miyakawa : “Eye Tracking for Low Vision Aids, --Toward Guiding of Gaze--” , Springer, Lecture Notes in Computer Science (LNCS), Vol.6180, pp.308-315, July 2010.
- [3] 村井, 河原, 巽, 関田, 宮川: “視線解析による弱視者の視認支援への取り組み”, FIT2010 (第9回情報科学技術フォーラム), K-055, pp.743-744, 2010.
- [4] P. Viola, M. Jones : “Rapid Object Detection using a Boosted Cascade of Simple Features” , Proc. 2001 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, vol.1, pp.511-518, 2001.
- [5] R. Lienhart, J. Maydt : “An Extended Set of Haar-like Features for Rapid Object Detection” , IEEE Proc. Int. Conf. Image Processing, Vol.1, pp.900-903, 2002.
- [6] <http://opencv.jp/>
- [7] 宮川: “半確定鎖の最適化問題”, 昭和59年度電子情報通信学会全国大会講演論文集, No.1502, 1984.
- [8] 村井, 巽, 関田, 宮川: “色相を用いたピクトグラムの認識と誘導支援”, FIT2011 (第10回情報科学技術フォーラム), K-060, pp.855-856, 2011.
- [9] <http://blogs.oregonstate.edu/hess/>

Eye Tracking Analysis and Visual Aids for Low Vision (Part 2) — A proposal for Vision Navigator and Pictogram Recognition —

TATSUMI Hisayuki¹⁾, MURAI Yasuyuki²⁾, SEKITA Iwao¹⁾, MIYAKAWA Masahiro³⁾,

¹⁾Faculty of Health Sciences, Tsukuba University of Technology

²⁾School of Pharmacy, Nihon Pharmaceutical University

³⁾Tsukuba University of Technology

Abstract: Public signs are created for the convenience of individuals. Often, pictograms are used rather than letters to catch individuals' attention (e.g., for guidance, notification, and other purposes). We experimentally examined the abilities of individuals with low vision to find pictograms by the use of an eye-tracking technique. We discovered that individuals with low vision rarely notice pictograms. The purpose of our project was to conduct a visibility study. We hoped to analyze low vision by the use of eye-tracking analysis. In our paper, we present the preliminary results of a test of pictogram recognition based on pattern matching. We hope to link these results to the development of a future system that might be used to guide an individual's gaze to missing pictograms. The individual might then be able to zoom to the image.

Keywords: Low vision, Visual aids, Eye tracking, Vision navigator, Pictogram recognition