

弱視者の文字読み取りのためのウェアラブル視力増強システム

Wearable super-vision system for low-vision aids in reading



河野 英昭 (Hideaki KAWANO Dr. Eng.)

九州工業大学大学院 工学研究院 電気電子工学研究系
准教授

(Associate Professor, Department of Electrical Engineering and
Electronics, Faculty of Engineering, Kyushu Institute of
Technology)

日本知能情報ファジィ学会 バイオメディカルファジィシステム学会

日本医用画像工学会 電子情報通信学会 情報処理学会 IEEE

著書: 基礎C言語プログラミング, 共立出版, 2012

研究専門分野: ソフトコンピューティング 生体補償制御学 医用画像
工学

あらまし 本研究では、弱視者などの視認活動の際の困難を解決することを目的として、目表情の認識を用いたユーザーインターフェースを提案する。ウェアラブルカメラを使用し、視認困難な際に目元に表出することが多い「目を細める」表情をカメラで認識し、その状態に応じてもう一台のカメラから得られる視野画像を拡大・縮小して、ユーザが装着したヘッドマウントディスプレイに提示する。本研究で認識対象とする「目を細める」行為は、目を半開きにすることをいう。この動作の意図は、像がぼやけて見える際に光量を絞りピント調節を行うことである。特に近視や老眼（遠視）の初期症状として無意識に行っていることが多い。システムが「見づらい」というユーザの意図を画像上で検出することで、必要な時に自動で視野画像をユーザにとって見やすく拡大して表示することが可能となる。本システムは弱視者の中でも、特に近眼や老眼といったピントが合わない症状の人に有効であると考えられる。

1. 研究の目的

近年、老眼や白内障のような視覚障害を含む弱視者は、高齢者を中心に多く存在し高齢化に伴いその数は増加している。その弱視者は、階段の昇降、路線図など案内物の読み取りや道路の歩行など様々な場面で困難が想定される。これらの困難に対して、それぞれの事象に応じた支援を行う必要がある。従来から弱視者支援を目的として、メガネやコンタクトレンズ、拡大鏡などの視覚補助機器が存在する。また、近年ではユーザに搭載した小型カメラから得られた映像を提示する電子メガネの開発が進んでいる。これらの装置の問題として、システムがユーザの「見づらさ」を考慮していないため、見たい映像を得るために手動による操作が必要となるため、使い勝手の向上が課題となっている。

本研究では、弱視者などの視認活動の際の困難を解決することを目的とする。そこで、目表情の認識を用いたユーザーインターフェースを提案する。ウェアラブルカメラを使用し、視認困難な際に目元に表出することが多い「目を細める」表情をカメラで認識し、その状態に応じてもう一台のカメラから得られる視野画像を拡大・縮小してユーザが装着したヘッドマウントディスプレイに提示する。本研究が認識対象とする「目を細める」という行為は、目を半開きにすることをいう。この動作の意図は、像がぼやけて見える際に、光量を絞りピント調節を行うことである。特に近視や老眼（遠視）の初期症状として無意識に行っていることが多い。システムが「見づらい」というユーザの意図を画像上で検出することで、必要な時に自動で視野画像をユーザにとって見やすく拡大して表示することが可能となる。本システムは、弱視者の中でも、特に近眼や老眼といったピントが合わない症状の人に有効であると考えられる。

2. 関連研究

近年、インターフェースの小型化や軽量化に伴い、ウェアラブルコンピューティングに関する研究や実用化への取り組みが盛んに行われている。ウェアラブルコンピューティングとは、コンピュータを常時装着し

弱視者の文字読み取りのためのウェアラブル視力増強システム

Wearable super-vision system for low-vision aids in reading

て利用するスタイルのことをいう。

ウェアラブルシステムを用いた弱視者や視覚障害者の支援を行う研究開発も進んでいる。例えば、Bryantら[1]は、赤外光投影できるビデオカメラを利用して近接した障害物を検出して、その位置をファイバースキャニング網膜ディスプレイに提示するウェアラブルシステムを開発している。Alvaroら[2]は、周囲の危険状況を超音波センサで読み取って、危険性を音声や振動で通知するシステムを提案している。

また、屈折異常を補正する器具として、古くから眼鏡、コンタクトレンズ、拡大鏡などが存在する。近年、小型カメラ付き電子メガネ[3]が開発されている。これは、ユーザに搭載された小型カメラから得られた映像を網膜に直接映像光を投射して、視野映像を提示する装置である。この装置の問題点としては、システムがユーザの「見づらさ」を考慮していないため、見たい映像を得るために手動による操作が必要となるため、

使い勝手の向上が課題となっている。また、ディスプレイ提示型のヘッドマウントディスプレイよりも鮮明に見ることが可能であるが、光路中の障害物がすべて結像するため、目の中のゴミやまつ毛が映ることが課題である。さらに瞳孔や眼球運動により、映像の欠損が生じることも改善が必要である。

3. 提案システム

本研究の目的である「目を細める」という目表情の認識を用いた視野画像の拡大システムについて述べる。提案手法の手順を図1に示す。本システムを起動する前に、ユーザの目表情を学習しておく「目表情の学習」、システム起動後目元画像を入力して認識を行う「目表情の認識」、「目を細める」と判定した後、視野画像を拡大・縮小する「視野画像の提示」という3つのパートに分けることができる。はじめに本システムの構成を述べた後、それぞれの手法を説明する。

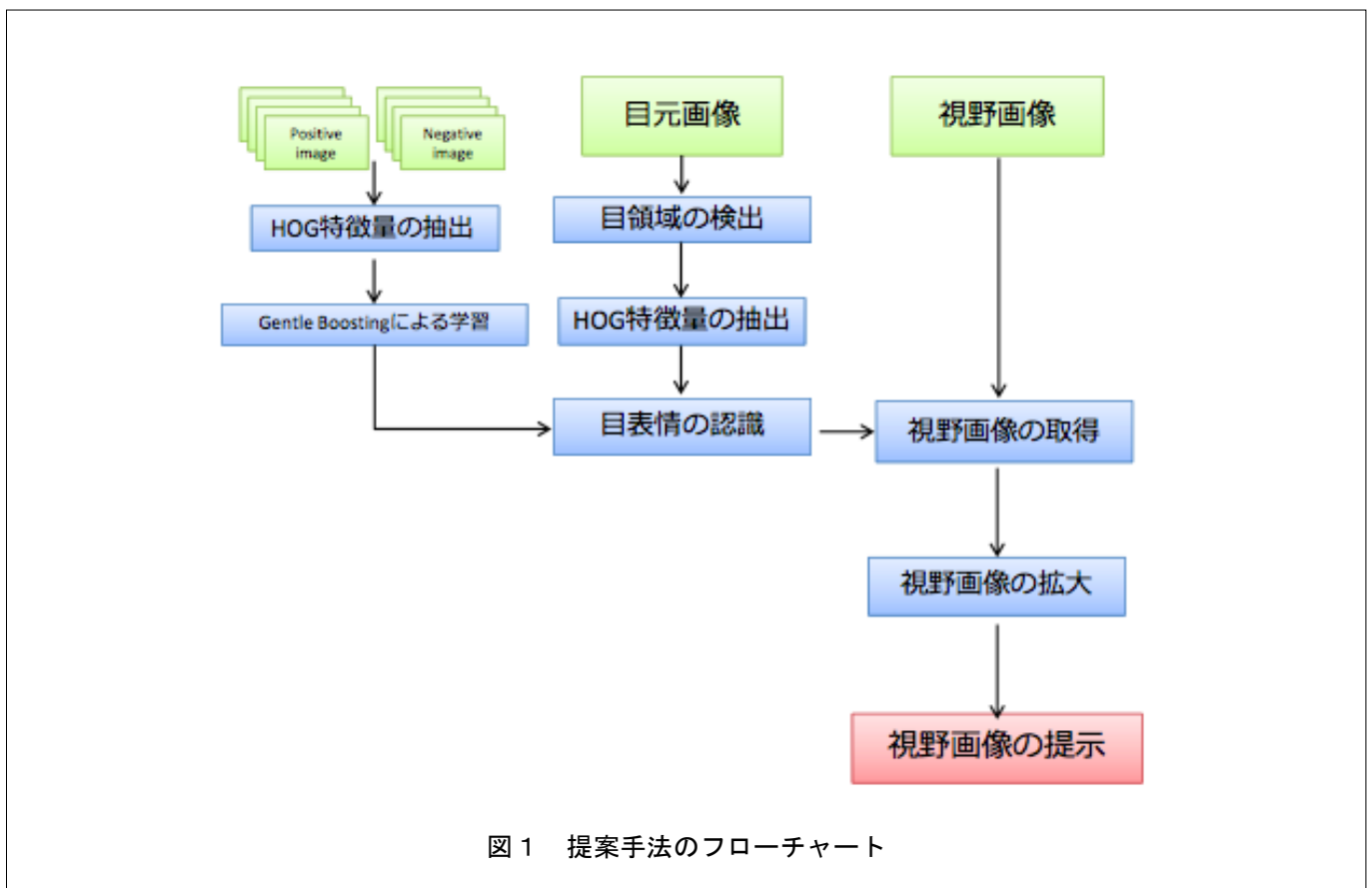


図1 提案手法のフローチャート

弱視者の文字読み取りのためのウェアラブル視力増強システム

Wearable super-vision system for low-vision aids in reading

3. 1 システムの構成

本研究において開発したウェアラブルシステムの外観を図 2 に示す。また、その構成要素は以下のとおりである。

- 目表情取得用カメラ
ユーザが視認困難な際に無意識に行う「目を細める表情」の検出を行うために、片目部分を捉えることができる小型カメラを使用する（カメラ 1）。
- 視野画像取得用カメラ
ユーザが見ている前方の視野画像を得るために、小型カメラを使用する（カメラ 2）。
- ヘッドマウントディスプレイ
拡大処理された視野画像をユーザに提示するために使用する。単眼式ヘッドマウントディスプレイを使用する。



図 2 開発したウェアラブルシステムの外観

3. 2 目表情の学習

本システムを使用するために、事前に各ユーザの目表情を学習しておく必要がある。本研究では、事前知識なしでも識別器を構成することが可能な Boosting 学習を用いて行う。以下に手順を示した後、詳細を説明する。

(1) 学習画像の準備

ウェアラブルシステムを装着し、カメラ 1 よりユーザの片目部分の目表情を撮影する。ユーザの目の部分がすべて写っていることが条件である。撮影された画像列から、目領域のみを切り出しておく。2 クラスの判別を行うために、正例には「目を細める」状態、負例には「それ以外」の状態の画像を準備する。入力画像列は、解像度 320×240 [pixel]、30 [frame/sec] で撮影されたものを使用する。切り出された学習画像は、すべて 260×153 [pixel] となる。

(2) HOG 特徴量による抽出

それぞれのクラスの学習画像に対して、画像の特徴抽出を行う。本研究で用いる HOG 特徴量 (Histograms of Oriented Gradient Feature) [5] は、輝度方向をヒストグラム化したものである。顔認識や一般物体認識などの分野で広く利用されている。画像サイズを 128×64 [pixel] にリサイズして行う。セルのサイズは 8×8 [pixel] で行う。この処理により、6,804 の特徴ベクトルが抽出される。

(3) Gentle Boosting による学習

HOG 特徴量によって抽出された特徴ベクトルに基づき、GentleBoosting [6] を用いた学習を行う。ブースティング学習は、重みを更新しながら学習画像を検出するのに適した特徴ベクトルを弱識別器として選択する。そして、弱識別器からより高精度な強識別器を構成していくアルゴリズムである。学習回数を 10 回とする。従来の AdaBoost より外れ値に対する頑健性が強い。

3. 3 目表情の認識

「目表情の学習」によって学習された結果に基づいて、本システムを起動して目表情の認識を行う。手順は以下ようになる。

(1) 目元画像の入力

ユーザがウェアラブルカメラシステムを装着し、本システムを実行させると、カメラ 1 よりユーザの片目部分が撮影された目元画像を入力する。解像度

弱視者の文字読み取りのためのウェアラブル視力増強システム

Wearable super-vision system for low-vision aids in reading

320×240 [pixel]、30 [frame/sec]である。

(2) 目領域の検出

入力された目元画像をそのまま認識に使用すると、装着時のずれの影響で正しく認識が行うことができない。前処理として、入力画像に対してテンプレートマッチングを用いて目領域の切り出しを行う。学習画像に用いられている画像をテンプレート画像として使用する。目元画像に対して、最も類似度が高い領域を推定し切り出す。本研究では、正規化相互相関関数を用いる。この値が1に近いほど類似度が高い。類似度 R は以下ようになる。

$$R_{NCC} = \frac{\sum_{j=0}^{N-1} \sum_{i=0}^{M-1} I(i,j)T(i,j)}{\sqrt{\sum_{j=0}^{N-1} \sum_{i=0}^{M-1} I(i,j)^2 \times \sum_{j=0}^{N-1} \sum_{i=0}^{M-1} T(i,j)^2}}$$

(3) HOG 特徴量による抽出

切り出された目元画像に対して、HOG 特徴量による抽出を行う。

(4) 目表情の認識

抽出した結果より正例か負例かを識別する。

3. 4 視野画像の提示

認識結果より、「目を細める」と判定された場合に、視野画像の拡大を行い提示する。以下に詳細を説明する。

(1) 視野画像の入力

カメラ 2 より得られた視野画像を入力する。解像度は 320×240 [pixel]である。

(2) 視野画像の拡大・提示

1 フレームごとの目表情の認識結果を出力として、そのまま利用するのではなく、認識結果を時系列的に評価してから行う。理由として、1 フレームの結果をそのまま利用すると、拡大・縮小がバタつき、ユーザの使用感が低下する。今回は、現在のフレームとそれより前の合計 10 フレーム分の出力値の平均をとる。その際、正例を+1、負例を 0 とする。その値が 0.5 以

下の場合、視野画像をそのままのスケールで提示し、それ以上の時は拡大処理を開始する。

ユーザが目を細めるということは、「対象物体が画像中心にある」と定義する。拡大方向は視野画像の中心方向に一定のスピードで行い、ヘッドマウントディスプレイに随時提示される。目を細める表情を止めた時（しきい値が 0.5 未満）は、「見えている」と判断して、拡大処理を停止する。よって、ユーザの任意の倍率に拡大することができる。一定時間後、元のスケールにズームアウトする。

4. 実験

4. 1 実験環境

本実験で使用した計算機の性能は、下表のとおりである。

表 1: 実験計算機の性能

型名	Dell Vostro430
CPU	Intel Core i7 2.93GHz
Memory	4GB
OS	Windows 7 Professional
プログラミング環境	Microsoft Visual Studio 2008 C++
画像ライブラリ	OpenCV 2.1

4. 2 実験 1

提案手法で述べたシステム全体を実装し、正しく目表情を認識し視野画像の拡大操作を行うことができたかを検証した。以下に詳細を述べる。

4. 2. 1 実験方法

1 人の被験者に対して実施した。被験者は晴眼であった。場所は室内で行った。学習画像については、予備実験の際に取得された学習画像データを使用した。被験者に対して、ウェアラブルカメラを装着してもらい、本システムを実行した。撮影条件として、カメラ 1 から入力される画像に、目がはみ出さないようにすることを条件とした。被験者は、正面前方にある物体に対して意図的に目を細める表情と、通常の表情を数秒毎に繰り返し行った。今回は、視野画像をヘッドマウントディスプレイに提示せず、それぞれの処理が正

弱視者の文字読み取りのためのウェアラブル視力増強システム

Wearable super-vision system for low-vision aids in reading

しく実行できているかを PC のモニタ上に表示し確認を行った。また、リアルタイム性を検証するために、処理時間を計測した。

4. 2. 2 実験結果

PC のモニタ上に出力した結果は、図 3 のようになった。上段左がウェアラブルカメラから入力された目表情画像、上段右がテンプレートマッチング後に切り出された画像、下段左が入力された視野画像、下段右が拡大処理された視野画像となっている。

図 4 は、視野画像が拡大して、元の倍率に縮小される様子を示している。左側が同時刻の目表情である。

このように、目を細める表情に合わせて、視野画像を拡大できていることを確認した。

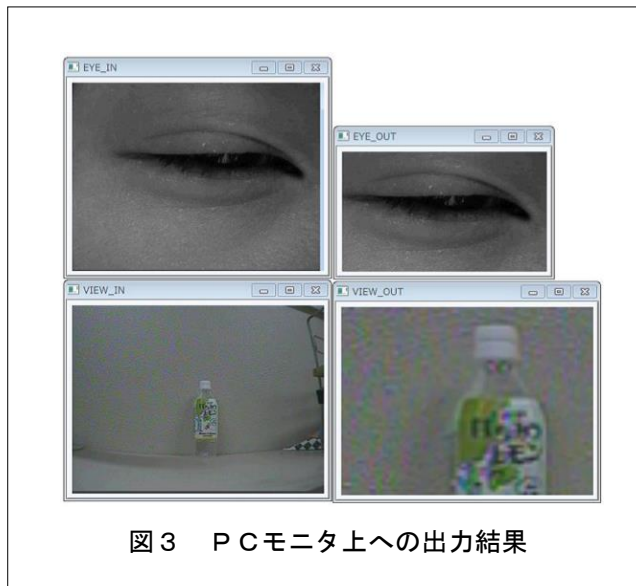


図 3 PCモニタ上への出力結果

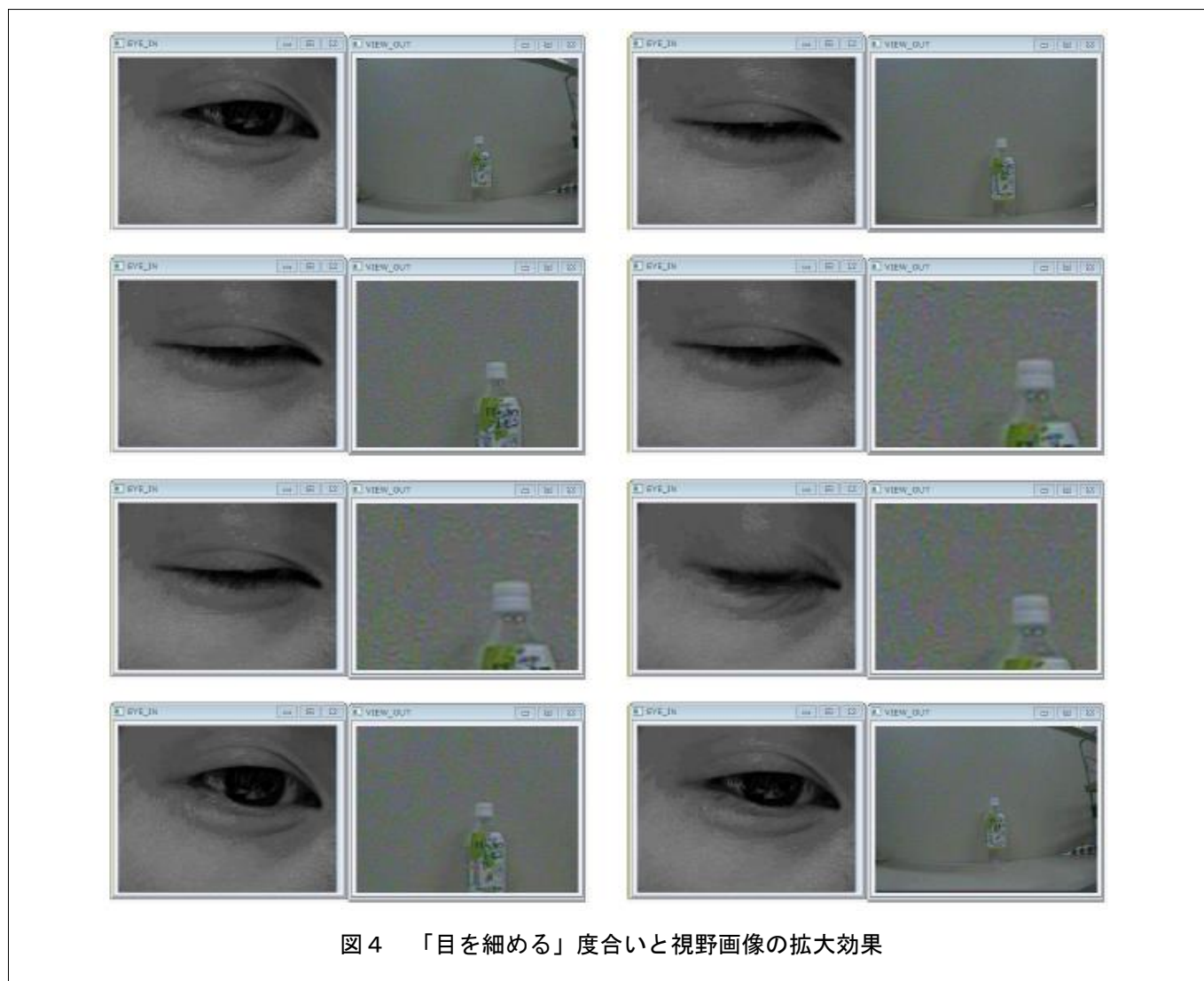


図 4 「目を細める」度合いと視野画像の拡大効果

弱視者の文字読み取りのためのウェアラブル視力増強システム

Wearable super-vision system for low-vision aids in reading

処理時間については、学習画像を Gentle Boosting によって学習するまでの計算時間が 60 秒程度であった。学習については、ユーザが実際にシステムを使用する前に実行することを想定しているため、特に問題はない。また、カメラ画像を入力して、1 フレームあたりの処理時間は平均で 30~40 [ms]であった。この処理時間は実用する上では問題ない。

本研究では、画像中心方向に拡大を行った。視認困難で目を細める時は、注視している対象が中心部にあることに基づいて行った。しかし、ユーザが見ている対象がどこなのかは特定していないため、意図しないものを拡大する可能性がある。このことは今後の課題である。

4. 3 実験 2

本実験では、被験者から約 50cm 程度前方の壁に掛けてある「時計」を視認したい対象とした。実験 1 と同様に、被験者は意図的に目を細める動作を繰り返し、システム動作を確認した。その他の実験条件についても、実験 1 と同様とした。拡大処理前と後の視野画像に対して、ガウシアンフィルタを施して弱視者の見え方（ピントが合わない場合）を再現した。図 5(a)と(b)はフィルタを施す前、図 5(c)と(d)はフィルタをかけた後の画像である。拡大することで、時計の文字や針から現在の時刻を確認することができた。



(a) 晴眼者の視野画像



(b) 晴眼者視野画像の拡大結果



(c) 弱視者の視野画像



(d) 弱視者視野画像の拡大結果

図 5 文字読み取りに対する拡大処理の効果

弱視者の文字読み取りのためのウェアラブル視力増強システム

Wearable super-vision system for low-vision aids in reading

5. 将来展望

本研究では、弱視者支援のためのウェアラブルカメラを用いた映像拡大インターフェースを提案した。システム全体を実装し、目表情に応じてリアルタイムで視野画像の拡大操作を行うことができることを確認した。

今後の課題については、以下のことが考えられる。今回、晴眼者に対して意図的に目を細める表情をしてもらい、動作を確認した。今後は、実際の弱視者による評価が必要となる。被験者を募って、本ウェアラブルシステムの使用感について調査・検討を行う必要がある。また、使用したカメラの制約により、計算機としてデスクトップ PC を用いた。実際の使用では、電源供給や重量の面から問題となる。実用化するためには、計算機やウェアラブルシステムの小型化・軽量化が必須である。

この研究の発展性として、弱視者の視認困難な場所の収集や困難状況の解析に用いることができると考えられる。例えば、本システムと現在地情報を記録・同期できる GPS ロガーや角速度を検出できるジャイロセンサーを利用することで、視認困難である「目を細める」時の視野画像を取得し、現在地情報を付加することが可能となる。そして、弱視者の不便な箇所を自動で地図上にマッピングすることができる。今後は、具体的な手法の検討を行う必要がある。

用語解説

弱視者：弱視者の定義は様々な見方があるが、世界保健機関(WHO)の定義では、「両眼の矯正視力が 0.05 から 0.3 の者、または視力以外の視機能障害があり、学習や日常生活に制約があるが、主として視覚における様々な行動ができる者」と言われている。「矯正視力が 0.05 から 0.3」とは、メガネやコンタクトレンズなどの矯正器具を用いても、視力が十分に出ない状態のことをいう。また、両眼の視力が 0.05 未満である場合を盲(blindness)と呼ばれる。日本眼科医会によると、弱視者を含む視覚障害者は、2007 年現在、約 164 万人に上ると推定されている[6]。そのうち、矯正視力が 0.1~0.5 である弱視者の人数は、144 万 9,000 人いる

と推定されている。また、高齢化に伴って、老眼や白内障の視覚障害の人数が増加している。今後も、増加が予想され、2030 年頃には 200 万人を超えると予想される。

参考文献、関連文献

- [1] R.C. Bryant, C.M. Lee, R.A. Burstein, and E.J. Seibel. "Engineering a low-cost wearable low vision aid based on retinal light scanning," In Proc. SID 2004, pp. 23-28, 2004.
- [2] A. Cassinelli, C. Reynolds, and M. Ishikawa, "Augmenting spatial awareness with haptic radar," 10th IEEE International Conference on Wearable Computers (ISWC 2006), pp.100-107, 2006.
- [3] ウェアビジョン URL: <http://wearvision.co.jp/index.html>
- [4] N. Dalal, B. Triggs, "Histograms of oriented gradients for human detection," Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2005.
- [5] J. Friedman, T. Hastie, and R. Tibshirani, "Additive logistic regression: a statistical view of boosting", The Annals. of Statistics, 28(2), 337-374, April 2000.
- [6] キャリアブレイン、「医療介護 CB ニュース」
<http://www.cabrain.net/news/article/newsId/24377.html> (2012.2.10 アクセス)

この研究は、平成 21 年度 S C A T 研究助成の対象として採用され、平成 22~24 年度に実施されたものです。