

口の動きを併用した意思伝達用視線入力システム

函館工業高等専門学校 ○藤田 智也, 浜 克己, 鈴木 学, 中村 尚彦

要 旨

筋萎縮性側索硬化症(ALS)などの神経難病によってコミュニケーションに問題を抱える人の意思伝達機能の向上を目的とし、本研究では選択画面における視線入力のみでの誤操作の改善, クリック法での注視や瞬き, ウィンクによる目の筋肉の疲労軽減, 円滑な意思伝達を図るための入力速度向上を目指して, カメラ画像による口の動きを併用した視線入力システムの拡張について報告する。

1. 緒言

筋萎縮性側索硬化症(ALS)をはじめとする神経難病においては, 筋力低下の進行が発声を困難にさせ, 意思伝達手段としてのコミュニケーションにも大きな支障をもたらす症例が少なくない。また, 病状の進行に伴い, スイッチ操作も困難となる場合が多いことから, 身体状況に合わせたさまざまな方式が検討されている。視線入力システムの利点は, マウス操作に比べて上半身が疲れにくく, 直感的な入力ができることである。しかし, クリック法での注視や瞬き, ウィンクでは, 誤操作が起こる, 選択に時間が掛かる, 目を意識的に動かすため疲労を感じやすいなどの問題があり, 改善の余地があると考えられる。

そこで本研究では, 作業効率や身体的使いやすさを考慮し, 視線のみの入力に対して口の動きを併用することで, 操作性の改善, 疲労軽減とともに, 画面表示方法の改善による誤操作の軽減を目指す。

本稿では, 作成予定の選択画面の概要と, 口を併用した視線入力システムの構想について報告する。

2. システム構成

2.1 使用機器

視線入力装置には, Tobii Technology 社製の Tobii Eye Tracker 4C を使用する。その外観を Fig.1 に示す。赤外線のマイクロプロジェクト, 光学センサ, 画像処理が一つになった装置である。

口の動作は画像処理を用いて解析するが, そのためのカメラとして, Logicool 社製の HD Webcam C270 を使用する。その外観を Fig.2 に示す。解像度 1280x720 ピクセルで, リアルタイムに映像を出力できる装置である。

2.2 全体構成

全体のシステム構成を Fig.3 に示す。視線の計測には, 角膜反射法を用いている。使用者の目に近赤外線を照射し, 角膜反射点を生成した後, 目の映像を記録する。その目の映像から反射点と瞳孔位置を検出し, これらの反射パターンから注視点を算出している。計測環境下の光の条件や目の特徴を考慮した測定方法にカスタマイズし, より正確に注視点を算出するためには, 使用前のキャリブレーションが重要となる [1]。

口の形状の取得は, Web カメラからの映像を基に, 画像処理ライブラリの OpenCV を用いて行う。詳細については後述する。



Fig.1 Tobii Eye Tracker 4C



Fig.2 HD Webcam C270

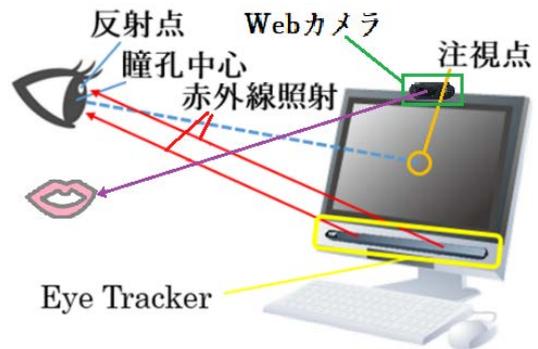


Fig.3 システム構成

3. 視線入力システムの概要

3.1 視線位置の取得

視線位置を画面に表示させる場合, 注視点を直接表示させると固視微動による誤操作を誘発してしまう可能性が高い。そこで, 視線の移動にスムージングを施し, 視線のブレを軽減することで誤操作の発生を抑制する。そのため, 注視点の検出周期は 1/60 秒とし, 連続する 10 点の移動平均値を用いることとした。注視点位置を (x_i, y_i) として, 視線位置 (x'_i, y'_i) の計算には(1)式を用いる。

$$(x'_i, y'_i) = \frac{1}{10} \left(\sum_{j=0}^9 x_{i-j}, \sum_{j=0}^9 y_{i-j} \right) \cdots (1)$$

3.2 選択画面イメージ

新たに選択画面を作成するにあたって, 先行研究にて作

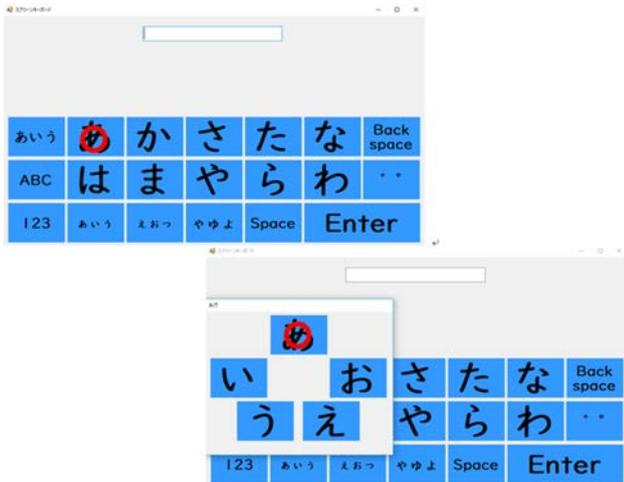


Fig. 4 先行研究の選択画面

Table. 1 先行研究の評価実験結果

	50音表示		二段階表示		50音表示 使いやすさ	二段階表示 使いやすさ
	誤操作回数/正確率	誤操作回数/正確率	誤操作回数/正確率	誤操作回数/正確率		
A	24回	46.7%	5回	83.9%	0人	3人
B	16回	56.8%	3回	89.7%		
C	23回	47.7%	5回	83.9%		

成された選択画面と、誤操作の観点から行われた評価実験の結果をそれぞれ Fig4, Table. 1 に示す. 具体的には、健常者 3 名の被験者に、“だい 19 かいけいそくじどうせいぎよがっかい”という文字列を注視 1 秒で入力してもらった.

Table. 1 から、50 音式よりも 2 段階式の方が誤操作が少ないことが見て取れる. しかし、「行を選択した後にフリック入力の感覚で文字を選ぶ」というフリック入力に慣れていない状況では、使用しにくいという点が課題となった. また、実用化に向けては正解率をさらに上げる必要があると考えた.

そこで、50 音式、2 段階式のどちらでもない新たな入力方式として、作成予定のスクリーンキーボードのイメージを Fig. 5 に示す. これは、2 段階式の利点である一文字あたりの面積の広さを残しつつ、口の動きを併用することで、表示する文字をそれぞれの行の頭文字のみとし、視線の移動量を減らして、誤操作を軽減しようと試みたものである. また、視線の操作としては頭文字を見る、という単純なものになるため、熟練者でなくとも容易に操作できると考えられる.

3.3 口の形状の利用法

2.2 で記した通り、口の形状の判別は OpneCV を用いて行う. 形状の判別の前に、まず口の位置を取得する必要がある. そのため、口を検知できるように、正解画像と不正解画像を用意し、それらをベクトルファイルに変換することで学習データを作成する. 口を検知する際、背景などによる誤検知を防ぐため、顔を検知したのちにその領域内で口の検知を行うこととする.

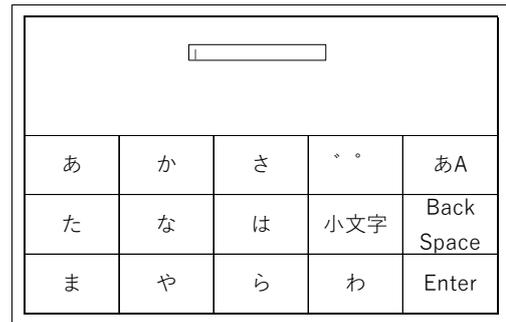
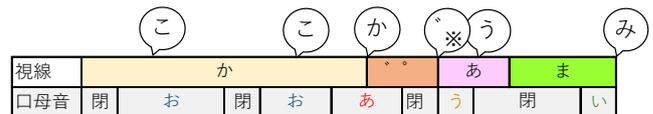


Fig. 5 作成予定スクリーンキーボード



入力文字:ここがうみ

※一部の入力は口の形状に関係なく、普通の文字入力の倍時間注視することで入力.

Fig. 6 入力の流れ例

口の検知後に行う形状の判別法として、口の上下左右の端点から長方形を描画し、その長辺、短辺の大きさよりどの母音かを判別する. この際、Web カメラとの距離によって大きさが変化するため、顔の認識をした後、その大きさと割合から判別を行うこととする.

文字入力の流れは、どの文字を見ているか、口がどの母音の形をしているか、という二つの状態から入力したい文字を判断する. それぞれの状態に移行してから一定時間が経過すると、文字が入力される. 同じ文字を連続して入力するときの誤操作を減らすため、視線と口の形状のいずれかの状態が変化しなければ次の文字は入力できないようにする. 流れのイメージを Fig. 6 に示す. また、口を閉じている状態を休憩に設定することで、誤操作を軽減する.

4. 結言

本研究では、アイトラッカーと Web カメラの併用により、既存のものに比べ快適な利用ができる視線入力システムの開発が目的である. そのため、作業効率や精神的なストレスを考慮するとともに、誤操作や疲労の軽減を目指して、視線入力操作性と画面表示の改善を図るためのシステムの拡張について考案した.

今後は、画像認識の手法を確立し、作成予定のシステムを実現する. さらに、そこからの発展として、ひらがな入力は口の形状を用いることで直感的な操作ができるが、アルファベットではそれができないため、その部分の改善を予定である.

5. 参考文献

- [1] tobii pro 「アイトラッキングとは」
 < https://www.tobiipro.com/ja/whats_eyetracking/>