

視線を利用したコミュニケーションシステムの開発

～ ALS とともに生きる方のために～

John Paulin HANSEN (IT University of Copenhagen, Denmark)
Håkon LUND (Royal School of Library and Information Science, Denmark)
青木洋貴 / 伊藤謙治 (東京工業大学)

1. はじめに

視線を入力手段として利用するコンピュータシステム（一般に注視インタクションシステムと呼ばれる）は、通常のマウス・キーボードなどによらず入力が可能となることから、きわめて有効なコミュニケーション手段となりうると認識されている。この注視インタクションに関連する技術開発は20年以上の歴史を有しており、現在においても着実に技術改良が進んでいる (Majaranta and Raiha, 2000)。これらの技術により、症状が進行し、運動機能に重大な障害を持つに至った ALS (Amyotrophic Lateral Sclerosis) 患者の意思伝達・会話コミュニケーションを支援することが可能である。この技術を用いれば、患者がコンピュータの画面上に表示されている文字や文章、コマンド（「テレビを消す」「エアコンをつける」など）をただ注視する、すなわち「じっと見る」だけで、家族や介護者とコミュニケーションをとる、あるいは自ら行いたいことを自らの意思で実行する、といったことが可能となる。

ここで簡単に注視インタクションシステムについて、その概要を説明する。このようなシステムは、「目の動きを追跡する装置」（一般に「眼球追跡システム」と呼ばれる）と、実際に文字や文章、コマンドを入力するための「入力インタフェース」という2つの要素から構成されている。以下にそれぞれの構成要素について説明する。

眼球追跡システムは、両眼あるいは片眼の動きを追跡するビデオカメラとそれに付随するソフトウェア等から構成されている。現時点で実装されているシステムの中には、このカメラがユーザから見えないように、コンピ

ュータディスプレイの中に設置されているものもある（このようなシステムでは、見かけは通常のパソコンとほとんど変わらない）。また、カメラをコンピュータとユーザの間に、別個に設置することが可能なタイプのシステムも存在する（このシステムでは、ALS 患者の状況に応じて設置方法を柔軟に変更することができる）。現在使用可能なほとんどのシステムでは、眼球に赤外光を当てることにより、画像処理に必要な鮮明な眼球の画像（瞳孔および角膜反射）を得ている。ユーザはシステムを使用するに当たって、「キャリブレーション (calibration)」と呼ばれる、目の動きと視線の対応をとるための調整・設定作業が必要となる。この作業は、使用を開始する前に画面に表示される5-12個の点を「じっと見る」ことによって実行され、通常は1-2分程度で完了する。このように、これまでの技術改良によって、実用に耐える眼球追跡システムはいくつか存在している。ただし、このような眼球追跡システムは依然として非常に高価であり、購入に当たっては10,000~20,000ドル/ユーロ（日本円で百数十万円~三百数十万円程度）程度の出費が必要となってしまう。

一方入力インタフェースについては、キーボードを模したものを画面に表示するオンスクリーンキーボード（たとえば、「Point for Windows」「Wivik」などが有名である）の利用が、注視インタクションシステムには適している。このようなオンスクリーンキーボードを利用した入力インタフェースは、身体障害者向けの会話コミュニケーションをサポートするシステム (type-to-talk システムと呼ばれる) において、10~20年前からすでに活用されている。

type-to-talk システムで実装されているオンスクリーンキーボードにおいては、文字を通常の QWERTY キー

配列に準じて配置するだけでなく、50音順・アルファベット順に配置する、あるいはよく使う文字順に配置するといったことも行われている。最近のtype-to-talkシステムの中には文字・単語予測機能（次にユーザが入力しそうな文字・単語を予測し、それらをキーボードに配置する機能）を備えているものもある。また、特定のコンテキスト（たとえば「夕食時」）に関連した単語辞書へのアクセスを可能にする機能を有しているシステムも存在している（このようなシステムでは、単語の意味が自動的に表示される）。

このようにオンスクリーンキーボードに関連する技術改良が進んでいるものの、これを用いた注視インタラクションシステムにおいても、口頭でのスムーズな会話と同等程度（1分当たり150英単語・かな300文字の入力）の効率的なコミュニケーションの実現は、残念ながら将来的にも難しいのが現状である。そこで、注視インタラクションシステムの開発における長期目標として、QWERTYキーボードによる通常の文字入力と同等のレベルを実現することを設定している。このレベルが達成されると、ユーザはオンラインチャットへ問題なく参加するのに十分な入力速度が得られる。なお、注視インタラクションシステムによる文字入力（すなわち、オンスクリーンキーボードを用いた視線による入力）は、現時点では携帯電話やPDAなどの携帯端末における文字入力方法と同等程度の効率が見られている。すなわち、1分当たり10英単語・かな20文字の入力が可能なレベルである。特にDasherシステム（Ward & Mackay, 2000）¹を用いた場合、視線による入力に十分に習熟した後で、モールス信号および手書き入力と同等（典型的には1分当たり25英単語・かな50文字程度の入力）の入力効率が達成できることが示されている。

追跡精度がきわめて高い眼球追跡システム（すなわち、購入費用はより高価になる）の中には、通常のWindows-OS環境下でマウスポインタをコントロールすることが可能になるような性能を有しているものもある。ただし、このようなことを実現しようとすると、画面の解像度は通常より下げる（たとえば600×800 pixels程度）、あるいはアイコンのサイズを通常よりも大きくするというような必要が出てくる場合がある（これらは、画面上の情報を大きく表示するための操作であ

る）。なお、このようなことを行わなくても済むように、通常の解像度環境でアイコンを拡大表示する機能を、いくつかのメーカーが提供している。

先進的な注視インタラクションシステムにおいては、アイコンやメニュー、リンク等を、その部分に視線をある決まった時間（滞留時間）以上維持することによって選択・実行することが可能となっている。この選択・実行方法は、画面上のオブジェクト（アイコン等）や文字を、事前に設定した時間（たとえば0.5秒間など）「じっと見続ける」ことで、それらの選択・実行・入力が行われるというものである。この方法を利用したシステムの使用当初においては、多くのユーザは1秒間以上を滞留時間として設定しているが、このようなシステムに十分に慣れたユーザの場合0.5秒間の滞留時間でもコントロールすることができる。なお、ほとんどのユーザは疲労した場合に、意図しないオブジェクトや文字の選択・実行を避けるために、滞留時間をより長く設定することを望ましいと考えるようである。このような方法とは別に、ユーザの瞬きによって選択・実行をすることが可能なシステムも存在しているが、一般論としてこのようなシステムは、文字の入力手続きが厄介で時間がかかるという問題をもたらし可能性もある。

ここまでは文字の入力を主とするコンピュータインタフェースへの注視インタラクションシステムについて論じてきた。ここではこのようなシステムの、代表的な福祉機器である車いすへの応用について、簡単に説明する。現時点においては、車いすへの応用（つまり視線により自らの意思でコントロールできるような車いす開発）に関する研究報告はほとんど見られず、このような応用については現時点でも研究開発が必要な重要分野である（すなわち、このような分野においてはまだ十分な技術の蓄積がない）。このことは、たとえば人々が回りで歩いているといったような、動的に変化する環境の中では、人間が視線を意図的にコントロールすることが難しいことが原因の一つである。個人がパソコンを利用するような場合であれば、意図しないオブジェクトや文字の選択・実行が発生しても、それらはユーザに不快な思いをさせるかもしれないが、それほど重大な損害・損失は招かない。しかしもしそのような選択・実行が、車いすの操作中に起こった場合、パソコンの場合に比べると、より重

大な結果（たとえば死傷事故など）を招いてしまうかもしれない。また、現時点では眼球追跡システムの多くは、変化が大きい照明環境（たとえば日光やネオン光がカメラに入り込んだり、ユーザの眼球に映り込んだりするような環境）の中ではうまく動作しないという問題も存在している。そして、車いすが移動する際に生じる振動も、眼球の追跡を困難にする原因となっている。すなわちこれらが、このような車いすの実現を難しくしている。

2. ALS の進行ステージとコミュニケーション能力

ALS と共に生きる方々に対して著者らが主として勧めたいこととして、注視インタラクショシステムに対する練習・準備を、これらが本当に必要になる前（すなわち手や足の運動機能がすべて失われる前）から始めるということがある。注視インタラクショシステムの使用に必要な視線のコントロールを習得するためには、ある程度の時間が必要であり、典型的には5時間程度の入力練習が必要である（Itoh et. al., 2006）。多くの注視インタラクショシステムでは、視線以外の他の入力方法もサポートしている。そのため、ユーザの手や足の運動が可能な時点で、文字や文章、コマンドを入力するための「入力インタフェース」をジョイスティックやマウス、ヘッドトラッキングシステム（顔の方向でカーソル

をコントロールする入力機器）などを用いてあらかじめ使用し、慣れておくことが可能である。もしこのようにユーザが早い段階で練習・準備することができれば、その間に作成した文章や頻繁に用いた単語などを記録しておき、たとえば文字予測機能などの効率的な文字入力に向けて利用することなどもできるかもしれない（少なくともヨーロッパ言語では可能である）。

ALS の病状は徐々に進行する。そして、コンピュータシステムの使用においてユーザが感じる疲労は重要な要因となり、これによって1日の中でもより楽な入力方法に切り替える必要がでてくる。表1は、ALS の進行に伴う障害の段階を示したものである。もちろん、同じ進行ステージにあったとしても、このような一般的なパターンとは異なるようなことは頻繁にあり、個人差はきわめて大きい。たとえば人によってはあるステージに長期にわたってとどまる、多くの運動・筋機能が失われても発声機能が残る、あるいは次のステージに進まずにすむといったようなことがありうる。すなわち、表1では、あくまでも一般的にはあるが、ALS の症状が次にどのような進行することがあり得て、そのようになった場合どのような入力方法が選択できるのかということを予め考慮しておくといった、事前の対策が賢明であることを説明しようとしている。

表1 ALS の典型的な進行（Johansen and Hansen [2006] を修正）

	症状	適した入力デバイス
ステージ1	疲労は顕著。手・腕の可動性および筋力の低下。しばしば発声が不明瞭。	ハンド・アームレストを利用したキーボード、および修正した操作手順（sticky shift の利用、繰り返し操作の排除）
ステージ2	疲労は要因。筋力不足により腕は動かせない。ただし、両腕もしくは片腕の可動性は保持。発声はきわめて不明瞭であり、他者から見ると意図・意志などを読み取ることが難しい。	マウス、ジョイスティック、あるいはキー数を減らしたキーボード（5-10 キー程度）。
ステージ3	ほぼ full lock-in. 発声機能は失われる。四肢の可動性がきわめて失われる。	眼球および頭部運動に基づく 1-2 個のスイッチによる入力
ステージ4	Full lock-in. 眼球運動は可能。	視線入力。たとえば EMG あるいは瞬きによる選択・入力
ステージ5	Full lock-in. 眼球運動も不可能。	生体電位スイッチ（たとえば EMG, EEG, EOG など）によるスキャン選択入力

あるALSの男性患者は、眼球運動が徐々に失われていってしまうこと（ステージ4から5へ進む段階）についての懸念を、メールの中で以下のように伝えてきた。「これはきわめて重大な問題です。なぜならば、ALS（PALS）が進行した患者の多くは、眼球運動の機能が弱まっていくことを経験するからです。これは多くのALSに関する一般的な学術書の記述とは相反するものです。でも、これは実際に起こることなのです。私は20年近く人工呼吸器とともに生きてきた日本のPALSの患者の何人かが、前をじっと見ることはできない状態にあることを知っています。さらに、症状が進行した患者の多くは、唾液の分泌を抑えるために、ScopodermやAtropin等の薬を投与されています。これらの薬は副作用として、視覚機能の精密さ・正確さを失わせるような影響を与えてしまうのです。」

著者らはこの問題に対しは、以下のように考えている。ステージ5においては、その前の段階で注視インタラクションによって慣れたコミュニケーションシステムについては、これを利用するためにはEMG（筋電位）やEEG（脳波）によるスイッチ動作を適用することで十分対応できる可能性がある。しかし、著者らはここでも強調したいのだが、このようなためにもEMGによるスイッチ動作について、早いタイミングで（すなわちこれらが差し迫って必要となる前に）導入しておくことを勧める。EMGによるスイッチを早く使い始めることを勧めることについては、次のような理由もある。それは、EMGによるクリック動作は、注視インタラクションシステムで用いられる視線の滞留時間による選択・実行動作（「1. はじめに」を参照されたい）に比べると、より頑健で迅速に行うことができる能力を潜在的に持っていることである。特に、患者の状況から、選択・実行に際して滞留時間を0.5秒間以上に設定しなければならないような場合は、このようなEMGを注視インタラクションに組み合わせる（すなわち、視線を選択、EMGスイッチを実行に使うような組み合わせを採用する）ことで、システムを効率的に使うことが可能となる。

3. 注視インタラクションシステム「GazeTalk」

「GazeTalk」は、著者らとALS患者グループとの密接なコラボレーションの中で開発されたALS患者向け

注視インタラクションシステムである。このシステムはALS患者が効率的に他者とコミュニケーションをとること、そして娯楽や情報収集・発信のためのコンピュータプログラムを操作することを可能にしている。文字やコマンド等を入力するための「入力インタフェース」は、きわめて少数の数のキーからなるオンスクリーンキーボードを採用している。このことから、眼球の追跡が困難であったり意図的な視線移動が困難なことが原因で十分満足な注視コントロールが得られなくとも、このインタフェースを操作することができる。このシステムは、合成音声を利用したtype-to-talk機能（会話コミュニケーションをサポートする機能）、文書作成、電子メール、web閲覧、映像再生、MP3ファイルの再生、そしてPDFファイルの迅速な閲覧機能を有している。さらにGazeTalkは、きわめて迅速なタイピングソフトウェアであるDasherシステムへのダイレクトアクセス機能を提供するとともに、使用言語の変更をシステムを使用する中でユーザ自身が行うことができる機能も有している。

この「GazeTalk」システムはフリーウェアであり、現時点では以下の言語で提供されている。それらの言語は、デンマーク語、英語、イタリア語、中国語、そして日本語である。これら以外の言語についても、きわめて近いうちに提供する予定である。これらのシステムは、次のウェブサイトwww.cogain.org/downloadsからフリーでダウンロードが可能であり、ぜひ試用していただきたい。このシステムはキーボード、マウス、トラックボール、ジョイスティック、ヘッドトラッキングシステム（顔の方向でカーソルをコントロールする入力機器）、そして注視による入力に対応している。さらにこのシステムは、スキャンによる選択・実行（システムが選択候補を順次提示して行き、ユーザが必要ところで実行のためのキューを与える方法）もサポートしており、これにより上記の入力以外にマウスのボタン、EMGスイッチ、およびこれ以外の各種スイッチによる入力も可能となっている。図1は、日本語版GazeTalkシステムの操作画面の例を示したものである。

オリジナル版のGazeTalkはヨーロッパ言語向けに開発されている。このシステムは入力文字の予測機能を有しており、この機能により最後に入力された文字の次に



図1 日本語版 GazeTalk における文章作成（左上）、電子メール（右上）、音楽再生（左下）およびWeb 閲覧（右下）

入力される可能性が高い6つのアルファベットと、ユーザが入力を意図している単語を予測することが可能となっている。このシステムでは、動的なメニュー構造を採用しており、この中ではどのキーにどの文字が割り当てられるかは固定されず、変化する。一方日本語版においては、階層的なメニュー構造が採用されている。すなわち、文字やコマンド等は常に同じ位置に固定されている。たとえば文字入力のトップ階層においては、50音の行（「あ行」など）が割り当てられている（図1参照）。たとえば「な行」が割り当てられているキーを予め設定された時間（滞留時間）以上じっと見続けると、各キーにな行のかなが割り当てられた、下位のメニュー階層が表示される。この階層（すなわちかなが割り当てられている階層）において、ユーザは各キーをじっと見ることによって、かなを入力することができる。さらに日本語版においては、かな漢字変換機能も提供しており、通常のかな漢字交じり文の作成も問題なく可能である。

4. おわりに

注視インタラクションは、手によってカーソルをコントロールできないユーザにとって、もっとも望ましい代替手段の一つであるように思われる。ALSとともに生きる患者の方に視線を利用したコミュニケーションシステムの使用に徐々に慣れていただくために、このような方々に対して注視インタラクションシステムを、症状が進行する早い段階のうちから導入すべきである。すでに利用可能なさまざまな注視インタラクションシステムが、存在している。本稿では、ALS患者の症状の段階に対応したコミュニケーションをサポートすることが可能な GazeTalk システムについて紹介した。この GazeTalk システムを用いることで、ユーザは「話す」「書く」「電子メールの送受信」「インターネットの閲覧」が可能になり、さらに音楽や映像も楽しむことができる。

参考文献

Majaranta, P., and Rih_, K. J. 2002. Twenty Years of

Eye Typing: Systems and Design Issues. In Proceedings of the Symposium on ETRA 2002: Eye Tracking Research & Applications Symposium 2002, New Orleans, LA, 15-22.

Johansen, A. S. and Hansen, J. P. 2006. Augmentative and alternative communication: the future of text on the move, Universal Access in the Information Society, Jan 2006, 1 - 25.

Itoh, K., Aoki, H., and Hansen, J. P. 2006. A comparative usability study of two Japanese gaze typing systems. In Proceedings of the 2006 Symposium on Eye Tracking Research & Applications (San Diego, California, March 27 - 29, 2006). ETRA '06. ACM Press, New York, NY, 59-66.

Ward, D. J., and MacKay, D. J. C. 2002. Fast Hands-Free Writing by Gaze Direction. Nature 418, p. 838, August 22.

¹ オンスクリーンキーボードを用いた type-to-talk システムの 1 種