

Fig. 6. Evaluation Results of the Font Recognition in All Characters.

with aging (3), (4). Our previous study also shows that the recognition ability of the tactile fonts of the elderly is lower than that of the young. It also shows the learning effect of the elderly is lower than that of the young (5). Since the subjects in this report include the elderly, other measurements including a special training program will be needed.

The reason why the correct answer of subject C shown in Figure 6 is higher than that of Figure 5 is seemed to be because Subject C has had a few experiences before this evaluation tests. The other subjects were also able to confirm the correct answer on each process. It is suggested that the recognition ability of subject C increased by learning effect while performing the evaluation tests.

With the regard to the contents of the mistakes, the results show as follows. Subject B mistook "B" for "8" twice. Subject B also mistook "D" for "0" twice. The other mistakes of subject B were once. Subject C mistook "5" for "S" twice. Subject C also mistook "D" for "O" twice and "O" for "c" twice. The other mistakes of subject C were once. It is supposed that we need to improve the design of several fonts that are observed conspicuous mistakes.

References

1. T. Watanabe and M. Kobayashi: Prototypes of the electronic tactile drawing system for blind persons, TVRSJ, Vol.7, No.1, pp.87-94, 2002. (In Japanese)
2. M. Shinohara, Y. Shimizu and A. Mochizuki: Three-dimensional tactile display for the blind, IEEE Transactions Rehabilitation Engineering, Vol.6, No.3, pp.249-256, 1998.
3. D. Kenshalo: Somesthetic sensitivity in young and elderly humans, Journal of Gerontology, Vol.41, No.6, pp.732-742, 1986.
4. J. Stevens: Aging and spatial acuity of touch, Journal of Gerontology, Vol.47, No.1, pp.35-40, 1992.
5. M. Sakajiri, K. Ito, S. Okada, H. Tomita and T. Ifukube: A development of the tactile display system to show the characters for individuals who are deaf-blind, Journal of Human Interface Society, Vol.5, No.4, pp.455-464, 2003. (In Japanese)

Appendix

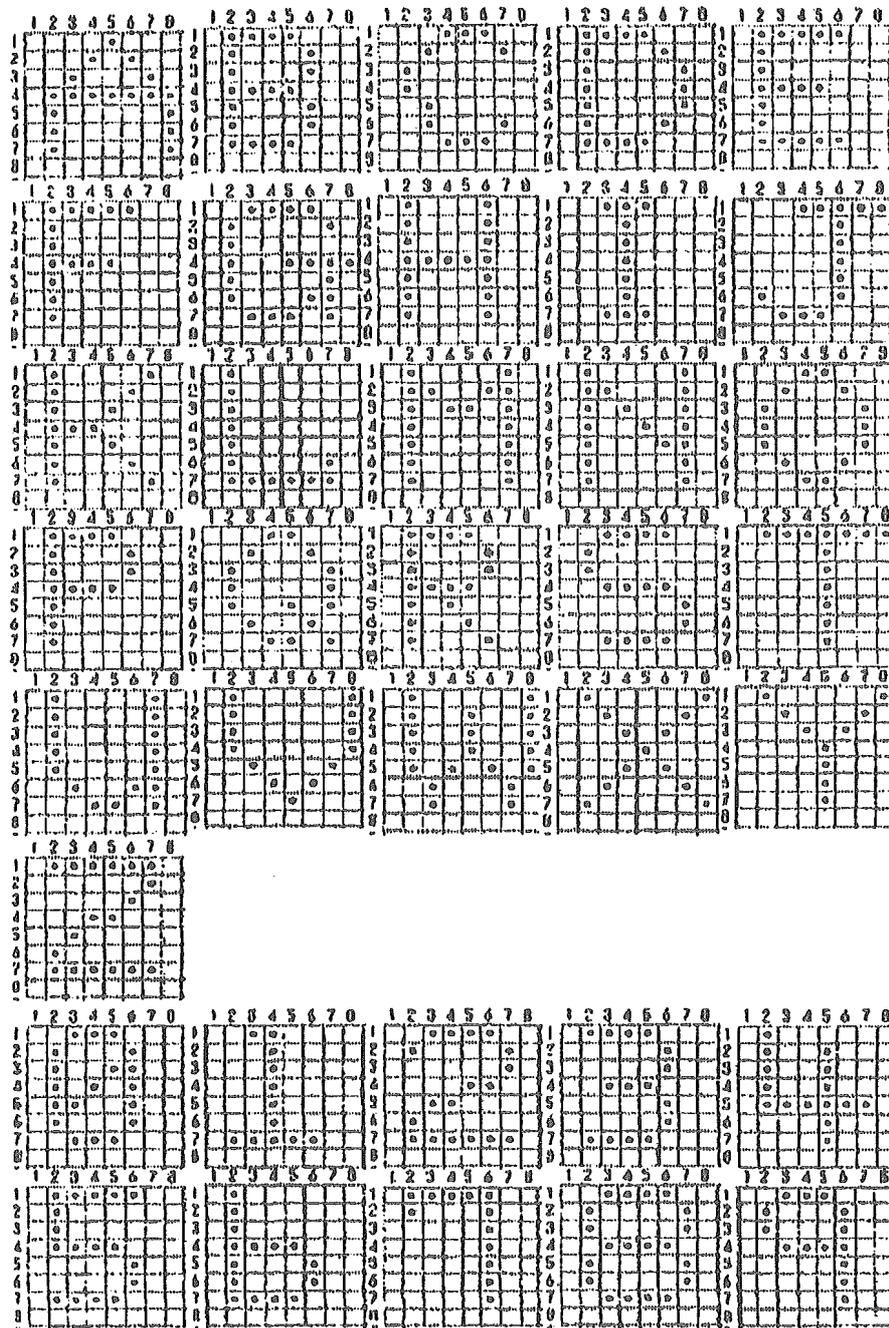


Fig. A-1. This figure illustrates "A", "B", "C", "D", "E", "F", "G", "H", "I", "J", "K", "L", "M", "N", "O", "P", "Q", "R", "S", "T", "U", "V", "W", "X", "Y", "Z", "0", "1", "2", "3", "4", "5", "6", "7", "8" and "9".

ビデオキャプチャ画像処理による視線検出及び意思伝達装置への応用

伊藤 和幸^{†a)} 伊福部 達^{††}

Image Processing with Video Capture for Eye Movement Measurement and Its Application to Eye Control System

Kazuyuki ITOH^{†a)} and Tohru IFUKUBE^{††}

あらまし 本研究では、重度肢体不自由者向けに視線を利用した意思伝達システム（視線入力式文字入力装置・環境制御装置）を安価に提供できるようなシステムを開発する。視線の検出には、画像処理ボードを利用した高精度・高機能なものが市販されているが、それらを使用するとコストダウンが困難であり実用に結び付かないという欠点があるため、本研究ではビデオキャプチャした画像をソフトウェア的に処理することで視線検出を行い、システムにかかるコストの削減を図った。近年のパソコン性能の向上により、ノートパソコンの使用も可能でありシステムのコンパクト化も実現できた。

キーワード ビデオキャプチャ、視線、支援機器、重度肢体不自由者

1. ま え が き

筆者らはこれまでに、重度肢体不自由者のQOL（生活の質）向上を目指し、眼球運動（視線）を利用した意思伝達装置（文字入力によるコミュニケーション装置や環境制御）の開発を行っている[1]~[3]。背景には、筋ジストロフィー患者や筋萎縮性側索硬化症患者（ALS患者）のような重度の障害者では、病状や症状の進行に伴い四肢運動機能に障害が生じ、文字盤への指差しや走査選択式装置[4]へのスイッチ操作が困難になるため、意思の伝達に利用できる身体部位が眼球だけになる場合も出てくるためである。

眼球運動を意思伝達機器の操作に利用するには、従来は上下左右への眼球の動きや目の開閉をスイッチのON-OFFとして検出し、1スイッチの走査選択式装置への入力に利用する方法が採用されている[5],[6]。しかし、これらの方法は走査選択式であるがゆえに操作効率の悪さがネックとなっている。一方、近年では文字やアイコンを見つめるだけでそれらを選択できる

視線入力方式のシステムが提供されるようになってきた[7]~[10]。視線入力方式は障害者が見つめた対象をそのまま選ぶことのできる直接的な選択方法で、操作は分かりやすく、文字入力に応用した場合入力効率にも優れている。しかし、セッティングの困難さや高額な価格が実用に結び付かない原因となっている。

市販の視線検出装置を利用した場合、これらは一般的には研究向けであるために高い性能（60Hzのサンプリング時間またはそれ以上の時間分解能と高精度な空間分解能）が要求されており、日常生活で利用できる程度の価格にはなりにくい。しかも、視線検出装置の需要は多いものではなく、ハードウェアを利用したシステムでは、大量生産によってコストダウンを実現するのは困難となる。

そこで本研究では、眼球運動を利用した障害者向け意思伝達システム（視線入力式文字入力装置と環境制御装置）の実用化を目指し、課題の一つであるシステムコストの削減を目標とする。具体的には、視線の検出を専用のハードウェア（視線検出専用機）に依存せず、ビデオキャプチャにより取り込んだ画像をソフトウェア的に処理するシステムの開発により、コストの削減を図る。以下、2. でビデオキャプチャによる画像処理の概要について述べ、3. で視線入力式コミュニケーション装置の開発内容、4. で環境制御装置の開発内容について述べる。5. でまとめを行う。

[†] 国立身体障害者リハビリテーションセンター研究所、所沢市
Research Institute, National Rehabilitation Center for Persons with Disabilities, Tokorozawa-shi, 359-8555 Japan

^{††} 東京大学先端科学技術研究センター、東京都
Research Center for Advanced Science and Technology, University of Tokyo, Meguro-ku, Tokyo, 153-8904 Japan
a) E-mail: ito@rehab.go.jp

2. ビデオキャプチャによる画像取込みと処理

本研究では、ビデオキャプチャにより画像をパーソナルコンピュータ（以下、パソコン、PC）に取り込み、その画像をソフトウェアにより処理して視線の検出を行うこととする。

本研究で開発するシステムで利用したハードウェアは、ノートパソコンのほかに視線入力式文字入力装置では安価なビデオキャプチャ用 PCI ボードと PCI ボード用拡張ボックス、PC カード、環境制御装置では USB カメラとテレビ用リモコンであり、画像処理専用ボードを利用しないことでシステムにかかるコストの削減を図っている。

ビデオキャプチャした画像で処理を行う方法を採用した理由には、視線の検出を行うにはビデオ信号を画像処理ボード上で処理して行うのが一般的であるものの、意思伝達向けには研究向けの性能を追求する必要はなく、10-15FPS (frame per second) 程度の処理能力と後述するような空間分解能があれば十分実用的となることが挙げられる。また、近年のパソコン性能の向上 (CPU 性能、大容量メモリの搭載) により、ソフトウェア的に画像処理を行ったとしても十分実用的となることも開発を進める上で有利な点となった。

システムは OS として Microsoft Windows XP、コンパイラに Visual C++ を利用してアプリケーションを作成する。そのアプリケーション上では Windows Driver Model を利用して画像を取り込み、その画像をソフトウェア的に処理して視線の計測を行う (図 1)。

3., 4. では、それぞれ視線入力式コミュニケーション装置と環境制御装置の概要について説明する。視線により文字入力を行うのか、簡単な環境制御で十分なのか、障害者のニーズにより利用するアプリケーショ

ンを決定する。

3. 視線入力式文字入力装置の概要

3.1 視線検出方法

眼球運動を意思伝達に使用するためには、機器の利用者がどこを見ているか (=視線) を検出する必要がある。人の視線を検出する試みは最近の話題ではなく、19 世紀の後半より視線検出技術が研究されている [11]。計測方法としては、EOG 法 [12]、強膜反射法 [13]、角膜反射法 [14], [15] 等が提案されており、本研究では角膜反射法を視線検出方法として採用した。角膜反射法は眼球に当たった光源 (一般的には赤外光) の光が角膜上で反射する点 (角膜反射点) と瞳孔中心点を検出し、これら二つの特徴点の相対的な位置関係を利用することで視線方向を検出する方法であり、市販の視線検出装置では画像処理ボードで二つの特徴点を抽出するのが一般的である。本研究では、画像処理ボードを使用せずにビデオキャプチャした画像から二つの特徴点を抽出することで視線の検出を行う。

3.2 画像処理手順

画像の処理手順は、以下のとおりである。

3.2.1 画像 (640 × 480 画素) の取込みと表示

眼球付近の画像を取り込み、その画像をアプリケーション上 (画面右下) に表示する。その結果、画像表示用のモニタを別途接続する必要がなくなり、システムがコンパクト化する。画像の取込み及び処理は、サンプリング速度が速ければ速いほど視線を移動させたときの追従性が良いが、現状ではソフトウェア的な処理能力と CPU への負担を考慮して 15FPS とした。

3.2.2 角膜反射点の特定

取り込んだ画像を高しきい値で 2 値化し、画像内の白い箇所を検索し、角膜反射点とみなす。したがって、眼鏡における反射が角膜反射点とみなされると正確な視線検出が行われないため、その場合は画像内に眼鏡の反射が入らないように取込み位置の工夫が必要となる。また、顔面における不要な反射を抑えるために必要以上に室内が明るくならないように考慮する。

3.2.3 瞳孔中心点の特定

しきい値を低く設定して再度 2 値化し、角膜反射点付近にある黒いパーツを瞳孔とみなす。そのパーツに楕円近似を施し瞳孔の中心点を算出する。

3.2.4 視線の検出

3.2.2, 3.2.3 の処理により二つの特徴点が検出されるので、画面上の基準点 (四隅) を見つめてキャリブ

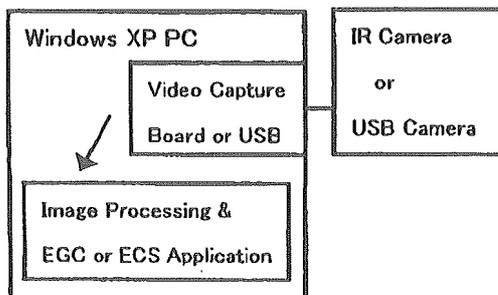


図 1 システム構成
Fig. 1 System configuration.

レーション（校正作業）を行うことで、画面上の視点が算出できる。

3.3 視線入力式文字入力装置

視線入力式文字入力装置は図2のように、ノートパソコンと赤外光源、赤外カメラ、PCIビデオキャプチャボード、PCIボード用拡張ボックス、拡張ボックス用PCカードで構成され、片足ベッドサイドテーブルに乗せられる程度となっている。現状ではUSBカメラに赤外対応のものがなく、本システムでは赤外カメラのビデオ信号をPCIキャプチャボード経由で取り込み利用した。図2中央の三脚に固定されているのが赤外光源と赤外カメラ、左の液晶ディスプレイ下部にあるのが拡張ボックスである。液晶ディスプレイは下部に赤外カメラと光源を固定しベッド上にアームで伸ばせば、仰臥位（仰向けの状態）でも利用可能となる。図3、図4にシステム構成と設置概要を示す。

3.4 文字の表示及び選択方法

日本語文字入力の場合、使用する文字は清音46文字のほかに濁点、半濁点、長音記号、編集・漢字変換用の選択肢などを加えると60個程度の選択肢が必要となる。これらすべてを一つの画面内に表示し、それぞれを見ている視線を検出できる性能（空間分解能）を

追求すると研究向けのシステムと同程度のコストがかかるために、市販の福祉用視線入力装置では、画面上の10個程度の選択肢を選択できる程度の性能としている。その場合の文字選択方法は、まず「あかさ…」を表示した画面から一つを確定し、その後選択した子音に対応する母音（例えば「か」を選択した後は「かきくけこ」）を表示して、最終的に目的の文字を確定する方法が採用されている[8],[10]。

一方、本システムでは図5に示すように初期画面に一度に60個の選択肢（文字及び編集・漢字変換用文字）を表示し、第1段階（図5左）では4個ごとのグループを選択し、第2段階（図5右）では選択したグループ中の文字をもとの選択個所の周囲に再表示して、最終的に目的の文字を選択する方法を提案する。これは、選択肢の近傍に確定用の注視個所を配置し、選択肢を見つめた上で確定個所へ視線を移動させることで誤選択なくかつ高速に選択を行うことを示した先行研究[16]の有効性を応用したものである。更に、初期画面に全文字を表示することで目的の文字を見つけやすくなるという利点に加え、画面上の15分割（3行5列）個所を見つめる視線の検出を行えばよいので、高い視線検出精度が要求されないという利点もある。15inchのディスプレイを視距離60cm程度で利用すると、一つの分割個所は5.92deg程度の視角となり、高性能な機器に比べてはるかに容易な視線検出機能の設計が可能となる。

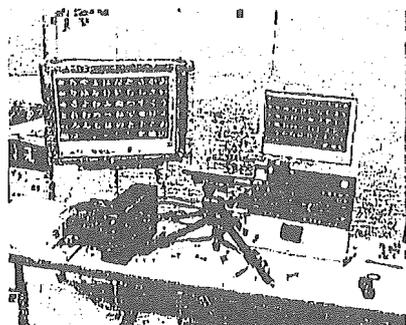


図2 視線入力式文字入力装置の概観
Fig.2 Eye gaze communication system.

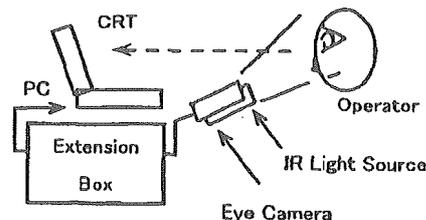


図4 装置の設置概要
Fig.4 Arrangement of measurement equipment.

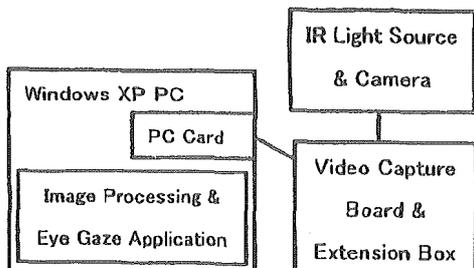


図3 視線入力式文字入力装置のシステム構成
Fig.3 System configuration of EGC system.

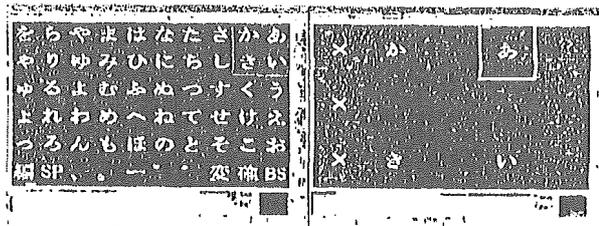


図5 初期画面と第2段階画面
Fig.5 Screen of 1st, 2nd phase.

文字グループの設定は、現状では4個ごととしたが6個(3行2列)ごととすることも可能である。この場合、分割数は10分割(2行5列)となり視線検出機能の設計は更に容易になるが、第1段階では縦並びだった文字列が第2段階では横並びとなる。本システムでは、画面上の15分割個所を見つめる視線検出精度が十分得られていること、第1・第2段階における文字の並びが統一されることを基本として、文字グループの設定を4個とした。文字並びの変更による混乱の有無や使用感の善しあしなど、他の設定による使用評価は今後の検討課題としたい。

また、第1段階で選択した個所は第2段階では空白となり、その個所を見続けても文字は選択されないため、いわば休憩個所となる。文字の確定は、瞬きもできなくなる患者も存在することを考慮して、目的の個所を設定時間見つめ続けることで行っているが、選択個所のすべてが文字の場合、選択を意図しない個所を見つめていても文字が確定されるとその訂正作業も必要になり疲労が増す、という意見がこれまでの臨床評価を通じて出されたため、休憩の意味合いも含めて空白個所を設けた。

図5では、第1段階では右上の「あいかき」個所に視線があり、ここが選択されると第2段階のような表示となる。第2段階では第1段階で選択した右上個所が空白となり、その個所周辺に「あいかき」が配置され、目的の文字を再度注視・確定することとなる。

文字の確定は、図6のように算出される画面上の視線移動速度と継続時間により視線の状態を移動・停留(目的の文字を見つけ、移動速度がしきい値以下で低下し始めた状態)・注視に分類し、注視状態において最も注視頻度の高い文字を注視文字としている[1]。停留の設定時間は文献[1]で議論したような初期値を設定するが、短すぎると感じる場合には利用者に合わせて調整することとし、第1、第2段階それぞれにおいて変更を可能としている。注視の設定時間も同様である。

図5右の「×」個所は、選択すると文字選択を行わずに初期画面へ戻るため、初期画面で誤った個所が選択されても修正作業の必要がなくなる。図7に示す漢字変換画面にある文字は、それぞれ「《,》」:変換範囲の変更、「→」:変換する文節の変更、「前」:変換の前候補へ戻る、「変」:再変換、「確」:変換候補の確定、「カナ」:片仮名へ変換、「×」:変換画面から初期画面へ戻る、機能を意味する。漢字変換はWindows標準のMS-IMEを利用し、単語登録(「おはようございま

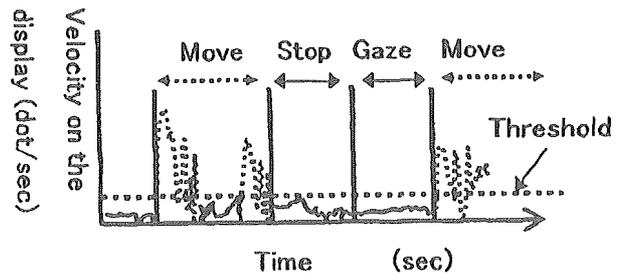


図6 移動速度による視線状態の分類
Fig. 6 Classification of eye movement.

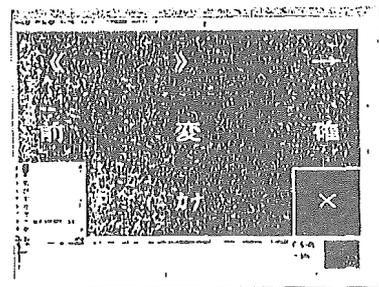


図7 漢字変換中の画面
Fig. 7 Kana-Kanji translation.

す」を「おは」などで登録し変換)により効率的な変換を行うことはできるが、単語予測機能は付加していない。

4. 環境制御装置

視線入力装置には画面上の15個所を見つめる視線の検出精度が要求されたが、眼球の左右への動きと円を描くような大きな動きを計測できる程度の精度であれば、システムを更に簡略化することができる。既に、Werdnig-Hoffmann (ウエルドニッヒ-ホフマン) 症患者自身によるテレビ制御システムをWindows 98をOSとした環境で開発しており[2],[3]、本研究では同様の機能をWindows XPをOSとしたシステムで構築し、USBカメラの利用も可能とした。

4.1 患者の状況

導入対象である患者の随意的な運動部位は眼球のみであり、意思の伝達は問いかけに対してYESの場合には円を描くように眼球を大きく回し(右回り)、NOまたは思案中には眼球を動かさないことで行っている。患者及び家族のニーズは、視線により文字入力を行うことよりはテレビの制御を患者自身で行えるようにしたい、ことである。

本患者に対しては数個のスイッチ制御が目的であり、病室内での実用性が重視される。数個のスイッチ制御

であれば EOG 法も有力だが、電極を皮膚に装着するための導電ペーストにかぶれるため採用できない経緯があった。そこで、可視光下で撮影した顔面付近の画像から虹彩と眉毛の重心位置を検出し、これらの相対的な位置関係を大まかな視線とすることで環境制御に結び付ける方法を採用した。

画像処理手順は以下のとおりである。

4.2 画像処理手順

4.2.1 画像 (640 × 480 画素) の取込み

アプリケーションの起動とともに顔面の画像を取り込み、アプリケーション上に表示する (図 8)。現状ではテレビの制御に必要な処理速度に合わせて 15FPS によるサンプリング及び処理を行う。

4.2.2 処理範囲の設定

虹彩と眉毛を処理するため、画像内に二つの方形フレームを定め処理範囲を設定する (図 9)。処理範囲の変更はマウスによるドラッグ操作で行い、ドラッグ & ドロップによりフレーム全体を移動できる。フレーム間の距離をロックすることでフレーム間の相対位置を保ったまま処理範囲全体を移動することも可能となっている。

4.2.3 グレースケール化及び 2 値化

取り込んだ画像を 256 階調のグレースケールに変換後、しきい値により 2 値化し (図 10)、しきい値以上

の画素のみ表示する。しきい値の調整はフレームごとに行う。

4.2.4 ラベリング処理

処理フレームごとに画素数によるマスクサイズ (M_{min} , M_{max}) を設定し、該当するパーツ ($M_{min} < M < M_{max}$) のみラベリングする。設定サイズより小さなパーツ (虹彩左にある目頭の影) や大きいパーツ (眉毛右にある頭髪の影) は処理対象から除外する。

4.2.5 形状分析

ラベリングされた各パーツの重心 (X , Y 座標)、画素数を算出する (図 11)。ラベリング処理同様、重心を中心としたサークルサイズ (S_{min} , S_{max}) を設定することで、該当するパーツ ($S_{min} < S < S_{max}$) のみ抽出することで、サイズ外のパーツ (右上の髪の毛など) は処理対象から除外される。また、眉毛に対して自動追従を実行すると、フレーム内の眉毛の重心位置が一定となるように処理フレーム範囲が自動的に移動するため、頭部の位置が CCD カメラに対して変動しても、ある程度は画像内での自動対応が可能となる。

4.2.6 処理結果によるリモコン制御

形状分析で処理された各値をもとに、環境制御用アプリケーションでテレビリモコンの制御を行う。

処理結果の概要を図 12 に示す。環境制御用には、眉毛の重心 (X_o , Y_o)、虹彩の重心 (X_i , Y_i)、各パー

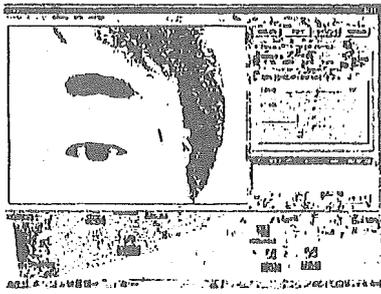


図 8 画像処理アプリケーションの起動
Fig. 8 Image processing application.

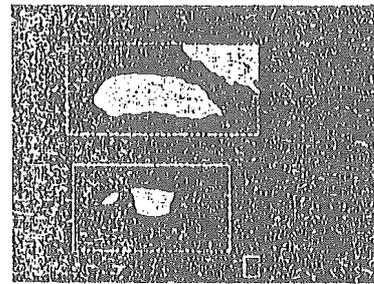


図 10 2 値化処理
Fig. 10 Binary processing.

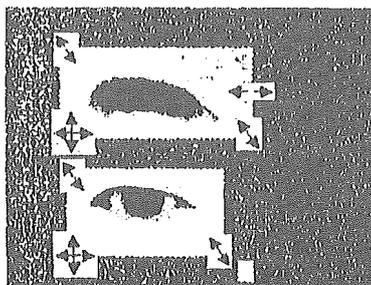


図 9 処理フレームの設定
Fig. 9 Arrangement of frame for image processing.

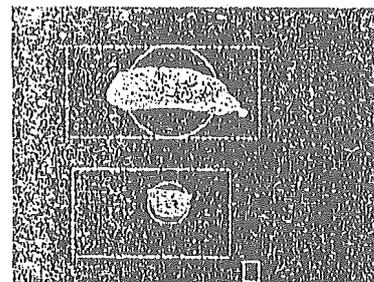


図 11 形状分析
Fig. 11 Analysis of shape.

ツの画素数値を利用し、

$$dX = X_o - X_i, dY = Y_o - Y_i \quad (1)$$

が相対値となり大まかな視線の推定が可能となる。

4.3 テレビの制御

テレビリモコン制御用の処理は、画像処理で求めた虹彩と眉毛それぞれの重心座標値を1秒間に10個保存し(10Hzのサンプリングとなる)、順次バッファ内でその値を更新しながら、2ch(テレビ電源とチャンネル操作)の制御を行うこととする。

4.3.1 電源操作

電源を操作するには、テレビを見ている状態と見ていない状態の眉毛と虹彩の相対値(大まかな視線方向)を検出し、図13のようにX方向の境界値(x1)による電源オン(テレビを見ている状態)とオフエリア(見ていない状態)を設定する。x1の調整は介護者が左右カーソルキーで値を変更するため、複数の基準点を見つめるキャリブレーション作業は不要となる。

電源スイッチは、電源オンエリアに虹彩が設定時間(例えば2.0秒とすれば20個のサンプリングデータ)以上得られていればオン、得られなければオフする。図13中の黒円(虹彩中は白円)はサンプリングごとの虹彩の重心点を示し、眼球の動きがなくなれば、重心の移動はほとんどなくなる。図では利用者が左を見た後、最終的に右を見ている。設定時間の経過後に機

器の制御を行うこととしたのは、画像の取込み状態による出力値のちらつきや瞬間的な眼球の移動(サッカー)等を電源操作のための動きと認識しないようにするためであり、テレビから目をそらしても設定時間内に再度テレビを見れば電源はオフされないこととなる。

4.3.2 チャンネル操作

また、テレビを見ながら円を描くように大きく眼球を動かすことでチャンネルをアップさせる機能も付加した。眼球が大きく動いたことを判定するには、1回の行為が1秒以内で終了するため、連続する10点の虹彩の座標値につきX、Y各方向の最大移動量(dx, dy)と移動軌跡による描画面積(A)を算出し、各値が設定範囲内(dx_min < dx < dx_max, dy_min < dy < dy_max, A_min < A < A_max)であった時点を大きな眼球運動と認識させた。dx_min, dy_min, A_minは、チャンネル変更を意図しない眼球運動に対する誤動作を排除するために設定している。dx_max, dy_max, A_maxは、サッカーなどによる大きな出力値や、カメラの前を人が横切った場合など、通常の画像と大きく異なる画像が得られ、結果的に眼球運動とはみなせない値が出力される際にも対応するために設定している。また、本システムでは眼球移動の方向も判断できるため、今後練習により左回りの運動ができればチャンネルダウンなど他の操作に対応させることも可能である。

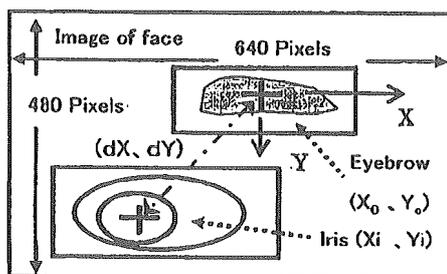


図12 画像処理結果の概要
Fig. 12 Result of image processing.

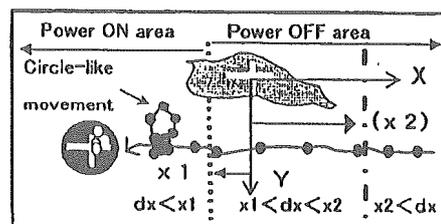


図13 エリアの設定例と虹彩(視線)の移動
Fig. 13 An example of area setting and movement of an iris.

図14に病室内の配置概要と装置の設置概要を示す。USBカメラ(Argo Lu055M-IOなど)に付属の望遠レンズにより患者の顔面付近を撮影し、画像処理で求めた値を処理してテレビの制御を行う。リモコン制御機(PATLITE PHC-100)はRS-232C経由でパソコンと接続し、本システムの利用時にテレビの正面に設置する。

4.4 画像処理の追加機能

本システムにおいて適切な処理結果を得るためには、2値化処理におけるしきい値設定を適切に行う必要がある。天候の変化(晴れ-曇り)や時間経過による日照量の変化により病室内の明るさが変わるため、図15のようにしきい値以上の部位が目尻付近にも出現ししきい値を調整する必要がある。単純に重心位置を算出するだけでは正しい虹彩の重心位置を算出できないため、未操作(電源がオンされない)になる場面もあった。

そこで、まず図16のように画像上のしきい値以上

の画素を縦方向へ加算して横方向のヒストグラムを求め、その最大値個所からヒストグラムの谷となる個所を検索し、谷個所から最大値方向の画像を虹彩とみなして重心を求める処理を追加した[3]。図16のヒストグラムでは中央部分が谷となり、左側の山部分が虹彩となる。この処理を加えて虹彩の重心を算出した結果

を図17に示す。

ヒストグラムを利用した処理の効果を検討するために、同一の2値化しきい値のもとで眼球運動を連続的に3時間程度(午後2時~5時)計測し、虹彩の重心位置が正しく算出されるかどうかをチェックした。その結果を表1に示す。表中左端は経過時間とその間の顔面付近の明るさの平均値を示す。総移動回数は患者がテレビから目をそらして図15, 図17のような画像が得られた回数、正処理数は図17のように虹彩の重心が正しく算出された回数、誤処理は図15のように

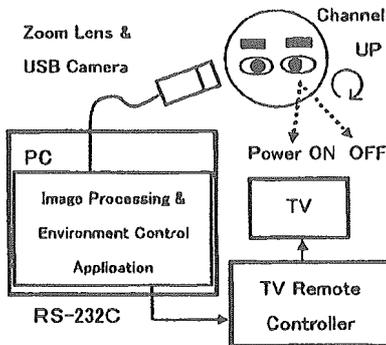
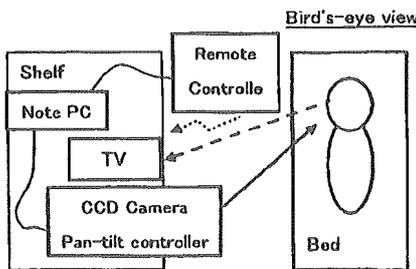
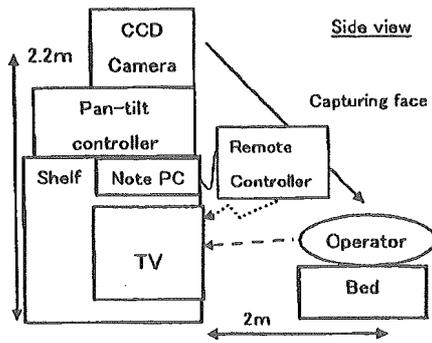


図14 病室内での配置と装置の設置概要
Fig. 14 Arrangement of measurement equipment.

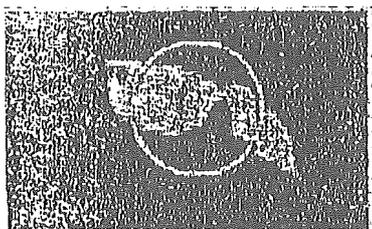


図15 虹彩以外の部分もしきい値以上になる場合
Fig. 15 Gravity was deviated from iris center.

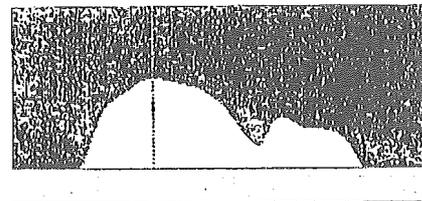


図16 しきい値以上の画素を縦方向に加算したヒストグラム

Fig. 16 Histogram means pixels more than the threshold are vertically added.

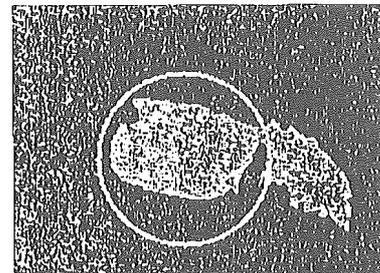


図17 ヒストグラム利用時の虹彩重心位置検出
Fig. 17 Accurate iris center of gravity is calculated with histogram.

表1 ヒストグラム処理の有無による処理結果
Table 1 Results of improved processing (Iris center of gravity is calculated with/without histogram).

経過時間(分)と明るさ(lux)	ヒストグラム処理	総移動回数(回)	正処理回数(回)	誤処理回数(回)
0~4(350)	有	11	11	0
4~6(260)	有	21	21	0
6~12(360)	無	14	0	14
12~15(350)	無	87	0	87
15~22(340)	有	89	68	21
108~123(330)	有	20	12	8
123~125(300)	無	24	0	24
161~171(260)	無	41	0	41
171~182(250)	有	78	78	0

正しく算出されない回数を示す。22~108分間は診察及び理学療法による計測の中断、及び患者が睡眠したためテレビを見ていない状態である。

ヒストグラム処理を適用すると、0~6, 171~182分間では虹彩の重心を正確に算出できている。ヒストグラム処理を適用しない6~15, 123~171分間ではすべて誤処理となっているが、虹彩と目尻の画像が連結した状態で谷個所を検出せずに重心を算出しているためである。しかし、この間ではどの移動に対しても谷個所は検出されており、ヒストグラム処理を適用すれば虹彩の重心は正しく算出できたことになる。一方、表中下線を付加した15~22, 108~123分間はヒストグラム処理を適用していても誤処理が見られる。これは、ヒストグラム中に谷個所が検出されず算出された重心が正しい位置からずれたためである。

全移動回数385回中ヒストグラム処理により谷個所を検出できなかった回数は15~22, 108~123分間の29回で、誤処理の割合は約7%程度となるが、ヒストグラム処理を適用すれば室内の明るさに合わせて常に2値化しきい値を変更する必要はなく、新規に追加した機能として有効であった。谷個所が検出できない場合には最大値個所を中心に虹彩個所を推定するなど最適な処理方法を検討する必要があるだろう。

5. む す び

比較的安価なビデオキャプチャ機器を利用し、重度身体障害者の自発的な意思表示を補助するシステムを簡易的に構築することが可能となった。本システムでは視線検出処理をソフトウェア的に行うために、処理機能とサンプリング速度の兼ね合いでどの程度の機能追加が有効かを考慮する必要がある。ヒストグラム処理の追加程度であれば大きな影響はなかったが、精度を高めるために更に複雑な処理を追加すると処理速度が犠牲になり、トータルでの操作性が悪くなることに留意しつつ追加する機能を考慮していく必要があるだろう。

パソコン本体の性能や周辺機器の性能は飛躍的に進歩しており、今後はデジタル画像による画像処理を行うことが可能になると予想される。XGA (1024×768画素)、SXGA (1280×1024画素) クラスの画像を利用し複雑な処理を行うとしても、十分実用的な処理が実現できると考えられるため、重度障害者向けの意思伝達手段に限らず、健常者の利用においても対応できるような技術が開発されることを期待したい。

今後、更なるシステムの高機能化・簡易化も視野に入れ、臨床例を増やしながら評価を継続していきたいと考えている。

文 献

- [1] 伊藤和幸, 数藤康雄, 伊福部達, “重度肢体不自由者向けの視線入力式コミュニケーション装置,” 信学論 (D-1), vol. J83-D-1, no. 5, pp. 495-503, May 2000.
- [2] 伊藤和幸, “画像センサを用いた眼球運動による環境制御システム,” 信学誌, vol. 85, no. 1, pp. 57-59, 2002.
- [3] 伊藤和幸, “ビデオキャプチャによる眼球運動計測および環境制御への応用,” ヒューマンインタフェース学会誌, vol. 5, no. 4, pp. 429-436, 2003.
- [4] 数藤康雄, コミュニケーション機器調査研究報告書, テクノエイド協会, 1991.
- [5] 山田光穂, 福田忠彦, “眼球運動による文章作成・周辺機器制御装置,” 信学論 (D), vol. J69-D, no. 7, pp. 1103-1107, July 1986.
- [6] 落合 積, 石松隆和, 高見 修, 松井稔治, “目の動きを利用した身障者用文字入力装置の試作,” 日本機械学会論文集 (C), vol. 63, no. 609, pp. 140-144, 1997.
- [7] R. John and C. Francis, “An eye movement communication control system for the disabled,” IEEE Trans. Biomed. Eng., vol. 37, no. 12, pp. 1215-1220, 1990.
- [8] <http://www.gen.co.jp/GENTECH.html>
- [9] <http://www.eyegaze.com/indexdis.htm>
- [10] <http://www.t-works.co.jp/9100.html>
- [11] L. Young and D. Sheena, “Survey of eye movement recording methods,” Behav. Res. Methods & Instrum., vol. 7, no. 5, pp. 397-429, 1975.
- [12] 久野悦章, 八木 透, 他, “EOGを用いた視線入力インタフェースの開発,” 情処学論, vol. 39, no. 5, pp. 1455-1462, 1998.
- [13] M. Mitsuho and F. Tadahiko, “Quantitative evaluation of eye movements as judged by sight-line displacements,” SMPTE J., pp. 1230-1241, 1986.
- [14] 奥山文雄, “角膜反射による眼球運動の測定,” VISION, vol. 3, no. 2, pp. 81-88, 1991.
- [15] 伴野 明, 飯田宗夫, 小林幸雄, “画像処理による非接触視線検出法に関する一検討,” 昭 63 信学秋季全大, D-I-121, 1988.
- [16] 大野健彦, “視線を用いた高速なメニュー選択作業,” 情処学論, vol. 40, no. 2, pp. 602-612, 1999.

(平成 16 年 5 月 21 日受付)



伊藤 和幸 (正員)

平元北大・工・応用物理卒。平3同大大学院工学研究科(生体工学)修士課程了。現在、国立身体障害者リハビリテーションセンター研究所福祉機器開発部に勤務。身体障害者用のコミュニケーション機器の研究・開発に従事。電子情報通信学会、ヒューマンインタフェース学会、ライフサポート学会、日本リハビリテーション工学協会各会員。



伊福部 達 (正員)

昭44北大・工・電子卒。昭46同大大学院修士課程了。同年北大・応用電気研究所助手。昭52同研究所・感覚情報工学部門教授(工博)。昭59米国スタンフォード大学・客員助教授(在外研究員)。平2北大・メデイカルエレクトロニクス部門教授。平14東京大学先端科学技術研究センター教授。約30年間にわたりヒトの感覚生理学・音響心理物理学に関する基礎研究に基づき、各種バリアフリー機器の開発を行う。著書「音声タイプライタの設計」、「音の福祉工学」、「人工現実感の評価」など。音響学会、音声言語医学会、ME学会、人口臓器学会、バーチャルリアリティ学会、ヒューマンインタフェース学会などの理事・評議員を務める。

透明文字盤 Etran 方式の PC 利用による代替について

伊藤 和幸[†] 山本 智子^{††}

[†] 国立身体障害者リハビリテーションセンター 研究所

^{††} 狭山神経内科病院

〒359-8555 埼玉県所沢市並木 4-1

Tel 042-995-3100, Fax 042-995-3132

E-mail: [†] ito@rehab.go.jp

あらまし 重度障害者が意思伝達を行う方法の一つに、透明なアクリル板や塩ビ板を用いた透明文字盤の利用がある。この方法は、介護者が障害者と透明文字盤をはさんで対面し、障害者が順次見つめる透明文字盤上の文字を介護者が透明文字盤の裏から推測して、意思の伝達を行うものである (Eyelink 方式と Etran 方式がある)。透明文字盤の利用は、視線を利用した意思伝達環境の中では安価で手軽であるが、意思伝達内容をメモ書きしなければならない等のデメリットも持ち合わせている。本報告では Etran 方式をパソコンで代用し、文章の保存や編集機能を追加して意思伝達環境を高機能化したのでその内容について発表する。

キーワード 重度肢体不自由者、支援機器、視線、透明文字盤、Etran、EyeLink

Emulate of Etran type Kana Board by use of PC

Kazuyuki ITOH[†] Tomoko YAMAMOTO^{††}

[†] Research Institute, National Rehabilitation Center for Persons with Disabilities (NRCPD)

^{††} Sayama Neurology Hospital

4-1 Namiki, Tokorozawa-shi, Saitama, 359-8555 Japan

Tel 04-2995-3100, Fax 04-2995-3132

E-mail: [†] ito@rehab.go.jp

Abstract An acrylic transparent Kana board is used when a person with severe physical disabilities communicate with caregiver. As for this method, a transparent Kana board is placed between caregiver and patient, and caregiver supposes letters on the transparent Kana board which patient gazes at sequentially (there are an Eyelink and an Etran method). Transparent Kana board is cheapest and simplest in the communication device that used eye movements, but it has some demerits that caregiver must write a memo for example. In this report, we describe about emulating of Etran method by use of PC. The PC system utilize video capture device to suppose patient's gazing letter.

Keyword Transparent Kana board, Assistive Technology, Etran, EyeLink

1. はじめに

ALS 患者や筋ジストロフィー患者などの重度障害者が介護者などとコミュニケーションをとる方法の一つに、透明なアクリル板や塩ビ版を使用した透明文字盤の利用がある。この方法は、透明文字盤には 50 音表や単語 (テレビ, 吸引 (痰の吸引), 体交 (体位交換) など) を記入し、介護者が障害者と透明文字盤をはさんで対面して、障害者が順次見つめる透明文字盤上の

文字や単語を介護者が推測して意思の伝達を行うものである。視線の読み取りにはある程度の慣れが必要であるものの、障害者の様子 (表情や目の様子) を見ながら作業を行うため、単純な文章作成作業という意味だけでなく、意思を伝達するというコミュニケーションの本質を実感できる方法でもある。

障害者の視線を利用して直接的に文字を選択していく方法は、50 音の文字盤を利用した走査式 [1]-[2]

(文字盤上の文字を介護者が指し示すか読み上げ、望みの箇所まで障害者が合図することで1文字を確定する方法)よりも効率の良い意思伝達が可能である。また、アクリル板のコストは数千円程度であり、意思伝達を行いたい際に透明文字盤を用意するだけなので複雑なセッティングが不要となる、というメリットもある。また、コミュニケーション場面により(文章の作成、定型句の選択等)、適宜文字盤を変えることで障害者の目的に即したコミュニケーション環境を素早く整えられるというメリットもある。

一方、デメリットとしては長い文章を作成したい場合には選択した文字を介護者が記憶するか、メモ書きして保存する必要があることである。さらに、一度作成した文章を後日編集したい場合には、ワープロなどに保存しておく作業も必要となる。つまり、透明文字盤はその場限りの短い意思伝達には有効であるものの、長い文章の作成や編集機能に関しては機能不足であることを否めない。

本報告では、この編集機能を補うためにパーソナル・コンピュータ(以下、パソコン)を利用し、USBカメラを利用したビデオキャプチャにより画面上に障害者の目の画像を表示して、視線を読み取ることのできるEtran形式のシステムを開発したのでその内容について記述する。

2. 透明文字盤を利用した意思伝達方法

障害者に発声機能があれば音声によるコミュニケーションを行うことが基本であるが、発音が不明瞭になったり気管切開で発声できない場合には代替手段の検討が必要となる。透明文字盤の利用も意思伝達の一手段であり、障害者のニーズ(意思の伝達、パソコンの利用など)や、介護状況などにより最適なものを選択することが重要である。

透明文字盤を利用し視線の方向で意思を伝える方法には、EyeLinkとEtranと呼ばれる方式がある[3]-[4]。EyeLink方式は障害者の見つめる文字と介護者の視線が一直線になるように(お互いに見つめあう: eye contact)、間にある透明文字盤を動かす方法で、図1のように障害者の意図する文字や単語・シンボルが目と目を結ぶ線上に移動してくると、次第に文字の向こう側に正面向きの相手の目が見えることになる。Etran方式は障害者の視線方向を利用する方法で、図2のように中央部を空白とした透明文字盤の周辺に選択肢を配置し、障害者の見つめる選択肢を介護者が推測していく方法である。文字を選択する場合、日本語であれば46音の清音、濁点、半濁点、長音記号などの選択肢が必要で、全てを表示すると視線方向がどの文字を示しているかが判断しづらい。そこで、EyeLink形式

では文字盤を大きくして1文字分の範囲を大きくするなどの工夫が必要となる。Etran形式では、文字の配置を上下左右に斜め方向を加えた8方向などとし、最初に「あかさたな・・・」(子音)を選択してから、母音を選択するなど、2段階で1文字を選択する方法が採用されている。母音の位置は予め決めておき(例えば図2内の○位置)、最初に選択した子音に合わせて介護者が母音を読み上げる方法を採用する。

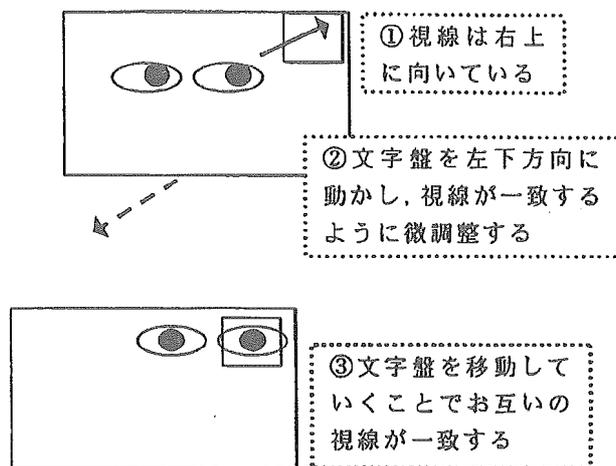


図1 EyeLink形式の目の動き

Fig.1 Eye movement in EyeLink method

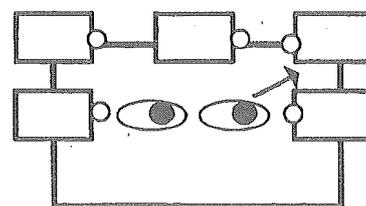


図2 Etran形式の目の動き(○は母音位置)

Fig.2 Eye movement in Etran method

障害者の目の動かし方と介護者の読み取り方は以下のようなになる。

EyeLink形式:

- 1 障害者: 選択したい箇所を見る。
介護者: 障害者の視線と自分の視線が一直線になるように文字盤をゆっくり動かす。
- 2 介護者: 障害者が見ていると思われる箇所を読み上げる。
障害者: 介護者の読み上げた箇所が合っていれば何らかの合図をして、次の文字を見る。
- 3 読み取られた箇所が違う場合
障害者: 選択したい文字を見続ける
介護者: 障害者の合図がない場合には、文字盤の位置を少しずつずらして視線を再度確かめる。

Etran 形式：

- 1 障害者：選択したい箇所を見る。
 介護者：障害者が見ていると思われる箇所を読み上げる。
- 2 障害者：介護者の読み上げた箇所が合っていれば何らかの合図をして、次の文字を見る。

Etran 形式では、選択したい箇所が分かりにくい場合には、障害者とその箇所の更に外側へ視線を移動させると視線の向きがはっきりするので読み取りが容易になる。または、介護者が視線方向付近の箇所を幾つか読み上げ、望みの選択肢が読み上げられた時に合図を送ることで目的の箇所を特定することができる。

3. パソコンを利用した Etran 形式の代替

パソコンを利用し、ビデオキャプチャした画像を画面上に表示させることで前述の方法のうち Etran 形式を代替することが可能となる。具体的には、眼球付近の画像をビデオキャプチャし、反転させた上で画面の中央に表示させるアプリケーションを作成する。選択する文字は画面の周辺に配置しておく。障害者は画面上に配置された文字を見つめ、介護者は画面中央に表示された目の向きを推測し、障害者が画面上のどの文字を見つめているかを読み取る、という方法となる。透明文字盤を利用する場合には障害者と介護者が対面していたが、介護者は障害者の脇に位置して障害者と介護者は同じ方向を向くことになる。

図3は透明文字盤を利用した場合の視線方向を表示したものであり、左が障害者側から見たもの、右は介護者側から見たものである。介護者側から見ると文字は裏側から見ることになる。

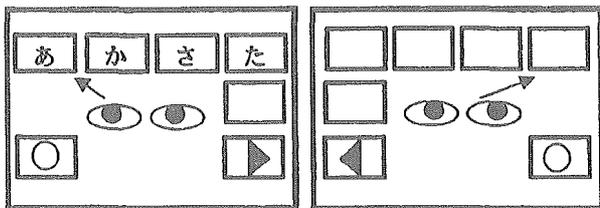


図3 障害者側 介護側から見た様子
 Fig.3 Line of eye on transparent Kana board

図4はパソコンを利用した場合の表示内容であり、起動しているアプリケーションは片目付近の画像をビデオキャプチャし画面中央に反転して表示する。眼球付近の画像を反転表示することで文字の位置と視線方向が一致し、透明文字盤を挟んで障害者と対面しなくても視線方向を推測することができる。介護者はマウスを使用して画面周囲に配置した文字上にカーソル（文字を囲む枠）を移動させ、介護者が判断している文字がどの文字であるかを障害者にフィードバックす

る。障害者は望みの文字がカーソルで囲まれていれば目を閉じる等のOKの合図を出せばよい。子音の選択後、周囲の文字は母音表示に換え、再度障害者に望みの文字を見つめてもらい、最終的に1文字が決定する。

アプリケーションでは、子音とその後の母音の選択は全て左クリックで行い、確定された文字が画面下部のワープロエリアに入力される。ビデオキャプチャ機能で目の画像を反転表示するとともに、文字編集機能（簡易テキストエディタ）によりアプリケーション上で選択した編集キー（漢字やカタカナへの変換、カーソルの移動、文字削除）を選択することで文字編集が可能となっている。簡易テキストエディタはキーボードからの直接操作も可能である。

文書の保存や読み込み、新規作成を行うことも可能であるが、メニュー操作中には文字盤が隠れるため、ファイル名などは予め打ち合わせる必要がある。

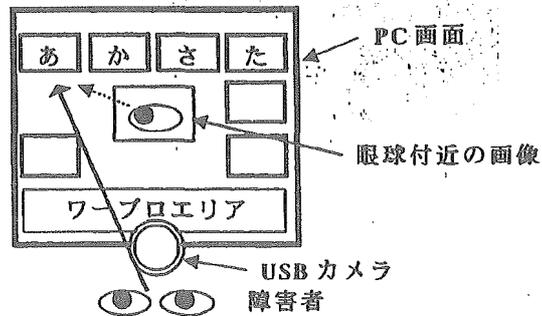


図4 P-C版 Etran 画面

Fig.4 Etran screen on PC

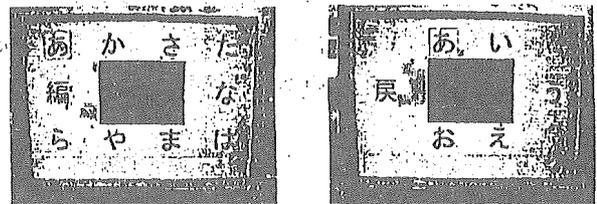


図5 子音「あ」の選択（左）と選択後の母音の表示（右）（「あ」の選択場面）

Fig.5 1st(consonant) and 2nd(vowel) screen on PC

システムはノート型の Windows パソコン（OS は Windows XP, Visual C++によりアプリケーションを作成）と USB カメラで構成され、短時間の利用であればバッテリー駆動により電源ケーブルが必要となり、コンパクトな構成となっている（図6、図7）。USBカメラは眼球付近の画像を拡大撮影する必要があることから Web 用の広角カメラではなく、望遠レンズや接写リングの選択が可能で C マウントタイプ（例：Argo, Lu055M-10）のものを利用した。適切な保持具が用意で

できれば、カメラと小型ノートパソコンか液晶モニターをベッド上に設置して仰向けのままでも利用が可能となる。

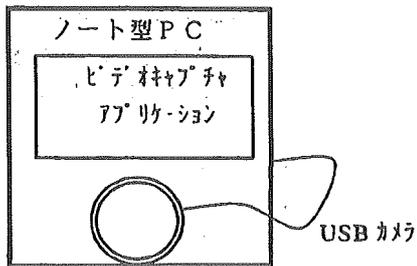


図6 システム構成図

Fig.6 System configuration

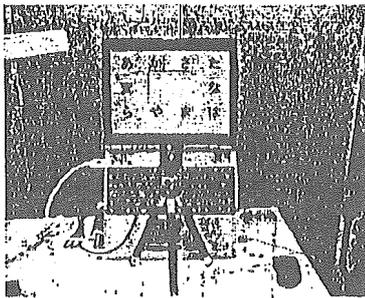


図7 システムの概観

Fig.7 General system view

4. まとめ

PCを利用し簡易な構成でEtran形式の透明文字盤を代替できるシステムを開発した。コスト的にはアクリル製透明文字盤の利用に勝ることはないが、透明文字盤の持つ欠点を補う機能が実現できたと考えられる。今後、臨床現場で評価を行い本システムに不足する機能を検証し、より実用性のあるものへと改善していきたいと考えている。

文 献

- [1] 数藤康雄, “コミュニケーション機器調査研究報告書”, テクノエイド協会, 1991
- [2] 小澤邦昭, 安藤研吾, 松田泰昌, 長谷川司, 岡高志, 安藤肇夫, 植野一政, 古和久幸, 長谷川一子, 斎藤豊和, “ALS患者向け意志伝達装置「伝の心」の開発” 第12回リハ工学カンファレンス講演論文集, pp.91-96, Aug.1997
- [3] 山本智子, “視線コミュニケーションの基礎 EyeGazeで文字を伝える EtranとeyeLink”, ATACカンファレンス2001テキスト, pp.26-27, Nov. 2001
- [4] 山本智子, 眼球運動が障害されたALS患者が使用

可能な透明文字盤の工夫, 第16回リハ工学カンファレンス講演論文集, Vol.16, pp.105-108, Aug.

重度身障者のための1ボタン自然言語入力システムに向けて

田中久美子[†] 伊藤和幸^{††}

[†] 東京大学情報基盤センター 〒113 文京区本郷 7-3-1

^{††} 国立身体障害者リハビリテーションセンター 〒359-8555 埼玉県所沢市並木 4 丁目 1 番地

E-mail: †kumiko@r.dl.itc.u-tokyo.ac.jp, ††ito@rehab.go.jp

あらまし 本稿では、モールス信号に基づく重度身障者のための1ボタン入力システムを報告する。単なるモールスコードに基づく入力システムに、単語の予測を導入し、さらに予測された単語候補のうち、候補選択により入力時間を削減しないものの刈り込みを行なったところ、入力効率の大きな向上が見られた。

キーワード 入力システム、モールス符号

Towards An Entry System For the Seriously Impaired

Kumiko TANAKA-ISHII[†] and Kazuyuki ITOH^{††}

[†] Information Technology Center, 7-3-1, Hongo, Bunkyo, Tokyo, 113 Japan

^{††} National Rehabilitation Center for Persons with Disabilities, 1, Namiki 4-chome, Tokorozawa Saitama, 359-8555, Japan.

E-mail: †kumiko@r.dl.itc.u-tokyo.ac.jp, ††ito@rehab.go.jp

Abstract We empirically demonstrate the efficacy of predictive text entry in a new context: the input of Morse code with a single key. Further, we present a new method of pruning, that outperforms simple text prediction. Our overall motivation is contributing to the development of efficient entry systems for the seriously disabled, but we also draw general conclusions with much wider implications.

Key words Text Entry System

1. はじめに

計算機技術は、健常者だけでなく、身障者の環境をも大きく変化させている。一般に、身障者の環境を改善する問題は、健常者用のそれよりも大きな制約があるため、難しいことが多い。しかし、それを解くことにより、身障者だけでなく、健常者の問題に対して新たな提言を与えることがある。本稿では、1ボタン入力という極度に制限された状況下での入力を考えることにより得た、一つのユニバーサルな技術となりうる言語処理例を示す。

本稿の問題は、1つのボタンだけを用いた入力である。このようなシステムは、身障者の中でも ALS 患者や、頸椎損傷を負った患者により実際に使われている。これまでも、国立身体障害者リハビリテーションセンターでの研究 [2] や、台湾 [4]、他に米国での研究も見られる。既存の研究では、モールスコードを用いたものがほとんどであるが、未だ予測入力は導入され

ておらず、ユーザは長い単語であっても最後までモールスコードを用いて入力しきらなければいけないという問題があった。そこで、本稿では既存の研究に、

- 予測入力機能を付け加え、
- 候補選択により入力時間が短縮されないものを刈り込む工夫を行なった。実際に英語に関してこのようなシステムを実装し、どの程度の入力効率の向上が見られるかを検証した。

2. システム概要

1ボタンで入力を行なうには、まずは1ボタンの出力を、言語のアルファベットなどに何らかの方法で展開する必要がある。その方法はさまざまに提案されているが、本稿では過去の多く試みと同様、モールスコードを用いるものとした。モールスコードはアルファベットの単純な文字頻度をもとに作られている。とはいえ、今でも無線通信において事実上の標準として使われている上、プロの入力は1分間あたり25単語と高速な

表 1. モールスコードにおける入力の意味

時間	入力	間
1 ユニット	●	入力間
3 ユニット	—	文字間
7 ユニット	未定義	単語間

表 2. モールスコードによるアルファベット

char	code	char	code
A	●—	N	—●
B	—●●●	O	— — —
C	—●—●	P	●— — ●
D	—●●	Q	— — ● —
E	●	R	●—●
F	●●—●	S	●●●●
G	— — ●	T	—
H	●●●●	U	●●—
I	●●	V	●●●—
J	●— — —	W	●— —
K	—●—	X	—●●—
L	●—●●	Y	—●— —
M	— —	Z	— — ●●

入力を可能とすることから、本入力においてもこれを基本とした。

モールスコードを用いた入力は、ボタンを異なる長さで押し分けることにより行なわれる。入力は最短の長さを1ユニットとすると、3ユニット分の長さで7ユニット分の長さが定義されており、その長さ分の入力と間(入力しないこと)それぞれに意味付けが行なわれている(表1参照)。通常のアルファベットの入力には、3ユニット分の長(本稿では—として記述)と短(●)を表2の表にしたがって押し分けることにより行なわれる。「間」に対する意味付けとしては、入力間には最低1ユニット分の間を空ける必要があるし、3ユニット分の間を空けると文字間の間、7ユニット分の間を空けると、単語間の間と解釈される。たとえば、「SOS」の入力は、

●(休み1)●(休み1)●(休み3)
 —(休み1)—(休み1)—(休み3)
 ●(休み1)●(休み1)●(休み7)

と入力する。以上はモールスコード上の定義である。

図1に、Java言語で実装された本入力システムのGUIを示す。下方のPressボタンを押すことにより入力が行なわれ、結果のテキストが上方の文書ボックスに、中央の候補ボックスにはユーザが選択可能なコマンドが、下方横長の窓(Pressボタンのすぐ上)には、現在ユーザが入力したモールス信号が表示される。左図においては、すでに“The human communications”という文章が入力された状態を示す。通常は、コマンドモードとなっており、候補ボックスにはコマンドが表示されている。ユーザは、ここで“revolution”の入力を行なうものとする。

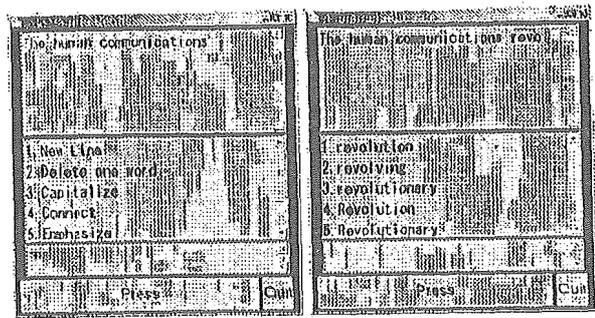


図 1 システムの GUI. コマンドモードと、補完モード

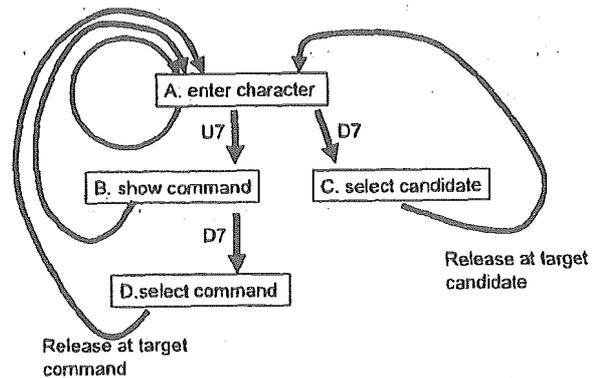


図 2 システムにおけるモード変化

表2にしたがって、アルファベットの入力を“revo”まで行なったところの図が右図である。候補ボックスには目的の候補が1位に挙がっている。このように、本システムでは、単語のprefix部分を入力すると単語全体が候補として挙がる予測を用いて入力を行なう。

候補ボックスから望みの候補を選択するには、ユニット7の長さでpressボタンを長押しする。すると、候補ボックスの候補のハイライトが始まり、一定の速さで候補を走査する。ハイライトされているときに、ボタンを放すと、その候補が入力される。

このやり方は、左の画面で、コマンドを選ぶことにおいても同様である。左の図でPressボタンを7ユニットの長さで押すと、候補ボックスのNew Lineコマンドがハイライトされる。さらに押しつづけると、つぎつぎにコマンドがハイライトされる。目的のコマンドのところではPressボタンを放すと、そのコマンドが実行される。モードを切り替えを7ユニットで押しつづけることにより行なうことは、それがもとのモールスコード上で未定義であったことを利用している(表1参照)。

ボタンを押すことによって、どのようにモードが変化するかを図示したのが、図2である。D7は7ユニット分ボタンを押している状態、U7は7ユニット分ボタンを放している状態を表している。一般に、モードには選択モードと、入力モードの二つがあり、通常は入力モードとなっている。選択モードに切り替えるには、Pressボタンを7ユニット以上押しつづける

る。選択モードから入力モードへは、候補ボックスから選択が行なわれると自動的に入力モードへと切り替わる。

本予測入力システムの特徴は、辞書にない単語であっても、入力が可能である点である。というのも、モールス信号で文字を入力することは、常に可能であるからである。一般に辞書にない単語をいかに入力するかという点は、予測入力やかな漢字変換においては最大の論点となるため、本システムはこの点の一つの利点となっている。とはいえ、特に長い単語などは予測を用いた方が入力の効率を向上させられる可能性があるし、本稿では重度身障者を対象としているため、予測を用いることによる効果が期待される。以下では予測の効果を確かめる。

3. 予測入力

3.1 基本モデル

予測では一般に入力 S に対して最大の条件付確率を与える語 W を最良の候補とする。

$$\operatorname{argmax}_W P(W|S) = \operatorname{argmax}_W P(S|W)P(W) \quad (1)$$

$P(S|W)$ は入力誤りがないと仮定すると 1.0 となるため、結局 $P(W)$ 、すなわち言語モデルが候補の良し悪しを決定する。 $P(W)$ の予測についてはさまざまなものがあり、単純な頻度から、高次の n -gram を利用するもの、動的に n -gram を推定する PPM を利用するもの、品詞列を利用した HMM を用いるものなどさまざまなものが提案されている。

本稿では、最も単純な頻度によるモデルを用いる。というのも、つぎに述べる刈り込みを論じるには、高次のモデルよりは、単純なモデルの方がそのふるまいが明確であるからである。とはいえ、次章で述べる刈り込みは、高次の言語モデルを用いたとしても併用できるものである。

3.2 候補の刈り込み

刈り込みによる工夫の基本アイデアは [3] に述べられているものである。一般に入力においては、それぞれの候補の入力にかかる時間は異なる。たとえば、“sh” が入力済みで、she, shy, should を入力したい場合について考えてみる。she の場合には、e を入力するためには、1 ユニット (e の入力分) しかかからない。一方で、候補から選択する場合には、モードの切り替えだけで最低 7 ユニットかかってしまう。したがって、she については、ユーザが e を直接入力してしまった方が早いことになる。

一方、shy であるが、y の入力は、13 ユニットである。というのも、—・—・— は、長に 9 ユニット、短に 1 ユニット、そして、長と短の間に 3 ユニットかかるからである。同じように考えれば、should は、47 ユニットかかる。というのも、o は —・—・— は 11 ユニット、n は —••— で 7 ユニット、l は —•— で 13 ユニット、d は —••7 ユニットで、さらに文字間に 3 ユニットづつおかなければならないので、 $3 \times 3 = 9$ ユニット足して計 47 ユニットである。とすると、should は予測により入力

を行なう方が直接入力を行なうよりも、時間を短縮できそうであることが予想される。

一般に、現在ユーザの入力した文字列を prefix とする単語 s を選択するのにかかる時間を、 $t_{select}(s)$ 、suffix 部分を入力する時間を $t_{enter}(s)$ とする。すると、

$$t_{enter}(s) < t_{select}(s) \quad (2)$$

が成り立つ場合には、候補リストから s を削除することにより、入力が困難な単語の候補順位が繰り上がり、さらに入力時間を短縮できる。そこで、ユーザに候補を提示するたびに、候補を高順位のものから走査して、上の条件を満たす候補を削除する。

$t_{enter}(s)$ については、should で示した例と同様に計算する。一方で、 $t_{select}(s)$ については、その候補が i 番目の候補として挙がっているものとする、

$$t_{select}(s) = 7 + (i - 1) \times dt \quad (3)$$

として計算する。ただし、7 ユニットはモード切替の時間、 dt を i 番目から $i+1$ 番目の候補に移動するユニット数としている。

一般に、刈り込みの効果は、ユーザが見る単語の数 ($NthMax$) によって異なる。 $NthMax$ が 1 である場合には、刈り込みの効果はあまり得られない。一方で、 $NthMax$ が大きい場合には、刈り込まれて繰り上がる候補が多くなる一方で、たくさんの候補をユーザが走査することになる。この点に関する詳細は、[3] において分析されている。

ここで述べた刈り込みの方法は、ある文字列を入力する際に、複数種の入力方法があるとき、その効率をシステムが計測し、効率の小さい方の入力を選択するようにユーザを仕向けることに相当する。実際、本システムにおいては、ある単語を入力するには、

- 直接モールス符号で入力を行う。
- 候補から選択する。

の二つの方法がある。前者の効率が後者の効率を下回るときには、前者でユーザが入力を行わないように、ユーザをガイドしていることになる。本稿の範囲では、モールス符号をそのまま用いた入力方法となっているが、今後、その点を変更したとしても、同様の考え方は用いることができる。

また、同様のアイデアは携帯電話における通常の英語入力や、かな文字列の日本語入力にも適用できることから、ユニバーサルな技術として発展させられる可能性があり、今後の課題である。

4. 実験

4.1 刈り込みの効果

本節では、予測入力の効果、および刈り込みの効果を計算機上の実験で確かめる。

表3 単なる予測入力、および刈り込みによる効率の向上 (ユーザに示す候補数が5のとき)

テスト コーパス	入力 単語数	入力時間の短縮率		キー入力数の減少率		単語の選択率	
		予測+刈込	予測のみ	予測+刈込	予測のみ	予測+刈込	予測のみ
ACL03の論文	4287	18.60%	14.67%	34.8%	35.8%	59.51%	77.96%
トムソーヤの冒険	4714	13.60%	8.08%	28.76%	29.99%	50.36%	71.64%
アメリカ合衆国憲法	4973	17.5%	13.59%	36.80%	37.35%	67.04%	82.99%
個人の電子メール	6567	18.04%	13.84%	35.68%	35.84%	59.08%	75.22%

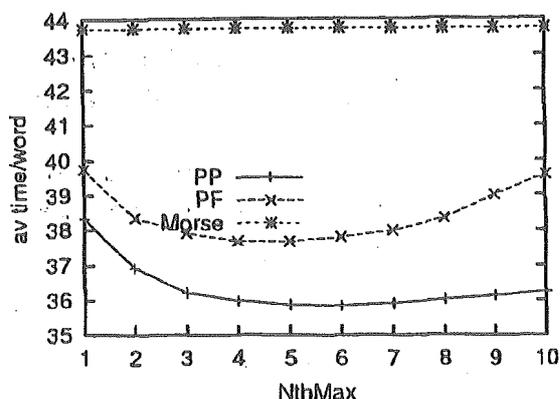


図3 ユーザに示す候補数に対する1単語あたり入力に必要な時間(電子メールの場合)

まず、予測に用いる辞書には、WSJの30Mバイト分の全単語を用いた。入力用のデータには、表3の第一列に示した4つのものを用意し、うち5000単語程度を入力した。候補をスキャンする速度 dt は3ユニットとして設定した。

ユーザに示す候補数を5とした場合の結果を、表3に示す。まず、第四列を見ると、予測だけでも、15%程度の時間上の効率向上が見られる場合がある。最終列を見ると、上位5単語を見せるだけで、実に8割近くの単語が候補から選択されていることがわかる。その前の列(第7列)を見ると、刈り込みを行なうと単語の選択率が減ることがわかる。刈り込まれた単語は候補には挙がらないため、当然の結果である。しかし、第三列と第四列を比較すると、入力時間は5%前後早くなっている上、第五、六列の比較から、この効果が、キー入力数をほとんど変更せずに得られていることがわかる。

つぎに、ユーザに示す候補数を変更した場合の時間効率を図3に示す。横軸は、ユーザに示す候補数、縦軸は1単語あたりの入力に必要なユニット数である。折れ線は、上からモールス信号のみの場合、予測のみの場合、予測と刈り込みを行なった場合である。予測のみでは、ユーザに示す候補数が増やすぎると、時間効率が悪くなることなることがわかる。一方で、刈り込みを用いると、時間効率が向上しない候補が削除されるので、示す候補を多くしても時間効率が悪くはならないことがわかる。

4.2 健常者による小実験

健常者による小実験を行った。30分を1セットとして、第一著者が2セット分の入力を行い、速度を測定した。1回目は

1分あたり4.5単語程度、2回目は1分あたり5.2単語程度となった。第一著者はモールス符号を記憶してはおらず、モールス符号の表を見ながらの入力実験となった。したがって、モールス符号初心者であれば、この程度の速度を出すこととなる。

一般に、アマチュア無線においては、(予測入力を用いずに)モールス符号での入力は1分あたり5単語程度と報告されている。今後はより詳細な実験を行い、予測入力の効果、刈り込みの効果などを入力実験によって確かめていきたい。

5. モールス符号の問題

モールスコードをそのまま踏襲すると、さまざまな問題が起こる。

第一の問題は、頻繁に使うコードが長いことがある点である。たとえば、一文字消去は `dot` を6回押すことになっている。モールスコードに慣れていない人が初めて本システムで入力をはじめるときには、頻繁に間違えるために、訂正を何回も押すことになる。その際、この割り当ては問題となる。

第二の問題は、単語の切れ目の同定である。モールス符号に従うならば、7ユニット以上の時間を空けると、単語の切れ目となってしまふ。この制限は、単語内ではユーザが7ユニットの間を空けることなく符号を打ち続けること強いられ、心理的負担が大きい。

一つの解決策として、1文字消去や単語切れ目を別の短い符号を割り当てることが考えられるものの、モールス符号では短い符号はすべて高い頻度で入力されるアルファベットに割り当てがあるため、再割り当てが難しい。

本稿の範囲では、モールス符号をそのまま踏襲する方向で論じており、これらの問題に対処するには到ってはいない。現在解決方法を模索中である。

6. 結 論

本稿では、モールス信号に基づく重度身障者のための1ボタン入力システムを報告した。既存のモールス信号に基づく入力システムの研究に加えて、本稿では新たに予測入力を導入し、さらに候補のうち入力時間の短縮をもたらさない候補を削除する刈り込みを行なったところ、入力効率の向上が見られた。本工夫は、現在のところは英語での効果のみ検証されているが、日本語のシステムや、一般の予測入力のシステムに適用可能であるため、この点が今後の課題となる。