

6 Summary and Outstanding Issues

With the present system, because a laser pointer is illuminated onto a site where a mouse cursor is to be moved, gain adjustment is not required. The main problem is to combat noise such as the reflection of indoor light captured on the screen, so the setting position of the USB camera and an appropriate threshold setting are important. However, in a usage situation the demand is to reduce as much as possible the time and effort associated with settings, so it is necessary to further stabilize rough setting environments and provide a function for detecting point sites. The next step is to conduct a clinical evaluation among SCI subjects and investigate ways of improving the system.

References

1. Yasuo. S., Tadayuki S.: Improvement and Clinical Evaluation of Lightspot Operated Keyboard, Nat. Rehab. Res. Bulletin, no.10, (1989)115-119 (in Japanese)
2. http://www.kokoroweb.org/index_device.html (in Japanese)
3. <http://www.madentec.com/>
4. <http://www.naturalpoint.com/trackir/>
5. Edmund. F. L. and David. M. B.: Adaptive Software for Head-Operated Computer Controls, IEEE Trans. on Neural System and Rehabilitation Engineering, vol.12, no.1, (2004)102-111

レーザー光線を利用した頸髄損傷者向けマウスポインティングデバイス

伊藤 和幸†^{a)}

Light Spot Operated Mouse Pointing Device for Cervical Spinal-Cord Injured PC Users

Kazuyuki ITOH†^{a)}

あらまし 当研究所では、これまでに頸髄損傷者向けのパソコン入力装置としてレーザーポインタを利用した光キーボードを開発している。しかしながら、開発コンセプトはあくまでもキーボード代用装置としての位置付けであり、このシステムを利用したマウス操作ではユーザインタフェースとして最適でないという課題が残っている。一方、頸髄損傷者向けのマウス代用装置として頭部の動きをマウスカーソルの動きへと連動するヘッドマウスポインタが市販されており、いくつかの市販品を光キーボードユーザに導入・使用評価を行った結果、ゲイン調整が十分でないために画面の周辺にマウスが移動しないというケースがあり、適切なマウスポインティングデバイスの開発が望まれていた。本研究では、キーボード入力に使用しているレーザーポインタを活用し、液晶画面に当たったレーザー光線の照射箇所にもマウスカーソルを移動させるシステムを開発することでマウスユーザインタフェースの改善を試みた。

キーワード 頸髄損傷者、光キーボード、偏光フィルタ、ヘッドマウスポインティングデバイス

1. ま え が き

情報通信技術 (Information Technology: IT) の進歩は近年特に著しく、パーソナルコンピュータ (以下、パソコン) や携帯電話を利用したメールやインターネットなどの新しいコミュニケーション手段の急速な普及によりマルチメディア社会が実現している。IT化の波は障害者を取り巻く環境にも着実に浸透し、移動における不都合が日常生活の支障となっている重度身体障害者であっても、IT 機器を活用することで時間や空間の制約を受けることなく、在宅就労による自立生活や社会への参加を実現している例が見られる。

IT 機器の代表的装置であるパソコンを利用するための標準入力装置はキーボードとマウスであるが、重度身体障害者にとってはその操作に困難を伴うことが多い。しかしながら、障害者の身体動作に対応した適切な代用装置が選択できれば、パソコンを使用することが可能となる。

高所からの転落や交通事故などにより脊髄に損傷を受ける外傷性脊髄損傷者のうち、高位の脊髄を損傷する頸髄損傷 (四肢まひ) では、脳から脊髄下部への信号伝達が遮断されることによる四肢の運動まひと、抹消における感覚受容器における知覚まひを呈し、身体の位置覚や触覚、痛覚などが喪失する。

頸髄損傷者がパソコンを利用するには、四肢によるキーボードやマウスの操作に困難を伴うため自助具や代用装置を使用することになる。ヘッドスティックやマウススティックによるキーボード操作は、頸髄損傷者の残存能力が高く、かつ頻繁な頷き動作によって頸部に起こる二次障害や歯の咬合への影響がなければ最も安価な対応策である [1]~[4]。

当研究所でも、これまでに頸髄損傷者向け MS-Windows パソコン入力装置としてレーザーポインタを利用した光キーボードを開発している [5]。しかし、この装置の開発に着手したのは OS が MS-DOS のころであり、開発コンセプトとしてはキーボード代用装置としての位置付けである。光キーボードは臨床評価・継続利用され、キーボード代用装置としての有効性は認められている [6]。しかし、その後 OS は MS-Windows に移行し、GUI (Graphical User Interface) が主流

† 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
a) E-mail: ito@rehab.go.jp

の利用環境ではマウスを利用する頻度が非常に高くなってきた。

光キーボードシステムにおけるマウスカーソルの移動操作は、後述するように各移動方向（上下左右と斜め方向の8方向）に対応する光センサを画面の周辺8箇所配置し、各光センサにレーザ光線を照射することで行っている（以下、光センサ式と呼ぶ）。そのため、移動の方向によっては顔（頭部）の向きと視線方向が一致しない場合があり、ユーザインタフェースとしては最適でないという課題が残っている。例えば、右端にあるマウスカーソルを若干左へ移動させたい場合、視線は右方向であるが左側のセンサにレーザ光線を当てるために顔は左を向かなければならない、というケースである。画面の周辺8箇所それぞれに8方向移動用のセンサ（合計64個のセンサ）を配置させるという解決策も現実的ではなく、また、カーソルの移動速度の設定によっては長い距離を移動させる場合には到達までに時間がかかる、という欠点もある。

そこで本研究では、上記の欠点を克服し、光キーボードユーザに対するマウスユーザインタフェースを改善することを目的とする。光キーボードユーザが使用しているレーザポインタをキーボード入力に利用すると同時に、マウスカーソルの移動にも活用する。

利用者はマウスカーソルを移動させたい箇所にレーザ光線を照射し、その照射箇所にマウスカーソルが移動するようなシステムであれば、利用者の意図（マウスカーソルの移動）は直接反映され、しかもマウスカーソルの移動に際して顔（頭部）の向きと視線は一致するので、前述したような光センサ式におけるユーザインタフェースを改善できる。レーザ光線を照射した箇所にマウスカーソルが直接移動するので、光センサ式による移動に比べてマウスカーソルの移動時間の短縮も見込むことができる。

以下、本論文では、2.で開発に対する背景について述べ、3.で画像処理による液晶画面上のレーザ光線照射箇所の検出方法、及びマウスカーソルの移動方法、4.で本システムの使用評価内容について述べる。次いで、5.で評価結果を示すとともに考察を行い、6.でまとめと今後の課題について述べる。

2. 開発に対する背景

2.1 既存のマウス代用装置

頸髄損傷者向けのマウス代用装置に関しては、いくつかの研究報告とともに[7]~[12]、市販品の利用も可

能である[13]。市販品には、光センサ式と同様に移動方向ごとのスイッチ操作によるマウスカーソル移動用デバイス、チンコントロールによるジョイスティック型のマウスコントローラ、ヘッドマウスポインタ（Track-IR, Tracker-Oneなど）のほか、スキャンタイプ（マウスカーソルの各移動方向、左右クリックなどが順にスキャンされ、目的箇所がスキャンされたときにタイミングを合わせてスイッチ操作を行うことで、各操作が実行される）も考案されている。

発声・発語機能に障害を受けなければ、音声認識技術による文字入力やコマンド操作、スイッチ操作が可能であり[14]、舌操作[15]、[16]や表情筋の筋電図[17]を利用することでマウスカーソル移動を実現している事例もある。

いずれにしても頸髄損傷者の残存機能に合わせた最適な代用装置を提供することが重要であるが、ヘッドマウスポインタ[7]、[8]、[10]~[12]は、マウスカーソルの移動が頭部の動きに連動して直接的に移動する方法であるため、頭部の制御性が良い場合にはマウスカーソルの移動速度面、マウス操作性において優れている。

しかし、ヘッドマウスポインタを使用する際にはゲインを設定し個々の頸髄損傷者の頭部の動きに対応したマウスカーソル移動量を調節する必要があり、利用者の頭部の可動範囲に対してゲインの設定が適切にできない場合には問題が生じる。

2.2 光キーボードユーザへの適合

本システムを開発した背景には、光キーボードユーザ（頸髄損傷者）に対して市販のジョイスティック型マウスコントローラ、Track-IR及びTracker-Oneの導入を検討したものの、以下の理由により導入できなかった経緯がある。

(1) 光キーボードユーザはパソコン使用中に度々リクライニングを行うため、ジョイスティック型マウスコントローラを使用した場合には、リクライニングのつどその姿勢に合わせてコントローラを固定し直す必要が生じ、介護負担が増える。

(2) Track-IR及びTracker-Oneは、利用開始時に厳密なキャリブレーション作業は必要なく導入は比較的容易である。画面の中央部付近における操作性は良いが、画面との視距離が遠い場合、画面の周辺、特に4隅へマウスカーソルを移動させるには頭部を大きく外側に向けないと移動しない（機器に備えてあるゲイン調整機能では対応できない）ケースがあり、光キーボードユーザの可動範囲を超えてしまうという問題

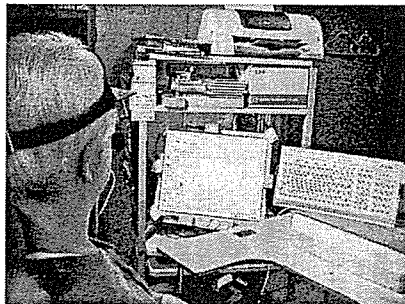


図1 光キーボードユーザによるパソコン使用場面
Fig.1 Utilizing scene of LSO-Keyboard.

がある。具体的には、光キーボードユーザは19インチ液晶画面を視距離約90cmで使用しており(図1)、Track-IRやTracker-One本体を液晶画面上に設置すると備え付けのゲイン調節機能では対応できない。

(3) Track-IRでは、液晶画面に正対した状態で規定のキー入力を行い、その位置を基準状態とする簡易キャリブレーション作業が必要となる。光キーボードによりキー入力を行う場合、キー入力の際には液晶画面には正対しない(光キーボードは液晶画面右位置にあり、頭部は右方向を向く)ため、簡易キャリブレーションそのものが不可能となる。

(4) Tracker-Oneでは、しばらく利用していると頭部の動きとマウスカーソルの動きにずれが生じるようになる。それを修正しようとする、利用者の頭部可動範囲を超えた動きが要求される(更に大きく外側を向くことで調整を行う必要がある)。

これらの理由に加えて、光キーボードユーザはこれまで光キーボードによるキーボード操作に習熟しているため[6]、レーザポインタを活用したマウスポインティング機能が実現できないかという要望があった。

3. 画像処理による液晶画面上のレーザ光線照射箇所の検出

本章では、ソフトウェア画像処理による液晶画面上のレーザ光線照射箇所の検出方法、及びマウスカーソルの移動方法について述べる。

3.1 偏光フィルタの利用

パソコン液晶画面には偏光フィルタが組み込まれており、同じフィルタを90度回転させたフィルタを通して撮影すると、液晶画面のみが黒く撮影される。図2、図3は本システム用のアプリケーションを起動させ、それぞれ偏光フィルタを0度、90度回転させた状態で

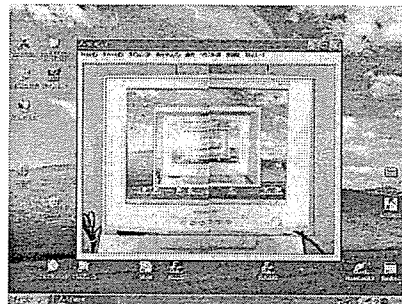


図2 偏光フィルタを0度回転させて撮影した画像
Fig.2 LCD image with 0 deg rotated PL-Filter.

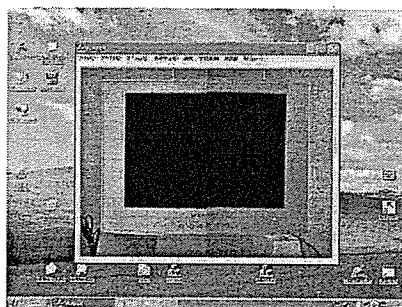


図3 偏光フィルタを90度回転させて撮影した画像
Fig.3 LCD image with 90 deg rotated PL-Filter.

液晶画面を撮影した画面である。図3のような状態であれば、液晶画面のみが黒く撮影されるため、標準状態で撮影した図2のような画像よりも液晶画面と白いフレームとの区別を付けやすく、液晶フレームの検出が容易になる。

3.2 USBカメラの利用

本システムでは、偏光フィルタを装着したUSBカメラ(トリニティ:uCamHi-130C,カラーCMOS,USB2.0)により液晶画面を撮影し、USBポート経由で画像を取り込んでソフトウェア的に処理している。画像処理ボードのようなハードウェアを利用しないことでシステムが簡略化する。画面周辺部にレーザ光線を照射するとレーザ光は外側へ反射し、カメラ1台による撮影では周辺部の照射箇所を検出できない可能性があるため、現状ではカメラを2台設置し、各々撮影した画像を合成した上で1枚の撮影画像(640×480画素)として処理を行う。

3.3 ソフトウェアによる画像処理

まず、取り込んだ画像を2値化し、液晶画面フレームを検出する。図4は2値化された画面データの左上部分を模式的に表しており、2重線内部の各セルが一つ

$Cy[m+1]-Cy[m]$	$Cy[m]$	m										
	8	0	1	1	1	1	1	1	1	1	1	1
0	8	1	1	1	1	1	1	1	1	1	1	1
3	5	2	1	1	0	1	0	1	1	1	0	
2	3	3	1	1	1	0	0	0	0	0	0	
1	2	4	1	1	0	0	0	0	0	0	0	
0	2	5	1	1	0	0	0	0	0	0	0	
1	3	6	1	1	1	0	0	0	0	0	0	
		n	0	1	2	3	4	5	6	7		
		$Cx[n]$	7	7	4	3	2	3	3	2		
		$Cx[n+1]-Cx[n]$	0	3	1	1	1	0	1			

図4 フレームの左上座標の求め方
Fig.4 Upper-left position of LCD frame.

の画素を示す。黒い液晶画面は「0」、白い液晶フレームは「1」となる。合成された画像内において液晶フレームは厳密な四角形とはならないため、以下の手順でフレームの左上、右下座標を算出し四角形とする。

図4中の $Cx[n]$ は $x = 0 \sim 640$ の各列において画素値が「1」である画素の合計値、 $Cy[m]$ も同様に $y = 0 \sim 480$ の各行における画素値が「1」である画素の合計値である。次に $Cx[n]$ と $Cx[n+1]$ 、 $Cy[m]$ と $Cy[m+1]$ の差分が一番大きい箇所を求め、交点をフレームの左上座標とする。図4の例では、左端 $n = 2$ 、上端 $m = 2$ が交点で、座標値は $X = 2$ 、 $Y = 2$ となる。右下座標についても同様である。

液晶フレームの検出後、しきい値を設定して液晶フレーム上の大きなノイズを除去する。しきい値は使用状況により調整することとし、現状では画面を120画素ずつ4行(上部、中央部上下、下部)に分割し各エリアごとにしきい値を設定する。しきい値は状況により異なるが、例えば上部を64、中央上部を54、中央下部を30、下部を25とするなど、各エリアごとに調整することで、液晶画面上に部分的に映る室内灯や窓の映り込みなどのノイズを除去しながらもレーザー光線の照射箇所の検出に最適なしきい値を設定できる。ここで液晶画面にレーザー光線を照射すると、しきい値以上の高輝度箇所が得られるため、その箇所をレーザーポインタの照射箇所とみなすことができる。

更に、細かいノイズを除去するため収縮・膨張処理を加える。収縮処理とは、ある画素の近傍に一つでも0値があればその画素を0値に、その他を1値にする処理である。膨張処理はその逆で、ある画素の近傍の一つでも1値があればその画素を1値に、その他を0値にする処理である。収縮、膨張の順に1回処理を行うと、白い孤立したノイズが収縮処理の際に除去され、レーザー光線の照射箇所だけを検出しやすくなる。

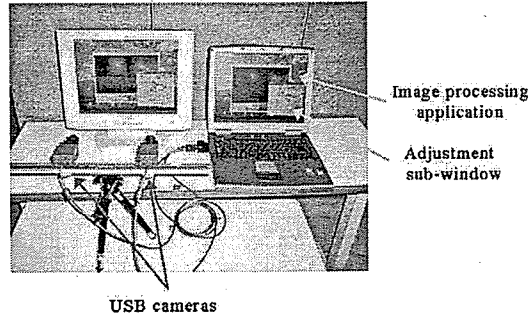


図5 システムの設置概要
Fig.5 System arrangement.

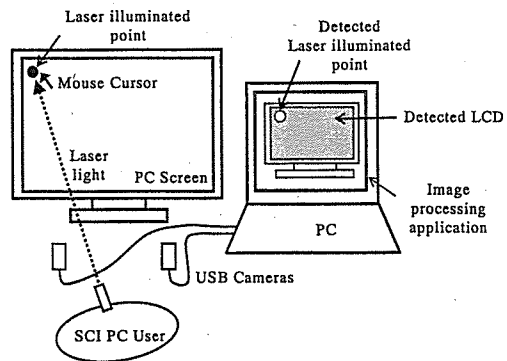


図6 システムの模式図
Fig.6 Schematic depiction of the system.

3.4 システムの設置概要

システム配置を図5に示す。図では拡張用の液晶ディスプレイを付属させて画像を取り込んでいるが、ノートパソコンの液晶画面を取り込み、処理しても同等である。図5ではアプリケーションウィンドウを開いているが、実際の利用時には最小化してタスクバーに収納しておくこととなる。

本システムでは、レーザー光線照射箇所の検出とマウスカーソルの移動はアプリケーションが行うため、システムの設置はUSBカメラ2台を液晶画面手前に配置するだけでよい。しかし、USBカメラは利用者の視界を遮らないように液晶画面の正面下部に配置すること、照明や窓明かりなどの映り込みができるだけ入らないようにすること、レーザー光線の反射光が撮影されるようにしきい値を調整しなければならないことなど、設置にはある程度の熟練が必要とされる。

図6はシステムの模式図で、利用者は左側の拡張用液晶ディスプレイに向かって操作するものとする。図では簡略化のため、左側の画面上にも表示されている

アプリケーションは記述せずに説明する。

利用者の見ている液晶画面は2台のUSBカメラで取り込まれ右側のパソコンで処理される。3.1で記述したように、フィルタの調節によりアプリケーション内では液晶画面のみが黒く撮影される。アプリケーション起動後の調整としては、画像の合成状況の確認、サブウィンドウにおけるしきい値の設定によるノイズの除去、がある。

3.5 マウスカーソルの移動

アプリケーション起動中は、取り込んだ画像内の液晶フレーム位置と液晶フレーム内に映るレーザ光線照射箇所との重心が検出されるので、両者の位置関係により液晶画面上のレーザ光線照射箇所を算出できる。ここでマウスカーソル追従機能を有効にすると、液晶画面においてレーザ光線を照射した箇所にマウスカーソルが移動する。

図6で説明すると、アプリケーション内において液晶フレーム(網掛けの長方形)とフレーム内左上にレーザ光線照射箇所が検出されており(液晶フレーム内の○)、左側の拡張用液晶画面上ではレーザ光線を照射した左上位置にマウスカーソル(矢印)が移動する。

画像の取込み及び処理は10FPS (frame/s) 程度で行い、レーザ光線照射箇所の座標値は随時バッファに蓄積される。マウスカーソルの移動は、フレームごとのマウスカーソルの移動量 m (dot/frame) を算出し、その移動量が規定値 M (dot/frame) 以下の状態で規定の継続時間 (T_s) 経過した後に行うこととした。この処理により、レーザ光線を素早く移動させたとき (M dot/frame 以上) には停止した後の T 秒後に、ゆっくり移動させたとき (M dot/frame 以下) には、照射箇所の移動に合わせてマウスカーソルがゆっくり追従する。

光キーボードユーザは左右クリック、ドラッグなどの操作を液晶画面の周辺に配置した専用センサにレーザ光線を照射することでやっている。上記のような移動処理を行うことで、クリック操作を行うために素早く頭部を移動させた場合でも、その動きに合わせてマウスカーソルが追従することはない。

また、アプリケーションの起動中は「F2」キーの入力によりマウスカーソル追従機能の On/Off を切り換えることが可能であり、 M 、 T の設定値も利用者の状況に合わせて変更可能としている。

4: 使用評価

本システムを利用したポインティング操作(以下、光マウス式と呼ぶ)と光センサ式によるポインティング操作を比較し、使用評価を行った。評価項目にはマウスカーソルがターゲットサークルへ移動するまでの時間 (T_1 s) とマウスカーソルがターゲットに接触してからクリックが完了するまでの時間 (T_2 s) を指標として利用した。

評価実験は、画面に図7のように中心点と周辺30箇所(8方向・3距離の計24箇所と左右画面端3箇所ずつ)の位置にランダムな順にターゲットサークルを表示させる。被験者は中心→ターゲット1→中心→ターゲット2...の順に表示されるターゲットにマウスカーソルを移動させ、左クリックによりターゲットを選択する作業を行う。画面上では、次に選択すべきターゲットだけが表示され、選択されたターゲットは消去される。移動時間 T_1 は、左クリックによるターゲット選択後から次のターゲット内にマウスカーソルが到達するまでの時間として計測する。

図8は光センサ式におけるマウスカーソル移動用センサを液晶画面の周辺に設置したもので、光キーボードユーザが利用している状況と同等である。

評価者として、頸髄損傷者2名(SCI-A:光センサ式3回, SCI-B:両方式3回ずつ)、障害のない被験者6名(C-H:両方式3回ずつ)に協力をお願いした。頸髄損傷者のうちSCI-Aは光キーボードを13年程度継続して利用しており、光センサ式によるマウスカーソルの移動には熟練している。SCI-B及びC-Hは各方式とも初体験である。評価計測は、SCI-Aは自宅において光センサを利用している状況で(図1)、SCI-B、C-Hは研究室において行った。

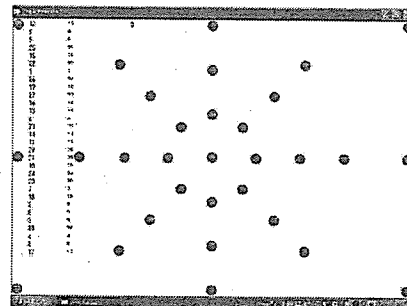


図7 ターゲットサークルの配置
Fig.7 Position of target circle.

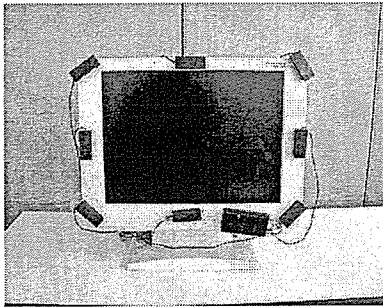


図 8 光センサの配置
Fig. 8 Light sensor arrangement.

使用ディスプレイは SCI-A は 19 インチ, 視距離 90 cm 程度, SCI-B, C-H は 15 インチ, 視距離 70 cm 程度, 画面の解像度は 1024 × 768 ドット, ターゲットサークルの大きさは直径 26 ドットであり, 標準的な画面表示における [最小化, 最大化, 閉じる] ボタンの 1 辺と同値である. マウスカーソルの移動速度は, 初期値は 80 dot/s で, 0.6 秒以上照射し続けると 300 dot/s に加速する設定であり, この値は頸髄損傷者 (SCI-A) が使い慣れている値と同値である. マウスカーソルの位置は 50 ms ごとにサンプリングを行いマウスカーソルの移動軌跡を取得し解析する. 光マウス式における M, T の初期設定値は, M = 10 dot/frame, T = 0.4 s とした.

5. 評価結果と考察

マウスカーソルのポインティングパフォーマンスを下記に定義されるようなスループット (一定時間に処理される仕事量) の値で評価する [18].

$$\text{Throughput} = \frac{\log 2(D/W + 1)}{T_1} \quad (1)$$

D は出発点からターゲット中心までの距離, W はターゲットの直径, T_1 は移動時間である. 単位は bits/s となる.

図 9, 図 10, 表 1 に Throughput 値の結果を示す (図 9 に SCI-A~D, 図 10 に E~H の結果を示す. LM は光マウスで中塗り印, LS は光センサで中抜き印で表す). 1 回の計測につきターゲットまでの移動 (往路) と, ターゲットから中心に戻る移動 (復路) に関して算出した. 図表では, 30 箇所のターゲット選択タスクにおける Throughput 値の平均値と標準偏差を示す (図 9 では SCI-A のみ標準偏差を ± 方向に付けている).

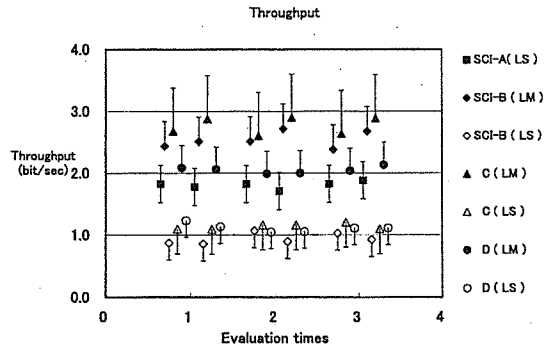


図 9 被験者 SCI-A~D の Throughput の結果 (平均値 (bit/s))

Fig. 9 Results of throughput of SCI-A~D (mean (bit/s)).

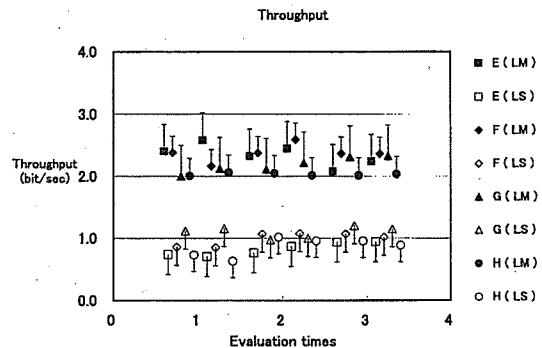


図 10 被験者 E~H の Throughput の結果 (平均値 (bit/s))

Fig. 10 Results of throughput of E~H (mean (bit/s)).

光マウス式では, ターゲットに対して頭部を向ける時間と追従遅れ時間を合計した程度の時間でマウスカーソルが移動するため, Throughput 値は光センサ式よりも高い値が得られている. 評価者によって復路の Throughput 値が大きいのは, 復路の場合ターゲットは常に中心に位置し往路に比べターゲットの位置を確認する必要がないことから, 左クリック後すぐにマウスカーソルの移動操作に移行できることが要因として考えられる.

光センサ式において SCI-A の Throughput 値が高いのは, 光センサ式によるマウスカーソルの移動操作に習熟しているためであり, マウスカーソルがターゲットに近づいた際の微調整操作 (ターゲット前でいったん高速移動中のマウスカーソルを止め, 低スピードでターゲット上にマウスカーソルを止める操作, 若しくは高速移動中であってもターゲット上でタイミングよくマウスカーソルを止める操作) がうまく, マウス

表 1 Throughput の結果 (平均値 ± 標準偏差 (bit/s))
Table 1 Results of throughput (mean ± SD (bit/s)).

被験者	評価回					
	1 (往)	1 (復)	2 (往)	2 (復)	3 (往)	3 (復)
SCI-A (LS)	1.80 ±0.3	1.78 ±0.3	1.82 ±0.2	1.71 ±0.3	1.82 ±0.2	1.87 ±0.2
SCI-B (LM)	2.41 ±0.4	2.53 ±0.3	2.51 ±0.7	2.74 ±0.5	2.34 ±0.5	2.65 ±0.5
SCI-B (LS)	0.87 ±0.5	0.86 ±0.3	1.07 ±0.5	0.90 ±0.4	1.03 ±0.4	0.93 ±0.3
C (LM)	2.68 ±0.7	2.88 ±0.4	2.61 ±0.7	2.90 ±0.4	2.64 ±0.7	2.89 ±0.4
C (LS)	1.10 ±0.4	1.09 ±0.4	1.16 ±0.4	1.16 ±0.4	1.21 ±0.4	1.10 ±0.3
D (LM)	2.08 ±0.7	2.06 ±0.4	1.99 ±0.3	2.00 ±0.6	2.04 ±0.4	2.13 ±0.5
D (LS)	1.23 ±0.5	1.13 ±0.3	1.05 ±0.3	1.05 ±0.2	1.11 ±0.3	1.11 ±0.3
E (LM)	2.40 ±0.4	2.58 ±0.8	2.32 ±0.3	2.45 ±0.5	2.08 ±0.4	2.24 ±0.3
E (LS)	0.74 ±0.4	0.71 ±0.3	0.76 ±0.3	0.87 ±0.5	0.94 ±0.4	0.94 ±0.3
F (LM)	2.38 ±0.3	2.16 ±0.3	2.38 ±0.4	2.59 ±0.4	2.37 ±0.4	2.36 ±0.3
F (LS)	0.86 ±0.7	0.85 ±0.3	1.07 ±0.6	1.08 ±0.6	1.07 ±0.7	1.02 ±0.6
G (LM)	2.00 ±0.6	2.13 ±0.3	2.11 ±0.4	2.22 ±0.4	2.31 ±0.4	2.33 ±0.4
G (LS)	1.12 ±0.8	1.16 ±0.6	0.98 ±0.6	1.00 ±0.5	1.20 ±0.8	1.15 ±0.7
H (LM)	2.01 ±0.5	2.06 ±0.4	2.05 ±0.3	2.02 ±0.2	2.02 ±0.4	2.03 ±0.3
H (LS)	0.73 ±0.5	0.63 ±0.4	1.02 ±0.3	0.96 ±0.5	0.96 ±0.4	0.89 ±0.3

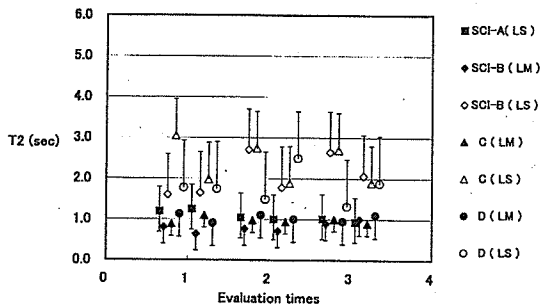


図 11 ターゲット到達からクリックまでの時間 (被験者 SCI-A~D, s)
Fig. 11 Results of T₂ (time from arrival to click (SCI-A~D, s)).

カーソルがターゲットを通過する回数が少ないためである。

T₂ の結果を図 11, 図 12, 表 2 に示す (図 11 が SCI-A~D, 図 12 が E~H). 光センサ式では SCI-A を除き 1.5~4.0 秒程度の操作時間となっていることに比べ, 光マウス式は 1 秒前後でクリックが完了しており

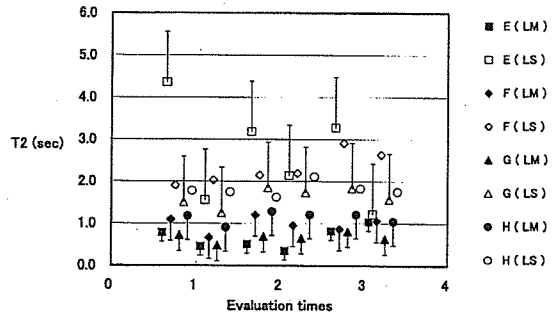


図 12 ターゲット到達からクリックまでの時間 (被験者 E~H, s)
Fig. 12 Results of T₂ (E~H, s).

表 2 ターゲット到達からクリックまでの時間 (平均値 ± 標準偏差 (s))
Table 2 Results of T₂ (time from arrival to click (mean ± SD (s))).

被験者	評価回					
	1 (往)	1 (復)	2 (往)	2 (復)	3 (往)	3 (復)
SCI-A (LS)	1.2 ±0.6	1.3 ±0.8	1.0 ±0.5	1.0 ±0.5	1.0 ±0.6	0.9 ±0.3
SCI-B (LM)	0.8 ±0.5	0.6 ±0.4	0.8 ±0.7	0.7 ±0.5	0.8 ±0.6	1.0 ±0.7
SCI-B (LS)	1.6 ±0.9	1.6 ±1.0	2.7 ±1.5	1.8 ±1.4	2.6 ±1.5	2.1 ±1.3
C (LM)	0.9 ±0.4	1.1 ±0.4	1.0 ±0.3	0.9 ±0.3	1.0 ±0.6	0.9 ±0.3
C (LS)	3.1 ±1.2	2.0 ±1.2	2.7 ±1.4	1.9 ±0.9	2.7 ±1.3	1.9 ±0.9
D (LM)	1.1 ±1.1	0.9 ±1.4	1.1 ±1.0	1.0 ±0.6	0.9 ±1.2	1.1 ±1.3
D (LS)	1.8 ±1.7	1.7 ±1.8	1.5 ±1.4	2.5 ±2.1	1.3 ±1.2	1.9 ±1.7
E (LM)	0.8 ±1.1	0.5 ±0.5	0.5 ±0.8	0.3 ±0.2	0.8 ±0.5	1.0 ±1.8
E (LS)	4.4 ±1.8	1.6 ±1.2	3.2 ±1.5	2.1 ±1.2	3.3 ±1.9	1.2 ±1.6
F (LM)	1.1 ±1.2	0.7 ±1.6	1.2 ±1.5	1.0 ±1.9	0.9 ±0.9	1.1 ±0.5
F (LS)	1.9 ±1.9	2.0 ±1.9	2.1 ±2.1	2.2 ±1.4	2.9 ±3.4	2.6 ±3.5
G (LM)	0.7 ±0.6	0.5 ±0.2	0.7 ±0.8	0.7 ±0.4	0.8 ±0.6	0.6 ±0.5
G (LS)	1.5 ±1.9	1.3 ±1.1	1.8 ±1.3	1.7 ±1.1	1.8 ±1.5	1.6 ±1.5
H (LM)	1.2 ±1.2	0.9 ±1.4	1.3 ±1.0	1.2 ±0.6	1.2 ±1.0	1.0 ±1.2
H (LS)	1.8 ±1.7	1.7 ±1.8	1.6 ±1.4	2.1 ±2.1	1.8 ±1.9	1.8 ±2.4

操作時間は 30~60%程度短縮している。Throughput 値と同様に, T₂ の平均値も SCI-A の方が短い結果となっているが, これは T₂ にはターゲットを通過して反対側に行き過ぎた場合のリカバリー時間も含まれるため, SCI-A の操作に対する習熟度が反映されている

表 3 光センサ式においてターゲットを通過した回数 (回)
Table 3 Results of target-over (times).

被験者	評価回					
	1 (往)	1 (復)	2 (往)	2 (復)	3 (往)	3 (復)
SCI-A	3	2	3	3	1	1
SCI-B	22	16	16	10	14	6
C	36	20	28	14	28	5
D	16	12	8	15	13	9
E	45	10	34	13	24	9
F	40	25	18	21	25	18
G	31	8	10	6	10	10
H	67	58	18	17	12	16

ためである。

参考までに、光センサ式においてターゲットを通り過ぎた回数の合計値を表 3 に示す。H が初回に最高 67 回、E、F が 40 回以上、C、G が 30 回以上を記録したのに比べ、SCI-A は 1~3 回程度しかターゲットを通過していないことから、SCI-A の習熟度がうかがえる。

しかし、どの被験者も回数をこなしていくとターゲット通過回数が少なくなっているため、操作を繰り返すことで SCI-A のような操作性を習得することは可能であろうと予想できる。

光マウス式によるマウスカーソルの移動を光センサ式と比較すると、図 9、図 10 に見られるように Throughput 値、ターゲットに到達してからクリックまでの時間において大きな改善が見られており、評価実験結果からは光マウス式を利用した方が操作性はよいことがうかがえる。

6. むすび

本システムでは、マウスカーソルを移動させたい箇所にレーザー光線を照射させることで移動が実現するためゲイン調節の必要はない。頭部に限らず液晶画面にレーザー光線を照射できる程度の身体動作（例えば、標準マウスは利用できなくともレーザーポインタを手に固定するなど）が得られれば、マウスカーソルを移動させることが可能となる。

課題としては、撮影画面上に存在する室内灯や窓明かりの映り込みなどに対するノイズ対策と、USB カメラを最適な位置に設置するための介護者負担の軽減である。ベッドサイドで、場合によっては仰向けの状態でパソコンを使用することなど、実際の利用場面ではカメラの設置に大きな制限が加わるケースも多い。また、実利用環境では介護者の手間を極力省くことが

要求されるため、しきい値の設定を自動で対応するなど、ラフな設置環境においても更に安定してレーザー光線の照射箇所を検出する機能を備える必要がある。

今後、光キーボードユーザや他の頸髄損傷者による実利用場面での臨床評価を行い、改良点を検討していく予定である。

文 献

- [1] <http://www.wisdomking.com/line255.html>
- [2] <http://homepage3.nifty.com/fuzawa/rc14.htm>
- [3] R. Smith, "Mouth stick design for the client with spinal cord injury," *The American Journal of Occupational Therapy*, vol.43, no.4, pp.251-255, 1989.
- [4] 本田典子, 早川泰詞, "頸髄損傷者 (C5) のマウス操作用自働具," 第 17 回リハビリ工学カンファレンス講演論文集, vol.17, pp.439-442, 2002.
- [5] 数藤康雄, 佐々木忠之, "光入力式キーボードの改良とその使用結果," 国立身体障害者リハビリテーションセンター研究紀要, no.10, pp.115-119, 1989.
- [6] 藤川 景, 五秒ほどの青空, 三五館, 東京, 1997.
- [7] 河村 洋, 土肥徳秀, "PSD 素子を用いた光マウスの試作 (第 2 報)—使用評価と改良," 第 6 回リハビリ工学カンファレンス講演論文集, vol.6, pp.149-152, 1991.
- [8] 小泉義樹, 畠山卓朗, "高位頸髄損傷者用パソコン入力装置の開発," 第 12 回リハビリ工学カンファレンス講演論文集, vol.12, pp.97-102, 1997.
- [9] 伊藤英一, 大橋正洋, "視線移動を考慮した頸髄損傷者用ペン型ポインティングデバイス," 情処学論, vol.39, no.5, pp.1440-1447, 1998.
- [10] 石濱裕規, "ジャイロ咬合型を用いた高位頸椎損傷者への入力支援," 第 15 回リハビリ工学カンファレンス講演論文集, vol.15, pp.323-326, 2000.
- [11] D.G. Evans, R. Drew, and P. Blenkhorn, "Controlling mouse pointer position using an infrared head-operated joystick," *IEEE Trans. Neural System and Rehabilitation Engineering*, vol.8, no.1, pp.107-117, 2000.
- [12] 田中博章, 蒔田憲和, "位置検出機能を備えたポインティングユニットの開発," 第 19 回リハビリ工学カンファレンス講演論文集, vol.19, pp.189-190, 2004.
- [13] こころリソースブック編集会 (編), 福祉情報技術 (e-AT) 製品ガイド こころリソースブック, こころリソースブック出版会, 東京, 2005.
- [14] 飯田朱美, 佐藤 登, 熊沢 透, 菅原 勉, "肢体不自由者へのコミュニケーション支援の実践," 信学技報, WIT2005-50, 2005.
- [15] C. Salem and S. Zhai, "An isometric tongue pointing device," *Proc. SIGCHI Conference on Human factors in computing systems*, pp.538-539, 1997.
- [16] N.S. Lotte, "A tongue based control for disabled people," *Proc. 10th International Conference on Computers Helping People with Special Needs*, pp.913-918, 2006.
- [17] A. Barreto, S. Scargle, and M. Adjouadi, "A placti-

cal EMG-based human-computer interface for users with motor disabilities,” J. Rehabilitation Research and Development, vol.37, pp.53-63, 2000.

- [18] F.L. Edmund and M.B. David, “Adaptive software for head-operated computer controls,” IEEE Trans. Neural System and Rehabilitation Engineering, vol.12, no.1, pp.102-111, 2004.

(平成 18 年 6 月 21 日受付, 10 月 5 日再受付)



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. ヒューマンインタフェース学会, 日本リハビリテーション工学協会各会員.

中途視覚障害者の学習支援を目的とした ペン入力学習ノート “Pen-Talker” の開発

清田 公保[†] 江崎 修央[‡] 伊藤 和之[¶] 伊藤 和幸[¶]

[†]熊本電波工業高等専門学校 〒861-1102 熊本県合志市須屋 2659-2

[‡]鳥羽商船高等専門学校 〒517-8501 三重県鳥羽市池上町 1-1

[¶]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 4-1

E-mail: [†] kkiyota@tc.knct.ac.jp, [‡] ezaki@toba-cmt.ac.jp, [¶] {kazu, ito}@rehab.go.jp

あらまし 中・高齢層を中心とした中途失明の視覚障害者が社会復帰するのに必要な職業教育に寄与するための中途視覚障害者用学習支援システムを提案する。我が国における中途視覚障害者の多くは、あん摩・マッサージ・指圧師、はり師、きゅう師の国家資格取得による職業的自立を目指して、国立施設の理療教育課程に在籍して専門教育を履修している。しかしながら、学習時において適切な文字入力手段を獲得できないために、授業の殆どが音声録音による音声機器に依存した記録に頼っているのが現状である。本研究では、授業や自学学習におけるノート・テイキングが学習の基本であるとの観点に立ち、中途視覚障害者による学習記録の支援を目的としたペン入力と音声録音を併用したノート・テイキングシステム “Pen-Talker” のプロトタイプを開発し、その有効性を検討した。PC 初心者の視覚障害者 8 名に対して、仮名漢字交じりの手紙例文を用いた日本語入力の評価実験の結果、簡単な説明のみで全ての被験者が日本語文字の入力が可能あることを確認した(平均文字認識率 93.3%, 平均文字入力速度 11.2 文字/分)。

キーワード ペン入力, ノートテイキング, 中途失明者, タブレット PC

Development of Pen-based Note-Taking System for Blind People

Kimiyasu KIYOTA[†] Nobuo EZAKI[‡] Kazuyuki Itoh and Kazuyuki Ito[¶]

[†] Kumamoto National College of Technology 2659-2 Suya Koushi Kumamoto, 861-1102 Japan

[‡] Toba National College of Maritime Technology 1-1 Ikegami Toba Mie, 517-8501 Japan

[¶] National Rehabilitation Center for Persons with Disabilities 4-1 Tokorozawa Saitama, 359-8555, Japan

E-mail: [†] kkiyota@tc.knct.ac.jp, [‡] ezaki@toba-cmt.ac.jp, [¶] {kazu, ito}@rehab.go.jp

Abstract When we attend class, we frequently take notes. The method of taking note means writing much information as a paper record. However, we cannot write it down in the notebook if there is no visual feedback information. We propose a pen-based note-taking system, which is called the ‘Pen-Talker’ on the ultra mobile PCs for blind person. A blind novice user with acquired blindness is able to input Japanese characters directly instead of a keyboard without well training by using the proposed system. The system can recognize 3,126 characters including JIS Level 1 characters. This paper describes system design concept of the pen-based note-taking system. We also investigate a higher recognition engine based on a fusion of an on-line and off-line recognition algorithm, and voice assistance function for blind person. Screen information of a display is given to the user via built-in a screen reader in the system. Our present prototype model can be used as a note pad function using a simple button operation. By applying the additional improvement method, recognition accuracy was 93.3% for eight blind subjects. This is an extremely high score compared with the previous our recognition system’s accuracies. Consequently, we confirmed the practical use possibility of the proposed system.

Keyword Pen input, Note taking, Blind people and Tablet PC

1. はじめに

厚生労働省の平成 13 年度の調べによると、我が国の視覚障害者の総数は 30.1 万人にも達する。しかし、点字利用率の実態は、視覚障害者全体の 9.2% に留まっており、10 代から 30 代までの利用率が高く、50 代、60 代と年代が高くなるにつれて利用率は低くなる傾向がある。一般に先天性の視覚障害者の点字利用率は高く、高齢の中途失明者ほど点字の習得率は低い。しかし近年では、網膜色素変性症、糖尿病性網膜症などの疾病や交通事故により、中途失明する後天的視覚障害者の割合が高くなり、高齢化社会の進行と共に深刻化している。このような中途視覚障害者の多くは、あん摩・マッサージ・指圧師、はり師、きゅう師の国家師の国家資格取得による職業的自立を目指しており、全国 5ヶ所に設置された国立施設の理療教育課程に在籍し、3 年若しくは 5 年にわたる専門教育を履修している。在籍する中途視覚障害者の多くは点字の習得・利用が難しく、弱視レンズ、拡大読書器、テープレコーダーをはじめ、PC、録音機能を有する DAISY 専用機、電子辞書や携帯電話など様々な機器を工夫して学習に活用する試みが行われている。しかし、入所時の年齢が高いことなどから、現実的には、音声録音機能を主流とした非効率な学習を余儀なくされるケースが多いのが現状である [1]。

一方、視覚障害者のための情報入力手段や音声出力による画面情報伝達の方法が各種研究機関で研究されており [2]-[4]、実際に盲学校などでも利用され始めている [5]。しかし、現在開発されている視覚障害者向けの日本語文書処理ソフトの殆どは、汎用フルキーボードを利用したタイピング方式が主流であり、キーの配置を覚えるか、6 点入力法と呼ばれる点字キーによる入力手法を学習しなければならない。従来のシステムの操作は、中途失明した高齢者にとっては非常に面倒であり、最初から利用を諦める人も少なくない。このような視覚障害者の日本語入力における問題に対する一手法として、筆者らは高齢者や視覚障害者にも有効なオンライン文字認識技術を用いることを提案し、中途失明者を対象としたペン型日本語入力システムを開発した [6]-[7]。この方式は、キー位置を学習するなどの事前学習が不要であり、ペンで紙に文字を書く要領で日本語文書を作成できるシステムである。オンライン手書き文字認識技術は、オフラインの手書き文字認識技術より比較的容易に実現しやすく、初心者や高齢者の日本語文字入力支援用として多くの市販ソフトが実用化されている。しかしながら、中途失明した視覚障害者に本方式を直接適用する場合、視覚情報のフィードバック無しで筆記される文字は図 1 に示すようなストローク相互の位置などが大きく変形し、誤認識

となる場合が多かった。このため視覚情報欠如によって筆記された文字に対しても、頑強なオンライン文字認識技術のアルゴリズムを構築するため、視覚障害者から採取した文字入力データから変形度の評価を行い [8]、これらの文字にも対応できる認識アルゴリズムを開発した [9]。この認識処理系は随時、文字種の拡大を行い、現在は JIS 第一水準漢字をはじめ 3,216 文字種まで認識可能としており、この処理系を利用して視覚障害者用のペン型電子メールシステムを開発した [10]。しかし、実際にプロトタイプを試作し視覚障害者に利用してもらった結果、日常で利用される日本語文章には平仮名や片仮名などの低画数の文字が非常に多く含まれ、漢字などの認識に比べて特徴量が不足するために正解文字を上位候補として常に出力することが難しいことが分かった。さらに、新たに①うる覚えで誤記入された漢字やあらかじめ学習したときと異なる筆順で記憶している文字に対して辞書に登録された文字データと異なるために誤認識となる場合があること、②視覚障害者の特有の入力ミスにはペン先が誤ってタブレット表面上に接触して、不要なストロークが含まれる場合があることなども新たに判明した。

本研究は、理療教育課程の入所者の就労学習支援を目的とした中途視覚障害者のためのノートテイキングシステムの開発を目指す。本システムでは、中途失明により理療教育課程での就学することになった高齢の学習者でも修練が不要でノートテイクが可能な手段としてペン入力方式を採用し、墨字利用が多い弱視者から中途失明の全盲者までを対象とする。

本報告では、はじめに、これまでに筆者らが開発してきた視覚障害者向けの認識アルゴリズムと、従来手法で十分な認識精度を上げることが出来なかった低画数の文字における特徴量の不足や冗長なストロークが付与された場合の文字に対して、オンラインとオフライン文字認識系を統合したハイブリッド文字認識手法

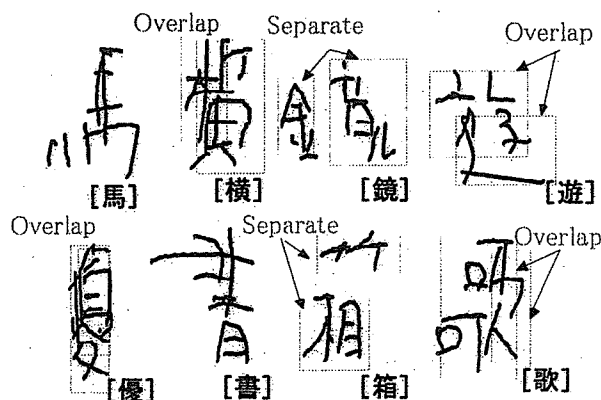


Figure 1 Examples of Japanese handwriting character written by blind users.

を新たに採用し、うろ覚えで筆順違いにより入力された文字や冗長ストロークの誤認識に対応する。次にこの認識エンジンを組み込んで試作したノートテイキングシステムの基本設計思想とシステムの概要について述べる。最後に、理療教育課程に在籍している学習者に利用してもらったときの評価実験について述べ、提案する手法の有効性を示す。

2. 手書き文字認識処理

2.1. 視覚障害者筆記による文字変形

国立施設の理療教育課程に在籍し、専門教育を履修している中途失明者の多くは、墨字による筆記を長年経験している者が多く、現在でも拡大読書機などを利用して紙に手書き入力をしながら学習している入所者（弱視者）も多い。従って、全盲者でも日本語文字の形状を殆どの人が記憶している。このような状況を踏まえて、これまでに筆者らは中途視覚障害者による視覚情報欠如で筆記された文字を複数収集し、その文字変形の特徴を明らかにすると共にオンライン文字認識技術の検討を行ってきた。その結果、視覚障害者による筆記文字の変形にも対応できる文字認識手法として、ストローク代表点移動ベクトル法（Relative Direction between two Strokes:以後、RDS法と略する）と、ストローク等分割法（Line Segment Directions in a Stroke:以後、LSDS法と略する）を提案した。RDS法は、隣り合ったストローク同士の始点、終点、中点間の移動方向を利用する方法で、外形を見ると大きく変形している文字でも隣り合ったストロークの相対位置は比較的安定しているという特徴を利用している。LSDS法は、各ストロークを等間隔で分割しそれぞれの方向成分を利用する、ストローク形状そのものを特徴とする方法である。これら2種類の認識処理系の特徴量は互いに独立であり、これらを併用することにより、視覚障害者による不安定な構造を持つ手書き文字に対しても安定した認識率を実現した。

しかし、本認識手法を用いて実用的な電子メールシステムに適用し、視覚障害者に利用してもらったところ、オンライン入力時の文字認識率は、我々が事前に採取した手書き文字による評価実験と比較すると大幅に低下する傾向がみられた[7]。この現象は、評価実験時には認識対象文字を漢字や平仮名、片仮名、英数字、記号など全字種平等に扱ったのに対して、電子メールなど一般的に利用される実際の文章入力文字には、平仮名などの低画数文字の出現頻度が非常に多くなることによる。すなわち、我々の提案した認識手法は、平仮名を多く含む文章に対しては低画数文字の頻度が増え、認識に必要な特徴量が少なくなるため認識精度が低下することが明らかになった。さらに、実際に中途

視覚障害者に筆記してもらった文字で誤認識された中には、普段は筆順や画数などが安定して書かれているものの、手書き時の状況によっては、不要な点の混入や（タブレットの入力棒とスタイラスペンのペン先が不用意に接触して入力されてしまったと思われる）、筆記中にストロークを一画、書き忘れてしまったなどのケースも見られた。この誤入力は、中途視覚障害者に多く見られる誤認識の大きな特徴であり、従来手法による文字認識処理では改善が困難であることが分かった。

2.2. ハイブリッド型文字認識への拡張

近年、オンライン文字認識技術は急速に発展しており、多くの研究機関で続け字、画数・筆順変動など多様な変形に対応できる手法が提案されている。中途視覚障害者の筆記する変形の大きい文字から精度良く認識処理を行うには、これらの複数の手法を併用して特徴量を増やし認識率の向上を実現することが望ましい。従って、突発的な筆順変動や画数変動も、晴眼者による筆記と比較すると遙かに発生する割合が高いため、筆順、画数フリーの文字認識技術としてオフライン文字認識処理の併用を検討した。すなわち、これまでに筆者らが開発したRDS法およびLSDS法を併用した認識手法のオンライン文字認識手法に加えて、今回は、2次元イメージの統計的マッチングに基づいたオフライン認識と、特徴点の非線形伸縮マッチングに基づいたオンライン認識とを併用したハイブリッド型文字認識処理[11]を併用した統合型の認識エンジンを

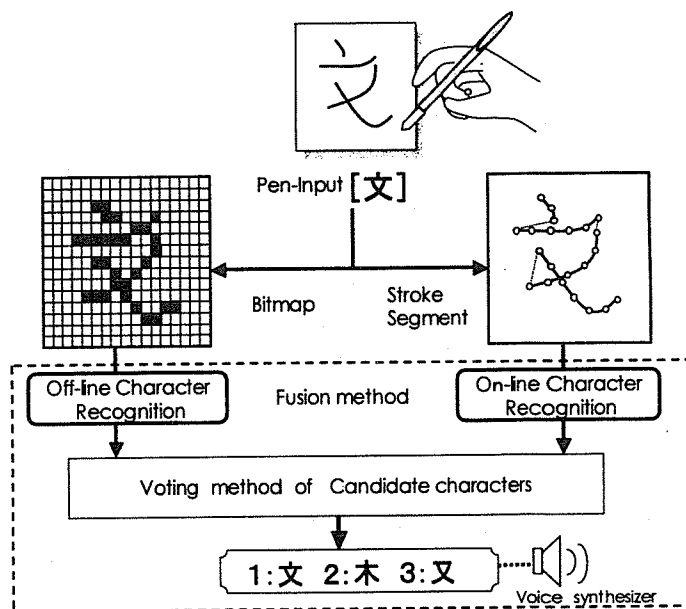


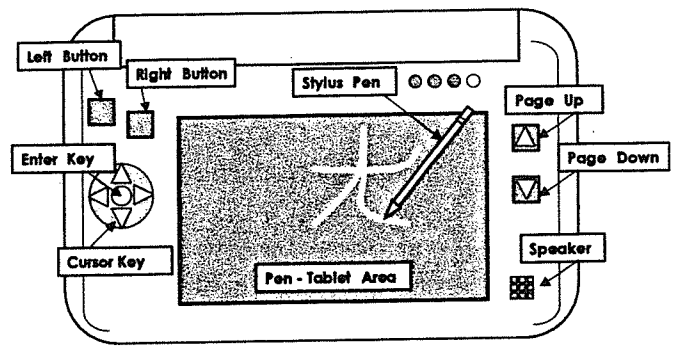
Figure 2 Integrated character recognition of two types: Off-line and On-line recognition.

ノートテイキングシステムのプロトタイプに組込むことにした。ここで複数の認識処理の統合には、図2に示すように別々に認識処理系を実行した後、候補文字の出現順位に着目した投票制を用いた。具体的には、オンライン系とオフライン系の文字認識処理を並列に行い、それぞれ独立に候補文字の順位に応じて得点を与える。そしてそれぞれの候補に含まれる文字毎に総得点を算出し、両者に出現した文字の中から総得点の高い文字から順位付けを行い第一候補文字から音声出力する方式を採用した。

3. “Pen-Talker” の基本設計

3.1. システム構成

図3に、試作した中途視覚障害者のためのペン型ノートテイキングシステムの外観を示す（以後、このシステムを“Pen-Talker”と呼ぶ）。本システムは、ウルトラ・モバイル PC (UM-PC) と呼ばれるタブレット PC (CPU:VIA C7-M 1.0GHz, RAM:512MB, HDD:40GB, 800×480 タッチスクリーンモニタ) をハードウェアとして用い、これらの表面パネルに装備された、マウス用の [左・右ボタン], [上下左右・カーソルキー], [PageUp/PageDown キー], [Enter キー] を入力操作のコマンドボタンとして利用する。但し、マウス用の [左ボタン] は、UM-PC のスタイラスペンのペンダウンと同じ作用を兼ねているために、直接的なコマンド操作は、主として [右ボタン] を用いて行うこととした。操作性に関して言えば、専用のコマンド操作用のハンドコントローラなどを装備した方が良くと考えら



Button Layout

Figure 3 Command button arrangement of the 'Pen-Talker' (Ultra-Mobile PC) with Windows XP Tablet Edition 2005.

れるが、理療教育課程に入所している中途視覚障害者の授業では、科目によって教室を移動する場合が多く、外付けの機器は持ち運びの観点から、今回は既存の操作パネルのボタンのみを利用することにした。

UM-PC に文字入力を制御するインターフェース機能、表示画面情報の音声出力を行うためのスクリーンリーダー (音声合成ソフト)、手書き入力文字の認識エンジンを組み込み、“Pen-Talker” のプロトタイプを開発した。基本的な操作は、ペンで紙に日本語を書く要領でコンピュータに文字を入力することが可能となっている。表示されている画面情報はスクリーンリーダーと呼ばれる画面読み上げソフトを利用して、文章作成時の候補文字列やファイル管理モードのメニュー読み上げ機能を付与している。

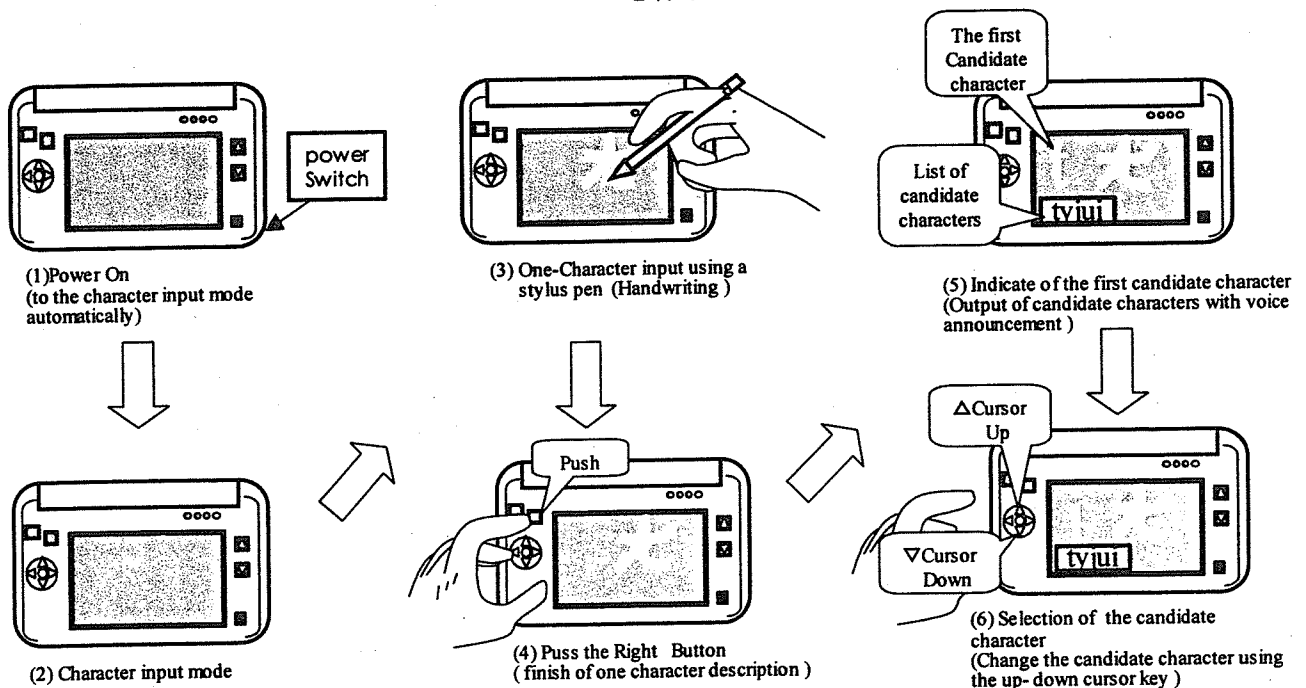


Figure 4 An example of screen monitor layout and character input procedure.

3.2. 文字入力操作法

“Pen-Talker”を用いた文字入力作成の流れを図4に示す。本体の右側サイドにある電源スイッチをスライドさせると本体の電源がON状態になり、しばらくするとシステムが自動的に起動し、文字入力待ち状態となる(図4(1),(2))。ユーザは、文字記入枠内(スクリーン画面枠内)に1文字筆記し(図4(3))、表面パネルの左上にあるマウス用の[右ボタン]を押す(1文字の区切りを指示)(図4(4))。このときシステムでは文字認識処理が実行され、第一候補文字が画面左に表示されると同時に、音声で読み上げられる(図4(5))。この後、読み上げられた候補文字が正解文字であれば、ユーザは続けて次の文字の入力作業へ移行する。もし違う候補が出力された場合は、表面パネルにある[下カーソルキー]を押して、次候補を繰り上げて読み上げる(図4(6))。このとき候補文字は最大第10候補まで用意しているが、もし仮に正解文字が候補中に含まれていなければ、[左カーソルキー]を押すことにより、一文字削除が行われる。また[右カーソルキー]を押すことにより、これまで入力した1文章を「なめらか読み」で読み上げる「音声読み上げ機能」が実行される。ここで、[Enterキー]を押すことにより「改行」が行われる。以下、上記の操作を繰り返すことで文章を作成することができる。

3.3. メニュー機能

“Pen-Talker”では、文書管理のサブメニュー機能を持っている。表面パネルの右側にある[PageDownキー]を押すことにより、画面右下にサブメニューが表示される。このメニューには、文章作成後の「ファイル保存」、「既存文書ファイルの読み込み」、「新規作成」、「システム終了」などが用意されている。ユーザは[上下カーソルキー]で項目を選択し、[右ボタン]で確定する。「システム終了」以外は、[PageUPキー]を押すことにより、文字入力モードへ戻る。また、「ファイル保存」時には、ファイル名は、文章の最初の8文字と作成時の日時を加えた文字列を自動的に付与して保存し、「既存文書ファイルの読み込み」では、ファイル名の一覧が表示されるので、[上下カーソルキー]を用いて音声出力により該当ファイルを選択し、[右ボタン]を押すことで既存文章が読み込まれて、自動的に文章内容の音声読み上げが行われる。

4. 評価実験

4.1. 入力サンプル文字と評価実験参加者

試作した“Pen-Talker”のプロトタイプを用いて、理療教育課程に在籍する入所者に対して、操作性と認識精度の評価実験を実施した。実験に参加していた中途視覚障害者は8名(男性6名、女性2名)で、

Table 1. Comparison of the input time per character for three type input methods.

[Unit: Number of character/min]

	First trial	Second trial	Average
Total Average.	13.2	16.0	14.6
Max.	17.8	20.5	19.1
Min.	9.6	12.8	11.2

Text data: 112 Japanese characters (Letter example)

Subjects: 8 blind persons (6 male and 2 female)

Table 2. Results of character recognition accuracy.

[Unit:%]

Recognition Accuracy	1 st candidate	2 nd candidates	3 rd candidates
First Trial	92.7	97.2	98.3
Second Trial	93.9	97.3	98.5
Ave. of Twice	93.3	97.3	98.4

1名を除いては、キーボードによるタッチタイピングなどの操作が不慣れなパソコンの初心者である。また、タブレットを用いたペン入力操作も全員が初めてであった。参加者の平均年齢は49歳(38歳~57歳)で、糖尿病性網膜症、白内障、緑内障、網膜色素変性症などの疾病により途中で視覚障害になられた方々である。評価用のサンプル文章として、認識率の低下の要因となっていた低画数文字が多く含まれる一般手紙例文集より無作為に選出した未学習文字112文字、6文例を用意した。

4.2. 評価実験結果

評価実験参加者には、はじめに実験の趣旨と実験内容について説明し、インフォームドコンセントを得た上で実験を行った。入力実験は各被験者毎に別々の時間に行い、担当者が入力する文章を読み上げて被験者に伝えて、6文例を入力してもらった。その後、時間を異にして合計2回ずつの筆記入力データを採取した。

表1に、“Pen-Talker”を用いて中途視覚障害者に日本語文章を入力してもらったときの1分あたりの平均入力時間を示す。1回目の試行では、最大で約18文字/分、最小で約10文字/分、2回目の試行では、それぞれ最大20文字、最小13文字に入力時間が早くなっていることが分かる。いずれも、音声読み上げ時間なども含めた時間であるために、タッチタイピングによるキーボード入力と比較すると、かなり遅いペースではあるが(1文字あたりの平均入力時間は約4秒、認識時間は平均約200[mS])、ペン操作やボタン操作の慣れによっては、さらに文字の入力速度の向上は期待できると考えられる。

次に今回の入力実験における文字認識率を表2に示す。平均文字認識率は、第1候補で93.3%、第3候補まで含めた累積認識率は98.4%であった。以前実施した手紙例文サンプルによる認識実験では、平仮名などの低画数文字の占める割合が多く、特徴量の減少と視覚情報欠如による文字筆記により冗長な文字や突発的な筆順変動が生じたことにより認識率が非常に低下した。しかし、今回のハイブリッド型認識エンジンの導入により、低画数文字の認識精度を向上させることができたことが確認できた。

なお、今回の投票制の導入によって、認識処理系のステップ数は約3倍に増加し、また、文字認識時間も単独手法で1文字あたり約50[ms]かかっていたものが、投票制では約200[ms]に増加した。しかしこれらの数字は文字入力時間や音声出力時間などに比べれば無視できる値で、認識処理系による処理時間の負担増は実用上問題ない時間であった。

5. まとめ

理療教育課程に在籍する中途視覚障害者の学習支援を目的としたペン型ノートテイキングシステム、“Pen-Talker”の基本設計思想とその有用性について報告した。筆者らが、従来開発した認識処理系を電子メールシステムに応用した場合、実際の手紙文などの入力時には平仮名などの低画数文字が文中の4割以上を占めるために、文字認識精度が低下する傾向が見られた。その原因として、低画数文字における特徴量不足と視覚情報欠如によって筆記されることによる突発的な冗長ストロークの混入、うる覚えによる画数、筆順変動などの要因があることが分かった。これらを解決するひとつの手段として、晴眼者向けに開発されたオンラインとオフライン認識手法を併用したハイブリッド型認識処理系を新たに加え、複数の認識手法の投票制により文字認識精度を改善することを試みた。中途視覚障害者による被験者実験の結果、本設計思想により試作した“Pen-Talker”のプロトタイプは、情報機器に不慣れた初心者でも簡単な説明で利用することができ、音声応答まで考慮しても1文字あたり平均4秒程度で入力ができる見通しを得た。また、文字認識精度は、新たな認識処理系の追加により、第1候補の平均認識率は93.3%、第3候補までの累積認識率は98.4%にまで向上し、音声読み上げ時間まで考慮すると候補文字3文字で殆ど正解文字を呼び出すことが可能な精度が得られることを確認した。実際に筆記入力実験に参加していただいた中途視覚障害者からの意見として、「漢字や文字の記憶維持にも有効である」、「ペン入力文字が記録できて便利」といった本システムの優位性を指摘していただいた。また、その他の要望として

は、「第3候補まで正解文字がなかった場合は、再入力した方が入力効率は良い」、「1文字入力ではなく、連続して認識して欲しい」、「確定ボタンと文字候補選択のボタンが離れていて使いにくい」などの意見もあり、Nグラムなどの誤り訂正処理などを加えて、今後の改善点としたい。今回は、文字入力を中心としたシステムの評価に留まったが、さらに今後は、授業などにおける本格的な板書を目的とした音声記録方式との併用型ノートテイキングシステムの確立を目指す予定である。

謝 辞

本研究を進めるにあたり、富士通研究所の石垣一司氏、田中宏氏、東京農工大・中川正樹教授には、認識処理系において多大なご支援とご指導をいただいた、ここに感謝の意を表す。本研究の一部は、文部科学省科研費・基盤研究(c) (17500387)、厚生労働科学研究費(長寿科学総合研究事業)の補助による。

文 献

- [1] 伊藤和之, 佐島 毅, 香川邦生, “理療教育課程入所者の学習手段の実態について,” 弱視教育, 第43巻, 第4号, pp.5-11, 2006
- [2] 篠原正美, 森川 浩, “障害者のためのインタフェース—視覚障害者を中心として—,” 情報処理学会誌, Vol.34-9, pp.118-119, 1993
- [3] 米沢義道, “視覚障害者の文書作業のためのエレクトロニクス,” BME, Vol.7, pp.2-7, 1993
- [4] 斎藤正夫, “視覚障害者支援ソフトウェアの製作,” 情報処理学会誌, Vol.6-12, pp.1116-1121, 1995
- [5] 小田島明, 山田信也, 加藤晴喜, 小田浩一, “視覚障害リソース・ネットワーク 視覚障害者とは,” <http://www.twcu.ac.jp/~k-oda/VIRN/>
- [6] 清田公保, 江崎修央, 柳井貴志, 山本眞司, “視覚障害者のためのオンライン日本語入力インタフェースの基本設計,” 信学論(A), J79-A-2, pp.310-317, 1996
- [7] 清田公保, 森田智之, 堂下万由美, 山本眞司, “視覚障害者のためのオンライン日本語入力システムの開発,” JJME(医用電子と生体工学), Vol.35-4, pp.42-49, 1997
- [8] 清田公保, 櫻井敏彦, 山本眞司, “視覚障害者によるオンライン手書き漢字の文字変形分析と画数情報を用いた分類,” 情処論, Vol.36-3, pp.636-644, 1995
- [9] 清田公保, 櫻井敏彦, 山本眞司, “ストローク代表点の相対位置関係に基づく視覚障害者用オンライン文字認識,” 信学論(D-II), J80-D-II-3, pp.715-723, 1997
- [10] 江崎修央, 清田公保, 引地 徹, 山本眞司, “中途失明者のためのオンライン日本語入力システム,” システム制情論, Vol.14-6, pp.316-321, 2001
- [11] 田中宏, 中島健次, 石垣一司, 秋山勝彦, 中川正樹, “オンライン認識とオフライン認識の候補統合によるハイブリッド型ペン入力文字認識エンジン,” 信学技報, PRMU pp.98-140, 1998

視覚障害者向けレストランメニュー画像からの メニュー読み上げシステム

江崎 修央[†] 砂崎 由樹[†] 中西 航[†] 清田 公保[‡]

[†] 鳥羽商船高等専門学校 〒517-8501 三重県鳥羽市池上町 1-1

[‡] 熊本電波工業高等専門学校 〒861-1102 熊本県合志市須屋 2659-2

E-mail: [†] ezaki@toba-cmt.ac.jp, [‡] kkiyota@tc.knct.ac.jp

あらまし 近年、画像中からの文字領域抽出・認識を利用したシステム開発が注目を集めている。情景画像中から文字領域を抽出し、認識処理を利用することで、旅行先での自動翻訳システムや図書の自動データベース化などに利用されつつある。情景画像中の文字情報を読み上げることができれば、視覚障害者にとって必要な情報が得られる可能性が広がる。そこで視覚障害者向けのシステム用技術としても注目されており、バスの行き先読み上げシステムなどが発表されている。我々もこれまでに文字領域抽出手法に重点を絞って研究を行ってきた。この文字情報読み上げシステムと上述のGPSによるヒューマンナビゲーションシステムを組み合わせることで、視覚障害者の単独外出を支援するシステムが構築できると考えている。

本稿では視覚障害者向けの情景画像中の文字情報読み上げシステムの実用化の第一歩としてレストランのメニュー読み上げを行うシステムの開発を目標とし、基本設計、文字領域の抽出手法の検討とメニューデータベースの作成およびメニュー選択に関する実験を行った。

キーワード 視覚障害者、文字領域抽出、文字認識、音声読み上げ

Restaurant Menu Reading System for Blind Person

Nobuo EZAKI[†] Yuuki SUSAKI[†] Wataru NAKANISHI[†] and Kimiyasu KIYOTA[‡]

[†] Toba National College of Maritime Technology 1-1 Ikegami Toba Mie, 517-8501 Japan

[‡] Kumamoto National College of Technology 2659-2 Suya Koushi Kumamoto, 861-1102 Japan

E-mail: [†] ezaki@toba-cmt.ac.jp, [‡] kkiyota@tc.knct.ac.jp

Abstract We have already proposed a document reading system by using micro camera. We apply our previous system to the menu reading system for visually impaired person. The proposed system can be used to check items of menu and prices. The system is composed with a laptop PC, Network Camera and a command input device. At first, the system extracts character areas automatically, and characters are recognized by offline-character recognition application. Next, the menu item database of the layered structure is composed from captured menu image. Then, a menu item database is created automatically by using offline character recognition. The database items are name of menu and its price. The system announces the menu items by voice synthesizer. The user is able to select the candidate item by using [Enter] or [Cancel] button on the control interface board in the system. This operation is very easy for the blind. In this paper, we describe about extraction method of character area for menu images. And we discuss an experiment result about selection time of hierarchy menu.

Keyword Blind person, character area extraction, character recognition, voice synthesizer

1. はじめに

平成 13 年の厚生労働省の調査によると、視覚障害者の数は約 30 万人であると報告されている。これから訪れる超高齢化社会においては、白内障や糖尿病などの疾病や交通事故による中途失明者のしめる割合はますます増加することが予想される。

近年の IT 化に伴う情報機器の普及はとどまるところを知らず、高齢者や障害者などの情報弱者にとってデジタルデバイドはますます深刻になりつつある。そ

の格差を埋めるべく多くの障害者向け情報機器が開発されている。視覚障害者を例にとって見ると、パーソナルコンピュータの利用に関しては、スクリーンリーダと呼ばれる画面読み上げ装置を利用することにより、ホームページの閲覧や電子メールの送受信が可能となっている。ただし、情報入力手段としてはフルキーボードが主流であるため、比較的若い視覚障害者であれば訓練により利用可能であるが、高齢の方の利用は難しいのが実情である。

視覚障害者であっても自立して生活したいという

のは重要な要望であり、携帯電話のGPS機能と音声出力機能を利用したヒューマンナビゲーションシステムなども開発されつつある[1]。しかし、GPSにより屋外での位置はつかめたとしても、建物の構内や店内での情報取得には根本的に無理がある。

ところで、近年、画像中からの文字領域抽出・認識を利用したシステム開発が注目を集めている。情景画像中から文字領域を抽出し、認識処理を利用することで、旅行先での自動翻訳システム[2]や図書の自動データベース化[3]などに利用されつつある。情景画像中の文字情報を読み上げることができれば、視覚障害者にとって必要な情報が得られる可能性が広がる。視覚障害者向けのシステム用技術としても注目されており、バスの行き先読み上げシステム[4]などが発表されている。我々もこれまでに文字領域抽出手法に重点を絞って研究を行ってきた[5-7]。この文字情報読み上げシステムと上述のGPSによるヒューマンナビゲーションシステムを組み合わせることで、視覚障害者の単独外出を支援するシステムが構築できると考えている。

本稿では視覚障害者向けの情景画像中の文字情報読み上げシステムの実用化の第一歩としてレストランのメニュー読み上げを行うシステムの開発を目標とし、2章でシステムの概要、3章で文字領域の抽出手法の検討、4章で多数決法を利用した認識処理、5章でメニューデータベースの作成およびメニュー選択時間に関する評価実験についてそれぞれ取り上げる。

2. メニュー読み上げシステムの概要

まず、レストランにおけるメニュー読み上げを行うにあたり、対象の店舗の形態を大きく「定食屋」系と「居酒屋」系に分けて考える。「定食屋」系とは、一人が一品頼む場合を主に想定している店舗形態、「居酒屋」系とは、多人数で複数の品を頼む形態であると定義する。「定食屋」系の店舗では、食べ物の注文は、一度で済むことが多いのに対して、「居酒屋」系の店舗では、食べ物や飲み物を複数回に渡って注文するという特徴がある。つまり、「定食屋」系の店舗では、視覚障害者であってもメニューの内容を一度店員に聞けばよいので、比較的質問しやすいが、「居酒屋」系の店では、メニューが多岐に渡る上に、注文するたびに質問することとなる。そこで、自動的にメニューを読み上げる本システムの利用により、食事をしたりお酒を飲みながら次に何を頼もうかとメニューを調べる楽しみも得ることができる。つまり、本システムは「居酒屋」系の店舗を対象としたときに有効に機能すると考えられる。

本システムの構成要素は、ノート型パソコンなどの小型携帯端末、ネットワークカメラ、コンピュータとへの情報入力を行うハンドコントローラである。動作

の概要は図1に示すとおりである。1. ユーザが読み上げさせたいメニューを前に広げてハンドコントローラのCaptureボタンを押すと画像を撮影する。2. 文字領域抽出処理によりメニュー画像から文字領域を抽出する。3. 切り分けられた文字領域に対して認識処理を行う。4. 認識結果から階層構造のメニューデータベースを作成する。5. システムはユーザに対して音声で順次選択を促す。6. メニューの選択にはハンドコントローラのボタンを利用する。メニューを本システムにより調べることで、その店にあるメニューが分かることはもちろん、値段やメニュー上に掲載されている説明に関する情報を得ることができる。

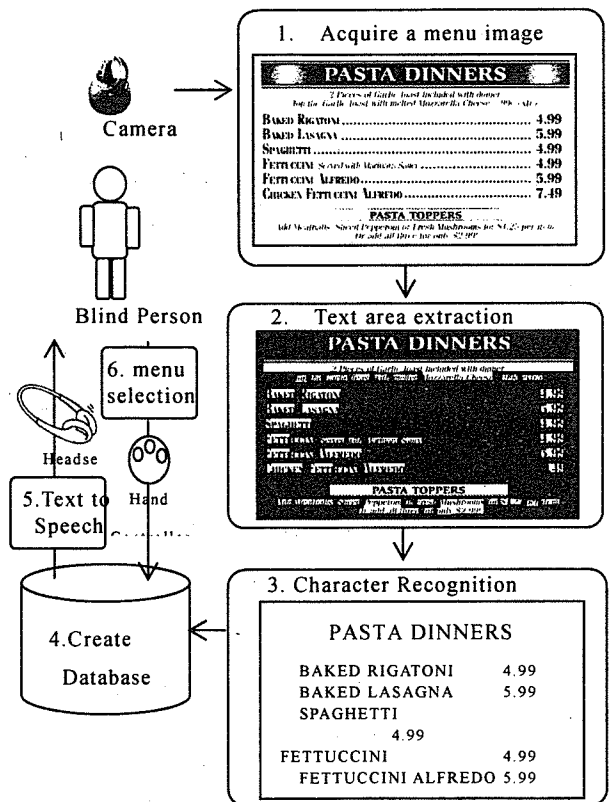


図1 システム概要図

Fig.1 System configuration.

3. メニュー画像からの文字領域抽出

3.1. 対象画像

今回対象としたメニューはインターネット上で取得した居酒屋のメニュー画像である。店舗数は10、画像数は35枚である。図2にメニュー画像の一例を示す。

一概にメニューといっても、文字のみからなる画像、食べ物写真が掲載されている中に文字列が記入されている画像などさまざまなものがある。従って、既存のOCRソフトウェアでは文字領域と非文字領域の切り分けができないので、そのままでは認識処理には利用できない。



図2 メニュー画像の例 [8]
Fig.2 Sample image of Menu

3.2. 文字領域の抽出手法

我々がこれまでに提案してきた文字抽出手法 [5-6]では、対象画像としてICDAR2003のRobust Reading Competition データ [9]としていた。この画像データにおいては Sobel エッジを用いた文字領域抽出手法がもっとも抽出率が高い結果が得られている。この Sobel エッジ手法をそのまま今回のメニュー画像に適用したところ思ったほど抽出率は高くなかった。そこで、対象としたメニュー画像における文字のサイズを調査した結果から、これまでに我々が提案してきた文字領域抽出手法の中で、比較的小さな文字領域を抽出するのに有効である 8color2 手法と、B-Binary2 手法、Morphology 演算の応用手法である Closing-Opening を利用した手法を対象メニュー画像に適用することとした。

8color2 手法は、RGB それぞれのカラーチャンネルに対して判別分析による2値化を行うことで、画像全体を8色に減色した後、文字領域を抽出する手法である。隣り合う文字領域は基本的に同じ色から構成されていることが多いという性質を利用した手法である。ただし、単純に各カラーチャンネルで2値化を行うと小さな文字領域は消失してしまう為、32×32ピクセルの小領域に対象画像を分割し、それぞれの領域のカラーチャンネルごとに2値化を行う手法を採用している。

B-Binary2 は RGB カラーチャンネルを利用するのではなく、HSB 表色系の明度に対して2値化を行う方法である。これも 8Color2 手法と同様に、小領域に分割したあと処理を行っている。メニュー

一に限らず文字領域は人間が認識しやすいように背景領域と文字領域のコントラストをはっきりさせるので、明度で2値化を行うことは、背景領域と文字領域の切り分けに有効な手段である。

Closing-Opening 処理は、Morphology 演算の基本処理手法である Closing を行った画像と Opening を行った画像の差分を取る処理である。この処理により、文字領域は水平に接続した領域として出力される。

3.3. 抽出結果

文字領域抽出の評価値として ICDAR2003 の Robust Reading Competition と同様に precision, recall を利用した。precision とは抽出した領域中に文字領域がどの程度含まれているかという指標である。recall は対象画像中の文字領域をどの程度抽出できたかという手法である。precision と recall はそれぞれ相反する指標であるので、それぞれを考慮した評価値として f も利用する。今、抽出した文字領域を E, 正解文字領域を T, E と T の共通領域を c とすると、precision と recall および f はそれぞれ下記のように計算される。

$$precision = c / |E| \quad (1)$$

$$recall = c / |T| \quad (2)$$

$$f = (0.5 / precision + 0.5 / recall)^{-1} \quad (3)$$

表1に抽出結果を示す。表中で“+”で表してある結果は、2つもしくは3つの手法で出力された領域の OR を取った領域である。表1の結果から、最も高い precision となったのは、B-Binary2 の手法で 78.5%であった。ただし、B-Binary2 だけでは recall の値が低く、総合的な評価値 f としては 54.8%にとどまった。f が最も高かったのは、B-Binary2 と Closing-Opening の OR をとった結果で 63.7%であった。メニューを完璧にデータベースに登録することを考えると、高い recall が必要である。多少 precision が下がったとしても、抽出領域が本当に文字領域かどうかを文字認識処理結果により判断できると考えるからである。

表1 各手法による文字領域抽出率

Table 1 Extraction Rate of Character Region

	precision	recall	f
8color2 (8)	69.4	44.1	53.9
B-Binary2 (B)	78.5	42.1	54.8
Closing-Opening(CO)	66.4	48.3	55.9
8 + CO	61.5	62.4	62.0
B + CO	66.1	61.4	63.7
8 + B	68.2	46.0	55.0
8 + B + CO	61.1	63.1	62.1

4. 投票制を用いた文字認識

4.1. 複数の認識処理系による投票制について

近年、文字認識手法において、複数の認識処理系を組み合わせることで高精度に認識を行う多数決法の有効性が確認されている[10].

図3に今回提案する投票制の概要を示す。3章で説明した文字領域の切り分け処理により、1文字ずつの文字画像データが入力として与えられる。その画像に対して認識処理を行う前に、複数の画像処理手法を組み合わせる前処理を行うことにより、いくつかの文字画像が生成される。これら別々の前処理を施された画像に対して文字認識処理を行うと、同じ文字画像にもかかわらず別の認識結果を得る可能性がある。つまり、元画像をそのまま認識処理した場合に正しい結果が得られなかった場合であっても、適切な前処理を行うことにより正しく認識される場合がある。ただし、前処理に関してはどの手法の組み合わせが適切かは元画像によって違う為、今回は下記の前処理の組み合わせを利用した認識形を構築した。

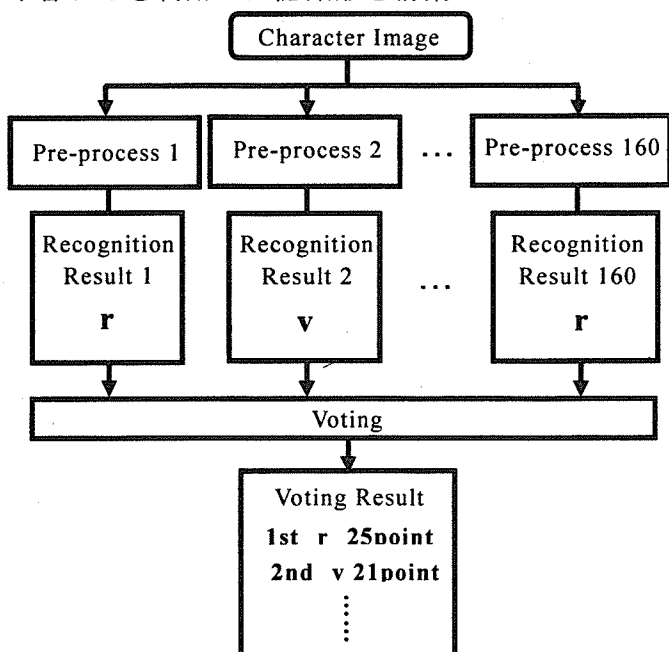


図3 前処理の違いを利用した投票制による文字認識
Fig.3 flow chart of character recognition by voting.

初期画像として用意する画像は次の5つとした。それらは、「元画像（カラー画像）」、「グレースケール画像」、「ブルーグレー画像（RGBの青要素のみの画像）」、「レッドグレー画像（RGBのRのみ）」、「グリーングレー画像（RGBのGのみ）」である。

上記で述べた5つの画像を初期画像として、それぞれに以下に示す5種類の前処理をすべての組

み合わせて施すこととする。5種類の前処理は「2値化」、「ノイズ除去（メディアンフィルタ）」、「階調反転」、「ガンマ補正」、「モルフォロジ演算によるトップハット処理」である。

これら5つの前処理の組み合わせの数は $2^5=32$ 通りとなる。つまり、5つの初期画像それぞれに対して、32通りのすべての組み合わせの前処理を施すことにする。これにより、5（初期画像数） \times 32（前処理の組み合わせの数）=160通りの画像が生成されることとなる。今回利用したOCR処理系はAIソフトの「活字文書OCRライブラリ」である。このソフトウェアは、1つの画像に対して1つの候補文字列を出力するので、この投票制を利用した認識処理系の投票総数は160票となる。最終的には、投票結果の多い順から上位10候補を出力する。

4.2. 文字認識実験結果

投票制による文字認識処理の結果が図4である。グラフの横軸は順位、縦軸は認識率を表している。1位の結果は28%だが、2位、3位と累積していくと認識率が上がっていった。1位ではまだ認識精度が低いものの、2位まで見るだけでも認識精度は38%、10位まで累積結果で認識率は50%程度に上がることを確認した。認識率はフォントによるばらつきが大きいので、今後はフォントに依存しない認識処理を検討する必要がある。

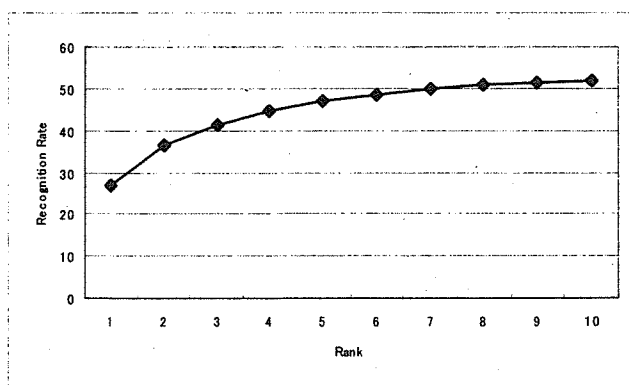


図4 投票制による文字認識処理の結果
Fig.4 result of the recognition rate by voting method.

5. メニューデータベースの作成と選択時間

5.1. メニューの階層構造による選択

多くのメニューは、「うどん」「丼」、「定食」などに分類されている上で「きつねうどん 400円」「肉うどん 600円」というメニュー項目などが記載されている。しかし、分類項目は店によって違う。特に最近新規開店する店舗では「一品料理」の分類項目に対して「すぐに出る料理」といった