

障害者のための携帯電話を入力デバイスとした 音声入力型介助支援システムの試作

宮崎 英一, 谷口 公彦*, 野田 知良**, 平野 和代***, 坂井 聡****
(技術教育) (香川県立高松養護学校教諭) * (香川県立聾学校教諭)**
(教育学部附属養護学校教諭) *** (障害児教育) ****

760-8522 高松市幸町1-1 香川大学教育学部

A Trial Production of the Voice Interface using Cellular Phone System for Impaired Person

Eiichi Miyazaki, Kimihiko Taniguti, Tomoyoshi Noda, Kazuyo Hirano and Satoshi Sakai

Faculty of Education, Kagawa University, 1-1 Saiwai-cho, Takamatsu 760-8522

要 旨 最近の携帯電話に代表されるように、情報機器は一般家庭に深く浸透している事は間違いなくであろう。しかし、これらは健常者の使用を標準として設計されており、様々な障害を持った情報弱者に対してはバリアーとなっている。そこで本研究では、携帯電話の発信操作を音声で介助するシステムを提案する。これは使用者の環境に合わせてカスタマイズが可能で、音声だけで目的に応じた携帯電話発信の補助が可能になる。

キーワード 障害者支援, 携帯電話, 音声制御, Visual Basic, USB

1. はじめに

近年のコンピュータの発展は、以前にもまして加速度的に進んでおり、携帯電話やインターネット環境といった様々な種類の情報インフラの発展には目覚ましいものがある。それに伴い、これらを利用するコンピュータ環境が広く一般家庭にまで普及しており、様々な障害者支援システム^{1,2,3}が提案されている。しかし大部分のシステムは研究室レベルでの提案であり、一般家庭への普及はシステム運用面から考えても、まだ時間がかかると思われる。そこで実際の応用面の観点から、ここではコンピュータおよび

携帯電話といったIT環境の普及に着目し、その情報インフラも含めて歴史的背景から、障害者との係わり合いに関する具体的な問題点に関して考察を行っていきたいと思う。

一般ユーザに対してコンピュータが導入され始めた、ごく初期の段階では、その使用目的もプログラミングやハードウェアの作成のような極めて専門的な目的に限られており、ネットワーク環境も現在とは比較にならないような狭い帯域を用いて、BBS等を介し、ごくローカルに行われていた。その後、Windowsに代表されるGUIを供えたOSの発売に伴い、マイクロソフト社のInternet Explorerやネットス

ケーブ社のNetscapeに代表されるWWWブラウザといった、キラー・アプリケーションの出現により、インターネットの相乗効果と相まって、これらが広く一般家庭にまで普及したのは記憶に新しい所であろう。この結果、インターネットと同義語として扱われるようになった、いわゆる「ホームページ」の閲覧環境が個人レベルにおいても爆発的に普及し、コンピュータ・システムも含めてこれらは現時点では日常生活における情報提供ツールとして大きな役割を持つようになって来ている。

さらには、コンピュータの低価格化および高性能化に伴い、比較的負荷の大きいマルチメディアの取り扱いが、家庭で使用される通常のコンピュータ程度の能力でも十分可能ようになってきた。その結果、TVやDVDの視聴だけに止まらず、ワンセグ等で既に普及し始めているデジタル放送の視聴にもコンピュータが使用されるようになって来ており、もはやコンピュータは一般家電と同程度のレベルで扱えるものとなっている。

またこれらの普及と前後するように携帯電話が広く一般ユーザまでも普及し始め、総務省による調査⁴では平成17年度において世帯当りの携帯電話保有率が89.6%となっており、その普及率は極めて高いものとなって来ている。これは携帯電話の持つポータビリティ性やマルチメディア等の処理能力が、旧来のコンピュータと同程度の能力に向上して来た事から、これを個人的な情報提供端末と見なす事が可能となった為である。さらに携帯電話におけるネットワークの高速化や低料金化および提供コンテンツの豊富さ等とあいまって、従来からの音声通話といった使用方法よりも、メールやWWWページの閲覧、さらには音楽配信といった分野にも重点が置かれるようになってきている。特に大学生にとっては、電話本来の音声通話機能よりも、メール、ブラウザ、音楽再生、スケジュール帳、時計、計算機、財布等の日常生活に必要不可欠なツールとして携帯電話が手放せない存在となっている。

このように普及が著しい携帯電話であるが、

この情報ツールは、健常者と比較して外出が困難な事が多い障害者にとってこそ他人との接触を持つ大変有効なコミュニケーション・ツールであり、潜在的に使用したいというユーザの数は大変に多いと言われている。これらは本研究に関連して坂井らが主催するATサポート研究会⁵における相談事項に関しても、サポート件数に数多く占められており、その要求のニーズが極めて高い事が示されている。

しかし、携帯電話自体のユーザ・インタフェース部分は、ハードウェアを開発するメーカーが、機器の持ち運びやデザインといった点を最優先することもあり、障害者にとっては筐体の小型化、また操作ボタンの大きさやその操作性等が問題となる場合が多く見られる。よってこれらの操作が困難な環境を持つユーザにとっては、携帯電話のインタフェース部分そのものが、情報弱者を生み出す原因になっているとも考えられる。これらにより、携帯電話は、障害者にとって潜在的には最も使用してみたいIT機器の一つであるが、使用するには大きな困難を伴う機器でもあると言え、この部分の改良が進めば、極めて有効な情報ツールとなる可能性を有している。そこで、本研究では障害者にとって最も身近で有効なコミュニケーション・ツールとなる可能性を持つ携帯電話のユーザ・インタフェース部分の改良を目的とし、障害を持ったユーザにおいても出来るだけ簡単にこれらの操作が行えるようなインタフェースの開発を行った。

本研究で試作するシステムは、一般家庭で使用されている通常のコンピュータをシステムの核とし、これに携帯電話と障害者でも利用しやすい音声認識システムや障害者用操作ボタン（ジェリービーン・スイッチ等）に代表される様々なユーザ・インタフェース機器を接続する事で、携帯電話のキーボードを全く操作せずに通話等の操作を行うシステムである。これにより、携帯電話を使用したくても、使用できなかったユーザのニーズを掘り起こし、現時点では、日常生活に大きなウエイトを占めているインターネット環境下においても情報弱者を生み

出さない、やさしいコミュニケーション・システムを構築できる可能性を示すものである。

2. 本システムの概略

2.1 FOMAに関して

現在、日本においては携帯電話を提供する複数のキャリアが存在する。携帯電話はIT機器の中でも最も進化が早い部類に存在し、短期間で新機種が次々と出されている。これにより、現時点では携帯電話において様々な規格（第2世代VS第3世代のような）が存在し、キャリア間の整合性は勿論、同一キャリア内においても、機種間で互換性が取れないというような場合が多数存在している。これらの環境を全てまとめて扱うようなシステムは極めて冗長的なシステムとなり、個人レベルでこのようなシステムの開発は困難である。そこで、本研究は現時点で最も普及台数が多いDocomoの提供するFOMA方式の携帯電話を利用する環境を開発ターゲットとして採用した。以後、本論文では特別な記述が無い場合にはFOMAに対応した携帯電話と見なす。

これは新規格である3世代（3G）の通信規格であり、i-modeに代表されるPDCやcdmaOne等に代表される旧規格ではない。このFOMAという規格は、携帯電話の規格として非常に高い通信ポテンシャルを持ち、今後は主流となる事が見込まれている。しかしその一方、障害者の利用においては、旧規格の携帯電話では実装されていた外部キーボード用の接続端子が存在しないという大きな問題点がある。

この接続端子があれば、携帯電話に実装されているキーボードの入力をコンピュータに接続しているキーボードや外部スイッチで操作するといった事が可能になり、携帯電話のキー操作が困難なユーザにとっては大きな助けとなっていた。その外付けキーボードを図1に示す。これはauのcdmaOneに対応したキーボード・デバイスである。これを用いることで、携帯電話のキー操作が困難なユーザでも、簡単にキー入力が可能になる。しかし、これらcdmaOneや

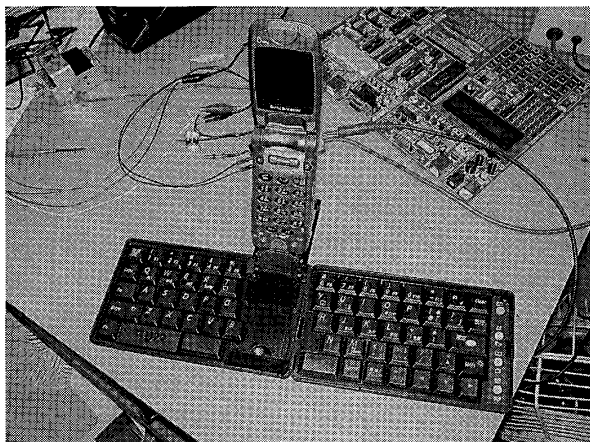


図1 外付けキーボード

i-modeのような旧規格の携帯電話は今後、新機種が発表される事が期待できないため、これらの操作環境を維持しようとするれば、実質上は新機種に買い替えができない状態であった。よって本研究で開発したシステムにより、FOMAに対応した携帯電話でこれらの操作が可能になるようであれば、障害を持ったユーザの利便性が大幅に向上する事が見込まれる。

そこで本研究ではNTTから公開されている内部資料⁶を元に、携帯電話を音声モデムと見なしてコンピュータから制御通信を行うシステムと音声認識システムを組み合わせた新しいシステムを試作した。この結果、各種制御コマンドの命令を発声する事で、通常のキー操作と同様の発信操作を行う事が可能となった。これにより、比較的キーボードの操作が困難なユーザでも、通話操作を音声で簡単に行う事が可能となる。

最後に注意点であるが、本研究では特定の機種（F881iES）でのみシステムの開発をおこなったが、その通信の規格上、コンピュータと携帯電話間の通信規格が完全に統一化されていないという問題が有る。よって本研究で開発したシステムは、その通信制御部分が携帯電話を開発した会社が作成したプログラムに大きく依存するため、他の機種では正常に動作しない事があるので、実際に使用される場合はこの点を注意されたい。

くある。この事前学習の事をエンロールと言
い、音声認識の精度を高めるために、必要な手
順の一つとなっている。しかしこの学習は認識
精度を向上させるため、かなり多量の文章を
ユーザが朗読し、更にそれをシステムに認識さ
せる事が必要なため、ユーザの負担となってい
る事も事実である。そのため、ユーザが長時間
のコンピュータの操作が困難な場合、これらの
エンロールを行わなくても、認識精度の高いシ
ステムが必要となる。

これらの条件を確認するため、本研究ではこ
のシステムを用いて音声認識率をテストした結
果、表1に示すような結果を得た。ここでは、
エンロールの有効性を検証するために、被験者
は事前に比較的明瞭な発声を行う者1名を選
択した。これは、エンロールは被験者が発声した
音声を認識しながら学習していくので、発生が
不明瞭な被験者では、エンロールの効果を確認
する事が困難な為である。ここではエンロール
を行っていない未学習の状態と、エンロール完
了後の2種類の状態において、決まった短いフ
レーズ(例「前へ」、「練習」等の9種類)を
ランダムに300回発声し、その音声認識の結果
を示している。本研究では、音声認識でコマン
ドの操作を行うため、長い文章の認識を行う必
要がないので、このようなコマンドに類するフ
レーズを認識テストのフレーズとしている。テ
ストの結果、エンロール後で、約5%程度の音
声認識結果の向上が見られたが、それよりも着
目すべきは、エンロールなしでも90%の音声認
識率を持つ事が示された事である。このシステ
ムでは、エンロールが無くても高い音声認識率
を持つため、ユーザの負担なしに使用を始める
事が可能である事が示された。これにより、エ
ンロールという長時間の運用が困難なことが多
い、障害を持ったユーザにおいても、負担が少

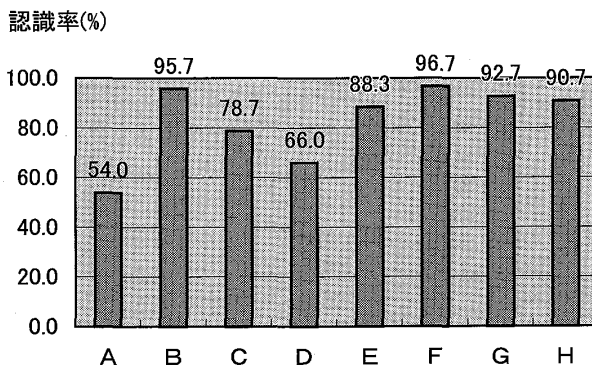
表1 エンロールにおける音声認識結果

エンロール	認識回数 (300回中)	認識率 (%)
無	272	90.7
有	287	95.7

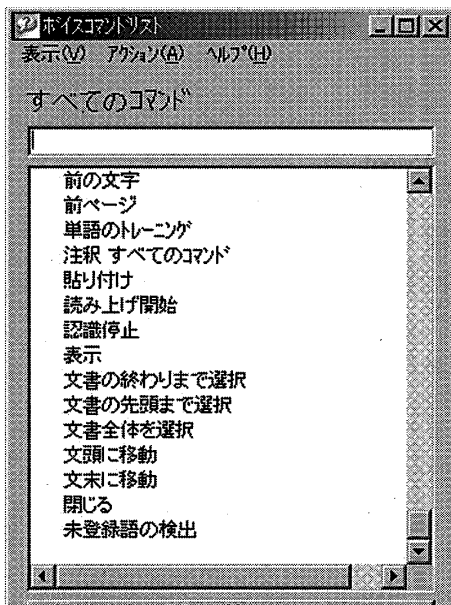
なく、導入が簡単なシステムであると考えられ
る。

更に、このシステムでは音声認識に不慣れな
ユーザが使用するために、様々な発声を行う可
能性のあるユーザを想定して音声認識率をテス
トする必要がある。ここでは、本教室の学生の
協力を仰ぎ、表1と同様のテストを行った。先
程のテスト同様に、エンロール無しで、ラン
ダムに300回の短いフレーズを発声し、その認
識率を確認した。その結果を図4に示す。ここ
ではAからHの合計8名(Hのユーザは表1のユ
ーザでもある)のユーザでテストした結果、最
高で96.7%、最低で54.0%、平均して82.9%の音
声認識結果を得た。実験結果より、特定のユー
ザにおいては高い認識率をもつ場合もあるが、半
分程度しか認識しない場合も存在し、ユーザ
(発声方法)によって認識精度に大きな偏差が
ある事が示された。この精度誤差発生の原因
は、ユーザの喋り方に大きく依存し、口ごも
るような喋り方をするユーザでは、低い音声認
識率(テストユーザAの場合)であった。以上の
結果から、本システムにおいて音声認識の精度
を向上させる場合は、喋り方をはっきりと発音
するように気をつけていれば、エンロール無し
でも、十分実用に耐えられる可能性が示され
た。

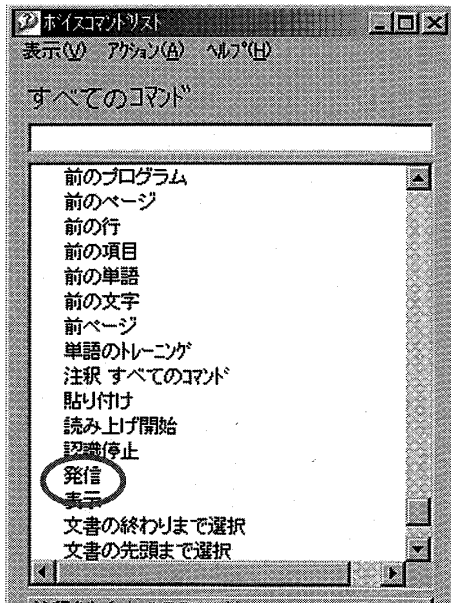
さらにこの音声認識システムはWindowsで
のユーザ・インタフェースのオブジェクトを自
動的に認識してくれるので、ユーザが作成した
プログラムにおいても、その操作コマンドを自
動的に認識してくれるという特徴がある。本研
究で作成したシステムでは、ユーザからの入力



に関して「発信」、「リセット」、「終了」の3種類のボタンを持っている。このプログラムを実行する前後において音声認識ソフトによって認識可能なボイスコマンドの認識リストを図5に示す。ここでは、携帯電話制御プログラムを実行前では、同図a)に示すように「発信」というコマンドがまだ認識されていない。しかし、本研究で開発した「発信」というボタン（ユーザ・インタフェースのオブジェクト）を持つプログラムが実行されると、同図b)に示したよ



a) 実行前



b) 実行後

図5 認識された音声コマンド

うに自動的に音声認識ソフトが「発信」というコマンドを認識している。このように、本研究で採用した音声認識ソフトは、自作したプログラムの操作ボタンでも自動的に認識してくれるので、プログラム作成の負担が極めて少ないという利点がある。

2.4 USBユーザ・インタフェース

さらに、本研究では、音声認識が操作困難なユーザでも、携帯電話の操作が行えるように、外部スイッチ制御のユーザ・インタフェースをシステムに実装している。本研究では、ユーザが使用する各種スイッチとの接続には汎用USBインタフェース・キット⁹を用いた。このインタフェースはKm2net社のキットであり、ハードウェアの構成から比較すると1台1000円程度と他のUSBインタフェースと比べて安価である事、PICマイコンを核とした数点の素子だけで構成されているので製作が容易である事、完成後に機器の調整等が必要無い事等の多くの利点がある。さらにソフトウェアから比較すれば、インタフェースのPIC内にベンダ番号を有しているので、コンピュータに接続するとWindows2000以降のOSならば自動的にUSB機器として認識してくれるので、ユーザ側でデバイス・ドライバーを新規にインストールする必要がない事、Visual Basicからこのインタフェースを制御するドライバー¹⁰がインターネット上でフリーとして広く出回っている事等の利点を持っている。よって本研究では、これ

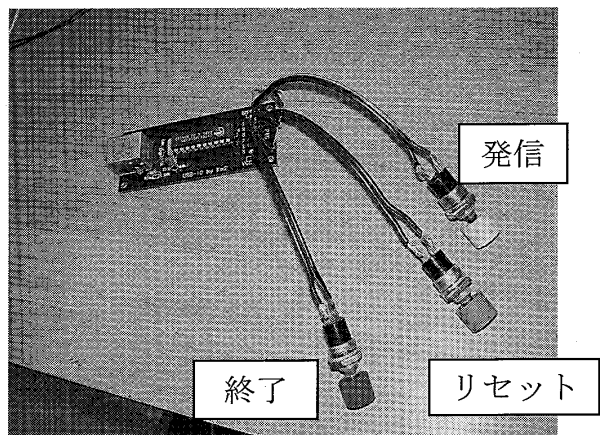


図6 USBインタフェースと制御ボタン

らの利点を踏まえて、このUSBインタフェースを採用し、図6に示したようにVisual Basicからボタン等の外部機器の入出力を制御している。このUSBインタフェースと各種スイッチを組み合わせる事で、音声認識に問題のあるユーザでも、携帯電話の発信が可能になり、障害を持ったユーザに対して、より多くの携帯電話の使用を促す助けになると考えられる。

2.5 システムコア部分

本研究で作成したシステム・プログラムのメイン・フローチャートを図7に示す。ここでは携帯電話との通信およびユーザからの入力制御の両者をリアルタイムで制御するシステムをVisual Basicを用いて作成している。最初にプログラムが起動されると、USBインタフェースを初めとするハードウェアの初期化を行い、コンピュータとこれらの接続を行う。その後、

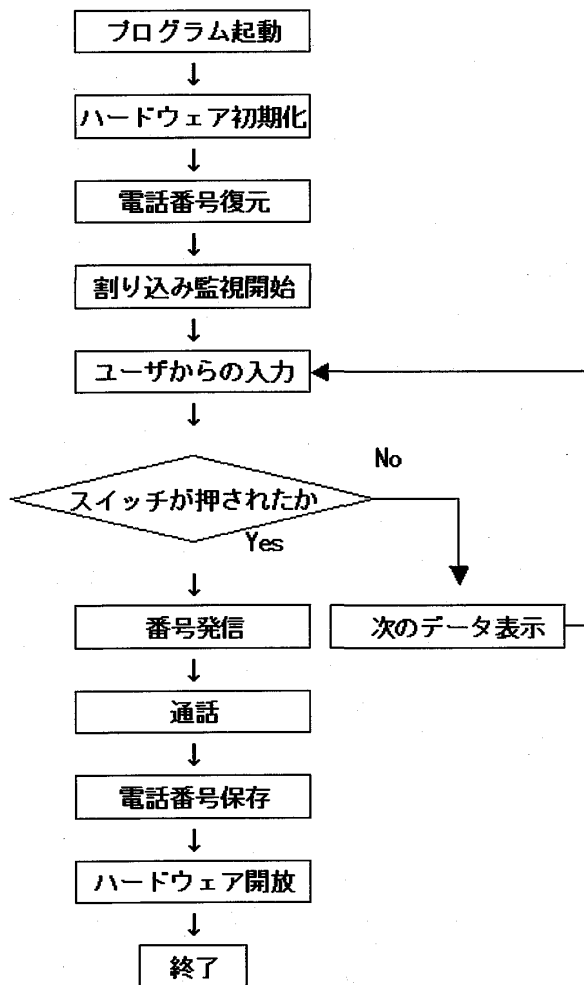


図7 システムのメイン・フローチャート

前回にこのシステムを利用した時の電話番号とその氏名の復元を自動的に行う。本システムは、情報弱者でもシステムの操作が簡単に行える事を目的としているので、ユーザに負担となる余分な操作を行う必要がないシステムとしている。これらが完了すれば、プログラムのユーザ・インタフェース部分に制御が移り、電話番号の表示が開始される。

このようなシステムでユーザが電話番号を選択する場合、登録された電話番号数だけスイッチを備え、そのスイッチで通話を行うという複ボタン型システムと電話番号の表示を次々とスクランし、発信したい番号が表示されたときに発信ボタンを押すという単ボタンスキャン・システムの2つが考えられる。これらのシステムではどちらも相反する特徴を持つが、本研究ではユーザが複数のボタンの制御が困難な場合が多い事を考慮して、単ボタンのスクラン・システムとした。この場合、電話番号をスクランしてユーザに表示する時間が問題になるが、これはユーザの能力に大きく依存するので、ユーザ側で自由に調整が出来るようにフォーム上に時間の調整項目を設けている。さらに本システムでは、電話番号の表示スクランと同期してビープ音を鳴らし、ユーザが発信ボタンのタイミングが取り易いようにしている。ユーザが発信したい電話番号がスクランされた時に、マイクから「発信」と発声する事により、USBインタフェースに接続された発信スイッチを押す事と同様の操作が可能になる。それにより電話発信プログラムに制御が移り、電話番号の発信を行う。ここでは図8に示すATコマンドを用いて携帯電話の発信を行っている。このコマンドはヘイズATコマンドであり、本来はモデムを利用して電話をかけるコマンドである。これを携帯電話の制御に応用する事で、携帯電話の音声通話制御が可能になった。

```
MSCComm1.Output = "ATDT " & 電話番号 & ";" & vbCr
```

図8 AT通信コマンド

2.6 通信プログラム

本研究で試作したシステムのプログラム画面を図9に示す。ユーザはこの画面を見ながら携帯電話の操作を行う。プログラムが起動されると、自動的に電話番号のスキャンが始まる。該当する電話番号の前に赤いマークが表示され、これが現在選択されている電話番号を示している。この赤いマークがフォーム上に表示されているスキャンタイム毎に移動を行う。スキャンタイムはユーザの使用感に大きな影響を与えるので、ユーザの操作スキルに応じて変更させる必要がある。ここではスキャンタイムの秒数を適宜変更後、フォーム上の「リセット」ボタンを押す（[リセット]と発声）と、設定した値でスキャンが始まる。ユーザは自分が発信したい電話番号にマークが移動した時に発信ボタンを押す（[発信]と発声）。その信号を受けて割り込み制御が行われ、最下段の表示欄に電話番号と氏名が表示され、携帯電話の発信が自動的に始まる。この時点でスキャンは自動停止している。この携帯電話での通話は割り込み制御下で行われているので通話中も音声認識が行われ

ており、通話と音声認識が並列的に動作している。通話が終るとユーザは再度「リセットボタン」を押す（[リセット]と発声）とスキャンが再度開始され、次の電話番号を選択して、再度通話を行うか、「終了」ボタン（[終了]と発声）を押してシステムを停止させる事が可能である。

3. おわりに

本研究で試作したシステムを使用して音声のみで携帯電話の発信が可能な事が確認された。これはキーボードの操作が困難なユーザにおいてユーザ・インタフェース操作性の向上をもたらす、情報弱者を生み出さないやさしいインタフェースであると言える。この後は実際のユーザに本システムを運用してもらい、実用面での問題点の洗い出しや、システム自体のブラッシュアップを行って行く予定である。

また更に今後は、このシステムを拡張し、もう一つの研究目的である携帯電話を入力デバイスとした障害者支援システムの構築を行っている

電話番号	氏名
<input type="radio"/> 0878321501	<input type="text" value="宮崎1"/>
<input type="radio"/> 0878321502	<input type="text" value="宮崎2"/>
<input checked="" type="radio"/> 0878321503	<input type="text" value="宮崎3"/>
<input type="radio"/> 0878321504	<input type="text" value="宮崎4"/>
<input type="radio"/> 0878321505	<input type="text" value="宮崎5"/>

リセット

発信

スキャンタイム(秒)

2

終了

0878321503 宮崎3

図9 通信プログラム画面

きたいと思う。

謝辞

本研究は平成18年度地域貢献推進経費「障害者の支援機器利用をサポートするための技術に関する研究」の一部として行われました。謹んでお礼申し上げます。

参 考 文 献

- 1 独立行政法人 情報通信研究機構「高齢者・障害者向けの新たな通信・放送サービスの研究開発係わるニーズ調査報告書」2006
- 2 平岡 茂夫, 宮本 一伸, 富松 潔, 高橋 広「Behind Touch 2: 視覚障害者のための触覚・音声による携帯電話インタフェース」FIT2004 第3回情報科学技術フォーラム 情報科学技術レターズ, Vol. 3, pp291-294, 2004
- 3 尾田継之, 宇根正美「携帯電話・PHSを利用した聴覚障害者のバリアフリーー聴覚障害者の情報入手についてー」第16回 リハ工学カンファレンス講演論文集, pp85-88, 2001
- 4 総務省情報通信統計データベース「主な情報通信機器の保有率(世帯)の推移」
<http://www.johotsusintokei.soumu.go.jp/field/kojin01.html>
- 5 坂井 聡, 武内 美子, 稲田 勤「障害のある人が情報を得るための家電製品の使用実態」日本特殊教育学会第44回大会発表論文集, pp.408, 2006
- 6 株式会社エヌ・テイ・テイ・ドコモ「FOMA USB インタフェースを利用するための技術参考資料」2004
- 7 後閑 哲也「PICで楽しむ USB機器自作のすすめ」技術評論社, 2006
- 8 林岳彦「ViaVoice 98完全活用ガイド」技術評論社, 1999
- 9 成松宏「USB-IOをLinuxで使う」Linux Japan, 第8号, pp.45-49, 2001
- 10 汎用USB-IO試用レポート
<http://homepage2.nifty.com/bake/usb001.htm>