

K-021

携帯メール機能を用いた事前登録型生活音自動識別システム  
The Living Sound Identification System with the Mail Function of Cellular Phone

猿舘 朝† 伊藤 憲三†

Ashita Sarudate Kenzo Itoh

## 1. はじめに

情報化社会に伴い、我々は様々な生活音を耳にするようになった。携帯電話の着信音はもちろん、電子レンジや洗濯機の終了音においてもユニークなものが増えてきた。その一方で、高齢者もまた増加傾向であり、現在では日本の全人口の20%が65歳以上の高齢者である。また、全難聴主催の「聴覚補償リハビリテーション」シンポジウム内での報告によると、難聴者の人口は自覚のある者、ない者を含め、全人口の15.4%(約1944万人)と推定されており、今後さらに増加すると考えられる<sup>[1-2]</sup>。しかし、このような聴力の低下した人のための機器はまだ十分とは言えず、生活するうえで不自由を感じる人は多い。

高齢者や重度難聴者などの耳が不自由な人たちは、補聴器等の力を借りて聴力を補助する。しかし、家庭内においては、わずらわしさや耳の疲労回復などの理由から補聴器を外してしまう。このため、彼らは家の中で生活するために必要な音情報を的確に得ることができない<sup>[3]</sup>。これに対処する支援システムはいくつか販売されているが、識別可能な音が限られていることや追加設備が必要など、価格や性能面から見ると十分とは言えない<sup>[4-10]</sup>。

本研究では、上述のような背景を考慮し、様々な生活音を高い精度で自動識別するため、事前登録方式の生活音識別システムについて提案し、検討・構築を行った<sup>[11]</sup>。さらに、多くの聴覚障害者が日常的に使用している携帯電話を識別結果呈示として用いた手法について検討を加えた。

† 岩手県立大学大学院 ソフトウェア情報学研究科  
† Iwate Prefectural University, Graduate School of Software and Information Science

## 2. 提案するシステム

## 2.1. 聴覚障害者支援機器の現状

現状調査は、インターネットを通じて聴覚障害者本人たちや民間団体に対し、生活時の不便さやシステムへの希望などを調査した。また、「全日本難聴者・中途失聴者団体連合会」他4つの団体で実施した「耳の不自由な人たちが感じているアンケート調査報告書(1993年実施)」、(財)共用品推進機構が作成した「不便さ調査データベース(2001)」も参考にしている。

表-1に生活時の不便さとシステムへの要望についてまとめたものを示した。例えば、ガス警報機音が聞こえない、水を出したまま止め忘れてしまう、といった不便さが挙げられる。これに対し、ガス漏れなどを知らせてくれるランプ(視覚情報)、家中どこにいてもわかるような振動装置(触覚情報)が欲しいといった要求がある。また、その他の意見として、音声を文字に変換して呈示してほしいなど、音声認識技術の重要性も改めて指摘された。

表-2に市販されている製品についての調査結果を示した。現在市販されている製品は、識別できる音の種類に制限があるものがほとんどである。このため、利用者にとって生活する上で重要な音であっても、知らせることができない製品も存在する。例えば、ランプで知らせる方式の場合、ランプが設置されていない所では情報を得ることができない。また、追加設備が必要な機器も多い。さらに、製品の認知度も低く、知っていた場合でも価格が高いという理由から購入できない人たちも多い。ただし、身体障害者福祉法により定められている、聴覚障害等級が2級以上(東京都は4級以上)の聴覚障害者であれば、日常生活用具の給

表-1 支援機器の必要性に関する調査結果

生活における不便さ	要望内容	現在の実現度
朝起きる時	振動式目覚まし時計	○
電話、FAXが鳴っても分からない	電話、FAXの着信を知らせる機器	○
ガス警報機音が聞こえない	ガス漏れを知らせるランプ、振動式装置	△
ドアチャイムが聞こえない	家中どこにいても分かるような光や振動式機器	△
水を止め忘れてしまう	タイマーをかけて振動で教える装置	×
電子レンジなどの終了音が聞こえない	作動の終了を振動と光で知らせる装置	×
ストーブの給油切れ等の音が聞こえない	警告メッセージが分かるようにして欲しい	×
赤ちゃんの鳴き声が聞こえない	家中どこにいても鳴き声が分かるような装置	△
その他	光、振動、文字を組み合わせた製品化 人の音声を文字に変える装置が欲しい	×

(○: 満足 △: 一部満足 ×: 実現なし)

表-2 市販されている支援機器

商品名	メーカー	機能/特徴	問題点
アラートマスター AM6000 (¥51,000~¥214,000)	㈱自立コム	<ul style="list-style-type: none"> <li>親機と子機で構成</li> <li>6種類の音を感知</li> <li>振動/点滅で知らせる</li> </ul>	<ul style="list-style-type: none"> <li>部屋ごとに子機が必要</li> <li>追加設備が必要 (携帯式バイブレータなど)</li> <li>追加設備によっては高価格</li> </ul>
おしらせらんぷ (¥95,000)	㈱リオン	<ul style="list-style-type: none"> <li>音センサと音声マイクで受信</li> <li>FM電波により信号を送信し、ランプで知らせる</li> </ul>	<ul style="list-style-type: none"> <li>高価格</li> <li>ランプが無いところでは情報入手が不可能</li> </ul>
システム5 (¥68,900~¥128,900)	㈱アシスト	<ul style="list-style-type: none"> <li>5種類の音をマイク入力・接点入力</li> <li>ランプ/振動で知らせる</li> </ul>	<ul style="list-style-type: none"> <li>追加設備が必要(ポケットパイプ、ドアノックセンサーなど)</li> </ul>
シルウォッチ (¥95,000)	㈱東京信友	<ul style="list-style-type: none"> <li>1セット4種類程度の音を感知</li> <li>時計型と携帯型の受信機</li> <li>光/振動/文字メッセージで知らせる</li> </ul>	<ul style="list-style-type: none"> <li>高価格</li> <li>必要に応じて別途送信機が必要</li> </ul>

付を受けることができる。しかしながら、ここでの2級というのは両耳の聴力が100 dB以上という障害の程度であり、これは大声で話しても聴き取れないほどの聴覚障害である。3級であっても両耳90 dB以上という障害の程度であり、生活音識別システムのような補助機器が必要な人々がほとんどであると考えられる<sup>[12]</sup>。以上のことから、現在市販されている製品のほとんどは、価格や性能面で使用者の要求を満たしていないといえる。また、システムに対する要望として、

- (1) 高精度である
- (2) 安価である (2~3万円程度)
- (3) システムの操作が簡便である
- (4) 家中どこにいてもわかるように身に付けることができ、光や振動で知らせる警報装置

といった要求があることがわかった。

## 2.2. 提案するシステムの概要

上記のような現状を受け、

- (1) 多くの音を高精度で識別できる

(2) 環境に依存しない

(3) どこにいても音情報がわかる情報呈示をする

これらをシステムの基本コンセプトとした。

まず、より多くの生活音を高精度に識別するため、識別したい生活音を予め登録しておく「事前登録方式」を採用した。図-1に提案するシステム全体の利用概要図を示す。処理の手順として、まず、使用者はシステムの登録モード時に識別させたい生活音を登録する。この時、登録した生活音は特徴パターンとして蓄積される。実際に識別する場合には、システムを識別モードにする。識別モードでは、後述する信号検出を行い、生活音信号であれば先に登録された特徴パターンと照合し、最も特徴が近いものを識別結果とする。

識別された生活音を聴覚障害者に知らせるには、視覚情報あるいは触覚情報に変換する必要がある。市販されている製品は、ランプやフラッシュなどの光による視覚情報、携帯バイブレータなどの振動による触覚情報がほとんどであり、電話の着信音や目覚ましなどといった、各生活音情報に合わせた呈示法となっている。本システムは事前登

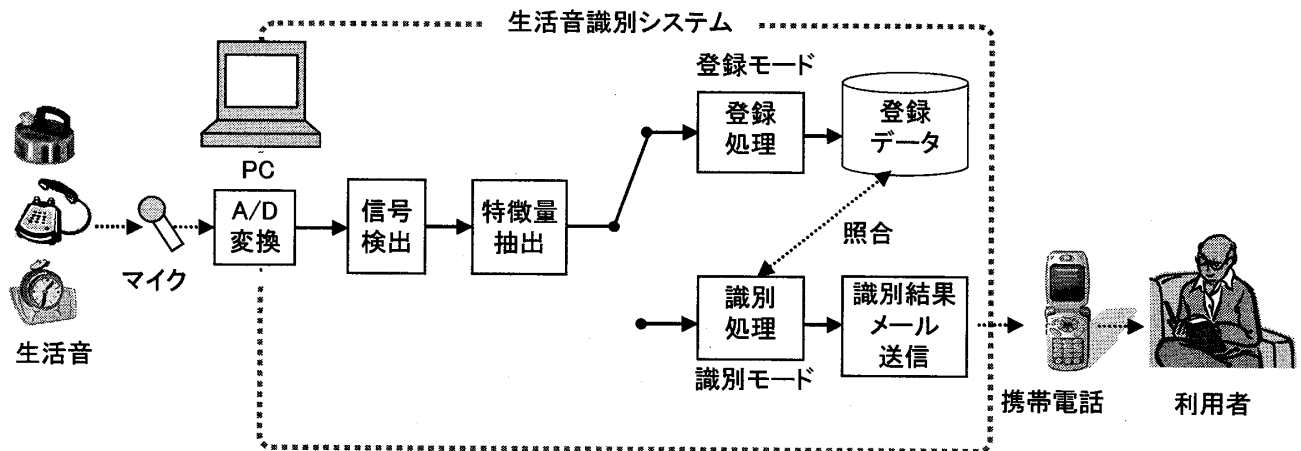


図-1 事前登録型生活音識別システムの利用概要図

録方式であり、識別したい生活音に制限は無い。このため、様々な生活音の特徴にあった情報に自動変換し、呈示する必要がある。また、調査では家中どこにいてもわかるように身に付けられる警報装置という要求が強い。

一方、聴覚障害者にとって、通話はもちろん、文字情報を送受信できる携帯電話は、今や生活になくてはならないものとなっている。そのため普及率も高く、一日中身につけている携帯機器である。また、マナーのためのバイブレーション機能も、聴覚障害者にとって非常に重要で、情報の行き違いをすることなく、即時に対応することができる。さらに、大々的ではないが、ほとんどの携帯電話にはLEDランプが備わっており、着信やアラームなど、複数色の光情報として取得できる。これらの情報は、健聴者、聴覚障害者問わず、呈示方法として利用できる<sup>[13-17]</sup>。

さらに、情報化社会に伴い、PC、インターネットも普及してきた。平成20年の総務省の調べでは、PCの普及率が85.9%、インターネットの普及率が75.3%と、若年層を中心に利用者数は確実に増えてきている。聴覚障害者もPCの操作性に難色を示す意見も多いが、メールや情報取得を目的として、81.1%の人が利用していることがわかっている<sup>[18-19]</sup>。

これらのことから、システムの実装方法として、生活音の登録、識別処理をPCで行い、呈示を携帯電話のメール機能にて行うこととした<sup>[20-22]</sup>。これはインターネット接続環境であれば、専用の機器を必要としない点が大きなメリットとしてあげられる。具体的には、PC、PC用マイク、携帯電話、そして生活音識別システムのアプリケーションソフトウェアがあれば実行可能となる。

### 3. システムの処理概要

#### 3.1. 信号の特徴量と識別処理

登録する特徴量として、時間平均FFTスペクトル特性を用いることとした。特徴量抽出の手順は、まず、ある一定の窓長で切り出した音声信号のフレーム平均パワーを求める。その平均パワーがある閾値以上のフレームに対し、周波数分析を行い、周波数スペクトルパターンを抽出する。このフレーム単位処理を登録音の終了まで行い、得られたスペクトルパターンを平均化したものを特徴量とした。

マッチング尺度として、スペクトル距離値を用いた。まず、スペクトルのパワーを統一するため、(1)式よりスペクトルの平均パワー差( $Pd$ )を求める。その後、(2)式よりパワー正規化を施しつつスペクトル距離値( $SD$ )を求める。

$$Pd = \frac{1}{f_h - f_l} \sum_{i=f_l}^{f_h} (S_x(i) - S_y(i)) \quad (1)$$

$$SD = \sqrt{\frac{1}{f_h - f_l} \sum_{i=f_l}^{f_h} \{S_x(i) - (S_y(i) + Pd)\}^2} \quad (2)$$

ここで、 $S_x$ は入力信号のスペクトル、 $S_y$ は登録信号の

スペクトル、 $f_l$ 、 $f_h$ は距離計算を行う周波数範囲を表す。この $SD$ 値が小さければ小さいほど、互いのスペクトルの特徴が近いこととなる。

さらに精度よく判別するために、スペクトル距離値及び平均パワー差を変量とした線形判別関数を用いた。この線形判別関数から得られる判別得点をもとに識別を行う<sup>[23]</sup>。

#### 3.2. 実時間処理

生活音は突発的に発生するものも多く、予測することは困難である。例えば、ドアをノックする音や、訪問時のチャイムの音などが挙げられる。このため、生活音識別システムでは常に識別すべき音かどうかを瞬時に判別し続けることが望ましい。そこで、本システムにおける実時間処理として、短時間フレーム単位で生活音信号を分析し、処理する手法を用いた<sup>[24]</sup>。手順として、ある一定の窓長で切り出した音声信号に対し、後述する手法により信号検出を行う。ここで、環境雑音ではない何らかの生活音信号であると判断した後、識別処理、即ちスペクトル距離値と平均パワー差を算出する。その後、登録時に求めた判別関数から判別得点を導き出す。識別結果は、まずフレーム単位で判別得点が最も高かったものをフレーム判定結果とし、出力する。そして、信号検出区間内でこのフレーム判定結果が最多であるものを総合識別結果とし、利用者に伝える。これら一連の識別処理は、基本的な信号分析であり、非常に簡便な処理であるため、高速に実時間処理を実現することができる。

#### 3.3. 信号検出処理

生活音を実時間で識別する際、環境に対する頑健性が問題点として挙げられる。これまで音声信号の検出は音声認識の研究領域においても、信号パワーを閾値判定する手法が用いられている事が多い<sup>[25-27]</sup>。パワー閾値を固定的に設定した場合、環境によっては環境雑音と生活音信号を区別することは困難である。また、生活音信号のパワーが閾値を下回った場合、生活音を全く検出できない。環境に合わせて適応的に閾値を設定することも利用者にとって非常に手間である。

一方、一般家庭内では様々な環境雑音がある中、ヒトは目的の音情報のみ抽出し、認知している。それが可能な要因の一つとして、周波数帯域ごとに設定された最小可聴値が挙げられる。最小可聴値は環境雑音の影響により変化し、その値を閾値として信号を検出している。そこで本研究では、ヒトの聴知覚処理をもとに、周波数帯域分析することで信号検出する手法を提案した<sup>[28]</sup>。図-2に信号検出処理の概要を示す。手順として、まず、予め環境雑音の時間平均スペクトルを雑音特性として登録しておく。識別モードでは、一定のフレーム窓長で切り出し、FFTスペクトルを求めた後、周波数帯域毎に分析を行うため、帯域分割を行う。帯域分割の尺度は様々な考えられるが、ここではメル尺度を

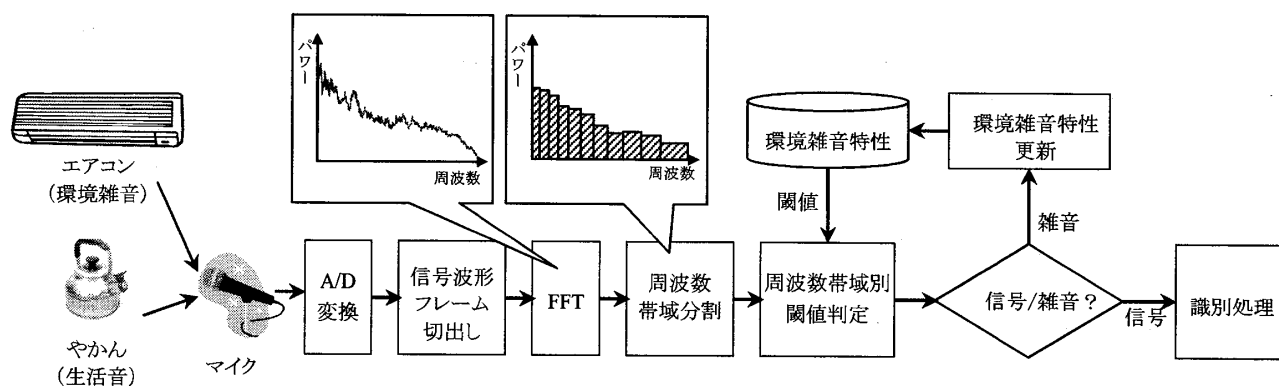


図-2 生活音信号検出法の処理概要図

用いた。これは、メル尺度と周波数および基底膜位置と周波数の関係がほぼ一致しているためである<sup>[29-30]</sup>。メル尺度と周波数  $f$  の関係は次式で表される。

$$mel = \left( \frac{1000}{\log 2} \right) \log \left( \frac{f}{1000} + 1 \right) \quad (3)$$

帯域分割後、予め登録されている環境雑音特性を閾値として判定を行う。判定は周波数帯域毎に行い、閾値を超えた場合には何らかの生活音信号の発生と見なし識別処理に移る。超えなかった場合は、環境雑音と見なし、登録データの環境雑音特性を再度求める。この手法を検証するため、PC上で予備実験を行った。実験ではメルバンド幅を250 melとし、周波数帯域の分割数を11帯域とした。この手法について、空調機音などの定常雑音下で実験を行った結果、SN比10 dBの環境においても良好に生活音信号が検出できることがわかっている。

### 3.4. 携帯メールを利用した呈示法

識別結果の呈示法は、2.2.でも述べたとおり、携帯電話のメール機能を用い、文字、光、振動により伝える。メール内容は生活音が発生した時刻と、識別した生活音名とし、文字情報として呈示することとした。想定される問題として、メールの送信から受信までの遅延時間が考えられる。そこで、携帯電話キャリア2社23種について呈示時間を測定する予備実験を行った。測定は岩手県滝沢村で平日16時頃に統一し、それぞれ20回の通信速度を測定した。その結果、各機種で結果呈示に6~18秒を要した。具体的には、最大通信速度384kbpsの機種で、20秒で伝わることは決して早いとはいえない。しかしながら、キャリア各社が現行規格の通信速度を高速化する動きもあり、呈示時間の問題はある程度解消されると考えられる。

## 4. 実環境識別実験

### 4.1. 実験概要

本システムが家庭内で使用されることを想定して、実環境下での動作実験を行った。図-3に、実験を行った部屋の見取図を示す。部屋には擬似的な室内騒音を発生させ、そ

のレベルを約48dBとした。これは一般家庭内の騒音レベルよりもやや高いレベルである。実験の手順として、まず初期設定として、識別する生活音の数、生活音の名前、登録サンプル数、メールの送信元と送信先情報を入力する。その後、登録モードを起動し、各場所におかれた生活音をサンプル数分発生させ、登録を行う。このときシステムでは、各生活音の平均パワー、時間平均スペクトルについてそれぞれ計算を行っている。登録終了後、線形判別関数を求め、識別尺度を算出する。これで、生活音の登録は終了となる。実際に識別する際には、識別モードを起動する。識別モードスタート時に、まず、環境雑音を測定する。その後、リアルタイムに入力された音に対し、生活音信号検出、識別、呈示処理をする流れとなっている。

実験には、聴覚障害者が特に必要としているとされる10

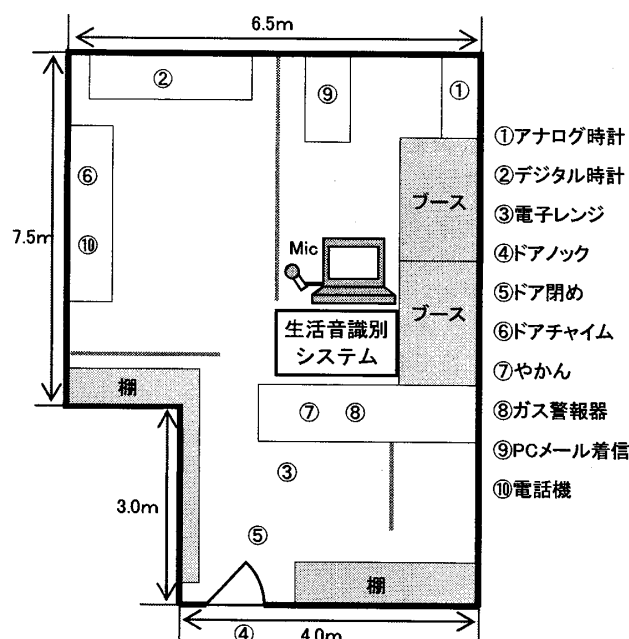


図-3 実験環境の見取図

種の生活音、即ち、アナログ式及びデジタル式時計のアラーム音、電子レンジの終了音、ドアのノック音、ドアの閉め音、ドアチャイム音、やかんの警笛音、ガス警報機音、PCメール着信音、電話の着信音を対象として用いた。実験では、登録モードで各生活音の特徴を登録し、識別モードで携帯電話に送信された識別結果の正答率を求めた。信号の分析条件は、以下の通りである。サンプリング周波数 = 16 kHz、量子化ビット数 = 16 bit、周波数帯域 200~7000Hz、フレーム窓長及びフレーム周期 = 64 ms、周波数スペクトル分析として(FFT) = 1024 Point、信号検出のための帯域分割数 = 11 帯域、メルバンド幅 = 250 mel、登録サンプル数 = 10 回。

図-4 に生活音識別システムの識別モード実行画面を示す。画面上段には信号波形と分析条件、下段に FFT スペクトルがそれぞれリアルタイムに表示される。また、スペクトル右上には識別結果が表示され、画面上で確認できるようになっている。実験では、識別結果のログをとり、識別率を求めた。

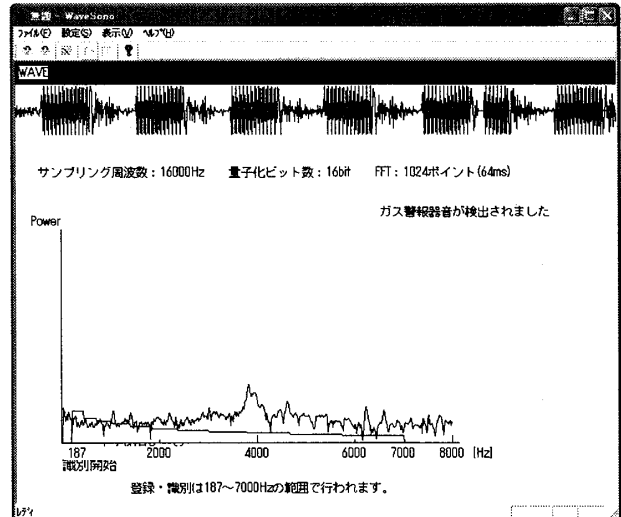


図-4 システム識別モード実行画面の一例

#### 4.2. 実験結果・考察

図-5 に携帯メール呈示画面の一例を示す。検出した生活音の名前と時刻を呈示し、いつ生活音が発生したかわかるようにした。前述したとおり、現状では遅延がややあるが、良好に携帯電話へ呈示されていることが確認できた。

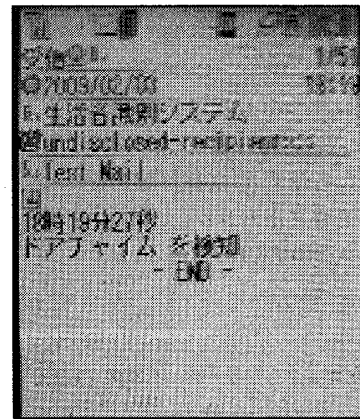


図-5 携帯電話へのメール呈示画面の一例

図-6 に各生活音の識別結果を示す。横軸は識別率、縦軸は生活音の種類を表し、一音における試行回数は 80 回である。この図から、実環境下でも平均 95%以上の高い識別率を示し、高い環境騒音レベルの一般家庭においても本システムが有効であることがわかった。特にこれまでの検討と同様、アナログ式時計アラーム音、デジタル式時計アラーム音、ガス警報機音など、周波数特性が安定している生活音については非常に高い識別率となった。また、図-3 の見取図から、識別システムの周囲 5m 程度であれば、生活音信号パワー減衰の影響もなく、精度よく識別できることがわかった。一方で、ドアを閉める音、ドアチャイム音などの識別率がやや低い。これらは、音の発生パターンが複数存在する生活音である。このような生活音については、その各パターンを登録することで、さらに精度よく識別できると考えられる。

#### 5. おわりに

本稿では、聴覚障害者の生活支援を目的とした生活音識別システムについて述べた。本システムは識別したい生活音の信号特徴量をあらかじめ事前に登録する方式であり、精度よく識別できることがわかっている。また、実時間システム構築の際に必要な、フレーム処理に基づく信号検出及び識別アルゴリズム等について検討を加えた。さらに、その識別結果の呈示法として、現在、健聴者、聴覚障害者問わず普及している携帯電話のメール機能を用い、ネット

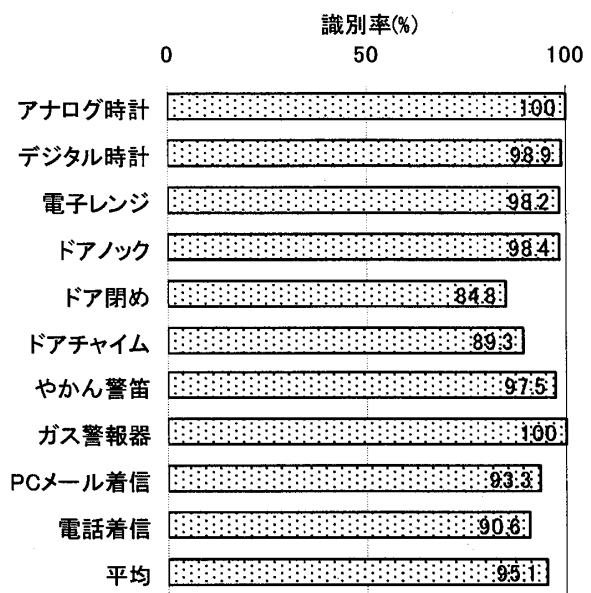


図-6 識別結果

ワーク環境、PC、PC用マイクがあれば、すぐに利用できるシステムとして構築した。そして、実環境下における識別実験を行い、環境雑音下であっても精度よく識別できることがわかった。

今後は、聴覚障害者を被験者として、本システムの評価を行う。その後、インタフェースなどの改良を行い、フリーソフトとして実現し、健聴者、聴覚障害者問わず、実際に利用してもらう予定である。

## 文 献

- [1] 総務省統計局, “人口推計結果”:  
<http://www.stat.go.jp/data/jinsui/index.htm>.
- [2] 社団法人全日本難聴者・中途失聴者団体連合会, “全難聴 行事案内” 2008年12月6,7日 シンポジウム【聴覚補償リハビリテーションの研究】:  
<http://www.zennancho.or.jp/info/event.html>.
- [3] Mark Ross, Communication Access for Persons with Hearing Loss, York Press, pp.197-226, (1994)
- [4] “耳の不自由な人たちが感じているアンケート調査書”, (社)全日本難聴者・中途失聴者団体連合会, 東京都聴覚障害者連盟, 東京都難聴児を持つ親の会, (社)聴力障害者情報文化センター, E&C プロジェクト(現(財)共用品推進機構), (1993).
- [5] “不便さ調査データベース”, (財)共用品推進機構, (2001).
- [6] (株)自立コム, “アラートマスター6000”:  
<http://www.jiritsu.com/site/goods/products/1.html>.
- [7] (株)リオン, “おしらせらんぶ”:  
<http://www.rion.co.jp/products/rionet/rionet20-8.html>.
- [8] (株)アシスト, “システム5”:  
<http://www.assist-hp.co.jp/>.
- [9] (株)東京信友, “シルウォッチ Type1”:  
<http://www.shinyu.co.jp/products/fukushi/index.html>.
- [10] 中途失聴・難聴者ガイドブック作成委員会, 耳のことで悩まないで!—中途失聴・難聴者のガイドブック—, (社)全日本難聴者・中途失聴者団体連合会, (2003, 3)
- [11] 具本榮, 伊藤憲三, “聴覚障害者支援を目的とした生活音識別法に関する検討”, 電子情報通信学会福祉情報工学会研究会資料, WIT2002-69, (2002, 12)
- [12] 後藤修二, リハビリテーション医学全書 13 聴覚障害第2版, pp.10-62, 医歯薬出版株式会社, (1984)
- [13] 徳田浩一, 駒谷和範, 尾形哲也, 奥乃博: “音源定位結果と音声認識結果をHMDに統合提示する聴覚障害者向け音環境理解支援システム”, 情報処理学会講演論文集, 5ZD-7, (2008, 3)
- [14] 高松 衛, 中嶋 芳雄, 中島 賛太郎, 三間 賢一: “照度レベルを考慮した短時間提示条件下における最適表示色数”, 電学論A, Vol. 125, No. 12, pp.1065-1066, (2005).
- [15] Recanzone, G. H. Auditory Influences on Visual Temporal Rate Perception., J. Neurophysiol. 89, 1078-1093, (2003)
- [16] Stuckless, R. Deaf and Hearing Students' Ability to Detect and Correct Word Errors in the Real-time Graphic Display of Spoken Lectures. RTGD working paper No.2. Rochester, NY: National Technical Institute for the Deaf. (1982)
- [17] 伊福部達, 日本音響学会編, 音の福祉工学, pp.47-74, pp.131-162, コロナ社, (1997, 6)
- [18] 総務省, “平成20年通信利用動向調査の結果(概要)”, pp.1-8, (2009)
- [19] 総務省, “障がいのある方々のインターネット等の利用に関する調査報告書—国民全般の情報環境との比較を通じて”, 総務省 情報通信政策研究所, (2002, 6)
- [20] 猿舘朝, 伊藤憲三, “事前登録型生活音識別システム”, 電子情報通信学会技術研究報告, pp.13-18, (2008, 7)
- [21] 猿舘朝, 伊藤憲三, “生活音識別システムにおける識別結果提示法に関する研究”, 電子情報通信学会技術研究報告, pp.89-93, (2008, 9)
- [22] 大野雄也, 猿舘朝, 上野育子, 伊藤憲三, “携帯電話を利用した生活音識別システムの実装”, 情報処理学会全国大会講演論文集, 3Q-5, (2009, 3)
- [23] 猿舘朝, 伊藤憲三, “聴覚障害者支援を目的とした生活音識別システム”, 社団法人 日本音響学会 聴覚研究会資料, (2006, 10)
- [24] 猿舘朝, 伊藤憲三, “生活音識別のための実時間処理に関する検討”, 東北地区若手研究者研究発表会講演資料, YS-436, (2006, 3)
- [25] 鹿野清宏, 伊藤克亘, 河原達也, 武田一哉, 山本幹雄編著, 音声認識システム, pp.1-15, オーム社, (2001, 5)
- [26] Waibel, A., and Lee, K.F. Why study speech recognition? In Readings in Speech Recognition, eds. A. Waibel and K. F. Lee. San Mateo, CA: Morgan Kaufmann Publishers. (1990)
- [27] 伊藤憲三, 水島昌英, 北脇信彦, “音声と非音声の識別処理に基づく定常雑音抑圧方式” 日本音響学会誌 VOL.61, NO8, (2005)
- [28] 猿舘朝, 伊藤憲三, “生活音識別システムにおける信号区間検出法の検討”, 日本音響学会講演論文集, 2-5-5, (2006, 9)
- [29] 三浦種敏, 聴覚と音声, 社団法人電子情報通信学会, pp.341-396, (1991, 6)
- [30] Fletcher, H. and Munson, W.A.: “Loudness, its definition, measurement, and calculation”, J. Acoust. Soc. Am., 5, pp.82-108, (1933)