

マイクロコンピュータを用いた単発音 入力インタフェースの開発

尾崎 晃[†], 宮島 千代美[‡], 西野 隆典^{*}, 北岡 教英[‡], 武田 一哉[‡]

[†]名古屋大学工学部電気電子・情報工学科, [‡]名古屋大学大学院情報科学研究科,

^{*}名古屋大学情報メディア教育センター

〒464-8603 名古屋市千種区不老町

^{† ‡ *} {ozaki,miyajima,nishino,kitaoka,takeda}@sp.m.is.nagoya-u.ac.jp

あらまし 音声認識技術を用いれば、道具を持たずに機械へ命令をすることができる。しかし音声は発話の仕方による変動が比較的大きく、認識処理のために高性能な計算機が必要である。そこで変動が少ない単発音を対象にした、マイクロコンピュータによる安価で小型の入力インタフェースを開発する。本研究では親指の付け根を中指で弾く音に注目して、登録した音のみに反応する家電製品の電源スイッチを試作した。

キーワード 非音声音, 単発音, マイクロコンピュータ, スイッチ

An Acoustic Interface Based on an Isolated Non-speech Sound Using a Microcomputer

Akira Ozaki[†], Chiyomi Miyajima[‡], Takanori Nishino^{*}, Norihide Kitaoka[‡],
and Kazuya Takeda[‡]

[†] Department of Electrical and Electronic Engineering and Information Engineering,
School of Engineering, Nagoya University,

[‡] Graduate School of Information Science, Nagoya University,

^{*} Center for Information Media Studies, Nagoya University

Furo-cho, Chikusa-ku, Nagoya 464-8603, JAPAN

^{† ‡ *} {ozaki,miyajima,nishino,kitaoka,takeda}@sp.m.is.nagoya-u.ac.jp

Abstract Speech interfaces enable hands-free remote control without looking at or touching the equipment. However, a relatively large amount of computation for speech signal processing requires high-performance devices. This paper presents an acoustic interface using a low-cost microcomputer for detecting an isolated non-speech sound. We developed a prototype power switch for home electric appliances, which was turned on or off only by finger snap sounds that were matched to the stored template pattern.

Keywords Non-speech Sound, Isolated Sound, Microcomputer, Switch

1 はじめに

音声は人間にとって最も手軽な情報伝達手段のひとつである。音声認識技術を用いれば、より直感的でユーザーへの負担が少ない、ハンズフリーの操作デバイスが開発できる。ここ数十年の研究によって音声認識性能は大幅に向上し、音声認識を利用した新しいアプリケーションの開発が盛んに進められている。しかし高性能の音声認識には高性能のコンピュータが必要であり、小型化や低価格化が難しく普及への障害となっている。

そこで音を用いたハンズフリーの新しい入力インタフェースとして、時間的な変動が少ない非音声を対象にすることを着想した。これによって計算処理を簡便にし、低性能なコンピュータを用いることでシステムの小型化を図る。手拍子などの大きな音に反応する装置は既に市販されているが、特定の非音声中にのみ反応する装置は少なかった。そこで本研究では、あらかじめ登録した音のみを判別して他の機器を操作する装置を、マイクロコンピュータを用いて開発した。先行研究として、環境音や非音声中の認識に関する研究 [1-4] がなされているが、マイクロホンアレーを用いるなどいずれも規模が大きい。

2 単発音の判別

音声を認識できれば利用者の発話をそのまま入力できるため、インタフェースとして多様な操作が可能である。しかし音声は時間的な変動が大きいため、音声認識システムでは隠れマルコフモデルや動的計画法マッチングなどの比較的高負荷な処理を行っている。一方単発音は生活環境下で類似した音が発生しやすいが、時間的な変動が少なく波形の長さが短いという特徴がある。よって隠れマルコフモデルのような高負荷な処理が必要なく、低性能プロセッサでも判別が可能だと考えられる。そこでマイクロコンピュータを用いることで、システムの小型化と低価格化が期待できる。

本節では単発音を判別する方法について検討する。

2.1 判別のアルゴリズム

信号処理を行うマイクロコンピュータには、判別させたい見本の単発音の波形 $x(t)$ をあらかじめ記録しておく。マイクロコンピュータは、マイクロホンから入力された波形 $y(t)$ との相互相関関数 (1) をサンプルを追加する毎に計算する。

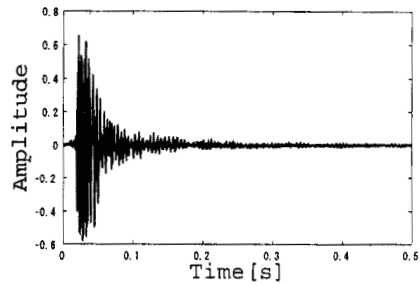


図 1: 単発音の波形例

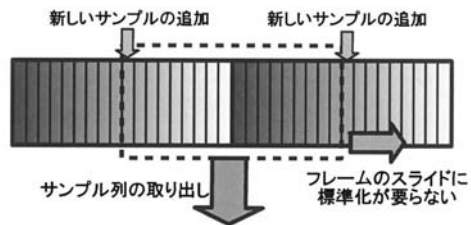


図 2: 2重リングバッファ

$$R = \sum_{n=0}^{N-1} x(n) * y(n) \quad (1)$$

計算能力が低いマイクロコンピュータでもリアルタイムで処理出来る必要があるが、単発音は波形の長さが短いため、比較的少ない計算量で相互相関関数を計算することが可能である。相互相関関数の計算結果が閾値を上回り、波形が短時間の間に観測された場合に、判別させたい音が入力されたものと判断する。

2.2 マイクロコンピュータへの実装

マイクロコンピュータで音声波形のサンプリングと信号処理を両立させるには、実装に工夫をしなければならぬ。アセンブラ命令を参照しながら、命令数が最小となるように記述を行った。サンプリングした波形は図 2 に示すような 2 重リングバッファに記録し、サンプリング毎に相関を計算するようにした。単純なリングバッファは、配列の先頭と末尾をつなげて環状に見立てたデータ構造であり、相関を計算するにはフレームを滑らすたびに標準化のための剰余算をしなければならぬ。そこで同じリングバッファを 2 個連続に接続することで、標準化処理の回数を少なくして計算を高速化した。

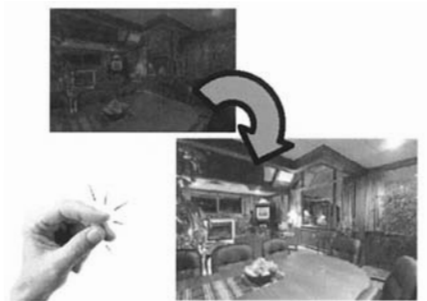


図 3: 単発音判別装置のイメージ

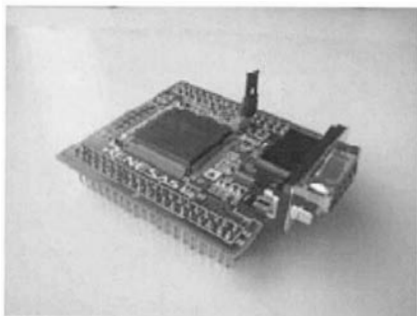


図 4: SH-2 搭載マイコン基板

3 単発音判別装置の構成

特定の単発音で入力を行う，家屋の壁へ埋め込みが可能なマイクロコンピュータを用いた小型のスイッチを考える．本スイッチは単発音を判別した場合に，接続された家電製品の電源を ON/OFF する．生活環境でも利用できるようにするため，他の大きな音を誤って判別しないことが求められる．

3.1 マイクロコンピュータ

日本の組み込み技術は世界で最も優れていると言われている．その中心となる技術がマイクロコンピュータであり，電機メーカー各社が製品を販売している．

単発音の判別に要求される性能を考慮し，部品や開発環境や資料が比較的手に入り易いことから，ルネサステクノロジ社の SH-2 マイコン (図 4) を選択した．相互相関関数の計算には変数同士の乗算を繰り返す必要があるが，SH-2 マイコンには乗算器が内蔵されているため，高速で計算ができる利点がある．SH-2 マイコンは RISC プロセッサであり，プログラミングにはルネサステクノロジ社の純正 C コンパイラを利用する．

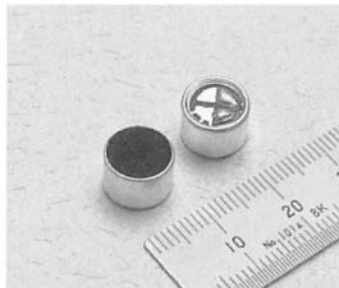


図 5: エレクトレットコンデンサマイクロホン



図 6: 単発音判別スイッチ

3.2 マイクロホン

入力音をマイクロコンピュータに取り込むには，小型のエレクトレットコンデンサマイクロホン (図 5) を用いる．この素子の出力信号は微弱なため増幅が必要である．大きな音が入力された場合に相互相関関数の値が大きくなって誤判別を起こすという問題に対応するため，信号の増幅には ALC (Auto Level Controller) が内蔵されたマイクロホンアンプを利用した．ALC は，大きな音を制限して，小さな音をより増幅する素子である．

3.3 その他の構成

SH-2 マイコンには，周辺回路として A/D 変換器やタイマユニットなどが内蔵されており，制御分野への応用性が高い．コンデンサマイクロホンの信号をアンプで増幅し，SH-2 マイコン内蔵の A/D 変換器で標準化する．あらかじめ登録した音に似ていると判別された場合には，入出力端子に接続したリレー回路を動作させて他機器を操作する．回路の電源は，操作する機器の AC100V 電源をスイッチングレギュレータで直流にして用いる．

表 1: 非音声音の判別率

音の種類	距離 50cm[%]	距離 100cm[%]
指パッチン (左手)	100	52
指パッチン (右手)	0	0
手拍子	8	0
ペットボトルを潰す音	57	47

4 インタフェースの性能評価

開発した単発音の判別スイッチを図6に示す。市販されている露出型のスイッチボックスの中にシステムをパッケージ化した。

この装置ではマイクロコンピュータのROM領域に、親指の付け根を中指で弾く音（通称：指パッチン）を記録させてある。指を鳴らす動作は、周りの注目を得るために用いられる行為として広く認識されており、家電製品への使役的な命令動作に適していると考えられる。また音の振幅が比較的大きいことと、時間的な減衰が早いこと計算処理がしやすい利点がある。あらかじめ記録する音を変えることで、他の単発音を判別することも可能である。

4.1 性能評価実験

開発した判別システムの判別率を表1に示す。判別率は、PCで収録した音をスピーカから再生して求めた。ペットボトルを潰す音のような、大きな単発音ほどより誤判別しやすい傾向がある。

PCで行ったシミュレーションでは、手拍子の誤判別率は0%であったが、これに比べて性能が落ちてしまった。これはマイクロホンの信号を増幅する電子回路のノイズの影響や、閾値の調整が難しいこと、スピーカからマイクまでの伝達特性の影響など、実装による障害が理由であると思われる。しかし、シミュレーション評価に比べて精度の低い環境においての性能としては比較的良好な結果であると考えている。

5 まとめと今後の課題

特定の単発的な非音声音を判別するスイッチの開発を行った。判別プログラムを記述し、信号処理にはSH-2マイコンを用いて小型のスイッチボックスにシステムを封入した。判別率としては挿入誤りが多く、今後の改善が求められる。しかし、従来の単発的な音に反応する装置が誤判別しやすかった、マイクロホン周辺のインパルス音に対して誤判別をしにくくなった。

現在は常に入力音に対して相関を計算しているため、マイクロコンピュータに更なる演算をさせることは難しい。さらに性能を向上させる方法として、信号処理を行う対象の波形を振幅の大きな波形に限定し、相互相関関数以外の周波数的な信号処理を行うことを考えている。

参考文献

- [1] 比屋根一雄, 澤部直太, 飯尾淳, “擬音語表現に基づく衝突音認識システム”, 日本音響学会秋季研究発表会, pp.135-136, Sept. 1998.
- [2] S. Nakamura, K. Hiyane, F. Asano, T. Nishiura, T. Yamada, “Acoustical sound database in real environments for sound scene understanding and hands-free speech recognition”, Proc. International Conference on Language Resources and Evaluation, 2000.
- [3] 山田武志, “マイクロホンアレーを用いた環境音の認識に関する研究”, 電気通信普及財団研究調査報告書 第16号, pp.674-679, 2001.
- [4] T. Ashiya, M. Hasegawa, M. Nakagawa, “IOSES: An Indoor Observation System Based on Environmental Sounds Recognition Using a Neural Network”, Transactions of the Institute of Electrical Engineers of Japan, Vol.116-C, No.3, pp.341-349, 1996.
- [5] ルネサステクノロジ, SH-1/SH-2/SH-DSP ソフトウェアマニュアル.
- [6] CQ 出版社, Interface 2006 年 6 月号.