

多変量解析を用いた警告音の識別

1 S-3

柘植 康彦[†]大西 昇^{§‡}[†]名古屋大学大学院工学研究科計算理工学専攻[§]理化学研究所 バイオミメティックコントロール研究センター[‡]名古屋大学大学院工学研究科情報工学専攻

1 はじめに

聴覚に障害をもった人は音源を識別することが困難なため、何か音がした時すぐさま回りの状況の変化に対処することができない。そのため、もしそれが危険を知らせる音の場合、例えば、火災報知器や踏み切りの遮断器の音などの場合は、命に関わる問題となってくる。このように、環境音の中から音源を識別し、それを聴覚に障害をもった人々に知らせるのは大変重要である。音の識別を支援する、従来の研究として文献 [1] の研究があるが、識別する音源に識別装置が固定される・使用が屋内に限られる、といった問題がある。そのため本研究では、多変量解析の手法を用いて、環境音の中から音源を識別し、それを聴覚障害者に知らせるシステムを開発する。

2 識別が必要とされる音源

本研究では、識別する音源を日常生活に必要な音とし、家の中の音と外での音に分類した。そして各々を警告音とそうでない音に分類した。分類した結果を表 1 に示す。当面对象とする音源は、警告音とし 5 種類の音を識別する。

表 1 音源の分類

	警告音	その他の音
屋内の音	火災報知器のベル	玄関のベル、電話のベル 人の呼ぶ声、赤ちゃんの泣き声 犬の泣き声、窓ガラスの割れる音 やかんの沸騰する音、時計のベル
屋外の音	踏み切りの音、電車の警笛 車のクラクション 警報のサイレン	救急車のサイレン、自転車のベル バトカーのサイレン、鐘の音 消防車のサイレン、人の呼ぶ声

Discrimination of Alarm Signals by Multivariate Statistical Analysis

Yasuhiko Tsuge[†], Noboru Ohnishi^{§‡}[†]Computational Science and Engineering, Nagoya University[‡]RIKEN BMC Research Center, RIKEN[§]Department of Information Engineering, Nagoya University

3 音源識別処理

3.1 処理の流れ

本システムでは、まず、音のパワーの急激な変化を検出し、警報装置に知らせ、その後に音源の識別処理を行ない、何の音源かを知らせる。音源テンプレートの作成には主成分分析を用い、音源識別には判別分析を用いる。処理の流れを図 1 に示す。以下では、音源テンプレートと音源識別の手法について述べる。

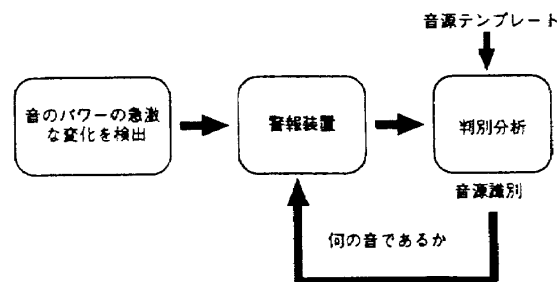


図 1 処理の流れ

3.2 音源テンプレート

本システムでは音源テンプレートとして、表 1 で警告音として分類した 5 種類の音源を市販の CD から 44.1KHz でサンプリングし、 $\Delta t=0.0029$ 秒、 $\Delta f=172$ Hz の短時間 FFT で時間周波数空間に変換した (例: 図 2)。次に、変数に周波数 (128 変数) をとり、サンプルに各音源の時間毎の短時間 FFT の結果 (各音源ごとのサンプル数は 135 個。音源の種類が 5 種類なので、675 個のサンプル) を使い主成分分析を行なった。サンプルを \mathbf{Z}_l ($l=1, \dots, N$) ($N=675, 128$ 次元ベクトル) で表し、また、全サンプルの平均を $\bar{\mathbf{Z}}$ とすると、サンプル \mathbf{Z}_l の共分散行列は次式 (1) で表される。

$$\mathbf{C} = \frac{1}{N} \sum_{l=1}^N (\mathbf{Z}_l - \bar{\mathbf{Z}})(\mathbf{Z}_l - \bar{\mathbf{Z}})^T \quad (1)$$

共分散行列 \mathbf{C} より、固有値と固有ベクトル (主成分ベクトル) \mathbf{P}_n ($n=1, \dots, P$) を求めた。そして、抽

出した主成分ベクトル P_n から、主成分得点を算出した。各サンプル Z_i の主成分得点 S_i は次式 (2) で表される。

$$S_i = (P_1, \dots, P_P)^T (Z_i - \bar{Z}_i) \quad (2)$$

得られた得点を第1および第4主成分間にプロットしたものを図3に示す。図中の記号は各音源の主成分得点の平均である。

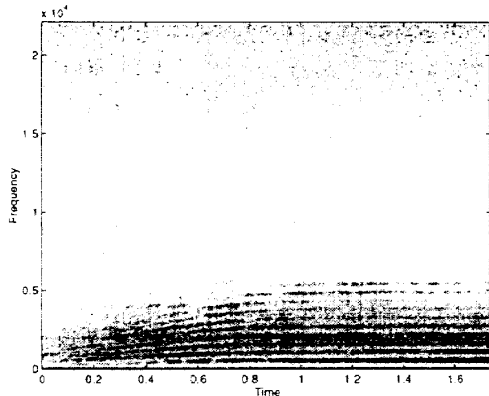


図2 サイレンのスペクトログラム

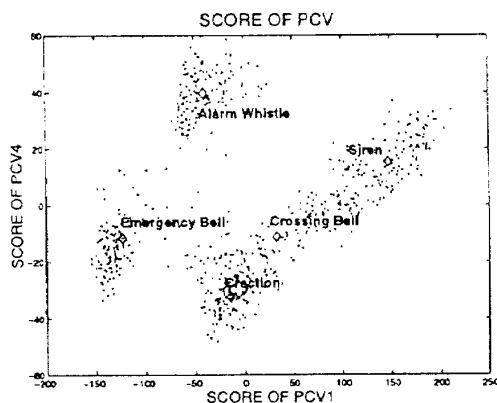


図3 主成分得点のグラフ

3.3 判別分析

音源を識別する方法として、判別分析を使う。各音源のテンプレートと入力音とのマハラノビス距離を求めて、その距離を最小にする音源が、識別された音源となる。ただし、閾値を設定し最小の距離がその閾値より大きいときは、どの音源にも属しないと判別するようにした。音源 i の主成分得点を $S_i^{(i)}$ ($i = 1, \dots, 5$)、その平均を $\bar{S}^{(i)}$ 、共分散行列 $C^{(i)}$ (式 (3)) とする。

$$C^{(i)} = \frac{1}{N^{(i)}} \sum_{l=1}^{N^{(i)}} (S_l^{(i)} - \bar{S}^{(i)})(S_l^{(i)} - \bar{S}^{(i)})^T \quad (3)$$

未知の入力音源の主成分得点を x とする時、 $\bar{S}^{(i)}$ と x とのマハラノビス距離 $D_{(i)}^2$ は次式 (4) であらわされる。

$$D_{(i)}^2 = (x - \bar{S}^{(i)})^T C^{(i)} (x - \bar{S}^{(i)}) \quad (4)$$

判別に使う距離は、マハラノビス距離を $S^{(i)}$ の次元数 P で割った距離 $D_{(i)}^2$ (式 (5)) を使う。

$$D_{(i)}^2 = \frac{D_{(i)}^2}{P} \quad (5)$$

4 判別実験

音源テンプレートを作成したデータとは別のデータを用い各音源の判別実験を行なった。各音源を3、2と同じ条件で主成分得点を求め、音源のテンプレートとの判別分析を行なった。判別実験を行なうデータの数は各音源135個ずつとした。マハラノビス距離の閾値を $D_{(i)}^2 \leq 100$ とした。表2により、第4主成分までを使えば90%以上の確率で判別を行なうことができ、また第10主成分までを使うと95%以上の確率で判別であることが分かった。この結果より、音源の識別に用いる主成分の数は、第10主成分までで十分であると思われる。

表2 判別率

主成分の数 (P)	判別率 (%)				
	2	3	4	10	20
音源					
車のクラクション	91	94	93	96	96
踏み切りの音	94	100	100	100	100
非常ベルの音	97	100	100	100	100
電車の警笛	91	90	100	100	100
警報のサイレン	75	79	99	98	100

5 まとめ

本研究では、警告音の識別処理について述べ、判別実験を行なった。しかし、まだ環境音から、音のパワーの急激な変化を検出するシステムの実装をしなければならない。この方法については、入力音にマスキングをかけて、新しく入ってきた音だけを抽出する事を考えている。また、環境音については雑音の問題があるが、このマスキング処理によってこの問題にも対処できるのではないかと考えている。

参考文献

- [1] Robert I. Demper *et al.*: A Multifunction Domestic Alert System for the Deaf-Blind, IEEE Trans. Rehabilitation Engineering, Vol.3, No.4, 354/359 (1995).
- [2] 大西 昇: ロボットと人間・外界との視聴触インタフェース, 計測と制御, vol.34, No.4, 267/273 (1995).