

信号持続曲線の特徴量を用いた生活音の識別 -雑音付加における検討-

Recognition of Living Sound using the feature value of the Signal Duration Curve: Study in noise addition

○中川昌幸¹, 外山紘之², 細野裕行³*Masayuki Nakagawa¹, Hiroyuki Toyama², Hiroyuki Hosono³

Abstract: In this study, we recognize Living Sound using the new feature value of the Signal Duration Curve. It is aimed at supporting the people who are hard of hearing. We assume the situation in the living environment of the actual by placing a noise in the test data. In the circumstance, average of the rate of recognition is calculated by the 14 category of recognized Living Sound. In addition, about each feature quantity which has deviation in the feature numerical value, it reexamines by normalizing feature quantity.

1. まえがき

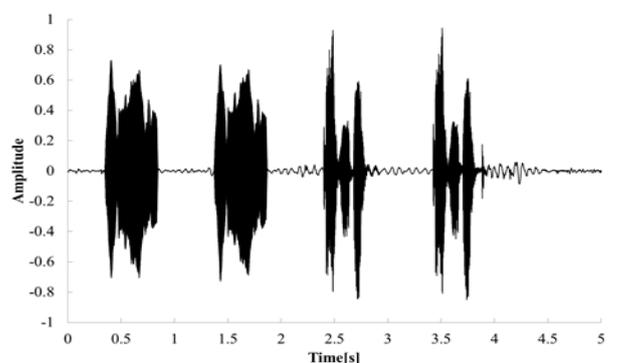
聴覚が不自由な人の生活補助を目的としたシステム構築における、雑音を考慮した場合の生活音の識別について検討する。これまでの報告で、生活音の識別をおこなう際の新たな特徴抽出方法として、信号持続曲線(SDC)を提案し、97.8%の平均識別率の結果を得てその有効性を示した[1]。

しかし、実生活には機械の駆動音、風の音などの様々な音が入り混じっており、システムの弊害となることが考えられる。そこで本報告では、標本に擬似的な雑音を付加することによって、本システムの雑音に対する耐性を評価する。また、特徴数値に偏りのある各特徴量について、特徴量を正規化することにより再検討をおこなう。

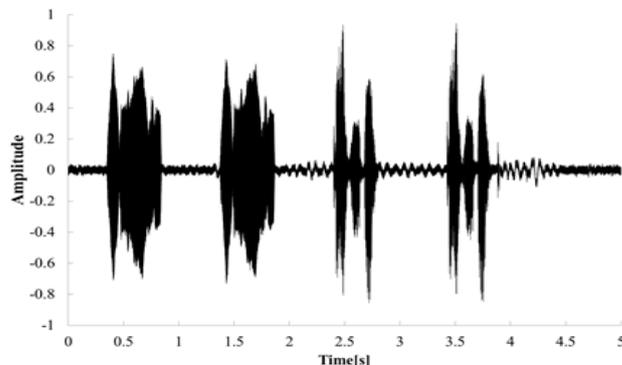
2. 雑音付加方法

文献[1]では、学習データとテストデータの両方に、雑音を付加させて識別をおこなった。本報告も同様に、雑音が全ての周波数帯域において、エネルギーが一律に分布しているガウス雑音を使用する。

ガウス雑音の付加の対象として、14種類の生活音に対して100個のサンプルを用意する。サンプルは全て録音時間5s、サンプリング周波数44.1kHz、量子化ビット16bitである。この100個のサンプルのうち、10個をテストデータとして試行毎にランダムで選出する。残りの90個は学習データとなる。このテストデータにガウス雑音を付加する。ガウス雑音は、標準偏差を各標本の振幅最大値の1%、平均を0とした。雑音を付加したのち、スケール $m=12$ まで離散ウェーブレット変換をおこなう。Figure 1(a), (b)に火災報知器音データのサンプル例と、ガウス雑音付加後のサンプル例を示す。



(a) Specimen



(b) Specimens were added noise

Figure 1. Sound data of fire alarm

3. 特徴量の正規化

本手法の識別では、雑音を付加した場合、雑音付加前と比較して平均識別率の低下が見られた。この原因として、文献[1]で使用した特徴量1~12と13~24の、取りうる数値の幅が違っていたことがあげられる。

特徴量1~12の周波数帯域幅はスケールの増大によって2進数レベルで低下する。また、特徴量13~24の信号減衰時間は0~2.5sの範囲の値に定まるのに対し、特徴量1~12はスケール0~11の間で変化する。そのため、下記の2種類の方法を用いて正規化をおこなう。

①特徴量 1~12 の周波数帯域幅を予め均等にするため、下記の式を用いる。

$$(\text{正規化エネルギー})_n = (\text{エネルギー})_n \times 2^n \quad (1)$$

②特徴量 1~12 と特徴量 13~24 の最大数値を同じにするために下記の式を用いる。

$$(\text{正規化特徴数値})_n = (\text{特徴数値})_n \times \frac{2.5}{11} \quad (2)$$

式(1), (2)共に($n=1, 2, \dots, 12$)であり n はスケールである。正規化無し, ①, ②, ①・②の 4 種類のパターンとスケール数を 8, 10, 12 に変更したときの計 12 パターンの識別をし、平均識別率の向上に対する検討をおこなう。

4. 識別方法

本報告では、学習データとテストデータの 2 種類からマハラノビス距離を用いた最短距離法によって生活音の識別をおこなう。また各正規化のパターンにおいて上記の識別を 1,000 回試行した平均識別率を求める。

5. 識別結果

12 パターンおこなったうち、一番平均識別率が高いのはスケール 12 で両方正規化したときである。Table 1 にスケール 12 で (I)ノイズなし, (II)正規化無しでの手法, (III)両方正規化したときの平均識別率結果を示す。

Table 1. A comparison of the average of the recognition rate

Category of living sound	The average of the recognition rate[%]		
	(I)	(II)	(III)
Opening a hinged door	97.9	0.0	0.0
Closing a hinged door	92.3	0.0	0.0
Opening a sliding door	97.3	0.0	0.0
Closing a sliding door	100.0	34.5	60.0
Phone ringtone1	98.0	0.0	0.0
Phone ringtone2	96.8	30.5	97.5
Flowing water	97.1	97.5	99.0
Voice(Ohayou)	97.7	0.0	0.0
Intercom	99.9	29.4	99.0
Bell of toaster	99.5	0.0	0.0
Tuning fork	99.8	89.6	99.1
Microwave oven	100.0	100.0	99.1
Fire alarm	97.8	52.1	66.5
Gas alarm	95.4	94.1	96.0
The average recognition rate	97.8	37.7	51.2

文献[1]と本手法の平均識別率を比較すると、本手法はテストデータのみ雑音を付加しているため、全体的に平均識別率が低下している。さらに、扉の開閉音やトースターの音などは平均識別率が 0% であった。これらの平均識別率低下の原因として、瞬間的な音が

雑音付加の影響によって、持続的な音と誤認識していることがあげられる。

次に、正規化の有無で平均識別率を比較すると、正規化によって 13.5% と識別率が向上した。この理由として、正規化をおこなったことにより周波数帯域幅が均等になったため特出した特徴量を捉えやすくなった。

また、雑音付加の影響が最も大きい特徴量 13 の学習データとテストデータの各カテゴリにおける平均値を比較した。Table 2 にその結果を示す。

Table 2. The average value of each category

Category of living sound	Learning data [s]	Test data [s]
Opening a hinged door	0.002	0.495
Closing a hinged door	0.002	0.383
Opening a sliding door	0.005	0.612
Closing a sliding door	0.149	0.597
Phone ringtone1	0.264	0.612
Phone ringtone2	0.471	0.561
Flowing water	0.544	0.550
Voice(Ohayou)	0.036	0.613
Intercom	0.561	0.613
Bell of toaster	0.012	0.529
Tuning fork	0.571	0.614
Microwave oven	0.221	0.245
Fire alarm	0.068	0.104
Gas alarm	0.472	0.535

この結果より特徴量 13~24 において、ノイズを付加することにより、学習データとテストデータの信号減衰時間にカテゴリによって差異があらわれていることがわかる。これが、平均識別率低下の原因になった可能性がある。

6. まとめ

本報告では、文献[1]の手法での雑音に対する評価をおこなった。しかし、雑音付加により瞬時的な音が持続的な音と誤認識してしまう問題点が生じた。そのため平均識別率向上にむけて特徴量の正規化をおこなった。このとき 13.5% の平均識別率の向上が見られた。

今後は、音が鳴っている場所のみに雑音を付加させる方法、雑音に影響されない特徴量の考案をおこなう予定である。

7. 参考文献

- [1] 外山紘之, 細野裕行: 「信号持続曲線を用いた生活音の識別の雑音評価」, 信学ソ大, A-4-8, 9.2014.
- [2] 池原雅章, 島村徹也: 「MATLAB マルチメディア信号処理」, 培風館, 2004.