

信号持続曲線の特徴量を用いた生活音の識別

Recognition of Living Sound Using The Feature Value of the signal duration curve.

○稲井孝行¹, 細野裕行², 伊藤彰義³*Takayuki Inai¹, Hiroyuki Hosono², Akiyoshi Itoh³

Abstract: In this study, we recognize “living sound” using the new feature value of the signal duration curve. This is a fundamental study aimed at supporting the people who are hard of hearing. We recognize 14 categories of “living sound” and calculate the average of the recognition rate. Additionally, we recognize 14 categories of “living sound” by using the method of the discrete Fourier transform and compare the average of the recognition rate with the previous one.

1. まえがき

聴覚が不自由な人の生活補助を目的とした研究の基礎的検討として、生活音の識別における新たな特徴抽出方法を提案する。一般的な音信号処理には短時間フーリエ変換が用いられている[1]。しかし、非定常な音信号に対して、短時間フーリエ変換に使用する窓の長さが固定であることから、近年では音信号処理にウェーブレット変換が使用されつつある[2]。

これまでの報告において、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、その有効性を示した[3]。

本報告では、SDC から得られる特徴量を変更して生活音の識別をおこない、14 種類の生活音サンプルにおいてこれまでの結果より良い識別結果が得られたことを示す。また、本手法の識別結果と離散フーリエ変換(DFT)を用いた生活音識別手法における識別結果を比較し、再現性が悪い音の識別結果、特徴量を新たに増やす際の容易さの観点から本手法の優位性を示す。

2. 特徴抽出方法

14 種類の生活音に対して、それぞれ 100 個のサンプルを用意する。1 つの音サンプルの長さは 5sec、サンプリング周波数は 44.1kHz とした。信号持続曲線は、音の元信号にウェーブレット変換をおこない、得られた高周波成分のエネルギー強度を求め、エネルギー強度の振幅を降順ソートすることにより作成する。本手法ではスケール番号 $m=12$ までウェーブレット変換をおこない、1 個の音信号に対して 12 個の SDC を求める。図 1 に求めた信号持続曲線の例を示す。

次に本手法で使用した特徴量について示す。特徴量 13~24 について、文献[3]で使用した特徴量を変更して識別をおこなった。

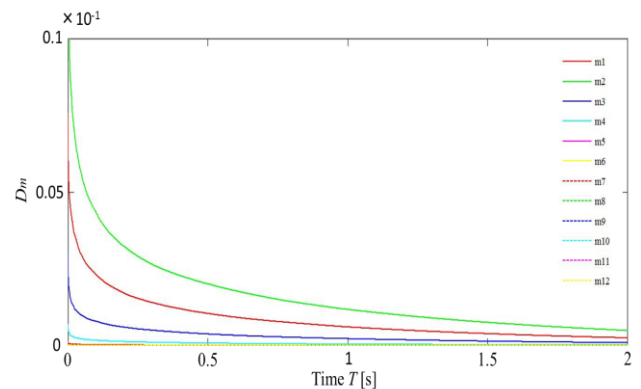


Figure 1. Signal duration curves

(1)SDC の特徴量 1~12

文献[3]と同様に、各スケール番号におけるウェーブレット変換結果から 12 個の SDC を求める。スケール番号を SDC のエネルギーが大きい順に順位付けし、順位に対応するスケール番号を特徴量 1~12 とする。

(2)SDC の特徴量 13~24

文献[3]では、SDC のエネルギーが大きい順に並び替えた信号減衰時間を、降順に特徴量 13~24 としていた。本手法では、信号減衰時間をエネルギーが大きい順には並び替えず、スケール番号の昇順に、各スケール番号における信号減衰時間を特徴量 13~24 とする。これにより、各周波数帯域ごとの信号減衰時間を比較することが可能となる。

DFT を用いた手法で使用した特徴量について示す。DFT の特徴量 1~12

各生活音サンプルを時間軸上で 12 等分の区間に分ける。1 区間約 0.4sec の各区間において DFT をおこなう。各区間でパワースペクトルが最大となる周波数を選び、パワースペクトルが大きい順に周波数を並び替え、特徴量 1~12 とする。

3. 識別方法

本報告では、各カテゴリにおいて各々10個のテストデータを無作為に抽出し、残りの90個を学習データと定める。これら2種類のデータからマハラノビスの距離を用いた最短距離法により生活音の識別をおこなう。前回の報告、本報告、DFTを用いた手法における特徴量それぞれにおいて上記の識別を100,000回おこない、平均識別率を求める。

4. 識別結果

表1に、文献[3]、本手法、DFTを用いた手法それぞれ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 [%]		
	Literature[3]	This study	DFT
Opening a hinged door	98.7	97.9	32.3
Closing a hinged door	90.9	92.3	74.7
Opening a sliding door	93.5	97.3	79.7
Closing a sliding door	98.7	100.0	58.8
Phone ringtone1	96.9	98.0	73.8
Phone ringtone2	92.4	96.8	85.2
Flowing water	92.5	97.1	97.9
Voice(Ohayou)	96.9	97.7	94.3
Intercom	98.0	99.9	97.4
Bell of toaster	93.6	99.5	95.9
Tuning fork	99.9	99.8	84.1
Microwave oven	99.0	100.0	79.2
Fire alarm	96.8	97.8	91.1
Gas alarm	94.6	95.4	96.7

文献[3]と本手法の平均識別率を比較すると、本手法の方が「開き戸を開ける音」、「音叉の音」を除き、同程度または良い平均識別率が得られ、全体で約1.9%向上した。「音叉の音」に関しては平均識別率算出の際の誤差であると考えられるが、「開き戸を開ける音」に関しては平均識別率が低下している。

DFTを用いた手法と本手法の平均識別率を比較すると、本手法の方が「蛇口から流れる水」、「ガス警報器」を除き、良い平均識別率が得られ、全体で約16.3%高い結果となった。特に、「開き戸を開ける音」や「開き戸を閉じる音」は、DFTを用いた手法では低い平均識別率となった。これは音のサンプルを録音する際に、人の力加減により各サンプルの音色が変化してしまったこと、さらに5sec間の中で同タイミングに音が録音できていなかったことが考えられる。このことから、本手法は再現性の悪い音に対する識別に優位であることが言える。また、新たに特徴量を増やす際に、容易に時間情報を特徴量として使用できることが言える。

使用したサンプルの標準偏差の変動から、変更した特徴量13~24を評価する。表2に文献[3]、本手法特徴量それぞれによる標準偏差の平均値を示す。

Table 2. The average of a standard deviation

Category of living sound	Literature[3]	This study
Opening a hinged door	0.03	0.01
Closing a hinged door	0.02	0.01
Opening a sliding door	0.04	0.01
Closing a sliding door	0.04	0.03
Phone ringtone1	0.07	0.05
Phone ringtone2	0.10	0.06
Flowing water	0.10	0.10
Voice(Ohayou)	0.04	0.02
Intercom	0.06	0.05
Bell of toaster	0.15	0.05
Tuning fork	0.13	0.11
Microwave oven	0.11	0.09
Fire alarm	0.06	0.05
Gas alarm	0.14	0.13

表2から、特徴量を変更したことにより、各音カテゴリの特徴量13~24の標準偏差が小さくなっていることがわかる。各音100個ずつのデータのばらつきが小さくなっていることが平均識別率向上の一因となったと考えられる。

5. まとめ

本報告では、ウェーブレット変換を用いた生活音の特徴抽出方法を提案し、14種類の生活音識別をおこなった。識別手法は前回の報告と同様であるが、使用する特徴量を変更したことにより、前回と比較して良い識別結果が得られた。

6. 参考文献

- [1] 猿舘朝, 伊藤憲三:「事前登録型生活音自動識別システム」, 電子情報通信学会技術研究報告, Vol.108, No.179, pp13-18, 2008.
- [2] 吉井圭吾, 木場俊暁, 金丸隆, 関根優年:「ウェーブレット変換を用いた音素マッチング処理」, 電子情報通信学会技術研究報告, Vol.102, No.529, pp123-128, 2002.
- [3] 山田貴俊, 細野裕行, 伊藤彰義:「ウェーブレット変換を用いた生活音の特徴抽出と識別」, 信学総大(通信), A-4-13, 3. 2013.
- [4] 新誠一, 中野和司:「図説 ウェーブレット変換ハンドブック」, 朝倉書店, 2005.