

ウェーブレット変換を用いた生活音の特徴抽出 Feature extraction of living sound using wavelet transforms.

○山田貴俊¹, 細野裕行², 伊藤彰義²*Takatoshi Yamada¹, Hiroyuki Hosono², Akiyoshi Itoh²

Abstract: In this study, we propose a new method of feature extraction from living sound which is one of the bases for constructing a new hearing aid. An automatic recognition system of “living sound” will be helpful for the people who are hard of hearing. We tried to utilize the wavelet transform for the features of “living sound”. We introduced and investigated about “Signal Duration Curves” as one of the features to distinguish live sound which are extracted by wavelet transform.

1. まえがき

本報告では、生活音の特徴抽出方法を提案する。これは、聴覚障害者の補助を目的とした研究の基礎的検討にあたる。生活音の特徴抽出方法にはウェーブレット変換を用いた。従来の生活音抽出法で使われているフーリエ変換^[1]と比較して、ウェーブレット変換は周波数成分と時間成分の両方を扱える。この利点から、音信号の持続性に着目し「信号持続曲線」を考案した。

2. 特徴抽出

検討した生活音の特徴は以下の二つである。

- ① ウェーブレット変換によって求めた元信号の周波数帯において、最も信号成分が多く存在している周波数帯。なお、今回は高周波数帯を扱う。
- ② ①で求めた周波数帯において、信号持続曲線が一定の基準まで減衰する時間。

本報告では、上記二つの特徴量から特徴空間を求め生活音の判別を試みる。

2-1 ウェーブレット変換

ウェーブレット変換^[2]はフーリエ変換と同じように解析対象の信号とウェーブレット関数(後述、窓関数に相当)を畳み込むことでおこなう。ウェーブレット変換は非定常な信号を時間と周波数の両方で解析できる。連続ウェーブレット変換と離散ウェーブレット変換のうち、今回は離散ウェーブレット変換を用いる。式を(1)、(2)に示す。

$$T_{m,n} = \int_{-\infty}^{\infty} x(t)\psi_{m,n}(t)dt \quad (1)$$

$$S_{m,n} = \int_{-\infty}^{\infty} x(t)\phi_{m,n}(t)dt \quad (2)$$

$T_{m,n}$ を詳細係数、 $S_{m,n}$ を近似係数、 $\psi_{m,n}$ をウェーブレット関数、 $\phi_{m,n}$ をスケーリング関数と呼ぶ。添字の

整数 m, n はウェーブレットの伸長と平行移動をそれぞれ調節するパラメータである。これら関数は多くの種類が存在し、関数ごとに特徴を持つ。今回は Coiflet N=5 の関数を選択した。

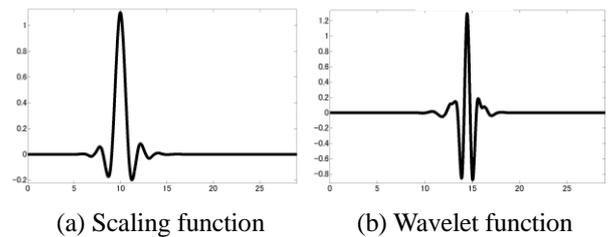


Figure 1. Used wavelet, coiflet N=5.

(1), (2)式を用いて元の信号 $x(t)$ を再構成する。変換式を(3)式に示す。

$$x(t) = \sum_{n=-\infty}^{\infty} S_{m_0,n} \phi_{m_0,n}(t) + \sum_{m=-\infty}^{m_0} \sum_{n=-\infty}^{\infty} T_{m,n} \psi_{m,n}(t) \quad (3)$$

上式は次式のように書くことができる。

$$x(t) = x_{m_0}(t) + \sum_{m=-\infty}^{m_0} d_m(t) \quad (4)$$

x_{m_0} を近似成分、 d_m を詳細成分とする。近似成分が低周波成分、詳細成分が高周波成分にあたる。この式が示しているのは、任意のスケールレベル(添字 m)における信号の詳細成分を同じスケールでの信号の近似成分を加えることで、より小さいスケールでの信号の近似が得られる。これを、多重解像度解析と呼ぶ。

2-2 信号持続曲線

ウェーブレット変換で得られた信号を以下の手順で処理し、信号持続曲線を求める。なお、今回は詳細成分を例とする。

- i. 詳細成分 d_m のパワースペクトル D_m を求める。

$$D_m(t) = |d_m(t)|^2 \quad (5)$$

- ii. D_m の振幅を降順ソートし、縦軸とする.
- iii. 降順ソートした結果から、横軸の時刻は順不同になる. つまり、時刻で連続したデータではない. しかし、サンプリング周波数は等しい(本報告は 44.1kHz)ので信号全体の時間の長さは不変である. 上記の手順で得られた信号を信号持続曲線と定義する. 信号持続曲線は信号全体から集めた大きな振幅がどのくらいの時間存在するかを意味する. 信号持続曲線の例を図 2 に示す.

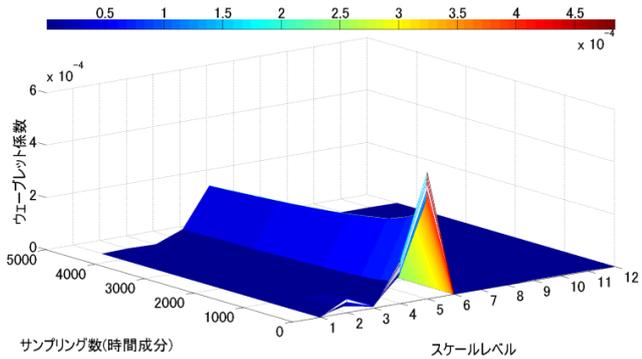


Figure 2. Signal duration curve.

2 - 3 特徴抽出方法

二つの特徴の抽出法について述べる. ウェーブレット変換から求まる詳細成分を用いた特徴と, 求めた詳細成分の減衰時間を用いた特徴である.

- 特徴 1: ウェーブレット変換により求めた信号の詳細成分(高周波成分にあたる)を用いる. 最も大きなエネルギーを持つ詳細成分を特徴量とする.
 - i. 各々のサンプル内の詳細成分 D_m において信号エネルギーをそれぞれ求める. ($1 \leq m \leq 12, m$ は自然数)
 - ii. i. で求めた詳細成分 D_m の信号エネルギーを比較し, 大きさの順に並べる.
- iii. 順位 1 位のスケールレベル m を特徴量 1 とする.
- 特徴 2: 方法 1 で求めた詳細成分において, 信号持続曲線のエネルギーが半分に減衰するまでの時間(信号減衰時間)を特徴量とする.
 - i. スケールレベル m の信号持続曲線を求める.
 - ii. 信号持続曲線のエネルギーが半分に減衰する (3db)までの時間(sec)を求める
 - iii. その時間の量を特徴量 2 とする.
 上記の特徴量を用いて, 生活音の特徴抽出を試みる.

2 - 4 生活音の特徴抽出

特徴抽出を行った生活音を表 1 に示す. 用いた生活音は録音時間 5 秒, サンプリング周波数 44.1kHz, 1 つの音につき 100 個である.

2 - 3 の特徴抽出法を用いて特徴空間を作った. 結果を図 3 に示す. 図 3 より二つの特徴量から 9 つの生活音をグループ分けできる. 特徴量 1 から生活音は大きくグループ分けでき, 信号減衰時間により更に分離できる. スケールレベルと信号減衰時間は生活音の識別に良い特徴量といえる.

Table 1. Living sound sample.

用いた生活音	生活音の説明
開き戸	開き戸を開く音
着信音	携帯電話の着信音
水が流れる音	蛇口から水が流れる音
人の声	人の声「おはよう」
インターホン	玄関のインターホン
トースターのベル	オーブントースターの終了音
火災警報器	火災警報器の警報音
ガス警報器	ガス漏れ警報器の警報音
音叉	440Hz 音叉を実際に鳴らした音

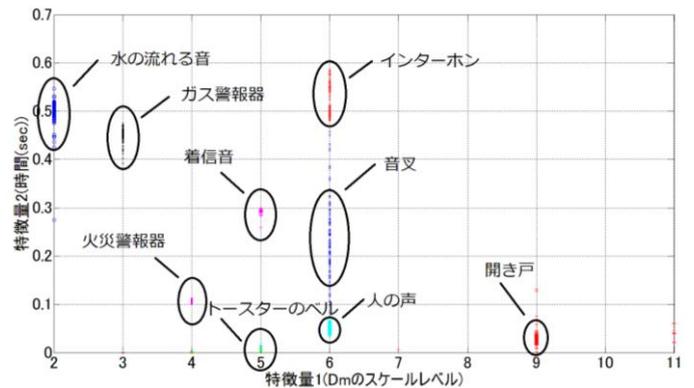


Figure 3. Distribution of "live sound" on the featurespace.

3. 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を試みた. また, その過程で信号持続曲線を考案した. 二つの特徴量を用いて特徴空間を作成し, 生活音判別の可能性を探った. 結果, 生活音をグループ分けすることができた.

今後は, サンプル音を追加する, 特徴量を追加する等の検討をおこなう予定である.

参考文献

[1] 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」電子情報通信学会技術研究報告. WIT, 福祉情報工学 No.102(493), pp.45-50, 2002.

[2] 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.