

エントロピーによる宅内音響の解析:人の活動有無の検出可能性

瀬戸 新一郎[‡] 浅井 紀久夫[‡]

Shinichiro Seto Kikuo Asai

1. はじめに

高齢化社会における課題の一つとして、独居老人の見守りが挙げられる。現在、企業からは種々の安否確認システムが商用化されている。他方、安否確認のみならず高齢者のフレイル検知を目的とした、ADL(Activities of Daily Living)や、IADL(Instrumental ADL)の検知に関する研究がなされている。ADLやIADLの特定は、学習を前提としており、その学習精度の向上が課題となっている。

本研究は、ADLやIADLの特定ではないが、学習レスで、独居者の生活音から、独居者が活動している状況にあるのか、そうでない非活動状況にあるか、その検出可能性を探るものである。学習レスにするのは、多様な対象音に簡単な仕組みで対応するためである。

宅内音響の音響シーンを周波数解析した上で、音の識別に用いる特徴量としてエントロピーを採用し、人の活動に由来する“活動音”か、それ以外の音である“非活動音”かを識別する。活動音/非活動音の識別の閾値は適切に設定する必要はあるが、特徴量抽出のための学習は不要である。

2. 宅内音響の分類と特徴

宅内音響を、次のA,B,Cに3分類する。Aは人が活動することにより発生する、活動音である。Bは電気や水圧により発生する音である。Bの場合、活動音の場合と、人の活動とは無関係に発生する非活動音の場合がある。Cは人の活動とは無関係に発生する、非活動音である。表1にその一例を示す。

2.1 多様性のある“活動音”と単調な“非活動音”

活動音と非活動音を区別するために、活動音には変化の多様性がみられ、非活動音は相対的に単調である、という仮説をたてる。

図2は、調理音(活動音)と、炊飯器音(非活動音)の、50秒間の信号波形とスペクトログラムである。調理音は音が間歇的であり波形も多種である。食材を水道で洗い、包丁でカットし、ガスを点火する等多種の動作音を含み、多様な音を発生し、複雑なスペクトログラムとなる。他方、炊飯器音は、周波数は広帯域であるが、時間的変化としての多様性はない。

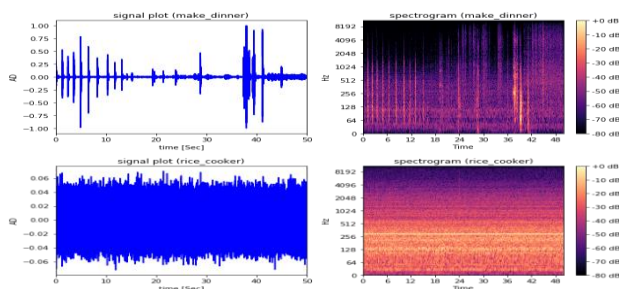


図2 調理音(上)、炊飯器音(下)の信号とスペクトログラム

このように、音響シーンにおける多様性の有無は、スペクトログラムを用いると直観的に把握が可能である。その判断を自動化することを、学習しない方式で実現することを考える。

2.2 音響シーンの多様性を表す特徴量

音響シーン全体の特徴を、2段階で求め、最終的に1変数特徴量で表す。特に重要なのが、2段階目のエントロピーである。これにより、以下に示す窓毎の特徴量の、時間的変化の多様性を捉えることができる。

第1段階で、音響シーンを細かい窓に分割し、窓内の音を対象に特徴量を求める。対象音が音声に限らない場合は、周波数分布と強度を表す特徴量として $\log_mel_energies[1]$ も候補となる。mel数に応じた特徴量ベクトルを得る。さらにこのベクトルを、mel数を階級数とし、各mel係数を度数とするヒストグラムとして捉える。そのヒストグラムから、エントロピーを求める。これが窓内の周波数分布の広がりを表す、1変数特徴量(E1)である。

第2段階で、窓毎の特徴量(E1)の、音響シーン全域における変化の度合い(変化の多様性)を把握する。音響シーン全体をn個の窓で分割すると、n個のE1の時系列データが得られる。このn個のE1のとり値によるヒストグラムを作成し、ヒストグラムの階級数と度数から、そのエントロピーを求める。これが第2段階のエントロピーE2である。E2は、E1即ち窓毎の周波数分布の、音響シーン全体での変化の多様性を示す。このE2を、Entropy of Entropy (EofE)と略称する。2段階エントロピーに関しては、心電図解析に適用した報告がある[2]。

EofEは、音響シーンにおける音の変化が多様であればあるほど、高い値を示す。活動音は非活動音に較べて変化の多様性が大きいという仮説が正しければ、活動音のEofEは、非活動音のEofEより高い値をとると期待できる。実際、炊飯器音のEofEは0.17で、調理音は0.78である。

次に、公開音源を用いて、より広い範囲で検証を行う。

3. 提案手法の有効性を評価する実験

公開 freesound 音源[3]の中から、検証対象として、A/B/Cに該当する53クラス、778音源を選択した(表1)。各音源は10秒である。音響シーン解析の窓長は93msec、mel数は32である。

3.1 A/B/CのEofE値

結果の各クラスの平均EofE値を棒グラフとして図3に示す。Aに属するクラスを赤色系として示す。これらはEofEが高くグラフの右側に位置する。Bに属するクラスは、青、緑、肌色で、EofEが低くグラフの左側に位置する。Cに属するクラス中、戸外からの音(濃緑)は低いEofEで、

[‡] 放送大学

ット(4 本の水色)は比較的高い EofE 値を示す。ペットは活動音の判断を行う際外、乱要因となる可能性がある。

表 1 検証に用いた音源のクラス一覧

| 種類 | クラス群 | クラス名 |
|----|-------|---|
| A | 咽喉音 | 会話、女性の声、笑い、男性の声、歌、叫び、絶叫、喉、あくび、嗅ぎ、くしゃみ、うなる |
| | 肉体音 | 指で叩く、拍手、歩く、走る |
| | 操作音 | スラム、引き戸、食器、引き出し、食器棚、きしみ、やかん、キー、ノック、はさみ、タイプライター、ジッパー、コンピューターのキーボード、コイン(落下)、歯ブラシ、シャッフルカード、梱包テープ、ミシン、チョッピング、ライティング |
| B | メディア音 | テレビ、ラジオ |
| | 機械音 | 扇風機、エアコン、電子レンジ、掃除機、ドライヤー |
| | 水音 | 給水栓、流し台、浴槽、洗面所 |
| C | 犬、猫 | ニャー、ハウル、イップ、ウィッパー、 |
| | 戸外 | 雨、雨滴 |

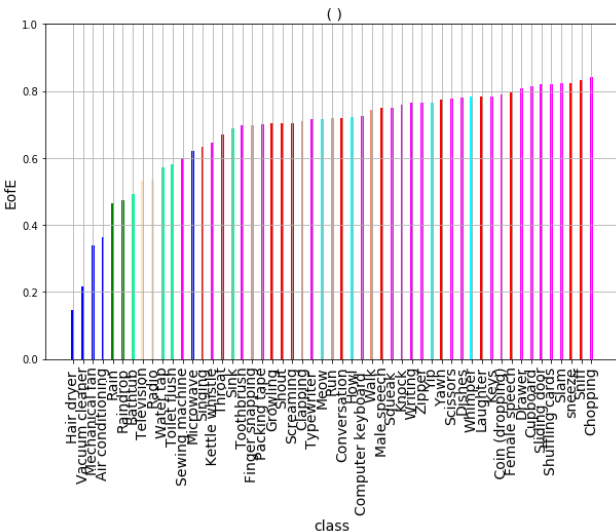


図 3 クラス平均の EofE

3.2 A/B/C と活動音、非活動音との対応

2.で述べたように、A は活動音、C は非活動音である。B は活動音と非活動音の場合がある。水道音を例にすると、水を流しっぱなしにしている場合の音は非活動音であるが、人が手を洗ったり、皿を洗ったり、人の行動が音に反映されている場合は活動音である。今回の B に限ると、電子レ

ンジ以外は非活動音である。電子レンジの場合、扉の開閉や設定等人的の操作がある場合は活動音となる。そこで電子レンジ音を、非活動音(青)と活動音(赤)に分離し、その EofE を図 4 に示す。活動音の EofE は、非活動音のそれよりも高いことがわかる。電子レンジの場合も、人の行動により、音の変化多様性が大きくなると考えられる。

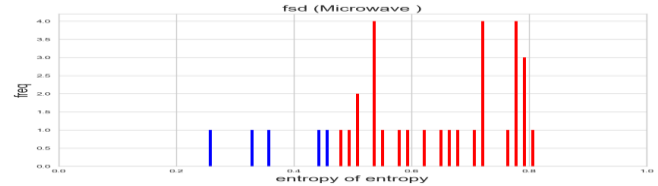


図 4 電子レンジの活動音(赤)と非活動音(青)の EofE(0~1)と度数

3.3 活動音と非活動音の EofE の差異

対象音源の EofE を求め、その度数分布を、活動音(赤)/非活動音(青)として図 5 に示す。活動音は相対的に EofE が高いことがわかる。EofE(活動音)>EofE(非活動音)と仮定し混同行列を求める。その ROC 曲線の AUC を求めると、AUC=0.79であった。また、非活動音で EofE が高いペットを除外した場合は、AUC=0.88 であった。これから、閾値を適切に設定することで、音響シーンの EofE が高い場合は、ある程度の確度で、人の活動を想定できるといえる。

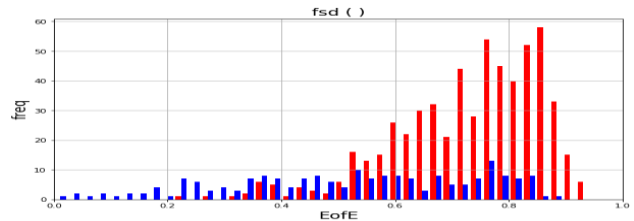


図 5 活動音(赤)と非活動音(青)の EofE(横軸 EofE、縦軸度数)

4. おわりに

本研究は、エントロピーをもちいた特徴量で音響シーンを捉え、人の活動に起因する活動音と、そうでない非活動音を識別するものである。例としては、定性的には炊飯器/調理の例と freesound の電子レンジの例、定量的には freesound 音源で(AUC=0.79)が成立することを示した。いずれも、相対的に高いエントロピーを示す場合、背景に人の活動がある。このように、活動音と非活動音の識別が、エントロピーを用いた特徴量で、学習レスでできる可能性があるが、さらに 24 時間連続実生活音を対象とした検証が必要である。

参考文献

[1] V. Wellekens and C. Tyagi, "On desensitizing the Mel-Cepstrum to spurious spectral components for Robust Speech Recognition," Proc_IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. 529-532. 2005
 [2] Hsu, C.F.; Wei, S.Y.; Huang, H.P.; Hsu, L.; Chi, S.; Peng, C.K. Entropy of Entropy: Measurement of Dynamical Complexity for Biological Systems. Entropy 2017, 19, 550.
 [3] <https://freesound.org/browse/> (2020.05.25 現在)