

聴取者の主観評価に基づく音地図作成のための環境音収録

Sound recording to construct sound-map on the basis of listener's subjective evaluation

原 直

Sunao Hara

岡山大学 阿部研究室

Abe Laboratory, Okayama University

概要 本研究では、「聴取者の主観を含んだ音の収集」と「大規模な音の物理量の収集」という二つのデータ収集を通して、聴取者の主観を伴った大規模な環境音データベース構築を目指す。Participatory sensing と Opportunistic sensing というクラウドセンシングの考え方に着目する。本報告では長時間にわたる特定地域での環境音収録を行い、収録された騒音レベルの可視化において聴取者の主観を反映させる方式について検討を行った。

1 はじめに

環境音は人が環境を理解するための重要な情報の一つであるが、聴取者の状況や主観的な印象によって環境音の印象が変化することがある。例えば音の大きさの物理量として騒音レベルが知られているが、同じ騒音レベルの環境音でも場の状況や聴取者により主観的な音の大きさは異なると考えられる。従って環境音の研究を進めるためには聴取者の主観を伴った大規模な環境音データベースが必要である。

大量のデータを集めるにはライフログ [1, 2] やクラウドソーシング [3, 4] のアプローチが有効である。しかし、ライフログとしての環境音の収集は聴取者の主観を含んだデータが収集できるが、プライバシーの問題から研究用データベースへの活用は容易ではない。一方で、クラウドソーシングのアプローチに基づく環境音収集においてプライバシーに配慮するならば、音そのものではなく音の特徴の一部を表した物理量を収集することは可能と考えられる。

本研究では、「聴取者の主観を含んだ音の収集」と「大規模な音の物理量の収集」という二つのデータ収集を通して、聴取者の主観を伴った大規模な環境音データベース構築を目指す。これまでにスマートデバイスによる環境音収録アプリケーションの開発と収録実験を行ってきた [6]。本報告では長時間にわたる特定地域での環境音収録を行い、収録された騒音レベルの可視化において聴取者の主観を反映させる方式について検討を行った。

2 環境音収録アプリケーション

本研究で実施する収録実験のために、環境音収録アプリケーション [6] を改良した。このアプリケーションでは、音を保存する際に、その場の音を主観的に評価するために騒音度と混雑度の 5 段階評価を与える事ができ

る。さらに、その場の状況や聞こえてくる音を記録するために、任意のテキスト入力エリアと 12 種類のあらかじめ設定された音の種類が用意されている。ユーザはツイートボタンを押すことで様々な情報を付与した音声を記録することができる。システムは、ツイートボタンが押されると、リングバッファから直近 10 秒の音声波形を RIFF wav ファイルとして作成すると同時に、主観評価の選択項目とテキスト入力の内容を表す文字列に時間をつけてログファイルに記録する。アプリケーションでは「音声」、「位置履歴ログ」、「騒音レベルログ」、「ツイートログ」がそれぞれファイルとして生成される。

音声は量子化ビット数 16 bit、サンプリングレート 32000 Hz で収録が行われる。1 秒毎に得られる音声サンプル $x[n], n = 1, \dots, N$ に対して A 特性 $A[k]$ に基づく等価騒音レベル L_{eq} を以下の計算式で計算する。

$$X[k] = \frac{1}{N} \sum_{n=0}^{N-1} x[n] \cdot e^{-2\pi jkn/N} \quad (1)$$

$$L_0 = 10 \log_{10} \frac{1}{K} \sum_{k=0}^{K-1} |A[k] \cdot X[k]|^2 \quad (2)$$

$$L_{eq} = B(L_0) \quad (3)$$

ここで、 N は計算に用いる信号長であり、 K は FFT 長 ($K > N$) であり、 $B(\cdot)$ は量子化された音声波形のパワーを音圧レベル値に変換する関数である。本報告では 1 秒間に相当する $N = 32000$ を利用し、 $B(\cdot)$ は騒音計 (RION NL-42) を用いて事前に収録された校正値の線形補間により関数を作成した。

3 環境音収録

3.1 収録の概要

6 名の実験参加者によって岡山市内の 4 地域を重点的に収録した。4 地域の主な属性は「閑静な住宅地」、「駅前」、「駅に近い商店街」、「駅からやや離れた商店街」である。収録は 2014 年 11 月 22 日から 29 日の間で、平日 2 日と土日 2 日の合計 4 日間として、1 日あたり 2 地点を 3 名ずつ巡回することで午前 8 時から午後 9 時までの環境音と等価騒音レベルを収録した。1 時間を 1 セッションとして地域毎に定めた経路を巡回する。巡回時には端末 (Google Nexus 7) を手に持ったまま収録しており、経路上の特定の位置で主観評価や環境音の収録を行うように指示した。

主観的な騒音度合いの選択肢は、 L_1 : とても静か、 L_2 : 比較的静か、 L_3 : やや騒がしい、 L_4 : かなり騒々しい、

表 1: 収録された環境音の出現頻度分布

Class	#data	Class	#data
T_1 人の声	1356	T_7 踏切の音	1
T_2 鳥の声	634	T_8 電車の音	97
T_3 虫の音	56	T_9 サイレン音	37
T_4 車の音	2267	T_{10} 音響信号機	788
T_5 風切音	185	T_{11} 音楽	374
T_6 バイクの音	727	T_{12} 動物の声	61

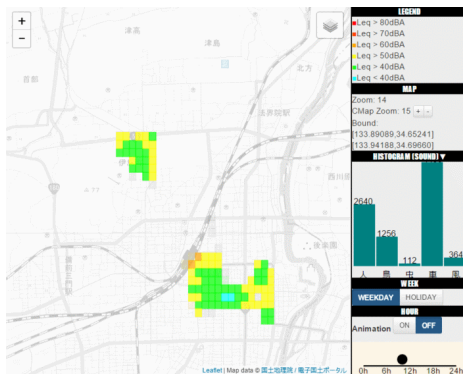


図 1: 地図上での環境音可視化システム (騒音表示)

L_5 : とても騒々しい, の 5 段階である. また, 主観的な混雑度合いの選択肢は, C_1 : 誰もいない, C_2 : まばら, C_3 : やや多い, C_4 : かなり多い, C_5 : 混雑, の 5 段階である. 環境音の種類は表 1 の 12 種類として, 複数選択を可能にした. 参加者には, すべての項目を埋める必要はないことを指示している.

3.2 収録データの予備分析

収録実験により収集された位置履歴ログファイルと騒音ログファイルの時間同期を行った結果, 緯度経度と等価騒音レベルをすべて持つデータは 367,242 組得られた. 音声データの総数は 2,740 サンプルであり, 全て 10 秒の音声ファイルである. 収録された音声データにつけられた環境音ラベル毎の頻度分布を表 1 に示す.

等価騒音レベルの可視化を行った例を図 1 に示す. この図は収録実験で得られたデータから緯度経度を座標軸とした二次元ヒストグラムによる集計を行い, 平均値をその地域の代表的な騒音レベルとして色分け表示している. 可視化システムではアニメーション表示をしており, 時間帯の移り変わりに伴う音の変化を見ることが観察できる.

3.3 等価騒音レベルと主観評価の比較

図 2 は主観評価に対する等価騒音レベルの平均値を示している. 収録データ数が少ないことから, 「かなり多い (C_4)」と「混雑 (C_5)」は一つのクラス $C_{4,5}$ とした. いずれの図についても主観的な感覚と客観的な指標がおおよそ一致していることから, 人による評価のばらつきは小さいと考えられる. ただし, 図 2a において「と

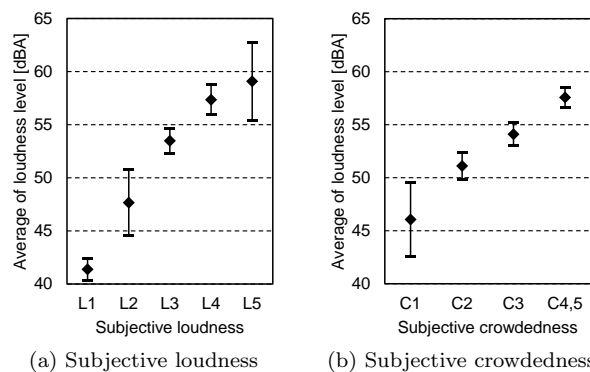


図 2: 主観評価に対する等価騒音レベルの平均値. エラーバーは 90% 信頼区間を示す.

ても騒々しい (L_5)」の信頼区間が広がっていることから, 「かなり騒々しい (L_4)」との弁別基準は人によるばらつきが大きいことが示唆される.

4 まとめ

本報告では, タブレット型端末を用いた岡山市 4 地域における環境音収録について述べた. また, 主観的な混雑度合いの可視化に向けた GMM-UBM による推定手法について検討した. 推定の有効性は示唆されたが, 今後は推定手法の詳細な検討が必要である.

参考文献

- [1] A. Mesaros, *et al.*, “Acoustic event detection in real life recordings,” Proc. of 18th European Signal Processing Conference (EUSIPCO-2010), pp. 1267–1271, Aug. 2010.
- [2] M. Shah, *et al.*, “Lifelogging: Archival and retrieval of continuously recorded audio using wearable devices,” Proc. of IEEE International Conference on Emerging Signal Processing Applications (ESPA-2012), pp. 99–102, Jan. 2012.
- [3] R. Rana, *et al.*, “Ear-Phone: An end-to-end participatory urban noise mapping system,” Proc. of The 9th ACM/IEEE International Conference on Information Processings in Sensor Networks (IPSN 2010), pp. 105–116, Apr. 2010.
- [4] E. D’Hondt, *et al.*, “Participatory noise mapping works! an evaluation of participatory sensing as an alternative to standard techniques for environmental monitoring,” Pervasive and Mobile Computing, vol. 9, no. 5, pp. 681–694, Oct. 2013.
- [5] “The HTK Book,” <http://htk.eng.cam.ac.uk/>.
- [6] 原他, “クラウドセンシングにより収集された環境音のシンボル表現を用いた音地図構築手法,” 音講論, pp. 1535–1538, sept. 2014.