

知らない音も書き起こす

～ ベイジアン半教師あり音響イベント認識 ～

どんな研究

膨大なメディアデータを自在に検索して活用するためには、メタデータに頼るだけでなく、音や映像自体から中身の情報を自動的に引き出す技術が必要不可欠です。この展示では、音響信号に含まれる**音響イベント（音によって人に認識される事象）**を書き起こす技術を紹介합니다。

どこが凄い

一部区間のラベルを手がかりとして、重なり合った音響イベントの全区間を特定します。さらに、音響信号に存在する音響イベントの数を学習することにより、ラベルに無い音を、「**知らない音（未知の音響イベント）**」として、書き起こすことができます。

めざす未来

人間の耳のように、どんな音が、どこで、どんな風に鳴っているかを聞き分けて認識することで、クラウドにおける大量の音コンテンツの可視化を目指します。これにより、**特定の音を含む動画検索や危険な音を知らせてくれる監視・行動支援システム**など、様々な応用が期待されます。

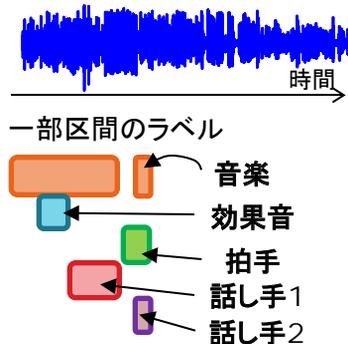
「どんな」音響イベントが、「いつ」鳴っているか認識したい



☹️ 難しさ

- ・音響イベントは重ね合わさっている
- ・音響イベントがどんな音かを、事前に学習しておかなくてはならない
- ・知らない音は認識できない

半教師あり学習によって、音響イベントの全区間を特定して認識する

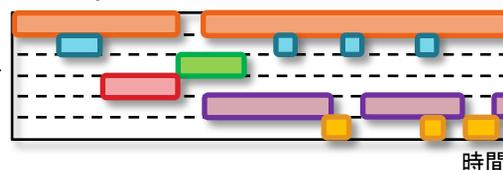


キーポイント：
音の重なりを考慮しながら、存在する音響イベントの数を学習すること

どんな特徴の音響イベント？

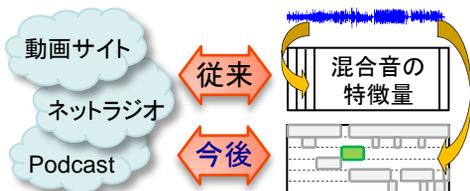
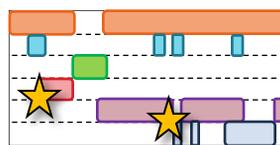


いつ？



適用例(期待される効果)

特定の音を含むコンテンツ検索

危険な音を知らせてくれる
監視・行動支援システム音声表現認識(話し方や
歌い方の分成合成)の前処理

関連文献

[1] Y. Ohishi, D. Mochihashi, T. Matsui, M. Nakano, H. Kameoka, T. Izumitani, K. Kashino, "Bayesian semi-supervised audio event transcription based on Markov Indian buffet process," in *Proc. 38th International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2013.

[2] 大石康智, 持橋大地, 松井知子, 中野允裕, 亀岡弘和, 泉谷知範, 柏野邦夫, "mIBPに基づくベイジアン半教師付き音響イベントダイアライゼーション," 日本音響学会2013年春季研究発表会, 1-9-6, pp. 19-22, 2013.

連絡先

大石 康智 (Yasunori Ohishi) メディア情報研究部 メディア認識研究グループ
E-mail : ohishi.yasunori@{at}lab.ntt.co.jp ({at} の部分を @ に置き換えてください)