

どんな研究

収録音声に含まれる雑音や残響を取り除く技術（音声強調技術）の研究は今までも沢山行われてきました。しかし、処理音声には歪み（残留雑音・残響、不自然な音）が残り、聞き取りにくい音となってしまうことがしばしばでした。この研究は、**処理音に歪みが全く残らない音声強調の確立**を目指した研究です。

どこが凄い

紹介技術は、膨大な量の学習用クリーン音声事例を背後に持ち、収録音に含まれるクリーン音声に最も合致する素片をその中からうまく探し出して、収録音声をその素片に置き換えます。クリーン音声素片の接続で作り出される出力音には、歪みが入りこむ余地はなく、**超高品質な音声強調が可能**となります。

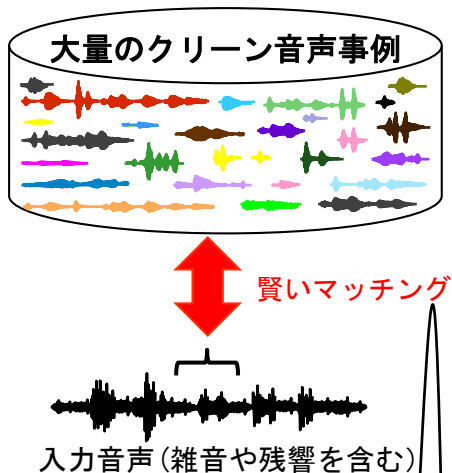
目指す未来

超高品質音声強調が可能となれば、例えば、今まで快適に行うことが難しかったスマートフォンやウェアラブル端末でのハンズフリー通話や、様々な場面での高品質音声録音等が可能となります。また、この技術は、音声認識性能の向上にも役に立ちます（雑音下音声認識コンペCHiMEチャレンジ2011でトップスコアを達成）。

事例ベース音声強調処理の概念

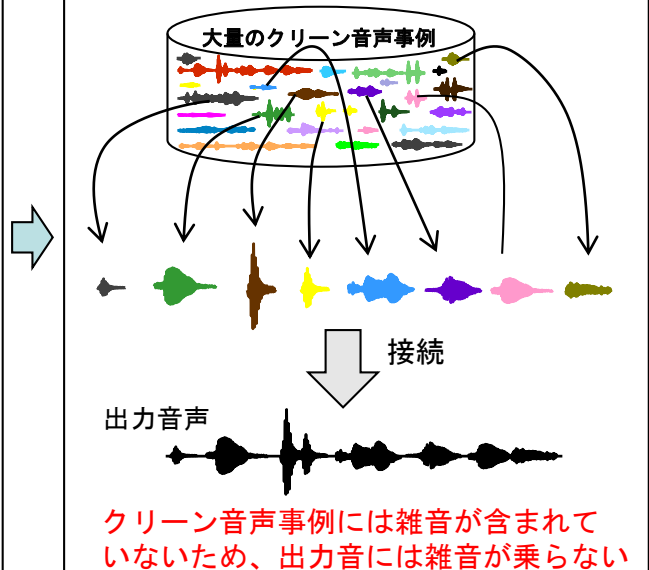
Step-1

雑音や残響を含む入力信号と大量のクリーン音声事例を賢くマッチング。入力信号内に含まれるクリーン音声成分と最もマッチする事例を探索・推定。



Step-2

マッチした中で最長のクリーン音声事例を選択・接続し、出力音を作成。(※)



入力信号に含まれるクリーン音声成分にマッチするクリーン音声事例を探すための「賢いマッチング」とは？

現在までに、以下の2つの方法（相補的な関係）を提案

- ・ マッチングの方策1[1]: マッチングの不確かさを確率的に判断しながら、最長類似事例を探索
- ・ マッチングの方策2[2]: クリーン音声事例と対を成す、雑音音声事例を人工的に大量に作成。そのデータと入力信号の照合を高速探索アルゴリズムを用い、最長類似事例を探索

(※) 事例接続により得られた信号を音声強調フィルタ計算に利用し、フィルタリング処理により音声強調を行うことも可能（不特定話者の処理への対応等に有効）

関連文献

- [1] K. Kinoshita, M. Souden, M. Delcroix, T. Nakatani, "Single channel dereverberation using example-based speech enhancement with uncertainty decoding technique," in *Proc. Interspeech*, pp. 197-200, 2011.
- [2] A. Ogawa, K. Kinoshita, T. Hori, T. Nakatani, A. Nakamura, "Fast segment search for corpus-based speech enhancement based on speech recognition technology," in *Proc. ICASSP*, pp. 1557-1561, 2014.

連絡先

木下慶介 (Keisuke Kinoshita) メディア情報研究部 信号処理研究グループ
E-mail : kinoshita.k(at)lab.ntt.co.jp

