

雑音抑圧音声の客観品質評価に用いる総合品質推定モデルの改良

An Improvement of Overall Quality Estimation Model for Objective Quality Evaluation of Noise-Reduced Speech

藤田悠希
Yuki Fujita

山田武志
Takeshi Yamada

牧野昭二
Shoji Makino

筑波大学
University of Tsukuba

1 はじめに

雑音抑圧音声の主観品質評価法を定めている ITU-T 勧告 P.835[1] では、被験者は雑音抑圧音声の音声成分のみに着目したときの音声品質、及び雑音成分のみに着目したときの雑音品質を各々評価し、その後に総合品質を評価する。我々はこの評価過程に注目し、音声品質と雑音品質から総合品質を推定する総合品質推定モデルを構築し、さらに本モデルを用いたフルリファレンス型、及びノンリファレンス型の客観品質評価法を開発した [2, 3]。

これらの手法により雑音抑圧音声の品質を高い精度で推定できるようになったものの、依然として実用レベルには達していないのが現状である。本稿では、まずは総合品質推定モデルの改良を行い、その有効性を検証する。

2 総合品質推定モデル

総合品質推定モデルを用いた客観品質評価法では、品質評価対象である雑音抑圧音声（とフルリファレンス型の場合はリファレンスとするクリーン音声）から各種の特徴量を得る。そして、これらの特徴量から音声品質 Q_s と雑音品質 Q_n を各々推定する。最後に音声品質と雑音品質の推定値を総合品質推定モデルに代入することにより総合品質 Q_o を得る。

これまでに我々は、総合品質推定モデルとして次式に示す単純な線形モデルを用いていた。

$$Q_o = 0.6303 \times Q_s + 0.6125 \times Q_n - 1.3917 \quad (1)$$

しかし、特に総合品質が低いときに十分な推定精度が得られないという問題があった。そこで本稿では、音声品質、雑音品質、総合品質の関係をより適切に表すモデルとして次式を提案する。

$$Q_o = \sum_{i=1}^2 c_{s_i} Q_s^i + \sum_{i=1}^2 c_{n_i} Q_n^i + c_{sn} Q_s Q_n + c \quad (2)$$

ここで、 $c_{s_1} = 0.3558$, $c_{s_2} = -0.0661$, $c_{n_1} = -0.0423$, $c_{n_2} = -0.0298$, $c_{sn} = 0.2199$, $c = 0.4202$ である。これらの値は、後述する主観品質評価実験の結果に基づいて決定した。

3 提案モデルの有効性の検証

P.835 により定められる主観品質評価実験を行うことにより、雑音抑圧音声の音声品質、雑音品質、総合品質を得た [3]。被験者は男性 22 名、女性 10 名の計 32 名であり、防音室内でヘッドホンにより音声サンプルを受聴した。本実験ではサンプリング周波数を 8kHz とした。音声サンプルとして男性 2 名、女性 2 名の計 4 発話を用意した。発話内容は連続した 2 つの日本語文である。これらの音声サンプルに、走行自動車内雑音、展示会場雑音、列車雑音、白色雑音を計算機上で加算することにより、雑音重畳音声を生成した。ここで、SNR は Clean、

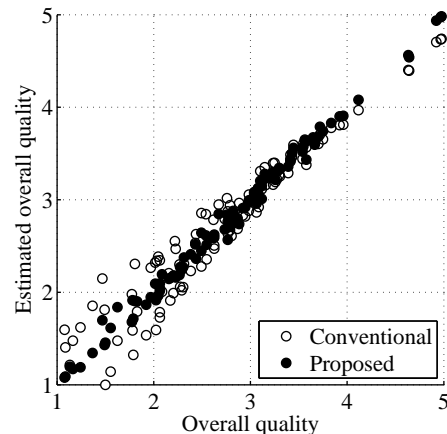


図 1 総合品質の推定結果

20, 15, 10, 5, 0 (dB) の 6 種類、雑音抑圧アルゴリズムは雑音抑圧を行わない場合を含めて 5 種類である。

本実験により得られた主観音声品質と主観雑音品質の値を式 (1) の従来モデルと式 (2) の提案モデルに各々代入し、総合品質を推定した結果を図 1 に示す。ここで、横軸は主観総合品質、縦軸は客観総合品質である。各点が表す品質は 128 個の評価値 (32 名 × 4 発話) の平均である。決定係数と RMSE は、従来モデルの場合は各々 0.93, 0.23, 提案モデルの場合は各々 0.99, 0.07 であり、提案モデルは従来モデルよりも高い精度で推定できていることが分かる。また、RMSE の目標値を主観評価により得られた総合品質の 95% 信頼区間片側幅とすると、RMSE の目標値は 0.12 となる。提案モデルの RMSE はこれをクリアしていることが確認できる。

4 おわりに

本稿では、新たな総合品質推定モデルを提案し、提案モデルにより実用に耐え得るレベルの推定精度が得られることを示した。今後は、提案モデルの入力となる音声品質と雑音品質の推定精度の改善に取り組む予定である。

謝辞 本研究の一部は科研費 19300271 の助成による。

参考文献

- [1] ITU-T Rec. P.835, "Subjective test methodology for evaluating speech communication systems that include noise suppression algorithm," Nov. 2003.
- [2] 篠原ら, "雑音抑圧音声の総合品質推定モデルを適用したフルリファレンス客観品質評価法," 信学総大 B-11-2, Mar. 2010.
- [3] T. Yamada *et al.*, "Non-reference objective quality evaluation for noise-reduced speech using overall quality estimation model," IEICE Trans. Comm., Vol. E93-B, No. 6, pp. 1367-1372, June 2010.