

# ミュージカルノイズを考慮した雑音抑圧音声の FR 型客観品質評価の検討\*

藤田悠希, 山田武志, 牧野昭二, 北脇信彦 (筑波大)

## 1 はじめに

駅構内のような騒々しい環境で音声通話を行うと、音声周囲の雑音に埋もれてしまい、正常に会話を行うことが困難となる。この問題を解決するためには、音声に重畳している雑音を信号処理により抑圧することが有効である。しかし、雑音の音量感が低減する一方で、音声成分にはひずみが生じ、また抑圧しきれなかった雑音成分が聞き苦しい音に変形するという副作用が起こる。その程度は、抑圧対象となる雑音の特性や雑音抑圧の原理によって異なり、ユーザ体感品質に大きな影響を及ぼす。したがって、雑音抑圧アルゴリズムの性能評価・性能比較はもちろんのこと、音声通話サービスの品質設計・品質管理を行うにあたって、雑音抑圧音声の品質を適切に評価する手法の確立が必要不可欠である。

音声の品質評価は、人間が実際に被評価信号を受聴し、その品質を主観的に判断することを基本とする。これを主観品質評価という。音声の総合的な品質の評価には、平均オピニオン評点 (MOS: Mean Opinion Score) がよく用いられる。これは、多数の被験者が被評価信号に付けた評点の平均値として定義される。雑音抑圧音声の主観品質評価法は、ITU-T 勧告 P.835[1] により定められている。しかし、主観品質評価を実施するには、専用の設備・機器や多大な時間・労力が必要となる。したがって、被評価信号から品質に対応する特徴量を抽出し、その特徴量から主観品質を推定する手法、すなわち客観品質評価法が求められる。

これまでに我々は、雑音抑圧音声の総合品質を推定するモデル (総合品質推定モデル) [2]、及びそれを用いたフルリファレンス (FR: Full Reference) 型客観品質評価法を提案している [3, 4]。提案手法は P.835 における品質評価過程を模擬したものであり、まず音声品質と雑音品質を個別に推定し、そして推定した音声品質と雑音品質から総合品質を推定する。提案手法によって従来よりも高精度に総合品質を推定できることを示しているものの、主観品質評価の代替として利用できるレベルには達していないのが現状で

Table 1 Conditions of the subjective test.

音声サンプル	男女各 2 名の計 4 発話
発話内容	連続した 2 つの日本語文
サンプリング周波数	8 kHz
雑音	走行自動車内, 展示会場, 列車走行音
SNR	Clean, 15, 10, 5, 0 (dB)
雑音抑圧手法	4 種類 (適用なしを含む)
受聴サンプル数	208 個
被験者	32 名
受聴環境	防音室内でのヘッドホン受聴

ある。

本稿では、提案手法の推定精度をさらに高めるために、P.835 において明示的に評価されていないミュージカルノイズに着目する。まず、ミュージカルノイズが音声品質と雑音品質に及ぼす影響を明らかにする。次に、ミュージカルノイズの影響を考慮した雑音品質推定モデルを提案し、その有効性を示す。最後に、提案する雑音品質推定モデルを FR 型客観品質評価法に適用する。

## 2 ミュージカルノイズが音声品質と雑音品質に及ぼす影響

### 2.1 主観品質評価試験

雑音抑圧の副作用として、ミュージカルノイズと呼ばれるトーン性の雑音が生じることが知られている。しかし、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] においては、ミュージカルノイズに関する明示的な評価はなされていない。そこで、ミュージカルノイズが音声品質と雑音品質に及ぼす影響を調査するために、主観品質評価を実施した [5]。

まず、P.835 により定められる主観品質評価を実施し、雑音抑圧音声の音声品質、雑音品質、総合品質を得た。次に、ミュージカルノイズに関する詳細な説明を行った上で、ミュージカルノイズにのみ注目したときのミュージカルノイズ品質

\* Full reference objective quality evaluation for noise-reduced speech considering effect of musical noise. by Yuki Fujita, Takeshi Yamada, Shoji Makino, and Nobuhiko Kitawaki (University of Tsukuba)

Table 2 Correlation coefficients among the three qualities.

	音声品質	雑音品質	MN 品質
音声品質	1.00	-	-
雑音品質	0.34	1.00	-
MN 品質	0.09	0.78	1.00

を得た．ここで，ミュージカルノイズ品質は5段階絶対品質評価尺度により表される（1が最低品質，5が最高品質）．

試験条件を Table 1 に示す．被験者は男性 19 名，女性 13 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．ここで，サンプリング周波数は 8 kHz である．音声サンプルは，男性 2 名，女性 2 名の計 4 発話であり，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，電子協騒音データベース [6] の走行自動車内雑音，展示会場雑音，列車走行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．SNR は Clean, 15, 10, 5, 0 (dB) の 5 種類である．雑音抑圧アルゴリズムとしては，Enhanced Variable Rate Codec に含まれている雑音抑圧法 [7]，スペクトル減算と振幅抑圧の相互制御に基づく雑音抑圧法 [8]，Wiener フィルタを用いた雑音抑圧法 [9] の 3 種類に加え，雑音抑圧を行わない場合の計 4 種類を用いた．その内の 2 種類は，ミュージカルノイズが顕著に発生するように調整した．音声サンプルの総数は 208 である．

## 2.2 主観品質評価試験の結果と考察

2.1 節の主観品質評価によって得られた音声品質，雑音品質，ミュージカルノイズ品質の相関係数を Table 2 に示す．ここで，MN はミュージカルノイズを表す．ミュージカルノイズ品質が 4 以上の場合はミュージカルノイズが含まれていないとみなせるので除外している．Table 2 より，ミュージカルノイズ品質と音声品質の相関はほぼ認められないことから，音声品質はミュージカルノイズの影響をほとんど受けていないと考えられる．一方，ミュージカルノイズ品質と雑音品質の間にはやや強い相関があることが分かった．

次に，ミュージカルノイズ品質と雑音品質の関係を Fig. 1 に示す．ここで，横軸はミュージカルノイズ品質，縦軸は雑音品質を表す．図中の ○ はミュージカルノイズ品質が 4 未満，\* は 4 以上で

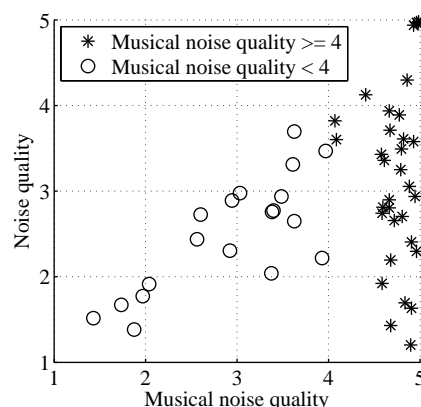


Fig. 1 Relationship between the musical noise quality and the noise quality.

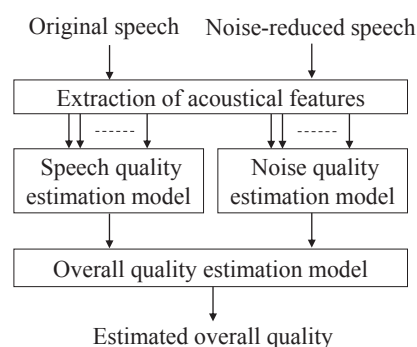


Fig. 2 Overview of the full-reference objective quality evaluation method.

あることを示す．また，個々のマーカは，雑音抑圧アルゴリズム，雑音，SNR の組合せの 1 つに対する MOS (32 名の被験者が 4 個の音声サンプルに付けた評点の平均) を表している．Fig. 1 において，ミュージカルノイズ品質が 4 未満の場合に注目すると，ミュージカルノイズ品質と雑音品質にやや強い相関があること，及び両者の関係にはばらつきがあることが確認できる．これは，被験者がミュージカルノイズを加味した上で雑音品質を評価していること，すなわちミュージカルノイズは雑音品質を決定する要因の 1 つであることを示唆している．

## 3 FR 型客観品質評価法の改良

### 3.1 FR 型客観品質評価法の概要

我々が提案している FR 型客観品質評価法 [3, 4] では，Fig. 2 のような二段階の処理によって雑音抑圧音声の品質評価を行う．まず，音声区間と非音声区間の各々から信号の成分が増えるタイプのひずみと信号の成分が減るタイプのひずみを求める．ここで，ひずみ尺度は耳内音圧スペクト

ルひずみ尺度である．これは，主に符号化音声を対象とする FR 型客観品質評価法である，ITU-T 勧告 P.862[10] に採用されているひずみ尺度である．また，雑音抑圧音声の非音声区間から残留雑音の平均対数パワーを求める．次に，これら 5 種類の特徴量から音声品質と雑音品質を各々推定する．音声品質と雑音品質の各推定式は，上述した特徴量の 1 次結合として定義している．最後に，推定した音声品質と雑音品質を総合品質推定モデルに代入することにより，総合品質を推定する．総合品質推定モデルを次式に示す [4]．

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

ここで， $Q_o$  は総合品質， $Q_s$  は音声品質， $Q_n$  は雑音品質， $c_{s_1} = 0.3582$ ， $c_{s_2} = -0.0696$ ， $c_{n_1} = -0.0751$ ， $c_{n_2} = -0.0271$ ， $c_{sn} = 0.2228$ ， $c = 0.5091$  である．

### 3.2 ミュージカルノイズを考慮した雑音品質推定モデル

2.2 節で述べたように，ミュージカルノイズは雑音品質を決定する要因の 1 つであると考えられる．したがって，FR 型客観品質評価法における雑音品質推定の際に，上述した 5 種類の特徴量に加えてミュージカルノイズ発生量を特徴量として用いることが有効であると期待できる．このことから，雑音品質推定モデルを次式のように定めた．

$$\text{雑音品質} = \sum_{i=1}^6 (\beta_i X_i) + \beta_7 \quad (2)$$

ここで， $X_1 \sim X_5$  は上述した 5 つの特徴量である．また， $X_6$  はミュージカルノイズの発生量に相当する特徴量であり，本稿では Uemura らが提案した対数カートシス比 [11] を採用した．なお， $\beta_1 \sim \beta_7$  は雑音品質の推定誤差を最小にするように決定される．

### 3.3 提案する雑音品質推定モデルの有効性の検証

2.1 節の主観品質評価に用いた音声サンプルから上述した 6 種類の特徴量を抽出し，雑音品質推定モデルにより雑音品質を推定した結果を Fig. 3 に示す．ここで，横軸は真の雑音品質，縦軸は推

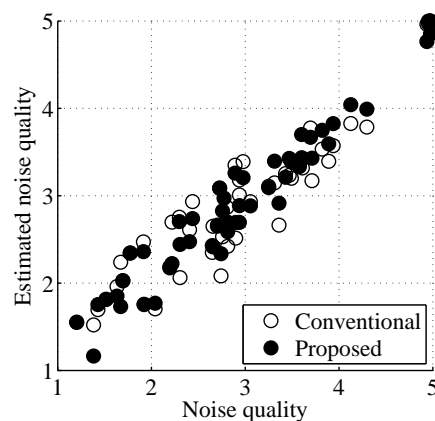


Fig. 3 Relationship between the true noise quality and the estimated noise quality.

定した雑音品質である．図中の  $\bullet$  は提案する雑音品質推定モデルを用いた場合， $\circ$  は従来の雑音品質推定モデルを用いた場合である．個々のマーカは，雑音抑圧アルゴリズム，雑音，SNR の組合せの 1 つに対応する．Fig. 3 より，提案する雑音品質推定モデルにより推定精度が改善していることが見て取れる．決定係数と RMSE は，従来の雑音品質推定モデルについては各々 0.88，0.34，提案する雑音品質推定モデルについては各々 0.94，0.24 であり，提案する雑音品質推定モデルの方が高精度であることが分かる．ここで，RMSE の目標値を主観品質評価によって得られた雑音品質の 95 % 信頼区間片側幅 [12] とするとき，RMSE の目標値は 0.13 となる．提案する雑音品質推定モデルはこれに迫る精度で雑音品質を推定できることが分かった．

### 3.4 提案する雑音品質推定モデルの FR 型客観品質評価法への適用

2.1 節の主観品質評価に用いた音声サンプルから音声品質と雑音品質をそれぞれ推定し，その推定値を総合品質推定モデルである式 (1) に代入することにより，総合品質を推定した結果を Fig. 4 に示す．ここで，横軸は真の総合品質，縦軸は推定した総合品質である．図中の  $\bullet$  は提案する雑音品質推定モデルを用いた場合， $\circ$  は従来の雑音品質推定モデルを用いた場合である．個々のマーカは，雑音抑圧アルゴリズム，雑音，SNR の組合せの 1 つに対応する．決定係数と RMSE は，従来の雑音品質推定モデルを用いた場合については各々 0.94，0.22，提案する雑音品質推定モデルを用いた場合については各々 0.91，0.26 であり，推定精度にさほど違いは見られない．これは，音声品質の推定精度が不十分であるからと考

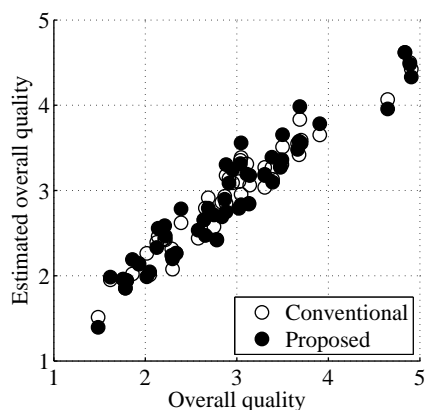


Fig. 4 Relationship between the true overall quality and the estimated overall quality.

えられる [4] .そこで, 推定した音声品質の代わりに主観品質評価で得た真の音声品質を用いて総合品質を推定した. 決定係数とRMSEは, 従来の雑音品質推定モデルを用いた場合には各々0.93, 0.22, 提案する雑音品質推定モデルを用いた場合には各々0.96, 0.16であり, 提案する雑音品質推定モデルの適用により, 推定精度が改善していることが分かる. したがって, さらなる推定精度の改善のためには音声品質の推定精度を高める必要がある.

#### 4 おわりに

本稿では, 我々が提案している雑音抑圧音声のFR型客観品質評価法の推定精度改善について述べた. まず, ミュージカルノイズが音声品質と雑音品質に及ぼす影響を調査し, ミュージカルノイズは音声品質ではなく, 雑音品質に大きな影響を及ぼすことを明らかにした. この事実に基づいて, 提案手法における雑音品質推定モデルの改良を行った. 雑音品質の推定に用いる特徴量に, ミュージカルノイズ発生量に相当するとされる特徴量を追加することにより, 雑音品質をこれまでよりも高い精度で推定できることを示した.

今後は, 音声品質推定モデルの推定精度を改善するための方策を検討する.

謝辞 本研究の一部は電気通信普及財団の助成による.

#### 参考文献

[1] ITU-T Rec. P.835, “Subjective test methodology for evaluating speech com-

munication systems that include noise suppression algorithm,” Nov. 2003.

- [2] T. Yamada *et al.*, “Non-reference objective quality evaluation for noise reduced speech using overall quality estimation model,” IEICE Transactions on Communications, Vol. E93-B, No. 6, pp. 1367-1372, June 2010.
- [3] 篠原佑基 ら, “雑音抑圧音声の総合品質モデルを用いたフルリファレンス客観品質評価法の検討,” 第7回 QoS ワークショップ, QW7-P-13, pp. 40-41, Nov. 2009.
- [4] 藤田悠希 ら, “雑音抑圧音声の総合品質推定モデルの改良とその客観品質評価への適用,” 日本音響学会秋季研究発表会, pp. 127-130, Sep. 2011.
- [5] 藤田悠希 ら, “雑音抑圧音声の主観品質評価におけるミュージカルノイズの影響,” 信学総大 D-14-1, Mar. 2012.
- [6] 電子協騒音データベース, <http://research.nii.ac.jp/src/list/detail.html/#JEIDA-NOISE>.
- [7] 3GPP2 C.S0014-A Version 1.0, “Enhanced variable rate codec, speech service option 3 for wideband spread spectrum digital systems,” Apr. 2004.
- [8] 古田訓 ら, “スペクトル減算と振幅抑圧の相互制御に基づく雑音抑圧法の検討,” 電子情報通信学会論文誌, Vol. J87-D-II, No. 2, pp. 464-474, Feb. 2004.
- [9] 池原雅章 ら, “MATLAB マルチメディア信号処理 下,” 培風館, 2004.
- [10] ITU-T Rec. P.862, “Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs,” Feb. 2001.
- [11] Y. Uemura *et al.*, “Automatic optimization scheme of spectral subtraction based on musical noise assessment via higher-order statistics,” Proc. of International Workshop on acoustic Echo and Noise Control, 2008.
- [12] 高橋玲, 北脇信彦, “符号化音声品質客観評価尺度の性能評価,” 電子情報通信学会論文誌, Vol. J80-B-I, No. 6, pp. 480-487, June 1997.