

非音声区間のスペクトル分布を利用した音声強調法

川村 新[†] タンヒカムウィーラウト[†] 飯國 洋二[†]

[†] 大阪大学大学院基礎工学研究科 〒560-8531 大阪府豊中市待兼山町1-3

E-mail: [†]{kawamura,iiguni}@sys.es.osaka-u.ac.jp, ^{††}weerawut@sip.sys.es.osaka-u.ac.jp

あらまし 雑音が重畳する音声に対して、音声スペクトルだけを推定し抽出する方法が多数提案されている。最近の有用な音声スペクトル推定法として、音声スペクトルおよび雑音スペクトルそれぞれの出現確率を表す分布関数を仮定し、MAP推定により音声スペクトルを求める方法がある。この方法では、音声スペクトル分布をあるひとつの関数で近似する。しかし、非音声区間における音声スペクトルの出現確率を表す分布は明らかにデルタ関数となることから、音声が存在する場合と存在しない場合とでは、音声スペクトルの出現確率がそれぞれ異なると考えられる。そこで本論文では、音声スペクトル分布を可変としたMAP推定に基づく音声強調法を提案する。提案法では、まず、音声の振幅スペクトルが小さくなる場合にデルタ関数、大きくなる場合にレイリー分布に近づくような分布関数を定める。そして、定めた分布関数を用いてMAP推定を実行し、雑音重畳音声から音声スペクトルを抽出する。音声強調シミュレーションの結果から、特に非音声区間において、提案法の雑音除去性能が従来法よりも優れていることを明らかにする。

キーワード 音声強調, MAP推定, 確率密度関数, 非音声区間

A speech enhancement method using spectral probability density function of non-speech segments

Arata KAWAMURA[†], Thanhikam WEERAWUT[†], and Youji IIGUNI[†]

[†] Department of Systems Innovation, Graduate School of Engineering Science, Osaka University
Machikaneyama 1-3, Toyonaka, Osaka, 560-8531 Japan

E-mail: [†]{kawamura,iiguni}@sys.es.osaka-u.ac.jp, ^{††}weerawut@sip.sys.es.osaka-u.ac.jp

Abstract In this paper, we propose a speech enhancement algorithm based on the MAP estimation using variable speech spectral distribution. To extract high quality speech signals, the proposed speech enhancer adaptively changes the speech spectral density used in the MAP estimation according to the observed spectral power. In speech segments, the speech spectral density approaches to Rayleigh distribution to keep the quality of the enhanced speech. In non-speech segments, it approaches to a delta function to reduce noise effectively. The proposed technique is effective in suppressing residual noise well. From computer experiments, we confirm the effectiveness of the proposed method.

Key words speech enhancement, MAP estimation, probability density function, non-speech segment

1. はじめに

携帯電話や音声認識装置の急速な普及により、これらの機器が様々な環境下で使用されている。しかし、周囲に雑音が存在する場合には、通話品質の劣化や、音声認識率の低下が生じるため、機器を快適に使用できないという問題がある。このような背景から、音声強調技術に関する研究が盛んに行われている。音声強調技術とは、観測信号として雑音重畳音声を与えられた場合に、観測信号から雑音を除去し、音声信号だけを抽出する

技術である。携帯電話に応用すれば図1に示すように円滑なコミュニケーションが実現できる。これまでに提案されている音声強調技術は、時間領域の処理と、周波数領域の処理とに大別できる。いずれの方法においても多数の研究結果が報告されているが、最近の携帯電話に搭載されている音声強調手法の多くは周波数領域の処理である。そこで、本研究においても特に周波数領域の音声強調技術に着目する。

代表的な周波数領域の音声強調法として、スペクトルサブトラクション法[1]がある。スペクトルサブトラクション法は、観

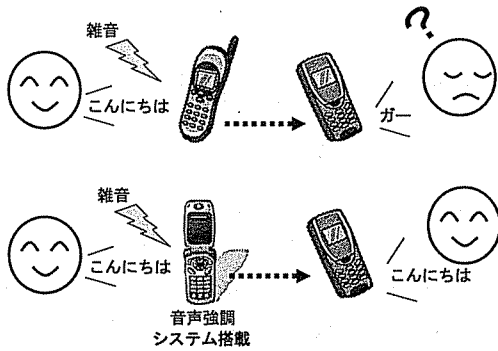


図1 音声強調技術の応用例

測信号の振幅スペクトルから雑音の振幅スペクトルを減算するという単純な方法である。しかし、スペクトルサブトラクション法には、ミュージカルノイズと呼ばれる人工的で耳障りな雑音が生じる問題が指摘されていた。その後、ミュージカルノイズが生じにくい方法として、MMSE-STSA(minimum mean square error-short time spectral amplitude)法 [2] や、MAP(maximum a posteriori) 推定に基づく方法 [3] が提案された。これらの方法は、雑音スペクトルと音声スペクトルを確率変数としてとらえ、それらが従う確率密度関数を仮定し、推定音声を求める方法である。一般に、雑音スペクトルはガウス分布、音声の位相スペクトルは一様分布に従うと仮定することが多い。一方、音声の振幅スペクトルが従う分布には様々なものが提案されており、当然ながら定めた分布関数ごとに異なる音声推定法が提案されている [3]~[9]。しかし、いずれの方法においても低 SNR 環境下では、ミュージカルノイズの発生を抑えることができず、より高性能な音声強調法の確立が望まれている。

さて、多くの文献では、音声が存在する音声区間と、存在しない非音声区間を区別せず、音声の確率密度関数があるひとつの分布関数で表現している。しかし、通常の会話音声では、非音声区間が全体の 2/3 程度存在しており [10]、電話で相手側が発話している場合には、さらに長時間の非音声区間が生じる。したがって、音声区間と、長時間存在する非音声区間とを同じ分布関数で表現することは現実的ではない。そこで我々は、音声区間と非音声区間とを異なる分布関数で表現し、それぞれの区間において MAP 推定による音声強調を行う。具体的には、文献 [9] で提案されたパラメトリック関数を利用し、音声区間と非音声区間で異なるパラメータの組を用いて可変分布関数を実現する。シミュレーションの結果から、低 SNR 環境下において、提案法が従来法よりも効果的にミュージカルノイズを抑圧できることを明らかにする。

2. 従来の周波数領域における音声強調法

本節では、従来の周波数領域における音声強調法について述べる。

2.1 音声強調システムの基本構成

観測信号が次式で与えられるとする。

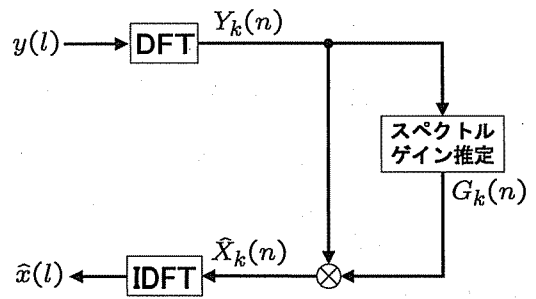


図2 音声強調システム

$$y(l) = x(l) + d(l) \quad (1)$$

ここで、 $y(l)$ は時刻 l における観測信号であり、 $x(l)$ と $d(l)$ はそれぞれ音声と雑音を表す。ただし、音声と雑音は独立であると仮定する。式 (1) の両辺に対して DFT を実行すると、第 n 分析フレームにおける k 番目のスペクトルは、次式で与えられる。

$$Y_k(n) = X_k(n) + D_k(n) \quad (2)$$

このとき、音声強調スペクトル $\hat{X}_k(n)$ は、

$$\hat{X}_k(n) = G_k(n)Y_k(n) \quad (3)$$

として表現できる。すなわち、観測信号スペクトル $Y_k(n)$ に適当なスペクトルゲイン $G_k(n)$ を乗じたものが、音声強調スペクトルとなる。式 (3) で与えられる音声強調システムを図 2 に示す。種々の周波数領域における音声強調法は、すべて図 2 の形で統一的に表現でき、スペクトルゲイン $G_k(n)$ の決定方法のみが異なる。ここで、式 (2), (3) からわかるように、理想的なスペクトルゲインは、

$$G_{k,opt}(n) = 1 - \frac{D_k(n)}{Y_k(n)} \quad (4)$$

である。このとき、観測信号から完全に雑音を取り除き、元の音声だけを抽出することができる。しかしながら、 $Y_k(n)$ の情報だけから $D_k(n)$ を正確に知る方法が明らかにされていないため、なんらかの方法で近似的にスペクトルゲイン $G_k(n)$ を決定しなければならない。

以下では、従来の音声強調法によって導出される $G_k(n)$ について説明する。ただし、以降では表記を簡単にするため、必要でない場合にはフレーム番号 n とスペクトル番号 k を省略する。

2.2 MMSE-STSA 法

最小化すべき評価関数として、次の条件付期待値を考える。

$$J = E \left[|X - \hat{X}|^2 \middle| Y \right] \quad (5)$$

これは、観測信号スペクトル Y が生じたときの、音声信号スペクトル X とその推定値 \hat{X} との平均 2 乗誤差を表している。問題は、式 (5) を最小化する音声スペクトル \hat{X} をどのように

選ぶかである。

MMSE-STSA 法 [2] では、 \hat{X} として音声スペクトルの平均値、すなわち X の条件付期待値 $E[X|Y]$ を選択する。 $E[X|Y]$ は、事後確率密度関数 $p(X|Y)$ から計算することができる [11]。さらに、 $p(X|Y)$ は、音声と雑音スペクトルそれぞれの確率密度関数から得られる。 MMSE-STSA 法では、音声の振幅スペクトル $|X|$ がレイリー分布 [12]、位相スペクトル $\angle X$ が一様分布、さらに雑音の複素スペクトルがガウス分布に従うと仮定する。このとき、条件付期待値 $E[X|Y]$ から得られるスペクトルゲインは、

$$G_{\text{MMSE}} = \frac{(\pi v)^{1/2}}{2\gamma} \exp\left(\frac{-v}{2}\right) \times \left[(1+v)I_0\left(\frac{v}{2}\right) + vI_1\left(\frac{v}{2}\right) \right] \quad (6)$$

となる [2]。ただし、 $I_0(\cdot)$ 、 $I_1(\cdot)$ は 0 次と 1 次の変形ベッセル関数を表す。また、

$$v = \frac{\xi}{1+\xi}\gamma, \quad \xi = \frac{\sigma_n^2}{\sigma_d^2}, \quad \gamma = \frac{|Y|^2}{\sigma_d^2} \quad (7)$$

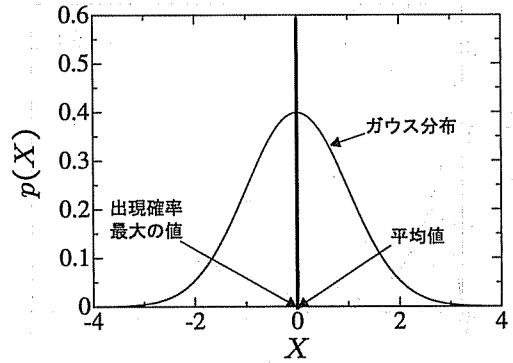
であり、 σ_d^2 、 σ_n^2 は雑音スペクトルおよび音声スペクトルの分散である。ここで特に、 ξ を事前 SNR、 γ を事後 SNR と呼ぶ。 MMSE-STSA 法は、スペクトルサブトラクション法 [1] よりもミュージカルノイズを抑制できることが報告されている [2]。しかし、式 (6) のスペクトルゲインを得るためには、特殊関数である \exp 関数、変形ベッセル関数を計算しなければならず、演算量が大きくなるという問題がある。

2.3 MAP 推定法

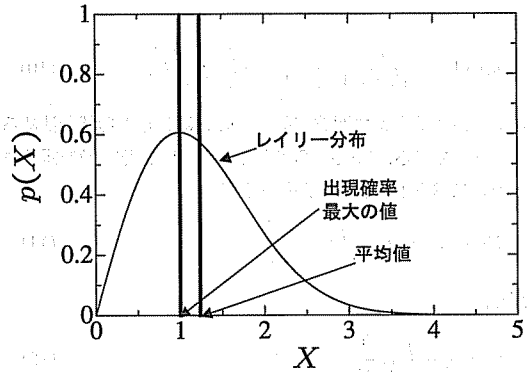
MMSE-STSA 法では、評価関数 J の最小化を目的とし、その最適値として条件付期待値 $E[X|Y]$ を選択した。これは、ガウス分布のような対称の確率密度関数を考える場合には最適な選択といえる。しかし、レイリー分布のような非対称の分布を扱う場合には、条件付期待値が最適な選択になるとは限らない。そこで、MAP 推定法 [3] では、評価関数 J を最小化するために、事後確率密度関数 $p(X|Y)$ を最大にする X を選択する。

ここで MMSE-STSA 法と MAP 推定法の違いを明確にするため、具体的な確率密度関数を例にして説明する。図 3(a)、(b) にガウス分布およびレイリー分布を示す。同図において、横軸は確率変数 X 、縦軸は確率密度関数 $p(X)$ である。また、 X の平均値と、最も出現確率の高い X の値それぞれに垂直線を引いている。前者が MMSE-STSA 法による X の推定値、後者が MAP 推定法による推定値に対応する。図 3(a) に示すガウス分布は対称な関数なので、MMSE-STSA 法と MAP 推定法の推定値は一致する。これに対して、図 3(b) に示すレイリー分布は非対称であるため、両者の推定値は異なる。明らかに、非対称の確率密度関数を考える場合には、MMSE-STSA 法で得られる平均値よりも、出現確率が最大となる MAP 推定値を採用する方が、評価関数 J の最小化に適している。

さて、実際に MAP 推定によりスペクトルゲインを得るために、事後確率密度関数 $p(X|Y)$ の最大化を考える。ベイズの定理 [13] より、



(a) ガウス分布の例



(b) レイリー分布の例

図 3 分布関数の最大値と平均値

$$p(X|Y) = \frac{p(Y|X)p(X)}{p(Y)} \propto p(Y|X)p(X) \quad (8)$$

であるから、右辺最後の式を最大化する $|X|$ および $\angle X$ を求めればよい。具体的な手順は、式 (8) の右辺最後の式の対数を取り、 $|X|$ 、または $\angle X$ で微分し、それらの結果を 0 とおく。そして得られた方程式を、 $|X|$ 、 $\angle X$ についてそれぞれ解けばよい。ここで、MMSE-STSA 法と同様のスペクトル分布関数を仮定すると、次のスペクトルゲインが得られる [3]。

$$G_{\text{MAP}} = \frac{\xi + \sqrt{\xi^2 + 2(1+\xi)(\xi/\gamma)}}{2(1+\xi)} \quad (9)$$

MAP 推定においては、位相スペクトルの最適値は、観測信号の位相スペクトルと一致するため、 G_{MAP} は実数となる。また、式 (9) は、 \exp 関数も変形ベッセル関数も含んでおらず、式 (6) よりも容易にスペクトルゲインを計算することができる。

ところが、T. Lotter と P. Vary は、実際の長時間音声を観察することで、その振幅スペクトルがレイリー分布に従わないことを指摘し、独自の音声スペクトル分布を提案した [9]。提案された音声スペクトル分布は、位相スペクトルが一様分布、振幅スペクトルが次式に従うとするものである。

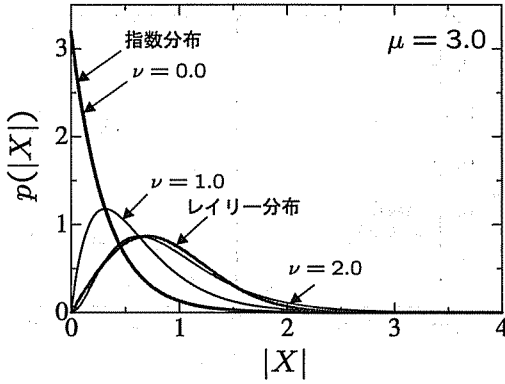


図4 ν の値による分布の変化

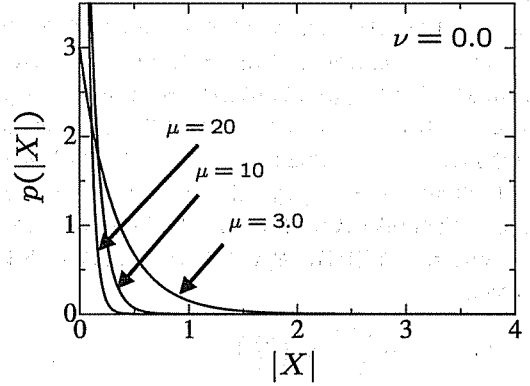


図5 μ の値による分布の変化

$$p(|X|) = \frac{\mu^{\nu+1}}{\Gamma(\nu+1)} \frac{|X|^\nu}{\sigma_x^{\nu+1}} \exp\left(-\mu \frac{|X|}{\sigma_x}\right) \quad (10)$$

ここで、 $\Gamma(\cdot)$ はガンマ関数であり、 μ, ν は分布の形状を決めるパラメータである。Lotter らの分布から MAP 推定値を求めれば、次のスペクトルゲインが得られる。

$$G_{L,MAP} = u + \sqrt{u^2 + \frac{\nu}{2\gamma}} \quad (11)$$

ただし、

$$u = \frac{1}{2} - \frac{\mu}{4} \sqrt{\frac{1}{\gamma\xi}} \quad (12)$$

である。ここで、文献[9]において、 $\mu = 1.74, \nu = 0.126$ が実際の音声スペクトル分布を最も良く近似する値であることが報告されている。また、本方式においても、位相スペクトルの MAP 推定値は、観測信号の位相スペクトルに一致する。

3. 可変音声分布に基づく音声強調法

Lotter らによって提案された分布関数は、パラメータ μ と ν を固定値としているため、その分布の形状は常に一定である。しかし、実際の音声には、長時間にわたる休止区間や、音声の始まりや終わり付近に見られる小さいレベルの音声成分が多く含まれている。そして、これらの区間においては、音声振幅スペクトルの分布は、0 だけが出現するデルタ関数か、または指数分布に近い形となるはずである。一方、音声が存在する区間だけに注目すると、音声振幅スペクトルの分布は、レイリー分布で近似できることが確認されている。これらの事実から、我々は以前、式(10)を利用して音声スペクトル分布の形状を適応的に変化させる方法を提案した[14]。

図4に示すように、式(10)は、 $\nu = 0.0$ のときに指数分布、 $\nu = 2.0$ のときにレイリー分布を近似できる。ここで、 $\mu = 3.0$ と固定している。文献[14]では、非音声区間では $\nu \rightarrow 0.0$ 、音声が存在する区間では $\nu \rightarrow 2.0$ となるように、次式を用いて可変音声分布を実現した。

$$G_{T,MAP} = u + \sqrt{u^2 + \frac{\bar{\nu}}{2\gamma}} \quad (13)$$

$$\bar{\nu} = \begin{cases} 2.0, & \hat{\nu} > 2.0 \\ \hat{\nu}, & 0.0 \leq \hat{\nu} \leq 2.0 \\ 0.0, & \hat{\nu} < 0.0 \end{cases} \quad (14)$$

$$\hat{\nu} = 0.05 \cdot 10 \log_{10} \frac{\sum_{k=0}^{M-1} |Y_k|^2}{\sum_{k=0}^{M-1} \sigma_{d,k}^2} \quad (15)$$

ここで、式(15)右辺の M は DFT 点数、添え字 k はスペクトル番号である。また、定数 0.05 は分布形状を調整するために導入されている。式(14)、(15)からわかるように、 $\bar{\nu}$ は、観測信号の事後 SNR を参照し、現在の音声の大きさに応じて適応的に変化する。しかし、この方法では式(10)に含まれる2つのパラメータ ν, μ のうち、 ν だけを可変としたため、非音声区間における音声分布を、本来のデルタ関数ではなく、立下りのゆるやかな指数分布でしか表現できないという問題がある。

そこで本論文では、もう一方のパラメータ μ についても適応的に変化させる方法を提案する。 μ は式(10)において指数関数のべきを定めるパラメータであり、値が大きいほど分布の急峻な立下りを実現できる。したがって、非音声区間ではデルタ関数を近似するために μ を大きくして立下りを急峻にし、逆に音声区間では μ を小さくして立下りのゆるやかなレイリー分布を近似する。図5に $\nu = 0$ と固定して μ の値を変化させた場合の分布の変化を示す。同図からわかるように、 $\nu = 0$ であっても μ の値を大きくすることでより急峻な分布関数が得られることがわかる。

可変音声分布を用いてスペクトルゲインを得るための提案アルゴリズムを次式に示す。

$$G_{P,MAP} = \bar{u} + \sqrt{\bar{u}^2 + \frac{\bar{\nu}}{2\gamma}} \quad (16)$$

$$\bar{u} = \frac{1}{2} - \frac{\bar{\mu}}{4} \sqrt{\frac{1}{\gamma\xi}} \quad (17)$$

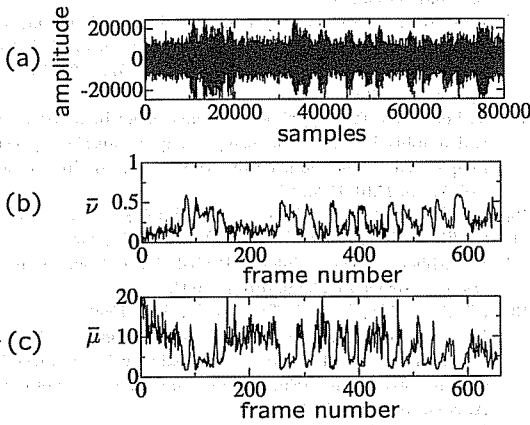


図6 観測信号に対するパラメータの変化 (a) 観測信号 (b) \bar{v} の変動 (c) $\bar{\mu}$ の変動

$$\bar{\mu} = \begin{cases} 20, & \hat{\mu} > 20 \\ \hat{\mu}, & 1.74 \leq \hat{\mu} \leq 20 \\ 1.74, & \hat{\mu} < 1.74 \end{cases} \quad (18)$$

$$\hat{\mu} = 20 \frac{\sum_{k=0}^{M-1} \sigma_{d,k}^2}{\sum_{k=0}^{M-1} |Y_k|^2} \quad (19)$$

ここで、式(19)右辺の定数20は分布関数を調整するために導入した。また、 $\bar{\mu}$ の下限はLotterらが文献[9]で提示した μ の値と同一とし、上限については実験的に定めた。式(18)、(19)からわかるように、 $\bar{\mu}$ は事後SNRに反比例して変化させた。

4. シミュレーション

本節では、提案した音声強調法の有効性を確認するため、シミュレーション実験を行う。比較のため、Lotterらの方法[9]と v だけを変変とした文献[14]の方法についても同様の実験を行う。観測信号として、高速道路トンネル内で収録された実環境雑音を、SNR=0dBとなるように、女声に付加したものをを用いた。ただし、サンプリング周波数8kHz、DFT点数を $M=256$ とした。また、雑音分散 σ_d^2 の推定は、携帯電話にも搭載され、有用な方式として注目されている、重み付き雑音推定法[15]を利用した。さらに、第 n フレームにおける k 番目のスペクトルの事前SNRは、次式のdecision-directed法[2]により求めた。

$$\xi_k(n) = \beta \gamma_k(n-1) G_k^2(n-1) + (1-\beta) \cdot \max[\gamma_k(n) - 1, 0] \quad (20)$$

ただし、 $\max[\cdot]$ は最大値を選択する演算子であり、 β は忘却係数である。実験では $\beta=0.98$ と設定した。

観測信号に対する可変パラメータ \bar{v} 、 $\bar{\mu}$ の動きを図6(a)-(c)に示す。ここで、同図(a)は観測信号、(b)は \bar{v} の変動、(c)は $\bar{\mu}$ の変動である。結果から、観測信号のSNRに比例して \bar{v} が、反比例して $\bar{\mu}$ が適応的に変化していることが確認できる。次に、音声強調結果の時間波形を図7(a)-(d)に示す。ここで、図7(a)は観測信号、(b)はLotterらのMAP推定法[9]による結

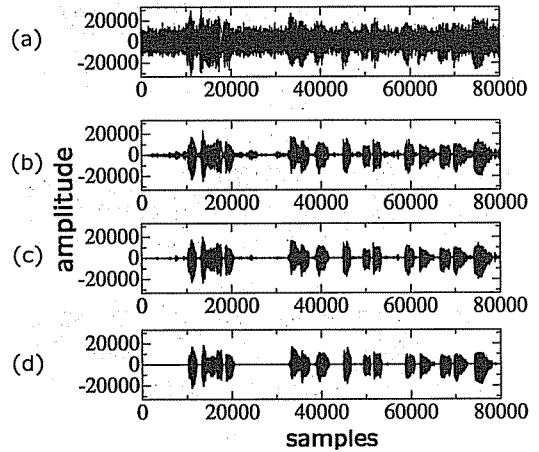


図7 シミュレーション結果 (a) 観測信号 (b) Lotterらの方法 (c) 文献[14]の方法 (d) 提案法

果、(c)は文献[14]の方法による結果、そして、(d)は提案法による結果である。同図から、いずれも優れた音声強調結果が得られていることがわかる。また、最も雑音除去性能が高いものは、提案法による結果であることが確認できる。これらの結果を試聴すると、図7(b)、(c)の非音声区間における残留雑音が耳障りなミュージカルノイズであることがわかる。このことを視覚的に確認するために、それぞれのスペクトログラムを図8(a)-(d)に示す。ここで、横軸は時間、縦軸は周波数を表す。図8(b)、(c)から、特に非音声区間において、残留雑音がスペクトルの孤立点として多数現れていることがわかる。このようにスペクトログラム上で孤立点として不規則に現れる雑音が聴感上不快なミュージカルノイズとなる。これに対して提案法では、図8(d)からわかるように、雑音が十分除去されており、試聴してもほとんどミュージカルノイズが知覚されなかった。しかしながら、音声区間においては多少の音質劣化が認められた。

提案法では、音声区間で分布関数がレイリー分布を近似すること、すなわち、 $\bar{v} \approx 2.0$ 、 $\bar{\mu} \approx 3.0$ を得ることが目的のひとつである。ところが、図6(b)に示した \bar{v} は、音声区間においても0.5付近にとどまっている。一方、今回提案した $\bar{\mu}$ は、音声区間において3.0付近の値をとっている。したがって、音質劣化の主な原因は、 \bar{v} の設定方法にあると考えられる。今後は、両パラメータの相互作用も考慮し、 \bar{v} のより適切な設定方法について検討する必要がある。

5. おわりに

本論文では、音声区間と非音声区間において、音声スペクトルが異なる分布関数に従うことを仮定し、音声スペクトル分布を変変としたMAP推定に基づく音声強調法を提案した。提案法は、音声の振幅スペクトルが小さくなる場合にはデルタ関数に、音声の振幅スペクトルが大きくなる場合にはレイリー分布に近づくような分布関数を導入することで、高性能な音声強調を実現している。音声強調シミュレーションの結果から、0dB

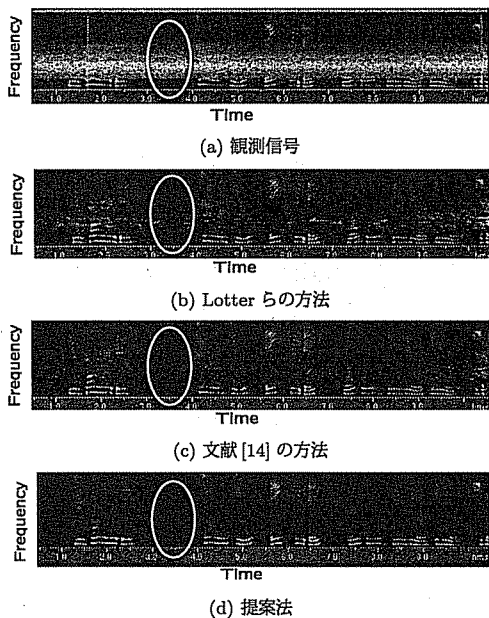


図 8 シミュレーション結果のスペクトログラム

という低 SNR 環境下においても、提案法が優れた音声強調性能を示すことが確認できた。また、特に非音声区間においては、従来法では抑圧できなかったミュージカルノイズが、提案法では知覚できない程度まで抑圧できることが明らかとなった。しかしながら、音声区間においては音質劣化が生じているため、今後は、両パラメータの相互作用を考慮して、より現実の音声に近い分布を実現するパラメータ設定法について検討する必要がある。

文 献

- [1] S.F. Boll, "Suppression of acoustic noise in speech using spectral subtraction," *IEEE Trans. Acoustics, Speech, and Signal Processing*, Vol. ASSP-27, No. 2, pp. 113-120, 1979.
- [2] Y. Ephraim, and D. Malah, "Speech enhancement using a minimum mean square error short-time spectral amplitude estimator," *IEEE Trans. Acoustics, Speech, and Signal Processing*, Vol. ASSP-32, No. 6, pp. 1109-1121, 1984.
- [3] P.J. Wolf and S.J. Godsill, "Efficient alternatives to the Ephraim and Malah suppression rule for audio signal enhancement," *EURASIP Journal on Applied Signal Processing*, Vol. 10, pp. 1043-1051, 2003.
- [4] B. Widrow, J.G.R. Glover, Jr., J.M. Mccool, J. Kaunitz, C.S. Williams, R. H. Hearin, J.R. Zeidler, E. Dong, Jr., and R.C. Goodlin, "Adaptive noise cancelling: Principles and applications," *Proceedings of The IEEE*, Vol. 63, No. 12, pp. 1692-1719, 1975.
- [5] S. Gazor and W. Zhang, "Speech enhancement employing Laplacian-Gaussian mixture," *IEEE Trans. Speech and Audio Processing*, Vol. 13, No. 5, pp. 896-904, 2005.
- [6] R.J. McAulay and M.L. Malpass, "Speech enhancement using a soft-decision noise suppression filter," *IEEE Trans. Acoustics, Speech, and Signal Processing*, Vol. ASSP-28, No. 2, pp. 137-145, 1980.
- [7] B. Chen and P.C. Loizou, "Speech enhancement using a MMSE short time spectral amplitude estimator with Laplacian speech modeling," *IEEE ICASSP2005*, Vol. 1,

- pp. 1097-1100, 2005.
- [8] R. Martin, "Speech enhancement based on minimum mean-square error estimation and super-Gaussian priors," *IEEE Trans. Speech and Audio Processing*, Vol. 13, No. 5, pp. 845-856, 2005.
- [9] T. Lotter and P. Vary, "Speech enhancement by MAP spectral amplitude estimation using a super-Gaussian speech model," *EURASIP Journal on Applied Signal Processing*, Vol. 7, pp. 1110-1126, 2005.
- [10] 古井貞照, デジタル音声処理, 東海大学出版社, 東京, 1985.
- [11] 飯國洋二, 基礎から学ぶ信号処理, 培風館, 東京, 2004.
- [12] S.L. Miller and D.G. Childers, *Probability and random processes*, Elsevier Academic Press, 2004.
- [13] 渡部洋, 基礎から学ぶ信号処理, 培風館, 東京, 1999.
- [14] Y. Tsukamoto, A. Kawamura, and Y. Iiguni, "Speech enhancement based on MAP estimation using a variable speech distribution," *IEICE Trans. Fundamentals*, Vol.E90-A, No.8, pp.1587-1593, 2007.
- [15] 加藤正徳, 杉山昭彦, 芹沢昌宏, "重み付き雑音推定と MMSE STSA 法に基づく高音質雑音抑圧," *電子情報通信学会論文誌 (A)*, Vol. J87-A, No. 7, pp. 851-860, 2004.