

# 周波数領域における超解像を導入したスペクトルサブレーション法

学籍番号：90145088 飯國研究室 鈴木 智之

## 1 はじめに

音声に重畳した雑音を除去する簡単で有効な手法としてスペクトルサブレーション法が知られている．しかしこの手法ではミュージカルノイズと呼ばれる人工的なノイズの発生や、音質の劣化などの問題が生じる．周波数領域における雑音除去法はスペクトル解像度が高いほど、少ない周波数に音声のエネルギーを集中させることができるために、高い雑音除去性能を達成できることが知られている [1]．

そこで本論文は、スペクトルサブレーション法の前処理として、周波数領域における超解像を導入することで雑音除去性能を向上させる手法を提案する．

## 2 周波数領域における超解像

$N$  点のフレームのスペクトル解像度を整数  $a$  倍に向上させるには、フレーム長を  $a$  倍に拡張する必要がある．そのためにフレームに  $(a-1)N$  個の点を与えることになるが、数列は自由に決めることができ、組み合わせは無数に存在する．そこでここでは高解像度化のための条件を設けることによってこの区間を一意に定めることを考える．

本問題で自由度があるのは数列を与える部分のみであり、中央の既知である  $N$  点は変わらない．すなわち  $m$  番目のフレームの時間  $n$  における与えられた信号  $x(m, n)$  に関して、拡張後の信号  $\hat{x}(m, n)$  は

$$\hat{x}(m, n) = x(m, n) \text{ for } (a-1)N/2 \leq n < (a+1)N/2 \quad (1)$$

を満たすことを条件とする．この条件の下で音声のエネルギーを少ない周波数に集中させるためのコスト関数を与え、最小化する．音声エネルギーを少ない周波数に集中させるためのコスト関数として、フーリエスペクトルの絶対値の総和を用いる．以上から、この問題を最適化問題に定式化すると、

$$\min \|x\|_1 \quad \text{subject to : } s = MAx \quad (2)$$

が得られる．ただし、 $x$  は信号の FFT 係数ベクトルを表し、 $\|x\|_1$  は複素数ベクトル  $x$  の要素の絶対値の総和を表す．また、 $A$  は逆 FFT 行列であり、 $s$  は中央の  $N$  点が入力信号となっている任意のベクトルである． $M$  はベクトルの中央の  $N$  点を取り出すための行列を表す．

音声の応用では、[2] より式 (2) を無拘束問題

$$\min : J = \frac{1}{2} \|A^T M(Ax - s)\|_2^2 + \lambda \|x\|_1 \quad (3)$$

に緩和する．ただし、 $\|\cdot\|_2$  は複素数ベクトルの要素の絶対値の 2 乗和の平方根をとったものを表す．また、 $A^T$  は順 FFT 行列であり、コスト関数を周波数領域に統一するために形式上導入した．

この処理を MAP 推定を用いた雑音除去法の前処理に導入する．

## 3 実験

男声、女声をそれぞれ 2 種類ずつ用意し、3 種類の雑音、2 種類の異なる入力セグメンタル SNR に対する雑音抑圧結果として、雑音スペクトルが既知の場合を表 1、未知の場合を表 2 に示す．表の (a) は雑音除去結果を表し、(b) は処理をクリーン音声に適用し、音質の劣化の程度を比較した結果である．また、結果は男声、女声で得られたセグメンタル SNR の平均値を示している．なお、比較のため、超解像を導入しない従来法の結果も同時に示す．

表 1: 雑音スペクトルが既知の場合の比較

(a) 雑音除去結果						
雑音の種類	white		babble		f16	
	入力 SNR	0	5	0	5	0
ゼロパディング	6.28	10.65	5.84	10.24	5.96	10.40
超解像	6.78	11.37	6.29	10.96	6.37	11.10
従来法	5.19	9.91	5.15	9.65	4.92	9.63

(b) 音質劣化推定結果						
雑音の種類	white		babble		f16	
	入力 SNR	0	5	0	5	0
ゼロパディング	22.40	23.67	21.89	23.13	21.82	23.10
超解像	22.61	23.12	21.97	22.64	22.01	22.61
従来法	8.92	12.40	8.95	13.19	9.12	13.33

表 2: 雑音スペクトルが未知の場合の比較

(a) 雑音除去結果						
雑音の種類	white		babble		f16	
	入力 SNR	0	5	0	5	0
ゼロパディング	5.31	9.86	2.15	7.22	3.95	9.03
超解像	5.40	9.94	2.25	7.41	4.10	9.31
従来法	4.78	9.45	1.90	7.13	3.75	8.93

(b) 音質劣化推定結果						
雑音の種類	white		babble		f16	
	入力 SNR	0	5	0	5	0
ゼロパディング	8.76	12.54	8.73	13.80	8.84	13.76
超解像	8.47	12.03	8.59	13.31	8.53	13.17
従来法	8.28	11.71	8.48	13.06	8.70	12.99

表から、提案法は雑音除去性能、音質劣化のいずれの観点からも従来法より優れていることがわかる．特に雑音スペクトルが未知の場合における音質の劣化を比較すると、従来法に比べて提案法は 10[dB] 以上優れており、音質の観点からも優れた手法であることが確認された．

## 4 まとめ

本論文では雑音除去法の前処理に周波数領域における超解像を導入することを提案した．従来のスペクトルサブレーション法と比較すると、提案法を導入することにより性能が向上することが確認された．

## 参考文献

- [1] I.Kauppinen and K.Roth, "Improved noise reduction in audio signals using spectral resolution enhancement with time-domain signal extrapolation," IEEE Trans. on Speech and Audio Processing, vol.13, no.6, pp.1210-1216, Nov.2005.
- [2] S.S.Chen, D.L.Donoho and M.A.Saunders, "Atomic decomposition by basis pursuit", SIAM Journal on Scientific Computing, vol.20, no.1, pp.33-61, 1998.