

補助関数法を用いたIダイバージェンス規準残響抑圧*

安良岡直希[†], 亀岡弘和[‡], 吉岡拓也[‡], 奥乃博[†][†] 京都大学 大学院情報学研究科 知能情報学専攻[‡] 日本電信電話(株)NTTコミュニケーション科学基礎研究所

1 はじめに

残響は音楽の演奏や鑑賞の場面では楽曲の印象を豊かにすべく積極的に活用されている半面, 楽曲解析のための音響信号処理の場面では弊害となることが多い. 既存楽曲から残響の影響を自動的に取り除く事ができれば, 楽曲検索などのための自動採譜や音源分離 [1] の精度向上が期待される. また, 既存楽曲の残響効果を自由に操作できるようになれば個人の嗜好に合わせた音楽鑑賞システム [2] へと応用できる.

一般に音楽信号は音声信号に比べ残響抑圧が困難な傾向にある. 主な理由は, 残響成分が比較的強い, 楽器個体の共鳴音は残響と区別することが非常に難しい, などである. 実際, 音声を対象とした既存の残響抑圧法 [3] をそのまま音楽信号に適用すると, 残響が十分に抑圧されない曲や楽器の共鳴音まで過剰に抑圧される曲がある. 最近では, 音楽信号の特徴である調波構造をうまく表現するスペクトルモデルを用いた手法 [4] も報告されているが, 残響の推定精度は依然として十分とは言えず, ピアノなどの撥弦楽器を含む楽曲において楽器音がスタッカート気味になるといった問題が残されている.

既存の残響抑圧法 [3, 4] が過剰な抑圧結果を導く原因について, 残響フィルタと音源 (非残響音を指す) スペクトルの同時推定に用いる最適化基準に問題があると我々は考える. 従来の最適化規準のもとでは, i) 残響フィルタは抑圧後信号の大きさをできるだけゼロに近づけるように推定される, ii) 音源スペクトルモデルのパラメータ推定が困難である, という不都合がある. その結果, 楽器の共鳴音や減衰音など残響に似た特徴を持つ成分は過剰に抑圧される, [4] では残響フィルタと音源スペクトルの推定にそれぞれ別の最適化規準を用いて問題 ii) に対応したが, この方法はアルゴリズムの収束性が保証されない.

本稿では, Iダイバージェンスと呼ばれる最適化規準を採用した新しい残響抑圧法を報告する. この規準は上記問題 i) を回避できる上に, 音楽音響信号の音源分離などの問題において良い性能を示すことが知られている [5]. また, 調波 GMM [6, 7] などの楽器音に適した音源モデルはIダイバージェンス規準での効率的なパラメータ推定アルゴリズムが存在するので, 上記問題 ii) も回避できる. 一方, 残響フィルタをIダイバージェンス規準で最適化するアルゴリズムは知られていないので, 本稿ではこれを新たに設計する. 本稿で述べるアルゴリズムは補助関数法 [8] に基づいて導出されるため, 収束性が保証される.

2 問題定義と手法の比較

本章では本稿で扱う残響抑圧問題を定義し, Iダイバージェンス規準残響抑圧法を既存法と比較する.

2.1 問題定義

本稿では, 残響に関する正確な事前情報を持たない条件下で, 残響が付加された入力信号から音源成分を推定する問題を扱う. 本稿では時間周波数領域での処理を考える. 入力信号と推定すべき未知の音源信号の時間周波数成分それぞれを $x_{n,l}, s_{n,l}$ と表記する. ここで n は時間フレーム, l は周波数ビンに対応するインデックスである. 本稿では入力信号はモノラルと仮定する. またパッチ処理を想定している.

本手法及び [3, 4] では逆畳み込みにより残響抑圧の過程をモデル化する. すなわち, 入力信号は音源信号によって駆動される自己回帰システムによって生成されると見なす.

$$x_{n,l} = \sum_{d=k}^K g_{d,l} x_{n-d,l} + s_{n,l}. \quad (1)$$

ここで, k 及び K はそれぞれ初期反射の時間遅れ及びフィルタ長と対応する. 我々の目標は周波数ビン毎に設定される自己回帰係数 $g_{d,l}$ の推定し $x_{n,l} - \sum_{d=k}^K g_{d,l} x_{n-d,l}$ によって音源信号を回復することである. 以下 $g_{d,l}$ を残響フィルタと呼ぶ.

また, 残響フィルタ $g_{d,l}$ の他に音源信号のスペクトルモデル (以下単に音源モデルと呼ぶ) $p_{n,l}$ を導入する. $p_{n,l}$ は何らかのパラメータで特徴付けられるとする. 残響フィルタは音源モデルが表すスペクトルの情報を利用して推定されるので, 音源モデルが正確であれば残響フィルタが正確に推定されるが, 両者はともに未知なので同時に推定を行う.

以上より, 残響をうまく推定し音源成分をうまく回復するには以下の2点の選択が重要と言える.

1. 逆畳み込み後信号 $r_{n,l} = x_{n,l} - \sum_{d=k}^K g_{d,l} x_{n-d,l}$ と音源モデル $p_{n,l}$ との間に定義する最適化規準.
2. 音源モデルの定義: 音楽音響信号に対しては調波構造を上手く表現できるかなどを基準にモデルを設計する. こちらは3.2節で詳説する.

2.2 ISダイバージェンス規準の従来法

吉岡らの手法 [3] では以下で定義される最適化規準の最小化により残響フィルタを推定する.

$$Q^{(IS)} = \sum_{n,l} \left(\log p_{n,l} + \frac{|r_{n,l}|^2}{p_{n,l}} \right), \quad (2)$$

* I-divergence-based Dereverberation with Auxiliary Function Approach. by Naoki Yasuraoka (Kyoto University), Hirokazu Kameoka, Takuya Yoshioka (NTT Communication Science Laboratories), and Hiroshi G. Okuno (Kyoto University)

このとき $p_{n,l}$ は音源信号のパワースペクトル成分に対応する $p_{n,l}$ の推定に対してこの最適化規準は $|r_{n,l}|^2$ と $p_{n,l}$ の間の板倉-斎藤 (IS) ダイバージェンスと等価である。そこで、 $(g_{d,l})$ の推定に対しては IS ダイバージェンスと等価ではないがこの最適化規準を以後 IS ダイバージェンスと呼ぶ。IS ダイバージェンスにおいて $g_{d,l}$ は $p_{n,l}$ により重み付けされた二次形式となっており、容易にパラメータが推定できる。

しかしながら、式 (2) に基づき推定された残響フィルタ $g_{d,l}$ は残響抑圧効果が過剰になることがある。この問題の原因は主に以下の 2 点にあると我々は考える。i) 式 (2) では逆畳み込み後信号 $r_{n,l}$ (自己回帰モデルの予測誤差に相当) の重みつき和が 0 に近づくように推定されるから、その結果、楽器音の減衰部分などのように過去信号から容易に説明できる成分は過剰に抑圧される。ii) IS ダイバージェンス規準での音源モデル $p_{n,l}$ の推定が困難であるから、例えば、楽器音の調波構造のモデルである調波 GMM[6] は IS ダイバージェンス規準では上手く動作しない [4]。

2.3 I ダイバージェンス規準の提案法

我々は、残響抑圧後信号 $|r_{n,l}|$ と音源モデル $p_{n,l}$ との間の I ダイバージェンスを最適化規準として用いる。

$$Q^{(l)} = \sum_{n,l} \left(|r_{n,l}| \log \frac{|r_{n,l}|}{p_{n,l}} - (|r_{n,l}| - p_{n,l}) \right). \quad (3)$$

上式において $p_{n,l}$ は今度は音源信号の振幅スペクトルに相当することに注意されたい。I ダイバージェンスを用いることで以下の 2 つの利点がある。i) $g_{d,l}$ の推定は $|r_{n,l}|$ が $p_{n,l}$ に近づくように行なわれるので、過剰な信号の抑圧が発生しにくい。ii) 音源モデルのパラメータ推定が容易で推定精度も良い。調波 GMM のパラメータは後述するように I ダイバージェンス規準で容易に推定可能であり、実際音楽音響信号の音源分離や基本周波数推定に利用されている。

この I ダイバージェンス規準残響抑圧法における一番の課題は $g_{d,l}$ の推定アルゴリズムの導出である。式 (3) において $g_{d,l}$ はもはや二次形式ではなく、式 (3) を最小化する $g_{d,l}$ は解析的に求まらない。

3 I ダイバージェンス規準残響抑圧

本章では I ダイバージェンス規準残響抑圧のための補助関数法を用いたパラメータ推定アルゴリズム (Table 1) について述べる。補助関数法 [8] とは、最小化したい最適化規準 $Q(\theta)$ に対して次の条件：

$$Q(\theta) = \min_{\vartheta} Q^+(\theta, \vartheta) \quad (4)$$

を満たす補助関数 $Q^+(\theta, \vartheta)$ を設計し、 Q^+ に対し補助変数 ϑ に関する最小化と本来の変数 θ に関する最小化を反復することで、間接的に本来の最適化規準を単調減少させる手法である。 $Q^+(\theta, \vartheta)$ を最小化させる θ, ϑ がともに解析的に解けるように Q^+ を設計すればパラメータ推定は簡単化される。

Table 1 Algorithm for parameter estimation

1. $\{g_{d,l}\} = 0$ とみなし音源モデル $p_{n,l}$ を初期化。また $p_{n,l} = |x_{n,l}|^2$ とみなし残響フィルタ $g_{d,l}$ を式 (2) に基づき IS ダイバージェンス規準で初期化。
2. 現在の $g_{d,l}$ を用いて音源モデル $p_{n,l}$ を更新。
3. 現在の $p_{n,l}$ を用いて残響フィルタ $g_{d,l}$ を更新:
 - 3.a 式 (14) により補助変数 $\{t_{n,l}, b_{n,l}, \omega_{n,l}\}$ を更新。
 - 3.b 式 (15) により残響フィルタ $g_{d,l}$ を更新。
 - 3.c 適当な回数 3.a から反復。
4. 収束するまで 2. から反復

3.1 残響フィルタ推定アルゴリズム

I ダイバージェンス規準で残響フィルタ $g_{d,l}$ を推定するための反復更新アルゴリズムについて述べる。以下で述べる方法は、式 (3) に対して補助関数を設計し、 $g_{d,l}$ を解析的に更新できる形すなわち二次形式に帰着させるという方針に基づく。二次形式に至るまで 2 段の補助関数を導入する。

まず、式 (3) 第 1 項の対数関数に対し、次の不等式を考える。

$$\log |r_{n,l}| \leq \log t_{n,l} + \frac{1}{t_{n,l}} (|r_{n,l}| - t_{n,l}), \quad (5)$$

ここで $t_{n,l} > 0$ は補助変数である。この不等式の右辺は対数関数 $\log |r_{n,l}|$ に接点 $t_{n,l}$ で接する接線であり、対数関数が凹関数であることから左辺が右辺を上回ることはない。また、右辺の最小化する $t_{n,l}$ は $t_{n,l} = |r_{n,l}|$ と容易に導出できる。そこで、式 (3) から $|r_{n,l}|$ の対数関数を置き換えた次の補助関数 Q^+ を立てる。

$$Q^+ = \sum_{n,l} \left(\frac{|r_{n,l}|^2}{t_{n,l}} + \left(\log \frac{t_{n,l}}{p_{n,l}} - 2 \right) |r_{n,l}| + p_{n,l} \right). \quad (6)$$

しかしながら、この式は第 2 項に $|r_{n,l}|$ が存在するので、 d で偏微分を 0 と置いた式は依然として d について解析的に解くことはできない。

続いて、式 (6) 第 2 項に対し、 $D_{n,l} = \log(t_{n,l}/p_{n,l}) - 2$ の符号に応じて以下の 2 つの不等式を考える。

$$D_{n,l} \geq 0 \text{ のとき: } |r_{n,l}| \leq \frac{|r_{n,l}|^2}{2b_{n,l}} + \frac{b_{n,l}}{2}, \quad b_{n,l} > 0 \quad (7)$$

$$D_{n,l} < 0 \text{ のとき: } |r_{n,l}| \geq \text{Re} [\omega_{n,l}^* r_{n,l}], \quad |\omega_{n,l}| = 1 \quad (8)$$

ここで $*$ は複素共役を表す。 $b_{n,l}$ は式 (7) に対し、また $\omega_{n,l}$ は式 (8) に対して導入される新たな補助変数である。式 (7) 右辺は一次関数 $|r_{n,l}|$ に接する二次関数である。一方式 (8) 右辺は $r_{n,l}$ と方向ベクトル $\omega_{n,l}$ との内積であり、絶対値関数に対する下限となっている。

不等式 (7, 8) により、式 (6) から $|r_{n,l}|$ の 1 次項を置き換えた以下の補助関数 Q^{++} が設計できる。

$$Q^{++} = \sum_{n,l} Q_{n,l}^{++} \quad (9)$$

$$Q_{n,l}^{++} = \begin{cases} \left(\frac{D_{n,l}}{2b_{n,l}} + \frac{1}{t_{n,l}} \right) |r_{n,l}|^2 + c & (D_{n,l} \geq 0) \\ \frac{1}{t_{n,l}} \left| r_{n,l} + \frac{t_{n,l}}{2} D_{n,l} \omega_{n,l} \right|^2 + c & (D_{n,l} < 0) \end{cases} \quad (10)$$

ここで c は $g_{d,l}$ に非依存の項である．簡単のため

$$\kappa_{n,l} = \begin{cases} x_{n,l} & (D_{n,l} \geq 0) \\ x_{n,l} - \frac{t_{n,l}}{2} D_{n,l} \omega_{n,l} & (D_{n,l} < 0) \end{cases} \quad (11)$$

$$\frac{1}{\rho_{n,l}} = \begin{cases} \frac{1}{t_{n,l}} + \frac{D_{n,l}}{2b_{n,l}} & (D_{n,l} \geq 0) \\ \frac{1}{t_{n,l}} & (D_{n,l} < 0) \end{cases} \quad (12)$$

と表記すると，式 (9) は

$$Q^{++} = \sum_{n,l} \frac{|\kappa_{n,l} - \sum_{d=k}^K g_{d,l} x_{n-d,l}|^2}{\rho_{n,l}} \quad (13)$$

と書き直せる．これは明らかに $g_{d,l}$ に対する二次形式であるため，IS ダイバージェンス基準の従来法と同様に $g_{d,l}$ を更新することができる．

式 (13) のもとでは $g_{d,l}$ は以下のように更新される．まず，3 種の補助変数 $\{t_{n,l}, b_{n,l}, \omega_{n,l}\}$ を次式で更新する．

$$t_{n,l} = |r_{n,l}|, \quad b_{n,l} = |r_{n,l}|, \quad \omega_{n,l} = \frac{r_{n,l}}{|r_{n,l}|} \quad (14)$$

これらはいずれも対応する不等式の等号成立条件である．次に式 (13) の $g_{d,l}$ に対する偏微分を 0 と置き連立方程式を解くことで，以下の更新式が得られる．

$$[g_{k,l}, \dots, g_{K,l}]^T = R_l^{-1} r_l \quad (15)$$

ここで， R_l 及び r_l はそれぞれ修正相関行列及び修正相関ベクトルと呼ばれ，次式で表される．

$$R_l = \begin{bmatrix} \sum_{n=0}^{N-1} \frac{x_{n-k,l}}{\rho_{n,l}} x_{n-k,l}^* & \dots & \sum_{n=0}^{N-1} \frac{x_{n-K,l}}{\rho_{n,l}} x_{n-k,l}^* \\ \vdots & \ddots & \vdots \\ \sum_{n=0}^{N-1} \frac{x_{n-k,l}}{\rho_{n,l}} x_{n-K,l}^* & \dots & \sum_{n=0}^{N-1} \frac{x_{n-K,l}}{\rho_{n,l}} x_{n-K,l}^* \end{bmatrix} \quad (16)$$

$$r_l = \begin{bmatrix} \sum_{n=0}^{N-1} \frac{\kappa_{n,l}}{\rho_{n,l}} x_{n-k,l}^* & \dots & \sum_{n=0}^{N-1} \frac{\kappa_{n,l}}{\rho_{n,l}} x_{n-K,l}^* \end{bmatrix}^T \quad (17)$$

なお R_l と r_l は FFT を用いて高速に計算できることが知られている [9]

3.2 音源モデルとそのパラメータ推定

本節では I ダイバージェンス基準の推定と相性の良い音源モデルを 2 種類紹介する．

3.2.1 調波 GMM (HGMM)

HGMM は楽器音の調波構造を上手く表現するスペクトルモデルである [4, 6, 7]．具体的には，振幅 (IS ダイバージェンス基準ではパワー) スペクトルを周波数軸に整数倍に並んだ調波構造用の鋭いガウス関数群と，拡がりの大きい非調波構造用ガウス関数群の 2 要素で表す．今，各時間フレームに最大 I 個の単音があるとし，各単音が J 個の倍音を持つとすると，HGMM は次式で定義される．

$$p_{n,l} = \sum_{i=1}^I \left(\sum_{j=1}^J u_{i,j,n} H_{i,j,n,l} \right) + \sum_{m=1}^M v_{m,n} I_{m,n,l}, \quad (18)$$

$$H_{i,j,n,l} = \exp \left[-\frac{(\omega_l - j\mu_{i,n})^2}{2\sigma^2} \right], \quad (19)$$

$$I_{m,n,l} = \exp \left[-\frac{(\omega_l - \nu_m)^2}{2\gamma^2} \right], \quad (20)$$

$u_{i,j,n}$ は第 i 音 j 倍音の強度， $\mu_{i,n}$ は基本周波数， σ^2 は調波成分の周波数方向拡がりである．式 (18) 第 2 項は非調波成分のモデルであり，平均 ν_m 及び周波数方向拡がり γ^2 の M 個のガウス関数の和である． ω_l は周波数ビン l に対応する周波数 (Hz) である．

HGMM の推定すべきパラメータ $\{u_{i,j,n}, v_{m,n}, \mu_{i,n}\}$ の I・IS 両ダイバージェンス基準による更新則を以下に示す．残響フィルタ推定と同様補助関数法を用いる．

- 補助変数 $\Psi_{i,j,n,l}^{(I)}$ 及び $\Psi_{i,j,n,l}^{(IS)}$ を以下で定める．

$$\Psi_{i,j,n,l}^{(I)} = \frac{u_{i,j,n} H_{i,j,n,l}}{p_{n,l}} |r_{n,l}| \quad (21)$$

$$\Psi_{i,j,n,l}^{(IS)} = u_{i,j,n} H_{i,j,n,l} \left[1 + u_{i,j,n} H_{i,j,n,l} \frac{|r_{n,l}|^2 - p_{n,l}}{p_{n,l}^2} \right] \quad (22)$$

- $u_{i,j,n}$ を次式で更新する．

$$u_{i,j,n} = \begin{cases} \frac{\sum_l \Psi_{i,j,n,l}^{(I)}}{\sum_l H_{i,j,n,l}} & \text{(I-divergence)} \\ \frac{1}{L} \sum_l \frac{\Psi_{i,j,n,l}^{(IS)}}{H_{i,j,n,l}} & \text{(IS-divergence)} \end{cases} \quad (23)$$

$v_{m,n}$ の更新式も同様である (L は周波数ビンの数)．

- I ダイバージェンス基準では $\mu_{i,n}$ を次式で更新する．

$$\mu_{i,n} = \frac{\sum_{j=1}^J \sum_l \omega_l j \Psi_{i,j,n,l}^{(I)}}{\sum_{j=1}^J \sum_l j^2 \Psi_{i,j,n,l}^{(I)}} \quad (24)$$

IS ダイバージェンス基準では $\mu_{i,n}$ は解析的更新が困難なのでニュートン法などを用いて更新する．

3.2.2 NMF 型モデル (NMF)

非負値行列因子分解 (NMF) は振幅 (パワー) スペクトログラムを含む非負二次元データを因子分解する手法であり，音楽分野では音源分離や自動採譜などに応用されている [1]．NMF のモデル定義及び更新則は [10] を参照されたい．なお NMF による音源分離性能の I・IS ダイバージェンス間の比較では I ダイバージェンス基準が優位であることが報告されている [5]

4 評価実験

本手法の有効性を確認するために行ったシミュレーション実験について述べる．

4.1 実験条件

5 種類の残響抑圧法を比較する．まず，最適化規準が I ダイバージェンスか IS ダイバージェンスか，及び音源モデルが HGMM か NMF か，の組み合わせで

Table 2 Experimental results: The left-hand side in each cell shows the LSD between the true source and the dereverberated signal. The right-hand is the LSD between the true source and the signal given by the source model. Smaller values indicate higher quality. The best performing values are shown in boldface.

		$K = 80$		$K = 160$		$K = 240$	
Derev.-I / HGMM-I (Proposed)	$k = 3$	8.72	8.68	8.68	8.71	8.75	8.78
	$k = 10$	8.65	8.70	8.63	8.72	8.68	8.69
	$k = 20$	8.66	8.68	8.63	8.66	8.62	8.65
Derev.-I / NMF-I (Proposed)	$k = 3$	8.74	9.06	8.78	8.96	8.88	9.03
	$k = 10$	8.64	8.98	8.62	9.08	8.64	8.82
	$k = 20$	8.64	8.78	8.62	8.78	8.63	8.78
Derev.-IS / HGMM-I [4]	$k = 3$	9.82	9.66	9.98	9.82	10.05	9.89
	$k = 10$	9.07	8.98	9.14	9.19	9.16	9.09
	$k = 20$	8.79	8.76	8.82	8.78	8.85	8.82
Derev.-IS / HGMM-IS	$k = 3$	10.06	13.53	10.27	13.61	10.34	13.64
	$k = 10$	9.24	12.69	9.34	12.80	9.32	13.37
	$k = 20$	8.91	13.37	8.98	13.38	9.01	13.41
Derev.-IS / NMF-IS	$k = 3$	10.14	10.06	10.34	10.21	10.40	10.26
	$k = 10$	9.24	9.80	9.36	10.15	9.22	9.32
	$k = 20$	8.90	9.16	8.94	9.22	8.96	9.10

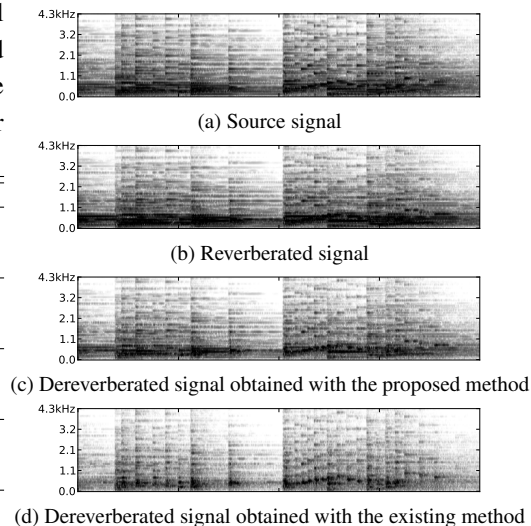


Fig. 1 Example of a dereverberated signal

4種の残響抑圧法を定義する。残り1つは、残響フィルタをISダイバージェンスで推定しHGMM音源モデルをIダイバージェンスで推定する方法[4]である。

RWC音楽データベース[11]のJazz, Classicから選出した12曲に対し、各曲のStandard MIDI Fileから音響信号を合成しこれを真の音源信号とする。この音響信号に2種類のインパルス応答(残響時間はそれぞれ約1, 1.5秒)を畳み込み入力信号 $x_{n,l}$ を合成した。残響を推定抑圧したのち、抑圧結果の1)逆畳み込み後信号 $\mathbb{R} = \{\hat{r}_{n,l}\}$ 、2)音源モデル振幅スペクトル $\mathbb{P} = \{\hat{p}_{n,l}\}$ 、がそれぞれ真の音源信号 $\mathbb{S} = \{s_{n,l}\}$ にどれだけ近いかを評価する。ただし $\hat{\cdot}$ は推定結果を示す。最適化規準は以下の対数スペクトル距離とする。

$$\text{LSD}(\mathbb{Y}, \mathbb{Z}) \equiv \sum_{n,l} \left| 20 \log_{10} \left| \frac{y_{n,l}}{z_{n,l}} \right| \right| / NL \quad (25)$$

ここで N 及び L は時間フレーム、周波数ピンの数である。LSD(\mathbb{S}, \mathbb{R})では残響抑圧後信号の正確さを評価し、一方LSD(\mathbb{S}, \mathbb{P})では音源スペクトルの推定精度を評価することに相当する。全ての音響信号は44.1kHzでサンプリングし、256点間隔で1024点ガウス窓を用いて時間周波数解析した。

4.2 実験結果と考察

Table 2に、初期反射時間遅れ k 及びフィルタ長 K に対して複数の条件の下で残響抑圧実験を行った結果の対数スペクトル距離の12曲による平均を示す。各セルの左の数値はLSD(\mathbb{S}, \mathbb{R})を、右の数値はLSD(\mathbb{S}, \mathbb{P})を表す。この結果から、Iダイバージェンス規準の提案法がISダイバージェンスを用いる従来法よりも高精度であることが分かる。また、[4]より提案法の方がLSD(\mathbb{S}, \mathbb{R})の値が小さいことから、音源モデルだけでなく残響フィルタもIダイバージェンス規準で推定することの有効性が分かる。

Fig. 1にはピアノ曲に対する従来法と提案法の残響抑圧結果のスペクトログラムを示す。従来法(d)に比べ提案法(c)が残響以外の過剰な抑圧を避けられていることが分かる。

5 まとめ

本稿では逆畳み込み後信号と音源モデルの間の最適化規準にIダイバージェンスを用いる新たな残響抑圧法を報告し、また補助関数法に基づく収束性が保証されたパラメータ推定アルゴリズムを示した。評価実験では従来用いられていたISダイバージェンス規準による手法と比較を行い、提案法の有効性を示した。

今後の課題として、音源分離や楽器音分析合成といった他の音楽音響信号処理課題への本残響抑圧法の統合や、マルチチャンネル処理、特にステレオ信号への拡張が挙げられる。

謝辞 本研究は科研費(S)、グローバルCOE、Crest-Museの支援を受けた。

参考文献

- [1] P. Smaragdis and J.C Brown. Non-negative matrix factorization for polyphonic music transcription. In *Proc. WASPAA*, pp. 170–180, 2003.
- [2] M. Goto. Active music listening interfaces based on signal processing. In *Proc. ICASSP*, Vol. 4, pp. 1441–1444, 2007.
- [3] T. Yoshioka, T. Nakatani, and M. Miyoshi. An integrated method for blind separation and dereverberation of convolutive audio mixtures. In *Proc. EUSIPCO*, 2008.
- [4] N. Yasuraoka, T. Yoshioka, T. Nakatani, A. Nakamura, and H. G. Okuno. Music dereverberation using harmonic structure source model and Wiener filter. In *Proc. ICASSP*, pp. 53–56, 2010.
- [5] D. FitzGerald, M. Cranitch, and E. Coyle. On the use of the beta divergence for musical source separation. In *Proc. ISSC*, pp. 1–6, 2009.
- [6] H. Kameoka, T. Nishimoto, and S. Sagayama. Extraction of multiple fundamental frequencies from polyphonic music using harmonic clustering. In *Proc. ICA*, pp. 1–59–62, 2004.
- [7] H. Kameoka, T. Nishimoto, and S. Sagayama. A multipitch analyzer based on harmonic temporal structured clustering. *IEEE Trans. Audio, Speech and Lang. Process.*, Vol. 15, No. 3, pp. 982–994, 2007.
- [8] D. D. Lee and H. S. Seung. Algorithms for non-negative matrix factorization. In *Proc. NIPS*, pp. 556–562, 2001.
- [9] T. Yoshioka, T. Nakatani, and M. Miyoshi. Fast algorithm for conditional separation and dereverberation. In *Proc. EUSIPCO*, 2009.
- [10] M. Nakano, H. Kameoka, J. Le Roux, Y. Kitano, N. Ono, and S. Sagayama. Convergence-guaranteed multiplicative algorithms for nonnegative matrix factorization with beta-divergence. In *Proc. MLSP*, 2010.
- [11] M. Goto, H. Hashiguchi, T. Nishimura, and R. Oka. RWC music database: Popular, classical, and jazz music databases. In *Proc. ISMIR*, pp. 287–288, 2002.