

リカレント構造を持つ複素制限ボルツマンマシンによる複素スペクトル系列モデリング

中鹿 亘^{1,a)} 高木 信二² 山岸 順一²

概要：識別器や特徴抽出など、様々な分野で広く用いられる制限ボルツマンマシン (restricted Boltzmann machine; RBM) はバイナリ値または実数値の入力データを仮定していたが、音声信号処理では複素スペクトルなど複素数値データを取り扱う場合が多い。そこで我々の先行研究では、複素数データから直接潜在的な特徴量 (隠れ素子) を抽出する RBM の拡張モデル (Complex-valued RBM; CRBM; 複素 RBM) を提案した。本研究ではさらに、音声信号が時系列データであることに着目し、隠れ素子間の時間的な依存関係を考慮したモデル (Complex-valued temporal RBM; CTRBM; 複素 TRBM) を新たに定義し、複素スペクトル系列モデリングの有効性を検証する。

Complex spectral sequence modeling using complex-valued restricted Boltzmann machine with recurrent structure

TORU NAKASHIKA^{1,a)} SHINJI TAKAKI² JUNICHI YAMAGISHI²

1. はじめに

近年、ディープラーニングを用いた手法が飛躍的に高い精度を上げ、画像認識や音声認識など、幅広い分野において盛んに研究されている [4]。これまでに数多くのディープラーニング手法が提案されているが、最も代表的なモデルとして、制限ボルツマンマシン (restricted Boltzmann machine; RBM) を多層に積み重ねた Deep Belief Net (DBN) が挙げられる [2]。また、RBM 単独でも特徴量抽出手法としてしばしば用いられ、様々な RBM の拡張モデルも提案されている [3], [7], [8], [9]。以上のように、RBM はこれまで様々な分野で用いられてきたにもかかわらず、いずれのアプローチでも入力特徴量はバイナリまたは実数値が仮定されてきた [1], [2], [5]。しかし、複素スペクトル、MRI 画像、音響インテンシティなど、実データに基づいた画像認識や音声信号処理では複素数データを取り扱う場合が多い。特

に音声信号処理では、位相スペクトルよりも振幅スペクトルの方が認識や合成に効果的であることが知られており、MFCC やメルケプストラム特徴量、STRAIGHT スペクトルなど振幅スペクトルに基づいた音響特徴量がしばしば利用されるが、これらは位相情報が欠落しており、元の音声データに対して少なからず情報の損失が存在する。そこで我々は、複素スペクトルから直接潜在的な特徴量 (隠れ素子) を抽出できる RBM の拡張モデル (Complex-valued RBM; CRBM; 複素 RBM) を提案した [6]。

しかしながら、従来の複素 RBM ではフレーム独立な複素スペクトルのモデリングを行っており、時間的に依存関係を持つ音声信号の複素スペクトル系列を正しく表現できるとは言い難い。そこで本研究では、時間的に隣接する隠れ素子間の依存関係を考慮した複素 RBM の拡張モデル (Complex-valued temporal RBM; CTRBM; 複素 TRBM) を新たに提案し、その有効性について検証する。複素 TRBM は、バイナリ値や実数値の可視素子に対して時間的な接続を考慮した TRBM (Temporal RBM [10]) の複素数拡張とみなすこともできる。

¹ 電気通信大学
1-5-1 Chofugaoka, Chofu, Tokyo 182-8585, Japan

² 国立情報学研究所
2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan

a) nakashika@uec.ac.jp

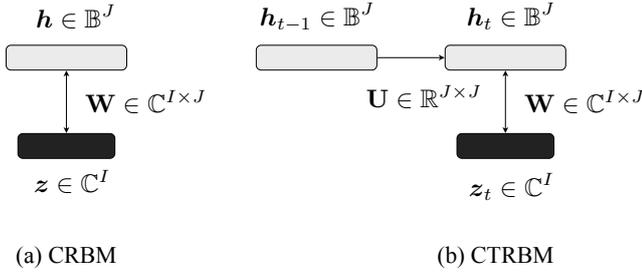


図 1 複素 RBM (a) および複素 TRBM (b) のグラフ表現 .

2. 従来モデル：複素 RBM

複素 RBM (Complex-valued restricted Boltzmann machine [6]) は, 図 1 (a) に示すように, I 次元の複素数の可視素子 $z \in \mathbb{C}^I$ と, J 次元の 2 値の隠れ素子 $h \in \mathbb{B}^J$ の関係性を記述する確率モデルであり, 次式で定義される .

$$p(z; \theta) = \sum_{\mathbf{h}} p(z, \mathbf{h}; \theta) \quad (1)$$

$$p(z, \mathbf{h}; \theta) = \frac{1}{U(\theta)} e^{-E(z, \mathbf{h}; \theta)} \quad (2)$$

$$E(z, \mathbf{h}; \theta) = \frac{1}{2} \begin{bmatrix} z \\ \bar{z} \end{bmatrix}^H \Phi^{-1} \begin{bmatrix} z \\ \bar{z} \end{bmatrix} - \begin{bmatrix} \mathbf{b} \\ \bar{\mathbf{b}} \end{bmatrix}^H \Phi^{-1} \begin{bmatrix} z \\ \bar{z} \end{bmatrix} - \mathbf{c}^T \mathbf{h} \quad (3)$$

$$U(\theta) = \int \sum_{\mathbf{h}} e^{-E(z, \mathbf{h}; \theta)} d\mathbf{z} \quad (4)$$

$$\Phi = \begin{bmatrix} \Gamma & \mathbf{C} \\ \mathbf{C}^H & \Gamma^H \end{bmatrix} \quad (5)$$

$$\Gamma = \text{diag}(\gamma) \quad (6)$$

$$\mathbf{C} = \text{diag}(\delta) \quad (7)$$

ただし, $\theta = \{b, c, \mathbf{W}, \gamma, \delta\}$ は複素 RBM のパラメータであり, 可視素子のバイアス $b \in \mathbb{C}^I$, 隠れ素子のバイアス $c \in \mathbb{R}^J$, 可視素子-隠れ素子間の複素結合重み $\mathbf{W} \in \mathbb{C}^{I \times J}$, 複素可視素子の分散 $\gamma \in \mathbb{R}_+^I$ と擬似分散 (共役複素数との共分散) $\delta \in \mathbb{C}^I$ を含む . また, オーバーラインは複素共役, \cdot^H はエルミート転置, \cdot^T は転置, $\text{diag}(\cdot)$ は入力されたベクトルを対角要素とした行列を出力する関数を表す .

以上の定義により, 隠れ素子が与えられたときの可視素子の条件付き確率, および可視素子が与えられたときの隠れ素子の条件付き確率をそれぞれ以下のように表すことができる .

$$p(z|\mathbf{h}) = \mathcal{N}_c(z; \mathbf{b} + \mathbf{W}\mathbf{h}, \Gamma, \mathbf{C}) \quad (8)$$

$$p(\mathbf{h}|z) = \mathcal{B}(\mathbf{h}; \rho(\mathbf{c} + 2\Re(\mathbf{W}^H z))) \quad (9)$$

ただし, $\mathcal{B}(\cdot; \pi)$ は成功確率 π の多次元ベルヌーイ分布, $\rho(\cdot)$ は要素毎のシグモイド関数, $\mathcal{N}_c(\cdot; \mu, \Gamma, \mathbf{C})$ は平均 μ , 分散共分散行列 Γ , 擬似分散共分散行列 \mathbf{C} の多変量複素正規分布 :

$$p(z) = \frac{1}{\pi^D \sqrt{\det(\Gamma)\det(\mathbf{C})}} \cdot \exp \left\{ -\frac{1}{2} \begin{bmatrix} z - \mu \\ \bar{z} - \bar{\mu} \end{bmatrix}^H \begin{bmatrix} \Gamma & \mathbf{C} \\ \mathbf{C}^H & \Gamma^H \end{bmatrix}^{-1} \begin{bmatrix} z - \mu \\ \bar{z} - \bar{\mu} \end{bmatrix} \right\} \quad (10)$$

$$\mathbf{Q} \triangleq \bar{\Gamma} - \mathbf{C}^H \Gamma^{-1} \mathbf{C} \quad (11)$$

を表す . また, (9) 式における \mathbf{W}' は,

$$\mathbf{W}' \triangleq \text{diag}(\mathbf{p})\mathbf{W} + \text{diag}(\mathbf{q})\bar{\mathbf{W}} \quad (12)$$

$$\mathbf{p} \triangleq \frac{\gamma}{\gamma^2 - |\delta|^2} \quad (13)$$

$$\mathbf{q} \triangleq -\frac{\delta}{\gamma^2 - |\delta|^2} \quad (14)$$

とおいた . ただし分数線は要素除算, \cdot^2 は要素ごとの 2 乗を表す .

3. 提案モデル：複素 TRBM

区間 (A, B) の系列変数を $\mathbf{x}_{A:B} = (\mathbf{x}_A, \mathbf{x}_{A+1}, \dots, \mathbf{x}_B)$ と表記する . 例えば T フレームの複素特徴量系列は $\mathbf{z}_{1:T}$ と表記される . 前述の複素 RBM では, 全てのフレームの複素特徴量 $\mathbf{z}_t \in \mathbb{C}^I$ ($1 \leq t \leq T$) について対応する隠れ素子 $\mathbf{h}_t \in \mathbb{B}^J$ を用いて同一の確率分布 $p(\mathbf{z}_t, \mathbf{h}_t)$ を与え, 各フレーム特徴量が独立, すなわち

$$p(\mathbf{z}_{1:T}, \mathbf{h}_{1:T}) = \prod_{t=1}^T p(\mathbf{z}_t, \mathbf{h}_t) \quad (15)$$

と仮定していた . しかし, 音声信号は時間的に相関のある系列データであり, そこから得られる複素スペクトルなどの特徴量も時間的に依存関係を持つが, 複素 RBM ではそのような関係性を考慮していなかった . そこで本研究では, 時間的に依存関係のある複素特徴量系列を表現するため, 複素 TRBM (complex-valued temporal restricted Boltzmann machine; CTRBM) を新たに提案する .

複素 TRBM では, T フレームの複素特徴量系列データ $\mathbf{z}_{1:T}$ および潜在変数系列 $\mathbf{h}_{1:T}$ をマルコフ過程と仮定し, 同時確率を

$$p(\mathbf{z}_{1:T}, \mathbf{h}_{1:T}) = \prod_{t=1}^T p(\mathbf{z}_t, \mathbf{h}_t | \mathbf{h}_{t-1}) \quad (16)$$

と表現する . ここで, 条件付き確率分布 $p(\mathbf{z}_t, \mathbf{h}_t | \mathbf{h}_{t-1})$, ($t = 2, \dots, T$) は, 図 1 (b) に示すように, 複素 TRBM のパラメータ θ を用いて以下の通り定義される .

$$p(\mathbf{z}_t, \mathbf{h}_t | \mathbf{h}_{t-1}; \boldsymbol{\theta}) = \frac{1}{U_{\mathbf{h}_{t-1}}(\boldsymbol{\theta})} e^{-E(\mathbf{z}_t, \mathbf{h}_t | \mathbf{h}_{t-1}; \boldsymbol{\theta})} \quad (17)$$

$$E(\mathbf{z}_t, \mathbf{h}_t | \mathbf{h}_{t-1}; \boldsymbol{\theta}) = \frac{1}{2} \begin{bmatrix} \mathbf{z}_t \\ \bar{\mathbf{z}}_t \end{bmatrix}^H \boldsymbol{\Phi}^{-1} \begin{bmatrix} \mathbf{z}_t \\ \bar{\mathbf{z}}_t \end{bmatrix} - \begin{bmatrix} \mathbf{b} \\ \bar{\mathbf{b}} \end{bmatrix}^H \boldsymbol{\Phi}^{-1} \begin{bmatrix} \mathbf{z}_t \\ \bar{\mathbf{z}}_t \end{bmatrix} - \mathbf{c}^\top \mathbf{h}_t - \mathbf{h}_{t-1}^\top \mathbf{U} \mathbf{h}_t \quad (18)$$

$$U_{\mathbf{h}_{t-1}}(\boldsymbol{\theta}) = \int \sum_{\mathbf{h}_t} e^{-E(\mathbf{z}_t, \mathbf{h}_t | \mathbf{h}_{t-1}; \boldsymbol{\theta})} d\mathbf{z}_t \quad (19)$$

$\mathbf{U} \in \mathbb{R}^{J \times J}$ は隣接する隠れ素子間の結合重みパラメータを表す．複素 RBM のエネルギー関数と比較すると，複素 TRBM では隠れ素子の再帰項 $-\mathbf{h}_{t-1}^\top \mathbf{U} \mathbf{h}_t$ が追加されている．式 (15) 及び式 (16) や図 1 を見ても分かるように，複素 TRBM では，複素 RBM と異なり，一つ前の時刻の隠れ素子に依存して現在の時刻の確率が定義される．また， $t = 1$ のときの確率分布を以下のように定義する．

$$p(\mathbf{z}_1, \mathbf{h}_1; \boldsymbol{\theta}) = \frac{1}{U_{\mathbf{h}_0}(\boldsymbol{\theta})} e^{-E(\mathbf{z}_1, \mathbf{h}_1; \boldsymbol{\theta})} \quad (20)$$

$$E(\mathbf{z}_1, \mathbf{h}_1; \boldsymbol{\theta}) = \frac{1}{2} \begin{bmatrix} \mathbf{z}_1 \\ \bar{\mathbf{z}}_1 \end{bmatrix}^H \boldsymbol{\Phi}^{-1} \begin{bmatrix} \mathbf{z}_1 \\ \bar{\mathbf{z}}_1 \end{bmatrix} - \begin{bmatrix} \mathbf{b} \\ \bar{\mathbf{b}} \end{bmatrix}^H \boldsymbol{\Phi}^{-1} \begin{bmatrix} \mathbf{z}_1 \\ \bar{\mathbf{z}}_1 \end{bmatrix} - \mathbf{c}_0^\top \mathbf{h}_1 \quad (21)$$

$$U_{\mathbf{h}_0}(\boldsymbol{\theta}) = \int \sum_{\mathbf{h}_1} e^{-E(\mathbf{z}_1, \mathbf{h}_1 | \mathbf{h}_0; \boldsymbol{\theta})} d\mathbf{z}_1 \quad (22)$$

ただし， $\mathbf{c}_0 \in \mathbb{R}^J$ は隠れ素子の初期バイアスパラメータである．ここで $\mathbf{h}_0 \triangleq \mathbf{U}^{-\top}(\mathbf{c}_0 - \mathbf{c})$ とすると， $t = 1$ のときも確率分布を式 (17)(18)(19) で表すことができる．

以上の定義式より，一つ前の時刻の隠れ素子が与えられたときの可視素子および隠れ素子の条件付き確率をそれぞれ以下の計算できる．

$$p(\mathbf{z}_t | \mathbf{h}_t, \mathbf{h}_{t-1}) = p(\mathbf{z}_t | \mathbf{h}_t) \quad (23)$$

$$= \mathcal{N}_c(\mathbf{z}_t; \mathbf{b} + \mathbf{W} \mathbf{h}_t, \boldsymbol{\Gamma}, \mathbf{C}) \quad (24)$$

$$p(\mathbf{h}_t | \mathbf{z}_t, \mathbf{h}_{t-1}) = \mathcal{B}(\mathbf{h}_t; \boldsymbol{\rho}(\mathbf{c} + \mathbf{U}^\top \mathbf{h}_{t-1} + 2\Re(\mathbf{W}'^H \mathbf{z}_t))) \quad (25)$$

3.1 パラメータ推定

複素 TRBM のパラメータは勾配法に基づく最尤推定によって最適化を行う．学習用の複素系列データ $\mathbf{z}_{1:T}$ の対

数尤度 L は

$$L = \log p(\mathbf{z}_{1:T}) = \log \sum_{\mathbf{h}_{1:T}} p(\mathbf{z}_{1:T}, \mathbf{h}_{1:T}) \quad (26)$$

と表される．ここで，どのフレームについても正規化項 $U_{\mathbf{h}_{t-1}}$ が等しい，すなわち $U_{\mathbf{h}_0} = U_{\mathbf{h}_1} = \dots = U_{\mathbf{h}_T} = U$ と条件を緩和すると，パラメータ $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ に関する偏微分は

$$\frac{\partial L}{\partial \boldsymbol{\theta}} = \mathbf{E}_{p(\mathbf{h}_{1:T} | \mathbf{z}_{1:T})} \left[-\frac{\partial E}{\partial \boldsymbol{\theta}} \right] - \mathbf{E}_{p(\mathbf{h}_{1:T}, \mathbf{z}_{1:T})} \left[-\frac{\partial E}{\partial \boldsymbol{\theta}} \right] \quad (27)$$

と計算できる．各パラメータに関するエネルギー関数の偏微分 $\frac{\partial E}{\partial \boldsymbol{\theta}}$ は解析的に求まるが，紙面の都合上省略する．右辺第一項は学習データが与えられた時の隠れ素子系列についての期待値で，

$$p(\mathbf{h}_{1:T} | \mathbf{z}_{1:T}) = \prod_{t=1}^T p(\mathbf{h}_t | \mathbf{z}_t, \mathbf{h}_{t-1}) \quad (28)$$

となり，確率自体は容易に導出できる．ただし，式 (25) に示すように，フレーム t における期待値は一つ前の時刻の隠れ素子の値 \mathbf{h}_{t-1} が必要となる．RTRBM (recurrent TRBM [11]) のようにバイナリ値の潜在変数と，その期待値を表す変数を用いて学習し，推論時に期待値を伝播させる方法も考えられるが，本研究では単純に推論時のみ隠れ素子の期待値を伝播させて近似計算する．また，式 (27) 右辺第二項は膨大な数の項が含まれ計算困難であるが，通常の RBM と同様 CD (contrastive divergence [2]) 法によって近似することができる．

具体的には，各フレーム $1 \leq t \leq T$ について，式 (27) 右辺第一項では

$$\hat{\mathbf{h}}_t \triangleq \mathbb{E}_{p(\mathbf{h}_t | \mathbf{z}_t, \hat{\mathbf{h}}_{t-1})} [\mathbf{h}_t] \quad (29)$$

$$= \boldsymbol{\rho}(\mathbf{c} + \mathbf{U}^\top \hat{\mathbf{h}}_{t-1} + 2\Re(\mathbf{W}'^H \mathbf{z}_t)) \quad (30)$$

から得られる隠れ素子を用い，第二項では

$$\tilde{\mathbf{h}}_t \sim p(\mathbf{h}_t | \mathbf{z}_t, \tilde{\mathbf{h}}_{t-1}) \quad (31)$$

$$\tilde{\mathbf{z}}_t \sim p(\mathbf{z}_t | \tilde{\mathbf{h}}_t) \quad (32)$$

から得られるサンプルおよび期待値

$$\hat{\tilde{\mathbf{h}}}_t \triangleq \mathbb{E}_{p(\mathbf{h}_t | \tilde{\mathbf{z}}_t, \hat{\tilde{\mathbf{h}}}_{t-1})} [\mathbf{h}_t] \quad (33)$$

$$= \boldsymbol{\rho}(\mathbf{c} + \mathbf{U}^\top \hat{\tilde{\mathbf{h}}}_{t-1} + 2\Re(\mathbf{W}'^H \tilde{\mathbf{z}}_t)) \quad (34)$$

を用いて勾配を計算する．

4. 評価実験

提案手法の有効性を確認するため，ATR 音声データベース・セット A から女性アナウンサー (FTK) の音声を用いて再構築音声の品質評価実験を行なった．同セットから 50 文音声 (約 4.2 分，サンプリングレート 20kHz を 16kHz にダウンサンプリング) を使用し，窓幅 256，64 サンプルオー

表 1 各手法による PESQ 評価 .

Method	TRBM	CRBM	CTRBM
PESQ	1.699	1.792	1.794

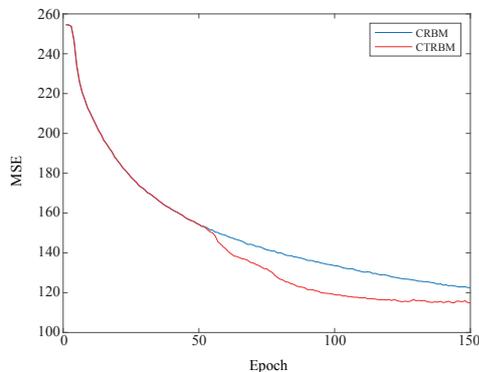


図 2 学習時における MSE の比較 .

オーバーラップの短時間フーリエ変換を施して得られた複素スペクトル (129 次元) を $T = 1,000$ フレーム連続して並べたものを複素特徴量系列とし、隠れ素子数 $J = 200$ の複素 TRBM の学習を行った . なお、このとき得られた総フレームは 64,438 であり、 $[64,438/1,000] = 64$ 個の系列データを用いて、学習率 0.001、減衰率 $\beta_1 = 0.9$ 、 $\beta_2 = 0.999$ 、エポック数 150 の複素 Adam に基づく確率的勾配法によりパラメータを推定した . なお、安定して学習を行うため、最初の 50 エポックは $U = O$ とし (複素 RBM とみなし) て学習し、その後の 100 エポックで全パラメータを更新した . また比較手法として、同じ複素入力特徴量の、実部と虚部を連結したベクトルを可視素子として同様の条件 (ただし、複素 Adam ではなく通常の Adam を用いている) で学習させた TRBM、系列でなく各フレーム特徴量とみなし、バッチサイズ 1,000 で学習させた複素 RBM を用いた . モデルの評価として、学習とは異なる別の 53 文の、同じ話者の音声データから計算される複素スペクトルを各モデルでエンコード (隠れ素子を計算) およびデコード (隠れ素子から可視素子を計算) し、復元された音声に対して算出される PESQ (Perceptual evaluation of speech quality) を算出した .

各手法による復元音声の品質結果を表 1 に示す . 表 1 より、実数値表現 (TRBM) よりも複素数表現 (複素 RBM、複素 TRBM) の方が精度が高く、さらに提案手法が僅かに従来の複素 RBM よりも良い精度を示している . また、学習中の MSE (図 2) から、50 エポック目から開始される複素 TRBM の学習の方が、従来の複素 RBM よりも収束が早く、効率的であることが分かる .

また、提案手法の複素 TRBM により推定されたパラメータ W (絶対値を取ったもの)、 U をそれぞれ図 3 (b)、図 4 に示す . 比較のため、従来の複素 RBM で推定されたパラメータ W (絶対値を取ったもの) を図 3 (a) に示す . 図 3

より、 W の各列ベクトルがそれぞれの隠れ素子に対応する複素スペクトル基底だとみたときに、複素 TRBM で得られる基底の方が、複素 RBM の基底よりも互いに異なるものになっている . このことから複素 TRBM は隠れ素子のアクティベーションをよりスパースに導くことができると推察される . また、図 4 より、推定された U の中で僅かに対角要素の値が他の要素の値よりも大きいことから、同じ隠れ素子を時間的に継続させる働きを持つことが分かる . 最後に、オリジナル音声の振幅スペクトルと、提案法により復元された複素スペクトルの振幅の例を図 5 に示す . 図 5 より、提案法の音声符号化では、バイナリ値の隠れ素子 200 個のみから複素スペクトルを復元したにも関わらず、音声の調波構造やフォルマントが十分表現できていることが確認できる .

5. おわりに

本研究では、時間的に相関のある 2 値の隠れ素子と、複素数の可視素子との相互の関係性を表現する新しい確率モデル : 複素 TRBM を提案し、その有効性を示した . 今後は、複素 RNN の事前学習として用いるなど、複素 TRBM の応用について検証していきたい .

参考文献

- [1] Freund, Y. and Haussler, D.: Unsupervised learning of distributions of binary vectors using two layer networks (1994).
- [2] Hinton, G. E., Osindero, S. and Teh, Y.-W.: A fast learning algorithm for deep belief nets, *Neural computation*, Vol. 18, No. 7, pp. 1527–1554 (2006).
- [3] Krizhevsky, A. and Hinton, G. E.: Factored 3-way restricted Boltzmann machines for modeling natural images, *Journal of Machine Learning Research* (2010).
- [4] LeCun, Y., Bengio, Y. and Hinton, G.: Deep learning, *Nature*, Vol. 521, No. 7553, pp. 436–444 (2015).
- [5] Lee, H., Ekanadham, C. and Ng, A. Y.: Sparse deep belief net model for visual area V2, pp. 873–880 (2008).
- [6] Nakashika, T., Takaki, S. and Yamagishi, J.: Complex-valued restricted Boltzmann machine for direct learning of frequency spectra, *Interspeech* (2017).
- [7] Nakashika, T., Takiguchi, T. and Minami, Y.: Non-Parallel Training in Voice Conversion Using an Adaptive Restricted Boltzmann Machine, *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, Vol. 24, No. 11, pp. 2032–2045 (2016).
- [8] Salakhutdinov, R. and Hinton, G. E.: Deep Boltzmann Machines, *AISTATS* (2009).
- [9] Sohn, K., Zhou, G., Lee, C. and Lee, H.: Learning and Selecting Features Jointly with Point-wise Gated Boltzmann Machines, *ICML (2)* (2013).
- [10] Sutskever, I. and Hinton, G.: Learning multilevel distributed representations for high-dimensional sequences, *Artificial Intelligence and Statistics*, pp. 548–555 (2007).
- [11] Sutskever, I., Hinton, G. E. and Taylor, G. W.: The recurrent temporal restricted Boltzmann machine, *Advances in Neural Information Processing Systems*, pp. 1601–1608 (2009).

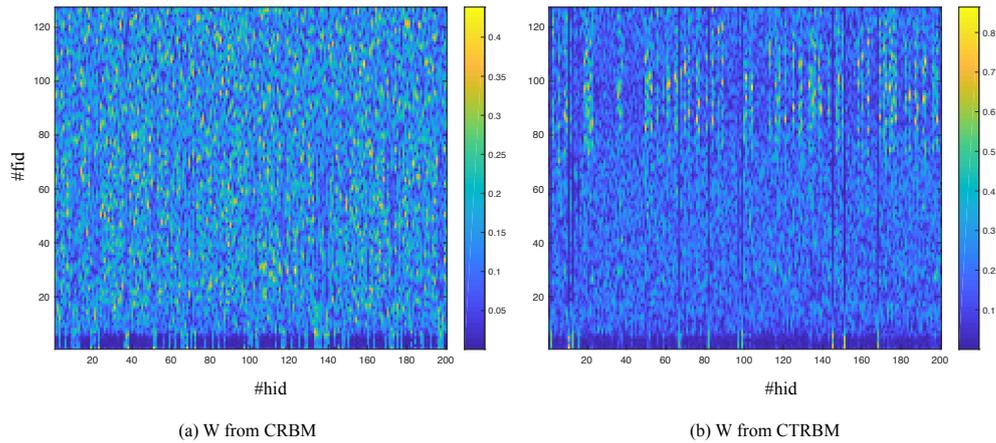


図 3 推定された、可視素子・隠れ素子間の結合重みパラメータ W .

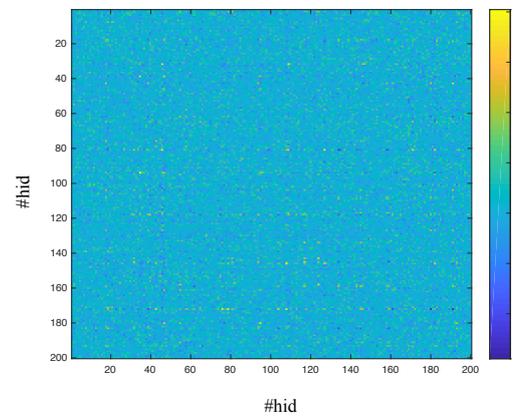


図 4 推定された、隣接する隠れ素子間の結合重みパラメータ U .

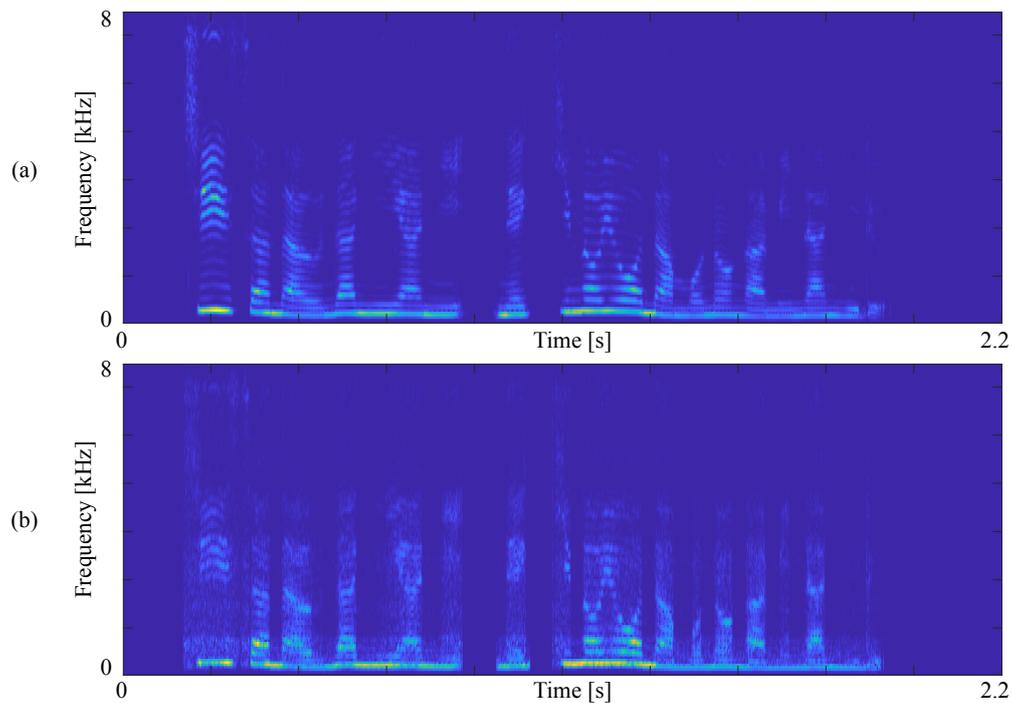


図 5 元のスペクトログラム (a) と、複素 TRBM によって復元されたスペクトログラム (b) .