

## エアコン音の聴感印象推定のための比較判断を考慮した脳活動特徴量抽出\*

☆矢野 肇 (神戸大/産総研), 滝口哲也, 有木康雄 (神戸大),  
神谷 勝 ((株)デンソー), 中川誠司 (千葉大/産総研)

### 1 はじめに

新たな音の印象評価の手法として, 脳活動計測に基づく印象評価の方法が提案されている<sup>[1-3]</sup>. これらの手法は, 従来の質問紙法などの心理的手法に比べて, 心理的バイアスの影響の低減や意識下の印象評価の可能性といったメリットを有している.

我々はこれまで, エアコン音の聴取によって惹起される印象の評価に取り組んできた. エアコン音に時間変動を与えた場合の涼しさ, 好ましさの印象を一对比較法で推定し, エアコン音聴取時の自発律動の時間的安定性, および事象関連同期/脱同期との関係を調査した. 印象と脳活動との間にいくつかの相関を見出した<sup>[4,5]</sup>. これらの知見は, 脳活動指標を用いた印象評価の可能性を示唆するものであるものの, 得られた印象と脳活動との間の相関は大きくなく, 現状ではこれらの指標による印象評価は困難であると言わざるをえない.

その一方で, 脳活動特徴量を入力として, エアコン音の聴感印象の尺度値を推定する印象予測モデルを構築し, より頑健なエアコン音の印象評価指標の構築, および印象と関連のある脳活動の抽出にも取り組んできた<sup>[6,7]</sup>. 前稿<sup>[7]</sup>では, エアコン音聴取時の脳磁界の時間一周波数特徴から, 非負値テンソル分解 (Non-negative Tensor Factorization: NTF) によって, 時間一周波数特徴の持つ多次元の構造を考慮した低次元の特徴量を抽出した. また, この特徴量を入力特徴量とし, 刺激対の涼しさ, 好ましさに関する一对比較判断を教師データとして印象予測モデルを学習した<sup>[7]</sup>. 一对比較判断の予測に基づいてモデルの評価を行った結果, 回帰モデルを用いた手法と比較して平均の正解率が向上した. しかしながら, その平均正解率は 60% に至っておらず, 脳活動指標を用いた印象評価には未だ多くの課題が残されている.

本稿では, 印象予測モデルの予測精度向上を目指し, 特徴量抽出手法の改善を試みた. NTF による特徴量抽出は, テンソルデータの多次元構造を捉えることができ, 特徴量の各次元に対応する脳活動を基底行列から読み取ることができるという利点がある. その一方で, NTF では近似されたテンソルと元のテンソルとの距離が小さくなるようにフィッティングさせている

に過ぎず, 必ずしも印象の予測に有効な特徴量が抽出されているとは言い難い. そこで, Discriminant Non-negative Tensor Factorization (DNMF) を用いて, 刺激対に対する優劣の判断の情報を NTF のコスト関数に組み込むことで, 印象の予測に有効な特徴量抽出を試みた. また, 得られた特徴量を, 前稿で提案した印象予測モデルに適用し, その予測精度から特徴量抽出手法の有効性を検証した.

### 2 脳磁界計測実験

#### 2.1 刺激音

乗用車内の運転者の着座位置 (ヘッドレスト付近) で計測されたエアコン音を基に, 線形予測符号化 (Linear Predictive Coding: LPC) によってスペクトル包絡を抽出することでモデル化されたエアコン音<sup>[8]</sup>を用いて刺激音を作成した. このモデル音に対し, 時間変動として正弦波による振幅変調を施し, 変調周波数を変化させて 7 種類の刺激音 (0 (変調なし), 0.2, 0.4, 0.6, 0.8, 1.6, 3.2 Hz) を作成した. なお, 刺激音の持続時間は 5 s, 変調度は 0.15 に固定し, 刺激音圧は被験者毎に, 変調周波数 0 Hz (変調なし) の刺激音が “はっきりと聞こえ, なおかつちょうど良いと感じる” 大きさに設定した.

#### 2.2 方法

実験参加の同意を得た聴覚健常者 8 名 (男性 6 名, 女性 2 名, 21-24 歳) を被験者とし, 刺激音聴取時の脳磁界を計測した. 刺激音は予め被験者ごとに 7 つの刺激音に対する涼しさ, および好ましさの尺度値を一对比較法によって調査し, 各印象の尺度値が最大, 最小となる刺激音を含む計 4 つの刺激音を選択した. 被験者に 4 つの刺激音のうち 2 つずつを連続的に呈示し, 涼しさ, 好ましさの各基準における優劣を評価させた. なお, 刺激音は挿入型イヤホンによって被験者の両耳に呈示された.

脳磁界計測は磁気シールドルーム内で 122ch 全頭型脳磁界計測システム (Neuromag-122<sup>TM</sup>, Neuromag Ltd.) を用いて行われた. 全 122ch で計測された脳磁界信号は通過帯域が 0.03-100 Hz のアナログバンドパスフィルタを通過した後, 400 Hz でサンプリングされた.

\*Extraction of brain cortical feature considering comparative judgment for estimation of subjective auditory impressions induced by HVAC sound. by YANO, Hajime (Kobe Univ./AIST), TAKIGUCHI, Tetsuya, ARIKI, Yasuo (Kobe Univ.), KAMIYA, Masaru (DENSO Corp.), NAKAGAWA, Seiji (Chiba Univ./AIST).

### 3 特徴量抽出

#### 3.1 時間一周波数特徴

各チャンネルにおける刺激呈示後 0–5000 ms の raw data に対し、連続ウェーブレット変換を用いて時間一周波数特徴を抽出した。時間信号を  $x(t)$  とすると、連続ウェーブレット変換は以下のように表される。

$$C(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \psi^* \left( -\frac{b-t}{a} \right) dt \quad (1)$$

$$\psi(t) = \pi^{-\frac{1}{4}} e^{j\omega_0 t} e^{-\frac{t^2}{2}} \quad (2)$$

$a$  は解析する周波数、 $b$  は解析する時間に対応するパラメータである。 $\psi(t)$  はウェーブレット関数で、本稿では複素モルレーウェーブレットを用いた。ウェーブレット変換の分析周波数は、4–95 Hz の区間を 2 を底とする対数スケールで 100 分割して得た。変換後の値は各時刻で複素数として得られるため、その大きさを取った後、20 ms 毎に平均した。

#### 3.2 NTF を用いた特徴量抽出

被験者 1 人の計測データから得られる時間一周波数特徴量は、時間、周波数、チャンネル、試行の 4 つの次元（添字）を持つ 4 階のテンソル（4 次元配列）で表現される。この非負の要素を持つ 4 階のテンソルを  $\underline{\mathbf{X}} = \{x_{i,j,k,l}\} \in \mathbb{R}_+^{I \times J \times K \times L}$  で表す。テンソル  $\underline{\mathbf{X}}$  は、NTF によって式 (3) のように、 $R$  個のランク 1 テンソルに分解される<sup>[9,10]</sup>。この時、テンソルの各要素は式 (4) のように表される。

$$\underline{\mathbf{X}} \approx \hat{\underline{\mathbf{X}}} = \sum_{r=1}^R \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r \circ \mathbf{t}_r \quad (3)$$

$$x_{i,j,k,l} \approx \sum_{r=1}^R a_{i,r} b_{j,r} c_{k,r} t_{l,r} \quad (4)$$

ここで、 $\mathbf{A} = \{a_{i,r}\} = [\mathbf{a}_1 \dots \mathbf{a}_R] \in \mathbb{R}_+^{I \times R}$ 、 $\mathbf{B} = \{b_{j,r}\} = [\mathbf{b}_1 \dots \mathbf{b}_R] \in \mathbb{R}_+^{J \times R}$ 、 $\mathbf{C} = \{c_{k,r}\} = [\mathbf{c}_1 \dots \mathbf{c}_R] \in \mathbb{R}_+^{K \times R}$ 、 $\mathbf{T} = \{t_{l,r}\} = [\mathbf{t}_1 \dots \mathbf{t}_R] \in \mathbb{R}_+^{L \times R}$  は、それぞれ非負の要素を持つ周波数、時間、チャンネル、試行に関する基底行列である。また、“ $\circ$ ” はベクトルの外積（直積）を表す。

本研究では、周波数、時間、チャンネルの基底から成る 3 階のランク 1 テンソル  $\mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r$  が、特定の周波数帯域、時間区間、領域の脳活動を表現すると仮定し、基底  $\mathbf{t}_r$  の要素  $t_{l,r}$  が、ランク 1 テンソルが表す脳活動の強度を表現していると考えた。したがって、基底行列  $\mathbf{T}$  の各行ベクトルに着目し、小さい  $R$  を選ぶことによって、膨大な要素数を持つ時間一周波数特徴量から低次元の特徴量を抽出した。

NTF による分解は、 $\underline{\mathbf{X}}$  と  $\hat{\underline{\mathbf{X}}}$  との間の Euclid 距離を最小化することによって求めた。基底行列は式 (5-8)

の更新式を順に繰り返し適用することで求めることができる。

$$a_{m,n} \leftarrow \frac{a_{m,n} \sum_{j,k,l} x_{m,j,k,l} b_{j,n} c_{k,n} t_{l,n}}{\sum_r a_{m,r} \sum_{j,k,l} b_{j,r} c_{k,r} t_{l,r} b_{j,n} c_{k,n} t_{l,n}} \quad (5)$$

$$b_{m,n} \leftarrow \frac{b_{m,n} \sum_{i,k,l} x_{i,m,k,l} a_{i,n} c_{k,n} t_{l,n}}{\sum_r b_{m,r} \sum_{i,k,l} a_{i,r} c_{k,r} t_{l,r} a_{i,n} c_{k,n} t_{l,n}} \quad (6)$$

$$c_{m,n} \leftarrow \frac{c_{m,n} \sum_{i,j,l} x_{i,j,m,l} a_{i,n} b_{j,n} t_{l,n}}{\sum_r c_{m,r} \sum_{i,j,l} a_{i,r} b_{j,r} t_{l,r} a_{i,n} b_{j,n} t_{l,n}} \quad (7)$$

$$t_{m,n} \leftarrow \frac{t_{m,n} \sum_{i,j,k} x_{i,j,k,m} a_{i,n} b_{j,n} c_{k,n}}{\sum_r t_{m,r} \sum_{i,j,k} a_{i,r} b_{j,r} c_{k,r} a_{i,n} b_{j,n} c_{k,n}} \quad (8)$$

また、基底行列にはスケールの任意性があるため、 $\mathbf{A}$ 、 $\mathbf{B}$ 、 $\mathbf{C}$  の各列ベクトルを更新の直後に L2 ノルムが 1 になるように正規化した。

#### 3.3 Discriminant NTF を用いた特徴量抽出

NTF では、時間一周波数特徴のテンソル  $\underline{\mathbf{X}}$  と近似したテンソル  $\hat{\underline{\mathbf{X}}}$  の間の Euclid 距離を最小化しているに過ぎず、刺激音に対する印象の情報は組み込まれていない。そこで本稿では、DNTF<sup>[11]</sup> を用いて、テンソルの分解時に刺激対に対する聴感印象の比較判断を取り入れた。DNTF はクラスラベルを教師データとして用いて、テンソルを式 (3) のモデルで分解する。DNTF のコスト関数は NTF のコスト関数に特徴量のクラス間の分離を促す正則化項を付加した次の式で表される。

$$D(\underline{\mathbf{X}} || \hat{\underline{\mathbf{X}}}) + \alpha \text{Tr}(\mathbf{S}_w) - \beta \text{Tr}(\mathbf{S}_b) \quad (9)$$

ここで、 $D(\underline{\mathbf{X}} || \hat{\underline{\mathbf{X}}})$  は NTF のコスト関数を表し、 $\mathbf{S}_w$ 、 $\mathbf{S}_b$  は、それぞれクラス内分散共分散行列とクラス間分散共分散行列を表す。また、 $\text{Tr}(\cdot)$  は行列のトレースを表す。 $\alpha, \beta > 0$  は特徴量の分離度を制御するパラメータである。DNTF の正則化項は判別分析のようにクラス内分散を小さく、クラス間分散を大きくするようにする働きがある。

本稿では、刺激対に対する比較判断を 2 つのクラスに対応させ、刺激対に対応する特徴量対の差のベクトル  $\mathbf{d}_p = t_{p,\cdot}^A - t_{p,\cdot}^B \in \mathbb{R}^{1 \times R}$  が分離しやすくなるように特徴量を抽出した。 $\mathbf{d}_p$  が行ベクトルであることに注意すると、分散共分散行列  $\mathbf{S}_w$ 、 $\mathbf{S}_b$  は次のように与えられる。

$$\mathbf{S}_w = \sum_{k=1}^2 \sum_{d \in C_k} (\mathbf{d} - \boldsymbol{\mu}^{(k)})^T (\mathbf{d} - \boldsymbol{\mu}^{(k)}) \quad (10)$$

$$\mathbf{S}_b = \sum_{k=1}^2 N_k (\boldsymbol{\mu}^{(k)} - \boldsymbol{\mu})^T (\boldsymbol{\mu}^{(k)} - \boldsymbol{\mu}) \quad (11)$$

$$\boldsymbol{\mu}^{(k)} = \frac{1}{N_k} \sum_{d \in C_k} \mathbf{d}, \quad \boldsymbol{\mu} = \frac{1}{N} \sum_{k=1}^2 N_k \boldsymbol{\mu}^{(k)} \quad (12)$$

ここで、 $C_k$  ( $k = 1, 2$ ) は比較判断に対応するクラス、 $N_k$ 、 $N$  はそれぞれクラス  $k$  に所属する特徴量対の数、および全特徴量対の数を表す。また、 $\mu^{(k)}$ 、 $\mu$  はそれぞれ、クラス  $k$  に所属する特徴量対の差の平均ベクトル、全ての特徴量対の差の平均ベクトルを表す。

式 (9) のコスト関数の正則化項は基底行列  $T$  のみに依存するので、基底行列  $A$ 、 $B$ 、 $C$  は NTF と同様に式 (5-7) で更新することができる。基底行列  $T$  の更新アルゴリズムは補助関数法を用いて求めた。補助関数法では、元のコスト関数を最小化する代わりに補助関数を最小化し、局所的最適解を得る。式 (9) の補助関数は、式 (9) の  $D(\underline{X}||\hat{\underline{X}})$  を Euclid 距離規準の NTF の補助関数に置き換えたものを用いた。補助関数を最小化することによって得られる更新式は式 (13) のようになる。

$$t_{m,n} \leftarrow t_{m,n} \frac{P_{m,n} + \gamma_m t_{pair(m),n} + ord(m) M_{m,n}}{Q_{m,n} + \gamma_m} \quad (13)$$

$$\gamma_m = \alpha - \frac{\alpha + \beta}{N_{cls(m)}} + \frac{\beta}{N} \quad (14)$$

$$M_{m,n} = \frac{\alpha + \beta}{N_{cls(m)}} \sum_{\substack{p \in C_{cls(m)} \\ p \neq ind(m)}} d_{p,n} - \frac{\beta}{N} \sum_{p \neq ind(m)} d_{p,n} \quad (15)$$

ここで、 $P_{m,n} = \sum_{i,j,k} x_{i,j,k} a_{i,n} b_{j,n} c_{k,n}$ 、 $Q_{m,n} = \sum_r t_{m,r} \sum_{i,j,k} a_{i,r} b_{j,r} c_{k,r} a_{i,n} b_{j,n} c_{k,n}$  である。また、 $m$  番目の試行と対になっている試行のインデックスを  $pair(m)$ 、 $m$  番目の試行が含まれている対に対応する比較判断を  $cls(m)$ 、 $m$  番目の試行が含まれている対のインデックスを  $ind(m)$  とする。 $ord(m)$  は、 $m$  番目の試行が刺激音 A の聴取か、刺激音 B の聴取であるかをそれぞれ 1、-1 で示す関数である。

特徴量対の差  $d$  や  $\gamma_m$  は負値を取りうるため、 $t_{m,n}$  が負の値となった場合、0 もしくは非常に小さい正の数にする必要がある。このため、補助関数の最適化ステップにおける最適性が失われ、基底行列  $T$  の更新において必ずしもコスト関数が減少することが保証されない。また、 $\alpha$ 、 $\beta$  が非常に大きくなると更新式 (13) の分母の  $\gamma_m$  が非常に大きくなり、 $t_{m,n}$  が 0 に近くなってしまふといった問題もある。しかしながら、4 章で述べる評価実験において、ある範囲の  $\alpha$ 、 $\beta$  では、式 (9) のコスト関数は単調減少し、アルゴリズムが収束すること確認した。

## 4 評価実験

### 4.1 印象予測モデルの構築

前稿 [7] と同様に、比較判断を考慮した脳磁界からのエアコン音印象予測モデルを、被験者ごとに構築した。この印象予測モデルは、線形モデル  $f(x) = \langle w, \phi(x) \rangle + b$  ( $w$  は重みベクトル、 $b$  はバイアス項)

で表され、与えられた学習データに対し、写像  $\phi$  で高次元空間上に写像された入力特徴量対 ( $x^A, x^B$ ) の差  $d = \phi(x^A) - \phi(x^B)$  を、比較判断に対応する 2 つのクラスに分類するように Support Vector Machine (SVM) の枠組みを用いて学習される [12]。また、SVM と同様にカーネル関数を用いることで、非線形なモデルを構築することができる。

まず、計測データごとに不良なチャンネルを除いて時間一周波数特徴を計算した後、NTF および DNTF を用いて低次元の特徴量を抽出した。NTF および DNTF によって抽出された特徴量の次元数は  $R = 30$  である。また、DNTF の 2 つのパラメータは  $\alpha = \beta$  とし、 $10^{-11}$ 、 $10^{-10}$ 、 $10^{-9}$  の 3 通りに設定した。これらのパラメータは非常に小さいが、これは脳磁界計測によって得られるデータのオーダーが非常に小さく、時間一周波数特徴のテンソルの各要素の値も小さくなったためである。次に、この特徴量と刺激対に対する比較判断を用いて、エアコン音の涼しさ、および好ましさの印象の尺度値を予測するモデルを学習した。なお、カーネル関数にガウシアンカーネルを用いたため、印象予測モデルは非線形なモデルとなった。

### 4.2 モデルの評価方法

印象予測モデルの評価は、5 分割の交差検証によって行った。まず、被験者ごとに全てのデータを分割して作った学習データセットを使って、基底行列  $A$ 、 $B$ 、 $C$ 、および特徴量  $T$  を抽出した。分割でできた残りのテストデータセットからの特徴量抽出は、学習データの分解で得られた基底行列  $A$ 、 $B$ 、 $C$  を使って計算した。テンソル  $\underline{X}$  の各要素を、その 4 番目の添字が行の添字になるように並び替えてできる行列を  $\underline{X}_{(4)} \in \mathbb{R}_+^{L \times IJK}$  とすると、式 (3) は式 (16) のように書き換えられる。したがって、基底行列  $A$ 、 $B$ 、 $C$  が与えられている場合、テストデータの特徴量  $T$  は逆行列を用いて式 (17) のように得られる。

$$\underline{X}_{(4)} \approx T(C \odot B \odot A)^T \quad (16)$$

$$T = \underline{X}_{(4)}[(C \odot B \odot A)^T(C \odot B \odot A)]^{-1} \quad (17)$$

ここで“ $\odot$ ”は、列のサイズが等しい 2 つの行列に対し列ごとのクロネッカー積をとることを表す (Khatrao-Rao 積)。なお、逆行列の計算の結果、負の値となった要素は 0 に置き換えた。

次に、学習したモデルを用いてテストデータの特徴量から印象の尺度値を予測した後、刺激対と対応する 2 つの予測尺度値の差を計算し、その符号から一対比較判断を予測した。最後に、脳磁界計測実験において得られた比較判断と予測した比較判断とを比べ、正解率を算出した。

Table 1 Prediction accuracies of paired-comparative judgment on coolness (%).

Sub.	NTF		DNTF	
	Closed	Open	Closed	Open
1	100.0	56.9	100.0	60.3
2	73.0	59.5	74.3	58.2
3	100.0	50.0	82.8	57.3
4	100.0	57.2	100.0	57.2
5	70.8	63.5	67.5	56.9
6	100.0	56.9	100.0	56.9
7	100.0	60.3	100.0	60.5
8	100.0	58.9	100.0	53.9
Ave.	93.0	<b>57.9</b>	90.6	57.7

Table 2 Prediction accuracies of paired-comparative judgment on preference (%).

Sub.	NTF		DNTF	
	Closed	Open	Closed	Open
1	100.0	57.4	100.0	62.8
2	100.0	59.5	72.9	62.7
3	100.0	59.1	96.7	57.8
4	100.0	57.5	100.0	53.8
5	74.3	56.9	65.5	55.6
6	69.9	59.7	67.4	58.6
7	74.4	54.0	94.4	62.1
8	66.0	62.1	94.1	62.1
Ave.	85.6	58.3	86.4	<b>59.4</b>

### 4.3 結果・考察

印象予測モデルによって予測した涼しさ、および好ましさに関する一対比較判断の正解率を Table 1, 2 にそれぞれ示す。DNTF を用いた場合の正解率は平均正解率が最も高かった時のパラメータの結果のみを示す。Closed は学習したモデルの評価を学習データで行った場合、Open は学習したモデルの評価をテストデータで行った場合を表している。

評価方法が Open の場合の正解率は全体として 60% 前後であり、DNTF の平均正解率が最も高かったパラメータは、涼しさ、好ましさ、いずれの評価の場合でも  $\alpha = \beta = 10^{-10}$  の時であった。涼しさの評価においては、NTF を用いた場合の平均正解率が DNTF を用いた場合よりも僅かに高い一方、好ましさの評価においては、DNTF を用いた場合が NTF を用いた場合よりも高い平均正解率を示した。いずれの結果からも、NTF と DNTF の平均正解率の間に有意な差は見られなかった。

次に、Fig. 1 に各反復における DNTF のコスト関数

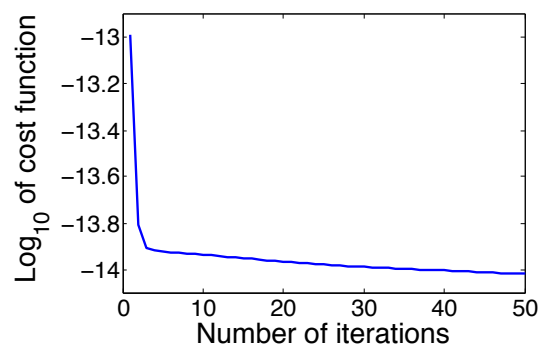


Fig. 1 An example of the cost function in Eq. (9) for each iteration.

の変化の例を示す。図より式 (9) のコスト関数が単調減少して収束していく様子が確認できる。

## 5 まとめ

エアコン音聴取時の脳磁界からの聴感印象予測モデルの精度を改善するために、DNTF を用いて、エアコン音の聴感印象に対する比較判断を取り入れた特徴量抽出手法を提案した。

DNTF、および前稿で提案した NTF を用いて抽出した特徴量から印象予測モデルを学習し、一対比較判断の予測の正解率によって、2 つの特徴量抽出手法の性能を評価した。その結果、好ましさの評価において、DNTF を用いた場合の方が高い平均正解率を示したが、NTF と DNTF の間に有意な差を見出すことはできなかった。アルゴリズムの改良など、さらなる検討が必要であると考えられる。

## 参考文献

- [1] Y. Soeta *et al.*, J.Sound Vib., 258, 419–428, 2002.
- [2] Y. Soeta *et al.*, J.Sound Vib., 277, 479–489, 2004.
- [3] S. Nakagawa *et al.*, ICIC Exp. Lett., 7, 1545–1550, 2013.
- [4] 矢野ら, 音講論 (春), 123–124, 2015.
- [5] 矢野ら, 音講論 (春), 511–514, 2016.
- [6] 矢野ら, 音講論 (秋), 673–676, 2016.
- [7] 矢野ら, 音講論 (春), 539–542, 2017.
- [8] T. Hotehama and S. Nakagawa, Proc. of INTER-NOISE 2015, in 15\_915.
- [9] T. Kolda and B. Bader, SIAM Review, 51(3), 455–500, 2009.
- [10] A. Cichocki *et al.*, “Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-Way Data Analysis and Blind Source Separation,” Wiley, 2009.
- [11] S. Zafeiriou, IEEE Transactions on Neural Networks, 20(2), 217–235, 2009.
- [12] R. Herbrich *et al.*, Proc. of AAI, 80–84, 1998.