

# 脳磁界計測を用いたエアコン音の聴感印象推定の試み

## - 非負値テンソル分解による関連脳活動の抽出 - \*

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. まず, エアコン音に時間変動を与えた場合の“涼しさ”および“好ましさ”の印象と, 自発律動の時間的安定性との関連を調査した<sup>[4]</sup>. その結果, 両者の間にいくつかの相関が認められた. 前稿では, エアコン音聴取に伴う事象関連同期および脱同期を調査し, エアコン音の“好ましさ”の印象と後頭部の 8-13 Hz の帯域のパワーとの間に正の相関があることを明らかにした<sup>[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はなく, 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得ない.

一方で, 線形回帰を用いて, 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ついても試みた<sup>[6]</sup>. 自発律動の時間-周波数特徴から抽出したテンソル特徴量と印象の尺度値から回帰モデルを学習することで, 一対比較判断を 7 割の精度で推定することができた. しかし, 印象と関連の高い特定の脳活動成分を見いだすには至らず, 適切なモデルや変数の選択が課題となった.

本稿では, エアコン音評価のための実用的な脳活動指標の構築を目的として, さらなる解析を試みた. まず, 皮質部位, 自発律動の周波数, 刺激オンセットからの時間ごとに時間-周波数特徴を抽出した. 次に非負値テンソル分解 (Non-negative Tensor Decomposition: NTD) によって, 時間周波数特徴を複数の成分に分解し, [6]におけるテンソル特徴量とは異なる低次元の特徴量を抽出した. NTD による特徴量からエアコン音の“涼しさ”および“好ましさ”の尺度値を予測する回帰モデルを学習し, その予測精度の評価を行った.

### 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>[7]</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 6 名 (男性 6 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”の尺度値を一対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, 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>) を用いて行った.

### 3 特徴量抽出

#### 3.1 時間-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-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.

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

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

\* Estimation of auditory impressions induced by HVAC sound using Non-negative Tensor Decomposition. by YANO, Hajime (Kobe Univ./AIST), TAKIGUCHI, Tetsuya, ARIKI, Yasuo (Kobe Univ.), KAMIYA, Masaru (DENSO Corp.), NAKAGAWA, Seiji (Chiba Univ./AIST).

$a$  は解析する周波数,  $b$  は解析する時刻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. 本稿では, 0-5000 ms の区間を 20 ms 毎, 4-100 Hz の区間を 1 Hz 毎にウェーブレット係数を計算した.

### 3.2 非負値テンソル分解

多チャンネル信号から得た時間 - 周波数特徴量は (周波数)  $\times$  (時間)  $\times$  (チャンネル) の 3 次元配列で表現される. さらに, この 3 次元配列に対し, 試行数の次元を加えると 4 次元配列で表現することができる. この特徴量をベクトル化してモデルの学習をすることが考えられるが, 特徴量の次元数が非常に大きくなり, 線形回帰では多重共線や結果の解釈の困難性といった問題が生じる. また一般的には次元の呪いといった問題も生じやすい. この問題に対しては, ベクトル化したテンソルに PCA などを適用して次元圧縮することが考えられるが, いずれにしてもベクトル化によって (周波数)  $\times$  (時間)  $\times$  (チャンネル) といったデータの構造が失われてしまう.

そこで本研究では, 非負値テンソル分解 (Non-negative Tensor Decomposition: NTD) に着目し, 特徴量抽出を行った. テンソル分解は多次元配列で表現されるようなデータに対し, その構造を保持したまま効果的な特徴を抽出・選択する手法として, 脳信号処理や画像処理等様々な分野で応用されている [8,9].

(周波数)  $\times$  (時間)  $\times$  (チャンネル)  $\times$  (試行数) のテンソル (4 次元配列),  $\underline{\mathbf{X}} = \{x_{i,j,k,l}\} \in \mathbb{R}_+^{I \times J \times K \times L}$ , に対する NTD は以下の式で表される.

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

ここで,  $\mathbf{A} = \{\mathbf{a}_{i,r}\} = [\mathbf{a}_1 \dots \mathbf{a}_R] \in \mathbb{R}_+^{I \times R}$ ,  $\mathbf{B} = \{\mathbf{b}_{j,r}\} = [\mathbf{b}_1 \dots \mathbf{b}_R] \in \mathbb{R}_+^{J \times R}$ ,  $\mathbf{C} = \{\mathbf{c}_{k,r}\} = [\mathbf{c}_1 \dots \mathbf{c}_R] \in \mathbb{R}_+^{K \times R}$ ,  $\mathbf{T} = \{\mathbf{t}_{l,r}\} = [\mathbf{t}_1 \dots \mathbf{t}_R] \in \mathbb{R}_+^{L \times R}$  であり, それぞれ周波数, 時間, チャンネル, 試行に関する基底行列である. またテンソル  $\underline{\mathbf{X}}$  の各要素は次のように表される.

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

NTD は, テンソル  $\underline{\mathbf{X}}$  を  $R$  個の非負値要素を持つランク 1 テンソル (4 つの基底ベクトルの直積 “ $\circ$ ” で表される. また, 本稿ではコンポーネントとも呼ぶ.) の線形和で表現する. また, 式 (3) の各ランク 1 テンソルが (周波数)  $\times$  (時間)  $\times$  (チャンネル) の 3 階のランク 1 テンソル  $\mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r$  に, 係数  $\mathbf{t}_{l,r}$  をかけたものと見なすことができる. したがって, 基底行列  $\mathbf{T}$  の行ベクトルが各コンポーネントの強さを表

すと考えられ, 本稿ではこれを特徴量として用いた. [6] におけるテンソル特徴量は最も寄与が大きいコンポーネントのみに着目しているのに対し, ここではコンポーネント全てに着目している.

基底行列  $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{T}$  は, 次の最適化問題を解くことによって得られる.

$$\begin{aligned} \min_{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{T}} \left\| \underline{\mathbf{X}} - \sum_{r=1}^R \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r \circ \mathbf{t}_r \right\|_F^2 \quad (5) \\ \text{s.t. } a_{i,r} \geq 0, b_{j,r} \geq 0, c_{k,r} \geq 0, t_{l,r} \geq 0, \\ \|\mathbf{a}_r\|_2 = 1, \|\mathbf{b}_r\|_2 = 1, \|\mathbf{c}_r\|_2 = 1 \end{aligned}$$

ここで,  $\|\underline{\mathbf{X}}\|_F^2 = \sum_{i,j,k,l} x_{i,j,k,l}^2$  はテンソルの Frobenius ノルムの 2 乗を表す. なお, 基底行列のスケールの任意性から,  $\mathbf{A}, \mathbf{B}, \mathbf{C}$  の各列ベクトルを  $L_2$  ノルムが 1 になるように正規化した.

本稿では式 (5) を NMF の乗法更新式 [10] を基にした更新式を用いて解く. まず, テンソルの各要素を行列の形に並び替えることを考える. テンソル  $\underline{\mathbf{X}}$  の 4 つの添字がそれぞれ行の添字となるように, テンソルの要素を行列の形に並びかえたものを,  $\mathbf{X}_{(1)} \in \mathbb{R}_+^{I \times JKL}$ ,  $\mathbf{X}_{(2)} \in \mathbb{R}_+^{J \times KLI}$ ,  $\mathbf{X}_{(3)} \in \mathbb{R}_+^{K \times LIJ}$ ,  $\mathbf{X}_{(4)} \in \mathbb{R}_+^{L \times IJK}$  とする. すると, 式 (3) は以下の 4 通りの形式に書き換えることができる [8].

$$\mathbf{X}_{(1)} \approx \mathbf{A}(\mathbf{T} \circ \mathbf{C} \circ \mathbf{B})^T = \mathbf{A}\mathbf{Z}^{(1)T} \quad (6)$$

$$\mathbf{X}_{(2)} \approx \mathbf{B}(\mathbf{A} \circ \mathbf{T} \circ \mathbf{C})^T = \mathbf{B}\mathbf{Z}^{(2)T} \quad (7)$$

$$\mathbf{X}_{(3)} \approx \mathbf{C}(\mathbf{B} \circ \mathbf{A} \circ \mathbf{T})^T = \mathbf{C}\mathbf{Z}^{(3)T} \quad (8)$$

$$\mathbf{X}_{(4)} \approx \mathbf{T}(\mathbf{C} \circ \mathbf{B} \circ \mathbf{A})^T = \mathbf{T}\mathbf{Z}^{(4)T} \quad (9)$$

ここで,  $\mathbf{Z}^{(1)} = \mathbf{T} \circ \mathbf{C} \circ \mathbf{B}$ ,  $\mathbf{Z}^{(2)} = \mathbf{A} \circ \mathbf{T} \circ \mathbf{C}$ ,  $\mathbf{Z}^{(3)} = \mathbf{B} \circ \mathbf{A} \circ \mathbf{T}$ ,  $\mathbf{Z}^{(4)} = \mathbf{C} \circ \mathbf{B} \circ \mathbf{A}$  である. “ $\circ$ ” は Khatri-Rao 積を表し, この行列積は列のサイズが等しい 2 つの行列について, 列ごとにクロネッカー積をとったものである.

式 (6-9) からテンソルの各要素を並べ替えてできた行列が非負値行列の積で表されることがわかる. したがって, 非負値行列因子分解の乗法更新式を基にした以下の更新式で, 行列  $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{T}$  を順に, 収束するまで繰り返し更新を行えば, 式 (5) の局所的最適解を求めることができる.

$$\mathbf{A} \leftarrow \mathbf{A} * \frac{\mathbf{X}_{(1)}\mathbf{Z}^{(1)}}{\mathbf{A}\mathbf{Z}^{(1)T}\mathbf{Z}^{(1)}} \quad (10)$$

$$\mathbf{B} \leftarrow \mathbf{B} * \frac{\mathbf{X}_{(2)}\mathbf{Z}^{(2)}}{\mathbf{B}\mathbf{Z}^{(2)T}\mathbf{Z}^{(2)}} \quad (11)$$

$$\mathbf{C} \leftarrow \mathbf{C} * \frac{\mathbf{X}_{(3)}\mathbf{Z}^{(3)}}{\mathbf{C}\mathbf{Z}^{(3)T}\mathbf{Z}^{(3)}} \quad (12)$$

$$\mathbf{T} \leftarrow \mathbf{T} * \frac{\mathbf{X}_{(4)}\mathbf{Z}^{(4)}}{\mathbf{T}\mathbf{Z}^{(4)T}\mathbf{Z}^{(4)}} \quad (13)$$

ここで，“\*”，“/”はそれぞれ要素ごとの積，商を表す．

## 4 印象評価モデル

### 4.1 モデルの構築

脳磁界計測に参加した被験者ごとに，脳活動からエアコン音の印象を予測するモデルを回帰モデルを用いて構築した．具体的には，応答変数をエアコン音の印象の尺度値，説明変数をNTDによって得た特徴量ベクトル（基底  $T$  の行ベクトル）の各次元とする回帰モデルを学習した．

計測データごとに不良なチャンネルを除いた．特徴量ベクトルの次元数，すなわちNTDのランク数は  $R = 50$  とした．エアコン音の“涼しさ”および“好ましさ”の評価尺度値は，実験で得た一対比較判断からThurstoneの一対比較法（ケースV）を用いて計算した．回帰モデルには，線形な通常の重回帰（Ordinary Least Square: OLS）とL1正則化重回帰（LASSO），および非線形なサポートベクター回帰（SVR）を用いた．なお，SVRのカーネル関数はガウシアンカーネルを用いた．

### 4.2 モデルの評価方法

学習した印象予測モデルを評価する方法として，刺激  $s (= 1, \dots, 4)$  を含まない刺激対を呈示した時のraw dataから得た時間-周波数特徴，および一対比較判断を学習データとし，刺激  $s$  を含む刺激対を呈示したときの時間-周波数特徴，および一対比較判断をテストデータとした．NTDによるテストデータの特徴量抽出は，学習データの分解時に得られる基底行列  $A, B, C$  をテストデータを分解する際に与えることによって計算した．

学習したモデルの評価はまず，学習データから構築した回帰モデルを用いて，テストデータから評価尺度値を予測した．次に，刺激対に対応した2つの評価尺度値の予測値の差を計算し，その符号から一対比較判断を予測した．実際に得られた比較判断と，予測した比較判断を比べて計算した正解率をすべての  $s$  で平均したものを最終的な正解率とした．

### 4.3 結果・考察

被験者ごとの“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，それぞれTable 1, 2に示す．Closedはテストデータに学習データを用いた場合，Openはテストデータに学習データを含まない場合を表している．

Openの場合の正解率は，“涼しさ”，“好ましさ”ともに平均で50%程度であり，SVRを用いた場合が最も高かった．

Table 1 The prediction accuracy of paired-comparative judgement on coolness [%].

Sub.	OLS		LASSO		SVR	
	Closed	Open	Closed	Open	Closed	Open
1	70.0	55.6	70.1	57.1	79.9	54.3
2	69.4	50.7	52.4	51.4	78.8	55.6
3	69.3	47.8	64.4	52.3	82.5	54.6
4	30.3	29.9	35.0	35.0	35.3	35.1
5	61.9	54.4	63.2	54.8	69.0	56.8
6	50.9	45.0	55.8	55.8	55.8	55.8
Ave.	58.6	47.3	56.8	51.1	66.9	52.2

Table 2 The prediction accuracy of paired-comparative judgement on preference [%].

Sub.	OLS		LASSO		SVR	
	Closed	Open	Closed	Open	Closed	Open
1	68.4	51.7	69.7	55.3	72.7	56.2
2	69.1	50.5	56.7	52.3	68.2	56.4
3	67.4	51.1	68.1	52.4	91.6	56.6
4	44.2	35.7	42.9	40.5	43.9	39.4
5	68.5	45.5	67.2	47.9	59.6	49.9
6	46.0	39.9	40.6	42.0	52.2	41.3
Ave.	60.6	45.7	57.5	48.4	64.7	50.0

“涼しさ”の評価における被験者4および，“好ましさ”の評価における被験者4, 6がClosed, Openともに他の被験者と比較して正解率が低い結果となった．Closedの正解率が低いことから，十分に回帰モデルの学習ができていないと考えられる．テンソルの近似のランク数  $R$  を大きくし，より細かな成分の情報の特徴量ベクトルに取り入れることで，より精度の高いモデルを学習できる可能性がある．

本稿の結果は，一対比較判断の予測の正解率が70%程度であった[6]における手法と比べて低い正解率を示している．この理由としては本手法が，学習データに含まれていない刺激に関するデータを評価に用いている，つまり[6]における評価手法より，モデルの汎化性能を評価する方法を用いているためであると考えられる．また，一つの刺激に関するデータを除去することによって，学習データの数が減ったことも原因の一つと考えられる．

NTDによって得られた基底行列  $A, B, C, T$  をFig. 1に示す．基底行列  $A$  は周波数方向の基底であり，テンソルが4-20 Hzの成分に多く分解されていることがわかる．基底行列  $B$  は時間方向の基底であり，特定の時間区間のみ高い値を示している基底がい

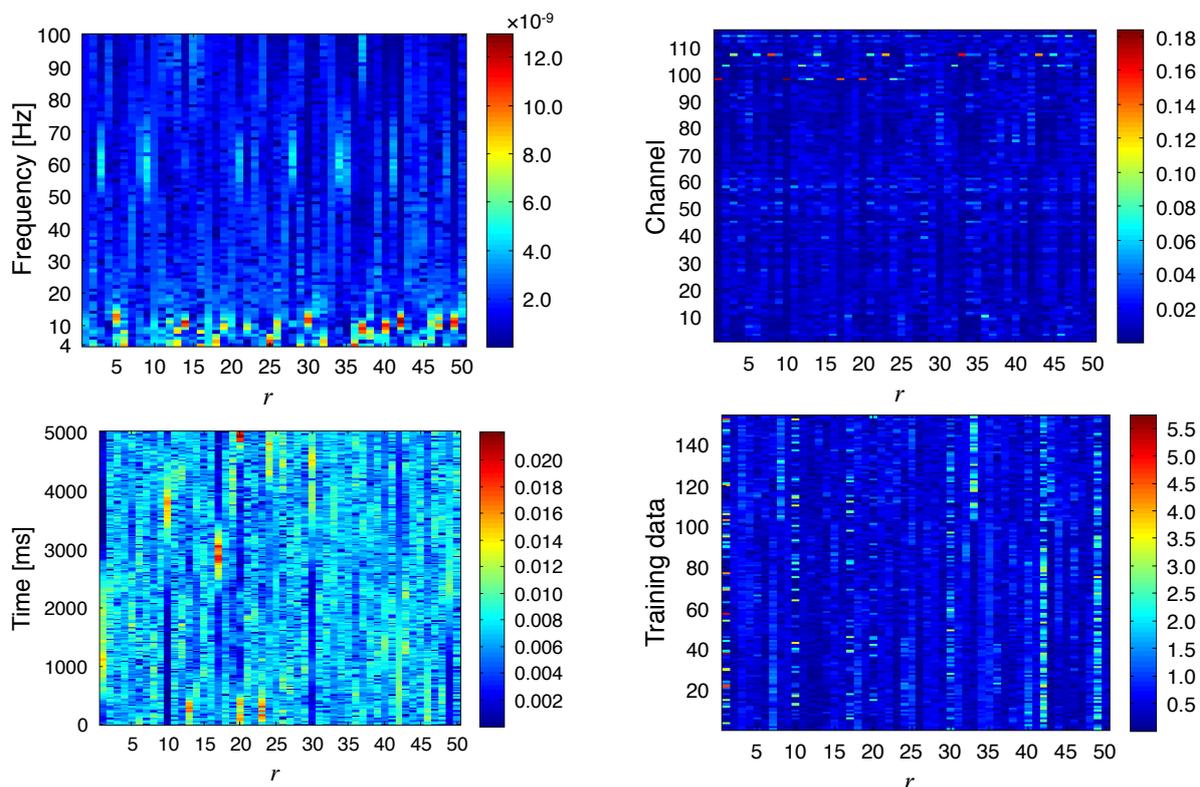


Fig. 1 Examples of basis matrix (Coolness, Subject1, LASSO), (top left: **A**, bottom left: **B**, top right: **C**, bottom right: **T**).

くつか見られる．これらの基底は特定の時間でのみ発生する律動を表現していると考えられる．基底行列 **C** はチャンネル方向の基底である．後半のチャンネルが高い値を示しており，これらの基底を持つコンポーネントが脳磁界信号に含まれるアーティファクトを表現している可能性がある．

本稿で用いた NTD は，あくまで式 (5) を最小化，すなわちランク 1 テンソルの和で近似したテンソルをもとのテンソルにフィッティングさせているに過ぎない．したがって，得られた特徴量ベクトルには必ずしも，印象の尺度値や一対比較判断の予測に有用な情報が含まれていない．実用的な予測モデルを構築するためにはそれらの情報を取り入れた手法を用いる必要があると考えられる．

## 5 まとめ

脳磁界の時間 - 周波数特徴から NTD を用いて抽出した特徴量と，エアコン音の“涼しさ”および“好ましさ”の尺度値から印象評価モデルを学習し，その評価を行った．しかしながら，一対比較判断の予測の正解率は 50%程度に過ぎず，有効な手法が提案できたとは言い難い結果となった．

NTD の基底行列を確認することによって，印象と関連する脳活動の抽出が期待される一方で，エアコン音の印象評価のための実用的な脳活動指標構築の

ためには，特徴量抽出手法の改善や学習データ数の確保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(春)，485-488，2015．
- [7] T. Hotehama and S. Nakagawa, Proc. of INTER-NOISE 2015, in 15\_915.
- [8] T. Kolda and B.Bader, SIAM Review, 51(3), 455-500, 2009.
- [9] A. Cichocki *et al.*, “Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-Way Data Analysis and Blind Source Separation,” Wiley, 2009.
- [10] D. D. Lee and H. S. Seung, Proc. NIPS, 2000.