

[ポスター講演] 非負値テンソル分解を用いたエアコン音の 印象関連脳活動の抽出

矢野 肇^{†,††} 滝口 哲也[†] 有木 康雄[†] 保手浜拓也^{††} 神谷 勝^{†††}
中川 誠司^{††††,††}

[†] 神戸大学大学院システム情報学研究科 〒657-8501 兵庫県神戸市灘区六甲台町 1-1

^{††} 国立研究開発法人産業技術総合研究所 〒563-8577 大阪府池田市緑丘 1-8-31

^{†††} 株式会社デンソー 〒448-8661 愛知県刈谷市昭和町 1-1

^{††††} 千葉大学フロンティア医工学センター 〒263-8522 千葉県千葉市稲毛区弥生町 1-33

E-mail: [†]pi314@me.cs.scitec.kobe-u.ac.jp, ^{††}{takigu,ariki}@kobe-u.ac.jp, ^{††††}s.nakagawa@chiba-u.ac.jp

あらまし 脳活動指標を用いたエアコン音の印象評価を目的として、エアコン音聴取時の脳活動から、エアコンの音の印象を予測するモデルを構築した。本研究ではまず、エアコン音聴取時の脳磁界の時間周波数特徴を計算した。次に、時間周波数特徴から非負値テンソル分解 (Non-negative Tensor Decomposition: NTD) によって低次元の特徴量を抽出した。特徴量から印象の尺度値を予測する回帰モデル学習し、その予測精度の評価を行った。

キーワード 聴感印象評価, 脳磁界計測, 非負値テンソル分解, プリファレンス, エアコン音

Extraction of brain activity related to auditory impressions induced by HVAC sound using non-negative tensor decomposition

Hajime YANO^{†,††}, Tetsuya TAKIGUCHI[†], Yasuo ARIKI[†], Takuya HOTEHAMA^{††}, Masaru
KAMIYA^{†††}, and Seiji NAKAGAWA^{††††,††}

[†] Graduate School of System Infomatics, Kobe University Rokkodai 1-1, Nada-ku, Kobe-shi, Hyogo,
657-8501 Japan

^{††} National Institute of Advanced Industrial Science and Technology (AIST) Midorigaoka 1-8-31,
Ikeda-shi, Osaka, 563-8577 Japan

^{†††} DENSO Corporation Showa-cho 1-1, Kariya-shi, Aichi 448-8661 Japan

^{††††} Center for Frontier Medical Engineering, Chiba University Yayoi-cho 1-33, Inage-ku, Chiba-shi, Chiba
263-8522 Japan

E-mail: [†]pi314@me.cs.scitec.kobe-u.ac.jp, ^{††}{takigu,ariki}@kobe-u.ac.jp, ^{††††}s.nakagawa@chiba-u.ac.jp

Abstract To evaluate auditory impressions induced by HVAC (heating, ventilation and air conditioning) sound using neurophysiological indices, a prediction model which estimates auditory impressions from brain activity was constructed. First, the time-frequency representation of the magnetoencephalographic (MEG) data was calculated. Second, the low-dimensional features were extracted from the time-frequency feature using non-negative tensor factorization (NTD). The regression model which predicts the scale values of auditory impressions from the feature vectors were learned, and its performance was evaluated.

Key words auditory impression, magnetoencephalography (MEG), non-negative tensor decompositon, preference, HVAC sound

1. まえがき

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 [1], [2]. 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”の印象と、自発律動の時間的安定性および、エアコン音聴取に伴う事象関連同期および脱同期との関連を調査し、印象と脳活動との間にいくつかの相関が認められた [3], [4]. 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えない.

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、エアコン音聴取時の脳磁界から、印象を予測するモデルの構築を試みた.

脳活動のデータは高次元の配列で表現されることが多く、例えば、多チャネルの計測データの時間周波数特徴は (周波数) \times (時間) \times (チャネル) の 3 次元配列で表現される. この特徴量をベクトル化してモデルの学習をすると特徴量の次元数が非常に大きくなり、次元の呪いの問題が生じやすい. この問題に対し、多次元配列をベクトル化し、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 (チャネル) といったデータの構造は失われてしまう.

多次元配列で表現されるようなデータはテンソルと呼ばれ、テンソルの持つ多次元な構造を考慮した効果的な特徴を抽出する手法としてテンソル分解が挙げられる [6]. 本稿では、非負値をとる時間周波数特徴を利用するため、非負値制約を付加した非負値テンソル分解 (Non-negative Tensor Decomposition: NTD) を用いて、特徴量抽出を行った. NTD は脳信号処理 [8] や画像処理 [7] 等様々な分野で応用されている [9].

本稿ではまず、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 NTD によって、時間周波数特徴を複数の成分に分解し、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で計測されたエアコン音を基に、線形予測符号化によってスペクトル包絡を抽出・モデル化されたエアコン音 [5] 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™) 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.

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

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

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. 複素モルレーウェーブレットによるウェーブレット変換の結果は複素数となるため、その大きさをとり時間周波数特徴量として用いた.

3.2 非負値テンソル分解

本稿では、以上の方法によって得られた非負値を要素に持つ (周波数) \times (時間) \times (チャネル) \times (試行数) の 4 階のテンソル, $\underline{\mathbf{X}} = \{x_{i,j,k,l}\} \in \mathbb{R}_+^{I \times J \times K \times L}$, を NTD によって分解し、特徴量を抽出した. $\underline{\mathbf{X}}$ に対する NTD は下の式で表される.

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

ここで、 $\mathbf{A} = [\mathbf{a}_1 \dots \mathbf{a}_R] \in \mathbb{R}_+^{I \times R}$, $\mathbf{B} = [\mathbf{b}_1 \dots \mathbf{b}_R] \in \mathbb{R}_+^{J \times R}$, $\mathbf{C} = [\mathbf{c}_1 \dots \mathbf{c}_R] \in \mathbb{R}_+^{K \times R}$, $\mathbf{T} = [\mathbf{t}_1 \dots \mathbf{t}_R] \in \mathbb{R}_+^{L \times R}$ であり、それぞれ周波数, 時間, チャネル, 試行に関する基底行列である.

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.) の重ねあわせで表現する. 基底行列 \mathbf{T} の行ベクトルの各要素は、試行ごとの (周波数) \times (時間) \times (チャネル) の 3 階のランク 1 テンソル $\mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r$ の重みを表すと考えられる. 本稿ではこの基底行列 \mathbf{T} を NTD による特徴量として用いた.

それぞれの基底行列は、元のテンソル $\underline{\mathbf{X}}$ と式 (3) で近似されたテンソルの差の Frobenius ノルムを最小化することによって求められる. 本稿では非負値行列因子分解 (Non-negative Matrix Factorization: NMF) の乗法更新式 [10] を基にした次

の更新式 [8] を用いて、行列 A , B , C , T について、収束するまで繰り返し更新を行い、基底行列を計算した

$$A \leftarrow A * \frac{X_{(1)} Z^{(1)}}{AZ^{(1)T} Z^{(1)}} \quad (4)$$

$$B \leftarrow B * \frac{X_{(2)} Z^{(2)}}{BZ^{(2)T} Z^{(2)}} \quad (5)$$

$$C \leftarrow C * \frac{X_{(3)} Z^{(3)}}{CZ^{(3)T} Z^{(3)}} \quad (6)$$

$$T \leftarrow T * \frac{X_{(4)} Z^{(4)}}{TZ^{(4)T} Z^{(4)}} \quad (7)$$

ただし、テンソル \underline{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、 $X_{(1)} \in \mathbb{R}_+^{I \times JKL}$, $X_{(2)} \in \mathbb{R}_+^{J \times KLI}$, $X_{(3)} \in \mathbb{R}_+^{K \times LIJ}$, $X_{(4)} \in \mathbb{R}_+^{L \times IJK}$ としている。また、 $Z^{(1)} = T \odot C \odot B$, $Z^{(2)} = A \odot T \odot C$, $Z^{(3)} = B \odot A \odot T$, $Z^{(4)} = C \odot B \odot A$ である。“*” および分数による記法はそれぞれ要素ごとの積および商を表す。“ \odot ” は Khatri-Rao 積 (2 つの行列の列ごとのクロネッカー積) を表す。

なお、基底行列のスケールの任意性から、 A , B , C の各列ベクトルを L2 ノルムが 1 になるように正規化した。

4. 印象評価モデル

4.1 モデルの構築

回帰モデルを用いて、脳活動からエアコン音の印象を予測するモデルを、脳磁界計測に参加した被験者ごとに構築した。具体的には、応答変数をエアコン音の印象の尺度値、説明変数を脳磁界データから抽出した特徴量ベクトルの各次元とする回帰モデルを学習した。

まず、計測データごとに不良なチャンネルを除いた raw data から時間周波数特徴を抽出した。次に、時間周波数特徴に対し NTD を行って得られた基底行列 T の行ベクトルを特徴量ベクトルとした。本稿では、NTD のランク数は $R = 50$ とした。

また、時間周波数特徴の (周波数) \times (時間) \times (チャンネル) のテンソルをベクトル化・次元圧縮して用いた場合として、NMF を用いた特徴量抽出も行った。NMF によって、ベクトル化した時間周波数特徴を列方向に並べた行列 $V \in \mathbb{R}_+^{IJK \times L}$ を $V \approx WH$ となるような 2 つの非負行列 $W \in \mathbb{R}_+^{IJK \times R}$, $H \in \mathbb{R}_+^{R \times L}$ に分解した。 H の列ベクトルを特徴量ベクトルとして用い、NMF の基底数は $R = 50$ 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 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、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

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。NMF および NTD

表 1 “涼しさ”の一対比較判断予測の正解率 [%].

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

Sub.	NTD-LASSO		NTD-SVR		NMF-SVR	
	Closed	Open	Closed	Open	Closed	Open
1	70.1	57.1	79.9	54.3	70.1	57.0
2	52.4	51.4	78.8	55.6	70.1	51.0
3	64.4	52.3	82.5	54.6	64.7	51.3
4	35.0	35.0	35.3	35.1	35.0	35.0
5	63.2	54.8	69.0	56.8	68.7	56.8
6	55.8	55.8	55.8	55.8	55.8	55.8
Ave.	56.8	51.1	66.9	52.2	60.7	51.2

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 W および、 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。

評価は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、実際に実験で呈示した刺激対に対する印象と対応させた 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。実験において記録した実際の比較判断と、予測した比較判断を比べて正解率を算出した。さらに刺激 s を除いてモデルを構築した時の正解率をすべての s について平均し、平均正解率を算出した。

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平均正解率を、それぞれ表 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。

いずれの場合においても、Open の場合の正解率は“涼しさ”、“好ましさ”ともに平均で 50% 程度であった。全被験者の結果を平均した場合、NTD・SVR の組み合わせで特徴量抽出・モデルの構築を行った場合が最も高かった。しかしながら、特徴量抽出において NTD を用いた場合と NMF を用いた場合の結果にはあまり差は見られず、本稿においては、NTD による特徴量抽出が NMF と比べて必ずしも良いとは言い難い。一対比較判断の予測の正解率が 50% 程度であった理由としては、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が原因の一つと考えられる。

“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。また、これらの被験者のデータがノイズが多く含んでいた可能性もあり、さらなる検討が必要である。

NTD によって得られた基底行列 A , B , C , T を図 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 はチャネ

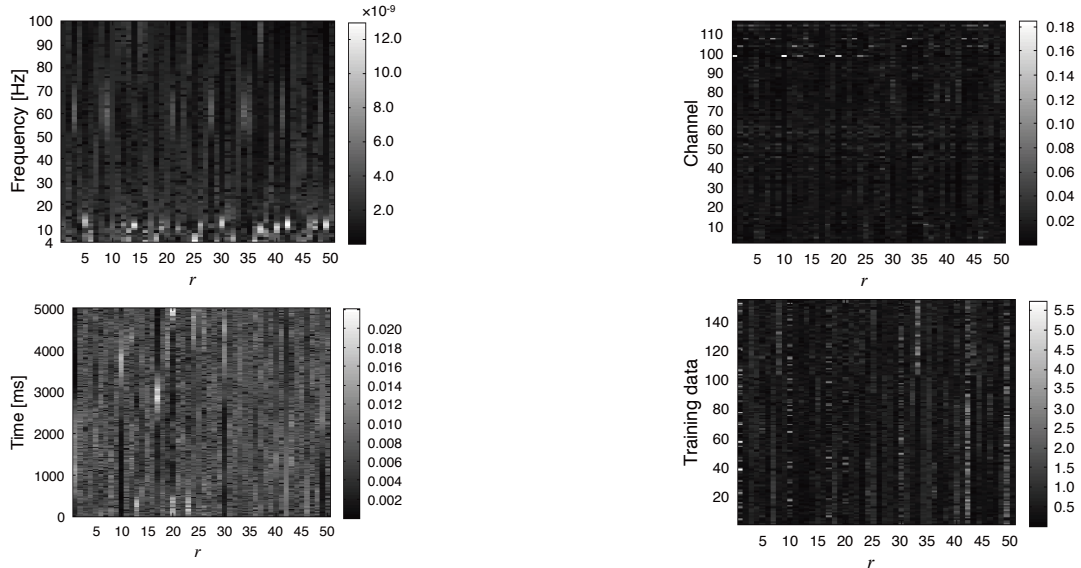


図1 NTDによって得られた基底行列の例 (Coolness, Subject1, NTD-LASSO). それぞれ, **A** (左上), **B** (左下), **C** (右上), **T** (右下).

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

表2 “好ましき”の対比較判断予測の正解率 [%].

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

Sub.	NTD-LASSO		NTD-SVR		NMF-SVR	
	Closed	Open	Closed	Open	Closed	Open
1	69.7	55.3	72.7	56.2	68.8	56.7
2	56.7	52.3	68.2	56.4	56.7	52.3
3	68.1	52.4	91.6	56.6	79.8	53.4
4	42.9	40.5	43.9	39.4	45.6	42.2
5	67.2	47.9	59.6	49.9	81.2	50.7
6	40.6	42.0	52.2	41.3	49.6	41.7
Ave.	57.5	48.4	64.7	50.0	63.6	49.5

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成分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。

NTDはランク1テンソルの和で近似したテンソルを元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

脳磁界の時間一周波数特徴から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“好ましき”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、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。

文献

- [1] Y. Soeta, S. Nakagawa, and M. Tonoike, “Magnetoencephalographic responses corresponding to individual subjective preference of sound fields,” *Journal of Sound and Vibration*, vol.258, pp.419–428, 2002.
- [2] Y. Soeta, S. Nakagawa, and M. Tonoike, “Magnetoencephalographic responses correspond to individual annoyance of bandpass noise,” *Journal of Sound and Vibration*, vol.277, pp.479–489, 2004.
- [3] 矢野肇, 滝口哲也, 有木康雄, 神谷勝, 保手浜拓也, 中川誠司, “エアコン音の時間変動が主観印象および大脳皮質活動に及ぼす影響,” 日本音響学会 2015 年春季研究発表会講演論文集, pp.123–124, 2015.
- [4] 矢野肇, 滝口哲也, 有木康雄, 保手浜拓也, 神谷勝, 中川誠司, “エアコン音の聴感印象関連領域の探索——脳磁界の時間周波数解析に基づく推定——,” 日本音響学会 2016 年春季研究発表会講演論文集, pp.511–514, 2016.
- [5] T. Hotehama, and S. Nakagawa, “Auditory impression of the coolness and warmness of automotive HVAC noise,” *Proceedings of INTER-NOISE 2015*, in 15_915, 2015.
- [6] T. Kolda, and B. Bader, “Tensor decompositions and applications,” *SIAM Review*, vol.51, no.3, pp.455–500, 2009.
- [7] A. Shashua, and T. Hazan, “Non-negative tensor factorization with applications to statistics and computer vision,” *Proceedings of ICML*, 2015.
- [8] H. Lee, Y. Kim, A. Cichocki, and S. Choi, “Nonnegative tensor factorization for continuous eeg classification,” *International Journal of Neural Systems*, vol.17, no.4, 2007.
- [9] A. Cichocki, R. Zdunek, A.H. Phan, and S. Amari, *Non-negative 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, “Algorithms for non-negative matrix factorization,” *Proceedings of NIPS*, 2000.