

脳磁界計測を用いたエアコン音の聴感印象推定の試み —線形回帰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), 保手浜拓也 (産総研),
神谷 勝 ((株) デンソー), 中川誠司 (産総研)

1 はじめに

音の印象評価には質問紙法などの心理的手法が用いられてきた。しかしながら、繰り返し判断を求めるために被験者の負担が大きいため、心理的バイアスが重畳しやすい、意識下で生じる聴感印象をとらえにくいといった問題点があった。一方で、近年、生理的計測、特に脳活動計測による印象評価の試みがなされている^[1-3]。生理指標を用いた手法には、心理的バイアスの影響を比較的受けにくい、被験者の負担を軽減可能といった利点が期待されている。また、意識下で生じる印象をも評価できる可能性もある。

我々は、エアコン音聴取時に惹起される“涼しさ”や“好ましさ”の印象の評価に取り組んで来た。エアコン音のスペクトル重心を変化させた際の“涼しさ”や“暖かさ”を一对比較で評価し、スペクトル重心が高い/低いほど、涼しい/暖かい印象を受けることを明らかにした^[4]。また、前稿ではエアコン音に時間変動を与えた場合の聴感印象と脳活動との関連を調査した^[5]。その結果、“涼しさ”印象と、 $\text{low-}\gamma$, β 帯域の自発律動の時間安定性、および“好ましさ”印象と、 β , θ 帯域自発律動の時間安定性、それぞれの間の有意な相関を見出した。この結果は脳機能指標による印象評価の実現の可能性を示唆するものであるが、現時点では実用的な指標を構築できたとは言いがたい。より頑健性の高い指標の構築が必要である。

本稿では、まず前稿で計測した脳磁界計測データの解析を進め、皮質部位、自発律動の帯域、刺激音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、線形回帰によって主観印象評価で得た“涼しさ”および“好ましさ”の尺度値と抽出した特徴量との関係を調査し、印象評価と相関の高い脳活動の抽出を試みた。また線形回帰を用いた印象評価指標についても検討した。

2 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

包絡を抽出した合成音を作成し、モデル音とした^[4]。LPC 次元は 150 とした。時間変動を加えるため、このモデル音に対して正弦波による振幅変調を施した。変調周波数を 0（変調なし）、0.2, 0.4, 0.6, 0.8, 1.6, 3.2 Hz として 7 種類の刺激音を作成した。刺激音の持続時間は 5 s、変調度は 0.15 に固定した。なお、刺激音圧は、変調周波数 0 Hz（変調なし）の刺激音を用いて、各被験者について“はっきりと聞こえ、なおかつちょうど良い感じる大きさ”に設定した。

3 主観印象評価実験

聴覚健全者 8 名（男性 7 名、女性 1 名、21-39 歳）を被験者とした。

主観印象は一对比較法にて評価した。上記の刺激音のうち、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およびその程度を評価させた。実験は防音室内で行われ、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。

“涼しさ”基準の実験は、防音室内の温度を 20, 25, 30°C と変化させて計 3 セッション行われた。“好ましさ”基準での実験は、防音室内の温度は 25°C に設定された。

各セッションにおいて、刺激対はランダムな順番で呈示され、それぞれの刺激の組み合わせに対して計 4 回の判断を行わせた。“涼しさ”、“好ましさ”の評価尺度値の算出には Scheffé の方法（中屋の変法）を用いた。

4 脳磁界計測

主観印象評価実験に参加した被験者のうち、6 名（男性 6 名、21-24 歳）を被験者とした。

主観印象評価実験で得られた評価値をもとに、評価値が最大および最小となる刺激音 2 つを含む、計 4 つの刺激音を選択した。そのうち 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。また、刺激音聴取時の脳磁界を計測した。

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™）を用いて

*Estimation of auditory impression associated with HVAC sound using magnetoencephalography, by YANO, Hajime (Kobe Univ./AIST), TAKIGUCHI, Tetsuya, ARIKI, Yasuo (Kobe Univ.), HOTEHAMA, Takuya (AIST), KAMIYA, Masaru (Denso Corp.), NAKAGAWA, Seiji (AIST)

行った。刺激音は、主観印象評価実験と同様のイヤホンを用いて、被験者の両耳に呈示した。なお、この実験の一対比較における評価尺度値を Thurstone の一対比較法（ケース V）によって推定し、線形回帰の応答変数として用いた。

5 特徴量抽出

5.1 時間-周波数特徴

各チャンネルにおける刺激呈示後 2-5 s の raw data に対し、連続ウェーブレット変換（Continuous Wavelet Transform: CWT）および短時間フーリエ変換（Short Time Fourier Transform: STFT）を施し、時間-周波数特徴を得た。ウェーブレットには複素モレーウェーブレットを用いた。周波数特徴は 43 次元（8-50 Hz, 1 Hz 間隔）、時間特徴は 5 次元（2.5-4.5 s, 0.5 s 間隔）とした。なお、本実験では明らかに不良な 3 つのチャンネルを除いた計 119ch について時間-周波数特徴を抽出した。つまり、得られる時間-周波数特徴量は $43 \times 5 \times 119 = 25585$ 次元となる。

5.2 テンソル分解

上記の方法で得た特徴量は（周波数） $\times$ （時間） $\times$ （チャンネル）の 3 次元配列で表され、特徴量の次元数は非常に大きくなる。この特徴量を基に線形回帰を行うと変数の数が膨大となるため、多重共線の発生や、結果の解釈の困難性といった問題が生じる。そこで、テンソル分解の一種である CANDECOMP/PARAFAC (CP) モデルを適用し、さらなる特徴量抽出を行った^[6,7]。テンソル分解は多次元配列で表現されるようなデータから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。

3 階のテンソル（3 次元配列）を $\underline{\mathbf{X}} \in \mathbb{R}^{I \times J \times K}$ とすると CP モデルによる分解は以下の式で表される。

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

ここで、 $\mathbf{a}_r \in \mathbb{R}^I, \mathbf{b}_r \in \mathbb{R}^J, \mathbf{c}_r \in \mathbb{R}^K (r = 1, \dots, R)$ で各ベクトルの大きさは 1 とする。このモデルはテンソル $\underline{\mathbf{X}}$ を R 個のランク 1 テンソル（3 つの基底ベクトルの直積 “ $\circ$ ” で表される）の線形和で近似している。また各要素 x_{ijk} は以下で表される。

$$x_{ijk} \approx \sum_{r=1}^R \lambda_r a_i b_j c_k \quad (2)$$

式 (1) は以下のように 3 通りの行列形式に書き表すこ

とができる。

$$\mathbf{X}_{(1)} \approx \mathbf{A}\mathbf{\Lambda}(\mathbf{C} \odot \mathbf{B})^T = \mathbf{A}\mathbf{S}_A \quad (3)$$

$$\mathbf{X}_{(2)} \approx \mathbf{B}\mathbf{\Lambda}(\mathbf{A} \odot \mathbf{C})^T = \mathbf{B}\mathbf{S}_B \quad (4)$$

$$\mathbf{X}_{(3)} \approx \mathbf{C}\mathbf{\Lambda}(\mathbf{B} \odot \mathbf{A})^T = \mathbf{C}\mathbf{S}_C \quad (5)$$

ここで、行列 $\mathbf{X}_{(1)} \in \mathbb{R}^{I \times JK}, \mathbf{X}_{(2)} \in \mathbb{R}^{J \times KI}, \mathbf{X}_{(3)} \in \mathbb{R}^{K \times IJ}$ は、それぞれテンソル $\underline{\mathbf{X}}$ の要素を行列の形に並べ替えたものである。行列 $\mathbf{A}, \mathbf{B}, \mathbf{C}$ はそれぞれ基底ベクトル $\mathbf{a}_r, \mathbf{b}_r, \mathbf{c}_r$ を $r = 1, \dots, R$ まで順に列方向に並べた基底行列である。 $\mathbf{\Lambda}$ は λ_r を対角成分に並べた対角行列である。また、演算子 “ $\odot$ ” 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。

時間-周波数特徴は非負値をとるため、行列 $\mathbf{A}, \mathbf{B}, \mathbf{C}$ に各要素および λ_r を非負値とする制約を与えた。分解は元のテンソル $\underline{\mathbf{X}}$ と近似されたテンソル $\underline{\hat{\mathbf{X}}}$ のカルバックライブラー情報量が最小となるように求めた。基底行列 $\mathbf{A}, \mathbf{B}, \mathbf{C}$ は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を繰り返し適用することによって計算される^[8]。

$$\mathbf{A} \leftarrow \mathbf{A} * \frac{[\mathbf{X}_{(1)} / (\mathbf{A}\mathbf{S}_A)] \mathbf{S}_A^T}{\mathbf{O}_A \mathbf{S}_A^T} \quad (6)$$

$$\mathbf{B} \leftarrow \mathbf{B} * \frac{[\mathbf{X}_{(2)} / (\mathbf{B}\mathbf{S}_B)] \mathbf{S}_B^T}{\mathbf{O}_B \mathbf{S}_B^T} \quad (7)$$

$$\mathbf{C} \leftarrow \mathbf{C} * \frac{[\mathbf{X}_{(3)} / (\mathbf{C}\mathbf{S}_C)] \mathbf{S}_C^T}{\mathbf{O}_C \mathbf{S}_C^T} \quad (8)$$

ここで、“ $*$ ”、“ $/$ ” はそれぞれ要素ごとの積、商を表す。また、 $\mathbf{O}_A \in \mathbb{R}^{I \times JK}, \mathbf{O}_B \in \mathbb{R}^{J \times KI}, \mathbf{O}_C \in \mathbb{R}^{K \times IJ}$ はすべての要素が 1 の行列である。なお、 λ_r は行列 $\mathbf{A}, \mathbf{B}, \mathbf{C}$ の各列ベクトルの大きさが 1 になるように求める。

5.3 テンソル特徴量抽出

（周波数） $\times$ （時間） $\times$ （チャンネル）の 3 階のテンソルを CP モデルによって分解することで、周波数、時間、チャンネルの特徴がそれぞれの基底行列に集約されることが期待される。

特徴量抽出の流れを Fig. 1 に示す。計 119ch について時間-周波数特徴を抽出し、CP モデルにより 2 つの成分に分解した ($R = 2$)。次に、2 つの成分のうち係数 λ_r が大きい成分を選択し、その成分を構成する周波数、時間、チャンネルの 3 つの基底ベクトルを連結して 1 つの特徴量ベクトルとした。この方法により得られる特徴量の次元は $43 + 5 + 119 = 167$ 次元となり、元の特徴量に比べて大幅に次元を削減できていることがわかる。

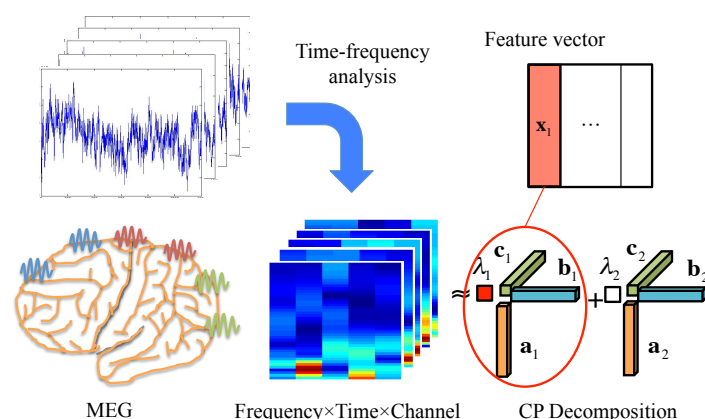


Fig. 1 Flowchart of tensor feature extraction

6 評価尺度推定

6.1 方法

脳磁界計測に参加した各被験者について、応答変数を脳磁界計測時の一対比較で得た評価尺度値、説明変数を特徴ベクトルの各次元として線形回帰を行った。

まず、抽出された特徴量ベクトルから刺激対に対応する特徴量対を一つ選択し、評価データとした。残りの特徴ベクトルを学習データとした。学習データを用いて回帰係数を推定した後、評価データ中の刺激対と対応させた二つの特徴量ベクトルからそれぞれの評価尺度値を予測した。次に、予測された評価尺度値の差を取ることによって一対比較判断を予測した。

交差検定によって各刺激対を評価データとした時の一対比較判断を予測し、実際の判断と比較して正解率を算出した。また、全ての変数を回帰モデルに含んだ状態から、正解率を減少させる変数を正解率が上昇しなくなるまで除去して行くことで変数の選択を行った。

6.2 結果

Table 1 に各被験者の“涼しさ”および“好ましさ”の判断の正解率をそれぞれ示す。また、時間-周波数特徴に CWT を用いた場合と STFT を用いた場合を示す。いずれの場合も 70% 程度の平均正解率が得られた。

Fig. 2 に被験者 5 に対して推定された回帰係数を示す。

0 となっている係数が除去された変数を表している。ほとんどの場合で時間特徴に対する回帰係数が小さい傾向が見られた。また、チャンネル特徴の係数をそれぞれチャンネル位置に模式的にプロットしたものを Fig. 3 に示す。

Table 1 Accuracy of prediction of paired-comparative judgement.

Subject	STFT		CWT	
	cool.	pref.	cool.	pref.
1	0.670	0.662	0.714	0.758
2	0.715	0.745	0.750	0.786
3	0.763	0.705	0.756	0.740
4	0.694	0.728	0.618	0.687
5	0.796	0.753	0.667	0.733
6	0.665	0.691	0.677	0.722
Ave.	0.717	0.714	0.697	0.738

7 考察

時間特徴に対応する回帰係数の値が小さいことから、テンソル分解によって時間方向の基底の寄与はあまり高くならないことがわかる。

Fig. 3 より左半球のチャンネルの係数が高い傾向が確認されることから、主に左半球でエアコン音に対する印象評価が営まれている可能性が伺える。一方、Fig. 2 に示されているように、周波数特徴、チャンネル特徴に関する回帰係数について、周波数やチャンネルが近接しているにもかかわらず符号が逆転し、大きく変化しているものがみられた。これらの原因として、変数の除去・選択が十分でない可能性が考えられる。

CP モデルにより、(周波数)×(時間)×(チャンネル)のテンソルを 2 つの成分に分解した時の基底行列の例を Fig. 4 に示す。もっとも、更新式 (6)-(8) による基底行列の計算結果が必ずしも大域的最適解となるわけではなく、局所最適解に陥る可能性がある。そのため各分解で同様な成分が得られるわけではない。また、係数 λ_1, λ_2 の大小関係も変化するので、データごとに異なる成分が抽出される可能性がある。

以上で挙げた問題を改善し、より高い精度の結果

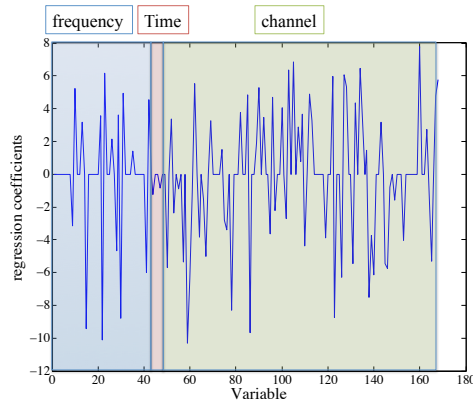


Fig. 2 Regression coefficients (Subject5, coolness, STFT).

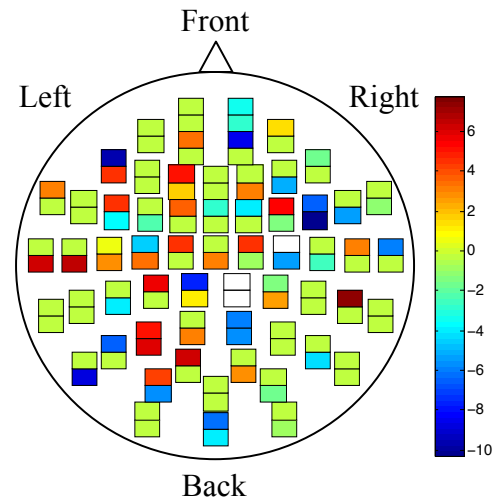


Fig. 3 Mapped regression coefficients (Subject5, coolness, STFT).

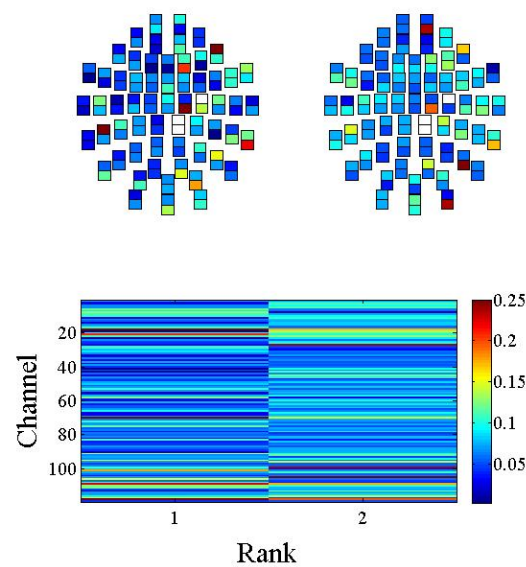
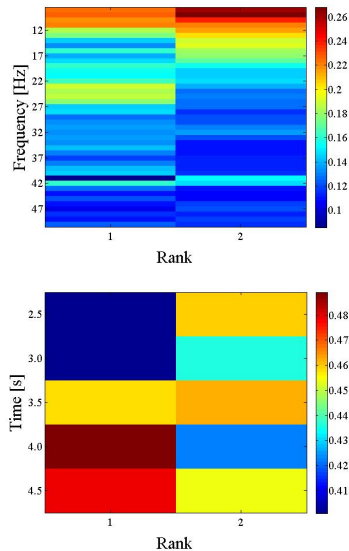


Fig. 4 Examples of factor Matrices (Subject5, coolness, STFT). $\lambda_1 = 0.378$, $\lambda_2 = 0.539$.

を得るために、異なる分解のモデルを使用する必要があると考えられる。本稿で用いた CP モデルはテンソル分解の中でも最も単純なモデルの一つであり、より複雑な構造を仮定したモデルや様々な制約を与えたモデルが多く提案されている [6,7]。また、本稿では解釈のしやすさから線形回帰を用いたが、より精度の高い印象評価値、比較判断の推定を目指すのであれば、サポートベクターマシン等の非線形な手法の利用を考慮する必要があるかもしれない。

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] T. Hotehama and S. Nakagawa, Proceedings of Inter-noise 2015, in press.
- [5] 矢野ら, 音講論 (春), 123-124, 2015.
- [6] T. Kolda and B. Bader, SIAM Review, 51(3), 455-500, 2009.
- [7] A. Cichocki *et al.*, "Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-Way Data Analysis and Blind Source Separation," Wiley, 2009.
- [8] D. D. Lee and H. S. Seung, Proc. NIPS, 2000.