

## 聴感印象推定のための脳活動特徴量抽出

### —テンソル分解による機能的結合の低ランク表現の検討—\*

☆矢野 肇 (神戸大/産総研), 高島遼一, 滝口哲也 (神戸大), 中川誠司 (千葉大)

#### 1 はじめに

脳活動計測に基づく印象評価は、質問紙法などの心理的手法に比べ、心理的バイアスの影響の低減や意識下で生じる印象の評価の可能性が期待されている。これまでにいくつかの音の印象と脳活動との間の相関 [1, 2, 3] が見出されているものの、それらの相関は小さく、現時点では実用的な印象評価指標であるとは言い難い。我々はこれまで、エアコン音の聴感印象を対象として、脳活動から抽出した特徴量を基に聴感印象の尺度値を推定する印象予測モデルを学習し、頑健なエアコン音の印象評価指標の構築に取り組んできた [4, 5, 6]。前稿 [6] では、印象の予測に有効な特徴を抽出するため、Common Spatial Pattern (CSP) を用いて印象の程度によって変化するような脳活動を抽出する空間フィルタを構成した。空間フィルタによって印象と関連のある脳活動が強調され、頑健な特徴量抽出ができる可能性が示唆されたが、現状の精度ではこれらを用いた印象評価は現実的ではない。

CSP 特徴量は脳活動の周波数成分ごとのパワーを元にした特徴量であり、主に個々の脳領域の活動強度に着目している。その一方で、印象評価のような高次な脳機能はいくつか脳領域が連動して関与していると考えられる。離れた領域間の神経活動の位相同期は機能的結合 (Functional Connectivity) と呼ばれ、様々な指標が提案されている [7]。個々の領域の活動強度に加えて、機能的結合の指標を利用することによって、印象予測モデルの精度を向上することができると考えられる。

我々は以前、機能的結合の指標の1つであるコヒーレンスに基づく特徴量 [5] を検討したが、コヒーレンスは近接するチャンネル間で計算すると、本来興味のない共通の信号源によって値が高くなってしまいう問題がある。そこで本稿では、離れた領域間の同期を多く検出することのできるコヒーレンスとは異なる機能的結合の指

標を用いて、すべてのチャンネル間・周波数帯域の位相同期を計算した。機能的結合の指標からなるテンソルを、機能的結合の持つ対称性を考慮した Nonnegative Tensor Factorization (NTF) によって低ランク近似し、すべての周波数帯域・大脳皮質全体にわたる結合を表現する低次元の特徴を抽出した。また、機能的結合に基づく特徴量が印象の予測に有効かどうかを検証した。

#### 2 機能的結合に基づく特徴量の抽出

##### 2.1 機能的結合の指標

フーリエ変換などによって得られる異なる2つの信号の複素表現を  $Z_1, Z_2 \in \mathbb{C}$  とすると、クロススペクトル  $X \in \mathbb{C}$  は以下の式で与えられる。

$$X = Z_1 Z_2^* \quad (1)$$

ここで、 $Z_2^*$  は  $Z_2$  の複素共役である。2つの信号間の位相差はクロススペクトルの位相で表される。コヒーレンスはクロススペクトルを用いて次の式で表される。

$$\text{Coh} = \frac{|\mathbb{E}[X]|}{\sqrt{\mathbb{E}[|Z_1|^2] \mathbb{E}[|Z_2|^2]}} \quad (2)$$

ここで、 $\mathbb{E}[\cdot]$  は期待値を取ることを表す。コヒーレンスは  $[0, 1]$  の範囲の値をとり、位相差が一定であるほど高い値を示す。しかしながら、近い位置で観測された信号間でコヒーレンスを計算すると、共通な信号源に由来する成分 (同相または逆相) が含まれるため値が高くなる傾向がある。このことは、頭外での計測データを用いて脳内の異なる活動源間の同期を調べる際に問題となる。

共通な信号源の成分の影響を受けにくい機能的結合の指標はいくつか提案されており、本研究では次の式で定義される Weighted Phase Lag Index (WPLI) [8] を用いる。

$$\text{WPLI} = \frac{|\mathbb{E}[\mathcal{J}(X)]|}{\mathbb{E}[|\mathcal{J}(X)|]} = \frac{\mathbb{E}[|\mathcal{J}(X)| \text{sgn}(\mathcal{J}(X))]}{\mathbb{E}[|\mathcal{J}(X)|]} \quad (3)$$

\*Low-rank representation of whole-head functional connectivity for prediction of subjective impressions. by YANO, Hajime (Kobe Univ./AIST), TAKASHIMA, Ryoichi, TAKIGUCHI, Tetsuya (Kobe Univ.), NAKAGAWA, Seiji (Chiba Univ.).

ここで,  $\text{sgn}(\cdot)$  は符号関数,  $\Im(X)$  は  $X$  の虚部である. WPLI も  $[0, 1]$  の範囲の値をとり, 位相差が一定であるほど値が大きくなるが, 同相や逆相の場合は  $X$  の虚部が 0 となるため値が 0 になる.

上で述べた機能的結合の指標はすべての異なる 2 つの計測チャンネル間で計算することができ, また, 複素スペクトルを用いれば周波数ごとに計算することができる. 計測チャンネルをグラフの節点, 機能的結合の指標をグラフの辺の重みと考えると, このグラフは脳内のネットワークを近似的に表現していると考えられる.

## 2.2 機能的結合の低ランク表現

グラフの非負の重み行列  $A$  からノードのクラスタ情報を得る手法として Symmetric Nonnegative Matrix Factorization (SymNMF) が知られている [9]. SymNMF は次の式 (4) で表される.

$$\min_{H \geq 0} \|A - HH^T\|_F^2 \quad (4)$$

$H$  は非負行列であり, その列数はクラスタ数 (< ノード数) に設定される. SymNMF でグラフの重み行列  $A$  を低ランク近似することで, グラフの低次元の埋め込み表現を得ることができる.

全チャンネル間の機能的結合を, すべての周波数・試行にわたって合わせると, すべての要素が非負の 4 階のテンソル  $C \in \mathbb{R}_+^{I \times J \times K}$  となる. ここで,  $I, J, K$  は, それぞれチャンネル数,  $J$  は周波数軸の分割数,  $K$  は試行回数である. 近い周波数の機能的結合は類似していること, SymNMF によってノードのクラスタが少ない基底で表現されることを踏まえ, テンソル  $C$  を NTF を用いて次のように分解することを考える.

$$C \approx \hat{C}_{\text{sym}} = \sum_{r=1}^R \mathbf{a}_r \circ \mathbf{a}_r \circ \mathbf{c}_r \circ \mathbf{d}_r \quad (5)$$

ここで,  $\mathbf{a}_r \in \mathbb{R}_+^I, \mathbf{c}_r \in \mathbb{R}_+^J, \mathbf{d}_r \in \mathbb{R}_+^K (r = 1, \dots, R)$  は, それぞれ  $r$  番目の成分に対応するチャンネル, 周波数, 試行に関する基底ベクトルであり, これらのベクトルの外積 (直積)  $\mathbf{a}_r \circ \mathbf{a}_r \circ \mathbf{c}_r \circ \mathbf{d}_r$  はランク 1 テンソルと呼ばれる. この分解によって, チャンネル基底に皮質間の結合パターン, 周波数基底に結合が類似している帯域, 試行基底に各サンプルに出現する同期の情報が反映されると考えられる.

式 (5) の分解を得るため, 各基底を交互に更新して,  $C$  と  $\hat{C}_{\text{sym}}$  の間の 2 乗距離を最小化しよう

とすると, 目的関数が基底  $\mathbf{a}_r$  に関する 4 次関数となっているため, 誤差の最小化は通常の NTF ほど容易ではない. そこで本稿では, 式 (6) のように 2 番目の方向に対応する基底を  $\mathbf{a}_r$  とは異なるベクトル  $\mathbf{b}_r \in \mathbb{R}_+^I$  で表現したモデル, すなわち通常の NTF による分解を考え, 2 乗誤差に  $\mathbf{a}_r = \mathbf{b}_r$  となるような正則化を加えることで式 (5) のような分解を得た.

$$C \approx \hat{C} = \sum_{r=1}^R \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r \circ \mathbf{d}_r \quad (6)$$

$$\min \frac{1}{2} \|C - \hat{C}\|_F^2 + \frac{\lambda}{2} \sum_{r=1}^R \|\mathbf{a}_r - \mathbf{b}_r\|_F^2 \quad (7)$$

ここで,  $\lambda$  は正則化パラメータであり, 十分大きな値を用いると  $\mathbf{a}_r = \mathbf{b}_r$  になると考えられる. このような正則化は SymNMF でも用いられている [10]. 式 (7) の目的関数は各基底ベクトルに関して 2 次関数となっているため, 本稿では Hierarchical Alternating Least Squares [11] を用いて最小化を行った. なお, NTF の基底のスケールの任意性を除くため  $\|\mathbf{a}_r\|_2 = \|\mathbf{b}_r\|_2 = \|\mathbf{c}_r\|_2 = 1$  となるように正規化した.

## 3 特徴量の評価

機能的結合に基づく特徴量が印象の予測に有効かどうかを検証するため, 被験者ごとに機能的結合の特徴量を用いて比較判断を予測するモデルを学習し, その予測精度の評価を行った.

### 3.1 脳磁界データ

特徴量の評価は, 時間変動を付与したエアコン音の対比較評価時に計測された脳磁界データ [12] を用いて行われた. 計測は聴覚健常者 6 名 (男性 4 名, 女性 2 名, 21–24 歳) を被験者として行われ, 被験者ごとに刺激音の涼しさの評価を行う場合と好ましさの評価を行う場合の計 2 回の計測が行われた. 刺激音は, 乗用車内で計測されたエアコン音を基にモデル化された音に, 正弦波による振幅変調を施したものをを用い, 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) が合成された. 計測では, これらの刺激音から被験者ごとに 4 つが選択され, 用いられた. 被験者は 4 つの刺激音のうち 2 つずつを連続的に呈示され, 2 つ目の刺激音の呈示終了後に, 各基準の優劣を評価するように教示された. なお, 各刺激音は挿入型

イヤホンによって5秒間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™, Neuromag Ltd.) を用いて行われた。計122chで計測された脳磁界信号は400 Hzでサンプリングされた。

データの前処理として、122 chのraw dataからデジタルフィルタと独立成分分析によって電源由来の60 Hzの成分、2 Hz以下の低周波成分、アーティファクト、眼球運動、および心磁図に由来するノイズ成分を取り除いた。

### 3.2 特徴量抽出

まず、刺激音呈示後500–4500 msの区間の脳磁界データに複素Morletウェーブレットによる連続ウェーブレット変換を行い、時間、周波数ごとの複素表現を得た。ウェーブレットのスケールは、周波数4–100 Hzを $\log_2$ の対数軸上で等間隔に分割し、それに対応するスケールを用いた。式(3)を用いて、周波数および刺激呈示(試行)ごとにすべての2つの異なるチャンネル間でWPLIを計算した。なお、期待値の計算は時間方向の平均で行った。計算されたWPLIから構成されたテンソルから、式(7)によって式(6)の分解を得た。分解のハイパーパラメータは $R = 10$ 、 $\lambda = 10^5$ と固定した。分解によって得られた10本の試行基底ベクトルの各要素はそれぞれ各試行に対応するため、試行ごとに10個の要素をまとめて10次元の特徴量ベクトルを得た。

### 3.3 評価方法

機能的結合に基づいた特徴量を用いて一対比較における2値の判断を予測するモデル(WPLI)を被験者ごとに学習し、その予測性能を評価した。機能的結合の特徴量との比較のため、周波数帯域ごとのパワーに基づく特徴量として、前稿で用いた14次元のCSP特徴量[6]を入力とするモデル(CSP)を学習した。さらに、機能的結合特徴量を加えた場合のモデルの精度も評価するため、CSP特徴量と機能的結合特徴量を連結した24次元の特徴量を入力とするモデル(Concat)も学習した。

比較判断の予測モデルはサポートベクターマシンに基づく印象予測モデル[4]を用いた。このモデルは刺激音聴取時の脳活動から刺激対に対応する2つの印象の尺度値を予測した後、それらの差の符号から比較判断を予測する。モデル

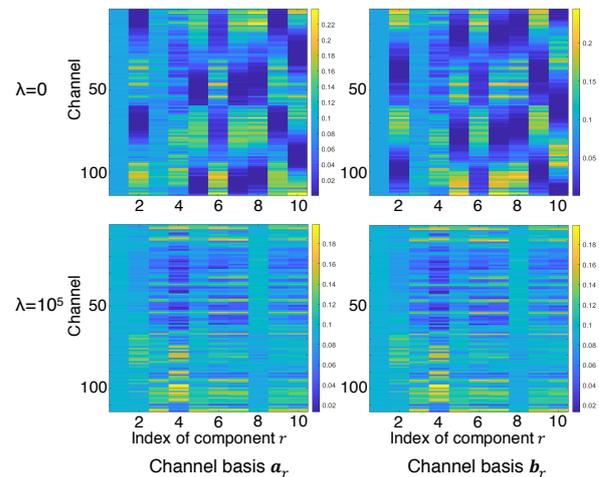


Fig. 1 Examples of channel basis vectors  $\{\mathbf{a}_r\}_{r=1,\dots,10}$  (left) and  $\{\mathbf{b}_r\}_{r=1,\dots,10}$  (right) obtained by Eq. (7) when  $\lambda = 0$  (top) and  $\lambda = 10^5$  (bottom).

の性能評価は評価尺度を一対比較判断の正解率とし、10分割交差検証によって行った。モデルのハイパーパラメータは、分割して得られた学習データに対し、さらに10分割交差検証を行って決定した。

### 3.4 結果・考察

式(7)の分解において、基底ベクトルが等しくなるような正則化を付与しなかった場合( $\lambda = 0$ )と付与した場合( $\lambda = 10^5$ )の2方向のチャンネル基底ベクトル $\{\mathbf{a}_r\}_{r=1,\dots,10}$ 、 $\{\mathbf{b}_r\}_{r=1,\dots,10}$ をFig. 1に示す。 $\lambda = 0$ の場合は対応する基底間で異なるベクトルが推定されているのに対し、 $\lambda = 10^5$ の場合はほぼ同じベクトルが推定されていることがわかる。この結果は、式(7)の正則化項を付与し正則化パラメータ $\lambda$ を大きくすることで $\mathbf{a}_r = \mathbf{b}_r$ となるように基底ベクトルを推定できることを示している。

被験者ごとのエアコン音の涼しさ、および好ましさに関する印象予測モデルの平均正解率をFig. 2に示す。いずれの評価においても半分以上の被験者で、WPLIに基づく特徴量だけを用いた場合のモデルが、CSP特徴量だけを用いたモデルと比べて精度が低い傾向を示した。特に好ましさに関する印象予測においてはその傾向が顕著に現れていることがわかる。この理由としてはCSP特徴量はその抽出に比較判断の情報を利用した識別的特徴量であるのに対し、WPLIは識別的な情報を用いていないことが考えられる。また、

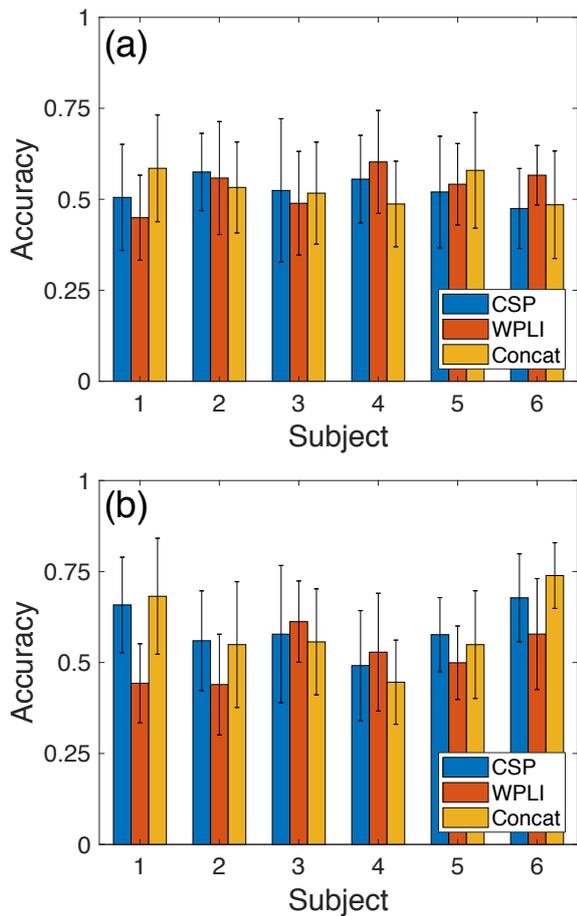


Fig. 2 Mean prediction accuracies of predicted comparative judgment on (a) coolness and (b) preference. Each error bar indicates standard deviation.

好ましさの評価で WPLI に基づく特徴量の精度が低い被検者については、領域間の情報伝達よりも個々の脳領域での活動が好ましさと関連していたことを示唆している。

その一方で、いずれの評価でも半分以上の被験者で、WPLI に基づく特徴量もしくは CSP 特徴量と連結した特徴量を用いたモデルの精度が最も高かった。このことは異なる脳領域間の活動の同期性の情報が印象の予測に有用である可能性を示唆している。しかしながら、被験者によっては、連結した特徴量で学習したモデルの精度が単体の特徴量で学習したモデルの精度より低い場合も見られる。この原因としてはモデルが過学習を引き起こしていることが考えられるため、より頑健なモデルの学習や特徴量の統合の手法が必要である。また、計算されたチャンネル間の機能的結合は、必ずしも明確なパターンを示しておらず、印象評価に直接関係のない脳活動やノイズの影響を受けていると考えられる。

## 4 まとめ

脳活動計測に基づく印象予測モデルの精度を向上させるため、異なる脳領域間の機能的結合の指標に基づく特徴量を検討した。すべてのチャンネル間の機能的結合からなるテンソルを、2つのチャンネル方向について対称な低ランクテンソルで近似し、周波数帯域・大脳皮質全体にわたる結合を表現する低次元の特徴を抽出した。

エアコン音の対比較評価時の脳磁界データから予測モデルを学習・評価した結果、被験者によっては機能的結合に基づく特徴量を用いた場合に最も高い予測精度を示し、機能的結合の情報が印象の予測に有効である可能性が示唆された。すべての被験者で高い予測精度を得るためには、より頑健な機能的結合の指標や特徴量選択の手法が必要である。

**謝辞** 本研究の一部は、JSPS 科研費 JP18K19820 の助成を受けたものである。

## 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), 539–542, 2017.
- [5] 矢野ら, 音講論(春), 755–758, 2018.
- [6] 矢野ら, 音講論(秋), 729–732, 2019.
- [7] M. Rubinov and O. Sporns, NeuroImage, 52, 1059–1069, 2010.
- [8] M. Vinck *et al.*, NeuroImage, 55, 1548–1565, 2011.
- [9] D. Kuang *et al.*, Proc. of SIAM Int. Conf. on Data Mining, 106–117, 2012.
- [10] D. Kuang *et al.*, J. Glob. Optim., 52, 545–574, 2015.
- [11] A. Cichocki *et al.*, “Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-Way Data Analysis and Blind Source Separation,” Wiley, 2009.
- [12] 矢野ら, 音講論(春), 123–124, 2015.