

空間フィルタによる印象関連脳活動の抽出の試み*

☆矢野 肇 (神戸大/産総研), 高島遼一, 滝口哲也 (神戸大),
神谷 勝 ((株)デンソー), 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]。印象の予測に有効な特徴を抽出するため、刺激対に対する比較判断を特徴量抽出に導入した [5, 6]。しかしながら、印象予測モデルによる比較判断の予測精度は高々60%と低く、現状ではこれを用いた印象評価は現実的ではないと言わざるをえない。

これまでは主に脳活動の周波数成分ごとのパワーを元にした特徴量の抽出を行ってきたが、計測データには印象と関係のない脳活動も多く含まれていると考えられ、印象と関連する脳活動信号の強度が変化したとしても、SN比が低いためにその変化が十分に取り出せない可能性がある。したがって、心理的な情報を利用して、印象の予測に有効な脳活動信号の強度を高めることが必要と考えられる。

脳波からの運動想起の識別によく用いられる手法として、Common Spatial Pattern (CSP) がよく知られている [7]。CSP によって2つの状態の識別に最適な空間フィルタを構築することができ、このフィルタは一方の状態の信号を強調/抑制と同時に他方の状態の信号を抑制/強調する。CSP を用いて主観的な判断の情報を考慮した空間フィルタを構築することで、印象と関連する脳活動のSN比を向上できる可能性がある。

本稿では、エアコン音の涼しさ・好ましさに関する一対比較タスク遂行時の脳活動に対し、脳活動特徴量を CSP を用いて抽出し、印象予測モデルの精度向上を試みた。CSP 特徴量を用いて被験者ごとの印象予測モデルを学習し、その性能を評価した。また、予測に寄与した CSP 特徴量の空間フィルタから、印象

に関連する脳活動を調査した。

2 脳磁界データ

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 [8] 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、刺激音圧は被験者自身に調整させた。

2.2 脳磁界計測実験

実験参加の同意を得た聴覚健全者6名 (男性4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。各被験者について、刺激音の涼しさの評価を行う場合と好ましきの評価を行う場合の計2回の計測が行われた。計測に先立ち、被験者ごとに7種類の刺激音に対する涼しさ、および好ましきの尺度値を一対比較法によって調査し、各印象について、その尺度値が最大、最小となる刺激音を含む計4つの刺激音を選択した。被験者には4つの刺激のうち2つずつを連続的に呈示し、2つ目の刺激の呈示終了後に、各基準の優劣を評価するように教示した。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-122TM, 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のアナログフィルタを通過した後、400 Hzでサンプリングされた。

2.3 データの前処理

まず、計122チャンネルで計測された raw data に重畳する電源由来の60 Hzの成分と2 Hz以下の低周波成分を、それぞれ2次および4次のバターワースフィルタで除去した。次に、不良チャンネルの信号と絶対値が1000 fT/cmを超える振幅が現れた試行を取り除いた。さらに、独立成分分析を適用し、眼球運動に由来

*Extraction of cortical activities associated with auditory impressions using spatial filter. by YANO, Hajime (Kobe Univ./AIST), TAKASHIMA, Ryoichi, TAKIGUCHI, Tetsuya (Kobe Univ.), KAMIYA, Masaru (DENSO Corp.), NAKAGAWA, Seiji (Chiba Univ.).

する成分や心磁図などのノイズと考えられる成分を取り除いた。なお、独立成分分析のアルゴリズムには FastICA [9] を用いた。

3 Common Spatial Pattern を用いた特徴量抽出

3.1 Common Spatial Pattern

長さ T の N チャンルの帯域制限信号からなる行列を $\mathbf{E} \in \mathbb{R}^{N \times T}$ で表す。多チャンネル信号 \mathbf{E} は2つの状態のいずれかに対応しているとし、状態を表すクラスラベルを $c \in \{1, 2\}$ とする。CSP では片方の状態の信号の分散を大きくする一方で、他方の状態の信号の分散を小さくするような空間フィルタ \mathbf{W} が構成される [7, 10]。フィルタリングは次式で行われる。

$$\mathbf{E}_{\text{CSP}} = \mathbf{W}^T \mathbf{E} \quad (1)$$

ここで、 \mathbf{E}_{CSP} は空間フィルタ後の多チャンネル信号である。状態 c の信号の分散共分散行列 \mathbf{C}_c は以下の式で計算される。

$$\mathbf{C}_c = \frac{1}{|\mathcal{I}_c|} \sum_{i \in \mathcal{I}_c} \frac{\mathbf{E}_i \mathbf{E}_i^T}{\text{tr}(\mathbf{E}_i \mathbf{E}_i^T)} \quad (2)$$

ここで、 \mathcal{I}_c はクラス c に属する信号のインデックスの集合を表し、 $\text{tr}(\cdot)$ は行列のトレースを表す。空間フィルタ行列 \mathbf{W} の各列ベクトルは、次の一般化固有値問題の解として得られる。

$$\mathbf{C}_1 \mathbf{w} = \lambda (\mathbf{C}_1 + \mathbf{C}_2) \mathbf{w} \quad (3)$$

ここで、 λ は一般化固有値で、状態 1 ($c = 1$) のフィルタ後の信号の分散に対応している。フィルタ後の状態 1 の分散と状態 2 の分散の和は常に 1 となる性質があり、状態 1 の分散が最大の時、状態 2 ($c = 2$) の分散は最小になり、状態 1 の分散が最小の時、状態 2 の分散は最大となる。

空間フィルタ行列 \mathbf{W} の逆行列の転置行列 ($\mathbf{W}^{-1})^T$ は空間パターンと呼ばれる。この行列の列ベクトルは、対応するフィルタで抽出された信号を元の空間に戻した時の信号の強度分布を表す。

3.2 CSP を用いた印象関連脳活動の抽出

本稿では、計測された脳活動を一対比較判断に基づき2つの状態に対応させて、CSP の空間フィルタを学習した。具体的には、一対比較において印象の程度が優れていると判断された刺激音の聴取時の脳活動を状態 1、そうでないと判断された刺激音の聴取時の脳活動を状態 2 として、空間フィルタを学習した。これによって、印象の程度に従って変化する脳活動を強調もしくは抑制するような空間フィルタの構築が期待される。

4 評価試験

被験者ごとに CSP 特徴量を用いて比較判断を予測するモデルを学習し、予測精度の評価を行った。また、空間フィルタを用いない特徴量でもモデルを学習し、予測性能を CSP 特徴量を用いた場合と比較した。

4.1 特徴量

前処理後の脳磁界データから 8 次のバターワースフィルタを用いて、4–8 Hz, 8–13 Hz, 13–20 Hz, 20–30 Hz, 30–50 Hz, 50–70 Hz, 70–100 Hz の 7 種類の帯域信号を抽出した。帯域ごとに刺激呈示後 500–4500 ms の区間の信号を用いて CSP の空間フィルタを学習し、一般化固有値が最大および最小の空間フィルタを特徴量抽出に用いた。空間フィルタ後の信号の 500–4500 ms 区間の平均パワーを対数変換したものを CSP 特徴量とした。したがって、この特徴量の次元数は 14 となる。

空間フィルタを用いない特徴量として、事象関連脱同期/同期 (Event-related Desynchronization/Synchronization: ERD/ERS) [11] に基づく特徴量に主成分分析 (Principal Component Analysis: PCA)、および線形判別分析 (Linear Discriminant Analysis: LDA) を適用した特徴量を抽出した。ERD/ERS 特徴量は各帯域の刺激呈示後 500–4500 ms の平均パワーを P_a 、刺激呈示前 500 ms 間の平均パワーを P_r とすると、 $ER = (P_a - P_r)/P_r$ でチャンネルごとに計算された。これに帯域ごとに PCA を適用し、ER-PCA 特徴量とした。次元数は CSP 特徴量と同じ 14 であった。また、PCA の代わりに LDA を帯域ごとに適用して次元圧縮した特徴量を ER-LDA 特徴量とした。LDA によって、比較判断に対応する 2 クラス間の分離が最大になるように入力特徴量の線形変換が学習される。各帯域で LDA を適用して得られる特徴量の次元数は (クラス数)–1 = 1 のため ER-LDA 特徴量の次元数は 7 となった。

4.2 評価方法

本稿では、比較判断の予測を 2 種類の方法で行った。1 つ目は、従来の印象予測モデル [4] によって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方法である。CSP 特徴量抽出時の 2 つクラスは、より涼しい (好ましい) と、より涼しくない (好ましくない) に対応し、ER-LDA 特徴量の抽出では、刺激対に対応する ERD/ERS 特徴量の差に LDA が適用された。印象予測モデルは線形 SVM (Support Vector Machine) と同様の枠組みで学習した。モデル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。

2つ目は、判断の直前に呈示された刺激、つまり一対比較において後続の刺激を聴取した時の脳活動のみから比較判断の予測を行う方法である。この場合は単純な2クラス分類問題となる。CSP特徴量の抽出では判断のラベルに対応した脳活動を強調/抑制するフィルタが構築され、ER-LDA特徴量の抽出では、後続の刺激聴取時のERD/ERS特徴量にそのままLDAが適用される。分類器には線形SVMを用いた。各クラスのデータ数に差がある場合を考慮して、SVMの学習時にデータ数に応じて重みづけを行った。交差検証における評価尺度に、クラスごとに算出したF1 scoreを平均したMacro F1 scoreを用いた。

4.3 結果・考察

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. 1に示す。涼しさの評価において、被験者3の印象予測モデルにおけるCSP特徴量の正解率が他の特徴量と比較して高くなった。一方、被験者2では、ER-LDA特徴量が他の特徴量と比較して高い平均正解率を示した。好ましきの評価においては、4人の被験者でCSP特徴量の平均正解率が最も高くなったが、他の特徴量と比較して大きな差は見られなかった。

次に、比較判断の直前の刺激を聴取した時の脳活動特徴量から、涼しさおよび好ましきに関する比較判断を予測した時のMacro F1 scoreの平均および平均正解率をFig. 2に示す。Macro F1 scoreはいずれの印象の評価でも0.5付近の値が多かったが、涼しさの評価における被験者3のCSP特徴量を用いた場合のMacro F1 scoreは他の特徴量と比べて高くなった(Fig. 2(a))。平均正解率は涼しさの評価において、被験者1, 3, 4, 6でCSP特徴量を用いた場合に約0.6–0.7と他の特徴量と比べて高い値となった(Fig. 2(b))。好ましきの評価では被験者1, 6でCSP特徴量を用いた場合に他の特徴量と比較して高い平均正解率となった(Fig. 2(d))。

CSP特徴量を用いて比較判断を単純なクラス分類で予測した場合、Macro F1 scoreよりも正解率のほうが大きくなる傾向が見られた。この理由として、判断の数に偏りがあり、多い方の判断が予測されやすくなっている可能性も考えられるが、各判断のデータ数で重み付けした学習を行っていることや、同じく識別的な特徴量であるER-LDA特徴量を用いた場合の正解率は高くないことから、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、頑健な特徴量が取

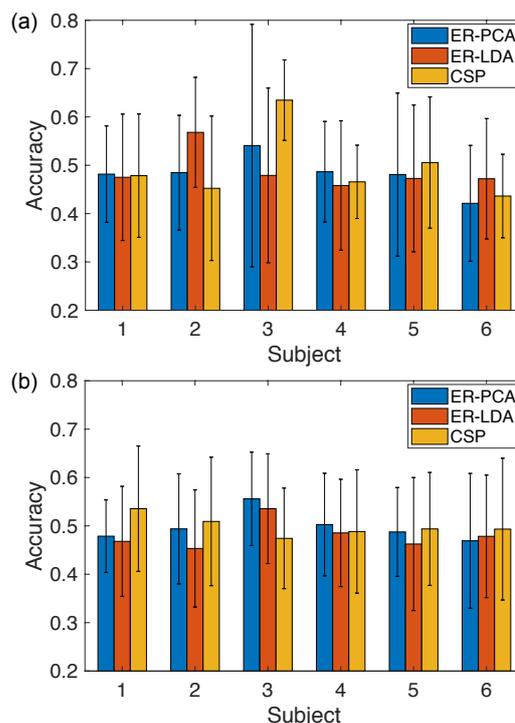


Fig. 1 Mean prediction accuracies of predicted comparative judgment on (a) coolness and (b) preference. Each error bar indicates standard deviation.

り出せたためと考えられる。その一方で、単純なクラス分類で予測した場合と比べて、印象予測モデルを用いて予測した場合の正解率は低い傾向があった。印象予測モデルでは刺激対を聴取した時の脳活動から得られる2つの特徴量を同等に扱うが、被験者は後続の刺激の聴取時に先行した刺激との比較を行う可能性があるため、同等の特徴量が得られている保証はない。後続の刺激音を聴取した時の脳活動に、先行した刺激音と比べた時の涼しさや好ましきの差が反映されたため、空間フィルタでその活動が取り出され、クラス分類器で予測した場合の正解率が高くなったと考えられる。

涼しさの評価において、被験者3のCSP特徴量を用いた場合のモデルは一貫して他の特徴量と比較して高い精度を示した。識別的な空間フィルタによって涼しさの印象に伴って変化する脳活動を抽出できたと考えられる。高い正解率の時の予測に最も寄与したCSP特徴量の空間パターンを元に計算した磁界パターンの例をFig. 3に示す。脳活動の帯域は50–70 Hz (Mid- γ 帯域) で、左前頭と右後頭付近に磁界の吸い込みと湧き出しのパターンが見られることから、左前頭と右後頭のMid- γ 帯域の活動が涼しさの印象に関連している可能性が考えられる。しかしながら、予測に寄与した特徴量の次元は複数の帯域にわたっており、印象と関連する脳活動の同定にはさらなる検討が必要である。

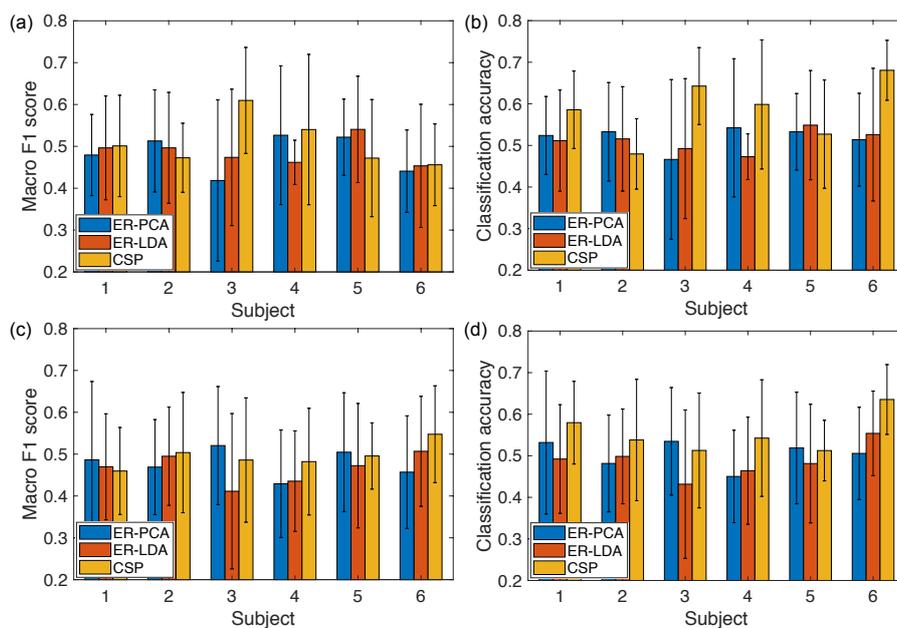


Fig. 2 (a) Mean macro F1 scores and (b) mean accuracies of predicted comparative judgment on coolness. (c) Mean macro F1 scores and (d) mean accuracies of predicted comparative judgment on preference. Each error bar indicates standard deviation.

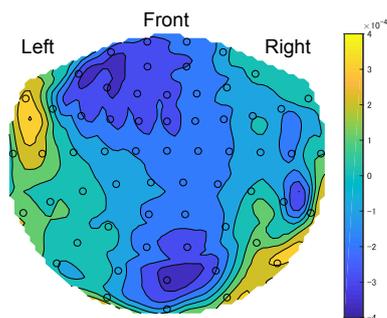


Fig. 3 An example of magnetic spatial pattern of the spatially filtered mid- γ activity.

5 まとめ

エアコン音に対する印象予測モデルの精度を向上させるため、CSPを利用して印象の程度によって変化する脳活動を反映する特徴量の抽出を試みた。比較判断の直前に聴取した刺激音に対する脳活動の特徴量から一対比較判断の予測を行った結果、CSP特徴量を用いた場合に他の特徴量と比較して高い正解率を示した。また、CSPの空間パターンを調べることで、印象の程度によって変化する脳活動を同定できる可能性がある。その一方で、従来の印象予測モデルを用いた予測では、ほとんどの被験者で大きな正解率の向上は見られず、印象予測モデルは刺激の呈示順の影響を受ける可能性が示唆された。そのため、呈示順による脳活動の違いを考慮して印象予測モデルを学習する必要がある。また、より良い特徴量抽出のためには、CSPにおける最適な帯域の選択や正則

化の導入 [12], 各周波数帯域のパワーとは異なる他の特徴量の統合などが必要だと考えられる。

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 (秋), 573–576, 2017.
- [6] 矢野ら, 音講論 (秋), 381–384, 2018.
- [7] B. Blankertz *et al.*, IEEE Signal Process. Mag., 25, 41–56, 2008.
- [8] T. Hotehama and S. Nakagawa, Proc. of INTER-NOISE 2015, in 15-915.
- [9] A. Hyvärinen and E.Oja, Neural Netw., 13(4-5), 411–430, 2000.
- [10] H. Ramoser *et al.*, IEEE Trans. Rehabil. Eng., 8, 441–446, 2000.
- [11] G. Pfurtscheller and F. H. Lopes, Clin. Neurophysiol., 110, 1842–1857, 1999.
- [12] F. Lotte and C. Guan, IEEE Trans. Biomed. Eng., 58, 355–362, 2011.