



基于多视角聚类分析的汉字字体审美偏好挖掘

张艳^{1,5†}, 谢源^{2†}, 洪辰¹, 曲延云^{1*}, 李睿³, 张俊松⁴, 李翠华¹

1. 厦门大学信息学院计算机科学系, 厦门 361005
2. 华东师范大学计算机科学与技术学院, 上海 200062
3. 华中师范大学教育大数据应用技术国家工程实验室, 武汉 430079
4. 厦门大学信息学院人工智能系艺术认知与计算实验室, 厦门 361005
5. 贵州师范大学数学科学学院, 贵阳 550001

* 通信作者. E-mail: yyqu@xmu.edu.cn

† 同等贡献

收稿日期: 2020-08-01; 修回日期: 2020-09-12; 接受日期: 2020-09-16; 网络出版日期: 2021-02-23

国家自然科学基金面上项目 (批准号: 61876161, 61772524, 61671397, U1065252, 61772440) 资助

摘要 在神经美学研究中已经证明, 中文字体审美偏好的情绪刺激可以通过观察 3 种偏好 (喜欢、不喜欢和中性) 之间的事件相关电位 (event related potential, ERP) 波动获得. 本文通过引入一种核化张量奇异值分解的多视角聚类方法分别构建了基于脑电图 (electroencephalogram, EEG) 和 ERP 的审美偏好识别模型, 通过这些模型首次确认了该结论. 本文方法将来自不同频段的数据视为描述中文字体审美偏好的不同视角, 通过张量多秩最小化的约束探索所有视角特征的一致性和关联性, 并通过之后的聚类获取审美偏好的识别结果. 采用多视角无监督聚类方法得到的识别精度达到 97.1%. 此外, 通过输入-扰动关联方法将电极的振幅与不同类型的审美偏好相关联, 可视化关键频段组合以及电极之间的关系, 分别取出与喜欢、不喜欢、中性最相关的 3 个电极, 包含次相关的 6 个电极, 包含第三相关的 9 个电极, 包含第四相关的 12 个电极, 分别形成 4 种不同组合的脑电特征. 通过比较实验, 验证了相对于 62 个电极信号, 上述 4 种组合方式在字体美学分类上更具有优势, 并且最相关的 3 个电极的组合特征对审美偏好最具判别性. 实验结果表明, 基于多视角聚类的方法能够解决神经信号与审美偏好的相关分析, 并能挖掘出与字体审美偏好最相关的电极.

关键词 中文字体, 审美评价, 计算美学, 事件相关电位, 核化张量奇异值分解, 数据挖掘

1 引言

汉字是世界上最古老的文字之一, 属于表意文字体系, 一个汉字通常表示汉语里的一个词或一个语素, 拥有音、形、义统一的特点. 直观上, 对于文字的感官, 除了理解文字本身的含义, 同时也会关注

引用格式: 张艳, 谢源, 洪辰, 等. 基于多视角聚类分析的汉字字体审美偏好挖掘. 中国科学: 信息科学, 2021, 51: 383–398, doi: 10.1360/SSI-2020-0234
Zhang Y, Xie Y, Hong C, et al. Esthetic preference mining of Chinese typefaces using multiview cluster analysis (in Chinese). Sci Sin Inform, 2021, 51: 383–398, doi: 10.1360/SSI-2020-0234

文字本身的字体. 字体是汉字的重要特征, 可以用不同的样式来表示, 例如标准的打印字体、手写体和书法家的作品. 以书法为例, 它是以汉字为载体, 以文房四宝“笔、墨、纸、砚”作为主要创作工具, 以笔法、章法为主要表现技巧, 以文字涵义和书法审美取向为表现内容的一种线条造型艺术. 随着计算机与信息技术的发展, 数字书法则将传统书法的创作工具、视觉艺术效果、书写技巧和书法创作等用数字化的方式再现. 而将汉字以符合人类审美偏好的方式呈现, 字形的审美评价与计算是一个不能回避的关键问题^[1]. 近年来, 有关审美认知的研究取得了长足的进展, 特别是审美偏好方面的研究已经针对审美认知评价进行了实验设计和客观度量方法的研究, 为构建符合审美认知机理的汉字字体的建模方法提供了理论支持.

在日常生活中, 即使没有主体的评估和决策指导, 某些常见的刺激 (例如徽标、图像和几何图形) 也可能触发人类内在的审美偏好^[2~4]. 以往的神经审美研究表明, 人的审美偏好与脑电图 (electroencephalogram, EEG) 信号之间存在很强的相关性. 而且, 由于 EEG 信号具有较高的时间分辨率, 具有和神经认知过程相同的时间尺度, 因此可以将其视为理解人的偏好并揭示潜在模式的有效工具. 事件相关电位 (event related potential, ERP)^[5] 是从 EEG 中发展而来的, 是特定刺激所诱发出的脑电信号, 通常可以通过将多个相同事件的 EEG 信号进行平均叠加获取, 用以反映认知过程中大脑的神经电生理改变. ERP 也被称为认知电位, 为研究大脑认知活动过程提供了新的方法和途径. 2000 年, Jacobsen 等^[6] 在人类的审美判断上进行了第一个 ERP 实验, 用于探讨审美判断的神经生理学基础. 此后, ERP 被广泛用于多种刺激的审美偏好研究, 比如, 人脸、图像等刺激. 除此以外, 还有部分研究者对英文字体的审美偏好做了初步的研究^[7,8], 但是其中关于汉字字体引起的审美偏好的研究非常少. 张俊松课题组^[9] 于 2015 年首次发现了 ERP 与字体的审美偏好有关联性. 他们首先构造了实验所用的汉字数据集, 记录每位参与者看到每个汉字的脑电信号. 然后让参与者为汉字标注偏好标签, 主观选择喜欢、不喜欢和中性的汉字. 形成了脑电信号和主观标签之间的对应关系. 最后通过对比喜欢字体刺激、不喜欢字体刺激和中性字体刺激下的脑电信号, 发现它们之间的信号有明显的差异, 从而得出汉字的审美与脑电有关联, 可以通过脑电信号发现人类对字体的审美偏好.

张俊松课题组^[9] 开辟了汉字字体审美偏好的神经美学研究, 但主要采用事件相关脑电位方法分析字形审美偏好的神经机制, 没有采用机器学习方法对脑电数据进行分析. 本文主要研究以下两个问题: (1) 能否通过对脑电数据的学习, 取得神经美学偏好的证明. 换句话说, 就是能否从脑电数据中挖掘出字体美学审美偏好的关联性, 即喜欢字体、不喜欢字体和中性字体的脑电信号具有聚合性, 能够形成 3 个不同的聚类; (2) 能否挖掘出与字体美学偏好关联的电极, 区分出具有判别性的电极与一般的电极, 以供后续的字体审美偏好分类. 为解决这些问题, 首先要对脑电数据进行特征表示. 在大多数方法中, 通常将脑电信号转换为特定频段后直接提取特征, 然后将所有频段的特征简单地串连为一个特征向量, 这种做法把所有频段的信号串连成一个一维信号, 通过一阶信号分析挖掘频段和电极之间的关联性. 该方式忽视了不同频段在审美关联性分析中的高阶关联性. 就好比判断一个学生优秀与否, 不能只看成绩, 而应该从德、智、体、美、劳等各方面全面地考察学生的综合素质. 得益于多视角聚类分析在图像聚类的成功应用, 本文针对脑电信号提出基于多视角聚类的字体美学偏好挖掘模型. 我们将不同频段的脑电信号当成不同的视角特征, 其中每个频段的字体偏好脑电信号形成了该频段所有字体偏好的特征描述矩阵. 由于不同频段可以看成不同的描述, 多种频段就形成了多个描述. 由于基于核化的张量奇异值分解方法是一种最新的被证明行之有效的多视角聚类方法, 本文将该方法用于多频段脑电信号聚类分析. 同时进一步利用脑电信号生成 ERP, 讨论与事件有关而与试者无关的电极与审美偏好之间的关系, 挖掘与字体审美偏好最相关的电极. 最后利用 ERP 挖掘的最具有判别性的电极进行字体偏好分类.

本文的主要贡献包括三方面.

(1) 提出了一种基于多视角聚类的字体美学偏好挖掘模型. 该模型将不同频段的脑电信号当成多个视角特征, 采用基于核化的张量奇异值分解 (kernelized tensor singular value decomposition, KT-SVD) 的算法对多视角聚类进行建模.

(2) 针对多视角脑电信号形成的多视角聚类模型, 设计了基于交替迭代的求解算法.

(3) 通过输入 - 扰动关联映射算法, 挖掘出字体审美偏好与脑电信号电极之间的关联性, 找到最具有判别性的电极.

2 相关工作

2.1 审美偏好

现有的大多数探索人类审美偏好的研究主要基于二维图像或三维形状. 文献 [10] 通过研究基于图像的空间结构、显著区域分布和突出线条的计算美学分数研究图像的构图美学. 文献 [11] 中的工作提出了一种方向梯度直方图金字塔 (pyramid histogram of oriented gradients, PHOG) 的表示形式表征审美图像. 还有一些神经美学研究^[12,13]集中在讨论单词的情感含义上. 文献 [9] 是第一个研究汉字字体是否会引起审美偏好的文章, 他们研究发现被试者在观察喜欢、不喜欢和中性字形的 ERP 波之间存在显著差异, 表明人类对汉字的字体具有快速、隐式的审美判断.

2.2 当前的 EEG 分类方法

传统的 EEG 识别方法可分为两部分: 特征提取和分类. 分类算法如: 支持向量机 (support vector machine, SVM) 算法或线性判别分析 (linear discriminant analysis, LDA) 算法等. 文献 [14] 通过组合眼动信号和 EEG 增强分类性能, 从而适应多模式下的情感识别框架. Zheng 等^[15]使用迁移学习技术来建立基于 EEG 的个性化情感模型. 另外随着深度学习技术研究的高速发展, 涌现了很多利用深度学习进行 EEG 分类的方法. 文献 [16,17] 应用卷积神经网络 (convolutional neural network, CNN) 来分析和分类 EEG 数据. Li 等^[18]根据情感大脑左右半球之间不对称性的神经科学发现, 构建了一个用于脑电图情感识别的双半球域对抗神经网络.

2.3 大脑活动模式

情感神经科学的各种研究旨在揭示个体之间共享的某些特定大脑活动的共同模式. Bos^[19]发现, 针对情感识别的 α 和 β 频段, 对于激活维 (arousal), 当以 β 频段选择 F3/F4 和 FPZ 的频段功率时可获得分类最佳性能; 而对于效价维 (valence), 获得分类最佳性能的组合是使用 F3/F4 和 FPZ 的 α 和 β 的频段功率. Nie 等^[20]发现与情绪有关的脑电图特征主要包括右枕叶和顶叶的 α 波段脑电图能量, 中心点的 β 频段脑电图能量以及左额叶和右颞叶的 γ 频段的脑电图能量. Wang 等^[21]指出对于正向和负向情绪, 与受试者无关的特征主要集中在右枕叶和顶叶的 α 频段, 顶叶和颞叶的 β 频段以及左额叶和右颞叶的 γ 频段.

本文旨在寻求一种用于汉字字体审美偏好识别 (喜欢、不喜欢、中性) 的模式, 并能够探索 EEG 关键频段组合以及电极之间关系的方法. 研究主要包括以下几个部分: EEG 数据的采集和预处理、多视角的表示方法、KT-SVD 方法的模型定义及优化、用于审美偏好识别的无监督多视角聚类方法和用于可视化分析关键频段组合和电极之间关系的输入 - 扰动关联映射方法.

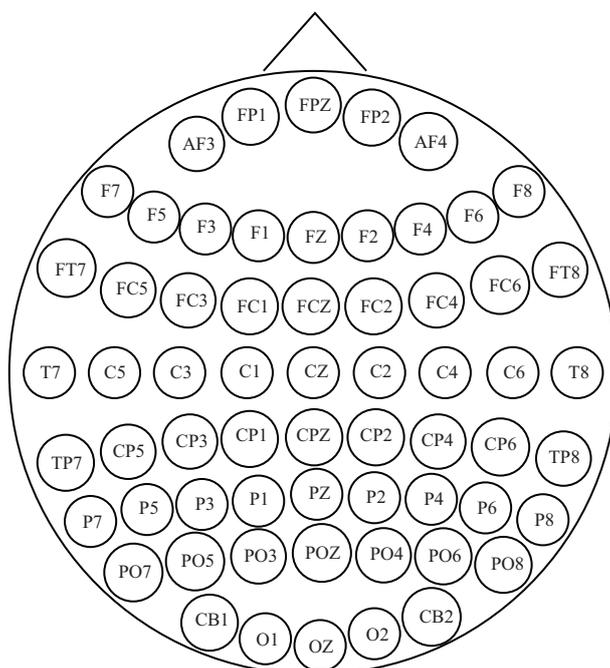


图 1 62 个电极的脑电波帽布局图
Figure 1 EEG cap layout for 62 electrodes

3 数据采集和预处理

3.1 获取 EEG 数据集

我们采用汉字字体作为刺激来激发审美偏好, 总体选择 175 个汉字作为刺激, 同时为了控制刺激的意义偏好, 所选汉字均为中性词. 实验涉及被试者 15 名, 每个被试者总共进行 5 轮实验. 在每轮实验中, 175 个字符中的每个字符仅显示一次. 在每轮实验期间, 首先在屏幕中央呈现一个固定十字, 让被试者的注意落在屏幕中央, 作为开始的提示. 实验中采用的注视交叉, 其持续时间在 1300 到 1500 ms 之间变化. 随后在屏幕中央显示一个汉字作为刺激, 持续 200 ms, 以调查被试者对汉字字体审美的快速反应, 然后再进行另一个注视交叉, 以进行下一个跟踪. 执行完每轮实验后, 要求每个被试者选择最喜欢的 12 个字符和最不喜欢的 12 个字符, 剩余 151 个字符被归类为中性. 根据国际 10-20 系统, 我们使用带有 62 通道电极帽的 Neuroscan 4.5 系统同时记录 EEG 数据. EEG 电极的布局如图 1 所示. 在记录期间, 所有电极的阻抗均保持低于 5 k Ω , 原始采样率为 1000 Hz.

3.2 数据预处理

由于 EEG 信号具有相对较低的信噪比, 也就是说, 与任务无关的信息源通常比任务相关的信息源对 EEG 信号的影响更大. 例如, 它们通常包含由肌肉活动引起的肌电图 (electromyogram, EMG) 信号污染和眼睛运动引起的眼电图 (electro-oculogram, EOG) 信号污染. 因此, 首先采用 0.05~100 Hz 的带通来放大 EEG 和 EOG 信号, 接着通过自动伪像排除功能消除包含眼球运动和肌肉活动引起的信号污染, 最后再将 EEG 信号下采样到 500 Hz, 以降低计算复杂性.

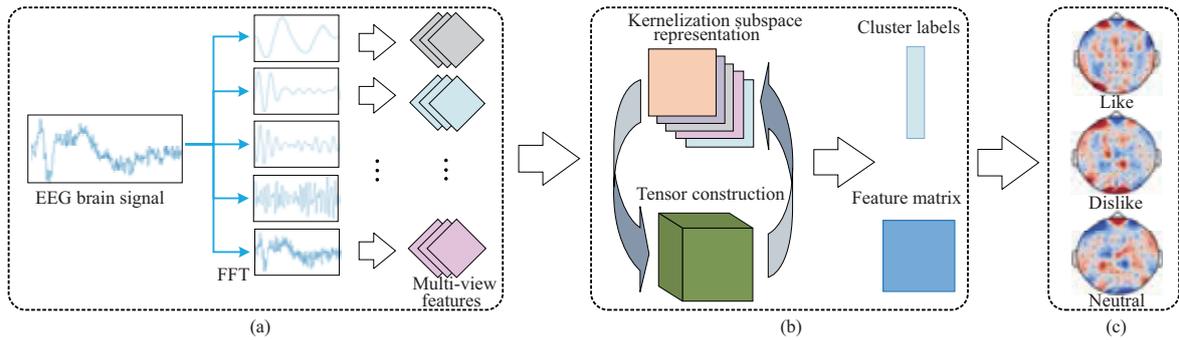


图 2 (网络版彩图) KT-SVD 模型概述. (a) 从 5 个频段 (δ : 1~3 Hz, θ : 4~7 Hz, α : 8~13 Hz, β : 14~30 Hz, γ : 31~50 Hz) 中提取 EEG 特征; (b) 基于叠加的张量结构, 进行核诱导特征空间中的多视角自表示的学习, 通过生成的特征自表示, 获取相似度矩阵, 用于谱聚类中得到特征矩阵和聚类标签, 聚类标签即为最后的审美偏好识别结果; (c) 基于前述过程处理 ERP 数据, 将结果送入输入 - 扰动关联映射方法中获取大脑关联图的可视化

Figure 2 (Color online) Overview of the proposed model. (a) EEG's feature extraction from five frequency bands (δ : 1~3 Hz, θ : 4~7 Hz, α : 8~13 Hz, β : 14~30 Hz, γ : 31~50 Hz) for multi-view representation; (b) multi-view self-representation in the kernel induced feature space and tensor construction by stacking the kernel multi-view self-representation; (c) visualization of brain correlation map through the input-perturbation correlation method to process the ERP-based feature matrix and cluster labels

4 方法

我们的审美偏好识别框架如图 2 所示, 主要包括 3 个部分: 多视角表示、核化子空间多视角的自表示及聚类 and 脑电图关联的可视化. 具体来说, 首先通过快速傅立叶变换 (fast Fourier transform, FFT) 将 EEG 信号变换到 5 个不同的频段, 分别提取该频段下 EEG 信号的特征, 将获得的数据作为基于不同频段的多视角表示. 其次利用 KT-SVD 算法对多视角表示进行无监督的一致化特征自表示过程学习, 然后基于得到的结果获取多视角的特征间的相似度矩阵, 再将相似度矩阵用于谱聚类算法中得到识别结果. 最后, 采用上述过程处理 ERP 信号, 并获取 ERP 信号的相似度矩阵, 通过给出的输入 - 扰动关联映射方法分析得到与 3 种审美偏好高度相关的关键电极组合方案.

4.1 多视角表示

在情感识别任务中存在许多手工设计特征以识别在不同情绪状态之间变化的时间序列的特征. 我们在实验中选取了时域参数特征 (time domain parameters, TDP) 和带功率特征 (band power, BP), 这两种特征都是基于 EEG 的情感识别中应用最广泛的特征. 我们遵循文献 [22, 23] 的定义, 首先采用 FFT 将原始信号变化到 5 个不同的频段, 然后分别提取特征. 即分别采用 δ (1~3 Hz), θ (4~7 Hz), α (8~13 Hz), β (14~30 Hz) 和 γ (31~50 Hz) 的 5 个不同频段中提取的特征作为多视角表示的原始数据.

4.2 KT-SVD

令 $\mathbf{X}^{(v)} = [x_1^{(v)}, x_2^{(v)}, \dots, x_n^{(v)}] \in \mathbb{R}^{d \times n}$ 表示第 v 个视角特征矩阵, 其列是来自第 v 个视角 (第 v 个频段) 的 EEG 信号的特征, 其中 d 表示原始特征空间的维度, n 表示信号的数量. 根据 t-SVD-MSC^[24] (t-SVD based multi-view subspace clustering) 在新的子空间表示矩阵中实现自表示的过程, 特征的自表示被定义为 $\mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}, \dots, \mathbf{Z}^{(V)}$, 随后通过合并不同的 $\mathbf{Z}^{(V)}$ 来构建 3 阶张量 $\tilde{\mathbf{Z}}$.

由于 t-SVD-MSC 算法主要用于处理线性子空间中的数据, 不适用于非线性情况. 因此, 在原始特

征空间上应用基于核化的方法^[25,26]来解决此问题. 如果定义 $\ker^{(v)}(x, y)$ 表示核函数, 则有

$$K_{ij}^{(v)} = \ker^{(v)}(x_i^{(v)}, x_j^{(v)}), \quad \forall i, j = 1, \dots, n, \quad (1)$$

其中, 核函数 $\ker^{(v)}(x, y)$ 中引入一个映射 $\psi^{(v)}: \mathbb{R}^{d^{(v)}} \rightarrow \mathcal{F}^{(v)}$ (通常隐式定义), $\mathcal{F}^{(v)}$ 被称为新的特征空间. 给定映射 $\psi^{(v)}$, 则函数 $\ker^{(v)}(\cdot, \cdot)$ 可以被重写为

$$\ker^{(v)}(x, y) = \psi^{(v)}(x)^T \psi^{(v)}(y). \quad (2)$$

定义 $\Psi^{(v)}(\mathbf{X}^{(v)}) = [\psi^{(v)}(x_1^{(v)}), \dots, \psi^{(v)}(x_n^{(v)})]$, 则第 v 个视角的核矩阵可以通过下式进行计算:

$$\mathbf{K}^{(v)} = \Psi^{(v)}(\mathbf{X}^{(v)})^T \Psi^{(v)}(\mathbf{X}^{(v)}). \quad (3)$$

此外, 特征空间 $\mathcal{F}^{(v)}$ 中的重构误差可以表示为

$$\mathbf{E}^{(v)} = \Psi^{(v)}(\mathbf{X}^{(v)}) - \Psi^{(v)}(\mathbf{X}^{(v)})\mathbf{Z}^{(v)}. \quad (4)$$

因此, t-SVD-MSC 的原始目标函数可以在特征子空间 $\mathcal{F}^{(v)}$ 中转换为非线性版本, 则 KT-SVD 模型可以写为如下形式:

$$\begin{aligned} \min_{\mathbf{Z}^{(v)}} \lambda \sum_{v=1}^V \|\Psi^{(v)}(\mathbf{X}^{(v)}) - \Psi^{(v)}(\mathbf{X}^{(v)})\mathbf{Z}^{(v)}\|_{2,1} + \|\tilde{\mathbf{Z}}\|_{\otimes}, \\ \tilde{\mathbf{Z}} = \Phi(\mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}, \dots, \mathbf{Z}^{(V)}), \end{aligned} \quad (5)$$

其中, 函数 $\Phi(\cdot)$ 表示聚合旋转形成张量 $\tilde{\mathbf{Z}}$ 的操作. 函数 $\|\cdot\|_{\otimes}$ 表示张量核范数计算.

对于方程 (5) 的优化问题, 我们使用增广拉格朗日 (Lagrange) 乘数^[27] 和交替方向最小化策略^[28], 通过引入辅助张量变量 \mathcal{G} , 以及拉格朗日乘子: 矩阵 \mathbf{Y}_v ($v = 1, 2, \dots, V$) 和张量 \mathcal{W} , 使得目标函数可分离. 同时, 令 $h^{(v)}(\mathbf{P}^{(v)}) = \sum_{i=1}^n \sqrt{\mathbf{p}_i^{(v)T} \mathbf{K}^{(v)} \mathbf{p}_i^{(v)}}$, $\mathbf{P}^{(v)} = \mathbf{I} - \mathbf{Z}^{(v)}$, $\mathbf{P}^{(v)} = [\mathbf{p}_1^{(v)}, \dots, \mathbf{p}_n^{(v)}]$. 可以将优化问题转换为最小化下列等效问题:

$$\begin{aligned} \mathcal{L}(\mathbf{Z}^{(1)}, \dots, \mathbf{Z}^{(V)}; \mathbf{P}^{(1)}, \dots, \mathbf{P}^{(V)}; \mathbf{Y}_1, \dots, \mathbf{Y}_V; \mathcal{W}; \mathcal{G}) \\ = \|\mathcal{G}\|_{\otimes} + \lambda \sum_{v=1}^V h^{(v)}(\mathbf{P}^{(v)}) \\ + \sum_{v=1}^V \left(\langle \mathbf{Y}_v, \mathbf{I}^{(v)} - \mathbf{Z}^{(v)} - \mathbf{P}^{(v)} \rangle + \frac{\mu}{2} \|\mathbf{I}^{(v)} - \mathbf{Z}^{(v)} - \mathbf{P}^{(v)}\|_F^2 \right) \\ + \langle \mathcal{W}, \tilde{\mathbf{Z}} - \mathcal{G} \rangle + \frac{\rho}{2} \|\tilde{\mathbf{Z}} - \mathcal{G}\|_F^2, \end{aligned} \quad (6)$$

其中 μ 和 ρ 实际上是惩罚参数, 这些参数通过使用文献 [27] 中建议的自适应更新策略进行调整. 我们根据文献 [29] 中提出的优化方法解决每个子问题.

更新的过程如下:

$\mathbf{Z}^{(v)}$ 子问题.

$$\mathbf{Z}^{(v)*} = (\mu \mathbf{I} + \mathbf{Y}_v + \rho \mathbf{G}^{(v)} - \mu \mathbf{P}^{(v)} - \mathbf{W}^{(v)}) / (\mu + \rho). \quad (7)$$

$\mathbf{P}^{(v)}$ 子问题.

根据第 v 个视图的 SVD 求解 $\mathbf{p}_i^{(v)*}$, 以获取 $\mathbf{P}^{(v)}$.

$$\mathbf{p}_i^{(v)*} = \begin{cases} \hat{\mathbf{p}}^{(v)}, & \text{如果 } \|[1/\sigma_1^{(v)}, \dots, 1/\sigma_{r_K}^{(v)}] \circ \mathbf{t}_u^{(v)}\| > 1/\tau, \\ \mathbf{c}_i^{(v)} - \mathbf{V}_K^{(v)} \mathbf{t}_u^{(v)}, & \text{其他情况.} \end{cases} \quad (8)$$

此处 $\mathbf{C}^{(v)} = \mathbf{I} - \mathbf{Z}^{(v)} + \mathbf{Y}_v/\mu$, $\mathbf{K}^{(v)} = \mathbf{V}^{(v)} \Sigma^{(v)2} (\mathbf{V}^{(v)})^T$, $\Sigma^{(v)} = \text{diag}([\sigma_1^{(v)}, \dots, \sigma_{r_K}^{(v)}, 0, \dots, 0]^T)$, 标量 $\{\sigma_i^{(v)}\}_{i=1}^{r_K^{(v)}}$ 为 $\mathbf{K}^{(v)}$ 的正奇异值, $r_K^{(v)}$ 为矩阵 $\mathbf{K}^{(v)}$ 的秩, $\mathbf{c}_i^{(v)}$ 和 $\mathbf{p}_i^{(v)}$ 分别为 $\mathbf{C}^{(v)}$ 和 $\mathbf{P}^{(v)}$ 的列, $\tau = \mu/\lambda$, $\mathbf{t}_u^{(v)} = \mathbf{V}_K^{(v)T} \mathbf{c}_i^{(v)}$, $\mathbf{V}_K^{(v)}$ 是 $\mathbf{V}^{(v)}$ 的前 $r_K^{(v)}$ 列形成的.

向量 $\hat{\mathbf{p}}^{(v)}$ 被定义为

$$\hat{\mathbf{p}}_i^{(v)} = \mathbf{c}_i^{(v)} - \mathbf{V}_K^{(v)} \left(\left[\frac{\sigma_1^{(v)2}}{\alpha^{(v)} + \sigma_1^{(v)2}}, \dots, \frac{\sigma_{r_K}^{(v)2}}{\alpha^{(v)} + \sigma_{r_K}^{(v)2}} \right]^T \circ \mathbf{t}_u^{(v)} \right), \quad (9)$$

$\alpha^{(v)}$ 是一个正标量, 满足以下等式:

$$\mathbf{t}_u^{(v)T} \text{diag} \left(\left\{ \frac{\sigma_i^{(v)2}}{(\alpha^{(v)} + \sigma_{r_i}^{(v)2})^2} \right\}_{1 \leq i \leq r_K^{(v)}} \right) \mathbf{t}_u^{(v)} = \frac{1}{\tau^2}. \quad (10)$$

当 $\|[1/\sigma_1^{(v)}, \dots, 1/\sigma_{r_K}^{(v)}] \circ \mathbf{t}_u^{(v)}\| > 1/\tau$ 时, 等式 (10) 中的 α 有唯一的正根, 可以用二分法求解得到. $\mathbf{G}^{(v)}$ 子问题.

$$\mathbf{g}^* = \arg \min_{\mathbf{g}} \|\mathbf{g}\|_{\otimes} + \frac{\rho}{2} \left\| \mathbf{g} - \left(\mathbf{z} + \frac{1}{\rho} \mathbf{W} \right) \right\|_F^2. \quad (11)$$

更新 $\mathbf{Y}^{(v)}$, \mathbf{W} , μ 和 ρ 的过程如下:

$$\mathbf{Y}_v^* = \mathbf{Y}_v + \mu(\mathbf{I} - \mathbf{Z}^{(v)} - \mathbf{P}^{(v)}), \quad (12)$$

$$\mathbf{W}^* = \mathbf{W} + \rho(\tilde{\mathbf{Z}} - \mathbf{G}), \quad (13)$$

$$\mu = \min(\eta\mu, \mu_{\max}), \quad (14)$$

$$\rho = \min(\eta\rho, \rho_{\max}), \quad (15)$$

其中 μ_{\max} 和 ρ_{\max} 是预设参数.

4.3 无监督多视角聚类方法

在解决了 4.2 小节中的 KT-SVD 问题之后, 相似度矩阵 \mathbf{A} 可通过以下方式获得:

$$\mathbf{A} = \frac{1}{V} \sum_{v=1}^V (|\mathbf{Z}^{(v)}| + |\mathbf{Z}^{(v)T}|)/2, \quad (16)$$

其中, $|\cdot|$ 表示取矩阵元素绝对值的操作. 接着, 通过下式使用相似度矩阵 \mathbf{A} 的谱聚类方法获得拉普拉斯 (Laplace) 矩阵 \mathbf{L} :

$$\mathbf{L} = \mathbf{D} - \mathbf{A}, \quad (17)$$

其中 \mathbf{D} 表示度矩阵, 定义为 $d_i = \sum_{j=1}^n a_{ij}$. 令 $U = [u_1, u_2, \dots, u_c] \in \mathbb{R}^{c \times n}$ 表示 \mathbf{L} 的前 c 个特征向量, 其中 c 是聚类个数. 最终得到的聚类标签即为审美偏好的识别结果.

4.4 输入 – 扰动关联映射方法

受文献 [17] 方法的启发, 我们提出了一种在无监督情况下将每个电极的特征与不同种类的审美偏好相关联的方法, 详细步骤如下: 为了得到频段组合和电极之间的关联关系, 我们采用同样的处理方法, 但是处理的数据由 EEG 信号变为 ERP 信号, 然后将得到的特征向量和聚类标签用于提出的输入 – 扰动关联映射方法中做进一步处理.

由于谱聚类获得的拉普拉斯矩阵 L 的特征向量携带着聚类信息, 我们将前 c 个特征向量组成的矩阵 U 视为特征矩阵, 用于描述 ERP 信号的特性.

定义噪声矩阵 $N_k \in \mathbb{R}^{s \times n}$, 其中 s 是电极数量 e 与 EEG 信号预处理完毕后的采样频率大小 f 的乘积, n 是 ERP 信号的数量. N_k 被用于第 k 次对 ERP 原始信号的扰动. 采用同样的方法处理扰动后的 ERP 信号, 获得本次扰动后的聚类标签以及特征矩阵 U_k . 为了将特征值的变化与审美偏好相关联, 定义矩阵 Q_k 为

$$Q_k = U_o^* - U_k^*, \quad (18)$$

其中, $U_o^* \in \mathbb{R}^{c \times c}$ 表示通过未经扰动的 ERP 信号获取的特征向量 U_o , 然后基于聚类标签的类内平均后得到的特征向量; 而 $U_k^* \in \mathbb{R}^{c \times c}$ 表示第 k 次对 ERP 信号扰动后, 通过扰动结果获取特征向量 U_k , 再进一步基于新的聚类标签的类内平均后得到的特征向量. 定义 $N_k^* \in \mathbb{R}^{s \times c}$ 表示 N_k 基于第 k 次聚类标签类内平均后的结果, 那么第 k 次噪声信号与特征值变化的关联矩阵 $C_k \in \mathbb{R}^{s \times c}$ 可以被定义为

$$C_k(N_k^*, Q_k) = \frac{\text{Cov}(N_k^*, Q_k)}{\sqrt{\text{Var}[N_k^*]\text{Var}[Q_k]}}, \quad (19)$$

最后, 平均所有的 C_k , 得到结果进一步在所有频率上叠加, 获得最终的电极与审美偏好关联矩阵 $C \in \mathbb{R}^{e \times c}$.

综上, 如果某个电极的特征在最终决策中起着至关重要的作用, 则其特征的变化也将反映在特征向量中. 通过扰动电极的特征, 结合 U 的变化, 我们可以在无监督的情况下研究电极与不同审美偏好之间的联系. 我们在算法 1 中总结了输入 – 扰动关联映射算法.

5 实验

本节设计了 3 类实验: (1) 基于 EEG 的分类实验. 为了验证所提出的 KT-SVD 在识别 EEG 数据方面表现良好, 我们首先将其应用于中文字体数据集的审美偏好识别中, 其中 KT-SVD 使用了基于 TDP 特征的 (δ , θ 和 α 频段) 组合视角获得了最高精度, 这意味着通过我们的方法获得的特征自表示向量可以很好地表达 EEG 数据. 同时, 我们将 TDP 特征和 BP 特征重新作为不同的视角数据, 在 EEG 的分类实验中进一步提升算法性能. (2) 输入 – 扰动关联映射的可视化实验. 为了以无监督的方式获得电极与中文字体激发的不同审美偏好之间的关系, 我们可视化了电极之间的关联关系. (3) 不同电极设置对于审美偏好识别性能的影响.

所有实验数据来源于 15 名受试者, 每个受试者总共进行 5 轮实验, 每轮实验中对 175 个汉字的刺激进行打分. 其中, 每轮实验中受试者挑出 12 个字符打分为喜欢, 挑出 12 个字符打分为不喜欢, 其余 151 个字符被归类为中性, 以得到最初的 EEG 数据. 所有 EEG 数据来源于 62 个电极, 每个汉字刺激产生的 EEG 数据的维度是 62×1200 . 所有信号在预处理后均下采样到 500 Hz 以降低复杂度. 去除无效数据后, 最终用于实验的汉字字体审美偏好数据集共计 8331 条. 在提取 BP 特征后数据维度变为 21600 维度, 提取 TDP 特征后数据维度变为 43200 维.

算法 1 Input-perturbation correlation mapping algorithm

Input: Multi-view feature matrices of ERP signal: $\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \dots, \mathbf{X}^{(V)}$, λ , cluster number c , electrodes number e , sampling frequency f , the best-combined result of frequency bands and time of perturbation K ;

Output: correlation matrix C

- 1: Initialize $\mathbf{Z}^{(v)} = \mathbf{Y}_v = 0$, $\mathbf{P}^{(v)} = I_n$, $i = 1, \dots, V$; $\mathcal{G} = \mathcal{W} = 0$;
- 2: $\mu = \rho = 10^{-5}$, $\mu_{\max} = 10^{13}$, $\varepsilon = 10^{-7}$;
- 3: **while** not converge **do**
- 4: Update subproblem $\mathbf{Z}^{(v)}$, \mathbf{P} and \mathcal{G} by using (7), (8) and (11) according to [29];
- 5: Obtain $\tilde{\mathbf{Z}} = \Phi(\mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}, \dots, \mathbf{Z}^{(V)})$;
- 6: Update $\mathbf{Y}^{(v)}$, \mathcal{W} and parameters μ , ρ by using (12)~(15), respectively;
- 7: **end while**
- 8: Obtain the affinity matrix \mathbf{A} by (16);
- 9: Obtain cluster labels and feature matrix U^o by applying the spectral clustering method with the affinity matrix \mathbf{A} , and obtain U_o^* at the same time;
- 10: **for** $k = 1 : K$ **do**
- 11: Perturbing EEG signals by applying the randomly generated noise matrix N_k ;
- 12: Replacing the multi-view feature matrix of the ERP signal with the perturbed result, and using steps 1~8 to obtain the affinity matrix \mathbf{A}_k ;
- 13: Obtain cluster labels and feature matrix U_k by applying the spectral clustering method with the affinity matrix \mathbf{A}_k , and obtain U_k^* at the same time;
- 14: Obtain Q_k by using (18), and obtain C_k by using (19);
- 15: **end for**
- 16: After taking the average of all C_k , and then superimpose it on all frequencies to generate the correlation matrix C ;
- 17: **return** correlation matrix C .

5.1 基于 EEG 的识别实验

为了验证算法的有效性,我们分别设计了两种不同的视角设定:以不同频段 EEG 的 TDP 特征作为多视角数据和以 EEG 的不同特征(TDP 特征和 BP 特征)作为多视角数据.识别实验数据的偏好标签来源于受试者标注,分为 3 类:喜欢、不喜欢和中性.我们将每个受试者的数据作为一个组,分别送入到模型中进行实验,最终的识别性能评估结果由 15 个组的结果进行平均后得到.

在第 1 种设定中,我们将预处理完毕后获得的汉字字体审美偏好 EEG 数据集通过 FFT 得到不同频段信号后提取特征,并将不同频段的特征数据视为不同的视角.将所有视角特征输入到 KT-SVD 的框架后得到每个视角的新的特征表示.在 EEG 识别实验中,将获得的特征表示送入我们的无监督多视角聚类方法中得到最终的审美偏好识别结果.表 1 显示了不同视角组合的性能,这有助于我们探索哪种频段组合最适合汉字字体审美偏好识别任务,可以观察到基于 TDP 特征的视角组合(δ , θ 和 α 频段)在各项指标中均获得了最高性能.

表 2 给出了单个频段的平均识别结果.首先将单个频段提取的特征数据计算出相似度矩阵,然后送入谱聚类中得到最终的识别结果.平均的识别性能同样来自于 15 个受试者数据的平均值.其中 γ 频段实现了最高的平均识别性能.在文献 [30] 中,EEG 的 γ 频段适用于以情感图像作为刺激的情感分类.文献 [31] 提出了更高频率的大脑活动反映情绪和认知过程.由于我们采用图片形式呈现不同种类的中文字体来获得中文字体数据集的审美偏好,因此我们的实验结果包含了之前的神经科学研究 [30,31] 结果.但是,从表 1 中,我们可以看到最佳组合是 δ (1~3 Hz), θ (4~7 Hz) 和 α (8~13 Hz) 频段,这似乎与我们从中得出的结论相矛盾.但是,这正好可以解释为什么最佳组合是较低的频段而不是较高的频段(β (14~30 Hz) 和 γ (31~50 Hz) 频段)的组合.这是因为,KT-SVD 算法通过将来自多

表 1 不同视角组合的平均识别性能

Table 1 Average recognition results for the combination of different views

View	$\delta\theta$	$\delta\alpha$	$\delta\beta$	$\delta\gamma$	$\theta\alpha$	$\theta\beta$	$\theta\gamma$	$\alpha\beta$	$\alpha\gamma$	$\beta\gamma$	$\delta\theta\alpha$	$\delta\theta\beta$	$\delta\theta\gamma$
NMI	0.759	0.800	0.710	0.797	0.681	0.621	0.734	0.503	0.749	0.666	0.891	0.757	0.823
ACC	0.894	0.938	0.860	0.934	0.856	0.835	0.914	0.770	0.921	0.869	0.971	0.884	0.945
AR	0.777	0.833	0.725	0.828	0.701	0.646	0.774	0.523	0.789	0.700	0.917	0.771	0.856
F-score	0.851	0.888	0.816	0.885	0.800	0.763	0.849	0.681	0.859	0.797	0.944	0.846	0.903
Precision	0.851	0.887	0.815	0.884	0.799	0.761	0.848	0.680	0.858	0.796	0.944	0.846	0.904
Recall	0.850	0.889	0.817	0.885	0.801	0.764	0.850	0.681	0.860	0.798	0.943	0.846	0.903

View	$\delta\alpha\beta$	$\delta\alpha\gamma$	$\delta\beta\gamma$	$\theta\alpha\beta$	$\theta\alpha\gamma$	$\theta\beta\gamma$	$\alpha\beta\gamma$	$\delta\theta\alpha\beta$	$\delta\theta\alpha\gamma$	$\delta\theta\beta\gamma$	$\delta\alpha\beta\gamma$	$\theta\alpha\beta\gamma$	$\delta\theta\alpha\beta\gamma$
NMI	0.797	0.840	0.808	0.767	0.804	0.714	0.739	0.845	0.858	0.833	0.842	0.777	0.875
ACC	0.916	0.950	0.939	0.909	0.938	0.892	0.906	0.937	0.957	0.948	0.952	0.920	0.963
AR	0.814	0.868	0.839	0.734	0.835	0.744	0.774	0.800	0.884	0.863	0.870	0.803	0.899
F-score	0.876	0.911	0.892	0.862	0.890	0.829	0.849	0.906	0.922	0.909	0.913	0.868	0.930
Precision	0.876	0.911	0.892	0.861	0.889	0.828	0.849	0.905	0.922	0.908	0.913	0.868	0.932
Recall	0.876	0.912	0.893	0.862	0.889	0.830	0.849	0.907	0.921	0.909	0.912	0.869	0.931

表 2 单个频段的平均识别性能

Table 2 Average recognition results for single band

View	δ	θ	α	β	γ
NMI	0.202	0.199	0.199	0.212	0.220
ACC	0.557	0.561	0.570	0.571	0.585
AR	0.124	0.130	0.133	0.138	0.157
F-score	0.472	0.472	0.469	0.471	0.487
Precision	0.392	0.396	0.399	0.402	0.416
Recall	0.596	0.587	0.573	0.573	0.579

个非线性子空间的多视图数据特征用核函数映射到可能线性分布的高维空间, 然后在该高维空间对低秩表示进行优化, 从而达到多视图子空间聚类的目的. KT-SVD 算法遵循在多种视角上寻求最大共识的共识原则, 正是由于它们组成的频率范围中较高的频段比较低的频段包含更多的信息, 因此在较高频段之间找到共识表示变得更加困难.

对于第 2 种设定, 在数据准备阶段, 将单个审美偏好 EEG 原始信号直接提取 TDP 特征或 BP 特征. 此时, TDP 特征数据和 BP 特征数据被视为基于不同特征的多视角数据用于本实验. 对于单个视角特征的评估, 采用类似单频段的方法. 对于多视角特征的评估, 将基于不同特征的多视角数据送入 KT-SVD 中获取新的特征表示, 再将该表示用于 EEG 的识别实验. 如表 3 所示, 基于 TDP 特征和 BP 特征的多视角组合比仅使用一种特征的识别性能显著提升. 虽然基于多特征的多视角组合比基于多频段的多视角组合的性能略有提升, 但是多特征的多视角表示中不能体现频段与审美偏好之间的联系, 所以在我们的总体框架中采用多频段作为多个视角.

5.2 可视化输入 – 扰动关联实验

为了找到跨受试者的共享神经信号, 我们对由受试者 1~15 合成的 ERP 数据进行了可视化输入 –

表 3 不同特征组合的平均识别性能

Table 3 Average recognition results for the combination of different features

View	TDP	BP	TDP&BP
NMI	0.220	0.227	0.920
ACC	0.597	0.573	0.979
AR	0.191	0.145	0.939
F-score	0.488	0.476	0.959
Precision	0.444	0.406	0.959
Recall	0.549	0.579	0.958

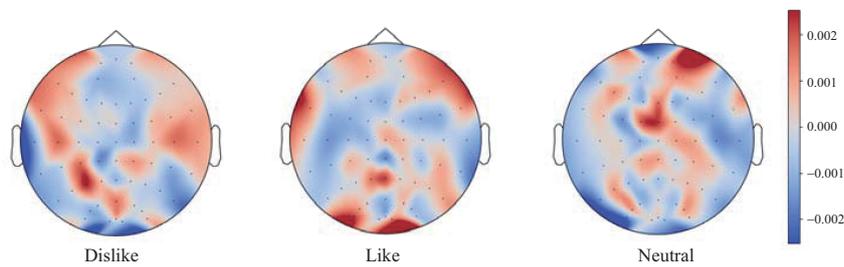


图 3 (网络版彩图) 喜欢、不喜欢和中性状态下 ERP 波的输入 - 扰动关联图

Figure 3 (Color online) Input-perturbation correlation map of ERP waves in terms of Dislike, Like and Neutral preference

扰动关联实验分析. 由于 ERP 波可以看作跨对象的通用波, 同时产生 ERP 波的过程也可以使得 EEG 噪声最小化, 这有助于揭示审美偏好的潜在模式. 一次刺激诱发的 ERP 波幅比 EEG 小得多, 其信息被淹没在 EEG 中, 两者构成了小信号与大噪音的关系. 所以几乎所有 ERP 研究都依靠某种平均处理过程来将 EEG 噪声降至最低. 同时, 又由于作为 ERP 背景的 EEG 波形与刺激间无固定的关系, 而其中所含的 ERP 波形在每次刺激后是相同的, 且 ERP 波形与刺激间的时间间隔 (潜伏期) 是固定的, 随着越来越多的平均实验, 平均波形中残留的噪声越来越小. 从数学上讲, 如果 R 是单次实验的噪声量, 而 M 是实验的次数, 则 M 次平均实验后的噪声大小等于 $(1/\sqrt{M}) \times R$. 此外, 由于假设信号不受平均过程的影响, 所以信噪比随实验次数的 \sqrt{M} 倍增加.

具体来说, 实验分为两部分: ERP 数据的形成和可视化输入 - 扰动关联实验. (1) ERP 数据的形成: 首先将所有受试者的 EEG 数据按照喜欢、不喜欢和中性偏好分为 3 个大组, 每个大组中保留 640 条数据, 然后将每个大组分为 50 个小组. 采用留一法处理平均过程, 生成 ERP 数据. 即对于每个大组, 该过程进行 50 次, 每次取出 1 组然后平均其余组以生成 ERP 数据. 最终形成的 ERP 数据集合计 150 条. (2) 可视化输入 - 扰动关联实验: 根据第 4 节中的方法, 将获得的每组数据的最佳频率组合 (δ , θ 和 α 频段) 特征作为多视角数据, 同时设置扰动次数参数 $K = 400$, 噪声函数定义为高斯 (Gauss) 噪声 (其中, 均值为 0, 方差为 1). 使用提出的输入 - 扰动关联算法获取的 ERP 波的输入 - 扰动关联图, 如图 3 所示. 从图 3 中, 我们可以观察到: 电极 PZ, PO7, F8 和 FT7 与喜欢的审美偏好相关, 电极 P3, C5, C6 和 CP3 与不喜欢的审美偏好相关. 就中性审美偏好而言, 电极 FCZ, FC1, FP2 和 FZ 显示出相对较高的关联性.

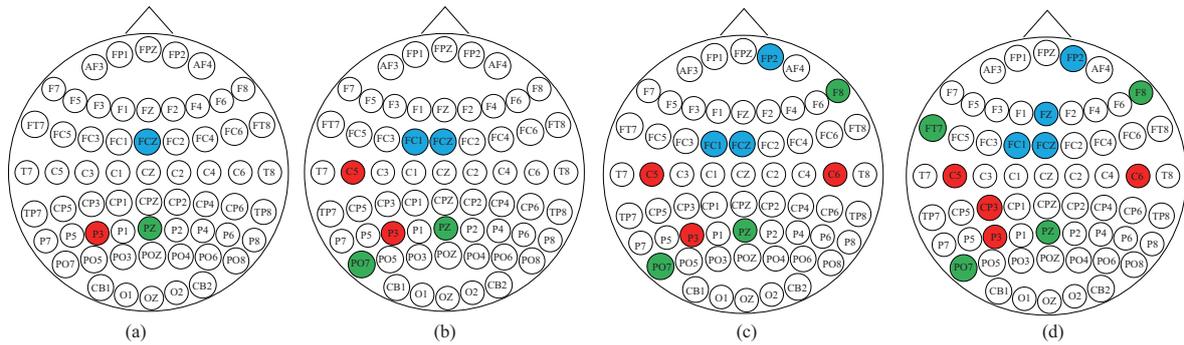


图 4 (网络版彩图) 4 种电极放置方式. 绿色、红色和蓝色分别表示与喜欢、不喜欢以及中性的审美状态相关
Figure 4 (Color online) The green color means this electrode correlates with like preference. The red color is more correlated with dislike preference and the blue one correlates with neutral preference

表 4 不同电极组合的 SVM 平均分类性能

Table 4 Classification accuracies of SVM for different electrode sets with 10-fold cross-validation

Selected electrode numbers	ACC (BP) (%)	ACC (TDP) (%)
62	74.71	70.25
12	79.13	84.58
9	84.13	83.50
6	80.29	82.83
3	89.17	89.16

5.3 电极设置的影响

文献 [19] 中的研究表明, 存在一些不相关的通道, 这些通道可能不仅会增加计算复杂度, 还会包含噪声. 基于上面的结果, 我们希望通过调整电极的设置进行如下探讨: (1) 进一步探讨能够反映与汉字字体的审美偏好相关的大脑活动模式中的关键电极; (2) 探讨能否可以通过较小的电极池反映潜在的审美偏好模式 (喜欢、不喜欢和中性) 而不会降低分类精度. 在实验中, 我们根据 ERP 波的输入 - 扰动关联图设计了 4 种电极设置方案, 如图 4 所示, 我们分别从喜欢、不喜欢和中性的输入 - 扰动关联图中选择了 Top-1, Top-2, Top-3 和 Top-4 的电极. 电极数为 62 的设置方案是指直接采用了 62 个电极的原始 EEG 数据. 为了实验的公平起见, 本小节实验过程都是在选定的电极池中采集到的原始 EEG 信号上分别直接提取 TDP 和 BP 特征, 再将它们送入到 SVM 分类模型中获取识别结果. 实验中, 我们将原始的 8331 条数据随机分成 10 份, 每次选择其中的 9 份作为训练和 1 份作为测试, 采用 10 折交叉验证的方式获得平均分类性能.

从表 4 中可以看出, 无论是 BP 特征, 还是 TDP 特征, 都是只采用了 3 个电极 (PZ, P3 和 FCZ) 的原始 EEG 信号的识别准确度最高, 分别为 89.17% 和 89.16%. 此外, 具有 6 个, 9 个和 12 个电极组的特征表示的识别精度均比 62 个电极的特征表示具有更好的性能, 这表明仅来自少数几个关键电极的 EEG 信号就能够提供足够的信息, 以区分不同的审美偏好. 并且通过减少电极数量不仅可以提高模型的鲁棒性, 还可以降低计算成本.

图 5 中显示了由喜欢字符、不喜欢字符和中性字符引起的选定电极的总平均 ERP 响应. 图中的第 1 行包括从喜欢的审美偏好中选择的关键电极, 以 PZ 和 PO7 为例, 我们可以观察到在 400~600 ms 的窗口中 3 个偏好响应之间存在显著差异, 这与晚正电位成分 (late positive potential, LPP) 有

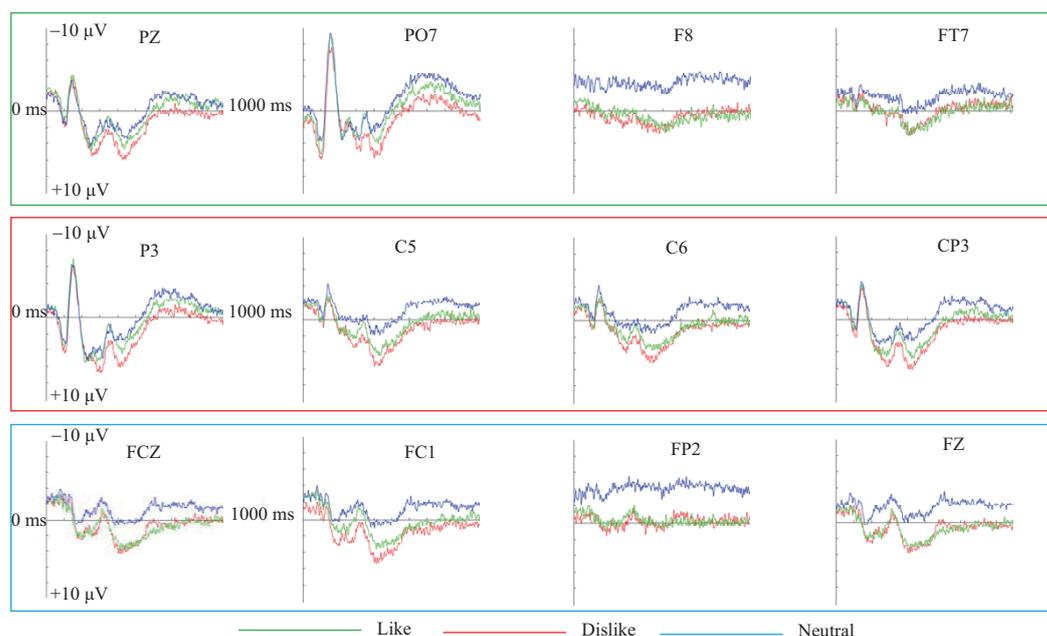


图 5 (网络版彩图) 从 3 种审美偏好 (喜欢/不喜欢/中性) 中选择的电极的总平均 ERP 响应

Figure 5 (Color online) The grand-average ERP responses of selected electrodes from three esthetic preferences (like/dislike/neutral)

关. 但是我们从电极 F8 和 FT7 中观察不到相同的现象, 这有可能是在使用 9 和 12 个电极时精度降低的原因. 图中的第 2 行由不喜欢的审美偏好中选择的关键电极组成, 在 400~600 ms 窗口中所有的关键电极均能观察到明显的差异. 图中的第 3 行由中性的审美偏好的关键电极组成, 在所有选择的电极中, 中性偏好的波形幅度比其余两种审美偏好明显具有区别性.

6 结论

我们提出了一种用于基于 EEG 的汉字字体审美偏好类别 (喜欢、不喜欢和中性) 识别的无监督核化张量奇异值分解的多视角框架: KT-SVD. KT-SVD 通过将不同频段的信号视为不同的视角探索最佳频段组合. 实验结果表明, 经过 δ (1~3 Hz), θ (4~7 Hz) 和 α (8~13 Hz) 的频段多视角组合达到了最高的识别性能. 最后, 我们还提出了一种输入 - 扰动关联策略来分析获取关键电极, 以降低计算成本并提高模型的鲁棒性, 采用 3 种关键电极组合配置的原始 EEG 信号特征获得了 89.17% (BP) 和 89.16% (TDP) 的平均识别精度, 比采用 62 种电极的原始 EEG 信号特征的识别精度 74.71% (BP) 和 70.25% (TDP) 高得多. 同时本文得出的结论与文献 [9] 中得出的结论基本一致, 即本文方法获取的关键电极主要分布在额叶和顶叶区域, 并且在刺激开始后的 400~600 ms 窗口中, 关键电极的各种偏好的区分更为明显.

参考文献

- 1 Zhang J S. A survey of digital calligraphy. *Sci Sin Inform*, 2019, 49: 143–158 [张俊松. 数字书法研究综述. 中国科学: 信息科学, 2019, 49: 143–158]
- 2 Bargh J A, Ferguson M J. Beyond behaviorism: on the automaticity of higher mental processes. *Psychological Bull*, 2000, 126: 925–945

- 3 Höfel L, Jacobsen T. Electrophysiological indices of processing aesthetics: spontaneous or intentional processes? *Int J Psychophysiol*, 2007, 65: 20–31
- 4 Handy T C, Smilek D, Geiger L, et al. ERP evidence for rapid hedonic evaluation of logos. *J Cognitive Neurosci*, 2010, 22: 124–138
- 5 Sutton S, Tuetting P, Zubin J, et al. Information delivery and the sensory evoked potential. *Science*, 1967, 155: 1436–1439
- 6 Jacobsen T, Höfel L. Aesthetics electrified: an analysis of descriptive symmetry and evaluative aesthetic judgment processes using event-related brain potentials. *Empirical Studies Arts*, 2001, 19: 177–190
- 7 Barton J J S, Sekunova A, Sheldon C, et al. Reading words, seeing style: the neuropsychology of word, font and handwriting perception. *Neuropsychologia*, 2010, 48: 3868–3877
- 8 Nakatsuka M, Thabit M N, Koganemaru S, et al. Writing’s shadow: corticospinal activation during letter observation. *J Cognitive Neurosci*, 2012, 24: 1138–1148
- 9 Li R, Qin R, Zhang J, et al. The esthetic preference of Chinese typefaces — an event-related potential study. *Brain Res*, 2015, 1598: 57–65
- 10 Liu L, Chen R, Wolf L, et al. Optimizing photo composition. *Comput Graph Forum*, 2010, 29: 469–478
- 11 Redies C, Amirshahi S A, Koch M, et al. PHOG-derived aesthetic measures applied to color photographs of artworks, natural scenes and objects. In: *Proceedings of European Conference on Computer Vision*. Berlin: Springer, 2012. 522–531
- 12 Hagoort P, Hald L, Bastiaansen M, et al. Integration of word meaning and world knowledge in language comprehension. *Science*, 2004, 304: 438–441
- 13 Hellige J B, Adamson M M. Hemispheric differences in processing handwritten cursive. *Brain Language*, 2007, 102: 215–227
- 14 Lu Y, Zheng W L, Li B, et al. Combining eye movements and EEG to enhance emotion recognition. In: *Proceedings of the 24th International Joint Conference on Artificial Intelligence*, 2015. 1170–1176
- 15 Zheng W L, Lu B L. Personalizing EEG-based affective models with transfer learning. In: *Proceedings of the 25th International Joint Conference on Artificial Intelligence*, 2016. 2732–2738
- 16 Stober S, Cameron D J, Grahn J A. Using convolutional neural networks to recognize rhythm stimuli from electroencephalography recordings. In: *Proceedings of the Neural Information Processing Systems*, 2014. 1449–1457
- 17 Schirrmester R T, Springenberg J T, Fiederer L D J, et al. Deep learning with convolutional neural networks for EEG decoding and visualization. *Hum Brain Mapp*, 2017, 38: 5391–5420
- 18 Li Y, Zheng W, Cui Z, et al. A novel neural network model based on cerebral hemispheric asymmetry for EEG emotion recognition. In: *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, 2018. 1561–1567
- 19 Bos D O. EEG-based emotion recognition. *Influence Visual Auditory Stimuli*, 2006, 56: 1–17
- 20 Nie D, Wang X W, Shi L C, et al. EEG-based emotion recognition during watching movies. In: *Proceedings of the 5th International IEEE/EMBS Conference on Neural Engineering*, 2011. 667–670
- 21 Wang X W, Nie D, Lu B L. Emotional state classification from EEG data using machine learning approach. *Neurocomputing*, 2014, 129: 94–106
- 22 Vidaurre C, Krämer N, Blankertz B, et al. Time domain parameters as a feature for EEG-based brain-computer interfaces. *Neural Netw*, 2009, 22: 1313–1319
- 23 Jenke R, Peer A, Buss M. Feature extraction and selection for emotion recognition from EEG. *IEEE Trans Affective Comput*, 2014, 5: 327–339
- 24 Xie Y, Tao D, Zhang W, et al. On unifying multi-view self-representations for clustering by tensor multi-rank minimization. *Int J Comput Vis*, 2018, 126: 1157–1179
- 25 Cristianini N, Shawe-Taylor J, Kandola J S. Spectral kernel methods for clustering. In: *Proceedings of the Neural Information Processing Systems*, 2002. 649–655
- 26 Hofmann T, Schölkopf B, Smola A J. Kernel methods in machine learning. *Ann Statist*, 2008, 36: 1171–1220
- 27 Lin Z, Chen M, Ma Y. The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices. 2010. ArXiv: 1009.5055
- 28 Lin Z, Liu R, Su Z. Linearized alternating direction method with adaptive penalty for low-rank representation. In: *Proceedings of the Neural Information Processing Systems*, 2011. 612–620

- 29 Qu Y, Liu J, Xie Y, et al. Robust kernelized multi-view self-representations for clustering by tensor multi-rank minimization. 2017. ArXiv: 1709.05083
- 30 Li M, Lu B L. Emotion classification based on Gamma-band EEG. In: Proceedings of the 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2009. 1223–1226
- 31 Ray W J, Cole H W. EEG Alpha activity reflects attentional demands, and Beta activity reflects emotional and cognitive processes. *Science*, 1985, 228: 750–752

Esthetic preference mining of Chinese typefaces using multiview cluster analysis

Yan ZHANG^{1,5†}, Yuan XIE^{2†}, Chen HONG¹, Yanyun QU^{1*}, Rui LI³, Junsong ZHANG⁴ & Cuihua LI¹

1. *Computer Science Department, School of Informatics, Xiamen University, Xiamen 361005, China;*
2. *School of Computer Science and Technology, East China Normal University, Shanghai 200062, China;*
3. *National Engineering Laboratory for Technology of Big Data Applications in Education, Central China Normal University, Wuhan 430079, China;*
4. *Mind, Art & Computation Group, Department of Artificial Intelligence, School of Informatics, Xiamen University, Xiamen 361005, China;*
5. *School of Mathematics Science, Guizhou Normal University, Guiyang 550001, China*

* Corresponding author. E-mail: yyqu@xmu.edu.cn

† Equal contribution

Abstract Previous neuroesthetic studies have proved that Chinese typefaces can be viewed as an esthetic preference stimulus by observing differences in event related potential (ERP) waves among three preferences, namely, like, dislike, and neutral. We first reconfirm this conclusion by introducing a multiview clustering method of kernelized tensor singular value decomposition (KT-SVD) to construct an esthetic preference recognition model based on electroencephalograms (EEGs). Our approach regards data from different frequency bands as different views describing the esthetic preferences of Chinese fonts, explore the relevance of all view features through the constraint of tensor multi-rank minimization, and obtains the esthetic preferences using the clustering results. Additionally, the input-perturbation correlation method is used to correlate the amplitude of the electrodes with different types of esthetic preferences and describe the relationship between the key frequency-band combinations and electrodes, and take out the electrodes most relevant to likes, dislikes, and neutrality, including 3 electrodes of Top-1, 6 electrodes of Top-2, 9 electrodes of Top-3, and 12 electrodes of Top-4, forming four different combinations of EEG features for esthetic preference recognition experiments. Experimental results show that the method based on multiview clustering can solve the correlation analysis of neural signals and esthetic preferences and mine the electrodes most relevant to the esthetic preferences of fonts.

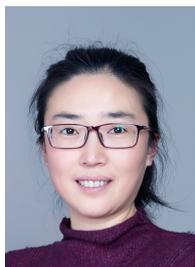
Keywords Chinese typeface, esthetic evaluation, computational esthetics, event-related potentials, kernelized tensor-SVD, data mining



Yan ZHANG received her B.S. degree in Computer Science and Technology from Guizhou Normal University, China, in 2001 and M.S. degree in Software Engineering from Guizhou University, China, in 2006. Since 2015, she has been pursuing her Ph.D. degree at Xiamen University in Computer Science and Technology. Her research interests include image processing, computer vision, object classification, and object detection algorithms.



Yuan XIE was born in 1985. He received his Ph.D. degree in Pattern Recognition and Intelligent Systems from the Institute of Automation, Chinese Academy of Sciences (CAS), in 2013. He is currently a full professor at the School of Computer Science and Technology, East China Normal University, Shanghai, China. His research interests include image processing, computer vision, machine learning, and pattern recognition.



Yanyun QU was born in 1972. She received her Ph.D. degree in Pattern Recognition and Intelligent Systems from the Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University, Xi'an, in 2006. Currently, she is a professor at the School of Informatics of Xiamen University. Her research interests include image processing, computer vision, machine learning, and pattern recognition.



Junsong ZHANG received his Ph.D. degree in Computer Science from the State Key Lab of CAD&CG, Zhejiang University, Hangzhou, China, in 2008. He is currently an associate professor in the Mind, Art and Computation Group, Cognitive Science Department, Xiamen University, Xiamen, China. His main research interests include computer graphics, computational aesthetics, and cognitive science.