

DOI: 10.19650/j.cnki.cjsi.J2108346

## 基于改进 MEDA 算法的脑电情绪识别\*

何群,李冉冉,付子豪,江国乾,谢平

(燕山大学电气工程学院 河北省测试计量技术及仪器重点实验室 秦皇岛 066004)

**摘要:**针对普通机器学习算法与迁移学习在应用方面的局限性,利用改进流形嵌入分布对齐算法(MEDA)算法解决跨被试情绪识别中准确率低的问题。其中MEDA通过流行特征变换来减小域之间的数据漂移,并能够自适应定量估计边缘分布和条件分布的权重大小。针对特征维度大且有可能存在不良特征的问题,提出改进MEDA算法,即引入改进最小冗余最大相关算法用于特征选择,并对多源域下的多组识别结果进行决策级融合,进一步提升迁移学习效果。在SEED数据集和实测数据对该算法验证,改进MEDA算法相比于支持向量机、迁移成分分析和联合分布适配算法,整体识别精度分别提升了8.97%、4.00%、2.89%,改进的MEDA算法相比于改进前,每个被试识别准确率均有提升的同时整体识别提升3.36%,验证了该方法的有效性。

**关键词:**情绪识别;特征选择;迁移学习;流行嵌入分布对齐算法

中图分类号:TH79 文献标识码:A 国家标准学科分类代码:510.40

## EEG emotion recognition based on the improved MEDA

He Qun, Li Ranran, Fu Zihao, Jiang Guoqian, Xie Ping

(Key Laboratory of Measurement Technology and Instrumentation Hebei Province, Institute of Electric Engineering, Yanshan University, Qinhuangdao 066004, China)

**Abstract:** The limited applications of the traditional machine learning algorithms and the transfer learning algorithm are considered in this study. The improved manifold embedded distribution alignment (MEDA) algorithm is utilized to improve the detection accuracy in the cross-subject emotion recognition. The MEDA algorithm in the manifold space could reduce the data drift between domains by popular feature transformation, which can adaptively and quantitatively estimate the weights of edge distribution and conditional distribution. This article proposes an improved manifold space distribution alignment algorithm to address the problems of large feature dimension and possible bad features. An improved minimum redundancy maximum correlation algorithm is introduced for feature selection. The computational complexity is reduced, the associated features are selected, and the decision-level fusion on multiple groups of recognition results in multi-source domain is performed to further improve the transfer learning effect. The analysis results of SEED data set and the measured data set show that the distribution alignment algorithm in the manifold space is better than those of the support vector machine, transfer component analysis and joint distribution adaptation. The overall recognition accuracy is improved by 8.97%, 4.00%, and 2.89%, respectively. The improved distribution alignment algorithm in manifold space has improved the recognition accuracy of each subject, and the overall recognition accuracy is improved by 3.36%. Therefore, the effectiveness of the proposed method is verified.

**Keywords:** emotion recognition; feature selection; transfer learning; manifold embedded distribution alignment algorithm

### 0 引言

情绪是综合了人的感觉、思想和行为的一种状态,在人们的日常工作和生活中,情绪的作用无处不在。情绪

识别在医疗护理、虚拟现实游戏、交通安全等方面具有重要的应用前景<sup>[1-2]</sup>。在安全交通领域<sup>[3]</sup>,情绪识别技术可用于保障操作者安全驾驶的辅助检测系统中;在医疗健康领域<sup>[4]</sup>,患病的情绪状态对恢复效果有很大的影响。传统的情感识别方法基本上是利用视觉和听觉来推测人

收稿日期:2021-07-31 Received Date: 2021-07-31

\* 基金项目:国家自然科学基金(U20A20192,62076216)项目资助

类的情绪状态,例如面部表情,语音和肢体手势等,许良凤等<sup>[5]</sup>提出一种完全局部二值模式和幂指数加权融合方法用于语音情感识别。但是这些外在表现对情绪的表征不够细腻,且易于伪装。由于脑电信号的自发性,使其能够反映真实的情绪状态,基于脑电信号的情绪识别成为情绪脑机接口研究的重要部分。在进行情绪识别时,若直接对脑电信号进行识别处理,结果往往较差,所以在进行识别分类前需要进行脑电信号特征的提取,以便获取能够有效反映情绪状态的特征,保证后续工作的准确性。目前在情绪脑电的特征选择上国内外研究者已经做了大量工作,并找到一些有效区分不同情绪状态的脑电特征。Nie等<sup>[6]</sup>首先通过带通滤波将脑电信号转换到5个频带上,然后通过短时傅里叶变换计算出每个通道上的各个频带的频谱能量,并以此作为脑电的频域特征来进行情绪识别。Stikic等<sup>[7]</sup>通过傅里叶变换提取5个频段上的功率谱密度作为脑电特征进行积极和消极的情绪识别。苗敏敏等<sup>[8]</sup>首先根据导联重要性进行空间电极的选择,然后通过稀疏回归算法对微分熵特征进行优化选择,从而实现情绪识别。同时,深度学习在情绪识别中也有所应用。柳长源等<sup>[9]</sup>通过多尺度注意力残差网络实现对脑电信号的情绪识别。赵亮等<sup>[10]</sup>拟将不同脑区通道、不同频段的EEG信号之间的Granger因果系数差值作为特征进行情感识别。

由于脑电信号的特征维数较高,导致模型训练需要较大的时间开销。为了有效地降低特征维数,Zheng等<sup>[11]</sup>在基于EEG的情绪识别任务中,提出使用群稀疏典型相关性分析(group sparse canonical correlation analysis, GSCCA)进行EEG通道的选择。Hu等<sup>[12]</sup>在基于EEG的注意力识别任务中使用了基于相关性的特征选择方法。

此外,由于脑电信号具有非平稳性和个体差异性,不同被试在同种情绪状态下的脑电信号分布存在较大差异,这种差异性往往较大而不能被忽略。传统基于脑电的情绪识别一部分没有考虑到差异性的存在,导致模型

的泛化能力较差;另一部分则是充分考虑到这种差异性的存在,针对各被试脑电信号构建独立识别分类模型,但是通过机器学习构建模型需要大量训练数据及其标签作为支撑,而采集数据尤其是为数据标定标签存在耗时大,成本高的问题,导致构建模型的可行性较差。由于传统的机器学习要求源域和目标域数据服从独立同分布的条件,所以传统机器学习具有很大的局限性,而迁移学习通过映射拉近源域和目标域分布、减小分布差异,可以将其他被试采集的源域数据应用到新被试的目标域数据上,构建适合目标域的识别分类模型。Zheng等<sup>[13]</sup>通过迁移成分分析法(transfer component analysis, TCA)和核主成分分析法(kernel principle component analysis, KPCA)进行3分类识别分别达到了69.44%和77.96%的识别精度,Ni等<sup>[14]</sup>利用提出的域自适应稀疏表示分类器进行脑电情绪识别。

实际中对目标域数据的学习往往有多个源域数据,考虑到单源域迁移学习不仅会限制识别性能,并且如果领域间相关性较低还会产生负迁移现象<sup>[15]</sup>,导致模型识别精度下降。针对现有算法的局限性,本文引入流形嵌入分布对齐算法<sup>[16]</sup>(manifold embedded distribution alignment, MEDA)并对其进行改进,即在流形特征变换后进行特征选择,之后进行自适应分布对齐,再将多源域下得到的多组识别结果送入决策级融合,降低整体计算复杂度的同时有效保障脑电情绪识别精度。

## 1 情绪识别整体框架

本文进行的迁移学习情绪识别流程如图1所示,首先将采集到的脑电信号进行200 Hz降采样,并通过0.5~70 Hz的带通滤波剔除噪声和伪迹。对预处理后的脑电信号提取分类特征,然后通过改进MEDA算法进行迁移学习识别分类得到多组结果,最后通过决策级融合算法得到最终识别结果。

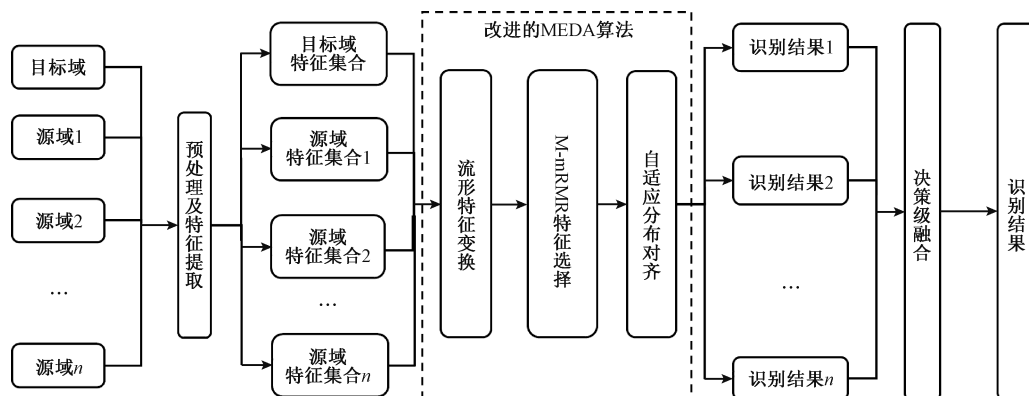


图1 情绪识别整体流程

Fig. 1 Flowchart of emotion recognition

如何对特征进行选择, 剔除不良特征, 并对高维度特征进行筛选降维, 得到相似性强且具有高可分性的特征集合, 并依此对传统 MEDA 算法进行有效改进, 使其具有良好的迁移学习能力, 提高跨个体情绪识别的准确性是本文算法的研究重点。

## 2 改进 MEDA 算法

考虑到源域和目标域经过特征提取后可能得到的特征集合维度较大, 其中不仅包含能够有效进行迁移学习的分类特征, 也存在一些无用特征及一些不利于迁移学习的不良特征, 同时特征集合维度过大会增加计算复杂度, 降低整体效率, 所以本文提出了一种改进的最小冗余最大相关 (mmd max-relevance and min-redundancy, M-mRMR) 算法, 该算法对源域数据和目标域数据的特征集合进行联合选择, 目的是选择出适合进行迁移学习的有效分类特征, 降低后续计算复杂度, 提高整体识别精度。

进一步, 在 MEDA 算法的基础上引入 M-mRMR 算法, 提出改进 MEDA 算法。考虑到流形特征变换可能会使得特征间关系产生变化, M-mRMR 算法应该添加到 MEDA 算法中流形特征变换之后, 特征选择后再进行自适应分布对齐。

### 2.1 传统 MEDA 算法

MEDA 算法的基本原理如图 2 所示。首先进行流行特征学习减小域间数据漂移, 应对退化特征变换的问题, 然后进行动态分布对齐对边缘和条件分布的重要性进行定量分析, 应对未评估的分布对齐的挑战, 最后通过结构化风险最小化 (structural risk minimization, SRM) 原则总结表示学习领域不变分类器  $f$ 。

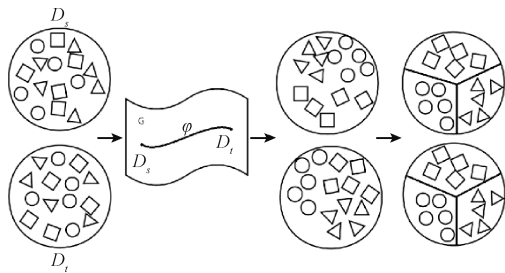


图 2 MEDA 原理图

Fig. 2 MEDA schematic diagram

#### 1) 流形特征变换

考虑到在流形空间中的特征通常都有很好的几何性质, 可以避免特征扭曲, MEDA 首先将原始空间下的特征变换到流形空间中, 在 MEDA 中选择的是测地线流式核方法 (geodesic flow kernel, GFK)<sup>[17]</sup>。在 GFK 方法中需

要确定子空间维度  $d$ , 其采用空间不一致度量 (subspace disagreement measure, SDM)<sup>[17]</sup> 确定源域目标域能保持最大一致性的子空间维度大小, 其计算方式如下:

首先针对给定的源域数据  $D_s$  和目标域数据  $D_t$ , 通过主成分分析分别对其进行处理得到  $P_s$  和  $P_t$ , 同时, 将  $D_s$  和  $D_t$  合并得到一个数据集  $D_{s+t}$ , 并对其进行主成分分析处理得到  $P_{s+t}$ 。如果两个数据集相似度高, 那么  $P_s, P_t$  和  $P_{s+t}$  间的距离应该比较小。

分别计算两个数据集和合并数据集的空间夹角, 记数据集  $P_s, P_t$  与  $P_{s+t}$  间的夹角分别  $\alpha_d$  和  $\beta_d$ , 用  $D(d) = \frac{1}{2}[\sin\alpha_d + \sin\beta_d]$  表示两个夹角的总度量。如果  $D(d)$  很小表示两个数据集距离很小,  $D(d)$  最大时表示两个空间垂直, 数据集间距离最大。然后通过贪心算法在不使两个空间正交的前提下尽可能取得最大的  $d$  值以保证得到更多的子空间个数。

#### 2) 动态分布对齐

当源域数据和目标域数据的特征分布情况不同时, 边缘分布适配和条件分布适配的重要性也是不同的。MEDA 自适应动态调节边缘分布和条件分布的重要性, 拉近两域间联合概率分布大小。其中自适应的分布适配  $\overline{D}_f$  可以被表示为:

$$\overline{D}_f(D_s, D_t) = (1 - \mu)D_f(P_s, P_t) + \mu \sum_{c=1}^C D_f^{(c)}(Q_s, Q_t) \quad (1)$$

其中,  $\mu \in [0, 1]$  是自适应因子调节边缘分布和条件分布的重要性, 当其趋于 0 时, 表示源域数据和目标域数据集特征分布存在较大的差异, 边缘分布适配更重要; 当其趋于 1 时, 表示源域数据和目标域数据集特征分布较为相似, 条件分布适配较为重要。  $c \in \{1, \dots, C\}$  是类别。  $D_f(P_s, P_t)$  代表边缘分布,  $D_f^{(c)}(Q_s, Q_t)$  代表  $c$  类的条件分布。

由于目标域数据  $D_t$  不携带标签, 无法直接得到其条件概率分布  $Q_t = Q_t(y_t | z_t)$ , 而在样本个数足够多的前提下类条件概率  $Q_t(z_t | y_t) \approx Q_t$ , 通过在源域数据  $D_s$  上训练一个弱分类器对目标域数据  $D_t$  进行预测得到目标域数据的软标签以近似  $Q_t(z_t | y_t)$ , 并通过不断迭代修正预测结果。

MEDA 通过 MMD 来衡量源域和目标域概率分布间的差异性。概率分布和间的 MMD 计算公式如下:

$$d^2(p, q) = (E_p[\varphi(z_s)] - E_q[\varphi(z_t)])_{HK}^2 \quad (2)$$

式中:  $HK$  为由特征映射  $\varphi(\cdot)$  所张成的再生核希尔伯特空间,  $E[\cdot]$  为嵌入样本的均值。

为了保证 MMD 与分类器间的一致性, 通过映射的 MMD 距离对边缘分布差异进行计算。

$$D_f(P_s, P_t) = \|E[f(z_s)] - E[f(z_t)]\|_{HK}^2 \quad (3)$$

同理,条件分布差异可以被表示为:

$$D_f^{(c)}(Q_s, Q_t) = \| \mathbf{E}[f(\mathbf{z}_s^{(c)})] - \mathbf{E}[f(\mathbf{z}_t^{(c)})] \|_{HK}^2 \quad (4)$$

然后,自适应分布适配可以被表示为:

$$\overline{D}_f(D_s, D_t) = (1 - \mu) \| \mathbf{E}[f(\mathbf{z}_s)] - \mathbf{E}[f(\mathbf{z}_t)] \|_{HK}^2 + \mu \sum_{c=1}^C \| \mathbf{E}[f(\mathbf{z}_s^{(c)})] - \mathbf{E}[f(\mathbf{z}_t^{(c)})] \|_{HK}^2 \quad (5)$$

MEDA 通过采用来估计不同分布之间的距离,进而估计  $\mu$ , 其中的计算如下:

$$\mu \approx 1 - \frac{d_M}{d_M + \sum_{c=1}^C d_c} \quad (6)$$

式中:  $d_M$  为边缘分布差异的 A-distance,  $d_c$  为标签类别  $c$  的数据条件分布差异的 A-distance。

然后根据 SRM 原则,则分类器  $f$  可以表示为:

$$f = \operatorname{argmin}_f \left( \sum_{i=1}^n H_k(f(g(x_i)), y_i) + \eta \|f\|_K^2 + \lambda \overline{D}_f(D_s, D_t) + \rho R_f(D_s, D_t) \right) \quad (7)$$

式中:  $\operatorname{argmin}_f \left( \sum_{i=1}^n H_k(f(g(x_i)), y_i) + \eta \|f\|_K^2 \right)$  为  $f$  在源域数据

上的损失,  $H_k$  为核函数构造的希尔伯特空间,  $g(\cdot)$  为特征流形变换过程,  $L$  为平方损失函数  $L(f(x), y) = (y - f(x))^2$ ,  $x_i$  和  $y_i$  为特征向量及其对应的类别标签,  $\|f\|_K^2$  为  $f$  的平方标准形式,  $\overline{D}_f(\cdot, \cdot)$  为动态分布对齐方式,  $R_f(\cdot, \cdot)$  为拉普拉斯正则化,以进一步利用流型  $G$  中最近点的相似几何特性,  $\eta$ ,  $\lambda$  和  $\rho$  是正则化参数。

## 2.2 改进的最小冗余最大相关算法

M-mRMR 算法在保证源域和目标域相似的基础上,通过对源域相关性的分析来获取目标域上相关性最大的 index。算法流程如图 3 所示,算法步骤如下:

1) 首先衡量源域数据和目标域数据各维度特征间的相似性大小,然后根据相似性大小对特征进行排序,考虑到相似性高的特征更有利于后续学习,所以选取源域和目标域数据间相似性高的前  $n$  维特征。

2) 然后根据 1) 中处理后的源域数据及其标签进行相关性分析,根据相关性大小进行排序,从  $n$  维特征中选取相关性高的前  $m$  维特征,组成最终的源域数据和目标域数据。

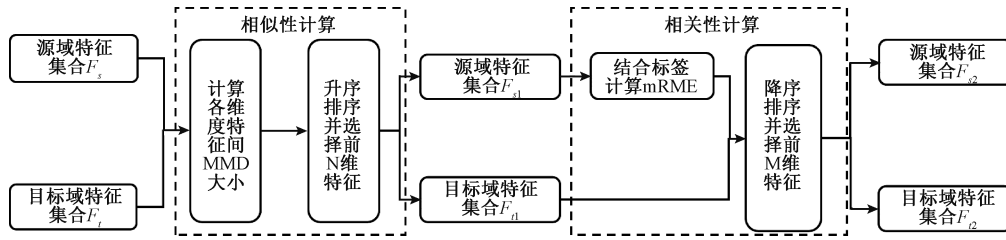


图 3 M-mRMR 算法流程

Fig. 3 Flow of M-mRMR algorithm

### (1) 脑电特征相似性分析

计算源域数据和目标域数据同一维度特征间的相似度,常见的分布度量函数有相对熵 (kullback-leibler divergence, KL)<sup>[18]</sup>、最大均值差异 (maximum mean discrepancy, MMD)<sup>[19]</sup>。相对熵的计算通常需要提前知晓样本分布,常常难以满足,而 MMD 可以用源域和目标域在再生希尔伯特空间中的均值差异来近似,计算较为便捷高效。本文通过源域数据和目标域数据各维度特征间的 MMD 大小作为其相似性的衡量标准。

假设源域数据和目标域数据的特征维度为  $N$ , 那么源域数据中的特征  $F_s^i (i < N)$  和目标域数据中的特征  $F_t^i (i < N)$  之间的 MMD 大小为:

$$MMD^2(F_s^i, F_t^i) = \left\| \frac{1}{N^s} \sum_{n=1}^{N^s} f(x_n^s) - \frac{1}{N^t} \sum_{j=1}^{N^t} f(x_j^t) \right\|_{HK}^2 = \operatorname{tr}(KL) \quad (8)$$

式中:  $N^s$  和  $N^t$  分别为源域数据和目标域数据的样本数量,  $f(x_n^s)$  和  $f(x_j^t)$  分别为样本  $x_n^s$  和  $x_j^t$  经的第  $i$  维特征过

核变换以后得特征,  $K_i$  为源域数据第  $i$  维特征和目标域数据第  $i$  维特征间的核矩阵,表达式如下:

$$K_i = \begin{bmatrix} K_i^{s,s} & K_i^{s,t} \\ K_i^{t,s} & K_i^{t,t} \end{bmatrix} \quad (9)$$

$L$  矩阵如下:

$$L = \begin{cases} \frac{1}{N^s N^s}, & x_n, x_j \in D^s \\ \frac{1}{N^t N^t}, & x_n, x_j \in D^t \\ -\frac{1}{N^s N^t}, & \text{其他} \end{cases} \quad (10)$$

通过上述计算方法对源域数据特征集  $F_s = \{f_1^s, f_2^s, \dots, f_{n-1}^s, f_n^s\}$  和目标域数据特征集  $F_t = \{f_1^t, f_2^t, \dots, f_{n-1}^t, f_n^t\}$  间各个维度的特征进行计算,然后可以得到特征相似性集合  $L_f = \{l_1, l_2, \dots, l_{n-1}, l_n\}$ 。其中,  $l_m$  为第  $m$  维特征间的相似性大小,值越小代表此维度特征越相似。

然后对相似性集合  $L_f$  从小到大进行排序,得到有序的相似性集合  $L_{fs} = \text{Sort}(L_f)$ ,从排序好的相似性集合  $L_{fs}$  中选择前  $N$  个相似性对应的特征维度  $index$ ,从源域数据特征集和目标域数据特征集中按上述  $index$  进行选择得到新源域数据特征集  $F_{s1}$  和新目标域数据特征集  $F_{t1}$ 。

## (2) 脑电特征相关性分析

在得到新源域数据特征集  $F_{s1}$  后,结合其数据标签通过最大相关最小冗余算法(max-relevance and min-redundancy, mRMR)计算进行第 2 次特征选择,保证所选特征与标签间具有强相关性。

mRMR 的核心思想是,通过计算特征与类别标签间的相关性  $D$  和特征与特征间的相关性  $R$ ,然后选择  $D$  相对较大、 $R$  相对较小的特征<sup>[20]</sup>。其中,利用互信息衡量特征与类别标签间以及特征与特征间的相关性,计算公式如下:

$$D(S, c) = \frac{1}{|S|} \sum_{x_i \in S} I(x_i; c) \quad (11)$$

$$R(S) = \frac{1}{|S|^2} \sum_{x_i, x_j \in S} I(x_i, x_j) \quad (12)$$

式中:  $S$  为特征集合,  $c$  为类别标签,  $I(x_i; c)$  为第  $i$  维特征和类别标签  $c$  之间的互信息大小,  $I(x_i, x_j)$  为第  $i$  维特征和第  $j$  维特征之间的互信息大小,  $D(S, c)$  为特征集  $S$  和类别标签  $c$  之间的相关性,  $R(S)$  为集合  $S$  中所有特征冗余。

给定两个随机变量  $x$  和  $y$ , 其互信息为:

$$I(x; y) = \iint p(x, y) \log \frac{p(x, y)}{p(x)p(y)} dx dy \quad (13)$$

其中,  $p(x)$ 、 $p(y)$  为  $x$ 、 $y$  的边缘概率分布函数,  $p(x, y)$  为  $x$  和  $y$  的联合概率分布函数。

mRMR 是特征与类别标签间相关性和不同特征间冗余度的相互协调,定义如下:

$$mRMR = \max_S (D - R) \quad (14)$$

使用增量搜索方法可以写成优化问题:

$$\max_{x_i \in S - S_m} \left[ I(x_i; c) - \frac{1}{m} \sum_{x_j \in S_m} I(x_i; x_j) \right] \quad (15)$$

通过 mRMR 计算源域数据特征集  $F_{s1}$  各维度特征和其标签  $Y_i$  间的相关性大小,得到相关性集合  $M = \{m_1, m_2, \dots, m_{N-1}, m_N\}$ ,然后对  $M$  从大到小进行排序得到  $M_s$ ,然后选择  $M_s$  前  $m$  个相关性大小对应的特征维度  $index$ ,从源域数据特征集  $F_{s1}$  和目标域数据特征集  $F_{t1}$  中按上述  $index$  进行选择得到新源域数据特征集  $F_{s2}$  和新目标域数据特征集  $F_{t2}$ 。

综上所述, M-mRMR 算法通过计算源域数据和目标域数据特征间相似度作为第 1 次判别条件,然后计算源域数据和其标签间的相关性作为第 2 次判别条件,得到的特征集合能够有效反映出源域数据和目标域数据共享信息,且具有更好的可分性。

## (3) 决策级融合

由于对所选的多组源域数据分别构建了迁移学习分类模型,会得到多组目标域数据的识别分类结果。直接对多组结果的准确精度作和取平均作为最终识别准确精度会忽略各结果中的有效信息,导致识别精度较差,所以本文采用投票法进行最终的决策级融合,其融合模型如图 4 所示。

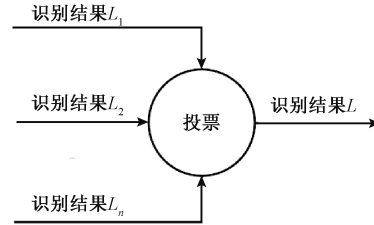


图 4 决策级融合

Fig. 4 Decision level fusion

决策级融合算法计算方式如下,即:

$$F = \text{MAX} \begin{cases} \sum_{i=1}^n (f_i), & f_i = 1 \\ \sum_{i=1}^n (f_i/2), & f_i = 2 \\ \sum_{i=1}^n (f_i/3), & f_i = 3 \end{cases} \quad (16)$$

式中:  $n$  为所选源域序号,  $f_i$  为第  $i$  个源域的识别结果(1: 消极, 2: 中性, 3: 积极),  $F$  代表最终识别结果,其中投票占比最大的即为最后得识别结果。

## 3 测试结果分析

### 3.1 测试数据集

本文使用的数据为公开数据集 SEED(SJTU emotion EEG dataset)<sup>[21-22]</sup>。根据国际 10~20 系统采集了 62 个通道的脑电信号,采样频率为 1 000 Hz。SEED 数据集记录了 7 名男生、8 名女生共 15 名被试观看不同类型视频时所采集的脑电数据。其中每名被试参加 3 次实验,每次间隔时间大约为 1 周,共得 45 组数据。单次实验包含 15 个情绪诱发任务(积极、中立及消极情绪各 5 个),每个任务均包含 5 s 的提示时间、大约 4 min 的视频观影时间、45 s 的自我评估时间和 15 s 的休息时间。将初始脑电信号降采样到 200 Hz,并将每 1 s 的实验数据作为一个样本,最终每名被试每次实验可以得到 3 394 个样本。在 5 个频段上分别提取脑电信号的微分熵,每个样本共有 310 维特征。

微分熵(differential entropy, DE)是对时不变连续源信息熵的度量。对预处理后的脑电信号各通道信号的

delta(1~4 Hz)、theta(4~8 Hz)、alpha(8~13 Hz)、beta(13~30 Hz)、gamma(36~44 Hz) 5个频段进行微分熵特征提取。其中微分熵计算的方法如下:

$$DE(X) = \frac{1}{2} \log(2\pi e \sigma^2) \quad (17)$$

式中:  $\sigma^2$  表示脑电信号时间序列的方差。

实测数据来自4位精神状态良好,身体健康的在校大学生,2名男生,2名女生,年龄在25~27岁之间。按照国际10~20标准佩戴neuracle脑电帽,采样频率为1000 Hz。情绪诱发视频包括3类,积极、中性和消极,所选素材都是

取自国语电影。每类视频有5段,每段影片播放时间为300 s。自我评估时间为45 s,休息时间为15 s。对实测数据的前期数据处理方式与SEED数据集一致。

### 3.2 M-mRMR 有效性分析

为了验证本文所提M-mRMR算法的有效性,将情绪脑电SEED数据集中的被试1的第1组数据作为目标域数据,被试15的第1组数据作为源域数据,通过M-mRMR算法进行处理,特征维度设为60,得到的特征曲线如图5所示。

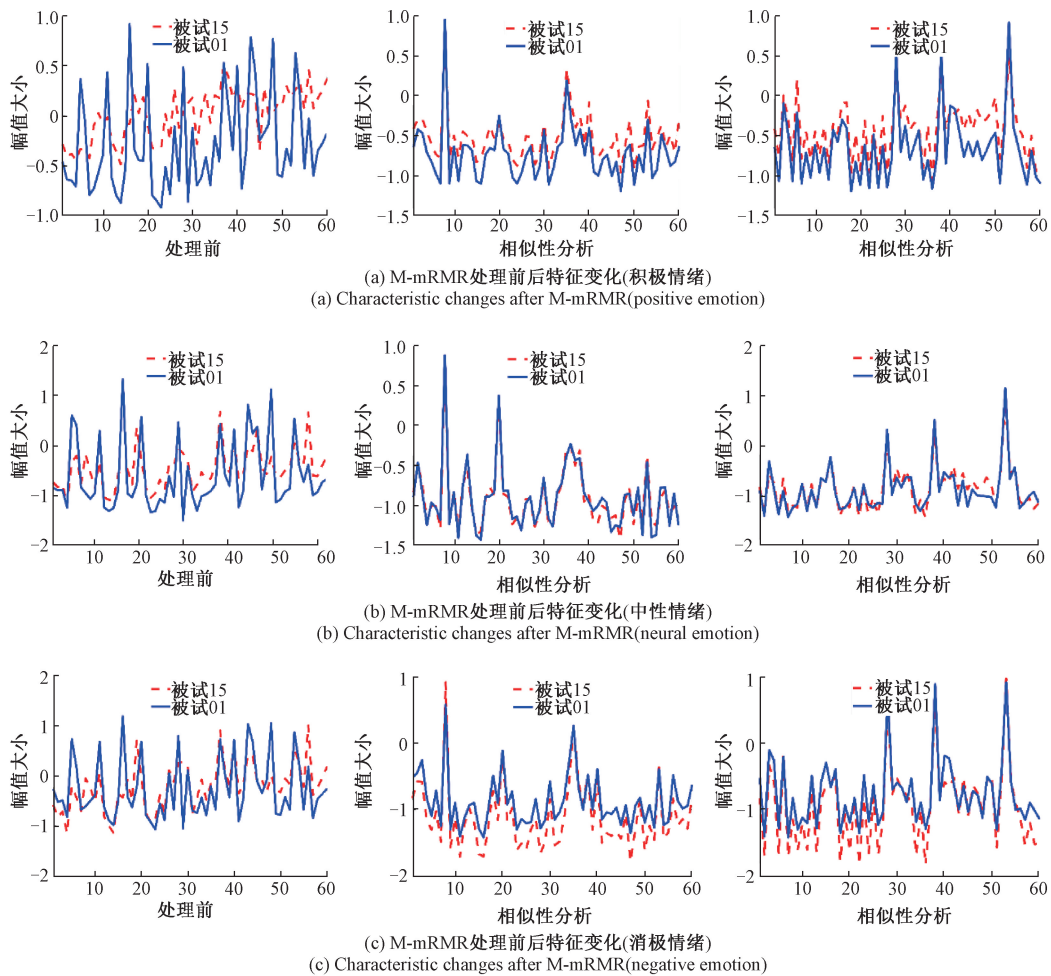


图5 M-mRMR 处理前后特征变化

Fig.5 Characteristic changes after M-mRMR

图5中包含积极、中性和消极3种情绪状态下的特征曲线。每横向3张为一组,每组中第1张为处理前的特征曲线,第2张为相似性处理后的特征曲线,第3张为相关性处理后得特征曲线,可以看出处理前源域和目标域特征分布差异较大,而处理后能够选择出相似性强、相关性高的特征,符合理论预期,验证了本文所提M-mRMR算法的有效性。

### 3.3 MEDA 性能分析

为验证本文所应用的MEDA算法的有效性,通过与支持向量机算法(support vector machine, SVM)<sup>[23]</sup>、迁移成分分析算法(transfer component analysis, TCA)<sup>[24]</sup>和联合分布适配算法(joint distribution adaptation, JDA)<sup>[25]</sup>进行对比分析。其中,SVM通过源域数据构建分类器并对目标域数据进行识别分类,TCA将源域和目标域变换到

相同空间,最小化其边缘概率分布差异,JDA 分别适配源域和目标域的边缘概率分布和条件概率分布。而 MEDA 结合流形特征变换和动态分布适配,能够减小域间数据漂移并自适应估计两种分布的权重。

不同被试同情绪状态间的数据分布存在很大差异,使用传统机器学习方法在一个被试上训练的分类模型对其他被试数据进行识别分类时准确率较低。在 SEED 数据集中依次选择每名被试的第 1 次实验数据作为目标域,其余的 14 名被试的第 1 次实验数据作为该目标域对应的不同源域,最终可以得到 14 组预测结果,通过决策级融合算法得到该目标域最终的预测标签及准确率。考虑到脑电信号的非稳定性,同被试不同时间的信号存在差异性,所以实验时只取每位被试的第 1 组数据用于实验。整个操作重复 15 次以保证每名被试都作为目标域进行过实验。以被试 15 为目标域为例进行决策级融合,结合 MEDA 算法实验结果如表 1 所示。

表 1 决策级融合与平均值对比

Table 1 Comparison of decision level fusion and average %

被试序号	准确率	被试序号	准确率
被试 1	95.70	被试 8	55.33
被试 2	62.17	被试 9	74.10
被试 3	83.94	被试 10	60.46
被试 4	78.85	被试 11	82.73
被试 5	77.17	被试 12	44.61
被试 6	54.90	被试 13	67.47
被试 7	92.99	被试 14	83.59
平均值			72.43
决策融合			91.40

由表 1 可知,以被试 15 为目标域时经过计算平均值得到结果为 72.43%,而经过决策级融合得到的结果为 91.40%,决策级融合得到的准确率比使用平均值高 18.97%,下面对全部 15 组数据进行测验,结果如图 6 所示。

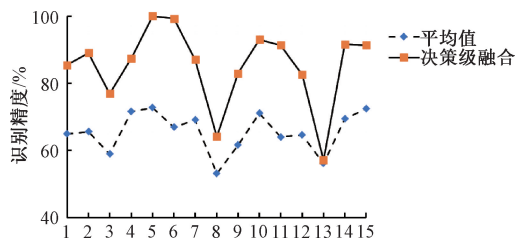


图 6 平均值与决策级融合结果对比

Fig. 6 Comparison of decision level fusion and average

由图 6 可以看出,经过决策级融合处理得到的识别精度高于直接求平均得到的识别精度,证明了决策级融合算法的优越性。

经过实验,SVM、TCA、JDA 和 MEDA 这 4 种算法的平均识别准确率如表 2 所示。其中 SVM 的参数  $c$  从 [0.001 0.01 0.1 1 10 100] 中寻优,TCA 和 JDA 核函数选择  $rbf$  核,参数  $gamma$  和  $lamda$  从 [0.01 0.1 1 10] 中寻优,特征维度  $d$  从 [20 40 60 80 100 120 140 160] 中寻优,在 MEDA 中流形变换时需要将原始的  $D$  维数据降低到  $d$  维度,结合空间不一致度量 (subspace disagreement measure, SDM) 计算  $d$  的大小,同时在 [20 40 60 80 100 120 140 155] 中寻优,参数  $lambda$  从 [0.01 0.1 1 10] 中寻优。

表 2 不同算法整体识别精度

Table 2 Recognition accuracy of different algorithms %

分类算法	处理方式	整体识别精度
SVM	归一化	66.11
	标准化	76.07
TCA	归一化	78.55
	标准化	81.04
JDA	归一化	80.22
	标准化	82.15
MEDA	归一化	81.03
	标准化	85.04

常用的数据变换方式有两种:归一化和标准化。其中,归一化是将数据变换到区间 [0,1] 之间,而标准化则是将数据变化为均值为 0,方差为 1 的分布。由表 2 可知,特征集合经过标准化后整体识别精度更高,同时由于 SVM 算法的前提是数据独立同分布,因此跨个体的识别分类效果较差,TCA 算法仅拉近源域和目标域的边缘分布差异,结果相较于 SVM 有所提升,JDA 算法以相同重要性同时拉近源域和目标域间的边缘分布和条件分布差异,结果相较于 TCA 有所提升,而 MEDA 算法结合流形特征变换减小域间数据漂移,同时进行自适应分布适配、最小化损失函数、流形正则化,具有较好的识别效果。

### 3.4 改进的 MEDA 性能分析

为验证本文所提的改进 MEDA 算法的有效性,使用 3.3 节中的数据,通过与传统 MEDA 算法进行对比分析,结果如表 3 所示。特征选择 M-mRMR 中特征选择维度从 [40 60 80 100 120 160] 中寻优。

由表 3 可得,改进的 MEDA 在不同被试作为目标域时,每个受试者的识别结果都有不同程度得提高,最高可提升 5.51%,整体平均识别精度提升了 3.36%,同时方

表3 改进 MEDA 与 MEDA 结果对比

Table 3 MEDA algorithm comparison of results before and after improvement %

被试序号	MEDA	M-mRMR+MEDA	准确率提升
被试 1	85.50	91.01	5.51
被试 2	89.07	93.46	4.39
被试 3	76.87	78.55	1.68
被试 4	87.48	89.04	1.56
被试 5	98.92	99.47	0.55
被试 6	98.46	98.83	0.37
被试 7	85.07	85.91	0.84
被试 8	64.11	70.71	6.60
被试 9	82.97	85.12	2.15
被试 10	93.08	96.70	3.62
被试 11	91.37	95.32	3.95
被试 12	82.59	84.24	1.65
被试 13	57.07	70.17	13.10
被试 14	91.66	96.08	4.42
被试 15	91.40	91.46	0.06
平均值	85.04	88.40	3.36
var	11.46	9.32	-2.14

差平均降低 2.14%，验证了本文所提改进 MEDA 算法的有效性，在保证提高准确度的同时提高稳定性。

针对实测数据进行分析，实测数据共包括 4 名被试，依次指定其中 1 名被试的实验数据作为目标域，其余 3 名被试的实验数据作为源域，以此构建分类模型进行情绪识别分类。对各类算法进行对比分析，结果如表 4 所示。

表4 实测数据结果分析表

Table 4 Analysis of measured data results %

被试	SVM	JDA	MEDA	M-mRMR-MEDA
被试 1	69.23	72.85	76.21	81.42
被试 2	58.75	63.20	67.58	75.63
被试 3	63.11	68.18	72.43	79.03
被试 4	66.38	72.17	78.17	83.88
平均值	64.38	69.27	73.60	79.85
Std	3.90	3.99	4.04	2.81

由表 4 可知，本文所提的改进 MEDA 算法不仅在各被试上具有较高的识别准确度，同时整体识别精度较高，标准差较小。可以看出本文所提出的改进 MEDA 算法在实测数据上也有较好的分类性能。

本次实验环境：Intel (R) Xeon (R) CPU E5-2630 v4 @ 2.20 GHz, 64 G 内存, 64 位 Windows 系统, 算法的实现采用 MATLAB2019a。计算时间成本如表 5 所示。

表5 计算时间成本对比

Table 5 Comparison of calculation time cost

算法	特征数	时间/s	精度/%
MEDA	300	319	85.04
改进 MEDA	60	310	88.40

由表 5 可知，改进的 MEDA 算法相比改进前，在提高精度的同时降低了计算成本，体现了该算法的优越性。

### 3.5 同类对比

将本文提出的方法与同样采用 SEED 数据集的其他方法的实验结果作对比。文献 [26] 从脑电信号中提取 18 种线性和非线性脑电信号特征，并从中自动选取与情感最相关的特征信息，利用支持向量机达到 83.33% 的识别率。文献 [27] 使用最大独立领域自适应算法 (maximum independence domain adaptation, MIDA)，以差分熵作为特征取得 72.31% 的分类精度。文献 [28] 将针对时间序列的卷积深度域自适应模型 (convolutional deep domain adaptation model for time series data, CoDATS) 引入跨被试情感脑电分类问题中，最高精度达到 80.99%。对比结果如表 6 所示。

表6 同类研究对比

Table 6 Comparison of similar studies

对比文献	精度/%	分类数
文献 [26]	83.33	3
文献 [27]	72.31	3
文献 [28]	80.99	3
本文	88.40	3

## 4 结 论

本文提出一种基于改进流形空间下分布对齐迁移算法的脑电情绪识别方法。首先选择 SEED 数据集中提取的微分熵作为分类特征，然后引入流形空间下的分布对齐算法进行识别分类，进一步针对特征维度大，存在不良特征的问题，提出改进最小冗余最大相关算法用于特征选择，并将其引入流形嵌入分布对齐算法中提出改进流形嵌入分布对齐算法用于进行识别分类，最后将各源域下的识别结果进行决策级融合。实验结果表明，流形空间下的分布对齐算法相较于支持向量机算法、迁移成分



分析算法和联合分布自适应算法识别精度分别提升了 8.97%、4.00%、2.89%,改进流形空间下分布对齐算法相较于改进前,每个被试识别精度均有不同程度的提升且整体提升了 3.36%。因此,本文提出的基于改进流形空间下的分布对齐迁移算法的脑电情绪识别方法能够有效提升跨个体情绪识别的准确性和鲁棒性,为基于脑电的情绪识别的实际应用提供新方法。

### 参考文献

- [ 1 ] KALLIPOLITIS A, GALLIAKIS M, MENYCHTAS A, et al. Affective analysis of patients in homecare video-assisted telemedicine using computational intelligence[J]. *Neural Computing and Applications*, 2020, 32(23): 17125-17136.
- [ 2 ] HALIM Z, REHAN M. On identification of driving-induced stress using electroencephalogram signals: A framework based on wearable safety-critical scheme and machine learning [J]. *Information Fusion*, 2020, 53: 66-79.
- [ 3 ] LIU W, QIAN J, YAO Z, et al. Convolutional two-stream network using multi-facial feature fusion for driver fatigue detection [J]. *Future Internet*, 2019, 11(5): 115.
- [ 4 ] HUANG H, XIE Q, PAN J, et al. An EEG-based brain computer interface for emotion recognition and its application in patients with disorder of consciousness[J]. *IEEE Transactions on Affective Computing*, 2019, 12(4): 832-842.
- [ 5 ] 许良凤, 刘泳海, 胡敏, 等. 语谱图改进完全局部二值模式的语音情感识别 [J]. *电子测量与仪器学报*, 2018, 32(5): 30-37.
- XU L F, LIU Y H, HU M, et al. Speech emotion recognition based on improved complete local binary pattern of spectrogram images [J]. *Journal of Electronic Measurement and Instrumentation*, 2018, 32(5): 30-37.
- [ 6 ] NIE D, WANG X W, SHI L C, et al. EEG-based emotion recognition during watching movies [C]. *International IEEE/EMBS Conference on Neural Engineering*, Cancun: IEEE, 2011: 667-670.
- [ 7 ] STIKIC M, JOHNSON R R, TAN V, et al. EEG-based classification of positive and negative affective states [J]. *Brain-Computer Interfaces*, 2014, 1(2): 99-112.
- [ 8 ] 苗敏敏, 徐宝国, 胡文军, 等. 基于自适应优化空频微分熵的情感脑电识别 [J]. *仪器仪表学报*, 2021, 42(3): 221-230.
- MIAO M M, XU B G, HU W J, et al. Emotion EEG recognition based on adaptive optimized space frequency differential entropy [J]. *Chinese Journal of Scientific Instrument*, 2021, 42(3): 221-230.
- [ 9 ] 柳长源, 孙雨涵, 李文强, 等. 基于 MAREsnet 的脑电情感识别研究 [J]. *仪器仪表学报*, 2020, 41(7): 235-242.
- LIU CH Y, SUN Y H, LI W Q, et al. Research on EEG emotion recognition based on MAREsnet [J]. *Chinese Journal of Scientific Instrument*, 2020, 41(7): 235-242.
- [ 10 ] 赵亮, 方芳, 王伟, 等. 基于脑电信号间 Granger 因果关系特征的情感识别 [J]. *电子测量与仪器学报*, 2018, 32(6): 87-95.
- ZHAO L, FANG F, WANG W, et al. Emotion recognition based on Granger causality feature between EEG signals [J]. *Journal of Electronic Measurement and Instrumentation*, 2018, 32(6): 87-95.
- [ 11 ] ZHENG W. Multichannel EEG-based emotion recognition via group sparse canonical correlation analysis [J]. *IEEE Transactions on Cognitive and Developmental Systems*, 2017, 9(3): 281-290.
- [ 12 ] HU B, LI X, SUN S, et al. Attention recognition in EEG-based affective learning research using CFS+KNN algorithm [J]. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2018, 15(1): 38-45.
- [ 13 ] ZHENG W L, ZHANG Y Q, ZHU J Y, et al. Transfer components between subjects for EEG-based emotion recognition [C]. *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, 2015: 917-922.
- [ 14 ] NI T, NI Y, XUE J, et al. A domain adaptation sparse representation classifier for cross-domain electroencephalogram-based emotion classification [J]. *Frontiers in Psychology*, 2021, 12: 721266.
- [ 15 ] PAN S J, YANG Q. A survey on transfer learning [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2010, 22(10): 1345-1359.
- [ 16 ] WANG J, FENG W, CHEN Y, et al. Visual domain adaptation with manifold embedded distribution alignment [J]. *ACM*, 2018: 402-410.
- [ 17 ] GONG B, YUAN S, FEI S, et al. Geodesic flow kernel for unsupervised domain adaptation [C]. *2012 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2012: 2066-2073.
- [ 18 ] KAPOOR R, GUPTA R, JHA S, et al. Boosting performance of power quality event identification with KL divergence measure and standard deviation [J]. *Measurement*, 2018, 126: 134-142.
- [ 19 ] YAN H, DING Y, LI P, et al. Mind the class weight bias: Weighted maximum mean discrepancy for

- unsupervised domain adaptation [ C ]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017;2272-2281.
- [20] PENG H, LONG F, DING C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy [ J ]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2005, 27(8): 1226-1238.
- [21] DUAN R N, ZHU J Y, LU B L. Differential entropy feature for EEG-based emotion classification [ J ]. IEEE, 2013: 81-84.
- [22] ZHENG W L, LU B L. Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks [ J ]. IEEE Transactions on Autonomous Mental Development, 2015, 7 ( 3 ): 162-175.
- [23] NOBLE W S. What is a support vector machine? [ J ]. Nature Biotechnology, 2006, 24(12): 1565-1567.
- [24] PAN S J, TSANG I W, KWOK J T, et al. Domain adaptation via transfer component analysis [ J ]. IEEE Transactions on Neural Networks, 2011, 22(2): 199-210.
- [25] LONG M, WANG J, DING G, et al. Transfer feature learning with joint distribution adaptation [ C ]. Proceedings of the 2013 IEEE International Conference on Computer Vision, 2013;2200-2207.
- [26] LI X, SONG D, ZHANG P, et al. Exploring EEG features in cross-subject emotion recognition [ J ]. Frontiers in Neuroscience, 2018, 12(1): 162.
- [27] LAN Z, SOURINA O, WANG L, et al. Domain adaptation techniques for EEG-based emotion recognition: A comparative study on two public datasets [ J ]. IEEE Transactions on Cognitive & Developmental Systems, 2019, 11(1): 85-94.
- [28] 高寒冰, 张驰, 靳明艳, 等. 基于多源域自适应的跨被试情感脑电识别 [ J ]. 智能科学与技术学报, 2021, 3(1): 59-64.
- GAO H B, ZHANG CH, JIN M Y, et al. Cross-subject emotional EEG recognition based on multi-source domain adaptation [ J ]. Chinese Journal of Intelligent Science and Technology, 2021, 3(1): 59-64.

### 作者简介



何群, 2001 年于燕山大学获得硕士学位, 2005 年于燕山大学获得博士学位, 现为燕山大学副教授, 主要研究方向为机械故障诊断、基于生理信号的辅助驾驶等。

E-mail: hq@ysu.edu.cn

**He Qun** received his M. Sc. degree and Ph. D. degree both from Yanshan University in 2001 and 2005, respectively. He is currently an associate professor at Yanshan University. His main research interests include mechanical fault diagnosis and assisted driving based on physiological signals.



谢平 (通信作者), 2006 年于燕山大学获得博士学位, 现为燕山大学教授, 主要研究方向为脑机接口、智能康复等。

E-mail: pingx@ysu.edu.cn

**Xie Ping** (Corresponding author) received her Ph. D. degree from Yanshan University in 2006. She is currently a professor at Yanshan University. Her main research interests include brain-computer interface and intelligent rehabilitation.