

分类号 \_\_\_\_\_  
学校代码 10487

学号 D201677605  
密级 \_\_\_\_\_

**华中科技大学**

# 博士学位论文

脑机接口中的迁移学习方法研究

学位申请人： 何赫

学科专业： 控制科学与工程

指导教师： 伍冬睿 教授

答辩日期： 2020年5月29日

**A Thesis Submitted in Partial Fulfillment of the Requirements  
For the Degree of Doctor of Philosophy in Engineering**

**Transfer Learning in Brain-Computer Interfaces**

**Candidate : He He**  
**Major : Control Science and Engineering**  
**Supervisor : Prof. Wu Dongrui**

Huazhong University of Science & Technology

Wuhan 430074, P. R. China

May, 2020

## 独创性声明

本人声明所呈交的学位论文是我个人在导师的指导下进行的研究工作及取得的研究成果。尽我所知，除文中已标明引用的内容外，本论文不包含任何其他人或集体已经发表或撰写过的研究成果。对本文的研究做出贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本人承担。

学位论文作者签名：何赫

日期：2020年 6 月 2 日

## 学位论文版权使用授权书

本学位论文作者完全了解学校有关保留、使用学位论文的规定，即：学校有权保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。本人授权华中科技大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

本论文属于  保密，在 \_\_\_\_\_ 年解密后适用本授权书。  
 不保密

(请在以上方框内打“√”)

学位论文作者签名：何赫

日期：2020年 6 月 2 日

指导教师签名：伍冬睿

日期：2020年 6 月 3 日

## 摘要

脑机接口可以为大脑和机器之间提供直接的信息交互通道，简单来讲，就是使用户通过意念控制外部设备，在生物、医疗和工程等领域都有着极广泛的应用场景。脑电图是脑机接口最常用的输入，通过在用户的头皮表面放置电极来采集脑电信号，然后基于算法模型从脑电信号中识别大脑的活动模式，从而读取大脑“意图”并转化为外部设备的控制命令。

但是，由于个体差异，即不同用户对相同的刺激会产生不同的神经反应，所以很难建立一个适用于所有用户的通用算法模型。因此，对每个新用户，在使用前通常都需要经过一段长时间的校准来采集标注数据，这在用户体验上不友好。迁移学习研究相邻领域间知识的迁移和复用，被用来解决这个问题。应用迁移学习可以弥补辅助用户（源用户）与新用户（目标用户）之间数据或者模型的差异，使得源用户的标注数据或者模型也可以用于目标用户，从而减少甚至消除目标用户的校准时间。

本文对脑机接口中的迁移学习方法进行了研究，由浅入深，由算法到数据，由简单场景到复杂场景，提出了几种迁移学习方法。首先考虑同构迁移学习场景，本文先提出一种结合迁移学习与空域滤波的算法，可以用于传统算法框架；然后提出两种优化方法，可以用于黎曼几何算法框架；再提出一种针对脑电数据的迁移学习方法，可以用于不同的算法和模型。最后考虑异构迁移学习场景，针对异构类别空间问题提出一种类别对齐的方法，不但可以用于不同的算法和模型，还可以与其它迁移学习方法结合使用。取得的主要成果如下：

1. 针对空域滤波算法里最常用的共同空间模式，提出一种结合迁移学习与共同空间模式的方法。具体有两种实现方式，分别以原始脑电信号和协方差矩阵的切空间向量为特征，通过对样本特征加权以减少分布差异；再用源用户的加权样本协方差矩阵帮助目标用户建立空域滤波器，加强目标用户标注样本较少时的滤波效果，实验显示该方法可以提升算法 9.34% 的性能，优于其它数种相关迁移学习方法。

2. 针对黎曼空间中的迁移学习方法，提出两种优化方法：方法一是通过选择脑电通道来减少样本协方差矩阵的维度，从而提高黎曼空间中矩阵计算的效率和准确度；方法二是基于样本协方差矩阵之间的黎曼距离进行聚类，挑选出位于目标用户高密度区域的源用户样本。实验表明这两种优化方法都能加强算法的性能，其中方法一能在平均分类正确率方面提升 15.39%，并在计算效率上提升约 30 倍，而在此基础上使用方法二能进一步在分类正确率上提升 2.71% 至 4.54%。

3. 针对脑电多通道时序信号，提出一种无监督的数据对齐方法。以用户的平均协方差矩阵作为参考矩阵，基于仿射不变性对用户进行映射，使得映射后的样本协

方差矩阵以单位矩阵为中心。通过对不同用户分别映射，其样本协方差矩阵的分布得以对齐。实验显示该方法的效果显著优于现有最好的迁移学习方法之一，在运动想象数据集上的平均性能提升约 11%，在 P300 数据集上的平均性能提升约 9%。

4. 针对异构类别空间的迁移学习场景，提出一种对齐异构类别空间的方法。首先为每一个源用户类别设置一个匹配的目标用户类别；然后基于仿射不变性，将源用户每个类别的样本映射到目标用户的匹配类别所在区域，并根据匹配规则更新源用户样本的标签。实验表明，该方法在多种异构类别空间场景下均有效，显著优于未进行对齐的方法，且可以与多种其它迁移学习算法结合使用来进一步提升效果。

综上，本文从不同角度提出了新的迁移学习方法，以实现脑机接口中不同用户、不同设备、不同任务之间的迁移。实验表明，上述方法不仅能够显著地提高算法的分类效果，而且非常高效。这些方法能够减少脑机接口中新用户的校准，提高脑机接口的实用性。

**关键词：**脑机接口    迁移学习    协方差矩阵    数据对齐    异构类别空间

## Abstract

Brain-computer interface (BCI) provides a direct communication pathway between the brain and a computer. In other words, it enables a user to control external devices by using brain signals, which can be applied in biology, medical treatment, and engineering. Electroencephalogram (EEG) may be the most popular BCI input signal, which is collected from the scalp. BCI recognizes the brain activity patterns and the user's intentions by decoding EEG, and translates them to control signals for external devices.

However, due to individual differences, i.e., different subjects have different neural responses to even the same stimulus, it is very difficult to build a generic model whose parameters fit all subjects. So, it usually needs a long calibration process to collect labeled subject-specific data for each individual subject, which is not user-friendly. Transfer learning, which transfers information from related domains to a target domain, is a promising approach for this problem. More specifically, transfer learning reduces the distribution gaps between the auxiliary subjects (source subjects) and the new subject (target subject), and leverages data or knowledge from source subjects to improve the learning performance for a target subject, so that the calibration process can be shorten or even eliminated.

This dissertation focuses on transfer learning in BCIs. We proceed from algorithms to data, and from homogeneous to heterogeneous scenarios, and propose several approaches for transfer learning. In homogeneous scenarios, transfer learning is first incorporated into spatial filtering algorithms. Then improvements over a state-of-the-art transfer learning algorithm are proposed. Afterwards, a data alignment approach is proposed to align EEG trials, which could be used for different algorithms and models. In heterogeneous scenarios, the problem of heterogeneous label spaces is studied and a label space alignment approach is proposed. The major contributions of this dissertation are:

1. For addressing the problem of incorporating transfer learning into common spatial pattern, which is the most widely used spatial filtering algorithm in BCI, an instance-based approach is proposed. More specifically, it reduces the distribution discrepancies between the source and target subjects by re-weighting the raw signals or the tangent space vectors of covariance matrices, and then uses the reweighted sample covariance matrices from source subjects to improve the computation of common spatial pattern filters for the target subject. Experimental results show the performance on two motor imagery datasets was improved by 9.34%, which is greater than that of several existing relevant approaches in literature.
2. Two approaches are proposed to improve the transfer learning algorithms designed

in a Riemannian framework. The first one is channel selection, which reduces the dimension of covariance matrices, and hence makes the Riemannian space computations more accurate and efficient. The second is trial selection, which resamples trials from source subjects using a Riemannian distance-based clustering, such that they become more consistent with those of the target subject. Experimental results show that both approaches enhance the performance of one of the state-of-the-art transfer learning algorithms. More specifically, the first approach improved the average accuracy across all subjects by 15.39%, and made the algorithm 30 times faster. And the average accuracy was further improved by from 2.71% to 4.54% when performing both approaches.

3. An unsupervised data alignment approach is proposed to directly align the EEG trials from different subjects. For each subject, it uses the mean covariance matrix as a reference to transform all trials, such that the mean covariance matrix of the transformed trials becomes an identity matrix. The data of different subjects become more similar as their covariance matrices are aligned. Experimental results show that this data alignment approach significantly outperforms state-of-the-art transfer learning algorithms. More specifically, the performance on motor imagery datasets was improved by around 11%, and that on P300 dataset was improved by around 9%.

4. A label alignment approach is proposed for heterogeneous label spaces transfer learning. More specifically, it first matches each source subject label with a target subject label, then moves the per-class covariance matrices of the source subject to re-center them at the corresponding class means of the target subject, and updates the source labels to the matched target labels. Experimental results show that this label alignment approach is effective in multiple scenarios of heterogeneous label spaces, and can be integrated with other transfer learning approaches to achieve even better performance.

In conclusion, this dissertation proposes several transfer learning approaches to address the problems of cross-subject, cross-headset and cross-task for brain-computer interfaces. The experimental results demonstrated the effectiveness and efficiency of the proposed approaches, which can be used to reduce the calibration effort of brain-computer interfaces.

**Key words:** Brain-computer interface; Transfer learning; Covariance matrix; Data alignment; Heterogeneous label spaces

## 目 录

摘 要 .....	I
Abstract .....	III
<b>1 绪论</b>	
1.1 脑机接口的研究背景及意义 .....	(1)
1.2 脑机接口的组成部分 .....	(7)
1.3 脑机接口中的机器学习 .....	(12)
1.4 迁移学习简介 .....	(20)
1.5 脑机接口的主要挑战 .....	(23)
1.6 本文的研究内容与组织结构 .....	(24)
<b>2 基于迁移学习的共同空间模式</b>	
2.1 引言 .....	(28)
2.2 基于用户的协方差矩阵迁移 .....	(29)
2.3 基于样本的协方差矩阵迁移 .....	(32)
2.4 共同空间模式中的目标函数研究 .....	(33)
2.5 实验比较 .....	(35)
2.6 小结 .....	(45)
<b>3 黎曼空间中迁移学习的优化方法</b>	
3.1 引言 .....	(46)
3.2 黎曼空间数据对齐方法 .....	(46)
3.3 通道选择 .....	(48)
3.4 源用户样本选择 .....	(48)
3.5 实验 .....	(49)
3.6 小结 .....	(56)
<b>4 欧式空间中的数据对齐方法</b>	
4.1 引言 .....	(57)
4.2 黎曼空间数据对齐方法 (RA) 的局限 .....	(57)
4.3 欧式空间中的数据对齐方法 .....	(58)
4.4 实验数据集 .....	(59)
4.5 离线无监督场景实验 .....	(64)
4.6 在线监督场景实验 .....	(71)
4.7 小结 .....	(75)

<b>5 异构类别空间的对齐方法</b>	
5.1 引言 .....	(76)
5.2 异构类别空间的场景 .....	(77)
5.3 类别空间对齐方法 .....	(78)
5.4 LA 与 EA 的比较 .....	(79)
5.5 实验 .....	(81)
5.6 小结 .....	(103)
<b>6 总结与展望</b>	
6.1 本文总结 .....	(104)
6.2 未来工作展望 .....	(105)
致 谢 .....	(107)
参考文献 .....	(108)
附录 1 攻读博士学位期间发表论文目录 .....	(116)
附录 2 攻读博士学位期间的其它研究成果 .....	(118)
附录 3 公开发表的学术论文与博士学位论文的关系 .....	(119)

## 1 绪论

### 1.1 脑机接口的研究背景及意义

人是怎样感知外界信息，以及与外界进行交流的？一般而言，人与外界的信息交互需要经过由肌肉系统、外周神经系统、中枢神经系统组成的复杂神经通路，如图1-1所示：首先由人体各处肌肉感受外界信息，然后由外周神经系统把信息传导到中枢神经系统，由中枢神经系统进行整合、加工、协调，形成思维和决策，再把决策信息通过外周神经系统传回到肌肉，最后由肌肉完成相应的反应和行为。在这个过程中，中枢神经系统负责指挥，外周神经系统负责信息传递，肌肉系统负责感受信息与执行命令。

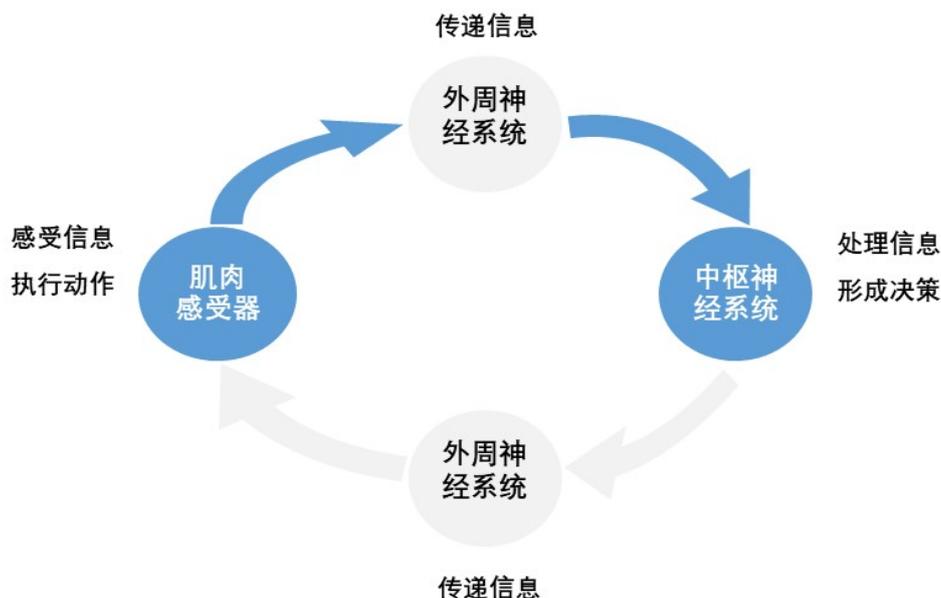


图 1-1 人体神经信息传递示意图。

然而，对于外周神经系统或者肌肉受损的病人，这种信息传递通路可能受阻甚至是中断，以致他们失去了正常的身体功能，比如感知、运动等等。除了修复受损的神经系统或者肌肉以外，科学家们也在寻找一种替代的交流通路来帮助这些病人，特别是当病人受到严重损伤以致无法修复时。脑机接口（**brain-computer interface, BCI**）最初就是为此而设计，它是一种在大脑与外部设备之间建立直接连接通路的系统，使得大脑可以与外部设备进行直接的信息交互和功能整合。这里大脑是中枢

神经系统的高级部分，负责信息的处理和思维的形成；而外部设备作为人体部位的替代与延伸，可以直接向大脑传递信息，或者直接接收大脑指令并完成相应动作。

脑机接口最重要的特点在于使得大脑与外界的信息交互绕过了外周神经系统与肌肉的传导，极大地扩展了人与外界交流的方式，对人类认知大脑、改造身体、乃至探索世界都有重要影响。近年来脑机接口也取得了很多的进展，更在 2019 年被美国商务部列为拟出口管制的新兴技术，下面介绍一些基于脑机接口的应用场景：

### (1) 脑控机器人

脑控机器人是指用户使用“意念”直接控制外部设备的动作，这里的“机器人”泛指机械臂、轮椅、无人机等智能运动设备。该项技术常被用于一些严重运动障碍患者，比如脊髓损伤、脑干中风、重症肌无力、甚至高位截瘫等，这些疾病会使大脑与身体失去联系，从而丧失执行意志运动的能力。而脑机接口可以帮助他们通过大脑信号直接控制辅助运动器械，比如机械外骨骼、轮椅等，从而重新获得一部分运动能力。例如，Hochberg 等人于 2012 年在《Nature》发表研究成果<sup>[1]</sup>，通过脑机接口技术使得长期瘫痪的病人可以用“意念”控制机械臂，并完成抓取水杯、喝水、放回水杯等动作，如图1-2 所示；在 2014 年巴西世界杯上，一位脊髓损伤后截瘫 9 年的患者应用了脑机接口系统，经过虚拟行走、踢球的训练后，学会了通过大脑控制外骨骼运动，成功开出了第一脚球，如图1-3 所示。

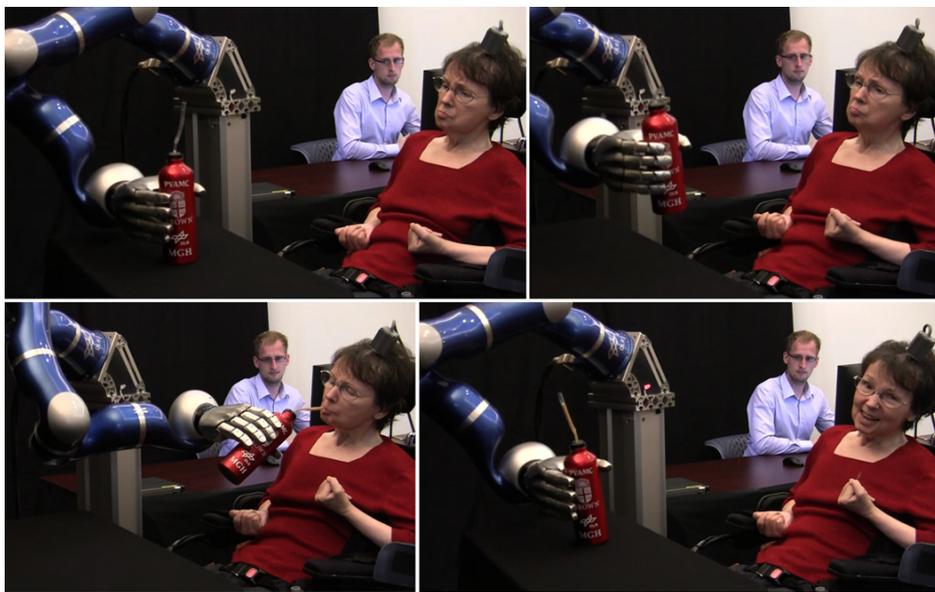


图 1-2 瘫痪病人使用“意念”控制机械臂，完成抓取水杯、喝水、放回水杯等动作<sup>[1]</sup>。

此外，随着无创式脑机接口技术的发展，脑控机器人技术也被推广到无身体障碍的普通消费者人群，使得人们可以用“意念”控制机器人、无人机等设备。尽管当前人们基于此项技术还只能完成一些简单的指令，但相信随着技术的发展，脑控



图 1-3 截瘫患者用“意念”控制外骨骼行走，并在2014巴西世界杯上开出第一球。

机器人的效果会越来越接近“如臂指使”。卡内基梅隆大学的贺斌教授<sup>[2]</sup>于 2019 年在《SCIENCE ROBOTICS》上发表了其团队的一项最新研究成果，基于非侵入式脑机接口，研发出了一种能让人用意念控制机器臂连续、快速运动的系统，其特点在于受控设备可以“连续”运动，更接近现实使用的情境。

## (2) 意念交流

2014 年的“冰桶挑战”使得很多人认识到并开始关心渐冻症病人这一群体，由于神经损伤，他们不仅四肢和躯干会逐渐萎缩，甚至会失去说话的能力。脑机接口系统可以帮助他们表达出自己的想法，目前最常用的就是“意念打字”系统，即让受试者看着电脑屏幕上闪烁的字符，然后通过脑电波判断受试者在看哪个字符来进行文字输入。2019 年中央电视台在《挑战不可能》节目中展示了由清华大学开发的一个文字输入系统，一名渐冻症患者通过脑机接口在电脑屏幕上打出了著名作家舒婷的诗句：“希望，而且为之奋斗，把这一切放在你的肩上”，如图1-4 所示。在 2019 年的世界机器人大会上，天津大学的选手以每分钟 691.55 比特的信息传输率获得冠军，创造了脑电打字的新纪录，该比特率等同于每 0.413 秒可以录入一个字母，接近普通人用手机打字的速度。

上述系统都是基于字符闪烁的视觉刺激引起的脑电反应进行文字输入，此外也有研究开始探索一些更加直接的意念交流方式。美国加州大学华裔科学家 Edward Chang<sup>[3]</sup>于 2019 年在《Nature》上发表了其团队基于脑机接口解码意念的研究成果，通过实时解码大脑信号，将受试者默念的语句转换成语音，但该研究仍然不是直接理解大脑的思维，而是先利用神经信号与人声道咬合部位运动的映射关系进行解码，再根据发声器官的活动和说话声的对应关系，最终判断人们想说的是什么。2017 年

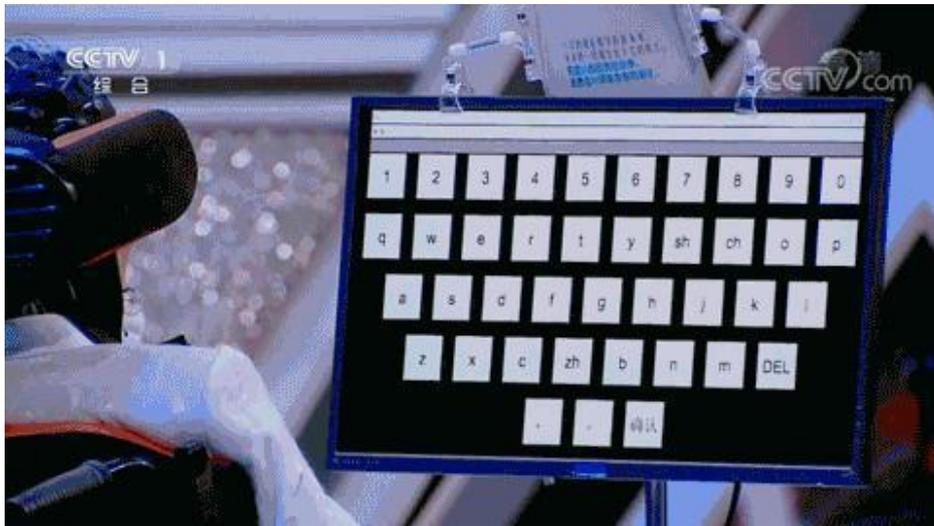


图 1-4 渐冻症患者通过脑机接口“意念打字”。

4 月，Facebook 在 F8 大会上称正在开发“脑机语音文本界面”的项目，目标是未来能通过大脑信号实现 100 字/分钟的输出速度。相比较现有的“意念打字”技术，Facebook 的目标打字速度有了极大的飞跃，因此也计划采用完全不同的技术，据报道，Facebook 不是基于视觉诱发的脑电反应，而是以极快的速度扫描大脑海马体里的语言信息。

除了大脑与计算机之间的通讯，科学家们也在研究大脑与大脑之间的通讯，图1-5所示是美国华盛顿大学的 Rajesh Rao 教授开发的两个人之间的通讯，其中脑信号输出者想象自己右手的运动，然后使用脑机接口读取此种“意图”，并通过颅内刺激把信号直接传递给另一方的大脑，脑信号接收者在信号抵达的瞬间，几乎是立刻地、无意识地运动了右手。

### (3) 神经通路修复

神经通路修复是脑机接口在康复领域的一个重要应用，可以帮助一些病人主动去修复一些受损的神经。例如中风引起的偏瘫病人，由于大脑的运动皮层受到了损伤，导致部分身体无法运动。现在医院普遍的做法是用康复训练机器人帮助活动瘫痪的肢体，但是这样做的效果是很差的，因为受损的部位是在大脑而不是四肢。所以科学家们设计了一套新的康复训练方法，通过脑机接口让大脑和肢体协同训练，简单来说，就是让患者先想像瘫痪肢体的运动，然后通过脑机接口监测病人的脑电反应，当发现病人正在想象运动时再去活动瘫痪的肢体。研究表明，这种需要患者主动参与的康复方法可以在患者受损的中枢神经中形成反馈，刺激脑神经的重塑，有助于断裂神经的再生从而提高康复疗效<sup>[4-6]</sup>。

除了刺激断裂神经的再生，脑机接口甚至可以在体内直接替代断裂的部位。2016 年，瑞士联邦理工学院的 Grégoire 教授等<sup>[7]</sup>在《Nature》发表了相关成果，研



图 1-5 通过脑机接口实现大脑与大脑之间的通讯。

究人员给一只脊髓半切断的猴子植入了“脑-脊髓接口”的神经假体，绕过猴子脊髓受损部分，重新在大脑与受损脊髓远端搭建起连接，让猴子重新获得了对瘫痪后肢的控制。

#### (4) 大脑状态解读与监测

脑机接口还可以用于解读用户的专注度、疲劳度、睡眠程度以及情绪等基础认知功能的状态。相对于传统的主观评测手段，脑机接口可以进行实时监测，且评测依据更为客观，在管理、安全、市场、人因工程等领域都有着很好的应用前景。例如，通过脑机接口技术对学生听课的注意水平进行实时评估，老师们就可以根据学生的状态调整教学内容与课程节奏<sup>[8]</sup>；评测驾驶员、化工厂操作员等一些较危险岗位人员的精神状态，及时对疲劳操作和开小差等情况进行干预，可以帮助减少意外事故的发生<sup>[9]</sup>；在产品设计和市场营销方面，脑机接口作为一种新型的人机交互技术，可以用来记录用户在体验产品过程中的感受或者情绪，从而帮助改进产品或者营销手段。

#### (5) 脑疾病辅助诊断

脑机接口技术还可以用于与大脑相关疾病的研究、诊断与治疗，比如癫痫、自闭症、帕金森症等等。癫痫是最常见的神经系统疾病之一，以脑神经元异常放电引起反复痫性发作为特征。约有三分之一的患者的病情无法用现有的抗癫痫药物控制，但其中一半可能受益于癫痫手术<sup>[10]</sup>。而脑机接口可以通过检测异常放电来定位癫痫发作区域，为手术提供帮助。除了定位病灶，脑机接口还可以通过对病灶施加电刺激来调节该处神经元兴奋性，从而缓解病情<sup>[11-13]</sup>。目前，美国 NeuroPace 公司研制的用于癫痫治疗的脑部植入装置 RNS Stimulator 已于 2013 年 11 月被 FDA 批准上市。基于类似原理开发的深部脑刺激也已经成为治疗帕金森症的一种有效手段，进

入了临床应用<sup>[14,15]</sup>阶段。

### (6) 感觉功能修复

基于脑机接口技术研发出视觉、听觉神经假体可以帮助恢复病人的视听觉神经功能。人工耳蜗是迄今为止最成功、临床应用最普及的脑机接口设备，与助听器将声音放大来辅助听力的原理不同，人工耳蜗可以把外界的声音变成听觉神经能理解的电流指令，然后用来刺激听觉神经的功能正常部分，使听力严重受损的患者恢复一部分感受声音的能力。

人工视网膜的工作原理与人工耳蜗相似，但又更加复杂。人工视网膜首先把接收到的光信号转化为电信号，然后通过脑机接口（微电极阵列）对视觉神经施加电刺激，使神经细胞产生视觉。William Doherty 是最早进行相关研究的科学家之一，他早在 1978 年就将视觉神经假体进行了临床应用，通过脑机接口技术在一位盲人的视觉皮层植入微电极阵列，使病人恢复了有限的视觉感受能力。总部位于美国的医疗设备公司 Second Sight Medical Products 研究的视觉假体系统也已经获得 FDA 批准进入美国市场，尽管目前还不能做到让盲人彻底恢复，该技术对患者来说仍然是极大的帮助。在未来，增加微电极阵列的电极数量可以提高人工视网膜的效果，必将为众多失去视力的患者带来光明与希望。

综上所述，脑机接口目前最主要的应用方向是医疗和康复，对个人和社会都能带来直接影响。随着社会的发展进步，越来越多的国家和地区开始进入老龄化社会，而我国已经事实上迈入老龄化并正在加深，根据中国老龄科学研究中心的报告，截至 2017 年底，中国 60 岁及以上老年人口已达 2.41 亿。不可避免地，中风、老年痴呆以及帕金森症等疾病的患者人数也会越来越多，这些疾病通常会使病人丧失正常生活的能力，不仅会给患者及其家庭带来痛苦，也会引发很多的社会问题。脑机接口在治疗这些疾病，以及提高患者的自主生活能力上都能发挥重要作用。

此外，脑机接口也被用到了军事、教育、交通、娱乐等诸多领域。例如，美国已实现了利用飞行员的意念对 F-35 战机飞行模拟器进行操控驾驶，美国国防部高级研究计划局（DARPA）也是较早开展脑机接口相关研究的机构之一。此外，现在市场上也已经发售了多款基于脑机接口训练人的专注度、帮助冥想、监测睡眠、游戏等产品。

鉴于脑科学研究的重要意义，近年来世界各国纷纷启动脑计划。美国在 2013 年率先宣布启动“创新性神经技术大脑研究”计划（Brain Research through Advancing Innovative Neurotechnologies（BRAIN）Initiative）；同年，欧盟推出了多国联合参与的“人类脑计划”（Human Brain Project, HBP）；2014 年，日本也宣布神经科学研究计划（Brain Mapping by Integrated Neurotechnologies for Disease Studies, Brain/MINDS）。在中国，“脑科学与类脑研究”已经被列入“十三五”规划纲要中的国家重大科技创新和工程项目，北京和上海也相继成立脑科学与类脑研究中心。

企业界也很重视脑机接口的发展，特斯拉创始人 Elon Musk 于 2016 年 7 月投资

创立了脑机接口公司 Neuralink，其短期目标是治愈严重的脑部疾病，如老年痴呆症和帕金森症，并且最终通过与人工智能的融合来增强大脑。Elon Musk 于 2019 年发布了第一款产品：柔性电线和缝纫机器人，该产品能够以很小的创伤在大脑内部植入电极。Facebook 也于 2017 年宣布展开一项使用非侵入式的脑机接口进行意念打字的项目，但 Facebook 打算采用与现有技术不同的方案，使得打字速度超过每分钟 100 字。此外还有 Emotiv 和 NeuroSky 等初创公司开展脑机接口研究并获得了资本市场的支持。

总体而言，如今脑机接口的研究还处于初期阶段。在未来，不仅前文所述应用会得到进一步发展，变得更加安全、可靠、方便，还会发展出许多新的应用，比如人脑增强，Elon Musk 设想基于脑机接口技术将微电极植入大脑，从而提高人类的思维能力、记忆力、决策能力，最终帮助人类保持对人工智能的竞争力。甚至像《阿凡达》、《阿丽塔》等影视作品里所描绘那样的意念控制、半机械人等也都有实现的可能，正如《科学美国人》当年所说的：我们所想象的一切，都将变为现实。

## 1.2 脑机接口的组成部分

脑机接口系统通常包含以下几个部分：信号采集模块、信号分析模块、控制模块，如图1-6所示。其中信号采集模块是把大脑活动信息通过某种信号介质采集出来，以便于后续的储存、分析、应用等等。信号分析模块是要解码大脑信号，从中获取大脑意图。控制模块是把大脑意图转化为外部设备的控制命令，最终实现大脑意图。

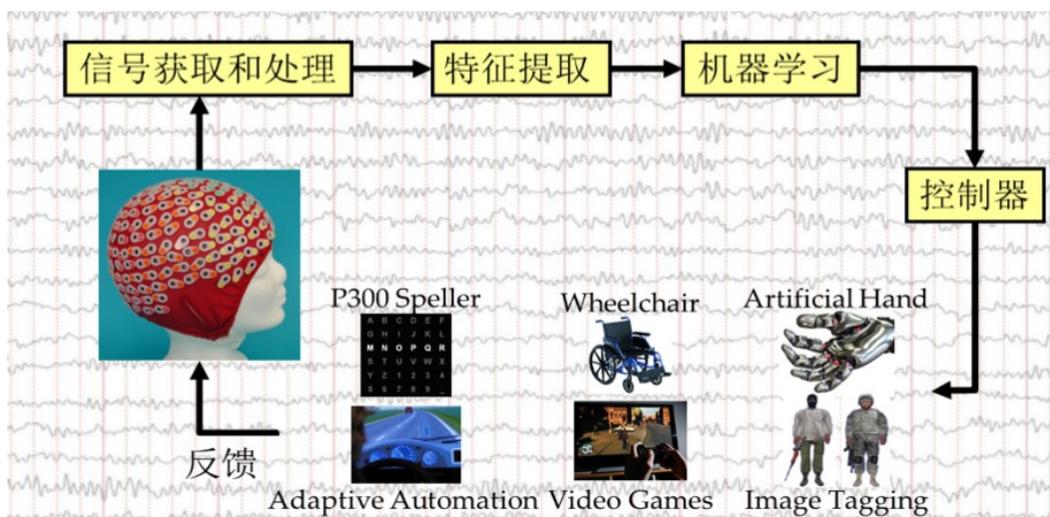


图 1-6 脑机接口系统示意图。

### 1.2.1 信号采集模块

脑机接口的核心是使得机器能够“读懂”大脑意图，所以首先需要用一种方式来记录、测量大脑的活动情况。目前用来测量脑活动的脑成像方式有很多种，本节作简要介绍。

从测量原理看，信号采集方式大致可以分为两大类<sup>[16]</sup>：(i) 基于电生理的；(ii) 基于血液动力学的。其中电生理活动是由神经元的电化学物质交换信息时产生的，人的大脑内部有百亿级的神经元，神经元之间又有着错综复杂的连接关系。研究发现，大脑活动会大量激活神经元，激活神经元会产生离子电流，离子电流在神经元内和神经元间流动。所以可以通过神经元活动产生的电信号来收集大脑活动的信息。皮层脑电、头皮脑电、脑磁图等都是基于电生理活动来监测大脑活动的信号。血液动力学反应是一个过程，在这个过程中，大脑通过血液的新陈代谢为神经元活动提供所需的氧，而血红蛋白是血液运送氧的载体，因此，在大脑神经元集体活动区域会出现血液中氧合血红蛋白浓度上升，脱氧血红蛋白浓度下降的现象。功能性近红外光谱技术和功能性磁共振成像都是基于血液动力学原理，用一种间接的方式来监测神经元的活动。

从信号采集方式看，脑机接口可分为侵入式、非侵入式和半侵入式。侵入式需要将信号采集设备（电极或者电极阵列）植入到大脑皮层，优点是不仅能够采集到质量更优、位置更精确的神经反应，还能对神经实施精确的电刺激；缺点是需要进行手术，存在着较高的风险。非侵入式无需在大脑内部植入设备，只要在头皮表面放置电极，甚至不用接触头部就可以采集到神经信号。其特点与侵入式相反，优点是避免了手术风险，安全且方便；缺点是信号质量不高，因为神经信号在传导过程中会衰减，并且伴随着大量的噪声。半侵入式则是介于侵入式和非侵入式之间，它需要通过手术将信号采集设备植入到颅腔内，但又不像侵入式脑机接口那样植入到大脑皮层，而是在大脑皮层之外。半侵入式的优缺点也是介于侵入式和非侵入式之间，一方面能获得中等质量的信号，另一方面也会带来中等的手术风险。不同的采集方式有着不同的应用场景，目前侵入式脑机接口一般是用于医疗方向，而非侵入式脑机接口则可以用于普通消费市场。

下面简要总结各种脑成像方式的特点：

#### (1) 头皮脑电

头皮脑电是脑机接口中最常用的信号之一，通常由头戴式的脑电帽通过电极从头皮表面采集，可以在头皮上监测到群体神经元的放电活动。它的优势是采集方便、无手术风险、响应快；缺点是空间分辨率低、信噪比低、易受外界噪声干扰。

#### (2) 皮层脑电

皮层脑电是一种半侵入性记录方式，需要开颅手术将电极网格放置在大脑表面来测量大脑皮层电活动。与脑电图相比，皮层脑电具有更高的时间和空间分辨率，

以及更高的振幅和更低的易损性。

### (3) 脑磁图

脑磁图检测的是脑沟内锥体细胞的细胞内电流产生的磁场，检测过程安全、简便，对人体无任何副作用。与脑电信号相比，脑磁图优势在于传导过程中介质的影响小，信号没有扭曲，所以空间分辨率高；劣势在于设备昂贵，对环境要求苛刻。

### (4) 功能性近红外光谱技术

功能性近红外光谱技术是近年来新兴的一种非侵入式、基于血液动力学的脑功能成像技术。当大脑区域激活时，脑供血系统会向大脑活动区域输送含氧合血红蛋白的血液，而氧合血红蛋白和脱氧血红蛋白对不同波长的近红外光吸收率有差异。功能性近红外光谱技术即是基于此原理，从血液动力学变化中推论大脑皮层的活动区域与情况。

### (5) 功能性磁共振成像

功能性磁共振成像也是基于血液动力学的原理，不过它是检测的是血液动力过程带来的磁场变化。前文提到，脑供血系统会向大脑活动区域输送含氧合血红蛋白的血液，而带氧血红素与去氧血红素之间磁导率有差异。功能性磁共振成像就是从磁场变化中推论大脑皮层的活动情况，其优点是空间分辨率高，缺点是时间分辨率相对低。

在上述脑成像方式中，头皮脑电（electroencephalographic, EEG）没有任何手术风险、装备最容易使用、成本低、相应快的特点使得它是当前最常被研究，也是未来最可能被广泛应用的。因此，本文的研究对象是基于 EEG 的脑机接口，并专注于对 EEG 信号的处理与分析算法研究。

## 1.2.2 信号分析模块

信号分析模块负责从采集的原始信号中读懂大脑的意图，是脑机接口系统中最为关键的环节，一般包括信号预处理、特征提取、分类（回归）几个步骤。

### (1) 预处理

预处理是信号分析模块的第一个步骤，其目的在于提高信号质量。EEG 是大脑神经元集体活动产生的电信号，再由放置在头皮上的电极采集而来。在这个过程中，脑电信号会受到颅骨等组织的阻隔，以及各种噪声干扰，所以通常信号微弱、非平稳、信噪比低。具体表现来看，未经处理的 EEG 信号在波形幅值上波动大且无规律，在空间上各通道之间相互影响，从中基本不能得到有意义的结果。

因此，首先需要对信号进行预处理，以减少各种噪声干扰，提高信噪比。人工伪迹是 EEG 里最常见的噪声，可分为生理伪迹和非生理伪迹。非生理伪迹主要是由环境与设备造成的干扰，比如工频干扰或电极阻抗的变化。中国的工频干扰为 50Hz，一般可采用 50Hz 陷波滤波器来抑制噪声；电极阻抗的变化是指电极与头皮

之间的接触阻抗可能会随着实验进行（出汗、导电膏变干等等）而发生变化，头部的晃动也可能导致电极的位置移动而引入伪迹。大部分的非生理伪迹可通过改善实验方法，以及滤波来加以控制。

生理伪迹是指用户的生理活动产生的伪迹，通常由肌肉、眼部和心脏活动引起，分别称为肌电、眼电和心电伪迹<sup>[17]</sup>。其中肌电伪迹的频率成分主要集中在较高频段，可用低通滤波器去除；眼电伪迹的特点是在空间分布上，主要影响的是靠近前额区域的通道，通道位置离眼部越远，受影响则越小；心电的频率成分则主要集中在较低频段。大部分的生理伪迹也可以通过滤波来加以控制，但仍有部分伪迹难以去除。比如眼电，它的幅值较大，在频率成分上也与 EEG 的低频部分重叠。为减少眼电的干扰，一方面应尽量减少眼电的引入，比如实验过程中引入间歇期；另一方面眼电也难以完全避免，所以需要一定的方法来将其降至最低，经典的方法是把得到的 EEG 信号减去一定比例的眼电参考信号<sup>[18]</sup>。

除了实验方法改进和时域滤波，空间滤波也是常用的提高信噪比的方法。比如共同平均参考，将每一个通道的信号减去所有通道的均值，可以降低远距离信号源的影响；Laplace 滤波则可以看作局部的共同平均参考，将每一个通道的信号减去其周围若干通道的均值；独立成分分析<sup>[19]</sup> 则试图从混合的 EEG 中分离出各个信号源，也称为盲源分析。

此外，伪迹去除也与 EEG 特征的频段有关，比如运动想象中的 EEG 特征频段主要是 8-13Hz 以及 14-30Hz 两个频段，所以对于低频伪迹可以用高通滤波器来去除。因此，不同的实验模式需要不同的预处理方法，本文采用的预处理方法则会在后续章节中的数据集介绍时作对应说明。

## (2) 特征提取

EEG 是多通道时域信号，数据维度很大，所以在预处理之后，通常还需要针对具体任务提取相应的脑电特征。常用的特征提取方法可分为以下几大类：

1. 时域分析：时域分析主要研究信号在时域上的特征，直观地理解就是观察波形幅值的规律。此类方法通常比较直观，物理意义也比较明确，比如峰值检测、平均值检测、方差分析等等。
2. 频域分析：脑电是有节律的神经电活动，所以其很多特征反映在频域上，比如  $\delta$  波（1-3Hz）可以在成年人极度疲劳和昏睡时检测到； $\alpha$  波则在人清醒、安静并闭眼时表现明显等等。频域分析主要通过分析功率谱得到特征，比如功率谱估计和傅里叶变换等等。
3. 时频分析：时频分析是指同时在时间和频率域中对信号进行分析。此类方法有利于表现非平稳信号和时变信号的特征，常用的方法有小波变换和短时傅里叶变换等等。
4. 空间分析：大脑的各个区域控制着不同的思维活动、身体部位和功能，所以区别 EEG 的信号源是进行信号分析的一个重要任务。空间分析即是研究脑

电的空间信息，对神经活动进行定位，是近年越来越受到重视的研究方向，常用的方法有独立成分分析和共同空间模式等等。

### (3) 分类方法

在提取特征之后，还需要使用机器学习模型对大脑活动的模式进行识别。线性判别分析、支持向量机、逻辑回归等算法也常常被用于 EEG 信号的分类，这部分是本文的主要研究内容，因此本文将在1.3节作进一步阐述。

## 1.2.3 控制模块

由于人类对大脑的研究与理解还比较粗浅，目前的技术尚不能使计算机理解人的高级思维，而是仅仅能够识别一部分简单的大脑活动模式。因此，研究者通过设计特定的脑机接口范式来控制外部设备：将当前技术可理解的简单大脑活动模式与某些特定的高级思维联系起来，然后用户只需执行计算机可识别的简单大脑活动，就可以表达对应的高级思维。这种间接“读取”意图的模式是现阶段脑机接口控制外部设备主要方法，目前脑机接口系统最常用的控制信号有视觉诱发电位、P300 诱发电位和感觉运动节律。

### (1) 视觉诱发电位

视觉诱发电位是受试者的视觉皮层在接受视觉刺激后的大脑活动调制<sup>[20]</sup>。当刺激靠近中央视野时，视觉诱发电位的振幅会大幅增加<sup>[21]</sup>，所以这些调制相对容易检测。视觉刺激的形态可以是闪光刺激或者图形模式。根据刺激频率不同，视觉诱发电位也可分为瞬态视觉诱发电位和稳态视觉诱发电位，其中瞬态视觉诱发电位发生在视觉刺激频率低于 6hz 时，而稳态视觉诱发电位发生在较高频率的刺激下<sup>[22]</sup>。

常用于脑机接口的是稳态视觉诱发电位，它是由相同的视觉刺激以高于 6hz 的频率变化所引起的。稳态视觉诱发电位根据使用中的具体刺激序列调制可分为三类<sup>[23]</sup>：时间调制、频率调制和伪随机码调制。在时间调制脑机接口中，不同目标的闪烁序列在时间上是正交的。在频率调制脑机接口中，每个目标都以一个独特的频率闪烁，产生一个与它的谐波相同基频的诱发响应周期序列。在伪随机码调制脑机接口中，每个目标的闪烁持续时间由伪随机序列决定。典型的基于稳态视觉诱发电位的脑机接口是在屏幕上显示闪烁的刺激，如数字或字母。用户可以将视线移动到闪烁的数字或字母上，然后通过脑电响应的频率来判断用户所看的刺激字符<sup>[24]</sup>。这种控制范式的优点是几乎不需要训练，但缺点是用户必须盯着屏幕且容易产生视觉疲劳。

### (2) P300诱发电位

P300 是一种事件相关电位，是在相关事件发生后约 300 毫秒左右出现的一个正的峰值<sup>[25]</sup>，相关事件的刺激来源可以是视觉、听觉或者是其它感觉。一些研究已经证明，刺激的可能性越小，反应峰值的振幅越大<sup>[26]</sup>。使用基于 P300 的脑机接口不

需要培训，但是用户习惯了不频繁的刺激可能会导致 P300 振幅减小，从而降低脑机接口的使用效果<sup>[27]</sup>。基于 P300 的脑机接口有两种常见的应用：文字输入和异常图片检测。

基于 P300 的文字输入是在屏幕上显示由字母、数字或其他符号或命令的矩阵<sup>[28,29]</sup>，然后使这个矩阵的行或列随机闪烁。用户注视所需的符号，只有当所需的行或列闪烁时才会触发 P300。由于脑电信号的信噪比较低，单次试验的目标符号检测非常困难。所以对于每个字符选择，行或列必须闪烁几次然后求平均值，以提高实验的准确性。

异常图片检测是把少数异常图片夹杂在多数普通图片中间，然后以较快的速度在受试者眼前闪过，当异常图片出现时，一般受试者的脑电响应中会出现 P300 信号。

### (3) 运动想象

人体不同部位的运动与感知，受到大脑皮层不同区域的控制，而当人想象自己身体部位的运动时，会在对应的大脑皮层区域引起感觉运动节律。基于运动想象的脑机接口就是在用户想象运动的部位与外部设备的控制命令之间建立映射关系，比如当检测到用户正在想象左手运动时，命令外部设备向左移动。

人在进行肢体运动或者想象运动时，会引起大脑皮层对应区域的感觉运动节律的调制，包括 mu 节律（8-12Hz）和 beta 节律（14-30Hz）<sup>[30]</sup>。当大脑执行运动相关的任务时，这些节律活动会因为运动、运动想象或者运动准备被抑制而下降，在运动结束或者进入休息状态之后上升<sup>[31,32]</sup>。因此，检测大脑皮层特定区域的感觉运动节律可用于“解读”受试者正在进行哪部分身体的运动想象，然后通过映射关系控制外部设备。

感觉运动节律在脑机接口中得到了广泛的研究，比如 Wadsworth<sup>[33]</sup>、Berlin<sup>[34]</sup>、以及 Graz<sup>[35]</sup> 脑机接口系统都使用感觉运动节律作为控制信号。基于感觉运动节律的调节，可以在人类自主运动发生之前预测其运动<sup>[36]</sup>。

### (4) 其它控制信号

除了上述常用的 3 种控制信号之外，研究者们也基于 EEG 开发了一些其它的实验范式，一个主要类别是识别大脑的基础认知或者情感状态，比如驾驶疲劳监测<sup>[37,38]</sup>、情感计算<sup>[39-41]</sup>、睡眠监测<sup>[42-44]</sup>等等。

## 1.3 脑机接口中的机器学习

本文在上一节中提到，信号分析模块是决定计算机能否理解大脑意图的关键阶段。在对数据进行降噪等预处理步骤后，信号分析模块还需要从原始脑电数据中提取与大脑活动模式相关的特征，然后基于历史数据，使计算机能够学习到特征与大

脑活动模式之间的内在规律，并能够自动识别未来的脑电数据的模式，这个过程也称为机器学习<sup>[45]</sup>。

我们首先介绍 EEG 信号的采集过程及数据特点，如图1-7所示：左图是 EEG 信号采集的示意图，即以脑电帽的形式在用户头皮表面不同位置放置电极，以从各个方位采集大脑的脑电信号；右图是一段 EEG 信号的示例，其每一条曲线代表一个电极（也称为“通道”）采集到的数据，也就是脑电幅值随时间变化的波形。

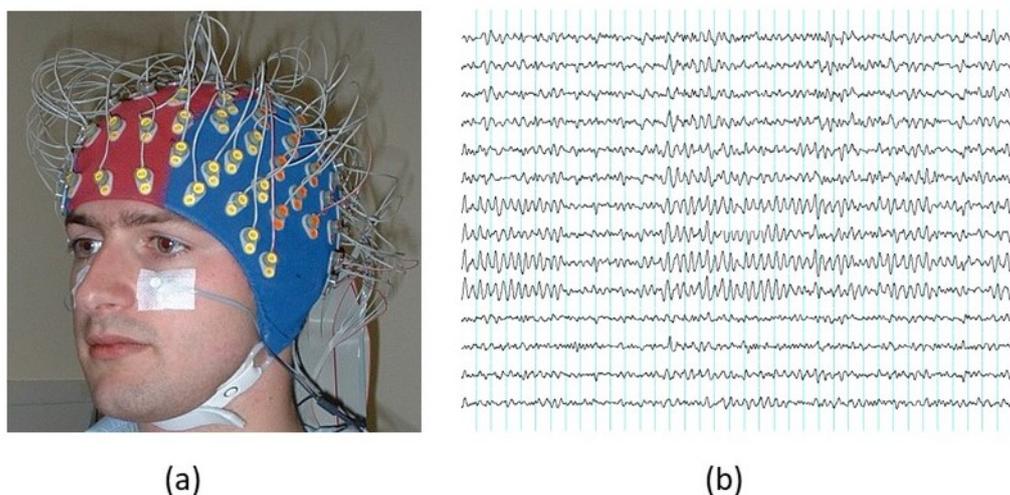


图 1-7 EEG信号采集与示例：(a) EEG信号采集示意图；(b) 一段EEG信号示例。

在实际应用中，一次脑机接口实验或者使用通常包含多个“事件”，我们需要把一段长时间的 EEG 信号按照事件发生的时间分割成多个小段，每一小段对应的是机器学习概念中的“样本”。以运动想象实验为例，用户在实验全程佩戴脑电帽，然后实验过程中会设置很多的事件来激发用户的神经反应，这些事件的类型对应着不同的控制命令。比如“想象左手运动”和“想象右手运动”即为两个类型的事件，分别对应着控制器“向左运动”和“向右运动”的命令，实验过程中会以一定的方式多次提醒用户进行“想象左手运动”和“想象右手运动”，并记录事件发生的时间。然后，我们根据每一个事件发生的时间将整个实验期间的 EEG 信号分割为多个片段，每个片段对应一个事件。分割完成后，EEG 片段就是机器学习概念中的样本，片段对应事件的类型即为该样本的“标签”。

令  $X_i \in \mathbb{R}^{n_c \times T}$  表示一段被截取的 EEG 信号（一个样本），这里  $n_c$  代表通道数目， $T$  代表采样点数目，即截取的 EEG 时间长度乘以采样频率；令  $y_i$  表示该样本的标签，即事件的类型。

脑机接口中的机器学习是指使用已知事件类型的 EEG 信号（已标注样本）训练模型，然后用于未知事件类型的 EEG 信号的识别。识别结果关乎到控制器是否接收到正确的指令，是脑机接口系统的核心。目前，脑机接口中常用的机器学习模型可

以分为两大框架:

1. 基于欧式空间的传统分类框架;
2. 基于黎曼几何的分类框架。

另一方面, 尽管深度学习在计算机视觉等领域表现出了卓越的性能, 也渐渐地被应用于脑机接口<sup>[46,47]</sup>。但在本文针对的运动想象和P300两种应用场景中, 尚未显示出令人信服的改进<sup>[48]</sup>。所以本文专注于非深度学习方法的研究, 本节将对上述两种分类框架分别作简要介绍。

### 1.3.1 基于欧式空间的传统分类框架

基于欧式空间的传统分类框架一般包括空域滤波、特征提取、以及分类几个步骤。

#### (1) 空域滤波

多通道 EEG 信号的生成与检测过程与图1-8所示: 大脑皮层的局部活动激发生物电信号 (源信号  $s$ ), 经脑容积传导到头皮, 由于脑容积的导电效应, 头皮电极检测到的其实是由多个源信号混合而成, 即观测信号  $x = As$ , 其中  $A$  表示脑容积传导模型。这使得从观测到的原始 EEG 信号不能直接回溯到源信号, 而源信号才是分析大脑活动的基础。Nunez 等研究者<sup>[49]</sup>通过模拟实验发现, 从头皮电极采集的 EEG 信号中, 只有不到一半来自信号源半径  $3cm$  的区域。

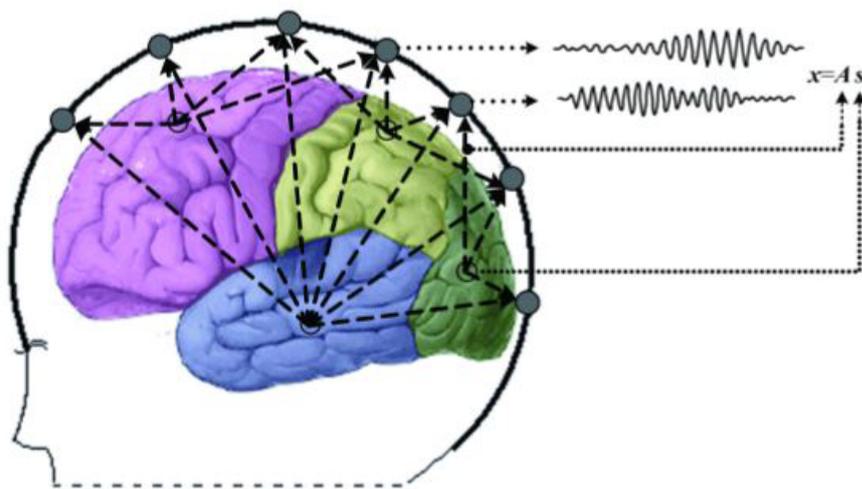


图 1-8 EEG的生成与检测示意图<sup>[50]</sup>。

同时, 一些干扰信号比如眼电、肌电等也会被头皮电极检测到。因此, 尽管 EEG 有安全、方便、响应快等优点, 但也存在空间分辨率和信噪比低的缺点。空域滤波器常常被用于提高 EEG 信号的信噪比, 比如共同平均参考、独立成分分析、共同空间模式等等。

共同空间模式 (common spatial pattern, CSP) [51-54] 是被最广泛应用的一种空域滤波器。它最开始是应用于二分类的运动想象任务, 基本思想是把一段多通道的信号映射到一个新的空间, 使得两个类别的方差尽量差别大。

CSP 是一种有监督滤波方法, 建立在已标注数据的基础之上。假设一个用户有  $N$  个已标注样本  $\{X_i, y_i\}_{i=1}^N$ , 其中样本  $X_i \in \mathbb{R}^{n_c \times T}$ , 标签  $y_i \in \{0, 1\}$ , 则首先计算每个样本的协方差矩阵:

$$C_i = X_i X_i^T \quad (1.1)$$

其中  $C_i \in \mathbb{R}^{n_c \times n_c}$  代表样本  $X_i$  的协方差矩阵。

然后分别计算每一个类别的平均协方差矩阵:

$$\bar{C}_m = \frac{1}{|\mathcal{I}_m|} \sum_{j \in \mathcal{I}_m} C_j, \quad m = 0, 1 \quad (1.2)$$

其中  $\mathcal{I}_m$  代表标签为  $m$  的样本集合,  $|\mathcal{I}_m|$  代表集合内样本个数。

CSP 基于两个类别的平均协方差矩阵  $\bar{C}_0$  和  $\bar{C}_1$  寻找一个映射矩阵  $W \in \mathbb{R}^{n_c \times f}$ , 将所有样本映射到一个新的低维空间:

$$\tilde{X}_i = W^T X_i \quad (1.3)$$

其中  $\tilde{X}_i \in \mathbb{R}^{f \times T}$  为映射后的新样本, 映射矩阵  $W$  的每一列代表一个滤波器。

前文提到过, CSP 的思路是使得映射后的 EEG 样本在两个类别的方差上表现出尽量大的差别, 所以映射矩阵的目标函数可以写为:

$$W_0 = \arg \max_W \frac{\text{tr}(W^T \bar{C}_0 W)}{\text{tr}(W^T \bar{C}_1 W)} \quad (1.4)$$

其中  $\text{tr}(\cdot)$  代表矩阵的迹。

$W_0$  的解是矩阵  $\bar{C}_1^{-1} \bar{C}_0$  最大的  $f$  个特征值对应的特征向量, 它使得映射后的 0 类样本方差最大, 同时 1 类样本的方差最小。

在实际应用中, 我们通常会寻找一个映射矩阵  $W = [W_0, W_1] \in \mathbb{R}^{n_c \times 2f}$ , 使得

$$W_1 = \arg \max_W \frac{\text{tr}(W^T \bar{C}_1 W)}{\text{tr}(W^T \bar{C}_0 W)} \quad (1.5)$$

即  $W_1$  最大化类别 1 的方差, 同时最小化类别 0 的方差。与  $W_0$  的解相似,  $W_1$  是矩阵  $\bar{C}_0^{-1} \bar{C}_1$  最大的  $f$  个特征值对应的特征向量。

由于  $\bar{C}_1^{-1} \bar{C}_0$  与  $\bar{C}_0^{-1} \bar{C}_1$  有相同的特征向量且对应特征值互为倒数,  $W_1$  也是矩阵  $\bar{C}_1^{-1} \bar{C}_0$  最小的  $f$  个特征值对应的特征向量。所以我们只需要对矩阵  $\bar{C}_1^{-1} \bar{C}_0$  (或者  $\bar{C}_0^{-1} \bar{C}_1$ ) 进行特征分解, 选取其最大以及最小的  $f$  个特征值对应的特征向量作为滤波器即可。

CSP 优秀的性能使得它得到了广泛的应用和研究，同时研究者们也在不断地寻找完善和优化 CSP 的方法。比如，Dornhege 等采用“一对多”的方式将 CSP 滤波器从二分类扩展到多分类<sup>[55]</sup>；Ang 等<sup>[56]</sup>提出滤波器组 CSP (filter bank CSP)，将 EEG 通过多个频带的带通滤波，再从每个频带分别提取 CSP 特征，然后使用特征选择算法挑选出最有益于分类的特征；Lotte 等<sup>[57]</sup>则通过正则化项来提升 CSP 的性能。

## (2) 特征提取与分类

在 CSP 进行空域滤波之后，一般的机器学习算法还要求提取 1-D 特征（向量），对于运动想象数据，通常使用方差的对数作为特征<sup>[58]</sup>：

$$\mathbf{x} = \log \left( \frac{\text{diag}(\tilde{X}\tilde{X}^T)}{\text{tr}(\tilde{X}\tilde{X}^T)} \right) \quad (1.6)$$

其中  $\text{diag}(\cdot)$  返回的是一个矩阵主对角线上的元素。

至此，EEG 样本从矩阵转换为了向量特征，可以用于常见的机器学习模型，比如线性判别分析、支持向量机以及逻辑回归等等。

### 1.3.2 基于黎曼几何的分类框架

Barachant 等<sup>[59]</sup>于 2012 年首次将黎曼几何引入到脑机接口，并提出一种新的分类框架，将上一小节所述传统方法流程中的空域滤波和特征提取合并为一个步骤，直接基于 EEG 样本的协方差矩阵之间的距离来对运动想象任务进行分类，发现这种方法具有很好的性能。随后，研究者们在其基础上进行完善，逐渐将黎曼几何方法的应用场景从运动想象范式推广到 P300 诱发电位以及疲劳度监测等范式，并且普遍取得了良好的效果<sup>[60-63]</sup>，本文将这类基于黎曼几何的分类方法统称为黎曼方法。

那么黎曼方法是如何能够将空域滤波和特征提取合为一个步骤的呢？关键在于对 EEG 样本的协方差矩阵进行充分利用。其实在传统方法中，CSP 滤波器就是基于样本的协方差矩阵计算而来。在多通道 EEG 信号中，样本的协方差矩阵代表了其各个通道之间的相关关系，意味着它包含了大脑活动的位置信息（因为通道布置在头皮表面不同位置）。而在脑机接口中，解码大脑的活动区域常常是非常关键的步骤，比如在运动想象任务中，对大脑皮层特定部位感觉运动节律的检测可用于“解读”受试者正在进行哪部分身体的运动想象，一个简单的例子就是左侧大脑控制右手的运动（想象），右侧大脑控制左手的运动（想象）。所以，黎曼方法以样本的协方差矩阵为特征，而无需进行空域滤波。

#### (1) 黎曼流形

黎曼几何是为分析协方差矩阵而引入的一个工具，研究局部可线性逼近的光滑曲线空间，这里曲线空间就是流形，在每一点上的线性近似就是切空间。由于协方

差矩阵具有对称正定的矩阵形式，可视为坐落于对称正定黎曼流形<sup>[61]</sup>上的点，如图1-9所示， $C_1, \dots, C_4$ 代表四个对称正定矩阵，它们之间既可以计算直线距离（欧氏距离），也可以计算曲线距离（黎曼距离）。研究发现，黎曼距离比欧式距离更能准确刻画协方差矩阵之间的距离关系<sup>[48]</sup>。

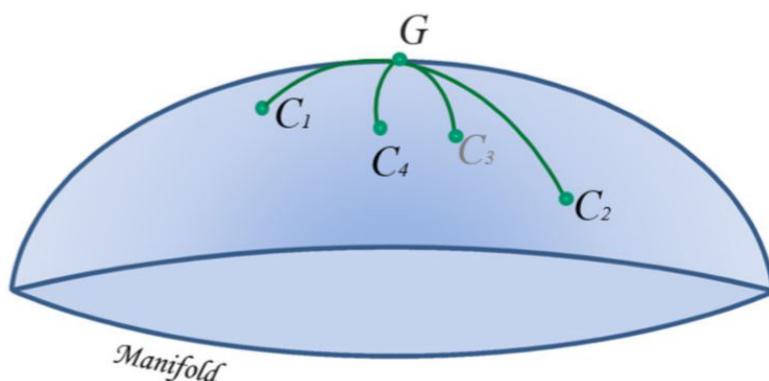


图 1-9 黎曼流形示意图： $C_1, \dots, C_4$ 为黎曼流形上的四个点， $G$ 为它们的中心<sup>[48]</sup>。

对于任意两点  $C_1$  和  $C_2$ ，它们之间的黎曼距离定义为在流形上连接两点的最短曲线的距离，也称为测地线距离：

$$\delta(C_1, C_2) = \|\log(C_1^{-1}C_2)\|_F = \left[ \sum_{r=1}^R \log^2 \lambda_r \right]^{\frac{1}{2}}, \quad (1.7)$$

其中下标  $F$  代表 Frobenius 范数， $\lambda_r$  ( $r = 1, 2, \dots, R$ ) 是矩阵  $C_1^{-1}C_2$  的实特征值。

关于黎曼距离，有一个非常重要的性质叫做仿射不变性（congruence invariance），即两点之间的黎曼距离在经过可逆矩阵转换后保持不变：

$$\delta(W^T C_1 W, W^T C_2 W) = \delta(C_1, C_2), \quad (1.8)$$

其中  $W$  为可逆矩阵，这个性质在本文后续章节会持续用到。

同距离相似，在计算一群对称正定矩阵的中心时，我们既可以计算它们的算术中心（欧式中心），即这些矩阵的算术平均；也可以计算它们的几何中心（黎曼中心）。如图1-9所示， $G$ 代表  $C_1, \dots, C_4$  的黎曼中心，其定义为使得  $G$  到其它点的黎曼距离之和最小的点。

黎曼中心的计算可公式化为：

$$\varrho(C_1, \dots, C_N) = \arg \min_G \sum_{n=1}^N \delta^2(G, C_n). \quad (1.9)$$

需要注意的是，式 (1.9) 没有解析解，而是需要使用梯度下降算法来迭代求解<sup>[64]</sup>。

### (2) MDRM分类器

Barachant 等<sup>[59]</sup>直接基于协方差矩阵之间的黎曼距离来对样本进行分类，称为 MDRM (Minimum Distance to Riemannian Mean)。图1-10表达的是一个二分类问题的分类过程，图中蓝色圆点和绿色圆点分别代表两类样本的协方差矩阵， $G_1$  和  $G_2$  分别代表两类协方差矩阵的黎曼中心，对于一个新的协方差矩阵  $C$  (测试集样本)，MDRM 分别计算它到  $G_1$  和  $G_2$  的黎曼距离，然后把它划分到距离更近的一类。即：

$$g(C) = \arg \min_m \delta(C, G_m), \quad (1.10)$$

其中  $g(C)$  为对新样本的预测类别。

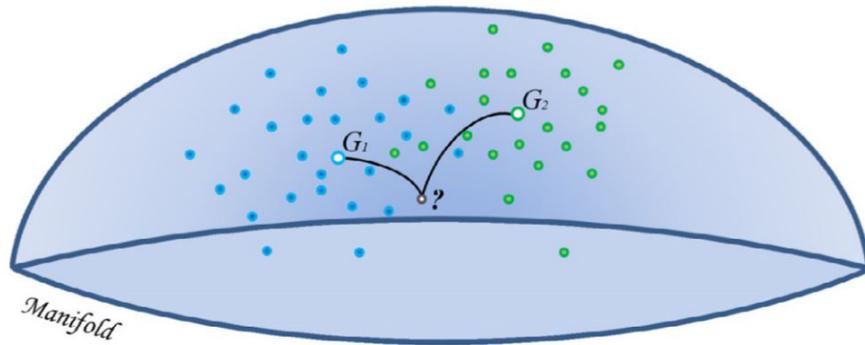


图 1-10 MDRM分类示意图<sup>[48]</sup>。

### (3) 黎曼切空间

除了直接在黎曼流形上对协方差矩阵进行分类，Barachant 等<sup>[59]</sup>还提出了一种把协方差矩阵映射到切空间的分类方法。对于黎曼流形上的任意一点  $C$ ，该点的切空间  $\mathcal{T}_C\mathcal{M}$  定义为以  $C$  为切点的切向量的集合。如图1-11所示， $C_i$  到  $C$  的测地线在切空间  $\mathcal{T}_C\mathcal{M}$  上的投影即为对应的切向量  $S_i$ ：

$$S_i = \text{Log}_C(C_i) = C^{\frac{1}{2}} \log(C^{-\frac{1}{2}} C_i C^{-\frac{1}{2}}) C^{\frac{1}{2}}. \quad (1.11)$$

反过来也可以把切向量  $S_i$  映射到黎曼流形：

$$C_i = \text{Exp}_C(S_i) = C^{\frac{1}{2}} \exp(C^{-\frac{1}{2}} S_i C^{-\frac{1}{2}}) C^{\frac{1}{2}}. \quad (1.12)$$

### (4) P300信号的增广协方差矩阵

前文提到，EEG 样本的协方差矩阵里包含了对解码大脑活动非常关键的空间信息。对于运动想象这样的任务，大脑活动的空间信息直接关系到任务类型的识别，

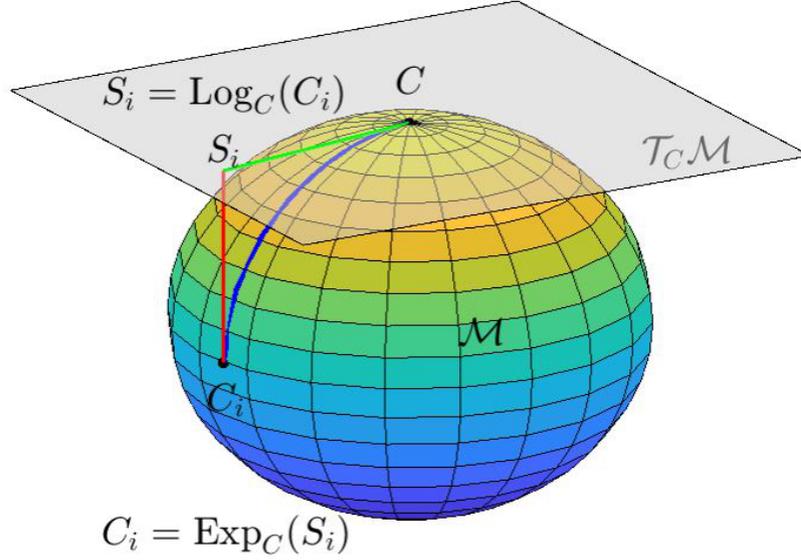


图 1-11 黎曼流形与其切空间之间的映射示意图。

所以可以直接求取协方差矩阵并应用上述分类方法。但是对于 P300 诱发电位，其样本类型可分为正类（目标样本）和负类（非目标样本），其中正类样本会在事件发生约 300 毫秒后出现一个幅值上的波峰，所以其最重要的信息不是关于空间，而是时域信息。因此，上述黎曼方法不直接适用 P300 诱发电位。

为了解决这个问题，Barachant 等<sup>[60]</sup>又于 2014 年采用增广矩阵的方式将时域信息也嵌入到样本的协方差矩阵中。对于已标注的训练数据，首先计算正类样本的均值：

$$\bar{X}_+ = \frac{1}{|\mathcal{I}_+|} \sum_{i \in \mathcal{I}_+} X_i, \quad (1.13)$$

其中  $\mathcal{I}_+$  是训练集正类样本集合， $|\mathcal{I}_+|$  是集合内样本数目。

然后对每一个样本  $X_i$ ，建立一个增广样本：

$$X_i^* = \begin{bmatrix} \bar{X}_+ \\ X_i \end{bmatrix} \quad (1.14)$$

样本  $X_i$  的增广协方差矩阵即为其增广样本的协方差矩阵：

$$\begin{aligned} C_i^* &= X_i^*(X_i^*)^T = \begin{bmatrix} \bar{X}_+ \\ X_i \end{bmatrix} \begin{bmatrix} \bar{X}_+^T & X_i^T \end{bmatrix} \\ &= \begin{bmatrix} C_{\bar{X}_+} & C_{\bar{X}_+, X_i}^T \\ C_{\bar{X}_+, X_i} & C_{X_i} \end{bmatrix} \end{aligned} \quad (1.15)$$

在上式中,  $C_{\bar{X}_+}$  是  $\bar{X}_+$  的协方差矩阵, 这部分对所有样本的增广协方差矩阵都一样, 不包含任何有用信息;  $C_{X_i}$  是原样本  $X_i$  的协方差矩阵, 如前面所解释的, 这部分包含一定的有用信息, 但不足以取得好的效果;  $C_{\bar{X}_+, X_i}$  是  $\bar{X}_+$  与  $X_i$  之间的互协方差矩阵, 这部分包含对样本分类起主要作用的信息。比如, 当  $X_i$  为正类样本时, 互协方差矩阵的对应项就会比较大, 反之当  $X_i$  为负类样本时, 互协方差矩阵的对应项就会接近于零。

## 1.4 迁移学习简介

传统的机器学习模型一般是基于两个假设: (1) 有足量的已标注训练数据; (2) 测试数据与训练数据服从独立同分布。但是现实中的问题却往往难以满足这两个假设, 我们经常会遇到这样的场景: 当前领域中没有或者只有少量已标注训练数据, 同时在相邻领域却有大量的已标注数据。例如在计算机视觉领域有很多开源的已标注数据集可以进行监督学习或者深度分析, 但在一些细分领域, 比如动作识别, 安防等, 仅能获得一些没有标注的视频, 高质量的标注数据相对难以获得。再例如在自然语言处理领域已经积累有大量已标注的英文文本, 但却缺少相应的中文标注数据。

在上述两个例子中, 若只使用当前领域的知识或者数据, 则会缺少标注数据来训练模型; 若使用其它的开源已标注数据, 则会由于测试数据与训练数据的分布不同导致模型泛化能力差。迁移学习<sup>[65]</sup> 研究的就是这个问题, 试图找出相邻领域与当前领域的相关性, 从而顺利地将相邻领域学习到的知识或者能力迁移到一个新的、标注数据缺乏的领域。这种迁移能力是人类天生具有的, 比如一个会骑自行车的人通常能够更快地学会骑摩托车; 一个会打网球的人, 他对球的判断能力可以帮助他学会乒乓球。因此, 有研究者认为, 迁移学习是实现更高水平人工智能的一个关键技术。

### 1.4.1 问题定义

首先介绍领域 (domain) 和任务 (task) 两个概念, 其中领域  $\mathcal{D}$  包含特征空间  $\mathcal{X}$  和边缘概率分布  $P(x)$  两部分, 即  $\mathcal{D} = \{\mathcal{X}, P(x)\}$ ; 给定一个领域, 任务  $\mathcal{T}$  则由类别空间  $\mathcal{Y}$  和预测函数  $f(\cdot)$  组成, 即  $\mathcal{T} = \{\mathcal{Y}, f(\cdot)\}$ , 其中  $\mathcal{Y}$  是所有类别的集合,  $f(\cdot)$  用来预测一个样本的标签。从概率论的角度看,  $f(\cdot)$  也可以写作条件概率分布  $P(y|x)$ 。

**定义 1.1:** 迁移学习: 给定一个已标注样本的源域  $\mathcal{D}_s = \{(X_i, y_i)\}_{i=1}^{N_s}$  和学习任务  $\mathcal{T}_s$ , 一个未标注的目标域  $\mathcal{D}_t = \{X_i\}_{i=1}^{N_t}$  和学习任务  $\mathcal{T}_t$ , 且  $\mathcal{D}_s \neq \mathcal{D}_t$  或者  $\mathcal{T}_s \neq \mathcal{T}_t$ 。迁移学习的目标是尽量降低目标域预测模型的泛化误差。

## 1.4.2 迁移学习类型

根据源域与目标域的特征空间  $\mathcal{X}$ 、边缘概率分布  $P(x)$ 、类别空间  $\mathcal{Y}$ 、条件概率分布  $P(y|x)$  的异同，迁移学习可以划分为不同类型<sup>[66]</sup>：

1. 特征空间相同，但边缘概率分布不同；
2. 特征空间与边缘概率分布都不相同；
3. 类别空间相同，但条件概率分布不同；
4. 类别空间和条件概率分布都不相同。

### (1) 同构迁移学习

同构迁移学习研究的场景是特征空间和类别空间都相同，但边缘概率分布或（且）条件概率分布不同，即  $\mathcal{X}_s = \mathcal{X}_t$ ， $\mathcal{Y}_s = \mathcal{Y}_t$ ，但  $P_s(x) \neq P_t(x)$  或（且） $P_s(y|x) \neq P_t(y|x)$ 。这种场景相对异构迁移学习简单，也是目前绝大多数迁移学习方法所研究的。

由于源域与目标域有相同的输入特征维度，分类目标也不变，所以同构迁移学习主要是弥补领域间分布的差异。这类方法通常是基于实例的或者基于特征表示的。Zadrozny 等<sup>[67]</sup>分别估计源域与目标域的边缘概率分布  $P_s(x)$  与  $P_t(x)$ ，然后通过采样来减少  $P_s(x)$  与  $P_t(x)$  之间的分布差异。文献<sup>[68,69]</sup>则省去对  $P_s(x)$  与  $P_t(x)$  的估计，通过对源域样本的赋权来显式减少它们的分布差异。Pan 等<sup>[70]</sup>则学习一个低维的特征空间表示，使得  $P_s(x)$  与  $P_t(x)$  在新空间的分布差异最小化。Long 等<sup>[71]</sup>同样学习一个低维的特征空间表示，但学习目标是使得新空间中源域与目标域之间的边缘概率分布和条件概率分布同时减少。

### (2) 异构特征空间迁移学习

异构特征空间迁移学习<sup>[66]</sup>是指源域与目标域的特征空间不同，这种情形下源域与目标域之间没有共享的特征或者特征维度不同，通常需要先对特征空间进行变换。这是一个比同构迁移学习更具挑战性的问题，因为域之间的共性更少。换句话说，知识可以从源数据获得，但其表示方式与目标数据不同，问题在于如何提取它。

异构特征空间迁移学习的典型应用场景是跨语言的自然语言处理，即源域与目标域的数据来自不同语种。Wei 和 Pal<sup>[72]</sup>首先采用机器翻译将源域文本翻译成目标域文本语种，将异构特征空间迁移学习转化为同构迁移学习问题，再处理源域与目标域之间的分布差异问题。另一个应用场景是动作识别，比如有两个装有摄像头的房间，由于它们的摄像头的数目和位置不同，导致数据的特征属性和维度不同，一个简单的解决办法是挑选出最具共性的摄像头作为共同特征空间。典型相关分析（Canonical Correlation Analysis，CCA）也常常被用来解决这类多视图数据之间的迁移问题，Wu 等<sup>[73]</sup>提出典型相关的异质转移判别分析（HTDCC），同时最小化类间样本的相关性和最大化类内样本的相关性。Samat<sup>[74]</sup>等则使用了多视图 CCA 进行集

成的方法。Yeh 等<sup>[75]</sup>提出了一种利用核 CCA 进行对称特征变换的跨域模式识别框架，并提出了一种改进的支持向量机。

大多数异构迁移学习解决方案在转换特征空间时可分为两类：对称转换和非对称转换<sup>[76]</sup>。对称变换学习源域特征空间  $\mathcal{X}_s$  和目标域特征空间  $\mathcal{X}_t$  的特征变换，将其投影到公共子空间  $\mathcal{X}_c$  上进行自适应。非对称转换将源域特征空间  $\mathcal{X}_s$  变换为与目标域特征空间  $\mathcal{X}_t$  对齐 ( $\mathcal{X}_s \rightarrow \mathcal{X}_t$ )。

### (3) 异构类别空间迁移学习

异构类别空间迁移学习是指源域与目标域的类别空间不同，这意味着源域与目标域的分类任务不同，是一个比异构特征空间更难的问题。综述文献<sup>[66]</sup>总结了数十种异构迁移学习方法，其中只有 3 种是针对异构类别空间问题的。

异构类别空间迁移学习通常依赖于领域特定的先验知识，比如多视图数据各个视图之间的关系等等。Shi 等<sup>[77]</sup>基于贝叶斯研究源域预测函数与目标域预测函数之间的相互关系。Moon 和 Carbonell<sup>[78]</sup>提出了一个旨在解决文本分类任务的异构迁移学习框架，该方法利用一个基于跳跃图的语言模型<sup>[79]</sup>，学习词向量的语义表示，从而将源域和目标域的类别映射到一个词向量嵌入空间。

当缺乏领域特定知识时，一般很难进行异构类别空间迁移学习。本文将在第五章提出一种类别空间对齐方法，用于 EEG 数据的异构类别空间迁移学习问题。另外，本文的“异构”不同于多模态数据的设定，源域和目标域依然是同一种类型的数据 (EEG)，甚至是同一种控制范式 (运动想象或者 P300)，其“异构”来源于不同脑电设备或者执行不同类型的任务。

## 1.4.3 迁移学习方法

本节介绍迁移学习领域的两类主流方法<sup>[80,81]</sup>：基于样本 (实例) 的方法与基于特征的方法。其中基于样本的方法着重于调整源域样本的权重，提升位于目标域高密度区域的源域样本的权重，减少位于目标域低密度区域的源域样本的权重，从而使源域数据与目标域数据在分布上匹配度更高。而基于特征的方法则着重于学习新的特征表示，使得源域数据与目标域数据在新的特征空间中分布更为接近。

### (1) 基于样本的迁移学习方法

基于样本的迁移学习方法是根据一定的权重生成规则，调整源域样本权重，简单来说，就是重用那些被认为更“重要”的源域样本。当然，有不同的方法来衡量源域样本的“重要性”。比如，早期的工作<sup>[67]</sup>估计源域与目标域的边缘概率分布比值  $P_s(x)/P_t(x)$ ，该比值就是源域样本的权重。但其实这种比值并不容易准确地测量，Huang 等<sup>[68]</sup>提出核均值匹配法对概率分布进行估计，目标是使得加权后的源域和目标域的概率分布尽可能相近。Dai 等<sup>[82]</sup>基于 AdaBoost 的思想，提出一种迭代更新源域样本权重的迁移学习方法 TrAdaBoost，该方法要求目标域有少量的已标注

数据，然后根据基分类器的结果增加误分类源域训练样本的权重，同时减少误分类目标域训练样本的权重，并理论证明了方法的泛化误差上界。

一般来说，基于样本的方法有较好的理论支撑但通常适用于源域与目标域分布差异较小的场景。

## (2) 基于特征的迁移学习方法

基于特征的迁移学习方法是把源域的特征映射到目标域的特征空间，或者是同时把源域和目标域映射到一个共同的子空间，其目的是使得源域与目标域在新的特征表示下分布更相似。

概率分布适配是最常见的基于特征的迁移学习方法，通常是先定义一个分布的距离函数，再通过新的特征表示来显式地减少源域与目标域之间的分布距离。最大均值差异 (Maximum Mean Discrepancy) [83] 是常用的度量概率分布差异的标准，Pan 等 [70] 提出迁移学习方法 MMDE，旨在最小化源域与目标域的边缘概率分布的 MMD 距离，以及最大化核空间中的数据嵌入方差。Long 等 [71] 提出联合分布适配 (joint distribution adaptation, JDA)，不仅最小化边缘概率分布的 MMD 距离，还定义了条件概率分布的 MMD 距离并使它们同时最小化。Wang 等 [84] 则在联合分布适配的基础上补充了流形正则化项。Persello 等 [85] 则提出了一种基于核的高光谱图像特征选择方法，定义一种新的条件分布偏移度量来评估特征的稳定性。

除了概率分布适配，Sun 等 [86] 提出一个简单的方法，以白化的方式对齐源域与目标域的二阶统计量。Herath 等 [87] 基于马氏距离把不同域映射到不变潜在空间，并同时最小化域的方差，与最大化类别区分能力。Zhang 等 [88] 提出一种联合几何分布对齐 (Joint Geometrical and Statistical Alignment, JGSA)，学习两个耦合投影，将源域和目标域数据投影到低维子空间中，同时减少几何位移和分布位移。Lu 等 [89] 只使用类均值来学习特定于类的线性投影，并形成了一个类似线性判别分析的框架，从而得到一个有效的解析解。Shu 等 [90] 利用内积度量的边际 Fisher 准则对源域进行特征映射，然后提出一种去除边界奇异点的算法来选择源域密集区域样本点，与目标域中的标记样本点组成训练样本对并得到特征映射矩阵。

Long 等 [91] 提出联合匹配迁移方法 (Transfer Joint Matching, TJM) 同时使用特征匹配和样本权重调整两种策略。

此外，综述文献 [65] 也介绍了基于模型的迁移学习和基于知识关系的迁移学习。且近几年来，随着深度学习的兴起，深度迁移学习也收到了更多的关注 [92,93]，本文第 4, 5 章都是直接对数据进行变换的方法，后续也可以使用深度学习模型。

## 1.5 脑机接口的主要挑战

由于脑机接口是一个具有高度学科交叉特性的研究领域。它一方面需要脑科学、神经科学的知识，另一方面也需要计算机，工程学领域的的能力。所以，脑机接口当

前面临的挑战大致来自脑科学、硬件、算法三个方面<sup>[94]</sup>。

要想实现高级的脑机接口，首先需要对大脑的高级思维形成有充分的理解，并能够把这种思维提取出来。目前的技术还仅仅能够理解一些简单、基础的部分。事实上，虽说现在以人工神经网络为代表的“人工智能”取得了很大的发展以及应用成果，但依然是基于数据去拟合一个函数模型，只能在特定的任务上起作用，没办法实现真正的生物的智能，即形成思考、思维的能力。有一个重要的派别认为要想实现“强人工智能”或“通用人工智能”，还需要从理解人的大脑的工作方式出发。但是，人的大脑内有百亿级的神经元，且以非常复杂的方式相互连接，理解大脑的难度可比拟理解宇宙的难度。

然后是硬件方面，主要是指需要有很好的信号采集装置以及控制模块。对于侵入式脑机接口，如何以极小的创伤在大脑内部安装电极，使用生物相容的材料来制作电极而不引起生物的免疫反应；对于非侵入式脑机接口，如何更高质量地采集到微弱的电信号都是研究者非常关注的问题。

最后是算法方面，脑机接口还需要优秀的信号处理和机器学习算法来识别大脑活动的模式。算法方面的一个主要挑战是，由于个体差异，即不同用户，甚至同一用户在不同时期，对相同刺激的神经反应不同，导致在脑机接口中很难建立一个适用于所有用户的通用模型。这意味着历史标注数据，无论是其他辅助用户的已标注数据还是当前用户过去的已标注数据，都没办法直接用于当前用户的模型的建立。所以在实际应用中，每个新用户在使用脑机接口时都需要一段较长时间的校准，校准过程是为了采集该用户的标注数据，从而能够为该用户建立相应的机器学习模型。

脑机接口的校准过程在用户体验上不友好，很大程度上影响了其在实际生活中的应用。针对这个问题，本文研究用迁移学习解决脑机接口中校准时间长的问题，即使用迁移学习来弥补辅助用户（本文称为“源用户”，对应迁移学习概念中的“源域”）与新用户（本文称为“目标用户”，对应迁移学习概念中的“目标域”）之间的数据差异，使得辅助用户的数据可以用于新用户模型的建立，从而减少甚至消除新用户所需要的校准数据。图1-12 为在脑机接口中使用传统机器学习与迁移学习的对比示意图。

如图1-12所示，传统机器学习只使用新用户自身的数据，所以校准时间长；迁移学习则充分利用其他用户的数据，从而减少新用户的校准时间。同时，由于脑机接口独特的原理、数据形式、特征方法，通常需要为不同类型的脑机接口设计独立的迁移学习方法<sup>[95-98]</sup>。

## 1.6 本文的研究内容与组织结构

本文针对脑机接口中校准时间长的问题，研究如何使用迁移学习来减少甚至消

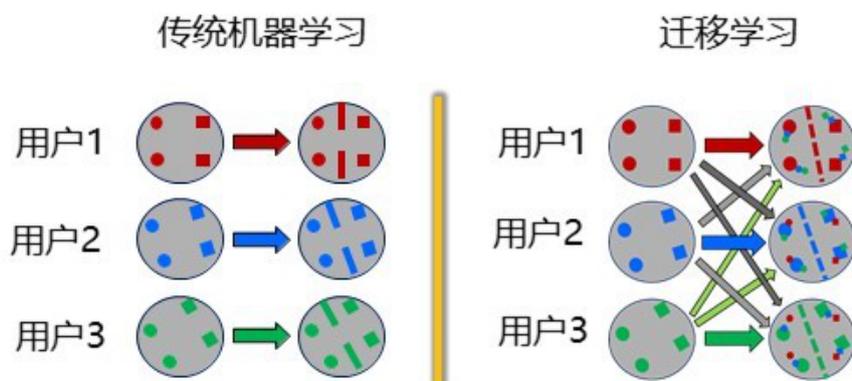


图 1-12 脑机接口中传统机器学习与迁移学习的对比。

除新用户所需要的校准数据。全文组织可看作“总-分-总”结构，即第 1 章介绍问题的背景与来源，第 2 至 5 章从不同的角度提出了新的解决问题的方法，最后第 6 章总结全文并展望未来的工作。

由浅入深，本文的第 2 至 5 章之间也存在一定的组织逻辑。它们之间的结构关系可以从两个角度来看：脑机接口与迁移学习。

从脑机接口的角度看，本文的组织架构如图1-13所示。1.3小节介绍了脑机接口中的两种基本分类框架，本文第 2, 3 章则分别研究这两种分类框架中的算法，其中第 2 章研究如何将迁移学习应用于传统算法框架中的共同空间模式，第 3 章研究如何将迁移学习应用于黎曼方法框架。此外，本文第 4, 5 章都是针对 EEG 数据开发新的迁移学习方法，数据在被对齐后既可以用于传统算法框架，也可以用于黎曼方法框架。

从迁移学习的角度看，本文的组织架构如图1-14 所示。其中第 2 至 4 章都是研究同构迁移学习场景，第 5 章研究的则是更具挑战性的异构类别空间迁移学习。

具体来看：

第 1 章为绪论部分，介绍脑机接口的研究背景、组成部分、基本算法框架与主要挑战，迁移学习的概念与基本方法，以及本文为何要研究脑机接口中的迁移学习方法。

第 2 章研究如何将迁移学习应用于共同空间模式算法。首先总结文献中相关方法的特点，即它们多是以用户为单位进行迁移；然后提出一种新的迁移学习方法，即以样本为单位进行迁移。该方法分别以脑电样本及其协方差矩阵的切空间向量为特征，基于最大均值差异对源用户的样本特征加权以减少其与目标用户特征之间的分布差异，然后对源用户样本的协方差矩阵进行独立赋权，并将赋权后的协方差矩阵用于新用户共同空间模式滤波器的计算。实验结果显示本文所提方法能够有效减

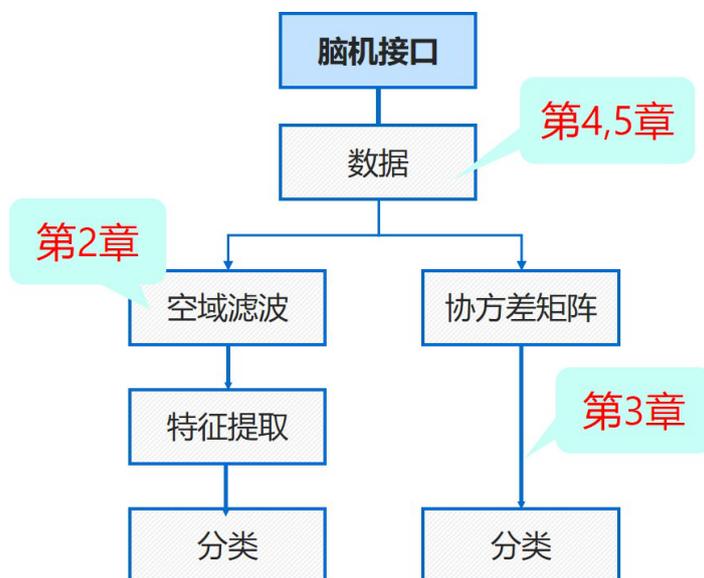


图 1-13 从脑机接口角度看文章组织结构。

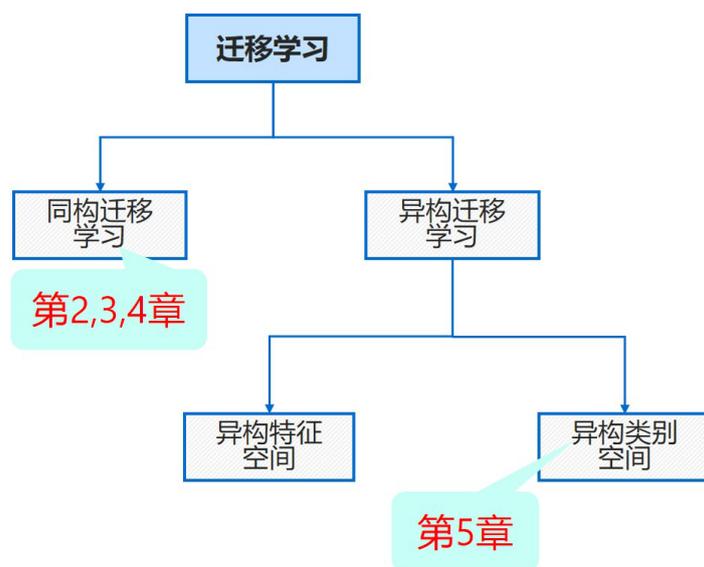


图 1-14 从迁移学习角度看文章组织结构。

少新用户建立共同空间模式滤波器所需要的标注数据。此外，第2章还针对共同空间模式的目标函数开展了研究，提出了一个新的目标函数并将之与传统目标函数进行了比较。

第3章首先介绍了迁移学习在脑机接口领域当前最先进的算法，然后针对该算法提出了两种优化方法。其一通过选择通道来减少协方差矩阵的维度，从而提高黎曼空间中有关协方差矩阵计算的效率和准确度；其二是基于样本协方差矩阵之间的黎曼距离进行聚类，挑选出位于目标用户高密度区域的源用户样本作为训练集。实验表明这两种优化方法都能提升当前最好迁移学习方法的效果，且当两种优化方法结合使用时，效果能进一步得到提升。

第4章针对脑电信号提出一种无监督的数据对齐方法，可以作为一种预处理步骤，并用于不同的特征提取和分类算法。该对齐方法以单个用户的平均协方差矩阵作为参考矩阵，然后基于仿射不变性对用户进行映射，使得映射后的样本协方差矩阵以单位矩阵为中心。通过对不同用户分别实施上述操作，它们的样本协方差矩阵的分布得以对齐。实验结果显示该方法的迁移效果和效率都优于现有最好的迁移学习方法之一。

第5章针对迁移学习中最具挑战性的异构类别空间问题，提出一种异构类别空间的对齐方法。首先通过聚类挑选出目标用户最具信息量的样本进行标注，并基于这些少量的标注样本衡量目标用户各个类别的中心。然后为每一个源用户类别设置一个匹配的目标用户类别，再基于仿射不变性，将源用户样本按类别进行映射，使其中心与目标用户匹配类别的中心重叠，从而将一个源用户类别对齐至一个目标用户类别。实验表明，该方法在多种异构类别空间场景中都有显著效果，且可以与其它迁移算法结合使用来进一步提升迁移效果。

第6章总结全文，并展望了未来可以继续研究的工作。

## 2 基于迁移学习的共同空间模式

### 2.1 引言

本文在上一章（1.3节）介绍了脑机接口中的机器学习，文中提到，目前脑机接口中常用的机器学习方法可以分为两个不同的框架：

1. 基于欧式空间的传统分类方法，一般包括空域滤波、特征提取、分类三个步骤；
2. 基于黎曼几何的方法，在黎曼空间直接针对 EEG 样本协方差矩阵进行操作，可以将传统方法中的空域滤波和特征提取合为一个步骤。

本章研究的是基于欧式空间的传统分类方法，具体来讲，研究如何将迁移学习应用于传统方法中最常用且重要的空域滤波器，即共同空间模式（common spatial pattern, CSP）。由于 CSP 是有监督算法，所以要求当前用户有足量的已标注数据。然而现实中的场景却往往不满足这个条件，而是目标用户只有少量，甚至没有已标注数据，但同时可以获得多个已标注的源用户的数据。所以，本章研究如何充分利用源用户的已标注数据，建立一组泛化性能好，可以有效用于目标用户滤波的 CSP 滤波器。

回顾1.3.1节对 CSP 的介绍，滤波器  $W = [W_0, W_1]$  的目标函数为：

$$W_0 = \arg \max_W \frac{\text{tr}(W^T \bar{C}_0 W)}{\text{tr}(W^T \bar{C}_1 W)} \quad (2.1)$$

$$W_1 = \arg \max_W \frac{\text{tr}(W^T \bar{C}_1 W)}{\text{tr}(W^T \bar{C}_0 W)} \quad (2.2)$$

其中  $\bar{C}_0$  和  $\bar{C}_1$  分别代表类别 0 和类别 1 的平均协方差矩阵。而  $W$  的解是矩阵  $\bar{C}_1^{-1} \bar{C}_0$ （或者  $\bar{C}_0^{-1} \bar{C}_1$ ）最大以及最小的  $f$  个特征值对应的特征向量。

也就是说，CSP 滤波器的计算完全是基于对当前用户每个类别的平均协方差矩阵的衡量。因此，在使用源用户数据辅助建立针对目标用户的 CSP 滤波器时，我们实际是使用源用户的协方差矩阵来辅助衡量目标用户的类别平均协方差矩阵。

鉴于源用户与目标用户之间的个体差异，该如何高效地利用源用户的协方差矩阵呢？一个简单的想法是计算每个源用户的平均协方差矩阵，再将多个源用户的结果平均，这种多用户的平均通常会导致一定的用户通用性，所以可以将平均之后计算得到的 CSP 滤波器用于目标用户的空域滤波。这种方法通常会有一定效果，却又效果有限，因为不同用户之间的个体差异也是不同的，但是这种方法同等地对待

不同的源用户，甚至同等地对待源用户与目标用户的已标注数据，显然有不合理之处。

因此，文献中提出了一些更精细的方法<sup>[99,100]</sup>来处理这个问题，其中最常见的就是在综合多个源用户时给它们以不同的权重。本文将这类方法统称为基于用户的协方差矩阵迁移，意思是迁移操作的对象在用户层面，即给每个源用户的平均协方差矩阵以不同的权重。

在文献已有的方法之外，本章提出一种新的思路来结合多个源用户的协方差矩阵，称为基于样本的协方差矩阵迁移。它将协方差矩阵的权重赋予操作从用户层面下沉到样本层面，即给每个源用户的每个样本协方差矩阵以不同的权重。此外，无论是基于用户的协方差矩阵迁移还是基于样本的协方差矩阵迁移，都会平衡源用户数据与目标域的已标注数据。

本章内容安排如下：2.2节介绍文献中两种基于用户的协方差矩阵迁移方法；2.3节提出一种新的基于样本的协方差矩阵迁移方法；2.4节提出一个新的 CSP 的目标函数；2.5节介绍本章实验，首先比较2.3节所提的基于样本的协方差矩阵迁移方法与基于用户的协方差矩阵迁移方法，然后比较2.4节所提的 CSP 目标函数与传统的目标函数。最后，2.6节总结本章内容。

## 2.2 基于用户的协方差矩阵迁移

本节介绍两种典型的基于用户的协方差矩阵迁移方法（Subject-based transfer Approach, SA），并以 SA1 和 SA2 代称。它们都是在用户层面进行协方差矩阵迁移，其中 SA1 的思路是挑选出一部分与目标用户相似的源用户，其它不相似的源用户则直接舍弃；SA2 的思路则是使用所有源用户，但是会基于相似度给每个源用户赋予不同的权重。并且 SA1 与 SA2 采取了不同的方法来衡量源用户与目标用户的“相似性”。

### 2.2.1 SA1

Lotte 和 Guan<sup>[99]</sup> 提出了针对 CSP 的迁移学习方法 SA1，其主要思路是先基于相似性挑选出一部分与目标用户相似的源用户，再平均这些已挑选源用户的数据，最后以一定的权重平衡源用户数据与目标用户的已标注数据。

SA1 在挑选源用户时需要一部分目标用户已标注数据，所以它考虑的是一个半监督学习场景：假设有一个目标用户  $D_T$ ，包含少量的已标注数据  $D_T^l$  和大部分的未标注数据  $D_T^u$ ，另有  $Z$  个已标注的源用户  $D_S^z$ ， $z = 1, 2, \dots, Z$ 。

首先介绍 SA1 是怎样选择源用户的：基于搜索的思想，SA1 比较不同源用户组合所训练的模型在目标用户已标注数据上的测试结果，若测试结果好，则说明对应源用户组合与目标用户相似度高，算法1 详细描述了源用户的选择过程。

**Algorithm 1:** SA1 的用户选择算法。

**Input:**  $D_T^l$ : 目标用户已标注数据;

$\Omega = \{D_S^z\}_{z=1}^Z$ :  $Z$  个已标注的源用户集合。

**Output:**  $S_t(\Omega)$ :  $\Omega$  的一个子集, 即被选中的源用户集合;

*selectedAcc*: 使用  $S_t(\Omega)$  训练的模型在  $D_T^l$  上取得的分类正确率。

初始化已被选择的源用户集合  $S_t(\Omega) = \{\}$ ;

初始化未被选择的源用户集合  $R_e(\Omega) = \Omega$ ;

初始化计数器  $n = 1$ ;

初始化分类正确率 *selectedAcc* = 0;

**while**  $n < Z$  **do**

**步骤一:**

**for**  $z \in R_e(\Omega)$  **do**

        从  $R_e(\Omega)$  中选一个源用户  $D_S^z$ , 加入到  $S_t(\Omega)$  中;

        使用  $S_t(\Omega)$  训练模型并在  $D_T^l$  上测试得到分类正确率  $A_z$ ;

        将源用户  $D_S^z$  从  $S_t(\Omega)$  中移除;

**end**

    选择正确率最高的  $A_z$ , 将对应源用户  $D_S^z$  从  $R_e(\Omega)$  中转移到  $S_t(\Omega)$ ;

*selectedAcc* =  $A_z$ ;

$n = n + 1$ ;

**步骤二:**

**for**  $z \in S_t(\Omega)$  **do**

        将用户  $D_S^z$  从  $S_t(\Omega)$  中移除;

        使用  $S_t(\Omega)$  训练模型并在  $D_T^l$  上测试得到分类正确率  $A_z$ ;

        将  $D_S^z$  加入到  $S_t(\Omega)$  中;

**end**

    选择正确率最高的  $A_z$ , 及其对应源用户  $D_S^z$ ;

**if**  $A_z > \textit{selectedAcc}$  **then**

        将  $D_S^z$  从  $S_t(\Omega)$  中转移到  $R_e(\Omega)$ ;

*selectedAcc* =  $A_z$ ;

$n = n + 1$ ;

        跳转到步骤二;

**else**

        跳转到步骤一;

**end**

**end**

**Return**  $S_t(\Omega)$ ; *selectedAcc*。

在源用户子集  $S_t(\Omega)$  选择完毕后，基于式 (1.1) 和式 (1.2)，首先计算每个被选择源用户的类别平均协方差矩阵  $\bar{C}_s^z$ ，再基于目标用户已标注数据计算类别平均协方差矩阵  $\bar{C}_t$ ，然后将  $\bar{C}_t$  与  $\bar{C}_s^z$  结合起来：

$$\tilde{C}_t = (1 - \lambda)\bar{C}_t + \frac{\lambda}{|S_t(\Omega)|} \sum_{z \in S_t(\Omega)} \bar{C}_s^z \quad (2.3)$$

其中  $\tilde{C}_t$  代表结合后的类别平均协方差矩阵，将被用于目标用户 CSP 滤波器的计算， $\lambda$  为权重因子，其计算过程为：

$$\lambda = \begin{cases} 1, & targetAcc \leq randAcc \\ 0, & targetAcc \geq selectedAcc \\ \frac{selectedAcc - targetAcc}{1 - randAcc}, & otherwise \end{cases} \quad (2.4)$$

其中  $targetAcc$  是在  $D_T^l$  上采用留一法交叉验证得到的正确率， $randAcc$  是随机的分类正确率， $selectedAcc$  代表使用  $S_t(\Omega)$  训练的模型在  $D_T^l$  上取得的分类正确率。

最后使用不同类别的  $\tilde{C}_t$  基于式 (2.1) 和 (2.2) 计算 CSP 滤波器。

## 2.2.2 SA2

SA1 选择了与目标用户关联更紧密的源用户，然后给这些源用户统一的权重。Kang 等<sup>[100]</sup> 则提出了一种基于 KL 散度的迁移学习方法 SA2，首先基于 KL 散度计算每个源用户与目标用户的相似度，然后根据相似度给每个源用户的平均协方差矩阵赋予不同的权重。不同于 SA1 的半监督场景，SA2 在计算源用户权重时不需要目标用户已标注样本，但是其未标注样本的协方差矩阵是可以使用的。

首先基于 KL 散度计算每个源用户与目标用户的分布差异。假设  $p_s^z$  代表源用户  $D_S^z$  的协方差矩阵分布， $p_t$  代表目标用户  $D_T$  的协方差矩阵分布，它们之间的 KL 散度定义为：

$$KL(p_s^z, p_t) = \frac{1}{2} \left\{ \log \left( \frac{|\bar{C}_t|}{|\bar{C}_s^z|} \right) + \text{tr}[\bar{C}_t^{-1} \bar{C}_s^z] - n_c \right\}, \quad z = 1, \dots, Z \quad (2.5)$$

其中  $\bar{C}_s^z$  代表源用户  $D_S^z$  的平均协方差矩阵， $\bar{C}_t$  代表目标用户的平均协方差矩阵， $\text{tr}(\cdot)$  代表矩阵的迹， $n_c$  则是通道数目，代表了协方差矩阵的维度。

然后基于 KL 散度判断源用户与目标用户的相似度，进而为源用户确定权重：

$$\alpha_z = \frac{1}{\gamma} \cdot \frac{1}{KL(p_s^z, p_t)} \quad (2.6)$$

其中  $\gamma = \sum_{z=1}^Z \frac{1}{KL(p_s^z, p_t)}$  是归一化因子。

再把权重赋予每一个源用户的类别平均协方差矩阵，并将它们与目标用户的类别平均协方差矩阵结合：

$$\tilde{C}_t = (1 - \lambda)\bar{C}_t + \lambda \sum_{z=1}^Z \alpha_z \bar{C}_s^z \quad (2.7)$$

其中  $\tilde{C}_t$  是结合后的类别平均协方差矩阵，将被用于目标用户 CSP 滤波器的计算； $\lambda$  为权重因子，SA2 希望根据实际情况（比如目标用户已标注的数据量）来调整权重，通常来讲， $\lambda = 0.5$  是一个不错的选择。

### 2.3 基于样本的协方差矩阵迁移

本节提出一种新的方法来处理源用户的协方差矩阵，与 SA 相比，它将协方差矩阵迁移从用户层面下沉到样本层面，所以本文称之为“基于样本的协方差矩阵迁移”（Instance-based transfer Approach, IA）。对每一个源用户，IA 试图给其中每个样本一个权重，以使得赋权后的源用户数据分布与目标用户更相似。

因此，首先需要有一个度量方式来衡量分布之间的距离，SA2 采用的是 KL 散度，IA 采取的则是“最大均值距离”（Maximum mean discrepancy, MMD）<sup>[101]</sup>。仍然假设源用户  $D_s^z$  的数据分布为  $p_s^z$ ，有  $N_s$  个样本，目标用户  $D_T$  的分布为  $p_t$ ，有  $N_t$  个样本。则源域与目标域之间的最大均值距离为：

$$MMD(p_s^z, p_t) = \left\| \frac{1}{N_s} \sum_{i=1}^{N_s} \Phi(X_i) - \frac{1}{N_t} \sum_{j=1}^{N_t} \Phi(X_j) \right\|_H^2 \quad (2.8)$$

其中  $\Phi(\cdot)$  代表映射，把原变量映射到可再生希尔伯特空间，可再生希尔伯特空间对函数内积是完备的。

简单地理解，MMD 是求两个分布在新空间里的均值之间的距离，是迁移学习中最常用的距离度量之一。Long 等<sup>[71]</sup> 曾基于此距离提出一种进行联合概率分布适配的迁移学习方法，Huang 等<sup>[68]</sup> 则基于此距离提出一种对样本的赋权方法来减少协变量偏移。本文沿用这个思想，将权重用于协方差矩阵的赋权以及平均协方差矩阵的计算，以期减少协方差矩阵的分布距离，优化的目标函数为：

$$\begin{aligned} \min_{\beta} & \left\| \frac{1}{N_s} \sum_{i=1}^{N_s} \beta_i \mathbf{x}_s^i - \frac{1}{N_t} \sum_{j=1}^{N_t} \mathbf{x}_t^j \right\|_H^2 \\ \text{s.t.} & \quad 0 \leq \beta_i \leq 1, \quad i = 1, \dots, N_s \\ & \quad \left| \sum_{i=1}^{N_s} \beta_i - N_s \right| \leq N_s \epsilon \end{aligned} \quad (2.9)$$

其中  $\mathbf{x}_s^i$  代表源用户第  $i$  个样本的特征向量,  $\beta_i$  是  $\mathbf{x}_s^i$  的权重,  $\mathbf{x}_t^j$  代表目标用户第  $j$  个样本的特征向量。

式 (2.9) 为一个二次规划问题, 并可以使用 MATLAB 函数 “*Quadratic*” 求解。但是由于 EEG 样本是矩阵形式, 所以需要先将 EEG 样本矩阵转化为向量, 本节考虑两种转化方式:

1. IA1: 直接将多个通道的数据首尾连接, 合并为一个向量, 但是这样通常会得到过大的特征维度, 所以需要通下采样或者减少样本时间来降低维度。
2. IA2: 分别将每个用户的协方差矩阵从黎曼空间映射到切空间, 然后采用切空间向量作为特征向量。

在对每个源用户样本求得权重后, 将权重用于协方差矩阵的结合:

$$\tilde{C}_t = (1 - \lambda)\bar{C}_t + \lambda \sum_{z=1}^Z \sum_{i=1}^{N_z} \beta_i C_i^z \quad (2.10)$$

其中  $\tilde{C}_t$  是结合后的类别平均协方差矩阵, 将被用于目标用户 CSP 滤波器的计算;  $\lambda$  为权重因子, 用来平衡源域与目标域的权重, 其取值与 SA2 中的一样;  $C_i^z$  是第  $z$  个源用户的第  $i$  个样本的协方差矩阵,  $N_z$  为第  $z$  个源用户的样本数。

## 2.4 共同空间模式中的目标函数研究

本节对 CSP 的目标函数开展讨论, 研究原目标函数可能存在的缺点, 从而提出一种新的目标函数以及解法, 最后再在实验部分进行比较。

首先把式 (2.1) 和 (2.2) 中的两个优化目标综合为一个:

$$Ratio1 = \arg \max_{W=[\mathbf{w}_1, \dots, \mathbf{w}_{2f}]} \sum_{i=1}^f \frac{\mathbf{w}_i^T \bar{C}_0 \mathbf{w}_i}{\mathbf{w}_i^T \bar{C}_1 \mathbf{w}_i} + \sum_{i=f+1}^{2f} \frac{\mathbf{w}_i^T \bar{C}_1 \mathbf{w}_i}{\mathbf{w}_i^T \bar{C}_0 \mathbf{w}_i} \quad (2.11)$$

其中  $w_i$  为第  $i$  个滤波器, 总共有  $2f$  个滤波器。

但是, 我们发现这样求解的 CSP 滤波器可能存在两个缺点:

1. 尽管  $C_0$  和  $C_1$  都是对称矩阵, 但是  $C_1^{-1}C_0$  或者  $C_0^{-1}C_1$  却通常不是对称的, 所以  $C_1^{-1}C_0$  和  $C_0^{-1}C_1$  的特征值并不相互正交<sup>[102]</sup>。也就是说, 使用 *Ratio1* 优化得到的 CSP 滤波器之间是相关的, 可能会导致滤波之后的特征之间有冗余信息。
2. *Ratio1* 与最终的分类正确率并不呈严格正相关关系。我们在实验中发现, 在对  $C_1^{-1}C_0$  进行特征分解时, 取最大的几个特征值对应的特征向量有时候效果不如取另外的特征向量。虽然现在也有研究开始探索其它挑选特征向量 (滤波器) 的准则<sup>[103,104]</sup>, 但效果都不甚如意。

### 2.4.1 一种新的CSP目标函数

因此，本节考虑使用另一个相关的优化函数来获取滤波器，以使得滤波器之间没有相关性：

$$Ratio2 = \arg \max_{W=[\mathbf{w}_1, \dots, \mathbf{w}_{2f}]} \frac{\sum_{i=1}^f \mathbf{w}_i^T \bar{C}_0 \mathbf{w}_i}{\sum_{i=1}^f \mathbf{w}_i^T \bar{C}_1 \mathbf{w}_i} + \frac{\sum_{i=f+1}^{2f} \mathbf{w}_i^T \bar{C}_1 \mathbf{w}_i}{\sum_{i=f+1}^{2f} \mathbf{w}_i^T \bar{C}_0 \mathbf{w}_i} \quad (2.12)$$

一方面， $Ratio2$  与  $Ratio1$  非常相似，它们在  $f = 1$  时是完全等价的；另一方面，我们在正交 Stiefel 流形上分别对 (2.12) 的前半部分和后半部分使用梯度下降法进行求解，以确保  $Ratio2$  解得的前  $f$  个滤波器内部，以及后  $f$  个滤波器内部都是正交的。

### 2.4.2 优化方法

本节在 Stiefel 流形 (Stiefel Manifold) 上分别对 (2.12) 的前半部分和后半部分使用梯度下降法进行求解<sup>[105]</sup>，优化过程如图2-1所示。

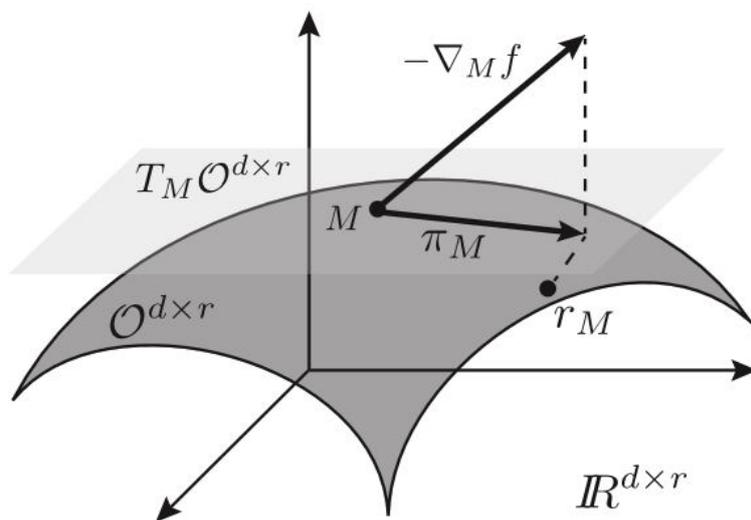


图 2-1 在 Stiefel 流形上进行梯度下降优化的示意图。

具体步骤如下：

1. 令  $t_1 = \text{tr}(W^T \bar{C}_0 W)$ ,  $t_2 = \text{tr}(W^T \bar{C}_1 W)$ ，然后需要优化下列函数

$$f_1(W) = -\frac{t_1}{t_2}, f_2(W) = -\frac{t_2}{t_1} \quad (2.13)$$

2. 在原空间  $R^{nc*2f}$  计算梯度  $\nabla_W f$ :

$$\nabla_W f_1 = -((2t_1)\bar{C}_1 W - (2t_2)\bar{C}_0 W) ./ (t_1)^2 \quad (2.14)$$

$$\nabla_W f_2 = -((2t_2)\bar{C}_0 W - (2t_1)\bar{C}_1 W) ./ (t_2)^2 \quad (2.15)$$

3. 将梯度  $\nabla_W f$  映射到切空间:

$$\pi_W(Z) = W \text{skew}(W^T Z) + (I - WW^T)Z \quad (2.16)$$

式中  $\text{skew}(W^T Z) = (M^T Z - Z^T M)/2$ 。这一步骤是寻找在流形上下降最快的方向，然后以步长  $\beta$  更新目标:  $W = W + \beta \pi_W(\nabla_W f)$ 。

4. 由于下降方向是在切空间上，优化目标此时会偏离 Stiefel 流形，所以需要把它重新映射回 Stiefel 流形:

$$r_W(Z) = \arg \min_{N \in O^{n*2r}} \|N - (M + Z)\|_F \quad (2.17)$$

5. 重复步骤2-3，直到  $f(W)$  收敛。

## 2.5 实验比较

本节在两个数据集上验证本章提出的算法的有效性，并将它的效果与现有方法对比。2.5.1节介绍本章所使用的数据集；2.5.2节比较本章所提的基于样本的协方差矩阵迁移方法（IA1，IA2）与基于用户的协方差矩阵迁移方法；2.5.3节比较上一小节提出的 CSP 滤波器与传统的 CSP 滤波器。

### 2.5.1 运动想象数据集

本章在两个运动想象数据集上进行实验，它们都是来自于“BCI Competition IV”<sup>①</sup>的公开数据集。两个数据集的实验过程相似，如图2-2所示：实验开始时，用户坐在一个舒适的椅子上，正对着一台电脑；然后需要用户根据提示多次执行不同的运动想象任务，每一个任务开始时，电脑屏幕上会出现一个正十字符号“+”，这个符号是提示用户做好准备（ $t = 0$ ）；稍后，正十字符号“+”处会出现一个箭头，箭头的方向对应着用户需要执行的运动想象任务，比如，左箭头指示用户开始想象左手运动，下箭头指示用户开始想象双脚运动，等等（ $t = 2$ ）；用户想象自己身体部位的运动，直到箭头消失（ $t = 6$ ）；然后正十字符号“+”也会消失，用户可以休息片刻，直到下一个任务开始（ $t = 8$ ）。

<sup>①</sup> <http://www.bbci.de/competition/iv/>.

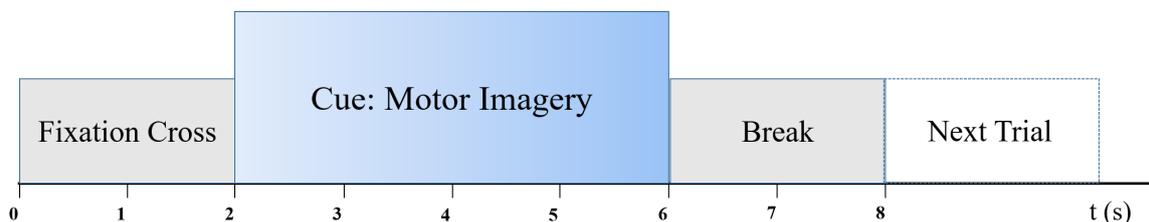


图 2-2 一个运动想象任务的时间线。

第一个数据集<sup>①</sup>称为 MI 1<sup>[106]</sup>。它包含 7 个健康用户的数据，所有数据都是由 59 个 EEG 通道以 100Hz 的采样频率采集而来。每个用户的采集都分为了 3 个阶段（3 个阶段之间有足够长时间的休息）：训练阶段，测试阶段，特殊特征阶段。由于只有训练阶段提供了完整的标签信息，所以本文只使用训练阶段的数据。每个用户的训练部分都有两个类别，这两个类别是“{左手，右手}”和“{左手，脚}”中的一组，每个用户的每个类别有 100 个样本。

第二个数据集<sup>②</sup>称为 MI 2<sup>[106]</sup>。它包含 9 个健康用户的数据，所有数据都是由 22 个 EEG 通道和 3 个眼电通道以 250Hz 的采样频率采集而来。每个用户的采集都分为了训练和测试两阶段，同样由于只有训练阶段数据提供了完整的标签信息，所以本文只采用训练阶段的数据。每个用户的训练数据都包含四个类别，分别是“左手”，“右手”，“舌”，“双脚”，每个用户的每个类别有 72 个样本。本章只取对应着“左手”和“右手”的 22 个 EEG 通道的数据。

本章采用 Matlab 中的 EEGLAB 工具箱<sup>[107]</sup>来预处理这两个数据集。首先对数据进行 [8, 30]Hz 的频域滤波，滤波器采用 Matlab 函数 *fir1*，其作用在于去除脑电中的伪迹，背景噪声等。然后把动作提示箭头出现后的 [0.5, 3.5] 秒之间（即图 2-2 中  $t = 2.5$  到  $t = 5.5$  之间）的数据作为一个 EEG 样本。对于数据集 MI 2，我们将数据降采样到 100Hz，两个数据集的相关总结如表格 2.1 所示。

表 2.1 数据集总结。

	数量				
	通道	采样点	用户	类别	样本数/类别
MI 1	59	300	7	2	100
MI 2	22	300	9	2	72

① [http://www.bbc.de/competition/iv/desc\\_1.html](http://www.bbc.de/competition/iv/desc_1.html).

② [http://www.bbc.de/competition/iv/desc\\_2a.pdf](http://www.bbc.de/competition/iv/desc_2a.pdf).

### 2.5.2 实验一：协方差矩阵迁移方法对比

本节在两个数据集上分别验证2.3节所提的基于样本的协方差矩阵迁移方法 (IA1, IA2) 的效果, 并将它们与文献中的迁移学习方法进行比较。

对每一个数据集, 首先用留一法依次选择目标用户与源用户, 即每次实验取一个用户为目标用户, 其它用户为源用户, 直到所有用户被选为一次目标用户。鉴于 SA1 需要少量的目标用户标注数据, 本节考虑一种半监督场景: 对目标用户, 先从每个类别中随机选择 20 个样本作为训练池, 剩余样本作为测试集; 然后从训练池中每次添加 4 个样本 (每个类别 2 个) 作为训练集, 与源用户样本结合训练一个模型, 用于测试集的预测, 直到训练池中的样本添加完毕。

本节比较以下 7 个算法的效果:

1. 基准1 (BL1): 直接合并所有源用户数据;
2. 基准2 (BL2): 只使用目标用户已标注数据;
3. 基准3 (BL3): 合并所有源用户数据与目标用户已标注数据;
4. SA1: 基于用户的协方差矩阵迁移方法, 如2.2.1节所描述;
5. SA2: 基于用户的协方差矩阵迁移方法, 如2.2.2节所描述;
6. IA1: 本章提出的基于样本的协方差矩阵迁移方法, 并使用通道拼接的方法得到特征向量进行优化, 如2.3节所描述。另外为了降低数据维度, 从而提升优化效率, 我们对每个样本取前两秒的数据;
7. IA2: 本章提出的基于样本的协方差矩阵迁移方法, 并使用样本的协方差矩阵的切向量进行优化, 如2.3节所描述。

上述所有算法采用相同的训练集和测试集, 亦采用相同的分类模型, 即首先使用 CSP 进行空域滤波并提取信号方差的对数作为特征, 然后用线性判别分析 (LDA) 作为分类器。图2-3和图2-4分别是在数据集 MI 1 和 MI 2 上的测试结果, 其中图2-3的前 7 个子图和图2-4的前 9 个子图代表使用单个用户作为目标用户的实验结果, 最后一个子图是所有目标用户的平均结果。在所有子图中, 横轴代表目标用户每个类别的已标注的样本数目  $k$ , 纵轴代表分类正确率。由图中可以看出:

1. 当目标用户已标注的样本数目  $k$  较小时, BL2 在数据集 MI 1 上的效果仅优于 BL1, 在数据集 MI 2 上的效果最差。这说明当目标用户只有少量的已标注样本时, 利用源用户的数据是很有必要的, 即使是最简单的用法, 即直接将所有源用户数据与目标用户已标注数据合并使用也能取得更好的效果。
2. 随着  $k$  增大, BL2 的效果上升最明显, 会渐渐接近甚至超过其它迁移学习方法。这说明当目标用户有足够的已标注样本时, 无需进行迁移学习, 直接使用当前用户的数据即可。
3. 当  $k$  较小时, 基于用户迁移的方法 (SA1 和 SA2) 的平均效果在两个数据集上都优于简单平均的方法 (BL3), 说明 SA1 和 SA2 对源用户和目标用户之间的相似度的衡量是有意义的。

4. 当  $k$  较小时, 本文提出方法 (IA1 和 IA2) 的平均效果在两个数据集上都优于其它方法, 验证了本文所提方法的有效性。若以正确率曲线以下的面积 ( $k \leq 10$ ) 来衡量对应算法在不同目标域带标注样本数目时的性能, 则 IA1 相对于 BL3 在两个数据集上的平均性能提升了 9.34%, 相对于 SA1 和 SA2 分别提升了 3.9% 和 3.12%。
5. 平均来看, IA1 的效果比 IA2 更好, 所以建议使用 IA 时直接将 EEG 样本按通道拼接即可。

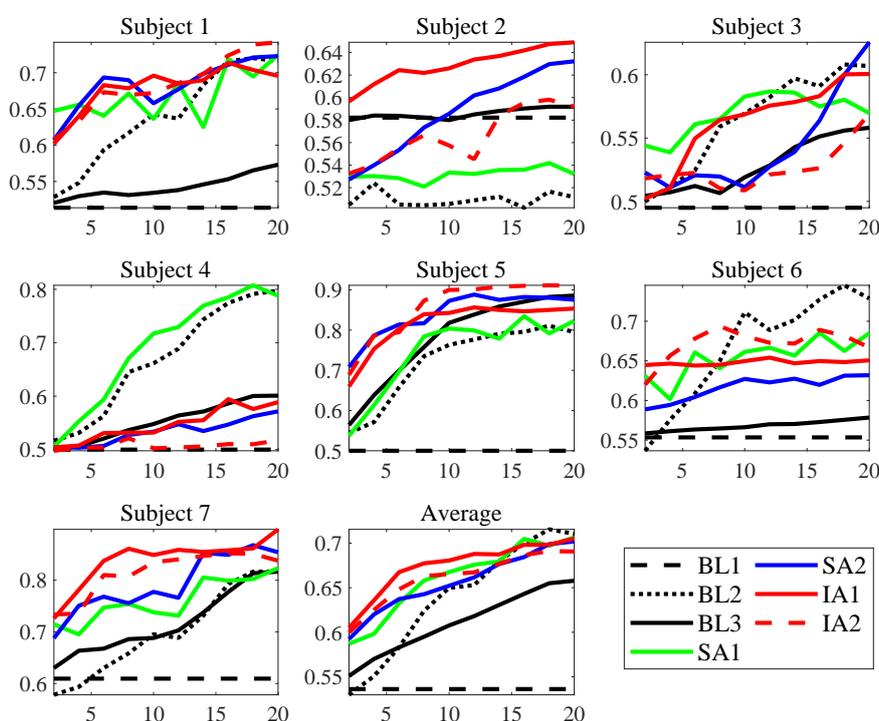


图 2-3 7 种不同 CSP 在数据集 MI 1 上的分类正确率。

### 2.5.3 CSP 的目标函数比较

本节在两个运动想象数据集上测试 2.4 节所提的新 CSP 目标函数 (用 SM 代称, 以与传统 CSP 区分), 并将之与传统 CSP 比较。本节实验不考虑迁移学习场景, 对每一个用户, 随机选取 50% 样本作为训练集, 另外 50% 样本作为测试集, 然后重复实验 30 次以获得在统计上有意义的结果。

使用 SM 的一个出发点是为了得到相关性更小的滤波器, 所以本节首先比较 SM 与 CSP 的滤波器之间的平均相关系数, 表格 2.2 和 2.3 分别展示了它们在数据集 MI 1 和 MI 2 上的结果, 可以观察到 SM 计算的滤波器之间的平均相关系数更小。

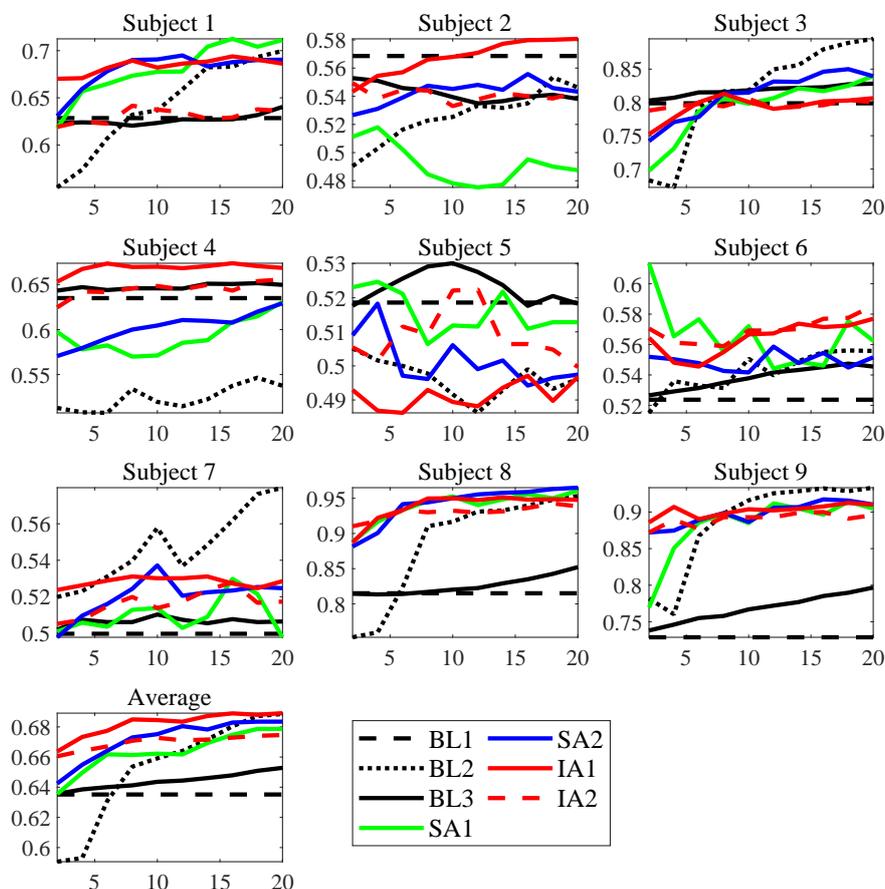


图 2-4 7 种不同 CSP 在数据集 MI 2 上的分类正确率。

表 2.2 单个滤波器之间的平均相关系数：数据集 MI 1。

Subject	1	2	3	4	5	6	7	平均
CSP	.1393	.0760	.1064	.0797	.0839	.1122	.1148	.1018
SM	.0907	.0477	.0399	.1373	.0738	.0943	.0886	.0817

然后比较 SM 和 CSP 在两个数据集上得到的实际  $Ratio1$  和  $Ratio2$ ，除了普通的 CSP 之外，我们还考虑正则化 CSP (regularized CSP, RCSP)。由式 (2.1) 和 (2.2) 可知，普通 CSP 的解是  $\bar{C}_1^{-1}\bar{C}_0$  (或者  $\bar{C}_0^{-1}\bar{C}_1$ ) 的最大  $f$  个以及最小  $f$  个特征值对应的特征向量。对于 RCSP，特征分解的对象变为  $(\bar{C}_1 + \lambda I)^{-1}\bar{C}_0$ ，其中  $I$  为单位矩阵， $\lambda$  为正则化项的调节因子，通过在训练数据上交叉验证得到，该方法已在文献<sup>[57]</sup>中表现出了比普通 CSP 更好的效果。

表 2.3 单个滤波器之间的平均相关系数：数据集 MI 2。

Subject	1	2	3	4	5	6	7	8	9	平均
CSP	.1807	.3224	.2571	.1776	.1416	.2116	.2419	.1904	.1665	.2100
SM	.1744	.1759	.1522	.1488	.1280	.1481	.1354	.1288	.1521	.1493

类似地，我们也对 SM 进行正则化（regularized SM, RSM），RSM 的前  $f$  个滤波器的目标函数变为：

$$\max \frac{\sum_{i=1}^f \mathbf{w}_i^T \bar{C}_0 \mathbf{w}_i}{\sum_{i=1}^f \mathbf{w}_i^T (\bar{C}_1 + \lambda I) \mathbf{w}_i} \quad (2.18)$$

后  $f$  个滤波器的目标函数变为：

$$\max \frac{\sum_{i=f+1}^{2f} \mathbf{w}_i^T \bar{C}_1 \mathbf{w}_i}{\sum_{i=f+1}^{2f} \mathbf{w}_i^T (\bar{C}_0 + \lambda I) \mathbf{w}_i} \quad (2.19)$$

其中  $I$  同样为正则化项， $\lambda$  为正则化项的调节因子，通过在训练数据上交叉验证得到。

图2-5展示了四种滤波方法（CSP, SM, RCSP, RSM）在滤波器个数分别为 4, 6, 8 时，*Ratio1* 和 *Ratio2* 在数据集 MI 1 上的结果比较。图中横轴数字代表用户索引，最后一组柱状图是在所有用户上的平均结果。图2-6则展示的是在数据集 MI 2 上的结果。可以观察到 CSP 总是比 SM 有更高的 *Ratio1*，而 SM 则有更高的 *Ratio2*，这点很容易理解，因为 CSP 和 SM 是分别以 *Ratio1* 和 *Ratio2* 为目标函数。

最后研究 SM 和 CSP 对分类正确率的影响，对上面提到的每一种滤波器，我们都考虑两种分类模型：

1. 提取方差对数作为特征，如式 (1.6) 所示，然后使用 LDA 分类；
2. 计算所有样本的协方差矩阵，然后使用 MDRM 分类，如1.3.2节所描述。

结合四种滤波器与两种分类模型，我们一共比较 8 种算法。图2-7展示了当滤波器个数分别为 4, 6, 8 时，所有算法在数据集 MI 1 上的分类正确率。图2-8则展示在数据集 MI 2 上的分类正确率。图中横轴数字代表用户索引，最后一组柱状图是在所有用户上的平均结果。可以观察到：

1. 总体来看，RCSP 比 CSP 更好，RSM 比 SM 更好，这不仅验证了文献<sup>[57]</sup>的结果，也证明了正则化对 SM 的正面效果；
2. SM 只在少数用户上的效果比 CSP 更好，而平均效果 SM 是稍差于 CSP 的。

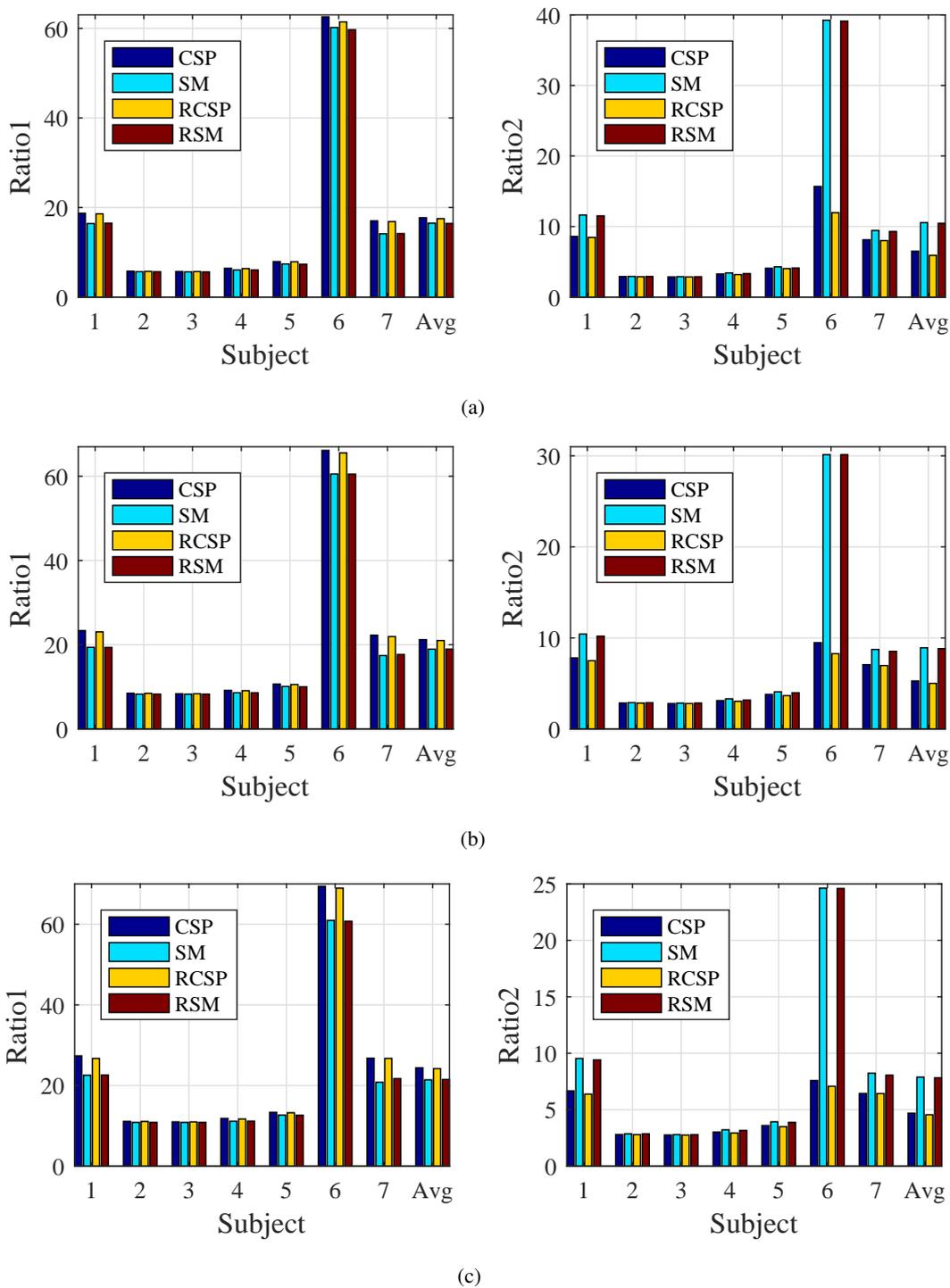


图 2-5 数据集 MI 1 上 *Ratio1* 和 *Ratio2* 的比较: a)  $f = 2$ ; b)  $f = 3$ ; c)  $f = 4$ .

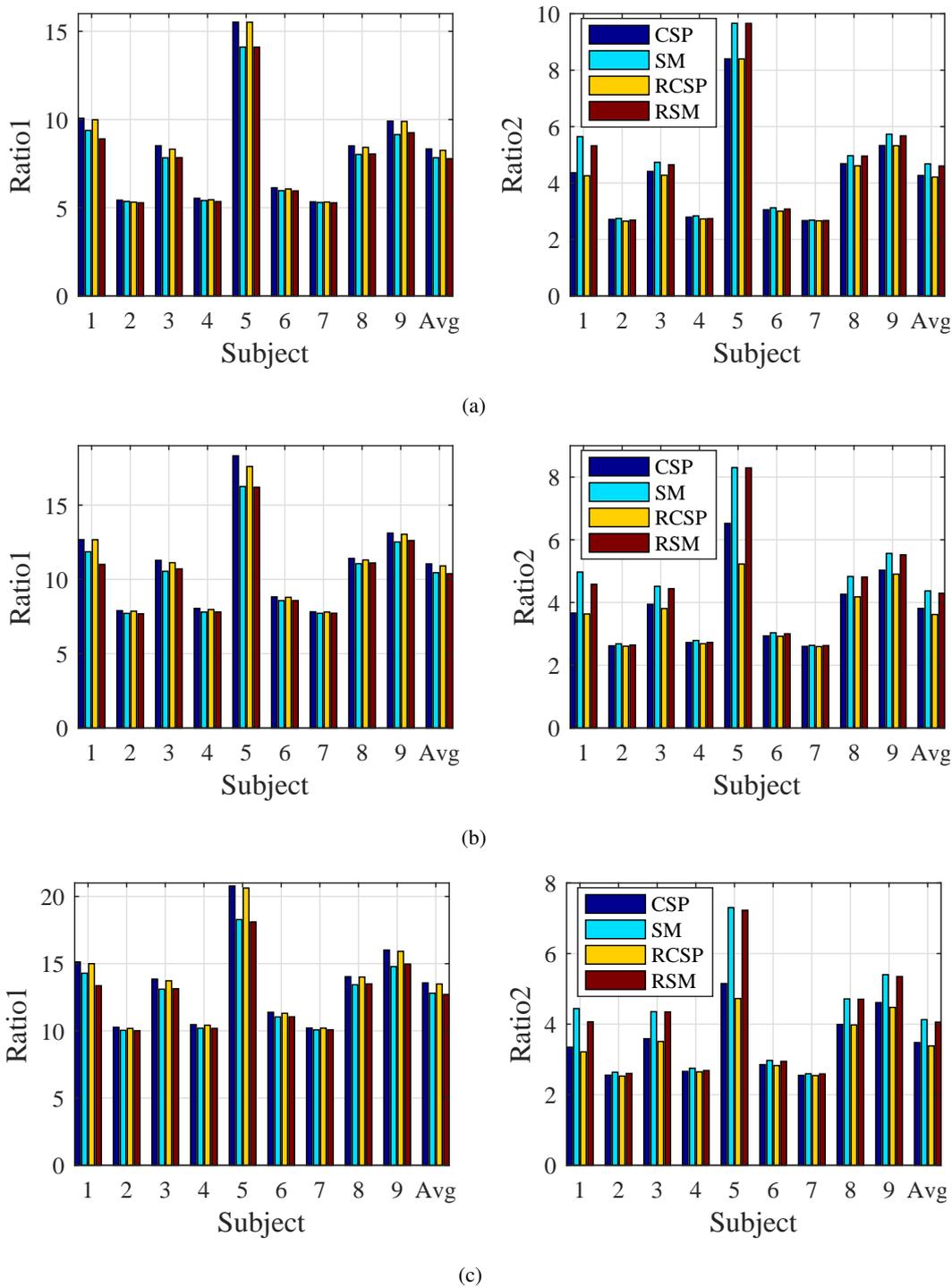
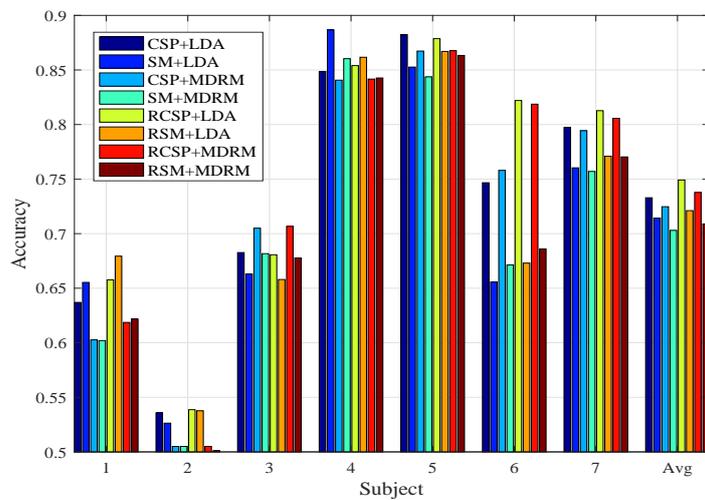
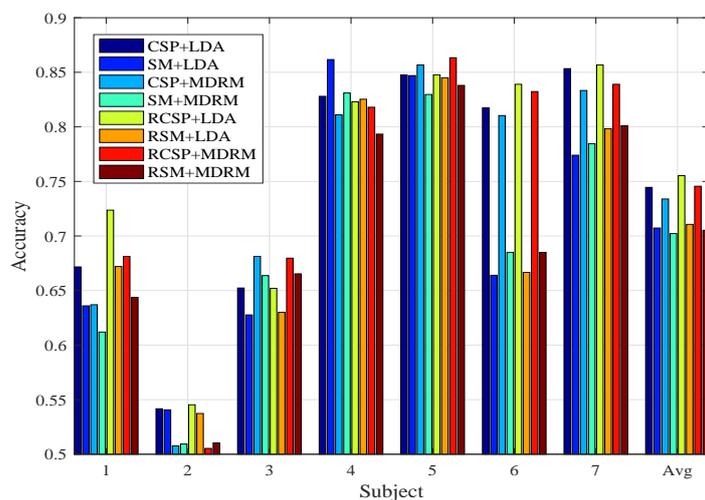


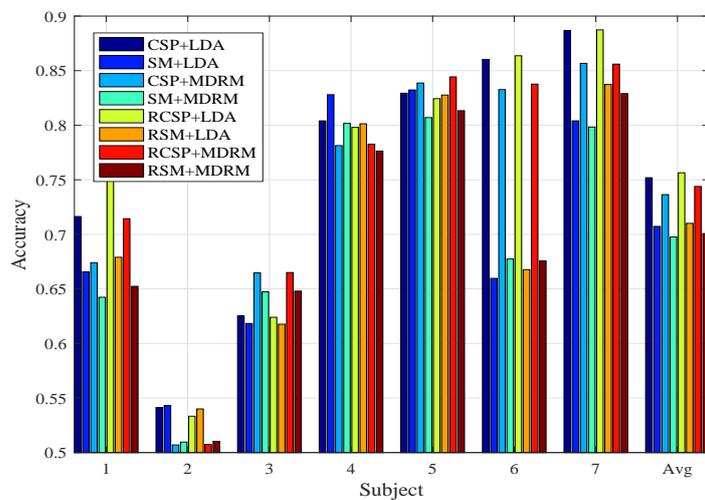
图 2-6 数据集 MI 2 上 *Ratio1* 和 *Ratio2* 的比较: a)  $f = 2$ ; b)  $f = 3$ ; c)  $f = 4$ 。



(a)

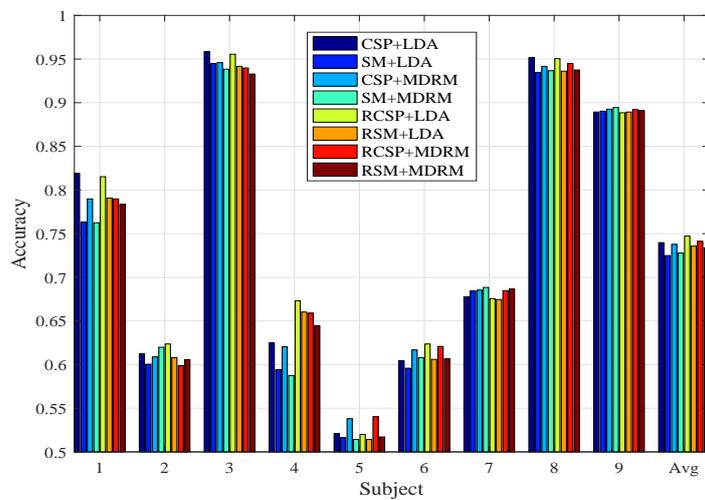


(b)

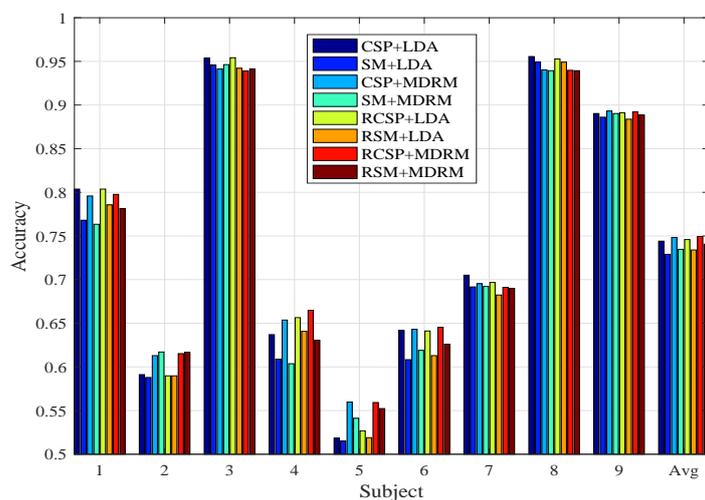


(c)

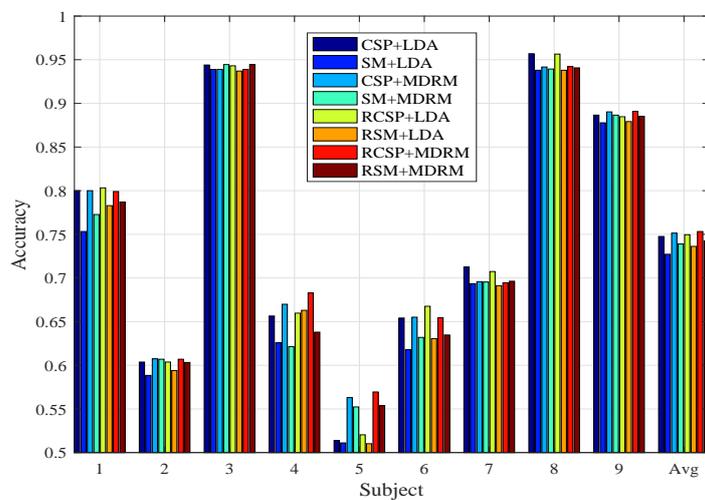
图 2-7 数据集 MI 1 上的分类正确率比较: a)  $f = 2$ ; b)  $f = 3$ ; c)  $f = 4$ 。



(a)



(b)



(c)

图 2-8 数据集 MI 2 上的分类正确率比较: a)  $f = 2$ ; b)  $f = 3$ ; c)  $f = 4$ 。

## 2.6 小结

本章对脑机接口中最重要的空域滤波方法，即共同空间模式（CSP）展开研究，其一是研究当新用户缺乏标注数据时，如何将迁移学习应用于 CSP；其二是在当前用户不缺乏标注数据时，讨论 CSP 的目标函数的特性以及可能存在的局限。

对于第一方面，本章首先总结了将迁移学习应用于 CSP 的基本思路，即使用源用户的协方差矩阵来帮助衡量目标用户的类别平均协方差矩阵；然后介绍了两种文献中的迁移学习方法，根据迁移特点把它们总结为“基于用户的协方差矩阵迁移”；接着本章提出一种新的迁移思路，称为“基于样本的协方差矩阵迁移”，并给出了两种具体使用方法。实验证明，本章提出的两种方法都是有效的，且效果比文献现有方法都更好。

对于第二方面，本章先研究了传统 CSP 的解，指出其可能存在的问题，即多个滤波器得到的特征之间可能存在冗余信息；然后针对这个问题，本章提出一种新的相关的 CSP 目标函数，并在正交空间使用梯度下降法进行求解。最后用实验比较两种不同的 CSP 目标函数，以及它们正则化之后的解。实验发现，尽管新目标函数的滤波器之间有更低的平均相关系数，但其分类效果总体略差于传统 CSP 的目标函数；另外，无论是哪种目标函数，正则化之后都能得到更好的分类效果。

### 3 黎曼空间中迁移学习的优化方法

#### 3.1 引言

上一章的研究对象是脑机接口中的传统分类框架，本章则针对基于黎曼几何的分类框架开展研究。

法国的 Barachant 等研究者<sup>[59]</sup>于 2012 年首次将黎曼几何方法引入到脑机接口领域，提出了一种新的、针对 EEG 样本协方差矩阵的分类框架，并依靠此方法获得数个国际脑信号分类比赛的冠军，比如“DecMEG2014”(<https://www.kaggle.com/c/decoding-the-human-brain>)和“BCI challenge 2015”(<http://neuro.embs.org/2015/bci-challenge/>)。此后，黎曼几何方法在脑机接口领域获得了很多的关注，研究者在文献<sup>[59]</sup>的基础上进行完善，逐渐将黎曼几何方法的应用场景从最开始的运动想象范式推广到了 P300 诱发电位以及疲劳度监测等范式，并且取得了良好的效果<sup>[60-62]</sup>。

随着黎曼几何方法应用的扩展，脑机接口中的迁移学习研究也从传统方法发展到了黎曼几何方法领域<sup>[98,108-112]</sup>。最近，Zanini<sup>[112]</sup>等研究者提出一种黎曼空间数据对齐方法 (Riemannian space Alignment, RA)，主要思想是在黎曼空间中把不同用户的样本协方差矩阵进行对齐，是脑机接口领域当前最先进的迁移学习方法之一。

进一步地，本章提出两种优化 RA 的思路：其一是计算优化，本章采用两种不同的方法来挑选通道，从而降低样本协方差矩阵的维度；其二是进行训练数据重采样，即基于黎曼距离挑选一小部分源用户样本作为训练集。内容安排如下：3.2 节介绍针对样本协方差矩阵的迁移学习方法 RA；3.3 节提出用挑选通道的方法来优化黎曼空间中的计算；3.4 节提出一种基于黎曼距离挑选源域样本的方法，从而在 RA 的基础上进一步减少源用户与目标用户之间的分布差异；3.5 节用实验验证本章所提方法对于 RA 效果提升的有效性；3.6 节总结本章内容；

#### 3.2 黎曼空间数据对齐方法

在脑机接口中，如 1.3.2 节所介绍的，黎曼几何方法是以样本的协方差矩阵为特征。因此，在黎曼几何框架下的迁移学习也需要在样本的协方差矩阵上进行迁移。Zanini 等研究者<sup>[112]</sup>于 2018 年特别针对 EEG 的样本协方差矩阵提出了一种黎曼空间数据对齐方法 RA，该方法在黎曼空间中对来自不同用户的协方差矩阵进行对齐，再使用对齐后的源用户协方差矩阵训练 MDRM 分类器，然后应用在对齐后的目标用户上。

在黎曼几何的框架下，如果把每个协方差矩阵看作黎曼流形上的一个点，则协方差矩阵的变换也可以看作“点”在黎曼流形上的移动。当在黎曼流形上观察来自

不同用户的协方差矩阵时，我们不能判定两点之间的偏移是由不同任务引起的，还是由个体差异引起的。这也是为什么不能直接在源用户上训练模型，而是需要先用迁移学习来减少个体差异的原因。

那么这种个体差异是怎么引起的呢？原因可能有很多种，比如不同用户之间的头部结构差异、电极在头皮上放置的位置差异，以及用户对同一刺激产生的神经反应差异等等。RA 主要考虑前两种因素，因为它们在整个实验过程中是固定存在的，不因大脑状态或所执行任务不同而变化。RA 就是基于这个特点，利用用户在休息时段的 EEG 信号来减少执行任务时段的个体差异。休息时段的 EEG 信号在其它算法中通常是被抛弃的，比如在运动想象实验中，每次执行完一次运动想象任务后，通常会休息 2 秒左右的时间，而这段休息时间不会被剪辑到该任务的 EEG 样本中。

对每一个用户，RA 首先截取该用户在休息时段的 EEG 信号，称为“参考状态”。头部结构和电极位置的差异会引起参考状态协方差矩阵的偏移，并且当大脑在执行一个指定任务时，协方差矩阵会在黎曼流形上从参考状态向特定的方向移动。因此，针对每一个用户，若有多个休息时段，RA 计算这些参考状态的协方差矩阵的黎曼中心，称为“参考矩阵”；然后整体移动所有样本协方差矩阵，使它们的中心点固定在参考矩阵。这样，我们在不同用户的样本协方差矩阵上观察到的偏移就大多是由任务不同引起的。

具体地，针对每一个用户，RA 首先获取该用户在休息时段的 EEG 信号，并计算它们的协方差矩阵，记为  $\{R_i\}_{i=1}^k$ ，然后计算这些协方差矩阵的黎曼中心作为参考矩阵：

$$\bar{R} = \arg \min_P \sum_{n=1}^k \delta^2(P, R_n). \quad (3.1)$$

假设该用户的样本协方差矩阵为  $\{C_i\}_{i=1}^N$ ，RA 对每一个样本协方差矩阵执行如下变换：

$$\tilde{C}_i = \bar{R}^{-1/2} C_i \bar{R}^{-1/2}, \quad (3.2)$$

这里  $\tilde{C}_i$  是对齐后的协方差矩阵。

这里回顾一个重要的、关于黎曼距离的仿射不变性（congruence invariance）定理：

**定理 3.1:** 假设  $P_1$  和  $P_2$  为两个正定对称矩阵， $W$  为一个可逆矩阵， $\delta(\cdot)$  为两个正定对称矩阵之间的黎曼距离，则：

$$\delta(W^T P_1 W, W^T P_2 W) = \delta(P_1, P_2), \quad (3.3)$$

在用户内部，样本协方差矩阵使用的转换矩阵  $\bar{R}$  是一样的，意味着在执行式 (3.2) 代表的 RA 变换后，用户内部样本协方差矩阵之间的黎曼距离是不变的，RA

变换相当于对该用户的样本协方差矩阵作了某种整体移动，这确保  $\mathbf{RA}$  不会破坏用户原本的模式。而对不同用户，它们使用了不同的转换矩阵  $\bar{R}$ ，意味着  $\mathbf{RA}$  拥有改变不同用户协方差矩阵分布的能力。

### 3.3 通道选择

$\mathbf{RA}$  最核心的一步是计算参考矩阵，即参考状态协方差矩阵的黎曼中心。然而，当数据的通道数  $n_c$  较大时，协方差矩阵的维度  $n_c \times n_c$  也会很大，此时关于协方差矩阵的计算效率会大大降低，准确度也会受到影响。特别是在计算多个协方差矩阵的黎曼中心时，可能多次迭代都无法收敛。尽管有一些基于图的方法能对黎曼空间进行降维<sup>[113]</sup>，但这些降维方法通常本身就不存在解析解，而是需要通过迭代优化来得到低维黎曼空间，且最后对分类结果的影响难以预测。

因此，本节采用简单的通道选择来减少协方差矩阵的维度，从而提高黎曼空间中计算的效率和准确度。事实上，文献中已有一些方法挑选通道，但在传统的机器学习算法中，这些方法并不总是有效的。具体地，本节介绍两种挑选通道的方法：

1. 人工挑选：EEG 数据的不同通道对应着头皮表面不同位置的电极。显然，挑选出距离当前任务相关脑区的电极会有利于数据的质量。虽然目前存在着不同规格的脑电帽，但其电极的位置大都是基于 10-20 国际标准导联系统，如图3-1所示。因此，人工挑选方法就是依据特定任务的神经活动区域和 10-20 国际标准导联系统，挑选出若干通道。
2. 基于 CSP 滤波器的挑选：每个 CSP 滤波器事实上是以不同的权重将所有通道进行一种重组，这里的权重其实也反映了各个通道对信号可分的重要性，所以也有文献基于 CSP 滤波器来选择通道<sup>[114]</sup>。具体地，假设  $W \in \mathbb{R}^{2f \times n_c}$  是已经求得的 CSP 映射矩阵，而  $w_i$  是  $W$  的第  $i$  行，对应着第  $i$  个通道。则第  $i$  个通道的得分为：

$$SC(i) = \frac{\|w_i\|_1}{\|W\|_1} \quad (3.4)$$

其中  $\|\cdot\|$  代表  $l_1$  范数。

然后，我们再将通道按照  $SC(i)$  值从大到小进行排列，再选择前任意个通道即可完成挑选。

### 3.4 源用户样本选择

协变量偏移，即源域与目标域的输入之间的分布差异，也是现实场景中最常见的分布差异。多数迁移学习方法会首先试图减少协变量偏移，对源域样本进行重采样或者赋权是常用的方式<sup>[67-69]</sup>。



**Algorithm 2:** 样本选择算法。

**Input:**  $N^t$  个未标注的目标用户样本,  $\{\mathbf{X}_i^t\}_{i=1}^{N^t}$ ;  
 $N^s$  个已标注的源用户样本,  $\{\mathbf{X}_j^s\}_{j=1}^{N^s}$ ;  
 $k$ , 对每一个目标用户样本执行聚类的邻居数目;

**Output:** 被选择的源用户样本

**for**  $i = 1, \dots, N^t$  **do**

**for**  $j = 1, \dots, N^s$  **do**

根据式 (1.1) 计算  $\mathbf{X}_i^t$  与  $\mathbf{X}_j^s$  的协方差矩阵, 以及根据式 (1.7) 计算协方差矩阵之间的黎曼距离;

**end**

对目标用户样本  $\mathbf{X}_i^t$ , 从所有源用户样本  $\{\mathbf{X}_j^s\}_{j=1}^{N^s}$  中挑选出  $k$  个黎曼距离最近的样本;

**end**

结合所有被选择的源用户样本并去掉其中被重复选择的;

**Return** 去重之后的源域样本。

表 3.1 数据集总结。

	数目					
	通道	样本采样点	参考状态采样点	用户	类别	样本数/类别
MI 1	59	300	100	7	2	100
MI 2	22	750	250	9	2	72

本章考虑无监督迁移学习场景: 首先用留一法依次选择每个用户为目标用户, 其它用户为源用户, 其中目标用户的 EEG 样本可以使用, 但不知道它们的标签, 源用户都是已标注数据。

为强调不同方法的效果对比, 本章将要用到的算法分为三个阶段:

1. **预处理阶段:** 主要包括对 EEG 数据进行时域滤波, 并且截取正在执行运动想象的时段作为样本, 如2.5 节所示。
2. **中间阶段:** 选择性地执行3.3节所描述的通道选择方法, 3.2节所描述的迁移学习方法 RA, 或者3.4 所描述的源用户样本选择方法。
3. **分类阶段:** 对处理后的数据应用分类模型, 获得测试效果。

对所有算法, 本章采用相同的预处理阶段和分类阶段, 然后采用不同的中间阶段, 以突显本章所提算法的影响。并且, 每次比较都会在分类阶段采用两种分类流程, 以验证算法在不同分类流程下的效果:

1. **MDRM:** 直接使用 MDRM 分类器进行分类;

2. CM: 先使用 CSP 对数据进行空域滤波, 再使用 MDRM 分类。

### 3.5.1 通道选择对RA的影响

本节针对3.3节所提方法开展实验, 即研究通道选择能否改善 RA 的计算效率与精确度, 从而提升算法整体的迁移效果。鉴于数据集 MI 1 有较多通道, 而 MI 2 通道数相对较少, 本节只在数据集 MI 1 上验证通道选择方法的效果。具体地, 本节在算法的中间阶段分别测试以下两种通道选择方法:

1. CS1 (channel selection 1): 基于脑电图中的 10-20 国际导联标准, 人工选择以下通道: {‘Fz’, ‘FC3’, ‘FC1’, ‘FCz’, ‘FC2’, ‘FC4’, ‘C5’, ‘C3’, ‘C1’, ‘Cz’, ‘C2’, ‘C4’, ‘C6’, ‘CP3’, ‘CP1’, ‘CPz’, ‘CP2’, ‘CP4’, ‘P1’, ‘Pz’, ‘P2’, ‘PO1’}。这些通道常被认为与“左手运动想象”和“右手运动想象”相关大脑活动区域的位置接近。
2. CS2 (channel selection 2): 基于 CSP 滤波器选择通道, 如3.3 节所描述, 并且将被选择的通道数目分别取为 {15, 20, 25, 30, 35}, 然后观察 CS2 在这些不同参数设置下的效果。

首先观察 CS1 对 RA 的影响, 无论采用分类流程 MDRM 或是 CM, 我们都比较以下四种算法:

1. MDRM/CM: 在中间阶段既不选择通道, 也不进行对齐;
2. CS1: 在中间阶段使用 CS1 挑选通道, 但不进行对齐;
3. RA: 在中间阶段使用所有通道, 进行 RA 对齐;
4. CS1-RA: 在中间阶段先使用通道选择方法 CS1 挑选通道, 再进行 RA 对齐。

实验结果结果如表3.2所示。可以观察到:

1. MDRM 的效果优于 CS1-MDRM, 且 CM 的效果优于 CS1-CM, 这说明直接使用 CS1 并不能直接提高分类模型的性能。
2. CS1-RA-MDRM 的平均效果优于 RA-MDRM, 且 CS1-RA-CM 的平均效果优于 RA-CM, 这说明在使用迁移学习算法 RA 后, CS1 能够提高模型的性能。与上一条的结果对比, 说明 CS1 有利于提升 RA 的迁移效果, 这验证了我们在3.3节的理论。

然后比较在使用 RA 的条件下, 经过 CS1 选择通道后的算法整体用时与使用全部通道的算法整体用时。实验平台为 Thinkpad 系列笔记本, 装载 Intel Core i5-6200U CPU@2.30GHz, 4GB 内存, 190GB 固态硬盘, 64位 Windows 10 以及 Matlab 2018b。实验结果如表格3.3 所示, 从平均用时看, CS1-RA-MDRM 比 RA-MDRM 快约 32 倍, CS1-RA-CM 比 RA-CM 快约 30.7 倍, 且经 CS1 选择通道后, 在不同用户上耗时的标准差也小了很多。这说明 CS1 极大地提升了 RA 的计算效率。

表 3.2 通道选择方法 CS1 在数据集 MI 1 上的结果比较。

用户	MDRM				CM			
	MDRM	CS1	RA	CS1-RA	CM	CS1	RA	CS1-RA
1	51.00	51.00	<b>77.50</b>	73.50	49.50	51.50	55.00	73.50
2	50.00	50.00	67.00	61.00	54.00	50.00	<b>68.00</b>	62.00
3	50.00	50.00	62.00	61.00	50.00	50.00	<b>68.00</b>	60.00
4	51.50	50.00	54.00	59.00	52.50	50.00	60.00	<b>69.50</b>
5	50.00	50.00	72.00	79.00	50.00	50.00	70.50	<b>91.00</b>
6	77.00	60.00	<b>80.00</b>	79.00	75.50	54.50	60.00	<b>80.00</b>
7	50.00	50.00	72.00	79.00	50.00	50.00	67.00	<b>81.50</b>
平均	54.21	51.57	69.21	70.21	54.50	50.86	64.07	<b>73.93</b>

表 3.3 算法耗时（秒）

	RA-MDRM	CS1-RA-MDRM	RA-CM	CS1-RA-CM
平均	132.6188	4.1392	122.3769	3.9822
标准差	1.4100	0.0134	0.0908	0.0057

接下来观察 CS2 对于 RA 的效果，我们比较同样的四种算法，只是将通道选择算法 CS1 替换为 CS2。由于本节重点是研究 CS2 对 RA 迁移效果的影响，这里只比较 RA-MDRM 与 CS2-RA-MDRM，以及 RA-CM 与 CS2-RA-CM，结果分别如表格 3.4 和 3.5 所示。可以观察到，无论是以 MDRM 分类还是 CM 分类，CS2 在一个很宽泛的参数设置下（被选择通道数目从 15 变化到 35），都可以提升 RA 的迁移效果。

表 3.4 以 MDRM 为分类模型，通道选择方法 CS2 在数据集 MI 1 上的结果比较。

用户	RA-MDRM	CS2-RA-MDRM				
	59 (All)	15	20	25	30	35
1	77.50	<b>79.00</b>	77.50	78.50	77.50	74.50
2	<b>67.00</b>	61.00	65.50	62.00	60.50	62.00
3	62.00	62.00	61.00	64.00	<b>68.00</b>	62.50
4	54.00	60.00	61.00	60.50	<b>62.50</b>	61.00
5	72.00	<b>86.50</b>	86.00	<b>86.50</b>	83.5	84.00
6	<b>80.00</b>	77.50	78.50	77.50	77.00	78.50
7	72.00	78.50	<b>83.50</b>	79.50	80.00	80.5
平均	69.21	72.07	<b>73.29</b>	72.64	72.71	71.86

表 3.5 以 CM 为分类模型，通道选择方法 CS2 在数据集 MI 1 上的结果比较。

用户	RA-CM	CS2-RA-CM				
	59 (All)	15	20	25	30	35
1	55.00	74.50	71.50	73.50	72.50	<b>75.00</b>
2	68.00	69.50	<b>77.00</b>	73.50	69.00	71.50
3	68.00	60.00	69.00	68.50	72.5	<b>75.00</b>
4	60.00	60.00	<b>67.50</b>	<b>67.50</b>	66.50	66.50
5	70.50	<b>95.50</b>	<b>95.50</b>	95.00	93.5	94.50
6	60.00	<b>83.00</b>	80.00	78.00	78.00	79.00
7	67.00	86.50	<b>88.00</b>	84.50	85.00	84.50
平均	64.07	76.43	<b>78.36</b>	77.21	76.71	78.00

### 3.5.2 样本选择算法的实验结果

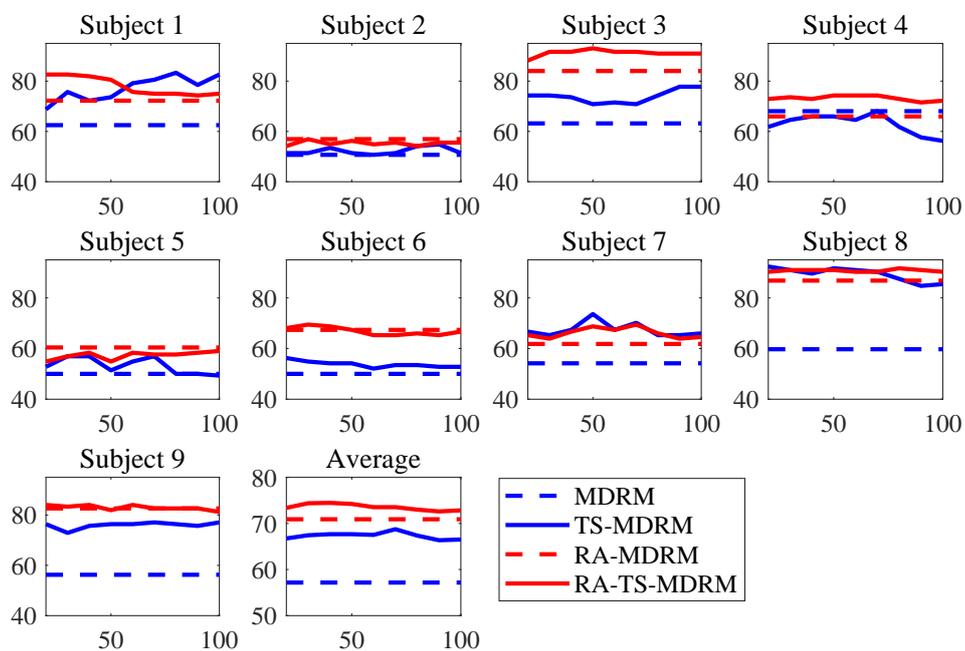
本节在数据集 MI 2 上验证样本选择算法对 RA 迁移效果的影响。具体地，本节分别在两种分类模型下比较以下四种算法：

1. **MDRM/CM**: 在中间阶段既不进行协方差矩阵对齐，也不挑选源用户样本；
2. **TS**: 在中间阶段使用3.4节所提方法挑选源用户样本，但不进行对齐；
3. **RA**: 在中间阶段使用所有源用户样本，进行 RA 对齐；
4. **RA-TS**: 在中间阶段先用 RA 对齐，再使用3.4节所提方法挑选源用户样本。

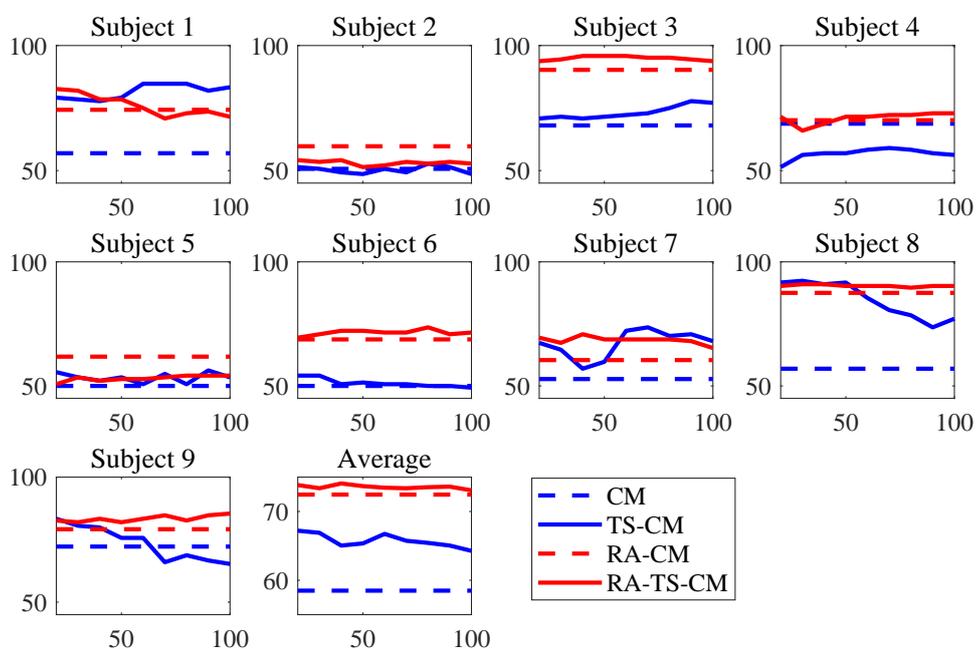
TS 有一个参数，即算法2中执行 KNN 聚类的邻居个数  $k$ ，为了观察 TS 在不同参数  $k$  下的性能，本节从  $k = 20$  开始，每次增加 10，直到  $k = 100$ 。实验结果如图3-2所示，其中前 9 个子图代表使用单个用户作为目标用户的结果，最后一个子图是平均结果。在所有子图中，横轴代表目标用户  $k$  的取值，纵轴代表分类正确率。

由图中可以看出：

1. 从平均效果看，TS-MDRM 的分类效果优于 MDRM，且 TS-CM 的分类效果优于 CM。这说明无论分类模型是 MDRM 还是 CM，对源用户数据进行采样能够减少协变量偏移。
2. 从平均效果看，RA-TS-MDRM 的分类效果优于 RA-MDRM，且 RA-TS-CM 的分类效果优于 CM。这说明无论分类模型是 MDRM 还是 CM，TS 都能进一步加强 RA 的迁移效果。
3. 从效果增幅看，TS 在 RA 之后的效果增幅不如 TS 直接在未对齐数据上的效果增幅，这可能是因为 RA 已经大大减少了分布差异。



(a)



(b)

图 3-2 样本选择算法 TS 在数据集 MI 2 上的实验结果：a) MDRM 分类；b) CM 分类。

### 3.5.3 综合使用通道选择与样本选择算法

前两节分别验证了通道选择和源用户样本选择的有效性，本节则观察同时使用这两种优化方法时的效果。具体地，本节比较以下三种算法：

1. CS1-RA: 在中间阶段先执行 CS1 挑选通道，再执行 RA 对齐；
2. RA-TS: 在中间阶段先执行 RA 对齐，再使用 TS 挑选源用户样本；
3. CS1-RA-TS: 在中间阶段先执行 CS1 挑选通道，再执行 RA 对齐，最后使用 TS 挑选源用户样本；

图3-3与图3-4表示在数据集 MI 1 上的实验结果，其中前 7 个子图代表在单个目标用户上的实验结果，最后一个子图是所有目标用户的平均结果。在所有子图中，横轴代表目标用户  $k$  的取值，纵轴代表分类正确率。可以观察到 CS1-RA-TS-MDRM 的分类效果优于 CS1-RA-MDRM 和 RA-TS-MDRM，CS1-RA-TS-CM 的分类效果优于 CS1-RA-CM 和 RA-TS-CM。这说明无论采用哪种分类模型，CS1 和 TS 可以互相兼容，进一步地提升 RA 的迁移效果。

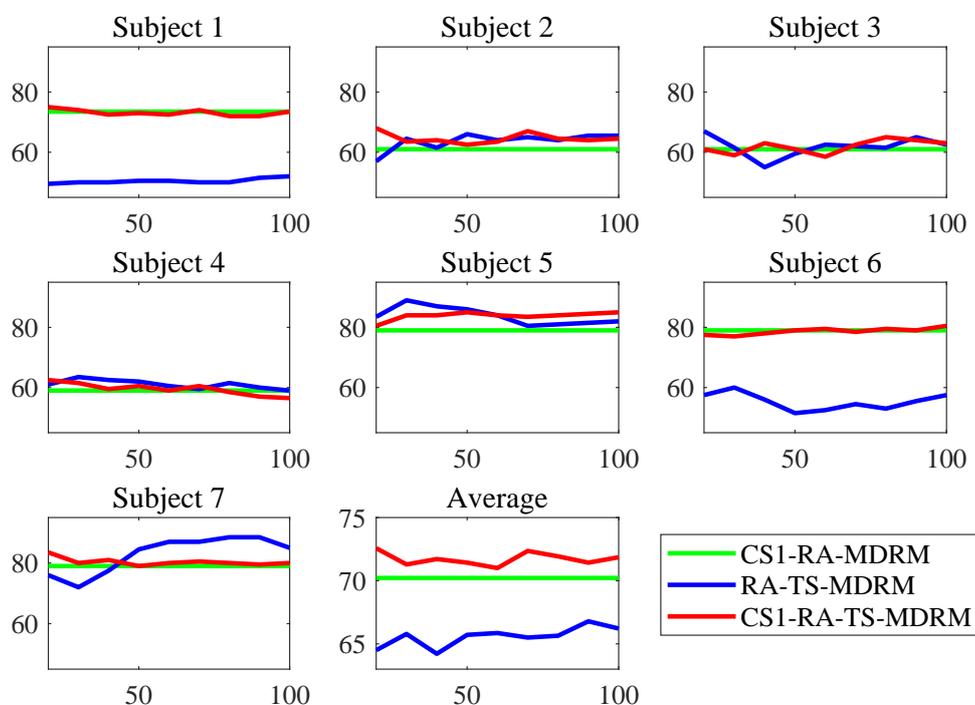


图 3-3 以 MDRM 为分类器，综合使用 CS1 与 TS 在数据集 MI 1 上的实验结果。

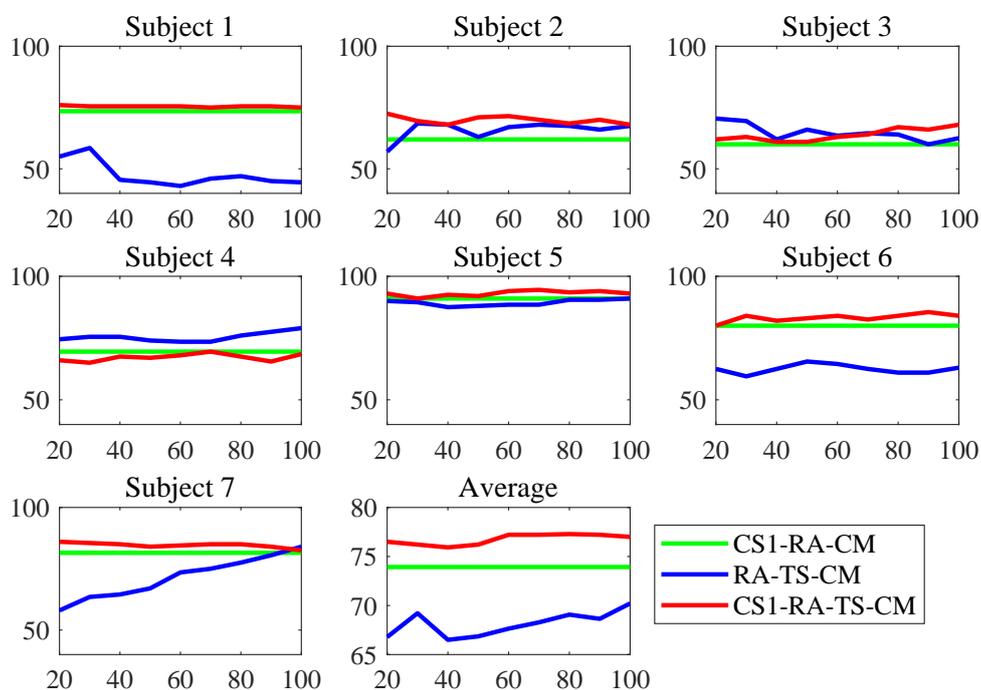


图 3-4 以 CM 为分类器，综合使用 CS1 与 TS 在数据集 MI 1 上的实验结果。

### 3.6 小结

黎曼几何方法作为脑机接口领域近年来兴起的一大类方法，已经被证明能够取得很好的效果。同时，RA 作为在黎曼空间框架下特别为 EEG 数据提出的一种迁移学习方法，代表了脑机接口领域迁移学习的当前最高水平。然而，RA 在通道数或者源用户样本过多时，会遭遇到一定的计算问题。因此，本章针对 RA 的计算困难提出了两种优化思路来分别降低协方差矩阵的维度和源用户样本数。实验显示这些优化方法可以提升 RA 的计算效率和迁移效果，而当这两种优化方法结合使用时，RA 的迁移效果能够得到进一步地提升。

## 4 欧式空间中的数据对齐方法

### 4.1 引言

在前两章中，本文分别针对脑机接口领域的两种机器学习框架，即欧式空间中的传统分类框架和基于黎曼几何的分类框架，提出了新的迁移学习算法或者优化方法。具体地，本文在第 2 章提出了一种新的针对 CSP 滤波器的迁移学习方法，在第 3 章提出了两种在黎曼空间中提升 RA 迁移效果的优化方法。总的来说，它们都是局限于当前框架下的算法，不能够很容易地嵌入到其它算法中。

本章则提出一种新的欧式空间中的数据对齐方法（Euclidean-space Alignment, EA）。该方法是一种更底层的迁移学习算法，其特点在于直接对齐不同用户 EEG 样本的原始数据，且不改变 EEG 样本的数据结构。因此，在 EA 对齐之后，既可以使用传统机器学习算法进行分类，也可以使用黎曼方法进行分类。

本章内容安排如下：4.2 节总结迁移方法 RA 的几个缺陷；4.3 节提出一种新的数据对齐方法 EA；4.4 节介绍数据集并进行数据可视化；4.5 节在离线实验场景中对比 EA 与 RA 的效果；4.6 节在在线实验场景中对比 EA 与 RA 的效果；4.7 节总结本章内容。

### 4.2 黎曼空间数据对齐方法（RA）的局限

尽管本文在上一章中提到，RA 代表了迁移学习在脑机接口领域的先进水平，而且进一步提出了两种优化的方法，但 RA 仍然存在一些上述优化方法没有解决的局限：

1. RA 的可解释性问题：RA 使用休息时段的 EEG 信号（参考状态）来对齐任务实施阶段的 EEG 信号（样本），尽管文献<sup>[112]</sup>从直观理解的角度描述了为什么要这样做，但仍缺乏相应的数学模型来说明参考状态与样本之间的关系；
2. RA 需要计算协方差矩阵的黎曼中心，这在通道数较多时会耗费大量计算时间。尽管本文在上一章提出了用挑选通道的方法来降低协方差矩阵的维度，但这样有可能会损失一定的信息量；
3. RA 只针对 EEG 样本的协方差矩阵进行对齐，所以后续需要使用针对协方差矩阵分类的算法，比如 MDRM，这限制了 RA 的应用场景；
4. 对于 P300 范式的 EEG 数据，两个实验样本之间不存在休息时段，所以 RA 需要用负类样本代替休息时段信号作为参考状态。也就是说，RA 需要一定数量的目标用户标注样本，不能用于完全无监督的场景。

### 4.3 欧式空间中的数据对齐方法

针对 RA 的局限，本节提出一种欧式空间中的数据对齐方法 EA，该方法在上述局限之处都作了改进：

1. EA 不使用休息时段的 EEG 信号，而是直接使用所有样本的平均协方差矩阵作为参考矩阵，这样的好处是原理清晰，下一小节将对此作详细阐述；
2. EA 在欧式空间对齐数据，所以只需要计算协方差矩阵的算术中心，计算效率会得到提升；
3. EA 对齐的对象是原始样本而非协方差矩阵，且不改变原始样本的维度。所以，在 EA 对齐之后，既可以使用传统机器学习算法进行分类，也可以使用黎曼方法进行分类。
4. 对于 P300 范式的 EEG 数据，EA 同样只使用样本而不需要任何标签，所以可以用于完全无监督的场景。

#### 4.3.1 EA 的原理与实现

本节介绍数据对齐方法 EA 的原理以及实现方式。

在机器学习中，常常会用到一种叫“白化”的预处理步骤，可以降低输入的冗余性。假设输入为  $X \in \mathfrak{R}^{N \times D}$ ，其中  $N$  为样本个数， $D$  每个样本的特征维度。则首先计算输入的协方差矩阵：

$$C = X^T X \quad (4.1)$$

白化是对输入  $X$  执行线性变换：

$$\tilde{X} = XW \quad (4.2)$$

其中转换矩阵  $W$  满足  $W^T W = C^{-1}$ ，常用的 ZCA 白化的转换矩阵为  $W = C^{-\frac{1}{2}}$ 。

则白化转换后的数据协方差矩阵为：

$$\tilde{C} = \tilde{X}^T \tilde{X} = W^T X^T X W = I \quad (4.3)$$

其中  $I$  代表单位矩阵，意味着新数据的特征之间不相关，且方差相同。

需要注意的是，上面提到的协方差矩阵与 EEG 样本的协方差矩阵不同：白化中的协方差矩阵针对的输入样本都是向量，协方差矩阵的元素是特征维度之间的协方差；但 EEG 样本是矩阵而非向量，其协方差矩阵是在单个样本而非所有样本上求得的，所以其协方差矩阵的元素是该样本通道之间的协方差。这意味着如果我们计算一个用户的平均协方差矩阵，再将其用于“白化”该用户的所有样本，则用户转换后的平均协方差矩阵为单位矩阵。如若所有用户的平均协方差矩阵都为单位矩阵，意味着它们非常重要的一个二阶统计量，即样本协方差矩阵的分布会变得更相似。

具体地，假设当前用户有  $N$  个样本  $\{X_i\}_{i=1}^N$ ，EA 首先计算其平均协方差矩阵：

$$\bar{R} = \frac{1}{N} \sum_{i=1}^n X_i X_i^T, \quad (4.4)$$

然后对所有样本执行如下变换：

$$\tilde{X}_i = \bar{R}^{-1/2} X_i. \quad (4.5)$$

在变换之后，该用户的平均协方差矩阵会变成单位矩阵：

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \tilde{X}_i \tilde{X}_i^T &= \frac{1}{N} \sum_{i=1}^N \bar{R}^{-1/2} X_i X_i^T \bar{R}^{-1/2} \\ &= \bar{R}^{-1/2} \left( \frac{1}{n} \sum_{i=1}^N X_i X_i^T \right) \bar{R}^{-1/2} \\ &= \bar{R}^{-1/2} \bar{R} \bar{R}^{-1/2} = I, \end{aligned} \quad (4.6)$$

当分别对源用户和目标用户执行上述变换后，它们的平均协方差矩阵都会变成单位矩阵。鉴于样本协方差矩阵是 EEG 信号的一个重要统计量，用户间的差异也会变小，这正是迁移学习想要达到的效果。

### 4.3.2 EA与RA的对比

EA 和 RA 都是对数据进行变换，它们在计算上的不同之处在于：

1. 在计算转换矩阵时，RA 是计算休息状态协方差矩阵的黎曼中心，而 EA 是计算样本协方差矩阵的算术平均；
2. RA 对齐的对象是样本协方差矩阵，而 EA 对齐的是样本自身（时域 EEG 信号）。
3. RA 对齐使得用户的休息状态协方差矩阵中心变为单位矩阵，而 EA 使得用户的样本协方差矩阵中心变为单位矩阵，显然 EA 是一种更直接的对齐。

## 4.4 实验数据集

本章使用两个运动想象数据集和一个 P300 诱发电位数据集来验证 EA 的有效性。其中运动想象数据集仍然为 MI 1 和 MI 2，并且采取与上一章同样的预处理步骤，即首先进行 [8,30]Hz 的频域滤波，然后截取实验试次开始后 [2.5, 5.5] 秒之间的数据作为样本，[6.25, 7.25] 秒之间的数据作为休息时段数据（RA 的参考状态）。

#### 4.4.1 RSVP 数据集介绍

P300 诱发电位数据集来自 PhysioNet<sup>①</sup> [115]。该数据集通过快速序列视觉呈现 (Rapid Serial Visual Presentation, RSVP) 来激发 P300 电位: 给用户快速地呈现一系列图片, 每当一个与众不同的图片或者用户感兴趣的图片出现, 用户的 EEG 信号会在该图片刺激约 300 毫秒之后呈现一个幅值的波峰, 这种会激发 P300 电位的图片称为“目标图片”(或正类样本), 通常只占系列图片的一小部分, 而大部分的普通图片称为“非目标图片”(或负类样本)。在本数据集中, 给用户呈现的系列图片都是伦敦的航拍图片, 其中目标图片里都包含一个以任意角度出现的飞机, 而非目标图片则不包含飞机。实验的任务是通过 EEG 信号判断用户当前看到的是目标图片还是非目标图片。

我们以实验范式“RSVP”来命名这个数据集, 它包含 11 个健康用户的 EEG 数据。实验过程中每个用户都坐在电脑屏幕前, 然后快速地在电脑屏幕上播放系列图片。实验还基于不同的图片播放速率采集了不同的数据版本, 包含 5Hz, 6Hz, 10Hz 三个不同的版本。每个版本又采集了“a”, “b”两个阶段, 两阶段之间有一段长时间的休息。本章则采用 5Hz 版本(每秒播放 5 张图片)的“a”阶段数据, 该阶段的数据都是由 8 个 EEG 通道以 2048Hz 的采样频率采集而来, 不同用户包含的样本数从 368 到 565 不等, 其中目标图片与非目标图片的数量之比在 1:9 左右。

数据在采集时已经由 [0.15,28]Hz 的频率滤波。此外, 为了降低运算代价, 本章将数据从 2048Hz 降采样到 64Hz, 然后截取每张图片出现后 [0, 0.7] 秒的数据作为一个样本。

表格 5.1总结了上述数据集的关键性质。

表 4.1 数据集总结。

	数量				
	通道	用户	样本数/用户	类别	类别比例
MI 1	59	7	200	2	1:1
MI 2	22	9	144	2	1:1
RSVP	8	11	368~565	2	1:9

#### 4.4.2 数据可视化

本节通过可视化技术在数据的不同阶段观察 EA 对数据的影响, 以及它是怎样起到迁移效果的。

① <https://www.physionet.org/physiobank/database/ltrsvp/>.

首先观察 EEG 信号幅值在 EA 对齐前后的对比，图4-1展示了两个随机选取的样本，它们来自数据集 MI 2 的用户 1，其中一个为左手运动想象样本，另一个为右手运动想象样本。在图4-1中，黑色和红色的曲线分别代表 EEG 信号在 EA 之前和 EA 之后的幅值波形，且每一条曲线对应着一个通道。图中横轴代表时间，纵轴上的数字代表对应红线和黑线之间的相关性。可以观察到，在 EA 对齐后，脑电信号的幅值变得普遍较低，且更加均匀。另外 EA 之前和 EA 之后的幅值相关性普遍较低，说明 EA 对数据的改变较大。

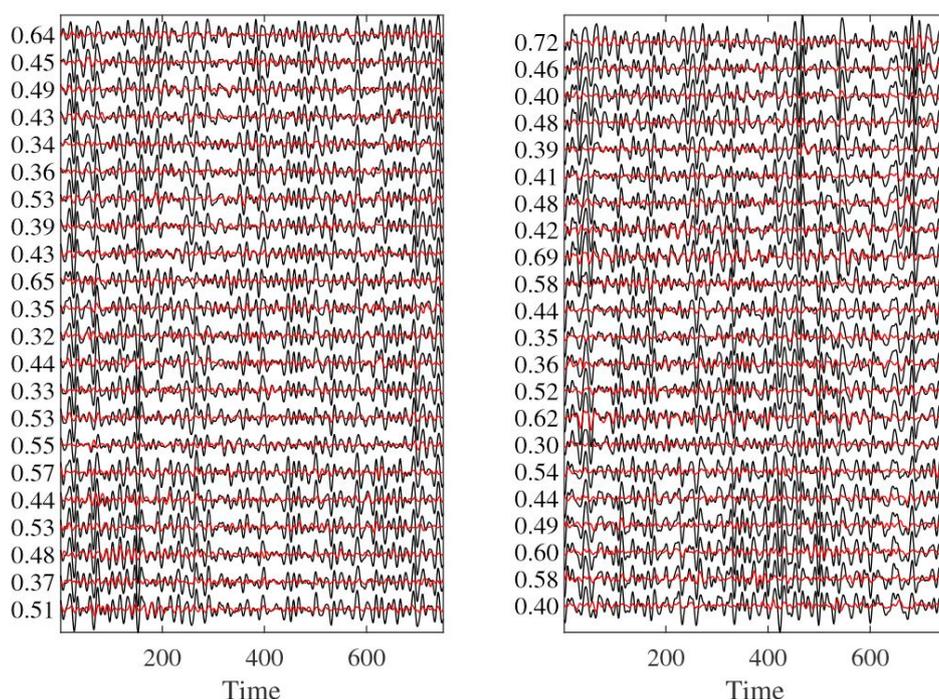
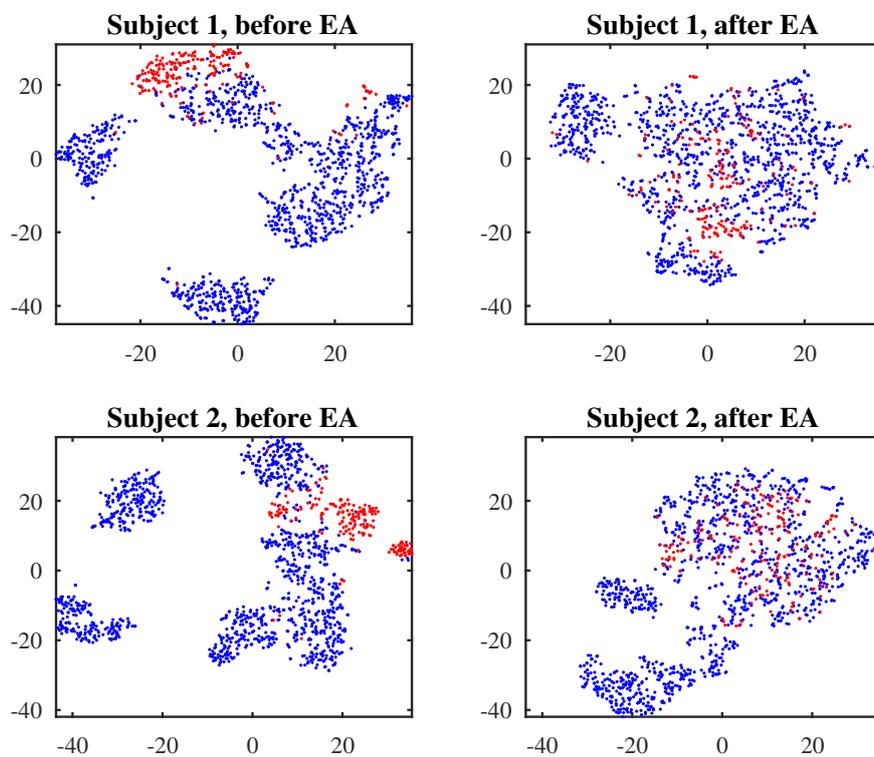


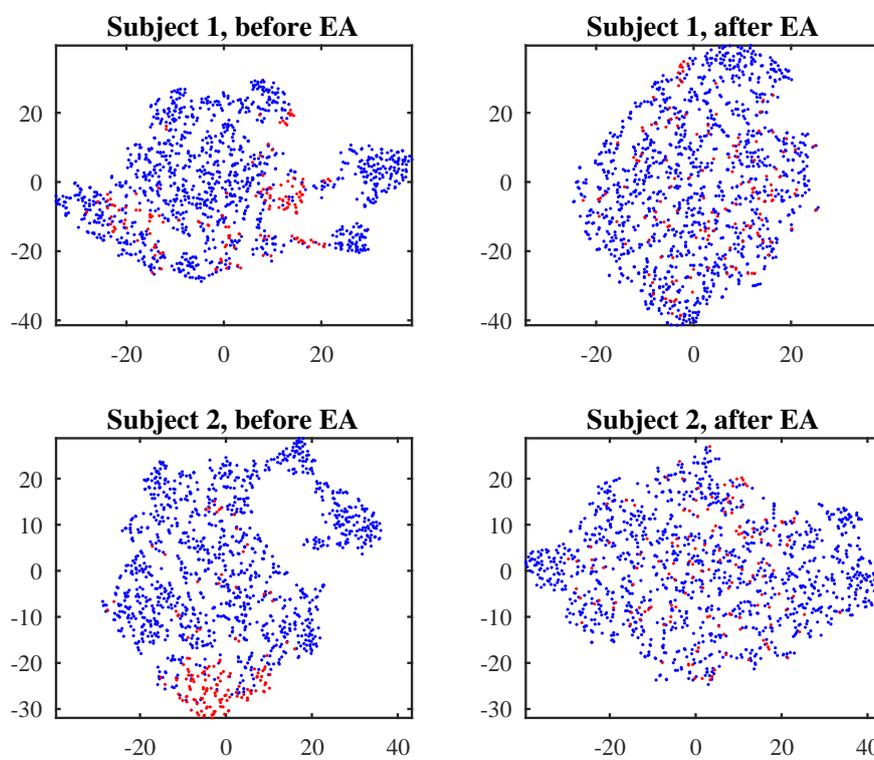
图 4-1 EA 之前（黑色曲线）和 EA 之后（红色曲线）的脑电信号幅值对比。

接下来观察 EA 对最后提取的特征有怎样的影响，以及 EA 是否减少了用户间的特征分布差异。为此，本节对每个用户执行 EA，保存 EA 之前和 EA 之后的数据，并对这两种数据分别执行以下可视化步骤，可视化结果如图4-2和图4-3所示：

1. 数据划分：采用留一法选择目标用户与源用户，即依次选择一个用户作为目标用户，其它用户作为源用户；
2. 特征提取：合并源用户数据作为训练集，对运动想象数据集，训练一个 CSP 滤波器，然后分别用于源用户和目标用户的滤波，并提取信号方差的对数作为特征；对 RSVP 数据集，将每个样本按通道合并为向量，然后使用主成分分析提取前 20 个特征；
3. 可视化：对提取的特征使用 t-SNE<sup>[116]</sup> 把特征降到 2 维，再在二维空间中将所有样本可视化。



(a)



(b)

图 4-2 对特征的可视化: a) 数据集 MI 1; b) 数据集 MI 2。

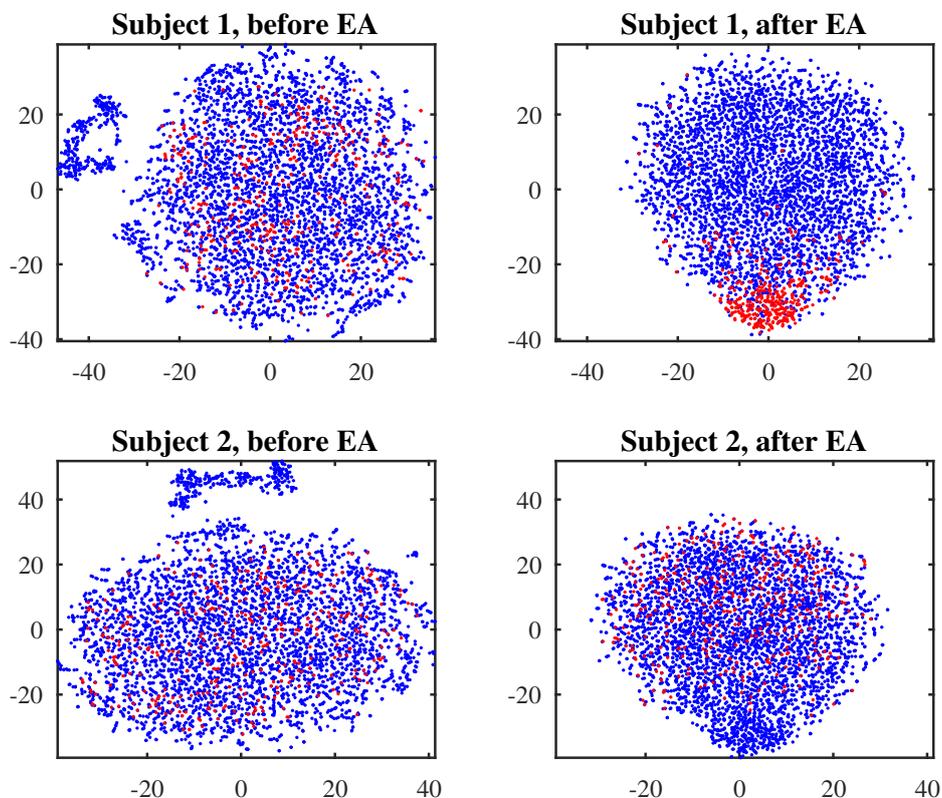


图 4-3 对特征的可视化: 数据集 RSVP

对每个数据集，我们都展示前两个用户分别为目标用户时的可视化结果。图4-2是对数据集 MI 1 和 MI 2 上的可视化结果，图4-3是对数据集 RSVP 的可视化结果。每一行代表一个目标用户，左图代表 EA 之前，右图代表 EA 之后，红色圆点代表目标用户样本，蓝色圆点代表源用户样本。可以观察到，源用户样本与目标用户样本在 EA 之前分布间隔很远，此时若将训练集上的模型直接应用于测试集，分类效果会很差。而在 EA 之后，红色圆点和蓝色圆点相互之间重叠的较多，这意味着它们之间特征分布差异变小了。

EA 在原理上是不同用户的平均协方差矩阵都对齐到单位矩阵，所以可视化 EA 前后的协方差矩阵能够更加直观地观察到 EA 的影响。因此，我们以同样的方法划分源用户和目标用户，计算所有样本的协方差矩阵，映射到切空间，然后对切空间向量使用 tSNE 降维到 2 维并可视化。图4-4以数据集 MI 2 的前 2 个用户为例，展示了 EA 的效果。同样地，左图代表 EA 之前，右图代表 EA 之后，红色圆点代表目标用户样本，蓝色圆点代表源用户样本。可以观察到，EA 之前各个用户的协方差矩阵分布间隔很远，各自聚成一团；而在 EA 之后，各个用户都以单位矩阵为中心，因此分布距离也变近了。

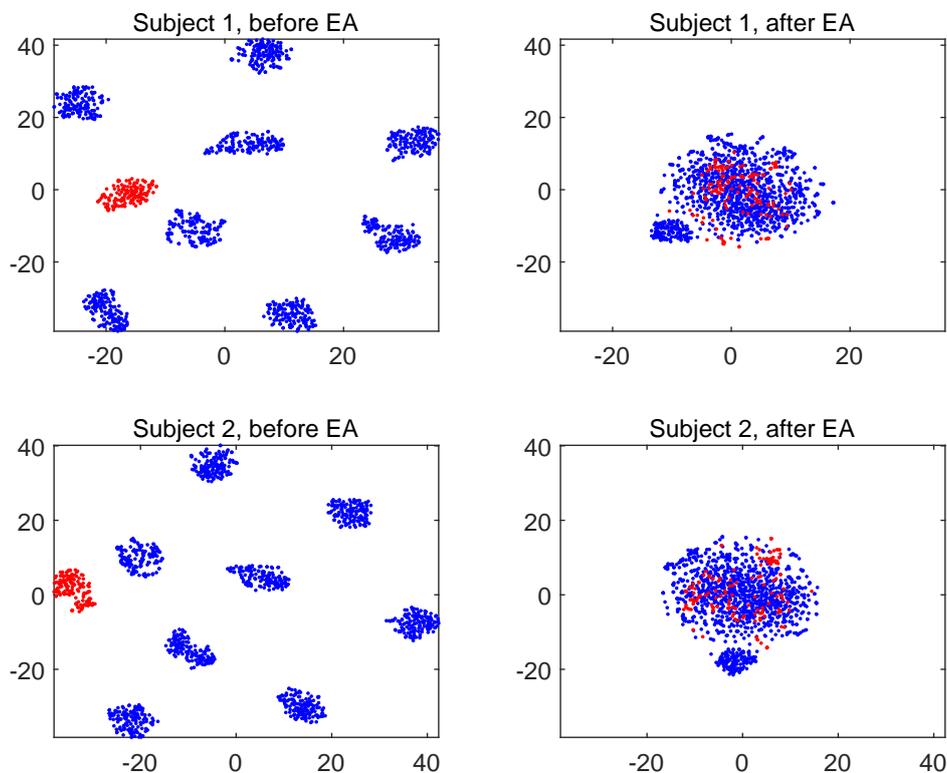


图 4-4 对源用户和目标用户的可视化。

## 4.5 离线无监督场景实验

本节在离线场景下测试 EA 对最后分类结果的影响，并将基于 EA 的算法与基于 RA 的算法进行对比。对每个数据集，实验采用留一法选择目标用户与源用户，即依次选择一个用户作为目标用户，其它用户作为源用户。离线场景假设目标用户的数据已经被采集，但是我们不知道这些数据的标签，而源用户都是已标注数据。无论是 EA 还是 RA，实验中都是先对每一个用户（包括源用户和目标用户）分别执行对齐操作，然后合并对齐后的源用户样本作为训练集，训练分类模型，最后用于对齐后目标用户数据的预测。

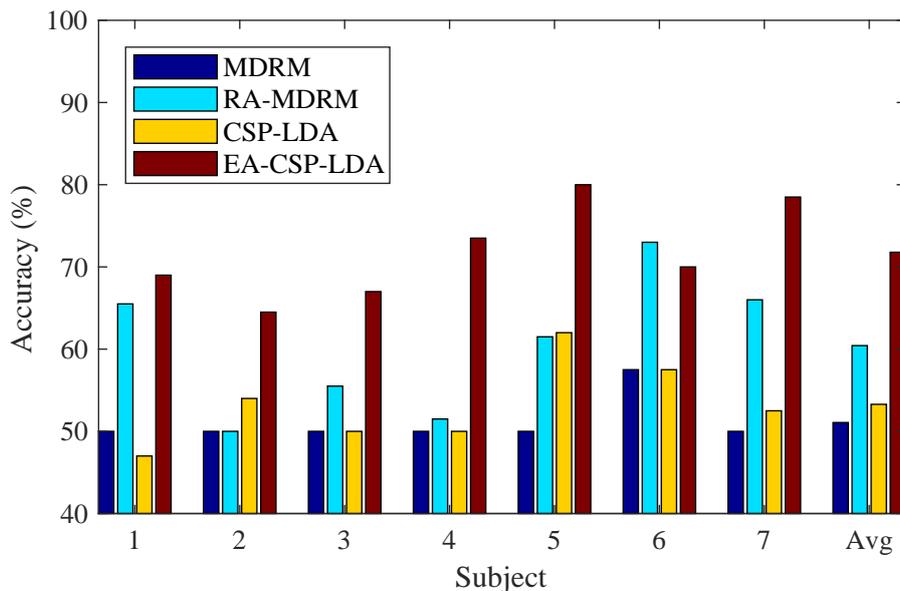
### 4.5.1 运动想象数据集实验结果

本节使用运动想象数据集 MI 1 和 MI 2 来测试 EA 的性能，并将之与 RA 的迁移效果对比。对于 RA，我们使用黎曼空间中的经典分类器 MDRM；对于 EA，本节也使用欧式空间中的经典分类算法：先使用共同空间模式（CSP）进行滤波，再使用线性判别分析（LDA）进行分类。具体地，本节比较以下算法：

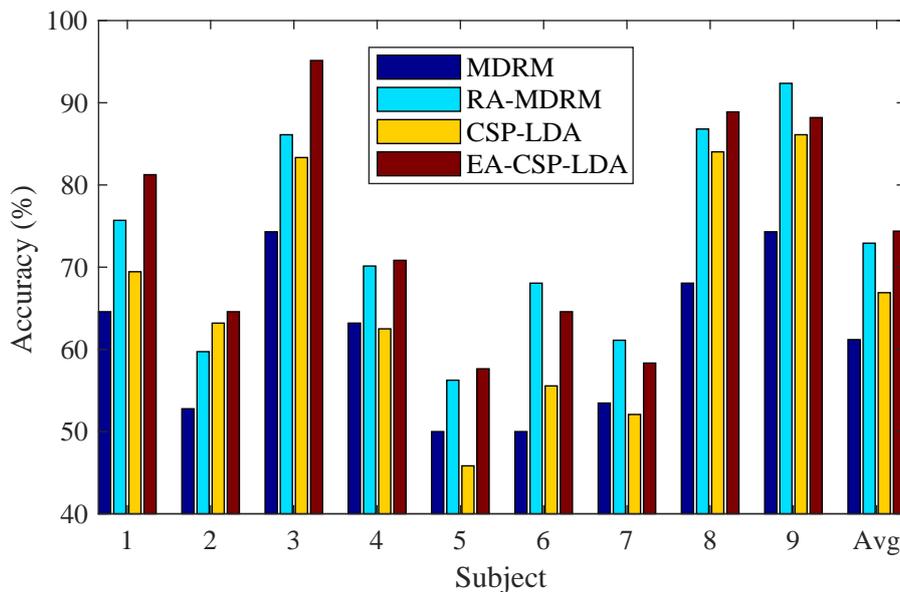
1. **MDRM**: 直接使用黎曼空间中的经典分类器 MDRM，而不使用任何数据对齐方法。
2. **RA-MDRM**: 先使用 RA 对齐，再用 MDRM 分类。

3. CSP-LDA: 先使用 CSP 进行滤波, 再使用 LDA 分类, 不包括任何数据对齐方法。
4. EA-CSP-LDA: 先用 EA 对齐不同用户, 再用 CSP 进行滤波, 最后用 LDA 分类。

以上四种算法在运动想象数据集上的分类结果如图4-5和表4.2所示。



(a)



(b)

图 4-5 离线场景下的分类正确率 (%)：a) 数据集 MI 1；b) 数据集 MI 2。

表 4.2 离线场景下的分类正确率 (%)。

Dataset	Subject	MDRM	RA-MDRM	CSP-LDA	EA-CSP-LDA
MI 1	1	50.0000	64.5000	47.0000	<b>69.0000</b>
	2	50.0000	49.0000	54.0000	<b>64.5000</b>
	3	50.0000	56.0000	50.0000	<b>67.0000</b>
	4	50.0000	54.0000	50.0000	<b>73.5000</b>
	5	50.0000	59.5000	62.0000	<b>80.0000</b>
	6	57.5000	69.0000	57.5000	<b>70.0000</b>
	7	50.0000	66.0000	52.5000	<b>78.5000</b>
	avg	51.0714	59.7143	53.2857	<b>71.7857</b>
MI 2	1	62.5000	72.9167	69.4444	<b>81.2500</b>
	2	50.6944	56.2500	50.6944	<b>63.7944</b>
	3	62.5000	84.0278	84.7222	<b>90.9722</b>
	4	63.8889	66.6667	61.8056	<b>70.1389</b>
	5	50.0000	<b>62.5000</b>	45.1389	53.4722
	6	50.0000	61.1111	52.7778	<b>61.8056</b>
	7	56.2500	<b>62.5000</b>	55.5556	60.4167
	8	56.9444	79.1667	<b>87.5000</b>	84.7222
	9	53.4722	<b>86.8056</b>	79.7667	75.6944
	avg	56.2500	70.2160	65.2006	<b>71.2963</b>

从图4-5和表4.2中可以观察到:

1. 对两个数据集共 16 个用户, RA-MDRM 在其中 15 个用户上的分类效果超过 MDRM, 说明 RA 是有效的。
2. 对两个数据集共 16 个用户, EA-CSP-LDA 在其中 14 个用户上的分类效果超过 CSP-LDA, 说明本文提出的 EA 也是有效的。
3. 对两个数据集共 16 个用户, EA-CSP-LDA 在其中 11 个用户上的分类效果超过 RA-MDRM, 说明 EA 比 RA 的迁移效果更好。

接下来验证 EA 与其它方法的测试结果是否具有统计上的显著差异。我们采用配对  $t$  检验, 原假设为匹配对象之差的均值为 0, 显著性水平设为  $\alpha = 0.05$ ,  $p$  值代表符合原假设的概率, 当  $p \leq \alpha$  时则拒绝原假设, 即存在显著性差异。在进行配对  $t$  检验之前, 先使用 Lilliefors 检验证实了观测数据符合正态分布, 最后对  $p$  值执行了 Holm-Bonferroni 校准来确保 I 类错误的概率不超过 0.05。表 4.3 展示了 EA-CSP-LDA 分别与其它三种方法的对比, 其中有显著性差异的结果用加粗表示。结果显示在数据集 MI 1 上 EA-CSP-LDA 显著优于其它所有方法, 在 MI 2 上 EA-CSP-LDA 显著优于 CSP-LDA 和 MDRM, 进一步证实了 EA 的迁移效果比 RA 更好。

表 4.3 表格4.2中不同算法在配对  $t$  检验下的  $p$  值,  $p < 0.05$  表示结果有显著性。

数据集MI 1			
	MDRM	RA-MDRM	CSP-LDA
EA-CSP-LDA	<b>0.0002</b>	<b>0.0047</b>	<b>0.0001</b>
数据集MI 2			
	MDRM	RA-MDRM	CSP-LDA
EA-CSP-LDA	<b>0.0016</b>	0.6621	<b>0.0129</b>

EA 的另一个相对于 RA 的优势是它的计算速度更快, 所以本文也比较了 EA-CSP-LDA 和 RA-MDRM 在计算时间上的差异, 本文在一台 Dell XPS15 笔记本上执行实验, 具体配置为 Intel Core i7-6700HQ CPU@2.60GHz, 16GB 内存, 512GB 固态硬盘, 运行 64 位 Windows 10 教育版以及 Matlab 2017a。计算耗费时间如表5.14所示, 可以观察到 EA-CSP-LDA 比 RA-MDRM 快 3.6-19.5 倍, 且不同实验耗费时间的方差更小。此外, RA-MDRM 在 MI 1 上耗费的时间比 MI 2 上大很多的原因是 MI 1 的通道数更多。

表 4.4 EA-CSP-LDA 和 RA-MDRM 的计算运行时间对比 (秒)。

	EA-CSP-LDA		RA-MDRM	
	Mean	std	Mean	std
MI 1	0.3864	0.0514	7.5326	0.2200
MI 2	0.2405	0.0322	0.8766	0.0729

#### 4.5.2 RSVP数据集

本节在 RSVP 数据集上测试 EA 的效果, 如前文所说, 在 RSVP 数据中, RA 和 MDRM 都需要一部分目标域带标签样本 (RA 需要负类样本来计算参考矩阵, MDRM 需要正类样本来计算增广协方差矩阵), 所以它们无法应用于无监督场景。

因此, 本节比较以下算法来验证 EA 的有效性:

1. SVM: 先使用主成分分析提取特征, 再使用 SVM 进行分类。
2. EA-SVM: 首先用 EA 对齐不同用户, 然后用主成分分析法来提取特征, 最后用 SVM 分类。
3. xDAWN-SVM: 首先使用 xDAWN 进行空域滤波, 然后用主成分分析法来提取特征, 最后用 SVM 分类。
4. EA-xDAWN-SVM: 首先用 EA 对齐不同用户, 再用 xDAWN 进行空域滤波, 然后用主成分分析法提取特征, 最后用 SVM 分类。

对上述所有算法，本文采取相同的特征提取方法和分类器：首先把每个样本的不同通道拼接成一个通道，以把 2-D 特征转变为 1-D 特征，再使用主成分分析提取前 20 个成分作为特征，然后将所有数据标准化到 [0,1] 之间，最后都使用 SVM 来进行分类。这里使用的是 LibSVM<sup>[117]</sup>，并且在训练集上使用 5 折交叉验证来选择最佳惩罚参数  $C$ 。

此外，由于 RSVP 数据有严重的类别不平衡问题（正负样本之比约为 1:9），本章采用平衡分类正确率（BCA）来作为算法的评价标准。假设  $m^+$  和  $m^-$  分别为正类样本和负类样本的真实数目， $n^+$  和  $n^-$  分别为被算法正确分类的正类样本和负类样本的数目。然后分别计算正类的分类正确率  $a_+$  和负类的分类正确率  $a_-$ ：

$$a_+ = \frac{n^+}{m^+}, \quad a_- = \frac{n^-}{m^-}, \quad (4.7)$$

则 BCA 为：

$$BCA = \frac{a_+ + a_-}{2}. \quad (4.8)$$

本节四种算法在 RSVP 上的测试结果如图 4-6 和表 4.5 所示，可以观察到：

1. 对所有 11 个用户，EA-SVM 在其中 9 个用户上的表现优于 SVM，说明 EA 在不使用空域滤波器的情形下是有效的。
2. 对所有 11 个用户，EA-xDAWN-SVM 在其中 8 个用户上表现优于 xDAWN-SVM，说明 EA 在使用空域滤波器的情形下同样是有效的。
3. 总体来说，xDAWN-SVM 与 SVM 的效果接近，但 EA-xDAWN-SVM 却表现优于 EA-SVM，说明 EA 可能有助于发掘 xDAWN 的潜在性能。

本节同样对算法之间的差异进行统计检验，以验证 EA 对 BCA 的提升是显著的。这里仍然使用配对  $t$  检验来比较 SVM 与 EA-SVM，以及 xDAWN-SVM 与 EA-xDAWN-SVM，显著性水平设为  $\alpha = 0.05$ 。检验结果如表 4.6 所示，其中符合显著性要求的被加粗表示。结果显示 EA-SVM 显著优于 SVM，且 EA-xDAWN-SVM 显著优于 xDAWN-SVM，证明了 EA 对于 RSVP 数据也是有效的。

### 4.5.3 不同参考矩阵的影响

EA 与 RA 的一个重要差异在于采用了不同的参考矩阵：EA 采用所有样本协方差矩阵的算术平均，而 RA 采用休息状态协方差矩阵的黎曼中心。相比而言，差别在于两点：

1. 计算主体不同，EA 采用所有样本的协方差矩阵，而 RA 采用休息状态的协方差矩阵。
2. 计算方式不同，EA 计算算术平均，而 RA 计算黎曼中心。

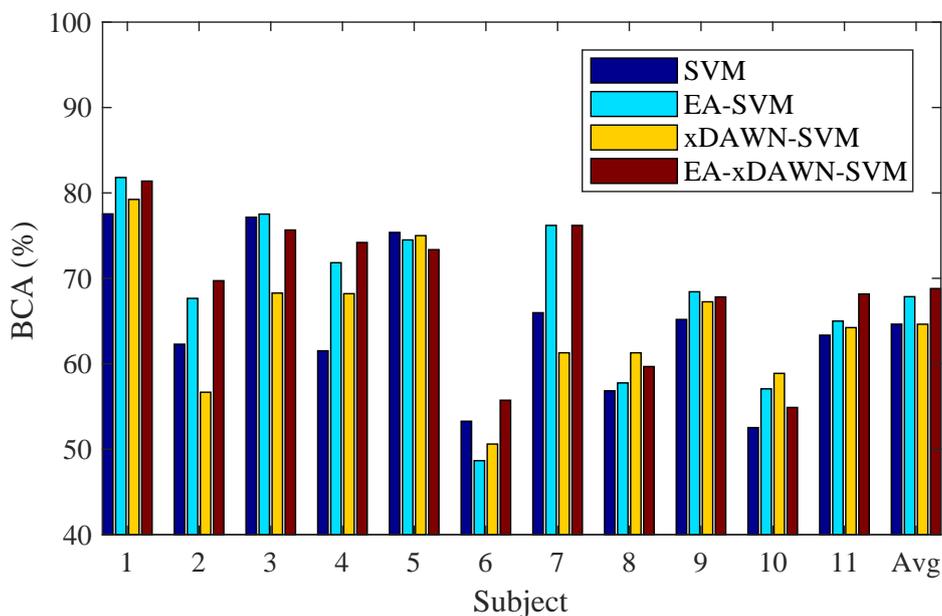


图 4-6 离线场景下所有算法在 RSVP 数据集上的 BCA(%)。

表 4.5 离线场景下所有算法在 RSVP 数据集上的 BCA(%)。

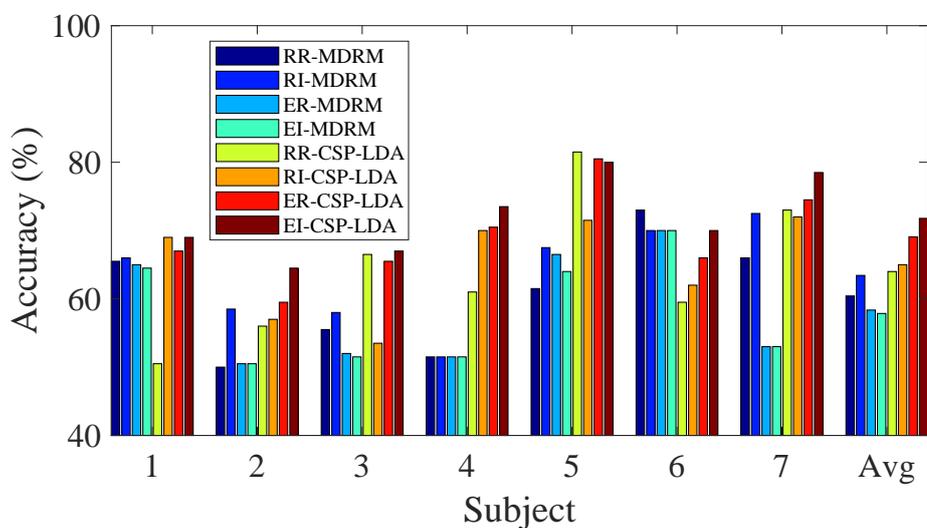
Subject	SVM	EA-SVM	xDAWN-SVM	EA-xDAWN-SVM
1	77.5445	<b>81.8049</b>	79.2408	81.3870
2	62.2882	67.6526	56.6652	<b>69.7218</b>
3	77.1559	<b>77.5210</b>	68.2753	75.6953
4	61.5170	71.8306	68.2058	<b>74.1983</b>
5	<b>75.3788</b>	74.4949	75.0000	73.3586
6	53.2692	48.6538	50.5983	<b>55.7265</b>
7	65.9753	<b>76.1983</b>	61.2844	<b>76.1983</b>
8	56.8242	57.7520	<b>61.2827</b>	59.6716
9	65.1835	<b>68.4287</b>	67.2532	67.8214
10	52.5242	57.0604	<b>58.8724</b>	54.8818
11	63.3459	64.9920	64.2316	<b>68.1676</b>
avg	64.6370	67.8536	64.6282	<b>68.8026</b>

表 4.6 表格4.5中不同算法在配对  $t$  检验下的  $p$  值,  $p < 0.05$  表示结果有显著性。

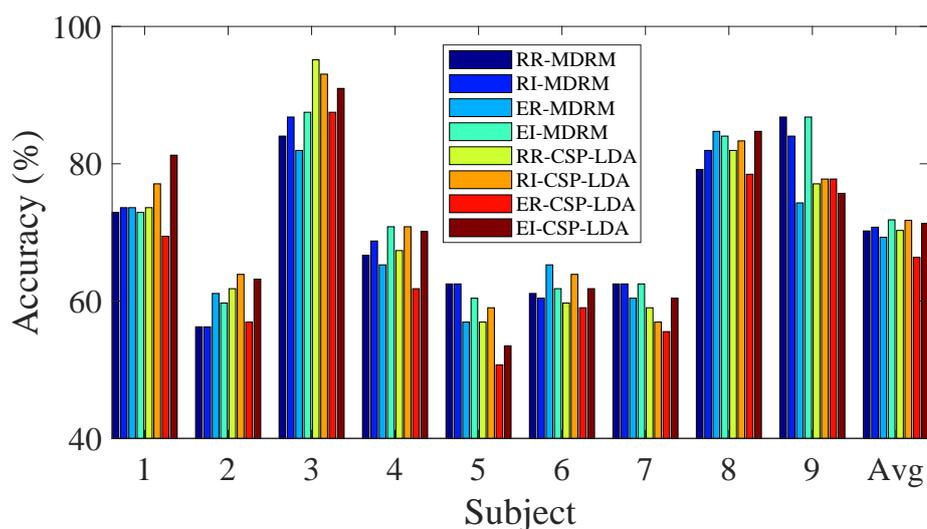
	SVM	xDAWN-SVM
EA-SVM	<b>0.0386</b>	
EA-xDAWN-SVM		<b>0.0438</b>

为了进一步比较参考矩阵的影响，我们基于计算主体和计算方式，比较四种不同的参考矩阵：样本协方差矩阵的算术平均 (*Euclidean mean of all imagery trials*, EI); 样本协方差矩阵的黎曼中心 (*Riemannian mean of all imagery trials*, RI); 休息状态协方差矩阵的算术平均 (*Euclidean mean of all resting trials*, ER); 休息状态协方差矩阵的黎曼中心 (*Riemannian mean of all resting trials*, RR);

图4-7(a) 和图4-7(b)分别表示这四种参考矩阵在数据集 MI 1 和 MI 2 上的实验结果。



(a)



(b)

图 4-7 不同参考矩阵的对比: a) 数据集 MI 1; b) 数据集 MI 2。

从图4-7中可以观察到:

1. 在两个数据集上, RI-MDRM 的平均效果优于 RR-MDRM, 且 EI-CSP-LDA 平均效果优于 ER-CSP-LDA, 说明使用样本协方差矩阵比使用休息状态协方差矩阵更好。
2. 总体来说, 使用 CSP-LDA 分类时, EI 的效果最好; 而使用 MDRM 分类时, RI 的效果最好。这一点与 EA 设计的初衷相符: RA 是为黎曼空间的分类算法所设计, 所以用黎曼中心更好; 而 EA 是为欧式空间的分类算法所设计, 所以用欧式中心更好。
3. EI-CSP-LDA 的效果比 RI-MDRM 更好, 这意味着即使我们在 RA 中使用与 EA 相同的参考矩阵, EA 仍然能取得更好的效果。

## 4.6 在线监督场景实验

本节使用同样的数据集, 但是考虑一种在线监督实验的场景: 实验开始时, 我们没有任何的目标用户的数据; 实验前期有一个校准过程, 可以逐渐获取少量的目标用户标注样本; 然后我们需要使用这些目标用户的已标注样本以及源用户数据训练模型, 用于接下来的目标用户数据(测试集)的预测。与离线实验相比, 区别主要有两点:

1. 在线实验无法使用测试集的任何信息; 而离线实验可以使用测试集的输入;
2. 在线实验可以逐渐获取少量的目标用户已标注样本(非测试集); 而离线实验的目标用户样本都是未标注的。

为了模拟这样的在线实验场景, 我们把目标用户数据的一小部分先划分为校准数据, 其它则作为测试集。以数据集 MI 1 为例, 我们首先仍然采用留一法选取目标用户与源用户; 目标用户总共有 200 个样本, 然后选取一个随机数  $n_0 \in [1, 200]$ , 将接下来的  $m$  个已标注样本  $\{X_{n_0+i}\}_{i=1}^m$  作为校准池(如果  $n_0+i$  大于 200, 则将  $n_0+i$  替换为  $n_0+i-200$ ), 剩下的  $200-m$  样本作为测试集; 数据划分完成后, 我们从零开始, 每次从校准池中选择  $r$  个已标注样本添加到训练集, 训练相应的模型并用于测试集的分类, 直到校准池的数据被全部添加完毕。

对于 EA 或者 RA, 由于源用户数据都是已经采集并标注完成的, 所以它们的对齐过程不受影响, 我们只需要依次对每一个源用户执行对齐操作。但是对于目标用户, 我们只能够使用少量的已标注样本(被从校准池添加到训练集)来计算转换矩阵(对于 EA, 是样本平均协方差矩阵; 对于 RA, 是休息时段的平均协方差矩阵)。

### 4.6.1 运动想象数据集

本节比较的算法与4.5.1节使用的一样, 即 MDRM, RA-MDRM, CSP-LDA, 以及 EA-CSP-LDA。但是在实际使用上稍有不同:

1. 无论是 EA 还是 RA，目标用户的转换矩阵计算方式不一样。离线实验使用全部的目标用户样本来计算转换矩阵（计算过程无需标签信息），而在线实验只使用被添加到训练集的目标用户校准数据来计算转换矩阵。
2. 模型（MDRM, CSP, LDA）的训练集不一样，离线实验的训练集是所有源用户，而在线实验的训练集是所有源用户加上目标用户的已标注数据。

本节将目标用户校准池的参数设置为  $m = 40$  和  $r = 4$ ，为了获得有统计意义的结果，每个目标用户的实验都重复 30 次（每次选择一个随机的  $n_0$ ）。图4-8展示了四个算法分别在数据集 MI 1 和 MI 2 的比较结果，其中图4-8(a)的前 7 个子图分别展示了在数据集 MI 1 的 7 个目标用户上的实验结果，最后一个子图代表在 7 个目标用户上的平均结果；图4-8(b)的前 9 个子图分别展示了在数据集 MI 2 的 9 个目标用户上的实验结果，最后一个子图代表在 9 个目标用户上的平均结果。在每一个子图中，横轴代表目标用户采集标注样本的次数（标注样本数=次数 $\times$  $r$ ），纵轴代表分类正确率（%），每一条曲线都是 30 次重复实验的平均结果，误差条代表 95% 置信区间。

从图4-8中可以观察到：

1. 对两个数据集共 16 个用户，RA-MDRM 在其中 15 个用户上的效果超过 MDRM，说明 RA 在在线场景中也是有效的。
2. 对两个数据集共 16 个用户，EA-CSP-LDA 在其中 14 个用户上的效果超过 CSP-LDA，说明 EA 在在线场景中也是有效的。
3. 对两个数据集共 16 个用户，EA-CSP-LDA 在其中 12 个用户上的效果超过 RA-MDRM，说明 EA 在在线场景中的效果仍然优于 RA。

接下来是对 EA 与其它方法的结果进行显著性检验，检验方法仍然是配对  $t$  检验以及  $\alpha = 0.05$  的显著性水平。为此，我们首先定义一个衡量算法性能的指标 AUC (area under the curve)，即用每条曲线以下的面积大小来衡量对应的算法性能，然后在 AUC 值上执行配对  $t$  检验，并对  $p$  值执行 Holm-Bonferroni 校准来确保 I 类错误的概率不超过 0.05。检验结果如表格4.7所示，其中有显著性的结果用粗体显示。结果显示 EA-CSP-LDA 在数据集 MI 1 上显著优于 RA-MDRM，在 MI 2 上有更高的平均值但不显著。

表 4.7 图4-8中不同算法在配对  $t$  检验下的  $p$  值， $p < 0.05$  表示结果有显著性。

数据集 MI 1			
	MDRM	RA-MDRM	CSP-LDA
EA-CSP-LDA	<b>0.0006</b>	<b>0.0006</b>	<b>0.0008</b>
数据集 MI 2			
	MDRM	RA-MDRM	CSP-LDA
EA-CSP-LDA	<b>0.0016</b>	0.8722	0.0598

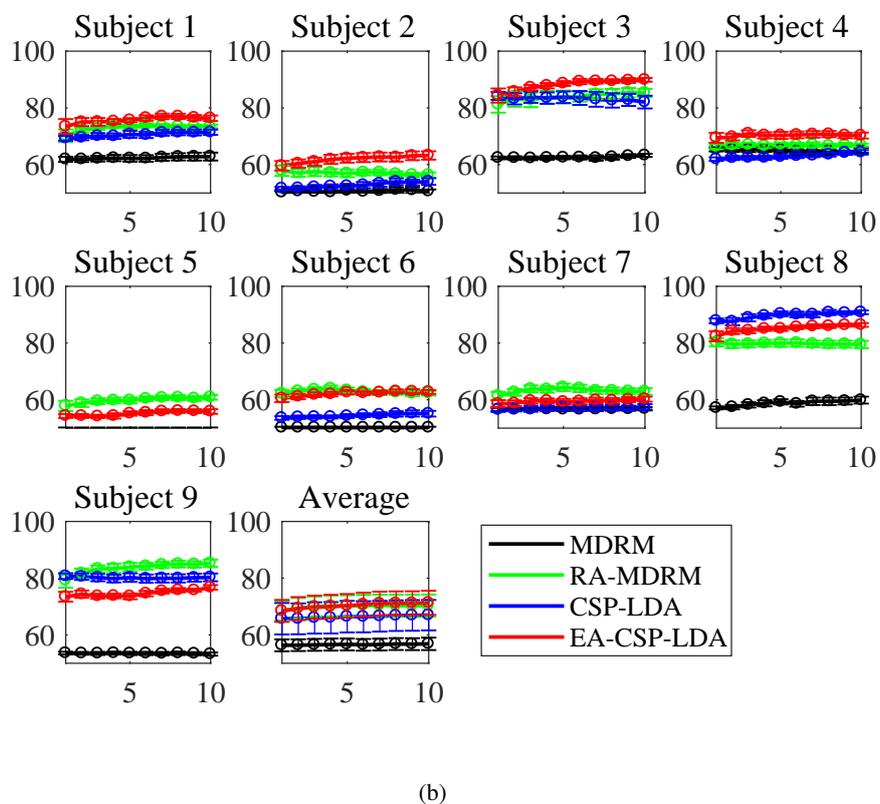
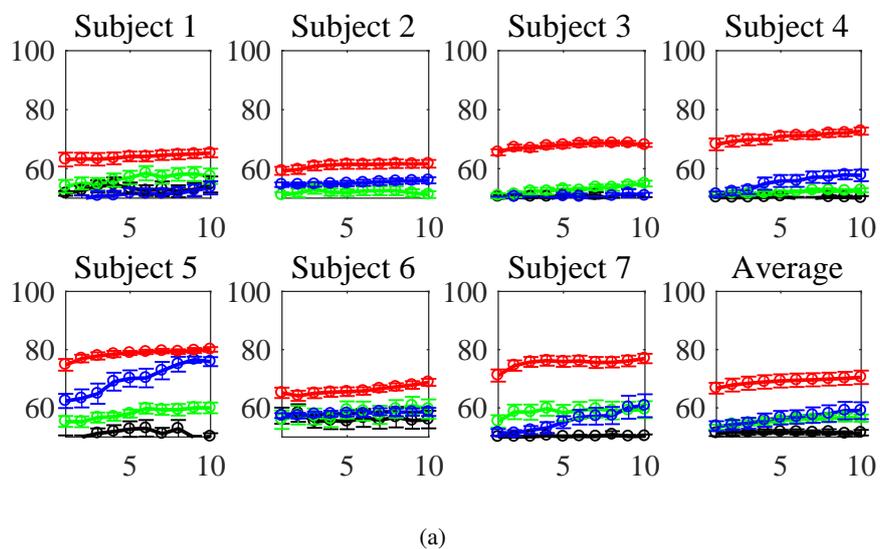


图 4-8 在线场景下的分类正确率 (%)：a) 数据集 MI 1；b) 数据集 MI 2。

### 4.6.2 RSVP数据集

接下来验证在线场景下 EA 在 RSVP数据集上的效果，需要注意的是，本节比较的算法与4.5.2节使用的不同。由于在线实验允许使用部分的目标用户已标注样本，所以我们加入了 MDRM 和 RA-MDRM 的对比：

1. MDRM: 不执行任何对齐操作，直接使用 MDRM 分类；
2. RA-MDRM: 先进行 RA 对齐，再使用 MDRM 分类；
3. xDAWN-SVM: 不执行任何对齐操作，先使用 xDAWN 进行空间滤波，然后用主成分分析法提取前 20 个成分作为特征，最后使用 SVM 分类；
4. EA-xDAWN-SVM: 先进行 EA 对齐，再使用 xDAWN 进行空间滤波，然后用主成分分析法提取前 20 个成分作为特征，最后使用 SVM 分类。

由于 RSVP 数据集每个用户的样本数较多且存在类别不平衡问题，我们将目标用户校准池的参数设置为  $m = 80$  和  $r = 10$ 。同样地，为了获得有统计意义的结果，每个目标用户的实验都重复 30 次（每次选择一个随机的  $n_0$ ）。实验结果如图4-9所示，在每一个子图中，横轴代表目标用户采集标注样本的次数（标注样本数=次数 $\times r$ ），纵轴代表分类正确率（%）。

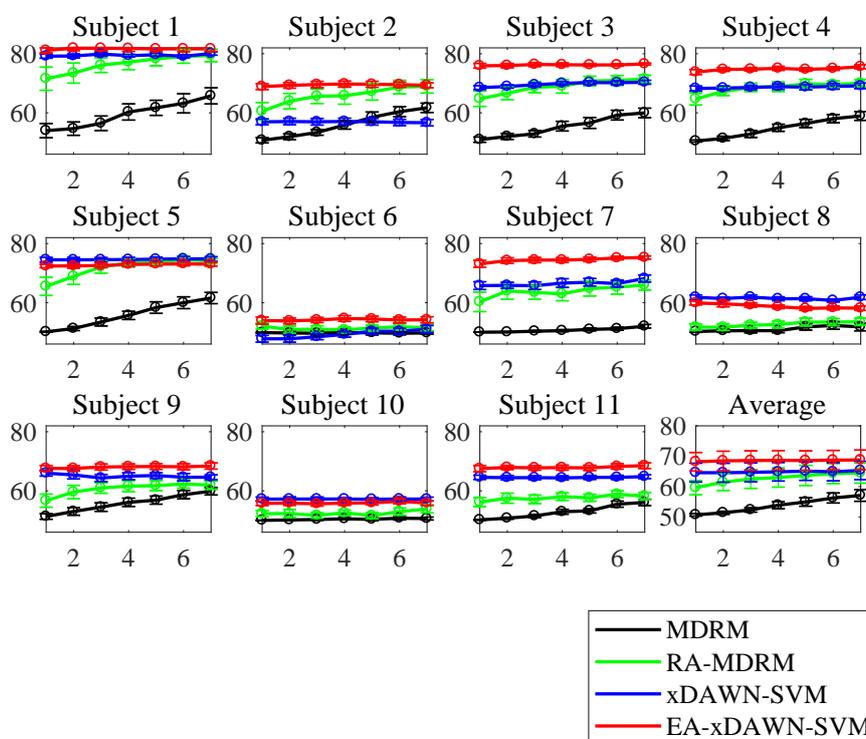


图 4-9 RSVP 数据集模拟在线场景的平衡分类正确率（%）。

从图4-9中可以观察到:

1. EA-xDAWN-SVM 在所有 11 个用户上的效果优于 RA-MDRM, 说明本文提出的 EA 在在线场景下相比 RA 也有更大的优势;
2. 对所有 11 个用户, EA-xDAWN-SVM 在其中 8 个用户上的效果优于其它所有算法, 再次证明了 EA 的有效性。

同样地, 我们在平均 AUC 上执行配对  $t$  检验以及  $\alpha = 0.05$  的显著性水平来对图4-9中的结果进行显著检验。检验结果如表格4.8 所示, 其中有显著性的结果用粗体显示。结果显示 EA-xDAWN-SVM 在数据集 RSVP 显著优于其它所有算法。

表 4.8 图4-9中不同算法在配对  $t$  检验下的  $p$  值,  $p < 0.05$  表示结果有显著性。

	MDRM	RA-MDRM	xDAWN-SVM
EA-xDAWN-SVM	<b>0.0000</b>	<b>0.0001</b>	<b>0.0214</b>

## 4.7 小结

本章提出一种针对脑机接口领域的对齐方法 EA, 该方法有如下优点: (1) EA 计算简单、容易实现; (2) EA 不需要任何目标用户标签信息, 可用于完全无监督的场景; (3) EA 对齐不影响原数据的结构, 在执行对齐后, 既可以使用传统机器学习算法进行分类, 也可以使用黎曼方法进行分类。通过在两个运动想象数据集和一个 P300 数据集上的实验, 我们证明 EA 在离线场景和在线场景下都有很好的迁移效果, 且显著性优于脑机接口领域当前最好的迁移学习方法之一。

本文作者基于 EA 算法获得了“2019 世界机器人大赛—BCI 脑控机器人大赛一等奖”。

## 5 异构类别空间的对齐方法

### 5.1 引言

本文在1.4.2节提到了迁移学习可以分为不同的类型，根据源域与目标域的特征空间  $\mathcal{X}$ 、边缘概率分布  $P(x)$ 、类别空间  $\mathcal{Y}$ 、条件概率分布  $P(y|x)$  的异同，迁移学习可以划分为：

1. 同构迁移学习：源域与目标域的特征空间和类别空间都相同，但是边缘概率分布与条件概率分布不同；
2. 异构迁移学习：异构迁移学习可以细分为异构特征空间迁移学习，即源域与目标域的特征空间和边缘概率分布都不同；以及异构类别空间迁移学习，即源域与目标域的类别空间和条件概率分布都不相同。

目前大多数迁移学习方法考虑的都是同构迁移学习类型，包括本文第2章到第4章，都是研究同构迁移学习。相较同构迁移学习来说，异构迁移学习是一个更困难的问题，因为它不仅需要减少源域和目标域之间分布的差异，还需要先对特征空间或者标签空间进行转化。而在异构迁移学习中，异构类别空间又是一个比异构特征空间更具挑战性，也是目前被研究得更少的问题。

在脑机接口领域，异构特征空间通常是由于用户使用了不同规格的脑电帽，这些脑电帽的电极数量或者在头皮表面安放的位置不一样，从而导致特征空间不一样。而异构类别空间通常是因为用户执行了不同数量或者类别的任务，比如在运动想象实验中，源用户是想象“左手”和“右手”的运动，而目标用户是想象“双脚”和“舌头”的运动。

在脑机接口领域目前很少有异构迁移学习的相关研究，据我们所知，只有 Wu 等<sup>[118]</sup>研究了异构特征空间的情形，并采取了一种直接的解决思路，即从源用户的数据中挑选出那些与目标用户最相近的通道。而关于异构类别空间的问题，还没有相关研究。

本章将针对异构类别空间这样一个最具挑战性的问题开展研究，提出一种新的异构类别空间的对齐方法（heterogeneous Label spaces Alignment, LA），并证明 LA 能够广泛地适用于多种不同的场景。内容安排如下：5.2节简要介绍文献中的异构类别空间迁移学习方法，并指出本文研究的场景与现有方法的应用场景的差别；5.3节针对 EEG 数据提出一种新的异构类别空间的迁移学习方法 LA；5.4节比较 LA 与上一章的数据对齐方法 EA 的异同；5.5节基于两个运动想象数据集模拟多种应用场景，并在这些场景下证明 LA 的有效性，以及 LA 能够与其它迁移学习方法兼容使用；5.6节总结本章内容。

## 5.2 异构类别空间的场景

异构类别空间在迁移学习中是一个较难的问题，相对而言被研究得较少，本节对文献中出现过的场景进行简单的总结。

当前在其它领域针对异构类别空间的迁移学习方法可以分为开集域适应、局部域适应和通用域适应，如图5-1所示。图中闭集域适应针对的是源域和目标域有相同的类别，即同构类别空间问题；Busto等<sup>[119]</sup>首先提出开集域适应（open set domain adaptation）的概念，针对的是源域和目标域包含一部分共同的类别，但也各自包含一些未知的类别；Saito等<sup>[120]</sup>则假设目标域包含源域的所有类别，以及一些其它的未知类别；You等<sup>[121]</sup>提出通用域适应（universal domain adaptation），假设目标域类别全都未知，且这些未知的类别既包含与源域类别相同的部分，也包含与之不同的部分。

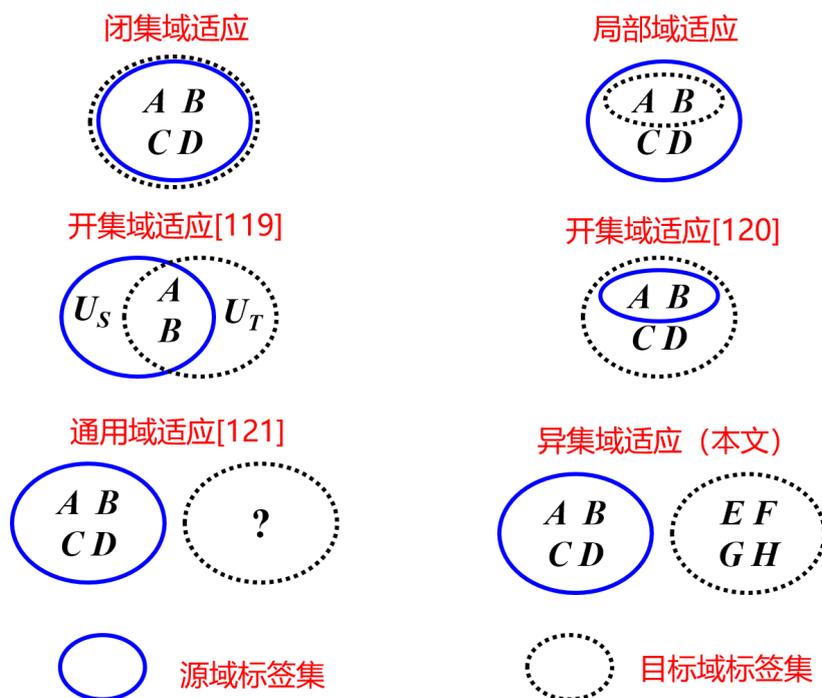


图 5-1 域自适应场景示意图：A, B, ..., H 代表不同的类别， $U_s$ 和 $U_t$ 分别是源域和目标域的未知类别。在闭集域适应中，源域和目标域有相同的类别；在文献<sup>[119]</sup>考虑的开集域适应中，源域和目标域包含一部分共同的类别，但也各自包含一些未知的类别；在文献<sup>[120]</sup>考虑的开集域适应中，源域类别集合是目标域类别集合的子集；在通用域适应<sup>[121]</sup>中，目标域类别集合未知；在本文提出的异集域适应中，源域类别与目标域类别既可以部分不同，也可以完全不同。

在分类目标上，开集域适应和通用域适应都是在源域数据上训练一个模型，应用在目标域样本上：若目标域样本的类别属于源域类别集合，则将之划分到对应类

别；若目标域样本不属于任何源域类别，则将之标记为“未知”。一个常用的应用场景是人脸识别，需要把不存在于训练数据库的测试样本标记出来，并拒绝其权限。

本章考虑的是异集域适应 (different set domain adaptation)，与前面提到的方法相比，一个关键区别在于对那些不存在于源域类别集合的目标域样本，不能只将其标记为“未知”，而是要标注为它真实所属的类别。

### 5.3 类别空间对齐方法

本节针对 EEG 数据提出一种新的类别空间对齐方法 LA，使得当源用户的类别空间与目标用户不同时，我们依然可以使用源用户的数据来辅助目标用户模型的建立。

假设源用户和目标用户都有  $M$  个类别，但是它们的类别集合有部分甚至全部不同。LA 首先对源用户和目标用户的类别进行一一配对：当有相同类别时，优先使相同的类别匹配；当没有相同类别时，可使它们随机匹配。

然后，LA 会为源用户每个类别的样本寻找一个转换矩阵  $A_m$ ：

$$A_m = \arg \min_A \|\bar{A} \bar{C}_{S,m} A^T - \bar{C}_{T,m}\|_F^2, \quad m = 1, 2, \dots, M. \quad (5.1)$$

其中  $\bar{C}_{S,m}$  是源用户第  $m$  个类别的平均协方差矩阵， $\bar{C}_{T,m}$  是目标用户对应匹配类别的平均协方差矩阵，这里我们使用 *Log-Euclidean* 均值<sup>[122]</sup>来计算平均协方差矩阵，它是在半正定矩阵的均值计算中常用的度量，并且比黎曼均值更简单快速。

式 (5.1) 的优化目标是使得转换后的源用户与目标用户在每对匹配类别的平均协方差矩阵都能够距离最小化，本节使用文献<sup>[86]</sup>中的优化方法来解目标函数 (5.1)，得到：

$$A_m = \bar{C}_{T,m}^{-\frac{1}{2}} \bar{C}_{S,m}^{-\frac{1}{2}}, \quad m = 1, 2, \dots, M. \quad (5.2)$$

把所有类别的转换矩阵  $\{A_m\}_{m=1}^M$  都求出后，对每一个源用户样本，我们基于它所属类别进行以下变换：

$$\tilde{X}_j = A_m X_j, \quad \text{if } X_j \in \text{Class } m \quad (5.3)$$

转换后，优化目标的损失变为：

$$\begin{aligned} \bar{A} \bar{C}_{S,m} A^T - \bar{C}_{T,m} &= \bar{C}_{T,m}^{-\frac{1}{2}} \bar{C}_{S,m}^{-\frac{1}{2}} \bar{C}_{S,m} \bar{C}_{S,m}^{-\frac{1}{2}} \bar{C}_{T,m}^{-\frac{1}{2}} - \bar{C}_{T,m} \\ &= \bar{C}_{T,m}^{-\frac{1}{2}} I \bar{C}_{T,m}^{-\frac{1}{2}} - \bar{C}_{T,m} = \mathbf{0}, \end{aligned} \quad (5.4)$$

其中  $\mathbf{0}$  指一个全零矩阵，即式 (5.1) 中的目标函数已经被最小化了。

LA 的一个关键在于当缺乏目标用户标签信息时如何衡量其各个类别的平均协方差矩阵，即式 (5.1) 中的  $\bar{C}_{T,m}$ 。本文考虑一种离线场景：假设目标用户的 EEG 信号已经被采集，且我们能够对其中一小部分进行标注。问题是该如何合理地选择出一小部分样本才能尽量准确地估计出  $\bar{C}_{T,m}$  呢？为了解决这个问题，LA 首先基于目标用户所有的样本协方差矩阵之间的黎曼距离进行 *k-medoids* 聚类，将目标用户样本划分为  $k$  个簇后，对这  $k$  个簇的中心点样本进行标注，然后使用这  $k$  个标注样本来估计类别中心  $\bar{C}_{T,m}$ 。在极少情形下， $k$  个标注样本所包含的类别个数小于  $M$ ，我们就使用 EA 来代替 LA。

---

**Algorithm 3:** LA 对齐过程。

---

**Input:**  $\{X_i\}_{i=1}^{N_T}$ : 目标用户数据;  
 $\{y_{T,m}\}_{m=1}^M$ : 目标用户的类别空间;  
 $\{X_j, y_j\}_{j=1}^{N_S}$ : 源用户数据;  
 $\{y_{S,m}\}_{m=1}^M$ : 源用户的类别空间;  
 $k$ : 对目标用户执行 *k-medoids* 聚类的簇数，也是对目标用户进行标注的样本数目。

**Output:**  $\{\tilde{X}_j, \tilde{y}_j\}_{j=1}^{N_S}$ : 转换后的源用户数据;  
 计算目标用户每个样本的协方差矩阵  $\{C_i\}_{i=1}^{N_T}$ ;  
 基于目标用户样本协方差矩阵之间的黎曼距离，对目标用户样本执行 *k-medoids* 聚类，将目标用户样本划分为  $k$  个簇;  
 对这  $k$  个簇的中心样本进行标注，得到  $k$  个标注样本;  
 对这  $k$  个标注样本按类别求中心，即为  $\{\bar{C}_{T,m}\}_{m=1}^M$ ;  
 计算源用户每个样本的协方差矩阵，并求每个类别的中心，即  $\{\bar{C}_{S,m}\}_{m=1}^M$ ;  
 将源用户的标签与目标域标签进行一一配对，假设  $y_{T,m}$  匹配  $y_{S,m}$ ;

**for**  $m = 1, 2, \dots, M$  **do**

根据式 (5.2) 计算转换矩阵  $A_m$ ;  
 根据式 (5.3) 对源用户标签为  $y_{S,m}$  的样本进行转换得到  $\tilde{X}_j$ ;  
 按照标签匹配规则更改源用户样本的标签:  $\tilde{y}_j = y_{T,m}$ ;

**end**

**Return**  $\{\tilde{X}_j, \tilde{y}_j\}_{j=1}^{N_S}$

---

## 5.4 LA 与 EA 的比较

LA 与本文在上一章提出的 EA 存在一定的关联，本节分析它们的异同。一般来讲，不同数据分布间的偏移可分为三种：

1. 协变量偏移<sup>[123]</sup>：输入特征的分布不同，即  $P_S(X) \neq P_T(X)$ 。

2. 先验概率偏移: 输出的分布不同, 即  $P_S(Y) \neq P_T(Y)$ 。

3. 概念偏移<sup>[124]</sup>: 输入与输出之间的关系不同, 即  $P_S(X | Y) \neq P_T(X | Y)$ 。

EA 针对的是协变量偏移, 它不考虑任何类别信息, 仅仅使源用户和目标用户的样本协方差矩阵分布尽量相近。而 LA 则基于源用户和目标用户的类别中心, 同时考虑了以上三种数据偏移, 其目标是使得源用户与目标用户配对类别之间的分布尽量靠近。图5-2 展示了 EA 与 LA 两种对齐思路的差异, 为方便展示, 我们假设有一个源用户和一个目标用户, 它们各自包含两个类别。图中蓝色点代表源用户的样本协方差矩阵, 黑色点代表目标用户的样本协方差矩阵, 每种形状代表一个类别。

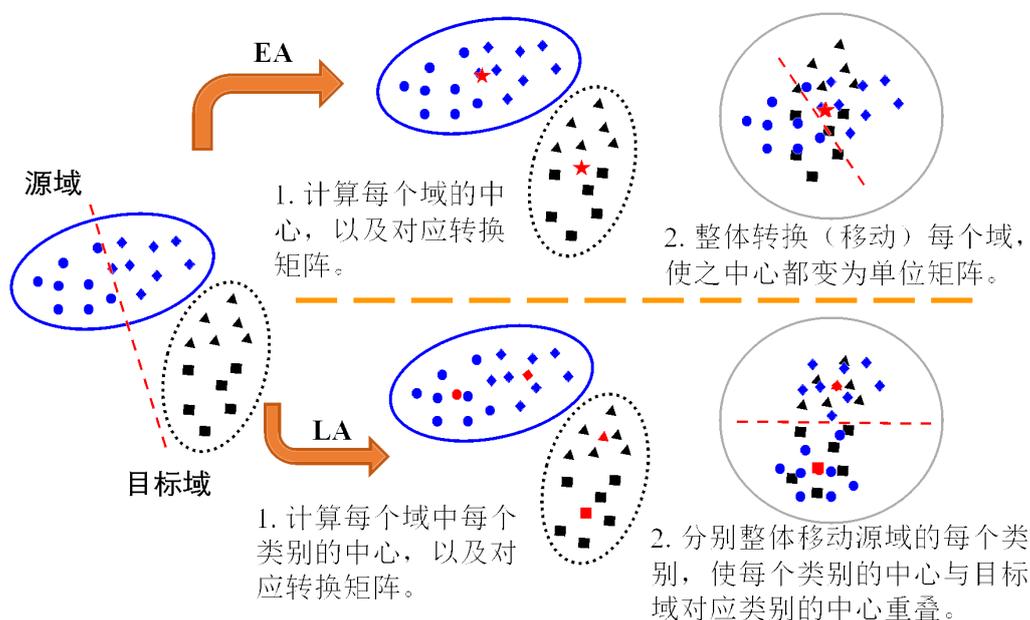


图 5-2 EA与LA的对齐过程示意图。

从图5-2中可以观察到, 由于特征分布差异和类别差异, 源用户与目标用户的点在对齐之前相隔较远。若此时在源用户上建立分类模型, 如图中红色虚线所示, 它可能对目标用户的分类没有任何帮助, 所以我们需要在建立模型之前先减少源用户与目标用户之间的分布差异:

1. EA 专注于减少特征分布差异而忽略类别差异。它对源用户和目标用户采取相同的处理步骤: (1) 计算用户的平均协方差矩阵 (图中用红色五角星表示); (2) 基于此平均协方差矩阵计算一个转换矩阵; (3) 使用转换矩阵对用户每一个样本进行映射, 这个映射相当于对域内所有点进行整体移动 (移动过程中域内点之间的距离保持不变), 使得域中心为单位矩阵。如此, 源用户和目标用户都以单位矩阵为中心, 意味着原本分布相隔很远的两个域在对齐后重叠度较高。此时再在源用户上训练一个分类模型如图中红色虚线所示, 它就能更好地应用于目标用户。

2. LA 同时考虑特征分布差异与类别差异。它对源用户和目标用户采取不同的处理步骤：（1）对目标用户，首先进行  $k$ -medoids 聚类，将目标用户划分为  $k$  个簇后，对这  $k$  个簇的中心点进行标注，然后使用这  $k$  个标注样本来估计每个类别的中心（分别用红色正方形与红色三角形表示）；（2）对源用户，直接根据标签计算每个类别的中心（分别用红色圆形与红色菱形表示）。然后将源用户与目标用户的类别一一配对，基于配对的类别中心计算转换矩阵，并用于源域样本的映射。这个映射相当于对源用户每个类别的样本进行整体移动，使得源用户每个类别的中心与目标用户配对类别的中心重叠。此时再在源域样本上训练一个分类器，它就能更好地分类目标用户样本。

对比 EA 与 LA 的示意图，可以推论当存在类别差异时，LA 对齐的效果会比 EA 更好。

## 5.5 实验

本节在两个运动想象数据集 MI 1 和 MI 2 上对 LA 进行验证，数据集已在 2.5 节介绍。特别地，对于数据集 MI 2，本节使用全部 4 个类别的数据并将它划分为多个子数据集，数据集总结如表格 5.1 所示。另外 P300 诱发电位数据都只有正负两个类别，所以不适合用于异构类别空间的实验。

表 5.1 数据集总结。

	数量				
	通道	采样点	用户	类别	样本数/类别
MI 1	59	300	7	2	100
MI 2	22	300	9	4	72

### 5.5.1 异构类别空间场景

本章研究的迁移学习场景是源用户与目标用户的类别空间不一致，为了充分验证 LA 在不同情形下的效果，我们进一步考虑了以下细分场景：

1. 场景 I-a: 源用户与目标用户的特征空间一致，类别空间部分不一致（二分类问题）。
2. 场景 I-b: 源用户与目标用户的特征空间一致，类别空间部分不一致（多分类问题）。
3. 场景 II-a: 源用户与目标用户的特征空间一致，类别空间完全不一致（二分类问题）。

4. 场景 II-b: 源用户与目标用户的特征空间一致, 类别空间完全不一致 (多分类问题)。
5. 场景 III: 源用户与目标用户的特征空间与类别空间都不一致。

为了创造场景 I-a, I-b, II-a, 以及 II-b, 我们将数据集 MI 2 根据类别进行拆分: 在每个实验中, 先将数据集 MI 2 拆分为两个子集, 一个作为源数据集, 另一个作为目标数据集。比如子集 “1, 2” 代表其只包含标签为 1 和 2 的样本, 而子集 “3, 4” 代表其只包含标签为 3 和 4 的样本 (这样子集 “1, 2” 就与子集 “3, 4” 有相同的特征空间, 以及完全不同的类别空间), 然后用 “1, 2→3, 4” 代表以子集 “1, 2” 作为源数据集, 以子集 “3, 4” 作为目标数据集。

上述细分场景的数据划分规则为:

1. 场景 I-a: 每次将 MI 2 划分为一个源数据集和一个目标数据集, 使得源数据集与目标数据集有一个相同的类别和一个不同的类别, 比如 “1, 2→1, 3”, 总共有 24 种满足要求的数据集组合。
2. 场景 I-b: 每次将 MI 2 划分为一个源数据集和一个目标数据集, 使得源数据集与目标数据集有两个相同的类别和一个不同的类别, 比如 “1, 2, 3→1, 2, 4”, 总共有 12 种满足要求的数据集组合。
3. 场景 II-a: 每次将 MI 2 划分为一个源数据集和一个目标数据集, 使得源数据集与目标数据集的两个类别完全不同, 比如 “1, 2→3, 4”。总共有 6 种满足要求的数据集组合。
4. 场景 II-b: 使用与场景 I-b 相同的数据集组合, 但是使所有标签都错位匹配, 比如 “1, 2, 3→2, 1, 4”, 总共有 12 种满足要求的数据集组合。
5. 场景 III: 以 MI 1 为源数据集, MI 2 的子数据集 “3, 4” 为目标数据集, 这样源数据集和目标数据集的特征空间和类别空间都不相同。

在经过上述数据集划分后, 我们在每个场景下都进行了多次实验来验证 LA 的效果, 各个场景下的实验数据如表格 5.2 所示:

表 5.2 异构类别空间场景, 其中  $k$  为目标用户带标注样本个数。

	数量			
	数据集组合	实验	训练集样本	测试集样本
场景 I-a	24	$24 \times 9$	$144 \times 8 + k$	$144 - k$
场景 I-b	12	$12 \times 9$	$216 \times 8 + k$	$216 - k$
场景 II-a	6	$6 \times 9$	$144 \times 8 + k$	$144 - k$
场景 II-b	12	$12 \times 9$	$216 \times 8 + k$	$216 - k$
场景 III	1	9	$1400 + k$	$144 - k$

在表格5.2中,  $k$  为  $k$ -medoids 聚类的簇数, 也是目标用户的标注样本数目。在每个数据集组合中, 我们都用留一法选择目标用户和源用户: 每次实验从目标数据集中选择一个用户作为目标用户, 源数据集中的其他用户作为源用户, 由于目标数据集有 9 个用户, 所以每个数据集组合都包含 9 个实验。

### 5.5.2 数据可视化

由于本文涉及的实验场景和数据集组合较多, 本节以数据集组合“1, 4→2, 3”为例, 使用可视化技术分析 LA 的作用, 以及它与 EA 的效果对比。

具体地, 分别选择数据集“2, 3”的前两个用户为目标用户, 数据集“1, 4”的其它用户为源用户。对每个用户, 计算其所有样本的协方差矩阵, 映射到切空间, 然后对切空间向量使用 tSNE 降维到 2 维并可视化, 如图5-3所示。图中第一列代表没有进行任何对齐, 第二列代表 EA 对齐之后, 第三列代表 LA 对齐之后, 其中红色圆点和黑色圆点代表目标用户的两个类别, 蓝色圆点和绿色圆点代表源用户的两个类别。可以观察到, 在没有进行对齐时, 各个用户的分布间隔很远, 各自聚成一团; 在 EA 之后, 各个用户都以单位矩阵为中心, 源用户和目标用户的样本分布重叠在一起, 但不同类别之间也有重叠; 而在 LA 之后, 不仅源用户和目标用户的分布距离变近了, 而且类别的可分性也加强了。

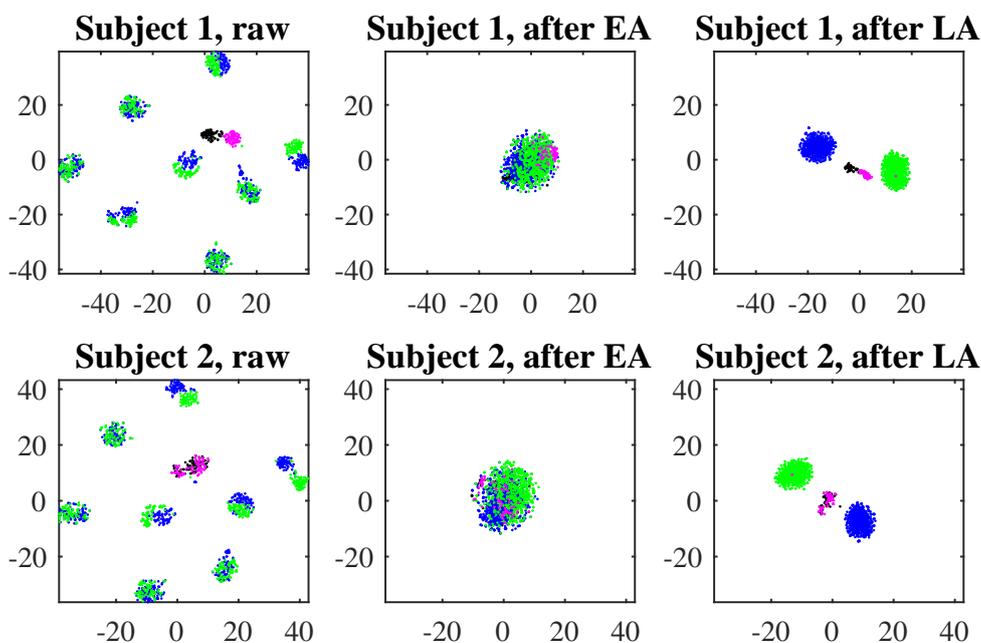


图 5-3 对协方差矩阵的可视化。

### 5.5.3 实验目标与设置

为了在不同算法中强调 LA 的效果，我们首先将所有算法都拆分为三个阶段：

1. **预处理**：对 EEG 信号进行预处理，如2.5.1节所描述。
2. **对齐**：对数据使用 EA 或者 LA 对齐。
3. **分类**：提取特征，选择性使用其它迁移学习方法，然后在训练集上建立分类模型，并用于测试集。

为了强调 LA 的迁移效果，我们在预处理阶段与分类阶段相同的条件下，比较以下三种对齐阶段的方法：

1. **Raw**：不进行任何对齐操作；
2. **EA**：分别对源用户和目标用户执行 EA 对齐；
3. **LA**：对源用户执行 LA 对齐。

并且在每个场景下，我们都讨论以下两个问题：

1. **问题1**：LA 是否适用于不同的特征提取与分类方法？
2. **问题2**：LA 是否可以与其它迁移学习方法结合使用，并进一步提高迁移效果？

### 5.5.4 将 LA 用于不同特征提取与分类方法

为了研究问题 1，即 LA 是否适用于不同的特征提取与分类方法，我们在分类阶段考虑以下两种分类框架：

1. **CSP-LDA**：先用 CSP 滤波器进行空间滤波，然后提取方差的对数作为特征，最后使用 LDA 分类器分类；
2. **TS-SVM**：先计算样本协方差矩阵，然后提取切空间特征，最后使用 SVM 分类器分类。

把这两种分类方法与对齐阶段的 Raw, EA 和 LA 分别结合，则共有 6 种方法需要比较，我们的目的是验证 LA 是否在两种分类框架下都能取得最好的效果。

#### 5.5.4.1 场景 I-a

场景 I-a 研究在二分类问题下，源用户与目标用户的类别空间部分不一致的情形。如5.5.1节所总结的，本节共考虑 24 种满足场景 I-a 的数据集组合，每种组合下源用户与目标用户都有一个相同的类别与一个不同的类别。

对每种数据集组合，我们先将相同的类别配对，再将另一组不同的类别配对。比如数据集组合“1, 2  $\rightarrow$  1, 3”，我们先将源用户类别“1”与目标用户类别“1”配对，再将源用户类别“2”与目标用户类别“3”配对。也就是说，在执行 LA 对齐时，我们把源用户标签为“1”的样本与目标用户标签为“1”的样本对齐，再把源

用户标签为“2”的样本与目标用户标签为“3”的样本对齐，并把源用户的标签“2”都换为“3”。

对于涉及 LA 的算法，有一个参数  $k$  需要设置，即执行  $k$ -medoids 聚类时的簇数。为了验证该参数的鲁棒性，我们将其设为  $k \in \{2, 4, \dots, 20\}$  并观察 LA 在每一个取值下的效果。为了公平比较，我们把所有算法的训练集都设为源用户样本加上目标用户的  $k$  个已标注样本，测试集都设定为目标用户的未标注样本。

我们在两种分类框架下分别比较对齐阶段的三种方法，实验结果如图5-4所示。

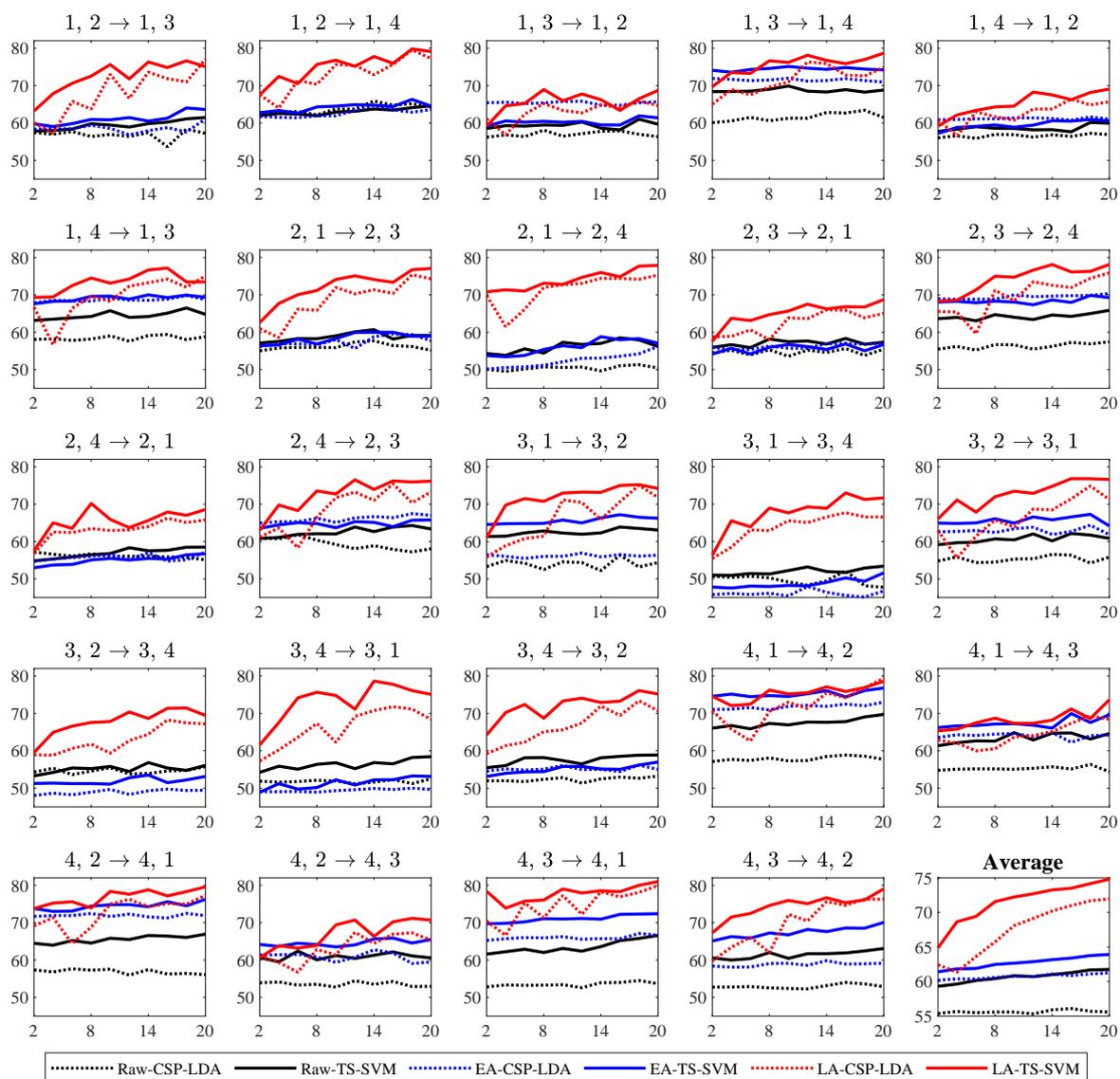


图 5-4 LA 用于不同特征提取和分类方法，在场景 I-a 下的分类正确率(%)。

在图5-4中，前 24 个子图分别代表在 24 个数据集组合上的实验结果（每个子图都是 9 个目标用户的平均结果），子图的标题即为数据集组合的方式，最后一个子图

是 24 个数据集组合的平均结果。在每一个子图中，横轴代表参数  $k$ ，即目标用户已标注样本的数量，纵轴代表分类正确率(%)，从图中可以观察到：

1. 对总共 24 种数据集组合，EA-CSP-LDA 在其中 20 个组合上的分类正确率高于 Raw-CSP-LDA，EA-TS-SVM 在其中 14 个组合上的分类正确率高于 Raw-TS-SVM。从平均效果看，EA-CSP-LDA 优于 Raw-CSP-LDA，EA-TS-SVM 优于 Raw-TS-SVM。这说明即使 EA 最初是为同构类别空间场景而设计的，它在异构类别空间场景下也能取得一定的迁移效果。
2. 当  $k$  取值稍大时，LA-CSP-LDA 在所有 24 个数据集组合上的分类正确率都高于 Raw-CSP-LDA，LA-TS-SVM 也在所有 24 个组合上的分类正确率都高于 Raw-TS-SVM，说明 LA 总是能够很好地处理类别空间部分不一致的场景。
3. 当  $k$  取值稍大时，LA-CSP-LDA 在所有 24 个数据集组合上的分类正确率都高于 EA-CSP-LDA，LA-TS-SVM 也在所有 24 个组合上的分类正确率高于 EA-TS-SVM，说明 LA 的迁移效果比 EA 更好、更稳定。
4. 总体来讲，LA-CSP-LDA 和 LA-TS-SVM 的分类正确率都随着  $k$  增大而上升。这是因为当  $k$  增大时，LA 有更多的标注样本可以用于目标用户类别中心的衡量，而目标用户的类别中心衡量得越准确，LA 的迁移效果越好。

接下来对图5-4中的结果进行统计检验，以验证基于 LA 的算法与其它算法之间的效果差异是否显著。我们采用配对  $t$  检验，显著性水平设为  $\alpha = 0.05$ 。原假设为匹配对象之差的均值为 0， $p$  值代表观测对象符合原假设的概率，当  $p \leq \alpha$  时拒绝原假设，意味着观测对象之间存在显著性差异。为此，我们使用 AUC (area under the curve)，即图中每条曲线以下的面积大小来作为评价对应算法的标准。场景 I-a 共进行了  $24 \times 9 = 216$  次实验（24 种数据集组合，每种组合选择 9 次目标用户），所以每个算法都有 216 个 AUC 值。我们在这 216 个 AUC 值上进行配对  $t$  检验，检验结果如表格5.3 所示，其中有显著性的结果用粗体显示。从表中可以看到 LA-CSP-LDA 显著优于 EA-CSP-LDA，LA-TS-SVM 显著优于 EA-TS-SVM，这些结果与从图5-4中得的结论一致：LA 在不同的特征提取和分类算法框架下都有显著的迁移效果。

表 5.3 图5-4中算法在配对  $t$  检验下的  $p$  值， $p < 0.05$  表示结果有显著性。

	EA-CSP-LDA	EA-TS -SVM
LA-CSP-LDA	<b>0.0000</b>	
LA-TS-SVM		<b>0.0000</b>

5.5.4.2 场景 I-b

场景 I-b 研究在多分类问题下，源用户与目标用户的标签空间部分不一致的情形。如5.5.1节所总结的，本节共考虑 12 种满足场景 I-b 数据集组合，每种组合下源用户与目标用户都有两个相同的类别和一个不同的类别。

对每种数据集组合，我们先将两组相同的类别配对，再将剩余一组不同的类别配对。比如数据集组合“1, 2, 3 → 1, 2, 4”，我们先将源用户类别“1”与目标用户类别“1”配对，源用户类别“2”与目标用户类别“2”配对，再将源用户类别“3”与目标用户类别“4”配对。

由于三分类问题中我们需要衡量更多的目标用户类别中心，所以在执行 *k-medoids* 聚类时将参数设为  $k \in \{3, 6, \dots, 30\}$ 。另外，在进行 CSP 滤波时，我们使用“一对多”的方式将 CSP 滤波器从二分类扩展到多分类<sup>[55]</sup>。然后我们继续在两种分类框架下比较 LA 与 Raw, EA。实验结果如图5-5所示，其中前 12 个子图分别代表在 12 个数据集组合上的实验结果（每个子图都是 9 个目标用户的平均结果），子图的标题即为数据集组合的方式，最后一个子图是 12 个数据集组合的平均结果。在每一个子图中，横轴代表 *k-medoids* 聚类的参数 *k*（即目标用户已标注样本数目），纵轴代表分类正确率(%)。

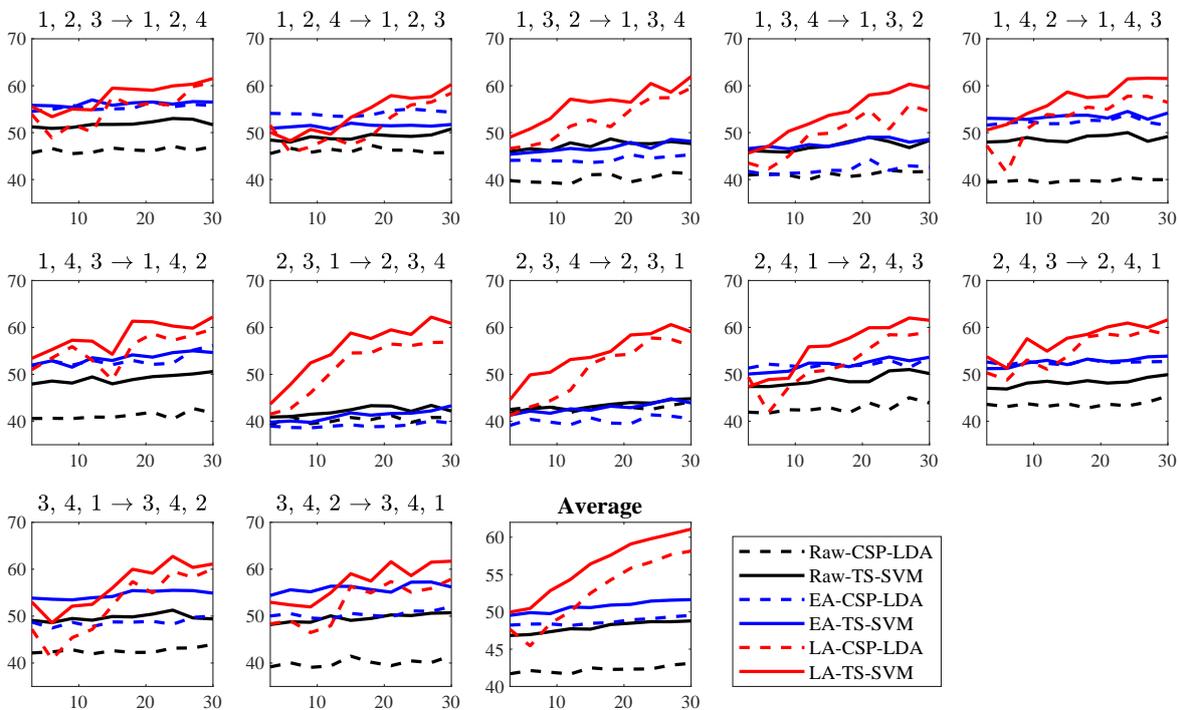


图 5-5 LA 用于不同特征提取和分类方法，在场景 I-b 下的分类正确率(%)。

从图5-5中可以观察到:

1. 对总共 12 种数据集组合，EA-CSP-LDA 在其中 10 个组合上的分类正确率高于 Raw-CSP-LDA；EA-TS-SVM 在其中 8 个组合上的分类正确率高于 Raw-TS-SVM。从平均效果看，EA-CSP-LDA 优于 Raw-CSP-LDA，EA-TS-SVM 优于 Raw-TS-SVM。这说明 EA 仍然在多数情况下是有效的，但在少数情况下也可能没有效果。
2. 当  $k$  较大时，LA-CSP-LDA 在所有 12 个数据集组合上的分类正确率都高于 Raw-CSP-LDA，LA-TS-SVM 也在所有 12 个组合上的分类正确率高于 Raw-TS-SVM，说明 LA 在多分类问题下也总是能够很好地处理类别空间部分不一致的场景。
3. 当  $k$  较大时，LA-CSP-LDA 在所有 12 个数据集组合上的分类正确率高于 EA-CSP-LDA；LA-TS-SVM 也在所有 12 个数据集组合上的分类正确率高于 EA-TS-SVM，再次说明 LA 的迁移效果比 EA 更好。
4. 当  $k$  较小时，LA 的迁移效果（对算法的提升幅度）不如场景 I-a（图5-4，二分类问题）中 LA 的迁移效果。这是因为在多分类的设置下，衡量目标用户的类别中心变得更困难了。

接下来采用同样的方法对图5-5的结果进行显著性检验，此时每个算法进行了  $12 \times 9 = 108$  次实验，所以我们在这 108 个 AUC 值上进行配对  $t$  检验，结果如表格5.4 所示，其中有显著性的结果已经用粗体显示。可以观察到 LA-CSP-LDA 显著优于 EA-CSP-LDA，LA-TS-SVM 显著优于 EA-TS-SVM。再次验证了 LA 可以作为一个有效的预处理步骤，然后用于不同的特征提取和分类算法。

表 5.4 图5-5中算法在配对  $t$  检验下的  $p$  值， $p < 0.05$  表示结果有显著性。

	EA-CSP-LDA	EA-TS-SVM
LA-CSP-LDA	<b>0.0196</b>	
LA-TS-SVM		<b>0.0010</b>

#### 5.5.4.3 场景 II-a

场景 II-a 研究在二分类问题下，源用户与目标用户的类别空间完全不一致的情形。如5.5.1节所总结的，本节共考虑了 6 种满足场景 II-a 数据集组合。

对每种数据集组合，由于源用户与目标用户的类别完全不一样，所以我们将它们随机配对。比如数据集组合“1, 2  $\rightarrow$  3, 4”，我们既可以将源用户类别“1”与目标用户类别“3”配对，源用户类别“2”与目标用户类别“4”配对；也可以将源用户类别“1”与目标用户类别“4”配对，源用户类别“2”与目标用户类别“3”配对。我们的实验结果显示 LA 在这两种情形下都是有效的。

场景 II-a 研究二分类问题，所以在执行 *k-medoids* 聚类时将参数设为  $k \in \{2, 4, \dots, 20\}$ ，然后我们继续在两种分类框架下比较 LA 与 Raw, EA。结果如图5-6所示，其中前 6 个子图分别代表在 6 个数据集组合上的实验结果（每个子图都是 9 个目标用户的平均结果），子图的标题即为数据集组合的方式，最后一个子图是 6 个数据集组合的平均结果。在每一个子图中，横轴代表 *k-medoids* 聚类的参数  $k$ （即目标用户已标注样本数目），纵轴代表分类正确率(%)。

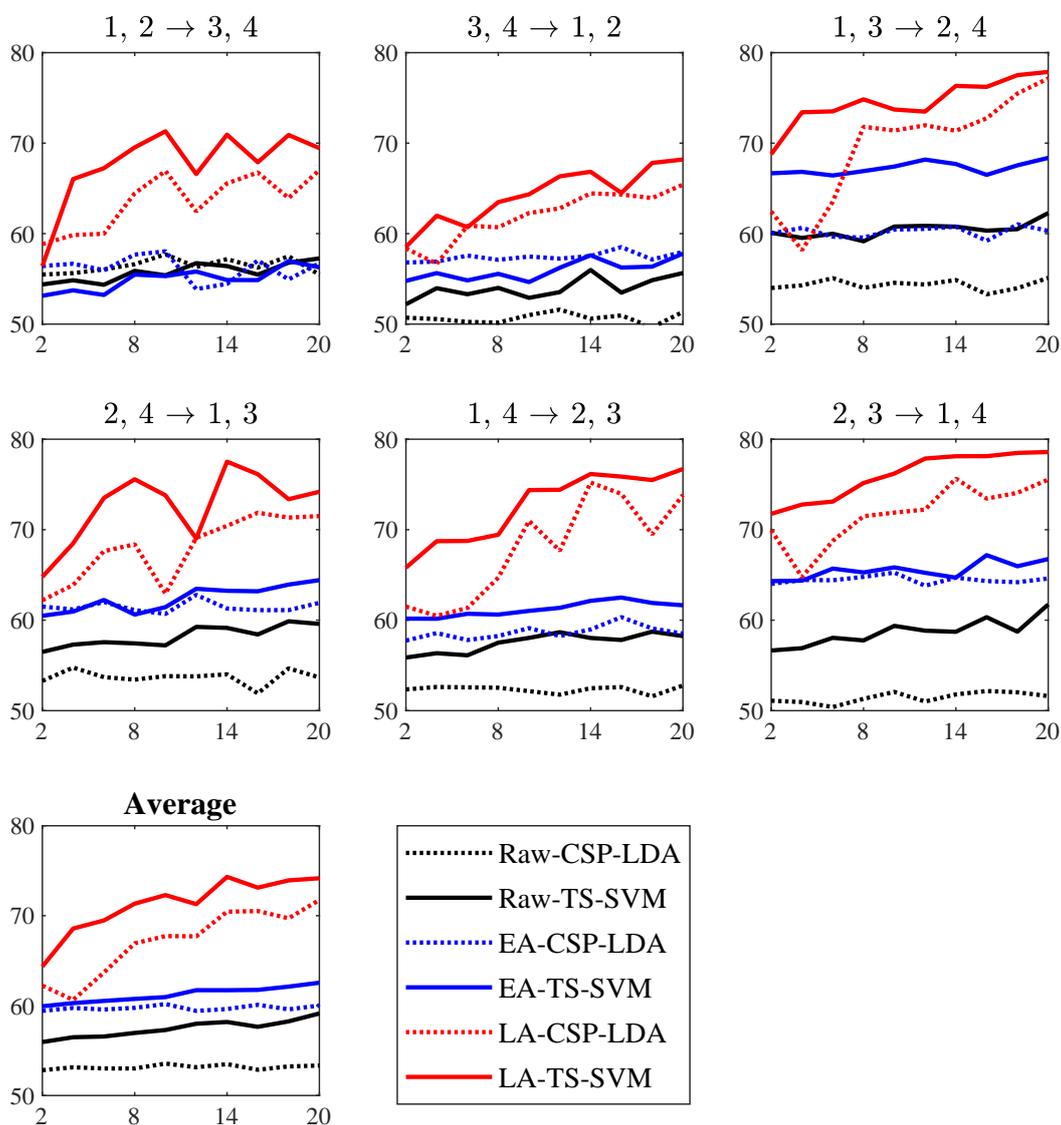


图 5-6 LA 用于不同特征提取和分类方法，在场景 II-a 下的分类正确率(%)。

从图5-6中可以观察到:

1. 对所有 6 个数据集组合，LA-CSP-LDA 的分类正确率总是高于 EA-CSP-LDA 和 Raw-CSP-LDA，LA-TS-SVM 的分类正确率总是高于 EA-TS-SVM 和

Raw-TS-SVM, 说明 LA 在两个分类框架下都是有效的, 且迁移效果比 EA 更好。

2. 比较图5-4与图5-6, 发现图5-6中 Raw-CSP-LDA 和 Raw-TS-SVM 的平均分类正确率比它们在图5-4 中的效果差, 这点很容易理解, 因为图5-6代表了类别空间差异更大的场景。但是图5-6中 LA-CSP-LDA 和 LA-TS-SVM 的效果与它们在图5-4中的表现却差不多, 说明 LA 能够很好地处理类别空间差异大的场景。

前面的实验发现 LA 的效果会随着目标用户已标注样本数量的增加而提升, 所以这里考虑当我们只能够标注很少量的目标用户样本时, LA 是否仍然有效呢? 为此, 我们将目标用户标注样本数量降至最低, 即每个类别只标注一个样本, 然后仍然比较上述算法。实验结果如表格5.5.4.3 所示, 其中每个场景下的最高分类正确率已经加粗显示。可以观察到, 即使目标用户标注样本数量很少, 基于 LA 的算法仍然在两个分类框架下都取得了最好的分类效果。

表 5.5 场景 II-a: 当目标域每个类别只有一个标注样本时的结果比较, 表中数值是在 9 个目标用户上的平均分类正确率(%)。

Experiment	Approach	Raw	EA	LA
1, 2 → 3, 4	CSP-LDA	55.48	56.42	<b>58.84</b>
	TS-SVM	54.38	53.13	<b>56.42</b>
3, 4 → 1, 2	CSP-LDA	50.70	56.81	<b>58.37</b>
	TS-SVM	52.19	54.77	<b>58.53</b>
1, 3 → 2, 4	CSP-LDA	53.99	60.02	<b>62.44</b>
	TS-SVM	60.09	66.67	<b>68.78</b>
2, 4 → 1, 3	CSP-LDA	53.29	61.50	<b>62.21</b>
	TS-SVM	56.49	60.49	<b>64.79</b>
1, 4 → 2, 3	CSP-LDA	52.35	57.75	<b>61.50</b>
	TS-SVM	55.87	60.17	<b>65.81</b>
2, 3 → 1, 4	CSP-LDA	51.10	64.01	<b>69.95</b>
	TS-SVM	56.65	64.32	<b>71.75</b>

接下来采用同样的统计检验方法对图5-6的结果进行显著性检验, 此时每个算法进行了  $6 \times 9 = 54$  次实验, 所以我们在这 54 个 AUC 值上进行配对  $t$  检验, 结果如表格5.6 所示, 其中有显著性的结果已经用粗体显示。可以观察到 LA-CSP-LDA 显著优于 EA-CSP-LDA, LA-TS-SVM 显著优于 EA-TS-SVM。验证了当类别空间完全不一致时, LA 仍然可以作为一个有效的预处理步骤, 然后用于不同的特征提取和分类算法。

表 5.6 图5-6中算法在配对  $t$  检验下的  $p$  值,  $p < 0.05$  表示结果有显著性。

	EA-CSP-LDA	EA-TS-SVM
LA-CSP-LDA	<b>0.0000</b>	
LA-TS-SVM		<b>0.0000</b>

#### 5.5.4.4 场景 II-b

场景 II-b 研究在多分类问题下, 源用户与目标用户的类别空间完全不一致的情形。如5.5.1节所总结的, 本节共考虑 12 种与场景 I-b 相同的数据集组合, 但使用不同的类别匹配方式。

理想情况下, 如果数据集 MI 2 有 6 个或更多的标签, 我们可以研究 LA 在数据集组合 “1, 2, 3  $\rightarrow$  4, 5, 6” 这样情形下的效果。但是现有数据集最多只有 4 个类别, 所以我们只能使用错位匹配来模拟类别空间完全不一致的情形。比如对数据集组合 “1, 2, 3  $\rightarrow$  1, 2, 4”, 我们故意不将相同的类别配对, 而是将源用户类别 “1” 与目标用户类别 “2” 配对, 源用户类别 “2” 与目标用户类别 “1” 配对, 再将源用户类别 “3” 与目标用户类别 “4” 配对。这种错位匹配在现实中的一个潜在的应用场景是我们知道源域哪些样本属于同一个类别, 但是不知道类别的标签。

与场景 I-b 一样, 本节在执行  $k$ -medoids 聚类时将参数设为  $k \in \{3, 6, \dots, 30\}$ , 并且使用 “一对多” 的方式将 CSP 从二分类扩展到多分类<sup>[55]</sup>。然后我们继续在两种分类框架下比较 LA 与 Raw, EA。实验结果如图5-7 所示, 其中前 12 个子图分别代表在 12 个数据集组合上的实验结果 (每个子图都是 9 个目标用户的平均结果), 子图的标题即为数据集组合的方式, 最后一个子图是 12 个数据集组合的平均结果。在每一个子图中, 横轴代表  $k$ -medoids 聚类的参数  $k$  (即目标用户已标注样本数目), 纵轴代表分类正确率(%). 从图中可以观察到:

1. 综合来看, 使用 EA 对齐的算法 (EA-CSP-LDA 和 EA-TS-SVM) 的结果与没有使用任何对齐的算法 (Raw-CSP-LDA 和 Raw-TS-SVM) 的结果差不多, 说明在类别错位匹配的场景下, EA 对齐没有任何提升算法效果的作用。
2. 对总共 6 个数据集组合, LA-CSP-LDA 的分类正确率总是高于 EA-CSP-LDA 和 Raw-CSP-LDA, LA-TS-SVM 的分类正确率也总是高于 EA-TS-SVM 和 Raw-TS-SVM, 说明 LA 在类别错位匹配的场景下仍然有很好的迁移效果。
3. 对比图5-5, 虽然使用了相同的数据集组合, 但是在类别错位匹配的设置下, 那些没有 LA 的算法 (EA-CSP-LDA, Raw-CSP-LDA, EA-TS-SVM 和 Raw-TS-SVM) 的效果都下降很多, 分类正确率接近随机。但是 LA-CSP-LDA 和 LA-TS-SVM 受影响较小, 依然表现出了很好的迁移效果, 说明 LA 能够很好地解决类别空间的各种差异。

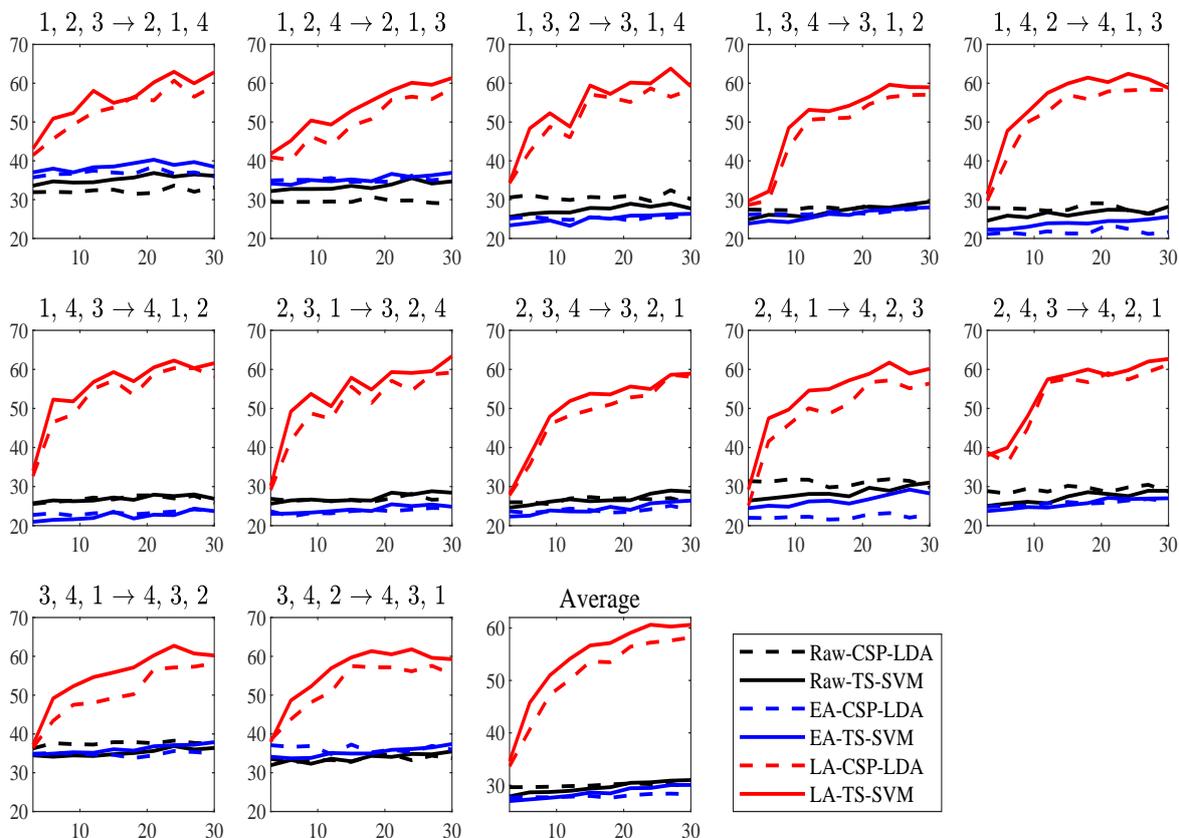


图 5-7 LA 用于不同特征提取和分类方法，在场景 II-b 下的分类正确率(%)。

接下来仍然对图5-7的结果进行显著性检验，此时每个算法进行了  $12 \times 9 = 108$  次实验，所以我们在这 108 个 AUC 值上进行配对  $t$  检验，结果如表格5.7 所示，其中有显著性的结果已经用粗体显示。可以观察到 LA-CSP-LDA 显著优于 EA-CSP-LDA，LA-TS-SVM 显著优于 EA-TS-SVM。验证了当类别错位匹配时，LA 仍然能够在不同特征提取和分类算法的框架下取得最好的效果。

表 5.7 图5-7中算法在配对  $t$  检验下的  $p$  值， $p < 0.05$  表示结果有显著性。

	EA-CSP-LDA	EA-TS-SVM
LA-CSP-LDA	<b>0.0000</b>	
LA-TS-SVM		<b>0.0000</b>

#### 5.5.4.5 场景 III

场景 III 研究特征空间与类别空间都不一致的情形，如5.5.1节所介绍的，本节使用数据集 MI 1 作为源数据集，MI 2 的子数据集“3, 4”作为目标数据集。这种组合代表了一个复杂的场景：

1. MI 1 包含 59 个通道的数据，而 MI 2 包含 22 个通道的数据，所以它们的特征空间不一致；
2. MI 1 的不同用户可能有不同的类别空间，其用户 1 和用户 6 的类别空间是{左手，右手}，其他用户类别空间是{左手，脚}；而 MI 2 的子数据集“3, 4”的类别空间是{脚，舌}，所以源用户与目标用户的类别空间有可能部分不一样，也有可能完全不一样。

每次我们从目标数据集“3, 4”选择一个用户作为目标用户，源数据集 MI 1 的所有 7 个用户作为源用户。然后从源数据集中挑选出 22 个与目标用户最接近的通道，再比较 LA 与其它方法分别在两种分类框架下的效果，实验结果如图 5-8 所示。

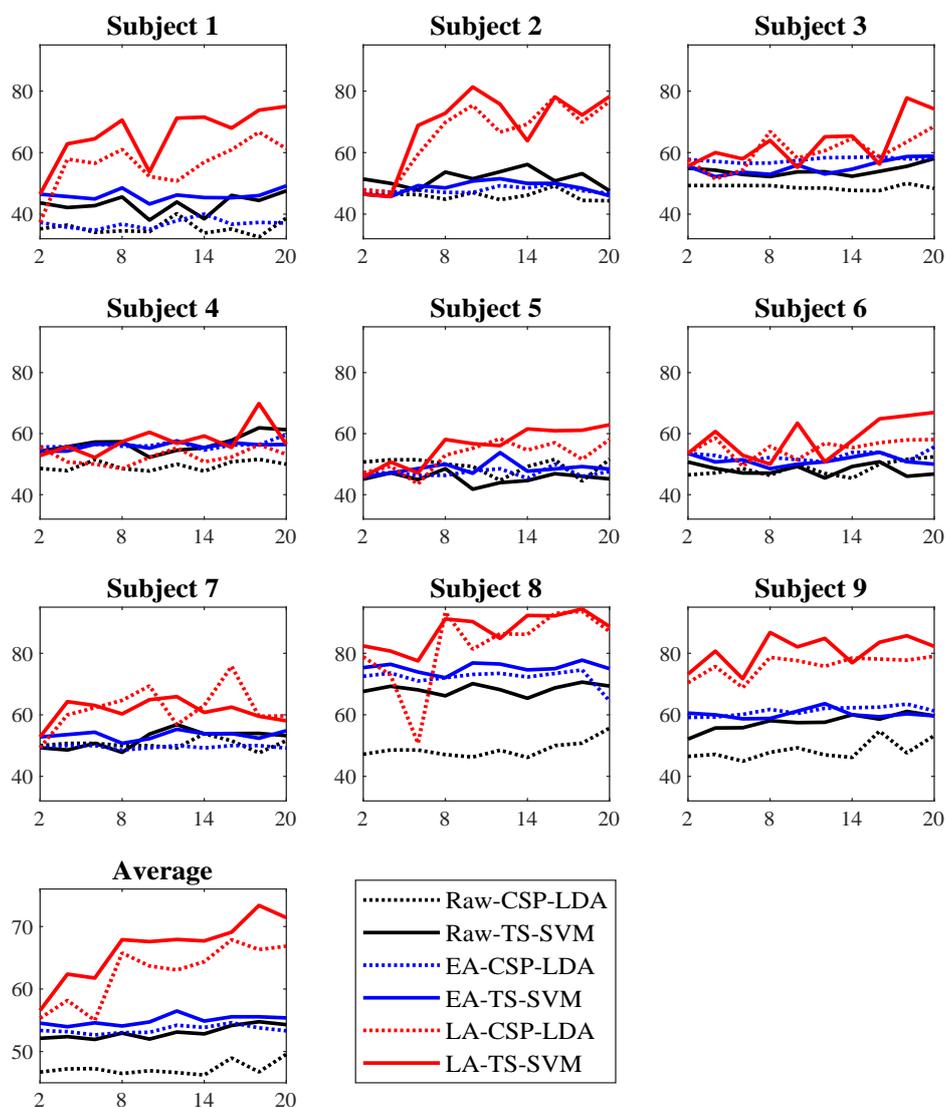


图 5-8 LA 用于不同特征提取和分类方法，在场景 III 下的分类正确率(%)。

在图5-8中，前9个子图分别代表在9个目标用户结果（不再是一种数据集组合的平均结果），最后一个子图是9个目标用户的平均结果。在每一个子图中，横轴代表 *k-medoids* 聚类的参数  $k$ （即目标用户已标注样本数目），纵轴代表分类正确率(%)，从图中可以观察到：

1. 对所有9个用户，EA-CSP-LDA在其中6个用户上的效果优于Raw-CSP-LDA，EA-TS-SVM在其中5个用户上的效果优于Raw-CSP-LDA，说明EA只能取得有限的迁移效果。
2. 对所有9个用户，LA-CSP-LDA的效果总是优于Raw-CSP-LDA，且在其中8个用户上优于EA-CSP-LDA；LA-TS-SVM在所有9个用户上都优于Raw-CSP-LDA和EA-CSP-LDA。这说明即使在一种很复杂的条件下，LA仍然取得了最好且稳定的迁移效果。

接下来仍然对图5-8的结果进行显著性检验，此时每个算法进行了9次实验，所以我们在这9个AUC值上进行配对  $t$  检验，结果如表格5.8所示，其中有显著性的结果已经用粗体显示。可以观察到LA-CSP-LDA显著优于EA-CSP-LDA，LA-TS-SVM显著优于EA-TS-SVM。这意味着即问题1再次得到了正面验证：LA对于不同特征提取和分类算法都能表现出好的迁移效果。

表 5.8 图5-8中算法在配对  $t$  检验下的  $p$  值， $p < 0.05$  表示结果有显著性。

	EA-CSP-LDA	EA-TS-SVM
LA-CSP-LDA	<b>0.0082</b>	
LA-TS-SVM		<b>0.0006</b>

### 5.5.5 LA 与其它迁移方法的结合

接下来我们在同样的场景下研究问题2：LA是否可以作为一种预处理步骤与其它迁移学习方法结合使用，并进一步提高迁移效果？

为此，我们在分类阶段比较以下几种算法：

1. BL：基准方法，即直接使用 SVM 分类器；
2. JDA：先用迁移学习方法 JDA<sup>[71]</sup> 进行数据迁移，再使用 SVM 分类；
3. JGSA：先用迁移学习方法 JGSA<sup>[88]</sup> 进行数据迁移，再使用 SVM 分类；
4. MEDA：使用迁移学习方法 MEDA<sup>[84]</sup> 进行数据迁移并分类。

上述四种方法都要求输入特征是向量，所以我们先求样本的协方差矩阵，然后计算切空间向量作为输入。把这四种分类方法与对齐阶段的 Raw, EA 和 LA 结合，则共有 12 种算法需要比较，目的是验证两个问题：

1. 只使用 LA 的迁移效果是否优于只使用 JDA, JGSA, 或者 MEDA 的迁移效果?
2. 同时使用 LA 和其它迁移方法的效果是否优于只使用 LA 的迁移效果?

### 5.5.5.1 场景 I-a

场景 I-a 研究在二分类问题下, 源用户与目标用户的类别空间部分不一致的情形, 本节使用与5.5.4.1节相同的数据集组合与参数设置, 然后比较上文提到的 12 种算法, 实验结果如图5-9所示。

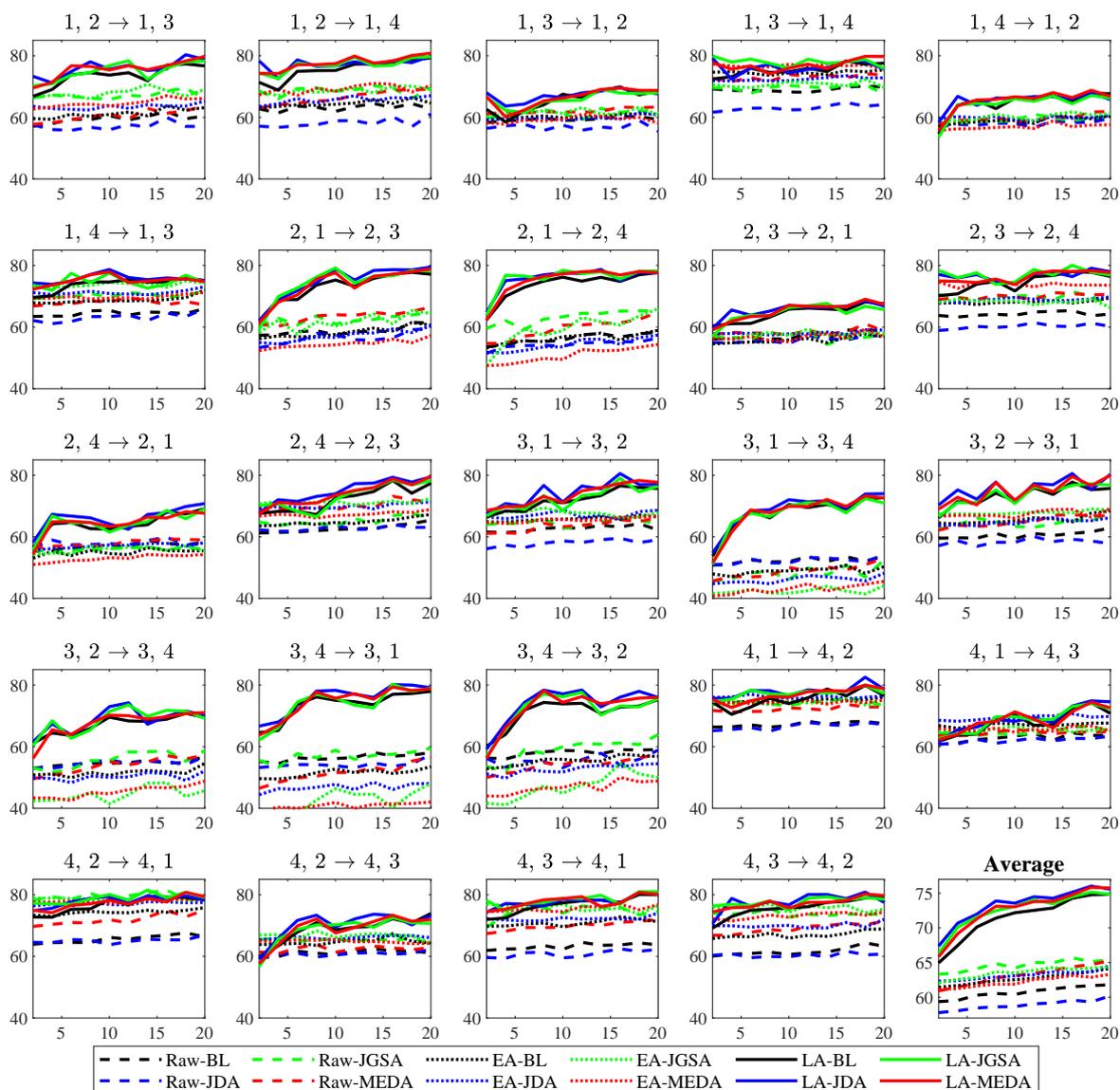


图 5-9 LA 与其它迁移学习方法结合, 在场景 I-a 下的分类正确率(%)。

在图5-9中，前 24 个子图分别代表在 24 个数据集组合上的实验结果（每个子图都是 9 个目标用户的平均结果），子图的标题即为数据集组合的方式，最后一个子图是 24 个数据集组合的平均结果。在每一个子图中，横轴代表 *k-medoids* 聚类的参数  $k$ （即目标用户已标注样本数目），纵轴代表分类正确率(%)，从图中可以观察到：

1. 当  $k$  较大时，LA-BL 的分类正确率总是高于 Raw-BL 和 EA-BL；LA-JDA 的分类正确率总是高于 Raw-JDA 和 EA-JDA；LA-JGSA 的分类正确率总是高于 Raw-JGSA 和 EA-JGSA；LA-MEDA 的分类正确率总是高于 Raw-MEDA 和 EA-MEDA。这说明无论是否使用额外的迁移学习方法，以及使用哪种迁移学习方法，在对齐阶段使用 LA 总能够提升分类效果。
2. 对总共 24 种数据集组合，LA-BL 在所有 24 个组合上的效果都优于 Raw-JDA 和 Raw-MEDA，在其中 23 个组合上的效果优于 Raw-JGSA。这说明当源用户与目标用户类别空间不一致时，LA 的效果比其它迁移学习方法的效果好很多。
3. 对总共 24 种数据集组合，LA-JDA 和 LA-MEDA 在所有组合上的分类正确率都高于 LA-BL，LA-JGSA 在其中 23 个组合上的分类正确率高于 LA-BL。这说明 LA 能够很好地与其它迁移学习方法兼容，且结合使用时，能够进一步提升算法的整体效果。

接下来采用配对  $t$  检验对图5-9的结果进行显著性检验，结果如表格5.9 所示，其中有显著性的结果已经用粗体显示。只使用 LA 的方法（LA-BL）显著优于不使用 LA 的方法（EA-JDA, EA-JGSA, EA-MEDA），而结合 LA 和其它迁移学习方法的算法（LA-JDA, LA-JGSA, LA-MEDA）又显著优于 LA-BL。

表 5.9 图5-9中算法在配对  $t$  检验下的  $p$  值， $p < 0.05$  表示结果有显著性。

	LA-BL	EA-JDA	EA-JGSA	EA-MEDA
LA-BL		<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>
LA-JDA	<b>0.0000</b>	<b>0.0000</b>		
LA-JGSA	<b>0.0001</b>		<b>0.0000</b>	
LA-MEDA	<b>0.0000</b>			<b>0.0000</b>

### 5.5.5.2 场景 I-b

场景 I-b 研究在多分类问题下，源用户与目标用户的类别空间部分不一致的情形，本节使用与5.5.4.2节相同的数据集组合与参数设置。实验结果如图5-9所示，其中前 12 个子图分别代表在 12 个数据集组合上的实验结果（每个子图都是 9 个目标用户的平均结果），子图的标题即为数据集组合的方式，最后一个子图是 12 个数据集组合的平均结果。在每一个子图中，横轴代表 *k-medoids* 聚类的参数  $k$ （即目标用

户已标注样本数目), 纵轴代表分类正确率(%). 从图中可以观察到:

1. 当  $k$  较大时, LA-BL 的分类正确率总是高于 Raw-BL 和 EA-BL; LA-JDA 的分类正确率总是高于 Raw-JDA 和 EA-JDA; LA-JGSA 的分类正确率总是高于 Raw-JGSA 和 EA-JGSA; LA-MEDA 的分类正确率总是高于 Raw-MEDA 和 EA-MEDA。这说明无论是否使用额外的迁移学习方法, 以及使用哪种迁移学习方法, 在对齐阶段使用 LA 总能够提升分类效果。
2. 对于总共 12 种数据集组合, LA-BL 在所有 12 个组合上的效果都优于 Raw-JDA, JGSA 和 Raw-MEDA, 说明当源用户与目标用户类别空间不一致时, LA 的效果比其它迁移学习方法的效果好很多。
3. 平均来看, LA-JDA, LA-JGSA 和 LA-MEDA 的分类正确率高于 LA-BL, 说明 LA 能够很好地与其它迁移学习方法兼容, 并进一步提升算法的效果。

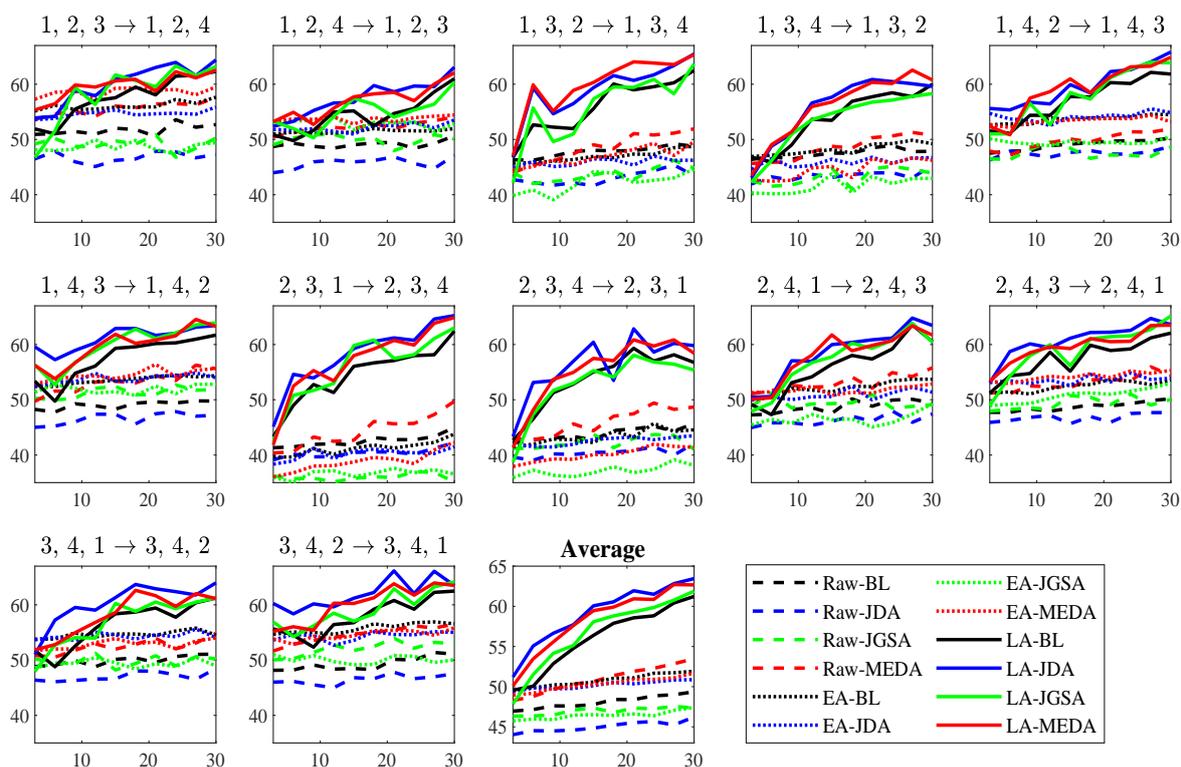


图 5-10 LA 与其它迁移学习方法结合, 在场景 I-b 下的分类正确率(%).

接下来对图5-10的结果进行显著性检验, 结果如表格5.10 所示, 其中有显著性的结果已经用粗体显示。结果显示只使用 LA 的方法 (LA-BL) 显著优于不使用 LA 的方法 (EA-JDA, EA-JGSA, EA-MEDA), 而结合 LA 和其它迁移学习方法的算法 (LA-JDA, LA-JGSA, LA-MEDA) 又显著优于 LA-BL, 再次验证了 LA 与其它迁移方法的兼容性。

表 5.10 图5-10中算法在配对  $t$  检验下的  $p$  值,  $p < 0.05$  表示结果有显著性。

	LA-BL	EA-JDA	EA-JGSA	EA-MEDA
LA-BL		<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>
LA-JDA	<b>0.0000</b>	<b>0.0000</b>		
LA-JGSA	<b>0.0295</b>		<b>0.0000</b>	
LA-MEDA	<b>0.0000</b>			<b>0.0000</b>

### 5.5.5.3 场景II-a

场景 II-a 研究在二分类问题下, 源用户与目标用户的类别空间完全不一致的情形。本节使用与5.5.4.3节相同的数据集组合与参数设置, 比较前文提到的 12 种算法。实验结果如图5-11所示, 其中前 6 个子图分别代表在 6 个数据集组合上的实验结果 (每个子图都是 9 个目标用户的平均结果), 子图的标题即为数据集组合的方式, 最后一个子图是 6 个数据集组合的平均结果。在每一个子图中, 横轴代表  $k$ -medoids 聚类的参数  $k$  (即目标用户已标注样本数目), 纵轴代表分类正确率(%). 从图中可以观察到:

1. 对于总共 6 种数据集组合, LA-BL 的分类正确率总是高于 Raw-BL 和 EA-BL; LA-JDA 的分类正确率总是高于 Raw-JDA 和 EA-JDA; LA-JGSA 的分类正确率总是高于 Raw-JGSA 和 EA-JGSA; LA-MEDA 的分类正确率总是高于 Raw-MEDA 和 EA-MEDA。这说明无论是否使用额外的迁移学习方法, 以及使用哪种迁移学习方法, 在对齐阶段使用 LA 总能够提升分类效果。
2. 对于总共 6 种数据集组合, LA-BL 在所有 6 个组合上的效果都优于 Raw-JDA, Raw-JGSA 和 Raw-MEDA, 说明当源用户与目标用户类别空间不一致时, LA 的效果比其它迁移学习方法的效果好很多。
3. 平均来看, LA-JDA, LA-JGSA 和 LA-MEDA 分类正确率高于 LA-BL, 说明 LA 能够很好地与其它迁移学习方法兼容, 并进一步提升算法的效果。
4. 与图5-9 (类别空间部分不一致, 二分类) 的结果对比, 其它迁移学习算法 (JDA, JGSA, MEDA) 在类别差异变大时效果下降很多; 但与 LA 结合后 (LA-JDA, LA-JGSA 和 LA-MEDA), 算法的效果并不受类别差异增大的影响。这说明 LA 能够作为预处理步骤很好地处理类别差异。

接下来对图5-11的结果进行显著性检验, 结果如表格5.11 所示, 其中有显著性的结果已经用粗体显示。结果表明只使用 LA 的方法 (LA-BL) 显著优于不使用 LA 的方法 (EA-JDA, EA-JGSA, EA-MEDA), 而结合 LA 和其它迁移学习方法的算法 (LA-JDA, LA-JGSA, LA-MEDA) 又显著优于 LA-BL。这证明 LA 能够与其它迁移方法结合使用, 并进一步提升算法效果。

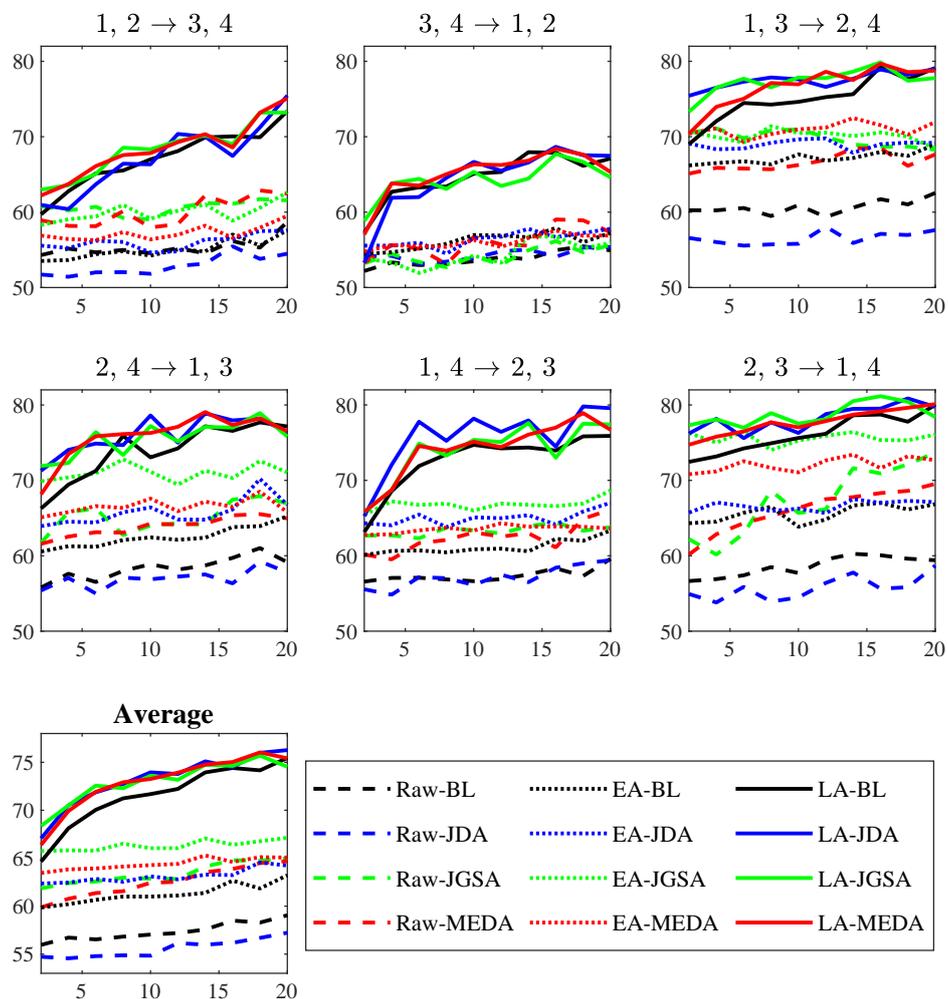


图 5-11 LA 与其它迁移学习方法结合，在场景 II-a 下的分类正确率(%)。

表 5.11 图5-11中算法在配对  $t$  检验下的  $p$  值， $p < 0.05$  表示结果有显著性。

	LA-BL	EA-JDA	EA-JGSA	EA-MEDA
LA-BL		<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>
LA-JDA	<b>0.0109</b>	<b>0.0000</b>		
LA-JGSA	<b>0.0116</b>		<b>0.0000</b>	
LA-MEDA	<b>0.0009</b>			<b>0.0000</b>

5.5.5.4 场景 II-b

场景 II-b 研究在多分类问题下，源用户与目标用户的类别空间完全不一致的情形。本节使用与5.5.4.4节相同的数据与设置，比较前文提到的 12 种算法，结果如图5-12所示，其中前 12 个子图分别代表在 12 个数据集组合上的实验结果（每个子图都是 9 个目标用户的平均结果），子图的标题即为数据集组合的方式，最后一个子图是 12 个数据集组合的平均结果。在每一个子图中，横轴代表  $k$ -medoids 聚类的参数  $k$ （即目标用户已标注样本数目），纵轴代表分类正确率(%)。

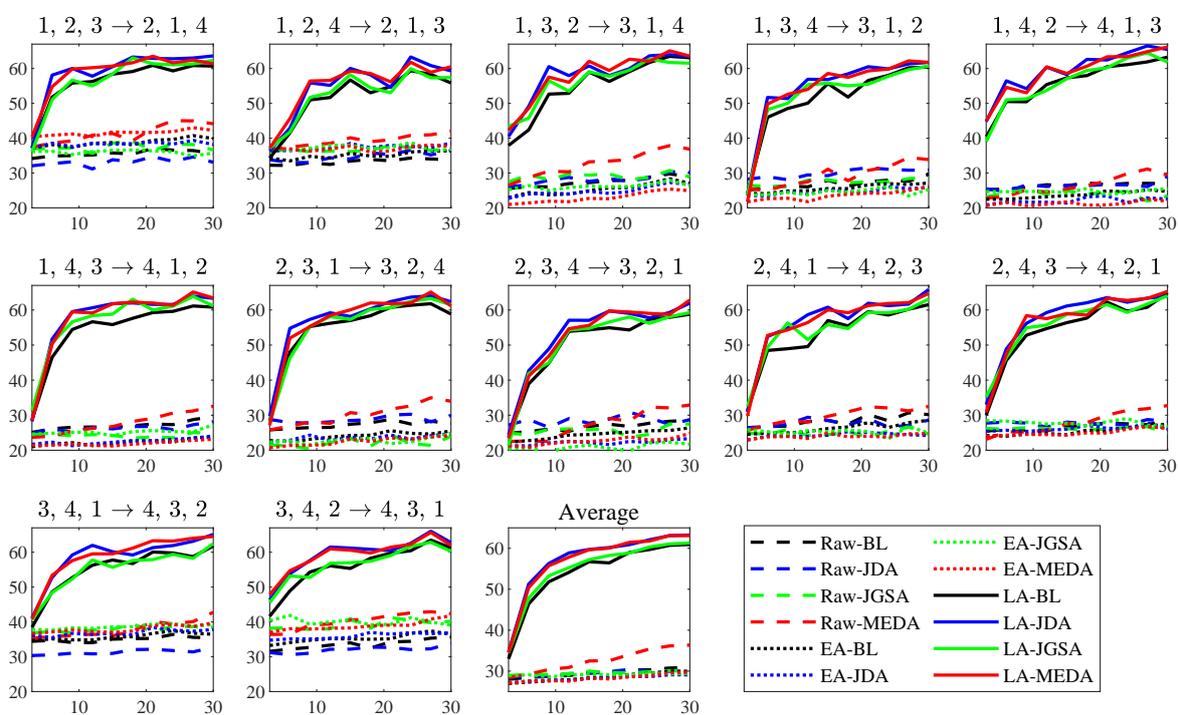


图 5-12 LA 与其它迁移学习方法结合，在场景 II-b 下的分类正确率(%)。

从图5-12中可以观察到:

1. 对总共 12 种数据集组合，LA-BL 的分类正确率总是高于 Raw-BL 和 EA-BL；LA-JDA 的分类正确率总是高于 Raw-JDA 和 EA-JDA；LA-JGSA 的分类正确率总是高于 Raw-JGSA 和 EA-JGSA；LA-MEDA 的分类正确率总是高于 Raw-MEDA 和 EA-MEDA。这说明无论是否使用额外的迁移学习方法，以及使用哪种迁移学习方法，在对齐阶段使用 LA 总能够提升分类效果。
2. 对总共 12 种数据集组合，LA-BL 在所有 12 个组合上的效果都优于 Raw-JDA, JGSA 和 Raw-MEDA，说明当源用户与目标用户类别空间不一致时，LA 的效果比其它迁移学习方法的效果好很多。

3. 平均来看, LA-JDA, LA-JGSA和 LA-MEDA 分类正确率高于 LA-BL, 说明 LA 能够很好地与其它迁移学习方法兼容, 并进一步提升算法的效果。
4. 对比图5-10, 虽然使用了相同的数据集组合, 但是在标签错位匹配的设置下, 那些没有 LA 的算法效果都下降很多。但是使用了 LA 的算法依然表现出了很好的迁移效果, 特别是当  $k$  较大时, 说明 LA 能够很好的解决类别空间的各种差异。

接下来对图 5-12的结果进行显著性检验, 结果如表格5.12 所示, 其中有显著性的结果已经用粗体显示。只使用 LA 的方法 (LA-BL) 显著优于不使用 LA 的方法 (EA-JDA, EA-JGSA, EA-MEDA), 而结合 LA 和其它迁移学习方法的算法 (LA-JDA, LA-JGSA, LA-MEDA) 又显著优于 LA-BL。

表 5.12 图5-12中算法在配对  $t$  检验下的  $p$  值,  $p < 0.05$  表示结果有显著性。

	LA-BL	EA-JDA	EA-JGSA	EA-MEDA
LA-BL		<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>
LA-JDA	<b>0.0000</b>			
LA-JGSA	<b>0.0009</b>			
LA-MEDA	<b>0.0000</b>			

### 5.5.5.5 场景 III

场景 III 研究源用户与目标用户的特征空间和类别空间都不一致的情形。与5.5.4.5节一样, 本节使用数据集 MI 1 作为源数据集, MI 2 的子数据集 “3, 4” 作为目标数据集。

然后比较 LA 与其它迁移学习算法, 以及 LA 与其它迁移学习方法结合的效果。每次实验从目标数据集选取一个用户作为目标用户, 源数据集 MI 1 的所有用户为源用户 (MI 1 的用户与 MI 2 不一样), 实验结果如图5-13 所示, 其中前 9 个子图分别代表 9 个目标用户的实验结果, 最后一个子图是 9 个目标用户的平均结果。在每一个子图中, 横轴代表  $k$ -medoids 聚类的参数  $k$  (即目标用户已标注样本数目), 纵轴代表分类正确率(%). 总的来看, 使用了 LA 的算法 (图中实现) 效果要好于没有使用 LA 的算法 (图中虚线)。

图5-13 对应的显著检验结果如表格5.13 所示, 可以看到 LA 在场景 III 中仍然优于其它迁移学习方法 (JDA, JGSA, MEDA), 但是 LA 与它们的结合 (LA-JDA, LA-JGSA, LA-MEDA) 不再显著优于 LA-BL。这说明在特征空间与类别空间都不同时, LA 与其它迁移学习方法的结合效果可能会受到影响, 但 LA 也不会对其它方法产生负作用。

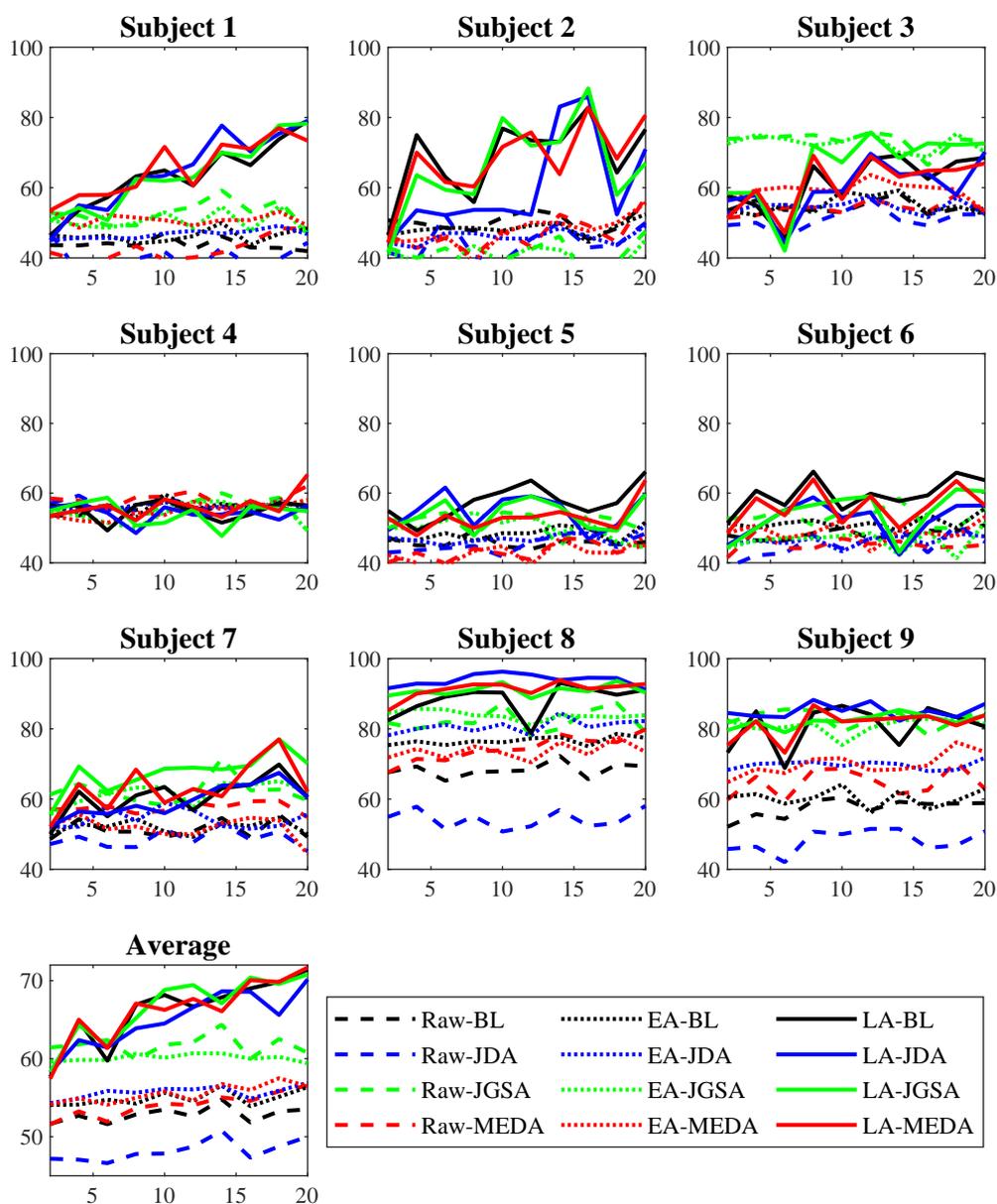


图 5-13 LA 与其它迁移学习方法结合，在场景 III 下的分类正确率(%)。

表 5.13 图5-13中算法在配对  $t$  检验下的  $p$  值， $p < 0.05$  表示结果有显著性。

	LA-BL	EA-JDA	EA-JGSA	EA-MEDA
LA-BL		<b>0.0017</b>	0.1335	<b>0.0011</b>
LA-JDA	0.3733	<b>0.0026</b>		
LA-JGSA	0.8449		0.0691	
LA-MEDA	0.9777			<b>0.0012</b>

### 5.5.6 LA的时间复杂度分析

LA在整体算法流程中可以当作一个独立的模块，或者预处理步骤，其计算的时间复杂度为 $O(N_T^2)$ ，其中 $N_T$ 为目标用户的样本数目。

本文也研究了LA在实际计算中对于算法耗时的影响，由于在不同场景，不同算法中LA的执行过程是一样的，所以这里只以场景 III 的 Raw-TS-SVM 和 LA-TS-SVM 两个算法的耗时为例，以作展示。执行算法的平台为一台联想 Thinkpad 笔记本，配置为 Intel Core i5-6200U CPU@2.30GHz，4GB 内存，以及 190 GB 固态硬盘，运行 64 位 Windows 10 系统以及 Matlab 2018b。实验结果如表格 5.14所示，展示的是两个算法在 9 个目标用户的平均以及标准差。从表中结果可以看出，LA 对算法实际耗时的影响很小。

表 5.14 场景 III 中算法 Raw-TS-SVM 和 LA-TS-SVM 的计算耗时（秒）。

	平均	标准差
Raw-TS-SVM	2.1963	0.1492
LA-TS-SVM	2.3669	0.3469

## 5.6 小结

本章提出一种针对异构类别空间的迁移学习方法 LA。对于特征空间相同，类别空间不同的情形，我们直接使用 LA 将源用户数据对齐到目标用户；对于特征空间与类别空间都不同的情形，我们先挑选出源用户与目标用户共有或位置接近的通道，再使用 LA 将源用户数据对齐到目标用户。

通过多种数据集组合的模拟，本章在不同异构类别空间场景下验证了 LA 的有效性，以及 LA 的几个很好的优点：（1）LA 只需要对很少量的目标用户样本进行标注；（2）经 LA 对齐后，后续分类可以使用不同的特征提取和分类算法；（3）LA 可以与其它迁移学习方法兼容，经 LA 对齐后，再使用其它迁移学习方法能够进一步提升算法的整体效果。

## 6 总结与展望

### 6.1 本文总结

本文的研究工作主要源自青年千人项目“脑机接口和情感计算中的机器学习研究”以及相关项目。具体来讲，本文研究脑机接口中的迁移学习，目标是从算法角度来减少新用户所需的校准时间，从而提高脑机接口的实用性。

本文章节之间采取的是“总-分-总”结构：第1章介绍问题的背景与意义，阐述为什么要研究脑机接口中的迁移学习；第2至5章则针对不同场景，从不同的角度提出了新的迁移学习方法；第6章总结全文内容，并展望未来的工作。

进一步地，第2至5章是本文提出的具体方法，也是论文的主体部分和创新所在，其章节之间的关联，或者说论文主体的架构可以从脑机接口和迁移学习两个角度来理解：

1) 从脑机接口的角度看，脑机接口中的机器学习方法分为两大框架：基于欧式空间的传统分类框架和基于黎曼几何的分类框架。本文第2章和第3章分别针对这两种分类框架提出了新的迁移学习算法或者优化方法；第4章和第5章则是针对EEG数据提出了新的迁移学习方法，数据进行迁移后可同时适用于两种分类框架。

2) 从迁移学习的角度看，迁移学习的场景可分为同构迁移学习和异构迁移学习。本文第2至4章研究的都是同构迁移学习；第5章研究的则是最具挑战性的异构类别空间问题。

取得的主要创新成果如下：

第2章的成果可分为两部分：(1) 提出了一种新的协方差矩阵迁移方法用于共同空间模式滤波器，该方法分别以EEG样本及其协方差矩阵的切空间向量为特征，基于最大均值差异对源用户的样本特征加权以减少其与目标用户特征之间的分布差异，然后将权重赋予源用户的样本协方差矩阵，并用来计算目标用户的共同空间模式滤波器。该方法采取了比文献中的方法更为细腻的操作，即将权重赋予的对象从用户的平均协方差矩阵下沉到单个样本的协方差矩阵。实验结果显示，该方法的迁移效果优于文献中的方法，能够有效减少新用户建立共同空间模式滤波器所需要的标注数据。(2) 针对共同空间模式滤波器之间的信息冗余问题，提出了一种新的共同空间模式目标函数，并在流形上使用梯度下降法求解。实验结果显示，新的目标函数能够获得相关性更小的滤波器，但不能获得更好的分类效果。

第3章针对基于黎曼几何的迁移学习算法提出了两种优化方法，其一是挑选通道，可以降低协方差矩阵的维度，从而提高有关协方差矩阵计算的效率与准确度；其二是基于源用户与目标用户的样本协方差矩阵之间的黎曼距离进行聚类，从而采样出位于目标用户高密度区域的源用户样本，可以在已经进行了协方差矩阵对齐的

情况下进一步减少协变量偏移。实验结果显示这两种优化方法都能提升脑机接口领域当前最优迁移学习方法的效果，且两种优化方法结合使用能进一步提升迁移效果。

第 4 章提出一种新的无监督数据对齐方法。对任意一个用户，使用其平均协方差矩阵作为参考矩阵对 EEG 样本进行转换，使得转换后的用户平均协方差矩阵为单位矩阵，且在转换过程中用户内部样本之间的距离不变。因此，若对不同用户分别执行上述转换，它们的样本协方差矩阵在分布上会变得更相似。由于样本的协方差矩阵是 EEG 数据的一种非常重要的特征和统计量，所以转换后的源用户数据可以用于目标用户模型的建立。此外，这种数据对齐方法直接作用于 EEG 样本，且不改变样本的维度结构，所以该方法可以作为一种预处理步骤，用于不同的特征提取和分类算法。实验结果显示，该方法在运动想象和 P300 数据集上都取得了比当前最先进算法更优秀的迁移效果。

第 5 章针对迁移学习中最具挑战性的异构类别空间问题，提出一种类别空间的对齐方法，只需要对目标用户的少量样本进行标注，就可以把源用户的一个类别对齐到目标用户的另一个不相关类别上。该方法有很多优秀的性质：（1）对目标用户标注数据的需求量小，最少一个类别只需要一个标注样本；（2）可以作为一种预处理步骤，用于不同特征提取和分类算法；（3）兼容性好，该方法可以与其它迁移学习方法结合使用，进一步提升迁移效果。最后在多种场景下的实验结果都证明了该方法的有效性。

本文的迁移学习方法都是针对脑机接口，特别是 EEG 数据的特点所提出的，其中第 2, 3 章针对的都是脑机接口中常用的算法，第 4, 5 章针对的是多通道 EEG 数据，要求每一个样本都是矩阵形式且其协方差矩阵或者增广协方差矩阵携有重要信息，所以这些方法目前还不能作为通用算法用于其它领域。此外，关于使用迁移学习方法的先决条件，整体来讲，条件是目标域的标注数据不足，且能够获取足够有质量的源域标注数据。特别地，本文第 2 至 4 章针对的是同构迁移学习场景，要求源用户与目标用户的特征空间和类别空间相同；第 5 章针对的则是异构类别空间场景，即源用户与目标用户的类别空间不同，且与第 4 章的无监督迁移方法不同，第 5 章要求能够对少量目标用户样本进行标注。

## 6.2 未来工作展望

本文针对脑机接口中的迁移学习方法开展研究，从算法和数据的角度都提出了新的迁移学习方法。虽然在理论、方法和应用方面取得了一些阶段性的成果，但迁移学习作为机器学习领域的前沿方向，特别是其在脑机接口这样一个特点鲜明的领域中的应用，仍然存在很多挑战。

基于全文的研究内容、方法和实验结果，本节对未来可能的研究方向作简要展

望:

1) 平衡源域与目标域的权重, 即当源域有大量的标注数据且目标域有少量的标注数据时, 有时需要根据实际情况给目标域的标注数据赋予不同的权重。比如本文第 2 章中, 需要以一定的权重因子来平衡目标域的平均协方差矩阵与源域的平均协方差矩阵。有的方法是将权重因子设为 0.5 (平均法), 效果可能不稳定; 有的方法采用交叉验证, 效果可能更好但会耗费更多的计算资源。所以如何快速地对权重因子进行动态调整是一个值得研究的问题。

2) 针对各种范式的脑机接口提出对应的迁移学习方法。各种范式的脑机接口建立自不同的神经学基础, 所以通常需要不同的算法来处理。比如本文第 4 章所提的数据对齐方法, 对 P300 诱发电位 (采用波形特征) 和运动想象 (采用时空特征) 的范式有显著效果, 但是对稳态视觉诱发电位 (采用频域特征) 范式则没有表现出好的效果。

3) 关于本文第 5 章所提出的异构类别空间对齐方法, 一个关键点在于衡量目标用户每个类别的平均协方差矩阵。目前本文是先使用聚类算法来选择少量最具代表性的目标用户样本进行标注, 然后根据标注样本计算各个类别的平均协方差矩阵, 那么是否存在更有效的方法呢?

4) 目前大多数的迁移学习研究都是针对某个问题或目标提出对应的迁移学习方法, 而关于迁移学习的基础性研究则相对较少。比如很多方法都基于某种距离度量来减少源域与目标域之间的分布差异, 那么哪种度量最能衡量域间的相似性呢? 如何鉴定源域的可迁移性, 从而避免负迁移? 迁移学习的能力也是有限的, 所以迁移学习的边界在哪里? 从实用的角度看, 目前已经出现了众多的迁移学习方法, 如何能够在统一的框架下评价它们?

## 致 谢

时光荏苒，恍然间博士生涯已接近尾声，回顾四年的博士生活，虽有遗憾，但更多的是收获。值此博士论文完成之际，颇多感慨，思绪万千，却不知如何下笔。静下心来，始终萦绕在心头的是老师的教导与恩情，同学的情谊与帮助，还有家人毫无保留的支持与关心。我心中的感激之情平时难以尽述，唯能在此重要时刻表达一二。

首先感谢我的恩师伍冬睿教授！从入学到毕业，伍老师始终关心我的成长，给予我悉心的指导与关怀。在生活里，伍老师总是尽最大努力给予我帮助，提供良好的科研环境，让我在学业压力下仍然安心、快乐地度过了四年，我对此感到非常庆幸。在学业上，伍老师引导我科研的方向，指导我的每一篇论文，教导我科研的态度与方法，才让我能够顺利完成学业，可以说这篇博士论文里处处有伍老师的痕迹。伍老师渊博的专业知识、严谨的治学态度和孜孜不倦的追求更是深深地影响了我，我也将在未来始终铭记教导，并期待能有继续向伍老师学习请教的机会。

伍老师不仅是我学业的导师，更是我人生的导师。曾经有一句电影台词风靡一时：“懂得了很多的大道理，但仍然过不好这一生”。在伍老师这里我明白了，重要的是践行而不是懂得。那些我从小学习的简单道理，比如勤奋、严谨、守时、可靠等等，在伍老师这里我都获得了更深的理解。伍老师的言传身教，让我收获了更多可贵的品质，受益匪浅。伍老师，我将在未来的工作中加倍努力，践行您做人做事的态度，也祝您未来的工作一帆风顺！

感谢脑机接口与机器学习实验室这个温暖的大家庭，和大家在一起的日子真的是快乐又难忘。感谢宋志康、崔雨琦、石振华、刘子涵、王阳几位同学，大家与我同为第一批进实验室的学生，在我最初开始一段新生活的日子里给予了很多的帮助，如今大家已经或者即将进入工作岗位，祝大家前途似锦。感谢谭显烽、张潇、郭陈凤、徐祎璠、赵昶铭、张稳、孟璐斌、蒋雪、彭睿旻等同学，你们在工作和生活里都给予我很多的配合和帮助，祝你们学业有成。还有即将正式进入实验室的各位新同学，相信你们未来会为实验室取得更大的成绩。

感谢黄剑老师在过去几年曾给予我的帮助，感谢胡若澜老师，博1601班的肖强、曹瑜等同学，以及其他所有帮助过我的老师和同学。

感谢本论文评审、答辩工作的各位专家、教授和老师们的辛勤劳动。

最后要感谢我的父母和家人，谢谢你们对我的支持和鼓励，你们的关爱是我前进的最大动力，我最大的愿望就是你们身体健康！

## 参考文献

- [1] Hochberg L R, Bacher D, Jarosiewicz B, et al. Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. *Nature*, 2012, 485(7398):372.
- [2] Edelman B, Meng J, Suma D, et al. Noninvasive neuroimaging enhances continuous neural tracking for robotic device control. *Science Robotics*, 2019, 4(31):eaaw6844.
- [3] Anumanchipalli G K, Chartier J, Chang E F. Speech synthesis from neural decoding of spoken sentences. *Nature*, 2019, 568(7753):493.
- [4] 刘小燮, 毕胜, 高小榕, 等. 基于运动想象的脑机交互康复训练新技术对脑卒中大脑可塑性影响. *中国康复医学杂志*, 2013, 28(2):97–102.
- [5] 郑宇新, 朱君明, 郑筱祥, 等. 脑机接口技术在神经修复中的应用. *国际神经病学神经外科学杂志*, 2014, 41(1):58–61.
- [6] 李明爱, 罗新勇, 崔燕, 等. 基于MI-BCI的上肢在线运动康复原型系统. *北京生物医学工程*, 2017, 36(3):273–278.
- [7] Capogrosso M, Milekovic T, Borton D, et al. A brain–spine interface alleviating gait deficits after spinal cord injury in primates. *Nature*, 2016, 539(7628):284.
- [8] 姜雷, 张海, 张岚, et al. 脑机接口研究之演化及教育应用趋势的知识图谱分析——基于1985-2018年SCI及SSCI期刊论文研究. *远程教育杂志*, 2018, (4):27–38.
- [9] Lin C T, Wu R C, Liang S F, et al. EEG-Based Drowsiness Estimation for Safety Driving Using Independent Component Analysis. *IEEE Trans. on Circuits and Systems*, 2005, 52:2726–2738.
- [10] Liang S F, Shaw F Z, Young C P, et al. A closed-loop brain computer interface for real-time seizure detection and control. in 2010 Annual Int'l Conf. of the IEEE Engineering in Medicine and Biology. IEEE, 2010, 4950–4953.
- [11] Sun F T, Morrell M J, Wharen R E. Responsive cortical stimulation for the treatment of epilepsy. *Neurotherapeutics*, 2008, 5(1):68–74.
- [12] Smith J R, Fountas K N, Murro A M, et al. Closed-loop stimulation in the control of focal epilepsy of insular origin. *Stereotactic and Functional Neurosurgery*, 2010, 88(5):281–287.
- [13] Anderson W S, Kossoff E H, Bergey G K, et al. Implantation of a responsive neurostimulator device in patients with refractory epilepsy. *Neurosurgical Focus*, 2008, 25(3):E12.
- [14] Odekerken V J, Laar T, Staal M J, et al. Subthalamic nucleus versus globus pallidus bilateral deep brain stimulation for advanced Parkinson's disease (NSTAPS study): a randomised controlled trial. *The Lancet Neurology*, 2013, 12(1):37–44.
- [15] Weaver F M, Follett K A, Stern M, et al. Randomized trial of deep brain stimulation for Parkinson disease: thirty-six-month outcomes. *Neurology*, 2012, 79(1):55–65.
- [16] Nicolas-Alonso L F, Gomez-Gil J. Brain computer interfaces, a review. *Sensors*, 2012, 12(2):1211–1279.

- [17] Fatourehchi M, Bashashati A, Ward R K, et al. EMG and EOG artifacts in brain computer interface systems: A survey. *Clinical neurophysiology*, 2007, 118(3):480–494.
- [18] Vijila C K S, Kanagasabapathy P, Johnson S, et al. Artifacts removal in EEG signal using adaptive neuro fuzzy inference system. in 2007 Int'l Conf. on Signal Processing, Communications and Networking, Chennai, India, February, 2007, 589–591.
- [19] Vigario R, Sarela J, Jousmiki V, et al. Independent component approach to the analysis of EEG and MEG recordings. *IEEE Trans. on Biomedical Engineering*, 2000, 47(5):589–593.
- [20] Regan D. Evoked potentials and evoked magnetic fields in science and medicine. *Human Brain Electrophysiology*, 1989, pages 59–61.
- [21] Wang Y, Wang R, Gao X, et al. A practical VEP-based brain-computer interface. *IEEE Trans. on Neural Systems and Rehabilitation Engineering*, 2006, 14(2):234–239.
- [22] Gao X, Xu D, Cheng M, et al. A BCI-based environmental controller for the motion-disabled. *IEEE Trans. on Neural Systems and Rehabilitation Engineering*, 2003, 11(2):137–140.
- [23] Bin G, Gao X, Wang Y, et al. VEP-based brain-computer interfaces: Time, frequency, and code modulations. *IEEE Computational Intelligence Magazine*, 2009, 4(4):22–26.
- [24] Lee P L, Hsieh J C, Wu C H, et al. Brain computer interface using flash onset and offset visual evoked potentials. *Clinical Neurophysiology*, 2008, 119(3):605–616.
- [25] Farwell L A, Donchin E. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and clinical Neurophysiology*, 1988, 70(6):510–523.
- [26] Polich J, Ellerson P C, Cohen J. P300, stimulus intensity, modality, and probability. *Int'l Journal of Psychophysiology*, 1996, 23(1-2):55–62.
- [27] Ravden D, Polich J. On P300 measurement stability: habituation, intra-trial block variation, and ultradian rhythms. *Biological psychology*, 1999, 51(1):59–76.
- [28] Tanaka K, Matsunaga K, Wang H O. Electroencephalogram-based control of an electric wheelchair. *IEEE Trans. on Robotics*, 2005, 21(4):762–766.
- [29] Mugler E M, Ruf C A, Halder S, et al. Design and implementation of a P300-based brain-computer interface for controlling an internet browser. *IEEE Trans. on Neural Systems and Rehabilitation Engineering*, 2010, 18(6):599–609.
- [30] Jeannerod M. Mental imagery in the motor context. *Neuropsychologia*, 1995, 33(11):1419–1432.
- [31] Pfurtscheller G, Neuper C, Flotzinger D, et al. EEG-based discrimination between imagination of right and left hand movement. *Electroencephalography and clinical Neurophysiology*, 1997, 103(6):642–651.
- [32] Blankertz B, Sannelli C, Halder S, et al. Neurophysiological predictor of SMR-based BCI performance. *Neuroimage*, 2010, 51(4):1303–1309.
- [33] Wolpaw J R, McFarland D J, Vaughan T M. Brain-computer interface research at the Wadsworth Center. *IEEE Trans. on Rehabilitation Engineering*, 2000, 8(2):222–226.

- [34] Blankertz B, Losch F, Krauledat M, et al. The Berlin Brain-Computer Interface: Accurate performance from first-session in BCI-naive subjects. *IEEE Trans. on Biomedical Engineering*, 2008, 55(10):2452–2462.
- [35] Pfurtscheller G, Neuper C, Muller G, et al. Graz-BCI: state of the art and clinical applications. *IEEE Trans. on Neural Systems and Rehabilitation Engineering*, 2003, 11(2):1–4.
- [36] Bai O, Rathi V, Lin P, et al. Prediction of human voluntary movement before it occurs. *Clinical Neurophysiology*, 2011, 122(2):364–372.
- [37] Lin C T, Wu R C, Jung T P, et al. Estimating Driving Performance Based on EEG Spectrum Analysis. *EURASIP Journal of Applied Signal Processing*, 2005, 19:3165–3174.
- [38] Lin C T, Chung I, Ko L W, et al. EEG-based assessment of driver cognitive responses in a dynamic virtual-reality driving environment. *IEEE Trans. on Biomedical Engineering*, 2007, 54(7):1349–1352.
- [39] Harmon-Jones E, Allen J J. Anger and frontal brain activity: EEG asymmetry consistent with approach motivation despite negative affective valence. *Journal of Personality and Social Psychology*, 1998, 74(5):1310.
- [40] Kupfer D J, Foster F G, Coble P, et al. The application of EEG sleep for the differential diagnosis of affective disorders. *The American Journal of Psychiatry*, 1978.
- [41] Gruzelier J H. EEG-neurofeedback for optimising performance. I: a review of cognitive and affective outcome in healthy participants. *Neuroscience & Biobehavioral Reviews*, 2014, 44:124–141.
- [42] Kattler H, DIJK D J, Borbely A A. Effect of unilateral somatosensory stimulation prior to sleep on the sleep EEG in humans. *Journal of Sleep Research*, 1994, 3(3):159–164.
- [43] Fell J, Röschke J, Mann K, et al. Discrimination of sleep stages: a comparison between spectral and nonlinear EEG measures. *Electroencephalography and Clinical Neurophysiology*, 1996, 98(5):401–410.
- [44] Šušmáková K. Human sleep and sleep EEG. *Measurement science review*, 2004, 4(2):59–74.
- [45] 周志华. 机器学习. 北京: 清华大学出版社, 2016.
- [46] Lawhern V J, Solon A J, Waytowich N R, et al. EEGNet: a compact convolutional neural network for EEG-based brain-computer interfaces. *Journal of neural engineering*, 2018, 15(5):056013.
- [47] Schirrneister R T, Springenberg J T, Fiederer L D J, et al. Deep learning with convolutional neural networks for EEG decoding and visualization. *Human brain mapping*, 2017, 38(11):5391–5420.
- [48] Lotte F, Bougrain L, Cichocki A, et al. A review of classification algorithms for EEG-based brain-computer interfaces: a 10 year update. *Journal of Neural Engineering*, 2018, 15(3):031005.
- [49] Nunez P L, Srinivasan R, Westdorp A F, et al. EEG coherency: I: statistics, reference electrode, volume conduction, Laplacians, cortical imaging, and interpretation at multiple scales. *Electroencephalography and clinical neurophysiology*, 1997, 103(5):499–515.
- [50] 吴小培, 周蚌艳, 张磊, 等. 脑-机接口中空域滤波技术现状与进展. *安徽大学学报自科版*, 2017, 41(2):14.

- [51] Koles Z J, Lazar M S, Zhou S Z. Spatial patterns underlying population differences in the background EEG. *Brain Topography*, 1990, 2(4):275–284.
- [52] Müller-Gerking J, Pfurtscheller G, Flyvbjerg H. Designing optimal spatial filters for single-trial EEG classification in a movement task. *Clinical Neurophysiology*, 1999, 110(5):787–798.
- [53] Blankertz B, Tomioka R, Lemm S, et al. Optimizing Spatial filters for Robust EEG Single-Trial Analysis. *IEEE Signal Processing Magazine*, 2008, 25(1):41–56.
- [54] Ramoser H, Muller-Gerking J, Pfurtscheller G. Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE Trans. on Rehabilitation Engineering*, 2000, 8(4):441–446.
- [55] Dornhege G, B. Blankertz G C, Muller K R. Boosting bit rates in non-invasive EEG single-trial classifications by feature combination and multi-class paradigms. *IEEE Trans. on Biomedical Engineering*, 2004, 51(6):993–1002.
- [56] Ang K K, Chin Z Y, Zhang H, et al. Filter Bank Common Spatial Pattern (FBCSP) in Brain-Computer Interface. in *Proc. IEEE World Congress on Computational Intelligence*, Hong Kong, June, 2008, 2390-2397.
- [57] Lotte F, Guan C. Regularizing Common Spatial Patterns to Improve BCI Designs: Unified Theory and New Algorithms. *IEEE Trans. on Biomedical Engineering*, 2011, 58(2):355–362.
- [58] Dalhoumi S, Dray G, Montmain J. Knowledge Transfer for Reducing Calibration Time in Brain-Computer Interfacing. in *Proc. 26th IEEE Int'l Conf. on Tools with Artificial Intelligence*, Limassol, Cyprus, November, 2014, 634-639.
- [59] Barachant A, Bonnet S, Congedo M, et al. Multiclass Brain-Computer Interface Classification by Riemannian Geometry. *IEEE Trans. on Biomedical Engineering*, 2012, 59(4):920–928.
- [60] Barachant A, Congedo M. A Plug & Play P300 BCI Using Information Geometry. *arXiv: 1409.0107*, 2014.
- [61] Yger F, Berar M, Lotte F. Riemannian approaches in brain-computer interfaces: a review. *IEEE Trans. on Neural Systems and Rehabilitation Engineering*, 2017, 25(10):1753–1762.
- [62] Wu D, Lawhern V J, Lance B J, et al. EEG-Based User Reaction Time Estimation Using Riemannian Geometry Features. *IEEE Trans. on Neural Systems and Rehabilitation Engineering*, 2017, 25(11):2157–2168.
- [63] Congedo M, Barachant A, Andreev A. A New Generation of Brain-Computer Interface Based on Riemannian Geometry. *arXiv: 1310.8115*, 2013.
- [64] Fletcher P T, Joshi S. Principal geodesic analysis on symmetric spaces: Statistics of diffusion tensors. *Lecture Notes in Computer Science*, 2004, 3117:87–98.
- [65] Pan S J, Yang Q. A survey on transfer learning. *IEEE Trans. on Knowledge and Data Engineering*, 2010, 22(10):1345–1359.
- [66] Day O, Khoshgoftaar T M. A survey on heterogeneous transfer learning. *Journal of Big Data*, 2017, 4(1):29.
- [67] Zadrozny B. Learning and evaluating classifiers under sample selection bias. in *Proc. 21st Int'l conf. on Machine learning*, Banff, Canada, July, 2004, ACM, 114.

- [68] Huang J, Smola A J, Gretton A, et al. Correcting Sample Selection Bias by Unlabeled Data. in Proc. Int'l. Conf. on Neural Information Processing Systems, Vancouver, Canada, December, 2006, 601-608.
- [69] Sugiyama M, Nakajima S, Kashima H, et al. Direct importance estimation with model selection and its application to covariate shift adaptation. in Proc. 32nd Annual Conf. on Advances in Neural Information Processing Systems, Vancouver, Canada, December, 2008, 1433-1440.
- [70] Pan S J, Kwok J T, Yang Q, et al. Transfer learning via dimensionality reduction. in Proc. of 23th AAAI Conf. on Artificial Intelligence, Chicago, Illinois, July, 2008, 677-682.
- [71] Long M, Wang J, Ding G, et al. Transfer feature learning with joint distribution adaptation. in Proc. IEEE Int'l. Conf. on Computer Vision, Sydney, Australia, Dec., 2013, 2200-2207.
- [72] Wei B, Pal C. Cross lingual adaptation: an experiment on sentiment classifications. in Proc. the 48th Annual Meeting of the Association of Computational Linguistics, Uppsala, Sweden, 2010, Association for Computational Linguistics, 258-262.
- [73] Wu X, Wang H, Liu C, et al. Cross-view action recognition over heterogeneous feature spaces. in Proc. of the IEEE Int'l Conf. on Computer Vision, Sydney, Australia, December, 2013, 609-616.
- [74] Samat A, Persello C, Gamba P, et al. Supervised and semi-supervised multi-view canonical correlation analysis ensemble for heterogeneous domain adaptation in remote sensing image classification. *Remote Sensing*, 2017, 9(4):337.
- [75] Yeh Y R, Huang C H, Wang Y C F. Heterogeneous Domain Adaptation and Classification by Exploiting the Correlation Subspace. *IEEE Trans. on Image Processing*, 2014, 23(5):2009-2018.
- [76] Friedjungová M, Jirina M. Asymmetric Heterogeneous Transfer Learning: A Survey. in DATA, Madrid, Spain, 2017, 17-27.
- [77] Shi X, Liu Q, Fan W, et al. Transfer learning on heterogeneous feature spaces via spectral transformation. in 2010 IEEE Int'l conf. on Data Mining, Sydney, Australia, December, 2010, IEEE, 1049-1054.
- [78] Moon S, Carbonell J. Proactive transfer learning for heterogeneous feature and label spaces. in Joint European Conf. on Machine Learning and Knowledge Discovery in Databases, Riva del Garda, Italy, September, 2016, Springer, 706-721.
- [79] Mikolov T, Chen K, Corrado G, et al. Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781, 2013.
- [80] 庄福振, 罗平, 何清, 等. 迁移学习研究进展. *软件学报*, 2015, 26(1):26-39.
- [81] 刘建伟, 孙正康, 罗雄麟. 域自适应学习研究进展. *自动化学报*, 2014, 40(8):1576-1600.
- [82] Dai W, Yang Q, Xue G R, et al. Boosting for transfer learning. in Proc. 24th Int'l Conf. on Machine Learning, Corvallis, OR, June, 2007.
- [83] Gretton A, Borgwardt K, Rasch M, et al. A kernel method for the two-sample-problem. in Proc. Advances in Neural Information Processing Systems, Vancouver, Canada, Dec., 2007, 513-520.
- [84] Wang J, Feng W, Chen Y, et al. Visual domain adaptation with manifold embedded distribution alignment. in Proc. 26th ACM Int'l Conf. on Multimedia, Seoul, Korea, Oct., 2018, 402-410.

- [85] Persello C, Bruzzone L. Kernel-based domain-invariant feature selection in hyperspectral images for transfer learning. *IEEE Trans. on Geoscience and Remote Sensing*, 2015, 54(5):2615–2626.
- [86] Sun B, Feng J, Saenko K. Return of frustratingly easy domain adaptation. in Proc. 30th AAAI Conf. on Artificial Intelligence, Phoenix, AZ, Feb., 2016, 2058-2065.
- [87] Herath S, Harandi M, Porikli F. Learning an invariant hilbert space for domain adaptation. in Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition, HI, USA, July, 2017, 3845–3854.
- [88] Zhang J, Li W, Ogunbona P. Joint geometrical and statistical alignment for visual domain adaptation. in Proc. the IEEE Conf. on Computer Vision and Pattern Recognition, Honolulu, HI, July, 2017, 1859–1867.
- [89] Lu H, Shen C, Cao Z, et al. An embarrassingly simple approach to visual domain adaptation. *IEEE Trans. on Image Processing*, 2018, 27(7):3403–3417.
- [90] 舒醒, 于慧敏, 郑伟伟, 等. 基于边际Fisher 准则和迁移学习的小样本集分类器设计算法. *自动化学报*, 2016, 42(9):1313–1321.
- [91] Long M, Wang J, Ding G, et al. Transfer joint matching for unsupervised domain adaptation. in Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition, 2014, 1410–1417.
- [92] Long M, Zhu H, Wang J, et al. Deep transfer learning with joint adaptation networks. in Proc. of the 34th Int'l Conf. on Machine Learning, Sydney, Australia, August, 2017, 2208–2217.
- [93] Wen L, Gao L, Li X. A new deep transfer learning based on sparse auto-encoder for fault diagnosis. *IEEE Trans. on Systems, Man, and Cybernetics: Systems*, 2017, 49(1):136–144.
- [94] 陈小刚, 王毅军, 张丹. 2018 年脑机接口研发热点回眸. *科技导报*, 2019, 37(1):173–179.
- [95] Jayaram V, Alamgir M, Altun Y, et al. Transfer learning in brain-computer interfaces. *IEEE Computational Intelligence Magazine*, 2016, 11(1):20–31.
- [96] Azab A M, Mihaylova L, Ang K K, et al. Weighted Transfer Learning for Improving Motor Imagery-Based Brain-Computer Interface. *IEEE Trans. on Neural Systems and Rehabilitation Engineering*, 2019, 27(7):1352–1359.
- [97] Wang P, Lu J, Zhang B, et al. A Review on Transfer Learning for Brain-Computer Interface Classification. in Proc. 5th Int'l Conf. on Information Science and Technology (IC1ST), Changsha, China, April, 2015.
- [98] Waytowich N R, Lawhern V J, Bohannon A W, et al. Spectral Transfer Learning Using Information Geometry for a User-Independent Brain-Computer Interface. *Frontiers in Neuroscience*, 2016, 10:430.
- [99] Lotte F, Guan C. Learning from other subjects helps reducing brain-computer interface calibration time. in Proc. IEEE Int'l. Conf. on Acoustics Speech and Signal Processing, Dallas, TX, March, 2010.
- [100] Kang H, Nam Y, Choi S. Composite common spatial pattern for subject-to-subject transfer. *Signal Processing Letters*, 2009, 16(8):683–686.

- [101] Belkin M, Niyogi P, Sindhvani V. Manifold Regularization: A Geometric Framework for Learning from Labeled and Unlabeled Examples. *Journal of Machine Learning Research*, 2006, 7:2399–2434.
- [102] Yan S, Tang X. Trace quotient problems revisited. in *European Conference on Computer Vision*, Graz, Austria, May, 2006, Springer, 232–244.
- [103] Suppakun N, Maneewongvatana S. Automatic and semi-automatic approaches for Selecting prominent spatial filters of CSP in BCI applications. in *Int’l Conf. on Brain Informatics*, Beijing, China, October, 2009, Springer, 203–213.
- [104] Wang X, Ma L, Li H, et al. CSP Based Extraction and F-Score Based Optimization of Time-Frequency Power Features for EEG Mental Task Classification. in *2015 Int’l Conf. on Instrumentation and Measurement, Computer, Communication and Control (IMCCC)*, Qinhuangdao, China, September, 2015, IEEE, 820–824.
- [105] Cunningham J P, Ghahramani Z. Linear dimensionality reduction: survey, insights, and generalizations. *Journal of Machine Learning Research*, 2015, 16(1):2859–2900.
- [106] Blankertz B, Dornhege G, Krauledat M, et al. The non-invasive Berlin Brain-Computer Interface: Fast acquisition of effective performance in untrained subjects. *NeuroImage*, 2007, 37(2):539–550.
- [107] Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 2004, 134:9–21.
- [108] Qi H, Xue Y, Xu L, et al. A speedy calibration method using Riemannian geometry measurement and other-subject samples on a P300 speller. *IEEE Trans. on Neural Systems and Rehabilitation Engineering*, 2018, 26(3):602–608.
- [109] Rodrigues P L C, Jutten C, Congedo M. Riemannian Procrustes Analysis: Transfer Learning for Brain-Computer Interfaces. *IEEE Trans. on Biomedical Engineering*, 2018, 66(8):2390–2401.
- [110] Congedo M, Barachant A, Bhatia R. Riemannian geometry for EEG-based brain-computer interfaces; a primer and a review. *Brain-Computer Interfaces*, 2017, 4(3):155–174.
- [111] Li F, Xia Y, Wang F, et al. Transfer Learning Algorithm of P300-EEG Signal Based on XDAWN Spatial Filter and Riemannian Geometry Classifier. *Applied Sciences*, 2020, 10(5):1804.
- [112] Zanini P, Congedo M, Jutten C, et al. Transfer Learning: a Riemannian geometry framework with applications to Brain-Computer Interfaces. *IEEE Trans. on Biomedical Engineering*, 2018, 65(5):1107–1116.
- [113] Xie X, Yu Z L, Gu Z, et al. Bilinear regularized locality preserving learning on Riemannian graph for motor imagery BCI. *IEEE Trans. on Neural Systems and Rehabilitation Engineering*, 2018, 26(3):698–708.
- [114] Meng J, Liu G, Huang G, et al. Automated selecting subset of channels based on CSP in motor imagery brain-computer interface system. in *Proc. IEEE Int’l conf. on robotics and biomimetics (ROBIO)*, Guilin, China, Dec., 2009, 2290–2294.

- [115] Goldberger A L, Amaral L A N, Glass L, et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation*, 2000, 101(23):e215–e220.
- [116] Maaten L, Hinton G. Visualizing data using t-SNE. *Journal of Machine Learning Research*, 2008, 9:2579–2605.
- [117] Chang C C, Lin C J. LIBSVM: A library for support vector machines. *ACM Trans. on Intelligent Systems and Technology*, 2011, 2(3):27:1–27:27.
- [118] Wu D, Lawhern V J, Hairston W D, et al. Switching EEG Headsets Made Easy: Reducing Offline Calibration Effort Using Active Weighted Adaptation Regularization. *IEEE Trans. on Neural Systems and Rehabilitation Engineering*, 2016, 24(11):1125–1137.
- [119] Panareda Busto P, Gall J. Open set domain adaptation. in *Proc. the IEEE Int’l Conf. on Computer Vision*, Venice, Italy, October, 2017, 754–763.
- [120] Saito K, Yamamoto S, Ushiku Y, et al. Open set domain adaptation by backpropagation. in *Proc. the European Conf. on Computer Vision*, Munich, Germany, September, 2018, 153–168.
- [121] You K, Long M, Cao Z, et al. Universal Domain Adaptation. in *Proc. the IEEE Conf. on Computer Vision and Pattern Recognition*, Long Beach, CA, June, 2019, 2720–2729.
- [122] Arsigny V, Fillard P, Pennec X, et al. Geometric means in a novel vector space structure on symmetric positive-definite matrices. *SIAM Journal on Matrix Analysis and Applications*, 2007, 29(1):328–347.
- [123] Shimodaira H. Improving predictive inference under covariate shift by weighting the log-likelihood function. *Journal of Statistical Planning and Inference*, 2000, 90(2):227–244.
- [124] Utgoff P E. *Machine learning: An artificial intelligence approach*, volume 2. San Mateo, CA: CA: Morgan Kaufmann, 1986, chapter Shift of bias for inductive concept learning, pages 107–148.

## 附录 1 攻读博士学位期间发表论文目录

- [1] **He H** and Wu D, Different Set Domain Adaptation for Brain-Computer Interfaces: A Label Alignment Approach, *IEEE Trans. on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 5, pp. 1091-1108, 2020. (SCI源, 影响因子3.478, 中科院1区, 自动化学会A类, 署名单位: 华中科技大学)
- [2] **He H** and Wu D, Transfer Learning for Brain-Computer Interfaces: A Euclidean Space Data Alignment Approach, *IEEE Trans. on Biomedical Engineering*, vol. 67, no. 2, pp. 399-410, 2020. (SCI源, 影响因子4.491, 中科院2区, 自动化学会A类, 署名单位: 华中科技大学)
- [3] **He H** and Wu D, Transfer learning enhanced common spatial pattern filtering for brain computer interfaces (BCIs): Overview and a new approach, in *Proc. 24th Int'l Conf. on Neural Information Processing (ICONIP)*, Guangzhou, China, November 2017, pp. 811-821. (EI收录, 最佳学生论文奖提名, 署名单位: 华中科技大学)
- [4] **He H** and Wu D, Channel and Trials Selection for Reducing Covariate Shift in EEG-based Brain-Computer Interfaces, in *Proc. 2019 IEEE Int'l Conf. on Systems, Man and Cybernetics (SMC)*, Bari, Italy, October 2019, pp. 3635-3640. (EI收录, 署名单位: 华中科技大学)
- [5] **He H** and Wu D, Spatial Filtering for Brain Computer Interfaces: A Comparison between the Common Spatial Pattern and Its Variant, in *Proc. 2018 IEEE Int'l Conf. on Signal Processing, Communications and Computing (ICSPCC)*, Qingdao, China, September 2018, pp. 1-6. (EI收录, 署名单位: 华中科技大学)
- [6] Shi Z, Chen X, Zhao C, **He H**, Stuphorn V and Wu D, Multi-view Broad Learning System for Monkey Movement Intension Classification from Local Field Potentials and Spikes, *IEEE Trans. on Neural System and Rehabilitation Engineering*. (已接收, SCI源, 影响因子3.478, 中科院1区, 自动化学会A类, 署名单位: 华中科技大学)
- [7] Yang B, **He H**, and Hu X. Detecting community structure in networks via consensus dynamics and spatial transformation, *Physica A Statistical Mechanics and Its Applications*, Vol. 483, pp.156-170, 2017. (SCI源, 影响因子2.500, 中科院2区, 自动化学会B类, 署名单位: 武汉理工大学)

- [8] Yang B, Li X, Liu X, **He H**, Chen W, Alternating between consensus and leader selection reveals community structure in networks, *Physica A: Statistical Mechanics and its Applications*, Vol. 515, pp. 693-706, 2019. (SCI 源, 影响因子2.500, 中科院2区, 自动化学会B类, 署名单位: 武汉理工大学)

## 附录 2 攻读博士学位期间的其它研究成果

### [已申请发明专利]

1. 伍冬睿, 何赫. 一种减少脑机接口校准时间的方法. 中国国家发明专利, 申请号: 201911100991, 专利号: CN110851783A。

### [获奖情况]

1. 2020届暑期优秀毕业研究生荣誉称号, 颁奖单位: 华中科技大学, 2020年。
2. 2019世界机器人大赛—BCI脑控机器人大赛一等奖, 颁奖单位: 国家自然科学基金委, 2019年。
3. 优秀博士学业奖学金, 颁奖单位: 华中科技大学, 2019年。
4. 第一届深圳医疗健康大数据创新应用国际大赛三等奖, 颁奖单位: 深圳市发改委、卫计委, 2018年。
5. 第24届国际神经信息处理会议 (ICONIP) 最佳学生论文提名奖, 颁奖单位: 亚太神经网络学会 (APNNS), 2017年。

附录 3 公开发表的学术论文与博士学位论文的关系

章节号	章节名称	发表论文题目	刊物名称	作者署名名次
2	基于迁移学习的共同空间模式	Transfer Learning Enhanced Common Spatial Pattern Filtering for Brain Computer Interfaces (BCIs): Overview and A New Approach	24th International Conference on Neural Information Processing (ICONIP)	第一
		Spatial Filtering for Brain Computer Interfaces: A Comparison between the Common Spatial Pattern and Its Variant	2018 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)	第一
3	黎曼空间中迁移学习的优化方法	Channel and Trials Selection for Reducing Covariate Shift in EEG-based Brain Computer Interfaces	2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)	第一
4	欧式空间中的数据对齐方法	Transfer Learning for Brain-Computer Interfaces: A Euclidean Space Data Alignment Approach	IEEE Transactions on Biomedical Engineering	第一
5	异构类别空间的对齐方法	Different Set Domain Adaptation for Brain-Computer Interfaces: A Label Alignment Approach	IEEE Transactions on Neural Systems and Rehabilitation Engineering	第一