

利用脑成像多体素模式分析解码认知的神经表征：原理和应用

雷威^{1,2} 杨志² 詹旻野² 李红¹ 翁旭初²

(¹西南大学心理学院, 重庆 400715) (²中国科学院心理学研究所脑高级功能实验室, 北京 100101)

摘要 多体素模式分析(multi-voxel pattern analysis, MVPA)是一种基于机器学习理论发展出来的新的功能磁共振数据分析技术。MVPA 技术通过训练分类器,对由不同认知状态引起的多体素信号模式进行分类。与传统的基于单个体素的分析方法相比,该技术可更敏感地检测脑对认知状态的表征,并使得从神经信号解码认知状态成为可能。文章介绍MVPA技术的基本原理、分析步骤以及可以用MVPA来解决的科学问题,并对应用中可能面临的问题提供了参考建议。

关键词 多体素模式分析; 表征; 功能磁共振成像; 分类

分类号 B841.1

如何从神经成像数据中提取出与特定认知状态相联系的神经表征信息,是认知神经科学研究的基本问题。传统功能磁共振成像(functional magnetic resonance imaging, fMRI)数据分析试图通过基于单个体素(voxel)的血氧依赖性水平(blood oxygenation level dependent, BOLD)信号变化描述神经活动如何随认知状态的不同而改变(Kanwisher, 2000; Norman, Polyn, Detre, & Haxby, 2006)。但由于fMRI数据信噪比(signal noise ratio, SNR)较低和必须对结果进行多重比较校正,基于单个体素的单变量分析方法(voxel-wise analysis)对数据中携带的信息的统计检测力不足(Norman et al., 2006)。

近年来,通过借鉴机器学习研究中的模式分类(pattern classification)技术,研究者们发展出一种新的多变量fMRI数据分析方法——“多体素模式分析”(multi-voxel pattern analysis, MVPA)。该方法将多个体素的信号看作一个多维变量,即空间模式(spatial pattern),并求解判别函数(分类器, classifier)用以区分不同认知状态下形成的空间模式。MVPA以多个体素信号形成的空间模式为基础进行分析,可以很大程度克服单个体素数据信噪比低、多重比较校正严格等因素带来的局限性,因而提高了从神经成像数据中检测不同认知状态间差异的灵敏度。运用该方法,研究者不但能以比传统方法更高的精度检测出与特定认

知状态相对应的神经信号模式(Davatzikos, Ruparel, Fan, Shen, & Acharyya, 2005; Hutchinson, 2006; MacEvoy & Epstein, 2009),还能从这些模式中解码被试所处的认知状态或其知觉内容(Kay, Naselaris, Prenger, & Gallant, 2008; Miyawaki et al., 2008)。

本文第一部分概述MVPA的基本原理以及它相对于传统分析方法的优势;第二部分介绍MVPA的实现方法和技术细节;第三部分通过已有的研究案例介绍可以用这种技术来解决的科学问题;最后对MVPA技术在研究和实践中的应用进行展望。

1 MVPA 原理

1.1 模式分析的基本原理

MVPA技术的核心原理是利用在不同认知状态下由多个体素信号形成的空间模式训练分类器,再用独立的实验数据测试分类器的性能(图1)。训练和测试分类器的过程称为模式分析,是机器学习理论中的重要技术,原本用于机器对模式进行识别(比如,手写或者语音识别)。

在神经成像研究中,分类器将不同特征(fMRI数据中的每一个体素可作为一个特征)组合成一个特征空间(如由多个体素变量形成的多维空间),特征空间的不同取值表现为不同的空间模式,称为样本(fMRI数据中特定认知状态下由多个体素信号形成的空间模式即为一个样本)。通过将已知类别的数据导入分类器可以求解出

描述特征空间的函数，并得到一个区分不同类别的决策边界(decision boundary)，这个过程就称为“学习”或者“训练分类器”。用于训练分类器的数据称为训练集。要确定训练后的分类器的性能，则需要通过将它用于新的样本来测试。用于测试分类器性能的数据叫做测试集。类别学习算法的

典型假设是：训练集和测试集满足独立同分布的性质，因此，如果分类器确实掌握了特征和类别间的关系，则应该能准确预测它从未见过但同属该类别的新模式(即测试集样本)。评价分类器的常用标准是预测率，即被分类器正确预测的样本在所有测试样本中所占的比例。

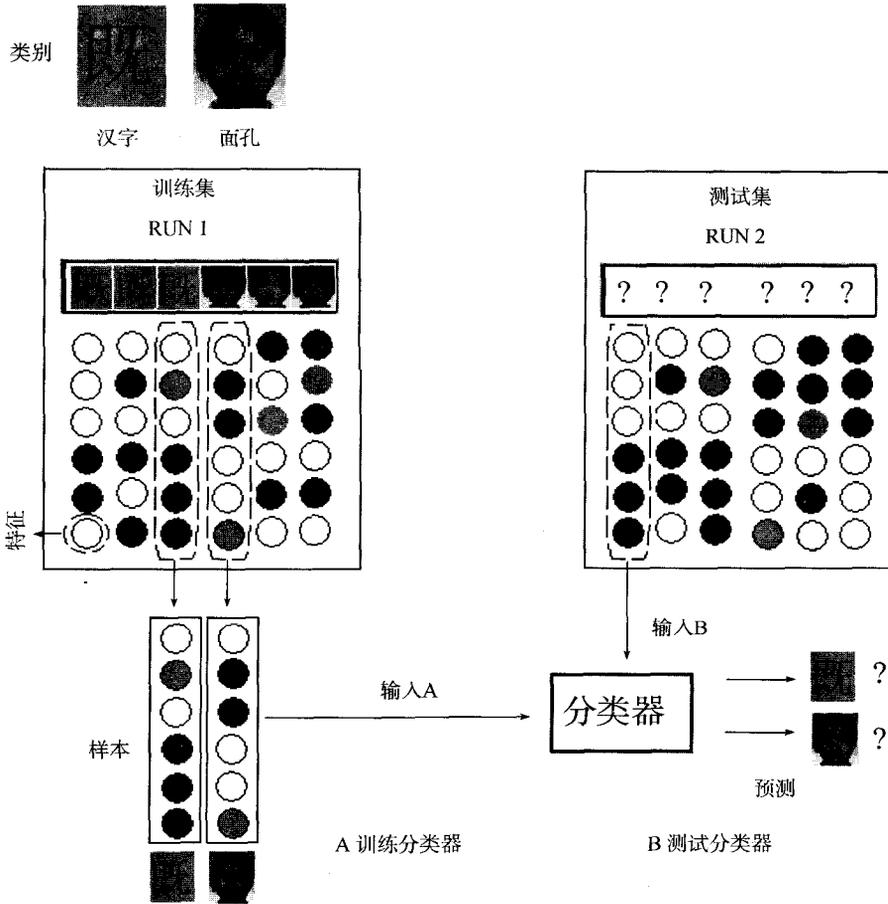


图 1 模式分析简图

图 1 为一个假设的实验中应用 MVPA 来进行分析的过程。向被试呈现两个类别的刺激(汉字和面孔)，以诱发不同的认知状态。通过特征选择程序(见本文 2.1)选出用于模式分析的特征(图中圆点)。依据实验顺序将这些特征在不同认知状态下的取值组合成不同的模式(样本，图中圆点列)，并标记对应的类别；将样本分为训练集和测试集；(A)将已知类别的样本(训练集)输入分类器(输入

A)使后者习得不同认知状态所对应的信号模式差异；(B)然后将新的信号模式(测试集样本)输入训练后的分类器(输入 B)来预测(预测)其所对应的类别，以测试分类器的性能。

1.2 分类器的种类

依据可以学习的关系种类，分类器可以分成线性分类器(linear classifier)和非线性分类器(nonlinear classifier)两种。

线性分类器可以学习特征和类别之间的线性关系。比如有 v 个特征的特征空间 y ，其分类器函数可以表示为特征值的加权和： $y = x_1w_1 + x_2w_2 + \dots + x_vw_v$ ，函数中每个特征对应的权重值“ w ”需要通过训练分类器来求解。现有的 MVPA 研究大都使用线性分类器。如没有隐藏层的神经网络 (neural networks) (Polyn, Natu, Cohen, & Norman, 2005)、线性支持向量机 (linear support vector machines, SVMs) (Pessoa & Padmala, 2007; Kamitani & Tong, 2005) 和朴素贝叶斯分类器 (Gaussian Naïve Bayes classifiers) (Mitchell *et al.*, 2004) 等。

非线性分类器可以学习特征和类别之间复杂的非线性关系，其求解过程也更为复杂。如非线性支持向量机 (nonlinear support vector machines) (Cox & Savoy, 2003) 和有隐藏层的神经网络 (Hanson, Matsuka, & Haxby, 2004) 等。线性和非线性分类器的关键区别是，非线性分类器可以学习特征间的交互作用，即特定认知状态出现时多个体素共同作用的信息 (Norman *et al.*, 2006)。尽管非线性分类器能学习更多的特征关系，但相对于线性分类器，非线性分类器还存在以下两个问题：(1) 尽管非线性分类器在实践中表现出相对于线性分类器的显著优势，但其原理并不清楚，有研究者认为这可能仅仅是由于可用的样本数太少导致非线性分类器未能很好的学习特征之间的交互作用关系 (Pereira, Mitchell, & Botvinick, 2009)；(2) 非线性分类器对特征间关系的描述和预测结果的解释都比线性分类器更复杂也更模糊 (Kamitani & Tong, 2005)。

1.3 MVPA 技术的优势

传统的基于单个体素的 fMRI 数据分析方法对各体素一一分析，测试其信号变化是否与认知状态的变化相关。由于磁共振数据本身的信噪比较低，传统方法通常要求纳入分析的体素信号高于某个统计阈值，并且为了提高检测敏感性，传统方法通常会在数据的预处理中将感兴趣区域 (region of interest, ROI) 内的信号进行空间平均，通过检测整个区域体素的信号在不同条件下的差异来推论该皮层区域是否“参与”了特定认知功能。而 MVPA 通过对多个体素信号形成的空间模式进行编码来区分不同的认知状态。作为一种多变量分析技术，MVPA 对单个体素的噪音较传统

方法更不敏感，即使包含了大量噪音的磁共振数据也可以通过模式分析来有效的分类，从而可以推论特定认知状态的神经表征。具体来说，MVPA 相对于传统数据分析方法的优势主要表现在以下 4 个方面：

(1) 传统方法只分析对实验条件有显著反应的体素，忽略了那些信号未能达到统计显著水平但仍可能很好地区分不同认知状态的体素。而 MVPA 技术同时分析多个体素信号形成的空间模式，模式中不仅包含单个体素反应的信息，也包含多个体素共同作用的信息 (即使其中某些体素的信号并没有达到显著水平) (Raizada, Tsao, Liu, & Kuhl, 2009)。比如，最近的一个研究使用模式分析表明知觉上不同的语音在右侧听觉皮层引起的信号水平相当，但信号模式却不同 (Mur, Bandettini, & Kriegeskorte, 2009)。

(2) 传统分析中常用的空间平均虽然降低了噪音，却也掩盖了某些对不同认知状态具有良好区分力的细微空间模式，而 MVPA 保留了这些信息，从而大大提高了检测细微模式差异的敏感性 (Norman *et al.*, 2006)。Haynes 和 Rees 通过初级视觉皮层的分布式信号模式解码刺激朝向的研究 (Haynes & Rees, 2005a) 就很好地说明了 MVPA 在解码细微表征差异方面的这种优势。

(3) 传统方法利用数据与认知状态在时间上的相关性来检测神经活动，因此只有进行多次重复测量才能获得可靠结果。而 MVPA 技术相对不依赖于信号模式的重复出现，因而允许我们仅仅从几个 trial 的数据来估计被试的认知状态 (Haynes & Rees, 2006)。已经有研究应用 MVPA 技术成功追踪了被试在双眼竞争中知觉内容的动态变化过程 (Haynes & Rees, 2005b)。

(4) 比起传统单变量分析，MVPA 技术同时考虑多个体素的信号变化，更符合 fMRI 数据内在的多变量特性 (Mur *et al.*, 2009)。因而模式分析对神经活动和认知状态之间关系的描述比传统方法更加完整。

总之，MVPA 技术保留和利用了比传统方法更完整的信息，从而大大提高了检测认知表征间差异的敏感性，使得它成为认知神经科学研究中非常有力的工具。因而 MVPA 受到越来越多研究者的青睐，并在应用中得到了许多令人兴奋的结果 (Haxby *et al.*, 2001; Cox & Savoy, 2003;

O'Toole, Jiang, Abdi, & Haxby, 2005; Kamitani & Tong, 2005; Kriegeskorte, Formisano, Sorger, & Goebel, 2007; Schurger, Pereira, Treisman, & Cohen, 2009)。比如, 研究者可以依据神经信号模式来推论被试正在想象何种刺激(Thirion et al., 2006)。或者, 即使在信号回到基线水平以后, 研究者也可以从初级视觉区的神经信号模式中以超过 80% 的正确率预测被试记住了两个不同朝向光栅中的哪一个(Harrison & Tong, 2009)等等。本文第三部分将更详细地介绍应用 MVPA 技术来解码的认知神经表征的更多研究案例。

2 MVPA 的分析步骤

MVPA 的分析步骤大致包含以下内容: 特征选择(feature selection)、创建样本(creating examples)、训练分类器(classifier training)和泛化测试(generalization testing)。这部分介绍每一步骤的内容和要点。

2.1 特征选择

典型的 fMRI 研究中, 数据中的特征(fMRI 研究中通常以体素为特征)常会比样本(多个体素信号形成的空间模式)多。若以全脑的体素为特征, 仅有几十个样本却有几万个特征的情况也并不少见。这种情况下分类器在训练集的性能可能会很好, 但在测试集中却很不稳定, 这个现象称为过拟合(over fitting) (对过拟合的详细解释见 Mitchell, 1997)。因此研究者通常不会使用神经成像数据所包含的全部体素, 而会选取某些符合特定条件的体素来组成用于分析的空间模式, 这个过程叫做特征选择。这样做不但能有效避免过拟合, 而且能使得到的模式更确切地反映认知状态在神经活动中的表征。

特征选择中有两个问题需要解决: 首先, 以什么作为特征? 研究者需要依据研究目的来确定用什么参数描述特征值。比如, 以体素在特定认知状态下的信号与基线条件对比所得的 t 值, 或者以体素在不同认知状态(如, 对 A 与 B\C\D)下信号差异测试的 F 值等。其次, 选取哪些特征? 在具体研究中, 根据研究目的有两种方法取向, 一种是 ROI 分析, 另一种是全脑分析。ROI 分析常会将在解剖上或者功能上事先定义好的 ROI 内的体素作为特征(Kamitani & Tong, 2005)。当研究问题明确指向某些脑区时, ROI 分析是最直接的方法。ROI 分析的局限性是: 只能测试有

限空间内的模式差异, 而忽视 ROI 以外对不同认知状态可能有良好区分的体素; 当研究问题是定位表征特定认知状态的体素时, 研究者可以在全脑范围内进行特征选择(用于全脑的特征选择方法见 De Martino et al., 2008)。全脑分析的优势是它不需要事先假设与特定认知状态相关的体素位置, 当认知状态与广泛分布的信号模式相联系时这种方法尤其有优势。但是当分析目标是细微的模式差异时, 全脑分析则可能不够准确(Haynes & Rees, 2005b)。

2.2 创建样本

完成特征选择后, 即可按照实验程序将这些特征在不同认知状态下的信号组合成不同的空间模式, 该过程称为创建样本。在 fMRI 实验中一个样本通常就是选出的体素在一个 trail 内的信号所形成的空间模式, 样本中每个体素的特征值通常以其在这个 trial 内的信号均值代表。根据研究需要也可以进一步区分每个 trial 中不同时间段的信号或者将几个 trial 的信号平均作为样本等等。选择样本时需注意每两个样本的抽取要相互独立, 训练集和测试集样本的选取也要相互独立。对于 block 设计或者慢速的事件相关设计数据而言, 这一条件通常都能满足, 而快速事件相关设计数据中则需要考虑刺激的延续效应对其后 trial 的影响。一般样本越多则分类器的性能越好。对训练集而言, 每个类别至少需要几十个样本, 以便分类器有效地确定所需的参数。对测试集而言, 样本数越多对分类器性能的测试越精确。

除了样本数量, 创建样本过程中还有一些重要问题需要考虑: (1)用于训练的数据必须包含所有的类别信息, 分类器不能预测未曾学习过的类别; (2)各个类别样本数应该平衡, 否则大多数判别函数会偏向于样本数较多的类别, 造成对结果的解释出现偏差(Pereira et al., 2009)。

2.3 训练分类器

训练分类器是将训练集样本输入分类器, 以便后者从这些数据中学习可区分不同类别的函数。受到特征数、样本数、实验条件等因素的影响, 不同分类器在敏感性上稍有差异, 但大都能有效地解码神经信号模式中所包含的认知表征信息(实践中分类器的选择见 Pereira et al., 2009)。需要注意的是, 尽管有些分类器能同时区分多个类别, 实践中一般先进行类别之间的两两对比

(如,类别 A 与类别 B 对比,类别 B 与类别 C 对比等),再将这些两两对比结合起来形成多类别比较。因为即使在包含了多个条件的实验中,同时对比多个类别的结果也不如两两对比稳定(Mur et al., 2009)。

为了提高分类器性能,我们希望训练集样本越多越好,但是又不能用同一组数据同时进行训练和测试。在样本数有限的条件下尽可能提高分类器性能的方法之一就是交叉验证(cross-validation)。交叉验证有以下步骤:(1)将全部样本分成 k 份(如 $k=10$,即每次取出 10%的数据, $k=5$ 即取出 20%的数据);(2)从中取出一份样本,将其余样本($k-1$ 份)用于训练分类器,再用这份样本进行测试;(3)轮流取出每一份样本,重复步骤(2);(4)将得到的所有预测率平均。这样尽管每次训练得到的分类器函数有细微的不同,但因大部分训练数据一致,我们可以认为这些分类器都很相似。因此可以将这样得到的平均预测率看作用全部样本训练分类器的预测率的无偏估计值。交叉验证不是 MVPA 中必须的步骤,但是大多数研究者都会在实际分析中使用交叉验证以提高分类器性能(Haynes & Rees, 2005b; Mitchell et al., 2008)。

2.4 泛化测试

完成训练的分类器被用于新的数据(测试集),以检测分类器在更广泛情境中预测特定神经信号模式所对应的认知状态的能力,这个过程称为泛化测试。泛化测试是训练的逆向操作,向分类器呈现未知类别的数据(空间模式)以检验分类器是否能正确预测该模式所属的类别,分类结果多以预测率来表示。fMRI 研究中训练分类器的目的,一般是从神经信号模式中将感兴趣的认知状态表征解码出来。也就是说,研究者希望训练之后的分类器能以高于随机水平(如 50%)的正确率对不同认知状态或者知觉内容进行分类。从统计学角度而言,在测试集上的预测率实际上是分类器真正预测率(即当测试集无限时分类器的预测率)的无偏估计值。这个估计值有多精确取决于测试集的大小,测试集越小估计值的变异就越大,因此交叉验证也可以适当提高泛化测试的准确性。最后研究者可能需要对分类结果进行统计检验,若样本抽取相互独立,则可以用二项检验来反映分类器的性能是否显著高于随机水平。若怀疑抽

样中存在某种选择偏差,也可用置换检验法(permutation test)进行检验。

3 MVPA 的应用

3.1 解码知觉表征的差异

无意识条件下对不同朝向的物体轮廓敏感的神经模块在初级视觉区的差异非常细微(仅有约 500 微米)(转引自 Haynes & Rees, 2005a)。现有的 fMRI 成像技术的空间解析度(通常大于 1 毫米)还无法达到这种水平,传统分析方法为了提高信号质量常常需要将信号进行空间平均,更难以区分此精度下的神经表征差异。而应用 MVPA 技术的研究不但成功解码出阈下呈现的不同朝向光栅在 V1 区的神经信号模式(Haynes & Rees, 2005a),还能够从模式差异中准确解码被试所见点的运动方向(Kamitani & Tong, 2006)。

由于对精细的模式差异敏感, MVPA 技术在许多研究中被用于解码被试的知觉内容。比如,从外侧纹状体视觉区的分布式信号模式中研究者可以准确地预测出被试看见了 12 种不同物体中的哪一种(Sayres, Rees, & Grill-Spector, 2005)。或者被试所见词的语义类别(Mitchell et al., 2008)。除了解码单独呈现的刺激, MVPA 技术还可以用于解码更为复杂的知觉内容,比如同时呈现的两个刺激的类别(MacEvoy & Epstein, 2009),甚至可解码被试对内容复杂的自然图片的知觉(Kay et al., 2008)。

与解码知觉内容相似,一些研究中 MVPA 也被用于解码被试的认知状态,如,玩牌的时候被试是否就某张牌的内容撒了谎(Davatzikos et al., 2005)。在阈下呈现的面孔情绪判断实验中,被试看见的是中性还是恐惧表情(Pessoa & Padmala, 2007)。或者在记忆提取任务中被试回忆起的刺激类别(Polyn et al., 2005)等。

3.2 定位认知表征

除了从神经信号模式中区分不同类别的认知信息,人们可能还想知道这些信息被表征在哪些大脑结构当中,以及是何时表征的。在 MVPA 中,表征定位问题可以表述为:哪些体素对分类器预测表现的影响最大。而由于在线性分类器中每个体素仅通过权重来影响预测结果,因此只要知道该特征的权重,就可以推论该体素对认知表征的贡献大小。运用 MVPA 技术,Reddy 等人在最近的一个研究中发现,经典的类别选择性区域

(如, 对面孔刺激敏感的梭状回面孔区 FFA, 见 Kanwisher, 2000)内的体素在偏好的刺激条件下的权重比在其他刺激条件下高出 30% (Reddy, Kanwisher, & VanRullen, 2009)。另一个研究中, Harrison 和 Tong 发现以 V1 区体素的信号模式训练的分类器比 V2-V4 区能更准确的预测被试所见线条的朝向, 从而证明线条朝向表征与 V1 区的联系更为紧密(Harrison & Tong, 2009)。

由于 fMRI 数据本身时间分辨率的限制, 从神经信号中解码出被试在短时间(数秒)内的认知状态变化在传统的分析方法中是非常困难的。而 MVPA 对不同认知状态差异的检测敏感性的提高, 使得在实验中追踪认知状态变化的时间进程成为可能。已经有研究者使用 MVPA 成功的追踪了被试认知状态改变的进程, 比如, 仅仅基于单个 trial 的神经信号模式, 研究者可以准确地预测出在双眼竞争中被试当前所见的优势刺激的类别 (Haynes & Rees, 2005b), 或者被试的决策过程 (Pessoa & Padmala, 2007)。

3.3 解码认知表征方式

解码认知变量在头脑中的具体表征方式是认知神经科学研究中最大的难题之一(Pereira et al., 2009)。MVPA 技术的应用为我们理解认知变量的表征方式提供了新的视角。MVPA 假设认知状态由多个维度组成, 特定维度上的不同取值由不同的神经信号模式来表示。这意味着我们可以通过测量特定区域 A 中的信号模式随着认知维度 x 变化的程度, 来测量有多大可能认知维度 x 是在脑区 A 表征的。因此, 测查脑区 A 是否表征了认知维度 x 的有效途径之一是: 将认知状态沿维度 x 逐步变化, 看分类器的性能是否会随着认知状态的相似性提高而逐步下降(Norman et al., 2006)。已经有研究者证明了这种方法的可行性, 比如, O'Toole 等发现当刺激间的视觉相似性上升时, 应用于腹侧颞叶的分类器的预测准确率则随之下降(O'Toole et al., 2005)。除了利用神经信号模式的相似性以外, 也有研究者直接通过对比神经信号模式间的空间关系来推论刺激表征的神经机制。如, Hutchinson 发现同时进行的多重认知加工与重叠的信号模式相联系 (Hutchinson, 2006)。

MacEvoy 和 Epstein 的实验很好的说明研究者是如何用模式分析来解码知觉刺激的表征方

式的(MacEvoy & Epstein, 2009)。该研究中实验者给被试呈现了四类刺激(刷子、鞋子、椅子和汽车), 呈现方式包括单独呈现和配对呈现(两个类别的刺激)两种。然后将数据分成了两半, 分别计算了每类刺激(和刺激对)引起的信号模式。这个研究中他们发现在外侧枕叶皮层(LOC)视觉刺激的信号模式是按照线性方式组合的, 配对刺激的信号模式可以用组成刺激对的两个类别刺激单独呈现时的信号模式平均来准确预测。

4 总结和展望

MVPA 技术将皮层神经信号的空间模式作为分析对象, 模式差异不仅包括了传统方法所关注的单个体素信号水平差异, 也包括不依赖单个体素信号水平的多体素模式差异, 比如信号模式的稳定性差异被认为反映了被试是否意识到了(同一个)刺激(Schurger et al., 2009)等。MVPA 利用了磁共振信号中更加完整的信息, 大大提高了通过神经成像数据检测认知表征的敏感性。

除了对认知表征差异的检测力度更强之外, MVPA 技术还使得从神经信号模式反过来推论被试的认知状态成为可能。这是一个意义重大的进步: 一方面, MVPA 已经超越了仅仅作为一种数据处理工具的功能, 且启发研究者从一个全新的角度看待认知过程, 即不仅可以通过对操纵认知状态来诱发相应的神经活动, 而且可以反过来从神经信号中推测被试的认知状态。这种预测功能将会为我们深入理解认知状态在大脑中的表征提供更多有趣的证据, 推动认知神经科学研究的发展。另一方面, 这种全新的预测功能具有相当的应用价值。比如, 最近的研究中 MVPA 技术已经被用于解码被试对复杂的自然情境的知觉 (Hasson, Nir, Levy, Fuhrmann, & Malach, 2004; Kay et al., 2008)或者测谎(Davatzikos et al., 2005)等。这意味着也许在不久的将来人们就能够通过神经信号完整地重建被试的知觉经验了。

MVPA 在今后研究中的应用有两个可能的发展方向: (1)将 MVPA 与其他技术结合使用。尽管检测敏感性大大提高, MVPA 对亚体素水平模式差异的区分还是有困难, 这就需要将 MVPA 和其他技术, 如高空间分辨率的成像序列或者新的实验设计方法(如相对不受体素大小限制的 fMRI 适应设计, 详细综述见: Grill-Spector & Malach, 2001)相结合来进行研究。最近已有研究尝试将

MVPA 和 fMRI 适应技术结合起来(Aguirre, 2007)。(2)改进分类器算法。由于线性分类器实际上是所有特征的加权求和,并不能表示特征之间的交互作用,而研究表明两块区域共同作用的信息要大于两块区域信息的简单相加(Pessoa & Padmala, 2007),因此改进非线性分类器算法以考察体素间的交互作用是不能忽视的问题。尽管还存在种种尚待解决的问题,可以相信随着分类算法的进步、fMRI 技术空间解析度的提高和计算机技术的进步, MVPA 技术带给我们的将不仅仅是 fMRI 数据分析的进步,还包括对认知神经表征理解的深入。

最后,现在已经有一些实用的 MVPA 工具箱供研究者选择,如 Brainvoyager 2.0 及以后的版本(www.brainvoyager.com/index.html),可以在 MATLAB 中使用的 Princeton-MVPA-Toolbox (code.google.com/p/Princeton-mvpa-tool-box)以及同时适用于 ERP 和 fMRI 数据的 PyMVPA 工具箱(www.pymvpa.org)等。

参考文献

- Aguirre, G. K. (2007). Continuous carry-over designs for fMRI. *Neuroimage*, 35(4), 1480-94.
- Cox, D. D., & Savoy, R. L. (2003). Functional magnetic resonance imaging (fMRI) 'brain reading': detecting and classifying distributed patterns of fMRI activity in human visual cortex. *Neuroimage*, 19(2), 261-270.
- Davatzikos, C., Ruparel, K., Fan, Y., Shen, D. G., & Acharyya, M. (2005). Classifying spatial patterns of brain activity with machine learning methods: application to lie detection. *Neuroimage*, 28(11), 663-668.
- De Martino, F., Valente, G., Staeren, N., Ashburner, J., Goebel, R., & Formisano, E. (2008). Combining multivariate voxels election and support vector machines for mapping and classification of fmri spatial patterns. *Neuroimage*, 43(1), 44-58.
- Grill-Spector, K., & Malach, R. (2001). fMRI-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychologica*, 107(1-3), 293-321.
- Hanson, S. J., Matsuka, T., & Haxby, J. V. (2004). Combinatorial codes in ventral temporal lobe for object recognition: Haxby (2001). revisited is there a 'face' area? *Neuroimage*, 23(1), 156-166.
- Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working Memory in early visual areas. *Nature*, 458(7238), 632-635.
- Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science* 303(5664), 1634-1640.
- Haynes, J. D., & Rees, G. (2005a). Predicting the orientation of invisible stimuli from activity in primary visual cortex. *Nature Neuroscience*, 8(5), 686-691.
- Haynes, J. D., & Rees, G. (2005b). Predicting the stream of consciousness from activity in human visual cortex. *CurrentBiology*, 15(14), 1301-1307.
- Haynes, J. D., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*, 7(7), 523-534.
- Hutchinson, R. (2006). Hidden process models. In Mitchell T M, Rustandi I. (Eds). *Proceedings of the 23rd international conference on Machine learning* (pp. 433-440). New York: ACM.
- Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience*, 8, 679-685.
- Kamitani, Y., & Tong, F. (2006). Decoding Seen and attended motion directions from activity in the human visual cortex. *Current Biology*, 16(11), 1096-1102.
- Kanwisher, N. (2000). Domain specificity in face perception. *Nature Neuroscience*, 3(8), 759-63.
- Kay, K. N., Naselaris, T., Prenger, R. J., & Gallant, J. L. (2008). Identifying natural images from human brain activity. *Nature*, 452, 352-355.
- Kriegeskorte, N., Formisano, E., Sorger, B., & Goebel, R. (2007). Individual faces elicit distinct response patterns in human anterior temporal cortex. *PNAS*, 104(51), 20600-20605.
- MacEvoy, S. P. & Epstein, R. A. (2009). Decoding the representation of multiple simultaneous objects in human occipitotemporal cortex. *CurrentBiology*, 19(11), 1-5.
- Mitchell T M. (1997). *Machine Learning*. New York: McGraw Hill Press.
- Mitchell, T. M., Hutchinson, R., Niculescu, R. S., Pereira, F., Wang, X., Just, M. et al. (2004). Learning to decode cognitive states from brain images. *Machine Learnin*, 5, 145-175.
- Mitchell, T. M., Shinkareva, S. V., Carlson, A. Chang, K. M., Malave, V. L., & Mason, R. A. (2008). Meanings of nouns predicting human brain activity associated with the meanings of nouns. *Science*, 320, 1191-1195.
- Miyawaki, Y., Uchida, H., Yamashita, O., Masa-akiSato. Morito, Y., & Tanabe, H. C. et al., (2008). Visual image reconstruction from human brain activity using a combination of multi scale local image decoders. *Neuron*, 60(5), 915-929.
- Mur, M., Bandettini, P. A., & Kriegeskorte, N. (2009). Revealing representational content with pattern- information

- fMRI an introductory guide. *SCAN*, 4(1), 101–109
- Norman, K., Polyn, S.M., Detre, G., & Haxby, J.V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fmri data. *Trends in Cognitive Sciences*, 10(9), 424–430.
- O'Toole, A. J., Jiang, F., Abdi, H., & Haxby, J. V. (2005). Partially distributed representations of objects and faces in ventral temporal cortex. *Journal of Cognitive Neuroscience*, 17(4), 580–590.
- Pereira, F., Mitchell, T., Botvinick, M. (2009). Machine learning classifiers and fMRI: A tutorial overview. *NeuroImage*, 45(1), S199–S209.
- Pessoa, L., & Padmala, S. (2007). Decoding Near-Threshold Perception of Fear from Distributed Single-Trial Brain Activation. *Cerebral Cortex*, 17(3), 691–701.
- Polyn, S. M., Natu, V. S., Cohen, J. D., & Norman, K. A. (2005). Category-specific cortical activity precedes recall during memory search. *Science*, 310(5756), 1963–1966.
- Raizada, R. D. S., Tsao, F. M., Liu, H. M., & Kuhl, P. K. (2009). Quantifying the adequacy of neural representations for a cross-language phonetic discrimination task: prediction of individual differences. *Cerebral Cortex*, 20(1), 1–12.
- Reddy, L., Kanwisher, N. G., & VanRullen, R. (2009). Attention and biased competition in multi-voxel Object representations, *PNAS*, 106(50), 1447–21452.
- Sayres, R., Ress, D., & Grill-Spector, K. (2005). Identifying distributed object representations in human extrastriate visual cortex. *Advances in Neural Information Processing Systems*. 1–8.
- Schurger, A., Pereira, F., Treisman, A., & Cohen, J. D. (2009). Reproducibility distinguishes conscious from nonconscious neural representations. *Science*, 2009, 327(5961), 97–99.
- Thirion, B., Duchesnay, E., Hubbard, E. *et al.* (2006). Inverse retinotopy: Inferring the visual content of images from brain. *NeuroImage*, 33(4), 1104–1116.

Decoding the Representation of Cognition: the Principles and Applications of MVPA

LEI Wei^{1, 2}; YANG Zhi²; ZHAN Min-Ye²; LI Hong¹; WENG Xu-Chu²

(¹Southwest University, School of Psychology, Chongqing 400715, China)

(² Lab for Higher Brain Function, Institute of Psychology, Chinese Academy of Science, Beijing 100101, China)

Abstract: Multi-voxel pattern analysis (MVPA), which is based on machine learning theories, has gained great popularity over the past years as a new approach for fMRI data analysis. By training a classifier, MVPA categorizes multi-voxel patterns tuned by different cognitive states. Compared to conventional voxel-wise methods, this new approach provide higher sensitivity for detecting cognitive representations in the brain. It opens up the possibility for “reading out” mental states of human beings from the non-invasive recordings of brain activities. This paper introduce the fundamental principles of MVPA and the basic realization procedures. Scientific questions that may be properly addressed with this new approach and potential problems in its applications are also discussed.

Key words: MVPA; representation; fMRI; classification