基于脑功能磁共振影像的视觉信息解读技术*

雷 煜, 童 莉, 李建新, 闫 镔
(国家数字交换系统工程技术研究中心, 郑州 450002)

摘 要: 介绍了以脑信号视觉计算模型为重点的视觉信息编码技术和以脑信号分类与识别为重点的视觉信息解码技术的发展现状;重点阐述了编码研究中的 Gabor 感受野模型、形式相似性分析方法和解码研究中的多体素分析方法,并讨论了分类器设计中的线性和非线性分类器以及特征选择问题;最后对该技术的发展趋势进行了展望。

关键词: 功能磁共振成像; 编码; 解码; 多体素分析; 计算模型
中图分类号: TP391 **文献标志码**: A **文章编号**: 1001-3695(2013)06-1611-06
doi:10.3969/j.issn.1001-3695.2013.06.003

Functional magnetic resonance imaging "brain reading" of visual information: a review

LEI Yu, TONG Li, LI Jian-xin, YAN Bin
(*China National Digital Switching System Engineering & Technological Research Center*, *Zhengzhou* 450002, *China*)

Abstract: This paper first reviewed the encoding teniques focused on computational models and decoding teniques focused on classification and identification, particullarly the Gabor receptive-field model, representational similarity analysis method in encoding and the multivariate pattern analysis(MVPA) approach in decoding. In addition, it also reviewed the linear and non-linear classifiers and feature selection approaches in decoding models. In the end, it suggested future research directions of "brain reading" with fMRI.

Key words: fMRI(functional magnetic resonance imaging); encoding; decoding; MVPA; computational model

0 引言

人脑在复杂视觉信息处理方面具有高效、鲁棒、抗噪等重要优点,而视觉信息是人类感知和认识外部世界最主要的途径之一[1]。近年来,随着现代神经科学关于脑认知过程信号获取、处理技术不断进步,采用脑电(electroencephalography/event-related potential, EEG/ERP)、脑磁(magnetoencephalography, MEG)、功能磁共振成像(fMRI)、近红外光学脑功能成像(functional near-infrared spectroscopy, fNIRS)等无创的神经影像手段进行人类视觉信息处理机制的研究取得了一系列的重要成果。其中,fMRI 作为同时具备良好的空间分辨率和时间分辨率的脑功能成像手段,已成为观察大脑活动的主要工具。近十年来,由于 fMRI 技术的性能不断提升,研究人员已经不再满足于通过功能磁共振成像技术探测大脑的激活区域,而是开始利用功能磁共振成像获得的数据进行大脑认知活动的解读,其中关于人脑视觉认知活动的解读研究又成为其最主要的研究方向[2~8]。

基于脑信号的视觉信息解读研究主要分为正向编码和逆向解码两个方向。其中,视觉信息的脑信号编码研究是以大脑视觉感知机制为基础,通过建立一个大脑视觉信息处理的可计算编码模型来描述大脑对外界刺激的响应过程,以实现对大脑活动的预测[9]。简而言之,正向编码描述了人类在感知视觉刺激时脑信号变化的过程,如图 1(a)所示。其中计算模型就是一个以人脑视觉感知的生理机制为基础的数学模型,模型的输入是图像的特征,模型的输出为大脑对图像的响应。研究视觉信息的 fMRI 脑信号编码计算模型,对于探索大脑视觉信息处理的过程和机制,利用机器智能实现人类视觉功能的模拟,提高机器智能对于自然界真实视觉信息的感知和认知能力具有重要意义。

视觉信息的脑信号解码研究则主要通过研究脑信号,从中找出关于外界刺激的对应关系,利用人类对于视觉信息的大脑响应信号,实现利用脑信号对外界刺激进行正确分类或识别,这是一个从脑信号到刺激的过程,如图 1(b)所示。逆向解码技术通常需要采用机器学习和模式分类方法,其中线性分类器又是各种分类器中用得最多的一种分类器[10~12]。在视觉信息的解码中,不同视觉刺激对应的脑信号被看成是不同的样本。至于样本的特征,可以有很多选择,现在用得最多的就是感兴趣区域中每一个体素的响应值,特征的维数就是与之对应的体素的个数。因此视觉信息的解码研究问题可以看成是一个高维空间的数据分类的问题。

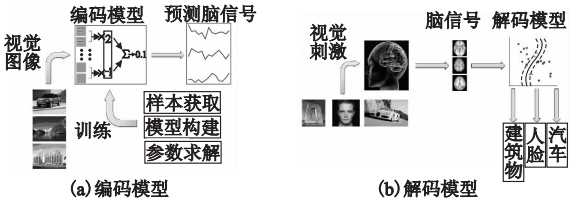

图1 视觉信号的编/解码模型

从图 1 可以看出,虽然编码模型研究的重点是建立计算模型,但是如果利用计算模型得到了预测脑信号,那么通过比较

收稿日期: 2012-09-14; **修回日期**: 2012-10-24 **基金项目**: 国家"863"计划资助项目(2012AA011603)
作者简介: 雷煜(1988-),男,贵州修文人,硕士研究生,主要研究方向为现代检测技术与信号处理(leiy8805@gmail.com)

真实脑信号和预测脑信号之间的关系,也可以实现对大脑信号的解读,也就是说不管是以解码还是编码为重点的研究,都能够实现对人脑视觉信号的解读。

1 基于 fMRI 的视觉信息编码

近几年来,在大脑视觉感知计算模型,也就是编码模型的研究中,取得了一些突破性的进展。2008 年,Mitchell 等人将一个海量语料库中的名词按照语义进行分类,然后给被试者看不同类别的名词,利用采集的全脑 fMRI 信号来训练一个大脑的计算模型,通过这个计算模型成功识别出了被试者看到的新的名词。同年,美国加州大学伯克利分校的神经科学家 Gallant 针对获取的 fMRI 数据,利用金字塔结构 Gabor 小波模型(Gabor wavelet pyramid)建立了从图片到脑激活模式的感受野模型(receptive-field model),成功预测了一个人看到的新图像信息[13]。2009 年,Gallant 团队的研究人员通过在 Gabor 小波金字塔模型中加入贝叶斯方法,并提供一些先验信息,实现了在一个包含六百万张图片的数据库中进行图像检索[14]。在 2011 年 9 月发表的最新文章中,Gallant 的研究团队通过在计算模型中加入运动能量信息,利用 Gabor 小波金字塔模型在动态图像(也就是视频图像)的预测上取得了令人振奋的成果,实现了对动态图像(视频)的重构[15],这是目前为止唯一能够对动态图像进行重构的工作。

基于 fMRI 的视觉信息编码研究,其基础是人脑视觉感知机制,重点是利用计算模型描述人脑的视觉感知活动,下面将分别从两方面介绍现有的研究成果。

1.1 人脑视觉感知机制

人脑视觉信息解读的本质就是对大脑对外界的响应信号的分析,因此,人脑视觉感知的基本原理和机制就是研究的基础。对于视觉信息的认知机制,科学家们提出了很多假说,例如诺贝尔奖获得者 Adrian 提出的同构编码假说[16],Barlow 提出的“老祖母细胞”假说,Abeles[17] 提出的同步发放(Synfire)假说。其中到目前为止最有影响力的假说是 Attneave 和 Barlow[18,19] 提出的有效编码假说。随着人类科技手段的不断进步,有效编码假说不断得到理论和实验证明,逐渐成为了视觉感知领域的一种主流假说。

基于有效编码假说主要有两种研究思路,一种是比较直接的想法,就是在自然刺激的条件下测试响应的统计特性,如 Vinje 等人[20] 认为,对于自然环境中的刺激,视皮层细胞的响应满足稀疏分布。另一种思路则是推测感知系统的信息处理模型:首先检测外界信号的统计特性,然后根据最优化理论推导一个转换工具,从而给出感知系统对外界刺激响应的最佳描述[21]。这种想法的缺点是模型与大脑神经系统的实际工作机理不一致,因为复杂细胞并不是以线性方式从刺激中抽取相位和位移的不变特征的。

上述方法本质上都需要一个提取特征的过程。实际上,在人类的神经系统中,即使不经过特征提取的预处理阶段,也能够处理不变的特征。神经系统能够获得不变特征的性质,已经得到了生物物理学家和神经生理学家的证明。因此,通过模拟人类神经系统的信息处理方式来提取不变特征,已经成为了另外一种解决有效编码问题的途径[22]。

1.2 大脑视觉信息处理的计算模型

编码方法的基本思想就是通过模拟大脑神经系统的信息处理方式来提取图像不变特征,为大脑视觉信息的解读提供依据。目前基于大脑视觉认知机制的信息处理计算模型的研究成果主要有体素感受野模型(voxel receptive-field model)[9,13,23] 和表现形式相似性分析方法(representational similarity analysis, RSA)[24,25]。体素感受野模型通过建立体素水平的计算模型来模拟大脑的信息处理机制,从而突破了传统的模式识别方法在自然图片识别问题上遇到的瓶颈。表现形式相似性分析方法是一种不依赖于数据获取方式的高阶数据分析方法,通过这种方法能够实现多模态数据甚至是跨物种数据的分析[25]。

1.2.1 感受野模型与 Gabor 小波

初级视觉皮层(V1)区的简单细胞的感受野结构和计算模型是近年来编码模型研究的重点。Hubel 和 Wiesel[26,27] 发现当具有一定朝向和宽度的条形刺激出现在简单细胞感受野内某个特定位置时,细胞的响应最强;当刺激偏离该朝向或者位置时,响应急剧下降,甚至消失。也就是说简单细胞对视觉刺激模式呈现出对方向、位置和空间频率选择性。基于此发现,Daugman[28] 提出用 Gabor 函数来模拟简单细胞的感受野结构,其中 Gabor 函数定义为高斯包络函数和正弦谐波函数的乘积,其空间域表示如下:

$$g(x,y) = K\exp\{-\pi[a^2(x-x_0)_r^2 + b^2(y-y_0)_r^2]\} \times \exp\{j[2\pi(u_0x+v_0y)+P]\}$$

其中:K 是归一化参数,P 表示相位[21]。

2008 年,Gallant 等人在自然图片的识别实验中,利用不同相位、空间频率和位置的 Gabor 小波基函数来模拟简单细胞的感受野模型,其所用的 Gabor 小波基函数如图 2 所示。

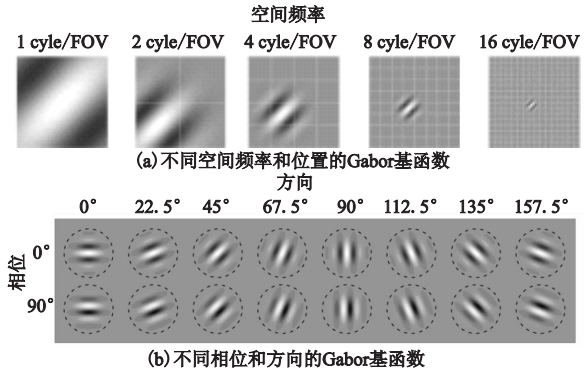

图2 Gabor小波基函数

在这个实验中,将每一个初级视觉皮层区域的体素的感受野模型看成是许多不同空间频率、不同方向、不同位置的 Gabor 基函数的线性组合,如下式所示:

$$r = \sum w^{\mathrm{T}} g + c$$

其中:w 代表权值向量;g 代表不同的 Gabor 基函数构成的一个向量;c 是一个常数偏置。其具体的计算过程如图 3 表示。

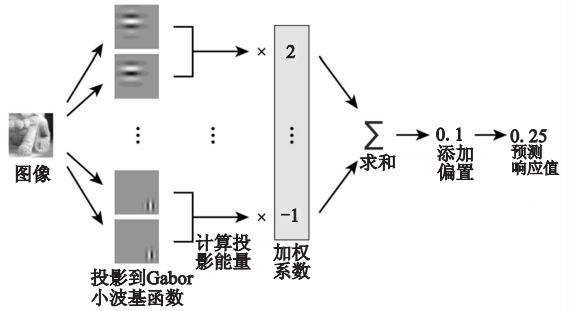

图3 Gabor感受野模型

通过这个模型就可以得到每一个体素对于任何视觉刺激

的模拟响应值,通过将实际测量的对一个视觉刺激的响应与很多模拟的视觉刺激响应进行比较,如计算相关性,就可以找出跟被试实际看到的场景类似的图片。

2009 年,Gallant 等人基于以上工作又在 Gabor 感受野模型中添加了基于贝叶斯理论的语义编码模型,从而使得识别的图片不仅在结构上相似,而且在语义上也极其相似,使得识别的结果更加接近被试真实看到的图片。2011 年,Gallant 的团队通过在 Gabor 小波感受野模型中加入运动能量的信息,使得他们的感受野模型能够对动态图像进行重构,将视觉信息解码研究推向了新的高度。

分析 Gallant 等人近年来取得的成果,不难发现其研究的基础仍然是 Gabor 感受野模型,而这实际上并不能完全描述初级视觉皮层的视觉感知活动,这也是 Gallant 等人研究的不足之处。

1.2.2　形式相似性分析方法

形式相似性分析(RSA)方法是 Kriegeskorte 等人[24] 在 2008 年提出来的一种计算模型。RSA 方法的核心是形式差异性矩阵(representational dissimilarity matrix, RDM)。在 RDM 中,元素 R_{ij} 代表第 i 个刺激和第 j 个刺激对应的脑响应之间的差异性。图 4 为一个四种刺激条件下的简单 RDM。

图4　形式差异性矩阵

从图 4 中可以看出,因为对角线元素对应的是相同刺激的脑响应信号,因此差异性为 0。从图中还可以看出,由于 RDM 是由不同刺激条件引发的脑响应信号的差异性构成,并不涉及大脑具体的响应信息,因此可以在不同模态或者不同模型之间方便地进行比较。

正是因为有了 RDM,就能够引入 RSA 这种新的分析方法。RSA 的核心思想是利用 RDM 作为一个脑区对于刺激的表现形式的一个独特"签名",通过分析 RDM 来分析不同脑区对于相同刺激的不同反应,或者不同计算模型之间的区别等。RSA 方法的应用主要体现在以下四个方面:a)通过计算真实脑信号和预测脑信号的 RDM,可以比较不同大脑计算模型性能的好坏;b)可以通过分析不同脑区、被试、物种和模态的数据,研究它们之间的关系;c)利用 RSA 方法还可以研究大脑信号和行为之间的关系;d)RSA 方法为多刺激条件的事件相关实验提供了一种新的分析方法。

RSA 方法的分析过程为:首先采集刺激对应的大脑信号,然后测量不同刺激对应的大脑信号之间的差异矩阵,再利用计算模型得到预测的不同刺激对应的大脑激活信号的差异性矩阵。对真实信号的 RDM 和预测信号的 RDM 进行分析,并利用随机化方法测试不同 RDM 之间的相关性,最后可以采用多维定标(MDS)的方法将不同 RDM 中的相似结构可视化。

2008 年,Kriegeskorte 等人[25] 利用 RSA 方法分析了人和猴子在看不同类别的图片时,大脑颞下皮层区域(inferior temporal, IT)的响应信号的关系。他们发现对于呈现的相同图片集中的不同类别的图片,人和猴子的颞下皮层区域的响应信号具有很强的相似性,这个实验揭示了一个现象:在灵长类动物中,颞下皮层在分辨不同类别图片的过程中可能发挥了相似的作用。

与 Gabor 感受野模型不同,RSA 方法是一种高阶的分析方法,因此能够在比预测脑信号更高的认知层面上研究视觉感知机制,但是因其不能对脑信号进行直接的预测,限制了该方法在视觉信息的解读中的应用范围。

2　基于 fMRI 的视觉信息解码

与基于 fMRI 的视觉信息编码研究相比,视觉信息的解码研究起步较早,2001 年 Haxby 等人[2] 利用脑功能磁共振成像信号,成功实现了对呈现给被试者的六个不同类别的图像进行分类,并且证明了表示人脸的图像和表示物体的图像在人脑中引起的响应具有分布性和混叠性。2005 年,Kamitani 等人[3] 将不同方向的条纹作为视觉刺激,从 fMRI 信号中实现了对不同条纹对应的脑响应信号进行分类,证明了在初级视觉皮层区域含有外界图像刺激的信息,并且为分析大脑意识的细节信息提供了一种方法框架。2006 年,Thirion 等人[29] 的研究进一步表明可以利用大脑视觉区的脑激活模式重构出被试实际观察到的图像,文献提供了一种基于显式的激励与响应之间计算模型的枕叶初级视皮层的解码机制,并成功识别了脑活动的诱发激励模式,不过它只使用了大脑初级视觉区的激活模式,重构出来的图像分辨率和精度都有待提高。2006 年,Miyawaki 等人[30] 利用 fMRI 信号和多尺度的局部图像基函数来重构视觉刺激图像,他们将每幅刺激图像表示为多尺度的局部图像基元的线性组合,每个局部图像基元的强度由 fMRI 信号通过多体素模式分析来进行预测。最后,所有的局部解码信息按最优参数组合成一幅图像,不过这种方法现在还只能重构简单的黑白几何图像。

在传统的 fMRI 数据分析中,主要是通过将刺激或者行为参数作为回归变量,利用广义线性模型(GLM)研究大脑的功能活动[31]。而在 fMRI 解码中,主要是通过机器学习的方法,同时考虑多个体素对刺激的响应,来预测被试的行为或者认知状态,这种方法也被称为多体素分析方法(multi-voxel pattern analysis, MVPA)。

2.1　多体素分析方法

在 fMRI 信号的解码中,多体素分析方法占据着重要的位置[6]。多体素分析方法把大脑中的一个或者几个体素看成一个特征,将每一个刺激对应的大脑响应看成是一个高维空间的数据,并且同时考虑多个体素对任务条件的贡献,然后利用模式分类的方法研究刺激和响应之间的关系。与传统的基于统计检验的单体素分析方法(如 GLM 模型)[32] 相比,多体素分析方法能挖掘更多的体素之间的互信息,因此更加有利于进行大脑信号的解读。其基本过程是:a)获得刺激对应的脑响应信号,并将数据分为训练数据集和测试数据集;b)确定要采用的分类器;c)在训练数据集中,每一个数据对应一个外部刺激的类别,利用这些数据训练给定的分类器;d)将测试数据应用于训练得到的分类器,将分类器的分类结果和真实结果比较,测试分类器的性能。其具体过程如图 5[33] 所示。

多体素分析方法的主要优点包括以下几点[6]:

a)对认知状态的变化更加敏感。传统的单体素分析方法为了提高体素对特定任务的敏感性,通常会对体素进行空间平滑,这样虽然提高了信号的信噪比,但是也会去除某些携带少量刺激信息的体素(也就是在单体素分析中并不显著的体素)

中含有的有用信息。此外,空间平滑还会“污染”一些对不同刺激条件具有可分性的体素[34]。然而,因为多体素分析方法考虑的是多个体素的响应,不用对体素进行空间平滑,能够保留更多的关于外界刺激的原始信息。

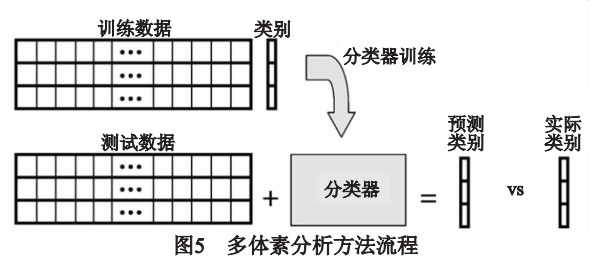

图5 多体素分析方法流程

b)能够将大脑活动状态和外部行为联系起来。多体素分析方法提供的高敏感性使得人们能够在较短的时间内判断大脑中某一种认知状态是否存在,而大脑的认知状态是与外部行为密切相关的,因此通过短时间的大脑扫描数据就能区分对应的外部行为[35~37]。此外,通过多体素分析方法计算不同行为和大脑体素信号之间的相关性,这也是多体素方法最重要的优点之一。

c)能够描述神经编码的结构。通过多体素分析方法,可以在一个特征集中任意地去除或者添加一些体素,从而观察体素的变化和任务条件之间的关系。通过这种分析,就能够发现特定体素集合和任务条件之间的联系,从而在一定程度上解释大脑的认知机制。

针对多体素分析方法,现在已经有一些较为成熟的软件包可以使用。在 MATLAB 环境下可以使用普林斯顿大学开发的 Princeton multi-voxel pattern analysis 工具箱(http://code.google.com/p/princeton-mvpa-toolbox/),或者使用一个由 Python 语言写成的 PyMVPA 工具包(http://www.pymvpa.org/)[38]。

2.2 视觉信息解码的分类器设计

目前常用于解码研究的分类器主要分为线性分类器和非线性分类器两类。线性分类器主要有最近邻分类器(Cor)、线性判决分析法(LDA)和线性支持向量机(SVM-lin);非线性分类器主要有 K 近邻分类器(KNN)、高斯贝叶斯判别法(GNB)和径向基函数支持向量机(SVM-RBF)。其主要区别是构成分类器的判决面是否是线性的,如图 6[39]所示。

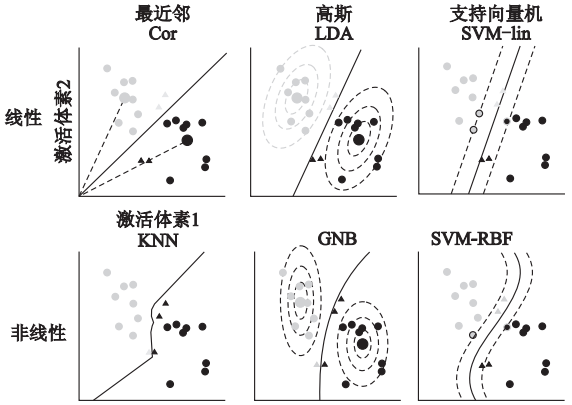

图6 不同分类器的判决面

上面提到的几种分类器都是从统计学或者机器学习领域中发展而来的很成熟的分类器,每种分类器的分类性能好坏与数据自身的结构有很大关系,因此具体使用哪种分类器就取决于对于 fMRI 信号的不同假设:如果假设刺激和响应之间具有线性关系,并且每种刺激对应的体素响应具有高斯分布的特点,那么采用 LDA 分类器的效果就会比较好;然而,如果仍然假设刺激和响应之间具有线性,但是刺激引起的体素响应不满足高斯分布的性质,那么 SVM-lin 的性能应该就会优于 LDA。Mitchell 等人[40]比较了 GNB、SVM 和 KNN 三种分类器在三个不同的数据集上的分类性能,得到的结论是 GNB 和 SVM 都能实现较好的分类效果,KNN 的分类正确率比较低[41]。Ku 等人[41]在 7 TB 的高场数据上比较了 LDA、Cor、GNB 和 SVM-lin 四种分类器的性能,得出的结论是三种线性分类器的性能不相上下,但是 GNB 的分类正确率要明显低于三种线性分类器。

如果一个分类器的判决面能够更好地拟合训练数据,那么这个判决面对于训练数据就会具有更好的可分性,但是由此带来的一个问题:由于 fMRI 实验通常采集的样本数都不会太多,而分析的体素一般都是几千个,从而导致相对于特征维数来说样本数严重不足,使得分类器的判决面产生过拟合现象,反而降低了分类器的泛化性能。相对于非线性分类器的非线性判决面,线性分类器的线性判决面能够在一定程度上克服过拟合现象的影响,这也是目前在大脑信号的解码研究中线性分类器的性能普遍优于非线性分类器的原因之一。当然,近些年来也有一些研究者在大脑信号的解码研究中使用非线性分类器[42~44],不过除了过拟合的影响之外,非线性分类器还有另一个较大的缺点,即不具有解释性。因为理论上来说至少存在一个非线性分类器能够从一个大脑区域中将需要的任何信息进行正确分类,也就是说对于一些不是明显包含在一个感兴趣区域(ROI)中的信息,非线性分类器也能够对其进行成功分类,这就使得结果的解释性大大降低[45]。例如可以给被试者分别呈现人脸和房屋两个类别的图像,并记录被试者初级视觉皮层的活动,然后会发现,利用非线性分类器可以实现对这两个类别图片的正确分类,虽然实际上初级视觉皮层区域并不含有与这两类图片的语义相关的信息。

由于过拟合现象对分类器的性能会产生很大的影响,如何减小过拟合是大脑信号解码研究的一个重要课题。其中的方法之一是减少潜在的判决面的数目,或者通过正则化方法来减少过拟合[46,47]。对数据进行降维处理是另一个常用的减少过拟合的方法,其中最简单的降维方法就是体素选择[40],在实际的数据处理过程中,多变量体素选择方法的应用非常普遍[10,34]。

鉴于体素选择在 fMRI 脑信号解码研究中的重要地位,很多研究者都对这个问题进行了深入研究[33,48],并提出了以下几种方法:

a)基于激活程度的体素选择方法。通过这种方法选取的体素至少要对一种刺激条件产生活跃反应。其具体做法是在不同的刺激条件下获得体素的平均激活程度,并与基线水平比较,采用 T 检验或者其他统计检验方法获得激活的体素。

b)基于分类正确率的体素选择方法。这种方法以一个体素为特征,利用训练数据训练产生分类器,并用交叉检验的方法获得分类正确率;然后以这个分类正确率作为是否选取该体素作为特征的指标。

c)Searchlight 方法。跟基于正确率的体素选择方法类似,不过这里不再用一个体素作为特征,而是以一个体素和与它在三维空间相邻的一些体素作为一个特征。总之该方法是选取一个相邻区域内的体素作为一个体素集,然后也是通过分类正确率决定是否选择该体素集作为特征集。

d)方差分析(ANOVA)。这种方法是寻找那些对于不同刺激条件响应有显著不同的体素。其零假设为不同刺激条件下体素值没有显著差别。在这种方法中,可以给每个体素一个得分,体素的得分就是 p 值,因此值越低越好。

e)稳定性分析方法。这种方法是利用交叉检验来评价不同训练数据下同一个体素对于不同测试集数据的预测稳定性。例如,假设有三个不同类别刺激 A、B、C,对于一组训练集来说,分类器的输出值关系是 $A > B > C$,那么就要找能在不同的训练数据集上依然能得到保持这样的大小关系的那些体素,并认为这样的体素具有良好的预测稳定性。

除了体素选择之外,一种类似的方法是特征提取,常用的特征提取方法包括奇异值分解(SVD)、主成分分析(PCA)和独立成分分析(ICA)等。

3　结束语

作为一个新兴的研究领域,人脑视觉信息的解读研究才刚刚起步,本文旨在介绍该领域的最新研究进展和主流研究方法。从分析中可以看出,了解人类视觉感知机制是实现人脑视觉信息解读的基础,随着神经影像技术的不断进步,视觉感知机制的研究将迎来新的时代。在视觉信息的编码模型研究中,目前还未形成较系统的理论,对于初级视觉皮层区域的神经元的视觉感知机制,即使是被广泛认可的 Gabor 感受野模型也只能解释其中的很小一部分[49],因此寻找能够更好地刻画初级视觉皮层视觉感知机制的计算模型的工作仍将是研究的重点。在视觉信息解码研究方面,合适的降维方法、分类器的选择仍然是很有潜力的研究方向。在利用体素空间激活模式的同时加入体素的时域变化信息进行解码也开始得到研究人员的关注。此外,通过提升算法的速度,实现实时的视觉信息解读将成为今后基于 fMRI 的脑机接口研究的热点。

参考文献:

[1] MARR D, VAINA L. Representation and recognition of the movements of shapes[J]. Proceedings of the Royal Society of London Series B, Biological Sciences,1982,214(1197):501.

[2] HAXBY J V, GOBBINI M I, FUREY M L, *et al*. Distributed and overlapping representations of faces and objects in ventral temporal cortex[J]. Science,2001,293(5539): 2425-2430.

[3] KAMITANI Y, TONG F. Decoding the visual and subjective contents of the human brain[J]. Nature Neuroscience,2005,8(5):679-685.

[4] KRIEGESKORTE N, BANDETTINI P. Analyzing for information, not activation, to exploit high-resolution fMRI[J]. Neuroimage, 2007,38(4):649-662.

[5] LaCONTE S M. Decoding fMRI brain states in real-time[J]. Neuroimage,2011,56(2):440-454.

[6] NORMAN K A, POLYN S M, DETRE G J, *et al*. Beyond mind-reading: multi-voxel pattern analysis of f MRI data[J]. Trends in Cognitive Sciences,2006,10(9):424-430.

[7] Van GERVEN M A J, KOK P, De LANGE F P, *et al*. Dynamic decoding of ongoing perception[J]. Neuroimage,2011,57(3): 950-957.

[8] HAYNES J D, REES G. Decoding mental states from brain activity in humans[J]. Nature Reviews Neuroscience, 2006, 7(7): 523-534.

[9] MITCHELL T M, SHINKAREVA S V, CARLSON A, *et al*. Predicting human brain activity associated with the meanings of nouns[J]. Science,2008,320(5880): 1191-1195.

[10] De MARTINO F, VALENTE G, STAEREN N, *et al*. Combining multivariate voxel selection and support vector machines for mapping and classification of fMRI spatial patterns[J]. Neuroimage,2008, 43(1):44-58.

[11] HANSEN L K. Multivariate strategies in functional magnetic resonance imaging[J]. Brain and Language,2007,102(2):186-191.

[12] HAYNES J D. Decoding visual consciousness from human brain signals[J]. Trends in Cognitive Sciences,2009,13(5):194-202.

[13] KAY K N, NASELARIS T, PRENGER R J, *et al*. Identifying natural images from human brain activity [J]. Nature, 2008, 452 (7185):352-355.

[14] NASELARIS T, PRENGER R J, KAY K N, *et al*. Bayesian reconstruction of natural images from human brain activity[J]. Neuron, 2009,63(6):902-915.

[15] NISHIMOTO S, VU A T, NASELARIS T, *et al*. Reconstructing visual experiences from brain activity evoked by natural movies[J]. Current Biology,2011,21(19): 1641-1646.

[16] ADRIAN E D. The impulses produced by sensory nerve endings[J]. Journal of Physiology,1926,61(1):49-72.

[17] ABELES M. Local cortical circuits: an electrophysiological study [M]. New York: Springer-Verlag,1982.

[18] ATTNEAVE F. Some informational aspects of visual perception[J]. Psychological Review,1954,61(3):183.

[19] BARLOW H B. Possible principles underlying the transformation of sensory messages [M]//Sensory Communication. Combridge: MIT Press,1961:217-234.

[20] VINJE W E, GALLANT J L. Sparse coding and decorrelation in primary visual cortex during natural vision[J]. Science, 2000, 287 (5456):1273-1276.

[21] 罗四维. 视觉信息认知计算理论[M]. 北京:科学出版社, 2010.

[22] 罗四维. 视觉感知系统信息处理理论[M]. 北京:电子工业出版社, 2006.

[23] DUMOULIN S O, WANDELL B A. Population receptive field estimates in human visual cortex[J]. Neuroimage,2008,39(2):647-660.

[24] KRIEGESKORTE N, MUR M, BANDETTINI P. Representational similarity analysis-connecting the branches of systems neuroscience [J]. Frontiers in Systems Neuroscience,2008,2(11):4.

[25] KRIEGESKORTE N, MUR M, RUFF D A, *et al*. Matching categorical object representations in inferior temporal cortex of man and monkey[J]. Neuron,2008,60(6):1126-1141.

[26] HUBEL D H, WIESEL T N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex[J]. Journal of Physiology,1962,160(1):106-154.

[27] HUIBEL D H, WIESEL T N. Receptive fields and functional architecture of monkey striate cortex[J]. Journal of Physiology,1968, 195(1):215-243.

[28] DAUGMAN J G. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters[J]. Journal of Optical Society of America A: Optics and Image Science,1985,2(7):1160-1169.

[29] THIRION B, DUCHESNAY E, HUBBARD E, *et al*. Inverse retinotopy: inferring the visual content of images from brain activation patterns[J]. Neuroimage,2006,33(4):1104-1116.

[30] MIYAWAKI Y, UCHIDA H, YAMASHITA O, *et al*. Visual image reconstruction from human brain activity using a combination of multi-

scale local image decoders[J]. Neuron,2008,60(5):915-929.

[31] WORSLEY K, LIAO C, ASTON J, *et al*. A general statistical analysis for fMRI data[J]. Neuroimage,2002,15(1):1-15.

[32] WORSLEY K J, FRISTON K J. Analysis of fMRI time-series revisited: again[J]. Neuroimage,1995,2(3):173-181.

[33] PEREIRA F, MITCHELL T, BOTVINICK M. Machine learning classifiers and fMRI: a tutorial overview[J]. Neuroimage,2009,45(1): 199-209.

[34] KRIEGESKORTE N, GOEBEL R, BANDETTINI P. Information-based functional brain mapping[J]. Proceedings of the National Academy of Sciences of the USA,2006,103(10):3863-3868.

[35] POLYN S M, NATU V S, COHEN J D, *et al*. Category-specific cortical activity precedes retrieval during memory search[J]. Science, 2005,310(5756):1963-1966.

[36] O'TOOLE A J, JIANG F, ABDI H, *et al*. Partially distributed representations of objects and faces in ventral temporal cortex[J]. Journal of Cognitive Neuroscience,2005,17(4):580-590.

[37] CARLSON T A, SCHRATER P, HE S. Patterns of activity in the categorical representations of objects[J]. Journal of Cognitive Neuroscience,2003,15(5):704-717.

[38] HANKE M, HALCHENKO Y O, SEDERBERG P B, *et al*. PyMVPA: a python toolbox for multivariate pattern analysis of fMRI data [J]. Neuroinformatics,2009,7(1):37-53.

[39] MISAKI M, KIM Y, BANDETTINI PA, *et al*. Comparison of multivariate classifiers and response normalizations for pattern-information fMRI[J]. Neuroimage,2010,53(1):103-118.

[40] MITCHELL T M, HUTCHINSON R, NICULESCU R S, *et al*. Learning to decode cognitive states from brain images[J]. Machine Learning,2004,57(1):145-175.

[41] KU S, GRETTON A, MACKE J, *et al*. Comparison of pattern recognition methods in classifying high-resolution bold signals obtained at high magnetic field in monkeys[J]. Magnetic Resonance Imaging,2008,26(7):1007-1014.

[42] DAVATZIKOS C, RUPAREL K, FAN Y, *et al*. Classifying spatial patterns of brain activity with machine learning methods: application to lie detection[J]. Neuroimage,2005,28(3):663-668.

[43] HANSON S J, MATSUKA T, HAXBY J V. Combinatorial codes in ventral temporal lobe for object recognition: Haxby (2001) revisited: is there a "face" area? [J]. Neuroimage,2004,23(1):156-166.

[44] COX D D, SAVOY R L. Functional magnetic resonance imaging (fMRI)[J]. Neuroimage,2003,19(2):261-270.

[45] NASELARIS T, KAY K N, NISHIMOTO S, *et al*. Encoding and decoding in fMRI[J]. Neuroimage,2011,56(2):400-410.

[46] HASTIE T, TIBSHIRANI R, FRIEDMAN J. The elements of statistical learning: data mining, inference, and prediction[M]. Beijing: Publishing House of Electronics Industry, 2004.

[47] KRISHNAPURAM B, CARIN L, FIGUEIREDO MAT, *et al*. Sparse multinomial logistic regression: fast algorithms and generalization bounds[J]. IEEE Trans on Pattern Analysis and Machine Intelligence,2005,27(6):957-968.

[48] PUJARA J. Understanding feature selection in functional magnetic resonance imaging [D]. Pittsburgh: Carnegie Mellon University, 2005.

[49] OLSHAUSEN B A, FIELD D J. What is the other 85% of V1 doing [M]//Problems in Systems Neuroscience. Oxford: Oxford University Press,2004:182-211.

(上接第 1610 页)

[24] 金晓鸥,钟宝燕,李翔. 基于 Rhino 的 JavaScript 动态页面解析研究实现[J]. 计算机技术与发展,2008,18(2):1-4,50.

[25] Mozilla. XULRunner[EB/OL]. [2010-10-10]. https://developer.mozilla.org/en/XULRunner.

[26] Webclient[EB/OL]. (2007-09-23). [2010-02-16]. http://www.mozilla.org/projects/blackwood/webclient/.

[27] Watir[EB/OL]. [2009-11-16]. http://watir.com/.

[28] rbNarcissus[EB/OL]. [2010-03-20]. http://code.google.com/p/rbnarcissus/.

[29] 罗兵. 支持 Ajax 的互联网搜索引擎爬虫设计与实现[D]. 杭州:浙江大学,2007.

[30] 肖卓磊. 基于 Ajax 技术的搜索引擎研究[D]. 武汉:武汉理工大学,2009.

[31] 战琴. 基于 Ajax 技术的 deep Web 爬虫实现方法研究[D]. 济南:山东科技大学,2009.

[32] 曾伟辉,李淼. 基于 JavaScript 切片的 Ajax 框架网络爬虫技术研究[J]. 计算机系统应用,2009,18(7):169-171.

[33] 袁小节. 基于协议驱动与事件驱动的综合聚焦爬虫研究与实现[D]. 长沙:国防科学技术大学研究生院,2009.

[34] REIS D C, GOLGHER P B, SILVA A S, *et al*. Automatic Web news extraction using tree edit distance[C]// Proc of the 13th International Conference on World Wide Web. New York: ACM Press, 2004:502-511.

[35] Levenshtein distance[EB/OL]. [2008-08-14]. http://en.wikipedia.org/wiki/Levenshtein_distance.

[36] MARZAL A, VIDAL E. Computation of normalized edit distance and applications[J]. IEEE Trans on Pattern Analysis and Machine Intelligence,1993,15(9):926-932.

[37] BUTTLER D. A short survey of document structure similarity algorithm[C]//Proc of the 5th International Conference on Internet Computing. 2004: 3-9.

[38] W3C. XPath 2.0 [EB/OL]. [2010-12-14]. http://www.w3.org/TR/xpath20/.

[39] GitHub. OXPath[EB/OL]. [2011-11-17]. https://github.com/diadem/OXPath.

[40] WebRenderer[EB/OL]. [2010-02-16]. http://www.webrenderer.com/.

[41] JRex: the Java browser component[EB/OL]. [2009-06-21]. http://jrex.mozdev.org/.

[42] JExplorer[EB/OL]. [2010-01-29]. http://www.Leamdev.com/jexplorer/.

[43] Watij[EB/OL]. [2009-11-16]. http://watij.com/.

[44] HtmlUnit[EB/OL]. [2010-02-09]. http://htmlunit.sourceforge.net/.

[45] XHTML renderer project[EB/OL]. [2009-07-01]. https://xhtml-renderer.dev.java.net/.

[46] CSS parser[EB/OL]. [2009-11-16]. http://cssparser.sourceforge.net/.

[47] ActiveXML documentation [EB/OL]. http://webdam.inria.fr/axml/index.axml.html

[48] Crowbar[EB/OL]. [2010-01-16]. http://simile.mit.edu/wiki/Crowbar.

[49] FireWatir[EB/OL]. [2010-01-14]. http://code.google.com/p/firewater/.