

上海交通大学

SHANGHAI JIAO TONG UNIVERSITY

学士学位论文

THESIS OF BACHELOR



论文题目： 基于EEG的性别分类

学生姓名： 顾中磊

学生学号： 5050309900

专 业： 计算机科学与技术

指导教师： 吕宝粮

学院(系)： 电子信息与电气工程学院

基于 EEG 的性别分类

摘要

人对于人脸的判断如同天赋一般，无须刻意的训练，就能够快速的通过脸部特征辨别性别、年龄、种族和基本情感。人脸识别是通过对图像或视频流进行相关图像处理及数学计算，并结合机器学习进行模式识别或分类。在心理学领域，随着对脑科学研究的深入，也试图使用 fMRI (功能核磁共振) 和 EEG (脑电) 的方式从本质上去探索人脑的工作方式。本文将心理学的研究方法与计算机科学相结合，试图对人进行性别判断时的 EEG 进行分类。

在设计并完成的 EEG 实验中，被试被要求在接收到短时间(300ms)的人脸刺激后，隐式或显式的判定人脸的性别。在实验后对实验数据分别进行时域分析、功率谱分析和小波分析，并利用 SVM 对抽取的特征进行分类，分类精度最高达到 68%。

本文取得的成果如下：

- (1) 验证了(Y. Mouchetant-Rostaing et al. 2000)^[2]、(Seeck et al., 1997)中提到的与性别相关的三个时间区间，分别为刺激后 50-90ms、145-185ms、200-250ms。
- (2) 提出关于大脑性别处理三个阶段的假设：
 - a) 50-90ms 阶段为性别处理无意识阶段，在此阶段大脑初步处理视觉信息，附带性别信息，但未影响到大脑有意识区域；
 - b) 145-185ms 阶段为结构编码阶段，将性别信息作为结构编码一部分；
 - c) 200-250ms 阶段为性别处理有意识阶段，人能够获取性别信息，并做出判断。
- (3) 根据实验数据的分析，大脑在性别处理过程中，各人之间有“个性化”存在，即每个人进行性别判断的 EEG 强弱存在差异，脑分布也存在差异。

关键词：EEG, 性别判断, 功率谱, 小波, SVM

Gender Classification by EEG

ABSTRACT

Human have a talent on face judging which enable us to tell gender, age, race and basic emotion swiftly without explicit training. Face Recognition commonly process with images or video stream to extract characteristics, and uses machine learning to recognize pattern and classify. In psychology field, with the development of brain science, methods like fMRI and EEG were tried to explore brain's working naturally. This article combine the method of psychology and computer science, and try to classify the EEG while people judging face's gender.

EEGs were recorded while subjects were involved in eight gender-processing tasks based on human faces and cartoon faces. During the tasks, subjects were asked to decide whether stimulus is male or not. Time-Domain, Power Spectrum and Wavelet was applied in data analysis, and the result was used to train the SVM whose result reach a accuracy of 68% max.

The major results are as follows:

- (1) Conclusion was reached which confirm the recent conclusion (Y.Mouchetant-Rostaing et al. 2000) ^[2] that gender-related time interval was 45-85ms, 145-185ms and 200-250ms. However, early signs of face processing in humans is at [50, 90] time interval.
- (2) Assumption of three period of gender-processing was established.
 - a) The time interval 50-90ms was No-Intentional period. In the period, brain process the perceptual input with gender information, but the information aren't reach intentional period of brain.
 - b) The time interval 145-185ms was structure-encoding period which means gender information was encoding to brain structure.
 - c) The timer interval 200-250ms was intentional period. The gender information is available for man and make a judge.
- (3) Other assumption about personality of gender-processing indicates that everyone's EEG is different, especially for his distribution or structure of brain.

Key words: EEG, Gender Processing, Gender Classification, Power Spectrum, Wavelet, SVM

目 录

第一章 绪论	1
1.1 研究背景	1
1.2 研究现状	1
1.3 研究意义	2
1.4 工作简介	2
1.5 结构安排	2
1.6 本章小结	2
第二章 背景知识	3
2.1 假设检验	3
2.2 T检验	3
2.3 功率谱	4
2.4 小波分析	4
2.4.1 连续小波变换 (CWT)	4
2.4.2 离散小波变换	5
2.5 支持向量机 (Support Vector Machine, SVM)	6
2.6 本章小结	6
第三章 脑电与ERP分析技术	7
3.1 实验相关术语	7
3.1.1 大脑的分区	7
3.1.2 脑电的基本特征	7
3.1.3 电极的分布	7
3.1.4 ERP实验	8
3.2 ERP实验分析技术	8
3.2.1 数据记录格式	8
3.2.2 ERP波形的叠加	8
3.2.3 独立成分分析 (ICA)	8
3.3 本章小结	9
第四章 实验设计	10
4.1 被试	10
4.2 刺激方式	10
4.2.1 实验一至实验四	10
4.2.2 实验五至实验八	11
4.3 实验过程	11
4.3.1 训练过程	11
4.3.2 实验一至实验四	11
4.3.3 实验五至实验八	12
4.4 实验设计的依据说明	12
4.4.1 灰度图片的使用	12
4.4.2 卡通图片的使用	12

4.4.3 间隔的随机	12
4.4.4 15 秒的休息	12
4.5 实验设计目的	12
4.6 EEG记录参数	13
4.7 本章小结	13
第五章 实验数据分析	14
5.1 分析前提	14
5.2 时域分析	14
5.2.1 检验男女刺激波形幅值的大小关系	15
5.2.2 检验部分电极的显著性差异	19
5.3 功率谱分析	19
5.3.1 检验各个刺激间功率谱的差	20
5.4 小波分析	21
5.4.1 小波系数分析(连续小波分析)	21
5.5 利用特征进行分类	22
5.5.1 利用时域信息进行分类	23
5.5.2 利用功率谱信息进行分类	25
5.5.3 结合时域与功率谱信息进行分类	25
5.6 本章小结	26
第六章 结论及下一步工作	27
6.1 结论	27
6.2 下一步工作	27
谢辞	30
译文及原文	31

第一章 绪论

1.1 研究背景

人脸识别是从数字图像或视频流中的人脸信息中自动识别和确认身份的技术,是计算机领域中重要的应用。对于正常人,这项技能如同天赋,无须刻意的训练,经过视觉接收和记忆检索,就能够准确的指出身份。同时,能够在更短时间内轻松地通过脸部特征辨别性别、年龄、种族和基本情绪,并能够判断脸部特征是否熟悉。因此,心理学领域也常以此为突破口进而研究人脑。

通过脸部特征人类足以分辨大部分的人脸性别。一直以来,科学研究也试图从各方面去了解和实现人脑这项功能,人脸识别主要通过对图像或视频流进行相关图像处理及数学计算,并结合机器学习进行模式识别或分类,但并非从人脑的角度来考虑性别分类问题。而对于心理学领域,随着 EEG(Electroencephalographic)即脑电信号研究的不断发展,更倾向于从大脑的电生理活动来观察这项功能的本质。

1.2 研究现状

对于性别分类,包括计算机科学、心理学和神经科学都有较为深入的研究。

在计算机科学领域,基于多角度人脸图像的性别分类是一项具有挑战性的研究课题,上海交通大学“基于 Gabor 二值映射和 SVM 的性别分类”研究结果使得在在相对短时间内,平均分类准确率达到了 95%^[4]。

在心理学和神经科学领域,针对面貌不能识别的病患的研究,提出了理论模型与假设。

当头部遭受损伤或神经疾病时,人会引起视觉识别的障碍,包括面貌不能识别。研究表明,绝大多数面貌不能识别的病患能够分辨不同表情,而能够识别人脸的病患却不能分辨表情;并且面貌不能识别的病患能够辨别年龄、性别、种族等特征,从而验证了 Bruce 和 Young 在 1986 年提出了人脸识别的理论模型,如图 1 所示。

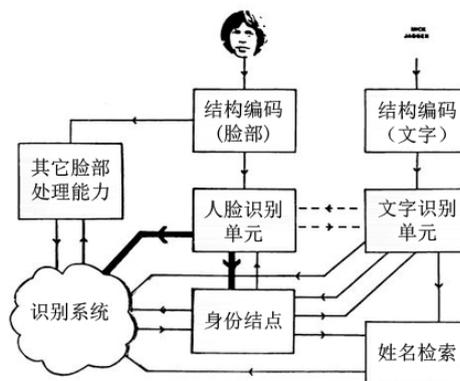


图 1 Bruce 和 Young 的人脸识别理论模型框图

而对于 Bruce 和 Young 的定义,年龄、性别、种族信息属于其它脸部信息的处理能力,因此,人脸识别与有关年龄、性别、种族信息的处理是并行的不相关的两个过程。

而最近的研究结果(Y. Mouchetant-Rostaing et al. 2000)表明, N170(ERP 实验中, 刺激后 170ms 后出现的负峰值)与性别信息处理没有联系; 对手部与脸部的判断时, 在 45-85ms 处波形有明显差异, 对脸在 145-185ms 也有差异。而结合年龄(Y. Mouchetant-Rostaing et al. 2003)^[11], 在 45-85ms、145-185ms、200-400ms 也存在差异。

但另有研究结果表明(Seeck et al., 1997), 人脑处理人脸信息的早期迹象在 50-90ms, 因此, 需要对 45-90ms 的时间区间进行进一步的验证。

尽管如此, 这些结果仅仅涉及在有性别任务与无性别任务之间, 大脑所诱发的 ERP 波形有显著的差异, 并未涉及辨认男性与女性时波形的差异。

1.3 研究意义

本文以大脑对人脸图像刺激的反映作为切入点, 观察人在看到人脸图像刺激时所诱发的 EEG, 试图从中找出人脑对于性别分类的规律, 并在规律的基础上分析研究人脑在进行视觉处理、性别分类的机制。

同时以对于人脸图像刺激的性别分类出发, 分析研究人类视觉系统的构成。

1.4 工作简介

本文设计并准备了有关性别分类的 EEG 脑电实验, 并组织了三名被试完成实验, 得到三名被试各八个实验, 共 24 组数据。并以 24 组数据为基础, 尝试了 ERP 分析(即时域特征分析)、功率谱特征分析、小波分析等分析手段, 经过特征抽取和选择后, 再利用 SVM(支持向量机)对特征进行分类。

1.5 结构安排

本论文分为六章, 各章的安排如下:

第一章: 介绍本文的研究背景、研究意义和所进行的工作简介。

第二章: 介绍文中讨论需要的数学基础知识, 便于后续内容的理解。

第三章: 介绍有关 EEG 的知识, 包括 ERP 实验以及处理信号时的数学工具。

第四章: 介绍本文设计并组织进行的 EEG 实验, 包括刺激类型、实验详细过程等。

第五章: 叙述了实验后对脑电数据所进行的处理, 并从中得到的相关结论。

第六章: 基于数据处理, 讨论实验数据分析结果, 并提出未来的研究计划。

1.6 本章小结

本章讨论了总体背景及研究意义, 介绍了进行研究的主要方法和各章节的安排

第二章 背景知识

2.1 假设检验

假设检验，它的作用在于，在总体的分布函数完全未知或只知其形式、但不知其参数的情况，为了推断总体的某些未知特性，提出某些关于总体的假设^[3]。

在研究中，使用假设检验来判断所观察到的结果是否是有统计学意义。而对于某一结果的假设通常有两种表述：

- (1) 研究期望的结果，如新药对于疾病有效。
- (2) 研究期望的结果的逻辑非，即新药对于疾病一点效果没有。

后者通常被称为“零假设”，记作 H_0 。而前者一般称为“实验假设”，记作 H_1 。

在进行假设检验时，首先选择合适的假设检验方法，即分布函数，并设定显著性水平 α ，然后定义 $P\{\text{拒绝 } H_0 | H_0 \text{ 为真}\} = \alpha$ 为拒绝域，最后设计合适的统计量并计算，若计算结果在拒绝域内，则 H_0 为假，即 H_1 为真，否则 H_0 为真

2.2 T 检验

T 检验（全称“学生 T 检验”，Students' T-Test），它通常用来回答形如“这两组变量是否有关”等问题。

例如：有两类音乐，假设是古典和摇滚，是否这两类音乐会影响大学生完成需要高度注意力集中的任务呢？从生活经验中得到的答案应当是肯定的，但当设计实验来证明时，就 T 检验来证明一组学生的表现优于另一组学生。

T 检验是假设检验的一种，检验两样本之间均值与方差的显著性水平。

假设两样本的样本容量分别为 N_a 与 N_b ，样本均值分别为 M_a 与 M_b ，样本方差为 S_a^2 与 S_b^2

设 $H_0: \mu_M - M = \delta$ ，其中显著性水平为 α ，则 T 检验需要得到的拒绝域。

而 T 检验的统计量为

$$t = \frac{(\overline{M}_a - \overline{M}_b) - \delta}{S_\omega \sqrt{\frac{1}{N_a} + \frac{1}{N_b}}} \quad (2-1)$$

其中

$$S_\omega^2 = \frac{(N_a - 1)S_a^2 + (N_b - 1)S_b^2}{N_a + N_b - 2} \quad (2-2)$$

$$S_\omega = \sqrt{S_\omega^2} \quad (2-3)$$

从而得到 T 检验的拒绝域为

$$|t| = \frac{|(\overline{X} - \overline{Y}) - \delta|}{S_\omega \left(\sqrt{\frac{1}{N_a} + \frac{1}{N_b}} \right)} \geq t_{\frac{\alpha}{2}}(N_a + N_b - 2) \quad (2-4)$$

2.3 功率谱

实时采样得到的数据，包括脑电数据，都是离散的点，且一般是非正弦函数，而任何有限时间内的波形均可视为连续的波形，且实时采样数据满足 Dirichlet 条件，即函数在 $(-\frac{T}{2}, \frac{T}{2})$ 上满足(1)连续或只有有限个第一类间断点，(2)只有有限个极值点^[5]。

则 $f(t)$ 就可以展开成傅氏级数：

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos n\omega t + b_n \sin n\omega t) \quad (2-5)$$

其中，

$$\omega = \frac{2\pi}{T} \quad (2-6)$$

令，

$$n\omega = \omega_n = \frac{2n\pi}{T} \quad (2-7)$$

从而，

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos \omega_n t + b_n \sin \omega_n t) \quad (2-8)$$

由此可以看出，Fourier 变换实质是将时域波形，通过线性变换，使用 sin 和 cos 函数叠加形成原来的波形，而频域的离散函数系数满足：

$$A_n = 2|c_n| = \sqrt{a_n^2 + b_n^2} \quad (2-9)$$

这样通过离散傅里叶变换得到的函数记作 $F(\omega)$ 。则功率谱密度为：

$$S(\omega) = |F(\omega)|^2 = F(\omega)\overline{F(\omega)} \quad (2-10)$$

2.4 小波分析

实时采样数据的采集均在时域内进行，可以很清楚的知道在某个时间点发生了什么事；同样，进行 Fourier 变换之后，在频域数据中，可以很清楚的知道在某个频率发生了什么事。但此时就出现了如同量子物理中的海森堡不确定性原理，即单看时域和频域无法知道某个时间点在某个特定频率的具体情况。

而小波变换恰能一定程度上解决这个问题。

如果说 Fourier 变换是使用 sin 和 cos 函数叠加形成原来的波形，则小波变换是使用特定的小波函数，通过拉伸、压缩和平移形成原来的波形。而特定的小波波形是有限的，非对称的波形。

2.4.1 连续小波变换 (CWT)

连续小波变换的一个重要的应用就是小波系数^[6]。

连续小波变换可以看作将小波函数以每个点为起始，在任意尺度上求与原函数的相似程度，系数越大，相似程度越高。

而尺度的大小就决定了频率的大小，当小波函数被压缩到小尺度时，变化的频率就变大，这样与原函数对比所反应出的信息就是信号的高频信息，反之则是低频信息。

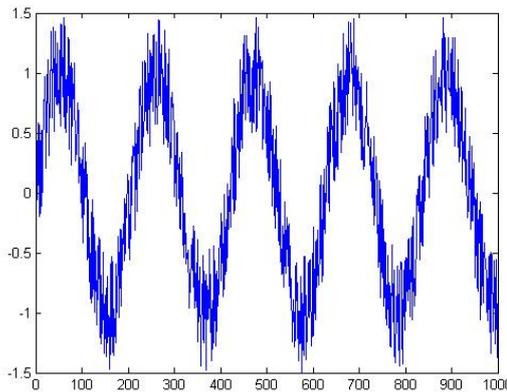


图 2 原始波形

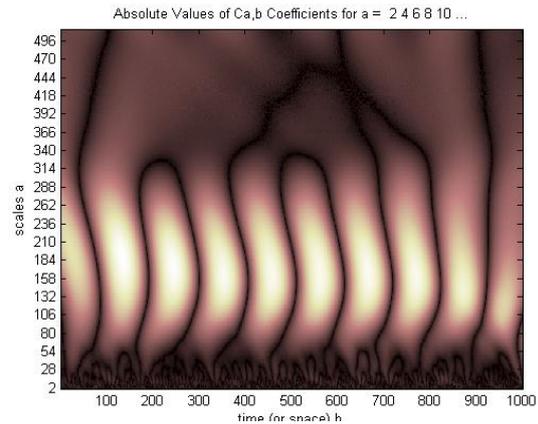


图 3 小波系数图

如图 2 所示，为原始信号波形，图 3 是使用 Daubechies4 小波函数进行连续小波变换后的小波系数图。从图中可以看到白色部分恰好刻画了原始波形的周期性，且周期在 100ms-200ms 左右。

而对于连续小波变换在 EEG 的应用^[9]，在 ERP 领域一般采用以下步骤：

(1) 对单个刺激的各个电极的波形进行 CWT 变换得到时间-尺度的系数表，并按需求分成几个组（如男性刺激与女性刺激）

(2) 将各组的系数表求平均后在各个电极间进行 T 检验，得到 T 检验分布图。

这样的手段在对 P300（刺激后 300ms 的正峰值）的研究中，能清楚地将 P300 的显著性以 T 检验分布图的方式表现出来。

2.4.2 离散小波变换

同连续小波变换相同，小尺度反应高频信息，大尺度反应低频信息，则就可以将原信号分解成近似信号（低频信号）和细节信号（高频信号）。

在 EEG 脑电信号中，这个特性不仅可以用来分析信号中高频部分和低频部分，而且可以用来去除原始信号中的肌电信号（即频率较高的部分）。

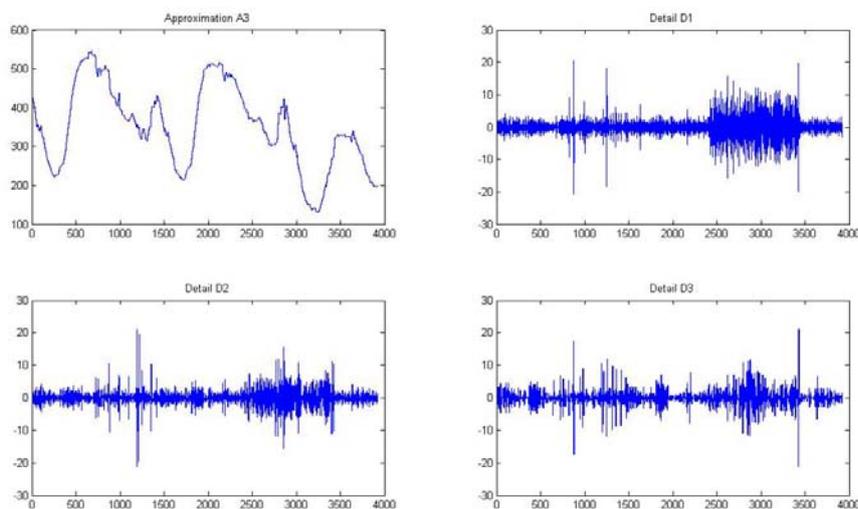


图 4 小波分析波形图

左上角为 3 层小波分解后重构的近似波形，右上角为 1 层小波系数重构的细节波形

左下角为 2 层小波系数重构细节波形，右下角为 3 层小波系数重构的细节波形

2.5 支持向量机(Support Vector Machine, SVM)

在机器学习中，对数据进行分类是重要的一项任务。已知给定一系列点，并且这些点分成两类，对于新的数据点，需要确定它属于两类中的哪一类。

在支持向量机中，一个数据点被表示成一个 p 维的向量，在 p 维空间中，需要确定一个超平面，将数据分为两类，并且两类到超平面的距离之和最大，从而考虑新的数据点落在超平面的哪一边。

从数学角度，支持向量机需要给出以下最优化问题的解^[8]。

$$\min_{w, b, \xi} \frac{1}{2} w^T w + C \sum_{i=1}^l \xi_i \quad (2-11)$$

其中，

$$y_i(w^T \phi(x_i) + b) \geq 1 - \xi_i, \quad \xi_i \geq 0 \quad (2-12)$$

而 $K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$ 称为核函数，本文将使用两类核函数：

线性核：

$$K(x_i, x_j) = x_i^T x_j \quad (2-13)$$

RBF 核：

$$K(x_i, x_j) = e^{-\gamma \|x_i - x_j\|^2} \quad (2-14)$$

2.6 本章小结

本章对本文中所需要的数学工具进行了介绍，为下文的阅读提供了基础帮助。

第三章 脑电与 ERP 分析技术

3.1 实验相关术语

3.1.1 大脑的分区

大脑表面被中央沟、顶枕裂及大脑外侧裂分成额叶、顶叶、枕叶和颞叶，如图 5 所示。其中颞叶以听觉功能为主，枕叶以视觉功能为主，顶叶是躯体感觉的高级中枢，额叶以躯体的运动功能为主^[1]。

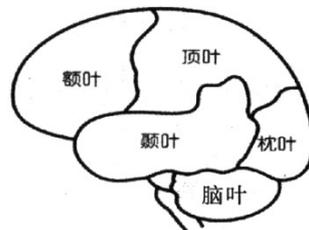


图 5 大脑基本分区图

3.1.2 脑电的基本特征

人的一举一动都离不开脑，无论是思维、行为，还是情感，都是脑 – 这个神经功能最高中枢控制的结果。

1924 年，德国精神科教授 Ham Berger 为了解释神经机能的生物学基础，开始研究人脑的电生理活动。他使用两根白金针状电极，从颅骨因外伤而缺失的部位，插入大脑皮层，并且成功记录得到有规则的电活动。然而，这种电生理活动并不需要将电极插入大脑皮层，而只需通过安置在头皮的电极即可。

其中，正常人在安静、闭眼时，主要在枕、顶部会出现 10Hz、振幅 50 μ V 左右的有规则的波叫做 α 波；当人睁眼时， α 波将消失并出现 18-20Hz、20-30 μ V 的波叫做 β 波等一系列波形。

这样的脑电活动总称为脑电图(electroencephalography, 简称 EEG)，因为脑电是通过具体实验采集的数据，所以 EEG 客观地记录了时时刻刻变化的脑机能状态。

但由于 EEG 所采集的数据为电极附近电场的均值，其中的电场包括真实 EEG 数据、肌肉电生理活动等一系列干扰，如前额附近的电极会受到眼睛眨动与转动的干扰，而所有电极都会受到身体肌肉运动的干扰，而前者一般称为眼电(EOG)，而后者称为肌电(EMG)。

3.1.3 电极的分布

EEG 实验一般采用 64 电极记录 EEG 数据，2 个电极记录 EOG 数据，电极的排布采用扩展的 10-20 系统，如图 6 和图 7 所示。

因此在脑电研究中，ICA 常常被用来去除眼电，效果非常显著。

3.3 本章小结

本章介绍了脑电和 ERP 实验中相关的术语，内容包括：大脑的分区、脑电的特征、电极的分布、ERP 实验、独立成分分析（ICA）。

第四章 实验设计

4.1 被试

3 个右手利的男性被试，2 名在校博士生及 1 位在校本科生，有正常的视力或矫正后正常的视力。在实验前，均了解 EGG 实验的记录技术，并知道实验是没有伤害性的。

对于单幅人脸，被试的平均按键速度为 504.1142ms；而对于三幅人脸，被试的的平均的按键速度为 870.3068ms。

4.2 刺激方式

所有刺激均可以明显区别人物的性别。刺激图片均位于屏幕中央，显示器分辨率为 1280×1024 像素，单张图片刺激的视觉角度为 6.9° 。

4.2.1 实验一至实验四

刺激分为真实人脸和卡通人脸两类，均为 300×300 灰度图片。

真实人脸总共 224 张（其中男性 122 张，女性 102 张），来自 AR 库 93 张（其中男性 45 张，女性 48 张）用于实验一，如图 2 所示；来自 FERET 库 131 张（其中男性 77 张，女性 54 张）用于实验三，如图 8 所示；

卡通人脸通过 Microsoft Cartoon Maker 软件在真实人脸中挑选 201 张照片转换卡通图片，用于实验二和实验四，如图 9 所示。

每张刺激持续 300ms，刺激与刺激之间间隔为 1100ms-1400ms 随机。



图 8 实验一与实验三真实人脸刺激



图 9 实验二与实验四卡通人脸刺激

4.2.2 实验五至实验八

刺激分为真实人脸和卡通人脸两类，均为 600×300 灰度图片。

真实人脸总共 312 组，每组图片均从 AR 库的 93 张图片中，随机抽取 3 张图片，3 张图片性别均相同分为一类（115 组），有一张不同性别的分为一类（197 组），如图 10 所示。

卡通人脸总共 314 组，每张图片均从单副卡通人脸中，随机抽取 3 张图片，3 张图片性别均相同分为一类（116 组），有一张不同性别的分为一类（197 组），如图 11 所示。

每张刺激持续 800ms，刺激与刺激之间间隔为 1400ms-2000ms 随机。



图 10 实验五与实验七 三张真实人脸刺激



图 11 实验六与实验八 三张卡通人脸刺激

4.3 实验过程

实验是一个连续的过程，刺激总长度为半小时。过程中包括 8 个实验，每个实验开始有 15 秒的实验指导，告知被试实验内容与实验中被试需要完成的任务；接下来每个实验包含 100 个刺激，50 个刺激后有 15 秒的休息。每个实验靶刺激与非靶刺激各占 50%，即各 50 张图片。

4.3.1 训练过程

实验前有二个训练过程，不记录 EEG 数据，为了让被试适应实验中刺激出现的速度，第一个过程使被试适应单张人脸照片的 300ms 的刺激时间，并做出按键判断；第二个过程使被试适应三张人脸照片的 800ms 的刺激时间，并做出按键判断。

4.3.2 实验一至实验四

第一个实验与第三个实验均为单张真实人脸照片的刺激。其中，实验一的刺激来源于 AR 库，实验三的刺激来源于 FERET 库。在实验一与实验三中，被试的任务是在男性刺激中分辨出女性刺激。与实验一相比，实验三需要被试通过按键来确定判断的结果，即当是女性刺激时，被试需要按键；当是男性刺激时，被试不需要按键。

第二个实验与第四个实验均为单张卡通人脸照片的刺激。在实验二与实验四中，被试的任务是在男性刺激中分辨出女性刺激。与实验二相比，实验四需要被试通过按键来确定判断的结果，即当是女性刺激时，被试需要按键；当是男性刺激时，被试不需要按键。

4.3.3 实验五至实验八

第五个实验与第七个实验的每个刺激均为一组照片，每组照片为三张并列真实人脸照片的刺激。刺激来源均来自于 AR 库。在实验五与实验七中，被试的任务是分辨每组照片中是否有不同性别的刺激。三张图片共有四种情况，三个男性、两个男性一个女性、一个男性两个女性、三个女性。其中三个男性和三个女性为性别相同的刺激，另外两个为性别不同的刺激。与实验五相比，实验七需要被试通过按键来确定判断的结果，即当是不同性别刺激时，被试需要按键；当是相同性别刺激时，被试不需要按键。

第六个实验与第八个实验的每个刺激均为一组照片，每组照片为三张并列卡通人脸照片的刺激。在实验六与实验八中，被试的任务是分辨每组照片中是否有不同性别的刺激。与实验六相比，实验八需要被试通过按键来确定判断的结果，即当是不同性别刺激时，被试需要按键；当是相同性别刺激时，被试不需要按键。

4.4 实验设计的依据说明

4.4.1 灰度图片的使用

在做 ERP 实验时，为了排除任何信息对被试刺激的干扰，在挑选和处理刺激时，要考虑照片的背景尽可能为白色背景或单色背景；而为了排除颜色信息对于被试的干扰，因此选用灰度图片作为刺激。

4.4.2 卡通图片的使用

卡通图片与真实图片相比，所包含的信息量更少。对于真实图片，作为细致的观察不难发现，许多的信息都会被处理到，如皮肤的光泽、环境的明暗、人物的表情等信息，虽然不能确定这些信息在视觉处理、尤其是人脸性别处理时是否需要这些信息，或什么时候需要这些信息，但为了排除这些因素，使用控制变量法，使得生成的卡通图片明暗相同，人物表情基本相同。

4.4.3 间隔的随机

在进行 ERP 设计时，刺激间隔和刺激序列要尽可能的做到随机或伪随机。根据研究表明，P300（即刺激产生后 300ms 左右正向的峰值）在规则序列靶刺激诱发成分比随机序列的靶的幅值显著减少^[1]。而本文所关注的时间区间在 45ms-85ms，200ms 以后的 ERP 成分，所以有必要尽可能的减少 P300 对于波形的影响。

4.4.4 15 秒的休息

在实验后，通过询问被试对于实验的情况，普遍反映整个实验过程中，视觉相当疲劳；而且由于实验任务相对枯燥，需要被试长时间的集中注意力，无论视觉疲劳，还是肌肉疲劳，都会对 EEG 记录时产生不良的影响，因此在实验中，设计 50 个刺激后有 15 秒的休息时间。

4.5 实验设计目的

实验设计总共分为八个实验，其中实验一至实验四为一组同类实验，实验五至实验八为一组同类实验。

对于实验一至实验四，刺激为单张人脸图片，分为真实人脸（实验一与实验三）和卡通

人脸（实验二与实验四）；实验一与实验三、实验二与实验四中，前者不需要被试按键，而后者需要被试按键。

其中实验一与实验二、实验三与实验四，期望能够发现处理人脸性别信息时，真实图片和卡通图片对于诱发的 ERP 的区别。

从实验三与实验四中，相较于实验一与实验二，需要被试保持更高的注意力，比较结果是否存在差异；同时，从按键情况可能得到被试的平均反应时间，并确定数据处理的时域跨度。

四个实验的整体目的是通过实验观察男性刺激与女性刺激所诱发的 ERP 波形是否存在稳定差异，是否可通过分类器加以区分。

同样，对于实验五至实验八，每组刺激包括三张人脸图片，分为真实人脸（实验五与实验七）和卡通人脸（实验六与实验八）；实验五与实验七、实验六与实验八中，前者不需要被试按键，而后者需要被试按键。

这四个实验的目的是：

- (1) 从复杂的输入中验证实验一至实验四得到的结论是否依然成立。
- (2) 人在判断性别时，即使是在刺激多而且时间间隔短的情况下，是否能够迅速辨认。
- (3) 人在观察多张人脸图像并做出判断时，EEG 是否作为一种简单的叠加。

4.6 EEG记录参数

EEG 的采样率为 1kHz，66 个电极（64 个 EEG 数据电极、2 个 EOG 数据电极），放大器 $\pm 200\text{mV}$ ，精度 29.80nV/LSB ，低通滤波频率为 100Hz ，高通滤波频率为 0.1Hz ，以顶骨中央作为参考点，眼电通过左眼上下电极记录水平眼电信号，两眼外侧电极记录垂直眼电信号。

EEG 记录后数据为 NeuroScan CNT 脑电波形数据，进行 DC 偏移校正后；使用 ICA 转换成 64 个独立分量，并手动去除眼电分量。

ERP 平均波形总长为 600ms ，包含刺激前 100ms 进行基线矫正。

4.7 本章小结

本章介绍了实验设计与进行情况，包括被试、实验刺激、实验过程、EEG 记录参数，还介绍了实验设计的目的与灰度图片、卡通图片、间隔随机、15 秒休息设置的理由，为下文数据分析做准备。

第五章 实验数据分析

实验数据的分析分为时域分析、功率谱分析和小波分析，并依据这三个分析抽取的特征，进行 SVM 训练。

5.1 分析前提

(1) 观察图 11 的波形可以发现，被试一的波形存在男性刺激趋势大于女性刺激的规律，但经过观察，被试二的波形恰恰相反，并且基于各人之间在思考、判断时大脑活动可能存在差异，因此假设各个被试间数据分析结果存在差异，但本质趋势保持不变。

(2) 使用 SVM 进行如下实验，

选取两次实验的数据（被试一），从中选出 100 个刺激，将前 100ms 至刺激后 499ms 的时域数据作为特征，共 600 维，前一次实验的标记记为 1，后一次实验的标记记为 2。总共 200 组数据，标记为 1 与标记为 2 的数据各 100 个，并随机排列。

使用线性核与 RBF 核进行分类，得到的结果如表 1 所示：

表 1 实验间数据分类精度表

	RBF 核	线性核
实验一与实验二	61.5%	61.5%
实验二与实验三	65.5%	57%
实验三与实验四	54%	50.5%

从表中的数据可以看到，实验间的数据随机排列后，700 维的数据进行分类的精度达到 55%-65% 左右。

实验的目的在于，用数据来说明，在对同一被试不同实验之间的刺激时域数据进行分类时，精度也能够达到 50% 以上，最高达到 65.5%，因此不进行不同实验间的数据分类。

(3) 假设对于性别分类时大脑反应的差异可能是阈值型或通道型的。

一方面，阈值型是指某些电极的波形幅值达到某一阈值时判断为一性别、未到达时则为另一性别。此时电极的位置则表明电极所在脑区用于处理性别。

另一方面，通道型指接收到人脸刺激后，对于男性与女性刺激，脑区的分别截然不同，刺激的视觉信息完全进入两条通道，彼此不相干。

本文主要依据阈值型进行验证。

5.2 时域分析

时域分析即 ERP 分析，根据已有的研究结果^[2]，通过进行有性别任务与无性别任务（实验分为两类，一类刺激均为同性别，称为无性别任务；另一类刺激将男女混合，称为有性别任务。），得到结论：被试接收刺激后 45-85ms、145-185ms、200-250ms 内，有性别任务与无性别任务的 ERP 波形有明显的差异。

但是并未提出男性与女性之间的波形是否有差异，因此，时域的分析主要依据 3 个被试

的实验数据，并且在这三个时间区域内进行统计检验，而检验的主要统计量为两个波形之间是否在幅值上有大小的差异。

如图 12 所示，是对单幅人脸实验数据进行 ERP 平均后的波形，初步观察可以发现位于 P5 电极 45-85ms 时间区域内男性波形明显高于女性波形，且差异明显，而这个现象在同一实验中出现在脑后大部分电极中。

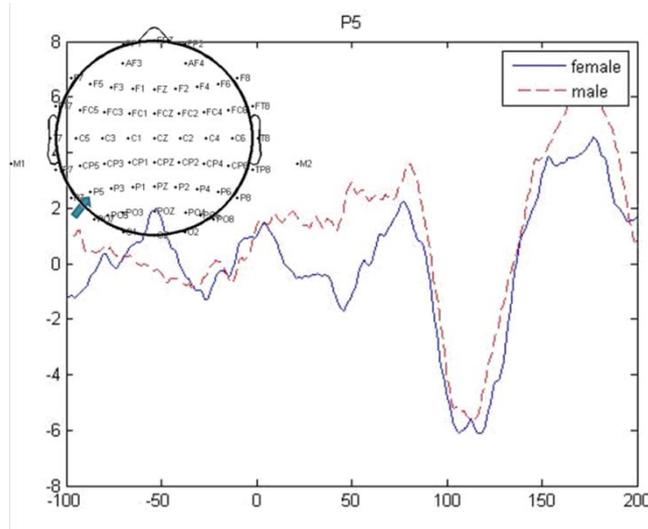


图 12 被试一实验一 单幅人脸 P5 电极男女刺激对比波形

5.2.1 检验男女刺激波形幅值的大小关系

这部分检验主要考察在各电极之间，是否存在绝对的男性波形高于女性波形，或者女性波形高于男性波形的情况。

(1) 检验区间内 ERP 整体波形

在考虑各次刺激间的关系前，先考察叠加平均后的 ERP 波形是否存在规律，具体流程如图 13 所示。

(a) 将实验过程记录的数据分成男性数据与女性数据两部分，其中各部分共 50 个刺激，即 50 组数据，每组数据开始时间为刺激前 100ms，结束时间为刺激后 499ms，即[-100, 499]，共 600 个数据点。

(b) 将男性 50 组数据叠加后取平均，得到男性的 ERP 波形；同样的方法得到女性的 ERP 波形，总共各 64 个电极的波形。

(c) 在各电极上将女性的波形减去男性的波形，并在相应时间区间内进行离散积分，并定义离散积分的值作为此区间内波形的**整体参数**。

定义：在时间区间 $[\alpha, \beta]$ 上的 f_n 的离散积分为

$$V = \sum_{i=\alpha}^{\beta-1} \frac{f_i + f_{i+1}}{2}$$

并称为此时间区间上 EEG 的整体参数。

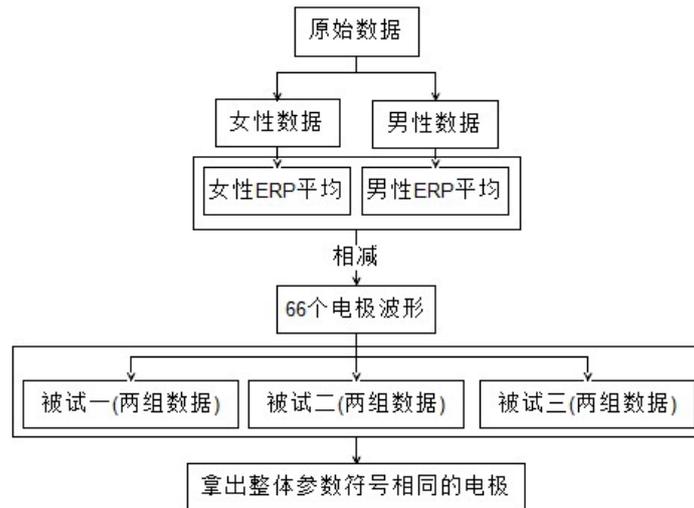


图 13 检验区间内 ERP 整体波形流程图

对实验一与实验三的数据在 45-85ms、145-185ms、200-250ms 时间区间内分别进行以上步骤的处理，对实验七与实验八的数据在 200-250ms 时间区间内分别进行以上步骤的处理，结果如表 2 所示，图 14 为结果中电极的分布图。

表 2 检验区间内 ERP 整体波形 结果

	所有通过电极	脑后电极
实验一与实验三(45-85ms)	6 19 45	45(P5)
实验一与实验三(145-185ms)	1 3 8 9 11 17 18 20 22 29 39 52 64	39(CP2) 52(P8) 64(O1)
实验一与实验三(200-250ms)	6 10 13 22 29 34 38 40 46 49 50	40(CP4) 49(P2) 50(P4)

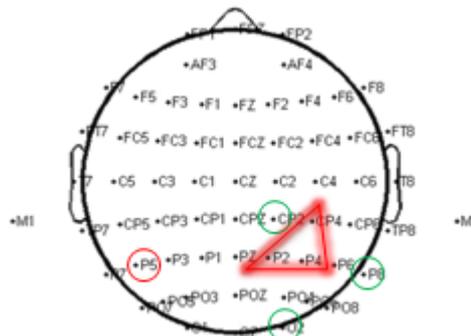


图 14 实验结果电极分布图

左侧红圈中的电极 P5 为 45-85ms 的检验结果

右侧绿圈中的电极 CP2、P8、O1 为 145-185ms 的检验结果

右侧红色三角形中的电极 CP4、P2、P4 为 200-250ms 的检验结果

每个被试各有两个真实人脸的实验，每个实验过有 66 个数据，对每个被试的数据，将两次实验符号不相同的电极去除，并求三个被试符合要求的电极的交集。

- [45, 85]时间区间内
得到 3 个电极，编号分别为 6、19、45。
其中 6 与 19 位于前额，受眼电干扰大，并且根据视觉理论，视觉信息的传递由脑后传递向前额，因此 2 个电极关联程序不大。
而 45 的电极名称为 P5，与之前结果^[2]所存在差异的电极 P3 的位置大致相同（两次实验同名电极位置不同）。
- [145, 185]时间区间内

得到 14 个电极，编号分别为 1、3、8、9、11、17、18、20、22、29、39、52、64。
考虑后三个电极：39 (CP2)、52 (P8)、64 (O1)。

- [200, 250]时间区间内
得到 11 个电极，编号分别为 6、10、13、22、29、34、38、40、46、49、50。
考虑脑后三个电极：CP4、P2、P4。三个电极相当的接近，说明 200-250ms 对于性别的处理可能位于右半球。

(2) 检验区间内单个时间点 ERP 波形

当区间内波形的整体参数没有显著性特征时，尝试在各电极间，在特定的时间点上寻找是否有明显的差异。具体的流程如图 15 所示。

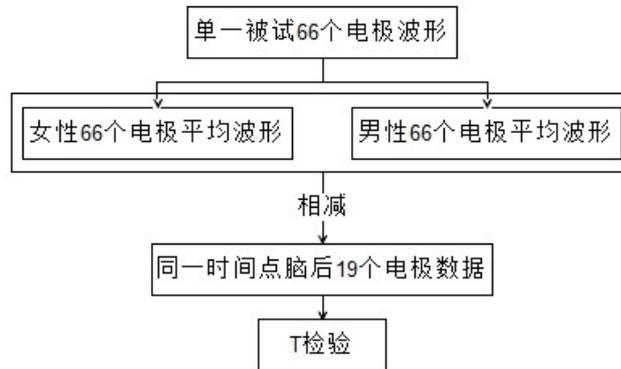


图 15 检验区间内单个时间点 ERP 波形

(a) 将实验过程记录的数据分成男性数据与女性数据两部分，其中各部分共 50 个刺激，即 50 组数据，每组数据开始时间为刺激前 100ms，结束时间为刺激后 499ms，即[-100, 499]，共 600 个数据点。

(b) 将男性 50 组数据叠加后取平均，得到男性的 ERP 波形；同样的方法得到女性的 ERP 波形，总共各 66 个电极的波形，抽取脑后 19 个电极（即 P7、P5、P3、PZ、P2、P4、P6、P8、PO7、PO5、PO3、POZ、PO4、PO6、PO8、O1、OZ、O2）的数据。

(c) 在某个时间点，如 45ms 处，以 19 个电极的相减波形的值作 T 检验 ($\alpha = 0.05$)。检验结果如表 3 所示。

表 3 检验区间内单个时间点 ERP 波形 结果 (显著性水平 $\alpha = 0.05$)

被试	实验	通过区间
被试一	实验一	76-89ms ($p < 0.04$)
	实验二	56-86ms ($p < 0.04$), 151-172ms ($p < 0.03$), 235-246ms ($p < 0.03$)
	实验三	248-350ms ($p < 0.04$) 263ms 后 ($p < 0.001$)
	实验四	224-274ms ($p < 0.04$) 281-350ms ($p < 0.04$)
被试二	实验一	无通过区间
	实验二	无通过区间
	实验三	无通过区间
	实验四	220ms 左右，但相当零散

比较被试一与被试二的情况，实验结果截然相反。从被试一的结果观察，可以在之前结

果^[2]的基础上得出, 在 45-85ms、145-185ms、200-250ms 不仅仅有性别与无性别任务有差异, 而且男性刺激与女性刺激也有差异。反观被试二的结果, 从中得出的结论是完全没有规律。这两个结果都可能是巧合, 有待调整实验再做验证。

(3) 检验各个刺激间的波形

这部分检验期望在各个刺激间, 能够在时域波形中观察到看到男性与看到女性的波形是否有明显的差异, 对于数据的处理的具体流程如图 16 所示。

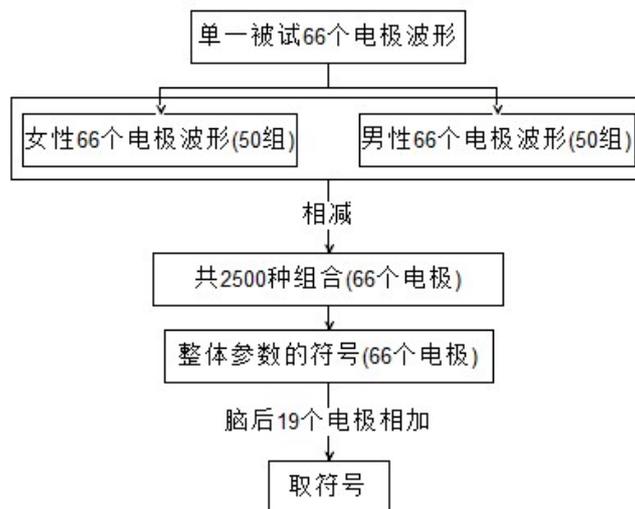


图 16 检验各个刺激间的波形 流程图

(a) 将实验过程记录的数据分成男性数据与女性数据两部分, 其中各部分共 50 个刺激, 即 50 组数据, 每组数据开始时间为刺激前 100ms, 结束时间为刺激后 499ms, 即 $[-100, 499]$, 共 600 个数据点。

(b) 各从男性刺激与女性刺激中取一组数据, 总共 2500 种排列, 共 66 个电极数据。

(c) 将一组男性刺激与一组女性刺激的波形相减, 并在相应时间区间内进行离散积分, 即整体参数, 取脑后 19 个电极积分值的符号相加后的符号为结果, 统计+1 与-1 出现的频率。

从而得到统计的结果如表 4 所示。

表 4 检验各个刺激间的波形结果总计

符号	出现次数	总体频率
1 (女性大于男性)	9833	49.2%
-1 (男性大于女性)	9459	47.3%
0	708	3.5%

从结果来看, 三个被试脑后各个电极积分和的符号从总体上看没有显著的差异, 各个实验的数据如表 5 所示。从表中的数据可以观察到, 有部分实验存在着显著的差异, 如被试一的实验二, 女性积分值较男性积分值高的比例占 60%, 但无法说明存在普遍性的规律。

表 5 检验各个刺激间的波形各个被试详细结果

被试	实验	女性大于男性的频率	男性大于女性的频率
被试一	实验一	53.3	44.5
	实验二	59.4	38.0
	实验三	48.3	48.9
	实验四	49.2	48.1
被试二	实验一	48.3	46.0

被试	实验	女性大于男性的频率	男性大于女性的频率
	实验二	49.9	46.4
	实验三	43.5	51.9
	实验四	50.2	45.8

5.2.2 检验部分电极的显著性差异

从以上的检验结果观察，综合脑后 19 个电极的数据，对于幅值并不存在显著的差异，但可能存在其中几个电极对于数据的幅值存在一定的规律，因此这部分的检验希望能够找到 19 个电极的子集，其中的电极对于幅值存在显著性的差异，具体数据处理步骤如图 17 所示。

(a) 将实验过程记录的数据分成男性数据与女性数据两部分，其中各部分共 50 个刺激，即 50 组数据，每组数据开始时间为刺激前 100ms，结束时间为刺激后 499ms，即[-100, 499]，共 600 个数据点。

(b) 将男性 50 组数据叠加后取平均，得到男性的 ERP 波形；同样的方法得到女性的 ERP 波形，总共各 66 个电极的波形。

(c) 将 66 个电极的女性刺激与男性刺激波形求差值，并在相应区间内进行离散积分。

(d) 枚举脑后 19 个电极的子集，对子集积分值进行 T 检验。

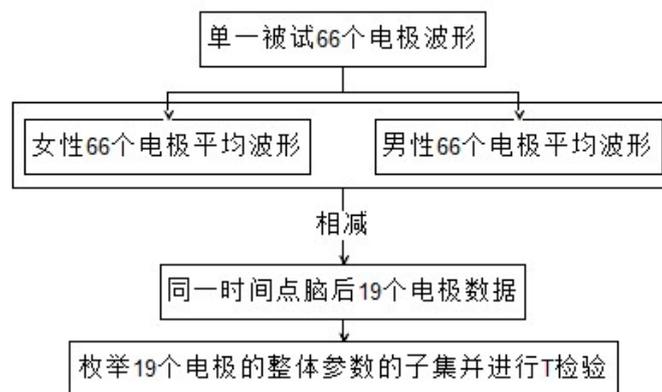


图 17 检验部分电极的显著性差异 流程图

从而得到统计结果如表 6 所示

表 6 检验部分电极的显著性差异 结果

被试	实验	通过 T 检验子集的个数
被试一	实验一	1010
	实验三	22

但各个实验间的子集并没有关系，但同样，被试一的实验二也超过其它实验子集个数许多，观察这个试验的数据发现，19 个电极的积分值中有 17 个电极的积分值是大于 0 的，且与 0 的差距大于 20。因此大部分子集能够通过 T 检验，并且观察数据可以发现，被试一实验二的数据在 45-85ms 区间内女性数据较男性数据的差异非常显著。

5.3 功率谱分析

对于频域数据的分析，主要依靠将相应区间内的时域数据转换为功率谱的手段，从而期

望验证时域分析中没有结果的部分。

5.3.1 检验各个刺激间功率谱的差

在时域分析中，将男性与女性刺激的波形相减后相应时间区间内的 2500 个整体参数的符号正负分布接近 50%:50%，从而得出没有显著性差异的结果，而对功率谱的检验针对相应时间区间的功率谱的整体参数，期望从功率谱中能够观察到显著性差异，具体流程如图 18 所示。

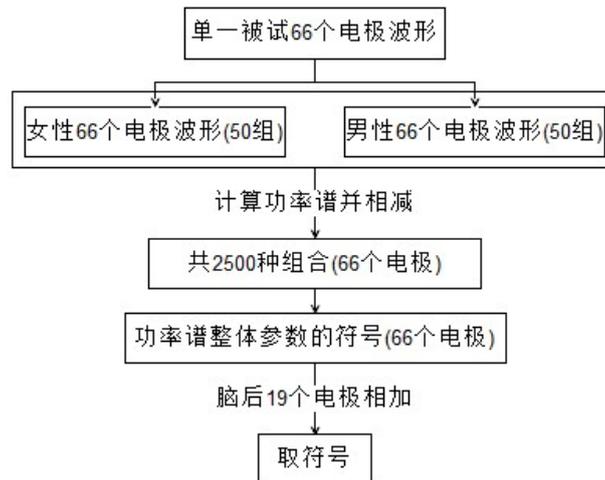


图 18 检验各个刺激间功率谱的差 流程图

从而得到统计结果如表 7 所示。

表 7 检验各个刺激间功率谱的差 结果总计

符号	出现次数	总体频率
1 (女性大于男性)	9761	48.8%
-1 (男性大于女性)	9016	45.1%
0	1223	6.1%

观察两个被试共 20000 个数据不难发现，与时域信息相同，在功率谱也不存在普遍的幅值大小规律，但更进一步观察表 8 的结果不难发现，大部分实验在功率谱中都有一定趋向性，即男性与女性的能量有存在大小关系，但在各实验、各个被试之间无普遍规律。

表 8 检验各个刺激间功率谱的差的各个被试详细结果

被试	实验	女性比男性能量高	男性比女性能量高
被试一	实验一	59.0%	36.2%
	实验二	47.2%	48.0%
	实验三	43.5%	52.7%
	实验四	43.2%	52.6%
被试二	实验一	31.1%	59.1%
	实验二	28.2%	65.2%
	实验三	57.0%	34.9%
	实验四	51.4%	41.8%

5.4 小波分析

选择小波变换的原因，一方面考虑前文时域与功率谱的分析的结果并没有进展，另一方面，考虑海森堡不确定性原理，在 40 个点的采样条件下，频域的信息并不能准确的体现出相应时间区间内的能量。

基于小波函数的选择与原始波形要尽可能的相似的依据，选择 Daubechies4 波形^[10]（即 db4），如图 19 所示。

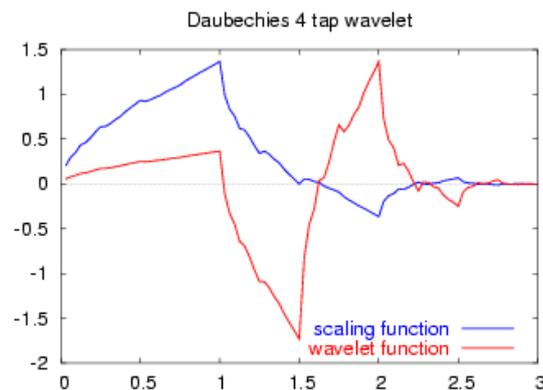


图 19 Daubechies4 小波波形

5.4.1 小波系数分析(连续小波分析)

小波以不同的尺度，对应不同的时间起点，对照时域波形求得的小波系数，表达的意义在于小波系数越大，波形与小波的相似程度越高。

根据一般作法^[9]进行统计，具体流程如图 20 所示。

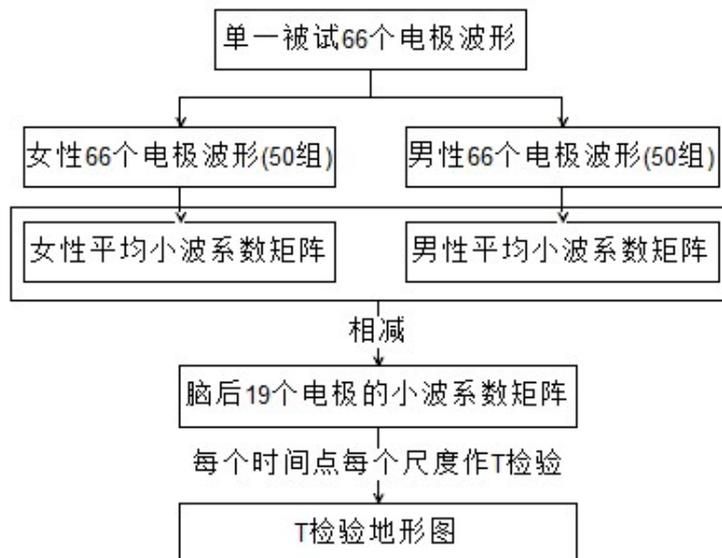


图 20 连续小波分析 流程图

统计结果如图 21 所示：

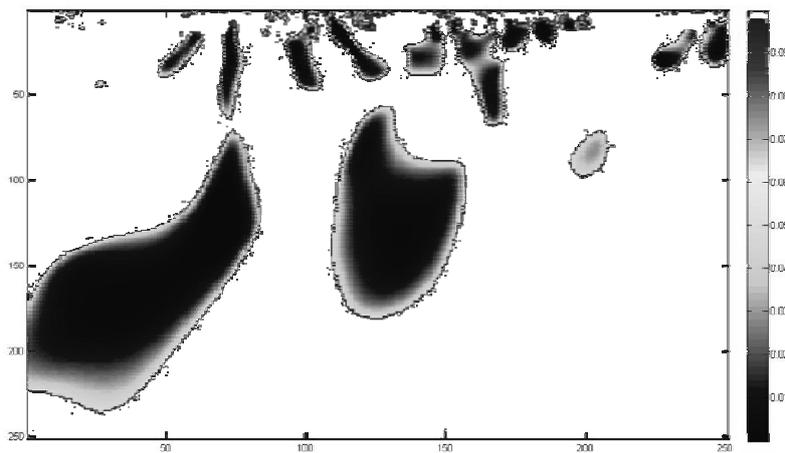


图 21 被试一实验一的 T 检验地形图

横轴为时间轴 0 - 250ms，纵轴为尺度由上至下 1-250

图中的空白部分为未通过 T 检验部分。而在其余的结果中：

三个区间 45-85ms、145-185ms、200-250ms 中，被试一的实验一至实验四没有发现显著相同的模式，而对于顶部的色块很好的验证了 5.2.1 中(2)的正确性，即被试一的实验一至实验四中，在 45-85ms、145-185ms、200-250ms 时间区间内，脑后 19 个电极在部分时间点中男女刺激差值通过 T 检验。

而对于四个实验，实验三在 45-85ms 与 145-185ms 内，几乎没有时间 - 尺度的点通过 T 检验，相比较 5.2.1 中(2)的结果，更进一步的是在任意的时间区间内，实验三的数据均不通过 T 检验。而在 200-250ms 区间内，4 个实验都有较多的时间 - 尺度点通过 T 检验，虽然没有相同的模式，但在一些公共区域内有接近的检验结果。

5.5 利用特征进行分类

之前分析的特征一部分从统计学上达到了显著性水平，另一部分根据统计结果进行猜测可能的关系，这一部分将各部分特征组织成为训练数据，使用 LibSVM 分别使用线性核和 RBF 核进行分类。

根据 5.1 的分析前提，各个实验间在训练时其余实验的数据，这就造成训练与测试样本数只有 100 个，但是特征可能上百甚至上千，因此使用线性核的结果可能更好一些^[7]。

每组实验 100 个数据，而刺激的顺序采用随机，因此在时间轴上刺激类型并不是均匀分布的，即按时间先后顺序，前 50 个与后 50 个出现的男性刺激与女性刺激的个数不一致。这样排布的原因是，较有顺序的刺激顺序会诱发 P300 的效果^[1]。

为了考察实时分类与训练数据的均衡，使用两种方式训练并预测。

(1) 按刺激时间顺序考虑，前 50 个刺激的波形作为训练集，后 50 刺激的波形作为测试集。

(2) 按刺激样本均衡考虑，即男女性刺激个数相同的角度，前 25 个男性刺激与前 25 个女性刺激为训练集，后 25 个男性刺激与后 25 个女性刺激为测试集。

对于线性核，使用 5 折交叉验证得到训练精度，用于考察特征的准确性，参数 C 的取值范围为集合 $\{2^{-5}, 2^{-3}, 2^{-1}, 2^1, 2^3, 2^5, 2^7, 2^9, 2^{11}, 2^{13}, 2^{15}\}$ 。

对于 RBF 核，使用 5 叠交叉验证得到训练精度，用于考察特征的准确性，参数 C 的取值范围为集合 $\{2^{-5}, 2^{-3}, 2^{-1}, 2^1, 2^3, 2^5, 2^7, 2^9, 2^{11}, 2^{13}, 2^{15}\}$ ，参数 G 的取值范围为集合 $\{2^{-15}, 2^{-13}, 2^{-11}, 2^{-9}, 2^{-7}, 2^{-5}, 2^{-3}, 2^{-1}, 2^1, 2^3\}$ 。

对于每次训练，将数据范围缩放至 $[-1, 1]$ 的范围。

5.5.1 利用时域信息进行分类

(1) 直接对时域信息进行分类

根据 5.2 的统计结果，对于性别任务较为敏感的时间区间为 45-85ms、145-185ms、200-250ms。

(a) 按时间顺序划分训练集和测试集

分别以脑后 19 个电极、三个时间区间的幅值信息，按时间顺序，前 50 个刺激作为训练集，后 50 个刺激作为测试集进行分类，结果分别如表 9、表 10、表 11 所示。

表 9 45-85ms 100 组数据 $19 \times 41 = 779$ 维特征 分类精度 (按时间顺序均衡)

被试	实验	线性核	RBF
被试一	实验一	30% (C=2⁻⁵)	30%(C=2⁻⁵, G=2⁻¹⁵)
	实验二	44% (C=2⁻³)	44%(C=2¹, G=2⁻⁵)
	实验三	52% (C=2 ³)	52%(C=2 ⁷ , G=2 ⁻¹¹)
	实验四	46% (C=2⁻⁵)	46%(C=2¹, G=2¹)
被试二	实验一	30% (C=2⁻⁵)	30%(C=2⁻⁵, G=2⁻¹⁵)
	实验二	44% (C=2⁻¹)	54%(C=2 ⁷ , G=2 ⁻¹¹)
	实验三	44% (C=2⁻¹)	52%(C=2 ⁷ , G=2 ⁻¹¹)
	实验四	46% (C=2⁻¹)	44%(C=2 ³ , G=2 ⁻⁵)

表 10 145-185ms 100 组数据 $19 \times 41 = 779$ 维特征 分类精度 (按时间顺序均衡)

被试	实验	线性核	RBF
被试一	实验一	30% (C=2⁻⁵)	46%(C=2 ³ , G=2 ⁻¹)
	实验二	44% (C=2⁻⁵)	52%(C=2 ¹ , G=2 ⁻¹)
	实验三	44% (C=2⁻⁵)	44%(C=2 ⁻⁵ , G=2 ⁻¹⁵)
	实验四	56% (C=2⁻¹)	46%(C=2 ¹ , G=2 ⁻¹)
被试二	实验一	30% (C=2⁻⁵)	30%(C=2⁻⁵, G=2⁻¹⁵)
	实验二	44% (C=2¹)	60%(C=2 ⁷ , G=2 ⁻¹¹)
	实验三	44% (C=2⁻⁵)	46%(C=2 ⁷ , G=2 ⁻³)
	实验四	48% (C=2 ⁷)	52%(C=2 ¹⁵ , G=2 ⁻⁹)

表 11 200-250ms 100 组数据 $19 \times 41 = 969$ 维特征 分类精度 (按时间顺序均衡)

被试	实验	线性核	RBF
被试一	实验一	36% (C=2 ³)	38%(C=2 ⁵ , G=2 ⁻³)
	实验二	44% (C=2¹)	46%(C=2 ¹ , G=2 ¹)
	实验三	44% (C=2⁻³)	50%(C=2 ⁷⁵ , G=2 ⁻⁹)
	实验四	46% (C=2⁻⁵)	46%(C=2 ⁻⁵ , G=2 ⁻¹⁵)
被试二	实验一	30% (C=2⁻⁵)	30%(C=2 ⁻⁵ , G=2 ⁻¹⁵)
	实验二	44% (C=2⁻³)	62%(C=2 ¹ , G=2 ¹)
	实验三	44% (C=2⁻¹)	36%(C=2 ³ , G=2 ⁻⁵)
	实验四	46% (C=2⁻⁵)	46%(C=2 ⁻⁵ , G=2 ⁻¹⁵)

表中黑体部分表示，对测试集中刺激类型的判断均为一种性别，即对于 50 个测试样本，训练得到的分类器的选择均为同一性别。考虑这样的分类器并没有分类的功能，而按时间顺序划分训练得到的分类器普遍存在这样的问题。分析原因是由样本分布不均造成的。

(b) 按刺激样本均衡划分训练集和测试集

分别以脑后 19 个电极、三个时间区间的幅值信息，作为训练数据，结果分别如表 12、表 13、表 14 所示。

表 12 45-85ms 100 组数据 $19 \times 41 = 779$ 维特征 分类精度 (按刺激样本均衡)

被试	实验	线性核	RBF
被试一	实验一	42% ($C=2^1$)	40% ($C=2^{-5}, G=2^{-9}$)
	实验二	58% ($C=2^{-5}$)	68% ($C=2^1, G=2^{-5}$)
	实验三	56% ($C=2^1$)	50% ($C=2^{-5}, G=2^1$)
	实验四	42% ($C=2^1$)	36% ($C=2^3, G=2^{-7}$)
被试二	实验一	48% ($C=2^{-5}$)	44% ($C=2^{11}, G=2^{-5}$)
	实验二	48% ($C=2^{-1}$)	58% ($C=2^1, G=2^{-5}$)
	实验三	50% ($C=2^{-3}$)	54% ($C=2^{-5}, G=2^1$)
	实验四	40% ($C=2^{-3}$)	42% ($C=2^{-5}, G=2^{-15}$)

表 13 145-185ms 100 组数据 $19 \times 41 = 779$ 维特征 分类精度 (按刺激样本均衡)

被试	实验	线性核	RBF
被试一	实验一	50% ($C=2^1$)	56% ($C=2^3, G=2^{-7}$)
	实验二	38% ($C=2^{-3}$)	46% ($C=2^{-5}, G=2^{-5}$)
	实验三	50% ($C=2^{-5}$)	58% ($C=2^{-5}, G=2^{-9}$)
	实验四	52% ($C=2^{-5}$)	54% ($C=2^{-5}, G=2^1$)
被试二	实验一	56% ($C=2^{-1}$)	64% ($C=2^3, G=2^{-7}$)
	实验二	56% ($C=2^{-5}$)	52% ($C=2^{-5}, G=2^{-5}$)
	实验三	52% ($C=2^{-1}$)	48% ($C=2^{-5}, G=2^{-7}$)
	实验四	54% ($C=2^{-5}$)	60% ($C=2^{-5}, G=2^{-5}$)

表 14 200-250ms 100 组数据 $19 \times 41 = 969$ 维特征 分类精度 (按刺激样本均衡)

被试	实验	线性核	RBF
被试一	实验一	44% ($C=2^{-1}$)	62% ($C=2^9, G=2^{-13}$)
	实验二	54% ($C=2^{-5}$)	64% ($C=2^{-5}, G=2^3$)
	实验三	58% ($C=2^{-3}$)	58% ($C=2^7, G=2^{11}$)
	实验四	46% ($C=2^{-5}$)	46% ($C=2^{-5}, G=2^{-13}$)
被试二	实验一	58% ($C=2^{-1}$)	58% ($C=2^1, G=2^{-1}$)
	实验二	46% ($C=2^{-1}$)	66% ($C=2^{-5}, G=2^{-5}$)
	实验三	52% ($C=2^{-3}$)	54% ($C=2^{15}, G=2^{-15}$)
	实验四	58% ($C=2^{-5}$)	60% ($C=2^1, G=2^{-9}$)

对于线性核，实验之间差异较大，三个时间区间 45-85ms、145-185ms、200-250ms 中，平均分类精度分别为 48%、51%、52%，而对于前两个区间的精度对于任意分类器只选择一种就能够达到这个精度，因此，分类结果没有太大意义，而最后一个区间分类精度较前两个更高，最高精度达到了 58%。

对于 RBF 核，实验之间相对稳定，三个时间区间中，平均分类精度为 49%、54.75%、58.5%，精度比起线性核的结果有了很大的提高，最高精度也达到了 68%。

(2) 对时域信息进行处理后进行分类

分析原因，可以发现因为样本数据相对较少，而且各个刺激之间的波形没有任何限制，因此考虑将各个刺激的各个电极中的波形的整体参数作为特征进行分类，因为特征少于样本

数量，只进行了 RBF 核的测试，结果如表 15 所示。

表 15 100 组数据 19 维特征 RBF 核分类精度

被试	实验	45-85ms	145-185ms	200-250ms
被试一	实验一	42% (C=2 ⁻⁵ , G=2 ⁻⁵)	44% (C=2 ⁻⁵ , G=2 ⁻¹)	54% (C=2 ¹³ , G=2 ⁻¹³)
	实验二	64% (C=2 ⁹ , G=2 ⁻¹³)	48% (C=2 ⁷ , G=2 ⁻³)	48% (C=2 ⁵ , G=2 ⁻³)
	实验三	50% (C=2 ⁻⁵ , G=2 ³)	54% (C=2 ⁻⁵ , G=2 ⁻¹⁵)	68% (C=2 ⁷ , G=2 ⁻⁹)
	实验四	40% (C=2 ⁻⁵ , G=2 ⁻¹⁵)	52% (C=2 ¹ , G=2 ⁻⁵)	46% (C=2 ⁻⁵ , G=2 ⁻⁵)
被试二	实验一	50% (C=2 ⁷ , G=2 ⁻³)	58% (C=2 ⁵ , G=2 ⁻³)	66% (C=2 ⁹ , G=2 ⁻¹³)
	实验二	50% (C=2 ⁻⁵ , G=2 ⁻¹⁵)	52% (C=2 ¹ , G=2 ⁻³)	56% (C=2 ⁻⁵ , G=2 ¹)
	实验三	54% (C=2 ⁻⁵ , G=2 ⁻¹)	52% (C=2 ⁻⁵ , G=2 ¹)	68% (C=2 ³ , G=2 ⁻¹)
	实验四	44% (C=2 ⁵ , G=2 ⁻¹⁵)	60% (C=2 ³ , G=2 ⁻¹³)	52% (C=2 ¹ , G=2 ⁻³)

5.5.2 利用功率谱信息进行分类

根据 5.3 的分析结果，对于功率谱信息进行分类。

分别以脑后 19 个电极、45-85ms、145-185ms、200-250ms 三个时间区的功率谱数据做为训练数据，结果分别如表 16 所示。

表 16 100 组数据 19 × 20 = 380 维特征 RBF 核分类精度

被试	实验	45-85ms	145-185ms	200-250ms
被试一	实验一	58% (C=2 ⁻⁵ , G=2 ⁻³)	48% (C=2 ⁹ , G=2 ⁻¹³)	62% (C=2 ⁻⁵ , G=2 ⁻⁵)
	实验二	46% (C=2 ³ , G=2 ⁻³)	50% (C=2 ¹ , G=2 ⁻¹)	50% (C=2 ⁻⁵ , G=2 ¹)
	实验三	46% (C=2 ⁻⁵ , G=2 ¹)	56% (C=2 ⁻⁵ , G=2 ⁻⁵)	60% (C=2 ¹ , G=2 ⁻¹¹)
	实验四	46% (C=2 ⁻⁵ , G=2 ⁻⁵)	46% (C=2 ¹ , G=2 ⁻¹)	40% (C=2 ¹¹ , G=2 ⁻¹⁵)
被试二	实验一	52% (C=2 ¹ , G=2 ¹)	48% (C=2 ⁻⁵ , G=2 ⁻¹³)	48% (C=2 ⁻⁵ , G=2 ⁻¹³)
	实验二	44% (C=2 ³ , G=2 ⁻⁵)	52% (C=2 ¹ , G=2 ⁻⁵)	50% (C=2 ⁻⁵ , G=2 ⁻¹)
	实验三	52% (C=2 ¹ , G=2 ⁻³)	54% (C=2 ⁻⁵ , G=2 ⁻¹¹)	54% (C=2 ³ , G=2 ⁻⁷)
	实验四	50% (C=2 ⁻⁵ , G=2 ⁻¹)	44% (C=2 ⁻⁵ , G=2 ⁻³)	54% (C=2 ⁻⁵ , G=2 ⁻¹⁵)

5.5.3 结合时域与功率谱信息进行分类

根据 5.3 的分析结果，功率谱信息对女性与男性相减能量差异上有一定的显著性，并且功率谱的分类精度更高，因此尝试将功率谱信息结合时域信息期望提高分类器的精度，结果如表 17 所示。

表 17 100 组数据 19 × 20 + 19 = 399 维特征 RBF 核分类精度

被试	实验	45-85ms	145-185ms	200-250ms
被试一	实验一	63% (C=2 ⁷ , G=2 ³)	65% (C=2 ³ , G=2 ⁵)	67% (C=2 ³ , G=2 ¹)
	实验二	55% (C=2 ¹⁵ , G=2 ⁻⁵)	49% (C=2 ⁻⁵ , G=2 ⁻⁵)	57% (C=2 ⁻⁵ , G=2 ⁷)
	实验三	60% (C=2 ¹³ , G=2 ⁻⁵)	55% (C=2 ⁵ , G=2 ¹)	59% (C=2 ⁹ , G=2 ⁻¹)
	实验四	54% (C=2 ⁻⁵ , G=2 ¹)	58% (C=2 ⁵ , G=2 ⁻¹)	56% (C=2 ¹¹ , G=2 ⁻³)
被试二	实验一	52% (C=2 ⁻⁵ , G=2 ⁻⁵)	63% (C=2 ¹ , G=2 ³)	59% (C=2 ⁹ , G=2 ⁻⁵)
	实验二	61% (C=2 ⁵ , G=2 ¹)	62% (C=2 ⁻⁵ , G=2 ⁵)	52% (C=2 ⁻⁵ , G=2 ¹³)
	实验三	64% (C=2 ¹¹ , G=2 ⁻⁵)	55% (C=2 ⁻⁵ , G=2 ³)	53% (C=2 ⁻⁵ , G=2 ¹³)
	实验四	56% (C=2 ⁻⁵ , G=2 ⁻¹)	50% (C=2 ⁵ , G=2 ³)	56% (C=2 ⁻⁵ , G=2 ⁻¹)

5.6 本章小结

本章围绕时域分析、功率谱分析、小波分析介绍了数据分析结果，并使用 SVM 进行分类，得到了以下初步的实验结果：

(1) 刺激后 45-85ms、145-185ms 时间区间内，脑后电极 P5、CP2、P8、O1 可能存在绝对性的结论，即在相应区间内，一种刺激幅值趋势高于另一种刺激。刺激后 200-250ms 时间区间内，对于被试右脑枕颞三个相近电极有相同趋势，但有可能是电极局部短路，需要进一步验证。

(2) 被试一对于 ERP 波形的 T 检验 ($\alpha = 0.05$)与之前结果^[2]提到的结果基本一致，即在 45-85ms、145-185ms、200-250ms 时间区间内女性刺激对于男性刺激有显著性差异；而被试二的 ERP 波形却完全不一致。

但观察被试一 45-85ms 区间内通过 T 检验的时间点，检验结果更吻合 50-90ms (Seeck et al., 1997)的时间区间。

(3) 在时域中，女性刺激与男性刺激间，对于脑后 19 个电极整体，各个刺激没有绝对性结论，即女性刺激与男性刺激所诱发的波形之间没有幅值大小之间的显著差异 ((A>B):(B>A)=50%:50%)。

(4) 在每个实验中，脑后 19 个电极都存在电极子集，女性刺激与男性刺激的波形之间的幅值关系达到统计学显著性水平，即子集中的电极的幅值差通过 T 检验 ($\alpha = 0.05$)。但各个实验间不存在相同的子集，子集的相似度需要进一步计算。

(5) 对于功率谱，进行与(3)相同的检验，对于脑后 19 个电极整体，功率谱更能体现出女性刺激与男性刺激的波形之间的幅值大小之间的显著差异((A>B):(B>A)=60%:40%)。

(6) 在对被试一的实验一至实验四的连续小波分析中，三个区间 45-85ms、145-185ms、200-250ms 都没有发现如 P300 实验中显著的相同模式，但很好的验证了被试一在这三个时间区间内女性刺激对于男性刺激的显著性差异，并在 200-250ms 中存在检验结果相同的时间-尺度点。

(7) 对观察到的结果使用 SVM 进行分类，其中使用时域的整体参数所得到的分类精度最高 68%，而无论时域、功率谱、两种方法结合，分类精度最高时间区域集中在 200-250ms，且平均精度最高。

按照时间顺序划分训练集的方法普遍存在分类器只选择一种性别的情况，分析原因可能由于样本分布不均造成，也从另一个方面说明，抽取的特征并不显著。

结果表明与 45-85ms、145-185ms 时间区间相比，200-250ms 各被试之间共性较大，为对人脸进行进一步性别处理的过程^[2]，而 45-85ms 出现在结构编码之前，而 145-185ms 为 N170 出现的区域，为结构编码阶段。

第六章 结论及下一步工作

6.1 结论

(1) 从而对于性别处理可以提出三个阶段的假设，即：

- 50-90ms 阶段为无意识阶段，即大脑对初步视觉信息进行处理时，有附带的处理性别信息，而这些信息仍驻留在视觉区，因此，人在这个阶段过后没有性别判断的意识。这个阶段过后脑已经对性别进行了简单初步的编码，但未影响到大脑有意识区域。

- 145-185ms 阶段为结构编码阶段，视觉信息两路并行被具体编码，处理识别信息（N170 相关）与其它脸部特征（性别等），且后者不影响前者编码^[2]。

- 200-250ms 阶段为有意识阶段，即人能够获取性别信息，并做出判断。

在之前的研究结果^[2]中，也根据当时的实验情况，提出了类似的假设。

(2) 每个人对于性别的判断存在差异性，即各人之间存在差异。

从分析与分类的结果观察，各人之间、各个实验之间，数据的特征都存在着差异。

一部分数据通过区间整体参数分类更能体现数据类别之间的差异，而另一部分数据通过功率谱分类更能提高分类精度，因此可以猜测每个人进行性别判断的脑分布甚至结构是有差异的。

6.2 下一步工作

(1) 验证结果

由于完成实验的被试共有三人，且被试三的数据由于两个电极实验的损坏、一个实验数据的丢失、以及实验数据中肌电伪迹很多，所以需要再进行实验并对实验进行调整后再作验证。

实验中的部分结论需要通过进一步实验进行验证：

- 对于 5.6 的结论(1)，被试在刺激后 200-250ms 区间内右脑枕颞三个相近电极有相同趋势，需要验证是否这个趋势是普遍的

- 对于 5.6 的结论(2)，通过 T 检验，被试一对于男性刺激与女性刺激在幅值上的假设：女性刺激的幅值与男性刺激的幅值有差异，达到了统计学显著性水平，并且与之前结果^[2]所提出的区间完全的吻合，说明这部分时间区间的波形不仅仅对于有性别任务与无性别任务存在差异，而且对于男性与女性也存在差异。然而，被试二的结果与被试一的截然不同，即没有区间显示出显著性差异。

- 进一步验证对于女性被试、左利手被试是否有前两个结论。即 200-250ms 区间内右脑枕颞三个相近电极有相同趋势，及女性刺激的幅值与男性刺激的幅值是否存在差异。

(2) 寻找子集间的关系

根据 5.6 的结论(4)，枚举了脑后 19 个电极的子集，期望在各个实验中找到是否有共同的子集，即同样的脑工作区间处理性别信息，但结果表明没有共同的子集。

对于这个问题进行进一步的考察，可以尝试定义并计算子集间的相似程度：构造电极间的图，图中结点间的权值定义为电极间物理距离，从而计算两子集间电极分布接近程度。

也可以尝试观察是否存在两个子集、或多个子集的并是全集的情况，即考虑这些子集之间是否存在协同工作。

(3) 进一步使用小波分析

本文使用了连续小波分析进行波形的分析，而小波分析非常适合时域与频域的切换，因此可以参阅一些文献，寻求更多小波分析在 EEG 分析中的应用，如离散小波分析。

(4) 对 SVM 分类结果进行分析

从 SVM 的分类结果，观察是哪些支持向量导致了超平面的移动，从支持向量出发分析性别分类的规律，从而验证或调整假设，并提高分类精度。

(5) 部分分类再使用线性核尝试

对性别分类的处理大部分采用 RBF 核，一部分原因是观察第一个分类结果发现，RBF 的分类精度优于线性核的结果；另一部分原因是线性核对于部分训练数据并不收敛。

(6) 进一步的寻求去除伪迹的方法

被试三的数据并没有应用到大部分的数据处理中，一部分原因是数据中存在两个位于脑后的坏电极，对于数据的处理结果影响较大；另一部分原因是伪迹较多充斥在整个波形中，干扰很多。因此需要进一步寻求去除伪迹的方法，如离散小波变换从一定程度可以分离高频信号等。

(7) 完成三幅人脸的分析工作

本文的工作涉及三幅人脸的工作较少，如图 22 所示，TP8 电极在 200-250ms 时间区域内不同性别波形明显高于女性波形，且差异明显，而这个现象在脑后大部分电极中出现，并且电极由后向前差异呈递增的趋势。

因此可以对 200-250ms 时间区间内进行相关检验与分类。

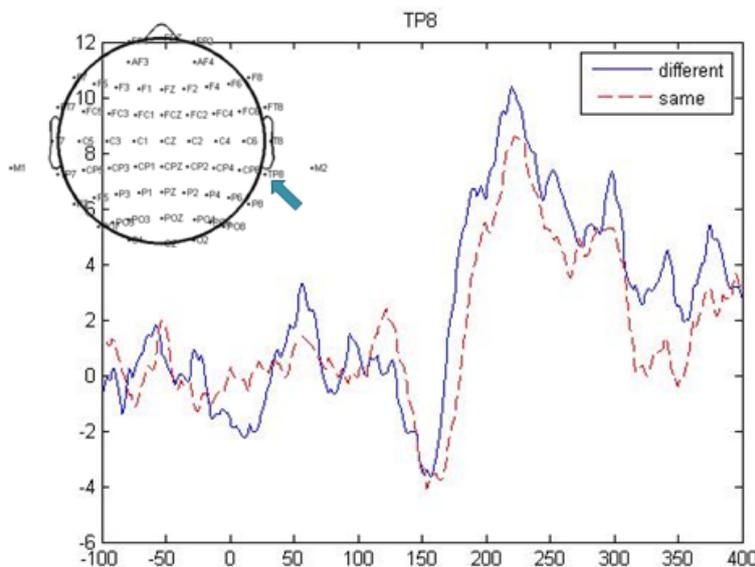


图 22 被试一实验五 三幅人脸 TP8 电极不同与相同刺激对比波形

参考文献

- [1] 赵化. ERP 实验教程[M]. 天津社会科学院出版社, 2004
- [2] Yolande Mouchetant-Rostaing, Meria-Helene Giard, Shlomo Bentin, Pierre-Emmanuel Aguera, Jacques Perneir. Neurophysiological correlates of face gender processing in humans[C]. European Journal of Neuroscience, Vol.12, pp. 303-310, 2000
- [3] 盛骤等. 概率论与数理统计[M]. 上海:高等教育出版社, 2003
- [4] 孙鹤, 吕宝粮. 基于局部 Gabor 二值映射和 SVM 的性别分类[J]. 计算机工程, 2009 35(2)
- [5] 冯卫国. 积分变换 上海:上海交通大学出版社, 2007
- [6] Michel Misiti, Yves Misiti, Georges Oppenheim, Jean-Michel Poggi, Matlab Wavelet Toolbox[M]. The MathWorks.1997
- [7] Arnaud Delorme, Toby Fernsler, Hilit Serby, Scott Makeig. EEGLAB Tutorial[M], 2006
- [8] Chih-Wei Hsu, Chih-Chung Chang and Chih-Jen Lin. A Practical Guide to Support Vector Classification[M],2009
- [9] Vladimir Bostanov. BCI Competition 2003 – Data Sets Ib and Iib: Feature Extraction From Event-Related Brain Potentials With the Continuous Wavelet Transform and the t-Value Scalogram[J], IEEE Transactions on Biomedical Engineering Vol. 51, No.6 June 2004
- [10] 杨帮华, 颜国正, 鄢波. 基于离散小波变换提取脑机接口中脑电特征[J], 中国生物医学工程学报. 2006.25 卷 5 期
- [11] Y. Mouchetant-Rostaing and M.H.Giard, Electrophysiological Correlates of Age and Gender Perception on Human Faces[J], MIT Press, Vol. 15, No.6, Pages 900-910.

谢辞

感谢吕宝粮老师在选题、进行实验、数据分析及论文写作过程中给予的指导，为我在论文题目可行性方面提供了许多宝贵意见，提供了许多相关资料让我阅读；在我没有头绪的时候，指点我下一步的工作，给我继续研究下去的信心；在论文写作过程中，为我指出论文中许多不恰当的表达。

感谢石立臣学长在选题、进行实验、数据分析过程中给予我的帮助，最初选题之后，为我讲解 EEG 实验的过程；对我设计的实验提出许多修改建议；在分析数据时提出许多全新分析的角度，帮助我开拓思路；在论文写作时，不仅仅指出我逻辑上存在的问题，还抓出我许多的语病。

感谢李沐学长、王晓韡学长、施少杰学长在实验设计阶段给我提出的修改建议，让我少走了许多弯路。

感谢石立臣学长、王晓韡学长、Georg、刘晨在闲暇之余完成我设计的实验，为我提供了原始的数据，并在每次实验过后，提出实验设计不恰当之处。

译文及原文

译文

人类人脸性别的神经生理学关联

摘要

事件相关电位(ERPs)在被试在进行三个基于人脸和人手背的性别处理实验时被记录。在第一个实验中,为了防止任何的性别差异,所有的刺激都是同一性别的;在第二个实验中,刺激中既有男性,也有女性的脸部,但被试要求完成的任务与性别无关,所以区分性别应当是附带的;在第三个实验中,任务需要被试明确的指出性别,因此此时区分性别是有意识的。位于枕颞处,性别处理约在刺激后 170ms 处的负电位(即 N170)没有任何作用,说明与人脸具体编码有关的神经机能不同于性别相关的特征提取。相反,与手部特征不同,附带的与有意识的人脸性别处理影响了 145ms 至 185ms 位于前头骨位置的电位。这个现象可以被 Bruce 和 Young 提出的人脸视觉处理模型解释。另外,性别区分效应在 45ms-85ms 时被观测到,无论是人脸还是手,但只有在附带实验中。这个区别可以尝试性的假设它反映了早期的粗糙视觉分类的机制。最后,有意识的性别处理对 ERP 后期有影响,开始于 200ms,对于人脸结束于 250ms,对于手则结束于 350ms。这个后期效果可能与集中注意力的性别分类或更加一般化的分类过程有关。

绪论

在关于视觉系统能够快速有效的在数千个样本中区分同一种类方面,人脸识别应该是一个最好的例子。人类不需要任何训练,就能够轻松准确地辨别人脸是否熟悉、是男是女、正确的估计他的情感、表情、年龄和种族等等。

人脸识别过程的复杂性和多样性在 Bruce 和 Young 在 1986 年提出的理论模型中很好的被表述出来。这个模型假设存在两条分离的通道来处理人脸:一条用来识别人;另一条并行的处理年龄、种族、性别或人脸所表达的情感。这两个不同过程的起始点是“结构编码”模块的提出,根据模型,在脸部特征被结构编码和形成直接视觉编码之后,年龄、种族和性别的判断会被激活。

这个模型的很多方面都根据经验主义和神经生理学的证据得到支持。如:面貌不能识别的病患会意识到自己看到到的是人脸,而且他们中的很多人都能够在没有任何识别象貌的意

识下,将同一个人的不同照片匹配,这个发现表明,一般事物的区分识别与人脸的区分识别,这两项的功能是独立的(Benton & Van Allen, 1972; Malone 1982; De Renzi 1991)。显然,因为人脸识别依赖于有效的结构编码(Bentin, 1999; Eimer & McCarthy 1999),神经生理学分离这两个过程是不可能的。然而,这两个过程的独立性最近在一次普通被试进行 ERP 实验时被发现(Bentin & Deouell, 1999)。人脸识别系统和另一个叫做面部情感表情偏好的系统被神经生理学证明是分离的。绝大多数的面貌不能识别的病患能够分辨不同表情,但有些能够识别熟悉人脸的病患却不能分辨不同表情(Kurukz & Feldmar, 1979)。相对于人脸识别和面部情感表情偏好,直接视觉编码模块很少被提及。但也有例外(NewCombe & Russel, 1969; De Renzi, 1989; Carlesimo & Caltagirone, 1995),面貌不能识别的病患能够通过观察人脸区分人的性别、种族和估计年龄。这个发现很大支持了 Bruce 和 Young 提出的这些过程和人脸识别过程的分离,因此即使结构编码是有缺陷的,但直接视觉编码也可能是完整的(Bentin, 1999)。这个结论表明,性别和年龄可能不依赖于结构编码的输出。当然,根据我们的知识,关于结构编码和有关人的外表的直接视觉编码的研究还没有出现过。这篇论文的目标就是填补这个空缺。

一个人脸特有的位于枕颞的 ERP 峰值分量(N170),在位于刺激发生后 170ms 左右。这个特征最近被发现与结构分析人脸特征有联系。N170 不受人脸相似度的影响,这表明 N170 与人脸识别无关,可能与不受自顶向下的信息的影响(Bentin & Deouell, 1999)。在这篇文章中将比较被试被要求进行性别处理时与性别无关的情况下 N170 的差异。准确的说,提出了以下问题:(1) 假设和人脸结构分析有关的 N170 分量是不是对性别判断敏感。如果根据性别处理的需要,直接视觉接收和结构编码有相同的神经机制,这将是 N170 所表现的一个任务;(2) 有没有可能将对于面部的结构编码的不同功能和性别处理分开。如果性别处理是通过不同于结构编码的机制实现的,比较两个不同情况实验下的 ERP 可能能够揭示这种活动是在 N170 之前还是之后出现。值得注意的是,如果关于性别处理的分量早于 N170 出现的话,那么脸部的直接视觉处理和结构编码无关就相当的有说服力。最后,做好可能性别处理与脸特异性无关的准备。这个假设通过比较对脸和手的 ERP 时进行检验。

素材和方法

被试

16 个右手利的被试,其中 8 位男性和 8 位女性,年龄在 18-32 岁,有正常的视力或矫正后正常的视力。所有都事先告知了纪录技术和无伤害性。

刺激

刺激是 1200 张通过数字扫描的灰度照片。其中有 4 种类型的刺激:不戴眼镜的人脸照片、戴眼镜的人脸照片、人的手部、人的躯体。人脸和躯体是大致正面的视图,手部是手背的照片,一半是左手、另一半是右手,从上向下俯视。一半的脸、手、躯体是男性的,另一半则是女性的。从光通量上,所有的照片是匹配的、且是一样的尺寸(100*120mm)。同时,刺激中央集中的,视觉的角度为 6.9 度,每个刺激持续 300ms,中间间隔 1400ms。每次开始之前在中心都会出现一个固定点。没有任何的刺激是重复的。

实验过程

实验包括两个连续的过程,每个过程包含五个实验情况。每次实验都进行 50 个刺激。在第一个过程(人脸过程)中,会出现面部刺激;同时,在第二个过程(控制过程)中,会出现手和躯体的刺激。在每个实验中,被试被要求进行特定刺激的检测任务,并在实验后,

口头的告知靶刺激的个数。靶刺激(20%)随机的分布在非靶刺激中。在情况一和情况二中,也就是无性别差异任务,靶刺激和非靶刺激都是来自于同一性别的。在人脸过程中,靶刺激是男性(情况一)和女性(情况二)戴眼镜的照片,非靶刺激是不戴眼镜的照片。而在控制过程中,靶刺激是男性(情况一)和女性(情况二)的躯体,非靶刺激是手的照片。情况三(附带的性别区分任务)与情况一和情况二相似,但不同点在于,男性和女性等价的混杂在一起。假设对被试反应的分析与情况一与情况二是同一性质的,但被试可以附带的处理性别信息。情况四和情况五(有意识的性别区分任务)的先后在所有被试中平均。在这两个情况中,被试必须有意识的区分刺激的性别:情况四要求被试从女性中找出男性,而情况五要求被试从男性中找出女性;而在控制组中,要求被试分辨手的性别。

EEG 记录

EEG 的采样率为 1kHz(0.03-320Hz 的模拟带宽), 32 个 Ag-AgCl 的电极, 以鼻尖作为参考点。其中 20 个电极根据 10-20 标准放置(Fz, Cz, Pz, Iz, Fp1, F7, F3, C3, P3, T3, T5, O1 和他们在右半脑的对称点)、2 个位于左右耳后的电极(M1, M2)。剩余的电极是 10-20 系统电极的中点: CP5(在 T3 和 P3 之间)、CP1(在 C3 和 Pz 之间)、PO3(在 O1 和 P3 之间)、POz(在 Pz 和 Oz 之间)、IM1(在 Iz 和 M1 之间)以及他们在右半脑的对称点。眼电右眼外眼角和 Fp1、Fp2 进行记录。电极的阻抗小于 $3k\Omega$ 。ERP 平均波形总长为 1000ms, 包含刺激前 100ms 进行基线矫正, 使用 0.03-20Hz 的滤波器进行滤波。任何一个电极的电压超过 $150\mu V$ 的波形自动被去除。

数据分析

只有对非靶刺激的反应进行分析。男性脸部和女性脸部和女性手(情况一和二、情况四和五)的 ERP 分成三个层次过行分析: 无性别差异、附带的性别差异、有意识的性别差异。每个层次的刺激数量是相同的。分析 ERP 分为两个步骤:

- (i) 性别处理在 N170 的滞后和幅值的变化用三路 ANOVA(F 检验), 三个参数分别为任务的类型(无性别差异、附带的性别差异、有意识的性别差异)、刺激的类型(脸、手)、左右手。N170 的滞后在 T5/T6 两个电极进行检测, 因为这两个电极的平均波形 N170 的幅值最大。N170 的幅值定义为 N170 峰值前后各 25ms 的窗口里电位的平均值。

相对于人脸, 空间平均的不同子集反应了更多前部电极对于手部刺激诱发的 ERP。对于人脸, 包括 IM1, M1, T5 和右半部的对称电极; 而对于手部, 包括 T5, P3, CP5 和右半部的对称电极。F 检验的差异度由 Greenhouse-Geisser 矫正。

- (ii) 另外一些性别处理的效果可以通过以下方式观察到。首先, 附带任务和有意识任务被量化为与无性别任务相减的结果。与幅度为 0 的波形的 T 检验被应用到每个电极的每个采样中。从而能够得到每个滞后的 T-Map。两个相近的电极在 20 个连续的 20ms 的采样中时空模式有稳定的拓扑关系和显著的幅值时, 能够被视作为显著的性别任务效果(Rugg, 1995; Thorpe, 1996)。

脑电地形图使用 two-dimensional spherical spline interpolation(Perrin, 1989), 颜色尺度在有效滞后中被归一化。

结果

对于所有的刺激类型和性别任务, ERP 波形有一系列正、负峰值组成, 在 O1-O2 的 100ms 左右(正峰值 P100)、T5-T6 的 165ms 左右(负峰值 N170)、PO3-PO4 的 220ms 左右(正

峰值), 这个正峰值之后, 有一个稳定的、持续长得 350ms-400ms, 形态根据性别任务和刺激类别而不同。

性别处理在 N170 分量的效果

图 2a 显示了在无性别差异任务, 刺激后 165ms, 人脸和手背诱发的 N170 分量的脑电地形图。两个刺激类型都激活了枕颞的电位, 但手背较人脸更靠前一些。性别区分任务对于 N170 的幅值和滞后的影响通过三路 ANOVAS (三路 F 检验), 其中任务类型、刺激类型和脑半区 (左脑和右脑) 作为检验的参数。对于这种类型任务, 没有特别主要的效果, 而这些参数之间、以及和其他参数之间在平均幅值和滞后上没有发现联系。唯一显著效果是相对于手背(170ms), 脸部(157ms)有更小的峰值($F_{1,15}=60.39, P<0.0001$)。

性别区分任务的其他 ERP 效果

对于面部和手部的刺激, 附带的和有意识的性别分类根据 ERP 波形与无性别任务的波形的差做为量化的标准。对于差的 T 检验揭示了三个关于时空特性的显著任务效果, 根据性别区分任务的本质 (附带的和有意识的)、刺激类型 (脸和手)。最早的分量差异出现在刺激后中间顶骨区域, 大约 40-90ms; 第二个在中央区域, 大约 140-190ms; 最后出现在枕骨和顶骨区域, 大约 200ms-350ms。图 3 和图 4 详细说明了电极和时空之间不同波形的统计学的显著性。图 2B 和 C 显示了不同波形的拓扑性和对每个波形滞后对应的 t-map。表 2 给出了全体波形的平均幅值和差的显著性。这些差异将在以下进行详细的分析。

在 40-90ms 的 ERP 波形

在 40ms 至 90ms 之间 (图 2B、C, 线 a), 附带的性别任务显著效果 (与无性别任务比较) 在面部和手部刺激中都被发现: 附带性别任务的 ERP 波形减去无性别任务的 ERP 波形的幅值在 45-70ms 的左侧顶骨处, 当刺激是人脸时, 统计学上显著性($P<0.01$)相当高 (图 3A 的左侧; 40-90ms 的平均幅值, $-0.74 \mu V$)。而对于手部, 40-90ms 在中央顶骨处 (图 3A 的右侧; 平均幅值为 $-0.79 \mu V$)。对于图 2B 中的线 a, 说明了在刺激后 65ms 的脑电地形图。没有观察到其他相关的模式出现在有意识的性别任务中, 无论是手或脸。

在 140-190ms 的 ERP 波形

在这个滞后时间范围中 (图 2B、C, 线 b), 附带的性别任务和有意性的性别任务与无性别任务的差的波形都表现出在大部分中央区域 (八个电极) 对于脸部有显著的正向幅值, 而对于手部刺激则没有。对于面部, 附带的性别任务与无性别任务的差的波形在 F3-C3 电极位置, 在位于 145ms 的滞后处到达了显著的统计学依据($P<0.01$), 而且在所有选定的电极 (统计学都达到了 $P<0.01$ 或 $P<0.001$), 在 155ms-185ms 处。在有意性的性别任务与无性别任务的差的波形中, 145ms 之后的幅值在大部分电极都达到了统计学意义($P<0.01$), 而且在所有选定的电极在 150-185ms 统计学方面都达到了 $P<0.01$ 或 $P<0.001$ 。对于手部, 没有或只找到边缘的显著效果。

在 200-250ms 的 ERP 波形

在 200-250ms 之间有明显的有意性的性别效果, 无论是人脸还是手, 对于枕颞电极。有意性性别实验与无性别实验之差的 ERP 幅值对于人脸和手在 200-250ms 在统计学上达到了显著性水平。这个差异对于手较脸的刺激持续更长的时间 ($\geq 250ms$)。在这个时间区间中只有边缘显著性在附带的性别实验中被观察到。

因为这个有意识性别任务的效果在后部时间区间里最大，它的脑电地形图与产生 N170 时的地形图进行了对比，观察是否这两个活动由同一个大脑结构产生。归一化数据通过两路（平均幅值在峰值前后 25ms，N170 和 200-250ms 的有意识性别效果）ANOVA，参数是条件（无性别实验、有意识性别实验）和电极（IM1, IM2, M1, M2, O1, O2, T5, T6）进行检验，结果表明在统计学上达到了显著性水平[F4,55, $P < 0.004$, $GG = 0.52$]，说明这两个波形不同的脑电地形图。

在 250-350ms 的 ERP 波形

在 200-250ms 观察到有意识的性别任务效果，而且对于手时间持续得更长。显著的幅值显著性($P < 0.001$)和稳定的脑电地形图在 350-400ms 被观察到。而在脸部刺激的波形中在这个时间区间内没有相应的模式。

讨论

结果可以总结为以下几点：

- (i) 性别处理对于 N170 的幅值和滞后没有影响。
- (ii) 附带的性别处理所诱发的电活动和完全一类性别诱发的完全不同。这个差异对于脸和手的刺激后 45-85ms、位于中部枕叶电极，而在 145-185ms 只对脸的刺激有不同。
- (iii) 有意识的性别处理所诱发的电活动与完全一类性别诱发的电位显著的不同。与附带的性别处理相似的是，这个不同只对脸部刺激在 145-185ms 有显著差异。然而，与附带的性别处理相反的是，有意识的并没有早期的反应。取而代之的是，有长滞后的效果，即对于脸和手的刺激，在 200-250ms 后对于后部电极有差异，这个效果对于手更长(至 350ms)。

对于脸部特征的结构编码与性别处理的分离

在枕颞出现针对脸部刺激的 N170 已经与脸部特征的结构编码的神经机制相联系。最近研究的一项重要结果表明，性别处理不影响 N170 的幅值和滞后。尽管例外的人脸性别区分的效果在相同时间区间内被观察到，但电极与有 N170 分量的电极不同。这个形式表明了人脸性别处理过程和由不同的大脑机制负责的结构编码是同时进行的，例如，性别处理过程可以在结构编码过程完成之前开始。这个结论和 PET 的发现一致，表明了人脸性别处理过程活跃在神经解剖学的区域，不同于那些和感知分析相关的人脸特性。因此，我们的结论为 Bruce 和 Young 的人脸处理模型提供了另外的支持，根据这个模型，感知分析和性别提取得益于两个功能上并不相同的模块—分别是结构模块和针对视觉的编码模块。和这个模型形成对比的是，对特定人脸的性别处理产生影响的时间区域（145-185ms）表明了直接感知并不紧随着结构编码发生。

早期时间的（45-85ms）附带区分影响

我们研究中的一个意外的结果是，在极早的时间段内（45-85ms），脸与手的附带区分实验与无性别实验的差存在差异。显然，这个结果表明，这个时间点有效的视觉信息已经被用来进行自动的性别区分。尽管对于猴子与人，脸部处理最早时间段分别为 20-50ms 和 50-90ms。然而，很难证明在 45-85ms 位于枕颞的效果就与性别处理相关。确实，这种滞后是由视觉刺激在纹状皮质诱发的第一皮层电位。另一个关于早期差异的解释是，附带的性别任务只是 5 个情况中的一个，并且都把两种类别的非靶刺激混杂在一起。因此，这个影响可能反映了一

个早期的、自动的过程，这个过程快速但粗糙地在视觉特性的基础上对两类刺激进行分类。最终，这个早期的分类过程可能只对生物刺激有反应。

人脸性别处理的特异性

性别处理对脸和对手的影响形式既有相同点，也有不同点。正如上面描述的，早期时间内附带性别处理的影响类似于两种类型的刺激。这个相似性说明了感知机制和这个活动相关，无论它们的属性是什么，都不是脸部特定的。在 145ms 到 185ms 间，从刺激开始，受影响的是脸部诱发的 ERP 而不是手诱发的。这些结果有两个含义。首先，附带的和有意识的性别区分 ERP 影响的相似性暗示了性别信息是从脸部自动提取的，在这个程度上，这个过程不受自顶向下战略因素的影响。其次，因为这些 ERP 影响没有在手上发现，它们可能指示了腔调人脸识别过程的生理操作上的特异性：从身体的其他部分例如手传来的性别判断可能通过其他的大脑机制被调解。

附加的性别区分效应

无论是对脸还是手，在有意识性别实验中，刺激后 200ms 后诱发的 ERP 与那些不需要明显区分性别不同。这个效果对于脸持续到 250ms，对于手则是 350ms。这个形式对于部分脸和手是共同存在的，但与 Bruce 和 Young 提出的直接视觉处理通路的模型不相容。然而，这个结果表明，性别处理的类型在这个层面上与刻意的判断性别有关。事实是，这个效果对手持续的影响比对脸长的原因可以被解释为从手区分性别比从脸区分难度大很多。这可能是因为脸比手携带了更多的性别特定信息，可能是人们被训练得更多的是以脸为基础来区分性别而不是以手为基础，可能是因为脸的性别已经被直接的视觉感知机制所决定，或者可能是上述所有原因。缺少一个在相似刺激下的无性别区分实验，我们不能毫不含糊的判断出是后期时间中观察到的活动有意识，而不是附带的性别区分特别地和性别处理过程相关，或和一个更为一般的区分活动类型相关。无论如何，这些区分的影响可能表达了，相较于身体其它部位，例如手，人们在对脸的信息处理上更为专业。

缩写

ERP, 动作相关电位;

N170, 刺激后 170ms 处的负电位。

原文

Neurophysiological correlates of face gender processing in humans

Abstract

Event-related potentials (ERPs) were recorded while subjects were involved in three gender-processing tasks based on human faces and on human hands. In one condition all stimuli were only of one gender, preventing any gender discrimination. In a second condition, faces (or hands) of men and women were intermixed but the gender was irrelevant for the subject's task; hence gender discrimination was assumed to be incidental. In the third condition, the task required explicit gender discrimination; gender processing was therefore assumed to be intentional. Gender processing had no effect on the occipito-temporal negative potential at ≈ 170 ms after stimulation (N170 component of the ERP), suggesting that the neural mechanisms involved in the structural encoding of faces are different from those involved in the extraction of gender-related facial features. In contrast, incidental and intentional processing of face (but not hand) gender affected the ERPs between 145 and 185 ms from stimulus onset at more anterior scalp locations. This effect was interpreted as evidence for the direct visual processing of faces as described in Bruce and Young's model [Bruce, V. & Young, A. (1986) *Br. J. Psychol.*, 77, 305-327]. Additional gender discrimination effects were observed for both faces and hands at mid-parietal sites around 45-85 ms latency, in the incidental task only. This difference was tentatively assumed to reflect an early mechanism of coarse visual categorization. Finally, intentional (but not incidental) gender processing affected the ERPs during a later epoch starting from ≈ 200 ms and ending at ≈ 250 ms for faces, and ≈ 350 ms for hands. This later effect might be related to attention-based gender categorization or to a more general categorization activity.

Introduction

Human face recognition is probably one of the best examples for the capacity of the visual system to rapidly and efficiently discriminate among thousands of exemplars of the same category. Perceiving a face, people are able, without explicit training, to decide accurately and effortlessly whether the face is familiar or unfamiliar, one of a male or female, and to correctly appreciate its emotional expression, age, race, etc.

The complexity and variety of the cognitive processes involved in face recognition have been well represented in the theoretical model suggested by Bruce & Young (1986). This model assumes the existence of separate face processing pathways: one is designed to identify the person, while

others, acting in parallel, are supposed to process the age, race or gender, or assess the emotion expressed in the face. The common starting point of these different pathways is the elaboration of a configural representation of faces through the 'structural encoding' module. According to the model, the mechanisms subtending age, race or gender judgments are activated following the structural encoding of facial features and form the 'directed-visual encoding' module.

Various aspects of this model have been supported, to differing extents, by empirical and neuropsychological evidence. The finding that prosopagnosic patients never fail to realize that they are looking at a face, and many of them are capable of matching different photographs of the same individual without showing any recognition of familiar faces, supports a distinction between a fairly normally functioning structural encoding module and a deficient face identification system (Benton & Van Allen, 1972; Malone et al., 1982; De Renzi et al., 1991). Obviously, given that face identification is dependent on efficient structural encoding (Bentin et al., 1999; see also Eimer & McCarthy, 1999), a neuropsychological double dissociation between the face identification and structural encoding mechanisms is impossible. However, the independence of these two mechanisms has been recently shown in normal subjects using event-related potentials (ERP) (Bentin & Deouell, 1999). The distinction between the face identification system and a system devoted to the appreciation of facial emotional expression has also been demonstrated by neuropsychological dissociation. The identification of emotional expressions is preserved in the great majority of prosopagnosics, while it may be impaired in patients who do not have problems in recognizing familiar faces (Kurucz & Feldmar, 1979). Relative to face identification and appreciation of emotional expression, the function of the direct encoding module has been less investigated. With a few exceptions (Newcombe & Russel, 1969; DeRenzi et al., 1989; Carlesimo & Caltagirone, 1995), prosopagnosic patients can discriminate a person's gender and race and can estimate his/her age by looking at his/her face. This finding supports Bruce and Young's (1986) suggested dissociation between these processes and those required for face recognition. However, the finding that the performance attributed to the direct visual encoding is intact even if the formation of the structural representation is deficient (Bentin et al., 1999) suggests that the perception of gender and age may not be based on the output of the structural encoder. Indeed, to our knowledge, studies of the relationship between structural encoding and direct visual perception of physiognomic information (age, race and gender) have not been reported. The goal of the present study was to bridge this gap.

A face-special scalp ERP component peaking on the posterior temporal areas at ~170 ms from stimulus onset (N170) has been recently associated with the mechanisms of structural analysis of facial features (Bentin et al., 1996; George et al., 1996). The N170 was not influenced by face familiarity, suggesting that it is not associated with face identification and probably not influenced by top-down information (Bentin & Deouell, 1999). In the present study, we have compared the N170 elicited while subjects were instructed to process the gender of the face with that generated while the gender was irrelevant. In particular, we asked the following questions: (i) Are the face structural analysis mechanisms, assumed to be associated with the N170 component, sensitive to gender judgments? If direct visual perception, as required for gender processing, involves the same neural mechanisms as those mobilized for structural encoding, we may find an effect of task on the N170. (ii) Is it possible to dissociate, on the basis of separate neural activities, different functional modules for structural encoding of a face and gender processing? If gender processing is performed by neural mechanisms other than those involved in structural encoding, comparing

the ERPs in the two task conditions might unveil such activity elicited earlier or later than the N170. Note that if components earlier than the N170 were modulated by gender processing, it would be conceivable that direct visual processing of the face is not based on the output of the structural encoding module. Finally, anticipating that such activity would be found, we examined the possibility that the gender processing is not a face-specific activity. This hypothesis was examined by comparing the ERPs associated with gender judgments based on faces with those associated with gender judgments based on human hands.

Materials and methods

Subjects

Sixteen right-handed subjects (eight males and eight females, 18-32 years old) with normal vision or corrected-to-normal vision participated as paid volunteers. All were fully informed of the recording technique, which was noninvasive.

Stimuli

Twelve hundred grey-scale photographs were digitally scanned. They included four stimulus types: human faces with no eyeglasses, human faces with eyeglasses, human hands and human torsos. The faces and torsos were presented in a roughly front view, whereas the backs of the hands, half of them being left and the other half right, were viewed from above. Half of the faces, half of the hands and half of the torsos were of men and the other half of women. Across categories, the stimuli were matched for luminance and were of equal size (100 × 120 mm). The stimuli were presented foveally, subtending a visual angle of 6.9° for 300 ms at a rate of one every 1.4 s. A fixation point was present at the centre of the screen between each presentation. No stimulus was repeated.

Procedure

The experiment consisted of two consecutive sessions, each including five experimental conditions. The stimuli in each condition were presented in blocks of 50. Face stimuli were presented during the first session (Face session), whereas during the second, we presented hand and torso stimuli (Control session). In all the conditions, subjects performed an oddball detection task reporting verbally, at the end of each block, the number of target items detected; target items (20% of total) were delivered randomly among non-target items (see Table 1 for details). In conditions 1 and 2, referred to as No-gender discrimination task, target as well as non-target stimuli were of only one gender. In the Face session, the targets were either male (condition 1) or female (condition 2) faces with glasses delivered among faces without glasses, while in the Control session the targets were male (condition 1) or female (condition 2) torsos delivered among hands. The order of presentation of conditions 1 and 2 was counterbalanced across subjects. Condition 3 (Incidental-gender discrimination task) was similar to conditions 1 and 2, except that the stimuli included male and female equally mixed. The subjects process the gender of the stimulus. Conditions 4 and 5 (Intentional-gender discrimination task) followed condition 3 with a counterbalanced order between subjects. In these two conditions, subjects had to process explicitly the stimuli according to their were assumed here to analyze the stimuli at the same level as in the two previous conditions, but they could also incidentally gender: the targets were either male

(condition 4) or female (condition 5) faces among female or male faces, respectively, in the Face session, and male or female hands among female or male hands, respectively, in the Control session.

EEG recording

EEG was continuously recorded at a sampling rate of 1 kHz (0.03-320 Hz analogue bandwidth) from 32 Ag-AgCl scalp electrodes referred to the nose, of which 20 were placed according to the international 10/20 system (Fz, Cz, Pz, Iz, Fp1, F7, F3, C3, P3, T3, T5, O1 and their counterparts on the right hemi scalp), and two at the left and right mastoids (M1 and M2). The remaining positions were midway between two positions of the 10-20 system: CP5 (between T3 and P3), CP1 (between C3 and Pz), PO3 (between O1 and P3), POz (between Pz and Oz), and IM1 (between Iz and M1), and their counterparts on the right hemi scalp. The EOG was monitored from the outer canthus of the right eye (YH) and the channels Fp1 and Fp2. Electrode impedances were < 3 kW. The ERPs were averaged off-line over a 1000-ms period including a 100-ms prestimulus baseline, and digitally filtered (0.03-320 Hz). Trials in which the potential exceeded 150 mV at any electrode were automatically rejected.

Data analysis

Only the responses to non-target stimuli were analyzed. The ERPs elicited by male and female faces or male and female hands (conditions 1 and 2, and conditions 4 and 5) were collapsed yielding three levels of gender processing (No-gender, Incidental-gender and Intentional-gender discrimination), with an equal number of stimuli presented at each level. Analysis of ERP involved two steps. (i) The effects of gender processing on the latency and amplitude of the N170 component (negativity at ≈ 170 ms) were tested using three-way within-subjects analyses of variance (ANOVAs) with factors: Type of task (No-, Incidental-, Intentional-gender discrimination), Stimulus type (Faces, Hands), and Hemisphere (left, right). Latency of N170 was measured at T5/T6 sites where the potential value was maximum in the grand average over all subjects. The N170 amplitude was quantized for each subject as the mean potential value in time over a 625 ms-window around the peak latency of the grand average. Reflecting the slightly more anterior scalp distribution of the N170 elicited by hands relative to faces, different subsets of electrodes were included in spatial averaging. For faces, the sites included were IM1, M1, T5 and their homologous sites on the right hemi scalp, and for hands, T5, P3 and CP5 and their homologous locations on the right hemi scalp. Significance levels for the F-values were adjusted with Greenhouse-Geisser correction whenever necessary.

(ii) Other possible effects of gender processing were assessed in the following way. Firstly, they were quantized in the difference waves obtained by subtracting the responses to the no-gender task from the responses to incidental (Incidental-No) and intentional (Intentional-No incidental-gender task (compared with the no-gender task) were No) gender tasks, respectively. Student's t-tests comparing the amplitude of the difference waves with zero were computed for each sample at each electrode. Student's t-maps could then be obtained at each latency. Spatio-temporal patterns having a stable topography with a significant amplitude ($P < 0.01$) at at least two adjacent electrodes for 20 consecutive 20-ms samples were considered as significant gender task effects (Rugg et al., 1995; see also Thorpe et al., 1996).

Scalp potential maps were generated using a two-dimensional spherical spline interpolation

(Perrin et al., 1989) with colour scale normalized to the peak voltage value at the considered latencies.

Results

For all stimulus types and gender tasks, the ERPs were characterized by a series of positive and negative deflections peaking successively around 100 ms at O1-O2 (a positive peak P100), around 165 ms at T5-T6 (a negative peak N170), and around 220 ms at PO3-PO4 (a positive peak) (Fig. 1). This latter peak was followed by a sustained positivity, lasting up to 350-400 ms, the morphology of which could differ according to the gender task and the stimulus type. Effects of gender processing on the N170 component Figure 2a shows the scalp potential distribution, at 165 ms latency, of the N170 component elicited by face and hand stimuli in the no-gender task. Both stimulus types activated occipito-temporal regions with, however, a distribution slightly more anterior for hands than for faces. Incidence of the gender discrimination tasks on the mean amplitude and peak latency of N170 was tested through three-way ANOVAs with the Type of task, Stimulus type and Hemisphere as factors (see Materials and Methods). No main effect of the Type of task, nor interaction between this factor and any other factors, were found on the mean amplitude or peak latency of N170. The only significant effect was a shorter peak latency for faces (157 ms) than for hands (170 ms; $F_{1,15} = 60.39$, $P < 0.0001$).

Other ERP effects of gender discrimination tasks

Both for face and for hand stimuli, the effects of incidental- and intentional-gender discrimination tasks were quantized in the difference between the ERPs elicited in each of these tasks and the ERPs elicited in the no-gender task (Incidental-No waves and Intentional-No waves, respectively). Student's t-maps computed on these difference waves revealed three spatiotemporal patterns of significant task effects (see Materials and Methods), that could vary with the nature of the gender discrimination task (incidental or intentional) and the stimulus type (face or hand). The earliest component of the difference pattern spread over the mid-parietal areas from 40 to 90 ms, the second was observed over the central areas from 140 to 190 ms, and the latest extended over the occipito-parietal regions from 200 ms to 350 ms. Figures 3 and 4 detail the statistical significance of the difference waves over the electrodes and periods characterizing these spatiotemporal patterns of interest. Figure 2B and C displays the topography of the difference waves and the corresponding t-map at an illustrative latency for each pattern. Table 2 gives the mean amplitude and significance of the difference waves over the whole patterns. These different patterns are analyzed in detail below.

Event-related potential at 40-90 ms

Between 40 and 90 ms (Fig. 2B and C, line a), significant effects of the found for both face and hand stimuli: the ERP amplitudes of the Incidental-No difference waves reached statistical significance ($P < 0.01$) from 45 to 70 ms latency at the left mid-parietal electrodes for faces (Fig. 3A left part; mean amplitude over 40-90 ms, -0.74 mV), and over 40-90 ms at the centro-parietal sites for hands (Fig. 3A right part; mean amplitude 0.79 mV; see Table 2). Figure 2B, line a, illustrates the topography of these effects at 65 ms latency. No corresponding patterns were observed for the intentional-gender task, for either faces or hands (Fig. 2C, line a; Fig. 4A; Table

2).

Event-related potential at 140-190 ms

In this latency range (Fig. 2B and C, line b), both the Incidental-No (Fig. 2B, line b) and the Intentional-No (Fig. 2C, line b) difference waves showed significant positive amplitudes over a wide central region (eight electrodes) for face but not for hand stimuli (Figs 3B and 4B). For faces, the amplitude of the Incidental-No wave reached statistical significance ($P < 0.01$) at F3-C3 from 145 ms latency, and

at all the selected electrodes ($P < 0.01$ or $P < 0.001$) from 155 to

185 ms (Fig. 3B left part; mean amplitude over 140-190 ms, -0.06 mV). In the Intentional-No difference wave, the amplitude reached statistical significance ($P < 0.01$) at most electrodes from 145 ms, and at all the selected electrodes ($P < 0.01$ or $P < 0.001$) from 150 to 185 ms (Fig. 4B left part; mean amplitude -0.26 mV). No or only marginally significant effects were found for hands during this epoch (see Figs 3B and 4B right side, and Table 2).

Event-related potential at 200-250ms

Significant intentional gender task effects were found between 200 and 250 ms latency (Fig. 2B and C, line c) at occipit-parietal electrodes for both face and hand stimuli. The ERP amplitude of the Intentional-No difference waves reached statistical significance ($P < 0.01$) from 200-250ms for faces (Fig. 4C left part; mean amplitude over 200-250 ms, 1.28 mV; see Table 2), and for hands (Fig. 4C right part; mean amplitude, 2.15 mV; see Table 2). This difference lasted longer (> 250 ms) for hands than for faces (see below). No or only marginally significant effects of the incidental-gender task were found over this period for face or hand stimuli (Fig. 3C). As this intentional gender task effect was maximal at posterior sites (Fig. 2C, line c), its topography was compared with that of the N170 component (recorded in the no-gender condition) in order to explore whether the two activities were generated within the same brain structures. Two-way ANOVA on normalized data (mean amplitude over 6 25 ms around the peak latency for N170, and over 200-250 ms latency for the intentional-gender effect) with the factors Condition (No-gender, Intentional-No) and Electrode (IM1, IM2, M1, M2, O1, O2, T5, T6) showed a significant interaction between the two factors [$F_{4,55} = 4.44$, $P < 0.004$, $GG = 0.52$], indicating different topographies for the two waves.

Event-related potential at 250-350 ms

The intentional gender task effect observed between 200 and 250 ms (Fig. 2B and C, line d) lasted longer for hands than for faces (Fig. 2C, lines c and d). Highly significant amplitude differences ($P < 0.001$) with a stable topography were observed up to 350-400 ms (Fig. 4C right part; mean amplitude over 250-350 ms, 4.56 mV; see Table 2). There was no corresponding pattern within the same period for faces (Fig. 4C left part; mean amplitude, 0.31)

Discussion

The results can be summarized in the following way: (i) There was no effect of gender processing on either the amplitude or latency of the N170 component. (ii) Incidental-gender processing elicited electrical activities that were significantly different from those elicited when all stimuli

were of one gender. These differences were observed around 45-85 ms from stimulus onset over mid-parietal sites for both face and hand stimuli, and around 145-185 ms over central sites for faces only. (iii) Intentional-gender processing elicited electrical activity that was significantly different from that elicited when stimuli of only one gender were presented. Similar to the activity generated by incidental-gender processing, this difference was significant only for faces over central sites between 145 and 185 ms from stimulus onset. In contrast to the activity specific to incidental processing, however, intentional processing did not show the early effect. Instead, there were longer-latency effects, between 200 and 250 ms, for both face and hand stimuli over posterior sites, and this effect lasted longer for hands than for faces (up to 350 ms).

Separate processes for structural encoding and gender processing of faces

The occipito-temporal N170 component responding to face stimuli has been associated with the neurophysiological mechanisms underlying structural encoding of facial features (Bentin et al., 1996; George et al., 1996; see also Allison et al., 1994 for intracranial ERP recordings). An important result of the present study is that gender processing influenced neither the latency nor the amplitude of N170. Although exclusive effects of face gender discrimination were observed during similar time ranges, the sites at which these effects were observed were different from those of the N170 component. This pattern indicates that face gender processing is performed (at least partly) in parallel with its structural encoding by different brain mechanisms; i.e. the gender processing stage can begin before the structural encoding stage is completed. This conclusion is also consistent with PET findings that have shown that face gender processing activates neuroanatomical regions different from those involved in the perceptual analysis of the facial features (Sergent et al., 1992; Kapur et al., 1995; Andreasen et al., 1996). Therefore, our results bring additional (electrophysiological) support to Bruce and Young's (1986) model of face processing, according to which perceptual analysis and gender extraction are subserved by two functionally distinct modules, the structural and the directed-visual encoding modules, respectively. In contrast to what was suggested in his model, however, the time-course of the face-specific gender processing effect (145-185 ms) suggests that direct perception is not subsequent to structural encoding.

Early (45-85 ms) incidental discrimination effects

An unpredicted result of our study was the very early epoch (45-85 ms) of the differences observed between the ERPs elicited by both faces and hands during the incidental-gender discrimination task as compared with the no-gender task. Apparently, this result indicates that the visual information available at this time is already used for automatic distinction between genders. Although early signs of face processing have been reported in monkeys (Tovee et al., 1993: 200-250ms) and in humans (Seeck et al., 1997: 50-90 ms), it is difficult, however, to assume that the effects observed at this latency (45-85 ms) over parietal regions are specifically related to the processing of gender information. Indeed, such latencies correspond to the first cortical potentials elicited by visual stimulation in the striate cortex (Clark et al., 1995). An alternative explanation of the early difference between the incidental- and no-gender processing conditions may refer to the fact that the incidental-gender task was the only one among the five experimental conditions in which the two gender categories were intermixed in the non-targets. It is therefore possible that this effect reflects an early, automatic process of rapid but coarse categorization between two

stimulus classes on the basis of visual characteristics. To this end, it would be interesting to know whether this early categorization process may occur only for biological stimuli indicates a more general ability of the visual system to cursorily categorize wide classes of stimuli at an early stage of sensory processing.

Specificity of face gender processing

The pattern of gender processing effects for faces and for hands showed both similarities and differences. As discussed above, the very early effect of incidental-gender processing was similar for both types of stimuli. This similarity suggests that the perceptual mechanisms associated with this activity, whatever their nature, are not face-specific. Between 145 and 185 ms from stimulus onset, however, gender processing (intentional- as well as incidental-) affected the ERPs elicited by faces but not those elicited by hands. These results have two implications. Firstly, the similarity of the incidental- and intentional-gender discrimination ERP effects imply that gender information is extracted from faces automatically and that, at this stage, this process is not influenced by top-down strategic factors. Secondly, because these ERP effects were not found for hands, they might index the specificity of the physiological operations underlying face gender processing: gender judgments from other encoding of body parts such as hands would be mediated through other (later) brain mechanisms.

Additional gender discrimination effects

Regardless of whether the stimuli were faces or hands, the ERPs elicited from 200 ms poststimulation in the intentional-gender discrimination task were significantly different from those elicited when explicit discrimination between the genders was not required. The effect was significant up to 250 ms for faces, and up to 350 ms for hands. This pattern, partly common for faces and hands, is not compatible with the activity of the directed-visual processing pathway stipulated by Bruce and Young (1986). Rather, it suggests that the type of gender processing reflected at this level is associated with the splines explicit process of discrimination between the two gender categories. The fact that this effect lasted longer for hands than for faces may be explained by the higher difficulty of discriminating gender from hands than from faces. This could be because faces carry more gender-specific information than hands, because humans may be better trained to distinguish gender on the basis of the face than on the basis of the hands, because the gender of the face has already been determined by direct visual perception mechanisms, or all of these reasons. In the absence of a nongender discrimination task for similar stimuli however, we cannot unequivocally determine whether the activity observed during the latter time epochs for intentional- but not incidental-gender discrimination is associated specifically with gender processing or with a more general categorical discrimination activity. In any case, these discrimination effects could express the expertise that humans have developed in processing faces compared with other body parts such as hands.

Abbreviations

ERP, event-related potential; N170, negative potential at 170 ms after stimulation.