

空间频率信息对面孔身份与表情识别的影响^{*}

汪亚珉 王志贤 黄雅梅 蒋 静 丁锦红

(首都师范大学心理系, 北京市“学习与认知”重点实验室, 北京 100048)

摘 要 已有面孔身份与表情识别研究提示, 高频空间信息可能选择性地与表情识别有关, 而低频空间信息则选择性地与身份识别有关。为验证这一假设, 操纵空间频率设计三个 Garner 效应测量实验。实验 1 测量全频条件下身份表情识别之间的 Garner 效应, 结果显示, 相互间的干扰均显著。实验 2 测量高频条件下的干扰效应, 发现表情识别的 Garner 效应不再显著而身份识别的 Garner 效应无明显变化, 出现分离。实验 3 测量低频条件下的 Garner 效应, 结果表明, 表情与身份识别的 Garner 效应仍显著, 未受高频过滤影响。基于 Garner 范式, 提出面孔识别的可分离度与难度双指标同时考察的方法, 对实验结果进行了分析, 并由此得出结论, 高频空间信息是面孔身份与表情信息分离的有效尺度。

关键词 面孔识别; 空间频率; Garner 效应; 可分离度; 难度

分类号 B842

1 引言

人能很容易地分别识别面部表情与面孔身份, 认知心理学家与人工智能专家相信, 面孔身份与表情信息识别之间存在着某种可分离特性, 而早期的面孔识别功能模型更是将面孔身份与面部表情识别设置为相互独立的加工路径(Bruce & Young, 1986)。然而, 无论如何假设, 寻找与验证不同类型的面孔信息的可分离尺度无疑是解决问题的关键。但是到目前为止, 采用某种尺度成功演示身份表情识别双分离的研究并不多。

1.1 身份表情识别在空间频率上的双分离推论

White (2002) 用构形或特征改变技术(configural or categorical /featural change technique)考察身份表情识别的双分离。实验中, White 以双眼位置同时移动作为构形改变, 单眼位置移动作为范畴改变, 用再认范式测试发现, 前者的改变更多影响身份识别, 而后者的改变更多影响表情识别。对此, White 基于构形改变与特征改变技术区分出构形关系(configural relation)与范畴/特征关系(categorical/featural relation)两类面孔信息, 并以此解释身份表

情信息之间的分离: 构形信息多用于身份识别, 而范畴信息多用于表情识别。然而, 我们认为, 这种基于构形与范畴改变技术的信息分类有其局限性(White, 2001, 2002)。就 White (2002)实验中的构形与范畴信息的操作而言, 构形既有整体构形也有局部构形, 那么, 双眼同时改变与单眼改变是否就一定对应构形改变与范畴改变呢? 表情识别会否也依赖于某种类型的构形信息? 此外, 构形与范畴信息的界定模糊, 人的视觉加工在抽取这两类信息时应该有更具普遍性的物理尺度作为基础。因此, White (2002)成功地演示了身份表情识别的双分离现象, 但这种构形信息与范畴信息的划分方式尚有局限性。

近来有研究同样对面孔信息进行了分类, 分为高频空间信息与低频空间信息。在空间频率分类的基础上, 研究者们采用构形改变与特征改变技术修改高频与低频面孔图片, 用修改后图片进行面孔匹配实验, 结果揭示低频空间信息主要影响构形加工(configural processing) (即低频信息选择性影响构形加工), 而高频空间信息主要影响特征加工(featural processing) (即调频信息选择性影响特征加

收稿日期: 2010-04-28

^{*} 国家自然科学基金项目(30970891)、北京市教委项目[2008]3 号及教育部博士学科点基金(20091108110001)支持。

通讯作者: 汪亚珉, E-mail: wangym@mail.cnu.edu.cn

工) (Goffaux, Hault, Michel, Vuongo, & Rossion, 2005)。不仅如此, Goffaux 和 Rossion (2006)进一步研究发现, 低频空间信息主要与面孔的整体加工有关, 在面孔识别中先于特征信息被加工, 并且, Goffaux 等(2006)认为, 低频信息加工可能是专家经验加工的基础。如果我们将 Goffaux 等(Goffaux et al., 2005; Goffaux & Rossion, 2006)的研究结论与 White (2002)的研究结论关联起来就会得出这样一个推论: 低频空间信息显著影响构形信息加工, 主要与身份识别相关, 而高频空间信息显著影响局部特征加工, 主要与表情识别相关, 身份表情识别在高低空间频率上出现双分离。为了便于后面表述我们称这一推论为**高低频双分离推论**。情况是否真的如此呢?

Deruelle, Rondan, Salle-Collemiche, Rosset 和 Da Fonseca (2008)研究孤独症儿童对高低空间频率面孔图片进行身份、表情、性别判断, 采用 ABX 匹配任务, 他们发现, 在性别判断上, 两类儿童对空间频率信息的使用无差异, 都倾向于使用低频信息, 但在表情判断上, 孤独症儿童显著偏向使用高频信息, 而正常儿童更多倾向于低频信息。总体而言, 两组都表现出在低频信息上对身份信息识别得更好。这一结果表明, 表情识别上正常儿童似乎比孤独症儿童更倾向于使用低频信息, 这与上述低频信息与身份识别相关联的推论似乎矛盾, 但若考虑到低频信息与专家经验之间的关联(Goffaux & Rossion, 2006), 这一结果可以理解为孤独症儿童在专家经验形成上的滞后。在身份识别上两组都表现出在低频信息条件下识别得更好, 支持上述双分离推论。总体上, Deruelle 等(2008)实验结果支持身份识别对低频信息的依赖, 但表情识别与高频信息之间的关联未能得到明确的验证。

此外, vuilleumier, Armony, Driver 和 Dolan (2003)采用脑成像技术研究发现, 杏仁核对恐惧面孔图片有很强的激活反应, 对只有低频信息的恐惧面孔也有较强的反应, 但对只有高频信息的恐惧面孔图片却不再有强烈的反应。后来实验者让被试对这些高频恐惧面孔图片进行恐惧性主观评价, 结果显示, 高频恐惧表情面孔的恐惧强度评价比低频的表情面孔的恐惧评价还要高。基于这些结果, 研究者们提出表情识别存在快慢两条神经通路, 一是表情加工早期基于低频面部信息的快速通路, 另一条是后续表情识别加工基于高频面部信息的慢通道。这一结论表明, 表情的早期侦测有赖于低频信息的

快速加工, 但对表情的进一步辨别加工则主要依赖高频空间信息, 这一结果支持双分离推论中关于表情识别与高频空间信息相关的假设。总结 Deruelle 等(2008)与 vuilleumier 等(2003)的研究可以发现, 双分离推论在神经心理学及认知神经科学研究中也能够得到支持, 但低频空间信息的作用表现得比较复杂, 与识别任务、学习及成熟等因素密切相关。

进一步综述更早期研究可见, 尽管系统探讨不同空间频率信息与不同面孔属性识别之间关系的研究主要开始于近期, 但探讨不同空间频率成分在面孔识别中的重要性问题很早就已开始。Ruiz-Soler 和 Beltran (2006)综述早期的研究揭示: 在面孔识别中, 高频空间信息在面孔识别中有重要作用, 中频空间信息被证明是视觉识别的敏感频段, 低频信息与识别任务及识别加工阶段有密切关联。综合早期与近期的这些研究可以发现, 高频空间信息在面孔识别中的关键作用得到较一致的肯定, 但双分离推论中关于高频空间信息有助于表情识别的假设尚需进一步的实验数据支持。对于低频空间信息, 无论是早期还是近期的研究都一致表明其作用复杂, 涉及专家经验与具体的加工阶段, 因此, 作为分离尺度, 它既可能与不同面孔属性加工(如身份或表情)有关, 又可能与不同加工阶段有关。高频空间信息在面孔识别中的稳定表现符合以 Marr 为代表的早期认知科学家的理论假设: 人的视觉识别中存在着一个根据亮度对比来抽取信息的基本机制, 以实现边缘侦测之类的功能(Marr & Hildreth, 1980; Marr, 1982)。相对于低频信息, 高频信息本身就具有放大边缘对比和促进边缘侦测的作用, 因此高频信息在不同属性的面孔信息抽取上更可能成为一个有效的分离尺度。低频空间信息以其粗略加工和资源占用少为特征, 在面孔识别的早期导航或定位以及自上而下的加工中起着重要作用, 因而在杏仁核的早期反应以及与面孔识别初期加工阶段相关的实验任务中体现出相应的实验效应(vuilleumier et al., 2003; Ruiz-Soler & Beltran, 2006)。

综上分析, 双分离推论得到了一些支持, 但这些支持尚不够直接与完整。为了系统地验证双分离推论, 我们提出双分离假设: 高频空间信息促进面部表情识别而不促进面孔身份识别, 低频空间信息促进身份识别而不促进表情识别。

1.2 身份表情信息分离的量化指标

已有研究较多采用构形信息与特征信息解释

身份表情识别,然而,如前所述,构形与特征只能是定性界定,难以量化研究。为了能更为量化地研究空间频率是否是合适的身份表情信息的分离尺度,我们选用近来身份表情识别研究中用到的 Garner 范式。在面孔身份与表情识别的独立与非独立加工研究中, Garner 干扰效应被证明是一个反映身份表情识别独立程度的敏感指标,是从测量身份或表情识别受对方信息干扰的程度上来反映两者的识别独立 (Ganel & Goshen-Gottstein, 2004; Schweinberger & Soukup, 1998; Schweinberger, Burton, & Kelly, 1999), 比如, 身份识别受表情信息干扰的量越小则表明身份识别越独立, 身份信息与表情信息的分离越容易。

Garner 效应测量是一种量化的测量指标, 因此, 它能充分地体现视觉识别中的灵活性原则。当前研究者基本形成共识, 身份与表情识别之间既不是绝对的独立也不是绝对的非独立, 而是体现了视觉识别上的灵活性原则 (Calder & Young, 2005)。因此, 无论是身份还是表情信息, 其分离都是相对的, 是量上的区分。从视觉识别现象上来说明就是: 在日常情境中, 当某一面孔属性信息的识别困难时人的视觉会自动、灵活地利用其它可能的面孔属性信息来帮助识别。在 Garner 范式测量上, 这种利用其它面孔属性信息的程度即是干扰效应。经典的 Garner 效应测量在于考察两个维度变量之间的交互影响。在面孔身份与表情识别的交互影响实验中, 典型的 Garner 任务设计包括一个基线与一个过滤线(或称正交线): 在基线中要求被试只根据一个维度变量(任务相关维度, 如表情)进行快速分类作业, 而另外一个维度变量(非任务相关维度, 如身份)保持恒定(称为基线组段); 在过滤线中要求被试与在基线中一样根据任务相关维度(如表情, 同基线一样)进行快速分类但非任务相关维度(如身份)则是变化的, 分别记录基线与过滤线中的反应时成绩, 以过滤线平均反应时减去基线平均反应时得到的即是非任务相关维度对任务相关维度的干扰效应。

在选择了 Garner 效应测量范式后, 进一步对这一范式进行了分析, 从中分解出两个指标: 可分离度(dissociability)与难度(difficulty)。就面孔身份与表情的识别而言, 可分离度与难度所反映的信息加工内涵不同, 可分离度主要反映不同属性信息之间的可区分程度, 如表情或身份信息彼此之间不受干扰的程度, 而难度则是同一属性信息内亚类区分的难易程度, 如身份不变时两种表情之间的区分。按

照我们的这一界定, 在 Garner 范式中, 基线反应时可以作为有效的难度指标, 测量的主要是属性内亚类区分的难度, 而干扰效应可以作为可分离度指标, 主要测量的是属性信息间的干扰程度。难度与可分离度不是相互独立的, 彼此之间存在着复杂的交互影响关系, 如 Ganel 等(2004)通过提高身份信息的难度(采用长得相似的弟兄面孔)实现了身份信息识别的可分离度的下降, 但这种可分离度的下降并不代表身份信息与表情信息的可分离性下降, 而主要是由于难度提高导致的, 因此, 同时考察难度与可分离度是检验某一身份表情信息分离尺度的关键。删除面孔图片上的高或低以及中间频率信息可能本身会造成相应的身份或表情分类判断难度的增加或减少, 此时单纯地以可分离度的升降来说明身份表情信息之间的可分离性变化就难以成立。可能频率信息的删除本身并没有改变两类信息之间的可分离性, 只是因为难度的改变而导致可分离度的改变。因此, 考察高低空间频率信息是否是身份表情信息的有效分离尺度, 需要同时考察难度与可分离度两个指标。

用可分离度与难度概念取代已有研究(Ganel & Goshen-Gottstein, 2004; Wang & Fu, 2007)中使用的区分度(discriminability)概念原因有二方面: 其一, 已有研究中将基线反应时作为区分度概念的度量, 从而说明 Garner 效应, 而实际上基线反应时能直接影响干扰效应但并不是唯一影响因素, 将它们分为两个指标有助于人们对这两个指标之间关系的理解。其二, 不同表情面孔上的身份表情信息的识别难度不同, 用这些图片作为探讨身份表情信息分离尺度的材料, 以干扰效应为度量的可分离度在不同条件下难以直接进行对比。

按照我们提出的可分离度与难度指标, 上述的双分离假设可以解释为: 如果高频空间信息真的是选择性地促进表情识别而不促进身份识别, 那么, 高频信息下的表情信息的可分离度会高于全频信息下的表情信息的可分离度, 同时两种条件下的难度无明显差异。相应地, 如果低频信息是选择性地促进身份识别而不促进表情识别, 那么, 在低频信息下身份识别的可分离度会高于全频条件下的身份识别可分离度, 同时, 难度无明显差异。

基于以上假设, 我们设计了如下实验: 采用 Garner 范式, 实验 1 使用正常面孔表情图片设计身份与表情识别的 Garner 判断任务, 测量基线区分难度与 Garner 干扰效应(可分离度)。实验 2 仍然使

用实验 1 中的面孔图片,但对这些图片进行低通处理;然后,测量基线区分难度与 Garner 干扰效应(可分离度)。实验 3 也同样使用实验 1 的图片作为材料,但对这些图片进行高通处理;然后,测量基线区分难度与 Garner 干扰效应(可分离度)。比较实验 1 至 3 的结果验证高低空间频率分离尺度假设。

考虑到身份与表情识别的可靠性,实验中空间频率的过滤采用了与已有一些研究相同的做法(Goffaux et al., 2005),高频空间信息面孔图片是通过高通处理获得,低频空间信息面孔图片是通过低通处理获得。因此,本研究中所使用的无论是高频还是低频空间频率面孔图片都实际上不包含中频空间信息。已有一些面孔识别研究表明,人在识别面孔时对中频空间信息较为敏感(Nasanen, 1999; Ojanpaa & Nasanen, 2003),但实际上中间频率更多是与其它频率叠加产生作用的(Liu, Collin, Rainville, & Chaudhuri, 2000)。因此,为了单独考察高、低空间频率信息在身份与表情识别中的作用,材料处理中只保留了高频或低频的空间信息。

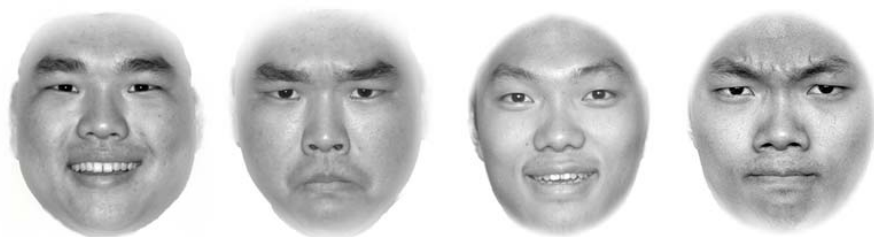


图 1 实验中使用的两位模特的表情图片,每位模特包含愤怒与快乐两种表情示例

2.1.3 实验设计 实验 1 包括表情识别与身份识别两个相对独立的 Garner 分类判断任务,表情与身份分别为 Garner 任务中的二个维度变量,自变量为“非相关”维度(irrelevant dimension),因变量为反应时。考虑 Garner 范式两实验组段中呈现的图片张数不同对反应时的影响,采用已有的做法(Ganel et al., 2004),将各实验组段中呈现的图片张数翻倍,即每个模特的每种表情图片有两张。

2.1.4 实验程序 在表情识别任务条件下,为了使被试熟悉实验中的愤怒与快乐两种表情,实验前给被试呈现两个模特的 8 张表情图片(愤怒与快乐表情图片各 4 张,这也是后面正式实验中用到的 8 张图片),接着进行表情辨别的快速反应测试,反应给予反馈,正确率达到 100%才通过测试,开始

2 实验一 全频面孔图片干扰效应测量

本实验采用新拍摄的中国人脸图材料,采用已有的 Garner 效应研究范式,考察身份信息对表情识别的干扰以及表情信息对身份识别的干扰。

2.1 方法

2.1.1 被试 32 名在校大学本科生,身份与表情任务各 16 名。被试年龄在 18~25 岁之间,男女各半。

2.1.2 实验材料 选取两名男性模特(确认参与研究的被试不认识他们),请他们分别表演 2 种(高兴的、愤怒的)面部表情,拍摄了 10 张数码图片(其中的两张中性表情图片只用于身份识别任务条件下被试预先熟悉模特身份),每个模特的每种表情拍摄两张,表情强度稍有不同;图片用 PhotoShop 软件进行处理,去掉头发、耳朵及面部明显的特征痣,柔化边缘,并将图片去色以减少肤色、光照、瑕疵等的影响。图片的大小为 350×465 像素,实验呈现时为 10°视角,示例见图 1。所选图片的表情类型经 24 名被试进行类型分类评定,这些被试不参加实验,认可百分比均在 80%以上。

表情识别的实验,个别长时间不能通过测试的被试不再进行后面的实验。身份识别任务中的练习程序与表情识别的相同,只是实验前给被试呈现的图片为 10 张,增加了两张中性面孔图片。

图片呈现在 17 吋彩色显示器屏幕中央,图片与屏幕背景均为白色,图片的图像部分与被试眼睛成约 10°视角。实验时要求被试对每次呈现的图片上的表情或身份做二择一判断。实验流程如下:首先出现注视点“+”500 ms,接着是 500 ms 的空屏,然后呈现图片,被试即可按键反应,图片呈现 2500ms 或被试按键消失,空屏 1000 ms,然后呈现下一张图片。

身份识别与表情识别任务均为简化型的 Garner 分类任务,即只包括基线与过滤线(或称作

正交线)两组段, 不包括相关线组段, 这与已有研究做法一致(Ganel & Goshen-Gottstei, 2004)。任务中的基线组与过滤组都包含全部的 8 张图片, 但基线组的 8 张图片根据“非相关”维度分成两部分, 每部分包含 4 张图片, 因此基线组与过滤组的差别在于基线组 8 张图片是根据“非相关”维度分为两部分, 然后分开呈现的, 而过滤组的 8 张图片则是混合在一起呈现。以表情识别为例, 基线组的 8 张图片就会根据“非相关”维度的身份分为模特 A 与模特 B 的表情图片各 4 张, 每个模特的 4 张表情图片各自混合随机呈现, 但在过滤线组, 两个模特的 8 张表情图片混合在一起随机呈现。基线组与过滤线组的每张图片均呈现 18 次, 每个任务共有 288 个试次。正式实验中, 基线组与过滤组图片呈现的前 8 次(非

相关维度“不变”的组段, 4 张图片各呈现 2 次)或 16 次(非相关维度“变”的组段, 8 张图片各呈现 2 次)为练习。

2.2 结果

分别统计比较实验中的表情识别与身份识别任务中的被试平均反应时, 结果见表 1。单因素重复测量方差分析表明, 表情识别受身份干扰显著, $F(1,15)=7.66, p<0.05$, 身份识别受表情干扰也显著, $F(1,15) =7.53, p<0.05$ 。表情识别任务中的 Garner 干扰效应为 20ms, 身份识别中的 Garner 干扰效应 17ms。对错误率进行的相应的方差分析结果表明, 不存在速度—准确性权衡问题($F(1,15) =5.05, p<0.05$; $F(1,15) =6.82, p<0.05$, 过滤线错误率均显著高于 基线)。

表 1 表情分类与身份分类的反应时(括号内为标准误)、错误率及干扰效应

实验及任务	无关维度不变		无关维度变化		干扰效应(ms)
	反应时(ms)	错误率(%)	反应时(ms)	错误率(%)	
表情识别	496(22)	1.4	516(24)	2.3	20
身份识别	496(13)	2.2	513(14)	3.2	17

这一结果表明, 身份表情识别之间出现相互干扰, 两者之间的加工为非独立加工关系, 这与 Ganel 和 Goshen-Gottstei (2004)的实验结果一致。

3 实验二 高频面孔图片干扰效应测量

为进一步考察高频信息在身份与表情识别中的作用, 实验 2 采用与实验 1 同样的程序, 但对实验 1 中的材料进行了空间频率的高通处理, 具体程序处理如下。



图 2 实验中使用的两位模特的高频空间频率表情图片示例

3.2 结果

分别统计比较实验中表情识别与身份识别两任务中被试的平均反应时, 结果见表 2。单因素重复测量方差分析表明, 表情识别受身份干扰不显著, $F(1,15) = 2.84, p > 0.05$, 身份识别受表情干扰显著,

3.1 方法

3.1.1 被试 32 名在校大学本科生, 未参加过前面的实验 1, 表情与身份识别任务各 16 名。被试年龄 在 18~25 岁之间, 男女各半。

3.1.2 程序与材料 程序同实验一, 材料是通过对实验 1 中的材料进行高通过滤处理获得。方法如下: 用 PhotoShop 软件进行空间频率的高通处理, 高斯过滤参数为 $\sigma=38$ 像素, 保证高通在 25c/fw 以上, 参照已有研究的做法(Goffaux et al., 2005; Collins, Therrien, Martin, & Rainville, 2006), 示例见图 2。

$F(1,15) = 12.54, p < 0.05$ 。表情识别受身份干扰的效应为 11ms, 身份识别受表情干扰的效应为 23ms。对错误率进行相应的方差分析表明, 不存在速度—准确性权衡问题($F(1,15) = 0.31, p > 0.05$; $F(1,15) = 0.12, p > 0.05$)。

表 2 表情分类与身份分类的反应时(括号内为标准误)、错误率及干扰效应

实验及任务	非相关维度不变		非相关维度变化		干扰效应 (ms)
	反应时(ms)	错误率(%)	反应时(ms)	错误率(%)	
表情识别	482(17)	2.1	493(19)	2.2	11
身份识别	603(21)	3.0	626(24)	3.2	23

与实验 1 相比, 干扰效应从 20 ms 下降至 11 ms, 表情识别受身份干扰的 Garner 效应下降明显, $t = 1.79$, $p < 0.1$; 两实验中基线与过滤线错误率差值之间无显著差异, $t = 1.58$, $p > 0.05$ 。从实验内看, 实验 1 中的表情识别显著受身份干扰, 实验 2 中的表情识别则未表现出受身份干扰。这些结果表明, 高频条件下表情信息的可分离度明显提高。在基线水平上, 实验 1 的基线水平为 496 ms, 实验 2 为 482 ms, 两者相差 14ms, 但这种差异不显著($t = 0.52$, $p > 0.05$), 并且实验 1 基线错误率高于实验 2 基线错误率, $t = 1.77$, $p < 0.1$ 。这些结果表明, 表情识别上的难度指标在两个实验之间无明显差异。综合可分离度与难度指标表明, 在身份表情信息中, 高频率空间信息是分离表情信息的有效尺度。

与实验 1 相比, 身份识别受表情干扰的程度无明显变化($t = 0.73$, $p > 0.05$), 基线与过滤线错误率差值无显著变化($t_{\text{错误率}} = 0.89$, $p > 0.05$)。从实验内看, 无论是在实验 1 还是实验 2 中, 表情都显著干扰身份识别。这些结果表明, 身份信息可分离度指标无明显变化。在基线水平上, 实验 2 中的身份识别在高频条件下变得更加困难, 反应时从 496ms 上升至 603ms, 差异显著($t = 4.31$, $p < 0.01$), 这表明身份识别的难度指标上升。难度指标的结果与 Ganel 和 Goshen-Gottstein (2004) 的研究发现一致。综合可分离度与难度指标表明, 在身份识别中, 高

频信息增加了身份识别的难度但并没有提高身份信息可分离度, 高频信息对身份而言并不是一个有效的分离尺度。

综合表情识别与身份识别结果表明: 1) 单纯高频信息提高表情信息的可分离度, 但不影响表情识别的难度; 2) 单纯高频信息不影响身份信息可分离度, 但却提高身份识别的难度。高频信息促进表情识别而不影响身份识别, 表现出了分离性影响, 符合双分离假设预期。

4 实验三 低频面孔图片干扰效应测量

为进一步探讨低频信息的影响, 实验 3 采用与实验 1 同样的程序, 但对实验 1 中的材料进行了低通过滤处理, 具体程序处理如下。

4.1 方法

4.1.1 被试 32 名在校大学本科生, 未参加过前面的实验 1 与 2, 表情与身份识别任务各 16 名。被试年龄在 18~25 岁之间, 男女各半。

4.1.2 程序与材料 程序同实验一, 材料是通过实验 1 中的材料进行低通过滤处理获得。方法如下: 用 PhotoShop 软件进行高斯低通过滤处理, 低通过滤参数为 $\sigma = 10$ 像素, 保证低通在 $8c/fw$ 以下, 处理参数的选择同时参照已有研究(Goffaux et al., 2005; Collins et al., 2006), 示例见图 3。



图 3 实验中使用的两位模特的低频空间频率表情图片示例

4.2 结果

分别统计比较实验中表情识别与身份识别两任务中被试的平均反应时, 结果见表 3。单因素重复测量方差分析表明, 表情识别受身份干扰效应边缘显著, $F(1,15) = 3.43$, $p < 0.1$, 身份识别受表情干扰显著,

$F(1,15) = 7.15$, $p < 0.05$ 。表情识别受身份干扰的效应为 25ms, 身份识别受表情干扰的效应为 22ms。对错误率进行相应的方差分析表明不存在速度-准确性权衡问题($F(1,15) = 0.06$, $p > 0.05$; $F(1,15) = 4.67$, $p < 0.05$, 身份识别上过滤线错误率显著高于基线)。

表 3 表情分类与身份分类的反应时(括号内为标准误)、错误率及干扰效应

实验及任务	无关维度不变		无关维度变化		干扰效应 (ms)
	反应时(ms)	错误率(%)	反应时(ms)	错误率(%)	
表情识别	535(29)	2.9	559(31)	3.0	25
身份识别	544(24)	3.5	566(23)	4.5	22

与实验 1 相比, 表情识别受身份干扰的效应无明显变化($t = 0.31, p > 0.05$)。尽管干扰效应值从 20ms 上升至 25ms, 但基线错误率显著增加($t_{\text{错误率}} = 2.86, p < 0.05$)。这表明, 表情信息的可分离度没有变化。在基线水平上, 实验 3 中的表情识别与实验 1 中的表情识别之间没有显著差异($t = 1.06, p > 0.05$), 这表明表情识别上的难度没有显著变化。综合可分离度与难度指标表明, 单纯低频信息没有明显增加表情的可分离度, 也没有明显增加表情识别的难度。

与实验 1 相比, 身份识别受表情干扰的效应没有变化($t = 0.53, p > 0.05$)。尽管干扰效应从 17ms 上升到 22ms, 但错误率显著上升($t_{\text{错误率}} = 2.10, p < 0.05$), 这表明身份信息的可分离度并没有变化。在基线水平上, 实验 3 明显高于实验 1 ($t = 1.73, p < 0.1$), 这表明低频条件下身份识别的难度增加。综合可分离度与难度指标表明, 单纯低频信息增加了身份识别的难度, 但并没有提高身份信息的可分离度。

综合表情与身份识别结果表明: 1) 单纯低频信息未影响表情识别的可分离度与难度; 2) 单纯低频信息增加了身份识别的难度但不影响身份信息的可分离度。结果不符合双分离假设预期。

5 讨论

综合实验 1 至 3 的结果表明, 以难度与可分离度为观察指标, 身份表情识别在高频空间信息条件下出现了分离现象而在低频空间信息条件下未表现出分离现象, 部分验证了前面提出的双分离假设, 即高频空间信息促进表情识别而不促进身份识别。低频空间信息上没有表现出分离现象, 这同引言中对已有研究的综述结果相近, 低频信息的作用复杂, 涉及专家经验与具体的加工阶段, 作为分离尺度它既可能与不同面孔属性加工(如身份或表情)有关又可能与不同加工阶段有关。此外, 低频信息上未出现分离现象的另外一个可能原因是实验因素。我们在实验过程当中以及事后对被试的访谈表明, 在呈现低频面孔图片时, 被试会不自觉地产生回避行为, 他们认为, 这种低频面孔图片反复观看会觉得很

晕。可分离度指标的总体结果见图 4 所示。

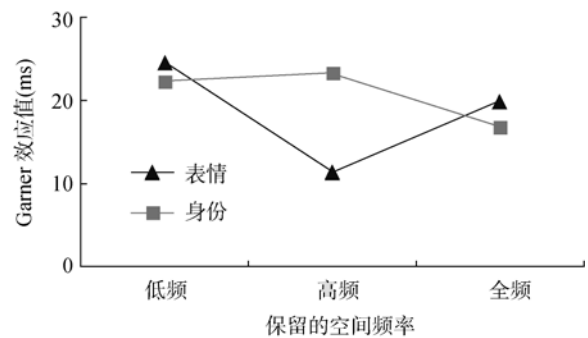


图 4 不同空间频率信息对 Garner 干扰效应的影响

5.1 高水平表情识别主要依赖边缘特征信息

从实验 2 探讨的高频信息结果来看, 高频空间信息是表情信息抽取的关键尺度。按照 Marr 等(1980, 1982)的边缘侦测观来解释就是高频空间信息有利于边缘信息的突显, 从而使得表情的特征提取容易。这一解释也同时告诉我们, 表情识别任务(至少是本实验中的两种表情分类判断)主要依赖于边缘特征的检测, 这与 White 等(2002)运用构形/特征改变研究的结果相一致。相反, 如果身份特征(尤指非特异标记性的身份特征)识别主要依赖于非边缘特征的构形信息, 那么高频信息可能会影响构形信息的抽取从而影响身份识别, 这与上述实验 2 的结果一致。

然而, 我们同样需要注意到认知神经科学以及神经心理学中的一些研究结论(vuilleumier et al., 2003; Deruelle et al., 2008), 即表情识别似乎对低频信息很敏感。在表情信息的视觉加工中, 越来越多的研究表明早期快速侦测与高水平的表情识别并非同一加工层次, 快速侦测是粗略加工, 主要处理低频空间信息, 而表情的类型与意义依赖于后期的精细加工(LeDoux, 2000; Hughes, Nozawa, & Kitterle, 1996)。认知神经科学研究中观察到的早期反应主要体现的是表情信息的早期侦测, 表现为无意识地自动化加工, 快速获取并处理粗略水平的表情信息, 完成简单水平上的表情侦测。复杂任务要求下的辨别区分则主要体现的是后期的精细加工,

主要依赖于高频空间信息提供的边缘信息来“确认或辨别”表情的类型或意义,属于高水平的表情识别,我们上述实验中的 Garner 分类判断任务可归于高水平的表情识别。两种加工水平的观点既能解释现有研究的结果又能体现视知觉加工的灵活性原则。

5.2 身份信息无特异空间频率特征

就实验 3 中探讨的低频信息而言,身份表情的快速分类判断在低频空间信息上未表现出分离。根据已有研究结论(Ruiz-Soler et al., 2006),在面孔识别中低频信息主要与初期加工阶段关联,其作用主要是对面孔信息(身份的或其它面孔属性的信息)进行粗略加工,以便为后续加工进行导航,提供精细加工的路径,这被认为是人的视觉识别的早期共同特征(Pylyshyn, 2001)。我们实验中的快速分类判断任务的完成主要涉及两种表情或两种身份的辨别,依赖于后期的精细加工,因而低频信息在两种任务上都只是表现为提高了识别的难度而没有提高身份信息的可分离度。

White (2002)采用构形改变与特征改变技术,研究提出身份识别主要与构形信息有关, Goffaux 等(2005)则采用空间频率上构形改变与特征改变技术,研究提出低频空间信息与构形加工关联而高频空间信息与特征加工关联。综合这两项研究能够简单地得到推论:低频空间信息主要与身份识别相关联。进而,低频空间信息与身份加工之间的关联同样得到 Deruelle 等(2008)有关孤独症儿童研究的支持。有趣的是,上述三项研究都采用的面孔匹配实验任务,这与本实验中的面孔信息区分任务有区别。按照大多数研究者的共识(Ruiz-Soler et al., 2006),低频信息的利用主要在早期加工阶段,那么采用面孔匹配任务可能主要涉及到的就是早期加工阶段的低频信息的利用。这一点我们可以直接从 Deruelle 等(2008)的研究结果中进行求证:孤独症儿童在识别表情时更倾向于利用高频信息而正常儿童则倾向于利用低频信息。考虑到匹配任务与早期加工阶段关联,而早期加工阶段又主要是低频信息的快速粗略加工,我们就能容易理解为什么正常儿童倾向于低频信息来匹配表情。可见,面孔识别加工的具体信息与实验任务所要求的加工深度有关(Costen, Parker, & Craw, 1996)。

综合加工阶段与实验任务差异这两方面因素,前人研究、本研究发现以及双分离推论之间的不一致就能够得到解释:低频信息主要在早期加工阶段

起作用,匹配任务更能体现这一加工阶段,因而采用匹配任务既能发现低频空间信息对身份识别的影响(Goffaux et al., 2005; Goffaux & Rossion, 2006),也能发现低频空间信息对表情识别的影响(Deruelle et al., 2008; vuilleumier et al., 2003);高频空间信息主要在精细加工要求的边缘信息侦测上起重要作用,而高水平的表情识别正是依赖于边缘信息的检测,因此,在分类判断任务中能揭示高频空间信息提高表情信息的可分离度(如本实验 2 的研究),而不能揭示高频信息提高身份信息的可分离度。

5.3 身份表情信息分离的多维尺度展示

面孔身份与表情信息的分离是视知觉的复杂功能,信息分离的尺度是多维的。人们很早就揭示低频空间信息对早期加工阶段的重要性(Navon, 1977; Sanocki, 1993),人们也很早发现高频空间信息对表情识别的重要性(Fiorentini, Maffei, & Sandini, 1983),但实际上,这些发现都有一个重要前提——加工阶段。一般情况下,这一阶段遵从由整体到局部,从粗略到精细这一顺序,然而近来的研究发现,这种阶段之间的顺序也似乎并不是固定不变的(Morrison & Schyns, 2000; Smith, Gosselin, & Schyns, 2004)。看来加工阶段与信息类型两者之间存在着复杂的交互关系,期望简单地用一种维度指标来分离抽取不同类的信息并不符合视知觉加工事实。与面孔表情相比,面孔身份这类信息似乎不具有简单维度分布上的特征,因此,研究者通常认为身份信息为个体化信息而表情信息是类型化信息。从信息表征的维度化来描述,个体化信息与类型化信息之间差异主要在于维数上的差异,个体化的身份信息需要很多维度才能有效地表征这些信息,而类型化的表情信息只需要有限的几个维度就能表征。这种维度数上的差异也决定身份信息与表情信息之间的区分可能是单向的,即类型化的表情可能稳定一致地在某一尺度上区别于身份信息,但个体化的身份信息就难于在单一的尺度上区分于表情信息。

目前有关面孔身份与表情信息的分离尺度最主要的有两个,一是主观性的特征与构形分类,另一个就是客观性的空间频率分类。构形与特征分类得到不少研究的支持,但构形与特征的界定困难,主观经验性强,难以量化操作。更为重要的是很难区分什么样的信息是构形信息,什么样的信息是特征信息。已有多数研究者认为,构形信息大概是面孔的几种构成部分,如眼睛、鼻子、嘴巴、眉毛之

间的空间比例关系, 以及面孔整体的形状特征(White, 2001, 2002; Goffaux et al., 2005)。这种界定还是非常模糊, 有些信息似乎像是人们感觉到的构形信息, 比如, 颧骨的弧度, 但按现有的构形信息界定很难说明它。同样地, 像颧骨的弧度特征也可以归到特征信息当中, 因为它反映的是一种局部特征。可见, 构形信息与特征信息之间的分类不是客观明确的。

空间频率是研究者们一直寄予厚望的一个客观性分离尺度, 研究者们也相信这一尺度更符合人们对眼睛生理机能的基本猜想: 有不同的细胞对不同的频率敏感, 这样就自然地进行了信息的分类, 有点把人的眼睛构造与某些低级生物的眼睛构造相类比的倾向。这种比较有助于人们对视觉的理解, 然而事实却并非如此。低等生物的视觉识别对象可能远简单于人类, 科学研究认为某些鸟类单凭分辨同伴的瞳孔就能够识别出它们, 这在人类简直是不可能。不同空间频率并没有对应于相应的神经通道(Hughes et al., 1996), 不仅是高频或低频信息还有那些中频信息都在识别中扮演着重要作用(Nasanen, 1999; Ojanpaa et al., 2003), 甚至是不同空间频率之间的重叠而不是分离也起重要作用(Liu et al., 2000)。

实际上, 人的面孔信息异常复杂多样, 信息量与信息类型都是难以计数的, 因此凭借单一维度很难进行有效的区分。空间频率可能是最理想的分离尺度, 但这些尺度因为太过容易而变得难以胜任复杂化的要求。当然, 越来越多的研究者也开始相信, 人的视知觉的识别能力可能并没有我们想象的那样好, 一个庞大的知识经验库恐怕才是这一识别系统最有力的支持, 这也是当前面孔识别知觉专家经验说(the perceptual expertise hypothesis)被大多数学者认可的关键原因(Gauthier, Skudlarski, Gore, & Anderson, 2000)。

与信息分离尺度多维相对应, 面孔识别研究中的实验范式也呈现多样性。正如研究者所揭示, 面孔信息的独立识别本身受实验任务影响, 任务上的差异导致选择的尺度不同, 同一尺度在不同任务上的表现也不同(Schyns & Oliva, 1999; Morrison et al., 2000; Smith et al., 2004)。现有实验范式多种多样, 各实验范式所涉及到的信息维度也有很大的不同。用得最普遍的匹配范式对识别早期加工阶段较为敏感(White, 2002), 合成脸范式(compound face paradigm)则对面孔整体构形信息维度敏感, 重复

启动范式主要对面孔身份信息敏感(Ellis, Young, & Flude, 1990), 微观分析范式(microgenetic approach paradigm)则主要用于探测可定量操作维度加工的时序过程, 在考察不同空间频率加工先后顺序上应用较多(Hughes et al., 1996), 叠加脸范式(hybrid face paradigm)探讨了多维选择上的灵活性问题(Schyns & Oliva, 1999), 本研究使用的 Garner 范式主要用于专门考察两维度之间的交互关系。因为面孔信息, 尤其是身份信息, 大多是多维的, 采用不同范式考察到的维度不同, 其结果之间也就不具有可比性, 如 Vuilleumier 等(2003)实验中表情识别的脑成像结果与被试的主观评定结果就完全相反。因此, 采用单一维度的考察方法考察这些面孔信息的分离与抽取显然不够全面。本研究选用的 Garner 范式虽然专门设计于考察两维度信息之间的交互关系, 但相对于多维的面孔身份或表情信息而言仍然相形见绌。未来能否发展出新的能够同时考察多个维度信息的实验范式是探讨面孔信息分离识别的关键。

本研究的创新之处在于首次从身份表情两维度交互干扰的角度考察了两类信息的分离尺度问题: 在难度一定的情况下, 高频空间信息选择性地降低了表情识别受身份信息干扰的程度, 揭示高频信息是表情信息区分于身份信息的有效尺度。基于 Garner 范式, 研究中还提出可分离度与难度两项重要指标, 使得利用 Garner 范式考察身份表情分离尺度的方法更加成熟。未来研究的目标应该集中在寻找能同时考察多个维度上的实验范式, 以期更好地探讨维度信息分离的问题。

6 结论

通过以上实验及分析, 本研究得出如下结论:

- 1) 高频空间频率是分离面孔身份与表情信息的一个有效尺度;
- 2) 低频空间信息在面孔身份与表情的辨别任务中既未表现出有利于身份, 也未表现出有利于表情识别, 其对身份表情信息的分离性影响可能与实验任务及加工阶段复杂关联。

参 考 文 献

- Bruce, V., & Young, A. W. (1986). Understanding face recognition. *British Journal of Psychology*, 77, 305-327.
- Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews*, 6, 641-651.
- Collin, C. A., Therrien, M. E., Martin, C., & Rainville, S. J. M. (2006). Spatial frequency thresholds for face recognition

- when comparison faces are filtered and unfiltered. *Perception & Psychophysics*, 68, 879–889.
- Costen, N. P., Parker, D. M., & Craw, I. (1996). Effects of high-pass and low-pass spatial filtering on face identification. *Perception & Psychophysics*, 58, 602–612.
- Deruelle, C., Rondan, C., Salle-Collemiche, X., Rosset, D.B., & Da Fonseca, D. (2008). Attention to low and high spatial frequencies in categorizing facial identities, emotions and gender in children with autism. *Brain and Cognition*, 66, 115–123.
- Ellis, A. W., Young, A. W., & Flude, B. M. (1990). Repetition priming and face processing: Priming occurs within the system that responds to the identity of a face. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 42, 495–512.
- Fiorentini, A., Maffei, L., & Sandini, G. (1983). The role of high spatial frequencies in face perception. *Perception*, 12, 195–201.
- Ganel, T., & Goshen-Gottstein, Y. (2004). Effects of familiarity on the perceptual integrality of the identity and expression of faces: The Parallel-Route Hypothesis revisited. *Journal of Experimental Psychology: Human Perception and Performance*, 30, 583–597.
- Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, 3, 191–197.
- Goffaux, V., Hault, B., Michel, C., Vuong, Q., & Rossion, B. (2005). The respective role of low and high spatial frequencies in supporting configural and featural processing of faces. *Perception*, 34, 77–86.
- Goffaux, V., & Rossion, B. (2006). Faces Are “Spatial”—Holistic face perception is supported by low spatial frequencies. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 1023–1039.
- Hughes, H. C., Nozawa, G., & Kitterle, F. (1996). Global precedence, spatial frequency channels, and the statics of natural images. *Journal of Cognitive Neuroscience*, 8, 197–230.
- LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience*, 23, 155–184.
- Liu, C. H., Collin, C., Rainville, S., & Chaudhuri, A. (2000). The effects of spatial frequency overlap on face recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 956–979.
- Marr, D., & Hildreth, E. C. (1980). Theory of edge detection. *Proceedings of the Royal Society of London B*, 207, 187–217.
- Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. Freeman, San Francisco.
- Morrison, D. J., & Schyns, P. G. (2001). Usage of spatial scales for the categorization of faces, objects, and scenes. *Psychonomic Bulletin & Review*, 8, 454–469.
- Nasanen, R. (1999). Spatial frequency bandwidth used in the recognition of facial images. *Vision Research*, 39, 3824–3833.
- Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, 9, 353–383.
- Ojanpaa, H., & Nasanen, R. (2003). Utilisation of spatial frequency information in face search. *Vision Research*, 43, 2505–2515.
- Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition*, 80, 127–158.
- Ruiz-Soler, M., & Beltran, F. S. (2006). Face perception: An integrative review of the role of spatial frequencies. *Psychological Research*, 70, 273–292.
- Sanocki, T. (1993). Time course of object identification: Evidence for a global-to-local contingency. *Journal of Experimental Psychology: Human Perception and Performance*, 19, 878–898.
- Schweinberger, S. R., & Soukup, G. R. (1998). Asymmetric relationships among perceptions of facial identity, emotion, and facial speech. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 1748–1765.
- Schweinberger, S. R., Burton, A. M., & Kelly, S. W. (1999). Asymmetric dependencies in perceiving identity and emotion: experiments with morphed faces. *Perception and Psychophysics*, 61, 1102–1115.
- Schyns, P. G., & Oliva, A. (1999). Dr. Angry and Mr. Smile: when categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition*, 69, 243–265.
- Smith, M. L., Gosselin, F., & Schyns, P. G. (2004). Receptive fields for flexible face categorizations. *Psychological Science*, 15, 753–761.
- Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. (2003). Distinct spatial frequency sensitive for processing facial and emotion expression. *Nature Neuroscience*, 6, 624–631.
- White, M. (2001). Effect of photographic negation on matching the expressions and identities of faces. *Perception*, 30, 969–981.
- White, M. (2002). Different spatial-relational information is used to recognize faces and emotional expressions. *Perception*, 31, 675–682.
- Wang, Y. M., & Fu, X. L. (2007). Effect of discriminability on interference between facial expression and facial identity recognition. *Acta Psychologica Sinica*, 39, 191–200.
- [汪亚珉, 傅小兰. (2007). 区分度在面部表情与面孔身份识别交互中的作用. *心理学报*, 39, 191–200.]

Effects of Spatial Frequencies on Recognition of Facial Identity and Facial Expression

WANG Ya-Min; WANG Zhi-Xian; HUANG Ya-Mei; JIANG Jing; DING Jin-Hong

(Beijing Key Laboratory of Learning and Cognition and Department of Psychology, Capital Normal University, Beijing 100048, China)

Abstract

By changing configural or featural/category information, White (2002) revealed that configural changes mainly interfered with facial identity processing while featural alterations largely reduced facial expression processing. With this technique, Goffaux, Hault, Michel, Vuongo, and Rossion (2005) presented that low spatial frequency played in configural changes detection, whereas featural changes detection depended on high spatial frequency. Based on these two studies, we can draw a conclusion that low spatial frequency plays an important role in facial identity recognition while high spatial frequency plays in facial expression recognition. Can this conclusion be really supported by experiments?

To test this hypothesis, we conducted three Garner experiments in current study. In terms of the hypothesis, high spatial frequency enhances facial expression recognition but not facial identity recognition, while low spatial frequency facilitates facial identity recognition but not facial expression recognition. Dissociation could be found in recognition of facial identity and facial expression.

Three Garner experiments were performed on 96 participants, with each 32 participants. In experiment 1, we measured Garner effect between facial identity and expression with full spatial face photos. In experiments 2, we complete the measurement on low spatial frequency filtered photos, and in experiment 3 the same measurement were made on high spatial frequency filtered face photos. 32 participants were recruited for each experiment.

Results of experiment 1 indicated that obvious Garner effects were observed in either identity or expression recognition. In experiment 2, high spatial frequency enhanced facial expression recognition by lowering Garner effect, but it had no influence on facial identity recognition. In experiment 3, low spatial frequency exerted no obvious influence on either identity or expression recognition. Analysis of dissociability (measurement of Garner effect) and difficulty (measurement of selective reaction time in base line) showed that high spatial frequency selectively affected expression detection.

In conclusion, high spatial frequencies were possibly used by us to distinguish facial expression information from identity information in hard-detected recognition task. Low spatial frequencies were easy to interact with both processing level and experimental task.

Key words facial recognition; spatial frequency; Garner effect; dissociability; difficulty