

## 场景的一致性效应及其机制\*

龚明亮 禩宇明\*\* 傅小兰

(中国科学院心理研究所, 脑与认知科学国家重点实验室, 北京 100101)

**摘要** 有关场景一致性效应的研究发现, 人们对与背景语义一致的前景物体的命名、分类、搜索和再认等都快于与背景不一致的物体. 和与情境一致的物体相比, 与情境不一致的物体在中央顶区等部位, 会诱发一个幅度更大的负波(N390). 旁海马皮层/旁海马位置区(PHC/PPA)和压部后皮层(RSC)是负责场景加工的重要脑区. 场景一致性效应的时程可能首先由低空间分辨率(LSF)信息激活眶额皮层(OFC)(130 ms左右)、PHC/PPA和RSC, 之后LSF信息与高空间分辨率(HSF)信息在颞叶进行整合. 在诸多理论模型中, 情境促进模型从生理角度对一致性效应作了较充分的解释.

**关键词** 一致性效应, 场景, 情境, 语义关系, 理论模型, 生理机制

**学科分类号** B845

**DOI:** 10.3724/SP.J.1206.2011.00032

在现实世界中, 我们看到的物体并不是孤立存在的, 它们总是出现在一定的场景之中. 因此, 许多对物体知觉的研究从以单个物体为对象扩展到以场景为对象. 场景包括情境(context)和前景物体(objects)两个重要的组成部分, 其中, 情境又可称为背景(background). 场景图式(scene schema)的存在表明, 随着经验的积累, 我们会逐渐地积累关于某一类物体出现的地点或场合的知识结构, 形成先验知识(prior knowledge). 比如, 我们一般只在海上见到轮船, 而微波炉则出现在厨房. 这就是说, 物体通常固定出现在特定的场景中, 甚至是场景的特定空间位置上. 物体出现在其通常出现的位置有利于对它的再认<sup>[1]</sup>, 而出现在一般不出现的情境中则会令人感到惊奇<sup>[2]</sup>.

我们识别正常场景中物体的绩效明显好于识别打乱场景中的物体, 表明情境会影响对物体的再认<sup>[3]</sup>. 其他研究还发现, 情境会影响人们对物体的命名、分类和回忆等<sup>[2, 4-6]</sup>. 大量研究都发现, 被试对与背景语义一致的前景物体的命名、分类、搜索和再认等都快于与背景不一致的物体, 这一现象称为“场景一致性效应(scene consistency effect)”.

本文主要从行为学、电生理学和脑成像等方面介绍场景一致性效应的研究证据, 并总结了解释这一效应的主要理论模型(尤其是Bar的情境促进模

型), 最后对当前研究现状进行小结并指出存在的不足.

### 1 场景一致性效应的行为学证据

研究者很早就发现, 物体所处的情境会影响对物体的加工. Biederman等<sup>[7]</sup>采用视觉搜索任务, 较早对场景和物体的关系进行了研究. 他们先给被试呈现一个物体, 再呈现场景, 要求判断之前呈现的物体是否在该场景中. 物体与场景的关系有三种: 物体一般只出现在该场景中; 物体可能出现在该场景中; 物体不出现在该场景中. 结果发现, 被试在第二种情况下反应最慢, 这表明物体与场景的语义关系会影响搜索的效率. 后来Palmer<sup>[8]</sup>在物体辨别任务中发现, 物体紧随与之语义一致的场景呈现时的辨别准确率明显高于紧随与之语义不一致场景呈现时的准确率. 更近的一个研究<sup>[9]</sup>采用了类似的搜索范式, 并以眼动情况为指标, 要求被试搜索直升机、吉普或是轻型飞艇, 并在被试执行任务的时候对眼动轨迹进行了记录, 发现被试搜索直升机

\* 国家重点基础研究发展计划(973)(2011CB302201)和国家自然科学基金(90820305, 61075042)资助项目.

\*\* 通讯联系人.

Tel: 010-64837040, E-mail: xuanyim@psych.ac.cn

收稿日期: 2011-01-18, 接受日期: 2011-03-28

所需的时间比搜索吉普和轻型飞艇更长, 并且搜索轻型飞艇时被试对天空有更多的注视, 停留时间(dwell time)也 longer. 相反, 搜索吉普时被试则对陆地有更多的注视, 停留时间 longer. 这说明先验知识会影响空间搜索: 轻型飞艇一般只出现在空中, 故要求搜索轻型飞艇时, 被试更多地注视空中; 而直升机可出现在陆地上, 也可出现在空中, 因而搜索直升机时需要搜索的空间范围更大, 从而使反应时间增加. 由此来看, 场景可以作为一种情境线索(contextual cueing), 自上而下地指导视觉搜索, 使物体与情境语义不一致时所需要的搜索时间 longer.

以前对一致性效应的定义强调的是情境对物体的命名、搜索和探测的影响<sup>[8]</sup>. 后来, Davenport 和 Potter<sup>[10]</sup>发现物体对情境存在同样的影响, 从而将一致性效应的范围进一步扩大. 在她们的实验中, 先呈现图片 80 ms, 接着出现一个掩蔽刺激, 要求对场景中的物体或背景进行命名任务. 实验结果发现, 前景物体和背景一致时对物体或背景的命名准确率都高于不一致的时候, 说明一致性效应并不只是场景对物体加工的单方面效应, 而是一种相互作用, 情境影响物体的加工, 反过来, 物体也影响情境的加工. Davenport<sup>[11]</sup>后来的一项研究还发现, 甚至物体间语义一致时也会加快对物体的加工, 表明在物体之间也存在一致性效应. 由此, 她们提出了视觉场景加工的相互作用模型(interactive model): 前景物体和背景在加工过程中相互影响, 背景为物体提供情境, 同时物体也为背景提供情境. Joubert 等<sup>[12]</sup>采用 go/no-go 分类任务范式, 发现不一致情境下正确率下降了 10%, 且反应时增加了 16 ms, 并且更换其他被试和材料重复实验后, 一致性效应仍然存在, 这在一定程度上表明了一致性效应的鲁棒性.

## 2 场景一致性效应的电生理学证据

N400 是指示语义违反的一个电生理指标, 它最初由 Kutas 和 Hillyard 于 1980 年发现<sup>[13]</sup>. 当单词逐个地在屏幕上呈现句子时, 最后一个是不合语境的单词会比符合语境的单词在 250 ms 开始有一个更大的波幅, 并在 400 ms 左右达到最大, 因而 Kutas 和 Hillyard 将这个成分命名为 N400. 后来 Ganis 和 Kutas<sup>[14]</sup>发现, 图片也会诱发与 N400 类似的成分. 其前提是图片中的物体出现情境违反(context violation), 即物体出现在不寻常的场景中. 由于波峰出现的时间在 390 ms 左右, 他们将

这个成分命名为 N390. 这个发现提供了第一个关于一致性效应的事件相关电位(event-related brain potentials, ERPs)证据, 表明一致性效应是物体与情境间语义违反所导致的结果. 他们的实验流程为: 先在屏幕的某个位置上呈现一个注视点, 接着出现一个情境图片 300 ms, 之后在注视点的位置出现一个物体 300 ms(可能与情境一致, 也可能不一致), 然后物体和情境一起消失, 要求被试尽快对物体进行辨别. 结果发现, 除了在行为水平上发现典型的一致性效应外, 在有些脑区, 尤其是中央顶区, 与情境一致的物体比与情境不一致的物体有一个更大的负走向波(N390), 其产生源(颞叶)、时间进程和头皮分布与词语研究中发现的 N400 极其相似, 表明两者具有相似的内在加工机制. 而之所以波峰一个在 390 ms 左右, 另一个在 400 ms 左右, 可能是由于图片材料比文字材料通达语义更快<sup>[15]</sup>, 因而图片的语义整合更早, 在脑电上表现为更早出现波形的差异. 因此, Ganis 和 Kutas 认为 N390 与 N400 一样, 反映了场景在语义分析水平上对物体加工的影响. 但先呈现情境再呈现物体条件下, 被试会先根据情境对可能的物体产生预期并由此产生预备反应, 因此, 产生的 N390 有可能是情境违反导致的, 但也可能是对物体的预期与实际物体的失匹配或预备反应与实际的正确反应之间的失匹配所引发的. 为了排除预期和反应失匹配的混淆, Mudrik 等<sup>[16]</sup>将场景与物体同时呈现, 结果依然有明显的类 N400 效应, 表明其确实与情境违反有关.

Sitnikova 等<sup>[17]</sup>第一次使用真实场景的视频为材料, 也发现了类似 N400 的成分. 他后来的一个研究<sup>[18]</sup>为此提供了进一步的证据. 在这个研究中, Sitnikova 等给被试观看一段视频, 其中有些视频的结尾符合我们的生活经验, 而有些则出乎意料(如首先观看到影片中的人取砧板, 然后把面包放在砧板上, 但最后出现的却是“用熨斗熨面包”). 脑电记录发现这些视频引发了类似 N400 的成分, 进一步表明了真实场景的语义一致性效应的存在. 由于观看视频片断所诱发的知觉经验与对现实世界中事件的知觉非常相似<sup>[19]</sup>, 因而这些研究结果具有更高的生态效度.

对时间进程的研究有助于我们理清场景一致性效应的产生过程和各相关脑区的激活顺序. 而目前对于场景加工的时间进程的研究还非常少, 但从已有的相关研究中, 我们可以得到一些启示.

Thorpe 等<sup>[20]</sup>首先将 ERPs 技术应用到场景分类

任务中去。他采用 go/no-go 的范式, 给被试呈现 20 ms 图片材料, 要求判断场景图片中是否有动物, 并且在实验的过程中用 ERPs 技术记录被试的脑电情况。结果发现, 额区 no-go 试次在刺激出现后 150 ms 时开始出现比 go 试次更大的负波, 表明对场景中物体的探测只需要 150 ms。考虑到探测 (detection) 和辨别 (identification) 区别——前者可以平行进行, 加工较浅; 后者需要特征的捆绑 (binding), 故需要注意资源的参与, 加工也更深<sup>[21-23]</sup>——对于分类和命名等任务的前提是对物体进行辨别, 因而可能需要更多的加工时间。事实也确实如此<sup>[14, 16-18]</sup>。大多数研究通过给被试呈现情境违反的图片或预料之外的视频, 发现大约从 250 ms 开始不一致场景与一致场景的波形出现差异 (前者有一个更大的负波), 并在 390 ms 时达到最大, 这就是 N390。它属于 ERP 的晚成分, 反映了知觉加工中的语义整合, 并可能与前颞叶的激活有关<sup>[14]</sup>。Bar 等<sup>[24]</sup>则结合脑磁 (MEG) 和功能性磁共振成像 (fMRI) 技术, 发现在位于前额叶的眶额皮层 (orbitofrontal cortex, OFC) 处, 再认成功的物体与再认失败的物体在刺激出现后的 130 ms 出现差异, 这比颞叶处的激活要早 50 ms。此外, 完整的正常图片与低空间分辨率 (LSF) 图片引起的 OFC 激活相似, 而与高空间分辨率 (high spatial frequency, HSF) 图片差异显著。这表明 OFC 先于颞叶得到激活, 而且是由图片的 LSF 信息导致的。Peyrin 等<sup>[25]</sup>直接以场景图片为材料, 也得到了相似的结果。

综合上述研究结果, 我们认为, 场景加工的时间进程大致为: 首先, 场景图片信息进入初级视皮层, 然后, 在 130 ms 左右 LSF 信息到达 OFC。与此同时, 场景图片信息通过小细胞通路也在向颞叶传递, 180 ms 左右开始激活颞叶, 之后不同通道的信息到达颞叶并进行整合 (包括语义整合), 使得场景一致和场景不一致的信息于大约 250 ms 开始出现差异, 并在 390 ms 达到最大, 表现为场景的一致性效应。

### 3 场景一致性效应的脑机制

随着技术的发展, 研究者也尝试采用 fMRI 等手段来探究物体与所处情境的关系对场景加工的影响, 这有助于我们更深入地了解场景一致性效应产生的生理基础。脑成像研究显示, 加工情境相关的信息时, 在旁海马皮层 (parahippocampal cortex, PHC), 尤其是 PHC 中的旁海马位置区 (parahippocampal

place area, PPA) 和位于后扣带皮层的压部后皮层 (retrosplenial cortex, RSC) 存在情境特异性的激活, 表明它们在情境加工中起着重要作用<sup>[6, 26-29]</sup>。

Epstein 等<sup>[30]</sup>最早定义了 PPA 这个脑区, 并发现 PPA 对没有物体的场景 (如没有任何东西的空房子) 有反应, 却对呈现在白背景上的物体没有反应, 表明 PPA 表征场景的空间位置信息。这种观点得到了他最近的一个研究<sup>[28]</sup>以及生理病理学方面<sup>[31]</sup>的证据支持。如有的海马旁回损伤的病人不能辨别出场景的整体结构, 却能再认场景中的物体, 表明海马旁回表征的是场景的空间结构信息。Bar 等<sup>[26, 29]</sup>在控制了场景图片的复杂度的前提下, 设置了高情境联结 (contextual associations) 的场景 (即场景中的物体与特定的空间情境有强烈联系) 与低情境联结的场景 (即场景中的物体可能出现在多种情境中) 两种条件, 发现与观看低情境联结的场景相比, 当被试观看高情境联结的场景时, PHC 和 RSC 的激活增加。情境联结的强度会影响 PHC 和 RSC 的激活, 表明 PHC 和 RSC 可能编码的是通过经验而储存在长时记忆中的情境联结的知识 (如“厨房中可能会出现些什么物体” )。

虽然 PHC/PPA 和 RSC 都与场景一致性效应有关, 但它们在编码的具体内容上可能存在一些差别<sup>[26-27, 29, 32]</sup>。Bar 等<sup>[26]</sup>发现, PHC 在高情境联系条件下的激活高于低情境联系条件, 而两种条件下 RSC 的激活则相同。这可能是由于 RSC 编码较抽象的内容, 而 PHC 则编码情境的物体属性及相关的联结成分等更具体的内容。Park 等<sup>[32]</sup>的研究表明, PPA 表征场景每个视角的特定细节, 而 RSC 表征包括多种视角的抽象场景, 因而从另一个角度验证了这个观点。另外有研究<sup>[27]</sup>发现, 在分类任务中对于倒置的场景, V1 的激活并没有减少, 而 PPA 的激活降低, 即存在场景倒置效应 (scene inversion effect)。这可能是由于倒置影响了我们对场景意义的理解, 并使整体的场景布局 (global scene layout) 意义不那么明确, 因而场景的情境性大大降低甚至接近消失, 从而使 PPA 的激活减少。但由于场景的物理特征本身并不受倒置的影响, 因而 V1 的激活并没有减少。但令人惊奇的是, RSC 的激活并不受倒置的影响。这似乎也论证了 RSC 与 PHC/PPA 负责不同的加工。

### 4 场景一致性效应的理论模型

针对场景的一致性效应, 许多研究者提出了相

关的理论解释, 如相互作用模型、情境引导模型和情境促进模型等。其中, 只有情境促进模型能够较好地解释为何与场景一致的物体相比于与场景不一致的物体更慢地吸引注意却能更快地得到加工。

#### 4.1 相互作用模型

早期的场景一致性效应单方面强调一致的场景有利于物体的加工, 不一致场景不利于物体的加工, 忽视了物体对其所在情境的影响。后来, Davenport 和 Potter 将一致性效应的范围扩大, 认为一致性效应并不只是情境对物体加工的单方面作用, 而是相互的, 背景为物体提供情境, 反过来物体也为背景提供情境, 这就是视觉场景加工的相互作用模型(interactive model)<sup>[10-11]</sup>。物体和情境一致时比不一致时对场景的加工更快, 是借助两种可能的机制实现的: 第一, 场景一致时促进对场景的知觉加工, 使知觉加工的速度更快; 第二, 场景一致时降低辨别所需要的阈值, 使我们作出判断的标准更低。这两种可能的机制都会使对一致场景加工所需要的时间减少, 从而表现出相应的优势。

相互作用模型简单明了, 符合直觉经验, 但却只是一种笼统的描述, 对一些细节并没有作出更具体的说明。比如, 物体和背景一致时使知觉加工速度更快的具体机制是怎样的, 或者判断阈值是如何受物体和背景关系影响的, 等等。

#### 4.2 情境引导模型

Torralba 和 Oliva 等<sup>[33-34]</sup>提出了情境引导模型(contextual guidance model), 认为信息加工存在两条并行的通路: 局部通路(local pathway)和全局通路(global pathway)。局部通路主要通过提取局部场景的颜色、亮度和方向等物理特征, 形成场景的显著图(saliency map), 其中突显的区域即为视觉优势区, 也就是最吸引注意的区域。具体算法可参照 Itti 等的研究<sup>[35-36]</sup>。全局通路主要表征整体场景的统计性信息和任务要求, 并激活已有的知识经验, 从而对注意起指引作用。

该模型就其实质来讲, 是一种强调自下而上和自上而下加工并行的理论, 其中局部通路强调的是自下而上的注意吸引, 而全局通路则是自上而下的注意指引。由于有具体算法的支撑, 情境引导模型是一种可以实际运用于人工智能的理论, 这是它的一大优势。但其主要问题在于, 在计算显著图时采用的是基于贝叶斯框架的计算机视觉方法, 虽然这种方法能在较大程度上拟合纯刺激的物理特性吸引注意的特点, 但毕竟只是一种理论上的算法, 不可

能真正符合视觉系统的实际特点。此外, 个体经验对人的知觉加工具有很大的影响, 而情境引导模型对此考虑有所欠缺。这些都是该模型在模拟人的视觉加工过程中要解决的问题。

#### 4.3 情境促进模型

一致性效应表明, 与情境一致的物体能更早地得到加工, 但绝大部分研究尤其是眼动研究却发现眼睛更早注视不一致物体<sup>[21, 37-39]</sup>, 虽然对这一点还存在一定的争议<sup>[40-41]</sup>。也就是说, 虽然更晚注视与情境一致的物体, 但它却能更快地得到加工, 这似乎与我们的直觉相矛盾。究竟是什么因素导致这种“矛盾”呢? 上述两种理论并没有对此作很好的解释, 而 Bar 的情境促进模型(context facilitation model)很好地说明了这一点<sup>[24, 42-43]</sup>。

Bar 认为存在背侧的大细胞通路(dorsal magnocellular pathway)和腹侧的小细胞通路(ventral parvocellular pathway)。大细胞通路能快速形成一个低空间分辨率(low spatial frequency, LSF)的表征(即一个模糊的图像)。此时物体只得到了部分的加工, 形成的是场景的粗糙表征(coarse representation), 这个表征能够通过某种捷径从初级视皮层直达前额叶皮层(prefrontal cortex, PFC)和 PHC。PFC 根据 LSF 信息形成一些可能物体的表征, 在 PHC 处则会产生“情境框架(context frames)”<sup>[11]</sup>, 它相当于图式, 包含了对一个特定的场景中可能会出现什么物体及出现在什么位置等信息的猜测(initial guesses)。小细胞通路则是自下而上的分析, 这条通路能很好地加工细节信息, 但速度较慢。然后产生的猜测信息通过前馈机制与正在进行的自下而上的分析在下颞叶皮层(inferior temporal cortex, IT)整合<sup>[24, 42-43]</sup>。因而, 如果物体与场景的语义是一致的, 情境框架对场景中可能出现的物体的预测就是有效的, 这种自上而下的信息能够减少需要考虑的物体表征的数量, 加快对物体的加工, 反之, 如果物体与场景的语义不一致, 这种预测就是无效的, 由于与预期不符, 反而会不利于对物体的辨别。

Bar 的模型得到了许多行为学研究的支持。首先, 我们确实可以在场景中的物体和其他细节信息并未得到充分加工的前提下快速(40~100 ms)地获得场景的要义(gist)<sup>[2, 44-49]</sup>, 也就是说可以“只见森林不见树木(seeing the forest without representing the trees)”<sup>[47]</sup>。这一点非常重要, 是背侧通路的基础, 甚至可以说是 Bar 的理论基础。正是基于快速提取场景的要义, 才使自上而下的这条通路成为可能,

也才使这个理论模型能够顺利解释前面提到的“矛盾”。其次, 尽管能在极短的时间内提取场景的要义, 但对细节的加工是一个相对较慢的过程, 因而要完成对场景的加工还需要额外的时间<sup>[48]</sup>, 而这是以视觉信息可以随着时间而积累<sup>[50-51]</sup>这一事实为基础的。

Bar 的模型也得到了脑成像和电生理学方面的证据支持。Bar 等发现 OFC 确实先于颞叶得到激活, 而且是由图片的 LSF 信息导致的<sup>[24]</sup>, 这与情境促进模型的解释一致。最新的一项研究则结合了 fMRI 和 ERPs 技术, 在实验中采用两种呈现序列: 先呈现 LSF 图片, 再呈现 HSF 的图片; 或者先呈现 HSF 图片, 再呈现 LSF 的图片。结果发现, 对于前一序列, LSF 首先会在前额叶和颞枕区产生激活(刺激呈现后的 140~160 ms), 接着 HSF 使初级视皮层激活, 而对后一序列则不存在这种情况<sup>[25]</sup>。这也证明了对场景图片加工时是先由 LSF 信息激活高级皮层, 然后再对 HSF 信息进行加工的。

同情境引导模型一样, 情境促进模型也是一种强调自上而下与自下而上两种加工方式共同起作用的理论。自上而下的加工通过 LSF 信息使我们对场景中的可能物体形成“初始猜测(initial guesses)”, 这样就限制了候选物体的范围, 然后通过自上而下的分析在下颞叶的整合, 对物体进行识别。该模型的一大特点是从生理学角度解释了场景一致效应, 也得到了相关的证据支持, 但最大的不足在于片面地强调情境对物体的单方面影响, 对物体影响情境加工这一点考虑不足<sup>[12]</sup>。

## 5 讨 论

我们无时无刻不处于一定的场景之中, 而同一个物体处于不同的场景中会影响我们的知觉加工。例如, 对于一个平时很文静的人, 若在图书馆看到他, 我们可以很快地确定是他; 而如果他出现在舞厅, 或许我们就不会轻易地断定是这个人。这种现象就是场景的一致性效应。这种效应获得了很多行为学以及电生理学的验证, 而且也有不少脑成像的研究探讨其内在的生理机制。

首先, 场景一致性效应获得了行为学和电生理学两方面证据的支持。行为学研究采用命名、分类、探测等任务都证明了一致效应的存在。ERPs 的研究也提供了另一个角度的证据。N390 与 N400 在加工进程与产生源上很相似, 因而一致场景和一致场景波形的差异反应了一致性效应确实是由于

物体和情境的语义匹配关系造成的。虽然这些研究都表明场景的一致性效应是一种稳定存在于场景知觉中的现象, 但它们在材料的选择上还存在的问题, 如有些研究采用线条画或三维图, 但它们与现实的场景还是不同的, 尤其是线条画存在较大的差别。还有一些研究采用真实场景的图片, 但对其物理特征的控制并不严格, 这也可能会影响实验的结果。目前关于场景一致性效应产生的时间进程的研究还非常少, 但我们能从已有的少量研究中得到一个大致的结论。首先由 LSF 信息激活 OFC (130 ms 左右)、PHC/PPA 和 RSC, 之后 LSF 信息与 HSF 信息在颞叶进行整合。这时如果物体与情境语义不一致(失匹配), 会在 250 ms 左右开始诱发出一个更大的负走向波(N390)。然而遗憾的是, 虽然 PHC/PPA 和 RSC 是与场景加工相关的重要脑区, 但却没有研究明确指出 PHC/PPA 和 RSC 是何时激活的。根据已有的研究, 我们推测, 它们的激活应该与 OFC 的激活时间大致相同。当然, 确切的激活时间还需要相关研究的论证。

其次, 场景一致性效应的生理基础方面仍有一些不清楚的地方。随着研究的深入, 越来越多的研究者借助脑成像技术探究情境对物体加工的影响, 这有助于我们更好地了解一致性效应的内在加工机制和产生基础。就目前已有的研究来讲, 发现主要由 PHC/PPA 和 RSC 两个脑区对场景的加工起作用。但还有两个问题没有明确。第一, 虽然这两个脑区都对场景一致性效应起调节作用, 但它们在具体分工上还可能存在一定的差别, 如 Bar 等<sup>[29]</sup>认为 RSC 负责较抽象的内容的加工, 而 PHC 主要负责较具体的内容的加工。对于这一点仍然不是很明确。第二, 除了上述两个比较确定的起作用的脑区外, 还可能其他脑区的参与, 如 Bar 等<sup>[29]</sup>发现 MPFC 可能在基于情境对物体的预测中起了重要作用。另一个更重要的问题是, 目前所有的研究都只是针对物体或情境或物体与情境一致的场景情况, 还没有研究关注不一致场景的脑机制问题。我们不禁要问, 如果物体与情境不一致时大脑的激活会是怎样的? 是和一致的场景一样, 还是这种不一致会削弱 PHC/PPA 与 RSC 的激活, 又或者是有其他不同的脑区负责这种场景的加工呢?

最后, 对于场景一致性效应的解释, 相互作用模型显得简单易懂, 但其内在机制究竟是通过促进对场景的知觉加工, 还是降低辨别所需要的阈值, 或者两者共同起作用呢? 对此还没有一个确定的答

案. 背景引导模型同时考虑了自下而上和自上而下的加工通路, 试图解释现实生活中人的视觉加工过程, 但其算法过于机械, 对人的主观经验等考虑还不够. 此外, 相互作用模型和背景引导模型只是单纯地解释了背景与物体一致时会促进对场景加工的原因, 对于即使不一致的物体先吸引注意却识别得更慢这一事实没有作出说明. 而 Bar 的情境促进理论可以较好地对此作出解释. 该理论模型以脑成像的发现为基础, 能比较好地说明情境对物体的影响, 但忽视了物体对情境的影响<sup>[12]</sup>. 对一些细节问题的描述显得有些笼统, 如通过情境框架的引导形成对物体的猜测, 但没有说明情境框架是如何形成的. 要理清这些问题, 还需要进行更进一步的研究.

## 6 小 结

有关场景一致性效应的研究发现, 人们对与背景语义一致的前景物体的命名、分类、搜索和再认等都快于与背景不一致的物体. 与一致的物体相比, 与情境不一致的物体在有些脑区, 尤其是中央顶区, 会诱发一个幅度更大的负波 (N390). PHC/PPA 和 RSC 是负责场景加工的重要脑区. 场景一致性效应的时间进程可能首先由 LSF 信息激活 OFC (130 ms 左右)、PHC/PPA 和 RSC, 之后 LSF 信息与 HSF 信息在颞叶进行整合. 在诸多理论模型中, 情境促进模型从生理学角度对一致性效应作了较充分的解释.

## 参 考 文 献

- [1] Bar M, Ullman S. Spatial context in recognition. *Perception*, 1996, **25**(3): 343-352
- [2] Biederman I R, Mezzanotte, Rabinowitz J. Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology*, 1982, **14**(2): 143-177
- [3] Biederman I. Perceiving real-world scenes. *Science*, 1972, **177**(4043): 77-80
- [4] Boyce S J, Pollatsek A. Identification of objects in scenes: the role of scene background in object naming. *J Exp Psychol Learn Mem Cogn*, 1992, **18**(3): 531-543
- [5] Joubert O R, Rousselet G A, Fize D, Fabre-Thorpe M. Processing scene context: Fast categorization and object interference. *Vision Research*, 2007, **47**(26): 3286-3297
- [6] Hayes S M, Nadel L, Ryan L. The effect of scene context on episodic object recognition: Parahippocampal cortex mediates memory encoding and retrieval success. *Hippocampus*, 2007, **17**(9): 873-889
- [7] Biederman I, Glass A, Stacy E. Searching for objects in real-world scenes. *J Exper Psychol*, 1973, **97**(1): 22-27
- [8] Palmer S E. Effects of contextual scenes on identification of objects. *Mem Cogn*, 1975, **3**(5): 519-526
- [9] Neider M B, Zelinsky G J. Scene context guides eye movements during visual search. *Vision Research*, 2006, **46**(5): 614-621
- [10] Davenport J, Potter M. Scene consistency in object and background perception. *Psychol Sci*, 2004, **15**(8): 559-564
- [11] Davenport J. Consistency effects between objects in scenes. *Mem Cogn*, 2007, **35**(3): 393
- [12] Joubert O R, Fize D, Rousselet G A, Fabre-Thorpe M. Early interference of context congruence on object processing in rapid visual categorization of natural scenes. *J Vision*, 2008, **8**(13): 1-18
- [13] Kutas M, Hillyard S A. Reading senseless sentences - brain potentials reflect semantic incongruity. *Science*, 1980, **207**(4427): 203-205
- [14] Ganis G, Kutas M. An electrophysiological study of scene effects on object identification. *Cogn Brain Res*, 2003, **16**(2): 123-144
- [15] Glaser W R, Glaser M O. Context effects in Stroop-like word and picture processing. *J Exper Psych: General*, 1989, **118**(1): 13
- [16] Mudrik L, Lamy D, Deouell L. ERP evidence for context congruity effects during simultaneous object-scene processing. *Neuropsychologia*, 2010, **48**(2): 507-517
- [17] Sitnikova T, Kuperberg G, Holcomb P J. Semantic integration in videos of real-world events: An electrophysiological investigation. *Psychophysiology*, 2003, **40**(1): 160-164
- [18] Sitnikova T, Holcomb P J, Kiyonaga K A, Kuperberg G R. Two neurocognitive mechanisms of semantic integration during the comprehension of visual real-world events. *J Cogn Neurosci*, 2008, **20**(11): 2037-2057
- [19] Levin D T, Simons D J. Perceiving stability in a changing world: Combining shots and integrating views in motion pictures and the real world. *Media Psychology*, 2000, **2**(4): 357-380
- [20] Thorpe S, Fize D, Marlot C. Speed of processing in the human visual system. *Nature*, 1996, **381**(6582): 520-522
- [21] Evans K K, Treisman A. Perception of objects in natural scenes: Is it really attention free?. *J Exper Psych-Human Perception and Performance*, 2005, **31**(6): 1476-1492
- [22] Underwood G, Templeman E, Lamming L, Foulsham T. Is attention necessary for object identification? Evidence from eye movements during the inspection of real-world scenes. *Consciousness Cogn*, 2008, **17**(1): 159-170
- [23] Potter M C, Fox L F. Detecting and remembering simultaneous pictures in a rapid serial visual presentation. *J Exper Psych-Human Perception and Performance*, 2009, **35**(1): 28-38
- [24] Bar M, Kassam K S, Ghuman A S, Boshyan J, Schmidt A M, Dale A M, Hamalainen MS, Marinkovic K, Schacter D L, Rosen B R, Halgren E. Top-down facilitation of visual recognition. *Proc Natl Acad Sci USA*, 2006, **103**(2): 449-454
- [25] Peyrin C, Michel C M, Schwartz S, Thut G, Seghier M, Landis T, Marendaz C, Vuilleumier P. The neural substrates and timing of top-down processes during coarse-to-fine categorization of visual

- scenes: A combined fMRI and ERP study. *J Cogn Neurosci*, 2010, **22**(12): 2768–2780
- [26] Bar M, Aminoff E. Cortical analysis of visual context. *Neuron*, 2003, **38**(2): 347–358
- [27] Walther D B, Caddigan E, Fei-Fei L, Beck D M. Natural scene categories revealed in distributed patterns of activity in the human brain. *J Neurosci*, 2009, **29**(34): 10573–10581
- [28] Epstein R A, Ward E J. How reliable are visual context effects in the parahippocampal place area?. *Cerebral Cortex*, 2010, **20**(2): 294–303
- [29] Bar M, Aminoff E, Schacter D L. Scenes unseen: The parahippocampal cortex intrinsically subserves contextual associations, not scenes or places per se. *J Neurosci*, 2008, **28**(34): 8539–8544
- [30] Epstein R, Kanwisher N. A cortical representation of the local visual environment. *Nature*, 1998, **392**(6676): 598–601
- [31] Mendez M F, Chierri M M. Agnosia for scenes in topographagnosia. *Neuropsychologia*, 2003, **41**(10): 1387–1395
- [32] Park S, Chun M M. Different roles of the parahippocampal place area (PPA) and retrosplenial cortex (RSC) in panoramic scene perception. *Neuroimage*, 2009, **47**(4): 1747–1756
- [33] Oliva A, Torralba A. The role of context in object recognition. *Trend Cogn Sci*, 2007, **11**(12): 520–527
- [34] Torralba A, Oliva A, Castelhana M S, Henderson J M. Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psych Rev*, 2006, **113**(4): 766–786
- [35] Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis. *Ieee Trans Patt Anal Mach Intell*, 1998, **20**(11): 1254–1259
- [36] Itti L, Koch C. A saliency-based search mechanism for overt and covert shifts of visual attention. *Vis Res*, 2000, **40**(10–12): 1489–1506
- [37] Brockmole J, Henderson J. Prioritizing new objects for eye fixation in real-world scenes: Effects of object-scene consistency. *Visual Cognition*, 2008, **16**(2/3): 375–390
- [38] Loftus G, Mackworth N. Cognitive determinants of fixation location during picture viewing. *J Exper Psych: Human Perception and Performance*, 1978, **4**(4): 565–572
- [39] Underwood G, Foulsham T. Visual saliency and semantic incongruity influence eye movements when inspecting pictures. *Quarterly J Exper Psych*, 2006, **59**(11): 1931–1949
- [40] Henderson J M, Weeks P A, Hollingworth A. The effects of semantic consistency on eye movements during complex scene viewing. *J Exper Psych-Human Perception and Performance*, 1999, **25**(1): 210–228
- [41] De Graef P, Christiaens D, d'Ydewalle G. Perceptual effects of scene context on object identification. *Psych Res*, 1990, **52**(4): 317–329
- [42] Bar M. A cortical mechanism for triggering top-down facilitation in visual object recognition. *J Cogn Neurosci*, 2003, **15**(4): 600–609
- [43] Bar M. Visual objects in context. *Nat Rev Neurosci*, 2004, **5**(8): 617–629
- [44] Oliva A. *Gist of the Scene*//Itti L, Rees G, Tsotsos J. *Neurobiology of Attention*. San Diego, CA: Elsevier, 2005: 251–256
- [45] Rousselet G, Joubert O, Fabre-Thorpe M. How long to get to the "gist" of real-world natural scenes?. *Vis Cogn*, 2005, **12**(6): 852–877
- [46] Oliva A, Torralba A. Building the gist of a scene: the role of global image features in recognition//Waxman S, Stein D G, Swaab D, Fields H. *Visual Perception, Pt 2: Fundamentals of Awareness: Multi-Sensory Integration and High-Order Perception*, 2006, **155**: 23–36
- [47] Greene M R, Oliva A. Recognition of natural scenes from global properties: Seeing the forest without representing the trees. *Cogn Psych*, 2009, **58**(2): 137–176
- [48] Rayner K, Smith T J, Malcolm G L, Henderson J M. Eye movements and visual encoding during scene perception. *Psychological Science: J American Psychological Society/APS*, 2009, **20**(1): 6–10
- [49] Castelhana M S, Henderson J M. The influence of color on the perception of scene gist. *J Exper Psych-Human Perception and Performance*, 2008, **34**(3): 660–675
- [50] Melcher D. Accumulation and persistence of memory for natural scenes. *J Vis*, 2006, **6**(1): 8–17
- [51] Goto S G, Ando Y, Huang C, Yee A, Lewis R S. Cultural differences in the visual processing of meaning: Detecting incongruities between background and foreground objects using the N400. *Soc Cogn Affect Neurosci*, 2009, **5**(2): 242–253

## Scene Consistency Effect and Its Mechanisms\*

GONG Ming-Liang, XUAN Yu-Ming\*\*, FU Xiao-Lan

(State Key Laboratory of Brain and Cognitive Science, Institute of Psychology, The Chinese Academy of Sciences, Beijing 100101, China)

**Abstract** The present paper reviewed the evidence for scene consistency effect accumulated recently. Behavioral studies showed that participants performed better in naming, categorizing, searching and recognizing consistent objects (which appeared in usual context) than inconsistent objects (which appeared in unusual context). ERP studies demonstrated that inconsistent objects, compared with consistent objects, induced a larger negative-going wave (N390) especially in centro-parietal sites. fMRI studies revealed that PHC/PPA and RSC played important roles in scene processing. Theories proposed to explain the scene consistency effect were also discussed.

**Key words** consistency effect, scene, context, semantic relationship, theoretical model, physiological mechanism

**DOI:** 10.3724/SP.J.1206.2011.00032

---

\* This work was supported by grants from National Basic Research Program of China (2011CB302201) and The National Natural Science Foundation of China (90820305, 61075042).

\*\*Corresponding author.

Tel: 86-10-64837040, E-mail: xuanym@psych.ac.cn

Received: January 18, 2011 Accepted: March 28, 2011