

# 场景知觉研究综述及展望

詹萍

内蒙古师范大学心理学院, 内蒙古 呼和浩特

收稿日期: 2023年4月12日; 录用日期: 2023年5月30日; 发布日期: 2023年6月7日

## 摘要

场景知觉是指人们如何感知和处理真实环境中的复杂信息。这些场景包括物体和背景两个重要的组成部分, 并根据加工水平不同, 可以分为三种不同的类型。对于场景信息的处理机制, 存在多种理论解释, 如直觉整合模型、特征整合模型、贝叶斯层级模型、导向搜索模型、背景引导模型和双通道加工模型。研究者已经尝试采用自上而下和自下而上两种方式来解释场景信息的提取和加工, 并尝试从二者交互的角度进行解释。此外, 不同的实验目的和技术, 如搜索范式、变化觉察范式、物体觉察范式、眼动范式和背景提示范式等, 被用于研究场景信息的知觉问题。但是, 场景的定义、不同理论模型间以及范式间的整合和不同加工阶段的加工方式仍需要进一步深入和探讨。

## 关键词

场景知觉, 理论模型, 加工方式, 研究范式, 展望

# A Review and Prospect of Scene Perception Research

Ping Zhan

School of Psychology, Inner Mongolia Normal University, Hohhot Inner Mongolia

Received: Apr. 12<sup>th</sup>, 2023; accepted: May 30<sup>th</sup>, 2023; published: Jun. 7<sup>th</sup>, 2023

## Abstract

Scene perception refers to how people perceive and process complex information in real-world environments. These scenes consist of two important components, objects and backgrounds, and can be categorized into three different types depending on the level of processing. There are multiple theoretical explanations for how scene information is processed, including intuitive integration models, feature integration models, Bayesian hierarchical models, guided search models, background guidance models, and dual-channel processing models. Researchers have attempted

to explain the extraction and processing of scene information using both top-down and bottom-up approaches, and have tried to explain the interaction between the two. In addition, different experimental paradigms and techniques, such as search paradigms, change detection paradigms, object perception paradigms, eye movement paradigms, and background cueing paradigms, have been used to study perception problems related to scene information. However, the definition of scenes, the integration of different theoretical models and paradigms, and the processing methods of different stages of processing still need further exploration and discussion.

## Keywords

Scene Perception, Theoretical Model, Processing Pattern, Research Paradigm, Prospect

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

近年来, 视知觉研究成为基础心理学和实验心理学中的重点。随着技术的不断进步和研究的深入, 越来越多的研究者开始关注人们如何知觉和加工周围的真实场景。早期的视知觉研究将人们的视觉加工过程划分为三个层级[1]。第一层级的研究关注于物理属性特征, 例如颜色、形状、大小、表面和边界等; 第二层级的研究者则主要研究刺激物与背景空间关系的提取, 但不涉及心理意义的获取; 第三层级的研究则主要研究物体在视觉系统中的表征过程及其心理意义的获取, 这个层级主要探究认知加工和知觉表征之间相互关系, 包括场景信息的提取、刺激物识别以及视觉工作记忆等内容。

## 2. 场景与场景知觉

### 2.1. 场景的界定

场景是由多个离散的物体及其背景构成的一种具有连贯性的图像, 同时具有语义一致性效应[1]。场景由两个部分组成: 背景和物体。背景通常指表面和结构范围较大的相对静止的物体, 如房间和地板等。物体则指范围较小且不连续的刺激物, 如书桌和挂钟等。在场景知觉中, 背景和物体之间的关系是相对的, 一个刺激物既可作为物体也可作为背景[2]。场景还可以根据构成方式分为真实场景和人工场景。真实场景是具有独特物理属性的离散刺激物和它们所处的空间背景相结合而形成的一种特定的空间层次结构。人工场景则是由简单的图形符号构成, 同时具有人为的结构特征和语义信息。

### 2.2. 场景知觉的概念及分类

场景知觉是人类视觉系统对真实场景中的刺激物和背景进行识别和加工的过程[3]。人类的视觉加工可以分为三个水平: 初级水平、中级水平和高级水平。初级水平是提取场景的基本内容; 中级水平是提取场景中独立、离散的刺激物的特征信息; 高级水平是加工场景的语义信息。而场景知觉研究的重点就在于探究人们视觉系统的高级水平在加工和识别真实场景中刺激物和背景时的工作模式。

目前, 场景知觉的研究利用三种类型的材料, 分别为低、中、高水平的视觉研究。低水平的研究以简单无意义场景为研究对象; 中等水平的研究则将多个物体简单放置在一起作为场景; 高级水平的研究以真实的情景为研究对象, 例如利用彩色图片、3D 模拟场景图或虚拟现实系统等, 其研究目的在于研究

客体的语义概念、知觉加工过程以及多种认知方式的相互作用等等[4] [5]。这三种类型的材料反映了场景知觉研究的不同层次和水平。

### 3. 场景知觉的理论模型

有关场景知觉研究的理论基础主要包括两个方面：感知组织理论和知觉加工流理论。感知组织理论认为，人类对视觉信息的处理是按照一定的规则和原则进行的，即人们能够将视觉信息分解成各个部分，而整合这些部分得到完整的视觉信息。这些规则和原则包括图形分离原则、相邻性原则、共形性原则等。知觉加工流理论则认为，人类对视觉信息的加工是一个分层次的过程，即先对信息进行初步的预处理，再进入后续的加工流程。这个理论认为，视觉信息的加工流程包括初级视觉加工(边界检测、色彩感知等)和高级视觉加工(目标识别、场景知觉等)两个阶段。这些理论为场景知觉研究提供了理论基础，同时在具体的场景知觉研究中，研究者也会结合具体的实验任务和方法对这些理论进行验证和完善。

#### 3.1. 直觉整合模型(Intuitive Visual Integration Model)

场景知觉中，人们通过视觉搜索捕获视觉信息。直觉整合模型认为每个注视点都有其独特的表征，即使在同一位置出现的系列刺激没有在视网膜的同一位置也可以进行整合[6]。在视觉搜索时，先捕获到的注视点信息会存储在一个缓存加工器中，并整合前面获得的信息，形成完整的视觉信息表征[7]。每次搜索都需要进行调整，以与真实情景中客观刺激物位置的内在模型相一致。

#### 3.2. 特征整合模型(Feature Integration Model)

特征整合模型是 Treisman 和 Gelade 在研究视觉加工过程中提出的，该模型认为视觉加工包含特征登记和特征整合两个过程[8]。在特征登记过程中，视觉系统自动化地加工场景中客体的特征，形成各个维度的特征地图；在特征整合过程中，知觉系统通过相关的知识和经验将之前独立的特征整合成完整的客体表征。场景知觉是视觉特征分析和识别网络共同作用的结果，但更强调知觉特征加工的自下而上的过程。

#### 3.3. 贝叶斯等级模型(Bayesian Hierarchical Model)

贝叶斯等级模型是基于特征整合理论提出的，可用于表征场景中的客体信息。关于场景知觉，人们对于场景中客体的识别持不同看法，Li 等人结合贝叶斯框架提出了层次模型。该模型由四个层次组成，包括图式子成分、图式、场景和类别[9]。场景中的每个区域表征为不同图式的子成分，对场景的表征是通过局部信息的累计完成的[10]。根据模型，场景图式表征不是自上而下的处理过程，而是自动完成的。在场景知觉中，语义约束、客体关系和任务约束等因素会影响场景信息的整合。

#### 3.4. 导向搜索模型(Guided Search Model)

根据 Wolfe 提出的导向搜索模型，在初级视觉加工阶段，我们通过平行处理各位置的视觉信息获取信息的基本特征，这些特征会激活各个视网膜投射图形成激活地图[11] [12]。这一激活地图引导着随后的搜索过程，该过程通过确定眼动的方向、合理分配注意资源和忽视不符合目标刺激特征的刺激，从而使视觉特征的整合更有效。

#### 3.5. 显著性地图模型(Saliency Map Model)

显著性地图模型旨在从视觉分析和认知加工视角来分析场景知觉的加工，主要关注注视位置和注视时间等因素[13]。模型中，视觉信息的初步分析会影响对目标的视觉表征；视觉空间会结合物理特点划分

出潜在兴趣范围，并赋予显著性权重，这个显著性权重结合空间位置关系构成显著性地图。显著性权重由客体的视觉刺激物理特点及目标任务带来的自上而下的主观经验共同决定。在视觉信息加工过程中，有两个基本规则：一是注意分布在最高意义的视觉场景的空间面积，二是眼睛总体需要的注视时间由两个完整的区域意识的分析和决策的认知加工时特征决定。

### 3.6. 背景引导模型(Contextual Guidance Model)

Torralla 等人提出了背景引导模型，该模型是在贝叶斯等级模型基础上，利用计算机算法模拟人类识别过程，研究背景如何影响场景和客体识别。模型为场景中目标搜索提供了概率框架，通过整合场景整体信息和局部信息来确定目标位置[14] [15]。背景引导模型认为，客体识别受场景背景信息和过去经验的共同影响。

### 3.7. 双通道加工模型(Dual-Path Model of Scene Processing)

Ungerleider 和 Haxby 在研究中发现，人类的视觉加工可以分为腹侧通路和背侧通路[16]。腹侧通路主要负责加工刺激信息，背侧通路则加工空间信息。虽然这两个通路并不是完全独立的，但双通道模型可以帮助我们理解情景识别以及相关研究结果。根据已有的研究结果，Chun 提出了场景识别的双通道加工模型，认为整体特征会引导背侧通路加工空间信息，腹侧通路则表征场景中的刺激物[17]。

不同的理论模型虽然对于理解场景知觉有重要意义，但需要综合这些理论并进一步探索。

## 4. 场景知觉的加工方式

人们的视觉系统每天接收着很多输入的刺激信息，他们通过自下而上和自上而下两种方式进行加工和识别。自下而上加工方式是以物体的物理特征驱动整个场景的加工过程，这种方式会吸引人们的注意力，且是被动接受场景本身的属性的引导。自上而下的加工方式是基于知觉者的知识、经验、期望和动机[18]，这种加工方式在场景非常模糊，物理特征并不明显无法识别出刺激物时依然可以识别一个场景[19]。相比于在人工场景中对刺激物表面特征的加工，真实场景中更多涉及语义信息的获得，这种加工方式是对低水平的物理特征与知识经验的整合[20]。在视觉搜索中，场景语义信息的引导比物理属性的引导作用更大，在搜索过程中对场景语义的理解也会影响视觉搜索。这两种方式对场景的加工和识别起到了不同的作用，其中自下而上的加工已经足够让人们识别一个场景，而自上而下的加工则能预先提示语义信息。

## 5. 场景知觉的理论模型

场景知觉研究区别于其他视觉认知研究的独特之处在于，它不仅考虑了背景和客体信息，还需要考虑它们之间的相互关系。然而，这些因素可能会成为实验中影响相关实验变量的额外变量，从而影响研究结果的纯净性。因此，针对不同的研究目的，研究者选取不同的研究范式来突出重点。常见的场景知觉研究范式包括搜索范式、变化觉察范式、物体觉察范式、眼动范式和背景提示范式等[21]。

### 5.1. 搜索范式

在场景信息识别的研究中，搜索任务范式被广泛使用。在这种范式下，被试需要搜索一个或多个混杂在非目标刺激中的目标刺激，并记录反应时和判断的准确性以考察识别的绩效。搜索绩效的标准测量指标是搜索函数的斜率，它是单位时间内目标搜索处理量的一个估计。截距也是参考指标，反映了任务成分，包括注意选择之前的视觉加工以及决策和运动等成分。在某些实验条件下，可以从错误率中获得更多的信息。例如，刺激物呈现时间非常短会导致错误率升高[22]。



## 5.2. 变化觉察范式

变化觉察范式旨在了解被试对场景中物体的变化进行探测和识别的能力。该范式通常要求被试观察一个场景，场景中只允许改变一个物体或一个物体的物理属性，观察被试是否能察觉到这种变化。变化觉察包括客体变化的探测和识别两个过程，客体变化的探测要求被试报告是否有变化发生，而客体变化的识别则要求被试报告变化客体的位置。目前常用的变化觉察范式有两种形式：迫选范式和闪烁范式。迫选范式在空屏幕间隔之后在存储器图像中呈现存储器图片，然后立即呈现检测图片，考察被试是否能够探测到前后的变化。而闪烁范式是在考察被试在多个视觉刺激中寻找目标刺激并快速做出反应，目的是模拟眼球的快速移动并探讨视觉灵敏度和在眼跳过程中场景认知和表征的过程。

## 5.3. 物体觉察范式

Biederman 在他的一系列研究中最先使用并命名了物体觉察范式[23] [24] [25]，该范式将场景中目标物体的觉察准确性作为物体觉察的指标，包括对目标信息的搜索和对干扰信息的抑制[26]。该范式的经典流程是呈现目标物体的名称、注视点、场景图片、掩蔽刺激，最后让被试判断。物体觉察范式中常用的比较方法是在打乱重排场景图片和正常场景图片中客体觉察的反应时进行比较，但忽略了这两类场景图片在视觉方面的复杂性是不对等的，从而降低了实验结果的可信性。为了解决这个问题，研究者改进了经典范式，采用信号检测方法控制测量敏感性的影响，改进实验设计来控制信号检测还不能完全剔除的反应偏向。此外，线索呈现的顺序也得到了改变，一种新的命名叫做场景后强迫判别程序[27] [28]，该程序把线索放置在场景图片中，发现了不同于先前研究的结果。

## 5.4. 眼动范式

随着眼动仪的使用，能够更精准地掌握场景知觉的眼动模式及视觉反应指标。眼动技术已被引入变化探测范式，使得变化觉察范式能够更精准地把握在场景知觉中的视觉过程。眼动范式可以记录被试对变化觉察的反应时和正确率的同时，搜集并记录场景知觉过程中被试的眼动指标。有关场景知觉的眼动分析指标，有研究者从注视和眼跳的视角出发，将其分为以下指标：注视时间、凝视时间、注视位置、扫描模式比；眼跳幅度、眼跳次数和眼跳距离[29]。其中重点关注的是眼睛的注视位置和注视时间，即在观察一个场景时眼睛的注视过程、注视区域，停留时间，眼跳和信息整合。场景知觉中的加工时间是对场景在空间加工和时间上注视的整体函数表达。眼动研究范式在场景知觉中很少单独使用，大多会与其他实验任务相结合，探讨场景知觉中的信息提取、客体识别以及将场景中的信息编码和贮存等问题[5]。因此，眼动范式需要与其他研究范式相结合来研究场景知觉过程。

## 5.5. 背景提示范式

背景在知觉场景中起着重要的作用。背景可以为我们提供不同种类的信息，例如物体位置、大小和与场景的关系等。视觉背景信息的处理能够帮助我们更有效地了解所看到的世界。背景提示范式是基于视觉搜索研究发展而来的一种实验方法。实验要求被试在不同的背景条件下搜索目标物体，并判断其特定属性，例如倾斜方向。结果表明，重复出现的背景布局能够提高被试的反应速度，并具有提示作用。这种学习过程是内隐的[30]，不需要有意识的编码和记忆。背景提示范式支持视觉自上而下的加工过程，其理论假设认为视觉背景信息能够促进视觉场景的加工。

## 6. 场景知觉的应用

研究者使用眼动技术研究了不同经验水平教师的教学反思认知加工机制。他们发现新手教师主要使

用自下而上的场景知觉方式，而熟手教师则同时使用自上而下与自下而上的方式。这种差异源于教师的课堂行为洞察能力和风险知觉能力的不同。因此，提高教师的课堂行为洞察力是教学反思的关键，这需要调整教师的关注点[31]。还有研究者利用眼动技术研究了课堂突发事件场景图片的情绪场景和不同类型教师的加工差异。结果显示，专家、熟手和新手教师在观察不同课堂突发事件场景中存在显著差异，专家型教师采取的是观看时间长，注视次数少的策略，这证明经验对于教师应对课堂突发事件具有重要影响[32]。研究者对幼儿对教师不同课堂教学行为场景知觉的眼动进行了研究。研究发现，随着年级的增长，幼儿对教师的注视增多，而在贫乏背景下对教师的注视更多。此外，丰富的背景会影响幼儿首次注视时间并影响其他信息的注视，而指示行为和提问行为有利于幼儿对场景中教师的识别，指导行为则有利于对多媒体背景的识别。而且，6岁幼儿在注视教师时受背景的影响更大。通过对不同年龄阶段的比较，这一研究有助于推动有关不同群体的场景知觉的理论发展，并促进课堂教学的实践和应用[33]。

还有研究者对于高低抑郁倾向的大学生进行有关场景图片的研究，发现与低抑郁倾向大学生相比，高抑郁倾向大学生观看非情绪图片时表现出兴趣减退的特征；且观看情绪图片时表现出对负性情绪信息的加工优势，并表现出回避眼睛的特征[34]。后续对于有高抑郁倾向的群体应多接触积极的情绪信息，引导他们向一个有前瞻性的方向发展。何虎应用眼动技术，比较了蒙古国蒙古族大学生、中国内蒙古蒙古族大学生和汉族大学生的场景知觉特点，研究文化对认知的作用。结果表明，内蒙古蒙古族和汉族大学生的眼动指标无显著差异，而蒙古国大学生更加关注前景物体的知觉，这反映了文化差异对认知的影响[35]。还有研究者还探讨了刺激物数量、刺激物之间关系、刺激物-背景一致性等对背景识别的影响，发现它们均会影响背景的识别。这一结果有助于进一步研究场景识别的影响因素，丰富了相关研究，为未来提供实证支持[36]。

场景知觉作为一种人类的感知和认知能力，在现实生活和社会中扮演着重要角色，影响着人们的学习、工作、交往等方方面面。首先，在教育领域，场景知觉能够帮助学生更好地理解教材内容，识别并学习重要的信息和概念。教学中通过引导学生注视教师、黑板、幻灯、电脑等信源，可以帮助学生提高场景知觉能力，更好地理解 and 掌握所学知识。其次，在职场工作中，场景知觉也是一项重要的能力。例如，在工厂、车间的生产线上，工人需要通过场景知觉来掌握生产过程中的物流、品质和卫生状况等信息，并及时发现问题和解决问题。在商务谈判或社交场合中，场景知觉也帮助人们更好地理解对方的意图和态度，从而更有效地交流和合作。最后，场景知觉的研究对于认知科学领域以及教育心理学发展等领域都具有重要意义。通过场景知觉的研究，可以更深入了解人类感知和认知的基本机制，进而为相关理论的建构、认知模型的制定以及教学策略的制定提供基础和指导。

## 7. 场景知觉的展望

人类视觉加工的一个显著的特点就是能够快速识别场景刺激并归类。对于场景识别研究，目前的争论主要体现在两个方面：局部信息和整体信息之争以及语义信息和知觉信息之争。对于整体信息在场景识别中的作用，研究者认为整体信息可以通过场景的空间层次结构来推断[37]。在真实场景识别研究中，场景的快速识别是基于整体信息而非局部信息。然而，也有研究者认为局部信息和知觉信息的加工对场景识别有影响[38]。另外，语义信息会影响场景中刺激物的注意优先性[39]，但也有研究发现场景统计特征和边缘密度会影响注视点位置[40]。场景识别同时受到场景知觉信息的影响并有自下而上的场景空间注意模型[41]。

关于场景信息加工存在自上而下和自下而上的争论[14][41]。这个问题涉及到局部和整体信息的争论以及场景知觉属性和语义属性的注意优先性的争论。然而，现实场景知觉过程中，视觉信息加工机制仍然存在许多未解决的问题。这些问题包括关于如何提取知识以及知识如何对真实场景知觉产生作用的问题

题, 以及个体在知识提取和知识对眼动控制方面是否会表现出差异的问题。此外, 自下而上的加工通路究竟会控制哪些场景知觉属性, 以及提取的场景知觉特征对于个体实际知觉加工的拟合程度如何都需要进一步研究。最后, 在场景识别的视觉加工过程中, 存在两个通道加工并且不完全分离的问题, 而对于这种情况下的加工整合条件和二者的协同作用还需要更深入地研究。

## 参考文献

- [1] Henderson, J.M. and Hollingworth, A. (1999) High-Level Scene Perception. *Annual Review of Psychology*, **50**, 243-271. <https://doi.org/10.1146/annurev.psych.50.1.243>
- [2] 关荐. 场景知觉中客体的属性及语义关系对变化盲视的影响[D]: [博士学位论文]. 西安: 陕西师范大学, 2016.
- [3] 田宏杰, 王福兴, 徐菲菲, 申继亮. 场景知觉中物体加工的背景效应[J]. *心理科学进展*, 2010, 18(6): 878-886.
- [4] Henderson, J.M. (2005) Introduction to Real-World Scene Perception. *Visual Cognition*, **12**, 849-851. <https://doi.org/10.1080/13506280444000544>
- [5] Henderson, J. M. (2007) Regarding Scenes. *Current Directions in Psychological Science*, **16**, 219-222. <https://doi.org/10.1111/j.1467-8721.2007.00507.x>
- [6] Rayner, K. and McConkie, G.W. (1976) What Guides a Reader's Eye Movements? *Vision Research*, **16**, 829-837. [https://doi.org/10.1016/0042-6989\(76\)90143-7](https://doi.org/10.1016/0042-6989(76)90143-7)
- [7] Peterson, M. and Rhodes, G. (2003) Perception of Faces, Objects and Scenes: Analytic and Holistic Processes. Oxford University Press, Oxford.
- [8] Treisman, A.M. and Gelade, G. (1980) A Feature-Integration Theory of Attention. *Cognitive Psychology*, **12**, 97-136. [https://doi.org/10.1016/0010-0285\(80\)90005-5](https://doi.org/10.1016/0010-0285(80)90005-5)
- [9] Li, F.F., Van Rullen, R., Koch, C. and Perona, P. (2002) Rapid Natural Scene Categorization in the Near Absence of Attention. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 9596-9601. <https://doi.org/10.1073/pnas.092277599>
- [10] Fei-Fei, L. and Perona, P. (2005) A Bayesian Hierarchical Model for Learning Natural Scene Categories. *Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR' 05)*, San Diego, 20-25 June 2005, 524-531.
- [11] Wolfe, J.M., Friedman-Hill, S.R. and Bilsky, A.B. (1994) Parallel Processing of Part-Whole Information in Visual Search Tasks. *Perception & Psychophysics*, **55**, 537-550. <https://doi.org/10.3758/BF03205311>
- [12] 任延涛, 韩玉昌, 隋雪. 视觉搜索过程中的眼跳及其机制[J]. *心理科学进展*, 2006(3): 340-345.
- [13] Henderson, J.M. (1992) Object Identification in Context: The Visual Processing of Scenes. *Canadian Journal of Psychology/Revue Natural Canadienne de Psychologie*, **46**, 319-341. <https://doi.org/10.1037/h0084325>
- [14] Torralba, A., Oliva, A., Castelhano, M.S. and Henderson, J.M. (2006) Contextual Guidance of Eye Movements and Attention in Real-World Scenes: The Role of Global Features in Object Search. *Psychological Review*, **113**, 766-786. <https://doi.org/10.1037/0033-295X.113.4.766>
- [15] 康廷虎, 白学军. 真实情景知觉中注视控制的研究进展[J]. *西北师范大学学报(社会科学版)*, 2008, 45(4), 107-111.
- [16] Ungerleider, L.G. and Haxby J.V. (1994) "What" and "Where" in the Human Brain. *Current Opinion in Neurobiology*, **4**, 157-165. [https://doi.org/10.1016/0959-4388\(94\)90066-3](https://doi.org/10.1016/0959-4388(94)90066-3)
- [17] Chun, M.M. (2003) Scene Perception and Memory. *Psychology of Learning and Motivation: Advances in Research and Theory*, **42**, 79-108. [https://doi.org/10.1016/S0079-7421\(03\)01003-X](https://doi.org/10.1016/S0079-7421(03)01003-X)
- [18] Henderson, J.M. (2003) Human Gaze Control during Real-World Scene Perception. *Trends in Cognitive Sciences*, **7**, 498-485. <https://doi.org/10.1016/j.tics.2003.09.006>
- [19] Oliva, A. and Schyns, P.G. (1997) Coarse Blobs or Fine Edges? Evidence that Information Diagnosticity Changes the Perception of Complex Visual Stimuli. *Cognitive Psychology*, **34**, 72-107. <https://doi.org/10.1006/cogp.1997.0667>
- [20] Wu, C.-C., Wick, F.A. and Pomplun, M. (2014) Guidance of Visual Attention by Semantic Information in Real-World Scenes. *Frontiers in Psychology*, **5**, Article 54. <https://doi.org/10.3389/fpsyg.2014.00054>
- [21] 王福兴, 田宏杰, 申继亮. 场景知觉及其研究范式[J]. *心理科学进展*, 2009, 17(2), 268-277.
- [22] Dukewich, K. and Klein, R. (2005) Implications of Search Accuracy for Serial Self-Terminating Models of Search. *Visual Cognition*, **12**, 1386-1403. <https://doi.org/10.1080/13506280444000788>
- [23] Biederman, I. (1972) Perceiving Real-World Scenes. *Science*, **177**, 77-80. <https://doi.org/10.1126/science.177.4043.77>

- [24] Biederman, I. (1981) Do Background Gradients Depth Facilitate Object Identification. *Perception*, **10**, 573-578. <https://doi.org/10.1068/p100573>
- [25] Biederman, I., Mezzanotte, R.J. and Rabinowitz, J.C. (1982) Scene Perception: Detecting and Judging Objects Undergoing Relational Violations. *Cognitive Psychology*, **14**, 143-177. [https://doi.org/10.1016/0010-0285\(82\)90007-X](https://doi.org/10.1016/0010-0285(82)90007-X)
- [26] Vergheze, P. (2001) Visual Search and Attention: A Signal Detection Theory Approach. *Neuron*, **31**, 523-535. [https://doi.org/10.1016/S0896-6273\(01\)00392-0](https://doi.org/10.1016/S0896-6273(01)00392-0)
- [27] Henderson, J.M. and Hollingworth, A. (1998) Chapter 12-Eye Movements during Scene Viewing: An Overview. In: G. Underwood., Ed., *Eye Guidance in Reading and Scene Perception*, Elsevier, Oxford, 269-283. <https://doi.org/10.1016/B978-008043361-5/50013-4>
- [28] Castelhamo, M.S. and Henderson, J.M. (2005) Incidental Visual Memory for Objects in Scenes. *Visual Cognition*, **12**, 1017-1040. <https://doi.org/10.1080/13506280444000634>
- [29] 康廷虎, 张会. 场景知觉中的眼动分析指标: 基于注视和眼跳视角[J]. 心理科学, 2020, 43(6): 1312-1318.
- [30] Chun, M.M. and Jiang, Y. (1998) Contextual Cueing: Implicit Learning and Memory of Visual Context Guides Spatial Attention. *Cognitive Psychology*, **36**, 28-71. <https://doi.org/10.1006/cogp.1998.0681>
- [31] 王碧梅. 不同经验水平教师教学反思认知加工机制研究—基于眼动和访谈的证据[J]. 教师教育研究, 2022, 34(04): 77-85. <https://doi.org/10.13445/j.cnki.t.e.r.2022.04.018>
- [32] 汪海彬, 程思, 姚本先. 新手-熟手-专家型教师对课堂突发事件加工的比较——基于眼动的研究[C]//中国心理学会. 第二十二届全国心理学学术会议摘要集. 2019: 581.
- [33] 王菲菲. 幼儿对教师不同课堂教学行为场景知觉的眼动研究[D]: [硕士学位论文]. 乌鲁木齐: 新疆师范大学, 2015.
- [34] 贾旖璠. 高、低抑郁倾向大学生图片加工的眼动特征差异[D]: [硕士学位论文]. 天津: 天津师范大学, 2022. <https://doi.org/10.27363/d.cnki.gtsfu.2022.000384>
- [35] 何虎, 李杰, 侯友, 包呼格吉乐图, 田颖, 刘敏. 文化差异与融合: 蒙古族大学生场景知觉的眼动研究[J]. 心理与行为研究, 2018, 16(6): 751-756.
- [36] 周肃军. 刺激物数量及刺激物之间不同关系对场景识别的影响[D]:[硕士学位论文]. 兰州: 西北师范大学, 2018.
- [37] Oliva, A. and Torralba, A. (2006) Chapter 2-Building the Gist of a Scene: The Role of Global Image Features in Recognition. *Progress in Brain Research: Visual Perception*, **155**, 23-36. [https://doi.org/10.1016/S0079-6123\(06\)55002-2](https://doi.org/10.1016/S0079-6123(06)55002-2)
- [38] Hollingworth, A. (2007) Object-Position Binding in Visual Memory for Natural Scenes and Object Arrays. *Journal of Experimental Psychology: Human Perception and Performance*, **33**, 31-47. <https://doi.org/10.1037/0096-1523.33.1.31>
- [39] Brockmole, J.R. and Henderson, J.M. (2008) Prioritizing New Objects for Eye Fixation in Real-World Scenes: Effects of Object-Scene Consistency. *Visual Cognition*, **16**, 375-390. <https://doi.org/10.1080/13506280701453623>
- [40] Mannan, S.K, Ruddock, K.H. and Wooding, D.S. (1997) Fixation Sequences Made during Visual Examination of Briefly Presented 2D Images. *Spatial Vision*, **11**, 157-178. <https://doi.org/10.1163/156856897X00177>
- [41] Einhaeuser, W., Mundhenk, T.N. and Baldi, P.F. (2007) A Bottom-Up Model of Spatialattention Predicts Human Error Patterns in Rapid Scene Recognition. *Journal of Vision*, **7**, 1-13. <https://doi.org/10.1167/7.10.6>