

复杂信息界面视觉搜索效能量化评估方法与模型

朱 奥¹, 王晓彬¹, 李晓鹤¹, 丁 新², 王家润², 张 丹^{1*}

(1. 清华大学心理学系, 北京, 100084; 2. 中国电子科技集团公司第十五研究所, 北京, 100080)

摘要 未来战争具有高复杂性, 来自战场复杂信息界面的大量视觉信息输入对决策人员带来巨大挑战。立足人类视觉信息感知特点和视觉显著性机制相关理论, 提出以色调、饱和度、明度、图像梯度等关键视觉信息感知特征为自变量的评估方法, 采用复杂影像地图下的目标识别任务, 建立以视觉搜索响应时为指标的视觉搜索效能量化评估模型。实验结果表明, 基于3种地貌(城市、海洋、高原)下不同对比度和亮度的441张影像地图实验数据生成的量化模型, 可以良好地拟合被试的行为响应时($R^2 = 0.232, p < 0.001$)。

关键词 视觉显著性; 视觉搜索; 视觉信息感知; 量化评估模型; 复杂信息界面

DOI 10.3969/j.issn.2097-1915.2022.06.012

中图分类号 TN974 **文献标志码** A **文章编号** 2097-1915(2022)06-0084-08

A Quantitative Evaluation Model of Visual Search Effectiveness for Complex Information Interface

ZHU Ao¹, WANG Xiaobin¹, LI Xiaohe¹, DING Xin², WANG Jiarun², ZHANG Dan^{1*}

(1. Department of Psychology, Tsinghua University, Beijing 100084, China;

2. The North China Institute of Computing Technology, Beijing 100080, China)

Abstract Future combats being extremely complex, and challenge to decision-makers taken by significant cognitive overload due to the massive input of visual information, an evaluation method with key visual information characteristics, i. e. hue, saturation, value and image gradient as independent variables is proposed by the relevant theories of human visual perception and visual saliency mechanism. A quantitative model of visual search efficiency is constructed via a target search task over a complicated image map, and the visual search response time is used as a dependent variable. Three landforms (city, ocean, and plateau) consisting of 441 image maps with different contrast and brightness are tested in the psychology experiment, and the model result show that the R^2 is equal to 0.232, $p < 0.001$. And the results have practical guiding evaluation and optimization to complex interfaces under different scenarios.

Key words visual saliency; visual search; visual perception; quantitative evaluation model; complex information interface

收稿日期: 2022-07-08

基金项目: 清华大学春风基金(2021Z99CFY037)

作者简介: 朱 奥(1993—), 男, 四川峨嵋山人, 博士生, 研究方向为工程心理学, 人机协作与互信。E-mail: za20@mails.tsinghua.edu.cn

通信作者: 张 丹(1983—), 男, 重庆人, 副教授, 研究方向为心理学。E-mail: dzhang@tsinghua.edu.cn

引用格式: 朱奥, 王晓彬, 李晓鹤, 等. 复杂信息界面视觉搜索效能量化评估方法与模型[J]. 空军工程大学学报, 2022, 23(6): 84-90. ZHU Ao, WANG Xiaobin, LI Xiaohe, et al. A Quantitative Evaluation Model of Visual Search Effectiveness for Complex Information Interface[J]. Journal of Air Force Engineering University, 2022, 23(6): 84-90.

现代化战争中的海量战场信息通过可视化的方式呈现,战场态势感知界面也趋于复杂,其承载的信息容量已经严重超过了决策人员的认知负荷^[1],如何提升复杂界面下的信息搜索效能成为需要解决的关键问题^[2]。面对复杂战场视觉信息情境,人类视觉系统通过初级处理过程筛选接收到的信息,进而通过中高级处理过程进一步加工,这被称为选择性注意机制,其中视觉显著性(visualsealiency)起到关键作用。

视觉显著性指场景中某个特定突出的物体在预注视处理阶段得到注意,并在视觉系统处理流程的早期阶段引起即时视觉唤醒。视觉显著性属于自下而上的视觉注意机制,可以粗略描述为场景中的元素吸引个体视觉注意力的能力,但本质上是物体带来的视觉刺激与人工视觉系统之间相互作用的结果^[3]。在较早的研究中,有研究者用视觉“凸显”(pop-out)形容这一过程,并探索了能够引起视觉“凸显”的视觉刺激特征^[4]。有研究者认为物体在某些特征上的特殊性会让它“凸显”,比如 Julesz 提出的纹理理论(texton theory)认为主要是纹理上的差异会引起我们更多的视觉注意^[5]。文献^[6]认为有大量研究证据可以表明大脑中存在一种注意力转移“地图”模型:在每个地图中,视觉场景中所有点的显著性都是并行编码的,通过每个地图内的竞争和地图合并的过程选择下一个注意目标,其中一种类型的地图是“显著性地图”,它根据低层次的视觉特征(如明度、颜色、定向边缘和运动)计算视觉上的显著点^[6]。越显著的物理刺激,越容易被注意选择,即发生注意捕获;但也有研究者认为,尽管显著性可能有正面作用但也可能阻碍正确知觉,这取决于使用的环境、场景、任务等。

受显著性地图启发,具备良好可应用性的视觉显著性模型应该具备以下特点:①模型输入参数应贴近人眼感知能力,同时人类对视觉显著目标的反应过程是快速的,所以模型也需要具备实时性;②模型应从低层视觉特征出发,同时也要防止过度拟合,应用尽量少的参数覆盖低层视觉特征;③模型不仅考虑视觉场景中前景因素(目标区域),也需要考虑背景因素;④最重要的一点,模型应能直接反映人眼搜索效能,如与响应时等指标直接关联。

1 视觉界面信息的量化

随着视觉显著性理论的积累,研究者在视觉显著性的量化计算方法方面不断取得重要进展,在20世纪末 Itti 基于 Koch 等人的视觉理论模型提出 ITTI 算法,其思路是在多尺度空间中以颜色、强度

和方向特征进行显著性检测,然后根据各组特征图像中心-周围像素差异性获取各尺度特征图,并进行合并和归一化处理得到最终结果;其优势是可以高效地计算出最具显著性的第一个目标,但在后续过程中需要在多个空域基本特征上进行操作,导致算法运算速度较慢,难以满足实时检测要求,以及对后续目标区域计算不精确^[7-8]。

针对 ITTI 方法的不足,Achanta 等提出了全分辨率算法(AC 算法),该算法强调快捷性,仅考虑输入图像的颜色和亮度特征,而没有考虑梯度等特征^[9]。Cheng 等人提出了基于色彩直方图(HC 算法)和局部对比度(RC 算法)的图像视觉显著性计算方法^[10]。其中,RC 算法考虑了空间关系在人类视觉注意方面的显著影响,将输入图像分为若干子区域,然后计算每个区域的颜色对比度与其他区域对比度加权和以定义其显著值,加权的权值由区域空间距离决定,距离与权值成正比,从而使与相邻区域存在高对比度关系的区域能获得更高的显著值。

研究者进一步通过分析频谱找到其与显著性特征的关系,把图像从空间域变换到频率域,从而更高效地完成显著性检测,比如 Hou 和 Zhang 提出的谱残余算法(SR 算法)方法^[11]。但因为相关方法缺乏生理基础的支持,所以本文不深入讨论^[12]。

在显著性计算技术发展的同时,研究者们也将其运用于预测人类视觉搜索过程,比如 Foulsham 和 Underwood 使用 Itti 的显著性模型预测人眼动态轨迹,发现预测准确率高于随机模型^[13];Ehinger 等人使用显著性等模型在眼动实验中预测人类注意力变化,最好的模型的 AUC 能达到 0.9^[14];还有研究者使用显著性模型用于预测视盲变化的检测^[15]。但现有的主流图像评估模式存在以下不足:

1)当前主流技术主要从计算机角度出发采用 RGB 等模型,忽视了人眼的感觉能力和对图像纹理的识别能力。在 HSV 空间中,H 表示一幅图像的色调,S 表示图像的饱和度,V 表示图像的明度。相对于传统 RGB 模型,HSV 模型更接近人类视觉系统的感知。RGB 彩色空间在综合的色彩呈现上具有优势,因而被广泛应用于数字图像的合成中,但不能直观地表现图像的明度和色彩饱和度等人眼可直观获取的信息^[16]。

2)人眼对于梯度信息十分敏感,图像梯度描述的是图像局部灰度变化的快慢与方向,是区分不同物体的最重要特征^[17]。图像像素与像素之间的变化程度能够自下而上的影响个体对图像信息的搜索,梯度特征能够传递重要的视觉信息,具有强大地捕捉结构信息和对对比度变化的能力,因此被广泛用

于图像质量评价等任务中。梯度特征是人眼进行识别的重要依据,而很多模型忽略了这一特性^[9,16]。

3)复杂信息搜索过程中不仅受目标属性影响,也受背景属性的影响,而部分相关技术仅考虑了单方面因素,比如 RC 算法过分放大局部因素的影响。

4)现有模型较难以直接应用^[3]。目前大部分相关技术方案是为了利用计算机来模拟人类视觉,而没有联系到人类的真实行为(如响应时等),面向应用的外部效度有限。已有的少量结合人类真实行为的研究,比如 Hou 等人的机器视觉模型对比研究^[18],仅考虑了视觉注视点、变化盲视等相对经典的心理学实验范式下的人类行为数据,更多面向理论机制探索,相关研究结论不能直接迁移到有更明确应用价值的复杂信息视觉搜索任务中。

根据国内外研究文献综述,我们结合颜色空间和图像梯度来对图像进行量化处理,并建立其与人类响应时的联系。在颜色空间方面,采用更符合人眼感知的 HSV 模型。在图像梯度方面,目前主流的图像梯度计算方法包括 Canny、LoG 等,其中 LoG(Laplacian of Gaussian)边缘检测算子是 Marr 和 Hildreth 共同提出的经典方法,也称为 Marr & Hildreth 算子,它根据图像的信噪比来求检测边缘的最优滤波器^[19]。相较于 Canny 的梯度极大值方法,LoG 引入二阶导数的 Laplace 算子,并在此之前为图像进行了高斯卷积滤波,消除高频噪声的干扰;由于对高频噪声的处理与引入二阶导数以计算梯度变化的快慢,LoG 被认为是较为接近人眼视觉生理的数学模型^[19],所以本研究将采用 LoG 方法进行图像梯度量化。

综上所述,我们将从人类心理学和认知生理学的角度建立复杂信息下视觉搜索模型,为复杂战场态势认知提供模型基础支撑。

2 本文方法

2.1 模型框架

基于上述讨论,我们基于视觉显著性计算技术中经典的 Itti 模型^[6,8],突破前人研究中眼动态轨迹等指标与行为绩效关系不够紧密的局限^[13-14],创新地提出以面向复杂信息界面视觉搜索行为响应时为因变量的多元模型。具体而言,我们构建以影像地图为代表的复杂信息界面,让被试开展影像地图中的特定图标视觉搜索,以视觉搜索的行为响应时为目标,通过心理学行为实验,建立视觉搜索效能的量化评估模型。模型假设如下:

搜索响应时 =

$\beta_1 \times$ 目标区域色调 + $\beta_2 \times$ 目标区域饱和度 +

$\beta_3 \times$ 目标区域明度 + $\beta_4 \times$ 目标区域图像梯度 +

$\beta_5 \times$ 背景色调 + $\beta_6 \times$ 背景饱和度 +

$\beta_7 \times$ 背景明度 + $\beta_8 \times$ 背景图像梯度 + β_0 。

模型中, $\beta_i (i = 1, 2, \dots, 8)$ 为使用多元线性模型与最小二乘法拟合后各变量的回归系数,它的大小与正负表示在该模型下此变量对搜索响应时的贡献大小与方向, β_0 表示回归方程的常数偏置量。整体模型构建的流程见图 1。

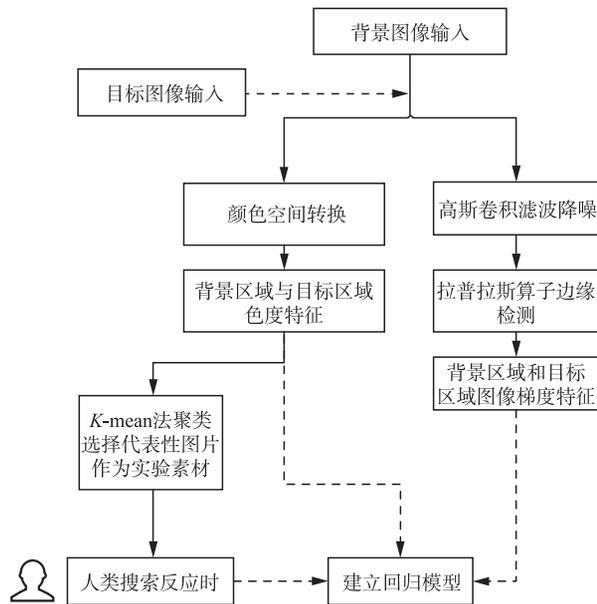


图 1 模型建立流程图

2.2 显著性特征提取

本文所提取的复杂信息界面颜色特征基于 HSV 模型框架。考虑到实际图像数据往往以 RGB 形式存储,色调(H)、饱和度(S)和明度(V)与 RGB 的转换公式为:

$$H = \begin{cases} \frac{60(G-B)}{S}, V=R \\ 180 - \frac{60(B-G)}{S}, V=G \\ 240 + \frac{60(R-G)}{S}, V=R \end{cases} \quad (1)$$

$$S = \begin{cases} \frac{(V - \min(R, G, B)) \times 255}{V}, V \neq 0 \\ 0 \end{cases} \quad (2)$$

$$V = \max(R, G, B) \quad (3)$$

LoG 方法见式(4),先通过平滑预处理, $G(x, y)$ 表示标准高斯函数, σ 表示标准差:

$$\text{LoG} = \Delta G_\sigma(x, y) = \frac{\partial^2 G_\sigma(x, y)}{\partial x^2} + \frac{\partial^2 G_\sigma(x, y)}{\partial y^2} = \frac{x^2 + y^2 - 2\sigma^2}{\sigma^4} e^{-(x^2 + y^2)/2\sigma^2} \quad (4)$$

在图像输入后,将背景和背景区域分为 RGB3

通道并依次计算所有像素点的 H、S、V 值,并求其平均值得到图像背景和区域 H、S、V 均值;之后,利用 LoG 方法计算背景和区域区域的平均梯度。

2.3 视觉界面参数化生成

本研究视觉界面所用影像地图包括来自城市、海洋、高原地貌的卫星影像素材图片各一张,图片尺寸为 $2\ 100 \times 2\ 100$ 像素。标识素材图片为 14 个军事地图常用标识图片,分为红蓝两色,共 28 个标识,图标尺寸为 160×160 像素。定义目标区域即图标覆盖在背景图上的区域,见图 2。

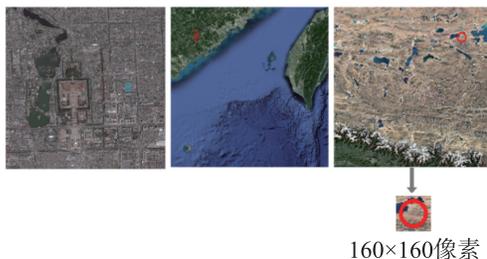


图 2 背景图片与目标

为了提升所考察的视觉界面丰富性和全面性,本研究通过点操作法对选定的影像地图进行亮度与对比度的量化变换^[20]。具体来说,图像亮度表示色彩本身因为光度、反射程度不同而产生的明暗差别。对基于 RGB 色彩空间的数字图像而言,R、G、B 3 个通道均被定义为 $0 \sim 255$ 的一个值,通过同时增加或减少 3 个通道的值即可调整亮度。图像的对比度用来描述图像不同像素之间属性的差异程度,对比度的增加,表示图像中亮处越亮、暗处越暗。在点操作法中,常人为地设置一个阈值来实现图片明暗分明的高对比度的效果^[20]。本文基于以下公式对原始图像进行变换:

$$L = L_0 + (L_0 - 127) \frac{(i_{\text{Con}} - 50)}{50} + 255 \frac{(i_{\text{Brt}} - 50)}{50} \quad (5)$$

式中: L_0 为原图像当前像素点在 RGB 通道下的值; i_{Con} 、 i_{Brt} 表示图像对比度、亮度的调整值。在本研究中,将原图像的对比度与亮度值的原始值均定义为 50,通过调整 i_{Con} 与 i_{Brt} ,对应地对图像的 3 个通道同时进行线性拉伸,得到当前像素点的值 L 。式中,系数 -127 表示本实验中将对比度的阈值定义为 0.5(即 0.5×255)。如式(5)所示,图像的点操作法基于图像的 RGB 空间定义图像的亮度与对比度,由于其在数字图像处理中较为方便的优势,适合规模化操作,本研究采用该图像特征调整方案。

对 3 个地貌的环境影像地图的对比度与亮度在普遍适用可视化参数范围内进行逐步调整。本研究

将普遍适用定义为基于经验的,每张影像地图按不同亮度(i_{Brt} 在 $50 \sim 70$ 范围)与对比度(i_{Con} 在 $40 \sim 60$ 范围)生成 441 张图片作为背景图片素材,在该参数范围下,图片不会因过暗、过亮或对比度差异过大导致图片失真。同时,为了尽可能提升计算效率,对目标位置采取了随机取样策略以减少取样总数:将每张图片均等划分为 $5 \times 5 = 25$ 个区域(420×420 像素),并在每个区域中随机取 3 个位置,将 28 个作为视觉目标的军事地图常用标识依次置入抽取中的位置并计算此时图标区域的 HSV 值与图像梯度数据,共得到 $441(\text{参数化影像地图}) \times 25(\text{区域/参数化影像地图}) \times 3(\text{目标位点/区域}) \times 28(\text{图标}) = 926\ 100$ 个不同复杂信息界面条件。进一步采用 K 均值(K-means)聚类方法,得到 24 个有代表性的复杂信息界面条件聚类。从每个聚类中随机抽取 50 个具体复杂信息界面条件,形成包含 $24(\text{聚类}) \times 50 = 1\ 200$ 个复杂信息界面条件(其中以城市、海洋、高原为背景的界面条件数分别为 385、452、363)用于后续行为学实验。K-means 算法所得聚类结果使用 t-SNE 算法映射到低维空间的可视化展示及代表性聚类结果见图 3。

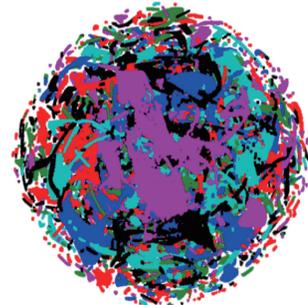


图 3 训练集聚类图

根据 t-SNE 图可见,这样基于特定影像地图参数化生成的视觉界面较为分布均匀、特征丰富,可以有效表征我们在现实生活中所感知到的视觉界面情况,同时便于有效控制视觉界面复杂性带来的可能额外影响,有利于构建稳健的量化模型。

2.4 实验材料与设备

实验设备包括实验电脑一台(用于实验程序执行和数据收集),型号为联想 ThinkPad S1(4th Gen)笔记本电脑;电脑屏幕为华硕 ROGPG248Q 型号 24 英寸显示器,屏幕刷新率 180 Hz,亮度 350 cd/m^2 。实验程序为 PEBL Version 2.1^[21]。

2.5 实验流程

实验采用经典的视觉搜索范式任务,要求被试在不同复杂信息界面中进行视觉目标(图标)搜索。视觉搜索任务包括上述聚类所得 1 200 个复杂信息

界面条件,并按照随机的方式呈现;单个试次中,被试被要求尽可能准确且快速地找到视觉目标,当发现并认清目标后按 F 键报告,从复杂信息界面图像出现到被试按下 F 键的时间记为该试次的响应时;之后被试需从 4 个不同图标中选择刚才看到的图标,以确认被试认清目标而非虚报。选择后程序会反馈被试的选择结果正误,并在 500 ms 间隔时间后进入到下一个试次,单试次的实验流程见图 4。

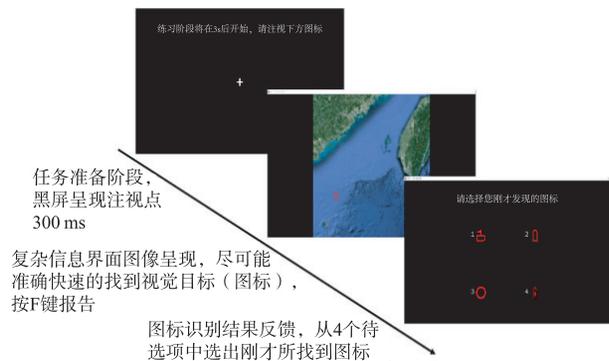


图 4 单试次实验流程图

在遵循自愿原则了解实验内容并签署知情同意书、登记基本个人信息后,被试首先阅读实验说明详细了解实验任务内容,并通过 10 个试次的练习训练以确认熟悉了解实验任务执行方式,进入正式实验后,被试总共需要完成 1 200 试次实验任务。其中,每完成 200 个试次后有 20 s 的额外休息时间。完成全部试次后实验结束,总时长约 50 min。

2.6 被试

本研究总共招募了 13 名大学生被试,排除一名被试超时次数情况过多(超过 10%)后,有效数据被试为 5 名男性,7 名女性,平均年龄 24.00 岁($SD = 0.72$ 岁)。所有被试均报告无色盲、色弱等情况,均为右利手,且实验前 24 小时内无饮酒或兴奋饮料的行为。12 名被试均完成全部 1 200 个试次,共获得 14 400 条数据;排除 3 个标准差外的异常值和图标识别结果错误的试次数据后,剩余 13 688 条,有效数据占比为 95.06%。

2.7 数据分析

对于实验中所用 1 200 个复杂信息界面条件,我们基于 2.1 节中的模型框架计算得到 1 200 组关键视觉信息感知特征(背景图像 H、S、V 均值与图像梯度、目标区域背景 H、S、V 均值与图像梯度,总共 8 个特征变量)作为自变量,以 12 名被试在相应条件下的行为响应时平均值为因变量,构建多元线性回归模型,实现对某给定复杂信息界面条件下视觉搜索效能的可靠群体趋势估计。通过实验数据拟合该回归模型,预期实现根据客观视觉信息的视觉搜索效能的量化

估计,回归模型的拟合优度 R^2 数值及其统计显著性可以有效表达模型的量化估计能力。

为进一步评估所得模型的外部效度,本研究采用留一法交叉验证(leave-one-out cross-validation)。具体来说,对于每一个复杂信息界面条件,基于剩余 1 199 个条件的关键视觉信息感知特征和对应的行为学数据构建回归模型,计算该条件下关键视觉信息感知特征的模型预测响应时并与实际实验所得响应时进行对比。遍历所有 1 200 个条件后,根据交叉验证所得模型预测结果与实际响应时得到交叉验证的 R^2 数值及其相关系数统计显著性,实现对模型推广应用能力的估计^[22]。

在此基础上,为探索所得模型跨地貌环境信息的预测性能,运用上述方法分别对 3 种地貌下的数据开展回归分析,所用数据量分别对应为城市、海洋、高原为背景界面条件数(385、452、363)。同时包括 3 种地貌影像地图的综合模型与分别独立运用单一地貌数据的模型结果对比可以为深入评估综合模型的性能提供依据。

最后,为避免上述关键视觉信息感知特征的共线性问题,上述模型实际所用因变量为关键视觉信息感知特征的主成分,本研究提取了 8 个关键视觉信息感知特征变量集合所得的 5 个主成分作为因变量,这 5 个主成分可表达原始数据 98.14% 的方差。

3 实验结果

对 8 个关键视觉信息感知特征变量进行主成分分析,提取得到的 5 个主成分($F_1 \sim F_5$)的表达式如下所示,其中 T 代表目标区域, B 代表背景区域; h 、 s 、 v 、 g 代表该色调、饱和度、明度和图像梯度的均值:

$$F_1 = 0.189T_h - 0.175T_s + 0.025T_v + 0.166T_g + 0.192B_h - 0.175B_s + 0.057B_v + 0.191B_g \quad (7)$$

$$F_2 = -0.038T_h + 0.167T_s + 0.467T_v + 0.070T_g - 0.042B_h + 0.171B_s + 0.475B_v + 0.130B_g \quad (8)$$

$$F_3 = 0.073T_h + 0.206T_s - 0.710T_v + 0.920T_g - 0.189B_h + 0.510B_s + 0.301B_v - 0.023B_g \quad (9)$$

$$F_4 = 0.846T_h + 1.085T_s + 0.167T_v - 0.134T_g + 0.709B_h + 0.599B_s - 0.692B_v + 0.296B_g \quad (10)$$

$$F_5 = -1.087T_h + 0.655T_s - 1.204T_v - 0.902T_g + 1.034B_h - 0.251B_s + 0.884B_v + 1.079B_g \quad (11)$$

根据各主成分系数可以观察到: F_1 中目标区域的图像梯度较高的载荷,而目标与背景区域的明度载荷较低,其可能主要反映的是图像梯度均值; F_2 中目标与背景的明度均值载荷较高,而色调均值占比较低,其可能主要反映明度均值; F_3 中目标与背

景饱和度载荷都较高,同时目标的明度与图像梯度载荷较高,而背景的图像梯度载荷较低,其主要反映目标区域相关物理属性; F_4 中目标与背景色调和饱和度占比较高,其主要反映的是色调与饱和度均值; F_5 体现的是除饱和度外的全局物理属性。

基于上述主成分做回归分析,结果见表1。

表1 主成分回归分析结果

	B	SE	Beta	t	p
常量	529.913	2.641		200.645	<0.01
F_1	27.350	2.642	0.262	10.351	<0.01
F_2	27.257	2.642	0.261	10.316	<0.01
F_3	17.570	2.642	0.168	6.650	<0.01
F_4	-17.134	2.642	-0.164	-6.485	<0.01
F_5	21.556	2.642	0.207	8.158	<0.01

多元线性回归分析结果显示 R^2 为 0.232, $F(5, 1194) = 73.281$, $p < 0.001$; 留一法交叉验证结果 $R^2 = 0.230$, $p < 0.001$ 。留一法结果可视化见图5,图中虚线表示 $y = x$ 的参考线,模型 R^2 数值对应的相关系数 $R = 0.480$ 。可以看到模型对实际响应时的总体拟合情况良好,模型预测响应时相比实际响应时略偏保守,特别是在实际响应时较大时。根据 $F_1 \sim F_5$ 表达式,还原系数为原始系数得方程为:

$$Y = 504.682 - 0.915T_h - 0.027T_s - 0.997T_v + 0.115T_g + 0.281B_h - 0.155B_s + 2.358B_v + 0.520B_g \quad (12)$$

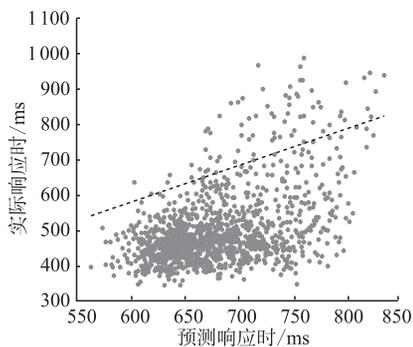


图5 模型拟合情况

基于所有地貌数据所得主成分对3种地貌进行回归分析,结果显示:城市地貌模型 R^2 为 0.205, $F(5, 379) = 20.727$, $p < 0.001$; 海洋地貌模型 R^2 为 0.140, $F(5, 446) = 15.705$, $p < 0.001$; 高原地貌模型 R^2 为 0.548, $F(5, 357) = 88.888$, $p < 0.001$ 。考虑到主成分分析对数据的可能依赖性,本研究还基于单一地貌数据分别进行了主成分分析,基于单一地貌主成分的回归分析结果显示:城市地貌模型 R^2 为 0.137, $F(5, 379) = 13.149$, $p < 0.001$; 海洋地貌模型 R^2 为 0.077, $F(5, 446)$

$= 8.546$, $p < 0.001$; 高原地貌模型 R^2 为 0.506, $F(5, 357) = 75.234$, $p < 0.001$ 。

4 讨论分析

HSV模型方面来看,首先目标区域的色调均值系数均为负数,而背景区域的色调、明度均值系数为正数,这说明目标与背景区域的色调和明度均值差异越大搜索响应时越短,这与前人的研究中认为颜色差异度是显著性计算中重要因素相呼应^[8,23]。其次,目标区域与背景区域的饱和度均值系数均为负数,但绝对值较小,说明在一定范围内人眼对高饱和度的图像差异更敏锐。图像梯度方面,本研究发现目标区域和背景区域的图像梯度均值系数均为正数,也就是说无论目标还是背景区域的图像越复杂,就越不利于视觉搜索,由于前人研究多数关注颜色特征而忽视梯度特征^[3],本研究结果拓展了相关研究发现。

在应用性方面,本研究提出了基于视觉显著性量化计算方案的应用思路,使其不仅可用于视觉搜索效能的评估,还可以作为优化视觉界面参数的工具。前人研究侧重于理论摸索,比如预测人眼在真实环境中的观察能力,尽管部分研究已经达到较高的拟合水平,但距离实际运用较远^[15-16];本研究依托反应时指标进行搜索效能预测,具有明确的应用价值。在视觉搜索效能评估方面,根据已有的模型公式预测不同目标物在每张图像不同位置时的响应时,并通过计算平均值得到评估此图像人眼搜索能力的图像视觉搜索效能;在视觉界面参数优化方面,通过对不同对比度与亮度组合下的视觉界面参数进行模型预测可以得到多个组合上的图像视觉搜索效能,从而评估出能达到最优视觉搜索效能的对比度与亮度组合并应用于当前图像。

值得指出的是,本文基于所有地貌数据所得综合模型结果($R^2 = 0.232$)表现出较好的跨地貌推广应用潜力。综合模型结果优于基于城市和海洋单一地貌所得模型结果($R^2 = 0.137, 0.077$),低于高原单一地貌模型结果($R^2 = 0.506$)。

5 结语

本研究面向复杂信息界面视觉搜索效能的量化评估应用需求,基于心理学理论与实验方法构建了良好拟合行为响应时的量化模型。探讨了HSV颜色空间和图像梯度对人眼视觉搜索的影响,验证了构建视觉效能模型的可行性。在方法上,本研究结合了心理学与信息学方法,基于图像处理的

方法建立了训练数据集,通过行为学实验的方式,采用视觉搜索任务的范式采集了人眼搜索数据并建立模型。研究发现,模型预测拟合结果良好,拟合优度 R^2 为 0.232。使用整体主成分对不同地貌进行预测的结果也优于 3 种地貌独立主成分预测的结果,这说明了根据符合人类视觉认知的 HSV 颜色空间和图像梯度建立的模型能在一定程度上有效预测搜索响应时,并具有一定的适用性,证明了基于人类视觉加工机理去构建人机协同方案的可行性与有效性。本研究成果不仅可应用于军事背景下的视觉搜索,也可以应用到其他民用场景中,如界面交互设计、复杂网页设计等,通过对特定应用情景界面视觉信息特征的定量计算和相关心理学行为数据分析,即可建立特定情境视觉搜索效能的量化评估模型。由于高原地貌相对视觉信息复杂度较低,本文所得综合模型结果优于复杂度较高的城市和海洋地貌结果说明,综合模型具有较好的跨地貌应用价值。将来可以考虑引入更多地貌影像数据开展更大规模的心理学实验,探索更具有普遍性应用价值的视觉搜索效能量化模型。

参考文献

- [1] 丁畅. 复杂海况环境下海面图像增强方法研究[D]. 大连:大连海事大学, 2018.
- [2] 贾奋励, 王光霞, 田江鹏, 等. 面向全域作战态势感知的战场全息地图[J]. 指挥与控制学报, 2022, 8(1): 30-36.
- [3] 敖欢欢. 视觉显著性应用研究[D]. 合肥:中国科学技术大学, 2013.
- [4] TREISMAN A. Preattentive Processing in Vision[J]. Computer Vision, Graphics, and Image Processing, 1985, 31(2): 156-177.
- [5] JULESZ B. Texton gradients: The Texton Theory Revisited[J]. Biological Cybernetics, 1986, 54(4-5): 245-251.
- [6] ITTI L, KOCH C. Computational Modelling of Visual Attention [J]. Nature Reviews Neuroscience, 2001, 2(3): 194-203.
- [7] 鲁海桢, 孙永荣, 赵伟, 等. 基于视觉显著性的机场跑道异物检测技术[J]. 现代电子技术, 2022, 45(11): 41-45.
- [8] ITTI L, KOCH C, NIENUR E. A Model of Saliency-Based Visual Attention for Rapid Scene Analysis [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(11): 1254-1259.
- [9] ACHANTA R, ESTRADA F, WILS P et al. Salient Region Detection and Segmentation[C]//International Conference on Computer Vision Systems. [S. l.]: Springer, 2008: 66-75.
- [10] CHENG M, ZHANG G X, MITRA N J et al. Global Contrast Based Salient Region Detection[J]. Computer Vision and Pattern Recognition, 2011, 37(3): 569-582.
- [11] HOU X, ZHANG L. Saliency Detection: A Spectral Residual Approach[C]//IEEE Conference on Computer Vision & Pattern Recognition. Minneapolis, MN:IEEE, 2007:1-8.
- [12] 孙晓飞, 潘文文, 王霞. 典型的图像显著性检测算法分析和比较[J]. 黔南民族师范学院学报, 2014, 34(5): 100-105.
- [13] FOULSHAM T, UNDERWOOD G. What Can Saliency Models Predict About Eye Movements? Spatial and Sequential Aspects of Fixations During Encoding and Recognition[J]. Journal of Vision, 2008, 8(2): 6.
- [14] EHINGER K A, HIDALOGO-SOTELO B, TORRALBA A et al. Modelling Search for People in 900 Scenes: A Combined Source Model of Eye Guidance [J]. Visual Cognition, 2009, 17(6-7): 945-978.
- [15] STIRK J A, UNDERWOOD G. Low-level Visual Saliency Does Not Predict Change Detection in Natural Scenes[J]. Journal of Vision, 2007, 7(10): 3-3.
- [16] LEE Y J, SONG J B. Autonomous Salient Visual Feature Detection Using Salient Points from Both RGB and HSV Color Spaces for Visual SLAM in Indoor Environments[C]//The 5th International Conference on Ubiquitous Robots and Ambient Intelligence. Korea:Korea Robotics Society, 2008:631-636.
- [17] 邓杰航, 袁仲鸣, 刘栋濠, 等. 基于视觉显著性的自适应图像质量评价方法[J]. 软件导刊, 2021, 20(11): 191-196.
- [18] HOU X, HAREL J, KOCH C. Image Signature: Highlighting Sparse Salient Regions[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 34(1): 194-201.
- [19] MARR D, HILDRETH E. Theory of Edge Detection [J]. Proceedings of the Royal Society of London. Series B. Biological Sciences, 1980, 207(1167): 187-217.
- [20] BURGER W, BURGER M. Digital Image Processing. Texts in Computer Science[Z]. Springer, 2008.
- [21] MUELLER S T, PIPER B J. The Psychology Experiment Building Language (PEBL) and PEBL Test Battery[J]. Journal of Neuroscience Methods, 2014, 222: 250-259.
- [22] WONG T T. Performance Evaluation of Classification Algorithms by K-fold and Leave-one-out Cross Validation [J]. Pattern Recognition, 2015, 48(9): 2839-2846.
- [23] HAREL J, KOCH C, PERONA P. Graph-based Visual Saliency[J]. Advances in Neural Information Processing Systems, 2006, 19:545-552.

(编辑:徐敏)