

融合眼动数据和文本信息的可视化阅读辅助*

程时伟, 胡屹凛, 孙煜杰



(浙江工业大学 计算机科学与技术学院, 浙江 杭州 310023)

通讯作者: 程时伟, E-mail: swc@zjut.edu.cn

摘要: 研究阅读过程中的视觉注意行为特征, 基于眼动数据和文本主题进行阅读行为特征和文档结构的分析, 设计了眼动热区图、圆环图、节点链接图、词云等可视化形式. 在此基础上开发了面向阅读辅助的可视分析原型系统, 该系统记录专家用户(如老师)的眼动数据, 然后通过可视化形式分享给新手用户(如学生). 用户实验结果表明, 实验组用户的阅读理解客观题和主观题得分平均值分别比对照组用户提高了 31.8% 和 55.0%, 阅读和答题总用时比对照组用户平均减少了 9.7%, 可以有效帮助读者提高阅读效率、快速抓住文章重点和更好地理解文章内容, 具有一定的有效性和可行性.

关键词: 眼动跟踪; 眼动数据; 可视分析; 热区图; 人机交互

中文引用格式: 程时伟, 胡屹凛, 孙煜杰. 融合眼动数据和文本信息的可视化阅读辅助. 软件学报, 2019, 30(Suppl.(12)):35-47. <http://www.jos.org.cn/1000-9825/19016.htm>

英文引用格式: Cheng SW, Hu YL, Sun YJ. Eye Tracking Data and Text Visual Analytics for Reading Assistance. Ruan Jian Xue Bao/Journal of Software, 2019, 30(Suppl.(12)):35-47 (in Chinese). <http://www.jos.org.cn/1000-9825/19016.htm>

Eye Tracking Data and Text Visual Analytics for Reading Assistance

CHENG Shi-Wei, HU Yi-Lin, SUN Yu-Jie

(School of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310023, China)

Abstract: This paper studied visual attention behavior characteristics during reading process. Several visualization such as eye movement heatmap, doughnut chart, node link graph, and word cloud were designed to extract eye movement data and text themes for analysis of reading behavior characteristics and document structure. A visual aid prototype system for reading assistance was developed to record the eye movement data of expert users (such as teachers), and the visualization can be shared to novice users (such as students). The user study results showed that the average scores of objective and subjective questions in the experimental group were increased by 31.8% and 55.0%, respectively, and the total reading and answering time was reduced by 9.7%. It can be seen that this system can effectively help readers improve reading efficiency as well as quickly grasp the focus of the article and better understand the content of the article, so that it has certain effectiveness and feasibility.

Key words: eye tracking; eye movement data; visual analytics; heat map; human-computer interaction

如何花费较少的时间快速把握阅读的重点、理解全文结构和段落关系, 进一步提高阅读效率和理解水平, 已成为一个重要的研究方向. 本文获取专家用户阅读时的眼动数据, 融合眼动跟踪数据与文本特征, 分析阅读行为及其相关的视觉注意力分布, 并将专家用户的眼动数据以可视化形式分享给新手用户, 例如, 通过热区图等让新手用户了解阅读重点, 使其直观、迅速地掌握文章重点内容和内在结构; 此外, 新手用户还可根据节点链接图了解专家用户在不同部分之间的眼动注视点转换, 帮助他们学习专家用户的阅读模式.

* 基金项目: 国家自然科学基金(61772468); 浙江省属高校基本科研业务费专项资金(RF-B2019001)

Foundation item: National Natural Science Foundation of China (61772468); Fundamental Research Funds for the Provincial Universities of Zhejiang (RF-B2019001)

收稿时间: 2019-07-15; 采用时间: 2019-11-04

1 相关研究

由于眼动跟踪技术具有对被试干扰小、使用方便、数据易于获取等优点,基于眼动跟踪技术的阅读研究已经成为研究热点之一.例如,Biedert 等人^[1]开发的 Text2.0 系统可以实时进行阅读困难检测,并以批注的形式显示该单词的翻译,以此来降低阅读难度.此外,根据对单词的注视时长生成个性化的文档摘要,还可以帮助用户对文档中的关键信息进行有效和快速的定位,但是传统的自动文档摘要技术很少将人的因素考虑在内.为此,Xu 等人^[2]提出了一种个性化文档摘要算法,它利用眼动跟踪技术记录并计算用户阅读时在每个单词上所花费的时间,以此为基础来抽取文档关键词,再结合机器学习的方法自动生成文档摘要.这样生成的文档摘要可以进一步满足用户的个性化需求.另一方面,分析阅读时的眼动数据可以检测和评估用户的语言技能和阅读理解水平^[3].例如,Augereau 等人^[4]通过提取眼动特征来估计阅读答题得分,评估用户的阅读理解水平.此外,还可以使用眼动跟踪系统来帮助用户提高阅读理解水平.例如,Cheng 等人^[5]研究了人们阅读时的眼动行为,将专家用户的阅读特征生成一系列可视化批注,帮助新手用户有效提高阅读理解水平.

2 眼动数据处理

2.1 眼动数据获取

实验邀请到 5 名计算机学科背景、长期阅读英文学术论文且具有较好英语水平的人士作为专家用户,进行阅读实验.所采用眼动仪的操作距离(用户眼睛与显示器屏幕之间的距离)为 40cm~60cm,并带有头部运动误差补偿功能,可有效降低用户在实验过程中由于头部运动带来的误差,平均注视精度为 0.91^[6].实验刺激材料选取自 ACM 全文数据库,这些英文论文逻辑结构清晰、布局层次分明,易于进行眼动数据的记录和分析.论文经加工编排后篇幅为 5 页,显示在分辨率为 1366×768 的 15.6 英寸液晶显示屏上.

为保证数据采集质量,实验在安静、舒适的环境中进行.被试用户距离显示器约 50cm,头戴眼动仪,目视前方显示器.正式实验开始前,向被试用户说明实验目的、操作流程等事项,并使用与正式实验无关,但是难度、篇幅适中的阅读材料进行短暂训练.正式实验过程中,被试用户的阅读时间没有限制,且可以自由翻页(如图 1 所示).同时眼动仪记录下阅读过程中的眼动数据,具体为注视点坐标、注视点注视时间.专家用户阅读完毕后,向实验员进行报告,停止眼动跟踪数据记录.然后实验员将实验材料全文进行感兴趣区(area of interest,简称 AOI)划分,让专家用户对所有 AOI 的重要性进行排序,并进行记录,以用于第 3.2 节 AOI 排序模型训练.



(a) 专家用户的眼动实验场景

(b) 眼动仪

Fig.1 Reading experiment and eye tracker

图 1 眼动实验和眼动仪

2.2 眼动数据处理

大脑无法在短暂的注视时间内对看到的信息进行有效的处理,一般只有当注视时间在 100ms 以上时,大脑才会对眼睛获取到的信息进行解释和加工,因此大多数人在阅读时的注视点持续时间一般为 100 ms~500ms^[7].另一方面,人眼在注视过程中不是稳定不动的,而是会有不自觉的微小颤动^[8].因此,本文对原始注视点作聚类处理,经过聚类后得到的注视点能够真实反映用户的认知情况,而且可以有效消除注视点颤动产生的误差^[9].原始

注视点聚类算法的主要思路是:首先找到一段连续的原始注视点序列,它们在空间上邻近(例如 30px)且在时间跨度上超过一定阈值(例如 80ms);然后计算它们的平均坐标,作为聚类后注视点的坐标^[10].图 2 是应用上述算法对专家用户原始注视点进行聚类后的结果,其中较小的点表示原始注视点,较大的点表示聚类后的注视点.可以发现聚类后的注视点主要分布在文字上,符合用户的实际阅读行为,并验证了聚类算法的有效.同时统计了聚类前后的注视点数据量,以其中一位专家用户为例,聚类前共有 43 828 个原始注视点,聚类后缩减为 8 727 个注视点,数据量只占原来的 19.91%.可见,经过聚类处理后,在一定程度上降低了眼动数据规模,有利于数据的存储和分析等后续操作.

另外,注视点数据在进一步分析使用之前还需进行过滤.例如,正常的阅读过程中几乎不存在持续时间在 1000ms 以上的注视点,如果存在,极有可能是由实验设备误差或其他干扰因素所导致,因此本文对这类极端数据进行剔除.此外,根据本文前期开展的实验发现,在文章前后页衔接的空白处会存在较多的注视点,这是由于用户在手动翻页时,无意识地浏览页末空白,进而产生相应的注视点,因此这些注视点也需要过滤.

3 基于眼动数据的阅读行为可视分析

3.1 阅读行为特征提取

本文以 AOI 为基本单位来计算用户在阅读时的眼动指标,进而提取阅读行为特征.如图 3 所示,为了便于操作,以及考虑到实验刺激材料文字的排版布局,本文选择使用矩形框来设置 AOI(标题、段落和图表等作为 AOI,其他区域,如空白部分,用户很少关注,所以不考虑在内).

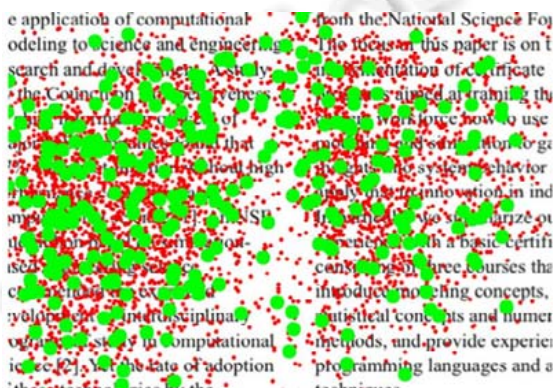


Fig.2 Fixation points before and after clustering
图 2 原始注视点和聚类后的注视点示例

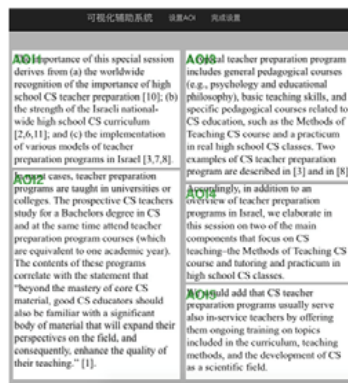


Fig.3 Settings of AOI
图 3 AOI 设置

假设有 m 个 AOI、 n 个注视点,为了统计各 AOI 范围内的注视点,传统方法需要遍历所有 AOI 和所有注视点,时间消耗为 $T(mn)$,所以当有较多 AOI 和注视点时,计算速度会明显降低;而使用基于 KD-Tree(K-Dimensional Tree)的查询算法,时间消耗降低到 $T(m \log n)$,可以大幅提高计算速度^[11].为此,本文基于 KD-Tree 的查询算法统计 AOI 内的注视点,可以加快计算速度,在有较多 AOI、大量注视点时效果尤为明显.

统计出各 AOI 内的注视点数据后就可以计算相关眼动指标,提取阅读行为特征.进一步定义两类阅读行为特征:一类是 AOI 内阅读特征,包括 AOI 内的阅读速度和该 AOI 的阅读次数;另一类是 AOI 间特征,具体指两个 AOI 之间的阅读转换次数(用于了解不同段落之间的语义联系,便于定位相关内容^[12]).各阅读行为特征具体描述如下:

(1) AOI 阅读速度

由当前 AOI 内的眼跳总距离与注视点驻留时间之比计算当前 AOI 的阅读速度,具体公式如下:

$$S_{AOI} = \frac{\sum_{k=1}^m L_k}{T_{AOI}} \quad (1)$$

式中 S_{AOI} 表示当前 AOI 的阅读速度, L_k 为当前 AOI 内的每一段眼跳距离(即两个时序相邻的注视点之间的距离), T_{AOI} 为当前 AOI 内的注视点驻留时间. 此外, 定义 AOI 内时间阈值为 p (根据前期实验经验可以取 $p=5s$), 如果 $T_{AOI} \geq p$, 那么记录下这段时间内的阅读速度, 因为过短的驻留时间通常意味着无效阅读, 例如读者常常会滚动浏览器, 在这个过程中虽然他们的视线会在一些 AOI 内进行非常短暂的停留, 但并不是有意识的阅读.

(2) AOI 阅读次数

通常重要的内容或者较难理解的内容会被读者反复阅读, 所以统计每个 AOI 的累计阅读次数有助于表征内容的重要性和可理解性. 然而, 因为自底向上的视觉认知干扰^[13], 如被显著的标题所吸引, 注视点会无意识地转移到其他 AOI, 停留短暂时间后马上又转换到原来的 AOI. 因此, 为了统计有效的阅读次数, 本文规定: 如果注视点首次进入某个 AOI, 且驻留时间超过阈值 p , 那么定义其为源 AOI, 并认为在该 AOI 内进行了有效阅读, 阅读次数加 1; 如果注视点离开该 AOI, 移动到其他 AOI 并进行了有效阅读, 那么定义其为目标 AOI, 则当注视点再次返回到该 AOI 时, 阅读次数加 1.

(3) 两个 AOI 之间的阅读转换次数

根据上文的规定, 本文将源 AOI 到目标 AOI 的注视点转换计为 1 次有效的阅读转换. 如果两个 AOI 之间阅读转换次数越多, 则其内容具有越紧密的联系.

3.2 AOI 排序模型训练

上文的专家用户阅读眼动实验记录了 5 位专家的眼动数据, 在此基础上提取相应的阅读行为特征, 包括 AOI 的阅读速度(readspeed)、AOI 的阅读次数(readcount)、AOI 内包含的关键词个数(keycount); 此外, 还记录了专家各自对所有 AOI 重要性的主观排序序号(order). 将上述数据作为数据样本, 进行支持向量回归模型的训练, 用以预测所有 AOI 的重要性排序, 进而确定文章的重点内容区域.

具体的, 样本数据可以分为两类, 一类是真实标签值(主观排序序号), 定义其集合为: $S = \{s_i = order_i, i=1, 2, \dots, n\}$, n 为 AOI 数目; 另一类是属性值(阅读行为特征), 定义其集合为: $E = \{e_i = (readspeed_i, readcount_i, keycount_i), i=1, 2, \dots, n\}$. 在支持向量回归模型的训练中, 得到 $readspeed, readcount, keycount$ 与 $order$ 之间的映射关系函数 $order_i = f(readspeed_i, readcount_i, keycount_i)$, 并对 f 进行求解. 模型在训练过程中还需要不断调整参数, 使得测试误差尽可能的小. 为此本文设置了多组备选参数进行训练, 逐一记录结果的精度, 并采用了 5 折交叉验证^[14]进行模型评估, 从而选取分类效果最好的参数作为模型最终的参数. 原始数据共含有 100 个样本(5 个专家, 每个专家 20 个样本), 本文选取其中的 80 个样本作为训练集, 20 个样本为测试集, 并经 5 折交叉验证后, 得到最终的 AOI 重要性排序结果 R (R 为排序结果集合), $R = \{AOI2, AOI19, AOI5, AOI4, AOI8, AOI10, AOI17, AOI11, AOI3, AOI13, AOI15, AOI14, AOI12, AOI7, AOI18, AOI6, AOI9, AOI6, AOI1, AOI20\}$.

3.3 基于关键词的文本主题可视分析

本文基于 TF-IDF (term frequency-inverse document frequency)^[15] 算法进行关键词提取. 词频 (term frequency, 简称 TF) 表示某单词在文档中出现的频率, 该单词在所选文档中出现的频率越高则代表该它与该文档的主题相关程度越高, 即重要性也越大. 逆文档频率 (inverse document frequency, 简称 IDF) 则是文档频率的倒数. 在使用 TF-IDF 算法提取文章关键词之前会建立一个语料库, 如果在语料库中含有某个词的文档数越多, 则该词的重要性和代表性就相应的下降. 提取关键词的算法主要可以分为 3 个部分:

(1) 文章预处理

定义需要提取关键词的文章为 d , 首先需要进行分词和剔除停用词 (stop words) 等预处理操作. 由于本文中所选实验材料为英文文章, 而英文的基本组成单位就是单词, 因此只需根据空格和标点符号进行分词. 停用词表示本身没有实际意义, 对文章内容也没有支撑性作用, 但是词频又往往比较高的词汇, 需要将此类词汇进行过

滤.经过预处理后得到文本 d_i ,即把文章 d 的内容看成是剔除了停用词后的单词组成的集合,可以表示为 $d_i=(t_1,t_2,t_3,\dots,t_j,t_D),1\leq j\leq D,D\in N$.

(2) 权重计算

根据各个单词 t_j 在文本 d_i 中的重要性并计算它们的权重值.

(3) 提取关键词

对文本 d_i 中的每个单词都进行上述权重计算,得到每个单词的权重值,然后按照权重值从大至小进行排序,选择前 r (例如 80)个单词作为实验中最终的关键词.本文得到的权重输出结果见表 1(篇幅有限仅列出一部分).

Table 1 Example result of weight output

表 1 权重输出结果示例

关键词	描述	排序
preparation	0.006 513	1
teacher	0.005 625	2
programs	0.005 033	3
teaching	0.004 440	4
session	0.003 552	5

4 应用实例

利用热区、圆环、节点链接等可视化视图形来展示专家用户的阅读行为特征,为新手用户阅读提供可视辅助,提高他们的阅读理解水平和效率.

4.1 原型系统

系统按功能和可视分析效果可以分为 5 个模块:AOI 设置模块、热区图模块、圆环图模块、节点链接图模块和词云模块.系统主界面如图 4 所示,下面对各个模块进行具体说明.



Fig.4 Main user interface of the prototype system

图 4 系统主界面

(1) AOI 设置模块

如图 5 所示,用户点击“设置 AOI”按钮后开始拖动鼠标进行 AOI 设置,并在左上角区域进行编号.当划分完所有 AOI 后,则点击“完成设置按”按钮.此外,由于实验刺激为图片格式,系统在每一个 AOI 划分的同时对该 AOI 区域进行截图和保存,并对截图进行光学字符识别(optical character recognition,简称 OCR)^[16],进而计算各 AOI 内包含的关键词数目.

(2) 热区图模块

将专家用户的注视点数量、注视点持续时间与颜色、透明度进行映射,绘制热区图^[17],并以半透明的方式在实验刺激材料上进行叠加,以此表示专家用户的视觉注意力分布情况.如图 6 所示,热区图中红色区域表示专家用户注视点较为集中、注视时间较长的区域,新手用户可以据此对文章的不同区域给予不同程度的重视.



Fig.5 An example of AOI setting
图5 AOI 设置界面示例

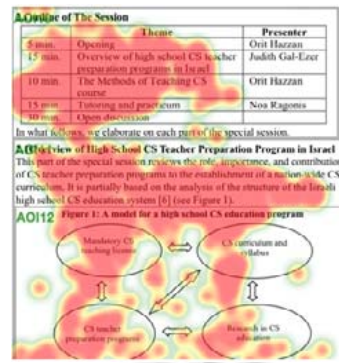


Fig.6 An example of heatmap
图6 热区图模块示例

(3) 圆环图模块

如图 7(a)所示,每段圆弧代表一个重点 AOI,重点 AOI 也根据重要性从高到低在圆环上顺时针排列,且越重要的 AOI 对应的圆弧弧长越长,反之则越短;同时,根据重要性排序从高到低,设置每段圆弧的颜色从亮红到暗红依次渐变.如图 7(b)所示,当用户需要了解更多信息时,可将鼠标移动到圆弧上,此时以浮动框形式显示该段圆弧所对应 AOI 的专家用户阅读速度、阅读次数,以及该 AOI 内包含的关键词数量,鼠标移开则浮动框消失.

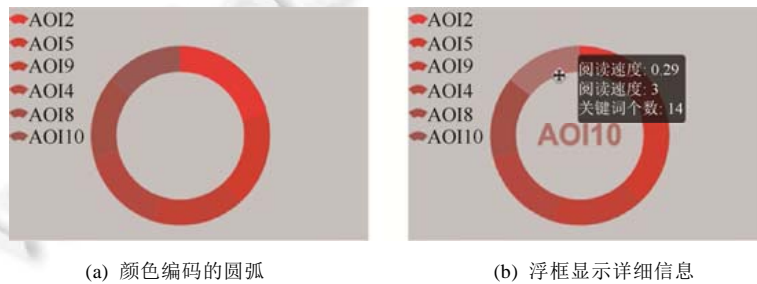


Fig.7 An example of doughnut chart
图7 圆环图模块示例

(4) 节点链接图模块

使用节点链接图展现专家用户对全文各段落的注意力分布情况.采用力导向(force-directed)^[18]布局算法进行节点布局,它将每个节点视为质点,通过在质点之间建立力的关系,形成一个物理的弹簧模型系统^[19,20].如图 8(a)所示,将每个 AOI 表示为一个节点,根据 AOI 的重要性排序结果设置节点半径,半径越大的节点表示该 AOI 的内容越重要,反之亦然.此外,使用颜色更加直观地区分重要性,例如,红色代表重点 AOI,蓝色代表普通 AOI.节点之间具有连线,用连线代表 AOI 之间的阅读转换,连线越短说明转换次数越多,即这两个 AOI 的内容之间具有越紧密的联系.如图 8(b)所示,当鼠标悬浮到某个节点上时,则显示该节点所对应的文字内容.如图 8(c)和 8(d)所示,为了减少视觉干扰,系统还允许用户通过视图左上角的按钮选择仅显示重点 AOI 或普通 AOI.

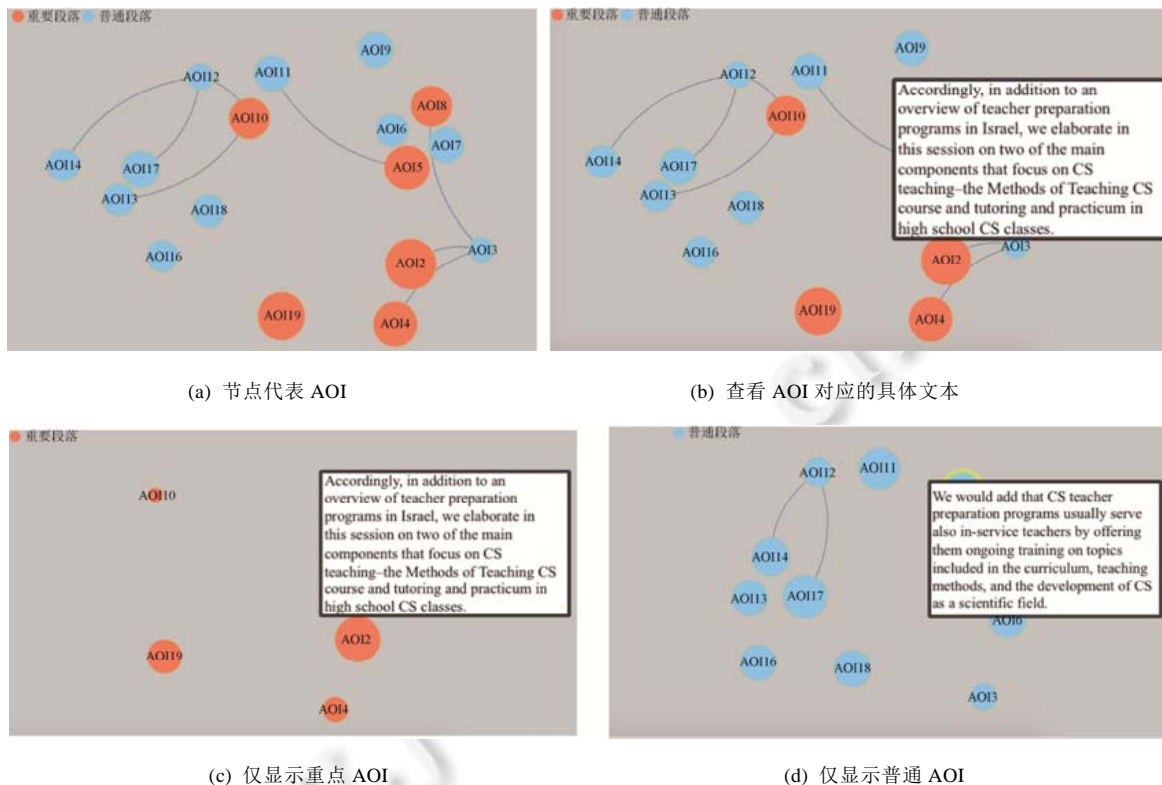


Fig.8 An example of node link diagram module
图 8 节点链接图模块示例

(5) 词云模块

关键词是一篇文档核心内容的体现,读者通过阅读关键词即可对整个文档的关键内容进行快速了解.本文以螺旋路径词云^[21]的方式对关键词进行可视化展示.如图 9 所示,根据权重从大至小的顺序从画布中心位置开始,按照螺旋路径逐渐往外层排列,并将关键词的权重与大小进行映射,即权重最大的词在词云中尺寸最大,反之亦然;且对关键词进行了不同的颜色编码,视觉效果上更加美观;此外,当鼠标悬停到某个关键词上时,还会显示该词的权重值.



Fig.9 An example of keyword cloud module
图 9 关键词词云模块示例

4.2 用户实验与结果分析

4.2.1 实验设计

本文实验邀请了 42 名被试用户(34 名男性和 8 名女性),年龄在 23 岁~27 岁之间,平均年龄 26 岁,均通过大学英语四级考试且分数相近,具有相近的英文阅读水平.所有被试用户视力矫正正常或裸视正常,且专业背景均为计算机.在正式实验前进行预实验,向用户讲解实验系统的使用方法和操作流程,并要求每个被试用户阅读一篇 1 页篇幅的英文学术论文(内容与正式实验中的阅读材料内容不同),然后剔除阅读速度过快(例如小于 5min)和过慢(例如 20min)的被试用户,降低被试用户在阅读经验和阅读能力方面之间的差异性,从而保证实验结果的可信度.预实验后被试用户剩下 40 名(剔除了 2 名男性),将这些用户随机分成实验组和对照组,然后交换个别用户,以平衡年龄和性别.

正式实验中,实验组和对照组用户阅读的论文与前文专家用户阅读的论文相同.对照组用户阅读时没有可视化及可视分析辅助,且佩戴眼动仪;实验组用户可以看到基于专家用户眼动数据生成的可视化视图,同时佩戴眼动仪,用于分析用户视觉行为是否受到这些视图的影响.两组用户均在舒适、无干扰的环境中进行实验,都不允许使用手机或其他设备.图 10 为实验组和对照组用户的实验场景.

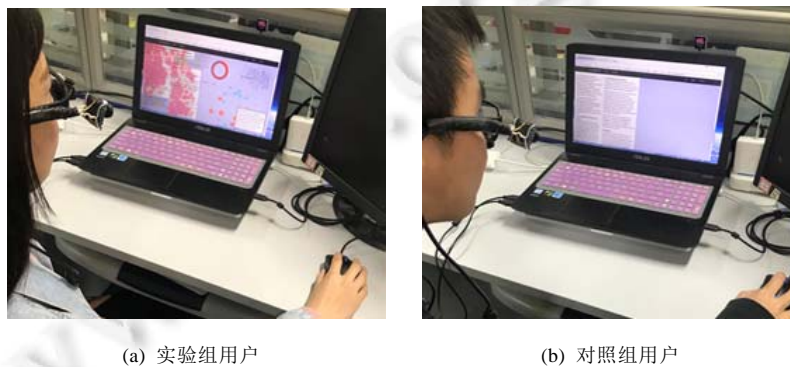


Fig.10 Experimental scene

图 10 实验场景

阅读实验结束后,采用答题的方式对用户进行阅读理解水平的评价.包括客观和主观评价两方面:客观评价是回答 5 个阅读理解问题(单项选择题,在论文章节中包含相应题目的答案),答对一题计 1 分;主观评价是让被试用户针对这篇文章写一段总结,然后由专家对该总结进行评价,衡量学生是否对论文内容有全面和准确的理解,并以 5 分制进行评分(1 分表示最差,5 分表示最好).

4.2.2 结果分析

(1) 任务完成时间分析

实验组和对照组的阅读任务平均完成时间分别为 552.2s(标准差 101.0s)和 567.5s(标准差 172.3s).配对 T 检验结果 $t=-0.342, p>0.05$,无显著性差异.原因在于,对两组用户的实验并没有设置时间限制,允许用户以正常阅读速度进行阅读,所以实验组用户可以充分浏览各可视化视图,所花费时间并不明显少于对照组.

实验组和对照组的答题(客观题和主观题合并统计)平均完成时间分别为 409.6s(标准差 64.2s)和 453.8s(标准差 69.3s).配对 T 检验分析结果 $t=-2.089, p<0.05$,存在显著性差异,即实验组用户明显要比对照组的答题时间更短,说明实验组用户在使用了本系统的帮助后,答题的效率更高.

(2) 答题得分情况分析

实验组和对照组的客观题答题平均值分别为 2.9 分(标准差 0.9 分)和 2.2 分(标准差 1.2 分).配对 T 检验分析结果 $t=2.185, p<0.05$,存在显著性差异,即实验组用户的客观题得分明显要高于对照组.类似的,对于主观题得分,实验组和对照组的答题平均分分别为 3.1 分(标准差 0.8 分)和 2.0 分(标准差 0.9 分).由于主观题的得分并

不具有正态分布特点,因此使用曼-惠特尼 U 检验,结果 $U=70.000, Z=-3.61, p<0.05$,存在显著性差异,即实验组用户的主观题得分明显要高于对照组.可见实验组用户通过本系统的帮助,对文章的理解程度得到了显著提高.

(3) 视觉行为分析

分析实验组用户是否受到可视化的影响,以及两组用户的视觉行为是否具有差异.图 11 展示了实验组所有用户的热区图、对照组用户的热区图和专家读者的热区图(限于篇幅,仅以文章某一页为例).



Fig.11 Heatmap for novice readers and expert readers

图 11 新手读者和专家读者的热区图

如图 11(b)所示,对照组用户的注视点分布比较均匀,基本没有密集和稀疏之分,说明不能较好地把握文章重点.从图 11(a)和图 11(c)可以发现,实验组用户的注视点分布与专家用户相似,在部分段落较为集中,主次较为分明.

进一步,对实验组单个用户的热区图进行对比(限于篇幅,此处仅以其中几个用户为例).图 12 中用户 3、用户 6 和用户 17 的热区图与专家用户的热区图比较相似,说明他们确实通过参考专家用户的阅读模式来阅读文章;而用户 8 和用户 10 的热区图与专家用户差异较大,说明他们可能并没有完全参考专家用户的阅读模式.



Fig.12 Example of comparison with the single user's heatmap from the experimental group

图 12 实验组单个用户热区图对比示例

另一方面,如图 13 所示,由于对照组用户无法参考专家用户的阅读行为模式,他们的眼动热区图与专家用户的热区图差异更大.

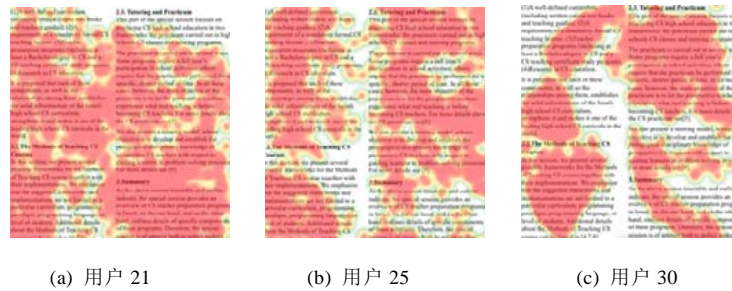


Fig.13 Example of comparison with the single user's heatmap from the control group

图 13 对照组单个用户热区图对比示例

本文还统计了实验组和对照组在各 AOI 内的平均阅读次数、平均阅读速度和 AOI 间的平均阅读转换次数.如图 14 所示,实验组用户在重点 AOI 内的平均阅读次数,较对照组更为接近专家用户.

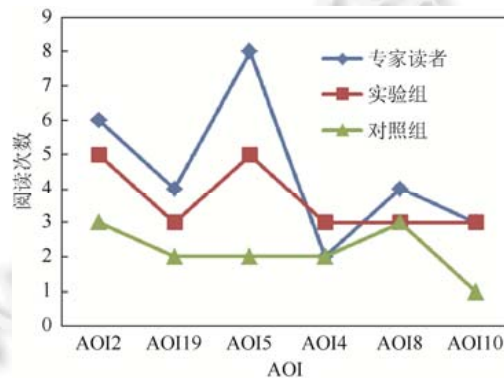


Fig.14 Reading time in key AOIs

图 14 重点 AOI 内的阅读次数

另一方面,利用矩阵分析用户在不同 AOI 之间的转换次数,图 15(a)、图 15(b)、图 15(c)分别为对照组、实验组、专家用户在不同 AOI 之间的平均转换次数.可以看出,对照组用户的 AOI 转换次数矩阵的数据分布较为随机,说明注意力的分布没有特定模式;而实验组和专家用户矩阵数据的分布具有较高的相似性,说明实验组用户借鉴了专家用户的阅读模式,在部分 AOI 之间具有较多的转换.

进一步,本文还基于眼动数据分析了实验组用户在原型系统各视图中的视觉注意行为,进而分析哪一个可视化视图对用户最为吸引.从图 16 中可以发现,由于热区图正好在刺激材料正文区域,因此用户的注视点在此区域最为集中.除此之外,注视点在节点链接图也较为集中(这与下文中用户的主观反馈也保持一致).

(4) 用户主观反馈

对实验组用户进行问卷调查和访谈.结果表明,大部分用户认为本系统及各模块很有帮助或非常有帮助,例如,45%的用户认为对他们最有帮助的是节点链接图,其次是热区图(40%的用户).大部分用户认为本系统实用性很强,对阅读帮助很大,能提高阅读效率,且具有较强的创新性.另一方面,用户也给出了完善建议,例如,热区图和 AOI 编号对文字所有遮挡,建议做适当调整.

4.2.3 讨论

在实验过程和用户访谈部分,发现不同专家用户的阅读模式存在一定的差异,例如有的专家对某些段落快速浏览,而有的专家则在相同段落花费了更多时间,为此,本文目前还是采取求平均值的方法进行差异性融合(例如对多个专家在同一个 AOI 的阅读速度求其平均值);另一方面,专家和新手之间由于在阅读习惯、水平等方面存在差异,新手在借鉴专家阅读模式时也会受到负面影响,例如专家对于专业名词可能不会停留过长时间

5 结论与下一步工作

本文面向阅读行为,融合眼动数据和文本信息进行分析,并设计了相关可视化形式展示文章重点段落、内在联系等,在此基础上设计与开发了一个面向阅读辅助的可视分析应用系统,帮助新手用户借鉴和参考专家用户的阅读模式,以便提高阅读效率和理解水平.下一步工作将对专家之间的差异性,以及专家和新手之间的差异性开展深入研究,并考虑用户的个人喜好,对具体的可视化策略和方法进行优化,进一步提高阅读辅助效果.

References:

- [1] Biedert R, Buscher G, Schwarz S, Hees J, Dengel A. Text 2.0. In: Proc. of the CHI'10 Extended Abstracts on Human Factors in Computing Systems. ACM, 2010. 4003–4008. [doi: 10.1145/1753846.1754093]
- [2] Xu SH, Jiang H, Lau F. User-oriented document summarization through vision-based eye-tracking. In: Proc. of the 14th Int'l Conf. on Intelligent User Interfaces. ACM, 2009. 7–16. [doi: 10.1145/1502650.1502656]
- [3] Martínez-Gómez P, Aizawa A. Recognition of understanding level and language skill using measurements of reading behavior. In Proc. of the 19th Int'l Conf. on Intelligent User Interfaces. ACM, 2014. 95–104. [doi: 10.1145/2557500.2557546]
- [4] Augereau O, Kunze K, Fujiyoshi H, Kise K. Estimation of english skill with a mobile eye tracker. In: Proc. of the 2016 ACM Int'l Joint Conf. on Pervasive and Ubiquitous Computing: Adjunct. ACM, 2016. 1777–1781. [doi: 10.1145/2968219.2968275]
- [5] Cheng SW, Sun ZQ, Sun Lw, Yee, K, Dey A K. Gaze-based annotations for reading comprehension. In Proc. of the 33rd Annual ACM Conference on Human Factors in Computing Systems. ACM, 2015. 1569–1572. [doi: 10.1145/2702123.2702271]
- [6] Cheng SW, Sun ZQ. An approach to eye tracking for mobile device based interaction. Journal of Computer-Aided Design & Computer Graphics, 2014,8:1354–1361(in Chinese with English abstract).
- [7] Inhoff AW, Briihl D, Schwartz J. Compound word effects differ in reading, on-line naming, and delayed naming tasks. Memory & Cognition, 1996,24(4):466–476. [doi: 10.3758/BF03200935]
- [8] Martinez-Conde S, Macknik S L, Troncoso X G, Hubel D H. Microsaccades: A neurophysiological analysis. Trends in neurosciences, 2009,32(9):463–475. [doi: 10.1016/j.tins.2009.05.006]
- [9] Cheng SW, Shi YW, Sun SQ. An approach to usability evaluation for mobile computing user interface based on eye-tracking. Acta Electronica Sinica, 2009,37(b04):146–150(in Chinese with English abstract).
- [10] Sun ZQ. Research on the human-computer interaction technique for multi-device based on eye tracking [MS. Thesis]. Hangzhou: Zhejiang University of Technology, 2016 (in Chinese with English abstract).
- [11] Li H. Statistical Learning Method. Beijing: Tsinghua University Press, 2012 (in Chinese).
- [12] Cheng SW, Sun YJ. Eye movement data visualization based annotation for reading teaching. Journal of Zhejiang University of Technology, 2017,45(6):610-614 (in Chinese with English abstract).
- [13] Connor CE, Egeth HE, Yantis S. Visual attention: bottom-up versus top-down. Current biology, 2004,14(19):R850–R852. [doi: 10.1016/j.cub.2004.09.041]
- [14] Browne MW. Cross-validation methods. Journal of mathematical psychology, 2000,44(1):108–132. [doi: 10.1006/jmps.1999.1279]
- [15] Shi CY, Xu CJ, Yang XJ. Study of TFIDF algorithm. Journal of Computer Applications, 2009,29(6):167–170. (in Chinese with English abstract).
- [16] Singh S. Optical character recognition techniques: a survey. Journal of emerging Trends in Computing and information Sciences, 2013,4(6):545–550.
- [17] Spakov O, Miniotas D. Visualization of eye gaze data using heat maps. Elektronika ir elektrotechnika, 2007:55–58.
- [18] Eades P. A heuristic for graph drawing. Congressus numerantium,1984,42:149–160.
- [19] Kobourov SG, Wampler K. Non-Euclidean spring embedders. IEEE Trans. on Visualization and Computer Graphics, 2005,11(6):757–767.
- [20] Kamada T, Kawai S. A general framework for visualizing abstract objects and relations. ACM Trans. on Graphics (TOG), 1991,10(1):1–39. [doi: 10.1145/99902.99903]
- [21] Viegas FB, Wattenberg M, Feinberg J. Participatory visualization with wordle. IEEE Trans. on Visualization and Computer Graphics, 2009,15(6):1137–1144. [doi: 10.1109/TVCG.2009.171]

附中文参考文献:

- [6] 程时伟,孙志强,陆煜华.面向多设备交互的眼动跟踪方法.计算机辅助设计与图形学学报,2016,28(7):1094-1104.
- [9] 程时伟,石元伍,孙守迁.移动计算用户界面可用性评估的眼动方法.电子学报,2009,37(b04):146-150.
- [10] 孙志强.基于眼动跟踪的多设备人机交互技术研究[硕士学位论文].杭州:浙江工业大学,2016.
- [11] 李航.统计学习方法.北京:清华大学出版社,2012.
- [12] 程时伟,孙煜杰.面向阅读教学的眼动数据可视化批注方法.浙江工业大学学报,2017,45(6):610-614.
- [15] 施聪莺,徐朝军,杨晓江.TFID 算法研究综述.计算机应用,2009,29(B06):167-170.



程时伟(1981—),男,湖北黄石人,博士,教授,博士生导师,CCF 高级会员,主要研究领域为人机交互,普适计算,协同计算.



孙煜杰(1992—),男,硕士,主要研究领域为人机交互.



胡屹凇(1994—),女,硕士生,主要研究领域为人机交互.

www.jos.org.cn

www.jos.org.cn