

# 面向满意度预测的滑鼠行为量化分析方法

洪 宇 王 剑 王 凯 康杨杨 孔 芳 姚建民 朱巧明 周国栋

(苏州大学计算机科学与技术学院 江苏 苏州 215006)

**摘 要** 现有信息检索研究领域中,衡量知识学习质量和信息获取精度的核心标准是信息与用户需求的相关性(量化指标为相关度),然而,这一测度往往无法直观反映用户对信息伪反馈的“满意度”。相比于多媒体(文字语言、图像、音频和视频)之间可测量的相关度,由用户主观认知驱动的满意度往往无法通过直观的量化方法予以获取和测量。针对这一问题,文中提出一种基于鼠标滑动(Mouse Movement,简称“滑鼠”)运动学规律的“满意度”量化分析和预测方法,该方法集中于人类肢体活动驱动下的滑鼠滑行轨迹分析,借助复杂滑行过程中滑鼠呈现出的动力学能量,间接预测人类思维活跃的程度,以此估计用户接触特定信息伪反馈时隐式反射出的满意度。实验验证,该方法能够有效辅助信息检索过程中的用户体验分析。

**关键词** 信息检索;满意度;滑鼠行为;滑鼠能量;社交网络;社会计算

**中图法分类号** TP391 **DOI号** 10.11897/SP.J.1016.2015.02064

## Satisfaction Prediction Oriented Quantitative Mouse Movement Analysis

HONG Yu WANG Jian WANG Kai KANG Yang-Yang KONG Fang YAO Jian-Min  
ZHU Qiao-Ming ZHOU Guo-Dong

(School of Computer Science and Technology, Soochow University, Suzhou, Jiangsu 215006)

**Abstract** In current study on Information Retrieval (IR), the determination of quality of knowledge learning and precision of information acquisition heavily depend on the relevance (quantity is named as relevance degree) between user's need and information. However, the quantity normally is incapable of reflecting the satisfactoriness degree of users to pseudo information feedback. Compared to the scalable relevance degree among multi-medias, such as linguistic texts, images, audios and videos, the satisfactoriness, which is triggered and driven by human's subjective recognition, is not easily reachable. In particular, it is difficult to directly measure the satisfactoriness degree. To solve the problem, the paper proposes a Mouse-Movement-Law based on satisfactoriness analysis and measurement method. The method concentrates on trajectory analysis of mouse movement that is driven by human's physical activity. More importantly, the method detects the impetus of mouse movement (i. e., momentum) during the course of sliding, by which it indirectly reflects the activity degree of the mind. The quantity of momentum, accordingly, is favorable to measurement of satisfactoriness toward pseudo information feedback. Experiments show that the method is effective in supporting analysis of user experience in the process of IR.

**Keywords** information retrieval; satisfaction; mouse movement; energy of mouse movement; social networks; social computing

收稿日期:2014-08-02;最终修改稿收到日期:2015-05-10. 本课题得到国家自然科学基金(61373097,61272259,61272260)资助。洪宇,男,1978年生,博士,副教授,中国计算机学会(CCF)会员,研究方向为信息检索、话题检测、信息抽取与观点挖掘。E-mail: tianxianer@gmail.com。王剑,男,1990年生,硕士研究生,中国计算机学会(CCF)会员,主要研究方向为个性化信息检索。王凯,男,1990年生,硕士研究生,中国计算机学会(CCF)会员,主要研究方向为个性化信息检索。康杨杨,男,1989年生,硕士研究生,主要研究方向为个性化信息检索。孔芳,女,1977年生,博士,副教授,研究方向为信息抽取和自然语言处理。姚建民,男,1971年生,博士,副教授,研究方向为机器翻译。朱巧明,男,1963年生,博士,教授,研究领域为中文信息处理。周国栋,男,1968年生,博士,教授,研究领域为自然语言处理。

## 1 引言

现代信息检索的雏形是一套以信息和知识获取为目标,以用户与信息载体的迭代交互为基本过程,以信息需求与信息本源内涵的相关性为核心度量标准的机器智能和处理技术.其中,相关性理论(相关论)的核心思想是评价信息中蕴含的知识、实体、媒体、数据和语义(整体称为“内容”)是否拟合于用户表述的查询内容与意图<sup>[1-2]</sup>.比如,衡量一篇“苹果公司上市”和一篇“进口苹果上市”的新闻报道分别与查询“苹果上市”的相关性(该例侧重意图相关性的判断,而非语言层面,前者相关,后者不相关).这一理论支撑了现有绝大部分的信息获取应用技术,包括以领域相关性为核心的垂直检索(Vertical Search)、以用户为中心的个性化搜索(Personal Search)<sup>[3]</sup>和广告搜索(Advertisement Search)<sup>[4]</sup>.相应地,诸如排序学习<sup>[5]</sup>、查询重构与扩展<sup>[6]</sup>、意图与偏好学习<sup>[7-8]</sup>等核心方法的研究集中于解释、表示和测量相关性.

然而,现有信息检索技术的应用与实践已然证明相关性仅仅代表了检索结果价值体系中的独特一元,而非全部.比如,Yu 等人<sup>[9]</sup>的研究显示,信息本源正确性在评定检索结果合理性中的价值,并证明了在检索结果推送、过滤和排序过程中区别对待相关性和正确性的必要性,比如,“地球是圆的”(本源正确)为值得反馈的合理结果,“地球是方的”(本源错误)为具有误导性的噪声,应予以屏蔽.

相对地,在用户一端,相关性也并非唯一有效的评价标准.恰如本文所述,用户对信息的满意度也是反映检索结果价值的重要标准.比如,尽管都在文字信息层面相关于查询意图“风衣”,但是男款“风衣”和女款“风衣”在两性之间有着截然不同的吸引力,产生明显相异的信息获取体验.由此,满意度可以定义为个体主观认知驱动的一种对信息质量的认可程度.满意度是一种难以直接探测与度量的指标.这一论断来源于如下认知发现:语言形式和信号形式的信息是一种可见信息,机器学习和处理可以借助语言和信号之间的逻辑、语义、结构和焦点关系衡量信息之间的关联程度<sup>[10]</sup>,而满意度是一种人类的感知反馈,属于不可见信息,缺少直接承载和表示这类感知信息的方法,从而难以在语言和感知(或信号和感知)之间建立相互关联的桥梁,关联度难以得到直接

且无偏的计算.比如,在线文本搜索过程中,隐藏于文字查询之后的用户情感、品味和智慧活动并没有显式且直接的表现形式和传输载体,难以估计与度量.

针对这一问题,现有基于用户行为特征的信息检索研究尝试间接的探测手段.比如,点击行为(Click)指向的检索结果往往间接体现用户的意图与偏好<sup>[10-13]</sup>.然而,这一特征更多地反映用户对现有多样性查询反馈的筛选原则,有助于缩小目标对象的概念范围,去除冗余和噪声,以及消除信息和信号中蕴含的歧义.比如,如果用户点击题为“苹果公司上市”的检索结果,则可以缩小“苹果”的词义范畴,摒除“水果类”结果的干扰.但是,这一特征仍然驻留于“什么相关于用户需求”的相关论范畴,仅仅是浅层的认知理解,无法解释用户在感念相同的同类检索结果中进行比较、对照和筛选的根本原理.比如,同为相关于“女款风衣”的检索结果,特定女性用户仍然有着截然不同的选择,并往往具有“优良”、“尚佳”、“一般”和“较差”等程度不同的价值判定结果,呈现一种“度”的区分性,而非“类”的区分性,反映出满意度和相关性截然不同的作用.

相对而言,现有利用视觉特征的用户学习有助于挖掘用户满意度生成过程中的生理线索,比如,瞳孔扩张的程度和视觉焦点驻停的时长等<sup>[14]</sup>.然而,视觉特征的采集和机器学习模型的训练往往对用户具有的设备(是否具有摄像头及其清晰度)、所处的操作环境(眼球与摄像头的距离与角度)以及个体生理特点(瞳距、尺寸和运动规律)等等,都有着极大程度的依赖性,其通用性和可靠性往往难以得到有效保证.

本文提出一种基于滑鼠运动学原理的检索满意度分析与测量技术.其核心思想是利用鼠标在屏幕上的滑行轨迹,计算其虚拟的物理动量,并根据滑鼠运动与人体上肢运动之间的协调关系,等比表示用户控制鼠标滑行和“趋向点击”过程中产生的生物能量.根据“意图愈强,则条件反射更为剧烈”的假设,该技术尝试将这一生物能量映射为认知驱动的满意度量化指标,辅助满意度的测量.这一技术的优点在于:(1)现有机械和触摸式鼠标已得到广泛应用;(2)滑鼠数据易于采集且在相同操作系统下往往具有统一的信号模式,从而,机器学习和训练所得的数学模型具有一定的通用性和可靠性.

本文主要内容为集中验证滑鼠动量与满意度之

间具有协同性的基本假设,重点开展了如下工作:

(1) 面向滑鼠数据和满意度数据的采集,建立了一套以图片检索为基础的标注平台,集成了图片搜索、滑鼠跟踪以及个体满意度评分技术;

(2) 提出并实施了滑鼠动量分析与计算方法,其中滑鼠动量涵盖无质量的动能与角动能指标,利用分段计算与合并法,对复杂滑鼠轨迹中的动量进行综合测量;

(3) 研究分析了滑鼠动量与用户满意度评分的拟合规律,侧重分析特定时间范围内且完整检索行为序列中,滑鼠动量与满意度趋势的整体拟合与互斥现象,挖掘协作演化的规律.

其中,个体满意度的客观存在性和差异性检验,滑鼠动量的分析与计算方法以及两者之间的协作关系检验是本文重点. 该研究是利用人体行为特征间探测与模拟用户认知反馈的一项全新尝试. 研究过程引入了简单的仿生学和物理学知识,辅助滑鼠行为分析和量化计算. 实验结果显示,特定用户在概念相同且属性类似的检索伪反馈中,仍然存在显著的满意度评判差异. 此外,滑鼠的动量能在一定程度上反映个体满意度的差异.

本文第 2 节简要回顾用户检索倾向性分析的相关研究;第 3 节给出本文的基本假设和主要思想;第 4 节给出滑鼠行为的量化分析方法;第 5 节介绍满意度标注及滑鼠跟踪平台;第 6 节介绍实验并给出结果分析;第 7 节总结全文.

## 2 相关工作

现有信息检索的基础方法理论可分为 4 类: 语言理解、概率统计、机器学习和行为分析,其中,前三者可合并为计算语言学,从语言特征和统计特征入手,以机器学习为辅助,建立信息关联关系的计算模型,侧重以文字相关性为准绳进行知识获取. 相对地,行为分析专注于学习用户在检索过程中的行为特征,借以探测用户意图和倾向,侧重于以行为相关性为准绳的知识获取和应用. 因而,面向信息检索的行为分析研究与本文最为相关,本节将重点回顾这一领域的相关工作.

前人已对用户的点击行为<sup>[10-12]</sup>进行了较为深入的分析. 其中,根据 Agichtein 等人<sup>[13]</sup>的研究成果,用户在浏览检索结果列表时进行的有序点击行为,能够在一定程度上反映用户主观判断相关性的标准

与原则,点击顺序的机器学习与应用有助于实时的检索结果重排序. 在此基础上,Dupret 等人<sup>[15]</sup>提出了基于点击顺序的相关信息检测模型,通过学习用户查询日志中检索记录组块中的查询与点击顺序关系,识别用户实时检索过程中的关联查询,并判定点击概率较高的潜在相关检索结果.

但是,上述研究无法回避如下问题:(1) 点击顺序受限于检索结果列表的先验排序,具有一定的局限性;(2) 点击顺序往往取决于用户的检索习惯,却并不绝对依赖于点击对象与用户需求在内容层面上的一致性 or 可比性,这对于反映用户更倾向于 or 更热衷于特定检索结果的判定,难以提供可靠的线索. 此外,Fan 等人<sup>[16]</sup>认为点击的目标是信息的抽象,比如,文本检索结果列表中,每项结果往往是一套文档的主题、快照(一种包含查询词的文摘或文字片段)和发布日期的可点击的套件,仅仅涵盖与查询关联的局部内容,并未覆盖全部内容. 其结果是点击前后的用户认知并不完全一致,换言之,点击前的用户体验与点击后的用户体验将有一定差异,从而,点击往往反映一种“第一印象”,点击顺序则反映“第一印象”的排序. 为此,Fan 等人提出将点击行为特征与实际被点击的信息的全局内容进行综合考量的思想,形成了融合模型. 值得注意的是,本文研究的对象(滑鼠轨迹与动量)发生于点击之前,因此,也存在“第一印象”的问题. 在实际应用中,应根据 Fan 等人的理论,对点击前的行为特征和后验的信息内容进行综合建模.

触发本文核心工作的相关研究来自于 Dupret 等人<sup>[17]</sup>,其提出了点击“效用衰减”理论,核心思想来源于如下假设: 用户点击和浏览信息的过程中,将从目标信息中获得一定程度的认知量,特定用户对知识的摄取存在先验的认知总量,比如,特定用户在一次查询中仅仅希望了解“效用衰减”理论,而不涉及具体实验和结果,那么“效用衰减”理论的定义、解释与分析即为当前用户的摄取总量,而非一篇完整的论文. 在摄取知识的过程中,特定检索结果提供的认知量,将削弱用户继续查询与检索的欲望,直至用户摄取到自我先期谋定的足量的认知水平,欲望衰减为零,检索过程即会停止. 这一过程中,用户每摄取一定认知量,新的查询结果能提供的认知量将衰减,简称“效用衰减”. 这一理论解释了用户迭代进行查询重构、浏览和点击浏览的根本原因. 同时,该研究也在认知层面上给出了“度”的概念,对于解释本文用户满意度及其演化的思想起到借鉴作用.

除了点击行为,前人还对诸如视觉等其它特征

进行了深入探索. 其中, Buscher 等人<sup>[18]</sup>检测并采集用户在检索过程中的视觉焦点特征, 并通过计算和统计视觉焦点的分布密度, 判定视觉焦点锁定的检索结果的相关性. 然而, 高性能的视觉采集设备往往价格昂贵, 在实际应用中并不普及. 此外, 计算机系统携带的便携式摄像头往往精度较差, 难以有效采集可靠的视觉信息. 更为重要的是, 视觉信息采集对操作环境和对象有着较高的要求, 抗干扰能力略低, 无法有效支持各类复杂环境下的视觉分析.

作为一种替代方法, 基于光标跟踪的用户行为分析与应用具有更高的可靠性和实用性. 研究基础起源于光标在屏幕上的位置与用户视觉覆盖区域的关联性<sup>[19]</sup>. 具体地, Chen 等人<sup>[20]</sup>证实视觉覆盖和焦点往往与光标的位置有着较小的距离差异. Huang 等人<sup>[21-22]</sup>证明, 现有检索行为特征之中, 视觉焦点与光标位置的关联性最为明显. Guo 等人<sup>[23]</sup>进一步证明, 光标位置与视觉焦点在半径为 100 个像素点的范围内, 偏差概率低于 77%, 并且纵向偏差低于横向偏差. 光标与视觉焦点在运动过程中的协同规律, 支持了利用肢体驱动的设备反馈跟踪和应用于用户行为特征的信息检索研究. 其中, Guo 等人<sup>[24]</sup>提出了后置点击行为(Post-Click Behavior)模型, 这一模型融合了多种行为特征和设备反馈, 包括光标的停滞与延迟时长、移动、交互模式与区域、翻页与滚动、上下文和顺序等特征.

综上所述, 现有利用行为特征的信息检索研究详细分析了认知、行为和反馈之间的关联性, 并将这一关联性应用于信息与用户需求和意图的相关性判定, 能够有效辅助用户偏好的发现. 但是, 对于用户在特定检索过程中呈现的偏好程度, 以及对符合偏好的同类检索结果给出的不同满意程度, 则难以有效区分. 详细而言, 仍存在如下问题:

(1) 相关度是否能够表征满意度, 即是否相关信息必然满足用户的检索需求?

(2) 如果满意度能表征用户对同类检索结果的不同价值评判, 什么特征能够有效探测满意度?

(3) 如果行为特征有助于相关性判定, 是否也能够辅助满意度判定, 如何使用和建模?

## 3 满意度假设与量化分析思想

### 3.1 检索满意度假设

围绕现有研究中暴露的上述问题, 本文围绕用户端的检索体验, 提出满意度这一全新的检索结果

评价标准. 其动机来源于如下基本假设: (1) 相关性为信息在内容和语义层面的一致性、可比较性或关联性, 可表征特定信息呈现查询目标和意图的能力; (2) 相关性仅仅是评价检索体验的一元, 源自信息端, 无法涵盖用户端植根于用户认知体系的主观体验; (3) 用户端的主观体验是一种独立的全新的评价标准, 往往需要脱离信息本身的价值而依赖不同用户独特的认知体系进行判断.

基于这一假设, 本文将用户端用户的主观检索体验定义为“满意度”, 其具有如下特点:

(1) 满意度仅仅注重用户审视信息价值的视角, 忽略信息本身的相关性. 换言之, 与当前查询相关度较低的检索结果依然存在符合用户满意度标准的可能性, 这一特征可用于解释用户查询和浏览过程中的目标转移与意图切换.

(2) 满意度反映用户自身具有的认知体系和价值权衡方式, 信息质量的高低需要考量用户的意识、品味、视角和直观感受等特征, 其量化标准、数值计算和数学建模均需要围绕这类特征予以实施, 而不能将信息本身的特征作为主要判定依据. 例如, 文本检索中涉及的关键词和图像检索中涉及的色彩与轮廓, 皆为信息本身的特征, 而视觉跟踪中对人类眼球收缩机能的描述与量化则属于刻画满意度的主观特征.

(3) 相关度与满意度是两种不同的量化指标, 不等价且无交叉, 在实际检索场景中应独立测量、分析和应用. 比如, 不能判定一个文字层面相关度略低的文本检索结果同样具有较低的用户满意度, 实际情况往往恰恰相反.

满意度与相关度在信息获取研究和应用领域有着截然不同的意义, 本节以电子商务中的信息推荐为实例场景, 分析两者在结果推送中的差异.

图 1 显示了一组针对查询“新款皮鞋”的信息检索与推荐结果. 其中, 图 1(a)表示的是利用通用检索系统反馈的初始检索结果, 并以平铺的方式予以呈现. 在这一基础上, 假设用户针对子图 1(a)中的 2 号图片进行点击, 表示该图片引起了用户的关注. 那么, 利用这一用户反馈为特征, 基于相关度和满意度进行检索结果重排序, 可获得两种截然不同的排序结果. 图 1(b)即为根据相关度进行结果筛选和重排序的结果, 表现为一组“新款女鞋”, 显示了相关度在识别属性相同和类别相同方面的优势.

相对地, 图 1(c)为根据满意度进行重排序和结果推荐, 可获得一系列“萝莉商品”, 比如, “萝莉春



图 1 “满意度”和“相关度”的应用对比

装”、“萝莉熊”和“萝莉家装”等,并不包含“女鞋”类的反馈信息.上述例证显示,用户对检索结果有着多样性的评价标准,比如类别与属性的一致性,以及品味和视角的一致性.而显然地,相关性更趋向挖掘和组织属性类别一致的信息检索进行推荐,从而推荐集中于“女鞋”类的信息;满意度则趋向探测用户对色彩和款式方面的品味与审美观点(“粉色萝莉”),从而推荐集中于“粉色萝莉”涉及的范畴,并不局限于特定类别与目标.因此,满意度有助于构建全新的信息检索和推荐模式.尤其,通过联合使用相关度和满意度,可形成既关联于目标类别又符合主观认知特点的综合检索结果,如图 1(d)所示,为一系列“粉色萝莉”款式的“新款女鞋”.

### 3.2 满意度的滑鼠量化分析思想

由于满意度植根于用户主观认知和价值体系,而非围绕信息的属性和外部关系,因此,针对信息满意度的处理手段需要摒弃传统以信息为中心的方法,转而围绕用户端的主观特性进行挖掘和学习.由此,本文提出一种基于滑鼠行为和动量的满意度测量方法.

滑鼠特指鼠标滑行与运动,受人类上肢(手臂与手部)控制,是配合视觉驱动光标滑行与驻停的根本

条件.作为一种运动形式,滑鼠可通过基本的物理学原理进行数值计算和数学建模,形成可计量的标准化数据.其中,与本文直接相关的量化数据是滑鼠的动量,包括滑鼠动能(Kinetic Energy, KE)和角动能(Angular Kinetic Energy, AKE)两种形式及其合成.

本文将这类能量作为间接反映用户满意度的量化的关键性数据,其动机与原理如下:满意度来源于主观认知,由复杂的脑部运动形成,对其直接探测具有一定难度.但是,人们脑部运动的变化往往来源于外部信号的刺激,且在肢体和生理层面产生相应的条件反射.比如,强光刺激可触发瞳孔缩小和上肢遮挡眼球的系列反应.就满意度而言,令用户满意(或极大触发获取欲望)的信息往往调动多样的生理反应,比如视觉驻停、瞳孔扩展、面部肌肉收紧、肾上腺素增加和肢体运动剧烈等等,而这类生理与肢体反应都是受脑部运动活跃性驱动,或称之为脑部响应外界刺激的条件反射.相比于探测和学习脑部复杂的运动形式与原理,采集和分析生理和肢体反应更为直接,恰如在实验室环境观测猿类对香蕉产生的反应一样直观而有效.尤其,对生理和肢体反应的采集与分析装置已在生物信息学领域得以设计和

使用。比如, 针对视觉和瞳孔运动<sup>[25-26]</sup>的分析研究, Google 和 Microsoft 等公司与研究机构已经设计并研制了视觉采集仪器<sup>[27]</sup>。针对信息检索而言, 用户与信息之间的主要媒介是由计算机装置为载体的操作系统, 对系统控制的主要途径是由人类上肢驱动外部设备, 鼠标即为其中一种。从而, 用户对信息的反应形式可以通过{脑部、肢体、鼠标和光标}形成的相互协调的行为予以展现。从而, 滑鼠运动也可定义为一种“肢体反映思维”的间接认知与意识反馈。尤其, 滑鼠运动远比表情、视觉和内分泌等数据更容易采集和分析, 所需的装置仅仅为通用且价格低廉的机械、光电和触摸式鼠标。因此, 本文将用户满意度的分析与测量集中于滑鼠这一间接的肢体反馈进行研究。

基于滑鼠运动能量的满意度测量遵循如下基本思想: 针对视觉锁定的满意信息, 用户的脑部运动将更加活跃, 与脑部运动相互协调的肢体反应也将更为剧烈, 比如手部推拉鼠标的力度与速度, 而光标直接受手部驱动, 其点击目标和运动至目标的轨迹都能反应手部行为的特征。以相反的顺序推理, 利用光标起点与点击点之间的距离、运动的加速度和角度在物理学范畴内进行量化计算, 所获得的能量指标可以等比地反映肢体行为的能耗(尽管这一能耗较为细微, 但仍然存在), 借以反映肢体的条件反射强度, 那么根据肢体反射与脑部运动的协调关系, 滑鼠能量也可等比地模拟脑部运动的活跃性。那么, 如果满意度是一种脑部运动的能量, 则滑鼠的能量可作为间接的测量手段, 量化地模拟满意度的高低。

总而言之, 来源于脑部运动的满意度与受肢体运动驱动的滑鼠能量形式, 存在如下相互转化的过程: (1) 外部信息通过视觉刺激大脑; (2) 脑部运动形成满意度意识; (3) 满意度意识决定肢体条件反射的强度; (4) 肢体响应意识并实现条件反射, 进行能耗等比的控制行为; (5) 鼠标接受肢体控制, 触发能耗等比的滑鼠行为; (6) 滑鼠运动过程呈现的能量形式可作为量化指标反映脑部运动的活跃性, 进而刻画满意度强弱。

根据上述动机与思想, 本文通过满意度标注系统, 收集了 4232 套滑鼠样本和相对应的满意度评分。利用物理学和机器人学的基本常识与模型, 形成了滑鼠运动的量化计算方法。实验重点集中于检测和验证滑鼠能量与满意度之间的协调关系, 分析利用行为动量仿真满意度的可行性。

## 4 滑鼠行为的量化方法

如前文所述, 满意度是由人脑部对事物的能动反应而产生的, 直观地反映人们对事物或信息的认可程度, 其量值由脑部活动的剧烈程度而决定。然而, 衡量脑部活动的剧烈程度极其困难, 寻找一种替代方案成为解决这个问题的可靠思路。随着人机交互技术的发展, 键盘与鼠标成为当今网络用户与计算机进行交互的主要形式。尤其是鼠标, 其简单直观的操作方式充分体现了机器感知用户在可视化界面中获取信息的具体过程。进一步地, 用户在获取信息过程中反映的心理及脑部活动自然地通过手部运动进行传递, 从而使得机器能够感知用户的信息需求。根据人机交互学的理论研究, 用户对鼠标的操作具有两种不同的运动状态: 平滑状态及转动状态。物质运动过程中, 以上的两种状态同时伴随着不同的物理量变化, 即动能与角动能变化。这两种能量来源于用户脑部活动对其神经系统的直接反馈, 同时直观地反映了用户对于所浏览信息的满意程度。因此, 本研究将通过计算用户在信息检索过程中产生的能量波动, 进一步地分析用户对所检索到信息的满意程度。下文内容将详述两种能量的产生过程及计算方法。

下面将通过分析用户在交互界面中的鼠标真实移动轨迹来具体阐述两种能量(动能和角动能)的计算方法。图 2 为用户在一个交互界面中的真实鼠标滑动轨迹, 图中数字表示每段滑鼠轨迹编号。

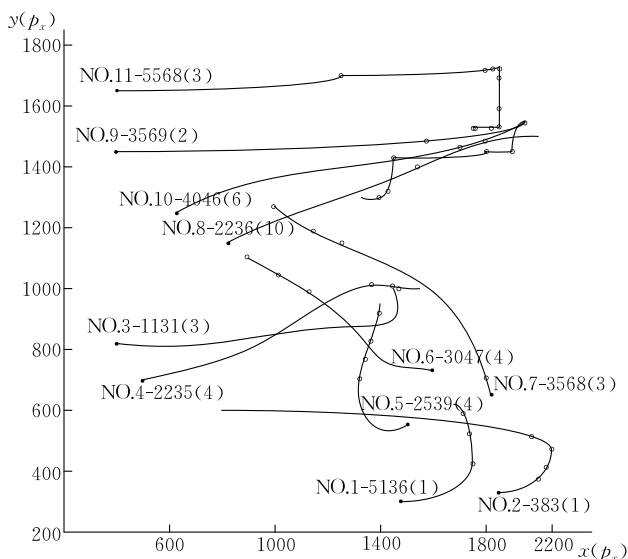


图 2 一个交互界面中鼠标移动轨迹

#### 4.1 直线运动过程中动能计算分析

如前文所述,鼠标的滑动过程具有平滑和转动两种状态.因此,可以抽象出两种运动方式,即直线运动和曲线运动.在直线运动过程中,鼠标的动能变化易于计算,具体方法如图3所述.

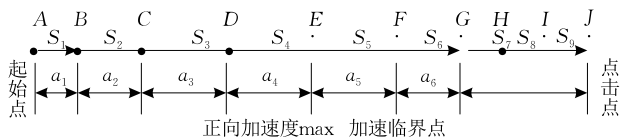


图3 直线运动状态示意

图3所示为直线运动状态的示意图.图中每个点(A、B、C…)均表示光标在当前时刻所在位置,相邻两个点间的时间间隔固定已知(由实验平台决定,下文将详细叙述该实验标注平台的原理及语料标注方法).由于鼠标滑动过程中,相邻两点之间的距离相对较小,且时间较短.因此,本实验将相邻两个点之间的运动状态看作是匀变速直线运动,其运动过程中的物理量性质可认为与真实匀变速直线运动状态下相同.同时,图中相邻两点间的距离通过实验平台记录的滑鼠坐标可精确计算出,用 $S_1, S_2, S_3 \dots$ 表示.根据经典物理学公式,在一段匀变速直线运动中,若起始点的初速度、运动时间及运动距离已知,则可以计算出该段运动中,终点的速度、该段运动过程的加速度,从而进一步计算出该运动过程中,物体的动能变化情况.以上图3中AB段为例,具体计算方法如下:鼠标在AB段的运动中,点A是鼠标的初始位置,因此初速度为0,用 $v_0$ 表示( $v_0=0$ ),AB段的距离用 $S_1$ 表示,时间用 $t$ 表示( $S_1$ 与 $t$ 均由实验平台精确记录并测量得到,可认为是常量).因此,根据下述式(1)、(2)可计算出AB段运动中的加速度 $a_1$ 及鼠标匀加速运动至B点处的速度 $v_b$ .

$$v_b = v_0 + a_1 \times t \quad (1)$$

$$s_1 = v_a \times t + \frac{1}{2} \times a_1 \times t^2 \quad (2)$$

式中: $v_0, S_1, t$ 已知,容易计算出 $v_b, a_1$ .

进一步地,由动能定理公式,如下:

$$\Delta E_{AB} = \frac{1}{2} \times m \times v_b^2 - \frac{1}{2} \times m \times v_0^2 \quad (3)$$

可计算出AB段运动中,光标滑动“消耗”的动能,这里将光标看做是质点,因此质量 $m$ 忽略不计.同理,BC段运动的初始速度为 $v_b$ ,距离为 $S_2$ ,时间同为 $t$ ,因此计算出BC段的动能变化量 $\Delta E_{BC}$ ,以此类推,计算出一段光标连续滑动过程中相邻两点间的动能变化 $\Delta E_{CD}, \Delta E_{DE}, \Delta E_{EF} \dots$ .

最后,通过累加相邻两点间的动能变化计算出整个连续滑动过程中的光标动能变化量,如下式(4):

$$E_k = \sum_{i=0} |\Delta E_{k_i}| \quad (4)$$

本研究将整段连续滑动过程中累加的动能变化量作为一次光标滑动轨迹中“消耗”的动能,作为表征满意度的一个测度.同理,由以上递推计算方法容易计算出角动能累加和用于作为表征满意度的一个测度.

#### 4.2 曲线运动过程中角动能变化分析

前文已述,光标的一次连续滑动轨迹中包含直线运动与曲线运动状态.直线运动状态中相邻两点间的光标移动可近似看作匀变速直线运动,其中包含着动能的变化.同理,光标的曲线运动状态包含着角动能的变化.

图4所示为光标移动轨迹中抽象出的曲线运动状态.由于网页信息的不规则分布,信息检索用户在从事信息检索活动时,光标的移动会呈现曲线运动状态,在此曲线运动状态中,通常伴随着角动能的变化,本节将以具体实例分析曲线运动过程中光标的角动能变化情况.

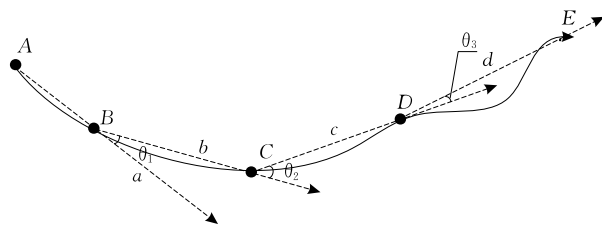


图4 曲线运动状态示意

如图4所示,图中的每个实心点同样代表当前时刻光标所在的位置(如A、B、C…),图4表示鼠标从A点移动到E点,其中相邻两点间的时间间隔固定已知(由实验平台决定).根据实验平台记录的光标所在点的实时坐标,计算出向量 $AB, BC$ 的夹角 $\Delta\theta_1$ ,用于表征光标从A点移动到B点旋转的角度,因此,类比角速度计算公式:

$$\omega_B = \Delta\theta_1 / t \quad (5)$$

可以计算出光标在B点处的角速度 $\omega_B$ ,同理,以此类推可以分别计算出光标在各个点处的角速度 $\omega_C, \omega_D, \omega_E, \dots$ .进一步地,计算光标在每个点处的角动能,公式如下:

$$E_\omega = \frac{1}{2} \times I \times \omega^2 \quad (6)$$

这里, $I$ 代表曲线运动中的转动惯量,通常为常量.因此,相邻两点间的角动能变化量可通过下式(7)计算:

$$\Delta E_{\omega} = \frac{1}{2} \times I \times E_{\omega_2}^2 - \frac{1}{2} \times I \times E_{\omega_1}^2 \quad (7)$$

同时, 类比动能的计算过程, 整个光标移动轨迹的角动能变化同样使用相邻两点间角动能的变化累加和来表征. 因此, 利用公式:

$$E_{\omega} = \sum |\Delta E_{\omega_i}| \quad (8)$$

计算出整条光标移动轨迹过程中的角动能变化累加和, 作为表征满意度的另外一个测度.

综上, 用户信息检索活动的光标移动轨迹中, 包含着不同状态的运动规律, 如直线运动状态和曲线运动状态, 详细分析两种运动过程, 抽象出其中的直线运动及曲线运动形式, 借助物理学经典公式计算出光标在移动过程中“消耗”的动能及角动能, 作为衡量用户检索满意度的测度, 下文将通过实验进一步论证两种能量作为衡量满意度测度的可靠性.

## 5 语料标注系统

### 5.1 系统描述及设计动机

本文研究旨在探测用户对检索结果的满意程度, 不同于传统信息检索的文本相关度, 其侧重于用户个人品味偏好对检索结果的影响. 前文已述, 鼠标滑动行为可以直观地反映用户对所接收信息的脑部活动状态, 从而得到用户对检索结果的满意性反馈. 因此, 采集用户在信息检索过程中的滑鼠行为成为该实验系统的核心任务.

实验系统是基于浏览器/服务器(Browse-Server, B/S)架构开发的交互式网站, 为了便于直观地反映用户对所浏览信息的兴趣度, 本实验系统选择人物图片作为信息载体. 语料标注人员登录到系统后对页面中的图片根据自己的满意度主观评分, 每张图片预设 1 到 10 个不同的分值. 实验系统记录下用户登录系统后的所有鼠标滑动及点击行为. 利用 JavaScript 脚本语言的采点机制, 实验系统在等间隔时间点内记录下光标的瞬时坐标, 形成光标移动轨迹(如图 2 所示), 同时将数据存储于后台数据库中, 用于光标移动过程中的能量“消耗”计算.

为了保证实验数据的真实有效, 本系统的语料采集目的及内容对标注人员保密, 保证了标注人员仅从主观满意性角度对图片评分, 作为其对检索结果满意度的标准.

### 5.2 语料标注方式

系统提供实时在线服务, 语料标注人员可以在任意时间段内通过计算机远程登录系统进行图片评

分. 为了提高用户对图片的评分热度, 系统设置不同的性别权限, 男性标注人员仅对“美女图片”进行评分, 女性标注人员仅对“帅哥图片”进行评分.

实验系统将不同类别的人物图片随机分布于交互页面中, 每张图片大小一致(格式化显示于每个交互页面中). 用户登录进系统后, 移动鼠标至感兴趣的图片缩略图上, 点击缩略图显示该图片的完整内容, 标注人员根据自己对该图片的喜好程度进行评分(点击图片下方星号, 共有 10 个), 评分结束后点击页面任意位置返回主界面, 至此为一次完整的语料标注过程. 实验系统记录下用户每次标注过程中的时间、鼠标移动轨迹、点击位置等信息. 用户在一个页面标注结束后, 可以点击翻页. 下一页面中将随机出现其余未标注过的人物图片. 标注人员可随时退出系统结束图片评分.

## 6 实验数据

### 6.1 图片数据

本研究的语料标注平台是基于图片评分的用户满意度标注系统, 因此, 图片数据是该系统的基本数据. 为了排除信息相关性对用户的直接影响, 本实验采用不同性别的公众人物图片作为信息载体, 共包含男女图片 10 404 张. 实验抓取了不同领域的明星人物图片, 按照男女比例接近 1:1 的自然分布嵌入标注平台中. 为了提升标注人员参与实验标注的体验, 实验过程中事先将从网络中抓取的图片人工过滤, 保留其中内容清晰、无重复且易于辨认的数据. 同时, 考虑到不同用户个性化偏好及兴趣点不同, 图片数据不均衡分布于包括体育明星、新闻主播、歌手、电影演员等共计 24 个领域内.

### 6.2 滑动轨迹数据

本实验的核心内容是基于用户的鼠标滑动轨迹, 计算其滑鼠过程中“消耗”的动能及角动能, 由此分析并度量用户对不同信息的满意度. 因此, 鼠标滑动轨迹数据是本实验的核心研究数据.

前文已述, 语料标注平台记录下标注人员登录系统后等时间间隔内鼠标所在的瞬时坐标. 其中, 若用户在某时刻产生鼠标点击行为, 系统将记录下用户的所点击图片的编号以及其主观评分, 具体格式如表 1 所示.

表 1 所示为某标注人员在一个标注页面中的部分鼠标滑动轨迹数据, 其中第 1 列  $x$  表示在某一时刻鼠标在页面上的横坐标信息; 第 2 列  $y$  表示某时



刻鼠标在页面上的纵坐标位置,根据横纵坐标信息,可以计算出鼠标相邻两个时间间隔内移动的水平位移、垂直位移以及滑过的距离,用于后续“消耗”的能量值计算;第 3 列 *sec* 指的是鼠标当前所在位置距离上一位置经过的时间(单位:ms),由于平台预设等间隔采点时间为 8ms,因此超过 8ms 的时间差即表示鼠标在此位置产生停顿,用于后续过滤噪音数据;第 4 列 *start* 是二值的标记数据,*T* 表示当前位置为鼠标起始位置,*F* 表示非起始位置;第 5 列

*click* 同为标记位,−1 表示鼠标在当前位置未产生点击,否则其具体数值表示当前点击的图片编号(如表 1 中的 182);第 6 列 *user id* 表示当前用户的编号;第 7 列 *time* 表示该用户登录系统起始时间,结合第 3 列时间差,可精确计算出鼠标在任意位置的具体时刻;第 8 列 *page id* 表示当前标注页面的编号,为了排除不同页面间的轨迹干扰,本实验将针对在同一标注页面中的多次点击构建实例对,分析滑鼠能量与满意度之间的关系。

表 1 滑鼠行为轨迹数据

<i>x</i>	<i>y</i>	<i>sec</i>	<i>start</i>	<i>click</i>	<i>user id</i>	<i>time</i>	<i>page id</i>
1319	213	0	<i>T</i>	−1	42	2014-4-2 21:07:15	a19b9cgfbc62db
1319	215	8	<i>F</i>	−1	42	2014-4-2 21:07:15	a19b9cgfbc62db
1320	304	8	<i>F</i>	−1	42	2014-4-2 21:07:15	a19b9cgfbc62db
...	...	...	...	...	...	...	...
1321	309	124	<i>F</i>	182	42	2014-4-2 21:07:15	a19b9cgfbc62db

### 6.3 能量数据

本文旨在验证用户检索满意度独立于相关度,由用户个性化的兴趣偏好及认知偏置产生,并首次提出使用滑鼠过程中“消耗”的能量来有效度量对不同检索结果的满意性差异.因此,鼠标移动轨迹中计算出的能量数据是本实验的直接分析数据.表 2 所示为一个标注页面中,过滤噪声后五次有效滑动点击轨迹对应的能量计算结果及所点击图片评分数据.第 1 列为对应点击图片评分;第 2 列为该条移动轨迹中“消耗”的动能量值;第 3 列表示该条移动轨迹中“消耗”的角动能量值.

表 2 一个标注页面中能量计算结果

<i>SCORE</i>	<i>KE</i>	<i>AKE</i>
7	0.0009	0.1972
8	0.1228	0.2967
4	0.0395	0.2989
7	0.0129	0.0199
7	0.2640	0.4502

为了排除不同页面中鼠标滑动轨迹的相互干扰,本实验将在同一标注页面中构建数据实例分析满意度与能量值之间的关系,后文将详述.

本实验参与标注语料的人员共有 60 名,标注周期为 1 个月.据统计,所有标注人员产生的有效滑动点击轨迹为 35 897 条,即平均每名标注者标注 596 张图片.共得到有效标注页面 1672 个,平均每名标注者对 28 个页面进行评分标注.

## 7 实验结果及分析

当前,搜索引擎检索结果主要以多反馈列表的

形式呈现,并结合查询相关度进行优化排序,检索结果的优劣评价核心在于多结果间的相关性程度对比.同样,检索结果的满意性度量及评价应聚焦于多结果的满意度高低对比.为此,本实验设置两种不同的数据实例:基于二元组的实例以及基于多元组的实例,为满意性的度量提供量化的分析方法.

基于二元组的实例分析,旨在通过构建成对结果的对比实例样本,分析两两滑动点击轨迹中“消耗”的动能与角动能与用户主观评分之间的大小关系来说明动能及角动能对量化满意性大小的有效性及可行性.表 3 表示一个二元组数据样例.其中,*NO* 表示编号;*SCORE* 表示用户对图片的主观评分,直观地表示用户对该图片的满意度;*KE* 表示计算出的动能量;*AKE* 表示计算出的角动能量.以表 3 中的二元组实例为例,分别比较动能和角动能与用户评分的大小关系,若一致则置此实例为正例,否则为负例.本实验采用正例的准确率来衡量动能/角动能在度量满意度过程中的不同性能,得到表 4 的实验结果.

表 3 二元组实例

<i>NO</i>	<i>SCORE</i>	<i>KE</i>	<i>AKE</i>
#1	7	0.0009	0.1972
#2	8	0.1228	0.2967

表 4 两种能量准确率对比

<i>ENERGY</i>	<i>ACCURACY RATE</i>
<i>KE</i>	0.5483
<i>AKE</i>	0.4517

表 4 中第 1 列表示两种能量;第 2 列为其对应的正实例准确率.从表中数据可以看出,动能及角动

能单独度量满意度存在一定的可行性,且使用动能准确性高于角动能.但是,单独使用两种能量中的一种仅能得到 50%左右的实验性能.为了提高实验性能,本研究进一步分析两种能量在度量满意度过程中的差异性,以期寻找有效的融合方法,为满意度度量提供更加合理有效的方法.

为了分析两种能量对于满意度度量的差异性,本研究提出错误相似度计算方法,如式(9)所示:

$$FL(KE, AKE) = \frac{|F(KE, AKE)|}{|F(KE)| + |F(AKE)| - |F(KE, AKE)|} \quad (9)$$

其中:  $|F(KE)|$  表示仅使用动能判断满意度大小错误的实例数;  $|F(AKE)|$  表示仅使用角动能判断满意度大小错误的实例数;  $|F(KE, AKE)|$  表示两种能量均不能有效判断的实例数.该错误相似度表示两种能量在度量满意度中的一致程度,其值越小表示两种能量之间的一致性越差,差异性越小.实验结果发现,两种能量在度量满意度中的一致性仅为 0.0734.进一步地,统计单独使用动能判断正确的实例集合中,有 23.20%的实例角动能判断错误,统计单独使用角动能判断正确的实例集合中,有 23.20%的实例动能判断错误.以上分析说明,两种能量在衡量满意度过程中具有不同角度的贡献.由此,寻找一种合理的融合方法融合两种能量,必然可以进一步提高两种能量度量满意度的准确性.

为此,本实验从用户浏览信息的行为习惯角度分析两种能量产生差异的原因.图 5 和图 6 分别表示用户在标注页面中点击图片的分布以及数据统计,图中的不同象限分布表示用户从鼠标起始滑动至点击图片旋转过的角度,图 5 中不同圆的半径表示用户对所点击图片的评分.从图中看出,用户滑动

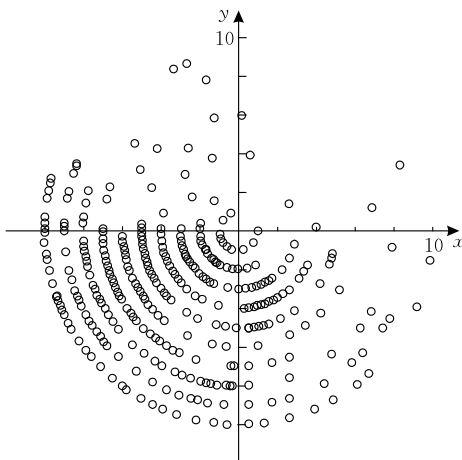


图 5 点击点象限分布

点击旋转过的角度集中于第三象限.由此说明,用户在浏览页面过程中旋转过的角度呈现一定的规律,且用户在不同满意度下旋转角度呈现不同的变化.因此,本实验进一步对比在多个实例中,在不同的用户满意度下,动能及角动能呈现的规律.尝试使用规则过滤的方法来融合两种能量.

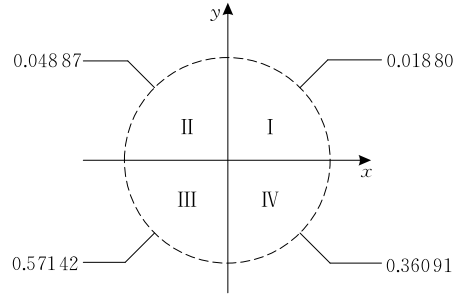


图 6 点击点象限分布统计

图 7 和图 8 分别表示多个实例中分数与动能、分数与角动能对比趋势图,两个图中的实例一一对应.以实例 2 为例,第 1 个点击点的动能值高于第 2 个点,其分数对比关系与角动能一致;实例 8 则相反,第 1 个点的动能值低于第 2 个点,其分数对比关系与角动能一致.通过大量观察数据发现,二元组中前后两个点不同的大小关系,其对应的动能与角动能与分数对应关系也不同.

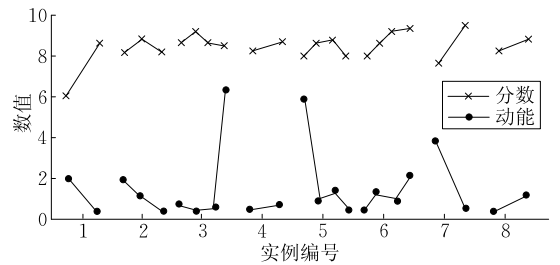


图 7 分数与动能对比趋势

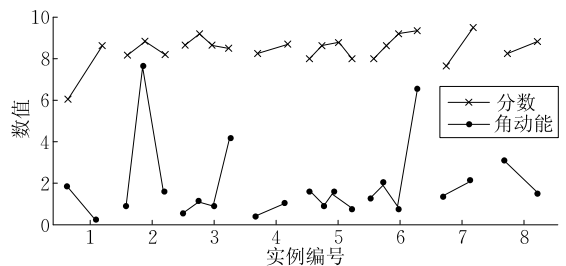


图 8 分数与角动能对比趋势

因此,根据上述现象,本实验通过人工设定规则来进一步量化满意度.规则 1:若前一个点动能高于后一个点,则使用动能度量满意度;反之,使用角动能度量.规则 2:若前一个点动能低于后一个则使用动能度量满意度;反之使用角动能度量.实验结果如

表 5 所示,从表中数据看出,使用规则 1 度量满意度明显高于规则 2,且使用规则融合的方法性能超过单独使用一种能量度量满意度方法,因此,本研究为用户检索满意度预测提供了有效的量化标准。

表 5 不同规则下满意度度量精确度

Rule	ACCURACY RATE
R1	<b>0.72</b>
R2	0.35

## 8 总 结

目前,对检索性能的评价仅停留在信息相关性及本源正确性的基础上,并未涉及用户主观认知和价值体系下检索结果质量的评判标准和度量方法。本文前瞻性地提出满意度是一种衡量检索性能的新标准。实验结果证明了满意度的客观存在性。在此基础上,本文尝试利用滑鼠行为和能量形式模拟源自脑部活动的满意度。实验证明了这一方法在模拟满意度计量和分布研究中的可行性。尤其,滑鼠行为的动能与角动能在度量检索结果满意度上具有各自的优势。

未来工作将集中于提出合理的机器学习模型,其研究重点将围绕融合动能与角动能,形成面向单一目标的满意度等比计量方法,并在此基础上,通过对序列点击行为中滑鼠行为及其能量变化的趋势,模拟用户检索过程中的满意度演化模式,形成针对特定用户序列检索行为的检索结果重排序模型。比如 CRF,该模型是一种全局线性模型,可以在全局数据中抽取特征,使得构建的模型充分考虑用户在多次点击之间的相互作用。

**致 谢** 苏州大学朱巧明教授和姚建民研究员对这一研究给予了长期支持与资助。国家自然科学基金委和教育部以及江苏省和苏州市自然科学基金委对本文研究给予了长期支持。在此一并感谢!

## 参 考 文 献

- [1] Manning C D, Raghavan P, Schütze H. Introduction to Information Retrieval. Cambridge: Cambridge University Press, 2008
- [2] Jansen B J. The effect of query complexity on Web searching results. Information Research, 2000, 6(1): 6-1
- [3] Liu Run-Ran, Jia Chun-Xiao, Zhou Tao, et al. Personal recommendation via modified collaborative filtering. Physical A: Statistical Mechanics and its Applications, 2009, 388(4):

462-468

- [4] Hillard D, Schroedl S, Manavoglu E, et al. Improving ad relevance in sponsored search//Proceedings of the 3rd ACM International Conference on Web Search and Data Mining. Banff, Canada, 2010: 361-370
- [5] Cao Z, Qin T, Liu T Y, et al. Learning to rank: From pairwise approach to listwise approach//Proceedings of the 24th International Conference on Machine Learning. Banff, Canada, 2007: 129-136
- [6] Qiu Y, Frei H P. Concept based query expansion//Proceedings of the 16th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. Pittsburgh, USA, 1993: 160-169
- [7] Hussein S E, Granat M H. Intention detection using a neuro-fuzzy EMG classifier. IEEE Engineering in Medicine and Biology Magazine, 2002, 21(6): 123-129
- [8] De Gemmis M, Iaquinta L, Lops P, et al. Preference learning in recommender systems//Proceedings of the ECML/KDPP-09 Workshop on Preference Learning (PL 09). Bled, Slovenia, 2009: 41-56
- [9] Hong Yu, Kang Yang-Yang, Yao Jian-Min, et al. Discovery and illustration of novel optimal retrieval result. Chinese Journal of Computers, 2013, 36(3): 643-653(in Chinese)  
(洪宇, 康杨杨, 姚建民等. 一种新型最优检索结果的发现与论证. 计算机学报, 2013, 36(3): 643-653)
- [10] Commentz-Walter B. A String Matching Algorithm Fast on the Average. Berlin: Springer Berlin Heidelberg, 1979
- [11] Craswell N, Zoeter O, Taylor M, Ramsey B. An experimental comparison of click position-bias models//Proceedings of the 2008 International Conference on Web Search and Data Mining. Banff, Canada, 2008: 87-94
- [12] Joachims T, Granka L, Pan B, Pan B. Accurately interpreting clickthrough data as implicit feedback//Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. Bermen, Germany, 2005: 154-161
- [13] Richardson M, Dominowska E, Ragno R. Predicting clicks: Estimating the click-through rate for new ads//Proceedings of the 16th International Conference on World Wide Web. Tokyo, Japan, 2007: 521-530
- [14] Guo Q, Agichtein E. Towards predicting Web searcher gaze position from mouse movements//Proceedings of the CHI'10 Extended Abstracts on Human Factors in Computing Systems. Harbin, China, 2010: 3601-3606
- [15] Dupret G, Murdock V, Piwowarski B. Web search engine evaluation using click-through data and a user model//Proceedings of the WWW2007 Workshop Query Log Analysis: Social and Technological Challenges. Banff, Canada, 2007
- [16] Guo F, Liu C, Wang Y M. Efficient multiple-click models in web search//Proceedings of the 2nd ACM International Conference on Web Search and Data Mining. Stockholm, Sweden, 2009: 124-131

- [17] Dupret G, Liao C. A model to estimate intrinsic document relevance from the clickthrough logs of a web search engine// Proceedings of the 3rd ACM International Conference on Web Search and Data Mining. Harbin, China, 2010: 181-190
- [18] Buscher G, Cutrell E, Morris M R. What do you see when you're surfing?: Using eye tracking to predict salient regions of web pages//Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Stockholm, Sweden, 2009: 21-30
- [19] Rodden K, Fu X. Exploring how mouse movements relate to eye movements on web search results pages//Proceedings of the Web Information Seeking and Interaction Workshop. Amsterdam, The Netherlands, 2006: 29-32
- [20] Chen M C, Anderson J R, Sohn M H. What can a mouse cursor tell us more?: Correlation of eye/mouse movements on web browsing//Proceedings of the CHI' 01 Extended Abstracts on Human Factors in Computing Systems. Seattle, USA, 2001: 281-282
- [21] Huang J, White R, Buscher G. User see, user point: Gaze and cursor alignment in web search//Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Warsaw, Poland, 2012: 1341-1350
- [22] Huang J, White R W, Dumais S. No clicks, no problem: Using cursor movements to understand and improve search// Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Orlando, Florida, 2011: 1225-1234
- [23] Agichtein E, Brill E, Dumais S. Improving web search ranking by incorporating user behavior information//Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. San Antonio, Texas, 2006: 19-26
- [24] Guo Q, Agichtein E. Beyond dwell time: Estimating document relevance from cursor movements and other post-click searcher behavior//Proceedings of the 21st International Conference on World Wide Web. Warsaw, Poland, 2012: 569-578
- [25] Ding Yan, Wang Yu-Ping. Emotional stimulus in visual cognitive processing. Chinese Journal of Clinical Rehabilitation, 2006, 9(28): 152-155(in Chinese)  
(丁岩, 王玉平. 情绪信息刺激对视觉认知过程的影响. 中国临床康复, 2006, 9(28): 152-155)
- [26] Wang Lu, Yang Ling, Ge Yan, et al. The development and application of lie detection in forensic science. Journal of Forensic Medicine, 2008, 24(5): 365-368(in Chinese)  
(王璐, 杨玲, 葛燕等. 测谎的发展历程及法医学应用前景. 法医学杂志, 2008, 24(5): 365-368)
- [27] Oliver C J. Visual display for communication network monitoring and troubleshooting, USA, 1989-12-26



**HONG Yu**, born in 1978, Ph. D., associate professor. His research interests focus on personal information retrieval, topic detection and tracking, information extraction and opinion mining.

**WANG Jian**, born in 1990, M. S. candidate. His research interests focus on personal information retrieval.

**WANG Kai**, born in 1990, M. S. candidate. His research interests focus on personal information retrieval.

**KANG Yang-Yang**, born in 1989, M. S. candidate. His research interests focus on personal information retrieval.

**KONG Fang**, born in 1977, Ph. D., associate professor. Her research interests focus on information extraction and natural language processing.

**YAO Jian-Min**, born in 1971, Ph. D., associate professor. His research interests focus on machine translation.

**ZHU Qiao-Ming**, born in 1963, Ph. D., professor. His research interests focus on Chinese information processing.

**ZHOU Guo-Dong**, born in 1968, Ph. D., professor. His research interests focus on natural language processing.

## Background

Current universal search engines focus on mining and recommending all possible relevant information to users. Further, the personal search engines are learning the real intention of users and recommending the information that extremely satisfies the personal requirements. But whether are all the relevant search results really make users satisfy? The answer should be not necessary. In detail, users tend to choose the retrieval results which fit their personal interest and taste.

The doubts mentioned above raise two questions:

- (1) what's the difference between the satisfaction and relevance?
- (2) How to measure the degree of satisfaction? So, in this paper, we propose a novel evaluation metric for information retrieval results, and we focus on the analysis of energy consumption during the travelling process of cursor, to measure the degree of satisfaction.

This research is supported by the National Natural Science Foundation of China (Nos. 61373097, 61272259 and 61272260).