

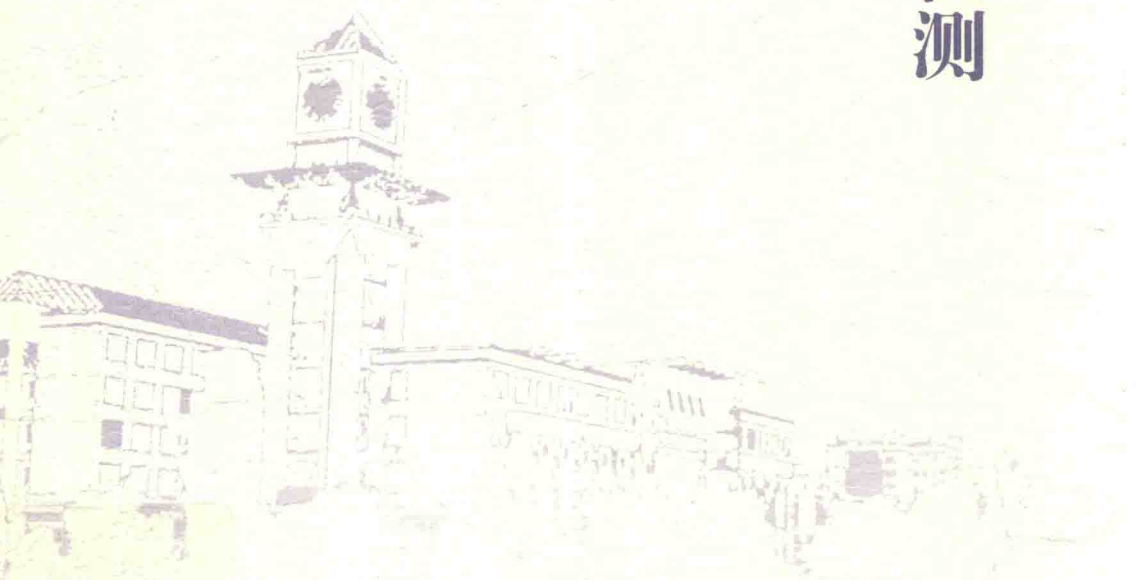


中南财经政法大学
青年学术文库

跨媒体网络事件检测 与跟踪研究

Cross-media Web Event Detection
and Tracking

张承德 ○ 著



中国社会科学出版社



中南财经政法大学
青年学术文库

跨媒体网络事件检测 与跟踪研究

Cross-media Web Event Detection
and Tracking

张承德 ○ 著

中国社会科学出版社

图书在版编目 (CIP) 数据

跨媒体网络事件检测与跟踪研究 / 张承德著. —北京:

中国社会科学出版社, 2018. 6

(中南财经政法大学青年学术文库)

ISBN 978 - 7 - 5203 - 2502 - 8

I. ①跨… II. ①张… III. ①情报检索 IV. ①G252. 7

中国版本图书馆 CIP 数据核字(2018)第 103404 号

出版人 赵剑英
责任编辑 徐沐熙
特约编辑 张 婉
责任校对 庞 彦
责任印制 戴 宽

出 版 **中国社会科学出版社**
社 址 北京鼓楼西大街甲 158 号
邮 编 100720
网 址 <http://www.csspw.cn>
发 行 部 010 - 84083685
门 市 部 010 - 84029450
经 销 新华书店及其他书店

印刷装订 北京君升印刷有限公司
版 次 2018 年 6 月第 1 版
印 次 2018 年 6 月第 1 次印刷

开 本 710 × 1000 1/16
印 张 13.5
插 页 2
字 数 170 千字
定 价 38.00 元

凡购买中国社会科学出版社图书,如有质量问题请与本社营销中心联系调换
电话:010 - 84083683

版权所有 侵权必究

《中南财经政法大学青年学术文库》

编辑委员会

主任：杨灿明

副主任：吴汉东 姚 莉

委员：（按姓氏笔画排序）

朱延福 朱新蓉 向书坚 刘可风 刘后振

张志宏 张新国 陈立华 陈景良 庞凤喜

姜 威 赵 曼 胡开忠 胡贤鑫 徐双敏

阎 伟 葛翔宇 董邦俊

主 编：姚 莉

序 言

感谢教育部人文社会科学研究青年基金项目：16YJC860026、跨媒体网络事件检测与跟踪研究，对本课题及本专著的支持与赞助，本书是该项目的研究成果。

辍笔在即，我是快乐的！虽然曾经失败过，迷茫过，放弃过，但每一次的转折都是我成长的动力，是我奔向成功的砝码。我的快乐成长离不开下面我要感谢的所有的人！细数这一路的点滴收获，除了源于自己的努力付出，更离不开你们给我的关心、帮助和支持。

首先，我要向我的导师吴晓老师表示崇高的敬意和衷心的感谢！本书中的大部分研究工作都是在吴晓老师的精心指导和严格要求下完成的，无论是在理论研究和实验方法的指导上，还是在最后的撰写、修改、定稿上，都凝聚着吴晓老师大量的心血。吴晓老师严谨求实的科研态度、敏锐的创新意识和无私的奉献精神深深地影响着我，对我在做人、处事和治学等方方面面的教导和影响使我受益匪浅，这将是我最宝贵的财富。

在整个科研过程中，我对吴晓老师言传身教的学者风格、平易近人的为人风范和科研上的钻研精神印象深刻。在学习上，吴晓老师对我精心地指导，倾囊相授、毫无保留。在日常生活中，吴晓老师事必躬亲，对我关怀备至。当我面临压力时，吴晓老师默默地关

注我，并及时地给我提供各种力所能及的帮助，让我感受着双重的爱与关怀。相比其他同学，我是最幸福的，也是最让师弟师妹们最羡慕的。正是在导师的指导、支持、鼓励和影响下，我才得以顺利完成本书，并且对科学研究产生了浓厚的兴趣。衷心地感谢我的导师，没有他的帮助，就没有我在科研道路上的成长。

非常感谢彭强和张家树老师提出的宝贵建议和意见，它们很好地弥补了我的研究中的一些不足之处，你们的激励也为我增添了科研动力，让我更有信心地开展论文的研究工作。

感谢彭强和陈俊周老师在本书完成过程中给予认真负责的指导。两位老师广博的学术造诣、敏锐的洞察力和严谨的科研态度是我学习的榜样。

非常感谢美国迈阿密大学的美籍华人徐美玲教授，为本书的完成提供了许多宝贵的建议和启迪。在美国迈阿密大学学习的1年中，徐美玲教授不管是在学习上还是生活上，一直都关照着，提点着我，鼓励着我，激励着我在科研的道路上不断进取。他谆谆教诲，严谨的科研态度使我受益匪浅，我将对此终生难忘。毕业在即，在此向徐美玲教授表示我最衷心的感谢！

感谢西南交通大学实验室的张庆明、范小九、张蕾、王广伟、陈洪瀚、孙广路、赵波、任健鹏、刘祥凯、高攀、闫川以及美国迈阿密大学实验室的刘典婷、朱秋莎、孟涛和黄安辉等博士研究生和陈川江、梁玲玲、鲁一杰、何俊彦、刘兆瑞、吴金鹏、张靓云、梁润亭、张宗峰、张扬阳、程治淇等硕士研究生，他们在学术和实验等方面给了我很多有益的建议和帮助。感谢实验室所有兄弟姐妹们带给我的美好回忆。

感谢我的父母和妻子肖霞，你们一直尊重我的选择，默默地支持我、帮助我、鼓励我，感谢你们一直以来对我的大力支持，正是你们的无私付出和积极鼓励给予了我战胜困难的信心和决心，才使

得我能够顺利完成本书。

最后，对所有在此尚未提及，但曾给予我关心和帮助的人们表示诚挚的谢意！

张承德

目 录

第一章 绪论	(1)
第一节 研究背景与意义	(1)
第二节 基本概念	(6)
一 大数据	(6)
二 信息检索	(11)
三 多媒体	(14)
四 多媒体检索	(18)
五 数据挖掘	(20)
第三节 国内外研究现状	(22)
一 话题检测与跟踪	(22)
二 特征轨迹	(30)
三 关联规则挖掘	(32)
四 多重对应分析	(34)
五 时空信息的事件挖掘	(36)
第四节 研究目标及拟解决的关键问题	(39)
一 课题研究目标	(39)
二 拟解决的关键问题	(40)
第五节 本书的主要研究内容及组织结构	(41)

第二章 基于视觉内容相关性与多重对应分析的网络视频

事件挖掘研究	(45)
第一节 视觉近似关键帧	(47)
第二节 共同发生与多重对应分析的融合	(50)
一 多重对应分析	(53)
二 共同发生	(55)
三 文本与视觉信息的融合	(57)
第三节 实验与分析	(57)
一 实验数据	(57)
二 实验分析	(59)
第四节 本章小结	(62)

第三章 基于视觉特征轨迹与文本分布特征的网络视频

事件挖掘研究	(64)
第一节 基于特征轨迹的事件挖掘	(66)
一 视觉特征轨迹	(66)
二 文本特征轨迹分析与研究	(68)
三 视觉特征轨迹分析与研究	(69)
四 文本与视觉特征轨迹的融合	(70)
第二节 实验与分析	(72)
一 实验数据	(72)
二 实验分析	(73)
第三节 基于内容的视觉特征轨迹与文本信息融合的 事件挖掘	(79)
一 文本相关性	(83)
二 基于内容的视觉特征轨迹	(87)
三 文本与视觉信息的融合	(92)

第四节	实验与分析	(93)
一	实验数据	(93)
二	实验分析	(94)
第五节	本章小结	(101)
第四章	基于动态关联规则与视觉近似片段的网络视频	
事件挖掘研究		(103)
第一节	基于动态关联规则的事件挖掘	(103)
一	动态关联规则	(105)
二	分类	(111)
第二节	实验对比与分析	(111)
一	实验数据	(111)
二	实验分析	(112)
第三节	基于视觉近似片段的事件挖掘	(117)
一	视觉近似片段	(120)
二	视觉近似片段与文本分布特征的融合	(123)
三	关联规则挖掘	(125)
四	文本信息挖掘	(128)
五	多模态融合	(130)
第四节	实验与分析	(131)
一	实验数据	(131)
二	实验分析	(132)
第五节	本章小结	(138)
第五章	结论与展望	(139)
第一节	主要研究成果	(139)
一	基于视觉内容相关性与多重对应分析的网络视频	

► 跨媒体网络事件检测与跟踪研究

事件挖掘	(139)
二 基于内容视觉特征轨迹与文本分布特征的网络视频 事件挖掘	(140)
三 基于动态关联规则与视觉近似片段的网络视频事件 挖掘研究	(140)
第二节 今后研究方向	(141)
一 文本突发性特征的研究	(141)
二 视觉突发性特征的研究	(141)
三 文本与视觉信息融合的研究	(142)
参考文献	(175)
发表论文	(196)
后 记	(199)

第一章

绪 论

第一节 研究背景与意义

最近几年，社交网络和网络视频网站的普及和快速发展，使得网络视频的数量以指数级的速度增长。随着互联网、计算机和移动技术的迅猛发展，互联网成为人们生活的重要组成部分。互联网网页、图像、网络视频和社交信息等不同类型数据的快速增长为用户提供了丰富的多媒体信息。2006年7月，用户每天上传6.5万个视频至视频分享网站优兔（YouTube），而且，每天的视频点击量超过1亿次。到了2009年9月，大约每分钟有20小时的新视频数据上传到优兔^①。2014年每月有超过10亿人的唯一身份用户访问优兔，用户每月在优兔上观看的视频总时长超过60亿小时，每分钟有长达100小时的视频上传到优兔，全球网站数量已超过10.6亿个。在国内，网页数量达1500亿，网站达320万个，中国网民数量超过6.32亿。优酷土豆月均覆盖用户数量超过5亿，日均播放

^① Zhang J, Fan X, Wang J, Et Al., "Keyword-Propagation-Based Information Enriching And Noise Removal For Web News Videos" In Proceedings of the 18th Acm Sigkdd International Conference on Knowledge Discovery And Data Mining (Kdd), Beijing, China, August 2012.

量超过7亿次。现代社会每天产生巨量的多媒体数据，全球已进入数据和信息爆炸性增长的时代——大数据时代（Big Data）。根据麦肯锡报告的定义，大数据是指其大小超出了典型数据库软件的采集、存储、管理和分析等能力的数据集。巨量不断更新的网络视频充分体现了大数据的特色：（1）数据量大：在互联网上，网络视频数量以亿为单位不断快速增长，且数据量异常庞大甚至有EB量级的数据需要及时的分析和处理；（2）快速应对：由于互联网数据和人们需求的快速变化，急切地需要本书能够及时根据需求做出快速准确的应对策略或动作以应对瞬息万变的外部环境，这不仅需要数据分析的速度快，而且对各种性能也提出了新的更高的要求，因此单位时间内要处理的数据量也非常大；（3）数据种类繁多：由于网络技术的飞速发展和数据类型的多种多样以及来源于不同媒体的数据越来越多，需要对这些多种数据源或多种数据类型进行必要的整理和筛选等操作；（4）数据的价值密度低：由于数据量非常庞大，且数据比较繁杂以及数据不完整等因素，数据的有效信息较少，甚至可能会失真，面对这样的情况可以通过大样本及多种类型的数据达到更真实地反映实际情况的效果。

在日常生活中，网络和有线电视等媒体的迅速发展，以及拍摄和存储多媒体设备价格的快速下降，使得视频等多媒体信息成为人们获取信息和享受娱乐的重要载体。由于多媒体所含信息具有信息量丰富、信息简单明了、易于理解、省时和省力等特点，普通用户不必再花费大量时间从大量文本信息中去寻找自己感兴趣的信息。这样，大量普通用户可以非常方便地通过手机、相机或者直接从网络上获取视频，然后经过编辑后上传到视频分享网站上。多媒体的上述特点使得信息检索变成了一种享受，而不再是耗时耗力的枯燥的事情。另外，国内外众多电视媒体为方便用户在线浏览视频，了解国内外重大事件，逐渐将大量视频节目通过互联网这种新兴媒体

进行传播。在中国过去的一段时间内，涌现出了一批视频分发网站如优酷、土豆等，并得到迅猛发展，这使得中国普通用户也可以很容易地获得数以百万计的网络视频。



图 1—1 在优兔上搜索“伦敦恐怖袭击”（London terrorist attack）时返回的搜索结果

目前，网络视频的管理方式依然比较简单，主要通过将视频分为几个通用的类型，例如：新闻、财经、体育、音乐、教育等。而最近比较流行的网络视频搜索引擎如：优兔、谷歌视频、雅虎视频和百度视频等，则主要通过搜索的关键词与视频中的标题或标签等文本信息进行相关度检测，搜索引擎返回的比较靠前的网络视频主要是关键字相似度或点击率较高的网络视频，但并不一定是用户所要检索的网络视频。这不仅使得自动挖掘网络视频中的主要事件更有意义，而且更方便普通用户浏览、搜索和并快速掌握事态的发展动态。

以话题“伦敦恐怖袭击”为例，如图 1—1 所示，当搜索关键词“伦敦恐怖袭击”时返回了 95000 个相关搜索结果，并且搜索结

果仅仅是根据其与搜索关键词的相关度进行排序的。然而，仅仅是一幅缩略图和寥寥几个文字很难对想要了解这一话题的用户有很大的帮助。面对海量的搜索结果，为了进一步了解该话题，用户需要点击更多搜索到的网络视频。甚至，需要收看大部分与该话题相关的视频，这不仅是非常耗时的而且仍然很难让用户找到他们真正想要的结果，尤其是对于他们来说完全陌生的话题。在现实生活中，用户通常只点击查看搜索结果的头几页，并不会耐心不断地查看全部搜索结果。很显然，现在的搜索引擎并不能帮助用户抓住主要事件，并从整体上了解整个话题。因此，面对搜索引擎返回的成千上万的，甚至不断增长的网络视频，普通用户往往显得无所适从甚至无能为力。这对本书研究如何从搜索引擎返回的海量网络视频中，快速有效地挖掘出主要事件提出了迫切要求。因此，有必要对搜索结果自动进行网络视频的相关性检测，并对相关的网络视频进行聚类，以减少用户的工作量，且有助于日后事件间相互关系建立的进一步研究工作。

然而，当本书将文本领域中比较成功的方法应用到网络视频时，遇到了新的问题和困难。这对基于初始关键字和视觉特征的事件挖掘提出了新的挑战。对于文本信息，一方面，网络视频仅仅有十几个单词来简要地描述视频的内容，相对于传统的文件拥有更少的文本信息；另一方面，由于用户间不同的习惯、不同的语言、不同的文化以及不同的教育背景等，使得不同用户描述同一内容时通常会使用不同的词语，甚至有些用户为了增加点击量会人为地增加一些与视频内容无关的热点词汇。对于视觉信息，一方面，由于不同用户的拍摄设备和水平参差不齐，不同用户上传的视频的质量便会差别较大；另一方面，有些用户会对相关视频做出相应的编辑比如删除部分内容，增加一些文字，甚至把某些人的头像换成自己的头像，等等。上述问题，无形之中为网络视频的事件挖掘增加了很大的

难度。

这就需要本书深入研究与分析网络视频所体现出来的特点，并充分利用文本与视觉信息的特征，以及这些特征间的相关性和互补性提出新的高效算法，达到高效地从海量网络视频中挖掘主要事件的目的。相对于传统文档，网络视频蕴含着丰富的视觉内容信息。同时，网络视频中存在着大量的近似图像或近似关键帧（Near-duplicate keyframe, NDK），即不同视频间相似或者相同的图像。它们可以将不同的网络视频链接在一起，不仅可以强调重要的场景，而且可以唤起对事件的回忆。为了更准确地获得视觉近似关键帧，目前，主要通过局部关键点（又称为局部点）对不同视频间的关键帧进行相似性检测。然而，由于关键帧的捕获条件、获取时间、图像变化（光线或亮度的变化），以及后期编辑（添加台标或字幕）等因素，搜索的结果中会出现内容相似但又存在一定差异的图像。以上现象往往会对视频间关键帧相似性检测造成一定的影响，从而使不同视频间相关性的研究更具有意义。

至今为止，不论是从信息的数量，还是从观点的数量上来看，互联网都成功压倒广播、报纸以及电视等传统媒体，日益成为引导社会舆论发生及发展的重要媒介。然而，海量网络视频中有效聚类敏感的热点事件，将对管理部门日后分析和提取舆情信息，甚至跟踪与监控这些信息都起到非常重要的作用。本书还可以通过将互联网中多种媒体的信息相结合的方式来进行事件挖掘，例如：利用互联网上的文本、图像以及视频等信息进行舆情监测以方便日后的重大事件跟踪，及为相关管理部门提供必要的参考信息，从而便于相关部门及时全面地了解整个事件的发展动态，及时处理正在发生的突发事件，并快速应对复杂且多变的事件。

综上所述，面对巨大的市场需求，海量网络视频的事件挖掘研究迫在眉睫，由此本书面临着许多挑战。本书的研究将有助于提高网络视

频搜索引擎的实用价值，同时节省普通用户大量时间，方便用户及时了解及掌握主要事件的发展脉络并对日后进一步的操作提供强有力的帮助和支持。

第二节 基本概念

一 大数据

在过去的 20 多年里，信息科技的发展日新月异，尤其是移动互联网的飞速崛起以及通信社交电子商务的风靡导致了数据量以无法估量的速度增长，大数据这一概念也就应运而生。所有领域似乎都不可避免地投身到大数据发展的浪潮中。大数据时代的悄然而至给当今社会带来了无限机遇，同时也附带着新的挑战 and 冲击。

大数据，简单来说，即传统软件工具无法捕获、存储、管理、分析的巨型数据的集合，是高速发展的信息时代的产物。在新时代的大数据背景下，大数据意义不在于“庞大”，而在于其中蕴藏的价值多少。和传统的数据集相比，大数据则需要通过更多的实时分析和加工处理充分挖掘其隐藏价值才能具有更强的洞察能力和决策能力。正如脸书（facebook）的副总工程师杰伊·帕瑞克所言，“如果不利用所收集的数据，那么你所拥有的只是一堆数据，而不是大数据”。

大数据具有 4 个 V 特征，即海量（Volume）、快速化（Velocity）、多样化（Variety）、价值（Value）。具体来说，（1）大数据体积巨大，常处于 PB 级别，其体积增长速度快，而且数据每时每刻都在更新变化。例如：一组名为“互联网上一天”的数据显示，一天内互联网产生的全部内容可以刻满 1.68 亿张 DVD；一天内发出的邮件达 2940 亿封之多；每天上传到脸书的照片堆起来有