

新闻视频故事单元关联分析技术研究综述

吴玲达¹ 文 军^{1,2} 陈丹雯¹ 袁志民¹

(国防科技大学信息系统与管理学院 长沙 410073)¹ (国防科技大学理学院 长沙 410073)²

摘 要 新闻视频故事单元关联分析是将新闻视频分割所得的故事单元按照报道的新闻事件所进行的一种特殊聚类。故事单元是新闻视频各层次结构中包含有新闻事件完整语义的结构单元,是对新闻视频进行分析、检索、组织等应用的最佳层次。对新闻视频故事单元关联分析技术进行了综述;按照方法所利用的信息类型将现有方法分成三类:基于文本相似度的方法、基于多模态信息相似度的方法和基于关键帧视觉重复性的方法;讨论了每种方法的特点并对相关技术的发展趋势进行了探讨。

关键词 新闻视频,故事单元,关联分析

中图法分类号 TP391 **文献标识码** A

Analyzing Relationship of News Video Stories: An Overview

WU Ling-da¹ WEN Jun^{1,2} CHEN Dan-wen¹ YUAN Zhi-min¹

(Department of Information Systems, National University of Defense Technology, Changsha 410073, China)¹

(Department of Science, National University of Defense Technology, Changsha 410073, China)²

Abstract Analyzing relationship of news video stories is a special clustering of stories based on news events. "Story" is the unit which contains full semantic information in hierarchy units. It is the best semantic unit for video analyzing, searching and organizing. This paper gave an overview of existing method and classified them into three categories: the approach based on text similarity, the approach based on multi-modal similarity and the approach based on keyframe visual duplicate. Their advantages and limitation were discussed and compared in detail. At last, this paper presented some trends in this field.

Keywords News video, Story, Analyzing relationship

1 引言

随着信息技术和网络技术的快速发展,新闻视频的来源日益广泛,数量日益庞大,获取更加便捷。每时每刻都有大量来自不同国家、不同语种、不同电视台以及不同新闻网站的大量新闻视频。在理论上,任何用户都可以随时访问自己需要的新闻信息;然而对于实际情况而言,为方便用户的检索和浏览,必须对新闻视频进行分析、组织和标注,目前这一系列的都完全依靠人工,是一项非常耗时的工作。

早期开展的视频数据库研究,主要利用人工输入各种属性,以支持数据的查询。但是,随着科技的发展,可获取的视频数据越来越多,数据库容量也不断增大,人工输入属性和标注的方法就暴露出自身的缺陷:一是人工标注需要大量人力;二是新闻视频包含了图像、音频和文字等各种类型的信息,有些类型的媒体信息在很多情况下难以用文字从各个层次标注清楚;三是用户对于新闻视频关注的重点是报道中包含的事件信息,一段新闻视频通常包括有多个事件的报道内容;同时,对于特定的重要事件,在一定时间范围内,又会出现在不同来源的多个新闻视频中。

当前研究通常把新闻视频结构层次化为帧-镜头-故事-视频,如图 1 所示。

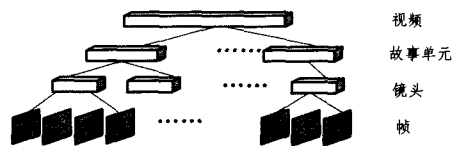


图 1 当前研究中新闻视频的层次结构

在上述的层次结构中,帧与镜头主要表示新闻事件的局部内容,单独的帧和镜头只能提供有限的语义信息,目前研究中对它们的分析主要基于底层的视觉特征。与之相比,故事单元是一个新闻事件的完整报道,能够提供关于事件的各种语义信息,因此围绕故事单元开展新闻视频研究不仅可在语义事件的层次上实现对新闻视频的组织和管理,也可结合故事单元及其包含的帧和镜头的各种信息,更加贴近实际应用中用户的各种服务需求。

2 新闻视频故事单元关联分析的任务

目前国内外新闻视频围绕故事单元所开展的研究绝大部

到稿日期:2009-07-09 返修日期:2009-11-09 本文受 863 国家高技术研究发展计划(2009AA01Z335),国家自然科学基金(60802080)资助。
吴玲达(1962-),女,博士生导师,主要研究方向为多媒体信息系统与虚拟现实技术,E-mail: danwen_chen@nudt.edu.cn;文 军(1976-),男,讲师,主要研究方向为多媒体信息系统;陈丹雯(1983-),女,博士生,主要研究方向为多媒体信息系统;袁志民(1982-),男,博士生,主要研究方向为多媒体信息系统。

分集中于故事单元的分割技术^[1-4],其中涉及关键帧特征分析、镜头分割等一系列相关内容。当前针对新闻事件开展的研究主要集中于文本类型的媒体中,例如话题识别与跟踪研究(Topic detection and tracking, TDT)^[5-8]。该研究针对文本类型的新闻报道展开,研究目的是实现按照话题查找组织和利用新闻信息。

考虑到用户对新闻视频中事件跟踪分析的需要,一部分研究试图通过自动语音识别(Automatic speech recognition, ASR)技术和机器翻译(machine translation, MT)技术获得新闻故事的演播稿,然后利用文本信息的话题识别与跟踪技术来实现新闻视频基于事件的组织。

近年来,随着图像处理技术的迅速发展,一部分研究试图以视觉信息为主,结合其它特征信息,利用新闻视频编辑中视觉素材的可重复性来实现对故事单元的相似程度分析,并在此基础上实现对故事单元的关联分析。与基于文本信息的分析技术相比,这种方法利用了更多类型的媒体特征,尤其是越来越重视视觉特征的应用,越来越接近实际应用中的各种可能情况。当前大部分研究将新闻视频故事单元的关联分析任务界定如下:

故事单元关联分析以故事单元分割为基础,识别新闻视频数据库中报道相同事件的故事单元。该过程可以看作对故事单元按照新闻事件的一种特殊形式的聚类。然而,由于新闻视频来源多样性所导致的视觉变化、语种差异等一系列问题,以及新闻事件关注重点程度变化所导致的密度差异,基于视觉特征和文本特征的传统聚类方法难以完成这一任务。

综上所述,新闻视频中的故事单元关联分析技术融合了文本 TDT 研究中的事件探测与跟踪技术和多媒体信息结构与内容分析技术的优势,密切贴近用户的实际需求,因此成为当前国内外新闻视频研究领域一个新的热点问题。

3 新闻视频故事单元关联与跟踪技术

对于新闻视频故事单元的关联分析而言,可以利用的信息主要有视频和文本信息,对于音频信息主要通过自动语音识别转换为文本信息进行利用。根据利用信息的角度可以将相关技术分为三类:基于文本相似度的关联分析技术、基于多模态信息相似度的关联分析技术以及基于关键帧视觉重复性的关联分析技术。

3.1 基于文本相似度的关联分析方法

新闻视频故事单元关联分析的目标与文本领域中 TDT 研究比较接近,因此最初开展的部分研究^[9-12]首先通过自动语音识别技术获得新闻视频相应的文本信息,通过文本信息中的故事单元分割和相似度分析,直接利用 TDT 的方法进行故事单元的话题探测与跟踪分析,如图 2 所示。

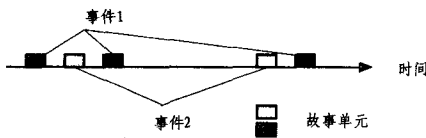


图 2 基于文本的 TDT 示意图

这种方法在故事分割和关联分析的各个阶段都只利用文本信息。对于文本信息的利用要求将文本的每个句子同步到新闻视频的音频轨迹,进行关联分析之前的文本预处理过程有以下步骤:

(1)对每个故事单元的文本数据提取名词组合;

(2)创建每个故事单元的关键词向量,关键词向量通常包括四个语义属性:一般名词、人名、地点/组织名词和时间,关键词向量通过对名词组合进行分析获取,并利用相关技术计算关键词的权重信息。在关联分析方法中,两个故事之间的关系按照二者关键词向量之间的余弦距离等方法来定义。如果这个故事单元之间相似度的值超过预定义阈值,则将故事单元看作具有关联关系。通过在所有故事单元之间将这种关联关系循环扩展,得到每个故事单元关联关系的树形结构。这种树形关联关系结构中,存在大量的重复信息,因此需要对这种复杂的关联关系进行过滤和合并,得到比较完备的关联关系线程结构。

基于文本的新闻视频故事单元关联分析方法最直接地利用了 TDT 研究的相关成果,明确了故事单元关联分析的基本任务和流程。然而,在实际研究和应用中,对于不同来源的多语种新闻视频而言,因为缺乏适当和有效的语音识别和机器翻译工具,通过音频轨迹得到的文本信息可能模糊不清,甚至很难得到有价值的文本信息,所以这类方法在新闻视频中的实际应用中存在很大局限性。

3.2 基于多模态信息相似度的关联分析方法

与文本数据相比,视频数据的最大特点在于多模态,即同时包括有视觉、听觉和文本等各种类型的信息。因此一部分研究将重点转移到不同模态信息的相似度融合分析,其中的多模态信息包括视觉特征、文本特征、语义概念特征等。通过各种模态信息的融合实现对新闻视频的故事单元分割,并衡量故事单元之间多模态信息的相似度,实现对新闻事件主题相关的故事单元进行关联分析,如图 3 所示。

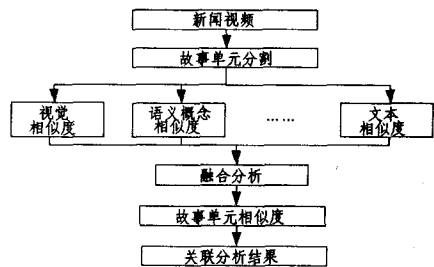


图 3 基于多模态信息相似度的关联分析流程

Xie 等^[13]使用层次 HMM 模型在低层视音频特征中发现时空模式,潜在语义分析发现文本簇,然后混合这些模态来发现潜在的故事主题。在文献^[14]中研究了人工标注视觉主题(例如:地点、人物、对象等)和主题标注技术在故事单元关联分析中的应用,使用图像分割技术进行故事单元聚类。在文献^[15]中讨论了同一天中两个不同的英语电视台之间新闻的联系问题,将关键帧分类为人物关键帧和不包含人物的关键帧,分别进行相似度计算,前者使用人物区域的相似度,后者使用的是关键帧的全局放射匹配视觉相似度,并融合了文本信息。在这些研究中,对于视觉重复性的研究非常有限,在语义概念识别中也没有使用自动标注的方法,而文本信息的利用也受到各种实际情况的限制。

凌坚^[16]在故事单元分割的基础上,利用关键帧典型场景分类、语音识别等技术,将视频场景信息和语音识别的文本信息相结合来实现对新闻视频的主题识别与跟踪。研究中对于各种典型场景分类有人脸特写、人群、野外和城市、室内/室外

几种类型来辅助语义要素分析,视觉特征选择主要利用了颜色和纹理特征;作为主题探测与跟踪的另一个重要基础是通过语音识别技术获得的文本信息。

基于多模态信息相似度分析的故事单元关联分析方法比基于文本信息的方法更好地体现了视频数据的特点,通过各种模态信息的融合分析,能够克服文本信息获取中的一些实际问题,具有更好的效率。但是也面临一些实际问题,对于视觉信息而言,底层特征中的颜色、纹理等特征难以克服视角、光照等变化的影响;而场景标注技术能够提供的语义信息比较有限,并且在当前技术条件下,大部分语义概念标注的效率与实际需求相比还有明显的差距;文本信息的获取也面临种种困难。

正是因为各种模态信息获取过程中的一些困难,以及当前技术条件下各模态信息的一些局限性,使得基于多模态信息相似度的故事单元关联分析方法并未能够完全体现多模态信息融合分析技术的优越性。

3.3 基于关键帧视觉重复性的关联分析方法

因为新闻报道中事件的延续性及对事件重点内容的强调,在新闻视频的编辑过程中,报道相同事件的故事单元关键帧中通常包含有重复或者近似的场景和对象,如图4所示。这种关键帧视觉上的重复性或者近似性,成为衡量故事单元是否具有关联关系的一个重要线索。文献[17,18]将用户选定的感兴趣故事单元中的镜头或场景作为输入条件,通过视觉特征的重复性来发现和跟踪相关的故事单元。早期的其他一些研究^[20,21]主要通过底层的全局特征来衡量图像相似度(例如:颜色直方图)。全局信号对包含相同内容的片段匹配比较有效。Odobez等^[22]使用颜色直方图来衡量视觉相似性,然后使用时间距离来扩展这种相关性。然而,因为新闻视频摄制和编辑过程中,关键帧可能出现各种复杂的变化,基于全局特征(如:HSV颜色直方图)的分析方法对于分析不同来源、不同时间的镜头相似性易于受到光照、编辑方式等各种因素的干扰,建立在全局特征基础上的镜头(关键帧)相似性衡量方法不够鲁棒^[23]。

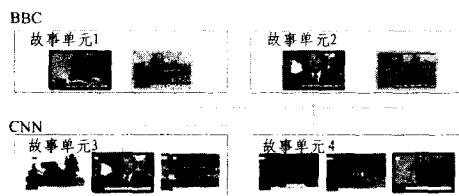


图4 故事单元之间的视觉重复与语义关联关系

Pinar Duygulu等^[19]在研究中认为,新闻视频制作中有两条规则:(1)重复使用特定镜头来提示特定的事件;(2)如果没有相同的,则选取相似的镜头或图标作为事件的标志。在此基础上设计了一种方法,以发现重复的视频片段并识别匹配相似的图标来探测和跟踪新闻事件。该方法首先通过对一幅关键帧排序和检查最近邻的 N 幅关键帧并研究关键帧视觉相似度的微分,在此基础上融合文本信息进行故事相似度分析。该方法是一种启发式方法,对一些经验参数的设置非常敏感。

随着图像处理技术研究中局部特征提取与分析技术的进步,最近的一些研究证明了使用基于局部关键点和局部特征来克服几何与光学变化对图像匹配分析的影响是有效

的^[24,25]。与全局特征不同的是,基于局部关键点的特征探测和描述能够避免全局特征的缺点,得到比较鲁棒的匹配结果。

文献[23]认为相似关键帧图像的探测可以辅助视频线程化分析,提出了基于局部特征统计随机属性关系图(attributed relational graph, ARG)的相似图像探测方法,图像之间的相似度通过计算 ARG 相似度来获得。ARG 是一个以探测到的局部关键点作为顶点的完全连通图,包括了局部特征和空间关系,使得该方法在相似图像的识别中具有较高的准确率。该研究的一项重要意义在于正式提出了“相似图像”(Image Near-Duplicate, IND)的概念。相似图像被定义为:一幅图像与另一幅图像在表现内容和场景上相同或者相似,只是在图像获得的时间、条件、编辑操作上有所差异。Hsu等^[26]在 ARG 相似度的基础上,通过视觉语义概念和文本信息的相似度辅助跟踪故事单元的主题,进一步证明了 IND 探测结果能够显著提高新闻视频中主题跟踪的性能。然而,基于 ARG 的方法受启发式学习参数的限制,并且匹配速度很慢,不适合于大容量数据库的应用。

随后一系列研究将“相似图像”的概念扩展为与研究背景更加贴近的“相似关键帧”(Near-duplicate keyframes, NDK),如图5所示。图5(a)为不同电视台在同一场景的拍摄,在拍摄时间上的不同导致视觉上有所差异;图5(b)为素材重复利用时采用了不同编辑效果导致视觉上出现差异。最近的研究基于局部关键点匹配的 NDK 探测方法来进行故事单元关联分析,其中目前最为有效的局部特征是 David G. Lowe^[27]提出的尺度不变特征转换(Scale Invariant Feature Transform, SIFT)理论与技术。该方法对于克服光照、旋转、尺度变化等问题具有良好的性能。

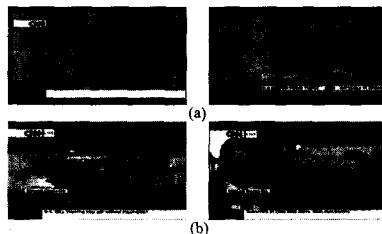


图5 相似关键帧的示例

香港城市大学的 Chong-Wah Ngo 等在新闻视频的 NDK 探测^[28-33]和故事单元关联^[34-39]等方面开展了一系列的研究,这些研究利用了视频关键帧的 SIFT 和 PCA-SIFT 等局部特征,并研究了提高匹配分析准确率的匹配点过滤方法和提高匹配分析速度的索引机制;提出了基于局部关键点匹配的 NDK 识别方法,研究了 NDK 与故事单元关联关系之间的联系。此外研究的一个重要贡献是提出了按照时间分割来建立关联分析子数据库,提高了匹配分析速度,并提出了相似关键帧传递性来实现子数据库之间 NDK 的识别。新加坡国立大学的 Shi-Yong Neo 等人^[41,42]在此基础上,进一步提出了提高速度效率的方法。方法对关联分析子数据库中的关键帧利用视觉特征聚类方法再将关键帧分为一系列更小的类型,只在每个类型进行 NDK 识别。这一方法进一步减少了每次进行匹配分析的关键帧数量,有效地提高了分析的速度。Wu xiao^[40]在最近的研究中将“视觉特征袋”(bag of visual features)与视频的文本信息相结合来进行视频检索和重复性探测,然而这种方法虽然有效提高了分析的速度,方法的查全率

比较高,可以在视频检索等研究领域提高分析处理的速度,但是准确率有限,难以满足新闻视频故事单元关联和跟踪的需要。

基于关键帧视觉重复性,尤其是局部特征匹配分析技术的故事单元关联分析方法由于其特定的优势,成为近年来故事单元关联分析技术的发展趋势。然而,该系列技术的一个突出问题在于匹配分析的时间效率,如何提高分析的时间效率成为最近的研究热点。当前的加速方法主要侧重于减少每次匹配分析的关键帧数量和利用索引方法来提高数据访问的速度。

3.4 不同关联分析技术的比较

上述3种不同的新闻视频故事单元关联分析方法具有各自的特点和优势。对于方法的比较如表1所列。

表1 关联分析方法对比表

	基于文本相似度的关联分析	基于多模态信息相似度的关联分析	基于关键帧视觉重复性的关联分析
信息模态	文本特征	视觉特征、语义概念、文本特征	视觉特征
信息获取难度	难	难	易
处理速度	快	较慢	慢
关联分析实现难度	一般	一般	一般
适用性能	差	较差	较好
量化程度	好	好	一般

从表中可以看出,基于文本相似度的关联分析方法虽然处理速度比较快,但是因为文本特征获取的难度,导致适应性比较差,难以处理实际应用中出现的多语种等实际问题;基于多模态信息相似度的关联分析方法虽然可能在各种模态的信息之间实现互补,适用性能有所提高,但是因为各种模态信息的获取都存在一些困难,对各种模态信息的计算也在一定程度上增加了复杂程度,也可能导致错误信息的传播等问题;基于关键帧视觉重复性的关联分析方法具有较好的适用性能,能够比较有效地处理各种不同来源的新闻视频,并且单纯的视觉特征信息获取比较容易,但是重复度判断的计算量比较大,使得处理速度比较慢。同时,目前研究的视觉重复性判断方法主要基于图像匹配的算法,因此在很多情况下难以提供有效的相似度量分析。

综合来看,基于关键帧视觉重复性的关联分析方法因为其较好的适用性能,成为目前研究的一种趋势。其主要不足在于分析速度和相似度的量化表示方面,针对这些不足进行改进将成为进一步研究的方向。

结束语 从目前新闻视频中故事单元关联分析的国内外研究情况来看,所采用的研究方法从初期基于纯文本信息的方法演变为融合视觉特征、文本语义特征等各种媒体信息的方法,尤其是结合实际应用出现的问题,为克服难以准确获取新闻视频语义信息的问题,最近的研究更加注重视觉信息的应用,研究的重点侧重于关键帧基于局部关键点和局部特征的NDK识别技术来进行故事单元关联分析。当前研究的技术已经可以有效克服视觉相似分析中出现的几何和光照变化等问题,并且针对视觉重复性与故事单元关联关系之间的联系开展的部分研究,取得了一定的进展。然而新闻视频故事单元关联分析方法并不完善,进一步研究的可能改进包括:

(1)基于局部关键点和局部特征的NDK识别方法中图像匹配分析的速度受到限制,难以满足大规模视频数据库的

需求,影响速度的因素来自两个方面:数据库中关键帧数量庞大和每幅关键帧中的局部关键点数量很多。当前研究虽然采取了一系列措施来提高分析速度,但是只是侧重于减少每次分散的关键帧数量,仍需要研究有效控制和减少局部关键点的方法来进一步提高分析的速度,从而在本质上提高匹配分析的速度。

(2)单独基于局部特征或者全局特征来衡量新闻故事的视觉相似度都具有各自的局限性,目前大部分研究中仍然缺乏对局部特征与全局特征的综合考虑,虽然ARG图等方法中涉及了空间信息,但是分析的速度较慢,因此可以在进一步研究中考虑局部特征与全局特征的融合分析方法,在提高效率的同时,保证有较快的分析速度。

(3)当前研究中提出了各种获取语义信息的途径和方法,虽然语义信息可以辅助故事单元的关联分析,然而语义信息的获取仍然面临种种困难。目前研究中的语义信息提取技术主要包括基于ASR的文本获取以及语义概念自动标注技术,前者在处理不同语种的视频时面临许多实际问题,后者作为近年来研究的热点。在目前技术水平下,大部分语义概念的标注效率有限,因此必须对需要标注的语义概念进行研究以提高标注的效率。

(4)当前研究的基于NDK的故事单元关联分析方法大部分只是简单研究了NDK与故事单元的关联关系,缺乏基于NDK的定量分析。相似关键帧的出现对于故事单元之间的视觉和语义相似度具有很强的信息提示,作为关联分析的一种重要线索,需要提供一种定量的分析和支持方法。

(5)当前研究结合实际应用情况,强调视觉信息的应用,目前比较有效的方法是基于关键帧匹配分析的方法,关键帧是一种相对比较宏观的视觉信息,而视觉语义概念、底层视觉特征等是相对比较微观层次的视觉信息,可以利用这些微观层次的视觉信息进行辅助分析,以提供更多的量化分析信息。

参考文献

- [1] Chua Tat-Seng, Chang Shih-Fu, Chaisorn Lekha, et al. Story boundary detection in large broadcast news video archives-techniques, experience and trends[C]//Proceedings of the 12th Annual ACM International Conference on Multimedia (ACM MM2004). New York, USA, October 2004; 656-659
- [2] 刘华咏. 基于音视频特征和文字信息自动分段新闻故事[J]. 系统仿真学报, 2004, 16(11): 2608-2610
- [3] 张春林, 张鹏林, 胡瑞敏. 新闻视频中基于主持人识别的新闻故事探测[J]. 计算机工程, 2003, 29(14): 20-26
- [4] Hsu W, Kennedy L, Huang C-W, et al. News video story segmentation using fusion of multi-level multi-modal features in trecvid 2003[C]// 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2004). Montreal, Canada, May 2004; 645-648
- [5] Allan J. Topic detection and tracking; event-based information retrieval[M]. Norvell, MA, USA: Kluwer Academic Publishers, 2002
- [6] 贾自艳, 何清, 张海俊, 等. 一种基于动态进化模型的事件探测和追踪算法[J]. 计算机研究与发展, 2004, 41(17): 1273-1280
- [7] 李保利, 俞士汶. 话题识别与跟踪研究[J]. 计算机工程与应用, 2003(17): 7-10
- [8] 于满泉, 骆卫华, 许洪波, 等. 话题识别与跟踪中的层次化话题识

- 别技术研究[J]. 计算机研究与发展, 2006, 43(3):485-495
- [9] Ide C, Mo H, Katayama N, et al. Threading news video topics [C]//Proceedings of the 5th ACM SIGMM International Workshop on Multimedia Information Retrieval (MIR2003). California, USA, November 2003; 239-246
- [10] Ide I, Mo H, Katayama N, et al. Topic threading for structuring a large-scale news video archive[C]//Image and Video Retrieval; Third International Conference (CIVR2004). Dublin, Ireland, July 2004; 123-131
- [11] Katayama N, Mo H, Ide I, et al. Mining large-scale broadcast video archives towards inter-video structuring[C]//Pacific Rim Conf. on Multimedia (PCM2004). Japan; Tokyo, 2004; 489-496
- [12] Ide I, Mo H, Katayama N, et al. Exploiting topic thread structures in a news video archive for the semi-automatic generation of video summaries[C]//2006 IEEE International Conference on Multimedia and Expo (ICME2006). Toronto, Canada, July 2006; 1473-1476
- [13] Xie L, et al. Discover meaningful multimedia patterns with audio-visual concepts and associated text [C]// IEEE International Conference on Image Processing (ICIP 2004). Singapore, October 2004; 2383-2386
- [14] Kender J R, Naphade M R. Visual concepts for news story tracking; analyzing and exploiting the NIST Trecvid video annotation experiment[C]//IEEE Computer Society Conference on Computer Vision and Pattern Recognition(CVPR 2005). San Diego, USA, June 2005
- [15] Zhai Y, Shah M. Tracking news stories across different sources [C]//Proceedings of the 13th Annual ACM International Conference on Multimedia (ACM MM2005). Singapore, November 2005; 2-10
- [16] 凌坚. 新闻视频主题识别与跟踪的研究[D]. 杭州: 浙江大学, 2007
- [17] Miadowicz J Z, Gauch J M, Shivadas A. Image Based Tracking of News Stories[C]//Seventh IEEE International Symposium on Multimedia (ISM 2005). Irvine, USA, December 2005; 545-550
- [18] Miadowicz J Z. Story Tracking in Video News Broadcasts[D]. University of Kansas, 2004
- [19] Duygulu P, Pan Jia-yu, Forsyth D A. Towards auto-documentary; tracking the evolution of news stories[C]//Proceedings of the 12th Annual ACM International Conference on Multimedia (ACM MM2004). New York, USA, October 2004; 820-827
- [20] Cheung S C, Zakhor A. Efficient video similarity measurement with video signature[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(1):59-74
- [21] Jain A K, Vailaya A, Xiong W. Query by video clip [J]. ACM Multimedia Systems, 1999, 7(5): 369-384
- [22] Odobez J-M, Gatica-Perez D, Guillemot M. Video shot clustering using spectral methods[C]//Third International Workshop on Content-Based Multimedia Indexing (CBMI 2003). Rennes, France, September 2003; 94-102
- [23] Zhang D-Q, Chang S-F. Detecting image near-duplicate by stochastic attributed relational graph matching with learning[C]//Proceedings of the 12th Annual ACM International Conference on Multimedia (ACM MM2004). New York, USA, October 2004; 877-884
- [24] Grauman K, Darrell T. Efficient image matching with distributions of local invariant features[C]//IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2005 (CVPR 2005). San Diego, USA, June 2005; 627-634
- [25] Dance C, Willamowski J, Fan Lixin, et al. Visual categorization with bags of keypoints[C]//ECCV International Workshop on Statistical Learning in Computer Vision, Prague, May 2004; 59-74
- [26] Hsu W, Chang S-F. Topic tracking across broadcast news videos with visual duplicates and semantic concepts[C]//IEEE International Conference on Image Processing. Atlanta, USA, October 2006; 141-144
- [27] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2):90-110
- [28] Wu Xiao, Zhao Wan-lei, Ngo C-W. Near-duplicate keyframe retrieval with visual keywords and semantic context[C]//Proceedings of the 6th ACM international conference on Image and video retrieval(CIVR 2007). Amsterdam, The Netherlands, July 2007; 162-169
- [29] Wu Xiao, Zhao Wan-lei, Ngo C-W. Efficient near-duplicate keyframe retrieval with visual language models[C]//Proceedings of IEEE International Conference on Multimedia & Expo (ICME 2007). Beijing, China, July 2007; 500-503
- [30] Wu Xiao, Hauptmann A G, Ngo C-W. Practical elimination of near-duplicates from web video search[C]//Proceedings of the 15th International Conference on Multimedia(ACM MM2007). Augsburg, Germany, September 2007; 218-227
- [31] Ng C-W, et al. Fast tracking of near-duplicate keyframes in broadcast domain with transitivity propagation [C]//Proceedings of the 14th Annual ACM International Conference on Multimedia (ACM MM2006). Santa Barbara, USA, October 2006; 845-854
- [32] Jiang Yu-gang, Wei Xiaoyong, Ngo C-W, et al. Modeling local interest points for semantic detection and video search at TRECVID 2006 [C]//Proceedings of TRECVID Workshop 2006. Gaithersburg, USA, November 2006
- [33] Ngo C-W, Pan Zailiang, Wei Xiaoyong, et al. Motion driven approaches to shot boundary detection; Low-Level Feature Extraction and BBC Rush Characterization at TRECVID 2005[C]//Proceedings of TRECVID Workshop 2005. Gaithersburg, USA, November 2005
- [34] Wu Xiao. Threading stories and generating topic structures in news videos across different sources[C]//Proceedings of the 13th Annual ACM International Conference on Multimedia (ACM MM2005). Singapore, November 2005; 1047-1048
- [35] Wu Xiao, Ngo C-W, Li Qing. Co-clustering of time-evolving news story with transcript and keyframe[C]//Proceedings of IEEE International Conference on Multimedia & Expo (ICME 2005). Amsterdam, The Netherlands, July 2005; 117-120
- [36] Wu Xiao, Lee V C S, Ng J K-Y. A preemptive scheduling algorithm for wireless real-time on-demand data broadcast[C]//Proceedings of the 11th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA 2005). Hongkong, China, August 2005; 17-22
- [37] Wu Xiao, Ngo C-W, Hauptmann A G. Multi-modal news story clustering with pairwise visual near-duplicate constraint [J].

- [38] Wu Xiao, Hauptmann A G, Ngo C-W. Novelty and redundancy detection with multimodalities in cross-lingual broadcast domain [J]. Computer Vision and Image Understanding, 2008, 110(3): 418-431
- [39] Lee V C S, Wu Xiao, Ng J K-Y. Scheduling real-time requests in on-demand data broadcast environments[J]. Real-Time Systems Journal, 2006, 34(2):83-99
- [40] Wu Xiao, Hauptmann A G, Ngo C-W. Novelty detection for cross-lingual news stories with visual duplicates and speech transcripts[C]// Proceedings of the 15th International Conference

on Multimedia (ACM MM2007). Augsburg, Germany, September 2007:168-177

- [41] Zheng Y-T, Neo S-Y, et al. Fast near-duplicate keyframe detection in large-scale video corpus for video search[C]// International Workshop on Advanced Image Technology 2007 (IWAIT 2007). Bangkok, Thailand, January 2007
- [42] Zheng Y-T, Neo S-Y, et al. The Use of Temporal, Semantic and visual partitioning model for efficient near duplicate detection in large scale news corpus[C]// International Conference on Image and Video Retrieval (CIVR2007). Amsterdam, The Netherlands, July 2007:409-416

(上接第4页)

管理系统及数据库等。通过这些设施,能够支撑物联网海量信息的处理、存储、管理等工作。

物联网的应用需要智能化信息处理技术的支撑,主要针对大量的数据通过深层次的数据挖掘,并结合特定行业的知识和前期科学成果,建立针对各种应用的专家系统、预测模型、内容和人机交互服务。专家系统利用业已成熟的某领域专家知识库,从终端获得数据,比对专家知识,从而解决某类特定的专业问题。预测模型和内容服务等基于物联网提供的对物理世界精确、全面的信息,可以对物理世界的规律(如洪水、地震、蓝藻)进行更加深入的认识和掌握,以做出准确的预测预警,以及应急联动管理。人机交互与服务也体现了物联网“为人类服务”的宗旨。人机交互提供了人与物理世界的互动接口。物联网能够为人类提供的各种便利也体现在服务之中。

3 物联网典型应用

下面以环境监测为例说明物联网的应用系统。环境监测是物联网的一个重要应用领域,物联网自动、智能的特点非常适合环境信息的监测。一般来讲,环境监测物联网系统结构包括以下几个部分:

感知层:该层的主要功能是通过传感器节点等感知设备,获取环境监测的信息,如温度、湿度、光照度等。由于环境监测需要感知的地理范围比较大,所包含的信息量也比较大,该层中的设备需要通过无线传感器网络技术组成一个自治网络,采用协同工作的方式,提取出有用的信息,并通过接入设备与互联网中的其它设备实现资源共享与交流互通。

接入层:该层的主要功能是通过现有的公用通信网(如有线 Internet 网络、WLAN 网络、GSM 网、TD-SCDMA 网)、卫星网等基础设施,将来自感知层的信息传送到互联网中。

互联网层:该层的主要功能是以 IPv6/IPv4 为核心建立的互联网平台,将网络内的信息资源整合成一个可以互联互通的大型智能网络,为上层服务管理和大规模环境监测应用建立起一个高效、可靠、可信的基础设施平台。

服务管理:该层的主要功能是通过大型的中心计算平台(如高性能并行计算平台等),对网络内的环境监测获取的海量信息进行实时的管理和控制,并为上层应用提供一个良好的用户接口。

应用:该层的主要功能是集成系统底层的功能,构建起面

向环境监测的行业实际应用,如生态环境与自然灾害实时监测、趋势预测、预警及应急联动等。

通过上面几个部分,环境监测物联网就可以实现协同感知环境信息,进行态势分析,并预测发展趋势。

结束语 物联网是由各种技术融合而成的新型技术体系。物联网在多种应用中具有潜在、显著的技术价值和应用需求,物联网的发展必将推动物物相联、人物互动的信息化社会建设。影响物联网大规模应用的关键问题是技术繁多,而且需要互联互通,因此物联网具有巨大的深入发展和提升的空间。

参考文献

- [1] <http://www.autoidlabs.org/page.html>
- [2] Kahn J M, Katz R H, Pister K S J. Next century challenges: mobile networking for “Smart Dust”[C]// Proceedings of the 5th annual ACM/IEEE international conference on Mobile computing and networking (MobiCom1999). August 1999
- [3] Neil G. When Things Start to Think[M]. New York: Henry Holt, 1999
- [4] ITU Strategy and Policy Unit (SPU). ITU Internet Reports 2005: The Internet of Things[R]. Geneva: International Telecommunication Union (ITU), 2005
- [5] Buckley J. From RFID to the Internet of Things—Pervasive networked systems[R]. Brussels: European Commission, DG Information Society and Media, Networks and Communication Technologies Directorate, 2006
- [6] http://ec.europa.eu/information_society/policy/rfid/documents/in_cerp.pdf
- [7] <http://www.ibm.com/smarterplanet/us/en/>
- [8] <http://space.tv.cctv.com/video/VIDE1268482063865885>
- [9] 崔莉, 鞠海玲, 苗勇, 等. 传感器网络研究进展[J]. 计算机研究与发展, 2005, 42(1):163-174
- [10] http://www.itu.int/net/TELECOM/World/2009/newsroom/speeches/wang20091005_oc.aspx
- [11] ITU NGN-GSI Rapporteur Group. Requirements for support of USN applications and services in NGN environment[R]. Geneva: International Telecommunication Union (ITU), 2007
- [12] http://news.xinhuanet.com/mrdx/2009-09/11/content_12035797.htm
- [13] <http://www.rfidchina.org/readinfos-37143-177.html>