

数字视频信息的元数据研究¹⁾

严 明 苏新宁

(南京大学信息管理系, 南京 210093)

摘要 利用元数据对各种资源(特别是数字化资源、网络资源)进行全面的描述,并据此对资源进行管理和检索,一直是最直接、最有效的手段。针对不同的资源类型、不同的用途,不同的机构建立了多种不同的元数据标准。本文首先探讨了数字视频信息的描述需求,然后分析了目前常见的数字视频信息的元数据和元数据标准,最后展望了数字视频信息元数据的发展趋势。

关键词 数字视频 元数据 元数据标准

Study on Metadata of Digital Video

Yan Ming and Su Xinning

(Department of Information Management, Nanjing University, Nanjing 210093)

Abstract It is one of the most direct and efficient methods to manage and retrieve all kinds of resources, especially digital resources and network resources, through metadata which is used to describe these resources. There are many different metadata standards designed by different organizations for different kinds of resources or different applications. This paper firstly discusses the requirement for describing digital video, then analyzes the digital video metadata and its standards which are used recently. Finally, the paper tries to give a prospect for the development of digital video metadata.

Keywords digital video, metadata, metadata standard.

1 引 言

随着各种影像设备的发展与普及,以及电影电视行业、多媒体技术和 Internet 的发展,人们面对的各种视频信息迅速增长。据国家统计局公布的 2001 年年度统计数据^[1],我国仅 2001 年一年制作的电视节目就达 989173 小时,电影达 154 部。面对如此巨量的视频信息,如何能快速地获取和有效地利用,这就需要对其进行科学、合理的组织与存储。元数据的诞生,无疑为视频信息的描述和组织提供了

有效的手段。

元数据(metadata)是关于数据的组织、数据域及其关系的信息^[2]。简言之,元数据是关于数据的(结构化)数据。元数据实质上是传统的文献资源的著录信息在电子资源的一种扩展^[7]。利用元数据对各种电子资源(特别是数字化资源、网络资源)进行描述,并据此对资源进行交流和检索,可以说是视频信息管理的最佳选择。

针对不同的资源对象、不同的用途,不同的机构建立了多种不同的元数据标准。例如,适用于书目数据的 MARC 元数据标准,适用于艺术品数据的

收稿日期:2003 年 10 月 16 日

作者简介:严明,南京大学信息管理系博士研究生,南京政治学院副教授,研究方向:数据库技术、多媒体信息的组织与检索。苏新宁,博士生导师,教授,研究方向:网络信息资源的研究与开发、情报检索算法、信息处理技术。

1) 教育部人文社科重点研究基地重大项目(项目编号:02JAZJD870004)、国家 863 课题(课题编号:2001AA113161)成果。

CDWA,适用于地理空间数据的 FGDC,适用于政府公用信息的 GILS,适用于档案和手稿的 EAD,等等。近年来,视频信息的元数据及其标准的研究也在不断地深入,出现了一些专用或可用于视频信息描述的元数据标准,特别是 1996 年 10 月 MPEG 组织提出的 MPEG-7 国际标准。MPEG-7 标准的正式名称是“多媒体内容描述接口”(Multimedia Content Description Interface),其目标是解决对多媒体信息(特别是视听信息)描述的标准问题,并将该描述与所描述的内容相联系,以实现快速有效的搜索与过滤。

本文将从数字视频信息的描述需求入手,分析目前常见的数字视频信息的元数据和元数据标准,并展望数字视频信息元数据的发展趋势。

2 数字视频信息的描述需求

视频信息是一种无结构的流式媒体,犹如一本无目录、无索引、无前言的书籍,且视频信息在存档格式、压缩编码方式、内容组织形式等方面呈现多种不同的类型。与文本信息相比,视频信息含有更丰富的信息,包括视觉信息、听觉信息和文本信息,且不同类型的视频信息在功能、用途、消费对象和方式等方面都有所不同,因而对其描述要求也不尽相同。此外,不同的用户、不同的应用对视频信息有不同的管理和检索要求,这也要求能提供多角度、多层次的视频信息的描述。

例如,对于有关足球比赛的视频信息来说,一般观众与足球教练的需求往往是不同的。一般观众通常只需要能基于视频信息外部特征的描述信息检索到某一场比赛,而对于足球教练来说,通常需要能基于视频内容的描述检索感兴趣的一组视频片段。

根据 MPEG-7 的目标要求^[3],对视频信息的描述至少应包括如下的信息:

- 有关内容的产生和发展进程的描述信息(如导演/作者、标题等);
- 与内容使用有关的信息(如版权、使用历程、宣传计划等);
- 有关内容存储特性的信息(如存储格式、编码等);
- 有关内容的低级特性的信息(如颜色、纹理、音质、音调描述等);
- 从内容捕捉到的实体的概念化信息(如对象和事件,对象间的交互作用等);
- 关于如何用有效方法浏览内容的信息(如概

要、变更、空间和频率等);

- 关于用户和内容交互作用的信息(如用户选择、使用历史等)。

这些描述可以分为一般性描述和基于内容的描述。一般性描述是指从视频文档外部导出的、不依赖于其内容的信息,用于视频文档的标识和检索,如标题、作者、时间、文档大小、存储格式与编码格式、使用信息(软硬件要求、使用要求、版权等)等等。基于内容的描述又可以分为结构性描述(用于描述针对视频内容的逻辑结构,包括节目、场景、镜头、帧之间的层次结构以及所包括的对象之间的结构与关系等)、低级特征描述(用于描述视频信息的物理特征信息,如颜色、纹理、运动、音质、音调描述等)和语义描述(用于描述视频中包括的对象和事件)。

3 数字视频信息的元数据标准

随着数字视频技术的发展,针对数字视频信息的元数据研究也在不断地深入,并已出现了一些基于不同目标的元数据模式或标准。

3.1 数字电视元数据标准

在数字电视的发展过程中,出现了一些数字电视元数据标准,如 DVB-SI、TV-Anytime 和 SMPTE 等。它们的特点是这些元数据标准均为行业标准,且仅适用于某个地区或应用领域。

DVB-SI 是指“数字视频广播业务信息系统标准”(Digital Video Broadcasting -Service Information)。该标准用于为用户提供广播业务指南。DVB-SI 由一系列的表和描述符构成,它们给出了有关 DVB 业务或节目的所有参数^[4]。SI 中有传送的节目信息,包括节目的种类,节目的时间,节目的来源等。这些信息主要包括在 4 个基本表和一系列可选送的表中。4 个基本表是网络信息表(NIT),业务描述表(Service Description Table: SDT),节目信息表(Event Information Table: EIT),时间及日期表(Time and Date Table: TDT)。除这 4 个表外,DVB-SI 信息表还有节目组相关表(Bouquet Association Table: BAT)作为给不同观众不同组合节目的一种方法,运行状态表(Running Status Table: RST)用于快速更新某节目或某些节目的运行参数等。该标准目前主要在欧洲使用。

从 1999 年 9 月至今,由 160 多个成员组成的即时电视论坛(TV-Anytime Forum)致力于研究和开发

即时电视规范 (TV-Anytime), 提供适合个人视频录制 (PVR) 的电子节目指南 (EPG) 信息标准和实现机制^[5]。TV-Anytime 有多个部分组成, 其中元数据定义了节目、节目组和电视节目安排表的描述。元数据主要有实例描述元数据和内容描述元数据。前者包含已发布的节目的时间安排信息 (但缺乏内容的描述), 实际的节目安排可能会有所变化; 后者给出了有关节目的描述信息。元数据使用 XML 形式, 在 DVB 上是压缩传输的。

SMPTE 是指电影与电视工程师协会 (Society of Moving Picture and Television Engineers)。该协会制定了一系列面向厂商的有关视频信息处理的标准。其中, SMPTE 元数据字典 (SMPTE Metadata Dictionary) 提供了一组视听描述符, 覆盖整个生产链, 包括前期制作、后期制作、成品、分发、广播、存储等各个环节^[6]。绝大多数元数据是特定媒体的特征, 如关键帧、关键语音、定时等等。由于它着重是面向整个生产流程、面向厂商, 所以不太适合于视频检索和最终用户。

3.2 数字视频软硬件开发商制定的元数据标准

目前, 各种针对数字视频信息的软硬件产品均以不同的形式支持元数据。这些元数据部分是在视频创建中自动地生成的, 有些是在编辑过程中人工给出的。

SONY 在 2003 CCBN 展览会上提出了“AV/IT Reality”的主题, 介绍了 AV 与 IT 不断融合的新技术和新产品, 展示了数字与网络的现实应用与解决方案。在展出的设备、系统当中, 有三项关键技术: MXF (媒体交换文件格式)、Metadata (元数据) 和 MIB (管理信息基础)。Sony 公司的许多数码视频产品均支持元数据管理。例如 Sony PDW 系列专业光盘摄录一体机支持元数据记录: UMID (唯一素材识别码)、扩展的 UMID、基本标识 (包括镜头标记) 和 GPS 信息等。

Apple 公司的 QuickTime 电影文件现已成为数字媒体领域的工业标准。QuickTime 电影文件格式定义了存储数字媒体内容的标准方法, 使用这种文件格式不仅可以存储单个的媒体内容 (如视频帧或音频采样), 而且能保存对该媒体作品的完整描述。QuickTime 文件格式中媒体描述 (元数据) 和媒体数据是分开存储的, 元数据 (称为“Movie”部分) 包含轨道 (track) 数目、视频压缩格式和时间信息以及媒体数据存储区域的索引等信息。轨道可以有許多不同

的类型, 其中 Movie track——版权、注释及其他概要信息, Video track——数字化视频、着色的 3D 动画或其他编辑图像的序列和可选的特殊效果, Text track——输入到 QuickTime 中的标题、片头字幕等字符信息, Hint track——允许流服务器通过实时流方式传输媒体轨道的信息。

Microsoft Windows Media 技术是一个能适应多种网络带宽条件的流式多媒体信息的发布平台, 提供密切结合的一系列服务和工具, 用以创造、管理、广播和接收通过 Internet 和企业 Intranet 传送的极其丰富的流式化多媒体演示内容, 包括了流式媒体的制作、发布、播放和管理的一整套解决方案。Windows Media 提供了全面的元数据属性支持, 其中核心属性包括标题 (Title)、作者 (Author)、描述 (Description)、版权 (Copyright) 和等级 (Rating)。对于视频文件来说, 常用的元数据属性分为三类: 8 个主要属性、13 个次要属性和 13 个第三级属性。有关 Windows Media 元数据的详细内容可参见文献[8]。

3.3 都柏林核心元数据集

都柏林核心元数据集 (Dublin Core Metadata Element Set, 缩写为 DCMES、简称为 DC) 是一个起源于图书馆界 (由美国 OCLC 公司发起)、基于内容描述的元数据模型。它以其简练性、灵活性、易于人机理解性、可扩展性以及能与其他元数据标准间的互操作性, 成为国际范围内通用的网络资源描述元数据标准, 主要用于博物馆、图书馆、政府机构、商业组织的电子资源的资源发现^[9]。

虽然 Dublin Core 主要面向网络资源的著录描述且仅包含 15 个基本元素, 但通过对 Dublin Core 进行扩展, 制定针对视频信息的修饰词, 也可以用于描述视频信息的结构信息和格式信息。Jane Hunter 等在文献[10]中介绍了针对视频信息的 Dublin Core 扩展建议, 给出了一个利用 Dublin Core 的多层次的视频元数据描述方案, 并给出一些示例。

ViDe (Video Development Initiative) 也在 Dublin Core 的扩展方面做了大量的工作。ViDe 是一个致力于数字视频研究与教育的国际协作项目。ViDe 的视频存取工作组 (Video Access Working Group) 在工作组的报告[11]和[12]中, 系统地介绍了针对 Dublin Core 而制定的 ViDe 修饰词及其扩展, 并给出了视频描述矩阵 (Video Description Matrix)。有关详细信息可参见网站 <http://www.vide.net>。

3.4 多媒体内容描述接口——MPEG-7

活动图像专家组(MPEG)是国际标准化组织(ISO)和国际电工委员会(IEC)联合技术委员会下属的一个工作组,负责数字视频、音频和其他媒体的压缩、解压缩、处理和表示等国际技术标准的制定工作。MPEG-7的工作于1996年启动,目标是制定一套标准的描述符和描述方案,用来描述各种类型的多媒体信息及它们之间的关系,以便更快更有效地检索信息。这些多媒体材料可包括静态图像、图形、3D模型、声音、语音、电视以及在多媒体演示中它们之间的组合关系。

MPEG-7 标准的核心内容是描述定义语言(DDL, Description Definition Language)、描述符(Ds, Descriptor)和描述规范(DSs, Description Scheme)。其中,描述符主要用于定义多媒体内容的特征信息的语法和语义属性;描述方案用于说明描述符与描述符、描述符与描述方案、描述方案与描述方案之间的结构和关系;描述定义语言用于定义描述符和描述方案,它以XML语言为基础并进行了扩展。它们之间的关系如图1所示。

MPEG-7 的描述符与描述方案主要用于多媒体内容的低层次视听特征的描述(如颜色、纹理、运动、声量以及位置、时间、品质等)和高层次视听特征的描述(如区域、分割、对象、事件以及与创作、作品、使用等相关的元数据)^[3]。它只规定了媒体特征的描述语法和语义,不包含怎样从原始数据中提取特征

和怎样通过搜索引擎或过滤器来利用这些特征描述。

现有的 MPEG-7 标准包括 8 个部分:系统、描述定义语言、视觉、音频、多媒体描述方案、参考软件、一致性测试以及抽取与使用描述。同时,MPEG-7 提供了一系列标准化工具,可以从不同的层次、不同的侧面描述视频信息。但 MPEG-7 标准还处于发展之中,目前有许多不同领域、学科和组织的学者从事这方面的研究^{[13][14][15][16]},并在特定的应用领域建立了原型系统。

4 数字视频元数据的发展趋势

随着数字广播电视、数字电影、流媒体的发展以及 Internet 的高速发展,数字视频的应用越来越普遍,应用要求越来越多样化。作为数字视频信息管理、共享、重用和检索等应用基础的视频元数据的研究,近年来成为了热点,也是难点。目前,无论是图书情报界、相关的标准化组织,还是有关的行业,都积极地进行数字视频元数据及其标准的研究,并形成了许多国际性的合作项目。

4.1 数字视频元数据向多模态方向发展

数字视频信息进行描述的元数据从功能上看,可以分为描述元数据、管理元数据、结构元数据和基于语义的元数据等等。目前研究的热点和难点是基于内容的元数据描述。

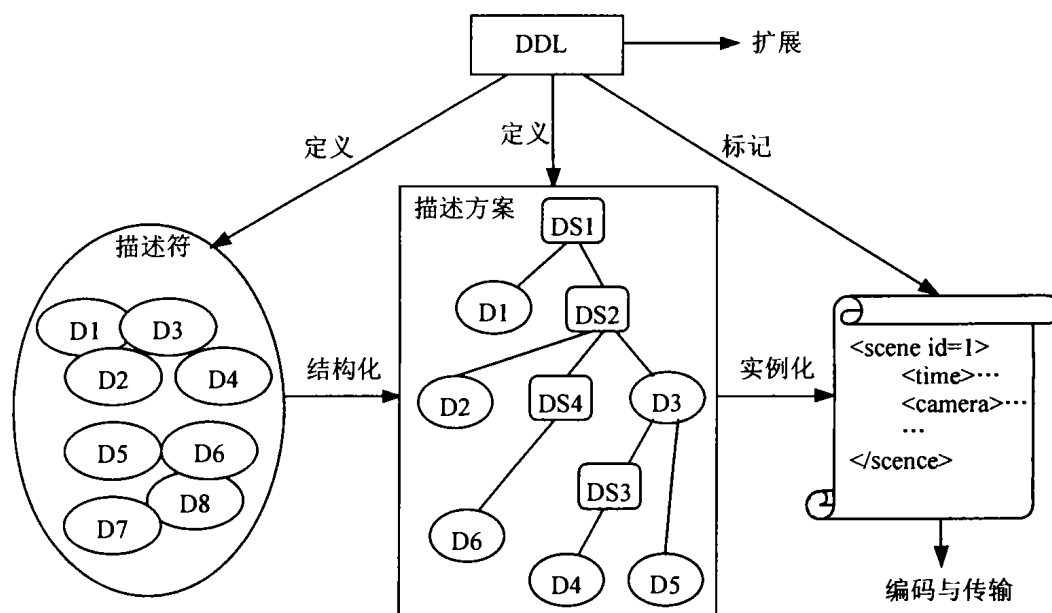


图1 MPEG-7 主要的组成部分及其关系

数字视频信息内容丰富,包含视觉特征、听觉特征、运动特征、文本特征等等。因此,基于内容的元数据描述呈现多态性,在设计应用系统时应综合利用视频信息的多种特征,即数字视频元数据应用向多模态(Multi-modality)方向发展。

4.2 数字视频元数据向混合模式方向发展

任何一种元数据标准都是基于特定领域发展起来的,即特定的元数据面向特定的应用,没有一种元数据标准能适应所有的应用背景。虽然 MPEG-7 是一个用于多媒体内容描述的国际标准,但它仅给出了一组描述符、描述方案和一套描述定义语言,而如何在具体的应用领域使用,以及如何如何进行视频内容的特征提取并未包含在 MPEG-7 标准中。因此,在具体的应用中,应根据视频信息的类型和应用需求,选择某种元数据标准或基于多种元数据标准混合的应用模型来进行视频信息的描述。

Jane Hunter 在文献[9]中给出了一种基于 Dublin Core 和 MPEG-7 的描述视频的一种应用规范,其基本模型如图 2 所示。

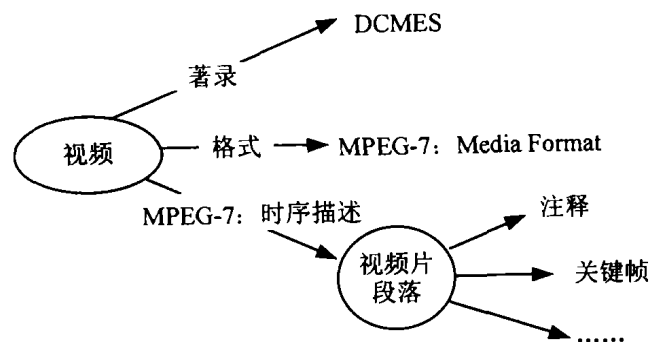


图 2 基于 Dublin Core 和 MPEG-7 的描述视频的基本模型

性(inter-operability)要求在由不同的组织制定与管理且技术规范不尽相同的元数据环境下,能够做到对用户一致性的服务。因此,一种可同时支持多种元数据来往于 Internet 和 WWW 上的架构,成为不可或缺的工具,资源描述框架(RDF, Resource Description Framework)借此而生。RDF 是一个描述和交换元数据的框架,提供了一种强有力的可转换的句法模式。RDF 本身只有很少定义,它采用 XML 作为交换和处理元数据的通用语法结构体系,其句法简单,采用“资源(Resource)——属性(Property)——字符串(Literal)”这样的“三元组”来表述,犹如自然语言中的“主语—谓语—宾语”。

可用于数字视频元数据表示的方案目前有多种。Jane Hunter 等人在文献[17]中针对可用于视频元数据表示的多种方案(RDF、XML DTDs、DCD、SOX 和 DDML)进行了比较,并总结出各自的优劣。从目前的技术发展状况来说,XML/RDF 模式是最有效的数字视频元数据表示模式。

5 结束语

数字视频信息将成为互联网上重要的信息资源,它将给人类的工作、生活、学习等带来重要影响,同时也预示着视频信息的广泛应用前景。针对不同的应用领域,制定出合适的元数据标准、研制出基于内容的元数据自动抽取方法,是实现数字视频信息有效存储、交换、共享、过滤、重用和检索的基础。因此,加强对视频信息的元数据描述和组织研究,将对视频信息的处理、共享和服务产生积极作用。

参 考 文 献

- 1 中国国家统计局 2001 年统计数据. <http://www.stats.gov.cn/tjsj/ndsj/index.htm>
- 2 李国辉,汤大权,武德峰. 信息组织与检索. 北京:科学出版社
- 3 José M. Martínez. MPEG-7 Overview. <http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm>
- 4 数字电视现行三种标准的比较. <http://www.lnnu.edu.cn/xdjy/jx/xsqy/gbds/szds.htm>
- 5 即时电视规范的研究及应用. <http://www.gmmchina.com/cn-artical5.htm>
- 6 Meng Yang, et al. A Multi-level, Multi-layer and Multi-channel Metadata Schema for the Open Video Digital Library. <http://www.ils.unc.edu/~mux/publications/ICADL2002.pdf>
- 7 Preservation Metadata for Digital Objects: a Review of the

4.3 以 XML 为基础、以 RDF 为框架对数字视频元数据编码

随着 XML 语言及其相关技术(XML 名称空间、XML 查询语言、XML 数据库等)的发展,XML 语言已逐渐成为描述网络资源的主流语言。将大量数字视频信息的描述存储在 XML 数据库中、利用 XML 查询语言进行查询,将成为数字视频信息(乃至各种多媒体信息)存储、管理和检索的必然趋势。

网络上的资源种类繁多、形式不一,元数据的种类复杂且用途殊异,多种元数据共存共荣的局面已经形成,这就要求能实现元数据的互操作。互操作

- State of the Art. http://www.oclc.org/research/pmwg/presmeta_wp.pdf
- 8 Jay Loomis. Windows Media 元数据使用指南. http://www.microsoft.com/china/msdn/library/dnwm/html/WM_Metadata_Usage.asp
- 9 Jane Hunter. An Application Profile which Combines Dublin Core and MPEG-7 Metadata Terms for Simple Video Description. http://metadata.net/harmony/video_appln_profile.html
- 10 Jane Hunter, et al. The application of metadata standards to video indexing. In: Second European Conference on Research and Advanced Technology for Digital Libraries, Crete, Greece, September 1998.
- 11 ViDe User's Guide: Dublin Core Application Profile for Digital Video. <http://www.vide.net/workgroups/videoaccess/resources.shtml>
- 12 Dublin Core Metadata Record for Digital Video. <http://www.vide.net/workgroups/videoaccess/resources.shtml>
- 13 卢坚等. 远程教学数字图书馆中基于 MPEG-7 的语音数据的组织与索引. 情报学报, 2002, 21(4)
- 14 Andrew Graves, Mounia Lalmas. Video Retrieval Using an MPEG-7 Based Inference Network. SIGIR'02
- 15 M. Bober. MPEG-7 Visual Shape Descriptors. <http://cv.cs.tu-berlin.de/lectures/MPEG7-3.pdf>
- 16 MPEG-7 Based Image/Video Retrieval System. http://vision.hanyang.ac.kr/research/mpeg7_retrieval/mpeg7_retrieval.htm
- 17 Jane Hunter, et al. A Comparison of Schemas for Dublin Core-based Video Metadata Representation. <http://archive.dstc.edu.au/RDU/staff/jane-hunter/mpeg7/contribution.html>

(责任编辑 许增棋)