

基于概念本体的视频内容分析框架

张 良, 周长胜

(北京信息科技大学 计算中心, 北京 100192)

摘 要:为解决视频内容分析过程中对视频组织结构划分的问题, 建立了一个基于概念本体的视频分析模型, 将视频以概念本体的方式进行划分, 归纳出视频中四种典型的概念: 类别概念、对象概念、属性概念以及情感概念。论述了各类概念之间的关系: 实例关系、属性关系、归类关系、组成关系。采用人工定义规则与学习方法相结合的方法, 实现对视频概念本体的识别和划分。以概念本体为中心划分视频结构, 符合人类认知事物过程和认知规律, 便于将底层特征向量与高层语义进行映射、组织和处理。

关键词: 视频; 内容分析; 概念本体

中图分类号: TP391

文献标识码: A

文章编号: 1673-629X(2011)12-0116-04

Video Content Analysis Framework Based on Concept Ontology

ZHANG Liang, ZHOU Chang-sheng

(Computer Center of Beijing University of Information & Science Technology, Beijing 100192, China)

Abstract: In order to solve the video structure problem of semantic analysis, created a video content analysis framework based on concept ontology. The video is divided by the concept ontology, it summarized the concept of four typical video: category concept, object concepts, property concepts and emotional concepts. It also discusses the relationship between various concepts: instance-of, attribute-of, kind-of, part-of. The video structured by concept ontology consistent with human cognitive processes and cognitive rules of things. The concept ontology as the center can establish good contact between the underlying feature vector and high-level semantic, to the easy mapping, organization and processing.

Key words: video; content analysis; concepts ontology

0 引 言

目前, 数字视频正大量的产生, 并不断在互联网上进行发布、传播。对比文本类型的数据, 视频数据能够提供事件最直接和真实的描述。但视频数据存在检索和管理困难的缺点。人工从大量的视频数据中检索出人们需要的数据是一项耗时耗力的工作, 因此需要一种自动高效的方式从视频中提取语义的方法。

视频数据的分析和检索存在以下难题:

视频数据不便于交流。文本数据以人类语言为基础, 是为人类交流而存在的工具, 在数据的管理和检索上有天然的优势。而在数据处理方面可以采用文本匹配、自动分词等技术, 可以基本满足数据检索的需求, 当然在文本包含的高层语义的挖掘方面有待进一步的研究。而视频数据是以人类直观感觉为基础, 记录的

是人的直观感知内容, 并不以交流为主要目的。数据视频的存储形式在计算机中一般以视频文件的方式, 即二进制格式, 以直接或者压缩格式记录了视频的像素信息, 主要是像素点的色彩、亮度信息。这些被称为低层信息, 即感知层信息, 而人们需要检索的是视频中包含的语义信息, 之间存在所谓的语义鸿沟^[1,2]。感知层的数据无法通过简单匹配、分词等方法进行数据的检索。

视频类型数据缺少有效的组织结构。文本数据有字、词、句、段落、文章等方式进行组织, 随着人类语言的发展演变, 已经有了相对成熟的组织结构。而视频数据, 目前还没有一个成熟的、权威的组织结构, 在视频的研究过程中, 人们一般采用视频帧、镜头、场景等划分方式, 但视频镜头和场景的划分算法并不成熟, 而镜头只是帧的集合, 对语义分析来说并没有特殊的意义。

由于视频的内容一般可分为三层: 感知层、认知层、情感层。三层之间的映射关系缺少一个统一的系统框架进行“语义鸿沟”的跨越。在三层内容中, 认知层和情感层与人类的主观感受相关, 与每个人的认知

收稿日期: 2011-05-16; 修回日期: 2011-08-21

基金项目: 北京市属高等学校人才强教计划资助项目 (PHR2010 08447)

作者简介: 张 良 (1980-), 男, 讲师, 研究方向为视频语义分析、数据挖掘。

能力、情感取向有很大关系,有不确定性^[3]。

目前业内主要研究方式是在视频底层特征与高层语义之间建立映射,桥接语义鸿沟,主要是通过视频帧的低层特征实现对高层信息的映射,基于对象的特征进行语义分析和提取,以对象检测到概念分类的步骤对视频进行分析^[4]。以采用目标跟踪等方法,实现了从视频中提取语义事件的方法,应用于交通事件的分析 and 检索。而交通事件一般相对简单,事件也相对固定^[5,6]。该框架很难应用于其他复杂的通用视频中去。

为解决以上问题,文中建立一个基于 Ontology 的视频分析模型。将视频以概念的方式进行划分,解决了视频的组织结构问题。以概念划分的视频结构,符合人类认知事物过程和认知规律。以概念为中心,可以在底层特征向量与高层语义间建立很好的联系,便于进行映射、组织和处理。

1 相关知识

1.1 人脑的认识过程

所有的视频信息都是经过人脑的“接收-过滤-筛选-反映”的过程。所以对视频的分析,也需要将人的主观接受过程考虑进去。人们在观看视频文件时,一般与人们对现实世界的认知过程是一致的,人脑中存在着一个概念模型,用于人们对现实世界事物的认知。

图1 为人脑识别视频的过程。

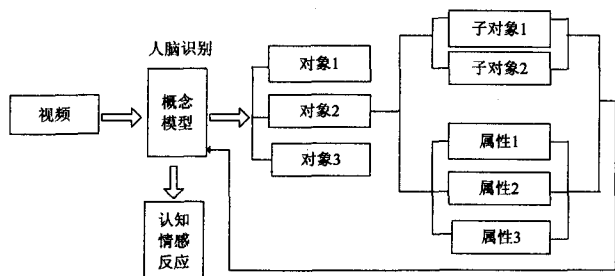


图1 人脑识别视频的过程

视频经过人脑中的概念模型的识别后,形成了若干个对象,这些对象可以再被细分为若干子对象。也可以识别出这些对象的若干属性。这些对象、子对象及其属性的变化与概念模型共同作用引发了人类的主观反应^[7]。从这个过程可以看出,视频对象识别最重要的工作,是建立一个模拟人脑的一个科学、合理的概念模型。视频语义识别和情感识别的工作都要围绕概念模型而展开。

1.2 Ontology(本体)模型

Ontology 原来是哲学词汇,一般将其译为“本体”或“存在”。主要是指用 Ontology 如何对现实世界客观存在的事物进行分类、描述、定义的过程。解决客观世界的抽象以及空间与时间的关系等问题。在计算机领域实现为面向对象的程序设计方法。

世界上存在着的一切事物,可以称之为实体。实体拥有各种属性,这些属性可以进行量化,并且会随着时间的变化而变化。实体是可以区别的,在客观世界中是唯一存在的。也就是说实体有属性会与其他属性相区别,以实现其唯一性。各种对象之间可能存在着各种关联。这些关联也与时间的变化相关。对象会参与各种事件的流程,这些事件和流程在不同的时间会有不同的变化。实体可以进行抽象,形成各种类别。

我们可以与那些对该领域的知识表示有类似需求的人分享这种知识代表语言,因此避免了重复知识分析流程。共享的 Ontology 形成了领域专门知识表示语言的基础。Ontology 包含5个基本的建模元语(Modeling Primitive)。这些元语分别为:概念(concepts)、关系(relations)、函数(functions)、公理(axioms)和实例(instances)。

基本的关系有4种:part-of(部分)、kind-of(归类)、instance-of(实例)和 attribute-of(属性)。Ontology 是对共享概念模型的规范说明,这里所说的“概念模型”指该模型中的概念是公认的,至少在某个特定的领域是公认的。一般情况下,Ontology 是面向特定领域,用于描述特定领域的概念模型^[8]。

视频数据是现实世界最真实的记录,对视频进行处理,首先要从现实世界以及人对现实世界的认知角度进行分析。从人的主观角度,现实世界的模型通常由实体(对象)构成。以本体论的分析方法可以将视频内容划分为各种本体。

2 概念本体

概念本体是人认知范围的一切事物的抽象。在人们认识世界的过程中,会在大脑中形成一系列的概念,这些概念大致分为以下几种:

2.1 类别概念

指人们在长期的生活实践中总结和抽象出来的对事物的分类,如人物、动物、植物、汽车、水果等。而现实世界中的事物分类方法多种多样,并没有统一的标准。在视频分析中,需要根据视频分析的具体应用领域确定该领域中的类别概念。如在足球运动视频分析中,可以划分视频类别概念集合为{场地、球员、裁判、观众、足球...},而在交通视频分析中可以划分视频类别概念集合为{公路、人、机动车、非机动车、交通信号灯、交通标志...}^[9]。

2.2 对象概念

指人们对某一具体的事物的识别,对象识别一直是图像和视频识别领域中的一个重要应用,如人脸识别、车牌号识别等等。通过对象识别,得到关于语义内容的第二个层次的内容,即认知内容^[10]。

2.3 属性概念

是指人们视觉对事物的直接感知内容。属性概念往往不是独立存在的,一般是附属于某一类别本体或对象本体,如汽车的颜色、汽车的形状等。属性本体是视频中唯一可以直接从文件中获取的信息。在视频分析中,对象本体的识别、情感概念的判定往往都是由属性本体的划分、计算、学习等方法得到的。

2.4 情感概念

是视频分析中最高层语义,是指人通过观看视频,得到的主观情感上的反应,如高兴、愤怒、悲伤等。情感概念分析是视频语义分析中最复杂、同时也是最重要的分析之一。通过对情感概念的分析,可以有效的对视频进行分类管理,进行符合人主观情感的筛选、摘要等工作。

这里进行视频分析的概念本体模型定义:视频中能够表现的事物,称之为概念本体,概念模型是这些概念本体的集合,记为 $C\{c_1, c_2, \dots, c_n\}$ 。类别概念本体集合记为 $K\{k_1, k_2, \dots, k_n\}$, 对象概念本体集合记为 $O\{o_1, o_2, \dots, o_n\}$, 属性概念本体集合记为 $A\{a_1, a_2, \dots, a_n\}$, 情感概念本体集合记为 $E\{e_1, e_2, \dots, e_n\}$ 。

这里,可以得到:

$$C = K \cup O \cup A \cup E \quad (1)$$

从公式(1)可以得到 K, O, A, E 四个集合是概念集合 C 的真子集。

3 概念本体间的关系

概念模型中, K, O, A, E 四个集合中各种概念本体之间存在各种关系(relations)。

3.1 实例关系(instance-of)

当对视频进行分析时,会在视频中分出若干个视频对象,这些对象属于概念模型中对象本体集合。同时,这些对象本体也是类别概念集合中的一个实例,它们应该纳入概念分类。该对象是经常被用来作为视频的基本单元,这些对象是作为一个概念分类。然后,它们获得概念属性。如视频中出现的一辆汽车属于某一对象本体,同时,它又是类别概念中的一个实例。从而可以得出从类别本体集合到对象本体集合的一个关系,即实例关系。这里用 $\text{instance-of}(K_i, O_j)$ 表示对象本体 O_j 是类别本体 K_i 的一个实例。

3.2 属性关系(attribute-of)

属性本体和本体的属性:

在视频数据中,指视频中存在的各种低层特征的集合,如颜色、形状。属性本体是将这些属性抽象出来,形成若干属性类别,称之为属性本体。

上一节提到,属性本体往往不是独立存在,是依附于某类别本体或对象本体,本体的属性通常包括颜色、

纹理的形状、位置和其他低级别的对象特征。这些数据是直接获得的视频数据。与此同时,它们也是从视频中提取视频对象的基本依据。该属性的关系包括概念属性和对象属性。这里用 $\text{Attribute-of}(C_i, A_j)$ 表示属性本体 A_j 与类别本体 C_i 构成属性关系,即 C_i 具有 A_j 所标识的属性。

如图 2 所示,视频中颜色、形状等称之为属性本体,而定义汽车为类别本体, $\text{Attribute-of}(\text{car}, \text{color})$ 标识汽车本体具有颜色本体所表示的属性。即汽车具有颜色属性。同理, $\text{Attribute-of}(\text{car}, \text{shape})$ 表示汽车具有形状属性。 $\text{attributes-of}(\text{car}) = \{\text{color}, \text{shape}, \dots\}$ 与汽车构成是属性关系的所有属性本体的集合。

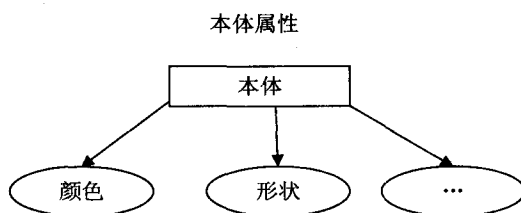


图 2 本体的属性

3.3 归类关系(kind-of)

视频中的本体,也可以按层次来进行管理。一些概念本体细分为另一些概念本体。其中,类别本体比较明显。在人们认识现实世界或进行视频分析的过程中,事物往往先被划分为若干大类,这些大类再进行细分,得到若干小类别,这些小类别可以继续细分。这就构成了视频本体分析中的归类关系。这里,用 $\text{kind-of}(K_i, K_j)$ 表示类别本体 K_j 是类别本体 K_i 的一个子类别。图 3 为层次关系举例。

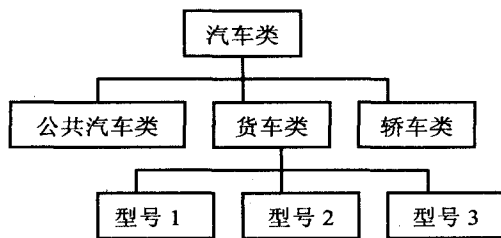


图 3 层次关系举例

人在识别该事物是一个从上至下的过程,汽车→货车→型号。具体的人对该事物的认知与大脑中固有的观念、知识结构是相关的。如果人不懂汽车的分类那么他的认识深度就认为该事物是汽车,如果懂汽车分类但对品牌型号不懂,他就认知该事物为货车,如果是专业人士,可能会认识到某一个品牌,某一个型号。人对事物的认识是逐层深入的,那么对视频的分析也应该是逐层深入的。

3.4 组成关系(part-of)

从概念本体的逻辑角度,得到本体间的层次关系,如果从本体间物理关系,可以得到对象间的组成关系。

即一些本体是另一些本体的组成部分。在类别本体和对象本体中,存在组成关系。这里使用 $\text{part-of}(K_i, K_j)$ 表示概念本体 K_j 是 K_i 的组成部分, $\text{part-of}(O_i, O_j)$ 表示对象本体 O_j 是 O_i 的组成部分。组成关系举例见图4。

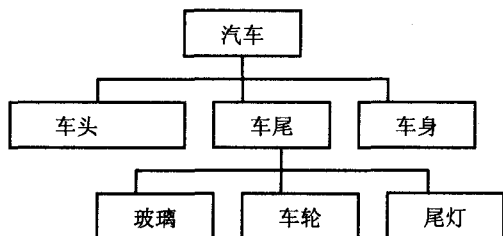


图4 组成关系举例

3.5 概念本体间的函数体系

在上一节中,定义了对象以及对象间的关系,下面进行对象间的推理体系:

推理1,如果 K_j 与 K_i 构成归类关系,那么 K_j 的属性本体集合是 K_i 属性本体集合的一个子集。

推理2,如果 C_k 是 C_j 的一个子类,且 C_j 是 C_i 的一个子类,那么 C_k 与 C_i 构成从属关系,即 C_k 是 C_i 的一个子类。

推理3,如果对象本体 O_k 是类别本体 C_j 的一个实例,且 C_j 是 C_i 的一个子类,那么, O_k 也是 C_i 的一个子类。

推理4,如果 C_j 是 C_i 的组成部分,且 C_k 是 C_j 的组成部分,那么,可以得到 C_k 是 C_i 的组成部分。

4 视频语义分析模型建立

基于概念本体的视频语义分析的核心是建立一个类似人脑识别反应系统的概念模型。该模型需要包含以下几点:

1) 合理的概念本体集合的建立和定义。这里的概念本体包括四种类型的概念本体。类型库相当于人脑中的知识系统。

2) 视频本体的识别方法,建立了视频本体集合。下面就需要对视频中出现的事物进行有效的识别,这里一般采用两种方法:

一是基于规则的方法,即人为根据对象的某些特征定义规则,然后根据规则,将对象进行归纳,形成一个概念本体集合,这里需要据特定应用领域的先验知识。但视频的种类有多种,而特定应用领域很多,就需要为每一种领域建立一个规则比较适合于较为简单的视频^[11,12]。

二是基于学习的方法,对通用视频,较为适合的就

是基于学习的算法^[13],算法流程如图5所示。

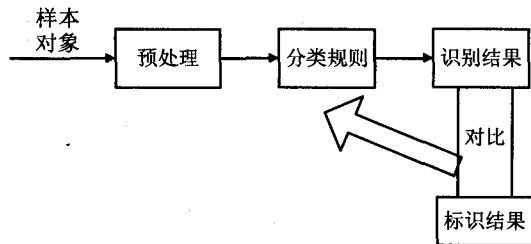


图5 概念本体识别流程

文中采用人工定义规则与学习方法相结合的办法,先人工定义规则,样本对象是人工标识后的样本,该样本采用预先设定的规则进行识别,得到识别结果,该识别结果与人工标识的结果进行对比,如果正确加强该规则,如果错误修改该规则。图6为概念本体识别算法。

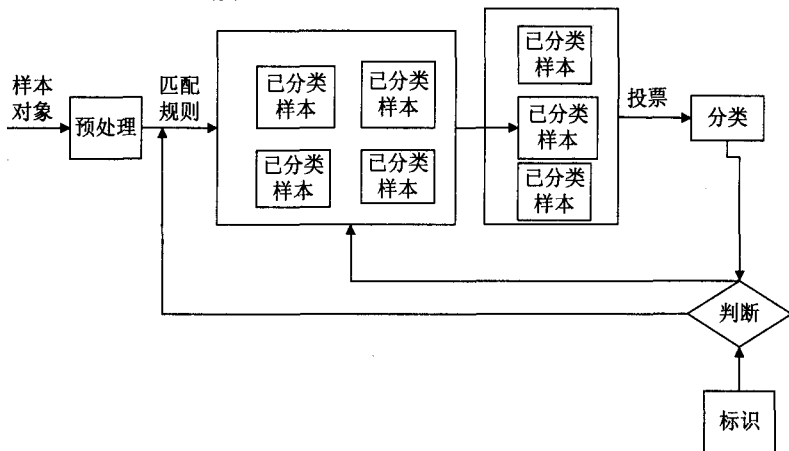


图6 概念本体识别算法

5 结束语

文中建立一个基于概念本体的视频分析模型。文中分析了视频数据在内容分析和检索方面存在的难题,将视频以概念的方式进行划分,解决了视频语义分析中的组织结构问题。论述了视频概念本体的定义、分类以及各类概念本体间的关系。建立了概念本体间的函数体系,以概念划分的视频结构,符合人类认知事物过程和认知规律。以概念本体为中心,便于将底层特征向量与高层语义进行映射、组织和处理。对通用视频,采用基于学习的算法,可以对视频中的概念本体进行合理的划分。

参考文献:

- [1] 林新棋. 基于模糊理论的电影情感识别[D]. 北京:北京邮电大学,2009.
- [2] Hanjalic A. Content-Based Analysis of Digital Video[M]. [s.l.]:Springer,2004.
- [3] Hanjalic A, Xu Liqun. Affective Video Content Representation

Hibernate 开发框架实现数据持久层处理的功能,关键是通过配置文件(hibernate. properties 或 hibernate. cfg. xml)和映射文件(. hbm. xml)把 java 对象或持久化对象映射到数据库中的数据表,然后通过操作持久化对象,对数据表中的数据进行增、删、改、查等操作,其中 hibernate. cfg. xml 配置文件设置如下:

```
<hibernate-configuration>
<session-factory>
.....
<propertyname="hibernate. connection. driver_class">
com. mysql. jdbc. Driver</property>
.....
<propertyname="hibernate. connection. url">
jdbc:mysql://localhost/hibernate</property>
.....
<property name="hibernate. connection. username">
SCIPMS</property>
<property name="hibernate. connection. password">
root</property>
<property name="connection. characterEncoding">UTF
-8</property>
<propertyname="hibernate. dialect">
org. hibernate. dialect. MySQLDialect</property>
.....
<!--Mapping files-->
< mapping resource="hibernate/SCIPMS/UserInfo.
hbm. xml"/>
.....
</session-factory>
</hibernate-configuration>
```

4 结束语

文中首先简要介绍 MVC 设计模式及其典型实现框架 Struts 的运行原理,其次介绍了 ORM 模型和 Hi-

bernate 的实现方式,并分析了两者的优缺点,最后将这两个框架组合应用到本系统实际开发中,实践证明采用 MVC 模式开发、分层架构,简化了 Web 应用程序的开发过程,提高开发效率,缩短了开发周期,提高系统的可扩展性。

参考文献:

- [1] 匡芳君. 基于 MVC 模式的网络教学平台设计[J]. 计算机技术与发展,2010,20(12):251-255.
- [2] 杨磊,陈凌云. 精通 Eclipse Web 开发—Java 体系结构、工具、框架及整合应用[M]. 北京:人民邮电出版社,2006.
- [3] Alur D, Crupi J, Malsk D. Core J2EE Patterns: Best Practices and Design Strategies[M]. American: Prentice Hall, 2003.
- [4] Johnson R. J2EE Development Frameworks[J]. IEEE Computer,2005,38(1):107-110.
- [5] 姬朝阳,唐红喜. 基于 SSH 的日志统计分析系统的分析与设计[J]. 计算机技术与发展,2010,20(8):212-217.
- [6] 付京周. Hibernate 3.0—Java 数据库持久层开发实践[M]. 北京:人民邮电出版社,2007.
- [7] 郭秀娟,王春光. 基于 B/S 模式的毕业设计管理系统开发与实现[J]. 计算机技术与发展,2010,20(3):239-244.
- [8] 张志杰. 基于分层结构的管理信息系统架构设计[J]. 计算机技术与发展,2010,20(10):146-151.
- [9] 孟斌,吴建波,冷文浩. 基于 Struts 和 Hibernate 架构的 Shids 系统设计与实现[J]. 计算机工程与设计,2006,27(21):4101-4104.
- [10] 许锋,周慧英. Java Web 整合开发全程指南:基础、进阶(Struts+Spring+Hibernate)、实战[M]. 北京:电子工业出版社,2009.
- [11] Beauchemin B, Berglund N, Sullivan D. A First Look at SQL Server 2005 for Developers[M]. Boston: Addison-Wesley, 2004.
- [12] 任晓鹏,赵文兵,张春平. 基于框架的 Web 系统开发研究[J]. 计算机工程与设计,2010,31(4):772-775.

(上接第 119 页)

- and Modeling[J]. IEEE transactions on multimedia,2005,17(1):143-154.
- [4] Lin Liang, Gong Haifeng, Wang Liang. Semantic event representation and recognition using syntactic attribute graph grammar[J]. Pattern Recognition Letters,2009,30:180-186.
- [5] 雷云,王夏黎,孙华. 基于视频的交通目标跟踪方法研究[J]. 计算机技术与发展,2010,20(7):44-47.
- [6] 侯伟,卢炎麟,郑河荣,等. 固定背景下的视频分割及在交通视频流的应用[J]. 计算机技术与发展,2008,18(9):191-193.
- [7] 郭旭,张丽杰. 运动目标检测视频监控软件的设计与实现[J]. 计算机技术与发展,2010,20(8):199-207.

- [8] 王海林. 基于 OWL DL 的学科知识本体建模[J]. 计算机技术与发展,2010,20(8):108-111.
- [9] 刘翔,吴谨,祝愿博. 基于视频序列的目标检测与跟踪技术研究[J]. 计算机技术与发展,2009,19(11):179-182.
- [10] 魏维,邹书蓉,刘凤玉. 多层视频语义概念分析与理解[J]. 计算机辅助设计与图形学学报,2008,20(1):85-92.
- [11] 刘骏伟. 视频对象检测及其在视频语义内容分析中的应用[D]. 杭州:浙江大学,2003.
- [12] 孙君顶,赵珊. 图像低层特征提取与检索技术[M]. 北京:电子工业出版社,2009.
- [13] 郭丽,龚声蓉. 基于内容的视频分割技术研究[J]. 计算机技术与发展,2008,18(9):33-36.