



(12) 发明专利

(10) 授权公告号 CN 108959323 B

(45) 授权公告日 2021.12.07

(21) 申请号 201710382359.3

CN 103226569 A, 2013.07.31

(22) 申请日 2017.05.25

CN 105868686 A, 2016.08.17

(65) 同一申请的已公布的文献号

CN 105912682 A, 2016.08.31

申请公布号 CN 108959323 A

CN 1298522 A, 2001.06.06

(43) 申请公布日 2018.12.07

CN 106294783 A, 2017.01.04

(73) 专利权人 腾讯科技(深圳)有限公司

CN 103714130 A, 2014.04.09

地址 518000 广东省深圳市南山区高新区

WO 2016190945 A1, 2016.12.01

科技中一路腾讯大厦35层

朱义明.“基于社交信息的网络视频分类”.《中国优秀硕士学位论文全文数据库信息科技辑》.2012, (第04期),

专利权人 山东财经大学

艾丽丽 等.“基于文本描述的视频分类系统建模”.《信息通信》.2013, (第4期),

(72) 发明人 聂秀山

(74) 专利代理机构 北京康信知识产权代理有限

Xiao Wu 等.“Towards Google Challenge: Combining Contextual and Social

责任公司 11240

代理人 赵囡囡 褚敏

Information for Web Video

(51) Int. Cl.

G06F 16/78 (2019.01)

G06K 9/62 (2006.01)

G06Q 50/00 (2012.01)

Categorization”.《MM ’09: Proceedings of the 17th ACM international conference on Multimedia》.2009,

李慧 等.“融合主题与语言模型的个性化标签推荐方法研究”.《计算机科学》.2015, 第42卷(第8期),

(56) 对比文件

CN 105913072 A, 2016.08.31

CN 106411679 A, 2017.02.15

CN 102160084 A, 2011.08.17

审查员 倪礼

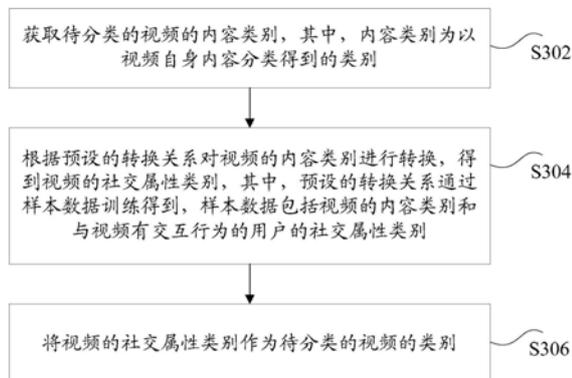
权利要求书3页 说明书12页 附图4页

(54) 发明名称

视频分类方法和装置

(57) 摘要

本发明公开了一种视频分类方法和装置。其中,该方法包括:获取待分类的视频的内容类别,其中,内容类别为以视频自身内容分类得到的类别;根据预设的转换关系对视频的内容类别进行转换,得到视频的社交属性类别,其中,预设的转换关系通过样本数据训练得到,样本数据包括视频的内容类别和与视频有交互行为的用户的社交属性类别;将视频的社交属性类别作为待分类的视频的类别。本发明解决了由于视频分类时仅考虑视频自身内容造成的视频分类结果不贴近用户的技术问题。



1. 一种视频分类方法,其特征在于,包括:

获取待分类的视频的内容类别,其中,所述内容类别为以视频自身内容分类得到的类别;

获取第一用户群中用户的社交属性类别,根据所述用户的社交属性类别为第一类视频添加社交属性类别,其中,所述第一用户群为所有与所述第一类视频具有交互行为的用户的集合;

获取样本集合中所有视频的社交属性类别,通过字典学习算法得到所述社交属性类别和所述视频的内容类别之间转换关系,将所述转换关系作为预设的转换关系;

根据所述预设的转换关系对所述视频的内容类别进行转换,得到视频的社交属性类别;

将所述视频的社交属性类别作为所述待分类的视频的类别。

2. 根据权利要求1所述的方法,其特征在于,用户的社交属性类别包括第一社交属性类别,获取所述第一用户群中用户的社交属性类别包括获取所述第一用户群中用户的第一社交属性类别,获取所述第一用户群中用户的第一社交属性类别包括:

建立所有用户群的超图模型,其中,所述超图的点由一个用户的一个社交属性类别构成,所述超图的边由一个用户的所有社交属性类别构成;

通过谱图聚类算法对所述所有用户群的超图模型进行聚类计算,得到所有用户群的第一社交属性类别;

从所有用户群的第一社交属性类别中查找到所述第一用户群的第一社交属性类别。

3. 根据权利要求1所述的方法,其特征在于,用户的社交属性类别包括第二社交属性类别,获取所述第一用户群中用户的社交属性类别包括获取所述第一用户群中用户的第二社交属性类别,获取所述第一用户群中用户的第二社交属性类别包括:

获取所述第一用户群在预设时间段内的社交信息;

通过在线LDA模型抽取所述社交信息的主题内容,得到所述第一用户群在所述预设时间段内的主题内容信息;

根据所述主题内容信息确定所述第一用户群中用户的第二社交属性类别。

4. 根据权利要求1所述的方法,其特征在于,

获取待分类的视频的内容类别包括:通过在ImageNet上预训练的VGG模型得到的全连接层的特征作为所述待分类的视频的内容类别,

通过字典学习算法在所述社交属性类别和所述视频的内容类别之间建立所述预设的转换关系包括:

通过函数 $\arg \min_{B,s} \frac{1}{2} \|V - Bs\|_2^2 + \lambda_1 \|s\|_1 + \lambda_2 \phi(s)$ 计算得到使所述视频的社交属性类别和

内容类别误差最小的转换关系,其中,V表示训练集中的视频内容类别和社交属性类别连接的向量,B和s分别代表一组基和V在所述基上的系数, $\phi(s)$ 代表训练集中不同向量的稀疏系数的距离函数,所述距离函数通过高斯核函数确定,

将所述转换关系作为所述预设的转换关系。

5. 一种视频分类装置,其特征在于,包括:

第一获取单元,用于获取待分类的视频的内容类别,其中,所述内容类别为以视频自身内容分类得到的类别;

第二获取单元,用于获取第一用户群中用户的社交属性类别,其中,所述第一用户群为所有与第一类视频具有交互行为的用户的集合;

添加单元,用于根据所述用户的社交属性类别为所述第一类视频添加社交属性类别;

第三获取单元,用于获取样本集合中所有视频的社交属性类别;

计算单元,用于通过字典学习算法得到所述社交属性类别和所述视频的内容类别之间转换关系,将所述转换关系作为预设的转换关系;

转换单元,用于根据所述预设的转换关系对所述视频的内容类别进行转换,得到视频的社交属性类别;

处理单元,用于将所述视频的社交属性类别作为所述待分类的视频的类别。

6. 根据权利要求5所述的装置,其特征在于,用户的社交属性类别包括第一社交属性类别,获取所述第一用户群中用户的社交属性类别包括获取所述第一用户群中用户的第一社交属性类别,所述第二获取单元包括:

建立模块,用于建立所有用户群的超图模型,其中,所述超图的点由一个用户的一个社交属性类别构成,所述超图的边由一个用户的所有社交属性类别构成;

计算模块,用于通过谱图聚类算法对所述所有用户群的超图模型进行聚类计算,得到所有用户群的第一社交属性类别;

查找模块,用于从所有用户群的第一社交属性类别中查找到所述第一用户群的第一社交属性类别。

7. 根据权利要求5所述的装置,其特征在于,用户的社交属性类别包括第二社交属性类别,获取所述第一用户群中用户的社交属性类别包括获取所述第一用户群中用户的第二社交属性类别,所述第二获取单元包括:

获取模块,用于获取所述第一用户群在预设时间段内的社交信息;

抽取模块,用于通过在线LDA模型抽取所述社交信息的主题内容,得到所述第一用户群在所述预设时间段内的主题内容信息;

确定模块,用于根据所述主题内容信息确定所述第一用户群中用户的第二社交属性类别。

8. 根据权利要求5所述的装置,其特征在于,

所述第一获取单元用于通过在ImageNet上预训练的VGG模型得到的全连接层的特征作为所述待分类的视频的内容类别,

所述计算单元用于:

通过函数 $\arg \min_{B,s} \frac{1}{2} \|V - Bs\|_2^2 + \lambda_1 \|s\|_1 + \lambda_2 \phi(s)$ 计算得到使所述视频的社交属性类别和

内容类别误差最小的转换关系,其中, V 表示训练集中的视频内容类别和社交属性类别连接的向量, B 和 s 分别代表一组基和 V 在所述基上的系数, $\phi(s)$ 代表训练集中不同向量的稀疏系数的距离函数,所述距离函数通过高斯核函数确定,将所述转换关系作为所述预设的转换关系。

9. 一种存储介质,其特征在于,所述存储介质包括存储的程序,其中,在所述程序运行

时控制所述存储介质所在设备执行权利要求1至4中任意一项所述的视频分类方法。

10. 一种处理器,其特征在于,所述处理器用于运行程序,其中,所述程序运行时执行权利要求1至4中任意一项所述的视频分类方法。

视频分类方法和装置

技术领域

[0001] 本发明涉及视频领域,具体而言,涉及一种视频分类方法和装置。

背景技术

[0002] 视频分类和内容识别系统以视频自身内容为中心,依赖特征的选取和构造进行视频分类,传统的方法使用静态视觉特征、声音特征和运动特征等进行视频内容的识别和分类。近年来,随着深度学习研究热度的不断提升,基于CNN网络等通过学习得到的特征也被用来进行视频内容的识别和分类。不管采用何种特征,现有视频内容分类技术都是基于视频本身内容,图1是现有技术的视频内容分类的框架图,如图1所示,通过人工设计特征或深度学习网络学习特征对视频内容进行分类,将视频分为多个类别。

[0003] 现有的视频内容分类技术主要是以“视频自身内容”为中心的思路,主要缺点如下:忽略了用户,即人的主观意愿和兴趣。互联网视频的消费者和散播者都是人,因此用户的兴趣和意愿,以及用户群体所处的环境和观看视频时的情绪等主观因素对视频内容的分类、个性化推荐等应用具有一定的影响,发明人经过研究发现,用户不同的知识背景、兴趣、关注点等主观因素会对相同视频内容产生不同的分类判别,这类比于,对于同一个公众人物,例如贝克汉姆,有足球背景的人会把他归为“球星”,但是,没有体育背景的用户可能会把他归为“娱乐明星”。对视频分类来说,同样存在这个问题,例如,比较著名的“李世石和阿尔法的围棋人机大战”的视频,具有围棋背景或在围棋等竞技类项目用户可能把此类视频归为“竞赛”或“围棋比赛”,而对信息技术比较感兴趣的用户可能会把此类视频归为“科技”或“人工智能”的类别。所以在视频内容分类技术中仅以视频自身内容分类会使得视频分类结果不够贴近用户。

[0004] 针对上述的问题,目前尚未提出有效的解决方案。

发明内容

[0005] 本发明实施例提供了一种视频分类方法和装置,以至少解决由于视频分类时仅考虑视频自身内容造成的视频分类结果不贴近用户的技术问题。

[0006] 根据本发明实施例的一个方面,提供了一种视频分类方法,包括:获取待分类的视频的内容类别,其中,所述内容类别为以视频自身内容分类得到的类别;根据预设的转换关系对所述视频的内容类别进行转换,得到视频的社交属性类别,其中,所述预设的转换关系通过样本数据训练得到,所述样本数据包括视频的内容类别和与所述视频有交互行为的用户的社交属性类别;将所述视频的社交属性类别作为所述待分类的视频的类别。

[0007] 根据本发明实施例的另一方面,还提供了一种视频分类装置,包括:第一获取单元,用于获取待分类的视频的内容类别,其中,所述内容类别为以视频自身内容分类得到的类别;转换单元,用于根据预设的转换关系对所述视频的内容类别进行转换,得到视频的社交属性类别,其中,所述预设的转换关系通过样本数据训练得到,所述样本数据包括视频的内容类别和与所述视频有交互行为的用户的社交属性类别;处理单元,用于将所述视频的

社交属性类别作为所述待分类的视频的类别。

[0008] 根据本发明实施例的另一方面,还提供了一种存储介质,该存储介质包括存储的程序,在程序运行时控制存储介质所在设备执行本发明实施例的视频分类方法。

[0009] 根据本发明实施例的另一方面,还提供了一种处理器,该处理器用于运行程序,所述程序运行时执行本发明实施例的视频分类方法。

[0010] 在本发明实施例中,采用根据预设的转换关系对视频的内容类别进行转换,得到视频的社交属性类别,根据视频的内容类别得到待分类的视频的社交属性类别的方式,达到了根据视频的社交属性对视频进行分类的目的,从而实现了使视频分类结果更贴近用户的技术效果,进而解决了由于视频分类时仅考虑视频自身内容造成的视频分类结果不贴近用户的技术问题。

附图说明

[0011] 此处所说明的附图用来提供对本发明的进一步理解,构成本申请的一部分,本发明的示意性实施例及其说明用于解释本发明,并不构成对本发明的不当限定。在附图中:

[0012] 图1是现有技术的视频内容分类的框架图;

[0013] 图2是根据本发明实施例的视频分类方法的硬件环境的示意图;

[0014] 图3是根据本发明实施例的一种可选的视频分类方法的流程图;

[0015] 图4是根据本发明实施例的一种视频分类方法的流程图;

[0016] 图5是根据本发明实施例的一种社交网络超图建模的示意图;

[0017] 图6是根据本发明实施例的一种用户关联关系的建立过程的流程图;

[0018] 图7是根据本发明实施例的动态社交属性抽取过程的流程图;

[0019] 图8是根据本发明实施例的社交属性与视频内容分类统一表达的流程图;

[0020] 图9是根据本发明实施例的一种可选的视频分类装置的示意图;以及

[0021] 图10是根据本发明实施例的一种终端的结构框图。

具体实施方式

[0022] 为了使本技术领域的人员更好地理解本发明方案,下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分的实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都应当属于本发明保护的范围。

[0023] 需要说明的是,本发明的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象,而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换,以便这里描述的本发明的实施例能够以除了在这里图示或描述的那些以外的顺序实施。此外,术语“包括”和“具有”以及他们的任何变形,意图在于覆盖不排他的包含,例如,包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元,而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

[0024] 实施例1

[0025] 根据本发明实施例,提供了一种视频分类的方法实施例。

[0026] 可选地,在本实施例中,上述视频分类方法可以应用于如图2所示的由服务器102和终端104所构成的硬件环境中。如图2所示,服务器102通过网络与终端104进行连接,上述网络包括但不限于:广域网、城域网或局域网,终端104并不限于PC、手机、平板电脑等。本发明实施例的视频分类方法可以由服务器102来执行,也可以由终端104来执行,还可以是由服务器102和终端104共同执行。其中,终端104执行本发明实施例的视频分类方法也可以是由安装在其上的客户端来执行。

[0027] 图3是根据本发明实施例的一种可选的视频分类方法的流程图,如图3所示,该方法可以包括以下步骤:

[0028] 步骤S302,获取待分类的视频的内容类别,其中,内容类别为以视频自身内容分类得到的类别。

[0029] 步骤S304,根据预设的转换关系对视频的内容类别进行转换,得到视频的社交属性类别,其中,预设的转换关系通过样本数据训练得到,样本数据包括视频的内容类别和与视频有交互行为的用户的社交属性类别。

[0030] 步骤S306,将视频的社交属性类别作为待分类的视频的类别。

[0031] 通过上述步骤S302至步骤S306,采用根据预设的转换关系对视频的内容类别进行转换,得到视频的社交属性类别,根据视频的内容类别得到待分类的视频的社交属性类别的方式,达到了根据视频的社交属性对视频进行分类的目的,解决了由于视频分类时仅考虑视频自身内容造成的视频分类结果不贴近用户的技术问题,进而达到使视频分类结果更贴近用户的技术效果。

[0032] 在步骤S302提供的技术方案中,本发明实施例的视频分类方法可以由网络服务器执行,也可以由一些功能比较强大的终端执行,可以通过特定的系统执行本发明实施例的各个步骤。待分类的视频可以来自是社交网站的视频,也可以是来自一些视频分享网站的视频,这些视频分享网站的视频具有用户信息,例如,用户转发、评论、点赞等具有交互行为的用户信息。对于视频的内容类别获取的过程可以通过网络爬虫爬取的方式获取待分类的视频的内容类别,视频的内容类别可以是网站上已经为视频分配的内容类别,或者,视频的内容类别也可以是由预设的系统重新对视频按照内容进行分类得到的内容类别。

[0033] 视频的内容类别为以视频自身内容分类得到的类别,也即,如果视频内容是科幻类别,则视频的内容类别就为科幻类别,如果视频的内容是喜剧类别,则视频的内容类别就为喜剧类别。

[0034] 在步骤S304提供的技术方案中,预设的转换关系是在对具体的视频进行分类之前预先设置好的,预设的转换关系可以通过一定的方式进行样本数据训练得到,在样本训练时,样本数据包括视频的内容类别和与视频有交互行为的用户的社交属性类别,例如,对于某类视频来说,所有转发、评论、点赞的用户都具有社交属性,根据所有的用户的社交属性的类别得到最有代表性的社交属性类别,这个社交属性类别就可以代表这类视频的社交属性,在实际处理过程中,社交属性类别可以有多个具体的项,例如,某类视频的社交属性可以为“科技”和“人工智能”两个社交属性类别。

[0035] 在步骤S306提供的技术方案中,在根据预设的转换关系对视频的内容类别进行转换,得到视频的社交属性类别之后,可以将得到的视频的社交属性类别作为待分类的视频

的类别,可以不再使用视频原有的内容类别对视频进行分类,而是采用新的社交属性类别作为视频的类别,或者也可以保留原有的内容类别,将内容类别和社交属性类别都作为视频的类别。

[0036] 本发明实施例通过在传统的视频内容分类之外,嵌入用户兴趣和意愿等因素,使得视频的分类更加贴近用户。对视频内容分类来说,对同一段视频内容,由于用户的兴趣、知识背景和观看目的不同,用户对视频类别的认知存在差异,不同于传统方法仅从视频自身内容出发来识别视频类别,在本发明实施例中,用户的意愿、兴趣、关注主题等主观因素统称为用户社交属性,从用户和视频内容两个角度入手,研究融合用户社交属性的视频内容分类问题。首先抽取用户的社交属性特征,然后利用机器学习算法,实现用户社交属性的转移,并和视频内容属性融合,进而利用融合属性实现视频类型的识别和分类,为社交媒体时代的互联网视频内容分类和识别问题提供可供参考的解决方案。

[0037] 作为一种可选的实施例,在根据预设的转换关系对视频的内容类别进行转换之前,获取第一用户群中用户的社交属性类别,其中,第一用户群为所有与第一类视频具有交互行为的用户的集合;根据用户的社交属性类别为第一类视频添加社交属性类别;获取样本集合中所有视频的社交属性类别;通过字典学习算法得到社交属性类别和视频的内容类别之间转换关系,将转换关系作为预设的转换关系。

[0038] 样本集合中包括多个类别的视频,每个视频具有很多用户群,用户群是与这个视频具有交互行为的用户组成的用户群,交互行为可以是对视频进行转发、点赞、评论等行为,对于一个视频的用户群,如果用户群的社交属性集中度很高,说明这个视频与这个社交属性类别具有很高的相关度,可以用这个社交属性类别为这个视频进行社交属性分类。例如,某篮球明星的比赛视频的相关用户群的社交属性类别为“90后”、“球迷”,则说明具有这种社交属性类别的用户关注这类视频,可以对这类视频添加社交属性类别为“90后”、“球迷”,这样,对于该篮球明星的其他比赛视频也可以添加社交属性“90后”、“球迷”,可以将具有这类社交属性的视频为这类社交属性的用户推荐。这样的视频分类和推荐方法能够更加贴近用户。

[0039] 作为一种可选的实施例,用户的社交属性类别包括第一社交属性类别,获取第一用户群中用户的社交属性类别包括获取第一用户群中用户的第一社交属性类别,获取第一用户群中用户的第一社交属性类别包括:建立所有用户群的超图模型,其中,超图的点由一个用户的一个社交属性类别构成,超图的边由一个用户的所有社交属性类别构成;通过谱图聚类算法对所有用户群的超图模型进行聚类计算,得到所有用户群的第一社交属性类别;从所有用户群的第一社交属性类别中查找到第一用户群的第一社交属性类别。

[0040] 作为一种可选的实施例,用户的社交属性类别包括第二社交属性类别,获取第一用户群中用户的社交属性类别包括获取第一用户群中用户的第二社交属性类别,获取第一用户群中用户的第二社交属性类别包括:获取第一用户群在预设时间段内的社交信息;通过在线LDA模型抽取社交信息的主题内容,得到第一用户群在预设时间段内的主题内容信息;根据主题内容信息确定第一用户群中用户的第二社交属性类别。

[0041] 第一社交属性类别可以是静态社交属性类别,第二社交属性类别可以是动态社交属性类别,静态社交属性类别可以是对于用户来说适用时间尺度比较大的类别,例如,性别、兴趣、自我标签等,而动态社交属性类别可以是对用户来说时间尺度比较小的类别,例

如,用户近期在社交网站等浏览过的事件、发表的评论、转发的社会问题等用户短期内关注的时间,例如,对于一些近期播放的综艺节目关注的用户的动态社交属性类别可以是“某综艺节目”。

[0042] 对于静态社交属性的获取可以是建立所有用户群的超图模型,其中,超图的点由一个用户的一个社交属性类别构成,超图的边由一个用户的所有社交属性类别构成;通过谱图聚类算法对所有用户群的超图模型进行聚类计算,通过聚类计算可以得到所有用户群的第一社交属性类别,例如,某个用户群中有超过90%的用户的静态社交属性类别为足球迷,则这个用户群的静态社交属性类别为足球迷,通过聚类计算可以得到每个用户群的第一社交属性类别,从所有用户群的第一社交属性类别中查找到第一用户群的第一社交属性类别,

[0043] 对于动态社交属性类别,可以通过主题学习模型来获取,例如LDA主题学习模型,可以爬取某一时间区间内用户群体在某社交网站的信息,通过LDA主题学习模型抽取主题内容信息,然后作为这个用户群的动态社交属性类别。

[0044] 作为一种可选的实施例,获取待分类的视频的内容类别可以通过在ImageNet上预训练的VGG模型得到的全连接层的特征作为待分类的视频的内容类别,通过字典学习算法在社交属性类别和视频的内容类别之间建立预设的转换关系包括:通过函数

$$\arg \min_{B,s} \frac{1}{2} \|V - Bs\|_2^2 + \lambda_1 \|s\|_1 + \lambda_2 \phi(s)$$
 计算得到使视频的社交属性类别和内容类别误差最小

的转换关系,其中, V 表示训练集中的视频内容类别和社交属性类别连接的向量, B 和 s 分别代表一组基和 V 在基上的系数, $\phi(s)$ 代表训练集中不同向量的稀疏系数的距离函数,距离函数通过高斯核函数确定,将转换关系作为预设的转换关系。

[0045] 字典学习算法可以是LASSO模型,通过字典学习算法和优化模型,可以将属于不同的“源”的用户的社交属性类别和视频的内容类别之间建立对应关系,从而把社交属性和视频内容融合,实现两者的统一表达。

[0046] 本发明实施例从用户和视频内容两个角度入手,主要研究融合用户社交属性的视频内容分类问题。首先抽取用户的社交属性特征,然后利用机器学习算法,实现用户社交属性的转移,并和视频内容属性融合,进而利用融合属性实现视频类型的识别和分类,为社交媒体时代的互联网视频内容分类和识别问题提供可供参考的解决方案。

[0047] 本发明还提供了一种优选实施例,该优选实施例包括以下组成部分:

[0048] 本发明实施例与传统方法的主要区别和改进在于本发明是依据用户的社交属性(social attribution)进行视频分类,进而应用在视频检索或视频推荐等系统中。

[0049] 本发明实施例的技术方案以互联网中高精度、高效率的视频的类型的识别和分类性能为基本目标,从视频内容本身和网络用户两个角度,融合社交属性,研究关键方法,形成一个比较系统的结合用户社交属性的全段视频的分类和识别方案,主要可以解决如下两个问题:

[0050] ①如何建立用户社交属性抽取模型,获取用户社交属性特征。

[0051] ②阐明用户社交属性和视频内容的统一表达机制。

[0052] 通过本实施例提出的方案对互联网、特别是社交媒体(例如微博、朋友圈)中的视频进行分类,得到的视频类别标签和用户的兴趣、喜好、关注点等因素高度相关,从而在视

频推荐中可以提高用户体验,为用户提供更合适的视频内容。图4是根据本发明实施例的一种视频分类方法的流程图,如图4所示,首先,利用超图(超图和谱图聚类)和在线LDA(Online Latent Dirichlet Allocation)主题模型分别抽取用户静态和动态社交属性,进而根据用户和视频的关系(例如是否转发等)给视频打上社交属性标记;然后利用字典学习实现用户社交属性和视频自身内容特征的统一表达,最后根据融合特征借助SVM或深度学习网络等分类器实现视频内容分类。

[0053] 具体地,对于详细的步骤,执行方案如下:

[0054] (1) 社交属性抽取

[0055] 本发明实施例研究的用户社交属性主要包括用户兴趣和在社交媒体中讨论的主题等特征。具体抽取过程如下:

[0056] 用户关联关系的超图建模

[0057] 在社交媒体中,由于单一用户对视频的喜好或点击具有一定的偶然性,因此,本发明实施例以用户群体的形式表示用户的喜好,把具有相似性的用户聚合为一个群体,以群体的喜好和兴趣作为用户的兴趣。本发明实施例采用超图对用户群体进行建模,超图是一般图模型的扩展。以用户组成超图的顶点集合,若多个用户转发过相同的视频,或者为相同的视频进行“点赞”等操作,则形成一个超边,也可以理解为一个点代表一个属性,一个超边代表一个用户,所有超边构成超图。图5是根据本发明实施例的一种社交网络超图建模的示意图,如图5所示,图中每个点 $v_1 \sim v_8$ 为顶点, $e_1 \sim e_4$ 表示超边。

[0058] 图6是根据本发明实施例的一种用户关联关系的建立过程的流程图,如图6所示,通过超图建立用户之间的关联关系的步骤可以包括:

[0059] ①利用相关技术和工具从互联网爬取包含社交属性的视频数据和相关用户信息数据,例如视频的转发计数、点赞数、视频转发人的社交信息,包括年龄、兴趣、自我标签等。结合现有数据库构造研究数据集(研究数据集可以作为样本数据);

[0060] ②构造用户超图模型,顶点为用户,若用户对相同视频进行转发(超过一定的次数)或对相同视频有相似的操作,则构造超边;

[0061] ③对获得的超图进行顶点聚类,利用谱图(spectral clustering)聚类算法进行聚类。

[0062] 用户静态社交属性特征抽取

[0063] 对于每个用户来说,其个人的静态社交属性包括年龄、教育程度、兴趣爱好以及自我标记等信息。对用户进行超图建模后,采用超图聚类算法把用户分为不同的用户群体。在每个群体中,把群体内用户共有的一些静态属性作为该群体的静态社交属性,表示此用户群体具有近似一致的兴趣和爱好。对于某个视频来说,如果该视频被某一群体内的多数用户“点赞”或转发过,则此视频加上该群体的静态社交属性信息标记。具体步骤如下:

[0064] ①提取每一个类的用户共同信息,例如“90后”、“球迷(football fans)”等,作为当前类用户的静态属性标签;

[0065] ②对每个类中的用户来说,如果有超过一定的比例(设定经验阈值)转发或点赞过 m (m 为一设定值)次视频,则此视频加上该用户类的属性标签。

[0066] 用户动态社交属性特征抽取

[0067] 社交媒体在某个时间区间的内所感兴趣或讨论的主题,能有效的反响这个群体的

属性,从而影响用户群体对视频内容类别的判断,因此要建立一个主题学习模型,用来抽取社交群体的话题或主题特征。本发明实施例采用在线LDA(Online Latent Dirichlet Allocation)模型来实现。每个主题用一组相关词来表示,LDA模型通过一个生成模型建立社交网络中的文本语言(例如帖子、评论、朋友圈等)和一个或多个主题的对应关系,进而从社交网络中某一个时间段内的讨论内容中学习主题,体现了相关用户群体的关注点,可以称为用户的动态社交属性。若转发过某个视频的用户群体参与了该主题的讨论,则该视频被加上此用户动态社交属性标记,图7是根据本发明实施例的动态社交属性抽取过程的流程图,如图7所示,具体步骤如下:

[0068] ①爬取某一时间区间内用户群体在微博等社交媒体上的信息,利用LDA主题模型抽取主题内容信息,作为该用户群体的动态社交属性;

[0069] ②给视频添加动态社交属性标签。根据视频的静态属性标签确定视频所属的用户类别,然后为该视频添加相应用户群体的动态属性标签。

[0070] (2) 社交属性和视频内容特征的统一表达

[0071] 社交属性特征来自于用户,而视频内容特征来自于视频自身,因此两者具有不同的“源”,不能简单的对其进行融合。为充分把社交属性融入到视频内容,必须把两者投影到同一个空间,实现统一表达。本发明实施例可以采用字典学习的方法(LASSO模型)实现,图8是根据本发明实施例的社交属性与视频内容分类统一表达的流程图,如图8所示,将视频社交属性标记和深度学习模型(VGG)进行学习,可以通过字典学习算法和优化模型,在训练集中学习一个统一的字典。然后,把视频获取的社交属性标记特征和内容特征用这一个字典进行统一表达,从而把社交属性和视频内容融合到由一个字典张成的空间中,实现了两者的统一表达。具体的:

[0072] ①获取视频数据的内容特征。本发明实施例采用视频内容的深度特征作为视频内容特征,利用在ImageNet上预训练的模型VGG模型,采用全连接层的特征作为视频内容特征;

[0073] ②通过LASSO模型获得视频融合特征的稀疏表示形式,LASSO模型在应用的主要思路是在特征空间内,利用基底来重构每一个元素,实现稀疏表达,具体通过如下函数优化实现:

$$[0074] \quad \arg \min_{\mathbf{B}, \mathbf{s}} \frac{1}{2} \|\mathbf{v} - \mathbf{B}\mathbf{s}\|_2^2 + \lambda_1 \|\mathbf{s}\|_1 + \lambda_2 \phi(\mathbf{s})$$

[0075] 其中, \mathbf{v} 表示训练集中的视频内容特征和社交属性连接的向量, \mathbf{B} 和 \mathbf{s} 分别代表一组基和 \mathbf{v} 在此基上的系数, $\phi(\mathbf{s})$ 代表训练集中不同向量的稀疏系数的距离函数,用来保证相似度的传递性,本方案拟结合高斯核函数来设计距离函数 $\phi(\mathbf{s})$,整个优化问题可以通过交替优化的方式来解决。

[0076] (3) 视频内容分类

[0077] 获取了融合用户社交属性和视频内容特征的统一表示之后,视频内容的分类和内容识别就可以通过常规分类方法进行,本发明实施例可以采用传统的SVM分类器进行初步试验,然后考虑采用深度学习网络进行更进一步的分类试验,通过深度学习等工具,实现一个视频内容分类系统。

[0078] 现有技术的视频分类方法在视频分类中主要以“视频内容本身”为中心,以视频帧

像素或某种视觉特征为输入,利用深度学习网络实现视频语义的分类,其主要通过弥补“语义鸿沟(Semantic gap)”进行视频内容分类。与现有的视频分类方法相比,本发明实施例的技术方案改进主要体现在:在视频的分类中,加入用户群体的社交属性,即通过用户的主观兴趣和意愿进行视频内容分类。这种分类方法的优势在于,视频的类别标签不仅依赖于视频内容,还和观看过此视频的用户的信息有关,即在视频分类中体现了用户的需求,不仅解决了视频分类中的“语义鸿沟(Semantic gap)”,还在一定程度上解决了“需求鸿沟(Need gap)”(即给用户的是否是用户想要的)的问题,即解决了客观视频数据与主观用户需求之间的鸿沟,对于视频内容分类和视频个性化推荐都有重要的参考作用。

[0079] 需要说明的是,对于前述的各方法实施例,为了简单描述,故将其都表述为一系列的动作组合,但是本领域技术人员应该知悉,本发明并不受所描述的动作顺序的限制,因为依据本发明,某些步骤可以采用其他顺序或者同时进行。其次,本领域技术人员也应该知悉,说明书中所描述的实施例均属于优选实施例,所涉及的动作和模块并不一定是本发明所必须的。

[0080] 通过以上的实施方式的描述,本领域的技术人员可以清楚地了解到根据上述实施例的方法可借助软件加必需的通用硬件平台的方式来实现,当然也可以通过硬件,但很多情况下前者是更佳的实施方式。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质(如ROM/RAM、磁碟、光盘)中,包括若干指令用以使得一台终端设备(可以是手机,计算机,服务器,或者网络设备等)执行本发明各个实施例所述的方法。

[0081] 实施例2

[0082] 根据本发明实施例,还提供了一种用于实施上述视频分类方法的视频分类装置。图9是根据本发明实施例的一种可选的视频分类装置的示意图,如图9所示,该装置可以包括:

[0083] 第一获取单元10,用于获取待分类的视频的内容类别,其中,内容类别为以视频自身内容分类得到的类别。

[0084] 转换单元20,用于根据预设的转换关系对视频的内容类别进行转换,得到视频的社交属性类别,其中,预设的转换关系通过样本数据训练得到,样本数据包括视频的内容类别和与视频有交互行为的用户的社交属性类别。

[0085] 处理单元30,用于将视频的社交属性类别作为待分类的视频的类别。

[0086] 需要说明的是,该实施例中的第一获取单元10可以用于执行本申请实施例1中的步骤S302,该实施例中的转换单元20可以用于执行本申请实施例1中的步骤S304,该实施例中的处理单元30可以用于执行本申请实施例1中的步骤S306。

[0087] 此处需要说明的是,上述模块与对应的步骤所实现的示例和应用场景相同,但不限于上述实施例1所公开的内容。需要说明的是,上述模块作为装置的一部分可以运行在如图2所示的硬件环境中,可以通过软件实现,也可以通过硬件实现。

[0088] 通过上述模块,可以解决由于视频分类时仅考虑视频自身内容造成的视频分类结果不贴近用户的技术问题,进而达到使视频分类结果更贴近用户的技术效果。

[0089] 可选地,该装置还包括:第二获取单元,用于在根据预设的转换关系对视频的内容类别进行转换之前,获取第一用户群中用户的社交属性类别,其中,第一用户群为所有与第

一类视频具有交互行为的用户的集合；添加单元，用于根据用户的社交属性类别为第一类视频添加社交属性类别；第三获取单元，用于获取样本集合中所有视频的社交属性类别；计算单元，用于通过字典学习算法得到社交属性类别和视频的内容类别之间转换关系，将转换关系作为预设的转换关系。

[0090] 可选地，用户的社交属性类别包括第一社交属性类别，获取第一用户群中用户的社交属性类别包括获取第一用户群中用户的第一社交属性类别，第二获取单元包括：建立模块，用于建立所有用户群的超图模型，其中，超图的点由一个用户的一个社交属性类别构成，超图的边由一个用户的所有社交属性类别构成；计算模块，用于通过谱图聚类算法对所有用户群的超图模型进行聚类计算，得到所有用户群的第一社交属性类别；查找模块，用于从所有用户群的第一社交属性类别中查找到第一用户群的第一社交属性类别。

[0091] 可选地，用户的社交属性类别包括第二社交属性类别，获取第一用户群中用户的社交属性类别包括获取第一用户群中用户的第二社交属性类别，第二获取单元包括：获取模块，用于获取第一用户群在预设时间段内的社交信息；抽取模块，用于通过在线LDA模型抽取社交信息的主题内容，得到第一用户群在预设时间段内的主题内容信息；确定模块，用于根据主题内容信息确定第一用户群中用户的第二社交属性类别。

[0092] 可选地，第一获取单元10用于通过在ImageNet上预训练的VGG模型得到的全连接层的特征作为待分类的视频的内容类别，计算单元用于：通过函数

$$\arg \min_{B,s} \frac{1}{2} \|V - Bs\|_2^2 + \lambda_1 \|s\|_1 + \lambda_2 \phi(s)$$
 计算得到使视频的社交属性类别和内容类别误差最

小的转换关系，其中，V表示训练集中的视频内容类别和社交属性类别连接的向量，B和s分别代表一组基和V在基上的系数， $\phi(s)$ 代表训练集中不同向量的稀疏系数的距离函数，距离函数通过高斯核函数确定，将转换关系作为预设的转换关系。

[0093] 此处需要说明的是，上述模块与对应的步骤所实现的示例和应用场景相同，但不限于上述实施例1所公开的内容。需要说明的是，上述模块作为装置的一部分可以运行在如图2所示的硬件环境中，可以通过软件实现，也可以通过硬件实现，其中，硬件环境包括网络环境。

[0094] 实施例3

[0095] 根据本发明实施例，还提供了一种用于实施上述视频分类方法的服务器或终端。

[0096] 图10是根据本发明实施例的一种终端的结构框图，如图10所示，该终端可以包括：一个或多个（图中仅示出一个）处理器201、存储器203、以及传输装置205（如上述实施例中的发送装置），如图10所示，该终端还可以包括输入输出设备207。

[0097] 其中，存储器203可用于存储软件程序以及模块，如本发明实施例中的视频分类方法和装置对应的程序指令/模块，处理器201通过运行存储在存储器203内的软件程序以及模块，从而执行各种功能应用以及数据处理，即实现上述的视频分类方法。存储器203可包括高速随机存储器，还可以包括非易失性存储器，如一个或者多个磁性存储装置、闪存、或者其他非易失性固态存储器。在一些实例中，存储器203可进一步包括相对于处理器201远程设置的存储器，这些远程存储器可以通过网络连接至终端。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

[0098] 上述的传输装置205用于经由一个网络接收或者发送数据，还可以用于处理器与

存储器之间的数据传输。上述的网络具体实例可包括有线网络及无线网络。在一个实例中,传输装置205包括一个网络适配器(Network Interface Controller,NIC),其可通过网线与其他网络设备与路由器相连从而可与互联网或局域网进行通讯。在一个实例中,传输装置205为射频(Radio Frequency,RF)模块,其用于通过无线方式与互联网进行通讯。

[0099] 其中,具体地,存储器203用于存储应用程序。

[0100] 处理器201可以通过传输装置205调用存储器203存储的应用程序,以执行下述步骤:获取待分类的视频的内容类别,其中,内容类别为以视频自身内容分类得到的类别;根据预设的转换关系对视频的内容类别进行转换,得到视频的社交属性类别,其中,预设的转换关系通过样本数据训练得到,样本数据包括视频的内容类别和与视频有交互行为的用户的社交属性类别;将视频的社交属性类别作为待分类的视频的类别。

[0101] 处理器201还用于执行下述步骤:在根据预设的转换关系对视频的内容类别进行转换之前,获取第一用户群中用户的社交属性类别,其中,第一用户群为所有与第一类视频具有交互行为的用户的集合;根据用户的社交属性类别为第一类视频添加社交属性类别;获取样本集合中所有视频的社交属性类别;通过字典学习算法得到社交属性类别和视频的内容类别之间转换关系,将转换关系作为预设的转换关系。

[0102] 处理器201还用于执行下述步骤:建立所有用户群的超图模型,其中,超图的点由一个用户的一个社交属性类别构成,超图的边由一个用户的所有社交属性类别构成;通过谱图聚类算法对所有用户群的超图模型进行聚类计算,得到所有用户群的第一社交属性类别;从所有用户群的第一社交属性类别中查找到第一用户群的第一社交属性类别。

[0103] 处理器201还用于执行下述步骤:获取第一用户群在预设时间段内的社交信息;通过在线LDA模型抽取社交信息的主题内容,得到第一用户群在预设时间段内的主题内容信息;根据主题内容信息确定第一用户群中用户的第二社交属性类别。

[0104] 处理器201还用于执行下述步骤:通过在ImageNet上预训练的VGG模型得到的全连接层的特征作为待分类的视频的内容类别,通过字典学习算法在社交属性类别和视频的内

容类别之间建立预设的转换关系包括:通过函数 $\arg \min_{B,s} \frac{1}{2} \|V - Bs\|_2^2 + \lambda_1 \|s\|_1 + \lambda_2 \phi(s)$ 计算

得到使视频的社交属性类别和内容类别误差最小的转换关系,其中,V表示训练集中的视频内容类别和社交属性类别连接的向量,B和s分别代表一组基和V在基上的系数, $\phi(s)$ 代表训练集中不同向量的稀疏系数的距离函数,距离函数通过高斯核函数确定,将转换关系作为预设的转换关系。

[0105] 采用本发明实施例,采用根据预设的转换关系对视频的内容类别进行转换,得到视频的社交属性类别,根据视频的内容类别得到待分类的视频的社交属性类别的方式,达到了根据视频的社交属性对视频进行分类的目的,从而实现了使视频分类结果更贴近用户的技术效果,进而解决了由于视频分类时仅考虑视频自身内容造成的视频分类结果不贴近用户的技术问题。

[0106] 可选地,本实施例中的具体示例可以参考上述实施例1和实施例2中所描述的示例,本实施例在此不再赘述。

[0107] 本领域普通技术人员可以理解,图10所示的结构仅为示意,终端可以是智能手机(如Android手机、iOS手机等)、平板电脑、掌上电脑以及移动互联网设备(Mobile Internet

Devices, MID)、PAD等终端设备。图10其并不对上述电子装置的结构造成限定。例如,终端还可包括比图10中所示更多或者更少的组件(如网络接口、显示装置等),或者具有与图10所示不同的配置。

[0108] 本领域普通技术人员可以理解上述实施例的各种方法中的全部或部分步骤是可以通程序来指令终端设备相关的硬件来完成,该程序可以存储于一计算机可读存储介质中,存储介质可以包括:闪存盘、只读存储器(Read-Only Memory, ROM)、随机存取器(Random Access Memory, RAM)、磁盘或光盘等。

[0109] 实施例4

[0110] 本发明的实施例还提供了一种存储介质。可选地,在本实施例中,上述存储介质可以用于执行视频分类方法的程序代码。

[0111] 可选地,在本实施例中,上述存储介质可以位于上述实施例所示的网络中的多个网络设备中的至少一个网络设备上。

[0112] 可选地,在本实施例中,存储介质被设置为存储用于执行以下步骤的程序代码:

[0113] S1,获取待分类的视频的内容类别,其中,内容类别为以视频自身内容分类得到的类别;

[0114] S2,根据预设的转换关系对视频的内容类别进行转换,得到视频的社交属性类别,其中,预设的转换关系通过样本数据训练得到,样本数据包括视频的内容类别和与视频有交互行为的用户的社交属性类别;

[0115] S3,将视频的社交属性类别作为待分类的视频的类别。

[0116] 可选地,存储介质还被设置为存储用于执行以下步骤的程序代码:在根据预设的转换关系对视频的内容类别进行转换之前,获取第一用户群中用户的社交属性类别,其中,第一用户群为所有与第一类视频具有交互行为的用户的集合;根据用户的社交属性类别为第一类视频添加社交属性类别;获取样本集合中所有视频的社交属性类别;通过字典学习算法得到社交属性类别和视频的内容类别之间转换关系,将转换关系作为预设的转换关系。

[0117] 可选地,存储介质还被设置为存储用于执行以下步骤的程序代码:建立所有用户群的超图模型,其中,超图的点由一个用户的一个社交属性类别构成,超图的边由一个用户的所有社交属性类别构成;通过谱图聚类算法对所有用户群的超图模型进行聚类计算,得到所有用户群的第一社交属性类别;从所有用户群的第一社交属性类别中查找到第一用户群的第一社交属性类别。

[0118] 可选地,存储介质还被设置为存储用于执行以下步骤的程序代码:获取第一用户群在预设时间段内的社交信息;通过在线LDA模型抽取社交信息的主题内容,得到第一用户群在预设时间段内的主题内容信息;根据主题内容信息确定第一用户群中用户的第二社交属性类别。

[0119] 可选地,存储介质还被设置为存储用于执行以下步骤的程序代码:通过在ImageNet上预训练的VGG模型得到的全连接层的特征作为待分类的视频的内容类别,通过字典学习算法在社交属性类别和视频的内容类别之间建立预设的转换关系包括:通过函数

$\arg \min_{B,s} \frac{1}{2} \|V - Bs\|_2^2 + \lambda_1 \|s\|_1 + \lambda_2 \phi(s)$ 计算得到使视频的社交属性类别和内容类别误差最小

的转换关系,其中, V 表示训练集中的视频内容类别和社交属性类别连接的向量, B 和 s 分别代表一组基和 V 在基上的系数, $\phi(s)$ 代表训练集中不同向量的稀疏系数的距离函数,距离函数通过高斯核函数确定,将转换关系作为预设的转换关系。

[0120] 可选地,本实施例中的具体示例可以参考上述实施例1和实施例2中所描述的示例,本实施例在此不再赘述。

[0121] 可选地,在本实施例中,上述存储介质可以包括但不限于:U盘、只读存储器(ROM, Read-Only Memory)、随机存取存储器(RAM, Random Access Memory)、移动硬盘、磁碟或者光盘等各种可以存储程序代码的介质。

[0122] 上述本发明实施例序号仅仅为了描述,不代表实施例的优劣。

[0123] 上述实施例中的集成的单元如果以软件功能单元的形式实现并作为独立的产品销售或使用,可以存储在上述计算机可读的存储介质中。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的全部或部分可以以软件产品的形式体现出来,该计算机软件产品存储在存储介质中,包括若干指令用以使得一台或多台计算机设备(可为个人计算机、服务器或者网络设备)执行本发明各个实施例所述方法的全部或部分步骤。

[0124] 在本发明的上述实施例中,对各个实施例的描述都各有侧重,某个实施例中未详述的部分,可以参见其他实施例的相关描述。

[0125] 在本申请所提供的几个实施例中,应该理解到,所揭露的客户端,可通过其它的方式实现。其中,以上所描述的装置实施例仅仅是示意性的,例如所述单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,例如多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口,单元或模块的间接耦合或通信连接,可以是电性或其它的形式。

[0126] 所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0127] 另外,在本发明各个实施例中的各功能单元可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现,也可以采用软件功能单元的形式实现。

[0128] 以上所述仅是本发明的优选实施方式,应当指出,对于本技术领域的普通技术人员来说,在不脱离本发明原理的前提下,还可以做出若干改进和润饰,这些改进和润饰也应视为本发明的保护范围。



图1

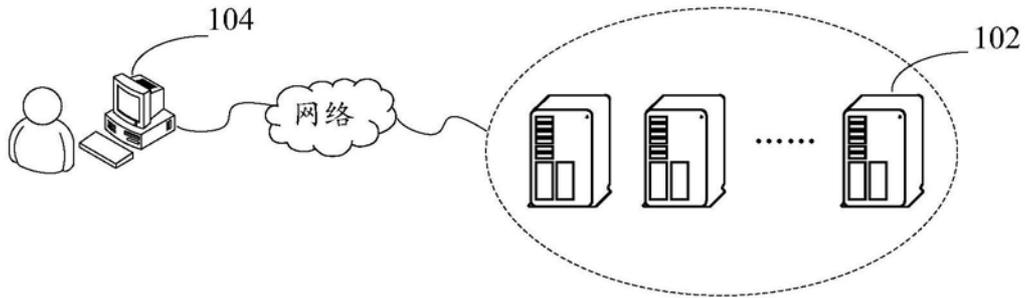


图2

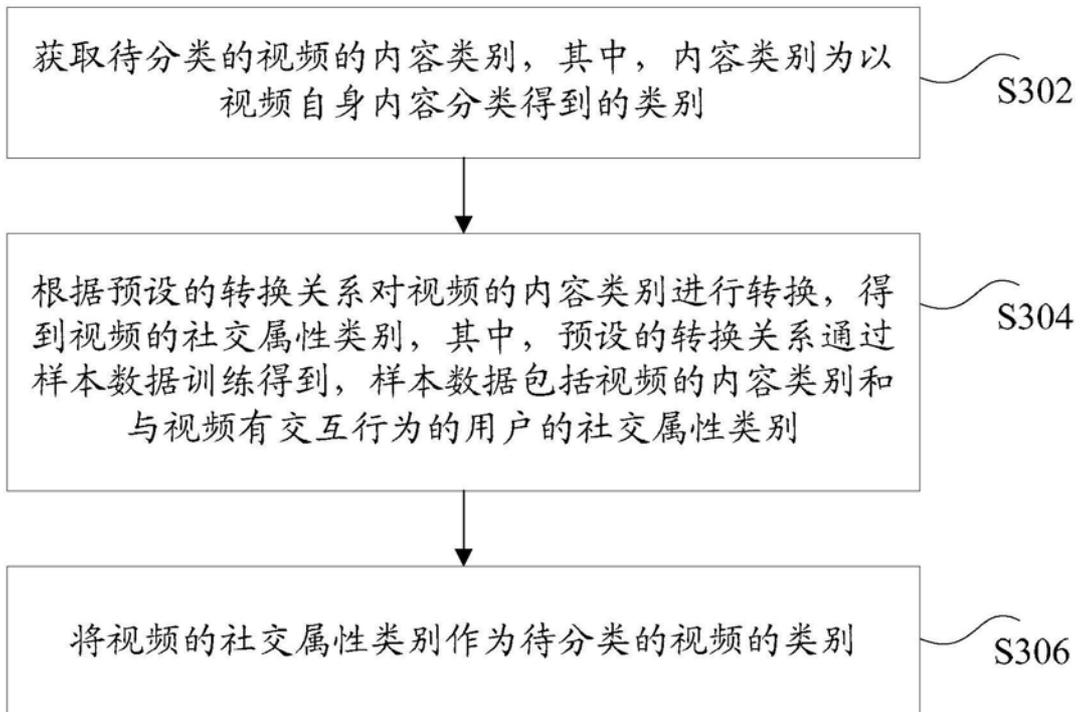


图3

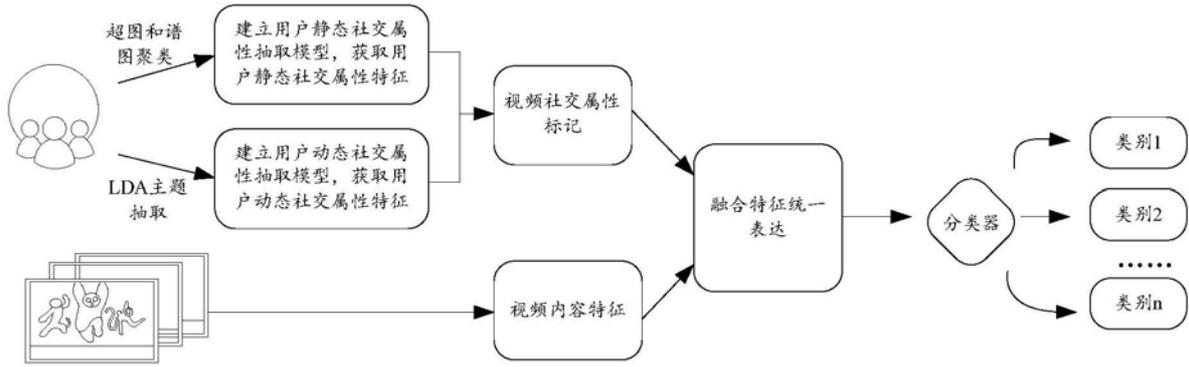


图4

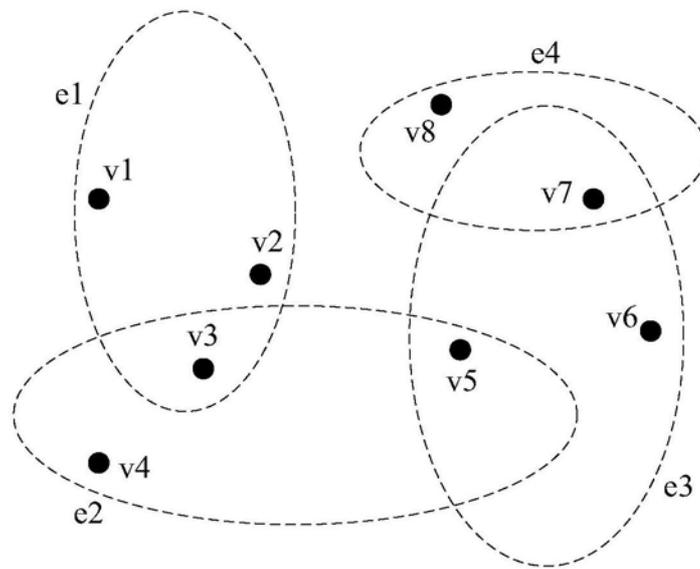


图5

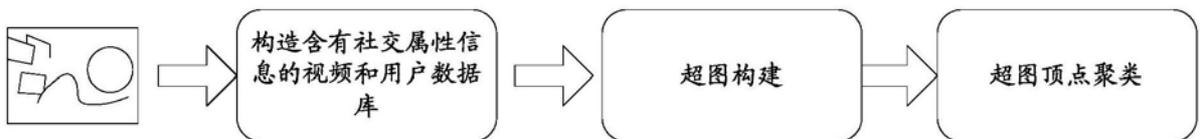


图6



图7

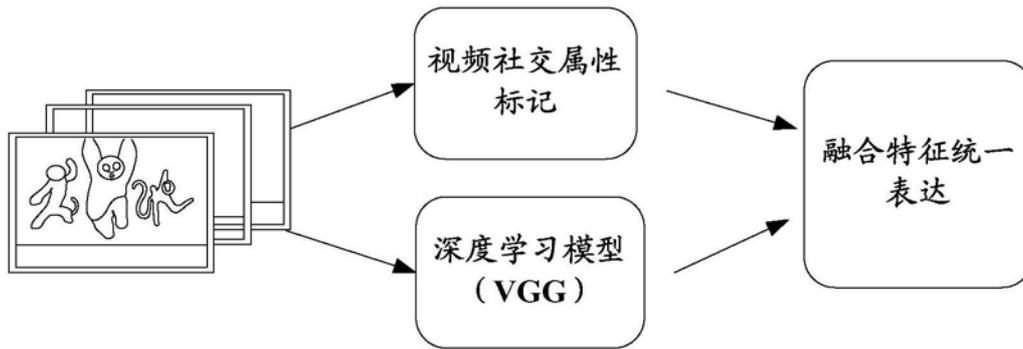


图8



图9

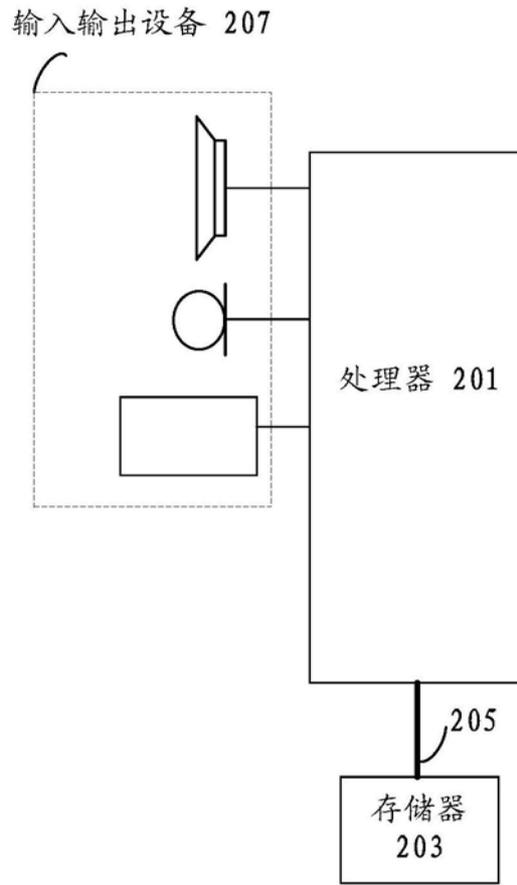


图10