

# 视频摘要技术综述

李 双 杨欣怡

(国家知识产权局专利局专利审查协作四川中心 四川 成都 610000)

**摘 要:** 视频摘要的主要目的是可以让用户快速、方便的了解整个视频的内容,通过它可以快捷地在庞大的数据库中找到用户所需要的视频。文章利用专利数据库对视频摘要技术进行了数据分析,对该领域的专利发展趋势进行分析,通过对视频摘要技术各个时期的专利文献进行梳理和分析,以不同阶段的视频摘要相关专利为例统计出该领域的技术演进路线。

**关键词:** 视频;摘要;概要;关键帧;专利;发展状况

中图分类号:T-18

文献标志码:A

文章编号:2095-2945(2018)25-0015-02

**Abstract:** The main purpose of video summary is to enable users to quickly and conveniently understand the content of the entire video, through which we can quickly find the video users need in the huge database. This paper makes use of patent database to analyze the data of video abstract technology, analyzes the patent development trend in this field, and sorts out and analyzes the patent literature of each period of video abstract technology. Taking the patents related to video abstracts in different stages as an example, the evolution route of technology in this field is analyzed.

**Keywords:** video; abstract; summary; keyframe; patent; development status

## 1 概述

在现代社会中,随着网络技术的飞速发展和多媒体技术的广泛应用,数字视频大量涌现,例如新闻、广告、监控视频、家庭视频、网络电影等数字视频的快速增加引发了许多新的技术,包括视频存档、编目、索引以及存取等<sup>[1][2]</sup>。视频摘要主要目的是可以让用户快速、方便的了解整个视频的内容,以决定是否详细观看这段视频,同时,还可以用于视频数据库的索引、查询等,通过它可以快捷地在庞大的数据库中找到用户所需要的视频<sup>[3][4]</sup>。

## 2 技术演进路线

通过对视频摘要技术各个时期的专利文献进行梳理和分析,可以得到该领域的专利技术演进路线如图1所示。

### 2.1 第一阶段(1987年-2000年)萌芽期

1989年-2000年是视频摘要技术的萌芽期,在这一阶段,文本类型的视频摘要主要是基于文本识别方法,图片类型、视频概要、以及特殊视频片段集合主要基于视觉信息生成方法生成的,同时也出现了一些关于简单的时间压缩或语音信号识别方法。

文字类型的视频摘要典型的代表专利,例如飞利浦公司于2001年申请的专利WO0241634 A2公开了一种基于副本信息提取节目摘要的方法,根据分类的节目类型在副本信息中确定出第一待识别提示语,然后根据在副本信息中识别的提示语,存储一部分副本信息,根据存储的这部分副本信息产生节目的摘要。

这一阶段主要研究方向是静态的图片类型的视频摘要,基于颜色和运动的关键帧提取进入了研究者的视线,

例如夏普1997年申请的专利US19970994558A,一种视频摘要生成方法,通过统计输入视频帧序列关键帧的直方图,基于得到的关键帧直方图检测到镜头边界,将镜头边界的视频帧作为视频摘要图片;关于视频概要以及特殊视频片段方面的研究刚刚起步,苹果公司于1994年申请的专利US5805733 A,一种检测视频场景和视频帧序列的方法,通过比较各个场景之间的颜色直方图、运动向量偏差等相似性检测场景的改变,对场景进行分割,然后依据它们的持续时间来选择重要的场景组合作为视频摘要,还有一些基于用户交互的特殊视频片段集合的应用,例如索尼公司1996年申请的专利HU224575 B1,提供了用户管理记录在记录介质的节目的激动人心或特征部分(作为摘要)的记录位置的管理区,所以可获得记录介质上的节目的概要。基于简单的时间压缩和仅采用语音信号生成视频摘要的方法代表有专利有日本松下公司1996年申请的专利JP1998032809A,一种基于简单时间压缩的视频摘要审查方法,通过以双倍的速度提取关键帧生成摘要。

### 2.2 第二阶段(2000年-2009年)快速发展期

2000年-2010年期间,基于图片的视频摘要发展迅速。涌现了一些更复杂的基于镜头聚类的关键帧提取方法。例如北京新岸线网络技术有限公司于2006年申请的专利CN101021904 A,一种视频内容分析的方法,设某个镜头包含的多个图像帧,预定义阈值控制聚类密度,从每个聚类中抽取离聚类质心最近的作为这个聚类的代表帧,所有聚类的代表帧就构成了镜头的关键帧,然后结合人脸识别对视频关键帧进行分析。

同期,基于视频概要和特殊视频片段集合也开始得到

作者简介:李双(1990.06-),性别:女,民族:汉,籍贯:黑龙江青冈县,学历:硕士研究生,职务:审查员,研究方向:图像通信;杨欣怡(1990.07-),性别:女,民族:汉,籍贯:江苏南京,学历:硕士研究生,职务:审查员,研究方向:图像通信。

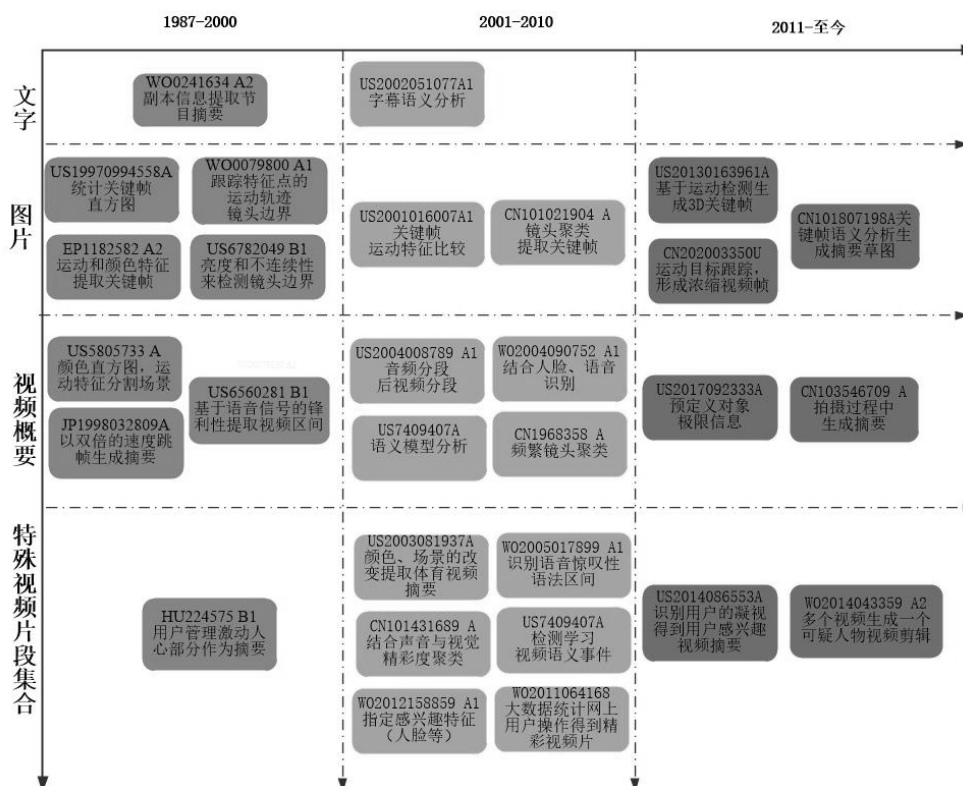


图 1 技术演进路线

研发重视,快速发展起来,基于视觉的视频概要生成也有了进一步的发展,例如夏普 2002 年申请的专利 US2003081937A1,一种体育视频摘要提取方法,基于颜色特征和场景的改变通过马尔科夫定理或贝叶斯模型识别体育事件,提取出体育事件区间组合为视频摘要。随着人脸识别以及文字、语音识别技术的发展,出现了融合了人脸识别技术、文字识别技术、语音识别技术等综合特征的视频概要以及特殊视频片段摘要生成方法。华为公司于 2007 年申请的专利 CN101431689 A,公开了一种生成视频摘要的方法及装置,通过对视频文件的视觉精彩度和声音精彩度进行处理,然后将满足视觉精彩度和声音精彩度的关联镜头组聚合成视频片段得到视频摘要。与此同时出现了一些基于视频语义的生成方法的视频摘要技术,具有代表性的专利有三菱公司 2004 年申请的专利 US7409407A,通过定义相似矩阵,聚类模型学习视频语义结构关系从而生成视频摘要;在用户交互方面,出现了通过统计用户行为,生成用户感兴趣的视频摘要生成方法,例如国际商业机器公司于 2009 年申请的专利 WO2011064168 A1,公开了一种用于提供网络视频访问热度的方法和装置,通过大数据统计网络用户在播放视频时的快进、回放、前拖、后拖等操作来生成热点视频摘要。

2.3 第三阶段(2010 年至今):多元化发展期,中国崛起

2011 年至今,视频摘要技术进入多元化发展时期。在图片类型的摘要方面出现了全景关键帧,例如三菱公司于 2011 年申请的专利 US2013163961 A1,通过多个关键帧中运动目标的识别与跟踪,把各个目标放入一个 3D 场景中,

形成 3D 场景的视频摘要。在视频概要方面,例如韩国电信研究所 2013 年申请的专利 US2014086553 A1,通过识别在观看视频时用户眼睛的凝视,来生成用户感兴趣的视频摘要;国际商业机器公司于 2015 年申请的专利 US2017092333A,通过视频语法结构识别预定义对象在标记帧内可识别信息的极限信息来确定关键帧。

此时,中国的视频摘要技术进入了快速发展时期,具有代表性的中国专利有,广东鑫程电子科技有限公司于 2010 年申请的专利 CN202003350U,公开了一种视频摘要系统,将视频中的运动目标提取为运动块,进行处理后,与前一帧中相对应的运动块进行匹配,实现目标的跟踪,并将运动块在不同时间的位置同时显示于同一帧画面中,这样就可以将一段时间之内的视频内容浓缩地显示于同一帧中,当这些浓缩后的视频帧连贯起来,便可以形成整段视频的摘要。

### 3 结束语

本文通过核心专利的解析详细梳理了视频摘要技术的发展情况,有助于全面了解视频摘要技术的基本发展态势,对其专利审查工作有较大帮助。

### 参考文献:

- [1]苏翠宁.视频摘要技术与系统实现[D].中南大学,2009.
- [2]余卫宇,曹燕,谢胜利.视频摘要的现状和研究进展[J].计算机应用研究,2008,25(7):1948-1952.
- [3]谢月飞.基于摘要的视频内容检索系统[D].桂林电子科技大学,2013.
- [4]陈添丁.基于内容特征的图像检索和整合性视频检索[D].浙江大学,2003.