

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.  
G06F 17/30 (2006.01)



# [12] 发明专利申请公布说明书

[21] 申请号 200810304677.9

[43] 公开日 2009 年 2 月 18 日

[11] 公开号 CN 101369287A

[22] 申请日 2008.9.25

[21] 申请号 200810304677.9

[71] 申请人 四川长虹电器股份有限公司

地址 621000 四川省绵阳市高新区绵兴东路  
35 号

[72] 发明人 刘 东 唐 军

[74] 专利代理机构 成都虹桥专利事务所  
代理人 李顺德

权利要求书 1 页 说明书 2 页

## [54] 发明名称

音视频数据挖掘的实现方法

## [57] 摘要

本发明涉及对数字电视中的音视频数据处理技术领域，尤其涉及数字电视的音视频数据挖掘的实现方法。本发明提出了一种音视频数据挖掘的实现方法，让用户很容易搜集到自己喜欢的作品。其技术方案的要点是：搜集音视频作品，并搜集每一个音视频作品的相关信息数据，再对这些信息数据进行充分的索引和挖掘，完成分类、估值、相关性分组、聚集等预处理；完成预处理后对数据进行格式转换及相关信息的提取，并存储于前端媒体服务器中，用户在客户终端输入搜索信息，通过广电网络回传于前端媒体服务器，前端媒体服务器根据该搜索信息对音视频作品进行排查并反馈给客户终端。本发明的方法能使用户很容易就能搜索到自己喜欢的作品。

【权利要求1】音视频数据挖掘的实现方法，其特征在于：包括以下步骤：

- a. 搜集音视频作品，将其制作为可供在数字电视网络中传送的数据格式，存储于前端媒体服务器中；
- b. 针对每一个音视频作品，搜集其相关信息数据；
- c. 对步骤b中搜索到的信息数据进行充分挖掘，完成分类、估值、相关性分组的预处理；
- d. 将所有经过预处理的信息数据根据其主关键字排序、索引建立媒体数据挖掘器所能识别的数据库，完成将信息数据转换为计算机程序所能识别的数据；
- e. 在客户终端输入搜索信息，通过广电网络回传到前端媒体服务器，前端媒体服务器根据该搜索信息对音视频作品进行排查，确定用户需要查找的具体资源，并反馈给客户终端。

【权利要求2】如权利要求1所述的音视频数据挖掘的实现方法，其特征在于：所述步骤b中，相关信息数据包括剧情、台词、歌词、导演名、演员名、视频片段。

【权利要求3】如权利要求1所述的音视频数据挖掘的实现方法，其特征在于：所述步骤e中，搜索信息为关键字。

【权利要求4】如权利要求1所述的音视频数据挖掘的实现方法，其特征在于：所述步骤e中，搜索信息为一段音视频。

## 音视频数据挖掘的实现方法

### 技术领域

本发明涉及对数字电视中的音视频数据处理技术领域，尤其涉及数字电视的音视频数据挖掘的实现方法。

### 背景技术

数据挖掘，又称为数据库中的知识发现，就是从大量数据中获取有效的、需要的知识。目前数字电视的发展越来越快，信息量越来越大，各种原来基于IP网络的应用逐渐转向广电单向网络。对于观众来说，有可能只记得某些电影或音乐的一点点细节，即引起自己注意力的一个微笑的画面、一段文字、一句台词、导演名字、演唱者、演员名字等等。要想把整部作品描述出来是相当困难的，可能除了亲身参与过该作品的人以外，观众一般是无法完整复述作品，从而也就会在想起来的时候没办法再次搜集到自己喜欢的作品。

### 发明内容

本发明所要解决的技术问题是：提出一种音视频数据挖掘的实现方法，让用户很容易搜集到自己喜欢的作品。

本发明解决上述技术问题采用的技术方案是：音视频数据挖掘的实现方法，包括以下步骤：

- a. 搜集音视频作品，将其制作为可供在数字电视网络中传送的数据格式，存储于前端媒体服务器中；
- b. 针对每一个音视频作品，搜集其相关信息数据；
- c. 对步骤b中搜索到的信息数据进行充分挖掘，完成分类、估值、相关性分组的预处理；
- d. 将所有经过预处理的信息数据根据其主关键字排序、索引建立媒体数据挖掘器所能识别的数据库，完成将信息数据转换为计算机程序所能识别的数据；
- e. 在客户终端输入搜索信息，通过广电网络回传到前端媒体服务器，前端媒体服务器根据该搜索信息对音视频作品进行排查，确定用户需要查找的具体资源，并反馈给客户端。

本发明的有益效果是：使用户很容易就能搜索到自己喜欢的作品。

### 具体实施方式

本发明中的音视频数据挖掘实现方法可以帮助用户较容易地找回自己想收看或收听的作品。

品，我们将音视频的数据挖掘应用于数字电视中，提供给用户各种个样的信息，用户可以通过自己曾记得的关于此作品的某个关键字或是通过其它资源途径找到一小段音视频，然后将其输入客户终端，通过一定的回传通道，将此信息反馈回前端媒体服务器，前端媒体服务器的应用系统根据此关键字在数据库中逐一排查，迅速找出用户希望找到的作品信息，并作关联，在广电网络中通过CAS (Conditional Access System) 或DRM(digital right management)接口，向其发送用户授权管理信息的基本数据，据此实现对用户收看电视节目的控制，从而实现用户的视频点播。

在具体实施上，本发明可采取以下步骤：a. 媒体资产库建立：搜集音视频作品，将其制作为可供在数字电视网络中传送的数据格式，如TS流或流媒体，然后存于前端媒体服务器中；

b. 数据信息搜集：针对每一个音视频作品，尽可能地搜集其相关信息，如剧情、台词、歌词、导演名、演员名、视频片段等各种各样的信息数据，并建立相应的数据库文件；

c. 数据挖掘：针对步骤b中所搜集到的各类信息数据，与相应音视频作品进行关联，对这些信息数据进行充分的索引和挖掘，完成分类、估值、相关性分组、聚集等预处理；

d. 数据处理：对所有经过预处理的信息数据根据其主关键字排序、索引建立媒体数据挖掘器所能识别的数据库，完成将信息数据转换为各计算机程序所能识别的数据；将上述数据进行相关信息提取，按照既定的协议数据格式，经由数据库或数据仓库的接口，最终将经过数据挖掘过程的数据全部存入硬件介质中；

e. 回传信息接收和处理：用户通过在客户终端输入关键字或者一段音视频，通过广电网络回传到前端媒体服务器，前端媒体服务器的应用系统根据此关键字或一段音视频对音视频作品进行排查，确定用户需要查找的具体资源，并通过广电网络反馈给客户终端；

f. 用户确认查找信息：查找结果反馈到客户终端后，终端接收设备提供给用户选择点播等业务的操作方法，如通过OSD (On Screen Display) 和RC (Remote Control) 的显示和控制确认点播业务等；

g. 前端媒体服务器实现授权：用户确认点播的信息回传到前端媒体服务器的应用系统后，在广电网络中通过CAS或DRM接口，向客户终端发送用户授权管理信息的基本数据，据此实现对用户收看电视节目的控制，从而实现用户的视频点播。