



(12) 发明专利

(10) 授权公告号 CN 102346768 B

(45) 授权公告日 2013. 01. 16

(21) 申请号 201110279361. 0

(22) 申请日 2011. 09. 20

(73) 专利权人 北京鹏润鸿途科技有限公司

地址 100085 北京市海淀区学清路 38 号金
码大厦 B 座顶 1-2322

(72) 发明人 何长龙 刘向宇 李勋

(74) 专利代理机构 深圳市威世博知识产权代理
事务所 (普通合伙) 44280

代理人 何青瓦 李庆波

(51) Int. Cl.

G06F 17/30 (2006. 01)

(56) 对比文件

CN 101540907 A, 2009. 09. 23, 权利要求 1.

US 2005/0149968 A1, 2005. 07. 07, 全文 .

CN 1543096 A, 2004. 11. 03, 全文 .

审查员 胡平

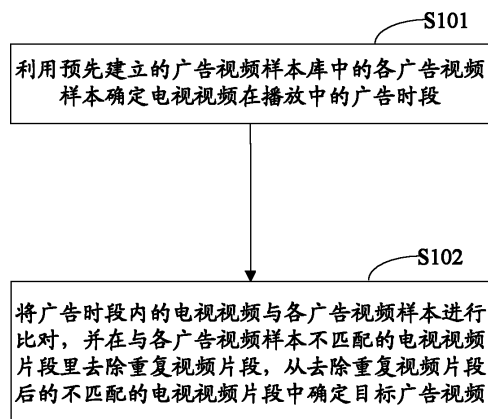
权利要求书 2 页 说明书 8 页 附图 5 页

(54) 发明名称

一种发现广告视频的方法及装置

(57) 摘要

本发明提供了一种发现广告视频的方法及装置, 其中发现广告的方法包括 :A. 利用预先建立的广告视频样本库中的各广告视频样本确定电视视频中的广告时段 ;B. 将所述广告时段内的电视视频与各广告视频样本进行比对, 并在与各广告视频样本不匹配的电视视频片段里去除重复视频片段, 从去除重复视频片段后的不匹配的电视视频片段中确定目标广告视频。通过上述方式, 可以提高广告监管的效率。



1. 一种发现广告的方法,其特征在于,所述方法包括:

A. 利用预先建立的广告视频样本库中的各广告视频样本确定电视视频中的广告时段;

B. 将所述广告时段内的电视视频与各广告视频样本进行比对,并在与各广告视频样本不匹配的电视视频片段里去除重复视频片段,从去除重复视频片段后的不匹配的电视视频片段中确定目标广告视频;

其中所述步骤 A 包括:

在电视视频中采集电视视频样本;

将所述电视视频样本与各广告视频样本逐一比对,将所述电视视频样本中与各广告视频样本比对上的视频片段确定为各匹配视频片段;

将各匹配视频片段按照在所述电视视频样本中的时间顺序排列,从中确定连续视频片段集与独立视频片段,其中所述连续视频片段集为相邻两个匹配视频片段之间的时间间隔不超过预定值的匹配视频片段集合,所述独立视频片段为所有匹配视频片段中除所述连续视频片段集之外的匹配视频片段;

将所述连续视频片段集在所述电视视频样本中开始时刻之前设定时长的第一冗余时刻与结束时刻之后设定时长的第二冗余时刻之间的时间区间,以及所述独立视频片段在所述电视视频样本中开始时刻之前设定时长的第三冗余时刻与结束时刻之后设定时长的第四冗余时刻之间的时间区间确定为所述广告时段。

2. 根据权利要求 1 所述的方法,其特征在于,所述方法在步骤 B 后进一步包括:将确定的目标广告视频放入所述广告视频样本库。

3. 根据权利要求 1 所述的方法,其特征在于,在所述步骤 B 之后还包括:将与各广告视频样本匹配的电视视频片段和所述目标广告视频作为所述电视视频包含的广告。

4. 根据权利要求 1 至 3 中任一权项所述的方法,其特征在于,从去除重复视频片段后的不匹配的电视视频片段中确定目标广告视频包括:

利用对所述去除重复视频片段后的不匹配的电视视频片段的人工审核结果,确定目标广告视频;或者,

通过大量广告视频数据训练得到的分类器,对所述去除重复视频片段后的不匹配的电视视频片段进行分类判断,确定目标广告视频。

5. 一种发现广告的装置,其特征在于,所述装置包括:

广告时段确定单元,用于利用预先建立的广告视频样本库中的各广告视频样本确定电视视频中的广告时段;

去重单元,用于将所述广告时段内的电视视频与各广告视频样本进行比对,并在与各广告视频样本不匹配的电视视频片段里去除重复视频片段;

目标广告确定单元,用于从去除重复视频片段后的不匹配的电视视频片段中确定目标广告视频;

其中所述广告时段确定单元包括:

第一采集子单元,用于在电视视频中采集电视视频样本;

第一比对子单元,用于将所述电视视频样本与各广告视频样本逐一比对,将所述电视视频样本中与各广告视频样本比对上的视频片段确定为各匹配视频片段;

匹配集确定子单元,用于将各匹配视频片段按照在所述电视视频样本中的时间顺序排列,从中确定连续视频片段集与独立视频片段,其中所述连续视频片段集为相邻两个匹配视频片段之间的时间间隔不超过预定值的匹配视频片段集合,所述独立视频片段为所有匹配视频片段中除所述连续视频片段集之外的匹配视频片段;

第一广告时段确定子单元,用于将所述连续视频片段集在所述电视视频样本中开始时刻之前设定时长的第一冗余时刻与结束时刻之后设定时长的第二冗余时刻之间的时间区间,以及所述独立视频片段在所述电视视频样本中开始时刻之前设定时长的第三冗余时刻与结束时刻之后设定时长的第四冗余时刻之间的时间区间确定为所述广告时段。

6. 根据权利要求5所述的装置,其特征在于,所述装置进一步包括插入单元,用于将确定的目标广告视频放入所述广告视频样本库。

7. 根据权利要求5所述的装置,其特征在于,所述装置进一步包括广告提取单元,用于将与各广告视频样本匹配的电视视频片段和所述目标广告视频作为所述电视视频包含的广告。

8. 根据权利要求5至7中任一权项所述的装置,其特征在于,所述目标广告确定单元利用对所述去除重复视频片段后的不匹配的电视视频片段的人工审核结果,确定目标广告视频;或者,所述目标广告确定单元通过大量广告视频数据训练得到的分类器,对所述去除重复视频片段后的不匹配的电视视频片段进行分类判断,确定目标广告视频。

一种发现广告视频的方法及装置

【技术领域】

[0001] 本发明涉及数据处理技术,特别涉及一种发现广告视频的方法及装置。

【背景技术】

[0002] 在电视视频中对广告进行监管是一项很重要的工作。在现有技术中,从电视视频中发现广告视频的方法比较简单,通常采用对全天的电视视频进行重复度比对,找出电视视频中重复出现次数较高的视频片段,确定为疑似广告,然后再通过人工审核,确定为广告样本。在这种方法下,存在两个问题:首先,这样做很难发现全天只播放了一次的广告视频,即发现的广告视频不够全面;其次,全天播放的其他重复视频也会被确定为疑似广告,例如全天多次播放的歌曲、电视剧的片头和片尾等等,这些本不属于广告的视频也需要通过人工审核,大大增加了审核的工作量。

【发明内容】

[0003] 本发明所要解决的技术问题是提供一种发现广告视频的方法及装置,以解决现有技术对广告监管时,发现的广告视频不全面、增加人工审核工作量的缺陷。

[0004] 本发明为解决技术问题而采用的技术方案是提供一种发现广告所述方法包括:

A. 利用预先建立的广告视频样本库中的各广告视频样本确定电视视频中的广告时段;
B. 将所述广告时段内的电视视频与各广告视频样本进行比对,并在与各广告视频样本不匹配的电视视频片段里去除重复视频片段,从去除重复视频片段后的不匹配的电视视频片段中确定目标广告视频,其中所述步骤 A 包括:在电视视频中采集电视视频样本;将所述电视视频样本与各广告视频样本逐一比对,将所述电视视频样本中与各广告视频样本比对上的视频片段确定为各匹配视频片段;将各匹配视频片段按照在所述电视视频样本中的时间顺序排列,从中确定连续视频片段集与独立视频片段,其中所述连续视频片段集为相邻两个匹配视频片段之间的时间间隔不超过预定值的匹配视频片段集合,所述独立视频片段为所有匹配视频片段中除所述连续视频片段集之外的匹配视频片段;将所述连续视频片段集在所述电视视频样本中开始时刻之前设定时长的第一冗余时刻与结束时刻之后设定时长的第二冗余时刻之间的时间区间,以及所述独立视频片段在所述电视视频样本中开始时刻之前设定时长的第三冗余时刻与结束时刻之后设定时长的第四冗余时刻之间的时间区间确定为所述广告时段。

[0005] 根据本发明之一优选实施例,所述方法在步骤 B 后进一步包括:将确定的目标广告视频放入所述广告视频样本库。根据本发明之一优选实施例,在所述步骤 B 之后还包括:将与各广告视频样本匹配的电视视频片段和所述目标广告视频作为所述电视视频包含的广告。

[0006] 根据本发明之一优选实施例,从去除重复视频片段后的不匹配的电视视频片段中确定目标广告视频包括:利用对所述去除重复视频片段后的不匹配的电视视频片段的人工审核结果,确定目标广告视频;或者,通过大量广告视频数据训练得到的分类器,对所述去

除重复视频片段后的不匹配的电视视频片段进行分类判断,确定目标广告视频。

[0007] 本发明还提供了一种发现广告的装置,包括:广告时段确定单元,用于利用预先建立的广告视频样本库中的各广告视频样本确定电视视频中的广告时段;去重单元,用于将所述广告时段内的电视视频与各广告视频样本进行比对,并在与各广告视频样本不匹配的电视视频片段里去除重复视频片段;目标广告确定单元,用于从去除重复视频片段后的不匹配的电视视频片段中确定目标广告视频;其中所述广告时段确定单元包括:第一采集子单元,用于在电视视频中采集电视视频样本;第一比对子单元,用于将所述电视视频样本与各广告视频样本逐一比对,将所述电视视频样本中与各广告视频样本比对上的视频片段确定为各匹配视频片段;匹配集确定子单元,用于将各匹配视频片段按照在所述电视视频样本中的时间顺序排列,从中确定连续视频片段集与独立视频片段,其中所述连续视频片段集为相邻两个匹配视频片段之间的时间间隔不超过预定值的匹配视频片段集合,所述独立视频片段为所有匹配视频片段中除所述连续视频片段集之外的匹配视频片段;第一广告时段确定子单元,用于将所述连续视频片段集在所述电视视频样本中开始时刻之前设定时长的第一冗余时刻与结束时刻之后设定时长的第二冗余时刻之间的时间区间,以及所述独立视频片段在所述电视视频样本中开始时刻之前设定时长的第三冗余时刻与结束时刻之后设定时长的第四冗余时刻之间的时间区间确定为所述广告时段。

[0008] 根据本发明之一优选实施例,所述装置进一步包括插入单元,用于将确定的目标广告视频放入所述广告视频样本库。

[0009] 根据本发明之一优选实施例,所述装置进一步包括广告提取单元,用于将与各广告视频样本匹配的电视视频片段和所述目标广告视频作为所述电视视频包含的广告。

[0010] 根据本发明之一优选实施例,所述目标广告确定单元利用对所述去除重复视频片段后的不匹配的电视视频片段的人工审核结果,确定目标广告视频;或者,所述目标广告确定单元通过大量广告视频数据训练得到的分类器,对所述去除重复视频片段后的不匹配的电视视频片段进行分类判断,确定目标广告视频。

[0011] 由以上技术方案可以看出,通过利用广告视频样本确定电视视频中的广告时段,再通过将广告时段内的电视视频与广告视频样本进行比对,可以很好地提高从电视视频中发现的广告视频数量,将电视视频中只播一次的广告视频提取出来,同时可以过滤电视视频中重复播放的非广告视频,提高广告监管的效率。

【附图说明】

[0012] 图1为本发明中发现广告视频的方法的实施例的流程示意图;

[0013] 图2为确定广告时段的方法的实施例一的流程示意图;

[0014] 图3为视频比对的方法的流程示意图;

[0015] 图4为广告时段的示意图;

[0016] 图5是确定广告时段的方法的实施例二的流程示意图

[0017] 图6为本发明中发现广告视频的装置的实施例一的结构示意框图;

[0018] 图7为本发明中发现广告视频的装置的实施例二的结构示意框图。

【具体实施方式】

[0019] 为了使本发明的目的、技术方案和优点更加清楚,下面结合附图和具体实施例对本发明进行详细描述。

[0020] 请参考图 1,图 1 为本发明中发现广告视频的方法实施例的流程示意图。如图 1 所示,所述方法包括:

[0021] 步骤 S101:利用预先建立的广告视频样本库中的各广告视频样本确定电视视频中的广告时段。

[0022] 步骤 S102:将广告时段内的电视视频与各广告视频样本进行比对,并在与各广告视频样本不匹配的电视视频片段里去除重复视频片段,从去除重复视频片段后的不匹配的电视视频片段中确定目标广告视频。

[0023] 下面对上述步骤进行具体说明。

[0024] 广告视频样本,是指单个广告的视频,广告视频样本库,就是包含多个广告视频的数据库。广告视频样本库中的广告视频样本,首先可通过收集既有广告视频得到,其次,从本发明的方法中发现的目标广告视频,最终也可以加入到广告视频样本库中作为广告视频样本。

[0025] 利用广告视频样本库中的各广告视频样本确定电视视频在播放中的广告时段,可采用以下两种方式进行。

[0026] 实施方式一:

[0027] 请参考图 2,图 2 为确定广告时段的方法的实施例一的流程示意图。如图 2 所示,确定广告时段的方法包括:

[0028] 步骤 S1011:在电视视频中采集电视视频样本。在本实施例中,可采集一天的电视视频作为电视视频样本。

[0029] 步骤 S1012:将电视视频样本与各广告视频样本逐一比对,从电视视频样本中确定与广告视频样本匹配的各视频片段。

[0030] 视频比对采用的方法可参考图 3,图 3 为视频比对的方法的流程示意图。电视视频样本即为图中的源视频,广告视频样本即为图中的目标视频。将源视频与目标视频进行解码后,就可按照图 3 中所示的方法,将两者进行比对。

[0031] 从图 3 可以看出,在将源视频与目标视频进行比对时,实际上是先取目标视频的第一帧,与源视频的每一帧比对,如果没有比对上,则取目标视频的下一帧与源视频的每一帧比对,直到找到源视频中与目标视频中能比对上的帧,这时,记录下该帧在源视频中的位置,再循环执行从源视频中取下一帧与目标视频中取的下一帧进行比对,直到比对失败,记录下最后一个比对成功的帧在源视频中的位置,此时,在源视频的两个记录的帧的位置之间的视频片段,即为匹配视频片段。在寻找源视频与目标视频中第一个相同帧时,也可以先取源视频的第一帧,与目标视频的每一帧比对,如果没有比对上,再取源视频的下一帧与目标视频的每一帧比对,以此类推。如果源视频与目标视频中没有一帧能比对成功,则说明源视频与该目标视频比对失败,源视频中没有该目标视频的匹配视频片段。

[0032] 将各个广告视频样本分别作为目标视频执行上述比对过程,就可以找到电视视频样本中的各个匹配视频片段了。

[0033] 请继续参考图 2。

[0034] 步骤 S1013:将各匹配视频片段按照在电视视频样本中的时间顺序排列,从中确

定连续视频片段集与独立视频片段。

[0035] 通常来说,在电视视频中播放的广告是有一定连续性的,这些广告都是在一段时间内集中播放,每个广告之间间隔的时间比较短,利用这个特点,可以找出连续视频片段集,该连续视频片段集是指相邻两个匹配视频片段之间的播放时间间隔不超过预定值的匹配视频片段集合。例如,有五个匹配视频片段,这五个匹配视频片段中两两相邻的匹配视频片段之间的播放时间间隔为 5 秒,则可以将这五个匹配视频片段作为一个连续视频片段集。所有匹配视频片段中除连续视频片段集之外的匹配视频片段就称之为独立视频片段。

[0036] 步骤 S1014 :将连续视频片段集在电视视频样本中的时间区间与独立视频片段在电视视频样本中的时间区间确定为广告时段。

[0037] 连续视频片段集在电视视频样本中的时间区间为第一冗余时刻至第二冗余时刻之间的时间区间,其中第一冗余时刻位于连续视频片段集中首个匹配视频片段在电视视频样本中的开始时刻之前,且该第一冗余时刻到连续视频片段集中首个匹配视频片段在电视视频样本中的开始时刻之间的时间长度不超过设定的第一阈值,第二冗余时刻位于连续视频片段集中最后一个匹配视频片段在电视视频样本中的结束时刻之后,且连续视频片段集中最后一个匹配视频片段在电视视频样本中的结束时刻到该第二冗余时刻之间的时间长度不超过设定的第二阈值。同理,独立视频片段在电视视频样本中的时间区间为第三冗余时刻至第四冗余时刻之间的时间区间,其中第三冗余时刻位于独立视频片段在电视视频样本中的开始时刻之前,且该第三冗余时刻到独立视频片段在电视视频样本中的开始时刻之间的时间长度不超过设定的第三阈值,第四冗余时刻位于独立视频片段在电视视频样本中的结束时刻之后,且独立视频片段在电视视频样本中的结束时刻到该第四冗余时刻之间的时间长度不超过设定的第四阈值。

[0038] 请参考图 4,图 4 为广告时段的示意图。从图上可以看出,在电视视频样本中,包含了连续视频片段集与独立视频片段,在确定广告时段时,广告时段是包含连续视频片段集的一段时间区间和包含独立视频片段的一段时间区间。广告时段中有一段可调节时间,由图中的第一阈值、第二阈值、第三阈值与第四阈值体现。第一阈值、第二阈值、第三阈值与第四阈值可以设置为相同,也可以设置为不同。例如连续视频片段集所处时间为第 10 分钟至第 20 分钟,广告时段可以是第 7 分钟至第 23 分钟,其中第一阈值与第二阈值相同,均为三分钟。

[0039] 利用广告视频样本库中的各广告视频样本确定电视视频在播放中的广告时段,还可以采用下面的方式实施。请参考图 5,图 5 是确定广告时段的方法的实施例二的流程示意图。

[0040] 实施方式二 :

[0041] 步骤 S101A :在电视视频中采集多个电视视频样本。在本实施例中,可采集一天的电视视频作为其中的一个电视视频样本。

[0042] 步骤 S101B :将多个电视视频样本与各广告视频样本逐一比对,确定在多个电视视频样本的相同时间段内均与同一广告视频样本比对上的视频片段为匹配视频片段。

[0043] 步骤 S101C :将匹配视频片段在电视视频样本中的时间区间确定为广告时段。

[0044] 匹配视频片段在电视视频样本中的时间区间为第五冗余时刻至第六冗余时刻之间的时间区间,所述第五冗余时刻位于匹配视频片段在电视视频样本中的开始时刻之前,

且该第五冗余时刻到匹配视频片段在电视视频样本中的开始时刻之间的时间长度不超过设定的第五阈值,第六冗余时刻位于匹配视频片段在电视视频样本中的结束时刻之后,且匹配视频片段在电视视频样本中的结束时刻到该第六冗余时刻之间的时间长度不超过设定的第六阈值。这里的第五阈值与第六阈值可以与实施例一中的第一阈值、第二阈值、第三阈值或第四阈值相同或不同,这里不再做具体说明。

[0045] 在确定广告时段的实施方式二中,通过电视视频的多个样本来确定匹配视频片段,从而找到广告时段,符合电视视频中广告出现的规律。可以理解,通常在电视视频中出现的广告具有时段重复性,即在电视视频中,很可能出现连续多天在同一时间段内播放的都是相同的广告,因此,采用实施方式二的方法,也能够很好地提取到广告时段。

[0046] 请继续参考图 1。在确定了广告时段后,就可以利用广告时段来确定电视视频中的目标广告视频。

[0047] 在步骤 S102 中,首先将广告时段内的电视视频与各广告视频样本进行比对,这时,广告时段内的电视视频就包括了与广告视频样本匹配的视频片段和与广告视频样本不匹配的视频片段,与广告视频样本匹配的视频片段就是广告视频样本库中的广告视频,而与广告视频样本不匹配的视频片段则需要进一步的处理,从中去除重复的视频片段。

[0048] 去除重复的视频片段就是将多个相同的视频片段进行归一化,保留其中的一个,其实施方法为:将与广告视频样本不匹配的视频片段进行两两比对,如果比对上则保留其中一个,将保留的这个视频片段与下一个未比对的视频片段进行比对。

[0049] 从去重处理后得到的不匹配视频片段中也可能确定出目标广告视频,也就是说,有些广告视频可能尚未被纳入广告视频样本库中。具体可采用的方法包括人工审核,审核人员从不匹配视频片段中选择出广告视频,根据审核人员的选择确定出目标广告视频。由于通过上述处理流程,只有这部分不匹配视频片段需要进行人工审核,大大降低了需要审核的可疑视频,因此极大地缩减了人力。此外,也可以采用机器学习算法,通过大量广告视频数据训练得到的分类器,对去重处理后得到的不匹配视频片段进行分类判断,以确定该不匹配视频片段是否为目标广告视频。

[0050] 经过上述处理流程得到的目标广告视频,可以进一步添加到广告视频样本库,也就是说,本发明的方法,既利用了广告视频样本来实现发现电视视频中仅出现一次的广告视频以及排除全天多次播放的非广告视频的目的,同时还可以利用最终的输出结果来扩充广告视频样本的数量,重复这个过程,最终可以大大提高在电视视频管理上的效率。

[0051] 此外,采用本发明的方法,还可以进一步地将步骤 102 中与各广告视频样本匹配的电视视频片段及目标广告视频作为电视视频中的广告。也就是说本发明的方法还可以用于从电视视频中提取广告。

[0052] 请参考图 6,图 6 为本发明中发现广告视频的装置的实施例一的结构示意框图。所述装置包括:广告时段确定单元 201、去重单元 202 及目标广告确定单元 203。

[0053] 广告时段确定单元 201,用于利用预先建立的广告视频样本库中的各广告视频样本确定电视视频中的广告时段。

[0054] 广告视频样本,是指单个广告的视频,广告视频样本库,就是包含多个广告视频的数据库。广告视频样本库中的广告视频样本,首先可通过收集既有广告视频得到,其次,利用本发明得到的目标广告视频,最终也可以加入到广告视频样本库中作为广告视频样本。

[0055] 广告时段确定单元 201 包括第一采集子单元 2011、第一比对子单元 2012、匹配集确定子单元 2013 以及第一广告时段确定子单元 2014。

[0056] 其中采集子单元 2011, 用于在电视视频中采集电视视频样本。在本实施例中, 可采集一天的电视视频作为电视视频样本。

[0057] 第一比对子单元 2012, 用于将电视视频样本与各广告视频样本逐一比对, 从电视视频样本中确定与广告视频样本匹配的各视频片段。

[0058] 视频比对采用的方法可参考图 3, 图 3 为视频比对的方法的流程示意图。电视视频样本即为图中的源视频, 广告视频样本即为图中的目标视频。将源视频与目标视频进行解码后, 就可按照图 3 中所示的方法, 将两者进行比对。

[0059] 从图 3 可以看出, 在将源视频与目标视频进行比对时, 实际上是先取目标视频的第一帧, 与源视频的每一帧比对, 如果没有比对上, 则取目标视频的下一帧与源视频的每一帧比对, 直到找到源视频中与目标视频中能比对上的帧, 这时, 记录下该帧在源视频中的位置, 再从源视频中取下一帧, 与目标视频中取的下一帧进行比对, 直到比对失败, 记录下最后一个比对成功的帧在源视频中的位置, 此时, 在源视频的两个记录的帧的位置之间的视频片段, 即为匹配视频片段。在寻找源视频与目标视频中第一个相同帧时, 也可以先取源视频的第一帧, 与目标视频的每一帧比对, 如果没有比对上, 再取源视频的下一帧与目标视频的每一帧比对, 以此类推。如果源视频与目标视频中没有一帧能比对成功, 则说明源视频与该目标视频比对失败, 源视频中没有该目标视频的匹配视频片段。

[0060] 将各个广告视频样本分别作为目标视频执行上述比对过程, 就可以找到电视视频样本中的各个匹配视频片段了。

[0061] 请继续参考图 6。

[0062] 匹配集确定子单元 2013, 用于将各匹配视频片段按照在电视视频样本中的时间顺序排列, 从中确定连续视频片段集与独立视频片段。

[0063] 通常来说, 在电视视频中播放的广告是有一定连续性的, 这些广告都是在一段时间内集中播放, 每个广告之间间隔的时间比较短, 利用这个特点, 可以找出连续视频片段集, 该连续视频片段集是指相邻两个匹配视频片段之间的播放时间间隔不超过预定值的匹配视频片段集合。例如, 有五个匹配视频片段, 这五个匹配视频片段中两两相邻的匹配视频片段之间的播放时间间隔为 5 秒, 则可以将这五个匹配视频片段作为一个连续视频片段集。所有匹配视频片段中除连续视频片段集之外的匹配视频片段就称之为独立视频片段。

[0064] 第一广告时段确定子单元 2014, 用于将连续视频片段集在电视视频样本中的时间区间与独立视频片段在电视视频样本中的时间区间确定为广告时段。

[0065] 连续视频片段集在电视视频样本中的时间区间为第一冗余时刻至第二冗余时刻之间的时间区间, 其中第一冗余时刻位于连续视频片段集中首个匹配视频片段在电视视频样本中的开始时刻之前, 且该第一冗余时刻到连续视频片段集中首个匹配视频片段在电视视频样本中的开始时刻之间的时间长度不超过设定的第一阈值, 第二冗余时刻位于连续视频片段集中最后一个匹配视频片段在电视视频样本中的结束时刻之后, 且连续视频片段集中最后一个匹配视频片段在电视视频样本中的结束时刻到该第二冗余时刻之间的时间长度不超过设定的第二阈值。同理, 独立视频片段在电视视频样本中的时间区间为第三冗余时刻至第四冗余时刻之间的时间区间, 其中第三冗余时刻位于独立视频片段在电视视频样

本中的开始时刻之前,且该第三冗余时刻到独立视频片段在电视视频样本中的开始时刻之间的时间长度不超过设定的第三阈值,第四冗余时刻位于独立视频片段在电视视频样本中的结束时刻之后,且独立视频片段在电视视频样本中的结束时刻到该第四冗余时刻之间的时间长度不超过设定的第四阈值。

[0066] 请参考图 4,图 4 为广告时段的示意图。从图上可以看出,在电视视频样本中,包含了连续视频片段集与独立视频片段,在确定广告时段时,广告时段是包含连续视频片段集的一段时间区间和包含独立视频片段的一段时间区间。广告时段中有一段可调节时间,由图中的第一阈值、第二阈值、第三阈值与第四阈值体现。第一阈值、第二阈值、第三阈值与第四阈值可以设置为相同,也可以设置为不同。例如连续视频片段集所处时间为第 10 分钟至第 20 分钟,则广告时段可以是第 7 分钟至第 23 分钟,其中第一阈值与第二阈值相同,均为三分钟。

[0067] 去重单元 202,用于将广告时段内播放的电视视频与各广告视频样本进行比对,并在与各广告视频样本不匹配的电视视频片段里去除重复视频片段。

[0068] 去除重复的视频片段就是将多个相同的视频片段进行归一化,保留其中的一个,其实施方法为:将与广告视频样本不匹配的视频片段进行两两比对,如果比对上则保留其中一个,将保留的这个视频片段与下一个未比对的视频片段进行比对。

[0069] 目标广告确定单元 203,用于从去除重复视频片段后的不匹配的电视视频片段中确定目标广告视频。

[0070] 目标广告确定单元 203 从去重处理后得到的不匹配视频片段中可以确定目标广告视频,包括两种方式:一种是利用对去除重复视频片段后的不匹配的电视视频片段的人工审核结果,确定目标广告视频。由于通过上述处理流程,只有这部分不匹配视频片段需要进行人工审核,大大降低了需要审核的可疑视频,因此极大地缩减了人力。第二种方式是目标广告确定单元 203 通过大量广告视频数据训练得到的分类器,对所述去除重复视频片段后的不匹配的电视视频片段进行分类判断,确定目标广告视频。

[0071] 请参考图 7,图 7 为本发明中发现广告视频的装置的实施例二的结构示意框图。如图 7 所示,所述装置包括:广告时段确定单元 301、去重单元 302、目标广告确定单元 303、插入单元 304 及广告提取单元 305。

[0072] 其中广告时段确定单元 301,用于利用预先建立的广告视频样本库中的各广告视频样本确定电视视频中的广告时段。广告时段确定单元 301,包括第二采集子单元 3011、第二比对子单元 3012 及第二广告时段确定子单元 3013。

[0073] 其中第二采集子单元 3011,用于在电视视频中采集多个电视视频样本。在本实施例中,可采集一天的电视视频作为其中的一个电视视频样本。

[0074] 第二比对子单元 3012,用于将多个电视视频样本与各广告视频样本逐一比对,确定在多个电视视频样本的相同时间段内均与同一广告视频样本比对上的视频片段为匹配视频片段。

[0075] 第二广告时段确定子单元 3013,用于将匹配视频片段在电视视频样本中的时间区间确定为广告时段。

[0076] 匹配视频片段在电视视频样本中的时间区间为第五冗余时刻至第六冗余时刻之间的时间区间,所述第五冗余时刻位于匹配视频片段在电视视频样本中的开始时刻之前,

且该第五冗余时刻到匹配视频片段在电视视频样本中的开始时刻之间的时间长度不超过设定的第五阈值,第六冗余时刻位于匹配视频片段在电视视频样本中的结束时刻之后,且匹配视频片段在电视视频样本中的结束时刻到该第六冗余时刻之间的时间长度不超过设定的第六阈值。

[0077] 去重单元 302,用于将广告时段内的电视视频与各广告视频样本进行比对,并在与各广告视频样本不匹配的电视视频片段里去除重复视频片段。

[0078] 目标广告确定单元 303,用于从去除重复视频片段后的不匹配的电视视频片段中确定目标广告视频。

[0079] 插入单元 304,用于将确定的目标广告视频放入广告视频样本库。本发明的装置,既利用了广告视频样本来实现发现电视视频中仅出现一次的广告视频以及排除全天多次播放的非广告视频的目的,同时还可以利用最终的输出结果来扩充广告视频样本的数量,重复这个过程,最终可以大大提高在电视视频管理上的效率。

[0080] 广告提取单元 305,用于将与各广告视频样本匹配的电视视频片段及目标广告视频作为电视视频中的广告。也就是说本发明的装置还可以用于从电视视频中提取广告。

[0081] 以上所述仅为本发明的较佳实施例而已,并不用以限制本发明,凡在本发明的精神和原则之内,所做的任何修改、等同替换、改进等,均应包含在本发明保护的范围之内。

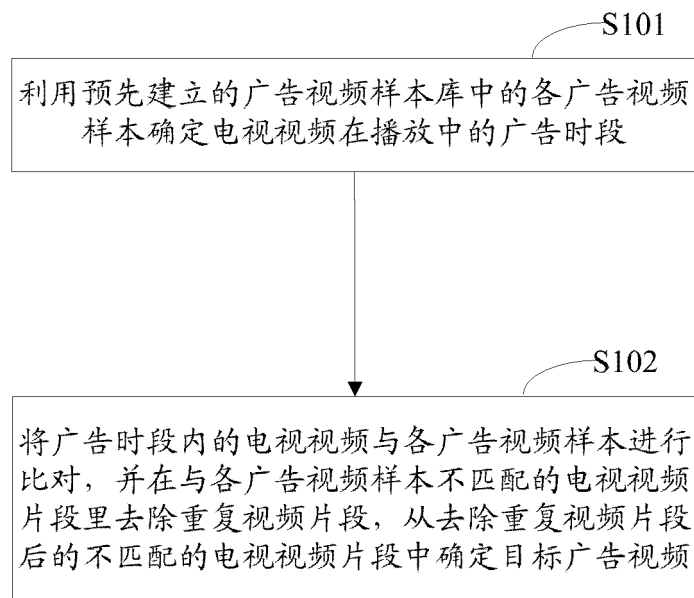


图 1

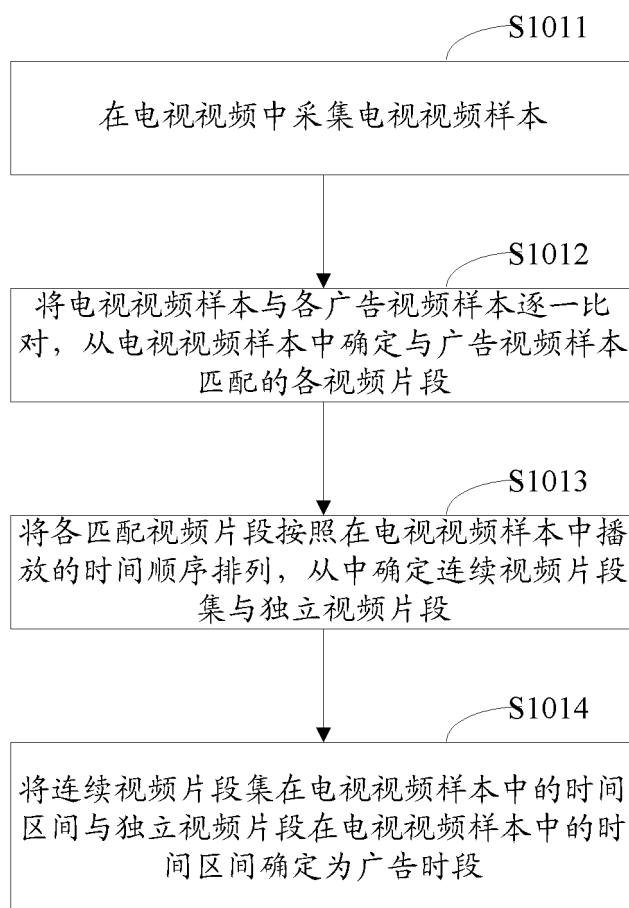


图 2

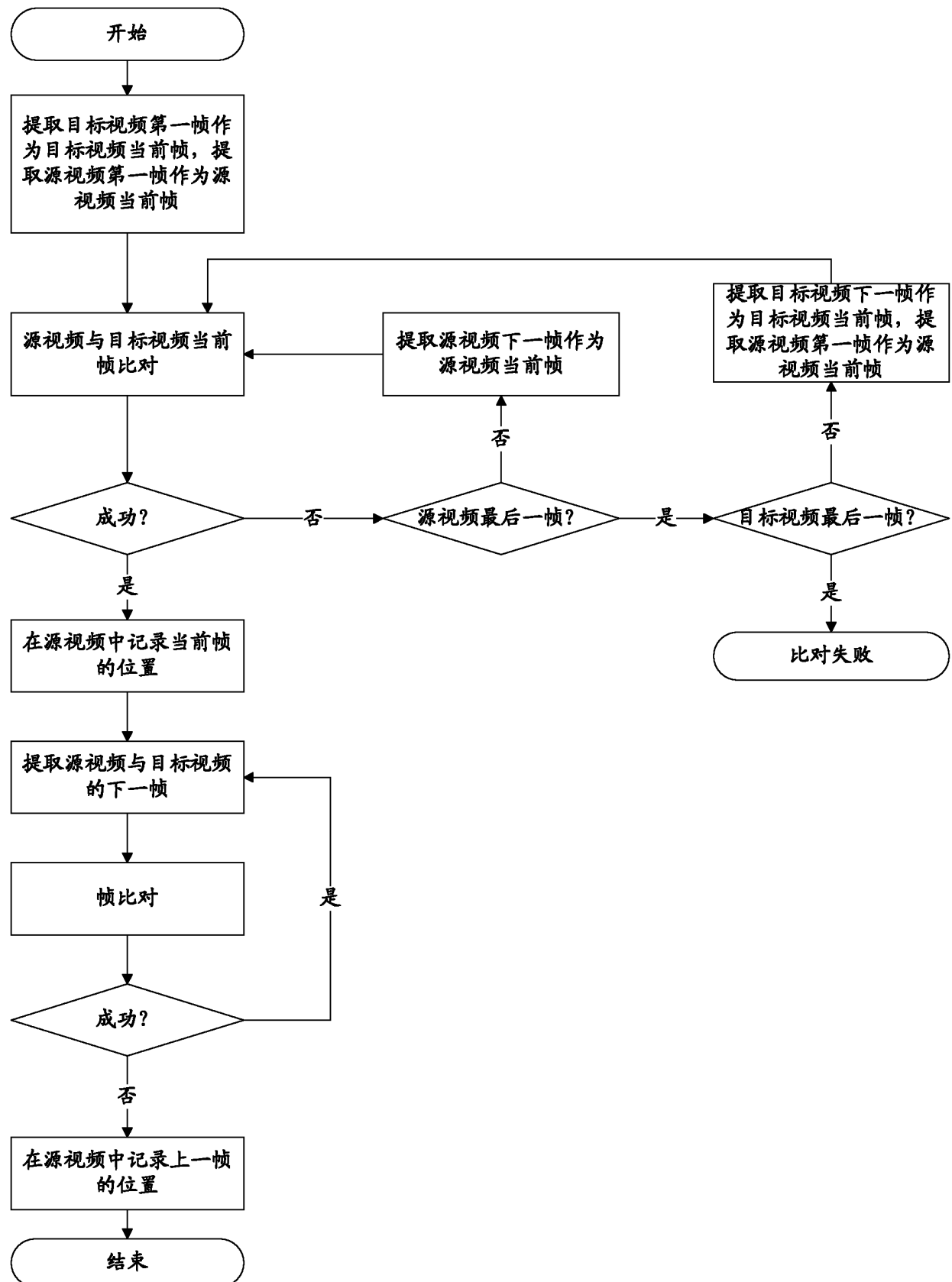


图 3

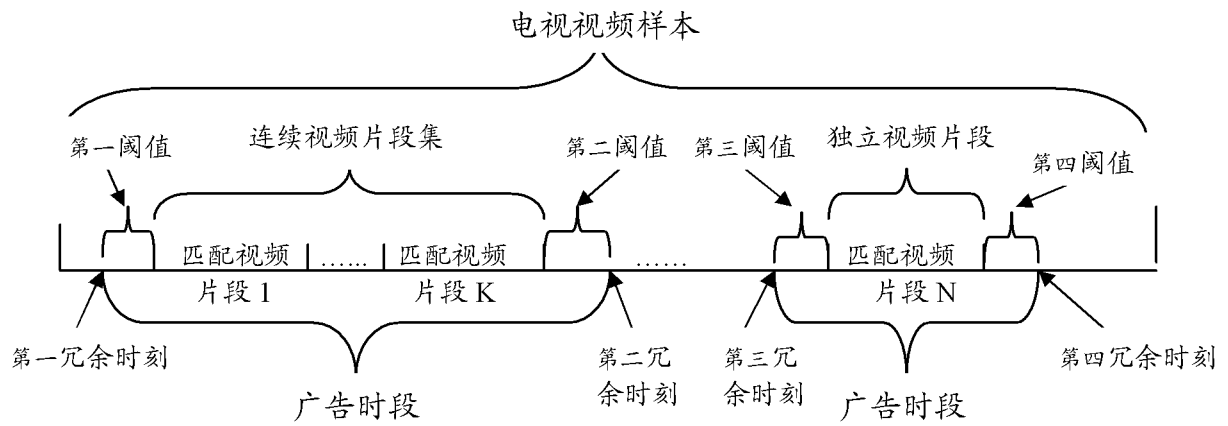


图 4

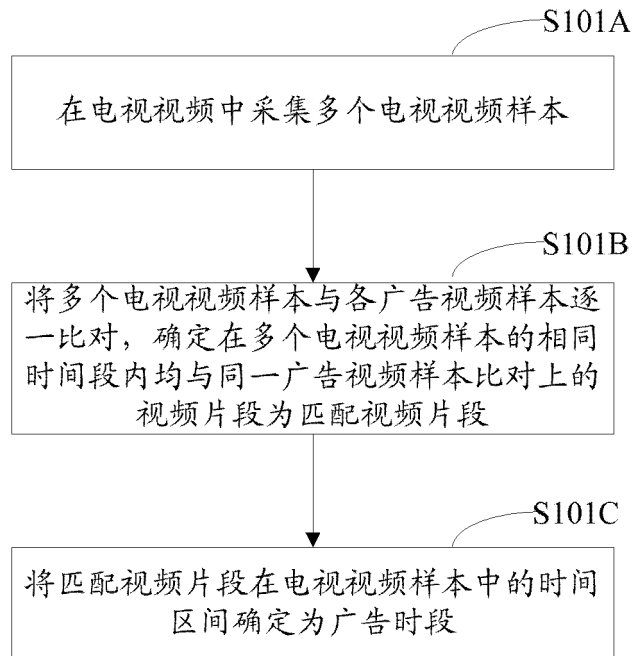


图 5

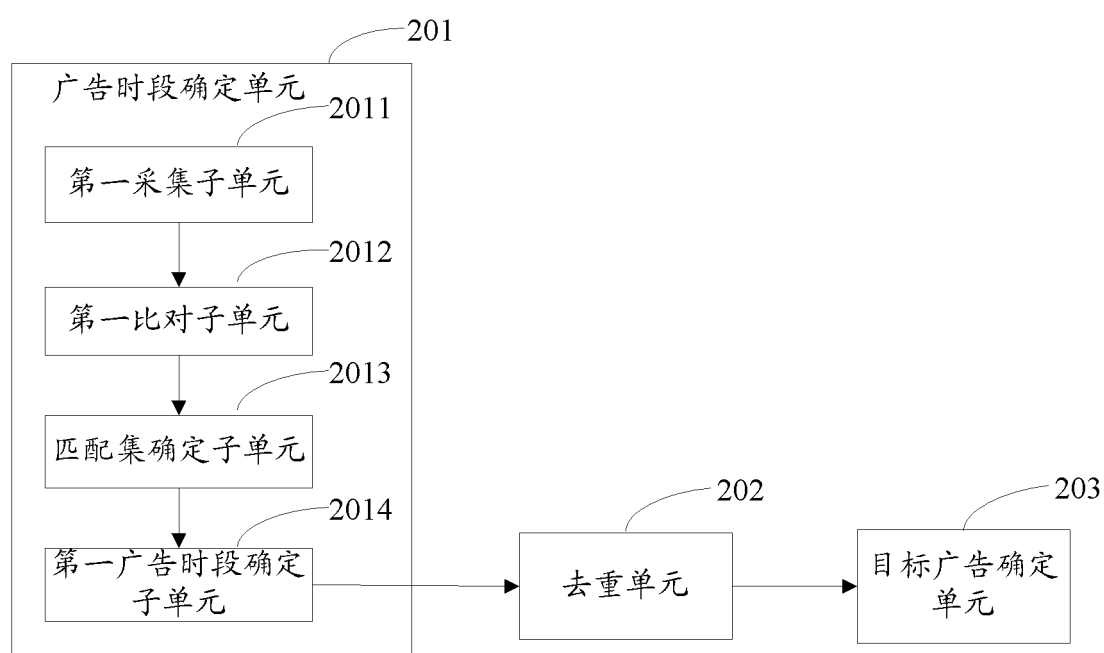


图 6

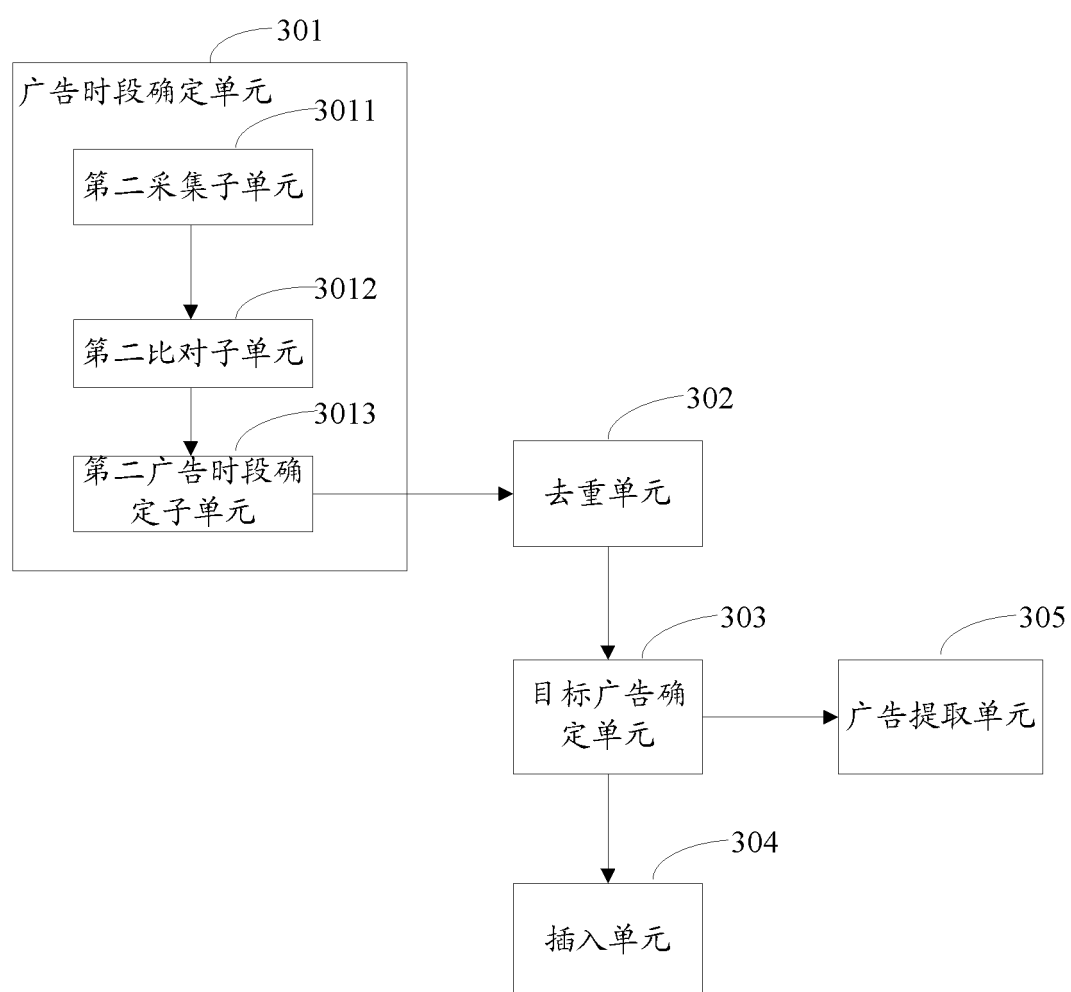


图 7