



(12)发明专利

(10)授权公告号 CN 106845390 B

(45)授权公告日 2019.09.20

(21)申请号 201710035420.7

审查员 陈飞

(22)申请日 2017.01.18

(65)同一申请的已公布的文献号

申请公布号 CN 106845390 A

(43)申请公布日 2017.06.13

(73)专利权人 腾讯科技(深圳)有限公司

地址 518000 广东省深圳市南山区高新区

科技中一路腾讯大厦35层

(72)发明人 曾佩玲

(74)专利代理机构 广州三环专利商标代理有限

公司 44202

代理人 郝传鑫 贾允

(51)Int.Cl.

G06K 9/00(2006.01)

G06F 16/738(2019.01)

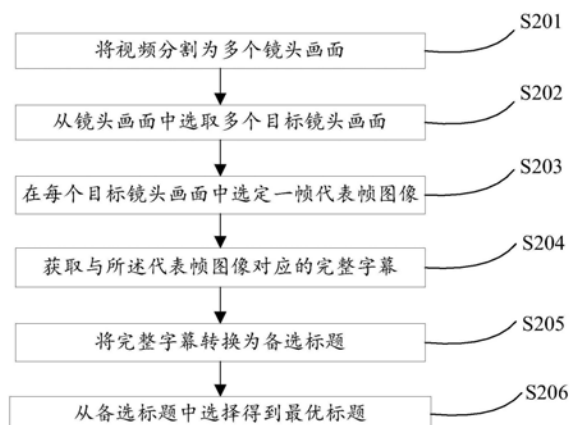
权利要求书3页 说明书13页 附图6页

(54)发明名称

视频标题生成方法及装置

(57)摘要

本发明公开了一种视频标题生成方法及装置,属于文本挖掘领域。所述视频标题生成方法包括:将视频分割为多个镜头画面;从镜头画面中选取多个目标镜头画面;在每个目标镜头画面中选定一帧图像,该帧图像作为相应目标镜头画面的代表帧图像;获取与所述代表帧图像对应的完整字幕,并将所述完整字幕转换为备选标题;从备选标题中选择得到最优标题,作为所述视频的标题。本发明通过从视频中提取代表帧的字幕生成备选标题,再对备选标题进行测试得到运营标题,无需人工制作标题,有利于更新的海量视频剧集自动生成运营标题,节约大量人力成本和时间成本。



1. 一种视频标题生成方法,其特征在于,所述方法包括:

将视频分割为多个镜头画面;

从镜头画面中选取多个目标镜头画面;

在每个目标镜头画面中选取一个或多个关键帧,计算所述关键帧的权重,选定权重值最大的关键帧作为目标镜头画面的代表帧图像;

获取与所述代表帧图像相邻的帧图像,提取代表帧图像和相邻的帧图像各自对应的字幕,得到与所述代表帧图像对应的具有上下文联系的完整字幕;

基于预先建立的自然语言转换模型对所述完整字幕进行识别和分割,所述自然语言转换模型基于自然语言库,所述自然语言库中预存有适作标题的短词短句;

以所述自然语言库为基础,对分割后的数据进行短词短句匹配查询,将得到的匹配结果进行筛选组装,得到备选标题;

从备选标题中选择得到最优标题,作为所述视频的标题。

2. 根据权利要求1所述的方法,其特征在于,所述从备选标题中选择得到最优标题包括:

将所述备选标题分别分配给相同数量的测试用户进行点击测试,得到点击率最高的备选标题,选定所述点击率最高的备选标题为最优标题。

3. 根据权利要求1所述的方法,其特征在于,所述从镜头画面中选取目标镜头画面包括:对镜头画面进行重要度评估,根据评估结果,选取符合重要度标准的目标镜头画面,

或者,设置特征标签,将命中特征标签的镜头画面选定为目标镜头画面。

4. 根据权利要求3所述的方法,其特征在于,

所述对镜头画面进行重要度评估包括:根据镜头画面包含的帧数量计算所述镜头画面的实际权重值;

所述根据评估结果,选取符合重要度标准的目标镜头画面包括:设定权重标准,选取实际权重值大于权重标准的镜头画面为目标镜头画面,或者,设定排名标准,选取实际权重值大小排名在所述排名标准内的镜头画面为目标镜头画面;

所述设置特征标签,将命中特征标签的镜头画面选定为目标镜头画面包括:将用户选中的人物设为特征标签,将所述镜头画面中的帧图像与所述特征标签比对,若比对结果为镜头画面中包括所述特征标签,则判定所述镜头画面为目标镜头,否则,判定所述镜头画面不是目标镜头。

5. 根据权利要求1所述的方法,其特征在于,

所述在每个目标镜头画面中选取一个或多个关键帧包括:对每个目标镜头画面中的所有帧图像进行聚类,得到一个或多个差异帧聚类,选取离所述差异帧聚类中心最近的帧图像作为所述目标镜头画面的关键帧;

所述计算所述关键帧的权重包括:选取多个权重参数,为每个权重参数设置相应的权重值,根据权重参数及相应的权重值构建权重模型,通过所述权重模型计算关键帧的权重值。

6. 根据权利要求1所述的方法,其特征在于,所述分割视频,得到镜头画面包括:

根据视频分割算法将视频分割为镜头画面,所述视频分割算法包括像素法、直方图法和边缘轮廓变化率法,

或者,还包括:将分割得到的镜头进一步分割成子镜头,所述镜头画面更新为所述子镜头。

7. 一种视频标题生成装置,其特征在于,包括:

分割模块,用于分割视频,得到镜头画面;

目标选取模块,用于从镜头画面中选取目标镜头画面;

帧模块,所述帧模块包括关键帧单元,所述关键帧单元用于在每个目标镜头画面中选取一个或多个关键帧,计算所述关键帧的权重,选定权重值最大的关键帧作为目标镜头画面的代表帧图像;

字幕模块,所述字幕模块包括提取单元和处理单元,所述提取单元用于获取与所述代表帧图像相邻的帧图像,提取代表帧图像和相邻的帧图像各自对应的字幕,得到与所述代表帧图像对应的具有上下文联系的完整字幕;所述处理单元用于基于预先建立的自然语言转换模型对所述完整字幕进行识别和分割,所述自然语言转换模型基于自然语言库,所述自然语言库中预存有适作标题的短词短句,

以所述自然语言库为基础,对分割后的数据进行短词短句匹配查询,将得到的匹配结果进行筛选组装,得到备选标题;

选择模块,用于从备选标题中选择得到最优标题,作为所述视频的标题。

8. 根据权利要求7所述的装置,其特征在于,所述选择模块包括点击测试单元,所述点击测试单元用于将所述备选标题分别分配给相同数量的测试用户进行点击测试,得到点击率最高的备选标题,选定所述点击率最高的备选标题为最优标题。

9. 根据权利要求7所述的装置,其特征在于,所述目标选取模块包括评估单元,所述评估单元用于对镜头画面进行重要度评估,根据评估结果,选取符合重要度标准的目标镜头画面,

或者,所述目标选取模块包括标签单元,所述标签单元用于将命中特征标签的镜头画面选定为目标镜头画面。

10. 根据权利要求9所述的装置,其特征在于,

所述评估单元根据镜头画面包含的帧数量计算所述镜头画面的实际权重值,并设定权重标准,选取实际权重值大于权重标准的镜头画面为目标镜头画面,或者,设定排名标准,选取实际权重值大小排名在所述排名标准内的镜头画面为目标镜头画面;

所述标签单元将用户选中的人物设为特征标签,将所述镜头画面中的帧图像与所述特征标签比对,若比对结果为镜头画面中包括所述特征标签,则判定所述镜头画面为目标镜头,否则,判定所述镜头画面不是目标镜头。

11. 根据权利要求7所述的装置,其特征在于,所述关键帧单元包括聚类子单元和权重子单元,

所述聚类子单元用于对每个目标镜头画面中的所有帧图像进行聚类,得到一个或多个差异帧聚类,选取离所述差异帧聚类中心最近的帧图像作为所述目标镜头画面的关键帧;

所述权重子单元用于选取多个权重参数,为每个权重参数设置相应的权重值,根据权重参数及相应的权重值构建权重模型,通过所述权重模型计算关键帧的权重值。

12. 根据权利要求7所述的装置,其特征在于,所述分割模块根据视频分割算法将视频分割为镜头画面,所述视频分割算法包括像素法、直方图法和边缘轮廓变化率法,

或者,所述分割模块还包括子镜头单元,所述子镜头单元用于将分割得到的镜头进一步分割成子镜头,所述镜头画面更新为所述子镜头。

视频标题生成方法及装置

技术领域

[0001] 本发明涉及文本挖掘领域,特别涉及一种视频标题生成方法及装置。

背景技术

[0002] 在视频播放浏览器或应用程序中,首页上显示的是推荐的视频,附带标注着该视频的运营标题,视频标题中包含描述该视频内容的文字标题,所述视频标题可以通过以下方法获得:运营人员通过浏览器获取全部视频信息,在全部视频中选出一张用于描述视频内容的视频图片,调整选出的视频图片的尺寸,并编辑对应的文字标题添加在该视频图片中作为视频的标题,显示在浏览器/播放App首页。

[0003] 视频产品运营位的标题是通过人工制作而成的,对于所有的用户展示相同的运营标题,甚至来不及随着剧情的更新而更新其视频标题,使得打开剧集选集页面时,仅出现数字选项,而没有相应的视频标题,无法帮助用户回忆上次的观看集数,精准找到续看的选集,用户体验较差;另一个方面,运营人员制作视频标题,消耗大量的人力成本,且容易造成人工误差,无法精准贴合视频内容,更无法实现为用户定制私人个性化标题。

[0004] 现有技术至少存在以下缺点:

[0005] 1、运营人员制作视频标题,耗费人力成本和时间成本,无法适应海量的视频更新速度;

[0006] 2、人工制作视频标题,误差性大;

[0007] 3、运营标题单一,无法实现个性化私人订制。

发明内容

[0008] 为了解决现有技术的问题,本发明提供了一种视频标题生成方法及装置,适应于海量视频的更新,节约人力成本。所述技术方案如下:

[0009] 一方面,本发明提供了一种视频标题生成方法,所述方法包括:

[0010] 将视频分割为多个镜头画面;

[0011] 从镜头画面中选取多个目标镜头画面;

[0012] 在每个目标镜头画面中选定一帧图像,该帧图像作为相应目标镜头画面的代表帧图像;

[0013] 获取与所述代表帧图像对应的完整字幕,并将所述完整字幕转换为备选标题;

[0014] 从备选标题中选择得到最优标题,作为所述视频的标题。

[0015] 具体地,所述从备选标题中选择得到最优标题包括:将所述备选标题分别分配给相同数量的测试用户进行点击测试,得到点击率最高的备选标题,选定所述点击率最高的备选标题为最优标题。

[0016] 可选地,所述从镜头画面中选取目标镜头画面包括:对镜头画面进行重要度评估,根据评估结果,选取符合重要度标准的目标镜头画面,

[0017] 或者,设置特征标签,将命中特征标签的镜头画面选定为目标镜头画面。

[0018] 进一步地,所述对镜头画面进行重要度评估包括:根据镜头画面包含的帧数量计算所述镜头画面的实际权重值;所述根据评估结果,选取符合重要度标准的目标镜头画面包括:设定权重标准,选取实际权重值大于权重标准的镜头画面为目标镜头画面,或者,设定排名标准,选取实际权重值大小排名在所述排名标准内的镜头画面为目标镜头画面。

[0019] 另一方面,所述设置特征标签,将命中特征标签的镜头画面选定为目标镜头画面包括:将用户选中的人物设为特征标签,将所述镜头画面中的帧图像与所述特征标签比对,若比对结果为镜头画面中包括所述特征标签,则判定所述镜头画面为目标镜头,否则,判定所述镜头画面不是目标镜头。

[0020] 进一步地,所述在每个目标镜头画面中选定一帧图像作为所述目标镜头画面的代表帧图像包括:在每个目标镜头画面中选取一个或多个关键帧,计算所述关键帧的权重,选定权重值最大的关键帧作为目标镜头画面的代表帧图像。

[0021] 具体地,所述在每个目标镜头画面中选取一个或多个关键帧包括:对每个目标镜头画面中的所有帧图像进行聚类,得到一个或多个差异帧聚类,选取离所述差异帧聚类中心最近的帧图像作为所述目标镜头画面的关键帧;

[0022] 所述计算所述关键帧的权重包括:选取多个权重参数,为每个权重参数设置相应的权重值,根据权重参数及相应的权重值构建权重模型,通过所述权重模型计算关键帧的权重值。

[0023] 进一步地,所述分割视频,得到镜头画面包括:根据视频分割算法将视频分割为镜头画面,所述视频分割算法包括像素法、直方图法、X2直方图法、X2直方图分块法和边缘轮廓变化率法,或者,还包括:将分割得到的镜头进一步分割成子镜头,所述镜头画面更新为所述子镜头。

[0024] 优选地,所述获取与所述代表帧图像对应的完整字幕,并将所述完整字幕转换为备选标题包括:获取与所述代表帧图像相邻的帧图像,提取代表帧图像和相邻的帧图像各自对应的字幕,得到与所述代表帧图像对应的完整字幕;对所述完整字幕进行自然语言处理,将处理结果作为所述备选标题。

[0025] 另一方面,本发明提供了一种视频标题生成装置,所述装置包括:

[0026] 分割模块,用于分割视频,得到镜头画面;

[0027] 目标选取模块,用于从镜头画面中选取目标镜头画面;

[0028] 帧模块,用于在每个目标镜头画面中选定一帧图像作为所述目标镜头画面的代表帧图像;

[0029] 字幕模块,用于获取与所述代表帧图像对应的完整字幕,并将所述完整字幕转换为备选标题;

[0030] 选择模块,用于从备选标题中选择得到最优标题,作为所述视频的标题。

[0031] 优选地,所述选择模块包括点击测试单元,所述点击测试单元用于将所述备选标题分别分配给相同数量的测试用户进行点击测试,得到点击率最高的备选标题,选定所述点击率最高的备选标题为最优标题。

[0032] 可选地,所述目标选取模块包括评估单元,所述评估单元用于对镜头画面进行重要度评估,根据评估结果,选取符合重要度标准的目标镜头画面,

[0033] 或者,所述目标选取模块包括标签单元,所述标签单元用于将命中特征标签的镜

头画面选定为目标镜头画面。

[0034] 进一步地,所述评估单元根据镜头画面包含的帧数量计算所述镜头画面的实际权重值,并设定权重标准,选取实际权重值大于权重标准的镜头画面为目标镜头画面,或者,设定排名标准,选取实际权重值大小排名在所述排名标准内的镜头画面为目标镜头画面;

[0035] 另一方面,所述标签单元将用户选中的人物设为特征标签,将所述镜头画面中的帧图像与所述特征标签比对,若比对结果为镜头画面中包括所述特征标签,则判定所述镜头画面为目标镜头,否则,判定所述镜头画面不是目标镜头。

[0036] 进一步地,所述帧模块包括关键帧单元,所述关键帧单元用于在每个目标镜头画面中选取一个或多个关键帧,计算所述关键帧的权重,选定权重值最大的关键帧作为目标镜头画面的代表帧图像。

[0037] 进一步地,所述关键帧单元包括聚类子单元和权重子单元,

[0038] 具体地,所述聚类子单元用于对每个目标镜头画面中的所有帧图像进行聚类,得到一个或多个差异帧聚类,选取离所述差异帧聚类中心最近的帧图像作为所述目标镜头画面的关键帧;

[0039] 所述权重子单元用于选取多个权重参数,为每个权重参数设置相应的权重值,根据权重参数及相应的权重值构建权重模型,通过所述权重模型计算关键帧的权重值。

[0040] 进一步地,所述分割模块根据视频分割算法将视频分割为镜头画面,所述视频分割算法包括像素法、直方图法、X2直方图法、X2直方图分块法和边缘轮廓变化率法,

[0041] 或者,所述分割模块还包括子镜头单元,所述子镜头单元用于将分割得到的镜头进一步分割成子镜头,所述镜头画面更新为所述子镜头。

[0042] 进一步地,所述字幕模块包括提取单元和处理单元,所述提取单元用于获取与所述代表帧图像相邻的帧图像,提取代表帧图像和相邻的帧图像各自对应的字幕,得到与所述代表帧图像对应的完整字幕;所述处理单元用于对所述完整字幕进行自然语言处理,将处理结果作为所述备选题。

[0043] 本发明提供的技术方案带来的有益效果如下:

[0044] 1) 视频运营标题能够随着视频剧集的更新而自动生成,无需人工制作,节约人力成本和时间成本;

[0045] 2) 能够根据用户的喜好,为用户私人订制符合其兴趣爱好的视频标题。

附图说明

[0046] 为了更清楚地说明本发明实施例中的技术方案,下面将对实施例描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0047] 图1是本发明实施例提供的视频标题生成方法及装置的实施环境示意图;

[0048] 图2是本发明实施例提供的视频标题生成方法的流程图;

[0049] 图3是本发明实施例提供的按照重要度生成视频标题的流程图;

[0050] 图4是本发明实施例提供的按照用户标签生成视频标题的流程图;

[0051] 图5是本发明实施例提供的视频标题生成装置的模块框图;

- [0052] 图6是本发明实施例提供的视频标题生成装置的计算机终端的硬件结构框图；
- [0053] 图7是本发明实施例提供的视频标题分类生成的流程图；
- [0054] 图8是本发明实施例提供的代表帧图像；
- [0055] 图9是图8中代表帧图像的相邻前一帧图像；
- [0056] 图10是图8中代表帧图像的相邻后一帧图像。

具体实施方式

[0057] 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。

[0058] 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、装置、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。

[0059] 本发明提供了一种视频标题生成方法及装置，参考图1可以看出，其示出了本发明实施例提供的视频标题生成方法及装置所涉及的实施环境的结构示意图。该实施环境包括配置有用户终端101、视频标题生成装置102和数据库103。

[0060] 其中，视频标题生成装置102如下述装置实施例所述，每个用户终端101通过视频标题生成装置102生成视频标题。视频标题生成装置102可以配置在用户终端101中，也可以独立于用户终端101而配置。实施环境中可以有一至多个视频标题生成装置102，组成一个视频标题生成装置的集群，当需要进行视频标题生成的用户终端101的个数较多时，多个视频标题生成装置102可以并行操作，从而提高视频标题生成的速度。

[0061] 数据库103中存储有视频标题生成装置102在视频标题生成过程中所需的数据，如目标视频等。

[0062] 数据库103除了可以独立于视频标题生成装置102和用户终端101配置之外，还可以配置在视频标题生成装置102中，使得视频标题生成装置102可以直接从自身配置的数据库中获取所需的数据，或者数据库103还可以配置在用户终端101中，使得视频标题生成装置102从用户终端101配置的数据库中获取所需的数据，本发明实施例在此不对数据库103的配置方式进行具体限定。

[0063] 用户终端101与视频标题生成装置102之间通过网络进行通信，视频标题生成装置102与数据库103之间也通过网络进行通信，该网络可以为无线网络或有线网络，本发明实施例在此不对通信方式进行具体限定。

[0064] 在本发明的一个实施例中提供了一种视频标题生成方法，参见图2，方法流程包括：

[0065] S101、将视频分割为多个镜头画面。

[0066] 具体地,从开机到关机所拍摄下来的一段连续的画面,叫做一个镜头画面,或两个剪接点之间的片段,也叫一个镜头画面。针对于前者,可以根据颜色空间的距离将视频分割为镜头画面;针对于后者,可以根据剪接点分割视频,得到多个镜头画面。除了上述两种方式,还可以根据视频分割算法将视频分割为镜头画面,所述视频分割算法包括但不限于像素法、直方图法、X2直方图法、X2直方图分块法和边缘轮廓变化率法。

[0067] S102、从镜头画面中选取多个目标镜头画面。

[0068] 参见图7,镜头画面有分重要镜头和不重要镜头,所述重要镜头为紧密贴合剧情的镜头画面,重要度特征具体为所述重要镜头包含的帧图像数量较多,选取所述重要镜头作为目标镜头画面;另一方面,镜头画面分为符合用户爱好的喜好镜头和非喜好镜头,所述喜好镜头的具体特征为包含用户标签内容,选取所述喜好镜头作为目标镜头画面,所述目标镜头画面的选取为生成视频标题提供了基础。

[0069] S103、在每个目标镜头画面中选定一帧图像,该帧图像作为相应目标镜头画面的代表帧图像。

[0070] 具体地,选取代表帧图像的方法可以是直接选取每个目标镜头画面中权重值最大的帧图像。代表帧图像的选取对生成视频标题的精确度具有重大意义,因此,优选地,可以选通过聚类得到若干个关键帧,然后再在所述关键帧中选取权重值最大的关键帧作为代表帧图像。

[0071] S104、获取与所述代表帧图像对应的完整字幕。

[0072] 具体地,所述完整字幕为以代表帧图像当前的字幕(如果所述代表帧图像当前没有字幕,则以代表帧图像为中心,获取离它最近的帧图像的字幕)为中心,获取前后若干帧图像的字幕,组成能够反映上下文联系的完整语境字幕。

[0073] S105、将所述完整字幕转换为备选标题。

[0074] 具体地,所述字幕为剧集中演员交流的台词,且完整字幕的字数比较多,不能够直接作为备选标题,因此,可以将所述完整字幕转换为自然语言的备选标题。

[0075] S106、从备选标题中选择得到最优标题,作为所述视频的标题。

[0076] 具体地,将多个备选标题分发给测试用户,按照一定的规则选取最优标题,所述最优标题即为该视频的运营标题。

[0077] 参见图7可知,有两种方式从镜头画面中选取目标镜头,在本发明的一个实施例中,提供了按照重要度模型生成视频标题的方法,参见图3,方法流程包括:

[0078] S301、将视频分割为多个镜头画面。

[0079] S302、将每个镜头画面分割为多个子镜头画面。

[0080] 为了提高自动生成的标题的贴合度,需要将视频分割的镜头画面越小越好,具体地,根据颜色空间的距离或剪接点将视频分割为镜头画面之后,再根据该镜头画面内相机运动进一步分割成子镜头画面。

[0081] S303、计算子镜头画面的权重。

[0082] 所述子镜头画面的权重取决于镜头的持续时间长度,即计算所述子镜头画面包含帧数量作为所述子镜头画面的权重大小。

[0083] S304、选取权重前N的子镜头画面。

[0084] 将所述子镜头画面按照权重由大到小排列,取权重较大的前N个子镜头画面作为目标镜头画面,一般重要的剧情都需要长镜头来描述,因此,视频标题从这些目标镜头画面的字幕中提取,准确性较高。

[0085] S305、遍历N个子镜头画面。

[0086] 依次对前N个子镜头画面执行S306-S311操作。

[0087] S306、将子镜头画面分割为帧图像。

[0088] 镜头的基本组成是帧图像,多个帧图像顺序播放形成镜头画面。

[0089] S307、对帧图像进行聚类,得到多个关键帧。

[0090] 本发明实施例中采用K-均值聚类算法对待聚类的帧图像进行聚类,对于帧图像集 $X = \{x_1, x_2, \dots, x_n\}$, n 为帧图像个数,设拟划分为 k 个聚类 $V = \{v_1, v_2, \dots, v_k\}$,先随机选取 K 个对象作为初始的聚类中心,然后计算每个对象与各个种子聚类中心之间的距离,把每个对象分配给距离它最近的聚类中心。聚类中心以及分配给它们的对象就代表一个聚类,一旦全部对象都被分配了,每个聚类的聚类中心会根据聚类中现有的对象被重新计算,这个过程将不断重复,直至聚类中心不再变化,算法终止,聚类得到 k 个差异比较大的帧类,相应地,每个帧类中离聚类中心最近的帧图像即作为关键帧(如果聚类中心本身就是帧图像,则聚类中心的帧图像即为关键帧),最终得到子镜头画面的 k 个关键帧。

[0091] 本发明不限定使用K-均值聚类法进行聚类,本发明也可以使用其他图像聚类算法,比如均值漂移(Mean Shift)算法,其算法原理是,在样本中随机选择一圆心为 o ,半径为 h 的区域,得出这个区域中所有样本点的平均值,圆心处的样本密度必然比均值处的样本密度小或者相等,将均值定为新的圆心重复以上步骤,直到收敛到密度极大值点。或者采用模糊C均值聚类算法,其工作原理是,算法将 n 个样本分为 c 个组,得到各个组的聚类中心,最终让非相似性指标的目标函数达到最小,算法给各个样本点赋予 $0 \sim 1$ 之间的隶属度,通过隶属度的值来判断样本归属于各个分类的程度,或者采用层次聚类算法等等。

[0092] S308、计算关键帧的权重。

[0093] 要从 k 个关键帧中选出一个代表帧,就得判断哪个关键帧最具有代表意义,比如,重要镜头都会给一个特写,这种情况下,人脸占比会比较大,具体地,通过建立关键帧的权重模型实现,所述权重模型中,选定计算参数,比如人脸占比、相机运动方向、相机摇摆动态、相机焦距等参数,为每个计算参数赋予不同的权重系数,得到权重模型,计算每个关键帧的权重值。

[0094] S309、选取权重最大的关键帧作为代表帧。

[0095] S310、提取附近的字幕。

[0096] 具体地,除了提取代表帧当前的字幕,还包括提取前后若干帧的字幕,使其形成具有上下文情景的完整字幕。

[0097] S311、转换成备选标题。

[0098] 字幕是镜头画面中演员的对话台词,需要将台词字幕进行自然语言转换操作,具体地,建立自然语言转换模型,所述自然语言转换模型基于自然语言库,自然语言库中预存有适作标题的短词短句,自然语言转换模型首先识别台词字幕,并对其进行智能分割,然后以自然语言库为基础进行匹配查询,将得到的匹配结果进行筛选组装,得到备选标题。参见图8为代表帧,字幕为“明明是在救你”,图9是代表帧之前的帧图像,字幕为“我说你怎么

动手打人呢”，图10是代表帧之后的帧图像，字幕为“莫非你是个女的”，查询到的匹配结果分别为“英雄救美”“不打不相识”“女扮男装”，再结合预输入的主演角色信息，转换成备选标题“北凉公主女扮男装，高阳王英雄救美”。

[0099] S312、判断是否完成遍历，若是，执行S313，若否，返回S305继续遍历下一个子镜头画面。

[0100] S313、对N个备选标题进行AB测试。

[0101] 每个子镜头画面会产生一个备选标题，完成遍历N个子镜头画面，得到N个备选标题。AB测试流程如下：确定测试人员数量，将其评价分为N组，将N个备选标题分配给N组测试人员。

[0102] S314、得到最优标题。

[0103] 将AB测试的最优测试结果作为最优标题，最优标题为所述视频的运营标题。

[0104] 在本发明的一个实施例中，提供了从镜头画面中选取目标镜头的第二种方式，即按照用户标签生成视频标题的方法，比如，用户喜爱的人物在该剧集中只是作为配角出现的，那么采用用户标签的方式可以为用户私人订制视频标题，参见图4，方法流程包括：

[0105] S401、将视频分割为多个镜头画面。

[0106] S402、将每个镜头画面分割为多个子镜头画面。

[0107] S403、比对用户标签，得到匹配子镜头。

[0108] 假设用户标签为某个明星，则建立对应于该明星的基准库，所述基准库中存储该明星的不同角度的照片，将子镜头画面中的帧图像与基准库中的照片进行比对，如果帧图像与照片匹配，则判定该子镜头画面为匹配子镜头。

[0109] S404、计算匹配子镜头的权重。

[0110] 镜头的权重取决于镜头的持续时间长度，即计算所述匹配子镜头包含的帧数量作为所述匹配子镜头的权重大小，这与上一种方式的计算子镜头画面的权重操作相同，包括以下操作步骤，均与上一种方式相同，在此不再赘述。

[0111] S405、选取权重前N的匹配子镜头。

[0112] 具体地，除了选取排名前N的方式来选取目标镜头以外，还可以设定一个权重标准，比如权重标准为100帧图像，则选取大于100帧图像的匹配子镜头作为选中的目标镜头。

[0113] S406、遍历N个匹配子镜头。

[0114] S407、将匹配子镜头分割为帧图像。

[0115] S408、对帧图像进行聚类，得到多个关键帧。

[0116] S409、计算关键帧的权重。

[0117] S410、选取权重最大的关键帧作为代表帧。

[0118] S411、提取附近的字幕。

[0119] S412、转换成备选标题。

[0120] S413、判断是否完成遍历，若是，执行S414，若否，返回S406继续遍历下一个子镜头画面。

[0121] S414、对N个备选标题进行AB测试。

[0122] S415、得到最优标题。

[0123] 除了上述的将AB测试的最优测试结果作为最优标题以外，也可以将测试得到的最

优的两个测试标题结合作为所述视频的运营标题。

[0124] 在本发明的一个实施例中,提供了一种视频标题生成装置,所述装置的模块架构参见图5,所述装置包括以下模块:

[0125] 分割模块510,用于分割视频,得到镜头画面;

[0126] 目标选取模块520,用于从镜头画面中选取目标镜头画面;

[0127] 帧模块530,用于在每个目标镜头画面中选定一帧图像作为所述目标镜头画面的代表帧图像;

[0128] 字幕模块540,用于获取与所述代表帧图像对应的完整字幕,并将所述完整字幕转换为备选标题;

[0129] 选择模块550,用于从备选标题中选择得到最优标题,作为所述视频的标题。

[0130] 具体地,所述分割模块510根据视频分割算法将视频分割为镜头画面,所述视频分割算法包括像素法、直方图法、X2直方图法、X2直方图分块法和边缘轮廓变化率法,除了视频分割算法,还可以根据颜色空间的距离或剪接点分割视频,本发明对于具体的分割方法不做限定。

[0131] 本实施例中,所述目标选取模块520包括评估单元521,所述评估单元521用于对镜头画面进行重要度评估,根据评估结果,选取符合重要度标准的目标镜头画面,所述评估单元521根据镜头画面包含的帧数量计算所述镜头画面的实际权重值,并设定权重标准,选取实际权重值大于权重标准的镜头画面为目标镜头画面,或者,设定排名标准,选取实际权重值大小排名在所述排名标准内的镜头画面为目标镜头画面。

[0132] 具体地,所述帧模块530包括关键帧单元531,所述关键帧单元531用于在每个目标镜头画面中选取一个或多个关键帧,计算所述关键帧的权重,选定权重值最大的关键帧作为目标镜头画面的代表帧图像。

[0133] 进一步地,所述关键帧单元531包括聚类子单元532和权重子单元533,

[0134] 所述聚类子单元532用于对每个目标镜头画面中的所有帧图像进行聚类,得到一个或多个差异帧聚类,选取离所述差异帧聚类中心最近的帧图像作为所述目标镜头画面的关键帧;

[0135] 所述权重子单元533用于选取多个权重参数,为每个权重参数设置相应的权重值,根据权重参数及相应的权重值构建权重模型,通过所述权重模型计算关键帧的权重值。

[0136] 具体地,所述字幕模块540包括提取单元541和处理单元542,所述提取单元541用于获取与所述代表帧图像相邻的帧图像,提取代表帧图像和相邻的帧图像各自对应的字幕,得到与所述代表帧图像对应的完整字幕;所述处理单元542用于对所述完整字幕进行自然语言处理,将处理结果作为所述备选标题。

[0137] 具体地,所述选择模块550包括点击测试单元551,所述点击测试单元551用于将所述备选标题分别分配给相同数量的测试用户进行点击测试,得到点击率最高的备选标题,选定所述点击率最高的备选标题为最优标题。

[0138] 在另一个实施例中,所述目标选取模块520包括标签单元522,所述标签单元522用于将命中特征标签的镜头画面选定为目标镜头画面,所述标签单元522将用户选中的人物设为特征标签,将所述镜头画面中的帧图像与所述特征标签比对,若比对结果为镜头画面中包括所述特征标签,则判定所述镜头画面为目标镜头,否则,判定所述镜头画面不是目标

镜头。

[0139] 在本发明的一个实施例中,所述分割模块510还包括子镜头单元511,所述子镜头单元511用于将分割得到的镜头进一步分割成子镜头,所述镜头画面更新为所述子镜头,即上述对镜头画面的操作转变为对子镜头的操作。

[0140] 需要说明的是:上述实施例提供的视频标题生成装置在生成视频标题时,仅以上述各功能模块的划分进行举例说明,实际应用中,可以根据需要而将上述功能分配由不同的功能模块完成,即将视频标题生成装置的内部结构划分成不同的功能模块,以完成以上描述的全部或者部分功能。另外,本实施例提供的视频标题生成装置实施例与上述实施例提供视频标题生成方法属于同一构思,其具体实现过程详见方法实施例,这里不再赘述。

[0141] 本发明实施例提供的方法实施例可以在移动终端、计算机终端或者类似的运算装置中执行,在本发明的一个实施例中,以运行在计算机终端上为例,图6是本发明实施例的视频标题生成装置的计算机终端的硬件结构框图。如图6所示,终端800可以包括RF (Radio Frequency, 射频) 电路110、包括有一个或一个以上计算机可读存储介质的存储器120、输入单元130、显示单元140、传感器150、音频电路160、WiFi (wireless fidelity, 无线保真) 模块170、包括有一个或者一个以上处理核心的处理器180、以及电源190等部件。本领域技术人员可以理解,图6中示出的终端结构并不构成对终端的限定,可以包括比图示更多或更少的部件,或者组合某些部件,或者不同的部件布置。其中:

[0142] RF电路110可用于收发信息或通话过程中,信号的接收和发送,特别地,将基站的下行信息接收后,交由一个或者一个以上处理器180处理;另外,将涉及上行的数据发送给基站。通常,RF电路110包括但不限于天线、至少一个放大器、调谐器、一个或多个振荡器、用户身份模块(SIM) 卡、收发信机、耦合器、LNA (Low Noise Amplifier, 低噪声放大器)、双工器等。此外,RF电路110还可以通过无线通信与网络和其他设备通信。所述无线通信可以使用任一通信标准或协议,包括但不限于GSM (Global System of Mobile communication, 全球移动通讯系统)、GPRS (General Packet Radio Service, 通用分组无线服务)、CDMA (Code Division Multiple Access, 码分多址)、WCDMA (Wideband Code Division Multiple Access, 宽带码分多址)、LTE (Long Term Evolution, 长期演进)、电子邮件、SMS (Short Messaging Service, 短消息服务) 等。

[0143] 存储器120可用于存储软件程序以及模块,处理器180通过运行存储在存储器120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。存储器120可主要包括存储程序区和存储数据区,其中,存储程序区可存储操作系统、功能所需的应用程序(比如声音播放功能、图像播放功能等)等;存储数据区可存储根据终端800的使用所创建的数据(比如音频数据、电话本等)等。此外,存储器120可以包括高速随机存取存储器,还可以包括非易失性存储器,例如至少一个磁盘存储器件、闪存器件、或其他易失性固态存储器件。相应地,存储器120还可以包括存储器控制器,以提供处理器180和输入单元130对存储器120的访问。

[0144] 输入单元130可用于接收输入的数字或字符信息,以及产生与用户设置以及功能控制有关的键盘、鼠标、操作杆、光学或者轨迹球信号输入。具体地,输入单元130可包括触敏表面131以及其他输入设备132。触敏表面131,也称为触摸显示屏或者触控板,可收集用户在其上或附近的触摸操作(比如用户使用手指、触笔等任何适合的物体或附件在触敏表面131上或在触敏表面131附近的操作),并根据预先设定的程式驱动相应的连接装置。可选

的,触敏表面131可包括触摸检测装置和触摸控制器两个部分。其中,触摸检测装置检测用户的触摸方位,并检测触摸操作带来的信号,将信号传送给触摸控制器;触摸控制器从触摸检测装置上接收触摸信息,并将它转换成触点坐标,再送给处理器180,并能接收处理器180发来的命令并加以执行。此外,可以采用电阻式、电容式、红外线以及表面声波等多种类型实现触敏表面131。除了触敏表面131,输入单元130还可以包括其他输入设备132。具体地,其他输入设备132可以包括但不限于物理键盘、功能键(比如音量控制按键、开关按键等)、轨迹球、鼠标、操作杆等中的一种或多种。

[0145] 显示单元140可用于显示由用户输入的信息或提供给用户的信息以及终端800的各种图形用户接口,这些图形用户接口可以由图形、文本、图标、视频和其任意组合来构成。显示单元140可包括显示面板141,可选的,可以采用LCD(Liquid Crystal Display,液晶显示器)、OLED(Organic Light-Emitting Diode,有机发光二极管)等形式来配置显示面板141。进一步的,触敏表面131可覆盖显示面板141,当触敏表面131检测到在其上或附近的触摸操作后,传送给处理器180以确定触摸事件的类型,随后处理器180根据触摸事件的类型在显示面板141上提供相应的视觉输出。虽然在图6中,触敏表面131与显示面板141是作为两个独立的部件来实现输入和输入功能,但是在某些实施例中,可以将触敏表面131与显示面板141集成而实现输入和输出功能。

[0146] 终端800还可包括至少一种传感器150,比如光传感器、运动传感器以及其他传感器。具体地,光传感器可包括环境光传感器及接近传感器,其中,环境光传感器可根据环境光线的明暗来调节显示面板141的亮度,接近传感器可在终端800移动到耳边时,关闭显示面板141和/或背光。作为运动传感器的一种,重力加速度传感器可检测各个方向上(一般为三轴)加速度的大小,静止时可检测出重力的大小及方向,可用于识别终端姿态的应用(比如横竖屏切换、相关游戏、磁力计姿态校准)、振动识别相关功能(比如计步器、敲击)等;至于终端800还可配置的陀螺仪、气压计、湿度计、温度计、红外线传感器等其他传感器,在此不再赘述。

[0147] 音频电路160、扬声器161,传声器162可提供用户与终端800之间的音频接口。音频电路160可将接收到的音频数据转换后的电信号,传输到扬声器161,由扬声器161转换为声音信号输出;另一方面,传声器162将收集的声音信号转换为电信号,由音频电路160接收后转换为音频数据,再将音频数据输出处理器180处理后,经RF电路110以发送给比如另一终端,或者将音频数据输出至存储器120以便进一步处理。音频电路160还可能包括耳塞插孔,以提供外设耳机与终端800的通信。

[0148] WiFi属于短距离无线传输技术,终端800通过WiFi模块170可以帮助用户收发电子邮件、浏览网页和访问流式媒体等,它为用户提供了无线的宽带互联网访问。虽然图6示出了WiFi模块170,但是可以理解的是,其并不属于终端800的必须构成,完全可以根据需要在不改变发明的本质的范围内而省略。

[0149] 处理器180是终端800的控制中心,利用各种接口和线路连接整个终端的各个部分,通过运行或执行存储在存储器120内的软件程序和/或模块,以及调用存储在存储器120内的数据,执行终端800的各种功能和处理数据,从而对终端进行整体监控。可选的,处理器180可包括一个或多个处理核心;优选地,处理器180可集成应用处理器和调制解调处理器,其中,应用处理器主要处理操作系统、用户界面和应用程序等,调制解调处理器主要处理无

线通信。可以理解的是,上述调制解调处理器也可以不集成到处理器180中。

[0150] 终端800还包括给各个部件供电的电源190(比如电池),优选的,电源可以通过电源管理系统与处理器180逻辑相连,从而通过电源管理系统实现管理充电、放电、以及功耗管理等功能。电源190还可以包括一个或一个以上的直流或交流电源、再充电系统、电源故障检测电路、电源转换器或者逆变器、电源状态指示器等任意组件。

[0151] 尽管未示出,终端800还可以包括摄像头、蓝牙模块等,在此不再赘述。具体在本实施例中,终端的显示单元是触摸屏显示器,终端还包括有存储器,以及一个或者一个以上的程序,其中一个或者一个以上程序存储于存储器中,且经配置以由一个或者一个以上处理器执行述一个或者一个以上程序包含用于进行以下操作的指令:

[0152] 将视频分割为多个镜头画面;

[0153] 从镜头画面中选取多个目标镜头画面;

[0154] 在每个目标镜头画面中选定一帧图像,该帧图像作为相应目标镜头画面的代表帧图像;

[0155] 获取与所述代表帧图像对应的完整字幕,并将所述完整字幕转换为备选标题;

[0156] 从备选标题中选择得到最优标题,作为所述视频的标题。

[0157] 具体地,终端的处理器还用于执行以下操作的指令:将所述备选标题分别分配给相同数量的测试用户进行点击测试,得到点击率最高的备选标题,选定所述点击率最高的备选标题为最优标题。

[0158] 具体地,终端的处理器还用于执行以下操作的指令:对镜头画面进行重要度评估,根据评估结果,选取符合重要度标准的目标镜头画面,

[0159] 或者,设置特征标签,将命中特征标签的镜头画面选定为目标镜头画面。

[0160] 具体地,终端的处理器还用于执行以下操作的指令:根据镜头画面包含的帧数量计算所述镜头画面的实际权重值;所述根据评估结果,选取符合重要度标准的目标镜头画面包括:设定权重标准,选取实际权重值大于权重标准的镜头画面为目标镜头画面,或者,设定排名标准,选取实际权重值大小排名在所述排名标准内的镜头画面为目标镜头画面。

[0161] 具体地,终端的处理器还用于执行以下操作的指令:将命中特征标签的镜头画面选定为目标镜头画面包括:将用户选中的人物设为特征标签,将所述镜头画面中的帧图像与所述特征标签比对,若比对结果为镜头画面中包括所述特征标签,则判定所述镜头画面为目标镜头,否则,判定所述镜头画面不是目标镜头。

[0162] 具体地,终端的处理器还用于执行以下操作的指令:在每个目标镜头画面中选取一个或多个关键帧,计算所述关键帧的权重,选定权重值最大的关键帧作为目标镜头画面的代表帧图像。

[0163] 具体地,终端的处理器还用于执行以下操作的指令:对每个目标镜头画面中的所有帧图像进行聚类,得到一个或多个差异帧聚类,选取离所述差异帧聚类中心最近的帧图像作为所述目标镜头画面的关键帧;

[0164] 具体地,终端的处理器还用于执行以下操作的指令:选取多个权重参数,为每个权重参数设置相应的权重值,根据权重参数及相应的权重值构建权重模型,通过所述权重模型计算关键帧的权重值。

[0165] 具体地,终端的处理器还用于执行以下操作的指令:根据视频分割算法将视频分

割为镜头画面,所述视频分割算法包括像素法、直方图法、X2直方图法、X2直方图分块法和边缘轮廓变化率法,或者,还包括:将分割得到的镜头进一步分割成子镜头,所述镜头画面更新为所述子镜头。

[0166] 具体地,终端的处理器还用于执行以下操作的指令:获取与所述代表帧图像相邻的帧图像,提取代表帧图像和相邻的帧图像各自对应的字幕,得到与所述代表帧图像对应的完整字幕;对所述完整字幕进行自然语言处理,将处理结果作为所述备选标题。

[0167] 通过以上实施方式的描述,本领域的技术人员可以清楚地了解到本发明提供的视频标题生成技术方案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。

[0168] 在本发明的一个实施例中,提供了一种计算机可读存储介质,该计算机可读存储介质可以是上述实施例中的存储器中所包含的计算机可读存储介质;也可以是单独存在,未装配入终端中的计算机可读存储介质。计算机可读存储介质存储有一个或者一个以上程序,一个或者一个以上程序被一个或者一个以上的处理器用来执行视频标题生成的方法,所述方法包括:

[0169] 将视频分割为多个镜头画面;

[0170] 从镜头画面中选取多个目标镜头画面;

[0171] 在每个目标镜头画面中选定一帧图像,该帧图像作为相应目标镜头画面的代表帧图像;

[0172] 获取与所述代表帧图像对应的完整字幕,并将所述完整字幕转换为备选标题;

[0173] 从备选标题中选择得到最优标题,作为所述视频的标题。

[0174] 具体地,所述从备选标题中选择得到最优标题包括:将所述备选标题分别分配给相同数量的测试用户进行点击测试,得到点击率最高的备选标题,选定所述点击率最高的备选标题为最优标题。

[0175] 可选地,所述从镜头画面中选取目标镜头画面包括:对镜头画面进行重要度评估,根据评估结果,选取符合重要度标准的目标镜头画面,

[0176] 或者,设置特征标签,将命中特征标签的镜头画面选定为目标镜头画面。

[0177] 进一步地,所述对镜头画面进行重要度评估包括:根据镜头画面包含的帧数量计算所述镜头画面的实际权重值;所述根据评估结果,选取符合重要度标准的目标镜头画面包括:设定权重标准,选取实际权重值大于权重标准的镜头画面为目标镜头画面,或者,设定排名标准,选取实际权重值大小排名在所述排名标准内的镜头画面为目标镜头画面。

[0178] 另一方面,所述设置特征标签,将命中特征标签的镜头画面选定为目标镜头画面包括:将用户选中的人物设为特征标签,将所述镜头画面中的帧图像与所述特征标签比对,若比对结果为镜头画面中包括所述特征标签,则判定所述镜头画面为目标镜头,否则,判定所述镜头画面不是目标镜头。

[0179] 进一步地,所述在每个目标镜头画面中选定一帧图像作为所述目标镜头画面的代表帧图像包括:在每个目标镜头画面中选取一个或多个关键帧,计算所述关键帧的权重,选

定权重值最大的关键帧作为目标镜头画面的代表帧图像。

[0180] 具体地,所述在每个目标镜头画面中选取一个或多个关键帧包括:对每个目标镜头画面中的所有帧图像进行聚类,得到一个或多个差异帧聚类,选取离所述差异帧聚类中心最近的帧图像作为所述目标镜头画面的关键帧;

[0181] 所述计算所述关键帧的权重包括:选取多个权重参数,为每个权重参数设置相应的权重值,根据权重参数及相应的权重值构建权重模型,通过所述权重模型计算关键帧的权重值。

[0182] 进一步地,所述分割视频,得到镜头画面包括:根据视频分割算法将视频分割为镜头画面,所述视频分割算法包括像素法、直方图法、X2直方图法、X2直方图分块法和边缘轮廓变化率法,或者,还包括:将分割得到的镜头进一步分割成子镜头,所述镜头画面更新为所述子镜头。

[0183] 优选地,所述获取与所述代表帧图像对应的完整字幕,并将所述完整字幕转换为备选标题包括:获取与所述代表帧图像相邻的帧图像,提取代表帧图像和相邻的帧图像各自对应的字幕,得到与所述代表帧图像对应的完整字幕;对所述完整字幕进行自然语言处理,将处理结果作为所述备选标题。

[0184] 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。

[0185] 本领域普通技术人员可以理解实现上述实施例的全部或部分步骤可以通过硬件来完成,也可以通过程序来指令相关的硬件完成,所述的程序可以存储于一种计算机可读存储介质中,上述提到的存储介质可以是只读存储器,磁盘或光盘等。

[0186] 以上所述仅为本发明的较佳实施例,并不用以限制本发明,凡在本发明的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本发明的保护范围之内。

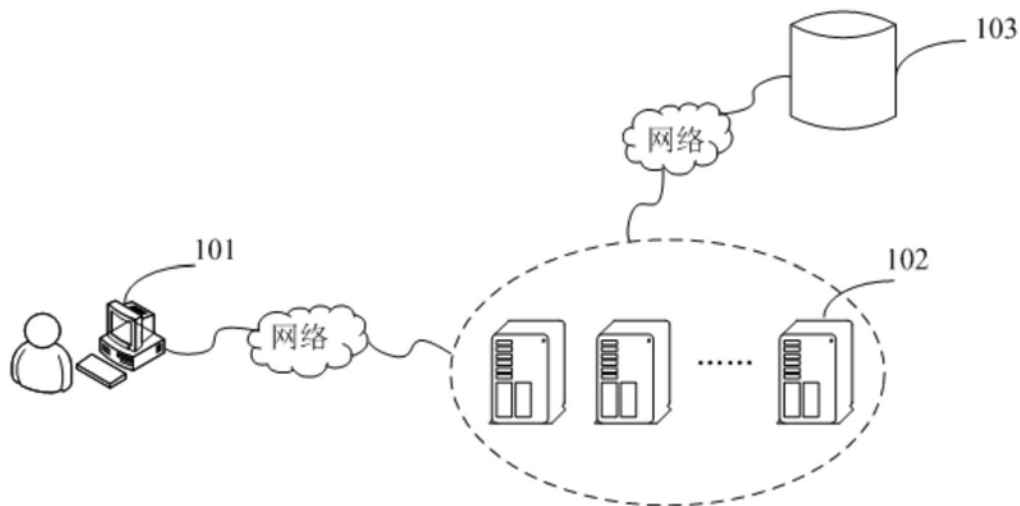


图1

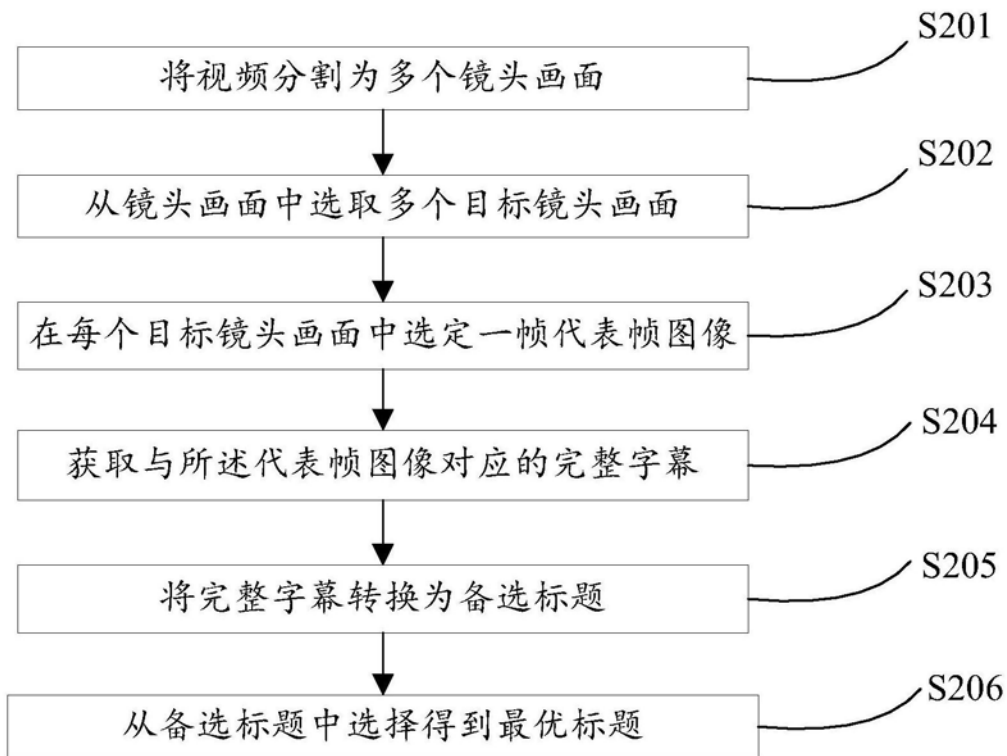


图2

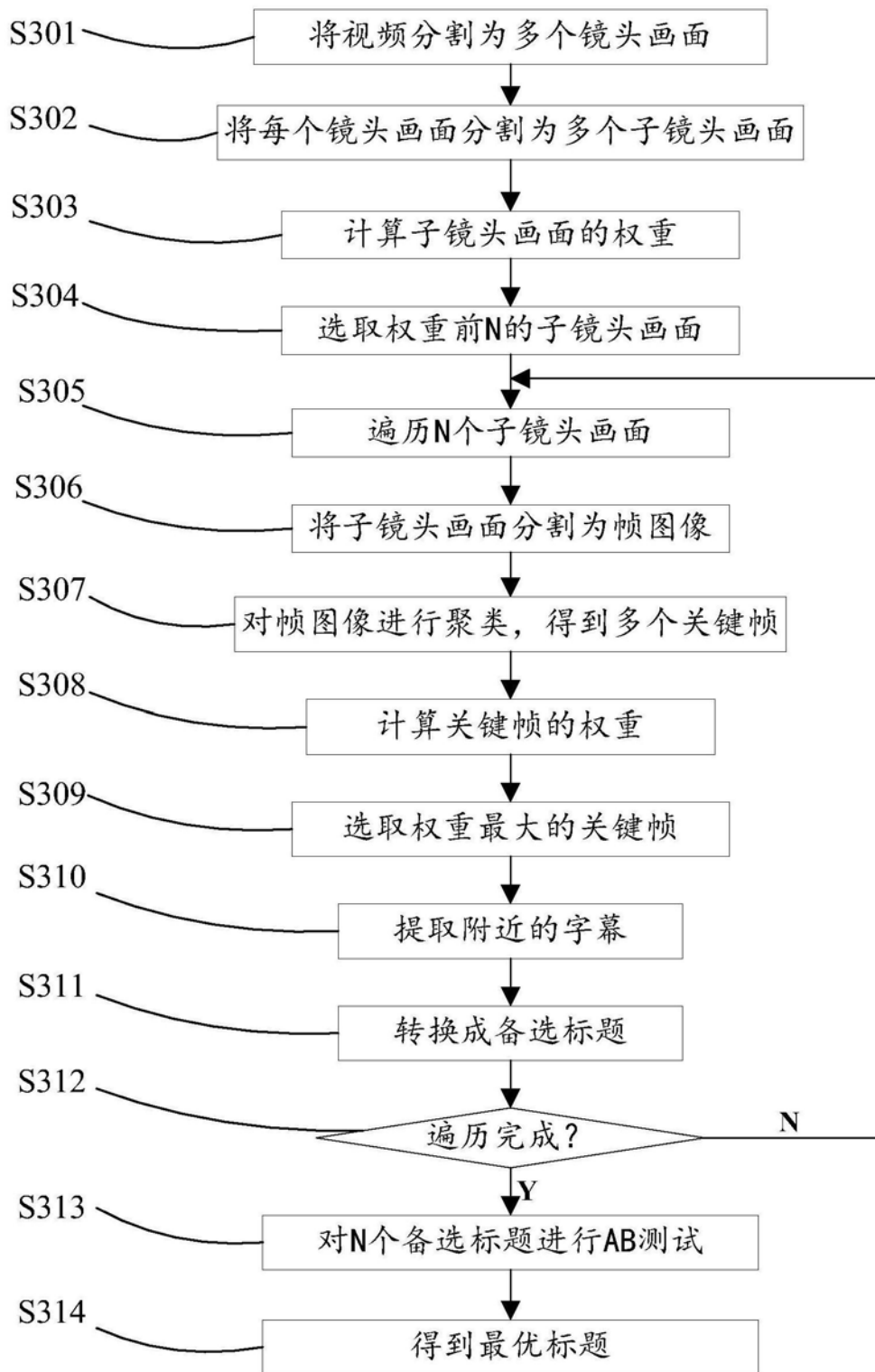


图3

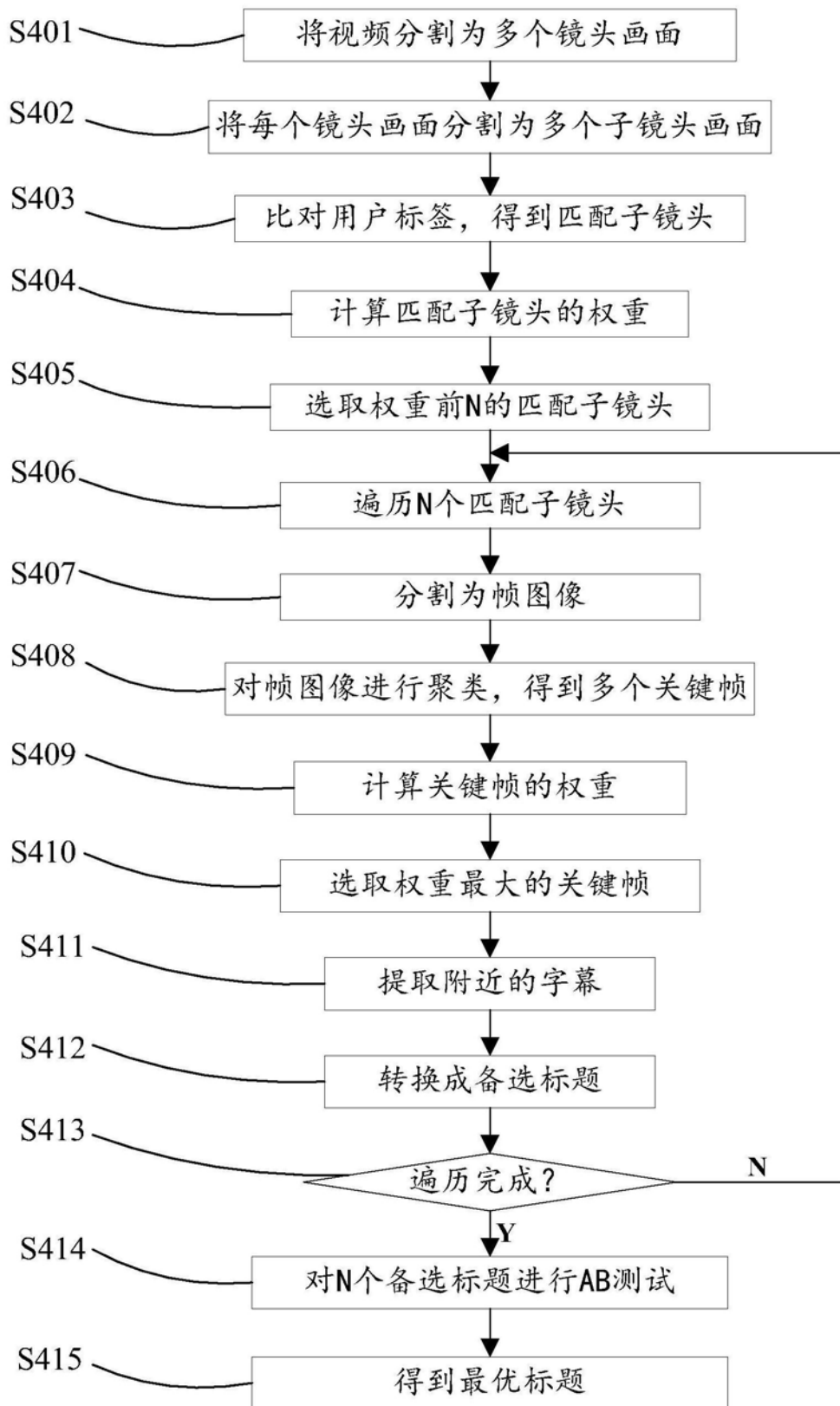


图4

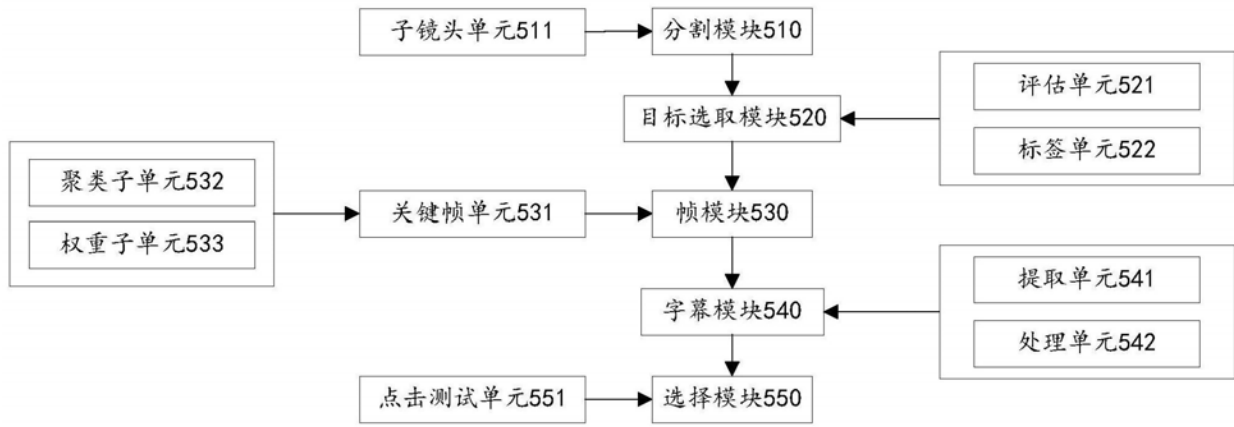


图5

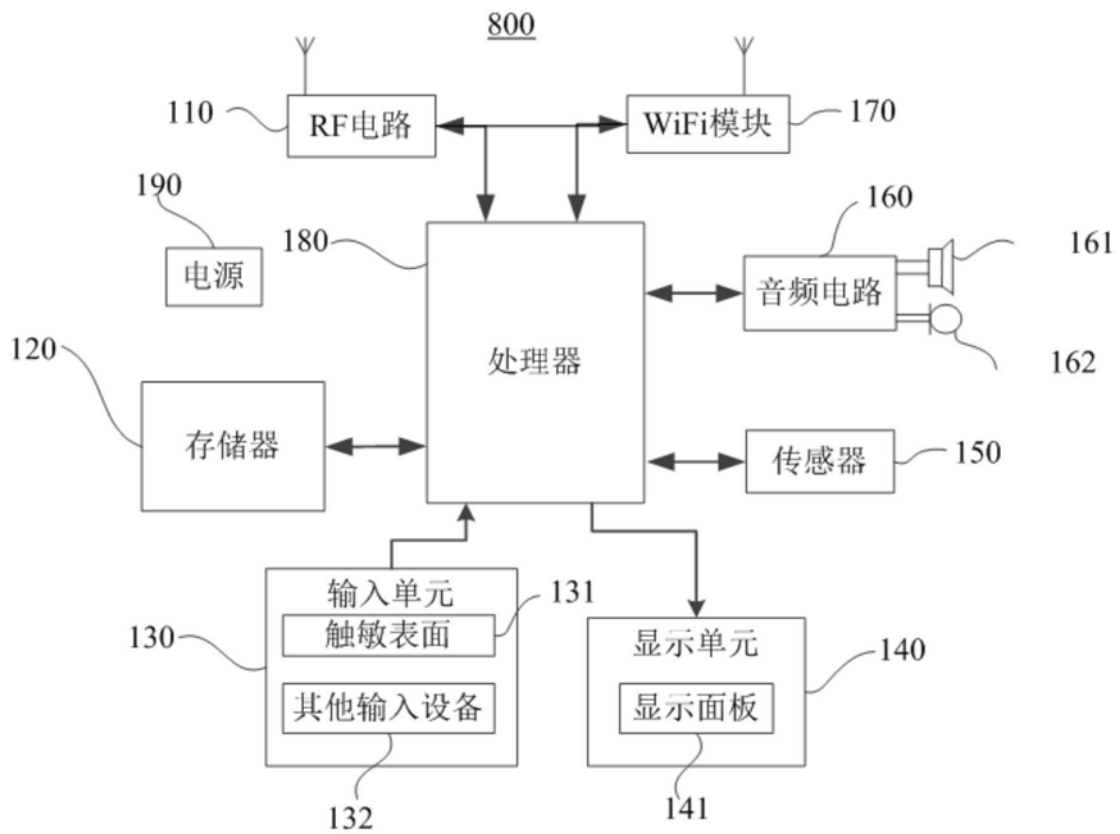


图6

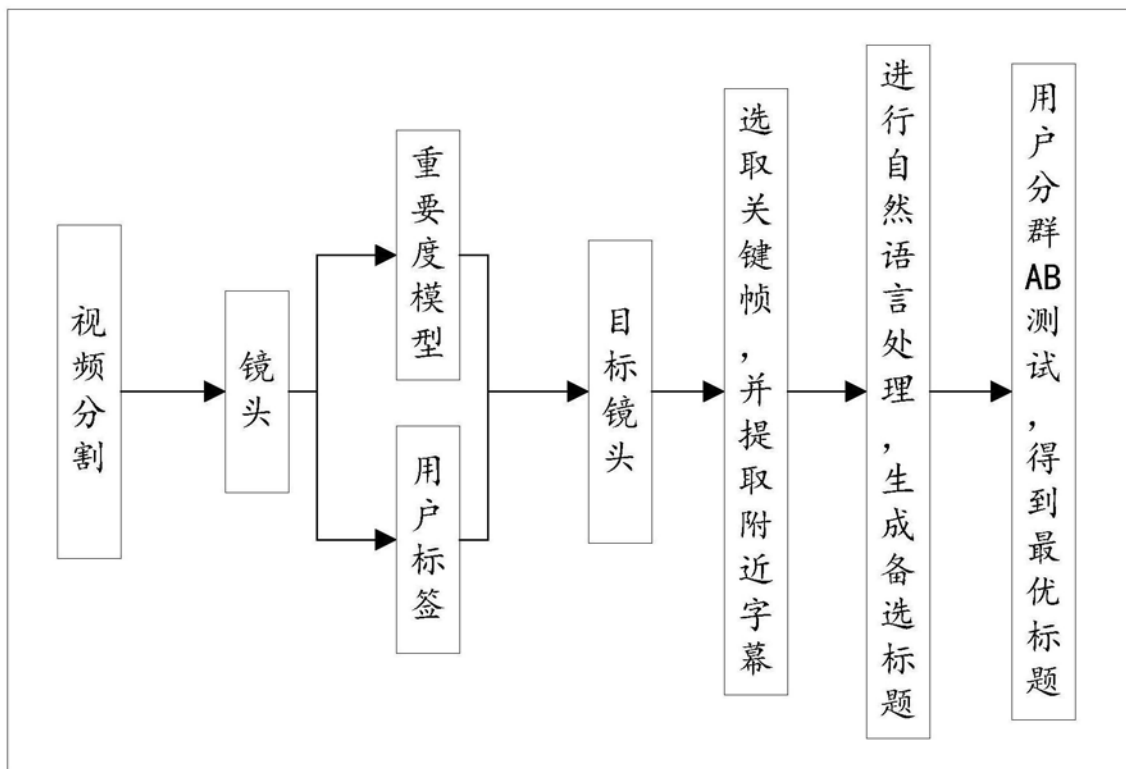


图7



图8



图9



图10