



(12)发明专利申请

(10)申请公布号 CN 108229476 A

(43)申请公布日 2018.06.29

(21)申请号 201810014862.8

(22)申请日 2018.01.08

(71)申请人 北京奇艺世纪科技有限公司
地址 100080 北京市海淀区北一街2号爱奇艺创新大厦10、11层

(72)发明人 刘楠

(74)专利代理机构 北京集佳知识产权代理有限公司 11227
代理人 王宝筠

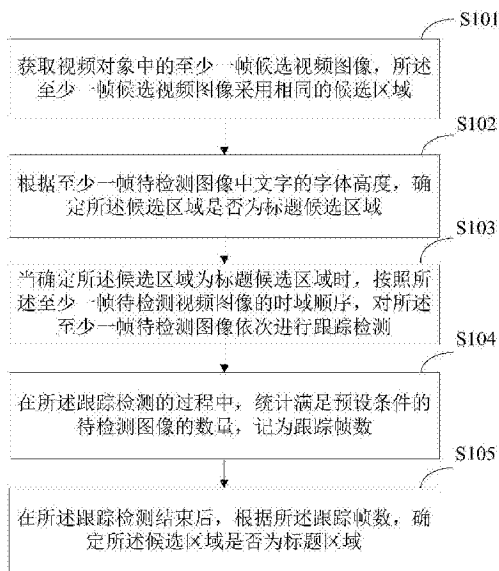
(51) Int. Cl.
G06K 9/32(2006.01)
G06K 9/00(2006.01)

权利要求书6页 说明书17页 附图5页

(54)发明名称
标题区域检测方法及系统

(57)摘要

本发明提供了一种标题区域检测方法及系统,获取视频对象中采用相同的候选区域的至少一帧候选视频图像,根据候选区域中的至少一帧待检测图像中文字的字体高度,确定候选区域是否为标题候选区域;当确定候选区域为标题候选区域时,按照至少一帧待检测视频图像的时域顺序,对候选区域进行跟踪检测;在跟踪检测的过程中,统计满足预设条件的待检测图像的跟踪帧数;最后根据跟踪帧数,确定候选区域是否为标题区域。本发明基于新闻标题的文字的字体高度特征,能够滤除滚动字幕的干扰,准确判断出标题候选区域,并通过对标题候选区域的跟踪检测,能够最终确定出标题候选区域是否为标题区域,提高标题区域的检测精度。



1. 一种标题区域检测方法,其特征在于,包括:

获取视频对象中的至少一帧候选视频图像,所述至少一帧候选视频图像采用相同的候选区域;

根据至少一帧待检测图像中文字的字体高度,确定所述候选区域是否为标题候选区域;其中,所述至少一帧待检测图像分别为所述至少一帧候选视频图像的候选区域内的图像;

当确定所述候选区域为标题候选区域时,按照所述至少一帧待检测视频图像的时域顺序,对所述至少一帧待检测图像依次进行跟踪检测;

在所述跟踪检测的过程中,统计满足预设条件的待检测图像的数量,记为跟踪帧数;

在所述跟踪检测结束后,根据所述跟踪帧数,确定所述候选区域是否为标题区域。

2. 根据权利要求1所述的方法,其特征在于,所述获取视频对象中的至少一帧候选视频图像包括:

获取视频对象中的任一帧视频图像;

获取所述任一帧图像的标题潜在区域内的图像,作为潜在图像;所述标题潜在区域为用户指定的标题所在区域;

计算所述潜在图像的边缘强度图;

根据所述潜在图像的边缘强度图在水平方向上投影得到的水平直方图,确定候选区域的上下边界;

根据所述潜在图像的边缘强度图在垂直方向上投影得到的垂直直方图,确定候选区域的左右边界;

根据所述上下边界与所述左右边界,确定候选区域;

当确定出的候选区域符合预设条件时,将所述任一帧视频图像确定为候选视频图像。

3. 根据权利要求1所述的方法,其特征在于,所述根据至少一帧待检测图像中文字的字体高度,确定所述候选区域是否为标题候选区域包括:

根据任一帧待检测图像,计算所述任一帧待检测图像的文本掩码;

根据所述文本掩码,确定所述任一帧待检测图像中文字的字体高度;

根据全部待检测图像中文字的字体高度,计算全部待检测图像的文字字体平均高度;

根据所述文字字体平均高度,确定所述候选区域是否为标题候选区域。

4. 根据权利要求3所述的方法,其特征在于,所述根据任一帧待检测图像,计算所述任一帧待检测图像的文本掩码包括:

根据任一帧待检测图像,生成所述任一帧待检测图像的灰度图像;

计算所述任一帧待检测图像的灰度图像的分割阈值,其中,所述任一帧待检测图像的灰度图像的分割阈值作为第三分割阈值;

利用所述第三分割阈值,对所述任一帧待检测图像的灰度图像进行二值化处理,得到灰度二值化图像;

根据所述灰度二值化图像,计算所述灰度二值化图像的平均像素值;

根据所述平均像素值,确定背景像素值;

根据所述灰度二值化图像与背景像素值,生成与所述任一帧待检测图像相对应的文本掩码,其中,所述文本掩码为所述任一帧待检测图像中文本区域的掩码。

5. 根据权利要求4所述的方法,其特征在于,所述根据所述灰度二值化图像,计算所述灰度二值化图像的平均像素值包括:

水平扫描所述灰度二值化图像的任一行像素;

将所述任一行中的首个像素值记为所述任一行的参考像素值;

统计所述任一行中像素值不等于所述任一行的参考像素值的像素个数,记为第一个数;

当所述第一个数小于预设个数阈值时,确定所述任一行符合第一预设条件;

统计符合第一预设条件的行数以及符合第一预设条件的所有行的参考像素值;

垂直扫描所述灰度二值化图像的任一列像素;

将所述任一列中的首个像素值记为所述任一列的参考像素值;

统计所述任一列中像素值不等于所述任一列的参考像素值的像素个数,记为第二个数;

当所述第二个数小于预设个数阈值时,确定所述任一列符合第二预设条件;

统计符合第二预设条件的列数以及符合第二预设条件的所有列的参考像素值;

计算符合第一预设条件的所有行的参考像素值与符合第二预设条件的所有列的参考像素值的像素值总和,以及,符合第一预设条件的行数与符合第二预设条件的列数的行列总数;

将所述像素值总和与所述行列总数的比值确定为所述灰度二值化图像的平均像素值。

6. 根据权利要求4所述的方法,其特征在于,所述根据所述平均像素值,确定背景像素值包括:

当所述平均像素值小于预设像素值时,确定背景像素值为第一像素值;

当所述平均像素值不小于预设像素值时,确定背景像素值为第二像素值;

所述第一像素值与所述第二像素值不同。

7. 根据权利要求4所述的方法,其特征在于,所述根据所述灰度二值化图像与背景像素值,生成与所述任一帧待检测图像相对应的文本掩码包括:

将所述灰度二值化图像中像素值与所述背景像素值相同的区域确定为背景区域;

将所述灰度二值化图像中像素值与所述背景像素值不同的区域确定为文字区域;

根据所述背景区域与所述文字区域,生成与所述待检测图像相对应的文本掩码。

8. 根据权利要求7所述的方法,其特征在于,所述方法还包括:

将所述背景区域的像素值设置为第三像素值;

将所述文字区域的像素值设置为第四像素值;

所述第三像素值与所述第四像素值不同。

9. 根据权利要求3所述的方法,其特征在于,所述根据所述文本掩码,确定所述任一帧待检测图像中文字的字体高度包括:

利用所述文本掩码在水平方向上的投影,生成掩码水平特征 H_b ;

计算所述任一帧待检测图像的灰度图像的边缘强度图 E_{a11} ;

对所述任一帧待检测图像的灰度图像的边缘强度图 E_{a11} 进行二值化处理,得到边缘二值化图像 E ;

利用所述边缘二值化图像 E 在水平方向上的投影,生成边缘水平特征 H_e ;

根据所述掩码水平特征 H_b 与所述边缘水平特征 H_e ,计算图像分割特征 H_s ;

在 $i \in [0, H-1]$ 的范围内,对所述图像分割特征 $H_s[i]$ 进行遍历;

在所述遍历的过程中,当所述图像分割特征 $H_s[i]$ 符合水平分割条件时,记录每一水平方向上的位置 $[i, i+n]$ 为文字行区域,并将所述位置 $[i, i+n]$ 存入预设列表 POS_{text} ;

根据所述预设列表 POS_{text} 中符合有效高度条件的位置 $[i, i+n]$,计算分割区域 $SEG(0, i, W, n)$;

将所有分割区域 $SEG(0, i, W, n)$ 中 n 值的平均值确定为所述任一帧待检测图像中文字的字体高度;

其中,所述水平分割条件为同时满足第一条件、第二条件与第三条件;第一条件为: $i=0$ 或 $H_s[i-1]=1$;第二条件为:对于每一个 j 值, $H_s[i+j]$ 均等于0, $0 \leq j \leq n$, j 与 n 为自然数;第三条件为: $H_s[i+n+1]=1$ 或 $i=H-1$;

所述有效高度条件为: $n \geq M_h$;其中, $M_h = H / (N+1)$, N 为所述预设列表 POS_{text} 中的位置 $[i, i+n]$ 的数量;

$$\text{其中, } H_b[i] = \begin{cases} 0, & \text{numb}[i] \geq TH1 \\ 1, & \text{numb}[i] < TH1 \end{cases}, \quad E(x,y) = \begin{cases} 0, & E_{all}(x,y) \leq TK4 \\ 1, & E_{all}(x,y) > TK4 \end{cases}, \quad H_e[i] = \begin{cases} 0, & \text{nume}[i] \geq TH2 \\ 1, & \text{nume}[i] < TH2 \end{cases},$$

$H_s[i] = H_b[i] \mid H_e[i]$; $\text{numb}[i]$ 为所述掩码图像第 $i+1$ 行中像素值不等于所述背景像素值的像素个数; $TH1$ 为预设的第一特征阈值; $\text{nume}[i]$ 为所述边缘二值化图像第 $i+1$ 行中像素值不等于0的像素个数, $TH2$ 为预设的第二特征阈值; $TK4$ 为第四分割阈值; $i \in [0, H-1]$, H 为所述待检测图像的高度, W 为所述待检测图像的宽度。

10. 根据权利要求1所述的方法,其特征在于,所述按照所述至少一帧待检测视频图像的时域顺序,对所述至少一帧待检测图像依次进行跟踪检测包括:

当所述标题候选区域为首次跟踪时,获取当前帧待检测图像在跟踪区域内的图像,作为候选跟踪图像;所述跟踪区域与所述标题候选区域相对应;

计算所述候选跟踪图像的参考颜色直方图与参考二值化图像;

当所述标题候选区域为非首次跟踪时,获取当前帧待检测图像在跟踪区域内的图像,作为当前跟踪图像;

计算所述当前跟踪图像的当前颜色直方图与当前二值化图像;

根据所述当前颜色直方图与所述参考颜色直方图的差异以及所述当前二值化图像与所述参考二值化图像的差异,确定当前帧待检测图像是否满足预设条件。

11. 根据权利要求10所述的方法,其特征在于,所述方法还包括:

在跟踪检测过程中,统计不满足预设条件的待检测图像的数量,记为丢失帧数;

当所述丢失帧数大于预设丢失阈值时,结束对所述至少一帧待检测图像依次进行跟踪检测的过程。

12. 一种标题区域检测系统,其特征为,包括:

候选图像获取单元,用于获取视频对象中的至少一帧候选视频图像,所述至少一帧候选视频图像采用相同的候选区域;

字体高度过滤单元,用于根据至少一帧待检测图像中文字的字体高度,确定所述候选区域是否为标题候选区域;其中,所述至少一帧待检测图像分别为所述至少一帧候选视频

图像的候选区域内的图像；

候选区域跟踪单元,用于当确定所述候选区域为标题候选区域时,按照所述至少一帧待检测视频图像的时域顺序,对所述至少一帧待检测图像依次进行跟踪检测；

所述候选区域跟踪单元,还用于在所述跟踪检测的过程中,统计满足预设条件的待检测图像的数量,记为跟踪帧数；

标题区域确定单元,用于在所述跟踪检测结束后,根据所述跟踪帧数,确定所述候选区域是否为标题区域。

13. 根据权利要求12所述的系统,其特征在于,所述候选图像获取单元包括:

第一获取单元,用于获取视频对象中的任一帧视频图像；

第二获取单元,用于获取所述任一帧图像的标题潜在区域内的图像,作为潜在图像;所述标题潜在区域为用户指定的标题所在区域；

第一计算单元,用于计算所述潜在图像的边缘强度图；

第一确定单元,用于根据所述潜在图像的边缘强度图在水平方向上投影得到的水平直方图,确定候选区域的上下边界；

第二确定单元,用于根据所述潜在图像的边缘强度图在垂直方向上投影得到的垂直直方图,确定候选区域的左右边界；

第三确定单元,用于根据所述上下边界与所述左右边界,确定候选区域；

第四确定单元,用于当确定出的候选区域符合预设条件时,将所述任一帧视频图像确定为候选视频图像。

14. 根据权利要求12所述的系统,其特征在于,所述字体高度过滤单元包括:

第二计算单元,用于根据任一帧待检测图像,计算所述任一帧待检测图像的文本掩码；

第五确定单元,用于根据所述文本掩码,确定所述任一帧待检测图像中文字的字体高度；

第三计算单元,用于根据全部待检测图像中文字的字体高度,计算全部待检测图像的文字字体平均高度；

第六确定单元,用于根据所述文字字体平均高度,确定所述候选区域是否为标题候选区域。

15. 根据权利要求14所述的系统,其特征在于,所述第二计算单元具体用于:

根据任一帧待检测图像,生成所述任一帧待检测图像的灰度图像；

计算所述任一帧待检测图像的灰度图像的分割阈值,其中,所述任一帧待检测图像的灰度图像的分割阈值作为第三分割阈值；

利用所述第三分割阈值,对所述任一帧待检测图像的灰度图像进行二值化处理,得到灰度二值化图像；

根据所述灰度二值化图像,计算所述灰度二值化图像的平均像素值；

根据所述平均像素值,确定背景像素值；

根据所述灰度二值化图像与背景像素值,生成与所述任一帧待检测图像相对应的文本掩码,其中,所述文本掩码为所述任一帧待检测图像中文本区域的掩码。

16. 根据权利要求15所述的系统,其特征在于,所述第二计算单元具体用于:

水平扫描所述灰度二值化图像的任一行像素；

将所述任一行中的首个像素值记为所述任一行的参考像素值；
统计所述任一行中像素值不等于所述任一行的参考像素值的像素个数，记为第一个数；
当所述第一个数小于预设个数阈值时，确定所述任一行符合第一预设条件；
统计符合第一预设条件的行数以及符合第一预设条件的所有行的参考像素值；
垂直扫描所述灰度二值化图像的任一列像素；
将所述任一列中的首个像素值记为所述任一列的参考像素值；
统计所述任一列中像素值不等于所述任一列的参考像素值的像素个数，记为第二个数；
当所述第二个数小于预设个数阈值时，确定所述任一列符合第二预设条件；
统计符合第二预设条件的列数以及符合第二预设条件的所有列的参考像素值；
计算符合第一预设条件的所有行的参考像素值与符合第二预设条件的所有列的参考像素值的像素值总和，以及，符合第一预设条件的行数与符合第二预设条件的列数的行列总数；
将所述像素值总和与所述行列总数的比值确定为所述灰度二值化图像的平均像素值。

17. 根据权利要求15所述的系统，其特征在于，所述第二计算单元具体用于：
当所述平均像素值小于预设像素值时，确定背景像素值为第一像素值；
当所述平均像素值不小于预设像素值时，确定背景像素值为第二像素值；
所述第一像素值与所述第二像素值不同。

18. 根据权利要求15所述的系统，其特征在于，所述第二计算单元具体用于：
将所述灰度二值化图像中像素值与所述背景像素值相同的区域确定为背景区域；
将所述灰度二值化图像中像素值与所述背景像素值不同的区域确定为文字区域；
根据所述背景区域与所述文字区域，生成与所述待检测图像相对应的文本掩码。

19. 根据权利要求18所述的系统，其特征在于，所述第二计算单元具体用于：
将所述背景区域的像素值设置为第三像素值；
将所述文字区域的像素值设置为第四像素值；
所述第三像素值与所述第四像素值不同。

20. 根据权利要求14所述的系统，其特征在于，所述第五确定单元具体用于：
利用所述文本掩码在水平方向上的投影，生成掩码水平特征 H_b ；
计算所述任一帧待检测图像的灰度图像的边缘强度图 E_{a11} ；
对所述任一帧待检测图像的灰度图像的边缘强度图 E_{a11} 进行二值化处理，得到边缘二值化图像 E ；
利用所述边缘二值化图像 E 在水平方向上的投影，生成边缘水平特征 H_e ；
根据所述掩码水平特征 H_b 与所述边缘水平特征 H_e ，计算图像分割特征 H_s ；
在 $i \in [0, H-1]$ 的范围内，对所述图像分割特征 $H_s[i]$ 进行遍历；
在所述遍历的过程中，当所述图像分割特征 $H_s[i]$ 符合水平分割条件时，记录每一水平方向上的位置 $[i, i+n]$ 为文字行区域，并将所述位置 $[i, i+n]$ 存入预设列表 POS_{text} ；
根据所述预设列表 POS_{text} 中符合有效高度条件的位置 $[i, i+n]$ ，计算分割区域 $SEG(0, i, W, n)$ ；

将所有分割区域SEG(0, i, W, n)中n值的平均值确定为所述任一帧待检测图像中文字的字体高度;

其中,所述水平分割条件为同时满足第一条件、第二条件与第三条件;第一条件为: $i=0$ 或 $H_s[i-1]=1$;第二条件为:对于每一个j值, $H_s[i+j]$ 均等于0, $0 \leq j \leq n$,j与n为自然数;第三条件为: $H_s[i+n+1]=1$ 或 $i=H-1$;

所述有效高度条件为: $n \geq M_h$;其中, $M_h=H/(N+1)$,N为所述预设列表POS_{text}中的位置[i, i+n]的数量;

$$\text{其中, } H_b[i]=\begin{cases} 0, & \text{numb}[i] \geq TH1 \\ 1, & \text{numb}[i] < TH1 \end{cases}, \quad E(x,y)=\begin{cases} 0, & E_{\text{all}}(x,y) \leq TK4 \\ 1, & E_{\text{all}}(x,y) > TK4 \end{cases}, \quad H_e[i]=\begin{cases} 0, & \text{nume}[i] \geq TH2 \\ 1, & \text{nume}[i] < TH2 \end{cases},$$

$H_s[i]=H_b[i] \mid H_e[i]$;numb[i]为所述掩码图像第i+1行中像素值不等于所述背景像素值的像素个数;TH1为预设的第一特征阈值;nume[i]为所述边缘二值化图像第i+1行中像素值不等于0的像素个数,TH2为预设的第二特征阈值;TK4为第四分割阈值; $i \in [0, H-1]$,H为所述待检测图像的高度,W为所述待检测图像的宽度。

21.根据权利要求12所述的系统,其特征在于,所述候选区域跟踪单元包括:

第三获取单元,用于当所述标题候选区域为首次跟踪时,获取当前帧待检测图像在跟踪区域内的图像,作为候选跟踪图像;所述跟踪区域与所述标题候选区域相对应;

第四计算单元,用于计算所述候选跟踪图像的参考颜色直方图与参考二值化图像;

第四获取单元,用于当所述标题候选区域为非首次跟踪时,获取当前帧待检测图像在跟踪区域内的图像,作为当前跟踪图像;

第五计算单元,用于计算所述当前跟踪图像的当前颜色直方图与当前二值化图像;

第七确定单元,用于根据所述当前颜色直方图与所述参考颜色直方图的差异以及所述当前二值化图像与所述参考二值化图像的差异,确定当前帧待检测图像是否满足预设条件。

22.根据权利要求21所述的系统,其特征在于,所述候选区域跟踪单元还用于:

在跟踪检测过程中,统计不满足预设条件的待检测图像的数量,记为丢失帧数;

当所述丢失帧数大于预设丢失阈值时,结束对所述至少一帧待检测图像依次进行跟踪检测的过程。

标题区域检测方法及系统

技术领域

[0001] 本发明涉及信息处理领域,更具体的说,涉及标题区域检测方法及系统。

背景技术

[0002] 新闻视频中蕴含着大量的最新资讯信息,现有的视频网站或新闻类的应用通常会对每日播出的整条新闻视频进行拆分、上线,以供用户对于其中感兴趣的每条新闻进行点击观看。由于新闻标题在新闻视频的拆分过程中是具有重大意义的语义线索,所以,需要先对新闻视频中的新闻标题进行定位、识别。但是,在不同电视台播出的新闻视频中,或者,在不同类型的新闻视频中,新闻标题的展示风格千差万别,而且,在新闻标题的附近位置常会长时间出现风格内容极为相似的滚动字幕,而滚动字幕与新闻标题的位置信息、时域信息、颜色信息都比较接近,使得新闻标题与滚动字幕之间很难区分,从而为新闻标题的检测带来了很大的难度。

[0003] 因此,目前迫切需要一种能够在新闻视频中能够准确检测新闻标题的技术方案。

发明内容

[0004] 有鉴于此,本发明提供了一种标题区域检测方法及系统,以解决目前新闻视频中新闻标题难以准确检测的技术问题。

[0005] 为实现上述目的,本发明提供如下技术方案:

[0006] 一种标题区域检测方法,包括:

[0007] 获取视频对象中的至少一帧候选视频图像,所述至少一帧候选视频图像采用相同的候选区域;

[0008] 根据至少一帧待检测图像中文字的字体高度,确定所述候选区域是否为标题候选区域;其中,所述至少一帧待检测图像分别为所述至少一帧候选视频图像的候选区域内的图像;

[0009] 当确定所述候选区域为标题候选区域时,按照所述至少一帧待检测视频图像的时域顺序,对所述至少一帧待检测图像依次进行跟踪检测;

[0010] 在所述跟踪检测的过程中,统计满足预设条件的待检测图像的数量,记为跟踪帧数;

[0011] 在所述跟踪检测结束后,根据所述跟踪帧数,确定所述候选区域是否为标题区域。

[0012] 一种标题区域检测系统,包括:

[0013] 候选图像获取单元,用于获取视频对象中的至少一帧候选视频图像,所述至少一帧候选视频图像采用相同的候选区域;

[0014] 字体高度过滤单元,用于根据至少一帧待检测图像中文字的字体高度,确定所述候选区域是否为标题候选区域;其中,所述至少一帧待检测图像分别为所述至少一帧候选视频图像的候选区域内的图像;

[0015] 候选区域跟踪单元,用于当确定所述候选区域为标题候选区域时,按照所述至少

一帧待检测视频图像的时域顺序,对所述至少一帧待检测图像依次进行跟踪检测;

[0016] 所述候选区域跟踪单元,还用于在所述跟踪检测的过程中,统计满足预设条件的待检测图像的数量,记为跟踪帧数;

[0017] 标题区域确定单元,用于在所述跟踪检测结束后,根据所述跟踪帧数,确定所述候选区域是否为标题区域。

[0018] 从上述的技术方案可以看出,本发明提供的标题区域检测方法及系统,获取视频对象中采用相同的候选区域的至少一帧候选视频图像,根据候选区域中的至少一帧待检测图像中文字的字体高度,确定候选区域是否为标题候选区域;当确定候选区域为标题候选区域时,按照至少一帧待检测视频图像的时域顺序,对候选区域进行跟踪检测;在跟踪检测的过程中,统计满足预设条件的待检测图像的跟踪帧数;最后根据跟踪帧数,确定候选区域是否为标题区域。本发明基于新闻标题的文字的字体高度特征,能够滤除滚动字幕的干扰,准确判断出标题候选区域,并通过对标题候选区域的跟踪检测,能够最终确定出标题候选区域是否为标题区域,提高标题区域的检测精度。

附图说明

[0019] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例或现有技术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,还可以根据提供的附图获得其他的附图。

[0020] 图1为本申请实施例提供的标题区域检测方法的一种流程图;

[0021] 图2为本申请实施例提供的候选视频图像获取方法的一种流程图;

[0022] 图3为本申请实施例提供的字体高度过滤方法的一种流程图;

[0023] 图4为本申请实施例提供的标题候选区域跟踪方法的一种流程图;

[0024] 图5为本申请实施例提供的标题区域检测系统的一种结构示意图。

具体实施方式

[0025] 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,都属于本发明保护的范围。

[0026] 请参阅图1,图1为本申请实施例提供的标题区域检测方法的一种流程图。

[0027] 如图1所示,所述方法包括:

[0028] S101:获取视频对象中的至少一帧候选视频图像,所述至少一帧候选视频图像采用相同的候选区域。

[0029] 在大部分新闻视频中,新闻标题通常位于视频图像中的特定区域,例如,位于视频图像底部的某一区域,所以,直接从视频图像的特定区域去检测标题,能够减小计算量,从而更有效率地确定出标题区域。

[0030] 所述候选区域即为视频图像中新闻标题可能存在且具有文字内容的区域,如果视频图像中新闻标题可能存在的区域内并不存在文字内容,则该区域也不能被称为候选区

域。所以,为了提高新闻标题的检测效率,本实施例只对具有候选区域的视频图像帧进行检测,对于不具有候选区域的视频图像帧,直接作丢弃处理,由此能够提高标题区域的检测效率。

[0031] S102:根据至少一帧待检测图像中文字的字体高度,确定所述候选区域是否为标题候选区域。

[0032] 其中,所述至少一帧待检测图像分别为所述至少一帧候选视频图像的候选区域内的图像。

[0033] 新闻标题的字体高度与滚动字幕的字体高度是不同的,根据两者的差异,设定相应的预设高度阈值,能够确定待检测图像中的文字是新闻标题还是滚动字幕。当根据待检测图像中文字的字体高度,确定出该待检测图像中的文字为新闻标题时,将候选区域确定为标题候选区域。

[0034] 具体地,可以根据一帧待检测图像中文字的字体高度来进行确定,也可以根据连续N帧待检测图像中文字的字体高度来进行确定。

[0035] 在一示例中,当一帧待检测图像中文字的字体高度大于预设高度阈值时,或者,连续N帧待检测图像中文字的字体高度的平均值大于预设高度阈值时,则确定所述候选区域为标题候选区域;否则,确定所述候选区域为非标题候选区域。

[0036] 在该示例中,新闻标题的字体高度大于预设高度阈值,滚动字幕的字体高度小于或等于预设高度阈值。

[0037] S103:当确定所述候选区域为标题候选区域时,按照所述至少一帧待检测视频图像的时域顺序,对所述至少一帧待检测图像依次进行跟踪检测。

[0038] 候选视频图像在视频对象中具有时域顺序,相应的,所述至少一帧待检测图像也具有同样的时域顺序,按照候选视频图像的时域顺序,即可对所述至少一帧待检测图像依次进行跟踪检测。

[0039] 由于所述至少一帧待检测图像分别为所述至少一帧候选视频图像的候选区域内的图像,所以,对所述至少一帧待检测图像依次进行跟踪检测,实际上就是在对所述标题候选区域进行跟踪检测。

[0040] S104:在所述跟踪检测的过程中,统计满足预设条件的待检测图像的数量,记为跟踪帧数。

[0041] 在一示例中,预设跟踪帧数的初始值为0,当当前帧待检测图像满足预设条件时,跟踪帧数加1。

[0042] S105:在所述跟踪检测结束后,根据所述跟踪帧数,确定所述候选区域是否为标题区域。

[0043] 在一示例中,在跟踪检测结束后,比较跟踪帧数tracking_num与预先设定的阈值 $Th_{tracking_num}$ 的大小关系,如果 $tracking_num \geq Th_{tracking_num}$,则判定标题候选区域为标题区域,否则,判定标题候选区域为非标题区域(如,滚动字幕区域)。

[0044] 本实施例提供的标题区域检测方法,基于新闻标题的文字的高度属性,能够滤除滚动字幕的干扰,准确判断出视频图像中的检测区域是否为标题区域;并且,针对不同风格的新闻视频中新闻标题文字的高度属性,本发明可以通过设定相应的预设高度阈值,以适用于各种风格的新闻视频中标题区域的准确定位,具有普遍适用性。

[0045] 请参阅图2,图2为本申请实施例提供的候选视频图像获取方法的一种流程图。

[0046] 本实施例提供的候选视频图像获取方法主要用于实现前述实施例中的步骤S101,具体对任一帧视频图像中的候选区域进行检测,进而确定出具有候选区域的至少一帧候选视频图像。

[0047] 如图2所示,所述候选视频图像获取方法包括:

[0048] S201:获取视频对象中的任一帧视频图像。

[0049] 在视频对象中,可以按照视频图像帧的时域顺序,依次获取任一帧视频图像。

[0050] S202:获取所述任一帧图像的标题潜在区域内的图像,作为潜在图像。

[0051] 所述标题潜在区域为用户指定的标题所在区域,具体地,可以是用户根据历史经验指定的标题所在区域。例如,在一些视频对象中,标题通常在视频图像的底部,在另一些视频对象中,标题通常在视频图像的右侧,那么,用户就可以指定视频图像中的相应区域为标题潜在区域。

[0052] 以标题潜在区域为视频图像的底部区域为例,假设视频图像的宽度为WW,高度为HH,则底部区域在视频帧中的位置为Rect(x,y,w,h),其中, $x=0$, $y=HH \times \text{cut_ratio}$, $w=WW$, $h=HH \times (1-\text{cut_ratio})$; (x,y)构成底部区域的起点坐标,w为底部区域的宽度,h为底部区域的高度,cut_ratio为预设参数.cut_ratio可以由历史经验得到,也可以是人为设置,用于表示视频图像顶端到底部区域之间的高度。

[0053] S203:计算所述潜在图像的边缘强度图。

[0054] 根据所述潜在图像,将所述潜在图像由RGB颜色空间转换为灰度或者任意亮度色彩分离空间,例如YUV颜色空间、HSV颜色空间、HSL颜色空间、LAB颜色空间等,以生成潜在图像的灰度图像U。

[0055] 对于从RGB颜色空间转换到灰度空间,转换公式为:

[0056] $U=R*0.299+G*0.587+B*0.114$;

[0057] 对于从RGB颜色空间转换到亮度色彩分离空间,以HSL颜色空间举例,转换公式为:

[0058] $U=(\max(R,G,B)+\min(R,G,B))/2$;

[0059] 对于灰度图像U,仍以Sobel算子为例。利用水平方向边缘梯度算子和垂直方向边缘梯度算子,与灰度图像进行卷积,获得水平边缘图 E_{uh} 和垂直边缘图 E_{uv} ,最终计算边缘强度图 E_{uall} ,即对于边缘强度图上任意一点 $E_{uall}(x,y)$,满足公式 $E_{uall}(x,y)=\sqrt{E_{uv}(x,y)^2+E_{uh}(x,y)^2}$ 。

[0060] S204:根据所述潜在图像的边缘强度图在水平方向上投影得到的水平直方图,确定候选区域的上下边界。

[0061] 首先,利用第一分割阈值TK1,对所述潜在图像的边缘强度图 E_{uall} 进行二值化处理,得到潜在二值化图像 E_u 。其中,潜在二值化图像 E_u 中像素点(x,y)的像素值 $E_u(x,y)$ 为:

[0062]
$$E_u(x,y)=\begin{cases} 0, & E_{uall}(x,y) \leq TK1 \\ 1, & E_{uall}(x,y) > TK1 \end{cases}$$
;

[0063] 第一分割阈值TK1是预先设定的。

[0064] 然后,对潜在二值化图像 E_u 进行水平方向的投影,统计每一行中符合预定条件的像素的数量Numedge,如果Numedge>Thnum,则令水平直方图 $H_u[i]=1$,否则,水平直方图 $H_u[i]=$

0。其中, Th_{num} 为针对Numedge设定的阈值。

[0065] 预定条件:该像素以及上下相邻像素中,存在至少一个像素为1的值。

[0066] 其中,该像素满足预定条件时,该像素的边缘值视为1,同时,统计该像素左右连续的像素边缘值为1,且,连续长度大于预设长度阈值 Th_{len} 的像素的总个数。

[0067] 在一示例中,在根据所述潜在二值化图像,获取所述潜在二值化图像在水平方向上的水平直方图的过程中,还可以对潜在二值化图像进行加强处理。

[0068] 例如,针对所述潜在图像的RGB中每个通道(R、G、B),分别获取相应的边缘强度图 E_r, E_g, E_b 。利用第二分割阈值TK2,对边缘强度图 E_r, E_g, E_b 分别进行二值化处理,得到相应的边缘二值化图 E_{2r}, E_{2g}, E_{2b} 。

[0069] 如果 $E_r(x, y) > TK2$,则令边缘二值化图 $E_{2r}(x, y) = 1$,否则, $E_{2r}(x, y) = 0$ 。

[0070] 如果 $E_g(x, y) > TK2$,则令边缘二值化图 $E_{2g}(x, y) = 1$,否则, $E_{2g}(x, y) = 0$ 。

[0071] 如果 $E_b(x, y) > TK2$,则令边缘二值化图 $E_{2b}(x, y) = 1$,否则, $E_{2b}(x, y) = 0$ 。

[0072] 通常情况下,第二分割阈值TK2小于第一分割阈值TK1。

[0073] 对于潜在二值化图像进行加强,可以选择某个通道进行加强,也可以选择全部三个通道进行加强。以选择全部三个通道进行加强为例,使 $E_u(x, y) = E_u(x, y) | E_{2r}(x, y) | E_{2g}(x, y) | E_{2b}(x, y)$,其中,“|”为“求或运算符”。通过对潜在二值化图像 E_u 的加强,能够防止图像中的字幕区域出现渐变导致检测失败的情况。

[0074] 对于水平直方图 H_u ,进行遍历,获取图像边缘为直线的相邻行的行间距。其中,当 $H_u[i] = 1$ 时,表明第 $i+1$ 行的图像边缘为直线。例如,当 $H_u[i] = 1$ 且 $H_u[i+1] = 1$ 时,表明第 $i+1$ 行与第 $i+2$ 行的图像边缘为直线,则获取第 $i+1$ 行与第 $i+2$ 行之间的行间距。

[0075] 在水平直方图 H_u 中,如果获取的相邻行的行间距大于行间距阈值 Th_{row} ,则将这两行之间的区域确定为第一候选区域;如果在潜在图像中不存在行间距大于行间距阈值 Th_{row} 的相邻行时,获取下一帧视频图像的潜在图像,重新执行上述步骤。

[0076] 其中,所述第一候选区域表征候选区域的上下边界。

[0077] S205:根据所述潜在图像的边缘强度图在垂直方向上投影得到的垂直直方图,确定候选区域的左右边界。

[0078] 在一示例中,可以根据所述潜在二值化图像在所述第一候选区域中(即所述上下边界范围内)的部分图像,获取所述部分图像在垂直方向上的垂直直方图 V_u 。

[0079] 在垂直直方图 V_u 中,对于任一系列,如果该列的边缘像素为1的数量大于像素数目阈值 Th_{uv} 时,则 $V_u[i] = 1$,否则 $V_u[i] = 0$,并强制设置 $V_u[0] = 1$ 且 $V_u[W-1] = 1$ 。在垂直直方图 V_u 中,将满足“ $V_u[i] = 1 \& \& V_u[j] = 1 \& \& V_u[k] = 0 \& \& \arg\max(i-j)$ ”这一条件的区域确定为第二候选区域。其中, i, j 均为列号, $k \in (i, j)$, $\arg\max(i-j)$ 表征使 $i-j$ 的值最大。

[0080] 其中,所述第二候选区域表征候选区域的左右边界。

[0081] S206:根据所述上下边界与所述左右边界,确定候选区域。

[0082] 在一示例中,候选区域为:在标题潜在区域中,由所述上下边界与所述左右边界组成的区域。

[0083] 在另一示例中,该步骤S206可包括:

[0084] 精细寻找第二候选区域所表征候选区域的左右边界,以一定长度的滑动窗口扫描第二候选区域的原图,计算每一个窗口内的颜色直方图,同时统计该窗口内颜色直方图中

非0位的个数 $\text{num}_{\text{color}}$,寻找单色区域或者颜色复杂的背景区域的位置,即符合 $\text{num}_{\text{color}} < \text{Th}_{\text{color}1}$ 或 $\text{num}_{\text{color}} > \text{Th}_{\text{color}2}$ 的窗口位置,将符合该条件的窗口的中心位置,作为新的垂直方向边界。其中, $\text{Th}_{\text{color}1}$ 为针对 $\text{num}_{\text{color}}$ 设定的上限阈值, $\text{Th}_{\text{color}2}$ 为针对 $\text{num}_{\text{color}}$ 设定的下限阈值, $\text{Th}_{\text{color}1}$ 与 $\text{Th}_{\text{color}2}$ 用于确定新的垂直方向的边界,进而确定出候选区域。

[0085] S207:当确定出的候选区域符合预设条件时,将所述任一帧视频图像确定为候选视频图像。

[0086] 预设条件可以是候选区域的起点的位置位于预设的图像范围中,也可以是候选区域的高度在预设范围中,还可以根据具体情况具体限定。如果候选区域符合预设条件,则将所述任一帧视频图像确定为候选视频图像;否则,判定所述任一帧视频图像不具有候选区域,将所述任一帧视频图像确定为非候选视频图像。

[0087] 本实施例提供的标题区域检测方法,获取视频对象中的任一帧视频图像;获取所述任一帧图像的标题潜在区域内的图像,作为潜在图像;根据潜在图像,确定候选区域;当确定出的候选区域符合预设条件时,将所述任一帧视频图像确定为候选视频图像,为标题区域的检测提供了充分有效的候选视频图像,为后续的标题区域检测过程,节省了大量的运算量,从而能够更有效率地确定出标题区域。

[0088] 请参阅图3,图3为本申请实施例提供的字体高度过滤方法的一种流程图。

[0089] 本实施例提供的字体高度过滤方法用于实现前述实施例中的步骤S102,根据至少一帧待检测图像中文字的字体高度,确定所述候选区域是否为标题候选区域。

[0090] 如图3所示,所述字体高度过滤方法包括:

[0091] S301:根据任一帧待检测图像,计算所述任一帧待检测图像的文本掩码。

[0092] 其中,该步骤S301具体包括:

[0093] S3011:根据任一帧待检测图像,生成所述任一帧待检测图像的灰度图像。

[0094] 根据所述待检测图像,将所述待分割图像由RGB颜色空间转换为灰度或者任意亮度色彩分离空间,例如YUV颜色空间、HSV颜色空间、HSL颜色空间、LAB颜色空间等,以生成所述待检测图像的灰度图像I。

[0095] 对于从RGB颜色空间转换到灰度空间,转换公式为:

[0096] $I = R * 0.299 + G * 0.587 + B * 0.114$;

[0097] 对于从RGB颜色空间转换到亮度色彩分离空间,以HSL颜色空间举例,转换公式为:

[0098] $I = (\max(R, G, B) + \min(R, G, B)) / 2$;

[0099] S3012:计算所述任一帧待检测图像的灰度图像的分割阈值。

[0100] 其中,所述任一帧待检测图像的灰度图像的分割阈值作为第三分割阈值。

[0101] 在一示例中,可以采用OTSU(最大类间方差法)方法计算所述第三分割阈值TK3。第三分割阈值是采用OTSU方法针对待检测图像的灰度图像计算出来的最佳阈值。

[0102] S3013:利用所述第三分割阈值,对所述任一帧待检测图像的灰度图像进行二值化处理,得到灰度二值化图像。

[0103] 灰度二值化图像B中像素点(x,y)的像素值 $B(x,y)$ 为:

[0104]
$$B(x,y) = \begin{cases} 0, & I(x,y) < TK3 \\ 255, & I(x,y) \geq TK3 \end{cases}$$

- [0105] 其中, $I(x, y)$ 为灰度图像 I 中像素点 (x, y) 的像素值, $TK3$ 为第三分割阈值。
- [0106] S3014: 根据所述灰度二值化图像, 计算所述灰度二值化图像的平均像素值;
- [0107] 其中, 该步骤 S3014 具体包括:
- [0108] (1)、水平扫描所述灰度二值化图像的任一行像素;
- [0109] (2)、将所述任一行中的首个像素值记为所述任一行的参考像素值;
- [0110] (3)、统计所述任一行中像素值不等于所述任一行的参考像素值的像素个数, 记为第一个数;
- [0111] (4)、当所述第一个数小于预设个数阈值时, 确定所述任一行符合第一预设条件;
- [0112] (5)、统计符合第一预设条件的行数以及符合第一预设条件的所有行的参考像素值;
- [0113] (6)、垂直扫描所述灰度二值化图像的任一列像素;
- [0114] (7)、将所述任一列中的首个像素值记为所述任一列的参考像素值;
- [0115] (8)、统计所述任一列中像素值不等于所述任一列的参考像素值的像素个数, 记为第二个数;
- [0116] (9)、当所述第二个数小于预设个数阈值时, 确定所述任一列符合第二预设条件;
- [0117] (10)、统计符合第二预设条件的列数以及符合第二预设条件的所有列的参考像素值;
- [0118] (11)、计算符合第一预设条件的所有行的参考像素值与符合第二预设条件的所有列的参考像素值的像素值总和, 以及, 符合第一预设条件的行数与符合第二预设条件的列数的行列总数;
- [0119] (12)、将所述像素值总和与所述行列总数的比值确定为所述灰度二值化图像的平均像素值。
- [0120] 在一示例中, 上述步骤 (1) ~ (12) 具体实现如下:
- [0121] 水平扫描灰度二值化图像 B 的每行像素, 取本行第一个像素值为 $color$, 统计本行像素值不等于 $color$ 的像素的个数 num , 如果 $num < h1$, 则执行 $back_color = back_color + color$ 以及 $count = count + 1$ 的累加过程。
- [0122] 垂直扫描图像 B 的每列像素, 取本列第一个像素值为 $color$, 统计本列像素值不等于 $color$ 的像素的个数 num , 如果 $num < h1$, 则执行 $back_color = back_color + color$ 以及 $count = count + 1$ 的累加过程。
- [0123] 计算平均像素值 $back_color_avr$, 其中, $back_color_avr = back_color / count$ 。
- [0124] 在该示例中, $h1$ 为针对 num 预设的个数阈值, $back_color$ 的初始值为 0, $count$ 的初始值为 0。
- [0125] S3015: 根据所述平均像素值, 确定背景像素值;
- [0126] 在一示例中, 当所述平均像素值小于预设像素值时, 确定背景像素值为第一像素值; 当所述平均像素值不小于预设像素值时, 确定背景像素值为第二像素值。其中, 所述第一像素值与所述第二像素值不同。
- [0127] 例如, 预设像素值为灰度二值化图像 B 中的像素值 0 和像素值 255 的中位数 128, 第一像素值为灰度二值化图像 B 中的像素值 0, 第二像素值为灰度二值化图像 B 中的另一像素值 255。如果最后得到平均像素值 $back_color_avr$ 满足 $back_color_avr < 128$, 则确定灰度二

值化图像B的背景像素值back_color_f=0,否则,确定灰度二值化图像B的背景像素值back_color_f=255。

[0128] S3016:根据所述灰度二值化图像与背景像素值,生成与所述任一帧待检测图像相对应的文本掩码。

[0129] 其中,所述文本掩码为所述任一帧待检测图像中文本区域的掩码。

[0130] 其中,步骤S3016具体包括:

[0131] (1)、将所述灰度二值化图像中像素值与所述背景像素值相同的区域的确定为背景区域;

[0132] (2)、将所述灰度二值化图像中像素值与所述背景像素值不同的区域确定为文字区域;

[0133] (3)、根据所述背景区域与所述文字区域,生成与所述待检测图像相对应的文本掩码。

[0134] 在一示例中,将灰度二值化图像B中像素值等于back_color_f的区域确定为背景区域,将背景区域之外区域确定为文字区域。

[0135] 在另一示例中,步骤S3016还可以包括:

[0136] 将所述背景区域的像素值设置为第三像素值,将所述文字区域的像素值设置为第四像素值,进而生成由第三像素值和第四像素值组成的文本掩码。

[0137] 其中,第三像素值与第四像素值不同。例如,第三像素值为0,第四像素值为255。

[0138] S302:根据所述文本掩码,确定所述任一帧待检测图像中文字的字体高度。

[0139] 其中,步骤S302包括:

[0140] S3021:利用所述文本掩码在水平方向上的投影,生成掩码水平特征H_b;

[0141] 根据所述文本掩码在水平方向上的投影,统计文本掩码每行中像素值不等于灰度二值化图像的背景像素值back_color_f的像素的个数numb,根据numb与第一特征阈值TH1的大小关系确定掩码水平特征H_b,具体如下:

$$[0142] \quad H_b[i] = \begin{cases} 0, & \text{numb}[i] \geq TH1 \\ 1, & \text{numb}[i] < TH1 \end{cases};$$

[0143] 其中, $i \in [0, H-1]$,H为待检测图像的高度(也即候选区域的高度),numb[i]为所述文本掩码第i+1行中像素值不等于所述背景像素值的像素个数。

[0144] S3022:计算所述任一帧待检测图像的灰度图像的边缘强度图E_{a11};

[0145] 对于灰度图像,提取图像边缘特征的方法有多种,如索贝尔(Sobel)边缘检测算子,Canny边缘检测算子等,均可用于计算所述待检测图像的灰度图像的边缘强度图。

[0146] 在一示例中,以Sobel算子为例。利用水平方向边缘梯度算子和垂直方向边缘梯度算子,与灰度图像进行卷积,获得水平边缘图E_h和垂直边缘图E_v,最终计算出边缘强度图E_{a11},其中,对于边缘强度图E_{a11}上任意一点(x,y)的值E_{a11}(x,y),满足公式E_{a11}(x,y) = sqrt(E_v(x,y)²+E_h(x,y)²)。

[0147] 在其他示例中,也可以采用其他算子来计算获得水平边缘图E_h和垂直边缘图E_v,在此不再赘述。

[0148] S3023:对所述任一帧待检测图像的灰度图像的边缘强度图E_{a11}进行二值化处理,

得到边缘二值化图像E;

[0149] 边缘二值化图像E中像素点(x,y)的像素值E(x,y)为:

$$[0150] \quad E(x,y) = \begin{cases} 0, & E_{all}(x,y) \leq TK4 \\ 1, & E_{all}(x,y) > TK4 \end{cases};$$

[0151] 其中,TK4为预设的第四分割阈值。

[0152] S3024:利用所述边缘二值化图像E在水平方向上的投影,生成边缘水平特征H_e;

[0153] 根据所述边缘二值化图像E在水平方向上的投影,统计边缘二值化图像E每行中像素值不等于0的像素的个数nume,根据nume与第二特征阈值TH2的大小关系确定边缘水平特征H_e,具体如下:

$$[0154] \quad H_e[i] = \begin{cases} 0, & nume[i] \geq TH2 \\ 1, & nume[i] < TH2 \end{cases};$$

[0155] 其中,i ∈ [0,H-1],H为待检测图像的高度(也即候选区域的高度),nume[i]为所述边缘二值化图像第i+1行中像素值不等于0的像素个数。

[0156] S3025:根据所述掩码水平特征H_b与所述边缘水平特征H_e,计算图像分割特征H_s;

[0157] 将所述掩码水平特征H_b与所述边缘水平特征H_e进行融合,得到图像分割特征H_s,其中,H_s[i] = H_b[i] | H_e[i],“|”为求或运算符。

[0158] 图像分割特征H_s是掩码水平特征H_b与边缘水平特征H_e的融合结果。融合的目的旨在增强拆分的通用性,如果仅使用边缘水平特征H_e进行文字图像的分割,会出现由于列间距过密导致H_e[i]全部为0,进而导致无法进行分割的情况;如果仅使用掩码水平特征H_b[i]进行文字图像的分割,则会出现由于图像的压缩效应导致标题与背景不能完整分离的效果,同时对于多颜色字体也会造成影响。所以,本实施例采取由掩码水平特征H_b与边缘水平特征H_e的相融合得到的图像分割特征H_s,从所述文本掩码中分割文字图像,能够得到更好的文字图像分割效果。

[0159] S3026:在i ∈ [0,H-1]的范围内,对所述图像分割特征H_s[i]进行遍历;

[0160] S3027:在所述遍历的过程中,当所述图像分割特征H_s[i]符合水平分割条件时,记录每一水平方向上的位置[i,i+n]为文字行区域,并将所述位置[i,i+n]存入预设列表POS_{text};

[0161] 其中,所述水平分割条件为同时满足第一条条件、第二条条件与第三条条件;第一条条件为:i=0或H_s[i-1]=1;第二条条件为:对于每一个j值,H_s[i+j]均等于0,0 ≤ j ≤ n,j与n为自然数;第三条条件为:H_s[i+n+1]=1或i=H-1;

[0162] S3028:根据所述预设列表POS_{text}中符合有效高度条件的位置[i,i+n],计算分割区域SEG(0,i,W,n);

[0163] 所述有效高度条件为:n ≥ M_h;其中,M_h = H/(N+1),N为所述预设列表POS_{text}中的位置[i,i+n]的数量。

[0164] 对于预设列表POS_{text}中符合有效高度条件的位置[i,i+n],计算其对应的分割区域SEG(0,i,W,n),W为所述待检测图像的宽度,n的值即为位置[i,i+n]所对应的文字行的行高值。进一步讲,分割区域SEG(0,i,W,n)与符合有效高度条件的位置[i,i+n]之间具有一一对应关系。当符合有效高度条件的位置[i,i+n]的有多个时,相应地,分割区域SEG(0,i,W,

n) 也有多个。

[0165] 需要说明的是,在一分割区域SEG(x,y,w,h)中包括x、y、w与h这4个参数,其中,x与y构成分割区域的起点坐标(x,y),w为分割区域的宽度,h为分割区域的高度。

[0166] 在其他示例中,根据不同的需求,也可以单独采用边缘垂直特征 H_e 或掩码垂直特征 H_b 进行文字图像的分割,在此不再赘述。

[0167] S3029:将所有分割区域SEG(0,i,W,n)中n值的平均值确定为所述任一帧待检测图像中文字的字体高度。

[0168] 根据预设列表POS_{text}中符合有效高度条件的位置[i,i+n],可能会计算得到多个分割区域SEG(0,i,W,n),每个分割区域SEG(0,i,W,n)中的n值即为分割区域SEG(0,i,W,n)所对应文字行的行高值,计算得到所有分割区域SEG(0,i,W,n)中的n值的平均值,便可作为所述任一帧待检测图像中文字的字体高度。

[0169] S303:根据全部待检测图像中文字的字体高度,计算全部待检测图像的文字字体平均高度。

[0170] S304:根据所述文字字体平均高度,确定所述候选区域是否为标题候选区域。

[0171] 本发明中的“候选区域”是指初步确定的新闻标题可能存在的区域,本实施例中的“标题候选区域”是根据文字的字体高度进一步确定的新闻标题可能存在的区域。

[0172] 本实施例提供的标题区域检测方法,根据任一帧待检测图像,计算所述任一帧待检测图像的文本掩码;根据所述文本掩码,确定所述任一帧待检测图像中文字的字体高度;根据全部待检测图像中文字的字体高度,计算全部待检测图像的文字字体平均高度;根据所述文字字体平均高度,确定所述候选区域是否为标题候选区域。本发明根据待检测图像中文字的字体高度,实现了对非标题候选区域的过滤,从而更加准确地为标题区域的检测提供有力依据。

[0173] 请参阅图4,图4为本申请实施例提供的标题候选区域跟踪方法的一种流程图。

[0174] 本实施例提供的标题候选区域跟踪方法主要用于实现前述实施例中的步骤S103,当确定所述候选区域为标题候选区域时,按照所述至少一帧待检测视频图像的时域顺序,对所述至少一帧待检测图像依次进行跟踪检测。

[0175] 如图4所示,所述标题候选区域跟踪方法包括:

[0176] S401:当所述标题候选区域为首次跟踪时,获取当前帧待检测图像在跟踪区域内的图像,作为候选跟踪图像;

[0177] 其中,所述跟踪区域与所述标题候选区域相对应。

[0178] 在一示例中,所述跟踪区域即为前述实施例中的所述标题候选区域。对标题候选区域进行跟踪,是为了进一步确定该标题候选区域是否为标题区域,提高标题区域检测的精准度。

[0179] 在另一示例中,假设标题候选区域为CandidateRect(CandidateRect.x,CandidateRect.y,CandidateRect.w,CandidateRect.h),其中,(CandidateRect.x,CandidateRect.y)为标题候选区域的起点坐标,CandidateRect.w为标题候选区域的宽度,CandidateRect.h为标题候选区域的高度。据此,设定跟踪区域track(track.x,track.y,track.w,track.h)满足:

[0180] $track.x = CandidateRect.x + CandidateRect.w \times Xratio1;$

[0181] $track.y = CandidateRect.y + CandidateRect.h \times Yratio1;$

[0182] $track.w = CandidateRect.w \times Xratio2;$

[0183] $track.h = CandidateRect.h \times Yratio2;$

[0184] 其中, $Xratio1, Xratio2, Yratio1, Yratio2$ 均为预先设定的参数。

[0185] S402: 计算所述候选跟踪图像的参考颜色直方图与参考二值化图像;

[0186] 首先, 根据所述候选跟踪图像, 将所述候选跟踪图像由RGB颜色空间转换为灰度或者任意亮度色彩分离空间, 例如YUV颜色空间、HSV颜色空间、HSL颜色空间、LAB颜色空间等, 以生成所述候选跟踪图像的灰度图像G1。

[0187] 对于从RGB颜色空间转换到灰度空间, 转换公式为:

[0188] $G1 = R * 0.299 + G * 0.587 + B * 0.114;$

[0189] 对于从RGB颜色空间转换到亮度色彩分离空间, 以HSL颜色空间举例, 转换公式为:

[0190] $G1 = (\max(R, G, B) + \min(R, G, B)) / 2;$

[0191] 计算所述候选跟踪图像的灰度图像的分割阈值。

[0192] 其中, 所述候选跟踪图像的灰度图像的分割阈值作为第五分割阈值。

[0193] 在一示例中, 采用OTSU (最大类间方差法) 方法计算所述第五分割阈值TK5。第五分割阈值为采用OTSU方法针对所述候选跟踪图像的灰度图像计算出来的最佳阈值。

[0194] 然后, 利用所述第五分割阈值TK5, 对所述候选跟踪图像的灰度图像进行二值化处理, 得到参考二值化图像。

[0195] 参考二值化图像 B_{ref} 中像素点 (x, y) 的像素值 $B_{ref}(x, y)$ 为:

[0196]
$$B_{ref}(x, y) = \begin{cases} 0, & G1(x, y) < TK5 \\ 255, & G1(x, y) \geq TK5 \end{cases};$$

[0197] 其中, $G1(x, y)$ 为灰度图像G1中像素点 (x, y) 的像素值, TK5为第五分割阈值。

[0198] 同时, 计算所述候选跟踪图像的参考颜色直方图, 所述候选跟踪图像的参考颜色直方图 H_{ref} 即为候选跟踪图像的颜色直方图。

[0199] S403: 当所述标题候选区域为非首次跟踪时, 获取当前帧待检测图像在跟踪区域内的图像, 作为当前跟踪图像;

[0200] S404: 计算所述当前跟踪图像的当前颜色直方图与当前二值化图像;

[0201] 根据所述当前跟踪图像, 获取所述当前跟踪图像的灰度图像。

[0202] 根据所述当前跟踪图像, 将所述当前跟踪图像由RGB颜色空间转换为灰度或者任意亮度色彩分离空间, 例如YUV颜色空间、HSV颜色空间、HSL颜色空间、LAB颜色空间等, 以生成所述当前跟踪图像的灰度图像G2。

[0203] 对于从RGB颜色空间转换到灰度空间, 转换公式为:

[0204] $G2 = R * 0.299 + G * 0.587 + B * 0.114;$

[0205] 对于从RGB颜色空间转换到亮度色彩分离空间, 以HSL颜色空间举例, 转换公式为:

[0206] $G2 = (\max(R, G, B) + \min(R, G, B)) / 2;$

[0207] 利用所述第五分割阈值, 对所述当前跟踪图像的灰度图像进行二值化处理, 得到跟踪二值化图像。

[0208] 跟踪二值化图像 B_{track} 中像素点 (x, y) 的像素值 $B_{track}(x, y)$ 为:

$$[0209] \quad B_{\text{track}}(x,y) = \begin{cases} 0, & G2(x,y) < TK5 \\ 255, & G2(x,y) \geq TK5 \end{cases};$$

[0210] 其中, $G2(x,y)$ 为灰度图像 $G2$ 中像素点 (x,y) 的像素值。

[0211] 同时, 计算所述当前跟踪图像的参考颜色直方图, 所述当前跟踪图像的参考颜色直方图 H_{track} 即为当前跟踪图像的颜色直方图。

[0212] S405: 根据所述当前颜色直方图与所述参考颜色直方图的差异以及所述当前二值化图像与所述参考二值化图像的差异, 确定当前帧待检测图像是否满足预设条件。

[0213] 对于差分平均值 $\text{Diff}_{\text{binary}}$ 与颜色差异值 $\text{Diff}_{\text{color}}$, 分别同各自预先设定的阈值 $\text{Th}_{\text{binary}}$ 与 Th_{color} 进行比较, 如果 $\text{Diff}_{\text{binary}} < \text{Th}_{\text{binary}}$ 且 $\text{Diff}_{\text{color}} < \text{Th}_{\text{color}}$, 则确定当前帧待检测图像是否满足预设条件。

[0214] 若当前帧待检测图像满足预设条件, 则表明在当前帧待检测图像中跟踪标题候选区域成功; 否则表明在当前帧待检测图像中跟踪标题候选区域失败。无论跟踪成功, 还是跟踪失败, 通常都会继续针对下一帧待检测图像继续进行跟踪过程。

[0215] 在一示例中, 所述标题候选区域跟踪方法还包括:

[0216] 在跟踪检测过程中, 统计不满足预设条件的待检测图像的数量, 记为丢失帧数;

[0217] 当所述丢失帧数大于预设丢失阈值时, 结束对所述至少一帧待检测图像依次进行跟踪检测的过程。

[0218] 在具体示例中, 预设丢失帧数 lost_num 的初始值为 0, 当前帧待检测图像不满足预设条件时, 丢失帧数 lost_num 加 1, 如果 $\text{lost_num} > \text{Th}_{\text{lost}}$, 则结束跟踪检测, 否则, 对下一帧待检测图像中的跟踪区域继续执行上述跟踪检测过程。

[0219] 其中, 设置 $\text{lost_num} > \text{Th}_{\text{lost}}$ 这一条件的目的是为了避开个别视频信号受到干扰, 引起图像出现失真, 导致匹配失败, 通过 $\text{lost_num} > \text{Th}_{\text{lost}}$ 这一条件的设置, 允许算法有个别数量的视频帧跟踪失败。其中, Th_{lost} 为针对 lost_num 设定的丢失计数阈值。

[0220] 本实施例提供的标题候选区域跟踪检测方法, 当所述标题候选区域为首次跟踪时, 获取当前帧待检测图像在跟踪区域内的图像, 作为候选跟踪图像; 当所述标题候选区域为非首次跟踪时, 获取当前帧待检测图像在跟踪区域内的图像, 作为当前跟踪图像; 根据所述当前跟踪图像的当前颜色直方图与所述候选跟踪图像的参考颜色直方图的差异以及所述当前跟踪图像的当前二值化图像与所述候选跟踪图像的参考二值化图像的差异, 确定当前帧待检测图像是否满足预设条件, 进而实现了按照待检测视频图像的时域顺序, 对标题候选区域进行跟踪检测的技术效果。

[0221] 对应于标题区域检测方法, 本发明实施例还提供了相应的标题区域检测系统。

[0222] 请参阅图 5, 图 5 为本申请实施例的标题区域检测系统的一种结构示意图。

[0223] 本实施例标题区域检测系统, 用于实施前述实施例的标题区域检测方法, 具体可参见前述的方法实施例。

[0224] 如图 5 所示, 所述系统包括:

[0225] 候选图像获取单元 U100, 用于获取视频对象中的至少一帧候选视频图像, 所述至少一帧候选视频图像采用相同的候选区域。

[0226] 字体高度过滤单元 U200, 用于根据至少一帧待检测图像中文字的字体高度, 确定

所述候选区域是否为标题候选区域。

[0227] 其中,所述至少一帧待检测图像分别为所述至少一帧候选视频图像的候选区域内的图像。

[0228] 具体地,可以根据一帧待检测图像中文字的字体高度来进行确定,也可以根据连续N帧待检测图像中文字的字体高度来进行确定。

[0229] 候选区域跟踪单元U300,用于当确定所述候选区域为标题候选区域时,按照所述至少一帧待检测视频图像的时域顺序,对所述至少一帧待检测图像依次进行跟踪检测;

[0230] 所述候选区域跟踪单元U300,还用于在所述跟踪检测的过程中,统计满足预设条件的待检测图像的数量,记为跟踪帧数;

[0231] 在一示例中,预设跟踪帧数的初始值为0,当当前帧待检测图像满足预设条件时,跟踪帧数加1。

[0232] 标题区域确定单元U400,用于在所述跟踪检测结束后,根据所述跟踪帧数,确定所述候选区域是否为标题区域。

[0233] 在一示例中,在跟踪检测结束后,比较跟踪帧数tracking_num与预先设定的阈值 $Th_{tracking_num}$ 的大小关系,如果 $tracking_num \geq Th_{tracking_num}$,则判定标题候选区域为标题区域,否则,判定标题候选区域为非标题区域(如,滚动字幕区域)。

[0234] 其中,所述候选图像获取单元U100包括:

[0235] 第一获取单元,用于获取视频对象中的任一帧视频图像;

[0236] 第二获取单元,用于获取所述任一帧图像的标题潜在区域内的图像,作为潜在图像;

[0237] 第一计算单元,用于计算所述潜在图像的边缘强度图;

[0238] 第一确定单元,用于根据所述潜在图像的边缘强度图在水平方向上投影得到的水平直方图,确定候选区域的上下边界;

[0239] 第二确定单元,用于根据所述潜在图像的边缘强度图在垂直方向上投影得到的垂直直方图,确定候选区域的左右边界;

[0240] 第三确定单元,用于根据所述上下边界与所述左右边界,确定候选区域;

[0241] 第四确定单元,用于当确定出的候选区域符合预设条件时,将所述任一帧视频图像确定为候选视频图像。

[0242] 其中,所述字体高度过滤单元U200包括:

[0243] 第二计算单元,用于根据任一帧待检测图像,计算所述任一帧待检测图像的文本掩码;

[0244] 第五确定单元,用于根据所述文本掩码,确定所述任一帧待检测图像中文字的字体高度;

[0245] 第三计算单元,用于根据全部待检测图像中文字的字体高度,计算全部待检测图像的文字字体平均高度;

[0246] 第六确定单元,用于根据所述文字字体平均高度,确定所述候选区域是否为标题候选区域。

[0247] 在一示例中,为了实现根据任一帧待检测图像,计算所述任一帧待检测图像的文本掩码,所述第二计算单元具体用于:

- [0248] 根据任一帧待检测图像,生成所述任一帧待检测图像的灰度图像;
- [0249] 计算所述任一帧待检测图像的灰度图像的分割阈值,其中,所述任一帧待检测图像的灰度图像的分割阈值作为第三分割阈值;
- [0250] 利用所述第三分割阈值,对所述任一帧待检测图像的灰度图像进行二值化处理,得到灰度二值化图像;
- [0251] 根据所述灰度二值化图像,计算所述灰度二值化图像的平均像素值;
- [0252] 根据所述平均像素值,确定背景像素值;
- [0253] 根据所述灰度二值化图像与背景像素值,生成与所述任一帧待检测图像相对应的文本掩码,其中,所述文本掩码为所述任一帧待检测图像中文本区域的掩码。
- [0254] 在一示例中,为了实现根据所述灰度二值化图像,计算所述灰度二值化图像的平均像素值,所述第二计算单元还具体用于:
- [0255] 水平扫描所述灰度二值化图像的任一行像素;
- [0256] 将所述任一行中的首个像素值记为所述任一行的参考像素值;
- [0257] 统计所述任一行中像素值不等于所述任一行的参考像素值的像素个数,记为第一个数;
- [0258] 当所述第一个数小于预设个数阈值时,确定所述任一行符合第一预设条件;
- [0259] 统计符合第一预设条件的行数以及符合第一预设条件的所有行的参考像素值;
- [0260] 垂直扫描所述灰度二值化图像的任一列像素;
- [0261] 将所述任一列中的首个像素值记为所述任一列的参考像素值;
- [0262] 统计所述任一列中像素值不等于所述任一列的参考像素值的像素个数,记为第二个数;
- [0263] 当所述第二个数小于预设个数阈值时,确定所述任一列符合第二预设条件;
- [0264] 统计符合第二预设条件的列数以及符合第二预设条件的所有列的参考像素值;
- [0265] 计算符合第一预设条件的所有行的参考像素值与符合第二预设条件的所有列的参考像素值的像素值总和,以及,符合第一预设条件的行数与符合第二预设条件的列数的行列总数;
- [0266] 将所述像素值总和与所述行列总数的比值确定为所述灰度二值化图像的平均像素值。
- [0267] 在一示例中,为了实现根据所述平均像素值,确定背景像素值,所述第二计算单元还具体用于:
- [0268] 当所述平均像素值小于预设像素值时,确定背景像素值为第一像素值;
- [0269] 当所述平均像素值不小于预设像素值时,确定背景像素值为第二像素值。
- [0270] 在一示例中,为了实现根据所述灰度二值化图像与背景像素值,生成与所述任一帧待检测图像相对应的文本掩码,所述第二计算单元具体用于:
- [0271] 将所述灰度二值化图像中像素值与所述背景像素值相同的区域的确定为背景区域;
- [0272] 将所述灰度二值化图像中像素值与所述背景像素值不同的区域确定为文字区域;
- [0273] 根据所述背景区域与所述文字区域,生成与所述待检测图像相对应的文本掩码。
- [0274] 在一示例中,所述第二计算单元还用于:

- [0275] 将所述背景区域的像素值设置为第三像素值；
- [0276] 将所述文字区域的像素值设置为第四像素值。
- [0277] 在一示例中,为了实现根据所述文本掩码,确定所述任一帧待检测图像中文字的字体高度,所述第五确定单元具体用于:
- [0278] 利用所述文本掩码在水平方向上的投影,生成掩码水平特征 H_b ;
- [0279] 计算所述任一帧待检测图像的灰度图像的边缘强度图 E_{a11} ;
- [0280] 对所述任一帧待检测图像的灰度图像的边缘强度图 E_{a11} 进行二值化处理,得到边缘二值化图像 E ;
- [0281] 利用所述边缘二值化图像 E 在水平方向上的投影,生成边缘水平特征 H_e ;
- [0282] 根据所述掩码水平特征 H_b 与所述边缘水平特征 H_e ,计算图像分割特征 H_s ;
- [0283] 在 $i \in [0, H-1]$ 的范围内,对所述图像分割特征 $H_s[i]$ 进行遍历;
- [0284] 在所述遍历的过程中,当所述图像分割特征 $H_s[i]$ 符合水平分割条件时,记录每一水平方向上的位置 $[i, i+n]$ 为文字行区域,并将所述位置 $[i, i+n]$ 存入预设列表 POS_{text} ;
- [0285] 根据所述预设列表 POS_{text} 中符合有效高度条件的位置 $[i, i+n]$,计算分割区域 $SEG(0, i, W, n)$;
- [0286] 将所有分割区域 $SEG(0, i, W, n)$ 中 n 值的平均值确定为所述任一帧待检测图像中文字的字体高度;
- [0287] 其中,所述水平分割条件为同时满足第一条条件、第二条条件与第三条条件;第一条条件为: $i=0$ 或 $H_s[i-1]=1$;第二条条件为:对于每一个 j 值, $H_s[i+j]$ 均等于0, $0 \leq j \leq n$, j 与 n 为自然数;第三条条件为: $H_s[i+n+1]=1$ 或 $i=H-1$;
- [0288] 所述有效高度条件为: $n \geq M_h$;其中, $M_h = H / (N+1)$, N 为所述预设列表 POS_{text} 中的位置 $[i, i+n]$ 的数量;

[0289] 其中, $H_s[i] = \begin{cases} 0, & \text{numb}[i] \geq TH1 \\ 1, & \text{numb}[i] < TH1 \end{cases}$, $E(x,y) = \begin{cases} 0, & E_{a11}(x,y) \leq TK4 \\ 1, & E_{a11}(x,y) > TK4 \end{cases}$, $H_e[i] = \begin{cases} 0, & \text{nume}[i] \geq TH2 \\ 1, & \text{nume}[i] < TH2 \end{cases}$,

$H_s[i] = H_b[i] | H_e[i]$; $\text{numb}[i]$ 为所述掩码图像第 $i+1$ 行中像素值不等于所述背景像素值的像素个数; $TH1$ 为预设的第一特征阈值; $\text{nume}[i]$ 为所述边缘二值化图像第 $i+1$ 行中像素值不等于0的像素个数, $TH2$ 为预设的第二特征阈值; $TK4$ 为第四分割阈值; $i \in [0, H-1]$, H 为所述待检测图像的高度, W 为所述待检测图像的宽度。

[0290] 在一示例中,为了实现当确定所述候选区域为标题候选区域时,按照所述至少一帧待检测视频图像的时域顺序,对所述至少一帧待检测图像依次进行跟踪检测,所述候选区域跟踪单元 $U300$ 包括:

[0291] 第三获取单元,用于当所述标题候选区域为首次跟踪时,获取当前帧待检测图像在跟踪区域内的图像,作为候选跟踪图像;所述跟踪区域与所述标题候选区域相对应;

[0292] 第四计算单元,用于计算所述候选跟踪图像的参考颜色直方图与参考二值化图像;

[0293] 第四获取单元,用于当所述标题候选区域为非首次跟踪时,获取当前帧待检测图像在跟踪区域内的图像,作为当前跟踪图像;

[0294] 第五计算单元,用于计算所述当前跟踪图像的当前颜色直方图与当前二值化图

像；

[0295] 第七确定单元,用于根据所述当前颜色直方图与所述参考颜色直方图的差异以及所述当前二值化图像与所述参考二值化图像的差异,确定当前帧待检测图像是否满足预设条件。

[0296] 在一示例中,所述候选区域跟踪单元U300还用于:

[0297] 在跟踪检测过程中,统计不满足预设条件的待检测图像的数量,记为丢失帧数;

[0298] 当所述丢失帧数大于预设丢失阈值时,结束对所述至少一帧待检测图像依次进行跟踪检测的过程。

[0299] 本实施例提供的标题区域检测系统,基于新闻标题的文字字体的高度特征,能够滤除滚动字幕的干扰,准确判断出视频图像中的检测区域是否为标题区域;并且,针对不同风格的新闻视频中新闻标题文字的高度特征,本发明可以通过设定相应的预设高度阈值,以适用于各种风格的新闻视频中标题区域的准确定位,具有普遍适用性。

[0300] 本发明提供的标题区域检测方法及系统,获取视频对象中采用相同的候选区域的至少一帧候选视频图像,根据候选区域中的至少一帧待检测图像中文字的字体高度,确定候选区域是否为标题候选区域;当确定候选区域为标题候选区域时,按照至少一帧待检测视频图像的时域顺序,对候选区域进行跟踪检测;在跟踪检测的过程中,统计满足预设条件的待检测图像的跟踪帧数;最后根据跟踪帧数,确定候选区域是否为标题区域。本发明基于新闻标题的文字的字体高度特征,能够滤除滚动字幕的干扰,准确判断出标题候选区域,并通过对标题候选区域的跟踪检测,能够最终确定出标题候选区域是否为标题区域,提高标题区域的检测精度。

[0301] 最后,还需要说明的是,在本文中,诸如第一和第一等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来,而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且,术语“包括”、“包含”或者任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、物品或者设备不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、物品或者设备所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括所述要素的过程、方法、物品或者设备中还存在另外的相同要素。

[0302] 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本申请可借助软件加必需的硬件平台的方式来实现,当然也可以全部通过硬件来实现,但很多情况下前者是更佳的实施方式。基于这样的理解,本申请的技术方案对背景技术做出贡献的全部或者部分可以以软件产品的形式体现出来,该计算机软件产品可以存储在存储介质中,如ROM/RAM、磁碟、光盘等,包括若干指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备等)执行本申请各个实施例或者实施例的某些部分所述的方法。

[0303] 本说明书中各个实施例采用递进的方式描述,每个实施例重点说明的都是与其他实施例的不同之处,各个实施例之间相同相似部分互相参见即可。对于实施例公开的系统而言,由于其与实施例公开的方法相对应,所以描述的比较简单,相关之处参见方法部分说明即可。

[0304] 本文中应用了具体个例对本申请的原理及实施方式进行了阐述,以上实施例的说明只是用于帮助理解本申请的方法及其核心思想;同时,对于本领域的一般技术人员,依据

本申请的思想,在具体实施方式及应用范围上均会有改变之处。综上所述,本说明书内容不应理解为对本申请的限制。

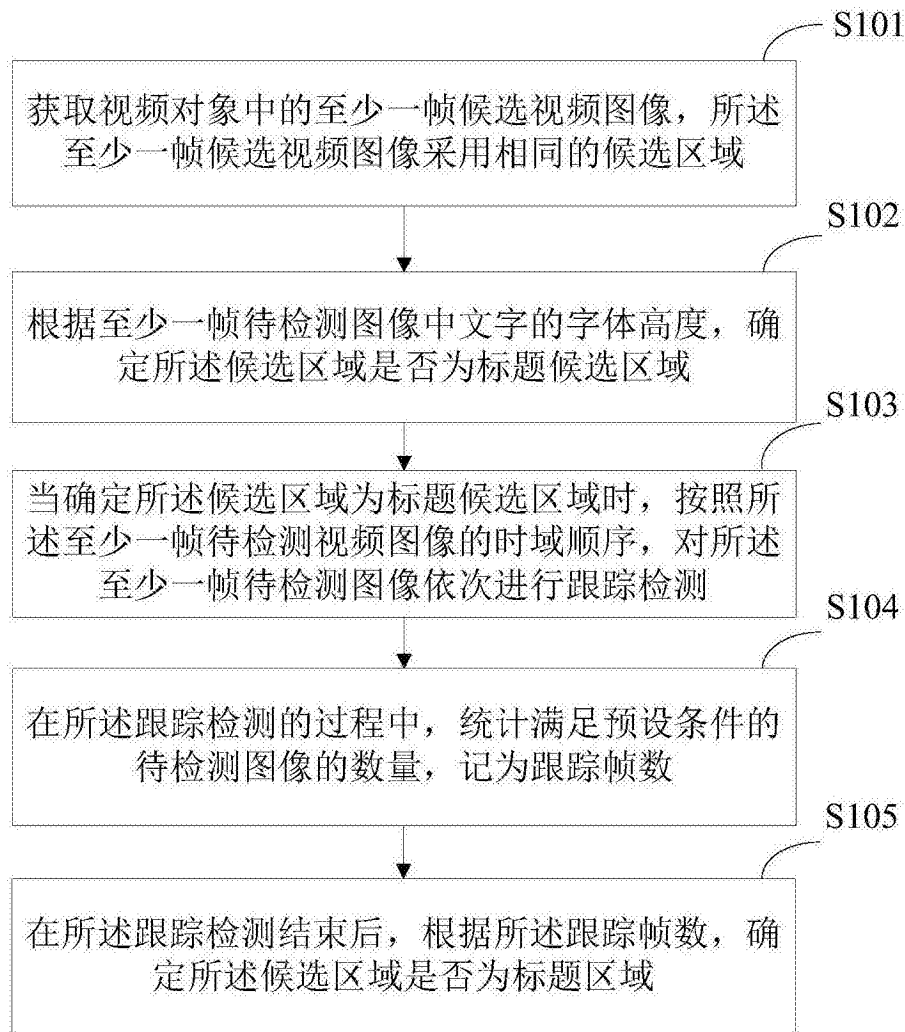


图1

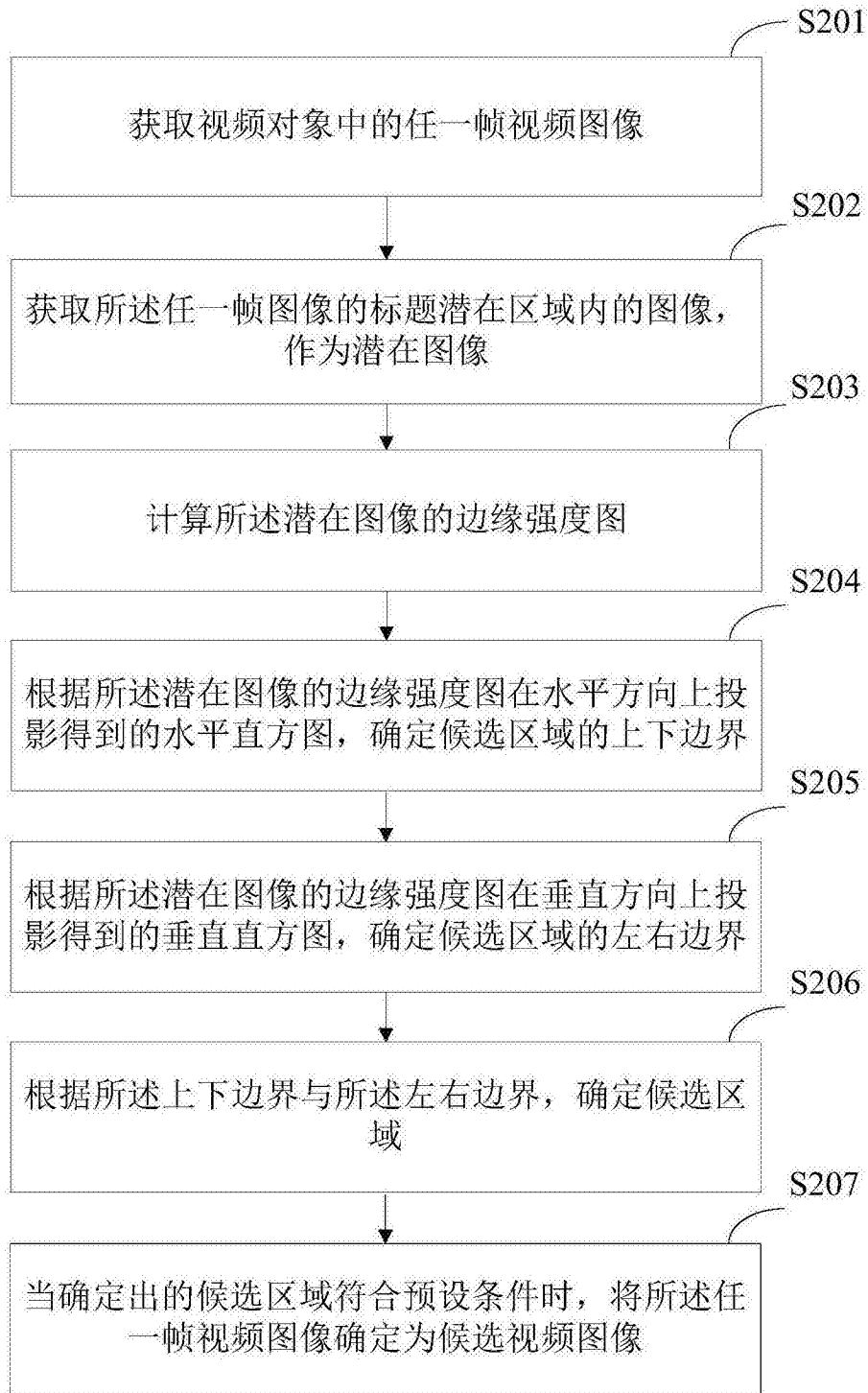


图2

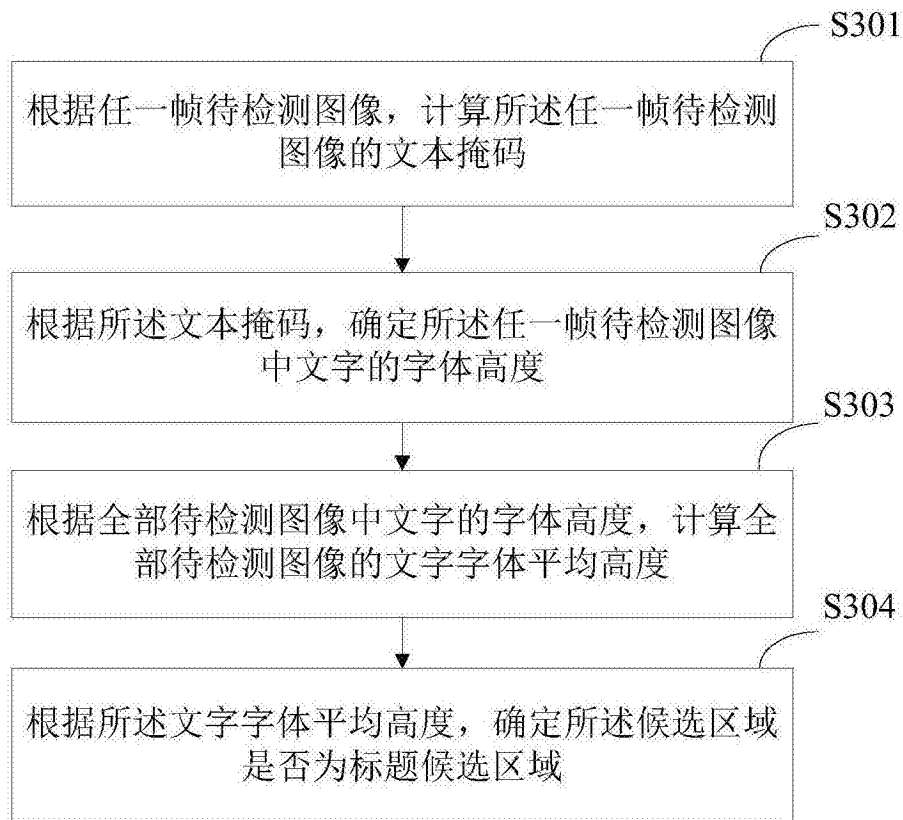


图3

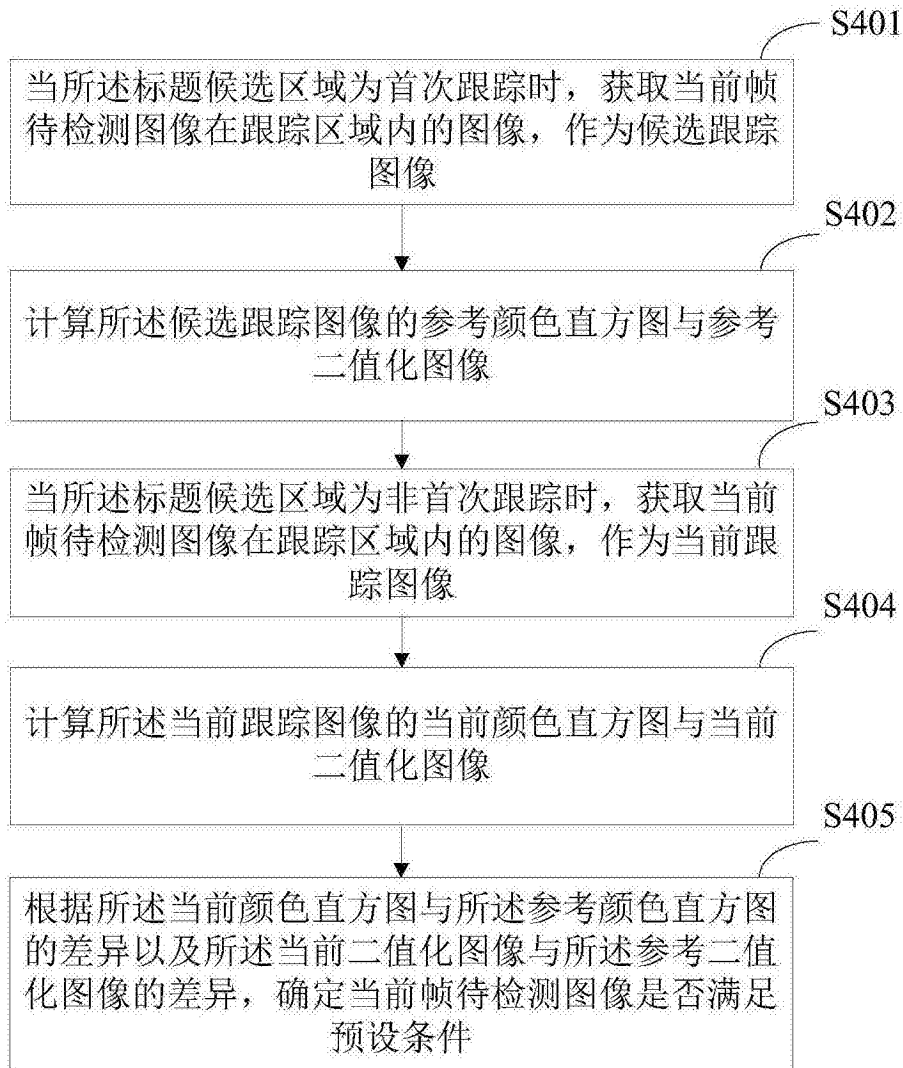


图4

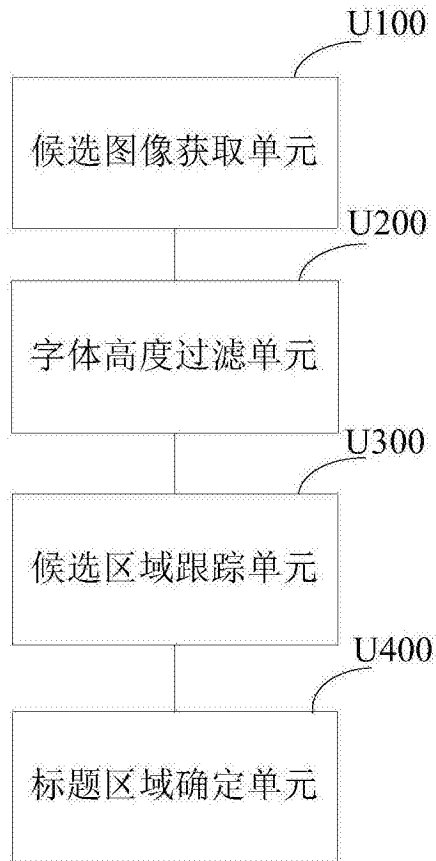


图5