

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2021年5月27日 (27.05.2021)



(10) 国际公布号
WO 2021/098861 A1

- (51) 国际专利分类号:
G06K 9/32 (2006.01)
- (21) 国际申请号: PCT/CN2020/130654
- (22) 国际申请日: 2020年11月20日 (20.11.2020)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201911147915.4 2019年11月21日 (21.11.2019) CN
- (71) 申请人: 上海高德威智能交通系统有限公司 (SHANGHAI GOLDWAY INTELLIGENT TRANSPORTATION SYSTEM CO., LTD.) [CN/CN]; 中国上海市嘉定区云谷路599弄6号2010室, Shanghai 201821 (CN)。
- (72) 发明人: 乔梁(QIAO, Liang); 中国上海市嘉定区云谷路599弄6号2010室, Shanghai 201821 (CN)。
- (74) 代理人: 北京三高永信知识产权代理有限公司(BEIJING SAN GAO YONG XIN INTELLECTUAL PROPERTY AGENCY CO., LTD.); 中国北京市

海淀区学院路蓟门里和景园A座1单元102室, Beijing 100088 (CN)。

- (81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。
- (84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,

(54) Title: TEXT RECOGNITION METHOD, APPARATUS, RECOGNITION DEVICE, AND STORAGE MEDIUM

(54) 发明名称: 识别文本的方法、装置、识别设备和存储介质

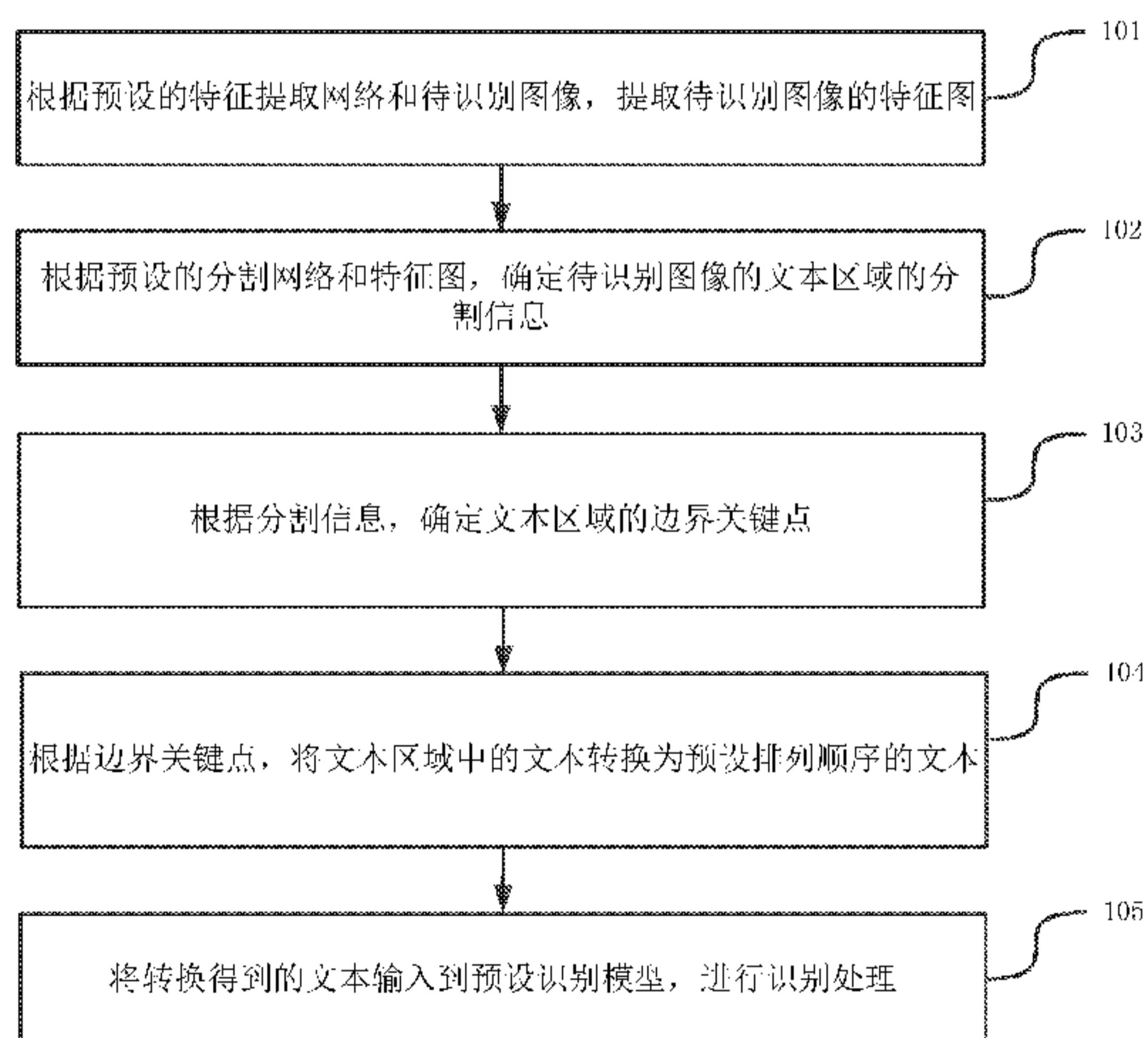


图 1

- 101 ACCORDING TO PRESET EXTRACTION FEATURE NETWORK AND IMAGE TO BE RECOGNIZED, EXTRACT FEATURE MAP OF IMAGE TO BE RECOGNIZED
- 102 ACCORDING TO PRESET SEGMENTATION NETWORK AND IMAGE TO BE RECOGNIZED, DETERMINE SEGMENTATION INFORMATION OF TEXT AREA OF IMAGE TO BE RECOGNIZED
- 103 ACCORDING TO SEGMENTATION INFORMATION, DETERMINE BOUNDARY KEY POINTS OF TEXT AREA
- 104 ACCORDING TO BOUNDARY KEY POINTS, CONVERT TEXT IN TEXT AREA INTO TEXT IN PRESET ARRANGEMENT ORDER
- 105 INPUT CONVERTED TEXT INTO PRESET RECOGNITION MODEL AND PERFORM RECOGNITION PROCESSING

(57) Abstract: The present disclosure provides a text recognition method, apparatus, recognition device, and storage medium, belonging to the field of text detection. The method comprises: according to a feature map of an extraction image to be recognized, the recognition device can, using a preset segmentation network and the image to be recognized, determine the segmentation information of the text area of the image to be recognized, and then, according to the segmentation information, determine the boundary key points of the text area, and use the boundary key points to convert the text in the text area into text in a target arrangement order, then input the converted text into a preset recognition model and perform recognition processing. The present invention improves the efficiency of recognition of curved text.

(57) 摘要: 本公开提供了一种识别文本的方法、装置、识别设备和存储介质, 属于文本检测领域。所述方法包括: 识别设备可以根据提取待识别图像的特征图, 然后使用预设的分割网络和待识别图像, 确定待识别图像的文本区域的分割信息, 然后根据分割信息, 确定文本区域的边界关键点, 使用边界关键点, 将文本区域中的文本转换为目标排列顺序的文本, 将转换得到的文本输入到预设识别模型, 进行识别处理。采用本公开, 可以提升弯曲文本的识别效率。

WO 2021/098861 A1

RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布：

— 包括国际检索报告(条约第21条(3))。

识别文本的方法、装置、识别设备和存储介质

本公开要求于 2019 年 11 月 21 日提交的申请号为 201911147915.4、发明名称为“识别文本的方法、装置、识别设备和存储介质”的中国专利申请的优先权，其全部内容通过引用结合在本公开中。

技术领域

本公开涉及文本检测领域，特别涉及一种识别文本的方法、装置、识别设备和存储介质。

背景技术

OCR (Optical Character Recognition, 光学字符识别) 技术作为一种重要的信息采集或录入的手段, 可以代替人工输入, 解放人力, 从而被应用于多个场景。例如, OCR 技术可以应用于车牌识别、身份证信息识别、PDF (Portable Document Format, 便携式文档格式) 文本解析等。OCR 技术的一般的流程是先进行检测后进行识别, 即先对传入的图像进行文本区域的检测, 再在检测到的文本区域内进行文本识别。

相关技术中, 在对传入的图像进行文本区域检测时, 得到的输出结果为一个四边形文本框 (如矩形文本框或凸四边形框), 然后对矩形文本框或凸四边形框中的内容进行识别。

由于在许多场景中均存在弯曲的文本, 例如, 发票章、广告牌等, 所以仅使用四边形作为检测结果, 这样, 四边形中有可能不会包含弯曲文本, 进而会导致弯曲文本的识别率比较低。

发明内容

为了解决弯曲文本的识别效率比较低的问题, 本公开实施例提供了一种识别文本的方法、装置、识别设备和存储介质。所述技术方案如下:

第一方面, 本公开提供了一种识别文本的方法, 所述方法包括:

根据预设的特征提取网络和待识别图像, 提取所述待识别图像的特征图;

根据预设的分割网络和所述特征图，确定所述待识别图像的文本区域的分割信息；

根据所述分割信息，确定所述文本区域的边界关键点；

根据所述边界关键点，将所述文本区域中的文本转换为目标排列顺序的文本；

将转换得到的文本输入到预设识别模型，进行识别处理。

可选的，所述根据所述分割信息，确定所述文本区域的边界关键点，包括：

根据所述分割信息中第一边界区域中各像素点与所述第一边界区域的两个边界关键点之间的偏移，确定所述第一边界区域的两个边界关键点的位置信息；并根据所述分割信息中第二边界区域中各像素点与所述第二边界区域的两个边界关键点之间的偏移，确定所述第二边界区域的两个边界关键点的位置信息，其中，所述第一边界区域位于所述文本区域的头部，所述第二边界区域位于所述文本区域的尾部；

根据所述第一边界区域的两个边界关键点的位置信息和所述第二边界区域的两个边界关键点的位置信息，确定所述文本区域中除所述第一边界区域与所述第二边界区域之外的其它边界关键点。

可选的，所述方法还包括：

确定所述文本区域中的文本的排布信息；

所述根据所述第一边界区域的两个边界关键点的位置信息和所述第二边界区域的两个边界关键点的位置信息，确定所述文本区域中除所述第一边界区域与所述第二边界区域之外的其它边界关键点，包括：

根据所述文本的排布信息、所述第一边界区域的两个边界关键点的位置信息和所述第二边界区域的两个边界关键点的位置信息，确定所述文本区域中除所述第一边界区域与所述第二边界区域之外的其它边界关键点。

可选的，所述根据所述文本的排布信息、所述第一边界区域的两个边界关键点的位置信息和所述第二边界区域的两个边界关键点的位置信息，确定所述文本区域中除所述第一边界区域与所述第二边界区域之外的其它边界关键点，包括：

如果所述排布信息为横向排布，则根据所述第一边界区域中第一边界关键点的位置信息和所述第二边界区域中第二边界关键点的位置信息，确定所述文

本区域的上边界上除所述第一边界关键点和所述第二边界关键点之外的其它边界关键点，并根据所述第一边界区域中第四边界关键点的位置信息和所述第二边界区域中第三边界关键点的位置信息，确定所述文本区域的下边界上除所述第三边界关键点和所述第三边界关键点之外的其它边界关键点；

如果所述排布信息为纵向排布，则根据所述第一边界区域中第一边界关键点的位置信息和所述第二边界区域中第二边界关键点的位置信息，确定所述文本区域的左边界上除所述第一边界关键点和所述第二边界关键点之外的其它边界关键点，并根据所述第一边界区域中第四边界关键点的位置信息和所述第二边界区域中第三边界关键点的位置信息，确定所述文本区域的右边界上除所述第三边界关键点和所述第四边界关键点之外的其它边界关键点。

可选的，所述根据所述第一边界区域中第一边界关键点的位置信息和所述第二边界区域中第二边界关键点的位置信息，确定所述文本区域的左边界上除所述第一边界关键点和所述第二边界关键点之外的其它边界关键点，包括：

根据所述第一边界区域中第一边界关键点的位置信息和所述第二边界区域中第二边界关键点的位置信息，确定第一预设数目个位置点的第一横坐标信息，其中，所述第一预设数目个位置点位于第一连线上，且是基于将所述第一连线等分为目标数值得到，所述目标数值等于所述第一预设数目加 1，所述第一连线为所述第一边界关键点与所述第二边界关键点的连线；

根据所述第一横坐标信息，确定所述文本区域的上边界上除所述第一边界关键点和所述第二边界关键点之外的其它边界关键点。

可选的，所述分割信息包括所述文本区域中属于文本的像素点与所述文本区域的上边界的偏移；

所述根据所述第一横坐标信息，确定所述文本区域的上边界上除所述第一边界关键点和所述第二边界关键点之外的其它边界关键点，包括：

对于所述第一预设数目个位置点中任一位置点的第一横坐标信息，根据所述位置点的第一横坐标信息对应的横坐标区间，确定所述文本区域中属于文本且横坐标位于所述横坐标区间的像素点；

将所述位置点的第一横坐标信息，确定为所述位置点对应的边界关键点的横坐标，并根据确定出的像素点与所述上边界的偏移，确定所述位置点对应的边界关键点的纵坐标，其中，所述位置点对应的边界关键点为位于所述文本区

域的上边界，且所述上边界包括所述第一边界关键点和所述第二边界关键点。

可选的，所述根据所述第一边界区域中第四边界关键点的位置信息和所述第二边界区域中第三边界关键点的位置信息，确定所述文本区域的下边界上除所述第三边界关键点和所述第三边界关键点之外的其它边界关键点，包括：

根据所述第一边界区域中第四边界关键点的位置信息和所述第二边界区域中第三边界关键点的位置信息，确定第一预设数目个位置点的第二横坐标信息，其中，所述第一预设数目个位置点位于第二连线上，且是基于将所述第二连线等分为目标数值得到，所述目标数值等于所述第一预设数目加 1，所述第二连线为所述第三边界关键点与所述第四边界关键点的连线；

根据所述第二横坐标信息，确定所述文本区域的下边界上除所述第三边界关键点和所述第三边界关键点之外的其它边界关键点。

可选的，所述分割信息包括所述文本区域中属于文本的像素点与所述文本区域的下边界的偏移；

所述根据所述第二横坐标信息，确定所述文本区域的下边界上除所述第三边界关键点和所述第三边界关键点之外的其它边界关键点，包括：

对于所述第一预设数目个位置点中任一位置点的第二横坐标信息，根据所述位置点的第二横坐标信息对应的横坐标区间，确定所述文本区域中属于文本且横坐标位于所述横坐标区间的像素点；

将所述位置点的第二横坐标信息，确定为所述位置点对应的边界关键点的横坐标，并根据确定出的像素点与所述下边界的偏移，确定所述位置点对应的边界关键点的纵坐标，其中，所述位置点对应的边界关键点为位于所述文本区域的下边界，且所述下边界包括所述第三边界关键点和所述第四边界关键点。

可选的，所述根据所述边界关键点，将所述文本区域中的文本转换为目标排列顺序的文本，包括：

使用薄板样条插值 TPS 对所述边界关键点进行模板匹配，将所述文本区域的文本转换为目标排列顺序的文本。

可选的，所述根据预设的特征提取网络和待识别图像，提取所述待识别图像的特征图，包括：

将待识别图像输入到预设的特征提取网络，提取所述待识别图像的预设尺寸的特征图；

将所述预设尺寸的特征图经过预设的卷积处理；

将卷积处理后的特征图进行卷积核矫正处理；

将卷积核矫正处理后的特征图经过上采样后与所述预设尺寸的特征图的上层特征图进行融合处理，得到所述待识别图像的特征图。

可选的，所述将卷积处理后的特征图进行卷积核矫正处理，包括：

将卷积处理后的特征图经过第一卷积核的卷积处理，得到第一特征图，将所述第一特征图进行第二卷积核的卷积处理，得到第二特征图；

将所述第二特征图进行第三卷积核的卷积处理，得到第三特征图，并将所述第二特征图进行第四卷积核的卷积处理，得到第四特征图；其中，所述第三卷积核的尺寸为 $m*n$ ，所述第四卷积核的尺寸为 $n*m$ ， m 和 n 为正整数，且 m 小于 n ；

将所述第一特征图、第二特征图、第三特征图，第四特征图进行通道拼接处理，得到第五特征图；

将所述第五特征图进行压缩激活处理，得到卷积核矫正处理后的特征图。

可选的，所述方法还包括：

获取样本集合，其中，所述样本集合中包括第二预设数目个标定文本区域的图像；

对于所述样本集合中的目标文本区域，确定所述目标文本区域的第一边界关键点、第二边界关键点、第三边界关键点和第四边界关键点；

对所述目标文本区域进行分割处理，得到所述目标文本区域的第一边界区域、第二边界区域、第三边界区域、第四边界区域、文本所在区域和背景区域，将所述第一边界区域、第二边界区域、第三边界区域、第四边界区域、文本所在区域和背景区域作为分割真值；

确定所述第一边界区域的各像素点与所述第一边界关键点、所述第三边界关键点的偏移，并确定第二边界区域的各像素点与所述第二边界关键点、所述第四边界关键点的偏移，作为边界关键点真值；

确定所述目标文本区域中文本所在区域的各像素点与所述目标文本区域的上边界、下边界、左边界和右边界的最小偏移，作为边界偏移真值；

根据所述样本集合中每个文本区域对应的分割真值、边界关键点真值和边界偏移真值，对分割网络训练模型进行训练，得到分割网络。

可选的，所述确定所述文本区域中的文本的排布信息，包括：

确定所述文本区域中的最大横坐标和最大纵坐标，并确定所述文本区域中的最小横坐标和最小纵坐标；

根据所述最大横坐标和最大纵坐标、所述最小横坐标和最小纵坐标，确定所述文本区域的外接矩形；

如果所述外接矩形的长度大于宽度，则确定所述文本区域中文本的排布为横向排布，如果所述外接矩形的长度小于宽度，则确定所述文本区域中文本的排布为纵向排布。

第二方面，本公开提供了一种识别文本的装置，所述装置包括：

提取模块，用于根据预设的特征提取网络和待识别图像，提取所述待识别图像的特征图；

确定模块，用于根据预设的分割网络和所述特征图，确定所述待识别图像的文本区域的分割信息；根据所述分割信息，确定所述文本区域的边界关键点；

转换模块，用于根据所述边界关键点，将所述文本区域中的文本转换为目标排列顺序的文本；

识别模块，用于将转换得到的文本输入到预设识别模型，进行识别处理。

可选的，所述确定模块，用于：

根据所述分割信息中第一边界区域中各像素点与所述第一边界区域的两个边界关键点之间的偏移，确定所述第一边界区域的两个边界关键点的位置信息；并根据所述分割信息中第二边界区域中各像素点与所述第二边界区域的两个边界关键点之间的偏移，确定所述第二边界区域的两个边界关键点的位置信息，其中，所述第一边界区域位于所述文本区域的头部，所述第二边界区域位于所述文本区域的尾部；

根据所述第一边界区域的两个边界关键点的位置信息和所述第二边界区域的两个边界关键点的位置信息，确定所述文本区域中除所述第一边界区域与所述第二边界区域之外的其它边界关键点。

可选的，所述确定模块，还用于：

确定所述文本区域中的文本的排布信息；

所述确定模块，用于：

根据所述文本的排布信息、所述第一边界区域的两个边界关键点的位置信息和所述第二边界区域的两个边界关键点的位置信息，确定所述文本区域中除所述第一边界区域与所述第二边界区域之外的其它边界关键点。

可选的，所述确定模块，用于：

如果所述排布信息为横向排布，则根据所述第一边界区域中第一边界关键点的位置信息和所述第二边界区域中第二边界关键点的位置信息，确定所述文本区域的上边界上除所述第一边界关键点和所述第二边界关键点之外的其它边界关键点，并根据所述第一边界区域中第四边界关键点的位置信息和所述第二边界区域中第三边界关键点的位置信息，确定所述文本区域的下边界上除所述第三边界关键点和所述第三边界关键点之外的其它边界关键点；

如果所述排布信息为纵向排布，则根据所述第一边界区域中第一边界关键点的位置信息和所述第二边界区域中第二边界关键点的位置信息，确定所述文本区域的左边界上除所述第一边界关键点和所述第二边界关键点之外的其它边界关键点，并根据所述第一边界区域中第四边界关键点的位置信息和所述第二边界区域中第三边界关键点的位置信息，确定所述文本区域的右边界上除所述第三边界关键点和所述第四边界关键点之外的其它边界关键点。

可选的，所述确定模块，用于：

根据所述第一边界区域中第一边界关键点的位置信息和所述第二边界区域中第二边界关键点的位置信息，确定第一预设数目个位置点的第一横坐标信息，其中，所述第一预设数目个位置点位于第一连线上，且是基于将所述第一连线等分为目标数值得到，所述目标数值等于所述第一预设数目加 1，所述第一连线为所述第一边界关键点与所述第二边界关键点的连线；

根据所述第一横坐标信息，确定所述文本区域的上边界上除所述第一边界关键点和所述第二边界关键点之外的其它边界关键点。

可选的，所述分割信息包括所述文本区域中属于文本的像素点与所述文本区域的上边界的偏移；

所述确定模块，用于：

对于所述第一预设数目个位置点中任一位置点的第一横坐标信息，根据所述位置点的第一横坐标信息对应的横坐标区间，确定所述文本区域中属于文本且横坐标位于所述横坐标区间的像素点；

将所述位置点的第一横坐标信息，确定为所述位置点对应的边界关键点的横坐标，并根据确定出的像素点与所述上边界的偏移，确定所述位置点对应的边界关键点的纵坐标，其中，所述位置点对应的边界关键点为位于所述文本区域的上边界，且所述上边界包括所述第一边界关键点和所述第二边界关键点。

可选的，所述确定模块，用于：

根据所述第一边界区域中第四边界关键点的位置信息和所述第二边界区域中第三边界关键点的位置信息，确定第一预设数目个位置点的第二横坐标信息，其中，所述第一预设数目个位置点位于第二连线上，且是基于将所述第二连线等分为目标数值得到，所述目标数值等于所述第一预设数目加 1，所述第二连线为所述第三边界关键点与所述第四边界关键点的连线；

根据所述第二横坐标信息，确定所述文本区域的下边界上除所述第三边界关键点和所述第三边界关键点之外的其它边界关键点。

可选的，所述分割信息包括所述文本区域中属于文本的像素点与所述文本区域的下边界的偏移；

所述确定模块，用于：

对于所述第一预设数目个位置点中任一位置点的第二横坐标信息，根据所述位置点的第二横坐标信息对应的横坐标区间，确定所述文本区域中属于文本且横坐标位于所述横坐标区间的像素点；

将所述位置点的第二横坐标信息，确定为所述位置点对应的边界关键点的横坐标，并根据确定出的像素点与所述下边界的偏移，确定所述位置点对应的边界关键点的纵坐标，其中，所述位置点对应的边界关键点为位于所述文本区域的下边界，且所述下边界包括所述第三边界关键点和所述第四边界关键点。

可选的，所述转换模块，用于：

使用薄板样条插值 TPS 对所述边界关键点进行模板匹配，将所述文本区域的文本转换为目标排列顺序的文本。

可选的，所述提取模块，用于：

将待识别图像输入到预设的特征提取网络，提取所述待识别图像的预设尺寸的特征图；

将所述预设尺寸的特征图经过预设的卷积处理；

将卷积处理后的特征图进行卷积核矫正处理；

将卷积核矫正处理后的特征图经过上采样后与所述预设尺寸的特征图的上层特征图进行融合处理，得到所述待识别图像的特征图。

可选的，所述提取模块，用于：

将卷积处理后的特征图经过第一卷积核的卷积处理，得到第一特征图，将所述第一特征图进行第二卷积核的卷积处理，得到第二特征图；

将所述第二特征图进行第三卷积核的卷积处理，得到第三特征图，并将所述第二特征图进行第四卷积核的卷积处理，得到第四特征图；其中，所述第三卷积核的尺寸为 $m*n$ ，所述第四卷积核的尺寸为 $n*m$ ， m 和 n 为正整数，且 m 小于 n ；

将所述第一特征图、第二特征图、第三特征图，第四特征图进行通道拼接处理，得到第五特征图；

将所述第五特征图进行压缩激活处理，得到卷积核矫正处理后的特征图。

可选的，所述装置还包括训练模块，用于：

获取样本集合，其中，所述样本集合中包括第二预设数目个标定文本区域的图像；

对于所述样本集合中的目标文本区域，确定所述目标文本区域的第一边界关键点、第二边界关键点、第三边界关键点和第四边界关键点；

对所述目标文本区域进行分割处理，得到所述目标文本区域的第一边界区域、第二边界区域、第三边界区域、第四边界区域、文本所在区域和背景区域，将所述第一边界区域、第二边界区域、第三边界区域、第四边界区域、文本所在区域和背景区域作为分割真值；

确定所述第一边界区域的各像素点与所述第一边界关键点、所述第三边界关键点的偏移，并确定第二边界区域的各像素点与所述第二边界关键点、所述第四边界关键点的偏移，作为边界关键点真值；

确定所述目标文本区域中文本所在区域的各像素点与所述目标文本区域的上边界、下边界、左边界和右边界的最小偏移，作为边界偏移真值；

根据所述样本集合中每个文本区域对应的分割真值、边界关键点真值和边界偏移真值，对分割网络训练模型进行训练，得到分割网络。

可选的，所述确定模块，用于：

确定所述文本区域中的最大横坐标和最大纵坐标，并确定所述文本区域中

的最小横坐标和最小纵坐标;

根据所述最大横坐标和最大纵坐标、所述最小横坐标和最小纵坐标, 确定所述文本区域的外接矩形;

如果所述外接矩形的长度大于宽度, 则确定所述文本区域中文本的排布为横向排布, 如果所述外接矩形的长度小于宽度, 则确定所述文本区域中文本的排布为纵向排布。

第三方面, 本公开提供了一种计算机可读存储介质, 所述存储介质内存储有计算机程序, 所述计算机程序被处理器执行时实现上述第一方面的方法步骤或者实现上述第二方面识别文本的装置的功能。

第四方面, 本公开提供了一种识别设备, 包括处理器和存储器, 其中, 所述存储器, 用于存放计算机程序; 所述处理器, 用于执行所述存储器上所存放的程序, 实现上述第一方面的方法步骤或者实现上述第二方面识别文本的装置的功能。

本公开实施例提供的技术方案带来的有益效果至少包括:

本公开实施例中, 识别设备可以根据提取待识别图像的特征图, 然后使用预设的分割网络和待识别图像, 确定待识别图像的文本区域的分割信息, 然后根据分割信息, 确定文本区域的边界关键点, 使用边界关键点, 将文本区域中的文本转换为目标排列顺序的文本, 将转换得到的文本输入到预设识别模型, 进行识别处理。这样, 由于是确定文本区域的边界关键点, 使用边界关键点, 将文本区域中的文本转换为目标排列顺序的文本, 再进行识别, 所以可以识别任意形状的弯曲文本, 进而可以提升弯曲文本的识别效率以及识别准确率。

而且本公开实施例中, 不需要对弯曲文本进行字符级别的标注, 也可以学习到文本阅读的语义信息(排布信息), 所以可以增加文本检测和识别的准确率。

而且本公开实施例中, 由于可以产生固定数目个边界关键点的检测框, 所以可以被应用于任何形状文本的标定工具, 适用范围广。

附图说明

图 1 是本公开实施例提供的一种识别文本的方法流程图；
图 2 是本公开实施例提供的一种提取特征图的示意图；
图 3 是本公开实施例提供的一种卷积核矫正处理的示意图；
图 4 是本公开实施例提供的一种确定文本转换的示意图；
图 5 是本公开实施例提供的一种 TPS 处理的示意图；
图 6 是本公开实施例提供的一种确定文本区域外接矩形的示意图；
图 7 是本公开实施例提供的一种确定边界关键点的示意图；
图 8 是本公开实施例提供的一种训练分割网络的方法流程图；
图 9 是本公开实施例提供的一种划分区域的示意图；
图 10 是本公开实施例提供的一种识别文本的装置的结构示意图；
图 11 是本公开实施例提供的一种识别文本的装置的结构示意图；
图 12 是本公开实施例提供的一种识别设备的结构示意图。

具体实施方式

为使本公开的目的、技术方案和优点更加清楚，下面将结合附图对本公开实施方式作进一步地详细描述。

本公开实施例提供了一种识别文本的方法，该方法的执行主体可以是识别设备。其中，识别设备可以是服务器或终端，服务器可以是识别文本程序的后台服务器，终端可以是电脑，终端上安装有用于识别文本的程序等。

识别设备中可以设置有处理器、存储器和收发器等。处理器可以是 CPU（Central Processing Unit，中央处理单元）等，可以用于进行识别文本过程中的处理，存储器可以为 RAM（Random Access Memory，随机存取存储器）、Flash（闪存）等，可以用于在识别文本过程中需要的数据以及产生的数据，收发器可以用于接收以及发送数据。在识别设备为终端时，识别设备还可以包括显示部件，显示部件用于显示对文本的识别结果等。

在进行实施前，首先介绍一下本公开实施例的应用场景以及涉及的名词：

OCR 技术作为一种重要的信息采集或录入的手段，可以代替人工输入，解放人力，从而被应用于多个场景。例如，OCR 技术可以应用于车牌识别、身份证信息识别、PDF 文本解析等。OCR 技术的一般的流程是先进行检测后进行识别，即先对传入的图像进行文本区域的检测，再在检测到的文本区域内进行文

本识别。

CNN (Convolutional Neural Network, 卷积神经网络), 一种前馈的人工神经网络, 通过权值共享方式提取图像的特征信息, 多层连接挖掘深层次的图像特征。

FPN (Feature Pyramid Network, 特征金字塔网络), 一种神经网络结构, 常用于通用目标检测, 主要是将不同层的特征进行融合, 使得网络可以适用于不同尺度的目标。

BFS (Breadth-First Search, 广度优先搜索), 一种连通图的遍历搜索策略。

TPS (Thin Plate Spline, 薄板样条插值), 一种插值算法, 可以将原图片按照对应模板进行形变。

本公开提供了一种识别文本的方法, 如图 1 所示, 该方法的处理流程可以如下, 以执行主体为识别设备为例进行说明:

步骤 101, 根据预设的特征提取网络和待识别图像, 提取待识别图像的特征图。

其中, 待识别图像为要进行文本识别的图像。预设的特征提取网络可以 FPN, 可以预先训练得到。

在实施中, 识别设备要对图像 (后续可以称为是待识别图像) 中的文本进行识别时, 可以将待识别图像输入到预设的特征提取网络中, 输出则为待识别图像的特征图。

可选的, 可以进行融合处理, 得到待识别图像的特征图, 相应的步骤 101 的处理可以如下:

识别设备将待识别图像输入到预设的特征提取网络, 提取待识别图像的预设尺寸的特征图, 将预设尺寸的特征图经过预设的卷积处理; 将卷积处理后的特征图进行卷积核矫正处理; 将卷积核矫正处理后的特征图经过上采样后与预设尺寸的特征图的上一层特征图进行融合处理, 得到待识别图像的特征图。

其中, 预设的特征提取网络可以是 FPN, 可以使用的是 FPN 的 RestNet-101。

在实施中, 识别设备可以将待识别图像输入到预设的特征提取网络, 经过特征提取后, 得到待识别图像的预设尺寸的特征图。然后将该预设尺寸的特征图经过预设的卷积处理, 再将卷积处理后的特征图进行卷积核矫正处理, 得到

卷积核矫正处理后的特征图。

将卷积核矫正处理后的特征图进行上采样，得到上采样后的特征图，将上采样后的特征图与预设尺寸的特征图的上一层特征图进行融合处理（即将特征图直接进行相加），得到的还是预设尺寸的特征图，这样，就得到待识别图像对应的特征图，用于后续处理。

例如，如图 2 所示，预设尺寸的特征图可以是 8 倍特征图，将 8 倍特征图经过 $1*1$ 的卷积核的卷积处理，然后再进行卷积核矫正处理，得到的还是 8 倍特征图。将该 8 倍特征图上采样后，得到 4 倍特征图。由于 8 倍特征图的上一层特征图是 4 倍特征图，所以可以将上采样后的 4 倍特征图与 8 倍特征图的上一层特征图进行融合，得到 4 倍特征图，用于后续处理。此处 x 倍特征图指的是原图尺寸 $1/x$ 的特征图，例如，4 倍特征图是原图尺寸 $1/4$ 的特征图，8 倍特征图是原图尺寸 $1/8$ 的特征图。所以上述 8 倍特征图上采样后得到一个与 4 倍特征图一样大小的特征图，然后将上采样后的 4 倍特征图与 8 倍特征图的上一层特征图进行融合，得到一个新的 4 倍特征图。

需要说明的是，上述预设尺寸的特征图可以基于实际场景进行设定，如果实际场景中包含较多的小尺度文本，则需要添加 2 倍特征图，如果实际场景中包含较多大尺度文本，则需要添加 32 倍特征图。

还需要说明的是，上述在预设的特征提取网络输入的待识别图像可以是 RGB（Red, Green, Blue, 红绿蓝）图像。

可选的，可以使用数个大卷积核进行卷积核矫正处理，相应的处理可以如下：

识别设备将卷积处理后的特征图经过第一卷积核的卷积处理，得到第一特征图，将第一特征图进行第二卷积核的卷积处理，得到第二特征图；将第二特征图进行第三卷积核的卷积处理，得到第三特征图，并将第二特征图进行第四卷积核的卷积处理，得到第四特征图；将第一特征图、第二特征图、第三特征图、第四特征图进行通道拼接处理，得到第五特征图；将第五特征图进行压缩激活处理，得到卷积核矫正处理后的特征图。

其中，第三卷积核的尺寸为 $m*n$ ，第四卷积核的尺寸为 $n*m$ ， m 和 n 为正整数，且 m 小于 n ， m 可以为 1， n 为大于 1 的正整数， n 的取值范围可以为 7~11。

在实施中，识别设备可以将卷积处理后的特征图经过第一卷积核的卷积处

理，得到第一特征图，将第一特征图进行第二卷积核的卷积处理，得到第二特征图，然后将第二特征图进行第三卷积核的卷积处理，得到第三特征图，并将第二特征图进行第四卷积核的卷积处理，得到第四特征图。

然后将第一特征图、第二特征图、第三特征图、第四特征图进行通道拼接处理，得到第五特征图。最后将第五特征图进行压缩激活处理，得到卷积核矫正处理后的特征图。

如图 3 所示，第一卷积核可以是 1×1 ，有 256 个通道，第二卷积核可以为 3×3 ，有 256 个通道，第三卷积核可以为 $1 \times n$ (n 可以取 7~11 中的一个正整数)，有 128 个通道，第四卷积核可以为 $n \times 1$ (n 可以取 7~11 中的一个正整数)，有 128 个通道。

这样，采用了第三卷积核和第四卷积核，第三卷积核为 $1 \times n$ ，可以提取到横向的长文本特征，第四卷积核为 $n \times 1$ ，可以提取到纵向的长文本特征。

步骤 102，根据预设的分割网络和特征图，确定待识别图像的文本区域的分割信息。

其中，分割网络可以预先训练得到，训练过程在后面进行描述。

在实施中，识别设备在得到待识别图像的特征图之后，可以获取预设的分割网络，将待识别图像的特征图输入到该分割网络中，输出则为待识别图像的文本区域的分割信息，分割信息包括文本区域中各类别的区域所在的位置，类别包括第一边界区域的类别（即头边界所在区域的类别）、第二边界区域的类别（尾边界所在区域的类别）、第三边界区域的类别（即上边界所在区域的类别）和第四边界区域的类别（即下边界所在区域的类别）、文本所在区域的类别、背景区域的类别。此处第一边界区域、第二边界区域、第三边界区域和第四边界区域在图 9 中有示例。

另外，对于待识别图像中每一个文本区域进行遍历搜索算法（如 BFS 等）的处理，搜索过程中记录下该文本区域所有相邻的头边界所在区域和尾边界所在区域，选取面积最大的一组头边界所在区域和尾边界所在区域分别作为该文本区域的第一边界区域和第二边界区域。

另外，对于一个文本区域，如果没有尾边界所在区域或者头边界所在区域，则将该文本区域删除，后续不再处理，这是由于大部分由于纹理造成的误检是缺少头边界所在区域或者尾边界所在区域。

需要说明的是，由于相近的文本区域是会被边界隔开，所以每一个独立的文本类区域连通域对应一个检出的文本实体。

此处还需要说明的是，第三边界区域和第四边界区域中的像素在分割网络训练和使用过程中可以作为同一类像素使用。例如，第三边界区域和第四边界区域中的像素在分割网络训练使用过程中，统称为第三边界区域。

此处还需要说明的是，这几种不同类别的边界在预测阶段是有顺序的，即先预测的是文本所在区域，然后是第一边界区域和第二边界区域（此处第一边界区域和第二边界区域有可能会覆盖文本所在区域的部分像素），最后是第三边界区域和第四边界区域（此处第三边界区域和第四边界区域有可能会覆盖第一边界区域和第二边界区域的部分像素）。

步骤 103，根据分割信息，确定文本区域的边界关键点。

在实施中，识别设备在确定出分割信息之后，可以使用该分割信息，确定文本区域的多个边界关键点。

步骤 104，根据边界关键点，将文本区域中的文本转换为目标排列顺序的文本。

其中，目标排列顺序可以是左向右的顺序。

在实施中，识别设备可以使用文本区域的边界关键点，将文本区域中包括的文本，转换为目标排列顺序的文本，也即将不同朝向的文本可以转换为按照目标排列顺序排列的文本。例如，如图 4 所示，目标排列顺序为从左向右，如果文本区域的文本为竖直排列，则转换为从左向右的文本。

可选的，在步骤 104 中，可以使用 TPS 进行转换，相应的步骤 104 的处理可以如下：

识别设备使用 TPS 对边界关键点进行模板匹配，将文本区域的文本转换为目标排列顺序的文本。

在实施中，假设步骤 103 的结果为 $2*N$ 个边界关键点，预设的识别模型的输入为 $H*W$ （ H 为图像的高， W 为图像的宽），可以将步骤 103 得到的 $2*N$ 个边界关键点的位置坐标一一映射到如下坐标：

$$P_i^* = \begin{cases} \left((i-1) * \frac{H-2*\Delta x}{N-1} + \Delta x, \Delta y \right), 1 \leq i \leq N \\ \left((2*N-i) * \frac{H-2*\Delta x}{N-1} + \Delta x, H - \Delta y \right), N+1 \leq i \leq 2*N \end{cases} \quad (1)$$

其中，在(1)中， i 为第 i 个边界关键点， $\Delta x, \Delta y$ 为边界预留空间， p_i^* 为第 i 个边界关键点转换得到的坐标，可以设定为5~15个像素点，以免形变过大造成文本区域被新的模板所截断。

这样，经过映射之后，不同朝向的文本，被转换为目标排列顺序的文本，以便于使用预设识别模型进行识别。如图5所示，将14个边界关键点的弯曲文本区域，经过TPS转换为从左向右的文本区域。

需要说明的是，由于使用了TPS，将文本的顺序统一为目标排列顺序的文本，所以可以使后续的预设识别模型无需学习复杂的形变参数。

步骤105，将转换得到的文本输入到预设识别模型，进行识别处理。

其中，预设识别模型可以是预先训练的CNN。

在实施中，识别设备可以将步骤104得到的文本输入到预设识别模型，则会输出对该文本的识别结果。

可选的，在步骤102中，确定边界关键点的过程可以如下：

根据分割信息中第一边界区域中各像素点与第一边界区域的两个边界关键点之间的偏移，确定第一边界区域的两个边界关键点的位置信息；并根据分割信息中第二边界区域中各像素点与第二边界区域的两个边界关键点之间的偏移，确定第二边界区域的两个边界关键点的位置信息；根据第一边界区域的两个边界关键点的位置信息和第二边界区域的两个边界关键点的位置信息，确定文本区域中除第一边界区域与第二边界区域之外的其它边界关键点。

其中，第一边界区域位于文本区域的头部，可以称为是头边界所在区域，第二边界区域位于文本区域的尾部，可以称为是尾边界所在区域。例如，对于文本从左向右的文本区域，第一边界区域位于文本区域的左侧，即左边界所在区域，第二边界区域位于文本区域的右侧，即右边界所在区域。对于文本从上向下的文本区域，第一边界区域位于文本区域的上侧，即上边界所在区域，第二边界区域位于文本区域的下侧，即下边界所在区域。

在实施中，分割信息中包括第一边界区域中各像素点与第一边界区域的两个边界关键点之间的偏移（如果边界关键点对应有横坐标和纵坐标，各像素点与边界关键点之间的偏移包括横坐标的偏移和纵坐标的偏移）。

识别设备可以使用第一边界区域中各像素点与第一边界区域中两个边界关键点之间的偏移，确定出第一边界区域的两个边界关键点的位置信息。第一边

界区域的两个边界关键点中的一个（使用 p_1 表示）的确定方法可以如下：

$$p_1 = \left(\frac{\sum_{(x_i, y_i) \in R_H} (x_i + \Delta dx_i)}{\|R_H\|}, \frac{\sum_{(x_i, y_i) \in R_H} (y_i + \Delta dy_i)}{\|R_H\|} \right) \quad (2)$$

其中， R_H 为第一边界区域， $\|R_H\|$ 表示代表 R_H 集合中像素点的数目， (x_i, y_i) 为属于第一边界区域的像素点的位置坐标， Δdx_i 为 x_i 与 p_1 之间在 x 方向上的偏移， Δdy_i 为 y_i 与 p_1 之间在 y 方向上的偏移。同理按照这种方式，可以确定第一边界区域的另一个边界关键点的位置坐标。

同理，按照上述式子（2），识别设备可以使用分割信息中第二边界区域中各像素点与第二边界区域的两个边界关键点之间的偏移，确定出第二边界区域的两个边界关键点的位置信息。

然后识别设备可以使用第一边界区域的两个边界关键点的位置信息和第二边界区域的两个边界关键点的位置信息，确定出文本区域中，第一边界区域与第二边界区域的其它边界关键点。

需要说明的是，一般是以待识别图像的左上角为坐标原点，向右为 x 的正向，向下为 y 的正方向，建立直角坐标系。

可选的，在确定其它边界点时，还考虑了文本区域中的文本的排布信息，相应的处理可以如下：

确定文本区域中的文本的排布信息，根据文本的排布信息、第一边界区域的两个边界关键点的位置信息和第二边界区域的两个边界关键点的位置信息，确定文本区域中除第一边界区域与第二边界区域之外的其它边界关键点。

其中，排布信息包括从左向右排布（即横向排布）和从上向下排布（即纵向排布）。

在实施中，识别设备可以识别文本区域中的文本的排布信息，然后使用该排布信息、第一边界区域的两个边界关键点的位置信息和第二边界区域的两个边界关键点的位置信息，确定文本区域中除第一边界区域中的边界关键点与第二边界区域中的边界关键点之外的其它边界关键点。

可选的，本公开实施例中，还提供了确定文本的排布信息的方式：

确定文本区域中的最大横坐标和最大纵坐标，并确定文本区域中的最小横坐标和最小纵坐标；根据最大横坐标和最大纵坐标、最小横坐标和最小纵坐标，确定文本区域的外接矩形；如果外接矩形的长度大于宽度，则确定文本区域中文本的排布为横向排布，如果外接矩形的长度小于宽度，则确定文本区域中文

本的排布为纵向排布。

在实施中，识别设备可以在文本区域中各像素点的位置坐标，确定最大横坐标和最大纵坐标，并且确定最小横坐标和最小纵坐标。使用最大横坐标和最小横坐标相减，得到文本区域的外接矩形的长度，并且将最大纵坐标和最小纵坐标相减，得到文本区域的外接矩形的宽度。例如，如图 6 所示，最大纵坐标为 5，最小纵坐标为 3，最大横坐标为 12，最小横坐标为 5，则长度为 7，宽度为 2。

然后判断外接矩形的长度和宽度的大小，如果长度大于宽度，则确定文本区域中文本的排布为横向排布，如果长度小于宽度，则确定文本区域中文本的排布为纵向排布。

另外，如果宽度等于高度，则确定为横向排布和纵向排布，后续分别进行识别处理。

可选的，对于不同的排布信息，按照不同的方式，识别设备确定其它边界关键点，相应的处理可以如下：

如果排布信息为横向排布，则根据第一边界区域中第一边界关键点的位置信息和第二边界区域中第二边界关键点的位置信息，确定文本区域的上边界上除第一边界关键点和第二边界关键点之外的其它边界关键点，并根据第一边界区域中第四边界关键点的位置信息和第二边界区域中第三边界关键点的位置信息，确定文本区域的下边界上除第三边界关键点和第三边界关键点之外的其它边界关键点；如果排布信息为纵向排布，则根据第一边界区域中第一边界关键点的位置信息和第二边界区域中第二边界关键点的位置信息，确定文本区域的左边界上除第一边界关键点和第二边界关键点之外的其它边界关键点，并根据第一边界区域中第四边界关键点的位置信息和第二边界区域中第三边界关键点的位置信息，确定文本区域的右边界上除第三边界关键点和第四边界关键点之外的其它边界关键点。

其中，第一边界关键点和第四边界关键点位于第一边界区域，第二边界关键点和第三边界关键点位于第二边界区域。排布信息为横向排布，第一边界关键点和第二边界关键点位于文本区域的上边界，第三边界关键点和第四边界关键点位于文本区域的下边界。排布信息为纵向排布，第一边界关键点和第二边界关键点位于文本区域的左边界，第三边界关键点和第四边界关键点位于文本

区域的右边界。

在实施中，如果排布信息为横向排布，则可以使用第一边界关键点的位置信息和第二边界点的位置信息，确定文本区域的上边界上除第一边界关键点和第二边界关键点之外的其它边界关键点的位置信息，并且可以使用第三边界关键点和第四边界关键点的位置信息，确定文本区域的下边界上除第三边界关键点和第四边界关键点之外的其它边界关键点的位置信息。

如果排布信息为纵向排布，则可以使用第一边界关键点的位置信息和第二边界点的位置信息，确定文本区域的左边界上除第一边界关键点和第二边界关键点之外的其它边界关键点的位置信息，并且可以使用第三边界关键点和第四边界关键点的位置信息，确定文本区域的右边界上除第三边界关键点和第四边界关键点之外的其它边界关键点的位置信息。

可选的，识别设备可以使用如下方式，确定边界关键点，相应的处理可以如下：

根据第一边界区域中第一边界关键点的位置信息和第二边界区域中第二边界关键点的位置信息，确定第一预设数目个位置点的第一横坐标信息，其中，第一预设数目个位置点位于第一连线上，且是基于将第一连线等分为目标数值得到，目标数值等于第一预设数目加 1，第一连线为第一边界关键点与第二边界关键点的连线；根据第一横坐标信息，确定文本区域的上边界上除第一边界关键点和第二边界关键点之外的其它边界关键点。

其中，第一预设数目可以预设，并且存储在识别设备上。

在实施中，排布信息为横向排布的情况下，识别设备可以将第一边界关键点和第二边界关键点连线，得到第一连线，然后将第一连线上等分为目标数值份（目标数值等于第一预设数目与 1 之和），得到第一预设数目个位置点。例如，如图 7 所示，第一预设数目为 5，目标数值为 6，将第一连线等分为 6 份，可以得到 5 个位置点。

由于第一边界关键点与第二边界关键点的位置信息已知，所以将第一连线等分为目标数值份后，使用第一边界关键点与第二边界关键点的位置信息，可以确定出这第一预设数目个位置点的横坐标，这样，即可得到第一横坐标信息。

然后可以使用第一横坐标信息，确定出文本区域的上边界上除第一边界关键点和第二边界关键点之外的其它边界关键点。

可选的，分割信息包括文本区域中属于文本的像素点与文本区域的上边界的偏移；确定文本区域的上边界上除第一边界关键点和第二边界关键点之外的其它边界关键点的方式可以如下：

对于第一预设数目个位置点中任一位置点的第一横坐标信息，根据该位置点的第一横坐标信息对应的横坐标区间，确定文本区域中属于文本且横坐标位于横坐标区间的像素点；将该位置点的第一横坐标信息，确定为该位置点对应的边界关键点的横坐标，并根据确定出的像素点与上边界的偏移，确定该位置点对应的边界关键点的纵坐标，其中，该位置点对应的边界关键点为位于文本区域的上边界，且上边界包括第一边界关键点和第二边界关键点。

在实施中，对于第一预设数目个位置点中任一位置点*i*的第一横坐标信息 x_i ，可以确定位置点*i*的第一横坐标信息对应的横坐标区间 $[x_i-a, x_i+a]$ （ a 可以预设并且存储至识别设备）。然后识别设备在文本区域中属于文本的像素点中，确定属于文本且横坐标位于横坐标区间 $[x_i-a, x_i+a]$ 的像素点。

然后将位置点*i*的第一横坐标信息 x_i ，确定为该位置点*i*对应的边界关键点的横坐标 x_i 。并且在分割信息中，获取属于文本且横坐标位于横坐标区间 $[x_i-a, x_i+a]$ 的像素点到上边界的最小偏移，使用该偏移中在*y*方向上的最小偏移，确定出该位置点*i*对应的边界关键点的纵坐标，该边界关键点的位置坐标用公式表示可以如下：

$$p_i = \left(x_i, \frac{\sum_{(x_i, y_i) \in B_i} (y_i + \Delta dy'_i)}{\|B_i\|} \right) \quad (3)$$

其中，在式（3）中， p_i 为位置点*i*对应的边界关键点的位置坐标， B_i 为第一边界区域， $\|B_i\|$ 表示横坐标区间 $[x_i-a, x_i+a]$ 中属于文本的像素点的数目， y_i 为横坐标区间 $[x_i-a, x_i+a]$ 中属于文本的像素点的纵坐标， $\Delta dy'_i$ 为横坐标区间 $[x_i-a, x_i+a]$ 中属于文本的像素点的纵坐标 y_i 与上边界在*y*方向上的最小偏移。

这样，使用与位置点*i*相同的方式，可以确定出第一预设数目个位置点分别对应的边界关键点的位置坐标。这样，就可以确定出上边界中除第一边界关键点和第二边界关键点之外的其它边界关键点。

可选的，可以使用以下方式确定下边界的其它边界关键点，相应的处理可以如下：

根据第一边界区域中第四边界关键点的位置信息和第二边界区域中第三边界关键点的位置信息，确定第一预设数目个位置点的第二横坐标信息，其中，

第一预设数目个位置点位于第二连线上，且是基于将第二连线等分为目标数值得到，目标数值等于第一预设数目加 1，第二连线为第三边界关键点与第四边界关键点的连线，根据第二横坐标信息，确定文本区域的下边界上除第三边界关键点和第三边界关键点之外的其它边界关键点。

其中，第一预设数目可以预设，并且存储在识别设备上。

在实施中，识别设备可以将第三边界关键点和第四边界关键点连线，得到第二连线，然后将第二连线上等分为目标数值份（目标数值等于第一预设数目与 1 之和），得到第一预设数目个位置点。

由于第三边界关键点与第四边界关键点的位置信息已知，所以将第二连线等分为目标数值份后，使用第三边界关键点与第四边界关键点的位置信息，可以确定出这第一预设数目个位置点的横坐标，这样，即可得到第二横坐标信息。

然后可以使用第二横坐标信息，确定出文本区域的下边界上除第三边界关键点和第四边界关键点之外的其它边界关键点。

可选的，识别设备可以使用属于文本的像素点与文本区域的下边界的偏移，确定下边界的其它边界关键点，处理可以如下：

对于第一预设数目个位置点中任一位置点的第二横坐标信息，根据位置点的第二横坐标信息对应的横坐标区间，确定文本区域中属于文本且横坐标位于横坐标区间的像素点；将位置点的第二横坐标信息，确定为位置点对应的边界关键点的横坐标，并根据确定出的像素点与下边界的偏移，确定位置点对应的边界关键点的纵坐标，其中，位置点对应的边界关键点为位于文本区域的下边界，且下边界包括第三边界关键点和第四边界关键点。

在实施中，对于第一预设数目个位置点中任一位置点 j 的第二横坐标信息 x_j ，可以确定位置点 j 的第二横坐标信息对应的横坐标区间 $[x_j - a, x_j + a]$ （ a 可以预设并且存储至识别设备），然后在文本区域中属于文本的像素点中，确定属于文本且横坐标位于横坐标区间 $[x_j - a, x_j + a]$ 的像素点。

然后将位置点 j 的第二横坐标信息 x_j ，确定为该位置点 j 对应的边界关键点的横坐标 x_j ，并且在分割信息中，获取属于文本且横坐标位于横坐标区间 $[x_j - a, x_j + a]$ 的像素点与下边界的最小偏移，使用 y 方向上的最小偏移，确定出该位置点 j 对应的边界关键点的纵坐标，该边界关键点的位置坐标用公式表示可以如下：

$$p_j = \left(x_j, \frac{\sum_{(x_j, y_j) \in B_j} (y_j + \Delta dy'_j)}{\|B_j\|} \right) \quad (4)$$

其中，在式(4)中， p_j 为位置点j对应的边界关键点的位置坐标， B_j 为第二边界区域， $\|B_j\|$ 表示横坐标区间 $[x_j-a, x_j+a]$ 中属于文本的像素点的数目， y_j 为横坐标区间 $[x_j-a, x_j+a]$ 中属于文本的像素点的纵坐标， $\Delta dy'_j$ 为横坐标区间 $[x_j-a, x_j+a]$ 中属于文本的像素点的纵坐标 y_j 与下边界在y方向上的最小偏移。

这样，使用与位置点j相同的方式，可以确定出第一预设数目个位置点对应的边界关键点的位置坐标。这样，就可以确定出下边界中除第三边界关键点和第四边界关键点之外的其它边界关键点。

可选的，对于排布信息为纵向排布，确定其它边界关键点的方式与横向排布基本类似，过程可以如下：

根据第一边界区域中第一边界关键点的位置信息和第二边界区域中第二边界关键点的位置信息，确定第一预设数目个位置点的第一纵坐标信息，其中，第一预设数目个位置点位于第三连线上，且是基于将第三连线等分为目标数值得到，目标数值等于第一预设数目加1，第三连线为第一边界关键点与第二边界关键点的连线，根据第一纵坐标信息，确定文本区域的左边界上除第一边界关键点和第二边界关键点之外的其它边界关键点。

其中，第一预设数目可以预设，并且存储在识别设备上。

在实施中，识别设备可以将第一边界关键点和第二边界关键点连线，得到第三连线，然后将第三连线上等分为目标数值份（目标数值等于第一预设数目与1之和），得到第一预设数目个位置点。

由于第一边界关键点与第二边界关键点的位置信息已知，所以将第一连线等分为目标数值份后，使用第一边界关键点与第二边界关键点的位置信息，可以确定出这第一预设数目个位置点的纵坐标，这样，即可得到第一纵坐标信息。

然后可以使用第一纵坐标信息，确定出文本区域的左边界上除第一边界关键点和第二边界关键点之外的其它边界关键点。

可选的，分割信息包括文本区域中属于文本的像素点与文本区域的左边界的偏移；确定文本区域的左边界上除第一边界关键点和第二边界关键点之外的其它边界关键点的方式可以如下：

对于第一预设数目个位置点中任一位置点的第一纵坐标信息，根据位置点

的第一纵坐标信息对应的纵坐标区间，确定文本区域中属于文本且纵坐标位于纵坐标区间的像素点；将位置点的第一纵坐标信息，确定为位置点对应的边界关键点的纵坐标，并根据确定出的像素点与左边界的偏移，确定位置点对应的边界关键点的横坐标，其中，位置点对应的边界关键点为位于文本区域的左边界，且左边界包括第一边界关键点和第二边界关键点。

在实施中，对于第一预设数目个位置点中任一位置点 k 的第一纵坐标信息 y_k ，可以确定位置点 k 的第一纵坐标信息对应的纵坐标区间 $[y_k - a, y_k + a]$ (a 可以预设并且存储至识别设备)，然后在文本区域中属于文本的像素点中，确定属于文本且纵坐标位于纵坐标区间 $[y_k - a, y_k + a]$ 的像素点。

然后将位置点 k 的第一纵坐标信息 y_k ，确定为该位置点 k 对应的边界关键点的纵坐标 y_k ，并且在分割信息中，获取属于文本且纵坐标位于纵坐标区 $[y_k - a, y_k + a]$ 的像素点对应的偏移，使用该偏移，确定出该位置点 k 对应的边界关键点的横坐标，该边界关键点的位置坐标用公式表示可以如下：

$$p_k = \left(\frac{\sum_{(x_i, y_i) \in B_k} (x_i + \Delta dx'_i)}{\|B_k\|}, y_k \right) \quad (5)$$

其中，在式 (5) 中， p_k 为位置点 k 对应的边界关键点的位置坐标， B_k 为第一边界区域， $\|B_k\|$ 表示纵坐标区间 $[y_k - a, y_k + a]$ 中属于文本的像素点的数目， x_i 为纵坐标区间 $[y_k - a, y_k + a]$ 中属于文本的像素点的横坐标， $\Delta dx'_i$ 为纵坐标区间 $[y_k - a, y_k + a]$ 中属于文本的像素点的横坐标 x_i 与上边界在 x 方向上的最小偏移。

这样，使用与位置点 k 相同的方式，可以确定出第一预设数目个位置点对应的边界关键点的位置坐标。这样，就可以确定出左边界上除第一边界关键点和第二边界关键点之外的其它边界关键点。

同理，可以确定出右边界上除第三边界关键点和第四边界关键点之外的其它边界关键点。

可选的，本公开实施例中，还提供了训练分割网络的过程，如图 8 所示，相应的处理可以如下：

步骤 801，获取样本集合。

在实施中，识别设备可以获取样本集合，样本集合中包括第二预设数目个标定文本区域的图像（第二预设数目可以预设，一般比较大）。

步骤 802，对于样本集合中的目标文本区域，确定目标文本区域的第一边界

关键点、第二边界关键点、第三边界关键点和第四边界关键点。

其中，目标文本区域是样本集中的任一文本区域。

在实施中，对于排布信息为横向排布的文本区域，此处可以是使用一个顺时针排布的多边形来标定一个文本，同时标定的起始顺序总是从文本的左上边界关键点开始，因为文本的标定总是单行的，所以可以确定最后一个边界关键点一定是这个文本区域的左下边界关键点。所以第一边界关键点为左上角边界关键点，第四边界关键点为左下角边界关键点，这两个边界关键点是起始点和结束点，所以可以直接获得。第二边界关键点和第三边界关键点可以使用以下公式确定：

$$\arg \min \left[\gamma \left(\left| \angle p_i - 90^\circ \right| + \left| \angle p_{i+1} - 90^\circ \right| \right) + \left| \angle p_i + \angle p_{i+1} - 180^\circ \right| \right] \quad (6)$$

其中，在式(6)中， $\angle p_i$ 为第二边界关键点作为顶点，且边为右边界和上边界的角， $\angle p_{i+1}$ 为第三边界关键点作为顶点，且边为右边界和下边界的角， γ 为权重系数，通常设置为0.5。式(6)表示 $\left[\gamma \left(\left| \angle p_i - 90^\circ \right| + \left| \angle p_{i+1} - 90^\circ \right| \right) + \left| \angle p_i + \angle p_{i+1} - 180^\circ \right| \right]$ 最小时，得到 $\angle p_i$ 和 $\angle p_{i+1}$ 。使用式(6)能确定第二边界关键点，这是由于右上边界关键点和右下边界关键点所在角是近乎平行的，且右边界与上边界的夹角，接近于90度，右边界与下边界的夹角接近于90度。第三边界关键点为第二边界关键点的下一个边界关键点。这样，按照该方式可以确定出第二边界关键点和第三边界关键点。这样，可以确定目标文本区域的任一边界关键点。如图9所示，1号位置为第一边界关键点，2号位置为第二边界关键点，3号位置为第三边界关键点，4号位置为第四边界关键点。

对于纵向排布的文本区域，将所确定的边界关键点按顺时针的方向顺移一位，即1号位置对应右上边界关键点，2号位置对应右下边界关键点，依次类推。

步骤803，对目标文本区域进行分割处理，得到目标文本区域的第一边界区域、第二边界区域、第三边界区域、第四边界区域、文本所在区域和背景区域，将第一边界区域、第二边界区域、第三边界区域、第四边界区域、文本所在区域和背景区域作为分割真值。

在实施中，如图9所示，对于横向排布的文本区域，在确定出第一边界关键点、第二边界关键点、第三边界关键点和第四边界关键点之后，可以将目标文本区域中文本的像素设置为类别1(文本类)，即文本所在区域，将文本区域的左上边界关键点、左下边界关键点连线并向内扩张一定距离设置为类别2(头边界类)，即第一边界区域，将文本区域的右下边界关键点、右上边界关键点连

线向内扩张一定距离设置为类别 3 (尾边界类), 即第二边界区域, 其余边界线向内外同时扩张生成区域设置为类别 4 (上下边界类), 即第三边界区域 (上边界所在的区域) 和第四边界区域 (下边界所在的区域), 其余设置为类别 0 (背景类)。

将第一边界区域、第二边界区域、第三边界区域、第四边界区域和文本所在区域作为分割真值, 用于后续训练。

需要说明的是, 这里的后一个类别的生成会覆盖前一种类别。

步骤 804, 确定第一边界区域的各像素点与第一边界关键点、第三边界关键点的偏移, 并确定第二边界区域的各像素点与第二边界关键点、第四边界关键点的偏移, 作为边界关键点真值。

在实施中, 识别设备可以确定第一边界区域的各像素点与左上边界关键点的纵向偏移和横向偏移, 并确定第一边界区域的各像素点与左下边界关键点的纵向偏移和横向偏移, 并确定第二边界区域的各像素点与右上边界关键点的纵向偏移和横向偏移, 并确定第二边界区域的各像素点与右下边界关键点的纵向偏移和横向偏移。

识别设备将确定出的偏移, 确定为边界关键点真值。

步骤 805, 确定目标文本区域中文本所在区域的各像素点与目标文本区域的上边界、下边界、左边界和右边界的最小偏移, 作为边界偏移真值。

在实施中, 识别设备可以确定文本所在区域的各像素点与目标文本的上边界、下边界、左边界和右边界最近的偏移, 对于任一边界 (上边界、下边界、左边界和右边界中的一种), 以及任一像素点, 可以使用该像素点的位置坐标和该边界上任一像素点的位置坐标, 求该像素点与该边界上任一像素点之间的距离, 确定与该边界上各像素点之间的距离的最小值, 即为该像素点与该边界的最小偏移。依此类推, 可以确定出目标文本区域中文本所在区域的各像素点与目标文本区域的上边界、下边界、左边界和右边界的最小偏移。

将确定出的偏移, 确定为边界偏移真值。

步骤 806, 根据样本集合中每个文本区域对应的分割真值、边界关键点真值和边界偏移真值, 对分割网络训练模型进行训练, 得到分割网络。

其中, 分割网络训练模型也是 FPN。

在实施中, 识别设备可以使用样本集合中每个文本区域对应的分割真值、

边界关键点真值和边界偏移真值，对预设的分割网络训练模型进行训练，得到分割网络训练模型的各参数值，将参数值代入到分割网络训练模型，即可得到分割网络。

需要说明的是，上述边界关键点真值和边界区域真值产生损失均为 L1 损失，表述为：

$$L = \begin{cases} 0.5(\sigma z)^2, & |z| < \frac{1}{\sigma} \\ |z| - \frac{0.5}{\sigma}, & |z| \geq \frac{1}{\sigma} \end{cases} \quad (7)$$

其中，在式 (7) 中，z 表示偏移距离， σ 为常数，取值可以为 0.3。

在训练时，分割真值、边界关键点真值和边界偏移真值对应的损失比值设为 10:1:1。

需要说明的是，在上述描述中主要是针对弯曲文本进行描述，实际上本公开实施例也可以应用于识别非弯曲文本（即任何条形码的检测和识别），处理过程与前文中对弯曲文本的处理过程相同，此处不再赘述。

本公开实施例中，识别设备可以根据提取待识别图像的特征图，然后使用预设的分割网络和待识别图像，确定待识别图像的文本区域的分割信息，然后根据分割信息，确定文本区域的边界关键点，使用边界关键点，将文本区域中的文本转换为目标排列顺序的文本，将转换得到的文本输入到预设识别模型，进行识别处理。这样，由于是确定文本区域的边界关键点，使用边界关键点，将文本区域中的文本转换为目标排列顺序的文本，再进行识别，所以可以识别任意形状的弯曲文本，进而可以提升弯曲文本的识别效率。

而且本公开实施例中，不需要对弯曲文本进行字符级别的标注，也可以学习到文本阅读的语义信息（排布信息），所以可以增加文本检测和识别的准确率。

而且本公开实施例中，由于可以产生固定数目个边界关键点的检测框，所以可以被应用于任何形状文本的标定工具，适用范围广。

而且本公开实施例中，基于分割的识别方法，除了增加边界区域的类别，还将边界区域分为头边界区域、尾边界区域，所以可以将属于不同文本区域的文本块分割开，而且可以学习到文本的排列顺序。

基于相同的技术构思，本公开实施例还提供了一种识别文本的装置，如图 10 所示，该装置包括：

提取模块 1010，用于根据预设的特征提取网络和待识别图像，提取所述待识别图像的特征图；

确定模块 1020，用于根据预设的分割网络和所述特征图，确定所述待识别图像的文本区域的分割信息；根据所述分割信息，确定所述文本区域的边界关键点；

转换模块 1030，用于根据所述边界关键点，将所述文本区域中的文本转换为目标排列顺序的文本；

识别模块 1040，用于将转换得到的文本输入到预设识别模型，进行识别处理。

可选的，所述确定模块 1020，用于：

根据所述分割信息中第一边界区域中各像素点与所述第一边界区域的两个边界关键点之间的偏移，确定所述第一边界区域的两个边界关键点的位置信息；并根据所述分割信息中第二边界区域中各像素点与所述第二边界区域的两个边界关键点之间的偏移，确定所述第二边界区域的两个边界关键点的位置信息，其中，所述第一边界区域位于所述文本区域的头部，所述第二边界区域位于所述文本区域的尾部；

根据所述第一边界区域的两个边界关键点的位置信息和所述第二边界区域的两个边界关键点的位置信息，确定所述文本区域中除所述第一边界区域与所述第二边界区域之外的其它边界关键点。

可选的，所述确定模块 1020，还用于：

确定所述文本区域中的文本的排布信息；

所述确定模块 1020，用于：

根据所述文本的排布信息、所述第一边界区域的两个边界关键点的位置信息和所述第二边界区域的两个边界关键点的位置信息，确定所述文本区域中除所述第一边界区域与所述第二边界区域之外的其它边界关键点。

可选的，所述确定模块 1020，用于：

如果所述排布信息为横向排布，则根据所述第一边界区域中第一边界关键点的位置信息和所述第二边界区域中第二边界关键点的位置信息，确定所述文本区域的上边界上除所述第一边界关键点和所述第二边界关键点之外的其它边界关键点，并根据所述第一边界区域中第四边界关键点的位置信息和所述第二

边界区域中第三边界关键点的位置信息，确定所述文本区域的下边界上除所述第三边界关键点和所述第三边界关键点之外的其它边界关键点；

如果所述排布信息为纵向排布，则根据所述第一边界区域中第一边界关键点的位置信息和所述第二边界区域中第二边界关键点的位置信息，确定所述文本区域的左边界上除所述第一边界关键点和所述第二边界关键点之外的其它边界关键点，并根据所述第一边界区域中第四边界关键点的位置信息和所述第二边界区域中第三边界关键点的位置信息，确定所述文本区域的右边界上除所述第三边界关键点和所述第四边界关键点之外的其它边界关键点。

可选的，所述确定模块 1020，用于：

根据所述第一边界区域中第一边界关键点的位置信息和所述第二边界区域中第二边界关键点的位置信息，确定第一预设数目个位置点的第一横坐标信息，其中，所述第一预设数目个位置点位于第一连线上，且是基于将所述第一连线等分为目标数值得到，所述目标数值等于所述第一预设数目加 1，所述第一连线为所述第一边界关键点与所述第二边界关键点的连线；

根据所述第一横坐标信息，确定所述文本区域的上边界上除所述第一边界关键点和所述第二边界关键点之外的其它边界关键点。

可选的，所述分割信息包括所述文本区域中属于文本的像素点与所述文本区域的上边界的偏移；

所述确定模块 1020，用于：

对于所述第一预设数目个位置点中任一位置点的第一横坐标信息，根据所述位置点的第一横坐标信息对应的横坐标区间，确定所述文本区域中属于文本且横坐标位于所述横坐标区间的像素点；

将所述位置点的第一横坐标信息，确定为所述位置点对应的边界关键点的横坐标，并根据确定出的像素点与所述上边界的偏移，确定所述位置点对应的边界关键点的纵坐标，其中，所述位置点对应的边界关键点为位于所述文本区域的上边界，且所述上边界包括所述第一边界关键点和所述第二边界关键点。

可选的，所述确定模块 1020，用于：

根据所述第一边界区域中第四边界关键点的位置信息和所述第二边界区域中第三边界关键点的位置信息，确定第一预设数目个位置点的第二横坐标信息，其中，所述第一预设数目个位置点位于第二连线上，且是基于将所述第二连线

等分为目标数值得到，所述目标数值等于所述第一预设数目加 1，所述第二连线为所述第三边界关键点与所述第四边界关键点的连线；

根据所述第二横坐标信息，确定所述文本区域的下边界上除所述第三边界关键点和所述第三边界关键点之外的其它边界关键点。

可选的，所述分割信息包括所述文本区域中属于文本的像素点与所述文本区域的下边界的偏移；

所述确定模块 1020，用于：

对于所述第一预设数目个位置点中任一位置点的第二横坐标信息，根据所述位置点的第二横坐标信息对应的横坐标区间，确定所述文本区域中属于文本且横坐标位于所述横坐标区间的像素点；

将所述位置点的第二横坐标信息，确定为所述位置点对应的边界关键点的横坐标，并根据确定出的像素点与所述下边界的偏移，确定所述位置点对应的边界关键点的纵坐标，其中，所述位置点对应的边界关键点为位于所述文本区域的下边界，且所述下边界包括所述第三边界关键点和所述第四边界关键点。

可选的，所述转换模块 1030，用于：

使用薄板样条插值 TPS 对所述边界关键点进行模板匹配，将所述文本区域的文本转换为目标排列顺序的文本。

可选的，所述提取模块 1010，用于：

将待识别图像输入到预设的特征提取网络，提取所述待识别图像的预设尺寸的特征图；

将所述预设尺寸的特征图经过预设的卷积处理；

将卷积处理后的特征图进行卷积核矫正处理；

将卷积核矫正处理后的特征图经过上采样后与所述预设尺寸的特征图的上一层特征图进行融合处理，得到所述待识别图像的特征图。

可选的，所述提取模块 1010，用于：

将卷积处理后的特征图经过第一卷积核的卷积处理，得到第一特征图，将所述第一特征图进行第二卷积核的卷积处理，得到第二特征图；

将所述第二特征图进行第三卷积核的卷积处理，得到第三特征图，并将所述第二特征图进行第四卷积核的卷积处理，得到第四特征图；其中，所述第三卷积核的尺寸为 $m*n$ ，所述第四卷积核的尺寸为 $n*m$ ， m 和 n 为正整数，且 m

小于 n ;

将所述第一特征图、第二特征图、第三特征图，第四特征图进行通道拼接处理，得到第五特征图;

将所述第五特征图进行压缩激活处理，得到卷积核矫正处理后的特征图。

可选的，如图 11 所示，所述装置还包括训练模块 1050，用于:

获取样本集合，其中，所述样本集合中包括第二预设数目个标定文本区域的图像;

对于所述样本集合中的目标文本区域，确定所述目标文本区域的第一边界关键点、第二边界关键点、第三边界关键点和第四边界关键点;

对所述目标文本区域进行分割处理，得到所述目标文本区域的第一边界区域、第二边界区域、第三边界区域、第四边界区域、文本所在区域和背景区域，将所述第一边界区域、第二边界区域、第三边界区域、第四边界区域、文本所在区域和背景区域作为分割真值;

确定所述第一边界区域的各像素点与所述第一边界关键点、所述第三边界关键点的偏移，并确定第二边界区域的各像素点与所述第二边界关键点、所述第四边界关键点的偏移，作为边界关键点真值;

确定所述目标文本区域中文本所在区域的各像素点与所述目标文本区域的上边界、下边界、左边界和右边界的最小偏移，作为边界偏移真值;

根据所述样本集合中每个文本区域对应的分割真值、边界关键点真值和边界偏移真值，对分割网络训练模型进行训练，得到分割网络。

可选的，所述确定模块 1020，用于:

确定所述文本区域中的最大横坐标和最大纵坐标，并确定所述文本区域中的最小横坐标和最小纵坐标;

根据所述最大横坐标和最大纵坐标、所述最小横坐标和最小纵坐标，确定所述文本区域的外接矩形;

如果所述外接矩形的长度大于宽度，则确定所述文本区域中文本的排布为横向排布，如果所述外接矩形的长度小于宽度，则确定所述文本区域中文本的排布为纵向排布。

本公开实施例中，识别设备可以根据提取待识别图像的特征图，然后使用预设的分割网络和待识别图像，确定待识别图像的文本区域的分割信息，然后

根据分割信息，确定文本区域的边界关键点，使用边界关键点，将文本区域中的文本转换为目标排列顺序的文本，将转换得到的文本输入到预设识别模型，进行识别处理。这样，由于是确定文本区域的边界关键点，使用边界关键点，将文本区域中的文本转换为目标排列顺序的文本，再进行识别，所以可以识别任意形状的弯曲文本，进而可以提升弯曲文本的识别效率。

需要说明的是：上述实施例提供的识别文本的装置在识别文本时，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，上述实施例提供的识别文本的装置与识别文本的方法实施例属于同一构思，其具体实现过程详见方法实施例，这里不再赘述。

图 12 是本公开实施例提供的一种识别设备的结构示意图，该识别设备 1200 可因配置或性能不同而产生比较大的差异，可以包括一个或一个以上处理器（central processing units, CPU）1201 和一个或一个以上的存储器 1202，其中，所述存储器 1202 中存储有至少一条计算机指令，所述至少一条计算机指令由所述处理器 1201 加载并执行以实现上述识别文本的方法的步骤。

本公开实施例中，还提供了一种计算机可读存储介质，所述存储介质内存储有计算机程序，所述计算机程序被处理器执行时实现上述识别文本的方法步骤。

本公开实施例中，还提供了一种识别文本的识别设备，包括处理器和存储器，其中，所述存储器，用于存放计算机程序；所述处理器，用于执行所述存储器上所存放的程序，实现上述识别文本的方法步骤。

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。

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

权 利 要 求 书

1、一种识别文本的方法，其特征在于，所述方法包括：

根据预设的特征提取网络和待识别图像，提取所述待识别图像的特征图；

根据预设的分割网络和所述特征图，确定所述待识别图像的文本区域的分割信息；

根据所述分割信息，确定所述文本区域的边界关键点；

根据所述边界关键点，将所述文本区域中的文本转换为目标排列顺序的文本；

将转换得到的文本输入到预设识别模型，进行识别处理。

2、根据权利要求1所述的方法，其特征在于，所述根据所述分割信息，确定所述文本区域的边界关键点，包括：

根据所述分割信息中第一边界区域中各像素点与所述第一边界区域的两个边界关键点之间的偏移，确定所述第一边界区域的两个边界关键点的位置信息；并根据所述分割信息中第二边界区域中各像素点与所述第二边界区域的两个边界关键点之间的偏移，确定所述第二边界区域的两个边界关键点的位置信息，其中，所述第一边界区域位于所述文本区域的头部，所述第二边界区域位于所述文本区域的尾部；

根据所述第一边界区域的两个边界关键点的位置信息和所述第二边界区域的两个边界关键点的位置信息，确定所述文本区域中除所述第一边界区域与所述第二边界区域之外的其它边界关键点。

3、根据权利要求2所述的方法，其特征在于，所述方法还包括：

确定所述文本区域中的文本的排布信息；

所述根据所述第一边界区域的两个边界关键点的位置信息和所述第二边界区域的两个边界关键点的位置信息，确定所述文本区域中除所述第一边界区域与所述第二边界区域之外的其它边界关键点，包括：

根据所述文本的排布信息、所述第一边界区域的两个边界关键点的位置信息和所述第二边界区域的两个边界关键点的位置信息，确定所述文本区域中除所述第一边界区域与所述第二边界区域之外的其它边界关键点。

4、根据权利要求3所述的方法，其特征在于，所述根据所述文本的排布信

息、所述第一边界区域的两个边界关键点的位置信息和所述第二边界区域的两个边界关键点的位置信息，确定所述文本区域中除所述第一边界区域与所述第二边界区域之外的其它边界关键点，包括：

如果所述排布信息为横向排布，则根据所述第一边界区域中第一边界关键点的位置信息和所述第二边界区域中第二边界关键点的位置信息，确定所述文本区域的上边界上除所述第一边界关键点和所述第二边界关键点之外的其它边界关键点，并根据所述第一边界区域中第四边界关键点的位置信息和所述第二边界区域中第三边界关键点的位置信息，确定所述文本区域的下边界上除所述第三边界关键点和所述第三边界关键点之外的其它边界关键点；

如果所述排布信息为纵向排布，则根据所述第一边界区域中第一边界关键点的位置信息和所述第二边界区域中第二边界关键点的位置信息，确定所述文本区域的左边界上除所述第一边界关键点和所述第二边界关键点之外的其它边界关键点，并根据所述第一边界区域中第四边界关键点的位置信息和所述第二边界区域中第三边界关键点的位置信息，确定所述文本区域的右边界上除所述第三边界关键点和所述第四边界关键点之外的其它边界关键点。

5、根据权利要求4所述的方法，其特征在于，所述根据所述第一边界区域中第一边界关键点的位置信息和所述第二边界区域中第二边界关键点的位置信息，确定所述文本区域的左边界上除所述第一边界关键点和所述第二边界关键点之外的其它边界关键点，包括：

根据所述第一边界区域中第一边界关键点的位置信息和所述第二边界区域中第二边界关键点的位置信息，确定第一预设数目个位置点的第一横坐标信息，其中，所述第一预设数目个位置点位于第一连线上，且是基于将所述第一连线等分为目标数值得到，所述目标数值等于所述第一预设数目加1，所述第一连线为所述第一边界关键点与所述第二边界关键点的连线；

根据所述第一横坐标信息，确定所述文本区域的上边界上除所述第一边界关键点和所述第二边界关键点之外的其它边界关键点。

6、根据权利要求5所述的方法，其特征在于，所述分割信息包括所述文本区域中属于文本的像素点与所述文本区域的上边界的偏移；

所述根据所述第一横坐标信息，确定所述文本区域的上边界上除所述第一边界关键点和所述第二边界关键点之外的其它边界关键点，包括：

对于所述第一预设数目个位置点中任一位置点的第一横坐标信息，根据所述位置点的第一横坐标信息对应的横坐标区间，确定所述文本区域中属于文本且横坐标位于所述横坐标区间的像素点；

将所述位置点的第一横坐标信息，确定为所述位置点对应的边界关键点的横坐标，并根据确定出的像素点与所述上边界的偏移，确定所述位置点对应的边界关键点的纵坐标，其中，所述位置点对应的边界关键点为位于所述文本区域的上边界，且所述上边界包括所述第一边界关键点和所述第二边界关键点。

7、根据权利要求4所述的方法，其特征在于，所述根据所述第一边界区域中第四边界关键点的位置信息和所述第二边界区域中第三边界关键点的位置信息，确定所述文本区域的下边界上除所述第三边界关键点和所述第三边界关键点之外的其它边界关键点，包括：

根据所述第一边界区域中第四边界关键点的位置信息和所述第二边界区域中第三边界关键点的位置信息，确定第一预设数目个位置点的第二横坐标信息，其中，所述第一预设数目个位置点位于第二连线上，且是基于将所述第二连线等分为目标数值得到，所述目标数值等于所述第一预设数目加1，所述第二连线为所述第三边界关键点与所述第四边界关键点的连线；

根据所述第二横坐标信息，确定所述文本区域的下边界上除所述第三边界关键点和所述第三边界关键点之外的其它边界关键点。

8、根据权利要求7所述的方法，其特征在于，所述分割信息包括所述文本区域中属于文本的像素点与所述文本区域的下边界的偏移；

所述根据所述第二横坐标信息，确定所述文本区域的下边界上除所述第三边界关键点和所述第三边界关键点之外的其它边界关键点，包括：

对于所述第一预设数目个位置点中任一位置点的第二横坐标信息，根据所述位置点的第二横坐标信息对应的横坐标区间，确定所述文本区域中属于文本且横坐标位于所述横坐标区间的像素点；

将所述位置点的第二横坐标信息，确定为所述位置点对应的边界关键点的横坐标，并根据确定出的像素点与所述下边界的偏移，确定所述位置点对应的边界关键点的纵坐标，其中，所述位置点对应的边界关键点为位于所述文本区域的下边界，且所述下边界包括所述第三边界关键点和所述第四边界关键点。

9、根据权利要求1至8任一项所述的方法，其特征在于，所述根据所述边

界关键点，将所述文本区域中的文本转换为目标排列顺序的文本，包括：

使用薄板样条插值 TPS 对所述边界关键点进行模板匹配，将所述文本区域的文本转换为目标排列顺序的文本。

10、根据权利要求 1 至 8 任一所述的方法，其特征在于，所述根据预设的特征提取网络和待识别图像，提取所述待识别图像的特征图，包括：

将待识别图像输入到预设的特征提取网络，提取所述待识别图像的预设尺寸的特征图；

将所述预设尺寸的特征图经过预设的卷积处理；

将卷积处理后的特征图进行卷积核矫正处理；

将卷积核矫正处理后的特征图经过上采样后与所述预设尺寸的特征图的上一层特征图进行融合处理，得到所述待识别图像的特征图。

11、根据权利要求 10 所述的方法，其特征在于，所述将卷积处理后的特征图进行卷积核矫正处理，包括：

将卷积处理后的特征图经过第一卷积核的卷积处理，得到第一特征图，将所述第一特征图进行第二卷积核的卷积处理，得到第二特征图；

将所述第二特征图进行第三卷积核的卷积处理，得到第三特征图，并将所述第二特征图进行第四卷积核的卷积处理，得到第四特征图；其中，所述第三卷积核的尺寸为 $m*n$ ，所述第四卷积核的尺寸为 $n*m$ ， m 和 n 为正整数，且 m 小于 n ；

将所述第一特征图、第二特征图、第三特征图，第四特征图进行通道拼接处理，得到第五特征图；

将所述第五特征图进行压缩激活处理，得到卷积核矫正处理后的特征图。

12、根据权利要求 1 至 8 任一项所述的方法，其特征在于，所述方法还包括：

获取样本集合，其中，所述样本集合中包括第二预设数目个标定文本区域的图像；

对于所述样本集合中的目标文本区域，确定所述目标文本区域的第一边界关键点、第二边界关键点、第三边界关键点和第四边界关键点；

对所述目标文本区域进行分割处理，得到所述目标文本区域的第一边界区域、第二边界区域、第三边界区域、第四边界区域、文本所在区域和背景区域，

将所述第一边界区域、第二边界区域、第三边界区域、第四边界区域、文本所在区域和背景区域作为分割真值；

确定所述第一边界区域的各像素点与所述第一边界关键点、所述第三边界关键点的偏移，并确定第二边界区域的各像素点与所述第二边界关键点、所述第四边界关键点的偏移，作为边界关键点真值；

确定所述目标文本区域中文本所在区域的各像素点与所述目标文本区域的上边界、下边界、左边界和右边界的最小偏移，作为边界偏移真值；

根据所述样本集合中每个文本区域对应的分割真值、边界关键点真值和边界偏移真值，对分割网络训练模型进行训练，得到分割网络。

13、根据权利要求 3 所述的方法，其特征在于，所述确定所述文本区域中的文本的排布信息，包括：

确定所述文本区域中的最大横坐标和最大纵坐标，并确定所述文本区域中的最小横坐标和最小纵坐标；

根据所述最大横坐标和最大纵坐标、所述最小横坐标和最小纵坐标，确定所述文本区域的外接矩形；

如果所述外接矩形的长度大于宽度，则确定所述文本区域中文本的排布为横向排布，如果所述外接矩形的长度小于宽度，则确定所述文本区域中文本的排布为纵向排布。

14、一种识别文本的装置，其特征在于，所述装置包括：

提取模块，用于根据预设的特征提取网络和待识别图像，提取所述待识别图像的特征图；

确定模块，用于根据预设的分割网络和所述特征图，确定所述待识别图像的文本区域的分割信息；根据所述分割信息，确定所述文本区域的边界关键点；

转换模块，用于根据所述边界关键点，将所述文本区域中的文本转换为目标排列顺序的文本；

识别模块，用于将转换得到的文本输入到预设识别模型，进行识别处理。

15、根据权利要求 14 所述的装置，其特征在于，所述确定模块，用于：

根据所述分割信息中第一边界区域中各像素点与所述第一边界区域的两个边界关键点之间的偏移，确定所述第一边界区域的两个边界关键点的位置信息；

并根据所述分割信息中第二边界区域中各像素点与所述第二边界区域的两个边界关键点之间的偏移，确定所述第二边界区域的两个边界关键点的位置信息，其中，所述第一边界区域位于所述文本区域的头部，所述第二边界区域位于所述文本区域的尾部；

根据所述第一边界区域的两个边界关键点的位置信息和所述第二边界区域的两个边界关键点的位置信息，确定所述文本区域中除所述第一边界区域与所述第二边界区域之外的其它边界关键点。

16、根据权利要求 15 所述的装置，其特征在于，所述确定模块，还用于：
确定所述文本区域中的文本的排布信息；

所述确定模块，用于：

根据所述文本的排布信息、所述第一边界区域的两个边界关键点的位置信息和所述第二边界区域的两个边界关键点的位置信息，确定所述文本区域中除所述第一边界区域与所述第二边界区域之外的其它边界关键点。

17、根据权利要求 16 所述的装置，其特征在于，所述确定模块，用于：

如果所述排布信息为横向排布，则根据所述第一边界区域中第一边界关键点的位置信息和所述第二边界区域中第二边界关键点的位置信息，确定所述文本区域的上边界上除所述第一边界关键点和所述第二边界关键点之外的其它边界关键点，并根据所述第一边界区域中第四边界关键点的位置信息和所述第二边界区域中第三边界关键点的位置信息，确定所述文本区域的下边界上除所述第三边界关键点和所述第四边界关键点之外的其它边界关键点；

如果所述排布信息为纵向排布，则根据所述第一边界区域中第一边界关键点的位置信息和所述第二边界区域中第二边界关键点的位置信息，确定所述文本区域的左边界上除所述第一边界关键点和所述第二边界关键点之外的其它边界关键点，并根据所述第一边界区域中第四边界关键点的位置信息和所述第二边界区域中第三边界关键点的位置信息，确定所述文本区域的右边界上除所述第三边界关键点和所述第四边界关键点之外的其它边界关键点。

18、根据权利要求 17 所述的装置，其特征在于，所述确定模块，用于：

根据所述第一边界区域中第一边界关键点的位置信息和所述第二边界区域中第二边界关键点的位置信息，确定第一预设数目个位置点的第一横坐标信息，其中，所述第一预设数目个位置点位于第一连线上，且是基于将所述第一连线

等分为目标数值得到，所述目标数值等于所述第一预设数目加 1，所述第一连线为所述第一边界关键点与所述第二边界关键点的连线；

根据所述第一横坐标信息，确定所述文本区域的上边界上除所述第一边界关键点和所述第二边界关键点之外的其它边界关键点。

19、根据权利要求 18 所述的装置，其特征在于，所述分割信息包括所述文本区域中属于文本的像素点与所述文本区域的上边界的偏移；

所述确定模块，用于：

对于所述第一预设数目个位置点中任一位置点的第一横坐标信息，根据所述位置点的第一横坐标信息对应的横坐标区间，确定所述文本区域中属于文本且横坐标位于所述横坐标区间的像素点；

将所述位置点的第一横坐标信息，确定为所述位置点对应的边界关键点的横坐标，并根据确定出的像素点与所述上边界的偏移，确定所述位置点对应的边界关键点的纵坐标，其中，所述位置点对应的边界关键点为位于所述文本区域的上边界，且所述上边界包括所述第一边界关键点和所述第二边界关键点。

20、根据权利要求 17 所述的装置，其特征在于，所述确定模块，用于：

根据所述第一边界区域中第四边界关键点的位置信息和所述第二边界区域中第三边界关键点的位置信息，确定第一预设数目个位置点的第二横坐标信息，其中，所述第一预设数目个位置点位于第二连线上，且是基于将所述第二连线等分为目标数值得到，所述目标数值等于所述第一预设数目加 1，所述第二连线为所述第三边界关键点与所述第四边界关键点的连线；

根据所述第二横坐标信息，确定所述文本区域的下边界上除所述第三边界关键点和所述第三边界关键点之外的其它边界关键点。

21、根据权利要求 20 所述的装置，其特征在于，所述分割信息包括所述文本区域中属于文本的像素点与所述文本区域的下边界的偏移；

所述确定模块，用于：

对于所述第一预设数目个位置点中任一位置点的第二横坐标信息，根据所述位置点的第二横坐标信息对应的横坐标区间，确定所述文本区域中属于文本且横坐标位于所述横坐标区间的像素点；

将所述位置点的第二横坐标信息，确定为所述位置点对应的边界关键点的横坐标，并根据确定出的像素点与所述下边界的偏移，确定所述位置点对应的

边界关键点的纵坐标，其中，所述位置点对应的边界关键点为位于所述文本区域的下边界，且所述下边界包括所述第三边界关键点和所述第四边界关键点。

22、根据权利要求 14 至 21 任一项所述的装置，其特征在于，所述转换模块，用于：

使用薄板样条插值 TPS 对所述边界关键点进行模板匹配，将所述文本区域的文本转换为目标排列顺序的文本。

23、根据权利要求 14 至 21 任一所述的装置，其特征在于，所述提取模块，用于：

将待识别图像输入到预设的特征提取网络，提取所述待识别图像的预设尺寸的特征图；

将所述预设尺寸的特征图经过预设的卷积处理；

将卷积处理后的特征图进行卷积核矫正处理；

将卷积核矫正处理后的特征图经过上采样后与所述预设尺寸的特征图的上一层特征图进行融合处理，得到所述待识别图像的特征图。

24、根据权利要求 23 所述的装置，其特征在于，所述提取模块，用于：

将卷积处理后的特征图经过第一卷积核的卷积处理，得到第一特征图，将所述第一特征图进行第二卷积核的卷积处理，得到第二特征图；

将所述第二特征图进行第三卷积核的卷积处理，得到第三特征图，并将所述第二特征图进行第四卷积核的卷积处理，得到第四特征图；其中，所述第三卷积核的尺寸为 $m*n$ ，所述第四卷积核的尺寸为 $n*m$ ， m 和 n 为正整数，且 m 小于 n ；

将所述第一特征图、第二特征图、第三特征图，第四特征图进行通道拼接处理，得到第五特征图；

将所述第五特征图进行压缩激活处理，得到卷积核矫正处理后的特征图。

25、根据权利要求 14 至 21 任一项所述的装置，其特征在于，所述装置还包括：

训练模块，用于获取样本集合，其中，所述样本集合中包括第二预设数目个标定文本区域的图像；

对于所述样本集合中的目标文本区域，确定所述目标文本区域的第一边界关键点、第二边界关键点、第三边界关键点和第四边界关键点；

对所述目标文本区域进行分割处理，得到所述目标文本区域的第一边界区域、第二边界区域、第三边界区域、第四边界区域、文本所在区域和背景区域，将所述第一边界区域、第二边界区域、第三边界区域、第四边界区域、文本所在区域和背景区域作为分割真值；

确定所述第一边界区域的各像素点与所述第一边界关键点、所述第三边界关键点的偏移，并确定第二边界区域的各像素点与所述第二边界关键点、所述第四边界关键点的偏移，作为边界关键点真值；

确定所述目标文本区域中文本所在区域的各像素点与所述目标文本区域的上边界、下边界、左边界和右边界的最小偏移，作为边界偏移真值；

根据所述样本集合中每个文本区域对应的分割真值、边界关键点真值和边界偏移真值，对分割网络训练模型进行训练，得到分割网络。

26、根据权利要求 16 所述的装置，其特征在于，所述确定模块，用于：

确定所述文本区域中的最大横坐标和最大纵坐标，并确定所述文本区域中的最小横坐标和最小纵坐标；

根据所述最大横坐标和最大纵坐标、所述最小横坐标和最小纵坐标，确定所述文本区域的外接矩形；

如果所述外接矩形的长度大于宽度，则确定所述文本区域中文本的排布为横向排布，如果所述外接矩形的长度小于宽度，则确定所述文本区域中文本的排布为纵向排布。

27、一种识别设备，其特征在于，所述识别设备包括处理器和存储器，所述存储器中存储有至少一条计算机指令，所述至少一条计算机指令由所述处理器加载并执行以实现如权利要求 1 至权利要求 13 任一项所述的识别文本的方法所执行的操作。

28、一种计算机可读存储介质，其特征在于，所述计算机存储介质中存储有至少一条计算机指令，所述至少一条计算机指令由处理器加载并执行以实现如权利要求 1 至权利要求 13 任一项所述的识别文本的方法所执行的操作。

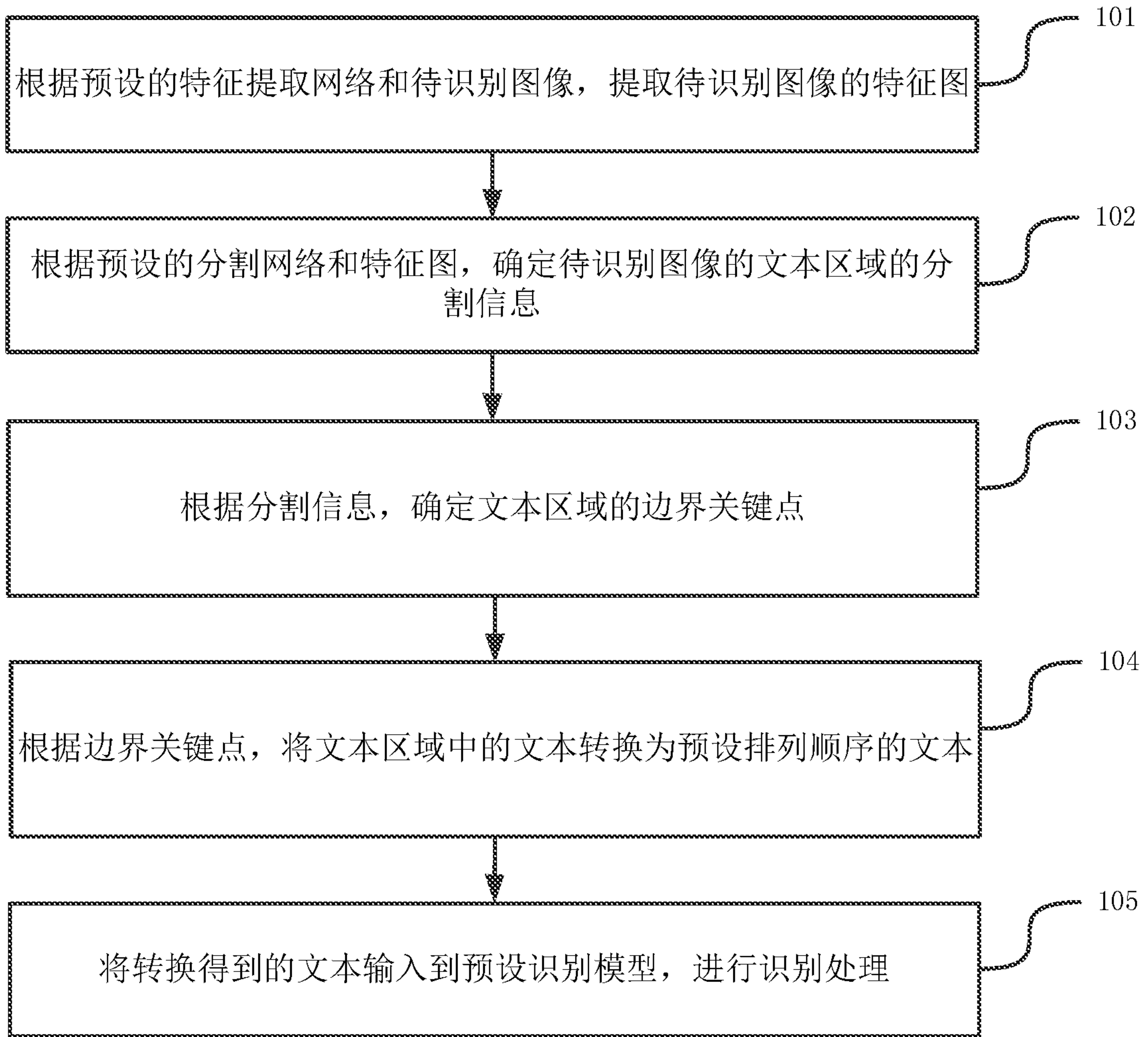


图 1

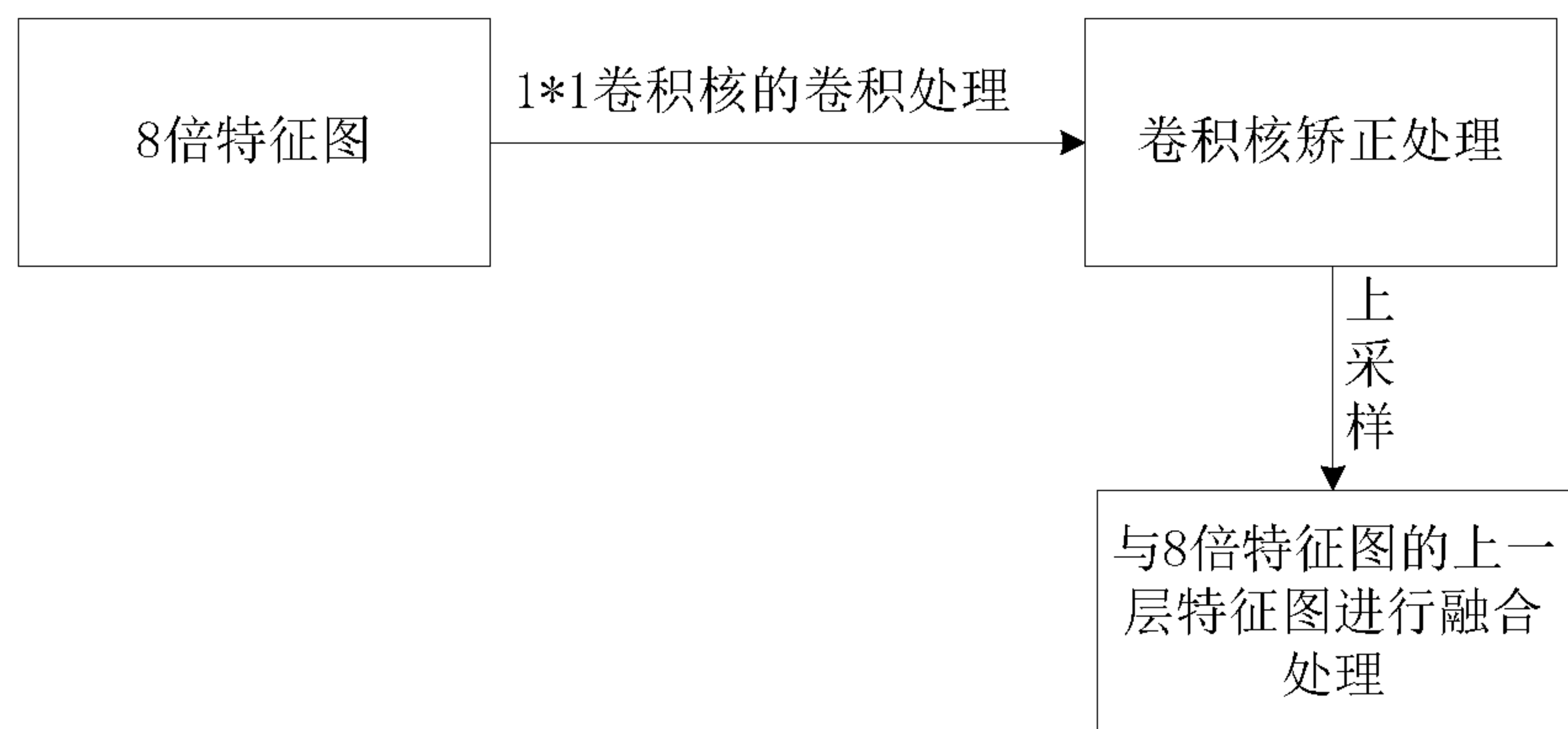


图 2

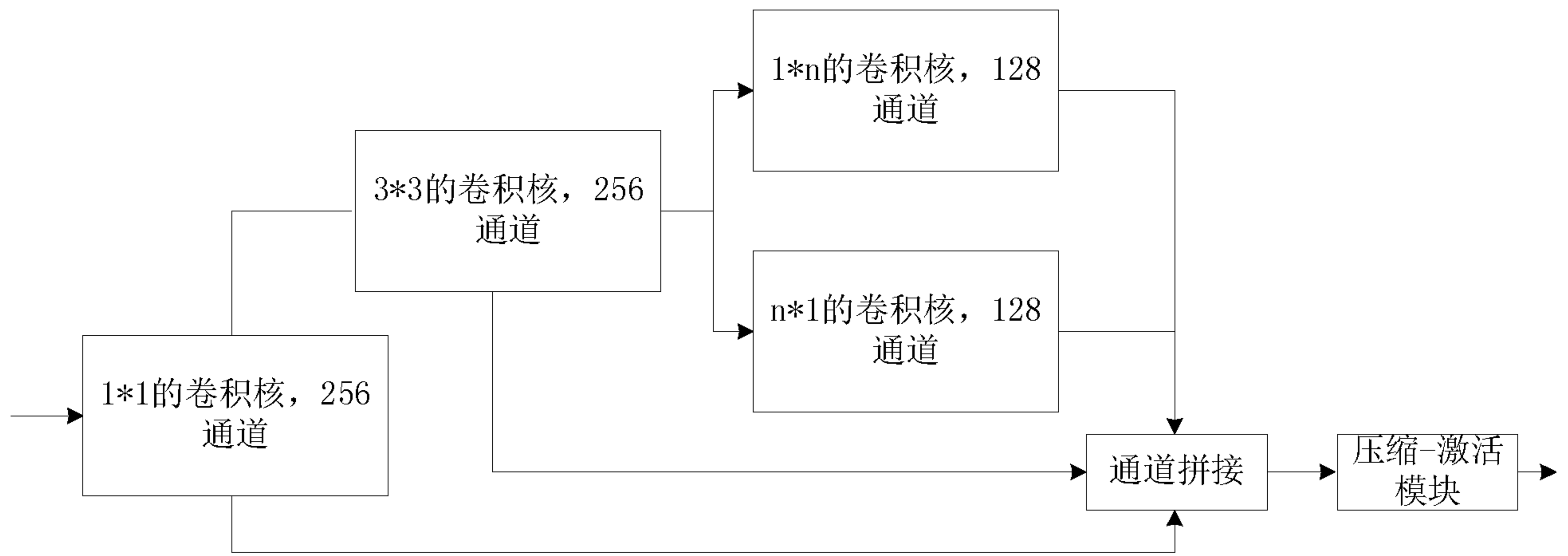


图 3

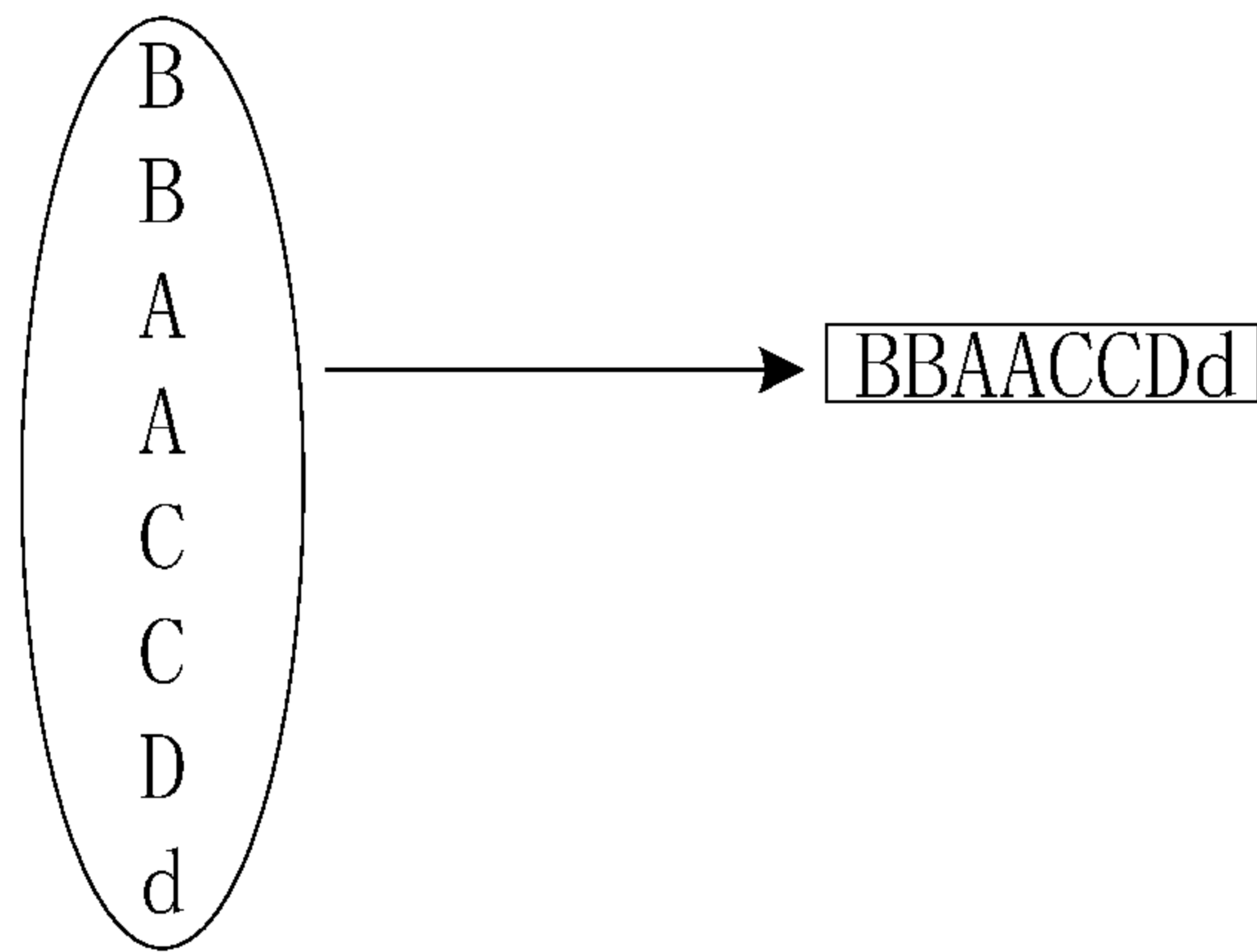


图 4

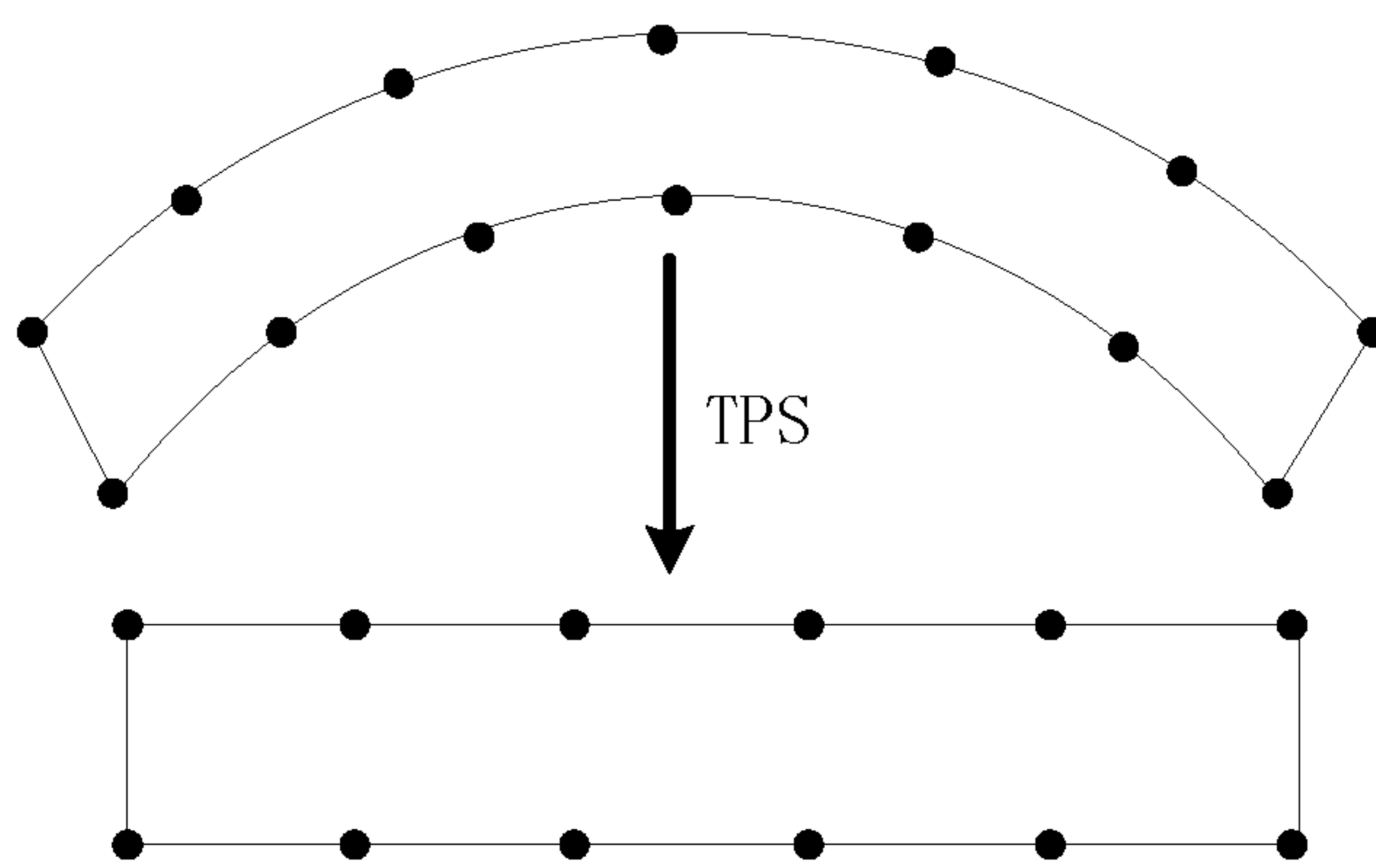


图 5

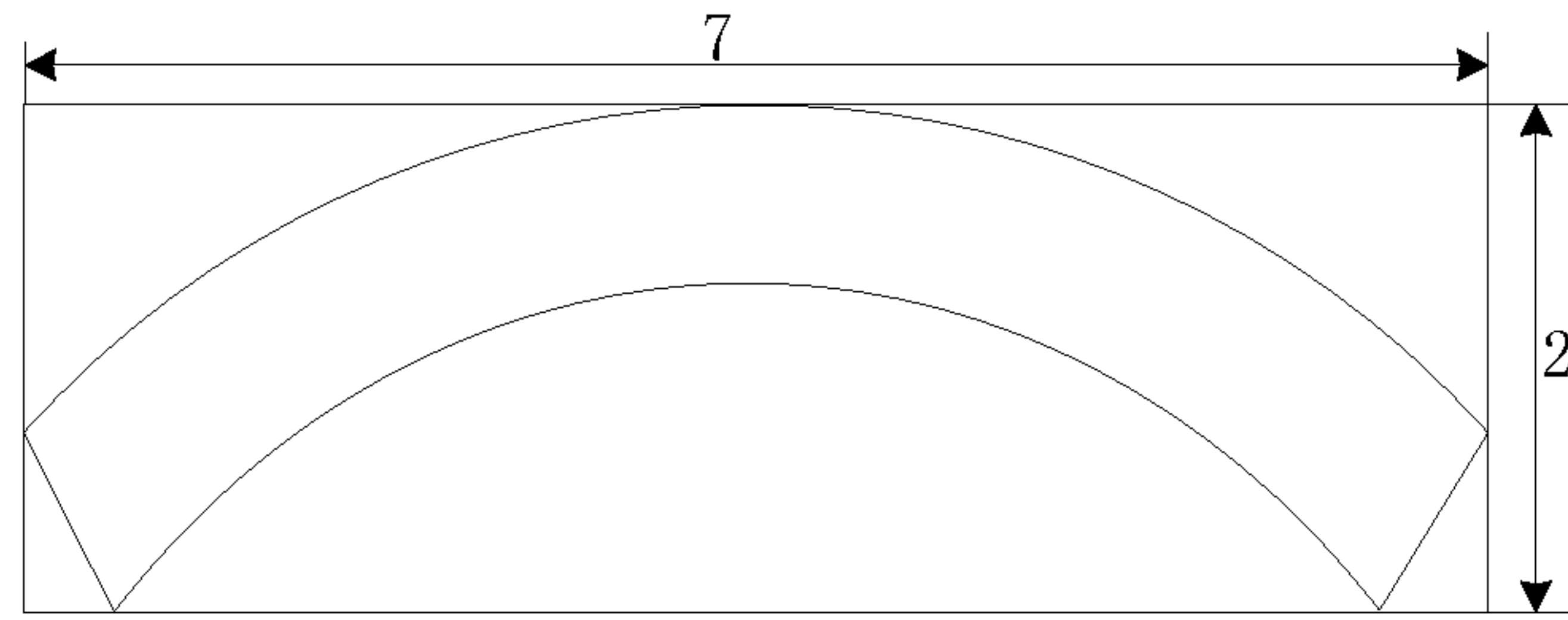


图 6

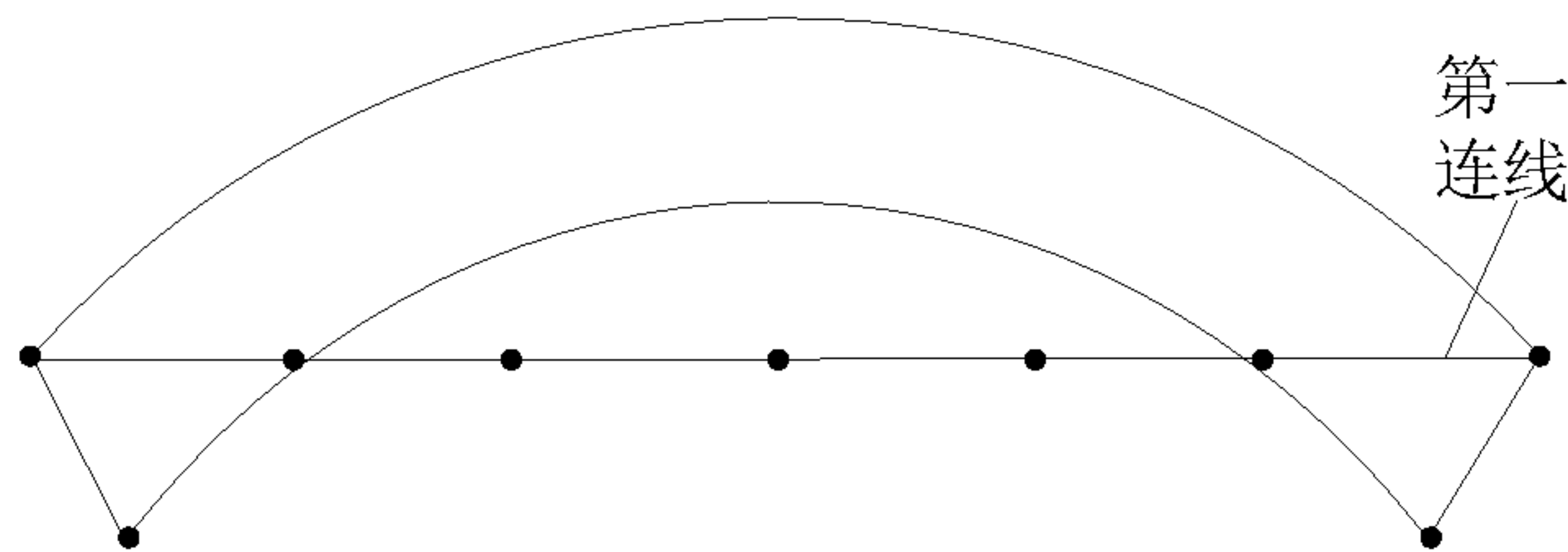


图 7

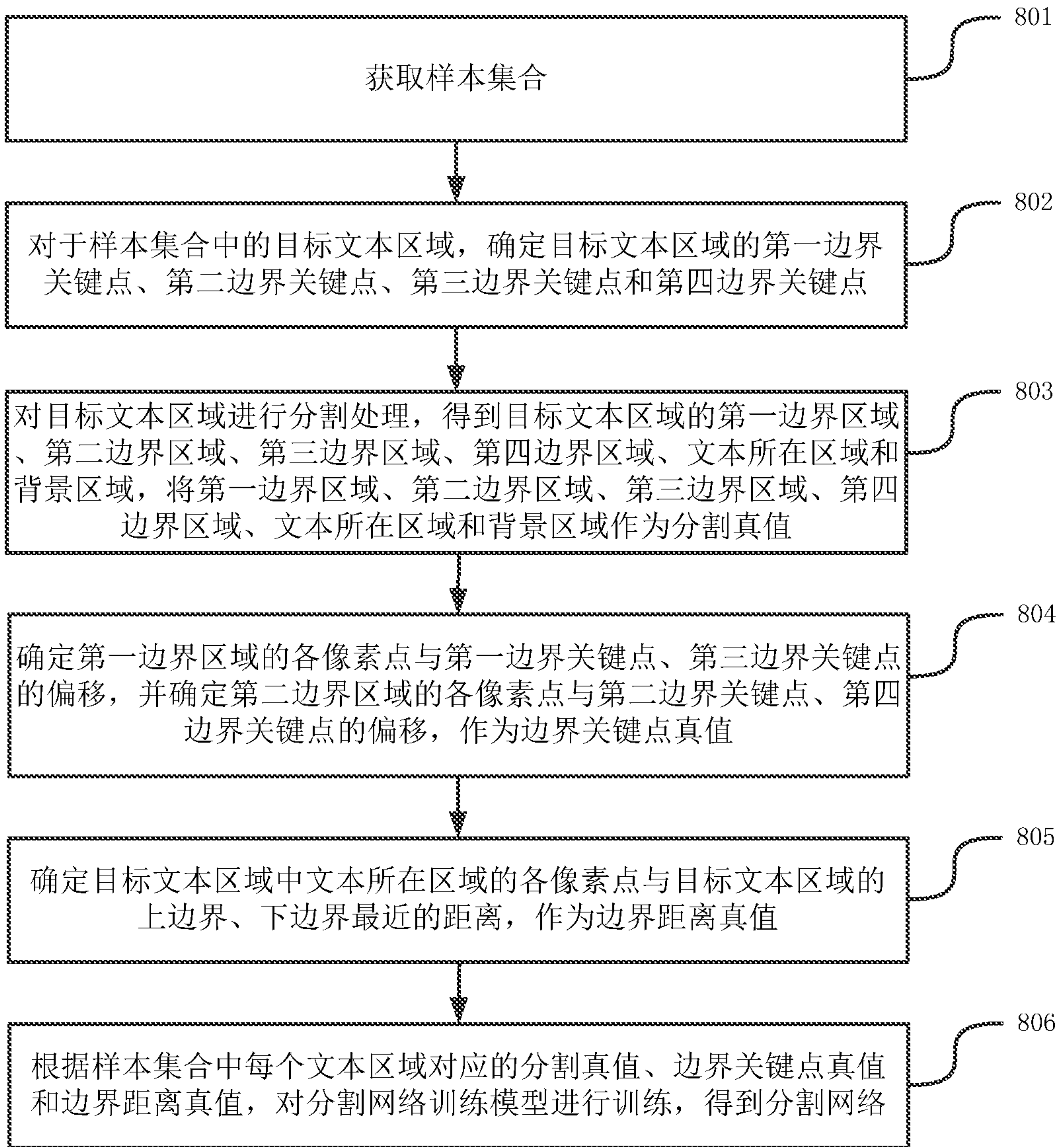


图 8

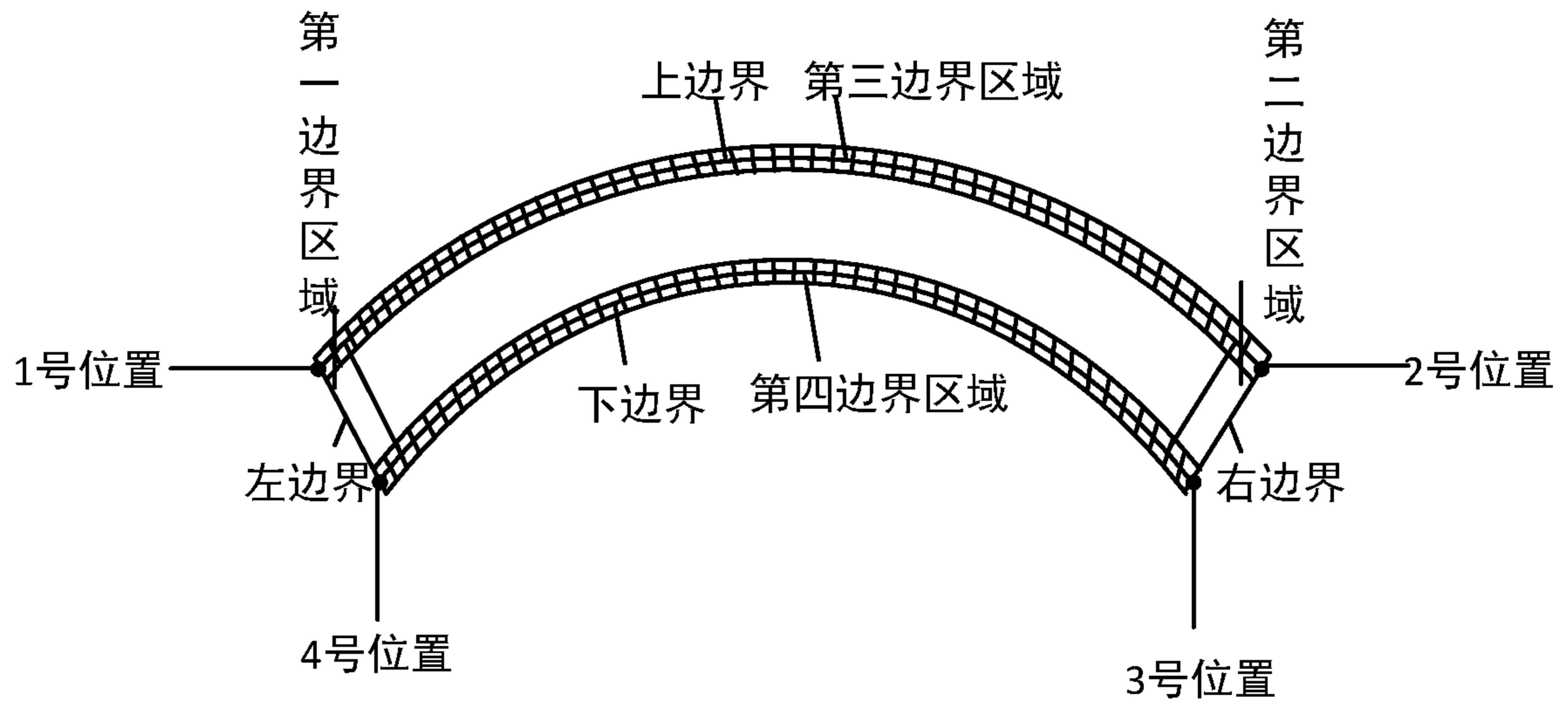


图 9

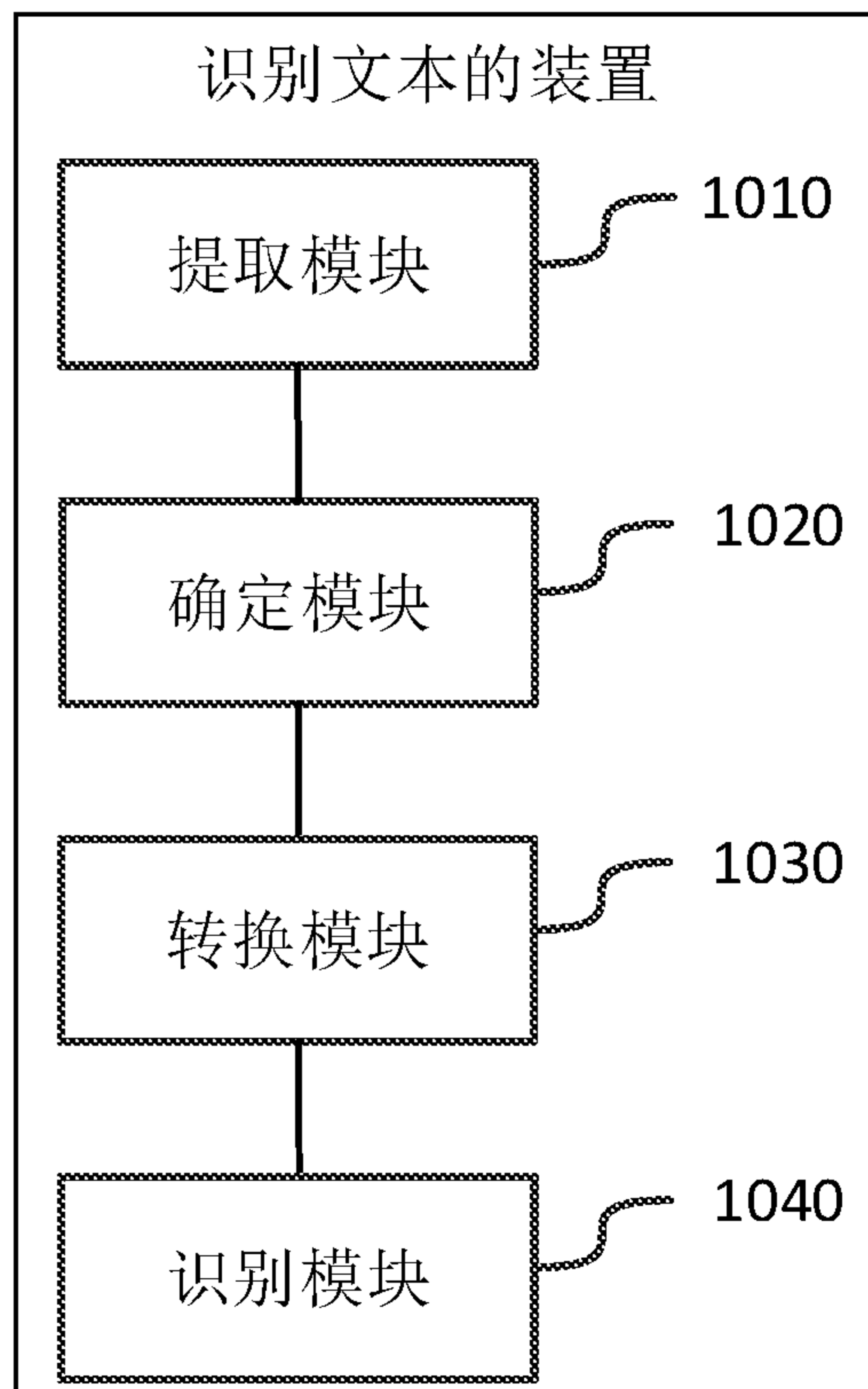


图 10

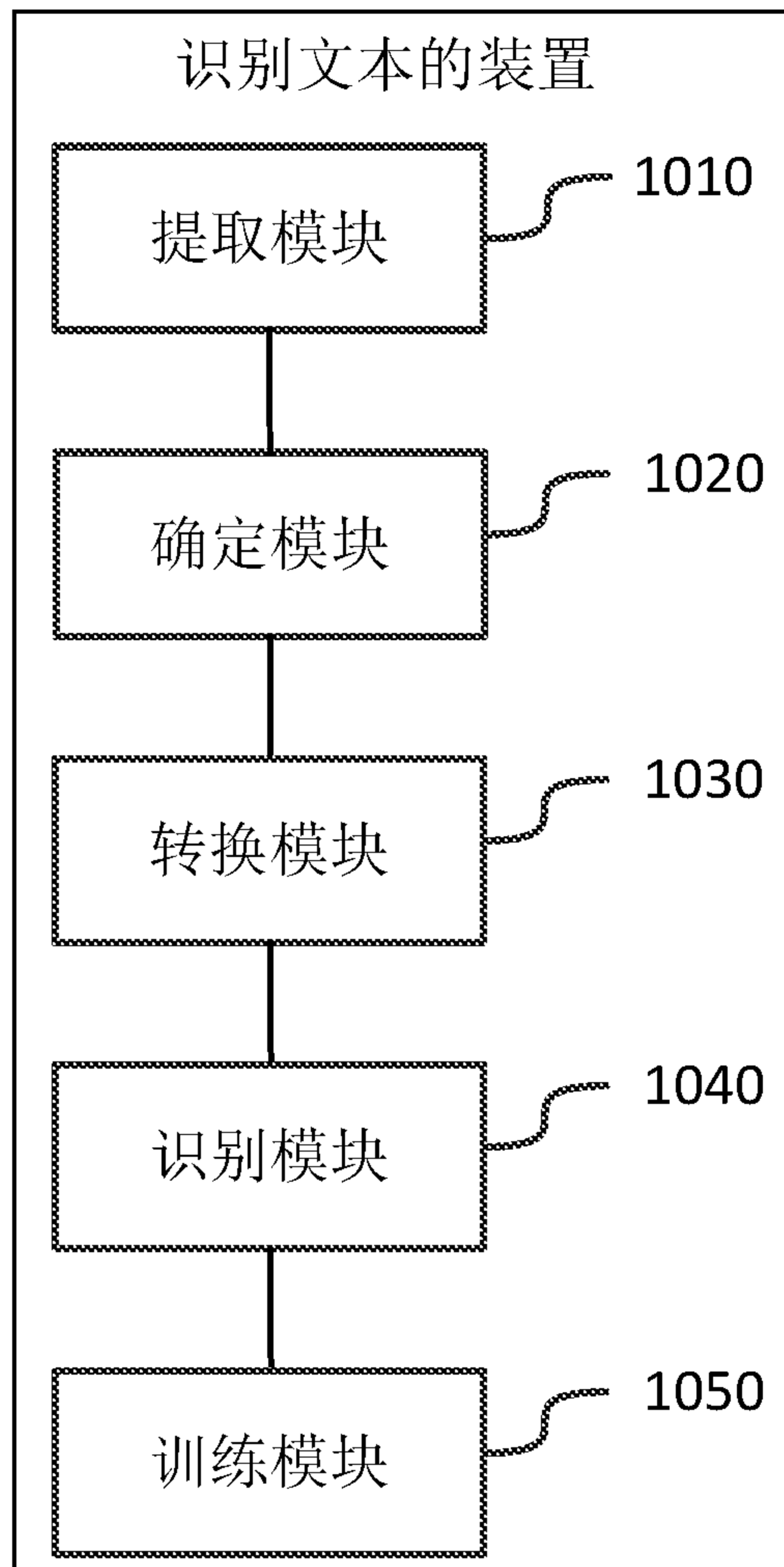


图 11

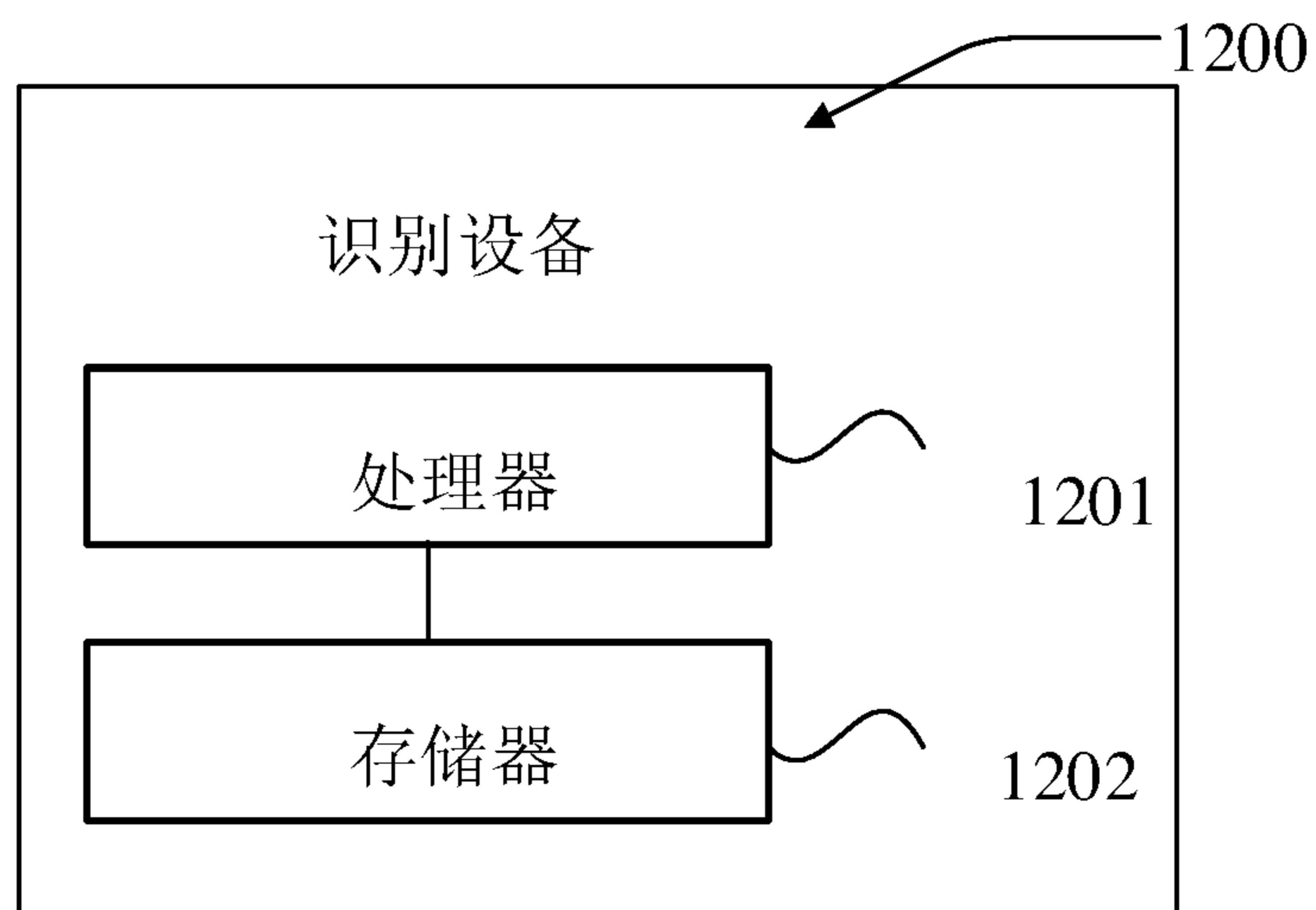


图 12

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/130654

A. CLASSIFICATION OF SUBJECT MATTER G06K 9/32(2006.01)i According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) G06K Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) CNABS; CNTXT; VEN; USTXT; EPTXT; WOTXT; CNKI: 弯曲, 文本, 文字, 字符, 识别, 检测, 边界, 边缘, 点, 头, 尾, curve, text, character, recognition, detection, boundary, edge, point, head, tail		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 111612009 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 01 September 2020 (2020-09-01) description, paragraphs 51-176, figures 1-3, 4a-4b, 5a-5c, 9-16	1, 9-10, 14, 22-23, 27-28
PX	CN 110837835 A (HUAZHONG UNIVERSITY OF SCIENCE AND TECHNOLOGY) 25 February 2020 (2020-02-25) description, paragraphs 63-111, and figures 1-4	1, 9-10, 14, 22-23, 27-28
Y	LIU, Yuliang et al. "Curved scene text detection via transverse and longitudinal sequence connection" 《Pattern Recognition》, 05 February 2019 (2019-02-05), ISSN: 0031-3203, sections 1-4	1-5, 7, 9-10, 13-18, 20, 22-23, 26-28
Y	CN 109829437 A (BEIJING KUANGSHI TECHNOLOGY CO., LTD.) 31 May 2019 (2019-05-31) description, paragraphs 39-176, and figures 1-10	1-5, 7, 9-10, 13-18, 20, 22-23, 26-28
A	CN 104809436 A (TIANJIN UNIVERSITY) 29 July 2015 (2015-07-29) entire document	1-28
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
<p>* Special categories of cited documents:</p> <p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p> <p>“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>“&” document member of the same patent family</p>		
Date of the actual completion of the international search 07 February 2021		Date of mailing of the international search report 22 February 2021
Name and mailing address of the ISA/CN China National Intellectual Property Administration (ISA/ CN) No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088 China		Authorized officer
Facsimile No. (86-10)62019451		Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/130654

C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 7043080 B1 (SHARP LABORATORIES OF AMERICA INC.) 09 May 2006 (2006-05-09) entire document	1-28
A	WANG, Xiaobing et al. "Arbitrary Shape Scene Text Detection with Adaptive Text Region Representation" 《2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)》 , 20 June 2019 (2019-06-20), sections 1-4	1-28

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No. PCT/CN2020/130654

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	111612009	A	01 September 2020	None	
CN	110837835	A	25 February 2020	None	
CN	109829437	A	31 May 2019	None	
CN	104809436	A	29 July 2015	CN 104809436	B 15 December 2017
US	7043080	B1	09 May 2006	None	

国际检索报告

国际申请号

PCT/CN2020/130654

<p>A. 主题的分类 G06K 9/32 (2006.01) i</p> <p>按照国际专利分类 (IPC) 或者同时按照国家分类和IPC两种分类</p>																							
<p>B. 检索领域</p> <p>检索的最低限度文献 (标明分类系统和分类号) G06K</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用)) CNABS; CNTXT; VEN; USTXT; EPTXT; WOTXT; CNKI: 弯曲, 文本, 文字, 字符, 识别, 检测, 边界, 边缘, 点, 头, 尾, curve, text, character, recognition, detection, boundary, edge, point, head, tail</p>																							
<p>C. 相关文件</p> <table border="1"> <thead> <tr> <th>类型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>PX</td> <td>CN 111612009 A (腾讯科技深圳有限公司) 2020年 9月 1日 (2020 - 09 - 01) 说明书第51-176段, 图1-3、4a-4b、5a-5c、9-16</td> <td>1、9-10、14、 22-23、27-28</td> </tr> <tr> <td>PX</td> <td>CN 110837835 A (华中科技大学) 2020年 2月 25日 (2020 - 02 - 25) 说明书第63-111段, 图1-4</td> <td>1、9-10、14、 22-23、27-28</td> </tr> <tr> <td>Y</td> <td>LIU, Yuliang 等. "Curved scene text detection via transverse and longitudinal sequence connection" 《Pattern Recognition》, 2019年 2月 5日 (2019 - 02 - 05), ISSN: 0031-3203, 第1-4节</td> <td>1-5、7、9-10、 13-18、20、 22-23、26-28</td> </tr> <tr> <td>Y</td> <td>CN 109829437 A (北京旷视科技有限公司) 2019年 5月 31日 (2019 - 05 - 31) 说明书第39-176段, 图1-10</td> <td>1-5、7、9-10、 13-18、20、 22-23、26-28</td> </tr> <tr> <td>A</td> <td>CN 104809436 A (天津大学) 2015年 7月 29日 (2015 - 07 - 29) 全文</td> <td>1-28</td> </tr> <tr> <td>A</td> <td>US 7043080 B1 (SHARP LABORATORIES OF AMERICA INC) 2006年 5月 9日 (2006 - 05 - 09) 全文</td> <td>1-28</td> </tr> </tbody> </table>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	PX	CN 111612009 A (腾讯科技深圳有限公司) 2020年 9月 1日 (2020 - 09 - 01) 说明书第51-176段, 图1-3、4a-4b、5a-5c、9-16	1、9-10、14、 22-23、27-28	PX	CN 110837835 A (华中科技大学) 2020年 2月 25日 (2020 - 02 - 25) 说明书第63-111段, 图1-4	1、9-10、14、 22-23、27-28	Y	LIU, Yuliang 等. "Curved scene text detection via transverse and longitudinal sequence connection" 《Pattern Recognition》, 2019年 2月 5日 (2019 - 02 - 05), ISSN: 0031-3203, 第1-4节	1-5、7、9-10、 13-18、20、 22-23、26-28	Y	CN 109829437 A (北京旷视科技有限公司) 2019年 5月 31日 (2019 - 05 - 31) 说明书第39-176段, 图1-10	1-5、7、9-10、 13-18、20、 22-23、26-28	A	CN 104809436 A (天津大学) 2015年 7月 29日 (2015 - 07 - 29) 全文	1-28	A	US 7043080 B1 (SHARP LABORATORIES OF AMERICA INC) 2006年 5月 9日 (2006 - 05 - 09) 全文	1-28
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求																					
PX	CN 111612009 A (腾讯科技深圳有限公司) 2020年 9月 1日 (2020 - 09 - 01) 说明书第51-176段, 图1-3、4a-4b、5a-5c、9-16	1、9-10、14、 22-23、27-28																					
PX	CN 110837835 A (华中科技大学) 2020年 2月 25日 (2020 - 02 - 25) 说明书第63-111段, 图1-4	1、9-10、14、 22-23、27-28																					
Y	LIU, Yuliang 等. "Curved scene text detection via transverse and longitudinal sequence connection" 《Pattern Recognition》, 2019年 2月 5日 (2019 - 02 - 05), ISSN: 0031-3203, 第1-4节	1-5、7、9-10、 13-18、20、 22-23、26-28																					
Y	CN 109829437 A (北京旷视科技有限公司) 2019年 5月 31日 (2019 - 05 - 31) 说明书第39-176段, 图1-10	1-5、7、9-10、 13-18、20、 22-23、26-28																					
A	CN 104809436 A (天津大学) 2015年 7月 29日 (2015 - 07 - 29) 全文	1-28																					
A	US 7043080 B1 (SHARP LABORATORIES OF AMERICA INC) 2006年 5月 9日 (2006 - 05 - 09) 全文	1-28																					
<p><input checked="" type="checkbox"/> 其余文件在C栏的续页中列出。 <input checked="" type="checkbox"/> 见同族专利附件。</p>																							
<p>* 引用文件的具体类型: "A" 认为不特别相关的表示了现有技术一般状态的文件 "E" 在国际申请日的当天或之后公布的在先申请或专利 "L" 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) "O" 涉及口头公开、使用、展览或其他方式公开的文件 "P" 公布日先于国际申请日但迟于所要求的优先权日的文件 "T" 在申请日或优先权日之后公布, 与申请不相抵触, 但为了解发明之理论或原理的在后文件 "X" 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 "Y" 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 "&" 同族专利的文件</p>																							
<p>国际检索实际完成的日期 2021年 2月 7日</p>		<p>国际检索报告邮寄日期 2021年 2月 22日</p>																					
<p>ISA/CN的名称和邮寄地址 中国国家知识产权局 (ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451</p>		<p>授权官员 马婷婷 电话号码 86-(20)-28958507</p>																					

C. 相关文件		
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求
A	WANG, Xiaobing 等. "Arbitrary Shape Scene Text Detection with Adaptive Text Region Representation" 《2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)》, 2019年 6月 20日 (2019 - 06 - 20), 第1-4节	1-28
.....		

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2020/130654

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	111612009	A	2020年 9月 1日	无	
CN	110837835	A	2020年 2月 25日	无	
CN	109829437	A	2019年 5月 31日	无	
CN	104809436	A	2015年 7月 29日	CN 104809436	B 2017年 12月 15日
US	7043080	B1	2006年 5月 9日	无	