



(12) 发明专利

(10) 授权公告号 CN 114332866 B

(45) 授权公告日 2024.10.18

(21) 申请号 202111490372.3

(22) 申请日 2021.12.08

(65) 同一申请的已公布的文献号

申请公布号 CN 114332866 A

(43) 申请公布日 2022.04.12

(73) 专利权人 上海大学

地址 200444 上海市宝山区上大路99号

(72) 发明人 韩越兴 陈侨川 王迎港 张瑞

潘登 夏锦桦 张家旺

(74) 专利代理机构 上海上大专利事务所(普通

合伙) 31205

专利代理师 何文欣

(51) Int. Cl.

G06V 30/148 (2022.01)

G06V 30/413 (2022.01)

(56) 对比文件

CN 105068807 A, 2015.11.18

CN 107909068 A, 2018.04.13

审查员 邓清清

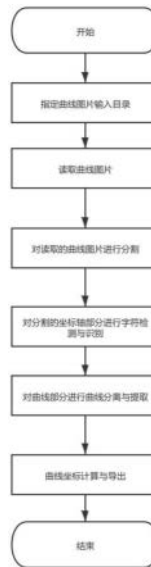
权利要求书2页 说明书6页 附图1页

(54) 发明名称

一种基于图像处理的文献曲线分离与坐标信息提取方法

(57) 摘要

本发明公开了一种基于图像处理的文献曲线分离与坐标信息提取方法,运用计算机图像检测方法,实现对文献曲线图片进行分割与曲线信息提取;首先读取一张曲线图片,利用图像处理方法分割曲线图片;然后利用字符检测算法找出字符区域,并获取字符区域具体数值与位置信息;随后根据颜色提取曲线图上不同颜色的曲线,在将曲线图分离成单一的曲线图之后就便于计算曲线的坐标了;本发明文献中曲线分离与坐标信息提取方法能够实现坐标分割、内容识别和坐标提取,提高文献信息提取速度,为构建相应学科数据库提供一种方法,促进对应学科的研究和发展。



1. 一种基于图像处理的文献曲线分离与坐标信息提取方法,其特征在于,包括如下步骤:

(1) 读入一篇文献,提取文献中曲线图像,利用图片格式进行保存,将图片的访问路径存入路径列表;

(2) 读取一张曲线图片,对曲线图像进行分割,分割成三个部分:X轴、Y轴和剩下的曲线内容部分,进行坐标信息获取,具体为:

首先对读取的曲线图像进行预处理,包括去噪,二值化,提高对比度;

然后根据坐标轴固定X轴和Y轴垂直的规律,获取到每张坐标轴垂直的坐标轴的位置信息,以此完成图片的分割过程;

(3) 字符检测与识别:对于截取的X轴和Y轴图片,对其进行字符检测与字符识别,以获取到所需要的数据信息,为像素坐标到真实坐标之间的转换提供数据来源;

(4) 曲线分离与提取:在完成了X轴Y轴的处理之后,对曲线图进行处理,对于有多种颜色区分的标准的曲线图,则根据曲线的颜色的不同将每条曲线单独的提取出来,获取曲线的像素坐标信息;

(5) 曲线坐标计算:在将曲线图像中不同颜色的曲线分离开,然后进行坐标计算并导出,将曲线上提取的数据保存到数据库中,完成坐标信息提取。

2. 根据权利要求1所述基于图像处理的文献曲线分离与坐标信息提取方法,其特征在于:在所述步骤(2)中,对曲线图像进行分割的具体步骤如下:

(2-1) 将原曲线图转换为灰度图,黑色的像素值为0,白色的像素值为255,其它灰色的像素值在0和255之间;

(2-2) 获取坐标系X轴信息:定义一个矩形,宽度为图片宽度,高度为一个像素,矩形内的像素值设定为1,用此矩形从上往下对图片进行滑动操作,与对应位置像素值进行相乘操作,然后将所有图片宽度个数的相乘结果相加,得到一组计算结果,比较并获取所有数值中的最小值,其数组序号就是所需的X轴所在的像素坐标;

(2-3) 获取坐标系Y轴信息:定义一个矩形,高度为图片高度,宽度为一个像素,矩形内的像素值设定为1,用此矩形从左往右对图片进行滑动操作,与对应位置像素值进行相乘操作,然后将所有图片高度个数的相乘结果相加,得到一组计算结果,比较并获取所有数值中的最小值,其数组序号就是所需的Y轴所在的像素坐标;

(2-4) 将图片分割成三个部分,分别包含X轴、包含Y轴和曲线内容部分,并对这三个部分进行保存。

3. 根据权利要求1所述基于图像处理的文献曲线分离与坐标信息提取方法,其特征在于:在所述步骤(3)中,进行字符检测与识别时,将X轴和Y轴图片上的有字符的区域从图片中检测出来,并获取字符的具体数值和其位置信息,具体步骤如下:

(3-1) 读取一张坐标轴的图片;

(3-2) 使用CRAFT文本检测模型,对输入的图片进行检测,识别出存在的字符文本并获取该文本的位置信息;

(3-3) 使用ResNet+LSTM+CTC组合网络结构,对检测完成的文本区域进行字符识别;

(3-4) 对在所述(3-4)步骤中获得的信息进行解码处理,获得输出结果,得到所需的文本信息,即,坐标轴上的坐标值信息;

将X轴和Y轴图片中的数字字符区域检测并识别出来,获取了字符的位置信息,为像素坐标到真实坐标之间的转换提供数据来源。

4.根据权利要求1所述基于图像处理的文献曲线分离与坐标信息提取方法,其特征在于:在所述步骤(4)中,进行曲线分离与提取,将曲线图片的曲线区域根据颜色的不同,将其分离了出来,生成多张单一曲线的图片,具体步骤如下:

(4-1)对分割出来的曲线坐标图中的像素进行原图颜色的复原;

(4-2)对在步骤(4-1)中获得的结果图像采取腐蚀操作,扩大曲线的颜色特征;

(4-3)对在步骤(4-2)中获取的腐蚀过后的图像进行划分,统计不同颜色的像素值出现的次数,当一种颜色出现次数到设定阈值,就认为整个图的曲线包含该类别颜色;

(4-4)对所获取的颜色RGB值进行筛选,删除其中相似的颜色,只保留一种主色调,作为依据提取图像中这种颜色的像素块。

5.根据权利要求1所述基于图像处理的文献曲线分离与坐标信息提取方法,其特征在于:在所述步骤(5)中,进行曲线坐标计算,在将图像中不同颜色的曲线分离开,进行坐标计算并导出,具体步骤如下:

(5-1)对提取出来的分离后曲线图像进行二值化操作,即可获取到像素值为255所对应的坐标,此坐标为像素坐标;

(5-2)将坐标原点转换到左下角;

(5-3)根据公式(1)和公式(2),将像素坐标转化为真实值的坐标,并导出;

$$X = X_{min} + pX \frac{X_j - X_i}{pX_j - pX_i} \quad (1)$$

$$Y = Y_{min} + pY \frac{Y_j - Y_i}{pY_j - pY_i} \quad (2)$$

其中X,Y代表最终转换过来的所需要的真实曲线点数据坐标值;pX,pY为在曲线上每个点的像素层次的坐标; X_{min}, Y_{min} 分别为X,Y轴上从原点开始的第一个检测到的数字坐标数据的数据值;在对坐标轴上的字符数值进行获取的时候,需要进行字符检测与识别,与此同时字符检测会对每个数字字符画出矩形检测框, pX_i, pX_j 代表X轴前后两个坐标轴数值检测过程中矩形框的左上角顶点像素的横坐标; X_i, X_j 为坐标检测进行字符识别后的具体数值; pY_i, pY_j 代表Y轴上下两个坐标轴数值检测过程中矩形框的左上角顶点像素的纵坐标; Y_i, Y_j 为坐标检测进行字符识别后的具体数值

(5-4)最终提取曲线图像中不同曲线的坐标信息,并生成Excel文件保存下来。

一种基于图像处理的文献曲线分离与坐标信息提取方法

技术领域

[0001] 本发明涉及一种从科学文献的曲线图像中提取信息的方法,包括曲线图像的分割、坐标的提取与信息提取等内容,应用于不同学科文献中曲线图数据提取和对应数据库构建等领域,在一定程度上提高该学科文献数据提取的速度,帮助提高该研究方向的研究进度和研究效率,促进相应学科快速发展。

背景技术

[0002] 在科学研究文献中,对于实验获取的数据信息多半常用曲线图或者是表格来进行描述,表格中的数据信息通常是显示的表示,但曲线图中各个点的坐标数据信息常常隐藏在图像中,因此获取曲线图像中的曲线数据信息显得尤其重要。早期的方法中,在Matlab工具上可以使用人工选择,获取的曲线上的点并导出,但是这种方法十分的低效,不能自动化获取大量曲线图像上的数据,无法满足日益增长的研究需要以及文献数据信息的爆炸性增长,因此需要一种可靠的检测手段来实现自动提取曲线坐标信息。随着计算机技术和图像处理技术的发展,运用计算机图像识别技术和图像处理方法能够实现曲线信息提取的目的,但如何将计算机信息分析与图像分析结合起来,从科学文献的曲线图像中提取信息还需要设计更加高效的方案,这成为亟待了解决的技术问题。

发明内容

[0003] 为了解决现有技术问题,本发明的目的在于克服已有技术存在的不足,提供一种基于图像处理的文献曲线分离与坐标信息提取方法,包括文献图像中曲线分离与坐标信息提取。为了提取大批量文献中曲线图像的坐标信息,本发明提出新的算法能够自动获取图片中不同颜色曲线的坐标信息,对于一张曲线图来说可以定位坐标轴,获取X轴坐标值,Y轴坐标值,分离曲线等问题。因此,本发明提出的这种新的方法可以实现对科学文献曲线图的信息提取,提高文献信息获取的速度,降低研究成本,促进相关学科的研究和发展。

[0004] 为达到上述发明创造目的,本发明采用如下技术方案:

[0005] 一种基于图像处理的文献曲线分离与坐标信息提取方法,包括如下步骤:

[0006] (1) 读入一篇文献,提取文献中曲线图像,利用图片格式进行保存,将图片的访问路径存入路径列表;

[0007] (2) 读取一张曲线图片,对曲线图像进行分割,分割成三个部分:X轴、Y轴和剩下的曲线内容部分,进行坐标信息获取,具体为:

[0008] 首先对读取的曲线图像进行预处理,包括去噪,二值化,提高对比度;

[0009] 然后根据坐标轴固定X轴和Y轴垂直的规律,获取到每张坐标轴垂直的坐标轴的位置信息,以此完成图片的分割过程;

[0010] (3) 字符检测与识别:对于截取的X轴和Y轴图片,对其进行字符检测与字符识别,以获取到所需要的数据信息,为像素坐标到真实坐标之间的转换提供数据来源;

[0011] (4) 曲线分离与提取:在完成了X轴Y轴的处理之后,对曲线图进行处理,对于有多

种颜色区分的标准的曲线图,则根据曲线的颜色的不同将每条曲线单独的提取出来,获取曲线的像素坐标信息;

[0012] (5) 曲线坐标计算:在将曲线图像中不同颜色的曲线分离开,然后进行坐标计算并导出,将曲线上提取的数据保存到数据库中,完成坐标信息提取。

[0013] 优选地,在所述步骤(2)中,对曲线图像进行分割的具体步骤如下:

[0014] (2-1) 将原曲线图转换为灰度图,黑色的像素值为0,白色的像素值为255,其它灰色的像素值在0和255之间;

[0015] (2-2) 获取坐标系X轴信息:定义一个矩形,宽度为图片宽度,高度为一个像素,矩形内的像素值设定为1,用此矩形从上往下对图片进行滑动操作,与对应位置像素值进行相乘操作,然后将所有图片宽度个数的相乘结果相加,得到一组计算结果,比较并获取所有数值中的最小值,其数组序号就是所需的X轴所在的像素坐标;

[0016] (2-3) 获取坐标系Y轴信息:定义一个矩形,高度为图片高度,宽度为一个像素,矩形内的像素值设定为1,用此矩形从左往右对图片进行滑动操作,与对应位置像素值进行相乘操作,然后将所有图片高度个数的相乘结果相加,得到一组计算结果,比较并获取所有数值中的最小值,其数组序号就是所需的Y轴所在的像素坐标;

[0017] (2-4) 将图片分割成三个部分,分别包含X轴、包含Y轴和曲线内容部分,并对这三个部分进行保存。

[0018] 优选地,在所述步骤(3)中,进行字符检测与识别时,将X轴和Y轴图片上的有字符的区域从图片中检测出来,并获取字符的具体数值和其位置信息,具体步骤如下:

[0019] (3-1) 读取一张坐标轴的图片;

[0020] (3-2) 使用CRAFT文本检测模型,对输入的图片进行检测,识别出存在的字符文本并获取该文本的位置信息;

[0021] (3-3) 使用ResNet+LSTM+CTC组合网络结构,对检测完成的文本区域进行字符识别;

[0022] (3-4) 对在所述(3-4)步骤中获得的信息进行解码处理,获得输出结果,得到所需的文本信息,即,坐标轴上的坐标值信息;

[0023] 将X轴和Y轴图片中的数字字符区域检测并识别出来,获取了字符的位置信息,为像素坐标到真实坐标之间的转换提供数据来源。

[0024] 优选地,在所述步骤(4)中,进行曲线分离与提取,将曲线图片的曲线区域根据颜色的不同,将其分离了出来,生成多张单一曲线的图片,具体步骤如下:

[0025] (4-1) 对分割出来的曲线坐标图中的像素进行原图颜色的复原;

[0026] (4-2) 对在步骤(4-1)中获得的结果图像采取腐蚀操作,扩大曲线的颜色特征;

[0027] (4-3) 对在步骤(4-2)中获取的腐蚀过后的图像进行划分,统计不同颜色的像素值出现的次数,当一种颜色出现次数到设定阈值,就认为整个图的曲线包含该类别颜色;

[0028] (4-4) 对所获取的颜色RGB值进行筛选,删除其中相似的颜色,只保留一种主色调,作为依据提取图像中这种颜色的像素块。

[0029] 优选地,在所述步骤(5)中,进行曲线坐标计算,在将图像中不同颜色的曲线分离开,进行坐标计算并导出,具体步骤如下:

[0030] (5-1) 对提取出来的分离后曲线图像进行二值化操作,即可获取到像素值为255所

对应的坐标,此坐标为像素坐标;

[0031] (5-2) 将坐标原点转换到左下角;

[0032] (5-3) 根据公式 (1) 和公式 (2), 将像素坐标转化为真实值的坐标, 并导出。

$$[0033] \quad X = X_{min} + pX \frac{X_j - X_i}{pX_j - pX_i} \quad (1)$$

$$[0034] \quad Y = Y_{min} + pY \frac{Y_j - Y_i}{pY_j - pY_i} \quad (2)$$

[0035] 其中X,Y代表最终转换过来的所需要的真实曲线点数据坐标值;pX,pY为在曲线上每个点的像素层次的坐标; X_{min}, Y_{min} 分别为X,Y轴上从原点开始的第一个检测到的数字坐标数据的数据值;在对坐标轴上的字符数值进行获取的时候,需要进行字符检测与识别,与此同时字符检测会对每个数字字符画出矩形检测框, pX_i, pX_j 代表前后两个坐标轴数值检测过程中矩形框的左上角顶点像素的横坐标; X_i, X_j 为坐标检测的具体数值; pY_i, pY_j 代表Y轴上下两个坐标轴数值检测过程中矩形框的左上角顶点像素的纵坐标; Y_i, Y_j 为坐标检测进行字符识别后的具体数值;

[0036] (5-4) 最终提取曲线图像中不同曲线的坐标信息, 并生成Excel文件保存下来。

[0037] 本发明与现有技术相比较, 具有如下显而易见的突出实质性特点和显著优点:

[0038] 1. 本发明能够分割出X轴和Y轴, 从而实现对特定的轴上数据进行提取;

[0039] 2. 本发明能够提取文献曲线图上坐标轴的数值;

[0040] 3. 本发明能够提取曲线图中曲线信息;

[0041] 4. 本发明文献中曲线分离与坐标信息提取方法能够实现坐标分割、内容识别和坐标提取, 提高文献信息提取速度, 为构建相应学科数据库提供一种方法, 促进对应学科的研究和发展。

附图说明

[0042] 图1为本发明优选实施例方法的流程图。

具体实施方式

[0043] 为了使本技术领域的人员更好地理解本发明方案, 下面将本发明优选实施例结合附图, 对本发明实施例中的技术方案进行清查、完整地描述。显然, 所描述的实施例仅仅是本发明的一部分实施案例, 而非全部实施案例。基于本发明中的实施例, 本领域普通技术人员在没有做过创造性劳动前提下所获得的所有其他实施例, 都应当属于本发明的保护范围。

[0044] 以下结合具体的实施例子对上述方案做进一步说明, 本发明的优选实施例详述如下:

[0045] 实施例一:

[0046] 在本实施例中, 参见图1, 一种基于图像处理的文献曲线分离与坐标信息提取方法, 包括如下步骤:

[0047] (1) 读入一篇文献, 提取文献中曲线图像, 利用图片格式进行保存, 将图片的访问路径存入路径列表;

[0048] (2) 读取一张曲线图片,对曲线图像进行分割,分割成三个部分:X轴、Y轴和剩下的曲线内容部分,进行坐标信息获取,具体为:

[0049] 首先对读取的曲线图像进行预处理,包括去噪,二值化,提高对比度;

[0050] 然后根据坐标轴固定X轴和Y轴垂直的规律,获取到每张坐标轴垂直的坐标轴的位置信息,以此完成图片的分割过程;

[0051] (3) 字符检测与识别:对于截取的X轴和Y轴图片,对其进行字符检测与字符识别,以获取到所需要的数据信息,为像素坐标到真实坐标之间的转换提供数据来源;

[0052] (4) 曲线分离与提取:在完成了X轴Y轴的处理之后,对曲线图进行处理,对于有多种颜色区分的标准的曲线图,则根据曲线的颜色的不同将每条曲线单独的提取出来,获取曲线的像素坐标信息;

[0053] (5) 曲线坐标计算:在将曲线图像中不同颜色的曲线分离开,然后进行坐标计算并导出,将曲线上提取的数据保存到数据库中,完成坐标信息提取。

[0054] 本实施例方法能够自动获取图片中不同颜色曲线的坐标信息,对于一张曲线图来说可以定位坐标轴,获取X轴坐标值,Y轴坐标值,分离曲线等问题。本实施例方法实现对科学文献曲线图的信息提取,提高文献信息获取的速度,降低研究成本。本实施例基于图像处理的文献曲线分离与坐标信息提取方法,运用计算机图像检测方法,实现对文献曲线图片进行分割与曲线信息提取;首先读取一张曲线图片,利用图像处理方法分割曲线图片;然后利用字符检测算法找出字符区域,并获取字符区域具体数值与位置信息;随后根据颜色提取曲线图上不同颜色的曲线,在将曲线图分离成单一的曲线图之后就便于计算曲线的坐标了;本实施例文献中曲线分离与坐标信息提取方法能够实现坐标分割、内容识别和坐标提取,提高文献信息提取速度,为构建相应学科数据库提供一种方法,促进对应学科的研究和发展。

[0055] 实施例二:

[0056] 本实施例与实施例一基本相同,特别之处在于:

[0057] 在本实施例中,在所述步骤(2)中,对曲线图像进行分割的具体步骤如下:

[0058] (2-1) 将原曲线图转换为灰度图,黑色的像素值为0,白色的像素值为255,其它灰色的像素值在0和255之间;

[0059] (2-2) 获取坐标系X轴信息:定义一个矩形,宽度为图片宽度,高度为一个像素,矩形内的像素值设定为1,用此矩形从上往下对图片进行滑动操作,与对应位置像素值进行相乘操作,然后将所有图片宽度个数的相乘结果相加,得到一组计算结果,比较并获取所有数值中的最小值,其数组序号就是所需的X轴所在的像素坐标;

[0060] (2-3) 获取坐标系Y轴信息:定义一个矩形,高度为图片高度,宽度为一个像素,矩形内的像素值设定为1,用此矩形从左往右对图片进行滑动操作,与对应位置像素值进行相乘操作,然后将所有图片高度个数的相乘结果相加,得到一组计算结果,比较并获取所有数值中的最小值,其数组序号就是所需的Y轴所在的像素坐标;

[0061] (2-4) 将图片分割成三个部分,分别包含X轴、包含Y轴和曲线内容部分,并对这三个部分进行保存。

[0062] 本实例将输入的曲线图片分割成三个部分,为后续的坐标值的识别,曲线的分离提供了基础。

[0063] 实施例三：

[0064] 本实施例与上述实施例基本相同，特别之处在于：

[0065] 在本实施例中，在所述步骤(3)中，进行字符检测与识别时，将X轴和Y轴图片上的有字符的区域从图片中检测出来，并获取字符的具体数值和其位置信息，具体步骤如下：

[0066] (3-1) 读取一张坐标轴的图片；

[0067] (3-2) 使用CRAFT文本检测模型，对输入的图片进行检测，识别出存在的字符文本并获取该文本的位置信息；

[0068] (3-3) 使用ResNet+LSTM+CTC组合网络结构，对检测完成的文本区域进行字符识别；

[0069] (3-4) 对在所述(3-4)步骤中获得的信息进行解码处理，获得输出结果，得到所需的文本信息，即，坐标轴上的坐标值信息；

[0070] 将X轴和Y轴图片中的数字字符区域检测并识别出来，获取了字符的位置信息，为像素坐标到真实坐标之间的转换提供数据来源。

[0071] 本实施例将X轴和Y轴图片中的数字字符区域检测并识别出来了，同时也获取了该字符的位置信息，为之后的坐标计算提供了基础。

[0072] 实施例四：

[0073] 本实施例与上述实施例基本相同，特别之处在于：

[0074] 在本实施例中，在所述步骤(4)中，进行曲线分离与提取，将曲线图片的曲线区域根据颜色的不同，将其分离了出来，生成多张单一曲线的图片，具体步骤如下：

[0075] (4-1) 对分割出来的曲线坐标图中的像素进行原图颜色的复原；

[0076] (4-2) 对在步骤(4-1)中获得的结果图像采取腐蚀操作，扩大曲线的颜色特征；

[0077] (4-3) 对在步骤(4-2)中获取的腐蚀过后的图像进行划分，统计不同颜色的像素值出现的次数，当一种颜色出现次数到设定阈值，就认为整个图的曲线包含该类别颜色；

[0078] (4-4) 对所获取的颜色RGB值进行筛选，删除其中相似的颜色，只保留一种主色调，作为依据提取图像中这种颜色的像素块。

[0079] 本实施例进行曲线分离与提取：在完成了X轴Y轴的处理之后，对曲线图进行处理，对于有多种颜色区分的标准的曲线图，则根据曲线的颜色的不同将每条曲线单独的提取出来，获取曲线的像素坐标信息。

[0080] 实施例五：

[0081] 本实施例与上述实施例基本相同，特别之处在于：

[0082] 在本实施例中，在所述步骤(5)中，进行曲线坐标计算，在将图像中不同颜色的曲线分离开，进行坐标计算并导出，具体步骤如下：

[0083] (5-1) 对提取出来的分离后曲线图像进行二值化操作，即可获取到像素值为255所对应的坐标，此坐标为像素坐标；

[0084] (5-2) 将坐标原点转换到左下角；

[0085] (5-3) 根据公式(1)和公式(2)，将像素坐标转化为真实值的坐标，并导出。

$$[0086] \quad X = X_{min} + pX \frac{X_j - X_i}{pX_j - pX_i} \quad (1)$$

$$[0087] \quad Y = Y_{min} + pY \frac{Y_j - Y_i}{pY_j - pY_i} \quad (2)$$

[0088] 其中X,Y代表最终转换过来的所需要的真实曲线点数据坐标值;pX,pY为在曲线上每个点的像素层次的坐标; X_{min}, Y_{min} 分别为X,Y轴上从原点开始的第一个检测到的数字坐标数据的数据值;在对坐标轴上的字符数值进行获取的时候,需要进行字符检测与识别,与此同时字符检测会对每个数字字符画出矩形检测框, pX_i, pX_j 代表前后两个坐标轴数值检测过程中矩形框的左上角顶点像素的横坐标; X_i, X_j 为坐标检测的具体数值; pY_i, pY_j 代表Y轴上下两个坐标轴数值检测过程中矩形框的左上角顶点像素的纵坐标; Y_i, Y_j 为坐标检测进行字符识别后的具体数值;

[0089] (5-4) 最终提取曲线图像中不同曲线的坐标信息,并生成Excel文件保存下来。

[0090] 本实施例进行曲线坐标计算,在将曲线图像中不同颜色的曲线分离开,然后进行坐标计算并导出,将曲线上提取的数据保存到数据库中,完成坐标信息提取。

[0091] 综上所述,图1是本发明上述实施例基于图像处理的文献曲线分离与坐标信息提取方法的流程图。本发明上述实施例方法首先,读入一篇文献,提取文献中曲线部分的内容,转化为图片格式保存,将图片的访问路径存入路径列表;读取一张曲线图片,需要对文献曲线图像进行分割,分割成三个部分,X轴,Y轴及曲线部分;然后读取一张图片,对读取的图像进行预处理,包括去噪,二值化,提高对比度等;之后根据坐标轴固定X轴和Y轴垂直的规律,获取到每张坐标轴垂直的坐标轴的位置信息,以此完成图片的分割过程;对于截取的X轴和Y轴图片,需要对其进行字符检测与字符识别以获取到所需要的数据信息,为之后的像素坐标到真实坐标之间的转换提供数据来源;对于有多种颜色区分的标准的曲线图,采取的主要方法是根据曲线的颜色的不同将每条曲线单独的提取出来,进而之后就可以采取普适的方法获取到曲线的像素坐标信息。

[0092] 上面对本发明实施例结合附图进行了说明,但本发明不限于上述实施例,还可以根据本发明的发明创造的目的做出多种变化,凡依据本发明技术方案的精神实质和原理下做的改变、修饰、替代、组合或简化,均应为等效的置换方式,只要符合本发明的发明目的,只要不背离本发明的技术原理和发明构思,都属于本发明的保护范围。

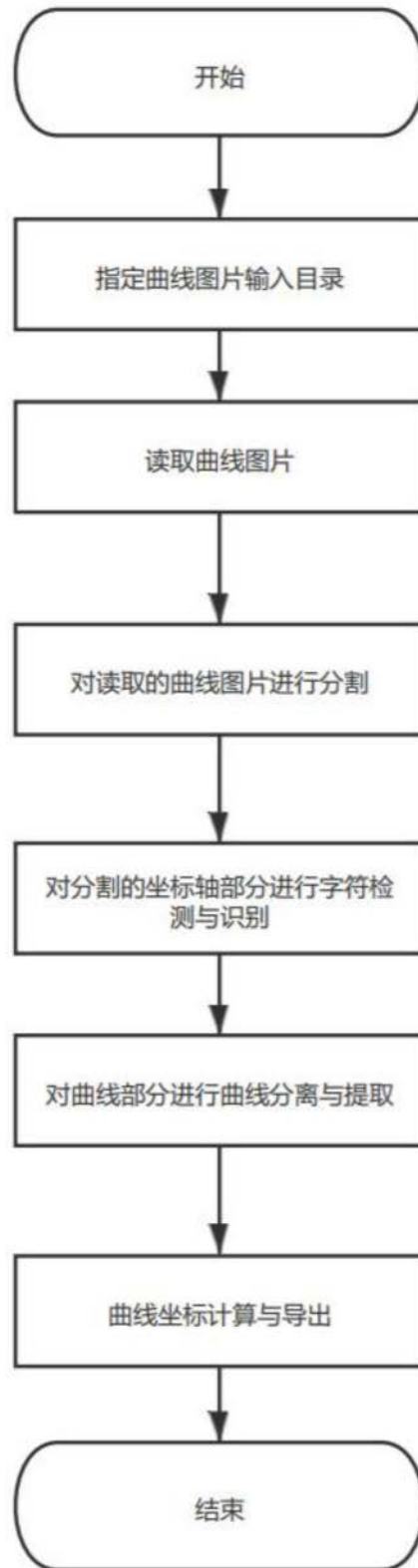


图1