



(12) 发明专利

(10) 授权公告号 CN 103093185 B

(45) 授权公告日 2016. 05. 25

(21) 申请号 201110347361. X

US 5212739 A, 1993. 05. 18,

(22) 申请日 2011. 11. 07

US 5054094 A, 1991. 10. 01,

(73) 专利权人 佳能株式会社  
地址 日本东京

钱稷. 基于图像处理的字符识别系统研究. 《中国优秀硕士学位论文全文数据库》. 2007,

审查员 单娟

(72) 发明人 姜涌 张琳琳

(74) 专利代理机构 中国国际贸易促进委员会专利商标事务所 11038

代理人 魏小微

(51) Int. Cl.

G06K 9/00(2006. 01)

G06K 9/20(2006. 01)

G06K 9/46(2006. 01)

(56) 对比文件

CN 101877050 A, 2010. 11. 03,

CN 101576956 A, 2009. 11. 11,

CN 101408933 A, 2009. 04. 15,

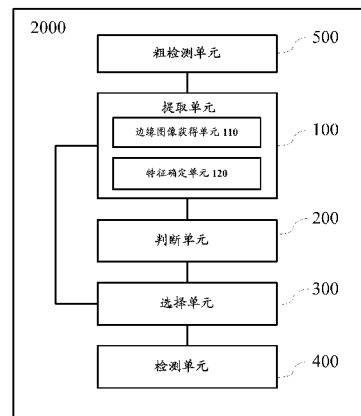
权利要求书4页 说明书18页 附图14页

(54) 发明名称

字符识别装置、图像处理装置及其方法

(57) 摘要

本发明提供字符识别装置、图像处理装置及其方法。该字符识别装置包括：提取单元，被配置为从图像中提取特征集，所述特征集包含一个或多个特征；判断单元，被配置为通过使用所提取的特征集，判断该图像是否有可能包含字符；选择单元，被配置为如果该图像被判断为有可能包含字符，则通过使用所提取的特征集，从一组具体字符检测器中选择要对该图像使用的具体字符检测器子集；以及检测单元，被配置为通过使用所选择的具体字符检测器子集中的具体字符检测器，对该图像进行检测，以识别该图像中的具体字符。



1. 一种字符识别装置,其特征在于包括:

提取单元,被配置为从图像中提取特征集,所述特征集包含一个或更多个特征;

判断单元,被配置为通过使用所提取的特征集,判断该图像是否有可能包含字符;

选择单元,被配置为如果该图像被判断为有可能包含字符,则通过使用所提取的特征集,从一组具体字符检测器中选择要对该图像使用的具体字符检测器子集;以及

检测单元,被配置为通过使用所选择的具体字符检测器子集中的具体字符检测器,对该图像进行检测,以识别该图像中的具体字符,

其中,所述特征集中的所述一个或更多个特征是沿预定方向被依次提取的一个或更多个边缘特征,所述一个或更多个边缘特征描述字符的基本结构,并且表示图像是否包含字符以及图像包含哪个字符的信息。

2. 如权利要求1所述的字符识别装置,其中

所述一个或更多个边缘特征中的每一个包含第一水平边缘、第二水平边缘、第一垂直边缘和第二垂直边缘中的至少一个,将所述预定方向的上游方向作为上方,所述第一水平边缘是潜在笔画的下边缘,所述第二水平边缘是潜在笔画的上边缘,所述第一垂直边缘是图像左部的潜在笔画的左边缘,所述第二垂直边缘是图像右部的潜在笔画的右边缘。

3. 如权利要求2所述的字符识别装置,其中

所述第一水平边缘是在从上向下的方向上从暗到亮的边缘,所述第二水平边缘是在从上向下的方向上从亮到暗的边缘,所述第一垂直边缘是在从左向右的方向上从亮到暗的边缘,所述第二垂直边缘是在从左向右的方向上从暗到亮的边缘。

4. 如权利要求3所述的字符识别装置,其中

所述提取单元包括:

边缘图像获得单元,被配置为获得所述图像的边缘图像;以及

特征确定单元,被配置为:

在该边缘图像中,沿从上向下的方向依次查找所述第一水平边缘和第二水平边缘,然后沿水平方向查找所述第一垂直边缘和第二垂直边缘,以便确定一个特征;以及

如果所述第二水平边缘不位于所述边缘图像的最底部,则继续向下查找下一个特征,

其中,对于一个特征:

如果该特征是该图像的第一个特征,则所述边缘图像的最顶部的预定点作为查找的起点并且作为所述第一个特征的第一水平边缘,

如果直到所述边缘图像的最底部的点仍未找到从上向下的方向上从亮到暗的边缘,则该特征确定单元将所述最底部的点作为该特征的第二水平边缘,

如果第一水平边缘位于边缘图像的最顶部,则该特征确定单元将该特征确定为不包含第一水平边缘,

如果第二水平边缘位于边缘图像的最底部,则该特征确定单元将该特征确定为不包含第二水平边缘。

5. 如权利要求4所述的字符识别装置,其中,对于一个特征:

所述特征确定单元被配置为:在该边缘图像中沿从上向下的方向依次查找所述第一水平边缘和第二水平边缘时,沿垂直中途线向下依次查找所述第一水平边缘和第二水平边缘,所述垂直中途线与等分该图像的垂直线之间的距离不大于图像宽度的1/8。

6. 如权利要求4或5所述的字符识别装置,其中,对于一个特征:

所述特征确定单元被配置为:在该边缘图像中沿水平方向查找所述第一垂直边缘和所述第二垂直边缘时,从所述第一水平边缘和第二水平边缘之间的中途点向左查找所述第一垂直边缘以及向右查找所述第二垂直边缘,所述中途点与所述第一水平边缘和第二水平边缘间的中点之间的距离不大于所述第一水平边缘和第二水平边缘之间的距离的1/8。

7. 如权利要求2所述的字符识别装置,其中

仅在满足以下条件1~4中的全部条件的情况下,所述判断单元将所述图像判断为有可能包含字符:

条件1:如果特征集仅具有一个特征,则该特征应属于D类,其中,所述D类是由不包含水平边缘的特征构成的;

条件2:如果特征集具有多于一个的特征,则第一个特征应属于A类,其中,所述A类是由包含第二水平边缘但不包含第一水平边缘的特征构成的;

条件3:如果特征集具有多于一个的特征,则最后一个特征应属于c类,其中,所述c类是由包含第一水平边缘但不包含第二水平边缘的特征构成的;

条件4:如果特征集具有多于一个的特征,则既不是第一个特征也不是最后一个特征的特征应属于B类,其中,所述B类是由包含第一水平边缘和第二水平边缘两者的特征构成的。

8. 如权利要求1所述的字符识别装置,其中

所述选择单元被配置为通过利用特征集与具体字符子集之间的映射关系,从所述一组具体字符检测器中选择要对该图像使用的具体字符检测器子集,所述具体字符检测器子集对应于所述具体字符子集。

9. 如权利要求1所述的字符识别装置,还包括:

粗检测单元,被配置为对该图像的边缘图像进行初步检测,以便检测该图像是否是潜在的包含字符的图像,

其中,仅在该粗检测单元确定该图像是潜在的包含字符的图像的情况下,所述提取单元、判断单元、选择单元和检测单元才对该图像进行相应操作。

10. 一种图像处理装置,其特征不在于包含:

滑动窗应用单元,被配置为对输入图像应用滑动窗以形成多个部分图像;

尺寸改变单元,被配置为如果所述滑动窗的尺寸小于预定尺寸,则增大所述滑动窗应用单元的滑动窗的尺寸,

其中,如果尺寸改变单元增大了滑动窗的尺寸,则所述输入图像被再次输入到所述滑动窗应用单元;以及

如权利要求1所述的字符识别装置,用于处理从所述滑动窗应用单元输出的所述多个部分图像。

11. 一种字符识别方法,其特征不在于包括:

提取步骤,从图像中提取特征集,所述特征集包含一个或更多个特征;

判断步骤,通过使用所提取的特征集,判断该图像是否有可能包含字符;

选择步骤,如果该图像被判断为有可能包含字符,则通过使用所提取的特征集,从一组具体字符检测器中选择要对该图像使用的具体字符检测器子集;以及

检测步骤,通过使用所选择的具体字符检测器子集中的具体字符检测器,对该图像进

行检测,以识别该图像中的具体字符,

其中,所述特征集中的所述一个或更多个特征是沿预定方向被依次提取的一个或更多个边缘特征,所述一个或更多个边缘特征描述字符的基本结构,并且表示图像是否包含字符以及图像包含哪个字符的信息。

12. 如权利要求11所述的字符识别方法,其中

所述一个或更多个边缘特征中的每一个包含第一水平边缘、第二水平边缘、第一垂直边缘和第二垂直边缘中的至少一个,将所述预定方向的上游方向作为上方,所述第一水平边缘是潜在笔画的下边缘,所述第二水平边缘是潜在笔画的上边缘,所述第一垂直边缘是图像左部的潜在笔画的左边缘,所述第二垂直边缘是图像右部的潜在笔画的右边缘。

13. 如权利要求12所述的字符识别方法,其中

所述第一水平边缘是在从上向下的方向上从暗到亮的边缘,所述第二水平边缘是在从上向下的方向上从亮到暗的边缘,所述第一垂直边缘是在从左向右的方向上从亮到暗的边缘,所述第二垂直边缘是在从左向右的方向上从暗到亮的边缘。

14. 如权利要求13所述的字符识别方法,其中

所述提取步骤包括:

边缘图像获得步骤,获得所述图像的边缘图像;以及

特征确定步骤:

在该边缘图像中,沿从上向下的方向依次查找所述第一水平边缘和第二水平边缘,然后沿水平方向查找所述第一垂直边缘和第二垂直边缘,以便确定一个特征;以及

如果所述第二水平边缘不位于所述边缘图像的最底部,则继续向下查找下一个特征,

其中,对于一个特征:

如果该特征是该图像的第一个特征,则所述边缘图像的最顶部的预定点作为查找的起点并且作为所述第一个特征的第一水平边缘,

如果直到所述边缘图像的最底部的点仍未找到从上向下的方向上从亮到暗的边缘,则将所述最底部的点作为该特征的第二水平边缘,

如果第一水平边缘位于边缘图像的最顶部,则将该特征确定为不包含第一水平边缘,

如果第二水平边缘位于边缘图像的最底部,则将该特征确定为不包含第二水平边缘。

15. 如权利要求14所述的字符识别方法,其中,对于一个特征:

在所述特征确定步骤中,在该边缘图像中沿从上向下的方向依次查找所述第一水平边缘和第二水平边缘时,沿垂直中途线向下依次查找所述第一水平边缘和第二水平边缘,所述垂直中途线与等分该图像的垂直线之间的距离不大于图像宽度的 $1/8$ 。

16. 如权利要求14或15所述的字符识别方法,其中,对于一个特征:

在所述特征确定步骤中,在该边缘图像中沿水平方向查找所述第一垂直边缘和所述第二垂直边缘时,从所述第一水平边缘和第二水平边缘之间的中途点向左查找所述第一垂直边缘以及向右查找所述第二垂直边缘,所述中途点与所述第一水平边缘和第二水平边缘间的中点之间的距离不大于所述第一水平边缘和第二水平边缘之间的距离的 $1/8$ 。

17. 如权利要求12所述的字符识别方法,其中

在所述判断步骤中,仅在满足以下条件1~4中的全部条件的情况下,将所述图像判断为有可能包含字符:

条件1:如果特征集仅具有一个特征,则该特征应属于D类,其中,所述D类是由不包含水平边缘的特征构成的;

条件2:如果特征集具有多于一个的特征,则第一个特征应属于A类,其中,所述A类是由包含第二水平边缘但不包含第一水平边缘的特征构成的;

条件3:如果特征集具有多于一个的特征,则最后一个特征应属于c类,其中,所述c类是由包含第一水平边缘但不包含第二水平边缘的特征构成的;

条件4:如果特征集具有多于一个的特征,则既不是第一个特征也不是最后一个特征的特征应属于B类,其中,所述B类是由包含第一水平边缘和第二水平边缘两者的特征构成的。

18.如权利要求11所述的字符识别方法,其中

在所述选择步骤中,通过利用特征集与具体字符子集之间的映射关系,从所述一组具体字符检测器中选择要对该图像使用的具体字符检测器子集,所述具体字符检测器子集对应于所述具体字符子集。

19.如权利要求11所述的字符识别方法,还包括:

粗检测步骤,对该图像的边缘图像进行初步检测,以便检测该图像是否是潜在的包含字符的图像,

其中,仅在该粗检测步骤中确定该图像是潜在的包含字符的图像的情况下,所述提取步骤、判断步骤、选择步骤和检测步骤才对该图像进行操作。

20.一种图像处理方法,其特征在于包含:

滑动窗应用步骤,对输入图像应用滑动窗以形成多个部分图像;

尺寸改变步骤,如果所述滑动窗的尺寸小于预定尺寸,则增大所述滑动窗的尺寸,

其中,如果在尺寸改变步骤中增大了滑动窗的尺寸,则所述输入图像被再次输入到所述滑动窗应用步骤;以及

执行如权利要求11所述的字符识别方法,以处理从所述滑动窗应用步骤输出的所述多个部分图像。

## 字符识别装置、图像处理装置及其方法

### 技术领域

[0001] 本发明一般涉及字符识别装置、图像处理装置及其方法。更特别地,本发明涉及用于检测和识别图像中的字符的字符识别装置和字符识别方法,以及用于检测和识别图像中的一个或更多个字符的图像处理装置和图像处理方法。

### 背景技术

[0002] 字符检测和识别在很多领域中被广泛使用。例如,字符检测和识别在诸如车牌照识别、邮政编码识别和票据识别等应用中是非常有用的。迄今为止已提出了很多用于字符检测和识别的技术。典型的技术包括基于连通分量分析的方法和基于神经网络的方法等等。

### 发明内容

[0003] 然而,本发明的发明人已发现,如上所述的已知的用于字符检测和识别的典型技术被诸如照明条件、模糊或复杂背景等的因素所限制。

[0004] 此外,根据已知的典型技术,如果一种具体字符检测器无法识别图像中的字符,则将由另一具体字符检测器处理该图像,直到有某个具体字符检测器成功识别该图像中的字符或者该图像已由所有具体字符检测器处理过为止。在最坏的情况下,将由所有具体字符检测器顺序地处理该图像,并且要对一幅图像使用的具体字符检测器的最大数量可以等于整个字符集中的字符的数量。因此,识别速度不得不是低的。

[0005] 为了解决以上技术问题中的至少一个,本发明提供一种字符识别装置,该字符识别装置包括:提取单元,被配置为从图像中提取特征集,所述特征集包含一个或更多个特征;判断单元,被配置为通过使用所提取的特征集,判断该图像是否有可能包含字符;选择单元,被配置为如果该图像被判断为有可能包含字符,则通过使用所提取的特征集,从一组具体字符检测器中选择要对该图像使用的具体字符检测器子集;以及检测单元,被配置为通过使用所选择的具体字符检测器子集中的具体字符检测器,对该图像进行检测,以识别该图像中的具体字符。

[0006] 此外,为了解决以上技术问题中的至少一个,本发明提供一种字符识别方法,该字符识别方法包括:提取步骤,从图像中提取特征集,所述特征集包含一个或更多个特征;判断步骤,通过使用所提取的特征集,判断该图像是否有可能包含字符;选择步骤,如果该图像被判断为有可能包含字符,则通过使用所提取的特征集,从一组具体字符检测器中选择要对该图像使用的具体字符检测器子集;以及检测步骤,通过使用所选择的具体字符检测器子集中的具体字符检测器,对该图像进行检测,以识别该图像中的具体字符。

[0007] 此外,为了解决以上技术问题中的至少一个,本发明提供一种图像处理装置,该图像处理装置包括:滑动窗应用单元,被配置为对输入图像应用滑动窗以形成多个部分图像;尺寸改变单元,被配置为如果所述滑动窗的尺寸小于预定尺寸,则增大所述滑动窗应用单元的滑动窗的尺寸,其中,如果尺寸改变单元增大了滑动窗的尺寸,则所述输入图像被再次

输入到所述滑动窗应用单元;以及如前所述的字符识别装置,用于处理从所述滑动窗应用单元输出的所述多个部分图像。

[0008] 此外,为了解决以上技术问题中的至少一个,本发明提供一种图像处理方法,该图像处理方法包括:滑动窗应用步骤,对输入图像应用滑动窗以形成多个部分图像;尺寸改变步骤,如果所述滑动窗的尺寸小于预定尺寸,则增大所述滑动窗的尺寸,其中,如果在尺寸改变步骤中增大了滑动窗的尺寸,则所述输入图像被再次输入到所述滑动窗应用步骤;以及执行如前所述的字符识别方法,以处理从所述滑动窗应用步骤输出的所述多个部分图像。

[0009] 通过使用根据本发明的字符识别装置和字符识别方法,能够快速检测和识别自然图像中的字符。此外,根据本发明的一些实施方式,还能够快速检测和识别自然图像中的字符,而不论与字符相关联的诸如缩放尺度、照明条件、模糊或复杂背景等的因素。

[0010] 从参照附图的以下描述,本发明的其它特性特征和优点将变得清晰。

## 附图说明

[0011] 并入说明书中并且构成说明书的一部分的附图图示本发明的实施例,并且与描述一起用于说明本发明的原理。

[0012] 图1是示出可实施本发明的实施例的计算机系统1000的硬件配置的框图。

[0013] 图2示出根据本发明的一个例子的字符识别装置2000的示意性功能框图。

[0014] 图3是示出根据本发明的从图像检测和识别字符的基本处理的流程图。

[0015] 图4是示出根据本发明的一个例子的用于提取特征集的示例性处理的流程图。

[0016] 图5是示出根据本发明的一个例子的用于查找一个特征的示例性处理的流程图。

[0017] 图6示出根据本发明的一个例子的在具体图像中确定特征集的示例性例子。

[0018] 图7是根据本发明的一个例子的用于对图像执行边缘检测方法以生成有向边缘图像的流程图。

[0019] 图8a和8b示出根据本发明的一个例子的用于检测梯度的边缘梯度检测器的示例性例子。

[0020] 图9a示出原始图像的例子,图9b示出该原始图像的有向边缘图像。

[0021] 图10示出根据本发明的一个例子的示例性金字塔形体系结构系统。

[0022] 图11是用于说明图10中所示的系统中的示例性工作流程的流程图。

[0023] 图12是根据本发明的字符识别装置可应用于的示例性图像处理装置的图。

[0024] 图13是示出由图像处理装置执行的处理的流程图。

[0025] 图14是根据本发明的一个例子的从照片识别车牌照号码的示意性流程图。

[0026] 图15a~15c是用于说明图14中所示的处理的示意性视图。

## 具体实施方式

[0027] 以下将参照附图详细描述本发明的实施例。

[0028] 请注意,类似的参考数字和字母指的是图中的类似的项目,因而一旦在一幅图中定义了一个项目,就不需要在之后的图中讨论了。

[0029] 首先,将说明本公开的上下文中的一些术语的含义。

[0030] 在本公开中,较亮的像素意味着具有较高像素值的像素。类似地,较暗的像素意味着具有较低像素值的像素。

[0031] 在本公开中,“字符”意味着相对简单的字符集中的元素。这种简单字符集的例子包括数字(0~9)的集合、预定符号(例如,+、-或#等等)的集合、字母表(例如,英文字母表或希腊字母表)。所述简单字符集中的元素数量一般小于100。图像中的字符不包括图像的背景。也就是说,当提及“字符”时,其表示构成该字符的笔画或元素。例如,如果存在在白色背景上书写黑色字符的图像,则字符或笔画意味着黑色的像素。

[0032] 在本公开中,表示方向的术语,诸如“上”、“下”、“左”、“右”、“顶部”、“底部”、“垂直”和“水平”并不被解释为限于字符正立的情况(即字符以正常方式被阅读的情况)。诸如“上”、“下”、“左”、“右”、“顶部”、“底部”、“垂直”和“水平”之类的方向是当预定方向被看作“向上”时的相对方向也是可以的。一旦预定方向被看作“向上”,则其它相对方向可被相应地确定。例如,这些方向可以是字符与被阅读的状态相反地设置的情况下观看的方向。术语“水平”意味着大致从左向右的方向或大致从右向左的方向。术语“垂直”意味着与水平方向垂直的方向。术语“x方向”意味着从左向右的方向。术语“y方向”意味着从上向下的方向。

[0033] 在本公开中,“边缘”意味着像素值具有梯度的像素。在某方向上从亮到暗的边缘意味着在该方向上变暗(在该方向上具有正梯度)的边缘。在某方向上从暗到亮的边缘意味着在该方向上变亮(在该方向上具有负梯度)的边缘。

[0034] 在本公开中,“水平边缘”意味着像素值具有大致垂直的梯度分量(即,像素值的垂直梯度分量的绝对值大于其水平梯度分量的绝对值)的边缘,“垂直边缘”意味着像素值具有大致水平的梯度分量(即,像素值的水平梯度分量的绝对值大于其垂直梯度分量的绝对值)的边缘。换言之,“水平边缘”意味着在大致水平方向上延伸的边缘,“垂直边缘”意味着在大致垂直方向上延伸的边缘。

[0035] 在本公开中,潜在字符意味着被假设为相对于背景构成字符的像素的集合。也就是说,有可能潜在字符并非真实字符,而是作为字符被对待或处理。

[0036] 在本公开中,潜在笔画意味着被假设为是字符的笔画的像素的连通集合。也就是说,有可能潜在笔画并非字符的真实笔画(例如,存在图像甚至根本不包含字符的情况),而是作为字符的笔画被对待或处理。

[0037] 在本公开中,边缘图像的边界与用于获得边缘图像的图像的边界是相同的。

[0038] 图1是示出实施本发明的实施例的计算机系统1000的硬件配置的框图。

[0039] 如图1中所示,计算机系统包括计算机1110。计算机1110包括经由系统总线1121连接的处理单元1120、系统存储器1130、固定非易失性存储器接口1140、可移动非易失性存储器接口1150、用户输入接口1160、网络接口1170、视频接口1190和输出外围接口1195。

[0040] 系统存储器1130包括ROM(只读存储器)1131和RAM(随机存取存储器)1132。BIOS(基本输入输出系统)1133驻留在ROM 1131中。操作系统1134、应用程序1135、其它程序模块1136和某些程序数据1137驻留在RAM 1132中。

[0041] 诸如硬盘之类的固定非易失性存储器1141连接到固定非易失性存储器接口1140。固定非易失性存储器1141例如可以存储操作系统1144、应用程序1145、其它程序模块1146和某些程序数据1147。

[0042] 诸如软盘驱动器1151和CD-ROM驱动器1155之类的可移动非易失性存储器连接到



可移动非易失性存储器接口1150。例如,软盘1152可以被插入到软盘驱动器1151中,以及CD(光盘)1156可以被插入到CD-ROM驱动器1155中。

[0043] 诸如鼠标1161和键盘1162之类的输入设备被连接到用户输入接口1160。

[0044] 计算机1110可以通过网络接口1170连接到远程计算机1180。例如,网络接口1170可以经由局域网1171连接到远程计算机1180。或者,网络接口1170可以连接到调制解调器(调制器-解调器)1172,以及调制解调器1172经由广域网1173连接到远程计算机1180。

[0045] 远程计算机1180可以包括诸如硬盘之类的存储器1181,其存储远程应用程序1185。

[0046] 视频接口1190连接到监视器1191。

[0047] 输出外围接口1195连接到打印机1196和扬声器1197。

[0048] 图1所示的计算机系统仅仅是说明性的并且决不意图对本发明、其应用或用途进行任何限制。

[0049] 图1所示的计算机系统可以被实施于任何实施例,可作为独立计算机,或者也可作为设备中的处理系统,可以移除一个或更多个不必要的组件,也可以向其添加一个或更多个附加的组件。

[0050] 图2示出根据本公开的字符识别装置2000的说明性功能框图。

[0051] 字符识别装置2000包括:提取单元100,被配置为从图像中提取特征集,所述特征集包含一个或更多个特征;判断单元200,被配置为通过使用所提取的特征集,判断该图像是否有可能包含字符;选择单元300,被配置为如果该图像被判断为有可能包含字符,则通过使用所提取的特征集,从一组具体字符检测器中选择要对该图像使用的具体字符检测器子集;以及检测单元400,被配置为通过使用所选择的具体字符检测器子集中的具体字符检测器,对该图像进行检测,以识别该图像中的具体字符。

[0052] 根据一个例子,所述特征集的所述一个或更多个特征可以是沿预定方向被依次提取的一个或更多个边缘特征。所述一个或更多个边缘特征中的每一个包含第一水平边缘、第二水平边缘、第一垂直边缘和第二垂直边缘中的至少一个。将所述预定方向的上游方向作为上方,所述第一水平边缘是潜在笔画的下边缘,所述第二水平边缘是潜在笔画的上边缘,所述第一垂直边缘是图像左部的潜在笔画的左边缘,所述第二垂直边缘是图像右部的潜在笔画的右边缘。

[0053] 根据一个例子,所述第一水平边缘可以是在从上向下的方向上从暗到亮的边缘,所述第二水平边缘可以是在从上向下的方向上从亮到暗的边缘,所述第一垂直边缘可以是在从左向右的方向上从亮到暗的边缘,所述第二垂直边缘可以是在从左向右的方向上从暗到亮的边缘。

[0054] 根据一个例子,所述提取单元100可包括:边缘图像获得单元110和特征确定单元120。该边缘图像获得单元110被配置为获得所述图像的边缘图像。所述特征确定单元120被配置为:

[0055] 在该边缘图像中,沿从上向下的方向依次查找所述第一水平边缘和第二水平边缘,然后沿水平方向查找所述第一垂直边缘和第二垂直边缘,以便确定一个特征;以及如果所述第二水平边缘不位于所述边缘图像的最底部,则继续向下查找下一个特征。其中,对于一个特征:如果该特征是该图像的第一个特征,则所述边缘图像的最顶部的预定点作为查

找的起点并且作为所述第一个特征的第一水平边缘;如果直到所述边缘图像的最底部的点仍未找到从上向下的方向上从亮到暗的边缘,则该特征确定单元将所述最底部的点作为该特征的第二水平边缘;如果第一水平边缘位于边缘图像的最顶部,则该特征确定单元将该特征确定为不包含第一水平边缘;如果第二水平边缘位于边缘图像的最底部,则该特征确定单元将该特征确定为不包含第二水平边缘。

[0056] 根据一个例子,字符识别装置2000可进一步包括:粗检测单元500,被配置为对该图像的边缘图像进行初步检测,以便检测该图像是否是潜在的包含字符的图像,其中,仅在该粗检测单元500确定该图像是潜在的包含字符的图像的情况下,所述提取单元100、判断单元200、选择单元300和检测单元400才对该图像进行相应操作。

[0057] 根据一个例子,图像处理装置3000可包含字符识别装置2000,将参照图12更加详细地描述该图像处理装置3000。

[0058] 以上描述的单元和以下要描述的单元是用于实施以下要描述的处理的示例性和/或优选的模块。这些模块可以是硬件单元(诸如场可编程门阵列、数字信号处理器或专用集成电路等)和/或软件模块(诸如计算机可读程序)。以上并未详尽地描述用于实施各个步骤的模块。然而,只要有执行某个处理的步骤,就可以有用于实施同一处理的对应的功能模块或单元(由硬件和/或软件实施)。通过以下描述的步骤以及与这些步骤对应的单元的所有组合限定的技术方案都被包括在本申请的公开内容中,只要它们构成的这些技术方案是完整并且可应用的。

[0059] 此外,由各种单元构成的上述设备和/或系统可以作为功能模块被并入到诸如计算机或照相机之类的硬件装置中。除了这些功能模块之外,计算机或照相机当然可以具有其他硬件或者软件部件。

[0060] 现在,将参照图3描述本发明的实施例。

[0061] 图3是示出根据本发明的从图像检测和识别字符的基本处理的流程图。

[0062] 在步骤S100中,从图像中提取特征集,所述特征集包含一个或更多个特征。所述特征集可描述图像的基本结构,并且,如果图像包含字符,则所述特征集可描述该字符的基本结构。

[0063] 根据本发明,一些特征集可对应于一个或更多个具体字符(具体字符子集),而其它特征集不对应于任何字符。因此,根据本发明的 图像的特征集披露关于图像是否可能包含字符以及图像可能包含哪些字符的信息。

[0064] 在步骤S200中,通过使用在步骤S100中提取的特征集,判断图像是否有可能包含字符。

[0065] 如果判断该图像有可能包含字符(步骤S200中为“是”),则处理进行到步骤S300,在步骤S300,通过使用在步骤S100中提取的特征集,从一组具体字符检测器中选择要用于该图像的具体字符检测器子集。

[0066] 在步骤S400中,通过使用在步骤S300中选择的具体字符检测器子集中的具体字符检测器,对图像(或图像的边缘图像)执行检测,直到识别出图像中包含的具体字符。

[0067] 如果判断该图像不可能包含字符(步骤S200中为“否”),则处理结束。

[0068] 根据一个例子,在步骤S300中,可以通过使用特征集和与具体字符检测器子集对应的具体字符子集之间的预定的映射来选择要用于该图像的具体字符检测器子集。

[0069] 根据图3中所示的以上处理,通过使用从图像提取的特征集,仅全部具体字符检测器的子集中的具体字符检测器要被应用于一个图像。在一个例子中,具体字符检测器子集中的具体字符检测器顺序地处理图像,直到识别出具体字符为止。因此,要对一幅图像所使用的具体字符检测器的最大数量可被降低,从而识别所需要的时间可减少。

[0070] 可由特征提取器803(见图10)实施步骤S100(以及提取单元100)和S200(以及判断单元200)。可由字符子集检测器805(见图10)实施步骤S300(以及选择单元300)。可由具体字符检测器806(见图10)实施步骤S400(以及检测单元400)。

[0071] 现在,将更加详细地描述步骤S100中的处理。

[0072] 根据一个例子,在步骤S100(提取单元100)中,沿预定方向依次提取一个或更多个边缘特征,作为一幅图像的特征集。所述一个或更多个边缘特征中的每一个包含第一水平边缘、第二水平边缘、第一垂直边缘和第二垂直边缘中的至少一个。当将所述预定方向的上游作为上侧时,对于一个边缘特征,第一水平边缘是(潜在)笔画的下边缘,第二水平边缘是(潜在)笔画的上边缘,第一垂直边缘是图像的左边部分中的(潜在)笔画的左边缘,第二垂直边缘是图像的右边部分中的(潜在)笔画的右边缘。

[0073] 也就是说,当图像包含较亮的背景和较暗的字符时,第一水平边缘是在从上向下的方向上从暗到亮的边缘,第二水平边缘是在从上向下的方向上从亮到暗的边缘,第一垂直边缘是在从左向右的方向上从亮到暗的边缘,第二垂直边缘是在从左向右的方向上从暗到亮的边缘。

[0074] 类似地,当图像包含较暗的背景和较亮的字符时,第一水平边缘是在从上向下的方向上从亮到暗的边缘,第二水平边缘是在从上向下的方向上从暗到亮的边缘,第一垂直边缘是在从左向右的方向上从暗到亮的边缘,第二垂直边缘是在从左向右的方向上从亮到暗的边缘。

[0075] 现在将参照图4描述步骤S100(提取单元100)的上述例子的示例性处理。

[0076] 在步骤S110中,获得图像的边缘图像。该边缘图像可使用边缘检测方法来获得,或者可从诸如存储器装置或有向边缘检测器之类的另一功能部件获得。

[0077] 在步骤S120中,在边缘图像中查找特征。如果存在特征,则确定该特征。

[0078] 在步骤S130中,判断是否要查找下一个特征。

[0079] 如果在步骤S130中判断要查找下一个特征(步骤S130中为“是”),则处理返回到步骤S120以查找下一个特征。

[0080] 如果在步骤S130中判断不要查找下一个特征(步骤S130中为“否”),则处理结束,并且所确定的一个或更多个特征构成该图像的特征集。

[0081] 通过使用上述的四种边缘特征,该方法对于诸如照明条件、模糊或复杂背景等等因素可以是稳健的。

[0082] 图5是示出用于在步骤S120中查找一个特征的示例性处理的流程图。

[0083] 在步骤S121中,对于当前特征,在从上向下的方向上搜索边缘图像(有向边缘图像),以便依次查找第一水平边缘作为起点然后查找第二水平边缘作为终点。如上所述,第一水平边缘是潜在笔画的下边缘,第二水平边缘是潜在笔画的上边缘。

[0084] 参照以下的(1)~(3)来描述用于在步骤S121中对于一个特征实施水平边缘查找的示例性处理。对于其中背景较亮而字符较暗的图像来描述此示例性处理。然而,可将类似

的处理应用于其中背景较暗而字符较亮的图像的情况。

[0085] (1)确定当前特征是否是图像的第一个特征。

[0086] (2)如果当前特征是该图像的第一个特征,则该图像(有向边缘图像)的顶边界上的水平方向上的中途点被作为当前特征的起点。水平方向上的中途点可以是作为顶边界的中点的点,或者可以是近似顶边界的中点的点。例如,顶边界的中途点与中点之间的距离可以不大于顶边界的长度(左边界和右边界之间的距离)的1/8。

[0087] 如果当前特征不是该图像的第一个特征,则在找到先前特征的终点之后,从所述先前特征的终点起在垂直方向上向下对边缘图像进行搜索。如果找到潜在笔画的下边缘(在本示例性处理中为从暗到亮的边缘),则将该边缘设定为当前特征的起点。查找水平边缘所沿的直线被称为垂直中途线。

[0088] (3)从起点开始在垂直方向上向下对边缘图像进行搜索。如果找到潜在笔画的上边缘(在本示例性处理中为在从上向下的方向上的从亮到暗的边缘),将该边缘设定为终点。如果直到图像(边缘图像)的底边界仍未找到终点,则垂直中途线与图像的底边界相交处的点被设定为终点。

[0089] 通过上述的(1)~(3),一个特征的起点和终点被确定。以上的(1)~(3)给出了用于在步骤S121中对于一个特征实施水平边缘查找的示例性处理,其中,在从上向下的方向上查找起点和终点。然而,如上所述,“上”和“下”仅是相对概念。任何预定方向可被看作“从上向下”的方向,在这种情况下,预定方向的上游将为“上”侧,诸如“下”、“左”和“右”之类的其它朝向可被相应地确定。

[0090] 在步骤S122中,在水平方向上对边缘图像(有向边缘图像)进行搜索,以查找作为左侧点的第一垂直边缘和作为右侧点的第二垂直边缘。如上所述,第一垂直边缘是图像的左侧部分上的潜在笔画的左边缘,第二垂直边缘是图像的右侧部分上的潜在笔画的右边缘。

[0091] 参照以下的(4)~(6)来描述用于在步骤S122中对于一个特征实施垂直边缘查找的示例性处理。对于其中背景较亮而字符较暗的图像来描述此示例性处理。然而,可将类似的处理应用于其中背景较暗而字符较亮的图像的情况。

[0092] (4)在找到了当前特征的起点和终点两者之后,设定垂直方向上的中途点,所述中途点是同一当前特征的起点与终点之间的中点或者近似该中点。在一个例子中,此中途点与起点和终点间的中点之间的距离不大于所述起点和中点间的距离的1/8。

[0093] (5)在(4)中设定中途点之后,对边缘图像进行搜索以查找第一垂直边缘和第二垂直边缘。更特别地,从所述中途点起在水平方向上向左查找潜在笔画的左边缘(在本示例性处理中为在从左向右的方向上的从亮到暗的边缘)作为左侧点,并且,从所述中途点起在水平方向上向右查找潜在笔画的右边缘(在本示例性处理中为在从左向右的方向上的从暗到亮的边缘)作为右侧点。

[0094] (6)如果直到图像(边缘图像)的左边界仍未找到左侧点,则将当前特征的左侧点设定为“无效”。类似地,如果直到图像(边缘图像)的右边界仍未找到右侧点,则将当前特征的右侧点设定为“无效”。

[0095] 在步骤S123中,根据起点、终点、左侧点和右侧点来确定特征。更特别地,如下来确定一个特征:

[0096] -如果起点不处于图像的顶边界,则该特征被确定为包含第一水平边缘。如果起点位于图像的顶边界,则该特征被确定为不包含第一水平边缘。

[0097] -如果终点不处于图像的底边界,则该特征被确定为包含第二水平边缘。如果终点位于图像的底边界,则该特征被确定为不包含第二水平边缘。

[0098] -如果找到了左侧点,则该特征被确定为包含第一垂直边缘。如果左侧点被设为了“无效”,则该特征被确定为不包含第一垂直边缘。

[0099] -如果找到了右侧点,则该特征被确定为包含第二垂直边缘。如果右侧点被设为了“无效”,则该特征被确定为不包含第二垂直边缘。

[0100] 到目前为止,已通过步骤S120(步骤S121、S122和S123)确定了一个特征。以下的表1示出了根据本公开的所有可能的特征。

[0101] 表1

	1	2	3	4
A				
B				
C				
D				

[0102] 在以上的表1中,上方的水平线表示第一水平边缘(即,由起点确定的潜在笔画的下边缘),下方的水平线表示第二水平边缘(即,由终点确定的潜在笔画的上边缘),左侧的垂直线表示第一垂直边缘,右侧的垂直线表示第二垂直边缘。如表1中所示,一个特征包括第一水平边缘、第二水平边缘、第一垂直边缘和第二垂直边缘中的至少一个。

[0103] 表1还示出用于对特征分类的方式。A类中的每个特征包括第二水平边缘但是不包括第一水平边缘。B类中的每个特征包括第一水平边缘和第二水平边缘两者。C类中的每个特征包括第一水平边缘但是

[0104] 不包括第二水平边缘。D类中的每个特征既不包括第一水平边缘也不包括第二水平边缘。

[0105] 然后,可以记录所确定的特征的类型、位置和大小以供之后使用。

[0106] 在步骤S130中,如果当前特征的终点不在图像的底边界(也就是说,如果当前特征包含第二水平边缘),则判断为要查找下一个特征(步骤S130中为“是”)。否则,如果当前特征的终点在图像的底边界(也就是说,如果当前特征不包含第二水平边缘),则判断为不要查找下一个特征(步骤S130中为“否”),并且特征集中的全部特征已被确定。

[0107] 特征集中的特征按照沿从上向下的方向的位置的次序被排序(即,按照其起点的位置的次序被排序)。

[0109] 图6示出确定具体图像中的特征集的示例性例子。在图6中,外部的矩形框是图像边界。图像的背景是较亮的并且潜在字符(潜在笔画)是较暗的。

[0110] 如图6中所示,首先查找第一个特征。如上所述,该图像的第一个特征的起点被确定为顶边界的中途点(以下被称为点SP1)。然后,沿垂直中途线(以下被称为线 $L_{VM}$ )查找第一个特征的终点,并且将该第一个特征的终点找到作为点EP1。然后,SP1和EP1之间的中途点被确定为 $P1_{VM}$ ,并且分别沿向左的方向和向右的方向查找第一个特征的左侧点和右侧点。如图6所示,左侧点被找到为点LP1,而直到图像的右边界仍未找到右侧点。因此,第一个特征的右侧点RP1被设为“无效”。从而,由于第一个特征的起点是图像的顶边界处的SP1而第一个特征的右侧点是“无效”,因此图像的第一个特征被确定为由第一垂直边缘和第二水平边缘构成(表1中的特征A3)。

[0111] 由于第一个特征的终点EP1不在图像的底边界处,因此还要沿着从点EP1起的线 $L_{VM}$ 向下查找下一个特征的起点和终点。第二个特征的起点被找到为点SP2,直到图像的底边界为止仍未找到第二个特征的终点。因此,第二个特征的终点被设为点EP2,该点EP2是线 $L_{VM}$ 与图像的底边界的交点。然后,从点SP2和点EP2之间的中途点 $P2_{VM}$ 开始查找第二个特征的左侧点和右侧点。直到图像的边界,既没有找到左侧点也没有找到右侧点,因此第二个特征的左侧点LP2和右侧点RP2均被设定为“无效”。由于第二个特征的终点是处于图像的底边界处的EP2并且第二个特征的左侧点和右侧点为“无效”,因此该图像的第二特征被确定为仅由第一水平边缘构成(表1中的特征C4)。

[0112] 由于第二个特征的终点是处于图像的底边界处的EP2,因此不存在下一个特征,特征查找处理结束。

[0113] 因此,图6中所示的图像的特征集被提取,该特征集包括两个边缘特征,它们是表1中所示的特征A3和特征C4。特征集中的特征的次序是按照沿从上向下的方向的位置的次序,即依次为A3和C4。

[0114] 现在将更详细地描述步骤S200中的处理。

[0115] 在对于图像确定了特征集之后,基于在步骤S100中或者由提取单元100提取的特征集,判断(步骤S200,判断单元200)该图像是否有可能包含字符。

[0116] 根据一个例子,仅当图像的特征集满足所有以下条件1~4时才判断为该图像有可能包含字符。

[0117] 条件1:如果特征集仅具有一个特征,则该特征应为表1中所示的D类中的一个。

[0118] 条件2:如果特征集具有多于一个的特征,则第一个特征应为表1中所示的A类中的一个。

[0119] 条件3:如果特征集具有多于一个的特征,则最后一个特征应为表1中所示的C类中的一个。

[0120] 条件4:如果特征集具有多于一个的特征,则既不是第一个特征也不是最后一个特征的特征应为表1中所示的B类中的一个。

[0121] 现在将更详细地描述步骤S300中的处理。

[0122] 如果在步骤S200中判断为存在该图像包含字符的可能性,则在步骤S300中,从一组具体字符检测器中选择具体字符检测器子集。

[0123] 根据一个例子,字符识别装置2000从一组具体字符中识别出一个字符。在这种情

况下,可以有一组具体字符检测器,所述一组具体字符检测器中的每一个对应于该组具体字符中的一个具体字符。因此,如果该组具体字符中的具体字符的数量为N,则具体字符检测器的数量也应为N。例如,如果该组具体字符是数字集合(数字“0”~“9”),则该组具体字符检测器包括分别对应于数字“0”~“9”的十个具体字符检测器。


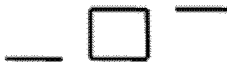


[0124] 另一方面,由于图像的特征集可描述图像中的潜在字符的基本结构,因此可根据该图像的特征集来确定包含该图像的可能候选的字符子集。

[0125] 根据一个示例性例子,按照上述的特征提取处理,具体字符集中的每个字符均具有确定的特征集。例如,如以上例示的那样,字符“L”具有依次由“A3”和“C4”构成的特征集。此外,按照以上的描述,可清楚地知晓,字符“8”具有依次由“A4”、“B1”、“B1”和“C4”构成的特征集。也就是说,根据如以上例示的特征提取处理,可以在特征集与具体字符子集之间建立映射。一个特征集有可能对应于一个或多个具体字符,在这种情况下,仅与所述一个或多个具体字符对应的具体字符检测器子集要被选择用于该图像。

[0126] 一个特征集不对应于任何具体字符也是可能的。如果特征集被确定为不对应于任何具体字符,则该图像被判断为不包含字符。

[0127] 根据一个示例性例子,字符和特征集之间的映射可以具有查找表的形式,该查找表描述各特征集与包括一个或多个字符的字符子集之间的映射。表2示出用于包括英文字母和数字的字符集的一个示例性映射(查找表)的一部分。

[0128] 表2

特征数量	特征代码	特征集	字符子集
2	A1 C1		H
3	A4 B1 C4		0 D
3	A4 B3 C3		F
4	A4 B3 B1 C4		6
4	A4 B3 B2 C4		5 S

[0130] 在以上的表2中,列“特征数量”表示特征集中的特征的数量,列“特征代码”表示特征集中的特征的特征代码(例如,如表1中所示)的序列(按照特征集中从第一个特征起的次序),列“特征集”表示特征集的特征的序列(按照特征集中从第一个特征起的次序),列“字符子集”表示可对应于该特征集的可能的字符(字符子集)。

[0131] 请注意,表2仅仅示出示例性映射的一部分,可以通过对字符集中的每一个具体字符执行步骤S100~S200中描述的特征提取处理而建立对于其它特征集(字符)的映射。此外,虽然表2示出了用于包括英文字母和数字的字符集的示例性映射,但是可以通过对字符集中的每一个具体字符执行步骤S100~S200中描述的特征提取处理而建立对于其它字符集(诸如对于希腊字母集或日文字母集)的映射。可在图3的整个流程之前事先建立该映射。

[0132] 此外,虽然通过使用特征集中的特征的数量和特征集的特征代码(见表1)的组合作为关键字并且使用相关联的字符子集作为值而构建了表2中示出的查找表,但是当然可以采用其它形式的查找表,只要该查找表示出特征集与字符子集之间的映射即可。此外,该

映射不必呈现查找表的形式。该映射当然可采取诸如函数之类的其他形式，只要该映射示出特征集和字符子集之间的关系即可。

[0133] 在步骤S300中，可使用特征集和具体字符子集之间的映射，从所述一组具体字符检测器中确定具体字符检测器子集。更具体地，如果根据映射，图像的特征集对应于一个字符子集，则对于该图像选择对应于所述字符子集的具体字符检测器子集。

[0134] 当使用以上所示的表2时，选择具体字符检测器子集的一个示例性处理例如可以如下。首先，在查找表(“特征数量”列)中查找图像的特征集中的正确的特征数量；在具有该正确的特征数量的行中查找正确的特征代码；然后确定与该正确的特征代码对应的正确的字符子集。在确定字符子集之后，可相应地选择具体字符检测器子集。

[0135] 具体字符检测器子集中的具体字符检测器的数量通常比所述一组具体字符检测器中的具体字符检测器的数量小得多。因此，可能用于图像的具体字符检测器的最大数量可大大地减小。

[0136] 现在，将更详细地描述步骤S400中的处理。

[0137] 存在多个具体字符检测器，每个所述具体字符检测器被用于检测相应的具体单个字符。所述多个具体字符检测器构成与该组字符对应的一组具体字符检测器。通过使用在步骤S300中选择的具体字符检测器子集中的具体字符检测器而对图像执行检测，直到识别出该图像中所包含的具体字符。在一个例子中，由具体字符检测器子集中的具体字符检测器顺序地处理图像，直到一个具体字符检测器检测出该图像包含一个具体字符。也就是说，如果具体字符检测器子集中的任何具体字符检测器将该图像看作包含字符，则该图像被确定为包含字符的图像，并且可由该具体字符检测器识别出该字符。如果具体字符检测器子集中的任何具体字符检测器都不将该图像看作包含字符，则该图像被确定为不包含字符的图像。

[0138] 在一个例子中，所有具体字符检测器被事先训练，并且每一个具体字符检测器可确定图像是否包含特定字符。许多用于生成具体字符检测器的方法可被用于这里。例如，这里可以使用诸如被称为“SVM”方法的方法(例如参阅Thorsten Joachims, “Transductive Inference for Text Classification using Support Vector Machines”, Proceedings of the 1999 International Conference on Machine Learning(ICML 1999), pp.200-209)和被称为“AdaBoost”方法的方法(例如参阅AdaBoost and the Super Bowl of Classifiers-A Tutorial on AdaBoost)之类的方法。

[0139] 根据一个例子，字符识别装置2000可包括粗检测单元500，所述粗检测单元500在步骤S100之前进行粗略检测。所述粗检测单元500被配置为对图像的边缘图像执行初步检测，以便检测该图像是否是潜在的包含字符的图像。如果粗检测单元500确定该图像不是潜在的包含字符的图像(即，如果该图像无论如何也不可能包含字符)，则该图像将不经受提取单元100、判断单元200、选择单元300和检测单元400进行的处理(步骤S100、S200、S300和S400中的处理)，并且将被看作不包含字符的图像(即，将被丢弃)。也就是说，仅在该粗检测单元500确定为该图像是潜在的包含字符的图像的情况下，所述提取单元100、判断单元200、选择单元300和检测单元400才执行相应操作(步骤S100、S200、S300和S400)。

[0140] 得益于粗检测单元500，即使是要对大量的图像识别字符，也可通过粗检测单元500过滤掉明显不包含字符的一定比例的图像。因此，单元100~400的工作负荷可被降低，



并且检测可被加速。

[0141] 用于实施粗检测单元500的一个示例性例子如下。

[0142] 更具体地,可通过边缘密度检测器来实现粗检测单元500,所述边缘密度检测器能够使用有向边缘图像而非常快地检测。如果图像的有向边缘图像的像素密度在预定范围之内,则该图像被确定为潜在的包含字符的图像。否则如果图像的有向边缘图像的像素密度在预定范围之外,则该图像被确定为不包含字符的图像。根据一个例子,边缘密度检测器可利用有向边缘图像的积分图图像以便加速检测。

[0143] 现在将描述对输入图像的边缘图像的检测。在这里,将以优选的边缘检测方法作为例子。然而,存在本领域技术人员已知的许多边缘检测方法,并且,本发明不被具体的边缘检测方法限制。这里描述的边缘检测方法可由边缘图像获得单元110在步骤S110中使用,或者可以由边缘图像获得单元110从其获得边缘图像的功能模块使用。

[0144] 根据一个例子,执行边缘检测方法以产生有向边缘图像。该边缘检测方法计算输入图像的每个像素处的边缘梯度向量,并且将边缘确定为四种有向边缘之一。前文中提及的边缘图像可以是这种有向边缘图像。

[0145] 图7是用于对图像执行边缘检测方法以产生有向边缘图像的流程图。

[0146] 在步骤S111中,使用具有两个掩模的边缘梯度检测器来检测图像中的各点的梯度。这两个掩模与原始图像卷积,以近似梯度向量。一个掩模(被称为第一掩模,其例子在图8a中所示)执行关于X方向的偏导,而另一掩模(被称为第二掩模,其例子在图8b中所示)执行关于Y方向的偏导。

[0147] Sobel检测器是梯度检测器的一个例子,其两个掩模说明性地示于图8a和图8b中。

[0148] 第一掩模与具有相应尺寸的原始图像像素区域的卷积被计算,作为该像素区域中的中心像素的X梯度( $P_{X\text{-gradient}}$ ,X方向上的梯度分量)。第二掩模与具有相应尺寸的原始图像像素区域的卷积被计算,作为该像素区域中的中心像素的Y梯度( $P_{Y\text{-gradient}}$ ,Y方向上的梯度分量)。

[0149] 对于原始图像中的每个点,通过使用这两个掩模计算边缘梯度向量( $P_{X\text{-gradient}}$ , $P_{Y\text{-gradient}}$ ),并且将其分别保存在X梯度图像和Y梯度图像中。通过下式来计算每个点的梯度的幅值 $P_{\text{magnitude}}$ :

[0150] 根据步骤S111,图像中的每个点的X梯度( $P_{X\text{-gradient}}$ )、Y梯度( $P_{Y\text{-gradient}}$ )和梯度的幅值可被计算以产生X梯度图像、Y梯度图像和幅值图像。X梯度图像、Y梯度图像和幅值图像构成梯度向量图像。

[0151] 这里的幅值图像是灰度级图像,无论原始图像是灰度级图像还是二值图像。当要在以下的步骤S113中执行细化处理时,对幅值图像的二值化处理将是必要的。在步骤S112中,通过执行幅值图像的二值化,产生二值边缘图像。由于梯度检测已大大抑制了原始图像中的噪声,因此二值化处理不必是复杂的。例如,可使用公知为“最大类间方差”方法(也称为“Otsu”方法)的二值化方法(例如参阅N.Otsu,A threshold selection method from gray level,IEEE Transactions On System,Man.And Cybernetics,Vol.9,No.1,pp62-66,1979)。此外,这里也可以使用诸如公知为“Niblack”方法的局部阈值法二值化方法(例如参阅W.Niblack,An Introduction to Image Processing,Prentice-Hall,1986,pp:115-116)或动态阈值法二值化方法(例如参阅Bernsen,Dynamic Thresholding of Gray

level Image, ICPR`86: Proc. Int. Conf. on Pattern Recognition, Berlin, Germany, 1986, pp:1251-1255)之类的其它方法。

[0152] 然后,在步骤S113中,二值边缘图像被细化,并且产生细化的边缘图像。细化的边缘图像意味着其中边缘由宽度为单个像素的直线描述的边缘图像。细化处理是优选的,因为在步骤S100(更特别地,在步骤S120中)从细化的边缘图像中确定上述的特征将是更有效率的。许多细化方法可被用于此步骤S113中。例如,公知的被称作“形态细化”的细化方法(例如参阅Edward Dougherty(ed.), *Mathematical morphology in image processing*, M. Dekker, New York, 1993)可在这里被使用。此外,诸如被称作“Rosenfeld”方法的方法(例如参阅R. Stefanelli and A. Rosenfeld, *Some parallel thinning algorithms for digital pictures*, *J. Assoc. Comput. March 18, No. 2, 1971, 255~264*)或被称作“Skeleton”方法的方法(例如参阅C.C. Han and K.C. Fan, *Skeleton generation of engineering drawings via contour matching*, *Pattern Recognition, Vol. 27, No. 2, pp. 261~275, 1994*)之类的其它方法也可在这里被使用。

[0153] 然后,在步骤S114中,将确定细化的边缘图像中的各点的边缘种类。“边缘种类”表示该边缘是垂直边缘(具有X方向的主导梯度的边缘)还是水平边缘(具有Y方向的主导梯度的边缘)。此外,“边缘种类”还表示边缘是X方向或Y方向上的从亮到暗的边缘还是从暗到亮的边缘。X(或Y)方向上的从亮到暗的边缘是指该边缘具有在X(或Y)方向上大于零的梯度。X(或Y)方向上的从暗到亮的边缘是指该边缘具有在X(或Y)方向上小于零的梯度,如8a和8b所示。也就是说,根据(细化的)边缘图像中的各点的边缘种类,可产生有向边缘图像。

[0154] 如上所述,边缘种类可以包括:

[0155] X\_L2D(相当于如上所述的第一垂直边缘):在X方向(从左向右的方向)上从亮到暗的边缘;

[0156] X\_D2L(相当于如上所述的第二垂直边缘):在X方向(从左向右的方向)上从暗到亮的边缘;

[0157] Y\_L2D(相当于如上所述的第二水平边缘):在Y方向(从上向下的方向)上从亮到暗的边缘;

[0158] Y\_D2L(相当于如上所述的第一水平边缘):在Y方向(从上向下的方向)上从暗到亮的边缘。

[0159] 以下是示出用于根据X梯度图像和Y梯度图像确定边缘种类的示例性处理的一个例子。

[0160] 更具体地,当细化的图像中的点是边缘点(具有梯度的点)时,将按照下式那样来评估相应点的在X梯度图像和Y梯度图像中的两个值。

$$[0161] \quad P_{evaluate} = \begin{cases} P_{X-gradient} & \text{if } |P_{X-gradient}| \geq |P_{Y-gradient}| \\ P_{Y-gradient} & \text{if } |P_{X-gradient}| < |P_{Y-gradient}| \end{cases}, \text{ 以及}$$

[0162] 如果 $P_{evaluate} = P_{X-gradient}$ 并且 $P_{X-gradient} \geq 0$ ,则该边缘是X\_L2D;

[0163] 如果 $P_{evaluate} = P_{X-gradient}$ 并且 $P_{X-gradient} < 0$ ,则该边缘是X\_D2L;

[0164] 如果 $P_{evaluate} = P_{Y-gradient}$ 并且 $P_{Y-gradient} \geq 0$ ,则该边缘是Y\_L2D;

[0165] 如果 $P_{evaluate} = P_{Y-gradient}$ 并且 $P_{Y-gradient} < 0$ ,则该边缘是Y\_D2L。

[0166] 然后,产生有向边缘图像,所述有向边缘图像包含每个点的有向边缘,每个所述有向边缘是上述四种边缘中的一个。也就是说,有向边缘图像中的每个边缘是水平方向的(X方向的,从亮到暗或者从暗到亮)或者是垂直方向的(Y方向的,从亮到暗或者从暗到亮)。

[0167] 图9a示出原始图像的例子,图9b示出原始图像の有向边缘图像。图9b中的箭头指示边缘的种类。由指向右的箭头指示的边缘是第一垂直边缘(边缘X\_L2D),由指向左的箭头指示的边缘是第二垂直边缘(边缘X\_D2L),由指向上的箭头指示的边缘是第一水平边缘(边缘Y\_D2L),由指向下的箭头指示的边缘是第二水平边缘(边缘Y\_L2D)。

[0168] 请注意,已参照潜在笔画较暗而背景较亮的情况说明了详细的例子。在潜在笔画较亮而背景较暗的相反情况下,可在实施在本公开中详细描述的和算法和方法时将“亮”和“暗”相互交换。

[0169] 参照图7描述的步骤可以是用于获得有向边缘图像作为步骤S110或边缘图像获得单元110以及通过粗检测单元500进行的粗略检测的步骤中的边缘图像的优选方法。例如,边缘检测步骤或单元可以是特征提取步骤或单元的一部分。作为替换方案,边缘检测步骤或单元也可以是粗检测步骤或单元的一部分。作为替换方案,边缘检测步骤或单元可以是分离的步骤或单元,而其输出被输入到特征提取步骤或单元和/或粗检测步骤或单元。

[0170] 现在将参照图10描述用于实施本发明的示例性实施例。

[0171] 图10示出用于实施本发明的示例性金字塔形体系结构系统。

[0172] 如图10所示,金字塔形体系结构系统包括布置为金字塔形体系结构的三层。一旦图像输入到该金字塔形体系结构系统,该系统将确定该图像是否包含字符以及它是哪个字符。

[0173] 层1包含边缘检测器801和粗检测器802。层2包含特征检测器803和字符子集检测器805。层3包含与该组具体字符的N个字符对应的多个具体字符检测器806(8061~806N)。

[0174] 该图像可被输入到边缘检测器801以产生有向边缘图像。边缘检测器801可操作作为执行参照图7所描述的操作。

[0175] 然后,有向边缘图像被输入到粗检测器802和特征检测器803。粗检测器802可由粗检测单元500来实施。特征检测器803可由提取单元100和判断单元200来实施。也就是说,特征检测器803可通过使用来自粗检测器802和边缘检测器801的输出而提取图像的特征集并且判断该图像是否有可能包含字符,如参照步骤S100和S200描述的那样。

[0176] 有可能包含字符的图像的特征集被输入到字符子集检测器805。字符子集检测器805可接收特征集查找表,在所述特征集查找表中记录了如参照步骤S300在以上描述的特征集与具体字符子集之间的映射。字符子集检测器805可由选择单元300实施或者可以执行上述步骤S300中的处理。

[0177] 然后,由字符子集检测器805从一组具体字符的所有N个字符中选择具体字符子集。

[0178] 虽然总共有与该组具体字符的N个字符对应的N个具体字符检测器806(8061~806N),但是仅由与所选择的具体字符子集对应的具体字符检测器子集中的具体字符检测器处理图像(有向边缘图像),直到所选择的具体字符检测器子集中的具体字符检测器之一识别出该图像实际上包含哪个字符。在图10中所示的例子中,具体字符检测器8061、8062和8063被示出为被选择(以实线箭头示出)作为具体字符检测器子集,其它具体字符检测器被

示出为未被选择(以虚线箭头示出)。在该例子中,图像(有向边缘图像)最多仅由三(小于N)个具体字符检测器处理。

[0179] 图11是用于说明图10中所示的系统中的示例性工作流程的流程图。

[0180] 层1确定输入图像是否是潜在的包含字符的图像。如果层1确定为输入图像不是潜在的包含字符的图像(分支“否”),则流程结束,以确认该图像不包含任何字符。如果层1确定为输入图像是潜在的包含字符的图像(分支“是”),则流程来到层2。

[0181] 层2确定是否存在对应于图像的特征集的具体字符子集。如果层2确定为不存在对应于图像的特征集的具体字符子集(分支“否”),则流程结束,以确认图像不包含任何字符。如果层2确定为存在对应于图像的特征集的具体字符子集(分支“是”),则流程来到层3,以使与所述具体字符子集对应的具体字符检测器处理该字符。

[0182] 层3确定该图像实际上包含哪个具体字符。在此例子中,假设具体字符检测器8061、8062和8063构成所选择的具体字符检测器子集。图像(或者有向边缘图像)依次经具体字符检测器8061、8062和8063的检测。如果具体字符检测器8061、8062和8063中的一个确定为图像(或者有向边缘图像)包含具体字符,则流程结束,以确认该图像包含该具体字符。如果具体字符检测器8061、8062和8063都没有从图像(或者有向边缘图像)中检测出具体字符,则图像(或有向边缘图像)不需要经历其它具体字符检测器的处理,并且流程结束以确认该图像不包含任何字符。

[0183] 图12是字符识别装置2000可被应用于的示例性图像处理装置3000的图。图像处理装置3000包含:滑动窗应用单元3001,被配置为对输入图像应用滑动窗(输入图像有可能包含多个字符)以形成多个部分图像;尺寸改变单元3003,被配置为如果所述滑动窗的尺寸小于预定尺寸,则增大所述滑动窗应用单元3001的滑动窗的尺寸。如果尺寸改变单元3003增大了滑动窗的尺寸,则所述输入图像被再次输入到所述滑动窗应用单元3001。如果滑动窗的尺寸已等于或大于预定尺寸,则输出对所有部分图像的处理结果。图12中所示的示例性图像处理装置3000还包括如上所述的字符识别装置2000,该字符识别装置2000用于处理从滑动窗应用单元3001输出的部分图像(每个部分图像作为输入到图3所示的流程图的“图像”以及作为输入到图7所示的流程图的“输入图像”而被处理,或者作为输入到图10所示的系统的“输入图像”而被处理)。字符识别装置2000可按照任何定时来处理所述部分图像。

[0184] 在本实施例中,字符识别装置2000也可采用图10和11中所示的体系结构和工作流程。

[0185] 图13是示出由图像处理装置3000执行的处理的流程图。

[0186] 在步骤S3001中,滑动窗被应用于输入图像以通过滑动窗应用单元3001形成多个部分图像。

[0187] 该滑动窗是尺寸近似等于预期存在于输入图像中并且期望被检测的最小字符的尺寸的搜索窗。换言之,搜索窗的原始尺寸可以是实验值,其是期望被检测的字符的最小尺寸。对于金字塔形体系结构检测器的性能,原始搜索窗的尺寸(高度和宽度)优选为10个像素或更大。

[0188] 搜索窗在输入图像上逐渐滑动。在每次滑动之前,包含在搜索窗内的像素被提取以创建部分图像。在输入图像上滑动搜索窗时,搜索窗的滑动间隙可被固定或者根据输入图像的内容按需要被设定。在整个输入图像上滑动了搜索窗之后,形成多个部分图像。

[0189] 在步骤S3002中,在步骤S3001中形成的所述多个部分图像由上述的字符识别装置2000(或者由图10所示的系统)处理。以上关于字符识别描述的任何例子或方法可被用于此步骤。作为此步骤的结果,部分图像中的一些将被识别为包含字符并且将记录这些具体字符,而其它部分图像将被识别为不包含字符并且将被丢弃。通过使用根据本公开的一些方法和装置,可快速检测出部分图像中包含的字符。此外,使用根据本公开的一些方法和装置,不太可能包含字符的部分图像可在较早阶段被丢弃。因此,具体字符检测器的工作负荷可显著降低。

[0190] 在步骤S3003中,判断搜索窗的当前尺寸是否小于预定尺寸。如果判断为搜索窗的当前尺寸小于该预定尺寸,则流程来到步骤S3004。否则,如果判断为搜索窗的当前尺寸不小于该预定尺寸,则流程结束以输出部分图像的认识结果。

[0191] 在步骤S3004中,以规定的缩放增量来增大搜索窗的尺寸。所述规定的缩放增量可以是固定的或者根据输入图像的内容而按需要被设定。在步骤S3004之后,流程回到步骤S3001,在步骤S3001中,具有增大的尺寸的滑动窗被应用于原始图像以形成另外的多个部分图像。

[0192] 如图13所示,步骤S3001~S3004的处理被重复,直到搜索窗(滑动窗)的尺寸不小于所述预定尺寸为止。

[0193] 虽然步骤S3002被示为在图13中的该流程图中的步骤S3001和步骤S3003之间执行,但是步骤S3002也可在步骤S3003(在“是”和“否”分支两者中)之后执行,或者甚至在步骤S3003中的判断结果为“是”的情况下在步骤S3004之后执行。此外,可在产生使用当前窗的这些部分图像之后并且在要使用调整大小的窗产生其它部分图像之前,由字符识别装置2000处理这些部分图像。作为替换方案,也可在具有各种尺寸的所有部分图像都已被产生之后,通过字符识别装置2000处理所述部分图像。也就是说,可按照任何定时来执行步骤S3002,只要所有产生的部分图像可被处理即可。

[0194] 本发明可被用于许多应用中。在这里描述用于应用本发明的例子中的一个。该例子是自动车牌照识别系统中的应用,如图14中所示。

[0195] 图14示出从照片中识别车牌照号码(车牌照号)的示意性流程图。

[0196] 在步骤S4001中,由照相机(例如,数字照相机或者摄像机)拍摄车辆。车的照片示于图15a。如图15a中所示,照片包括包含字符的部分也包括不包含字符的部分。

[0197] 在步骤S4002中,根据本发明的方法或装置被应用于该照片以检测字符。图15b示出照片中的被确定为包含字符的部分图像。

[0198] 在步骤S4003中,如图15c中所示,所检测的字符以文本方式被布局,然后作为识别出的车牌照号码(例如包括数字和英文字母)被输出。

[0199] 现在,将讨论根据本公开的方法的技术效果。本发明的发明人通过使用包含数字(0~9)的车牌照而对图10中示出的方法进行了评价。

[0200] 在该评价中,通过具有所有三层的图10中的方法以及通过除了不具有层2以外与图10中所示的方法相同的比较方法来处理一组测试图像。对于每个处理后的图像,将检测出的数字的坐标与事先保存的真实车牌照的坐标进行比较。如果检测出的数字的坐标在真实车牌照的领域内,则将该数字认为是正确的数字。

[0201] 如下计算检测率:

- [0202] 检测率=(实际检测的数字的数量/图像中的真实数字的数量)×100%
- [0203] 这样的数字将被看作是错误检测数字:
- [0204] 如果检测出的数字的坐标不在真实车牌照的领域内,或者
- [0205] 如果检测出的数字的坐标在真实车牌照的领域内但是被认为是正 确数字的数字数量大于实况文件中的数字数量。
- [0206] 如下计算错误检测率:
- [0207] 错误检测率=(错误检测的数字的数量/图像中的真实数字的数量)×100%
- [0208] 表3示出图10中所示的方法的性能,表4示出比较方法的性能。
- [0209] 表3
- [0210]

样本		检测率	错误检测率	速度
XGA 样本 (1024*768) 数量: 模糊:100 正常:500	模糊: 数字的尺寸: (20,30)~(50,75) 倾斜角 <30 度	0.913273	0.485216	平均 295ms
	正常: 数字的尺寸: (20,30)~(50,75) 倾斜角 <30 度	0.950237	0.544347	平均 314ms
在 XGA 上: 倾斜角 <30 度 样本尺寸: (2500*2000)~(4000*3000) 数量: 65		0.926027	0.350685	平均 423ms

[0211] 表4

[0212]

样本		检测率	错误检测率	速度
XGA 样本 (1024*768) 数量: 模糊:100 正常:500	模糊: 数字的尺寸: (20,30)~(50,75) 倾斜角<30 度	0.923567	1.065817	平均 1232ms
	正常: 数字的尺寸: (20,30)~(50,75) 倾斜角 <30 度	0.968517	1.190589	平均 1341ms
在 XGA 上: 倾斜角 <30 度 样本的尺寸: (2500*2000)~(4000*3000) 数量: 65		0.930108	0.556452	平均 1653ms

[0213] 从以上的表3可见,相对于表4中所示的比较方法,根据本发明的例子可获得更好的性能,尤其是在输入图像被模糊时。

[0214] 此外,可从以上的表3看出,相对于表4中所示的比较方法,根据本发明的例子可显著减少字符检测所花费的时间,因此能够以高得多的速度执行字符检测。

[0215] 可以通过许多方式来实施本发明的方法和设备。例如,可以通过软件、硬件、固件、或其任何组合来实施本发明的方法和设备。上述的方法步骤的次序仅是说明性的,本发明的方法步骤不限于以上具体描述的次序,除非以其他方式明确说明。此外,在一些实施例中,本发明还可以被实施为记录在记录介质中的程序,其包括用于实现根据本发明的方法的机器可读指令。因而,本发明还覆盖存储用于实现根据本发明的方法的程序的记录介质。

[0216] 虽然已通过示例详细展示了本发明的一些具体实施例,但是本领域技术人员应当理解,上述示例仅意图是说明性的而不限制本发明的范围。本领域技术人员应该理解,上述实施例可以被修改而不脱离本发明的范围和实质。本发明的范围是通过所附的权利要求限定的。

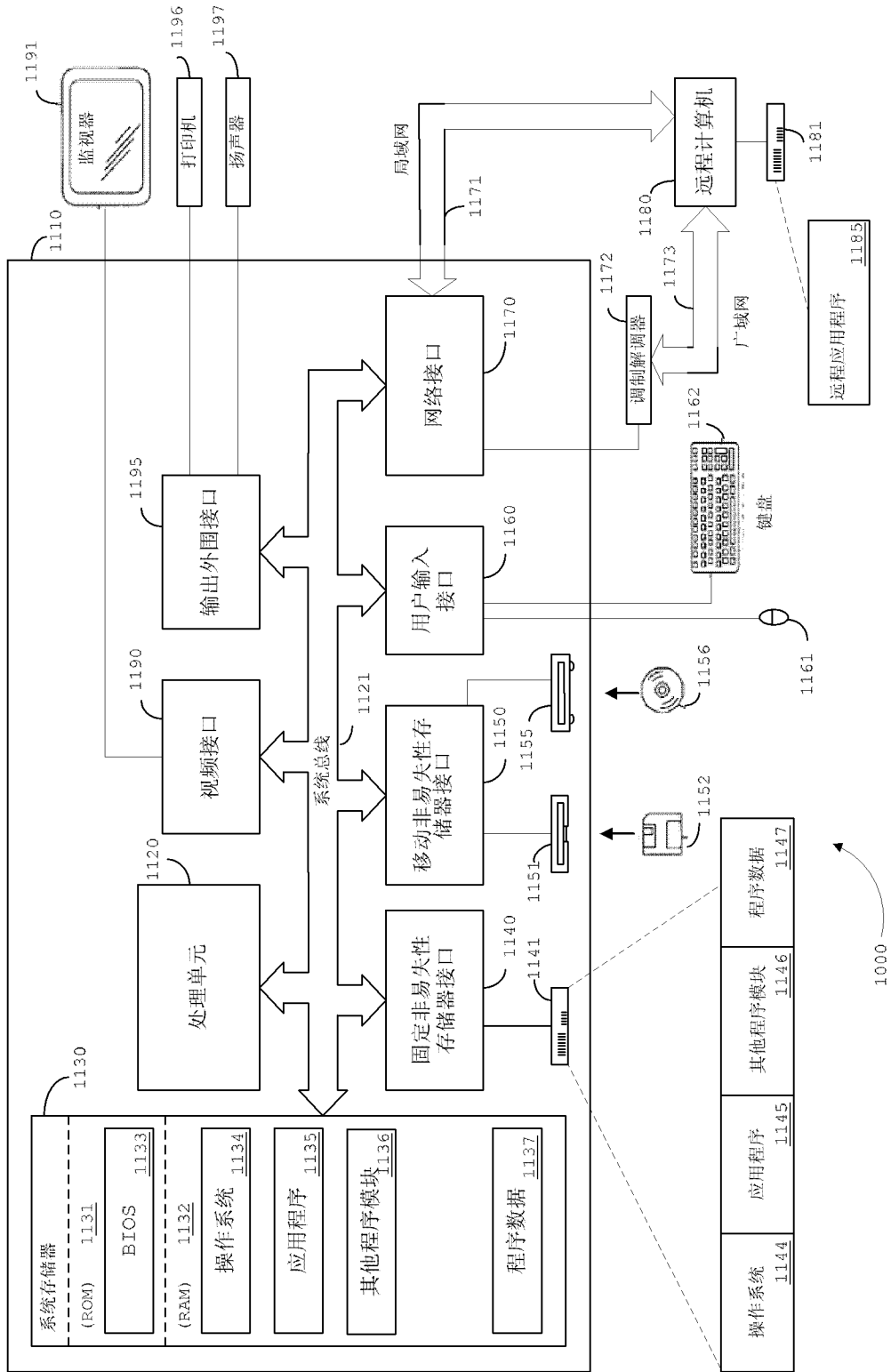


图1



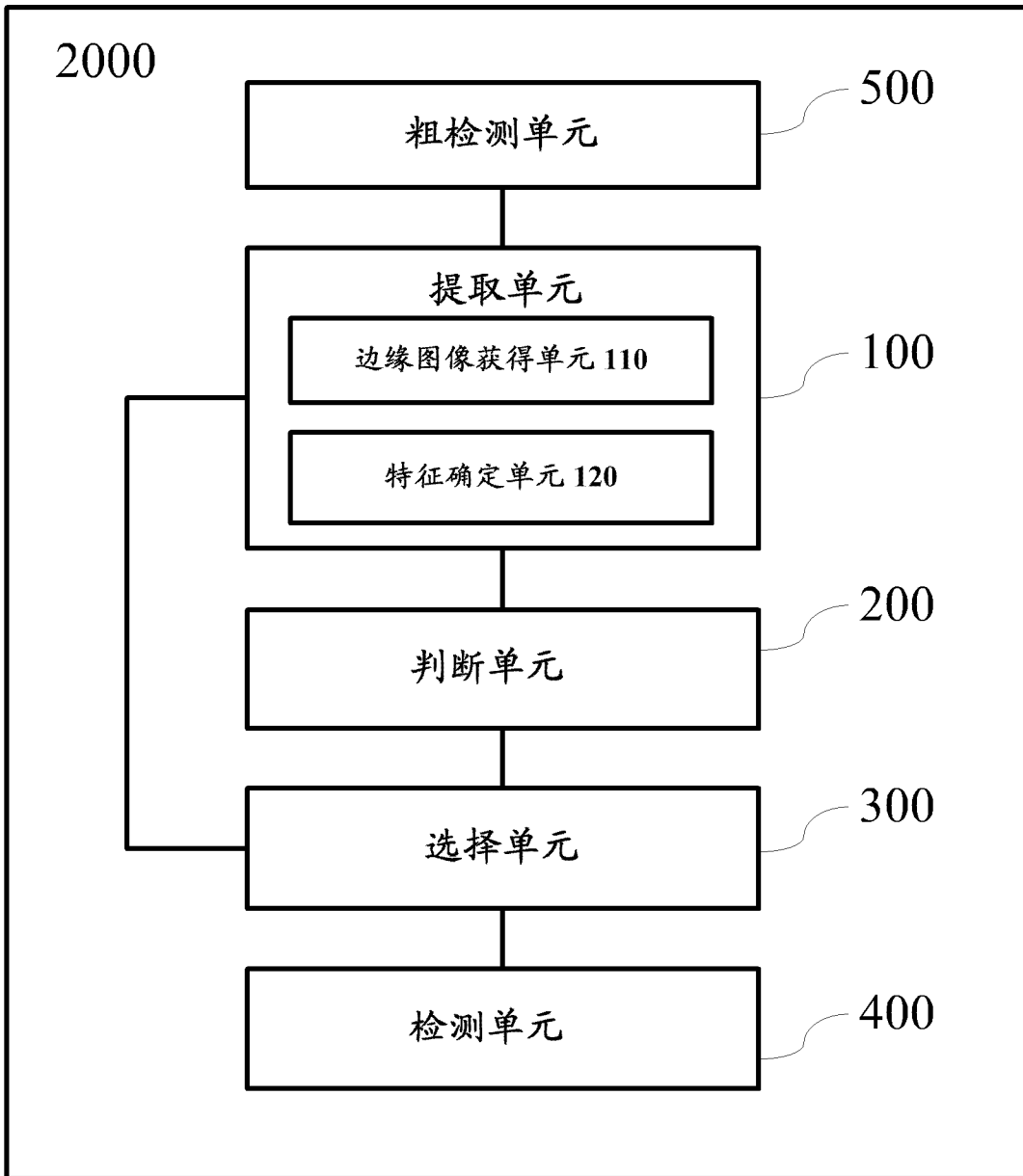


图2

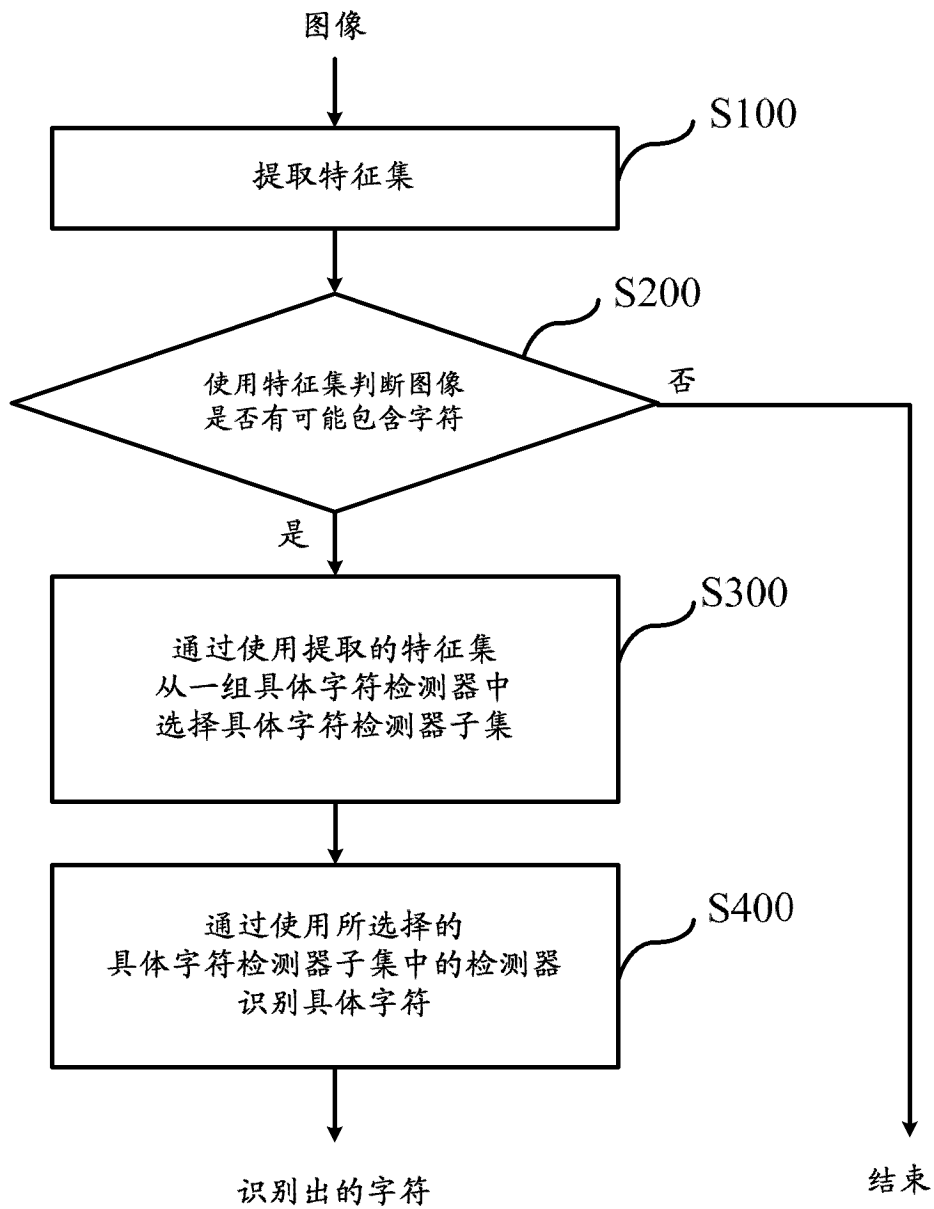


图3

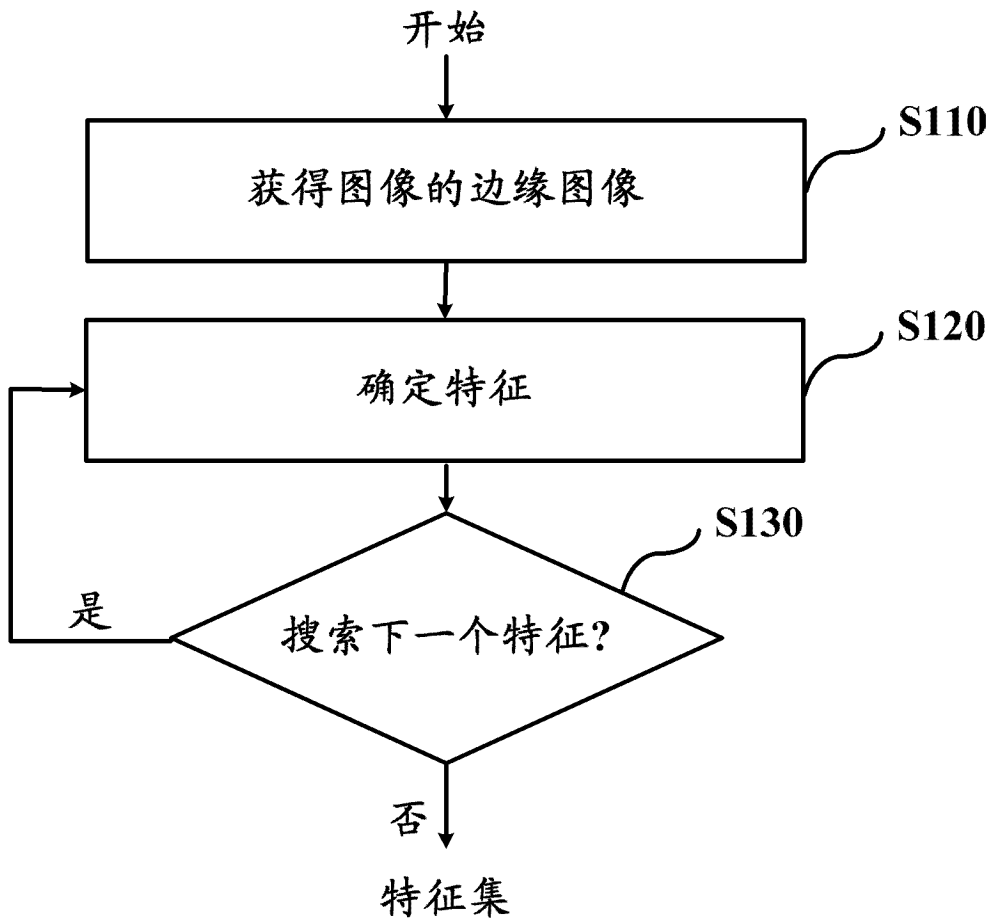


图4

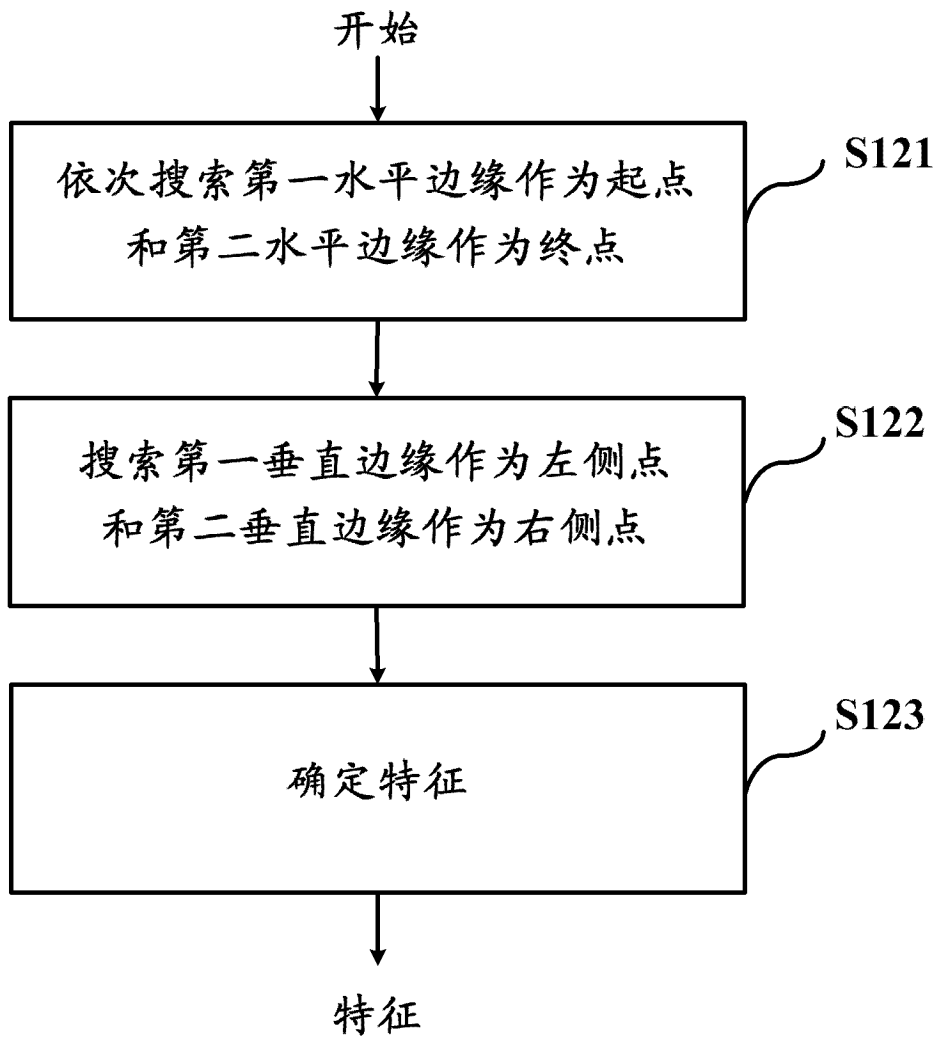


图5

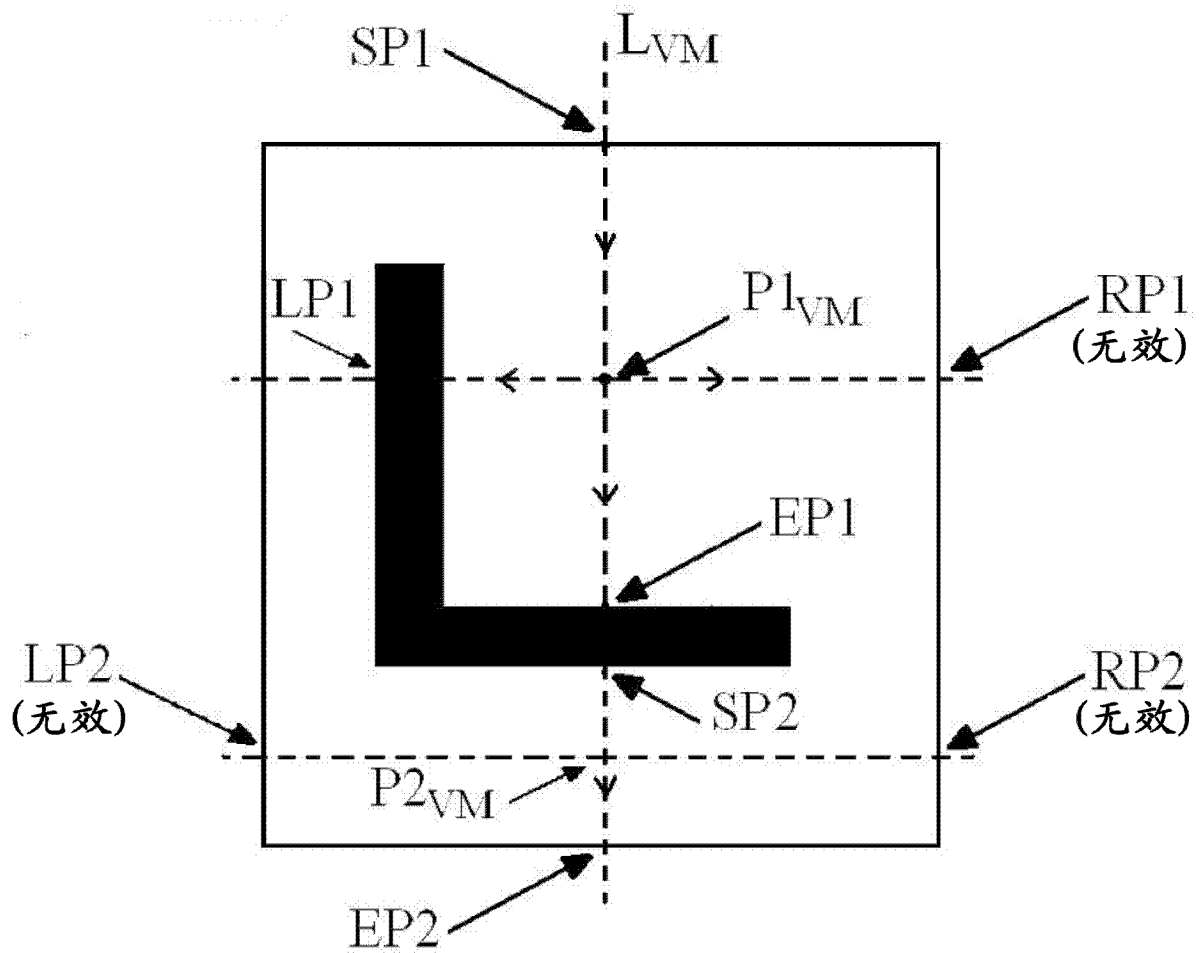


图6

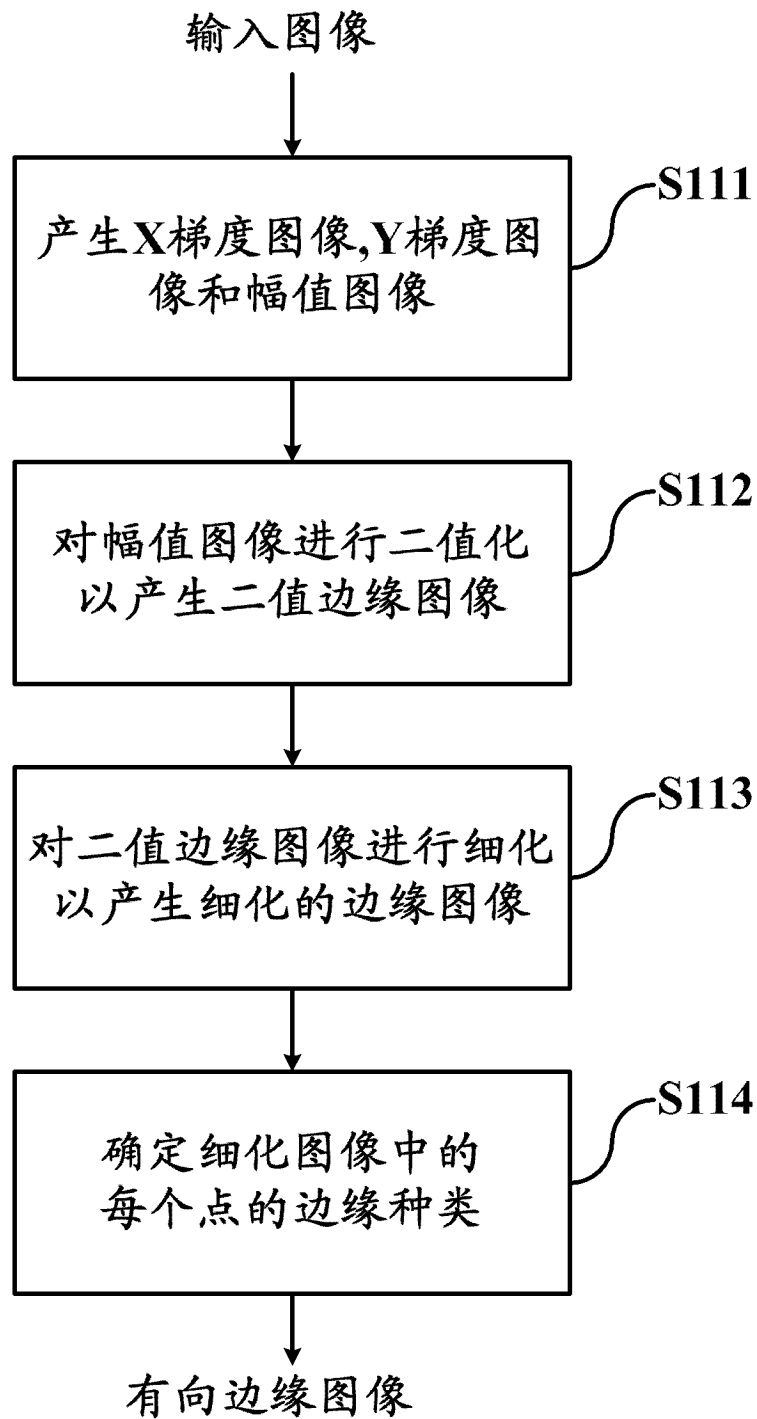


图7

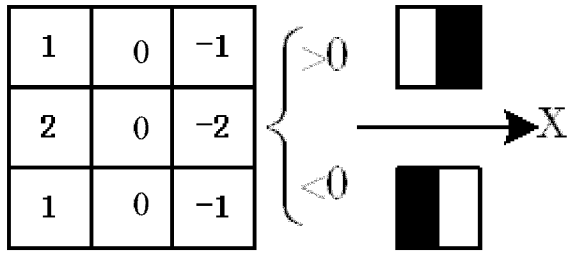


图8a

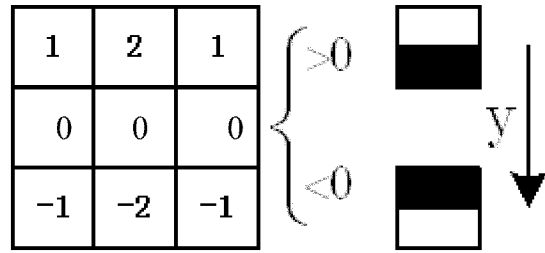


图8b

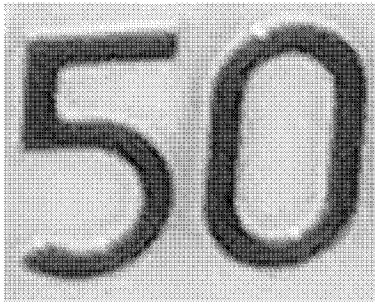


图9a

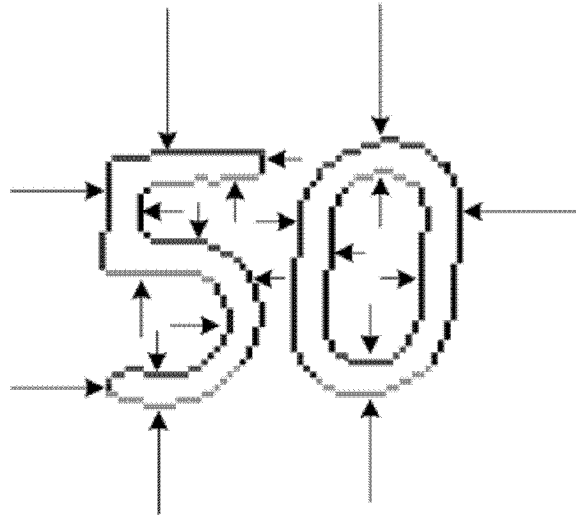


图9b

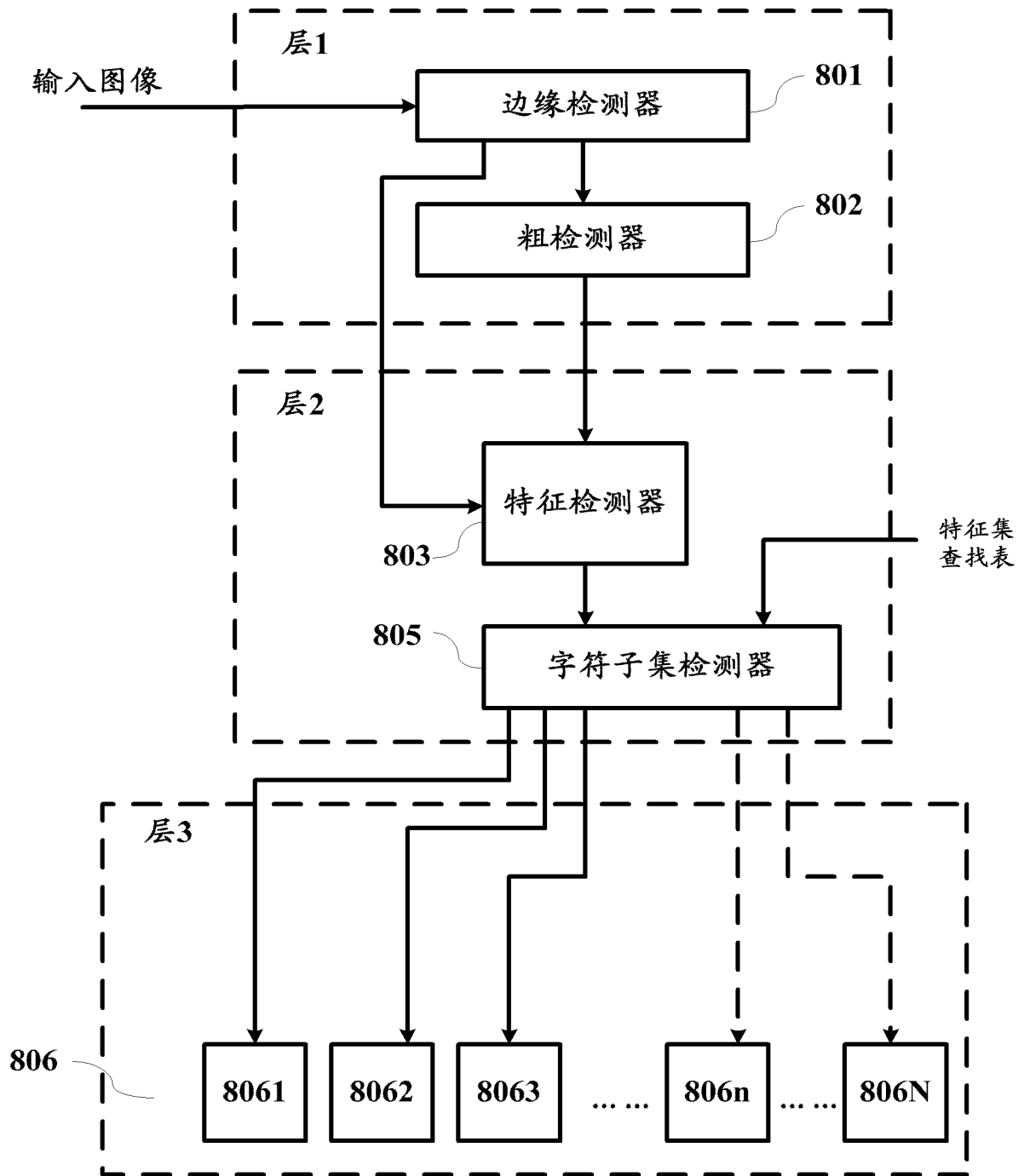


图10



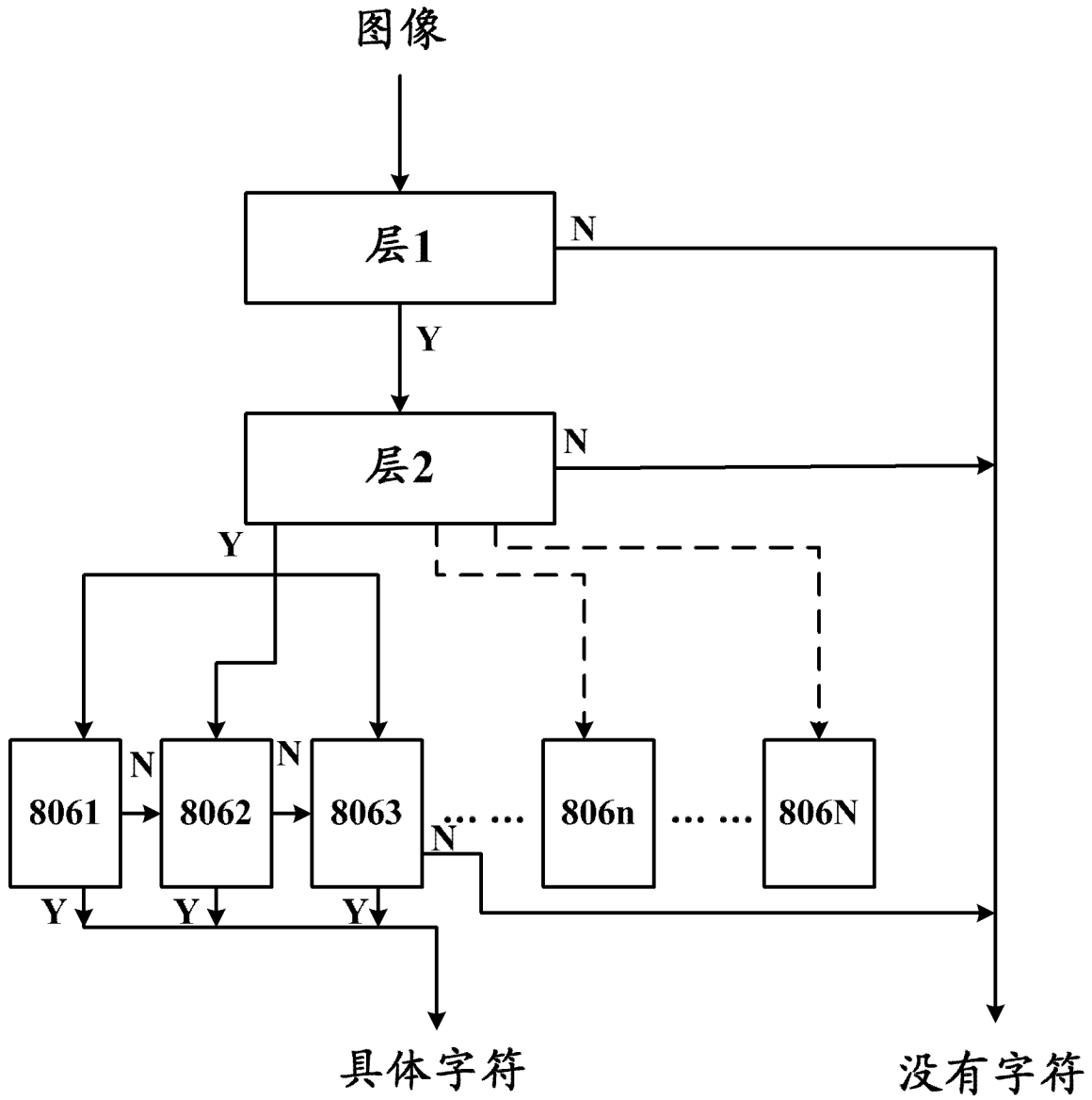


图11

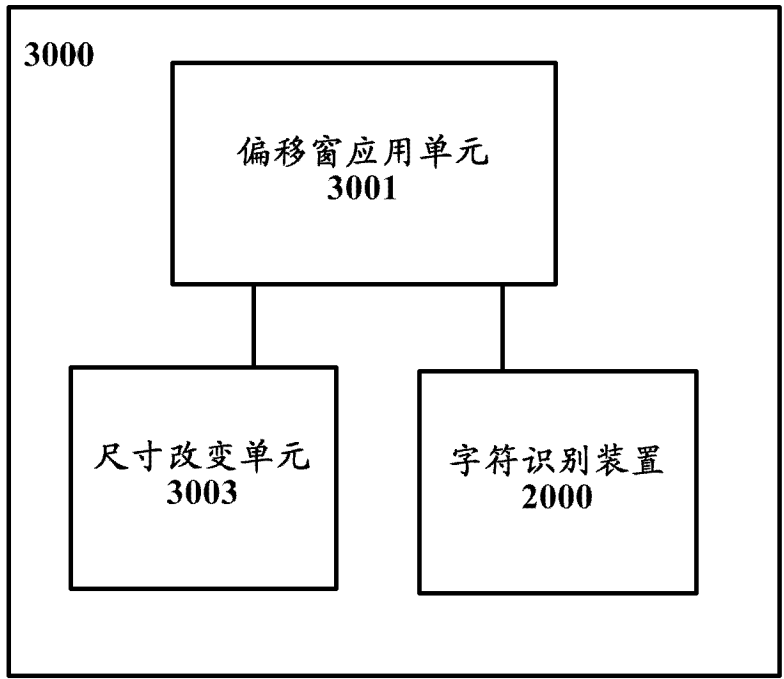


图12

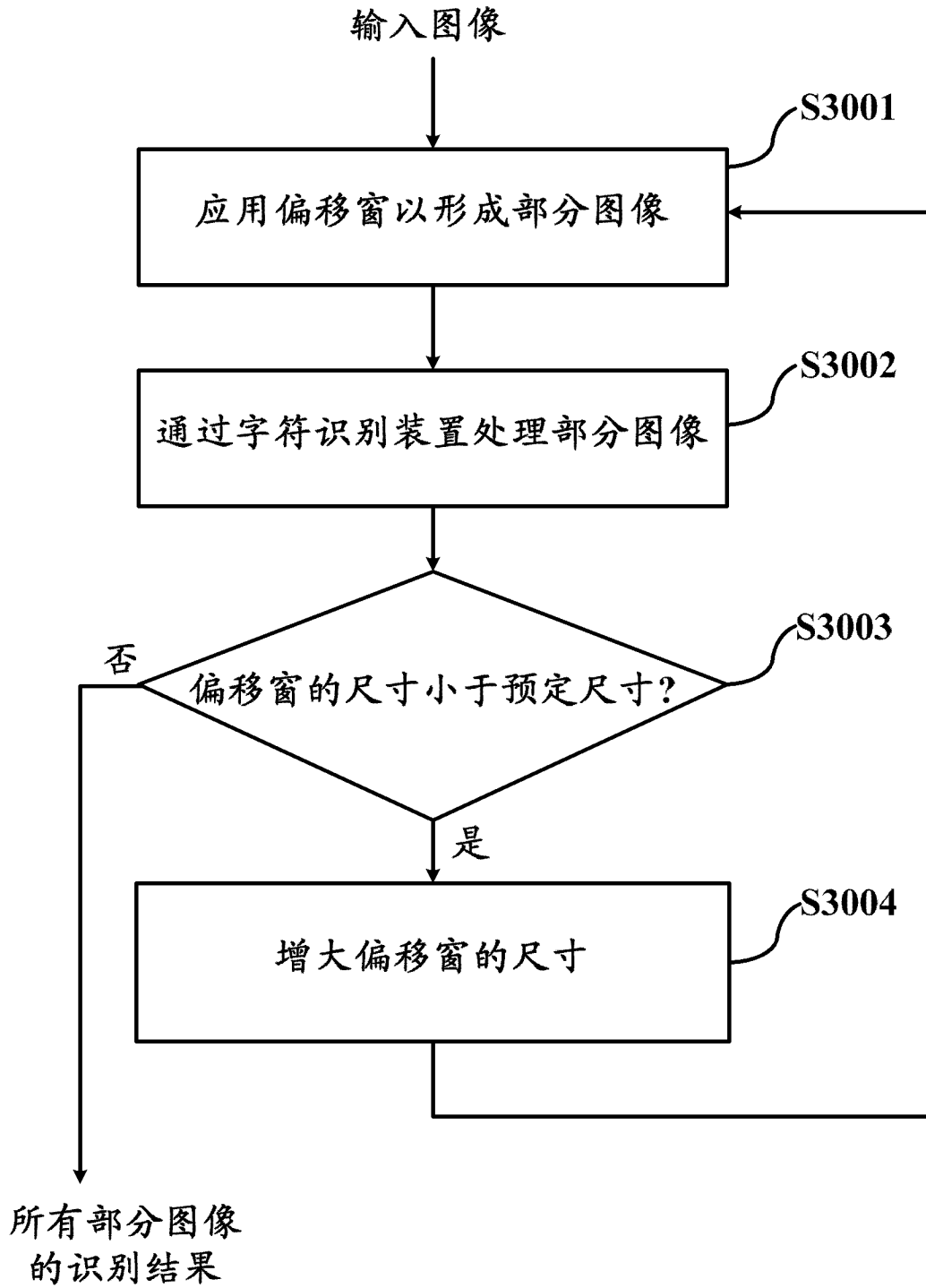


图13

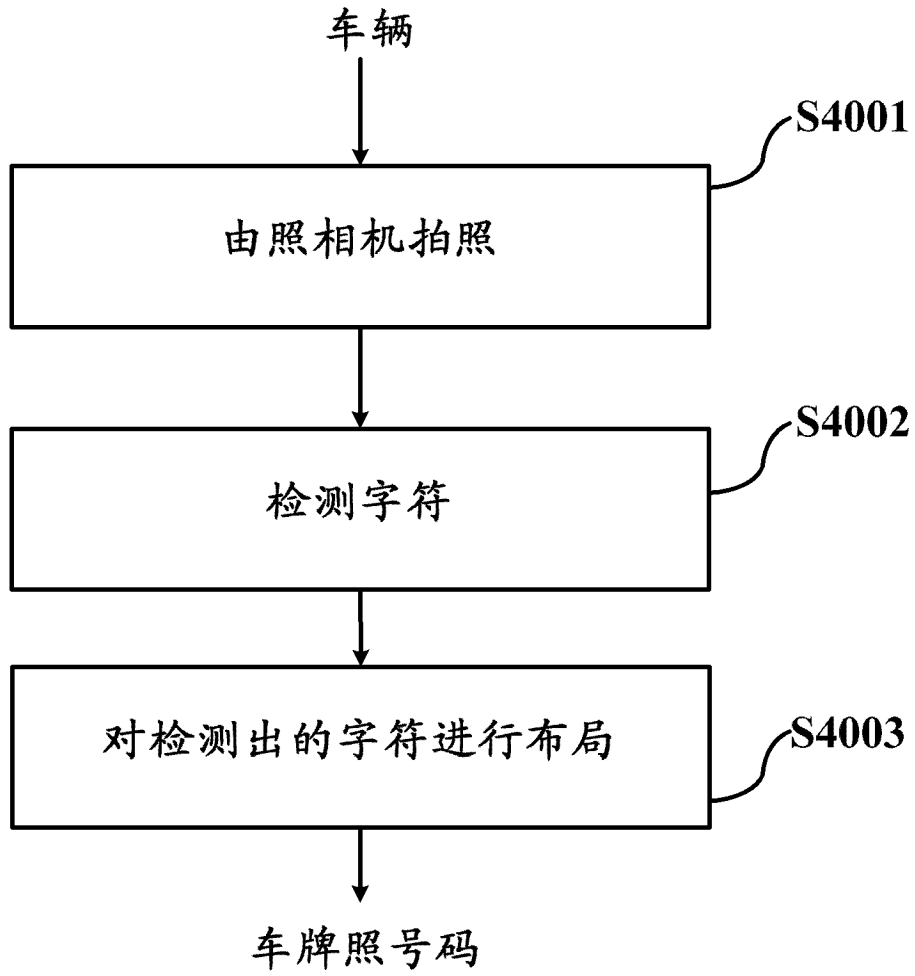


图14



图15a



图15b



**M JA 7229**

图15c