



(12) 发明专利

(10) 授权公告号 CN 113506297 B

(45) 授权公告日 2021.12.03

(21) 申请号 202111063256.3

G06T 7/13 (2017.01)

(22) 申请日 2021.09.10

G06T 5/00 (2006.01)

(65) 同一申请的已公布的文献号

G06N 3/04 (2006.01)

申请公布号 CN 113506297 A

G06N 3/08 (2006.01)

(43) 申请公布日 2021.10.15

审查员 丁娇

(73) 专利权人 南通天成包装有限公司

地址 226000 江苏省南通市通州区西亭镇
西亭居委会31组

(72) 发明人 葛峰

(74) 专利代理机构 郑州知倍通知识产权代理事
务所(普通合伙) 41191

代理人 夏开松

(51) Int. Cl.

G06T 7/00 (2017.01)

G06T 7/12 (2017.01)

权利要求书3页 说明书6页 附图1页

(54) 发明名称

基于大数据处理的印刷数据识别方法

(57) 摘要

本发明公开了一种基于大数据处理的印刷数据识别方法,方法包括以下步骤,步骤一:使用语义分割技术将采集到的RGB图像分割得到印刷品图像;步骤二:对标准图像和印刷图像进行处理,得到各自的图像描述;步骤三:对比标准图像与印刷图像的图像描述,判断印刷异常的情况。本发明与现有技术相比,其有益效果为:本发明先通过边缘检测提取图案内的边缘,再利用各个连通域描述的变化差异来进行缺陷检测,避免了光照的干扰,提高了结果的可靠性,通过使用图像内容的连通域描述来判断异常,而非对应像素的响应值。



1. 一种基于大数据处理的印刷数据识别方法,该方法包括以下步骤,步骤一:使用语义分割技术将采集到的RGB图像分割得到印刷品图像;步骤二:对标准图像和印刷图像进行处理,得到各自的图像描述;包括:对印刷品图像进行边缘检测,获得对应的边缘图像;提取边缘图像内闭合的连通域;设置初始参数,印刷图像的各连通域 P_i 分别设置对应的初始值,包括像素个数 $N=0$,层数 $L=0$;极限坐标: $x_{max}=0, x_{min}=0, y_{max}=0, y_{min}=0$;

由于不同图像的像素大小不同,故而连通域的面积参数用该连通域的像素个数与整个图像的总像素个数的比值来表示,即该连通域面积参数 S :

$$S = \frac{N}{N_{总}};$$

式中, N 为连通域的像素个数, $N_{总}$ 为整个图像的像素个数;

逐行遍历图像的像素点,获得:图像中每行的像素点都有其对应的连通域标签序列,形式如下:

[00100302000203000010004050050600604007000070000];其中0为背景像素即非连通域像素,0之外的数字为对应连通域的标签号,对标签值进行处理,获得该行下连通域的层次信息,由于一个连通域是一个闭合区域,所以在图像的一行像素上从左到右遍历时,它的标签号要至少出现两次,第一次为开始进入这个连通域,第二次为离开这个连通域,存在嵌套结构的连通域,必然是大连通域包含小连通域,所以连通域的嵌套层数一旦确定是不会改变的,如果 $L=0$,表示对应连通域的嵌套层数未确定,当 $L \neq 0$ 时,不需要更改 L 的值,设置一个临时变量 $C=0$,从左向右进行遍历,记录第一个非0数字,上述连通域标签序列中对应的数字为1,其对应的 C 设置为1,此由于标签号为1的连通域的嵌套层数 $L=0$,所以对对应的连通域的嵌套层数 L 的值更新为 C ,表示该连通域的最大嵌套层数为1,此时记录下的非0标签序列为{1};第二个非0数字,上述连通域标签序列中对应的数字为3,已记录的标签序列中不存在数字3,记录该数字,所得标签序列为{1, 3},此时令 $C = C + 1 = 2$,表示进入更深一层的嵌套连通区域中,由于标签号为3的连通域最大嵌套层次 $L=0$,更新 $L = C = 2$,表示标签号为3的连通域的最大嵌套层数为2;第三个非0数字为2,不在已记录的非0标签序列内,将数字2记录进标签序列,此时的非0数字标签序列为{1, 3, 2},由于标签号2的连通域的嵌套层数 $L=0$,令 $L = C = 3$,表示标签号为2的连通域的最大嵌套层数为3;继续遍历,第四个遍历的非0数字为2,由于之前记录的数字序列{1, 3, 2}中已经存在2,表示标签为2的连通域遍历已经结束,不再将其记录进标签序列;令 C 减1,此时 $C=2$,即此时遍历的像素点位于嵌套层次为2的连通域中,以此类推,每遇到一个新的连通域,即记录的非0数字标签序列中不存在的标签号,使 C 加1,表示进入到更深一层的嵌套区域中;每离开一个连通域,令 C 减1,表示回到上一层的嵌套区域中;另外,在每次对 C 进行加1操作时,需要判断对应标签号的嵌套层次 L 是否为0,直到遍历完该行像素,每遍历一个带标签的像素点,就使对应标签的像素个数 $N+1$;比较该带标签的像素点坐标与已有的横纵坐标的最大值和最小值: $x_{max}, x_{min}, y_{max}, y_{min}$ 进行比较,包括:某遍历的标签像素点坐标为 (x_0, y_0) ,若 $x_0 < x_{min}$,则 $x_{min} = x_0$,否则 x_{min} 保持不变;若 $x_0 > x_{max}$,则 $x_{max} = x_0$,否则 x_{max} 保持不

变;对 y_{min}, y_{max} 同理;遍历结束后,得到中心点坐标 $Z(x_z, y_z)$,其中

$x_z = \frac{x_{min} + x_{max}}{2}, y_z = \frac{y_{min} + y_{max}}{2}$;得到 $S = \frac{N}{N_{\#}}$,以及各自的嵌套层数 L ;计算各个连通域整体特征值 M ,用连通域中心点坐标到原点距离以及连通域面积表示:

$$M = 0.5 \times \sqrt{x_z^2 + y_z^2} + 0.5 \times S;$$

得到每个连通域 P_i 的描述情况 (M, S, L, T) ;

步骤三:对比标准图像与印刷图像的图像描述,判断印刷异常的情况。

2. 根据权利要求1所述的一种基于大数据处理的印刷数据识别方法,其特征就在于对印刷品图像进行边缘检测,获得对应的边缘图像具体包括:输入印刷图像,将图像灰度化处理后,使用Canny算子对图像进行边缘检测,得到梯度边缘,即为印刷区域中图案的边缘。

3. 根据权利要求1所述的一种基于大数据处理的印刷数据识别方法,其特征就在于,提取边缘图像内闭合的连通域具体包括:对印刷区域中图案的边缘使用种子填充法进行连通域分析,得到具有不同标签的连通域 P_i ,并得到最大的标签号的值 M ,即总共的数量。

4. 根据权利要求1所述的一种基于大数据处理的印刷数据识别方法,其特征就在于,连通域描述组合得到该印刷品图像的描述具体为:整个图像的描述表现为

$$Img_{\#} = \{P_1, P_2, P_3, \dots, P_m\};$$

同理,对标准图像重复上述操作,得到标准图像的图像描述数据

$$Img_{\#} = \{Q_1, Q_2, Q_3, \dots, Q_n\};$$

其中 Q_i 为标准图像中的连通域, n 为连通域个数。

5. 根据权利要求1所述的一种基于大数据处理的印刷数据识别方法,其特征就在于,步骤三具体为:遍历标准图像和印刷图像的连通域,将其按照层数 L 进行分组,并分别计算各个层级对应的个数;比较标准图像和印刷图像各个层级的对应个数,个数一致为正常情况,个数不一致的层级需要对该层级内的连通域进一步检测:设置标准图像中的对应连通域标志 $A = 0$,整体特征值差距 $B = 0$;印刷图像中的对应连通域标志 $E = 0$;寻找该层级各连通域之间对应关系:查找印刷图像某个连通域与标准图像中各个连通域的整体特征值 M 之间的差值:

$$D = |M_p - M_{Q_i}|$$

差值最小,即 D_{min} 时,对相关的两个连通域即为印刷图像对应标签号 $T_{p'}$ 的对应连通域 P' 与标准图像对应标签号 T_Q 的对应的连通域;

若标准图像连通域标志 $A = 0$,表示此时的标准图像中的连通域无对应连通域 P ,更新 A, B, E 的值:

$$A = T_{p'}, B = D, E_{p'} = T_Q;$$

若标准图像连通域标志 $A \neq 0, (A = T_p)$,表示标准图像中的该连通域当前在印刷图像中已存在一个对应连通域 P ,比较该标准图像连通域的 B 与对应连通域 P' 对应的差值 D 的大小:

若 $B > D$,更新 A, B, E 的值,令

$$A = T_{P'}, B = D, E_{P'} = T_Q, E_P = 0;$$

其中, $E_{P'}$ 为对应连通域 P' 的连通域标志, T_Q 为标签号, E_P 为对应连通域 P 的连通域标志;

若 $B = D$, 获得该印刷图像中的对应连通域 P 与标准图像连通域的面积差距; 获得对应连通域 P' 与标准图像连通域的面积差距; 如果对应连通域 P' 对应的面积差距相对于对应连通域 P 的面积差距小, 则更新 A, B, E 的值; 否则不更新 A, B, E 的值;

若 $B < D$, 则不更新 A, B, E 的值;

遍历结束后, 印刷图像中连通域的对应连通域标志 E 存在两种情况: $E = 0, E \neq 0$;

$E = 0$: 表示该连通域在标准图像中不存在对应连通域, 即缺陷区域, 这种缺陷存在的位置为图案内部和图案外部空白区域, 该区域缺陷的整体影响情况 Y_0 为:

$$Y_0 = \sum \frac{1}{L} \times S$$

$E \neq 0$: 在标准图像中存在对应连通域, 比较二者之间的面积差距, 由于存在统计误差, 误差在标准图像总面积的3%以内均为正常, 超过3%的认为存在异常, 存在图案缺印或多印的状况, 则这些区域的缺陷整体影响情况为 Y_1 , 表示如下:

$$Y_1 = \sum |S_P - S_Q|$$

印刷图像整体存在的缺陷 Y 为:

$$Y = Y_0 + Y_1。$$

基于大数据处理的印刷数据识别方法

技术领域

[0001] 本发明涉及大数据处理领域,尤其涉及一种基于大数据处理的印刷数据识别方法。

背景技术

[0002] 现有的对印刷品缺陷检测的方法除主观目测方法之外,往往是将待检测的印刷品数据与标准模板数据进行对比分析,比如色度检测法中将一束光投射到印刷品上,通过仪器获取颜色的三刺激值,换算成可以对比的数值,然后与样本的值比对,出现异常的地方即为存在缺陷的位置;以及基于图像的将拍摄图像与标准图像做差对比,但这些方法容易受到光照等因素的干扰,最终都会导致结果不准确。因为很多时候标准模板都只是电子版文件,不受环境的干扰,而印刷出的产品其图像是通过相机真实采集获得的,而真实场景中的光源强度、方向等都是随机的,所以直接比较时,很容易受到光照影响,即难以确定差异的地方是真实的印刷缺陷,还是光照的差异。

发明内容

[0003] 本发明的目的是为了解决现有技术中存在的缺点,而提出的一种基于大数据处理的印刷数据识别方法。

[0004] 为了实现上述目的,本发明采用了如下技术方案:

[0005] 一种基于大数据处理的印刷数据识别方法,该方法包括以下步骤,步骤一:使用语义分割技术将采集到的RGB图像分割得到印刷品图像;步骤二:对标准图像和印刷图像进行处理,得到各自的图像描述;步骤三:对比标准图像与印刷图像的图像描述,判断印刷异常的情况。

[0006] 进一步,所述步骤二具体为:对印刷品图像进行边缘检测,获得对应的边缘图像;提取边缘图像内闭合的连通域;由连通域分析来获得其描述;连通域描述组合得到该印刷品图像的描述。

[0007] 进一步,对印刷品图像进行边缘检测,获得对应的边缘图像具体包括:输入印刷图像,将图像灰度化处理后,使用Canny算子对图像进行边缘检测,得到梯度边缘,即为印刷区域中图案的边缘。

[0008] 进一步,提取边缘图像内闭合的连通域具体包括:对印刷区域中图案的边缘使用种子填充法进行连通域分析,得到具有不同标签的连通域 P_i ,并得到最大的标签号的值 m ,即总共的数量。

[0009] 进一步,由连通域分析来获得其描述具体为:设置初始参数,印刷图像的各连通域 P_i (共 m 个)分别设置对应的初始值,包括像素个数 $N=0$,层数 $L=0$ 。极限坐标: $x_{max}=0, x_{min}=0, y_{max}=0, y_{min}=0$;

[0010] 由于不同图像的像素大小不同,故而连通域的面积参数用该连通域的像素个数与整个图像的总像素个数的比值来表示,即该连通域面积参数 S :

$$[0011] \quad S = \frac{N}{N_{\#}};$$

[0012] 式中, N 为连通域的像素个数, $N_{\#}$ 为整个图像的像素个数;

[0013] 逐行遍历图像的像素点, 获得: 图像中每行的像素点都有其对应的连通域标签, 形式如下:

[0014] [00100302000203000010004050050600604007000070000]; 其中0为背景像素即非连通域像素, 0之外的数字为对应连通域的标签号, 对标签值进行处理, 获得该行下连通域的层次信息, 由于一个连通域是一个闭合区域, 所以在图像的一行像素上从左到右遍历时, 它的标签号要至少出现两次, 第一次为开始进入这个连通域, 第二次为离开这个连通域, 存在嵌套结构的连通域, 必然是大连通域包含小连通域, 所以连通域的嵌套层数一旦确定是不会改变的, 如果 $L = 0$, 表示对应连通域的嵌套层数未确定, 当 $L \neq 0$ 时, 不需要更改 L 的值, 设置一个临时变量 $C=0$, 从左向右进行遍历, 记录第一个非0数字, 上述序列中对应的数字为1, 其对应的 C 设置为1, 此由于标签号为1的连通域的嵌套层数 $L = 0$, 所以对应的连通域的嵌套层数 L 的值更新为 C , 表示该连通域的最大嵌套层数为1, 此时记录下的非0标签序列为{1}; 第二个非0数字, 上述序列中对应的数字为3, 已记录的标签序列中不存在数字3, 记录该数字, 所得标签序列为{1, 3}, 此时令 $C = C + 1 = 2$, 表示进入更深一层的嵌套连通域中, 由于标签号为3的连通域最大嵌套层次 $L = 0$, 更新 $L = C = 2$, 表示标签号为3的连通域的最大嵌套层数为2; 第三个非0数字为2, 不在已记录的非0标签序列内, 将数字2记录进标签序列, 此时的非0数字标签序列为{1, 3, 2}, 由于标签号2的连通域的嵌套层数 $L = 0$, 令 $L = C = 3$, 表示标签号为2的连通域的最大嵌套层数为3; 继续遍历, 第四个遍历的非0数字为2, 由于之前记录的数字序列{1, 3, 2}中已经存在2, 表示标签为2的连通域遍历已经结束, 不再将其记录进标签序列。令 C 减1, 此时 $C=2$, 即此时遍历的像素点位于嵌套层次为2的连通域中, 以此类推, 每遇到一个新的连通域, 即记录的非0数字标签序列中不存在的标签号, 使 C 加1, 表示进入到更深一层的嵌套区域中; 每离开一个连通域, 令 C 减1, 表示回到上一层的嵌套区域中; 另外, 在每次对 C 进行加1操作时, 需要判断对应标签号的嵌套层次 L 是否为0, 直到遍历完该行像素, 每遍历一个带标签的像素点, 就使对应标签的像素个数 $N + 1$; 比较带标签的像素点坐标 (x_0, y_0) 与已有的 $x_{max}, x_{min}, y_{max}, y_{min}$, 若 $x_0 < x_{min}$, 则 $x_{min} = x_0$, 否则 x_{min} 保持不变; 若 $x_0 > x_{max}$, 则 $x_{max} = x_0$, 否则 x_{max} 保持不变; 对 y_{min}, y_{max} 同理; 遍历结束后, 得到中心点坐标 $Z(x_z, y_z)$, 其中

$$x_z = \frac{x_{min} + x_{max}}{2}, y_z = \frac{y_{min} + y_{max}}{2}; \text{得到 } S = \frac{N}{N_{\#}}, \text{ 以及各自的嵌套层数 } L; \text{ 计算各个连通域整体}$$

特征值 M , 用连通域中心点坐标到原点距离以及连通域面积表示:

$$[0015] \quad M = 0.5 \times \sqrt{x_z^2 + y_z^2} + 0.5 \times S;$$

[0016] 得到每个连通域 P_i 的描述情况 (M, S, L, T) 。

[0017] 进一步, 连通域描述组合得到该印刷品图像的描述具体为: 整个图像的描述表现为

$$[0018] \quad \text{Img}_{\text{印}} = \{P_1, P_2, P_3, \dots, P_m\};$$

[0019] 同理, 对标准图像重复上述操作, 得到标准图像的图像描述数据

$$[0020] \quad \text{Img}_{\text{标}} = \{Q_1, Q_2, Q_3, \dots, Q_n\};$$

[0021] 其中 Q_i 为标准图像中的连通域, n 为连通域个数。

[0022] 进一步,步骤三具体为:遍历标准图像和印刷图像的连通域,将其按照层数 C 进行分组,并分别计算各个层级对应的个数;比较标准图像和印刷图像各个层级的对应个数,个数一致为正常情况,个数不一致的层级需要对该层级内的连通域进一步检测:设置标准图像中的对应连通域标志 $A = 0$,整体特征值差距 $B = 0$;印刷图像中的对应连通域标志 $E = 0$;寻找该层级各连通域之间对应关系:查找印刷图像某个连通域与标准图像中各个连通域的整体特征值 M 之间的差值:

$$[0023] \quad D = |M_p - M_{Q_i}|$$

[0024] 差值最小,即 D_{min} 时,对相关的两个连通域即为印刷图像对应标签号 $T_{p'}$ 的对应连通域 P' 与标准图像对应标签号 T_Q 的对应的连通域;

[0025] 若标准图像连通域标志 $A = 0$,表示此时的标准图像中的连通域无对应连通域,记 $A = T_{p'}, B = D, E_{p'} = T_Q$;

[0026] 若标准图像连通域标志 $A \neq 0, (A = T_p)$,表示标准图像中的该连通域当前在印刷图像中已存在一个对应连通域 P ,比较该标准图像连通域的 B 与对应连通域 P' 对应的差值 D 的大小:

[0027] 若 $B > D$,更新 A, B, E 的值,令

$$[0028] \quad A = T_{p'}, B = D, E_{p'} = T_Q, E_p = 0;$$

[0029] 若 $B = D$,获得该印刷图像中的对应连通域 P 与标准图像连通域的面积差距;获得对应连通域 P' 与标准图像连通域的面积差距,如果对应连通域 P' 对应的面积差距较小,则更新 A, B, E 的值;否则不更新 A, B, E 的值;

[0030] 若 $B < D$,则不更新 A, B, E 的值;遍历结束后,印刷图像中连通域的对应连通域标志 E 存在两种情况: $E = 0, E \neq 0$;

[0031] $E = 0$:表示该连通域在标准图像中不存在对应连通域,即缺陷区域,这种缺陷存在的位置为图案内部和图案外部空白区域,前者嵌套层次较多,不影响印刷产品的整体内容,故其缺陷影响程度较小,而后者的后者程度较大,所以该区域缺陷的整体影响情况 Y_0 为:

$$[0032] \quad Y_0 = \sum \frac{1}{L} \times S$$

[0033] $E \neq 0$:在标准图像中存在对应连通域,比较二者之间的面积差距,由于存在统计误差,误差在标准图像总面积的3%以内均为正常,超过3%的认为存在异常,存在图案缺印或多印的状况,则这些区域的缺陷整体影响情况为 Y_1 ,表示如下:

$$[0034] \quad Y_1 = \sum |S_p - S_{Q_i}|$$

[0035] 综上所述,印刷图像整体存在的缺陷 Y 为:

$$[0036] \quad Y = Y_0 + Y_1。$$

[0037] 本发明与现有技术相比,其有益效果为:本发明先通过边缘检测提取图案内的边缘,再利用各个连通域描述的变化差异来进行缺陷检测,避免了光照的干扰,提高了结果的可靠性,通过使用图像内容的连通域描述来判断异常,而非对应像素的响应值。

附图说明

- [0038] 图1为系统流程图；
 [0039] 图2为连通域标签形式图；
 [0040] 图3为嵌套示意图；

具体实施方式

[0041] 下面将结合本发明实施例中的附图，对本发明实施例中的技术方案进行清楚、完整地描述。

[0042] 参见图1。本发明的主要目的是检测印刷品存在缺陷；主要针对连通域比较少数的海报等印刷品。

[0043] 步骤一：使用语义分割技术，提取采集到的RGB图像中的印刷图像；首先需要使用DNN来识别采集图像中的印刷品，即采集图像中有背景等复杂工况，同时也有待判断印刷品。通过语义分割技术分理出待检测图像：输入相机采集到的印刷纸张的图像，使用DNN网络对其进行语义分割，网络结构为Encoder-Decoder结构，数据集为各种类型的印刷纸张图像；标签分为两类，印刷产品和背景。该方式为像素级分类，即需要给图像中所有像素标注上对应的标签。属于印刷纸张的像素，其值标注为1，属于背景的像素，其值标注为0；网络所用的loss函数为交叉熵损失函数。得到印刷品连通域后进行下面操作：以语义分割得到的结果为遮罩，从RGB图像中提取出对应的印刷品图像。根据遮罩的长轴和短轴夹角，进行图像的旋转。得到转正后的印刷品图像。该部分的遮罩分割，旋转为现有惯用技术。

[0044] 至此，步骤一完成，能够实现从采集的RGB图像中分割出印刷品图像。

[0045] 步骤二：对标准图像和印刷图像进行连通域分析，得到各自的图像描述，光照的变化很容易带来图像内像素值的变化，所以为了避免光照的影响，本发明需要不使用灰度值来得到图像的描述。本发明获得印刷品描述的过程为：对印刷品图像进行边缘检测，获得对应的边缘图像。提取边缘图像内闭合的连通域。由连通域分析来获得其描述。连通域描述组合得到该印刷品图像的描述。

[0046] 对印刷品图像进行边缘检测，获得对应的边缘图像：输入印刷图像，将图像灰度化处理后，使用Canny算子对图像进行边缘检测，得到梯度边缘，即为印刷区域中图案的边缘；

[0047] 提取边缘图像内闭合的连通域：对上一步所得结果使用Seed Filling(种子填充法)进行连通域分析，得到具有不同标签(label)的连通域 P_i ，并得到最大的标签号的值 m (即总共的数量)，连通域标签形式如图2所示。原理：<https://blog.csdn.net/liangchunjiang/article/details/79431339>。

[0048] 由连通域分析来获得其描述：设置初始参数，首先，印刷图像的各连通域 P_i (共 m 个)分别设置对应的初始值，包括像素个数 $N=0$ ，层数 $L=0$ 。极限坐标： $x_{max}=0, x_{min}=0, y_{max}=0, y_{min}=0$ ；由于不同图像的像素大小不同，故而连通域的面积参数用该连通域的像素个数与整个图像的总像素个数的比值来表示，即该连通域面积参数 S ：

$$[0049] \quad S = \frac{N}{N_{\text{总}}}$$

[0050] 式中， N 为连通域的像素个数， $N_{\text{总}}$ 为整个图像的像素个数。

[0051] 逐行遍历图像的像素点,获得:图像中每行的像素点都有其对应的连通域标签,形式如下:[00100302000203000010004050050600604007000070000]

[0052] 其中0为背景(非连通域)像素,0之外的数字为对应连通域的标签号。对标签值进行处理,获得该行下连通域的层次信息。由于一个连通域是一个闭合区域,所以在图像的一行像素上从左到右遍历时,它的标签号要至少出现两次,第一次为开始进入这个连通域,第二次为离开这个连通域。此外,存在嵌套结构的连通域,必然是大连通域包含小连通域,所以连通域的嵌套层数一旦确定是不会改变的,如果 $L = 0$,表示对应连通域的嵌套层数未确定,当 $L \neq 0$ 时,不需要更改L的值。首先,设置一个临时变量 $C=0$,从左向右进行遍历,记录第一个非0数字,上述序列中对应的数字为1,其对应的C设置为1,此时,由于标签号为1的连通域的嵌套层数 $L = 0$,所以对应的连通域的嵌套层数L的值更新为C,表示该连通域的最大嵌套层数为1,此时记录下的非0标签序列为{1};第二个非0数字,上述序列中对应的数字为3,已记录的标签序列中不存在数字3,记录该数字,所得标签序列为{1, 3},此时令 $C = C + 1 = 2$,表示进入更深一层的嵌套连通区域中,由于标签号为3的连通域最大嵌套层次 $L = 0$,更新 $L = C = 2$,表示标签号为3的连通域的最大嵌套层数为2;同理,第三个非0数字为2,不在已记录的非0标签序列内,将数字2记录进标签序列,此时的非0数字标签序列为{1, 3, 2},由于标签号2的连通域的嵌套层数 $L = 0$,令 $L = C = 3$,表示标签号为2的连通域的最大嵌套层数为3;继续遍历,第四个遍历的非0数字为2,由于之前记录的数字序列{1, 3, 2}中已经存在2,表示标签为2的连通域遍历已经结束,不再将其记录进标签序列。令C减1,此时 $C=2$,即此时遍历的像素点位于嵌套层次为2的连通域中。以此类推,每遇到一个新的连通域,即记录的非0数字标签序列中不存在的标签号,使C加1,表示进入到更深一层的嵌套区域中;每离开一个连通域,令C减1,表示回到上一层的嵌套区域中;另外,在每次对C进行加1操作时,需要判断对应标签号的嵌套层次L是否为0,直到遍历完该行像素。整个流程如下图3所示。

[0053] 每遍历一个带标签的像素点,就使对应标签的像素个数 $N + 1$;比较该带标签的像素点坐标与已有的 $x_{max}, x_{min}, y_{max}, y_{min}$ 进行比较,比如:某遍历的标签像素点坐标为 (x_0, y_0) ,若 $x_0 < x_{min}$,则 $x_{min} = x_0$,否则 x_{min} 保持不变;若 $x_0 > x_{max}$,则 $x_{max} = x_0$,否则 x_{max} 保持不变;对 y_{min}, y_{max} 同理;遍历结束后,得到中心点坐标 $Z(x_z, y_z)$,其中 $x_z = \frac{x_{min} + x_{max}}{2}, y_z = \frac{y_{min} + y_{max}}{2}$;得到 $S = \frac{N}{N_{总}}$,以及各自的嵌套层数L;计算各个连通域整体特征值M,用连通域中心点坐标到原点距离以及连通域面积表示:

$$[0054] \quad M = 0.5 \times \sqrt{x_z^2 + y_z^2} + 0.5 \times S$$

[0055] 得到每个连通域 P_i 的描述情况 (M, S, L, T) ;

[0056] 连通域描述组合得到该印刷品图像的描述:则整个图像的描述表现为

$$[0057] \quad Img_{印} = \{P_1, P_2, P_3, \dots, P_m\};$$

[0058] 同理,对标准图像重复上述操作,得到标准图像的图像描述数据

$$[0059] \quad Img_{标} = \{Q_1, Q_2, Q_3, \dots, Q_n\};$$

[0060] 其中 Q_i 为标准图像中的连通域, n 为连通域个数;

[0061] 至此,步骤二完成,

[0062] 步骤三:对比标准图像与印刷图像的图像描述,判断异常情况;遍历标准图像和印

刷图像的连通域,将其按照层数 C 进行分组,并分别计算各个层级对应的个数;

[0063] 比较标准图像和印刷图像各个层级的对应个数,个数一致为正常情况,个数不一致的层级需要对该层级内的连通域进一步检测:设置标准图像中的对应连通域标志 $A = 0$,整体特征值差距 $B = 0$;印刷图像中的对应连通域标志 $E = 0$;寻找该层级各连通域之间对应关系:查找印刷图像某个连通域与标准图像中各个连通域的整体特征值 M 之间的差值:

$$[0064] \quad D = |M_p - M_{Q_i}|$$

[0065] 原则上,差值最小,即 D_{min} 时,对相关的两个连通域即为印刷图像(对应标签号 $T_{p'}$)与标准图像(对应标签号 T_Q)的对应的连通域 P' ;若标准图像连通域标志 $A = 0$,表示此时的标准图像中的连通域无对应连通域 P ,更新 A, B, E 的值:

$$[0066] \quad A = T_{p'}, B = D, E_{p'} = T_Q;$$

[0067] 若标准图像连通域标志 $A \neq 0$, ($A = T_p$),表示标准图像中的该连通域当前在印刷图像中已存在一个对应连通域 P ,比较该标准图像连通域的 B 与对应连通域 P' 对应差值 D 的大小:

[0068] 若 $B > D$,更新 A, B, E 的值,令

$$[0069] \quad A = T_{p'}, B = D, E_{p'} = T_Q, E_p = 0;$$

[0070] 若 $B = D$,获得该印刷图像中的对应连通域 P 与标准图像连通域的面积差距;获得对应连通域 P' 与标准图像该连通域的面积差距;如果对应连通域 P' 对应的面积差距较小,则更新 A, B, E 的值;否则不更新 A, B, E 的值;

[0071] 若 $B < D$,则不更新 A, B, E 的值。

[0072] 遍历结束后,印刷图像中连通域的对应连通域标志 E 存在两种情况: $E = 0, E \neq 0$;
 $E = 0$:表示该连通域在标准图像中不存在对应连通域,即缺陷区域,这种缺陷存在的位置为图案内部和图案外部空白区域,前者嵌套层次较多,不影响印刷产品的整体内容,故其缺陷影响程度较小,而后者的后者程度较大,所以该区域缺陷的整体影响情况 Y_0 为:

$$[0073] \quad Y_0 = \sum \frac{1}{L} \times S$$

[0074] $E \neq 0$:在标准图像中存在对应连通域,比较二者之间的面积差距,由于存在统计误差,误差在标准图像总面积的3%以内均为正常,超过3%的认为存在异常,存在图案缺印或多印的状况,则这些区域的缺陷整体影响情况为 Y_1 ,表示如下:

$$[0075] \quad Y_1 = \sum |S_p - S_q|$$

[0076] 综上所述,印刷图像整体存在的缺陷 Y 为: $Y = Y_0 + Y_1$

[0077] 以上所述,仅为本发明较佳的具体实施方式,但本发明的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本发明揭露的技术范围内,根据本发明的技术方案及其发明构思加以等同替换或改变,都应涵盖在本发明的保护范围之内。

