



# 基于机器视觉的化妆品纸质标签缺陷检测方法

刘 鹏<sup>1</sup>, 戴文战<sup>2</sup>

(1. 浙江理工大学自动化研究所, 杭州 310018;

2. 浙江工商大学信息与电子工程学院, 杭州 310018)

**摘 要:** 针对化妆品纸质标签生产中出现的不干胶标贴漏贴、偏移、倾斜、叠加, 生产日期字符漏喷、偏移、倾斜、多喷, 日期喷印错误等缺陷, 提出一种化妆品纸质标签缺陷视觉检测方法。首先, 利用数字形态学与连通域分析技术, 从倾斜校正后的标签本体图像中分别提取不干胶标贴和日期字符区域; 其次, 分别提取区域重心坐标与方向角检测不干胶标贴和日期字符的位置缺陷; 进而, 在日期点阵字符垂直校正基础上分割字符, 提取网格特征、垂直投影、投影宽度及字符占空比等字符特征, 利用这些特征训练 BP 神经网络并实现对日期点阵字符的识别; 最后, 在样机上采集了大量正常及具有不同类型缺陷的化妆品标签图像, 进行实验验证。实验结果表明: 该算法准确率高、稳定性好, 能够快速检测位置缺陷和准确识别日期字符, 检测准确率可达 94.4%。

**关键词:** 纸质标签; 质量检测; 特征提取; 点阵字符识别; 机器视觉

**中图分类号:** TP391.41

**文献标志码:** A

**文章编号:** 1673-3851 (2019) 03-0231-08

## Method for detecting cosmetic paper label defects based on machine vision

LIU Peng<sup>1</sup>, DAI Wenzhan<sup>2</sup>

(1. Institute of Automation, Zhejiang Sci-Tech University, Hangzhou 310018, China;

2. School of Information and Electronic Engineering, Zhejiang

Gongshang University, Hangzhou 310018, China)

**Abstract:** In view of the defects in the production of cosmetic paper labels, such as the leakage, offset, tilt and overlay of adhesive sticker labels as well as the leakage, offset, tilt, multiple spraying and date printing errors of the production date characters, this paper proposes a visual inspection method for the defects of cosmetic paper labels. Firstly, by using the digital morphology and the connected component analysis technology, we extracted the adhesive sticker label and date character area respectively from the image of the body after tilt correction. Secondly, in order to detect the position defects of the label and date characters, the coordinates and directional angles of the center of gravity of the region were respectively extracted. Thirdly, on the basis of vertical correction of the dot matrix characters of date, characters were segmented to extract grid features, vertical projection, projection width and character duty ratio, and these features were used to train the BP neural network and identify the date dot matrix characters. Finally, a large number of normal cosmetic label images and those with different types of defects were collected on the prototype for experimental verification. Experimental results showed that the algorithm had high

收稿日期: 2018-09-14 网络出版日期: 2018-12-01

基金项目: 国家自然科学基金项目 (61374022); 浙江省基础公益研究计划项目 (LGG18F030001); 金华市科学技术研究计划重点项目 (2018-1-027)

作者简介: 刘 鹏 (1992-), 男, 安徽淮北人, 硕士研究生, 主要从事图像处理、产品缺陷检测等方面的研究。

通信作者: 戴文战, Email: dwz@zjsu.edu.cn

accuracy and good stability. It could not only quickly detect the position defects, but also accurately identify the date characters, and the detection accuracy could reach 94.4%.

**Key words:** paper label; quality detection; feature extraction; dot matrix character identification; machine vision

## 0 引言

化妆品标签承载着产品的价格、功效、成分、日期等一些消费者及经销商关注的重要信息,其质量影响着消费者的购物心理及经销商对产品的第一印象。不合格的标签可能导致客户对企业的产品质量及信誉产生怀疑,从而给企业带来经济损失。因此,在标签与化妆品产品组装之前必须对其质量进行检测,剔除不合格标签。化妆品纸质标签上的内容大部分是由自动印刷机在线印刷而成,在快速印刷环境下,标签在生产过程中会受到生产工艺、环境、设备等众多不确定因素的影响,生产出来的标签存在各种缺陷及瑕疵。传统的化妆品标签缺陷检测主要依靠人工进行,这种方式检测速度难以满足目前生产速度和要求,检验员长时间集中注意力作业极易产生眼疲劳导致误检和漏检;而且人工成本逐年提高,人员流动性大,企业招工难。

随着计算机技术和数字图像处理技术的发展,基于机器视觉的自动化技术逐渐代替人工进行作业。机器视觉技术具有快速获取大量信息、快速自动化处理信息等特点,使得产品表面缺陷在线检测得到快速发展<sup>[1-2]</sup>,国内外学者针对不同研究对象提出了各种缺陷检测方法和图像处理算法,一些研究成果已经成功应用于铁轨、玻璃、纺织布、皮革等产品的缺陷检测<sup>[3-6]</sup>。目前,国内外学者对基于机器视觉的印刷品缺陷自动检测进行了研究。陈丽等<sup>[7]</sup>提出了印刷品缺陷检测算法,利用 DP (Dynamic partitioning) 方法对标准样本和缺陷样本进行分割和匹配,根据图像不同分辨率及检测精度来设计分割方式,可以有效地进行图形颜色信息的检测和分析,从而检测出缺陷特征及缺陷类型。但是该算法处理数据量较大,只能通过减少分区检测的像素数,才能提高检测速度。刑堃等<sup>[8]</sup>提出了一种基于自适应的 Canny 边缘掩模的方法,该方法可以有效消除因拍摄的非同步性和传送带抖动所产生的轮廓伪影,但是同时也会出现标签边缘缺陷漏检的情况。Mehenni 等<sup>[9]</sup>提出了一种速度快且参数输出多的逐像素比较和 *n-tuple* 相结合的检测方法,但是该检测方法只能获得缺陷位置,不能获得缺陷具体形状

及类型等信息,而且它需要专门的硬件设备才能完成自动质量检测任务。胡方尚等<sup>[10]</sup>提出了一种基于改进多类支持向量机的印刷缺陷检测方法,该方法首先将图像配准后的印刷图像进行动态阈值差分运算,接着通过分析统计缺陷几何特征和形状特征,利用改进的多类支持向量机实现印刷缺陷的识别。鹿武先<sup>[11]</sup>提出了一种食品日期针孔字符识别算法,通过对膨胀运算后的点阵字符进行连通域面积的均等分割,利用 BP 神经网络识别分割后的字符。该方法对于处理字符倾斜、粘连的效果较差,鲁棒性较低。陈宇等<sup>[12]</sup>提出一种基于数学形态学的彩色商标印刷缺陷特征提取算法,它通过对待检商标彩色图像分量与模板商标彩色图像分量的差分图进行多种形态学运算,然后提取商标印刷时产生的缺陷形状、位置等信息,但是对于反光材质商标、运动状态情况下的分量差分图效果不尽理想。

本文检测的化妆品标签生产工艺主要由不干胶标贴粘贴及日期点阵字符<sup>[13]</sup>喷墨组成。针对生产过程中出现的不干胶标贴粘贴位置缺陷、日期字符喷印位置缺陷、日期内容错误等不良现象,本文提出一种基于机器视觉的化妆品纸质标签缺陷检测方法。首先,从倾斜校正后的标签本体图像提取出不干胶标贴及日期字符区域;其次,根据不干胶标贴和日期可能出现的漏贴、漏喷、叠加、偏移、倾斜等缺陷现象,设计相应的位置缺陷检测算法;进而,提取网格特征、垂直投影、投影宽度及字符占空比等字符特征训练 BP 神经网络<sup>[14]</sup>,并进行生产日期点阵字符识别;最后,在样机上进行标贴位置缺陷、日期位置缺陷、日期字符喷印错误等不同缺陷类型的化妆品标签检测实验。

## 1 化妆品纸质标签缺陷及形成原因

对化妆品标签生产包装公司的现场考察和调研情况发现,标签缺陷的主要类型及其产生原因主要有:a)在传输过程中,由于传送机台的抖动及人为放置偏斜,导致不干胶标贴粘贴和日期字符的喷印发生位置偏移和倾斜;b)当控制粘贴和喷印的颜色传感器感应失灵时,不干胶标贴和日期字符的喷印可能出现漏喷、漏贴、叠加等现象;c)

当操作员将日期喷印装置设置错误时,会出现生产日期喷印错误。当不干胶标贴位置发生倾斜或偏移时,日期字符的喷印也会发生不同程度地偏移或倾斜。但是由于日期字符喷印区域范围较

大,轻微的倾斜及偏移并不影响整个标签的质量,化妆品标签生产包装公司只要求检测出偏移较大且超出允许范围的缺陷。合格及存在主要缺陷的化妆品标签图像如图 1 所示。



图 1  合格及存在主要缺陷的化妆品标签图像

2  化妆品纸质标签缺陷视觉检测方法

根据化妆品标签生产包装厂家的检测要求,本文提出的化妆品标签缺陷视觉检测流程如图 2 所示。该标签缺陷检测分为四个步骤:a)标签本体图像倾斜校正。提取标签本体图像重心及方向角,构造齐次旋转变换矩阵对标签本体图像进行倾斜校正; b)不干胶标贴位置检测。通过数字形态学开运

算与连通域面积提取出不干胶标贴区域,并根据其与标签本体图像区域的重心坐标差及方向角度差检测不干胶标贴位置;c)日期字符位置检测。通过形态学闭运算与连通域面积精确提取出日期字符区域,并根据其与标签本体图像区域的重心坐标差及方向角度差检测日期字符位置;d)日期字符识别。对倾斜校正后的日期字符进行分割,并提取归一化后的字符特征训练 BP 神经网络进行日期点阵字符的识别。

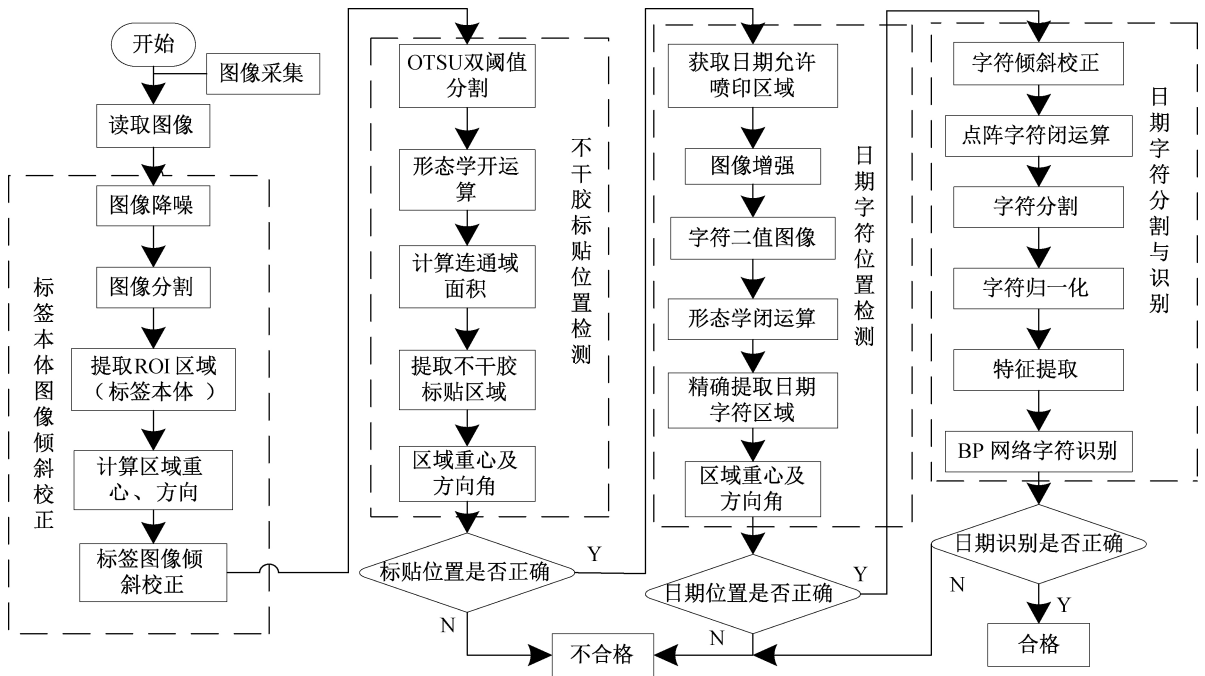


图 2  化妆品纸质标签缺陷检测流程

2.1  标签本体图像的倾斜校正

由于机台抖动及周围环境等因素的影响,标签在成像时会出现不同程度的倾斜现象,因此在进行不干胶标贴及日期字符位置检测之前要对

采集的图像进行倾斜校正。本文的倾斜校正原理如图 3 所示,主要包括图像的滤波降噪、阈值分割、标签本体图像的提取以及图像倾斜校正过程。

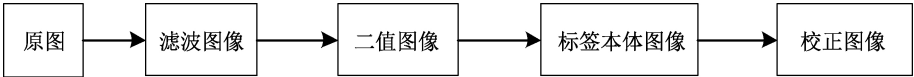


图3 倾斜校正原理

倾斜校正的具体过程如下：

a) 结合 Sigma 滤波<sup>[15]</sup>与形态学开运算，去除标签本体图像边缘噪声及其它细微噪。

b) 由于标签中存在文字、不干胶标贴、条形码、商标、打印码等，成像背景比较复杂，利用自适应OTSU 双阈值分割降噪后的图像，并通过连通域面积筛选提取 ROI 区域，即标签本体图像。

将图像 F 的灰度值分为  $m$  个等级，再将  $m$  个等级的灰度值分为  $C_0 \in [0, T_L]$ 、 $C_1 \in [T_L + 1, T_H]$ 、 $C_2 \in [T_H + 1, m]$  三类，图像分割时的高低阈值  $T_H$ 、 $T_L$  根据式(1)确定：

$$g(T_L, T_H) = \arg \max_{0 < T_L < T_H < m} \omega_0(T_1)(\mu - \mu_0(T_1))^2 + \omega_1(T_2)(\mu_1(T_2) - \mu)^2 + (1 - \omega_0(T_1) - \omega_1(T_2)) \left[ \frac{\mu\omega_0(T_1)\mu_0(T_1) - \omega_1(T_2)\mu_1(T_2)}{1 - \omega_0(T_1) - \omega_1(T_2)} - \mu \right]^2 \quad (1)$$

其中： $T_L \in [0, m - 1]$ ， $T_H \in (T_L, m]$ ； $\omega_0$  为  $C_0$  出现的概率， $\mu_0$  为  $C_0$  灰度平均值； $\omega_1$  为  $C_1$  出现的概率， $\mu_1$  为  $C_1$  灰度平均值； $C_2$  出现的概率为  $(1 - \omega_0 - \omega_1)$ ，灰度平均值为  $\frac{\mu - \omega_0\mu_0 - \omega_1\mu_1}{1 - \omega_0 - \omega_1}$ ； $\mu$  为整幅图像 F 的灰度平均值。当  $g(T_L, T_H)$  取最大值时对应的  $(T_L, T_H)$  即为最佳阈值。

分析分割图像的连通域面积，并利用连通域面积提取 ROI 区域，即化妆品标签本体图像。

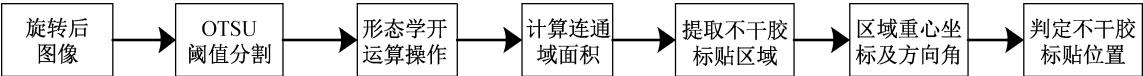


图4 不干胶标贴位置检测流程

由于倾斜校正后的标签前景图像经过 OTSU 双阈值分割后包含多个区域，为了提高不干胶标贴位置检测算法的准确率及效率，本文在对分割后区域的形态学开操作基础上，通过连通域面积提取不干胶标贴图像区域(如果没有满足面积条件的区域，则视为漏贴或缺损)，并在计算不干胶标贴图像区域的面积  $S_L$ 、重心坐标  $(x_2, y_2)$  及倾斜角度  $\theta_2$  的基础上，按如下程序对不干胶标贴位置进行检测：

```
if(abs( SL - S1)<2000)
    if(abs( θv - θ2) <= λ)
        if(abs( y1 - y2) <= εY)
            if(abs( x1 - x2) <= εX)
```

c) 计算标签本体图像的重心坐标及方向角，并构造齐次旋转变换矩阵对标签本体图像进行倾斜校正。

构造的齐次旋转变换矩阵  $H_L$  为：

$$H_L = \begin{bmatrix} 1 & 0 & x_1 \\ 0 & 1 & y_1 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \cos(\theta_v - \theta_1) & -\sin(\theta_v - \theta_1) & 0 \\ \sin(\theta_v - \theta_1) & \cos(\theta_v - \theta_1) & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & -x_1 \\ 0 & 1 & -y_1 \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

其中： $(x_1, y_1)$  为标签本体图像的重心坐标； $\theta_1 \in (-90^\circ, 90^\circ)$  为标签本体图像的方向角，即标签本体图像的最小外接椭圆长轴与水平方向的夹角； $\theta_v$  为旋转后的标签本体图像与水平方向的夹角。

利用  $H_L$  对标签本体图像  $I(X, Y)$  进行倾斜校正：

$$Q(X, Y) = H_L \cdot I(X, Y) \quad (3)$$

其中： $Q(X, Y)$  为倾斜校正后的标签本体图像， $X$ 、 $Y$  为图像像素的横、纵坐标。

2.2 不干胶标贴位置检测

不干胶标贴位置检测流程如图 4 所示，主要包括图像分割、不干胶标贴区域提取、区域重心坐标及方向角的计算以及不干胶标贴位置判定过程。

```
cout<<“不干胶标贴位置合格”<<endl;
else
    cout<<“标贴位置上下偏移”<<endl;
else
    cout<<“不干胶标贴位置左右移”<<endl;
else
    cout<<“不干胶标贴角度倾斜”<<endl;
else
    cout<<“不干胶标贴出现叠加情况”<<endl;
其中：S1 表示不干胶标贴的面积参数；λ 表示容许角度的偏差范围；εY 表示竖直方向的偏差范围；εX 表示水平方向上的位置偏差范围。
```

## 2.3 日期字符位置检测

日期字符位置检测流程如图 5 所示,包括日期允许喷印区域图像的增强、阈值分割、日期字符区域

的精确提取、字符区域重心坐标及方向角的计算以及日期字符位置判定过程。

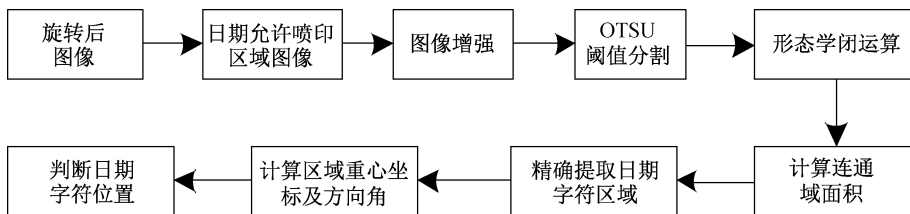


图 5 日期字符位置检测流程

日期字符位置检测的具体过程如下:

a) 根据日期字符所处标签本体图像的位置及标签本体图像重心坐标等先验知识,从倾斜校正后的标签本体图像提取出日期字符允许喷印范围区域图像,并利用 DoG (Different-of-Gaussian) 算子<sup>[16]</sup>增强图像突出感兴趣字符区域;

b) 对增强后的图像进行 OTSU 双阈值分割,并利用数学形态学闭运算将所有字符连接成一个连通区域;

c) 利用连通域面积提取出精确的日期字符区域,并记为面积  $S_D$  (如果没有满足面积条件的区域,则视为日期漏喷或喷印不全),在计算字符区域的重心  $(x_3, y_3)$  及水平方向倾斜角度  $\theta_3$  的基础上,按照如下程序对日期字符位置进行检测:

```

if(abs(  $S_D - S_2$  ) < 300)
    if (  $\theta_3 \leq \beta$  )
        if(abs(  $y_1 - y_3$  ) <=  $\delta_Y$  )
            if(abs(  $x_1 - x_3$  ) <=  $\delta_X$  )
                cout<<"日期字符位置合格"<<endl;
            else
                cout<<"日期字符位置上下偏移"<<endl;
        else
            cout<<"日期字符位置左右偏移"<<endl;
    else
        cout<<"日期字符角度倾斜"<<endl;
    else
        cout<<"日期字符出现叠加情况"<<endl;
  
```

其中:  $S_2$  表示日期字符的面积参数;  $\beta$  表示容许角度的偏差范围;  $\delta_Y$  表示竖直方向的偏差范围;  $\delta_X$  表示水平方向上的位置偏差范围。

## 2.4 化妆品标签的日期字符识别

### 2.4.1 字符倾斜校正

传送带速度不均、机台抖动导致化妆品日期在喷印的过程中出现不同程度的水平倾斜及垂直倾

斜,在进行日期字符识别前需要对其进行倾斜校正。水平倾斜是指字符上下扭曲变形,即水平方向坐标不变,只改变竖直方向坐标的变形方式;而垂直倾斜是指字符左右扭曲变形,即垂直方向坐标不变,只改变水平方向坐标的变形方式。本文日期字符倾斜校正过程如下:

a) 计算字符区域重心  $(x_3, y_3)$  和水平倾斜角  $\theta_3$ ,在此基础上构造齐次变换矩阵  $H_D$  对字符图像进行水平校正;

b) 计算字符区域垂直倾斜角  $\varphi$ ,根据字符区域重心  $(x_3, y_3)$  和垂直倾斜角  $\varphi$  构造齐次变换矩阵  $H_V$  对字符图像进行垂直倾斜校正。

### 2.4.2 日期字符分割

由于化妆品喷印日期为点阵字符,单个字符中的点仍处于分离状态。为了完整有效地将单个字符分割开,提高字符的识别效率,本文将二值字符图像进行形态学闭操作,把单个点阵字符笔画连接起来,并利用水平投影法进行字符行分割,再利用垂直投影法将字符分割成单个字符。具体日期字符分割过程如下:

a) 字符行分割:对输入的二值化字符图像从上到下逐行扫描,同时统计每个扫描行像素值为 0 的像素点数,则该像素行的黑点数的总和就是该行的投影值。把所有行的投影值统计出来得到整个二值化字符图像的水平投影,并利用字符行与行之间的空白间隔(投影图像的波谷)将各行字符分割开。

b) 单个字符分割:从左向右逐列扫描行分割后的二值字符图像,统计每列投影图像的投影数值记为  $N(j)$ ;遍历投影数值,第一次出现满足条件  $((N(j) = 0) \text{ and } (N(j+1) \neq 0))$  的点  $j_{sl}$ ,使其作为第一个字符的左边界。满足条件  $((N(j) \neq 0) \text{ and } (N(j+1) = 0))$  点  $j_{el}$ ,使其作为第一个字符的右边界。以此类推,找出所有字符的起止位置  $(j_{si}, j_{ei})$  作为分割位置;剔除符号“:”,由字符的特

征可以看出,“:”所占黑色像素值最少且在水平方向上的投影宽度也最小,因此根据“:”的垂直投影数值 $N(j)$ 和字符宽度( $j_{ei}-j_{si}$ )将其剔除。

2.4.3 日期字符识别

日期字符识别主要包括分割后日期字符的特征提取以及应用BP神经网络对字符识别过程。日期字符识别具体过程如下:

a) 特征提取:根据化妆品喷印日期的点阵字符特点,为了提高字符识别的准确率及效率,本文提出将字符网格统计数值及单个字符的垂直投影数值、投影宽度、字符占空比等作为特征进行字符识别。各特征参数计算具体如下:

a1) 将分割后的归一化二值字符划分为48个像素数相同的网格小块,统计每个网格小块内像素值为“0”的像素个数,可以得到一个包含48个整数值的数组,将这些网格数值作为字符识别的特征值。字符“E”的点阵图像如图6所示。

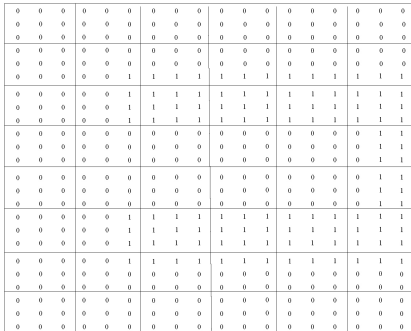


图6 字符“E”点阵图像

a2) 单个字符垂直投影数值 $N_s(i)$ 及投影宽度 $W_{si}$ 为:

$$W_{si}=j_{ei}-j_{si} \quad (i=1,\cdots,23) \tag{4}$$

$$N_s(i)=\sum_{j=1}^{W_{si}}N(j) \tag{5}$$

a3) 字符占空比 $D_s$ 为:

$$D_s=\frac{N_s(i)}{18\times 24}=\frac{\sum_{j=1}^{j_{ei}-j_{si}}N(j)}{432} \tag{6}$$

b) 字符识别:BP神经网络是一种由输入层、隐藏层、输出层组成的多层前馈神经网络,无论是从性能还是原理上都已非常成熟,该网络具有很强的非线性映射能力可以逼近任意非线性函数。化妆品标签日期由固定的4个英文字符“E”、“L”、“P”、“X”及阿拉伯数字“0,1,...,9”共14个字符组成,是一个多样本类分类问题,由于所需识别量较小,因此选用三层BP神经网络构造一个数字和字符混合的神经网络。把提取出的51个特征值组成51行1列的向量当作输入层的输入向量,即输入层的节点数设为51;输出层节点数根据所需识别日期字符的样本种类来确定,网络的输出端采用一个14行1列的输出向量,所以输出层的节点选定为14个;对于隐藏层节点数,在此根据公式 $n=\sqrt{a\times(b+1)+1}$ 计算隐层节点数。其中 $n$ 表示隐层节点数, $a$ 表示输入层节点数, $b$ 表示输出层节点数,计算得出隐藏层的节点数为28个。

3 实验与结果分析

3.1 不干胶标贴及日期位置参数的设定

实际生产中,允许不干胶标贴粘贴和日期字符喷印在一定范围内有倾斜及位置偏移现象发生。为确定各有效参数,在实验中采集大量图像,分别计算出不干胶标贴和日期字符的面积、重心坐标差及角度差,然后将计算结果进行统计,最终确定合格标签的不干胶标贴及日期各位置参数范围。表1为不干胶标贴位置参数,表2为日期字符位置参数。

表1 不干胶标贴位置参数

缺陷类型	面积 $S_L$ /像素	重心坐标差/像素	角度差(弧度)
合格标贴	$S_1:130000\pm2000$	$\epsilon_X:[168.00,192.00],\epsilon_Y:[0,11.00]$	$\lambda:[0,0.0150]$
叠加	138297	(166.82,14.50)	0.0045
漏贴	0	—	—
倾斜	130245	(173.50,11.93)	0.0165
偏移	130325	(174.91,13.85)	0.0061

表2 日期字符位置参数

缺陷类型	面积 $S_D$ /像素	重心坐标差/像素	角度差(弧度)
合格字符	$S_2:7500\pm300$	$\delta_X:[505.00,575.00],\delta_Y:[0,20.00]$	$\beta:[0,0.0200]$
叠加	8528	(553.54,14.50)	0.0125
漏喷	0	—	—
倾斜	7620	(546.44,27.78)	0.0239
偏移	7581	(542.50,45.00)	0.0138



根据以上数据可得,叠加缺陷的面积明显大于合格标签;标贴漏贴和日期漏喷则表现出同样的特征,用面积即可将其筛选出来;倾斜缺陷的特征除了方向角度差明显大于固定参数范围,还伴着重心纵坐标的偏移;偏移缺陷的重心坐标差相较于合格的标签有明显的差异,但其面积和角度在参数允许范围内。因此可以将面积、重心坐标差、角度差这 3 个参数作为判断不干胶标贴及日期字符位置是否合

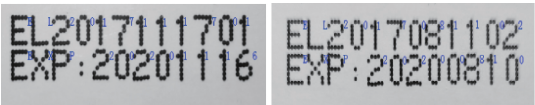
格的依据。

### 3.2 日期点阵字符识别结果比较

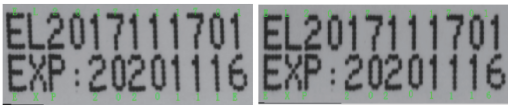
字符识别过程中的每一步都会对最终的字符识别率产生影响,为明确识别率的具体影响因素及测试本算法的有效性,以 125 枚合格标签图像进行实验验证并与文献[11]算法进行比较。本文算法与文献[11]算法部分标签日期字符识别结果如图 7 所示,表 3 为字符识别过程的数据统计。

表 3 字符识别过程数据统计

统计过程	待测样本数		错误次数		正确次数		正确率/%	
	本文	文献[11]	本文	文献[11]	本文	文献[11]	本文	文献[11]
字符定位	125	125	1	0	124	125	99.2	100
字符分割	124	125	4	11	120	114	96.8	91.2
字符识别	120×23=2760	114×23=2622	25	59	2735	2563	99.1	97.7
整个字符识别过程	125×23=2875	125×23=2875	140	312	2735	2563	95.1	89.1



(a) 本文算法识别结果



(b) 文献[11]算法识别结果

图 7 部分日期字符识别结果

由表 3 可以看出,本文算法根据标签重心及标签特征等先验知识可以准确定位日期字符。在字符分割方面,分割不准确主要是受图像采集质量的影响,分割错误的字符图像都呈现出字符间有异物或者字符成像较模糊导致字符不完全分割。文献[11]采用连通域分割与字符固定宽高分割相结合的分割方法,该方法对粘连字符、倾斜字符及宽度不同字符分割效果不尽理想;在字符识别效果方面,本文将提出的字符特征作为神经网络的输入,并经大量样本训练可以达到对日期字符的准确识别。文献[11]直

接将待检图片输入神经网络,由神经网络自提取字符特征以此实现日期字符识别。实验结果表明本文算法在字符分割及字符识别过程均优于文献[11],整个字符识别过程的识别效率达到 95.1%。

### 3.3 实验检测结果

为了检测整个算法的实用性与准确性,本文采集了 232 张标签图像,合格标签图像 75 张,不合格标签图像 157 张,其中不合格标签包括标贴位置缺陷、日期字符位置缺陷以及日期喷印错误。最后实验检测结果见表 4、表 5。

表 4 实验检测结果

检出数/合格标签	合格标签错检率/%	检出数/不合格标签	不合格标签漏检率/%	正确检测产品数/标签总数	检测率/%
71/75	5.3	148/157	5.7	219/232	94.4

表 5 不合格标签检测结果

缺陷类型	缺陷检出数/总数	检出率/%
标贴位置缺陷	73/78	93.6
日期位置缺陷	50/53	94.3
字符喷印错误	25/26	96.2

## 4 结 论

本文根据标签制造工艺将标签缺陷检测分为标签本体图像的倾斜校正、不干胶标贴位置检测、日期字符喷印位置检测及日期字符识别四个过程。根据

OTSU 双阈值分割及连通域面积筛选方法从倾斜校正后的标签本体图像中提取出不干胶标贴及日期字符区域,并通过计算面积、重心坐标、方向角来检测其位置缺陷;利用水平投影及垂直投影分割字符,并将提取出的字符网格特征、垂直投影、字符占空比等字符特征输入到 BP 神经网络完成对日期字符的识别。实验结果表明,本文算法不仅可以快速有效地检测出位置缺陷,而且可以完成对日期字符的识别,整个算法的检测率可达到 94.4%。但是,在对标签日期进行提取定位时,由于不同类型标签的先

验知识不同,所以要手动设置日期字符所允许喷印的区域范围,不能达到完全自动化,还需要进一步改善。

### 参考文献:

- [1] 徐浪,曾忠,刘金赞,等. 机器视觉在印刷缺陷在线检测中的应用与研究[J]. 计算机系统应用, 2013, 22(3): 186-190.
- [2] Truong M T N, Kim, Hoon S. Automatic image thresholding using Otsu's method and entropy weighting scheme for surface defect detection[J]. Soft Computing, 2018, 22(13): 4197-4203.
- [3] Zhang H, Jin X T, Jonathan Wu Q M, et al. Jonathan. Automatic visual detection system of railway surface defects with curvature filter and improved gaussian mixture model [J]. IEEE Transactions on Instrumentation and Measurement, 2018, 67(7): 1593-1608.
- [4] 简川霞,高健. 手机玻璃屏表面缺陷视觉检测方法研究[J]. 包装工程, 2018, 39(5): 16-21.
- [5] 冀永乐,张团善,颜亚飞,等. 基于机器视觉的纺织布品色差检测研究[J]. 西安工程大学学报, 2018, 32(4): 438-442.
- [6] 孙毅,张明,高延鑫. 皮革表面缺陷检测的研究现状及展望[J]. 西部皮革, 2015, 37(20): 28-31.
- [7] 陈丽,唐万有. 基于 DP 方法的印刷品图像检测技术研

究[J]. 包装工程 2014, 35(5): 116-120.

- [8] 邢堃,韩汉光,吴怡之. 基于机器视觉的印刷标签检测系统的改进[J]. 计算机工程与应用, 2014, 50(11): 197-201.
- [9] Mehenni B, Wahab M A. APRIS: Automatic pattern recognition and inspection system[C]// Compeuro '93. 'computers in Design, Manufacturing, and Production', Proceedings. IEEE, 1993: 23-28.
- [10] 胡方尚,郭慧. 基于改进多类支持向量机的印刷缺陷检测[J]. 华东理工大学学报(自然科学版), 2017, 43(1): 143-148.
- [11] 鹿武先. 基于 Halcon 的食品生产日期针孔光学字符检测[J]. 软件导刊, 2017, 16(4): 168-170.
- [12] 陈宇,何邦贵,顾文娟. 一种商标印刷缺陷特征提取算法的研究[J]. 制造业自动化, 2014, 36(18): 37-38.
- [13] 黄斌,郭海华,江进兵. 基于机器视觉的点阵字符识别系统设计[J]. 测控技术, 2014, 33(12): 24-26.
- [14] 焦敬品,李勇强,吴斌,等. 基于 BP 神经网络的管道泄漏声信号识别方法研究[J]. 仪器仪表学报, 2016, 37(11): 2588-2596.
- [15] 邓少平,李平湘,张继贤,等. 一种改进的极化 SAR Sigma 滤波法[J]. 测绘信息与工程, 2010, 35(6): 1-3.
- [16] 刘吉磊,李强子,杜鑫. 基于高斯差分模型的雪地扰动痕迹遥感识别[J]. 遥感技术与应用, 2015, 30(1): 140-147.

(责任编辑:康 锋)