

[自控·检测]

DOI:10.3969/j.issn.1005-2895.2020.01.012

# 基于 HOG 局部双线性插值的机械 零部件检测与识别

王子阳, 魏丹, 胡晓强, 罗一平, 方轶

(上海工程技术大学机械与汽车工程学院, 上海 201620)

**摘要:**传统方向梯度直方图(HOG)特征方法用于机械零部件检测时,对零部件的微小缺陷点的检测和识别不是很有效,因此课题组提出了一种局部双线性插值的 HOG 特征提取改进算法,将其用于小缺陷点的零件的检测和识别。首先,在图像内部的某个块上实现局部双线性插值;然后对局部双线性插值图像块进行梯度计算,提取新的梯度方向直方图;最后利用神经网络分类方法对具有微小缺陷点的零件进行检测。实验结果表明:局部双线性插值 HOG 特征提取方法比传统的 HOG 特征提取方法具有更好的检测性能;增强图像的抗混叠识别效果。

**关键词:**零件检测;方向梯度直方图(HOG);局部双线性插值;神经网络

中图分类号:TH164;TP274 文献标志码:A 文章编号:1005-2895(2020)01-0065-06

## Mechanical Parts Detection and Recognition Based on Local Bilinear Interpolation of HOG

WANG Ziyang, WEI Dan, HU Xiaoqiang, LUO Yiping, FANG Yi

(School of Mechanical and Automotive Engineering, Shanghai University of Engineering Science, Shanghai 201620, China)

**Abstract:** The traditional HOG feature method in the detection of mechanical parts are not very effective for the detection and recognition of small defect points of industrial parts. Therefore, an improved HOG feature extraction algorithm based on local bilinear interpolation was proposed for the detection and identification of mechanical parts which have small defects. Firstly, local bilinear interpolation was implemented on some internal block inside the image. Then the gradient of the local bilinear interpolation image block was calculated and the histogram of the new oriented gradient were abstracted. Finally, the components with tiny flaw point by neural network classification were detected. The experimental results demonstrate that the local bilinear interpolation HOG features extraction offers better detection performance than the conventional HOG feature extraction method and it can also enhance the effect of anti-aliasing image recognition.

**Keywords:** mechanical parts detection; HOG (histogram of oriented gradient); local bilinear interpolation; neural network

随着计算机技术的飞速发展,目标检测已成为近年来模式识别领域中最具挑战性的问题之一<sup>[1]</sup>。目标检测包括2个主要任务:特征提取和分类。经典的特征提取方法包括主成分分析<sup>[2]</sup>和梯度直方图(HOG)<sup>[3]</sup>等。在分类方法上,常用的分类方法有最近邻<sup>[4]</sup>、神经网络<sup>[5]</sup>和支持向量机<sup>[6]</sup>等。在HOG图像处理方法中,图像被分割成小块,它是一种有效的预处理

步骤,用于检测具有准规则结构(如行人)的目标。与尺度不变特征变换(scale-invariant feature transform, SIFT)描述符类似, HOG 描述符由一组在空间偏移的局部块上计算的归一化直方图组成;结果是一个描述符,它捕获了粗糙的空间结构,但对局部的小变形是不变的。

Datal 等<sup>[7]</sup>提出了一种基于梯度直方图的人体检测方法,在人体检测方面取得了良好的效果。但是对

收稿日期:2019-07-14;修回日期:2019-11-05

基金项目:国家自然科学基金项目(51805312);上海市地方能力建设资助(19030501100)。

第一作者简介:王子阳(1991),男,安徽阜阳人,硕士研究生,主要研究方向为机器视觉、无人驾驶技术。E-mail: wangziyangwilliam@163.com

于工业部件检测与识别中的一个微小缺陷点,会出现边缘混叠现象,因此,课题组提出了一种局部双线性插值的梯度直方图特征提取算法,将其用于存在小缺陷点的零件上,并进行检测和识别。首先,在图像内部的某个块上实现局部双线性插值;然后对局部双线性插值图像块进行梯度计算,提取新的梯度方向直方图;最后利用神经网络分类方法对具有微小缺陷点的零件进行检测。数据库上的实验结果验证了该算法的鲁棒性和改进效果,增强了图像的抗混叠识别效果。

### 1 梯度直方图特征提取

梯度直方图(HOG)描述符试图用一个小的图像窗口来构造一个更详细的空间结构特征,这是一个非常有效的检测对象准规则结构预处理步骤的算法(如行人再识别)。与任何其他描述符一样,梯度直方图描述符由一组在空间偏移补片上计算的标准化直方图组成,结果是一个描述符,它捕获了粗糙的空间结构,但对局部的小变形是不变的。计算梯度直方图描述符的过程适合于行人检测,它包括以下几个阶段:首先在一个(64×64) dpi 范围中计算每个像素处图像梯度的方向和幅度;然后通过式(1)得到字符值。

对于垂直边缘和水平边缘,如伽马压缩如式(1)所示, $\gamma$ 常取 1/2。

$$I(x,y) = I(x,y)^\gamma. \tag{1}$$

在图像中,一阶像素滤波器和二阶滤波器<sup>[8]</sup>的梯度计算如下:

$$\left. \begin{aligned} G_x(x,y) &= H[(x+1),y] - H[(x-1),y]; \\ G_y(x,y) &= H[x,(y+1)] - H[x,(y-1)]. \end{aligned} \right\} \tag{2}$$

$$\left. \begin{aligned} G_x(x,y) &= H[(x+2),y] - 2H(x,y) + H[(x-2),y]; \\ G_y(x,y) &= H[x,(y+2)] - 2H(x,y) + H[x,(y-2)]. \end{aligned} \right\} \tag{3}$$

式中: $G_x(x,y)$ ,  $G_y(x,y)$ ,  $H(x,y)$  分别表示输入图像像素(x,y)的水平梯度、垂直梯度和像素值。

计算像素(x,y)的梯度幅值和梯度方向:

$$\left. \begin{aligned} G(x,y) &= \sqrt{G_x(x,y)^2 + G_y(x,y)^2}; \\ \alpha(x,y) &= \tan^{-1} \left[ \frac{G_y(x,y)}{G_x(x,y)} \right]. \end{aligned} \right\} \tag{4}$$

方位量化为分布在 180 dpi 范围内的 9 个块。大小为(64×64) dpi 的探测器区域被划分成一个由重叠的 16 个单元组成的规则网格。在每个单元内计算一个九维方向直方图,其中直方图的贡献由梯度振幅和距单元中心的距离加权确定,从而使更多的中心像素贡献更多。对于每个(2×2) dpi 单元格块,将描述符

连接起来并进行规范化,以形成块描述符。单元格和块在图 1 所示的(64×64) dpi 图像中分割。

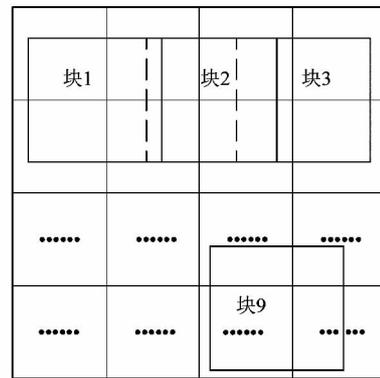


图 1 (64×64) dpi 图像分割成单元格和块

Figure 1 Cell and block divided in (64×64) dpi image

所有的块描述符都连接起来,形成最终的 HOG 描述符。最后的描述符包含关于局部梯度(在每个单元内)的空间池化信息,还维护一些空间分辨率(因为有许多单元)。HOG 描述符只使用梯度大小来创建极性对比的不变性;HOG 描述符通过对每个块进行归一化来创建不变性来对比局部强度。HOG 描述符在本质上类似于 SIFT 描述符,区别在于 HOG 描述符的对比度保持不变,具有较高的空间分辨率,并且更容易对局部执行标准化。

HOG 描述符主要应用于计算机视觉和图像处理领域,是一种用于目标检测的特征描述符,可用于计算局部图像梯度方向信息的统计值。这种方法与边缘方向直方图和尺度不变特征变换描述符以及形状上下文方法有很多相似之处,但是也有区别:HOG 描述符计算密集的网格的均匀间隔的细胞来提高性能,还采用了重叠的局部对比度归一化技术。

对于(64×64) dpi 图像,我们可以获得所有的描述符为  $3 \times 3 \times 36 = 324$  dpi,其直方图和描述符如图 2 和图 3 所示。

## 2 利用双线性插值的 HOG 特征提取

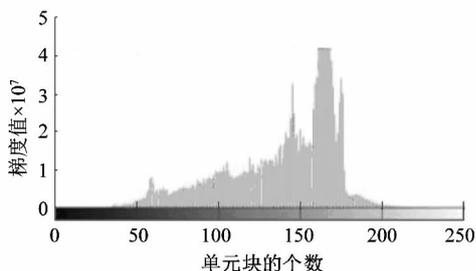
### 2.1 双线性插值

双线性插值算法由于计算效率高和去除损伤图片等特点,在各种插值方法中得到了广泛的应用<sup>[9]</sup>。

在数学中,双线性插值是 2 变量线性插值函数的插值扩展,其核心思想是有 2 个方向可以分别进行线性插值。如图 4 所示,点 P 的值已知,采用双线性插值的方法将按 4 个点  $Q_{11}(x_1,y_1)$ ,  $Q_{12}(x_1,y_2)$ ,  $Q_{21}(x_2,y_1)$  和  $Q_{22}(x_2,y_2)$  的比例进行分配。



(a) 零件



(b) 直方图

图2 (64 × 64) dpi 图像的原图及其梯度直方图

Figure 2 Image and histogram of its (64 × 64) dpi size

插值计算如下:

1)  $y$  方向上的插值

$$\left. \begin{aligned} f(Q_1) &= (y_2 - y)(y_2 - y_1) \times f(P); \\ f(Q_2) &= (y - y_1)(y_2 - y_1) \times f(P)。 \end{aligned} \right\} \quad (5)$$

2)  $x$  方向上的插值

$$\left. \begin{aligned} f(Q_{11}) &= (x_2 - x)(x_2 - x_1) \times f(Q_1) = (x_2 - x)(y_2 - y)(x_2 - x_1)(y_2 - y_1) \times f(P); \\ f(Q_{12}) &= (x_2 - x)(x_2 - x_1) \times f(Q_2) = (x_2 - x)(y - y_1)(x_2 - x_1)(y_2 - y_1) \times f(P); \\ f(Q_{21}) &= (x - x_1)(x_2 - x_1) \times f(Q_1) = (x - x_1)(y_2 - y)(x_2 - x_1)(y_2 - y_1) \times f(P); \\ f(Q_{22}) &= (x - x_1)(x_2 - x_1) \times f(Q_2) = (x - x_1)(y - y_1)(x_2 - x_1)(y_2 - y_1) \times f(P)。 \end{aligned} \right\} \quad (6)$$

该算法减少了图像调整为非积分缩放因子所造成的一些视觉失真。

## 2.2 使用双线性插值的梯度直方图特征值

对于梯度的性质,描述符对噪声相当敏感。当使用 HOG 算法时,它通常被用来固定一个划分为几个区域的变量范围。因为边界变量的邻域也存在相关性,只对一个区域进行变量投影,邻域与其他区域完全无关,就会产生混叠效应。这种混叠现象在图像轮廓上比原始图像更加明显。为了提高特征向量的抗混叠效果,我们将通过 4 个单元的 4 个中心来统计梯度,然后对 4 个单元中心进行双线性插值,插值区域如图 5 所示。

对图像进行插值计算时,每个细胞的直方图都是统计的;直方图方向相同,4 个角度停留在像素点上;

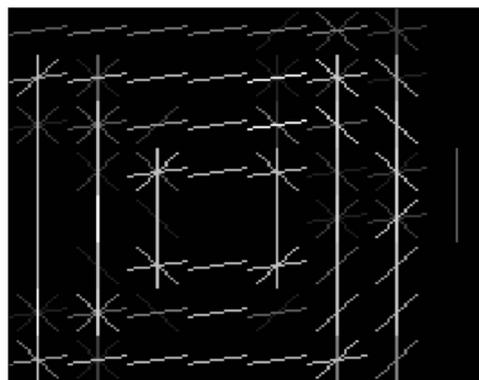


图3 计算的 HOG 描述符

Figure 3 Computed HOG descriptor

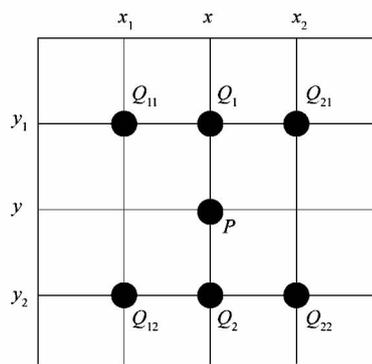


图4 双线性插值计算

Figure 4 Bilinear interpolation calculation

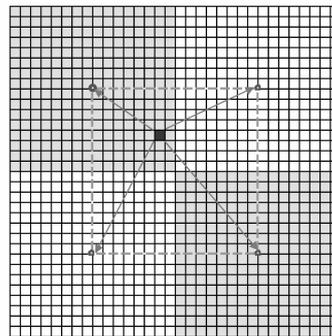


图5 像素在单元格之间分配权重

Figure 5 Pixels weights allocation between cells

梯度值的大小应该作为权重投影到单元格中。另一方面,双线性插值算法使用了大量的乘法,随着图像尺寸

的增大,计算量也随之增大,最终导致功耗增大,处理速度变慢。笔者对图像本身相邻像素间的内在关系采用简化算法,并对插值算法采用迭代累加和移位的方法。实验结果表明该算法的性能得到了提高。

### 2.3 神经网络分类

距离判别法、贝叶斯判别法和费雪判别法等传统的判别分析方法,在判别分析中有较高的流产率。因此,引入概率神经网络 (probalistic neural networks, PNN) 判别分析是十分必要的。神经网络具有较高的精度,并具有较强的分布存储和学习能力,它还具有较强的噪声和容错能力,能够完全逼近复杂的非线性关系,还具有联想记忆功能。实验采用概率神经网络<sup>[10]</sup>对图像进行分类,因为概率神经网络处理大规模图像具有明显的优势。

神经网络的结构一般由一个输入层、一个或多个隐藏层和一个输出层组成,如图 6 所示。概率神经网络的基本过程是:首先计算输入向量与目标向量之间的距离,形成一个新的向量,新的向量表示输入与目标输出之间的差异。第 2 层对每个输入权重求和,计算每个输入的输出,得到一个概率向量。第 3 层的竞争传递函数很可能从 2 层的概率输出向量中挑出最大的概率值,使最大的概率归入相应的类别。

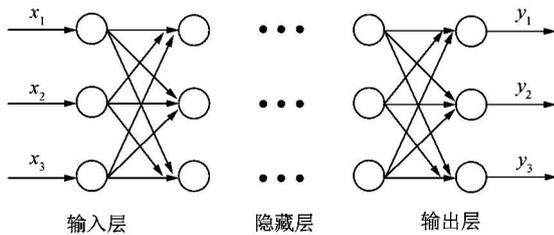


图 6 神经网络结构

Figure 6 Neural network

PNN 解决了分类与识别的问题,可以应用于大部分领域。只要有一定的样本容量,PNN 就会收敛到贝叶斯分类。而 PNN 算法也有不足:①计算分类的时间较长。当然,如果将 PNN 用于经济分析,而不是硬件产品设计,计算时间可以忽略。②研究者不能根据样本得到训练分类判别函数,即不能对输入和输出进行结构分析。

### 3 实验及其结果

在实验中,我们的主要任务是识别出 3 种类型零件的图像:无缺陷零件、焊接点零件和划痕零件。这些图像是在实际生产中通过工业相机获得的。同时,本实验共使用了 900 幅图像,每种类型的图像数量为

300 幅。3 种汽车零部件图像的实例如图 7 所示。

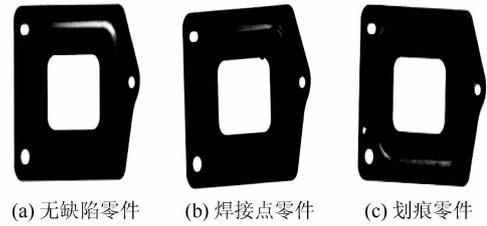


图 7 汽车零部件的 3 种种类的举例说明

Figure 7 Example of three kinds of automobile parts image

为使图像处理更加方便,有必要对原始图像<sup>[11]</sup>进行预处理。笔者对所有图像进行归一化处理,并将其分割成(128 × 128) dpi 像素的图像,同时比较了该方法与非双线性插值方法在不同尺寸图像上的实验结果。实验进行了 10 次,最终精度取 10 次结果的平均值。所有实验均在 Intel core i3 CPU M350 @ 2.27 GHz (2 核 4 线程) 计算机上使用 MATLAB 实现。课题组使用不同阶滤波模板得到梯度,检测分类过程如下:

1) 随机选取每种图像中的 100 幅作为训练样本,其余图像作为测试样本。每一种图像的提取量相同,实验结果如表 1 所示。从表 1 可以看出,当图像尺寸为(32 × 32) dpi 时,采用 2 阶滤波模板和双线性插值得到了最高的精度。同时,对于所有其他项目,准确率都超过 90%。可以看出,课题组提出的方法具有较好的检测精度,无论图像的大小,2 阶滤波器都取得了较好的检测精度。

表 1 采用 100 个训练样本非双线性插值和双线性插值的神经网络分类的精度

Table 1 Accuracy of neural network classification using bilinear interpolation and non-bilinear interpolation with 100 training samples

图像尺寸/ (dpi × dpi)	模板	非双线性 插值/%	双线性 插值/%
128 × 128	2 阶滤波器	95.8	97.5
	1 阶滤波器	91.3	92.3
64 × 64	2 阶滤波器	94.5	95.2
	1 阶滤波器	94.1	97.0
32 × 32	2 阶滤波器	98.3	98.7
	1 阶滤波器	89.2	96.7

2) 随机选取每种图像中的 200 幅作为训练样本,其余图像作为测试样本。每一种图像的提取量都是相同的,实验结果如图 8 ~ 11 所示。由表 2 可以看出,当

图像尺寸为(64×64) dpi 时,使用 2 阶滤波模板和双线性插值得到的精度最高。同时,对于所有其他项目,准确率都超过 90%。可以看出,文中提出的方法具有较好的检测精度。由于 2 阶滤波器可以达到更好的检测精度,在接下来的实验中使用了该滤波器。

表 2 采用 200 个训练样本非双线性插值和双线性插值神经网络分类的精度

Table 1 Accuracy of neural network classification using bilinear interpolation and non-bilinear interpolation with 200 training samples

图像尺寸/ (dpi × dpi)	模板	非双线性 插值/%	双线性 插值/%
128 × 128	2 阶滤波器	97.2	98.7
	1 阶滤波器	93.4	94.8
64 × 64	2 阶滤波器	98.3	99.7
	1 阶滤波器	97.3	98.5
32 × 32	2 阶滤波器	98.2	99.1
	1 阶滤波器	95.6	95.7

笔者比较了使用双线性插值和非双线性插值的提取特征,以及使用神经网络和 SVM 分类的结果。图 8 和图 9 为使用不同样本训练的 1 阶插值结果,图 10 和图 11 为使用不同训练样本的 2 阶插值结果。其中, Hog + Network 是指通过非双线性插值提取梯度直方图特征,并通过神经网络对图像进行分类; BiHog + Network 是指用双线性插值法提取梯度直方图特征,并通过神经网络对图像进行分类; Hog + SVM 指的是通过非双线性插值提取梯度直方图特征,并通过支持向量机对图像进行分类; BiHog + SVM 指的是通过双线性插值提取梯度直方图特征,并通过支持向量机对图像进行分类。

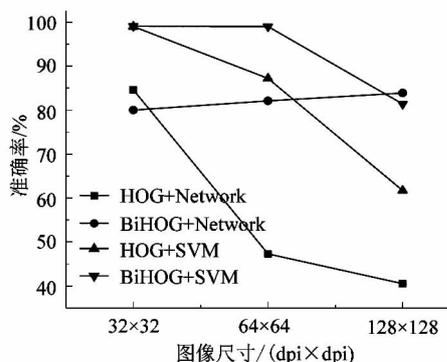


图 8 采用 100 个训练样本的 1 阶插值结果  
Figure 8 Results of 1st order interpolation with 100 training samples

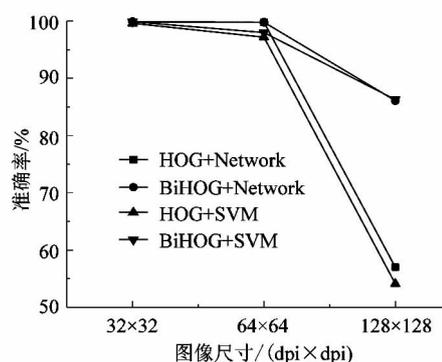


图 9 采用 200 个训练样本的 1 阶插值结果  
Figure 9 Results of 1st order interpolation with 200 training samples

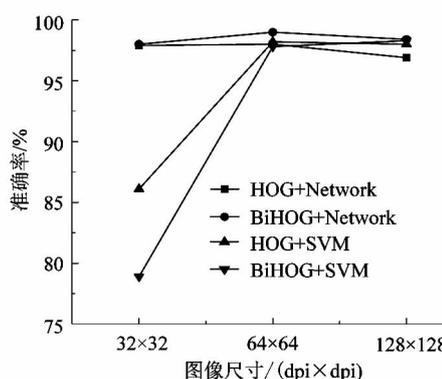


图 10 采用 100 个训练样本的 2 阶插值结果  
Figure 10 Results of 2nd order interpolation with 100 training samples

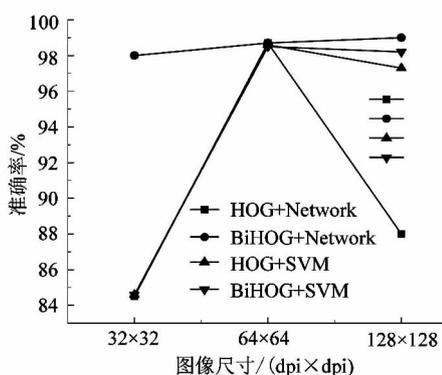


图 11 采用 200 个训练样本的 2 阶插值结果  
Figure 11 Results of 2nd order interpolation with 200 training samples

#### 4 结论

课题组提出了一种局部双线性插值的梯度直方图特征提取算法,用于对零件中存在小缺陷点的零件进行检测和识别。对局部双线性插值图像块进行梯度计

算,提取新的梯度方向直方图。采用神经网络分类方法对微小裂纹点的成分进行检测。实验结果表明:局部双线性插值梯度直方图特征提取方法比传统的梯度直方图特征提取方法具有更好的检测性能;增强了图像的抗混叠识别效果。但是,从计算结果可以看出,该方案具有较高的维特征描述符,且精度始终没有达到100%。因此,在降维和选择不同特征值方面还有待进一步研究。

#### 参考文献:

- [1] HJELMSA E, LOW B K. Face detection: a survey [J]. Computer Vision and Image Understanding, 2001, 83(3): 236 - 274.
- [2] WEN Chenglin, HU Ting, WANG Tianzhen. Relative principle component and relative principle component analysis algorithm [C]// Advances in Neural Networks 2007. Nanjing: Springer-verlag Berlin Heidelberg, 2007: 985 - 993.
- [3] HUANG Chenjie, QIN Zheng, XU Kaiping, et al. Beyond HOG: learning local parts for object detection [C]// Advances in Multimedia Information Processing. [S. l.]: Springer, 2015: 285 - 295.
- [4] SAVCHENKO A V. Nonlinear Transformation of the distance function in the nearest neighbor image recognition [C]// Computational Modeling of Objects Presented in Images, Fundamentals, Methods, and Applications. [S. l.]: Springer, 261 - 266.
- [5] LIAO Yong, DAI Xuewu, SHI Huanyu. An optimal channel allocation algorithm for SU-MIMO WLAN cognitive network [J]. Journal of Computers, 2015, 26(1): 1 - 17.
- [6] LIU Lijuan, SHEN Bo, WANG Xing. Research on kernel function of support vector machine [J]. Advanced Technologies, Embedded and Nuctimedia for Human-centric Computing, 2013: 827 - 834.
- [7] DATAL N, TRIGGS B. Histograms of oriented gradients for human detection [C]// 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, CA, USA: IEEE, 2005: 886 - 893.
- [8] CAO Hui, YAMAGUCHI K, NAITO T, et al. Pedestrian recognition using second-order HOG feature [C]// Asian Conference on Computer Vision. [S. l.]: Springer Berlin Heidelberg, 2009: 628 - 634.
- [9] HURTIK P, MADRID Bilinear interpolation over fuzzified images: enlargement [C]// IEEE International Conference on Fuzzy Systems. Istanbul, Turkey: IEEE, 2015: 3.
- [10] KOCHE K, PATIL V, CHAUDHARI K. Study of probabilistic neural network and feed forward back propagation neural network for identification of characters in license plate [J]. Advances in Power Electronics and Instrumentation Engineering, 2011, 148: 7 - 12.
- [11] WU Xiaoyang, LIU Yun, QI Dapeng, et al. A New method of subway tunnel crack image compression based on ROI and motion estimation [C]// 2014 Tenth International Conference on Intelligent Information Hiding and Multimedia Signal Processing. Kitakyushu, Japan: IEEE, 2014: 27 - 29.

[信息·简讯]

· 产品信息 ·

## 西门子发布支持持续软件开发的工程平台

西门子全新 TIA 博途 V16 工程平台拓展了众多实用新功能,全面覆盖从规划、工程到调试的各个阶段。针对开发过程的持续集成不断创新,包括标准化、跨团队集成工程以及集成功能测试,可帮助用户显著提高其软件质量,减少调试次数,降低工程成本。

对于开发过程而言,分布式团队协作是关键。现在,在 TIA 博途项目服务器上,用户可通过全新 Exclusive Engineering 功能以独占方式访问服务器项目,该功能具有项目修订和变更记录以及自动报告和归档等功能。这意味着在设备、对象或功能导向的基础上组织自动化任务时没有功能限制,而且使用 Exclusive Engineering 无需特殊许可。使用 TIA 博途 V16 扩展的“异步调试”模式,可通过后台运行的第 2 个 TIA 博途实例执行,将数据加载到 Simatic S7-1500 控制器,从而由团队对项目进行调试。而且还可立即再次运行前台 TIA 博途实例,从而大大缩短加载时间。此外,TIA 博途 V16 还提供有全新版本控制接口(VCI),用于诸如 GIT、SVN 和 TFS 等外部版本控制系统。通过 TIA 博途界面导入和导出软件对象,将版本控制无缝集成到了开发过程中,从而可对 TIA 博途外部的所有软件对象实现完全透明的对象粒度版本控制。除了可自动创建和检查编程准则的遵守情况(即 Styleguide Check)外,还可使用虚拟 S7-PLCSIM Advanced 控制器生成并执行应用程序测试,从而大幅减少工程和调试时间,提高软件质量。

通过 TIA 博途 V16,现在所有 Simatic S7-1200 控制器都具有了 OPC UA 服务器功能,可连接上位系统(如制造执行系统),进行垂直数据集成。而且,通过 OPC UA,还可实现所有 Simatic 控制器间的通信。使用 TIA 博途,可轻松导入标准化接口(配套规范),从而易于在生产线或工厂中集成机器设备。通过使用西门子 OPC UA 建模编辑器(SiOME)工具,自动化工程师们可实现过程变量互连,简化到 Simatic 机器控制系统中的导入。

(田实)