

[经营·管理]

DOI:10.3969/j.issn.1005-2895.2017.04.022

# 基于熵权法和马氏田口法的关键质量特性识别研究

马丽莎, 茅 健\*

(上海工程技术大学 机械工程学院, 上海 201620)

**摘 要:**针对机械产品制造检测过程中质量特性之间存在相关性以及检测效率低的问题,以实现高效、准确关键质量特性(key quality characteristics, KQCs)识别为目标,提出了基于熵权法(entropy weight method)和马氏田口法(Mahalanbis-Taguchi system method, MTS)的识别方法。在介绍产品设计阶段以及制造过程关键质量特性研究的基础上,分析了熵权法中熵值熵权的获取,以及马氏田口法中马氏距离和田口方法在质量特性识别中的应用,提出了两种模型结合识别机械产品关键质量特性的方法。以六轴五联动机床的皮带传动轴为例介绍了该方法的应用。结果表明该方法能有效处理质量特性之间的相关性问题,提高KQCs识别效率。

**关 键 词:**质量检测;关键质量特性;马氏田口法;熵权法;信噪比

中图分类号:TH165+.4 文献标志码:A 文章编号:1005-2895(2017)04-0101-05

## Research on Key Quality Characteristics Identification Based on Entropy Weight Method and Mahalanbis-Taguchi System

MA Lisha, MAO Jian\*

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

**Abstract:** Aiming at the problem of low detection efficiency and correlation between quality characteristics of mechanical manufacturing in the process of detection, a recognition method based on entropy weight method and MTS (Mahalanbis-Taguchi system method) was put forward under the target of achieving efficient and accurate identification. Based on the introduction of key quality characteristics in the product design stage and manufacturing process, the acquisition of entropy was analyzed. Then the application of Markov distance and Taguchi method of Mahalanbis-Taguchi system were proposed. The process of identifying the key quality characteristics of mechanical products with these two modes was presented. The application of this method was introduced by taking the belt drive shaft of six axis five linkage machine tool as an example. The result shows that the method can effectively solve the correlation between quality characteristics, and improve the efficiency of KQCs identification.

**Keywords:** quality inspection; key quality characteristics; Mahalanbis-Taguchi system method; entropy weight method; signal noise ratio

产品质量控制与企业生产成本、生产效率息息相关,企业为降低成本、提高效益需要重点控制关键的质量特性,以减少潜在的质量问题。质量特性根据重要度可以划分为关键、重要和一般质量特性<sup>[1]</sup>。关键质量特性识别有利于明确产品研制阶段质量保证的重点,更有利于理清产品设计、制造以及质量保证各过程

之间的关系,从而提高质量控制水平,降低产品研制成本。

在产品设计阶段,研究人员主要用质量功能展开法(quality function deployment, QFD)识别KQCs。黄铁群等<sup>[2]</sup>介绍了这一阶段工艺过程的质量特性模型。由于产品的复杂性以及设计过程与实际生产质检过程

收稿日期:2016-11-24;修回日期:2016-02-08

基金项目:国家自然科学基金(51175322)。

第一作者简介:马丽莎(1993),女,江苏东台人,硕士研究生,主要研究方向为质量控制。E-mail:2460169412@qq.com

的差异, 只从设计层面进行 KQCs 识别远不能达到要求。对于制造过程 KQCs 识别的研究, Boeing 公司最早基于质量损失的程度界定 KQCs。近年来国内较有代表性的研究有: 何益海等<sup>[3]</sup>提出了基于过程质量数据的制造系统可靠性建模分析方法; 张根保等<sup>[4]</sup>提出了基于人工神经网络(ANNs)技术的复杂机电产品 KQCs 提取模型和制造过程 KQCs 优化模型。

制造过程中的 KQCs 识别可以归结为对高维数据进行降维处理, 两种基本方法是特征选择<sup>[5]</sup>和特征提取<sup>[6]</sup>。针对这两种方法, 闫伟<sup>[7-8]</sup>引入 EM(expectation maximization)算法以及信息增益(information gain, IG)方法进行 KQCs 识别。谢荣琦等<sup>[9]</sup>引入数据挖掘领域中特征聚类方法, 得到相关性较强且相互之间冗余性较小的 KQCs 子集。何桢等<sup>[10]</sup>提出基于遗传模拟退火算法(genetic simulated annealing, GSA)的特征选择算法, 可以有效过滤无关、冗余质量特性。特征提取/选择本质上是个优化问题, 马氏田口法(MTS)是解决优化问题的重要工具, 而且具有预测/诊断功能。国外学者 Pal A 等<sup>[11]</sup>提出了马氏田口法和粒子群算法结合的混合优化算法, 指出了 MTS 的应用优势。Soylemezoglu A 等<sup>[12]</sup>提出了一种基于马氏田口法的故障检测方法, 并以滚动轴承内的一个数控主轴箱为例验证了其故障检测、隔离和预测的有效性。Iquebal A S 等<sup>[13]</sup>提出了一种基于马氏田口法和粗糙集的分类方法, 将粗糙集用于评价变量, 马氏田口法用于选择重要特征, 并以实例说明这种方法在分类精度上明显优于传统马氏田口法。国内学者韩亚娟等<sup>[14]</sup>对传统马氏田口方法进行了介绍, 针对多元系统中的强相关问题提出马氏田口广义逆矩阵法。伍建军等<sup>[15]</sup>提出基于田口方法的稳健优化设计方法, 能够选出最佳水平组合。

作者基于现有的识别方法, 针对制造阶段产品关键质量特性识别时各个指标之间存在相关性以及识别效率不高的问题, 提出了将熵权法和马氏田口法结合识别机械产品 KQCs 的方法。应用熵权法原理计算出各质量特性的熵权, 剔除熵权较小的指标, 实现第 1 轮降维; 计算马氏距离和信噪比, 剔除质量评价体系中贡献不大的指标, 进行第 2 轮降维; 最后, 判断优化后的样本马氏距离是否比原来小, 检验关键质量特性识别的正确性。

## 1 熵权法

### 1.1 信息熵基本定义

当产品有  $m$  个样本,  $n$  个质量特性指标时, 形成质

量特性矩阵  $D = (d_{ij})_{m \times n}$ , 其中  $d_{ij}$  表示第  $i$  个样本中第  $j$  个质量特性的数值。对于某个质量特性指标  $d_j$  有信息熵

$$e_j = - \sum_{i=1}^m P_{ij} \ln P_{ij} \quad (1)$$

式中  $P_{ij}$  为第  $i$  个样本中第  $j$  个特性指标值的比重。

质量特性指标的熵值越小, 表示不同样本中指标值变异程度越大, 提供的信息量越多, 其权重越大, 在质量评价中该特性指标起的作用越大。

### 1.2 熵权法计算权重

根据质量特性指标变化程度, 利用信息熵计算出各特性指标的熵权。 $n$  个质量特性指标在  $m$  个样本中形成的数据矩阵, 计算各个质量特性指标权重:

第  $i$  个样本中第  $j$  个特性指标值的比重

$$P_{ij} = d_{ij} / \sum_{i=1}^m d_{ij} \quad (2)$$

第  $j$  个质量特性指标的熵值

$$e_j = -k \sum_{i=1}^m P_{ij} \ln P_{ij}, k = 1/\ln m \quad (3)$$

第  $j$  个质量特性指标的熵权

$$w_j = (1 - e_j) / \sum_{j=1}^n (1 - e_j) \quad (4)$$

## 2 马氏田口法

马氏田口法结合了田口方法和统计学中的马氏距离, 通过优化后的指标对系统进行科学的诊断和评价。不仅真正实现数据降维, 而且能够诊断异常样本, 对质量特性进行宏观监控。

### 2.1 马氏距离

马氏距离是印度统计学家 Mahalanobis 提出的一种有效、可靠的样本数据相似度衡量方法, 多用于特征变量的综合评价, 计算每个样本与样本集重心的距离, 据此判别正常样本和异常样本。

2 组样本数据组成的列向量为  $x_i$  和  $x_j$ , 则数据之间的马氏距离可以定义为:

$$d_{ij} = \sqrt{(x_i - x_j)^T S^{-1} (x_i - x_j)} \quad (5)$$

式中  $S$  为向量  $x_i$  和  $x_j$  的协方差矩阵, 且

$$S = (x_{ij} - x_0)(x_{ij} - x_0)^T \quad (6)$$

式中:  $x_{ij} = (x_i, x_j)$ ;  $x_0 = 0.5[(x_i, x_j) + (x_j, x_i)]$ 。

在计算马氏距离的过程中, 若样本的维数小于样本数, 则样本协方差矩阵的逆矩阵不存在, 马氏距离不稳定。为了增强马氏距离的稳定性, 用广义逆矩阵对其进行改进, 可以表示为:

$$d_{ij} = \frac{1}{\lambda} \sqrt{(x_i - x_j)^T S^+ (x_i - x_j)} \quad (7)$$

式中: $S^+$ 表示协方差矩阵的广义逆矩阵, $\lambda$ 表协方差矩阵 $S$ 的最大特征值。

改进后的马氏距离避免了协方差矩阵的逆矩阵不可求的情况,不仅稳定性增强,而且简化了计算,抑制了无用信号的干扰。

## 2.2 田口方法

田口玄一于二十世纪七十年代初创立了田口方法(Taguchi methods),该方法以正交设计为基础,引入信噪比(signal noise ratio, SNR)降低误差对试验结果的影响,用尽可能少的试验次数获得尽可能准确的试验结果。田口方法的基本思想是把产品的稳健性融入到产品设计和制造过程中,使产品质量稳定,波动性小。

### 2.2.1 正交表

正交表是正交设计的主要工具,一般表示为 $L_n(t^c)$ ,这种特制的表格可以排出主次因素,确定最优方案。 $L$ 表示正交表, $n$ 表示试验的次数, $t$ 表示各因素的水平, $c$ 表示因素个数。用马氏田口法识别产品KQCs时,采用2水平正交表,依据正常样本中各质量特性的均值 $m_i$ 、标准差 $s_i$ 和相关矩阵 $C$ 生成相应的马氏空间,用以计算各个因素在每个水平的信噪比。

### 2.2.2 信噪比

信噪比是一种测量质量的尺度,通常用 $\eta$ 表示。 $\eta$ 值越大,产品质量越稳定,损失越小。信噪比计算方法分为望大型、望目型和望小型3类,在机械产品关键质量特性识别中选用望大型信噪比,计算公式如下:

$$\eta = -10 \lg \left( \frac{1}{N} \sum_{j=1}^N \frac{1}{d_{Mj}} \right) \quad (8)$$

式中 $d_{Mj}$ 表示第 $j$ 个噪声条件下质量特性的马氏距离。

## 3 关键质量特性识别方法

基于上述分析,本节结合马氏田口法和熵权法解决机械产品KQCs识别问题。步骤如下:

首先,输入采集到的质量特性数据,利用熵权法计算每个质量特性的熵值和熵权,将指标的熵权从大到小排序,设定合理的阈值,剔除熵权较小的质量特性,得到初次筛选结果;然后,构建质量特性测量表,计算每个样本的马氏距离,判别异常样本并将其剔除;再者,设计2水平正交表,分别计算质量特性在2个水平下的信噪比及其极差,将各个质量特性的信噪比极差进行排序,剔除质量评价体系中贡献率较小的特性指标;最后,确定产品的关键质量特性,计算优化后的样本马氏距离,验证识别的正确性。

关键质量特性识别方法的流程图如图1所示。

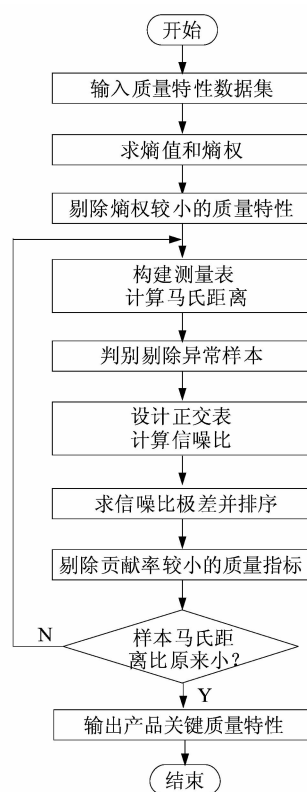


图1 关键质量特性识别流程图

Figure 1 Flow chart of key quality characteristic identification

## 4 实例分析

以六轴五联动机床中皮带传动轴为例,验证本识别方法的有效性。皮带传动是机床中重要的机械传动部分,其三维装配图以及二维设计图如图2~3所示。

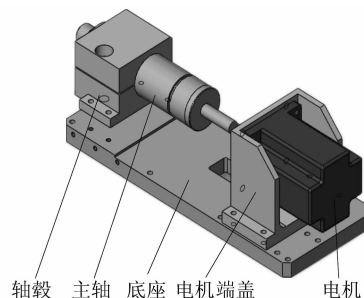


图2 皮带传动模型装配图

Figure 2 Assembly drawing of belt transmission model

从一批产品中选择30个样本分别编号1~30,分别检测30个样本的尺寸变量 $a_1 \sim a_6$ 、尺寸变量 $b_1 \sim b_7$ 、平面度变量 $c_1$ 、平行度变量 $c_2$ 、垂直度变量 $c_3$ 、同轴度变量 $c_4$ 以及粗糙度变量 $d_1 \sim d_4$ ,分析数据识别其制造过程中的关键质量特性,各变量含义如图3所示。

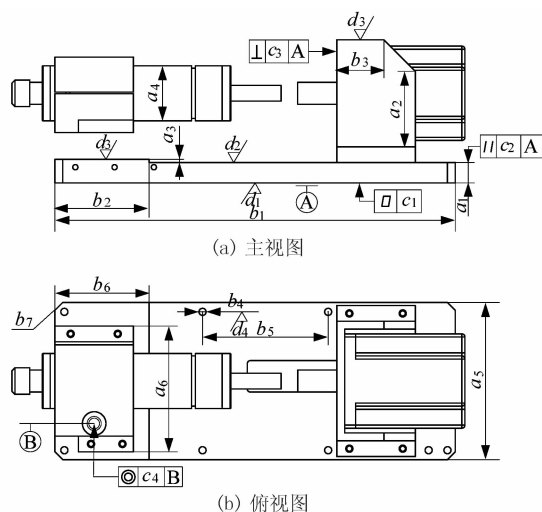


图3 皮带传动模型设计图

Figure 3 Design drawing of belt transmission model

#### 4.1 求解熵值和熵权

由公式(2)~(4)可得各质量特性的熵值和熵权,如表1所示。

表1 质量特性的熵值和熵权

Table 1 Entropy and entropy weight of quality characteristics

| 质量特性  | 熵值    | 熵权    | 质量特性  | 熵值    | 熵权    |
|-------|-------|-------|-------|-------|-------|
| $a_1$ | 0.853 | 0.059 | $b_6$ | 0.941 | 0.024 |
| $a_2$ | 0.492 | 0.204 | $b_7$ | 0.972 | 0.011 |
| $a_3$ | 0.912 | 0.035 | $c_1$ | 0.982 | 0.007 |
| $a_4$ | 0.581 | 0.168 | $c_2$ | 0.951 | 0.020 |
| $a_5$ | 0.928 | 0.029 | $c_3$ | 0.924 | 0.031 |
| $a_6$ | 0.498 | 0.201 | $c_4$ | 0.971 | 0.012 |
| $b_1$ | 0.936 | 0.026 | $d_1$ | 0.967 | 0.013 |
| $b_2$ | 0.956 | 0.018 | $d_2$ | 0.959 | 0.016 |
| $b_3$ | 0.925 | 0.030 | $d_3$ | 0.877 | 0.049 |
| $b_4$ | 0.976 | 0.009 | $d_4$ | 0.881 | 0.048 |
| $b_5$ | 0.953 | 0.019 |       |       |       |

将21个质量特性变量按熵权从大到小依次排列,结果为 $a_2 > a_6 > a_4 > a_1 > d_4 > a_3 > c_3 > b_3 > a_5 > b_1 > b_6 > c_2 > b_5 > b_2 > c_4 > d_3 > d_2 > d_1 > b_7 > b_4 > c_1$ 。其中 $c_1$ 和 $b_4$ 这2个质量特性熵权均小于0.01,远低于平均值。说明在质量特性综合评价中这2个质量特性作用较小,不同样本中的检测数据提供的信息较少。在实际生产过程中,相邻工序之间平行度 $c_1$ 和尺寸 $b_4$ 这2个质量特性没必要重复检测,它们不属于精密机床皮带传动部分制造过程中的关键质量特性。

#### 4.2 分析马氏距离

由公式(7)可得样本中其余19个质量特性的马

氏距离,如表2所示。

表2 样本马氏距离

Table 2 Markov distance of normal samples

| 样本 | 马氏距离  | 样本 | 马氏距离  | 样本 | 马氏距离  |
|----|-------|----|-------|----|-------|
| 1  | 4.247 | 11 | 3.593 | 21 | 4.959 |
| 2  | 4.316 | 12 | 4.341 | 22 | 4.668 |
| 3  | 3.879 | 13 | 4.312 | 23 | 4.599 |
| 4  | 4.670 | 14 | 4.472 | 24 | 4.081 |
| 5  | 4.098 | 15 | 5.235 | 25 | 4.185 |
| 6  | 3.490 | 16 | 3.897 | 26 | 3.759 |
| 7  | 5.311 | 17 | 3.987 | 27 | 4.286 |
| 8  | 4.164 | 18 | 3.390 | 28 | 4.818 |
| 9  | 3.800 | 19 | 3.901 | 29 | 4.392 |
| 10 | 4.292 | 20 | 5.426 | 30 | 3.763 |

划分样本类别时,具有相似性质的样本聚集为一类。识别关键质量特性过程中,规定马氏距离大于5的样本为异常样本,小于5的样本为正常样本。由表2中数据可以发现,样本7,15和20的马氏距离大于5,所以这3个样本为异常样本,将其剔除,避免异常值对整体评价的影响。

#### 4.3 分析信噪比

选用 $L_{32}(2^{31})$ 型正交表,计算每个质量特性2个水平的信噪比以及信噪比极差,结果如表3所示。

表3 信噪比分析

Table 3 Signal to noise ratio

| 质量特性  | 水平1 信噪比 | 水平2 信噪比 | 信噪比之差 | 排序 |
|-------|---------|---------|-------|----|
| $a_1$ | 10.935  | 9.203   | 1.732 | 1  |
| $a_2$ | 9.472   | 8.418   | 1.054 | 3  |
| $a_3$ | 9.122   | 9.156   | 0.034 | 17 |
| $a_4$ | 8.973   | 8.601   | 0.372 | 11 |
| $a_5$ | 9.962   | 9.394   | 0.568 | 8  |
| $a_6$ | 9.245   | 8.792   | 0.453 | 9  |
| $b_1$ | 9.640   | 9.253   | 0.387 | 10 |
| $b_2$ | 9.405   | 9.441   | 0.036 | 16 |
| $b_3$ | 10.124  | 9.488   | 0.636 | 7  |
| $b_5$ | 9.568   | 9.324   | 0.244 | 13 |
| $b_6$ | 10.680  | 10.711  | 0.031 | 18 |
| $b_7$ | 10.962  | 10.150  | 0.812 | 5  |
| $c_2$ | 8.473   | 7.786   | 0.687 | 6  |
| $c_3$ | 9.562   | 9.208   | 0.354 | 12 |
| $c_4$ | 8.986   | 8.770   | 0.216 | 14 |
| $d_1$ | 10.483  | 10.651  | 0.168 | 15 |
| $d_2$ | 9.598   | 8.171   | 1.427 | 2  |
| $d_3$ | 10.922  | 10.935  | 0.013 | 19 |
| $d_4$ | 7.965   | 7.139   | 0.826 | 4  |

#### 4.4 讨论

由表3可知, $a_3$ , $b_2$ , $b_6$ 和 $d_3$ 这4个质量特性的信

噪比极差均小于 0.1,在质量评价体系中贡献率较小,将其剔除。经过熵权法以及马氏田口法 2 次筛选后,共留下 15 个质量特性。对样本进行优化后,15 个质量特性的样本马氏距离均值变成 3.162,小于原先的 4.278,表明筛选出的关键质量特性是有效的。因此,最终确定在制造过程中产品关键质量特性为尺寸  $a_1$ 、 $a_2$ 、 $a_4$ 、 $a_5$ 、 $a_6$ ,尺寸  $b_1$ 、 $b_3$ 、 $b_4$ 、 $b_5$ 、 $b_7$ ,平行度  $c_2$ ,垂直度  $c_3$ ,同轴度  $c_4$ ,粗糙度  $d_1$ 、 $d_2$ 、 $d_4$ 。

## 5 结语

笔者基于熵权法和马氏田口法的应用,探索了一种识别产品关键质量特性的方法。用熵权法对各个质量特性进行赋权,能够得到较为客观的评价结果,避免了权重确定过程中的主观性,而且只需计算一次就能得到所有质量特性的权重,大大减少了评价过程的工作量。马氏距离可以判别异常样本,剔除异常样本能够减少异常值对评价体系的影响。用协方差矩阵的广义逆矩阵对马氏距离进行改进,抑制了无用信号的干扰,同时增强了马氏距离的稳定性。在田口方法中,用信噪比作为衡量产品质量稳健性的指标,确定产品性能最稳定的方案。应用文中提出的识别方法,可以提高产品关键质量特性识别效率,有效解决质量特性之间的相关性问题的同时能够优化、诊断多元系统,为进一步质量控制奠定了基础。

## 参考文献:

- [1] 庞继红. 复杂机电产品设计质量若干关键技术研究[D]. 重庆:重庆大学,2011:7-8.
- [2] 黄铁群,徐新胜. 基于质量功能展开的工艺规程主模板评价[J]. 计算机集成制造系统,2010,16(6):1183-1188.
- [3] 何益海,沈珍,尹超,等. 基于过程质量数据的制造系统可靠性建模分析[J]. 北京航空航天大学学报,2014,40(8):1027-1032.
- [4] 张根保,纪富义,任显林,等. 复杂机电产品关键质量特性提取模型[J]. 重庆大学学报(自然科学版),2010,32(2):8-14.
- [5] 姚旭,王晓丹,张玉玺,等. 特征选择方法综述[J]. 控制与决策,2012,27(2):161-166.
- [6] 柴瑞敏,曹振基. 基于 Gabor 小波与深度信念网络的人脸识别方法[J]. 计算机应用,2014,34(9):2590-2594.
- [7] 闫伟,何桢,田文萌,等. 基于 EM 的不平衡数据关键质量特性识别[J]. 工业工程与管理,2012,17(4):38-42.
- [8] 闫伟,何桢,田文萌,等. 基于 IG 的复杂产品关键质量特性识别[J]. 工业工程与管理,2012,17(1):70-74.
- [9] 谢荣琦,何桢,何曙光. 基于 ReliefF 和 k-modes 聚类的复杂产品关键质量特性识别[J]. 工程工业与管理,2014,19(1):30-34.
- [10] 李岸达,何桢,何曙光,等. 基于 GSA 的复杂产品关键质量特性识别[J]. 系统工程与电子技术,2015,37(9):2073-2079.
- [11] PAL A, MAITI J. Development of a hybrid methodology for dimensionality reduction in Mahalanobis-Taguchi system using Mahalanobis distance and binary particle swarm optimization[J]. Expert systems with applications,2010,37(2):1286-1293.
- [12] SOYLEMEZOGLU A, JAGANNATHAN S, SAYGIN C. Mahalanobis-Taguchi System (MTS) as a prognostics tool for rolling element bearing failures[J]. Journal of manufacturing science and engineering,2010,132(5):1014-1026.
- [13] IQUEBAL A S, PAL A, CEGLAREK D, et al. Enhancement of Mahalanobis-Taguchi System via rough sets based feature selection[J]. Expert systems with applications,2014,41(17):8003-8015.
- [14] 韩亚娟,何桢,宋国防,等. 多维系统优化中马氏田口强相关问题研究[J]. 工业工程,2012,15(2):71-77.
- [15] 伍建军,万良琪,吴事浪,等. 基于田口方法的柔性铰链柔度稳健优化设计[J]. 工程设计学报,2015,22(3):224-229.

## [信息·简讯]

### · 产品信息 ·

## 西门子推出高性价比基本型工控机

西门子近期再次扩展其工控机系列,推出了 Simatic IPC327E 箱式 PC 和 Simatic IPC377E 面板式 PC。这 2 款无风扇基本型工控机针对诸多工业自动化应用进行了优化,性价比更高,尤其适用于 HMI(人机界面)应用、组装工作站、仓储物流、数据采集(如生产和能源数据),以及生产网络和简单的图像处理。与已可供货的 Simatic IPC347E 机架式 PC 一起,专门设计用于机器制造商、分销商和系统集成商。该系列工控机免维护,可连续 24 h 运行,采用最新 PC 技术,配有多种实用接口。目前,可现货交付 8 款预配置型号:带 12、15 或 19 英寸单点触摸式宽屏显示器的面板式 PC 和箱式 PC。用户可选择预装 Windows 7 Ultimate 操作系统。

如果对工业环境中的证书、温度范围或保护等级无特定要求,使用全新 Simatic IPC327E 和 IPC377E 基本型工控机,可快速实施相关应用。与所有 Simatic 工控机一样,均适用于工业环境,提供有丰富选件,部件之间可完美协同,无风扇设计,配有密封机箱,坚固可靠,使用寿命长。采用最先进 PC 技术,配有多种接口,可轻松集成到新的和现有自动化解决方案中。均配有 1 个 Intel Celeron 4 核处理器、1 个 VGA 和 DisplayPort 口、1 个 PCIe 迷你接口、2 个千兆以太网接口和多达 6 个 COM 口和 USB 3.0 接口。

(朱建芸)