Journal of Manufacturing Systems 32 (2013) 550–555

Contents lists available at ScienceDirect

Journal of Manufacturing Systems

journal homepage: www.elsevier.com/locate/jmansys

Technical paper

# Feature selection for manufacturing process monitoring using cross-validation

Chenhui Shao [a,*], Kamran Paynabar [b], Tae Hyung Kim [a], Jionghua (Judy) Jin [c], S. Jack Hu [a], J. Patrick Spicer [d], Hui Wang [d], Jeffrey A. Abell [d]

[a] Department of Mechanical Engineering, The University of Michigan, Ann Arbor, MI, USA
[b] H. Milton Stewart School of Industrial and Systems Engineering, The Georgia Institute of Technology, Atlanta, GA, USA
[c] Department of Industrial and Operations Engineering, The University of Michigan, Ann Arbor, MI, USA
[d] Manufacturing Systems Research Lab, General Motors Technical Center, Warren, MI, USA

ARTICLE INFO

*Article history:*
Received 27 March 2013
Accepted 15 May 2013
Available online 25 June 2013

*Keywords:*
Feature selection
Parameter tuning
Cross-validation
SPC monitoring
Ultrasonic metal welding

ABSTRACT

A novel algorithm is developed for feature selection and parameter tuning in quality monitoring of manufacturing processes using cross-validation. Due to the recent development in sensing technology, many on-line signals are collected for manufacturing process monitoring and feature extraction is then performed to extract critical features related to product/process quality. However, lack of precise process knowledge may result in many irrelevant or redundant features. Therefore, a systematic procedure is needed to select a parsimonious set of features which provide sufficient information for process monitoring. In this study, a new method for selecting features and tuning SPC limits is proposed by applying $k$-fold cross-validation to simultaneously select important features and set the monitoring limits using Type I and Type II errors obtained from cross-validation. The monitoring performance for production data collected from ultrasonic metal welding of batteries demonstrates that the proposed algorithm is able to select the most efficient features and control limits and thus leading to satisfactory monitoring performance.



## 1. Introduction

On-line process monitoring is crucial for product quality and process stability in manufacturing [1]. For example, in electric vehicle battery manufacturing, quality monitoring for battery joining is of great importance because any low-quality joints may result in a failure of the entire battery pack, causing high production loss. Thus, on-line process monitoring has received great attention over the past several decades.

Among various monitoring methods, the classical statistical process control (SPC) method has been widely used in monitoring manufacturing processes [2]. Control charts are the main SPC tools to determine whether a manufacturing process is in a state of statistical control. Two of the most popular types of control charts are the univariate Shewhart control chart and Hotelling $T^2$ control chart [3]. Exemplary applications of control charts in manufacturing process monitoring can be found in [4,5].

In order to monitor manufacturing processes, various sensor signals, such as force, acceleration, temperature, pressure and acoustic emission, are collected on-line to gather process information. Due to the large volume of data, feature extraction is often carried out to reduce the dimensionality of data. Efficient application-dependent features are constructed when expert knowledge about manufacturing processes is available. Whereas, if a lack of expert knowledge is encountered, some general data-driven dimensionality reduction techniques can help. Examples of such techniques include Principal Component Analysis (PCA) [6], kernel PCA [7], semidefinite embedding [8], and wavelets analysis [9].

In manufacturing, when a new process is initially implemented for production, it often occurs that a thorough physical understanding of the process is not available. For example, ultrasonic metal welding is recently utilized to join lithium-ion batteries, but there is insufficient expert knowledge about this process. Thus, signal features without good physical understanding may be irrelevant or redundant. Under this circumstance, feature selection is commonly applied to pick a minimally sized subset of features for monitoring. By removing a large number of irrelevant and redundant features, feature selection is able to help avoid overfitting, improve model performance, provide more efficient and cost-effective process monitoring, and acquire better insights into the underlying processes that generated the data.

Generally speaking, feature selection techniques can be divided into three categories in terms of means of combining feature subset selection search with the classification model construction: filter methods, wrapper methods and embedded methods [10]. Filter techniques determine the relevance of features by looking only at the intrinsic properties of the data. In wrapper methods, the model hypothesis search is embedded within the feature subset

* Corresponding author. Tel.: +1 734 764 9103.
*E-mail address:* chshao@umich.edu (C. Shao).

0278-6125/$ – see front matter © 2013 The Society of Manufacturing Engineers. Published by Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jmsy.2013.05.006

search. Embedded techniques build the feature subset search into the classifier construction. A summary of the advantages and disadvantages of each type of method and some examples of these methods can be found in [10].

In this study, a new feature selection algorithm based on cross-validation is developed for quality monitoring of manufacturing processes. The method belongs to the category of wrapper methods. Cross-validation is a common statistical technique for evaluating how the results of a statistical analysis will generalize to an independent data set [11]. It is mainly used to evaluate how accurately a predictive model will perform independent of the training dataset. In this paper, cross-validation is applied to selecting significant features and setting monitoring limits simultaneously, based on Type I ($\alpha$) and Type II ($\beta$) error rates calculated from validation tests.

The rest of this paper is organized as follows. Section 2 presents the details of the proposed feature selection algorithm. In Section 3, the proposed scheme is applied for feature selection and control limits tuning for monitoring of ultrasonic metal welding in battery assembly processes. Finally, Section 4 concludes the paper.

## 2. Feature selection and parameter tuning based on cross-validation

In the proposed feature selection and parameter tuning algorithm, we adopt the stepwise forward feature selection to select the optimal feature subset from candidate features. Forward selection is a greedy search strategy and is particularly computationally advantageous and robust against overfitting [12]. In some cases, this search strategy may alleviate the problem of overfitting, as illustrated in [13]. Forward selection was first utilized in [14] for measurement/feature selection to determine the best subset of measurements/features for pattern classification, and it is still widely used as a feature selection scheme [15]. Forward feature selection starts the search with an empty feature subset. First, all the features are considered for possible selection, and the one feature that performs the classification the best is included in a subset. Then a new step is started, and the remaining features are considered for inclusion. This is repeated until a prespecified number of features have been included in the subset. Usually the search is repeated until all features are included for comparison purpose.

Cross-validation is a statistical technique for evaluating and comparing learning algorithms by partitioning data into two sets: one used for model training and the other used for model validation. This method is applicable for the performance comparison of different predictive modeling procedures [16], as well as for variable selection [17].

In this study, the $k$-fold cross-validation is employed for simultaneous feature selection and SPC parameter tuning. The original sample is randomly partitioned into $k$ mutually exclusive subsamples/folds of equal (or approximately equal) size. Then $k$ iterations of training and validation are performed such that within each iteration one different subsample is held-out for validation while the remaining $k-1$ subsamples are used for training. After the $k$ iterations are finished, the $k$ results can be averaged (or otherwise combined) to give a single estimation. In this method, all observations are used for both training and validation, and each observation is used for validation exactly once. In practice, 10-fold cross-validation is widely used.

In the algorithm, candidate features are denoted by $f_1, f_2, \ldots, f_N$, and the total number of features is $N$. The percentile limits are used as control limits. It is assumed that the total number of candidate percentile limit sets is $M$, and the $m$th set is denoted by $\mathbf{P}(m)$, where $m = 1, 2, \ldots, M$. Each percentile limit set includes a lower limit and an upper limit, namely,

$$\mathbf{P}(m) = [\, p_{ml} \quad p_{mu} \,], \tag{1}$$

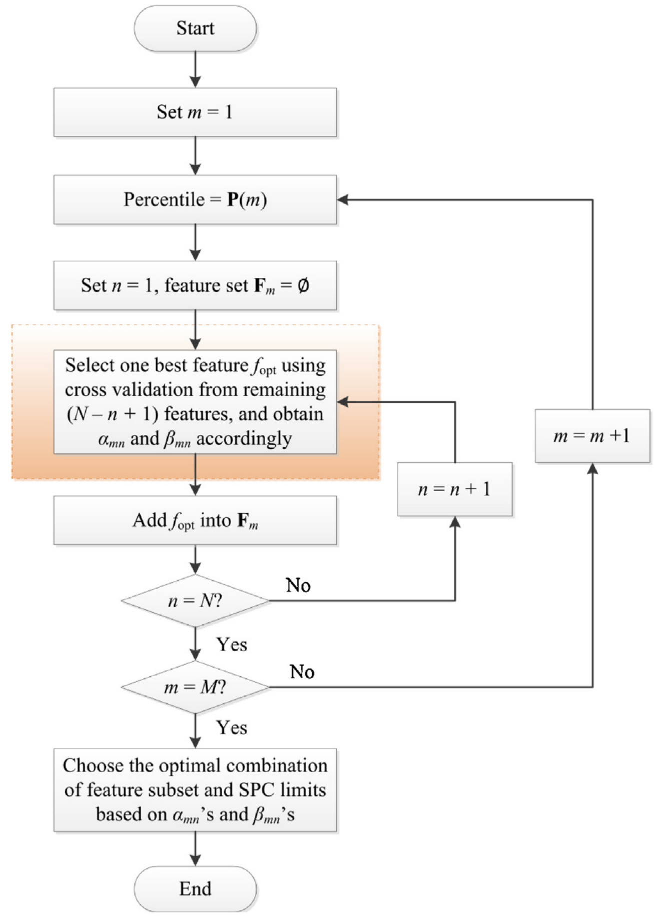


**Fig. 1.** Feature selection and SPC limits tuning.

where $p_{ml}$ and $p_{mu}$ are lower and upper percentile limits, respectively.

Fig. 1 shows the proposed algorithm for forward feature selection and SPC limits tuning, and each forward feature selection step is performed using cross-validation. Fig. 2 illustrates how to use cross-validation to select the $n$th feature from remaining $N-n+1$ features in forward feature selection for the $m$th percentile limit set.

For each set of percentile limits $\mathbf{P}(m)$, we perform forward feature selection using cross-validation, as illustrated in Fig. 2. The forward selection criterion is given by

$$\min R_{mn} = A\alpha_{mn} + B\beta_{mn}, \tag{2}$$

where $m = 1, 2, \ldots, M$; $n = 1, 2, \ldots, N$; $A$ and $B$ are penalty coefficients for $\alpha$ error rate and $\beta$ error rate, and they can be tuned according to different monitoring schemes. For example, if $\beta$ error rate is of higher concern, and then $B$ can be set higher correspondingly.

For each limit set, an arrangement of candidate features is obtained, as given by Eq. (3).

$$\mathbf{F}(m) = [f_{m(1)}, f_{m(2)}, \cdots, f_{m(N)}]. \tag{3}$$

Meanwhile, corresponding $\alpha$ error rates as well as $\beta$ error rates are also calculated stepwise, and we record them in vectors, as shown by Eqs. (4) and (5).

$$[\alpha_{m1}, \alpha_{m2}, \cdots, \alpha_{mN}]. \tag{4}$$

$$[\beta_{m1}, \beta_{m2}, \cdots, \beta_{mN}]. \tag{5}$$

After performing forward feature selection for all percentile limit sets, we can select from 1 to $N$ features for each set, and