Applied Soft Computing 14 (2014) 4–11

Contents lists available at ScienceDirect

Applied Soft Computing

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

Fault detection and identification spanning multiple processes by integrating PCA with neural network

Jing Zhou [a,b,*], Aihuang Guo [b], Branko Celler [c], Steven Su [a]

[a] Faculty of Engineering and Information Technology, University of Technology, Sydney, Australia
[b] College of Electronics and Information Engineering, Tongji University, China
[c] CSIRO ICT Centre, Australia

ARTICLE INFO

Article history:
Received 31 March 2013
Received in revised form 27 August 2013
Accepted 24 September 2013
Available online 23 October 2013

Keywords:
BP neural network
Exercising monitoring
Fault detection and identification
Principal component analysis

ABSTRACT

This paper proposes an effective fault detection and identification method for systems which perform in multiple processes. One such type of system investigated in this paper is COSMED K4b^2. K4b^2 is a standard portable electrical device designed to test pulmonary functions in various applications, such as athlete training, sports medicine and health monitoring. However, its actual sensor outputs and received data may be disturbed by Electromagnetic Interference (EMI), body artifacts, and device malfunctions/faults, which might cause misinterpretations of activities or statuses to people being monitored. Although some research is reported to detect faults in specific steady state, normal approach may yield false alarms in multi-processes applications. In this paper, a novel and comprehensive method, which merges statistical analysis and intelligent computational model, is proposed to detect and identify faults of K4b^2 during exercise monitoring. Firstly the principal component analysis (PCA) is utilized to acquire main features of measured data and then *K*-means is combined to cluster various processes for abnormalities detection. When faults are detected, a back propagation (BP) neural network is constructed to identify and isolate faults. The effectiveness and feasibility of the proposed model method is finally verified with experimental data.

1. Introduction

Cardio-respiratory parameters, such as oxygen consumption rate, carbon dioxide production rate and heart beat are major indicators for the evaluation of human's physical status and sports performances [1–5]. These parameters are traditionally tested by indirect calorimeter with a metabolic cart. However, most metabolic carts are restricted in laboratories due to their size and weight limitations. Although many activities can be simulated in the laboratory (e.g. walking on a treadmill), some activities involving in occupational, recreational activities and health monitoring are not available in the laboratory. Using heavy equipments to collect expired air in fields can often disturb activities under investigation. Because of the improvement of miniaturized metabolic measurement systems oxygen consumption can be measured in real circumstances by small gadgets. Recently lightweight, portable telemetric gas analysis systems are used to acquire parameters in daily activities or sports for further analysis.

Among several gas analysis devices COSMED K4b^2 metabolic measurement system is a typical and popular instrument. It is designed to measure parameters of ventilatory, oxygen consumption and carbon dioxide production with several sensors such as flow meter, oxygen sensor, carbon dioxide sensor, environmental sensor and so on [6–9]. This system is one of the latest portable devices for cardiopulmonary gas exchange analysis base on true breath-by-breath without limitation. Along with the widely usage of COSMED K4b^2, the reliability is very crucial for measurement in metabolic testing or health monitoring. R Duffield assessed the validity and reliability of a COSMED K4b^2 portable telemetric gas analysis system by experiments and reliability is formed in specific steady state and sustained exercise [10]. However, in free living environment, the sensor outputs and data received by base station may be abnormal, which can be caused by impulsive noise, low batteries or environment interference. These abnormal data might cause misinterpretations of exercises or living activities and lead to unreliable results. Although some methods are available to detect faults, normal approach may yield false alarms for multi-processes applications especially for wireless transmission. Therefore, an effective and feasible method is necessary to detect abnormal situation and identify faults in multi-statuses for monitoring.

* Corresponding author at: Faculty of Engineering and Information Technology, University of Technology, Sydney, Australia. Tel.: +61 295142376.
E-mail address: Jing.Zhou-2@student.uts.edu.au (J. Zhou).

1568-4946/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.asoc.2013.09.024

In this paper, a novel fault detection and identification method is proposed for systems performing in multiple processes by optimally integrating principal component analysis (PCA) [11], *K*-means [12] and neural networks [13]. Firstly, the PCA is used to detect the potential faults. At the same time, it decreases the dimension of the data set so that the energy consumption in transmission is consequently reduced. Thereafter, in the reduced data space, *K*-means is applied to cluster data into different groups and detect faults with statistic criteria. Then a neural network is constructed and trained to identify faulty by reconstruction. Based on real data experiments and simulations it is demonstrated that the proposed fault detection and identification (FDI) method is able to correctly detect and identify faults.

The main contributions of this paper exist in three aspects: (1) It applies the dimension reducing technique to decrease energy consumption in transmission which is very important in wireless body sensors networks; (2) With clustering it can correctly detect faults in different processes which is usually ignored in traditional model; (3) This model works beyond fault detection and it can identify crashed sensors according to measured data.

The remaining sections of this paper are presented as follows: Section 2 introduces the portable medical device: COSMED K4b^2. Section 3 explains the principles of PCA and *K*-means for clustering and fault detection. In Section 4 the back propagation (BP) neural networks for prediction is outlined. Section 5 verifies the FDI model with demonstration of experimental results. Section 6 makes a conclusion.

2. Overview of K4b^2 system

A K4b^2 system is a COSMED portable medical instrument used for testing of pulmonary functions. It can be worn by people during activities and is capable of delivering real-time measurements into a PC base station. Due to its convenience, it is applied in many fields such as: sports medicine research, gait lab, occupational health, cardiology, cardiac rehabilitation, clinical nutrition and so on. It can measure physiological response to exercise in the field without limitation. This telemetric gas analysis system contains a soft, flexible face mask to sample expired air and a sensor unit to test ventilation (VE), oxygen (O_2) and carbon dioxide (CO_2) concentrations in the expired air [14]. These measurements represent energy cost of activities and are highly correlated. So it is available to extract main principal components for reducing transmission amount. The meaning of measured symbols of K4b^2 in breath by breath exercise testing are showed in Table 1 [6].

Table 1
Measured parameters in K4b^2.

Symbol	UM	Parameter
VO_2	Ml/min	Oxygen uptake
VCO_2	Ml/min	Carbon dioxide production
Vt	1	Tidal volume
$FetO_2$	%	End tidal O_2
$FetCO_2$	%	End tidal CO_2
R	–	Respiratory quotient
VE	l/min	Ventilation
HR	l/min	Heart rate
Qt	l	Cardiac output
AT	–	Anaerobic threshold
VE	l/min	Ventilation
SV	l/min	Stroke volume
RF	l/min	Respiratory frequency
FeO_2, $FeCO_2$	%	Averaged expiratory concentration of O_2 or CO_2
VE/VO_2	–	Ventilatory equivalent for O_2
VE/VCO_2	–	Ventilatory equivalent for CO_2
VO_2/HR	Ml/beat	Oxygen pulse
VO_2/kg	ml/min/kg	VO_2 per kg
Ti, Te, Ti/Ttot	s	Time breaths
Vd/Vt	–	Vd/Vt ratio
$PaCo_2$	mmHg	Arterial PCO_2 (estimated)
P(a-et)CO_2	mmHg	Delta PaCO-PetCO_2

3. Principles of PCA and *K*-means

3.1. Principle of PCA

PCA is often used to transform multivariable space into a subspace which preserves maximum variance of the original space in minimum dimensions [15]. The measured process variables are usually correlated to each other. The measured process variables are usually correlated to each other and data can be disposed into the significant patterns. Then the abnormalities, such as noises or outliers in residual subspace can be identified. PCA transforms the original correlated data into a new set of uncorrelated data set that represent the trend of the process. It is highly useful in analyzing state data which contains relationships between variables. It is proved successfully in many applications such as reducing dimensionality, data compression, and fault detection [16–20].

In normal condition, the PCA is established with a collected $n \times m$ data matrix, where n is the number of samples and m is the number of variables. This matrix must be standardized as $\overline{X}$ to eliminate different units' effects of variables. Then the covariance matrix is constructed as Eq. (1):

$$R = \frac{1}{n-1}\overline{X}^T\,\overline{X} \tag{1}$$

And then perform the SVD (Single Variable decomposition) on R:

$$R = UD_\lambda\ U^T \tag{2}$$

where $U_{m\times m}$ is a unitary matrix, and $D_\lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is a diagonal matrix contains the eigenvalues of R sorted in decreasing order ($\lambda_1 > \lambda_2 > \ldots > \lambda_n$). The transformation matrix $P \in R^{n\times k}$ is generated by choosing k eigenvectors. Matrix P transforms the space of measured variables into reduced dimension space. Elements of T, called as scores, are calculated by columns of matrix P.

$$T = XP \tag{3}$$

Scores are values of original measured variables that have been transformed into the reduced dimension space.

The subspace, which is formed with the first k ($k < n$) vectors without correlation, is called principal component subspace $\hat{S} = [u_1, u_2, \ldots, u_k]$. Another subspace, which is formed with the n–k vectors $\tilde{S} = [u_{k+1}, u_{k+2}, \ldots, u_n]$ are called residual subspace $\tilde{S}$. So the database X with m dimensions is replaced by the principal subspace $\tilde{S}$ with k dimension and the residual subspace $\tilde{S}$ with n–k dimension, where $\hat{X} = TPT$, $E = X - \hat{X}$. At last raw data space X can be calculated as:

$$X = TPT + E \tag{4}$$

where TPT represents the principal elements of variability in the process and E stands for the variability related to the process noise. It is critical to select the number of principal components. The popular method for choosing components is Cumulative Percent Variance (CPV) approach. It captures the $CPV(i) \geq 90\%$ by the first k principal components:

$$CPV(i) = \sum_{i=1}^{k}\lambda_i\left(\sum_{j=1}^{n}\lambda_j\right)^{-1} \tag{5}$$