

**Simultaneous Fault Diagnosis of Automotive
Engine Ignition Systems using Pairwise Coupled
Relevance Vector Machine, Extracted Pattern
Features and Decision Threshold Optimization**

by

Zhang Zaiyong



MSc in Electromechanical Engineering

2011



**Faculty of Science and Technology
University of Macau**

Simultaneous Fault Diagnosis of Automotive Engine
Ignition Systems using Pairwise Coupled Relevance Vector
Machine, Extracted Pattern Features and Decision Threshold
Optimization

by

Zhang Zaiyong

M-A9-6593-0

A thesis submitted in partial fulfillment of the
requirements for the degree of

MSc in Electromechanical Engineering

Faculty of Science and Technology
University of Macau

2011

Approved by _____
Supervisor

Program Authorized
to Offer Degree _____

Date _____

In presenting this thesis in partial fulfillment of the requirements for a Master's degree at the University of Macau, I agree that the Library and the Faculty of Science and Technology shall make its copies freely available for inspection. However, reproduction of this thesis for any purposes or by any means shall not be allowed without my written permission. Authorization is sought by contacting the author at

Address: 3W, Block Hoi Wan, Hoi Wan Garden, Paipa, Macau

Telephone: (853)62112738

E-mail: zhangqibiao177@hotmail.com

Signature _____

Date _____



University of Macau

Abstract

Simultaneous Fault Diagnosis of Automotive Engine
Ignition Systems using Pairwise Coupled Relevance Vector
Machine, Extracted Pattern Features and Decision Threshold
Optimization

by Zhang Zaiyong

M-A9-6593-0

Thesis Supervisor: Prof. Wong Pak Kin

Department of Electromechanical Engineering, Faculty of Science and Technology

Whenever there is any fault in an automotive engine ignition system or change of engine conditions, a mechanic can analyze the engine ignition patterns to identify the engine fault according to both the specific domain knowledge and the shape features of the patterns. One of the challenges in ignition system diagnosis is that more than one fault may appear at a time. This kind of problem refers to simultaneous fault diagnosis. Another challenge is the acquisition of large amount of costly simultaneous-fault ignition patterns for constructing the diagnostic system because the number of the training patterns depends on the combination of different single faults. The above problems could be resolved by the proposed framework combining feature extraction, probabilistic classification, and decision threshold optimization. For feature extraction, the thesis proposes to use wavelet packet transform and principal component analysis together with domain knowledge. With the proposed framework, the features of the single faults in a simultaneous-fault pattern are extracted and then detected using a new probabilistic classifier namely pairwise coupling relevance vector machine, which is trained with single-fault patterns only. Therefore, the training dataset of simultaneous-fault patterns is not necessary. Experimental results show that the proposed

framework performs well for both single-fault and simultaneous-fault diagnoses, and is superior to the existing approach.



TABLE OF CONTENTS

LIST OF FIGURES	III
LIST OF TABLES	IV
LIST OF ABBREVIATIONS	V
ACKNOWLEDGMENTS	IX
CHAPTER 1 INTRODUCTION	1
1.1 BACKGROUND OF AUTOMOTIVE ENGINE IGNITION PATTERNS	1
1.2 FEATURE EXTRACTION METHODS	2
1.3 CLASSIFICATION METHODS	3
1.4 DECISION THRESHOLD OPTIMIZATION	5
1.5 RESEARCH OBJECTIVES AND THE PROPOSED FRAMEWORK	6
CHAPTER 2 PROPOSED FRAMEWORK AND RELATED TECHNIQUES	8
2.1 FORMULATION OF THE PROPOSED FRAMEWORK	8
2.2 EXTRACTION OF DOMAIN KNOWLEDGE FEATURES FOR IGNITION PATTERNS	10
2.3 FEATURE EXTRACTION USING WPT AND PCA AND COMBINED FEATURE VECTOR	12
2.4 RELEVANCE VECTOR MACHINE	13
2.5 PAIRWISE COUPLED RELEVANCE VECTOR MACHINE	15
2.6 DECISION THRESHOLD OPTIMIZATION AND F-MEASURE	17
2.7 SUMMARY OF PROPOSED FRAMEWORK AND TECHNIQUES	19
CHAPTER 3 EXPERIMENTAL SETUP	22
3.1 DATA SAMPLING	22
3.2 DATA NORMALIZATION	24

3.3 ALLOCATION OF DATASETS	25
CHAPTER 4 EXPERIMENTAL RESULTS	29
4.1 RESULTS OF VARIOUS COMBINATIONS OF FEATURE EXTRACTION AND CLASSIFICATION TECHNIQUES.....	29
4.2 RESULTS WITH THRESHOLD OPTIMIZATION.....	31
4.3 INDIVIDUAL RESULT OF SINGLE AND SIMULTANEOUS FAULT DIAGNOSIS.....	34
4.4 RESULTS COMPARISON WITH THE LATEST TECHNIQUE.....	36
CHAPTER 5 DISCUSSION OF RESULTS	37
5.1 EFFECT OF FEATURE EXTRACTION AND PAIRWISE PROBABILISTIC CLASSIFICATION.....	37
5.2 EFFECT OF DECISION THRESHOLD OPTIMIZATION.....	38
5.3 DIAGNOSIS OF SIMULTANEOUS FAULTS	38
5.4 COMPARISON WITH THE LATEST APPROACH.....	39
CHAPTER 6 CONCLUSIONS AND RECOMMENDATIONS	40
6.1 CONCLUSIONS	40
6.2 ORIGINALITIES AND CONTRIBUTIONS OF RESEARCH	42
6.3 RECOMMENDATIONS FOR FUTURE WORK.....	43
REFERENCES	44
PUBLICATIONS.....	48

LIST OF FIGURES

<i>Number</i>	<i>Page</i>
Figure 1 Proposed framework of the simultaneous engine ignition fault-diagnosis system and its evaluation	9
Figure 2 Decision function based on threshold ε	10
Figure 3 Key domain knowledge features of the normal engine ignition signal	11
Figure 4 Procedure for calculating the burn time (F_2)	12
Figure 5 Indecisive regions (shaded area) using 1vA (left) and pairwise coupling (1v1) (right)	16
Figure 6 Architecture of 1vA classifier (left) and 1v1 classifier (right)	16
Figure 7 Workflow of feature extraction, decision threshold optimization, diagnostic system training and testing	21
Figure 8 Sample ignition patterns and their corresponding single engine faults	26
Figure 9 Sample ignition patterns and their corresponding simultaneous engine faults	27
Figure 10 Collection of ignition patterns from a test engine using a computer-linked automotive scope meter	28

LIST OF TABLES

<i>Number</i>	<i>Page</i>
Table 1 Algorithm of the proposed framework for simultaneous-fault diagnosis of time-dependent ignition patterns	19
Table 2 Sample single faults of engine trouble reflected by ignition patterns.....	23
Table 3 Sample possible simultaneous faults of engine trouble reflected by ignition patterns	23
Table 4 Evaluation of different combinations of techniques	31
Table 5 GA operators and parameters	32
Table 6 PSO operators and parameters.....	32
Table 7 Evaluation of different combinations of techniques using GA-optimized threshold .	33
Table 8 Evaluation of different combinations of techniques using PSO-optimized threshold	33
Table 9 Single fault diagnosis under F-measure (With DK+WPT+PCA, $\varepsilon_{opt} = 0.7147$)	34
Table 10 Simultaneous fault diagnosis under F-measure with classifiers trained with single faults (With DK+WPT+PCA and $\varepsilon_{opt} = 0.7147$)	35
Table 11 Evaluation results of the binarization approach and the proposed framework for fault diagnosis	36

LIST OF ABBREVIATIONS

A: diagonal matrix of hyperparameters

a : end point of burning time

B: diagonal matrix in RVM

$Bf_j(\cdot)$: j th binary classifier

C_i : i th probabilistic classifier

$C_i(\mathbf{f})$: probability of \mathbf{f} belonging to the i th label

C_{ij} : pairwise classifier

$C_{ij}(\mathbf{f})$: pairwise probability of \mathbf{f} belonging to the i th label against the j th label

c_1 : cognitive parameter of PSO

c_2 : social parameter of PSO

D: sample dataset

d : number of labels (faults)

e_j : j th eigen value

e_j' : j th normalized eigen value

F: set of feature vectors

F': set of feature vectors created by WPT and PCA

F_1 : firing voltage

F_2 : burn time

F_3 : average spark voltage of spark line

F_{me} : F-measure

f: feature vector

f': feature vector created by WPT and PCA

\mathbf{f}_h : h th feature vector

f_{class} : probabilistic classifier

\mathbf{H} : PCA transformation matrix

\mathbf{h}_j : j th eigen vector

J : decomposition level of WPT

$K(\cdot)$: kernel function in RVM

L_P : length of ignition pattern (i.e. number of data point in ignition pattern)

\mathbf{l} : true label vector

\mathbf{l}_i : i th true label vector

l_i : i th label in \mathbf{l}

l_{ig} : g th label in \mathbf{l}_i

l_i^j : i th label in the j th test data

N : number of training data

N_D : number of cases in sample dataset

N_T : number of test data

n_{ij} : number of training data with the i^{th} and j^{th} labels

$P(l_j|\mathbf{f})$: probability of \mathbf{f} belonging to l_j

$P(\cdot)$: probability

s : input dimension of classifier to be evaluated

\mathbf{t} : set of faulty labels in training dataset

t_n : faulty label of the n th training case

$TEST$: original test dataset

$TEST_1$: single fault cases in test dataset

$TEST_s$: simultaneous fault cases in test dataset

$TEST_F$: test dataset after feature extraction

$TRAIN$: original training dataset

$TRAIN_F$: training dataset after feature extraction

\mathbf{V} : set of coefficient vectors

$VALID$: original validation dataset

$VALID_F$: validation dataset after feature extraction

\mathbf{v} : coefficient vector

\mathbf{w} : optimal vector in RVM

w_n : n th optimal parameter in RVM

\mathbf{w}_{MP} : most probable weight vector in RVM

w_c : inertial weight of PSO

$WPT(\cdot)$: wavelet packet transform function

\mathbf{X} : set of ignition pattern vectors

\mathbf{x} : unseen ignition pattern

x_i : i th data point in \mathbf{x}

\mathbf{y} : predicted label vector

y_i : i th predicted label

y_i^j : i th predicted label in the j th test data

$z(\mathbf{f})$: predicted decision

$\boldsymbol{\alpha}$: hyperparameter vector of RVM

α_n : n th hyperparameter of RVM

ε : decision threshold

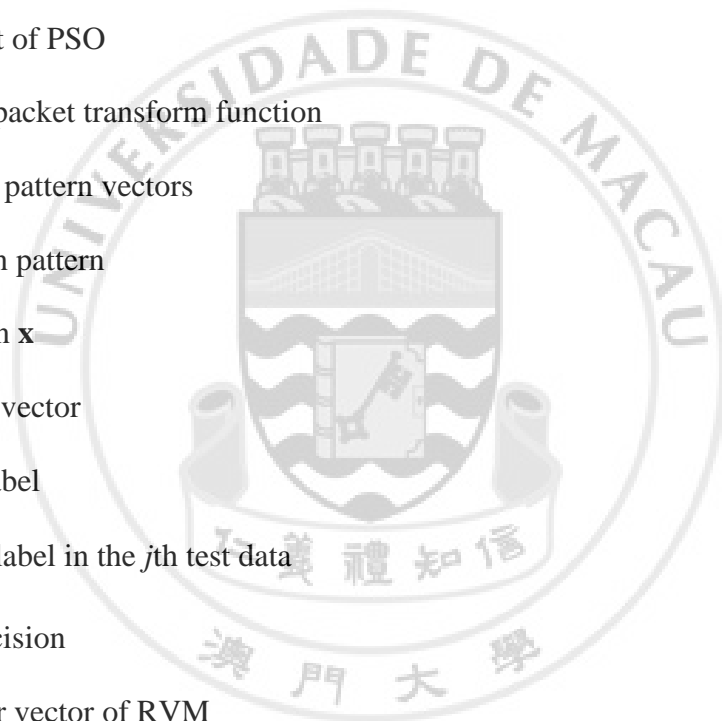
ε_k : k th tentative threshold produced in optimization process

ε_{opt} : optimized threshold

$\theta(\cdot)$: decision function of binarization approach

π : precision

$\boldsymbol{\rho}$: probability vector



ρ_i : probability of the i th label

ρ_{ij} : pairwise probability of the i th label against the j th label

Σ : covariance matrix in RVM

Σ_{ii} : i th diagonal element of covariance matrix Σ

$\sigma(\cdot)$: logistic sigmoid function

τ : recall

\mathcal{E} : initial population

Φ : design matrix in RVM



ACKNOWLEDGMENTS

The author wishes to thank Prof. Wong Pak Kin and Dr. Vong Chi Man for their patient instruction and valuable comments on the thesis.

