A COUSTIC PSEUDOSPECTRUM BASED FAULT LOCALIZATION IN MOTORCYCLES

Veerappa B. Pagì¹, Basavaraj S. Anami² and Ramesh S. Wadawadagi³

¹,³Dept. of Computer Applications, Basaveshwar Engg. College, Bagalkot, India
veereshpagi@gmail.com, rswbec@gmail.com
²Dept. of Computer Science and Enng., KLE Institute of Technology, Hubli, India
anami_basu@hotmail.com

ABSTRACT

Vehicles generate dissimilar sound patterns under different health conditions. The sound generated by the vehicles gives a clue of some of the faults. Automotive experts diagnose the faults in vehicles based on the produced sound. This paper presents a methodology for fault source localization in motorcycles using the estimated pseudospectra of the sound signals. The pseudospectra are traced to generate a chaincode. The generated chaincode is transformed into an eigenvector, which is used as the feature vector for classification using artificial neural network (ANN). The overall classification accuracy is 88%. The proposed work finds applications in traffic census of the vehicles, traffic rule observance, machine fault detection and localization, automatic acoustic surveillance and the like.

KEYWORDS

Fault Localization, Chaincode, Pseudospectrum, Eigen Vectors, Neural Classifier

1. INTRODUCTION

The sound patterns produced by the vehicles convey information necessary for fault diagnosis. Vehicle sound sources include parts of the engine subsystem, timing chain, clutch plate and exhaust subsystem. We rely on automotive experts for the repair and maintenance of vehicles. Automated fault diagnostic systems are essential when the vehicle is in remote places and in places of scarce expertise. Motorcycles dominate the Indian automobile market, with nearly 77% of total vehicle sales. Society of Indian Automobile Manufacturers (SIAM) has forecast the two-wheeler segment to register a growth of 6%-8% in 2013-14 [1].

This work is an attempt to localize the fault source in motorcycles based on the produced sound. The features are extracted from pseudospectra of the sound signals. The chaincode of the estimated pseudospectra are constructed and transformed into matrices. The eigenvectors [9] computed over the matrices are considered as features. These features are later subjected to classification by ANN classifier. The present work leaves scope for further investigation of localization of combinations of faults. The research contribution is significant since it indicates a fault well earlier, reducing a possible future accident. The service station experts can use the findings of this work for preliminary fault diagnosis.

The literature survey is organized into two parts: engine fault diagnosis and gearbox fault diagnosis. Following is the gist of the studied literature.

DOI:10.5121/ijcsa.2013.3307
The 1D central contour moments and their invariants of approximation coefficients of DWT are used as feature inputs for DTW classifier for determining the health condition of motorcycles [2]. The sound intensity method is used to identify the noise sources of engine front of a diesel engine [3]. A continuous wavelet transform (CWT) algorithm is combined with an ANN and generalized regression for analyzing fault signals in a scooter fault diagnosis system [4]. Engine fault diagnosis uses empirical mode decomposition (EMD) and wavelet packet BP neural network [5].

RMS and Power Spectral Density (PSD) of Massey Ferguson gearbox are calculated to detect different faults [6]. An adaptive wavelet filter based on Morlet wavelet is used for detection of symptoms from vibration signals of a gearbox with early fatigue tooth crack [7]. Vibration signals are decomposed into a finite number of intrinsic mode functions (IMF) and then the autoregressive (AR) model of each IMF component is established; finally, the corresponding AR parameters and the variance of remnant are regarded as the characteristic vectors and used as input of SVM classifier to classify the working condition of gear [8].

From the literature survey, it is evident that reasonable amount of research is reported at the wider range of applications, ranging from fault classification in machines to gearboxes. Since there is no work reported on fault source localization of motorcycles based on their sound patterns, we have taken up a study. The remainder of the paper is organized into 3 sections. The proposed methodology of the work, along with a brief on tools and techniques, is discussed in Section 2; the experimental results are discussed in Section 3. Finally, the Section 4 concludes the work.

2. PROPOSED METHODOLOGY

The overview of the methodology is depicted in figure 1. It comprises five stages namely, sound signal acquisition, segmentation, spectral analysis, feature extraction and classification.

Figure 1. Block diagram of the proposed method

Legend: VS: Valve setting; FC: Faulty crank; CK: Cylinder kit problem; TC: Timing chain fault; ML: Muffler leakage; SL: Silencer leakage
The following subsections briefly explain each of the important stages in the methodology.

2.1. Acquisition of Sound samples

The automotive guidelines suggest the recording frequency in the range of 9 kHz to 30 kHz. The sound signals of the motorcycles are recorded using Sony ICD-PX720 digital voice recorder, with sampling frequency of 44.1 kHz and quantized with 16 bits. The recording of the sound samples is carried out in service stations under the supervision of service experts. The selected healthy motorcycles are less than one year old, not run more than 6000 km, well-maintained and regularly serviced in authorized service stations. The motorcycles having faulty crank, damaged timing chain, unset valve, muffler leakage, silencer leakage, and faulty cylinder kit are considered in this work. The scope of the work is limited to the fault source localization in motorcycles with single fault. Figure 2 depicts the recording environment.

The recording environment has disturbances from human speech, sound of other vehicles being serviced, air-compressor and auto-repair tools. The recorder is held 500 mm from the centerline of the exhaust end, and at the angle of 45° measured from the centerline of the exhaust end, and at the height of the exhaust pipe. The motorcycle is in neutral state, and stationary. The engine runs in idle state and the throttle is controlled by the expert, while recording the sound signals.

A brief description of the faults considered in this work is given as under.
Timing Chain: Timing chain operates the valves. A loose chain vibrates and results in improper timing for valve operation. This results in abnormal operation of engine leading to noise.
Valve setting: For smooth functioning of engine, correct operation of valves is necessary. Any deviation of 5 to 10 degrees in valve opening/closing will cause considerable rise in peak combustion chamber pressures, leading to change in sound.

Crank fault: It may occur due to wear and tear of either oil ring, first ring or second ring.

Muffler Leakage: The exhaust gases coming out of combustion chamber passes through muffler. The main function of muffler is to reduce the noise and filter exhaust gases. Due to the reactive gases in the residual exhaust, mixed with water vapour, result in corrosion reactions. This creates minute holes in the muffler and changes the firing sound coming out of engine.

Silencer leakage: If there is a hole inside the silencer filter pipe or the damaged gasket, it causes silencer leakage.

2.2. Segmentation

The acquired sound samples are segmented into samples of one-second each for uniformity in processing. The portion of the signal of duration one second, beginning from the local maxima is considered as a segment. The next segment begins at local maxima in the next 50 ms duration from the end of the previous segment.

2.3. Feature Extraction

The chaincode of the pseudospectrum is constructed by tracing the gradient changes at each point. The eigenvectors computed based on the constructed chaincodes are used as feature vectors.

2.3.1. Pseudospectrum estimation

The multiple signal classification (MUSIC) algorithm [10] is used for estimating the pseudospectrum of the sound signals. The pseudospectrum is calculated using estimates of the eigenvectors of a correlation matrix associated with the input data. The MUSIC estimate is given by the Equation (1).

\[
P_{\text{music}}(f) = \frac{1}{e^{H}(f) \sum_{k=p+1}^{N} v_k v_k^H e(f) + \sum_{k=p+1}^{N} |v_k e(f)|^2} \quad \ldots (1)
\]

where \( N \) is the dimension of the eigenvectors and \( v_k \) is the \( k \)th eigenvector of the correlation matrix of the input signal. The integer \( p \) is the dimension of the signal subspace, so the eigenvectors \( v_k \) used in the sum correspond to the smallest eigenvalues. The vector \( e(f) \) consists of complex exponentials, so the inner product \( v_k^H e(f) \) amounts to a Fourier transform. In the eigenvector method, the summation is weighted by the eigenvalues \( \lambda_k \) of the correlation matrix, as shown in Equation (2).

\[
P_{\text{ev}}(f) = \frac{1}{\sum_{k=p+1}^{N} |v_k e(f)|^2 / \lambda_k} \quad \ldots (2)
\]
Figure 3 shows the spectra of sound signals of healthy and faulty motorcycles.

![Figure 3. Spectra of sound signatures of healthy and faulty motorcycles](image)

In case of the spectra of healthy samples the spectral peaks decrease monotonically and exhibit regular variations. But, in case of faulty motorcycles, the degraded harmonicity, non-monotonic decrease in spectral peaks and spurious peaks at higher frequencies, are observed. Figure 4 combines the spectra of different faults under consideration. The variations in peaks and harmonicity are larger at low frequencies. The features based on this part of the spectrum aid for better classification.

![Figure 4. Spectra of different faults](image)

**2.3.2 Chaincode of a spectral segment**

The chaincode of a spectral segment is generated based on the gradient changes, tracing the pseudospectrum from left to right.

The chaincode vector of a spectral trace is computed as:

\[
\text{ChCode}(i) = \begin{cases} 
0 & x(i) = x(i+1) \\
7 & x(i) > x(i+1) \\
1 & x(i) < x(i+1)
\end{cases} \quad \ldots(3)
\]
Where, \(1 \leq i \leq 128\)

Figure 5 shows the directions in which the spectral gradients change.

![Figure 5. Chaincode directions](image)

Since the gradient changes can be observed only in three directions, the constructed chaincodes comprise 0 (right), 1 (top-right), and 7 (bottom-right) as shown in Figure 5. The reference feature vector for healthy motorcycles is constructed by taking the mode of the respective chaincodes. Similarly the reference feature vector for faulty motorcycles is constructed.

### 2.3.3. Dimensionality reduction

Figure 6 summarizes the dimensionality reduction process. The 128-digit chaincode is transformed into a matrix of size \(16 \times 8\). The adjacent values in each row are averaged, converting the matrix into a square matrix of order \(8 \times 8\). The eigenvector of the reduced square matrix is computed, resulting in a vector of length eight. The eigenvector is considered as the feature vector input to the neural classifier with eight inputs.

![Figure 6. Dimensionality reduction](image)

### 2.3.4. Eigenvectors

Let \(A\) be a complex square matrix. Then if \(\lambda\) is a complex number and \(X\) a non–zero complex column vector satisfying \(AX = \lambda X\), we call \(X\) an eigenvector of \(A\), while \(\lambda\) is called an eigenvalue of \(A\). We also say that \(X\) is an eigenvector corresponding to the eigenvalue \(\lambda\). Table 1 summarizes the mean eigenvectors for different types of faults.
Table 1. Mean eigenvector values for different types of faults.

<table>
<thead>
<tr>
<th>Fault</th>
<th>Eig1</th>
<th>Eig2</th>
<th>Eig3</th>
<th>Eig4</th>
<th>Eig5</th>
<th>Eig6</th>
<th>Eig7</th>
<th>Eig8</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1 (VS)</td>
<td>9.1235</td>
<td>0.8387</td>
<td>0.4893</td>
<td>0.4535</td>
<td>0.2013</td>
<td>0.0940</td>
<td>0.0210</td>
<td>0.0272</td>
</tr>
<tr>
<td>F2 (FC)</td>
<td>9.3850</td>
<td>0.7107</td>
<td>0.6083</td>
<td>0.4536</td>
<td>0.1943</td>
<td>0.0585</td>
<td>0.0000</td>
<td>0.1897</td>
</tr>
<tr>
<td>F3 (CK)</td>
<td>8.2913</td>
<td>1.4210</td>
<td>0.4033</td>
<td>0.2424</td>
<td>0.1008</td>
<td>0.0883</td>
<td>0.0313</td>
<td>0.0000</td>
</tr>
<tr>
<td>F4 (ML)</td>
<td>9.3886</td>
<td>0.4682</td>
<td>0.2182</td>
<td>0.1493</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>F5 (SL)</td>
<td>7.7834</td>
<td>1.8341</td>
<td>0.3541</td>
<td>0.2058</td>
<td>0.1021</td>
<td>0.0925</td>
<td>0.0085</td>
<td>0.2258</td>
</tr>
<tr>
<td>F6 (TC)</td>
<td>8.0524</td>
<td>1.3658</td>
<td>0.4363</td>
<td>0.2718</td>
<td>0.0597</td>
<td>0.0261</td>
<td>0.0000</td>
<td>0.0167</td>
</tr>
</tbody>
</table>

2.3.5. ANN classifier

Figure 7 shows the overview of the architecture of the ANN. The eight features, extracted from the chaincode, are input to the neural network with 8 input nodes. The six output nodes correspond to the six-bit output vector indicating the type of the fault in the motorcycle. The hidden layer contains 10 nodes. The neural network is trained using backpropagation-learning algorithm [11]. The stabilized weights are reloaded and test vectors are input during testing. The optimal number of hidden layer neurons is chosen using the criterion [12]:

\[ n = C \sqrt[4]{\frac{N}{d \log N}} \]  \( \ldots (4) \)

where, \( n \)=number of hidden layer neurons, \( C \)=constant to yield optimal performance, \( d \)=number of features, and \( N \)=number of rows in the training sample matrix.

The mean squared error (MSE) is computed for the sample sets with 70% of the samples used for training, 15% for validation and 15% for testing.

![Figure 7. Architecture of ANN](image)

Figure 8 shows a typical learning process of ANN where the goal is met. In this case, goal for the error is set as 0.00001, the neural network is trained with learning parameter of 0.3, and the algorithm halted after 731 epochs. If the goal is met and the same feature vectors are used for testing, the classification accuracy will be 100%. If the user terminates the testing process before the goal is met, normally for large test sets, the classification accuracy suffers. As the number of training samples is increased, the learning rate is increased accordingly.
Figure 9 shows the validation MSE plotted for varying number of neurons in the hidden layer. The minimum MSE is observed when 10 neurons are present in the hidden layer. Hence, the neural network is designed with eight input nodes, 10 hidden layer nodes and six output nodes. The performance of the built neural network is analyzed for different combinations of training, validation and testing sets.

3. RESULTS AND DISCUSSION

The faults present in motorcycles of different models are considered. Models considered are Passion plus, Extreme, Street 100, CD 100, CD Dawn, CD Deluxe, CBZ, Splendour, Splendour plus, Super splendour, all from Hero Honda (Presently known as Hero MotoCorp). The training sample set contains 360 sound samples of motorcycles. The samples are partitioned into disjoint sets for training and testing based on random selection. The size of the training and test sample sets is varied from 10 to 60, for each type of fault. Confusion matrix for a total of 360 test samples, containing 60 samples of each type of fault, is given in Table 2.
From the Table 2 it the false acceptance rate is observed to be 0.1167, which is admissible in real-world fault diagnosis. Hence, the overall classification accuracy of the approach is 0.8834. Classification accuracy for different types of faults is shown in Table 3, when the number of test samples for each type of fault is varied from 10 to 60. The classification accuracy is 100% when 10 and 20 samples of each type of fault are used for testing.

Table 3. Classification accuracy for different faults on varying sizes of inputs

<table>
<thead>
<tr>
<th>No. of test samples</th>
<th>Mis-set Valve</th>
<th>Faulty crank</th>
<th>Faulty cylinder kit</th>
<th>Muffler leakage</th>
<th>Silencer leakage</th>
<th>Faulty timing chain</th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>120</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.9000</td>
<td>0.9500</td>
<td>0.9000</td>
<td>1.0000</td>
</tr>
<tr>
<td>180</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.8000</td>
<td>0.9667</td>
<td>0.8667</td>
<td>0.9667</td>
</tr>
<tr>
<td>240</td>
<td>0.8750</td>
<td>1.0000</td>
<td>0.7750</td>
<td>0.9750</td>
<td>0.8750</td>
<td>0.9250</td>
</tr>
<tr>
<td>300</td>
<td>0.7600</td>
<td>1.0000</td>
<td>0.8000</td>
<td>0.9800</td>
<td>0.9000</td>
<td>0.8400</td>
</tr>
<tr>
<td>360</td>
<td>0.8167</td>
<td>1.0000</td>
<td>0.8500</td>
<td>0.9667</td>
<td>0.7834</td>
<td>0.8834</td>
</tr>
</tbody>
</table>

It is evident from Table 3 that the classification accuracy suffers with increase in the size of the samples sets.

4. CONCLUSION

The presented work identifies the source of the fault in motorcycles, based on the acoustic signals. Eigenvectors of the transformed chaincodes generated from the pseudospectra are used as feature vectors. The extracted features are classified into one of the six types of faults using ANN classifier. The classification accuracy is over 88% for testing with 60 samples of each type of fault. Since these results are obtained without denoising, the approach is suitable for real-world implementation. The method is not apt for on-ride fault diagnosis since it requires considerable
amount of time for spectral computations and ANN training. The fault localization in presence of multiple faults will be attempted in the future work.

ACKNOWLEDGEMENTS

We the authors thank Shrinivas M. S. and Basavaraj Kumbar for recording and testing help, We thank Prof. Vishwanath Kagawade and Prof. Sangamesh Magi for language corrections. The service stations in Bagalkot city need special mention for sharing the fault diagnostic expertise.

REFERENCES

Authors

Basavaraj S. Anami  B. S. Anami, received the Ph.D. from the University of Mysore, and the M. Tech. from Indian Institute of Technology, Chennai, both in Computer Science, in 2003 and 1986, respectively. From 1983 to 2008, he was with the Basaveshwar Engineering College, Bagalkot, India, where he worked on several research projects in image processing. Presently he is with KLE Institute of Technology, Hubli, India, as Principal. He visited Saginaw Valley State University, USA, as Visiting Professor, from August to December 2006. He has authored two books and over 50 refereed papers. He is a fellow of Institution of Engineers, India, and member of Indian Society of Technical Association and Computer Society of India. He has guided five students for Ph. D.

Veerappa B. Pagi  V. B. Pagi, received the M. E. degree from the Bangalore University, Bangalore, and B. E. Degree from Karnatak University, Dharwad, India, both in Computer Science, in 2007 and 1994, respectively. From 1994 to 1999, he was with Sir M. Visvesvaraya Institute of Technology, Bangalore, where he worked as Lecturer in the Department of Computer Science. In 1999, he joined Basaveshwar Engineering College, Bagalkot, where he is currently working as Assistant Professor, in the Department of Computer Applications. He has published 9 papers in International journals and 3 papers in conferences. He has guided over 10 projects in the area of Image processing and 3 projects in the area of sound signal processing.

Ramesh S. Wadawadagi  R. S. Wadawadagi received the B.Sc., M.Sc. degrees in Computer Science from Karnatak University Dharwad, and the M. Tech. degree in Computer Science from University of Mysore, in 2001, 2004, and 2009, respectively. From 2009 he is working as Assistant Professor in the department of Computer Applications, Basaveshwar Engineering College, Bagalkot, Karnataka State, India. His areas of interest include pattern recognition, data mining and system simulation. He has guided over 10 projects for the postgraduate students. His current research activities include vehicle engine diagnostics, neural networks, and web content mining.