Chapter 6

Fault Diagnosis Framework

'*New frameworks are like climbing a mountain - the larger view encompasses rather than rejects the more restricted view.*'

-Albert Einstein

Highlights:

- A framework for diagnosis of bearing faults under variable and constant speed conditions is proposed.
- This chapter includes comparative analysis of three existing and two proposed signal processing algorithms.
    1. Hilbert transform based enveloping
    2. Wavelet packet transform and enveloping
    3. Variational mode decomposition
    4. Fractional enveloping
    5. Cumulative distribution sharpness profiling
- Two time-frequency domain fault features - *Prominence* and *Compliance* - are defined for diagnosis.
- The proposed CDSP method is found to be superior under different objectives of fault diagnosis.

In previous chapters, two signal processing methods - fractional enveloping (FE) and cumulative distribution sharpness profiling (CDSP) - are proposed and validated on publicly available datasets. These methods effectively enhanced the fault signature in the signal. The objective of this chapter is to verify these algorithms under proposed framework. A comparative analysis is carried out with well-established signal processing algorithms.

6.1 Proposed Framework

A common framework for diagnosis of bearing faults under constant as well as variable speed conditions is proposed in Fig. 6.1. Fractional enveloping and cumulative distribution sharpness profiling are the proposed signal processing methods, which are discussed, in detail, in previous chapters. Available MTFCE algorithm [153] is used for extraction of the ridge curves. Two time-frequency domain features - *Prominence* and *Compliance* are also proposed for diagnosis.

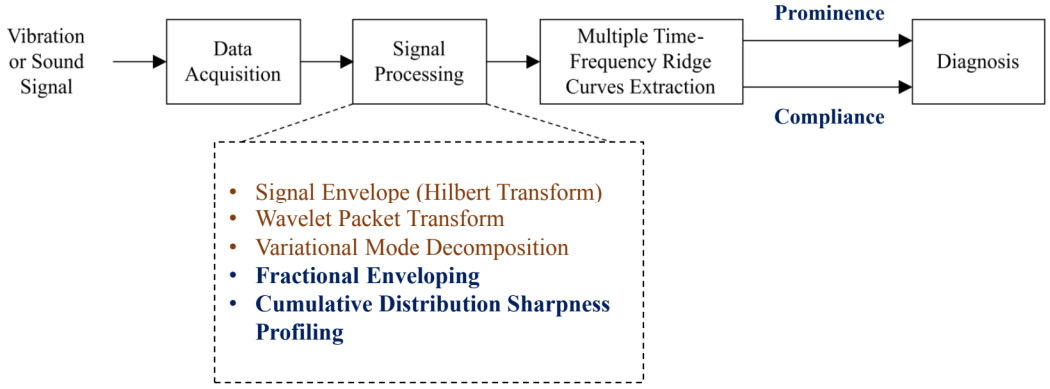

Figure 6.1: Proposed framework for diagnosis of bearing faults under constant as well as variable speed conditions.

6.1.1 Signal Processing

Following signal processing methods are used in the proposed framework and their results are compared.

1. Method 1: This is the conventional Hilbert transform method to calculate the signal envelope.
2. Method 2: As discussed in the literature review, signal enveloping alone is not enough, but has to be preceded by band-pass filtering or signal decomposition. Since wavelet transform is a well-established method for bearing fault diagnosis, it is used with signal enveloping. The discrete wavelet packet transform (dual-tree, maximum overlap) is used as the decomposition technique with the mother wavelets 'db8' and 'fk4'. Maximum decomposition level is 3. For selection of optimum band and optimum mother wavelet, the diagnosis percentages, calculated from two proposed features - *Prominence* and *Compliance* - are maximised for each case.
3. Method 3: As discussed in the literature review, variational mode decomposition (VMD) is a modification of the empirical mode decomposition. Using the VMD, the signals are decomposed into the intrinsic mode functions (IMF). For selection of optimum IMF, the diagnosis percentages, calculated from two proposed features - *Prominence* and *Compliance* - are maximised for each case.
4. Method 4: This is the fractional enveloping method, in which the fractional Fourier transform (FrFT) is applied followed by the wavelet packet transform. The optimum fractional parameter is found by maximising the diagnosis percentages, calculated from two proposed features - *Prominence* and *Compliance*, over the range $[0, 1]$ with an interval of 0.1. The wavelet decomposition is optimised as discussed in Method 2.
5. Method 5: The cumulative distribution sharpness profile (CDSP) of the signal is calculated in this method. The procedure for calculating CDSP is discussed, in detail, in Chapter 5.

6.1.2 Multiple Time-Frequency Ridge Curves Extraction

The signal envelope obtained from each of the above methods is represented using the short-time Fourier transform (STFT). Although, the time-frequency representation using STFT is not necessary for constant speed conditions, it provides a uniform way of analysis for different operating conditions. That is, using STFT both the constant as well as the varying speed conditions can be processed. The window type, window length and overlap are empirically selected during validation as hamming, 4.5% and 0.1% of the signal length respectively. First four TFR ridge curves are then extracted from the STFT using the algorithm given in [153].

6.1.3 Time-Frequency Features and Diagnosis

Two proposed features - *Prominence* and *Compliance* - are defined for the extracted TFR ridge curves. Out of the four TFR ridge curves $(f_1, f_2, f_3 \ \& \ f_4)$, curve with minimum mean squared error (MSE) with theoretical fault frequency is selected as the estimated fault frequency. *Prominence* is defined as a measure of how prominent is the estimated frequency curve compared to other extracted curves.

Let A_{f_i} be the average amplitude of the first four extracted curves and $A_{f'}$ be the average amplitude of the estimated fault frequency curve. Then the *Prominence* is proposed as $Pr = 100 \times (A_{f'} - \mu)/\mu$, where $\mu = \Sigma\, A_{f_i}/4$.

Larger value of *Prominence* indicates the presence of fault. Depending on closeness of prominent frequency with the fault frequency, the location of fault can also be identified. But, in case of variable speed conditions, *Prominence* is not sufficient for diagnosis.

Compliance is defined as a measure of closeness between the estimated frequency curve and the theoretical fault frequency. The *Compliance* is proposed as $Co = 100 \times n_c/N$. Where, the estimated fault curve (f') has N samples and for n_c samples it satisfies $|f' - f_f| \leq \varepsilon$, where f_f is the theoretical fault frequency for inner race or outer race and ε is the threshold for closeness between these curves.

This deviation is estimated to be 2% [182] to 5% for the dataset under study [152]. This 5% tolerance is in fault characteristic coefficient (FCC), which is nothing but a multiplication factor to calculate the fault frequency. Considering, the range of operating speed upto 2000 rotations per minute, that is 33.33 Hz, and the FCC of inner race fault to be 5.43, the 5% tolerance becomes $\varepsilon = 33.33 \times 5.43 \times 5/100 = 9.05\ Hz$. Thus, the value of ε is chosen as 10 Hz.

Larger value of *Compliance* indicates that the estimated fault curve follows the theoretical fault frequency closely for larger duration of time. *Compliance* is mainly important under variable speed conditions and is helpful in fault localization. For example, if the estimated fault curve has larger *Compliance* with inner race frequency than outer race frequency, then inner race fault is detected.

The fault detection thus depends on proposed features - Prominence and Compliance and the diagnosis percentages (Di) are calculated as a product of *Prominence* (Pr) and *Compliance* (Co) scaled down to 100.

6.2 Results

CWRU (2013), Set 1

Description:	Diagnosis of *small (7mil)*, *medium (14 mil)* and *large (21 mil)* size inner race and outer race faults under *constant* operational speed using *primary* accelerometer located near the BUT.
Operating Speeds:	Constant; around 1700 rotations per minute
Data Acquisition:	Vibration signals; for 10 sec; 48 KHz (Down sampled to 20 KHz for processing)
Results:	Maximum accuracy of 91% using proposed method of *CDSP*

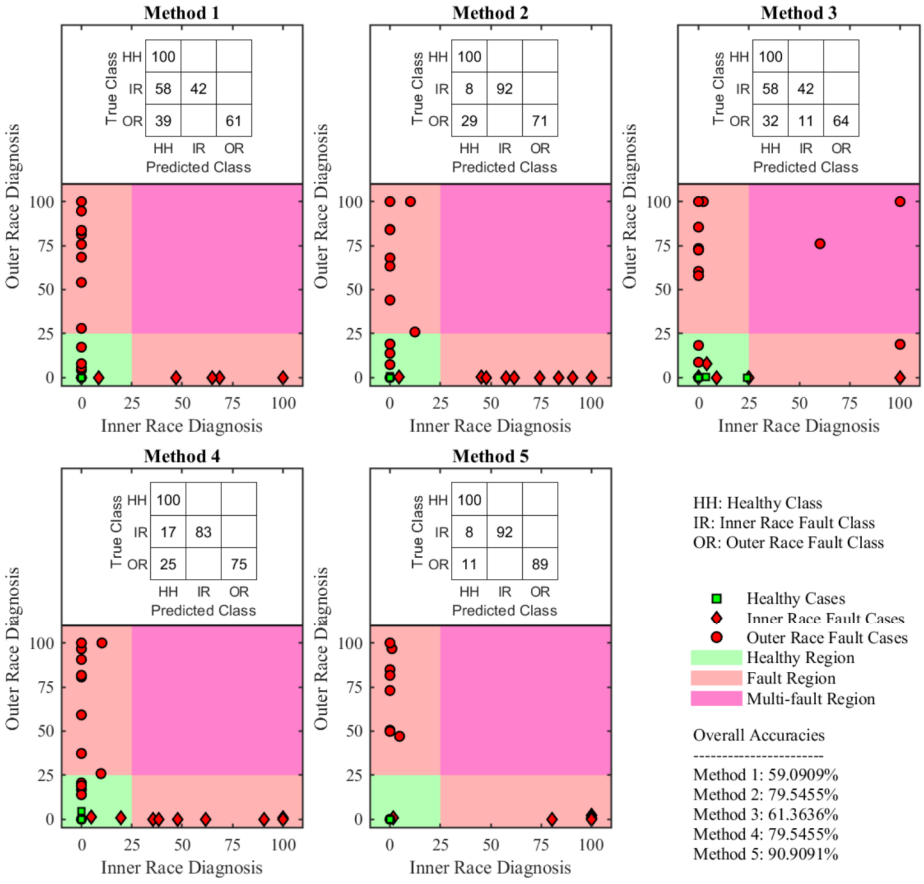

Figure 6.2: Comparative analysis of results for bearing dataset by Case Western Reserve University (2013), Set 1. (Note: The confusion matrices are row normalised, that is, the rows should add up to 100. If this exceeds 100, then some cases are wrongly diagnosed as *multiple* fault cases. Method 1 - Hilbert Transform, Method 2 - Wavelet Packet Transform, Method 3 - Variational Mode Decomposition, Method 4 - Fractional Enveloping, Method 5 - Cumulative Distribution Sharpness Profiling.)

CWRU (2013), Set 2

Description:	Diagnosis of *small (7mil)*, *medium (14 mil)* and *large (21 mil)* size inner race and outer race faults under *constant* operational speed using *secondary* accelerometer located away from the BUT.
Operating Speeds:	Constant; around 1700 rotations per minute
Data Acquisition:	Vibration signal; for 10 sec; 48 KHz (Down sampled to 20 KHz for processing)
Results:	Maximum accuracy of 86% using proposed method of *CDSP*

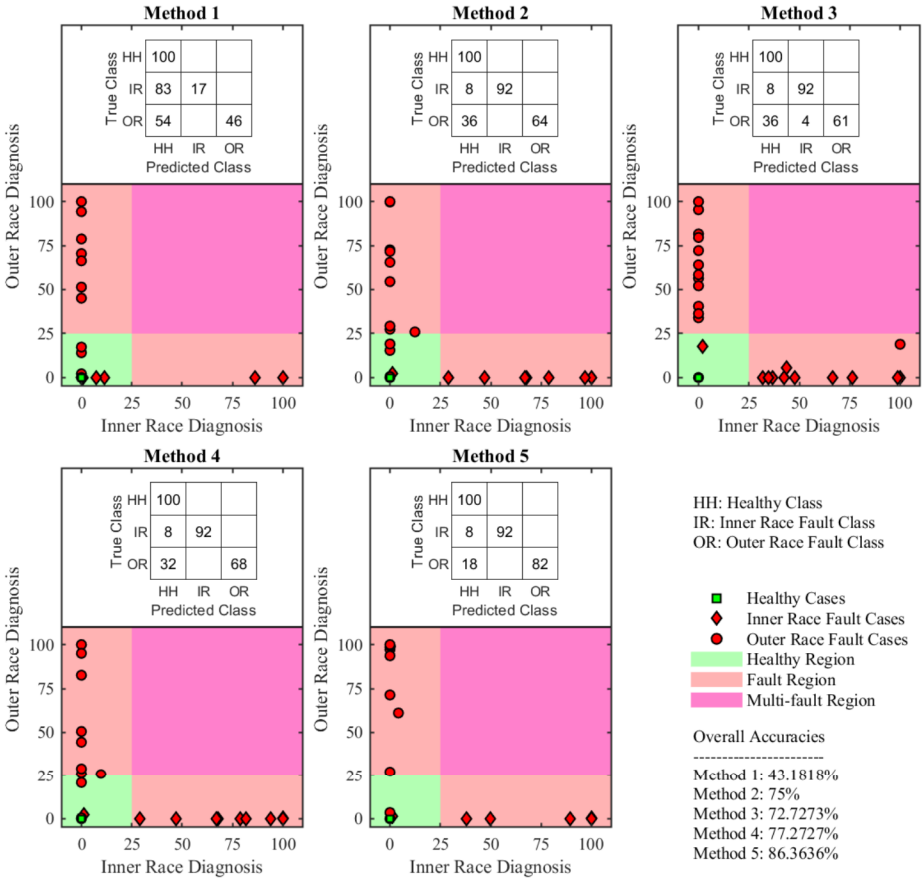

Figure 6.3: Comparative analysis of results for bearing dataset by Case Western Reserve University (2013), Set 2. (Note: The confusion matrices are row normalised, that is, the rows should add up to 100. If this exceeds 100, then some cases are wrongly diagnosed as *multiple* fault cases. Method 1 - Hilbert Transform, Method 2 - Wavelet Packet Transform, Method 3 - Variational Mode Decomposition, Method 4 - Fractional Enveloping, Method 5 - Cumulative Distribution Sharpness Profiling.)

UOO (2018)

Description:	Diagnosis of inner race and outer race faults under *varying* speed using *primary* accelerometer located near the BUT.
Operating Speeds:	Variable; around 1700 rotations per minute
Data Acquisition:	Vibration signal; for 10 sec; 200 KHz (Down sampled to 20 KHz for processing)
Results:	Maximum accuracy of 100% using proposed method of *CDSP*

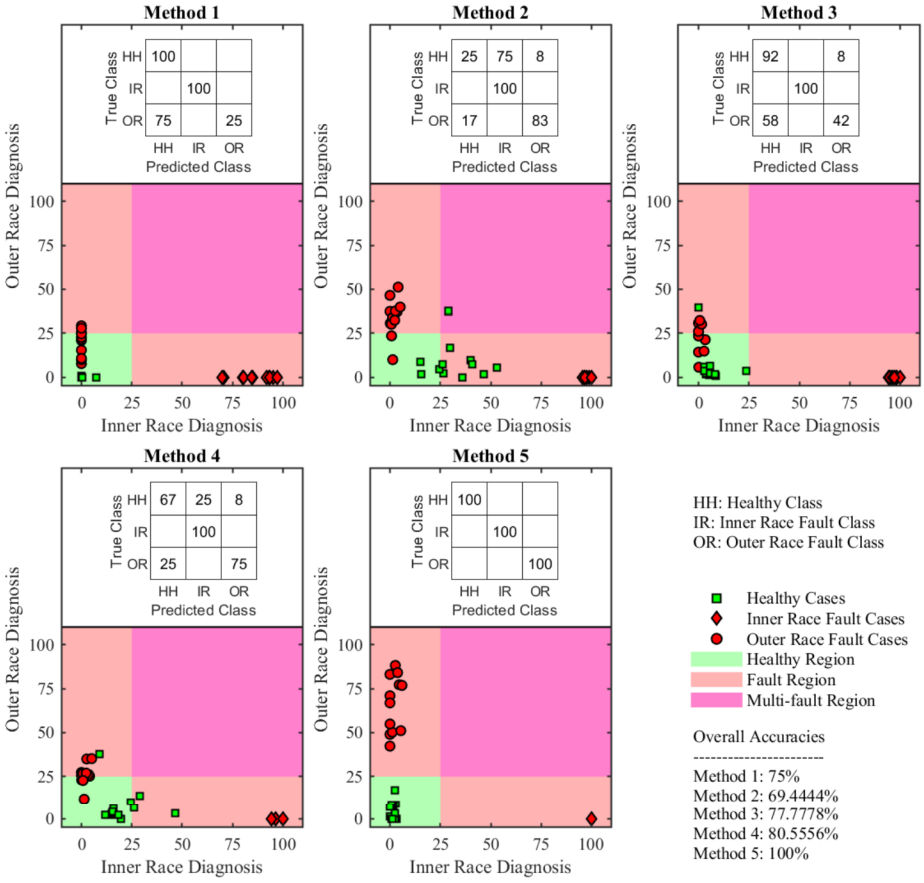

Figure 6.4: Comparative analysis of results for bearing dataset by University of Ottawa (2018). (Note: The confusion matrices are row normalised, that is, the rows should add up to 100. If this exceeds 100, then some cases are wrongly diagnosed as *multiple* fault cases. Method 1 - Hilbert Transform, Method 2 - Wavelet Packet Transform, Method 3 - Variational Mode Decomposition, Method 4 - Fractional Enveloping, Method 5 - Cumulative Distribution Sharpness Profiling.)

BITS (2021), Set 1

Description:	Diagnosis of *medium (0.6 mm)* and *large (0.7 mm)* inner race and outer race faults under *constant* and *varying* speed using primary *accelerometer* located near the BUT.
Operating Speeds:	Constant; approximately 1000, 1200, 1400, 1600, 1800 and 2000 rotations per minute Variable; approximately 1200-1500, 1500-1200, 1700-2000, 2000-1700 rotations per minute
Data Acquisition:	Vibration signal; for 10 sec; 100 KHz (Down sampled to 20 KHz for processing)
Results:	Maximum accuracy of 100% using proposed method of *CDSP*

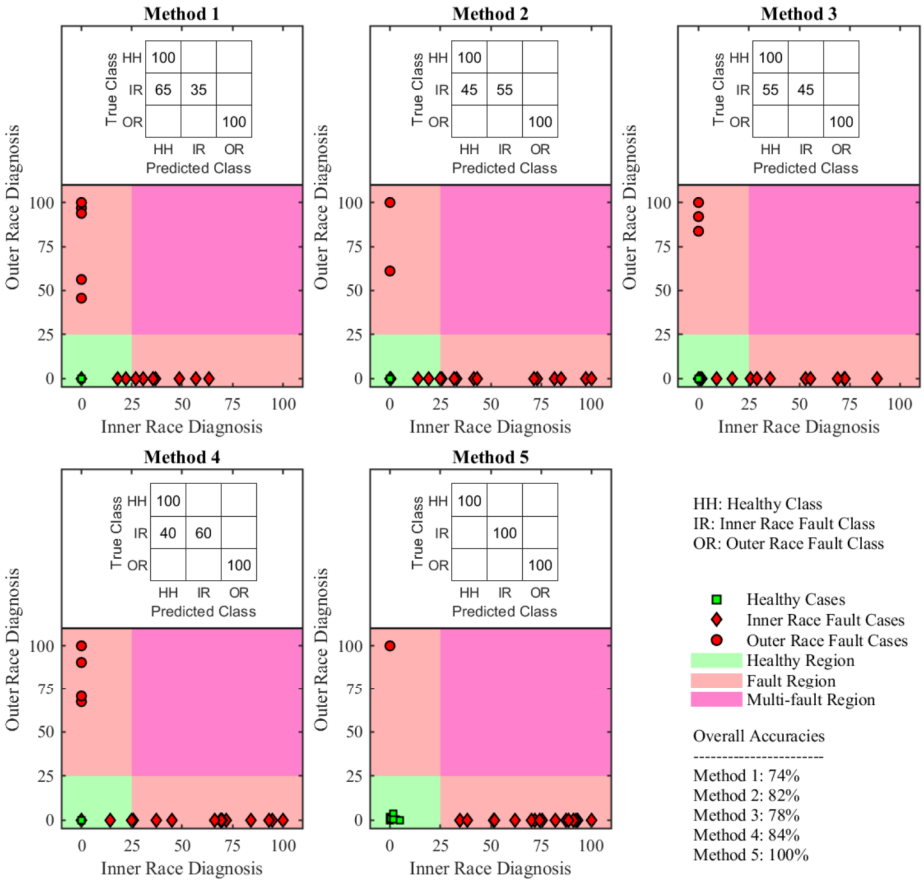

Figure 6.5: Comparative analysis of results for experiments performed at Birla Institute of Technology and Science (2021), Set 1. (Note: The confusion matrices are row normalised, that is, the rows should add up to 100. If this exceeds 100, then some cases are wrongly diagnosed as *multiple* fault cases. Method 1 - Hilbert Transform, Method 2 - Wavelet Packet Transform, Method 3 - Variational Mode Decomposition, Method 4 - Fractional Enveloping, Method 5 - Cumulative Distribution Sharpness Profiling.)

BITS (2021), Set 2

Description:	Diagnosis of *medium (0.6 mm)* and *large (0.7 mm)* inner race and outer race faults under *constant* and *varying* speed using primary *microphone* located near the BUT.
Operating Speeds:	Constant; approximately 1000, 1200, 1400, 1600, 1800 and 2000 rotations per minute Variable; approximately 1200-1500, 1500-1200, 1700-2000, 2000-1700 rotations per minute
Data Acquisition:	Sound signal; for 10 sec; 100 KHz (Down sampled to 20 KHz for processing)
Results:	Maximum accuracy of 94% using proposed method of *CDSP*

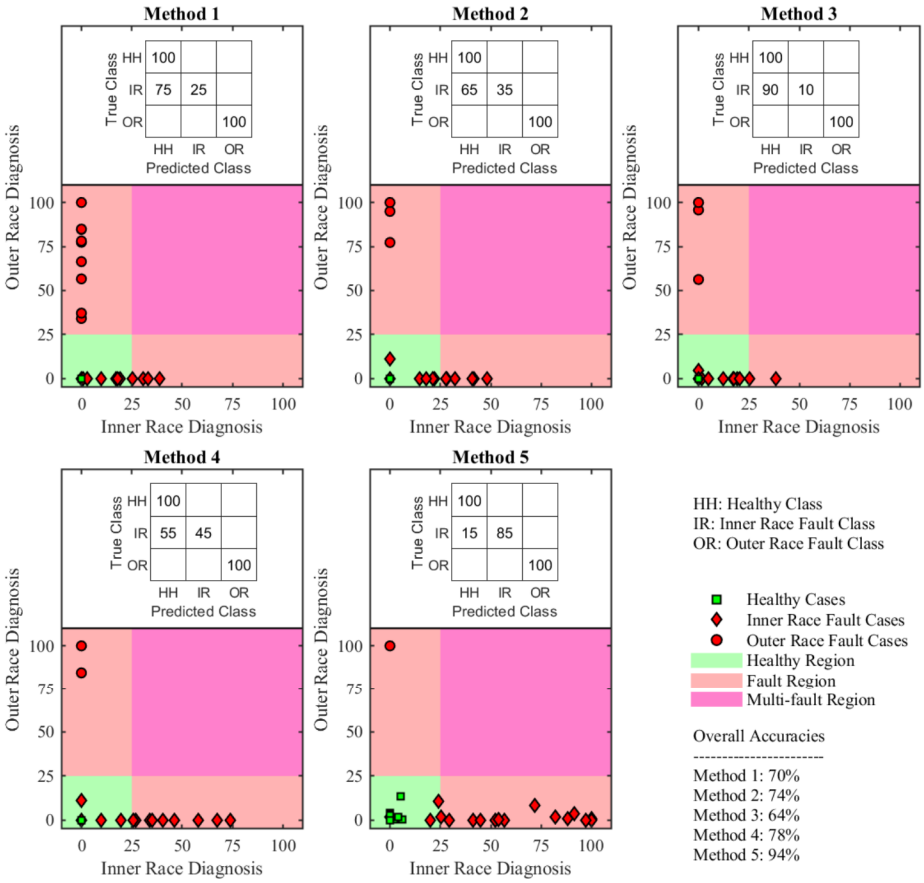

Figure 6.6: Comparative analysis of results for experiments performed at Birla Institute of Technology and Science (2021), Set 2. (Note: The confusion matrices are row normalised, that is, the rows should add up to 100. If this exceeds 100, then some cases are wrongly diagnosed as *multiple* fault cases. Method 1 - Hilbert Transform, Method 2 - Wavelet Packet Transform, Method 3 - Variational Mode Decomposition, Method 4 - Fractional Enveloping, Method 5 - Cumulative Distribution Sharpness Profiling.)

6.2.1 CWRU-2013

It is a widely studied dataset for bearing fault diagnosis. The details of this dataset can be found in Chapter 3. The analysis for different subsets is carried out under the framework discussed in Fig. 6.1 and the results are shown in Fig. 6.2 and Fig. 6.3.

Diagnosis Using Primary Accelerometer

Fig. 6.2 shows results for the CWRU-2013 dataset when the signal is acquired using the primary accelerometer located near the bearing under test (BUT). The conventional Hilbert transform method has poorest accuracy for both inner race and outer race faults. Although all the healthy cases lie below the 25% diagnosis threshold, several faulty cases also lie in this region and, thus, wrongly identified as healthy cases. The medium and large size inner race faults are misclassified by this method. The accuracy of detecting outer race faults is better compared to that of the inner race faults. However, the overall accuracy of 59% makes this method unsuitable for fault diagnosis.

As highlighted in the literature review in Chapter 2, conventional signal enveloping alone is not useful. But, when preceded by a wavelet packet transform, it shows significant improvement in the overall accuracy - from 59% to 79.5%. This improvement is mainly due to maximisation of diagnosis percentages in the wavelet domain. As the calculation of diagnosis percentages involves computation of short-time Fourier transform, it takes more amount of time. This limits the maximum level at which the signal can be decomposed. However, as further diagnosis depends on these percentages, proposed maximisation achieves better results over kurtosis or Hoyer index optimised level 7 decomposition. The accuracy of detecting healthy and inner race fault cases, using this method, is at par with the best performing method. Whereas, some of the outer race faults are wrongly diagnosed as healthy cases. This method still outperforms the variational mode decomposition.

Variational mode decomposition (VMD) method is modification of the empirical mode decomposition and is advantageous over the wavelet transform because the basis function

need not be chosen for this method. In wavelet based methods (Method 2 and 4), choice of mother wavelet is a critical issue, but in VMD only the optimum intrinsic mode function (IMF) needs to be selected. This is proposed to be done by maximising the diagnosis percentages. Selection of IMFs based on kurtosis, Hoyer index or other sparsity based measures is found to be less effective to the proposed maximisation. However, the accuracy of this method is poor and few outer race fault cases are wrongly diagnosed as multiple fault cases, that is, cases having both inner and outer race faults.

Application of fractional enveloping is discussed, in detail, in Chapter 4. This method is found to improve the accuracy compared to conventional enveloping. However, for this dataset, accuracy of Method 2 and Method 4 is equal. This method improves detection of outer race faults, but it also loses on the inner race fault detection. The overall accuracy of 79.5% is still a significant improvement.

The proposed cumulative distribution sharpness profiling (CDSP) method outperforms other methods in this framework. With overall accuracy of 90.9%, only few inner and outer race fault cases are misclassified as healthy cases. This is mainly due to two reasons - shorter duration of data acquisition and severe misalignment or imbalance.

The effect of misalignment or imbalance is discussed in detail in [169] and many cases in this dataset are non-diagnosable because of this. The CDSP method also fails in such cases. From signal processing perspective, this problem can be tackled by pre-processing the signal using cepstral editing to remove the speed harmonics. However, this is not possible in variable speed condition. Two misclassified cases are found to have shorter time duration. This affects the time resolution of the short-time Fourier transform and thus such cases are not properly diagnosed. It is therefore recommended that the signal should be acquired for a duration of at least 10 sec.

Improvement in the diagnosis percentages is also seen in this method. Thus, CDSP creates better separation between the fault types and allows diagnosis with increased confidence.

Diagnosis Using Secondary Accelerometer

Fig. 6.3 shows results for the CWRU-2013 dataset when the signal is acquired using the secondary accelerometer. As the secondary accelerometer is located away from the bearing under test, the acquired signal has more interference and noise from other machine components. Thus, diagnosis of faults is difficult for this set, compared to that of the primary accelerometer data.

Similar to the results from previous set, the accuracy of conventional Hilbert transform is poorest. Only 17% inner race fault cases are correctly diagnosed, remaining are misclassified as healthy.

This method shows significant improvement over Method 1, but diagnosis of outer race faults is poor compared to the results of primary accelerometer.

The accuracy of VMD is improved for secondary accelerometer compared to the primary accelerometer. It is evident that the effect of interfering signal components can be reduced using this method. Because of this reason it is an efficient signal processing method.

For secondary accelerometer also, the fractional domain maximisation shows significant improvement in the accuracy of diagnosis. This method gives better result compared to the wavelet or VMD alone.

CDSP is again found to be an efficient fault diagnosis method compared to other signal processing methods. Overall, inner race fault diagnosis is not affected due to the location of the accelerometer, but the accuracy of outer race fault diagnosis drops. The reasons of misclassification for this method are same as discussed for previous set.

6.2.2 UOO-2018

The objective of this dataset is diagnosis of bearing faults under varying speed conditions. The details are discussed in Chapter 3 and can also be referred to in the original paper [168]. The analysis is carried out under the framework discussed in Fig. 6.1.

Performance of Method 1 is much better compared to the CWRU (2013) dataset, as shown in Fig. 6.4. It even outperforms the wavelet based Method 2. In fact, although

accuracy of Method 3 and Method 4 is improved, the diagnosis percentages are very less and thus the confidence in the diagnosis is deteriorated. This shows that, decomposition based methods perform poorly under variable speed conditions. This is due to selection of improper node of decomposition. To improve this performance, deeper levels of decomposition need to be calculated, but it increases the computational cost. The resonance band is also likely to be spread over multiple nodes in variable speed conditions. Thus appropriate node selection criteria need to be applied. CDSP, on the other hand, achieves 100% accuracy as it is made adaptive with the varying rotational speed.

6.2.3 BITS-2021

Based on the publicly available datasets, similar sets of experiments are performed during this research. The details of these experiments are discussed in Chapter 3. The analysis is carried out under the framework discussed in Fig. 6.1.

Two sets of experiments are performed for testing the proposed algorithms. In the first set, vibration data from bearing under test (BUT) is acquired using an accelerometer mounted on the housing of that bearing. These experiments include both constant as well as variable speed conditions. The proposed CDSP (Method 5) is found to correctly classify all the cases, as shown in Fig. 6.5-6.6. The fractional enveloping also outperforms the wavelet and VMD based methods. All the outer race fault cases are easily diagnosed using each of these methods. Lower size (0.6 mm) inner race faults are somewhat difficult to diagnose and can be correctly diagnosed only by Method 5 at lower confidence. Large size faults comparatively easy to detect.

Similar trend is observed in second set of experiments using microphone. Sound is, conventionally, an indicator of sever fault in a machine. Sound signal is prone to surrounding noise and is difficult from signal processing perspective. This is reflected in the results of this set. CDSP is still the best performing method, but its overall accuracy is slightly dropped.

6.3 Comparison of Computational Time

Computational time is an important aspect of real-time, online fault diagnosis. Critical machines are often continuously monitored to detect any presence of fault as early as possible. If the diagnosis algorithm is computationally demanding, this task of continuous processing of machine signals becomes cumbersome. It is always desirable for the fault diagnosis algorithm to have less computational requirements.

For the proposed analysis framework, the average time taken by each method for processing a 10 sec, 200000 sample signal (sampling rate of 20 KHz) is calculated. The time taken by the conventional Hilbert transform is the least (3.37 sec), whereas proposed fractional enveloping takes highest time (142.87 sec). The proposed cumulative distribution sharpness profiling (CDSP) takes around 41 sec, which is better compared to the wavelet and variational mode decomposition based methods. The wavelet analysis takes around 98 sec for level 3 decomposition because of the calculation of multiple time-frequency ridge curves extraction (MTFCE) at each node. If kurtosis or some other sparsity index is used for selection of the optimum node, then wavelet analysis becomes much faster, but the accuracy reduces.

6.4 Conclusion

The proposed CDSP and MTFCE framework is faster and has better accuracy compared to the Hilbert transform, wavelet transform, variational mode decomposition and fractional enveloping. This framework does not free from training as opposed to the machine learning algorithms. Also, the classification is based on simple thresholding of diagnosis percentages. This makes the CDSP, MTFCE framework suitable for automatic fault diagnosis under constant as well as varying speed conditions. Next chapter gives concluding remarks on the thesis by discussing the motivation, need, limitations and future scope of the proposed approaches.