Wavelets

Robert X. Gao • Ruqiang Yan

Wavelets

Theory and Applications for Manufacturing

Springer

Robert X. Gao
Department of Mechanical Engineering
University of Connecticut
Storrs, CT 06269-3139, USA
rgao@engr.uconn.edu

Ruqiang Yan
School of Instrument Science and Engineering
Southeast University
Nanjing, Jiangsu
China, People's Republic
ruqiang@seu.edu.cn

ISBN 978-1-4419-1544-3 e-ISBN 978-1-4419-1545-0
DOI 10.1007/978-1-4419-1545-0
Springer New York Dordrecht Heidelberg London

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Since the publication of Alfred Haar's work on orthogonal function systems a hundred years ago, the world has witnessed a tremendous growth in the theory and practice of *wavelet*, even though a reported, systematic study of the subject field and its applications to engineering did not occur until the 1980s. Over the last 3 decades, a plethora of literature has been published, describing advancement in the wavelet theory and its successful applications in various fields of engineering: from image processing in biomedical engineering to signal processing in meteorology to bridge monitoring in civil engineering. The adaptive, multiresolution capability of the wavelet transform has also made it a powerful mathematical tool for the diagnostics of equipment operation conditions in manufacturing, such as tool breakage.

Past research on wavelets has been translated into a large volume of publications and significantly impacted the state-of-the-technology. These papers, together with a series of classic books, have taught generations of engineers the theory and applications of wavelets. Nevertheless, there exists a gap in the literature that is particularly dedicated to graduate students and practicing engineers in manufacturing who are interested in learning about and applying the theory of wavelet transform to solving problems related to equipment and process monitoring, diagnostics, and prognostics in manufacturing.

The book is intended to bridge such a gap by presenting a systematic yet easily accessible treatment of the mathematics of wavelet transform and illustrate, in concrete terms, how wavelet transform as a mathematical tool can be realized for applications in manufacturing. Contributing to the understating and adoption of wavelets by the manufacturing community is the primary motivation for this book, and the 12 chapters included herein provide an overview of some of the latest efforts in this vibrant field.

To establish a common ground for the treatment of signals, which is the focal point of this book, Chap. 1 starts by introducing a general classification scheme of signals typically countered in mechanical systems, from the point of view of their statistical behavior – deterministic and nondeterministic signals. Using a mass–spring–damper system as a physical embodiment, the analytical expression, waveform, and solution of deterministic signals are first illustrated. These are then contrasted against the nondeterministic family of signals, and the concept of nonstationary, which provides the fundamental motivation for dedicating this book to wavelet transform in manufacturing, is introduced. Taking signals measured in two representative manufacturing processes as a realistic example, the link between manufacturing and signal processing, as well as the need for properly treating nonstationary signals, are established, motivating the dedication of the book to this subject matter.

Chapter 2 reviews several major events occurred in the field of signal processing since the invention of the Fourier transform in the nineteenth century, thereby recognizing the historical significance of spectral analysis. Such events have initiated and accompanied the conceptualization, formulation, and growth of the theory of wavelet transform. Based on the concept that signal transformation (for revealing the information content of the signal) can generally be represented by a convolution operation between the signal and a known *template* function, we sought to illustrate the common ground shared by the Fourier transform as well as its enhanced version (the short-time Fouriertransform), which has a fixed length of the analysis window, and the wavelet transform, which features an analysis window of variable length.

The next three chapters, Chaps. 3–5, are devoted to introducing the fundamental mathematics involved in understanding what wavelet transform is and does, and how to apply it to decompose nonstationary signals as typically encountered in manufacturing. Aware of the existence of many excellent books on wavelets and at the same time, the recognized need by many graduate students and practicing engineers for a step-by-step treatment of some of the mathematical procedures involved to implement the wavelet transform, in terminologies familiar to engineers, we tried to take a balanced approach when writing these chapters. Specifically, we introduced the continuous version of the wavelet transform in Chap. 3, by first drawing the resemblance between a continuous, sinusoidal wave and a time-localized wavelet that is essentially a linear, integral transformation satisfying the *admissibility* condition. To provide the readers with a handy access to some of the most often encountered properties of the continuous wavelet transform (CWT) in one place, we included descriptions of concepts such as superposition, covariance under translation and dilation, and the Mayol principle, together with a mathematical *proof*, for each of these properties. By providing detailed proofs, we wish to encourage readers who might have initially felt intimidated by the wavelet mathematics to gain some confidence in approaching the topic from a practical yet mathematically rigorous perspective, instead of resorting to a strictly recipe type of operations. We then proceeded to give a step-for-step procedure for implementing the CWT, in two ways, such that readers can see, in concrete terms, where all the background information finally leads to, in terms of performing CWT on some representative signals.

Chapter 4 introduces the discrete version of the wavelet transform, or DWT. The chapter is motivated by the recognition that CWT, while enabling a 2D decomposition of signals in the time–frequency (via the scale) domain with high resolution, is computationally complex due to the generation of redundant data. In comparison, the DWT is computationally more efficient, thus it is better suited for image compression and real-time applications. Using logarithmic discretization as an example, we first discussed how parameters are discretized to guarantee correct information retrieval as a result of the DWT process. Several derivation details are provided to illustrate the thought process. We then moved to the *dyadic* discretization method that allows for *orthogonal* wavelet basis to be constructed, based on the theory of multiresolution analysis (MRA). To satisfy readers who may be interested in knowing a bit more about the "why" and "how" related to MRA, we supplemented the explanation with several mathematical details, and illustrated why the process of DWT will lead to the generation of *detailed* and *approximate* information. In this context, we demonstrated that DWT, in essence, is about performing a series of low-pass and high-pass filtering operations, which can be implemented by following Mallat's algorithm. Mirroring the structure of Chap. 3, we presented several commonly used wavelets for DWT and illustrated how they can be used for applications such as signal denoising, by means of soft and hard thresholding.

While enabling flexible time–frequency resolution in signal decomposition, the relatively low resolution of DWT in analyzing the high-frequency region gives rise to the wavelet packet transform (WPT), which is the focus of Chap. 5. After a brief coverage of its definition and basic properties, two algorithms for implementing the WPT – the recursive algorithm developed by Mallat and a Fourier transform-based algorithm that leads to the harmonic wavelet packet transform – are introduced. We then illustrated how a signal's time–frequency composition of a vibration signal, which relates directly to the working state of manufacturing equipment, can be revealed by the WPT, and how WPT can be applied to removing Gaussian noise from a chirp signal. These applications exemplify how the enhanced resolution of WPT can provide an attractive tool for detecting and differentiating transient elements with high-frequency characteristics.

With the fundamentals of wavelet transform covered, Chaps. 6–8 describe several application scenarios where the effectiveness of wavelet transforms are demonstrated. The first application relates to signal *enveloping*, a technique commonly used for nondestructive testing and structural defect identification. Addressing the limitation of enveloping in requiring a priori knowledge for choosing the filtering band to extract a signal's envelope, an adaptive, multiscale enveloping method (MuSEnS) that is rooted in the wavelet transform is described in Chap. 6, which effectively overcomes the limitation. Taking advantage of the Hilbert transform in extracting the envelop of an *analytic* signal and the fact that performing wavelet transform on a signal using a complex-valued base wavelet will result in an *analytic* signal, the chapter illustrates how a signal's envelope can be readily calculated from the modulus of the corresponding wavelet coefficients. To illustrate the effectiveness of this technique in signal decomposition, two manufacturing-related applications – differentiation of ultrasonic pulses that are timely overlapped and spectrally adjacent for wireless pressure measurement in injection molding and bearing defect diagnosis in rotary machines – are demonstrated, using both experimentally measured signals and synthetic signals for quantitative evaluation.

While the localized signal decomposition capability of wavelet transform is particularly useful for transient events identification, the result of wavelet transform does not explicitly reveal distinct characteristic frequencies that are often times indicative of defective modes of a machine, e.g., the ball passing frequency at the inner raceway of a rolling element, when a localized spalling is present. In such situations, the effectiveness of wavelet transform can be leveraged by the Fourier transform in identifying a signal's frequency components. This leads to the formulation of a *unified* time-scale-frequency analysis technique that adds spectral post-processing to the data set extracted by wavelet transform, for enhanced defect diagnosis. In Chap. 7, we demonstrate such an integrated method, in the context of a *generalized* signal transformation frame. An expression for both the Fourier transform and wavelet transform in the generalized frame is first presented, establishing the basis for crossdomain unification of the two transforms. Next, the viability of postspectral analysis of wavelet processed data is analytically justified, and the effectiveness of the technique in identifying the bearing defects under various operating conditions is demonstrated.

A question that naturally arises upon defect detection is the severity of the defect, which affects the proper scheduling of maintenance. To answer this question, we demonstrate in Chap. 8 how WPT can be applied in classifying machine defect severity, using vibration signals from rolling bearings as an example. We start the discussion by associating *features* (e.g., energy content or Kurtosis value) of a signal with the subfrequency bands of its decomposition, enabled by the WPT, and demonstrate how WPT can flexibly extract features from the subfrequency bands of the decomposed signal where the features are concentrated. The chapter further contains a discussion on how to process the features, once they are obtained, for classification purpose. Relevant techniques for selecting best-suited features using the Fisher linear discriminant analysis and principal component analysis, and classifying features to quantify defect severity levels are described. Two case studies presented toward the end of the chapter on ball and roller bearings confirm the validity of WPT for defect severity classification.

Chapter 9 continues the discussion on signal *classification*, with a focus on how it can be applied to differentiate different working conditions of a machine, for the purpose of diagnosis. The concept of *discriminant* features is first introduced, and a technique called the local discriminant bases (LDB) is described in detail. In a nutshell, the LDB algorithm determines an optimal set of wavelet packet nodes, each of which corresponding to a wavelet packet basis, to represent signals acquired under different machine states as different *classes*. Similar to the Shannon entropy feature introduced in Chap. 5 for signal compression, several features suited for diagnosis of rotating machines, e.g., relative entropy or correlation index, are identified in this chapter. We provided a step-by-step description of the LDB algorithm, for readers to see how the algorithm can be implemented. Using three synthetic signals with added white noise and vibration signals measured on a gearbox under different states of wear, we quantitatively demonstrated how the wavelet packet bases constructed using the LDB algorithm can more successfully differentiate and classify these signals than without using the LDB.

Given the abundance of the base wavelets in the published literature, it is natural to ask the question as to how to choose an optimal base wavelet for analyzing a particular type of signal. This is based on the understanding that (1) the choice of base wavelet made in the first place will affect the result obtained at the end, and

(2) each base wavelet may be developed for different purposes and emphasis; therefore, an educated approach to their selection is needed when solving a specific type of engineering problems. In this book, we have tried to address this issue of significant intellectual interest in two ways. First, in Chap. 10, we introduced a general strategy for base wavelet selection, using both *qualitative* measures (e.g., orthogonality and compact support) and *quantitative* measures (e.g., Shannon entropy and discrimination power). Subsequently, we presented several criteria for base wavelet selection, including the *energy-to-Shannon entropy ratio* and the *maximum information measure*. Using both real-valued and complex-valued base wavelets, we demonstrated how these criteria can be applied to selecting the best-suited base wavelet from a pool of candidates to decompose both a numerically formulated Gaussian-modulated sinusoidal test signal and a vibration signal measured on a defective ball bearing, thus confirming the effectiveness of these criteria.

Besides investigating how to choose an appropriate base wavelet from the existing library, another approach is to design a *customized* wavelet that is adapted to a specific type of application to maximize the degree of matching with the signal of interest, thus improving the effectiveness of feature extraction. Such a complementary technique is the focus of Chap. 11. After reviewing the fundamental issues involved in the wavelet design process and several customized wavelets, we described in detail the process of designing an *impulse* wavelet for bearing vibration analysis, based on the impulse response of the mechanical structure where the bearing is housed. The importance of satisfying the *dilation* equation to avoid information loss in the signal reconstruction is stressed, and the procedure of meeting this requirement is illustrated. Using the designed impulse wavelet, vibration signals from a defective bearing are analyzed, and the result is compared with that from using five standard wavelets available in the library, using the signal-to-noise ratio for the defect-characteristic frequency as the measure for comparison. The good performance of the impulse wavelet confirms the validity of the analytical procedure described in developing customized wavelets for enhanced signal analysis in a broad range of applications in engineering.

The last chapter of the book provides a brief survey of some new advancement reported in recent years that goes beyond the classical wavelet transform. These latest developments address some of the fundamental limitations inherent to the wavelet transform, e.g., when it is used to analyze signals of finite length and/or limited duration, or for capturing and defining image boundaries. We started the survey by introducing the second generation wavelet transform, or SGWT, which uses the so-called *lifting scheme* to replace the traditional mechanism of wavelet construction that uses translation and dilation. Major operation steps for realizing the SGWT, such as splitting, prediction, and updating, are described, and the effectiveness of the technique for separation and reconstruction of an intermittent linear chirp signal is demonstrated. Addressing the inherent limitations of classical wavelets (e.g., *isotropic*) and the specific challenges in image processing (e.g., in resolving image boundaries), we then introduced the ridgelet and curvelet transforms. The former was developed to address the need for analyzing anisotropic features in images, whereas the latter enables improved representation of curved boundaries in images. For each transform, we have presented the definition and basic properties, and demonstrated a representative application in manufacturing.

As is true with any book, the writing reflects upon the authors' understanding of and knowledge about the subject matter. While we have strived to present to the readers a composition that is both rigorous in the mathematical treatment and relevant in the examples chosen to complement the theory, it is inherently difficult for a work of this size to be completely free of errors. We bear the responsibility for anything that is not correctly stated in the book and would greatly appreciate hearing from our readers about any mistakes they found such that we can correct them in future printings.

We thank the anonymous reviewers for their insightful and constructive comments that have both sharpened our thinking and provided clues to improving the pedagogic presentation. We are indebted to former graduate students at the Electromechanical Systems (EMS) Laboratory, whose intellectual contributions have made the book a reality. In particular, we thank Drs. Brian Holm-Hansen, Changting Wang, and Li Zhang, who dedicated a substantial part of their doctoral research to the study of wavelet for diagnosis in manufacturing equipment and processes, and Dr. Qingbo He, who spent a year as a postdoctoral research fellow at the EMS Laboratory, working on the characterization of physical activities, for their dedication to and enthusiasm in exploring the world of wavelets, which has laid the foundation for this book. We also thank the US National Science Foundation for funding a number of relevant projects, which has allowed us to systematically study this fascinating subject.

This book-writing project was initially planned to be completed in 1 year, but a number of events that took place in between have delayed the writing and made the time needed for ultimately completing the book nearly twice as long. We take this opportunity to express our sincere appreciation to the publisher for supporting this project. In particular, we thank Mr. Stephen Elliot, Senior Editor for Engineering, and Mr. Andrew Leigh, editorial assistant, for their earnest cooperation, editorial assistance, and above all, patience, which has created a relaxed environment for us to finish the writing while juggling many other deadline issues. Last but not the least, we sincerely thank our respective families for their understanding and carrying the load for us during the course of this project so that we could devote as much time as possible to the writing of the book. It is our hope that the book is worth their selfless support, and our readers will find in it something that is of value to their research.

Contents