The State-of-the-Art on Framework of Vibration-Based Structural Damage Identification for Decision Making

Xuan Kong, Chun-Sheng Cai * and Jiexuan Hu

Department of Civil and Environmental Engineering, Louisiana State University, Baton Rouge, LA 70803, USA; xuankong119@gmail.com (X.K.); jhu11@tigers.lsu.edu (J.H.)

* Correspondence: cscai@lsu.edu; Tel.: +1-225-578-8898

Academic Editor: Dimitrios G. Aggelis
Received: 10 April 2017; Accepted: 4 May 2017; Published: 11 May 2017

Abstract: Research on detecting structural damage at the earliest possible stage has been an interesting topic for decades. Among them, the vibration-based damage detection method as a global technique is especially pervasive. The present study reviewed the state-of-the-art on the framework of vibration-based damage identification in different levels including the prediction of the remaining useful life of structures and the decision making for proper actions. This framework consists of several major parts including the detection of damage occurrence using response-based methods, building reasonable structural models, selecting damage parameters and constructing objective functions with sensitivity analysis, adopting optimization techniques to solve the problem, predicting the remaining useful life of structures, and making decisions for the next actions. For each part, the commonly used methods were reviewed and the merits and drawbacks were summarized to give recommendations. This framework is aimed to guide the researchers and engineers to implement step by step the structure damage identification using vibration measurements. Finally, the future research work in this field is recommended.

Keywords: vibration-based damage identification; response-based method; model-based method; objective function; optimization algorithm; residual useful life; damage Prognosis; decision making

1. Introduction

Deterioration and degradation of structures are of great concerns worldwide. The deficient structures can potentially endanger the safety and economical use of the infrastructure system, and even develop into structure failures if the damage is not detected in the early stage. The research need on detecting damage at the earliest possible stage is universal throughout the fields of civil, mechanical, and aerospace engineering. Most non-destructive damage detection methods can be categorized as either local or global techniques. The non-destructive testing (NDT) methods, such as the ultrasound, X-ray, dye penetrates, magnetic particle, and acoustic emission [1,2], are local techniques that can complement visual inspections as objective methods leading to quantifiable results. However, all these experimental methods require that the vicinity of the damage is known a priori and that the portion of the structure being inspected is readily accessible. Subjected to these limitations, the local NDT methods can only detect damage on or near the surface of the structure. The need for global damage detection methods that can be applied to complex structures has led to the development of vibration-based methods.
The research on the vibration-based damage identification was initiated in the late 1970s, particularly in the aerospace and offshore oil industries, and has been rapidly expanding to civil engineering industries over recent decades [3]. The basic idea behind this technology is that modal parameters (e.g., natural frequencies, modal shapes, modal damping, etc.) are functions of the physical properties of the structure (e.g., mass, damping, and stiffness) [4]. Therefore, changes or damages in the physical properties will cause detectable changes in the modal properties. There have been a few review publications on this topic. Doebling et al. [4] presented an extensive review of vibration-based damage identification methods up to 1997. Sohn et al. [5] then presented an updated version containing new technical developments up to 2001. Carden and Fanning [3] presented a review with particular emphasis on structural engineering applications. Yan et al. [6] introduced the development of modern-type vibration-based methods using modern signal-processing techniques and artificial intelligence. Wang and Chan [7] reviewed the recent development in statistical damage detection and condition assessment techniques for bridge structures. Fan and Qiao [8] comprehensively reviewed the modal parameter-based damage identification methods for beam- or plate-type structures and conducted a comparative study of five algorithms. Moughty and Casas [9] reviewed the developments with particular focus on the utilization of advanced computational methods. Das et al. [10] presented a review through a comparative study among different vibration-based damage detection methods. All the above literature reviews concentrate primarily on the damage identification Levels 1 to 3, i.e., determination of damage occurrence, location, and severity. However, they did not cover research topics such as the prediction of the remaining service life of structures and the decision making for maintenance, which is the ultimate objective of damage detection.

Therefore, this study is aimed to more systematically review the state-of-the-art on the framework of vibration-based damage identification in different levels, that is, the damage occurrence, damage location, damage severity, remaining life prediction, and decision making, to give a general guideline for researchers and engineers. Several major parts of the framework are to be discussed: (1) detecting damage occurrences using response-based methods; (2) building reasonable structural models; (3) damage characterization, objective functions, and optimization algorithms; (4) remaining life prediction; and (5) decision making.

2. Framework of Vibration-Based Damage Identification

Generally speaking, damage is defined as changes introduced into a system, which adversely affects the current or future performance [4]. The definition implies that it is only meaningful when two different states of the system are compared. In addition, the damage in structural and mechanical systems is referred to changes to the material and/or geometric properties of these systems and changes to the boundary conditions and system connectivity [11].

Damage identification is always carried out in conjunction with five closely related disciplines including structural health monitoring (SHM), condition monitoring, non-destructive evaluation (NDE), statistical process control (SPC), and damage prognosis [11]. SHM is a process of implementing a damage identification strategy in structures, focusing on the online global damage identification. Similar to SHM, condition monitoring mainly addresses damage identification in rotating and reciprocating machinery. NDE is primarily used for damage characterization and as a severity check in an offline local manner after the damage location is determined. SPC is a process using a variety of sensors to monitor changes due to structural damage. Once the damage has been identified, damage prognosis is adopted to predict the remaining useful life of a system.
The main objective of damage identification is to identify structural damage at a very early stage and predict the remaining service life of structures for the decision making. Based on the amount of information provided regarding the damage state, the structure damage identification can be classified into four levels [12]: determination of damage occurrence (Level 1), determination of damage geometric location (Level 2), quantification of damage severity (Level 3), and prediction of the remaining service life and decision making (Level 4). The vibration-based methods can be classified as response-based and model-based methods. As for the response-based method, by directly or indirectly interpreting the structural dynamic response, the changes of modal parameters between the intact and damaged states of the structure are used to identify damage in structures. Model-based methods assume that the structure responds in some predetermined manner that can be simulated by an analytical or numerical model such as the usually used finite element model. By comparing the updated model correlated with the measured data of the damaged structure to the original model, the damage location and severity can be identified. In terms of identification levels, the response-based detection methods are mostly able to detect damage in Level 1 and in some cases in Level 2, which is sufficient for many practical applications. Challenges on this level include how to select sensitive features, detect damage in a rather early stage without getting false alarms, eliminate the environmental and operational effects, and quantify the probability of detection [13]. In order to identify damage in Level 2 and Level 3, the model-based methods are usually adopted. In most cases, the location and extent of damage are determined simultaneously. The most sophisticated level is Level 4, i.e., predicting the structural remaining lifetime and making decisions based on damage diagnosis results. Level 4 is generally associated with the fields of fracture mechanics, fatigue-life analysis, and structural assessment, which requires the combination of global structural models with local damage models to predict the evolution of damage, or the probabilistic failure models based on previously failed or damaged components [14], etc.

Figure 1 demonstrates the procedure of damage identification for the four levels. The first step is to take measurements from the structure and apply the response-based methods to determine the occurrence of structure damage. After the structure damage is confirmed, a structure model is created using the finite element or other methods. The model should be verified first to reflect the structural behavior and the global parameters of the model are then updated to assure the applicability of the model. To determine the damage location and dimension, a proper damage index or parameter should be carefully selected and an objective function needs to be constructed with the purpose to minimize the difference between the measured response and the numerical prediction. Herein the sensitivity analysis is used to ensure the sensitivities of damage parameters and objective functions. Once the relationship between the damage parameter and the objective function is established, an effective optimization algorithm is introduced to solve the problem and obtain the solution. The optimization solution also needs to be verified to make sure that the result is physically meaningful. The above procedure could be repeated several times to obtain a reasonable result. Once the damage is quantified, the remaining life of the structure can be assessed and an action can be recommended for the decision makers.
3. Detection of Damage Occurrence

The first level of damage identification is to determine the damage occurrence in structures according to the global change of structural integrity. The most popular and effective method is the response-based method that directly or indirectly interprets the measured response data of the structure. Usually, the data available are the experimental data before and after damage occurred, and the occurrence of damage is determined through finding their difference. The vibration responses can be easily measured by a variety of sensors, for example, accelerometers, velocity transducer, displacement sensors, strain gauges, etc., and the measured time domain data can be converted into frequency or modal domain data by transforming techniques. Hence, the response-based damage detection methods can be grouped into three types in terms of the response data used: time domain, frequency domain, and modal domain.

3.1. Time Domain Methods

It is well known that the measurements are always made in the time domain, and therefore it is straightforward to use the time domain response for damage detection. The measurement contains
abundant information about the structure and preserves the possible nonlinear behavior as well. In addition, it eliminates the need to perform frequency transformation and the associated errors such as leakage and truncation, and thus enhances the accuracy and practicality of the time domain approaches [15]. Fassois and Sakellariou [16] reviewed the time-domain methods used in the context of structural diagnosis and vibration problems.

3.1.1. Output-Only Methods

In order to identify the structure dynamic parameters from the time domain data, at the early stage several techniques were proposed in the time domain on the basis of the experimentally measured responses of the structure: the Ibrahim method [17], the random decrement method [18], the least-squares complex exponential method [19], and the eigensystem realization algorithm [20]. Cattarius and Inman [21] used the beat phenomenon to characterize small differences. Peeters and De Roeck [22] reviewed the stochastic system identification methods used to estimate the modal parameters of vibrating structures in operational conditions, including the covariance-drive stochastic subspace identification (SSI-COV) and data-driven stochastic subspace identification (SSI-DATA) methods. Later, Loh et al. [23] compared the SSI-COV with the second-order blind identification (SOBI) approaches to identify the dynamic characteristics of the bridge directly from the vibration measurement. Loh et al. [24] also introduced the recursive stochastic subspace identification (RSSI) technique to identify the natural frequencies and Ruocci et al. [25] employed the kernel density estimation (KDE) to characterize the correlation between the vibration signatures acquired in the time domain. These methods eliminate the need of measuring or identifying inputs, and thus are valuable tools for identifying modal parameters from time domain measurements. The extracted vibration features and modal properties can then be used for detecting damage occurrence and/or location by comparing the identified modal properties with the original values. Meanwhile, there are methods that directly detect damage based on the measured data only, for example, using the moving time window technique to trace abrupt changes in the structural parameters [26], detecting novelties in the recorded seismic response by using expansion in a basis of bi-orthogonal wavelets [27], detecting damage based on changes in wave travel times between selected sections of a structure [28,29].

3.1.2. Modern Signal Processing Methods

In recent years, many signal processing techniques and artificial intelligence as analysis tools have been introduced to investigate the vibration signals and further extract features to represent the signal characteristics. Those methods, such as wavelet analysis, empirical modes decomposition (EMD), and Hilbert-Huang Transform (HHT), are more universal and less dependent on the structural shape. The main concept of using wavelet analysis is to break down the dynamic signal of a structural response into a series of local basis function called wavelets, so as to detect the special characteristics of the structure using scaling and transformation property of wavelets [30–33]. The method based on the EMD and HHT is intended to extract damage spikes due to a sudden change of structural stiffness from the measured data thereby detecting the damage occurring time instants and locations, and determining the natural frequencies and damping ratios of the structure before and after damage [34]. Zhou and Zhang [35] developed a recurrence plot-based fault detection method by integrating the statistical process control technique to analyze the vibration signal for damage detection.

3.1.3. Data-Based Statistical Methods

Structural damage affects the dynamic properties of a structure resulting in a change in the statistical characteristics of the measured responses. Over the past two decades, the data-driven time series methods utilizing statistical signal processing techniques have formed an important and rapidly evolving class of vibration based methods [16,36]. These methods rely on the well-established statistical concepts instead of human expertise to extract features that change with the onset of damage [37], which eliminates possible individual biases and requires very few assumptions regarding the physical
structure. This method is represented by the Auto-Regressive (AR) family method such as the AR or Auto-Regressive Moving Average (ARMA) models [37–40], AR model with exogenous inputs (ARX) [41,42], Auto-Regressive Moving Average Vector (ARMAV) model [43], vector autoregressive (ARV) models [44], and the time frequency autoregressive moving average (TFARMA) model [45]. The ARMA or other similar models are used to model the time-domain vibration signals obtained from the structure, and then the model coefficients estimated using statistical methods are used to identify the system dynamic parameters and to extract features that indicate the damage occurrence. The AR approach is solely based on signal analysis of the measured vibration data, making this approach very attractive.

3.1.4. Methods for Changing Environmental Conditions

In order to distinguish the influences on the signal driven by damage itself from those driven by environmental changes, the AR model is also used for damage detection under changing environmental and operational conditions [46–48]. The effect of environmental conditions has been investigated [49] and various methods are adopted to remove this effect, such as principal component analysis (PCA) [50,51], factor analysis [52], nonlinear PCA based on unsupervised support vector machine [53], the Cointegration concept for non-stationary time series [54], PCA method in time-varying systems [55], etc. Moreover, the novelty detection and outlier analysis are used to perform the lowest level of damage identification in operational conditions [56,57], and the multivariate statistical analysis method becomes popular for automatically revealing the damage existence using vibration data under changing environmental and operational conditions [58–60].

3.2. Frequency Domain Methods

Frequency domain responses are obtained from time series responses by non-parametric estimation and signal processing techniques, relying on the application of the Fourier transformation. The transformation can drastically reduce the volume of data and compensate the loss of information by averaging the effects of random noise [25].

One of the most common methods is based on the frequency response function (FRF) that is a ratio between the structure response such as displacement, velocity, or acceleration in the frequency domain to an applied force that is a known harmonic force generally. The FRFs are used as the data source to estimate modal parameters through curve fitting methods. The commonly selected features are the shift in the resonance and anti-resonances or changes affecting the amplitudes. By applying FRF data directly in damage detection problems, Maia et al. [61] compared the method based on FRF curvatures versus the method based on modal shape curvatures, and illustrated with a lumped-mass system and a real bridge with experimental data [62]. The main drawback is that this FRF curvature method works better for a frequency range before the first resonance or anti-resonance, whichever comes first. The FRF curvature method has been successfully studied in the simulation and experiment [63,64]. Meanwhile, the direct use of FRF shapes has been studied by many researchers. For example, Liu et al. [65] made use of the residual imaginary parts of FRF shapes to locate the structural damage and the abnormality caused by damage in a cantilever beam. Salehi et al. [66] presented a damage detection technique based on both the real and imaginary parts of measured FRF shape. Mohan et al. [67] evaluated the use of FRFs with the help of the Particle Swarm Optimization technique. Dilena et al. [68] presented the interpolation damage detection method using FRF measurements. Lee and Kim [69] proposed the signal anomaly index to express the amount of changes in the shape of FRFs or strain frequency response function (SFRF). Kim and Eun [70] compared the FRF-based approach extracted from dynamic measurements of a truss structure and the flexibility-based approach extracted from the modal data, and showed that the FRF-based approach can be utilized more explicitly than the flexibility-based approach. The FRF based method is very promising because the FRF data can be directly obtained without any further data extraction and processing. The main advantage of using the FRF shape in damage detection is that it contains higher frequency information, while the mode shape discussed later constructs a truncated frequency space.
The FRF requires the simultaneous measurement of both the excitation loads and the structural responses. When the input force is unknown, the computation of the FRF is substituted by the power spectral density or the transmissibility functions. The transmissibility that is calculated from the FRFs of a system is a transformation matrix between two sets of output responses in the frequency domain. The motivation of using transmissibility for damage detection relies on the fact that the transmissibility is a local quantity, suggesting a higher sensitivity than the modal parameters to detect structure changes [71]. For example, Worden et al. [72] detected damage using the novelty detection algorithm based on the measured transmissibility. Canales et al. [73] studied the output-only modal identification using transmissibility under variable load conditions. Devriendt et al. [74–76] used transmissibility to identify modal parameters from output-only transmissibility measurements and multivariable transmissibility measurements. Chesne and Deraemaeker [77] reviewed the state-of-the-art of using transmissibility functions for damage localization of mechanical systems. Kong et al. [78] studied the feasibility of detecting bridge damage using transmissibility of the vehicle response in a vehicle-bridge coupled system. Zhou et al. [79] proposed to detect structure damage employing transmissibility functions compressed by principal component analysis. Some other output-only methods in the frequency domain are developed based on auto/cross-spectral densities, such as the peak picking method, the frequency domain decomposition techniques (FDD) and the enhanced technique [80].

In addition, Yang et al. [81] proposed a frequency contour lines method to identify damage in plate-type structure and the Fourier spectrum-based strain energy damage detection method for beam-like structures in noisy conditions [82]. Cao and Ouyang [83] combined the joint approximate diagonalization (JAD) technique and gapped smoothing method (GSM) to form a sensitive and robust damage index in frequency domain.

3.3. Modal Domain Methods

Modal information can be extracted from input-output measurements by means of the classical modal analysis methods or from output-only data measured under the ambient excitation from wind, traffic loads, etc. without the artificial forces. The modal domain methods evolve along with the rapid development of experimental modal analysis technique and gain the popularity because the modal properties (i.e., natural frequencies, modal damping, modal shapes, etc.) have physical meanings. Thus, they are easier to be interpreted or interrogated than those abstract mathematical features extracted from the time or frequency domain. Response-based damage detection methods using modal information are relatively straightforward. That is, the changes between modal parameters of the intact structure and those of the damaged structure are directly used, or correlated with other relevant information, to develop the damage indicators for locating damage in the structure.

The modal domain provides a large set of damage sensitive features, such as the shifts of natural frequencies [84], changes in the Modal Assurance Criteria (MAC) that correlates modal shapes of the damaged and undamaged structures [85], changes in the Coordinate Modal Assurance Criteria (COMAC) [86], changes in the Multiple Damage Location Assurance Criterion (MDLAC) [87], changes in modal strain energy [88,89], changes in modal shape curvature [90], strain modal shapes [91], dynamic flexibility matrix [92], and dynamic flexibility curvature [93]. Sinou [94] presented a comprehensive review on the direct use of modal parameters for damage detection in linear systems. Since the effect of small damage on modal properties, such as natural frequency shifts and local changes in the mode shapes, are likely to be masked by experimental uncertainties and data reduction [95], the response-based methods that directly use modal parameters are not efficient to detect small damage.

In order to enhance the sensitivity and accuracy of damage detection in modal domain, a simple and effective approach is to use the higher-order derivatives of modal shapes, such as rotations (first derivative), curvatures (second derivative) and third and four derivatives. The behavior of higher-order derivatives has been numerically studied and experimentally validated in damaged beam-like structures [96,97]. Especially, the mode shape derivatives have been widely used to localize damage in laminated composite structures. For example, Cao et al. [98] explored the synergy between a wavelet
transform and a Teager energy operator (TEO) to ameliorate the curvature mode shape, and also used
two-dimensional curvature mode shape method to detect damage in plates [99]. Moreno-Garcia [100]
proposed a Ritz based method to improve the computation of mode shape derivatives.

In addition, the incorporation of advanced signal processing procedures has been widely studied.
Ratcliffe [101] applied a modified Laplacian operator solely on modal shapes of damaged structure
and successfully identifies the damage location. Chandrashekhar and Ganguli [102] used Gaussian
fuzzy sets to fuzzify changes in the modal curvatures due to damage and used a fuzzy logic system to
accurately locate the damage. Sazonov and Klinkhachorn [103] presented an analysis to determine
the optimal sampling interval of strain energy mode shapes. Yoon et al. [104] used a gapped smoothing
method to extract local features in vibration curvature shapes. Li et al. [105] proposed a fractal
dimension based method from the measured displacement mode shape in beam structures. In recent
years, the wavelet transform (WT) technique was used as an effective signal analysis tool to filter
out noise and to enhance robustness of damage detection. This approach involves the application of
wavelet analysis to modal data and shows high sensitivity to non-continuities and singularities in
the measurement signals. Cao et al. [106] developed a multi-scale pseudo-force model over wavelet
domain, in which WT was used as a multi-scale differential operator to acquire higher-order derivatives.
Katunin [107] presented an overview of the wavelet-based methods for structural damage identification.
Attractive features of these methods are the high sensitivity to damage and instant determination of
the damage location.

3.4. Summary of Response-Based Methods

In general, the response-based methods are straightforward, easy to use, and not computational
intensive, but at the same time they do not give a quantitative assessment of damage. It is a valuable
tool for damage identification in the preliminary stage. A guideline for using the response-based
methods is given in Table 1 based on the pros and cons of those methods.

Table 1. Guideline of Response-based Methods.

<table>
<thead>
<tr>
<th>Data Domain</th>
<th>Representative Methods</th>
<th>Advantages and Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time domain</td>
<td>Ibrahim method, Random decrem. ERA, ARMA family, Wavelet analysis, EMD and HHT, RSI, SSI-DATA, SSI-COV, KDE</td>
<td>Advantages: straightforward to use the time domain responses; eliminate the need to perform frequency transformation; preserved nonlinear behavior</td>
</tr>
<tr>
<td>Frequency domain</td>
<td>FRF shapes, FRF curvature, Transmissibility, Peak picking, FDD</td>
<td>Advantages: abundant information on structure dynamic behavior; contains frequency information in a wide range without further extraction and processing</td>
</tr>
<tr>
<td>Modal domain</td>
<td>Natural frequency, Mode shapes and curvatures, Modal strain energy, Strain mode shapes, Dynamically flexibility, High-order derivatives, Signal processing based</td>
<td>Advantages: modal properties are physically meaningful, easy to be interpreted or interrogated from output-only data and not require artificial forces</td>
</tr>
</tbody>
</table>

4. Structural Models

After the damage occurrence is confirmed and for a further detection, the model-based methods
are always introduced for localizing and quantifying the damage based on the measured responses.
To this end, it is necessary to have a structural model to represent the damage and the key characteristics of the structure system. Generally, two models are required, namely, a well-correlated model of the structure in its initial state, which is used as a reference model, and an updated model that reproduces the measured data under the damaged state through modifying the structural model parameters such as mass, stiffness, damping, etc. Comparisons of those two models provide the damage location and its severity. The structure models usually used for the model-based damage identification are as follows.

4.1. Structural Matrix

This method uses the stiffness, damping, or mass matrix of the structure to represent the structure and reproduce the measured results. To update the matrix of the structure, two ways are available. One way is the direct methods \[108,109\] that directly update stiffness and mass matrices of the elements in a one-step procedure to exactly reproduce the measured data, while the other way is the iterative methods \[110,111\] that use the sensitivity matrix of the parameters and based on a penalty function to maximize the correlation between the measured and analytical results. As for the direct methods, the updated results are seldom physically meaningful even though they are capable of accurately replicating the experimental measurements. With respect to the iterative methods, they preserve the structural connectivity in the updated model and provide a direct physical explanation for the updated results. However, the calculation of the sensitivity matrices is usually computationally intensive and the matrices are often ill-conditioned.

4.2. Finite Element (FE) Model

For large or complicated problems, finite element model is typically used. In order to acquire a reasonable FE model, the FE model updating technique is usually adopted to correct the FE model when the FE prediction is in conflict with the test results. One common approach is to consider an objective function that quantifies the differences between the analytical and experimental results, and to minimize the discrepancy by adjusting the unknowns of the FE model \[112\]. Mottershead and Friswell \[113\] gave a thorough review on model updating methods using structural dynamic responses. A major issue in FE model updating process is the possible ill-conditioning since there is no unique solution to the inverse problem \[114\]. There are several methods proposed to deal with this issue, such as the Bayesian approach \[115\], regularization \[110\], regularized Lanczos method \[116\], etc.

4.3. Mapping Model and Meta-Model

For a large structure with enormous elements, a simplified model is always necessary in the inverse problem. One type of models is the mapping model that is used to build a mapping relationship between the structural parameters and the structure responses. The representative method is the neural network (NN) method that is inspired by the architecture of the neurons and the human brain function \[117\] and has been widely applied in structural analysis due to its strong nonlinear mapping ability. The NN method is found capable of replacing conventional finite element updating methods in solving the inverse problems without utilizing the sensitivity matrix. However, the training process could be very time-consuming, and the accuracy of the predictions of this trained NN model is very dependent on the training data. Meanwhile, the other type of models is the meta-model (surrogate model) that replaces the structure model with an approximation model. This model is traditionally a polynomial function describing the relation between the structural responses and the model parameters. Meta-models have been extensively used in the design and optimization of computationally expensive problems. Different meta-models have been shown to perform well in different conditions, such as the response surface method (RSM) \[118–125\], Kriging method \[126,127\], radial basis function method (RBF) \[128\], and feed-forward neural network \[129\]. Li and Padula \[130\] and Queipo et al. \[131\] recently reviewed different surrogate models used in the aerospace industry, and Chen et al. \[132\] summarized the characteristics of different meta-model types and the experiment designs for fitting them. It is
recommended to use RBF for high-order nonlinear problems, Kriging for low-order nonlinear problems in high dimension spaces, and polynomial RSM for low-order nonlinear problems [133].

Although there are many methods available for structural models, the construction of those models usually results in model errors from simplified assumptions. In order to detect damage rather than the artificial errors from the model construction, a good quality model that could accurately predict the behavior of the structure is required, but is usually a quite challenge. As for damage detection in different levels, the complexity of the structure model and corresponding damage parameters depends on the structure type and required detection level. In the preliminary stage, a simplified FE model might be sufficient. If the damage is already localized in a limited range, a mapping/meta model or a detailed model could be built for the vicinity of the damage while other part of the structure could be simplified. The substructure method is a good alternative to build a fine model for the vicinity of the damage [134]. The compromise between the accuracy and simplicity of the model specifically depends on the problem.

The above conventional model-based methods always require a baseline data of vibration response from a healthy structure. However, in many cases the baseline model cannot be obtained for structures such as those with preexisting damage. Therefore, many studies have been devoted on the baseline-free method with no information about the original undamaged structure. Ratcliff [101] introduced a modified Laplacian operator for non-baseline damage detection from the mode shape data. Sazonov et al. [135] suggested a method to extract localized changes (damage peaks) from strain energy modal shapes based on Fourier analysis of the strain energy distribution. Sohn et al. [136] developed an instantaneous damage detection scheme based on time reversal acoustics and consecutive outlier analysis. Kim and Sohn [137] developed a guided-wave-based nondestructive testing to detect crack damage in a thin metal structure without using prior baseline data. Zhong and Oyadiji [138,139] presented a crack detection method based on finding the difference between two sets of detail coefficients obtained by using the wavelet transform of two sets of mode shape data of a cracked beam. Rucevskis et al. [140] described a method that requires only the mode shape curvature data of the damaged plate-like structures, in which the damage index is defined as the absolute difference between the measured curvature of the damaged structure and the smoothed polynomial representing the curvature of the healthy structure. In addition, the wavelet-based method is a major type of baseline-free method using the advanced signal processing techniques without the reference of healthy structure [107,141,142].

5. Damage Characterization, Objective Functions, and Optimization Algorithm

In addition to the structure model, there are three key components for the model-based damage detection method: the damage characterization, the selection of objective features and functions, and the optimization algorithms to solve the problem.

5.1. Damage Characterization

The first important step of a model-based method is the parameterization of the damage. The effect of damage on the structure model is introduced with a scalar variable or index as [143],

$$d_e = 1 - \frac{k^d_e}{k_e}$$

where $d_e$ is the damage index for a certain element, $k_e$ is the structure property value of the element in its intact status, and $k^d_e$ is the value of the element damaged. With respect to a single damage or discrete ones, the stiffness matrix of the element is assumed to decrease uniformly [144]. As for distributed damage, instead of adjusting the stiffness properties of all the elements separately, Teughels et al. [145] proposed the damage function to reduce the number of unknown variables and to determine the stiffness distribution over the FE model. The damage parameters are the multiplication factors of the damage functions. This method is useful for structures whose damage pattern can be represented by a reduction factor of the element stiffness.
The parameter definition and selection depend on the level of damage identification and uncertainty level. In the preliminary level, the parameters can be defined as global physical properties such as the Young’s modulus, material density, section area, and the moment inertia since they are directly related to the element stiffness. In this level, several elements can share the same parameter value or one super element can be used in the case that several close elements are expected to have similar values [146, 147]. As the identification becomes quite local, the damage parameters could be the width of cracks, the stiffness of joints, and physical properties of the uncertain elements. When the damage is located at one element, this specific element can have multiple parameters by dividing the element into small ones with different properties. The parameter selection requires a considerable physical insight into the target structures. In general, only sensitive parameters are selected, otherwise the process may be ill-conditioned since insufficient information is available to estimate the parameters accurately.

5.2. Objective Features and Functions

Since damage detection becomes an optimization problem eventually, the capability of a damage identification method strongly depends on the selected features and constructed objective functions that have a profound impact on the behavior of the optimization algorithm. Usually, we do not directly use the raw data in the time domain to compare two system states. Instead, we extract so-called features, for example, mean values, variances, maximum values, crest factors or other statistical characteristics of signals, correlation functions, spectral densities, frequency response functions, impedance spectra, eigenfrequencies, modal shapes, or any other features [13]. Meanwhile, the objective function is normally built up using the residuals between the measured responses and numerical predictions, such as a sum of squared difference [148] as,

\[
f(\theta) = \frac{1}{2} \sum_{j=1}^{m} |z_j(\theta) - \tilde{z}_j|^2 = \frac{1}{2} \sum_{j=1}^{m} r_j(\theta)^2 = \frac{1}{2} \|r(\theta)\|^2
\]  

(1)

where each \(z_j(\theta)\) represents an analytical quantity that is a nonlinear function of the variables \(\theta\), \(\tilde{z}\) refers to the measured value of the quantity \(z\), and \(\|\cdot\|\) denotes the Euclidean norm. In order to obtain a unique solution, the number \(m\) of residuals \(r_j = z_j - \tilde{z}_j\) should be greater than the number \(n\) of unknowns \(\theta\). There are various expressions that can be potentially used for this purpose, such as the frequency residual, mode shape related function, modal flexibility residual, and a combination of them [149].

The best features for damage detection are typically application specific, which can come from the time, frequency, or modal domains depending on the employed level of data consideration. Research into methods based on all the three domains is unlikely to stop because no single method has yet been able to identify every type of damage in every type of structures [3].

Time domain features may be represented by the peak accelerations, the root-mean squares of responses, the temporal moments, etc. The direct use of time response in a model-based identification was proposed by [150, 151] that used a least-squares sensitivity-based iterative parameter estimation routine. Masri et al. [152] proposed a neural network based detection scheme that uses time responses both in the training and detection phases. Bu et al. [153] and Lu and Liu [154] studied the direct use of time histories of vehicle or bridge acceleration in sensitivity-based model updating method.

Frequency domain features are commonly the shift in the resonance and anti-resonances or changes affecting the amplitudes. If the excitation is known, the FRF can provide abundant information on damage in a desired frequency range. With the objective of eliminating the modal extraction error, damage detection using measured FRF has been studied by a number of researchers [155–158] and is very promising because the FRF data can be obtained directly without any further extraction and processing. Meanwhile, the transmissibility functions employ only the measurement of the structural responses and do not require knowledge of the excitation input. This operational advantage is coupled with remarkable damage sensitivity as proven in the work by [71, 72, 78, 159]. Sampaio et al. [160]
published a comparative study of using the FRF curvature, transmissibility, and the strain energy obtained directly from FRFs as damage indicators.

Modal domain features are physically meaningful and easier to be interpreted than the time or frequency domain features. The modal domain provides a large set of damage sensitive features, such as the natural frequencies, MAC or COMAC of modal shapes and derivatives [144], modal strain energy [147,161], modal strain [162], damping [163], dynamic flexibility matrix, residual force vector, and a combination of some features, for instance, a combination of the measured modal damping and the computed strain energy distribution [164]. It is found that the modal curvatures are useful parameters for damage detection when using the response-based approach [165] but not as useful when using the model-based approach [166].

Among the three types of features, the time domain features have advantages of avoiding the data reduction that may cause the loss of important information and not being limited by the assumption of linearity as the modal analysis approach is. Therefore, when the structures show high nonlinearity, the time domain features are more suitable than the frequency and modal domain features. The main disadvantage of time domain features is the considerable computational effort for the calculation of time responses. On the other hand, the frequency feature can provide more information on damage in a desired frequency range when compared to the modal feature that is extracted from a very limited range around the resonances. Nevertheless, the modal features can be identified from output-only data measured under operational conditions and does not require the structure to be excited by artificial forces as the frequency domain feature requires. It is worth noting that the feature used for damage occurrence identification may be a global one, while it is better if the one used to determine damage locations and extents is a local one, and must be sensitive to structural local damage.

5.3. Sensitivity Study

As discussed earlier, the selection of objective features is application specific, i.e., none of those features has proven to be satisfactory for every structure and every type of damage. The selection of objective features resides in a trade-off between their damage sensitivity and the possibility to ensure their consistency with the expected structural response. For that purpose, a sensitivity study is typically recommended as a fast way to tell which input parameters are important to a particular output feature in order to reduce the parameter space and select appropriate objective features. Generally speaking, there are two groups of sensitivity analysis (SA) techniques, i.e., local SA and global SA. The first historical approach is the local SA, where the impact of small input perturbations on the model output is studied [167]. In order to overcome the limitations of local methods, such as the linearity and normality assumptions, and local variations, the global SA methods have been developed in a statistical framework to consider the whole variation range of the inputs [168]. Hamby [169] and Iooss and Lemaître [170] presented a comprehensive review of various global sensitivity analyses, ranging from the differential sensitivity analysis, one-at-a-time sensitivity measures, standardized regression coefficients to statistical tests. The most engineering-used screening method is based on the so-called “One At a Time” design, where each input is varied while fixing the others [171].

Moreover, when working with complex models, a screening tool may be necessary to reduce the number of parameters to a manageable size. The significant effects variable screening or linear variable screening method analyzes the contribution (linear effects only) of a particular input parameter to the total model variance [172]. Significant effects method provides an advantage over the sensitivity analysis because it is a probabilistic assessment of variable importance that is obtained through the analysis of variance [121]. In small parameter spaces, the design of experiment methods that systematically calculate a single parameter and combinations of parameters is used [133,173]. If the dimension or the nonlinearity of the parameter space increases, stochastic sampling strategies [174] are preferred for scanning the design space.

In summary, the sensitivity study is very helpful for the parameter selection and the adequate formulation of objective functions. By doing the sensitivity study, the parameter space could be
adjusted or reduced for subsequent optimization problems, and the adequate starting points for optimization can be obtained as well for a successful optimization.

5.4. Optimization Algorithms

The next step is to solve the problem using optimization techniques. The standard form for a single-objective, nonlinear, constrained optimization problem is provided as [175],

Minimize : \( f(x) \)
Subject to : \( g_j(x) \leq 0 \quad j = 1, m \)
\( h_k(x) = 0 \quad k = 1, p \)
\( x_{iL} \leq x_i \leq x_{iU} \quad i = 1, n \)

where \( f(x) \) is the objective function, \( g_j(x) \) is an inequality constraint, and \( h_k(x) \) is an equality constraint function. The \( x \) vector represents the \( n \) design variables that are changed to obtain the optimum. The searchable design space is defined by the lower and upper bounds, \( x_{iL} \) and \( x_{iU} \) of the design variables. Three major types of optimization algorithms are described below.

5.4.1. Gradient-Based Optimization Algorithms

One type of mathematical optimization methods is the gradient-based method that determines the search direction by using the gradient information, offering the best convergence behavior [176]. The gradient-based algorithm starts from a predefined initial parameter set and uses line search algorithms such as quasi-Newton methods [177] to ensure the stepwise convergence to an optimum. For unconstrained optimizations, two popular methods are the Fletcher–Reeves and the Broyden–Fletcher–Goldfarb–Shanno (BFGS) methods. For nonlinear constrained optimization problems, the three algorithms commonly encountered in engineering are the sequential linear programming (SLP), the modified method of feasible directions (MMFD), and the sequential quadratic programming (SQP). Among all the gradient-based methods, the SQP algorithm is the state-of-the-art algorithm for solving differentiable nonlinear programs and has gained great attention in the recent years due to its superior rate of convergence [177–179]. The formulation of an SQP is based on Newton’s method and Karush–Kuhn–Tucker optimality conditions for constrained problems that are changed to the unconstrained optimization [180]. With respect to the finite element based problems that involve relatively large number of degrees of freedom and design variables, the SQP is a quite effective option.

5.4.2. Evolutionary Optimization Algorithms

In contrast to the gradient-based methods, evolutionary algorithms that belong to the class of stochastic search methods have become very popular in the last two decades. The evolutionary algorithms do not require any gradient information and can employ a set of design points to find the optimum. This type of method is typically inspired by natural phenomena and is extremely robust. Two most popular evolutionary algorithms are the more established genetic algorithm (GA) [181] inspired by Darwin’s principle of survival of the fittest, and particle swarm optimization (PSO) [182] based on a simplified social model. Other algorithms include evolutionary programming [183], genetic programming [184], differential evolution [185], simulated annealing [186], tabu search [187], ant colony optimization [188], harmony search [189], etc. Fister et al. [190] presented a comprehensive review on the nature-inspired metaheuristic algorithms. The most important process of evolutionary algorithms is the mutation (random change) of single genes of a parental design to produce a descendant. Genetic algorithms are thereby especially useful for a relatively wide-ranged search in the design space. In general, evolutionary algorithms perform much better in finding the global or near global optimum in the presence of multiple local extrema than gradient-based methods and are well suited for discrete optimization problems [175]. The main drawback is the high numerical effort since a large number of problem solutions are required. Evolutionary algorithms should be applied to
optimization problems where discrete or binary variables dominate the response and problems with a very high number of variables and/or constraints.

5.4.3. Multi-Objective Optimization Algorithms

All the above optimization algorithms are referred to the single-objective problem. Sometimes, the damage detection problems are lack of a clear objective function and need simultaneous optimizations of several objectives, which is referred to the multi-objective optimization problem. These problems do not have a unique optimal solution. Instead, they have a set of alternative solutions named the Pareto Frontier or Functional Efficient Boundary. Therefore, special optimization techniques are needed to solve this type of problems. In general, the two main optimization strategies (i.e., gradient-based and evolutionary algorithms) can be used for solving the multi-objective problems. Coello et al. [191] and Ishibuchi et al. [192] presented a detailed introduction to the multi-objective optimization techniques. For the application in structure damage detection, Jaishi and Ren [147] used a multi-objective optimization technique to optimize two objective functions simultaneously, which overcomes the difficulty of weighing the individual objective function in conventional FE model updating procedure. Perera et al. [143,193,194] developed an evolutionary (GA and PSO) multi-objective framework for structural damage localization and quantification with objective functions based on the modal flexibilities, frequencies, and modal shapes. Cha and Buyukozturk [195] proposed hybrid multi-objective optimization algorithms based on modal strain energy to detect damages in various three-dimensional steel structures. In general, the gradient-based multi-objective algorithms are recommended for smooth (differentiable) problems but not suitable for explicit time integration, while the evolutionary-based algorithms are recommended for all other cases [174]. The advantages of using evolutionary algorithms to solve multi-objective problems are their ability of dealing simultaneously with a set of possible solutions in a single run, instead of having to perform a series of separate runs as in the case of the gradient-based techniques, and the feature of less susceptible to the shape or continuity of the Pareto front [191].

5.4.4. Guideline for Algorithm Selection

In summary, no single optimization algorithm can solve all optimization problems. A basic knowledge of different algorithms helps the selection of the appropriate algorithm for the problem at hand. The guideline for the selection of optimization algorithms is listed in Table 2 below. As for damage detection problems that always result in numerous local extrema, convergence to the correct estimates (not known a priori) will be strongly dependent on a good initial guess. Therefore, the evolutionary algorithm is the better option unless a good initial value is available.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Guidelines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gradient-based Algorithms</td>
<td>suited for problems with many design variables (more than 50)</td>
</tr>
<tr>
<td></td>
<td>where the analysis is computationally expensive</td>
</tr>
<tr>
<td></td>
<td>when numerical noise is not severe</td>
</tr>
<tr>
<td></td>
<td>when gradients are available and local extrema is not an issue</td>
</tr>
<tr>
<td>Evolutionary Algorithms</td>
<td>suited for problems with fewer design variables (less than 50)</td>
</tr>
<tr>
<td></td>
<td>where the analysis is computationally inexpensive</td>
</tr>
<tr>
<td></td>
<td>when numerical noise is severe</td>
</tr>
<tr>
<td></td>
<td>when the gradient does not exists and a global optimum is required</td>
</tr>
<tr>
<td></td>
<td>for discrete and combinatorial optimization problems when the objective</td>
</tr>
<tr>
<td></td>
<td>and/or constraint functions are discontinuous</td>
</tr>
<tr>
<td>Multi-objective Algorithms</td>
<td>for problems with more than one conflict objective functions</td>
</tr>
<tr>
<td></td>
<td>gradient based Pareto algorithm for differentiable problems</td>
</tr>
<tr>
<td></td>
<td>evolutionary based Pareto algorithm for all other cases</td>
</tr>
</tbody>
</table>

Table 2. Guidelines of Optimization Algorithms [175].
6. Prediction of Remaining Useful Life

Damage diagnosis allows for the early detection of damage in structures and also indicates early deterioration in performance levels. The process of implementing a damage identification strategy in structures, which is termed as structural health monitoring (SHM), involves the observation of a structure in real-time using various sensors, the extraction of damage-sensitive features from measured data, and the performance of statistical analysis of these features to determine the current state of structure health [11]. Both damage diagnosis and SHM are required for identifying damage in structures, while the damage prognosis (DP) is an extension of damage diagnosis and SHM and is defined as “the estimation of an engineered system’s remaining useful life” [196]. That is, based on the current state determined by damage diagnosis and SHM, DP is aimed to predict the structural future performance by estimating future loading environments and then predicting the remaining useful life through simulations and past experience.

6.1. Damage Prognosis Methodologies

The main goal of damage prognosis is to predict the remaining useful life (RUL) of the system or structure components before it requires for maintenance or the structure fails. Due to the uncertainty and nonlinearity of the predictive models when damage accumulates, an alternative goal is to estimate the RUL that the system can perform in a safe status under the future loading before one no longer has confidence in the prognosis model [196]. Damage prognosis is initiated and widely studied on the rotating machinery, as reviewed by [197,198]. Si et al. [199] reviewed the recent modeling developments for estimating the RUL, which is centered on statistical data driven approaches. Sikorska et al. [200] and Ahmadzadeh and Lundberg [201] reviewed the advances in RUL estimation from the industry point of view. Tsui et al. [202] provided a concise review of mainstream methods in major aspects of the prognostic and health management framework. The distinction between a damage detection model and a damage prognosis model is that the prognosis model needs to have the ability to estimate the damage evolution in the future, while the detection model only infers the current state of damage and do not necessarily have the ability to model progressive damage mechanisms such as crack growth [203]. In general, the two major categories are physics-based (i.e., model-based) and data-based prognosis methodologies, as detailed below [204].

Data-based techniques are based on statistical and learning techniques typically from the theory of pattern recognition. Data-based approaches [204,205] consist of multivariate statistical methods (e.g., static and dynamic principle component, linear and quadratic discriminant, partial least squares, and canonical variate analysis), machine learning methods (e.g., support vector machine, relevant vector machines, radial basis functions, and neural networks), and graphical models such as dynamic Bayesian networks and hidden Markov models. Although the data-based techniques can indicate a change in the presence of the new loading conditions or system configurations, they perform poorly in classifying the nature of the change.

Physics-based (i.e., model-based) approaches make use of the system models to estimate the RUL or other relevant metrics, which rely on relatively accurate physics-based models for prediction, such as physical failure models [206], filtering models [207], and statistical models. The advantage of physics-based approaches is that the physical knowledge of the system is incorporated into the monitoring process, which is especially useful for predicting system responses to new loading conditions and system configurations such as the damage status. Nevertheless, physics-based approaches are typically more computationally intensive than data-based techniques.

There are also some hybrid methodologies combining multiple approaches as reviewed by [208]. The selection of damage prognosis model depends on the required accuracy and data availability. The data-driven method is preferred for quick estimations that are less accurate, while the physics-based approach is applied when the accuracy of estimation is important. Typically, the balance between these two approaches depends on the amount of available data and the confidence level of the physics-based model in the predictive accuracy.
6.2. Prediction Model and Model Verification

The RUL distribution is sensitive to many aspects [204]: the initial condition supplied by damage diagnosis having a significant impact; the accuracy of predictive models for future loadings and physics of failure models as another major uncertainty source; the sensitivity of system measurements to the damage; the noise in the training and measurement data; and the sampling frequency of measurements. Therefore, besides the RUL distribution that accounts for the uncertainty coming from many sources, a prognosis methodology should have several other important components, namely, the prediction model for future loadings, the transition between situations when measurements are available and unavailable, and the prognosis verification and validation.

The prediction of structure deterioration when subjected to future loadings is another important component of a successful damage prognosis. Future loadings are always forecasted based on the previous loading histories using various data-driven prediction modeling techniques. For example, meta-models such as the state-space representation and multivariate ARMA models can be employed to track previous loadings and to predict future loadings [196]. Zhong et al. [209] used the wavelet neural network method to predict the future system loading based on data from the operational and environmental sensors.

Since the damage prognosis is affected by many sources of uncertainty, the validation of prognosis models is very essential to establish the confidence in the RUL estimate. The process of establishing this confidence in the predictive capabilities is termed as verification and validation (V&V) [196]. The performance can be measured by evaluating the errors between the predicted and the actual RULs, and computing metrics based on accuracy and precision such as standard deviation, mean absolute deviation, mean squared error (MSE), mean absolute percentage error (MAPE), etc. These metrics provide statistical information about the variations in RUL distributions. Saxena et al. [210] introduced several evaluation metrics (e.g., prediction horizon, α-λ performance, relative accuracy, and convergence) tailored for prognosis and presented a detailed discussion on how to interpret and use these metrics. They proposed a standard offline metric test to compare and evaluate prognosis algorithms, while Liu and Sun [211] discussed the online prognostics performance evaluation. In the area of prognosis validation, a significant challenge is to validate nonlinear models, although most of the current research is still at the stage of validating linear dynamic and stress models [212].

7. Decision Making

Once the remaining useful life of a structure or component is determined, a decision-making strategy is used to help decision makers choose a proper action. Decision making under uncertainties is the act of choosing between two or more actions when the outcome of those actions is uncertain [213]. In engineering system, when advanced diagnostics and prognostics provide vital information, a good decision making methodology is required to effectively utilize the available information. The theory of decision making in business, economics, and politics is well established and widely applied in practice [214]. Howard [215] describes the decision analysis as a logical procedure for balancing the factors that influence a decision. The essence of the procedure is to incorporate information about uncertainties, values, and preferences into a structural model of the decision that is suitable for computation and manipulation [213]. The goal of decision analysis is to help decision makers make defensible decisions.

7.1. Value of Information and Uncertainty Quantification

Structural health monitoring is a process aimed at providing accurate and real-time information concerning the condition and performance of structures, which can serve as an objective basis for making decisions regarding operation and maintenance. Nevertheless, in reality structure managers often make decisions based on their experience or on common sense, regardless of the action suggested by instrumental damage detection [216]. One reason is that damage detection is normally affected
by modeling errors and environmental noise such that the damage features are not deterministically related to the actual state and thus decision makers will differently weigh the detection outcomes. Another reason is that managers are very concerned with the consequences of wrong action and thus will make decisions considering the possible effects of the action they can undertake [217]. Wong and Yao [218] realized the need to evaluate the benefit of SHM on the decision making of owners or managers. Pozzi and Der Kiureghian [219] were the first to calculate the value of information (VOI) for SHM by simulating an observable linear degradation law. Zonta et al. [217] introduced a rational framework to assess the impact of SHM on decision making, which address the effect of prior perception and quantify the economic impact of a wrong action.

As discussed earlier, SHM is less used on real structures and one reason is the lack of understanding of the value of information obtained from SHM. Meanwhile, most of the decisions made in damage detection in civil structures are subject to a significant level of uncertainty. Due to these uncertainties, there is a risk of making an inappropriate decision, e.g., too little, too much, too soon, or too late. Therefore, uncertainty quantification plays a major role in the value of information and decision making [220]. There are various uncertainty quantification (UQ) methods developed for dealing with the uncertain parameters, which can be divided into two types as follows [221,222].

Probabilistic approaches assume that the input parameters of the model are random variables with a known probability density function. Three probabilistic UQ techniques are the Monte Carlo simulation method that considers the aleatory uncertainty, the Point estimate method that is based on the concept of moments of uncertain input parameters, and the scenario based method.

Possibilistic approaches describe the input parameters using the membership function. This function only characterizes the bounds where all parameters lie in without knowing the actual distribution, which is able to consider the epistemic uncertainty and applies well in finite element model updating. The possibilistic UQ techniques include the evidence theory, interval analysis, fuzzy theory, Bayesian theory, convex model theory, and information gap decision theory for severe uncertainty.

The main difference between these methods is in line with the different techniques used for describing the uncertainty of input parameters. The probabilistic UQ methods use probability density functions while the possibilistic UQ methods use membership functions to describe an uncertain parameter. The objective of UQ research is to reduce the computation complexity and corresponding costs, and to provide robust decision-making information.

7.2. Decision-Making Methods

Any damage detection, SHM, or damage prognosis strategy ultimately requires a decision-making process. In fact, the realm of decision making under uncertainty is of particular importance to the practicing engineer. In recent years, the decision-making method is gradually introduced to the structural engineering, especially in the areas of structural diagnosis, monitoring, prognosis, and life-cycle assessment. A few currently available decision-making methods are introduced below.

7.2.1. Decision Trees

Decision tree is a method based on graphical models, which uses a tree-like predictive model to achieve the final conclusion regarding the outcome of the desired function. Decision trees provide a detailed structure of the decision problem, graphically showing the different possible scenarios or decision paths as sequential branches in the order the events occur [223]. In general, decision trees contain three types of nodes, i.e., decision, chance, and utility, and paths directed from left to right between nodes. Nodes in a decision tree represent different types of variables. Decision nodes are determinate variables controlled by the decision maker, chance nodes are random variables that represent uncertain quantities in the decision model, and utility nodes correspond to the leaves of the decision tree and represent the value or utility of an outcome in the decision process [214]. Decision trees clearly outline the uncertainty associated with each decision path and provide a comprehensive organization of alternative strategies, which is intuitive and advantageous in calculation efficiency.
7.2.2. Bayesian Decision Theory

Bayesian network is a powerful tool for probabilistic modeling, inference, and decision-making, which offers assistance to decision makers working in complex and uncertain domains [224]. As introduced in the textbooks by [225–227], the Bayesian network is a directed acyclic graph composed of vertices, edges, and conditional probability distribution. Bayesian network models use Bayes’ rules and conditional probability to describe the joint probability of the network, and it considers both the distribution of the variables and their dependence. Due to the use of the prior probability, the Bayesian approach is considered as rational, making it easier for decision makers to accept the mathematical decision making. The Bayesian decision analysis is first introduced by [228], and later Benjamin and Cornell [229] introduced it into engineering decision-making to provide civil engineers with a theoretical framework for making decisions. The Bayesian network has been utilized for decision making in most real-life problems such as forecasting, safety control, risk management [230,231], etc. For engineering problems, Broglio and Der Kiureghian [232] proposed a Bayesian network framework for processing evolving information to assess the condition and make decisions for strategic structures immediately after an earthquake, Blaser et al. [233] applied Bayesian networks to near real-time tsunami warning based on observations of earthquakes, and the Bayesian decision approach has also been practiced for several decades for offshore inspection planning [234]. For modeling of deterioration processes, dynamic Bayesian networks (DBN) as a special class of Bayesian networks can be used [234].

7.2.3. Influence Diagrams

An influence diagram was originally a compact representation of a decision tree for a symmetric decision scenario. Nowadays, the influence diagram is a generalization of Bayesian networks extended with utility functions and decision variables, which can represent and solve decision problems under uncertainties. Influence diagrams contain three types of nodes, i.e., decision, chance, and utility, and two types of arcs, i.e., influences and informational arcs [214]. The influence diagram provides efficient updating of a deterioration model when indirect information is available and in addition it includes the possibility to find expected utilities for decision alternatives [234]. Bensi et al. [235] used the Bayesian network and influence diagram to develop a framework for post-earthquake risk assessment and decision making for infrastructure systems. In order to reduce the number of memory links, an alternative is to use the limited memory influence diagram (LIMID) that includes only direct parents [236–238].

7.2.4. Fuzzy Logic Method

Fuzzy set theory and fuzzy logic are very useful for solving decision-making problems with insufficient knowledge or imprecise data. Different from the probability theory, fuzzy logic theory expresses the uncertainty of truth in an explicit way. In the traditional sets an element is either included in the set or is not, while an element in a fuzzy set is included with a degree of truth normally ranging from 0 to 1 [239]. Meanwhile, fuzzy logic recognizes the lack of knowledge or absence of precise data, and describes most variables in linguistic terms, which make fuzzy logic models very intuitively similar to human reasoning. Ever since the first introduction to risk management [240], fuzzy logic has been widely applied in structure damage identification and decision making [102,241,242]. Ganguli [243] developed a fuzzy logic system for health monitoring of a helicopter rotor blade, which is able to reduce the possibility of false alarms and help the maintenance engineer by roughly locating the damage area but accurate for further nondestructive inspections. Zeng et al. [244] presented a fuzzy based decision making methodology to handle the uncertainties and subjectivities so as to assess risks in complicated construction situations. Meanwhile, due to the feature in describing an event/result using natural language, the fuzzy logic approach would be particularly useful for remedying the uncertainties and imprecision in bridge inspectors’ observations [245].
7.2.5. Summary of Decision-Making Methods

Table 3 summarized the strength and weakness of the above discussed decision-making methods to lay out a general guideline. (1) Decision trees are intuitive, easy to understand, and useful for classification because of the one-to-one linkage, and they work well with insufficient data if all the inference rules can be defined based on expertise. However, due to several disadvantages [214], decision trees are only appropriate for a small system with discrete variables. (2) Bayesian networks are relatively simple since they are graphical models capturing cause and effect relationships through influence diagrams. The use of Bayes’ theorem provides a formalized process to update models when new knowledge or data becomes available and Bayesian networks can readily incorporate uncertain information. Those features make Bayesian networks suitable for small-sized problems with sufficient knowledge of the relationships [239]. Meanwhile, with the combination of Bayesian networks, the influence diagram is a very powerful decision tool. (3) Fuzzy logic models are very convenient in incorporating different expert opinions and adapted to cases with insufficient and imprecise data. Due to the ability of readily incorporating data or opinions expressed by natural language, fuzzy logic models are suited for problems that are not well understood and without sufficient knowledge.

There are also many applications of combining different types of decision-making methods, for example, fuzzy logic models can be used with Bayesian networks, hidden Markov, or decision tree models. These combined models have the potential to solve difficult problems. For example, Ren et al. [246] developed fuzzy Bayesian networks by incorporating fuzzy logic into Bayesian network models such that the variables can be both discrete and continuous.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Guidelines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision Trees</td>
<td>Advantages: suitable for simple issues and easy to understand</td>
</tr>
<tr>
<td></td>
<td>good at dealing with discrete variables</td>
</tr>
<tr>
<td></td>
<td>straightforward for decision-making with limited choices</td>
</tr>
<tr>
<td></td>
<td>Disadvantages: not suitable for complex issues that require many factors and relationships</td>
</tr>
<tr>
<td></td>
<td>exponential growth in number of branches</td>
</tr>
<tr>
<td></td>
<td>all variables must be treated as discrete even if they are continuous</td>
</tr>
<tr>
<td></td>
<td>weak at identifying linear relationship due to its discreteness</td>
</tr>
<tr>
<td>Bayesian Networks and Influence Diagrams</td>
<td>Advantages: suitable for simple issues and the cause-and-effect relationships are known</td>
</tr>
<tr>
<td></td>
<td>can estimate the conditional probability and distribution</td>
</tr>
<tr>
<td></td>
<td>easy to update models when new knowledge or data becomes available</td>
</tr>
<tr>
<td></td>
<td>Disadvantages: not suitable for complex issues involving many variables</td>
</tr>
<tr>
<td></td>
<td>expensive to determine relationships and conditional probability functions</td>
</tr>
<tr>
<td></td>
<td>require expertise about cause-and-effect relationships</td>
</tr>
<tr>
<td>Fuzzy Logic Models</td>
<td>Advantages: suitable for complex problems with insufficient and imprecise data</td>
</tr>
<tr>
<td></td>
<td>easy to incorporate data or opinions expressed by natural language</td>
</tr>
<tr>
<td></td>
<td>useful for problems that are not well understood</td>
</tr>
<tr>
<td></td>
<td>Disadvantages: tedious to develop fuzzy rules and membership functions</td>
</tr>
<tr>
<td></td>
<td>outputs can be interpreted in a number of ways making the analysis difficult</td>
</tr>
<tr>
<td></td>
<td>require a lot of data and expertise to develop a fuzzy system</td>
</tr>
</tbody>
</table>

8. Summary and Future work

This study reviewed the state-of-the-art on the framework of vibration-based damage identification for decision making. This framework consists of several major parts including the detection of damage occurrence using response-based methods, building reasonable structural models,
damage characterization and objective functions, optimization algorithms, remaining useful life prediction, and decision making. The framework is able to give a guideline for researchers and engineers to implement the structure damage detection based on vibration measurements. Through the review of recent developments in vibration-based damage identification, the following conclusions can be drawn and corresponding suggestions for the future research are proposed below.

8.1. Damage Diagnosis

It can be concluded that there is no outright consensus among researchers regarding which vibration-based damage indicator or detection method is the most suitable: (1) The response-based methods are straightforward, easy for application, and not computationally intensive, but at the same time they do not give a quantitative assessment of damage. It is a valuable tool for damage identification in the preliminary stage. (2) The model-based methods assume that a detailed well-correlated numerical model of the structure is available for damage identification. Nevertheless, the construction of such models usually results in modeling errors from simplified assumptions. In order to detect damage rather than the artificial errors from the model construction, a good quality model that could accurately predict the behavior of the structure is required, but is usually a quite difficult challenge.

Although new techniques and extensive studies have been undertaken in the recent decades, there are still many challenges for the future research in structural damage identification to increase the reliability and safety of structures in civil engineering.

1. Most current damage detection techniques are based on linear theory and measurements, which can be used as a rapid estimation of the damage presence if the damage is not very small. However, if the intact structure is nonlinear by nature or a nonlinear behavior is presented due to damage accumulation, the development of nonlinear based damage indicators or detection methods could be more effective for robust damage detection.

2. The effect of various factors on the vibrational measurements needs to be considered for damage detection, for instance, the ambient conditions (e.g., temperature, traffic, etc.), the evolution of boundary conditions, and the variations in the mass and stiffness in real structures. All these variations can introduce uncertainties in both the linear and nonlinear vibrational measurements [94].

3. The development of detection/monitoring techniques that can take into account the modeling errors and uncertainties is desirable to improve the correlation between experimental and numerical results.

4. It is desirable to develop more effective structural models, sensitive damage indicators, and efficient optimization algorithms for the damage detection.

8.2. Damage Prognosis

As the extension of damage diagnosis and health monitoring, damage prognosis plays an important role in predicting the future loading environment and structural performance and eventually determining the remaining life of structures. The anticipated objective of damage prognosis is to provide a reliable and cost-effective approach to the structure operation, assessment, and maintenance, which consists of estimating the probability of mission completion, determining the optimal time for preventive maintenance, and developing the appropriate operation modifications preventing observed damage propagation [196]. The multidisciplinary nature of the damage prognosis problem makes it a grand challenge for researchers and engineers.

1. The remaining useful life (RUL) distribution is sensitive to many aspects as discussed in the paper, and thus one challenge of prognosis is to minimize the uncertainty in the estimation of RUL distribution.
2. The evolution of damage such as cracking, wear, and corrosion plays a critical role in determining the structure health. It is strongly essential to study damage mechanisms and develop accurate and comprehensive failure models. The prediction model of future loadings is another important component of a successful damage prognosis.

3. Since the damage prognosis is affected by many sources of uncertainties, the verification and validation of prognosis model is very essential to establish the confidence in the predictive capabilities. In particular, a significant challenge is to validate nonlinear models.

4. The probabilistic reliability analysis will be necessarily applied to quantify the confidence of the estimated RUL, which assists the decision makers for the proper action.

5. The future research of damage prognosis also includes the life cycle modeling considering the effect of deterioration processes, assessment of consequences associated with possible events, and modeling of preferences to cover the society performance [247].

8.3. Decision Making

When advanced diagnostics and prognostics provide vital information, a decision-making strategy is used to effectively utilize the available information and to help decision makers choose the proper action. Many well-established decision making tools are currently widely applied in business, economics and politics, but not yet in engineering. Decision making methods should have a wide range of potential applications and significant impacts in structural engineering, especially in structure assessment and maintenance.

1. Decision making theories are of little practice to most structural managers and they often make decisions based on their experience or on common sense. It is highly desirable to evaluate the value of information and assess the impact on decision making, which promotes decision-making methods being accepted by the structural managers in practice.

2. Due to the uncertainties of damage detection in civil structures, there is a risk of making an inappropriate decision. The uncertainty quantification plays a major role in decision making and requires further research.

3. The output from the decision-making system should make it apparent to the decision maker, i.e., what the recommendation is and why, including information about the expected cost for the decision alternatives [248].

4. It is necessary to develop methods to evaluate the decision-making performance such as detection/localization rate, false alarm rate, and the optimal trade-off.

5. For the application to real-life infrastructure systems, additional work is needed to develop more efficient approaches to improve the feasibility of near-real-time applications involving large infrastructure systems [235].

In conclusion, the damage identification methods still have limited applications in the practice application [9]. Only damage detection in Level 1 and maybe Level 2 are practically achieved for full-scale structures, while the damage quantification in Level 3 and prognosis in Level 4 are still a great challenge due to the scalability and other issues. Much research work is needed in the area of damage quantification, prognosis, and decision making for full-scale real structures in practice.

Author Contributions: X.K. and C.-S.C. conceived the idea and syllabus; J.H. collected and analyzed the references; X.K. and J.H. summarized the references; X.K. wrote the paper; C.-S.C. and J.H. revised the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References


27. Todorovska, M.I.; Trifunac, M.D. Earthquake damage detection in the Imperial County Services Building II: Analysis of novelties via wavelets. Struct. Control Health Monit. 2010, 17, 895–917. [CrossRef]


35. Zhou, C.; Zhang, W.D. Recurrence Plot Based Damage Detection Method by Integrating T2 Control Chart. Entropy 2015, 17, 2624–2641. [CrossRef]


44. Kraemer, P. Damage Diagnosis Approaches for Structural Health and Condition Monitoring of Offshore Wind Energy Plants; University of Siegen: Siegen, Germany, 2011.

45. Fan, X.; Li, J.; Hao, H. Piezoelectric impedance based damage detection in truss bridges based on time frequency ARMA model. Smart Struct. Syst. 2016, 19, 501–523. [CrossRef]


84. Cawley, P.; Adams, R.D. The location of defects in structures from measurements of natural frequencies. J. Strain Anal. 1979, 14, 49–57. [CrossRef]


87. Lu, Q.; Ren, G.; Zhao, Y. Multiple damage location with flexibility curvature and relative frequency change for beam structures. J. Sound Vib. 2002, 253, 1101–1114. [CrossRef]


99. Xu, W.; Cao, M.; Ostachowicz, W.; Radzieński, M.; Xia, N. Two-dimensional curvature mode shape method based on wavelets and Teager energy for damage detection in plates. *J. Sound Vib.* 2015, 347, 266–278. [CrossRef]

100. Moreno-García, P.; dos Santos, J.A.; Lopes, H. A new technique to optimize the use of mode shape derivatives to localize damage in laminated composite plates. *Compos. Struct.* 2013, 108, 548–554. [CrossRef]

101. Ratcliffe, C.P. Damage detection using a modified laplacian operator on mode shape data. *J. Sound Vib.* 1997, 204, 505–517. [CrossRef]


103. Sazonov, E.; Klinkhachorn, P. Optimal spatial sampling interval for damage detection by curvature or strain energy mode shapes. *J. Sound Vib.* 2005, 285, 783–801. [CrossRef]


108. Sazonov, E.; Klinkhachorn, P. Optimal spatial sampling interval for damage detection by curvature or strain energy mode shapes. *J. Sound Vib.* 2005, 285, 783–801. [CrossRef]


121. Fang, S.E.; Perera, R. A response surface methodology based damage identification technique. *Smart Mater. Struct.* 2009, 18, 065009. [CrossRef]


171. Saltelli, A.; Annoni, P. How to avoid a perfunctory sensitivity analysis. *Environ. Model. Softw.* 2010, 25, 1508–1517. [CrossRef]


© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).