Automated Detection of Geometric Defects on Connecting Rod via Acoustic Resonance Testing

Yun Zheng[*], Matthias Heinrich[*†], Ahmad Osman[*†], Bernd Valeske[*†]

[*]NDT of Comopnents and Assemblies, Fraunhofer Institute for Nondestructive Testing IZFP, 66123 Saarbrücken, Germany

[†]htw saar, University of Applied Sciences, 66117 Saarbrücken, Germany

Email: yun.zheng@izfp.fraunhofer.de

Abstract—**Fully automated defect detection and classification of automobile components are crucial for solving quality and efficiency problems for automotive manufacturers, due to the rising wage, production costs and warranty claims. However, metrological deviations in form still represent unsolved problems using state-of-the-art techniques, especially for forged or casted components with complex geometry. In this paper, we attempt to overcome these challenges by using an acoustic resonance testing model that combines features extraction with defect classification from acoustic natural vibration signals. In this case the study doesn't focus on typical defects like cracks but on defective components in the sense of geometric configurations which are out of tolerance range. With an optimal feature extraction algorithm and a classifier training step, the proposed approach significantly accelerates the detection speed of unacceptable deviations in dimensions and parallely enhances the accuracy. The main contribution of this paper is that an optimal feature from acoustic signals is found which represents the geometric parameters most appropriately, meanwhile, the most appropriate classifier is obtained which significantly improves the efficiency and accuracy in defect classification.**

I. INTRODUCTION

With the accelerating pace of globalization, quality control has been regarded as one of the most important parameters in automobile manufacturing industry to surmount current and future challenges. Inefficient and unreliable quality inspection methods lead not only to a significant incremental of costs, but also to serious failure and security risks, thus damaging the commercial image of the components manufacturers and supplier as well as the car manufactures. The testing of metallic components of complex geometries still mainly relies on visual inspection or destructive testing, which are time consuming, expensive and require trained operators. Thus there is a major need to replace the current methods with automated, efficient and low-cost testing methods.

The automated sound test, which is also known as acoustic resonance testing (ART), is a simple but robust nondestructive test method which is mostly used in the examination of serial parts which are produced in large quantities and with low cycle times, e.g. forged or casted metallic parts. A big advantage of ART compared to other volume-oriented methods is its fast in-line quality assessment of an entire specimen within a few seconds. The principle of ART is based on the fact that a sample's natural vibration behavior and its eigenfrequencies mainly depend on geometry, structural defects and material properties. Generally, the exact geometric dimensions and material properties of single parts in a serial production vary randomly within acceptable ranges. Intolerable component variations, like size out of standard range, leads to shifts in eigenfrequencies. A method to differ between eigenfrequency shifts caused by acceptable as well as intolerable variations is required. The goal of this paper is to evaluate the possibility of using certain spectral components, i.e. eigenfrequencies as inputs for classifiers to decide whether sizes of simplified connecting rod is "OK" (acceptable variations in geometry) or "not OK" (size out of range).

The remainder of this paper is organized as follows: in Sec. II.A, a review of state-of-the-art nondestructive testing techniques is presented. An overview of the proposed approach for classification of geometric defect on connecting rod is clarified in section. II.B, while the implementation details are provided in Sec. III. To assess the performance of the proposed approach, a comprehensive evaluation is conducted on a well-defined dataset using appropriate measures in Sec. IV. In the last section, some general conclusions are presented.

II. PROPOSED APPROACH

A. Related Work

Nondestructive testing ART is a well established quality control method. Some publications focus on the algorithms to differentiate the influences to the acoustic vibration behavior caused by acceptable scatter and defects respectively [1], [2]. Those approaches take advantage of simulated vibration data of samples with a geometric complexity near to real serial components. Other more recent publications consider different approaches to improve the quality assessment based on natural vibration data with the aim of identifying defect locations and types. These publications focused only on defect detection but ignored tolerable variations from part to part [3]–[5]. Due to the improvement of classification algorithms and significant advances in processing power nowadays via capable computer hardware, more intelligent and comprehensive solutions may be now employed to offer satisfactory and robust automated test results.

Within this sense, some preliminary works about calculating the geometry parameters with eigenfrequencies were introduced in [6] and [7]. The eigenfrequencies calculated by finite element method [8] are only slightly different from measured frequencies, especially when the input data like geometry and material information used for the calculation are sufficient

compared to real properties of the samples [9]. Within those contributions in [6] and [7] a similar approach was followed using a linear regression model to estimate the exact dimensions of components from their eigenfrequencies which led to very promising results when using a sufficient number of frequencies. However, small deviation of simulation, geometry and material parameters might lead to bigger deviation in the calculated geometric values (such as thickness and diameter). How to eliminate the number of needed eigenfrequencies while maintaining the robustness and accuracy of the model is still an open topic.

B. Deployed Workflow

The study is based on simulated vibration data because a big amount of samples is needed. At first, many similar samples with marginal random geometry variations were defined. Then, the samples' eigenmodes and eigenfrequencies were calculated using numerical finite element simulations. In the next step, the feature extraction algorithm is applied to find the most representative information of geometric parameters from the eigenfrequencies. At last, a classifier is trained on a dataset in terms of extracted features to separate the samples for the samples actual size into different categories: "OK", "under range" and "over range".

III. IMPLEMENTATION

In this section each stage of our proposed classification algorithm is described in detail. To keep the deployed workflow more general, all considered processing steps for classifying the geometric defects are presented. Further information on the evaluation of optimal features and classifiers are provided in Sec. IV.

A. Data Generation

Fig. 1 illustrates the geometry of connecting rod with flat and top view. Eight parameters like thickness or diameter, which are to be separated into the above mentioned categories ("OK", "under range"and "over range"), are defined with small random geometry variation. The eigenfrequencies and eigenmodes (like shown in Fig. 2) are then calculated with finite element simulation using a software named COMSOL Multiphysics (version 5.1 and 5.2a). The difference between ART measured eigenfrequencies of a manufactured component and those from the finite element calculations were less than 3 percent for all mode shapes. Those variations can be explained with simplifications in the finite element model respectively not precisely known properties of the real sample, the influence of discretization when using finite element method, measurement noise etc.

B. Feature Extraction

Feature extraction is an essential method, which has been widely used in applications like image processing, classification etc. to reduce the redundancy of sample data. Principal Components Analysis and Linear Discriminant Analysis are two of the most commonly used algorithms to conduct the
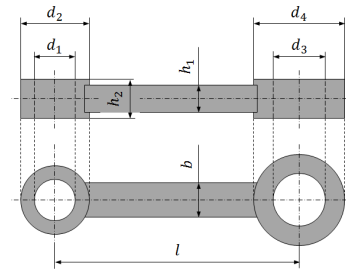

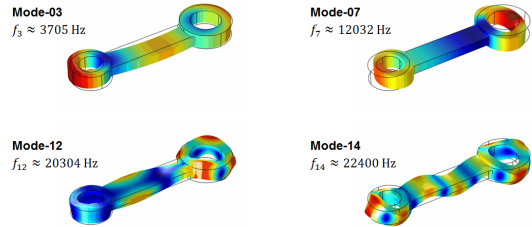Fig. 1. Geometry of connecting rod with flat view (upper) and top view (lower).

Fig. 2. Examples of eigenmodes of connecting rod with a certain geometry

classification more simply and precisely by attempting to projecting the samples into a lower dimensional space.

1) Principal component analysis (PCA): To solve the problem of redundant features, PCA [10] is applied to convert the possibly correlated high-dimensional data space into a set of linearly uncorrelated features by weighting the datasets in a balanced way. This leads finally to more robust classifier in general, but not necessarily to the one with best accuracy [11].

2) Linear discriminant analysis (LDA): Being similar to PCA, LDA looks for a linear combination of variables which best explain the data. The most obvious difference between LDA and PCA is that the features extraction performed with LDA is supervised (with class information). However, PCA is label agnostic. The purpose of LDA is to find a subspace which gathers the samples from same class and meanwhile enlarges the margin from different classes [12] .

C. Classification

1) Support Vector Machines (SVM): SVM are one of the standard tools for machine learning and data mining based on advances in statistical learning theory. They work in a two step process. The first step is the training (with representative sample data) where the support vectors are generated determining the optimal separating hyperplane or set of hyperplanes with the maximum distance to these support vectors. The second step is the regression/classification of unknown data in the feature space. A Support Vector Machine is a maximal margin hyperplane in feature space, built by using a kernel function in input space. In this work, the used kernel is a Gaussian Radial-Basis Function (RBF) [13]:

$$K(p, q) = \exp\left(-\gamma ||p - q||^2\right) \tag{1}$$

where $p = [p_1, p_2, \cdots, p_d]$ and $q = [q_1, q_2, \cdots, q_d] \in \mathbb{R}^d$ are feature vectors. The variable γ is the width of the RBF. The kernel function measures the similarity in feature space by working like an inner product in this space.

2) Bag of decision trees: Decision tree is a kind of weak classifier that has a flowchart-like structure, in which each node stands for a "test" on attribute and each branch represents the outcome of the test. In the practice, overfitting is a significant problem for decision tree, because the full growth of a decision tree may lead to loss of generalization capability. Bootstrap-aggregated (bagged) decision trees [14] combine the results of many decision trees, which eliminate the influence of overfitting and the final prediction of the compound classifier is given as a weighted combination of individual classifier predictions [15]:

$$H(d_i) = sign(\Sigma_{m=1}^M \alpha_m H_m(d_i)) \qquad (2)$$

where an example d_i is classified with a classifier H using a sequence of weak classifiers $H_m, m = 1, ...M$. Parameters α_m are determined in such way that more precise classifiers have stronger influence on the final prediction than less precise classifiers [15].

3) Random Forest: As an extension of bag of decision trees, random forests [16] differs in only one way from this general scheme: special process named "feature bagging" is applied in the learning process to select a random subset of the features at each candidate split. The purpose of this process is to prevent the features, which have a strong influence to the predictor in the training process, from causing the correlated feature space [17].

4) Adaptive Boosting (AdaBoost): AdaBoost is a machine learning algorithm referring to a particular method of training a boosted classifier [18]. One of the main ideas of the algorithm is to maintain a distribution or set of weights over the weak classifiers (here decision tree). The initial weights for all classifiers are set equally at first. With iterations in the training process, each weights will be updated according to the results wether the examples are misclassified. The weights of incorrectly classified samples are increased so that the learning could be focused on the misclassified samples.

IV. EVALUATION

A. Dataset and Measures

Dataset with 10000 samples for training and 2000 samples for evaluation were generated, which contains the actual sizes and the corresponding eigenfrequencies of the connecting rods. The actual sizes of the samples are shown in TABLE I. The uniformly distributed training value ranges 20% broader than evaluation data to obtain more robust classifiers. There are also critical values in transition class which ranges ±0.025mm (with width of 0.05mm) between "OK" and "not OK" parts (see the bottom figure in Fig. 3). In the evaluation process, classifiers are trained respectively with and without transition class to evaluate how samples in the transition class influences the accuracy of the classification. Fig. 3 visualizes the values

TABLE I
EXPECTED VALUE E AND THEORETIC VALUE RANGES OF THE "OK" SAMPLES

Parameter	E [mm]	Value ranges [mm] (Training data)		Value ranges [mm] (Evaluation data)	
		min	max	min	max
d_1	18	17.3250	18.6750	17.4375	18.5625
d_2	30	29.1000	30.9000	29.2500	30.7500
d_3	24	23.3250	24.6750	23.4375	24.5625
d_4	40	39.1000	40.9000	39.2500	40.7500
h_1	7	6.3250	7.6750	6.4375	7.5625
h_2	12	11.3250	12.6750	11.4375	12.5625
l	100	98.2000	101.8000	98.5000	101.5000
b	18	17.1000	18.9000	17.2500	18.7500

ranges of the connecting rod's sizes which are uniformly distributed and also gives approximate amount of "OK" and "not OK" sizes calculated from theoretical probability distribution.

Additionally, all data contain the eigenfrequencies up to about 30 kHz, which can be used as input values for classification. For each sample there are 18 eigenfrequencies sorted by the corresponding modes shapes.

Cross validation with 10 folders are applied during the training process to evaluate the generalizability of trained models.
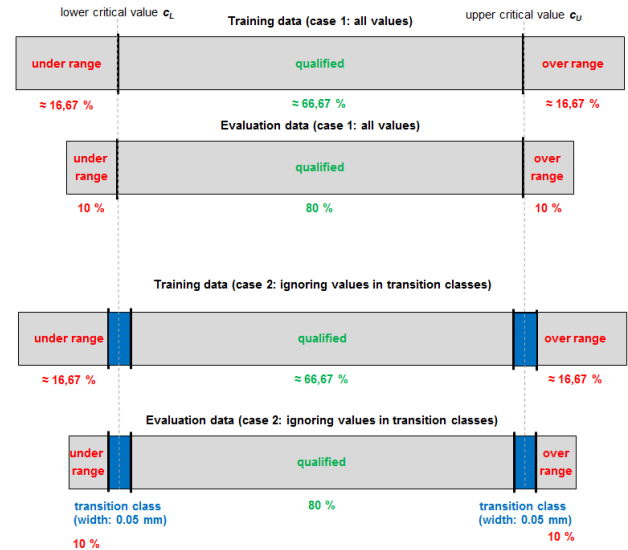

Fig. 3. Distribution of training and evaluation dataset

B. Results and Discussions

1) Training: For the similarity between geometric parameters, the evaluation below uses parameter d_1 as an example.

a) Feature Extraction: Fig. 4 indicates that model of random forest significantly outperforms the model of bag of decision trees using raw eigenfrequencies as input. As mentioned in section III-C3, random forest has the advantage in dealing with overfitting data, which explains why it performs more stable and accurate compared with bag of decision trees.
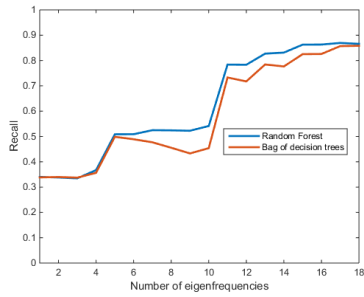
Fig. 4. Average recall of all classes with random forest and bag of decision trees using raw data

TABLE II
CONFUSION MATRIX OF RANDOM FOREST OF d_1 USING LDA FEATURES

$N = 10000$	Predicted: OK	Predicted: Under range	Predicted: Over range
OK	6566	41	39
Under range	46	1645	0
Over range	50	0	1613

where N stands for number of samples.

TABLE III
SUMMARIZED PERFORMANCE OF RANDOM FOREST OF d_1 USING LDA

Parameter	With transition classes	OK	Under range	Over range
Recall	Yes	98.70%	95.98%	90.56%
	No	100%	100%	100%
Precision	Yes	98.46%	93.63%	95.32%
	No	100%	100%	100%

In consideration of the redundant input eigenfrequencies, the feature extraction process is critical to find the features which represent the geometric length most accurately. As introduced before (see section III-B), PCA and LDA, as two of the most representative feature extraction algorithms, are used to convert the correlated eigenfrequencies into a set of independent variables.

Fig. 5 shows the recall of each group for the model learned with random forest using raw data, PCA and LDA respectively. An increasing number of eigenfrequencies are used as input to study how many eigenfrequencies required to represent the geometric parameters. Compared with training directly using raw data, PCA obviously needs fewer eigenfrequencies to get a stable performance. However, the recall of "not OK" samples is not significantly enhanced. LDA, on the other hand, needs minimal number of eigenfrequencies to achieve a stable performance with higher accuracy (98.89%, 97.75% and 97.53% respectively of "OK", "under range" and "over range" samples are correctly classified). The results shows that LDA, as a supervised feature extraction method, is more appropriate in extracting the uncorrelated features from eigenfrequencies between different classes.

b) Classifier Comparison: The performance of each classifiers using LDA is illustrated in Fig. 6 which shows the average recall of three classes. If sufficient number of eigenfrequencies (more than 10) are given, the performances are almost the same. The highest accuracy is achieved by SVM with recall of 97.98%. For the trained model with limited number of eigenfrequencies, random forest and bag of decision trees outperforms the others with more stable performance. Since the extracted LDA features are already uncorrelated with each other, no major difference in recall is found between random forest and bag of decision trees, which further proves that LDA features overcome the problem of redundancy in eigenfrequencies successfully. In general, random forest and bag of decision trees are the most suitable classifiers in our case.

c) Geometric Parameters: The confusion matrix of random forest is shown Table II, from which it can be seen that among the 10000 samples, no misclassification between "under range" and "over range" groups is found, which indicates that features of "under range" and "over range" samples are

more obvious to be distinguished from each other.

The graph in Fig. 7 shows that for each geometric parameter (length, diameter, depth etc.), the classifiers performs almost identically. Recall and precision for parameter with higher deviation like d_2, d_4 and l are slightly lower than the others. In general, the trained models have performed precisely with both high recall and precision for the three classes.

2) Evaluation: With the trained model obtained with random forest using LDA features, two set of tests are carried out to see how the transition data influences the performance of classifiers. In TABLE III it is shown that if the test data includes the transition data, which lies between groups of "OK" and "under range", or between "OK" and "over range", both recall and precision can be achieved for over 90%. For the dataset in which the transition classes were ignored, the recall and precision are even 100% for all groups. The result shows that the data in transition class is the most important factor, which strongly influences the performance.

V. CONCLUSION

Classification of geometric defect is one of the most critical step towards qualified production of forged or casted components in automobile manufacturing industry. However, the efficiency and accuracy in detecting geometric defects is still an open question so far. In this paper, a solution to overcome this problem is proposed by using a ART model that combines feature extraction with defect classification from acoustic signals. Except the connecting rod, this method can also be applied to other components consist of one homogeneous material.

By using the measured eigenfrequencies and appropriate features, LDA for instance, the most representative information of geometric parameters are obtained in comparison to state-of-the-art approaches. Further improvement by finding the most robust classifiers is observed, which helps to accomplish the whole detection pipeline in detecting geometric defects. Random forest is preferred for more robustness against redundancy of eigenfrequencies in geometric defect detection.

(a) Recall of OK samples (b) Recall of under range samples (c) Recall of over range samples
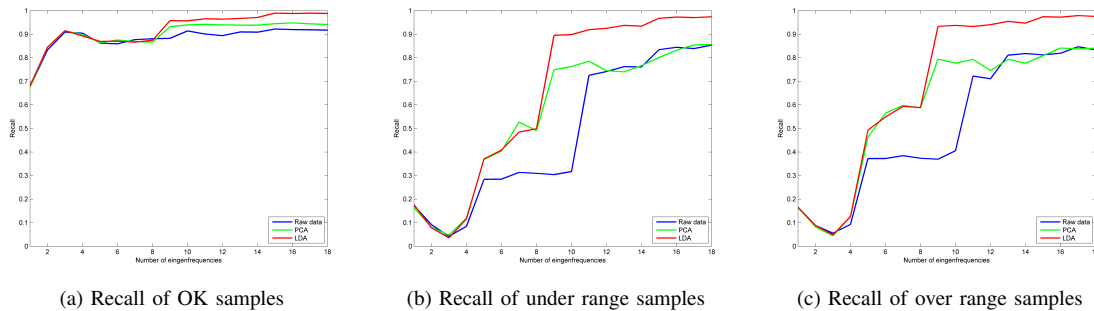
Fig. 5. Recall of each class with random forest using raw data (blue), PCA (green) and LDA (red).
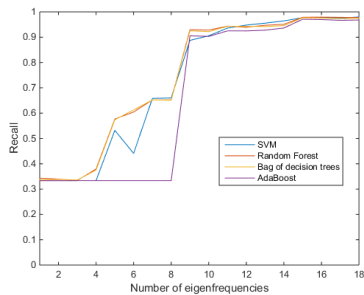

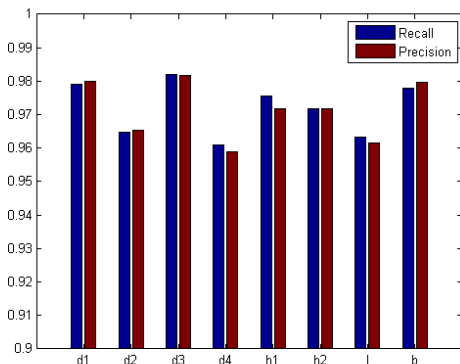Fig. 6. Average recall of three classes with different classifiers using LDA features

Fig. 7. Average recall and precision of three classes for each geometric parameter with random forest using LDA features

REFERENCES

[1] L. Schmidt and R. Steinbuch, "Improved Interpretation of the Acoustic Response Spectrum to Identify Types of Component Deviations," *Research in Nondestructive Evaluation*, vol.14(2), pp. 95–110, 2002.

[2] R. Steinbuch, "Scatter or defect? Some remarks on the interpretation of acoustic spectral shift," *Research in Nondestructive Evaluation*, vol. 15(4), pp. 173–189, 2004.

[3] C. Lai, W. Xu and X.Sun, "Development of an Inverse Algorithm for Resonance Inspection," *Journal of Vibration and Acoustics*, vol. 134(5), pp. 051017.1-051017.10, 2012.

[4] C. Lai and X.Sun, "Predicting flaw-induced resonance spectrum shift with theoretical perturbation analysis," *Journal of Vibration and Acoustics*, Vol 332(22), pp. 5953-5964, 2013.

[5] W. Xu, C. Lai and X. Sun, "Identify structural flaw location and type with an inverse algorithm of resonance inspection," *Journal of Vibration and Acoustics*, vol 21(13), pp. 2685-2696, 2015.

[6] M. Heinrich, U. Rabe, B. Grabowski and B. Valeske. "Acoustic resonance testing for nondestructive detection of forged or casted serial parts with intolerable geometric variations – experimental model adaption," *DGZfP-Proceedings BB 158*, 2016.

[7] M. Heinrich, U. Rabe, B. Grabowski and B. Valeske, "Methodische Entwicklung der akustischen Resonanzanalyse zur zerstörungsfreien Erkennung von ur- oder umgeformten Serienteilen mit unzulässigen Geometrieabweichungen – simulationsbasierte Modellüberlegungen,"*DGZfP-Proceeding BB 152*, Salzburg, Germany, 2015.

[8] J.N.Reddy, *An Introduction to the Finite Element Method*, third ed. New York: McGraw-Hill, 2005.

[9] C. Lai, X. Sun, C. Dasch, G. Harmon and M. Jones, "Quantify Resonance Inspection With Finite Element-Based Modal Analyses," *Journal of Vibration and Acoustics*, vol. 133(3), pp. 031004.1-031004.9, 2011.

[10] H. Hotelling. "Analysis of a complex of statistical variables into principal components", *Journal of Educational Psychology*, Vol. 24(7), pp. 498–520, 1933.

[11] D. Geller, G. Platsch, J. Kornhuber, T. Kuwert and D. Merhof, "Quantile-Based Classification of Alzheimer's Disease, Frontotemporal Dementia and Asymptomatic Controls from SPECT Data", *IEEE Nuclear Science Symposium and Medical Imaging Conference*, 2016.

[12] R. Fisher, "The Statistical Utilization of Multiple Measurements", *Annals of Eugenics*, vol. 8, pp. 376-386, 1938.

[13] W. Press, S. Teukolsky, W. Vetterling and B. Flannery, *Section 16.5. Support Vector Machines"*. , Numerical Recipes: The Art of Scientific Computing (3rd ed.). New York: Cambridge University Press. ISBN 978-0-521-88068-8, 2007.

[14] L. Breiman, "Bagging predictors", *Technical Report No. 421*, Department of Statistics, University of California at Berkeley, 1994.

[15] K. Machová, F. Barčák and P. Bednár, "A Bagging Method using Decision Trees in the Role of Base Classifiers", *Acta Polytechnica Hungarica*, Vol. 3, No. 2, pp. 121-132, 2006.

[16] L. Breiman, "Random Forests", *Machine Learning*, vol. 45(1), pp. 5–32, 2001.

[17] T. Ho, "A Data Complexity Analysis of Comparative Advantages of Decision Forest Constructors", *Pattern Analysis and Applications*, pp. 102–112, 2002.

[18] Y. Freund and R. Schapire, "A Short Introduction to Boosting", *Journal of Japanese Society for Artificial Intelligence*, vol. 14(5), pp. 771-780, 1999.