Contents

Abstract

Roller bearing data has already been used in many publications, to evaluate and proof *condition monitoring* concepts (Marwala, 2012), (Nelvamondo, Marwala, & Mahola, 2006), (Li, Chow, Tipsuwan, & Hung, 2000). Most of these papers used both normal and outlier training data to evaluate the efficiency of *supervised* classifiers such as *neural networks* or *support vector machines*. Goal of this thesis is the design and evaluation of an exemplary *semi-supervised* condition monitoring approach, using the roller bearing data from (Case Western University) as benchmark.

My contribution involves

- Research on *condition monitoring* and a concise presentation of a general condition monitoring approach (1)
- Research into *feature extraction* methods and selection of suitable methods with respect to the benchmark data set (2,4.2)
- Implementation of a *Higuchi fractal dimension (HFD) feature* extraction method (4.2.2)
- Research on *one-class-classification* and selection of suitable classifiers in relation with the benchmark data set and the available features (3,4)
- Design and implementation of a new *one-class-classifier* based on *random forest*
- Implementation and evaluation of a complete *semi-supervised condition monitoring* approach, including *feature extraction*, *feature selection*, classification and tests (4).

The first part of this thesis introduces several techniques, starting with *feature extraction* methods in chapter 2. Chapter 3 provides a general approach for the construction of *one-class-classifiers* and continues with an introduction of various *semi-supervised* classifiers which can be found in literature. The final part of this chapter discusses several aspects of the *random forest* classifier (3.7), of which a *semi-supervised* version was developed as part of this thesis (3.7.6), (4.3).

The techniques introduced in the first part, are evaluated in the second part of this thesis (4), which starts with a description of the *roller bearing* data set in 4.1. The *feature extraction* results are briefly discussed in 4.2 and 4.3 provides a proof of concept of the *semi-supervised random forest* classifier introduced in 3.7.6. Two experiments where designed to evaluate the complete *outlier detection* scenario, of which the first one (4.4.1) compares classifier performances, while the second analyzes the applicability of *random forest* for feature reduction purposes.