An Investigation of the Application of Sound Spectrum Features in Classifying Impeccable and Defective Gears

Isaack Adidas Kamanga

Dar es Salaam Institute of Technology (DIT), Dar es Salaam, Tanzania
Electronic and Telecommunications Engineering Department
Isaackkamanga[at]gmail.com
Dar es Salaam, Tanzania

Abstract: ***The purpose of this study is to provide a systematic approach to address the industrial challenge of recognizing impeccable and defective gears by analyzing the spectrum of sound waves produced by the gears when in operation if attached to a testing circuit during sorting. This work classifies the gears into two classes, impeccable and defective classes. The spectra from several samples from both impeccable and defective gears were analyzed and five audio features were extracted from their spectra namely short-time energy, zero-crossing rate, Spectral entropy, pitch, and block energy entropy. It was found that there is a significant difference between the two classes. In training the algorithm, 5D features vectors from 20 feature vectors from impeccable gears and another 5D features vector from 20 defective gears as training samples to determine the discriminating point. In testing the algorithm, 20 samples were extracted randomly from the impeccable and defective gears but whose status was clearly known by visual inspection. The results of gear status given by the algorithm were compared to that of visual inspection. The samples were classified by using the Support Vector Machine (SVM) learning classification approach. A promising efficiency of 95% was obtained.***

Keywords: Gear**,** Short-Time Energy, Zero crossing rates, Energy Entropy, Spectral entropy

1. Introduction

Image pattern recognition, ANN and machine learning, and sound spectrum analysis can all be used to tackle the challenge of distinguishing impeccable and defective gears. This research work explains how impeccable and defective gears can be classified with the help of features extracted from sound spectrum of a piece of sound produced by a respective gear when attached to a testing circuit as shown in Figure 3. This problem was inspired after a factory visit to a factory that manufactures these gears shown in Figure 2 in Tianjin China. I experienced the challenges faced by the factory workers in sorting the two groups of gears. The gears are quite small causing wastage of time and low accuracy in sorting the gears. This work is primarily aimed at overcoming the wastage of time and low accuracy due to human error challenges in sorting the gears. This research work treated this problem as a pure machine learning and classification problem with two classes namely impeccable and defective classes, hence the entire report is about how a gear sound was categorized as produced by impeccable or defective gear.

2. Literature Survey

In the field of gear classification in industrial manufacturing and monitoring, numerous signal processing approaches such as time-domain, spectrum, and continuous wavelet transform analyses have been seen in the literature. Machine learning methods have been utilized by few scholars to classify gears. However, earlier researches did not adequately examine its applicability. As a result, there is potential for establishing the best way for detecting various defects in gears using machine learning techniques. As part of a novel research project, vibration signals and machine learning techniques were used to diagnose faults in the gears while it was operating in a signal acquisition system as shown in Figure 3.

3. Problem Definition

Defects in the gear teeth have a significant effect on transferring the torque from one gear to another. When two gears are in contact, the teeth of one gear exert force on the teeth of the other. This force will have both a radial and circumferential component because the touching teeth produce a line connecting the shaft axes of both gears. If one tooth or more gears have defects, there will be a failure in the calculated force transfer between the two gears. For this reason, there is a need to have a smart system to sort the impeccable gears from defective ones before selling them to other manufacturers who use gears to make other machines such as gearboxes. At the factory I visited they use manual methods in sorting the gears, this wastes a significant amount of time. Therefore, there is a need for an automatic system to help to automate the process.

,

4. Methodologies

The initial stage in any audio classification system is to extract features, which entails identifying the components of the audio signal that are useful for classification while eliminating everything else, such as background noise and so on. A variety of features were extracted using this method, but those that showed no significant difference between excellent and defective gears during training, such as the Mel frequency cepstral coefficients (MFCC) feature were not taken into account. Short-time energy, signal block entropy, spectrum entropy, zero-crossing rate, and pitch are the features extracted and employed for training in this study. Each of these characteristics is briefly explored in this article. A clear line has been constructed to approximate the discriminating point between the two classes after features extraction and training, as well as an adequate approach for classifying the new characteristics vector extracted from the sound produced by the gear whose state is unknown (defective or impeccable is to be determined). Several sounds were collected from ten impeccable gears and ten defective gears for testing purposes, and they were then classified using our approach. Figure 1 shows the steps followed in identifying the status of a given gear.

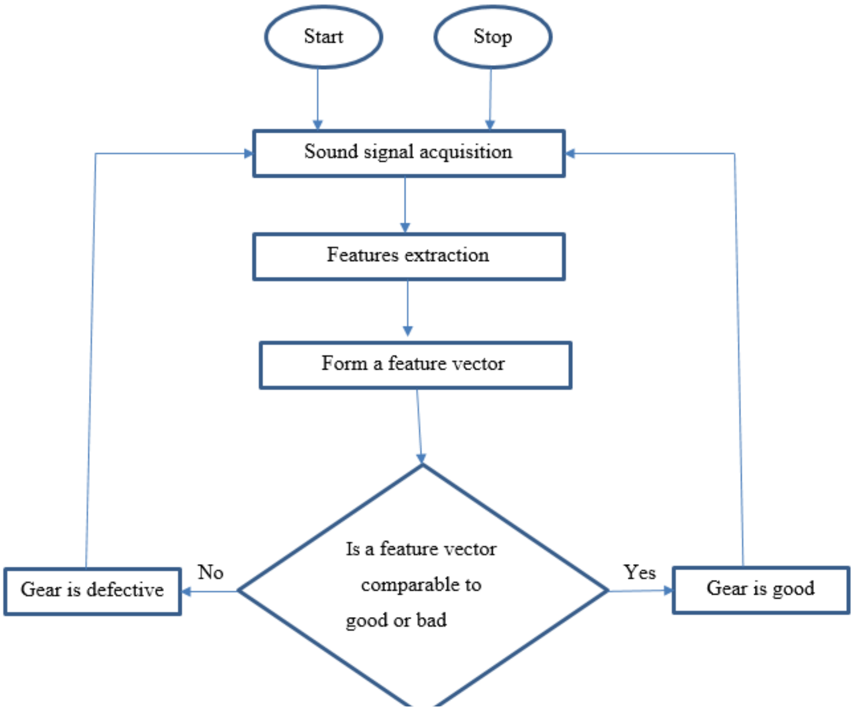

Figure 1: Flowchart for identifying the status of a given gear

4.1 A Gear

A gear, often known as a cog, is a portion of a spinning machine that has cut teeth that mesh with another toothed element to transmit torque [3]. After fabrication, a impeccablegear must have all of the required features, such as impeccablecut teeth, the suitable thickness and diameter; in our case, the number of cut teeth on a reducer must be 50, and the number of cut teeth on the inner section must be 11. Any piece of equipment that has features that are incompatible with those listed above will be regarded a dud. Figure 2 depicts the differences in cut tooth arrangement between the impeccableand defective gears, as illustrated in pictures.

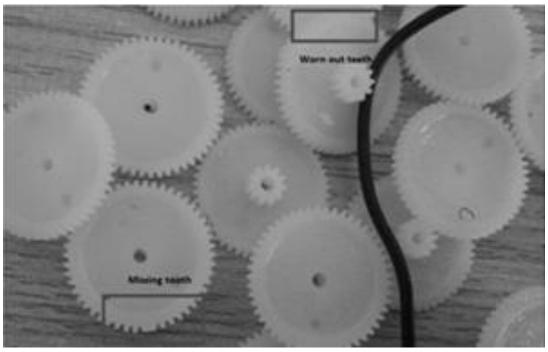

Figure 2: Image illustrating good and bad gears by cut teeth arrangement and shape

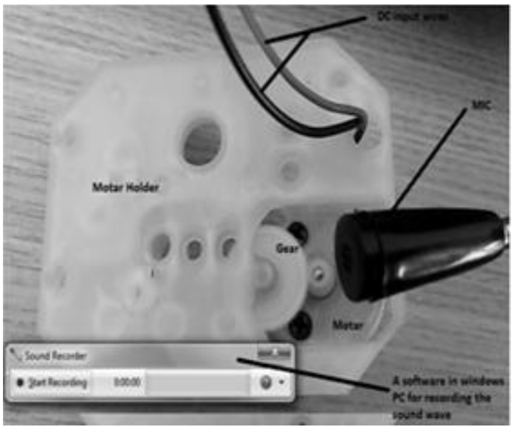

Figure 3: Arrangement used to input sound waves generated from several gears

4.2 The Short-Time Fourier Transform

The Fourier transform tells us how much of each frequency is present in a gear signal. If the signal's spectral content doesn't change much over time, this works well; however, if the signal changes over time, such as in a song where different notes are played one after another, the Fourier transform won't be able to distinguish between them, and the Fourier representation will display information about all of the notes together. The short-time Fourier transform (STFT) is an attempt to correct the standard Fourier transform's lack of time resolution. The input data is divided into several little sequential chunks known as frames or windows, and the Fourier transform is done to each of these frames one by one. The result is a time-dependent depiction that shows how the harmonic spectrum evolves as the signal develops. STFT originally operates on infinity waves, to apply on short waves we need to use an idea of windowing function. Windowing function reduces the discontinuity at frame boundaries by gently scaling the amplitude of the signal to zero at each end. Using no windowing function is the same as using a square-shaped windowing function. This is referred to as a square or boxcar window. The windowing functions do not totally eliminate frame boundary effects, but they do significantly minimize them. Figure 4 depicts a simple sine wave windowed using three distinct windowing functions, as well as the Fourier representations for each. A single sine wave should have a Fourier representation of a singleton component, and no STFT window totally eliminates boundary effects, as shown in Figure 4, but some do better than others.

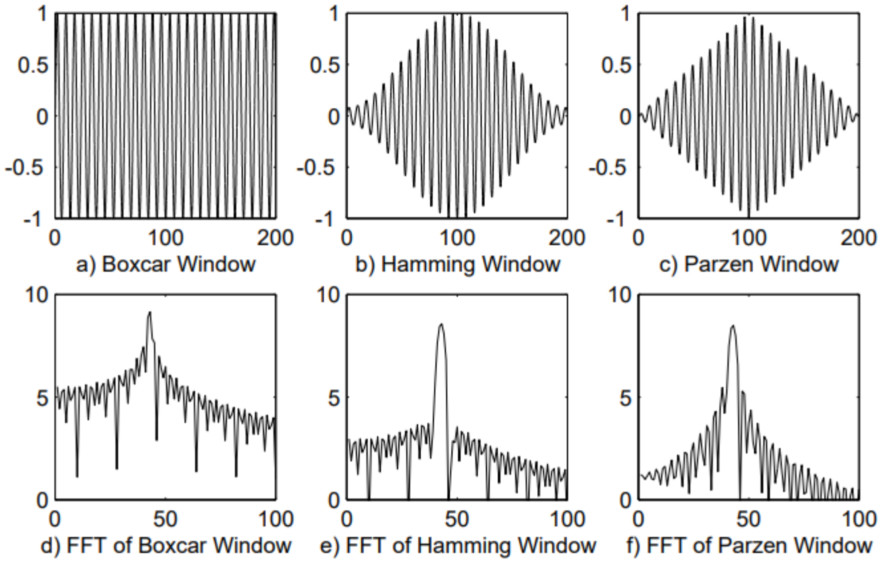

Figure 4: An impact of windowing functions on a sine wave

5. Sound Wave Acquisition

The audio recording setup is depicted in Figure 4 below. The sound wave was recorded for 10 impeccable gears before repeating the process for defective gears using the MIC linked to the PC. I had to convert the wave from.wma to.wav file for smooth analysis because I was using an earlier version of MATLAB program. I used the recorded sound from the working directory for analysis. The first 250000 samples have a flat spectrum, which makes them less valuable for our investigation. These samples may be the result of a time lag between starting the recording program and hearing the sound (ringing) from the gear, and should not be considered typical. Because the stereo recording is double channeled, I can only use the left channel for our analysis; in many cases, the left channel is selected because mixing stereo channels can generate odd results. Because the phase-shift varies by frequency, some frequencies may be lost when averaging. Figure 5 shows how the acquired sample (Figure 5(a)) was modified to remove unwanted part and remained with the useful part (Figure 5(b)).

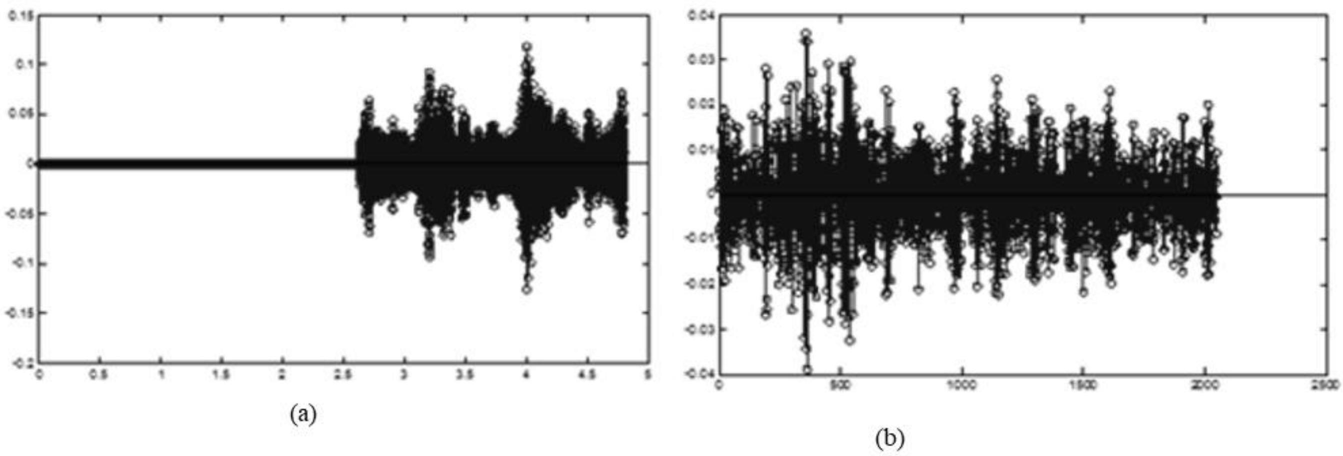

Figure 5: (a) Input sound wave and **(b)** the cut version of it

The usage of only one channel is recommended, especially to avoid needless memory waste. I lowered the number of samples before extracting features. Figure 6 shows the final period of the input signal I examined for analysis.

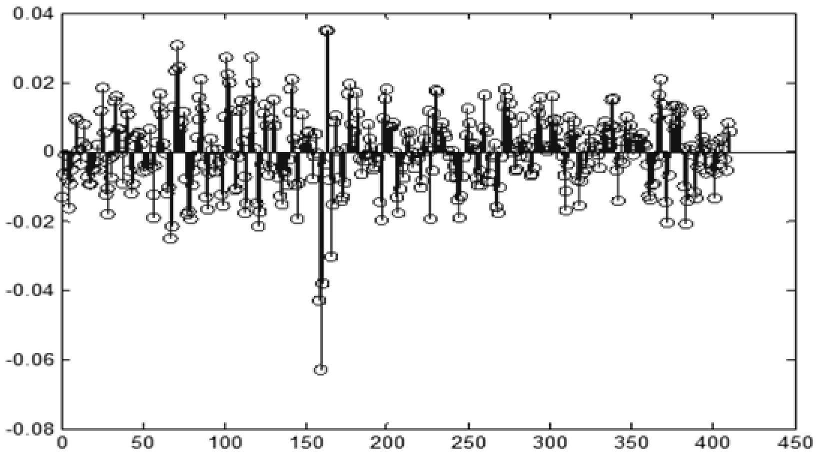

Figure 6: Samples considered for features extraction

6. Features Extraction

6.1. Definition of features used

6.1.1 Short-Time Energy

The energy E of a discrete temporal signal $x(n)$ is defined by Equation 1 below. Such a measurement is less important for many audio signals because it provides little information about the signals' time-dependent characteristics.

$$E = \sum_{n=-\infty}^{\infty} x^2(n) \quad \ldots\ldots(1)$$

Instead of evaluating the energy for the entire length, which makes little sense for the short duration signal, I analyze the short time energy, which has a large and notable value. The amplitude of an audio broadcast varies over time as well. The signal's brief temporal energy provides a meaningful depiction of these amplitude variations. In general, the short time energy is described by Equation 2.

$$E_m = \sum_n [x(n)w(m-n)]^2 \quad \ldots\ldots (2)$$

The above expression can be rewritten to form Equation 3

$$E_m = \sum_n (x(n))^2 h(m-n) \quad \ldots\ldots (3)$$

Where by $h(m) = w^2(m)$

The term is understood as the impulse response $h(m)$ of a linear filter in the above expression. The nature of short-time energy representation is determined by the impulse response chosen. The bandwidth of a hamming window is double that of a rectangular window of equal length. Furthermore, the hamming window has significantly higher attenuation outside the bandwidth than the rectangular window. Expanding the window length, on the other hand, reduces bandwidth in all cases. Short-term energy is used in a variety of audio categorization difficulties. It establishes a framework for distinguishing voiced and unvoiced speech fragments in speech signals [4]. In the case of particularly high-quality speech, short-term energy features are used to distinguish speech from silence.

Implementation

Because it demonstrated a difference between impeccable and defective gear, the highest value of this is considered a feature. A MATLAB function that takes in a brief chunk of the signal collected from the gear and returns a matrix of short-time energy of length L with a user-defined window length and sample step has been created (the same step must be used for all samples). The calculated values for all samples are shown in Table 1.

6.1.2 Zero crossing rates

The zero-crossing rate (ZCR) is the rate at which a signal [2] switches from positive to negative or back, i.e. the rate at which the signal changes from positive to negative or return. This property is useful for identifying percussive sounds and has been used in speech recognition and music retrieval. The zero-crossing rate (ZCR) is a solid classification method that has lately gained traction in the literature. Its usefulness has been questioned since it became popular in [36], but it has recently been resurrected. Simply said, the ZCR is a measurement of how often a signal crosses zero in a given amount of time. The ZCR is supposed to provide information about the signal's spectral composition. A zero-crossing happens in discrete-time signals when the algebraic signs of successive samples differ. As shown in Figure 6 the rate at which zero-crossings occur can be used to determine a signal's frequency content. In a given time, interval/frame, the zero-crossing rate is the number of times the amplitude of the gear sound signals passes through zero. A representation based on the short-time average zero-crossing rate [2] can be used to get rough estimations of spectral parameters.

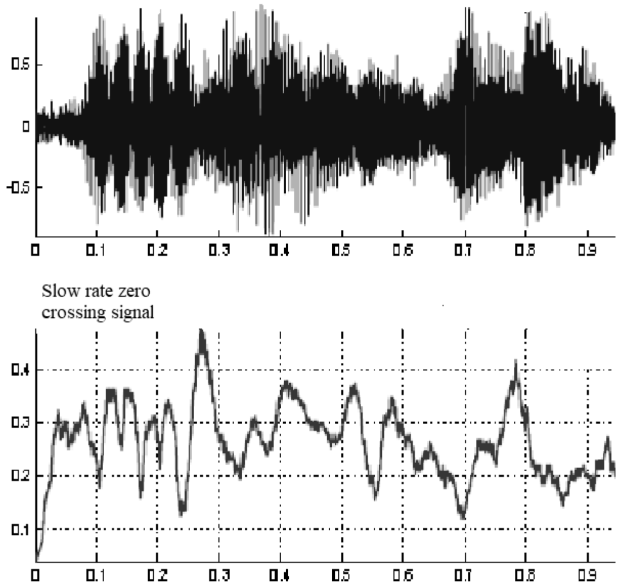

Figure 6: Definition of zero crossing

Mathematically zero crossing rates are defined by Equation 4 below;

$$z_n = \sum_{m=-\infty}^{\infty} \left|\operatorname{sgn}[x(m)] - \operatorname{sgn}[x(m-1)]\right| w(n-m) \text{ ....(4)}$$

Where

$$\operatorname{sgn}[x(m)] = \begin{cases} 1, x(m) \geq 0 \\ -1, x(m) < 0 \end{cases}$$

And

$$w(n) = \begin{cases} 0, otherwise \\ \frac{1}{2N} \; for, 0 \leq n \leq N-1 \end{cases}$$

Implementation

A MATLAB function has been programmed to calculate the zero crossing rates of a particular gear signal, Zn the mean value from Zn matrix is computed and taken as a feature; the results are included in Table 1 of results.

6.1.3 Energy Entropy

In the second half of the 18th century, German physicist Rudolf Claudius established entropy as a thermodynamic state variable [7]. Originally, it was described as reversibly received elementary heat at absolute temperature [7]. Of course, such a definition is meaningless in the context of signal processing. However, it sparked the spread of entropy as a concept in other fields. The earliest mention of entropy as a measure of system disorder was in connection with the First Postulate of Thermodynamics: "Any macroscopic system in time t_0 in certain time-invariant outside conditions will attain the so-called thermodynamic equilibrium after a relaxation period''. It's a state in which no macroscopic operations are taking place and the system's state variables have constant time-invariant values." When a system reaches thermodynamic equilibrium, its entropy reaches its maximum. Entropy became a generic measure of system disorder as a result of the main notion presented above. Entropy (or an entropy-based characteristic) can be calculated from any finite set of values, such as a parametric vector, a discrete spectral density estimate, or a digital signal segment [7]. Later, a different definition of entropy as represented by Equation 5 was given for use in mathematics, particularly statistics:

$$dS = -\sum_{k=1}^{N} p(x_k)\log_2 p(x_k) \text{ .... (5)}$$

Where $\{x_1, x_2 .... x_N\}$ is a set of random phenomena, and $p(x_k)$ is a probability of a random phenomenon x_k.The relationship between entropy and sound from gears is based on our discovery that disorganization of the sound wave from defective gears has the highest entropy value, whereas impeccable gears have a significantly lower entropy value because they are more organized and require more energy to produce in such an organized form. The entropy can also be utilized in signal processing to separate the valuable signal from background noise, according to the research.

The probability $p(x_k)$ is approximated by the difference of the spectrum component and the mean value as per Equation 6

$$p(x_k) = \left|s^{-}[k] - s[k]\right| \text{ ....(6)}$$

6.1.4 Spectral entropy

The measure of spectral flatness, also known as Wiener entropy, is used in digital signal processing to define an audio spectrum [8]. Spectral flatness is a metric for determining how noise-like a sound is. It is usually expressed in dB. Equation 7 can be used to calculate the spectral flatness of a sampled digital signal.

$$W.E = \frac{\exp\left(\frac{1}{N}\sum_{n=1}^{N} \ln x(n)\right)}{\frac{1}{N}\sum_{n=1}^{N} x(n)} \text{ .... (7)}$$

I utilized N =400 at a step of 5 for the first 2000 samples of 2048 samples for the computation, where $x(n)$ is a sequence and N is the number of samples used. In the appendix to this study, there is a function for computing spectral entropy. As a result, the geometric mean of the power spectrum is divided by the arithmetic mean of the power spectrum, as shown in equation 8. In general, Figure 7 shows the results I obtained. I discovered that impeccable gears had low spectral entropy while defective gears have greater values following our analysis. Given the notion that a badly cut teethed gear will produce more noise at the bad teeth location, the spectral entropy for defective gears should be larger than for gears. Table 1 in Chapter 2 of this study summarizes the findings. The entropy of Wiener is a pure number, meaning it has no units. White noise has an entropy value of 1 and total order on a scale of 0-1, while pure tone has an entropy value of 0.

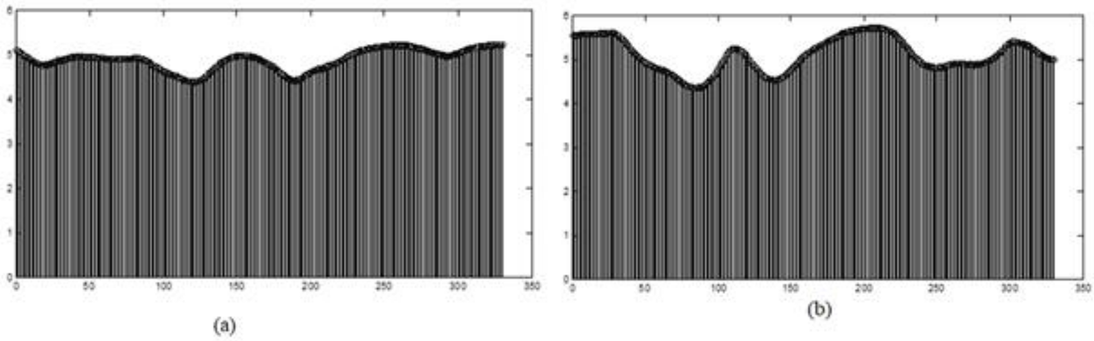

Figure 7: Typical spectral flatness for (a) Defective gear and (b) Impeccable gear

6.1.5 Pitch

Pitch is an auditory sense in which a listener allocates musical tones to relative positions on a musical scale based mostly on their impression of vibration frequency [1][2]. Pitch and frequency are linked but not synonymous. Frequency is a scientific quality that may be assessed objectively. Pitch is the subjective perception of a sound wave by each individual, which cannot be tested physically. This does not, however, imply that most individuals will disagree on which notes are higher and which are lower. I estimated the maximum of the absolute value of the spectrum of a sound wave because measuring pitch of a sound wave is challenging. Table 1 summarizes the data for both impeccable and defective gears. In general, I discovered that the value for defective gear sound is substantially higher than for impeccable gear sound. The spectrum of a piece of sound wave (not less than one revolution of a gear) is obtained using a quick Fourier transform as shown by Equation 8, and the maximum value of its absolute value is employed for training. The Fourier transform is described by equation 8, I used the maximum value of the spectrum instead of pitch, which is theoretically impossible to estimate.

$$X(k)=\sum_{n=0}^{N-1} x[n]e^{-j2\pi kn/N}, 0 \le k \le N-1 \quad (8)$$

These maximum values are complex exponential sequences, as a result, the DFT coefficients, $X(k)$ are complex numbers even if $x[n]$ are real, that is why for training, the maximum of the absolute value was computed.

6.2 Results of features extraction

Table 1 show the results obtained from 40 samples, 20 samples being impeccable gears and another 20 samples are from gears with defects. In Table 1, b stands for gear with defects and g stands for gear without defects.

Table1: Extracted sound spectrum features from 40 samples

S/N	Signal	Block Energy Entropy	Short Time Energy (Max)	Zero Crossing Rate (Mean)	Spectral Entropy (max)	Max of Spectrum
1	b1	9.732	0.1155	0.111	0.5234	6.717
2	b2	9.630	0.0807	0.124	0.4892	8.454
3	b3	9.820	0.1074	0.096	0.4123	3.081
4	b4	8.890	0.0948	0.132	0.3800	3.876
5	b5	9.821	0.0942	0.112	0.5127	6.516
6	b6	8.890	0.0628	0.098	0.4652	10.71
7	b7	9.821	0.0700	0.118	0.4653	18.29
8	b8	9.679	0.1010	0.121	0.4123	5.037
9	b9	9.693	0.1038	0.109	0.550	6.230
10	b10	9.798	0.0758	0.122	0.518	6.456
11	b11	8.912	0.1165	0.121	0.5124	6.755
12	b12	8.768	0.0817	0.134	0.4834	8.674
13	b13	9.211	0.1084	0.099	0.4223	3.671
14	b14	9.321	0.0848	0.112	0.3980	3.996
15	b15	9.672	0.0842	0.152	0.5357	6.676
16	b16	8.982	0.0728	0.098	0.4722	10.872
17	b17	8.934	0.0900	0.128	0.4853	18.291
18	b18	9.763	0.1010	0.161	0.4443	5.467
19	b19	9.643	0.1028	0.119	0.5672	6.780
20	b20	8.993	0.0858	0.132	0.5244	6.766
21	g1	10.055	0.1874	0.181	0.3281	2.670
22	g2	10.120	0.1739	0.211	0.2331	2.291
23	g3	10.040	0.1802	0.196	0.1822	2.015
24	g4	10.100	0.1689	0.187	0.2221	2.367
25	g5	9.997	0.1787	0.221	0.3101	2.120
26	g6	10.110	0.1910	0.218	0.1837	1.893
27	g7	10.130	0.1720	0.187	0.1789	2.132

28	g8	10.112	0.1598	0.204	0.1789	2.891
29	g9	10.123	0.1766	0.198	0.2362	2.891
30	g10	10.131	0.1824	0.214	0.2168	2.534
31	g11	9.978	0.1774	0.191	0.3421	2.634
32	g12	9.987	0.1639	0.221	0.2561	2.287
33	g13	10.006	0.1812	0.186	0.1982	2.023
34	g14	10.002	0.1789	0.197	0.2241	2.324
35	g15	10.010	0.1887	0.231	0.3131	2.137
36	g16	10.010	0.1810	0.208	0.1845	1.886
37	g17	10.100	0.1820	0.199	0.1879	2.142
38	g18	10.006	0.1798	0.206	0.1979	2.887
39	g19	9.899	0.1866	0.197	0.2242	2.881
40	g20	9.999	0.1924	0.202	0.2458	2.551

6.2.1 Training

I used artificial neural network (ANN) to train our data. The neural network is a computational technique based on a model of a biological neuron that accepts a bunch of electrical impulses as input and outputs an electrical pulse if and only if the sum of the incoming impulses exceeds a threshold. Groups of these modelled neurons form neural networks, which react to input and produce output. An input layer of neurons accepts an incoming parameter vector, one or more hidden layers process the data, and an output layer delivers a classification. The neural network's strength comes from the connections between the neurons, not from the neurons themselves.

Figure 8 depicts the structure of a neural network designed.

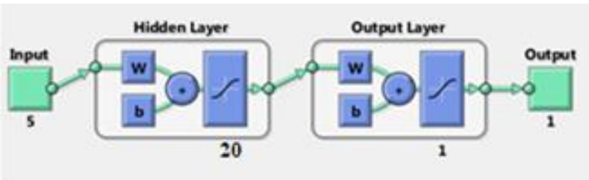

Figure 8 Structure of neural network used for training data in table 1

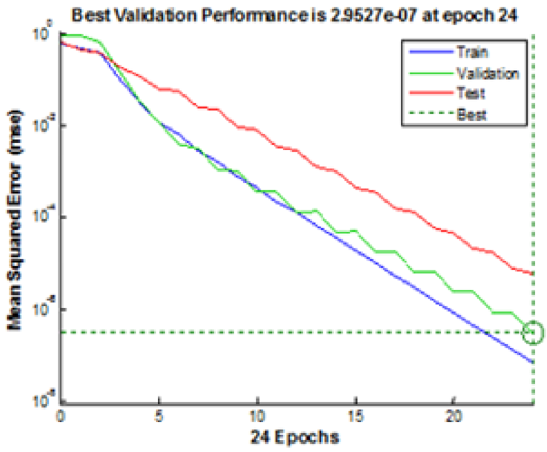

Figure 9: Training performance plot of the table 1

6.2.1.1 Training inputs

The neural network's input consists of five inputs (I have five features) for a total of twenty samples, forming a 5x20 matrix. Table 1's five columns of feature values create the input's five rows, with each row representing one feature.

6.2.1.2 Training target

Training target consists up of a 1x40 matrix as shown in Table 2, whereby each column represents the known class of an input sample, I used 0 to represent the class defective gear and 1 to represent the class impeccable gear.

Table 2: Training target

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

6.2.1.3 Training results

The training performance plot shows very good results, with little error for testing samples.

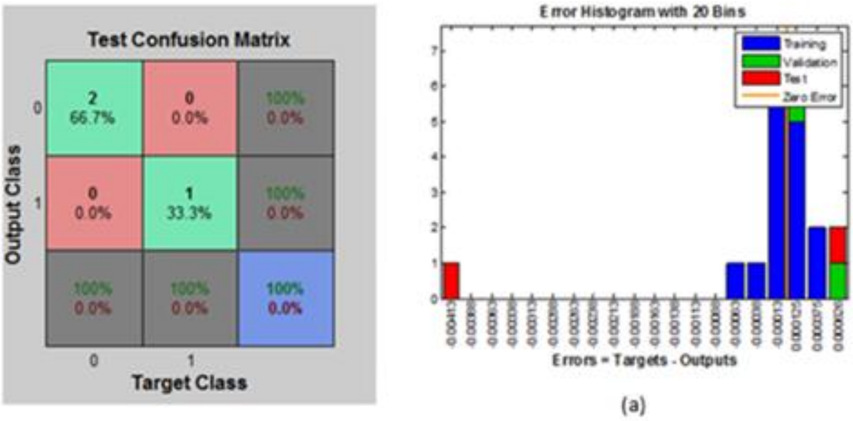

Figure 10: Training error histogram and the test confusion matrix of the table 1

Errors histogram and test confusion matrix below suggest that the training and testing results are very good.

7. Classification

7.1Definition

It's the method through which we allocate an individual object to one of several groups or classes based on its qualities [7]. The items are the noises produced by the gear; the characteristics are the attributes I extract from them (Short time energy, zero crossing rate, Spectral entropy pitch, and block energy entropy); and the classes (impeccable or defective gear) are appropriate for the challenge. Finding a proper link between features and classes [7] is the difficult part. To adequately account for between/within class changes, class models must be learned on a large number of sound samples. Over fitting occurs when the natural range of features is not well represented on the sample population: training data only covers a sub-region of the natural range, and class models are insufficient for fresh data [7].

7.2 Classification methods categories

There are two types of classification algorithms: supervised and unsupervised. A labelled set of training samples is used to "train" the algorithm in supervised classification, whereas data is sorted into some groups without the usage of a labelled training set in unsupervised classification. Another technique to categorize classification algorithms is to use parametric and nonparametric classification. In parametric approaches, the functional form of the probability density of the feature vectors of each class is known. Nonparametric approaches, on the other hand, do not assume a precise functional form in advance, instead approximating the probability density locally depending on the training data [6].

7.3 Support vector machine

Support vector machines are supervised learning models that examine data for classification and regression analysis. They come with associated learning algorithms. An SVM training algorithm creates a model that assigns new examples to one of two categories, making it a non-probabilistic binary linear classifier, given a series of training examples that are individually designated as belonging to one of two categories. [6] [7], I utilized this method for our experimentation, which was implemented in MATLAB, and the input was a 5-D feature vector, with the output being the percentage of match, with the class with the highest percentage being declared the winner class.

8. Results and Discussion

8.1 Experiment set up

I captured the sounds of 5 excellent gears and 5 defective gears from factory samples whose true state was known (figure 11), and I used the SVM classification algorithm in MATLAB to classify each of the input signals from table 2 below for experimentation only. The gear's status from SVM classier was noted, as well as the gear's actual class.

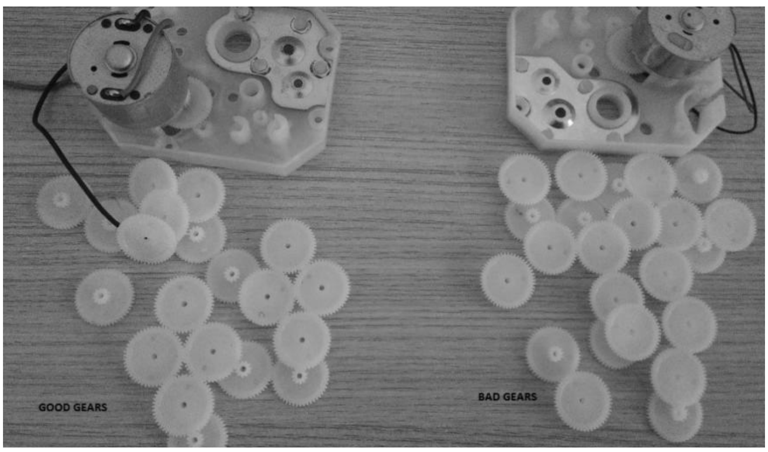

Figure 11: Impeccable and Defective gears used to extract sounds for testing

8.2 Result Discussion

The algorithm shows promising results, the efficiency can be improved by increasing the quantity of training samples and using as many features as feasible, as shown in Table 3 and our training results. To test the algorithm, a set of 20 randomly extracted sounds from gears were used to test the algorithm whereby the status of gear by visual inspection was known in advance. Table 3 shows the results, out of 20 gears only one gear was wrongly classified into the defective gear class while by visual inspection it was a impeccable gear.

Table 3: Table of experiment results

S/N	Input signal	Experimental results	Actual status
1	Defective	Defective	Defective
2	Defective	Defective	Defective
3	Good	Good	Good
4	Good	Good	Good
5	Defective	Defective	Defective
6	Good	Good	Good
7	Defective	Defective	Defective
8	Defective	Defective	Defective
9	Good	Good	Good
10	Defective	Defective	Defective
11	Good	Good	Good
12	Good	Good	Good
13	Defective	Defective	Defective
14	Good	Good	Good
15	Defective	Defective	Defective
16	Good	Good	Good
17	Good	Good	Good
18	Defective	Defective	Defective
19	Good	Good	Good
20	Good	Good	Good

Efficient of our approach is given by equation 6 below

$$efficiency = \frac{Total_{gears} - Wrongly_{classified\_gears}}{Total_{gears}} x100\% \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots (9)$$

$$efficiency = \frac{19-1}{20} x100\%$$

$$efficiency = 95\%$$

9. Conclusion

The achieved results are excellent. The efficiency can be further raised by using more samples. According to Bayes classification theories, the amount of straining samples should be very high, and the dimension of the training vector should be very huge, for any classification to approach ideal (Bayes method is an optimal one asking the statistics of the event/events to be completely known). Also, because this is a two-class classification problem, a likely hood ratio test could be used to reduce the average risk of categorization.

10. Future Scope

In the future, the algorithm will be written in python language and run in microcomputers such as raspberry pi, this is because code in MATLAB cannot be run in standalone machines, they are for research purposes only. I also recommend more training in the algorithm using as many samples as possible to approach the Bayesian theorem.

References

[1] Sanjit K. M., (2006). Digital Signal Processing: A Computer Based Approach (3rd ed). Publishing House of Electronics Industry Beijing.
[2] Theodoros , G, & Aggelos , P (2014), Introduction to Audio Analysis: A MATLAB Approach, Academic Press Elsevier Ltd, San Francisco.
[3] Erling, W., Thom, B., Douglas, K., & James, W.(1996). Content-Based Classification Search and Retrieval of Audio, IEEE Multimedia Mag., vol. 3, pp. 27-36, 1996.
[4] Brian, C. & Alex, P.(1999). Unsupervised clustering of ambulatory audio and video. In International Conference on Acoustics, Speech and Signal Processing, volume VI, pages 3037-3040. IEEE, 1999.
[5] Stanley, C., Lawrence, M.W., & James, T. E. (1994). Sensation and Perception. Harcourt Brace College Publishers, Toronto, 1994.
[6] Erkan, D. & Hamid, N. (1994). Improved musical pitch tracking using principal decomposition analysis. In icassp, volume II, pages 217-220. IEEE, 1994.
[7] Isaack Adidas Kamanga, & Peter Mwita Joseph (2018). Acceptance and Classification of Gears Based on Sound Signals Spectrum Analysis. International Journal of Advances in Scientific Research and Engineering (IJASRE), ISSN:2454-8006, DOI: 10.31695/IJASRE, 4(5), 88-98. https://doi.org/10.31695/IJASRE. 2018.32718.
[8] Shlomo, D., Gerard, A., & Ran, E.. Universal classification applied to musical signals. In ICMC, pages 333-340, 199.
pp. 193-196.

Author Profile

Isaack Adidas Kamanga is a registered professional engineer by Tanzania Engineers Registration Board (ERB), he received the MEng. at Tianjin University of Technology and Education in 20018, Tianjin China P.R. Obtained B.S.in Telecommunications Engineering from University of Dar Es Salaam in 2012, Tanzania, currently working as assistant lecturer at Dar Es Salaam Institute of Technology (DIT) as well as field engineer in fields of RF, VSAT, Optic fiber and computer networks.