DEVELOPMENT OF VISUAL ODOMETRY BASED MACHINERY MOTION ASSESSMENT SYSTEM

LOW SHEE TENG

RESEARCH REPORT SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF MECHANICAL ENGINEERING

FACULTY OF ENGINERRING
UNIVERSITY OF MALAYA
KUALA LUMPUR

2022

UNIVERSITY OF MALAYA
ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: Low Shee Teng

Matric No: S2039851

Name of Degree: Master of Mechanical Engineering

Title of Project Paper/Research Report/Dissertation/Thesis (“this Work”):

Development Of Visual Odometry Based Machinery Motion Assessment System

Field of Study: Vibration, Computer Vision

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya (“UM”), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate’s Signature Date: 9/3/2022

Subscribed and solemnly declared before,

Witness’s Signature Date: 9/3/2022

Name:

Designation:

DEVELOPMENT OF VISUAL ODOMETRY BASED MACHINERY MOTION ASSESSMENT SYSTEM

ABSTRACT

Monitoring the vibrations of a machine's mechanical components is critical to its proper operation as for performing preventive maintenance. Recently, a sizable number of the study approaches in vibration analysis are based on non-contacting vibration measuring equipment that offering various advantages than the conventional sensors. New methods for gathering information about the vibration of the machine have evolved simultaneously with the constant improvement of the visual odometry (VO) systems due to the rapid development of computer vision (CV) field. Digital image analysis, video analysis, and other visual inputs are all examples of CV, which is a branch of artificial intelligence (AI) that empowers computers or systems to obtain significant information from digital images, videos, as well as other visual inputs and to take actions or make recommendations based on this information. Research laboratories to actual industrial installations were made possible because of their actual effectiveness. The use of visualization tools can often be a useful addition to vibration analysis or even a complete replacement for more traditional approaches. The non-contacting attributes and the ability to simultaneously observe several spots in the defined region are the most important factors. Motion magnification (MM), an image processing technique that provides the visual observation of vibration processes that are not visible in their native state, is an image processing technology. Four types of methodologies involving optical flow (OF), motion amplification and MM have been implemented and linear based Eulerian Video Magnification (EVM) have been implemented as a benchmark. Method 1 include the calculation of OF follow by motion amplification on the video. Method 2 would be the same as Method 1 but including the insertion of the cut-off frequency. Method 3 would be combining Method 2 with linear based EVM, and Method 4 would be purely linear based EVM. These algorithms are implemented in terms of their computational complexity and visual quality as well as how they provide the amplified motion of video output. Machine diagnostics can be improved by using visual methods that magnify motion. Motion amplification aids in the visualization of complex vibration problems that are otherwise inaccessible to the human eye. When used in conjunction with other tools, this instrument can save time and money in the areas of routine condition monitoring programs, troubleshooting, vibration analysis and root cause analysis. In this research, the output of the video amplifying and magnifying algorithm have been compared. According to the findings, EVM is the most appropriate VO for a machinery motion assessment system because it has performed the best magnification work in this project. The EVM technique has the best magnification when comparing the data acquired from these approaches; nonetheless, the EVM method exhibits superior noise characteristics than Method 3. Method 2 outperforms Method 1 in terms of edge distortions, but the results are foggy at the end of the system because of the blurring that occurs at the end of the system.

Keywords: Vibration, Computer Vision (CV), Visual Odometry (VO), Optical Flow (OF), Motion Magnification (MM).

PEMBANGUNAN SISTEM PENILAIAN GERAKAN MESIN BERASASKAN ODOMETRI VISUAL

ABSTRAK

Memantau getaran komponen mekanikal mesin adalah penting untuk operasi yang betul seperti untuk melaksanakan penyelenggaraan pencegahan. Baru-baru ini, sejumlah besar pendekatan kajian dalam analisis getaran adalah berdasarkan peralatan pengukur getaran tidak bersentuhan yang menawarkan pelbagai kelebihan berbanding penderia konvensional. Kaedah baharu untuk mengumpul maklumat tentang getaran mesin telah berkembang serentak dengan peningkatan berterusan sistem odometri visual (VO) disebabkan perkembangan pesat bidang penglihatan komputer (CV). Penglihatan komputer ialah bidang kecerdasan buatan (AI) yang membolehkan komputer dan sistem memperoleh maklumat yang bermakna daripada imej digital, video dan input visual lain dan mengambil tindakan atau membuat cadangan berdasarkan maklumat tersebut. Makmal penyelidikan ke pemasangan industri sebenar telah dibuat kerana keberkesanan sebenar mereka. Penggunaan alat visualisasi selalunya boleh menjadi tambahan berguna kepada analisis getaran atau malah pengganti lengkap untuk pendekatan yang lebih tradisional. Sifat tidak bersentuhan dan keupayaan untuk memerhati beberapa tempat secara serentak di rantau yang ditentukan adalah faktor yang paling penting. Pembesaran gerakan (MM), teknik pemprosesan imej yang membolehkan pemerhatian visual proses getaran yang tidak kelihatan dalam keadaan asalnya, ialah teknologi pemprosesan imej. Empat jenis metodologi yang melibatkan aliran optik (OF), penguatan gerakan dan MM telah dilaksanakan dan Pembesaran Video Eulerian (EVM) berasaskan linear telah dilaksanakan sebagai penanda aras. Kaedah 1 termasuk pengiraan OF diikuti oleh penguatan gerakan pada video. Kaedah 2 akan sama dengan Kaedah 1 tetapi termasuk sisipan frekuensi potong. Kaedah 3 akan menggabungkan Kaedah 2 dengan EVM berasaskan linear, dan Kaedah 4 adalah EVM berasaskan linear semata-mata. Algoritma ini dilaksanakan dari segi kerumitan pengiraan dan kualiti visualnya serta cara ia menyediakan gerakan diperkuatkan output video. Diagnostik mesin boleh dipertingkatkan dengan menggunakan kaedah visual yang membesarkan gerakan. Penguatan gerakan membantu dalam visualisasi masalah getaran kompleks yang sebaliknya tidak boleh diakses oleh mata manusia. Apabila digunakan bersama alat lain, instrumen ini boleh menjimatkan masa dan wang dalam bidang analisis getaran, program pemantauan keadaan rutin, penyelesaian masalah dan analisis punca. Dalam penyelidikan ini, pengeluaran dari algoritma penguatan dan pembesaran video telah dibandingkan. Mengikut penemuan, EVM adalah VO yang paling sesuai untuk sistem penilaian gerakan jentera kerana ia telah melakukan kerja pembesaran terbaik dalam projek ini. Teknik EVM mempunyai pembesaran terbaik apabila membandingkan data yang diperoleh daripada pendekatan ini; namun begitu, kaedah EVM mempamerkan ciri hingar yang unggul daripada Kaedah 3. Kaedah 2 mengatasi Kaedah 1 dari segi herotan tepi, tetapi hasilnya berkabus pada penghujung sistem kerana kekaburan yang berlaku pada penghujung sistem.

Kata Kunci: Getaran, Penglihatan Komputer (CV), odometri visual (VO), aliran optik (OF), Pembesaran gerakan (MM)

ACKNOWLEDGEMENTS

I would want to convey my gratitude and respect to all those who offered me the opportunities and chance to complete this report. A very special thanks to my supervisor, Ir. Dr. Alex Ong Zhi Chao whose help, giving suggestion and encouragement, help me to have a clear mindset on how to carry out my work progress efficiently and finish my project on time. I would only be able to be successful as he provided me with gentle direction and suitable guidance for this undertaking. His supervision and coaching were invaluable throughout this project, and he helped me to write this report flawlessly.

Finally, a special thanks I would also like to give to my course mates who gave me so much motivation and cooperation mentally in completing this report. With the supports and encourages that given by my friends and my family, I was able to overcome the tensions and the problems I faced when I am completing this report. Finally, I would like to show my highest gratitude to everybody who was important to this research project, I would want to express my regret for not being able to personally thank each one of you.

TABLE OF CONTENTS

Abstract		III
Abstrak		IV
Acknowledgements		V
Table of Contents		VI-VIII
List of Figures		IX-XII
List of Tables		XIII
List of Symbols and Abbreviations		XIV-XVII
List of Appendices		XVIII

CHAPTER 1: INTRODUCTION

1.1	Background of Study	1-4
1.2	Problem Statement	4-5
1.3	Objective	5-6

CHAPTER 2: LITERATURE REVIEW

2.1	Maintenance Philosophies and The Role of Vibration Analysis	7-10
	2.1.1 Vibration Characteristic	10-12
2.2	Vibration Analysis Overview	12-15
	2.2.1 Vibration Analysis Profiles	15-16
	2.2.1.1 Theoretical Vibration Profiles	16-17
	2.2.1.2 Actual Vibration Profiles	18
	2.2.1.2.1 Time-Domain Analysis	19-20
	2.2.1.2.1.1 Visual Inspection of Time-Domain	20-22
	2.2.1.2.1.2 Feature-Based Inspection	22-26
	2.2.1.2.2 Frequency-Domain Analysis	26-30
	2.2.2 Vibration Measurement Parameters and Vibration Severity Criteria	30-32
	2.2.3 Vibration Analysis and Measuring Equipment	32
	2.2.3.1 Contacting and Non-Contacting Vibration Measurement Method	33-35

	2.2.3.2 Existing Vibration Measurement and Limitations	35-41
2.3	Visual Based Measurement System	42-43
	2.3.1 Visual Odometry System	43-48
	2.3.1.1 Appearance-Based Technique	48-50
	2.3.1.1.1 Optical Flow Method	50-52
	2.3.1.1.1.1 Lucas-Kanade and Gunnar Farnebäck Algorithm	52-53
	2.3.2 Improvement of Visual Odometry (VO) System	53-54
2.4	Motion Magnification (MM) Technique	54-58

CHAPTER 3: METHODOLOGY

3.1	Introduction	59-62
3.2	Calculation of Optical Flow (OF) Method Using Gunnar-Farnebäck Algorithm	62-63
3.3	Power Spectral Density (PSD) Analysis	63-66
3.4	Video Warping	66
3.5	Eulerian Video Magnification (EVM)	67-69
	3.5.1 Spatial-Temporal Information Processing	69-70
	3.5.2 Relation Between Temporal Filtering and Magnification	70-72
	3.5.3 Temporal Filter	72
	3.5.4 Pyramid Decomposition	73
3.6	Different Methods Implemented for Video Processing	74-80
3.7	Implementation of Software and Programming Code	80
	3.7.1 OpenCV Software	80-82
	3.7.2 Python Language	83-84

CHAPTER 4: RESULTS AND DISCUSSION

4.1	Results of Sleeping Baby Video	85
	4.1.1 Results of Method 1	85-86
	4.1.2 Results of Method 2	86-87
	4.1.3 Results of Method 3	88

	4.1.4 Results of Method 4	89
4.2	Results of Vibrating Guitar String Video	89-93
	4.2.1 Results of Method 1	89
	4.2.2 Results of Method 2	90
	4.2.3 Results of Method 3	92-93
	4.2.4 Results of Method 4	93
4.3	Results of Vibrating Phone Video	94
	4.3.1 Results of Method 1	94
	4.3.2 Results of Method 2	95-96
	4.3.3 Results of Method 3	96-97
	4.3.4 Results of Method 4	97

CHAPTER 5: CONCLUSION AND RECOMMENDATION

5.1	Conclusion	98
5.2	Future Recommendations	99-100
5.3	Sustainability	100-101
5.4	Complexity	101-102
5.5	Lifelong Learning	102

References	103-117
Appendix	118-119

LIST OF FIGURES

Figure 2.1	Vibration Profile	17
Figure 2.2	Small Oscillations of a Simple Pendulum, Harmonic Function	17
Figure 2.3	Roller bearing's time-domain vibration signal	
(A)	Brand-New Condition	21
(B)	Inner Race Fault Condition	21
Figure 2.4	Roller bearing's time-domain vibration signal	
(A)	Worn But Undamaged Condition	22
(B)	Outer Race Fault Condition	22
Figure 2.5	Frequency-Domain Components Versus Time-Domain Measurements	27
Figure 2.6	An Example of Frequency-Domain Representations	28
Figure 2.7	Basic Accelerometer	36
Figure 2.8	Photo of Accelerometer Sensor	36
Figure 2.9	Magnetic Field of Proximity Sensor	37
Figure 2.10	Photo of Different Sizes of Proximity Sensors	38
Figure 2.11	Beam Deflection Method for LDV Sensor	39
Figure 2.12	Photo of LDV Sensor	39
Figure 3.1	Flow Chart of Project	60
Figure 3.2	PSD Obtained from Phone Video	66
Figure 3.3	Overall Structure of The Linear-Based-EVM	69
Figure 3.4	Eulerian Motion Magnification Process Flow	70
Figure 3.5	Spatial Translation Can Be Approximated Using Temporal Filtering	71

Figure 3.6	Sleeping Baby	75
Figure 3.7	Guitar	75
Figure 3.8	Vibrating Phone	75
Figure 3.9	Process Flow Chart for Method 1	77
Figure 3.10	Process Flow Chart for Method 2	78
Figure 3.11	Process Flow Chart for Method 3	79
Figure 3.12	Four Main Components of OpenCV	81
Figure 3.13	Processing Algorithm Using OpenCV	83
Figure 3.14	Logo of Python Language	84
Figure 4.1	The filtered OF baby video screenshot	86
Figure 4.2	Motion amplification of the OF baby video Amplification factor of (a) 10 (b) 100	86
Figure 4.3	PSD of pixel intensity of original baby video	87
Figure 4.4	PSD of pixel intensity, magnitude, and angular component of original baby video.	87
Figure 4.5	Screenshot of output baby video for Method 2 using amplification factor of (A) 10 (B) 100	88
Figure 4.6	Screenshot of Output Baby Video for Method 3	88
Figure 4.7	Screenshot of Output Baby Video for Method 4	89
Figure 4.8	The filtered OF guitar video screenshot	90
Figure 4.9	Motion amplification of the OF guitar video using amplification factor of	90

(a) 50

(b) 100.

Figure 4.10 PSD of pixel intensity of original guitar video. 91

Figure 4.11 PSD of pixel intensity, magnitude, and angular component of original guitar video. 91

Figure 4.12 Screenshot of output guitar video for Method 3. 92

Figure 4.13 Screenshot of output guitar video for Method 4. 93

(a) amplification factor=50, cut-off frequency= 72Hz, 92 Hz

(b) amplification factor=100, cut-off frequency =100Hz, 120 Hz

Figure 4.14 The Filtered OF Phone Video Screenshot 93

Figure 4.15 Motion amplification of the OF Phone Video. Amplification factor of

(a) 20

(b) 100 94

Figure 4.16 PSD of pixel intensity of original phone video 95

Figure 4.17 PSD of pixel intensity, magnitude, and angular component of original phone video 96

Figure 4.18 Screenshot of output phone video for Method 2 96

Figure 4.19 Screenshot of output phone video for Method 3 97

Figure 4.20 Screenshot of output phone video for Method 4 97

LIST OF TABLES

Table 2.1	Frequent Use Function of Time-Domain and Frequency-Domain Analysis	18
Table 2.2	Summarization of Contacting and Non-Contacting Vibration Measurement Method Equipment	34
Table 2.3	Comparison of Base Constrains of Accelerometer and Visual Motion Sensor	41
Table 2.4	Comparison of Monocular and Stereo VO Configurations	44-45
Table 2.5	Comparison of Feature-Based, Appearance-Based, And Hybrid-Based VO	46-47
Table 3.1	Video Properties	76

LIST OF SYMBOLS AND ABBREVIATIONS

%	-	percentage
X_0	-	maximum amplitude
X_p	-	peak amplitude
g^2	-	spectrum of amplitude
g^2/Hz	-	acceleration per Hertz (Power Spectral Density Function)
x_{CF}	-	crest factor (CF)
x_{IF}	-	vibration signal impulse factor
x_{KURT}	-	fourth normalized central statistical moment
x_{RMS}	-	root mean square amplitude
x_{SF}	-	shape factor
x_{SK}	-	third normalized central statistical moment
x_i	-	element of signal x
x_{max}	-	maximum positive peak amplitude
x_{min}	-	maximum negative peak amplitude
1D	-	One-Dimensional
2D	-	Two-Dimensional

3D	-	Three-Dimensional
CCD	-	Charged-Coupled Device
cpm	-	cycles per minute
cps	-	cycles per second
CWT	-	continuous wavelet transforms
DFT	-	Direct Fourier Transform
EVM	-	Eulerian Video Magnification
FA	-	Factional Anisotropy
FFT	-	Fast Fourier Transform
GPS	-	Global Positioning Systems
high GUI	-	high-level graphical user interface
Hz	-	Hertz
I	-	Intensity
inch/s2	-	inches per square second
inch/sec	-	inches per second
kcpm	-	kilocycles per minute
LDV	-	Laser-Doppler Vibrometer

LVDT	-	Linear Variable Differential Transformer
MEMS	-	Micro Electromechanical Systems
MM	-	Motion Magnification
mm/s2	-	millimeters per square second
mm/sec	-	millimeters per second
NPT	-	Non-destructive testing
OF	-	Optical Flow
OpenCV	-	Open-Source Computer Vision Library
PDF	-	Probability Density Function
PSD	-	Power Spectral Density
PVM	-	Phased-based Video Magnification
RGB-D	-	Combination of a RGB image and its corresponding depth image
rpm	-	revolutions per minute
SHM	-	Structural Health Monitoring
SNR	-	signal-to-noise ratio
T	-	period
VO	-	Visual Odometry

XML	-	extensible markup language
L	-	overlapping segments
N	-	number of sampled points
X	-	vibration displacement
a	-	acceleration
r	-	magnitude of flow
t	-	Time (seconds)
v	-	velocity
$\bar{x}$	-	mean amplitude
θ	-	direction of flow
ω	-	circular frequency

LIST OF APPENDICES

(I) Gantt Chart 118-119

CHAPTER 1: INTRODUCTION

1.1 Background of Study

Material or system breakdown is inevitable when a machine is subjected to high levels of vibration (Schwarz *et al*., 2020). The physical parameters of any mechanical system or structure, such as mass, stiffness, and damping, are strongly linked to its vibration. In fact, its vibration is strongly influenced by changes in its physical properties. If any physical property changes, the structure will vibrate in a different way. It is also possible for the structure to vibrate in different ways if a boundary condition, such as a loose mounting, changes. Structure vibration analysis is extremely important in machinery motion assessment system including the evaluation of performance and the detection of faults in structures (Li *et al*., 2015). Traditional vibration analysis methods for the identification of vibration characteristics, which were used prior to the usage of vibration video records, were often based on one or more sensors. A structure's vibration is normally detected by attaching vibration sensors to its surface which known as contacting sensor or by detecting the structure's surface motion with a non-contacting sensor such as a laser vibrometer, depending on the method used (Li *et al*., 2016). Surface vibration has also been measured using non-contacting sensors, such as laser vibrometers, optical sensors, and among others. Every one of these ways is more expensive and time demanding than just recording a video, and their use may even be prohibitive in some situations when shooting a video is still feasible. Those approaches required the use of complex detection systems and have the potential to alter the structure's basic dynamic features to a certain extent. Nowadays, visual vibration measurement system is becoming increasingly used as a non-contact, wide-range vibration measurement approach corresponding to development in camera technology and image processing (Yu *et al.,* 2017). A significant amount of the visual measurement outcome is influenced by the camera performance and picture feature extraction, both of which are important factors.

The vision-based system offers several benefits over conventional sensors, including remote monitoring, cheap cost, and the ability to measure many points (Zhu *et al.*, 2020). An evaluation of computer vision-based systems such as vision-based detection systems, deformation measuring methods, and tracking approaches has recently been conducted. (Xu and Brown, 2018). Laboratory tests have proven that the vision-based system incorporates a cameras and computers for recording and calculating (Fukuda, Feng, & Shinozuka, 2010). As a result of its simplicity and cheap resources, camera-based motion estimation has acquired a lot of appeal in recent years. When it comes to motion estimation, visual odometry (VO) refers to a sort of approaches that utilize visual cues to estimate motion. They may be monocular, stereo, RGB-D, or any combination of these three types of sensors, with a variety of algorithm frameworks in each case (Poddar *et al.*, 2018). From gaming and virtual reality to wearable computing, industrial production, healthcare to underwater and aerial robotics to driver assistance systems and agricultural field robots to automotive and pedestrian navigation, visual odometry may be used in a broad range of contexts (Scaramuzza and Fraundorfer, 2011). For estimating a moving object's speed in space, the optical flow method employs the pixels moving speed under the picture grey model. Optical flow analysis relies on the Horn-Schunck algorithm and a modified version of it. VO systems using optical flow and deep learning have been developed and advocated in several papers. (Muller and Savakis, 2017). As an input to a convolutional network, optical flow pictures provide rotation and displacement values for each pixel in the image. A small-amplitude oscillation that isn't visible to the human eye was documented in some of the research (Hyatt and Lee, 2019). Using an amplification factor to multiply the optical flow, the researcher warps the video, resulting in a large-amplitude depiction of the original moving. One approach to reduce the amount of noise and isolating frequencies in a video is to analyses the video's power spectral density (PSD).

For vibration analysis, the small vibration motion that was previously imperceptible to the human eye may now be seen using video motion magnification methods, such as the swaying of buildings in the wind or the vibrations of aero plane wings. Understanding the structural health of a building (Cha *et al.*, 2013) and taking a person's vital signs (Balakrishnan *et al.,* 2013) are both made possible by being able to detect small motions. As a result of the magnification, there is a greater chance noise or excessive blurring. Visual motion magnification (MM) methods make it possible for humans to observe fast-moving objects in detail. MM is a technique for seeing motion that works similarly to a microscope. It can enhance small movements in a video sequence, allowing deformations that would otherwise be imperceptible to be seen. As the motions are so small, it is impossible to tell them apart from background noise. This leads to noisy outcomes and excessive blurring in existing video magnification algorithms when the magnification factor is high (Wu *et al*., 2012). Filters created by hand are still used to extract representations in current technology. Motion magnifying techniques have an important part which is multi-frame temporal filtering, which helps separate important motions and reduces the magnifying effect of background noise (Oh *et al*., 2018). A prevalent technique in modern video magnification systems is the decomposition of video frames into representations that enable them to magnify motion. Hand-designed filters, such the complicated steerable filters (Freeman and Adelson, 1991) are often used to decompose them, although this isn't always the best option.

Techniques for motion magnification fall into two categories: Lagrangian and Eulerian methods (Oh *et al.,* 2018). The Lagrangian technique takes the motion field (optical flow) openly and utilizes it to directly move the pixels (Liu *et al*., 2017). Furthermore, Eulerian techniques divide video frames into representations that permit motion manipulation without needing explicit tracking (Wu *et al*., 2012). These approaches are often divided into three procedures: deconstructing frames into an alter representation, modifying the representation, and rebuilding the altered representation into enlarged frames (Oh *et al*., 2018). As a result, they should be sensitive to noise and typically display significant blurring. In this study, the method with the fewest edge distortions and the best noise characteristics will be chosen. Distortion can be described as when straight lines appear bent or curvy in photographs and it is a form of optical aberration after motion magnification. Multi-frame temporal filtering applied to structures is a fundamental component of the best MM techniques, helping to isolate significant movements and reduce noise.

1.2 Problem Statement

Typically, traditional vibration analysis methods identifying vibration characteristics are utilizing on one or more sensors. Surface vibration has also been measured using non-contact sensors including optical sensors and laser vibrometers. All these approaches are more expensive and time consuming than video recording, and their use may even be prohibitively expensive in some circumstances when video recording is still available (Schwarz *et al*., 2020). In some buildings, laser displacement meter as non-contacting sensors are employed. While laser sensors eliminate the danger and complexity associated with the repairing procedure, their limited range and expensive cost restrict their use. Additionally, these non-contact traditional techniques are only suitable for single-point measurement. In comparison to conventional sensors, several benefits have been provided by vision-based system including remote monitoring, cheap cost, and multi-point measurement (Zhu *et al*., 2020). As a result, research on cost-effective measurement methods using non-contact sensors such as visual odometry system is gaining momentum. Proximity sensors comparing to optical cameras are inexpensive sensors that capture a wealth of useful data. Additionally, they are passive; visual localization systems are immune to the interferences that frequently occur during utilization of active ultrasonic or laser proximity sensors (Frontoni 2006). Image analysis and processing methods of the modern era have expanded the already-existing capabilities of vibration monitoring. Vibration measurement, which involves the use of sensors and instruments to convert the read-off signal to a form that a researcher can analyze, also an indirect approach of monitoring physical processes in action (Śmieja *et al*., 2021). To observe vibration processes that are not apparent in their original state, visual MM deforms input pictures in a way that allows for their visual observation. One of the basic goals of MM is to solve or overcome the barrier of human perception of changes in the viewed objects. Small movements with amplitudes below a threshold are difficult for humans to detect. The observation of small dynamic movements is nearly impossible, despite the existence of optical devices for seeing small static physical characteristics. MM enables us to observe and comprehend crucial biological and physical motions. MM also renders small motions in a video sequence such that they can be seen clearly and quantifies the motions so that they can be researched. This requires rigorous noise analysis to minimize the amplification of spurious signals. Quantitative estimate of the kinematic properties of the seen objects may be performed using the established algorithms and their additional features (Wadhwa *et al*., 2017). As depicted by the visualizations in this study, movements that are imperceptible to the human eye can be seen.

1.3 Objective

1. To develop the visual odometry system as non-contact vibration measurement method for machinery motion assessment

2. To investigate the suitable visual odometry technique for machinery motion assessment system

3. To investigate the suitable motion magnification algorithms that gives the fewest edge distortions and the best noise characteristics

CHAPTER 2: LITERATURE REVIEW

2.1 Maintenance Philosophies and the role of Vibration Analysis

An industrial machinery maintenance worker must be aware of and correct any problems as soon as they arise so that the machinery can operate safely and productively and avoid any financial disaster or personal injury or death. The primary goal of maintenance is to reduce downtime, improve safety, and increase availability of production equipment. (Senapaty and Rao, 2018). Vibration analysis may help prevent costly downtime due to a machine failure by allowing planned downtime to be used to fix or replace malfunctioning components. (Saied *et al.,* 2015). Because of this, both time and money can be saved. The following are some of the most common maintenance philosophies in use by industries (Senapaty and Rao, 2018).:

- Preventive maintenance: Time-based or fixed time maintenance are other terms used to describe this sort of maintenance. Equipment's maintenance is performed on a predetermined schedule, whether it is based on the calendar, the number of hours the equipment is in operation, or the amount of equipment cycles. This maintenance aids in the reduction of failures and the planning of work, spare parts, and labor resources. It is more expensive and will cause frequent downtimes because of the nature of the maintenance, despite its advantages. Therefore, it should only be utilized on machines that are showing signs of ageing; otherwise, superfluous parts and components may need to be replaced.
- Reactive maintenance: It is sometimes referred to as On-failure or Break-down maintenance when it occurs during a breakdown. This form of maintenance is carried out during equipment shutdowns after a failure of the machinery. This ideology is characterized by a lack of warning signs of failure, unplanned plant outages and

production losses or delays, as well as a high inventory of replacement parts that must be kept on hand in case of an emergency.

- Proactive maintenance: Predictive maintenance, also known as condition-based maintenance, is another term for this sort of maintenance. Since machine defects may be discovered before failure occurs, this is extremely beneficial to the plant, since it eliminates the need for regular planned maintenance. This approach to machine maintenance needs constant and continuous inspections of the machine's overall health status. Vibration analysis, infrared thermography and ultrasound may be used to identify machine health degradation and component issues before they cause the plant to be shut down, preventing costly downtime. Typically, this type of thinking is applied to critical industrial machinery. In addition, it aids in reducing downtime, identifying the underlying cause of the failure, and optimizing the use of resources such as spare parts and workers.

Predictive maintenance philosophy is known as a good approach to manage and sustain most of the materials and devices in industry or plant, but only when effectively implemented and in the right areas; otherwise, it may be highly costly. It is critical to identify equipment problems before they fail so that their failure does not impact the condition of other equipment that is working together and causes further harm to the machine. Predictive maintenance includes tracking and evaluating machine performance measures to spot potential issues before they have a chance to become catastrophic. A significant failure can be avoided if worn-out components are replaced in a timely manner. It may reveal the presence of contaminants such as moisture, dirt, and oil in systems, allowing the appropriate action to be taken (Mobley, 2002). Conditions can be monitored by using many factors as a preliminary step.

An initial collection of parameters can be utilized to establish the baseline for a condition monitoring system. Vibration in rotating machinery, particularly high-speed devices like turbines and compressors, is one of these factors. A wide range of industries have learned that vibration monitoring provides significant advantages, including cost savings, reduced maintenance, and increased equipment reliability (Thomson, 1993). Furthermore, the vibration produced by rotating machines is not advantageous. In addition to causing excessive wear, cracking, and fastener loosening, it can also produce excessive noise among other complications. Severe vibration can cause catastrophic failure in aircrafts, resulting in the death of passengers and crew. For detecting vibration patterns that will result in failure of machines, vibration monitoring is used.

Many academics have investigated the monitoring and analysis of vibrations in many applications. There are lot of research that target to build a self-adjusting and integrative monitoring system that could function that under wide range of working conditions while spending the least number of resources. Hansen and Gao explored the vibrational behavior of a deep groove ball bearing with a structurally integrated force sensor (Hansen and Gao, 2000). Their experimental experiments on a ball bearing were carried out for check and prove the numerical and analytical answers. The advancement in condition-based maintenance and vibration measurements over the course of seventy years was discussed in a technical essay (Mitchell, 2007). Shaft misalignment was examined as a possible cause of centrifugal blower vibration. (Alzadjali and Rameshkumar, 2013). Sheng and Veers discovered that for condition monitoring of wind turbines in the wind energy generation area, the most efficient instrument would be the vibration analysis (Sheng and Veers, 2011). Using a particle filtering approach, the researchers built an online lubrication monitoring software for a wind turbine to minimize energy costs while increasing the availability and dependability of the turbine

(Zuhn *et al*., 2011). Predictive monitoring using vibration analysis is the focus of this paper, which also includes illustrative case studies.

2.1.1 Vibration Characteristic

Violating its equilibrium state, a mechanism or a part of a machine is described as vibrating. Using the basic harmonic motion theory, one can see how a mass on the end of a spring may duplicate this vibratory motion (Inman, 1994). Vibration can be described using a wide range of terminology. The vibration's frequency is expressed in cycles per second and is commonly expressed in Hertz units. The distance that a vibrating element moves back and forth is what we mean by the term "amplitude." When the amplitude increases, so do the mechanical problems appears. For determination of amplitude, displacement, acceleration, and velocity can be used. A peak-to-peak displacement is used, and it can be expressed as a measurement in mils, millimeters, or inches. Velocity peak is the rate of change of the displacement against time, measured in millimeters per second (mm/sec) or inches per second (inch/sec). In mm/s2 or inch/s2, the peak acceleration represents the rate of change in velocity with respect to time. Each of the vibration characteristics exposes a unique aspect of the vibration. The following are the reasons why vibration measuring becomes necessary in practice (Rao, 2011):

1. In many situations, the ability of a structure or machine to withstand a given level of vibration is a crucial consideration. If the building or machine can carry out the intended task even after testing has been completed in the specified vibration environment, it is predicted to withstand the circumstances.
2. Structural engineers use data on earthquake-induced ground vibrations, wind speed fluctuations, random variations in ocean wave heights, and road surface roughness to design better machines.

3. Due to the increasing demands for increased productivity and more cost-effective design, higher working speeds of machinery and more efficient use of materials are being achieved using lightweight constructions. With these developments, resonance circumstances become increasingly common during the operation of machinery, hence decreasing the system's dependability, and increasing its vulnerability. To guarantee proper safety margins, it becomes necessary to measure the vibration characteristics of machinery and structures on a regular basis. Any detected shift in the natural frequencies or other vibration characteristics will signal either a breakdown of the equipment or the need for repair.
4. For the sake of simplicity, multidegree-of-freedom systems is continuous system. If a system's measured natural frequencies and modes match up with the model's computed natural frequencies and modes, the approximation can be accurate.
5. A machine or structure's theoretically calculated vibration characteristics may differ from their real values because of analysis assumptions.
6. To properly design and operate active vibration-isolation systems, the frequencies of vibration and the forces that have been induced must be measured.
7. To detect a system's mass, stiffness and damping by monitoring the system's input and output vibration characteristics is conceivable.
8. The determination of the natural frequencies from the machine is essential in determining the operational speeds of adjacent machinery to avoid resonance circumstances.

2.2 Vibration Analysis Overview

It is necessary to do vibration analysis on equipment to determine its mechanical and functioning conditions. A significant advantage of vibration analysis is that it can perform detection on developing issues before they become severe and create unscheduled downtime. Monitoring machine vibrations on a regular or predefined basis may aid in the achievement of this purpose. Mechanical looseness, damaged bearings and broken gears can be detected using vibration monitoring when performed on a regular basis (Alsalaet, 2012). Detection of misalignment and unbalance condition prior to the shaft or bearing failure can be implemented by vibration analysis. Imprecise rotor balancing, inaccurate shaft alignment, and false bearing installation are all examples of inadequate maintenance procedures that can be identified by trending vibration levels.

Problems are produced by all rotating machines because of the machine's dynamics, which includes factors such as balancing or alignment of the spinning parts. At frequencies, vibration measurements may provide important information about the quality of shaft alignment and balance, the condition of bearing, and the influence on the machinery produced by reverberation from pipes, enclosures, and other structures on the machine. (Alsalaet, 2012). A non-intrusive approach of monitoring machine status during start-up, shutdown, and normal operation, vibration measuring has proven to be effective and cost-efficient. When it comes to rotating equipment, vibration analysis is most employed on motors, pumps, compressors, paper machines, rolling mills, machine tools, and gearboxes, among other things. Reciprocating machinery, such as large diesel engines and reciprocating compressors, may now be studied in a limited capacity due to recent technical breakthroughs. Other procedures are also required for these machines to completely monitor their operation.

Vibration analysis method in detection mode is typically composed of four fundamental components:

1. A transducer

2. A signal analyzer

3. Analysis software

4. A computer for data analysis and storage

Three different types of systems can be built using these essential components: continuous online systems, periodic analysis systems utilizing portable equipment, and a multiplexed system that checks a group of transducers at specified time intervals. Because of the extra wiring and multiplexing required, multiplexed, and hardwired systems are more costly per point of measurement. The criticality of the equipment and the usefulness of continuous or semi-continuous measurement data play a role in determining the most practical and suitable design for a specific application.

In diagnosis mode of vibration analysis, it is not uncommon for workers in factories and shops to hear or feel strange noises or vibrations (Vishwakarma *et al.,* 2017). Vibration analysis could be used to find out whether there is a significant issue. If a problem is found, more spectrum analysis could be implemented to precisely describe the issue and predict the machine operation time can run before a catastrophic failure occurs. Low-priority equipment can save money and resources by analysis (diagnosis) mode of vibration measurements. Someone must be attentive to unexpected noises or vibration levels for the device to be effective. In the case of large or sophisticated machinery, particularly in noisy areas of a facility, this strategy may not be reliable at all. Furthermore, by the time an issue is discovered, it is possible that significant deterioration or damage has already happened.

To ensure that a machine repair was completed correctly, vibration analysis can also be used as an acceptance test. Bearings and gears can be inspected to verify if they have been properly installed or if the alignment and balancing have been conducted within acceptable tolerances (Saruhan *et al.,* 2014). Periodic machinery inspections, such as every month or quarter, may yield extra information. Based on periodic vibration monitoring and trends in vibration levels, it is feasible to anticipate the future condition of a machine's health. If a machine failure causes unplanned downtime, repairs can be scheduled to begin when the machine is next scheduled to shut down.

Analyzing vibrations can reveal ineffective methods of maintenance and repair. A few examples are rotor imbalance, faulty shaft alignment, and poor bearing installation and replacement. Because misalignment and unbalance account for about 80% of all common rotating equipment faults, vibration analysis is a crucial technique for reducing or eliminating recurrent machine issues (Alsalaet, 2012). Additionally, the specialties in vibration levels could be utilized to figure out poor production techniques, including overusing equipment (higher temperatures, speeds, or loads). Similar machines from various manufacturers can be compared using these patterns to see if any design advantages or flaws are apparent in increased or decreased performance. When used in combination with other programmers, vibration analysis can help improve equipment performance. There are numerous ways in which these improvements can be achieved. These include improved alignments and balancing, better installs and maintenance, and reduced average vibration levels in industrial equipment.

When it comes to new mechanical, process, and industrial equipment, vibration analysis has known to be a successful approach of comparing actual performance to design parameters. It is possible to employ preacceptance tests done at the manufacturer and shortly after installation to verify that functions of new equipment perform the highest possible efficiency and at the lowest possible life-cycle cost (Mobley, 2002). Design flaws, including the possibility of harm during transport or installation, may be identified, and remedied prior to long-term harm and unpredicted expenditures arise.

2.2.1 Vibration Analysis Profiles

Equipment in motion produces vibrations, which may be used to determine its present working status. Regardless of the machine's speed, whether it's rotating, reciprocating, or moving in a straight line, this remains true (Mobley, 2002). All mechanical equipment can benefit from vibration analysis, despite the popular but incorrect idea that it is only applicable to easy rotating machinery with running velocity more than 600 revolutions per minute (rpm). Vibration-profile analysis known as a valuable equipment for a variety of applications, including predictive maintenance, diagnostics, and many more.

Association of predictive maintenance with monitoring the vibration characteristics of rotating equipment in order to anticipate emerging problems and avoid catastrophic failure; however, the data needed for analyzing electrical tool, condition of lubricating oil did not offer by vibration analysis which frequently involved in the maintenance management program; therefore, predictive maintenance is not synonymous with vibration analysis. To be effective, a comprehensive plant predictive maintenance program must incorporate a variety of methodologies, each of which is designed to offer specific knowledge on the equipment.

2.2.1.1 Theoretical Vibration Profiles

New equipment and industrial systems must include vibration data into their design and development. Preliminary designs can be based on information gleaned from similar or already existing technology. Evolution of new machinery and systems prototype testing provide the refinement of basic designs and the addition of vibration data to the database of design.

The term "vibration" refers to a recurring motion or a motion that happens at a predetermined interval. Period T represents the length of time between two consecutive occurrences of a vibration. Vibrations are illustrated in Figure 2.1 as a graph or profile that indicate the period, T, and the maximum amplitude, X_0. The vibration's frequency, f, which is the inverse of period $\frac{1}{T}$ is measured in cycles per second (cps) or Hertz and is equal to the period's inverse (Hz). Figure 2.2 shows the harmonic function for the modest oscillations of a basic pendulum, which is the easiest sort of periodic motion. From its starting point, mass moves in a circular motion to the top limit of travel, back via its neutral position to the bottom limit of travel, and back to its neutral position. There is enough information in this single cycle of motion to precisely quantify the vibration of this system. The mass will simply continue to move in the same direction, resulting in the same cycle. Known as periodic and harmonic motion, this type of motion can be defined mathematically as a relationship between the displacement of the mass and time written in the form of a sinusoidal equation. The equation that expresses this relationship is (Alsalaet, 2012):

$$X = X_0 \sin(\omega \mathrm{t}) \qquad \text{(Equation 2.1)}$$

where:

X = Vibration displacement (thousandths of an inch, or mils)

X_0 = Maximum displacement or amplitude (mils)

$\omega = 2\pi f$, Circular frequency (radians per second)

t = Time (seconds)

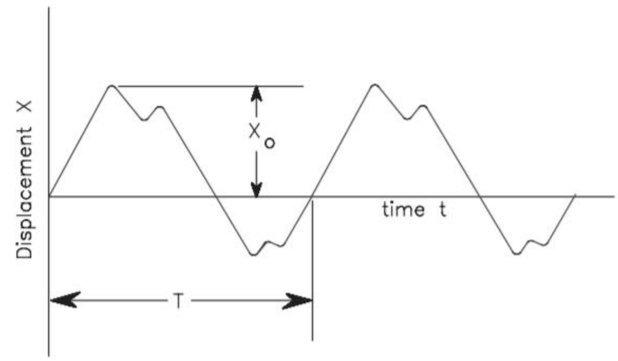


Figure 2.1: Vibration profile (Alsalaet, 2012).

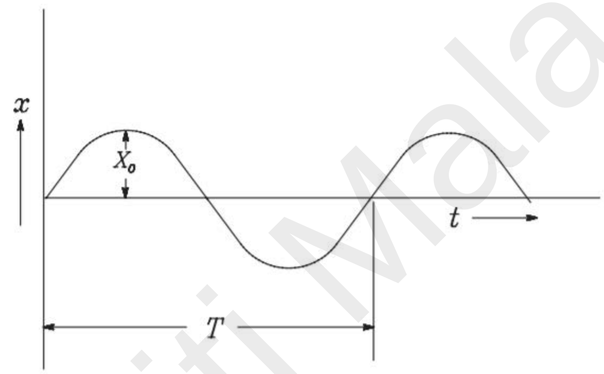


Figure 2.2: Small oscillations of a simple pendulum, harmonic function (Mobley, 2002).

2.2.1.2 Actual Vibration Profiles

It is necessary to collect and interpret complex machine data to do vibration analysis. Figures 2.1 and 2.2 represent simple theoretical vibration curves, but the profile of a single equipment is much more complicated because generally multiple sorts of vibration to account for occurs. To create a composite profile, the individual profiles from each source are multiplied together and shown. There are two ways to present these profiles which is time and frequency domains. Fault detection and monitoring often make use of time- and frequency-domain properties. (Goreczka and Strackeljan).

Two sorts of data analysis which is the time-domain and the frequency-domain modalities that can be used. They are widely employed in a variety of industries, including electronics, acoustics, telecommunications, and many more (Ambaye, 2020). Table 2.1 below have shown the frequent use function of time-domain and frequency-domain analysis.

Table 2.1: Frequent use function of time-domain and frequency-domain analysis (Ambaye, 2020).

Time-Domain Analysis	Frequency-Domain Analysis
Used in conditions where processes such as filtering, amplifying, and mixing are required.	This allows predictions and regression models for the signal, and it generates the behavior of the signal over time.
Useful in creating desired wave patterns, including binary bit ways of a computer.	Used to understand data sent in such bit patterns over time.

2.2.1.2.1 Time-Domain Analysis

A time-domain data profile is a graph that shows the amplitude of vibration over time. (Mobley, 2002). All linear and reciprocating motion machinery usually be plotted using time-domain graphs. For the general study of machine-trains, they can be quite beneficial, but time-domain data can be a problem. As a result of the combined display of all vibration data, it is difficult to discern the commitment of any one vibration source while looking at this chart.

Vibration signals are collected using computerized esteems that show closeness, speed, or acceleration in the alteration of time graph known as time space. As a component of time, it displays or investigates vibration information (Kim *et al*., 2007). According to the

Time Space Flag arrangement, almost no data are lost before the vibration signal can be reviewed. Even after standardizing, vitality disposal, and the sifting of time-space information, the vibration signals from abnormal machines and normal machines show distinct vibration signs (Goreczka and Strackeljan). The instrument's vibration or acoustic information can be dissected for fault and disappointment analysis using time-space signals. The arbitrary features of a vibrations signal generated by a physical structure are often studied using factual methodologies and highlight abstraction. It is possible to determine the crests, shapes and unflexible of a period-varying sign using the upper subordinates of the time area sign's numerical attributes (Douguer and Strackeljan, 2009).

Transducers that capture time-domain vibration data from rotating machinery are common. Using transducers that measure acceleration, velocity, or proximity, these data points are classified across time and reflect the corresponding acceleration, velocity, or closeness in relation to each other (Nandi and Ahmed, 2019). It is common for the vibration signals to involve a significant collection of feedbacks from various spinning machine components and some background noise. This led to difficulty to immediately use the obtained vibration signals for fault diagnostics of the machine, whether it is done manually or automatically. For a more basic description of vibration signals, it is usual for researchers to compute specific characteristics of the raw signal that can be used to describe the signal in general. In the area of machine learning, these attributes are referred to as characteristics, signatures, or features. Starting with raw vibration data and ending with mature findings, a variety of approaches must be employed to identify integrated faults. For example, vibration analysis techniques which can obtain relevant information from raw vibration datasets, which may be used for defect diagnosis, are included in this list. Being partial reason that dominate fault diagnosis of time-domain, visual and feature-based inspections of vibration signals can be divided as two basic forms of inspection.

2.2.1.2.1.1 Visual inspection of time-domain

Using this method, state of machine could be determined by making comparison of the observed vibration signal to a previously measured vibration signal from an unaffected or healthy machine. Both signals should be analyzed in the same frequency band in this situation. If vibration measures are higher than typical, this indicates an issue with the equipment, which causes it to vibrate more. Time domain vibration signals for fresh roller bearings may be seen in Figure 2.3a, while Figure 2.3b depicts an inner race (IR) fault state for roller bearings (Nandi and Ahmed, 2019). For example, A high measures of amplitude is clearly visible in some regions of the vibration signal on the time waveform shown in Figure 2.3b. However, the amplitude remains lower in other parts of this signal than in the usual condition shown in Figure 2.3 a. A problem exists with the machine because of this. An oscilloscope or computer-based assistance can be used to gather data and record or display information using this technique, which is easy and cost-effective.

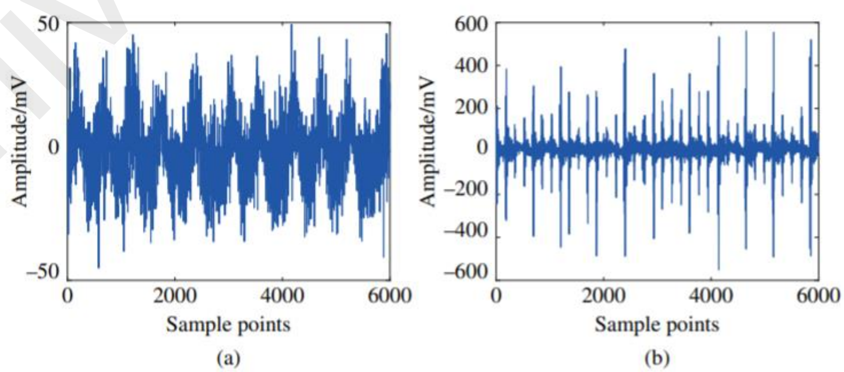


Figure 2.3: Roller bearing's time-domain vibration signal: (a) brand-new condition; (b) inner race fault condition (Nandi and Ahmed, 2019).

There are four reasons that indicated this form of examination is not dependable when it comes to keeping track of rotating equipment' health:

(i) not all rotating machine time waveform signals are visually distinct (Guo *et al.,* 2005)

(ii) There is a lot of background noise in the actual vibration signals that we deal with in practice.

(iii) Low-amplitude signals with noisy backgrounds are occasionally a problem.

(iv) Because of the need for early failure detection, manual inspection of all gathered signals is just not feasible.

Using Figure 2.4 as an example, it is difficult to determine whether a machine is malfunctioning by looking at the time waveform characteristics alone. Figure 2.4a shows two typical vibration signals from roller bearings that are both worn but undamaged and Figure 2.4b shows an outer race (OR) fault condition.

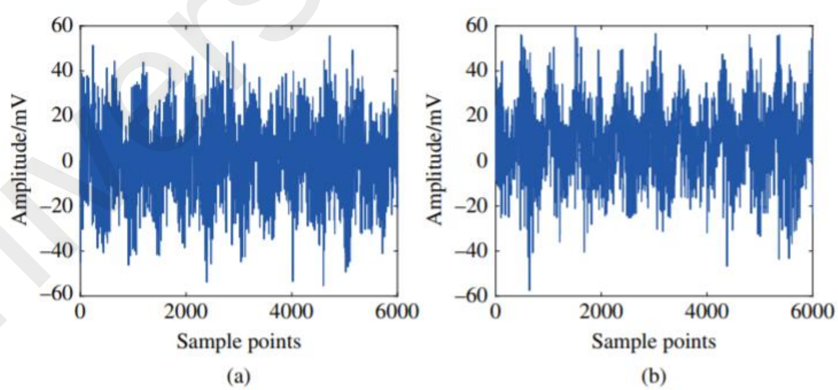


Figure 2.4: Roller bearing's time-domain vibration signal: (a) worn but undamaged condition; (b) outer race fault condition (Nandi and Ahmed, 2019).

2.2.1.2.1.2 Feature-Based Inspection

The raw vibration signal may be used to characterize the signal in its most basic form, and some characteristics of it can be computed to estimate the overall health of a machine using this sort of examination. It is possible to use these properties to discriminate between two comparable vibration signals. Machine learning classifiers may then be used to determine the proper condition type from the signal, which can be the raw vibration signal or the calculated features of the raw vibration signal in the time domain, depending on the application.

When working in the time-domain field, vibration signal processing can be accomplished using features extraction, which can be accomplished using statistical functions and other approaches that can offers extraction of features from time-indexed raw vibration datasets that accurately show machine health. Acquired vibration signals are often gathered from a variety of sources in a spinning machine that exhibits unpredictable behavior. A direct mathematical formula cannot be used to characterize these vibration signals because of their unpredictability features. Rather, statistical approaches must be used to analyses them in the context of time. As a result, it should come as no surprise that past research in this field has concentrated on time domain descriptive statistics-based features that may be utilized for both manual inspection and automatic monitoring, as well as for automatic monitoring. For extraction of features from vibration signals in the time domain based on signal amplitude, many various sorts of statistical functions have been employed extensively. The frequently used statistical functions covered in feature-based inspections are discussed in further depth as below (Kumar and Manjunath, 2017) (Nandi and Ahmed, 2019).

1. Peak Amplitude

   This is also known as the peak amplitude x_p.The peak amplitude is the maximum positive amplitude of a vibration signal; it can also be defined as half of the difference between the maximum and minimum vibration amplitudes, that is, the maximum positive peak amplitude and the maximum negative peak amplitude, respectively. This can be expressed mathematically in the following way:

   $$x_p = \frac{1}{2}[x_{\max}(\mathrm{t})\text{-}x_{min}(\mathrm{t})] \qquad \text{(Equation 2.2)}$$

2. Mean Amplitude

   The average of the vibration signal across a sampled interval is the mean amplitude, x̅ which may be calculated using Equation 2.3,

   $$\bar{x}=\frac{1}{T}\int x(t)dt \qquad \text{(Equation 2.3)}$$

   and T is the sampled signal duration and x(t) is the vibration signal. For a discrete sampled signal, can be rewritten as:

   $$\bar{x} = \frac{1}{N}\sum_{i=1}^{n} x_i \qquad \text{(Equation 2.4)}$$

   where N represent number of sampled points and x_i is an element of signal x.

3. Root Mean Square Amplitude

   The variation of the vibration signal magnitude is represented by the root mean square (RMS) amplitude, denoted by the symbol x_{RMS} According to Equation 2.5, the mathematical equation for x_{RMS} is

   $$x_{RMS} = \sqrt{\frac{1}{T}} \; |x(t)|^2 dt \qquad \text{(Equation 2.5)}$$

   T is the duration of the sampled signal, and x(t) is the vibration signal. In the steady-state operating condition, the RMS amplitude is resistant to false peaks. Equation 2.6 can be simplified as follows if the vibration signal is discrete:

$$x_{RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}|x_i|^2} \qquad \text{(Equation 2.6)}$$

4. Peak-to-Peak Amplitude

The peak-to-peak amplitude, also known as the range, x_{p-p} is the difference between the maximum positive peak amplitude and the maximum negative peak amplitude, $x_{max}(t) - x_{\min}(t)$.

5. Crest Factor (CF)

The crest factor (CF), x_{CF}, is defined as the ratio of the vibration signal's peak amplitude, x_p, to its RMS amplitude, x_{RMS}. This can be calculated using the following formula:

$$x_{CF} = \frac{x_p}{x_{RMS}} \qquad \text{(Equation 2.7)}$$

6. Skewness

A measure of the asymmetrical behavior of a vibration signal through its probability density function (PDF) is known as skewness, which is also known as the third normalized central statistical moment, x_{SK}. It determines whether the vibration signal is skewed to the left or right side of the normal state of the vibration signal distribution, respectively. In the case of a signal with N sample points, the value of x_{SK} can be represented as Eq (2.8),

$$x_{SK} = \frac{\sum_{i=1}^{N}(x_i - \bar{x})^3}{N{\sigma_x}^3} \qquad \text{(Equation 2.8)}$$

The value of x_{SK} for a normal condition is zero.

7. Kurtosis

It is known as the fourth normalized central statistical moment, x_{KURT}, and it is measuring the input vibration signal's peak value. It is calculated using its PDF. It determines if the peak of the distribution is higher or lower than the peak of the distribution corresponding to a normal state of the vibration signal, in other words, whether the peak is higher or lower. In the case of a signal with N sample points, x_{KURT} can be expressed in the form shown in Eq (2.9),

$$x_{KURT} = \sum_{i=1}^{N} \frac{(x_i - \bar{x})^4}{N{\sigma_x}^4} \qquad \text{(Equation 2.9)}$$

8. Impulse Factor

Vibration signal impulse factor x_{IF} is known as the ratio of the peak value to the average signal average, it can be computed as Eq (2.10),

$$x_{IF} = \frac{x_{peak}}{\frac{1}{N}\sum_{i=1}^{N} |x_i|} \qquad \text{(Equation 2.10)}$$

9. Shape Factor

The RMS value to the average of the absolute value of the vibration signal's form factor, x_{SF} can be stated as follows:

$$x_{SF} = \frac{x_{RMS}}{\frac{1}{N}\sum_{i=1}^{N} |x_i|} \qquad \text{(Equation 2.11)}$$

10. Clearance Factor

According to the clearance factor x_{CLF},the ratio of the maximum value of the input vibration signal to its mean square root (the absolute value of the input vibration signal) can be stated using Eq 2.12 as follows:

$$x_{CLF} = \frac{x_{max}}{(\frac{1}{N}\sum_{i=1}^{N} \sqrt{|x_i|})^2} \qquad \text{(Equation 2.12)}$$

2.2.1.2.2 Frequency-Domain Analysis

Machine condition monitoring techniques including such frequency analysis, also known as spectrum analysis, are one of the frequent used techniques for observing the state of machinery. Communications, geology, remote sensing, and image processing are just a few of the fields in which frequency-domain analysis is frequently utilized today (Ramirez *et al*., 2017). As a matter of fact, frequency-domain analytic methods may disclose information that is difficult to discern in the time domain by analyzing frequency features. For example, bearings, shafts, and fans all contributed to the recorded time-domain vibration signals by creating a single sine wave with a single frequency and amplitude, which is then multiplied by subsequent motions of other components. For better or worse, each component of a spinning machine generates only one type of frequency. Despite this, the sensor's distinct frequencies are seldom visible in the measured signal; instead, it sees an addition of the signals that were measured by the sensor. The frequency components created from the time domain waveforms are represented as a spectrum, which makes it easier to distinguish between the many sources of vibration. Based on figure 2.5, the frequency components spectrum derived according to time domain waveforms makes it simpler to identify the vibration's source (Nandi and Ahmed, 2019). The circular frequencies of the moving parts are related to the simple harmonic vibration functions, which are useful from a practical perspective. As a result, the frequencies are a amplification of the machine-basic train's operating speed, in terms of revolutions per minute (rpm) or cycles per minute (cpm), either (cpm). To determine these frequencies is the first and most fundamental procedure in determining the functioning state of the machine-train system (Mobley, 2002).

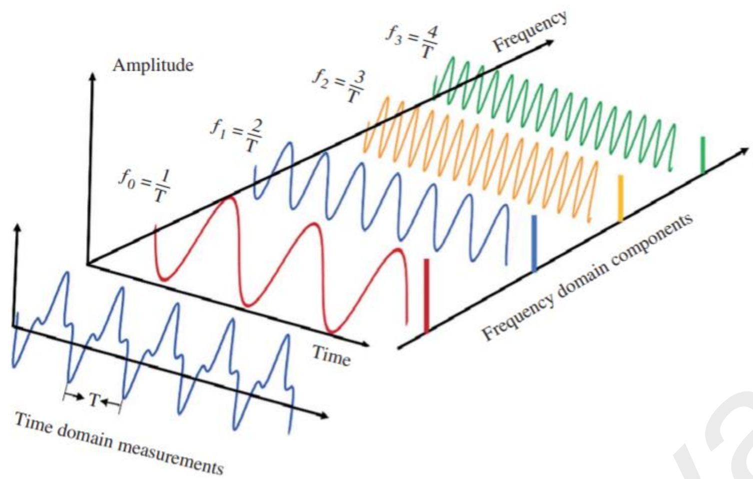


Figure 2.5: Frequency-domain components versus Time-domain measurements (Brandt, 2011).

Frequency-domain analysis is a critical instrument that cannot be overlooked in signal processing applications. Frequency-domain analysis, as opposed to time-domain analysis, indicates the way of energy of a signal is divided across a large frequency range. As part of the frequency representation, each frequency component must be shifted in order to recover the original time signal from a composite of all of its individual components in the frequency domain. A signal can be translated between the time and frequency domains using a transform, which is a pair of mathematical operations that converts between the two domains. Figure 2.6 have shown example of frequency-domain representations. The conversion of time-domain data into frequency-domain data is accomplished using a mathematical technique known as the Fast Fourier Transform (FFT) (Nandi and Ahmed, 2019). When a complicated machine-train spectrum is analyzed using FFT, each vibration component can be represented as a discrete frequency peak. Displacement per unit time associated with a specific frequency is the frequency-domain amplitude, which is displayed as the Y-axis against the frequency as the X-axis in the frequency-domain plotting. The opposite of this is true for time-domain spectrums, which total the velocities of all frequencies and plot that sum as the Y-axis versus time as the X-axis to represent the frequency spectrum.

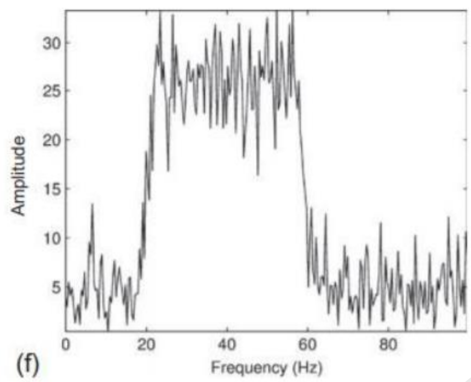


Figure 2.6: An example of frequency-domain representations (Ambaye, 2020)

By comparing the obtained data with standard or ordinary state frequency range information and comparing it to the obtained information in frequency space attributes of vibration analysis, one of the most significant advantages includes unquestionably summarizing initial data of fault proliferation and various disputes. This sort of life prediction research is very successful as a pattern assessment. With regard to the presentation or analysis of vibration information, "frequency space" refers to the frequency at which it occurs. The rapid FFT calculation, is typically used to translate the time space vibration flag into the frequency area. FFT computes the direct Fourier transform (DFT) and its inverse for a stationary time series signal with a significant reduction in complexity. The DFT of a signal with length N is computed using $N\ log_2 N$ complex multiplications rather than N^2complex multiplications when utilizing FFT (Nandi and Ahmed, 2019). The technique's most important advantage is that the gloomy concept of the vibration signal is plainly dislodged as tops in the frequency range at the frequency where the reiteration occurs. When investigating vibrational information in the frequency domain, it is beneficial to make use of the power range as a starting point. Suppose that a discrete time signal x(t) corresponds to an investigated occasional capacity with period T. The Fourier arrangement development of x(t) can be obtained from the Fourier integral if it is positive (Kumar and Manjunath, 2017).

$$X(f) = \frac{1}{T}\int_{\frac{-T}{2}}^{\frac{T}{2}} X(T)e^{-|\bar{s}|ft}dt \qquad \text{(Equation 2.13)}$$

In this case, f refers to distinct frequencies with comparable dispersion that are products of the complements of the period T. It is possible to characterize the power range P(f) as the greatness or power obtained from the Fourier integral, which can be expressed as follows:

$$p(f) = E[X(f)X * (F)] \qquad \text{(Equation 2.14)}$$

Where * refers to complex conjugation and E [] refers to the normal value, respectively. For vibration information, the FFT provides fast and benefits techniques to calculate the DFT, and a window capacity can be used to confine the vibration information to seem occasional, so reducing leakage starting with one recurrence segment and progressing to the next. When working with stationary arbitrary information, a smooth gauge of energy range can be obtained by averaging the range acquired by condition over several windowed information records. This is referred to as the Welch approach for control range estimation. One of the most important considerations in processing power spectra of vibrating information is to identify significant frequency parts that may be established in the range and then utilize these segments and their sufficiency for the slanting goals. The spectra analysis would provide frequency space amplitude in the context of removal, speed, increasing speed, or stage, among other factors. Auto-range or power range estimation is like frequency estimation in that the amplitudes are conveyed as the square of their respective sizes.

There are also some other frequency-domain vibration analyses that used in different research. One of the methods of using frequency-domain analysis is known as spectrogram.

To depict the way of spectrum (frequency-domain) differs with time, a spectrogram is created by overlapping a sequence of FFT. Spectrograms can be quite useful for performing vibration analysis in a dynamic setting since they can show exactly how the spectrum of the vibration varies over time (Yan and Gao, 2009). Another approach of utilizing frequency-domain is the Power Spectral Density (PSD). For vibration signals with a finite number of dominant frequency components, fast Fourier transforms (FFTs) are excellent tools; nevertheless, PSD are employed to characterize random vibration signals. A PSD could be calculated by multiplying the frequency bin in an FFT by its complex conjugate, which produce the real only spectrum of amplitude in g^2. The fundamental feature of a PSD that makes it more benefits for typical vibration analysis than an FFT is that the amplitude value is normalized to the frequency bin width, resulting in units of g^2/Hz being obtained (Ahmadi and Salami, 2010).

2.2.2 Vibration Measurement Parameters and Vibration Severity Criteria

There are a few of measurement parameters of vibration including acceleration, velocity, and displacement (Alsalaet, 2012). Referring to Figure 2.1 and 2.2, the velocity at the zero position, the oscillating mass's velocity reaches its maximum value, while at the lowest and highest points, it reaches its minimum value (zero). Velocity is the most important parameters that related to the destructive force of vibration. Typically, the RMS value of velocity which measured from 10 to 10000 Hz provide the top sign of vibration severity. As it turns out, velocity is a function of displacement, which is:

$$v = \frac{dx}{dt} = X_0\omega \ \cos \omega t \qquad \text{(Equation 2.15)}$$

For the acceleration, it is different from velocity. It goes from maximum and minimum values at the highest and lowest points to the zero point. High frequencies is important to acceleration. The acceleration signal can be converted to velocity or displacement. Derivative of velocity is acceleration and can be represented by:

$$a = \frac{dv}{dt} = -X_0\omega^2 \sin \omega t \qquad \text{(Equation 2.16)}$$

When the frequency ω is high, even the displacement is small, both velocity and acceleration would be high. The inverse is true when at low frequency. There are two factors that would affect the decision on choosing which of three parameters, acceleration, velocity, and displacement to measure the vibration. The first factor is to decide based on the purpose of taking the measurement depending on vibration analysis, periodic check or balancing. Another factor is the rely on the processing speed and sorts of machine element such as anti-friction bearing, gear etc. The amount of deformation (displacement) that a machine part undergoes, and the frequency of deformation, determine the amount of time it takes for a machine part to break. To summaries, vibration severity is determined by displacement and frequencies. Vibration severity may be directly measured by velocity since it is a function of these two quantities. Vibration severity can also be assessed using displacement and acceleration, although this requires knowledge of the vibration frequency. Although just a few instances, displacement may serve as a benchmark of vibration severity under conditions of dynamic stress, where properties like brittleness tend to accelerate failure, or when stress (deformation) approaches an established limit. As a result, low-frequency applications benefit from displacement measurement. Attenuation, rather than velocity, is typically considered to be the best indicator of how severe the vibrations are at higher frequencies, such as 1000 Hz (60 kcpm). Velocity is often employed for vibration monitoring and analysis because most typical rotating machinery (and their flaws) operate in the 10–1000 Hz range. Dynamic forces have a big role in acceleration, and even while the displacement and velocity may be tiny, they can nonetheless produce large forces at higher frequencies. For high-frequency vibration (above 1000 Hz), acceleration measurement is a good predictor of severity (Alsalaet, 2012).

2.2.3 Vibration Analysis and Measuring Equipment

It is possible to detect vibrations in a system using various mechanical or optical approaches that are referred to as vibration sensors. There are numerous sorts of sensors that may be used to measure vibrations. Although there are no direct vibration sensors, vibrations can be monitored indirectly by extracting values from classical mechanical or optical quantities. A few characteristics distinguish these sensors from one another. Among other things, they can be separated into two categories: those that measure relative behavior and those that measure absolute behavior. Frequency range, signal dynamics, and measurement data quality are all further distinguishing characteristics. The sensors listed below were originally organized into two groups: contacting and non-contacting, and within these groups, into sub-items such as displacement, measurement, velocity measurement, and acceleration measurement, among other things (Guo, 2014).

2.2.3.1 Contacting and Non-contacting Vibration Measurement Method

A. Contacting vibration Method

The sensors will be attached to the surface of the mechanical structure to measure its vibration. Based on research, some examples of contacting vibration equipment through different measurements have shown as below.

I. Displacement Measurement

- Potentiometric Transmitter
- Linear Variable Differential Transformer (LVDT)

II. Speed Measurement

- Seismometer

III. Acceleration Measurement

- Piezoelectric sensor
- Piezo-resistive sensor
- Resistive sensor
- Inductive sensor

B. Non-contacting Vibration Method

Non-contact measurement methods measure the vibration without contacting the surface of the mechanical machine structure. Normally, non-contact method needs visual access to the surface of detail.

i. Path Measurement

- Capacitive Principle
- Eddy current sensor
- Hall sensor
- Optical sensor

ii. Speed Measurement

- Laser-Doppler vibrometer (LDV)

Depending on the sensor type, each form of processing principle and analyzing stage would be distinct from the others. There are two types of sensors: active and passive. Some of these sensors detect relative behaviors, while others monitor absolute behaviors. In the following table, the characteristics of both the contacting and non-contacting vibration measurement methods that covered in the previous section have been summarized.

Table 2.2: Summarization of contacting and non-contacting vibration measurement method equipment

Sensors	**Contact/Non-contact**	**Active/Passive**	**Absolute/relative**	**Displacement/Speed/ Acceleration**
Potentiometric transmitter	contact	passive	relative	displacement
LVDT	contact	passive	relative	displacement
Principle of electrodynamics	contact	active	relative	speed
Seismometer	contact	active/passive	absolute	speed
Piezoelectric sensor	contact	active	absolute	acceleration
Piezo-resistive sensor	contact	active	absolute	acceleration
Resistive sensor	contact	active	absolute	acceleration
Inductive sensor	contact	active	absolute	acceleration

Capacitive principle	non-contact	passive	relative	displacement
Eddy current sensor	non-contact	passive	relative	displacement
Hall sensor	non-contact	active	relative	Displacement
Optical sensor	non-contact	passive	relative	Displacement
LDV	non-contact	passive	relative	speed

2.2.3.2 Existing vibration measurement and limitations

To date, several different methods for measuring vibration in rotating machinery have been devised and implemented. For many years, industries have used vibration sensors in rotating equipment to detect abnormal vibrations. The accelerometer sensor is the most widely used and is regarded to be the simplest of the sensors (Jamal and Rasheed, 2021). By using a piezoelectric transducer, accelerometers can measure acceleration forces. To get vibration measurements, the accelerometer is the transducer that is most usually utilized. To transform mechanical energy into electrical impulses, it employs piezoelectric film technology. In most cases, the device consists of a weight that is suspended between two piezoelectric film layers. Every time the weight squeezes the piezoelectric films, an electrical signal is generated as a result of the vibration (Mobley, 2002). Simple, cheap, precise, and responsive represent accelerometers when used at high frequencies. In most cases, it is mounted to the machine's casing and the vibration transducer is of the contact type. Accelerometers, on the other hand, do not measure the vibration that is immediately acting on the spinning shaft and are best suited for describing bearing fatigue and casing resonance.

It is somehow constraint on the different of measurements that they could go through (Jamal and Rashed, 2021).

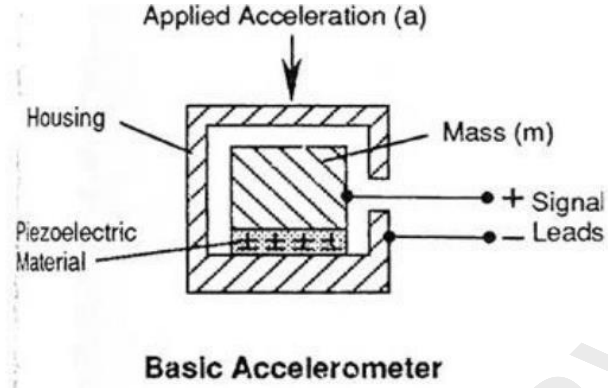


Figure 2.7: Basic Accelerometer

(http://www.intertechnology.com/Kistler/indexAcceleration.htm., 2015)

Figure 2.8: Photo of Accelerometer Sensor

(http://www.industrialelectronics.com/DAQ/industrial_electronics/input_devices_sensors_transducers_transmitters_measurement/Accelerometers.html., 2015)

Another well-known type of sensor for measuring vibration in rotating machinery is the non-contacting type one that is based on the displacement technique. When a rotating element is subjected to a magnetic field, Eddy current sensors being employed to estimate the displacement of the element. The internal electronic circuit generates an alternating magnetic field that is directed against the object that is being sensed can be demonstrate in figure 2.9. The induced Eddy current on the sensing item opposes the field produced by the probe, resulting in the generation of an output voltage on the probe. The voltage produced by the signal is proportional to the amount of movement or length between the probe tip as well as the surface being measured. This method is used to detect rotors that are misaligned or imbalanced (Jamia *et al.,* 2018). The probes are differentiated as non-contacting transducers since they do neglect any physical contact with the rotating element. They are modest in size, relatively inexpensive, and require little maintenance. Despite this, installation is hard, operate at a high frequency, and are extremely sensitive to mechanical and electrical disturbance (Jamal and Rashed, 2021).

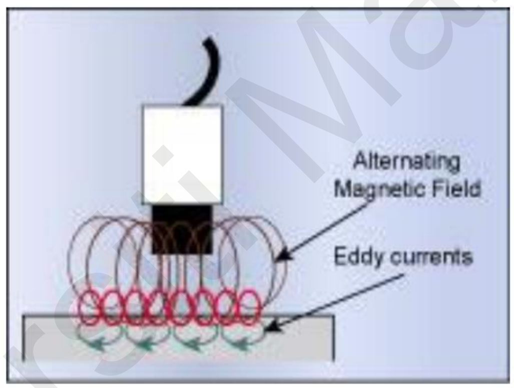


Figure 2.9: Magnetic field of Proximity Sensor (http://www.lionprecision.com/eddy-current -sensors/, 2015)

Figure 2.10: Photo of different sizes of proximity sensors (http://www.globalspec.com/learnmore/sensors_transducers_detectors/proximity_presence_sensing/eddy_current_proximity_sensors, 2015)

Based on Laser-Doppler (LD) technology, metrology-inspired vibration sensors have been created in recent years (Schwenke, 2002). In some ways, the Laser Doppler Vibrometer (LDV) and Laser Doppler Velocimeter (LDV) stands out from conventional accelerometers. One of the non-invasive sorts of measurement is LDV because they do not require calibration and can measure across the entire frequency range (Esposito, 2008). As shown in Figure 2.11, LDV employ a laser beam deflection method to determine the angular velocity and displacement of a rotating element. The transfer of laser power to the sensor's head is accomplished using a fiber optic. Furthermore, the bounced laser signal is converted to an electrical signal using a digital decoding technique. LDVs are commonly used in rotating machines to measure Torsional Vibration (Xiang *et al.,* 2012). Although they are extremely expensive, they still have some limitations such as if the laser spot scans the structure too quickly, there is an additional laser speckle noise (Johansmann *et al*., 2005).

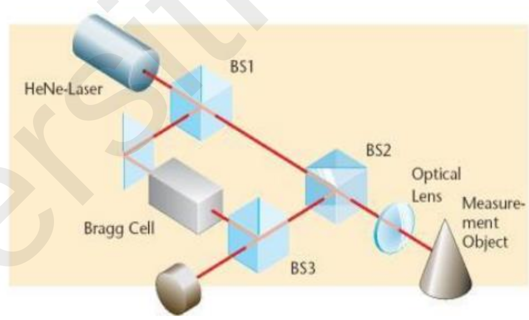


Figure 2.11: Beam Deflection Method for LDV Sensor

(http://www-cs.ccny.cuny.edu/~zhu/LDV/FinalReportsHTML/CCNY-LDV-Tech-Report-html.htm., 2015)

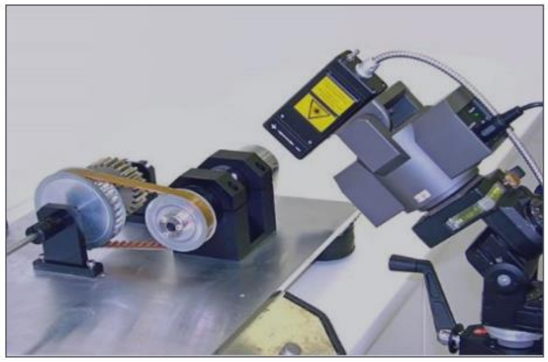


Figure 2.12: Photo of LDV Sensor

(http://acoutronic.se/images/vibration/Polytec_picture_RSV-150.png., 2015)

Almost every industry seeks sensors with a combination of high accuracy, low cost, ease of installation, and rigidity. Accelerometers and proximity probes as the vibration sensorsfall into these categories, but they are still having their own set of restrictions and drawbacks. While LDVs are a cutting-edge laser beam technology, they are at the beginning stages of development and are prohibitively expensive. Experimental data collection is inevitably influenced by the characteristics and limitations of the currently available vibration analysis technologies. Contact point sensors such as velocity, acceleration or strain sensors are used to transfer data to the data acquisition hardware via wired connections in civil engineering. When it comes to operational modal analysis (OMA) data acquisition, accelerometers with high sensitivity and protected connections are widely used. Wireless sensors, can be used to avoid the use of connection cables, given the difficulty and effort involved in cabling large structures (Li *et al.,* 2016). To fix the pointwise nature of conventional approaches, optical fiber sensors can be used to provide distributed monitoring (Bastianini, 2007). Engineers have recently been particularly interested in non-contact monitoring technologies (Narasimhan and Wang, 2020).In civil engineering, various approaches to vibration testing have been investigated, such as laser LDV(Rothberg *et al.*,2017), microwave radar interferometry (Gentile, 2010), infrasound (Lobo-Aguilar *et al.,* 2019), global positioning system (GPS) sensing (Moschas and Stiros, 2011), satellite remote sensing (Lazecky *et al*., 2016), optical methods based on the moiré effect (Ri *et al*.,2012) and optical vision based methods using digital imaging (Chen and Chang, 2019). The idea of eliminating physical installations of sensors for vibration analysis is particularly enticing of structures that may not be conveniently or safely accessible but require rapid assessment of their status, such as after major occurrences like large earthquakes, explosions or floods (Zona *et al*., 2020).

Traditional non-contact methods, on the other hand, can only be used to measure one point at a time. In addition, the use of conventional measuring sensors has additional downsides, such as harsh weather circumstances such as high temperatures and high humidity, difficulty in mounting the sensor on an item, and the influence of sensor weight on object dynamic characteristics (Cakar and Sanliturk, 2005). Using any of these methods is more expensive and time consuming than video recording, and it may even be prohibitively expensive in some cases even if video recording is still an option (Schwarz *et al*., 2020). The vision-based system has several advantages over conventional sensors, including remote monitoring, low cost, and the ability to measure multiple points at once (Zhu *et al*., 2020). Table 2.2 presents an illustrative comparison of the constraints imposed by a classical sensor (accelerometer) and a visual-based sensor (camera) in terms of performance (Śmieja *et al.,* 2021). Even with low-cost consumer-grade instruments, several applications involving items ranging from the smallest dimensions to huge buildings have shown remarkable promise (Ye *et al.,* 2016) (Spencer *et al*., 2019) (Dong and Catbas, 2020), so the focus on this research project is on vision-based methods among contactless technologies.

Table 2.3: Comparison of base constrains of accelerometer and visual motion sensor (Śmieja *et al*., 2021).

Accelerometer	Video Camera
Contact	Contactless
Sparse discrete pointwise measurements	Location of many points at the same time (Quasi continuous)
Impact on the immediate neighborhood in the assembly point (sensitivity to temperature, chemicals, etc.)	sensitivity to disturbances in the line of sight (Sensitivity to lighting, fog, smoke, etc.)
Measurement of absolute values	Measurement of relative values (Relative to camera base)
Direct acquisition	In plane 3D to 2D projection (In case of single camera)

2.3 Visual Based Measurement System

When it comes to structural vibration analysis, identifying vibration parameters is critical. Vibration characteristics can be identified using one or more sensors that are connected in series, which are commonly used in traditional approaches. Those methods necessitate the use of complex detection systems and have the potential to alter the structure's inherent dynamic characteristics to a certain extent (Li *et al.,* 2016). Visual vibration measurement is becoming increasingly popular as a non-contact, wide-range vibration measurement method (Yu *et al.,* 2017). The previously accessible capabilities for vibration observation have been greatly expanded by modern picture analysis and processing technologies. In this paper, a triangulation technique and point-tracking strategy for observing static and dynamic stresses in a civil structure are discussed in detail (Yang, 2019).

Structured light, single camera, and multi-view stereo vision are just a few of the visual assessment technologies that have undergone substantial study and use (Wang *et al.,* 2012). In a cantilever beam fitted with micro electromechanical systems (MEMS) and high frequency vibration, researchers were able to precisely quantify the in-plane motion displacement using a single high-speed CCD camera and microscope imaging system (Teyssieux, 2011). Determination of the two-dimensional motion parameters of a special mark can use a line-scan camera (Lim and Lim, 2008). About watching the movement of pile rebound and penetration, an observation system using a high-speed line-scan camera has been suggested. Measurement against composite board in full field based and analytical on the features of image using adaptive moment descriptors for getting structural modal have been performed. (Wang and Mottershead, 2013). Geometric moment has been used to effectively construct an image motion blur information extraction system and an algorithm for measuring harmonic vibration based on dynamically blurred picture sequences (Guan, 2005). By applying 3D digital image method for predicting in a full-field vibration analysis, the geometry and deformation of a mechanical shaker can be detected. When looking at small motions in video, Davies and others research developed a method for inferring material properties of an object (Davies, 2015). Motion magnification to extract displacements from high-speed video have been implemented, and they demonstrated that the algorithm was capable of qualitatively identifying the operational deflection shapes of simple structures by using motion magnification (Chen *et al*., 2015). These approaches need a complicated algorithm, costly imaging equipment, and perfect tracking of the target feature in order to operate. Consequently, camera and image feature extraction performance have a significant impact on the visual measurement result.

2.3.1 Visual Odometry System

Whenever it comes to motion compensation, visual odometry (VO) refers to a set of methods are using visual cues to estimate motion. VO utilization is by determining the pose of an agent (such as a vehicle, a person, or a robot) that relies solely on a stream of images acquired by a single or multiple cameras attached to the agent (Scaramuzza and Fraundorfer 2011). VO is a combination of a certain camera configuration, the programming model, and the hardware device that provides the camera's posture at each time instant. VO algorithm can be classified into two categories: stereo and monocular configurations. Several images are captured simultaneously from different perspectives in stereo configurations, which use a multi-camera array (or a moving camera) to take the different images at the same time from different perspectives (Howard, 2008). Detected features can be projected into 3D space and tracked over time to estimate vehicle motion if the baseline is known. In comparison to monocular VO, stereo VO mimics the human vision system and can instantly predict the image scale. Stereo camera systems, on the other hand, require additional calibration effort and strict camera synchronization, without which the error propagates over time (Poddar e*t al.*, 2018). Single-camera configurations, which are essentially bearing-only sensors, are used in monocular configurations. Structure-from-motion can be used to estimate the baseline (which in this case is comparable to camera translation and rotation) if a series of photos taken at various positions is provided. (Tomasi and Zhang, 1995). In monocular vision odometry, it is well-known that absolute scale cannot be recovered, a problem that has been addressed in (Scaramuzza *et al.,* 2009) of unique case of nonholonomic constraints. An affordable and small-form-factor monocular camera is useful when installing two cameras with a defined standard is really not practical, such as a phone or laptop (Poddar *et al*., 2018). Table 2.3 has shown the comparison between the pros and cons of stereo and monocular configurations.

Table 2.4: Comparison of Monocular and Stereo VO configurations (Aqel *et al*., 2016)

Types of Configurations	Pros	Cons
Monocular	• Cheap and simple to set up • Light weight: best for tiny robotics • Easy calibration	• Suffer from image scale uncertainty
Stereo	• It is simple to get information about the size and depth of an image. • Provide 3D vision	• More expensive and needs more calibration effort than monocular cameras • When the stereo baseline is substantially less than the distances between the camera and the scene, it degrades to monocular viewing.

		• Problematic interface and synchronization problems.

A feature-based, appearance-based, or a hybrid approach to VO's geometric design has all been distinguished. (Poddar *et al*., 2018) (Aqel *et al*., 2016).

a. Feature-based approach

A mathematical technique is used to extract visual features like corners, lines, and curves from a series of image frames. It is necessary to match or track the distinguishing characteristics within the features extraction before predicting the motion. This method compares each feature in two photos and calculates the Euclidean distance between feature vectors to find the candidate matching features in the two images, which is then used to find a match. The displacement is calculated by calculating the velocity vector between the pairs of points that have been identified.

b. Appearance-based approach

It is based on optimizing the photometric error rather than on sparse features and estimates motion by optimizing the photometric error. Instead of extracting and tracking features, this technique monitors changes in the appearance of acquired images as well as the intensity of pixel information contained therein. It focuses on the information that can be extracted from pixel intensity measurements. Optical flow (OF) as the most famous appearance-based can be used to estimate the motion of the camera and the speed of the vehicle. The displacement of brightness patterns from one image frame to another is computed by the OF algorithm, which uses the intensity values of neighboring pixels to compute the displacement. The template matching method is one of the most used methods in the appearance-based approach.

c. Hybrid of feature-based and appearance-based approach

In some cases, a hybrid approach, is a addition of feature- and appearance-based approaches, is the best solution, and this is the case in some cases. Using the image's pixel intensity information as well as tracking key elements over frames is a part of the process. Direct (feature-based) and indirect distance and angle calculations can be used in hybrid algorithms for VO (appearance-based). When using feature-based schemes, reliable data is obtained at the expense of a certain loss of available information, whereas appearance-based schemes produce dense reconstructions by utilizing all available data but with errors associated with only a few areas.

Table 2.5: Comparison of feature-based, appearance-based, and hybrid-based VO. (Poddar *et al.*, 2018) (Aqel *et al.,* 2016) (Guizilini and Ramos, 2011)

Feature-based	**Appearance-based**	**Hybrid of feature and appearance-based**
Suitable for a wide range of locations, including harsh and urban settings.	In low-textured environments, appearance-based tracking is reliable and more accurate than feature tracking.	Hybrid algorithms employ both benefits of direct (feature-based) and indirect approaches (appearance-based).

Failed to deal with environments with no or minimal texture.	Since this method allows for the use of a big template in the matching process, the chance of a successful match between two consecutive image frames is high when using this method.	The primary goal is to incorporate all available data and improve motion prediction.
In these low-textured situations, the feature-based technique is deemed ineffective because of the restricted number of prominent features that can be detected and tracked.	It is possible to estimate motion by working with intensity values directly and matching templates of sub-images over two frames or the optical flow values.	A hybrid registration technique that combines feature tracking data and optical flow limitations into a single framework.

After making comparison among the different type of VO approach, it can be found that appearance-based is the best choice to implement in this research report. Instead of relying on sparse features to estimate motion, appearance-based visual odometry predicts motion by computing photometric error. Feature-based approaches are deemed noisy when applied to smoothly altering landscapes such as a foggy environment or a sandy terrain, and the features do not need to be recognizable from their surroundings to be effective. Since appearance-based techniques use information from the entire image, they can provide a reliable estimate of ego-motion even in low-textured environments. It is faster to integrate entire photos rather than just a few locations since it lowers aliasing issues associated with similar-looking locales. It is also less expensive and works with smooth shifting landscapes.

2.3.1.1 Appearance-based Technique

Regional-based matching and optical flow-based matching are the two types of appearance-based techniques that are most used. Depending on the situation, region-based matching can be achieved through correlation (template matching) or with the assistance of global appearance-based descriptors and picture alignment techniques (Poddar *et al*., 2018). For aligning images, correlation-based techniques (also known as template matching) have been extensively researched in the past, with global invariant image representations or similarity measures being used as the basis for the research. Some of these schemes' drawbacks were addressed by employing a locally invariant similarity measure and global constraints which allowed them to be more effective. Irani and Anandan proposed an image alignment technique that can estimate a parametric 2D motion model for images acquired by sensors of various modalities (Irani and Anandan, 1998). It is built on the quadrifocal relationship between picture intensities as its foundation and is hence impervious to background clutter, inter-frame displacements, and lighting alterations, among other things. (Comport *et al.,* 2007). In a later study, an approach that reduces the intensity error across the entire image while also overcoming the inter-frame overlap problem associated with region-based approaches have been developed. To use one of these region-based schemes, a detailed interest area in the image must be defined in the first image, and the two images must be sufficiently overlapped for the region-based scheme to work. Furthermore, the image registration process necessitates the use of an optimization technique to minimize an objective function, which is typically subject to issues such as local minima and divergence which are common in optimization techniques (Comport *et al.*, 2010). Optical flow-based

VO schemes can avoid typical issues include the use of an inappropriate argument minimization criterion and the presence of independently moving objects (Adiv *et al.,* 1985).

Optical Flow (OF) is a fundamental concept in object movement that critical for predicting the motion of objects based on a sequence of images. OF is a term that refers to the velocity field of an image that is produced because of the transformation of one image into the next image in orderly (Horn and Schuck, 1981). OF method accurately depicts the movement of objects in three-dimensional space and has found widespread application in the field of vision and image processing for object segmentation, recognition, tracking, and robot navigation, among other applications (Yu *et al*., 2017). New optical flow algorithms that optimized acceleration by finding for an expected position based on the assumption of fixed velocity have been developed (Denman, 2007). There is a new optical flow algorithm for the detection and tracking of moving objects (Ya *et al*., 2013). Using optical flow, the researchers have developed a method for detecting head movements (Hui *et al*., 2014). The optical flow field adjacent to the robot body was also discovered, paving the way for the construction of autonomous robots that can avoid obstacles on their own accord. (Souhila and Karim, 2007). To aid in robot navigation, this research looked at the effects of various OF algorithms in conjunction with spatiotemporal filters (McCarthy and Barnes, 2014). In the above-mentioned method, expensive equipment or complicated operations or feature tracking are still required. However. Another proposed of using OF method for measuring vibrations in out-of-plane vision is of great importance as well (Yu *et al*., 2017). Based on the research, it has shown that when using appearance-based technique, it is more common that various vibration analysis can be done based on OF method and this research would be using OF method.

2.3.1.1.1 Optical Flow Method

Using the OF method, which is a technique that employs temporal variation and correlation of pixel intensity data from image sequences to detect the movement of pixel positions and to produce the three-dimensional motion field, the movement of pixel positions can be determined (Song *et al.,* 2011). In general, it is based on three basic assumptions (Beauchemin and Barron, 1995):

1)Between adjacent frames, the object movement is minimal, and the brightness remains constant.

2) Continuous image extraction.

3) The object movement maintains spatial consistency, meaning that pixels in the same sub-image move in unison.

Given that each image point has two velocity components, it is not feasible to calculate optical flow at any given location in an image plane without imposing additional constraints. In this case, one restriction is imposed on the image brightness at a particular location in the image plane due to motion, while two constraints are applied to the change in image brightness at the same position in the image plane due to motion. The optical flow approach operates under the assumption that the intensity I of moving points remains constant over a period of time in its most fundamental component. The assumption of brightness consistency is referred to as the brightness constancy assumption (Siong *et al.,* 2009). Set $I(x, y, t)$ as the luminance of an image pixel (x, y) at time t According to the constant brightness of optical flow and the assumption of tiny motion, there is (Yu *et al.,* 2017):

$$I(x + \Delta x, y + \Delta y, t + \Delta t) = I\ (x, y, t) \quad \text{(Equation 2.17)}$$

According to the Taylor series,

$$I(x + \Delta x, y + \Delta y, t + \Delta t = I(x, y, t) + \Delta x \frac{\partial I}{\partial x} + \Delta y \frac{\partial I}{\partial y} + \Delta t \frac{\partial I}{\partial t} + e \qquad \text{(Equation 2.18)}$$

Where e is a high-order error term on $\Delta x, \Delta y, \Delta t.$ From Equation 2.17 and 2.18,

$$\frac{\Delta x}{\Delta t}\frac{\partial I}{\partial x} + \frac{\Delta y}{\Delta t}\frac{\partial I}{\partial y} + \frac{\partial I}{\partial t} = 0 \qquad \text{(Equation 2.19)}$$

That is:

$$I_x u + I_y v + I_t = 0, \qquad \text{(Equation 2.20)}$$

where $I_y = \frac{\partial I}{\partial y}, I_t = \frac{\partial I}{\partial t}$ are the gradient of the image in space and time respectively and:

$$u = \frac{\Delta x}{\Delta t}, v = \frac{\Delta y}{\Delta t}$$

shows the optical flow velocity in x and y components respectively. It is also possible to see how the object is moving by looking at its equations. Based on Equation 2.20, $u \; and \; v$ represent the optical flow vectors and I_x, I_y, I_t represents the derivative of images intensities at coordinate (x, y, t).

Most optical flow approaches are based on the assumptions of brightness constancy and spatial smoothness (Black and Anandan, 1993). This assumption assumes that points maintain their intensity over the course of a sequence of frames, which is known as the brightness constancy assumption. Following the article's assertion, spatial smoothness is assumed to exist since neighboring pixels are typically connected with the same surface and so move in a comparable manner (Sun *et al.*, 2008). There are different optical flow computations method that could categorize to sparse and dense optical flow (Boer and

Kalksma, 2015). Lucas-Kanade algorithm is sparse method and Gunnar Farnebäck is the dense method which is a two-frame motion estimation algorithm.

2.3.1.1.1.1 Lucas-Kanade and Gunnar Farnebäck algorithm

The Lucas-Kanade method describes an image registration technique that searches for the best match based on spatial intensity gradient information (Lucas and Kanade, 1981). The program accomplishes this by considering additional information about the image. As a result, the method can find the best match while requiring significantly fewer computations than other techniques that search in a fixed order. The technique takes advantage of the fact that, in most circumstances, the two images are already relatively close to one another when the algorithm is performed. To solve the registration problem, one must first determine which vectors F(x) and G(x) have the smallest distance between them, and then determine which vectors F (x + h) and G(x) have the smallest distance between them in a region of interest R. For distortions such as image rotation, Lucas-Kanade suggests a generalization that can be used to deal with the problem. Lucas-Kanade is a sparse optical flow algorithm since it only employs specific pixels of the image to measure the optical flow. The Lucas-Kanade method is used to calculate optical flow given a sparse feature set of data (e.g., corners detected using Shi-Tomasi algorithm). This approach, developed by Gunnar Farnebäck, is a motion estimate algorithm that makes use of two consecutive frames of data. Gunnar Farnebäck approximates the motion between frames by using quadratic polynomials, which he developed himself (Farnebäck, 2003). Using the polynomial expansion transform, this can be accomplished quickly and efficiently. This algorithm computes the optical flow for all the points in the frame. VO techniques allow the estimation of motion by taking cues from the images. A moving object in space can be approximated by using the pixel motion velocity under the image grey model, which is implemented using the image grey model. The machinery vibration or deformation is generally very small and high frequency. The pixel motion shall undergo some motion magnification methods to amplify the tiny motions that were previously undetectable to the human eye. The magnification results are usually prone to noise or excessive blurring.

2.3.2 Improvement of Visual Odometry (VO) System

VO techniques, which use visual cues to estimate motion, allow for the estimation of motion to be done in real time. A moving object in space can be estimated by using the pixel motion velocity under the image grey model, which is implemented using the image grey model. The frequency and size of the machinery's vibration and deformation are generally small and high. Various motion magnification and amplification techniques will be applied to the pixel movement to amplify the small movements that were previously undetectable by the human eye. The results of magnification are typically prone to noise or excessive blurring when magnified. Because of this, it is necessary to select the MM method that is most appropriate for eliminating noise and overcoming the difficulties associated with the different nature of the object's movement and its environmental context. This is also true for the monitoring of the condition of technological objects. In this case, it is considerably more difficult to establish a clear criterion for evaluating the processing result accomplished as well as the methodology that was used. The original goal of video MM approaches is to overcome or break the boundary between human perception of changes (displacement) in the things observed and their perception of the same objects (Śmieja *et al.,* 2021).

2.4 Motion magnification (MM) Technique

Using modern picture analysis and processing technologies, the options for vibration detection that were previously possible have been greatly expanded. They devised and implemented a system for observing static and dynamic strains in a civil structure employing triangulation and point-tracking technologies (Yang, 2019). According to the name of this approach, multi-point measurement (MM) is a substantial collection of visual techniques that provide considerable benefits, including being able to simultaneously capture data on several locations as well as operating in a non-contact mode. It is possible to observe physical processes indirectly via the use of sensors and equipment that transform read-off signals into a form that can be evaluated by researchers. Allows one to go beyond the normal threshold of visual perception, which is the point when even slight changes in an observer's field of vision may be reflected in consciousness, or "consciously perceived," using this approach. Accordingly, when discussing the real geometric quantities that determine an object's location, "minimum" does not hold. In this method, the findings are presented in a quantitative way, and the visualizations are made utilizing graphs like histograms, charts, and phasors. It is possible to express the information contained in the vibrations of an object as displacements or mutual displacements of its chosen places (deformations) through time using an alternative, direct technique that is descriptive and only to a limited degree reliant on the observed real picture. The following are some of the things to keep in mind while creating a spatial-temporal data visualization (Wu *et al.,* 2018). Providing that analysis of vibrations a upgoing process and that the goal is for gaining information on potential consequences or causes for the phenomenon, a direct approach can be far more effective than a methodical approach.

Various of image processing techniques can be used to quantify motion in video. To measure any structural motion more easily, the past research have used edge detection, target objects, or light (Patsias and Staszewskiy, 2002). The utilization of computer vision techniques, including measurement of OF to find the displacements of structures, in more recent methods, which are link to methods described in this paper, has become increasingly popular (Caetano, 2011). When using a single camera, only in-plane motion can be measured; however, a stereo camera setup is capable of measuring both in-plane and out-of-plane motion which is more accurate. Currently, the time-of-flight camera is a more recent method of in-plane and out-of-plane measurement with cameras, but it does not yet provide sufficient resolution or speed for most vibration measurement applications currently (Kolb *et al.*, 2009). Small motions in videos can be magnified using new computer vision techniques, collectively known as MM which were introduced recently (Liu e*t al.*, 2005). A signal processing approach is used by the most recent motion magnification techniques to investigate image motions in a manner analogous to that of a Eulerian framework used in fluid flow analysis. Using temporal filtering, they are particularly well suited for visualizing mode shapes because they can detect small subpixel motions that are present in videos of vibrating structures and because they can distinguish between different modal motions. Because of the assumption of a non-moving camera and an immobile framework, this works very well in this situation. For non-destructive testing (NDT) and structural health monitoring (SHM), it is considered that using video cameras for vibration measurement represents a fresh capacity that would integrate various measurement and sensor systems that are either presently in use or being explored (Chen *et al.*, 2015).

While extensive measuring processes are required for many technological things, it is far simpler to make conclusions about them using a comprehensible depiction of vibrations rather than complex measurement procedures; this is especially true for mechanical objects. As part of the investigation into how to extend visual observation of vibrations while considering the extra element between the observer and moving object, a camera, operations were carried out to address this issue. These investigations in image analysis and processing were supported by a substantial rise in the computing power of widely accessible electronic devices. It is possible to extract information from a raw image sequence recorded on a camera sensor, such as the positions of 3D object points projected onto 2D planes at various points in time during the duration of the series' recording. When the original recording is reproduced, the observer will not notice any small differences between adjacent images that correspond to the object's movement because the differences are too small. There are several obvious reasons why it is not possible to change one's natural way of seeing. As a result, strategies for modifying a series of images captured by a camera in such a way that significant variations between the changing frames of the images that reach the observer are positioned above their visual perception threshold have been created in order to produce the intended effect. In video sequences, MM are used to shift picture changes that correlate to the motion of the item under observation into the field of visual perception. (Wadhwa *et al.,*2 017)

Video motion magnification is a technique for magnifying video motion. MM techniques can be divided into two categories (Oh *et al*., 2018): Lagrangian approaches and Eulerian approaches. Lagrangian approaches are the most widely used. When the Lagrangian approach is used, the motion field (optical flow) can be explicitly extracted and used to directly move the pixels. Instead of explicitly tracking motions, Eulerian approaches break down video frames into representations that can be used to manipulate them (Wu *et al*., 2012). These techniques are typically divided into three stages: decomposing frames into an alternative representation, manipulating the representation, and reconstructing the manipulated representation to magnified frames if applicable. When it comes to extracting a phase-based representation, the researcher makes use of a spatial decomposition that is motivated by the first-order Taylor expansion (Wu *et al*., 2018), and other researcher make use of the complex steerable pyramid (Freeman, 1991). Current Eulerian approaches are effective at revealing tiny motions, but they are labor-intensive to develop by hand (Wu *et al*., 2018) (Wang *et al.,* 2017) (Wu *et al.,* 2012) and do not consider many issues such as occlusion. As a result, they are prone to noise and are frequently plagued by blurring that is extreme. The research will belong to the Eulerian approach, decomposition is directly so it has fewer edge artifacts and better noise characteristics.

Eulerian approach, which are terminology borrowed from fluid mechanics, are the most frequently proposed methods in recent years (Wu *et al.,* 2012) (Wadhwa *et al.,* 2013) (Wadhwa *et al.,* 2014). Against Lagrangian approaches, Eulerian approaches can amplify small displacements or variations that evolve over time without the need for explicit optical flow computation (Liu *et al*., 2005). When it comes to color visual representation of face video and small motion magnification, Eulerian video magnification (EVM) has produced impressive results (Wu *et al*., 2012). The linear EVM has the disadvantage of being able to support only small magnification factors in regions with high spatial frequencies, which is a drawback. Furthermore, when the magnification factor is increased, it has the potential to significantly amplify noise. As a result, the phase-based Eulerian motion magnification techniques are recommended as an alternative to conventional techniques. Using techniques inspired by the Fourier shift theorem, it is possible to establish a connection between phase variations and motions in three-dimensional space (Wadhwa *et al.*, 2013) (Wadhwa *et al*., 2014). The methods outperform the EVM in terms of noise handling characteristics, and they also support higher magnification factors. Unfortunately, when it comes to magnifying minute differences in the presence of significant amounts of motion, these strategies are useless because of the large quantity of motion involved. A substantial amount of motion will result in significant blurring artefacts and will completely overpower the modest temporal changes that will be enhanced in video due to the compression.

CHAPTER 3: METHODOLOGY

3.1 Introduction

Chapter 3 discusses the methodology of this project. The discussions in this methodology emphasized clearly about the method, knowledge, and procedures of the visual odometry system. The methodology constructed by the project planning, schedule planning, designing the experiments and the methods that use to complete this project. It is important to construct a flow chart for project planning. Flow chart is a tool that used to visualize and illustrate the process sequences from the first step to the end of the project. The description of each of the step are clearly arranged and followed the sequences. The flow chart as shown as Figure 3.1 is made for clearly understanding the detailed processes and processes sequences in this project. From the literature review stage, some discussions and comparisons are made to identify the most methods to be selected and implement in this project. For the visual odometry system, the video processing algorithm is being planned to amplify or magnify the motion of the video. The data will be obtained after the video processing algorithm run successfully and end this project.

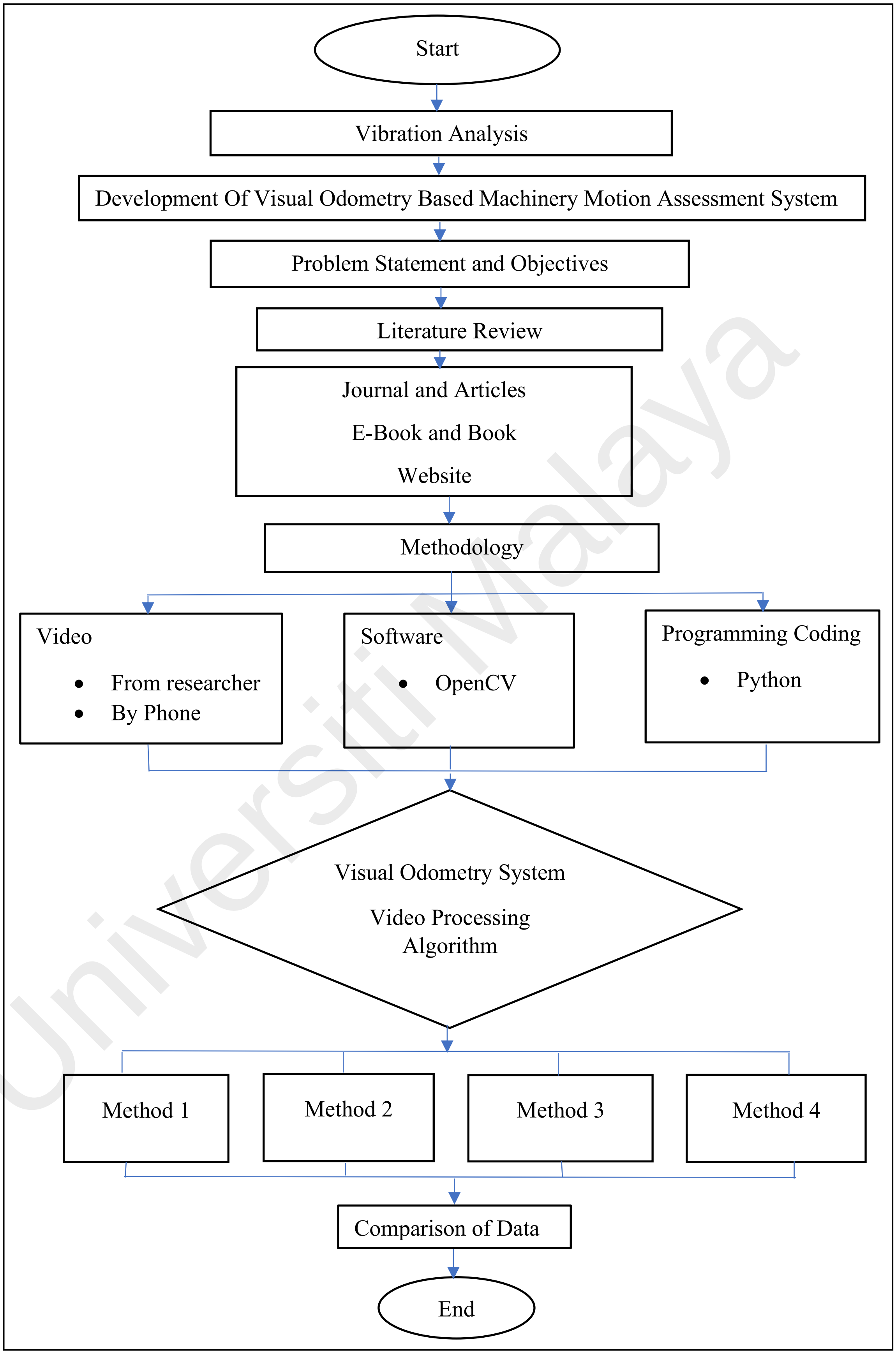


Figure 3.1: Flow Chart of Project

This research report is to implement motion amplification method on video using optical flow. The system is implemented by using the demonstration on how to use VO method to have motion amplification in video using optical flow. First, the optical flow between adjacent frames of the video footage of objects undergoing low-amplitude oscillations that are invisible to the naked eye have been calculated using Gunnar-Farnebäck algorithm. Furthermore, by analyzing the PSD through Welch method, the PSD of the angular and magnitude components of the optical flow should be found. The video is warped with a large-amplitude representation of the original small-amplitude motion by multiplying the optical flow by an amplification factor. It is also possible to reduce noise and isolate individual modes by looking at the PSD of a video (Hyatt *et al.*, 2019).

Previously, high-quality motion amplification of video data was accomplished using Eulerian techniques and commercially marketed for the purpose of detecting flaws in structures and equipment. These methods do not account for optical flow, which can be used to achieve arbitrarily large amplifications via image warping. Similarly, optical flow calculations are well established, but are almost always applied to large displacements, such as those encountered in object tracking. Calculating the optical flow enables us to disentangle the observed motion's components such as magnitude and angular. Additionally, it opens the possibility that the process could be significantly accelerated using a machine learning model capable of extracting optical flow from video data. The system briefly investigated machine learning-based optical flow for motion amplification, but without immediate success. This could be because, as with conventional optical flow calculators, the available pre-trained machine learning models are optimized for large displacements to be used for object tracking on motion. In this research, four different types of methods have been used, and the results obtained from each method will be compared in the next chapter. The following sessions have covered the theories that have been used to develop these methods.

3.2 Calculation of Optical Flow (OF) method using Gunnar-Farnebäck Algorithm

Motion detection is implement using optical flow method. Algorithms for estimating optical flow are capable of tracking points across two images. These techniques assume of brightness constancy and spatial smoothness. It maintains the same intensity between frames under the brightness constancy assumption. In the spatial smoothness assumption, adjacent pixels are assumed to belong to the same surface and thus have similar motion. OF is the distribution of an image's apparent velocities. the velocities of objects in the video are determine by estimating optical flow between video frames. To determine the optical flow between two images, it is necessary to solve the following OF equation (Wójcik *et al.,* 2014):

$$I_x u + I_y v + I_t = 0 \qquad \text{(Equation 3.1)}$$

Where:

- Ix, Iy, It are the spatiotemporal image brightness derivatives,
- u is the horizontal optical flow,
- v is the vertical optical flow.

More details of the formula of OF have been discussed in 2.3.1.1.1 sessions**.** OF is the apparent velocity of pixel intensity in a video, defined by

$$(\nabla I) \cdot V = -\partial_t I \ . \qquad \text{(Equation 3.2)}$$

I (x, y, t) is the intensity of a pixel at coordinates (x, y) and time t, and V (x, y, t) = $V_x\hat{x} + V_y\hat{y}$ is the optical flow. This is an ill-posed problem as there are two unknowns, V_x and V_y**.** In this research, the optical flow would be using Gunnar-Farnebäck method. The Gunnar

Farnebäck algorithm was designed to work as dense OF technique results that is on a dense grid of points. The first step is to approximate each neighborhood of both frames by quadratic polynomials. Gunnar-Farnebäck approximates the motion between the frames via quadratic polynomials, which he developed himself. Using the polynomial expansion transform, this may be accomplished quickly and efficiently. According to Gunnar-Farnebäck the point of interest is quadratic polynomials, which provide the local signal model stated in a local coordinate system, as shown in Farnebäck (2003):

$$f(x) \approx x^T A x + b^T x_c \qquad \text{(Equation 3.3)}$$

Where A is a symmetric matrix, b is a vector and c are a scalar.

Different methods for calculating optical flow make different assumptions to constrain the equation, but if using Farnebäck method, it will efficiently compute the optical flow over the entire frame. Specifically, the pyramidal implementation calcOpticalFlowFarneback in OpenCV software would be use. Rather than only calculating the flow once, across the base image pair, it repeatedly downs samples the images, obtains the flow at the lowest resolution, and updates it with the residual flow obtained at each successively higher resolution. The algorithm outputs the optical flow in the (x, y)-coordinate system, but it is more useful to visualize in the polar (r, θ)-system, where r is the magnitude and θ is the direction of the flow.

3.3 Power Spectral Density (PSD) Analysis

The calculated optical flow frequently contains a significant amount of noise in addition to any noise that may have been present in the original video, and it may also contain multiple distinct motions that must be amplify separately from one another. The PSD of the pixel intensity in the original video can be reduced and isolated using Welch's method, which is implemented in SciPy as signal.welch. This method is used to reduce noise and isolate specific frequency components. A popular non-parametric method in spectral analysis, the Welch PSD method takes advantage of the FFT to produce a spectral analysis result and makes the frequency spectrum smoother than raw FFT output (Parhi and Ayinala, 2014). The primary advantage of this method is that it reduces the number of computations required as well as the amount of core storage required. This procedure includes segmenting a signal, obtaining modified periodograms of each section, and averaging the modified periodograms obtained from each section (Thomas *et al*., 2015).

The complete algorithm can be described as follows (Parhi and Ayinala, 2014):

- Input signal $x[n]$ is divided into L overlapping segments.
- Each segment is being applied for the specified window.
- FFT is applied to the windowed data.
- Each windowed segment's periodogram is computed, which is referred to as the modified periodogram.
- Modified periodograms are averaged for obtaining the spectral estimate $S(k)$.

Welch method can be described in a mathematical form, let

$$x_1(n) = x(n + (l-1)M), n = 0, \ldots, N-1 \;\; l = 1, \ldots, L \qquad \text{(Equation 3.4)}$$

Denote the lth data segment. $(l-1)M$ is starting point for the lth sequence of observations. In Welch method, value recommended for M is $M = N/2$, the data segments contain 50% overlap between successive segments (Stoica and Moses, 1997). The windowed periodogram corresponding to $x_1(n)$ is computed as

$$A_l(k) = \sum_{n=0}^{N-1} x_1(n)\omega(n)e^{-j\frac{2\pi}{N}nk}$$

$$\emptyset_l(k) = \frac{1}{NP}|A_l(k)|^2, l = 1, \dots, L$$

where A_l is the FFT of windowed segment, $\emptyset_l$ is the periodogram and P denotes to power of the window $(\omega(n))$:

$$P = \frac{1}{N}\sum_{n=0}^{N-1}|\omega(n)|^2$$

The Welch method estimate PSD is the average of these periodogram:

$$S(k) = \frac{1}{L}\sum_{l=1}^{L}\emptyset_l k$$

(Equation 3.5)

Welch's method, which is implemented as a signal in SciPy, can be used to reduce and isolate the PSD of the pixel intensity in the original video and then isolate it. Using an average of all pixels and frames, PSD were created and then vertically scaled to improve clarity. An example of PSD graph that contains the pixel intensity, optical flow angular and magnitude component of a vibrating phone is as shown in Figure 3.1. The pixel intensity PSD is plotted in red, while the angular and magnitude components of optical flow are plotted in blue and green, respectively, on the same plotting plane. Gray regions indicate the frequency bands that will be amplified. In the intensity PSD, the two peaks that can be observe are reproduced in the PSD of the optical flow's angular component, but not in the PSD of the magnitude component. In the video, this indicates that the corresponding features are moving back and forth across the frame, rather than brightening and darkening, as one would expect from a vibrating phone. Rather than brightening and darkening as would be expected from a vibrating phone, the corresponding features in the video appear to be moving back and forth across the frame. The angular PSD is dominated by noise at high frequencies, which is a negative sign. Band-pass filters will be used to eliminate this problem and separate the motion at the peak frequencies.

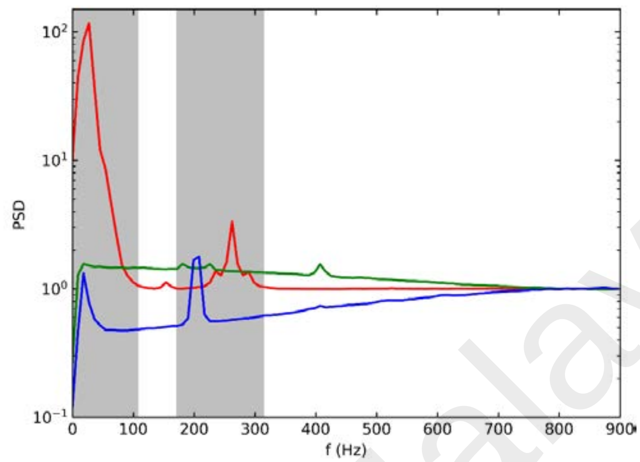


Figure 3.2: PSD obtained from phone video (Hyatt *et al.,* 2019)

3.4 Video Warping

The OF is multiplied by an amplification factor and then use that vector field to warp the original video using the OpenCV function remap to create a distorted version of the original. The calculated optical flow approximates the motion observed between two frames of video, whereas the amplification factor exaggerates this motion to the point where it can be seen with the naked eye between the frames of video. This must be determined on a case-by-case basis; if the amplification factor is too small, motion is not sufficiently amplified; if it is too large, the video is warped beyond recognition of its original objects. In addition, the amplification factor must account for the reduction in optical flow magnitude caused by the frequency filtering operation.

3.5 Eulerian Video Magnification (EVM)

Video analysis software such as EVM can detect and exaggerate even the tiniest movements and changes in a video recording. Because the purpose is to amplify the pyramid levels that contain movement frequencies, EVM is applied to every level within a pyramid rather than over the original images, as is the case with other types of image augmentation techniques. The desired frequencies are multiplied by a quantity known as the magnification factor, which is specified by the user once the desired frequencies have been determined. Because EVM enhances the real motion, it is possible to see movements that would otherwise be invisible to the naked eye with EVM. It is necessary to add back the magnified values of the desired frequency to the non-magnified values of the same level to create the final movie with exaggerated motion, which is presented on the screen.

The EVM approach can be separated into two categories: linear-based EVM and phase-based EVM. The linear-based EVM technique is the more common of the two. In linear techniques, the variation in intensity over a first-order extension of the Taylor series is linearly proportional to the variation in motion in video, and the reverse is true. A video sequence is taken into consideration as input, after which spatial decomposition is used, and frames are filtered by a temporal filter before being produced in the final product. Increasing the amplitude of the created temporal region allows for the discovery of previously unseen information (Shahadi *et al.*, 2020). This approach is basic and capable of detecting small motion fluctuations in a short amount of time; but, when the magnification factor is increased sufficiently, the Taylor approximation becomes imprecise, and the procedure is liable to failure. For this reason, phase-based magnification (PVM) is used to address the difficulty. It replaces the linear approximation with a Fourier decomposition by a complex steerable pyramid, which is a Fourier decomposition by a complex steerable pyramid (Wadhwa *et al.*, 2013). Pyramid coefficients move in a manner proportionate to the motion of the pyramid coefficients in different video frames over time, and this is demonstrated by the fluctuations in phase over time.

It is possible to analyses these fluctuations in real time and then amplify them to create a visual depiction of the motion. If we compare it with the linear-based technique, the phase-based method has a higher level of complexity and needs a longer amount of processing time, but it can accommodate higher levels of movement augmentation. Visualizations enlarged using Eulerian linear and phase-based algorithms are both faster and produce less noise than visualizations amplified using Lagrangian-based video magnification techniques (Shahadi *et al.,* 2020). It is possible to achieve small motion amplification through EVM. The first-order Taylor series expansions can be used to produce motion magnification in optical flow by using temporal processing in conjunction with the optical flow (Horn and Schunck, 1981). Aiming to process the time series of colour values for each pixel in the spatial domain independently by applying a standard 1D temporal signal processing to each time series to amplify a specific band of interest temporal frequencies in each time series is the goal of EVM's targeting process. The input video frame is dissected into various spatial frequency bands using a full Laplacian pyramid, which decomposes the frame into distinct spatial frequency bands. (Wu *et al*., 2012). which is composed of three layers. When an image is down sampled at successively sparser densities until no further down sampling is possible, the Laplacian pyramid is used as a data structure. It is necessary to have a video analysis pyramid that is based on a Gaussian pyramid for the Laplacian pyramid to work (Sahadi *et al*., 2020). It is becoming less popular to use the Laplacian pyramid method for analyzing video processing time than it was previously. In fact, as the magnification factor increases, this method becomes increasingly ineffective because when noises increase, magnification factor also increased (Wu *et al.,* 2012). If small magnification factor is utilized, the approach would be suitable for magnifying colors. Figure 3.3 depicts a functional LVM mechanism acquired from the research.

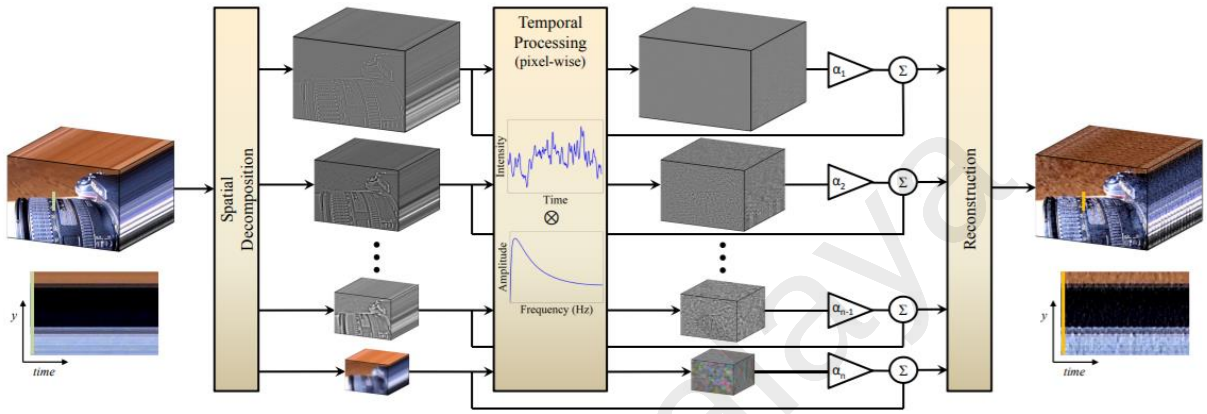


Figure 3.3: Overall structure of the linear-based-EVM (Wu *et al*., 2012)

3.5.1 Spatial-Temporal Information Processing

When it comes to motion estimation, spatial temporal processing is one of the most promising approaches. Motion selectivity can be achieved using spatial-temporal filtering based on the continuous wavelet transform (CWT*)* (Rui *et al*., 2011). Temporal processing is utilized at the past to extract signals that were not visible to the human eye. Temporal filters have been shown to reduce the occurrence of temporal aliasing of motions in videos (Fuchs *et al*., 2010). The Eulerian method, which makes use of spatial temporal filtering, can be used to extract small and subtle motions from video (Wu *et al.,* 2012). As illustrated in Figure 3.3, spatial processing is the first step in the Eulerian motion magnification process. Using a pyramid structure, a video sequence is decomposed into different spatial frequency bands. Depending on the signal-to-noise ratio (SNR), these frequency bands may contain a variety of different spatial frequencies. Low pass filtering is used to suppress artefacts in these bands later in the recording process. Afterwards, frames of the video are down sampled to improve the computational efficiency. To detect motion, these selective pixel bands are used in conjunction with temporal filtering.

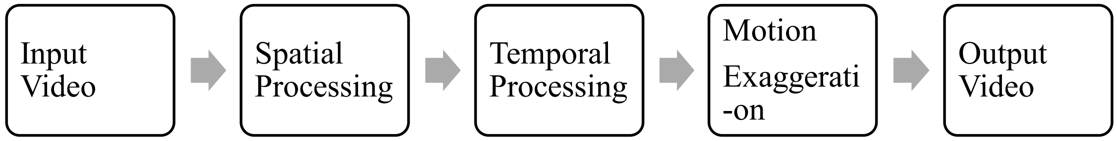


Figure 3.4: Eulerian Motion Magnification Process Flow.

3.5.2 Relation between Temporal Filtering and Magnification

A video is subjected to temporal processing after it has been spatially processed, as depicted by the diagram in Figure 3.4. It is possible to demonstrate the relationship between temporal filtering and motion magnification using the Taylor series expansion in equations Equation 2.17 and 2.18. Taylor series expansion for 1D images can also be derived, and this can then be applied to 2D images to achieve the desired result (Wu *et al.*, 2012).

Let $I(x,t)$ denote the image intensity at position x and time t. The image undergoes translation motion, the observed intensities can be express respect to a displacement function $\delta(t)$, such that $I(x,t) = f(x + \delta(t))$ and $I(x,0) = f(x)$. The aim of motion magnification is to synthesize the signal

$$\tilde{I}(x,t) = f(x + (1+\alpha)\delta(t)) \qquad \text{(Equation 3.3)}$$

for some amplification factor α.

Assuming the image can be approximated by a first-order Taylor series expansion, the image at time t, $f(x + \delta(t))$ in a first-order Taylor expansion about x can be written as

$$I(x,t) \approx f(x) + \delta t \, \frac{\partial f(x)}{\partial x} \qquad \text{(Equation 3.4)}$$

Let the result of applying a broadband temporal bandpass filter to $I(x,t)$ at every position x to be $B(x,t)$. A temporal band pass filter is considering with assumption that motion signal δt is within the band pass.

$$B(x,t) = \delta t \, \frac{\partial f(x)}{\partial x} \qquad \text{(Equation 3.5)}$$

The band pass signal is amplifying by exaggeration or amplification factor α and adding back to $I(x,t)$ gives the processed signal

$$\tilde{I}(x,t) = I(x,t) + \alpha \, B(x,t) \qquad \text{(Equation 3.6)}$$

By combining Equation 3.4,3.5 and 3.6 it can be get that

$$\tilde{I}(x,t) \approx f(x + (1+\alpha)\delta(t)\frac{\partial f(x)}{\partial x} \qquad \text{(Equation 3.7)}$$

Given that the first-order Taylor expansion applies to the amplified larger perturbation, $(1+\alpha)\delta(t)$, the amplification of the temporally bandpassed signal can relate to motion magnification. The process output is

$$\tilde{I} \approx f(x + (1+\alpha)\delta(t)) \qquad \text{(Equation 3.8)}$$

This shows that spatial displacement$\delta(t)$ of the local image $f(x)$ at time t,has been amplified to a magnitude of $(1+\alpha)$. In figure below, the process is illustrated.

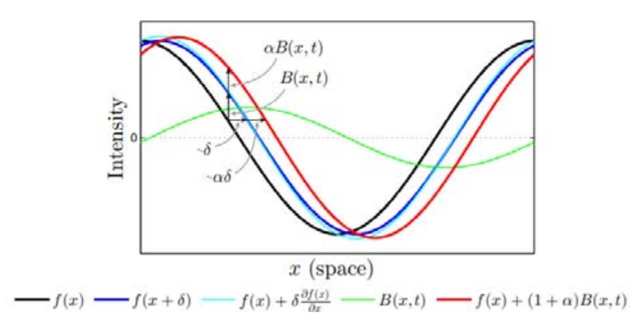


Figure 3.5: Spatial translation can be approximated using temporal filtering.

Based on Figure 3.5, this effect is shown here on a 1D signal, but it is applicable to 2D signals as well. The input signal is shown at two different time instants: $I(x,t) = f(x)$ at time t and $I(x,t) = f(x + \delta)$ at time $t + 1$. When $I(x, t + 1)$ is expanded about x, the first order Taylor series expansion approximates the translated signal very well. The temporal bandpass is amplified and combined with the original signal to generate a larger translational range. In this example, = 1, which magnifies the motion by 100% and the temporal filter is a finite difference filter, which subtracts the two curves.

For a low frequency cosine wave and a relatively small displacement, $\delta(t)$, the first-order Taylor series expansion serves as a good approximation for the translated signal at time $t + 1$. When boosting the temporal signal by α and adding it back to $I(x,t)$, the wave translated by $(1 + \alpha)\delta$. If for more general case where $\delta(t)$ is not entirely within the passband for temporal filter, let $\delta_k(t)$, indexed by k, represent the different temporal spectral components of $\delta(t)$. Every $\delta_k(t)$ will be attenuated by temporal filtering by a factor γ_k. Bandpassed signal results in

$$B(x,t) = \sum_k \gamma_{k\delta_k}(t) \frac{\partial f(x)}{\partial x} \qquad \text{(Equation 3.9)}$$

As reason of multiplication in Equation 3.6, the temporal frequency dependent attenuation can equivalently be interpreted as a frequency-dependent motion magnification factor, $\alpha_k = \gamma_k \alpha$, resulting in a motion magnified output,

$$\bar{I}(x,t) \approx f(x + \sum_k (1 + \alpha_k)\delta_k(t)) \qquad \text{(Equation 3.10)}$$

The result is the predicted for a linear analysis where the modulation of spectral components of the motion signal becomes modulation factor in motion amplification factor, α_k, for each temporal sub band, δ_k of motion signal.

3.5.3 Temporal Filter

The goal of this procedure is to isolate motions occurring at specific temporal frequencies so that they can be magnified. While building the pyramid, the phases on each spatial scale and orientation are isolated from one another. Once the differences between them have been determined, the process of temporal bandpass filtering is carried out on the data. Enhancing either color alterations or local phase variations is essential for improving filter performance, and this might be accomplished by increasing the signal-to-noise ratio of the temporal variations. Temporal and spatial filtering are applied to the variations, allowing for the elimination of noise-related components and the preservation of signal-related components, thereby improving the signal to noise ratios. As a result of the fact that different motions occur at different temporal frequencies, temporal filtering can be used to isolate a signal of interest Using a spatial pyramid construction, a temporal filtering process is applied to a series of temporal pixels within each spatial band to extract the frequency bands that are of interest. To extract motions or signals that are intended to be amplified, a temporal bandpass filter is used in conjunction with an amplifier. Users should be able to control the frequency bands of interest, according to the application that was used in the algorithm. Certain cases, however, allow for the selection of a frequency band to be automated. The type of application being used has an impact on the selection of filters as well as other factors. Consider the following examples: a wide pass band filter is frequently used for motion magnification, while narrow pass band filters are frequently used for color enhancement, such as blood flow, because the latter produces less noise distortion (Shahadi *et al.,* 2020). In this research, the bandpass filter is used for motion magnification of videos.

3.5.4 Pyramid Decomposition

The process of breaking down images into different spatial scales is known as a pyramiding process. In the field of pyramid decomposition, the Gaussian and Laplacian approaches are the two most widely used methods. It is possible to create a Gaussian pyramid by first smoothing the original image with a Gaussian filter and then scaling it down to the appropriate size. A Gaussian pyramid is a series of lowpass, down sampled images that are stacked on top of each other. The Gaussian pyramid technique is a type of image processing technique that is used to enhance contrast in images (Yamana *et al.,* 2000). This is like the Laplacian pyramid in that at a certain level of each picture of the Laplacian pyramid, the distance between different corresponding nearby levels of the Gaussian pyramid is equal to the difference between two input surrounding levels of the Laplacian pyramid. The tiniest level is maintained. Therefore, the difference images can be used in conjunction with the original image to reconstruct the original image. A sequence of bandpass-down sampled images can be assumed to be the Laplacian pyramid (Burt and Adelson, 1983).

3.6 Different Methods Implemented for Video Processing

Different methods have been implemented and tested in this research project for different videos to develop the visual odometry system for the demonstration of a non-contacting vibration analysis method. Three videos have been tested in this research report for the purpose of better illustration and comparison of the motion amplification from the result that could be obtained from different type of methods. The screenshots from the videos, which include a sleeping baby, a vibrating guitar string, and a vibrating phone, are depicted in the figures below. The baby and guitars video are extracted from others research and have been magnified for visualizing the vibration motion (Wadha *et al.*, 2017). The vibrating phone which is the model of Huawei Nova 3i is taken by the phone Redmi Note 9 Pro.

Figure 3.6: Sleeping baby

Figure 3.7: Guitar

Figure 3.8: Vibrating Phone

The different properties of videos have been shown in the table below.

Table 3.1: Video Properties.

Video	Length (s)	Frame width	Frame height	Frame rate (frames/second)
Sleeping Baby	10	960	544	30
Guitar	10	432	192	600
Vibrating Phone	10	1920	1080	30

There are three proposed methods implemented in this research to develop the visual odometry based for motion assessment system which is Method 1,2,3 and 4. Figures below have shown the different flow charts for each method. For the first and second method which OF would be the first step implemented to estimate the motion between frames of the video. Then the next step would be getting the PSD intensity graph from the original video and getting PSD angular and magnitude component from the optical flow, then multiply the OF by an amplification factor and use that vector field to warp the original video with the OpenCV function remap. The calculated OF approximates the motion observed between two frames of video, while the amplification factor exaggerates this motion so that it is visible to the naked eye. The major difference of method 1 and method 2 is in method 2 bandpass filter have been applied to obtain a better output of amplified video. For method 3, it is proposed to use EVM method to obtain motion magnification video output. Method 4 is purely linear based EVM method that reproduce from the article by Wu *et al*. to act as a benchmark (Wu *et al.,* 2012).

Method 1 algorithm is as follows:

1) Create a mask for the parts of the video where no motion occurs.

2) Obtain the OF between frames of the video.

3) Plotted the PSD of pixel intensity from the original video.

4) Plotted the PSD of angular and magnitude component of original video.

5) Magnify the amplitude by multiply the amplification factor to OF video and warping.

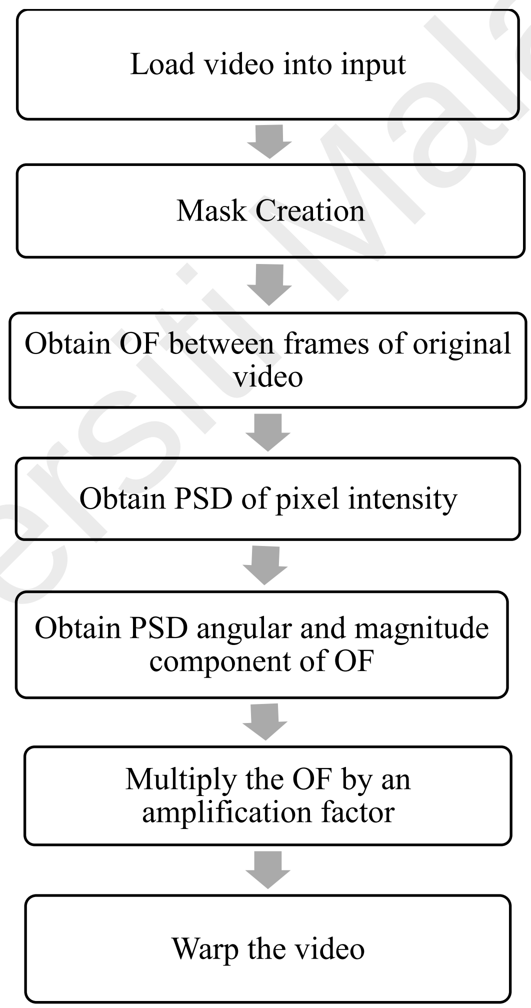


Figure 3.9: Process Flow Chart for Method 1.

Method 2 algorithm is as follows:

1) Create a mask for the parts of the video where no motion occurs.

2) Obtain the OF between frames of the video.

3) Plotted the PSD of pixel intensity from the original video.

4) Plotted the PSD of angular and magnitude component of original video.

5) Filter the cut-off frequency.

6) Magnify the amplitude by multiply the amplification factor to OF video and warping.

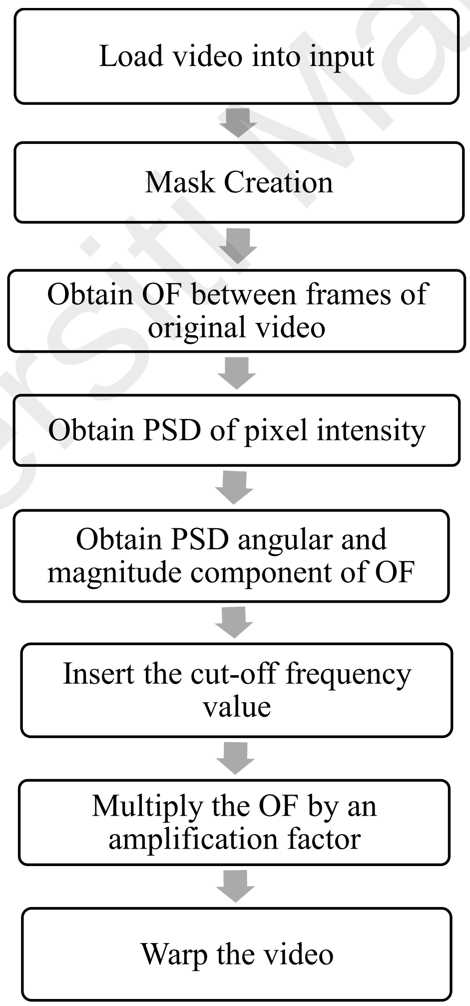


Figure 3.10: Process Flow Chart for Method 2.

Method 3 algorithm is as follows:

1) Create a mask for the parts of the video where no motion occurs.

2) Obtain the OF between frames of the video.

3) Plotted the PSD of pixel intensity from the OF filtered video.

4) Plotted the PSD of angular and magnitude component of OF filtered video.

5) Magnify the OF filtered video by EVM.

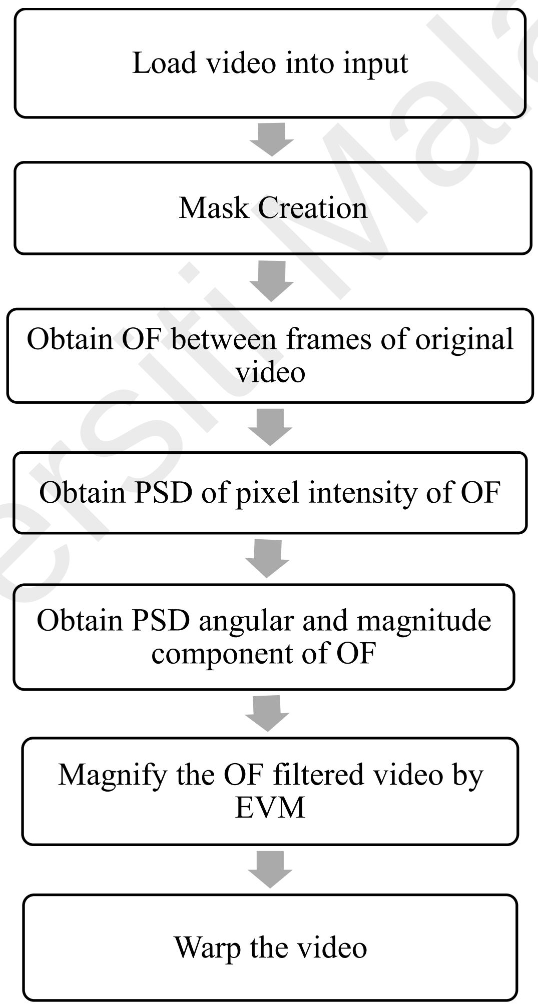


Figure 3.11: Process Flow chart for Method 3.

For method 4, processing an input video using EVM, have four procedures as below:

(1) Select a temporal bandpass filter.

(2) Select an amplification factor, α.

(3) Select a spatial frequency cutoff (specified by spatial wavelength, λc) beyond which an attenuated version of α is used.

(4) Select the form of the attenuation for αeither force α to zero for all $\lambda < \lambda c$, or linearly scale α down to zero. The frequency band of interest can indeed be chosen automatically in some circumstances, but it is typically vital for users to be able to control the frequency band matching to their application. The user has complete control over the amplification and cutoff frequencies in real-time program.

3.7 Implementation of Software and Programming Code

For developing the visual odometry system, software to process the video and produce the output for further motion magnification is essential for video processing. For this project, the software using for video processing is OpenCV and the coding language for implementation of pseudocode is Python Language. To complete the visual odometry system for amplification and magnification of the vibration amplitude of the videos, OpenCV-Python which is a library of Python bindings is designed and use in this project to solve the computer vision problems. There are different syntaxes that written in the video processing code to develop various processing function through OpenCV-Python library.

3.7.1 OpenCV Software

OpenCV is the video processing software is chosen to use for this project. OpenCV is an acronym for the Open-Source Computer Vision Center. Comparing to others available software, OpenCV is the software that most suitable to use in this project because it is free of charge to utilize. Intel was originally designed to provide access to the image processing technology needed to create computer vision applications. It is a collection with numerous built-in functions that are mostly designed to work in real time. As the software is open source, the features are invented and modified constantly. The built-in library features require multi-core computing to be utilized. It is free of charge for industrial as well as non-commercial use (Oza and Joshi, 2017).

OpenCV has four main elements, four of which appear in the following figure. The CV portion contains the critical image treatment and computer vision algorithms at the higher levels; ML is the machine learning library that consists of several statistical classification and clustering tools. The HighGUI provides I / O routines and features to store and load video and images, and CXCore comprises the basic data and content structures.

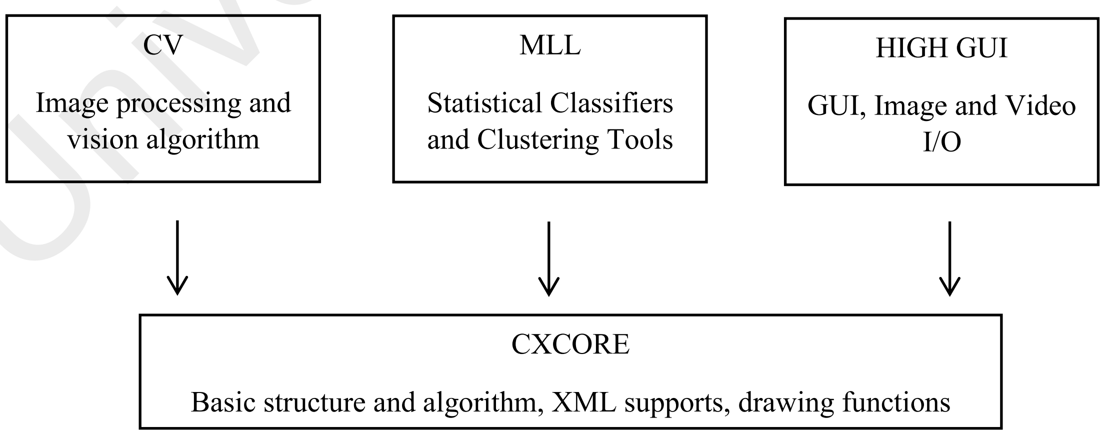


Figure 3.12: Four Main Components of OpenCV (Neelima & Saravanan, 2014).

The OpenCV feature that allows the user to communicate with the operating system, file system, and devices, such as cameras, is contained in a library called HighGUI ("high-level graphical user interface"). HighGUI offers users with the probability of opening windows, viewing images, reading, and writing files related to graphical images and videos, and managing basic mouse, cursor, and keyboard events. The OpenCV HighGUI library can be split into three parts: the hardware portion, the device component, and the GUI component. The hardware component focuses on the activity of cameras mainly. Interacting with a camera is a boring and frustrating job in most operating systems. However, HighGUI can be easily accessed.

OpenCV is an open-source library which provides us with the tools to perform almost any kind of image and video processing. OpenCV video processing algorithms are constructed by writing programming script files. The quality of video should be tested after the algorithms have been used. This results in visual and statistical analyses to test the accuracy of the images processed. In this, video processing algorithms are used and the results of OpenCV are evaluated. OpenCV video processing algorithms have been developed to detect performance of the built algorithms with analyzed statistical parameters. For creation of image and video processing algorithms, OpenCV's computer vision library is used. The proposed methodology of image and video processing algorithms using OpenCV is show in figure below.

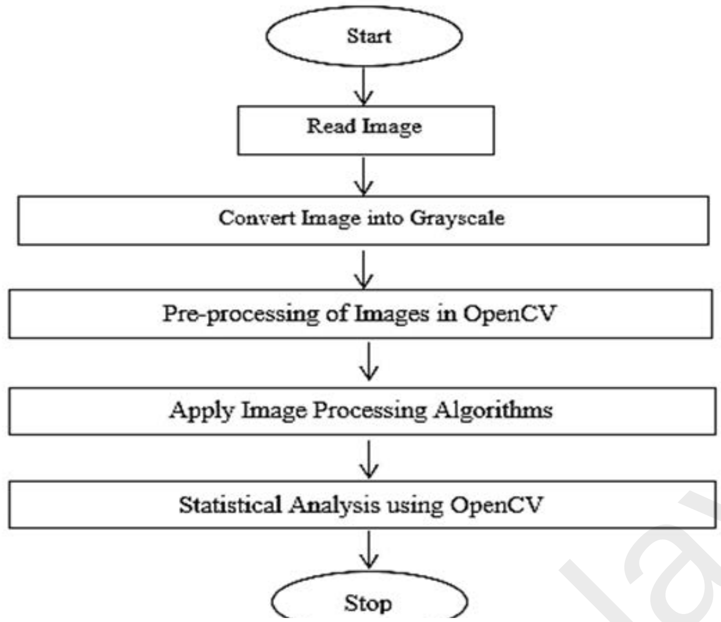


Figure 3.13: Processing Algorithm Using OpenCV (Neelima & Saravanan, 2014).

3.7.2 Python Language

The first language for beginner programmers is Python programming language, because it has powerful tools which represent people's way of thinking and implementation of the code. This also restricts the need to write syntactically appropriate programs to additional keywords. It seems more successful than teaching a lot of C++ or Java languages terms and elements related to language details instead of an algorithm success (Bogdanchikov *et al.*, 2013). It is common practice to build a Python interpreter, which is useful for quick testing and experimentation. Python code also has a wide standard library. Just to select a few random instances, Python ships with a range of XML parsers, csv & zip file readers & authors, libraries that use almost any internet protocol and data form, etc. Python coding also has great support for creating web apps. The language interpreted is Python, so it is easily verifying how the operators or functions operate by using the commandline interpreter. Python interpreter has an integrated support module that can enhance the comprehension of various language aspects.

Figure 3.14: Logo of Python Language.

CHAPTER 4: RESULTS AND DISCUSSION

This chapter describes the results achieved after applying various methods. In this section, the screenshot of the output that involving various motion amplification processing algorithms from Method 1,2,3 and 4 for three different types of videos have been recorded.

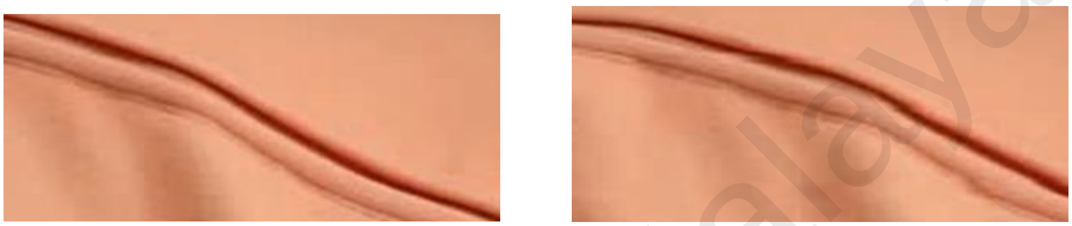
4.1 Results of Sleeping Baby Video

The findings of various studies have demonstrated that it is possible to enhance the subtle movements around the chest of a breathing baby using a variety of methods. When watching the original video, it is difficult to visualize the chest movement caused by the baby's breathing with the naked eye. However, by utilizing VO and MM systems, it is simple to visualize the motion with amplified amplitude and frequency response. The PSDs of the pixel intensity, magnitude, and angular components of the original sleeping baby video are depicted in the figures below, respectively.

4.1.1 Results of Method 1

After creation of mask where no motion occurs in the video, the OF video is filtered with amplification factor of 10 and 100. The figures below have shown the output achieved by multiplying the OF with factor of 10. Based on the output video, there is small amplification movement can be seen due to breathing of baby chest using factor of 10. But when using amplification factor of 100, the amplification of the breathing chest for the output video of the sleeping baby is more significant and it is free from noise. To ensure that the amplification of motion works for magnifying the motion, it is important to make sure to trial and error on the amplification factor to obtain high quality amplified result. However, high amplification factor caused the output video of magnified motion to be blurry.

Figure 4.1: The filtered OF baby video screenshot.



(a) **(b)**

Figure 4.2: Motion amplification of the OF baby video. Amplification factor of (a) 10 and (b) 100.

4.1.2 Results of Method 2

Based on the paper by Wu *et al*. (2012), the PSD analysis is implemented in method 2 to obtain the cut-off frequencies of 0.4 Hz and 3 Hz, which are used in the design of this technique. Because of this, when using the cut-off frequency with an amplification factor of 10, the output video chest vibration that is seen appears to be less exaggerated than when using an amplification factor of 100. While comparing Method 2 to Method 1, it is evident that this method amplifies the vibration more effectively, and the movement of the chest is visible when viewing the resulting video.

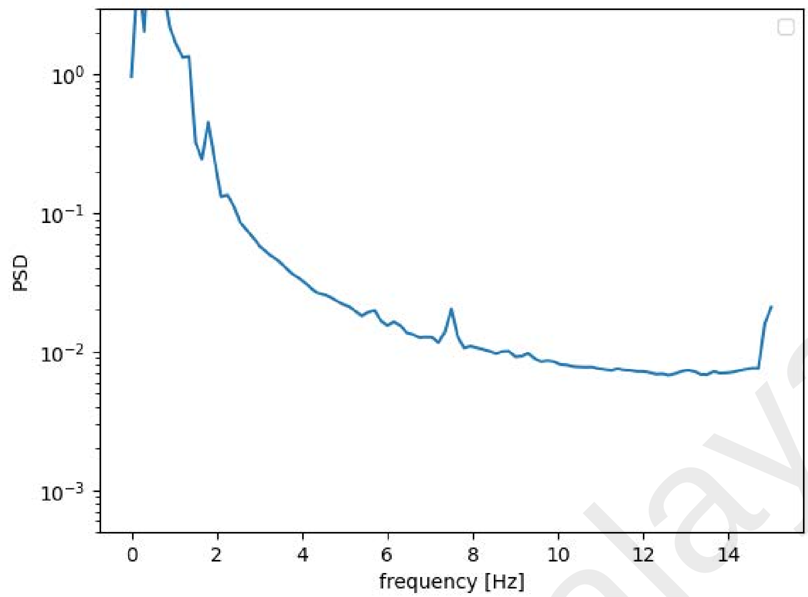


Figure 4.3: PSD of pixel intensity of original baby video.

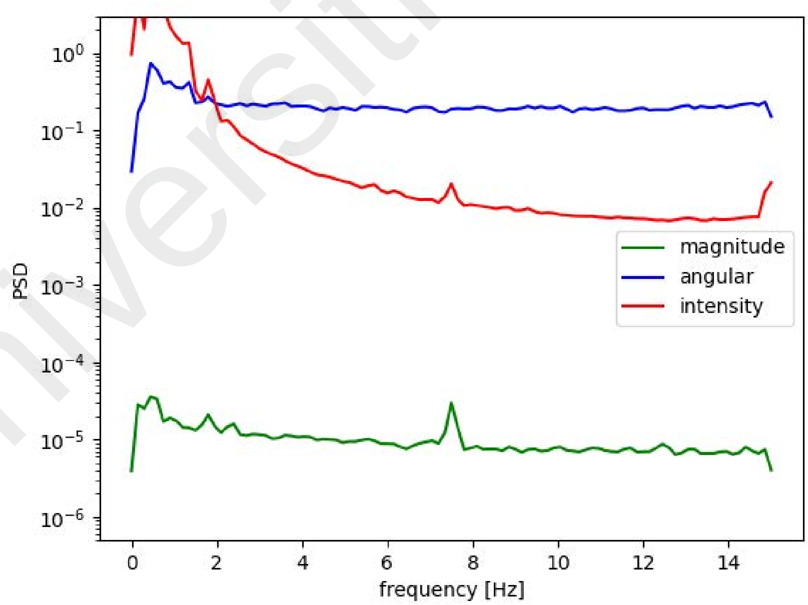


Figure 4.4: PSD of pixel intensity, magnitude, and angular component of original baby video.

(a) **(b)**

Figure 4.5: Screenshot of output baby video for Method 2 using amplification factor of (a)10 (b)100

4.1.3 Results of Method 3

In method 3, based on the value from article Wu *et al.,* 2012, the amplification factor that used in this method is 20 and the cut-of frequency is equal to 0.4 Hz and 3Hz. Based on the video output obtained from Method 3, the video output motion magnification shows a good magnification comparing to Method 1 and 2 which can easily visualize the amplification with bright intensity. However, this method result suffer from some of the noises and artifacts affects around the motion videos.

Figure 4.6: Screenshot of output baby video for Method 3.

4.1.4 Results for Method 4

Figure 4.7: Screenshot of Output Baby Video for Method 4

As evidenced by the bright pixels, Method 4 has shown the motion of the chest, amplifies both signal and noise, and produces artefacts for higher spatial frequencies and bigger motions, which are indicative of this. When compared to the previous Method 3, the magnification result of Method 4 has demonstrated that it is the most effective magnification for visualizing motion.

4.2 Results of Vibrating Guitar String Video

When viewing the original video of the vibrating guitar string, it is possible to see that the first and second strings of the guitar are vibrating with a small amplitude. To obtain a more accurate study of the guitar string's vibration, however, the video amplitude must be magnified using VO and MM systems. The guitar string can be amplified based on its cut-off frequency to magnify the various strings.

4.2.1 Results of Method 1

From the results of Method 1, it can be shown that when using amplification factor of 100, the output videos will show more magnified amplification of guitar string as comparing to using the amplification factor of 50. Each results output is free from noise and artifacts. However, the amplification output of the method using factor of 100 shows blurry on the video.

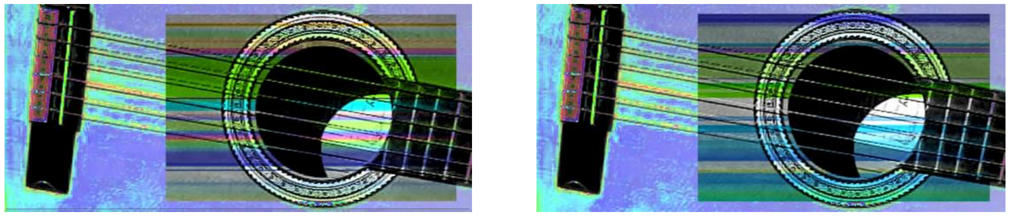


Figure 4.8: The filtered OF guitar video screenshot.

(a) **(b)**

Figure 4.9: Motion amplification of the OF guitar video using amplification factor of (a) 50 and (b) 100.

4.2.2 Results of Method 2

In Method 2, PSD analysis have been made to visualize the cut-off frequency. This method has been implemented on different cut-off frequencies including 72Hz,92Hz which is the cut-off frequency of guitar first string, string E and 100Hz-120Hz which is the cut-off frequency of guitar second string, string A. The methods are being implemented by using amplification factor of 10 in this method. The results obtained for both outputs have amplified the motion, but the visible result is blurry.

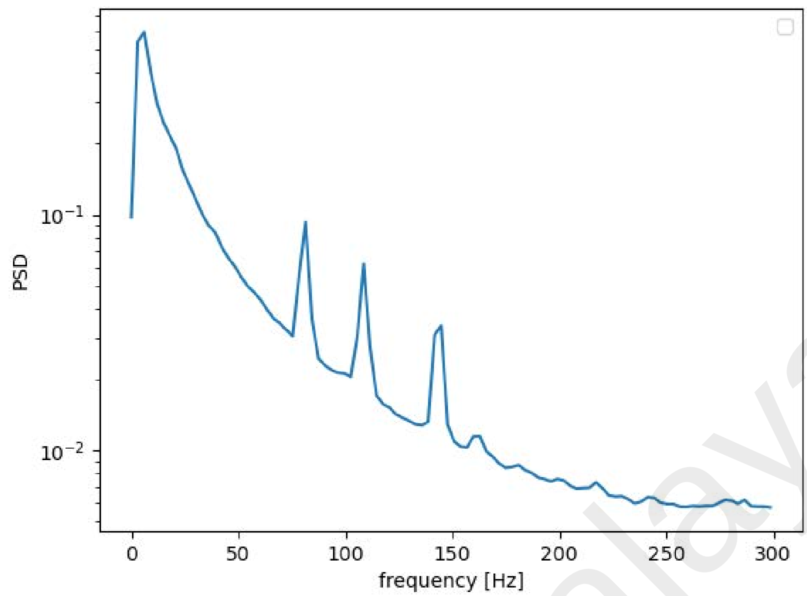


Figure 4.10: PSD of pixel intensity of original guitar video.

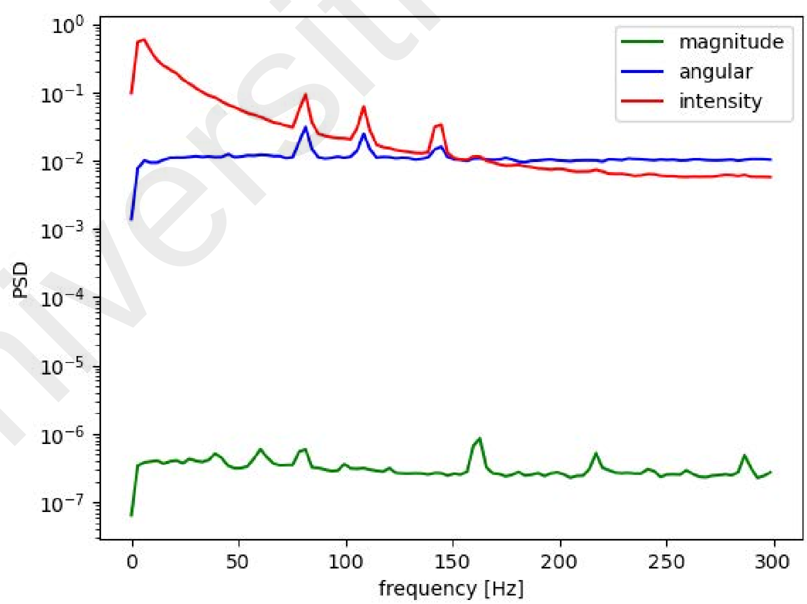
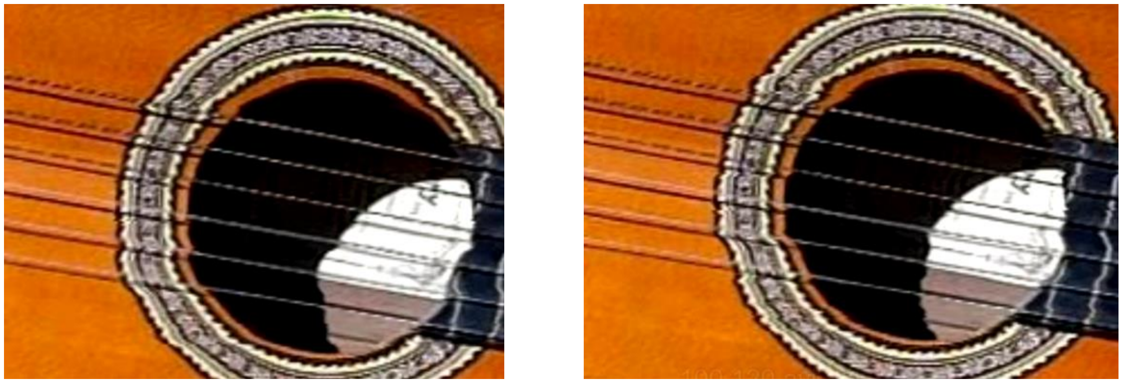


Figure 4.11: PSD of pixel intensity, magnitude, and angular component of original guitar video.



(a) (b)

Figure 4.12: Screenshot of output guitar video for Method 3.

(a) amplification factor =50, cut-off frequency= 72Hz, 92 Hz and (b)amplification factor=100, cut-off frequency =100,120.

4.2.3 Results of Method 3

For the visualization of video output using method 3, it has successfully amplified the motion of vibrating guitar with clear and bright intensities. For Figure 4.13 (a), is it filtering to amplify the first string, string E and significant vibration have been shown in the output. For Figure 4.13(b), the motion has been amplified mainly around the second string, and the results for both outputs have successfully magnified the vibration of string. However, the results produced have more noises compare to previous Method 4. Comparing to Method 1 and 2, it does not produce blurry appearance.

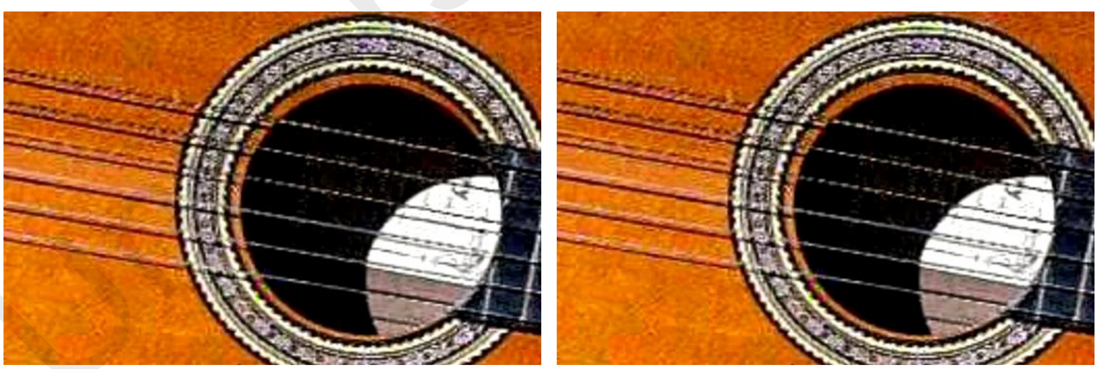
(a) (b)

Figure 4.13: Screenshot of output guitar video for Method 3.

(a) amplification factor =50, cut-off frequency= 72Hz,92 Hz and (b)amplification factor=100, cut-off frequency =100,120.

4.2.4 Results of Method 4

When compared to other methods, the results produced using this method demonstrate that it is carried out with the correct filter, which magnifies the motion and has the best magnifications for visualization.



(a) (b)

Figure 4.13: Screenshot of output guitar video for Method 4. (a) amplification factor=50, cut-off frequency= 72Hz,92 Hz and (b)amplification factor=100, cut-off frequency =100Hz, 120Hz

4.3 Results of Vibrating Phone Video

An experiment involving a vibrating phone video has been designed specifically for this study to serve as a motion assessment system for the machinery. In the case of vibrating phones, it is difficult for a human being to visualize the motion of any vibrating motion presence when looking at the phones or the original video of the vibrating phone. This testing experiment is intended to serve as a demonstration of how to analyses the vibration of machines that are used in factories as part of preventive maintenance program.

4.3.1 Results of Method 1

By using amplification factor of 100, it can be shown that the magnified motion of vibrating phone has been successfully. The video output of the result has shown the motion of the vibrating phone have been amplified without noises whereas when utilizing the amplification factor of 10, the result of vibration is not significant. The results produced is very blurry and showing edge distortions.

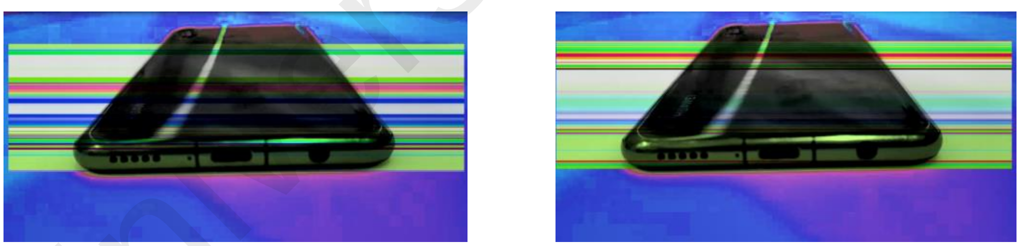
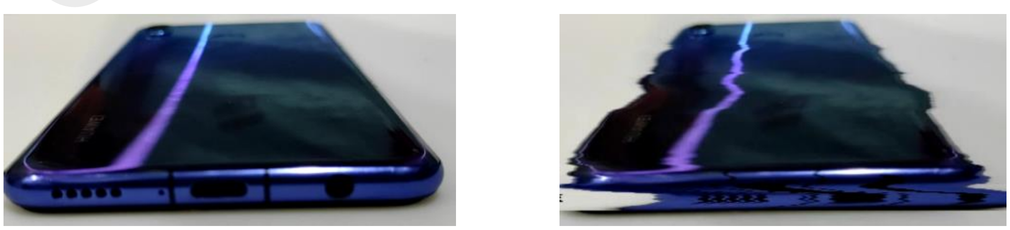


Figure 4.14: The filtered OF phone video screenshot.



(a) **(b)**

Figure 4.15: Motion amplification of the OF Phone Video. Amplification factor of

(a) 20 and (b) 100.

4.3.2 Results of Method 2

In Method 2, the PSD analysis is used to determine the cut-off frequency, which is not being used in method 1. In contrast, the result of the PSD analysis shows that the line is rather flat, which makes it impossible to select a specific cut-off frequency. Instead, a frequency cut-off in the range of 1Hz to 12Hz is employed as the cut-off frequency. By experimenting with an amplification factor of 20 and a cut-off frequency ranging from 1 Hz to 12 Hz, it was discovered that the output video of the vibrating phone had substantially more motion enhanced in it than before. By comparing Method 2 to Method 1, when utilizing an amplification factor of 20, Method 2 exhibits better amplified motion and produces less edge distortions than Method 1. The phones are vibrating and moving back and forth across the frame, which is to be expected from a vibrating phone, as shown in the video.

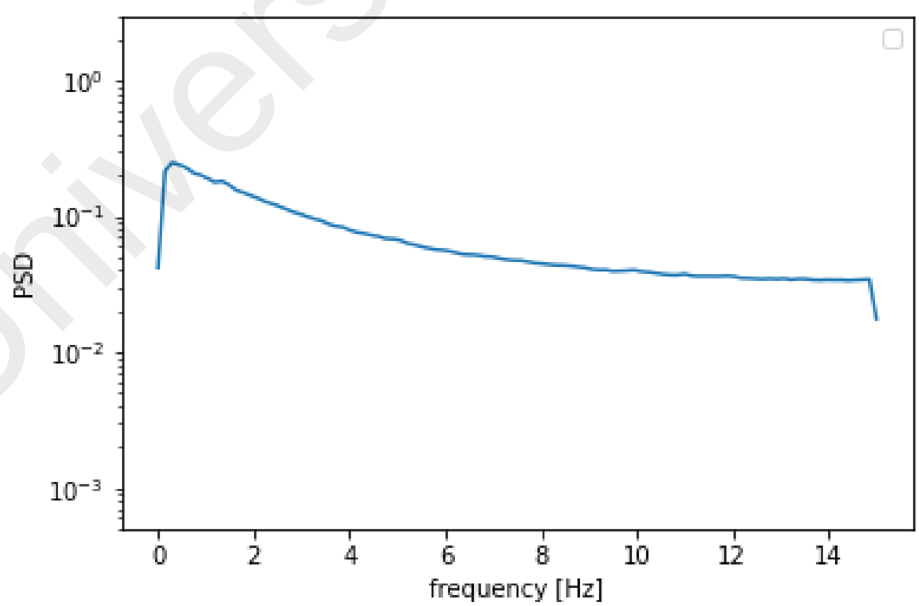


Figure 4.16: PSD of pixel intensity of original phone video.

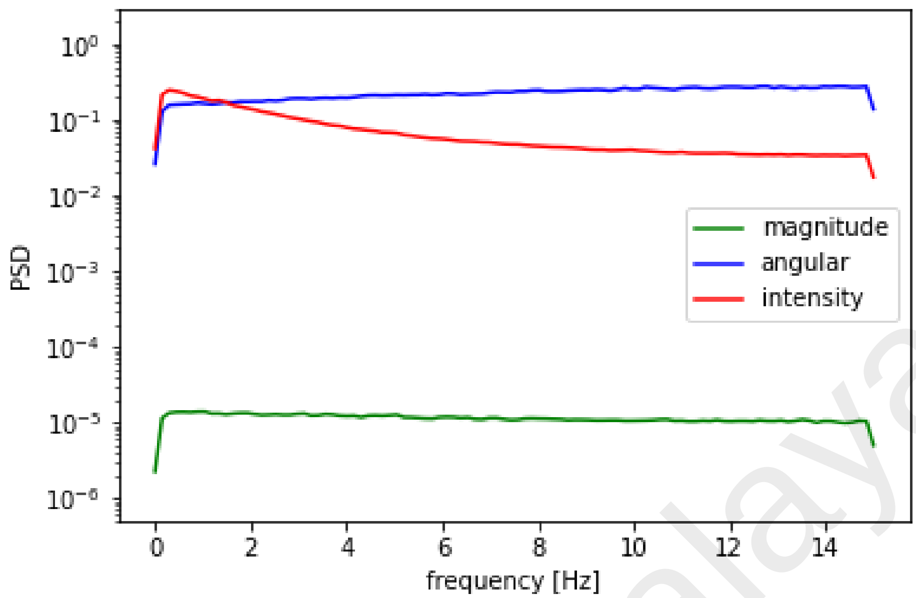


Figure 4.17: PSD of pixel intensity, magnitude, and angular component of original phone video.

Figure 4.18: Screenshot of output phone video for Method 2.

4.3.3 Results of Method 3

The video output has successfully formed artifacts at the boundaries of the phone, which can be visualized clearly when using Method 3. However, the video output has a lot of noise when comparing to other methods. When utilizing an amplification factor of 20 and magnifying with 1Hz and 12Hz frequencies, the motion magnification is more noticeable comparing to Method 1 and 2.

Figure 4.19: Screenshot of output phone video for Method 3.

4.3.4 Results of Method 4

Figure 4.20: Screenshot of output phone video for Method 4.

The output of the result of method 4 for have shown the vibration of phone is being magnified but it is mainly fill with noise and artifacts. To have better result of motion magnified output that can reduces noise, phased-based EVM could be implemented to solve the noises and artifacts issues with complex steerable pyramid.

CHAPTER 5: CONCLUSION AND RECOMMENDATION

5.1 Conclusion

The use of a VO system to detect motion or vibration in a machine is one of the most effective non-contacting methods available for performing vibration analysis on a variety of machines as part of predictive maintenance. The VO system and MM methods have been developed to aid in the amplification of the amplitude of motion, allowing humans to see motion that is otherwise invisible to the naked eye as a non-contact vibration measurement method for machinery motion assessment. The implementation of the video processing framework, which includes OF, PSD, amplification, filtering, and warping, has been completed as non-contacting vibration measurement. The suitable visual odometry technique have been investigated and successfully applied in this project. Four different types of methods of motion amplification algorithms have been implemented in this project, all of which make use of the VO system, which helps to amplify the motion of videos including magnifying the breathing of a sleeping baby, the vibrating guitar string, and the vibrating phone. Based on comparison on the result and data obtained, the most suitable VO for machinery motion assessment system is Method 4 which have perform the best magnification work in this project. For comparing the results obtained from these methods, the EVM method shows the noise characteristics, but it has the better noise characteristics than Method 3. Method 2 is better than method 1 in term on edge distortions, but they cause blurry outcomes at the end of the system.

However, each stage of the process must be carried out with caution and great care to ensure that the progression from loading the input of the video path to the final phase of amplifying and warping would result in producing good, magnified video output. When adjusting the amplification factor and cut off frequency, it is critical to ensure that the value inserted for each video is appropriate and produces the highest quality of amplified and magnified video output for better illustration purposes. Different codes and functions are checked and corrected repeatedly to achieve the best results for motion magnification in videos when developing the algorithm for video processing.

5.2 Future Recommendations

On behalf of successfully implemented this project, there is still some limitation and recommendations that can be implemented after getting the motion amplified and magnified video outputs. Apart from getting comparison with the visualizing output of the methods, quantitative approaches should be implemented to compare the various methods. It is the capability to recognize sub-pixel motion at large magnification factors while remaining noise resistant that distinguishes motion magnification techniques from other types of imaging. The method could be quantitatively analyzing and compare it to the different method based on many parameters to quantify their strengths and understand their limitations. Sub-pixel performance and noise performance with small and large input motion is some of the parameters that can be used to do quantitative comparison on motion magnified videos to let us choose for the best approach for motion magnified algorithm (Oh *et al*., 2018). On the other hand, physical accuracy test for verification of results can be carried out for justification on the result. It can be obtained by implementing the hammer sequence from the other researchers who work on the same video with this research, where accelerometer measurement is available. It can be integrated twice the accelerometer signal and used a zero-phase high-pass filter to remove drifts. The resulting signal of proposed methods should match up well with the other researchers result, that would suggest the methods implemented is physically accurate.

Furthermore, detection of the presence of large motions in an video or image, is not the only potential problem that could affect the final affect's quality of MM. Natural phenomena that occur in conjunction with the recording of an image from a light source or the performance of a sensor may have the effect of causing the appearance of changes that are likely to changes that are important in terms of the condition or behavior of the object under observation in terms of their real information. It is the mixing of meaningful and non-meaningful minor changes that results in the appearance of considerable noise in the magnified image. A method based that known as factional anisotropy (FA) was proposed in addition to the well-known techniques for reducing this effect, which are related to manual intervention into the processing process (Takeda *et al.*, 2019). This type of filter, which is commonly used in neuroscience, was used in the cited study to design a filter that eliminated non-meaningful differences. On the observation of temporal distribution of changes, it is hypothesized that FA can be used to show an anisotropic nature of diffusion for meaningful changes, whereas FA cannot be used for non-meaningful changes.

5.3 Sustainability

Using motion amplification, it is possible to see complicated vibration problems that are otherwise inaccessible to the naked eye. When used in conjunction with other tools, this instrument can save time and money in the areas of vibration analysis, routine condition monitoring program troubleshooting, and root cause analysis. Precision maintenance has a significant impact on the environmental sustainability of an organization, primarily through the improvement in product quality, which allows for a reduction in the consumption of raw materials. Predictive maintenance has a significant impact on the environmental sustainability of an organization. When it comes to reciprocating and rotary machines, vibration analysis is a predictive technique that is particularly well adapted to the task. In this technique, the vibration level of the machine is recorded on a periodic basis; the amount of vibration increases when abnormalities such as misalignment, unbalance, and so on are present; this technique is particularly well-suited for reciprocating and rotary machines. When it comes to the manufacture of bearings for vehicles, one of the quality parameters that is evaluated is the final vibration of the assembled bearing. This is done to ensure that the bearings have the desired quality. Vibrations appear during external grinding processes that are caused by the process itself rather than by flaws in the machine tool. These vibrations have the potential to cause defects that lower the overall quality of the workpieces produced. Despite this, the analysis of process-induced vibrations has received little attention in the scientific literature. Vibration spectra are used to distinguish between vibration frequencies that are problematic and those that are not problematic. In this case, intercomparisons of spectra between similar machines, as well as studies of the change in spectra over time, such as waterfall plots, are used. When designing components, this method takes into consideration the possibility of establishing separate vibration limits to ensure the quality of the finished product after it is assembled. By increasing the efficiency with which resources are used, the methodology presented can help to improve the environmental sustainability of an industrial organization's operations. Consequently, the machine's lifespan and efficiency are both improved. Reduced breakdown frequency and averted large-scale equipment rework can be achieved through the replacement of small parts.

5.4 Complexity

After completed this project, there were few complexities found when carried out the project from start to the end. To ensure that the VO system and MM systems runs smoothly, the theoretical knowledge and functions of each flow of process must be thoroughly understood before the project can be implemented. When implementing the video processing algorithm, there are several coding and functions that must be applied. Connecting code for each type of video processing flow was different. There were no specific instructions in previous research on how to construct the coding for connecting the OF with MM methods, so a self-learning practice is required to create the code and run it without error. For the output to be accurate and correct, the coding of each set of video processing must be written in the appropriate values and commands.

5.5 Lifelong learning

This project may serve as a VO system and MM method on machinery motion assessment system in the sense of lifelong learning in this project, which as an important milestone in performing the vibration analysis as the preventive maintenance in industry field. The development of non-contacting method vibration analysis will minimize the cost utilities of the company, increase production efficiencies, and increase the safety of the employees. The successful management of a facility relies heavily on a preventive maintenance strategy. Having a preventive maintenance program in working place minimizes the machine downtime. As an engineer, it is important for us to know if the facility at the company do not operating at optimal levels is a loss to the company because it is impossible to predict when equipment may break, reactive maintenance results in unplanned downtime, which results in idle personnel, a stoppage in manufacturing, and production delays. These unanticipated problems will almost always need paying a technician overtime and overnighting replacement components in order to get the firm back up and operating as quickly as feasible. To minimize the impact on day-to-day production, time and money can be saved in the long run if take proper care of the equipment and keep it running at peak efficiency. Every piece of equipment eventually wears out. Additionally, performing routine preventative maintenance can help to extend the life of your machines. Routine maintenance, such as the replacement of parts, the replacement of fluids and oils, and the performance of quality checks, should not be neglected.

REFERENCES

ADIV, G. (1985). Determining Three-Dimensional Motion and Structure from Optical Flow Generated by Several Moving Objects. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 384 - 401.

Ahmadi, H., & Salami, P. (2010). Using of Power Spectral Density for Condition Monitoring of Fan . *Modern Applied Science*, 54-59.

Ahmed, H., & Nandi, A. (2019). Frequency Domain Analysis. In *Condition Monitoring with Vibration Signals (Compressive Sampling and Learning Algorithms for Rotating Machine)* (pp. 63-77). London: John Wiley & Sons Ltd.

Ahmed, H., & Nandi, A. (2019). Time-Frequency Domain Analysis. In *Condition Monitoring with Vibration Signals (Compressive Sampling and Learning Algorithms for Rotating Machine)* (pp. 79-114). London: John Wiley & Sons Ltd.

Al Zadjali, A., & Rameshkumar, G. (2013). Condition Monitoring of Centrifugal Blower Using Vibration Analysis. *INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY SCIENCES AND ENGINEERING*, 50-59.

Alsalaet, J. (2012). *Vibration Analysis and Diagnostic Guide.* Iraq.

Ambaye, G. A. (2020). Time and Frequency Domain Analysis of Signals: A Review. *International Journal of Engineering Research & Technology (IJERT)*, 271-276.

Aqel, M., Marhaban, M., Saripan, M., & Ismail, N. (2016). Review of visual odometry: types, approaches, challenges, and applications. *SpringerPlus*, 1-26.

Balakrishnan, G., Durand, F., & Guttag, J. (2013). Detecting Pulse from Head Motions in Video. *Computer Vision Foundation*, 3430-3437.

Basic Accelerometer . (2015, September 11). Retrieved from INTERTECHNOLGY INC.: http://www.intertechnology.com/Kistler/indexAcceleration.htm.

Basic Accelerometer. (2015, September 11). Retrieved from Industrial Electronics: http://www.industrial-electronics.com/DAQ/industrial_electronics/input_devices_sensors_transducers_transmitters_measurement/Accelerometers.html.

Bastianini, F., Matta, F., Rizzo, A., Galati, N., & Nanni, A. (2007). Overview of Recent Bridge Monitoring Applications using Distributed Brillouin Fiber Optic Sensors. *Journal Nondestruct Test*, 269-276.

Black, M., & Anandan, P. (1993). A framework for the robust estimation of optical flow. *1993 (4th) International Conference on Computer Vision* (pp. 231-236). Berlin, Germany: IEEE.

Boer, J., & Kalksma, M. (2015). Choosing between optical flow algorithms for UAV position change measurement.

Brandt, A. (2011). *NOISE AND VIBRATION ANALYSIS: SIGNAL ANALYSIS AND EXPERIMENTAL PROCEDURES.* Southern Denmark: John Wiley & Sons, Ltd.

BURT, P., & ADELSON, E. (1983). The Laplacian Pyramid as a Compact Image Code. *IEEE TRANSACTIONS ON COMMUNICATIONS*, 532-540.

Caetano, E., Silva, S., & Bateira, J. (2010). A VISION SYSTEM FOR VIBRATION MONITORING OF CIVIL ENGINEERING STRUCTURES. *Experimental Techniques*, 74-82.

Cakar, O., & Sanliturk, K. (2005). Elimination of transducer mass loading effects from frequency response functions. *Mechanical Systems and Signal Processing*, 87-104.

Cha, Y.-J., Chen, J., & Büyüköztürk, O. (2017). Output-only computer vision based damage detection using phase-based optical flow and unscented Kalman filters. *Engineering Structures*, 300-313.

Chen, J., Wadhwa, N., Cha, Y.-J., Durand, F., Freeman, W., & Buyukozturk, O. (2015). Modal identification of simple structures with high-speed video using motion magnification. *Journal of Sound and Vibration*, 58-71.

Chen, J., Wadhwa, N., Cha, Y.-J., Durand, F., Freeman, W., & Buyukozturk, O. (2015). Modal identification of simple structures with high-speed video using motion magnification. *Journal of Sound and Vibration*, 58-71.

Chen, X., & Chang, C.-C. (2019). In-Plane Movement Measurement Technique Using Digital Sampling Moire Method. *J. Bridge Eng.*, 04019013-1 - 04019013-12.

Comport, A., Malis, E., & Rives, P. (2007). Accurate Quadrifocal Tracking for Robust 3D Visual Odometry . *Proceedings 2007 IEEE International Conference on Robotics and Automation* (pp. 40-45). Rome, Italy: IEEE.

Comport, A., Malis, E., & Rives, P. (2010). Real-time Quadrifocal Visual Odometry. *The International Journal of Robotics Research*, 245-266.

Cryer, J., & Chan, K.-S. (2008). Introduction To Spectral Analysis. In *Time Series Analysis* (pp. 319-350). New York: Springer, New York, NY.

Davis, A., Bouman, K., Chen, J., Rubinstein, M., Buyukozturk, O., Durand, F., & Freeman, W. (2017). Visual Vibrometry: Estimating Material Properties from Small Motions in Video. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 732-745.

Denman, S., Chandran, V., & Sridharan, S. (2007). An adaptive optical flow technique for person tracking systems. *Pattern Recognition Letters*, 1232-1239.

Doguer, T., & Strackeljan, J. (2009). Vibration Analysis using Time Domain Methods for the Detection of small Roller Bearing Defects. *SIRM 2009 - 8th International Conference on Vibrations in Rotating Machines, Vienna, Austria*, (pp. 1-12). Germany.

Dong, C.-Z., & Catbas, F. (2020). A review of computer vision-based structural health monitoring at local and global levels . *Structural Health Monitoring*, 1-106.

Esmaeilian, B., Wang, B., Lewis, K., Duarte, F., Ratti, C., & Behdad, S. (2018). The future of waste management in smart and sustainable cities: A review and concept paper. *Waste Management*, 177-195.

Farnebäck, G. (2003). Two-Frame Motion Estimation Based on Polynomial Expansion. *Scandinavian Conference on Image Analysis* (pp. 363-370). Springer, Berlin, Heidelberg.

Freeman, W., & Adelson, E. (1991). The Design and Use of Streerable Filters. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 891-906.

Frontoni, E. (2006). *Vision Based Mobile Robotics: Mobile Robot Localization Using Vision Sensors And Active Probabilistic Approaches.* Ancona, Italy: lulu.com.

Fuchs, M., Chen, T., Wang, O., Raskar, R., Seidel, H.-P., & Lensch, H. (2010). Real-time temporal shaping of high-speed video streams. *Computers & Graphics*, 575-584.

Fukuda, Y., Feng, M., & Shinozuka, M. (2010). Cost-effective vision-based system for monitoring dynamic response of civil engineering structures. *STRUCTURAL CONTROL AND HEALTH MONITORING*, 918-936.

Garcıa, D., Rojo, L., Aparicio, A., Castello, L., & Garcia, O. (2012). Visual Odometry through Appearance- and Feature-Based Method with Omnidirectional Images. *Journal of Robotics*, 1-13.

Gentile, C. (2010). Deflection measurement on vibrating stay cables by non-contact microwave interferometer. *NDT&E International*, 231-240.

Goreczka, S., & Strackeljan, J. (n.d.). Optimization of time domain features for rolling bearing fault diagnostics . *The Sixth International Conference on Condition*

Monitoring and Machinery Failure Prevention Technologies, (pp. 642-652). Germany.

Guan, B., Wang, S., & Wang, G. (2005). A biologically inspired method for estimating 2D high-speed translational motion. *Pattern Recognition Letters*, 2450-2462.

Guizilini, V., & Ramos, F. (2011). Visual odometry learning for unmanned aerial vehicles. *2011 IEEE International Conference on Robotics and Automation* (pp. 6213-6220). Shanghai, China: IEEE.

Guo, A. (2014). A Vibration Sensor Design Research. *Sensors & Transducers*, 228-234.

Guo, Hong; Jack, Lindsay B.; Nandi, Asoke K.; Senior Member, IEEE. (2005). Feature Generation Using Genetic Programming With Application to Fault Classification. *IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART B: CYBERNETICS, VOL. 35, NO. 1*, 89-99.

Holm-Hansen, B., & Gao, R. (2000). Vibration Analysis of a Sensor-Integrated Ball Bearing. *Transactions of the ASME*, 384-392.

Horn, B., & Schunck, B. (1981). Determining optical flow. *Artificial Intelligence*, 185-203.

Howard, A. (2008). Real-Time Stereo Visual Odometry for Autonomous Ground Vehicles. *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 3946-3952). Nice, France: IEEE.

Hyatt, J., Edwards, S., & Lee, M. (2019). *Motion Amplification in Video Using Optical Flow.* United States.

Irani, M., & Anandan, P. (1998). Robust multi-sensor image alignment. *Sixth International Conference on Computer Vision* (pp. 959-966). Bombay, India: IEEE.

Jamal, M., & Rasheed, K. (2021). Vibration Measurement of a Rotating Shaft using Electrostatic Sensor. *International Journal of Recent Technology and Engineering (IJRTE)*, 97-105.

Jamia, N., Friswell, M., El-Borg, S., & Fernandes, R. (2018). Simulating eddy current sensor outputs for blade tip timing. *Advances in Mechanical Engineering*, 1-12.

Johansmann, M., Siegmund, G., & Pineda, M. (2005). Targeting the Limits of Laser Doppler Vibrometry . 1-12.

K.M., A., & MANJUNATH, T. (2017). Vibration Signal Analysis using Time and Timefrequency Domain: Review. *IEEE International Conference on Power, Control, Signals and Instrumentation Engineering*, 1808-1811.

Kim, E., Tan, A., Yang, B.-S., & Kosse, V. (2007). Experimental Study on Condition Monitoring of Low Speed Bearings : Time Domain Analysis. *Proceedings of the Fifth Australasian Congress of Applied Mechanics* (pp. 108-113). Australia: Engineers Australia.

Kolb, A., Barth, E., Koch, R., & Larsen, R. (2009). Time-of-Flight Sensors in Computer Graphics. *EUROGRAPHICS*, 119-134.

Laser Deflection method. (2015, September 11). Retrieved from The City College of New York: http://www.ccny.cuny.edu/~zhu/LDV/FinalReportsHTML/CCNY-LDVTech-Report-html.htm.

Lazecky, M., Hlavacova, I., Bakon, M., Sousa, J., Perissin, D., & Patricio, G. (2016). Bridge Displacements Monitoring Using Space-Borne X-Band SAR Interferometry. *IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING*, 205-210.

LDV Sensor. (2015, September 11). Retrieved from ACOUTRONIC: http://acoutronic.se/images/vibration/Polytec_picture_RSV-150.png

Li , C., Cabrena, D., Oliveira, J., Sanchez, R.-V., Cerrada, M., & Zurita, G. (2016). Extracting repetitive transients for rotating machinery diagnosis using multiscale clustered grey infogram. *Mechanical Systems and Signal Processing*, 157-173.

Li, C., Cabrera, D., Oliveira, J., Sanchez, R.-V., Cerrada, M., & Zurita, G. (2016). Extracting repetitive transients for rotating machinery diagnosis using multiscale clustered grey infogram. *Mechanical Systems and Signal Processing* , 157-173.

Li, C., Sanchez, R.-V., Zurita, G., Cerrada, M., Cabrera, D., & Vásquez, R. (2015). Multimodal deep support vector classification with homologous features and its application to gearbox fault diagnosis. *Neurocomputing*, 119-127.

Li, J., Mechitov, K., Kim, R., & Spencer, Jr., B. (2016). Efficient time synchronization for structural health monitoring using wireless smart sensor networks. *STRUCTURAL CONTROL AND HEALTH MONITORING*, 470-486.

Lim, M.-S., & Lim, J. (2008). Visual measurement of pile movements for the foundation work using a high-speed line-scan camera. *Pattern Recognition*, 2025-2033.

Liu, C., Torralba, A., Freeman, W., Durand, F., & Adelson, E. (2005). Motion Magnification. 519-526.

Liu, Y., Lu, Y., Shi, Q., & Ding, J. (2013). Optical Flow Based Urban Road Vehicle Tracking. *2013 Ninth International Conference on Computational Intelligence and Security* (pp. 391-395). Emeishan, China: IEEE.

Liu, Z., Yeh, R., Tang, X., Liu, Y., & Agarwala, A. (2017). Video Frame Synthesis using Deep Voxel Flow. *2017 IEEE International Conference on Computer Vision*, 4473-4481.

Lobo-Aguilar, S., Jiang, Z., Zhang, Z., & Christenson, R. (2019). Infrasound-Based Noncontact Sensing for Bridge Structural Health Monitoring. *Journal of Bridge Engineering*.

Lucas, B., & Kanade, T. (1981). An Iterative Image Registration Technique with an Application to Stereo Vision. *Proceedings of Imaging Understanding Workshop*, 121-130.

McCarthy, C., & Barnes, N. (2004). Performance of optical flow techniques for indoor navigation with a mobile robot. *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004* (pp. 5093-5098). New Orleans, LA, USA: IEEE.

Mitchell, J. S. (2007, January). *SOUND AND VIBRATION.* Retrieved from From Vibration Measurements to Condition Based Maintenance Seventy Years of Continuous Progress: http://www.sandv.com/

Mobley, R. K. (2002). *An Introduction to Predictive Maintenance (Second Edition).* USA: Butterworth-Heinemann.

Moschas, F., & Stiros, S. (2011). Measurement of the dynamic displacements and of the modal frequencies of a short-span pedestrian bridge using GPS and an accelerometer. *Engineering Structures*, 10-17.

Muller, P., & Savakis, A. (2017). Flowdometry: An Optical Flow and Deep Learning Based Approach to Visual Odometry. *IEEE Winter Conference on Applications of Computer Vision* (pp. 624-631). Rochester Institute of Technology.

Narasimhan, S., & Wang, Y. (2020). Noncontact Sensing Technologies for Bridge Structural Health Assessment. *J. Bridge Eng.*, 02020001-1 - 02020001-2.

Neelima, K., & Saravanan, D. (2014). Image Detection and Count Using Open Computer Vision (Opencv). *Journal of Engineering Research and Applications*, 107-109.

Oh, T.-H., Jaroensri, R., Kim, C., Elgharib, M., Durand, F., Freeman, W., & Matusik, W. (2018). Learning-based Video Motion Magnification. *COmputer Vision Foundation*, 1-16.

Oza, S., & Joshi, D. (2017). Using OPENCV over MATLAB for Implementing Image Processing Application on CUDA GPU to Achieve Better Execution Speedup. *International Journal of Engineering Research & Technology (IJERT)*, 620-623.

Parhi, K., & Ayinala, M. (2014). Low-Complexity Welch Power Spectral Density Computation. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 172-182.

Patsias, S., & Staszewski, W. (2002). Damage Detection Using Optical Measurements and Wavelets. *Structural Health Monitoring*, 5-22.

Photo of different sizes of Eddy Current type proximity sensors. (2015, September 11). Retrieved from GlobalSpec: http://www.globalspec.com/learnmore/sensors_transducers_detectors/proximity_presence_sensing/eddy_current_proximity_sensors.

Poddar, S., Kottath, R., & Karar, V. (2018). Evolution of Visual Odometry Techniques. 1-13.

Proximity Sensor. (2015, September 11). Retrieved from Lion Precision: http://www.lionprecision.com/eddy-current-sensors/

Ramirez, A., Abdel-Rahaman, M., & Noda, T. (2017). Frequency-domain modeling of time-periodic switched electrical networks: A review. *Ain Shams Engineering Journal*, 1-7.

Rao, S. S. (2011). *Mechanical Vibrations, 5th Edition.* United States: Pearson.

Rasool, A., Javaid, H., Babar, T., & Saghir, R. (2013). VIDEO COLOUR VARIATION DETECTION AND MOTION MAGNIFICATION TO OBSERVE SUBTLE CHANGES. 1-57.

Ri, S., Muramatsu, T., Saka, M., Nanbara, K., & Kobayashi, D. (2012). Accuracy of the Sampling Moiré Method and its Application to Deflection Measurements of Large-Scale Structures. *Experimental Mechanics*, 331-340.

Rothberg, S., Allen, M., Castellini, P., Maio, D., Dirckx, J., Ewins, D., . . . Vignola, J. (2016). An international review of laser Doppler vibrometry: Making light work of vibration measurement. *Optics and Lasers in Engineering*, 1-12.

Rui, W., Yan, Z., Yujun, T., & Yan, Y. (2011). A spatio-temporal filtering method for motion estimation. *2011 6th International Conference on Computer Science & Education (ICCSE)* (pp. 830-834). Singapore: IEEE.

Saied, R., Mostafa, M., & Hussein, H. (2015). Predictive Maintenance Program Based on Vibration Monitoring . *Design and Modeling of Mechanical Systems - II*, 651-660.

Saruhan, H., Sarıdemir, S., Çiçek, A., & Uygur, İ. (2014). Vibration Analysis of Rolling Element Bearings Defects. *Journal of Applied Research and Technology*, 384-395.

Scaramuzza, D., & Fraundorfer, F. (2011). Part I: The First 30 Years and Fundamentals. In *Visual Odometry* (pp. 80-92). IEEE ROBOTICS & AUTOMATION MAGAZINE.

Scaramuzza, D., Fraundorfer, F., Pollefey, M., & Siegwart, R. (2009). Absolute Scale in Structure from Motion from a Single Vehicle Mounted Camera by Exploiting Nonholonomic Constraints. *2009 IEEE 12th International Conference on Computer Vision (ICCV)* (pp. 1413-1419). Kyoto, Japan: IEEE.

Schreiner, M., & Strlič, M. (2008). *Handbook On The Use Of Laser In Conservation And Conservation Science.* Brussels, Belgium: COST office.

Schwarz, B., Richardson, S., Tyler, J., Spears, R., & Richardson, M. (2020, February 10-13). *Post-Processing ODS Data from a Vibration Video.* Retrieved from VIBRANT: http://papers.vibetech.com/PPT_PDFs/Paper78.pdf

Schwenke, H., Neuschaefer-Rube, U., Pfeifer, T., & Kunzmann, H. (2002). Optical Methods for Dimensional Metrology in Production Engineering. *CIRP Annals*, 685-699.

Senapaty, G., & U, S. (2018). Vibration based condition monitoring of rotating machinery. *MATEC Web of Conferences*, 1-18.

Shahadi, H., Albattat, H., Al-allaq, Z., & Thahab, A. (2020). Eulerian video magnification: a review. *Indonesian Journal of Electrical Engineering and Computer Science*, 799-811.

Shao, H., Zhao, Q., & Tu, D. (2014). Head Behavior Detection Method Based on Optical Flow Theory. *2014 International Conference on Intelligent Environments* (pp. 102-106). Shanghai, China: IEEE.

Sheng, S., & Veers, P. (2011). Wind Turbine Drivetrain Condition Monitoring – An Overview. *National Renewable Energy Laboratory*, 1-19.

Siong, L., Mokri, S., Hussain, A., Ibrahim, N., & Mustafa, M. (2009). Motion detection using Lucas Kanade algorithm and application enhancement. *2009 International Conference on Electrical Engineering and Informatics* (pp. 537-542). Bangi, Malaysia: IEEE.

Smieja, M., Mamala, J., Praznowski, K., Cieplinski, T., & Szumilas, L. (2021). Motion Magnification of Vibration Image in Estimation of Technical Object Condition-Review. *Sensors*, 1-21.

Song, X., Seneviratne, L., & Althoefer, K. (2011). A Kalman Filter-Integrated Optical Flow Method for Velocity Sensing of Mobile Robots. *IEEE/ASME TRANSACTIONS ON MECHATRONICS*, 551-563.

Souhila, K., & Karim, A. (2007). Optical Flow based robot obstacle avoidance. *International Journal of Advanced Robotic Systems*, 13-16.

Spencer Jr., B., Horskere, V., & Narazaki, Y. (2019). Advances in Computer Vision-Based Civil Infrastructure Inspection and Monitoring. *Engineering*, 199-222.

Sun, D., Roth, S., Lewis, J., & Black, M. (2008). Learning Optical Flow. *European Conference on Computer Vision* (pp. 83-97). Springer, Berlin, Heidelberg.

Takeda, S., Akagi, Y., Okami, K., Isogai, M., & Kimata, H. (2019). Video Magnification in the Wild Using Fractional Anisotropy in Temporal Distribution. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 1614-1622). Long Beach, CA, USA: IEEE.

Teyssieux, D., Euphrasie, S., & Cretin, B. (2011). MEMS in-plane motion/vibration measurement system based CCD camera. *Measurement*, 2205-2216.

Thomas, B., John, S., & Abe, S. (2015). Power Spectral Density Computation using Modified Welch Method. *IJSTE - International Journal of Science Technology & Engineering*, 145-152.

Thomson, W. (2003). *Theory of Vibration with Appiications (Fourth Edition)*. New York: Taylor & Francis.

Tomasi, C., & Zhang, J. (1995). Is Structure-From-Motion Worth Pursuing? *Seventh International Symposium on Robotics Research*, 391-400.

Vishwakarmaa, M., Purohit, R., Harshlata, V., & Rajput, P. (2017). Vibration Analysis & Condition Monitoring for Rotating Machines: A Review. *5th International Conference of Materials Processing and Characterization (ICMPC 2016)* (pp. 2659-2664). India: Elsevier Ltd.

Wadhwa, N., Chen, J., Sellon, J., Wei, D., Rubinstein, M., Ghaffari, R., . . . Freeman, W. (2017). Motion microscopy for visualizing and quantifying small motions. 11639-11644.

Wadhwa, N., Rubinstein, M., Durand, F., & Freeman, W. (2013). Phase-Based Video Motion Processing. *ACM Transactions on Graphics*, 80:1-80:10.

Wadhwa, N., Rubinstein, M., Durand, F., & Freeman, W. (2013). Phase-Based Video Motion Processing. *ACM Transactions on Graphics*, 80:1-80:10.

Wadhwa, N., Rubinstein, M., Durand, F., & Freeman, W. (2014). Riesz pyramids for fast phase-based video magnification. *2014 IEEE International Conference on Computational Photography (ICCP)* (pp. 1-10). Santa Clara, CA, USA: IEEE.

Wadhwa, N., Wu, H.-Y., Davis, A., & Rubinstein, M. (2017). Eulerian video magnification and analysis. *Communications of the ACM*, 87-95.

Wadhwaa, N., Chen, J., Sellon, J., Wei, D., Rubinstein, M., Ghaffari, R., . . . Freeman, W. (2017). Motion microscopy for visualizing and quantifying small motions. *PNAS*, 11639-11644.

Wang, H., Vasseur, C., Christov, N., & Koncar, V. (2012). Vision servoing of robot systems using piecewise continuous controllers and observers. *Mechanical Systems and Signal Processing*, 132-141.

WANG, T.-C., ZHU, J.-Y., KALANTARI, N., EFROS, A., & RAMAMOORTHI, R. (2017). Light Field Video Capture Using a Learning-Based Hybrid Imaging System. *ACM Transactions on Graphics*, 1-13.

Wang, W., & Mottershead, J. (2013). Adaptive moment descriptors for full-field strain and displacement measurements. *The Journal of Strain Analysis for Engineering Design*, 16-35.

WÓJCIK, W., GROMASZEK, K., SHEGEBAYEVA, Z., SULEIMENOV, B., & ARON, B. (2014). Optimal Control for Combustion Process. *PRZEGLĄD ELEKTROTECHNICZNY*, 157-160.

Wu, H.-Y., Rubinstein, M., Shih, E., Guttag, J., Durand, F., & Freeman, W. (2012). Eulerian Video Magnification for Revealing Subtle Changes in the World. *ACM Transactions on Graphics*, 65:1-65:8.

Wu, X., Chen, Z., Gu, Y., Chen, W., & Fang, M.-e. (2018). Illustrative visualization of time-varying features in spatio-temporal data. *Journal of Visual Languages & Computing*, 157-168.

Xiang, L., Yang, S., & Gan, C. (2012). Torsional vibration measurements on rotating shaft system using laser doppler vibrometer. *Optics and Lasers in Engineering*, 1596-1601.

YAMANA, F., SUGIYAMA, T., & ABE, K. (2000). Edge plane detection in spatio-temporal images by using edge vector and edge reliability. *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000* (pp. 660-663). Barcelona, Spain: IEEE.

Yan, R., & Gao, R. (2009). Multi-scale enveloping spectrogram for vibration analysis. *Tribology International*, 293-302.

Yang, Y.-S. (2019). Measurement of Dynamic Responses from Large Structural Tests by Analyzing Non-Synchronized Videos. *Sensor 2019*, 1-24.

Ye, X., Dong, C., & Liu, T. (2016). A Review of Machine Vision-Based Structural Health Monitoring: Methodologies and Applications. *Journal of Sensors*, 1-10.

Yu, Q., Yin, A., Zhang, Q., & Ma, S. (2017). Optical flow tracking method for vibration identification of out-of-plane vision. *Journal of Vibroengineering*, 2363-2374.

Zhu, J., Lu, Z., & Zhang, C. (2020). A marker-free method for structural dynamic displacement measurement based on optical flow. *Structure and Infrastructure*, 1-13.

Zhu, J., Yoon, J., He, D., Qu, Y., & Bechhoefer, E. (2013). Lubrication Oil Condition Monitoring and Remaining Useful Life Prediction with Particle Filtering . *International Journal of Prognostics and Health Management*, 1-15.

Zona, A. (2020). Vision-Based Vibration Monitoring of Structures and Infrastructures: An Overview of Recent Applications. *Infrastructures*, 1-22.