Segmentation of Mechanical Coupling Using Phased Based Motion Amplification of Sub-Pixel Variation

Yi Cheng Zhu, Cheng Xu

1 Introduction

In industrial settings, oftentimes issues can manifest symptoms of low amplitude vibrations. For example, a misaligned bearing can cause such movements. These vibrations are known to cause long-term failures like loose fasteners and therefore needs to be caught early. The current standard method of detection for these vibrations involves accelerometers, but it is not practical to deploy such sensors all over the infrastructures. Thus, it would be extremely beneficial to be able to diagnose these vibrations directly from a video of the subject. These vibrations are normally too small to see with the naked eye. However, they are still captured in the subtle sub-pixel intensity variations. Previous works have demonstrated good results from using local phase-based Eulerian motion amplifications.

In a mechanical system, the vibration of components can be physically linked to a malformed or mispositioned part. Inadvertent coupling of two components will cause both components to synchronize their vibrations. Therefore, such vibrations can propagate along a mechanical chain and make the source of the vibration not apparent. In this paper, we present a novel method to segment a video into mechanically coupled components based on the temporal frequency of the spatial phase variations.

2 Related Work

Studies Fleet and Jepson [1990] as early as 1990 have shown that there is sub-pixel-level information that can be extracted from local phase information to create component image velocity field. Recent works Wadhwa et al. [2013] Wadhwa et al. [2014] have expanded on that idea to implement Eulerian motion magnifications. Those approaches use complex-valued steerable pyramids and in follow up papers, Riesz Pyramids, to amplify the space-domain phases of each pixel. Although these approaches amplified the motions into the ranges detectable to human eyes, They still require manual processing by experts to isolate the undesired motions. We aim to provide an additional layer of automation that clusters/segments the different types of vibrations by their phase and amplitude.

3 Project Overview

In this paper we will explore various methods to extract the frequency information of the vibrations. Our goal is to cluster and segment this frequency information into a color overlay (outline only) of the original video to allow easy visual identification of mechanical coupling.

3.1 Frequency Information

3.1.1 Single Pixel Temporal FFT

The first approach is to take the per pixel FFT over time. This is the most simple approach. We expect this approach to be able to capture some of the vibrations as local lighting changes. However, this approach is expected to be significantly impacted by pixel and quantization noise.

3.1.2 Spacial Blur + Temporal FFT

In an attempt to mitigate the pixel noise, we will apply a Gaussian blur of the image before computing the temporal FFT. We expect this method to perform significantly better than the first approach. However this approach is still susceptible to spatially constant, yet temporally variable noise such as lighting variation as seen in AC lighting flickers.

3.1.3 Spacial Blur + Temporal Filtering + Temporal FFT

This approach further introduces a temporal filtering step to minimize temporal noise and limiting to frequencies that our vibrations are in. We will be applying a temporal sinc filter.

3.1.4 Reiz Pyramid Based Spacial Phase FFT

In addition to the previous approaches, the spacial phases of each image extract additional information in the form of the orientation of the changes. The Reiz transform is applied to a Laplacian Pyramid to extract phase information from all scales of an input image. We would like to apply FFT over the time axis to extract the time-axis frequency domain informations and leverage clustering to isolate sources of vibrations.

3.2 Clustering

We will first filter out all pixels with a temporal FFT amplitude below a fixed threshold. These low frequency pixels will be transparent in the final video.

We will stack all of the per frequency FFT amplitudes and frequencies into a vector. We will then perform PCA to reduce the dimensionality of the vectors before clustering using K-Means.

Each cluster will be assigned a different color in the final image.

4 Milestones, Timeline & Goals

We set the following timeline for our project.

4.1 Naive Approaches (0.5 weeks)

We will complete the three naive approaches in half a week. The three approaches build on each other and is relatively simple to implement.

4.2 Reiz Pyramid (1 week)

We want to complete the implementation of the Reiz Pyramid based method in the first week. We will be modifying an existing implementation from Wadhwa et al. [2013] for our project.

4.3 Clustering (1 week)

We allocate a week for the clustering method. We presented our initial method for clustering in this report but we reserve time in case this approach performs poorly in this scenario.

4.4 Reserve (0.5 week)

We reserve half a week to solving any unforeseen issues in this project.

References

David J Fleet and Allan D Jepson. Computation of component image velocity from local phase information. *International journal of computer vision*, 5(1):77–104, 1990.

Neal Wadhwa, Michael Rubinstein, Frédo Durand, and William T Freeman. Phase-based video motion processing. *ACM Transactions on Graphics (TOG)*, 32(4):1–10, 2013.

Neal Wadhwa, Michael Rubinstein, Frédo Durand, and William T Freeman. Riesz pyramids for fast phase-based video magnification. In *2014 IEEE International Conference on Computational Photography (ICCP)*, pages 1–10. IEEE, 2014.