First immersive week - Deliverables

Gonçalo Ribeiro
Student Number 93193

Doctoral Program in Mechanical Engineering

November 30, 2024

Contents

D5.1 Title, Keywords and Abstract

- **Title:** Identification techniques of operational vibration modes with MotionScope assisted by stereoscopy
- **Keywords:** "3D Reconstruction", "(3D) Phase-based Motion Magnification", "Camera Calibration", "Feature-matching", "Dense Pixel Matching", "Uncalibrated Stereo"
- **Abstract:**

When analyzing vibrations using cameras with motion magnification, valuable information is lost due to the projection of the 3D scene onto a 2D image. This loss can easily result in misinterpreted data and incorrect conclusions. The proposed research aims to address challenges in analyzing vibrations using cameras by exploring innovative methodologies for 3D motion magnification and reconstruction. Current techniques often require extensive calibration, multiple camera setups, and are limited in their ability to present results effectively. This study aims to overcome these limitations by developing more accessible, efficient, and user-friendly approaches.

The key research objectives include investigating whether combining 2D motion magnification from multiple viewpoints or directly computing 3D motion magnification is more effective, exploring alternatives to multi-camera setups using a single moving camera, and developing methods for accurate 3D reconstruction without pre-calibration. The study also aims to propose novel visualization techniques for presenting 3D motion data.

By advancing computer vision solutions and using machine learning, this research seeks to simplify experimental setups, reduce hardware demands, and improve data interpretation, enabling broader accessibility and applicability in vibration analysis and related fields.

D5.2 Draft of a graphical abstract

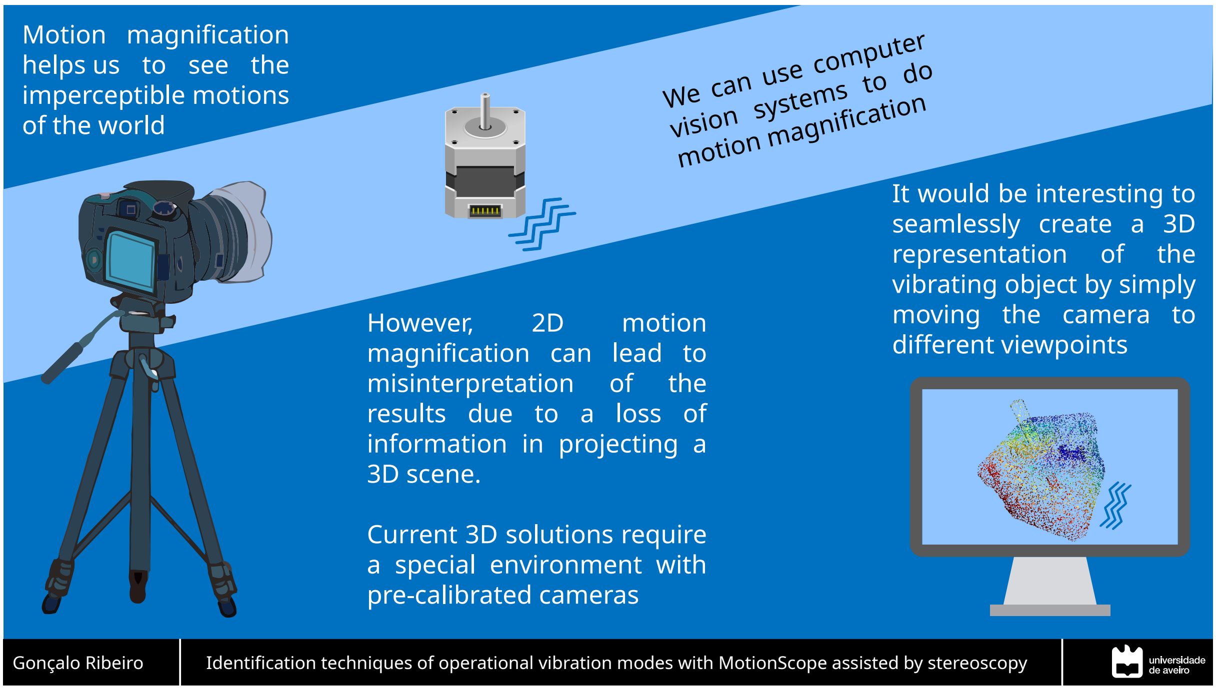


Figure 1: Graphical abstract (The full-page version is available in Appendix A)

D5.3 Research design canvas

Research Design Canvas

Name: 93193 - Gonçalo Ribeiro

Title of Thesis: Identification techniques of operational vibration modes with MotionScope assisted by stereoscopy

Date / Version: 30/11/2024

Literature

• 3D reconstruction: This is a big and well established body of literature and the overriding focus is the development of algorithms to accurately model and reproduce the geometry, texture, and appearance of real-world objects or environments in three dimensions.
Some relevant papers are: Image-Based 3D Object Reconstruction: State-of-the-Art and Trends in the Deep Learning Era, LightGlue: Local Feature Matching at Light Speed, NeRF: representing scenes as neural radiance fields for view synthesis

• Phase-based motion magnification: This is a big and well established body of literature and the overriding focus is the development of techniques to amplify subtle motions in image sequences by manipulating phase information in the frequency domain.
Some relevant papers are: 3D Motion Magnification: Visualizing Subtle Motions with Time Varying Radiance Fields, Target-free 3D tiny structural vibration measurement based on deep learning and motion magnification, Feasibility of extracting operating shapes using phase-based motion magnification technique and stereo-photogrammetry, 3D mode shapes characterisation using phase-based motion magnification in large structures using stereoscopic DIC

• Patternless camera calibration: This is a big and well established body of literature and the overriding focus is the development of techniques and methodologies to estimate a camera's intrinsic and/or extrinsic parameters without relying on pre-designed calibration patterns, such as checkerboards or grids.
Some relevant papers are: Stereo from uncalibrated cameras, ATOM: A general calibration framework for multi-modal, multi-sensor systems, Dense 3D Reconstruction with an Uncalibrated Stereo System using Coded Structured Light

Observations & Arguments

There is already extensive research on using multiple camera viewpoints to extract 3D information about the subtle motions of objects.

Some studies emphasize qualitative results, aiming to provide a better visual representation of the data. These approaches often generate videos that simulate moving the camera around the object under analysis. The main issues in this category include the large number of required viewpoints, the need for rigorously pre-calibrated systems, and the limitation of presenting results as videos, which limits the viewpoints from which the motions can be observed to the ones included in the video.

Other studies take a quantitative approach, focusing on using multi-view systems to improve the accuracy of 2D results or get 3D information. These methods rely on highly pre-calibrated camera setups and typically analyze only a small set of points. Results are usually presented by highlighting a point in an image and associating it with a graph showing its displacement over time along different axes. The primary limitations in this approach are the restricted number of points that can be analyzed and the reliance on precise camera calibration using patterns like checkerboards.

RQ / Hypotheses

• Is it more effective to combine multiple viewpoints to compute 3D motion magnification directly, or to first compute 2D motion magnification for each viewpoint and then combine them for 3D motion magnification?

• Does the latter approach reduce the need for highly accurate calibration?

• Can a single moving camera, synchronized across viewpoints, replace the use of multiple cameras for 3D motion magnification?

• Is it possible to perform 3D reconstruction using only a small number of viewpoints and without pre-calibrating the cameras?

• Is there a more effective way to present the results of 3D motion magnification?

Contributions

• Improved Methodology for 3D Motion Magnification: By reducing calibration requirements, making the process more practical and accessible, and by using a single moving camera instead of multiple cameras, which could lower costs and simplify experimental setups.

• Efficient 3D Reconstruction with Minimal Viewpoints: By developing methods for accurate 3D reconstruction using fewer viewpoints, reducing hardware requirements and computational complexity, and exploring the potential to bypass pre-calibration, which could streamline the workflow and make the technology more user-friendly.

• Enhanced Visualization of Results: By proposing new ways to display 3D motion magnification results, such as interactive 3D meshes or point clouds, which could offer better insights and flexibility compared to traditional video-based outputs.

Sample / Context

The sample will consist of raw video footage of objects subjected to vibrations induced by a shaker, as well as post-processed videos generated using the well-established 2D motion magnification software, *MotionScope*. It will also include footage of calibration patterns to evaluate traditional approaches. To address the large data requirements for training machine learning algorithms, the sample will incorporate simulator-generated data, reducing the time needed for data acquisition.

Methodology / Design / Methods

In the first stage of the research, the focus will be on replicating and evaluating state-of-the-art results. This involves rigorously calibrating both the cameras' intrinsic and extrinsic parameters and identifying suitable metrics to compare different techniques. Additionally, it is important to become familiar with tools like MotionScope, which have their workflows and limitations.

Once the state-of-the-art methods are replicated and MotionScope is operational, all the necessary tools will be in place to address the first and second research questions. The next stage will focus on answering the fourth question, which involves identifying or improving feature-matching algorithms to eliminate the need for pre-calibration. This stage will also focus on developing an effective method for performing dense matching between pixels from two viewpoints, enabling 3D reconstruction from two viewpoints, and using additional viewpoints to fill occlusions.

With 3D reconstruction and motion magnification combined, the next task will be to develop a more effective way to present the results, thus addressing the final research question. The third research question will be addressed last, as its answer depends on the effectiveness of the feature-matching algorithm developed in the previous stage.

Theory & Concepts

• Computer vision - For data manipulation and deterministic solutions
• Machine learning - For understanding and to adapt or develop machine learning architectures
• Phase-based motion magnification - Understanding the actual magnification of pixel displacement is crucial for critically interpreting results.
The methods used in this process may also be applicable to other steps in computer vision, and combining both methods could lead to greater computational efficiency.

Problem / Phenomenon

When analyzing vibrations using cameras with motion magnification, valuable information is lost due to the projection of the 3D scene onto a 2D image. This loss can easily result in misinterpreted data and incorrect conclusions.

Philosophical Assumptions / Research Paradigm

I believe this research is not grounded in any specific philosophical assumption.

Figure 2: Research design canvas (The full-page version is available in Appendix B)

D5.4 Detailed timeline

- **Year 1**
  - **Focus:** Literature review and replication of state-of-the-art results.
  - **Tasks:**
    * Learn the theory
    * Conduct a thorough literature review on 3D reconstruction, phase-based motion magnification, and patternless camera calibration.
    * Replicate state-of-the-art methods for 3D motion magnification and 3D reconstruction.
    * Familiarization with tools like MotionScope.
    * Identify suitable metrics to compare different techniques.
    * Prepare and submit the review paper.
  - **Output:** Review paper.
- **Year 2**
  - **Focus:** Addressing the first, second, and fourth research questions.
  - **Tasks:**
    * Determine if computing 2D motion magnification first reduces the need for highly accurate calibration. If so, this could be a new approach for 3D motion magnification and could result in the first research paper.
    * Write the first research paper
    * Identify or improve feature-matching algorithms to eliminate the need for pre-calibration.
    * Develop a method for dense matching between pixels from two viewpoints for 3D reconstruction.
    * Experiment with using additional viewpoints to fill occlusions in the 3D reconstruction.
    * Write the second research paper.
  - **Output:** First and second research papers.
- **Year 3**
  - **Focus:** Addressing the third and final research question.
  - **Tasks:**
    * Combine the findings of the two first research papers to create a dense 3D reconstruction of motion magnification
    * Develop a more effective way to present 3D motion magnification results (e.g., interactive 3D meshes or point clouds).
    * Explore the feasibility of using a single moving camera for 3D motion magnification.
    * Write third research paper.
  - **Output:** Third research paper.

D5.5 Abstract and paper structure (proposal) for the next research summit

- **Title:** The use of multi-view systems in phase-based motion magnification
- **Abstract:**

Phase-based motion magnification is a valuable tool for visualizing and measuring subtle motions. It utilizes complex algorithms to decompose images into amplitude and phase components at different spatial resolutions and magnifies motion information within specific frequency bands. While effective, single-camera motion magnification is limited in capturing 3D motion.

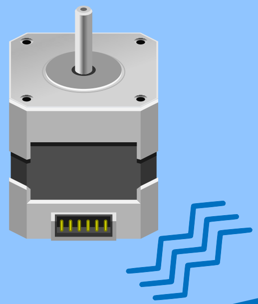
Multi-view systems, such as stereo-photogrammetry, provide 3D displacement data with sub-pixel accuracy. However, their accuracy is limited by camera resolution, noise, and calibration errors.

Integrating phase-based motion magnification with multi-view systems offers a promising solution for measuring tiny 3D vibrations, particularly at high frequencies. This combination leverages the strengths of both techniques to overcome their individual limitations. This review paper will explore the principles, methodology, advantages, and challenges of using multi-view systems in phase-based motion magnification. It will examine various applications, highlighting successful case studies in analyzing different structures.

- **Paper structure:**

  1. Introduction
  2. Background
     - Phase-based motion magnification
     - Multi-view systems
  3. Integration of Multi-view Systems and Phase-Based Motion Magnification
  4. Tests and results
  5. Conclusions

Motion magnification helps us to see the imperceptible motions of the world

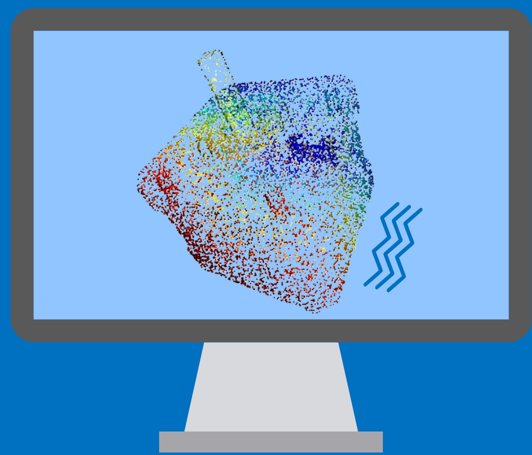


We can use computer vision systems to do motion magnification

However, 2D motion magnification can lead to misinterpretation of the results due to a loss of information in projecting a 3D scene.

Current 3D solutions require a special environment with pre-calibrated cameras

It would be interesting to seamlessly create a 3D representation of the vibrating object by simply moving the camera to different viewpoints



Name: 93193 - Gonçalo Ribeiro

Title of Thesis: Identification techniques of operational vibration modes with MotionScope assisted by stereoscopy

Date / Version: 30/11/2024

Research Design Canvas

Literature

• 3D reconstruction: This is a big and well established body of literature and the overriding focus is the development of algorithms to accurately model and reproduce the geometry, texture, and appearance of real-world objects or environments in three dimensions.
Some relevant papers are: Image-Based 3D Object Reconstruction: State-of-the-Art and Trends in the Deep Learning Era, LightGlue: Local Feature Matching at Light Speed, NeRF: representing scenes as neural radiance fields for view synthesis

• Phase-based motion magnification: This is a big and well established body of literature and the overriding focus is the development of techniques to amplify subtle motions in image sequences by manipulating phase information in the frequency domain.
Some relevant papers are: 3D Motion Magnification: Visualizing Subtle Motions with Time Varying Radiance Fields, Target-free 3D tiny structural vibration measurement based on deep learning and motion magnification, Feasibility of extracting operating shapes using phase-based motion magnification technique and stereo-photogrammetry, 3D mode shapes characterisation using phase-based motion magnification in large structures using stereoscopic DIC

• Patternless camera calibration: This is a big and well established body of literature and the overriding focus is the development of techniques and methodologies to estimate a camera's intrinsic and/or extrinsic parameters without relying on pre-designed calibration patterns, such as checkerboards or grids.
Some relevant papers are: Stereo from uncalibrated cameras, ATOM: A general calibration framework for multi-modal, multi-sensor systems, Dense 3D Reconstruction with an Uncalibrated Stereo System using Coded Structured Light

Observations & Arguments

There is already extensive research on using multiple camera viewpoints to extract 3D information about the subtle motions of objects.

Some studies emphasize qualitative results, aiming to provide a better visual representation of the data. These approaches often generate videos that simulate moving the camera around the object under analysis. The main issues in this category include the large number of required viewpoints, the need for rigorously pre-calibrated systems, and the limitation of presenting results as videos, which limits the viewpoints from which the motions can be observed to the ones included in the video.

Other studies take a quantitative approach, focusing on using multi-view systems to improve the accuracy of 2D results or get 3D information. These methods rely on highly pre-calibrated camera setups and typically analyze only a small set of points. Results are usually presented by highlighting a point in an image and associating it with a graph showing its displacement over time along different axes. The primary limitations in this approach are the restricted number of points that can be analyzed and the reliance on precise camera calibration using patterns like checkerboards.

RQ / Hypotheses

• Is it more effective to combine multiple viewpoints to compute 3D motion magnification directly, or to first compute 2D motion magnification for each viewpoint and then combine them for 3D motion magnification?

• Does the latter approach reduce the need for highly accurate calibration?

• Can a single moving camera, synchronized across viewpoints, replace the use of multiple cameras for 3D motion magnification?

• Is it possible to perform 3D reconstruction using only a small number of viewpoints and without pre-calibrating the cameras?

• Is there a more effective way to present the results of 3D motion magnification?

Contributions

• Improved Methodology for 3D Motion Magnification: By reducing calibration requirements, making the process more practical and accessible, and by using a single moving camera instead of multiple cameras, which could lower costs and simplify experimental setups.

• Efficient 3D Reconstruction with Minimal Viewpoints: By developing methods for accurate 3D reconstruction using fewer viewpoints, reducing hardware requirements and computational complexity, and exploring the potential to bypass pre-calibration, which could streamline the workflow and make the technology more user-friendly.

• Enhanced Visualization of Results: By proposing new ways to display 3D motion magnification results, such as interactive 3D meshes or point clouds, which could offer better insights and flexibility compared to traditional video-based outputs.

Sample / Context

The sample will consist of raw video footage of objects subjected to vibrations induced by a shaker, as well as post-processed videos generated using the well-established 2D motion magnification software, *MotionScope*. It will also include footage of calibration patterns to evaluate traditional approaches. To address the large data requirements for training machine learning algorithms, the sample will incorporate simulator-generated data, reducing the time needed for data acquisition.

Methodology / Design / Methods

In the first stage of the research, the focus will be on replicating and evaluating state-of-the-art results. This involves rigorously calibrating both the cameras' intrinsic and extrinsic parameters and identifying suitable metrics to compare different techniques. Additionally, it is important to become familiar with tools like MotionScope, which have their workflows and limitations.

Once the state-of-the-art methods are replicated and MotionScope is operational, all the necessary tools will be in place to address the first and second research questions. The next stage will focus on answering the fourth question, which involves identifying or improving feature-matching algorithms to eliminate the need for pre-calibration. This stage will also focus on developing an effective method for performing dense matching between pixels from two viewpoints, enabling 3D reconstruction from two viewpoints, and using additional viewpoints to fill occlusions.

With 3D reconstruction and motion magnification combined, the next task will be to develop a more effective way to present the results, thus addressing the final research question. The third research question will be addressed last, as its answer depends on the effectiveness of the feature-matching algorithm developed in the previous stage.

Theory & Concepts

• Computer vision - For data manipulation and deterministic solutions
• Machine learning - For understanding and to adapt or develop machine learning architectures
• Phase-based motion magnification - Understanding the actual magnification of pixel displacement is crucial for critically interpreting results. The methods used in this process may also be applicable to other steps in computer vision, and combining both methods could lead to greater computational efficiency.

Problem / Phenomenon

When analyzing vibrations using cameras with motion magnification, valuable information is lost due to the projection of the 3D scene onto a 2D image. This loss can easily result in misinterpreted data and incorrect conclusions.

Philosophical Assumptions / Research Paradigm

I believe this research is not grounded in any specific philosophical assumption.