A review on 3D motion magnification

Gonçalo Ribeiro

gribeiro@ua.pt

Department of Mechanical Engineering

University of Aveiro

January 19, 2025

***Keywords* —** 3D Reconstruction, (3D) Motion Magnification, Camera Calibration, Feature-matching, Dense Pixel Matching, Digital Image Correlation

***Target Journal* —** Mechanical Systems and Signal Processing

Abstract

This review paper explores the field of 3D motion magnification, a computer vision technique that enhances the visibility of subtle motions and vibrations in three dimensions. This technology is crucial for applications such as structural health monitoring, damage detection, and the design of vibration control systems. Current methods for 3D motion magnification include using time-varying radiance fields with a moving camera, combining stereo-photogrammetry with phase-based motion magnification, and using a single camera with a Finite Element model. Other methods include combining Fringe Projection with DIC, using binocular vision with deep learning, and extending phase-based motion magnification to 3D volumetric data.

While these methods show promise, significant challenges remain. Both 2D and 3D Digital Image Correlation are sensitive to lighting variations and speckle pattern quality, which affects measurement accuracy. Feature-matching algorithms have limitations in the number of points that can be analyzed, impacting the quality of generated meshes. Some methods are computationally intensive or require multiple viewpoints, limiting their practical application. Many methods also rely on highly calibrated camera setups, complicating their implementation. The lack of methods for using affordable DSLR cameras in 3D applications also presents an obstacle. Future work should focus on developing robust, efficient, and flexible techniques that address these limitations.

1 Introduction

Vibration measurement plays a critical role in structural and mechanical analysis. It facilitates modal identification and characterization by revealing the dynamic properties of structures, such as natural frequencies, mode shapes, and damping ratios. These insights are vital for understanding how structures respond to loads and excitations. Additionally, vibration measurement is used for operational deflection shape identification, which combines multiple mode shapes and proves valuable when determining individual mode shapes is challenging [1, 2].

This technique is indispensable for structural health monitoring and damage detection, as changes in vibration patterns, frequencies, or mode shapes can signal issues like cracks, loose connections, or other forms of structural damage. Furthermore, understanding vibration characteristics is essential for designing vibration control systems to reduce excessive vibrations, enhancing the performance and safety of structures and machinery. Vibration data also aids in refining and calibrating numerical models of structures, ensuring they accurately represent real-world behavior and support reliable predictive analysis [1, 2].

Traditional methods to measure vibrations use contact sensors, which have drawbacks such as difficult installation, high cost, and low efficiency. Non-contact methods like laser Doppler vibrometers and computer vision based methods have emerged as alternatives. While laser Doppler vibrometers can achieve high precision, they are expensive and difficult to operate. Computer vision based methods are advantageous because they are non-contact, provide full-field measurements, and are easy to operate. However, when vibration amplitudes are small, special computer vision techniques like motion magnification are needed [1, 2].

Motion magnification is a computer vision technique that makes subtle motions or vibrations, often imperceptible, visible. It can be based on Lagrangian or Eulerian approaches, which are used in fluid mechanics to describe motion, or based on learning approaches [1, 2].

However, motion magnification is limited to 2D, capturing only in-plane motions and failing to represent out-of-plane or depth movements. This creates issues in scenarios like 3D structural vibration analysis, depth-dependent motion, or accurate modal analysis, where full spatial motion is critical. 3D motion magnification solves this by amplifying motion in all three dimensions, using techniques like multi-camera systems to accurately capture and visualize complex 3D dynamics.

3D motion magnification is particularly useful in fields where understanding complex, three-dimensional motion is critical. For example, in aerospace engineering, it can reveal torsional vibrations in aircraft wings that are invisible to 2D analysis, ensuring structural integrity and aerodynamic performance. Similarly, in structural health monitoring, it allows engineers to detect out-of-plane bending or twisting in bridges and buildings, providing a complete picture of how these structures respond to stress. In robotics, 3D motion magnification is invaluable for analyzing the intricate motion of robotic arms, capturing all degrees of freedom, including subtle depth-related movements. It is also essential for monitoring wind turbine blades, where depth-dependent deflections caused by wind loads may go unnoticed with 2D techniques. By amplifying motion in all three dimensions, 3D motion magnification provides precise insights, making it a powerful tool for solving challenges in industrial diagnostics, biomechanics, and beyond.

Current solutions often rely on expensive equipment, such as high-speed cameras, and some also require professional-grade GPUs. Moreover, these solutions are typically designed for controlled environments, requiring precise camera calibration and, in some cases, painting the object being analyzed. These requirements are often impractical in real-world applications, where analyzing operational vibration modes is particularly important.

Despite the importance of 3D motion magnification, there isn’t a review article on the topic. Therefore, this document fills that gap by analyzing the works on 3D motion magnification.

2 Methodology

The topic of this review article includes two components: ”3D” and ”motion magnification”. I searched for works on 2D motion magnification, general 3D reconstructions, and consequentially feature-matching algorithms to better understand the topic. However, the main focus of this document is ”3D Motion Magnification”. Therefore, the majority of the works cited in the document are about this topic.

Some databases and tools were used to find the literature on these topics:

- **Web of Science**[1] : To find the first articles, I used the Web of Science using the following keywords: ”Motion magnification”, ”3D motion magnification”, ”Multi view motion magnification”, ”3D vibration”, ”Multi camera motion magnification”, ”Stereo motion magnification”, ”Phase-based motion magnification”, ”Feature matching”, ”Camera calibration”, ”Uncalibrated Stereo”, ”Digital Image Correlation”, ”3D Digital Image Correlation”;
- **Google Scholar**[2] : I used Google Scholar with the same keywords as in Web of Science;
- **Google Patents**[3] : I didn’t spend much time with Google patents and I didn’t find anything relevant;
- **Supervisors**: My supervisors showed me that the Department of Mechanical Engineering of the University of Aveiro has already used software for phase-based motion magnification using a single viewpoint. The papers published by the software developer are important for understanding the base theory used to create the software.
- **Online Courses**: My supervisors assigned me challenges that required learning about computer vision, machine learning, and computer science. These courses provided valuable references and, more importantly, introduced me to targeted keywords like ”Uncalibrated Stereo” and ”Epipolar Constraint”. These courses allowed me to search beyond what was already done on the subject of 3D motion magnification.
- **Litmaps**[4] : This tool helped me identify streams of thought because it links the papers to the references and citations. With this tool, it was also easier to identify the foundations of a topic, thus the more relevant old papers. Although the old papers are not relevant for the review, they are sometimes important to understand newer papers.

To filter the relevant papers, I used the following order of preference:

- Availability through b-on;

[1] https://www.webofscience.com/wos/woscc/basic-search
[2] https://scholar.google.com/
[3] https://patents.google.com/
[4] https://www.litmaps.com/

- Number of citations;
- Connection with other papers within a stream of thought;
- Relevance of the journal;

The main limitations of this research were the unavailability of source code from the referenced papers, which will make testing and comparing results significantly more challenging, and the paywall restrictions on some papers that were not accessible through B-on.

To select papers on techniques for achieving 3D motion magnification, I included those whose titles and mini summaries in Web of Science and Google Scholar closely matched the concept of 3D motion magnification. Due to time constraints, I excluded techniques that did not focus on mechanical systems, such as those used in deep-fake detection. I also disregarded papers on motion magnification for 3D printers, as they primarily involve 2D motion magnification despite including ”3D” in the title.

Given the limited number of relevant papers, I did not apply strict criteria for selection. For instance, a search for ”3D” and ”motion magnification” in Web of Science returned only 24 papers.

3 Main body

3D motion magnification has been achieved using different approaches. Some focus more on the quantitative results, and others also provide qualitative results, more graphical results for an intuitive understanding of motion.

Feng *et al.* [3] proposed a method for magnifying subtle motions in 3D scenes captured by a moving camera. It uses time-varying radiance fields to represent the scene and applies the Eulerian principle for motion magnification. The method works by extracting and amplifying the variation of the embedding of a fixed point over time and uses both implicit and tri-plane-based radiance fields as the underlying 3D scene representation. Feng *et al.* [3] evaluated this method quantitatively on synthetic scenes and qualitatively on real-world scenes with various camera setups. This method generates a video moving around the object with the motion magnified, providing a really intuitive understanding of motion. The source code of this method is publicly available.

Another way to achieve 3D motion magnification is by using stereo-photogrammetry techniques like 3D Digital Image Correlation (DIC) and 3D Point Tracking (3DPT), combined with phase-based motion magnification to measure 3D displacements and vibrations. This approach improves the Signal-to-Noise Ratio (SNR) in optical measurements, allowing for the detection of subtle motions at higher frequencies. Poozesh *et al.* [4], Molina-Viedma *et al.* [5] and Yan and Zhang [6] based their works in this approach.

Poozesh *et al.* [4] evaluated the use of motion magnification with 3D DIC and 3DPT to extract high-frequency operating shapes of structures. It shows that the method is effective for measuring small vibrations and extracting 3D operating shapes with low SNR. This was tested on a cantilever beam and a wind turbine blade.

Molina-Viedma *et al.* [5] explored measurements using 3D DIC in stereoscopic sets of magnified images for 3D Operational Deflection Shapes (ODSs) characterization. It validated that motion magnification helps to reveal ODSss from low amplitude tests and at high frequencies. This was evaluated on a cantilever beam and a curved panel.

Yan and Zhang [6] combined phase-based motion magnification and 3D-DIC to obtain high-frequency vibration modes. The method is validated by separating the first five out-of-plane vibration mode shapes of a cantilever beam under a single hammer excitation, and also applied to identify the out-of-plane vibration modes of a real engine pipe.

Xuan Le *et al.* [7] proposed a framework for 3D noncontact stereovision-based cable vibration measurement using Phase-based Video Motion Magnification. It magnifies micro-vibrations of a cable and uses a centroid-based bounding-box tracking technique. This method has been validated by comparing its results with accelerometer-based measurements.

Another method, presented by Renaud *et al.* [8], uses a single camera and a Finite Element model of the structure to measure the vibrations in 3D. The method involves identifying the camera's intrinsic and extrinsic parameters, projecting numerical deflection shapes and normal modes onto the camera's image frame, and comparing the motion of targets seen by the camera with the motion of the Finite Element model. By comparing the motion of targets in the video with the projected FE model motion, the method estimates the time evolution of modal amplitudes.

Felipe-Sesé *et al.* [9] combined Fringe Projection, which measures out-of-plane shape, with DIC, which tracks in-plane motion. Phase-Based Motion Magnification is then applied to enhance the visibility of subtle movements in the image sequences. The method independently magnifies both the fringe patterns used for FP and the speckle patterns used for 2D DIC. This allows for the measurement of small vibrations and deformations. By correcting for in-plane distortions caused by out-of-plane displacement, accurate 3D maps can be obtained. The technique has been validated against 3D DIC and Scanning Laser Doppler Vibrometer, showing its ability to measure operational deflection shapes and improve the SNR. The approach also provides simultaneous qualitative and quantitative information of the motion and displacement.

Shao *et al.* [10] introduced a novel method for measuring 3D vibrations without the need for artificial targets. The system uses a binocular vision system with two cameras to capture the motion of the structure. A phase-based motion magnification algorithm is then applied to the recorded videos to amplify small motions. This enhances the SNR and makes it possible to measure vibrations that would otherwise be imperceptible. Deep learning techniques are used to achieve target-free measurements and to detect and match key points. The SuperPoint network detects key points, and the SuperGlue algorithm matches these key points between images. Then, a polynomial triangulation approach is used to calculate the 3D vibration displacements. The accuracy of the system has been validated through experimental tests, including tests on a steel cantilever beam and an in-field test on a pedestrian bridge. The results have been compared with measurements from physical sensors like LVDTs, laser displacement sensors, and accelerometers showing good agreement. The study also includes an analysis of the system's sensitivity and finds that motion magnification is beneficial when the peak vibration displacement is smaller than one pixel.

Southwick *et al.* [11] presented a method called Volumetric Motion Magnification, which extends phase-based motion magnification into three dimensions for the analysis of volumetric data over time. Volumetric Motion Magnification uses a 3D complex steerable pyramid to decompose volumetric frames, enabling the filtering and amplification of subtle motions within 4D datasets (3D motion magnification outside and inside the object using, for example, XRays). This technique has been tested on synthetic data, demonstrating its ability to extract subtle motions, such as modal deformations of a cylinder.

Table 1 summarizes the techniques available for 3D motion magnification and the studies in which they were implemented.

Table 1: The available techniques for 3D motion magnification

Technique	Works on the development and/or use
Motion magnification on time-varying radiance fields	Feng *et al.* [3]
Phase-based motion magnification with 3D DIC and 3DPT	Poozesh *et al.* [4], Molina-Viedma *et al.* [5] and Yan and Zhang [6]
Finite element model combined with 2D motion magnification	Renaud *et al.* [8]
Fringe projection with DIC for 3D motion magnification	Felipe-Sesé *et al.* [9]
Pixel matching and triangulation for 3D motion magnification	Shao *et al.* [10] and Xuan Le *et al.* [7]
Volumetric motion magnificarion	Southwick *et al.* [11]

This review paper focuses specifically on the use of cameras for 3D motion magnification. In contrast, volume motion magnification employs X-ray Computed Tomography and Magnetic Resonance Imaging data. Consequently, volume motion magnification is excluded from comparisons in this review.

As of the time of writing, no standardized metric exists for evaluating 3D motion magnification techniques. Therefore, the authors of the studies tested their methods using their own evaluation approaches:

- **Motion magnification on time-varying radiance fields**: Evaluated using both synthetic and real-world scenes. Synthetic datasets allowed for quality measurement against ground-truth videos, while real-world scenes validated the method's practicality under various camera setups, scene compositions, and subject motions. The quantitative evaluation employed Structural Similarity

Index Measure (SSIM) and Perceptual Image Patch Similarity (LPIPS) metrics against ground-truth magnified frames.

- **Phase-based motion magnification with 3D DIC and 3DPT**: Evaluated by comparing results with numerical models and experimental data from Scanning Laser Doppler Vibrometer (SLDV) or accelerometers. Tests included objects such as cantilever beams and large curved composite panels.

- **Finite element modeling combined with 2D motion magnification**: This method does not evaluate motion magnification itself, as it is primarily used to estimate damping parameters and refine finite element models. Therefore, there is no ground-truth data. Renaud *et al.* [8] noted that they couldn't use accelerometers due to technical constraints.

- **Fringe projection with DIC for 3D motion magnification**: Validated through solid-rigid tests, with results compared against 3D DIC and SLDV data in cantilever beam experiments.

- **Pixel matching and triangulation for 3D motion magnification**: Validated through wind tunnel tests of cable vibrations. Results were compared with conventional accelerometer-based measurements by Xuan Le *et al.* [7]. In another approach to the technique, Shao *et al.* [10] compared the results with accelerometers,Linear Variable Differential Transformers (LVDTs), and lasers in a cantilever beam.

Most of these methods used are not directly comparable. The more comparable ones used cantilever beams, but even those were in different conditions. For example, the Phase-based motion magnification with 3D DIC and 3DPT technique was evaluated using a speckle pattern, and Shao *et al.* [10] evaluated the Pixel matching and triangulation for 3D motion magnification technique using a normal beam.

In other fields, such as perception in computer vision, well-established datasets and metrics enable consistent evaluation of different approaches to similar problems. For example, the BDD100k dataset includes ground-truth data for object detection, instance segmentation, and other tasks, along with standardized evaluation methods [12]. In contrast, the field of 3D motion magnification lacks a comprehensive dataset that incorporates both synthetic and real-world data, as well as standardized metrics, to facilitate fair comparisons between methods. Addressing this gap would significantly advance the field.

Despite the absence of unified evaluation metrics, it is still possible to analyze each approach to see their differences, compare their advantages and disadvantages when possible, and identify areas for improvement.

2D motion magnification, a foundational concept for 3D techniques, can be achieved through three primary methods [13]:

- **Lagrangian methods** compute motion explicitly and warp video frames based on magnified velocity vectors, but can be computationally intensive and prone to errors.

- **Eulerian methods** process video in space and time separately, avoiding the need for optical flow computation. However, linear Eulerian methods can be limited by small magnification factors and amplified noise.

- **Phase-based methods** are a newer type of Eulerian approach that analyzes phase variations over time in complex steerable pyramids. This method allows for larger magnification factors and is less sensitive to noise compared to previous Eulerian techniques. This approach manipulates phases, rather than amplitudes which avoids increasing the magnitude of spatial noise. Also, phase-based methods are exact for sinusoidal waves.

Authors exploring 3D motion magnification have predominantly adopted phase-based motion magnification, often finding it to yield the best results. Most techniques perform motion magnification on 2D images followed by triangulation to achieve 3D results. An exception is the motion magnification on time-varying radiance fields, which directly applies magnification within the feature embeddings of individual 3D points.

The concept of ”3D” is broad and can be interpreted in various ways. In the works discussed earlier, 3D representations range from a limited set of discrete points with corresponding displacement plots for each axis, to fully realized 3D visualizations. These advanced visualizations take the form of videos that illustrate the displacement of every point in the scene, dynamically changing the viewpoint smoothly over time.

The process of recovering the 3D information, either for a limited set of points or for a full 3D representation, can be done with traditional methods that involve the selection of matching pixels between images and triangulation based on the intrinsic and extrinsic parameters of the camera [14].

The Pixel matching and triangulation for 3D motion magnification technique uses the traditional triangulation method for 3D reconstruction. For point selection, Xuan Le *et al.* [7] manually selected points, while Shao *et al.* [10] employed an automated approach using the feature-matching algorithm SuperGlue [15] to identify pixel matches between images. In both cases, the 3D data is presented as three displacement plots per point for a selected set of points on the object.

The Phase-based motion magnification with 3D DIC and 3DPT technique also relies on triangulation, but indirectly. This method uses a closed-source software, Vic-3D[5], which incorporates triangulation. To present the results, the software outputs a 2D motion-magnified video where color coding is used to represent magnified motion along the third axis.

[5]https://www.correlatedsolutions.com/vic-3d

In contrast, the Fringe projection with DIC for 3D motion magnification and Finite element modeling combined with 2D motion magnification techniques do not use triangulation. These are single-camera, single-view methods. The fringe projection technique retrieves depth information, while the finite element modeling approach relies on a digital model.

The Motion magnification on time-varying radiance fields technique employs a completely different approach, leveraging machine learning, specifically the NERF [16] algorithm. This method produces a video that moves around the object, showcasing the magnified motion. Although the work by Feng *et al.* [3] does not directly provide displacement plots, it is theoretically possible to extract them with additional effort due to the algorithm's NERF-based foundation. When using the NERF network, we input the spatial coordinates of a point. NERF then outputs color information in three channels (red, green, and blue) along with an opacity value. If the point is positioned in mid-air, NERF will output an opacity value of 0. The big downside of using NERF is the high computational power needed to train this deep-learning network, requiring professional-grade GPUs with at least 16GB of VRAM.

Despite the advancements in these techniques, several challenges remain to be addressed.

Both 2D and 3D DIC are highly sensitive to lighting variations and rely heavily on the quality of the speckle pattern [17, 18]. In real-world scenarios, applying a speckle pattern in the object in analysis and controlling the lighting conditions is often impractical.

While feature-matching algorithms like SuperGlue deliver impressive results [15], faster and more efficient networks, such as LightGlue, have emerged with even better performance [19].

For instance, Shao *et al.* [10] implemented SuperGlue to identify matching points, eliminating the need for a speckle pattern. However, this approach limits the number of points that can be analyzed. The number of points matched by state-of-the-art feature-matching methods remains insufficient to generate high-quality meshes. As shown in Figure 1, a mesh created using only LightGlue is very sparse, whereas Figure 2 illustrates the results of employing a simple L^2 norm for dense matching.

The method used to produce the results shown in Figure 2 leverages the stereo system characteristics of the image pair. In a stereo system, the epipolar geometry, which defines a plane containing the centers of both cameras and a point in space, provides a useful constraint. Specifically, when a pixel is selected in one image, the epipolar constraint ensures that the corresponding pixel in the other image lies on the epipolar line [14]. Figure 3 depicts the epipolar geometry. Feature matching algorithms, which are used to identify corresponding points between images, typically produce confidence values that are filtered using Lowe's ratio [20]. The performance of these algorithms could potentially be enhanced by incorporating information about the epipolar line to refine the filtering of matches.

The only method capable of handling occlusions effectively is the one proposed by Feng *et al.* [3]. However, this approach is computationally intensive and requires a significant number of viewpoints, making it less practical in many scenarios. Feng *et al.* [3] implemented motion magnification on NERF.

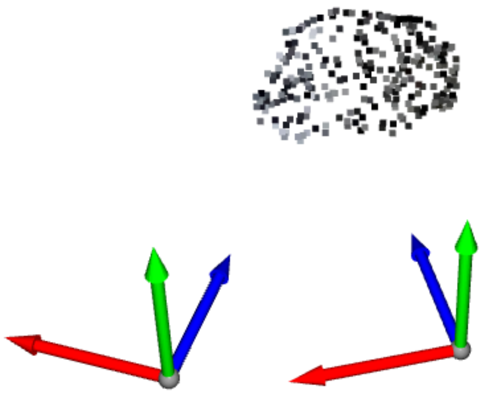

Figure 1: Implementation of LightGlue for pixel matching

However, there are already researchers improving the NERF algorithm. These improvements are mainly on computational efficiency rather than the reduction of necessary viewpoints.

- Recursive-NeRF improves upon NeRF by using a recursive, multi-stage approach that adapts to scene complexity. It starts with a small network and only forwards uncertain query coordinates to deeper, more powerful networks. This is determined by an uncertainty value predicted at each level. Low uncertainty coordinates are finalized, while high uncertainty coordinates go to deeper networks. This dynamic approach focuses resources on complex areas, avoiding unnecessary computations in simpler regions. The final image is a composite of all levels. Recursive-NeRF also uses k-means clustering to divide uncertain points and assign them to different branch networks for targeted training [21].

- FastNeRF significantly speeds up rendering by factorizing the neural network into position-dependent and direction-dependent networks. The position-dependent network produces a deep radiance map, while the direction-dependent network outputs weights. Their inner product gives a color value. This factorization allows separate caching of network outputs, dramatically improving inference times. FastNeRF looks up pre-computed values, unlike NeRF's multiple forward passes. FastNeRF's memory complexity is $O(k^3 * (1 + D \times 3) + l^2 \times D)$, much lower than NeRF's $O(k^3 l^2)$ (where O() is the Big O notation, k is the number of bins for positions, l is the number of bins for ray directions and D is the number of components in the deep radiance map). This enables real-time rendering on consumer GPUs. FastNeRF uses OpenVDB for a sparse data structure and hardware-accelerated ray tracing to skip empty space. FastNeRF achieves rendering speeds that

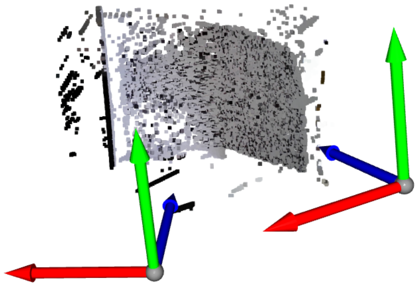

Figure 2: Implementation of L^2 norm across every epipolar line for dense matching

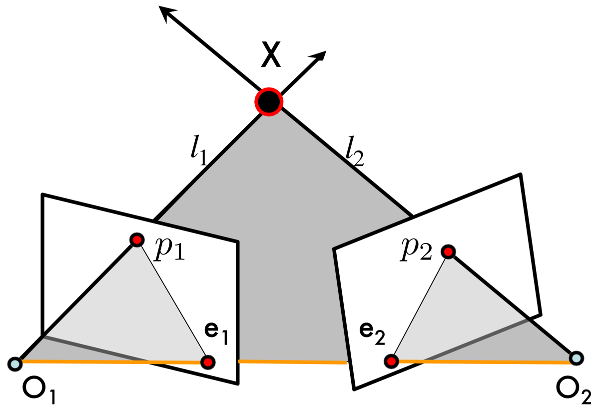

Figure 3: The epipolar geometry

are thousands of times faster than NeRF by avoiding calls to an MLP at inference time using its caching mechanism [22].

- Decomposed Radiance Fields (DeRF) enhances NeRF using spatial decomposition, assigning smaller networks to different parts of the scene. DeRF addresses diminishing returns from increasing network capacity. It uses independent ”head” networks for specific areas. The final image is rendered using the Painter’s Algorithm to combine the head outputs. This allows independent rendering of scene parts. DeRF uses a differentiable Voronoi diagram for decomposition, which is also GPU-friendly. The Voronoi diagram is learned to optimize scene representation. Using multiple smaller networks instead of one large network achieves comparable or better quality with reduced inference time. DeRF is complementary to other speedup approaches like Neural Sparse Voxel Fields [23].

Another limitation of these methods is their reliance on highly calibrated camera setups. Developing a more flexible calibration process, potentially one that could be performed on the fly, would greatly simplify the setup and broaden the applicability of these techniques. One field that could be explored in this regard is the uncalibrated stereo. If the intrinsics are calibrated, the process of calibrating an uncalibrated stereo setup involves detecting and matching corresponding points in both images. Using these matches and the known intrinsic parameters, the essential matrix is estimated. The essential matrix is then decomposed into rotation and translation to determine the relative pose between the two cameras. Finally, stereo rectification is performed, aligning the image planes to ensure epipolar consistency, which enables depth estimation [24]. This is essentially an online calibration method that, if successful, would significantly simplify the setup process. It would enable the use of a single camera to capture data from one viewpoint at a time while also leveraging periodic motions. Such an approach would be much easier to implement in real-world scenarios, particularly in industrial environments.

In the realm of 2D motion magnification, advancements have been made to bypass the need for high-speed cameras—which remain prohibitively expensive—by using DSLR cameras instead [25]. However, a similar approach has yet to be developed for 3D applications.

Javh *et al.* [25] used Spectral Optical Flow Imaging (SOFI) to measure vibrations beyond a camera's frame rate. SOFI works by harmonically blinking a light source at a specific frequency, which modulates the object's displacement. The camera then integrates the light intensity over its exposure time. This process acts as an analogue Fourier transform, where each image, taken with a different light-blinking frequency, produces separate Fourier coefficients representing specific spectral components of the displacement. A reference image, taken with constant light, is subtracted from an image taken with blinking light. This difference is scaled using the reference image gradient and illumination scaling to get a displacement spectral component for every pixel. Different spectral components can be measured by changing the light's blinking frequency, effectively measuring vibrations at frequencies far exceeding the camera's frame rate. SOFI captures individual spectral components with each image, unlike stroboscopic methods that capture instantaneous still frames [25].

There are also proprietary solutions, such as MotionScope[6], which work with DSLR cameras. However, this solution depends on a shaker that vibrates the object while being connected to a computer and employs a proprietary technique known as ”Smart aliasing”.

4 Conclusion

3D motion magnification is a valuable tool for revealing subtle motions and vibrations that are not visible to the naked eye, making it useful in structural and mechanical analysis. This review has explored various methods for achieving 3D motion magnification, including techniques that use time-varying radiance fields with a moving camera, stereo-photogrammetry combined with phase-based motion magnification,

[6]https://www.motion-scope.com/

single camera setups with Finite Element models, combinations of Fringe Projection and DIC, stereo vision with deep learning, and the extension of phase-based motion magnification to 3D volumetric data.

While these techniques have shown great promise, several challenges remain to be addressed. Both 2D and 3D DIC are highly sensitive to lighting variations and the quality of the speckle pattern. Feature-matching algorithms like SuperGlue, while effective, still have limitations in the number of points that can be analyzed, which affects the quality of generated meshes. Some methods, such as the one proposed by Feng *et al.* [3], are computationally intensive and require multiple viewpoints, making them less practical in many scenarios. Additionally, many methods rely on highly calibrated camera setups, which complicates their implementation and limits their applicability. Finally, there is a need to develop methods for using DSLR cameras, which are more affordable than high-speed cameras, in 3D applications.

Advancements are already underway in the techniques underlying 3D motion magnification. For instance, NERF, a computationally intensive algorithm, has inspired similar methods that require significantly less computing power. Feature matching algorithms have also seen substantial improvements, and the researchers in 3D motion magnification have not yet taken advantage of the epipolar restriction coming from stereographic systems to improve the reliability of feature matching. In the 2D domain, researchers have already transitioned to using DSLR cameras instead of high-speed cameras, paving the way for more cost-effective and accessible implementations.

Future research should focus on creating more robust, efficient, and flexible methods. This includes developing techniques that are less sensitive to lighting and speckle pattern quality, improving the density of matched points, reducing computational costs, simplifying camera calibration, and enabling the use of DSLR cameras for 3D motion magnification. A good starting point is the implementation of innovations in the techniques underlying 3D motion magnification.

These advancements would make 3D motion magnification more affordable and user-friendly. The cost would be reduced by utilizing a single DSLR camera instead of multiple high-speed cameras and consumer-grade GPUs instead of professional-grade hardware. Usability would also improve by eliminating the need to apply a speckle pattern to the object being analyzed and avoiding complex camera calibration methods. Instead, a single DSLR camera could simply be moved between viewpoints. These improvements are critical for broader adoption, enabling the technology to be used in most industrial settings. The reduced costs would also make full-time, real-time monitoring of important mechanical structures feasible.

Beyond cost and accessibility, 3D motion magnification has the potential to benefit various industries. Fields such as structural health monitoring, robotics, aerospace, and biomechanics could see improvements in diagnostics, safety, and performance optimization as the technology continues to evolve. Interdisciplinary collaboration between computer vision, machine learning, and engineering disciplines could play a key role in addressing current limitations and fostering innovation. As the techniques become more efficient and adaptable, they may find applications beyond industrial settings, including areas like medical diagnostics, and potentially even consumer-grade devices, making subtle motion analysis more widely available.

5 References

[1] K. Luo, X. Kong, J. Li, J. Hu, and L. Deng, "Motion magnification for video-based vibration measurement of civil structures: A review," *Mechanical Systems and Signal Processing*, vol. 220, p. 111 681, Nov. 1, 2024. DOI: 10.1016/j.ymssp.2024.111681.

[2] M. Śmieja, J. Mamala, K. Prażnowski, T. Ciepliński, and Ł. Szumilas, "Motion magnification of vibration image in estimation of technical object condition-review," *Sensors*, vol. 21, no. 19, p. 6572, Jan. 2021. DOI: 10.3390/s21196572.

[3] B. Y. Feng, H. Alzayer, M. Rubinstein, W. T. Freeman, and J.-B. Huang, "3d motion magnification: Visualizing subtle motions with time-varying radiance fields," in *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, Paris, France: IEEE, Oct. 1, 2023, pp. 9803–9812. DOI: 10.1109/ICCV51070.2023.00902.

[4] P. Poozesh, A. Sarrafi, Z. Mao, P. Avitabile, and C. Niezrecki, "Feasibility of extracting operating shapes using phase-based motion magnification technique and stereo-photogrammetry," *Journal of Sound and Vibration*, vol. 407, pp. 350–366, Oct. 27, 2017. DOI: 10.1016/j.jsv.2017.06.003.

[5] A. J. Molina-Viedma, L. Felipe-Sesé, E. López-Alba, and F. A. Díaz, "3d mode shapes characterisation using phase-based motion magnification in large structures using stereoscopic DIC," *Mechanical Systems and Signal Processing*, vol. 108, pp. 140–155, Aug. 1, 2018. DOI: 10.1016/j.ymssp.2018.02.006.

[6] S. Yan and Z. Zhang, "3d mode shapes characterization under hammer impact using 3d-DIC and phase-based motion magnification," *Engineering Research Express*, vol. 6, no. 3, p. 035 544, Aug. 2024. DOI: 10.1088/2631-8695/ad6e53.

[7] L. Xuan Le, D. M. Siringoringo, H. Katsuchi, Y. Fujino, and B. Xuan Luong, "Stereovision-based vibration measurement of stay cable using synchronized multi-camera setup and video motion magnification," *Engineering Structures*, vol. 296, p. 116 938, Dec. 1, 2023. DOI: 10.1016/j.engstruct.2023.116938.

[8] F. Renaud, S. Lo Feudo, J.-L. Dion, and A. Goeller, "3d vibrations reconstruction with only one camera," *Mechanical Systems and Signal Processing*, vol. 162, p. 108 032, Jan. 1, 2022. DOI: 10.1016/j.ymssp.2021.108032.

[9] L. Felipe-Sesé, A. J. Molina-Viedma, M. Pastor-Cintas, E. López-Alba, and F. A. Díaz, "Exploiting phase-based motion magnification for the measurement of subtle 3d deformation maps with FP + 2d-DIC," *Measurement*, vol. 195, p. 111 122, May 31, 2022. DOI: 10.1016/j.measurement.2022.111122.

[10] Y. Shao, L. Li, J. Li, S. An, and H. Hao, "Target-free 3d tiny structural vibration measurement based on deep learning and motion magnification," *Journal of Sound and Vibration*, vol. 538, p. 117 244, Nov. 10, 2022. DOI: 10.1016/j.jsv.2022.117244.

[11] M. Southwick, Z. Mao, and C. Niezrecki, "Volumetric motion magnification: Subtle motion extraction from 4d data," *Measurement*, vol. 176, p. 109 211, May 1, 2021. DOI: 10.1016/j.measurement.2021.109211.

[12] F. Yu *et al.*, "BDD100k: A diverse driving dataset for heterogeneous multitask learning," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA: IEEE, Jun. 2020, pp. 2633–2642. DOI: 10.1109/CVPR42600.2020.00271.

[13] N. Wadhwa, M. Rubinstein, F. Durand, and W. T. Freeman, "Phase-based video motion processing," *ACM Trans. Graph.*, vol. 32, no. 4, 80:1–80:10, Jul. 21, 2013. DOI: 10.1145/2461912.2461966.

[14] R. I. Hartley and P. Sturm, "Triangulation," *Computer Vision and Image Understanding*, vol. 68, no. 2, pp. 146–157, Nov. 1, 1997. DOI: 10.1006/cviu.1997.0547.

[15] P.-E. Sarlin, D. DeTone, T. Malisiewicz, and A. Rabinovich, "SuperGlue: Learning feature matching with graph neural networks," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2020, pp. 4937–4946. DOI: 10.1109/CVPR42600.2020.00499.

[16] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "NeRF: Representing scenes as neural radiance fields for view synthesis," *Commun. ACM*, vol. 65, no. 1, pp. 99–106, 2021. DOI: 10.1145/3503250.

[17] B. Pan, K. Qian, H. Xie, and A. Asundi, "Two-dimensional digital image correlation for in-plane displacement and strain measurement: A review," *Measurement Science and Technology*, vol. 20, no. 6, p. 062 001, Apr. 2009. DOI: 10.1088/0957-0233/20/6/062001.

[18] M. N. Helfrick, C. Niezrecki, P. Avitabile, and T. Schmidt, "3d digital image correlation methods for full-field vibration measurement," *Mechanical Systems and Signal Processing*, vol. 25, no. 3, pp. 917–927, Apr. 1, 2011. DOI: 10.1016/j.ymssp.2010.08.013.

[19] P. Lindenberger, P.-E. Sarlin, and M. Pollefeys, "LightGlue: Local feature matching at light speed," in *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, Oct. 2023, pp. 17 581–17 592. DOI: 10.1109/ICCV51070.2023.01616.

[20] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, Nov. 1, 2004. DOI: 10.1023/B:VISI.0000029664.99615.94.

[21] G.-W. Yang, W.-Y. Zhou, H.-Y. Peng, D. Liang, T.-J. Mu, and S.-M. Hu, "Recursive-NeRF: An efficient and dynamically growing NeRF," *IEEE Transactions on Visualization and Computer Graphics*, vol. 29, no. 12, pp. 5124–5136, Dec. 2023. DOI: 10.1109/TVCG.2022.3204608.

[22] S. J. Garbin, M. Kowalski, M. Johnson, J. Shotton, and J. Valentin, "FastNeRF: High-fidelity neural rendering at 200fps," presented at the 2021 IEEE/CVF International Conference on Computer Vision (ICCV), IEEE Computer Society, Oct. 1, 2021, pp. 14 326–14 335. DOI: 10.1109/ICCV48922.2021.01408.

[23] D. Rebain, W. Jiang, S. Yazdani, K. Li, K. M. Yi, and A. Tagliasacchi, "DeRF: Decomposed radiance fields," in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2021, pp. 14 148–14 156. DOI: 10.1109/CVPR46437.2021.01393.

[24] R. I. Hartley and A. Zisserman, *Multiple view geometry in computer vision*, 2nd ed. Cambridge University Press, ISBN: 0521540518, 2004.

[25] J. Javh, J. Slavič, and M. Boltežar, "Measuring full-field displacement spectral components using photographs taken with a DSLR camera via an analogue fourier integral," *Mechanical Systems and Signal Processing*, vol. 100, pp. 17–27, Feb. 1, 2018. DOI: `10.1016/j.ymssp.2017.07.024`.