Chapter 3
Towards Experimental Modal Analysis of Structural Systems with High-Speed X-Ray Radiography

Anthony LoRe Starleaf and David Mascareñas

Abstract This work introduces a novel data analysis technique for performing 3D modal analysis of structures with optically inaccessible internal geometry using radiographic data acquired via high-speed X-ray imaging systems. We reconstruct the static configuration of the structure by employing Neural Radiance Field (NeRF)-inspired inverse rendering, resulting in a volumetric representation defined by a three-dimensional grid of attenuation coefficients. To characterize the structure's response, we employ an isoparametric finite-element (FE) framework. The core innovation lies in directly learning deflection vectors for each time step using a fixed-point iterative grid warping algorithm, which estimates nodal deflections from dynamic radiographs. By comparing reconstructed radiographs against actual captured frames, the deflection estimates are refined until they implicitly represent the true response of the structure (up to the fidelity of the FE mesh). The final deflection estimates serve as inputs to a simple Rayleigh quotient–based modal analysis procedure, enabling accurate extraction of modal parameters. It should be noted that the estimation of modal parameters is entirely decoupled from the displacement measurement algorithm. There is a wide range of literature which addresses the problem of obtaining modal parameters from FE response data; in principle, any of these methods could be used with our approach. We selected the Rayleigh quotient due to its simplicity and straightforward implementation. Finally, the proposed method offers significant advantages for non-contact, high-resolution modal analysis. By considering X-ray radiography, we overcome a key limitation of traditional vision-based modal analysis – the requirement that the structure's surfaces must be externally visible.

Introduction

Experimental modal analysis (EMA) is a widely used technique for updating models of civil, mechanical, and aerospace structures [1]. Typically, EMA is performed by applying contact sensors, such as accelerometers and strain gauges, to the surface of a structure, then exciting the structure with a known force and recording its response. While contact sensors are well-studied and reliable, they come with a few limitations. First, when applied to thin or light (low mass) structures, the sensors themselves significantly impact the dynamics. Second, they can be expensive and time-consuming to apply. Third, they only provide pointwise measurements of the response, necessitating a large number of them to characterize the behavior of the structure.

In response to the limitations of contact sensors, several vision-based approaches to EMA have been proposed. Phase-based motion magnification and mode shape identification [2, 3] is a technique that extracts mode shapes and frequencies from video footage of vibrating structures. The authors of [4] proposed an x-ray-based digital volume correlation (DVC) method for extracting motion from stereo x-ray images. Recent work from Sandia National Laboratory [5] proposed the QuintX system, which is a 5-view stereo x-ray imaging system designed for EMA. Inspired by this work, we seek to develop a computational method of extracting full-field deflection measurements of a vibrating structure from dynamic, multi-view x-ray images.

Anthony LoRe Starleaf · David Mascareñas
e-mail: aglor@lanl.gov; dmascarenas@lanl.gov

Kristin B. Zimmerman, PhD (eds.), *Computer Vision and SLDV for Structural Dynamics, 2026, Vol. 6*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.1007/978-8-743-81445-0_3

Methods

Our methodology is broken into three major steps: (1) static volume reconstruction, (2) deflection field characterization, and (3) system identification. In the first two steps, we represent the attenuation field associated with the object of interest by a 3D grid of attenuation values (Figure 2). In the static reconstruction phase, we use the Adam optimizer [6] to iteratively update the 3D attenuation grid until radiographs rendered by our differentiable x-ray rendering engine match those captured experimentally. This results in a volumetric representation of the attenuation field associated with the structure of interest which implicitly encodes its geometry. In the deflection field characterization phase, we represent structural deformations using nodal displacement vectors on a finite element (FE) mesh. Using a differentiable grid warping procedure, we deform the static attenuation field according to the deflection field specified by the nodal displacement vectors and render the new volume. Again, the Adam optimizer is used to find a set of nodal displacements that minimizes the difference between the rendered and experimentally captured set of x-ray images at each time-step. This results in a set of nodal deflections for each time-step that can be input into a system identification algorithm. A flowchart describing this process is shown below (Figure 1).

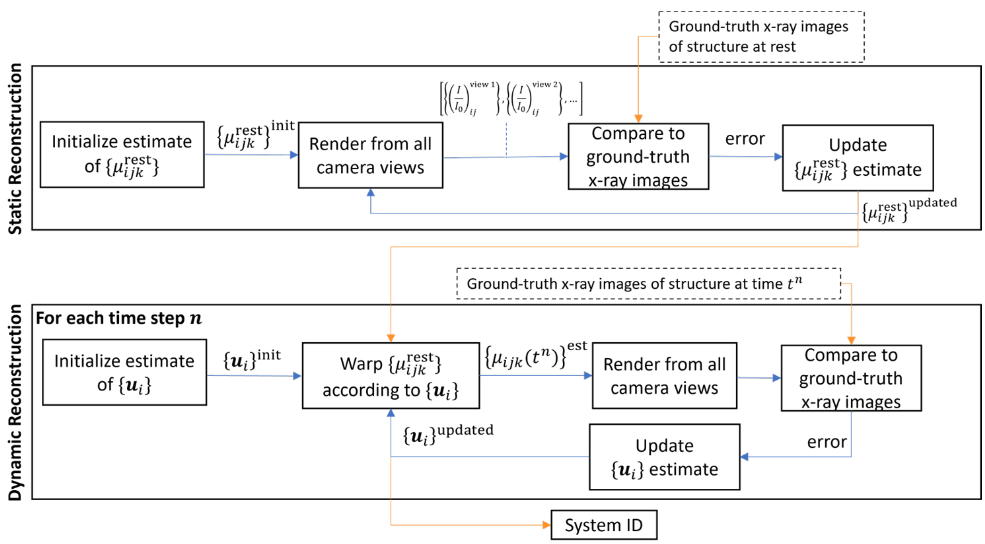

Fig. 1 Flowchart describing our methodology. Blue arrows represent the flow of information within a single step of our procedure. Orange arrows represent the flow of information between steps. Gradients are not propagated along orange arrows. The remainder of this section discusses each component of the process in more detail.

3D Voxel Grid Representation

We represent the structure's attenuation field using a 3D grid of attenuation coefficients μ_{ijk}, interpolated according to the standard trilinear interpolation formula for a uniform grid given by

$$\overline{\mu}\left(x,y,z\right)=\sum_{ijk}w_{ijk}\left(x,y,z\right)\mu_{ijk}$$

$$w_{ijk}\left(x,y,z\right)=\frac{1}{\Delta x\Delta y\Delta z}\begin{cases}|x-x_{ijk}|\cdot|y-y_{ijk}|\cdot|z-z_{ijk}| \text{ if } \left\|\left(\boldsymbol{x}-\boldsymbol{x}_{ijk}\right)/\Delta\boldsymbol{x}\right\|_{\infty}\leq 1\\ 0 \text{ otherwise}\end{cases}$$

where $\|\cdot\|_{\infty}$ denotes the supremum norm. A visualization of this interpolation procedure is given below in Figure 2. The grid is represented in 2D for the sake of visual clarity.

Applying this interpolation formula allows a continuous function $\mu:\mathbb{R}^3\rightarrow\mathbb{R}$ to be represented by a discrete set of points $\{\mu_{ijk}\}_{0,0,0}^{N_x,N_y,N_z}$. Similarly to Figure 2, Figure 3 below shows a 2D analog of this procedure.

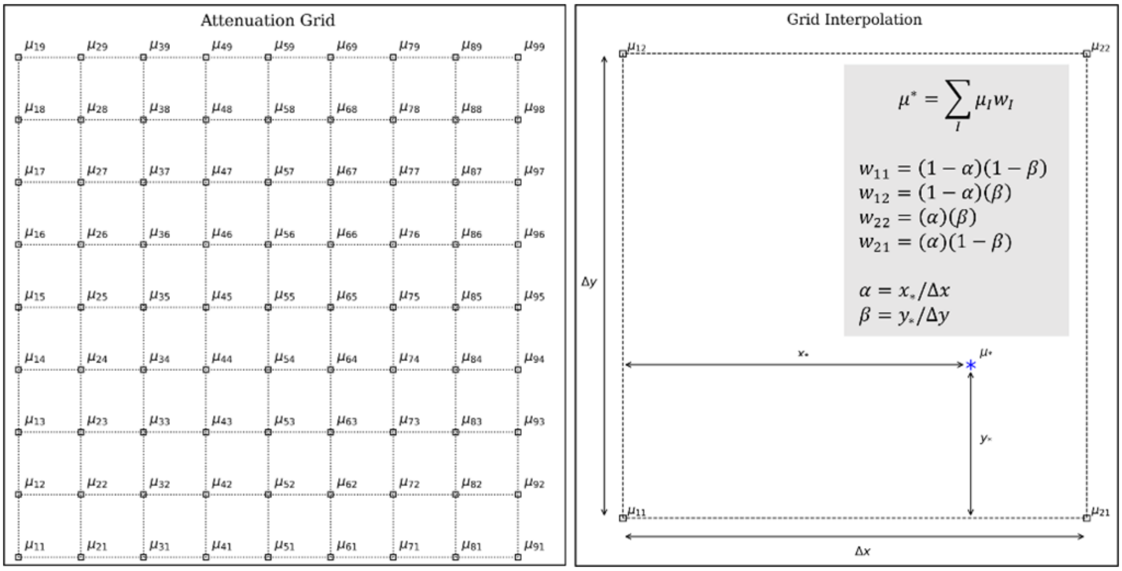

Fig. 2 Left: 2D representation of the 3D voxel grid. Right: bilinear interpolation (2D analog of trilinear interpolation) formula. In 3D, the grid would have an additional axis extending into/out of the page.

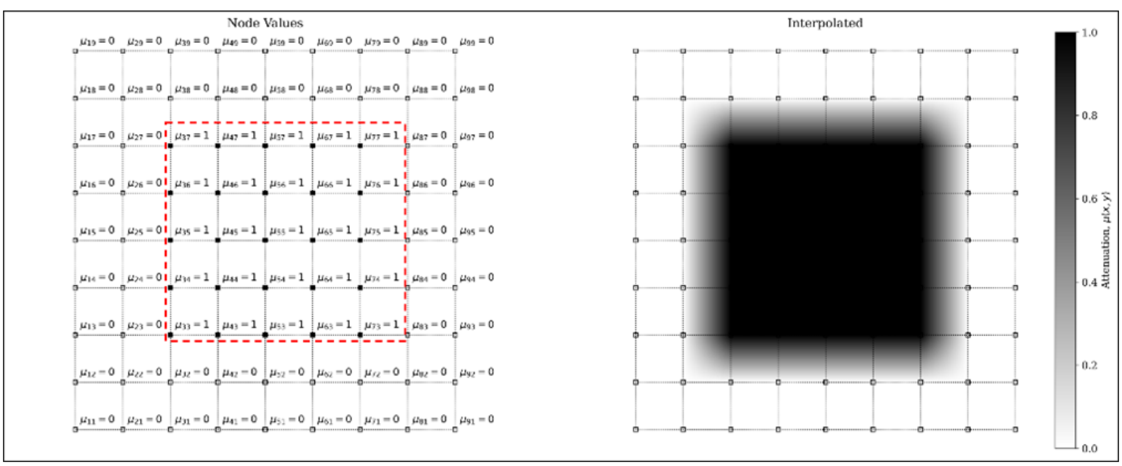

Fig. 3 Left: 2D voxel grid; The red square encloses nodes whose attenuation values are set to 1. Right: The interpolated continuous attenuation field $\mu(x, y)$.

Differentiable X-Ray Rendering

Since we use gradient-based optimization to perform both the static reconstruction and deflection field characterization, it is important for the x-ray rendering procedure to be differentiable. Fortunately, our parameterization of the attenuation field is continuous with respect to the node values μ_{ijk}. To render the attenuation field, we simulate an isotropic x-ray source casting rays through the voxel grid and striking a scintillator plate with a photodiode array behind it. For simplicity, we do not differentiate between the photodiode array and scintillator plate, and assume that the scintillator plate itself has pixels whose intensities can be recorded to form an image. A 2D analog of the rendering configuration is shown in Figure 4.

In this configuration, x-rays are emitted by the source at a uniform intensity I_0 and are attenuated by the object (represented by the 3D voxel grid). The rays then strike a scintillator plate/pixel array with received intensity I. The relationship between I, I_0, and μ_{ijk} is governed by Beer's Law.

$$\log \frac{I}{I_0} = - \int \mu\left(x, y, z; \mu_{ijk}\right) ds$$

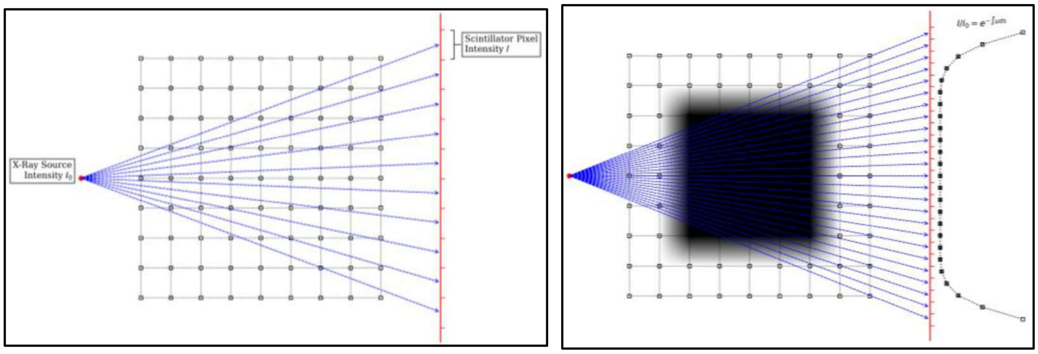

Fig. 4 Left: 2D analog of the rendering configuration. In 3D, the scintillator plate would be a bounded 2D plane. Right: Example of a 1D radiograph rendered from a 2D attenuation grid. The rightmost graph shows received intensity I/I_0 as a function of pixel position.

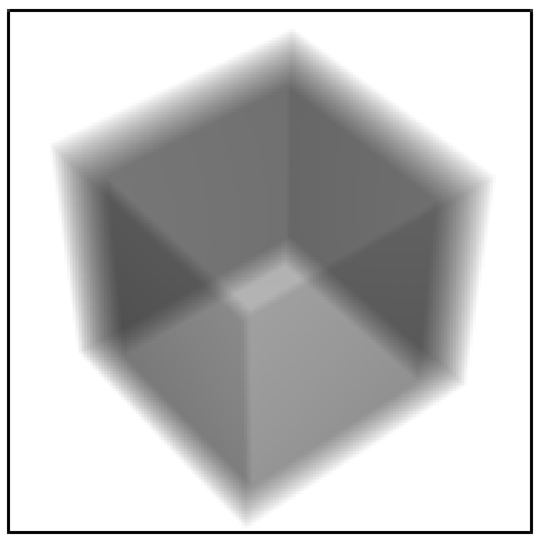

Fig. 5 Example of a 2D radiograph rendered from a 3D attenuation field.

where ds indicates that the integral is carried out lengthwise along each ray. Applying this procedure to a 3D volume would yield a 2D array of values $\left(\frac{I}{I_0}\right)_{ij}$ where ij indexes pixels. An example of such an image is shown in Figure 5. This is a simplification of the method of x-ray rendering employed in [7]. As is evident from the lack of dependency on wavelength in the equations above, we assume monochromatic x-ray generation.

Differentiable Grid Warping

The concept of "grid warping" may at first seem rather unintuitive. Its purpose in our methodology is to allow the statically reconstructed attenuation field representing the structure's rest configuration to be deformed according to estimates of the structure's deflection field at a given time t. Since the attenuation field is represented by a 3D voxel grid, the most straightforward way to represent the new configuration of the structure at time t is to construct a new voxel grid μ^t_{ijk} such that

$$\mathrm{Interp}\left(\mu^t_{ijk}, \left[\begin{array}{ccc} x+u^t(x,y,z) & y+v^t(x,y,z) & z+w^{\ t}(x,y,z)\end{array}\right]\right) = \mathrm{Interp}(\mu_{ijk}, \left[\begin{array}{ccc} x & y & z \end{array}\right])$$

where 'Interp' represents the interpolation operator. The central challenge in accomplishing this is determining the original position (i.e. where the node was located before the deformation was applied) of each node in the new grid μ^t_{ijk}. Once the original position $(x_{ijk}, y_{ijk}, z_{ijk})$ of each node in the new grid has been determined, the nodal attenuation values μ^t_{ijk} can be obtained by interpolating in the original grid at $(x_{ijk}, y_{ijk}, z_{ijk})$. Figure 6 demonstrates this idea in 2D.

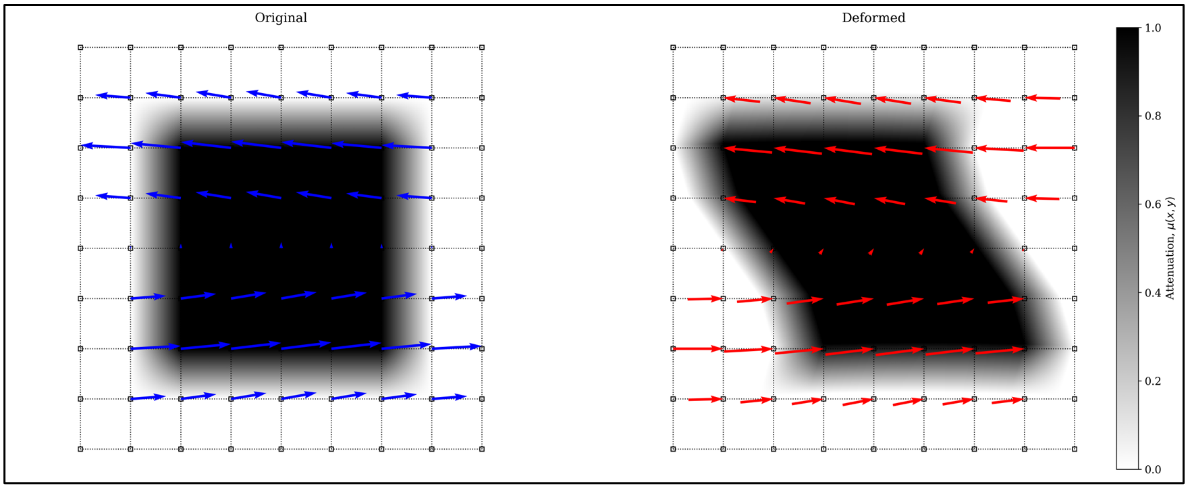

Fig. 6 Left: statically reconstructed grid μ_{ij} with blue arrows defining a forward deformation field for the object. Right: new grid μ_{ij}^t at time t representing the object after deformation. Red arrows point from the original position of each grid node to its current position.

In order to determine the original positions (x, y, z) of points in deformed space $(X, Y, Z) = (x, y, z) + (u^t, v^t, w^t)$, we apply an α-damped fixed-point iteration [8] with the mapping

$$g(\boldsymbol{x}) = (1 - \alpha)\,\boldsymbol{x} + \alpha\,(\boldsymbol{X} - \boldsymbol{\delta}(\boldsymbol{x}))$$

$$\boldsymbol{x}_{\text{est}}^{k+1} = g(\boldsymbol{x}_{\text{est}}^{k})$$

where $\boldsymbol{x} = (x, y, z)$, $\boldsymbol{X} = (x + u^t(\boldsymbol{x}), y + v^t(\boldsymbol{x}), z + w^t(\boldsymbol{x}))$, and $\boldsymbol{\delta}(\boldsymbol{x}) = (u^t(\boldsymbol{x}),\ v^t(\boldsymbol{x}),\ w^t(\boldsymbol{x}))$. This is guaranteed to converge as long as g is a contraction on $\mathbb{R}^3$. That is, as long as g satisfies

$$\|g(\boldsymbol{x}) - g(\boldsymbol{y})\|_2 \leq \|\boldsymbol{x} - \boldsymbol{y}\|_2 \qquad \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^3$$

In the $\alpha = 0$ case, $g(\boldsymbol{x}) = \boldsymbol{x}$ and convergence is guaranteed. However, this does not incorporate any knowledge of the deformation field $\boldsymbol{\delta}(\boldsymbol{x})$ and will "converge" to the wrong solution unless $\boldsymbol{x}_{\text{est}}^0 = \boldsymbol{x}$. In practice, we initialize $\boldsymbol{x}_{\text{est}}^0 = \boldsymbol{X}$, so this only works if $\boldsymbol{\delta}(\boldsymbol{x}) = \boldsymbol{0}$. In the $\alpha = 1$ case, the iteration only converges if $\boldsymbol{\delta}(\boldsymbol{x})$ is a contraction on $\mathbb{R}^3$, which is usually the case. In our implementation, we initialize α to 1 and halve it when $\boldsymbol{x}_{\text{est}}^{k+1}$ stops improving over $\boldsymbol{x}_{\text{est}}^{k}$. Importantly, this grid warping procedure is differentiable with respect to $\boldsymbol{\delta}(\boldsymbol{x})$.

We use PyTorch's backward-mode auto-differentiation to perform all gradient-based optimizations discussed in this paper. Backward-mode auto-differentiation stores intermediate results of all computations to be used for computing derivatives. This means that the memory usage associated with our procedure grows linearly with the number of fixed-point iterations needed to deform the grid. Empirically, we notice that the fixed-point iteration procedure tends to converge in approx. 10-20 iterations. This is not enough to cause a problem on most mid-grade consumer GPUs (i.e. the NVIDIA GeForce RTX 3070 Ti Laptop GPU) but may cause issues if the method is applied with very fine voxel grids or highly irregular deflection fields.

System Identification

Once the nodal displacement vector $\boldsymbol{u}(t)$ of the FE mesh has been estimated for each time step, we use a version of the Rayleigh quotient to obtain estimates of the quantity E/ρ. For completeness, we provide a short derivation of the quotient below. Begin with the discrete equations of motion for the structure in free vibration.

$$\boldsymbol{M}\ddot{\boldsymbol{u}}(t) + \boldsymbol{K}\boldsymbol{u}(t) = \boldsymbol{0}$$

Assuming material homogeneity gives $\boldsymbol{M} = \rho\widehat{\boldsymbol{M}}$ and $\boldsymbol{K} = E\widehat{\boldsymbol{K}}$. Rewriting,

$$\widehat{\boldsymbol{M}}\ddot{\boldsymbol{u}}(t) + \frac{E}{\rho}\widehat{\boldsymbol{K}}\boldsymbol{u}(t) = \boldsymbol{0}$$

Applying the Fourier transform and denoting the Fourier transform of $\boldsymbol{u}(t)$ by $\boldsymbol{\phi}(\omega)$,

$$\frac{E}{\rho}\widehat{\boldsymbol{K}}\boldsymbol{\phi}\left(\omega\right)-\omega^{2}\widehat{\boldsymbol{M}}\boldsymbol{\phi}\left(\omega\right)=\mathbf{0}$$

Moving the second term to the right and left multiplying by $\boldsymbol{\phi}^{T}$,

$$\frac{E}{\rho}\boldsymbol{\phi}^{T}\widehat{\boldsymbol{K}}\boldsymbol{\phi}=\omega^{2}\boldsymbol{\phi}^{T}\widehat{\boldsymbol{M}}\boldsymbol{\phi}$$

Finally, this may be rearranged to obtain

$$\frac{E}{\rho}=\frac{\boldsymbol{\phi}^{T}\widehat{\boldsymbol{M}}\boldsymbol{\phi}}{\boldsymbol{\phi}^{T}\widehat{\boldsymbol{K}}\boldsymbol{\phi}}\omega^{2}$$

In practice, we obtain the modal frequencies ω_i from the time-series values of $q_i(t)$ estimated in the dynamic reconstruction phase. We then estimate E/ρ according to the following square-amplitude-weighted average.

$$\frac{E}{\rho}=\frac{\sum_{i}\left(\frac{\boldsymbol{\phi}_{i}^{T}\widehat{\boldsymbol{M}}\boldsymbol{\phi}_{i}}{\boldsymbol{\phi}_{i}^{T}\widehat{\boldsymbol{K}}\boldsymbol{\phi}_{i}}\omega_{i}^{2}\right)\left(\max_{t}q_{i}^{2}\left(t\right)\right)}{\sum_{i}\max_{t}q_{i}^{2}\left(t\right)}$$

It should be noted that once the nodal deflections have been determined in the dynamic reconstruction step, any method of system identification may be used. We only use this one because it is simple.

Results & Discussion

We applied the above procedure to a synthetic experiment involving a hollow 3D box in free vibration. For the sake of expediency in evaluating this approach, an excessive number of imager views and samples were used. In the future we intend to repeat this experiment with fewer views and a lower sampling rate. Table 1 defines numerical values for several of the parameters which define our experiment.

Table 1 Numerical values of experimental parameters.

Experiment Parameter	Value	Units
Number of Imagers	32	-
Sampling Rate	3.33	kHz
Elastic Modulus, E	70	GPa
Density, ρ	2700	kg·m^3
Poisson's Ratio, ν	0.3	-
Structure Length/Width/Height	1.0 / 1.0 / 1.0	m
Voxel Grid Size	[128, 128, 128]	-
Voxel Grid Resolution	10.94	mm

Synthetic Data Generation

To validate our procedure, we applied it to a simulated hollow box in free vibration. A hollow box was selected because it is a simple structure, and the cavity in the middle allows us to leverage the x-ray imagers' ability to see internal surfaces. The box was initially deformed, then released and allowed to vibrate freely. Figure 7 shows the box in its rest configuration, as well as its shape at $t = 0$.

The box was meshed with linear tetrahedral (TET4) elements, and its motion was simulated according to the standard damping-free FE equation $\boldsymbol{M}\ddot{\boldsymbol{u}}\left(t\right)+\boldsymbol{K}\boldsymbol{u}\left(t\right)=\mathbf{0}$ using an implicit Newmark-Beta scheme with $\Delta t = 3\times10^{-4}$. Figure 8 shows the configuration of x-ray imagers around the box.

The rendering procedure described in Section II-B was used to generate simulated x-ray images of the box in its rest configuration (Figure 8), and in its deformed configuration at each time step.

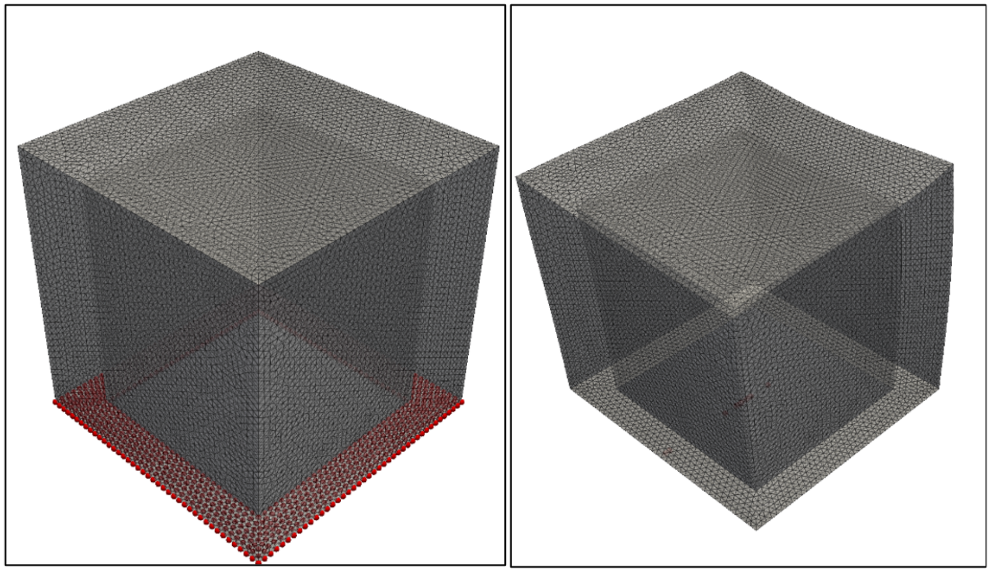

Fig. 7 Left: Resting (unloaded) configuration of the box. Red dots denote fixed nodes. Right: Deformed configuration of the box at time $t = 0$.

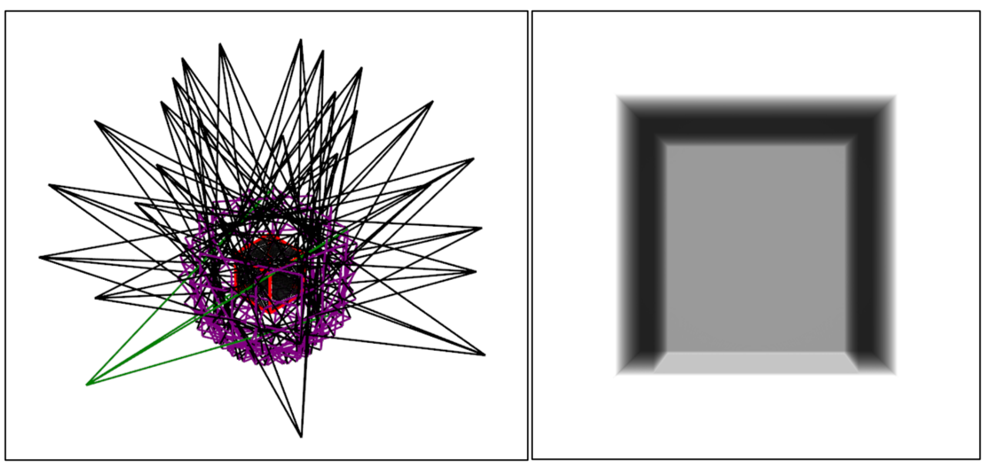

Fig. 8 Left: Imaging configuration. Pyramids represent imagers with x-ray sources at the tip and scintillator plates (purple) at the base. Red cube represents the voxel grid used to render images. Right: Radiograph rendered from x-ray imager picture in green in the left image.

Static Reconstruction

The static reconstruction procedure described in Figure 1 was applied to the simulated radiographs of the box in its rest configuration. Figure 9 shows two novel viewpoints of the reconstructed static attenuation grid, demonstrating that the geometry was recovered effectively.

Deflection Field Characterization / Dynamic Reconstruction

When attempting to learn the nodal displacement vector $\boldsymbol{u}(t)$ directly, we quickly noticed that the optimizer converges to high-spatial-frequency local minima that do not accurately reflect the behavior of the structure. As such, some degree of spatial smoothing was required in order to allow the optimization to converge to a set of nodal displacements that effectively characterized the motion of the structure. Since the dynamics of linear structures are typically sparse with respect to a modal basis, we assume that $\boldsymbol{u}\left(t\right) = \sum_{i=1}^{8} \boldsymbol{\phi}_i q_i(t)$ where $\boldsymbol{\phi}_i$ are the first 8 modes of the structure. For homogeneous structures, the mode shapes can be obtained with knowledge only of the structure's geometry. Note that these mode shapes are only used

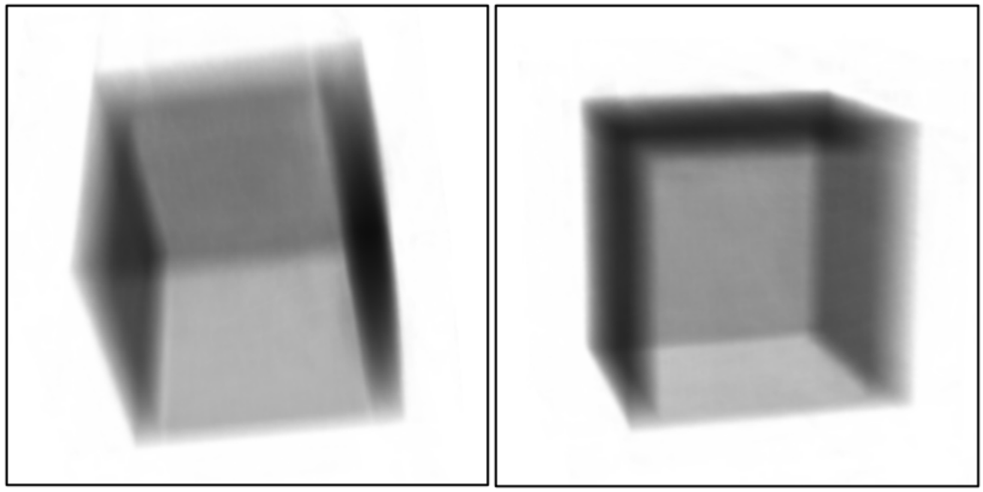

Fig. 9 Images rendered from the reconstructed attenuation field. Neither viewpoint was present in the data. Left: camera positioned beneath the box, looking up. Right: camera positioned beside the box, looking slightly down.

to smooth the learned displacement fields. In principle, for an inhomogeneous structure, the mode shapes corresponding to a materially homogeneous structure of the same shape could theoretically serve the same purpose, so long as enough of these modes are included to describe the observed dynamics of the structure.

Figure 10 below compares the true deflection at two timesteps to the deflection recovered using our method. Figure 11 performs a similar comparison using the modal coordinates of the structure.

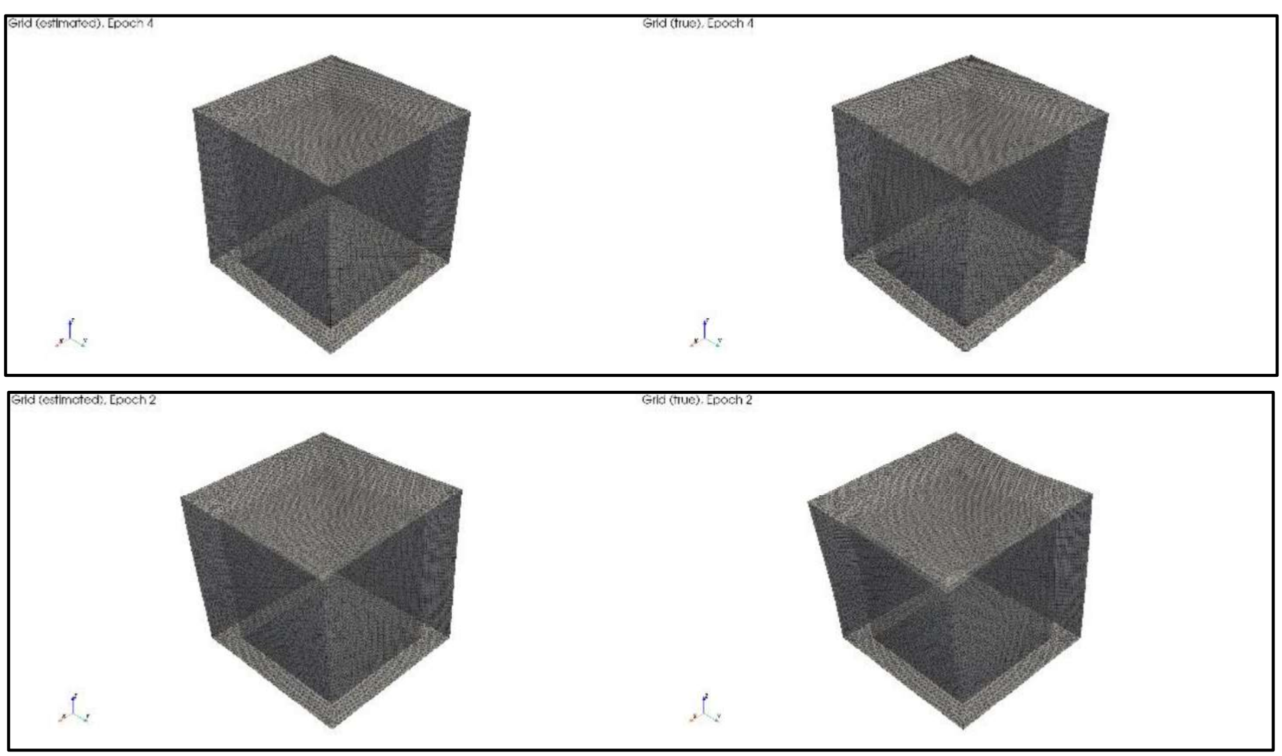

Fig. 10 True deflection (right) vs. recovered deflection (left) at timesteps 6 (top) and 17 (bottom).

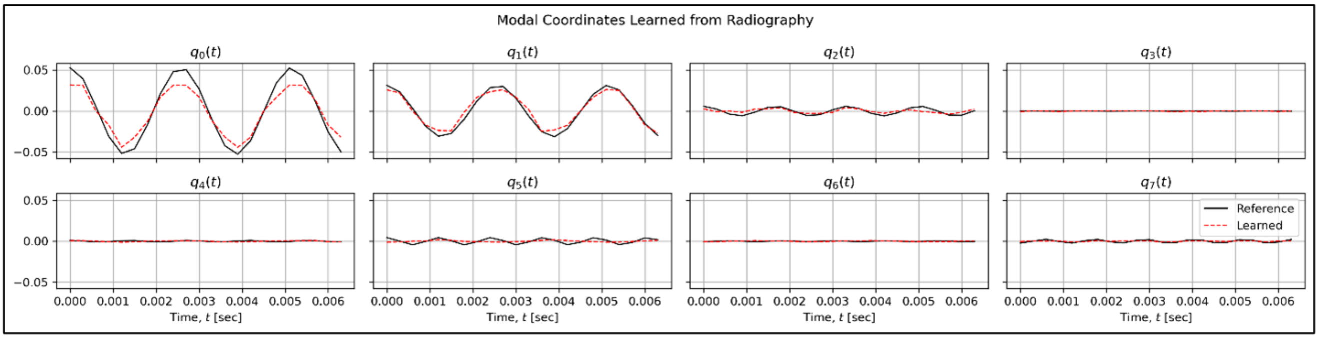

Fig. 11 True vs. recovered modal coordinates. Black line represents modal coordinates obtained by projecting the simulated nodal deflection history onto a modal basis. Red line represents modal coordinates recovered using our method.

System Identification

Having obtained the structure's modal coordinates at each time step, we apply the Rayleigh quotient method described in Section II-D to estimate the structure's Young's modulus (E), assuming that we know its density (ρ). We obtained $E_{\text{est}} = 67.16$ GPa. This is only a 4.05% difference with respect to the true value of $E_{\text{true}} = 70.00$ GPa.

Conclusion & Future Work

We have developed a framework for characterizing dynamic structural deformations from time-series radiography data. Additionally, we have demonstrated the framework's effectiveness on a synthetic experiment, assuming material linearity and homogeneity. While these initial results are promising, further work is still required to develop the framework to the point where it can be used on real experiments. Namely, steps must be taken to reduce the number of imager views necessary to characterize the displacement field, and to reduce the frequency at which these imagers must fire.

Due to time constraints, we did not experiment with different numbers of imagers and their placements. We expect that the current framework would likely work just as well with far fewer imager positions and intend to investigate this as future work. However, with fewer imagers present, it would likely become increasingly important to optimize their positioning such that they can observe the dynamics of the structure.

We believe that the second point may be addressed by introducing a Sinusoidal Representation Network (SIREN) to represent the modal coordinate vector. In the current procedure, the deflection vector associated with a given time step is only informed by images captured at that time step. By implicitly representing the deflection vector as a function of time using a SIREN, images captured at all time steps could influence the predicted nodal deflection vector at all other timesteps. This would introduce some complications in training the neural network, such as memory consumption and gradient variance (which we observed when we tried to implement this idea ourselves). However, if these issues could be effectively addressed, it would likely succeed in reducing the required firing rate of the x-ray sources by allowing for asynchronous imaging. For example, suppose there were 10 x-ray sources, each firing at 10 Hz. The sample overall sample rate could be achieved by firing each source at 1 Hz, with each source firing 0.1 seconds after the last. As an additional bonus, this method of representing nodal displacements would allow for less strict synchronization between imagers.

Finally, the underlying method of 3D representation proposed in this work shares many similarities with NeRF-style methods, specifically Direct Voxel Grid Optimization [9]. Our approach (with a few modifications) would likely work quite well with visible-light imagers for tasks that do not require knowledge of the behavior of internal structures.

References

1. S. W. Doebling, C. R. Farrar and M. B. Prime, "A summary review of vibration-based damage identification methods," *Shock and vibration digest*, pp. 91-105, 1998.
2. Y. Yang, C. Dorn, T. Mancini, Z. Talken, G. Kenyon, C. Farrar and D. Mascareñas, "Blind identification of full-field vibration modes from video measurements with phase-based video motion magnification," *Mechanical Systems and Signal Processing*, pp. 567-590, 2017.
3. A. S. Poozesh, Z. Mao, C. Niezrecki and Peyman, "Vibration-based damage detection in wind turbine blades using Phase-based Motion Estimation and motion magnification," *Journal of Sound and Vibration*, pp. 300-318, 2018.

4. M. Southwick, Z. Mao and C. Niezrecki, "Volumetric Motion Magnification: Subtle Motion Extraction from 4D Data," *Measurement*, p. 109211, 2021.
5. D. P. Rohe, J. Miers and Z. M. P. Avitabile, "Seeing Beyond the Surface – Structural Dynamics Identification through Multi-View Radiographic Imaging," in *International Modal Analysis Conference*, 2025.
6. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
7. F. P. Vidal and P. F. Villard, "Development and validation of real-time simulation of X-ray imaging with respiratory motion," *Computerized Medical Imaging and Graphics*, pp. 1-15, 2016.
8. J. D. Frankel and S. Hoffman, "Numerical Methods for Engineers and Scientists," in *Numerical Methods for Engineers and Scientists*, New York, CRC Press, 2001, pp. 141-145.
9. C. Sun, M. Sun and H.-T. Chen, "Direct Voxel Grid Optimization: Super-fast Convergence for Radiance," CoRR, 2021.