Supplementary Material for Visual Vibration Tomography

Contents

1. Supplementary

1.1. Optimization Runtime

Recall that the optimal material properties and 3D modes are found by iteratively updating the solutions to the following optimization problem:

$$w^*, v^* = \underset{\substack{w, v \in \mathbb{R}^m \\ K, M \in \mathbb{R}^{n \times n} \\ u_i \in \mathbb{R}^n, i=1,\dots,k}}{\arg \min} \left\{ \frac{1}{2k} \sum_{i=1}^{k} y_i \left\| K u_i - \widehat{\omega}_i^2 M u_i \right\|_2^2 \right. \tag{1}$$

$$+ \frac{\alpha_u}{2k} \sum_{i=1}^{k} \| P u_i - \widehat{\gamma}_i \|_2^2$$

$$\left. + \frac{\alpha_w}{2m} \left\| \nabla^2 w \right\|_2^2 + \frac{\alpha_v}{2m} \left\| \nabla^2 v \right\|_2^2 + \left(\sum_{e=1}^{m} w_e/m - \bar{w} \right)^2 \right\}$$

$$\text{s.t. } K = \sum_{e=1}^{m} w_e K_e, \quad M = \sum_{e=1}^{m} v_e M_e.$$

For an 8x8x8 cube with linear hexahedral elements, one iteration takes 2–3 seconds (tested on an 8-core Intel Core i9, 32 GB RAM), and convergence usually happens within 100 iterations.
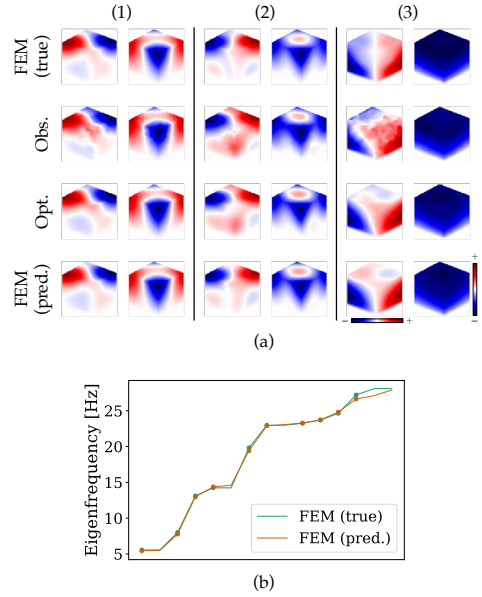

Figure 1. (a) Similarity of predicted image-space modes to true, observed, and optimized image-space modes for the cube sample shown in Fig. 3. "Obs." mode is often a noisy version of "FEM (true)." "Opt." refers to the optimized solution U^* in Eq. 1. "FEM (pred.)" is the image-space mode resulting from the estimated material properties. (b) Predicted eigenfrequencies vs. true eigenfrequencies. The frequencies of the 10 given motion-extracted image-space modes are marked by scatter dots.

1.2. Simulated Experiments

1.2.1 Predicted Image-Space Modes

In addition to normalized correlation, a way to assess estimated material properties is to verify that they produce the same image-space modes and natural frequencies as the true properties. Recall that 3D modes are a decision variable in our optimization scheme (Eq. 1). As Fig. 1 shows, it is informative to compare the true FEM modes, observed modes, optimized modes, and predicted FEM modes, in image-space. The optimization process usually de-noises observed modes. For some modes, spatially correlated noise may make it difficult to recover the true mode, but it is possible for the predicted FEM modes to still be similar to the truth (see example (2) in Fig. 1).

1.2.2 Effect of Regularization

The strength of spatial regularization on material properties affects the smoothness of the estimation. In Eq. 1, the regularization weights for Young's modulus and density are α_w and α_v, respectively. In Fig. 2, we show how the estimated density image becomes sharper as α_v decreases (keeping all else fixed). This can result in a crisper picture of the defect, but can also make reconstruction more sensitive to noise.
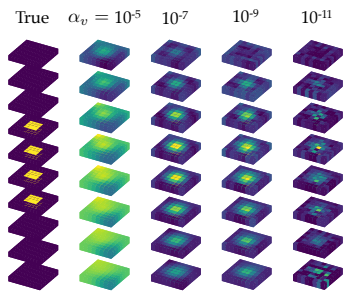

Figure 2. Effect of regularization. Here we show the estimated density for a cube sample (same as in Fig. 3). Keeping all else fixed, as α_v decreases, the image of the defect becomes crisper, but more sensitive to noise. Each estimation uses the same 10 motion-extracted image-space modes.

1.2.3 Intrinsic Resolution

Fig. 3 demonstrates how, as the number of input image-space modes increases, the intrinsic resolution of the reconstruction improves.
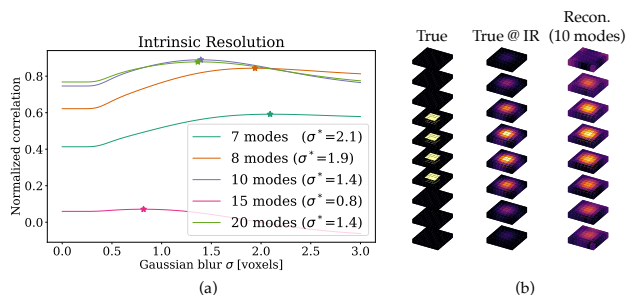

Figure 3. Intrinsic resolution of reconstructed volumes. Based on normalized correlation with the ground-truth material properties smoothed at different scales, one can approximate the intrinsic resolution of the reconstructed material properties. In (a), we plot normalized correlation versus Gaussian blur standard deviation σ, for the reconstruction of Young's modulus using different numbers of image-space modes (keeping all other hyperparameters fixed). As the number of observed modes increases, the reconstructed resolution also increases (i.e., smaller σ). (b) shows the true Young's modulus image blurred at the intrinsic resolution (IR) of the reconstruction given 10 image-space modes ($\sigma^* = 1.4$ voxels).

1.3. Simulated Experiments: Model Mismatch

In the main material, we discussed geometry mismatch. Here, we investigate the effects of model mismatch in the Poisson's ratio and the mesh element order.

1.3.1 Poisson's Ratio Mismatch

The Poisson's ratio is a measure of the deformation of a material perpendicular to an applied force and ranges from 0.1 to 0.45 [1]. Since our optimization formulation only estimates Young's modulus and density, we assume that every voxel has the same Poisson's ratio. Fig. 4 shows that this assumption does not significantly hurt the reconstruction of a defect, especially when assuming a Poisson's ratio that is closer to that of the main material.
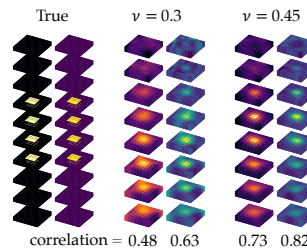

Figure 4. The effect of assuming a homogeneous Poisson's ratio. In the true cube, the main material has a Poisson's ratio of $\nu = 0.45$, while the defect material has a Poisson's ratio of 0.3 (roughly corresponding to the values for Jello and clay, resp.). When inferring material properties, we assume a uniform Poisson's ratio across the entire cube. We find that this assumption does not hurt the reconstruction much, especially when ν is set to the Poisson's ratio of the main material. Both estimations use the same 20 motion-extracted image-space modes.

1.3.2 Mesh Element Order Mismatch

In general, it is better to use higher-order elements to model a real-life object. However, there is a tradeoff in efficiency. A quadratic element approximates node displacements more accurately, but contains more DOFs than a linear element. Fig. 5 shows what happens when the forward model uses quadratic elements, while the inference model uses linear elements. For real-world objects, one should choose the order that strikes the right balance between approximation accuracy and computational cost.

1.4. Drum Experiment

1.4.1 Drum Construction

The drums were constructed by fixing a thin rubber sheet over a 4"x4" PVC adaptor with a rubber band. We tested defects of two materials: nail hardening gel and acrylic plastic
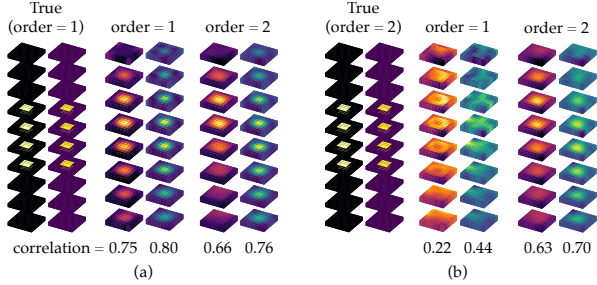
Figure 5. Element order mismatch. As (a) shows, when using linear elements in the forward model, any element order ≥ 1 suffices in the inference model. However, in (b), reconstruction quality degrades when attempting to model a quadratic-element cube with linear elements.

circles. For each defect, we recorded a video of the homogeneous drum before the defect was applied for comparison. We drew a speckle pattern on the drum head for texture.

1.4.2 Vibration-Capture Setup

Fig. 6 shows a schematic of the setup. We taped the drum onto an optical table, with the high-speed camera standing on the same optical table. The excitation source was a PreSonus Sceptre S8 loudspeaker, which sat on a platform separate from the optical table and was pointed at the drum. For each video, we recorded the drum head's vibration in response to a 3.5-second linear frequency sweep (50–1000 Hz) played by the speaker.
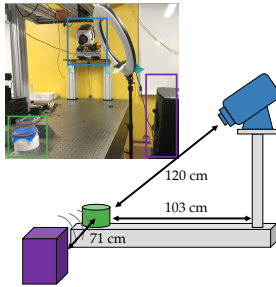

Figure 6. Experimental setup for real drums. Vibrations were induced by a loudspeaker and recorded with a high-speed camera.

1.4.3 Video Capture

Our camera was a Phantom V1610 high-speed camera. Each video was captured at 6000 FPS at an image resolution of 288×384. To reduce noise, we averaged every two frames for a resulting temporal frequency of 3000 FPS. Note that in Fig. 9 in the main material, the drums vibrate at frequencies below 120 Hz. While we chose to first demonstrate our approach using a high-speed camera, where com-

pression and camera noise are less challenging, many modal frequencies can be captured on a consumer camera.

1.4.4 Extracting Image-Space Modes

We found that in real videos, some level of manual selection was necessary to verify peaks in the motion amplitude spectrum as modal motion. For instance, spurious camera motion would often appear as spikes in the spectrum. Verification was done by visually inspecting the magnified motion in the video at the frequency in question (following the method proposed in [2]). We believe that in the future this step could be automated. The number of extracted modes ranged from 12 to 31, depending on the video.

1.4.5 Inference Details

We modeled each drum as a triangular membrane mesh with 1530 linear elements and inferred material properties on a 20x20 pixel grid. In the presented results, the hyperparameter values are $\alpha_u = 10^{12}$, $\eta = 1$, $\alpha_w = 0.1$, $\alpha_v = 0.1$, and $\bar{w} = 10^6$. w and v are initialized to uniform values of 10^6 [Pa] and 10^3 [kg/m3], respectively, and reflect the estimated stiffness and density of latex.

1.5. Jello Cube Experiment

1.5.1 Inference Details

Our inference model was a 10x10x10 hexadral mesh with linear elements. The optimization hyperparameters were $\alpha_u = 0.1$, $\eta = 1$, $\alpha_w = 10^{-10}$, $\alpha_v = 10^{-8}$, and $\bar{w} = 10000$. w and v were initialized to 10000 [Pa] and 1500 [kg/m3], which are the estimated Young's modulus and measured density values of Jello.

References

[1] Hoss Belyadi, Ebrahim Fathi, and Fatemeh Belyadi. *Hydraulic Fracturing in Unconventional Reservoirs*. Elsevier, Amsterdam, The Netherlands, 2019. 2

[2] 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. 3