Acoustic Non-Line-of-Sight Imaging Supplementary Material

David B. Lindell
Stanford University

Gordon Wetzstein
Stanford University

Vladlen Koltun
Intel Labs

1. Reconstruction Algorithm

1.1. Discretized Image Formation

The acoustic array captures the response of a non-line-of-sight scene to a modulated acoustic signal emitted at location $(x_t, y_t, z = 0)$ and measured at location $(x_r, y_r, z = 0)$ over time t. After a pre-processing step which effectively collapses the frequency modulated transmit signal to an impulse (detailed in the main paper), the captured measurements are described as

$$\tilde{\tau}(x_t, y_t, x_r, y_r, t) = \iiint_{\Omega} \frac{1}{(r_t + r_r)^2} \rho(x, y, z)\, f(\boldsymbol{\omega_t}, \boldsymbol{\omega_r})\, \delta((r_t + r_r) - tc)\, dx\, dy\, dz \tag{1}$$

where ρ is the spatially varying albedo we hope to recover, f is the acoustic bidirectional reflectance distribution function, δ is a Dirac delta function, and $\Omega = \{(x, y, z) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R} \,|\, z > 0\}$.

In practice, we capture and process a discretized version of these measurements, $\tilde{\boldsymbol{\tau}} \in \mathbb{R}^{n_{x_t} \times n_{y_t} \times n_{x_r} \times n_{y_r} \times n_t}$ and recover a discretized volume $\boldsymbol{\rho} \in \mathbb{R}^{n_x \times n_y \times n_z}$. For a single transmit and receive position (*i.e.* $n_{x_t} = n_{y_t} = n_{x_r} = n_{y_r} = 1$), the discretized captured measurement $(\tilde{\boldsymbol{\tau}})_i$ is given by

$$(\tilde{\boldsymbol{\tau}})_i = \iint_{\Omega_{x_t, y_t}} \iint_{\Omega_{x_r, y_r}} \int_{t_{i-1}}^{t_i} \tilde{\tau}(x_t, y_t, x_r, y_r, t)\, dx_t\, dy_t\, dx_r\, dy_r\, dt, \tag{2}$$

where Ω_{x_t, y_t} and Ω_{x_r, y_r} describe the regions sampled by the transmitter and receiver, and the range of time values is uniformly divided between $[a, b] \subset (0, \infty)$ such that $a = t_0 < t_1 < \cdots < t_{n_t} = b$, with $1 \leq i \leq n_t$. The values n_{x_t}, n_{y_t}, n_{x_r}, and n_{y_r} similarly describe the number of uniformly discretized spatial samples along each spatial dimension.

We can also define an element in the discretized acoustic albedo volume, $(\boldsymbol{\rho})_{i,j,k}$.

$$(\boldsymbol{\rho})_{i,j,k} = \int_{z_{k-1}}^{z_k} \int_{y_{j-1}}^{y_j} \int_{x_{i-1}}^{x_i} \rho(x, y, z)\, dx\, dy\, dz, \tag{3}$$

where the number of uniformly discretized spatial samples is n_x, n_y, and n_z.

1.2. Confocalizing the Measurements

In order to reconstruct the 3D albedo volume $\boldsymbol{\rho}$, we process the captured non-confocal measurements $\tilde{\boldsymbol{\tau}}$ to emulate measurements captured on a grid of confocal samples, $\tilde{\boldsymbol{\tau}}_{\boldsymbol{c}}^*$. This reconstruction procedure consists of three substeps including a reparameterization procedure, normal moveout correction (NMO), and dip moveout correction (DMO), where the latter two steps are common to seismic image processing pipelines [8]. The confocalization procedure is outlined in Algorithm 1 for the discrete captured measurements.

Reparameterization The reparameterization procedure takes the measurements parameterized by transmit and receive coordinates and reparameterizes them by midpoint and offset coordinates. For all pairs of transmit and receive coordinates (x_t, y_t) and (x_r, y_r) we calculate the new midpoint and offset coordinates as $x_m = (x_t + x_r)/2$, $y_m = (y_t + y_r)/2$ and $h_x = |x_t - x_r|/2$, $h_y = |y_t - y_r|/2$. We call the reparameterized measurements $\tilde{\boldsymbol{\tau}}_{\boldsymbol{m}}$. See Algorithm 1 for a pseudocode implementation of this procedure.

NMO Correction NMO correction takes time resolved measurements at a given midpoint and offset and re-interpolates along the time dimension to remove any time delay relative to confocal measurements with a zero offset. For measurements at offset h_x, h_y, the adjusted time values $\boldsymbol{t_n}$ are given by

$$(\boldsymbol{t_n})_i = \sqrt{(\boldsymbol{t})_i^2 - \frac{4h_x^2}{c^2} - \frac{4h_y^2}{c^2}} \quad \forall i \in \{1, \ldots, n_t\}. \tag{4}$$

NMO correction, however, assumes that the surface normals of scatters are oriented with surface normal perpendicular to the acoustic array. For off-angle scatterers, an additional DMO correction adjusts the time dimension and spatial dimension to emulate the confocal measurements.

DMO Correction To perform the DMO correction, we use a Fourier domain approach called log-stretch DMO [9]. We extend the DMO correction to 3 dimensions as detailed in Algorithm 1 using the approach of Vlad and Biondi [7]. Additional motivation for the Fourier domain approach and derivations are also included below.

1.3. ADMM-based Reconstruction

Given the measurements $\tilde{\boldsymbol{\tau}}_{\boldsymbol{c}}^*$ which emulate confocally captured samples, we solve an optimization problem to mitigate spatial blur and impose priors on the reconstructed volume. We formulate the objective as

$$\arg\min_{\{\boldsymbol{\rho}\}} \frac{1}{2}\|A\boldsymbol{\rho} - \tilde{\boldsymbol{\tau}}_{\boldsymbol{c}}^*\| + \lambda_1\Gamma_1(\boldsymbol{\rho}) + \lambda_2\Gamma_2(\boldsymbol{\rho}) \tag{5}$$

where $\boldsymbol{\rho} \in \mathbb{R}^{n_x n_y n_z}$ and $\tilde{\boldsymbol{\tau}}_{\boldsymbol{c}}^* \in \mathbb{R}^{n_{x_m} n_{y_m} n_t}$ are vectorized, and A implements convolution with the spatial blur kernel and the light-cone transform kernel if reconstructing retroreflective or diffusely scattering objects [4]. The additional terms Γ_1 and Γ_2 are priors on the reconstructed volume with penalty terms λ_1 and λ_2. In our implementation, we choose 3D isotropic total variation (TV) and sparsity priors.

To solve this problem, we use the alternating direction method of multipliers [2] which allows us to split the problem into sub-problems and enforce consensus in the constraints. We reformulate the objective as

$$\begin{aligned} &\arg\min_{\{\boldsymbol{\rho}\}} \frac{1}{2}\|A\boldsymbol{\rho} - \tilde{\boldsymbol{\tau}}_{\boldsymbol{c}}^*\| + \lambda_1\|\boldsymbol{z_1}\|_1 + \lambda_2 \sum_i \left\| \begin{bmatrix} (\boldsymbol{z_2})_i \\ (\boldsymbol{z_3})_i \\ (\boldsymbol{z_4})_i \end{bmatrix} \right\|_2 \\ &\text{subject to} \\ &\underbrace{\begin{bmatrix} I \\ \boldsymbol{D_x} \\ \boldsymbol{D_y} \\ \boldsymbol{D_z} \end{bmatrix}}_{\boldsymbol{K}} \boldsymbol{\rho} = \underbrace{\begin{bmatrix} \boldsymbol{z_1} \\ \boldsymbol{z_2} \\ \boldsymbol{z_3} \\ \boldsymbol{z_4} \end{bmatrix}}_{\boldsymbol{z}}, \end{aligned} \tag{6}$$

where we have substituted the sparsity and 3D TV penalties for Γ_1 and Γ_2. The $\boldsymbol{D_x}, \boldsymbol{D_y}$, and $\boldsymbol{D_z}$ operators indicate first order difference operators along the x, y, and z dimensions. The augmented Lagrangian is given as

$$L_\rho(\boldsymbol{\rho}, \boldsymbol{z}, \boldsymbol{y}) = \frac{1}{2}\|A\boldsymbol{\rho} - \tilde{\boldsymbol{\tau}}_{\boldsymbol{c}}^*\| + \lambda_1\|\boldsymbol{z_1}\|_1 + \lambda_2 \sum_i \left\| \begin{bmatrix} (\boldsymbol{z_2})_i \\ (\boldsymbol{z_3})_i \\ (\boldsymbol{z_4})_i \end{bmatrix} \right\|_2 + \boldsymbol{y}^T(\boldsymbol{K\rho} - \boldsymbol{z}) + \frac{\rho}{2}\|\boldsymbol{K\rho} - \boldsymbol{z}\|_2^2 \tag{7}$$

with dual variable $\boldsymbol{y}$ and penalty parameter ρ [2]. For convenience, we combine the linear and quadratic terms and use the scaled form of ADMM, yielding a slightly rewritten form of the augmented Lagrangian:

$$L_\rho(\boldsymbol{\rho}, \boldsymbol{z}, \boldsymbol{u}) = \frac{1}{2}\|A\boldsymbol{\rho} - \tilde{\boldsymbol{\tau}}_{\boldsymbol{c}}^*\| + \lambda_1\|\boldsymbol{z_1}\|_1 + \lambda_2 \sum_i \left\| \begin{bmatrix} (\boldsymbol{z_2})_i \\ (\boldsymbol{z_3})_i \\ (\boldsymbol{z_4})_i \end{bmatrix} \right\|_2 + \frac{\rho}{2}\|\boldsymbol{K\rho} - \boldsymbol{z} + \boldsymbol{u}\|_2^2 - \frac{\rho}{2}\|\boldsymbol{u}\|_2^2. \tag{8}$$

Here $\boldsymbol{u} = (1/\rho)\boldsymbol{y}$ is the scaled dual variable [2]. This leads to the following iterative updates.

$\boldsymbol{\rho} \leftarrow \boldsymbol{\rho_0},\ \boldsymbol{u} \leftarrow \mathbf{0},\ \boldsymbol{z} \leftarrow \mathbf{0}$
for $k = 1$ **to** $maxiter$

$$\boldsymbol{z_1} \leftarrow \mathbf{prox}_{\|\cdot\|_1}(\boldsymbol{v}) = \underset{\{\boldsymbol{z_1}\}}{\arg\min}\, L_\rho(\boldsymbol{\rho}, \boldsymbol{z}, \boldsymbol{u}) = \underset{\{\boldsymbol{z_1}\}}{\arg\min}\, \lambda_1 \|\boldsymbol{z_1}\|_1 + \frac{\rho}{2}\|\boldsymbol{v} - \boldsymbol{z_1}\|_2^2, \quad \boldsymbol{v} = \boldsymbol{\rho} + \boldsymbol{u_1} \tag{9}$$

$$\begin{bmatrix}\boldsymbol{z_2}\\ \boldsymbol{z_3}\\ \boldsymbol{z_4}\end{bmatrix} \leftarrow \mathbf{prox}_{\|\cdot\|_{2,\mathrm{TV}}}(\boldsymbol{v}) = \underset{\{\boldsymbol{z_{2/3/4}}\}}{\arg\min}\, L_\rho(\boldsymbol{\rho}, \boldsymbol{z}, \boldsymbol{u}) = \underset{\{\boldsymbol{z_{2/3/4}}\}}{\arg\min}\, \lambda_2 \sum_i \left\| \begin{bmatrix}(\boldsymbol{z_2})_i\\ (\boldsymbol{z_3})_i\\ (\boldsymbol{z_4})_i\end{bmatrix} \right\|_2 + \frac{\rho}{2} \left\| \underbrace{\begin{bmatrix}\boldsymbol{v_1}\\ \boldsymbol{v_2}\\ \boldsymbol{v_3}\end{bmatrix}}_{\boldsymbol{v}} - \begin{bmatrix}\boldsymbol{z_2}\\ \boldsymbol{z_3}\\ \boldsymbol{z_4}\end{bmatrix} \right\|_2^2, \tag{10}$$

$$\boldsymbol{v} = \begin{bmatrix}\boldsymbol{D_x}\\ \boldsymbol{D_y}\\ \boldsymbol{D_z}\end{bmatrix} \boldsymbol{\rho} + \begin{bmatrix}\boldsymbol{u_2}\\ \boldsymbol{u_3}\\ \boldsymbol{u_4}\end{bmatrix}$$

$$\begin{bmatrix}\boldsymbol{u_1}\\ \boldsymbol{u_2}\\ \boldsymbol{u_3}\\ \boldsymbol{u_4}\end{bmatrix} \leftarrow \boldsymbol{u} + \boldsymbol{K}\boldsymbol{\rho} - \boldsymbol{z} \tag{11}$$

$$\boldsymbol{\rho} \leftarrow \mathbf{prox}_{\|\cdot\|_2}(\boldsymbol{v}) = \underset{\{\boldsymbol{\rho}\}}{\arg\min}\, L_\rho(\boldsymbol{\rho}, \boldsymbol{z}, \boldsymbol{u}) = \underset{\{\boldsymbol{\rho}\}}{\arg\min}\, \frac{1}{2}\|\boldsymbol{K}\boldsymbol{\rho} - \boldsymbol{v}\|_2^2, \quad \boldsymbol{v} = \boldsymbol{z} - \boldsymbol{u} \tag{12}$$

end for

Before the iterations, we initialize the unknown volume with the processed measurements $\boldsymbol{\rho_0} = \tilde{\boldsymbol{\tau}}_{\boldsymbol{c}}^*$. If the scene is retroreflective or diffusely scattering we apply the first stage of the LCT to yield the initial reconstruction, and then apply the final step of the LCT after the ADMM algorithm has converged. Concretely, the LCT image formation is given by

$$\boldsymbol{\rho} = \boldsymbol{R}_z^{-1}\mathcal{F}^{-1}\boldsymbol{H}\mathcal{F}\boldsymbol{R}_t\tilde{\boldsymbol{\tau}}_{\boldsymbol{c}}^* \tag{13}$$

with $\mathcal{F}$ being the Fourier transform, $\boldsymbol{R}_z^{-1}$ and $\boldsymbol{R}_t$ being quadratic reinterpolation operators, and $\boldsymbol{H}$ representing Wiener filter deconvolution with the LCT kernel [4]. We initialize the volume such that $\boldsymbol{\rho_0} = \mathcal{F}^{-1}\boldsymbol{H}\mathcal{F}\boldsymbol{R}_t\tilde{\boldsymbol{\tau}}_{\boldsymbol{c}}^*$ and solve the ADMM iterations within the quadratically resampled space so as to avoid the computational complexity of the resampling operator within every ADMM iteration. Once ADMM is converged, we apply the inverse quadratic resampling operator to yield the final output as $\boldsymbol{\rho}^* = \boldsymbol{R}_z^{-1}\boldsymbol{\rho}$.

We repeat the iterations until the primal and dual stopping criteria described by Boyd et al. [2] are met. Specifically, we iterate until $\|\boldsymbol{K}\boldsymbol{\rho}^k - \boldsymbol{z}^k\|_2 \le \epsilon_{\text{pri}}$ and $\|\rho\boldsymbol{K}^T(\boldsymbol{z}^{k+1} - \boldsymbol{z}^k)\|_2 \le \epsilon_{\text{dual}}$, with $\epsilon_{\text{pri}} = \epsilon_{\text{dual}} =$ 1e-3.

1.3.1 Proximal Operator for the Sparsity Prior (Eq. 9)

A closed form solution of the proximal operator for the sparsity term takes the form of the elementwise soft thresholding operator, $\mathcal{S}_\kappa$ [5].

$$\mathbf{prox}_{\|\cdot\|_1}(\boldsymbol{v}) = \mathcal{S}_{\lambda_1/\rho}(\boldsymbol{v}) \tag{14}$$

$$= (\boldsymbol{v} - \lambda_1/\rho)_+ - (-\boldsymbol{v} - \lambda_1/\rho)_+ \tag{15}$$

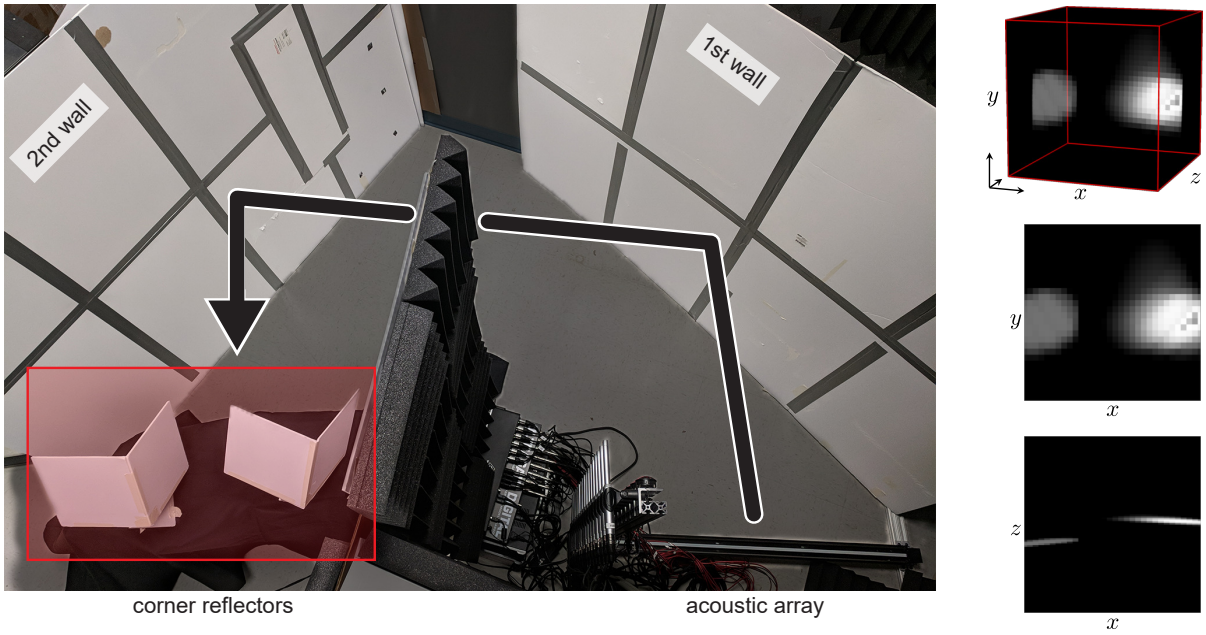

Figure 1. Double reflection experimental setup and results. Two corner reflectors are placed around two corners such that sound reflects off of two surfaces before being scattered back to the acoustic array (left). The reconstructed results show the relative positions of the corner reflectors (right).

1.3.2 Proximal Operator for the TV Prior (Eq. 10)

Again, the solution to the TV proximal operator is closed-form and is given by block soft thresholding [5].

$$\mathbf{prox}_{\|\cdot\|_{2,\mathrm{TV}}}(\boldsymbol{v}) = \left(1 - \frac{\lambda_2/\rho}{\left\| \begin{bmatrix} (\boldsymbol{v_2})_i \\ (\boldsymbol{v_3})_i \\ (\boldsymbol{v_4})_i \end{bmatrix} \right\|_2}\right)_+ \begin{bmatrix} (\boldsymbol{v_2})_i \\ (\boldsymbol{v_3})_i \\ (\boldsymbol{v_4})_i \end{bmatrix}, \quad \forall i \tag{16}$$

1.3.3 Proximal Operator for the Quadratic Term (Eq. 12)

Finally, the closed-form solution for the proximal operator for the quadratic term is given by the normal equations as

$$\mathbf{prox}_{\|\cdot\|_2}(\boldsymbol{v}) = \arg\min_{\{\boldsymbol{\rho}\}} \frac{1}{2}\|\boldsymbol{A\rho} - \tilde{\boldsymbol{\tau}}_{\boldsymbol{c}}^*\|_2^2 + \frac{\rho}{2}\|\boldsymbol{K\rho} - \boldsymbol{v}\|_2^2 \tag{17}$$

$$= (\boldsymbol{A}^T\boldsymbol{A} + \rho\boldsymbol{K}^T\boldsymbol{K})^{-1}(\boldsymbol{A}^T\tilde{\boldsymbol{\tau}}_{\boldsymbol{c}}^* + \rho\boldsymbol{K}^T\boldsymbol{v}) \tag{18}$$

$$= \mathcal{F}^{-1}\left\{\frac{\mathcal{F}^*\{a\}\cdot\mathcal{F}\{\tilde{\tau}_c^*\} + \rho\left(\mathcal{F}^*\{v_1\} + \mathcal{F}^*\{d_x\}\cdot\mathcal{F}\{v_2\} + \mathcal{F}^*\{d_y\}\cdot\mathcal{F}\{v_3\} + \mathcal{F}^*\{d_z\}\cdot\mathcal{F}\{v_4\}\right)}{\mathcal{F}^*\{a\}\cdot\mathcal{F}\{a\} + \rho\left(1 + \mathcal{F}^*\{d_x\}\cdot\mathcal{F}\{d_x\} + \mathcal{F}^*\{d_y\}\cdot\mathcal{F}\{d_y\} + \mathcal{F}^*\{d_z\}\cdot\mathcal{F}\{d_z\}\right)}\right\} \tag{19}$$

Here, the closed form solution consists of elementwise products $\cdot$ and divisions $\div$ computed on the Fourier transforms of the blur kernel a, first-order difference kernels d_x, d_y, and d_z, and volumes $\tilde{\tau}_c^*$, v_1, v_2, v_3, and v_4.

2. Supplementary Experiments

In this section, additional experimental results are described including imaging around multiple corners and evaluating the resolution of the acoustic array.

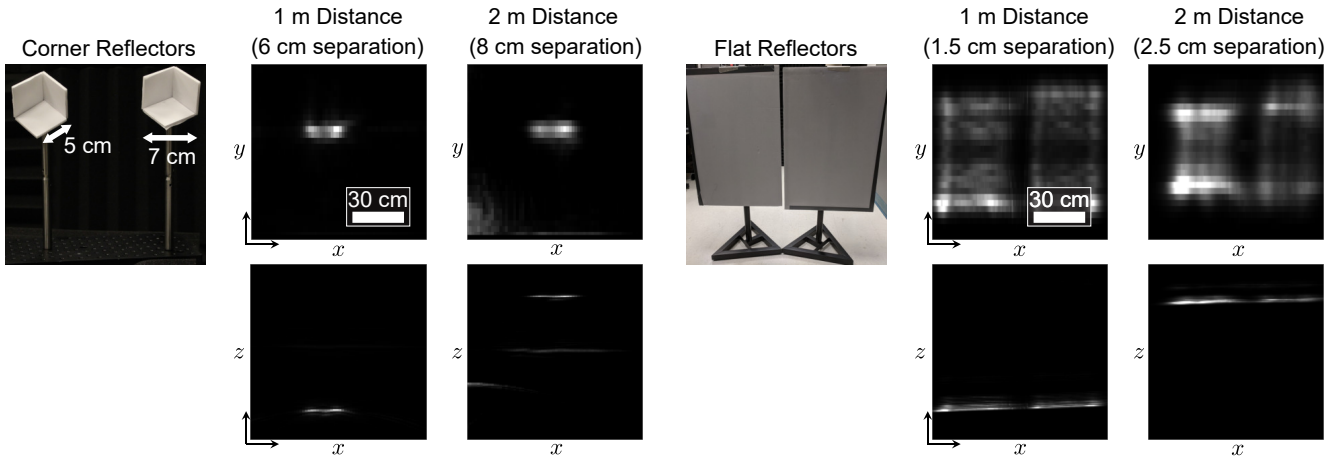

Figure 2. Resolution experimental results. The resolution limits of the acoustic array are tested by placing two corner reflectors (left) or two flat reflectors (right) close together until they are just separately resolveable. For the small corner reflectors, long wavelengths in the transmit signal tend to diffract around the reflectors, resulting in less backscattered signal and degraded resolution of approximately 13 cm and 15 cm at 1 m and 2 m distances. The flat reflectors reflect more of the signal and are just separable at spacings of 1.5 cm and 2.5 cm at distances of 1 m and 2 m.

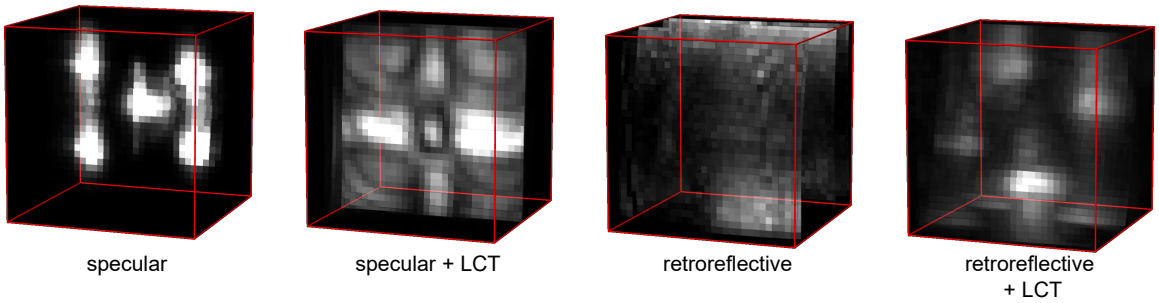

Figure 3. Effect of LCT on reconstruction of specularly and retrorefectively scattering objects. Applying the LCT to the specularly reflecting *Letter "H"* scene produces a blurry result because the LCT does not properly model specular image formation models. Applying the LCT to the retroreflectors in the *"Corner Reflectors"* scene focuses the scattered energy, recovering the locations of the scatterers.

Imaging Around Two Corners To capture an image around two corners, we place two corner reflectors with a side length of 25 cm in the scene, shown in Fig. 1. Here, the sound bounces of an initial wall, then another wall before scattering from the hidden objects and returning to the acoustic array. The total indirect path length is approximately 3.5 m, and we reconstruct the relative positions of the corner reflectors as shown in Fig. 1.

Resolution Experiments We test the maximum lateral resolution limits of the acoustic system by setting up two different line-of-sight scenes consisting of two small corner reflectors or two flat reflectors. We reconstruct the measurements on a grid of 64 horizontal $\times$ 30 vertical spatial samples for the corner reflectors and 128×30 samples for the flat reflectors. The resolution is determined at distances of 1 m and 2 m by decreasing the spacing between the scatterers until they are just separately resolveable in the reconstructed image (see Fig. 2).

For the small corner reflectors, wavelengths much longer than the side-length of 5 cm diffract around the object rather than being strongly backscattered. This reduces the effective bandwidth of the returning signal and results in worse resolution (roughly 6 cm at 1 m and 8 cm at 2 m) relative to the theoretical lower bounds of approximately 1 cm at 1 m and 2 cm at 2 m with a transmit signal bandwidth of 18 kHz. Theoretical resolution bounds are derived below.

The surface of the flat reflectors is larger than the longest wavelengths in the transmit signal (20 cm), and so a spacing of roughly 1.5 cm and 2.5 cm between the flat panels can be resolved at distances of 1 m and 2 m respectively.

Effect of LCT On Specular Measurements We examine the effect of the LCT on the reconstruction of retroreflective and specular objects in Fig. 3. Omitting the LCT for retroreflective objects results in a poorer result; however, we can still

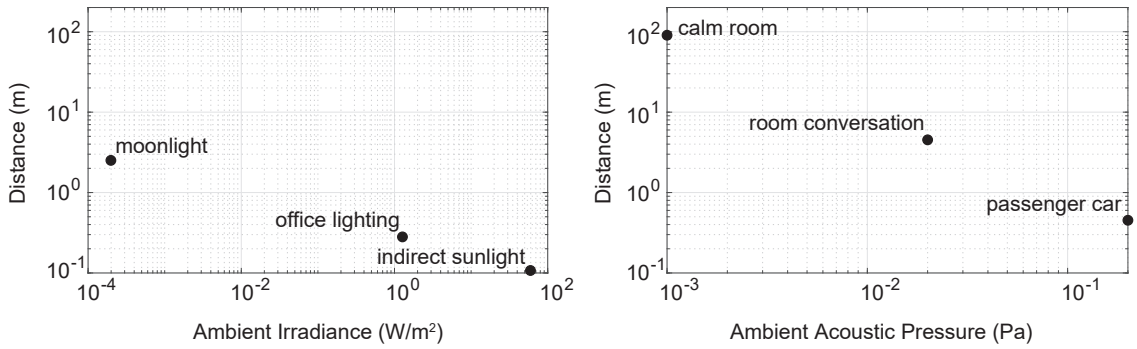

Figure 4. Simulated distance at which diffusely backscattered optical energy or specularly backscattered acoustic pressure equals the noise floor of incident ambient light or sound.

identify scatterers based on sound arrival times. Applying the LCT to specular objects blurs the reconstructed volume. A priori information about the scattering properties of the hidden objects should therefore be used to inform the choice of whether or not to use the LCT in the reconstruction.

Energy-Based Comparison to Optical Methods To provide additional comparison between acoustic and optical modalities, we analyze the acoustic and optical energies used in each hardware setup. The acoustic system emits approximately 1 mW of acoustic power for 1/16 second at each array position, and the optical setup uses a laser with 0.1 mW average power and 6 s exposure, resulting in 0.0625 mJ acoustic energy and 0.6 mJ optical energy per scan position.

A simulated comparison of the detection range of optical and acoustic techniques for a fixed transmit power and varying levels of ambient noise is shown in Fig. 4. In this comparison, we consider an optical source with 1 mW average power, 10 MHz repetition rate, and 80 ps pulse width. The detection distance is the distance from the wall to a diffusely scattering hidden object at which the backscattered energy from the laser pulse is equivalent to the energy of ambient light. For the acoustic case, we calculate the distance from the acoustic array to a specular hidden object such that the backscattered pressure wave from a 1 mW source is equivalent to the ambient pressure. We note the increased detection range of acoustic systems compared to optical systems as shown in Fig. 4.

3. Supplementary Derivations

Here we provide additional derivations which describe the resolution limits of the system and the geometry of the DMO correction as well as its Fourier domain relationships.

Derivation of Resolution Limits To derive an lower bound on the resolution of the optical system, we assume a perfectly specularly reflecting wall and couple the lateral resolution to the axial resolution of the system. Consider an acoustic signal emitted from a transmitter at location $(x_t, y_t, z = 0)$ which propagates to a reflector at a point (x, y, z) and back to a receiver at $(x_r, y_r, z = 0)$. The round trip propagation time, t_0 is calculated by adding the distance from the transmitter to the reflector r_t to the distance from the reflector to the receiver r_r and dividing by the speed of light c. This is expressed as

$$\begin{aligned} t_0 &= \frac{1}{c}(r_t + r_r) \\ &= \frac{1}{c}\left(\sqrt{(x_t - x)^2 + (y_t - y)^2 + z^2} + \sqrt{(x_r - x)^2 + (y_r - y)^2 + z^2}\right). \end{aligned} \tag{20}$$

We can determine how the time of flight changes with a small change in lateral position by taking the derivative with respect to the spatial position x:

$$\frac{dt_0}{dx} = \frac{1}{c}\left(\frac{x_t - x}{\sqrt{(x_t - x)^2 + (y_r - y)^2 + z^2}} + \frac{x_r - x}{\sqrt{(x_r - x)^2 + (y_r - y)^2 + z^2}}\right) \tag{21}$$

Substituting in the temporal resolution of the system γ for dt_0, yields an expression for the lateral resolution of the system. For our system, which uses a frequency modulated continuous wave (FMCW) signal, $\gamma = 1/(2B)$ [1], where B is the

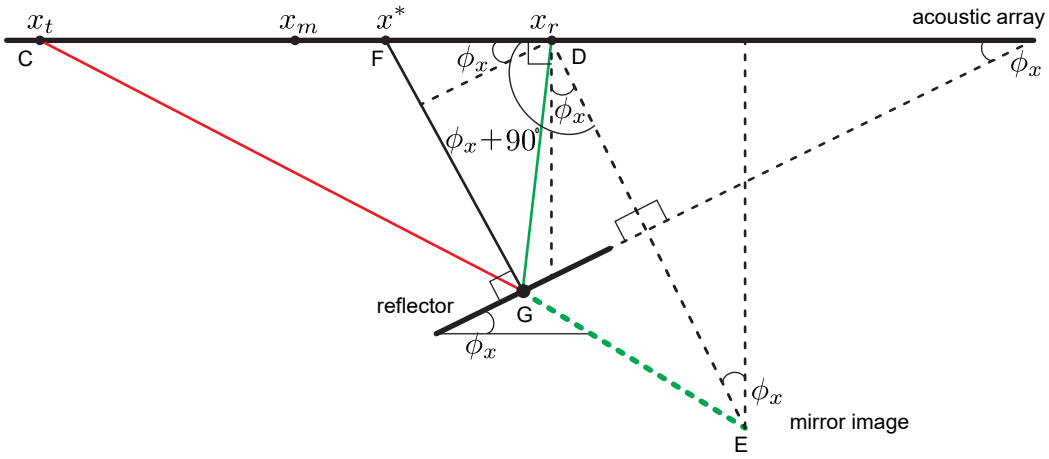

Figure 5. Geometry of acoustic array and angled reflector.

bandwidth of the FMCW signal.

$$dt_0 = \gamma = \frac{1}{2B} \tag{22}$$

$$\Rightarrow dx = \frac{\gamma c}{\frac{x_t - x}{r_t} + \frac{x_r - x}{r_r}}. \tag{23}$$

Note that this expression reduces to the form derived by O'Toole et al. [4] in the confocal case where $x_t = x_r$ and $y_t = y_r$. This expression also shows that confocal samples yield the best resolution, or the lowest value of dx.

Additional DMO Derivations DMO correction is applied after NMO correction to account for reflectors whose normal vector points away from the direction perpendicular to the acoustic array. We parameterize the orientation of the reflector by the angle ϕ_x which relates to the normal vector $\boldsymbol{n} \in \mathbb{R}^2$ by $n_x = \sin\phi_x$ and $n_z = -\cos\phi_x$. The DMO correction is given by

$$t_{\text{dmo}} = \sqrt{t_n^2 + \frac{4h_x^2 \sin^2\phi_x}{c^2}}. \tag{24}$$

where t_n is the NMO corrected time, h_x is the measurement offset, and t_{dmo} is the DMO corrected time which emulates measurements taken from a confocal sampling position. We derive this equation relying on the geometry of Fig. 5. Parts of this derivation are given by Slotnick [6] and Levin [3], but we include the entire derivation here for completeness.

Our intent is to find the time of propagation t^* along the confocal path from x^* to the scatterer and back. First, note from Fig. 5 the following relationships

$$\overline{CD} = 2h_x \tag{25}$$

$$\overline{CE} = r_r + r_t = t_0 c \tag{26}$$

$$\overline{DE} = t_1 c \tag{27}$$

where t_0 and t_1 indicate propagation time. By the law of cosines, we have that

$$\overline{CE}^2 = \overline{CD}^2 + \overline{DE}^2 - 2\,\overline{CD}\,\overline{DE}\cos\angle CDE. \tag{28}$$

Again, given the geometry of Fig. 5 we have that $\angle CDE = \phi_x + 90^0$. Then, substituting Eqs. 25, 26, and 27 into Eq. 28 gives

$$\begin{aligned} t_0^2 c^2 &= 4h_x^2 + t_1^2 c^2 - 4h_x t_1 c\,\cos(\phi_x + 90^0) \\ &= 4h_x^2 + t_1^2 c^2 + 4h_x t_1 c\,\sin\phi_x. \end{aligned} \tag{29}$$

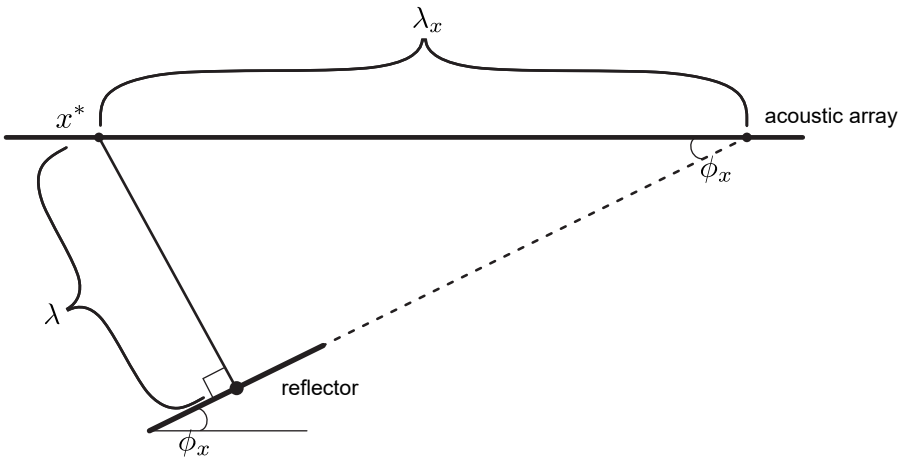

Figure 6. Geometry of angled reflector illustrating Fourier domain relationship. Assume the reflector is separated from confocal sample position x^* by a distance equivalent to a wavelength of the transmit signal. Then the observed spatial frequency at the acoustic array corresponds to a wavelength of λ_x dependent on the reflector angle ϕ_x.

Next, observe that

$$\begin{aligned} 2\overline{FG} = t^* c &= t_1 c + 2h_x \sin\phi_x \\ \Rightarrow t_1 c &= t^* c - 2h_x \sin\phi_x, \end{aligned} \tag{30}$$

and substitute Eq. 30 into Eq. 29.

$$t_0^2 c^2 = 4h_x^2 + c^2 (t^*)^2 + 4h_x^2 (1 - \sin^2\phi_x) \tag{31}$$

$$\Rightarrow t_0^2 = (t^*)^2 + \frac{4h_x^2 \cos^2\phi_x}{c^2}. \tag{32}$$

Recall the NMO equation and rearrange as

$$t_0^2 = t_n^2 + \frac{4h_x^2}{c^2}. \tag{33}$$

Then substitute into Eq. 32 to yield

$$\begin{aligned} t_n^2 + \frac{4h_x^2}{c^2} &= (t^*)^2 + \frac{4h_x^2 \cos^2\phi_x}{c^2} \\ \Rightarrow (t^*)^2 &= t_n^2 + \frac{4h_x^2}{c^2}(1 - \cos^2\phi_x) \\ \Rightarrow t^* &= \sqrt{t_n^2 + \frac{4h_x^2 \sin^2\phi_x}{c^2}} \end{aligned} \tag{34}$$

which is the DMO correction.

Relationship to Fourier Domain Generally we do not have knowledge of the angle of the reflector ϕ_x, and so we employ a Fourier domain approach which relates ϕ_x to the spatial and temporal frequencies k_x and ω. We briefly illustrate the relationship between ϕ_x and the Fourier parameters [8] and refer the interested reader to extended derivations of log-stretch DMO correction given by Yilmaz [8] or Zhou et al. [9].

In the geometry of Fig. 6 let the distance between the reflector and x^* be equal to one wavelength λ of the reflected signal at some transmit frequency. Then, the wavelength of the signal measured across the intersection with the acoustic array is given by λ_x and we have that

$$\sin\phi_x = \frac{\lambda}{\lambda_x} \tag{35}$$

The wavelengths can also be written using the following relationships

$$\lambda = \frac{2\pi c}{\omega} \tag{36}$$

$$\lambda_x = \frac{2\pi}{k_x}, \tag{37}$$

where ω is temporal frequency and k_x is spatial frequency. Substituting Eqs. 36 and 37 into Eq. 35 gives

$$\sin \phi_x = \frac{ck_x}{\omega} \tag{38}$$

which allows us to rewrite Eq. 34 without explicit dependence on ϕ_x via a ratio of the measured spatial and temporal frequencies.

References

[1] C. Baker and S. Piper. Continuous wave radar. In *Principles of Modern Radar: Volume 3: Radar Applications*, pages 17–85. Institution of Engineering and Technology, 2013. 6

[2] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 3(1):1–122, 2011. 2, 3

[3] F. K. Levin. Apparent velocity from dipping interface reflections. *Geophysics*, 36(3):510–516, 1971. 7

[4] M. O'Toole, D. Lindell, and G. Wetzstein. Confocal non-line-of-sight imaging based on the light cone transform. *Nature*, 555(7696):338–341, 2018. 2, 3, 7

[5] N. Parikh and S. Boyd. Proximal algorithms. *Foundations and Trends in Optimization*, 1(3):123–231, 2013. 3, 4

[6] M. Slotnick. Lessons in seismic computing. *Soc. Expl. Geophys*, 268, 1959. 7

[7] I. Vlad and B. Biondi. Effective AMO implementation in the logstretch, frequency-wavenumber domain. *Stanford Exploration Project*, SEP-110:1–8, 2001. 2

[8] Ö. Yilmaz. *Seismic Data Analysis: Processing, Inversion, and Interpretation of Seismic Data*. Society of Exploration Geophysicists, 2001. 1, 8

[9] B. Zhou, I. M. Mason, and S. A. Greenhalgh. An accurate formulation of log-stretch dip moveout in the frequency-wavenumber domain. *Geophysics*, 61(3):815–820, 1996. 2, 8

Algorithm 1 Measurement Confocalization Procedure

Require:

$\tilde{\boldsymbol{\tau}} \in \mathbb{R}^{n_{x_t} \times n_{y_t} \times n_{x_r} \times n_{y_r} \times n_t}$ ▷ Measurements

$\boldsymbol{x_t} \in \mathbb{R}^{n_{x_t}}, \boldsymbol{y_t} \in \mathbb{R}^{n_{y_t}}, \boldsymbol{x_r} \in \mathbb{R}^{n_{x_r}}, \boldsymbol{y_r} \in \mathbb{R}^{n_{y_r}}, \boldsymbol{t} \in \mathbb{R}^{n_t}$ ▷ List of transmit, receive, and time coordinates

$\boldsymbol{x_m} \in \mathbb{R}^{n_{x_m}}, \boldsymbol{y_m} \in \mathbb{R}^{n_{y_m}}, \boldsymbol{h_x} \in \mathbb{R}^{n_{h_x}}, \boldsymbol{h_y} \in \mathbb{R}^{n_{h_y}}$ ▷ List of midpoint and offset coordinates for measurement locations $\boldsymbol{x_t}, \boldsymbol{y_t}, \boldsymbol{x_r}, \boldsymbol{y_r}$

procedure COORD2IDX(x_t, y_t, x_r, y_r) ▷ Find midpoint, offset coordinate indices for given transmit, receive coordinates

  Find i s.t. $(x_t + x_r)/2 = (\boldsymbol{x_m})_i$

  Find j s.t. $(y_t + y_r)/2 = (\boldsymbol{y_m})_j$

  Find k s.t. $|x_t - x_r|/2 = (\boldsymbol{h_x})_k$

  Find ℓ s.t. $|y_t - y_r|/2 = (\boldsymbol{h_y})_\ell$

  Return i, j, k, ℓ

end procedure

$\tilde{\boldsymbol{\tau}}_{\boldsymbol{m}} \leftarrow \mathbf{0} \in \mathbb{R}^{n_{x_m} \times n_{y_m} \times n_{h_x} \times n_{h_y} \times n_t}$ ▷ **Reparameterize measurements by midpoint and offset**

for $p = (1 : n_{x_t}), q = (1 : n_{x_r}), r = (1 : n_{y_t}), s = (1 : n_{y_r})$ **do**

  $(i, j, k, \ell) \leftarrow$ COORD2IDX$((\boldsymbol{x_t})_p, (\boldsymbol{x_r})_q, (\boldsymbol{y_t})_r, (\boldsymbol{y_r})_s)$

  $(\tilde{\boldsymbol{\tau}}_{\boldsymbol{m}})_{i,j,k,\ell} \in \mathbb{R}^{n_t} \leftarrow (\tilde{\boldsymbol{\tau}})_{p,q,r,s} \in \mathbb{R}^{n_t}$

end for

for $i = (1 : n_{x_m}), j = (1 : n_{y_m}), k = (1 : n_{h_x}), \ell = (1 : n_{h_y})$ **do** ▷ **Apply NMO correction**

  $\boldsymbol{t_n} = \sqrt{\boldsymbol{t}^2 - \frac{4(\boldsymbol{h_x})_k^2}{c^2} - \frac{4(\boldsymbol{h_y})_\ell^2}{c^2}}$

  $(\tilde{\boldsymbol{\tau}}_{\boldsymbol{m}})_{i,j,k,\ell} \leftarrow$ INTERPOLATE$((\tilde{\boldsymbol{\tau}}_{\boldsymbol{m}})_{i,j,k,\ell}, \boldsymbol{t}, \boldsymbol{t_n})$ ▷ Reinterpolate along time dimension to NMO-corrected time coordinates

end for

$\tilde{\boldsymbol{\tau}}^* \leftarrow \mathbf{0} \in \mathbb{R}^{n_{x_m} \times n_{y_m} \times n_{h_x} \times n_{h_y} \times n_t}$ ▷ **Apply DMO correction**

$\boldsymbol{t'_n} = \ln \boldsymbol{t_n}$

for $i = (1 : n_{x_m}), j = (1 : n_{y_m}), k = (1 : n_{h_x}), \ell = (1 : n_{h_y})$ **do**

  $(\tilde{\boldsymbol{\tau}}^*)_{i,j,k,\ell} \leftarrow$ INTERPOLATE$((\tilde{\boldsymbol{\tau}}_{\boldsymbol{m}})_{i,j,k,\ell}, \boldsymbol{t_n}, \boldsymbol{t'_n})$ ▷ Reinterpolate along time dimension to log-stretch time coordinates

end for

$\boldsymbol{T}_{\text{dmo}} \leftarrow \mathcal{F}_{\boldsymbol{x_m}, \boldsymbol{y_m}, \boldsymbol{t}}\{\tilde{\boldsymbol{\tau}}^*\}(\boldsymbol{k_x}, \boldsymbol{k_y}, \boldsymbol{W}) \in \mathbb{R}^{n_{k_x} \times n_{k_y} \times n_{h_x} \times n_{h_y} \times n_W}$ ▷ Fourier transform along the midpoint and time coordinates $\boldsymbol{x_m}, \boldsymbol{y_m}, \boldsymbol{t}$, and calculate the respective dual frequency coordinates $\boldsymbol{k_x}, \boldsymbol{k_y}, \boldsymbol{W}$

for $i = (1 : n_{k_x}), j = (1 : n_{k_y}), k = (1 : n_{h_x}), \ell = (1 : n_{h_y}), p = (1 : n_W)$ **do**

  Let $\boldsymbol{k} \cdot \boldsymbol{h} = (\boldsymbol{k_x})_i (\boldsymbol{h_x})_k + (\boldsymbol{k_y})_j (\boldsymbol{h_y})_\ell$

  $$\Phi \leftarrow \begin{cases} 0, & \boldsymbol{k} \cdot \boldsymbol{h} = 0 \\ \boldsymbol{k} \cdot \boldsymbol{h} & (\boldsymbol{W})_p = 0 \\ \dfrac{(\boldsymbol{W})_p}{2} \left\{ \sqrt{1 + \left(\dfrac{2\boldsymbol{k} \cdot \boldsymbol{h}}{(\boldsymbol{W})_p} \right)^2} - 1 - \ln \left[\dfrac{\sqrt{1 + \left(\frac{2\boldsymbol{k} \cdot \boldsymbol{h}}{(\boldsymbol{W})_p} \right)} + 1}{2} \right] \right\}, & (\boldsymbol{W})_p \neq 0 \end{cases}$$

  $(\boldsymbol{T}_{\text{dmo}})_{i,j,k,\ell,p} \leftarrow e^{j\Phi} (\boldsymbol{T}_{\text{dmo}})_{i,j,k,\ell,p}$ ▷ Apply phase shift in Fourier domain

end for

$\tilde{\boldsymbol{\tau}}^* \leftarrow \mathcal{F}^{-1}_{\boldsymbol{k_x}, \boldsymbol{k_y}, \boldsymbol{W}}\{\boldsymbol{T}_{\text{dmo}}\}(\boldsymbol{x_m}, \boldsymbol{y_m}, \boldsymbol{t'_n})$ ▷ Inverse Fourier transform

for $i = (1 : n_{x_m}), j = (1 : n_{y_m}), k = (1 : n_{h_x}), \ell = (1 : n_{h_y})$ **do**

  $(\tilde{\boldsymbol{\tau}}^*)_{i,j,k,\ell} \leftarrow$ INTERPOLATE$((\tilde{\boldsymbol{\tau}}^*)_{i,j,k,\ell}, \boldsymbol{t'_n}, \boldsymbol{t_n})$ ▷ Reinterpolate along time dimension to undo log-stretch time coordinates to produce output, $\boldsymbol{\tau}^*$

end for

$\tilde{\boldsymbol{\tau}}^*_{\boldsymbol{c}} = \sum_{h_x, h_y} \tilde{\boldsymbol{\tau}}^*$ ▷ Sum over the offset dimensions to produce the final output