DOI 10.5162/iCCC2026/P1

Analysis of Ultrasound Echoes for Object Properties Characterisation Using Interpretable Features

Priscile Suawa, BTU Cottbus-Senftenberg, Cottbus, Germany, suawapri@b-tu.de
Marcel Jongmanns, Fraunhofer IPMS, Dresden, Germany, marcel.jongmanns@ipms.fraunhofer.de
Christian Herglotz, BTU Cottbus-Senftenberg, Cottbus, Germany, herglchr@b-tu.de

Abstract

This work analyzes ultrasound echoes across multiple objects using a compact set of physically interpretable time- and frequency-domain features, in addition to the raw waveform structure. Unlike black-box models such as deep neural networks or ensembles, which prioritize accuracy over transparency, our approach emphasizes descriptors with direct physical meaning. We extract features that capture geometry, reflectivity, spectral richness, and structural complexity to provide classifiers with interpretable inputs that explain the physical basis for object recognition.

Echoes from five objects were measured across varying orientations and placements. The Time-of-Flight (ToF) decreased near normal incidence and increased at oblique angles, reflecting the propagation geometry. Energy differed by a factor of five between dense-filled and hollow objects, while bandwidth broadened by several hundred hertz under specific orientations. Complementary visualizations using Short-Time Fourier Transform (STFT), Continuous Wavelet Transform (CWT), and Principal Component Analysis (PCA) revealed object-dependent patterns consistent with the scalar features. These results demonstrate that ultrasound echoes encode rich, interpretable information, supporting lightweight feature-based analysis for transparent and low-power embedded sensing.

1 Introduction

Ultrasound sensing is widely used in medical imaging, robotics, and industrial inspection due to its non-invasive nature and sensitivity to both surface and internal structures [1]. While recent machine learning approaches enable object classification from echoes, they often rely on opaque, compute-heavy models with limited physical interpretability, such as deep neural networks [2], [3], which are problematic in safety-critical domains. While these methods can be accurate, they often suffer from high computational complexity and limited interpretability. Attempts to mitigate this, such as GradCAM visualizations [4], provide image-level explanations in diagnostic contexts but do not establish direct links to the physical echo descriptors. Our case study, on the other hand, focuses on directly connecting interpretable features to geometry and material properties.

This work investigates the extent to which information can be directly extracted from echo signals through physics-driven, explainable signal processing. Rather than training deep networks, we extract descriptors such as time-of-flight (ToF), bandwidth, energy, and spectral centroid from ultrasound echoes to characterize material and geometry. These are supplemented with time-frequency visualizations and PCA to reveal structural patterns in the raw signals. Our study focuses on: (1) how angular orientation affects key acoustic features, and (2) how material and structure influence echo signatures across object types. Using five representative objects with varying rigidity, shape, and internal composition at multiple positions, we find that echoes encode rich, interpretable patterns: dense objects yield strong, narrowband reflections, while hollow and porous ones generate resonant responses. These findings highlight that ultrasound echoes contain discriminative and physically meaningful information, eliminating the need for complex computations and providing a foundation for interpretable, low-power embedded sensing.

2 Experimental Setup

Ultrasound echoes were recorded using a custom air-coupled transducer pair centred at 42 kHz, sampled at 2 MHz to preserve acceptable time resolution. The speed of sound in air was 343 m/s (room temperature). Each pulse-echo acquisition produced a 21×6000 matrix, where each row corresponds to an A-scan (Amplitude scan), i.e., a time-domain waveform captured at a specific incidence angle.

Five objects were selected to vary in material (glass, plastic, cardboard), geometry, and fill state (filled vs. hollow). They include a dense glass jar, a hollow plastic flask, a hollow glass bottle, and cardboard boxes of different dimensions. Echoes were collected at three lateral offsets (Left, Middle, Right) and two axial distances (20 cm, 40 cm), with angular variation from −30° to +30°. This resulted in 30 distinct measurement conditions. The test setup ensured consistent placement using a fixed ruler and protractor (see Figure 1). Throughout this paper, we define the axial distance d (object-to-transducer), lateral offset (Left, Middle, Right), and angle α (object rotation) as key geometric parameters.

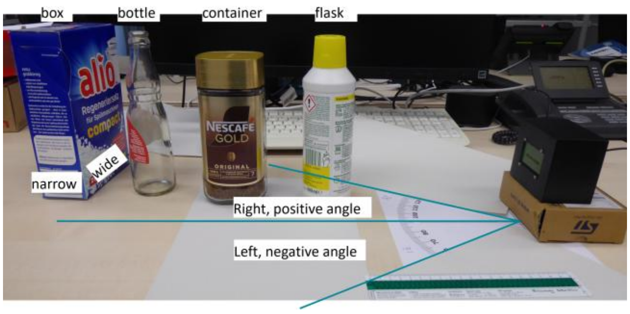

Figure 1. Experimental setup with objects (left) and transducers (right).

3 Methodology

First, the recorded echo measurements undergo pre-processing to remove artifacts and standardize the signal format, before feature extraction. Each A-scan comprises 6000 samples; however, the first 1500 samples are typically dominated by direct coupling and transducer cross-talk. These initial samples are discarded, leaving the remaining 4500 samples for analysis. A zero-phase bandpass filter [5], operating in the range of 35 - 60 kHz, is applied to each signal to isolate the dominant echo frequencies while attenuating low-frequency noise and high-frequency artifacts. This frequency window encompasses the transducer's centre frequency (42 kHz) and its effective bandwidth. Then, to characterize the acoustic responses of different objects across spatial positions and angular orientations, we extract a compact set of physically interpretable features from each A-scan. These include:

- ToF is estimated via peak detection, identifying the first significant reflection as the maximum absolute amplitude in the A-scan [6]. Formally,

$$\mathrm{ToF} = \frac{1}{f_s} \underset{k}{\operatorname{argmax}} |x[k]|$$

where $x[k]$ is the discrete-time echo signal, f_s is the sampling frequency (Hz), and k is the sample index of the first significant reflection.

- Echo energy quantifies the total reflected power as the sum of squared amplitudes [7]:

$$E = \sum_{k=0}^{N-1} x\,[k]^2$$

where $x[k]$ is the discrete-time echo signal after preprocessing and N is the number of samples ($N = 4500$ in this work). This corresponds to the time-integrated reflected power, i.e., the signal energy.

- Spectral centroid represents the frequency centre of mass of the echo signal. It is computed from the magnitude spectrum obtained by applying the Fast Fourier Transform (FFT) to the discrete-time signal [8], [9]:

$$C = \frac{\sum_{m=0}^{M-1} f\,[m]\,|X[m]|}{\sum_{m=0}^{M-1} |X[m]|}$$

where $X[m] = \mathrm{FFT}_M\{x[k]\}$ denotes the M-point Fast Fourier Transform of $x[k]$, $|X[m]|$ is its magnitude spectrum, $f[m] = m f_s/M$ is the frequency of bin m, and M is the FFT length (here equal to the number of samples, $N = 4500$).

- Bandwidth quantifies the spectral spread of the signal around its centroid, measuring how concentrated or dispersed the frequency content is:

$$B = \sqrt{\frac{\sum_{m=0}^{M-1} (f[m] - C)^2\ |X[m]|}{\sum_{m=0}^{M-1} |X[m]|}}$$

where $|X[m]|$ is the magnitude spectrum of the signal, $f[m]$ the frequency of bin m, and C the spectral centroid.

Finally, to investigate echo structure beyond scalar features, we apply both the Short-Time Fourier Transform (STFT) and Continuous Wavelet Transform (CWT) [10], [11] to A-scans. The STFT reveals the evolution of spectral energy over time with a fixed spectral resolution, while the CWT provides a multi-resolution analysis, offering finer temporal resolution at high frequencies and higher frequency resolution at low frequencies.

To further explore latent signal structure, PCA [12] is applied directly to the raw A-scan waveforms. PCA projects high-dimensional signals into an orthogonal basis of uncorrelated components, highlighting dominant modes of variation. The leading principal components are retained to reveal object-dependent clustering or geometric separability in feature space.

4 Echo Signature Analysis

In this section, we analysed how ultrasound echoes vary with object geometry, material, and orientation, using scalar features and visualization tools to uncover discriminative patterns.

To evaluate how object orientation influences acoustic signatures, we examine Obj3 (a hollow glass bottle) at a fixed axial distance of 40 cm, with angular sweeps from −30° to +30° in 3° steps across three lateral positions (Left, Middle, Right). We analyse two features: Time-of-Flight (ToF) and Spectral Bandwidth, both sensitive to object geometry and alignment.

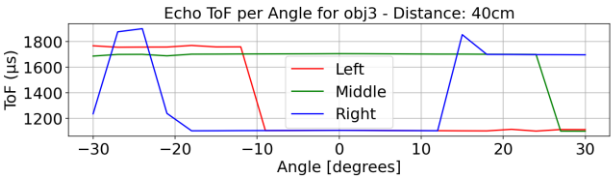

Figure 2. Time-of-Flight (ToF) vs. angle for Obj3 at 40 cm, across three lateral positions.

Time-of-Flight (ToF) measures the delay of the first significant echo. As shown in Figure 2, ToF is shortest near normal incidence (α ≈ 0°) and increases toward oblique angles due to longer propagation paths. Jagged, step-like variations, especially at off-centre positions, reflect the influence of internal reflections within the hollow structure. Values range from 1100 to 1900 µs, consistent with expectations for a 40 cm round trip in air (c = 343 m/s), after accounting for the discarded 1500-sample pre-echo window. ToF provides a reliable indicator of alignment and echo path length. Variability at oblique angles reveals the role of internal structure, especially in hollow objects like Obj3.

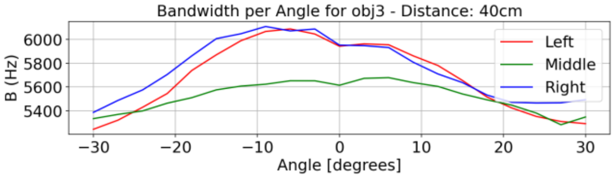

Figure 3. Spectral bandwidth vs. angle for Obj3 at 40 cm.

Spectral Bandwidth, shown in Figure 3, varies more subtly with angle. The Left and Right positions exhibit clear peaks near −10°, reaching up to 5600 Hz, while the Middle position remains relatively flat around 5200 Hz. This suggests that specific orientations enhance high-frequency scattering, possibly due to curvature-induced resonances. Bandwidth, therefore, complements ToF by capturing orientation-dependent spectral richness. Together, these two features reflect complementary aspects of acoustic interaction: ToF traces geometric alignment, while bandwidth encodes frequency-domain effects linked to surface and internal structure. Next, we compared all five objects at a fixed geometry (Middle lateral position, 40 cm axial distance) to assess how material and structure influence echo features. Two key descriptors were selected: energy and spectral centroid.

Echo Energy, as shown in Figure 4, indicates that Obj1 (a filled glass jar) produces the strongest echoes, peaking near 9 × 10^7 a.u. at central angles. This level is nearly four times higher than the hollow glass bottle (Obj3) and plastic flask (Obj2), which reach only 2-2.5×10^7. The cardboard containers (Obj4 narrow box, Obj5 wide box) fall in between, around 4-5×10^7. These contrasts confirm that echo energy is strongly tied to material density and structural fill: dense solids dominate, while hollow or lightweight objects return weaker reflections.

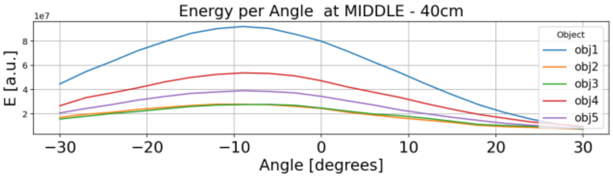

Figure 4. Echo energy (E) in arbitrary units (a.u.) across angle for all objects at Middle position, 40 cm.

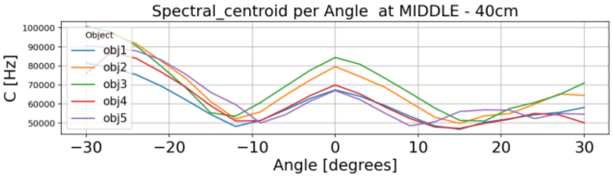

Figure 5. Spectral centroid (C) as a function of angle at Middle position.

Spectral centroid, shown in Figure 5, separates objects by geometry and rigidity. Obj3 reaches up to 100 kHz at oblique angles, approximately 20-30% higher than the cardboard boxes, which remain at around 50-60 kHz. Obj2 also exhibits elevated centroids (up to 80 kHz), reflecting stiff plastic surfaces. These trends suggest that narrow, hollow, or rigid structures concentrate energy in higher frequency bands, while broad, softer materials shift reflections downward.

These two features capture complementary aspects of acoustic behaviour: energy reflects reflectivity and mass, and centroid captures resonance and stiffness. Together, they allow physically interpretable separation of object types.

Table 3 summarizes the physical relevance and observed trends of the retained features, consolidating insights from the prior analyses.

Feature	**Physical Link**	**Observed Trend**
Time-of-Flight (ToF)	Propagation geometry	Shortest at normal incidence; increases at oblique angles
Energy	Material density, reflectivity	Highest for dense, filled objects
Spectral Centroid	Resonance, stiffness	Higher for hollow, narrow, or rigid objects
Bandwidth	Spectral richness (angular only)	Peaks at off-normal angles; flatter at centre

Table 3. Summary of features with their physical interpretation and observed trends.

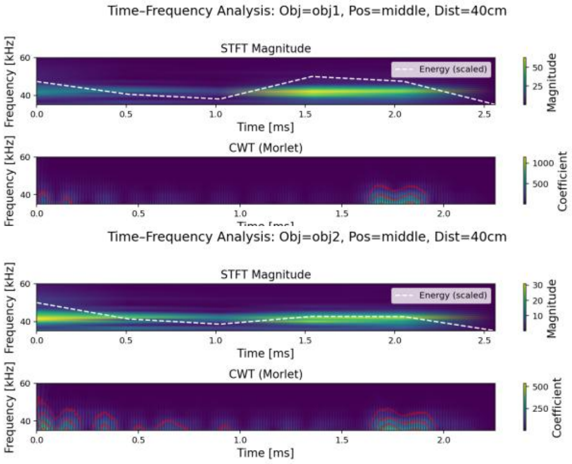

Figure 6. Time-frequency analysis of echoes at d = 40 cm, middle position. STFT magnitude spectrogram with overlaid normalized energy trace (white dashed line). CWT, with red contours indicating regions of high spectral energy.

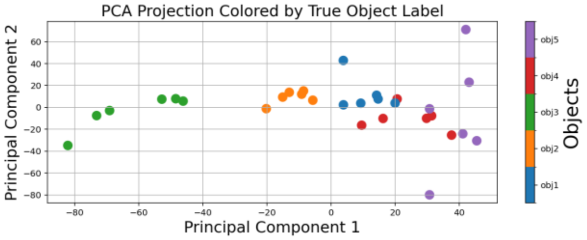

Figure 7. PCA projection of raw waveforms onto the first two principal components.

To complement scalar features, we visualized representative A-scans using time-frequency analysis and dimensionality reduction. STFT and CWT plots (Figure 6) compare Obj1 (filled glass) and Obj2 (hollow plastic). In the STFT panels, Obj1 exhibits a single sharp, short-lived burst of energy visible around 1.5-1.7 *ms*, concentrated in the 38-45 kHz band, consistent with a dense, filled structure producing strong specular reflections. In contrast, Obj2 shows a more prolonged response: its STFT reveals energy distributed over a broader window from 0.5-2.0 *ms*, with frequency content spanning roughly 35-45 kHz, suggesting reverberation within the hollow cavity. The CWT panels reinforce these differences. For Obj1, the red contours highlight a localized ridge near 40-45 kHz, indicating a compact narrowband reflection. For Obj2, multiple oscillatory ridges persist over time, characteristic of resonant ringing and sustained energy trapping. These complementary views illustrate how hollow and filled objects differ not only in scalar descriptors but also in their temporal and spectral evolution.

Additionally, a PCA projection of raw waveforms (Figure 7) highlights object separability in low-dimensional space. Obj2 and Obj3 form tight, distinct clusters, while Obj1 and Obj4 show partial overlap, indicating that hollow objects produce more consistent and distinctive waveform patterns. Even without explicit feature extraction, raw ultrasound signals exhibit structure that reflects object identity. Time-frequency analysis reveals where energy is concentrated, while PCA confirms waveform-level distinctions.

5 Conclusion and Future Work

This work presents a signal-processing-based framework for ultrasound echo analysis, emphasizing interpretable features directly linked to material, geometry, and structure. Using descriptors such as ToF, energy, spectral centroid, and bandwidth, we demonstrated that even lightweight metrics reveal systematic differences: dense filled objects produce strong, narrowband echoes. At the same time, hollow or porous structures yield weaker and resonant responses.Time-frequency visualizations (STFT, CWT) clarified the origins of scalar features, and PCA revealed separability of object types. Together, these complementary analyses highlight that ultrasound echoes encode rich, object-specific signatures that are both discriminative and physically interpretable.

Future work will first increase the feature set and extend this foundation toward classification by benchmarking raw and engineered features on larger datasets. In parallel, lightweight feature-based classifiers will be compared against opaque black-box models such as CNNs and ensembles, clarifying the trade-off between accuracy and interpretability for low-power embedded sensing.

Such lightweight, interpretable ultrasound characterization could support applications in embedded robotics, packaging inspection, and smart environments, where distinguishing between object types (e.g., filled vs. hollow containers, rigid vs. porous surfaces) must be achieved efficiently under power and computational constraints.

6 Literature

[1] Alexander L. Bowler, Michael P. Pound, Nicholas J. Watson, A review of ultrasonic sensing and machine learning methods to monitor industrial processes, Ultrasonics, Volume 124, 2022, 106776, ISSN 0041-624X, https://doi.org/10.1016/j.ultras.2022.106776.

[2] Shafiei Alavijeh M, Scott R, Seviaryn F, Maev RG. Using machine learning to automate ultrasound-based classification of butt-fused joints in medium-density polyethylene gas pipes. J Acoust Soc Am. 2021 Jul;150(1):561. doi: 10.1121/10.0005656. PMID: 34340509.

[3] Ning Mao, Amirah Nabilah Azman, Guangxin Ding, Yubo Jin, Can Kang, Hyoung-Bum Kim, Black-box real-time identification of sub-regime of gas-liquid flow using Ultrasound Doppler Velocimetry with deep learning, Energy, Vol 239, Part D, 2022, 122319, https://doi.org/10.1016/j.energy.2021.122319.

[4] R. Singh, A. S. Bansal, and M. R. Bansal, "Plaque classification using ultrasound images with machine learning and explainable AI," *Biomedical Signal Processing and Control*, vol. 93, p. 106063, 2024. https://doi.org/10.1016/j.bspc.2023.106063.

[5] Steven W. Smith, The Scientist and Engineer's Guide to Digital Signal Processing, 1st ed. California Technical Publishing, 1997. [Online]. Available: https://www.dspguide.com/ch19/1.htm.

[6] Cherif Othmani, Najmeh Sadat Dokhanchi, Sebastian Merchel, Albert Vogel, M. Ercan Altinsoy, Conrad Voelker, A review of the state-of-the-art approaches in detecting time-of-flight in room impulse responses, Sensors and Actuators A: Physical, Volume 374, 2024, 115467, ISSN 0924-4247, https://doi.org/10.1016/j.sna.2024.115467.

[7] Bowler, A.; Escrig, J.; Pound, M.; Watson, N. Predicting Alcohol Concentration during Beer Fermentation Using Ultrasonic Measurements and Machine Learning. Fermentation 2021, 7, 34. https://doi.org/10.3390/fermentation7010034.

[8] Nilesh Kulkarni, Vinayak Bairagi, Chapter Four - Use of Complexity Features for Diagnosis of Alzheimer Disease, Editor(s): Nilesh Kulkarni, Vinayak Bairagi, EEG-Based Diagnosis of Alzheimer Disease, Academic Press, 2018, Pages 47-59, ISBN 9780128153925, https://doi.org/10.1016/B978-0-12-815392-5.00004-6.

[9] Lv, D.; Zi, J.; Huang, X.; Gao, M.; Xi, R.; Li, W.; Wang, Z. Feature Extraction on the Difference of Plant Stem Structure Based on Ultrasound Energy. Agriculture 2023, 13, 52. https://doi.org/10.3390/agriculture13010052.

[10] Yang Yang, Zhike Peng, Wenming Zhang, Guang Meng, Parameterised time-frequency analysis methods and their engineering applications: A review of recent advances, Mechanical Systems and Signal Processing, Volume 119, 2019, Pages 182-221, ISSN 0888-3270, https://doi.org/10.1016/j.ymssp.2018.07.039.

[11] Lupea, I.; Lupea, M. Continuous Wavelet Transform and CNN for Fault Detection in a Helical Gearbox. Appl. Sci. 2025, 15, 950. https://doi.org/10.3390/app15020950.

[12] Azim, M. R., & Gül, M. (2020). Data-driven damage identification technique for steel truss railroad bridges utilizing principal component analysis of strain response. Structure and Infrastructure Engineering, 17(8),1019–1035. https://doi.org/10.1080/15732479.2020.1785512.