# Acoustic Impulse Response Anomaly Detection

UDK 621.395.7 (043.2) Oleksandr Terletskyi[1], Valerii Trushevskyi[2]

*Ivan Franko National University of Lviv,*
*[1]oleksandr.terletskyi@lnu.edu.ua, [2]valeriy.trushevsky@lnu.edu.ua*

Indoor intrusion detection traditionally relies on passive infrared sensors (PIR), magnetic door contacts and video surveillance. These suffer well-known limitations: PIR requires line of sight and is blocked by furniture, contact sensors monitor only specific entry points, and cameras depend on lighting and raise privacy concerns. This work converts the protected room itself into a sensor by continuously monitoring its acoustic impulse response (IR) in the ultrasonic band, where the probe remains inaudible to occupants [1].

The room is treated as a linear time-invariant acoustic channel whose impulse response h(t) is its geometric fingerprint. A fixed ultrasonic emitter periodically transmits a known probe s(t) in the inaudible band (20–48 kHz); a co-located microphone captures $y(t)=s(t)*h(t)+n(t)$. The estimated h(t) describes the direct path, early reflections and the late reverberant tail. While the room is undisturbed h(t) is approximately stationary and serves as a baseline $h_0(t)$; an intruder, opened door or displaced object alters the reflective geometry and yields a measurable deviation $\Delta h(t)=h(t)-h_0(t)$ [2].

Table 1

Detection performance for representative intrusion scenarios. Office 5.0×4.0×2.7 m, T = 23 ± 1 °C, 50 trials per scenario. Baseline $d^2 = 28 \pm 7$; threshold $\tau = 60$ (FAR ≈ $10^{-3}$, $\chi^2_{30}$).

| *Scenario* | *Mean $d^2$* | *Std of $d^2$* | *Pd at τ=60* | *Latency, ms* |
|---|---|---|---|---|
| **Door opened, 3 m** | 850 | 180 | 1.00 | 250 |
| **Person standing, 4 m** | 220 | 55 | 0.998 | 250 |
| **Chair displaced 30 cm** | 75 | 24 | 0.74 | 500 |

Among candidate probes — maximum-length sequences, Golay codes and exponentially swept sines (ESS) — ESS is preferred: deconvolution with the time-reversed sweep separates the linear room response from emitter harmonic distortion. A 100 ms sweep over 22–44 kHz, repeated at 4 Hz, sampled at 96 kHz, yields a 40 ms IR with effective SNR above 25 dB even with low-cost piezoelectric transducers.

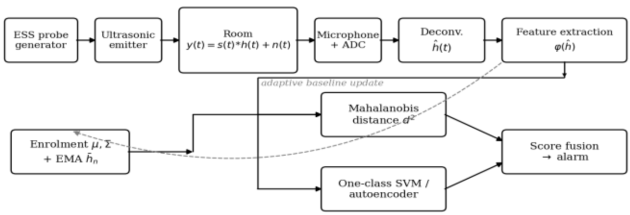


Fig. 1. Block diagram of the acoustic impulse response anomaly detection system.

The pipeline has four stages: probe emission, synchronous capture with deconvolution to recover the raw IR, feature extraction, and a two-stage classifier whose output is fused into a single anomaly score. Each stage runs in real time on a microcontroller-class device.

$$d^2(h_n) = (\varphi(h_n) - \mu)^T \Sigma^{-1} (\varphi(h_n) - \mu) \quad (1)$$

where $\varphi(h_n)$ – feature vector extracted from the n-th impulse response, $\mu$, $\Sigma$ – mean and covariance of the baseline feature distribution.

The reference is a statistical model of the room under no-event conditions. During a short enrolment phase the system captures 100–500 IRs and estimates $\mu$ and $\Sigma$.

Air temperature, humidity and airflow slowly modify the reflective geometry, so an exponential moving average $\bar{h}_n=(1-\alpha)\bar{h}_{n-1}+\alpha\, h_n$ with $\alpha\approx10^{-3}$ absorbs slow drift while leaving transient events fully visible.

Detection operates on a feature vector $\varphi(h)$ rather than on the raw high-dimensional IR.

The informative features are the energy decay curve and T60, per-band spectral magnitudes in one-third-octave bands, time-of-arrival and amplitude of the first 5–10 reflection peaks, and the short-time correlation between consecutive IRs — a coarse Doppler estimator that is highly sensitive to motion [3].

Two detectors run in parallel on $\varphi(h)$. A statistical detector uses the Mahalanobis distance (1) and fires when it exceeds a threshold $\tau$ tuned from the enrolment distribution for a target false-positive rate.

A learning-based detector — one-class SVM or shallow autoencoder — provides a non-linear decision boundary and reduces false alarms under simultaneous drift of several features.

Ambient ultrasonic emitters (fans, fluorescent ballasts) can mask the probe; frequency hopping and coherent averaging mitigate this. Slow legitimate changes are absorbed by the adaptive baseline, while sudden anomalies survive averaging and stay above threshold.

The system turns the protected room itself into a sensor: no line of sight, works in the dark and in the inaudible band, on inexpensive hardware.

Pilot data (Table 1) show human-scale intrusions produce IR deviations well above the baseline noise floor, a credible complement to PIR and contact sensors.

1. Schroeder M.R. Integrated-impulse method for measuring sound decay without using impulses. Journal of the Acoustical Society of America. – 1979. – V. 66, No. 2. – p. 497–500.
2. Farina A. Simultaneous measurement of impulse response and distortion with a swept-sine technique. *Journal of the Audio Engineering Society*. – 2000. – Preprint 5093 of the 108th AES Convention, Paris. – 24 p.
3. Stowell D. et al. Detection and classification of acoustic scenes and events. IEEE Transactions on Multimedia. – 2015. – V. 17, No. 10. – p. 1733–1746.