


Photo: Noise camera, Taichung, January 2025 · Saimmx · CC0 · Wikimedia Commons

OSINT · TECHNICAL ASSESSMENT

# THE EAR POINTS THE EYE

Optical sensors are precise and blind to almost everything outside their narrow view. Microphone arrays cover the whole sky and point only roughly. Engineers have been pairing them for years in a pattern called slew-to-cue: the ear finds the target, the eye confirms and pins it down. This first part of Sound and Sight takes one well-documented system apart to show how that works, what it achieved, and where it stops.

REPORT DETAILS

| PUBLISHED | SUBJECT | SOURCE BASE |
|---|---|---|
| November 2026 | Acoustic cueing of optical sensors (slew-to-cue) for small-drone detection and tracking | Southern University of Science and Technology field study (2023), Embry-Riddle Drone Net, National Technical University of Athens array study (2022), counter-drone technology review (2024) |
| **METHOD** | **VERDICT** | **CONFIDENCE** |
| Compare field of view, direction error and range of each sensor alone and fused, from published field tests | Acoustic arrays are the natural wide-area cue for narrow optics; ask how often the cue lands inside the optic's view | Moderate |
| **AUTHOR** | | |
| Ghulam Mujtaba, Ph.D. | | |

## BOTTOM LINE

**A camera good enough to identify a small drone at 500 metres sees about one per cent of the sky. A microphone array sees none of it but hears all of it.** That mismatch is why acoustic sensing and optics are increasingly built to work as a pair. The microphones find the drone's rough direction anywhere around the site; a gimbal swings the camera and a laser rangefinder (lidar) to that direction; the optics confirm the target and track it far more precisely than sound alone. In a 2023 field test of one such system, built at the Southern University of Science and Technology in Shenzhen, the acoustic estimate alone was off by about 4.4° in azimuth and 5.8° in elevation; with camera and lidar fused in, the errors fell to 0.8° and 0.25°, and the drone's position was known to within about 3 metres at 200 metres' range. The principle is simple and old. The design rule that follows from it is the important part: the ear only has to be accurate enough to land the target inside the eye's narrow window.

**~1%**

OF THE SKY SEEN BY A DRONE-RESOLVING CAMERA

◆ 15° × 10° TELEPHOTO VIEW

**4.4° → 0.8°**

DIRECTION ERROR, EAR ALONE VS EAR PLUS EYE

◆ AZIMUTH, ONE 2023 FIELD TEST

**481 m**

DRONE PICKED OUT OF NOISE BY THE ARRAY

◆ BEAMFORMED, 64 CHANNELS

**6 s**

FROM SWITCH-ON TO OPTICS LOCKED ON

◆ IN THE SAME TEST

SERIES · SOUND AND SIGHT · PART 1 OF 3

| | | |
|---|---|---|
| PART 1 | The Ear Points the Eye | YOU ARE HERE |
| PART 2 | Three Senses Are Better Than One | |
| PART 3 | The Layered Picture | |

## KEY QUESTIONS

### What is slew-to-cue in drone detection?

A wide-area sensor, such as a microphone array or radar, detects a possible drone and estimates its direction. That estimate 'cues' a narrow, high-resolution sensor, such as a telephoto camera or lidar on a gimbal, to 'slew' (turn) towards it, confirm what it is and track it precisely.

### Why pair microphones with cameras for drone detection?

They fail in opposite ways. A telephoto camera that can resolve a small drone at 500 m sees under 1% of the sky. A microphone array hears in every direction, day or night, but its direction estimate is several degrees off. In one 2023 system, fusing the two (plus lidar) cut direction error from about 4.4° to 0.8° in azimuth and from 5.8° to 0.25° in elevation.

### How accurate is acoustic drone localisation?

In a 2023 test of a 64-element microphone array against a drone orbiting at 180 m altitude, the acoustic direction error was about 4.4° in azimuth and 5.8° in elevation. That is enough to aim a camera with a 15° by 10° field of view, not enough on its own to guide an interceptor.

THE RULER — THE WHOLE ARGUMENT

## What each sensor brings, and what it lacks

Each sensor in a slew-to-cue chain does one job well and one job badly. The pairing works because the gaps do not overlap.

| SENSOR — WHAT IT IS | WHAT IT DOES WELL | WHAT IT LACKS |
|---|---|---|
| **MICROPHONE ARRAY** DOZENS OF MICROPHONES ON ONE BOARD | Hears in every direction, day or night; emits nothing | Direction only to a few degrees; no range; confused by loud nearby sounds |
| **TELEPHOTO CAMERA** NARROW VIEW ON A GIMBAL | Sharp direction (under 1°); shows what the object is | Sees ~1% of the sky at a time; needs light; no range |
| **LIDAR** LASER RANGING ON THE SAME GIMBAL | Range to centimetres; works in poor light | Same narrow view as the camera; must be pointed |
| **ALL THREE, FUSED** NEW | Wide search, sharp direction and range in one track | One gimbal follows one target at a time |
| **NETWORKED ACOUSTIC NODES** DROSONIC'S APPROACH | Many small passive sensors classifying sound on board; coverage grows by adding nodes | A bearing, not an image: they cue the camera rather than replace it |

● THE COMBINATION TESTED IN THE 2023 STUDY DISCUSSED BELOW

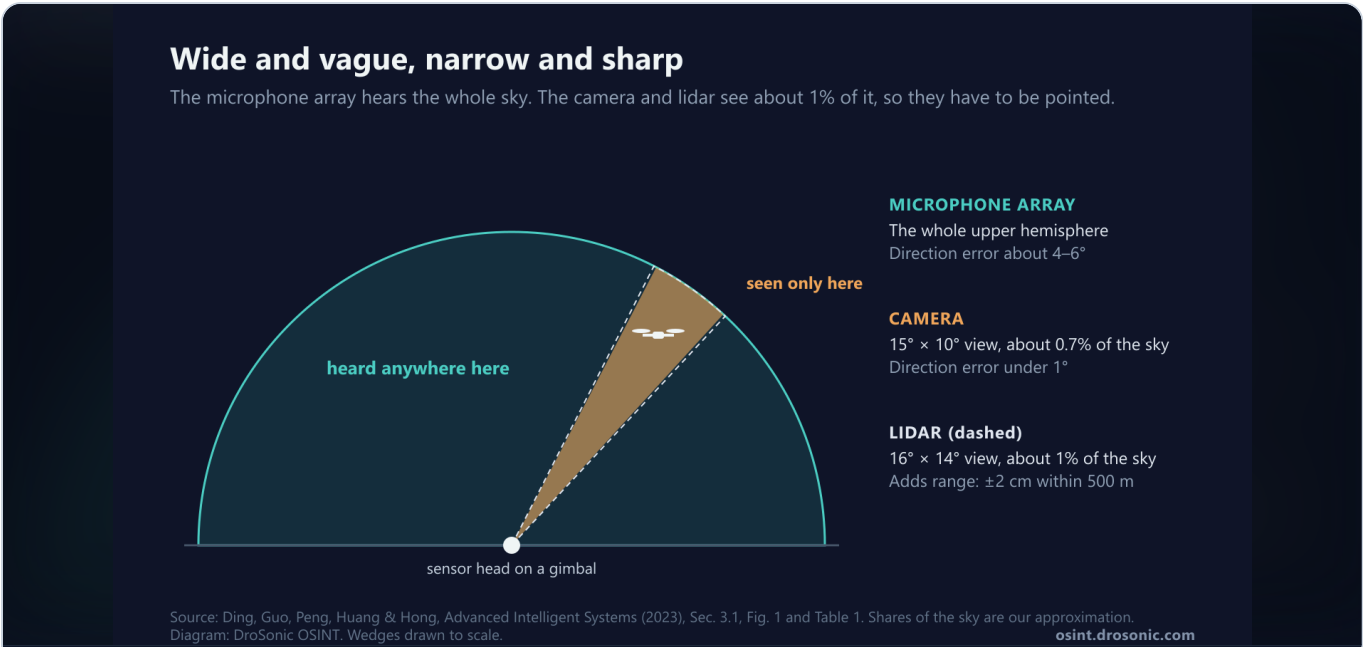
FIGURES FROM DING ET AL. (2023)

FIELD 01 — THE STRAW PROBLEM

# Why sharp sensors need pointing

Every camera trades field of view against detail. An image sensor has a fixed number of pixels, so widening the view spreads them thinner and a small, distant drone shrinks to a handful of pixels. To see far, systems use telephoto lenses, which narrows the view further. Advanced Intelligent Systems↗ The camera in the Shenzhen system covers 15° by 10° and puts more than 100 pixels on a 30-centimetre drone at 500 metres. Its lidar covers about 16° by 14° and measures distance to 2 centimetres within 500 metres. Advanced Intelligent Systems↗

By our approximation, each of those windows is about 1% of the sky above the horizon. Scanning the whole sky with them would be like searching a stadium through a drinking straw. Spinning lidars do not solve it either: in one study cited by the authors, a 64-line automotive lidar picked up a small drone at up to 50 metres with only about three points, and missed it entirely when it flew between the scan lines. Advanced Intelligent Systems↗




**COVERAGE AGAINST PRECISION — THE GAP THE PAIRING CLOSES**

The microphone array covers everything roughly; the optics cover a sliver precisely. Wedges drawn to scale.

DIAGRAM: DROSONIC OSINT · ORIGINAL · FROM DING, GUO, PENG, HUANG & HONG, ADVANCED INTELLIGENT SYSTEMS (2023)

FIELD 02 — AN OLD IDEA

## Hear it, then look at it

Cueing a narrow sensor from a wide one is how air defence has long worked; the Sound and Sight question is which sensor does the cueing. Before radar, acoustic locators and searchlights worked in pairs, as our report Listening to the Sky describes. For small drones the idea returned in research. In 2018 Embry-Riddle Aeronautical University described "Drone Net", a network of passive rooftop sensors meant as a cheaper alternative to radar at small airports: when its all-sky camera and acoustic network detect something moving, "it sends a message to begin electro-optical and infrared sensing by tilting and panning the camera", over an area of about a square kilometre. Unmanned Airspace↗ The team published the design at the AIAA SciTech Forum in January 2019. AIAA↗

A French–German team took it to an extreme, pairing a four-microphone array with a short-wave infrared camera whose view was under 1.2° wide; without the acoustic estimate, that camera would have been close to useless. Advanced Intelligent Systems↗

The same pattern is already on city streets. Since January 2021 Taiwan has used roadside "noise cameras" that photograph vehicles louder than 86 decibels on slower roads and 90 decibels on faster ones; the sound level decides when the camera fires. Taiwan News↗ The one in the photo above, in Taichung, is a microphone on a pole above a camera.

Commercial counter-drone systems mostly pair radar with electro-optical and thermal cameras. A 2024 review of the market found that "few incorporate acoustic sensors, or at least, their presence in the specifications is not clearly indicated", and that lidar is rarely included at all. Preprints.org↗

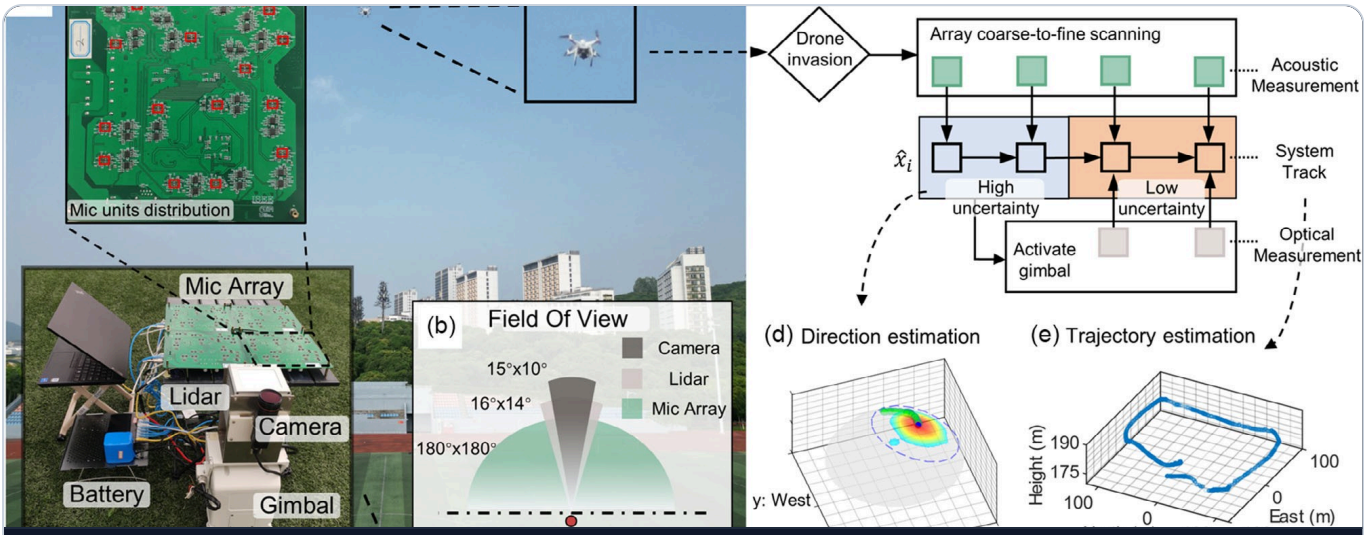
---

FIELD 03 — INSIDE ONE SYSTEM

## Microphones, a gimbal, a camera and a laser

The Shenzhen team's system is unusually well documented, which is why we use it here. Its array is a 40-centimetre square board carrying 64 elements of four microphones each, laid out in an irregular pattern chosen by a genetic algorithm to suppress false "side" directions. The board's size sets its angular resolution: about 8.7° for a 5 kHz tone. It listens over the whole hemisphere. A neural network trained on drone recordings mixed with wind, birds and insects cleans each frame of sound before the array estimates direction, raising the test signal-to-noise ratio from 6.5 to 11.7. In a recorded trial, a conventional beamformer was pulled off target by a nearby bird; the filtered one was not. Advanced Intelligent Systems↗

When the array hears a drone, the gimbal turns the camera and lidar to the estimated direction. Once the target is in view, a tracking filter merges all three sensors' measurements into one track and keeps the gimbal on it. Advanced Intelligent Systems↗




**THE SYSTEM IN THE FIELD — COARSE TO FINE**

The 64-channel array (top of the stack) hears the drone; the gimbal below turns the camera and lidar onto it. The flowchart shows uncertainty falling once optical measurements arrive.

IMAGE: DING, GUO, PENG, HUANG & HONG, ADVANCED INTELLIGENT SYSTEMS (2023), FIG. 1 · CC BY 4.0 · DOI.ORG/10.1002/AISY.202300251

At the other end of the cost scale, a team at the National Technical University of Athens built a cross of four off-the-shelf microphones that estimates direction and runs a machine-learning check on whether the sound is a drone. In outdoor trials it detected drones reliably at more than 70 metres.[Sensors↗] Even a cue that rough is enough to point a camera with a wide-enough lens.



## FOUR MICROPHONES — THE LOW-COST END

The Athens array being calibrated in an anechoic chamber. Four elements are enough for a direction estimate.

IMAGE: KYRITSIS, MAKRI & UZUNOGLU, SENSORS (2022), FIG. 1 · CC BY 4.0 · DOI.ORG/10.3390/S22228659




## LIDAR — RANGE, ONCE POINTED

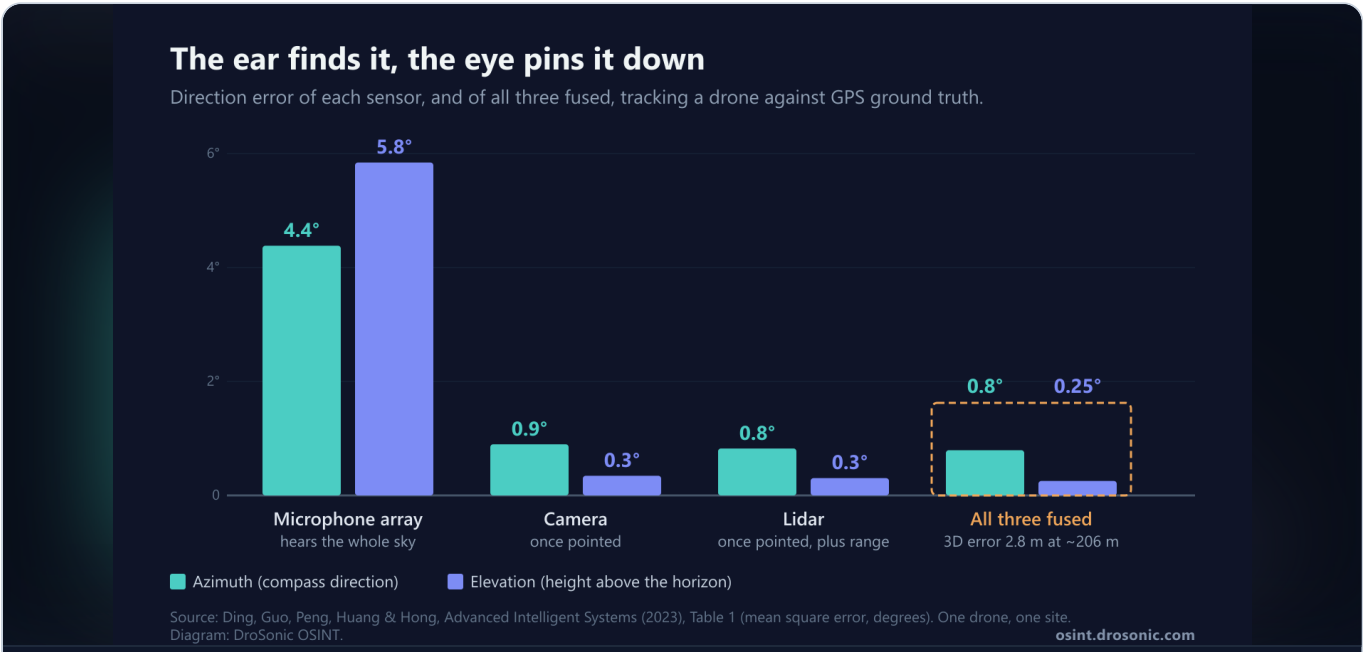
Livox lidars at a trade show; the Tele-15 on the right is the long-range model used in the Shenzhen system.

PHOTO: DLLU · CC BY-SA 4.0 · WIKIMEDIA COMMONS

FIELD 04 — THE NUMBERS

# From degrees to a fraction of one

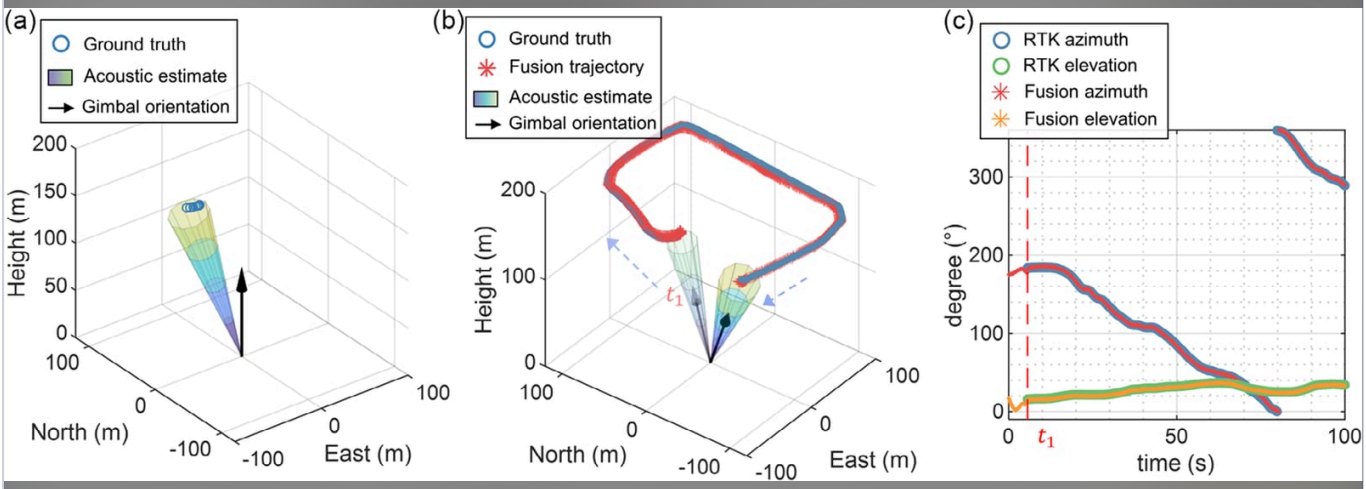
The team flew a drone carrying a high-precision satellite positioning receiver in an orbit 180 metres above an open field, with the system at the centre and its gimbal pointing straight up. The array found the drone and swung the gimbal; the optics locked on and began measuring six seconds after switch-on. Over the flight, the acoustic direction error averaged 4.4° in azimuth and 5.8° in elevation. Camera and lidar alone, once pointed, were each under 1°. Fusing all three gave the best result: 0.8° and 0.25°, and a position error of 2.8 metres at an average range of 206 metres. Advanced Intelligent Systems↗




**THE EAR FINDS IT, THE EYE PINS IT DOWN — DIRECTION ERROR BY SENSOR**

Sound alone is good enough to aim; optics make the track precise. Fusion adds a little more on top of the best single optic.

DIAGRAM: DROSONIC OSINT · ORIGINAL · FROM DING, GUO, PENG, HUANG & HONG, ADVANCED INTELLIGENT SYSTEMS (2023), TABLE 1

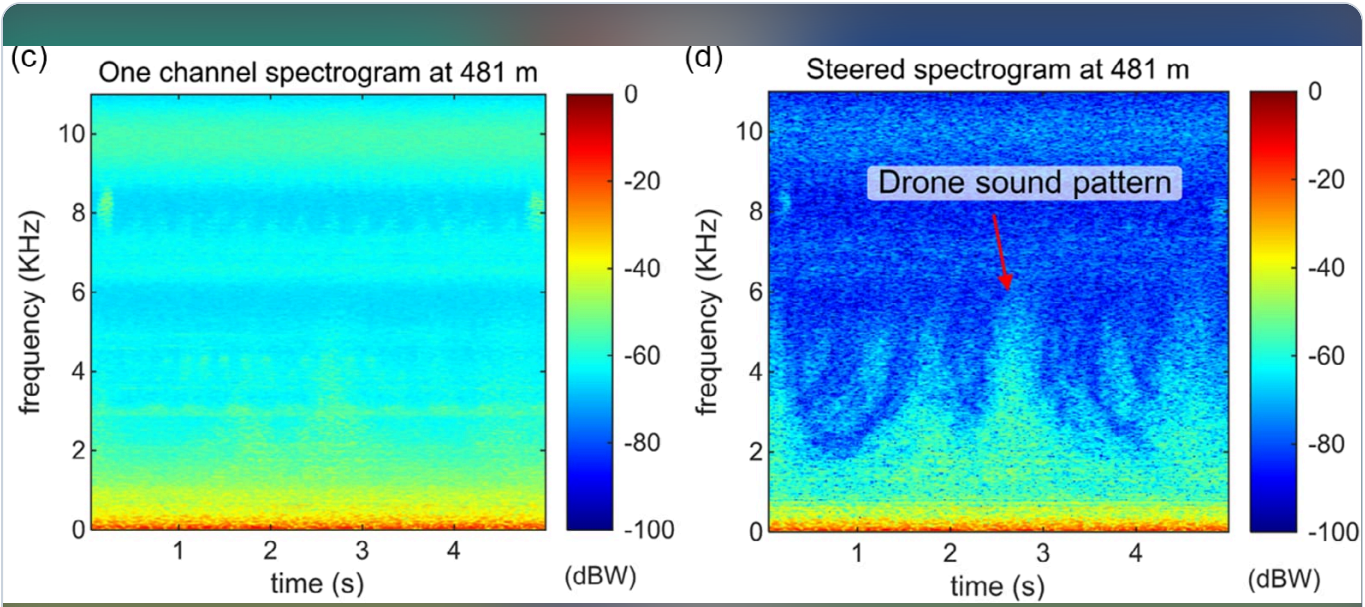



**CUE, SLEW, TRACK — THE FLIGHT IN THREE PANELS**

Left: the acoustic cone before the gimbal moves. Centre: the fused track on the true path. Right: after the optics lock on (t1), the estimate follows the truth.

IMAGE: DING, GUO, PENG, HUANG & HONG, ADVANCED INTELLIGENT SYSTEMS (2023), FIG. 4 · CC BY 4.0 · DOI.ORG/10.1002/AISY.202300251

Range is the other half of the case. Combining 64 channels lowers the background noise in proportion to the number of channels, which is what lets the array hear further than any single microphone. At 481 metres the drone was invisible in the spectrogram of one microphone and clear in the combined signal. Regulations capped the tests at 500 metres; by extending the measured trend, the authors estimate that the signal would sink below the background at about 1.3 kilometres. Advanced Intelligent Systems↗




**481 METRES — ONE MICROPHONE HEARS NOISE, SIXTY-FOUR HEAR A DRONE**

The same moment recorded by a single channel (left) and by the whole array steered at the drone (right).

IMAGE: DING, GUO, PENG, HUANG & HONG, ADVANCED INTELLIGENT SYSTEMS (2023), FIG. 3C-D · CC BY 4.0 · DOI.ORG/10.1002/AISY.202300251

> **The ear does not need to be precise.** It needs to be precise enough to land the target inside the eye's window.

FIELD 05 — THE DESIGN RULE

## Match the cue to the window

Put the numbers side by side and a practical rule appears. This is our analysis, not the authors'. The camera's view is 15° wide and 10° tall, so a target is in frame if the cue is within about 7.5° left or right and 5° up or down. The array's average azimuth error, 4.4°, sits comfortably inside the first; its 5.8° elevation error is at the edge of the second. In this system, then, the cue usually lands in frame across, but sometimes needs a short search up or down. A sensor whose view is ten times narrower, like the infrared camera in the French–German study, needs a cue ten times better, or a search pattern. This is the rule DroSonic designs its acoustic cue around: a bearing good enough to put a camera, or a crew's eyes, on the right patch of sky. DroSonic↗

The study's limits are equally clear. It flew one drone, at one site, against a clear sky, where the authors note that simple image processing was enough to pick out the target; cluttered backgrounds, birds and night would be harder for the camera. Its array can follow several drones at once, but one gimbal can only watch one; the authors suggest more gimbals or faster slewing. Advanced Intelligent Systems↗ A swarm is the case where a wide-area acoustic layer matters most and a single optical turret matters least.

For a buyer or integrator, the questions follow:

1. **How large is the acoustic direction error, in azimuth and elevation, at the ranges that matter?** Compare it with half the optic's field of view.
2. **How long from first sound to optics on target?** Six seconds is one test; ask for the distribution.
3. **How many targets can the optics follow at once?** The ear can hear several; one turret can look at one.
4. **What happens when the optic cannot confirm?** At night, in cloud or against clutter, does the acoustic track still reach the operator?
5. **Was it tested against noise, not just at range?** Birds, traffic and wind are where acoustic cues go wrong.

CONFIDENCE

### How sure we are, and what would change it

**High confidence** in the principle: wide-area acoustic sensing can cue narrow optics, and fusing them gives a more precise track than either alone. Independent research groups, a university airport project and civic noise enforcement all use the pattern.

**Moderate confidence** in the specific figures. They come mainly from one peer-reviewed field study with one drone, one site and clear skies, with ranges capped at 500 metres; the 1.3-kilometre figure is the authors' extrapolation, not a measurement. The design rule in Field 05 is our reading of their numbers.

**What would change it:** published trials of acoustic cueing against several drones at once, at night and in cluttered or noisy environments, and data from fielded counter-drone systems, which rarely publish their acoustic performance.

---

DROSONIC — OUR APPROACH

## Built to cue, not to replace

DroSonic builds the ear in this pairing: networked passive acoustic sensors that detect, classify and locate drones by the sound of their propellers, running neural networks on low-cost edge hardware and emitting nothing. DroSonic↗ Its U.S. provisional patent application, filed in September 2026, covers detection, classification, localisation and tiered defeat cueing, the step of handing a track to whatever looks or acts next. DroSonic↗

That is the division of labour this report describes. The acoustic layer does not need a camera's precision, only enough to point one. A DroSonic prototype was tested against live drones at the DIRT Days 26.04 field exercise in September 2026; DroSonic reports that it heard drones other sensors missed. DroSonic↗ No performance figures from that test have been published, so we quote none here.



**THE EAR IN THE FIELD — BEFORE ANY CAMERA LOOKS**

DroSonic's test position at DIRT Days 26.04. In a layered system, this is the sensor that hears first and points the others.

PHOTO: DROSONIC · FIELD-TEST WRITE-UP

CLOSE — WHAT IT MEANS

## The ear is the first layer

Acoustic sensing is often compared with cameras and radar as if they were rivals. The evidence here points the other way: a microphone array is the natural wide-area trigger for sensors that are precise but blind outside a narrow window. It covers the whole sky, works day and night and emits nothing, and all it has to do is be right to within half the width of the camera's view. Ukraine's national acoustic network is the large-scale version of the same logic, cueing mobile fire teams rather than turrets. Defense One↗

DroSonic builds passive acoustic detection to play exactly that role: hear the drone first and hand that cue to whatever sensor or team acts next. Part 2 of Sound and Sight looks at the next step: when several sensors report at once, how their evidence is actually combined, and what each one adds when the others fail.

## ABOUT DROSONIC

DroSonic builds passive acoustic drone detection: networked sensors that listen for drones instead of their radio signals, using acoustic AI and edge computing. DroSonic OSINT is our open-source research into the threat that technology is built to meet.

## THE RECORD SO FAR

| | |
|---|---|
| OCT 2026 | Joined the SOF RACER consortium, which connects USSOCOM with commercial technology companies |
| SEPT 2026 | Prototype tested against live drones at DIRT Days 26.04, a field exercise supported by CMI2 with U.S. Central Command and U.S. Army DEVCOM |
| SEPT 2026 | Technical presentation at NDIA's Emerging Technologies for Defense conference on detecting fibre-optic and radio-silent drones |
| SEPT 2026 | Filed U.S. provisional patent 64/148,522 on passive acoustic detection, classification, localisation and defeat cueing |
| JUL 2026 | Pitched to the U.S. Space Force at Space Systems Command's Reverse Industry Engagement Day, Colorado Springs |

How DroSonic detection works → · Where it is used → · Talk to the team →

## THIS REPORT ONLINE

https://osint.drosonic.com/assessments/acoustic-cueing-cameras-lidar.html

**Cite as:** Mujtaba, G. (2026, November 13). *The Ear Points the Eye: Acoustic Cueing: How Microphones Aim Cameras and Lidar at Drones*. DroSonic OSINT. https://osint.drosonic.com/assessments/acoustic-cueing-cameras-lidar.html

**Reusing this report.** Text and original diagrams © DroSonic. You may quote this report and share its link with credit to DroSonic OSINT; to republish it in full, reuse our diagrams or use it commercially, contact us first. Photos and third-party figures keep the licences shown in their credits.

DroSonic OSINT: open-source assessments on drones and counter-UAS from DroSonic, passive acoustic drone detection. drosonic.com · sudo@drosonic.com



## ABOUT THE AUTHOR

**Ghulam Mujtaba, Ph.D.**

FOUNDER AND CEO, DROSONIC

Ghulam Mujtaba, Ph.D., is the founder and CEO of DroSonic and leads its work on passive acoustic drone detection. A computer engineer with 13 years of experience building machine-learning systems for edge hardware, he holds a Ph.D. in computer engineering from Gachon University in South Korea, was an NSF postdoctoral fellow at West Virginia University in the United States, and is an Assistant Professor of AI and Data Science at Regis University in Denver, Colorado. An IEEE Senior Member, he has published more than 20 papers on deep learning, computer vision and multimedia for edge devices.

PROFILE ↗ LINKEDIN ↗ ORCID ↗

## ALL SOURCES (10)

1. Advanced Intelligent Systems — Ding et al. 2023, Sec. 2.1: fixed pixel count; wider FOV, lower angular resolution; telescope lenses reduce FOV
2. Unmanned Airspace — Oct 2018: all-sky + acoustic detect movement, cue EO/IR camera to slew and track; ~1 km2
3. AIAA — Siewert et al., Slew-to-cue electro-optical and infrared sensor network for small UAS detection, tracking, and identification, AIAA SciTech 2019
4. Taiwan News — Dec 2020: noise cameras nationwide from Jan 1; snapshot of vehicles over 86 dB (<50 km/h) or 90 dB (50-90 km/h)
5. Preprints.org — Gonzalez-Jorge et al. 2024 review (preprint): most C-UAS combine radar and EO/IR; few incorporate acoustic sensors; LiDAR not typically included
6. Sensors — Kyritsis, Makri & Uzunoglu 2022: 4-element cross-shaped array, moderate cost; DOA + ML identification; reliable detection beyond 70 m
7. DroSonic — Passive acoustic detection: neural networks on low-cost edge hardware, no emissions
8. DroSonic — U.S. provisional patent application 64/148,522, September 2026
9. DroSonic — DIRT Days 26.04 field test, September 2026
10. Defense One — Ukraine's cheap acoustic sensors helping troops fight waves of drones, July 2024