Acoustic Imaging with High-gain Microphone Arrays

Donald W. Miklovic, Joseph Church, Jeremy Banning
Science Applications International Corporation, Acoustic and Marine Systems Operation
San Diego, Calif.
donald.w.miklovic@saic.com, joseph.r.church@saic.com , jeremy.f.banning@saic.com

ABSTRACT

Sound is a potentially valuable but little used sensing modality for intelligence, surveillance, and reconnaissance (ISR) in the urban environment. It is potentially valuable because it can provide information difficult or impossible to obtain with other modalities. For example, sound can provide non-line-of-sight sensing beyond tree lines, around corners, and into buildings, something impossible to do with optical, electro-optical or radar imaging systems. When there is direct line-of-sight, sound can also be important for activities that have an acoustic signal component. These include weapons fire, the movement of small unmanned aerial vehicles (UAVs) or ground vehicles, and voice communications. And unlike many ISR sensors, sound is just as effective night or day, through clouds, smoke, fog or dust. However, sound is little used for ISR because conventional acoustic sensors suffer from poor range coverage in noisy environments and have no inherent ability to separate, locate or track multiple contacts – something that conventional imaging sensors do very well.

This paper presents innovative low-cost microphone array and acoustic imaging technology that provides the best of both worlds – it has the above mentioned advantages of sound exploitation, combined with the long-range coverage and precise location and tracking capability typically associated with the conventional non-acoustic ISR imaging modalities. This is essentially a new sensor modality – a new view of the battle space – which will substantially enhance ISR capabilities in a variety of scenarios.

Results are presented for a prototype system that utilizes an array of 128 microphones. The prototype can produce multiple (simultaneous) electronically-steered broadband audio beams in real time, as well as images of the incoming sound over ± 90 degrees in azimuth and elevation. Performance is characterized in terms of system noise, array gain, angular resolution, and beam-pattern response. Examples of acoustic images and broadband audio results illustrating the potential for ISR are presented.

ABOUT THE AUTHORS

Donald Miklovic, Ph.D., is a chief scientist at SAIC with over 30 years experience in cutting-edge remote sensing R&D. He is a specialist in broadband acoustics for both underwater and in-air applications, but also has expertise in radar and optical systems. His latest initiatives include R&D for real-time high-gain microphone arrays and acoustic imaging systems for improved situational awareness in the urban environment. He holds a Ph.D. in applied mathematics from the California Institute of Technology.

Joseph Church is an electrical engineer at SAIC with four years experience in the development of advanced acoustic prototypes and digital signal processing. He holds an M.S. in electrical engineering from San Diego State University.

Jeremy Banning is an electrical engineer at SAIC with over three years experience in the development of advanced acoustic prototypes and experimental acoustics. He holds a B.S. in electrical engineering from University of California, San Diego.

Acoustic Imaging with High-gain Microphone Arrays

Donald W. Miklovic, Joseph Church, Jeremy Banning
Science Applications International Corporation, Acoustic and Marine Systems Operation
San Diego, Calif.
donald.w.miklovic@saic.com, joseph.r.church@saic.com , jeremy.f.banning@saic.com

INTRODUCTION

Passive acoustics is a potentially valuable sensing modality for intelligence, surveillance, and reconnaissance (ISR) because it can provide important information that is difficult or impossible to obtain with other modalities. First of all, it can provide information about activities that are nearby but outside the line-of-sight. Two operationally relevant situations where this would be of value are illustrated in Figure 1. In the riverine environment, line-of-sight to potential threats is often obscured by riparian foliage. On land, *invisible* threats can be inside buildings, around corners, or beyond shrub or tree lines. In either case, smoke, haze or fog may preclude visibility to the point where acoustics is the only effective sensing modality.

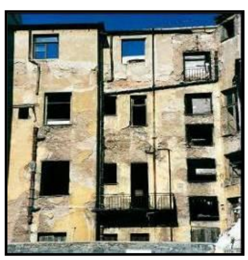

Figure 1. Situations where acoustics can provide information about activities that are invisible to other sensing modalities.

In addition, acoustics can be important even when there is a clear line-of-sight to the target (cf. Figure 2). A speaker may be readily seen, but what he is saying, the language or dialect he is using, and his biometric voice characteristics can only be determined through sound. Or the approach of a small unmanned aerial vehicle (UAV), which could be carrying an improvised explosive device (IED) or biological weapon of mass destruction (WMD), may be identified and tracked much sooner with sound than by any other means.

Figure 2. Situations where acoustics can be an important compliment to other other sensing modalities. Left: speakers; Right: potential hobby UAV threat.

Unfortunately, conventional microphones suffer from poor range coverage in noisy environments; large aperture listening devices such as the parabolic dish are cumbersome to use; and neither sensor type has the inherent ability to separate, locate or track multiple contacts – something that is important for effective ISR. On the other hand, typical ISR imaging sensors that exploit optical, electromagnetic, infrared (IR) or radar signals have excellent location and tracking capability, but fail in the situations described above where acoustics excels.

However, a revolutionary leap in acoustic sensing capability is possible with a spatial array of a large number of microphones. Such an array-sensor would have longer range and better coverage, be able to precisely locate and track many simultaneous sources of sound, and even provide passive acoustic images. This would essentially be a new modality – a new view of the battle space – which would substantially enhance ISR capabilities in a variety of scenarios.

This paper describes the design and development of low-cost prototype microphone array and acoustic imaging technology that can be transitioned to a variety of ISR applications.

LIMITATIONS OF CONVENTIONAL ACOUSTIC SENSING TECHNOLOGY

There are several kinds of devices that are available commercially for sound exploitation. Their primary limitations for ISR are discussed below.

Microphones

The simplest and most common sensor for exploiting sound is the microphone. One of its major limitations is that it has a very short range for all but the loudest of sounds. Its effective range for monitoring voice is about 20 feet in typical urban environments, and even less in chaotic aural environments. Its second major limitation is that it has no capability to determine the direction of transient sounds or to separate multiple sources of sound.

Direction Finders

Acoustic direction finders, often used in gunshot location devices, are typically small arrays of four to seven microphones. These can work well for loud sounds from a single source, but cannot detect and locate weaker sounds or even locate loud sounds from multiple, simultaneous sources.

Large Aperture Devices

To detect hard-to-hear sounds in distributed noise, a device that can preferentially collect sound from a specific direction is required. Due to the fundamental limit imposed by the diffraction of sound, performance is proportion to aperture size relative to the wavelength, and substantial improvements require a large aperture vis-à-vis microphones and direction finders. There are two distinct approaches to the aperture design: a continuous aperture, such as a parabolic dish, or a discretely sampled aperture, such as an array of microphones.

The most common example of a large-aperture listening device is the parabolic dish (Figure 3), which can be purchased commercially (e.g. see Crystal Partners, Inc). This device uses the geometry of the dish to preferentially collect sounds that are coming from the direction normal to the dish (i.e. the pointing direction). A microphone at the focus of the parabola provides an *analog sum* of the sound that arrives at the dish and reflects to the microphone. Waves coming into the dish from the pointing direction will be summed in phase, while sounds from other directions will be summed out of phase, giving a net enhancement of sound from the pointing direction.

Figure 3. A commercial parabolic dish microphone used to improve hearing.

Although the parabolic dish improves hearing remarkably well in many cases, it has several serious drawbacks that preclude it from being effective in many ISR applications. These are:

1) Cumbersome footprint. The shape dictated by the physics of the solution is essentially 3D. The microphone must be placed at the focus of the parabolic dish, which makes it vulnerable to damage.
2) Mechanical or manual steering. To change the listening direction, the dish has to be physically pointed to a new direction. Automation of steering requires substantial moving parts, motors, and supporting structure.
3) One direction only. A parabolic dish can only improve hearing in one direction at a time; continuous coverage of 10 directions would require 10 dishes.
4) No inherent directional information. A detected transient sound could be a weak sound in the pointing direction, or a strong sound from another direction, with no way to sort out the ambiguity without further information.

Microphone arrays can overcome these limitations, and are discussed in detail in the next section.

ADAVANTAGES OF THE MICROPHONE ARRAY

A revolutionary leap in capability over continuous aperture devices such as the parabolic dish is possible with an array of microphones. An example of such an array of 24 microphones (produced by G.R.A.S. Sound and Vibration) is shown in Figure 4.

A microphone array collects sound from a spatial aperture, as does the parabolic array; but in contrast, a microphone array can collect all the spatial acoustic information that arrives at the aperture – not just a single channel representing the sum of the acoustic field over the aperture. Through digital signal processing of this field, substantially more information can be provided to the user. Roughly, the increase in information available is in proportion to the number of microphone channels – a 100-channel array can potentially provide 100 times more information than a parabolic dish.

Figure 4. A commercial off-the-shelf array of 24 microphones. Sensor and preamp elements (inset) cost about $500 each.

If there is sufficient memory and computing power, many listening directions can be computed from the array simultaneously in real-time. By displaying the sound intensity as a function of elevation and azimuth for a large number of directions, the acoustic analogue of an image or a video can be formed.

In contrast to the limitations of the parabolic dish, microphone arrays have the following advantages:

1) <u>Electronic steering.</u> Listening beams can be redirected electronically in a small fraction of a second. Steering can be software-controlled and adapted to a specific application and the immediate acoustic environment. This allows, for example, scanning of large areas and tracking of targets.
2) <u>Many directions simultaneously.</u> The number of possible simultaneous directions is determined by the computer and memory resources available. Many targets can be monitored continuously with the same sensor.
3) <u>Directional information.</u> By comparing the intensity on various beams, or by other direction-finding techniques using multiple sensors, the direction of the sound can be precisely determined.
4) <u>Array shading and space-time adaptive processing (STAP)</u>. Advanced beamforming techniques such as array shading or STAP can be applied to better control side lobes and provide even greater hearing enhancement for specific sources or directions, especially in chaotic aural environments where the improvement is needed the most.

However, commercial off-the-shelf (COTS) array technology such as depicted in Figure 4 has several major disadvantages which make it unsuited for military operations:

- A cumbersome 3D footprint due to the design of COTS-array microphones, which are large in the direction normal to the plane of the array (Figure 4, inset).
- Little or no real-time capability. For a large number of channels, data is typically collected and processed off-line with a standard PC. For operational situational awareness, results are needed immediately.
- Expensive, due to the cost of individual microphones/preamps. The COTS array mics/preamps shown in Figure 4 costs about $500 per element, or $64,000 for microphones alone to assemble a 128-element array. This precludes their use in large numbers in an environment where sensors may often be damaged.

In the next section, we investigate the design of arrays better suited for military operations.

DESIGN OF LOW-COST HIGH-GAIN ARRAYS FOR MILITARY OPERATIONS

Sensors

Microphone array performance is generally proportional to the number of microphones employed. If the unit cost of microphones is high, arrays with a large number of microphones become economically impractical.

With this in mind, various companies build what are called *array microphones*. These microphones are specifically designed to keep costs low so that they can be affordable in large numbers, while still providing accurate low-noise measurements of the acoustic pressure with a flat response over a broad band of frequencies. For example, G.R.A.S., a world leader in sound and vibration measurements, provides an excellent condenser *array microphone* with integrated preamp (e.g., G.R.A.S. Type 40 PR, shown in the inset of Figure 4) at a cost of about $500 each. This is a major cost reduction when compared to a typical free-field reference microphone and preamp at a unit cost of $2,000, but still too costly for mass production. Furthermore, the footprint of such high-end microphones is long in the direction that would be normal to an array surface in a typical mounting configuration, making it awkward for array use.

Therefore, custom-designed low-cost alternatives are highly desirable. Since our purpose is to sense sound for situational awareness, the accurate measurement of pressure, or a flat response over frequency is not required, and possibly not even desirable. We do, however, want adequate bandwidth and sufficiently low sensor self-noise so that the inherent ISR capabilities of a low-cost alternative is not significantly less than what can be achieved with the best microphones.

One viable alternative is an electret condenser microphone (ECM) such as the Panasonic® WM-64 (Panasonic Corporation of North America) shown in Figure 5. This microphone element, without supporting preamp electronics, has a unit cost less than $2 and can operate on as little as 2 VDC.

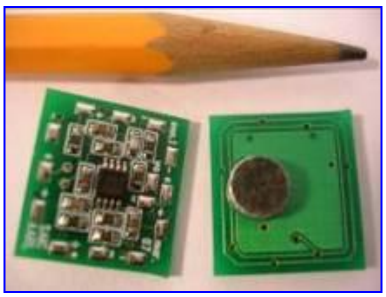

Figure 5. A low-cost electret condenser microphone with integrated preamp and filter circuitry produced by SAIC

We have made careful measurements to characterize the performance of these sensors, and they are more than adequate for use in high-performance arrays. The performance metric that we consider most important is the microphone electronic self-noise as a function of frequency. This is an important metric because the limiting noise is quite often the sensor self-noise, especially at higher frequencies or in quiet environments. Acoustic noise generally decreases rapidly with frequency, while the self-noise is approximately independent of frequency, and so the self-noise becomes the performance limiter beyond some frequency, which is typically in the range of 3-10 kHz.

A key result on self-noise is shown in Figure 6. The self-noise spectra in equivalent sound pressure level (SPL, dB re 20 uPa) are shown for both the Panasonic WM-64 ECM and a free-field measurement condenser microphone (Earthworks, Inc.'s M50, $1,200 unit cost, see Earthworks Precision Audio). The ECM was surface-mounted as it would be in a conformal array, while the measurement microphone was used in its normal freestanding configuration. We see that the low-cost ECM does as well or better than the expensive measurement microphone in terms of self-noise over the frequency range of 2-20 kHz. The somewhat higher self-noise of the ECM below 2 kHz is irrelevant because ambient sound levels are almost always higher than this at these frequencies.

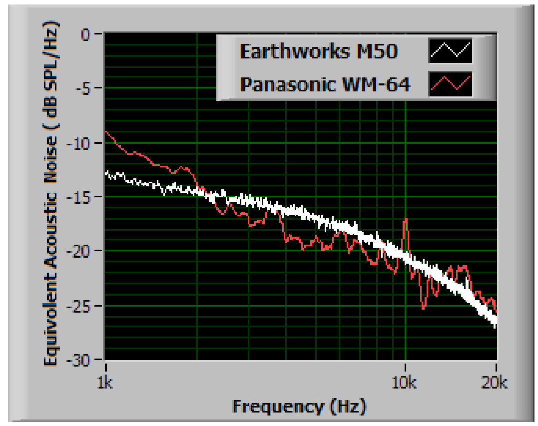

Figure 6. Self-noise for an expensive condenser measurement microphone (white) and an inexpensive electret condenser microphone (red).

Another measurement of interest is the microphone sensitivity as a function of frequency. Measurement microphones are essentially flat across their operating band, a feature that results in their high cost, but the ECMs are not, even though factory specificationd suggest that they are. The factory specification for sensitivity of the WM-64 is -45 dB (re Volt/Pa) ±4 dB at 1 kHz, with a frequency dependence that is essentially flat as shown in Figure 7 (Panasonic). The surface mounting would add 6 dB at the lower frequencies to the factory specification for sensitivity, giving an expected surface mounted sensitivity of -39 dB (re Volt/Pa) at 1 kHz. The surface mounting would also increase sensitivity at the higher frequencies, but less so than at the lower frequencies, so the effect of surface mounting is to decrease sensitivity as a function of frequency.

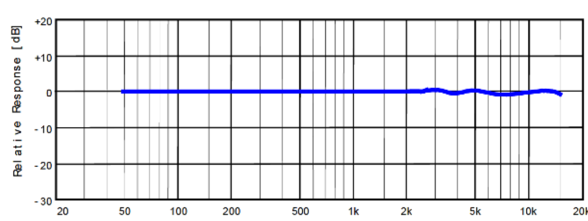

Figure 7. Relative sensitivity of the WM-64 as found in the product literature.

Figure 8 shows our measurement of sensitivity for the Panasonic WM-64 in the surface mount configuration. We get good agreement with the expectation of -39 dB (re Volt/Pa) at 1 kHz but see an increase of about 8 dB from 1 kHz to 20 kHz, which would not be expected based on the factory specification. This increase in sensitivity should not degrade performance, and may even be advantageous, but it needs to be considered in the system design and interpretation of data.

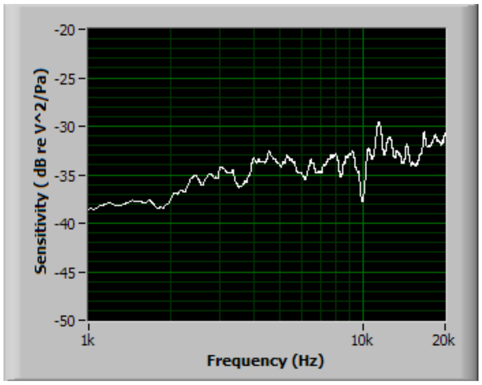

Figure 8. Measured sensitivity of the WM-64 in a surface mount configuration.

Analog Electronics

The ECMs discussed above need high input-impedance preamplifier circuits to increase the signal voltage prior to digitization. Such circuits are easily fabricated using operational amplifiers (Horowitz and Hill, 1989), and can be readily customized to match any array mounting application (see Figure 5 for one of our configurations). These circuits can be assembled in large quantities at a unit cost below $50.

A simple design for such a circuit is shown in Figure 9. The microphone element is effectively a current source. The 2.2 k Ω resistor converts the current signal to a voltage, which is AC-coupled to a non-inverting operational amplifier based on the low noise LM833 op-amp. The 33 k Ω and 1.1 k Ω resistors provide a voltage gain of 30 dB; the 120 nF capacitor and 22 k Ω resistor provide a high pass filter at 60 Hz to remove low frequency noise and sound prior to amplification, which provides better dynamic range at the higher frequencies; and the 160 pF capacitor and the 33 k Ω resistor provide a low-pass filter, cutting off the amplifier gain at 30 kHz, reducing potential for parasitic oscillation and providing some anti-alias filtering. This signal can then be sent to additional preamp and other analog filtering stages.

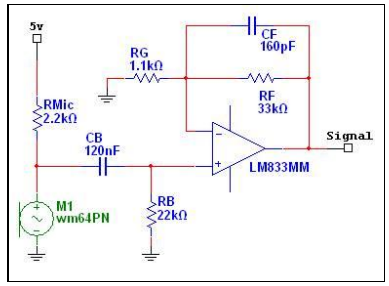

Figure 9. A simple preamp circuit design for use with low-cost electret condenser microphones.

To aide in the circuit design process, Figure 10 provides the equivalent input electronic noise spectrum for the ECM. Our design keeps the noise from the electronics, also shown in Figure 10, well below the electronic self-noise of the microphone itself across the entire band of interest.

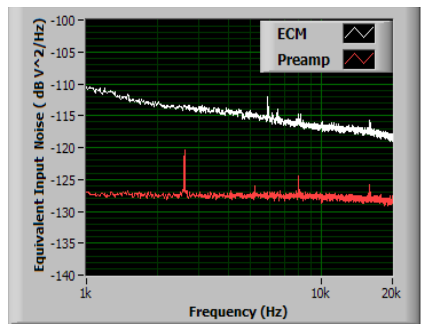

Figure 10. Equivalent input electronic noise for the ECM WM-64 microphone alone (white) and the custom preamp circuit alone (red).

Beamforming

Beamforming converts signals from the array of microphones to highly directional listening beams in one or more directions. There are a variety of signal-processing techniques of varying sophistication to do this. It is beyond the scope of this paper to cover this topic adequately, and the reader should refer to one of the many excellent references on the beamforming process (e.g., Brandstein and Ward, 2001). However, a few comments with regard to microphone array applications are in order.

In general, sound from a certain direction θ will arrive at microphone *i* with a relative time delay $\Delta t_i(\theta)$ that is determined by the direction θ and the relative position of the microphone as shown in Figure 11. A common of beamforming method is called *time-delay-and-sum* beamforming. In this process, the data is digitized with an analog-to-digital converter (ADC), each microphone signal is advanced to undo the time delay caused by the geometry, and then the signals are summed to form a single audio channel or beam. In this sum, sounds from the chosen direction will be added *in phase*, while sounds from other directions will be added *out of phase*, with the net result being an emphasis of the sound from the chosen direction.

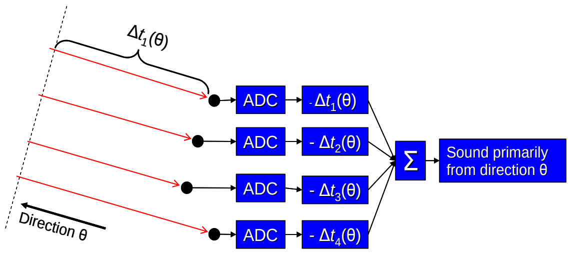

Figure 11. Illustration of the beamforming process for four microphones (●).

More formally, if the measurement on the *i* th microphone is m_i (t), the beamformer output for the chosen direction is

$$B(t) = \sum_i m_i(t + \Delta t_i) .$$

Clearly this process can be repeated for as many directions a desired.

For narrow band signals, the time-delay process is equivalent to a single change in phase for each sensor. In such cases the time delay operation can best be carried out computationally as a complex multiply on the incoming data to produce the desired phase change (e.g. the *phased* array). A major advantage to this is that the beam computation at time *t* only involves data at time *t*, thus eliminating the need to store and access previously acquired data to form beams.

However, audio data of interest for situational awareness typically spans several octaves of bandwidth, and the time delay cannot be modeled as a single phase change. In this case, actual time delay and sum as outlined above may be the best method of implementation. This method requires that previously acquired data be stored and accessed to form beams, and so the process may be limited more by the ability to move data in and out of memory than algorithms based on the narrow band beamforming case.

In general, the required time delays will not correspond to exact multiples of the sampling rate and data will have to be accurately interpolated between samples. Often, the fastest way to do this interpolation is to analog low-pass filter the data at the highest frequency of interest, over-sample the data in the ADC process, and then interpolate between samples using nearest neighbor interpolation. To do this without degradation in performance, the data should be sampled at least eight times per cycle for the highest frequency of interest, or four times the Nyquist sampling criterion. The beams however, only need to be computed at the Nyquist sampling rate. For example, for beams valid up to 8 kHz, the data should be low-pass filtered at 8 kHz, sampled at 64 ksps, and then beams computed using nearest neighbor interpolation at 16 ksps

The number of additions that are required in the time delay and sum process are proportional to the number microphones, the number of beams, and the highest frequency of interest:

$$N_{adds} = 2 f_{\max} N_{sensors} N_{beams} .$$

Real-time Processing

For a large number of microphones (~100), real-time computation of multiple high-quality audio beams is beyond the capabilities of a high-performance PC. To solve this problem, we have developed an FPGA solution (cf. Maxfield, 2004) that can provide the real-time computation for multiple simultaneous beams in a relatively low-cost and mobile platform. Current capabilities can provide 16 simultaneous beams from 128 microphones for audio frequencies up to 8 kHz using a single 3M gate FPGA running at 40 MHz. This will improve significantly as more capable FPGA hardware is incorporated into our system.

Our FPGA processor interfaces to a PC laptop controller via a 50 MB/sec communications link, providing convenient interaction with the FPGA processor using high-level development tools. As a result, interactive beam-steering strategies can be implemented in high-level software to best use available beams in any of a variety of scenarios. One example is to use a cluster of beams to locate and lock onto an intermittent source of sound (e.g. a human voice) as illustrated in Figure 12. Since the beams are simultaneous, the same sound can be measured on all beams when it occurs and the angular location of the sound within the cluster of beams determined precisely. Software can then put a listening beam exactly on target, and adaptively keep it there as the array and target move. This would be impossible with a single scanned beam, since an intermittent sound would change during the scan time, and it would not be possible to determine the position of the source with accuracy.

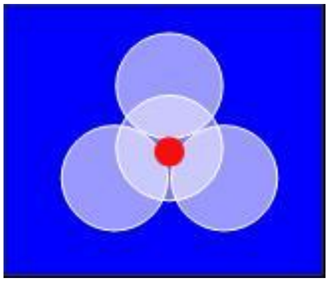

Figure 12. Location of intermittent sounds using a cluster of beams. Beam coverage in azimuth and elevation is indicated in white. Center beam is placed precisely on target (red) using information from the side beams.

A PROTOTYPE LOW-COST HIGH-GAIN MICROPHONE ARRAY SYSTEM

To demonstrate this microphone array technology, as well as the potential for high gain microphone arrays in general, we have developed and tested a 128-channel real-time microphone array system. The prototype system is an array of 128 independently digitized WM-64 microphones, which can be vehicle-mounted, as shown in Figure 13, for highly mobile operation, or tripod-mounted for fixed operation.

Figure 13. An array of 128 microphones mounted on a vehicle passenger door.

This system has FPGA processing and a Microsoft Corporation's Windows®-based controller that can form up to 16 simultaneous audio beams for frequencies up to 8 kHz in real time, and can form thousands of beams in near-real time to provide high-resolution acoustic images. Specific features of the prototype system are

- Broadband operation from 500 Hz to 8 kHz
- Wide field-of-view hearing capability, covering 180 degrees in both azimuth and elevation
- Custom analog pre-amp and signal-shaping circuits integrated with each microphone to minimize noise
- 128 digitally-controlled high-performance (8 th order) anti-alias filters with cutoff frequency adjustable in 10 kHz increments to minimize aliasing from ambient ultrasound
- 128 digitally-controlled second-stage pre-amps and spectral shaping filters with variable gain of 0-24 dB to accommodate a wide range of noise conditions
- Digital sampling: 128 synchronized channels at 62.5 ksps each
- 3M Gate Xilinx FPGA hardware processor in a ruggedized PXI bus chassis, which provides 16 simultaneous user-steerable digital-audio beams
- 50 MB/sec communications link from the FPGA PXI chassis to the host laptop computer
- Custom graphical user interface (GUI) software for operation from any Microsoft Corporation Windows XP® system
- Custom D-connectors, cabling, and tripod mounting for quick set up and takedown
- Removable windscreen to prevent noise from wind induced turbulence at the microphones

PROTOTYPE PERFORMANCE

The prototype described above has been field tested under a variety of conditions. Several relevant measures of performance are described in this section.

Array Gain

A good quantitative measure of array performance is the *array gain* as a function of frequency (cf. Van Trees, 2002). Array gain is a direct measure of how much the signal-to-noise ratio (SNR) is improved relative to an omni directional microphone in a specific acoustic environment. We have made array gain measurements outside in an open area where this is no significant multi-path effect, and where the noise is naturally coming from many sources in a typical urban environment. A point source of white noise in the far field, broadside to the array and at zero elevation was used as the signal. Array gain was measured relative to an omni-direction measurement microphone (Earthworks M50).

Figure 14 shows the measured array gain as a function of frequency. The measured gain (white) varies from 16 dB at 1.5 kHz to 22 dB at 5 kHz. Outside, where propagation loss is essentially geometric spreading in 2D, a 20 dB gain corresponds to a factor of 10 improvement in range, and a factor of 100 improvement in total area coverage.

Figure 14 also shows two predictions for this array. The upper prediction (green) is the directivity index, which would be the array gain if the ambient noise where truly isotropic in 2D, and the signal were indeed coherent across the array. The lower prediction (red) is for array gain in vertically impulsive noise (i.e. the noise arrives at only one angle in the vertical), with the signal arriving at the same angle as the noise. For impulsive noise, the problem is essentially one-dimensional, the vertical aperture is ineffective in reducing noise, and the array gain is substantially less than for the 2D isotropic noise case. As we see, the actual array gain is bounded by these two extremes, indicating that the noise is not truly vertically isotropic, but not vertically impulsive either.

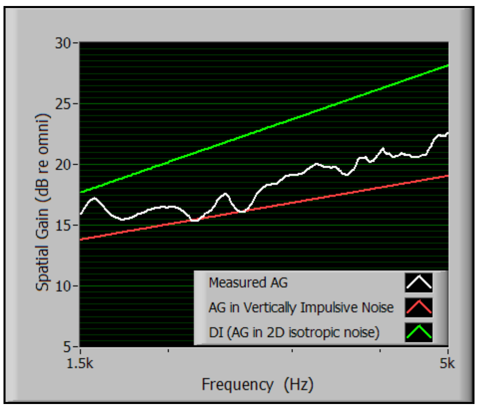

Figure 14. Measured array gain (AG, white) at zero elevation compared with AG predictions for 2D isotropic noise (green) and vertically impulsive noise (red).

It is important to note that if the noise were arriving near zero elevation, as seems to be the case for this array gain measurement, and the signal of interest were significantly above zero elevation, such as a sound from a nearby upper story window, then the array gain would be greater, possibly even greater than the upper prediction of Figure 14. Furthermore, array gain in other environments could be significantly different, either higher or lower. For example, in situations where there is substantial overhead noise, such as wind noise produced by overhead trees, the vertical aperture would be more effective, and the gain would be higher.

Beam Patterns

The array beam pattern (cf. Van Trees, 2002) is a measure of how well various discrete sources of sound can be separated. Figure 15 shows the 2D broadband beam response to a point source of white noise in the far field of the array at broadside (0 deg azimuth and 0 deg elevation). This response pattern is computed using the intensity of sound in each beam in the 2-6 kHz band. We see that the response is as much as 30 dB down, meaning that discrete sounds away from the target of interest can be reduced by as much as 30 dB.

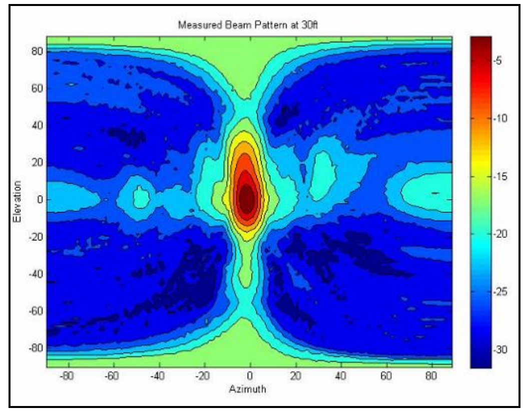

Figure 15. Array beam pattern response (dB) to a white noise point source of sound at zero elevation and azimuth. The sound can be attenuated by as much as 30 dB for beams away from the sound.

Audio Capabilities

Being able to monitor speech in situations where the speaker expects to have privacy is an excellent means of obtaining Humint (human intelligence) – note that this could be in violation of U.S. wiretapping laws and should only be done under appropriate circumstances. The key here is to have adequate standoff distance so that the speaker is not alerted to the possibility of monitoring.

As one example, we looked at the problem of monitoring a private cell phone conversation in an open suburban area where the speaker would not expect to be monitored. The speaker-to-sensor distance was 300 feet. Ambient conditions were quiet (0 dB SPL/Hz at 2 kHz). The cell phone user wanted not to be overheard and only spoke loudly enough for adequate communications. The experimental setup from the view of the sensor is shown in Figure 16. With the unaided ear, it is impossible to hear anything related to this activity, but with the array one can hear almost every word of the conversation. Since this paper does not allow for incorporation of sound files to illustrate this capability, we show instead a plot of the acoustic power versus time as received on the target beam for a brief part of the conversation (Figure 17).

Figure 16. The sound target (person in circle) standing in the open in a suburban environment as seen from the sensor at 300 feet standoff distance.

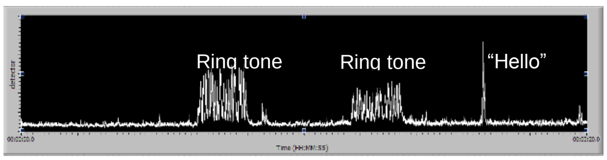

Figure 17. Acoustic power versus time for the beam pointed at the cell phone (see Figure 17) at 300 feet standoff.

Acoustic Imaging

Multi-beam acoustic arrays have the unique ability to make passive acoustic images, i.e., pictures of where the sound is coming from. This provides a new window on the acoustic environment that is unavailable with other hearing technology. Such imaging is useful for not only locating sounds of interest, but also for determining the directionality of ambient noise, which can then be used to better place microphone arrays to monitor specific areas.

Figure 18 shows an acoustic image of an urban setting in San Diego. The image shows ambient sound (noise) coming mainly from four discrete directions, with elevation angles of 0, 30, 60 and 80 degrees. By examining the accompanying picture of the location, we conclude that the point sources of ambient sound at high elevation are due to the sound from the wind moving the nearby palm trees, while the point source at zero elevation is due to sound coming down the street.

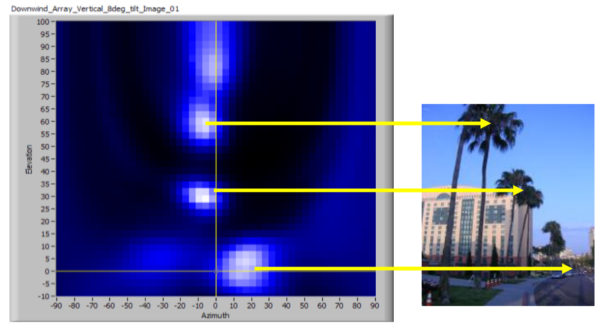

Figure 18. An acoustic image of sound coming from nearby trees and down the street

From this image, one can see that in this situation it is important to have sufficient vertical aperture to eliminate the wind/tree noise, and that the array should be placed so that targets of interest are not in-line with the street as seen from the array if at all possible.

CONCEPT OF OPERATIONS

The technology presented here could be used to improve situational awareness for military operations in a variety of modes. Arrays could be mounted on unmanned surface vessels, ground vehicles (GV), unmanned ground vehicles (UGV), and unattended ground sensors (UGS), as well as used in man-portable operations. A general operational overview is provided in Figure 19, and some specific concepts for a few possible modes of operation are described below.

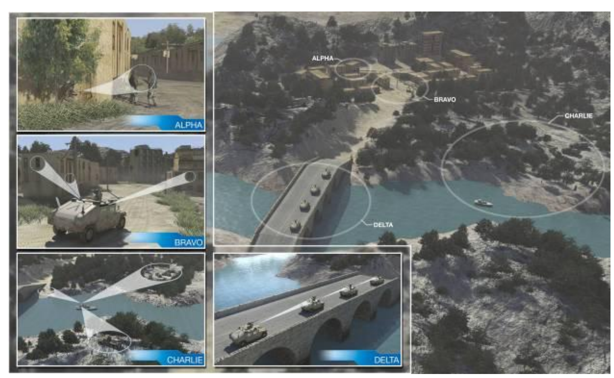

Figure 19. Operational overview for the use of acoustic arrays in the urban environment.

Cordon and Search

Buildings and surrounding areas need to be secured before they can be entered and searched. Multiple buildings could be monitored for subtle human activity by using a single vehicle-mounted array. An operator would set listening beams on windows, doors, air vents or other openings of opportunity. Anomalous sounds would be detected automatically and presented to the operator, and the operator could review the audio from these detections to better evaluate the situation.

Targeting

An acoustic beam from an *off-board* array could automatically slew with a weapon, allowing the operator to have both audio and visual queues about potential targets. When the weapon is fired, a beam could lock onto the target location for some time to monitor for damage assessment or other activities that may be of interest.

Physical Security

Vehicles, outposts, buildings, barracks, etc. are vulnerable to attack by close-in hostile personnel carrying explosives. Such an attack is difficult to detect with conventional sensors when the surrounding area is obscured by vegetation, structures, debris or other clutter. Listening beams could automatically scan the perimeter and provide an alert of any approaching traffic based on subtle nearby sounds.

Gunshot Detection

Current acoustic gunshot location devices, which use just a few microphones, are often inaccurate in noisy environments. A microphone array could augment or replace these sensors by eliminating the ambient noise for a better estimate of both the shock wave and the muzzle blast.

CONCLUSIONS

Microphone arrays can provide a revolutionary leap in the ability to exploit sound for improved situational awareness in a wide variety of applications. However, current COTS microphone array technology is not well suited for military operations. This paper presents core technology that allows development of highly capable real-time microphone array systems in a relatively low-cost and portable package better suited for military operations.

Microphone array performance is determined to a large extent by the physical size of the aperture and the number of sensors. For the prototype system described in this paper, with an aperture of about 1 m x ½ m and 128 sensors, we have been able to achieve 17-23 dB of array gain for audio frequencies in typical urban noise environments, and suppression of unwanted point sources of noise by up to 30 dB. This roughly allows a factor of 10 improvement in range, and a factor of 100 improvement in area coverage over a single omni-directional microphone. An example is shown were we can easily detect and adequately monitor a cell phone conversation at 300 feet in a quiet suburban environment.

A key feature for effective situational awareness in military operations is real-time capability. With a single FPGA, real-time computing of a large number of simultaneous listening beams is possible. The prototype presented here is able to compute 16 simultaneous beams at 8 kHz from 128 microphones. This is adequate to continuously track and monitor several targets simultaneously, or to quickly scan a large area.

This technology could provide enhanced hearing capability important for a wide variety of applications, including: urban combat, counterterrorism, intelligence, biometrics, border security, first responders, law enforcement, machinery health, and physical security.

REFERENCES

Brandstein, M., &Ward, D. (Eds.) (2001), Microphone Arrays, Signal Processing Techniques and Applications, Berlin Heidelberg New York, Springer –Verlag

Crystal Partners, Inc. Carol Stream, Illinois. *Big Ears* parabolic microphone at http://www.parabs.com/Broadcast/BroadcastBigEars.htm

Earthworks Precision Audio at http://www.earthworksaudio.com/3.html

G.R.A.S. Sound and Vibration, Holte, Denmark http://www.gras.dk/

Horiwitz, P., & Hill, W. (1989), The Art of Electronics, Second Edition, Chapter 4. Cambridge, U.K., Cambridge University Press

Maxfield, C. (2004), The Design Warriors Guide to FPGAs, Burlington, MA, Elsevier Newnes

Panasonic, Product specification sheet for WM-64PC at http://www.panasonic.com/industrial/components/pdf/ABA5000CE10.pdf

Van Trees, H. Optimal Array Processing, Chapter 2, New York, John Wiley & Son