Estimation of Aircraft Sound Direction of Arrival Using Directional-Microphone Arrays

Meritxell Genescà1, U. Peter Svensson1, Gunnar Taraldsen2

1) Acoustics Research Center, Department of Electronics and Telecommunications, Norwegian University of Science and Technology, 7491 Trondheim, Norway. meritxell.genesca@ntnu.no
2) Acoustics Research Center, SINTEF Information and Communication Technology, 7465 Trondheim, Norway

Summary

In the presence of ground reflections, the traditional methods for aircraft noise direction of arrival estimation based on the time differences between the microphones of a compact array lose accuracy in the vertical angle. This paper studies the benefits of using a tetrahedral array geometry of first order directional microphones instead of omnidirectional ones to reduce this error. It shows that the ground reflections introduce a systematic error in the time delay estimates, which is considerably reduced by using directional microphones, and that the selection of the optimum directivity pattern depends on the height of the array over the ground. The mean square estimation errors of the proposed approach are compared to the Cramer-Rao bound of those of a minimum variance estimator for a tetrahedral array of omnidirectional microphones in free field. Finally, computer simulations using real jet and propeller signals are used to assess the performance of the proposed method in the presence of microphone self noise and wind induced noise.

PACS no. 43.28.Tc, 43.60.Fg, 43.60.Jn

1. Introduction

This paper deals with the development of a compact sensor array capable of estimating the aircraft sound direction of arrival (DOA) in the surroundings of an airport considering ground reflections and wind induced and sensor self noise.

The existing methods to estimate the DOA by using an array of acoustic sensors can be classified as methods based on either the steered response power or the time difference of arrival (TDOA) estimation. The former category can be subdivided into spectral estimation methods and eigenvector based techniques. The first of these two calculates the power of the output of a beamformer as a function of the DOA and takes the argument of the highest peak(s) in the power output as the DOA of the source(s) [1]. One example is the Bartlett algorithm which uses the delay and sum beamformer requiring large microphone arrays to obtain good directivity when applied to aircraft sound DOA estimation [2, 3, 4]. Another example is the maximum-likelihood algorithm which uses the minimum variance distortionless beamformer [5] resulting in enhanced resolution but deteriorated performance in the presence of correlated sources (e.g. ground reflections). Previous studies have used the maximum-likelihood algorithm to track aircraft operations on the ground [6] and military vehicles [7]. There are also adaptations of the Bartlett and maximum-likelihood algorithms for arrays of vector sensors instead of pressure microphones aiming at reducing the array aperture in acoustic applications [8, 9].

The most popular eigenvector based techniques are the MUSIC [10] and the ESPRIT [11] algorithms which take benefit of the eigenstructure of the correlation matrix. Both algorithms are high resolution techniques and have been used in previous studies for ground vehicle DOA estimation [12, 13]. The performance of both algorithms decreases in the presence of correlated sources because such sources reduce the rank of the correlation matrix, but spatial smoothing techniques exist to overcome this problem [14].

The use of TDOA methods and compact arrays in aircraft sound DOA estimation is popular [15, 16, 17, 18]. They are attractive because of the simplicity of the signal processing algorithm, the low number of microphones required and the relatively small spatial extent of the final array. Wide aperture arrays require the differential Doppler to be taken into account [19, 20, 21] which increases the computation time. In free field conditions, the argument of the maximum of the generalized cross correlation function of the signals of two microphones equals the TDOA between them [22], which is related to the DOA. In the presence of ground reflections the correlation function contains multiple peaks and the highest may not represent that generated by the correlation of the direct sound. This mainly reverts into poor estimates of the elevation angle [23, 24]. Traditionally, this problem is handled using peak
picking algorithms to identify the peak related to the direct sound and disregard the rest. However, a recent study [25] shows that the flight parameters of an aircraft with a linear trajectory at constant altitude and speed can be more accurately estimated by also considering the other peaks. This paper presents a different approach which uses directional (instead of omnidirectional) microphones with the aim of reducing the amplitude of the reflected sound so that the problem can be treated as in free field conditions. This approach assumes no limitations or previous knowledge of the aircraft trajectory.

The rest of this paper is organized as follows: Section 2 describes the relationship between the TDOA and the DOA for the particular array geometry used here. Section 3 theoretically discusses the benefits of using directional microphones for TDOA estimation in the presence of ground reflections, section 4 describes the computer simulation procedure used to evaluate the performance of the method. Section 5 presents and discusses the results of the simulations and section 6 summarizes the main contributions of this work.

2. DOA estimation

Consider a pair of microphones \( m \) and \( n \) and a plane wave impinging on them as depicted in Figure 1. The sound DOA can be represented by the unitary vector \( r \) of Cartesian coordinates, or by the angular coordinates \( \theta \) (vertical angle) and \( \phi \) (horizontal angle). From Figure 1 it can be seen that the relationship between \( r \) and the extra time \( r_{mn} \) that the sound waves take to travel to a microphone \( n \) relative to the travel time to a microphone \( m \) is

\[
r_{mn} \cdot r = -r_{mn}c,
\]

where \( c \) is the speed of sound and \( r_{mn} \) is a vector from microphone \( m \) to \( n \). Note that \( r_{mn} = -r_{nm} \).

Assuming that the time delay \( r_{mn} \) is known, the unknowns in equation 1 are the components of the vector \( r \) (disregarding the non-linear restriction \( r = 1 \)). To obtain the unambiguous DOA, it is necessary to use at least 3 microphone pairs in a 3D configuration. A natural geometry has four microphones placed at the vertices of a regular tetrahedron as in Figure 2. This is the most symmetric distribution attainable with the minimum possible number of microphones and it has been used in previous studies [15, 16].

Due to atmospheric attenuation, most of the aircraft sound received at the ground – at distances greater than 1 km – is concentrated below 200 Hz. Moreover, below 20 Hz the acoustic signal may be severely corrupted by wind noise [26]. As a consequence, only the range within 20 Hz and 200 Hz is considered (with corresponding wavelengths in the range 1.7–17 m) for the considered application. Therefore, the distance \( a \) between the microphones is set to 0.85 m to guarantee compliance with the Nyquist-Shannon sampling theorem over this frequency range. This restriction is introduced to estimate \( r_{mn} \) correctly and may not be necessary in the case of jet noise but it can be relevant in the case of propeller noise because of the strong harmonic components [27].

For this geometry, the matrix form of equation 1 considering all the possible microphone pairs in Figure 2 is

\[
T = Xr^T,
\]

where

\[
T = \begin{bmatrix} r_{21} & r_{31} & r_{41} & r_{32} & r_{42} & r_{43} \end{bmatrix}^T,
\]

\[
X = \frac{1}{c} \begin{bmatrix} r_{21}^T & r_{31}^T & r_{41}^T & r_{32}^T & r_{42}^T & r_{43}^T \end{bmatrix}^T
\]

and \( r \) and \( r_{mn} \) are row vectors.

Therefore, from a set of TDOA estimates \( \hat{r}_{mn} \) it is possible to obtain an estimate of \( r \), denoted as \( \hat{r} \), as a least squares solution of equation 2 for instance.

Note that there is no guarantee that \( \hat{r} = [\hat{r}_x, \hat{r}_y, \hat{r}_z] \) is a unitary vector. Thus the DOA estimate in angular coordinates is calculated as

\[
\hat{\theta} = \cos^{-1} \left( \frac{\hat{r}_z}{\hat{r}} \right),
\]

\[
\hat{\phi} = \tan^{-1} \left( \frac{\hat{r}_x}{\hat{r}_y} \right).
\]

The vertical angle estimation error \( \epsilon_\theta \), can be defined directly as the difference between the estimated and the nominal value

\[
\epsilon_\theta = \hat{\theta} - \theta.
\]

On the contrary, the amount \( \hat{\phi} - \phi \) is not representative of the relevance of the horizontal error. For a nominal vertical angle \( \theta \) close to 0°, a difference \( \hat{\phi} - \phi \) as large as 180° has
little importance while for θ close to 90° such a difference means a really bad estimation. Therefore, the horizontal angle estimation error $ε_φ$ is defined as

$$ε_φ = (\hat{φ} - φ) \sin θ.$$  \hspace{1cm} (6)

3. TDOA estimation

The TDOA is estimated here as the argument of the maximum of the generalized correlation function. This section studies the effect of the ground reflections and the microphone directivity on this function in the case of a single source over a hard ground and assuming plane waves. Under these circumstances the two sided frequency spectrum received at the directional microphone $m$ is

$$P_m(ω) = P_m^d(ω)(D^d + D^e)e^{-jωτ_m^d},$$  \hspace{1cm} (7)

where $P_m^d(ω)$ is the two sided frequency spectrum of the direct sound component at the microphone $m$, $D^d$ and $D^e$ are the directivity gains associated with the direct and ground reflected sound DOA respectively, and $τ_m^d$ is the time delay between the direct sound and the reflected sound at microphone $m$.

$$τ_m^d = \frac{2H_m \cos θ}{c},$$  \hspace{1cm} (8)

where $H_m$ is the vertical distance between the microphone and the ground.

In the same way, the two sided frequency spectrum received at microphone $n$ becomes

$$P_n(ω) = P_n^d(ω)e^{-jωτ_m^e}(D^d + D^e)e^{-jωτ_m^e}. $$  \hspace{1cm} (9)

where the term $P_n^d(ω)e^{-jωτ_m^e}$ is again the two sided frequency spectrum of the direct sound component.

The generalized correlation function $R_{mn}(τ)$ between the microphones $m$ and $n$ is [22]

$$R_{mn}(τ) = \int_{-∞}^{∞} ψ_δ(ω) S_m(ω)e^{jωτ} dω,$$  \hspace{1cm} (10)

where $ψ_δ(ω)$ is a weighting function and $S_m(ω)$ is the cross spectrum

$$S_m(ω) = P_m(ω) P^*_m(ω).$$  \hspace{1cm} (11)

with $^*$ indicating complex-conjugate.

This study uses the phase transformation (PHAT) weighting function

$$ψ_δ(ω) = \frac{1}{||S_m(ω)||}$$  \hspace{1cm} (12)

to consider only the phase information of the cross spectrum in equation 10 so that the resulting function is independent of the magnitude. Any dependence on the magnitude of the cross spectrum is inconvenient here because the frequency spectrum varies with the position and speed of the aircraft due to the Doppler effect and the atmospheric attenuation, and also with the type of aircraft making the performance of the method quite variable. However, the PHAT generalization is sensitive to uncorrelated noise because all frequency bins have an equal weight regardless of the signal to noise ratio (SNR).

The combination of equations (7) and (9) into equation (10) leads to

$$R_{mn}(τ) = D^d D^d R_{mn}^d(τ + τ_m^d) + D^d D^e R_{mn}^e(τ + τ_m^d) + D^e D^e R_{mn}^e(τ - τ_m^d) + D^e D^d R_{mn}^d(τ - τ_m^d).$$  \hspace{1cm} (13)

where $R_{mn}^d(τ)$, which denotes the generalized cross correlation between the direct sound components in equations (7) and (9), is

$$R_{mn}^d(τ) = \int_{-∞}^{∞} ψ_δ(ω) \left| P_m^d(ω) \right|^{2} e^{jωτ} dω.$$  \hspace{1cm} (14)

The time delay $τ_m$ equals the argument of the maximum of $R_{mn}^d(τ)$, however in the present case, only $R_{mn}(τ)$ is available. The risks of estimating $τ_m$ as

$$τ_m = \arg \max_{τ} R_{mn}(τ),$$  \hspace{1cm} (15)

using omnidirectional microphones ($D^d = D^e = 1$) come up when realizing that equation 13 is the sum of four correlation functions. Thus $R_{mn}(τ)$ can have up to four peaks, and there is no guarantee that the peak associated with the first term in the right side of equation 13 is the maximum. Even when using peak picking algorithms to identify this peak, the argument of its maximum may not coincide with that of $R_{mn}^d(τ)$ due to the influence of the other terms in equation 13.

These problems are demonstrated in Figure 3a which shows the correlation function $R_{41}(τ)$ between two omnidirectional microphones at positions 1 and 4 in Figure 2 and the different component terms in equation (13). The array is placed at a height of 2.7 m over a hard ground. The aircraft signal is modeled as a band pass random noise signal with a flat spectrum between 20 and 200 Hz and impinges on the array with $θ = 60°$ and $φ = 0°$.

Considering now the use of directional microphones, the directivity gains in equation (13) act as weighting factors for each term. A good choice of microphone directivity would attenuate the terms other than the first making the peak associated with the first term in equation (13) be the highest and the argument of its maximum be closer to that of $R_{mn}^d(τ)$. As a consequence, the estimation of $τ_m$ obtained is more accurate than when using omnidirectional microphones and a peak picking technique.

Figure 3b illustrates the advantages of using directional microphones. It shows the same case as in Figure 3a but using cardioid microphones with the axes of the microphones vertically oriented.

4. Simulation procedure

For a set of nominal DOA, described in section 4.1, the signals received at the microphones of the array are simulated as explained in section 4.2. The method presented in sections 2 and 3 is implemented as described in section 4.3 and applied to these signals.
coordinates; \( N_m(\omega) \) is the discrete two sided frequency spectra of the additional noise; \( r_m^2 \) is the time difference between the direct sound and the ground reflection of the aircraft signal at microphone \( m \) calculated as in equation 8 with \( H_m = h + r_m \), where \( h \) is the vertical distance from the ground to the base of the tetrahedron (set by default to 2.7 m); and \( r_m \) is the extra time that the direct signal takes to reach microphone \( m \) relative to the time it takes to reach the origin of coordinates and is calculated using equation (1) with \( r = \sqrt{\sin \theta \sin \phi \sin \theta \cos \phi \cos \theta} \).

The simulations in this study consider only microphones with first order directivity patterns, thus

\[
D^d = (1 - b) + b \cos \theta, \\
D^f = (1 - b) + b \cos(180^\circ - \theta).
\]

(17)

where \( b \in [0 1] \) is a directivity parameter that controls the directivity pattern.

The way that \( P_a(\omega) \) and \( N_m(\omega) \) are modeled depends on the goal of the simulation. In section 5.1, \( P_a(\omega) \) is white gaussian noise generated, for each single frame, as the Fourier transform of a 0.5 second long series of pseudorandom values drawn from a zero mean normal distribution, and \( N_m(\omega) = 0 \) to achieve an infinite SNR. In section 5.2, both \( N_m(\omega) \) and \( P_a(\omega) \) are modeled as white gaussian noise generated as before, and \( N_\text{omni}(\omega) \) is uncorrelated from microphone to microphone. In section 5.3, for successive frames, \( P_a(\omega) \) equals the frequency spectrum of non overlapping successsive signal blocks of 0.5 s of two real measurements of jet and propeller sound, and \( N_m(\omega) \) is alternatively white (modeled as before), wind induced or microphone self noise. For each frame, the wind induced noise has been modeled as partially correlated noise [28] between the microphones following the Sandia method [29] assuming that the power density spectrum follows a \( \omega^{-11/3} \) law [26, 28] and that the coherence function varies as \( e^{-2\pi r_m^2 \cos(2\pi r_m \sin \theta) / RM_m^2} \), where \( r_m^2 \) is the distance between the microphones in wavelengths in the downwind direction and \( RM_m^2 \) in the crosswind or vertical direction [30]. For each frame, the directional microphone self noise is obtained by combining the noise of two omnidirectional microphones as [31]

\[
N_{\text{omni}}^\text{self}(\omega) = (1 - b) \frac{N_a^\text{omni} + N_b^\text{omni}}{2} + b \frac{N_a^\text{omni} - N_b^\text{omni}}{\text{d/c}}.
\]

(18)

where \( d \) is the distance between the omnidirectional microphone pair, which has to be much smaller than the smallest wavelength (here \( d=0.3 \) m), and \( N_a^\text{omni} \) and \( N_b^\text{omni} \) are two independent noise realizations generated as

\[
N_{\text{omni}}^i(\omega) = N^{\text{white}}(\omega) \cdot \frac{1}{\alpha^{n+1}}, \\
\alpha = 1.8592 \quad \text{because between 20 and 200 Hz the pre-amplifier noise, which is the main component, decreases by approximately 0.86 dB per third of octave [32].}
\]

4.1. Sampling mesh

A set of 3550 points (\( \theta, \phi \)) have been distributed on one sixth of an upper unitary hemisphere (\( \phi \in [0^\circ 60^\circ] \) and \( \theta \in [0^\circ 90^\circ] \)) centered on the coordinate origin. Due to the inherent symmetry of a regular tetrahedron, one sixth of the upper hemisphere is representative of the whole upper hemisphere. The angle \( \theta \) has been discretized at intervals of \( 1^\circ \) and \( \phi \) at intervals of \( 1^\circ \) when \( \theta = 90^\circ \). For any other value of \( \theta \), the number of nominal values of \( \phi \) considered is proportional to the length of the segment of the parallel enclosed within the one sixth hemisphere.

4.2. Signal model

For each nominal direction of the source, defined by \( \theta \) and \( \phi \), and for each microphone \( m \), the discrete two sided frequency spectrums of the sound pressure corresponding to a 0.5 second signal frame is assembled in the frequency domain as the sum of the following terms

\[
P_m(\omega) = P_a(\omega)e^{-j\omega r_m^2} \left( D^d + D^f e^{-j\omega r_m^2} \right) + N_m(\omega),
\]

(16)

where \( P_a(\omega) \) represents the discrete two sided frequency spectrums of the direct noise of the aircraft at the origin of
In all cases, the sampling frequency of the simulated signals is set to \( f_s = 5 \text{kHz} \), which avoids aliasing in the frequency range of interest.

Note that the effect of the atmospheric turbulence along the propagation path from the aircraft to the microphone array is not included in the signal model described in equation (16). This is because the coherence loss between the microphones of the array due to atmospheric turbulence is expected to be small for the particular geometry used here. Such a conclusion has been drawn from the coherence plots in [33]. Assuming a value of 1 m for the scale of turbulence, and a variance of sound speed fluctuations of \( 10^{-6} \) [34], and knowing that the maximum transversal and longitudinal distance between the microphones is 0.85 m, [33] shows that the total coherence loss over the array aperture is dominated by the transversal coherence loss. Then, for a worst case scenario with a maximum propagation distance from the aircraft to the array of 5 km and assuming a Gaussian turbulence spectrum, the top curve in Figure 2 in [33] applies, showing that the expected coherence between any two microphone signals is higher than 0.9.

4.3. Implementation of the method

An estimate of the aircraft sound DOA is calculated for each successive non-overlapped 0.5 second signal frame. For all the possible microphone pairs, the weighted cross spectra are calculated as

\[
\psi_{\tilde{g}}(\omega) \tilde{S}_{\text{mn}}(\omega) = \begin{cases} 
P_m(\omega)^* P_n(\omega) & \text{if } 40 \pi \text{ rad/s} \leq |\omega| \leq 400\pi \text{ rad/s}, \\
0 & \text{otherwise.}
\end{cases}
\]

Note that the use of a rectangular filter in equation (20) guarantees that the PHAT generalization depends only on the unweighted phase of the cross spectrum within 20 and 200 Hz.

The PHAT generalization is estimated as

\[
\hat{R}_{\text{mn}}(\tau) = \text{Re} \left( \text{ifft}(\psi_{\tilde{g}}(\omega) \tilde{S}_{\text{mn}}(\omega)) \right),
\]

where ifft stands for inverse Fourier transform, and \( \tau \) are now discrete values of time delay.

The time difference estimates \( \hat{\tau}_{\text{mn}} \) result from

\[
\hat{\tau}_{\text{mn}} = \arg \max \_{\tau \in [-\frac{\alpha}{\sin \epsilon}, \frac{\alpha}{\sin \epsilon}]} \hat{R}_{\text{mn}}(\tau),
\]

where \( \alpha \) is the angle between \( r_{\text{mn}} \) and the positive semi-axis \( z \). The search boundaries of \( \tau \) derive from equation 1 and from the fact that in this particular application the target source vertical angle lies within \( \theta \in [0^\circ \ 90^\circ] \). These boundaries make it possible to automatically discard, in certain cases, peaks caused by the ground reflections.

The precision of \( \hat{\tau}_{\text{mn}} \) is initially given by the selected sampling frequency, but it can be increased by interpolating \( \hat{R}_{\text{mn}}(\tau) \). Here, the applied cubic spline interpolation leads to a precision of \( 2 \cdot 10^{-6} \). This fine discretization introduces numerical errors \( \epsilon_\theta \) and \( \epsilon_\phi \) lower than 0.039°.

By substituting \( T \) for \( \hat{T} = [\hat{t}_{21} \ \hat{t}_{31} \ \hat{t}_{32} \ \hat{t}_{42} \ \hat{t}_{43}]^T \) in equation (2), an estimate \( \hat{r} \) of the direction vector is obtained. Finally the estimates of the vertical \( \hat{\theta} \) and horizontal \( \hat{\phi} \) angles are calculated from equations (3) and (4) and the associate estimation errors \( \epsilon_\theta \) and \( \epsilon_\phi \) from equations (5) and (6).

5. Simulation results

Section 5.1 studies the influence of the ground reflection as a function of the directivity of the microphones considering an infinite SNR, section 5.2 compares the performance of the present method to that of a minimum variance unbiased DOA estimator (MVUE), and section 5.3 studies the influence of the noise spectrum and the SNR on the method’s performance.

Note that the signal model presented in equation (16) corresponds to a hard ground situation as a worst case scenario so that, in any other conditions, the method is expected to perform at least with the accuracy predicted in here.

5.1. Influence of the ground reflection and the directivity of the microphones

In this section an infinite SNR is considered, under these circumstances the PHAT generalization does not depend on the aircraft signal, therefore the conclusions drawn here apply to any aircraft signal type.

For the case of using supercardioid microphones (\( b = 0.65 \)), Figure 4 shows \( \epsilon_\theta \) (a) and \( \epsilon_\phi \) (b) as a function of the sound nominal direction \([\theta, \phi]\) plotted on a sixth of the hemisphere seen from above. The results reveal that \( \epsilon_\theta \) basically only depends on \( \theta \) while \( \epsilon_\phi \) is very small, and depends on \( \theta \) and \( \phi \). As a consequence, for the rest of the paper all results are averaged over \( \phi \).

Figure 5 illustrates the advantages of using microphones of appropriate directivity instead of omnidirectional ones. It shows \( \epsilon_\theta \) (a) and \( \epsilon_\phi \) (b) when using omnidirectional (\( b = 0 \)), subcardioid (\( b = 0.3 \)), cardioid (\( b = 0.5 \)), supercardioid (\( b = 0.65 \)) and hypercardioid (\( b = 0.75 \)) microphones.

Figure 5b reveals that \( \epsilon_\phi \) is very small for all directivity patterns except for the case of omnidirectional microphones. The error peaks that appear in this case are a consequence of the introduction of the search domain limits in equation 22. When the search domain is unbounded the TDOA estimation error \( \hat{\tau}_{\text{mn}} - \tau_{\text{mn}} \) caused by the ground reflections is the same for \( mn = 41, 42, 43 \). This guarantees that the horizontal coordinates of \( \hat{r} \) obtained as a least squares solution of equation 2 are error free. The search domain limits may reduce the TDOA estimation error in some of the terms \( \hat{\tau}_{\text{mn}} \) with \( mn = 41, 42, 43 \) but not necessarily in all of them introducing then error in the horizontal coordinates of \( \hat{r} \). However, this phenomena is mainly observed only in the case of omnidirectional microphones, and in general the reduction introduced by the
search domain limits on the estimation of the vertical angle is greater than the error increase perceived on the horizontal angle. The curves in Figure 5b also show a small ripple as a function of $\theta$ which is in agreement with the small variations of $\varepsilon_{\phi}$ with the DOA shown in Figure 4b. These variations are due to the geometry and intrinsic symmetries of the tetrahedral array.

The curves in Figure 5a show a relevant reduction of $\varepsilon_{\theta}$ for directional microphones compared to omnidirectional ones. A given directivity pattern provides the lowest error of all directivity patterns around the value of $\theta$ for which the gain in the direction $180^\circ - \theta$ is zero. This happens at $\theta = 0^\circ$ for $b = 0.5$, at $\theta = 55^\circ$ for $b = 0.65$ and at $\theta = 71^\circ$ for $b = 0.75$. For $b = 0.3$ this gain is never zero. It follows that prior knowledge on the expected vertical angles of the aircraft allows for an optimal choice of directivity pattern.

Apart from these there are other minima in the curves. This peculiar bias pattern is caused by the existence of a systematic error on the time delay estimates of the non-horizontal microphone pairs. This error is due to the interference of the second, third and fourth terms the right side of in equation 13 with the first term. To support this statement Figure 6 shows, for $b = 0.65$, both $\varepsilon_{\theta}$ and the function arg max $R_{\text{dir}} - \tau_{\text{dir}}$ which represents the systematic component of the error $\delta\tau_{\text{dir}}$. Note that, for a given $\theta$, the relative time delay between the terms in equation 13 depends only on the height of the two microphones (see equation 8), therefore arg max $R_{\text{dir}} - \tau_{\text{dir}}$ is the same for $m=1,2$ or 3.

The pattern shown in Figures 5a and 6 changes with the height over the ground of the array because the height influences the time delay of the second and third terms in equation (13) relative to the first. Figure 7 shows (in black) the optimal directivity parameter, $b_{\text{optimal}}$, defined here as the value of $b$ for which $\varepsilon_{\theta} \leq 1^\circ$ for a largest continuous range of nominal vertical angles $\theta$ starting at
\( \theta = \theta_0 \). It also shows in the form of gray dots which are the nominal \( \theta \) with \( \epsilon_\theta \leq 1^\circ \) for the correspondent optimal directivity parameter and height over the ground.

\( \theta \) does depend both on \( \phi \) and \( \epsilon_\theta \), depends both on \( \theta \) and \( \phi \). For a more quantitative comparison, Figure 8 shows the root mean square estimation errors of the MVUE and those obtained with the current TDOA method considering \( P_s(\omega) \) and \( N_\omega(\omega) \) as white gaussian noise with different SNR ranging from 0 to 40 dB, and both the case of free field (thin lines) and ground reflections (thick lines) with \( b = 0.65 \) and \( h = 2.7 \text{ m} \). The expression of the root mean square estimation error of the current method calculated over \( Q = 500 \) signal frames for a given SNR is

\[
\langle \text{RMSE}(\epsilon) \rangle \phi = \left\langle \sqrt{\frac{\sum_{N=1}^{Q} \epsilon(\theta, \phi, SNR)^2}{Q}} \right\rangle \phi. \tag{24}
\]

where \( \langle \rangle \phi \) means average over \( \phi \).

Figure 8 shows that the performance of the current method clearly differs from that of a MVUE in at least one order of magnitude. In the case of \( \epsilon_\theta \), the systematic component, or bias of the error, introduced by the reflections dominates at high SNR while the random component dominates at low SNR conditions. For a SNR of 0 dB or lower the behavior is similar to that of the free field case. In the case of \( \epsilon_\phi \), the behavior is similar to the free field case except for angles around 80° confirming that the influence of ground reflections is minimal.
5.3. Influence of the noise spectrum and the SNR

This section delves into the performance of the method in the presence of noise by simulating a more realistic situation in which $P_s(\omega)$ is real jet or propeller aircraft sound, alternatively, and $N_m(\omega)$ is white, wind induced or microphone system self noise.

In this section, as opposed to the case studied in section 5.2, the SNR may experience great variations as a function of the frequency due to the non flatness of the frequency spectra of both noise and signal. To illustrate this, Figure 9 shows the normalized power spectral density function (PSD) of a sample of both jet and propeller sound, and also for a realization of white, wind induced and microphone self noise.

Figure 10 shows the root mean square error of the DOA estimation method for each of the six possible combinations of aircraft sound and noise with overall (between 20 and 200 Hz) SNR ranging from 0 to 40 dB for $b = 0.65$ and $h = 2.7$ m. The expression used to calculate the root mean square error is

$$\langle \text{RMSE}(\epsilon) \rangle_{\phi, \theta} = \sqrt{\frac{1}{Q} \sum_{q=1}^{Q} \epsilon(\theta, \phi, \text{SNR})^2}_{\phi, \theta},$$

with $\theta \in [0^\circ, 67^\circ]$, where $\langle \rangle_{\phi, \theta}$ means average over the nominal directions and it is restricted to $\theta \in [0^\circ, 67^\circ]$ to account only for the vertical angles with $\epsilon_\theta \leq 1^\circ$ when the SNR is infinite. $Q$ is 176 and 140 for the jet and the propeller case respectively, because the signals used to generate the aircraft sound are 88 and 70 seconds long, and the DOA is estimated for each non overlapping frame of 0.5 s of signal.

Figure 10 shows, first, that the errors due to microphone self noise are much smaller than those due to white or wind induced noise. Comparing white and self noise the cause can be that in the case of white noise the SNR is constant over the whole frequency range, while in the case of self noise it may be low at low frequencies but increases rapidly for most parts of the spectrum. This is also applicable when comparing wind induced and white noise, however it does not result in lower errors for wind induced noise probably due to the fact that wind induced noise is correlated between the microphones.

Second, it can be observed that the noise influence is lower for jet noise than for propeller noise. A possible cause can be that the SNR may be quite low for all the frequencies other than the tonal components of the propeller noise.

Third, the results show that, regarding microphone self noise, the overall SNR only needs to be higher than 15 dB to obtain root mean square estimation errors lower than 1° for both $\theta$ and $\phi$ in the range $\theta \in [0^\circ, 67^\circ]$. There are high quality omnidirectional measurement microphone systems whose self noise can be as low as 5 dB in the range between 20 to 200 Hz [32], therefore it seems reasonable to think that it is possible to obtain directional microphones with self noise levels low enough by combination of two of the former, or by a combination of one of the recently developed particle velocity sensors [35] and a pressure sensor element.

Finally, Regarding wind induced noise, a high SNR is critical for the performance of the method. According to the formula introduced in [26], a wind speed of 1 m/s (light air) induces a sound pressure level around 27.8 dB on a microphone with a 10 cm diameter wind screen, consider-
ing only frequencies from 20 to 200 Hz. In this case the SNR should be high enough for the method to perform in an acceptable way. However, a wind speed of 2 m/s (light breeze) induces a sound pressure level around 47.6 dB which will hinder the performance of the method in many cases. Given these values, and the fact that for the same SNR Figure 10 shows that the influence of wind induced noise is higher than that of the microphone self noise, it could be assumed that the wind induced noise level will typically be the limiting one.

6. Conclusions

This paper explores the idea of using a tetrahedral array of directional microphones to estimate the DOA of aircraft sound in the presence of ground reflections, and shows that the use of directional microphones noticeably reduces the DOA estimation error with respect to the use of omnidirectional microphones. It is shown that this improvement is due to the fact that the use of an appropriate first order directivity reduces the amplitude of the components of the generalized cross correlation function related to the reflected sound so that their influence on the component related to the direct sound is smaller. As a consequence the direct sound TDOA estimate, obtained as the argument of the maximum of this function, is estimated with significantly lower error. It is also proven that in the presence of ground reflection the TDOA estimates for a microphone pair which is not parallel to the ground suffer from systematic errors even when directional microphones are used.

Computer simulations show that, for example, for an array placed 2.7 m above the ground, in absence of uncorrelated noise, the range of nominal $\theta$ with estimation error lower than 1° can be extended from $\theta \in [0^\circ, 10^\circ]$ when using omnidirectional microphones to $\theta \in [0^\circ, 67^\circ]$ when using first order microphones with $b = 0.65$. However, the directivity parameter that maximizes the range changes with the array height over the ground. Nevertheless, this optimal directivity parameter always tends to be in the supercardioid range for array heights of 2–7 meters.

It is also shown that the vertical and horizontal angle estimation errors of the presented method are at least one order of magnitude higher than the Cramer-Rao bound of those of a minimum variance estimator in free field. In the case of $\epsilon_\theta$, the systematic error component introduced by the reflections dominates at high SNR, while at low SNR it is the random component introduced by the uncorrelated noise that dominates. In the case of $\epsilon_\phi$ the influence of the ground reflections is negligible. Moreover, it is shown that, for the presented method $\epsilon_\phi$ depends on both the nominal $\theta$ and $\phi$, and $\epsilon_\theta$ depends only on $\theta$, while in the case of a MVUE they are independent from the nominal direction.

Finally, computer simulations considering real jet and propeller aircraft signal show that overall SNRs higher than 30 dB and 15 dB are required to obtain root-mean square-errors lower than 1° for the range $\theta \in [0^\circ, 67^\circ]$ in the presence of wind induced and microphone self noise respectively. Therefore it is concluded that the system performance should be acceptable for wind conditions under 1 m/s (light air) and that it is reasonable to expect that a directional microphone self noise could be low enough to not limit the performance of the method.

Although the case when other ground borne sources are present is not discussed here, the use of directional microphones may also be useful to reduce their influence.

Appendix

Cramer Rao Lower Bound

The following Cramer Rao Lower Bound (CRLB) calculation considers estimation in an Euclidean space, however a Riemannian manifold [36] would be a better approach since DOA estimation is equivalent to the estimation of position on a sphere. Nevertheless, assuming small estimation errors the procedure chosen here is still reasonable.

The signal model considered to calculate the CRLB assumes a single source in free field conditions. Therefore, the two sided frequency spectrum of the signal $p_n(t)$ received at microphone $n$ is

$$P_n(\omega) = P_0(\omega)e^{-j\omega t_n} + N_n(\omega).$$

(A1)

In the case that the signal and noise are ideal band-pass filtered white noise processes, the autospectrum of the signal at the origin is

$$E[P_0(\omega)^*P_n(\omega)] = \begin{cases} S_0 \text{ if } \omega_1 \leq |\omega| \leq \omega_2, \\ 0 \text{ otherwise,} \end{cases}$$

(A2)

and the autospectrum of the uncorrelated noise is

$$E[N_n(\omega)^*N_m(\omega)] = \begin{cases} M \text{ if } \omega_1 \leq |\omega| \leq \omega_2, \\ 0 \text{ otherwise,} \end{cases}$$

(A3)

where $S_0$ and $M$ are constants. The CRLB is given by the matrix $F$ defined as $F = J^{-1}$, where $J$ is the Fisher matrix. For a zero mean-signal, the elements of $J$ can be calculated as [37]

$$J_{uv} = \frac{T}{4\pi} \int_{-\omega_2}^{u \omega} \text{tr} \left\{ \frac{\partial S(\omega)}{\partial q_u} \frac{\partial S(\omega)^{-1}}{\partial q_v} S(\omega)^{-1} \right\} d\omega,$$

(A4)

with $u, v = 1, 2$, where $q = [\theta, \phi]$ is the parameter vector to estimate, tr denotes the trace of the matrix, $T$ is the time length of the signals $p_n(t)$, and the elements of the spectral matrix $S$ are

$$S_{uv} = \begin{cases} S_0 e^{-j\pi r_n(r_m-r_n)} \text{ if } \omega_1 \leq |\omega| \leq \omega_2, \\ S_0 + M \text{ if } |\omega| \leq \omega_2, \\ 0 \text{ otherwise,} \end{cases}$$

(A5)

The resulting expressions of the elements of the inverse of the Fisher matrix for the case of a regular tetrahedron geometry are

$$F_{11} = -\frac{3\pi e^2}{2T \sqrt{\frac{3}{2}}} \left(1 + \frac{2\omega_1^3}{\omega_2^3} \right),$$

$$F_{12} = F_{21} = 0,$$

$$F_{22} = -\frac{3\pi e^2}{2T \sqrt{\frac{3}{2}}} \sin(\theta)^2 \left(\omega_1^3 - \omega_2^3\right),$$

(A6)
According to the error definitions given in equation 5 and 6, the root mean square errors of a MVUE are RMSE($\varepsilon_\theta$) = $\sqrt{F_1}$ and RMSE($\varepsilon_\phi$) = $\sqrt{F_2}$sin$^2\theta$. The fact that $F_1 = 0$ indicates that $\varepsilon_\theta$ and $\varepsilon_\phi$ are uncorrelated.

Acknowledgments

The work presented here is part of the project SOUNDTRACK funded by the European Union’s Seventh Framework Programme (FP7/2007–2013) under grant agreement no. 299620.

References


