Application of the short-time Fourier transform and the Wigner–Ville distribution to the acoustic localization of aircraft

Brian G. Ferguson and Barry G. Quinn
Maritime Operations Division, Defence Science and Technology Organisation, P. O. Box 44, Pyrmont, NSW 2009, Australia

(Received 12 July 1993; accepted for publication 29 January 1994)

The dominant feature in the acoustic spectrum of a propeller-driven aircraft is the spectral line corresponding to the propeller blade rate that is equal to the product of the propeller shaft rotation rate and the number of blades on the propeller. The frequency of this line, when measured by a stationary observer on the ground, changes with time due to the acoustical Doppler effect. In this paper, the short-time Fourier transform and the Wigner–Ville distribution are used to estimate the propeller blade rate at short time intervals for a turbo-prop aircraft flying at a constant altitude and speed over an acoustic sensor located just above ground level. The temporal variation in the observed blade rate is then used to estimate the speed and altitude of the aircraft, together with the source (or rest) frequency of the blade rate. Finally, the estimated values for these parameters are compared with the actual values recorded onboard the aircraft during each of the eighteen transits formed by pairing each element of a set of speeds: 150, 200, and 250 kn, with each element of a set of aircraft altitudes: 250, 500, 750, 1000, 1250, and 1500 ft.

PACS numbers: 43.60.Gk, 43.50.Lj

INTRODUCTION

The acoustic spectrum of a transitting aircraft, when received by a stationary observer, changes with time due to the acoustical Doppler effect. Although standard Fourier analysis enables a signal to be decomposed into its individual frequency components and establishes the relative intensity of each component, it does not provide any information on when these frequency components occurred. The traditional method for analyzing signals whose frequency content changes with time is the short-time Fourier transform, which selects only a short segment of the signal (or window of data) for spectral analysis at any one time. This method relies on the frequency content of the signal being stationary during the time interval corresponding to the window length. If this is not the case, then the frequency spectrum becomes smeared; the energy of a spectral line whose frequency is rapidly varying with time, for instance, will be spread over a number of adjacent frequency bins with no indication of how the spectrum evolved during the window interval.

I. SHORT-TIME FOURIER TRANSFORM

A. Continuous time signal

The magnitude squared of the Fourier transform is the classical method used to represent the frequency domain information, or spectrum, of a stationary signal. For a continuous time signal $x(t)$, the Fourier transform is defined as

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft} dt.$$  (1)

However, if the signal is nonstationary, the variation with time of the signal’s frequency content will be obscured because the frequency domain information is defined in (1) to be over an infinite period of time. The short-time Fourier transform approaches the problem of determining when a particular frequency occurs by partitioning the nonstationary signal into short time segments and then applying a weighting function to the signal within each segment, prior to evaluating the Fourier transform. The short-time Fourier transform is given by

$$X_s(t,f) = \int_{-\infty}^{\infty} \omega(t-\tau)x(\tau)e^{-j2\pi ft} d\tau.$$  (2)

where the analysis window $\omega(\tau)$ is centered on the time $t$. The window function satisfies $\omega(\tau) = 0$ for $|\tau|>T/2$, where $T$ is the duration of the time segment.
The power spectrum of the modified signal \( \omega(x(t)) \) is given by

\[ \rho_x(t,f) = |X_x(t,f)|^2. \]

For each time segment of the signal there is a corresponding spectrum and the totality of these spectra gives a time-frequency distribution known as a spectrogram, which is equivalent to a lofargram in sonar applications.

B. Discrete time signal

Let \( x(n) \) represent a discrete-time signal defined for all \( n \), then the short-time Fourier transform at time \( n \) and frequency \( f \) is defined as

\[ X_s(n,f) = \sum_{m=-\infty}^{\infty} \omega(n-m)x(m)e^{-j2\pi fm}, \]

where \( \omega(n) \) is the analysis window that selectively determines the portion of \( x(n) \) for analysis. (The convention of normalizing the sampling period, \( \Delta t \), to unity is adopted in this paper.) The short-time Fourier transform is a function of the time index \( n \) and represents only the local behavior of \( x(n) \) as selected by the analysis window \( \omega(n-m) \).

If \( X_s(n,f) \) is considered to be a function of \( n \) for a fixed \( f \), then \( X_s(n,f) \) can be expressed as the convolution of the signal \( x(n)e^{-j2\pi fn} \) with the impulse response \( \omega(n) \), that is,

\[ X_s(n,f) = \left[ x(n)e^{-j2\pi fn} \right] * \omega(n), \]

where \( * \) denotes the convolution operator with respect to \( n \), and \( \omega(n) \) is a low-pass filter being applied to the signal \( x(n)e^{-j2\pi fn} \). The modulation of \( x(n) \) by \( e^{-j2\pi fn} \) shifts the frequency spectrum of \( x(n) \) at frequency \( f \) to zero frequency. Thus the short-time Fourier transform can be interpreted as the output of a low-pass filter that is applied to the signal after its spectrum has been translated from a frequency \( f \) to zero frequency.

II. WIGNER–VILLE DISTRIBUTION

A. Continuous time signal

The Wigner–Ville distribution is given by

\[ W(t,f) = \int_{-\infty}^{\infty} z(t+\tau/2)z^*(t-\tau/2)e^{-j2\pi f\tau} d\tau, \]

where \( z(t) \) is the analytic signal associated with the real signal \( x(t) \).

In the time domain, the analytic signal is given by

\[ z(t) = x(t) + jH[x(t)], \]

where \( H[x(t)] \) is the Hilbert transform of the real signal \( x(t) \), which is defined by

\[ H[x(t)] = \mathcal{A} \left( \int_{-\infty}^{\infty} \frac{s(u)}{\pi(t-u)} du \right), \]

where \( \mathcal{A} \) denotes the Cauchy principal value of the integral.

In the frequency domain, the analytic signal is given by

\[ Z(f) = \begin{cases} 2X(f), & \text{if } f>0, \\ X(f), & \text{if } f=0, \\ 0, & \text{if } f<0. \end{cases} \]

Thus the analytic signal \( z(t) \) can be readily evaluated by calculating the Fourier transform \( X(f) \) of the real signal \( x(t) \), then in accordance with (9), setting the negative frequency spectral components to zero before calculating the inverse Fourier transform. The factor of 2 in (9) compensates for halving the total energy of the signal by elimination of that half of the spectrum associated with the negative frequency components of the signal.

B. Discrete time signal

The Wigner–Ville distribution for a discrete time signal is given by

\[ W(n,f) = \sum_{m=-\infty}^{\infty} z(n+m)z^*(n-m)e^{-j4\pi fn/m}, \]

where the sampling period has been normalized to unity.

The windowed version of the discrete time Wigner–Ville distribution is given by

\[ W_w(n,f) = \sum_{m=-M}^{M} w(m)w^*(-m)z(n+m) \]

\[ \times z^*(n-m)e^{-j4\pi fm}, \]

where \( w(m) = 0 \) for \( |m| > M \). Thus the window function is centered on the time index \( n \) and the duration of the window is equal to \( 2M+1 \) sampling periods.

Note that the periodicity in the frequency variable of the Wigner–Ville distribution is half that of the Fourier transform. Hence, even when the sampling of \( x(t) \) satisfies the Nyquist criterion \( (f_s > 2f_{\text{max}} \) where \( f_{\text{max}} \) is the maximum frequency of interest and \( f_s \) is the sampling rate), there will be aliasing components in the time-frequency distribution, if the real signal \( x(n) \), rather than the analytic signal \( z(n) \), is used in (11). Use of the analytic signal eliminates the need to sample at twice the Nyquist rate.

If the real signal is used, the resulting time-frequency distribution is commonly referred to as the Wigner distribution, and to avoid aliasing, either \( x(t) \) is sampled at double the Nyquist rate \( (f_s > 4f_{\text{max}}) \) or the additional data points are interpolated. In addition, if the real signal is used then there is a positive and a negative part to the spectrum that will give rise to cross terms between them. These cross terms manifest themselves as low-frequency artifacts in the time-frequency distribution at low frequencies. In this paper the analytic signal is used to calculate the time-frequency distribution; the resulting joint distribution is commonly referred to as the Wigner–Ville distribution. The Nyquist sampling rate is sufficient when the analytic signal is used because the frequency spectrum of the analytic signal vanishes for negative frequencies and so only the positive frequency components of the signal are used to calculate the time-frequency distribution.
FIG. 1. Schematic diagram of the experiment in which an aircraft flies at a constant altitude (h) and speed (v) so that it transits directly over an acoustic receiver located on the ground. The horizontal range is denoted by r and the separation distance between the aircraft and the receiver (or slant range) is denoted by d; both r and d are functions of time.

III. ACOUSTICAL DOPPLER EFFECT

The Doppler effect is the change in the observed frequency of an acoustic or electromagnetic wave due to the relative motion of source and observer. Consider the present case of an aircraft flying along a straight line with a constant speed v and at a constant altitude h. The flight path of the aircraft passes directly over a stationary receiver located on the ground—see Fig. 1. Now an acoustic tone of constant frequency $f_0$ emitted by the aircraft at a time $t$ arrives at the receiver at a later time $t$ that is given by

$$t = r + d/c,$$  \hspace{1cm} (12)

where $d/c$ is the time delay for the sound to propagate at a constant speed $c$ between the aircraft and the receiver that are separated by a distance $d$.

The separation distance (or slant range) is given by

$$d = (h^2 + r^2)^{1/2} = \left[ h^2 + v^2(t - t_c)^2 \right]^{1/2}, \hspace{1cm} (13)$$

where $t_c$ is the time at the closest point of approach, which occurs when the aircraft is directly above the receiver, that is, when the horizontal range $r$ is zero. The aircraft’s altitude $h$ is equal to the distance of closest approach.

Substituting (13) into (12), then solving for $r$, it can be shown that

$$r = \frac{c^2 t - v^2 t_c - [h^2(c^2 - v^2) + v^2 c^2(t - t_c)^2]^{1/2}}{c^2 - v^2}. \hspace{1cm} (14)$$

Now the instantaneous frequency of the signal received at time $t$ is given by

$$f_i = \frac{d f}{dt} = \frac{f_0 c^2}{c^2 - v^2} \left( 1 - \frac{v^2(t - t_c)}{h^2(c^2 - v^2) + v^2 c^2(t - t_c)^2} \right) \left( \frac{1}{1 + \frac{v^2(t - t_c)}{h^2(c^2 - v^2) + v^2 c^2(t - t_c)^2}} \right), \hspace{1cm} (15)$$

or

$$f_i = \alpha + \beta z_i(t_c, s), \hspace{1cm} (16)$$

where

$$\alpha = f_0 c^2/(c^2 - v^2),$$

$$\beta = -f_0 c v/(c^2 - v^2).$$

IV. PARAMETER ESTIMATION

The parameters $v$, $h$, $f_0$, and $t_c$ are estimated by minimizing the sum of the squared deviations of the frequency estimates from their predicted values, that is,

$$\sum_{j=1}^{N} (g_{ij} - f_i)^2 = \sum_{j=1}^{N} [g_{ij} - \alpha - \beta z_i(t_c, s)]^2, \hspace{1cm} (17)$$

where $g_{ij}$ is the frequency estimate at time $t_j$ obtained using the short-time Fourier transform or the Wigner–Ville distribution, and $f_i$ is the predicted frequency given by (16).

For a fixed $t_c$ and $s$, (17) is minimized with respect to $\alpha$ and $\beta$, when

$$\alpha = \hat{\alpha} = g - \hat{\beta} \bar{z}, \hspace{1cm} \beta = \hat{\beta} = \frac{\sum_{j=1}^{N} (g_{ij} - \bar{g}) z_i}{\sum_{j=1}^{N} \bar{z}^2}, \hspace{1cm} (18)$$

where $\bar{g} = (1/N) \sum_{j=1}^{N} g_{ij}$, $\bar{z} = (1/N) \sum_{j=1}^{N} z_i$, and $z_i = z_i(t_c, s)$.

Thus the following expression is required to be minimized:

$$\sum_{j=1}^{N} (g_{ij} - \bar{g})^2 = \left( \frac{\sum_{j=1}^{N} (g_{ij} - \bar{g}) z_i}{\sum_{j=1}^{N} \bar{z}^2} \right)^2 \hspace{1cm} (19)$$

or equivalently, the following expression is maximized:

$$Q = \sum_{j=1}^{N} (g_{ij} - \bar{g} z_i)^2 \hspace{1cm} (20)$$

The following iterative procedure, which requires initial estimates of $t_c$ and $s$, is used to maximize $Q$.

Step 1: Determine the initial estimate of $t_c$, which is chosen to be

$$t_c = \frac{1}{2} (t_k + t_{k+1}), \hspace{1cm} (21)$$

where

$$k = \arg \max_j \frac{(j \bar{g} - \sum_{i=1}^{j-1} g_{ij})^2}{j(N-j)}, \hspace{1cm} \text{for} \hspace{0.5cm} 1 \leq j \leq N-1.$$

Step 2: Determine the initial estimate for s by performing a coarse binary search of $Q$ over the range $0 < s < s_{\text{max}}$, where $s_{\text{max}}$ is equal to the maximum height divided by the minimum speed of interest. In this paper, $s_{\text{max}} = 10$.

Step 3: Refine the estimates of $t_c$ and $s$ by maximizing $Q$ using the Gauss–Newton method on the vector of partial derivatives of $Q$ with respect to $t_c$ and $s$. The initial estimates of $t_c$ and $s$, given above in steps 1 and 2, respectively, are used as the starting point for the Gauss–Newton method.
FIG. 2. (a) The variation with time of the Doppler-shifted blade rate observed during transit number 2. The observed variation is compared with the variation predicted by a simple model that provides estimates of the aircraft's speed and altitude, together with the source (or rest) frequency of the propeller blade rate. The frequency estimates are obtained using the short-time Fourier transform (STFT). (b) Similar to (a) but the Wigner–Ville distribution (WVD) is used for the time-frequency signal analysis.

Put the resulting estimates \( \hat{\nu} \) and \( \hat{h} \) equal to the final values of \( \nu_c \) and \( s \), and \( \alpha \) and \( \beta \) equal to the values of \( \alpha \) and \( \beta \) given by (18).

The estimate of the aircraft's speed is given by

\[
\hat{\nu} = \frac{1}{\hat{\beta}^2} \nu_c.
\]  

(22)

The estimate of the aircraft's altitude is given by

\[
\hat{h} = \frac{\hat{\nu} \nu_c}{(c^2 - \hat{\nu}^2)^{1/2}}.
\]  

(23)

The estimate of the source frequency of the propeller blade rate is given by

\[
\hat{f}_0 = \hat{\alpha}(1 - \hat{\nu}^2/c^2).
\]  

(24)

V. EXPERIMENT AND DATA PROCESSING

A turbo-prop aircraft having a constant propeller blade rate of 68 Hz is flown at a constant speed and altitude so that its flight path passes directly over a microphone located just above the ground. Altogether, there are 18 transits formed by combining each of the nominal aircraft speeds: 150, 200, and 250 kn with each of the nominal aircraft altitudes: 250, 500, 750, 1000, 1250, and 1500 ft.

During each aircraft transit, the acoustic data from the microphone are recorded as a digital time series for subsequent time-frequency signal analysis using the short-time Fourier transform and the Wigner–Ville distribution. The sampling rate is 889 Hz and the Doppler-shifted blade rate is estimated every 0.072 s (64 samples). A rectangular window function is selected for the analysis in which \( w(m) = 1 \) for \( |m| < M \), where \( M = 128 \); \( w(m) = 0 \) elsewhere. Hence there is a 75% overlap in the processing of the data. The size of the fast Fourier transform is 32768 points.

Various techniques for estimating the instantaneous frequency of a nonstationary signal have been considered by Boashash and Jones.\(^5\) The peak detection method using the time-frequency distribution is adopted in this paper. The frequency estimate for the instantaneous blade rate corresponds to the frequency at which the power spectrum (calculated using the short-time Fourier transform) is a maximum. Similarly, a frequency estimate for each window of data is also obtained by finding the frequency at which the Wigner–Ville distribution is a maximum.
FIG. 4. (a) Comparison of the estimated and actual ground speed of the aircraft for each of the 18 transits made by the aircraft. The short-time Fourier transform (STFT) was used to provide estimates of the Doppler-shifted blade rate as a function of time during each transit. (b) Similar to (a) but the Wigner–Ville distribution (WVD) was used for the time-frequency signal analysis.

FIG. 5. (a) Similar to Fig. 4(a) but for the aircraft altitude. (b) Similar to Fig. 4(b) but for the aircraft altitude.
The time sequence of spectral estimates is then used to estimate the aircraft’s speed and altitude, together with the rest frequency of the propeller blade rate, and the time of closest approach. The speed at which sound propagates in the atmosphere is assumed to have a constant value of 660 km.

Two examples of the results of the data processing are shown in Figs. 2 and 3, which apply to transits 2 and 15, respectively. The diagrams show the time variation of the frequency estimates (indicated by circles and labeled “OBSERVATION”) obtained using the short-time Fourier transform (STFT) and Wigner–Ville distribution (WVD). Also shown are the actual values of the speed, altitude, and blade rate recorded onboard the aircraft, together with the corresponding values estimated by processing the acoustic data; the solid line (labeled “MODEL”) is the predicted variation with time of the Doppler-shifted blade rate using the estimates for $v$, $h$, $f_0$, and $t_c$. A comparison of Figs. 2 and 3 shows that the magnitude of the Doppler shift in the propeller blade rate is much greater at the higher aircraft speed.

VI. RESULTS

The results of processing the acoustic data for all eighteen transits are shown in Figs. 4–6 in which the estimated values are compared with the actual values for the aircraft’s speed [see Fig. 4(a) and (b)], altitude [see Fig. 5(a) and (b)], and propeller blade rate (see Fig. 6).

The root-mean-square error between the estimated and actual values for all 18 transits is: 5 km (both for STFT and WVD) for the aircraft’s speed; 90 ft (for STFT) and 60 ft (for WVD) for the aircraft’s altitude; 0.4 Hz (for STFT) and 0.5 Hz (for WVD) for the propeller blade rate. Thus this groundbased passive acoustic technique provides estimates for the aircraft’s speed, altitude, and propeller blade rate that are in good agreement with the actual values.

VII. APPLICATION TO ANOTHER AIRCRAFT

On a separate occasion, the pilot of an aircraft of opportunity was requested to fly over the microphone at a speed of 150 km. The results of the time-frequency signal analysis are shown in Fig. 7(a) for the short-time Fourier transform, and in Fig. 7(b) for the Wigner–Ville distribution. Also, shown are the estimated values for the aircraft speed, altitude, and propeller blade rate.
VIII. CONCLUSIONS

The speed and altitude of a propeller-driven aircraft, together with its propeller blade rate, can be estimated by processing the acoustic data from a microphone during the aircraft’s transit overhead. The short-time Fourier transform and the Wigner–Ville distribution produced similar results when used to estimate the Doppler shift in the propeller blade rate as a function of time.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the technical discussions with Professor Boualem Boashash, Queensland University of Technology, on time-frequency signal analysis, and the operational advice and guidance provided by Commander C. W. Donald, RAN. This work was supported by the Cooperative Research Centre for Robust and Adaptive Systems which is funded by the Australian Government under the Cooperative Research Centres Program.