inter.noise 2004
Prague • Czech Republic • August 22-25

The 33rd International Congress and Exposition
on Noise Control Engineering

Aircraft sound level measurements in residential areas using sound source separation.

T.C. Andringa [a,c], P.W.J. van Hengel [a], R. Muchall[b], M.M. Nillesen[a]

[a]Sound Intelligence, Sint Jansstraat 2, 9712 JN, the Netherlands
[b]Engineering Company Omegam-Geluid, Haverkamp 15, 3828 HK Amersfoort, the Netherlands
[c]University of Groningen, Dep. of Artificial Intelligence, Groningen, The Netherlands

[a] [t.andringa;p.van.hengel;m.nillesen]@soundintel.com
[b]RM.Omegam-Geluid@planet.nl

Abstract [460] Measuring the contribution of a particular sound source to the ambient sound level at an arbitrary location is impossible without some form of sound source separation. This made it difficult, if not impossible, to design automated systems that measure the contribution of a target sound to the ambient sound level. This paper introduces sound source separation technology that can be used to measure the contribution of a sound source, a passing plane, in environments where planes are not the dominant sound source. This sound source separation and classification technology, developed by Sound Intelligence makes it, in principle, possible to monitor the temporal development of any soundscape.

The plane detection technology is based on Continuity Preserving Signal Processing (CPSP): signal processing technology mimicking the tracking and streaming ability of the human auditory system. It relies on an efficient implementation of the human cochlea which, like the natural system, preserves the continuous development of sound sources better than is possible with traditional systems. Detection and classification are based on the characteristic development of the target sound source. In the case of airplanes the characteristic property set includes spectral cues in combination with a characteristic temporal development and techniques to deal with prominent transmission effects (wind and reflections of buildings).

This airplane sound detection technology has not been published before. Experiments at a residential location about 25 km from Amsterdam-Airport demonstrate that it is possible to estimate the contribution of passing planes to the ambient sound. Even in the case when the target signal is only marginally (5 dB) above background noise and degraded due to wind and reflections, a detection performance comparable to human listeners is achieved.

KEYWORDS: Airplane detection, sound source detection.

1 INTRODUCTION

The human auditory system is still the best system for the analysis and recognition of arbitrary mixtures of input sounds. In many cases it is possible to develop sound recognition systems (such as speech recognition systems) that function adequately in some limited acoustic environment, but these systems cannot be generalized to more unconstrained environments. This entails that it has not been possible to develop sound source detection systems that can detect sound sources if and only if they are actually present in the input sound. Sound Intelligence develops sound analysis and detection technology that approximates this ideal.

A challenging and useful application of this technology is in the field of environmental monitoring, where the task is to determine what contribution to the overall sound level can be assigned to a certain sound source. Around airports, the area in which aircraft sound nuisance is experienced is much wider than the perimeter in which commercial sound detection systems can determine the contribution of aircraft sound (typically averaged over a certain period with a measure like the L_{den}). To allow reliable measurements of the contribution of aircraft sound in the urban areas, Sound Intelligence and Omegam-Geluid decided to cooperate. Sound Intelligence provided a commercial jet detection system based on the analysis of monaural input. Omegam-geluid provided a complete measurement set-up (of type Luistervink) and performed a field-test in an urban area (in Krommenie, near Amsterdam) about 25 km from Schiphol-airport. This publication presents the results from the first deployed system without any further optimization.

2 GENERAL APPROACH

The approach to sound source separation and sound source recognition pursued by Sound Intelligence is to form signal components [2]. These are representations that are very likely to contain information about a single sound source. Signal components can be combined into streams of information stemming from a single source if and only if they comply with the properties of a certain target class (speech, vehicle sounds, jets, etc.). Components that do not comply with the properties of the target classes are ignored. The notion of signal components stems from the observation that all physical sounds sources show characteristic *continuous* developments. Depending on the properties of the source and the transmission medium, these signal components might be individual harmonics, noise bursts or noise bands. Each of these signal components is coupled to some physical process that as a rule is restricted by some form of inertia and hence reflects a single continuous development, or a sequence of continuous developments separated by qualitative changes. Studies on auditory perception and auditory scene analysis are based on the assumption that signal components are tracked by the auditory system and combined into higher level representations [3].

Traditional frame-based signal analysis technology, such as the Fast Fourier Transform and Wavelet analysis, are unsuitable for tracking signal components. In fact they actually reduce the continuous nature of the signal components by assuming that an interval (window) must be analyzed in isolation from other (possibly even overlapping) intervals. The mammalian hearing system on the other hand conserves the continuity of individual signal components to a very high degree. The basilar membrane[1] behaves as a single physical structure that is sensitive to different frequencies along its length, with a continuous place-frequency relation. This leads to a

[1] The basilar membrane is located in the cochlea (inner ear) and plays the most important role in the frequency-time analysis in human hearing.

representation in which a continuous development in both time and frequency is conserved for further processing. When a mixture of signal components excites the basilar membrane, an excitation pattern arises in which different basilar membrane positions are dominated by different signal components. Each signal component enforces its oscillation upon the region it dominates. Since signal components originate from a single physical process it is possible to measure its spectro-temporal properties of this process from this region.

The estimation of all perceptible source properties from an unknown mixture of input sounds is a very difficult process that is not yet fully understood. But a number of techniques have been developed by Sound Intelligence and the University of Groningen that lead to the forming of representations that contain information of a single sound source. Together these techniques are known as Continuity Preserving Signal Processing. Initially CPSP was applied mostly on quasi-periodic signal components. This publication presents a first example application to target and background sounds that are predominantly aperiodic.

After the forming of signal components, the signal must be analyzed further to determine whether or not the acoustic evidence is consistent with knowledge of properties of the target class. In the case of jet detection this entails that a subset of the acoustic evidence must comply with the characteristic spectro-temporal development of passing jets. This requires knowledge about the physical properties of passages, as well as knowledge about the acoustic influence of the environment. Without detailed information on these two aspects, it can be estimated from the analysis of a number of examples of passing jets. Because the physical properties of jets do not vary much, nor do the environmental acoustic influences, this number can be small (in this case less then 10). The classification process is therefore knowledge-driven and not data-driven.

3 AIRPLANE DETECTION SYSTEM

To detect airplanes, the characteristic developments of both spectral and temporal cues are used. The general detection system consists of three steps. First a continuity preserving preprocessing step is performed by processing the incoming sound with a cochlea model. The second step is a selection step that selects only parts of the incoming sound that are likely to originate from the target source. Finally a classification step compares the selected parts of the sound with a reference distribution of sounds of the target class. These three steps will be described in more detail below.

3.1 Preprocessing: the cochleogram

The frequency analysis performed by the cochlea model (see [4]) results in a cochleogram [2], which can be thought of as the response of the human inner ear. This cochleogram shows some resemblance with a spectrogram obtained with short-term FFT or wavelet analysis, but has substantial advantages over these. The most important advantage is that in the cochleogram continuity in both time and frequency can be guaranteed. Therefore coherent signal components stemming from a single sound source can easily be tracked in time and frequency. A cochleogram typically has a logarithmic frequency scale.

Figure 1 shows the cochleogram of a passing airplane. The sound is part of the recordings made with the ‘Luistervink’ system developed by Omegam-Geluid [5]. The coherent frequency components of the airplane can easily be distinguished visually. The cochleogram also shows the noisy structure of the background sound and the airplane sound itself.

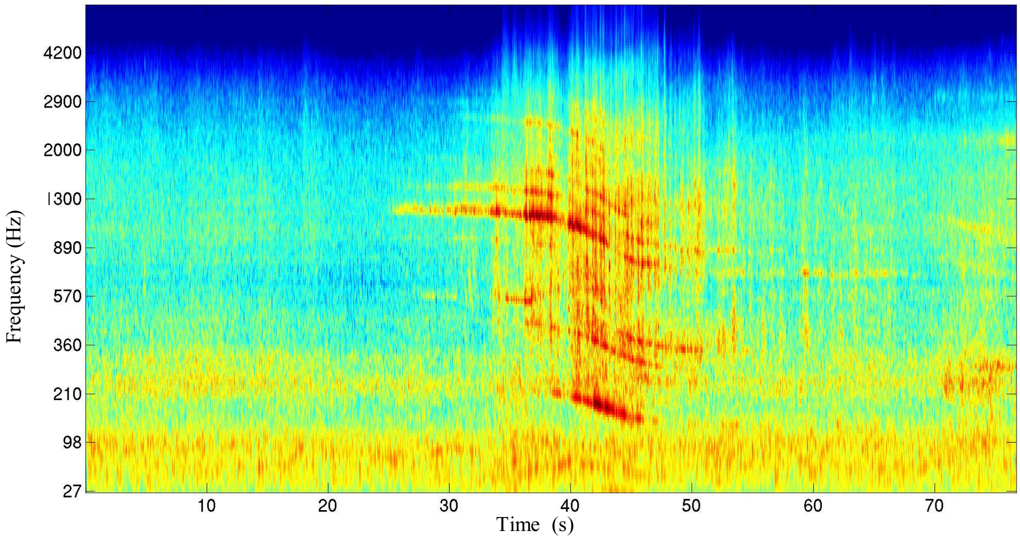

Figure 1: Cochleogram of the sound of an airplane passage recorded with the Omegam Luistervink system

3.2 Selection

This step mainly uses temporal characteristics of the target sound.
A sound source is partly characterized by its dynamics in time. These characteristic temporal dynamics are used to attenuate signals corresponding to parts of the cochleogram that disagree with these dynamics. This process causes a major improvement in the ratio between the target sound and the sounds originating from other sound sources.
The airplane detection system selects sounds with temporal dynamics faster than 6 dB per 50 seconds and slower than 6 dB per 500 ms. Figure 2 illustrates the result of this process: it shows only parts of the cochleogram that meet the desired temporal dynamics. It can be seen that the airplane sound (fore ground signal) is preserved and hardly modified, whereas background sounds are suppressed.

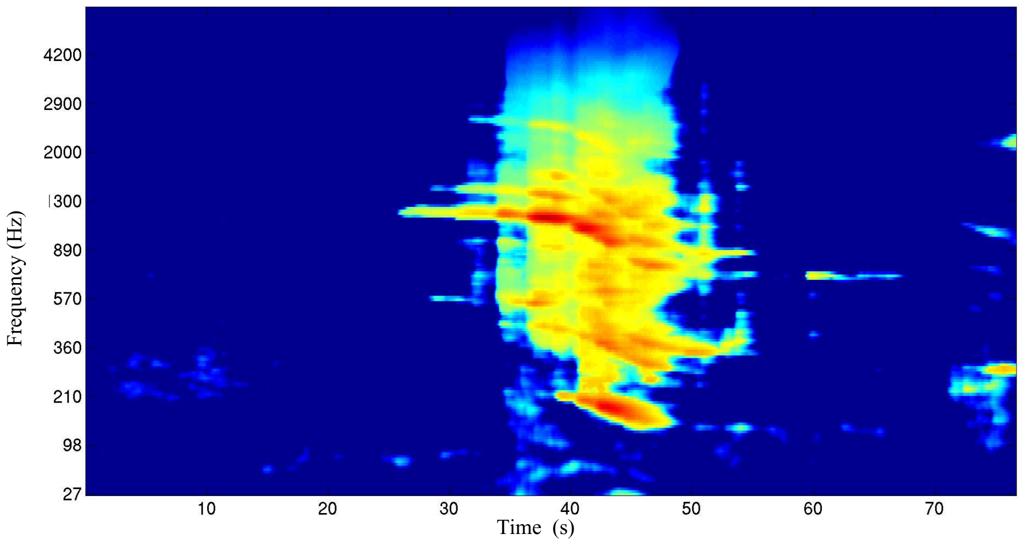

Figure 2: Subset of the cochleogram (from figure 1), which contains only parts of the cochleogram that satisfy the temporal dynamics of an airplane passage.

Additionally, a background model is computed which contains only sounds with a temporal dynamics slower than 6 dB per 50 seconds. This background model is used to decide whether the selected parts of the cochleogram can indeed be attributed to an airplane passage.
An airplane passage must satisfy the following criteria:
The average level of the foreground signal must exceed the average level of the background level by at least 5 dB at the top of the passage, where level is measured as energy in the (filtered) cochleogram. Figure 3 shows the average energy value of the foreground (compare figure 2) above the background (dark blue line). Positive values indicate that the foreground exceeds the background. The boundaries of a passage are then determined on the basis of the zero-crossing points. As can be seen, at the top of the passage, the foreground exceeds the background by about 14 dB, therefore satisfying the 5 dB criterion mentioned earlier.
Further requirements are that, at the top of the passage, the percentage of segments of the basilar membrane where the foreground level exceeds the background should be least 70% and that the minimum required passage length is 4 seconds.

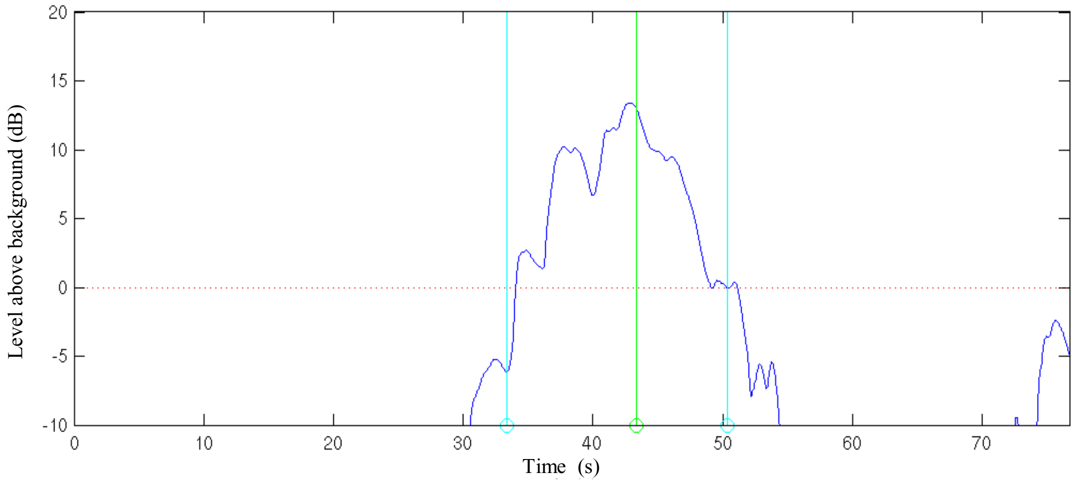

Figure 3: Average energy value of the filtered cochleogram exceeding the background. Starting point, end and maximum of the detection are indicated.

3.3 Classification

Despite the improved ratio between target sounds and non-target sounds obtained with the above described time-domain selection method, not every selection will be a passage. A final classification step is required to ensure that a selection is indeed the result of a passing airplane. To achieve this the spectrum at the top of the passage (i.e., when it reaches its maximum value above the background) is analyzed. This is done by comparing this spectrum with a so-called reference spectrum. Figure 4 shows the spectrum of the airplane sound at t=43.4 sec, corresponding to the maximum in figure 3.

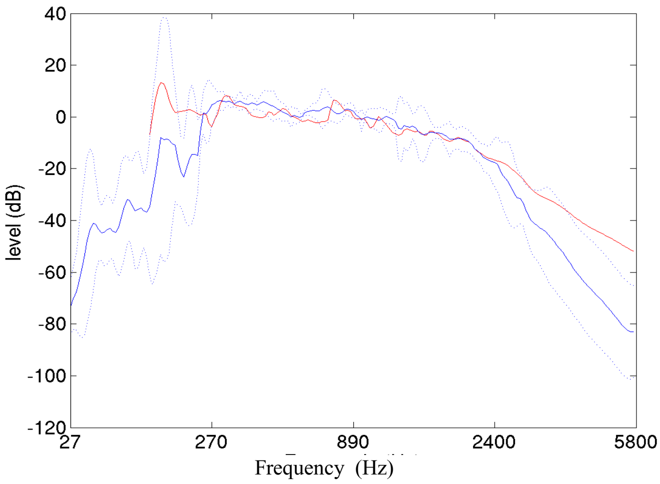

Figure 4: Comparing the target spectrum with the reference spectrum. The blue line denotes the reference spectrum for airplanes. This spectrum is obtained by calculating an average spectrum of several airplane passages. The dotted lines denote the standard deviation of the different passages. The red line shows the spectrum at the top of the detected airplane passage. Only reliable parts (exceeding the background level) of the target spectrum are shown.

The reference spectrum is determined on the basis of other airplane passages. The spectrum of the measured airplane passage matches well with the reference spectrum. Using the deviation of the target spectrum relative to the reference spectrum, the probability that this sound corresponds to the target class is calculated.
Other parameters that contribute to the probability of being an airplane are the relative energy value at the top of the passage, the percentage of segments above the background at the top of the passage and the passage length. This probability increases linearly from 0 to 1 for respectively 5 dB to 7 dB (top value), 70% to 90% (percentage of segments) and 4 seconds to 10 seconds (passage length). The total probability of detection an airplane is now obtained by multiplying all four probabilities. If this total probability is higher than a detection threshold, an airplane is detected.

4 DETECTION TEST

The airplane detection system was evaluated by Omegam-Geluid, which has extensive knowledge of and experience with automatic aircraft sound level measurements. The measurements described here were made with a new version of the Luistervink system [5] that was placed in an urban area in Krommenie, northwest of Amsterdam and about 25 km from the runways where the passing aircrafts take off and land. The system is placed on the roof of a sports club canteen. East of the measurement site are a number of high apartment buildings that reflect aircraft sounds and therefore contribute to beats in the signal.
.

4.1 Measurement setup

This system consists of a noise level meter with an outdoor microphone and wind speed meter. Compared to the previous setup [5] a new development is the use of a microphone house. This is a 2.5 x 2.5 m wide and 60 cm high casing covered with a 50 % open weather proof clothing and is furnished with a sound absorbing bottom with rock wool baffles. The unit is normally placed on the flat roof of a building.

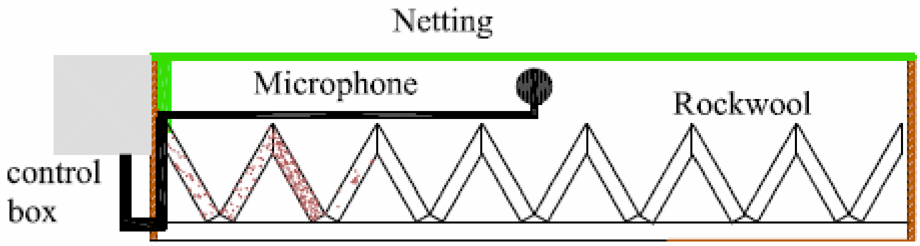

Figure 5: section of microphone house

This construction has a number of advantages compared to a 6 m high microphone mast. The 2.5 m open clothing produces much less wind noise than an ordinary windscreen. The casing shields the microphone from disturbing noise from ground level sources. The sound-absorbing bottom eliminates the ground reflection and the monitor unit is practically invisible from the ground level because of its low dimensions.

The plane detector software is the recognition part of the measurement software of the system. Another software part records parameters of every noise event like L_{max}, L_{eq}, event duration, rise time, and wind speed. After the validation procedure with the plane detector a number of noise climates are calculated like L_{den}, L_{night}, etc.

4.2 Test procedure and results

A test procedure was followed to establish the reliability and the accuracy of the recognition system. Therefore another noise monitor system, type *Symphonie* from manufacturer 01dB was installed for a week. This system continuously recorded the A-weighted noise level with a sampling speed of 10 per second. It also made a sound recording of all noises which exceeded the background level with more than 10 dB. Each recorded noise event was classified with the human ear and compared with the classification of the Sound Intelligence plane detector. This led to the following results.

Item	Value
Total manually classified plane events Lmax > 50 dB(A)	591
Events by detector classified as plane > 50 dB(A)	580
Incorrect classified noise events Lmax > 55 dB(A)	6
Incorrect classified noise events with Lmax 50-55 dB(A)	17
Missed aircraft noise events > 50 dB(A) not classified as plane	34
Correct classified aircraft noise events	557
Average Lmax aircraft noise events	55 dB(A)
Average L95 background level	42 dB(A)
Leq test period by manual classified aircraft noise events	43.1 dB(A)
Leq test period by detector classified aircraft noise events	43.27 dB(A)
Total measurement correction	0.17 dB(A)
Measurement correction by incorrect plane classification	-0.1 dB(A)
Measurement correction by missed plane classification	0.27 dB(A)

In total 94.3% of the passing jets were detected correctly. During the measurement period of one week, 23 events were incorrectly classified as jets, where the total number of plane events was almost 600. The main disturbances were train noise from a railway at 1 km distance and wind noise. Two loud aircraft noise events were missed because of disturbance noise prior to the noise event. The recognition accuracy of 0.17 dB(A) is better than the 1 dB measurement accuracy of the type 1 noise level meter and meets the required accuracy of the monitor installation.

5 CONCLUSION

A number of conclusions can be drawn:

- It is possible to measure reliable L_{eq} values for passing airplanes in urban areas at 25 km from airports.
- A little effort on optimization is likely to improve the performance considerably. Especially the probability for correct detection will improve and the number of incorrect classifications will decrease after improvement.
- Optimization will have only a minimal influence on the L_{eq}-value since the estimation of this value is very good.
- Because the measurement error is well below the acceptance level, it is very likely that this approach will also work at greater distances.
- The detection system has not been optimized for this location. Therefore, only a minimal optimization will suffice for other locations.

CPSP and the classification methodology have not been developed specifically for this task. On the contrary, the detection and recognition approach aims to be as general as possible, and the performance of this detection system merely is an indication of the versatility and power of the general approach.

6 REFERENCES

[1] Andringa, T.C. et al., 1999, International Patent Application WO 01/33547, "Method and Apparatuses for Signal Processing".

[2] Andringa, T.C. 2002, "Continuity Preserving Signal Processing", (PhD Thesis), University of Groningen, Groningen, http://www.ub.rug.nl/eldoc/dis/science/t.c.andringa/.

[3] Bregman, A.S., 1990, "Auditory Scene Analysis: The perceptual organisation of sound", Cambridge (MA), Bradford Books, MIT Press.

[4] Duifhuis, H., Hoogstraten, H.W., van Netten, S.M., Diependaal, R.J. en Bialek, W., 1985, "Modelling the cochlear partition with coupled Van der Pol oscillators", Cochlear Mechanisms: Structure, Function and Models, eds. J.W. Wilson and D.T. Kemp, Plenum, New York, pp 395-404.

[5] Muchall, R.C., 1994, "Ontwikkeling, ervaringen en onderzoeksresultaten van een permanent monitorsysteem voor vliegtuiglawaai uitgerust met patroonherkenning", NAG-journaal 124, pp 19-30.