Energy mobility method for vibroacoustic coupling

Myong-Sok RYU, Christian CACCIOLATI and Jean-Louis GUYADER

The Energy Mobility method was developed to predict the vibration behavior of structures in the medium frequency range. This method must only be applied to structures coupled by several points. It is a predictive method based on the classic mobility, and using data of uncoupled subsystems. It establishes the relationship between the frequency average of the active injected power at the excited points and the frequency average of the quadratic velocity at any point of the assembly in a frequency band. In this paper, it is extended to incorporate the case of vibroacoustic coupling through a surface between a structure and a fluid medium. The method is numerically tested using a simply supported rectangular plate located on a open face of a rectangular cavity with five rigid walls. The kinetic energy on the plate, the radiated power into the cavity and the radiation efficiency are investigated.

INTRODUCTION

Classic mobility \(Y_{ij}\) is defined as the ratio of the velocity \(V_i\) at a point of reception \(i\) to the force \(F_j\) at a point of excitation \(j\) on the linear and elastic structure in the frequency domain \((f)\) [1]. For vibrational and acoustical problems, this is used as a predictive method to the coupled or global system between substructures or subelements [2]. Energy Mobility \(H_{ij}\) is based on the classic mobility in a frequency average domain \(\langle \mathcal{A} \rangle\) [3, 4]. It is defined as

\[
H_{ij} = \left\langle |V_i|^2 \right\rangle_f / \left\langle Re \{ Y_{ij} \} \right\rangle_f
\]

where \(\langle \cdot \rangle_f\) designates the frequency average in the frequency band \(\mathcal{A}\). By using the time-frequency averages of the quadratic velocity \(|V_i|^2\) and of the active injected power \(\Pi_f\), we can write

\[
\left\langle |V_i|^2 \right\rangle_f = \sum_j H_{ij} \Pi_j
\]

The method may be used on a single structure, and in the assemblies connected by the points. The studied points \(i\) and \(j\) must be sufficiently spaced to have:

\[
\left\langle Re \{ Y_{ij} \} \right\rangle_f << \left\langle |V_i|^2 \right\rangle_f
\]

In vibroacoustic problems, the structures are coupled through their surface to the fluid. The existing Energy Mobility method can not be directly used because of surface coupling between the substructures. The aim of this paper is to extend the Energy Mobility method to the vibroacoustic coupling system. A simply supported rectangular plate located on a open face of a rectangular cavity with five rigid walls, will be considered.

THEORETICAL APPROACH

The coupling surface between the plate and the cavity is meshed in patches. By integrating the classic mobility over the areas of patches \(i\) and \(j\), we define the equivalent classic mobility \(Y_{ij-eq}\) between the patches :

\[
Y_{ij-eq} = \left\langle \Pi_j / \Delta s_j \right\rangle_f
\]

where \(\langle \cdot \rangle_f / \Delta s_j = \text{mean value over the surface of patch } j\). This is resulting from the assumption of constant pressure and velocity at each patch. Accordingly, the equivalent energy mobility \(H_{ij-eq}\) between the patches is defined by using the equation (1),

\[
H_{ij-eq} = \left\langle \left\langle |V_i|^2 \right\rangle_f / \left\langle Re \{ Y_{ij} \} \right\rangle_f \right\rangle_f / \Delta s_i / \Delta s_j
\]

The frequency average of the quadratic velocity is then obtained by the equation (2),

\[
\left\langle |V_i|^2 \right\rangle_f = \left[ H_{i-k-eq} + \left[ \left\langle |V_i|^2 \right\rangle_f / \Pi_i \right] \right] / \left[ H_{i-k-eq} + \left[ \left\langle |V_i|^2 \right\rangle_f / \Pi_i \right] \right]
\]

where the symbol ~ signifies before coupling. The frequency average of the exchanged power over the coupled patch \(c\), \(\left\langle \Pi_{i-c} \right\rangle_f\) is obtained by using the connectivity conditions at coupling surface:

\[
\left\langle \Pi_{i-c} \right\rangle_f = \left[ H_{i-c-eq} + \left[ \left\langle |V_i|^2 \right\rangle_f / \Pi_i \right] \right] / \left[ H_{i-c-eq} + \left[ \left\langle |V_i|^2 \right\rangle_f / \Pi_i \right] \right]
\]

where the subscript \(k\) indicates the coupled patch. In the following numerical test, we suppose that the active injected powers \(\left\langle \Pi_i \right\rangle_f\) and \(\left\langle \Pi_j \right\rangle_f\) are not modified after coupling.

APPLICATION

Figure 1 shows the test model, which consists of a simple rectangular homogeneous plate and a cavity which has one open face and five rigidly closed faces. A point force is applied to the plate. The characteristics of subsystems are given in Table 1.
After coupling of subsystems, the total kinetic energy on the plate \( E_{\text{kin}} \), the transmitted power into the cavity \( \Pi_{\text{tr}} \) and the radiation efficiency \( \sigma \) were investigated. This was done using both the Energy Mobility method and the exact calculations for different numbers of patches (Figure 2).

![Test model: a point force, a plate and a cavity.](image)

**FIGURE 1.** Test model: a point force, a plate and a cavity.

**Table 1.** Characteristics of the subsystems.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Plate</th>
<th>Cavity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimension ((m))</td>
<td>0.5\times0.6\times0.005</td>
<td>0.5\times0.6\times0.7</td>
</tr>
<tr>
<td>Mass per volume ((Kg/m^3))</td>
<td>7800</td>
<td>1.25</td>
</tr>
<tr>
<td>Damping factor</td>
<td>0.05</td>
<td>0.02</td>
</tr>
<tr>
<td>Young’s modulus ((Pa))</td>
<td>(2.11\times10^{11})</td>
<td>0.33</td>
</tr>
<tr>
<td>Poisson’s ratio</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speed of sound ((m/s))</td>
<td></td>
<td>344.8</td>
</tr>
</tbody>
</table>

Whatever the number of patches was, the \( E_{\text{kin}} \) calculated by the present approach matches well the exact calculation within the whole frequency range. In the considered case, coupling is weak, \( E_{\text{kin}} \) is almost unaffected by the cavity and the number of patches has a small influence on the plate vibration. The fluid-structure coupling is characterized by the transmitted power into the cavity.

![Graphs of E_{kin}, \Pi_{tr}, and \sigma against frequency.](image)

**FIGURE 2.** Total kinetic energy on the plate \((a)\), transmitted power into the cavity \((b)\) and radiation efficiency \((c)\) for the surface coupling system excited by a harmonic point force of 1N at a point \( j \) (L_x/2,L_y/6) on the plate. \( \Delta f = 400\text{Hz} \), Exact calculation \((-\bullet-\)) , Energy Mobility calculations with \(15\times15\) patches \((\square)\) and with \(3\times3\) patches \((\Delta)\).

The values of \( \Pi_{\text{tr}} \) and \( \sigma \) calculated with \(3\times3\) patches by the Energy Mobility method are a good approximation of the exact results, contrary to the calculation with \(15\times15\) patches. This strange result is due to the fact that equation (2) is a good approximation when the meshing points are sufficiently far to have weak correlation. At low frequency using \(15\times15\) patches violates the underlying assumption of equation (3). In this study we observed that it is convenient to have less than 4 patches over one acoustical wavelength to obtain a good prediction.

**CONCLUSIONS**

A plate-cavity system was studied. The Energy Mobility method can give a good prediction up to medium frequency with a small number of patches. The advantages of the presented method are to mesh the coupling surface with a small number of large patches and to use mobilities of uncoupled subsystems.

**REFERENCES**

Acoustic Characteristic of Elliptic Cavity Having a Non-uniformly Perforated Pipe

N. Tsuayoshi, E. Hiroyuki, N. Sohei and A. Tadashi

Department of Engineering, Sojo University, 4-22-1 Ikeda, Kumamoto 860-0082, Japan

The non-uniformly perforated pipes which these orifices are opened in a belt-shaped manner is widely used in muffler systems. The resonance of such perforated pipe when located inside the elliptic cavity is investigated experimentally in this paper. The results obtained are also compared with the predictions of our theoretical analysis method for an uniformly perforated pipe in order to clarify the resonance mechanism. It was found that the difference between the value measured and the theoretical value is less than 5Hz for a short pipe.

**EXPERIMENT AND CALCULATION**

Seven perforated pipes in elliptical form of the same size was used in the experimental. Each pipe has the dimension of length $L=20$cm, eccentricity $e_w=0.86$ and haft length of major axis $a_w=8.7$cm. These holes of diameter $d=6$mm and thickness $t=0.8$mm are punched at 5 rows with constant axial spacing 2.5cm. Distance $x$ from the end of pipe to the center of belt-shaped were varied among 4, 6, 8 and 10cm as shown in Fig. 1. Each perforated pipe is mounted inside the elliptical chamber which has eccentricity $e_w=0.5$ and haft length of major axis $a_w=15$cm. A block diagram of the experimental apparatus is shown in Fig. 2.

When the uniformly perforated pipe is installed inside the elliptical chamber, the sound pressure at the output side is determined as follows [1]

$$P_{out} = jZ_wU_0\left(-\frac{1}{\sin kL} + \frac{kQ_{m,j}^2e_m(\zeta, s_{m,j})e_m(\eta, s_{m,j})}{\gamma_{m,j}}\right)$$

where

$$\gamma_{m,j} = \mu_{m,j}\sinh\mu_{m,j}L + \bar{Z}_{m,j}\pi_{m+1,j}\sinh\pi_{m+1,j}L$$

$$\mu_{m,j} = \frac{1}{a_w}\sqrt{(ka_w)^2 - \bar{Z}_{m,j}^2}$$

$$\pi_{m+1,j} = \frac{1}{a_w}\sqrt{(ka_w)^2 - \bar{Z}_{m+1,j}^2}$$

$k$ is the wave number, $\bar{Z}_{m,j}$ is a function of perforated pipe impedance, the other symbols are defined by [1].

First term of Eq. 1 represents the sound pressure of plane wave and the second one denotes the sound component of the higher-order modes. These resonances correspond to the frequencies at which the denominators $\sin kL$ and $\gamma_{m,j}$ are zero.

Relationship between experiment value and $P_{out}$ is described by

$$20\log\left|\frac{P_B}{P_A}\right| = 20\log\left|\frac{P_{out}/U_0}{Z_0\sin kL}\right|$$

where $P_A$ and $P_B$ are the sound pressure of microphone (1) and microphone (2), $U_0$, $Z_0$ are the volume velocity supplied from the input pipe and the characteristic impedance of input pipe, respectively.

**Discussion**

Figure 3(a) shows the experiment result. A horizontal axis is shown as substitution of the frequency by the product of $a_w$ and the wave number $k$. Figure 3(b) and Fig. 3(c) are the calculated results of Eq. (1) and its denominator $\sin kL$, $\gamma_{m,0}$, respectively. shows a point at which $\sin kL=0$, show the resonances of higher-order mode (1,0), show

**FIGURE 1.** Shape of perforated pipe.

**FIGURE 2.** Experimental apparatus.
the resonances of mode \((2,0)\) and show the resonances of mode \((0,1)\), respectively. Figure 4 is the measurement results when the porosity is 0.75%. It is seen that difference of the resonance frequencies between the measured and the calculated is less than 5Hz. Therefore, it is thinkable that the theoretical calculation of the uniformly perforated pipe can be used to determine the resonance frequencies of the non-uniformly pipe for the short pipe.

REFERENCES

1. T. Ikeda, S. Nishimura, T. Nitta and T. Nishimura
   "Analysis of sound propagation in perforated tube muffler"
   in Proceedings of 17th International Congress on Acoustics.
Estimation of the Directions of Incident Plane Waves in Case That the Sound Sources Are Complex Pure Tone -Basic study of the error of the directions-

Noboru Kamiakitoa, Yasuhiro Yamashitaa, Masanao Owakib, Takefumi Zaimab, Takeshi Sugiyamac

aFaculty of Engineering, Univ. of Shinshu "4-17-1 Wakasato, Nagano city 380-8553, Japan
bTechnical Research & Development Institute, Kumagaigumi, CO., LTD."1043 Onigakubo, Tukuba city 300-2651, Japan
cElectric Power Research Center of Chubu Electric Power CO., INC. 20-1 Kitasekiyama, Ohdaka-cho, Midori-ku, Nagoya city 459-8522, Japan

We have investigated the estimation of incidence direction of plane wave in the case where the sound sources are complex pure tones like transformers. In this case, we show some difficulty in estimating the angle error of incidence of plane waves by reflection compared with the random noise signal. From the basic experiment to the pattern of characteristic of the error by reflection, we could confirm the influence of reflection theoretically, in the case of calculation, and the measurement outdoors.

INTRODUCTION

There are noise problems in complex pure tones generated factories. It is important that we estimate the direction of incident plane waves for the purpose of the improvement of the noise, but generally, it is difficult to estimate the directions of incident plane waves. This paper describes the way the estimation of the directions of incident plane waves in the case that the sound sources are complex pure tones.

THEORY

Location of microphones is shown in Figure 1. $\theta$ is the angle of sound direction. The estimated equation is given by

$$\theta = \tan^{-1} \frac{D_{x2}}{D_{y3}}$$

where Dij is delayed from Mi to Mj.[1][2]

ERROR BY REFLECTION

Position of calculation for the reflection is shown in Figure 2. Sound source is a complex pure tone in 60Hz, 120Hz, 240Hz, 360Hz. Thus characteristic of the error in frequency of the sound sources are shown in Figure 3. This characteristic is explained as follows.

1) Assuming that signals on the point of real sound source and virtual sound source are in a phase, that calculate a phase on the receiving point. It is converted to the delayed Dij. And the direction of the calculated sound by equation 1.

2) The angle of incident in Figure 2 rotates every 10 degrees until 360 degrees, and calculates 36 directions of sound to calculate the difference of the directions given by real sound source only, is regarded as the directions of the error.

3) Averaging of the errors of 36 directions given in each frequency, is regarded as the error in the frequency. Averaging error is calculated every 1Hz, from 10Hz to 404Hz.

The error characteristic is uneven on both sides when there is an influence of the reflection. If the sound is
in a continual frequency such as a wide band of noise, the averaging directions can approach true direction of sound exactly. In the case where the noise is a complex pure tone, it is difficult to reduce the error by simple averaging. That is to say, in the case where the noise is a complex pure tone, there is a unique problem. Therefore, we study reflection influence outdoors.

**EXPERIMENT OUTDOORS**

We experiment outside in order to be able to estimate all directions of practical sound in the case where there is a reflection. In a case that there is the influence of the reflection, angles of the direction of sound sources are uneven. There is the error of the direction of the sound in each frequency. We investigate the area that estimates the sound source, which is given by the angles of the direction of the sound sources from several receiving points. The measuring point is from the courtyard shown in figure 4. There are 7 receiving points which give the area by plotting the width of the angle of direction of sound source around the intersection points by each measuring point. Sound sources 1 is a complex pure tone in 120Hz, 240Hz, Sound source 2 is the complex pure tone in 120.5Hz, 241Hz, and 361.5Hz. These sound sources generate at the same time. Sound volume is setting 90dB at 1m from the loudspeaker. And, analysis frequency is extended to 20dB and 20data.

**RESULTS**

Table 1 shows the angle width at each measuring point. The angle width is given by the difference of the maximum angle and the minimum angle for every measuring point. This measurement changes from 3.6° to 36.0°. The direction lines of sound is shown in Figure 5. In the case where the area is made by the wide angle lines (example, measuring point 3 of source 1, measuring point 4 of source 2), extent of the area of sound source increases. On the other hand, the estimated area of the sound source by selecting narrow angle lines are shown in figure 6. As a result, the estimated area of the sound is near the point of real sound source. And, as for the selecting measuring points, we are able to realize identification of sound source practically (example, measuring point 1 and measuring point 7 of source 1, measuring point 3 and point 6 of source 2).

**CONCLUSIONS**

We theoretically showed error the behavior of the direction estimated of sound source by the reflection. The experiment outside showed that the estimated area of the sound is near the point of real sound source according to selected measuring points.

**REFERENCES**

1. I. Yamada, K. Makino, N. Hayashi, "Identification of Noise Source by Detecting 3 Dimensional Changes of Sound Incidence Angle", TECHNICAL REPORT OF IEICE. EA90-21


![Figure 4. Position of Measurement outdoors](image)

![Table 1. Angle for every receiving points, unit[°]](image)

![Figure 5. Direction lines to sound source](image)

![Figure 6. Estimated area by two receiving points](image)
Supersonic acoustic intensity for arbitrarily shaped sources

Marcelo B.S. Magalhães, Roberto A. Tenenbaum, Moysés Zindeluk

Acoustics and Vibration Lab., Federal University of Rio de Janeiro, Rio de Janeiro, Brazil
mbsmagalhaes@hotmail.com

Acoustic intensity may, in some situations, not give a precise insight about how much energy is in fact carried into the far field of a vibrating source. A new parameter was then proposed, the so-called supersonic acoustic intensity, which takes into account only the intensity generated by components having a smaller wave number than the acoustic one (supersonic ones). To the authors’ knowledge, this new parameter is, up to now, defined only for sources with separable geometries.

This work presents a new approach, based on the Boundary Elements Method and Singular Value Decomposition (SVD) to compute the supersonic acoustic intensity for arbitrarily shaped sources. A numerical example is also shown and briefly discussed.

INTRODUCTION

Acoustic intensity is the most usually employed quantity when sound radiation analysis is concerned. Numerical evaluation or experimental measurement of its real part gives a portrait of how much energy is carried by propagation, while its imaginary part shows how much energy is conveyed by hydrodynamic phenomena. A trap is however hidden behind the frequencies lower than a local coincidence criterion. Williams [1], based on a spatial Fourier analysis of the sound field generated by cylindrical radiators, shows that the interaction between components with wavelength shorter than the acoustic one (subsonic components) can indeed generate a not negligible amount of active nearfield intensity.

In this article, the outline of an approach developed in order to evaluate the supersonic acoustical intensity for arbitrarily shaped sources is described. The basic idea is to discard the lowest singular values present in the SVD of a radiation operator obtained via the Boundary Elements Method, which are well-known [3, 4, 5] to be related to very inefficient radiation modes. The pressure obtained by the application of the modified radiation operator to a filtered version of the normal velocity is termed as supersonic pressure, and the real part of its product by the conjugate of the velocity is then the supersonic acoustic intensity.

THEORETICAL BACKGROUND

Although one cannot, in a strict point of view, speak of evanescent waves when non-separable geometries are dealt with, the use of a numerical approach such as the boundary elements method (BEM) together with the singular value decomposition (SVD), well know to yield a quite suitable modal-like decomposition for radiation problems, can lead to interesting results.

After discretizing the source’s surface and creating a grid of points where the sound pressure is to be evaluated, the direct BEM yields a relationship between pressure and velocity on the source’s surface, $\mathbf{H}\mathbf{p} = \hat{\mathbf{G}}\hat{\mathbf{v}}$, where $\mathbf{H}$ and $\mathbf{G}$ are matrices containing integrals of the product of the free-field Green’s function and its normal derivative by the interpolation function used in the boundary mesh. In order to finally evaluate the sound pressure at the field point the relation $\mathbf{p}_\Omega = \mathcal{H}\hat{\mathbf{p}} - \mathcal{G}\hat{\mathbf{v}}$ is used, where the subscript $\Omega$ denotes a point in the field and the matrices $\mathcal{H}$ and $\mathcal{G}$ are matrices containing integrals of the Green functions relating points on the source’s boundary to points in the field. Simple algebraic manipulation with the relations above yields $\mathbf{p}_\Omega = \mathcal{R}\hat{\mathbf{v}}$, with $\mathcal{R} = \mathcal{G} - \mathcal{H}\mathbf{H}^{-1}\mathcal{G}$ being the radiation operator. It is well-known that the SVD yields a set of basis vector for both the surface velocity and the field pressure. These vectors can be thought of as a kind of modal basis for those two quantities. Moreover, the singular values are closely related to the radiation efficiency of each of the surface velocity modes. As already pointed out by some authors [3, 4, 5], the singular values can be used to set a distinction between those modes which radiate efficiently and those with low radiation efficiency. A possible use of the technique is to obtain a modified acoustic intensity, taking into account only those modes having radiation efficiency greater than a minimum threshold. In order to get this new quantity, we write the velocity at the source’s boundary as, $\hat{\mathbf{v}}(s) = \sum_{i=1}^{L} c_i \mathbf{V}_i$, where $\hat{\mathbf{v}}_n$ is a vector containing the nodal values of velocity on the mesh nodes, $\mathbf{V}_i$ are the basis vectors for the velocity and $c_i$ are weighting coefficients, which act as a kind of "modal participation" function and $L < N$ is the number of velocity modes retained, with $N$ being the total number of modes. $\hat{\mathbf{v}}(s)$ is a modified velocity containing only the modes which radiate efficiently and can, therefore, be thought of as an analog of the supersonic velocity [6, 1, 2] for arbitrarily shaped sources.
FIGURE 1. Active and supersonic intensity on the cylindrical source. As can be seen, supersonic intensity eliminates the regions of recirculation, showing only the regions which really contribute to far field.

Using this modified velocity, one is able to obtain an analog of the supersonic pressure on the source’s surface, that is \( \hat{p}(s) = H^{-1}G\hat{v}(s) \). The supersonic acoustic intensity on the source’s surface can then be obtained by the product \( \hat{I}(s) = \frac{1}{2} \text{Re}(\hat{p}(s)\hat{v}(s)^*) \).

NUMERICAL EXAMPLE

A simple numerical example was carried out so as to clarify the usefulness of supersonic acoustic intensity. A flat-capped, unbaffled cylinder with unitary diameter and aspect ratio \( r = 5 \), vibrating in a \((1,6)\)-mode is analyzed. The source was discretized in 820 quadrangular linear boundary elements. The supersonic acoustic intensity was evaluated at \( \eta = 0.2 \), with \( \eta \) being the ratio of the frequency in use to the coincidence frequency where the structural wavelength equals the acoustic one, which for this example is around 232 Hz. The tested mode has then a small radiation efficiency, with the peaks and valleys of the surface velocity canceling a considerable amount of the sound pressure generated by each other. On the regions near the flat caps, however, this cancellation is less complete due to the presence of the boundary. These regions become then a kind of hot spots, from which energy, instead of being reabsorbed by cancellation, is transported to the farfield, having therefore significant influence on noise generation. The following example shows that on the ”hot spots” the supersonic acoustic is significantly higher than on other regions.

CONCLUSIONS

The numerical example shows that for modes below coincidence, the conventional active intensity is masked by local energy recirculation. Supersonic intensity, however, could clearly locate the regions where the cancellation is weaker, and which in fact inject energy in the far field. It is worthwhile noticing that for frequencies above coincidence both active and supersonic intensity become equivalent, since energy recirculation due to cancellation becomes less significant as the acoustical wavelength decreases.

REFERENCES

A New Method for Active Absorption of Sound Based on the Kirchhoff-Helmholtz Integral Equation

S. Takane and T. Sone

Department of Electronics and Information Systems, Faculty of Systems Science and Technology, Akita Prefectural University, 84–4 Ebinokuchi, Tsuchiya, Honjo, 015–0055 Japan

In this paper, a new method for achieving the active suppression of reflected sound waves in sound field is proposed with its theoretical basis on Kirchhoff-Helmholtz integral equation. In order to actively suppress and absorb the unwanted reflections in the target sound field, the proposed method execute the control so as to equalize the target sound field to that of the free field. For this purpose, Kirchhoff-Helmholtz boundary integral equation gives an effective means. The validity of this method is demonstrated via a computer simulation.

INTRODUCTION

Approaches to active suppression of reflected sound waves in sound field can be divided into two classes. One of these approaches is intended to make the control sound source absorb the energy of the reflected sound waves. Numerous methods have been proposed based on this approach [1, 2, 3, 4]. However, the performance of these methods is limited depending on their scale such as the number of control sources and the absorbing area covered by the control sources. The other approach is based on equalizing the target sound field to that of the free field. Ise showed the principle of sound field control based on equalizing the target sound field to that of the free field. Considering the active suppression of the reflected sound waves in rooms with the arbitrary position of the primary sound source, this may be the serious drawback.

In this paper, a new method for active equalization of the target sound field to that of the free field is introduced based on the Kirchhoff-Helmholtz integral equation and the inverse system theory[5]. This principle, however, is not valid if the primary sound source is inside the target sound field. Considering the active suppression of the reflected sound waves in rooms with the arbitrary position of the primary sound source, this may be the serious drawback.

In this paper, a new method for active equalizaton of the target sound field to that of the free field is introduced based on the Kirchhoff-Helmholtz integral equation and the theory of the inverse system. Our proposed method can be applied to the control of sound field with the primary sound source inside it.

OUTLINE OF THE PRINCIPLE

Kirchhoff-Helmholtz Integral Equation in the Free Field

Sound pressure \( P(r, \omega) \) at the point \( r \) in the arbitrary sound field can be expressed by the sum of the incident component \( P_i(r, \omega) \) and the reflected one \( P_k(r, \omega) \) as:

\[
P(r, \omega) = P_i(r, \omega) + P_k(r, \omega),
\]

(1)

where \( \omega \) denotes the angular frequency of sound. As shown in Figure 1, the sound field in the enclosed region \( \Omega \) with its boundary \( \Gamma \) can be represented by the Kirchhoff-Helmholtz boundary integral equation as follows:

\[
P(r_p, \omega) = P_i(r_p, \omega) + \int_{\Gamma(r_q)} \left\{ G_F(r_p, r_q, \omega) \frac{\partial P(r_q, \omega)}{\partial n_q} - P(r_q, \omega) \frac{\partial G_F(r_p, r_q, \omega)}{\partial n_q} \right\} d\Gamma,
\]

(2)

where \( r_p \) is the arbitrary point in \( \Omega \), and \( G_F(r_p, r_q, \omega) \) represents the Green function of the Helmholtz equation. If any acoustical obstacle does not exist, i.e., the sound field is the free field, the following equation is obviously valid.

\[
P(r_p, \omega) = P_i(r_p, \omega).
\]

Hence the following equation is obtained for the arbitrary point \( r_p \in \Omega \).

\[
\int_{\Gamma(r_q)} \left\{ G_F(r_p, r_q, \omega) \frac{\partial P(r_q, \omega)}{\partial n_q} - P(r_q, \omega) \frac{\partial G_F(r_p, r_q, \omega)}{\partial n_q} \right\} d\Gamma = 0.
\]

(3)
Proposed Control System

The control method proposed here is illustrated in Figure 2. In the target sound field with its boundary represented by the dashed line, Eq. (2) can be applied. Since the sound field depicted in Figure 2 is not a free field, the integration term in Eq. (2) is not zero. In order to produce the free field in \( \Omega \), the secondary (control) sources are located outside \( \Omega \), and the output from them are controlled so as to realize the condition represented by Eq. (3). When the control sources are driven, sound pressure at the arbitrary point \( r \) is expressed as follows:

\[
P(r, \omega) = P_l(r, \omega) + P_h(r, \omega) + \sum_{j=1}^{M} H_j(r, \omega)X_j(\omega),
\]

where \( M \) indicates the number of control sources, \( H_j(r, \omega) \) denotes the transfer function from the \( j \)-th control source to the point \( r \), and \( X_j(\omega) \) is the input to the \( j \)-th control source. The reference points (corresponding to \( r_p \) in Eq. (3)) are chosen in \( \Omega \), and the left-hand side of Eq. (3) is computed at each reference point. \( X_j(\omega) \) is determined by substituting Eq. (4) to Eq. (3).

COMPUTER SIMULATION

As an example demonstrating the validity of the proposed method, computer simulation was executed for two dimensional sound field. As shown in Figure 3, reflection from a single rigid plane at \( y = 0 \) is suppressed by using the proposed method. *, \( \bigcirc \), and + in Figure 3 denote the sensor point, control source, and reference point, respectively. The area to be controlled is a rectangle with the size of \( 2.6 \times 1.8 \) [m\(^2\)], and the frequency is 200 Hz. Figure 3(a) indicates the equal sound pressure level contour when the control is off, whereas Figure 3(b) shows that when the control is on. It is clearly seen in Figure 3(b) that the sound pressure distribution produced by the sound wave radiated from the primary source is revived in the target sound field.

SUMMARY

In this paper, a new method for the active suppression of reflected sound waves in a sound field is proposed. The principle of the method was outlined, and the validity was demonstrated via computer simulation.

ACKNOWLEDGMENTS  This research was partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research, 11650367, 1999-2000, and was also executed under Cooperative Research Projects of Res. Inst. Electr. Comm., Tohoku University (H12/A10).

REFERENCES

Comparison of Two Semi-analytical Methods of Sound Radiation Calculation

B. Zeitler, M. Möser and R. Piscoya

Institut für Technische Akustik, Technische Universität Berlin, Sekr. TA7, Einsteinufer 25, 10587 Berlin, Germany
berndt.zeitler@tu-berlin.de

The broader goal of this study is to find the limitations of the multipole method for calculating the radiated power of discontinuous structures. In other words in which frequency band and under which conditions an error criteria can be satisfied. In the presented work an infinitely long radiator with the cross section of a slice of pie is chosen as the discontinuous subject of investigation. Thus far two distinctively different methods where used to calculate radiation results, the Wave-Approach and the Multipole-Approach. In cases where results are similar or equal, the probability of results being correct are high. Parametric studies where carried out to provide good agreement in large bands. Some such results will be discussed in form of modal radiation efficiency and directivity pattern under strongly radiating and weakly radiating circumstances.

DESCRIPTION OF THE PROBLEM AND APPROACHES

An infinitely long radiator with the cross section of a slice of pie, radius $b$ and enclosing the angle $\varphi_0$, has a given normal velocity distribution on the surface. The sides are rigid ($v_n = 0$) for $r \leq b$ and $\varphi = -\varphi_0/2 \wedge \varphi = \varphi_0/2$. The normal velocity distribution on the surface at radius $r = b$ for $-\varphi_0/2 \leq \varphi \leq \varphi_0/2$ will be treated in modal components, which can be described as $v_m = v_0 \sin m\varphi (\varphi/\varphi_0)$ or as cosine components. The results presented here will follow the former modal distribution and be presented as radiated sound, or more exactly by the modal radiation efficiency and the directivity pattern.

Wave Approach

In this case the problem is solved by splitting the field into two regions (See Fig. 1a). The pressure and normal velocity in regions 1 ($0 \leq r \leq b; \varphi_0 \leq \varphi \leq 2\pi - \varphi_0$) and 2 ($r \geq b; 0 \leq \varphi \leq 2\pi$) are described as follows:

\begin{equation}
    p_1 = \sum_{n=0}^{N} A_n J_0^2(kr) \cos \left(\frac{n\pi}{\alpha} (\varphi - \frac{\varphi_0}{2})\right) \tag{1}
\end{equation}

\begin{equation}
    p_2 = \sum_{m=0}^{M} H_1^2 (kr) (B_m \cos (m\varphi) + C_m \sin (m\varphi)) \tag{2}
\end{equation}

\begin{equation}
    v_{1,2} = \frac{j}{\omega \rho} \hat{n} \nabla p_{1,2}, \quad \text{with} \quad \alpha = 2\pi - \varphi_0 \tag{3}
\end{equation}

The ansatz for the pressure $p_1$ was chosen to meet the boundary condition ($v = 0$) on the rigid edges. The other boundary conditions, a transition of pressure and normal particle velocity between the two regions 1 and 2 ($p_1 = p_2$ and $v_1 = v_2$ for $r = b \wedge \frac{\varphi_0}{2} \leq \varphi \leq 2\pi - \frac{\varphi_0}{2}$) and satisfying the velocity distribution on the "membrane" ($v_2 = v_m$) lead to the coefficients. With these, the modal radiation efficiency as well as the directivity pattern can easily calculated.

\begin{figure}
(a) \hspace{1cm} (b)
\end{figure}

FIGURE 1. Description of methods. a) Wave-Approach, b) Multipole-Approach. ● - boundary conditions, ○ - multipoles

Multipole Approach

In this approach the sound radiation of the vibrating structure is calculated by approximating the normal velocity on the surface with the velocity produced by multipole sources placed within the slice (see Fig. 1b). Their amplitudes are gained by minimizing the surface error by method of least squares. The sound radiation is then the superposition of the pressures generated by all such multipole sources.

\begin{equation}
    p = \sum_{m=1}^{M} A_m H_1^2 (kr_m) + B_m H_2^2 (kr_m) \cos \varphi_m + C_m H_2^2 (kr_m) \sin \varphi_m \tag{5}
\end{equation}

\begin{equation}
    v_n = \frac{j}{\omega \rho} \hat{n} \nabla p \tag{6}
\end{equation}

In this case one monopole and two perpendicular dipoles where place in each point. Again knowing the pressure
coefficients, leads to the modal radiation efficiency and
directivity pattern.

**MODAL RADIATION EFFICIENCY**

The power radiated by the source can be quantified by a dimensionless magnitude, the modal radiation efficiency $\sigma_m$. The definition of $\sigma_m$ is given by:

$$\sigma_m = \frac{P_m}{\bar{v}^2 \ell}$$

where $P_m$ is the power radiated by mode $m$, $\bar{v}^2$ is the mean square velocity of the curved area and $\ell$ is its length, $\ell = b\varphi_0$. Two arrangements have been chosen to study the methods closer with. The modal radiation efficiency is displayed, firstly for a stronger radiator with a mode 1 velocity distribution and an opening angle $\varphi_0$ of 60° (see Fig. 2), and secondly, for a weaker radiator arrangement possessing a mode 2 velocity distribution yet a larger opening angle of 90° (see Fig. 3). As can be seen in Figure 2 and 3, the results of both approaches only differ by a few decibels within the range of $0.1 \leq b/\lambda \leq 10$, which increases the probability of the results being correct. That the results are as expected also leads to that presumption. The mode 2 arrangement radiates much less relatively than the other for $b/\lambda \leq 1$, which can be explained by the dipole similarities the radiator possesses for small $b/\lambda$; two sources of opposite amplitude positioned near one another. This can also be seen clearly in the modal directivity pattern.

**FIGURE 2.** Modal radiation efficiency for mode 1 with nodes at edges and $\varphi_0 = 60^\circ$. + Wave-Approach, o Multipole-Approach.

**MODAL DIRECTIVITY PATTERN**

The modal directivity pattern $R(\varphi)$ is the angle dependent component of the intensity in the far field $R(\varphi) \sim |p|^2$. Figure 4b depicts the modal directivity pattern in decibel of the structure with an opening angle $\varphi_0 = 90^\circ$, $b/\lambda = 4.0$ and a mode 1 normal velocity distribution. The dipole effect of zero pressure in the far field for $\varphi = 0^\circ$ and $\varphi = 180^\circ$ can be seen well. Here too, the results of both methods coincide strongly. Finally, a case with a high modal radiation efficiency shall be looked at, namely with an opening angle $\varphi_0 = 60^\circ$, $b/\lambda = 4$ and a mode 1 velocity distribution. Both methods calculate a very bundled radiation pattern.

**FIGURE 3.** Modal radiation efficiency for mode 2 with nodes at edges and $\varphi_0 = 90^\circ$. + Wave-Approach, o Multipole-Approach.

**FIGURE 4.** Modal directivity pattern in dB. a) Mode 1 with nodes at edges, $b/\lambda = 0.2$ and $\varphi_0 = 60^\circ$ b) Mode 2 with nodes at edges, $b/\lambda = 4$ and $\varphi_0 = 90^\circ$. + Wave-Approach, o Multipole-Approach.

**CONCLUSION AND OUTLOOK**

In many cases the results of both the Wave-Approach and Multipole-Approach were very similar for modal radiation efficiency and directivity pattern. Probably more interesting to investigate, are the cases where the results do differ, and why, and which is more probable to be correct. The next steps are to add one or two more methods (BEM and IFEM), fine tune the old ones, and of course to validate the methods with experimental measurements.
FUNDAMENTAL CONDITIONS FOR
THE ANALYSIS

The medium of the ultrasonic wave propagation is assumed to be isotropic and non-absorptive. The ultrasonic wave being investigated obeys Huygens principle. The ultrasonic diffraction is computed by the repeated trapezoidal rule principally with automatic interval halving. The computed quantities are both the relative amplitude and the phase delay of the sound pressure, the sound particle velocity and the specific acoustic impedance.

CIRCULAR FLAT TRANSDUCER AND SQUARE FLAT TRANSDUCER

Historically circular flat transducers are widely used in every application of ultrasonic wave. But, the feature of the ultrasonic diffraction by the circular flat transmitting transducer indicates much abnormality especially when the ultrasonic field on the central axis of the transmitting transducer is investigated. If square flat transducers are used in the applications, the abnormality of the ultrasonic field on the central axis of the transmitting transducer disappears. Above it, the ultrasonic field by the square flat transducers can be computed by numerical quadruple integration including the case when the receiving transducer being tilted out of parallel condition from the transmitting transducer.

PROCEDURE OF COMPUTATION

The ultrasonic field by the circular flat transducers is computed using the axial symmetry of the ultrasonic field. The ultrasonic field by the circular flat transducer at a point in the medium is derived by numerical single integration. The mean value of it is derived by numerical double integration. For the purpose to compute the ultrasonic field by the square flat transducers, axial symmetry cannot be used as it does not exist principally. So, the ultrasonic field by the square flat transducer at a point in the medium is derived by numerical double integration. And, the mean value of it is derived by numerical quadruple integration. Using the numerical quadruple integration, mean sound pressure, mean sound particle velocity and mean specific acoustic impedance are all computed and compared. One example of the computation is shown in Fig. 1 – Fig. 3, each of which shows the mean sound pressure distribution, the distribution of the mean sound particle velocity or the distribution of the mean specific acoustic impedance, respectively. The tilted angle of the investigated receiving transducer against the transmitting transducer is from zero to \( \pi/8 \). The value of \( a/\lambda \) is 2.5, and the distance of the receiving transducer from the transmitting transducer is \( (a) \).

CONCLUSION

Using the numerical integration, various features of the ultrasonic field by the square flat transducers are being computed. They are compared together with the results by the circular flat transducers. The typical abnormality on the central axis of the circular flat transducer disappears being investigated with the square flat transducers. The ultrasonic field by the square flat transducers is computed including the case when the receiving transducer tilts against the transmitting transducer.
FIGURE 1. Mean sound pressure ($a/\lambda=2.5$, $z=a$).

FIGURE 2. Mean sound particle velocity ($a/\lambda=2.5$, $z=a$).

FIGURE 3. Mean specific acoustic impedance ($a/\lambda=2.5$, $z=a$).
A New Ray Travel-time Calculation Approach Based on Perturbation Method

S. Piao and S.-e Yang

Ocean Engineering College, Northwestern Polytechnical University, Xi’an, 710072, P.R. China

In this paper, a new ray travel-time calculation method based on perturbation method is presented. Numerical studies show that the new algorithm can give enough accuracy for the ocean acoustic tomography based on ray travel-time inversion scheme. Compared with the conventional methods, the new algorithm is indeed extremely efficient.

INTRODUCTION

The ability to efficiently predict the travel time of sound wave for a particular ocean environment forms the basis of ocean acoustic tomography. Though ray theory has been widely used for studying tomographic pulse propagation for its simplicity and fast computational speed, since sound travel time along the ray paths must be calculated many times for ocean acoustic tomography, it is still necessary to improve the calculation speed.

RAY TRAVEL TIME CALCULATION USING PERTURBATION METHOD

In the ray theory, the ray travel time \( \tau(\vec{X}) \) obeys the eikonal equation

\[ | \nabla \tau |^2 = c^{-2}(\vec{X}) \]  (1)

in which \( \vec{X}(s) \) is the ray trajectory, and the eikonal equation is readily solved to yield

\[ \tau(s) = \tau(0) + \int_0^s \frac{1}{c(s')} ds' \]  (2)

The integral in Eq. (2) gives the travel time along the ray, so from physical point of view the phase of the wave is simply delayed in accordance with its travel time.

A modified ray method derived from perturbation theory was presented by Finn B. Jensen[1] to treat the volume attenuation in the ocean. Because a sound-speed profile can be described as a sum of only two or three EOFs (Empirical orthogonal functions) \( f_i(z_i), \ldots, f_K(z_i) \) with high accuracy, we can similarly write the perturbed sound speed as

\[ c(z) = \bar{c}(z) + \alpha_1 f_1(z) + \alpha_2 f_2(z) + \cdots \]  (3)

and seek a solution in the form

\[ \tau = \tau_0 + \alpha_1 \tau_1 + \alpha_2 \tau_2 + \cdots \]  (4)

The zero degree approximation is just the eikonal equation for average sound-speed, which can be solved as described previously. \( \tau_1(s) \) and \( \tau_2(s) \) are immediately solved using the ray paths for the average sound-speed profile.

\[ \tau_1(s) = -\int_0^s \frac{f_1(s')}{c_0^2(s')} ds' \]  (5)

\[ \tau_2(s) = -\int_0^s \frac{f_2(s')}{c_0^2(s')} ds' \]  (6)

Then, the solution of the eikonal equation for the sound-speed profile \( c(z) \) can be written as

\[ \tau = \tau_0 + \alpha_1 \tau_1(s) + \alpha_2 \tau_2(s) \]  (7)

This means that the ray travel time can be obtained by adding some modification to the ray travel time \( \tau_0 \) for the average sound speed-profile, provide the sound-speed profile can be expanded by EOFs. As the average sound-speed profile can be obtained from archival data, the average ray travel time \( \tau_0 \) and the ray trajectory can be calculated previously. Then it can
be expected that the ray travel time calculation method using Eq. (7) will be indeed extremely efficient.

A NUMERICAL EXAMPLE

We test the validity of the new ray travel-time calculation method with an example involving a sound propagation problem in the deep sound channel. An average sound-speed profile and three major EOFs in the deep ocean have been given by Tolstoy and Diachok [3], which are shown in Fig. 1 and Fig. 2. In order to simplify the problem, we only consider the first two major EOF and expand the sound-speed profile as

\[ c(z) = \bar{c}(z) + \alpha_1 f_1(z) + \alpha_2 f_2(z) \]  

(8)

Dependence of sound speed profile on the first EOF coefficient \( \alpha_1 \) is shown in figure 3. The source and the receiver are lying at 200 meters depth. The ray travel times are calculated using Eq. (7) along the rays, which shoot upward and reflect or refract above the sound-speed axis and reflect on the bottom boundary at 4500 meters depth, and finally arrive at the receivers. The distances between source and receivers are changed from 100 meters to 6000 meters. The dependence of the percentage of the relative errors for ray travel-times on coefficients \( \alpha_1 \) and \( \alpha_2 \) of EOFs are shown in Fig. 4 and Fig. 5 respectively. It is shown that the calculation error of the Eq.(7) are no more than 0.5 ms and the new approach can give enough accuracy for the ocean acoustic tomography.

REFERENCES

Method for solving acoustic boundary problems in thermo-viscous fluids

Romain Bossart\textsuperscript{a}, Nicolas Joly\textsuperscript{a}, Michel Bruneau\textsuperscript{a}

\textsuperscript{a}Laboratoire d’Acoustique UMR-CNRS 6613, Université du Maine, 72085 Le Mans cedex 09.

The present work contributes to methods for solving some classes of problems of acoustic propagation in thermoviscous fluids. The focus is here on reactive and dissipative phenomena which have to be taken into account in order to secure an accurate description of the acoustic field in small fluid-filled cavities or ducts. However, these dissipative and reactive phenomena are usually not included in existing numerical packagings since these packagings are usually not based on the whole set of equations involved. The difficulty to model such phenomena is circumvented using an original hybrid method which combines numerical solutions (obtained with a BEM code) and analytical solutions (for the field near the boundaries). A detailed application is presented to validate the process using a boundary element method.

INTRODUCTION

Acoustic propagation in unbounded domains in thermoviscous fluids involves dissipative and reactive phenomena which are characterized in the frequency domain by a complex wavenumber whose imaginary part is proportional to the viscosity coefficient and the thermal conductivity coefficient.

In bounded domains (ducts or cavities), these dissipative and reactive phenomena arise at the boundaries, due to interactions between the acoustic movement, the entropic movement (heat diffusion in the fluid), and the vortical movement (shear waves diffusion in the fluid). The amplitudes of the entropic and vortical movements decrease (according to a diffusion process) from the boundary where they are created, towards the bulk of the fluid, and die out at distances denoted by $\delta_h$ and $\delta_v$, the thicknesses of the thermal and the viscous boundary layers (the complex wavenumber which characterizes these movements has an imaginary part proportional to the square root of the thermal conductivity coefficient and the shear viscosity coefficient).

Moreover, one must distinguish the two following classes of domains. The first class consists of domains whose dimensions are much greater than the boundary layers thicknesses; the reactive and dissipative phenomena being located near the boundaries, the effects of those phenomena on the acoustic field in the bulk of the fluid can be expressed using an impedance-like boundary condition. The second class is the capillary domains, where dimensions are smaller or equal than these boundary layers thicknesses; in those domains, the entropic and vortical movements have the same order of magnitude as the acoustic movement (which provides energy), and then an accurate description of the acoustic field is requested.

The aim of this paper is to provide a method for solving the acoustic propagation problem in the three classes of domains mentioned: unbounded domains, bounded domains whose dimensions are greater than the boundary layers thicknesses, and capillary domains. We present a hybrid method based on any numerical packaging which can include both admittance boundary conditions and analytical solutions. As an example, this method is applied to calculate the acoustic pressure field in a thermoviscous fluid-filled large tube, using a classical BEM code (used in industry) and the results are compared with the corresponding analytical results [2].

ACOUSTIC FIELDS IN THERMOVISCOUS FLUIDS

The variables which describe the mechanical and thermodynamical state of the fluid are the pressure variation $p$, the particle velocity $v$, the density variation $\rho'$, the entropy variation $\sigma$ and the temperature variation $\tau$. The complete set of linear homogeneous equations governing small amplitude disturbances of the fluid (created by external acoustic sources) includes the Navier-Stokes equation, the conservation of mass equation and the Fourier equation for heat conduction.

In bounded domains (open or closed), whose dimensions are greater than the thermoviscous boundary layers thicknesses, the problem of acoustic propagation can be solved using a Helmholtz equation and an admittance-like condition on the boundaries [2]:

\[(\Delta + k_0^2) p_a = 0, \quad (\partial_n + ik\beta) p_a = 0, \quad \beta = k_0(k_0^2 + \frac{1}{k_h}) \]

where $k_0 = \sqrt{\frac{\mu}{\rho}}$, $k_h = \sqrt{\frac{\mu}{\rho_h}}$, and $\gamma = \frac{\kappa}{C_v}$ (1c) is a small admittance which is appropriate to express adequately the viscous and thermal dissipative and reactive phenomena.
processes near the rigid walls (the dissipative process in the bulk of the domain being accounted for in the complex wavenumber \( k_a \)), the factor \( k_a^2 / k_0^2 \) usually characterizing the direction of the acoustic velocities on the rigid wall. One must emphasize that the admittance \( \beta \) depends on the acoustic field that we want to calculate; then, because \( \beta \) is an unknown function, we used the following hybrid method to overlap this difficulty.

**METHOD**

First, the problem (eq. 1) is solved by setting arbitrarily the factor \( k_a^2 / k_0^2 \) to a value (for example zero). Afterwards, the solution obtained for acoustic pressure is post-processed for calculating the factor \( k_a^2 / k_0^2 \) on the rigid wall, mainly by using equation \( (k_a^2 / k_0^2) = -k_0 p_m \partial^2 p_m / \partial x^2 \).

Finally, the problem (eq. 1) is solved again, using the estimated values of factor \( k_a^2 / k_0^2 \).

It is noteworthy that the method immediately converges, mainly because the spatial distribution of the acoustic pressure outside the boundary layers does nearly not depend on the modulus of admittance \( \beta \).

Hence, an accurate description of acoustic movements and temperature variations in the boundary layers is obtained by using the numerical solution \( p_n \) calculated above; namely, for the temperature variations, the solution take the form

\[
\tau = \frac{\dot{\gamma}}{\gamma^0} p_n (1 - e^{-ik_a u}). \tag{2}
\]

**RESULTS**

In order to show the efficiency of the method when using the impedance-like function \( \beta \) an application is given in the case of a “large” rigid-walled waveguide, closed at one end by a plane piston source and at the other end by a rigid wall, “large” meaning that the transverse dimensions are much greater than the boundary layer thicknesses. The frequency range of the study lie over the first axial resonance of the waveguide in order to emphasize the role played by the dissipation process in the boundary layers. For this frequency range, the transverse dimensions are such as the field in the waveguide is a plane wave (under the first cut off frequency). All these requirements are here achieved using a waveguide 170mm long and 5mm large (square cross section) in the frequency range 900Hz to 1100Hz.

Figure 1 shows the analytical and numerical results for pressure variations computed at one end of the tube. Curves (3) and (4) show a very good agreement between exact analytical results and numerical one obtained here, emphasizing that the method presented in this paper is accurate to numerically solve problems which need to take into account dissipative processes (note that in the example chosen here, dissipation losses play an important and very sensitive role).

**REFERENCES**

Sound source separation by decorrelation of 2-point microphone signals

Y. Takahashi\(^a\), M. Toyama\(^a\), and M. Iwaki\(^b\)

\(^a\)Department of Informatics, Kogakuin University, Hachioji-shi, Tokyo, 192-0015 Japan
\(^b\)NHK Science and Technical Research Laboratories, Tokyo, Japan

A decorrelation algorithm for separation of a blind source under 2S2M conditions; that is, two sources are composed of a target and a competing source is arranged on a line that includes two microphones in free space. This decorrelation algorithm requires that the two source signals are uncorrelated and the time delay between the two-microphone signals can be estimated. The decorrelation matrix for separating source signals can then be obtained by solving a pair of simultaneous equations which require that the two short-time correlation records must be zeros. This pair of equations can be made every two frames, which are cut from the observed microphone signals. The frame-averaged solutions converge and reach the ideal solution. Computer simulations and experiments conducted in an anechic room confirmed that the proposed algorithm can successfully separate male and female voices.

**INTRODUCTION**

Extracting and separating a target signal from other competing sources is the focus of this study. Such target signal separation is a fundamental issue concerning signal processing for a wide variety of applications including recording and communication tools. Blind source separation (BSS) without knowledge of signal mixing structures is a promising technology for target signal tracking. BSS requires higher-order signal statistics than normal second-order statistics [1]. Weinstein [2], however, developed a decorrelation method (using two microphone) based on second-order statistics. Iwaki et al. [3] also recently proposed an adaptive method that finds the decorrelation matrix frame by frame according to second-order statistics. This method works well for separating signals under 2S2M conditions, where two sources and two microphones are located on a line. In the present work, the authors develop the algorithm suggested by Weinstein [2]. The proposed algorithm seeks the decorrelation matrix every two frames by solving a set of linear simultaneous equations with two unknown variables. It is shown to have a fast convergent property in getting the solutions. And it is confirmed that it works well under a similar 2S2M condition to that used by Iwaki [2].

**INTER-FLAME SIMULTANEOUS EQUATIONS FOR DECORRELATION**

We suppose that there are two uncorrelated sound sources, \( s_1 \) and \( s_2 \), under the 2S2M condition as shown in Fig. 1. Here, \( u_1 \) and \( u_2 \) denote the received signals through the microphones, where the time delay, \( \tau \), between the microphones is assumed to be compensated. The matrix containing signals \( \hat{s}_1 \) and \( \hat{s}_2 \) that must be separated from the received signals is given as

\[
\begin{bmatrix}
\hat{s}_1 \\
\hat{s}_2
\end{bmatrix} =
\begin{bmatrix}
1 & c_{12} \\
1 & c_{21}
\end{bmatrix}
\begin{bmatrix}
u_1 \\
u_2
\end{bmatrix}, \quad (1)
\]

where \( c_{12} \) and \( c_{21} \) are the separating solutions, which must satisfy the next equation,

\[
(u_1 + c_{12}u_2) \cdot (u_2 + c_{21}u_1) = 0, \quad (2)
\]

since \( \hat{s}_1 \) and \( \hat{s}_2 \) must be uncorrelated in order to be separated. Equation (2) has two unknown variables. Thus, if it takes two frames to build the simultaneous equation shown in Fig. 2, then we can get the solutions for the two unknowns \( c_{12} \) and \( c_{21} \). We can also get the frame-averaged solutions instead of solutions every two frames.
SIMULATIONS AND EXPERIMENTS

Computer simulations and experiments in an anechoic room were carried out under the 2S2M condition as shown in Fig. 1, where \( L_1 = 3m \), \( L_2 = 1m \), \( l_1 = 0.6m \), \( s_1 \) represents male speech, and \( s_2 \) denotes female speech. The ideal solutions are set to be \( c_{12} = -1.16 \) and \( c_{21} = -0.62 \). Figure 3 illustrates the frame-averaged solutions obtained by computer simulation, as a frame goes on shown in Fig. 2. Figure 3 clearly shows that the frame-averaged solutions reach the ideal solutions. Figure 4 illustrates the waveforms of the source signals \( s_1 \) and \( s_2 \), the received signals \( u_1 \) and \( u_2 \), and the separated signals \( \hat{s}_1 \) and \( \hat{s}_2 \). In both cases (simulations and experiments) we assumed the time-delay between the two microphones is known. In the anechoic-room experiments we estimated and equalized the time-delay characteristics by the cross-spectrum data of the two microphone signals (of a speech by one of the male or the female) having frequency characteristics.

SUMMARY

A BSS algorithm based on the second-order statistics under 2S2M conditions; that is, where two sources and two microphones are located on a line. The decorrelation matrix could be estimated every two frames by solving a pair of simultaneous equations stating that the separated signals must be mutually uncorrelated. Computer simulations and experiments in an anechoic room confirmed that the algorithm works well. Estimating the time-delay between the two microphones, including the frequency characteristics, and the effects of reverberation in the mixing structure on the algorithm are still under study.

REFERENCES

A new interpolation method of HRTF based on the Common Pole-Zero model

K. Watanabe\textsuperscript{a}, S. Takane\textsuperscript{b} and Y. Suzuki\textsuperscript{a}

\textsuperscript{a}Research Institute of Electrical Communication, Tohoku University, 2-1-1 Katahira, Aoba-ku, Sendai, Japan
\textsuperscript{b}Faculty of Systems Science and Technology, Akita Prefectural University, 84-4 Ebinoguchi, Tsuchiya, Honjo, Akita, Japan

A new interpolation method of HRTFs based on the Common-Acoustical-Pole and Zero (CAPZ) model \cite{1} is investigated. This method is efficient to represent a set of HRTFs because the poles are assumed to be independent of the source directions. The interpolation of HRTFs is achieved by interpolating MA coefficients of the CAPZ model. The results of interpolation of HRTFs in the horizontal plane with this method showed better accuracy overall than that with the simple linear interpolation of the impulse responses of HRTFs.

\section{INTRODUCTION}

In the control of sound localization, the synthesis of Head-Related Transfer Functions (HRTFs) is one of the most important techniques. Since HRTFs vary as a function of source location and have strong individuality, systems based on HRTF synthesis should have HRTFs for each listener covering all directions around the listener. However, it is not practical to obtain HRTFs through measurement because of the enormous amount of time and effort needed. One possible solution to this problem is to obtain HRTFs in arbitrary directions from measured HRTFs. In this paper, a new method for interpolating HRTFs in the horizontal plane is investigated. The method is based on the Common-Acoustical-Pole and Zero (CAPZ) model \cite{1}. This model is efficient, i.e., a set of HRTFs may be represented by fewer parameters (coefficients) because common poles independent of source directions are assumed.

\section{THE CAPZ MODEL}

HRTF $H(\theta, z)$ ($\theta$ denotes source direction) is modeled as follows by use of the CAPZ model:

$$H_{\text{CAPZ}}(\theta_m, z) = \frac{C z^{-Q_1} \prod_{i=1}^{Q_2} [1 - q_i(\theta_m) z^{-1}]}{\prod_{i=1}^{P} (1 - p_i z^{-1})},$$

$$= \frac{\sum_{i=0}^{Q} b_i(\theta_m) z^{-i}}{1 - \sum_{i=1}^{P} a_i z^{-i}},$$

where $p_i$ denotes the common acoustical poles and $q_i(\theta)$ denotes the $i$-th zero dependent on source direction $\theta[1]$. $P$ and $Q$ are, respectively, the order of the common poles and of the zeros, $a_i$ denotes the common autoregressive (AR) coefficients corresponding to the common acoustical poles, and $b_i(\theta)$ denotes the moving-average (MA) coefficients corresponding to the zeros. This model seems reasonable for HRTFs because common acoustical poles correspond to the resonance of the pinna and its surrounding area and may be independent of source direction.

In this paper, the HRTFs of a KEMAR dummy head measured by the Media Laboratory at MIT \cite{2} were used, in which HRTFs in 72 azimuthal directions in the horizontal plane are included. As preprocessing before interpolation, these HRTFs were divided into the following two parts \cite{3}. One is referred to as pure delay, and the other is the main response function. The length of the latter was set at 128 with a sampling frequency of 44.1 kHz. Orders of the poles and zeros of the CAPZ model were set at 16 and 40, respectively.

\section{METHOD}

\subsection{Interpolation of MA Coefficients}

By modifying Eq. (1), MA coefficients are expressed as

$$b_k(\theta_m) = h_{\text{CAPZ}}(\theta_m, k) - \sum_{i=1}^{P} a_{CI} h_{\text{CAPZ}}(\theta_m, k - i).$$

Since the common AR coefficients are independent of source direction, it is possible to interpolate HRTFs by interpolating MA coefficients. In this paper, the pure delay involved in each impulse response of HRTF was extracted before the interpolation of MA coefficients since the pure delay does not affect its spectral shape. Then a fine positioning of MA coefficients was performed before the in-
terpolation so as to maximize the following cross-correlation function:

$$C_k = \sum_{i=0}^{Q} b_i(\theta_1)b_{i+k}(\theta_2),$$

where $b_i(\theta_1)$ and $b_i(\theta_2)$ are $k$-th MA coefficients for two adjacent directions $\theta_1$ and $\theta_2$, respectively. After this treatment, MA coefficients for an arbitrary direction were obtained by linearly interpolating the corresponding coefficients.

**Condition and Evaluation Criterion**

The HRTFs were interpolated at 8 intervals. The intervals were $10^\circ$, $20^\circ$, $30^\circ$, $40^\circ$, $45^\circ$, $50^\circ$, $60^\circ$, and $90^\circ$, and the HRTF of $0^\circ$ (front) was always included. The interpolation was performed so that the total number of HRTFs would be 72 with the interval of $5^\circ$.

The performance of the interpolation was evaluated by the Spectral Distortion (SD) given by

$$SD = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(20 \log \frac{|H(\omega_k)|}{|\hat{H}(\omega_k)|} \right)^2} \text{[dB]},$$

where $|H(\omega_k)|$ is the magnitude response of the measured HRTF, $|\hat{H}(\omega_k)|$ is that of the interpolated HRTF and $N$ is the length of HRTF. The smaller SD is, the more accurate the interpolation can be regarded.

**RESULTS AND DISCUSSION**

Figure 1 shows the average SD of the interpolated HRTFs of the left ear. Results from the use of our proposed method as well as those by the simple linear interpolation of the main responses of the impulse responses [3] are shown. As the interval decreases, the accuracy of interpolation becomes better in both methods. At intervals less than or equal to $45^\circ$, the proposed method shows a slightly better accuracy than the conventional linear interpolation.

Figure 2 shows the SD for each direction at the interval of $10^\circ$. Accuracies when the source is in the contralateral directions are degraded in both methods. However, the proposed method provides better characteristics overall.

These results suggest that the separation of impulse responses into the pure delay and the main response as well as the fine positioning of the MA coefficients were effective for realizing the good performance of the proposed method. Moreover, the CAPZ model requires a significantly fewer number of parameters for the interpolation. This is also an advantage for efficient modeling of a set of HRTFs for arbitrary directions.

**CONCLUSIONS**

A new interpolation method of HRTFs based on the CAPZ model was proposed. This method showed better accuracies than that with the linear interpolation of the impulse responses. In summary, with our method, a set of HRTFs may be represented for all directions more accurately and efficiently than is possible with conventional methods.

**ACKNOWLEDGMENTS**

This research was partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research (B) (2), 10450143 and Scientific Research (C) (2), 11650367 and under Cooperative Research Projects of R. I. E. C., Tohoku University (H12/A10).

**REFERENCES**

Chosen methods of digital analyse of signals from acoustic measurement heads

L. JODLOWSKI\textsuperscript{a}, M. SZUSTAKOWSKI\textsuperscript{a}, M. PISZCZEK\textsuperscript{b}, K. FOKOW\textsuperscript{b}

\textsuperscript{a} Institute of Optoelectronics, Military University of Technology, 2 Kaliskiego Str., 00-908 Warszawa, Poland
\textsuperscript{b} Department of Electronics, Military University of Technology, 2 Kaliskiego Str., 00-908 Warszawa, Poland

Implementation of digital synthesis of probe pulses and conversion of echo from acoustic heads into digital form, gave the possibility to extract some interesting information, existing in acoustic echo reflected by real objects. Usage of filtration and convolution algorithms, or similar analysis is relatively simple and growth of computing power makes counting time short. In this work experimental results of information detection in case of long acoustic pulses: mono-frequency and with linear frequency modulation (LFM) is presented. Theoretical description, experimental results and possible ways of technical implementation will be discussed here.

QUASI-CONTINUOUS WAVE PACKET

The measurement system utilising quasi-continuous wave packet can be used to observe the changes of object’s size or its physical properties e.g. density or temperature. The basis of that method is assigning the changes of the phase of acoustic receiving signal.

Figure 1 shows the scheme of the measurement. The electric form of the transmitted signal is formed by a generator or DA conversion card (from the data stored in cash memory). After amplifying that signal to satisfy power level it is lead to a piezoelectric transducer. The acoustic signal propagates through the measurement space L and is received by an acoustic head. The positions of transmitter and receiver in investigated volume can be different. It is also possible to use only one transducer as transmitter as well as receiver. As it is shown on picture fig.2 the conversion of signal into digital form can be done using frequencies higher or lower than signal frequency own.

The minimum of error is observed when the samples are located even on one period of signal \cite{1}.

If the harmonic signal of echo is described as follows:

\[ g(t) = S + A \sin(\omega t + \varphi) \]  \hspace{1cm} (1)

then it is possible to find the parameters of signal, and the searched value of phase \(\varphi\), immediately from digital data \cite{2}.

The theoretical and experimental investigations of the measurement system working on quasi-continuous frequency 1.466 MHz were done. The receiving signal was converted into digital form by a 12-bit AD converter manufactured by AMBEX. Theoretically calculate precision of assigning the phase of harmonic signal was 0.2°. Checked on a signal produced by a generator the accuracy was 0.4° and in laboratory testing set-up 2.5°. In our experiment the measurement distance L was 300 mm and as tested medium water was used. In this case the resolution of density \(3 \times 10^{-7}\)
was obtained and detection of changes acoustic velocity was 3.7 mm/s. After some modification that method makes it possible to measure the transmit time similar to the pulse method [3].

ANALYSE OF THE PHASE MODULATED SIGNALS

Digital synthesis of transmitting signal and conversion into digital form of echo signals from acoustic heads can be done using signals with frequency and phase modulation. In more detail the signals with linear frequency modulation (LFM) and implementation convolution and wavelet analyse was investigated.

Piezoelectric transducer is driven by an electric signal with LFM modulation:

\[ N = f_0 + \frac{\Delta f}{t_i} \tag{2} \]

where \( f_0 \): beginning frequency of generator pulse, \( \Delta f \): deviation of frequency, and \( t_i \): time of pulse.

Then the signal of echo, received by acoustic head, is a sum \( s_n \) of signals reflected by object borders:

\[ S = \sum g(f) \cdot s_n(t) \tag{3} \]

where \( g(f) \) is the unknown function characterising the reflection proprieties of object and conversion efficiency of acoustic head.

After sampling, that signal can be analysed by convolution with reference function or by making a wavelet analyse. On figure 3 the results of convolution of the signal received from a real object is shown.

As send signals, packet with time delay 200 µs and changes of frequency from 1 to 2 MHz was used. The obtained resolution was comparable with frequencies about 5 MHz in pulse method. The real amplifying of the signal using convolution process was about 2000 times. Implementation of additional wavelet analyse makes the resolution 5-6 times better.

SUMMARY

The frequency analyse of acoustic echo is a basis for classification of objects. Nowadays pulse methods are used [4] but quasi-continuous or the phase modulated signals are more convenient. This work shows possibilities of different analyse techniques and it is an introduction for characterisation of objects due to its acoustic response.

The system can be applied also in the following fields of science and technology:
1. Fundamental research - measurements of phase velocity of an acoustic wave in material media (solid state, liquids and gases).
2. Detectors of geometrical, physical, and chemical magnitudes in analysis systems, automatics and control (measurements of temperature, density, concentration, phase changes of chemical reactions, measurements of objects dimensions, and level of liquids).
3. Medical diagnostics concerning resultant acoustic properties of tissues.

FIGURE3. The real signal of LFM, its convolution and details of compressed pulse.

REFERENCES
1. L. Jodlowski, M. Szustakowski, M. Piszczek: „Analysis of the measurement possibility of the long time wave bunches method for the possibility of technical application” Molecular and quantum acoustics vol.20 pp.103-108
2. M. Szustakowski, L. Jodlowski, M. Piszczek: „Analysis of measurement possibilities of acoustic wave packets for technical and medical applications” confirmed to publication in Archive of Acoustics
3. M. Szustakowski L. Jodlowski, , M. Piszczek „Investigations of medium acoustic proprieties, utilising the measurements of quasi-continuous waves phase” Proceedings of this Congress.
Blind source (signal) separation has gained much attention recently. The goal of this contribution is to optimize FastICA [2,3] algorithm for separation of segmented signals. Segmentation is necessary especially in the case of acoustic signal processing. Several FastICA parameters has to be tuned to obtain desired solution. Crosstalk suppression is used as a performance measure.

**INTRODUCTION**

The contribution addresses intensively studied problem of blind source (signal) separation [1,4]. The effort is devoted to the optimization of FastICA [2,3] algorithm for separation of segmented signals. Segmentation is required in case of acoustic signal processing. Following problems have to be analysed:

P1 FastICA parameters: type of non-linearity, initial guess of mixing matrix, deflation or symmetric approach, number of iterations for FastICA to converge to desired solution,

P2 segmentation parameters: window, minimal segment length and overlap.

**Definitions**

Scalar mixture of independent signals is defined as

\[ x = As, \]  

where \( x \) are the observed signals and \( s \) original ones.

The goal is to find a separation matrix \( W \) so that separated signals \( y \) are as close to the original as possible

\[ y = Wx \]  

Global system matrix, used for the separation quality evaluation, is given by

\[ G = WA. \]

Separation quality is measured by the performance index \( PI \) corresponding to cross-talk suppression, defined as

\[
PI = \sum_{i=1}^{n} \left[ \sum_{k=1}^{n} \frac{|g_{ik}|^2}{\max(\text{row}_k(G))} - 1 \right] + \sum_{i=1}^{n} \left[ \sum_{k=1}^{n} \frac{|g_{ki}|^2}{\max(\text{col}_k(G))} - 1 \right],
\]

where \( g_{ij} \) are the individual elements of the global system matrix, \( \text{col}_k(\cdot) \) and \( \text{row}_k(\cdot) \) denote \( k \)-th matrix column and row respectively.

**EXPERIMENTAL RESULTS**

Extensive experiments with artificial as well as real signals were performed. The minimization of the \( PI \) (4) led to the conclusions given in the next section.

First, the dependence of algorithm performance on nonlinearity and deflation/symmetric approach was tested. Cubic, hyperbolic tangents and logistic function (denoted as gaus) were used. Median, lower and upper quartile values of \( PI \) are depicted for segment length 1000 samples (Fig. 1a) and 2000 samples (Fig. 1b). The lines extending from each end of the box show the extent of the rest of the data. Outliers are data with values beyond the ends of the whiskers and are shown by plus signs. Cubic nonlinearity means nothing more than independence measure based on kurtosis, which is not much robust against outliers and has largest \( PI \) variance. Hyperbolic tangents and logistic function usually reach better results but they are more demanding to calculate. The ability to separate different signals (super-gaussian or sub-gaussian) with different probability distribution functions confirms that the chosen nonlinearity is not critical for successful separation (the necessary constraint is to estimate only super-gaussianity or sub-gaussianity correctly). All nonlinear functions are able to separate all signals. Experiments also revealed the symmetric approach reaches better performance than deflation approach.

Second, performance index dependence on segment length was evaluated. Typical result is illustrated in Fig. 2a for segment lengths increasing from 200 to 5000 samples. As one can see the critical value is about 500 samples.

Third, the \( PI \) sensitivity to window position is not critical. For the illustration, see Fig. 2b. Cubic nonlinearity, and three different segment lengths were used together with segment window moving about one sample.

Fourth, segment overlap has no significant influence on algorithm performance, therefore number of iterations were almost constant.

Similar results were obtained for different conditions and this can be summarized as follows:
CONCLUSIONS

FastICA algorithm can be used for separation of real segmented acoustic signals. If the segment length is greater than 1000 samples, no overlap, random initial matrix, and tanh nonlinearity is used then $PI$ is lower than $10^{-2}$, or equivalently, crosstalk suppression between separated signals is greater than 20 dB. The same results were obtained independently of the type of signal: synthetic signals, real speech and music.

ACKNOWLEDGMENTS

The research was supported by the grant GACR 102/96/KO87 Theory and Application of Voice Communication in Czech.

REFERENCES

Prediction of sound radiated by vibrating structures in the time domain

D. Habault, P.O. Mattei and P.J.T. Filippi

CNRS, Laboratoire de Mécanique et d’Acoustique, 13402 Marseille Cedex 20, France

The paper presents a method to evaluate the sound pressure radiated by a vibrating thin elastic structure immersed in a fluid and submitted to a transient excitation. The method consists in expressing the displacement on the structure as a series of the resonance modes of the fluid-loaded structure. These resonance modes correspond to the natural free oscillations of the fluid-loaded structure. They are therefore well suited to describe the response of the system to a transient excitation.

INTRODUCTION

Prediction of the response of a structure immersed in a fluid and submitted to a transient excitation is quite an essential problem with numerous applications, specially in the transportation domain.

The method presented here to evaluate the displacement on the structure and the sound radiated in the fluid is based on expansions of these two unknown functions in terms of the resonance modes of the coupled system structure / fluid. Because the resonance modes correspond to the free oscillations of the structure in the fluid, they provide an efficient tool to describe the behaviour of the displacement and the sound pressure in the time domain. This method can be used to describe the response of a coupled structure to various kinds of excitation (mechanical force, acoustical incident wave, turbulent boundary layer, ...).

An essential step is to compute the resonance frequencies and modes. Some useful approximations can be found in particular cases. For example, in a light fluid (such as air), a perturbation technique is used to compute the resonance frequencies and resonance modes as approximations of the eigenfrequencies and eigenmodes of the in vacuo structure. Let us point out that these eigenfrequencies and modes can be calculated by any kind of methods (finite elements in the case of a complicated structure).

A short description of the method is presented here, for the particular case of a thin elastic plate. All the details can be found in [2] and [3] for more general cases, along with applications to thin plates and shells.

SHORT DESCRIPTION OF THE METHOD

The response of a fluid-loaded thin elastic plate is described by two functions, the displacement and the sound pressure field. The equation for the displacement \( U(M,t) \) on the plate has the general form:

\[
AU(M,t) = F(M,t) + [P(M,t)] \quad \text{for } t > 0 \quad (1)
\]

\( M \) is a point on the plate \( \Sigma \). \( A \) is the differential operator for the plate. \( F \) is a mechanical excitation. \([P(M,t)]\) represents the jump of sound pressure exerted on both sides of the structure. If \( Q \) denotes a point in the fluid, the sound pressure \( p(Q,t) \) in the fluid is the solution to a wave equation:

\[
(\Delta - \frac{1}{c^2} \frac{\partial^2}{\partial t^2})p(Q,t) = S(Q,t) \quad \text{for } t > 0 \quad (2)
\]

where \( c \) is the sound speed in the fluid and \( S \) the acoustic source, if any. To these equations, conditions must be added such as:

- continuity of the accelerations on the plate;
- boundary conditions at the boundaries of the plate;
- initial conditions at time \( t = 0 \).

The two functions \( U \) and \( p \) are solutions to a coupled system of equations. However, for a baffled plate, the sound pressure can be related to the displacement by using a Rayleigh integral. Therefore, the method can be decomposed into three steps.

Resonance frequencies and resonance modes

The first step is to compute the resonance frequencies \( \omega_n \) and modes \( w_n(t) \) (displacement) of the coupled problem (structure / fluid). They correspond to the non zero solutions of the corresponding homogeneous system of equations, written in the frequency domain [2]. The resonance frequencies are complex. Their imaginary part, directly related to the energy radiated in the fluid, is always negative. This leads to an exponentially decreasing amplitude of the displacement and the sound pressure, as functions of time.

In the particular case of a light fluid, a perturbation technique can be used [3]. As a first order approximation, the resonance modes \( w_n(M) \) can be replaced by the
eigenmodes $w_n^0(M)$ of the in vacuo structure. Each resonance frequency is approximated by the sum of the corresponding eigenfrequency of the in vacuo structure and a small term which gives the damping ratio corresponding to the radiation of the energy in the fluid.

**Expression of the displacement**

The second step is to express the displacement as a series of resonance modes:

$$U(M,t) = -i \sum_{n=1}^{\infty} [\alpha_n w_n(M) \exp(-i\omega_n t) + \alpha_n^* w_n^+(M) \exp(+i\omega_n t)] \exp(-\tau_n t)$$

(3)

where $\omega_n = \tilde{\omega}_n - \tau_n$ and $f^*$ denotes the complex conjugate. The unknown coefficients $\alpha_n$ can be obtained by solving the system of equations in the time domain, with the excitation in the right-hand member. However in the case of a thin plate, they are obtained analytically [3] and in the case of a light fluid, they can be expressed by this simple first order approximation:

$$\alpha_n = \frac{\omega_n}{2} < F - \tilde{P}_h, w_n^0 > a(w_n^0, w_n^0)$$

(4)

where $P_0$ stands for the jump of the incident acoustic pressure (equal to zero if there is no acoustic source), $F$ is the Fourier transform of $F$. The scalar product $< f, g >$ of 2 real functions $f$ and $g$ is defined as $\int f(M)g(M)d\sigma(M)$. $a(w_n^0, w_n^0)$ represents a norm of the mode $w_n^0$.

**Expression of the sound pressure**

The third step is to compute the sound pressure radiated in the fluid. In the particular case of a baffled plate, the sound pressure is related to the displacement by a Rayleigh integral:

$$p(Q,t) = \nu \int \frac{\ddot{U}(M,t-R/c)}{R} d\sigma(M)$$

(5)

$\ddot{U}$ is the second derivative of $U$ with respect to $t$. $\nu$ is a coefficient related to the fluid density and $R = R(Q,M)$ is the distance between points $Q$ and $M$. With such simple formulas (4) and (5) it is therefore possible to evaluate the influence of some parameters such as the damping ratio or the excitation.

**SOME RESULTS**

Figures 1 and 2 present two examples of spectra of sound pressure radiated. They correspond to the sound pressure radiated by two clamped plates excited by the same mechanical force: impulse at $t=0$ and at a point close to a corner of the plate. Only the first 48 modes are taken into account. The observation point $Q$ is chosen at 10cm from the plate. Although the computation is made in the time domain, we present the spectra (computed by FFT) where differences appear more clearly. Both plates are made of steel, with dimensions equal to 1mx0.70m. In fig. 1, the plate has a constant thickness (h=5mm). In fig. 2, the plate has a varying thickness (h=5 to 7.5mm).

**REFERENCES**

A Unified Numerical Approach to Modeling Propagation and Scattering Phenomena in Ocean Waveguides

Natalia A. Sidorovskaia

Department of Physics, University of Louisiana at Lafayette, P.O. Box 44210, Lafayette, LA 70504-4210, USA.
E-mail: nsidorovskaia@louisiana.edu

Sound propagation simulation codes are widely used in underwater acoustics to predict the effects of propagation media on a received signal structure. The accuracy of an answer depends on how close the numerical model of a propagation channel is to a real environment. A newly developed numerical algorithm and originated normal mode range-dependent acoustical propagation code “SWAMP-V2” (Shallow Water Acoustic Modal Propagation, Version 2) are discussed. The new version of the code allows an inclusion of a multi-layered fluid-elastic sediment model and targeted for encompassing the procedure for the inclusion of an elastic scattering object as a part of propagation channel properties. The benchmarking results related to the model accuracy and computational performance are present.

INTRODUCTION

A study of an acoustic pulse structure detected by a vertical or horizontal hydrophone array plays an important role in understanding ocean environmental processes, in predicting the parameters of ocean waveguides and in detecting source and scatterer properties. There are only a few relatively simple waveguides which allow one to obtain a closed analytical form of the solution describing sound propagation in them. All other environments require a use of numerical acoustic models based on mathematical approximations. There are four major numerical solution strategies, which have been implemented in different numerical models. These include the ray, normal mode (NM), wave number integration (Fast Field Program - FFP) and parabolic equation (PE) methods. The normal mode approach, based on separation of variables, is one of the most useful tools to describe the propagation of sound in range-independent ocean waveguides. There are now several one-way and two-way coupled modal schemes implemented in computer codes for range-dependent cases. The most important and computational-time-consuming part of any NM-algorithm is a solution of the depth-separated differential equation - the modal equation. Most existing NM codes employ the Finite-Difference Methods (FDM) to solve the modal equation. This procedure could be very computationally intensive. This fact puts existing NM acoustic models far behind PE algorithms in the computational-time competition. From this point of view, a new procedure for obtaining the normal mode solution for an arbitrary vertical sound speed profile, discussed in this paper and implemented in the first version of SWAMP (Shallow Water Acoustic Modal Propagation), made an important contribution to the standard numerical modeling tools used in underwater acoustics. [1,2]

A GENERALIZED EIGENVALUE PROBLEM

The main idea of the new numerical approach is to represent the solution of the depth-separated wave equation for an arbitrary sound speed profile in a water column as a linear combination of the basis functions corresponding to the modes of an appropriate isovelocity waveguide. At this point we examine two ocean models with the same bottom properties. One model has an iso-velocity water column while the other has a variable velocity profile. The bottom properties are assumed to be identical for both cases. The solution for the depth-separated wave equation for the iso-velocity case in a water column is a set of modes, each one of which should satisfy the wave equation for the displacement potential:

$$\frac{d^2 \varphi_{0m}}{dz^2} + \left(k_0^2 - \lambda_m^0\right) \varphi_{0m} = \frac{d^2 \varphi_{0m}}{dz^2} + \sigma_m^2 \varphi_{0m} = 0,$$

$$\alpha_m^2 = k_0^2 - \lambda_m^0, \quad (1)$$

where \( k_0 = \omega / c_0 \) is the medium wavenumber at angular frequency \( \omega \) for the iso-velocity reference waveguide characterized by constant sound speed, \( c_0 \), \( \lambda_m^0 \) is the separation constant which may be also viewed as the horizontal wavenumber squared, \( \varphi_{0m} \)
are modal eigenfuntions for the iso-velocity waveguide.

The equation that governs the solution for mode $U_m$ in a water column for an arbitrary sound speed profile is [1]:

$$\frac{d^2 U_m}{dz^2} + \left\{ k^2(z) - \lambda_m^0 \right\} U_m = 0 \quad (2).$$

The terms in the last equation can be rearranged in such a way that the left-hand side of the equation contains the operator for the iso-velocity solution:

$$\frac{d^2 U_m}{dz^2} + \left\{ k_0^2 - \lambda_m^0 \right\} U_m =QU_m, \quad (3)$$

where $Q = k^2_0 - k^2(z) - \lambda_m^0 + \lambda_m^0 = q(z) + \Delta \lambda$.

We can assume completeness of the space, namely, that the eigenfunctions of the iso-velocity space span the variable velocity space. Then, in particular, within the water column we can decompose the solution for the variable velocity space as: $U_m = \sum_{j=1}^{L} a_{mj} \varphi_{0j}$.

Substituting this sum in Eq. (3), multiplying both sides by $\rho_0 \varphi_{0k}$ and integrating over only the water column depth, $h_0$, we obtain:

$$\lambda_m \sum_{j=1}^{L} a_{mj} b_{jk} = \lambda_m^0 \sum_{j=1}^{L} a_{mj} b_{jk} - \sum_{j=1}^{L} a_{mj} q_{jk} + \alpha_m^2 \sum_{j=1}^{L} a_{mj} b_{jk} - \sum_{j=1}^{L} a_{mj} \alpha_j^2 b_{jk}, \quad (4)$$

where $q_{jk} = \int_{0}^{h_0} \rho_0 q(z) \varphi_{0j}(z) \varphi_{0k}(z) dz$ and $b_{jk} = \int_{0}^{h_0} \rho_0 \varphi_{0j}(z) \varphi_{0k}(z) dz$, which can be immediately calculated using the known solutions of the iso-velocity problem. This set of equations can be reduced to a generalized eigenvalue problem written in matrix form as follows:

$$\Lambda \Phi = \Lambda^0 \Phi - \Phi + \Phi K - \Phi K, \quad (5)$$

$$K_{mj} = \alpha_m^2 \delta_{mj}, \quad \Lambda^0 + K = k_0^2 I.$$

$$\Lambda \Phi = k_0^2 \Phi - \Phi K - \Phi Q = \Phi \Lambda^0 \Phi - \Phi Q.$$

The above approach was implemented in SWAMP. The results of the SWAMP simulation of the signal propagation losses over the 50 km range in the range-dependent ocean are in a good agreement with the results obtained by using two Navy standard codes: RAM based on PE-approach (see Figure 1) and KRAKEN based on normal mode solution. [1] KRAKEN uses a finite-difference scheme for the solution of the modal equation, a scheme which is considerably slower than SWAMP.

FIGURE 1. Benchmark results of SWAMP vs. RAM.

**SUMMARY**

The new algorithm has already given the code the advantages in comparison with the models currently used by Navy as the standard propagation simulation tools. First, the modal eigenfunctions in this method can easily be transformed to a spherical representation suitable for an incorporation of the target scattering into the code. Second, the mode-coupling scheme is efficiently implemented, since transfer matrices across interfaces may be constructed analytically. This allows the code to be very competitive with the PE-models in computational speed for range-dependent environments. The introduction of a fluid-elastic layered sediment model allows SWAMP to become a flexible tool for evaluating scattering and reverberation scenarios present in experimental data. That would reduce the reported discrepancies between simulated and experimental data.

**ACKNOWLEDGMENTS**

I wish to recognize the original contribution of Dr. Michael F. Werby into the development and benchmarking of the first version of SWAMP and thank him for releasing the code for the future independent development.

**REFERENCES**

On the Behaviour of Fractional Derivative Maxwell Model

T. Pritz

Acoustics Laboratory, Szikkti Labs, 1301 Budapest, Pf. 81, Hungary

The introduction of fractional order time derivatives into the constitutive equation of the conventional Maxwell model of viscoelastic materials results in the fractional derivative Maxwell model. This model exhibits a more realistic behaviour than the conventional one does. The behaviour of the fractional derivative Maxwell model in the frequency domain is investigated in this paper. The role of the values of fractional derivatives of stress and strain in the model behaviour is studied and constraints are made on the permissible values. Two versions of the fractional derivative Maxwell model, characterized by three and four parameters, respectively, are investigated, and their behaviours are compared. It is shown that the three-parameter model has a dual nature. Furthermore, it is shown that the four-parameter model may be an effective tool to describe the viscoelastic behaviour of real fluids over a wide frequency range.

INTRODUCTION

The fractional derivative Maxwell model is derived from the conventional Maxwell model of viscoelastic materials by replacing the integer order time derivatives in the model constitutive equation with fractional order ones. The quantitative behaviour of this model is more realistic than that of the conventional Maxwell model [1,2]. The behaviour of the fractional derivative Maxwell model has been investigated in the time domain in detail[1], but such investigation has not been made in the frequency domain. The aim of this paper is to present the results of an theoretical investigation of the behaviour of the fractional derivative Maxwell model in the frequency domain. As a result of this investigation restrictions on the permissible values of the fractional order derivatives are made, and the usefulness of the model is pointed out.

THE MODEL BEHAVIOUR

The most general form of the constitutive equation for the fractional derivative Maxwell model is[1]

\[
\sigma(t) + \tau^\beta D^\beta[\sigma(t)] = G_m \tau^\alpha D^\alpha[\varepsilon(t)],
\]

where \( \sigma \) is the shear stress, \( \varepsilon \) is the strain, \( t \) is the time, \( \tau \) is the relaxation time, \( G_m \) is a material constant of modulus dimension, moreover \( 0<\alpha<1 \), \( 0<\beta<1 \) and \( D \) denotes the operator of fractional derivation defined as [3]

\[
D^\alpha[\varepsilon(t)] = \frac{1}{\Gamma(1-\alpha)} \frac{d}{dt} \int_0^t (t-\tau)^{-\alpha} \varepsilon(\tau) d\tau,
\]

in which \( \Gamma \) is the gamma function and \( \tau \) a dummy variable. Eq.(1) yields the conventional Maxwell model if \( \alpha=\beta=1 \).

In order to investigate the model behaviour in the frequency domain, the complex modulus of elasticity is derived from Eq.(1), bearing in mind that

\[
D^\alpha[\varepsilon(t)] = (j\omega)^\alpha \varepsilon(t)
\]

for harmonic vibration of \( \omega \) frequency, where \( j=\sqrt{-1} \). The complex shear modulus is

\[
G(j\omega) = G_d(j\omega) + jG_l(j\omega) = G_m \frac{(j\omega \tau_r)^\alpha}{1 + (j\omega \tau_r)^\beta},
\]

where \( G_d \) is the dynamic shear modulus, \( G_l \) is the loss modulus and \( \eta(\omega)=\frac{G_l(\omega)}{G_d(\omega)} \) is the loss factor. It can be derived from Eq.(4) that

\[
G_d(\omega) = \frac{\cos(\alpha \pi / 2) \omega_n^\alpha + \cos[(\alpha - \beta)\pi / 2] \omega_n^{\alpha+\beta}}{1 + 2 \cos(\beta \pi / 2) \omega_n^\beta + \omega_n^{2\beta}}
\]

and

\[
G_l(\omega) = \frac{\sin(\alpha \pi / 2) \omega_n^\alpha + \sin[(\alpha - \beta)\pi / 2] \omega_n^{\alpha+\beta}}{1 + 2 \cos(\beta \pi / 2) \omega_n^\beta + \omega_n^{2\beta}}
\]

where \( \omega_n=\omega \tau_r \) is the normalized frequency.

The model behaviour is investigated by analytical and numerical study of \( G_d(\omega) \), \( G_l(\omega) \) and \( \eta(\omega) \) functions. The most important conclusions can be read from Eq.(6), namely that \( \alpha\geq\beta \) must be satisfied to insure that \( G_l>0 \), i.e. the energy loss is positive for all frequencies. This conclusion is in accordance with that of the time domain investigation[1].

The frequency variations of the dynamic and loss moduli, and loss factor are shown in Fig.1a to 1d for some permissible values of \( \alpha \) and \( \beta \). It can be seen in Fig.1a that the conventional Maxwell model (\( \alpha=\beta=1 \)) has a dual nature; it exhibits fluid-like behaviour at low frequencies, and solid-like behaviour at high frequencies. Fig.1b illustrates that the fractional model...
has dual nature too if \(\alpha=\beta<1\) (three-parameter model). Nevertheless, the fluid-like behaviour of this fractional model is less realistic, and the solid-like behaviour is more realistic than those of the conventional model. In contrast to these, the fractional Maxwell model exhibits quite realistic fluid-like behaviour for all frequencies if \(\alpha=1\) and \(\beta<1\) (four-parameter model). The loss modulus has a peak if the value of \(\beta\) is close to 1 (Fig.1c), but the peak disappears by decreasing the value of \(\beta\) (Fig.1d). The frequency variations of dynamic and loss moduli seen in Figure 1c,d are in good accordance with those experienced for polymer solutions, melts and gels [2,4].

CONCLUSIONS

The behaviour of the fractional Maxwell model in the frequency domain has been investigated in this paper. It has been shown that the order of fractional derivation of strain (\(\alpha\)) must be greater than that of the stress (\(\beta\)) in order to have a physically meaningful model. Moreover, it has been shown that the fractional derivative Maxwell model with parameters of \(\alpha=1\) and \(\beta<1\) may be an effective tool to describe the viscoelastic behaviour of real fluids.

ACKNOWLEDGEMENTS

This work has been supported by the Hungarian Research Fund (OTKA) under contract T 30151.

REFERENCES

Special Wave Finite and Infinite Elements for Helmholtz Equation

R. Sugimoto, P. Bettess and O. Laghouche

School of Engineering, University of Durham, South Road, Durham DH1 3LE, U.K.

This work deals with numerical modelling using special finite and infinite elements to solve Helmholtz equation efficiently. The approach is based on the partition of unity method. The potential at each node is expanded in a discrete series of approximating plane waves propagating in different directions so that the nodal shape functions are multiplied by trigonometric functions. Because of this choice of basis function a single finite element can contain many wavelengths, unlike conventional elements. The mesh model can thus remain unchanged even though the wave number increases. This approach greatly reduces the dimension of a problem for short waves in high frequency range.

INTRODUCTION

The aim is to develop special finite elements which can contain many wavelengths to solve wave problems satisfying Helmholtz equation for short waves and in unbounded domains efficiently. The authors have developed the elements using shape functions containing trigonometric functions representing plane waves in multiple directions based on the Partition of Unity Finite Element Method by Melenk and Babuška [1]. In this paper, the formulation of the finite element model is described and then a numerical example is shown.

FORMULATION

Governing Equations and Residual Scheme

For the time-harmonic wave problem in domain $\Omega$, the velocity potential $\phi$ satisfies the Helmholtz equation

$$ (\nabla^2 + k^2)\phi = 0 \quad \text{in} \, \Omega $$

(1)

where $k$ is the wavenumber. By using the weighted residual method and introducing the boundary conditions

$$ \nabla \phi \cdot \mathbf{n} = -\mathbf{v} \quad \text{on} \, \Gamma_1 $$

(2)

$$ \nabla \phi \cdot \mathbf{n} - ik\phi = 0 \quad \text{on} \, \Gamma_2 $$

(3)

where $\mathbf{n}$ is the outward normal of boundary $\Gamma$, the weak form

$$ \int_\Omega (\nabla W \nabla \phi - k^2 W \phi) \, d\Omega - ik \int W \phi d\Gamma = - \int_{\Gamma_1} W \nabla \phi \cdot \mathbf{n} d\Gamma $$

(4)

is obtained where $W$ is the weighting function.

Finite Element Model

The domain $\Omega$ is divided into $n$-noded finite elements. The unknown function $\phi$ within each element is described by using polynomial shape functions $N_j$ and the nodal values of the potential $\phi_j$. If $\phi_j$ is approximated in terms of plane waves propagating in the direction $\theta_j$ where $l = 1, \ldots, m$ with amplitudes $A_j$, as follows,

$$ \phi_j = \sum_{l=1}^{m} A_j e^{ik(x\cos \theta_j + y\sin \theta_j)} $$

(5)

then $\phi$ is given by

$$ \phi = \sum_{j=1}^{n} N_j \phi_j = \sum_{l=1}^{m} \sum_{j=1}^{n} N_j A_j e^{ik(x\cos \theta_j + y\sin \theta_j)} $$

(6)

Let us take

$$ P_{(j+1)m+l} = N_j e^{ik(x\cos \theta_j + y\sin \theta_j)} $$

(7)

as ‘new shape functions’, then the amplitudes $A_j$ become new unknown functions. Therefore the degree of freedom of the element matrices becomes $(n \times m) \times (n \times m)$ unlike $n \times n$ in case of conventional method. However the special wave elements can be larger than the wave length, which will lead to a great reduction in the dimensions of the problem. By the Galerkin procedure, a set of discrete equations for each element is obtained as follows.

$$ [[K] - k^2[M] - ik[C]] \{A\} = \{F\} $$

(8)

where

$$ K_{rs} = \int_{\Omega_e} \nabla W_r \nabla P_s d\Omega_e $$

(9)

$$ M_{rs} = \int_{\Omega_e} W_r P_s d\Omega_e $$

(10)

$$ C_{rs} = \int_{\Gamma_e} W_r P_s d\Gamma_e $$

(11)

$$ F_r = - \int_{\Gamma_e} W_r \nabla \mathbf{v}_r d\Gamma_e $$

(12)

where $r$ and $s$ are integers equal to $1, 2, \ldots, (n \times m)$. 
Numerical Integration

To calculate the element matrices, the integrals encountered are of the form

\[ \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} f(\xi, \eta) e^{i(kx\cos\theta + y\sin\theta)} e^{i(lx\cos\theta + y\sin\theta)} \, d\xi \, d\eta \]  

(13)

where \( f(\xi, \eta) \) involves the product of the original shape functions, their derivatives, the determinant of the Jacobian and its inverse. These integrations can be evaluated by Gauss-Legendre integration, where about 10 integration points per wavelength are needed [2]. The authors have also developed a new integration scheme for the special wave elements [3][4], with which the integrations appear to be independent of wavelength so that it becomes more efficient as the wave becomes shorter.

NUMERICAL EXAMPLE

Scattering of a plane wave by a cylinder of radius \( a \) with rigid surface was solved for wavenumber \( k = 16 \).

For the finite element calculations, the cylindrical domain from \( r = a = 1 \) to \( r = 5 \) is divided into 36 sectors of 10 degrees each, and each sector is divided into 4 quadrilateral elements in radial direction. On the inner boundary \( \Gamma_1 \) the normal derivatives of scattered potential were given as boundary condition (2), and on the outer boundary \( \Gamma_2 \) plane wave damper boundary condition (3) were applied. 18 equally spaced plane waves at each node were used.

Figure 2 shows the calculated scattered potential. There are virtually no differences between the values, and sufficient accuracy was achieved although the dimensions of the largest elements are more than double the wavelength.

CONCLUSION

Finite element formulations for Helmholtz equation by using special wave elements and its application for a scattering wave problem were presented. This method allows the finite elements to contain many wavelengths per element rather than having many finite elements per wavelength, which is supposed to be a great advantage for short wave problems. Development of infinite elements with the same concept is in progress to model the unbounded domain problems more accurately.

ACKNOWLEDGMENTS

This work has been funded by the Engineering and Physical Sciences Research Council, under grants GR/K81652, GR/M41025, GR/M41018 and SF/000169.

REFERENCES

A finite element approach combined with analytical wave functions for acoustic radiation from elastic structures

H. Zimmer\textsuperscript{a}, M. Ochmann\textsuperscript{b} and Ch. Holzheuer\textsuperscript{a}

\textsuperscript{a}SFE Gesellschaft für Strukturanalyse in Forschung und Entwicklung mbH, Voltastrasse 5; D-13355 Berlin; Germany; E-mail: sfe@sfe-berlin.de; http://www.sfe-berlin.de

\textsuperscript{b}TFH Berlin - University of Applied Sciences; Department of Mathematics, Physics, and Chemistry; Luxemburger Strasse 10; D-13353 Berlin, Germany; E-mail: ochmann@tfh-berlin.de

A method will be presented for treating the coupled fluid-structure problem of an elastic structure immersed in an infinite acoustic fluid medium, developed by SFE [1] and funded by the Brite EuRam Project BRPR CT-97 0481. The vibrations of the elastic body are excited by internal forcing functions and lead to radiation of acoustic waves into the fluid. The structure and a closely surrounding part of the fluid are presented by a finite element model. At the artificial boundary of this finite volume a non-reflecting boundary condition must be introduced for applying the finite element method in the bounded domain. If the enclosing fluid has a canonical shape such as a sphere or a spheroid, this boundary condition can be given explicitly by an analytical expansion of spherical or spheroidal wave functions. For a compact, sphere-like structure a surrounding sphere is chosen. For structures of large aspect ratio, it is recommended to use better suited domains. For example, a circumscribing cylinder can be used instead of spheroids leading to complicated spheroid functions. The capacity of the method will be demonstrated by applying it to the calculation of the radiation from an elastic thin-walled spherical shell and from a complex cavity-like structure.

\section*{THEORY}

The sound radiation from a vibrating elastic structure into an infinite, three-dimensional fluid is considered. Following [2], the finite element method is used for calculating the vibrations of the elastic body \( B \) together with a part of the fluid which closely surrounds the body. If the surface \( S_1 \) of the enclosing part of the fluid is chosen to be spherical, the fluid pressure outside of \( S_1 \) can be expanded into spherical wave functions ([2-3])

\[ \psi_{\text{oc}}(x) = \Gamma_{mn} h_n^{(1)}(kr) P_{n}^m(\cos \vartheta) \left\{ \cos m \varphi \over \sin m \varphi \right\}, \]  

where the \( P_n^m \) are the associated Legendre polynomials. Here, spherical coordinates are introduced by \( x = (r \sin \vartheta \cos \varphi, r \sin \vartheta \sin \varphi, r \cos \vartheta)^T \), where \( T \) denotes transposition. The superscript \( c \) (or \( s \)) indicates that the cosine (or sine) is used. The cylindrical functions \( h_n^{(1)} \) are the spherical Hankel functions of the first kind. The normalizing factors \( \Gamma_{mn} \) are chosen in such a way that the corresponding spherical harmonics

\[ \gamma_{\text{oc}}(x) = \Gamma_{mn} P_{n}^m(\cos \vartheta) \left\{ \cos m \varphi \over \sin m \varphi \right\} \]  

are orthonormal with respect to the integration over the unit sphere [3]. The coupling between the FE-equations for the structure and the acoustic fluid loading caused by the sound pressure \( p \) is performed by transforming the finite element formulation on the surface to a functional basis consisting of a finite number \( L \) of spherical harmonics Eq. (2). Details can be found in [1,2]. Often, \( L \) is equal to the number of nodal points \( N \) at the artificial surface \( S_1 \). However, it is also possible to choose \( L < N \). For example, a pulsating elastic sphere can be simulated by only one spherical wave function corresponding to a monopole.

\begin{figure}[h]
   \centering
   \includegraphics[width=\textwidth]{sphere.png}
   \caption{FE model of an elastic sphere embedded in a fluid sphere.}
\end{figure}
NUMERICAL EXAMPLES

Radiation from an elastic sphere

Figure 1 shows the FE-model of an elastic thin-walled spherical shell embedded in a surrounding fluid sphere. If the spherical shell is uniformly driven with a spherically symmetric, time-harmonic pressure load, a closed form solution can be obtained by adding the structural impedance for the undamped spherical shell (without fluid loading) and the impedance of the surrounding fluid.

A frequency sweep from 20 Hz to 200 Hz is shown in Figure 2 for different boundary conditions at $S_1$. Clearly, the most important case is the totally absorbing boundary condition simulated by the use of the spherical wave functions. For this case, the analytical solution and our combined approach (= SFE radiation, SFE AKUSMOD in Fig. 2) give nearly identical results. Details about SFE AKUSMOD can be found in [4]. Sometimes, it is suggested to use the normal impedance $Z = \rho c$ ($\rho$ = fluid density, $c$ = speed of sound) at $S_1$ for approximating an absorbing boundary condition. However, such a choice leads to reflections at $S_1$ and to significant differences to the analytical, non-reflecting solution. Clearly, a rigid boundary condition gives strong reflections, too.

Radiation from a complex cavity into a half-space

Using the combined approach, it is also possible to calculate the sound field radiated into a half-space from extremely complicated structures like the box with external cavity shown in Fig. 3. The influence of the rigid plane, on which the structure is sitting, is taken into account by using only spherical wave functions with corresponding symmetry properties. In Fig. 4, the measured and calculated sound pressure level radiated into the half-space is shown at a certain fixed position in the cavity on the right hand sight. Both curves agree well, especially at the resonance at 120 Hz.

REFERENCES

1. H. Zimmer, Internal report Brite EuRam BRPR CT-97 0481 (Overall Interior Noise Design Tool for Road Vehicles “ONDROVE”)
Direct Computation of the Sound Radiated by a Temporal Mixing Layer

V. Fortuné, E. Lamballais and Y. Gervais

Laboratoire d’Études Aérodynamiques, Université de Poitiers, 40 av. Recteur Pineau, 86022 Poitiers, France

In this study, sound from a temporally evolving mixing layer is computed directly by DNS on a computational domain which includes both aerodynamic and acoustic fields. To ensure a high precision of the computed acoustic field, we use a numerical code based on very accurate schemes. Two- and three-dimensional simulations of mixing layers are performed, for various Reynolds numbers. The temporal approach allows to identify the acoustic radiation from each phase of the mixing layer transition. At first, results show that the vortex pairing process generates the largest acoustic radiation, in all cases considered. Secondly, we observe that the intensity of sound emission increases with the Reynolds number. Finally, 3D effects are found to give rise to an overall reduction of the acoustic radiation from the mixing layer.

INTRODUCTION

Direct Numerical Simulation (DNS) is an helpful tool to investigate physical mechanisms of noise generation in turbulent shear flows. Improvement of such a knowledge is of interest in aeroacoustics for developing hybrid approaches of sound prediction and for working out noise reduction strategies. In the present work, sound from temporally evolving mixing layer is computed directly by DNS on a computational domain which includes the aerodynamic and acoustic fields. The first section describes briefly the simulation techniques. The numerical code is based on very accurate schemes to ensure a very precise direct computation of sound. Then results of two-dimensional (2D) and three-dimensional (3D) simulations are presented. The acoustic radiation of each phase of the mixing layer transition is investigated and the Reynolds number effects are analyzed. Finally, specific effects of 3D motions in the flow on sound emissions are discussed.

SIMULATION TECHNIQUES

In the present work, we are interested in the temporal evolution of a plane mixing layer, which is formed between two streams of initial velocities $U_1$ and $U_2$. The two streams have the same mean temperatures and densities ($T_1 = T_2, \rho_1 = \rho_2$). The numerical code solves the compressible Navier-Stokes equations, on a computational domain which is large enough to give access to the far acoustic field. We use sixth-order compact schemes [1] to estimate the spatial derivatives and time advancement is made by a fourth-order Runge Kutta scheme. Periodic boundary conditions are imposed in the streamwise $x$ and spanwise $z$ directions, while a low order non-reflecting boundary condition [2] is used in the transverse $y$ direction.

The initial mean velocity is given by an hyperbolic-tangent profile. An incompressible, sinusoidal disturbance field ($\tilde{u}, \tilde{v}$) is added to the mean velocity profile to initiate the transition process. These 2D fluctuations contain the fundamental and two sub-harmonic disturbances. In 3D-simulations, random incompressible velocity fluctuations ($\tilde{u'}, \tilde{v'}, \tilde{w'}$), with a prescribed velocity spectrum, are also added. The Reynolds number of the flow is based on the initial vorticity thickness $\delta_\omega$ ($Re = \rho_2(U_1 - U_2)\delta_\omega/\mu$). The Prandtl number is defined by $Pr = \mu c_p/\kappa = 0.7$ and the Mach number by $M = (U_1 - U_2)/c_2$. Throughout this paper, we use the normalisation $(U_1 - U_2)/2 = 1, \delta_\omega = 1, \rho_2 = 1$ and $c_0 = 1$.

DIRECT COMPUTATIONS OF SOUND

2D and 3D simulations of mixing layers are performed for different values of Mach and Reynolds numbers [3, 4]. At first, the parameters of the simulations presented here are precised. The box length in the streamwise direction is chosen to match the development of four primary structures: $L_x = 4\lambda_h = 30.7$, where $\lambda_h$ is the most unstable wavelength given by the linear inviscid stability theory [5]. For simplicity, we impose $L_z = L_y$, in 3D simulations. The box length in the transverse direction is chosen in order to give access to the acoustic far field [3], with a common length $L_y = 60$. The grid resolutions are $(n_x, n_y) = (257, 321)$, and $(n_x, n_y, n_z) = (150, 257, 150)$, in 2D and 3D simulations respectively.

Now the temporal evolution of the flow structure is briefly described. In all cases, the time evolution of the vortical field (not shown here, see [3, 4]) shows the beginning of the transition process with a roll-up phase, leading to the formation of four Kelvin-Helmholtz vortices. This stage is followed by two successive vortex pairing processes. All these observations are in agreement with results from previous similar studies [6, 8]. In addition, we notice that strong 3D motions appear after the first vortex pairing.

Acoustic radiation from mixing layer

This section deals with the acoustic radiation from a temporal mixing layer. We consider the acoustic intensity $I_{ac}$, defined at the bottom boundary $y = -L_y/2$ as the product of the mean acoustic velocity and pressure $I_{ac} = (\overline{p} - p_0)(v)$, where $\overline{}$ represents a quantity averaged over the periodic $x$ and $z$ directions. As already noted by previous authors [6, 7], this intensity based on mean quantities is relevant to analyse the acoustic data from temporal simulations in terms of far field (see also [3, 4] for details).

Figure 1 shows the time evolution of acoustic intensity generated by a $M = 0.8$ and $Re = 400$ mixing layer, from 2D simulation. Time scale is shifted in order to take the propagation delay for an acoustic wave to travel across the domain ($L_x = t - L_y/2c_0$) into account.
Reynolds number effects

Now attention is focused on Reynolds number effects on sound emission from mixing layers. Figure 2 compares time evolutions of $I_{ac}$ for different values of Reynolds number. First, the time histories of $I_{ac}$ are qualitatively similar for the three $Re$ cases considered. Particularly, note that sound emissions resulting from the second vortex pairing are the noisiest ones, in all $Re$ cases. Secondly, we observe that peak amplitudes increase with the Reynolds number. Another point is that the relaxation peak amplitude becomes comparable to the compression peak amplitude, during the second vortex pairing for the $Re = 800$ case. Hence present results highlight that acoustic radiation is very sensitive to Reynolds number variations.

Thus we can identify directly the acoustic contribution of each stage of the transition. We observe a first peak associated with the roll-up process, followed by two peaks of larger amplitude associated with each pairing process. The analysis of time history of the pressure (not presented here) shows that these two peaks result from successive compression/relaxation waves during the pairing process. Notice that the intensity radiated during the first vortex pairing is lower than the intensity radiated during the second one. This last behaviour conforms to results from previous DNS of spatially evolving mixing layer [8].

3D effects

In this section, we compare acoustic radiations obtained from 2D and 3D simulations. Figure 3 shows the time evolution of $I_{ac}$ for $M = 0.8$ mixing layer cases. For one, we note that the intensities are identical for 2D and 3D cases during the roll-up phase. For another, the intensity of acoustic emissions associated with pairing processes are strongly reduced in 3D case, especially during the second vortex pairing. We outline that the intensity decrease corresponds to the occurrence of 3D motions in the mixing layer, during the first pairing. Finally, contrary to the 2D case, the first pairing is found to be the noisiest event in 3D case.

CONCLUSIONS

In this paper, we presented direct simulations of temporal mixing layer, for which the flow is assumed to be periodic in the streamwise direction. This temporal approach allows us to carry out direct 3D computations of sound for a moderate computational cost. Present results point out the ability of this idealized model to investigate physical mechanisms of sound generation by turbulent flows. In addition, the acoustic database obtained in this work could be very useful to develop and to validate hybrid approaches of sound prediction.

ACKNOWLEDGMENTS

Present calculations were carried out on the NEC-SX5 supercomputer at the Institut du Développement et des Ressources en Informatique Scientifique, to which we are thankful.

REFERENCES

Signal saturation and its effects on time domain parameters

Lorenzo Cioni

Laboratorio di Linguistica, Scuola Normale Superiore, Pisa, Italy

The present paper examines the effects of saturation on the time domain parameters that can be extracted from recordings of speech. The kinds of saturation we are going to discuss in the paper include clipping, zeroing and two’s complement whereas the parameters we will take into considerations include root mean square (RMS, as correlated with loudness), zero crossing rate (ZCR), auto correlation (as a tool to extract F0 from speech samples), voice onset time (VOT) and rise time (RT).

INTRODUCTION

Saturation is a physical phenomenon that occurs whenever a quantity exceeds the maximum value that can be managed on the physical system holding that quantity. In case of speech analysis systems saturation can occur either during the recording phase, during the acquisition phase or within the analysis software owing to wrong operations on the acquired signal. In all such cases saturation is a non linearity superimposed on the signal that reflects in a distortion of the signal itself [1, 3]. In many cases there is no way to correct such distortion [2] (owing to the unavailability of either the speakers or the tape recordings) and so what is needed is a way to evaluate the degree of saturation of a signal and the correctness of the (time domain) parameters that can be extracted from such a signal.

THREE MODELS OF SATURATION

Saturation can be modeled with three kinds of non linearity [2, 3]: clipping, zeroing and two’s complement (eqs. (1), (2) and (3)).

Such models can be described respectively by the following relations (where \( x = x(nT) \) and \( y = y(nT) \) are respectively the input and output signal of the non linear system, \( T \) is the sampling period and the values \( x_0, y_0 \) are the absolute values that define the range of linearity):

(1) \( y = k x [u(x+x_0) – u(x-x_0)] + y_0 [u(x-x_0) – u(-x-x_0)] \)
with \( u(x) = 1 \) if \( x \geq 0 \) and \( u(x) = 0 \) if \( x < 0 \) while \( k = y_0 / x_0 \)
(usually we have \( k = 1 \));

(2) \( y = k x [u(x+x_0) – u(x-x_0)] \)
and

(3) \( y = k (x-jx_0) \) if \( (j-1)x_0 \leq x \leq (j+1)x_0 \)
where \( j = 2m, m \in \mathbb{Z} \). Equation (3) can be written as:

\[ (3') y = [kx [u(x+x_0) – u(x-x_0)]]_{T'} \]
where \( T' = 2x_0 \) is the period of the function.

THE TIME DOMAIN PARAMETERS

A speech signal can be described both in the time domain and in the frequency domain with the use of parameters extracted from the signal itself. As to the time domain parameters in this paper we will limit our analysis to RMS, ZCR, autocorrelation, VOT and RT. Such parameters are correlated with physical quantities and can provide segmental cues for the analysis of speech [4].

RMS [1, 4] is a measure of the energy of a speech signal and when applied to successive windows gives a measure of its changes in amplitude (and so in loudness) over time. It depends on the sum of the squared values of the samples averaged over the length \( N \) of a window.

ZCR [4] depends on the number of times a signal crosses the zero line and gives us a measure of the dominant frequency in a signal. It can be used in differentiating between voiced and unvoiced sounds and, together with RMS, can be used to make a simple speech/no speech distinction. Autocorrelation [4] can be used to extract the pitch from a signal by comparing the signal with a delayed version of the same signal. VOT [4] is a measure of the duration of the burst in case of stops followed by a vowel. Last but not least RT [4] can be used as a cue for the distinction between the affricate and fricative sounds and depends on the time interval over which a maximum amplitude is reached.

AN EXPERIMENTAL SETTING

In order to examine the effects of saturation over speech signals (and the derived time domain parameters) an experimental setting has been conceived [2] (Figure 1). According to such a setting
speech signals are acquired in digital format so that streams of samples \( x(nT) \) are available. Such streams are then passed through the non linear system \((NL)\) characterized by relations (1), (2) and (3’) so to produce the new streams \( y(nT) \).

The effects of saturation over the signals can be examined by analyzing the signals \( x(nT), y(nT) \) and the “error signal” \( \varepsilon(nT) = x(nT) - y(nT) \) (Diff) and the analysis can be carried out by varying the values \( x_0 \) and \( y_0 \) so to simulate lower (with higher values) or higher (with lower values) degrees of saturation. The main aim of the analysis is the definition of criteria that, when applied to a generic signal, allow us to understand if the signal has suffered saturation, which kind of saturation and which was the degree of saturation (the famous “how and how much”).

FIGURE 1. Experimental setting, symbols explained in the text.

The programs used during the acquisition, saturation and analysis phases are Goldwave™ (NL and Diff) and Xwaves™.

**INFLUENCE OF SATURATION ON TIME DOMAIN PARAMETERS**

The three kinds of saturation (1), (2) and (3’) turn into a modification of the waveform of the speech sounds so that in all the cases the shape of the waveform is changed since some of the samples are substituted either with clipping value or with zeroes or gets their sign switched [2]. The extent of the saturation, in the model of figure 1, is determined by the values \( x_0 \) and \( y_0 \) since they determine, for the same signal \( x(nT) \), the energy and the length of the energy periods of the signal \( \varepsilon(nT) \), i.e. the length of the time intervals over which \( \varepsilon(nT) \) shows a non null amplitude.

The following table 1 summarizes the qualitative connections among the three kinds of saturation and the time domain parameters. Column headings are the three kinds of saturation presented in the paper and row headings are the time domain parameters we consider as more significant in the case of time domain analysis of speech.

<table>
<thead>
<tr>
<th></th>
<th>Clipping</th>
<th>Zeroing</th>
<th>Two’s complement</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMS</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>ZCR</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>AC</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>VOT</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>RT</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

A “-” sign means that a certain kind of saturation exerts (almost) no influence over the corresponding time domain parameter whereas a “+” sign means that the higher is a certain kind of saturation the stronger is the effect on the corresponding time domain parameter. It is obvious that since RMS, ZCR and AC are evaluated as a function of the length \( N \) of a window the shorter the more accurate are the graph of the parameters versus (discrete) time \( nT \) and so the more evident is the influence of saturation on those parameters.

We note that the use of the signals \( x(nT), y(nT) \) and \( \varepsilon(nT) \) is conceivable only within a theoretical model. In case of real life signals the use of such model should allow us to understand if a given signal has suffered saturation, which kind of saturation and at which degree and all this by simply examining the graphs of the waveforms and of the time domain parameters extracted by the signals themselves.

**CONCLUSIONS**

The analysis presented in this paper has been kept to an introductory level but further (and more accurate) details will be provided during the presentation. The topic of the paper represents one of the current research topics of the author and is going to be the subject of further investigations both on the theoretical and on the empirical ground.

**REFERENCES**

Modifying the frequency content of speech signals: a new approach

G. Evangelista\textsuperscript{a,b}, S. Cavaliere\textsuperscript{a}

\textsuperscript{a}Department of Physical Sciences, Naples University “Federico II”, 80126 Naples, Italy
\textsuperscript{b}Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

In this paper the authors experiment an already established tool for the modification of the frequency content of signals, which makes use of a recently introduced extension of Laguerre Transform allowing for time-varying shaping of the warping law. In the paper some applications of the transform are proposed and others are outlined. The first one allows the modification of the pitch of a speech signal, in order to change its intonation. An arbitrary intonation law, when properly detected, may be compensated for, obtaining a perfectly "flat" pitch contour, while, in turn, properly acting on a speech segment, any desired intonation law may be readily added. A second application concerns analysis: on the regularized speech signal separation of the excitation pulse from resonance is much more effective, allowing improved speech modeling. Finally other possible applications are outlined.

FREQUENCY WARPING VIA THE LAGUERRE TRANSFORM

Frequency warping by means of a transform related to the Laguerre transform was introduced by Oppenheim, Johnson and Steiglitz in 1971 [1]. In their pioneering paper the transformation was used to allow for improved frequency resolution in specified frequency regions, particularly in the context of speech processing. Starting from the above idea, the authors of this paper generalized the transformation to provide an entire new range of applications by using it in conjunction with an actual DSP paradigm, mainly that of using signal adapted transforms in order to detect special features of the signal under analysis [2].

Furthermore, by cascading the Laguerre transform to other orthogonal or biorthogonal transforms, mainly the Wavelet Transform and the Pitch Synchronous or Comb Wavelet Transform, the authors introduced an entire new range of useful transformation showing very interesting features, particularly in the field of acoustical DSP [3]. Finally the authors defined a set of biorthogonal sequences implementing a time-varying version of frequency warping [3]. By projecting a signal over this set a second signal is obtained, whose frequency characteristics are locally frequency warped version of the source signal: the warping law is, in this general case, time-varying, therefore it can be adapted to the signal. Even though the transform is computationally intensive, it can be computed by means of ordinary DSP operations, mainly time-reversal and digital filtering [1,2]. Finally in a recent work one of the authors [4] introduced a short-time version of the transform that allows for real-time operation with short delay and reduced computation.

FIGURE 1. Spectrogram of vowel /a/. Intonation is flat: the constant pitch feature may be verified in higher order partials.

FIGURE 2. Spectrogram of vowel /a/ after the transformation. The added ascending intonation law may be seen in the higher formants.
This feature enables modification of the frequency content of a signal in a time-varying fashion following a programmed frequency law. For example, given a frequency swept signal we may detect its frequency law and use time-varying frequency warping to reduce it to a constant pitch version, which may be easily analyzed or coded. Vice-versa, starting from a constant pitch signal, a desired frequency law may be superimposed, adding, for example, vibrato-like effects in the case of musical signal [3].

**MODIFYING INTONATION**

The above transformation will be shown in action by means of a first example: if we are willing to impart a desired intonation to vocal sounds (fig.1) what we can do is to modify the pitch of the signal over time.

![Figure 3](image_url)

**FIGURE 3.** In this case the added intonation law is descending.

This can be done by means of frequency warping: we have simply to draw the desired frequency law, possibly obtained from a spoken vowel with the requested intonation, thus recovering the proper sequence of parameters to be used in the transform (for the final result see spectrograms in fig.2 and 3).

**IMPROVED FORMANT ANALYSIS**

The proposed transformation may also be used for the purpose of analysis of spoken vowels. As shown in fig 4 referred to a vocal sound synthesized by means of a multi-resonant filter, detection of formants as well as detection of formant bandwidths is much more robust when analysis is carried on a properly regularized version of the signal, showing constant pitch. In fig.4 we have superimposed the frequency response of the AR coefficients of LPC analysis in subsequent time intervals, in the region of the first formant. Regularization was obtained as shown above, by means of time-varying frequency warping. In the same figure, to the right we show the same response obtained from the source signal before its regularization.

**OTHER APPLICATIONS**

A further application regards coding: regularizing the pitch of the vocal segments of speech allows us reducing the spoken vowel to a perfectly periodic part plus a specific law for the pitch contour. The former can simply be coded by the waveform of a single period, while the latter can be coded by the pitch contour itself, possibly including an additional random component for improved speech quality.

Another application of the proposed transformation is that of denoising. The key idea is that of regularizing the signal, reverting it to its perfectly periodic version. Slow random pitch variations as well as intonation or modulation may be thus eliminated from the signal. On this ‘periodic’ version of the signal, denoising, by means of a simple low pass filters as well as by means of wavelet techniques will be more effective since the signal will be confined to a narrower frequency region.

**REFERENCES**


"GAMAK"; Information System for the Designs with Low Noise Factor

Babalik F.C. (1), Cavdar K. (1)

(1) Uludağ Üniversitesi Mühendislik Mimarlık Fakültesi, Makina Mühendisliği Bölümü, 16059 Bursa, Turkey

Technical structure working with low noise is the main request in the comfort characteristics. In order to compensate this deficiency an information system called "GAMAK" which has parts of the basic knowledge about the acoustics, the principles of acoustics, step by step analytical solutions and the constructive solution samples, has been prepared by us. One of the important characteristics of the system is that the design engineer or the firm may add their experience to our database system. This information system gives utilisation to the designers without experience and prevents them giving wrong decisions.

DESIGN OF THE LOW NOISE PRODUCTS

When an engineer is asked about a technical problem, generally in order to find a technical solution, which is called a constructive process, he will put his whole creative mental activities based on his knowledge of basic sciences, his engineering knowledge and his professional experience. The most important step of creating a new technical product is called “construction”.

The solution of old days, which was assumed too good, can not be accepted as good today. In order to make machines more humane (ergonomic) “construction of machines with low noise” comes before. Especially in automotive and machine production sector, studies in this field are denser. The main responsibility for the acoustic behavior of machinery belongs to the constructor. If the constructor who knows much about the machinery acoustic is directed properly, he will be able to design machinery with lower noise.

The formation “knowledge system to support the constructor for the construction of low noise machinery” will be obviously helpful for the constructor who will be able to reach the available data easily. Such as, if the constructor is directed to construction with low noise, previously tested, appropriate confirmed solutions can be advised and as a result the construction time saving can be achieved. But it must be kept in mind that this sort of information should be changed by time.

The purpose to investigate lie construction of low noise machinery methodically: to make sure that approximation is done step by step and at each step to give the constructor necessary information on machinery acoustic and during the time period. First-Aid support should be given at each step. Only, like this, the noise problem can be solved on a large scale at conceptual-design step.

A construction of low noised productions takes place on many standards and statues. An international technique standard (ISO/TR 11688) presents a general method to low noised product construction and it mentions number of mechanism of noise origination. It includes some hindrance precaution, which is possible.

The Support Area of the Information System "GAMAK"

The techniques, which are used in order to solve noise problem, are collected in three main groups.

1. Noise problem is solved on its source
2. If not, the noisy part is insulated from outside.
3. If not, measures are taken on people who can be damaged, as a last way. Thus, health risk is tried to reduce at minimum level.

The studies, which the noise neutralizes on construction phase, contain the first group. Constructive studies, which are named “primal precaution”, are most effective solution and they bring a most appropriate conclusion financially.

All the elements, which produce functions on machines, interact each other. A noise of machine body which occurs at a component (vibration) is transferred to other components, which are in touch. In the same way, with impression air noises, which echo from surfaces, vibrated surfaces may cause to secondary noise of a machine body. For example, it is quite difficult to characterize noisy behavior of a component on a machine tool, so noisy behavior of the whole of machine and bases on certain rules; even it is impossible.

After the classic engineering education, it is very hard for a young engineer or for experienced constructor knowledge to have later on. It will cause a waste of
time and labor. Because, a knowledge system which has low noised machine construction guides and supports, it will be subsidiary of constructor. Moreover, all the low noised construction directives will provide safer, stronger and longer life. Also these directives are suggestions positive on construction.

On reducing noise, cost of solution will determine the direction. Expensive solutions can make the machine quite but they cannot be sold. That's why the constructor must discover the most solution which has most acceptable cost. Reducing noise level in the entire noise source is an illogical solution and its cost is also very high. On practice, reducing noise level of noise source doesn't cause reduction on total noise level of machine. Moreover, the machine can begin to work much more noisy always special situations. Because of this, the component which will be taken measure on constructions must be selected by constructor more carefully.

**CONSTRUCTION OF THE GAMAK**

GAMAK is one of the information sources, which supports the constructor. The technical support is given to constructor which contains four main title.

1. The Theoretical Acoustic Knowledge part which in theoretical knowledge on noised-machine acoustic area is given.
2. Acoustic Principles part. This part facilitates understanding of noise event by the constructor and this part maintains explanatory examples.
3. Step by step to solution part. This part guides step by step and supports the constructor with examples during from describing the period between the task of machine acoustic and the manufacturing of the product.
4. Constructive Solution Search part. This part guides the constructor to solution, which the constructor searches in large solution source. If there is no searching solution, it presents solutions. It allows to the constructor to add his own solutions.

Main program of the knowledge system had been prepared in DELPHI computer program language. Preface page and some sample pages from knowledge bank is given in supplement section. Web browser had been used to offer the knowledge system.

Those informations may be offered to give an idea about GAMAK:

placed 60 Mbytes in hard disc

placed 30,6 Mbytes for AcusticWEB
GAMAK.exe is 634 Kbytes
In 158 web pages 342 color picture
In solution database 51 constructive solutions

**CONCLUSION**

In this study, second opinion was admitted and an approach has been improved to attain low noise machine designs that are known a little by a number of designers. This approach includes the definition of problem, the test of solution and last decision.

Main approach of this study is to remove small problem, after the problem is divided into parts. On every step of program required theoretical knowledge has been given to a designer. Thus, the approach has been extended. The designer has been benefited from an extensive database system that had been prepared about machine acoustic.

In this system, in addition to theory of acoustic prepared practical solutions in field of machine acoustic are offered. Thus, solution has been facilitated for the designer. Therefore, it is taught that solution ability of the system will be better than other expert systems that exist in literature. Human is still at the center of the system and solution that is obtained is related to experience of the designer. As a result, it has been taught that an obtained approach will be enough to solve problems of the field of machine acoustic, but that the addition of information and examples to the system will be unlimited.

**REFERENCES**

Acquisition of Acoustic Signals Assisted by Recurrent Neural Networks

A. Czyzewski, R. Krolikowski

Sound & Vision Engineering Department, Technical University of Gdansk, 80-952 Gdansk, Poland

The issue of localisation of sound sources for videoconferencing is addressed in the paper, where a new method for estimating speaker locations is introduced. It is based on exploitation of temporal relationships between signals received by an array of microphones, and thereby recurrent neural networks are employed. Additionally, a parametrisation of the time-domain audio signals prior to the neural processing is performed. Some of the results of the experiments are briefly presented in the paper.

INTRODUCTION

Localisation of sound sources is vital to contemporary teleconferencing systems. Such a localisation improves considerably the efficiency of the source signal acquisition, since it reduces influences of other sources on the chosen one, improves the signal-to-noise ratio and in result influences the overall efficiency of noise reduction & dereverberation algorithms. Neural networks were already applied to purposes of sound localisation, however these attempts were based on feedforward structures [1]. The feedforward neural networks do not offer such feasibility as recurrent ones do, especially in the field of time series modelling [2]. Meanwhile, the physiopsychological theories explain human perception of the sound source position basing on temporal relationships [3]. Therefore RNNs were put to use in the approach proposed in this paper.

GENERAL SCHEME OF SOUND SOURCE LOCALISATION METHOD

The way of exploitation of RNNs with the purpose of sound localisation is presented in Fig. 1.

PARAMETRISATION OF AUDIO

The proposed parametrisation is based on the correlation between time-domain audio in channels. The vital information on the temporal relationship between the signals: $x_i(m)$ in the $i$-th channel and $x_j(m)$ in the $j$-th one can be obtained by taking into account the max values of the correlation coefficient $\rho_{ij}(\Delta)$ and the respective lag $\Delta_{ij}$ between these signals as follows:

$$\rho_{ij} = \rho_{ij}^{\text{max}}(\Delta) = \max_{L+1, \ldots, L-1} \{\rho_{ij}(\Delta)\}$$

$$\Delta_{ij} = \arg \rho_{ij}^{\text{max}}(\Delta)$$

which are computed in the $L$-point analysis window. Thus, the vector of parameters is composed of the
pairs \( (\rho_{ij}, \Delta_{ij}) \) for a given combination of channels. In the case of the presented approach, only combinations between opposite microphones are considered, which results in 8-element vector of parameters.

**RNN FOR SOUND LOCALISATION**

In a given moment of time \( t \), the multichannel signal is described by the 8-element vector \( u(t) \). In order to get a time series of such vectors, the parametrisation occurs over last \( N \) different points in time, which provides the \( N \times 8 \)-dimensional input vector \( x(t) \).

The employed RNN (Fig. 2), trained by means of the error backpropagation algorithm [4], is composed of:
- **Input Layer**, consisting of \( N \times 8 \) units and a bias,
- **Hidden Layer**, consisting of \( M \) neurons,
- **Context Layer**, consisting of \( M \) units which outputs are delayed by a single cycle,
- **Output Layer**, consisting of \( K \) neurons, determined by the number of directions of sound sources.

Figure 2. General Structure of RNN

**EXPERIMENTS**

The experiments were carried out in an acoustically adopted chamber. The array was fixed 1.58 m from the floor. There was one male speaker, distanced 1.5 m from the array, and used to be positioned in 5 places representing the sound directivity from -30° to +30° every 15°. 8 mono tracks (16 bit/sample, 48 kHz) were recorded simultaneously. As far as the correlation parameters are considered, the max value of the lag \( \Delta^{\text{max}} \) is strictly related to the time delay between those signals received by the opposite microphones, and thereby was set to 50. Besides, 3 values for the length of time sequences \( N \) were tested: \( N \in \{1,2,3\} \). In turn, the number of hidden neurons in RNNs was arbitrary set to 10, whereas the output layer was composed of 5 units. The training and testing vectors were different and selected randomly. The number of the vectors per class was equal to 100, which yielded totally 500 vectors for a training- and testing phase. In order to obtain statistically valid results, computations were repeated 10 times per a given survey.

The experiments revealed that the total efficiency of the sound localisation was equal to 94.7%. In turn, in Tab. 1 the results for each direction are assembled.

<table>
<thead>
<tr>
<th>Direction</th>
<th>( N = 1 )</th>
<th>( N = 2 )</th>
<th>( N = 3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>-30°</td>
<td>94.6 %</td>
<td>96.1 %</td>
<td>94.1 %</td>
</tr>
<tr>
<td>-15°</td>
<td>93.8 %</td>
<td>94.7 %</td>
<td>95.0 %</td>
</tr>
<tr>
<td>0°</td>
<td>95.6 %</td>
<td>96.4 %</td>
<td>92.1 %</td>
</tr>
<tr>
<td>+15°</td>
<td>91.3 %</td>
<td>94.4 %</td>
<td>94.3 %</td>
</tr>
<tr>
<td>+30°</td>
<td>96.0 %</td>
<td>95.9 %</td>
<td>96.2 %</td>
</tr>
</tbody>
</table>

**SUMMARY**

The results of the experiments on employment of RNNs to localisation of sound sources suggest that these neural structures can be efficient in such tasks. The obtained and presented results are encouraging for further more detailed experiments.

**ACKNOWLEDGEMENTS**

The research is sponsored by the Committee for Scientific Research, Warsaw, Poland. Grant No. 8 T11D 00218.

**REFERENCES**

Creation and Properties of Constant Envelope Sweeps for Transfer Function Measurements

Massarani, P.M, Müller, S.

Laboratório de Ensaios Acústicos (LAENA), INMETRO, Av. N. Sra. das Graças, 50 - Xerém - 25250-020 - Brazil

Pseudorandom noise has been used predominantly for transfer function measurements in the past decade. However, sweeps are far better suited for this purpose due to the possibility to entirely reject all harmonic distortion products from the recovered impulse response. In acoustical measurements, much more power can thus be fed to the loudspeaker and a significantly higher S/N will be obtained. As with pseudo-random noise such as MLS, an emphasis matched to the prevailing ambient noise should be given to the excitation sweep. This work shows how an arbitrary magnitude spectrum can be used as template to create sweeps which almost maintain a constant temporal envelope. Their properties will be shown using simple means such as the group delay display along with time-frequency tools.

INTRODUCTION

Recently, a new method of transfer function measurements using swept sine signals as excitation [1] has been presented. Compared to using pseudorandom noise signals, the sweeps have some distinct characteristics that bear some crucial advantages. Mainly, the measurements have higher immunity against distortion and time variance that could occur in acoustic systems.

Most of the advantages originate from the time-frequency behaviour of the sweeps. By properly controlling the sweep rate, it is possible to generate excitation signals with low crest factor, but also with arbitrary spectral distribution. The transfer function of the system under test is retrieved by rearranging the time-frequency location of the response. The reference [1] shows in detail some successful use of the technique in acoustic measurement.

This work presents some aspects concerning the creation of the constant amplitude sweep signals and show how the loudspeaker distortion can be separated from the impulse response measurement in a room. To illustrate the process, time-frequency descriptions of the sweeps, such as group delay and spectrogram, are being used.

CREATING SWEEPS

The signals presented in [1] are sines sweeping continuously from low to high frequencies, in a finite time. The sweep monocomponent nature permits a direct linking between the time and frequency domains. By controlling the sweep rate in a frequency-dependent fashion, an infinite number of different sweep signals can be created for acoustic measurements. Here, input signals for power amplifiers and loudspeakers with minimum crest factor are considered, but also giving an appropriated spectral coloration to keep a good S/N ratio over a broadband range. The loudspeaker response and the ambient noise spectrum demand an appropriate frequency emphasis.

Three parameters can be used for the generation of sweeps: the instantaneous amplitude, the spectral amplitude and the sweep rate. The instantaneous amplitude has to be constant, \( A(t) = A \), to achieve low crest factors. A desirable arbitrary spectral amplitude, \( B(\omega) \), will depend from the sweep rate, that can be represented by the group delay \( g(\omega) = -d\psi(\omega)/d\omega \), where \( \psi(\omega) \) is the spectral phase. The excitation signal of the method in [1] is synthesised from the inverse Fourier transform of \( B(\omega)e^{i\psi(\omega)} \).

To determine the sweep group delay from the conditions \( A(t) = A \) and \( B(\omega) \), the energy balance in a small frequency interval \( d\omega \) has to be kept, i.e.

\[
2A^2\{g(\omega + d\omega) - g(\omega)\} = B(\omega)^2d\omega,
\]

giving the following equation

\[
\frac{dg(\omega)}{d\omega} = \frac{B(\omega)^2}{A^2}.
\]  \hspace{1cm} (1)

The spectral phase is obtained by integration. By the eq. (1), it can be shown that around the more emphasised frequencies, the group delay slope must be steeper which lets the instantaneous frequency rise slower.

The practical signals have finite energy limited in a duration \( G \) and bandwidth \( F \). For a power spectrum density constant (white), \( B(\omega) = B \), the group delay is a straight line with \( dg(\omega)/d\omega = G/F \), for the energy balance \( E = A^2G = B^2F \).

Figure 1 shows the spectral amplitudes, (above) and the group delays (below) of a constant envelope sweep with 20 dB emphasis at low frequencies and 10 dB of high frequency emphasis, compared with a
linear sweep with same duration and bandwidth. The crest factors are 3.4 dB for the linear sweep and 3.9 dB for the sweep with emphasis.

**SEGREGATING THE DISTORTIONS**

After the system response to the sweep has been obtained, the transfer function is found by convoluting it with a reference signal. According to the method in [1], this is achieved by multiplying the captured response by an inverse version of the excitation sweep spectra, or \( 1/B(\omega)e^{i\psi(\omega)} \). The best way to get this reference spectrum, is to connect the input and output of the measurement device to include the amplifier and A/D-D/A converter responses.

The reference group delay is \(-g(\omega)\), corresponding to negative time values. At the time domain representation, the reference energy appears at the right side as time aliasing effect. By applying the reference, the sweep formerly spread out in time is compressed, yielding the impulse response of the system under test. As the reverberation time at low frequencies of acoustic responses is longer, the sweep rate optimises the excitation signal length. Anyway, some extra silence has to be included at the end of the sweep to wait for all delayed components of the system response and to be able to isolate the distortion.

Due to the well arranged time-frequency sweep properties, the distortion from electro-acoustic devices can be segregated from the acoustic response. Fig.2 presents the spectrograms of the sweep response, (above) and the impulse response after the deconvolution (below) of a loudspeaker and microphone arranged in a room. The sweep rate, identified by the stronger line in the upper spectrogram, divides the energy distribution into two regions. Behind the sweep time arrival, there are some reflections and the room reverberation, and in front of it, there are loudspeaker artefacts. In the spectrogram, second and third order harmonic distortion can be identified, along with some transient noise caused by a not entirely air-tight closed box. The circular FFT deconvolution process make these components appear at the right end of the time axis, far away from the actual impulse response so that windowing can be applied to reject noise and the distortion entirely without corrupting the impulse response.

**CONCLUSIONS**

The well-defined time-frequency properties of the sweep sine allows measurements with high dynamic range, due to the versatility in creation and the possibility to exclude distortion100%. In random noise excitation signals the distortion products are randomly spread out over the complete impulse response period.

**REFERENCES**

Time Delay Estimation with Binary Gradient Adaptive Algorithm

A. Makar

Polish Naval Academy, Smidowicza 69, 81-103 Gdynia, Poland

Time delay estimation between signals, which are radiated by an underwater source and received in different points is possible by means of the method with gradient type adaptive algorithm. It is the method, which works on the basis of moving and synchronization received signals to the moment, when the estimation error has the minimum value. It is the effective method for signals, which have comparable intensity values, it means when the distance between receivers is not far. When receivers are located far away, their intensity values could distinguish as a result of propagation wastes. In this paper has been shown the method of time delay estimation which is effective, when the distance between hydrophones is far for location of underwater moving objects.

INTRODUCTION

The problem of estimating and tracking the delay between two signals arises in many areas such as speed measurements, localisation and tracking of signal sources and sonar or radar detection. The adaptive methods for solving this problem gained great popularity because they do not generally require a priori statistics about signals. This method of time delay estimation can be used for localisation of underwater objects by means of acoustics method on the basis of signals radiated by these objects in hyperbolic navigational system.

TIME DELAY ESTIMATION WITH GRADIENT ADAPTIVE ALGORITHM

The functional system of time delay estimation consists of two components: a sum function and a changeable delay. Input signals are as a sequence of signals samples. Minimizing of the error $e(k)$ in the output of the system is achieved by modification of the changeable delay until the value, which gives us the minimum of the error $e(k)$. This value of time delay is the estimate of time delay between input signals. Then this form of the variable time delay is the estimate of time delay between input signals. The issue of time delay estimation can be treated as the problem of synchronisation of two signals, when the accuracy of this synchronisation is given by $E[e(k)^2]$.

One of methods, how to get minimum of the output error is to use the gradient adaptive algorithm. The synchronisation process begins from a value of time delay, which is chosen before setting in the algorithm. Next a gradient of the error function is calculated, which in another step of iteration increases the accuracy of determination of time delay. The procedure is repeated for obtain the minimum of the output error. The general formula of optimisation algorithm, which uses the gradient type adaptive algorithm, can be written as follows:

$$\hat{\tau}(k+1) = \hat{\tau}(k) - \mu \cdot \nabla(k)$$

where:

$\mu$ - convergence coefficient, $\mu > 0$;

$\nabla(k)$ - gradient of the error with respect to $\hat{\tau}$.

Described algorithm of time delay estimation works correct for signals, which have comparable values of intensity, i.e. during receiving signals by hydrophones located in close distance. For long distance between receivers, as a result of absorption a signal by the environment, it is necessary to use modification of the method.

MODIFICATION OF THE METHOD

For modification of time delay estimation method has been used digitising of received by hydrophones signals for the binary form. Binary signals are two-value signals $u(t) = A \cdot a_n(t)$ which consist of sequence of impulses which can have only two values: 0 and A or −A and A. In described method has been used binary formula:

$$A = 1 \text{ for } u(t) \geq 0$$
$$A = 1 \text{ for } u(t) < 0$$

Binary signals, despite different signals levels, which are result of absorption in water, can be estimated by means of gradient adaptive method.
RESULTS

For investigation have been used signals radiated by a moving ship and recorded using Bruel & Kjaer measuring apparatus. Received signals have been shown in Fig. 1 and a binary signal and delay between received signals have been shown in Fig. 2.

CONCLUSIONS

For maximising the operation zone of the passive underwater source localisation system it is necessary to place hydrophones in long distances. The acoustic wave which is radiated by a moving ship covers different distances to hydrophones, so received signals have different value of the intensity. It is difficulty in working the time delay estimation algorithm with gradient adaptive method. This difficulty has been reduced digitising received signals for the binary form.

ACKNOWLEDGEMENTS

The research was supported by the State Committee of Scientific Research grant No 7 T07A 002 19

REFERENCES

Noise Reduction of Speech Signals Using MUSIC Algorithm

Takahiro Murakami, Munehiro Namba, Tetsuya Hoya and Yoshihisa Ishida

"Department of Electronics and Communication, Meiji University
1-1-1, Higashi-mita, Tama-ku, Kawasaki, 214-8571, Japan

bDepartment of Mathematics and Informatics, Tokyo Gakugei University
Koganei-City, Tokyo, 184-8501, Japan

Laboratory for Advanced Brain Signal Processing, BSI-RIKEN
2-1, Hirosawa, Wakoh-City, Saitama, 351-0198, Japan

This paper presents a new method for noise reduction of speech signals using the MUSIC (MUltiple SIgnal Classification) algorithm. The proposed MUSIC algorithm based on a subspace principle is capable of estimating the power spectrum of the speech signal corrupted by noise and available for the voice activity detector. In this paper, the noise components are efficiently eliminated since noise reduction is carried out only to the segment where speech is presence. Experimental results show that the proposed method greatly improves the quality of the noisy speech signal in low SNR environment.

1. INTRODUCTION

In various speech processing applications, noise reduction is very important as pre-processing, such as in automatic speech recognition or hearing aids for impaired people. For noise reduction of speech signals, a number of methods based on subspace principles, especially the utility of singular value decomposition (SVD), have recently been developed [1]-[3]. This paper proposes a new method for noise reduction of speech signals using the MUSIC (MUltiple SIgnal Classification) algorithm, which is one of the well-known methods based on subspace decomposition [4]. The MUSIC algorithm exploits noise subspace for estimation of the frequencies of complex sinusoids in additive white noise.

2. APPROACH

2.1 Principle of MUSIC Algorithm

Assuming that a noisy signal vector

\[ y = y(m) \quad y(m + 1) \cdots y(m + N - 1) \]  \( \text{as} \)

\[ y = Sa + n \]  \( \text{where} \)

\[ S = [s_1 \quad \cdots \quad s_p] \]  \( \text{and} \)

\[ s_k = [1 \quad e^{i2\pi k/N} \quad \cdots \quad e^{i2\pi (N-1)k/N}]^T \]  \( \text{where} \)

\[ P \] is the number of sinusoidal components of the signal, \( f_k \) is the frequency of the \( k \)-th complex sinusoid, \( X_k \) is the complex amplitude of the \( k \)-th complex sinusoid, and \( n \) is a zero-mean Gaussian noise vector with variance \( \sigma_n^2 \). The autocorrelation matrix of the noisy signal is given by

\[ R_{yy} = E[yy^H] \]  \( \text{where} \)

\[ E[\cdot] \] denotes the expectation operator, the exponent \( H \) is a Hermitian transpose.

Then, the eigen-decomposition of \( R_{yy} \) may be expressed as

\[ R_{yy} = \sum_{k=1}^{P} \mu_k v_k v_k^H \]  \( \text{where} \)

\( \mu_k \) and \( v_k \) are the eigenvalues and the eigenvectors of \( R_{yy} \), respectively. \( \mu_k \) are all real numbers and satisfy

\[ \mu_1 \geq \cdots \geq \mu_p > \mu_{p+1} = \cdots = \mu_{N} = \sigma_n^2 \]  \( \text{and} \)

\( v_k \) are partitioned into \( \{v_1 \quad \cdots \quad v_p\} \) span the signal subspace and \( \{v_{p+1} \quad \cdots \quad v_N\} \) span the noise subspace. The signal and noise subspace are mutually orthogonal, and the signal matrix \( S \), given by Eq.(2), is orthogonal to the noise subspace. Therefore, the MUSIC spectrum is given as

\[ P_{xx}^{\text{MUSIC}}(f) = \frac{1}{\sum_{k=1}^{p} \mu_k e^{i2\pi f_k/N} s(f)} \]  \( \text{where} \)

\[ s(f) = [1 \quad e^{i2\pi f/N} \quad \cdots \quad e^{i2\pi (N-1)f/N}]^T. \]

2.2 Noise Reduction of Speech Signal

The MUSIC spectrum is available for estimating the power spectrum of the speech signal buried in noise. However, the estimated spectrum has no phase information since the MUSIC spectrum is obtained from the autocorrelation matrix of the speech signal. Therefore, the phase spectrum of speech signals can be
calculated by the FFT spectrum of the observed signals as
\[
\phi(f) = \tan^{-1} \frac{\text{Im}\{\hat{X}(f)\}}{\text{Re}\{\hat{X}(f)\}}
\]
where \(\hat{X}(f)\) is the complex amplitude of the observed signal. Then, the MUSIC spectrum is combined with the phase spectrum as
\[
\hat{P}^{\text{MUSIC}}_{xx}(f) = P^{\text{MUSIC}}_{xx}(f) \times e^{i\phi(f)}
\]
and finally the estimated speech signal is obtained by taking the IFFT of the complex MUSIC spectrum.

2.3 Voice Activity Detector

The eigenvalues of the autocorrelation matrix of the speech signal are distributed as Eq.(7). On the other hand, in case of the observed signal where no speech is present, the eigenvalues of the autocorrelation matrix may be written as
\[
\mu_k = \sigma^2_k \quad (k = 1, \ldots, N).
\]
Therefore, the distribution of the eigenvalues is utilized for the voice activity detection as
\[
\frac{\max(\mu_k)}{\text{mean}(\mu_k)} > L_{th}
\]
where \(\max()\) and \(\text{mean}()\) respectively indicate the maximum and mean operators, and \(L_{th}\) is the threshold level.

3. EXPERIMENTAL RESULTS

The proposed method for noise reduction is summarized as follows:
1) The analyzed speech signal is sampled at 11.025[kHz] and divided into overlapping frames of 256 samples.
2) Each frame is hamming-windowed and analyzed using the proposed method. The threshold level of Eq.(13) is set as \(L_{th} = 20dB\).
3) The estimated speech signal via MUSIC algorithm is normalized to the variance of the noisy signal in a single frame, and then each signal frame is overlapped and added to the preceding and succeeding frames.

In the experiments, the speech signal used is uttered by Japanese female. The sentence is “Sakura ga saita” (in Japanese). The additive noise is assumed Gaussian, SNR=0.03dB.

Figs.1-3 respectively show the original speech signal, the noisy signal and the estimated speech signal. As shown in Fig.3, the noise components is greatly eliminated. As in the figure, the proposed method improves the SNR by 6.8dB.

4. CONCLUSION

In this paper, we have proposed a new method for noise reduction of speech signals. In the experimental results, the SNR of the noisy signal has been improved at around 6dB.

REFERENCES

Comparing Data Acquisition Techniques

L. Cioni

Laboratorio di Linguistica, Scuola Normale Superiore, Pisa, Italy

The main aim of this paper is to present a qualitative and quantitative comparison among some of the techniques used at the present to acquire samples of speech. Such techniques share the source of the data as well as their final destination but they differ as to the ways by which the stream of speech is captured and stored as a stream of digital data. In the paper we will examine briefly some of the available techniques. Such techniques will be briefly examined on their own and then compared so to provide users with suggestions about the preferable technique for a certain situation.

INTRODUCTION

At the present many techniques are available for the acquisition of samples of speech as well as many are the purposes for such acquisitions but all share the source of the data (generally speaking such a source is a speaker who produces utterances that are captured by a microphone) as well as their final destination (a set of files stored on a magnetic medium) but they differ as to the ways by which the stream of speech is captured and stored as a stream of digital data.

In this paper we present the following techniques:

1. recording with a DAT and acquisition on a computer with an either analog or digital connection with an audio board driven by commercial or custom software;
2. recording with a minidisk recorder (MD) and acquisition on a computer with an either analog or digital connection with an audio board driven by commercial or custom software and
3. recording with a direct (analog) connection between a microphone and an audio board driven by commercial or custom software.

All the techniques presented in the paper make use of a Sound Blaster Live! Platinum 5.1 audio board that is equipped with both analog and digital connections [1]. It is worth mentioning that a minidisk once recorded with a MD can be loaded on a computer via an MDS-PC2 board [2, 3] but, since such a board is equipped with both digital and analog I/O connections, this case can be assimilated to the number 2 listed above.

BRIEF DESCRIPTION OF THE TECHNIQUES AND THEIR CHARACTERISTICS

In all the three cases mentioned above we have an analog phase (the microphone), an analog to digital conversion (microphone → DAT, microphone → MD or microphone → audio board) and possibly two conversions digital to analog and analog to digital if we use an analog connection between the DAT or MD and the audio board. The simplest technique is the number 3 where we have only a connection and a single analog to digital conversion. The quality of the acquisition process is determined by the quality of the microphone and of the audio board. The others are two steps techniques since they are characterized by a recording phase and an acquisition phase so that both phases determine the quality of the acquired signals.

The use of a DAT [4] allows the recording of speech in digital format usually as 16 bit linear coding with sampling frequencies of either 48kHz or 44.1 kHz with a S/N ratio better than 87 dB and a Dynamic Range better than 87 dB. The use of a MD for the recording of speech at a default sampling frequency of 4.1 kHz imposes some cautions and requires a deeper investigation since MD uses a special technique of data compression called Adaptive Transform Acoustic Coding (ATRAC).

ATRAC [2, 5] exploits some psychoacoustic principles plus sub-band coding and transform coding so to act on the frequency components of the signal and “discard” those that are perceptually irrelevant.

In this way the algorithm extracts and codes only the frequency components that are audible by the human ear. While this technique is guaranteed not to reduce sound quality and not to be perceived by listeners it is still to be proved that the frequency transformations have no influence over the parameters extracted from the speech signal.

THE EXPERIMENTAL SETTING

In order to compare the three techniques an experimental setting has been conceived. The setting is shown in the following figures 1 and 2. Figure 1 shows the first phase (recording phase). From such a figure we see how each speaker has been recorded with the three recorders in parallel so that the analysis (and the comparison) can be performed on the same utterances and under the same boundary conditions.
On the other hand figure 2 shows how the utterances recorded both with the DAT and the MD are stored on a Personal Computer via an audio board and either an analog or a digital (optical) connection.

The acquisitions are performed with a Sound Blaster audio board driven by Goldwave using both the analog mic-in (figure 1) and line-in (figure 2) connections and the digital SPDIF-In connection (figure 2).

The corpus of utterances used for the frequency analysis and comparison if formed by a small set of trisyllabic and disyllabic words of the form CVCCVCV, CVCCVCVCV (e.g. “canguro”), CVCCVCVV (e.g. “fantasia”) and CVCCV /e.g. “canto”) uttered in isolation.

CONCLUSIONS

In this paper we have sketched out three acquisition techniques of which we showed some characteristics. As shown in figure 1 and 2 we examined have two recordings styles and three acquisition methods. The use of a direct microphonic analog connection is well suited mainly in case of laboratory recordings with one speaker at a time. Portable DT and MD, on the other hand, can be used also in open air recordings as well as in laboratory recordings and DAT is equipped with a device that allows the use of two microphones, one on each channel, so to record two speakers at the same time. The main problem with a DAT is that speakers are aware of being recorded so that they produce at the best semi-spontaneous speech.

In such cases the use of a MD can be of help since the medium is smaller and less intrusive. The only problem is that is not yet clear if the ATRAC algorithm has a negative influence on the parameters extracted from the acquired signals.

At the end, we underline that the analysis presented in the paper is mainly qualitative and has been performed on a very small corpus produced by a single speaker. Future goals are the extension of the analysis in a more quantitative perspective and on wider corpora both acquired ad hoc or produced within other and wider research projects.

ACKNOWLEDGEMENTS

The author acknowledges the ownership of the trademarks of the products mentioned in the paper to their owners even if not explicitly cited.

REFERENCES

Optimisation of Settling Signal in MLS Technique

D. G. Ciric

Faculty of Electronic Engineering, University of Nis, 18000 Nis, Yugoslavia

In the measurement of the room impulse response by Maximum Length Sequence (MLS) technique, measured system should already be stabilized at the moment of measurement beginning, that is, the steady-state conditions should be settled. For this purpose, the settling signal preceding excitation MLS signal is used. The influence of various settling signals on improvement of Peak-to-Noise Ratio (PNR) of room impulse response is investigated in this paper on experimental basis. The results show that excitation MLS signal as settling signal gives the best PNR improvement. However, the same improvement can be obtained by considerably shorter signal representing optimal settling signal generated from corresponding number of the last samples of excitation MLS signal. Using optimal settling signal can considerably reduce measurement time.

INTRODUCTION

The advantages of MLS technique emerge only with the adequate implementation of the technique. One of the necessary requirements for this implementation is that measured system should already be stabilized at the moment of measurement beginning [1, 2]. In other words, measured system should already be settled to steady-state conditions. As this settling can not be performed at the moment, the solution is found in the excitation of the room by additional signal, here called settling signal, prior to MLS signal. It is usual that excitation MLS signal is used as settling signal. The influence of settling signal on improvement of measurement results is investigated in this paper with the purpose of showing the importance of settling signal and of obtaining the optimal settling signal leading to the best results. As a measure of result improvement, PNRs of broad-band impulse responses of a particular room measured by developed system [3] are observed.

INFLUENCE OF SETTLING SIGNAL

The impulse responses were measured in the room at the Faculty of Electronic Engineering in Nis [3]. The data were firstly sampled by frequency of 22.05 kHz. MLS signals of orders of 15 (MLS15) and of 17 (MLS17) were used as excitation signal, while various signals were used as settling signal: MLS signals of lower orders, then excitation MLS signal and signals generated from corresponding numbers of samples of excitation MLS signal taken from its beginning or from its end (signals ls1k, ls2k, ls5k, ls10k, ls20k and ls50k generated from the last 1k, 2k, 5k, 10k, 20k and 50k samples, respectively, where k represents 1000). PNRs of impulse responses measured by using settling signal prior to MLS signal are compared to PNR of response measured by using only MLS signal (the settling signal is 0). Thus, it is obtained that signals generated from the first samples of excitation MLS signal and MLS signals of lower orders as settling signal do not lead to the improvement of PNR, but to PNR decreasing. On the other hand, using settling signals generated from the last samples of excitation MLS signal leads to PNR improvement, Fig. 1.

FIGURE 1. PNRs of impulse responses measured using sampling frequency of 22.05 kHz.
value of improvement increases with the increasing of settling signal length up to the length of about 10k samples. The settling signals of this and bigger lengths yield approximately equal improvement to that one obtained for excitation MLS signal as the settling signal. Similar dependence of PNR improvement appears in MLS15 and MLS17 excitation signals. The same investigations are also performed using two other sampling frequencies of 11.025 and 44.1 kHz. The results again show that adequate settling signal can be excitation MLS signal and signals generated from corresponding numbers of the last samples of excitation MLS signal, Fig. 2. Nevertheless, the sampling frequency affects this corresponding number of the last samples, that is, the length of settling signal in samples. So, the length of the shortest settling signal leading to approximately the same PNR improvement as when excitation MLS signal is used as settling signal is 5k samples for sampling frequency of 11.025 kHz, that is 20k for sampling frequency of 44.1 kHz. This points out that duration of settling signal (in time) is an important factor.

![FIGURE 2. PNRs of impulse responses measured using MLS15 as excitation signal a), the dependence of ∆PNR versus settling signal duration b).](image)

In order to compare PNR improvements for various sampling frequencies obtained using corresponding settling signals of the same duration generated from the last samples of excitation MLS signal, the dependence of difference of these improvements and the one obtained using excitation MLS signal as settling signal (ΔPNR) upon the duration of settling signal is plotted, Fig. 2. b). For this purpose, the duration of every settling signal is calculated. The presented plot shows that settling signal generated from the last samples of excitation MLS signal whose duration is about 0.5 s or longer yields approximately equal PNR improvement as excitation MLS signal used as settling signal. Since this signal of duration of 0.5 s is the shortest one leading to the mentioned (biggest) PNR improvement, it can be called the optimal settling signal. By theory, the minimum duration of settling signal should be equal to the time for which the impulse response decays to a negligible value [1]. The duration of optimal settling signal obtained here is approximately equal to this theoretical minimum duration.

**CONCLUSIONS**

The settling of steady-state conditions already at the starting moment of measurement of room impulse response by implementation of MLS technique considerably improves PNR of the response. Due to this, MLS signal should be preceded by adequate settling signal. The excitation MLS signal as well as the optimal settling signal generated from corresponding number of the last samples of excitation MLS signal represent such signals leading to the same PNR improvement. As optimal settling signal is considerably shorter than excitation MLS signal, its using considerably reduces measurement time. Namely, in presented examples, the overall duration of corresponding optimal settling signal and the excitation MLS17 signal was 6.4 s for sampling frequency of 22.05 kHz. This duration is considerably shorter than that one of two MLS17 signals (settling and excitation signals), which is 11.9 s. If the averaging is implemented, the reduction of measurement time becomes bigger.

**REFERENCES**

REALIZATION OF HIGH QUALITY BAND STOP FILTER USING FPGA

W. Boonkumklao\textsuperscript{a}, Y. Miyanaga\textsuperscript{a} and N. Miki\textsuperscript{b}

\textsuperscript{a}Graduate School of Engineering, Hokkaido University, Sapporo, Japan
\textsuperscript{b}Hakodate Future College, Hakodate, Japan

In this report, we propose a new design method which can realize good flexibility for high speed processing with small number of gates. In addition, a new hardware architecture to realize a high quality band elimination filter using the proposed design method is introduced. The band elimination filter is usually required in a communication system to get speech data with good quality. In other words, when we consider a wire-less communication in which speech waveform and digital data (DTMF) are mixed in the same band, the band elimination filter with quite high quality should be required. The designed filter is useful to eliminate the DTMF signals only. It is also shown that an actually designed hardware can realize the low number of gates and high processing speed. In this report the hardware is realized with FPGA. The FPGA (Field Programmable Gate Array) is well known as one of programmable devices. In the case of the small number of production for a special purpose digital signal processing, the fabrication cost can be considerably reduced by using FPGA. Both the new design method on FPGA and a filter design as an example are explored in this report.

INTRODUCTION

A generic DSP has flexibility to design the certain application system including band stop filters\cite{1}. However, it is not easy task to program a complex software using a DSP assembler because of the difficulty on debugging in run-time environment and also DSP has the limitation of an architecture for high order filters. On the other hand, a systolic array architecture has been proposed for high performance filtering in real time systems\cite{2}. This approach is attractive to realize a high speed system. However many arithmetic units have to be realized on a chip, the cost of design and implementation is quite expensive and its architecture has no large flexibility. The architecture of Fig.1 is one of the solution of the flexibility for high speed processing\cite{3}.

REPROGRAMMABLE ARCHITECTURE

As in Fig.1, the proposed hardware core has some arithmetic components, several multipliers adders and subtracters. All of them are connected with the multiplexers controlled by a sequencer. In the block of the arithmetic components, the access dead lock to bus does not happen since there is no common bus into them. The blocks of memories are allocated for all data, e.g., coefficients and its states in digital filters. Each of them has two ports for read and write operations. The sequencer controls the flow of data and repeats the cycle of the processing.

ARCHITECTURE OF 2ND CASCADE FILTER

In a mobile radio communication system, an accurate filter which extracts only voice signals from the data including both speech and DTMF(Dual Tone Multiple Frequency) has been required. Since the DTMF signal exists in 500-2000Hz and the voice information does not have to lose as much as possible, the filter which has the characteristics such as shown in the Fig.4 becomes necessary.
In our design method, the 2nd cascade filter is employed. In this case, the architecture required two delay elements, three multipliers and two adders. For the construct of a filter section multiplexers are controlled.

By using this architecture, the high order cascade filter can be realized as shown in Fig.3. Let us explain the behavior on this architecture in Fig.3. The first session of the 2nd cascade filter is calculated with 2 clocks. When the first clock is raising, the address is calculated by the sequencer and it is sent off to the memory. Then a coefficient in the given address of ROM is loaded. A control signal is fed to the multiplexer and the register to read data. Then the left of the 1st section is computed. Next, memory RAMs (D0, D1) are used to memorize calculation results as w1 where w1 is defined as Fig.2. At the second clock, a filter coefficient and w1 are read from ROMs in the same way as the first step. Then the right of the 1st section is computed. From the second session, the excursion mechanism is the same as the first section. Note that the number of multipliers used for the left of each session and the right one as shown in the Fig.2 are not equal. Multiplexers (mux0, mux1) are used to change the structure of a circuit. It indicates the novel characteristic of our proposed architecture in Fig.3.

In our design system of the synthesizer, the order of the 2nd cascade filter has been set as 12 and the sampling frequency as 11.025kHz. Using this specification, we can see the output waveform in which DTMF can be erased. We performed the logic simulation of our VHDL program for the designed filter. The same impulse response can be calculated as the result given by C++ programming software system.

Our VHDL program for the DSP design is compiled and installed on the chip of FPGA (Altera FLEX10K) which has 4992 LUT and 9216 bit memory; this corresponds to 100K gates. The total LUT of the designed system is only 4722 and however the system has a programmable architecture. The characteristics of the realized filter is show in Fig.4.

We have proposed and discussed a new digital signal processing system with programmable architecture. This designed method is used for the 2nd cascade filter. The architecture is optimized as to the bit length of fixed point computation. In addition, the band stop filter which is useful to eliminate DTMF is realized on the chip FPGA and it is verified that the performance of the total hardware system satisfies the required specification.

REFERENCES

Using symbolic dynamic and change-point to isolate transition between vowel and consonant phonemes in connected speech

A. Guillamón ¹, F. Martínez ¹, A. García-Sánchez ¹, M.C. Ruiz ¹ and J.C. Alcaraz ¹

¹Department of Mathematics Applied and Statistics, University Politécnica de Cartagena, 30203 Cartagena, ESPAÑA

The aim of this paper is to present an automatic segmentation method based on nonlinear dynamics in order to characterize the transition between vowel and consonant phonemes in connected speech. The “symbolic dynamic” is an alternative and new method to “approximate entropy” quantifying the amount of regularity in time series. The segmentation process is carried out in three stages, the estimation of the approximate entropy, smoothness of the obtained signal by means of moving average technique and detection of change point using the Chung-Bow’s statistic. Chung-Bow’s statistic provides a nonparametric method that doesn’t need the stationary condition, hypothesis commonly used in classical theory of speech analysis. Through these techniques, we obtain an automatic method which can be used to isolate the section of the speech signal corresponding to transitions between vowel and consonant phonemes in connected speech. Our experiments have been computed on signals from a specific speech Spanish database (AHUMADA Database) designed for speaker recognition tasks in Spanish language.

Keywords: Signal processing; symbolic dynamics; change-point.

INTRODUCTION

Recently, it has been developed different segmentation methods of speech signals, but always under the hypothesis of the stationary condition of the studied signal. The stationary condition of the speech signal is a hypothesis commonly used in classical speech analysis [6], when we study frames of the recorded signals corresponding to voiced phonemes (vowels or nasal consonants). However, in the study of unvoiced sounds and transitions between vowel and consonant phonemes, the time series obtained are non-stationary [4]. Moreover, the phenomenon known as “nasal coarticulation” [7] appears in the transitions between vowel and nasal consonant. Therefore, those methods are not appropriate in this case. In this paper, we propose a segmentation method of speech signals based on nonlinear dynamics, in order to obtain frames corresponding to the transitions between vowel and nasal consonant and vice versa. The segmentation procedure is carried out in three stages: first, the series corresponding to the symbolic dynamics of the speech signals are obtained; next, these symbolic series are smoothed by means of moving average technique; and finally, the change-points corresponding to the start and end of the transition are located using the Chung-Bow statistic [2].

METHODOLOGY

Symbolic Dynamics

The symbolic dynamics technique [5] converts the time series of the speech signal \( \{x_t\} \) \( t = 1, \ldots, N \) in a sequence of symbols \( \{s_j\} \) based on a window of size \( M \), according to the following expression:

\[
s_j = \begin{cases} 
0 & \text{if } |x_j - x_{j-1}| \geq a \cdot \sigma_j \\
1 & \text{if } |x_j - x_{j-1}| < a \cdot \sigma_j
\end{cases}
\]

for \( j = 1, \ldots, N-M+1 \), where \( \sigma_j \) is the std. deviation of the recorded signal \( x_i \) over the \( j \)th window and \( a \) is a tolerance measure.

Smooth the Symbolic Series

The symbolic series \( \{s_j\} \) shows random fluctuations which impede to apply a change-point technique directly. This random fluctuations, can be removed by means of moving average method [1]. Thus, we generate a smoothed series \( \{\tilde{s}_i\} \) from a given series \( \{s_j\} \) in the following way:

\[
\tilde{s}_i = \frac{1}{T} \sum_{j=i-(T-1)/2}^{i+(T-1)/2} s_j
\]

where \( i = (T-1)/2, \ldots, M-(T-1)/2 \) and \( T \) is the window size.
Nonparametric change-point Estimation

Detection of change-points corresponding to the start and end of the transitions are carried out by means of Chung-Bow’s statistic [2]. This method, based on the weighted empirical measures, detects two different nonparametric distributions over a window of the series \( \hat{s}_i \) and then runs the windows over the full series.

In [2], it is considered a sequence of random variables \( S_1, \ldots, S_n \) having the following structure:

\[
L(S_j) = \begin{cases} 
    P_1 & \text{if } 1 \leq i \leq j_1 \\
    P_r & \text{if } j_{r-1} < i \leq j_r \\
    P_{k+1} & \text{if } j_k < i \leq n
\end{cases}
\]

where \( L(S_j) \) denotes the distribution of \( S_i \), \( P_j \)'s are unknown probability measures, \( k \) is the number of change-points and \( j_1, \ldots, j_k \) are the locations of the change-points.

Then, the difference \( D_j \) of two weighted empirical measures is taken for each possible location \( j \) of change-point:

\[
D_j = \frac{1}{\sqrt{c}} \sum_{i=1}^{A_n} c_i [I_{S_{j-r+1}} - I_{S_{j-r-1}}]
\]

where \( A_n \) is the length of the windows, \( I \) is the indicator function \( (I_S(s) = 1 \text{ if } s \in S) \) and \( c = (c_1, \ldots, c_{A_n}) \) is a positive vector. If we assume that the difference between two successive change-point is larger than \( 2A_n \), the expectation of the measure \( D_j \) will be:

\[
E[D_j] = \rho_n(j) [P_r - P_{r+1}] \quad r = 1, 2, \ldots, k
\]

where:

\[
\rho_n(j) = \begin{cases} 
    \frac{1}{\sqrt{c}} \sum_{i=1}^{A_n} c_i [1 - j_i] c_i & \text{if } |j - \hat{j}_k| \leq A_n - 1 \\
    0 & \text{if } |j - \hat{j}_k| > A_n - 1
\end{cases}
\]

In [2], it is proved that the function \( \rho_n(j) \) attains its local maximum values at each true location change-point and hence the same for \( |D_j| \).

RESULTS AND DISCUSSION

As a first result we can confirm that the speech signal corresponding to transition between vowel and nasal consonant verifies the nonstationary condition. On the other hand, the method proposed in this paper gives an efficient segmentation technique of speech signals for voiced and unvoiced phonemes and we obtain an automatic method that can be used to isolate the section of the speech signal corresponding to transitions between vowel and consonant phonemes in connected speech.

Our experiments have been applied on signals from a specific speech Spanish database (AHUMADA Database [3]) designed for speaker recognition tasks in Spanish language. Figure 1 shows the results obtained for the signal \( a/n/a \) of the AHUMADA Database (with \( T = 129, a = 0.3, M = 256 \) and \( A_n = 128 \)).

FIGURE 1. Transition a/n/a (AHUMADA Database)

REFERENCES

4. F. Martínez, A. Guillamón y J.C. Alcaraz, ”Caracterización de transiciones entre fonemas en el habla natural mediante la entropía aproximada”, Proceedings of the II Latin American Congress on Biomedical Engineering (2001)
A new metric based on ARMA models: Application to speech signals

A. Guillamón, F. Martínez, A. García, M.C. Ruiz, J.C. Alcaraz

Departamento de Matemática Aplicada y Estadística, Universidad Politécnica de Cartagena (España)

Working with ARMA models as complex rational functions, it is possible to define a metric \(d(M,M')\) between two stable ARMA models \(M, M'\) by means of the cepstrum coefficients of the models. As this metric verifies the property of invariability under translations, we can obtain two ARMA models from two ARMA models, and thus the metric may be calculated algorithmically as a finite product in the pole-zero domain or from the ARMA coefficients by means of determinants. In this paper we propose a speaker recognition method based on the use of cepstral features obtained from ARMA models. We use Shank's method to determine ARMA coefficients. The analysis is carried out using frames of 150 ms duration. The feature vectors are computed in transitions between consonant and vowel phonemes in connected speech. We propose this method as an alternative to classical discriminant analysis for speech signals in order to identify and verify the source in a set of speakers' signals. Our experiments have been computed over signals from a specific speech database (AHUMADA Database) designed for speaker recognition tasks in Spanish language.

INTRODUCTION

Speaker Recognition is the task of identifying a speaker by his or her voice. In a closed system performing speaker recognition, a particular speaker is identified as one in the finite set of reference speakers, [1]. The overall system that we consider consists of two stages: the ARMA analysis of the speech and the classification for taking decision.

It is said that the time series \(\{x(n)\}\) follows an ARMA \((p,q)\) model if:

\[
x(n) = \sum_{k=0}^{p} a_k \cdot x(n-k) + G \cdot \sum_{l=0}^{q} b_l \cdot u(n-l), b_0=1 \quad (1)
\]

where \(\{u(n)\}\) are unknown input elements, \(G\) is the gain of the model and \((a_k)\) and \((b_l)\) are the ARMA parameters with \(b_0=1\). Alternatively, if we work in the \(z\)-domain the system function is given by:

\[
H(z) = G \cdot \frac{\sum_{j=0}^{q} b_j \cdot z^{-j}}{\sum_{i=0}^{p} a_i \cdot z^{-i}} = G \cdot \frac{\prod_{j=0}^{q} (1-\beta_j z^{-1})}{\prod_{i=0}^{p} (1-\alpha_i z^{-1})} \quad (2)
\]

where \(a=1\) and \((\alpha_k)\) and \((\beta)\) are the poles and zeros of the model, respectively.

Most of speaker recognition systems use some type of short-time spectral analysis [2]. Thus, the speech analysis is carried out using frames of 150 ms duration. The feature vectors are computed in frames corresponding to transitions between nasal sounds and vowels. It should be noted that the choice of these voiced frames presents nasal coarticulation. The nasal coarticulation can be used as an acoustic clue for identifying speakers, since it concerns a rapid event, it is not likely to be consciously modified in natural speech, [3].

PARAMETERS ESTIMATION

We use Shank's method to determine the ARMA model coefficients, [4]. In this approach, if an ARMA model has a system function \(B(z)\), a minimum phase \(A(z)\) is first determined by AR analysis. Besides, if \(h(n)\) denoted the impulse response corresponding to the partial system \(1/A(z)\), then the error in the time domain is given by \(x(n)-h(n)*b(n)\).

Using least-squares method, the coefficients of \(B(z)\) are obtained through a system of linear equations:

\[
\begin{pmatrix}
  h(0) & 0 & \cdots & 0 \\
  h(1) & h(0) & \cdots & 0 \\
  \vdots & \vdots & \ddots & \vdots \\
  h(q) & h(q-1) & \cdots & h(0) \\
  \vdots & \vdots & \ddots & \vdots \\
  h(N-1) & h(N-2) & \cdots & h(N-q-1)
\end{pmatrix}
\begin{pmatrix}
  b(0) \\
  b(1) \\
  \vdots \\
  b(q)
\end{pmatrix}
= \begin{pmatrix}
  x(0) \\
  x(1) \\
  \vdots \\
  x(N-1)
\end{pmatrix} \quad (3)
\]

where the sequence \(h(n)\) results from:

\[
h(n) = -\sum_{k=1}^{p} a_k \cdot h(n-k) + \delta(n).
\]
Applying the AIC, we obtain that the optimal number of poles \( p \) and zeros \( q \) are 12 and 2, respectively. [5].

**METRIC**

We can associate an ARMA model with a nonzero complex rational function. It is known that the cepstrum can be obtained by the inverse Fourier Transform of the logarithm of the power spectrum. Then, for ARMA models, it reduces to

\[
\log[H(z)H'(1/z)] = \sum_{n=-\infty}^{\infty} c_n z^{-n}
\]

(4)

where \( H(z) \) is the system function of the ARMA model [6]. \((c_n)\) are the cepstrum coefficients that form a hermitian sequence \((c^*_n = c_{-n})\).

**Definition 1:** For two ARMA models \( M \) and \( M' \) with cepstrum coefficients \((c_0)\) and \((c'_0)\), we define:

\[
d(M, M') = \left( \sum_{n=0}^{\infty} |c_n - c'_n| \right)^{1/2}
\]

(5)

Note that \( d \) is a pseudometric, because if \((c_n), (c'_n)\) are such that \( c_n = c'_n, \forall n \in \mathbb{Z} - \{0\}\), then \( d(M, M') = 0 \), (but \( c_0 \) and \( c'_0 \) can be different). However, there is a standard method of turning a pseudometric space into a metric space, [7].

As \( d \) is invariant under translations, from two ARMA models we may construct two AR models. This fact permit us to calculate algorithmically the metric from the coefficients by using determinants, [8]. The following definition provides an algorithm to calculate \( d \) without obtaining the cepstrum coefficients.

**Definition 2:** For two stable AR models \( M \) and \( M' \) with orders \( p, p' \) and coefficients \((a_j)\), \((a'_j)\), respectively. If we denote:

\[
A(z) = \sum_{j=0}^{p} a_j z^{-j}, \quad A'(z) = \sum_{j=0}^{p'} a'_j z^{-j}
\]

(analogously \( A' \) and \( A'' \) for the \( M' \) model), then we can define:

\[
d(M, M') = \left[ \log \frac{\text{res}(A, A') \cdot \text{res}(A', A'')}{\text{res}(A, A') \cdot \text{res}(A', A'')} \right]^{1/2}
\]

(6)

where \( \text{res}(A, B) \) denotes the resultant of \( A \) and \( B \), [8].

**RESULTS AND DISCUSSION**

To illustrate the behavior of the method proposed in this paper, we have applied it to three different frames of speech signals corresponding to three situations with nasal coarticulation (m/a, n/o and m/i), over signals from a specific speech Spanish database (AHUMADA Database [9]) designed for speaker recognition tasks in Spanish language. The test determines if the source of a speech signal belongs to a specific speaker or to the rest of speakers, in a set of 15 speakers.

<table>
<thead>
<tr>
<th>Table 1. Results obtained over AHUMADA Database</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Sensibility</td>
</tr>
<tr>
<td>Specificity</td>
</tr>
<tr>
<td>Negative Predicted value</td>
</tr>
<tr>
<td>Efficiency</td>
</tr>
</tbody>
</table>

Sensibility represents the proportion of specific speakers verified, specificity is the ability of the methods to discard a speakers, negative predicted value is the proportion of successes when the test discards a specific speaker and efficiency shows the proportion of the test successes.

**REFERENCES**

Visualising acoustic surface properties, using colours

Claus Lynge Christensen

Ørsted-DTU, Acoustic Technology, Technical University of Denmark, building 352, DK-2800 Lyngby, Denmark.

The OpenGL™ rendering engine, which comes with the Windows PC operating system, allows display of room geometries, which have been modeled for room acoustics prediction purposes. Viewing a 3D geometry as a surface geometry rather than as a wire frame, allows easy detection of geometrical errors like holes and missing surfaces. However with an appropriate surface colouration it is also possible to view material properties. This paper suggests that the acoustic reflectance of acoustic materials is converted directly into a visual representation, where red represents low frequency, green represents middle frequency and blue represents high frequency reflectance. In this way total absorbing surfaces appear black whereas total reflecting surfaces appear white and indeed surface colours, which display acoustic reflectance appears in a visually plausible way and it is very easy to read the acoustics properties of the surfaces in a room model.

OpenGL for visualisation

Today’s computer configurations allow rendering of surface geometries using the OpenGL rendering engine, which comes with the Windows operating system. OpenGL is usually used for geometry rendering in computer games, but may indeed also be used visualisation of geometries in scientific programs – in fact Silicon Graphics™ originally created OpenGL for this purpose.

Acoustic properties to visualize

In room acoustics prediction programs based on geometrical assumptions, a number of surface properties may be applied to the surfaces in the room geometry. The properties are absorption coefficients, scattering coefficients and for some prediction programs transparency coefficients as well.

Transparency coefficients

In some room acoustics programs it is possible to assign a frequency independent transparency coefficients in order to model surfaces, which have a degree of transparency or to model an area covered by many small surfaces in an easy way, just modeling one large surface with a transparency coefficient. Examples of such surfaces are banisters, rails, installations bridges etc. In OpenGL it is possible to assign a visual transparency coefficient to each surface, so using this option is an obvious way to visualize a frequency independent transparency coefficient.

Scattering coefficients

Most room acoustics programs use scattering coefficients, which are assigned to each surface in the room geometries in order to describe to what degree reflected sound is scattered by a surface. In OpenGL it is also possible to assign a kind of scattering coefficient to surfaces, which determines the amount of shininess (white specular highlights) the surface will have.

Absorption Coefficients

The main material properties in a room acoustics modelling are the absorption properties. In a room acoustic program like ODEON, absorption coefficients, one for each frequency band, are assigned to each surface. In OpenGL, it is possible to assign the surface colours in terms of Red, Green and Blue intensities (the human eye is capable to distinguishing three different frequency bands of light because of different tap-cells in the eye, which are sensitive to each frequency range). To convert the acoustic absorption coefficients into a visual RGB intensity, an average absorption coefficient is calculated for the lower, the middle and the upper frequency range of the material. The average absorption coefficients are then converted into acoustic reflectance \( r = 1 - \alpha \), which are then used for the specification of colour intensities, one
intensity for each of the colours red, green and blue of the surfaces in the visual OpenGL representation.

Using the rendered room geometries

A visual 3D representation of the room geometry with the surfaces coloured in the acoustics colours as suggested above, allows a very efficient check on the assigned materials e.g. where are the absorbing or hard materials, have some of the surfaces been assigned a wrong material etc. As a bonus the colours usually looks quite natural, thus offers a nice way of presenting the room geometries built for acoustic prediction without having to spend time on preparing a sensible visual representation. The colour mapping is as follows:

- Low frequency reflectance is mapped into Red (which is low frequency light)
- Middle frequency reflectance is mapped into Green (which is middle frequency light)
- High frequency reflectance is mapped into Blue (which is high frequency light)

Some examples on the visual representation of different acoustics materials:

- White; total reflecting at all frequencies
- Black; total absorbing at all frequencies
- Grey scale; equal absorption at all frequencies
- Strong colour; strong colouration of the sound

The examples displayed here are in black and white even so with intensity based colour mapping of acoustic materials, one gets a reasonable impression of the material properties (dark surfaces reflecting less sound than white ones). If the pictures were reproduced in colours one would also get some information on the acoustic colours of the materials.

Some examples on the visual representation of different acoustics materials:

- White; total reflecting at all frequencies
- Black; total absorbing at all frequencies
- Grey scale; equal absorption at all frequencies
- Strong colour; strong colouration of the sound

Examples on rendered rooms

FIGURE 1. Visualisation of a computer model of the Royal Festival Hall, London.

FIGURE 2. Interior of a mosque, the acoustic colour of the carpet appear in a redish colour while windows appear blue.

CONCLUSION

OpenGL offers an easy yet very powerful way of providing information to users of a room acoustic prediction programs. In particular if a comprehensive colours scheme is chosen, it is very easy to read the material properties of a room.

REFERENCES

3. The Odeon home page, www.dat.dtu.dk/~odeon
Streaming of AAC data over Internet by using RTP/RTCP

Taejin Lee, Jose Soler Lucas, and Jinwoo Hong

Broadcasting Media Technology Department, Electronics and Telecommunications Research Institute
161 Gajeong-dong, Yuseong-gu, Daejeon, 305-350, KOREA, Tel +82-42-860-5713, Fax +82-42-860-6465
tjlee@etri.re.kr, jsoler@etri.re.kr, jwhong@etri.re.kr

Abstract: In this paper, we propose a method for AAC streaming over Internet using RTP/RTCP over UDP/IP and TCP/IP protocols. We use RTP for the real-time transmission of streaming data, RTCP for the control of real-time data, TCP/IP for some important control data exchange between server and clients, and the IETF’s AAC payload format for RTP is applied for the packet of real-time streaming data. We implemented and tested our streaming system using ETRI’s intranet, which is connected by about 2000 users. As a result, our streaming system works properly in our intranet environment, even though there is some packet loss or jitters, so our streaming system can be used for AAC audio broadcasting service or other applications. For the demonstration of our system, we made an exhibition through the KOBA 2001 in Korea.

MPEG-2 AAC

Because of its extreme qualities, AAC (Advanced Audio Coding) has been adopted as the audio coding standard in many applications. We choose AAC for its “stand alone” raw data block decoding property. Internet, as a lossy transmission environment, makes possible that a packet of the streamed data can be lost. So every packet should have “stand alone” decoding property, as it is the case of the AAC data [1].

In our application, we use 96kbps AAC bitstreams for streaming data.

INTERNET PROTOCOLS

As depicted in Figure 1, we use RTP/RTCP over UDP/IP protocols for the real-time transmission and control of the AAC bitstream over Internet. We also use TCP/IP protocols for reliable exchange of control information.

![Packet structure of AAC bitstream](image)

FIGURE 1. Packet structure of AAC bitstream.

RTP (Real-time Transport Protocol) offers end-to-end transmission of real-time data over multicast or unicast networks. RTCP (Real-Time Control Protocol) allows monitoring of client status, like fraction lost, cumulative packet loss and jitter information [2]. UDP (User Datagram Protocol) is the transport layer protocol. No other services are offered by UDP to the upper layers, so the characteristics of the transmission based on it same as the IP: connectionless and unreliable. IP (Internet Protocol) defines the basic unit of data transfer, so a TCP/IP Internet is sometimes called an IP-based technology [3].

SYSTEM ARCHITECTURE

The client’s H/W specifications of our streaming system, in Table 1, are the minimum requirements for our system [4].

<table>
<thead>
<tr>
<th></th>
<th>Server</th>
<th>Client</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processor</td>
<td>PIII Xeon 800MHz</td>
<td>Above PII 200MHz</td>
</tr>
<tr>
<td>Memory</td>
<td>512MB SDRAM</td>
<td>More than 128MB</td>
</tr>
<tr>
<td>Network card</td>
<td>32bit NIC 10/100 TX</td>
<td>Don’t care</td>
</tr>
<tr>
<td>Hard disk</td>
<td>90GB disk array</td>
<td>More than 10MB</td>
</tr>
<tr>
<td>OS</td>
<td>Win 2000 Server</td>
<td>Win 95/98/ME/2000</td>
</tr>
</tbody>
</table>

Figure 2 shows the sequence of procedures for AAC streaming in Internet. After insertion of audio data to Watermark insertion block, watermark data are added to raw audio data. Our watermark insertion and detection algorithm are based on a PN sequence. We insert 1 bit to each 1024 samples of audio data to insert the copyright information of the AAC bitstream. The AAC Encoder block encodes the AAC bitstream by using watermarked audio data and produces the AAC bitstream. Payload packing block uses the AAC bitstream and makes payload packed data and send it to RTP packing block. At Payload packing block, we implemented IETF draft payload and our modified payload format [5]. Both payload formats are based on priority of AAC bitstream structure. The modified payload format is easier to implement at the decoder part. RTP packing block makes RTP packet by adding RTP header to payload packet data and transfer to UDP/IP block. UDP/IP block add UDP/IP header to RTP packet and constructs IP packet. The final IP packets are sent to each client through IP network.
FIGURE 2. The architecture of the MPEG-2 AAC bitstream construction between client and server.

CONCLUSIONS AND FUTURE WORKS

In this paper, we presented an implementation of a real-time AAC streaming system. We can conclude that our streaming system works properly in intranet and Internet environment. Using this system, high quality audio broadcasting is possible. To improve it, we are considering the including of RTSP (Real-Time Streaming Protocol) protocol for allowing the control capability and researching QoS mechanisms.

ACKNOWLEDGEMENTS

This work is supported by the Ministry of Science and Technology of the Korean Government, under the title of “Digital Audio Broadcasting Technology Development”.

REFERENCES

1. ISO/IEC IS 13818-7 Generic Coding of Moving Pictures and Associated Audio. Part-7 Advanced Audio Coding (AAC)
Integral equation method of solving transmission problems for Helmholtz equation in angular domains.

Yu. Podlipenko

Faculty of Cybernetics, Kiev University, Vladymyrska st. 64, Kyiv, Ukraine

We are concerned with the following transmission problem that arise when studying diffraction of time harmonic acoustic waves by a penetrable homogeneous infinite cylinder with cross-section of an arbitrary shape located inside the homogeneous wedge with the edge parallel to the axis of the cylinder. We develop the potential theory enabling the reduction of the above problem to a system of two one-dimensional Fredholm integral equations on the boundary of the cross-section of the cylinder with the kernels having at most logarithmic singularities and establish existence and uniqueness of solutions both to this system of integral equations and to the transmission problem. We represent the kernels of integral equations as rapidly convergent series. This approach also provides a numerical method for solving this problem.

The time harmonic acoustic scattering by an arbitrary penetrable body contained within the wedge leads to a corresponding three-dimensional transmission problem which is also studied. Such investigation is based in particular on reduction of this problem to some system of two integral equations with weakly singular kernels on the body’s surface by using a corresponding linear combination of single and double layer potentials.

Let us consider in detail only the first one of the above problems. Introduce in the space $\mathbb{R}^3$ the cylindrical coordinate system $r, \phi, z$ and denote by $W = \{(r, \phi, z) \in \mathbb{R}^3 \mid r > 0, 0 < \phi < \Phi, -\infty < z < +\infty\}$ a wedge with sound-soft sides and the spread angle $\Phi$, $0 < \Phi \leq 2\pi$ and with the edge coinciding with the $z$-axis. Denote by $C$ an infinite cylinder with the axis parallel to the edge of the wedge. The domains $W \setminus C$ and $C$ are supposed to be filled by homogeneous isotropic mediums which are characterized by positive wavenumbers $k_1$ and $k_2$, densities $\rho_1$ and $\rho_2$ and sound velocities $c_1$ and $c_2$ respectively.

Let $\Omega$ and $D$ be the domains obtained as a result of intersection of the wedge $W$ and the cylinder $C$ with the plane $z = 0$ respectively. We make the assumption that $\overline{D} \subset \Omega$, and the domain $D$ is bounded, simply connected and has $C^2$ boundary $\partial D$. Denote by $\partial \Omega$ the boundary of the angular domain $\Omega$.

Let a line source of a time harmonic cylindrical wave be an infinitely long thin pulsing cylinder which is located in the domain $W \setminus C$ parallel to the edge of the wedge.

Consider the problem of diffraction of the acoustic waves excited by this source on the cylinder $C$ contained in the wedge with sound soft sides (the case of sound hard wedge’s sides or the case when one side is sound soft and another side is sound hard may be considered in a similar way). This problem is to determine the functions $p^{(s)}(M), M \in \Omega \setminus D$ and $p^{(i)}(M), M \in D$, that do not depend on $z$ and have the meaning of the space part of a time harmonic scattered and transmitted acoustic pressure respectively which we now relabel $p_1(M)$ and $p_2(M)$ such that

\[
\Delta p_1(M) + k_1^2 p_1(M) = 0, \quad M \in \Omega \setminus D, \quad (1)
\]
\[
\Delta p_2(M) + k_2^2 p_2(M) = 0, \quad M \in D, \quad (2)
\]
\[
p_1 - p_2 = f, \quad \lambda_1 \frac{\partial p_1}{\partial \nu} - \lambda_2 \frac{\partial p_2}{\partial \nu} = g \quad \text{on} \quad \partial D, \quad (3)
\]
\[
p_1 = 0 \quad \text{on} \quad \partial \Omega, \quad (4)
\]
\[
\frac{\partial p_1(r_M, \phi_M)}{\partial r_M} - ik_1 p_1(r_M, \phi_M) = 0 \left( \frac{1}{\sqrt{r_M}} \right), \quad r_M \to \infty, \quad (5)
\]

uniformly in $\phi_M$;

\[
\int_{\partial \Omega \setminus \delta} (|p_1|^2 + |\nabla p_1|^2) \, dS < \infty. \quad (6)
\]

Here $\lambda_j = \frac{1}{\rho_j}, k_j = \frac{c_j}{\rho_j}, j = 1, 2$, $\omega$ is a frequency of change of the field (we assume that the wave process depends on time in the form $e^{-i\omega t}$), $\Delta_{M, \phi_M} = \frac{1}{r_M^2} \times \frac{\partial}{\partial r_M} \left( r_M \frac{\partial}{\partial r_M} \right) + \frac{1}{r_M^2} \frac{\partial^2}{\partial \phi_M^2}$ is the Laplace operator in polar coordinates, $\delta$ is a certain neighbourhood of the origin of the polar coordinate system, $\nu$ is the unit normal to $\partial D$ drawn in the direction from $D$ to $\Omega \setminus D$, $f = -p^{(s)}|_{\partial D}$, $g = -\lambda_1 \frac{\partial p^{(s)}}{\partial \nu}$, $p^{(i)}(M) = -ik_1 p_1 c_1 G_{k_1}(M, M^*)$, $M \in \Omega$, where the function $G_{k_1}(M, N^*)$ is defined by the formula $(\cdot, N^*) = (r_{N^*}, \phi_{N^*}) \in \Omega \setminus D$ is the point of intersection of
the line source with the plane \( z = 0 \), \( Q \) is a productivity of the source.

Note that the function \( p^{(1)}(M) \) is the space part of acoustic pressure of the field excited by the line source in the wedge \( W \) (i.e., in the absence of cylindrical inhomogeneity), the radiation condition (5) excludes the waves coming from infinity, the condition (6) ensures the absence of energy flux from the edge of the wedge, the boundary condition (4) expresses the case when wedge’s sides are sound–soft and the transmission condition (3) guarantees continuity of the space parts of pressure \( p^{(1)}(M) + p_1(M) \) and \( p_2(M) \) and the corresponding normal components of velocity of total acoustic field in the domains \( W \setminus C \) and \( C \) respectively while crossing the lateral area of the cylinder \( \partial C \).

In all what follows we consider a more general transmission problem: find functions \( p_1 \) and \( p_2 \) satisfying (1)-(6) where \( f \) and \( g \) are prescribed continuous functions on \( \partial D \), and \( k_1, k_2, \lambda_1, \lambda_2 \) are non-zero complex numbers with \( 0 \leq \arg k_1 < \pi/2 \), \( 0 \leq \arg k_2 < \pi/2 \). This allows to embrace the case of absorbing mediums as well as the case when an incident field is excited by a distribution of line sources located in the wedge parallel to its edge.

Note that the classical transmission problem in free space has been investigated by Kupradze, Werner, Kress and Roach, Kleinman and Kittappa and others.

Our aim is to reduce the transmission problem (1)-(6) to a system of two one-dimensional Fredholm integral equations that is suitable for numerical computation. To this end, we look for a solution of this problem in the form of potentials

\[
p_1(M) = \lambda_1^{-1} \int_{\partial D} \left( \frac{\partial G_{k_1}(M, N)}{\partial \nu_N} \phi(N) + c_1 G_{k_1}(M, N) \psi(N) \right) d\nu, \quad M \in \Omega \setminus \overline{D} \tag{7}
\]

\[
p_2(M) = \lambda_2^{-1} \int_{\partial D} \left( \frac{\partial G_{k_2}(M, N)}{\partial \nu_N} \phi(N) + c_2 G_{k_2}(M, N) \psi(P) \right) d\nu, \quad M \in \Omega \tag{8}
\]

with densities \( \phi \in C^{1,0}(\partial D), \psi \in C^{0,0}(\partial D) \) where \( c_1, c_2 \in \mathbb{C} \setminus \{0\} \) are fixed constants and the functions \( G_{k_j}(M, N), \quad j = 1, 2 \), are defined by the formulas

\[
G_{k_j}(M, N) = \frac{i\pi}{\Phi} \sum_{m=1}^{\infty} J_{\nu_m}(k_1 \min(r_M, r_N)) \times \frac{H^{(1)}_{\nu_m}(k_1 \max(r_M, r_N)) \sin(\nu_m \phi_M) \sin(\nu_m \phi_N)}{r_M^2 + r_N^2 - 2r_Mr_N \cos(\phi_M - \phi_N)}, \tag{9}
\]

\[
G_{k_2}(M, N) = \frac{i}{4} H^{(1)}_0(k_2 \sqrt{r_M^2 + r_N^2 - 2r_Mr_N \cos(\phi_M - \phi_N)}),
\]

in which \( J_\nu(z) \) and \( H^{(1)}_\nu(z) \) are the Bessel and the Hankel functions of order \( \nu \), \( \nu_m = m\pi/\Phi, \quad m = 1, 2, \ldots, M = (r_M, \phi_M) \in \Omega, \quad N = (r_N, \phi_N) \in \Omega, \quad M \neq N \).

Using the asymptotic expansions of Bessel and Hankel function of large order, we extract a logarithmic singularity from the functions \( G_{k_i}(M, N) \) and \( \frac{\partial^2 G_{k_i}(M, N)}{\partial \nu_M \partial \nu_N} \), when their arguments coincide, and simultaneously obtain the representations that will make it possible, on the one hand, to compute these functions as well as the functions \( \frac{\partial G_{k_i}(M, N)}{\partial \nu_N} \) and \( \frac{\partial G_{k_i}(M, N)}{\partial \nu_M} \) in the case when their arguments are close and, on the other hand, to apply the theorems on the jump and regularity properties for harmonic potentials.

We prove that the potential (7) satisfies the radiation condition (5), the condition on the edge (6) and vanishes on \( \partial \Omega \).

Applying the potential theory arguments, we establish the following results:

1. The potentials \( p_1 \) and \( p_2 \) defined by formulas (7) and (8) solve the transmission problem (1)-(6) provided the densities \( \phi \) and \( \psi \) solve the following system of integral equations with weakly singular kernels

\[
\frac{1}{2}(\lambda_1^{-1} + \lambda_2^{-1})\phi(M) + \int_{\partial \Omega} \left( \lambda_1^{-1} \frac{\partial G_{k_1}(M, N)}{\partial \nu_N} - \lambda_2^{-1} \frac{\partial G_{k_2}(M, N)}{\partial \nu_N} \right) \phi(N) d\nu_N - \int_{\partial \Omega} (\lambda_1^{-1} c_1 G_{k_1}(M, N) - \lambda_2^{-1} c_2 G_{k_2}(M, N)) \psi(N) d\nu_N = f(M), \tag{11}
\]

\[
\frac{1}{2}(\lambda_1^{-1} + \lambda_2^{-1})\psi(M) - \int_{\partial \Omega} \left( \frac{\partial^2 G_{k_1}(M, N)}{\partial \nu_M \partial \nu_N} - \frac{\partial^2 G_{k_2}(M, N)}{\partial \nu_M \partial \nu_N} \right) \psi(N) d\nu_N - \int_{\partial \Omega} (c_1 G_{k_1}(M, N) - c_2 G_{k_2}(M, N)) \phi(N) d\nu_N = -g(M), \quad M \in \partial D. \tag{12}
\]

2. Under certain restrictions on the numbers \( \lambda_1, \lambda_2, k_1, k_2, c_1 \) and \( c_2 \), we show that the system of integral equations (11), (12) has a unique solution \( \phi \) and \( \psi \) and the potentials (7), (8) with densities \( \phi \) and \( \psi \) solve the transmission problem (1)-(6).