LOUDSPEAKER FAULT DETECTION USING TIME-FREQUENCY REPRESENTATIONS

Manuel Davy

Signal Processing Group, Univ of Cambridge
Dept of Engineering, Trumpington Street
CB2 1PZ Cambridge, UK
md283@eng.cam.ac.uk

Hélène Cottereau, Christian Doncarli

IRCCyN, 1 rue de la noë, BP92101
44321 Nantes Cedex 3 - FRANCE
{cotter,doncarli}@irccyn.ec-nantes.fr

ABSTRACT

This paper addresses the problem of loudspeaker test. Two applications are investigated: loudspeaker manufacturing fault detection and maintenance. In order to comply with practical requirements, we propose a new, brief, nonstationary test signal. Recorded loudspeaker responses are processed using an improved time-frequency decision algorithm. The fault detection procedure is tested with real data. Results show its accuracy and practical interest.

1. INTRODUCTION

[1]Sound system loudspeaker fault detection is a twofold problem: manufacturing and maintenance. Manufacturers need to detect any fault before selling the product; users have to maintain the loudspeaker (e.g. replace a faulty tweeter). The fault detection problem can be stated as follows: is it possible to detect any loudspeaker fault (manufacturing test), and measure its importance (maintenance test)? In this paper, we propose a satisfactory solution, consisting in comparing – in the Time-Frequency (TF) domain – the response of a given loudspeaker to a set of ”correct” loudspeaker responses.

Classical loudspeaker tests are e.g. on-axis, impedance or impulse responses, and lead to the estimation of Thiele and Small parameters [1]. This implicitly relies on the linearity assumption: loudspeakers can be modelled as linear filters, which are completely characterized by a transfer function. These measurement techniques are efficient for loudspeaker design, besides which additional techniques such as membrane displacement laser measurement can be used.

Many fault detection tests use the same kind of measurements: on-axis and impedance response, but also human earing. However, the aim is different from loudspeaker design, and, in this context, these techniques have drawbacks. On-axis, impedance and impulse (or step) responses can actually be interpreted *only* in case of:

- stationary signals: the test signal has to be a stationary tone of given frequency (or a collection of tones).
- linearity.

However, a loudspeaker fault can result in nonlinearity, in which case the interpretation of these measurements fails. An important conclusion is then that design measurements are not, from a theoretical point of view, adapted to fault detection. This point is experimentally confirmed in the classical problem of detecting the vibration caused by a small object (e.g. a wire) touching a membrane (a problem known as *rub and buzz*): most classical tests fail in detecting this noisy vibration!

In this paper we propose a new fault detection test based on a nonstationary test signal and TF signal classification/detection. This non-parametric technique is completely model-free and has been applied to, e.g., the analysis of impulse responses [2]. (Waterfall charts are also basic Time-Frequency Representations (TFRs).) Using TFRs does not require the definition of a signal model ; the linearity assumption is then not necessary. Moreover, TFRs are suited to the analysis of nonstationary signals, and we will show all the loudspeaker bandwidth can be processed using a short duration ($<$ 100 ms) test signal.

In section 2, we present a new, multicomponent FM test signal. In section 3, an improved TF signal classification/detection method is proposed [3]. In section 4, results obtained with real data are exposed. We emphasize the robustness of the proposed procedure — in particular, it can be implemented in a noisy environment. Finally, some conclusions are given, and future developments explained.

2. A TEST SIGNAL

Classical on-axis response measurements consist in measuring the acoustic pressure at many stationary frequencies in the loudspeaker frequency range (20Hz-20kHz), and plot the log-log response. The proposed TF signal classification/detection method enables the use of nonstationary test signals, which can explore the loudspeaker frequency range within a short duration.

2.1. Multicomponent FM test signal

The proposed test signal is made of 4 FM components of constant amplitude, in the range 20Hz–10kHz. (The range 10kHz-20kHz does not contain crucial information for sound system loudspeaker fault detection. The choice of this upper frequency permits a slower sampling rate and limits the number of data to be processed.) Figure 1 (left) represents the spectrogram of such a test signal. Signal duration is 92.9ms (2048 points at sampling rate 22050Hz). Frequency ranges of FM components do not overlap. The slope of each component depends on its frequency: low frequencies (i.e. long periods) require more time: the slope is smaller.

Direct use of the test signal may however cause problems in its beginning and end, caused by amplitude jumps. These problems can be avoided by merging begin and end parts. The final signal

[1]This work was mainly performed while M. Davy was still at IRCCyN.

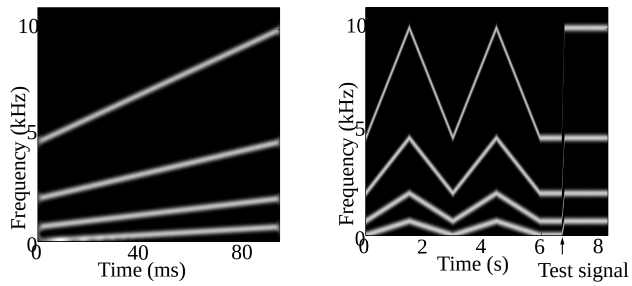

Fig. 1. Left: Spectrogram of the proposed test signal (Hamming window, 101 points). Right: Spectrogram of the complete test signal. The test part (see left) location is indicated by the arrow.

is then composed of three parts: heating part, test signal itself, and ending part. Figure 1 (right) displays the complete signal.

2.2. Recording procedure

The test procedure consists in recording the response of the loudspeaker emitting the amplified test signal. The response is recorded using a microphone and a computer sound card.

The recorded loudspeaker response is preprocessed in order to extract the test part (i.e. eliminate heating and end parts). Since the complete signal is perfectly known, we proceed as follows:

- To detect the beginning of the complete emitted signal (i.e. the *initial point* of the heating part);
- To extract the test part by counting the time points after the initial point.

The initial point is detected by comparing the instantaneous amplitude of the response to a threshold $\mathcal{E}$. The time point when the amplitude becomes higher than $\mathcal{E}$ is decided to be the initial point.

Figure 2 displays the example response of a sound system loudspeaker. Note that the amplitude of the FM components is made time-varying by the loudspeaker. The amplitude modulation can be interpreted as the loudspeaker signature. Processing such a response reveals possible faults, as shown in Section 4. In particular, the presence of energy aside the FM components (provoked by, e.g., mechanical vibrations) will be detected.

3. TF LOUDSPEAKER FAULT DETECTION

Processing loudspeaker responses requires specific methods, since e.g. the frequency spectrum may not be a relevant (it does not discriminate correct loudspeakers responses from faulty ones in presence of, e.g., rub and buzz problems).

TFRs are particularly suited to the analysis of multicomponent FM signals with time-varying amplitude: Firstly, TFRs are non-parametric which implies that no amplitude model has to be defined; Secondly, Cohen's class [4, 5] TFRs are adapted to FM signals[2]. In this section, we first introduce some elements of TF classification/detection, then we explain how this technique can be applied to loudspeaker fault detection.

[2]The Wigner-Ville distribution (a well-known TFR in Cohen's class) is perfectly localized for FM signals.

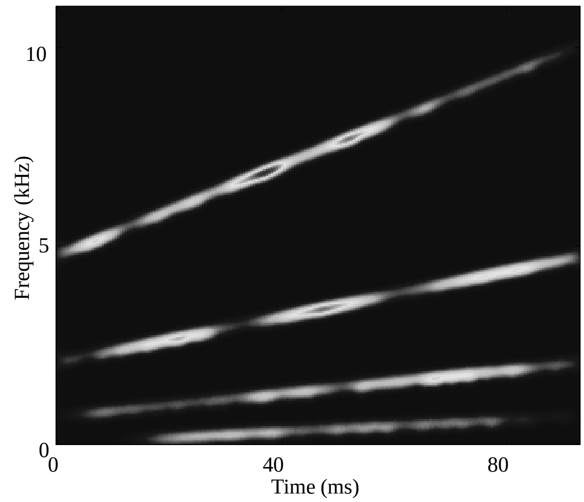

Fig. 2. Spectrogram of the recorded response of a NEXO-PS15 loudspeaker emitting the multicomponent FM test signal.

3.1. TF classification/detection

TF classification/detection theory has known recent advances [3, 6] that lead to a higher efficiency. The principle of TF methods consists of comparing TFRs (defined by a given kernel within the Cohen's class) using a TF distance measure. A learning set composed of labelled signals is used to define a representative element. We denote by $\mathcal{X}_i = \{x_1^i, \ldots, x_{n_i}^i\}$ the learning set corresponding to the class ω_i, composed of n_i learning signals. There are c classes. In the particular case of loudspeaker fault detection, the learning set is composed of several classes:

- The class of correct loudspeaker responses;
- the classes corresponding to given loudspeaker faults.

Both classes are identified by an expert.

In the following, the TFR of a signal $x(t)$, defined by a kernel ϕ, is denoted $\mathcal{C}_x^\phi(t,f)$ where t and f are the time and frequency variables.

3.1.1. TF decision rules

Given a TFR kernel ϕ, the classification of a given signal in one of the c classes relies on the following decision rule:

$$x \text{ is assigned to } \omega_i \;\Leftrightarrow\; i = \operatorname*{argmin}_{j=1,\ldots,c} d(\mathcal{C}_x^\phi(t,f), \overline{\mathcal{C}}_j^\phi(t,f)) \quad (1)$$

where d is a TF distance as defined below, $\mathcal{C}_x^\phi(t,f)$ is the TFR of the signal to be classified and $\overline{\mathcal{C}}_i^\phi(t,f)$ is the representative TFR:

$$\overline{\mathcal{C}}_j^\phi(t,f) \;=\; \frac{1}{n_j}\sum_{k=1}^{n_j} \mathcal{C}_{x_k^j}^\phi(t,f) \quad (2)$$

In other words, x is assigned to the class corresponding to the nearest representative TFR.

Sometimes, however, the learning set is only composed of the class of *correct* loudspeaker responses[3], and the general class of

[3]There are many more correct loudspeakers than faulty ones.

not correct loudspeaker responses. The latter is rather a *reject class* than a real class. In this case, there is only one representative (denoted $\overline{\mathcal{C}}^{\phi}_{\text{correct}}(t,f)$, and the decision rule consists of deciding whether a given loudspeaker is correct, or not:

$$\begin{array}{lll} x \text{ is decided } \textit{correct} & \Leftrightarrow & d(\mathcal{C}^{\phi}_x(t,f), \overline{\mathcal{C}}^{\phi}_{\text{correct}}(t,f)) < \eta \\ x \text{ is decided } \textit{not correct} & & \text{otherwise} \end{array} \tag{3}$$

where η is a given threshold. The distance $d(\mathcal{C}^{\phi}_x(t,f), \overline{\mathcal{C}}^{\phi}(t,f))$ indicates the importance of the loudspeaker fault.

3.1.2. *TF distance measures*

There exists many TF distance measures. In this paper, we consider the Kolmogorov distance:

$$d_{\text{Kolm}}(\mathcal{C}^{\phi}_x, \overline{\mathcal{C}}^{\phi}_j) = \iint \left| N\mathcal{C}^{\phi}_x(t,f) - N\overline{\mathcal{C}}^{\phi}_j(t,f) \right| dtdf \tag{4}$$

where the notation $N\mathcal{C}^{\phi}_x(t,f)$ emphasizes the prior normalization of the TFR:

$$N\mathcal{C}^{\phi}_x(t,f) = \frac{\left|\mathcal{C}^{\phi}_x(t,f)\right|}{\iint \left|\mathcal{C}^{\phi}_x(t,f)\right| dtdf} \tag{5}$$

This normalization enables the use of distance measures initially devoted to probability densities comparison. Other distance measures can be found in [7]. The correlation distance is often proposed for TFR comparison and is written:

$$d_{\text{Corr}}(\mathcal{C}^{\phi}_x, \overline{\mathcal{C}}^{\phi}_j) = \iint \mathcal{C}^{\phi}_x(t,f)\, \overline{\mathcal{C}}^{\phi}_j(t,f)\, dtdf \tag{6}$$

The choice of the distance measure can result from an optimization procedure as explained in [3]. However, the Kolmogorov distance is preferred here for its simplicity and accuracy. From a similar point of view, the chosen TFR kernel is the Spectrogram, with Hamming window (101 points), denoted $\mathcal{S}_x(t,f)$ when computed for a signal x.

3.2. Application to loudspeaker fault detection

The resulting detection algorithm is composed of two steps. The *learning step* consists in:

1. Recording *correct* and *not correct* loudspeaker responses;
2. Preprocessing these records in order to extract the test part;
3. Computing the TFRs of the *correct* responses and computing the representative TFR $\overline{\mathcal{C}}^{\phi}_{\text{correct}}(t,f)$ of the *correct* loudspeaker responses, using eq. (2).

The *test step* consists in comparing the response x of a given loudspeaker to the *correct* representative TFR, as follows:

1. Preprocess the recorded response in order to extract the test part, denoted x;
2. Compute its TFR $\mathcal{C}^{\phi}_x(t,f)$;
3. Compute the Kolmogorov distance $d_{\text{Kolm}}(\mathcal{C}^{\phi}_x, \overline{\mathcal{C}}^{\phi}_{\text{correct}})$ using eq. (4);
4. Compare this distance to a threshold η and decide (using eq. (3)). This distance also indicates the importance of the fault.

In the following section, results obtained using the proposed test signal and TF fault detection procedure are given.

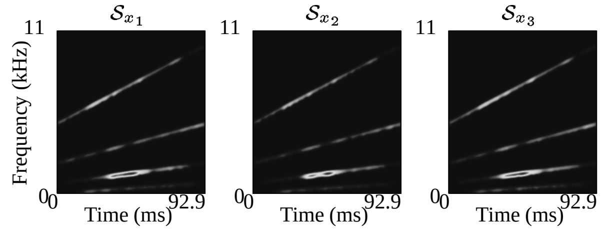

Fig. 3. Spectrograms of three responses of the same loudspeaker.

4. RESULTS

We present results obtained with NEXO-PS15 (2-way active) loudspeakers. The test signal was emitted by a DAT reader, and recorded using a microphone LEM EMU 4535 and a SoundBlaster sound card. The responses have been recorded in a hall, without special acoustic environment: noise and speech may corrupt the records, with a relative small energy however. The microphone position is 80 cm from the loudspeaker, its heights corresponds to the middle point between the tweeter and the bass cone drive.

In order to evaluate the robustness, and the practical interest of the method, three tests have been run. The questions investigated were: First, is the recorded response stable from one test to another? Second, is the microphone position crucial? Third, is it possible to discriminate new loudspeakers from old ones, and evaluate old loudspeakers fatigue?

4.1. Stability

Stability is a crucial point. Actually, a test procedure that would lead to different results for the same loudspeaker would not be accurate.

This point is investigated as follows. A loudspeaker is tested three times, the responses are denoted (x_1, x_2, x_3). Spectrograms $(\mathcal{S}_{x_1}, \mathcal{S}_{x_2}, \mathcal{S}_{x_3})$ and the representative TFR $(\overline{\mathcal{S}}_{\text{same}} = (\mathcal{S}_{x_1} + \mathcal{S}_{x_2} + \mathcal{S}_{x_3}))/3)$ are computed. Figure 3 displays $\mathcal{S}_{x_1}$, $\mathcal{S}_{x_2}$ and $\mathcal{S}_{x_3}$. Table 1 displays the Kolmogorov distance between these three responses and the representative TFR. These distance are very small (compared with the distances plotted Figure 4), which confirms the stability of the proposed procedure.

$d_{\text{Kolm}}(\mathcal{S}_{x_1}, \overline{\mathcal{S}}_{\text{same}})$	$d_{\text{Kolm}}(\mathcal{S}_{x_2}, \overline{\mathcal{S}}_{\text{same}})$	$d_{\text{Kolm}}(\mathcal{S}_{x_3}, \overline{\mathcal{S}}_{\text{same}})$
0.0675	0.0754	0.0590

Table 1. Distance between three response spectrograms and the representative TFR.

4.2. Microphone positioning errors

The next important question concerns the robustness of the test procedure in front of microphone positioning errors. We have tested different microphone locations around a central location, in the six directions (up, down, left, right, forward, backward) of 1cm, then of 5cm. Table 2 displays the test results (distance between the spectrogram of the response recorded at a given location, and the representative spectrogram). For each displacement length (1cm and 5cm), the representative TFR is computed as the average

Displacement	Forward	Backward	Left	Right	Up	Down
1cm	0.0469	0.0549	0.0609	0.1280	0.0865	0.1209
5cm	0.1697	0.1713	0.1566	0.1757	0.2029	0.2887

Table 2. Distance between loudspeaker responses recorded at different microphone locations around a central position, and the representative spectrogram.

of the six spectrograms, corresponding to the six records. A 1cm positioning error results in a small variation (compare with results of Table 1 and Figure 4). Displacements of 5cm causes more important errors. In practice, positioning the microphone within 1cm is easy (using e.g. a calibrated test bench) and enables accurate loudspeaker response processing.

4.3. Fault detection

A set of 7 new and 14 old loudspeakers is tested. The aim is here to recognize old loudspeakers and evaluate their state of fatigue. The loudspeaker responses are respectively denoted $(x_1, x_2, \ldots, x_7)$ and $(y_1, y_2, \ldots, y_{14})$. Of course, 21 responses is not a large enough set to enable straightforward processing, however a *leave one out* procedure can be implemented as follows:

1. Choose one of the 7 new loudspeaker responses. This response is denoted $\widehat{x}$. The learning set consists of the 6 remaining responses. The representative $\overline{\mathcal{S}}_{\text{new}}$ is computed as the average of the 6 corresponding spectrograms.
2. Compute $d_{\text{kolm}}(\mathcal{S}_{\widehat{x}}, \overline{\mathcal{S}}_{\text{new}})$;
3. Compute $d_{\text{kolm}}(\mathcal{S}_{y_j}, \overline{\mathcal{S}}_{\text{new}})$, $j = 1 \ldots 14$;
4. Choose another new loudspeaker (another $\widehat{x}$) and iterate from step 1.

Figure 4 displays the computed distances as plots. Each sub-plot corresponds to a different $\widehat{x}$. Its distance to the representative is plotted as a black square. The 14 distances corresponding to old loudspeakers are plotted as black circles.

Three main observations are: First, it is possible to define a threshold $\eta = 0.15$, i.e. it is possible to discriminate between new and old loudspeakers. Second, one old loudspeaker is positioned at the distance $d_{\text{Kolm}} \simeq 0.5$. This loudspeaker reveals to have a blown fuse. Third, new loudspeaker #5 is above the threshold (classified as old). Further examination reveals a fault (slightly De-centered tweeter membrane), which explains this classification result.

5. CONCLUSION

The proposed nonstationary test signal, combined with the TF classification procedure, enables the detection of loudspeaker faults as well as an evaluation of loudspeaker fatigue. Its efficiency and practical interest is shown: In particular, price test components and any acoustic environment match. This procedure could also be applied to loudspeaker objective quality evaluation (by comparing to a reference). Moreover, the accuracy can be improved by TFR kernel optimization, as explained in [3].

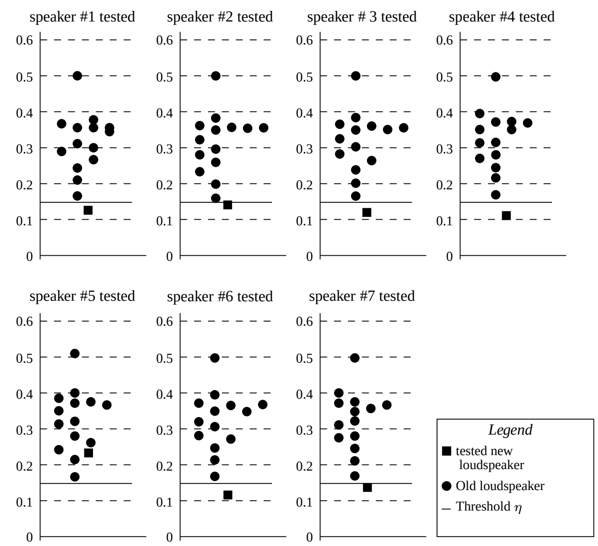

Fig. 4. Results of the fault detection test applied to 7 new loudspeakers and 14 old ones. Vertical axis : Kolmogorov distance to the representative TFR. Horizontal axis: no particular information (used to improve the readability).

6. REFERENCES

[1] J. D'appolito, *Testing loudspeakers*, Audio Amateur, 1998.

[2] C.P. Janse and A.M. Kaiser, "Time-frequency distributions of loudspeakers: The application of the Wigner distribution," in *Audio Engineering Society 71st Convention*, Montreux, Switzerland, 1982.

[3] M. Davy and C. Doncarli, "Improved optimization of time-frequency based classifiers," *IEEE Signal Processing letters*, To be published, January 2001.

[4] F. Hlawatsch and G.F. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations," *IEEE SP Magazine*, pp. 21 – 67, April 1992.

[5] P. Flandrin, *Time-Frequency/Time-Scale Analysis*, Academic Press, 1999.

[6] C. Richard and R. Lengellé, "Data driven design and complexity control of time frequency detectors," *Signal Processing*, vol. 77, no. 1, pp. 37–48, 1999.

[7] M. Basseville, "Distance measures for signal processing and pattern recognition," *Signal Processing*, vol. 18, no. 4, pp. 349 – 369, 1989.