SOUNDS DYNAMIC FILTERING DEPENDING ON SOURCE PLACEMENT

Cătălin Chivu

"Transilvania" University of Brasov, Romania, Department of Economic Engineering and Manufacturing Systems, Str. Parcul Mic, nr.13, bl.110, ap.31, Brasov, Romania,
e-mail: chivuc@acasa.ro

Abstract: This paper tries to remind to those that develop speech recognition engines the possibility and importance of separating sounds depending on the spatial position of the sources that generate them.

Keywords: signal processing, directional microphones, sound localization

1. SEPARATING SOUNDS DEVICE

The original idea presented here represents a device designed to be capable to separate sounds depending on the spatial position of the generator sources or, more precise, depending on the direction where they come from.

The proposed assembly is compound from the following modules:

a) Microphones module
b) Analogical module that amplifies and optional filters the signals
c) Signal acquisition module
d) Signal digital processing module

a) Microphone module is an assembly that has four microphones: one omnidirectional and three bipolar. Their spatial displacement is extremely important in determining the direction of sound propagation (figure 1). Distance between microphones is recommended to be as small as possible, thus the phase difference being minimized.

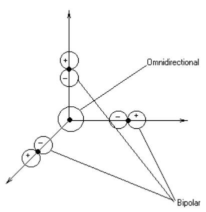

Fig. 1 Microphone module

b) Amplifying analogical module is represented by four identical operational amplifiers that perform a signal level adaptation between microphones and acquisition module. The frequency filters (optional ones) are designed so that they are capable to do the spectral correction of the signals obtained from the four microphones (especially those from the three bipolar microphones).

c) The signals acquisition module uses four acquisition channels and includes antialiasing filters that generate a band limiting. The needed sampling frequency results from the Nyquist condition applied to the upper limit of the desired spectrum (the cut-off frequency of the antialiasing filters). The analog-digital conversion is linear one using converters as precise as possible.

d) The signal digital processing module is represented by a calculus system (a PC with computing power high enough to execute, in real time, all the needed operations) that has oriented programs for signal processing and speech recognition. This digital signal processing module has to determine the propagation direction of the sounds and to extract, from the main spectrogram, the undesired spectral energies (disturbances) based on the criteria that says that the undesired signal comes from distinctive directions.

1.1. System description

The system is in fact a dynamic frequency filter that is capable to eliminate the undesired spectral components depending on their incidence angle (depending on their source placement). Consequently, the filter is capable to correct the main signal spectrogram thus this spectrogram will characterized as well as possible the signal came from a known direction.

For a sinusoidal sound signal with constant amplitude it can be determined: signal energy, **sa**, received by the omni-directional microphone **Esa0** and the energy of **sa0-sax, sa0+sax, sa0-say, sa0+say, sa0-saz, sa0+saz** signals that are **E(sa0-sax), E(sa0+sax), E(sa0-say), E(sa0+say), E(sa0-saz), E(sa0+saz)** where **sax, say, saz** are the signal received by the three bipolar microphones.

By norming the **E(sa0-sax)** signal energy to **Esa0** results a parameter for energy deflection, which is specific to the bipolar microphone that has an particular incidence angle **αa+** for the sound signal. This angle corresponds to the positive sense of Ox axis. Similarly, by norming the **E(sa0+sax)** signal energy to **Esa0** results a parameter that corresponds to an particular incidence angle **αa-** (negative sense of Ox axis.)

Depending on **E(sa0-sax)** and **E(sa0+sax)** it can be decided which value, **αa+** or **αa-**, corresponds to reality (if **E(sa0+sax)>E(sa0-sax)** then the signal is emitted from the positive sense of Ox and if **E(sa0+sax)<E(sa0-sax)** - from the negative sense).

The above parameter for the energy deflection results also from the rejection characteristic of the bipolar microphone depending on the signal frequency and its incidence angle. The rejection characteristic of the microphone depends on the microphone model and it is given in the producer catalogue. Otherwise, it can be obtained by lab determinations.

By norming the spectrograms of **s0-sx, s0+sx, s0-sy, s0+sy, s0-sz, s0+sz** signals to that of **s0** signal result the deflections of spectral energy specific to the bipolar microphones for **α, β** and **γ** incidence angles. In the end it can be determined the azimuth and elongation of **sa** signal source.

The system doesn't distinct direction of the source of two sinusoidal signals that have the same frequency but different sources. The result of the above calculus will give the azimuth and elongation of the gravity center of the system determined by the two signal sources.

2. CONCLUSIONS

If there are known the azimuth and elongation of the three incidence angles **α, β** and **γ** of the analyzed signal it can be deduce a correction rule for the main signal spectrogram (received by the central omni-directional microphone) depending on the direction deflection

$$(cos(\alpha-\alpha_a))^2+(cos(\beta-\beta_a))^2+(cos(\gamma-\gamma_a))^2 \qquad (1)$$

For this correction can be chosen different types of characteristics (figure 2)

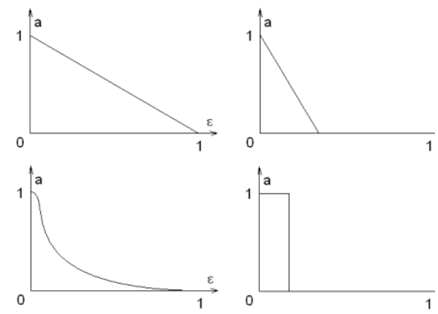

Fig.2 Characteristics of correction

where $$\varepsilon=(cos(\alpha-\alpha_a))^2+(cos(\beta-\beta_a))^2+(cos(\gamma-\gamma_a))^2 \qquad (2)$$

is the incidence angular error of the **fa** frequency signal and **a** is the correction coefficient applied as amplification to **fa** frequency energy from the main signal spectrogram.

Depending on the chosen characteristic is obtained a directional filter specific to the application where is used or to a specific environment of one application. For example, for the speech recognition applications there can be deduced characteristics that give better results in conditions of street noises rather than for noises from a conference hall or otherwise. There can be obtained filters that have better results for spatial dynamic noises than for spatial static ones or vice-versa.

There can be also considered characteristics or types of distinctive characteristics for particular frequency domains.

Also as a conclusion it should be remarked that the position of the analyzed signal source can be dynamically changed thus a fixed system like this, can separate the voice of a person that is moving around the system from noise. This can be done by changing the parameters, **α, β** and **γ**, by following the source of the main signal (the speaker voice).

The sounds filtering method that use the direction from where the sounds come offers to the specialists that develop the speech recognition engines a very precious tool because it gives the opportunity to select only the desired signals, by rejecting the disturbances that have, generally speaking, a source different from that of the desired sound.

REFERENCES

[1] Alvarado, V. M. and Silverman, H. F. (1990). Experimental results showing the effects of optimal spacing between elements of a linear microphone array. In Proceedings of the 1990 International Conference on Acoustics, Speech, and Signal Processing, pages 837–840, Albuquerque, New Mexico. Institute of Electrical and Electronic Engineers.
[2] Che, C., Lin, J., Pearson, J., de Vries, B., and Flanagan, J. (1994). Microphones arrays and neural networks for robust speech recognition. In Proceedings of the 1994 ARPA Human Language Technology Workshop, Princeton, New Jersey. Advanced Research Projects Agency, Morgan Kaufmann.
[3] Peterson, P. M. (1989). Adaptive array processing for multiple microphone hearing aids. Technical Report 541, Research Laboratory of Electronics, MIT, Cambridge, Massachusetts.