M. Reuter & R.H. Kluwe
Inst. für Kognitionsforschung, Universität Bw Hamburg
D-2000 Hamburg 70, Holstenhofweg 85, Tel. 040/6541 2876
Telefax 040/6530413

FD-spectrums and their simulation by neural networks

F - AI,connectionist,networks

Introduction

The research reported here is directed at the elaboration of an efficient system for the classification and identification of acoustical signals. In this context,a new frequency representation of acoustical signatures has been developed which is invariant under fade out of frequency bands or under masking of single signature shares. This "metaform" of the signatures results from the projection of all frequencies that are parts of the single oscillations on their fundamental frequencies.

The spectral presentation in which one can find the metaforms is called FD-spectrum and can be generated by rather simple neural networks. Preliminary experiments with these networks suggest that neural modules are of central importance for the computation of these spectrums. This led to the assumption that specific neural substructures rather than single neurons are the fundamental components of an efficient network in pattern recognition.

The modification of the operators that generate socalled eFD- and mFD-spectrums allow the construction of frequency filters and of higher order presentations which can be chosen to select individual qualities for the classification and identification of acoustical sources.

Theoretical frame

We assume that a situation- and problemoriented presentation of acoustical signatures is a necessary processing stage for a classoriented recognition system, which is independent of the acoustical background. The problemoriented presentation is termed 'metaform' of the signature, because it depends only on the situation-independent information of the noise source.

The metaform of the signature will be elaborated by an invariant, domain-independent neural network, whereby the problem is represented in the structure of the network .

The desired feature-specific presentation of acoustical signatures should provide a correct classification and identification of the sound sources even if superposition-effects, scattering of frequencies or from frequencybands (telephon-effect) exist.

As all those effects can be conceived of as a kind of scattering it is just necessary to find a presentation where parts of the signature-frequency-spectrum are redundant. This will be the case if there exists a well-defined point on which all the parts of the signature of one oscillation can be duplicated. As the signature of an oscillation is composed of the fundamental oscillation frequency value and its harmonics, each of these frequency values can be the "representation-point". As all frequency-values can be represented by a frequency-distance one can scan the autopowerspectrum by this "clasp" to elaborate a size of the modulation of the spectrum by a frequency. Contrary to the usual procedures we chose for a "clasp-procedure" not a special starting point,but start successively from all values of the spectrum. The cumulated feature-component amplitudes then represent a hypothetical "total modulation" of the spectrum done by a frequency. This form of the feature-component is termed a second order feature-component .

From this special form of the new feature-component follows that the correlation of the amplitude-values with their frequency-values are no longer important for the computation of the feature-component what means that the system must not be able to count. This will be an important fact for the construction of a neural network.

If we consider spectrums which are nearly free from modulations created by the ground-noise or other acoustical jammings, than we can compute the second order feature-components not only for one special frequency-value but for an entire spectrum of frequency-values. The result is a frequency distribution which delineates in terms of a specific representation the fundamental frequency and their harmonics that occur in the signature in a way that is nearly invariant under frequency scatterings.

This special presentation is termed ***FD-spectrum***.

The FD-spectrum which is based one the additive combination of the amplitude values of the DLS is called simple FD-spectrum (eFD).

The FD-spectrum where a multiplicative combination of the amplitude values from the DLS is the case is called multiplicative FD-spectrum (mFD).

The computation rule for the eFD- and mFD-spectrums is given by the operators:

$$eFD := \Gamma^{+}(f,i,M,I)\ Amp(f_i)$$

respectively:

$$mFD := \Gamma^{*}(f,i,M,I)\ Amp(f_i)$$

where f is the smallest and M the largest clasp-width; i is the starting value and I the endpoint of the clasp-procedure.

To ensure that the FD-spectrum comprehends only information of the different acoustical sources one has to find a class-specific representation which divide the autopower-spectrum in a situation-specific and a source-specific spectrum.

This can be done if one calculates the graph of the medians adapted by a function of the form:

$$A(f) = (1-\exp(-t_1 * f)) * (a + b * \exp(-t_2 * f))$$

where the parameters t_1, t_2, a and b are situation-specific /h1,r1/.

If the approach is directed at a model in terms of neural networks a classorientied computation of the local medians has been applied.

This graph can be interpreted as the seperation-line of the two information units; therefore we can define it also as the zeroline of a new spectrum,which represents a feature-specific presentation of the autopower-spectrum.

This spectrum is termed ***difference-autopower-sepctrum (DLS)*** (see figure 1).

If only the fundamental frequencies shall be shown, one has to find again a feature-specific presentation. This can be done by the calculation of the eFD-spectrum of the FD-spectrums.

To point out the fundamental frequencies the operator has to be modified in the way that the higher harmonics are eliminated after the eFD is calculated. So the operator has the form:

$$\Gamma + (f_f,f_f,M,I) := \Gamma + (f_f,f_f,M,I)\ Amp(f_i) \Big|_{Amp(n^* f_f) = 0}$$

where f_f should be the fundamental freqnency value.

This special spectrum is termed ***TTY-spectrum*** here.

In order to calculate the TTY-spectrum from a DLS we will have the operator-combinations:

$$TTY(eFD) := \Gamma + (f',f,M,I)\Gamma + (f,i,M,I)\ Amp(f_i)$$

resp.

$$TTY(mFD) := \Gamma + (f',f,M,I)\Gamma^* (f,i,M,I)\ Amp(f_i)\ .$$

The computation of the FD-spectrum by neural networks

The process on thined above can be simulated by means of a neural networks.

Three kinds of neurons will be defined. They are distinguished according to their function:

(a) The first type symbolized by refers to a frequency-selective neuron;

(b) the second type symbolized by refers to an intensity-selective neuron;

(c) the third type, symbolized by refers to neuron with combined functions, i.e. an intensity-and-frequency-selective neuron.

In addition, the last neuron shall be able to treat incoming data alternatively: multiplicative or additive.

These specific neurons are hypothetical components, but our research so far suggests that those connections of a defined number of neurons in special formations or "working modules" seems to be a fundamental organization principle similar to the feature-specific presentation

By definition the positive autopower-spectrum (pLS) represents the incoming pattern which corresponds to the first level of the network. The neurons on this level are frequency-selective; therefore the pLS will be divided into several pLS-parts. This provides the necessary classes for the computation of the DLS.

The next level will be constituted by m intensity-selective neurons which are interconnected by inhibitor-axons. On this level the computation of the local mean-values of the amplitudes of the pLS-part is achieved. The neuron with the highest activation resulting from the pLS-part-amplitudes will have the lowest threshold of all the neurons on that level. By that way one ensure that only those intensities of the signature which are equal to or above the mean value will be passed to the next level. This is achieved by intensity-selective neurons that constitute also the last neural layer. The threshold of these neurons is determined temporarily by the neuron-specific connections of the neuron with the highest activation on the level before. This variable intensity filter will be the classorientied new zeroline of the DLS.

The following network has to compute the eFD- or mFD-components. This requires frequency-and-intensity-selective neurons. On the first level of the FD-neural-level the DLS will be scanned for the several frequency values. Their amplitudes are stored for the next neural level which can combine them in an additive (eFD-) or multiplicative (mFD-) way.

On the next level the procedure will be applied by special coupling of the neurons. The special neurons knot together the amplitude values, and the result will be stored in the neuron again.

The computed second-order feature-components will be combined additively by a neuron of the next neural level which represents the eFD- or mFD-spectral-component of the pLS-part.

The total mFD-component will then be the additive cummulation of all mFD-parts-components of the other network-parts which calculate the several classes of the DLS (see figure 2).

In both cases parts of the network may fail without causing a break down of the fundamental structure of the FD-spectrum or reducing the strength of the statement.

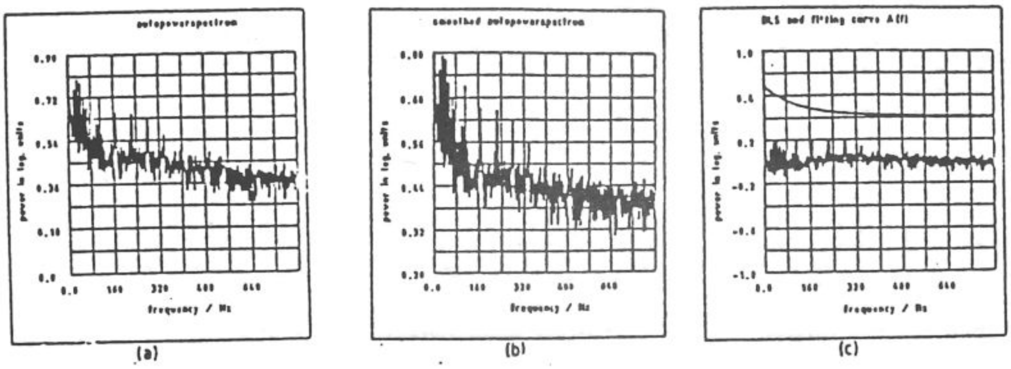

figure 1 : the calculationsteps of the DLS