

**MODELLING THE SOUND ENVELOPE AUDITORY PROCESSING
USING THE NON-NEGATIVE-IMPULSE-RESPONSE
MODULATION FILTERS CONCEPT
I. INITIAL SIMULATIONS**

D. KUTZNER

Adam Mickiewicz University, Institute of Acoustics
Department of Room Acoustics and Psychoacoustics
Umultowska 85, 61-614 Poznań, Poland
e-mail: konsbol@wp.pl

(received May 25, 2005; accepted December 9, 2005)

This article concerns with a new model of the sound envelope processing in the auditory system. The so-called non-negative-impulse-response (NNIR) modulation filters concept argues that if any form of the acoustic signal envelope filtering took place in the auditory pathway, this process should not be described in terms of a band-pass filtration. This modification of the traditional model of the auditory system temporal resolution, based on the modulation filterbank (MFB) activity, results from the cardinal property of the sound envelope and its neural representation, i.e. neural discharges period histogram, which are unavoidably unipolar signals of non-negative values. It has been assumed that if hypothetical modulation filters existed, they should be characterised by a non-negative-impulse-response and, consequently, the frequency characteristics of such filters might not reveal the band-pass properties. The results of the model investigations are compared with selected psychophysical and physiological data.

Key words: amplitude modulation, the Hilbert transform, modulation filterbank, impulse response, variance.

1. Introduction

The structure of most environmental sounds is characterised by continuous changes of their physical parameters. The temporal fluctuation of a signal envelope and its specific spectral structure are common features of the environmental signals and is a crucial factor enabling to distinguish between various natural and unnatural sounds. The auditory analysis of the sound envelope is thought to play an important role in speech recognition [1], speech perception [2] and music perception [3]. Under noisy conditions, the coherent amplitude modulation of the masker across different frequency bands enhances

the audibility of the masked sound [4]. Thus, experiments concerned with perception of amplitude and frequency (or both) modulated signals are important trends in the psychology of hearing.

There are, generally, two fundamental models of modulation processing in the auditory system. Both of them assume an initial transformation in bandpass filters reflecting the frequency selectivity of the basilar membrane (the auditory filters), a nonlinear unit (compression or half-wave rectification), the second filtering stage and the so-called decision unit. The main difference between the concepts is related to the nature of processing that takes place at the mentioned second filtering level. According to VIEMEISTER [5], the signal occurring at the output of the auditory filter, i.e. the sound envelope in a given auditory channel, is nonlinearly transformed and then low-pass filtered. The frequency characteristic of the low-pass filter is assumed to be reflected by the Temporal Modulation Transfer Function (TMTF), which is, generally, low-pass shaped. An alternative approach to the sound envelope processing at the higher stages of the auditory system, the so-called modulation filterbank (MFB) concept, presupposes the existence of a set of linear, overlapping, bandpass filters tuned to different modulation rates [6–12]. It is assumed that the MFB analyses the temporal structure of the sound by means of spectral decomposition of the signal envelope. Psychophysical data supporting the MFB concept come mainly from experiments concerned with masking in the modulation rate domain. In most of these experiments [8, 13, 14] tuning, similar to that in the audible frequency domain was observed. It ought to be stated that the hypothetical modulation filters are assumed to have a much lower quality factor (Q about 1 [10]) than the auditory filters (Q about 7–8 [15]). Although, the MFB concept, being actually an implementation of the auditory filters idea into the modulation rate domain, has attracted a considerable attention over the last years, it still remains somehow controversial.

The above mentioned concepts of the amplitude fluctuation processing postulate different processes applied to the sound envelope and, therefore, can be regarded as opposite models of temporal resolution of the auditory system. The main purpose of the present paper is to introduce and evaluate a new model of the signal envelope processing, which is aimed to interpret experimental results suggesting various mechanisms underlying the envelope analysis. Furthermore, the new model, called the non-negative-impulse-response (NNIR) modulation filters concept, provides a much more accurate description of physiology of the envelope processing than the previous models.

2. General approach to modelling auditory envelope processing

The envelope of a signal $x(t)$, i.e. the function describing its amplitude fluctuation, is defined as an absolute value of the so-called analytic signal $z(t)$:

$$|z(t)| = \sqrt{x^2(t) + \hat{x}^2(t)}, \quad (1)$$

where $\hat{x}(t)$ is the Hilbert transform of the original signal $x(t)$:

$$\hat{x}(t) = H[x(t)] = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(\tau) d\tau}{t - \tau}. \quad (2)$$

The signal $|z(t)|$ is usually called the Hilbert envelope of $x(t)$. Equation (2) describes a convolution of the original signal and the function $1/\pi t$. If $X(\omega)$ is the Fourier transform of $x(t)$, then Eq. (2) can be written as a simple multiplication in the frequency domain:

$$\hat{X}(\omega) = -j \operatorname{sgn}(\omega) X(\omega), \quad (3)$$

where $\hat{X}(\omega)$ is the Fourier transform of $\hat{x}(t)$. As results from Eq. (3), in the time domain, the result of the Hilbert transform is the original signal shifted in phase by $\pi/2$.

Results of many physiological experiments revealed that amplitude modulation (as well as frequency modulation [15]) of a stimulus is transformed into the fluctuation of action potentials generation rate observed in the auditory nerve and at the higher stages of the auditory system [16]. In other words, the amplitude modulation is (non-linearly) transformed to the modulation of temporal intervals between consecutive neural discharges.

An amount of neural spikes registered, for example, in the auditory nerve fibres in adjacent time intervals gives the so-called period histogram which reflects directly the envelope of a signal, Fig. 1. It should be emphasised that both the functions, i.e. the Hilbert envelope and period histogram, are always non-negative signals because neither an amount of action potential in a given analysis interval nor the envelope described by means of Eq. (1) can be negative. The traditional modelling modulation auditory processing is concerned, however, with determination of the signal envelope and application of the nonlinear transformation of the signal at the outputs of the auditory filters [17, 18]. It should be stressed that at this level frequency characteristics of the auditory filters are taken into account in the simulations only. Due to the fact that a phase response of the filter influences an amount of amplitude fluctuations at its output [19], it is highly probable that all components of the filter transfer function will be employed in computations in the future.

The obtained signal is filtered in a suitable filter [5] or in a set of filters [8, 10, 20, 21]⁽¹⁾. Instead of the envelope, a DC-coupled envelope is presupposed to be filtered. In the model proposed by DAU *et al.* [22], a linear combination of the DC-coupled envelope and a ‘*venelope*’ (i.e. the envelope of the DC coupled envelope) is assumed to be processed at the higher stages of the auditory pathway. In other words, the ‘*venelope*’ model assumes that in modulation processing the auditory system combines information coming from amplitude fluctuation of the acoustic signal and amplitude changes

⁽¹⁾ It should be stressed that “combined” models were also suggested. For example DAU [20] and DAU *et al.* [23, 24] proposed a model containing a low-pass filter that followed by a set of modulation filters. Another model, introduced by CHI *et al.* [25], contains five modulation filter banks of various filter bandwidth.

in the sound envelope. Nevertheless, very little physiological evidence supporting this approach to the modelling modulation perception has been found so far.

Consider now, Fig. 1. The upper part shows a 1000-Hz cosinusoidal carrier modulated by a 5-Hz cosinusoidal modulator with modulation depth of 80%. The middle and bottom parts of Fig. 1 depict a comparison between the above mentioned representations of the signal amplitude fluctuation: the envelope, the DC-coupled envelope and the “venelope”.

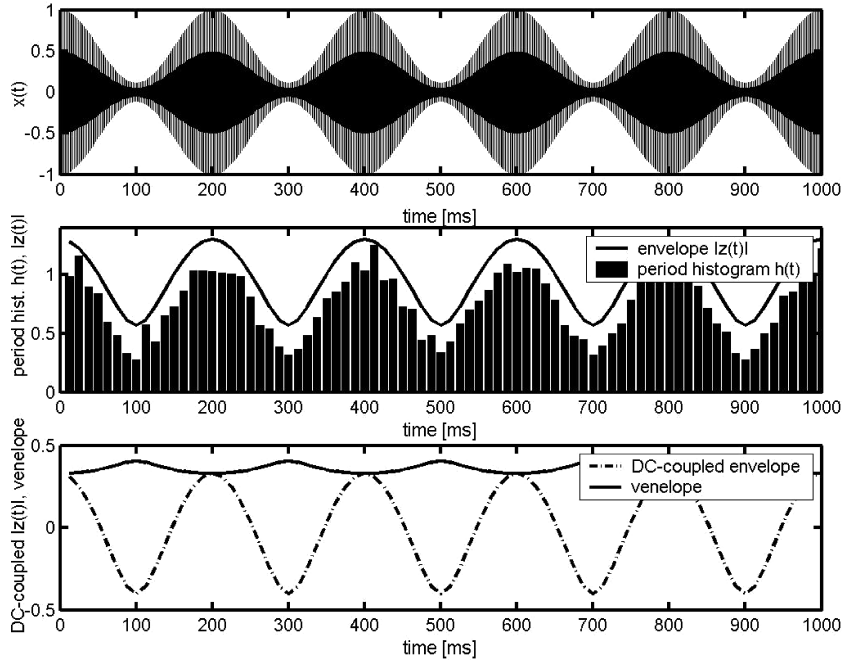


Fig. 1. The upper part presents an exemplary AM signal ($f_c = 1$ kHz, $f_m = 5$ Hz, $m = 80\%$). The middle and the bottom parts show a comparison of the signal envelope and neural discharges period histogram that could be observed in the auditory nerve (the middle part), the DC-coupled envelope and the venelope (the bottom part). The envelope and the period histogram are highly correlated (a 6 ms-time shift between the auditory filter input and output signal is not shown).

It is clear that the period histogram of action potentials and the Hilbert envelope are highly correlated, even if the period histogram is distorted by spontaneous neural discharges. Though the DC-coupled Hilbert envelope reflects the original signal amplitude fluctuation, it appears to be an inadequate representation of amplitude modulation. In the simplest case it predicts negative rates of action potential generation. The venelope seems to be the most abstractive sound envelope representation since it does not correlate with the original sound envelope. It is also inconsistent with physiological investigations. Therefore, the above analysis suggests that the Hilbert envelope and period histogram should be used in the modelling. Otherwise an agreement with physiological results will not be maintained. Henceforth, a nonlinearly transformed (compressed)

Hilbert envelope of the signal occurring at the output of a given auditory filter will be used in computations.

3. The non-negative-impulse-response modulation filters concept

As mentioned earlier, the histogram of the action potentials rate observed in the auditory nerve and at the higher stages of the auditory pathway reflects the fluctuation of the signal instantaneous power, i.e. its amplitude modulation.

The cochlea performs three basic transformations with respect to acoustic stimuli: band-pass filtering that originates from the multi-channel processing in the auditory system, compression enhancing sensitivity to low-level sounds and transformation of the basilar membrane oscillations into fluctuation of the action potential generation rate. Since the result of such a transformation correlates with the analytic signal modulus, the organ of Corti can be treated as an element extracting $|z(t)|$ (see the middle part of Fig. 1). The new model proposed in this paper is aimed to provide a high-precision description of neural transformation of the sound envelope. The starting point of the analysis is a comparison of the signals occurring at the outputs of elements associated with the above-mentioned models of the auditory system temporal resolution. The upper part of Fig. 2 presents the frequency characteristic of an exemplary traditional band-pass modulation filter and its impulse response. The impulse response is a bipolar function and non-negative values of neural discharges generation rate, as well as negative values of $|z(t)|$, are predicted.

The broadband low-pass filter (reflecting modulation processing properties of the auditory nerve, AN, [26]) can be described by a non-negative-impulse-response (the middle parts of Fig. 2). However, low-pass filter of a lower cut-off frequency (the bottom parts), i.e. of higher order, produces a bipolar impulse response. This remark, being concerned with unipolarity and bipolarity of the impulse response of the broadband and narrowband low-pass filters will be referred to in Sec. 3.2.

The above analysis suggests that the hitherto existing approaches to the auditory envelope processing, i.e. the traditional VIEMEISTER's model [5] and the MFB concept, are partially invalid since they do not fully correlate with physiological nor psychophysical data. Though the broadband low-pass modulation filtering [5] (that can be described by a non-negative-impulse-response) might be realised in the auditory system, it does not reflect the bandpass auditory selectivity in the modulation rate domain. On the other hand, the auditory temporal resolution model containing bandpass modulation filters unavoidably produces bipolar impulse responses and, consequently, predicts at some points of time negative rates of neural discharges. Hence, if any form of the sound envelope filtering took place in the real auditory pathway, the filters should produce non-negative impulse responses. Such hypothetical filters will be referred to as non-negative-impulse-response (or NNIR) modulation filters. As it will be shown, the above statement will cause some crucial changes of the assumptions of the traditional MFB model.

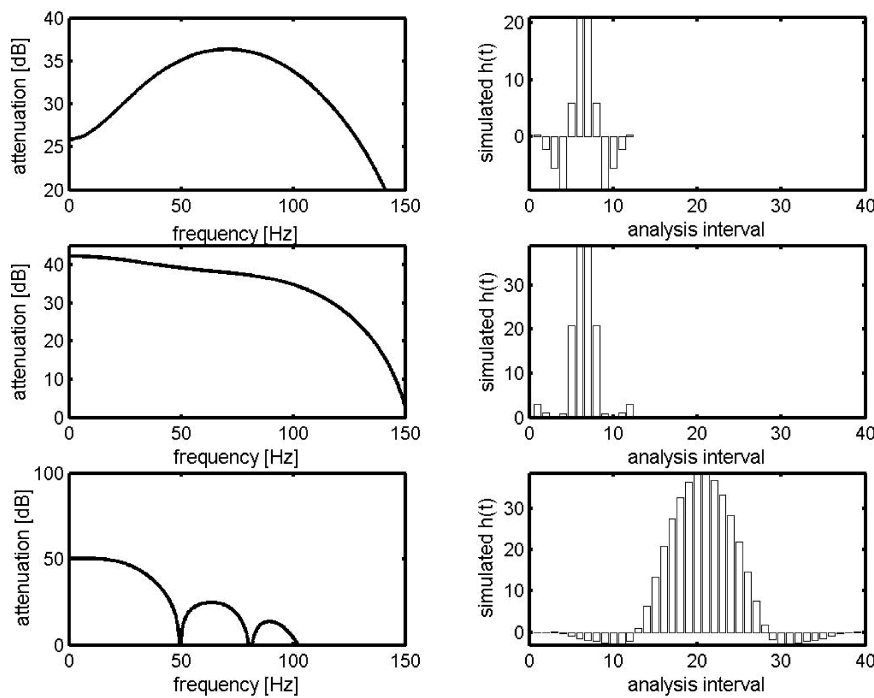


Fig. 2. The frequency characteristics of a traditional band-pass modulation filter (the left upper part), a broadband low-pass filter (the left middle part) and a narrow-band low-pass filter (the left bottom part). The right parts present impulse responses of the filters, respectively. Only the broadband low-pass filtering might be realised with respect to non-negative period histograms since such filter does not predict negative neural discharges generation rates (the right middle part).

3.1. The frequency characteristic of the NNIR modulation filters

A detailed analysis of the NNIR modulation filters model is depicted in Fig. 3. The left upper part of Fig. 3 presents the sound envelope and a simulation of the action potentials period histogram registered in the AN fibres. In this case, a white noise with a constant power spectrum density (left bottom part of Fig. 3) was used as a modulator. The input signal was previously processed by a band-pass filter reflecting the activity of the auditory filter tuned to the signal carrier frequency (1000 Hz) and passed through a non-linear transformation block simulating the non-linear properties of the cochlea. To make the picture more readable, the sound envelope presented in the left and middle parts of Fig. 3 (dashed lines) was shifted along the Y -axis by 1 unit (the right bottom part) and by 5 units (the middle bottom part) with respect to the simulated action potential histograms (bars).

The signal representation satisfying the assumptions of the NNIR modulation filters concept ought to be a function of non-negative values, otherwise the model would not correlate with the neural representation of the sound envelope. The middle upper part of Fig. 3 depicts the response of the hypothetical modulation filter (centre frequency

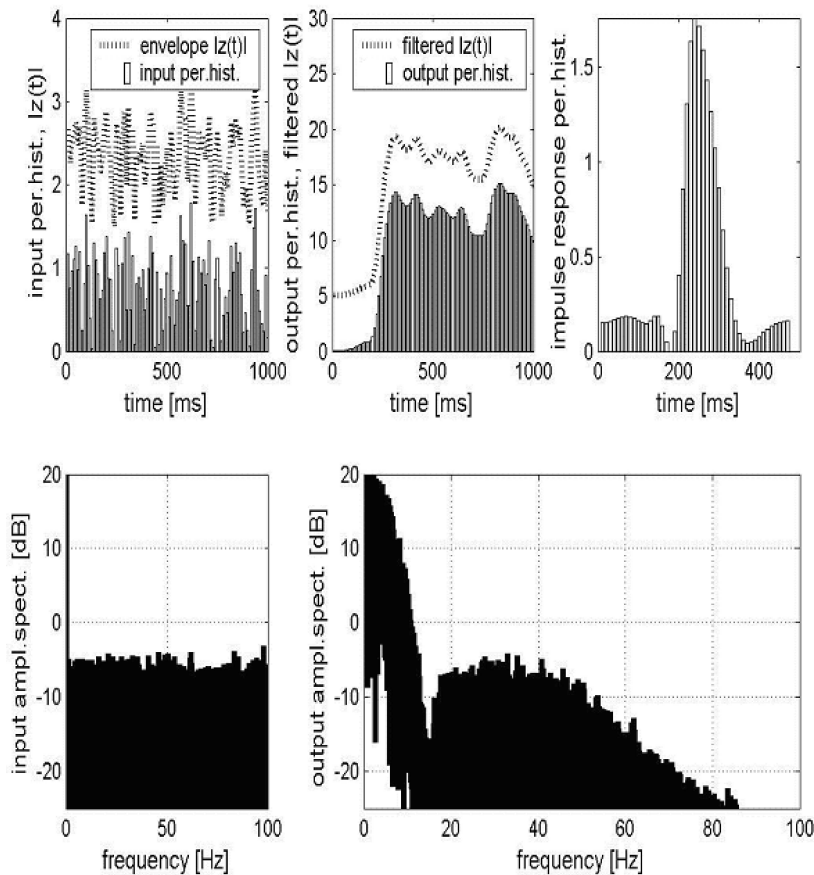


Fig. 3. Analysis of neural envelope processing that might occur in the auditory system: input period histogram correlated with $Iz(t)I$ (the bars the dotted line, respectively; the left upper part) and its long time average amplitude spectrum (the left bottom part), output period histogram correlated with the filtered $Iz(t)I$ (the bars the dotted line, respectively; the middle upper part) and its long time average amplitude spectrum (the right bottom part), the impulse response of the filter processing non-negative signals, i.e. output histogram deconvolved from the input histogram (the right upper part). The right bottom part demonstrates the shape of the NNIR modulation filter tuned to the frequency of 35 Hz (*per. hist-period histogram, ampl. spec. – amplitude spectrum*).

of 35 Hz) being stimulated by an input period histogram presented in the left upper part of Fig. 3, while the right upper part depicts the filter impulse response, i.e. the output period histogram deconvolved from the input period histogram. Taking into considerations the initial parameters of the filter (a band-pass frequency characteristics, centre frequency of 35 Hz), a special iterative procedure was employed with respect to the output signal $y(t)$ in order to obtain an impulse response of non-negative values. The DC-component of $y(t)$ had been optimised until the traditional band-pass modulation filter became the NNIR modulation filter. The right bottom part of Fig. 3 shows an average amplitude spectrum of the modulator after passing through the discussed NNIR

modulation filter. As the period histogram, reflecting a modulator of a constant power spectrum density was applied to the filter, the amplitude spectrum presented in the bottom right part of Fig. 3 demonstrates the frequency characteristic of the hypothetical NNIR modulation filter. Since the estimated frequency characteristic is no longer band-pass, the computation results lead inevitably to the following conclusion: if an envelope spectral structure is analysed in some way at the higher stages of the auditory system, this process could not be described in terms of the band-pass filtration only. The frequency characteristic of the NNIR modulation filter is characterised by two distinct local maxima. The first one corresponds to the frequency of 0 Hz and reflects a strong DC-component noticeable in the NNIR filter impulse response. The second band is related to the characteristic frequency of the NNIR filter and falls at the geometric centre of the band-pass part of the characteristic. Henceforth, the transfer function range with maximum falling at frequency of 0 Hz will be called the DC-band, while the frequency region corresponding to its characteristic frequency (CF) will be referred to the CF-band.

3.2. Variance-excitation pattern of NNIR modulation filters

The new model of the auditory temporal resolution requires a new approach to the description of the NNIR modulation filterbank activity. Figure 4 shows frequency characteristics of selected NNIR modulation filters and the RMS-based excitation pattern⁽²⁾ in response to the AM signal of $f_{\text{mod}} = 8$ Hz, $m = 100\%$. Frequency of the carrier was $f_c = 1000$ Hz.

It is obvious that the conventional RMS-based excitation pattern in the case of NNIR modulation filters is an inadequate way of the neural activity description because it represents the filters' (or modulation filters' [21]) response magnitude as a function of characteristic frequency of the filters [15]. It is related to an ambiguous excitation of the low-pass band of the NNIR modulation filters as all the low-pass bands fall at the same frequency of 0 Hz. Furthermore, if the auditory system determined a signal magnitude at the outputs of the NNIR modulation filters, the envelope spectral structure analysis could not be performed unequivocally since a sinusoidal modulator of frequency f_{mod} is unavoidably transmitted by the low-pass bands of the NNIR filters tuned to modulation rates that are higher than f_{mod} . The excitation pattern for a narrow-band modulation signal, excluding the DC-component, is not a band-pass function, thus this activity representation appears to be inadequate.

Presumably, in order to perform a band-pass spectral envelope analysis, which is suggested by a large majority of experimental data concerning the modulation masking effect [8, 11, 12, 21], the auditory system 'neglects' the modulator energy passing

⁽²⁾ Excitation pattern describes both peripheral auditory processing and hypothetical envelope processing (modulation excitation pattern). It is defined as power (expressed in terms of RMS and converted to dB scale) of the signals occurring at the outputs of the auditory (modulation) filters as a function of the characteristic frequency (CF) of the filters.

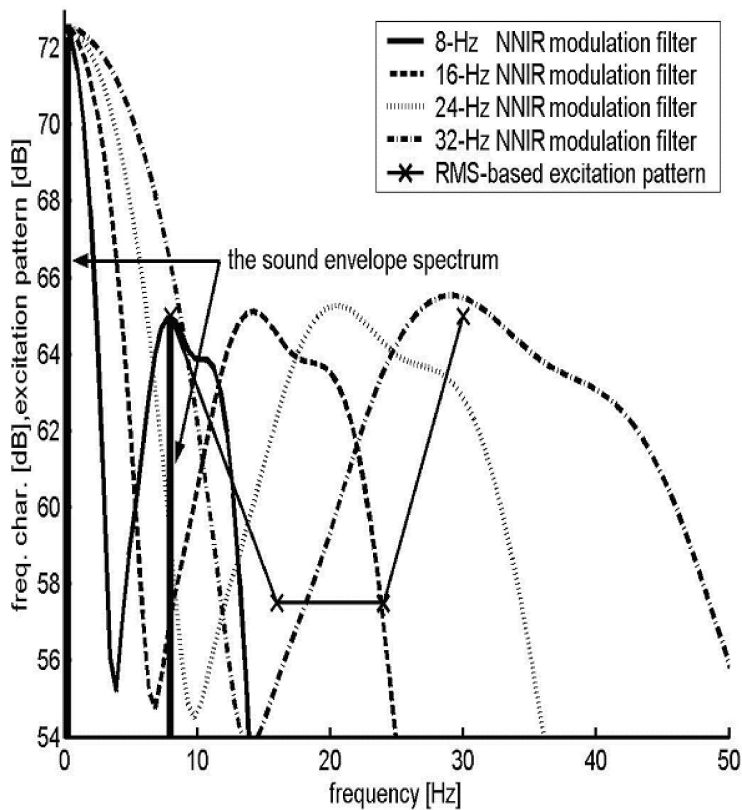


Fig. 4. The frequency characteristics of selected NNIR modulation filters tuned to modulation frequencies of 8, 16, 24 and 32 Hz. The vertical solid lines, falling at the frequencies of 0 and 8 Hz, indicate the sound envelope spectral structure; the crosses depict the RMS-based excitation pattern of the filters (*freq. char.:frequency characteristics*).

through the DC-band of the NNIR modulation filters. It seems that only a variability of the sound envelope (corresponding to modulation spectrum for $f_{\text{mod}} > 0$ Hz) plays a primary role in the amplitude envelope processing. Hence, an optimal-interval variance meter has been proposed as a unit following every NNIR modulation filter. Every NNIR modulation filter is assumed to be followed by a separate variance meter, which measures the variance of a signal appearing at the output of the filter. The time interval in which the variance σ^2 is measured is presupposed to be a function of the CF (characteristic frequency) of the analysed NNIR filter and increases as the CF of the filters increases. For a given time interval of analysis, the envelope spectral components of frequencies falling below some frequency, determined by the mentioned analysis interval, are characterised by periods of duration shorter than the interval. Thus, low-frequency fluctuations of a modulator are reduced in the variance-excitation pattern. Figure 5 presents the output of an exemplary optimal-variance meter for the analysis interval equal to 60 ms.

As it can be seen, although all the spectral components of the input signal are of the same variance, the result reflects some weighting in the modulation rate domain (especially for the DC-component and low-modulation frequencies).

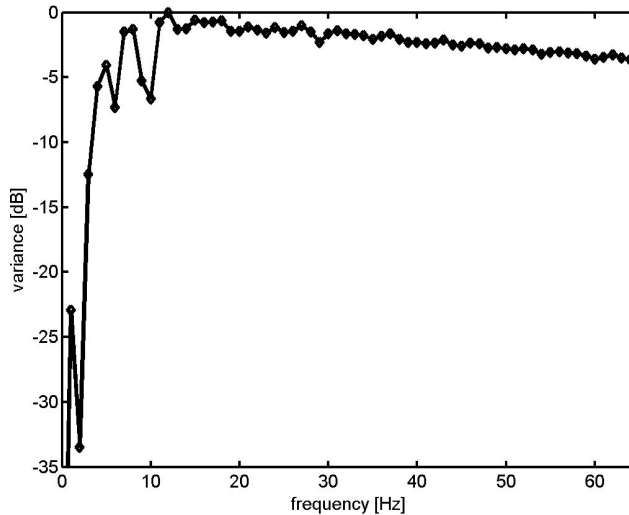


Fig. 5. The optimal-time-interval variance measurement for the analysis interval of 60 ms.

It ought to be stated that Fig. 5 does not depict the frequency characteristic of the optimal-variance meter since it is not a high-pass filter; it should not be treated as a transmitting device at all. The optimal-variance meters are elements describing a neural activity in respective modulation channels, rather. The range that low-frequency modulation energy is not taken into consideration in the calculations, i.e. is neglected, rises as the analysis interval increases. As seen from Fig. 5, the optimal-variance meter introduces a sort of weighting function in the modulation rate domain, so the final excitation pattern of the NNIR modulation filters is a band-pass function, (Fig. 6). Therefore, such an excitation pattern is called the variance-excitation pattern.

To conclude, the neural spikes period histogram, reflecting directly the sound envelope in a given auditory channel, as well as this period histogram determined after the envelope filtration in the hypothetical NNIR modulation filter are non-negative functions. The DC-component is, consequently, noticeable in the amplitude spectra of the signals (Fig. 4, solid lines). Nevertheless, due to the assumed activity of the optimal-interval variance meters, the band-pass (variance) excitation pattern is obtained as shown in Fig. 7. As mentioned in Sec. 3, the low-pass filters of higher orders, i.e. of narrower bandwidths, generate bipolar impulse responses. Accordingly, the band-pass selectivity for the lower modulation rates is not reproduced by a composition of a low-pass filter (the bottom parts of Fig. 2) and the optimal-interval variance meter. The NNIR modulation filters, revealing combined frequency characteristics, appear to be an adequate way of modelling the sound envelope auditory analysis.

4. The NNIR modulation filters concept in the light of experimental data

4.1. Psychophysical correlations

4.1.1. Masking in the modulation rate domain

The new auditory modulation processing model containing a bank of NNIR modulation filters followed by the optimal variance meters directly reflects the results of psychophysical experiments concerned with modulation masking. In a traditional modulation detection/discrimination interference paradigm, MDI, [12, 14, 27], the signal being detected (the probe signal), i.e. the amplitude modulation of one carrier, is masked by the so-called masker modulator, i.e. the amplitude modulation applied to the same carrier [14] or to the second carrier [27]. The probe thresholds are determined for various spectral separations between the masker and the probe, whereas the frequency of the latter is kept constant. It is assumed that in the detection task the subject analyses the output of a single modulation filter tuned to the frequency of the probe signal. The shape of the MDI curve, therefore, describes changes of S/N ratio at the output of the filter and may be regarded as an approximation of the frequency characteristic of the filter tuned to the frequency of the probe. Figure 6 presents a comparison of the NNIR modulation filters model predictions and typical MDI curves. The model predictions are expressed as $10 \log \sigma^2$, where σ^2 is the optimal-interval variation of a neural spikes period histogram of the neural signal appearing at the output of the NNIR modulation filter as a function a modulator rate. The entire AM signal was processed by the band-pass filter tuned to the carrier frequency (the auditory filter) and then, non-linearly transformed. It can be seen from Fig. 6 that the model predictions and the psychophysical data are coherent.

In the modulation masking paradigm, the probe detection is disturbed by the masking modulator, but the frequency of the masker remains constant [21, 28, 29]. Like in the audible frequency domain, the modulation masking patterns reveal band-pass characteristics and may be treated, therefore, as reflecting an excitation pattern of some post-cochlear filters. It should be stressed that although such masking patterns are narrowly tuned in the modulation rate domain, they reveal in a general case broadband tuning in the audible frequency domain [15]. It means that an increase in a spectral separation between the probe signal carrier and the modulation masker carrier does not affect the shape of the patterns. The band-pass modulation masking patterns were obtained for binaural presentation of the probe and the masker [29], which also confirms a post-cochlear nature of the sound envelope processing.

Figure 7 shows a comparison of various masking patterns in the modulation rate domain and the variance-excitation pattern produced by the hypothetical NNIR modulation filter bank. It should be noticed that though the NNIR modulation filters have a compound frequency characteristic (the low-pass band and the pass-band) due to the activity of the optimal-interval variance meters, the final variance-excitation pattern is approximately a band-pass function. Like in the comparison of MDI results, the neural response of the new model correlates very well with the psychophysical data.

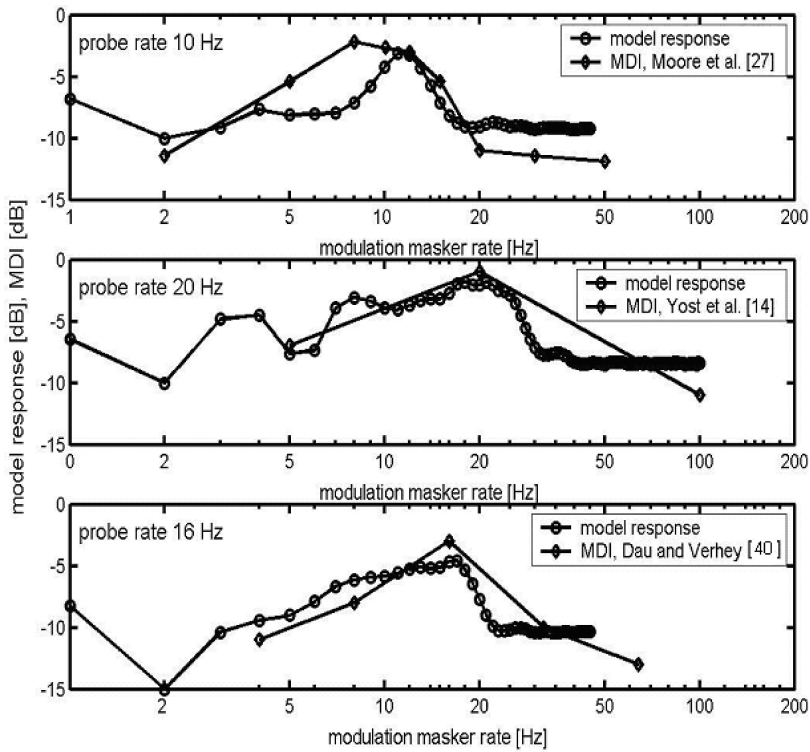


Fig. 6. Comparison of the model response and shape of MDI curves obtained by MOORE *et al.* [27], YOST *et al.* [14] and DAU and VERHEY [40].

The main difference between the traditional modulation filters and the NNIR modulation filters is related to the shape of the frequency characteristics. It is due to the DC-component of the envelope of non-negative representations that must be noticeable in the frequency characteristic of the NNIR modulation filters. It is assumed that to reproduce the band-pass selectivity in the modulation rate domain, the auditory system disregards the modulator frequency components passing through the low-pass band of a given NNIR modulation filter. It should be emphasised that in some situations this mechanism appears to fail partially. The activity of the low-pass band may have been well observed by BACON and GRANTHAM [28], who showed two local maxima in the modulation masking pattern (Fig. 7). The low-pass band might have been also noticeable in MDI curves obtained by MOORE *et al.* [27], although there was a significant dispersion in the data collected across the subjects. It should be stressed that the carriers used in the investigations mentioned were, in general, complex signals, so the multi-channel processing might be the factor that impairs the performance of the optimal-interval variance meters. Accordingly, the modulation masking investigations using complex carriers, that might disclose the activity of the low-pass bands of NNIR modulation filters, should be carried out to verify the new auditory processing model.

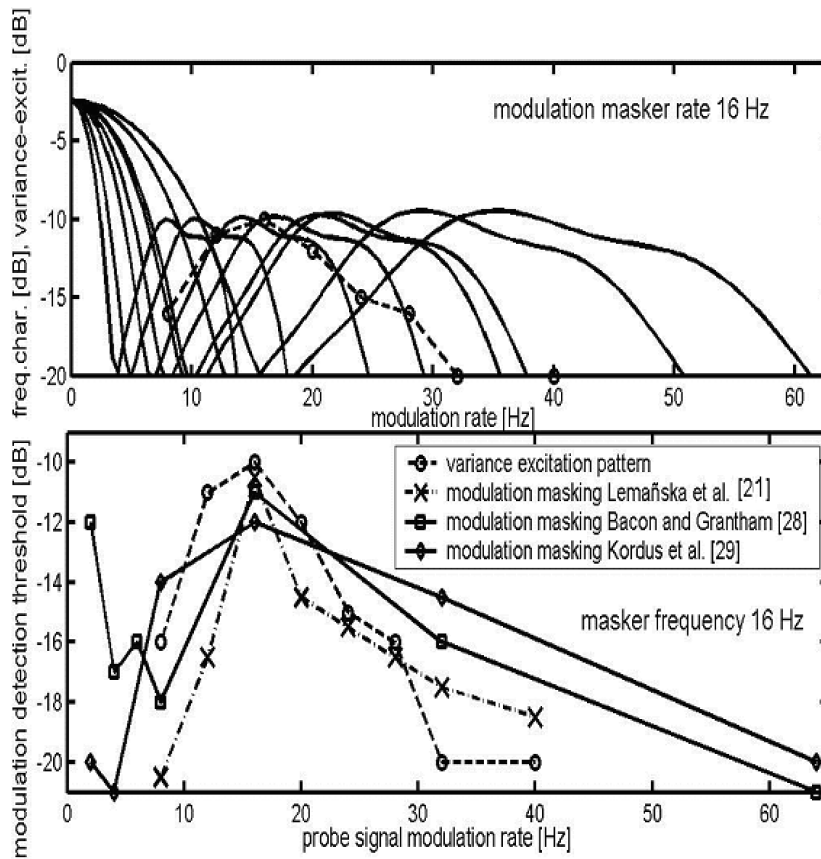


Fig. 7. The upper part: frequency characteristics of the NNIR modulation filters and the optimal-time-interval excitation pattern obtained for a 16-Hz modulation signal. The bottom part: the variance excitation pattern (replotted from the upper part) and modulation masking patterns obtained by LEMAŃSKA *et al.* [21], BACON and GRANTHAM [28] and KORDUS *et al.* [29] for a 16-Hz masking modulator.

4.1.2. The optimal-time-interval variance meter and other experimental data

In order to reproduce the band-pass frequency selectivity suggested by a large majority of experiments concerned with masking in the modulation domain phenomenon, the NNIR modulation filterbank is assumed to be followed by a set of the so-called optimal-variance detectors. Although the respective NNIR modulation filters reveal combined frequency characteristics (Fig. 7), an optimal-time interval measurement of the variance of a signal appearing at the output of a given NNIR modulation filter reduces the influence of the modulator low-frequency components. Consequently, the optimal-variance excitation pattern is a band-pass function. This assumption is in line with other psychophysical results. Measurements of modulation detection of complex modulator waveforms characterised by various crest factors suggest that the modulation detection should be thought in the category of a variance-excitation rather than as the traditional

RMS-excitation. The crest factor C_f , defined as the ratio of the peak value to the RMS value of a signal, strongly depends on the phase spectrum of the signal. In other words, by a manipulation of initial phases of complex modulator spectral components, modulation signals of various C_f can be obtained. According to the lack of a relationship between the signal power and its phase spectrum, the modification of C_f does not influence the RMS value of the modulating waveform, even though the signal variance may be altered. RYBICKA *et al.* [30–32] proved that, for modulators of the same RMS value, the listeners' performance (modulation detection) was better for modulators with larger C_f . This observation leads to the conclusion that the modulation detection and, presumably, masking in the modulation domain phenomenon reflects the activity of some filters being followed by variance detectors that analyse the magnitude of the modulator temporal fluctuation.

4.2. Physiological correlations

The idea suggesting that the sound envelope is analysed in a set of filters existing at the higher stages of the auditory system has been strongly supported by the results of physiological investigations [33–39]. The neurones, that are narrowly tuned in the modulation rate domain and revealed a broadband response to the audible frequency domain, were found in the cochlear nucleus (CN) of the rat [34, 35, 37], the gerbil [36], the inferior colliculus of the cat [39] and the guinea pig [38]. MØLLER [34, 35] and FRISINA *et al.* [36] demonstrated that the variability (depth) of the neural discharges rate modulation registered in the CN of the rat is by 30% larger than the modulation depth of the sound applied to animal's ear. This enhancement or gain is predicted by the NNIR modulation filters concept. This is related to the fact that the convolution variance of two positive-value signals (the input period histogram and the filter impulse response period histogram) are larger than the variance of the input period histogram. The filtering process is, in fact, an amplification of the modulator spectral components that pass the bands of a NNIR modulation filter. It is noteworthy that although the AN fibres have low-pass-shaped frequency characteristics, the measurements of the functions for modulation rates up to 6400 Hz reveal two-band frequency characteristics of the AN with the roll-off frequency (the local minimum) falling at approximately 800 Hz. Such AN modulation transfer functions were found for instance in the guinea pig [26]. The shape of the modulation transfer function and the frequency characteristic of the hypothetical NNIR modulation filter are, excluding the frequency range, nearly identical. Therefore, the respective AN fibres might have been the evolutionary origin of the mammalian auditory system frequency selectivity in the modulation rate domain. Consequently, due to the simplicity of the AN fibres, which are in fact uncomplicated axons, the envelope filtration stage could have been evolutionally transposed to the higher stage of the auditory system. The CN neurones of more complexity than the AN fibres appear to be a more adequate biological environment where high-order neural NNIR modulation filters might exist.

5. Conclusions

In this article a new model of the sound envelope auditory analysis has been proposed. The NNIR modulation filters concept presupposes that if any form of the amplitude envelope filtering takes place at higher stages of the auditory system, it should not be described as an analysis of the modulation waveform in a set of band-pass filters (the traditional MFB model). Due to the unipolarity of the input and output signals, the frequency characteristics of the NNIR modulation filters, in comparison to the traditional MFB, reveals two local maxima. The first maximum, associated with a filter DC-band, is related to the DC-component of its non-negative-impulse-response of the filter. The second one reflects the filter selectivity to a given part of the modulation spectrum for $f_{\text{mod}} > 0$. In order to reproduce the band-pass tuning in the modulation rate domain (suggested, for example, by a large majority of experiments concerned with modulation masking), the optimal-time-interval variance meter has been introduced. The predictions of the new envelope auditory processing model correlate with the physiological data as well. Further modulation masking experiments that may verify the assumptions of the NNIR modulation filters concept are required.

Acknowledgments

The author would like to thank professor Aleksander Søk for his helpful comments on the earlier version of this manuscript.

The State Committee for Scientific Research supported this work, project number 4T11E01425.

References

- [1] EISENBERG L. S., DIRKS D. D., DIRKS, BELL T. S., *Speech recognition in amplitude-modulated noise of listeners with normal and listeners with impaired hearing*, Journal of Speech and Hearing Research, **38**, 222–233 (1995).
- [2] LORENZI C., BERTHOMMIER F., APOUX F., BACRI N., *Effects of envelope expansion on speech recognition*, Hearing Research, **136**, 1–2, 131–138 (1999).
- [3] BURNS E. M., VIEMEISTER N. F., *Nonspectral pitch*, Journal of the Acoustical Society of America, **60**, 863–869 (1976).
- [4] HALL J.W., TYLER R. S., FERNANDES M. A., *Factors influencing the masking level difference in cochlear hearing-impaired and normal-hearing listeners*, Journal of Speech and Hearing Research, **27**, 145–154 (1984).
- [5] VIEMEISTER N. F., *Temporal modulation transfer functions based on modulation thresholds*, Journal of the Acoustical Society of America, **66**, 1364–1380 (1979).
- [6] LORENZI C., SOARES C., VONNER T., *Second-order temporal modulation transfer functions*, Journal of the Acoustical Society of America, **110**, 2, 1030–1038 (2001).

-
- [7] DERLETH R. P., DAU T., *On the role of envelope fluctuation processing in spectral masking*, Journal of the Acoustical Society of America, **108**, 1, 285–296 (2000).
- [8] SEK A. P., SKRODZKA E. B., *Tuning in the amplitude modulation rate domain*, Archives of Acoustics, **23**, 4, 491–498 (1998).
- [9] SEK A., SKRODZKA E., *Maskowanie w dziedzinie modulacji. Krzywe strojenia modulacji*, [in:] *Open Seminary on Acoustics*, Gdańsk, Poland 1997.
- [10] SEK A., MOORE B. C. J., *Testing the concept of a modulation filter bank: The audibility of component modulation and detection of phase change in three-component modulators*, Journal of the Acoustical Society of America, **113**, 5, 2803–2811 (2003).
- [11] HOUTGAST T., *Frequency selectivity in amplitude-modulation detection*, Journal of the Acoustical Society of America, **85**, 1676–1680 (1989).
- [12] EWERT S. D., DAU T., *Psychophysical tuning in AM processing*, [in:] *Psychophysics, Physiology and Models of Hearing*, DAU T., HOHMANN V., KOLLMEIER B. [Eds.], pp. 73–76, World Scientific, Singapore 1999.
- [13] BACON S. P., GRANTHAM W., CAMPEN L. E., *Modulation masking patterns*, Journal of the Acoustical Society of America, **83**, S35 (1988).
- [14] YOST W. A., SHEFT S., OPIE J., *Modulation interference in detection and discrimination of amplitude modulation*, Journal of the Acoustical Society of America, **86**, 2138–2147 (1989).
- [15] MOORE B. C. J., *An introduction to the psychology of hearing*, 5th Ed., Academic Press, London 2003.
- [16] SCHREINER C. E., LANGNER G., *Coding of temporal patterns in the central auditory system*, [in:] *Auditory Function: Neurobiological Bases of Hearing*, EDELMAN G., GALL W., COWAN W. [Eds.], pp. 337–361, Wiley, New York 1988.
- [17] LEMAŃSKA J., SEK A., *Maskowanie w dziedzinie modulacji amplitudowej*, [in:] *Wave methods and mechanics in biomedical engineering*, Polish Acoustical Society, Kraków–Zakopane 2000.
- [18] SEK A., MOORE B. C. J., *Mechanisms of modulation gap detection*, Journal of the Acoustical Society of America, **111**, 6, 2783–2792 (2002).
- [19] OXENHAM A., DAU T., *Towards a measure of auditory-filter phase response*, Journal of the Acoustical Society of America, **110**, 6, 3169–3178 (2001).
- [20] DAU T., *Modeling auditory processing of amplitude modulation*, University of Oldenburg, 1996.
- [21] LEMAŃSKA J., SEK A., RYBICKA W., *Masking in the amplitude modulation rate domain*, Archives of Acoustics, **28**, 3, 151–159 (2003).
- [22] DAU T., VERHEY J. L., KOHLRAUSCH A., *Intrinsic envelope fluctuations and modulation-detection thresholds for narrow-band noise carriers*, Journal of the Acoustical Society of America, **106**, 2752–2760 (1999).
- [23] DAU T., KOLLMEIER B., KOHLRAUSCH A., *Modeling auditory processing of amplitude modulation: I. Detection and masking with narrowband carriers*, Journal of the Acoustical Society of America, **102**, 2892–2905 (1997).
- [24] DAU T., KOLLMEIER B., KOHLRAUSCH A., *Modeling auditory processing of amplitude modulation: II. Spectral and temporal integration*, Journal of the Acoustical Society of America, **102**, 2906–2919 (1997).

-
- [25] CHI T., GAO T., GUYTON M. C., SAMMA S., *Spectro-temporal modulation transfer functions and speech intelligibility*, Journal of the Acoustical Society of America, **106**, 2719–2731 (1999).
- [26] PALMER A. R., *Encoding of rapid amplitude fluctuations by cochlear nerve in the guinea pig*, Archives of Otolaryngology, **236**, 197–202 (1982).
- [27] MOORE B. C. J., GLASBER B., GAUNT T., CHILD T., *Across-channel masking of changes in modulation depth for amplitude- and frequency-modulated signals*, Q. J. Exp. Psychol., **43A**, 3, 327–347 (1991).
- [28] BACON S. P. GRANTHAM D. W., *Modulation masking: effects of modulation frequency, depth and phase*, Journal of the Acoustical Society of America, **85**, 2575–2580 (1989).
- [29] KORDUS M., SEK A. P., KOCIŃSKI J., *Bianural masking of amplitude modulation*, Archives of Acoustics, **30**, 1, 3–17 (2005).
- [30] RYBICKA W., SEK A., *Detection of the amplitude modulation for modulating signals characterized by different crest factors*, Archives of Acoustics, **28**, 4, 203–222 (2003).
- [31] RYBICKA W., SEK A., SUZUKI Y., *Detection and modelling of AM thresholds for modulators with different crest factors*, [in:] *Structures-Waves-Human Health, Biomedical Engineering*, Kraków–Zakopane, Poland 2003.
- [32] SEK A., SUZUKI Y., RYBICKA W., *Amplitude modulation thresholds for modulators with different crest factor*, [in:] *The 2003 Spring Meeting of the Acoustical Society of Japan*, Waseda University, Tokyo, Japan 2003.
- [33] MOLLER A. R., *Dynamic properties of primary auditory fibers compared with cells in the cochlear nucleus*, Acta Physiologica Scandinavica, **98**, 157–167 (1976).
- [34] MOLLER A. R., *Coding of amplitude and frequency modulated sounds in the cochlear nucleus of the rat*, Acta Physiologica Scandinavica, **86**, 223–238 (1972).
- [35] MOLLER A. R., *Coding of time varying sounds in the cochlear nucleus*, Audiology, **17**, 446–468 (1977).
- [36] FRISINA R. D., YOUND E. D., COSTALUPES J. A., *Encoding of amplitude modulation in the gerbil cochlear nucleus. I. A hierarchy of enhancement*, Hearing Research, **44**, 99–122 (1990).
- [37] REES A., MOLLER A. R., *Responses of neurons in the inferior colliculus of the rat to AM and FM tones*, Hearing Research, **10**, 301–310 (1983).
- [38] REES A., PALMER A. R., *Neuronal responses to amplitude-modulated and pure-tone stimuli in the guinea pig inferior colliculus, and their modification by broadband noise*, Journal of the Acoustical Society of America, **85**, 5, 1978–1994 (1989).
- [39] LANGNER G., SCHREINER C. E., *Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms*, Journal of Neurophysiology, **60**, 1799–1822 (1988).
- [40] DAU T., VERHEY J. L., *Modelling across-frequency processing of amplitude modulation*, [in:] *Psychophysics, Physiology and Models of Hearing*, DAU T., HOHMANN V., KOLLMEIER B. [Eds.], pp. 229–234, World Scientific, Singapore 1999.