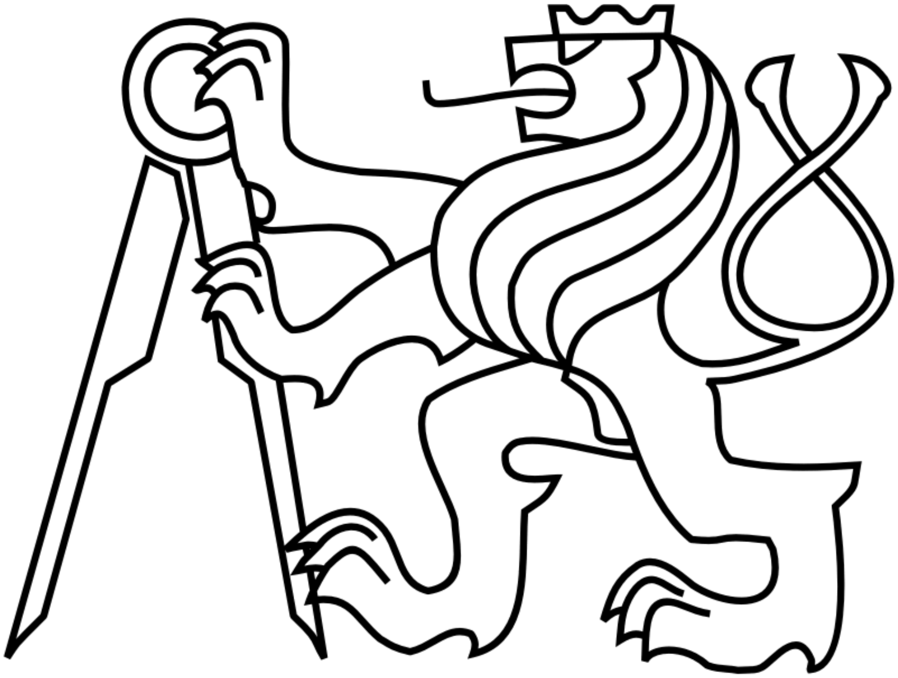
CZECH TECHNICAL UNIVERSITY IN PRAGUE



DOCTORAL THESIS STATEMENT

Czech Technical University in Prague
Faculty of Electrical Engineering
Department of Radioelectronics

Václav Vencovský

Assessment of Sound by means of an Auditory Model - Prediction of Roughness

Ph.D. Programme: Electrical Engineering and Information Technology
Branch of study: Acoustics

Doctoral thesis statement for obtaining the academic title of “Doctor”, abbreviated to “Ph.D.”

Prague, February 2015

The doctoral thesis was produced in part-time manner

Ph.D. study at the department of Radioelectronics of the Faculty of Electrical Engineering of the CTU in Prague

Candidate: Václav Vencovský
Department of Radioelectronics
Faculty of Electrical Engineering of the CTU in Prague
Technická 2, 166 27 Prague 6

Supervisor: Zdeněk Otčenášek
Musical Acoustics Research Center
Academy of Performing Arts in Prague
Malostranské nám. 13, 118 00 Prague 1

Opponents:

The doctoral thesis statement was distributed on: ....

The defence of the doctoral thesis will be held on .... at ...... a.m./p.m. before the Board for the Defence of the Doctoral Thesis in the branch of study of acoustics in the meeting room No. ...... of the Faculty of Electrical Engineering of the CTU in Prague.

Those interested may get acquainted with the doctoral thesis concerned at the Dean Office of the Faculty of Electrical Engineering of the CTU in Prague, at the Department for Science and Research, Technická 2, Prague 6.

....
Chairman of the Board for the Defence of the Doctoral Thesis
in the branch of study Acoustics
Faculty of Electrical Engineering of the CTU in Prague
Technická 2, 166 27 Prague 6.

Obsah

1 CURRENT SITUATION OF THE STUDIED PROBLEM

Roughness is a specific – harsh, buzzy and rattling – sound sensation. It accompanies stimuli with fast temporal fluctuations, for example, amplitude- or frequency-modulated tones [1]. This sensation can be measured by means of listening tests and was given a ratio-scale unit called **asper**. Since to conduct listening tests is time consuming and expensive, objective methods to measure roughness are being sought.

Researchers have developed a number of various roughness models in the last decades but none of them have been internationally or nationally standardized. This contrasts with models for other psychoacoustical parameters, for example, the Zwicker loudness model is described in the international standard, ISO 532B, and a model predicting sharpness in the German national standard, DIN 45692 [2].

So called **curve mapping** roughness models calculate roughness from the spectrum of analyzed stimuli. The models detect spectral components of the sound and predict roughness from their level and frequency position. This approach goes back to von Helmholtz [3] who estimated maximal roughness of two adjacent pure tones when the frequency difference between them was 33 Hz. These models were later adjusted and improved (e.g. [4, 5]) Disadvantage of these models is that they cannot process stimuli with continuous spectra, for example, noises. Moreover, the models which use only amplitude spectrum cannot predict the effect of phase of the spectral components and the shape of the signal envelope on roughness [6].

Some of the developed roughness models employ algorithms simulating the function of the peripheral ear. This stems from researcher's assumption that the roughness perception is caused by a limited frequency resolution of the peripheral ear (e.g. [4]). Among the models is, for example, the Daniel and Weber model [7] – designed by adjusting the Aures model [8], and the synchronization index (SI) model [6]. These models were implemented into softwares analyzing sound (e.g. [9]) and were further adjusted and used to predict roughness of vehicle noise (see [2, 10]). Kohlrausch *et al.* [11] showed that the Daniel and Weber, and the SI roughness model cannot predict (in agreement with results of listening tests given by [15]) the effect of phase of the spectral components and the shape of the signal envelope on roughness.

2 AIMS OF THE DOCTORAL THESIS

The specific aim of this thesis is to design a roughness model which employs algorithms simulating the function of the peripheral ear. The model should predict roughness of various types of acoustic stimuli in agreement with results of listening tests. The roughness model developed in this thesis is composed of two successive stages: a peripheral stage and a central stage. The peripheral stage employs an auditory model (a model of the peripheral ear). The central stage then processes the outputs of the peripherals stage and predicts roughness.

3 WORKING METHODS

3.1 Roughness model

Roughness model designed within the framework of this thesis is composed of two successive stages: a peripheral stage and a central stage.

3.1.1 Peripheral stage

The peripheral stage is an auditory model simulating the function of the outer- and middle-ear, cochlear mechanics, inner hair cells (IHCs) and auditory nerve (AN) synapse. I have adapted the algorithms from various studies and composed them into one auditory model. I have changed some of the parameters of the auditory model. Since the roughness perception is assumed to be caused by the limited frequency resolution of the peripheral ear [3], the model of cochlear mechanics is very important for the function of the roughness model. I have chosen a physical model designed by Mammano and Nobili [12, 13, 14] (the model is in the thesis called the Nobili *et al.* cochlear model). As is shown in the thesis, responses of the Nobili *et al.* cochlear model are qualitatively similar to responses measured in the cochlea of live mammals and the model can predict some of the known psychophysical phenoma (masking by pure tone and harmonic complex tone maskers). I have used psychophysical data (reproduced from the literature) showing frequency selectivity of the human hearing system to change the frequency selectivity of the Nobili *et al.* cochlear model. The data were measured using harmonic complex tone maskers.

3.1.2 Central stage

The central stage of the roughness model processes the output signal of the peripheral stage. I have designed the central stage within the framework of this thesis. It first extracts the envelope of the signal at the output of the peripheral stage, and then detects the rising slopes of the envelope – increasing parts of the envelope. The central stage then detects minimum and maximum of the rising slopes and calculates the duration of the rising slope and the modulation depth. These two parameters are then further processed and used – together with crosscorrelation coefficients calculated between the envelopes in adjacent channels and root mean square values of the envelopes – to calculate the predicted roughness. The central stage predicts roughness in 30-ms long time frames of the output signal of the peripheral stage. The overall predicted roughness of an analyzed stimulus is calculated as the median of the calculated roughness in 30-ms long successive time frames.

Pressnitzer and McAdams [15] processed stimuli with equal amplitude but not phase spectra (which were perceived with different roughness) by a model of cochlear mechanics. They showed that the processed stimuli differ in the shape of the filtered envelope (although the root mean square (RMS) values of the filtered envelopes are equal). They thus suggested that also the shape of the filtered envelope should be taken into account for roughness prediction. I have designed the central stage in order to take into account the shape of the envelope of the signal at the output of the peripheral stage (after filtering by the peripheral ear). I do not know another roughness model which can predict roughness of these stimuli in a good agreement with subjective data. I have set parameters of the roughness model in order to predict (in a quantitative agreement with results of listening tests) the dependence of roughness of sinusoidally amplitude modulated (SAM) tones on the modulation frequency. The results of the listening tests were reproduced from [1].

3.2 Model performance

The thesis compares the predicted and subjective (results of listening tests) roughness of sinusoidally amplitude-modulated (SAM) tones, two tone stimuli (dyads) composed of pure tones and harmonic complex tones, stimuli with temporal envelopes that are not sinusoidal – pseudo amplitude-modulated (pAM) tones and stimuli with asymmetrical temporal envelopes, sinusoidally frequency-modulated (SFM) tones, unmodulated broadband noise stimuli, amplitude-modulated (AM) harmonic complex tones, and synthetic and real vowels /a/. The subjective data of roughness of these stimuli were – except for AM complexes, synthetic and real vowels – reproduced from the literature. The subjective data of roughness of the AM complexes and vowels were measured by means of the listening tests conducted within the framework of this thesis.

3.3 Listening tests

I have conducted rating listening tests to measure the roughness of AM complexes, synthetic and real vowels. The method used to conduct the listening tests and to process the results is described below.

3.3.1 AM harmonic complexes

Stimuli: The stimuli were harmonic complexes composed of the first three harmonics ($N = 3$) as is given by

$$p(t) = [1 + m \cdot \cos(2\pi f_m)] \sum_{n=1}^{N} A(n) \cdot \cos(2\pi n f_0 t), \tag{1}$$

where f_m is the modulation frequency, m is the modulation index, $A(n)$ is the amplitude of the harmonics and f_0 is the fundamental frequency of the complexes. The fundamental frequency of the harmonics, f_0, was 300 Hz, the modulation frequency, f_m, was 30, 40, 50, 60, and 70 Hz, the modulation index, m, was set to 0, -3, -6, -9 and -12 dB given by the relation $20 \log_{10} m$, and the amplitude, $A(n)$, of the first, second and third spectral component was 0, -10 and -20 dB, respectively. The duration of the stimuli was 600 ms and they were ramped on and off with 30-ms raised-cosine ramps. Level of the stimuli was 75 dB SPL. The combinations of the modulation frequencies and the modulation depths led to 25 different stimuli.

Listeners: Five experienced listeners – four men, age ranging between 25 and 44 years, including the author – participated in the experiment. The listeners had normal hearing: pure-tone thresholds below 20 dB HL for frequencies between 250 Hz and 8 kHz.

Procedure: The listeners rated the roughness of the stimuli on a discrete scale from 1 to 7 in steps of 1, where 1 was for the lowest and 7 for the highest roughness. All 25 stimuli were presented in random order and the listeners rated each stimulus ten times which gave 250 ratings from each listener. The listeners could hear each stimulus as many times as they desired, assign a roughness rating and then proceed to the next stimulus. The test was conducted on a computer. The stimuli were presented diotically – the same signal to both ears – via Sennheisser HD-600 headphones. The procedure was inspired by the study [16] where the listeners rated the roughness of pathological voice samples on a 5 point discrete scale. Here, a 7 point scale was used instead. The reason for this was that the just noticeable difference of roughness corresponds to about 10% change of the modulation index, m, of a SAM tone [1]. Five chosen values of the modulation depths (0, -3, -6, -9 and -12 dB) of the AM complexes should thus cause perceptible changes of the roughness. Moreover, the roughness of the stimuli depends as well on the modulation frequency [1]. Hence, for the AM complex stimuli, a 5-point scale seemed to be too course.

Processing of results: Intra- and inter-rater reliability was estimated by Cronbach's alpha. The mean values across the listeners were taken as results.

3.3.2 Synthetic vowels /a/

Stimuli: Ten vowels /a/ varying in roughness were generated by means of the Klatt synthesizer [17]. The Klatt synthesizer first generates unit impulses which are then filtered by a glottal filter in order to create a glottal signal. The amplitude of the impulses and a time sequence between them was modulated which affected the amount of the perceived roughness. Ten different vowels were generated: nine of which had the frequency of the glottal pulses – fundamental frequency of the vowels – equal to 125 Hz, and one of the stimuli had the fundamental frequency equal to 63 Hz. The duration of the stimuli was 400 ms and it was ramped on and off with 30-ms raised-cosine ramps. The level of the stimuli was 75 dB SPL.

Listeners: Four normal-hearing experienced listeners participated in the experiment. Their pure-tone hearing thresholds were within a range of 15 dB hearing level (HL) for frequencies between 250 Hz and 8 kHz. The listeners were men aged between 25 and 36 years. The author was among the listeners.

Procedure and equipment: The listeners rated the roughness on a discrete 5-point scale from 1 to 5 in steps of 1, where 1 was for the lowest and 5 for the highest roughness. The same scale was used in the study [16] estimating the roughness of real pathological voice samples of a sustained vowel /a/. The procedure and equipment were the same as in the previous experiment with AM complexes. Randomly ordered 10 stimuli were rated 10 times, giving 100 stimuli per the listening test.

Processing of results: Same as in the previous experiment.

3.3.3 Real vowels /a/

Stimuli: 11 real pathological voice samples of a sustained vowel /a/ were used as stimuli. The vowels were extracted from the stimuli recorded from 11 different subjects during the scale singing. The subjects had a pathology affecting their larynx. The stimuli differed in the pitch and in the amount of roughness. The duration of the stimuli was 300 ms and they were ramped on and off by 30 ms raised-cosine ramps. The level of the stimuli was 75 dB SPL.

Listeners: Six experienced listeners – men aged between 25 and 36 years, including the author – participated in the experiment. The listeners had normal hearing; pure-tone thresholds below 20 dB HL for frequencies between 250 Hz and 8 kHz.

Procedure and equipment: Same as in the previous experiment with the synthetic vowels unless otherwise stated. Randomly ordered 11 stimuli were rated 10 times, giving 110 stimuli per the listening test.

Processing of results: Same as in the previous two experiments.

4 RESULTS

- The thesis describes a new roughness model designed within the framework of the thesis. The roughness model is composed of two successive stages: a peripheral and a central stage. The peripheral stage simulates the function of peripheral ear – algorithms simulating individual parts of the peripheral ear were adapted from the literature and composed into one model. The central stage was designed by the author. It predicts roughness from the output signal of the peripheral stage.

- The peripheral stage of the roughness model contains a physical model of the basilar membrane (BM) response and cochlear hydrodynamics (the Nobili *et al.* cochlear model). The thesis shows that the responses of the Nobili *et al.* cochlear model agree with similar responses measured in live mammalian cochlea – isointensity responses are level dependent, input/output (I/O) functions are compressively nonlinear and impulse responses are level near-invariant. The model was verified also using psychophysical masking thresholds for pure tone and harmonic complex maskers. The Nobili *et al.* cochlear model predicted the upward spread of masking thresholds observed during tone on tone masking. It also qualitatively predicted the phase effects in masking experiments with Schroeder phase maskers. These phenomena are not accounted for by many cochlear models [18, 19]. These results thus show that the Nobili *et al.* cochlear model could be applicable as a front end in the roughness model and also in other possible applications.

- The thesis shows the results of roughness listening tests conducted with sinusoidally amplitude-modulated (SAM) harmonic complexes, synthetic vowels /a/, and samples of real vowels /a/. The tests were conducted within the framework of the thesis.

- The described roughness model was used to predict the roughness of a large number of various stimuli: sinusoidally amplitude-modulated (SAM) tones; two tone stimuli (dyads) composed of pure tones and harmonic complex tones; stimuli with envelopes that are not sinusoidal – pseudo amplitude-modulated (pAM) tones and stimuli with asymmetrical temporal envelopes; sinusoidally frequency-modulated (SFM) tones; unmodulated bandpass noise stimuli; SAM harmonic complexes; and synthetic and real vowels /a/.

- The predicted roughness agreed with the results of the listening tests conducted within the framework of the thesis or reproduced from the literature. Since the central stage contains a new algorithm which takes into account the shape of the signal envelope at the output of the peripheral stage, the roughness model also predicted the effect of phase of the spectral components and the shape of the temporal envelope on roughness. These effects are not well covered by the roughness models known to the author. The worst agreement between the predicted and subjective roughness was for unmodulated bandpass noise and real vowels /a/. These stimuli contained noise.

5 CONCLUSION

The thesis described a model which predicts roughness of acoustic stimuli. The roughness model was in the thesis used to predict roughness of a large number of various types of acoustic stimuli.

The roughness model is composed of two successive stages: a peripheral stage and a central stage. The peripheral stage simulates the function of the peripheral ear: outer-/middle-ear, cochlear mechanics, inner hair cells and auditory nerve synapse. Algorithms simulating the function of the individual parts of the peripheral ear were adapted from the literature and composed into one model. The central stage – designed within the framework of the thesis – predicts roughness from the stimuli processed by the peripheral stage.

The peripheral stage of the roughness model contains a model of the basilar membrane (BM) response and cochlear hydrodynamics designed by Mammano and Nobili [12, 13, 14]. Specifically, the model variant with parameters described in the study [14] was used. The model is in this thesis called the Nobili *et al.* cochlear model. Since this model simulates the frequency resolution of the peripheral ear, it is important for prediction of roughness. The Nobili *et al.* cochlear model is a physical model which can simulate otoacoustic emissions [14] and is, in agreement with the experimental data observed in the mammalian cochlea, active, showing level dependent isointensity responses, compressively nonlinear input/output (I/O) functions of the responses and level near-invariant impulse responses. The model was verified using subjective data from masking experiments – tone on tone masking and harmonic complex maskers. The psychophysically measured frequency selectivity using harmonic complex tones was used to set frequency selectivity of the cochlear model. The model qualitatively accounted for the phenomena observed with pure tone and harmonic complex maskers. These physiological and psychophysical phenomena put a strong constraint on models of cochlear mechanics and thus limit many cochlear models [18, 19].

The roughness model was used in the thesis to predict roughness of a large number of various acoustic stimuli. Listening tests measuring roughness of these stimuli were conducted in the literature or – for amplitude modulated complex tones, synthetic and real vowels /a/ – by the author of this thesis. The predicted roughness agree with the subjective data for most of the used stimuli. The model covers the effect of phase of the spectral components and the shape of the temporal envelope on roughness which is its advantage in comparison to roughness models known to the author. The roughness model covers the roughness of theses stimuli because the central stage allows to take into account the shape of the signal envelope after it is processed by the peripheral stage. The largest discrepancies between the model predictions and the subjective data were for the unmodulated bandpass noise stimuli and for real vowels. These stimuli contain noise which probably worsened the roughness model performance. This disadvantage of the roughness model should be fixed in the future research.

REFERENCES

List of literature used in the thesis statement

[1] FASTL, H. & ZWICKER, E. (2007). Psychoacoustics: Facts and Models. Springer, Berlin, Heidelberg.

[2] WANG, Y., SHEN, G.Q., GUO, H. TANG, X.L. & HAMADE T. (2013) Roughness modelling based on human auditory perception for sound quality evaluation of vehicle interior noise. *J. Sound and Vibration*, **332**, 3893–3904.

[3] von HELMHOLTZ, H.L.F. (1877). On the Sensation of Tone as the Physiological Basis for the Theory of Music. 2nd ed., translated by A.J. Ellis (1885) from German, 4th ed. (Dover, New York, 1954).

[4] VASSILAKIS, P. (2001). Perceptual and physical properties of amplitude fluctuation and their musical significance. PhD thesis, University of California, Los Angeles.

[5] KAMEOKA, A. & KURIYAGAWA, M. (1996). Consonance Theory Part II: Consonance of Complex Tones and Its Calculation Method. *J. Acoust. Soc. Am.*, **45**(6), 1460–1469.

[6] LEMAN, M. (2000) Visualization and calculation of the roughness of acoustical musical signals using the synchronization index model (SIM). In: *Proceedings of the COST G-6 Conference on Digital Audio Effects (DAFX-00)*, Verona, Italy, DAFX 1 – DAFX 6.

[7] DANIEL, P. & WEBER, R. (1997). Psychoacoustical Roughness: Implementation of an Optimized Model. *Acustica*, **83**, 113–123.

[8] AURES W. (1984). Berechnungsverfahren für den Wohlklang beliebiger Schallsignale, ein Beitrag zur gehörbezogenen Schallanalyse. PhD Thesis, TU München, Germany.

[9] IPEM toolbox for perception-based music analysis `http://www.ipem.ugent.be/?q=node/27`

[10] WANG, Y. (2009). A Study on Sound Roughness Evaluation Based on an Auditory Synchronization Index Model. In: *Proceedings of the 2nd International Congress on Image and Signal Processing, IEEE*, **105**, 3612–3616.

[11] KOHLRAUSCH, A., HERMES, D. & DUISTERS, R. (2005) Modeling roughness perception for sounds with ramped and damped temporal envelopes. In: *Proceedings of Forum Acusticum*, Budapest, Hungary, 1719–1302.

[12] MAMMANO, F. & NOBILI, R. (1993). Biophysics of the cochlea: Linear approximation. *J. Acoust. Soc. Am.*, **93**(6), 3320–3332.

[13] NOBILI, R. & MAMMANO, F. (1996). Biophysics of the cochlea II: Stationary nonlinear phenomenology. *J. Acoust. Soc. Am.*, **99**(4), 2244–2255.

[14] NOBILI, R., VETEŠNÍK, A., TURICCHIA, L., & MAMMANO, F. (2003). Otoacoustic emissions from residual oscillations of the cochlear basilar membrane in a human ear model. *J. Assoc. Res. Otolaryngol.*, **4**, 478–494.

[15] PRESSNITZER, D. & McADAMS, S. (1999). Two phase effects in roughness perception. *J. Acoust. Soc. Am.*, **105**, 2773–2782.

[16] PATEL, S.A., SHRIVASTAV, R. & EDDINS D.A. (2012) Identifying a Comparison for Matching Rough Voice Quality. *J. Speech Lang. Hear. Res.*, **55**, 2012, 1407–1422.

[17] KLATT, D.H. (1980). Software for a cascade/parallel formant synthesizer. *J. Acoust. Soc. Am.*, **67**, 971–995.

[18] SHERA, C.A. (2001). Intensity-invariance of fine time structure in basilar-membrane click responses: Implications for cochlear mechanics. *J. Acoust. Soc. Am.*, **110**, 332–348.

[19] OXENHAM, A.J. & DAU, T. (2001) Reconciling frequency selectivity and phase effects in masking. *J. Acoust. Soc. Am.*, **110**, 1525–1538.

List of candidate's works relating to the doctoral thesis

Journal articles

Peer reviewed journals

[20] VENCOVSKÝ, V. (2014) Roughness prediction for complex acoustic stimuli. *Akustické listy*, **20**(3-4), 19–26.

Submitted journal articles

Journals with impact factor

[21] VENCOVSKÝ, V. Roughness prediction based on a model of the cochlear hydrodynamics. To be submitted in March 2015.

Conference papers

[22] VENCOVSKÝ, V. (2008). Influence of ear model parameter setting on simulated tone-on-tone masking patterns. In: *Poster 2008*, Prague, CTU, 1–8.

[23] VENCOVSKÝ, V. (2009). A Physiological Auditory Model. In: *126th AES Convention Papers*, New York: Audio Engineering Society, 1–6.

[24] VENCOVSKÝ, V. (2014) Modeling roughness perception for complex stimuli using a model of cochlear hydrodynamics. In: *Proceedings of Forum Acusticum*, Krakow, Poland, 1–6.

[25] VENCOVSKÝ, V. (2014) Modeling roughness perception for complex stimuli using a model of cochlear hydrodynamics. In: *Proceedings of International Symposium on Musical Acoustics (ISMA2014)*, Le Mans, France, 483–488.

Submitted conference papers

[26] VENCOVSKÝ, V. (2015) Prediction of masking thresholds for Schroeder phase maskers: masker level effect. In: *Proceedings of DAGA*, Nürnberg, Germany, submitted in February 2015.

SUMMARY

The term roughness describes a specific – harsh, buzzy and rattling – sound sensation which may occur when listening to stimuli with fast temporal fluctuations, for example, amplitude- or frequency-modulated tones. Roughness is, as well as loudness or pitch, an important psychoacoustic parameter. The thesis describes a computational model predicting roughness of acoustic stimuli, composed of two successive stages: a peripheral and a central stage. The peripheral stage is composed of an auditory model which transforms the input acoustic stimulus into the simulated neural signal. The auditory model contains a set of algorithms simulating the function of the outer- and middle-ear, cochlear mechanics, the inner hair cells and auditory nerve synapse. The algorithms were adapted from the literature and composed into one model. The central stage – designed within the framework of this thesis – predicts roughness from the envelope of the simulated neural signal.

The peripheral stage of the roughness model employs a physical model of the basilar membrane (BM) response and cochlear hydrodynamics (the Nobili *et al.* cochlear model). Since this model simulates the limited frequency resolution of the peripheral ear, it is important for prediction of roughness. The author of this thesis adjusted the frequency selectivity of the model according to psychophysical masking data for harmonic complexes. The isointensity responses, input/output functions and impulse responses of the cochlear model are compared with data (reproduced from the literature) measured in the cochleae of live mammals. The model was as well verified using psychophysical masking data (reproduced from the literature) for pure tones or harmonic complexes. The results show that the Nobili textitet al. cochlear model can account for physiological and psychophysical phenomena which limits many cochlear models.

The roughness model is in this thesis used to predict the roughness of a large number of acoustic stimuli. The predicted roughness was compared with results of listening tests. The subjective data of roughness of the stimuli were reproduced from the literature or obtained by means of the listening tests conducted within the framework of the thesis. The roughness was reproduced from the literature for: sinusoidally amplitude-modulated tones, two tone stimuli composed of pure tones and harmonic complex tones, pseudo amplitude-modulated tones, stimuli with assymetrical temporal envelopes, sinusoidally frequency-modulated tones, unmodulated bandpass noise stimuli. The listening tests were conducted within the framework of this thesis for: sinusoidally amplitude-modulated harmonic complex tones, synthetic and real vowels /a/. The predicted roughness agrees with the subjective data for most of the used stimuli. The largest discrepancies between the model predictions and the subjective data are for unmodulated bandpass noise stimuli and for real vowels. Both stimuli contain noise which complicates the roughness model performance. The roughness model covers both the effect of phase and the shape of the temporal envelope on roughness. This is its advantage in comparison to the roughness models known to the author.

RESUMÉ

Pojem drsnost označuje specifický – hrubý, rachotivý – vjem který může nastat při poslechu zvuků s rychlými časovými fluktucemi, například amplitudově modulovaných tónů. Drsnost je, obdobně jako výška nebo hlasitost, důležitým psychoakustickým parametrem. Tato disertační práce popisuje model který predikuje drsnost zvukových signálů. Model je složen ze dvou na sebe navazujících částí: periferální části a centrální části. Periferální část představuje model periferálního ucha, který transformuje vstupní akustický signál na simulovaný neurální signál. Model periferálního ucha je složen z algoritmů modelujících činnost vnějšího, středního a vnitřního ucha. Tyto algoritmy byly převzaty z literatury. Centrální část – navržená v rámci disertační práce – detekuje obálku simulovaného neurálního signálu a z ní predikuje drsnost. Algoritmy v centrální části umožňují vzít v potaz tvar obálky signálu na výstupu periferální části a tak predikovat drsnost signálů s nesinusovými obálkami.

Periferální část modelu drsnosti obsahuje model kochley. Model je v této práci nazván Nobili *et al.* kochleární model (po jeho autorech). Protože tato část simuluje frekvenční selektivitu periferálního ucha, je velmi důležitá pro činnost modelu drsnosti. V práci jsou uvedeny odezvy kochleárního modelu při stimulaci tóny o stejné intenzitě, vstupně výstupní funkce těchto odezev a také impulsové odezvy. Tyto charakteristiky jsou srovnány s obdobnými odezvami naměřenými v kochlee živých savců. Kochleární model je také ověřen pomocí psychofyzikálních dat naměřených na lidech – maskovacích prahů pro čisté a harmonické komplexní tóny. Psychofyzikálně zjištěná frekvenční selektivita je použita pro nastavení frekvenční selektivity kochleárního modelu. Kochleární model dovede simulovat řadu fyziologických a psychofyzikálních fenoménů, které limitují mnoho jiných kochleárních modelů.

Navržený model drsnosti je v této práci použit pro predikci drsnosti velkého množství různých typů signálů. Predikovaná drsnost je srovnána s výsledky poslechových testů, které byly převzaty z literatury nebo získány poslechovými testy provedenými v rámci disertační práce. Z literatury byla převzata data s výsledky poslechových testů pro amplitudově modulované tóny, dvoutónové stimuly složené z dvojice čistých tónů nebo dvojice harmonicých tónů, pseudo amplitudově modulované tóny, stimuly s asymetrickou obálkou a nemodulované šumy o různé šířce pásma. V rámci disertační práce byly provedeny poslechové testy pro amplitudově modulované harmonické komplexní tóny, syntetické a reálné samohlásky /a/. Modelem predikovaná data více či méně souhlasila se subjektivními daty pro použité stimuly. Nejhorší shoda byla v případě nemodulovaných šumů a také u reálných samohlásek. Oba tyto typy stimulů obsahují šum. To zřejmě znamená, že model špatně predikuje drnost pro signály s šumy.