An overview of technically oriented psychoacoustics in Finland

Matti Karjalainen
Helsinki University of Technology
Laboratory of Acoustics and Audio Signal Processing
Otakaari 5 A, FIN-02150 Espoo, Finland

This paper presents a short summary of educational and research activities in Finland related to psychoacoustics. Particularly, activities at Helsinki University of Technology (HUT), Laboratory of Acoustics and Audio Signal Processing, are described in more detail. This includes courses given and how they are organized. Some computer based psychoacoustics experimentation principles used in the course 'Communication Acoustics' are characterized. Examples of successful research projects, such as psychoacoustically oriented auditory modeling for sound quality measurements, computational modeling of spatial hearing, and psychoacoustic principles in signal processing, are referenced. A list of other psychoacoustics or more general hearing research in Finland is compiled.

1 Introduction

Until recently, the level of academic psychoacoustics in Finland has been relatively modest and its industrial application practically non-existing. The only unit that has had regular educational and research activities since 1970s is the Laboratory of Acoustics and Audio Signal Processing at the Helsinki University of Technology (HUT). Within few recent years psychoacoustically oriented academic research has been growing also elsewhere, see section 4.

Until very recently, psychoacoustics has found practically no applications in the Finnish industry. Now the interest is rapidly growing since, e.g., some leading companies in mobile communications are operating in Finland. At HUT Acoustics laboratory we believe that basic psychoacoustic knowledge is necessary in acoustic and audio R&D and more advanced application of auditory and psychoacoustic modeling is becoming one of the key issues in many research projects.

This paper presents a short summary of psychoacoustics and related hearing research in Finland. HUT activities, including courses given and how they are organized, research topics and most successful results, are metioned in somewhat more detail. E.g., some computer-based psychoacoustics experiments used in the course 'Communication Acoustics' are shortly characterized. Examples of successful research projects, such as psychoacoustically based auditory modeling for sound quality measurements, computational modeling of spatial hearing, and psychoacoustical principles in signal processing, are referenced.

2 Educational Activities

Since 1970s there has been regular educational activities at the Acoustics laboratory of the Helsinki University of Technology. The main form has been a course entitled 'Communication acoustics' (Kommunikaatioakustiikka). The contents of this course has varied but the main parts have been:

- Short introduction to the properties of sound sources, especially speech production and musical instruments.
- General description of the human auditory system.
- Principal concepts of psychoacoustics, in major parts following the formulation by Zwicker (Zwicker and Fastl, 1990), including the theory of critical bands and loudness formation, also pitch and timbre.
- Other concepts of psychoacoustics such as sharpness, roughness, fluctuation strength, consonance and dissonance, auditory streaming, perception of modulation and phase, etc.
- Spatial and directional hearing, the precedence effect, HRTFs, auralization.
- Issues of sound quality evaluation by means of psychoacoustics.
- Computational auditory models.
- Fundamentals of audiology and hearing aids.

Some aspects of psychoacoustics are discussed also in other courses, e.g., 'Speech processing' and 'Advanced course in speech processing'. There are also seminars and workshops of varying contents that in some cases are related psychoacoustics and its applications.

2.1 Computer Based Exercises

Psychoacoustics is a challenging topic for education since it can only be aquired well through practicing with listening experiments and, on the other hand, it is only partly formulated mathematically. The listening experiments in traditional forms are very tedious and mathematical excercises have often an artficial flavor. Thus, new and improved ways to carry out such educational activities are higly desirable.

One very promising approach to learning psychoacoustics is to use computers to carry out listening experiments. Modern computers often include high-quality sound output (16 bit sound quality) and a good graphical user interface. This reveals very powerful environments for exploration with psychoacoustic and other hearing related phenomena. Easy-to-do listening experiments are possible in order to illustrate the basic auditory phenomena., such as:

- pitch perception
- loudness formation
- masking and critical bands
- consonance and dissonance
- spatial hearing, etc.

We have developed a hypermedia environment for some psychoacoustic experiments (Karjalainen and Rahkila, 1995). It is implemented using the QuickSig signal processing environment that yields a relatively efficient support for DSP operations, a high-level programming environment (Common Lisp + CLOS language) and scriptable user interface specifications. As an example of the system, in the following we will illustrate a listening experiment that yields the frequency domain masking patterns of the student. The experiment consists of the following steps:

- show basic concepts
- give instructions
- do listening experiment
- display the results

The following series of figures (screen snapshots) illustrate the main phases of the masking experiment. Figure 1 shows the page that presents some background information and concepts.

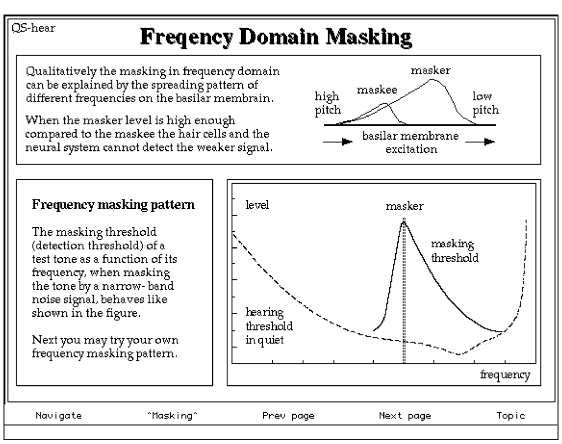

Fig. 1. Preliminaries page for frequency domain masking effect experiment.

The next page, Fig. 2, shows the experiment setup where the user moves mouse over a trackpad area in order to specify the level of the test tone by the vertical position of the cursor. Clicking the mouse means that the just noticeable level of the test tone in the presence of the masking narrow-band noise is found. The experimet automatically steps through a span of test signal frequencies.

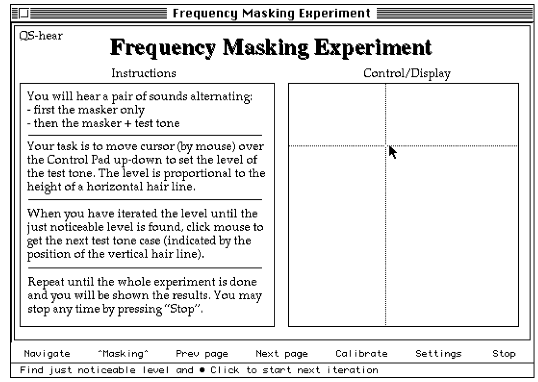

Fig. 2. Test panel for the frequency domain masking experiment.

Finally, a new window appears, Fig. 3, graphing the result of the experiment so that the masking pattern of the subject can be viewed.

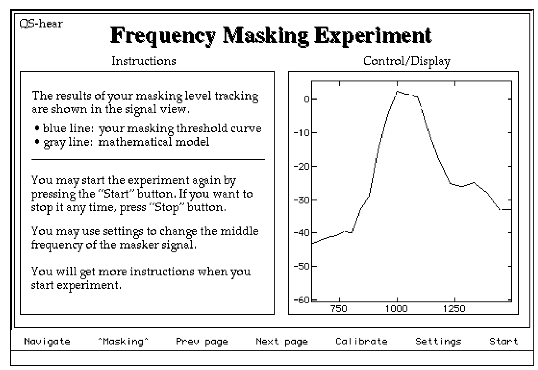

Fig. 3. Result and explanation page for the frequency domain masking experiment.

3 Research at HUT

In this section we refer to some psychoacoustically related projects that represent well the reasearch activities at the HUT Laboratory of Acoustics and Audio Signal Processing. The general approach is from the point of view of computational modeling.

3.1 Auditory Models and Applications

Since 1981 there has been efforts to develop psychoacoustically based and application-oriented computational auditory models (Karjalainen, 1982, 1985, 1987). The models follow generally known guidelines by implementing such features as equal loudness contour sensitivity, critical band frequency resolution, specific loudness scale, masking in the frequency and time domain, etc.

One of the best known results from these studies has been the application of computational auditory models to objective estimation of sound quality in speech and audio systems (Karjalainen, 1982, 1985). It is refered to as the first formulation where auditory model outputs of the original and the distorted signal are compared in the auditory spectrum domain. This yields objective measures of sound quality that correlate well with subjective listening tests. This general approach has since then been succesfully developed further, e.g., by Beerends and Stemerdink (1992).

Other applications for our auditory models have been within speech analysis and speech recognition.

3.2 Modeling of Spatial and Directional Hearing

A more recent form of auditory modeling at HUT has been the binaural modeling of directional and spatial hearing (Backman and Karjalainen, 1993), (Karjalainen, 1996). The approach was based on the use of a dummy head and a binaural auditory model as well as a neural network for the estimation of the sound source direction. The system is trained by sound examples in an anechoic chamber or in more reverberant environments and tested with similar sounds. It turned out that in simple environments the system outperforms human listeners in directional accuracy while in more complex environments it doesn't show ability to generalize its skills.

Such an approach to sound direction estimation and related possibilities to do sound source separation are considered to have many potential applications, such as noise-robust speech recognition and evaluation of audio system sound quality.

3.3 Auralization and 3-D Sound

A recent research activity at HUT is 'virtual acoustics' including the modeling of the full communication chain from sound sources—especially model-based sound synthesis of musical instruments—to room simulation and auralization of the results for a listener. Here the final part, the directional and spatial hearing, is directly related to psychoacoustics since, in addition to HRTF measurements and filter implementations (Huopaniemi and Karjalainen, 1996), listening experiments and evaluation of the quality of different parts in the system must be done. As a common effort with the HUT multimedia laboratory an integrated virtual environment DIVA (Takala et al., 1996) is developed which includes animation and auralization of musical instruments and performers in a virtual concert hall.

3.4 Auditory DSP Techniques

One of the main current interests at HUT Acoustics laboratory is to develop digital signal processing methods that can be used to implement some major auditory features in a wide range of applications without sacrificing computational efficiency or even by performing better than the traditional non-auditory methods do.

Probably the most fundamental difference between 'normal' DSP, such as FFT-based spectrum analysis, and the auditory spectrum analysis in the human auditory system is a different frequency scale. Digital signal processing is inherently fixed to uniform frequency resolution corresponding to the linear Hz scale. This is due to the use of the unit delay as a fundamental building block in DSP. Techniques such as wavelets are used to escape this necessity. We have been studying other alternatives using 'warped DSP techniques' where the unit delay is replaced by a frequency-dependent delay, such as the first order allpass filter. This leads to a good approximation of the Bark scale, as has been shown, e.g., by Strube (1980), as well as by Smith and Abel (1986).

Our effors were started from so called fam functions and famlets (Laine and Altosaar, 1990) that are a very general set of orthogonal functions. They were found also to be an efficient tool in the study of the temporal resolution of the human auditory system. Based on these functions and related warped digital filters (Karjalainen et al., 1996), we have developed several applications that utilize psychoacoustically motivated frequency scales and other auditory modeling features, such as auditory spectrograms, auditory models (Karjalainen, 1996), audio coding techniques based on warped linear prediction (Laine et al., 1994), (Härmä et al., 1996), HRTF filters for auralization (Huopaniemi and Karjalainen, 1996), etc. This approach is not limited only to the Bark scale warping: currently we are exploring to apply the same technique to efficient ERB scale processing.

4 Other Research in Finland

The following list is a short collection of units or groups that carry out research related to psychoacoustics or hearing research in general. The list is not comprehensive; other institutions could also be found. There is only one industrial research center (NRC) mentioned in this list.

- *Helsinki University Central Hospital, Departement of Audiology.* Audiological research. (*Dr. Tapani Jauhiainen*).
- *Helsinki University of Technology, Low Temperature Laboratory.* Neuromagnetic brain research, auditory perception vs. EEG and EMG

correlates. (*Prof. Riitta Hari*). http://boojum.hut.fi/

- *Nokia Research Centre (NRC), Audio group,* Tampere. Evaluation of audio system performance.

- *University of Helsinki, Cognitive Brain Research Unit.* Auditory sensory memory (memory/perception), Mismatch negativity (MMN). Methods: EEG (ERP), MEG, FMRI, PET, etc. (*Prof. Risto Näätanen*).

- *University of Kuopio, Dept. of Environmental Sciences.* Psychophysiological responses and annoyance. Annoyance and environmental impact analysis (EIA). (*Dr. Erkki Björk*). http://www.uku.fi/ bjork/

- *University of Tampere, Deparment of Psychology.* Audiovisual integration in humans: The mechanisms of human audiovisual integration are studied using both behavioral methods and functional brain imaging. Special interest in integration of visible and acoustic speech. (*Prof. Mikko Sams*) http://www.uta.fi/laitokset/psyk/

- *Center for Cognitive Neuroscience, University of Turku.* E.g., different aspects of auditory and speech processing. (*Prof. Heikki Lang*). http://www.utu.fi/research/ccn/

5 Summary

A short presentation of educational and research activities in Finland, related to psychoacoustics, has been presented, main focus being on the Helsinki University of Technology (HUT), Laboratory of Acoustics and Audio Signal Processing. Courses in psychoacoustics and how they are organized were described, e.g., how computer based psychoacoustics experimentation principles have been used in the course 'Communication Acoustics'. Examples of successful research projects, such as psychoacoustically oriented auditory modeling for sound quality measurements, computational modeling of spatial hearing, and psychoacoustic principles in signal processing, have been referenced. A list of psychoacoustics and hearing research activities in other institutions in Finland was compiled.

References

Backman J., and Karjalainen M., (1993). Modelling of Human Directional and Spatial Hearing Using Neural Networks. *Proc. of IEEE ICASSP-93*, Minneapolis.

Beerends, J. G., and Stemerdink, J. A., (1992). A Perceptual Audio Quality Measure Based on a Psychoacoustic Sound Representation, *J. Audio Eng. Soc.*,vol 40, No 12, pp. 963-978.

Huopaniemi, J., and Karjalainen, M., (1996). Comparison of Digital Filter Design Methods for 3-D Sound. *Proc. NorSig-96*, pp. 131-134, Espoo, Finland.

Härmä A., Laine U.K., and Karjalainen M., (1996). Warped Linear Prediction (WLP) for Audio Coding, *Proc. NorSig-96*, Sept. 24-27, pp. 447-450. Espoo, Finland.

Karjalainen, M., (1982). Measurement of Distortion in an Audio Signal Channel Based on Psychoacoustic Models. em Proc. of NAS-82, Stockholm, p. 141 - 144.

Karjalainen, M., (1985). A New Auditory Model for the Evaluation of Sound Quality of Audio Systems. *Proc. IEEE ICASSP-85*, Tampa.

Karjalainen, M., (1987). Auditory Models for Speech Processing. Invited paper, *Proc. of Int. Congr. of Phonetic Sciences*, Tallinn.

Karjalainen, M, and Rahkila, M., (1995). Learning Signal Processing Concepts and Psychoacoustics in the QuickSig DSP Environment, *Proc. IEEE ICASSP-95*, Detroit, USA.

Karjalainen, M., (1996). A Binaural Auditory Model for Sound Quality Measurements and Spatial Hearing Studies. *Proc. IEEE ICASSP-96*, Atlanta.

Karjalainen, M., Härmä, A., and Laine, U. K., (1996). Realizable Warped IIR Filter Structures. *Proc. NorSig-96*, pp. 483-486, Espoo, Finland.

Laine, U. K., and Altosaar, T., (1990). An Orthogonal Set of Frequency and Amplitude Modulated (FAM) Functions for Variable Resolution Signal Analysis. *Proc. IEEE ICASSP-90*, Vol. 3, pp. 1615-1618, Albuquerque, New Mexico.

Laine, U. K., and Huotilainen, M., (1996). A Study of Auditory Temporal Fine Resolution by Using Short Bark-Famplet Clicks. *Proc. NAM-96*, pp. 267-270, Espoo, Finland.

Laine U. K., Karjalainen M., Altosaar T., (1994). Warped Linear Prediction (WLP) in Speech and Audio Processing, *Proc. IEEE ICASSP -94*, pp. III-349-352, Adelaide, Australia.

Smith, J.O., and Abel, J.S. (1995). The Bark Bilinear Transform, *Proc. IEEE ASSP Workshop*, Mohonk, New Paltz.

Strube H. W., (1980). Linear Prediction on a Warped Frequency Scale, *J. Acoust. Soc. Am.*, vol. 68, no. 4, pp. 1071-1076.

Takala, T., (1996). An Integrated System for Virtual Audio Reality. Preprent 4229 (M-4) of *AES 100th Convention*, Copenhagen.

Zwicker, E., and Fastl, H., (1990). *Psychoacoustics—Facts and Models*. Springer-Verlag.