Psychoacoustical measurements, auditory models and their application
Outline

- Introduction of MMTG FEE CTU
  - MMTG
  - Audio group

- Psychoacoustical measurements, auditory models and their application
  - Human hearing system and models
  - Measurement
  - Applications
Multimedia Technology Group

Radio Systems Research and Development Center

Digital RAdio Communications Research Group

Broadcast Technology and Measurement Group

CAD for RF circuits

RF Embedded Systems
Multimedia technology group

- http://mmtg.fel.cvut.cz
- Imaging systems and image processing in astronomy
- Image/video and sound quality assessment, QoS, QoE
- Vision and hearing modeling, study of psychovisual and psychoacoustic processes
- Image processing in biology
- Image/video and audio compression and coding
- Multimedia information processing (audio, image/video)
- 3D audiovisual systems
- BCI and man-machine interface applications in multimedia technology
- Imaging systems and image processing in security
- Multimedia applications in assistive technologies
- Measurement and analysis of electro-acoustic and audio systems
Audio group

- **Libor Husník**
  - Digital transducers
  - Spatial sound synthesis
  - (subjective) quality of archive records

- **František Rund**
  - Auditory models and measurement
  - (objective) quality of archive records

- **Thomas Lavergne + Zdeněk Škvor**
  - Omnidirectional ultrasound PVDF transducer
  - Optoacoustic transducer
Auditory models project team

- **MMTG staff**
  - František Rund
  - Stanislav Vítek
  - Petr Maršálek

- **Ph.D. students**
  - Dominik Štorek
  - Václav Vencovský
  - Jaroslav Bouše

- **M.Eng. students**
  - Tomáš Lindner
  - Jan Stuchlík
  - Jan Štemberg
  - Ondřej Šupka
  - Sanyia Dyussekenova
Human hearing system

External, middle and inner ear
Auditory models

- **External and middle ear**
  - Cascade of simple filters

- **Inner ear – filterbank**
  - Transmission line model
  - Gamatone filters
  - Gamachirp filters
  - Biophysical models
  - ...
Binaural hearing

- Binaural summation
- Binaural masking
  - Coctail party effect
- Extraction of a signal in reverberant environment
- Localization of sound source
Basic localization cues

- **Interaural differences**
  - Time (phase) – ITD/IPD
  - Level – ILD
- **Laterazation**
- **Filtration cues (HRTF)**
- **Head movement**
- **Experience (memory)**
- **Visual information**
HRTF modeling
Olivary complex

- Decoding of spatial information
- Lateral superior olive (LSO)
  - ILD
- Medial superior olive (MSO)
  - ITD/IPD

- Models
  - Jeffress‘s delay line
  - Equalization-cancelation
  - Count-comparison
  - …
- **Monaural part**
  - Divided into 36 bands (CF 0.1 – 15 kHz)
  - Output: probability of spike firing

- **Designed binaural part**
  - Count-comparison principle
  - Output: Perceived lateralization

---

Dau et. al.: A quantitative model of the effective signal processing in the auditory system..., ASA 1996

1 model for each hemisphere
Output -> Count-comparison - subtraction

Lateral superior olive (LSO)

Yost, W. A.: Lateral position of sinusoids presented with interaural intensive and temporal differences., JASA 1981

<table>
<thead>
<tr>
<th>Pure tone frequency [Hz]</th>
<th>Correlation score [-]</th>
<th>p-value [-]</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>0.9957</td>
<td>9.7473e-13</td>
</tr>
<tr>
<td>1000</td>
<td>0.9960</td>
<td>6.8233e-13</td>
</tr>
<tr>
<td>2000</td>
<td>0.9969</td>
<td>1.7756e-13</td>
</tr>
<tr>
<td>5000</td>
<td>0.9960</td>
<td>7.1924e-13</td>
</tr>
</tbody>
</table>
- 1 model for each hemisphere
- Output -> Count-comparison - comparing ratio

Medial superior olive MSO

<table>
<thead>
<tr>
<th>Pure tone frequency [Hz]</th>
<th>Correlation score [-]</th>
<th>p-value [-]</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>0.7894</td>
<td>0.0023</td>
</tr>
<tr>
<td>500</td>
<td>0.9396</td>
<td>5.7183e-6</td>
</tr>
<tr>
<td>750</td>
<td>0.9349</td>
<td>8.2347e-6</td>
</tr>
<tr>
<td>1000</td>
<td>0.9330</td>
<td>9.5074e-6</td>
</tr>
</tbody>
</table>

Yost, W. A.: Lateral position of sinusoids presented with interaural intensive and temporal differences., JASA 1981
Medial superior olive MSO

<table>
<thead>
<tr>
<th>Central frequency [Hz]</th>
<th>Correlation score [-]</th>
<th>p-value [-]</th>
</tr>
</thead>
<tbody>
<tr>
<td>350</td>
<td>0.9809</td>
<td>3.650e-09</td>
</tr>
<tr>
<td>760</td>
<td>0.9626</td>
<td>1.402e-07</td>
</tr>
</tbody>
</table>
PSYCHOACOUSTIC EXPERIMENTS
HRTF measurement

- **Measurement chain**
  - A method for IR/TF measurement
  - Loudspeaker
  - Microphones
  - Azimuth & elevation

- HRTF – transfer from the source to the ear
- Time & instrument demanding
- Directions:
- Equalization
HRTF measurement
HRTF measurement optimization

- Automatized system for measurement
- Computer controlled rotation of the chair
  - Matlab, Arduino, step motor driver
- IR/TF measurement method implemented in Matlab
  - Sweep sine method
  - MLS based method
- Testing implementation – microphone polar pattern measurement
How to test orientation in 3D?

- Standard method

- New method
  - This method provides good results and subjects are more comfortable with understanding and completing the task
  - The task is to turn the head until the two sources overlap
HRTF test

- The same setup as for HRTF measurement
- Horizontal plane (azimuth)
- User switches real and virtual sources
- Improved azimuth detection algorithm
- Automatized data export
Psychoacoustic experiments

- Pure tone lateralization
- Narrow-band noise (1 ERB) lateralization
- Influence of headphones transfer function to lateralization
- Headphones calibration – audiology
- Dichotic pitch lateralization
- Lateralization – ITD vs IPD
Externalization experiment

- Augmented reality – mixed real and virtual
- Binaural earphones-microphones
- Set of experiments
  - Real source
  - Real source with mic-earphones
  - Recorded real source
  - Virtual source
  - ...

- With/without visual control
Auditory models applications

- Lossy compression
- Quality assessment
- Virtual reality, augmented reality
- Room acoustics,
- Cochlear implants,
- ...
Digitalization and restoration of archive records

Necessary to assess the quality

- Comparison with (old) standards – if the original tests are available
- Calculation of a measures from the signal
  - Bandwidth
  - Number of clicks
  - Dynamical range
  - ...
- Using of auditory models
  - Summation of the artifacts – is not linear
  - E.g. masking takes place
  - Algorithms like PEAQ, PEMO-Q is not optimized for archive records
Testing is complicate and time demanding
Verification of VAS algorithms using models?
Conclusions

- Human hearing and auditory models
- New findings necessary to implement (binaural hearing)
- Auditory models are used for wide range of applications
- Simple models, not accounting for some effects (e.g. binaural)
Thank you for your attention

QUESTIONS?