Psychoacoustics

Notes on Speech and Audio Processing

Chia-Ping Chen

Department of Computer Science and Engineering
National Sun Yat-Sen University
Kaohsiung, Taiwan ROC

Introduction

- Psychoacoustics is the science in which we quantify the human perception of sounds. We aim to derive a quantitative model that explains the results of auditory experiments.
- We relate psychoacoustic phenomena to physiological or physical measurements.
- Roughly speaking, a perceptive variable has its corresponding objective variable. But it’s a little more complicated than that.

Objective and Perceptual Variables

- Objective variables can be measured by instruments.
- For example, frequency can be measured by zero-crossing rate. Spectrum can be defined by Fourier transforms.
- The value of a perceptual variable can only be determined via psychoacoustic experiments. It is subjective.
- An objective variables has a corresponding perceptual variable.

Correspondence

- An approximate correspondence between objective and perceptive variables is

$$\begin{cases} \text{frequency} \Leftrightarrow \text{pitch} \\ \text{intensity} \Leftrightarrow \text{loudness} \\ \text{spectrum} \Leftrightarrow \text{timbre} \end{cases}$$

- The relation is non-linear: if the intensity is doubled, a subject does not “percept” a doubled loudness.
- In reality, the measure of a perceptive variable can depend on more than a single objective variable.

Psychoacoustic Tests

- It is natural to ask the following questions:
  - How does ear respond to different **intensities**?
  - How does ear respond to different **frequencies**?
  - How well does the ear focus on a given sound of interest in the presence of interfering sounds?
- Such questions can be quantitatively addressed by conducting psychoacoustic tests.
- The design of psychoacoustic tests, as well as the explanation of the results, is a tricky matter.

Sound Pressure Level

- Sensation level often varies logarithmically with the stimulus.
- The sound pressure level of pressure p is defined by

$$\mathrm{SPL} = 20 \log \frac{p}{p_0} \text{ dB},$$

where $p_0 = 2 \times 10^{-5} \frac{\mathrm{N}}{\mathrm{m}^2}$ is the threshold of hearing at 1000 Hz.

Loudness and SPL

- An empirical relation is found on loudness:

$$S \propto p^{0.6} (\propto I^{0.3}).$$

  - It is a **cubic-root** law.
  - The proportional constant depends on frequency.
  - Human ears are most sensitive at 4 kHz.
- The equal-loudness curves, as shown in Figure 15.1, display the dependence of intensity on frequency for given loudness levels.
- A loudness measure is *phon*, defined to be the SPL in decibels at 1000 Hz.

Loudness and Duration

- Experiments show that if the duration of a sound is less than 200 ms, then it will appear less loud than if it were longer than 200 ms.
- Figure 15.2 illustrates this point. Basically, the minimum audible level of intensity increases as the duration decreases below 200 ms.
- Apparently, some form of temporal integration is at work when perceiving loudness.

Critical Bands

- Critical-band experiment
  - have a tone audible in band-limited white noise.
  - decrease the tone intensity until it is inaudible. Record the tone intensity.
  - repeat with a reduced bandwidth of noise.
- It is found that when the band-width is reduced beyond a critical value (called critical band), the listener's response monotonically increased. So what counts is the SNR within the critical band.
- The critical band is larger for higher-frequency tone, implying that the human auditory filters have greater bandwidths with higher center frequencies.

Masking

- When two tones are presented simultaneously, the weaker one may be masked (not heard).
- Closer tones have a greater masking effect, and a louder tone affects tones further away.
- A tone masks a higher tone easier than a lower tone.
  - Figures 15.6 and 15.7 illustrate the **asymmetry** of masking. Here the target and masker signals are reversed and the number of times by which the lower tone masks the higher tone than the converse is shown.

Summary

- Loudness is roughly proportional to the cubic root of sound intensity.
- Loudness is dependent on frequency, as demonstrated by the equal-loudness curves.
- Loudness is dependent on the duration below 200 ms, implying some kinds of temporal integration.
- Critical-band filtering of the auditory systems does exist.
- Tones mask one another in an asymmetrical way.