Speech Signal Processing (SSV)

Examiners	Prof. Dr. Timo Gerkmann
Exam date	07/16/2020
Department	Informatik
Labels	mündlich

The oral exam was via Zoom. I needed a webcam + microphone as well as a second device with a camera but without audio to face my desk with empty paper which I could write on during the exam or a device like a tablet where you could directly write on.

Other than that, the exam was topic-wise very closely related to previous years' protocols which really helped me to prepare. So, I just shortly sketch my exam. First, he gave an introduction to the topic of speech processing before he started with his questions.

- sketch the source-filter model (source -> filter -> speech signal)
- then he asked for more details (unvoiced/voiced):
    - noise generator
        - e(n) -> vocal tract filter (h(n))-> speech signal (s(n))
    - T0, pulse-train
- write down the filter output: $s(n) = e(n) * h(n)$
- transform it to frequency domain: $S(z) = E(z) \cdot H(z)$
- derive the filter from time in z-domain
    - $s(n) = \sum h(m) \cdot e(n-m) = \sum b_m e(n-m) - \sum a_v s(n-v)$
    - $S(z) = E(z) \sum b_m z^{-m} - S(z) \sum a_v z^{-v}$ I did not have to derive that further but better be prepared
- add b and a in the previous drawn source-filter model (ARMA)

After talking about the source-filter model in detail, we switched to sampling/quantization:

- digitization steps, the difference between sampling and quantization, which can be reconstructed perfectly?
- sampling theorem?
- he showed me slide 206 - SNR for Uniform Quantization
    - form factor?
    - overload?
    - SNR?
    - $P_N = \frac{\Delta x^2}{12}$

Speech Enhancement:

- $S_k(l) = G_k(l) \cdot Y_k(l), Y = S + N$

- $G_k(l) = \dfrac{\sigma_S^2}{\sigma_S^2 + \sigma_N^2}$
- some questions regarding these formulas