

Question Set 1

Basic Questions

1. Describe the basic mechanism of speech production.
 - Speech signal is the result (convolution) of two phenomena; what are they?
 - Define (a) Pitch and (b) Formants
2. What is the goal of speech coding? What is the input to the coding system and what is the output? What is the “standard” bit rate after speech coding (calculate an example)?
3. What is the goal of speech recognition? What is the goal of text-to-speech (speech synthesis)? In each case, what is the input and what is the output? In both cases, which phenomena makes the problem difficult (definition, why, solutions, etc)?
4. What is the main difference between speech recognition and speaker verification?
 - Definition of each problem (goal, input, output)
 - The different characteristics/properties of the speech signal that is extracted from the speech signal for each case.
 - The “theoretical” criteria for each case?
5. Given that the speech signal is sampled at the rate of 8kHz, what is the pitch frequency if the time period between two pitch pulses (in case of a voiced sound) is 5 msec?

What is the pitch frequency in the case when the number of samples between two pitch impulses is 64 samples?

Of the two pitch frequency values, which one of the speech signal is probably produced by a female speaker?

6. Describe and illustrate, what is
 - An analysis window: definition, typical length, for what etc.
 - A power spectrum
 - A spectrogram
 - On power spectrum as well as on spectrogram, what are the typical properties of the speech signal that can be observed?
7. Briefly describe different methods (in time domain and frequency domain) to extract pitch frequency.