AUDIO ACQUISITION

RECORDING

Sound recording is an electrical or mechanical inscription and re-creation of sound waves, such as spoken voice, singing, instrumental music, or sound effects. The two main classes of sound recording technology are analog recording and digital recording. Acoustic analog recording is achieved by a small microphone diaphragm that can detect changes in atmospheric pressure (acoustic sound waves) and record them as a graphic representation of the sound waves on a medium such as a phonograph. In magnetic tape recording, the sound waves vibrate the microphone diaphragm and are converted into a varying electric current, which is then converted to a varying magnetic field by an electromagnet, which makes a representation of the sound as magnetized areas on a plastic tape with a magnetic coating on it.

The purpose of recording is to capture the best possible signal. Regardless of the recording situation, one should consider at least three important acoustic parameters: frequency response, dynamic range, and signal-to-noise ratio(SNR).
Frequency response describes the range of frequencies captured in the recording, measured in Hertz (Hz).
Dynamic range is the ratio of the loudest to the softest part of the signal, measured in Decibels (dB).
SNR is the ratio of signal amplitude to the amplitude of noise usually generated by the circuit, measured in Decibels (dB). In the audio digitization process, signal-to-noise ratio refers to the ratio of the maximum signal power to the quantization noise generated by the analog-to-digital converter.
Digital recording and reproduction converts the analog sound signal picked up by the microphone to a digital form by the process of digitization. This lets the audio data be stored and transmitted by a wider variety of media. Digital recording stores audio as a series of binary numbers representing samples of the amplitude of the audio signal at equal time intervals, at a sample rate high enough to convey all sounds capable of being heard. Digital recordings are considered higher quality than analog recordings not necessarily because they have higher fidelity (wider frequency response or dynamic range), but because the digital format can prevent much loss of quality found in analog recording due to noise and electromagnetic interference in playback, and mechanical deterioration or damage to the storage medium.

QUANTIZATION

Quantization is the process of mapping a large set of input values to a (countable) smaller set, such as rounding values to some unit of precision. A device or algorithmic function that

performs quantization is called a quantizer. The round-off error introduced by quantization is referred to as quantization error.

SAMPLING

In signal processing, sampling is the reduction of a continuous signal to a discrete signal. A common example is the conversion of a sound wave (a continuous signal) to a sequence of samples (a discrete-time signal). When it is necessary to capture audio covering the entire 20–20,000 Hz range of human hearing, such as when recording music or many types of acoustic events, audio waveforms are typically sampled at 44.1 kHz (CD), 48 kHz, 88.2 kHz, or 96 kHz. The approximately double-rate requirement is a consequence of the Nyquist theorem.

DIGITIZATION

Digitization is the representation of an object, image, sound, document or a signal (usually an analog signal) by a discrete set of its points or samples. The result is called digital representation or, more specifically, a digital image, for the object, and digital form, for the signal.

Pre-amplifiers

The main function of a pre-amplifier is to accept a very low-level signal (such as that from a microphone) and amplify it without adding noise. Good pre-amplifiers are not easy to build, as they have to be immune to all kinds of potential noise and signal distortion.

Recording devices

There are several types of recording devices used today. Some of the most common include analog cassette recorders, digital audio tape (DAT) recorders, minidisk recorders, and hard disk recorders.

Analog recorders

The analog tape is able to provide the necessary frequency response and dynamic range to capture a fair amount of detail in the speech signal.

Digital recorders

Digital recorders are becoming more and more common. They come in a variety of flavors, such as DAT, minidisk, solid state PC card, CD, and hard disk. What distinguishes them from one another are the recording medium and the recording format.

DAT recorders

DAT recorders use magnetic tape as a storage medium. Typical DAT tapes allow the storage of 120 minutes of uncompressed, high-quality mono recording at 48,000 Hz/16-bit.

Analog-to-Digital conversion and digital data transfer

In order to perform acoustic analysis on recorded speech data, the audio signal has to be converted into a digital audio file format, such as Wav or Aiff. Analog recordings have to be

digitized and digital recordings need to be transferred to a personal computer via a digital audio file transfer interface. This is an important, yet often underestimated, stage in the process of preparing audio data for analysis. The main goal of A/D conversion (digitization) is to obtain the best possible digital representation of the original analog waveform. Without going into too much technical detail of the digitization process, one should choose a sample rate that will capture a broad range of frequencies and a bit-depth that will allow a wide dynamic range and a negligible amount of quantization noise. These goals can be achieved by means of a premium-quality, stand-alone A/D converter operating at the sample rate of at least 48,000 Hz and a 24-bit resolution. It is absolutely crucial to not to use a PCI multimedia sound card, as they are built from inferior-quality electronic components and, more importantly, allow electrostatic noise and distortion to leak into the captured acoustic signal.

Python libraries to be used (tentatively)

Wave - Provide an interface to the WAV sound format

PyAudio - Python bindings for PortAudio audio input and output