Authors: Nina Khayyami & Jens Nilsson (E10)

Automatic Gain Control & Psychoacoustic Modeling for Near End Listening Enhancement

When communicating through a loudspeaker in a noisy environment, one solution to avoid degraded speech comprehension is to use an adaptive gain control, together with a voice activity detector and power estimation. The adaptive gain control adjusts the gain of the far end speech signal according to the loudness of the noise. An optional psychoacoustic filter can be applied to enhance the intelligibility of the speech signal on top of the adjusted gain.

When listening to a speech signal through a loudspeaker in a noisy environment, it can be difficult to comprehend what is being said if the background noise is unusually loud or if sudden undesired noise peaks appear. For instance, in door station devices, the noise in the environment in which the device is placed might mask important frequency components of the speech signal. The effect of this is degraded comprehension of the speech signal which derives from simultaneous masking and low signal to noise ratios, due to noise that is added to the speech signal.

One solution to this is a digital adaptive gain control which analyzes the loudness of the noise and automatically adjusts the gain of the far end speech signal. The gain is computed by using power estimation of the surrounding noise in the frequency domain. By doing this we can increase the signal to noise ratio by the listener’s ear in near end and increase the speech intelligibility.

The system has the possibility to add a feature consisting of a psychoacoustic filter which analyzes the frequency components of the incoming noise and computes a suitable filter for each frame of the speech signal. The filter is based on a psychoacoustic model which locates masking tones and noise in the noisy environment and compensates the masked frequencies in the far end speech signal.

Testing the system with initial signal to noise ratio values between -10 dB to 20 dB has shown that the signal to noise ratio can be improved by up to 19.2 dB without affecting speech quality, and the intelligibility can be improved by up to 43%. Speech quality and intelligibility is evaluated with a standard measurement called PESQ.

The thesis has been carried out at AXIS Communication AB and the presented system could for instance be used in devices such as AXIS’s new door station product, A8004-VE.