A MULTIRESOLUTION SPECTRAL SUBTRACTION ALGORITHM FOR NOISE SUPPRESSION IN AUDIO SIGNALS

Tkachenko M., Lukin A.

Moscow State University, Faculty of Computational Mathematics and Cybernetics,
Laboratory of Mathematical Methods of Image Processing

Noise is an undesirable signal appearing during transmission or measurement of another clean signal. There are two categories of noise, by spectral properties: stationary (the one that does not change in time) and non-stationary.

Additive noise is summed with the clean signal $y[t]$ and does not depend on it: $x[t] = y[t] + noise[t]$, where t is time and $x[t]$ is the observed signal. Constant hiss from a microphone or an amplifier, electric hum are the examples of additive stationary noise.

In this paper, we propose a method for suppression of additive stationary noise using a multiresolution STFT (short-time Fourier transform). This approach is able to improve quality of many audio processing algorithms (such as noise reduction) by adaptively varying STFT time-frequency resolution based on local properties of the signal: estimates of spectrogram sparsity.

The simple method of spectral subtraction (SS) is widely used for reduction of additive stationary noises. It works with a fixed STFT window size. Long windows provide good frequency resolution and achieve accurate separation of closely spaced noise and signal harmonics. However long windows sizes also lead to poor time resolution and increase smearing of transients (sharp attacks or fast changes in the signal). On the other hand, short-window STFT processing is inefficient in terms of frequency resolution and possible depth of noise suppression.

A filter bank with a fixed window size cannot provide good frequency and time resolution simultaneously. It is required to select best resolution for each local part of the signal during the processing. We propose using a multiresolution STFT to solve this problem.

The algorithm of MR STFT consists of three parts:

1. Calculation of several copies of the signal processed with different STFT window sizes;
2. Estimation of smearing (or sparsity) for local spectrogram areas for each resolution;
3. Mixing of the resulting spectrograms based on spectrogram sparsity estimates.

The effect of this adaptive algorithm is selection of high frequency resolution for tonal signals and selection of high time resolution for transients.

A noise reduction system based on MR STFT has been created. Several modifications of spectral subtraction rules have been implemented, including a highly effective method of Non-Local Means for smoothing of a “musical noise” artifact. During algorithm testing we have found optimal range of STFT sizes for the MT STFT frameworks, sizes of spectrogram mixing areas and other key parameters.

We have compared the performance of a simple spectral subtraction (SS) with several STFT window sizes, SS based on MR STFT (SS MR STFT), and SS with a Non-Local Means based on MR STFT. It has been shown that SS MR STFT produces results significantly different with a simple SS: after processing with a simple SS noise power decreased by 13.00 dB, while with SS MR FFT – by 16.10 dB. The application of a Non-Local Means smoothing has removed musical noise and further improved overall processing quality.

ОБ ОДНОМ СПОСОБЕ ЧАСТОТНОГО АНАЛИЗА ПЕРИОДИЧЕСКИХ СИГНАЛОВ

Колоколов А.С.

Институт проблем управления им. В.А.Трапезникова РАН, Москва

В настоящее время цифровой спектральный анализ является важным инструментом в практике анализа гармонических, периодических и квазипериодических сигналов. Примерами таких сигналов являются музыкальные звуки, вокализованные фрагменты речи, а также разного рода сигналы в задачах акустической диагностики.

Нахождению спектра сигнала $s(t)$ цифровыми методами предшествует его дискретизация с частотой f_S, в результате чего непрерывный сигнал представляется последовательностью дискретных отсчётов $s(n\Delta t)$, где n – номер отсчёта, принимающий значения $n = \dots -2, -1, 0, 1, 2, \dots$, $\Delta t = 1/f_S$ – интервал квантования. В этом случае комплексный спектр отрезка дискретного сигнала $s(n)$, заданного на интервале $n = 0, 1, \dots, N-1$, определяется выражением

$$\widetilde{S}(m) = \sum_{n=0}^{n=N-1} s(n)w(n)e^{-2\pi jmn\Delta f\Delta t} = a(m) + jb(m)\ ,$$