
NOISE REDUCTION TECHNOLOGY WHITE PAPER

August 2012

SCOPE

SpeechPro offers professional real-time noise reduction and speech enhancement software and hardware that uniquely combine all state-of-the-art algorithms and speech enhancement tools. Being a world leader in developing the technologies of audio analysis and speech enhancement, SpeechPro has supplied its noise reduction solutions to law enforcement agencies, military and intelligence services of many European, Latin American and Asian countries.

INTRODUCTION

The intelligibility of human speech plays an important part in communication. It is both a measure of comfort and comprehension.

The quality and intelligibility of the speech are not only determined by physical characteristics of the speech itself but also by communication conditions and information capacity, the ability to get the information from context, mimics and gestures.

When discussing intelligibility it is important to understand the difference between a real and recorded speech.

During a real conversation a person can recognize the surrounding sounds and concentrate on the speech of another person thus filtering the desired information out of various audio environments. Therefore the ability of a human to recognize and filter sounds significantly increases the intelligibility and comprehension of the speech even if a communication takes place in a noisy environment, situation or condition.

Listening to recorded speech is a different situation. The recording equipment doesn't focus on certain audio streams (unless it is a specialized shotgun microphone) and impartially record everything that happens in the audio spectrum. As a result we receive a "flat picture" of all recorded sounds which often makes the speech unintelligible, quiet and buried in the noises.

Additional reasons why speech recordings may be indistinct and distorted can be due to technical limitations of recording equipment, poorly placed or defective microphones and objective difficulties to record high quality "clean" sound.

As audio recording technologies achieved wider use since the middle of the 20th century the demand for audio processing and noise reduction has also increased exponentially. Even now when audio equipment has less limitations and allows for better quality the need for noise suppression is still of the utmost importance especially in the area of security and law enforcement.

Police departments, military and national security services largely use overt and covert recordings of speech communications that can be a crucial element in investigations and intelligence operations. Needless to say that sometimes audio recording may be the only evidence of a security threat or crime and therefore may become a key element in the case analysis or subsequent court trial. In these cases it is important for speech to be clear and easily understandable to ensure no vital information is lost. Moreover the intelligibility of audio evidence is a must for court proceedings as otherwise they might be eliminated from consideration.

Improving the intelligibility of a speech signal, reducing the noises and compensating for distortions is a main task of noise reduction technology that is currently available through different software and hardware products.

As a global leader in speech technologies SpeechPro Inc. has been developing specialized tools for efficient noise reduction and text transcription of low quality recordings for over 20 years. Various studies on the perception of poor audio recordings and noisy speech signals carried out by SpeechPro have resulted in the creation of the unique sound filtering algorithms that are now presented in the software and hardware products like Sound Cleaner, ANF II and The Denoiser Box.

This white paper is aimed at discussing the basics of noise reduction technology, its methods, goals and where SpeechPro Inc. is moving in this area.

CLASSIFICATION OF AUDIO HINDRANCES

To understand the basics of noise reduction technology and successfully use its methods in practice it is important to know what audio hindrances there are, how they differ and their specific unique attributes.

Generally all audio hindrances are divided into two main categories: noises and distortions. If we consider an original human speech in a recording as **a useful signal** all the additional information which decreases the quality of a useful signal are **noises**. Everything that changes the original useful signal itself are considered **distortions**.

Noises are mostly characterized by time and frequency (domains). In time domain noises can be:

- ▶ **Continuous**, slowly changing noises, like the sound of an office, industrial equipment, sound of the wind, traffic, hiss of an old record or a bad phone line.
- ▶ **Discontinuous**, repeated, usually tonal noises like honks, beeps or bells.
- ▶ **Pulse like**, abrupt, usually unharmonious and sometimes loud noises like clicks, taps of the steps, gunshots, bangs and thumps.

In frequency domains noises can be:

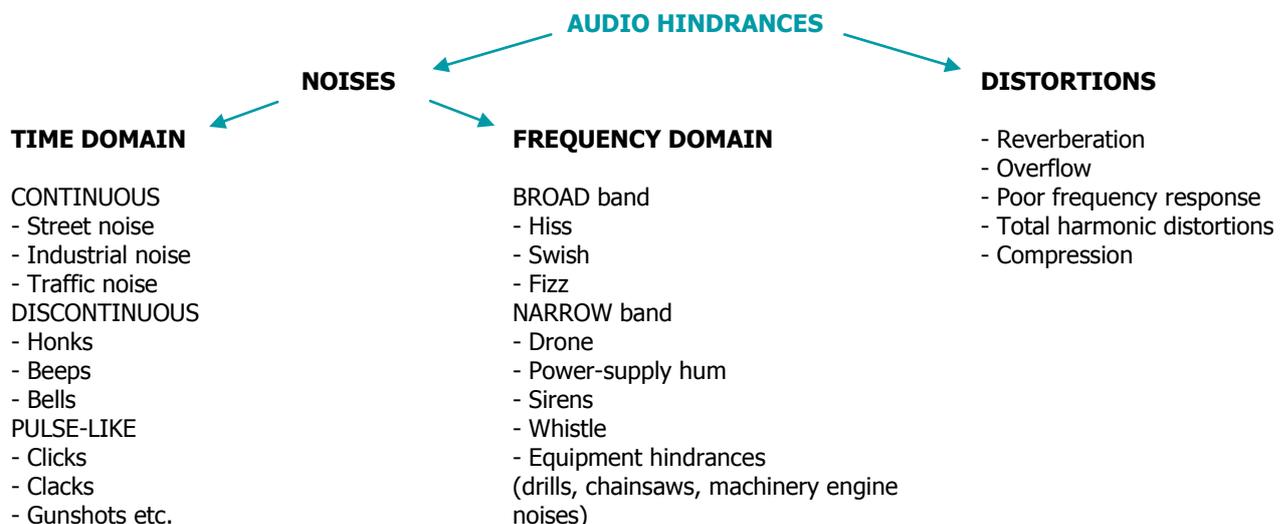
- ▶ **Broad band** noises which present at many frequencies like background hiss or fizzing sounds.
- ▶ **Narrow band** noises which represent a set of certain frequencies, fairly stable tonal sine waves (sinusoid): drones, power-supply hums, equipment hindrances (drills, chainsaws) machinery engine noises.

Distortions are modifications of the useful speech signal that decrease its quality. When distortions occur, parts or the whole speech signal changes and become new and sometimes can sound unacceptable. Typical distortions at acoustical level are reverberation and echo effects.

Distortions also occur when the acoustic signal (speech) transforms into electrical signal and meets various technical limitations like:

- ▶ Filtration of the audio signal caused by poor frequency response (FR) of the recording equipment or communication channel.
- ▶ Loss of the useful data caused by narrow dynamic range.
- ▶ Overflow effect which occurs when the amplitude of the acoustic signal is higher than the amplitude that a microphone can process.
- ▶ Total harmonic distortions which are the additional tones (harmonics) that mask a real signal components and make it indistinct and incomprehensible.
- ▶ Recording audio data in a compressed lossy format.

Generally the noise reduction technology helps to deal with such kinds of distortions however some types of distortions may completely destroy the useful information and cannot be restored during further signal processing.



NOISE REDUCTION METHODS

The process of noise reduction touches on a lot of questions regarding different fields of science (digital signal processing, acoustics, psychoacoustics and physiology) and engineering (programming, constructing etc.).

Its effectiveness depends on the correspondence between the method of processing and the type of audio interferences. Each digital filtration method is more effective for a specific kind of noise. This is why it is necessary to know at least generally what kind of audio hindrances changed an audio recording in order to choose an appropriate processing method. One can identify the audio hindrance in the recording by either the specific sound of the noisy signal or by the analysis of its spectrum and waveform.

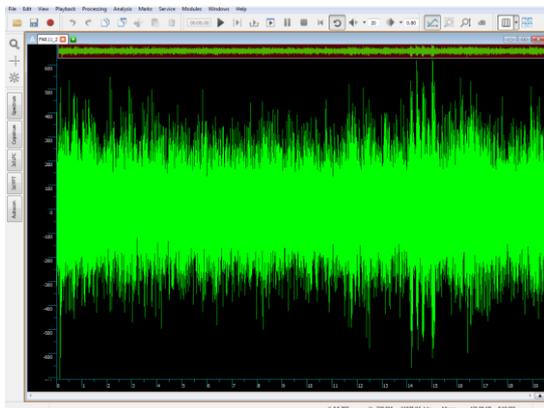
However various noises and distortions sometimes may sound similar therefore the most popular method to identify an audio hindrance is the analysis of the spectrum and the waveform. As noise characteristics usually change over time, it is necessary to use the special processing method that provides an automatic adjusting to noise characteristics. Digital filtration algorithms that may adjust to a certain type of audio hindrance are called adaptive filtration algorithms.

SpeechPro Inc. extensively uses adaptive algorithms of a new generation in its hardware and software products:

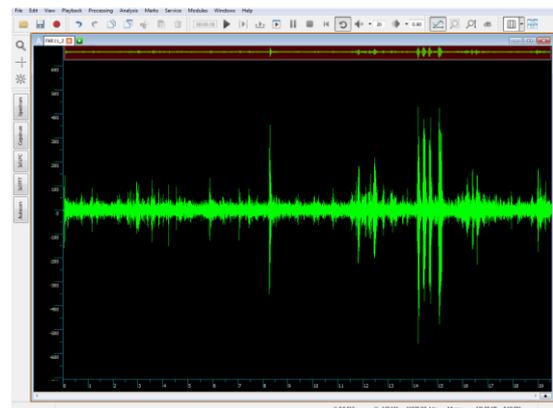
Adaptive broadband filtration
Adaptive inverse filtration
Frequency compensation
Impulse filtration
Dynamic processing
Stereo processing

Adaptive broadband filtration is based upon an adaptive frequency algorithm. This algorithm is designed to suppress broadband and periodic noises due to electric pick-ups or mechanic vibrations, room and street noise, communication channel or recording equipment interferences. You may hear these noises as hum, rumbling, hisses or roars. The broadband filtration method usually consists of two processing procedures of adaptive spectral noise subtraction that allows to enhance the speech and adaptive background extraction that separates background acoustic environment from the useful signal. It is nearly impossible to remove such noises with other methods, such as one-channel adaptive filtration, spectrum smoothing or equalizer, because the noises are spread across the whole spectrum and intersect with the speech signal.

Recorded conversation between two people in the noisy street



Before reduction



After reduction

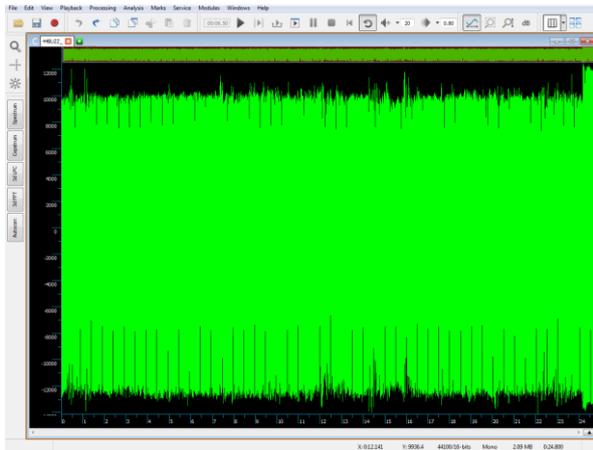
Adaptive inverse filtration process is based upon the adaptive spectral correction algorithm, sometimes also called adaptive spectral smoothing.

Adaptive inverse filtration effectively suppresses strong periodic noises from electrical pick-ups or mechanical vibrations thus recovering speech and equalizing the signal. It amplifies weaker signal components and suppresses the stronger ones at the same time. The average spectrum therefore tends to approach the flat spectrum enhancing the speech signal and improving its intelligibility. Broadband noises, however, usually become stronger making signal perception less comfortable. It means that you should try to reach a compromise between noise reduction and speech perception.

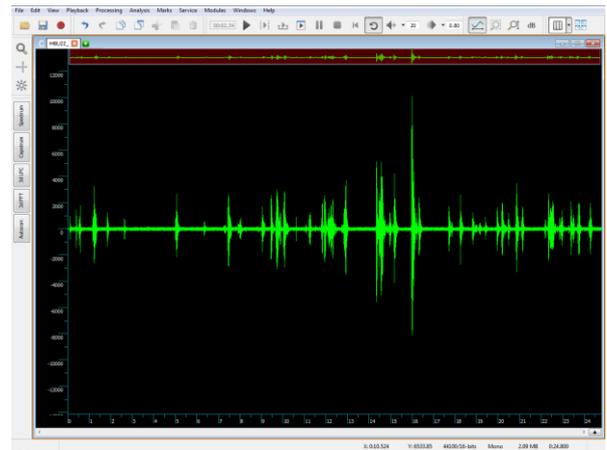
Frequency compensation process uses the Widrow–Hoff adaptive filtering algorithm of one-channel adaptive compensation. It is most effective for narrow-band stationary interferences. The filter adjusts itself smoothly maintaining good quality of the speech. The frequency compensation in this process also provides adaptive compensation in time domain.

Frequency compensation enables the ability to remove both narrowband stationary interferences as well as regular ones (vibrations, power-line pickups, electrical device noises, steady music, room, traffic and water noises, reverberation etc.). The main advantage of this process is its capability to preserve the speech signal much better than other filters usually do. Since the audio interferences in some cases may be removed only partially, it is possible to use frequency compensation method more than once.

Power-line buzz masking the conversation between two people



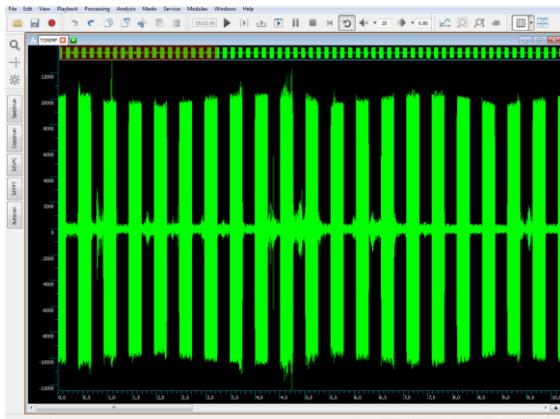
Before reduction



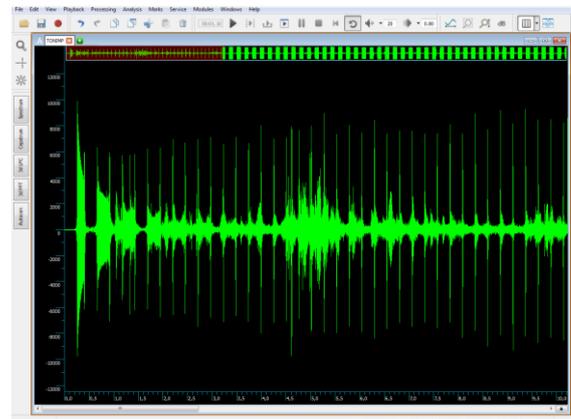
After reduction

Adaptive impulse filter automatically restores speech or musical fragments distorted and masked by various pulse interferences such as clicks, radio noises, knocks, gunshots etc. Adaptive impulse filtering algorithms improve the quality of the signal suppressing powerful signal impulses and thus unmasking the useful audio signal and increasing its intelligibility. During impulse filtration it substitutes impulses with smoothed and weakened interpolated signals. If the algorithm does not detect an impulse, it leaves the fragment intact. It also does not suppress tonal interferences and broadband noises. Impulse detection is based upon the information about the differences between the useful signal and an interference that the algorithms automatically detects.

Tapped phone conversation interfered by another line's beeping



Before reduction



After reduction

Dynamic signal processing improves the intelligibility of the speech if the signal fragments greatly differ in level, in the case of resonant knocks (i.e. long impulses) and room noises. Dynamic processing algorithms improve and unmask the audio signal suppressing the powerful impulses and clicks and reducing the listener's fatigue in case of long audio recordings.

Stereo filtration is one of the latest innovations in the field of noise reduction technologies. In some cases the problem of removing the noises can be resolved with the help of dual-channel audio information monitoring and further dual-channel adaptive filtration (stereo filtration). This method however is more sensitive to the audio recording process and its quality because it requires the more accurate use of two or more microphones. There are two methods of stereo filtration available: two-channel signal processing and adaptive stereo filtering. In the first case the sound in each channel is processed independently, while in the second case data acquired from one channel (reference channel) is used for filtering the signal in the second one (primary channel). Stereo filtration is the most effective way to control the audio environment full of various hindrances. This method effectively reduces background music and crowd noises enhancing the useful speech signal and is perfect for the recordings in big-sized rooms like halls, restaurants, theaters etc.

SpeechPro's product line includes solutions suitable both for real-time noise reduction during field operations and processing of the recorded data at audio lab or other location after recording. Providing automatic and manual modes, SpeechPro's products may be of a great value for both non-experts as well as audio specialists and experts.

SOLUTIONS

Being a leader in the field of speech technologies and sound analysis SpeechPro Inc. is constantly working on the improvement of its products and technologies in the area of noise reduction and speech enhancement. The noise reduction solutions that SpeechPro offers to the market currently are heading in three directions:

1. Expert systems
2. Automatic systems
3. Research and development

1. Expert systems are the professional hardware and software sets for advanced audio speech signal analysis that include powerful noise reduction tools that provide an excellent solution for solving the following tasks:

- ▶ speech enhancement and audio restoration,
- ▶ transcription of low quality recordings
- ▶ speaker identification
- ▶ authenticity analysis of analog or digital audio recordings
- ▶ testing and identification of audio equipment
- ▶ analysis of noise, acoustic environment and recording conditions

Some of the key elements of SpeechPro's expert systems is a unique sound cleaning software application that received the first prize in audio enhancement contest organized by AES (Audio Engineering Society) in 2008 and our present professional real-time noise cancellation and speech enhancement software.

(AES logo as in catalog and brochures)

SpeechPro's expert systems were highly evaluated by world-class experts in forensic audio analysis and were adopted by law-enforcement agencies throughout the USA, Europe and Latin America.

2. Automatic systems are mostly characterized by compact real-time noise filtering and speech enhancement devices that may be of great value for police, surveillance teams, private investigators, forensic labs and other law enforcement agencies. They can be used in real-time sound and speech quality improvement while recording or listening in field conditions. Moreover SpeechPro's hardware solutions can be of great interest for audio engineers working in the area of mobile processing of recorded audio data and broadcasting in terms of "live" mastering of the interviews and reports.

Being mobile and compact SpeechPro's hardware is effective against different noise sources: communication channel interferences, office equipment, industrial and vehicle engines, street traffic, environmental noises, background music, hiss and rumbling, reverberations and echo effects. They also provide the original methods for stereo processing using algorithms by reference channel.

3. Research and development solutions in the area of noise reduction are mostly presented as cross platform libraries, automatic/manual algorithm adjustment and real-time/post-processing embedded/workstation implementation

SpeechPro's SDK noise reduction features:

- ▶ Broadband Noise Filter/Canceler
- ▶ Equalizer (EQ), Graphical EQ, Adaptive EQ, Parametric EQ
- ▶ Dynamic Range Control, Sound Level Limiter
- ▶ Automatic Gain Control
- ▶ Level Control, Speech Level Enhancement
- ▶ Punch & Crunch Dynamics Processing
- ▶ Acoustic Shock Protection, Adaptive Shock
- ▶ Attenuator/Limiter [DSP-factory]
- ▶ Harmonic Reject Filter : Adaptive & Fixed COMB
- ▶ Hiss Filtration

CONCLUSION

Generally noise reduction methods were developed to unmask the useful signal hidden in different types of audio hindrances.

The noise reduction standard approach lies in the principle of the most practical removal of unnecessary extraneous sound components and returning the distorted parameters to their typical values. The most typical noise suppression goal is useful signal unmasking, i.e. to suppress noisy signal components in the areas where the hindrances are powerful and the useful signal is weak and to intensify those components where the useful signal is at maximum.

Thus the basic principles of noise reduction technologies are:

1. Unmasking the useful speech signal in time and frequency domains taking in consideration psycho-acoustic properties of human speech hearing.
2. Removing different kinds of background noises to decrease tiredness during listening.
3. Decreasing the frequency pass band of the signal to and removing low-frequency drones and high-frequency hisses.
4. Smoothing the high peaks and decreasing the audio signal amplitude in pauses without the speech.
5. Removing or decreasing pulse-like interference amplitude and other intensive outside sounds.
6. Removing regular slowly changing hindrances: music, traffic and industrial noises, decreasing reverberation (echo effects).
7. Smoothing signal spectrum.
8. Additional subtraction of narrow band interferences.
9. Removing additive broadband noises (tape, radio, phone and microphone hiss).