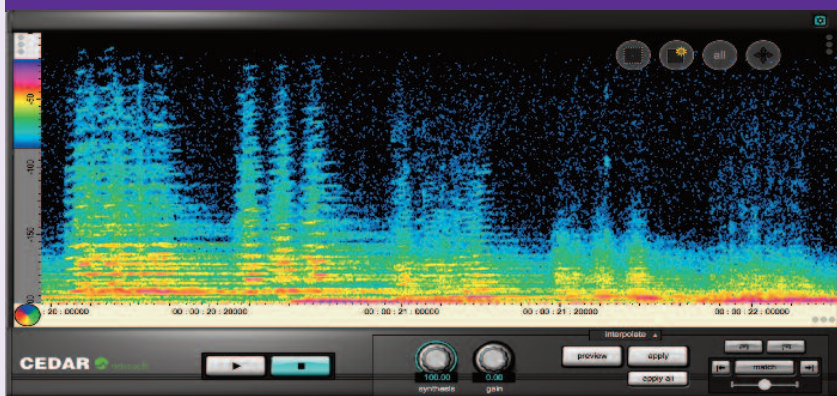


Cedar Audio: committed to noise suppression, speech enhancement and audio restoration



Gordon Reid
Managing
Director, CEDAR
Audio Ltd

CEDAR Audio is a UK based company committed to noise suppression, speech enhancement and audio restoration. It has focussed exclusively in these areas for more than three decades and is the recipient of numerous accolades including an IABM Design & Innovation Award, two Cinema Audio Society Awards and an Academy Award for services to the movie industry.



When CEDAR was established in 1989, several universities were researching what soon became known as digital audio restoration – the science of removing unwanted sounds such as clicks, crackle and hiss from existing recordings. These ‘single-ended’ processes were quite different from existing noise reduction methods that encoded and then decoded the audio to limit the noise added by the medium; they attempted to identify and remove unwanted sounds that already existed in the signal without adversely affecting the wanted sound.

Early processes were limited by the state-of-the-art of digital signal processing (which had only recently been applied to audio) and the processing power of the available hardware. At the time, there were just two companies active in the field. One

chose to implement all of its processes outside of real-time, thus allowing more computing power to be applied to each sample of the audio. In contrast, CEDAR chose to adopt newer, more powerful processors and to optimise its algorithms so that they could be applied in real-time. This immediately became CEDAR’s trademark; whatever we did, we did it in real-time so that the user could hear the effect of the processing as it was occurring. This is of much greater benefit than it might seem. If you can tweak a process while it’s running, you can soon identify the parameters needed to obtain optimum results. If you have to come back the following morning to listen to what you’ve done, you cannot. Real-time processing also removed the need for extensive hard disk storage, which was hideously expensive at that time.

Within a few years, it was apparent that the philosophy of real-time audio restoration was leading the company far beyond the bounds of libraries, archives and remastering for CDs and DVDs, and into areas such as broadcast, post-production and audio forensics. At the same time, solutions to other problems were being developed, and it was soon possible to remove complex buzzes, clipping distortion, timing errors between tracks, speed changes during a recording, and more.

However, neither the techniques nor the hardware of the 1990s were suitable for live broadcast because of the constraints on latency. Humans are very sensitive to any loss of synchronisation between lip movement and heard speech, and many (especially older) people are unaware of the degree to which they rely upon the former to aid comprehension. Consequently, the latency of any noise reduction process used for live broadcast has to be as close to zero as possible. The breakthrough in this area came in 2000 when CEDAR invented the digital ‘dialogue noise suppressor’ (DNS) and incorporated this within dedicated hardware so that the latency could be kept below 0.2ms at 48kHz. Although the earliest products were designed for post-production, units soon started to appear in areas such as newsrooms, reality TV, games shows and sports commentating.

Early versions of DNS were reasonably benign with regard to over-processing and the generation of unwanted artefacts, but they still required a degree of understanding and manual control to obtain optimum results. So the hunt began for a more autonomous version.

In 1994, CEDAR had released a product called the DH-1 dehisser, which used a very early implementation of machine learning to identify, track and remove

"Any sufficiently advanced technology is indistinguishable from magic"
Arthur C Clarke once wrote



the broadband noise contained within a signal. Common wisdom at that time suggested that this task was impossible without the aid of a noise fingerprint, but the DH-1 and its successors proved to be remarkably successful and remained in production until 2016. But in 2015, CEDAR refined its latest machine learning technology (often, but erroneously called 'AI') to create the Learn capabilities of a new generation of noise suppressors that offered the performance and near-zero latency of DNS while eliminating the need for complex controls. This means that products could be made smaller, lighter, simpler to use, and at a lower cost.

Of course, no process addresses all problems, and there is still much work to be done to cope with situations such as rapidly varying noise, noise that is too highly tuned for a broadband noise reduction system, and noise that reaches or even exceeds the level of the wanted signal. There are existing solutions for each of these cases, but with trade-offs. In particular, the algorithms capable of removing high levels of noise from signals obtained in extreme environments introduce a degree of tonal change that make their output unsuitable for broadcast. The CEDAR SE 1 Speech Enhancer (which was developed specifically for the surveillance community) uses these, but its ability to increase intelligibility is not the same thing as increasing listenability. Indeed, the two are often mutually exclusive.

So what of the future? New sources of noise and new requirements for noise suppression are forever being encountered. In 1989, nobody sat in a noisy office while talking to dozens of people worldwide using a laptop with a whirring fan as the communications device. Today, tens of millions of people do so every day, and sophisticated noise reduction and echo cancellation algorithms are running constantly on



the servers ("in the cloud") that allow them to do so. Similarly, a telephone call made on the London Underground would be unlistenable without similar technologies being employed. Elsewhere, perhaps as a consequence of improved delivery mechanisms, old problems are being readdressed with renewed energy, while new, speculative developments are being vigorously pursued in fields such as blind source separation. Isolating a single voice from the babble in a restaurant or club and simultaneously cleaning the resulting signal has long been deemed desirable and (perhaps) impossible, but current advances are bringing this ever closer.

To combat the ever-increasing noise in our lives, we are seeing more and more noise suppression products, whether for recording, mastering, podcasting, broadcasting, communications or security. Yet the Holy Grail remains what it has always been; a magic box that removes all unwanted sounds without human intervention, does so instantly without introducing artefacts, and leaves the wanted signal sounding totally clean but exactly as the listener originally perceived it. Is this possible?

It would be unwise to say that it's not. Today's processes are vastly more effective than those of 30 years ago, and users nonchalantly expect results that would have seemed unlikely when CEDAR was established. A good example of this is the spectral editor, which we invented in 2002. The ability to remove, move or correct a single sound within a recording without damaging the surrounding audio was a huge breakthrough, yet there's already a new generation of audio engineers for whom it has always existed. Arthur C Clarke once wrote that, "Any sufficiently advanced technology is indistinguishable from magic". He forgot to add that, once it enters common usage it soon becomes accepted, if not mundane.

