VoiceAge®
www.voiceage.com
The World's Premier Supplier of Speech and Audio Codecs





Enabling High-Quality Wideband Speech Across Networks

VoiceAge®
www.voiceage.com
The World's Premier Supplier of Speech and Audio Codecs

Enabling High-Quality Wideband Speech Across Networks

The increasing penetration of end-to-end digital networks, such as second-generation (2G) and third-generation (3G) wireless systems and wireline broadband packet networks, is freeing the telecommunications industry from the archaic restrictions imposed by the over-100-year-old technology foundations of traditional telephony. In addition to their more widely recognized benefits of being noise resistant, resilient to errors, flexible and cost-effective, pervasive digital communication systems are creating an opportunity to introduce a superior standard of wideband speech quality that promises to usher in an era of next-generation hi-fi telephony.

Constraints of traditional telephony

Traditional telephony based on *narrowband* speech, nominally limited to about 200-3400 Hz, imposes a serious limitation on communication quality and intelligibility. This is because speech intelligibility is particularly dependent on consonants, and most of the speech energy that allows people to distinguish consonants occurs above 3000 Hz, while the best conventional phone calls over the public switched telephony network don't pass any energy above 3400 Hz. This is why the person at the far end often has difficulty distinguishing between letters such as "s" and "f."

This bandwidth limitation of traditional telephony dates back to the early days of wireline telephony, around a hundred years ago. At that time, it was difficult to build inexpensive good handset microphones, and in any case higher frequencies were lost as calls passed over long lengths of copper wire. Later, when digital multiplexing was introduced in the 1960s, it was designed to be compatible and readily interoperable with the constraints imposed by the then-current telephone systems that were optimized over many decades to carry narrowband speech. This ensured a smooth evolution and incremental migration to digital systems, where they could add the most value in terms of reducing transmission errors and system costs in the network, without a costly and disruptive wholesale upgrade to a large and complex global telephone system that had become vital to everyday commerce and living. In contrast, the transformation of telecommunications networks today into all-digital packet-based systems removes the traditional constraint on speech bandwidth and will support seamless migration to much better speech quality.

Improving speech quality

The energy of speech signals can extend up to 14 kHz, particularly on unvoiced sounds. However, extending the telephony signal bandwidth from around 3400 Hz up to 7 kHz already results in substantial voice quality improvements over narrowband, where significant energy is lost at both ends of the spectrum. Figure 1 (next page), showing the spectrogram of the speech sentence: *"Everyone looked extremely confused about the news,"* illustrates how much more of a speech signal can be captured between 50 Hz and 7 kHz.

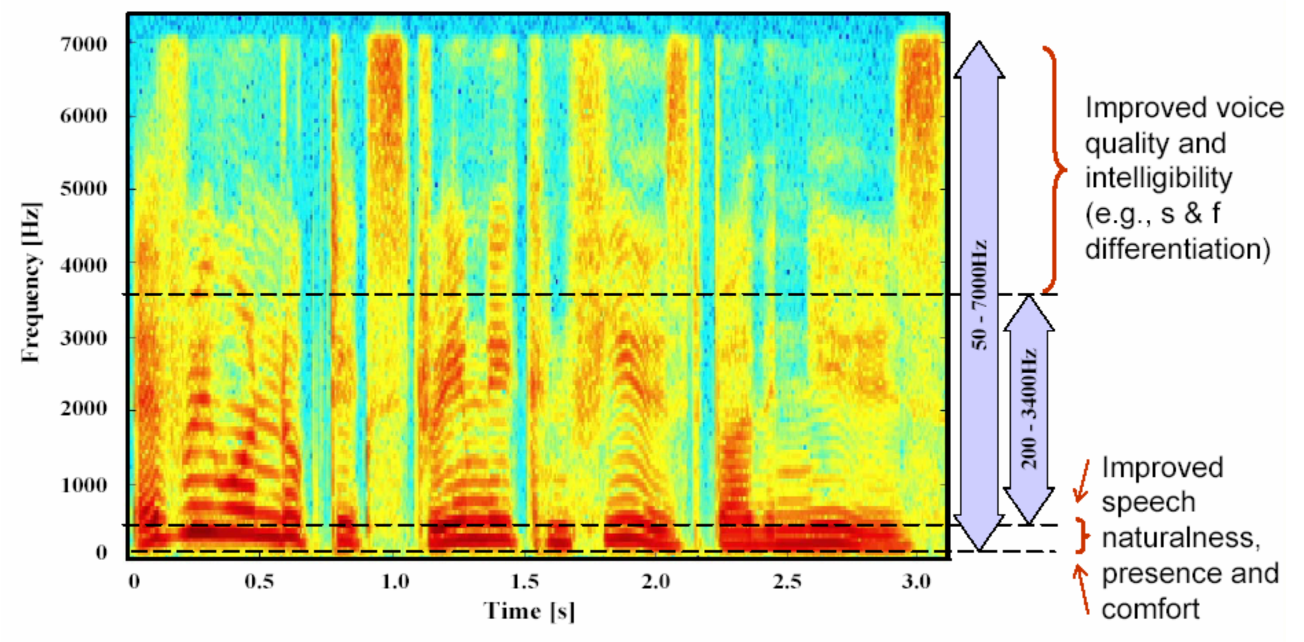

Figure 1: Spectrogram of *"Everyone looked extremely confused about the news."*

Wideband telephony refers to transmitting speech signals with bandwidth in the range of 50 to 7000 Hz, effectively doubling the narrowband speech signal of traditional telephony. Compared to narrowband telephony, the low-frequency enhancement from 50 to 200 Hz contributes to a sense of presence, adding increased naturalness and comfort to conversations. The high-frequency extension from 3400 to 7000 Hz provides better differentiation of certain consonants and therefore improves intelligibility. Wideband speech communications also ease speaker recognition and reduce listener fatigue (probably because they include more speech information so there are fewer gaps for the listener to fill in). Thus, test results clearly demonstrate that using a wider frequency band in speech coding results in major subjective improvements in speech quality.

Applications for wideband speech

The improved voice quality and naturalness of wideband speech are key features emerging in some current value-added telecommunications services such as audio and audiovisual teleconferencing, intra-site enterprise VoIP over high-speed LANs and next-generation services like multimedia messaging. They can also enhance Internet applications such as streaming, broadcasting, chat, multimedia real-time collaboration tools, on-line gaming and virtual reality immersion environments, archiving and distribution of narrative content, and network-based language learning applications. Essentially, any multimedia Internet application that can benefit from more intelligible and natural sounding speech can be enhanced by using wideband speech. As increasing numbers of services and applications use high-quality audio, user expectations for speech communications quality for telephony conversations are rising as well.

Efficient new standard codecs deliver much higher quality wideband speech at bit rates comparable to those required by narrowband speech

Wideband speech is being enabled by high-performance wideband-capable terminal devices that are interconnected over end-to-end digital networks. These terminals incorporate wideband-capable microphones and speakers, and either come equipped with or can download wideband codecs, acoustic echo cancellation, and noise reduction, as needed. Moreover, efficient new wideband coding standards make possible the deployment of wideband speech services at bit rates close to those required by narrowband speech, so pioneering service providers can roll out wideband services without incurring extensive investment costs to upgrade their network infrastructure. Service providers who capitalize on these developments to introduce wideband services will gain a clear differentiating competitive advantage in voice quality over competitors offering narrowband-only speech services.

Coordination of wideband standardization

Interoperability is fundamental to efficient communications both within and between networks, and it becomes even more important given the current goal among international standards organizations of bringing about worldwide access to converged wireless and wireline systems. Therefore work has been done to coordinate standardization activities across international bodies.

Following a rigorous comparative competition with other candidate codecs, the Nokia/Ericsson/VoiceAge codec known as AMR-WB (adaptive multi-rate wideband) was selected for wideband coding in GSM and 3G WCDMA wireless systems by 3GPP/ETSI, and its specifications were subsequently finalized and approved in March 2001. ITU-T recognized the importance of harmonizing their efforts with the 3GPP to eliminate the quality degradation caused by transcoding communication signals traversing different networks with disparate speech codecs. In 2001, following their own competitive process, they selected the AMR-WB 3GPP codec standard as ITU-T recommendation G.722.2 for a wideband coder operating at rates of 13 to 24 kbps. The AMR-WB/G.722.2 standard codec, which actually operates at rates from 6.6 to 23.85 kbps, has been extensively tested in these two standard bodies in selection and characterization phases.

These standard wideband codecs are best in class, are rigorously tested in real-world network conditions and provide interoperability without transcoding

In parallel activities, 3GPP2 selected the variable-rate multi-mode wideband codec (VMR-WB) codec as a standard for wideband speech coding in cdma2000® wireless systems. VMR-WB was designed based on the AMR-WB/G.722.2 codec and was specifically conceived to be interoperable with it in one of its operational modes.

Compatible codec standards eliminate the need for transcoding – even across networks

The adoption of AMR-WB/G.722.2 by 3GPP and ITU-T and of VMR-WB by 3GPP2 is an important milestone in the evolution of speech coding standardization because for the first time compatible codecs based on the same fundamental technology have been adopted to deliver GSM/WCDMA wireless, CDMA wireless, and wireline services. This coordination by the standards-setting bodies will obviate the need for transcoding and ease the implementation of wideband voice applications and services across a wide range of communication systems, networks and platforms.

Transcoding is analogous to the use of a simultaneous translator to mediate conversations between humans who are not able to speak the same language. As in the language analogy, transcoding introduces additional delays, cost and potential errors or loss of meaningful information and therefore should be avoided whenever possible.

Speech signals need to be transcoded to establish communication across networks and/or end-devices that use different codec technologies. The main reason for the wide variety of codecs currently deployed on communications equipment is the diverse requirements imposed by the different types of legacy and emerging networking technologies, both wired and wireless, found around the world. These networks all offer different speech signal throughput bit-rate capacities as well as various tolerances for delay, jitter, noise, errors and real-time performance characteristics.

Transcoding degrades quality, increases latency/delay and raises costs

Therefore, to enable clear and intelligible communication between all publicly accessible network end-points, the networking infrastructure and signal control schemes must be able to provide reliable mechanisms to convert signals encoded at the originating point to one of the representations valid at the terminating point or end-device. Obviously, for duplex human speech conversation this signal transformation needs to occur in both directions, as needed, in real time during a call session. This function of converting speech representations from one form to another is what is meant by transcoding. Typically, this transcoding function is performed at either a gateway or programmable switch that sits at the boundary of disparate networks and mediates the flow of information between them. The specific codecs to be used at each end-point of a conversation are decided during the call session establishment phase of the call setup.

Wideband service deployments across networks

Figure 2 shows a deployment scenario of AMR-WB/G.722.2 and VMR-WB across several wireless and wired networks. In this scenario, it is possible to carry wideband telephone conversations between wireless GSM/WCDMA and cdma2000 networks without the need to transcode between two disparate codecs. Wireline phones, Wi-Fi phones and softphones supporting wideband telephony through a G.722.2 codec will also be able to interoperate with wireless networks equipped with AMR-WB or VMR-WB codecs to deliver high-quality end-to-end next-generation telephony without transcoding.

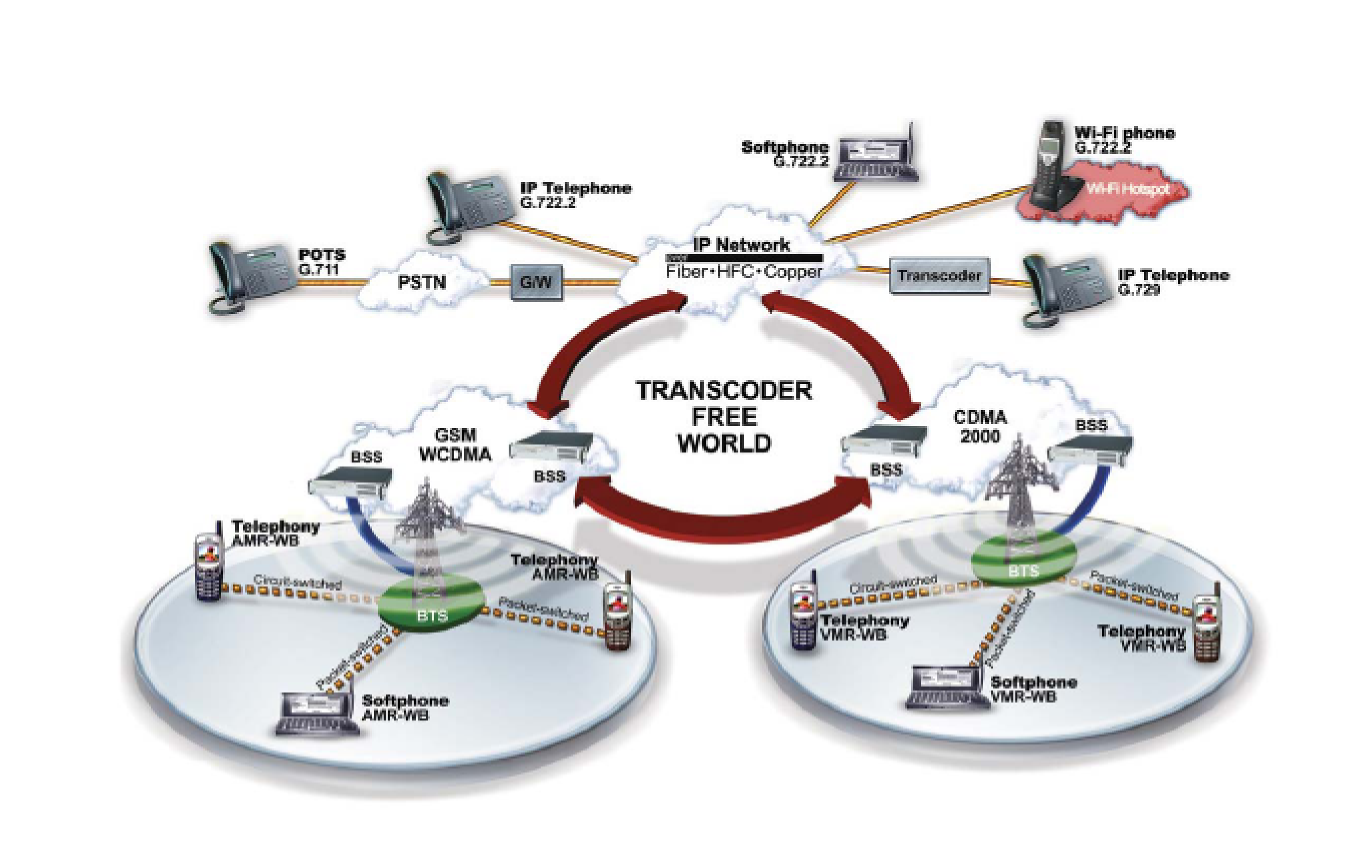

Figure 2: Transcoder-free Interoperability of Wideband Speech Communications Across Disparate Networks

This highly desirable feature helps to preserve service quality because the elimination of transcoding has several benefits. Transcoding not only adds to the cost of providing connectivity; by requiring additional transcoding equipment, it also adds delay, jitter and transcoding impairments to the speech signal. Furthermore, avoiding transcoding helps preserve audio quality because each encoding and decoding stage results in the loss of some signal information and distorts the signal which introduces perceived codec artifacts.

Interoperability for converging networks

The seamless interoperability of the AMR-WB/G.722.2 and VMR-WB wideband codec standards will ease the implementation of wideband applications and services because service providers will be able to offer one set of services across several networks without having to make separate investments in each one. As wideband services become available to greater numbers of users, their improved audio quality will become the norm, raising user expectations and creating even more demand for wideband.

Interoperability of the wideband codec standards will facilitate wideband service adoption

Wideband speech coding is a strong quality differentiator for service providers on all-digital networks. The new standardized wideband speech codecs operate at bit rates comparable to narrowband speech so they don't use up significantly more of the available bandwidth, and forward-thinking standards bodies have made sure that the wideband coding standards for their diverse networks are interoperable, saving investment costs while preserving audio quality. The flow of technology, networks and market evolution are joining forces to create a ripe environment for the widespread deployment of wideband speech communication – are you ready to join the leaders on the vanguard of this movement?