WHITE PAPER

Texas Instruments

Wideband voice challenges

By
Tom Flanagan
Director, Technical Strategy,
DSP Systems

Frank Fruth
Director, Multimedia Application
Software Development, DSP Systems

Dr. Bogdan Kosanovic
Manager/SMTS, Core DSP Components,
DSP Systems

Dunling Li
Senior software engineer, DSP Systems

Abstract

To achieve a richer and more accurate audio experience over the telephone, standard-definition (SD) voice could soon be replaced with high-definition (HD) voice based on Voice over IP (VoIP) technology. However, more issues are involved in HD telephony than simply the implementation of wideband codecs in IP phones or media gateways. A lack of agreement on a standard wideband codec is a serious issue, affecting such factors as processing demands, memory, and IP costs.

Other challenges include achieving realizable performance across existing home wiring; hardening performance to an acceptable level; and the availability of analog wideband handsets, well-designed speakers and microphones, and properly designed handset enclosures for both home and business teleconferencing applications.

Introduction

On a typical phone call it's virtually impossible to feel the excitement in a voice or to fully grasp the subtle meaning in the spoken word. Traditional telecommunications technology has been limited to a sound spectrum much narrower than the full range of the human voice. The limitations of technology, imposed by the constraints of the local loop, sacrifice the richness and subtleties of the voice for basic communications.

Now, the less nuanced standard-definition (SD) telephone voice can be phased out for a richer and fuller high-definition (HD) audio experience based on Voice over IP (VoIP) technology. Because of the advancements in digital signal processors (DSP), the technological encumbrances on voice quality are being removed. For example, conserving network bandwidth was previously a major concern for network operators. Moreover, the processing power of the voice engines in the network was limited; as a result, the computationally intensive tasks of HD telephony could not be contemplated.

Thanks to the widespread deployment of advanced DSP technology in the telecommunications infrastructure, the digital bandwidth available to subscribers has increased considerably. The days when the typical speeds for DSL and cable modems were 700 kilobits per second (kbps) upstream and 1.5 megabits per second (Mbps) downstream have been replaced by typical broadband rates of 2 Mbps and 5 Mbps, respectively. Higher data rates from passive optical networking (PON) and other new infrastructure technologies are also on the horizon.

Broadband service eliminates the need to adhere to the restrictions of the traditional public switched telephone network (PSTN) local loop, which allows end-to-end VoIP and wideband high-definition codecs (see Figure 1 on the following page). Likewise, the processing power of the DSPs in IP phones, residential gateways, media gateways, wireless base stations, and other telecommunications systems has escalated dramatically. Consequently, the wideband codecs that implement HD telephony can and will soon be deployed throughout the telecommunications network.

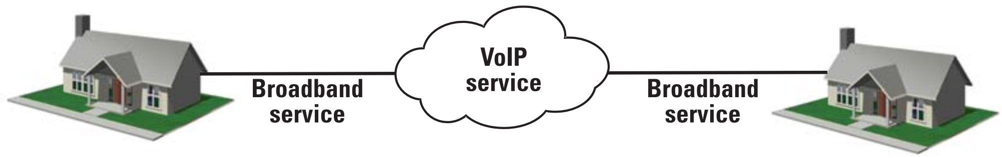

Figure 1: End-to-end VoIP

HD voice

In contrast to SD or carrier-grade voice quality, HD voice is based on advanced DSP technology that can capture and transmit almost CD-quality sound. Today, a typical digital telephone call mimics the lower quality of analog telecommunications technology by sampling an analog signal at a rate of 8,000 samples per second. This is converted into sound in the range of approximately 200 Hz to 3.3 KHz. Wideband HD voice is based on 16,000 samples per second, double the rate of SD voice (see Figure 2). The greater sampling rate capabilities of an HD voice channel account for sound quality that is significantly richer and extended at both ends of the spectrum, ranging from 50 Hz to 7 KHz.

This extended wideband range produces a perceptible difference in the resonance and intelligibility of the voice. For example, many fricative "s" sounds and plosive "p" sounds spread their energy above the 3 KHz upper limit of narrowband SD voice. Studies have shown that the lower end of the wideband range (50 to 200 Hz) contributes to the naturalness of a voice while the higher end (3,400 to 7,000 Hz) provides better differentiation.

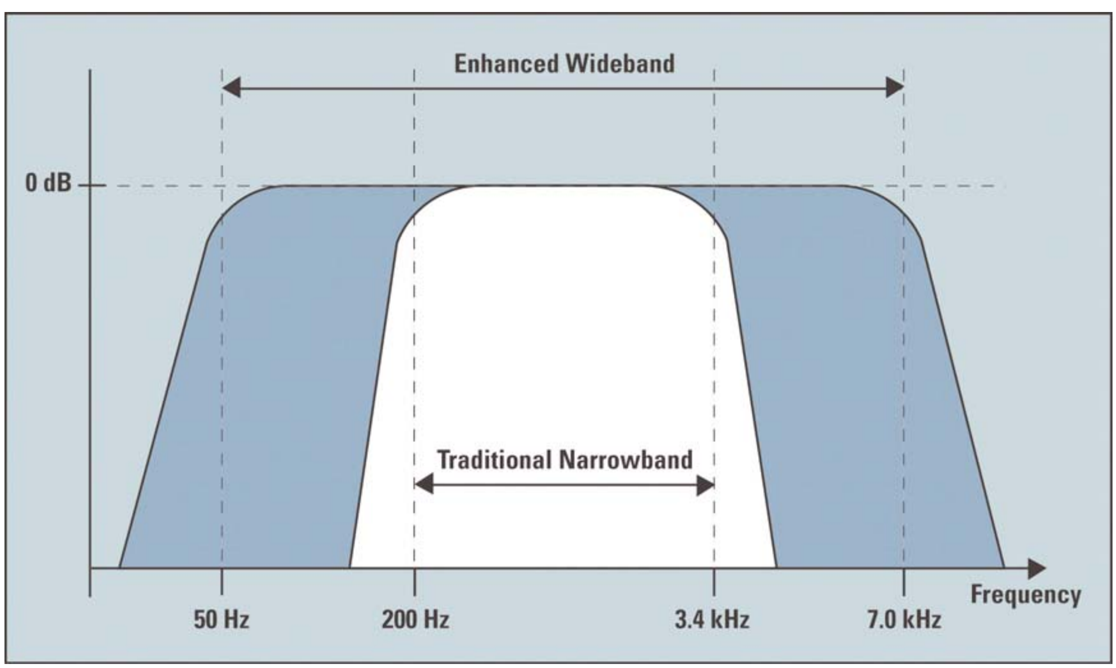

Figure 2: Narrowband vs. wideband

Wideband codecs: The workhorses of HD voice

Central to the deployment of HD voice are codec software modules that transform the voice's analog signals into compressed digital bitstreams for transmission over a network transport, converting the compressed digital bitstreams back into analog signals audible to the human ear.

Surprisingly, the first wideband codec was developed almost two decades ago. The G.722 wideband codec, which supports compressed data rates of 48, 56 and 64 kbps, was standardized by the International Telecommunications Union (ITU) in 1988 for ISDN connections. G.722 was the first codec to employ a 16-K sampling rate. It was developed mainly for teleconferencing applications.

More than a decade later, the second wideband codec appeared. Prompted by a need for lower bit error rates and lower bandwidth usage, the G.722.1 codec was introduced in 1999. G.722.1 halved the bandwidth requirements of its predecessor while maintaining comparable voice quality. It supports data streams of 32 and 24 kbps.

The first wideband codec for both GSM wireless and wireline networks, G.722.2 or AMR-WB, was developed by 3GPPP and the ITU and introduced in 2002. Based on adaptive multi-rate (AMR) algorithms, AMR-WB can eliminate the need for

(Continued on next page)

Today, several wideband codecs have been developed and are in use (see sidebar, "Wideband codecs: The workhorses of HD voice"). However, more issues are involved in HD telephony than the mere implementation of wideband codecs in IP phones or media gateways. That's a start, but it's not the complete solution.

Topological issues

A typical HD voice telephone call could pass through many different levels in the digital telecommunications network (see Figure 1). Beginning with endpoint devices like mobile phones and IP phones, an HD conversation may undergo processing by base stations, media gateways, and switches in the network infrastructure, residential gateways, and/or enterprise systems.

To complete an HD voice call, one or even several wideband codecs may be required at each processing node in the network. With no agreement on a standard wideband codec, several codecs will likely be deployed. Ideally, the endpoints of the conversation negotiate the codec used for the call, but if they can't agree, additional processing will be required to transcode the mismatched codecs. Thus, a processing node would have to convert inputs from one wideband codec to one the destination device supports. Excessive transcoding can introduce delays and degrade voice quality. In addition, the resources needed to support the possibility of several codecs adds to the cost of all systems in the call path – and the processing requirements for some codecs are considerably more computationally intensive than others. Providing adequate processing power for executing any one of several wideband codecs thus adds to the cost of processing resources in every node in the network.

It remains to be seen whether one or more wideband codecs will be widely deployed by the industry. Deciding which wideband codecs to deploy will be affected by solution cost criteria such as the processing demands of each particular codec, the amount of memory needed to store the codec, any intellectual property (IP) costs associated with the module, and other factors.

The last hundred feet

In the telecom infrastructure, "the last mile" refers to the local loop connecting the home or business to the network. Historically, the last mile received a great deal of attention because in many ways the quality of service of the entire infrastructure depended on the quality of the connection in the local loop. Similar logic applies to the deployment of wideband HD voice.

Wideband codecs: The workhorses of HD voice

(Continued from previous page)

transcoding HD voice channels that traverse wireless and wireline networks. By taking advantage of the advanced processing capabilities of DSPs, AMR-WB provides high-quality voice while consuming less bandwidth. For example, an AMR-WB 12.65-kbps channel has better audio performance than a G.722 48-kbps channel. The AMR-WB codec has been extended for mono or stereo applications as high as 19 KHz. This AMR-WB+ codec supports bit rates between 6 – 48 kbps.

In 2004, a new wideband codec was introduced for wireless CDMA2000 networks. The variable-rate multimode wideband (VMR-WB) codec was developed by 3GPP to be interoperable with the AMRWB codec at 12.65, 8.85 and 6.60 kbps. This allows GSM, CDMA2000 and wireline networks to intercommunicate HD voice with no transcoding, which can introduce delays and degrade the voice quality.

In most residential applications, HD voice will be delivered by a media gateway in the network to a residential gateway (RG) device in the home, which will pass the HD signals along to endpoint instruments such as IP phones or other telephony handsets. The last few hundred feet in this channel – from the RG to the telephony handsets – is where the quality of the HD voice connection could be challenged.

In the United States, most homes have been wired with standard two-wire telephone cabling; most handset instruments can only support the limited bandwidth of today's standard two-wire connection. Several options exist for transporting HD voice from an RG to endpoint instruments in the home.

Reuse two-wire cabling

One of the primary challenges to reusing the existing two-wire cabling plant in the home will be the removal of echo. It is well known in the telecom industry that the use of the two-wire transport introduces the primary source of echo because of impedance mismatching at the hybrid that performs the conversion between four- and two-wire interfaces. RG platforms today convert the native four-wire voice signals from the broadband connection for transmission to standard two-wire telephony devices. The RG platforms are responsible for the elimination of echo introduced by the two-wire conversion. Removing echo from a narrowband SD channel is difficult in its own right. The migration to HD voice will require RG devices to remove echo from a wider bandwidth. Technically, the cancellation of echo over a wider bandwidth should be feasible, but the challenge will be a realizable performance across the currently deployed home wiring and the investment in field deployment and trials to harden performance to an acceptable level. An immediate risk with this option is the potential impacts of deployment of HD voice if initial user experience is poor.

In addition to the wideband echo cancellation challenges, there is still the issue of availability of analog wideband handsets. Most analog handsets today cater toward the PSTN market and are limited to the 200 – 3.3 KHz range, which means that they cannot support HD voice. To deploy wideband voice using existing two-wire cabling, a service provider would need to ensure the availability of analog handsets that take advantage of wideband capabilities in the gateway.

Install four-wire cabling

Removing old two-wire residential wiring and installing four-wire cabling would certainly eliminate the echo produced by the hybrid four- to two-wire interface at the RG, but the economics of this alternative are not favorable.

Cordless telephones

Standards such as the W-DECT specification already have been developed for wideband cordless connections between an RG and handsets. In the future, a wideband cordless phone and base station with integral broadband access and VoIP gateway features will be a common platform for the deployment of HD voice (see Figure 3). The cordless air interface bypasses home wiring and reduces the echo cancellation burden that might otherwise be placed on the RG. This solution will need to combat echo, however, from both electrical coupling and acoustic sources. Although processor-intensive, these echo-canceling algorithm technologies are well established from existing IP phones and hands-free speaker phones and can be successfully leveraged. This seems like a promising near-term option for the immediate deployment of HD voice capabilities.

Figure 3: Wideband cordless phone and base station with integrated broadband access and VoIP gateway features

Business deployment issues

Small- and medium-sized businesses as well as global enterprises may encounter issues relating to HD voice that are application-specific. For example, businesses will be attracted to the more life-like communications that are possible with HD teleconferencing services. Unfortunately, participants in a teleconference could have any of several wideband codecs – or no wideband codec at all – in their telephone equipment. When this occurs, the simplest (but not the best) solution would be to “dumb-down” the call to the lowest common denominator of narrowband codec. Alternatively, the teleconferencing system could provide the full wideband HD channel to each participant’s local gateway or handset where the necessary decoding and transcoding would be performed.

The last few inches

No portion of a wideband channel is exempt from scrutiny if the richness of HD voice is to be fully experienced. This includes several design aspects of handsets that affect their acoustic performance. Without well-designed speakers and microphones and a properly designed handset enclosure, the full vibrancy of HD voice will be squandered in the last few inches of a wideband channel.

Most microphones found in telephony handsets perform adequately in the 80 Hz to 8 KHz range, which maps well with the wideband range of HD voice. For very low frequencies below 80 Hz, high-pass filtering is usually necessary in the handset design to reduce low-frequency noise.

Unfortunately, most speakers in traditional telephone instruments do not perform well above 3.4 KHz. Terminating a wideband HD voice channel at such a speaker would eliminate much of the sound in the wideband range from 3.4 KHz to 7 KHz. Once a speaker performs well over the entire wideband range, its enclosure must be carefully designed to optimize acoustic performance and, again, not minimize the gains of the wideband signal.

Reducing the acoustic coupling between speakers and a microphone is particularly critical when a speakerphone feature is engaged for hands-free operations, a frequent mode of operation for high-end handsets capable of HD voice. The spatial relationship of the microphone to the speakers in a hands-free speakerphone will affect the instrument's overall acoustic performance. To minimize acoustic coupling between them, the speakers and microphone should be placed as far apart as possible. In addition, a well designed speaker enclosure can help achieve better isolation from the microphone and improve the frequency response of the speakers.

Delivering HD voice

There are many benefits that accompany HD voice, such as better, more precise comprehension from clearer verbal communication. Of course, there's always the pure enjoyment of hearing the full range of emotions and subtly nuanced meanings of which the human voice is capable. Delivering on these and many other potential benefits will require attention from the beginning of a HD voice channel to its completion. In other words, system-level solutions are needed at every level throughout the infrastructure network and customer premise equipment.

Mailing Address: Texas Instruments, Post Office Box 655303, Dallas, Texas 75265