apt-X Scalable

a technical white paper

by Dr. David Trainor, APTX

www.aptx.com

INTRODUCTION

An area of increasing relevance for developers, manufacturers and users of equipment containing audio coding technologies is the concept of coding scalability. In the most general sense, scalable coding implies the existence of a unified audio coding scheme that is flexible enough to allow different codec performance measures to be traded off to some degree. This trade off can be as straightforward as exchanging audio quality for coded audio bit-rate or, as this paper will discuss, can be a complex and adaptive process whereby many codec performance measures are tuned simultaneously.

User expectations regarding the features, performance and price of devices capable of audio coding and communication are continuing to increase. This is true in many audio markets, from consumer electronics to professional recording and on-stage equipment to broadcast and telecommunications infrastructure. Increased expectations have, in turn, driven technology convergence and integration around a number of relevant device platforms. These platforms range from the smartphone and the netbook in consumer electronics up to powerful PC servers and dedicated rack-mounted processors in audio broadcast and Voice-over-IP telecommunications. As technology convergence continues, device platforms need to become more sophisticated and more flexible, whilst increasing system performance and lowering cost. Real-time system adaptation is a key mechanism in achieving more flexible operation, whilst making most efficient use of the hardware and software resources of the device. Essentially the device will adapt how it processes data based on the characteristics of the data itself, the status of any communications network to which the device is attached and the number of type of processing tasks that the device is required to execute concurrently.

Whilst a high degree of real-time adaptation and optimization is exhibited by modern software applications, operating systems and communications protocols, audio coding techniques have not embraced this trend to the same degree. A number of conventional audio coding algorithms have the ability to adapt some aspects of their operation, for example the compressed bit rate, and may offer some optional functions to alter certain aspects of the behavior of the codec. However, the degree of flexibility offered by conventional audio codecs falls short of enabling very fine-grained adjustment of the many relevant codec performance measures in response to changing system conditions. As a result audio coding algorithms tend to be somewhat prescriptive, relying on the flexibility of other processes such as network communications to efficiently provide high-quality audio under a range of system operating conditions. This approach can work tolerably, but the levels or performance and efficiency are limited by the fixed nature of the audio coding. The fairly prescriptive nature of audio coding techniques is also at least partly responsible for the large number of different codec algorithms in existence and the current need for devices to support multiple codecs.

The remainder of this paper discusses some of these trends and issues more fully and introduces a novel real-time adaptive audio coding technology called **apt-X Scalable**. This technology allows all relevant audio coding performance measures to be traded off with a fine degree of control. The operation of **apt-X Scalable** changes, in real-time, in response to changes in the system environment, such as the short-term network bandwidth, network quality of service, network delay, available processing resources, battery life on mobile devices and the level of audio quality that is deemed acceptable.

USER EXPECTATIONS FROM NEXT-GENERATION AUDIO SYSTEMS

Audio coding pervades a wide variety of consumer electronics devices, studio equipment and the audio broadcast and telecommunications infrastructure. With successive generations of such equipment, user expectations are becoming more demanding and are widening in scope. Devices are migrating from standalone to networked operation and the throughput, reliability and the ease of configuration of the networked audio communications is expected to increase. More intuitive user interfaces and interaction methods are being demanded, including graphical display, touch screen, speech-driven interfaces, haptic feedback, etc. Battery-powered mobile devices are expected to perform an increasingly sophisticated set of audio processing tasks whilst continuing to extend the operational time between charges. The audio quality itself is expected to improve, or at least be maintained at CD quality or better. Through the networking capability, devices are also expected to access an increased variety and quality of audio-related services. Finally, all these improvements need to be delivered at lower system cost!

TECHNICAL TRENDS ENABLING ADAPTIVE AUDIO SYSTEMS

In order for new audio devices to keep pace with all the increased and complex user demands device flexibility and adaptive processing is extremely important. Using the smartphone platform as a relevant example in the consumer electronics market, adaptive processing can be demonstrated in many aspects and at many levels. It is commonplace for a smartphone to contain multiple networking technologies e.g. 3G, WiFi, Bluetooth, etc and boast numerous interaction mechanisms e.g. touch screen, voice commands, keypad, etc. Third-party customizable software applications execute within sophisticated modular and multi-tasking operating systems, which provide control over the partitioning and switching of software tasks and hence promote more agile and adaptive processing. All this adaptive processing in software can be complex and hence the importance of faster and lower-power chipsets brought about by advances in fundamental silicon technology and hardware initiatives such as multi-core processors.

Other more general technical initiatives have recently come into existence to promote flexibility and adaptive processing modern electronic equipment. A good example in the domain of wireless communications and networking is the concept of cognitive radio [1]. The basic premise of cognitive radio is that a wireless device automatically adapts its transmission or reception parameters e.g. modulation scheme, transmission scheme, etc to skillfully avoid interference with and from other wireless users and also to promote the most reliable wireless link. Key aspects of this process are the automated nature of the adaptation and the fact that varying network and user demands trigger adaptive updates. The use of the term “cognitive” becomes obvious given the ability of the communications process to acquire information, apply knowledge and change preferences in a manner similar to a mental process. This is an important concept for the **apt-X Scalable** technology, which applies a similar cognitive approach to the area of adaptive audio coding.

BENEFITS AND APPLICATIONS OF SCALABLE OR “COGNITIVE” AUDIO CODING

Conventional audio codecs such as MP3 [2], AAC [3], FLAC [4] and others are capable audio compression techniques. They are very focused on obtaining low coded bit-rates and maintaining good audio quality. Several codecs have performance parameters that can be adapted dynamically, such as the ability to trade off bit rate and audio quality. Some codecs also have optional functions that can be included or omitted to give a coarse mechanism to trade off coded bit-rate and computational complexity. So what additional benefits can more widely scalable and cognitive coding bring to what already sounds like quite a healthy situation?

Firstly, a significant number of widely-used conventional audio codecs are better-suited for encoding on powerful computing platforms, with few real-time processing constraints, and storing coded output as a file for subsequent playback by a much simpler decoder algorithm, which can be easily realized as a real-time decoder, even on a low-power embedded device. They are much more challenging to deploy in an audio streaming application, for example between a smartphone and a wireless headset, where both encoding and decoding are real-time and being executed on more computationally-limited and power-sensitive devices. For this type of deployment scenario, highly scalable audio coding based on a cognitive coding concept can select and adapt the coding strategy to better suit the capabilities and limitations of the two devices and the communications network in question.

Secondly, the cognitive coding concept ensures that the adaptation of the codec operation is automated within the codec itself, in response to changes in easily-obtained system metrics such as available network capacity, reliability of service statistics, available processing resources, etc. Conventional audio codecs may provide adaptable options, but they don't generally provide any automated selection or control of those options. The system outside the audio codec then has to supply this intelligence and this is a difficult task for the system designers, especially if they are not audio coding experts.

A further benefit of the scalability provided by the cognitive coding concept is that it can adapt and optimize multiple codec performance measures concurrently. Rather than independently adapting codec parameters such as compressed bit rate, computational complexity, coding delay, error resilience of the coded audio format, etc, they are all adapted jointly, thereby reaching a higher-performance operating point. This mechanism also means that the codec can optimally adapt to a set of performance targets when it is supplied with the relative priority of the different performance targets and the fundamental constraints of the device and the network at any given time. This is a level of codec adaptation to the short-term system conditions that simply cannot be matched by more conventional audio codecs and results in practical benefits such as higher coded audio quality, more reliable audio transmission during periods of poor network connectivity, lower coded bit rates for more challenging audio material and better power management for mobile audio devices.

There are a wide variety of audio coding applications and use cases where scalable audio coding based on the cognitive coding idea can bring significant improvements. For the smartphone or netbook platform the technology offers automated means of better managing situations such as low remaining battery life, poor WiFi or Bluetooth connectivity and processing peaks caused by multi-tasking applications. For a professional wireless microphone and monitoring solution the effects of other interfering devices can be minimized and the coded bit-rate can precisely track capacity fluctuations in the wireless channel. For a broadcaster or telecommunications operator, the technology allows higher numbers of audio channels from fixed hardware and network resources and a wider heterogeneous mix of audio traffic.

OVERVIEW OF THE APT-X SCALABLE AUDIO CODING TECHNOLOGY

In order to implement the scalable and cognitive coding approach described in this paper, the **apt-X Scalable** audio coding technology consists of two principal components. The first of these is a “toolbox” of signal-processing functions from which collections of functions can be selected, connected and configured in many different ways to realize many different types of audio codec. The codec functions can be reconfigured or completely changed in real-time and, where necessary, the reconfiguration can occur smoothly over a period of time to avoid abrupt and unpleasant audible effects on the audio. Each function within the toolbox consists not just of code to execute on a processor or a piece of hardware, but also of metadata that quantitatively characterizes how that function influences the important performance measures of an audio codec e.g. coded audio quality, codec delay, computational complexity, robustness to errors, etc.

The second principal component of the **apt-X Scalable** technology is an optimization engine, whose responsibility is to select and configure functions from the toolbox in real-time, thereby creating updated audio coding solutions. This engine can be configured with a set of priorities that indicate the relative importance of the different performance measures of the audio codec at that particular time. It can also be configured with a set of system constraints that must be respected by the codec composed from functions in the toolbox at that particular time e.g. maximum average coded bit rate, maximum number of CPU cycles, etc. Based on these inputs, the optimization engine then carries out a mathematical optimization process that efficiently examines the metadata of the various functions in the toolbox and determines an optimum combination of functions that best meets the system constraints and codec performance prioritizations.

A diagram of the **apt-X Scalable** encoder is shown in Figure 1 below. Note that the optimization engine is providing the directives for the toolbox to create the required audio codec to transform the audio data from its original to its coded representation.

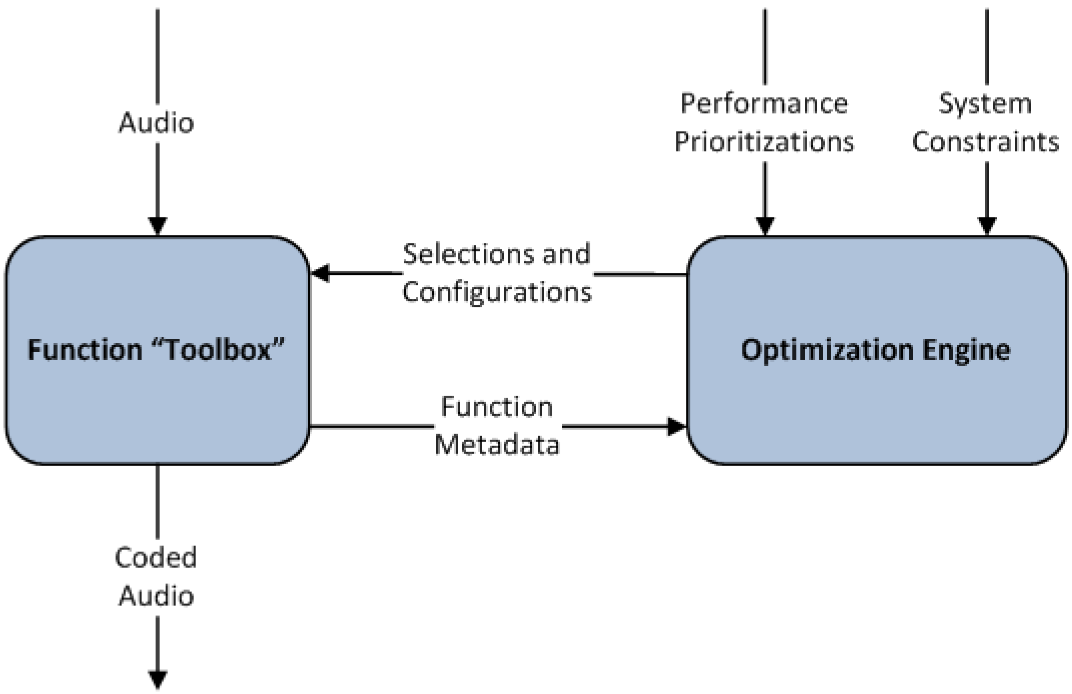

Figure 1 : apt-X Scalable Encoder

A diagram of the **apt-X Scalable** encoder is shown in Figure 2 below. At this high level of abstraction the decoder looks very similar to the encoder, with the obvious exception that coded audio is accepted and decoded audio is produced.

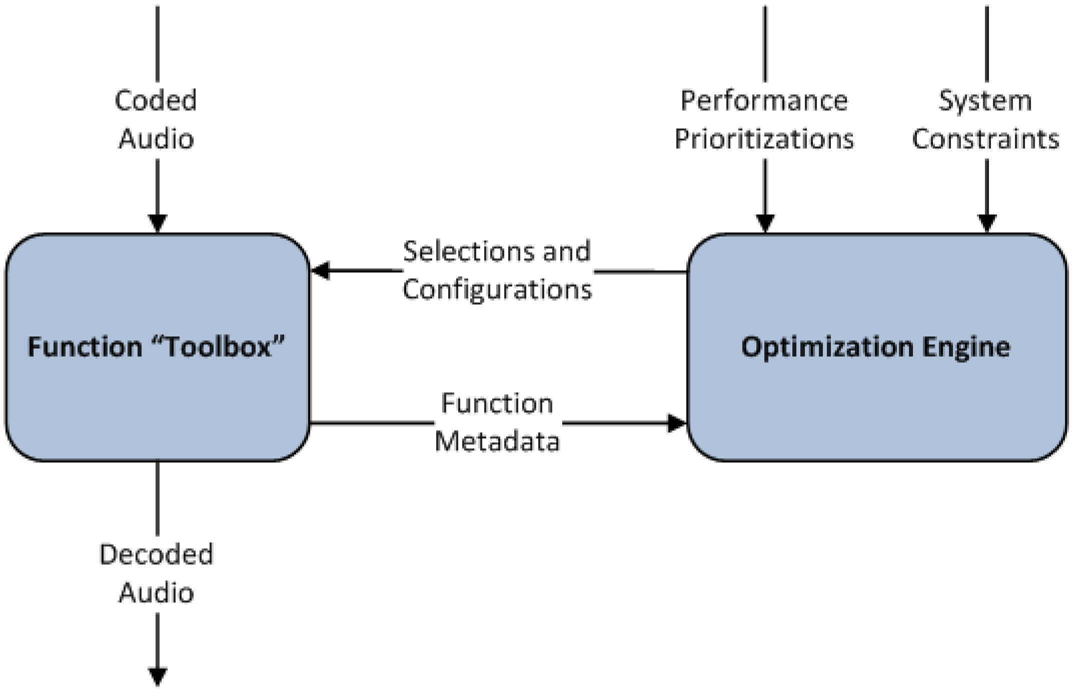

Figure 2 : apt-X Scalable Decoder

An important area of consideration is synchronization between the **apt-X Scalable** encoder and decoder. Clearly if the optimization process were to work completely independently for the encoder and decoder it would be possible for the encoder and decoder to be constructed from incompatible functions from their respective toolboxes. This would result in the audio being encoded in a format that the corresponding decoder could not decode! **apt-X Scalable** handles this issue by a communication of optimization decisions and system constraints from the encoder to the decoder. This information can be communicated from encoder to decoder in two ways, either as control information in the coded audio data itself or in the form of additional system constraints sent via any other logical communications channel existing between the encoding device and the decoding device. In the case of bidirectional communications systems, it is also possible to communicate decoder optimization decisions and system constraints back to the encoder. By giving the decoder visibility of optimization decisions made by the peer encoder, it is possible to further constrain the optimization process of the decoder to ensure that the optimum set of functions selected from the decoder function toolbox must also be able to successfully decode the encoded format currently produced by the peer encoder. If bidirectional communication exists, it is also possible to further constrain the optimization process of the encoder to ensure that the optimum set of functions selected from the encoder function toolbox must create an encoded format that can be decoded by the currently-selected decoder structure. Therefore, depending on the network infrastructure, **apt-X Scalable** allows encoder optimization to have priority over decoder optimization or vice versa. This can be very beneficial when audio is being sent between two devices with very different sets of capabilities.

APT-X SCALABLE DEVELOPMENT AND ROLLOUT

The **apt-X Scalable** audio coding technology is the result of nearly 2 years research effort on the part of APTX. Elements of the technology have previously been incorporated in the **apt-X Lossless** codec, first presented in early 2009 and now available for commercial license. The concepts of **apt-X Scalable** were initially publicized at the 126th AES Convention and the technology has now reached the point of practical demonstrability on real-world systems. The initial release of **apt-X Scalable** as a packaged licensable technology for a range of platforms will occur during Q4 2010.

CONCLUSIONS

Next-generation audio processing systems in consumer, professional and broadcast markets will require audio coding solutions as flexible and adaptive as the hardware and software used in the devices themselves and the networking technologies allowing devices to communicate. Significant performance gains can be realized via more flexible and reactive approaches to audio coding, but conventional audio codecs are still too prescriptive in their operation to fully embrace this concept. The novel approach adopted by **apt-X Scalable** creates a form of cognitive audio coding, whereby codec operation is automatically tuned and optimized based on system-level performance demands. This elevates the performance and flexibility of the audio device or equipment as a whole, which is a key requisite to allow next-generation devices to continue to satisfy increasing, and sometimes conflicting, user demands.

REFERENCES

[1] **The Wireless Innovation Forum** - http://www.wirelessinnovation.org/

[2] **MP3** – ISO/IEC Standards 11172-3 and 13818-3

[3] **AAC** – ISO/IEC Standards 13818-7 and 14496-3

[4] **FLAC** - http://flac.sourceforge.net

MORE INFORMATION

More information about **apt-X Scalable** and other **apt-X®** series audio codecs, including comprehensive technical data and specimen commercial licensing documentation, is available from the appropriate contacts given or send a request e-mail to *licensing@aptx.com*

NORTH AMERICA, REPUBLIC OF KOREA

Mike Canevaro, VP Business Development, APTX

mcanevaro@aptx.com

+1 714 731-1507

EUROPE, JAPAN

Stephen Wray, VP Licensing, APTX

swray@aptx.com

+44 28 9067-7200

REST OF WORLD

Jonny McClintock, Sales President, APTX

jmcclintock@aptx.com

+44 28 9067-7200

http://www.aptx.com

APTX – *the apt-X® licensing company*

APT Licensing Limited

Whiterock Business Park

729 Springfield Road

Belfast, BT12 7FP

Northern Ireland, UK