apt-X™

Digital Audio Data Compression

Technical Overview

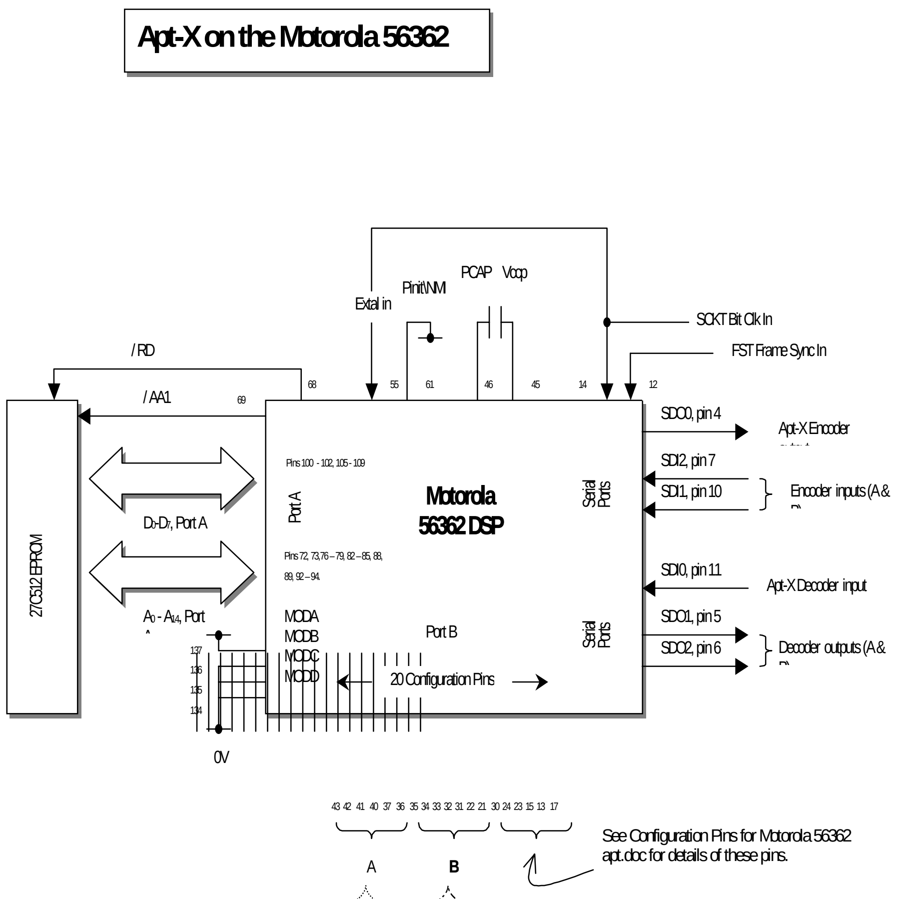
Index

Abstract

Audio Processing Technology developed the proprietary **apt-X™** algorithm, a real-time digital audio data reduction system that compresses 16-bit PCM audio samples by a factor of 4:1 with no perceptible audible degradation. It is suitable for a wide range of audio applications including digital audio storage and digital audio network distribution.

apt-X™ is a low complexity compression algorithm offering unsurpassed audio quality and many unique operating features capable of running on a cost-effective single DSP chip. The system is operable at any sampling rate up to 48 kHz, is inherently robust to random bit errors, extremely tolerant to tandem coding and has an exceptionally short coding delay. For example a dual channel encoder/decoder algorithm may be currently implemented on a single Motorola DSP56362, figure 1

Figure 1 Building blocks of the Motorola 56362 Implementation



Set Mode

EncdeX		Function
01 (pin 43,34)	02 (pin 42,33)	
0	0	Disabled
0	1	Encode
1	0	Decode
1	1	Reserved

Set Resolution

AlgResX		Function
01 (pin 41, 32))	02 (pin 40,31)	
0	0	Old 16 Bit
0	1	New 16 Bit
1	0	20 Bit
1	1	24 bit

Detected Resolution

DetResX		Function
01(pin 37,22)	02(pin 36,21)	
0	0	Old 16 Bit
0	1	New 16 Bit
1	0	20 Bit
1	1	24 bit

Detect Lock

DetLockX (pin 35,30)	Function
0	No Lock
1	Lock

In addition to the Motorola part , OEM's have the ability to utilise the apt-X algorithm, as it has been made available for Licensing for seem less integration into an infinite range of applications .

Under licensing terms, the algorithm can be made available as:

- SOFT apt-X™ - a Windows and Linux based library
- DSP object code for popular devices such as the Motorola 563XX , Texas Instruments and Analog Devices parts
- VHDL / Verilog

Key Features of the apt-X™ system

- Real-time and up to 16 times real time 4:1:4 data compression and expansion
- Low hardware complexity
- Single chip implementation
- Mono/stereo audio encoder/decoder
- Up to 22.5kHz stereo Duplex with a single device.
- Sampling rate up to 48 kHz
- Extremely low coding delay
- Highly robust to bit errors
- Tolerant to tandem coding
- Linear phase response
- Integrated AUTOSYNC™ encoder/decoder synchronisation
- Embedded auxiliary data transmission up to 12kbit/s
- Low power consumption requirements
- No external RAM or ROM requirements, device dependant

Introduction

Consumer expectation of high quality audio has greatly increased over recent years. The challenge has been largely met with, for example, the Compact Disc, NICAM coding for domestic television, FM and most recently DAB radio broadcasts. Digital audio has now become synonymous with high fidelity audio but the storage and transmission of CD-quality digital audio can be expensive. In the telecommunications industries considerable research has been conducted for many years to develop systems capable of transmitting music and speech at lower bit-rates than PCM. The apt-X™ coding algorithm is an extension of this work and was designed to allow CD quality audio to be exchanged in real-time at lower bit rates over ISDN, Kilostream and Megastream and other digital terrestrial and satellite telecommunications networks, and now with the availability of SOFT apt-X™ audio conversion and transport can be run entirely within the PC domain.

apt-X™

The **apt-X™** system is based on an implementation of subband Adaptive Differential Pulse Code Modulation, (Subband ADPCM). The CD and other high end digital audio systems on the other hand use 16-bit linear pulse code modulation, (linear PCM). As it's name implies ADPCM is a technique which re-codes the **difference** between two digital audio samples, using quantisation step-sizes that **adapt** to the energy of the input audio signal. In this way ADPCM can provide a similar audio quality to linear PCM but at a much reduced bit-rate. The **apt-X™** system aims to transparently code 16-bit PCM audio with a fixed compression ratio of 4:1.

Encoding (compression)

Figure 2 illustrates the encoding stages of the **apt-X™** algorithm. The analogue audio signal must first be converted to a 16 bits per sample PCM digital signal. This signal is the input to the **apt-X™** encoding algorithm where successive time blocks of four PCM samples are first filtered into four equal bandwidth frequency subbands. These signals, still in 16 bit format, are then simultaneously processed in four separate signal chains each incorporating a backward linear prediction loop that provides an estimation of the input signal. The prediction, based on the history of previous PCM samples, is subtracted from the input to yield a difference signal which is commonly termed the error signal. It is this 16 bit error signal which is then re-quantised using a backward adaptive Laplacian quantiser whose step sizes adapt to the magnitude of the error signal, again based on an analysis of preceding samples. The bit rate resolution of each of the four subband quantisers is different and much lower than the PCM bit rate. This reflects the non-linear frequency sensitivity characteristic of the human ear, the only reference made in the algorithm to psychoacoustic masking. The four code-words from each subband are then multiplexed into a single 16 bit apt-X™ code word suitable for storage or transmission. This 16 bit apt-X™ code word therefore represents the content of the original 4×16 bit linear PCM samples (64 bits) and is therefore ¼ of the PCM data rate, a rate reduction of 4:1.

Decoding (reconstruction)

Figure 3 illustrates the decoding stages of the **apt-X™** algorithm. The input to the decoder is a 16-bit apt-X™ compressed word. This word is de-multiplexed into the four low bit resolution code-words which are fed into four separate subband inverse quantiser chains. The output from each quantiser is a 16 bit error signal which then must be added to a predicted signal generated from the history of previously reconstructed samples. Finally the four reconstructed 16 bit, bandwidth limited samples are inverse filtered and leave the decoder in a serial data stream output as full bandwidth, 16-bit PCM samples. This has the same data rate as the original PCM signal at the input to the encoder. If required these samples can be converted to an analogue audio signal using an external D/A converter.

Figure 2 apt-X™ Coding process

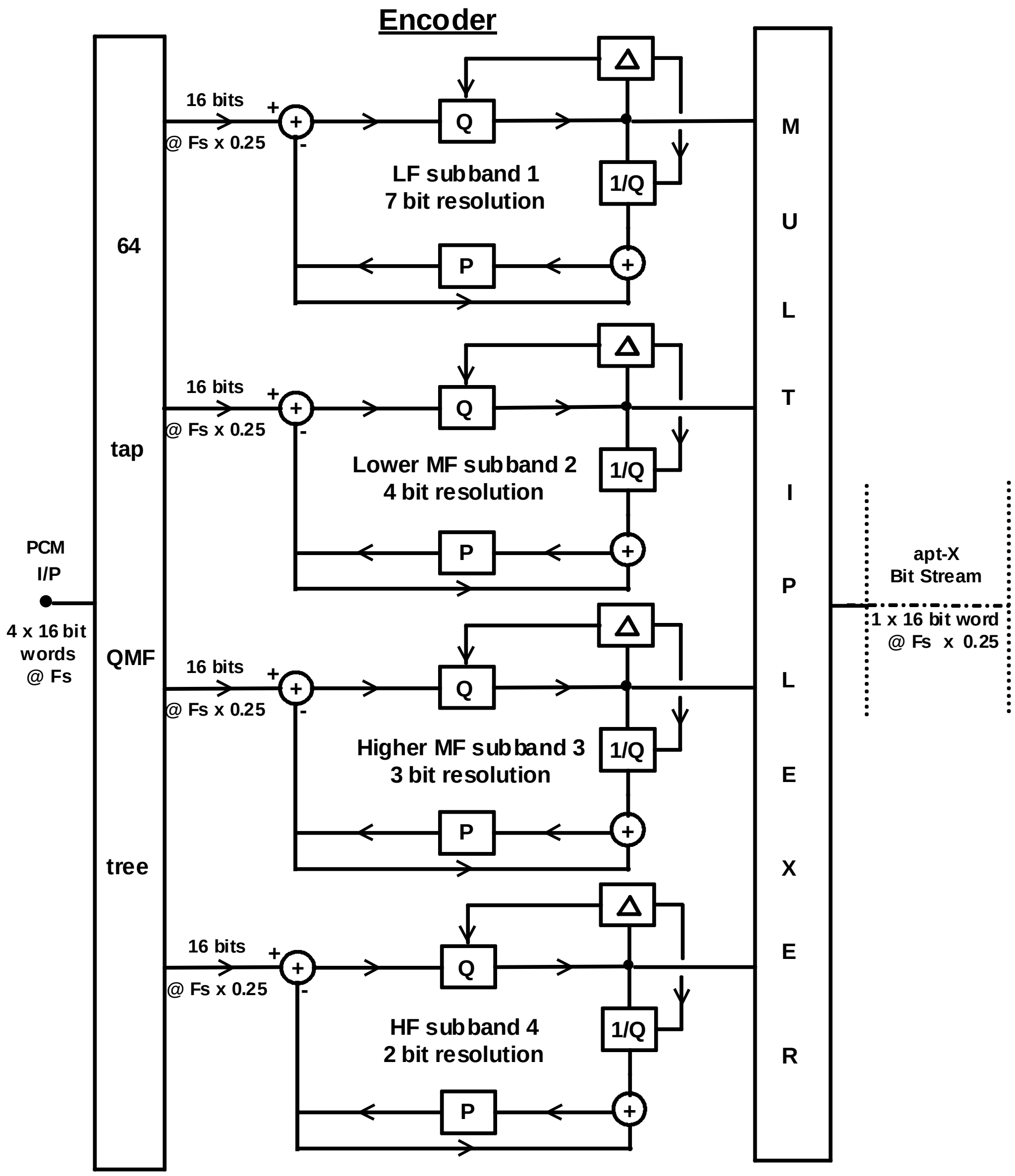
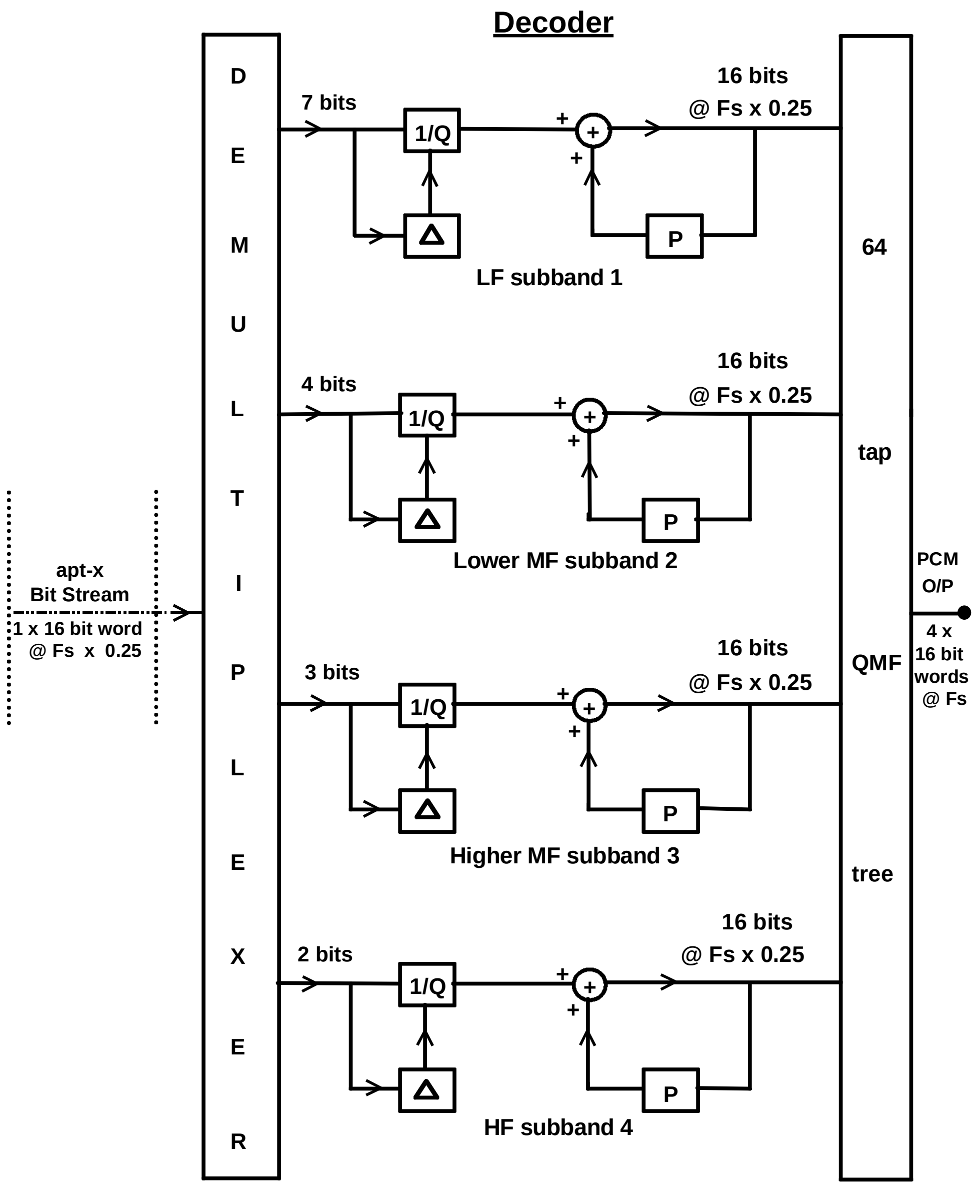


Figure 3 apt-X™ Decoding process



Key elements of apt-X™ coding system

There are three key elements of the algorithm which collectively achieve the bit rate reduction:

- **Subband encoding**
- **Linear prediction**
- **Adaptive quantisation.**

Subband encoding

Subband coding within **apt-X™** splits the audio signal into four frequency bands which are processed separately. By altering the short-term coding resolution in each band according to the energy of the signal in the band, spectral redundancies in the full-band audio signal are exploited. Spectral redundancy can be defined as the degree to which a stationary audio spectrum differs from a full amplitude white noise spectrum.

Musical instruments, such as the tin whistle, generate audio signals that are mostly tonal in appearance. Such signals contain a high level of redundant information in that most of the energy of the signal is concentrated in a small range of frequencies, principally at the fundamental note and higher frequency harmonics. On the other hand the signal from a cymbal is more like a noise signal in appearance and contains very little redundancy in that the energy of the signal is evenly distributed across a wide range of frequencies. Another phenomenon is that the fundamental frequency of audible sound, even a complex musical sound falls into the spectrum around and below 4kHz.

The spectral redundancy in an audio signal is exploited in **apt-X™** by allowing the subbands which contain high energy signals, particularly those around the fundamental reference, to be coded at a higher resolution than the subbands that contain low energy, higher frequency signals.

By utilising a quadrature mirror filter to generate the four subbands a further subjective advantage is gained in that any further quantisation noise generated within each subband is constrained within that frequency band. The subbands are therefore effectively de-coupled from each other.

Linear prediction

Linear prediction aims to further remove spectral redundancies from the subband signals. Redundancy is reduced by comparing the level of the incoming subband sample with a prediction that the algorithm had made for that sample. Whether the prediction is greater or smaller than the actual sample it nevertheless creates a difference, or error signal which can then be re-quantised. If the prediction has been very accurate then the difference signal is much smaller than the original subband sample, allowing the quantisation resolution to be reduced. The prediction is based on a short-term analysis of reconstructed previous samples, otherwise known as backward prediction. This technique ensures that the encoder and decoder can generate identical predictions without the need to communicate any side information. The same amount of redundancy removed at the encoder can then be added back at the decoder.

The success of linear predictive coding is highly dependent on the short-term periodicity of most musical signals, and in the case of **apt-X™** has been optimally matched to the non-linear auditory response of the human ear. Pure sinusoidal signals which are critically perceived are highly periodic and hence highly predictable. The difference signal generated by **apt-X™** is thus small and the overall sine wave is coded accurately. Noisy random signals, which are not critically perceived, exhibit little or no short-term periodicity and prediction is very difficult. Hence the difference signal for a noise-like signal input is by definition much larger and is therefore coded less accurately.

Adaptive quantisation

Adaptive quantisation exploits the relatively slowly time-varying energy fluctuations which audio exhibits by continually adjusting the size of the quantiser step to match the signal level. The step-size is determined from an analysis of previous quantiser step-sizes. If the level of a subband signal remains fairly constant over time then the quantiser step size also remains stable thus increasing the efficiency of the quantiser. Conversely if the signal level changes dramatically then the step-size adjuster tries to accommodate that sudden change, with a subsequent decrease in quantiser efficiency.

This feature exploits the phenomenon of temporal masking, whereby sharp impulse-like audio signals are known to mask out, for a substantial period of time, other audio signals which appear before (pre-echo) and after the impulse signal. The quantisation efficiency for signals immediately following impulsive sounds can thus be reduced without audible effect.

Operational Features

Hardware Implementation

Real-time operation has been made possible due to advances made in the speed of digital signal processors. The original target platform for the algorithm was that of the Motorola 563XX family , which can accommodate simultaneously both encode and decode operations in stereo, up to 22.5 kHz at a sample rate up to 48kHz.In addition requirements for memory are low to reduce external dependency(fig 4) as are the operational requirements (Fig 5).

Fig 4 Memory requirements

	P	X	Y	L
Standard apt-X™	1745	49	548	200
Enhanced apt-X™	2192	2033	19	132
Per channel pair		256	160	135

Figure 5 MIPS requirements

Algorithm	Time taken (μs)	Cycles/lo-rate sample/channel pair	MIPS/channel pair @ 48kHz
Standard apt-X™ encode	32	3146	37.75
Enh apt-X™ 16-bit encode	34.6	3401	40.81
Enh apt-X™ 20-bit encode	35.4	3478	41.74
Enh apt-X™ 24-bit encode	36.2	3559	42.71
Standard apt-X™ decode	30.4	2988	35.86
Enh apt-X™ 16-bit decode	29.8	2930	35.16
Enh apt-X™ 20-bit decode	29.8	2930	35.16
Enh apt-X™ 24-bit decode	29.8	2930	35.16

Coding Delay

The time delay for a complete encode/decode processing cycle is entirely due to the use of quadrature mirror filters, and is inversely proportional to the sampling frequency.

Sampling Rate (kHz)	**Standard Coding Delay (m.sec)**	**Enhanced coding delay (m.sec)**
48	2.5	1.9
44.1	2.9	2.04
32	3.8	2.8
16	7.6	5.6

Bit Error Response

The overall framework of the **apt-X™** algorithm ensures a high degree of resilience to random bit errors. No protection of the transmitted audio data is therefore included in the algorithm. The bit error response is well matched to the auditory response of the human ear, traditionally critical signals being relatively immune to errors. A bit error rate of 1:10,000 is normally inaudible for general programme material. This inherent resilience can be attributed to the adaptive differential coding operating within the de-coupled subbands. Distortions introduced by bit errors are constrained within a subband. In addition the backward adaptive prediction and quantisation tend to reduce the significance of random errors by spreading their effect over the trailing window of samples used for the adaptation. Furthermore, the magnitude of the effect of a bit error is proportional to the magnitude of the differential signal being decoded at that instant. Thus if the transmitted differential signal is small, which will be the case for a low-level signal or a resonant highly predictable signal, any bit error will have a very little effect on either the predictor or quantiser and hence should be inaudible.

Tandem coding and post-processing

Tandem coding is generally defined as multiple encode/decode cycles and may involve both analogue and digital signals. In a tandem coding situation any coding algorithm such as other compression algorithms and even including 16-bit linear PCM coding will cause degradation of the audio signal. In general all real-time compression systems will degrade the audio at a faster rate than linear PCM coding. Any form of coding, compression or linear will unavoidably inject quantisation noise into the original audio signal, but usually at levels and frequencies that cause the audio signal to mask that noise. The amount of inaudible noise that can be injected into the signal is obviously fixed implying that, once the threshold limit has been reached, any further attempt at coding, especially compression will cause the noise to eventually break through the threshold and become audible.

The **apt-X™ 100** algorithm is unique in that predictable elements of the audio signal are attenuated before quantisation. Therefore the level of quantisation noise around the critical harmonic frequencies is low, leading to a high degree of tolerance to tandem coding. Whilst the level of tolerance is dependent on the input signal, normal programme material shows little degradation even after ten digital encode/decode passes through the algorithm. Another feature of the **apt-X™** algorithm which improves the tolerance to both tandem coding and post-processing is the use of a fixed output bandwidth, irrespective of the complexity of the input signal. This ensures that the high frequency content is always coded.

Additional features

- **Automatic synchronisation**
- **Automatic de-multiplexing**

- **Auxiliary data transmission**

Automatic synchronisation

Automatic synchronisation (**AutoSync™**) is a ON/OFF facility which enables the **apt-X™** compressed audio data stream to be decoded without prior knowledge of the 16-bit compressed word boundaries. With **autoSync** ON this enables the compressed audio data to be handled at both the transmitter and receiver using only bit timing, no word clock being required. No bandwidth overheads are incurred through the use of any of the **autoSync** facilities. Synchronisation is obtained by periodically inserting a unique 10 bit sync word into the compressed audio data stream, each of the bits spread over the first ten 16 bit apt-X™ code words in a data frame of 128 samples. This sequence is searched for at the decoder and, once three successive sync words have been found, establishes the compressed word boundaries. Synchronism is normally lost only under very adverse channel error conditions (BER of 1:100) at which point the output is muted until synchronisation is restored. If a loss of lock is flagged synchronisation will be re-established through **autoSync** in approximately 50 msec.

Automatic de-multiplexing

In addition to synchronising a single channel, **autoSync** also enables two, four, or eight channels to be automatically demultiplexed, again with no bandwidth reduction. **AutoSync** combines powerfully with the 16-bit word format of the compressed audio data stream to allow, for example, the direct replacement of a 1024 kbit/s linear PCM stereo audio link with four **apt-X™** compressed stereo audio channels, with no change in the timing circuitry.

Auxiliary data transmission

Additional auxiliary user data may be transmitted between the encoder and decoder by embedding it within the compressed audio data stream. This data has a maximum transmission rate of 1/16 of the total compressed audio bit rate, allowing for example 16 kbit/s within a 256 kbit/s, 15 kHz stereo audio link. The encoder can remotely enable the auxiliary data mode at the decoder, allowing auxiliary data transmission to be dynamically switched on and off as demanded.

Enhanced **apt-X™**

Audio Processing Technology has released further developments of their world renowned 16 bit apt-X™ 4:1 data compression algorithm. On a new platform provided by the Motorola 56362 DSP chip, the original 16-bit algorithm has been enhanced to provide higher audio quality and to operate at 20 and 24-bit audio resolution. The improvements in part have been made possible with the re-working of the algorithm to make use of the 24-bit architecture available within this device.

The 56362 chip is a very low power CMOS device capable of operation at 100 million instructions per second (MIPS). The implementation of the apt-X™-algorithm on this platform therefore enables it to deliver four discrete audio channels with bandwidths up to 22kHz at 48kHz sampling frequency.

In **SIMPLEX** mode these can be configured as: -

- Four discrete mono encode or decode channels
  or
- Two stereo encode or two stereo decode pairs
  or

In full **DUPLEX** mode as: -

- Two encode channels plus two decode channels

Enabling any two channels to be configured with a different bit resolution further increases flexibility of operation.

The utilisation of the Enhanced algorithms offers several key advantages for system integrators and users alike:

1. **Delay**
   Reduction in encode/decode delay cycle – this can be illustrated comparatively as follows:

Sample rate (kHz)	Standard apt-X™ (ms)	Enhanced apt-X™ (ms)
48	2.5	1.9
44.1	2.9	2.04
32	3.8	2.8

2. **Dynamic Range**
   The dynamic range now stretches to 116dB and 120 dB with 20 and 24 bit enhanced apt-X™ respectively

3. **Total Harmonic Distortion**
   The techniques adopted by the Enhanced algorithms have shown an improvement in the THD+N figures to – 69 dB and – 72 dB for 20 and 24 bit respectively.

4. **Autosync™**
   One of the unique features of the standard algorithm was that of the ability to include a synchronisation word embedded with no loss of audio quality over the compressed bit stream. Enhanced apt-X™ takes this ability much further, in that the Autosync algorithm has been revised to embed a synchronisation pattern in every codeword. This gives the algorithm the ability to return sync much faster, and indeed is permanently present within the audio, thereby presenting a much more robust system which can pass configuration data from encoder to decoder.

5. **Aux data**

Auxiliary data within Standard apt-X was present on one channel of audio per stereo pair, at a rate of 12kBits/s. The advent of the Enhanced versions extends this in that the data bits may now be spread over both channels to reduce the ‘bit stealing’ of the compressed codewords or alternately to provide two discrete auxiliary data paths.

Data rate requirements

The increase in audio resolution leads to a requirement for an increase in bandwidth, which can be illustrated in the table below, showing the flavours of the algorithm and the corresponding DS0 slots required to achieve the audio performance:

Audio Bandwidth	Standard 16 bit (DS0 Slots)	Enhanced 16 bit (DS0 Slots)	Enhanced 20 bit (DS0 slots)	Enhanced 24 bit (DS0 Slots)
7.5 kHz Mono	1	1	N/A	N/A
7.5 kHz stereo	2	2	N/A	N/A
15 kHz mono	2	2	3	3
15 kHz stereo	4	4	5	6
20 kHz Mono	3	3	4	5
20 kHz Stereo	6	6	8	9

N.B 1 DS0 Slot equals 56/64 kbit/s

Contacts and Addresses

For further information Licensing and technicalities of the apt-X™ algorithm, please contact the APT sales staff at one of our three offices or visit our website, **http://www.aptx.com**

Corporate Headquarters:

Audio Processing Technology Ltd.
Edgewater Road
Belfast BT3 9JQ
Northern Ireland
Tel; +44 (0)28 9037 1110
Fax; +44 (0)28 9037 1137
e-mail : pmcgrath@aptx.com

North America Office:

APT Inc.
6255 Sunset Boulevard, Suite 1025
Los Angeles, CA 90028
Tel; +1 323 463 2963
Fax; +1 323 463 8878
Email : sfactor@aptx.com

Japan Office:

APT KK
2-14-4 Shinomome, Koutou-Ku
Tokyo 135
Tel; +81 (0) 3 3520 1020
Fax; +81 (0) 3 3520 1022