Techno World

Audio Technology for the Network Age

ATRAC3 High-Quality Audio Encoding Technology

- High-quality audio
- Affinity with ATRAC
- Supports miniaturization
- Applicable to a wide range of applications

Last year Sony released the Memory Stick Walkman, which uses the Memory Stick as its recording media. The Memory Stick Walkman adopts the ATRAC3 (Adaptive Transform Acoustic Coding 3) high-quality audio encoding technology and achieves high-quality audio reproduction equivalent to that of MD players.

ATRAC3 is a further development of the ATRAC technology adopted by the MiniDisc (MD), which was released in 1992. ATRAC3 can be effective not only in packaged media applications such as MD and Memory Stick, but also in networked environments such as music distribution services.

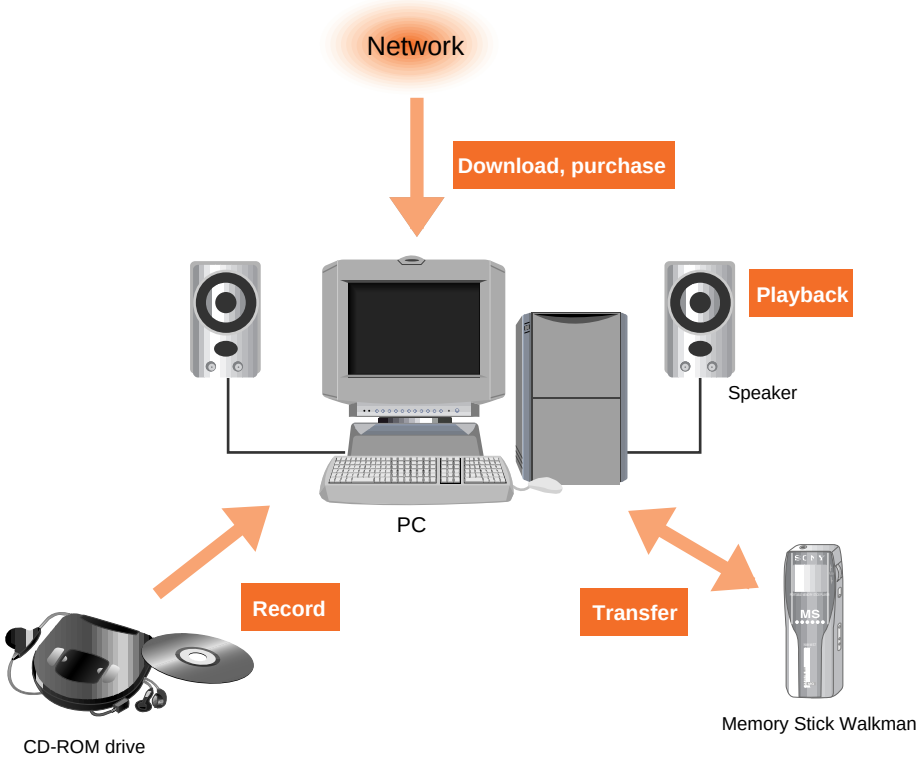
This article presents the ATRAC3 high-quality audio encoding technology and LSIs that include that technology.

ATRAC3 Features

The following are the four main features of the ATRAC3 technology.

- Achieves twice the compression ratio of ATRAC while maintaining the same high audio quality.
- ATRAC3 is similar to ATRAC, allowing the creation of systems that include closely coupled MD functionality.
- Both an encoder and a decoder can be implemented on a single LSI at low cost. Thus ATRAC3 can be used in a wide range of applications, including portable equipment.
- ATRAC3 allows switching to an optimal bit rate depending on the required audio quality, the media used, and the network conditions.

In this article, we describe the basic technologies used to implement.



■ Figure 1 Audio Technology for the Network Age

ATRAC3 Overview

Figures 2 and 3 present an overview of ATRAC3. ATRAC3 adopts a technique called transform coding, in which an audio signal is converted to a frequency domain signal and the data is compressed using the characteristics of human hearing.

The audio signal input to the encoder is split into four bands by a band splitting filter. Gain control is then applied to each band and the results are converted to frequency domain signals using MDCT (Modified Discrete Cosine Transform). The frequency domain signals are split into their tonal and non-tonal components and bitstream is generated using different procedures for those components.

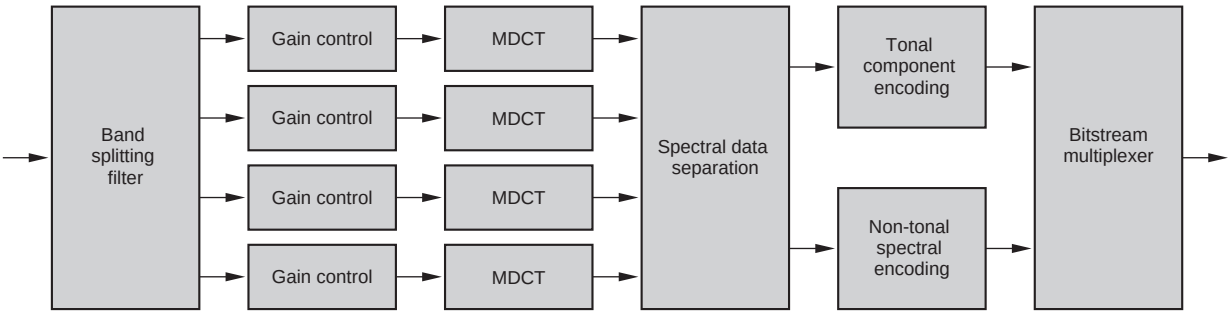
The decoder first recovers the spectral components from the bitstream and then applies an inverse transformation for each band using IMDCT (Inverse MDCT). It then applies gain adjustment and then generates the audio signal using a band synthesis filter.

Band Splitting Filter and MDCT

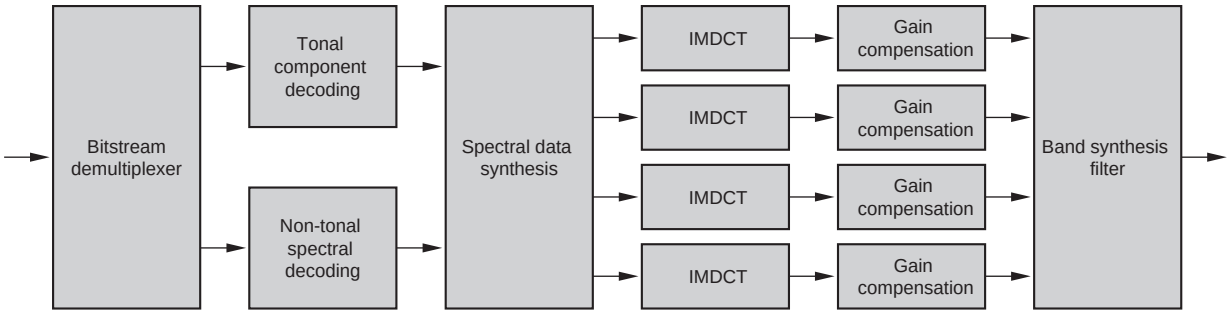
To provide similarity to the ATRAC format, the same QMF (Quadrature Mirror Filter) is adopted as the basic filter module for the band splitting filter. Figure 4 shows the structure of the band splitting filters used in ATRAC and ATRAC3. In ATRAC, after the audio signal is split into two bands by a QMF, the low band is split again into two bands by another QMF. Thus the audio signal is split into three bands. In ATRAC3, on the other hand, after the audio signal is split into two bands by a QMF, each of those bands is again split into another two bands. Thus the audio signal is split into four equal bands.

To achieve twice the frequency resolution of ATRAC, the MDCT conversion blocks are made longer and the number of output frequency domain signals for each band is increased to 256, so that there are a total of 1024 for all four bands.

One advantage of using band splitting filters in conjunction with MDCT is that a smaller memory working area is required. The memory area, which is a large part of the hardware scale of the LSI, can be made significantly smaller, thus allowing ATRAC3 to be implemented by an inexpensive LSI.



■ **Figure 2 ATRAC3 Encoder**



■ **Figure 3 ATRAC3 Decoder**

Gain Adjustment

Gain adjustment technology is being developed as a basic technology to achieve high audio quality in ATRAC3. Figure 5 shows an overview of ATRAC3 gain adjustment. In the transform coding technique, in which the audio signal is converted to spectral components (subbands) and then coded, one important issue is how well audio quality can be assured in sections where an attack, which is a place where the level of the audio signal changes rapidly, occurs. In particular, the problem is to prevent a type of noise called "pre-echo," which occurs immediately before an attack.

Gain adjustment technology was developed to prevent pre-echo. During encoding, this technology detects attacks and, before spectral conversion, increases the amplitude of the section of the signal preceding the attack. Then during decoding, this technology returns the signal to its previous level. Since gain adjustment technology reduces the quantization noise level in the sections preceding attacks during decoding, pre-echo is suppressed. This technology also has the advantage that it can be implemented without greatly increasing the scale of the hardware.

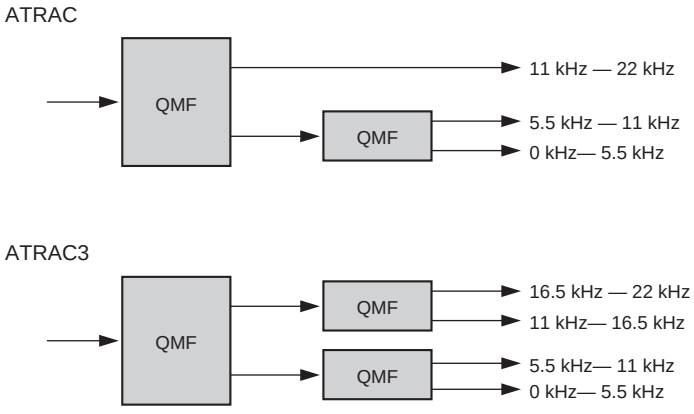
Tonal Component Separation Technology

Human hearing is not very sensitive to audio signals that have flat frequency domain signals, and a high signal-to-noise ratio is not required for such signals. However, human hearing is extremely sensitive to audio signals that include tonal components in their frequency domain signals. In this latter case, a high signal-to-noise ratio is required.

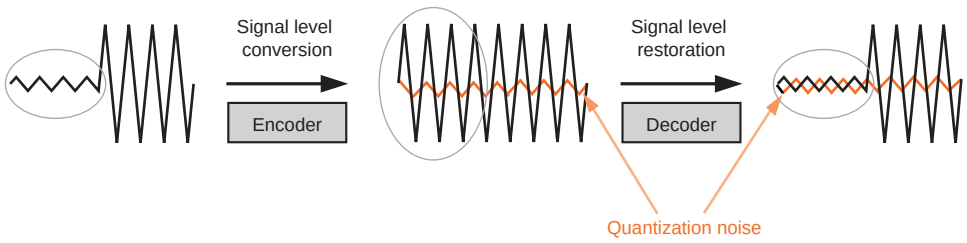
In ATRAC3, the tonal components are separated, quantized separately from other spectral components, and then coded using entropy coding. This separation technology is particularly effective for coding audio signals in which the energy is concentrated in a relatively small number of frequency components, and thus assures an extremely high signal-to-noise ratio and achieves excellent audio quality.

Joint Stereo Technology

ATRAC3 adopts (a technology called) joint stereo technology in which the correlations between the left and right channels in the stereo signal are used to compress the signal. Whether or not joint stereo technology is used is determined adaptively for each band, and it allows extremely efficient coding. Joint stereo coding technology is effective for assuring audio quality at low bit rates.



■ **Figure 4 Band Splitting Filter**



■ **Figure 5 Gain Adjustment**

CXD1859GA ATRAC3 LSI

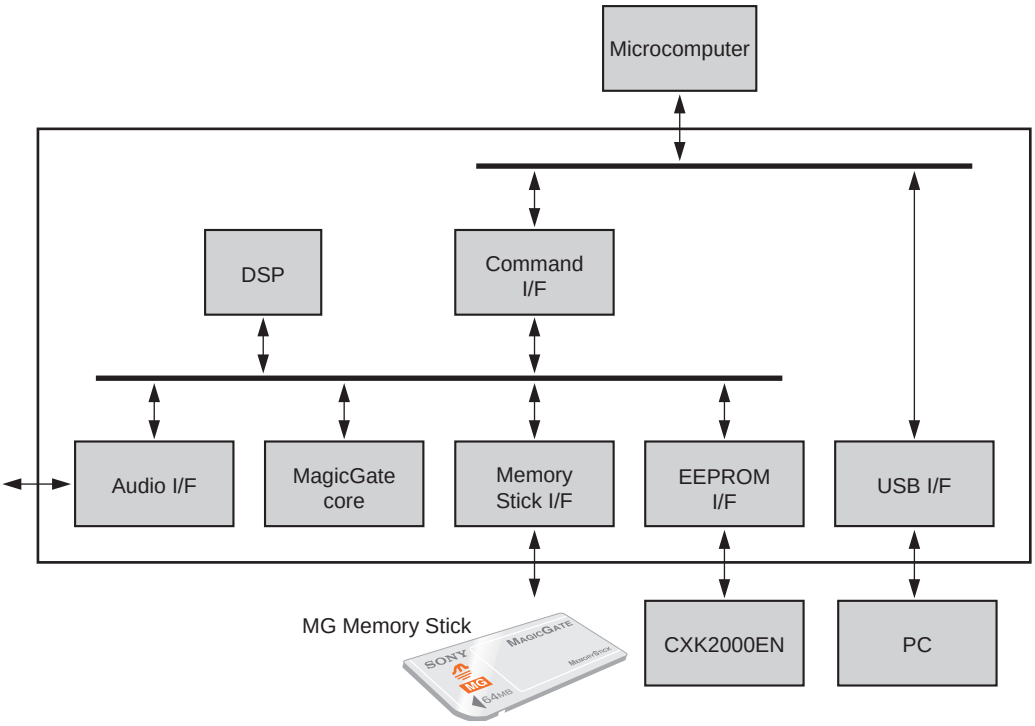
The CXD1859GA is a low-power LSI that implements both encoder and decoder functions for the ATRAC3 technology presented above. It performs ATRAC3 encoding and decoding using a built-in digital signal processor. At the same time as providing, in a single chip, all of the hardware modules required in Memory Stick audio equipment, including a Memory Stick interface, MagicGate, and a USB interface, it also provides both software and an API for the Memory Stick interface and MagicGate functions. Memory Stick audio equipment can be implemented easily by combining the CXD1859GA with a host microcomputer. Since the software, including the audio signal-processing components, are downloadable, flexible system structures are possible. (See figure 6.)

Main Features of the CXD1859GA

- ATRAC3 encoder and decoder
- Audio interface
- Memory Stick interface
- USB interface
- MagicGate functions
- CPU interface
- Package: 120-pin LFLGA

Future Developments

Sony has implemented an electronic music distribution system that protects copyrights by combining the ATRAC3 audio encoding technology, which features high-quality audio and high compression rates, with the MagicGate and OpenMG encryption technologies and electronic watermarking.
With networks becoming even more convenient and accessible due to the increasing availability of networks to the home and the growth in mobile computing environments, the adoption of ATRAC3 as a core digital audio technology is currently being considered, not only for Memory Stick applications, but for a wide range of other systems as well.



■ **Figure 6 CXD1859GA Block Diagram**