Audio Engineering Society

Convention Paper

Presented at the 118th Convention
2005 May 28–31 Barcelona, Spain

An Introduction to the KOZ Scalable Audio Compression Technology

Kevin M. Short, Ricardo A. Garcia, Michelle Daniels, John Curley, and Mike Glover

Chaoticom Technologies*, 3 Riverside Dr., Andover, MA 01810
{kevin, rago, michelle, john, mglover}@chaoticom.com

ABSTRACT

An overview of the high quality, full-bandwidth, low-bitrate, scalable KOZ audio codec technology is presented. This new compression method grew out of developments in the control of chaotic systems that allow for the creation of broad spectral components with only a few bits of information. These elements are combined with a high-resolution analysis of the audio signal that allows the signal to be decomposed into tonal, noise-like, and transient objects. Psychoacoustic principles have been adapted to prioritize and quantize these objects, and the reconstructed signal is built up in layers from the prioritized objects, resulting in a scalable format. Metadata and built-in Digital Rights Management are present in the digital filestream. The decoder is a very low-complexity algorithm that is implemented on a wide variety of portable devices such as cell phones in a software-only solution running on fixed-point processors without DSP support.

1. INTRODUCTION

Many formats for perceptually compressed audio have become popular over the past fifteen years. In most of these codecs, coefficients of a time-frequency transform are directly quantized based on a variety of psychoacoustic models. Most transform-based codecs are not easily scalable, so bitrate choices are usually limited. In the past few years, delivery over mobile networks has created a demand for a digital audio format that can offer high quality at very low bitrates and can scale to higher bitrates when bandwidth is available. From the start, development of the KOZ codec has been motivated by the stringent requirements of mobile music applications while providing a broad range of bitrates for a wide variety of applications.

A mobile music application places extensive demands on a codec, including the need for Digital Rights Management (DRM) to protect files and avoid unauthorized distribution. Portable devices for playing and storing the audio files, such as cell phones, music players, and video game devices, rarely have hardware DSP support and are often very limited in storage space, processor power, and battery life. Additionally, delivery of music over mobile networks is greatly limited by available bandwidth. Because of these

* Chaoticom Technologies is a division of Groove Mobile (formerly Chaoticom)

restrictions, the KOZ codec uses an asymmetric approach where the complexity is focused in the encoder while the decoder is very lightweight. The low-complexity decoder uses a limited number of processor cycles and reduces battery drain. In addition, its implementation is software-only, facilitating the portability of the codec to different devices. More importantly, the KOZ codec has been developed to provide high quality compression of both music and speech to meet the demands of the public at 32kbps and below, making it possible to distribute music over GPRS mobile phone networks and store hundreds of files on portable devices. The fine-grained scalability of the format allows distribution of higher bitrate and higher quality files as desired. Also, robust DRM technology that allows streaming playback over networks is integrated into the KOZ file format with very little bitrate overheard.

KOZ encoding goes beyond the bounds of traditional transform-based encoding, and uses a combination of high resolution front-end analysis and audio elements derived from chaos theory to achieve a high quality representation of the audio signal. Because these audio elements can have broad spectral characteristics, the KOZ codec does not band-limit the audio signal at any point and instead aims to accurately reproduce the entire spectrum at all bitrates. By combining these elements in layers, the KOZ codec is naturally scalable, as increases in bitrate allow the introduction of more layers to create a more accurate signal. Because the KOZ format can successfully meet the needs of demanding applications like mobile music, it is expected that the combined benefits of high quality encoding, scalability to a broad range of bitrates, preservation of the full audio signal bandwidth, and integrated DRM will continue to make it beneficial as a next-generation codec.

2. OVERALL STRUCTURE

As shown in Figure 1, the KOZ encoder uses a multi-stage process to compress raw audio data. The first stage involves pre-processing of the audio signal. Sections of the audio file are extracted in data windows of approximately 46 milliseconds with an overlap of approximately 12 milliseconds. Once the data is extracted, a lossless transformation converts the independent channels into a "Unified Domain" where all information describing the signal is encapsulated in a single multi-dimensional channel. This transformation can be used for stereo, surround sound and multi-channel audio streams. Next, a method for high-resolution frequency analysis is utilized, transient analysis is performed, and a customized psychoacoustic model is applied. Each of these processes occurs in the Unified Domain.

Following the initial signal processing, psychoacoustic principles and detailed knowledge of the signal content are used to decompose the Unified Domain signal into discrete objects such as steady tones, noise-like

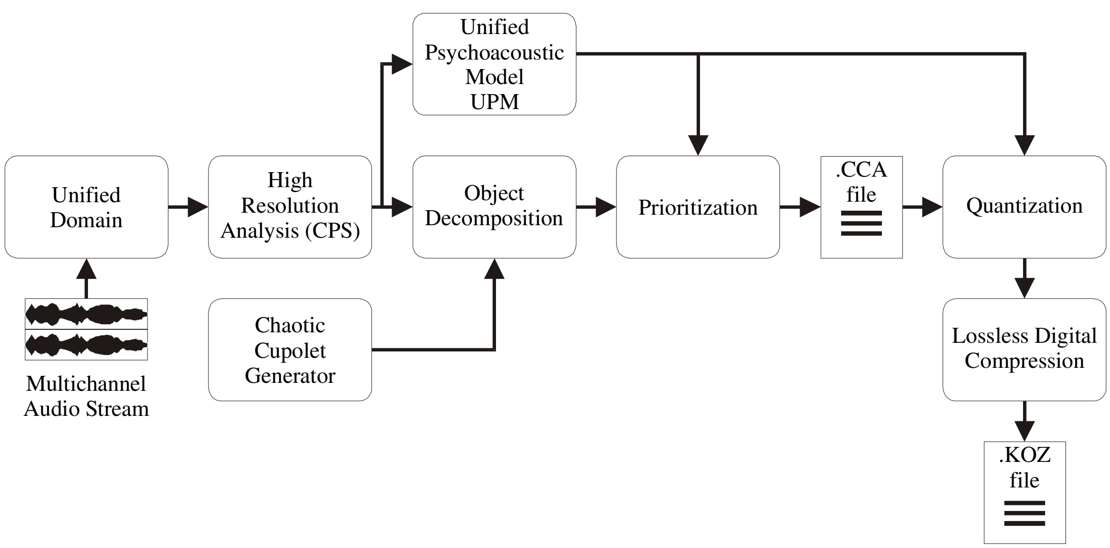

Figure 1 Overall structure of KOZ encoding

elements, and transient events. Similar viewpoints for signal decomposition have been proposed by Serra and Smith [1], and McAulay and Quatieri [2], among others. Once created, these objects can be modeled using chaotic waveforms to produce an accurate approximation of the original audio. At this stage in the process, the floating-point parameters describing the objects are written to an intermediate .CCA file. This file, with its structured groupings of objects, contains all of the information needed to reconstruct the modeled signal during the decoding process.

The next stage of the encoding process uses the psychoacoustic importance of the individual objects in order to detect and eliminate redundant information while quantizing the floating-point values from the .CCA file. Customized variations on traditional lossless coding techniques are then applied to the quantized objects, and the result is a coded audio stream in the KOZ format.

3. DETAILED INFORMATION

3.1. Background on Chaos

The ability to control chaotic systems in order to create complexity from a few bits of information is the key feature that motivated the development of the KOZ codec. It was discovered that under a fairly straightforward control procedure, chaotic systems could be induced to settle onto periodic orbits that would otherwise be unstable and difficult to use for practical applications. The control process requires only on the order of 16 bits of information, but the resulting periodic waveforms, called *cupolets*, can be as simple as a sine wave or so complex that they have more than 200 harmonics in their spectrum. Furthermore, cupolets naturally carry the structures seen in speech and music, including sequences of harmonics and formant structures. An analysis and comparison of the cupolet spectra to spectra of various musical instruments was conducted by M. K. Johnson [3]. The class of waveforms that can be produced by chaotic control is sufficiently diverse that the spectral signatures appearing in audio signals are well represented, and cupolets can be combined in such a way that they can be used to model these spectra.

One of the most useful chaotic systems for waveform generation and audio signal modeling is the double-scroll oscillator [4] described by the set of coupled nonlinear differential equations:

$$
\begin{aligned}
C_1 \frac{dV_{C1}}{dt} &= G(V_{C2} - V_{C1}) - g(V_{C1}) \\
C_2 \frac{dV_{C2}}{dt} &= G(V_{C1} - V_{C2}) + i_L \\
L \frac{di_L}{dt} &= -V_{C2}
\end{aligned}
\tag{1}
$$

where

$$
g(V) = \begin{cases} m_1 V, & -B_p \le V \le B_p \\ m_0(V + B_p) - m_1 B_p, & V \le -B_p \\ m_0(V - B_p) + m_1 B_p, & V \ge B_p \end{cases}
\tag{2}
$$

Here *g(V)* represents a nonlinear negative resistance component, and C_1, C_2, L, G, m_0, m_1, and B_p are constant parameters. These equations can be used to build an analog circuit, or the equations can be simulated on a computer. If a circuit is built, the variables V_{C1} and V_{C2} are voltages, and i_L is a current [5]. In the equations, the variables are real and continuous, while the output of a computer simulation produces a sampled waveform.

When a chaotic circuit is built or a chaotic system of equations is simulated, the system settles down onto a complicated structure called an attractor, and it will (generally) settle down onto the same attractor no matter what initial conditions are used. For a 3-variable system such as the one above, these attractors are usually ribbon-like structures that stretch and fold upon themselves and remain confined to a box. The actual state of the system is determined by the instantaneous value of the system variables, V_{C1}, V_{C2}, and i_L in this case. For a chaotic system, the values of these variables will never repeat, so a chaotic system in its natural state is aperiodic. Also, two states that are initially very close will diverge over time, a property that is called *sensitive dependence on initial conditions.*

While the chaotic attractors are aperiodic structures, the attractors usually have an infinite number of unstable periodic orbits embedded within them. A simple control scheme can be used to stabilize these orbits by simply perturbing the state of the system in certain fixed locations by a tiny amount. Using the equations above as an example, the attractor that results from a numerical simulation using the parameters C_1 = 1/9, C_2 = 1, L = 1/7, G = 0.7, m_0 = -0.5, m_1 = -0.8, and B_p = 1 has two lobes, as shown in Figure 2. To set up the controls, a control half-plane is passed through the center of each lobe and passes outward to intersect the outer part of each lobe. Since the attractor is ribbon-like, the intersection of the attractor with the control plane is essentially a line. When the state of the system passes through the control line, the control scheme allows perturbations of order 10^{-3} to be applied. The controls are defined by a bit string, generally of 16 bits, where a zero bit means that no perturbation is applied at an intersection with the control line, while a 1 bit means to apply a perturbation. These controls are applied repeatedly at intersections with the control line, and a single bit at a time is read from the control string to determine if a perturbation is to be applied (looping back to the beginning of the control string when the last bit has been read). After an exhaustive search of all possible control strings (of some fixed length), it was discovered that the majority of control strings cause the chaotic system to stabilize onto a periodic orbit, and these periodic orbits are in one-to-one correspondence with the control string used, *independent of the initial state of the system.* Tens of thousands of cupolets can be produced from the same circuit or simulated equations simply by varying a control string of a few bits.

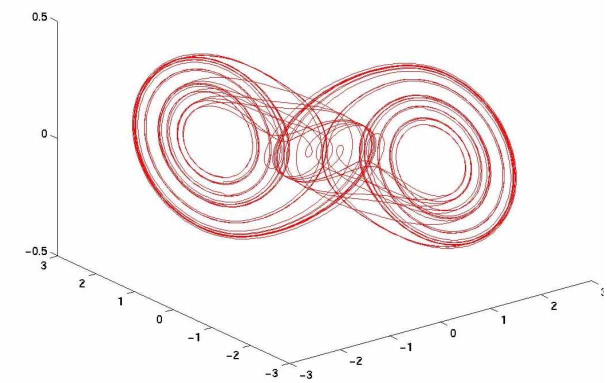

Figure 2 Double Scroll Oscillator

Once a cupolet is stabilized, it forms a closed loop that tracks around the attractor and is defined by the three state variables. The conversion to a one dimensional waveform can be done in a circuit implementation by taking the output of one of the voltages or current measurements; in the computer simulation case a digitized waveform is produced by sampling one of the state variables. In practice, the term *cupolet* is often used to represent both the periodic orbit in three dimensions and the one-dimensional waveforms that it produces.

One easy way to characterize the spectra of the cupolets is to take one of the associated one-dimensional waveforms and look at the magnitude of the FFT of a single period of oscillation. The single-period spectral representation reveals the number of harmonics as well as the envelope or formant structure of the cupolet. Figure 3 illustrates the huge variation in spectral signatures found among cupolets and shows the spectra for a simple cupolet, a cupolet of intermediate complexity, a complex cupolet, and a highly complex cupolet. In each case, the cupolets have been produced by the same chaotic system with the same parameters but different control strings. The corresponding signals in the time domain can be seen in Figure 4. It is clear from these figures that the simplest cupolet in Figure 3a and Figure 4a is essentially sinusoidal, so it is quite possible to use cupolets to produce signals by using single cupolets to model each bin in the transform domain. However, in Figure 3(b-d), it is apparent that much richer structure is available. In Figure 3b, multiple peaks at low frequencies are visible along with a significant second formant above the 20th harmonic. In Figure 3c, the spectral power is diminished in the lower frequencies, only to rise in multiple peaks in the high harmonics. Figure 3d represents an extremely complex tone with hundreds of harmonics and a complicated spectral envelope. The corresponding time domain signals in Figure 4 reveal the expected diversity of behavior.

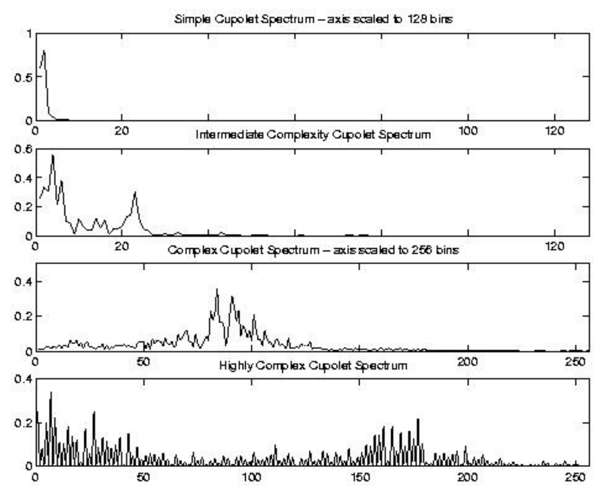

Figure 3 Spectral variation among cupolets

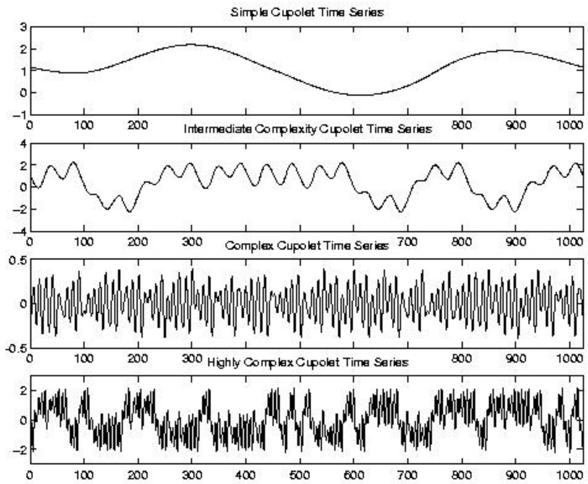

Figure 4 Time domain variation among cupolets

The developments mentioned in the control of chaotic systems are valuable in their own right and may have an impact beyond audio applications. Some of the mathematical properties that make chaotic systems and cupolets useful include the following:

- Many chaotic systems are produced through systems of a few coupled nonlinear differential or difference equations.
- The nonlinear equations are simple; only their time domain behavior is complex.
- The stabilization properties that produce the cupolets are extremely robust.
- Both continuous and discrete waveforms can be produced.
- Geometric transformations can be used to alter the cupolet waveforms in order to optimize them for a given application.

Other application areas include [6]:

- Cupolets can be used as a waveform synthesizer.
- Cupolets can be used for image and video compression.
- Cupolets can be used in chaotic methods of remote key generation, watermarking and secure communication.

3.2. Pre-processing

In analyzing a section of music there is often a great deal of redundancy between the different channels, and taking advantage of this property can reduce the information required in compressed music. Other codecs have used techniques such as joint stereo and parametric stereo to exploit this redundancy [7]. The KOZ codec uses a transformation to a Unified Domain, which retains information about the magnitudes, frequencies, internal phases, and spatial locations of the signal components. The Unified Domain transformation is a completely invertible technique that converts multiple channels of music into a representation that involves a single magnitude component multiplied by an element of the complex Special Unitary group SU(N), for N-channels [8]. The SU(N) group has many representations, but the most useful are representations as complex matrices. In the case of stereo input, N=2, and the representation amounts to a single magnitude component multiplied by a 2 x 2 complex matrix. This representation has also been used in other fields such as optics [9].
The transformation of a multi-channel audio stream is represented as:

$$\begin{aligned} &T : C^N \leftrightarrow mag * SU(N) \equiv U^N \\ &\left[audio_{ch0} \quad audio_{ch1} \quad \cdots \quad audio_{chN-1} \right] \leftrightarrow U^N \end{aligned} \qquad (3)$$

where the magnitude is a function of frequency, N channels are input, and U represents the Unified Domain.

For a conventional two channel audio stream (such as Left/Right) the representation becomes:

$$\left[L \quad R \right] \leftrightarrow U^2 \qquad (4)$$

This representation is a one to one mapping and is lossless. Any manipulations done in one domain have an equivalent counterpart in the other domain. Unified Domain manipulations are similar to traditional frequency domain operations, but they have the advantage of operating on all of the channels at the same time, thus keeping them synchronized.

Once the frame of music is converted to the Unified representation, it is passed through a modified complex cross power spectral (CPS) analysis to compute a super-resolution map of the frequency components. This transformation was originally published in [10] and a modified approach has been adopted for use in music. It analyzes the phase evolution of the spectral elements in a standard FFT and uses this evolution to remap the frequencies to a much finer scale. The transformation generally gives signal accuracies on the order of 0.01Hz for stable signals at CD sample rates analyzed in 46ms windows. In the paper [10], it is shown that more detailed processing can be used to get higher signal accuracy even in the presence of significant noise or overlapping signals, but for analysis of music the modified approach is sufficient.

An example for a signal composed of three sinusoids is given in Figure 5. The original FFT spectrum and the remapped spectrum are shown; the remapped spectrum is effectively a line spectrum. For this example, the exact frequencies are 28.7965317, 51.3764239, and 65.56498312, while the estimated frequencies are 28.7960955, 51.3771794, and 65.5644420.

In real-world music the data is not as clean and stable, and the accuracy of the computed high-resolution spectrum is affected by the presence of nearby signals that interfere, modulations of the frequencies, and noise-like signals that have a broadband spectrum. Even so, in these situations, the high-resolution analysis generally gives signal accuracy on the order of 0.1Hz for any signal component that is relatively stable over the sample window. An example is given for a window of data taken from a track by Norah Jones in Figure 6. A variation of the algorithm has been developed that will provide similar resolution for a linearly modulating signal component. This algorithm returns a high-resolution estimate of the initial signal frequency in the window along, with the modulation rate.

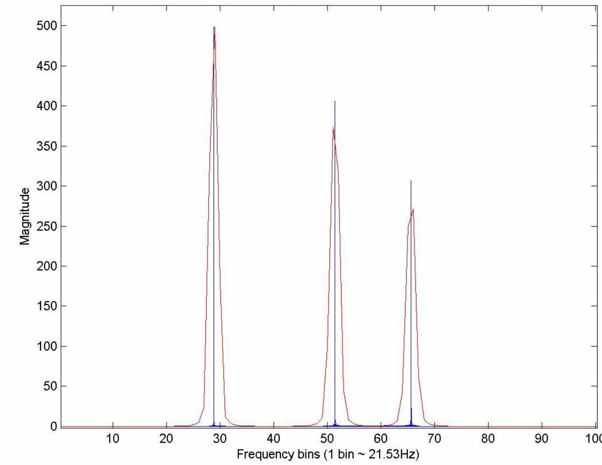

Figure 5 Comparison of high resolution spectrum to original spectrum for stable signal

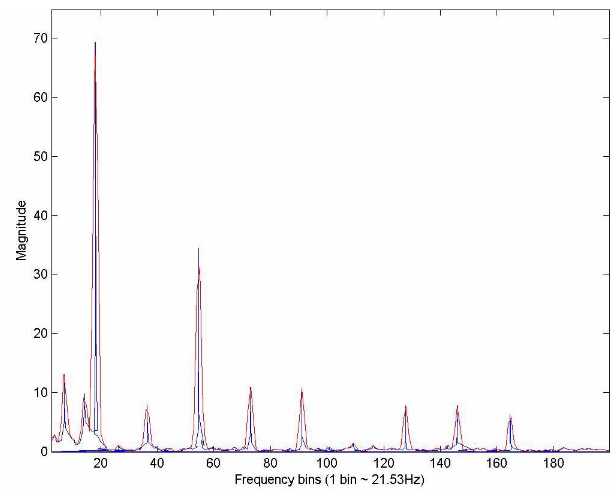

Figure 6 Comparison of high resolution spectrum to original spectrum for Norah Jones track

In the encoding process, after the signal is converted to the Unified Domain and its high-resolution CPS spectrum has been computed, principles of psychoacoustics are used to prioritize the components of the signal in order of perceptual importance. Since most psychoacoustic models used in other codecs are closely tied to the methods of signal analysis used, such models were found to be inadequate for the KOZ encoding process. However, the CPS analysis enabled the development of an effective and consistent psychoacoustic model directly in the Unified Domain. Unlike traditional psychoacoustic models, which require independent computations for each channel of data [7], the KOZ Unified Psychoacoustic Model (UPM) is designed to efficiently incorporate the effects of spectral, spatial, and temporal aspects of a signal into one algorithm. The UPM also has the benefit of being highly flexible. It can easily be customized for specific applications and facilitates psychoacoustic research by defining all functions relative to a few adjustable parameters.

The UPM computation can be separated into three steps. The first step uses the high-resolution signal analysis method described above to distinguish between tone-like and noise-like signal components in a frame of data. Following the classification of tone-like and noise-like components, the masking effect of each element is calculated based on frequency, sound pressure level, and spatial location. This computation is based on traditional models for the spreading of masking and psychoacoustic principles of spatial hearing [11][12]. Finally, the masking effects of each element are combined and projected to create the masking curve in the Unified Domain which is defined locally for each signal component in the subsequent object decomposition and quantization stages of the KOZ encoding process. This masking curve can easily be extended to create a masking "surface" defined over the entire spatial field. In addition, for stereo audio signals, traditional distinct left and right channel masking curves can be obtained with a simple transformation from the Unified Domain.

3.3. Decomposition and Representation

At this stage in the process, the frame of music has been converted to the Unified Domain and has been passed through the high-resolution spectral analysis. The high-resolution analysis converts relatively stable, tone-like signal components to line spectra with well-defined frequencies, while more noise-like signal bands do not take on structure. This allows the signal to be roughly segregated into tone-like and noise-like components. These components can be further analyzed to detect if there is the presence of a transient signal component in the frame of music, and also to test for and aggregate harmonic groupings of frequencies.

Once the separate signal components are isolated, the signal is rebuilt in an additive synthesis approach. The dominant part of this process is the step of selecting the cupolets that are a best match to the signal elements. First it is necessary to select a set of cupolets with the correct spectral characteristics for a given signal component. In this selection process, a vector of significant frequencies is determined for each component. This vector is then compared to the cupolets through a modified inner product, and the cupolet with the best psychoacoustic fit is chosen. The cupolet is then adjusted in phase and amplitude to match the original signal. A residual is computed, and the process continues in an iterative fashion until all signal components are represented.

After a good estimate is found for each of the signal components, it is possible to recreate the signal with high accuracy. For the same frame of data from the Norah Jones track in Figure 6, the signal can be reconstructed as in Figure 7 at 32kbps. Here the approximated signal appears dotted and nearly overlays the original given by the solid line A zoomed view is shown in Figure 8, and reveals some of the differences more clearly. The differences that appear are imperceptible relative to the applied psychoacoustic error bounds.

3.4. Back-end Compression

Once the analysis is complete, the input music signal is in a fully decomposed state. Oscillatory structures have been identified by their high-resolution frequency, phase and amplitude, while broader noise-like components and transient objects are represented separately. Classes of objects are then sorted in order of their "perceptual relevance" using the UPM. This ordering is a requirement for effective scalability. Finally, objects are segregated and written to the floating-point .CCA file format.

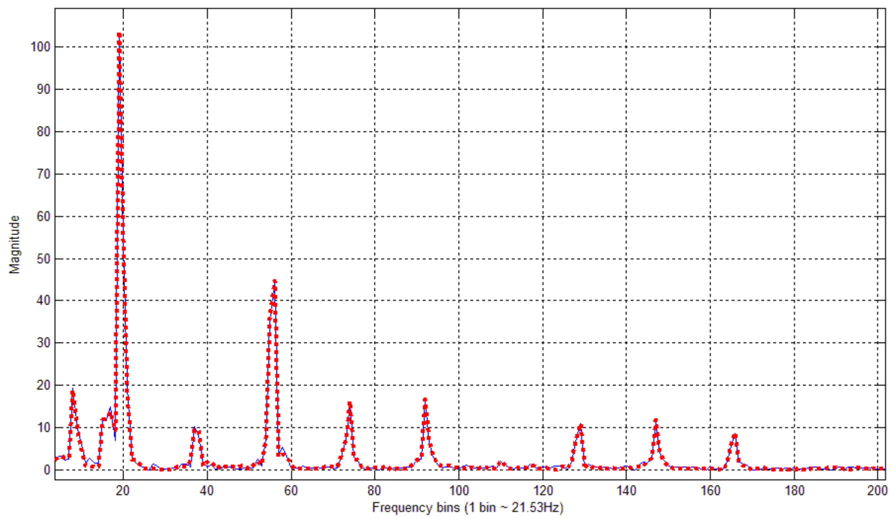

Figure 7 Comparison of 32 kbps approximated spectrum (solid) to original spectrum (dotted)

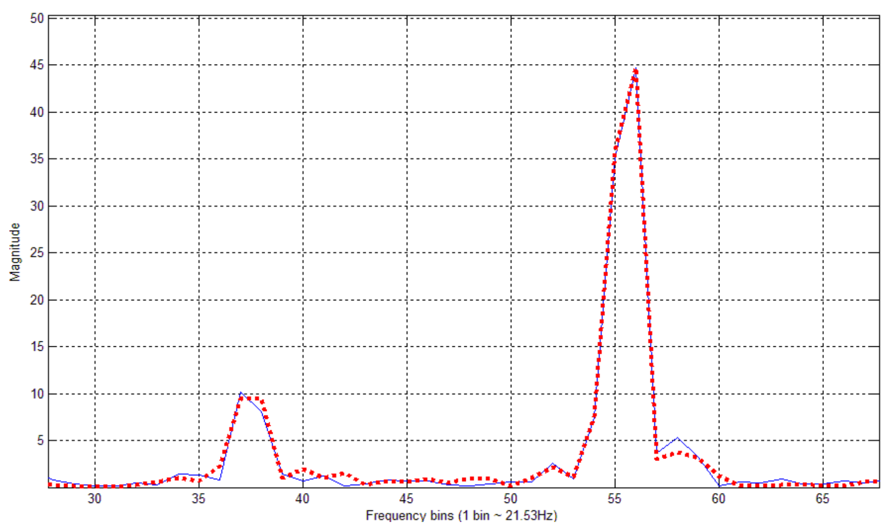

Figure 8 Comparison of 32 kbps approximated spectrum (solid) to original spectrum (dotted) [close up]

The last stage of encoding then quantizes the parameters of each signal component, based on a "sensitivity" measure derived from the UPM. This initial quantization distributes quantized values to maximize the efficiency of the lossless digital compression that is applied as the final step in reducing the size of the KOZ file. The output bitstream is then formed and packed for distribution. In this process, elements are packed so that the perceptually most relevant elements are in "lower layers". This allows very easy scalability of the file during the distribution stage by simply removing the least significant layers in order to reduce the bandwidth of the bitstream. Similarly, if more bandwidth becomes available, omitted layers can be added to the bitstream according to their psychoacoustic priority. In other cases, the application itself may dictate varying quality levels. For example, fewer layers may be used for previewing music over mobile networks, but more layers may be provided if a purchase is made.

3.5. Decoding

While the KOZ encoder is fairly complex, the KOZ decoder has been intentionally designed to be as simple and efficient as possible. During the decoding process, the quantized parameters needed to reconstruct each object are extracted from the KOZ file. These objects represent information in the Unified Domain, but they have a direct translation into either the frequency or time domains. This allows great flexibility when choosing the most efficient approach for the implementation of the decoder on a given device. Because components are reconstructed independently, the KOZ codec's structure gives the ability to alter the computational load associated with each one. Once each component is resynthesized in either the time or frequency domain, individual components are added together and the resulting frame of audio is written to the output audio buffer. The processors for the targeted devices in mobile music applications usually employ fixed-point math operations, where rounding errors add quickly and can introduce audible artifacts in the audio. For this reason, adaptive scaling of coefficients in the KOZ decoder is used to maintain a high signal-to-noise ratio while minimizing rounding error noise throughout the decoding process.

4. DEPLOYMENT IN MOBILE DELIVERY SERVICES

Deployment of the KOZ compression technology on mobile networks demands high-quality compression, the ability to run on restricted devices, and the permission of the record labels to distribute their copyrighted content, thus requiring the development of a robust DRM capability that was accepted by the music industry. Launching a mobile music service also required that billing integration and royalty settlement be included. All of these elements are part of the deployed services that currently use the KOZ technology.

Transmitting a KOZ file as part of a mobile music service involves several steps. The first step is to apply the digital rights management tools. For distribution to a wireless device, unique identifying information for the customer is present at the transmitter end on the server as well as at the receiver end on the wireless device. This unique identifier is used along with auxiliary information to feed into an encryption scheme to lock the compressed file so that it can be played on only the approved mobile device for which the transaction is recorded. The content is locked at the time of purchase, with only milliseconds of delay, and it is then transmitted over the network. A user cannot redistribute the content without first getting the content unlocked for a new approved device or user.

On the client side, the receiving device buffers the incoming data, unlocks the data using the unique identifier information resident on the client, and plays back the music, as shown in Figure 9. At this time, meta-data such as artist name, album and song titles, and album art can be extracted from the standard ID3 file embedded in the KOZ file and displayed on the client device as desired.

The KOZ technology has been a key enabler for mobile music download services that have been launched on GPRS networks in the UK, Norway, Hungary, the Czech Republic and Singapore. As of May 2005, the KOZ decoder has been implemented in both C++ and Java (including J2ME) and runs in over forty mobile phone handsets without the need for customized assembler programming. These handsets generally have ARM-family processors and run various operating systems including Windows CE, Windows Smartphone and Pocket PC Editions, and Symbian.

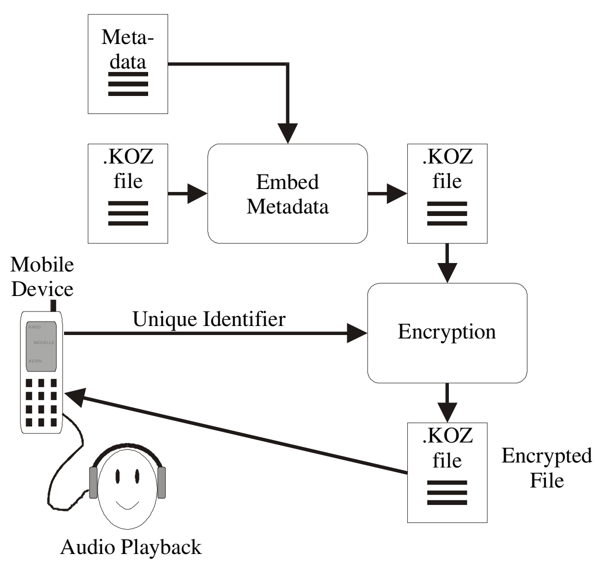

Figure 9 KOZ delivery

5. CONCLUSIONS

The KOZ compression format and the related DRM and transmission methods used in the deployed mobile music download services have all been designed to maximize the ability to transmit high-quality audio in an optimal fashion over a broad range of networks. The technology has allowed the development of a scalable, low complexity format that preserves the full CD bandwidth, and allows transmission over GPRS networks at 32Kbps for storage and playback on mobile phones and PDAs. The DRM is seamlessly integrated so that the user never notices its presence unless unauthorized redistribution is attempted, and the DRM permits the music to be streamed so that the user can listen while the download is in progress. Since the signal reconstruction methodology is additive, extra layers can be added to the stream on 3G networks to provide even higher quality. For broadband distribution, all of the signal components that were detected at the analysis and decomposition stage can be included in the transmission. The end result is a flexible encoding technology enabling users to *encode once, but access at any bitrate*.

A number of powerful tools have contributed to the development of this flexible model. The most notable among these tools are the Unified Domain representation, Unified Psychoacoustic Model, Cross-Power Spectral analysis, and chaotic cupolet generation. The ability to categorize and aggregate the signal components allows back-end quantization and lossless compression techniques that do not interfere with the capability of accessing the different layers in the file.

Since the first live demonstrations of KOZ file downloads over GPRS networks at the PopKomm Conference in Germany in the summer of 2002, the integration of the KOZ compression technology with effective DRM, billing, and royalty settlement and reporting, has allowed the development of full-featured mobile music download services in both Europe and Asia. To date, the services are available to over 25 million customers using a large variety of mobile handsets and PDAs. Continuing the development of these technologies will lead toward a future where music is available anytime, anywhere, over almost any network.

6. REFERENCES

[1] X. Serra and J. O. Smith, “Spectral modeling synthesis: A sound analysis/synthesis system based on a deterministic plus stochastic decomposition," Computer Music Journal, vol. 14, pp. 12{24}, (1990)

[2] R. J. McAulay, and T. F. Quatieri. “Speech Analysis/Synthesis based on a Sinusoidal Representation.” IEEE Transactions on Acoustics, Speech and Signal Processing 34(4):744-754, (1986)

[3] M. K. Johnson, "Controlled Chaos and Other Sound Synthesis Techniques," Thesis for the Degree of Bachelor of Science, University of New Hampshire, May 2000

[4] S. Hayes, C. Grebogi, and E. Ott, Communicating with Chaos, Phys, Rev. Lett. 70, 3031 (1993)

[5] M. Takashi, “Chaos in Electronic Circuits”, Proceedings IEEE, Vol. 75, No. 8, pp. 1033-1057, Aug. 1987

[6] US patents 6,137,045 and 6,363,153. Other patents pending.

[7] M. Bosi and R. E. Goldberg, *Introduction to Digital Audio Coding and Standards,* Kluwer Academic Publishers. (2003)

[8] Eric W. Weisstein. "Special Unitary Group." From MathWorld--A Wolfram Web Resource. http://mathworld.wolfram.com/SpecialUnitaryGroup.html

[9] Bhandari, “Polarization of Light and Topological Phases,” Physics Reports 281 (1997) 1-64.

[10] D. J. Nelson and K. M. Short. A channelized cross spectral method for improved frequency resolution. Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis. IEEE Press, October 1998.

[11] E. Zwicker and H. Fastl, *Psychoacoustics: Facts and Models,* Second Edition, Springer. (1999)

[12] P. R. Cook (Ed), *Music, Cognition and Computerized Sound: An Introduction to Psychoacoustics,* MIT Press, (1999)