Building a New Era of Open Multimedia
Overview of the Xiph.org Foundation Projects

Jack Moffitt

Xiph.org Foundation
`jack@xiph.org`

Abstract. The Xiph.org Foundation is a not-for-profit research and standards organization creating royalty-free, open standards for audio, video, streaming, and related forms of multimedia. The main motivation for the organization is to replace encumbered standards such as MP3, Quicktime, and RealAudio/Video, with technically superior unencumbered ones, suitable for free implementation in all environments, and to create a unified base platform in place of the fragmented set of technologies from other organizations that is currently dominating Internet multimedia. Our projects include the Vorbis audio codec, the Theora video codec, the Speex voice audio codec, FLAC (Free Lossless Audio Codec), the Icecast streaming media server, and the Ogg container and transport format. A high level technical overview of the main technologies will be presented as well as how each of them fits into the larger framework. Additionally a few fledging projects will be mentioned such as Writ, a subtitling system, and MNG over Ogg for graphic overlays.

Key words: Open Standards, Multimedia, Audio, Video, Streaming

1   Introduction

In the late 90s, an audio encoding technology called MP3 became very popular. There were dozens of free encoders, players, and utilities developed by interested programmers all over the world, and online music became accessible to everyone. In September of 1998, Fraunhofer, the German company which developed much of the technology, sent letters out to MP3 encoder developers everywhere demanding patent royalties. Many of the MP3 encoders disappeared as a result. This was only the first time the royalty structure around the MP3 technology had been changed.

In that same September of 1998, work on a replacement technology that didn't suffer from such licensing restrictions began, and a few years later, the Vorbis audio codec was completed.

Those of us working on the Vorbis project had even bigger aspirations, and so in 2000, we started the Xiph.org Foundation (Xiph). Xiph's mission is to design, develop, and promote royalty-free multimedia standards. Such standards are needed to provide a solid foundation on which to build future multimedia technology, just as the Internet Protocol (IP) was needed before we could have the World Wide Web.

Xiph has many projects, some mature, some only research. This paper will give a brief overview and some technical details of several of these projects.

2   Vorbis Audio Codec

The Vorbis audio codec transforms audio, be it speech, music, or both, into a compressed bitstream and back. It was specifically designed to compete against codecs like MP3, VQF, and RealAudio at bitrates from 8 to 768 kbps.

Vorbis uses a mixture of well known techniques from the early days of signal processing and novel ideas in order to perform better than its competitors. These include psychoacoustic properties such as the Absolute Threshhold of Hearing (ATH) and frequency masking, the MDCT, vector quantization of residues, as high level features such as multichannel coupling. The codec also preserves sample accurate seeking and editing, which was missing from a lot of the other, similar technologies.

Vorbis also stores special tables, called codebooks, with each bitstream, instead of preselecting a fixed set of codebooks that every implementation uses. This allows Vorbis to have vastly superier longevity compared to other codecs, as changes to the encoder can be incorporated without breaking any existing implementations.

These features give Vorbis a technical advantage that is realized in smaller bistreams, higher perceived quality by listeners, gapless playback, and powerful, accurate seeking.

Research is already underway for future improvements to the technology, incorporating wavelets and improved psycoacoustics.

3   Theora Video Codec

The Theora video codec is an extension of the VP3 video codec developed by On2 Technologies. On2 graciously donated the code to Xiph and granted royalty-free use of their patents related to the technology to everyone. Theora is a general purpose video codec, similar in design to other MDCT based codecs such as MPEG 4.

Xiph rewrote the implementation of VP3, made some minor improvements, and added the ability to tweak codebooks and format parameters into the individual bitstreams, similar to the Vorbis codec.

A new experimental encoder is being built to exploit these enhancements, based on a perceptual model of human vision. This new encoder should improve coding efficiency significantly over the current standard encoders, resulting in much higher quality output.

4   Speex Voice Codec

The Speex project started out independently in 2002. In 2003, Jean-Marc Valin, the lead developer, joined the Xiph.org Foundation and donated the Speex codec. Speex is an audio codec designed primarily for voice content at very low bitrates, from 2 to 44kbps. It is based on the CELP family of algorithms.

Speex is robust to lost packets, but not to corrupted ones, as the codec was designed from the outset for use in VoIP as opposed to most other similar codecs, which were designed for cellphones. Unlike Vorbis or FLAC, the Speex encoder is designed to run in realtime with very little CPU power, as its intended use is for two way communication.

Speex also uses voice activity detection to significantly improve coding efficiency when there is only background noise or near silence. In these non-vocal periods, only the minimum number of bits are used to reproduce 'comfort noise generation'.

5   Free Lossless Audio Codec (FLAC)

FLAC is the most recent addition to the Xiph.org Foundation's projects. It is a completely lossless audio codec which achieves com-

pression ratios on general music files of about 1.5:1 or more. This makes FLAC less ideal for Internet transmission, but very well suited to long term storage.

Since FLAC is a lossless codec, it makes every effort to check for errors in the data. It has both a CRC check in every frame, as well as a global MD5 sum check. In addition, errors in encoding affect only one frame, giving FLAC a big technical advantage over similar codecs.

FLAC is also fully seekable, streaming, and supports flexible metadata (similarly to Vorbis). It is also one of the very few lossless audio codecs to be implemented on embedded platforms, due to its ability to be decoded very fast — much faster than most perceptual lossy audio codecs.

6 Ogg Bitstream Format

In order to provide high level support for seeking, editing, and streaming, codec bitstreams are encapsulated into an Ogg bitstream. Ogg provides the infrastructure for multi-codec bitstreams, bitstream chains, framing information, error correction, and seeking structure.

The Ogg bitstream is designed so that it can be created in only one pass. This means that seeking is done with hints and binary search as opposed to having a global index.

Bitstreams are broken up into pages, which can hold partial, or many, packets. Each page holds packets for only one codec type. Pages from different codecs will be interleaved in synchronized multi-codec files. Additionally, concatenating Ogg bitstreams will produce another, valid Ogg bitstream, allowing one to easily combine bistreams. This is exactly how the HTTP based radio style streaming works.

7 Icecast Streaming Server

The Icecast streaming server provides live feed and static file streaming for the Xiph.org Foundation codecs. Using Icecast and its related tools, an Internet user can run their own radio station, capable of supporting as many clients as their bandwidth allows.

Unlike some other similar servers, Icecast uses a protocol heavily based on HTTP and transmits data over TCP. This alliviates many problems related to corporate and personal firewalls blocking UDP and also the lack of support in most places for multicast.

Icecast supports broadcast relays, remote administration, centralized directory servers for publishing broadcast information, and sophisticated statistics gathering and processing. It also does automatic playlist generation to facilitate linking from webservers, and contains its own small webserver capable of serving static files and doing XSLT transformations of statistics data for presentation.

8 Upcoming Projects

The Xiph.org Foundation has many fledgling projects that have not yet reached alpha status, or are still in research stages. These projects include MNG over Ogg, Vorbis over RTP, subtitling formats, and Tarkin/W3D.

MNG over Ogg is an embedding of the Motion PNG codec into an Ogg bitstream. This could be used to provide graphic overlays or subtitles in a video stream, create slideshows, or as a video codec of its own using motion JPEG.

Vorbis over RTP is an Internet Draft describing the encapsulation of Vorbis codec data in the RTP protocol. Since RTP provides many of the same features as Ogg encapsulation, Ogg is redundant in this context, and a native encapsulation desireable. This will allow multicast streaming of Vorbis format data.

Several ideas at subtitling formats are being researched, including using MNG over Ogg. The main thrust in this area is the tentatively named Writ project, which provides a new subtitling codec for text based subtitles — as opposed to graphics based subtitles — in Ogg format bitstreams. Xiph hopes to have work in this area complete by the 1.0 release of the Theora codec.

Tarkin and its successor project, W3D, are next generation 3D wavelet based video codecs. These use fundamentally different algorithms from currente video codecs which are based heavily on the use of the DCT. These projects are currently in research and experimentation stages, although they both produce bitstreams and contain example players.

9 Conclusion

The Xiph.org Foundation is dedicated to creating open, royalty-free multimedia standards and implementations that are state of the art. Our primary codecs, Vorbis, Theora, Speex, and FLAC, raise the bar for multimedia, in addition to providing an open platform for the growth and experimentation of Internet multimedia technology. Our other projects complement and extend this base, and our research efforts give an idea of what is possible now that such a base exists. The Xiph.org Foundation will continue to develop, research, and maintain these technologies in order to ensure that the future of Internet multimedia is bright and fruitful.