codec
Internet-Draft
Intended status: Standards Track
Expires: November 25, 2016

M. Graczyk
Google Inc.

May 24, 2016

Ambisonics in an Ogg Opus Container
draft-graczyk-codec-ambisonics-01

Abstract

This document defines an extension to the Ogg format to encapsulate ambisonics coded using the Opus audio codec.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on November 25, 2016.

Copyright Notice

Copyright (c) 2016 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction

Ambisonics is a representation format for three dimensional sound fields which can be used for surround sound and immersive virtual reality playback. See [gerzon75] and [daniel04] for technical details on the ambisonics format. For the purposes of the this document, ambisonics can be considered a multichannel audio stream. Ogg is a general purpose container, supporting audio, video, and other media. It can be used to encapsulate audio streams coded using the Opus codec. See [RFC6716] and [RFC7845] for technical details on the Opus codec and its encapsulation in the Ogg container respectively.

This document extends the Ogg format by defining a new channel mapping family for encoding ambisonics. The Ogg Opus format is extended indirectly by adding an item with value 2 to the IANA "Opus Channel Mapping Families" registry. When 2 is used as the Channel Mapping Family Number in an Ogg stream, the semantic meaning of the channels in the multichannel Opus stream is the ambisonics layout defined in this document. This mapping can also be used in other contexts which make use of the channel mappings defined by the Opus Channel Mapping Families registry.

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

3. Ambisonics With Ogg Opus

Ambisonics MAY be encapsulated in the Ogg format by encoding with the Opus codec and setting the Channel Mapping Family value to 2 in the Ogg Identification Header. A demuxer implmentation encountering Channel Mapping Family 2 MUST interpret the Opus stream as containing ambisonics with the format described in Section 3.1.

3.1. Channel Mapping Family 2

Allowed numbers of channels: (1 + n)^2 for n = 0...14. Explicitly 1, 4, 9, 16, 25, 36, 49, 64, 81, 100, 121, 144, 169, 196, 225. Ambisonics from zeroth to fourteenth order.

This channel mapping uses the same channel mapping table format used by channel mapping families 1 and 255. Each output channel is assigned to an ambisonic component in Ambisonic Channel Number (ACN) order. The ambisonic component with order n and degree m corresponds to channel (n * (n + 1) + m). The reverse correspondence can also be computed for a channel with index k.

```
order   n = ceil(sqrt(k)) - 1,
degree  m = k - n * (n + 1).
```

Channels are normalized with Schmidt Semi-Normalization (SN3D). The interpretation of the ambisonics signal as well as detailed definitions of ACN channel ordering and SN3D normalization are described in [ambix] Section 2.1.

3.2. Downmixing

Implementations MAY use the matrix in Figure 1 to implement downmixing from multichannel files using Channel Mapping Family 2 Section 3.1, which is known to give acceptable results for stereo. The first and second ambisonic channels are known as "W" and "Y" respectively.

```
/   \   /                 \ /  W  \
| L |   | 0.5  0.5 0.0 ... | |  Y  |
| R | = | 0.5 -0.5 0.0 ... | | ... |
\   /   \                 / \ ... /
```

Figure 1: Stereo Downmixing Matrix

4. Security Considerations

Implementations of the Ogg container need take appropriate security considerations into account, as outlined in Section 10 of [RFC7845]. The extension definied in this document requires that semantic meaning be assigned to more channels than the existing Ogg format requires. Since more allocations will be required to encode and decode these semantically meaningful channels, care should be taken in any new allocation paths. Implementations MUST NOT overrun their allocated memory nor read from uninitialized memory when managing the ambisonic channel mapping.

5. IANA Considerations

This document updates the IANA Media Types registry "Opus Channel Mapping Families" to add a new assignment.

Value	Reference
2	This Document Section 3.1

6. Acknowledgments

Thanks to Timothy Terriberry and Marcin Gorzel for their guidance and valuable contributions to this document.

7. References

7.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC6716] Valin, JM., Vos, K., and T. Terriberry, "Definition of the Opus Audio Codec", RFC 6716, DOI 10.17487/RFC6716, September 2012, <http://www.rfc-editor.org/info/rfc6716>.

[RFC7845] Terriberry, T., Lee, R., and R. Giles, "Ogg Encapsulation for the Opus Audio Codec", RFC 7845, DOI 10.17487/RFC7845, April 2016, <http://www.rfc-editor.org/info/rfc7845>.

[ambix] Nachbar, C., Zotter, F., Deleflie, E., and A. Sontacchi, "AMBIX - A SUGGESTED AMBISONICS FORMAT", June 2011, <http://iem.kug.ac.at/fileadmin/media/iem/projects/2011/ambisonics11_nachbar_zotter_sontacchi_deleflie.pdf>.

7.2. Informative References

[gerzon75] Gerzon, M., "Ambisonics. Part one: General system description", August 1975, <http://www.michaelgerzonphotos.org.uk/articles/Ambisonics%201.pdf>.

[daniel04] Daniel, J. and S. Moreau, "Further Study of Sound Field Coding with Higher Order Ambisonics", May 2004, <http://pcfarina.eng.unipr.it/Public/phd-thesis/aes116%20high-passed%20hoa.pdf>.

Author's Address

Michael Graczyk
Google Inc.
1600 Amphitheatre Parkway
Mountain View, CA 94043
USA

Email: mgraczyk@google.com