swissaudec®
Sound science

ECMA-407 "Instant HD to UHD Audio" White Paper

Superior UHD in a nutshell, by Clemens Par, CEO, Swissaudec

Clemens Par, CEO, Swissaudec

CLEMENS PAR has introduced inverse problems and invariants to the world of audio coding. ECMA-407 standardizes these results as the world's first UHD 3D audio codec. Clemens Par is founder and CEO of Swissaudec, a young Swiss codec company located near Lausanne in the Canton of Vaud, highly active in international standardization at Ecma International and inside ISO/ MPEG. He has served as elected convenor to Ecma TC32-TG22 since 2012. This publication is a tribute to professor Paul Kleihues, prominent oncologist and pioneer of medical science.

ECMA-407 broadcasts UHD audio up to NHK 22.2 inside backwards-compatible HD. The ECMA-407 payload < 2kbps travels by internal multiplex over satellite, antennae or IP, to Smartphones, tablets, computers and TVs - at NO cost for the HD broadcaster who wishes to maintain e.g. AC3, DTS, MPEG-4, MPEG-D, or Opus.

A virtual reality for UHD audio?
4K and 8K pictures are luring the broadcaster expert and amazed viewer whilst the consumer electronics industry has and is preparing for visual revolution with advanced ATSC 3.0 and MPEG-H (HEVC) video codecs. **IHS expects 1'053 UHD channels till 2025, UHD TV sets are forecasted to exceed a quarter billion already in 2017**.[1] Will they all incorporate UHD sound, too? A crucial question to which UHD audio codec manufacturers currently seem not to contribute:

When BBC stopped research in 3D video, it had become clear that unrestricted immersiveness would be delegated to UHD audio, ideally NHK 22.2, which is able to reproduce sound on a hemisphere with highest accuracy, by enhancing granularity, as the human ear is not equally susceptible to localization (sound source direction), with eleven speakers at the front. Any speaker subset can be derived from NHK 22.2 by means of downmix (the adding up of neighbouring speakers). **The ideal state-of-the-art would be unrestricted NHK 22.2 transmission regardless bitrate requiring compression rates up to almost 600 times**.

Three UHD audio standards are capable to transmit NHK 22.2, i.e. MPEG-H, ATSC 3.0 and the Ecma S5 standards family, in particular ECMA-407. **However, only ECMA-407 is capable to achieve sufficient compression at lowest bitrates to transmit NHK 22.2 whilst, for instance, MPEG-H is restricted to 9.1**.

In terms of computational complexity, the Ecma S5 family saves up to more the half with respect to its competition, which is most important for Smartphones and tablet use.

▶
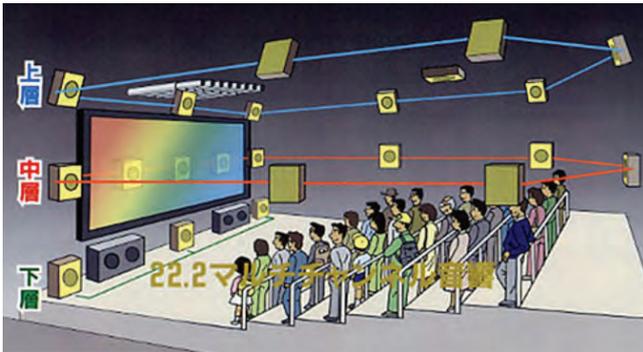
Fig 1: NHK (Hamasaki) 22.2 loudspeaker setup, as broadcasted by Japanese ARIB from 2016 onwards.[2]

► **Are you fit for UHD sound?**
UHD audio codecs, however, is awaiting bad news. Affording sensors with higher video resolution and subsequent infrastructure for processing and post-processing is primarily an investment in terms of money. Creating UHD audio content requires severely augmented skills, bulky workflows in terms of recording, mixing and processing. Some broadcasters declare to have not yet accomplished HD workflow in practice. Mixing in Surround sound is reserved to a handful of engineers, some with considerable technical skills, some with a tendency to experimental recording outside meaningful electroacoustic experience (often called "the alchemists ").

Broadcasters seems to have made up their choice: their UHD is 4K together with HD 5.1 Surround. Too big a risk to renew already costly infrastructure and further stretch the know-how inside daily workflows. UHD audio seems to remain a cinema domain.

UHD unreconciled
Too bad for the UHD audio codec world! Clients are announced prior to meaningful market-relevant projects and in the end turn out to be vanishing interest groups or studies for future broadcast environments, or even a standardization project where industry has declared "interest" in, whilst the proponents are subsequently kicked out by their competitors.

In such situation one might think that UHD audio proponents work together, for mutually wetting customer's appetite for an immersive sound experience, in forming one single interest group. Whilst I have been pleading for such attitude in industry since being active in UHD 3D audio research as a scientist and expert to ISO and Ecma International, the contrary has happened, two giant lobbies beating themselves up and attacking the third fair competitor, which is a small Swiss start-up, currently raising its voice in favour of science, reason and fair market practices.

UHD hide-and-seek inside HD
As far as science is concerned, Swissaudec has already laid down its ideas with Intercomms latest two issues (see

"*Taming the Beast in Mankind – Telecommunications in the 21st Century*", and "*Rationalism versus Empirism. A Crash Course in Invariant Theory and a Tribute to Rudolf E. Kálmán*"). As far as reason is concerned, **Swissaudec pleads for a backwards-compatible UHD solution to highly successful HD standards and technologies like AC-3, DTS, MPEG-4, and Opus, instead of urging the market towards global replacement of all devices, without any gain in subjective quality or computational efficiency with respect to smart HD extensions. Smart HD extensions mean "green" codecs, which by their technological structure help to reduce global electronic waste**.

Some broadcasters might never change from HD audio to UHD audio. This last remnant is considerable large and in our estimate represents 75% of global markets. Why should this remnant pay license fees despite nothing is going to happen? UHD audio codec designers are completely ignorant to this silent majority. Alas, 3D sound is too luring and tasty for engineers to care about their customers.

Hope for an UHD sound hype?
Unfortunately, I cannot advocate my competing colleagues' attempt to bully broadcasting markets, themselves and their competition by all means. I have worked too long as a director, executive producer and moderator for public broadcasting. Technology change outside the range of meaningful transition is a too big ask. This is why UHD audio seems to fail, despite its unequalled fascination to the engineer and the end consumer.

*Fig 2: Instant UHD audio inside HD – ECMA-407 on satellite and mobile devices in IBC 2015's "Future Zone". **The "Full UHD" satellite carrier of France Télévisions with SES with ECMA-407 was chosen to represent the technology assets of more than 1'700 exhibitors of IBC 2015 in "What caught my eyes. New technology, new content."***

We assume for the technology:
1. **Unrestricted parallel availability of UHD audio anytime, anyplace over satellite, antennae and the Internet** should not compromise current HD workflows and infrastructures.
2. Existing audio codecs, e.g. **AC-3, MPEG-4, DTS, or Opus** on the marketplace show sufficient performance to merit the immediate extension to UHD.
3. The switch to advanced audio codecs like **MPEG-D (USAC)** ►

►   should not impact already existing UHD workflow.

4. **Automatic HD to UHD conversion will be the primary asset for the successful future of UHD broadcasting**.
5. Broadcasters wish to change to **backwards-compatible** UHD by instant **plug-and-play** solutions.
6. UHD should not require any further technical exigencies.

We assume for the marketplace that HD broadcasters will remain a 75% majority, unless all exigencies above can be equally met. **However, the label "Full UHD" remains the instant wish of the broadcaster**.

Conventions and market reality

The technology requirements for such assumptions are too strict to be met by parametric coding approaches:

1. A major MPEG-H proponent did not expect MPEG-H to happen outside OTT, due to its heavy-load structure, which is sad news for the satellite world and for broadcasters' costly terrestrial networks.
2. When complemented with parametric 3D coding approaches neither AC-3, MPEG-4, DTS, nor Opus might meet broadcasters' quality requirements, due to high quantization noise.
3. 2D becomes a small world inside 3D. 3D inside 2D is neither offered by ATSC 3.0 nor by MPEG-H, due to bulky spatial payloads.
4. Upscaling to UHD is a highly advanced technology. Upmixing from HD to UHD audio, particularly to NHK 22.2, so far only occurs interactively (see, for instance, McGill University's Space Builder designed for Japanese broadcaster NHK).
5. UHD audio means to renew all infrastructures and devices with ATSC 3.0 or MPEG-H.
6. This is a too huge ask, in Swissaudec's opinion.

Swissaudec's White Paper for instant "Full UHD"

The only standardized non-parametric alternative is the Ecma S5 standards family, which insert an **invisible UHD Ecma S5 payload of less than 2kb/s for NHK 22.2 or less for all other known industrial loudspeaker configurations via internal multiplex** (e.g. via the data_stream_element of MPEG-4, or the UsacExtElement of MPEG-D) with **perfectly equal subjective performance to parametric approaches and significantly reduced computational complexity on the decoder's side**.[3] This internal multiplex solves all problems at once and addresses 100% of the current HD market.[4]

1. An internal multiplex, though standardized, does neither alter workflows nor infrastructure. **It comes for free**.
2. **The Ecma S5 family is codec-agnostic**, i.e. any known base audio codec can be used without any impact on quality. The reason is that the spatiality of the original signal is mimicked by invariant analysis and can consequently be reduced to mere gains and delays, without any subjective decrease of the UHD signal in comparison with parametric spatial approaches.[5] Even
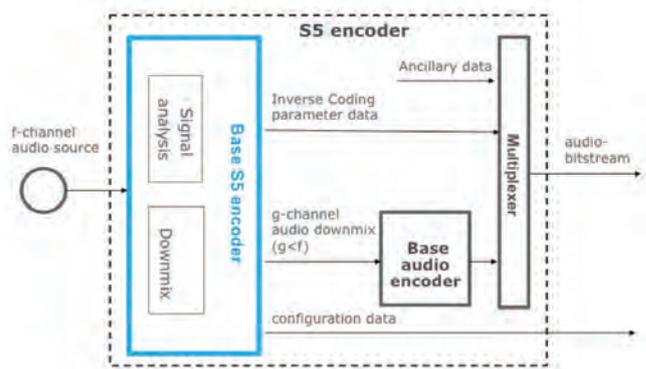
Fig 3: Swissaudec's White Paper in practice. "Plug-and-play" broadcaster unit for parallel satellite, antennae and OTT transmission with automatic invariant-driven HD to UHD audio conversion ("upmix"), as an extension to the encoder's "Signal analysis". Further workflow inside the "Base S5 encoder" is lossless. The broadcaster's HD "Base audio encoder" remains unchanged.

different base audio codecs can be used at the same time, e.g. MPEG-D on satellite and antennae, and MPEG-4 or even MP3 on the Internet!

3. **Any existing 2D audio infrastructure can be used to have UHD travel inside. If no Ecma S5 decoder is present, ordinary HD will be decoded in a backwards-compatible way**. The Ecma S5 decoder, e.g. ECMA-407 can even be configured on the encoder side, i.e. the broadcaster can himself update the decoder within the ECMA-407 payload by using simple "Polish notation". This has never happened in the codec world to such extent, and, in particular not for the decoding of highly complex UHD audio signals.[6]
4. The same technology, as already is used with the Ecma S5 family, in particular, ECMA-407, can be used for "upmixing", i.e. for **automatic HD to UHD conversion**. ECMA-407, for instance, is an upmixing technology itself, which mimics a given signal by invariant analysis for one single frame of a signal of several minutes' length at incredible speed. The same principle can be used to create fascinating UHD 3D sound up to NHK 22.2 out of mono, stereo and Surround. After "upmixing" HD to UHD, the Ecma S5 family takes over by using precisely the same technologies for encoding. **Hence the encoding remains lossless with respect to the UHD "upmix"**. The UHD "upmix", technically spoken, is a minimal proprietary extension of Swissaudec to the standardized Ecma S5 encoder.
5. The broadcaster connects a simple hardware unit (with the UHD Ecma S5 encoder plus its UHD extension followed by the HD base audio codec of his choice) with his mono, stereo or HD Surround audio infrastructure. Transition from HD to UHD invisibly takes place. **If no Ecma S5 decoder is present in the consumer's devices, backwards-compatible HD is decoded, otherwise UHD audio in optimum quality is decoded**.
6. With Ecma S5, UHD audio comes instantly for free as a **"green" codec to already existing devices**.

►

► The technology offers only advantages without disadvantages (including **energy signature** and state-of-the-art **loudness** inside the Ecma S5 bitstream). This is why Swissaudec is currently fought intensely by its competition, as a 100% market share can be immediately achieved. We wish to release the broadcaster from the need to pay any attention to UHD audio without having his "golden ear" sound engineers and his broadcasting engineers complain *unisono*. He plugs a box. That's all.

Fig 4: Two ECMA-407 "boxes" in IBC 2014's "Future Zone". Mayah Communication's Centauri-IV 5000 automatically converts HD to UHD audio and encodes and decodes UHD with NHK 22.2 (hidden inside MPEG-4 and Opus as internal multiplex) **over an HEVC/MPEG-4 satellite carrier of SES, in real-time**. The "Full UHD" program was crafted by France Télévisions.

UHD audio for all

There are multiple ways towards rendering UHD audio:
    Apart from headphones, future UHD TV devices will incorporate highly directive dipole speakers with crosstalk cancellation mimicking virtual headphones to the ears of the listeners. LAFs (Loudspeaker Array Frames), as developed by NHK and Japanese academia will achieve a similar effect, as showcased at NAB 2015 and IBC 2016. Such devices will face a dramatic increase in living rooms, as will already existing smart devices be able to decode UHD audio with Ecma S5, in particular ECMA-407, **at highest speed with lowest computational complexity** - by means of a downloadable app.

Swissaudec - science overrules industry

We wish to shape the UHD future by reason and unequalled scientific approaches. **Swissaudec was the world's first company to introduce an UHD audio standard with highest spatial compression ever achieved (up to 99%)**, by introducing inverse problems (well known in medical computer tomography) to electroacoustics and by discovering the first coefficient functions (invariants) with random processes, used with ECMA-407's "Signal analysis", **for fastest encoding with increased coding accuracy**, and with a new coding method, which can retrieve the

double number of input channels with zero side information ("inverse frequency component extraction").[7]
    Reason acts outside jealousy, greed and dominance. Our wish to instantly change the HD world to UHD is not only to further broadcasting industries but also to introduce 3D medical ultrasound in the same way as computer tomography works today.[8] This may save millions of lives in a cost-efficient way and bring wealth to developing nations, which are financially eaten up by their disproportional healthcare expenses.

Acknowledgements:
This article hence is dedicated to professor Paul Kleihues who has (together with professor Rudolf E. Kálmán) encouraged us to remain faithful to science and human welfare.

Footnote references:
1. SES White Paper "Ultra HD". 09/2015.
2. Courtesy of NHK.
3. See ISO/IEC JTC1/SC29/WG11 MPEG 2016/M37529.
4. International standard ECMA-407. 1st edition. June 2014.
5. C. Par. Internationaler Standard für modularen 3D-Audio-Transport. Teil 2. FKT 05/2015.
6. See Annex B of international standard ECMA-407. 1st edition. June 2014.
7. See Ecma TC32-TG22. ECMA-XXX v.1.
8. "...calculating for example an image in computer tomography or a source reconstruction in acoustics." (Inverse Problems International Association, see http://www.inverseproblems.info/start)

For more information please visit:
www.swissaudec.com