

きれいな音で伝えます。携帯電話も放送も。

携帯電話向け音声音響符号化EVSとハイレゾ放送向け音響ロスレス符号化ALS

How to transmit high-quality sound via network

Research and deployment of speech codec EVS for VoLTE and lossless audio codec ALS for broadcasting



守谷特別研究室
鎌本 優
Yutaka Kamamoto

プロフィール

NTT コミュニケーション科学基礎研究所 守谷特別研究室 研究主任。情報処理学会論文賞、日本音響学会粟屋潔学術奨励賞、電気通信普及財団テレコムシステム技術奨励賞・技術学生賞、各受賞。電波産業会 (ARIB) 音声符号化作業班 委員。IEC MT 61937-10 Convenor。超臨場感コミュニケーション産学官フォーラム (URCF) 超臨場感音響技術に関する調査WG幹事。

携帯電話向け音声音響符号化EVS

移動通信システムの国際標準規格を策定する3GPP (The 3rd Generation Partnership Project) では将来の携帯電話やVoLTE (Voice over Long-Term Evolution) 向け音声音響統合符号化方式として、Enhanced Voice Services (EVS) を新たに策定しました [1]。EVSコーデックは音声向けの時間領域の符号化方式と音楽向けの周波数領域の符号化を組み合わせたものであり、これまで実現できなかった、背景雑音や背景音楽を含む音声信号や音楽信号そのものを高音質のまま低遅延で伝送することが可能となっています [2]。

EVSコーデックは8 kHz (Narrow-band; NB), 16 kHz (Wide-band; WB), 32 kHz (Super-WB; SWB), 48 kHz (Full-band; FB) というように幅広いサンプリングレートをサポートし、1フレーム20 msごとに伝送し、アルゴリズム遅延32 msという通話可能な遅延で処理を行います。また、VoLTE向けということもあり、ビット列の誤り耐性ではなくパケット消失を隠蔽する方策がとられ、パケットベースの伝送に適した方式となっています。さらに、対応する

ビットレートもNBは5.9 kbpsから24.4 kbps、WBは5.9 kbpsから128 kbps、SWBは9.6 kbpsから128 kbps、FBは16.4 kbpsから128 kbpsというように幅広く対応し、任意の音声帯域・任意のビットレートにフレームごとに切り替えることができます。そして、現在普及しつつあるAMR-WB (Adaptive Multi-Rate WB)との互換性も具備するため、シームレスなコーデックの置き換えが可能です。

各社の提案する技術を正当に評価するために、標準化の過程において、様々な条件・音源・言語での大規模な主観品質評価実験が第三者機関によって行われ、実験結果が評価レポートに示されており、従来の音声符号化方式・音響符号化方式よりも高性能であることが確認されています。技術を提案した12社が協力し、統一方式を改善することにより、従来の符号化方式よりも高性能のEVSを策定することに貢献しました。

ハイレゾ放送向け音響ロスレス符号化ALS

放送の音声は、放送局内では48 kHz, 24 bitのいわゆるハイレゾ品質で制作されています。ところが、放送を家庭まで届けるために使われる電波は有限な資源であり節約

しなければならないため、仕方なく歪のある圧縮を行い皆様の元へ届けられていました。これから始まる4K/8K放送では、映像で使われる情報量が大きくなるため、それに伴い、音声に割り当てられる情報量も大きくなるのが期待されます。音声を伝送するための符号化方式の一つとして、日本の放送規格であるARIB STD-B32では音響ロスレス符号化のMPEG-4 ALSが使えるようになりました。ALSにより放送局のスタジオで制作された音声をそのまま家庭まで届けることができます。電波で届く放送に先駆けてIPTVでの放送コンテンツ伝送においても、ALSが使われる可能性が高まっています。

ALSは圧縮しても元に戻る国際標準技術で [3]、NTTも標準化に貢献してきました [4]。圧縮率は入力信号に依存しますが、おおよそ30%~70%の大きさのビットレートへ圧縮することができます。NTTでは標準策定後も周辺分野の標準化に取り組み、より使いやすいツールに限定したMPEG-4 ALS Simple Profile規格、日本の放送規格であるARIB STD-B32、デジタルアンプへの接続規格であるIEC 61937-10の標準化やそれらの改訂作業を進めてきました。今後、ALSが普及することにより、皆様が意識せずとも放送の音質が向上することが期待できます。

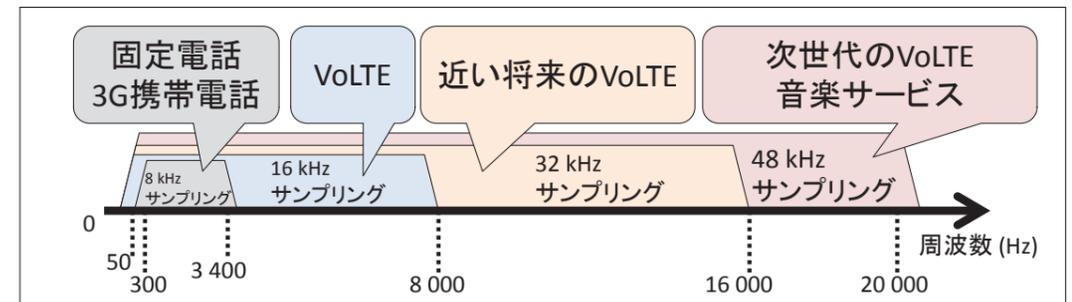


図1: 電話帯域の高音質化

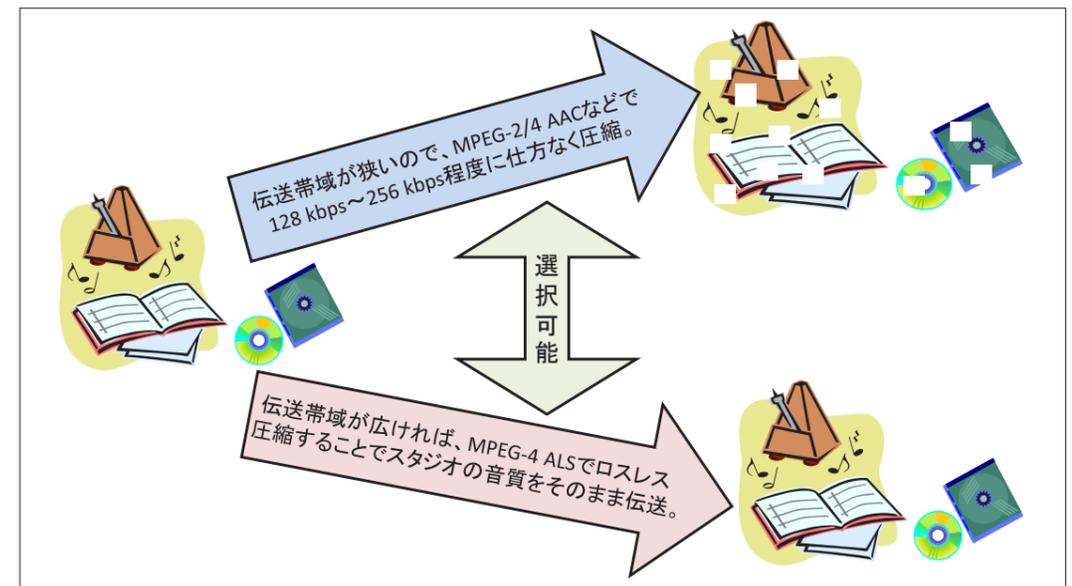


図2: 放送音声の高音質化

【関連文献】

[1] 3GPP TS 26.441, "Codec for Enhanced Voice Services (EVS); General overview,"
 [2] M. Dietz, M. Multrus, V. Eksler, V. Malenovsky, E. Norvell, H. Pobloth, L. Miao, Z. Wang, L. Laaksonen, A. Vasilache, Y. Kamamoto, K. Kikuri, S. Ragot, J. Faure, H. Ehara, V. Rajendran, V. Atti, H. Sung, E. Oh, H. Yuan, C. Zhu, "Overview of the EVS codec architecture," in Proc. 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2015), pp. 5698 - 5702, 2015.
 [3] ISO/IEC 14496-3, "Information technology - Coding of audio-visual objects - Part 3: Audio," fourth edition, 2009.
 [4] T. Liebchen, T. Moriya, N. Harada, Y. Kamamoto, Y. A. Reznik, "The MPEG-4 Audio Lossless Coding (ALS) Standard -Technology and Applications," in Proc. AES 119th Convention, 2005.