ABSTRACT

The rapid expansion of low-power embedded systems, edge devices, and real-time communication infrastructures has renewed interest in classical low-compute audio compression techniques. Although modern perceptual codecs and neural audio models such as Opus, SoundStream, EnCodec, and AGC have set new standards for high-fidelity compression, their compute and memory requirements make them unsuitable for microcontroller-grade environments or latency-critical deployments. In such contexts, traditional waveform-based methods—especially μ-law companding and Adaptive Differential Pulse Code Modulation (ADPCM)—continue to offer a reliable balance between computational efficiency and perceptual quality. This project revisits μ-law companding within modern evaluation frameworks to determine its relevance under present-day constraints, focusing on scenarios involving low bit-depth quantization (6–10 bits) at 16 kHz sampling rates where bandwidth, compute, and energy budgets are extremely limited.

We develop a fully reproducible benchmarking pipeline that evaluates four baseline methods: uniform PCM quantization, μ-law PCM, non-adaptive first-order ADPCM, and μ-law–enhanced ADPCM. Using speech-like audio signals, we compute objective audio-quality metrics including global SNR, segmental SNR, STOI, PESQ, and symbol entropy for estimating achievable compressed bitrates. Our findings reveal that μ-law companding provides significant improvements in local intelligibility, particularly reflected in segmental SNR gains of 12–18 dB over uniform PCM at 8-bit quantization, while achieving near-transparent intelligibility (STOI ≈ 1.0). Although μ-law flattens amplitude distributions—resulting in higher entropy and therefore reduced gains from entropy coders—the perceptual enhancements outweigh these bitrate trade-offs for speech-focused applications. Furthermore, we compare the classical pipelines against modern neural latent codecs and demonstrate that while learned models achieve dramatically lower bitrates at high quality, μ-law preprocessing provides no benefits to latent representations.

Overall, this project offers a contemporary technical re-evaluation of μ-law companding and ADPCM, establishing clear guidance on when these classical techniques remain effective. The results highlight that μ-law is best suited for ultra-low-compute speech compression, IoT audio transmission, and low-power real-time systems, whereas music and high-fidelity content favor perceptual or neural codecs. By combining reproducible experimentation, entropy analysis, and rate–distortion evaluation, this work provides a valuable blueprint for deploying efficient and interpretable audio compression pipelines in modern constrained environments.