Journal of Multidisciplinary Engineering
Published by DigitoPub | Digitodontics International Academy
DigitoPub
Open Access & Peer-reviewed
Indexed in: Crossref, ICI, Google Scholar

Codecs in Voice Over IP: A Review and Comparative Analysis of Signal Fidelity

Tahani E. Alqunsul[1] | Ibrahim E. Alwasif[2]

[1]Department of Electrical and Computer Engineering, Engineering College, Nalut University, Jado, Libya.
[2]Department of Electrical and Computer Engineering, Engineering College, Nalut University, Badr, Libya.

ABSTRACT

Purpose: To review the main speech codecs used in Voice over IP (VoIP) and to compare how closely selected codecs reproduce a reference speech signal in the time and frequency domains.

Materials and Methods: Nine codecs (ITU-T G.711, G.722, G.723.1, G.726, G.728, G.729, iLBC, iSAC and Opus) were reviewed. An 11.33 s speech recording and its versions coded with G.711 (A-law and μ-law), G.729, iLBC (8 kHz), iSAC and Opus (48, 24, 16 and 8 kHz) were compared in MATLAB using waveform and spectrum plots, descriptive statistics, SNR, PSNR and MSE.

Results: All coded signals preserved the temporal structure and spectral envelope of the original. PSNR ranged from 23.44 to 23.98 dB and MSE from 0.0040 to 0.0045, and the original-to-coded power ratio (SNR index) from 2.13 dB (Opus 48 kHz) to 5.01 dB (G.729). Opus 48 kHz retained the highest spectral peak (78% of the original) and G.711 μ-law the lowest (43%).

Conclusions: The codecs differed little in signal fidelity, so none can be ranked superior on these measures alone. Selection should rest on bandwidth, delay, packet-loss robustness, licensing and interoperability, ideally confirmed with subjective tests. Results come from a single sample and are indicative only. *(J Multidiscip Eng 2026;1:102-109)*

Keywords: Voice over IP, speech codec, G.711, G.729, Opus, speech quality.

ARTICLE INFO

*Corresponding author:
Tahani E. Alqunsul
Email: t.alqunsul@nu.edu.ly

Received: 23-09-2026
Accepted: 27-10-2026
Published: 29-10-2026

DOI:
https://doi.org/10.26629/jme.2026.13

Copyright: © 2026 by the authors.
Articles published in the *J Multidiscip Eng* are licensed under a Creative Commons Attribution-Noncommercial 4.0 International License.
https://creativecommons.org/licenses/by-nc/4.0/

Audioabstract:

How to cite this article | Alqunsul TE, Alwasif IE. Codecs in voice over IP: A review and comparative analysis of signal fidelity. J Multidiscip Eng. 2026;1:102-109.

1. INTRODUCTION

The rapid growth of the Internet has opened many commercial opportunities for network applications. Internet telephony software able to carry the voice services previously provided by traditional telephone networks first appeared in 1995, and although the voice quality of these early systems was poor, public interest in voice over the Internet Protocol was already considerable.[1] Voice over IP (VoIP) converts the voice signal into IP packets and delivers them across IP networks, which allows a service provider to offer telephony and conventional data services over the same infrastructure; this lowers costs for customers and raises revenue for providers.[1] VoIP also underpins IP telephony, in which higher-level features such as advanced call routing, voicemail and contact centres are built on the VoIP standards.[2] Driven by commercial solutions and network upgrades, VoIP

has grown rapidly and has attracted sustained attention from network-engineering research groups.[1]

The quality of a VoIP call depends on both the network and the end systems. Internet latency ranges from a few milliseconds to much larger values depending on hop count, congestion, and geographic distance, and one-way propagation and queuing delays are generally estimated at 20–100 ms.[1] Packet loss, jitter, out-of-order arrival, available bandwidth, hardware calibration and the choice of codec also affect a call.[3] The codec is particularly important because it determines the capacity of a VoIP system, influences quality-of-service (QoS) parameters such as jitter and delay, and sets the bandwidth requirement.[3] QoS describes the guaranteed performance of information transfer during a call and is governed mainly by latency, jitter and packet loss.[4]

Voice quality is commonly expressed through the Mean Opinion Score (MOS), standardised in ITU-T Recommendation P.800.[5,] In a MOS test, a panel of listeners rates voice samples, and the scores are averaged to give a value between 1 (bad) and 5 (excellent).[6]

A codec (coder–decoder) is the component that makes high-quality voice calls possible without excessive bandwidth. The coder digitises the analogue speech signal and compresses it for transmission, and the decoder reconstructs the signal at the receiver.[3,4] Compression is feasible because, although the human voice can produce frequencies between about 20 Hz and 20 kHz, most of the energy in everyday speech lies between roughly 350 Hz and 4 kHz; narrowband telephony therefore covers only 300–3,400 Hz.[7] Codecs differ in sampling rate, bit rate, coding algorithm, coding delay, required bandwidth and sound quality, and the most widely used are G.711 (A-law and μ-law), G.722, G.723.1, G.726, G.728, G.729, iLBC, iSAC and Opus.[3]

This paper reviews these codecs and compares a representative subset of them experimentally. A speech recording and its versions coded with G.711, G.729, iLBC, iSAC and Opus were analysed in the time and frequency domains in MATLAB, and the deviation of each coded signal from the original was quantified with the signal-to-noise ratio (SNR), peak signal-to-noise ratio (PSNR) and mean squared error (MSE). The aim is to show how closely each codec reproduces the original signal and to guide the selection of a codec that matches the requirements of a VoIP network.

2. MATERIALS AND METHODS

The study combined a literature review of VoIP codecs with an objective signal analysis of one speech recording. No human participants were involved, so ethical approval was not required.

2.1 Literature review of codecs

The characteristics of each codec (developer, bit rate, sampling frequency, frame length, coding algorithm and nominal MOS) were compiled from standards summaries, textbooks, vendor documentation and published comparative studies[8,9] and are summarised in Table 1. No new listening tests were conducted, so the MOS values in Table 1 are nominal values taken from the literature.

2.2 Test signals

The audio samples, in WAV format, were obtained from the Cisco website. They comprise an original speech recording of 11.33 s duration, containing two speech segments separated by a pause (Fig. 1a), and the same recording coded with G.711 A-law and μ-law, G.729, iLBC (8 kHz), iSAC and Opus at 48, 24, 16 and 8 kHz.

2.3 Analysis procedure

The analysis was carried out in MATLAB (version 2022). Each signal was plotted against time and as a magnitude spectrum, and descriptive statistics (minimum, maximum, median and range) were extracted from both representations. Three measures between the original signal x and each coded signal y were computed with the built-in functions snr(x,y), psnr(x,y) and immse(x,y), which are defined in Section 3.4.

The function snr(x,y) treats its second argument as a noise estimate and returns the power ratio, in dB, of the first argument to the second.[10] Because the coded signal was supplied as the second argument, the SNR values reported here are original-to-coded power ratios rather than ratios of signal power to coding-error power, and they equal

0 dB when the two signals are identical. This quantity is referred to below as the SNR index. PSNR and MSE were computed from the sample-by-sample difference between the signals, with PSNR values consistent with a peak value of 1 (normalised full scale).

3. THEORY AND CALCULATION

This section summarises the coding principles and characteristics of the codecs and defines the quality measures used in the analysis.

3.1 Codec classes

VoIP codecs fall into two groups according to their coding mechanism. In the first group, the encoding and decoding principle does not depend on the type of signal being coded; in the second, it is based on modelling the human speech-production apparatus and process. Codecs based on speech modelling do not work well on non-speech signals such as harmonic tones and music, whereas first-group codecs have no such limitation but require a higher bit rate.

3.2 ITU-T voice codecs

G.711. G.711 samples speech at 8,000 samples per second and quantises each sample non-uniformly with 8 bits, giving a 64 kbps PCM stream.[8] Two versions exist: μ-law, used mainly in North America, and A-law, used in most other countries, which differ in their companding characteristics. G.711 is supported by almost all VoIP equipment and adds negligible coding delay, but it offers no compression beyond companding and therefore needs a relatively high bit rate.

G.722. G.722 is a wideband codec covering 50–7,000 Hz that uses sub-band adaptive differential pulse code modulation (SB-ADPCM) at 64 kbps and targets higher-quality speech such as HD voice over VoIP.[8.] It offers a wider audio range and a good MOS, but it is less widely supported than G.711.

G.723.1. G.723.1 compresses speech and other audio components of multimedia services at 5.3 or 6.3 kbps as part of the H.324 family of standards. The higher rate uses MP-MLQ and gives better quality, whereas the lower rate uses CELP-based coding, gives good quality and offers system designers additional flexibility.[9] It is not well suited to music or sound effects and gives lower quality than many codecs at similar data rates.

G.726. G.726 converts a 64 kbps A-law or μ-law PCM channel to and from a 40, 32, 24 or 16 kbps channel using ADPCM transcoding.[8] At 32 kbps, half the rate of G.711, it provides near toll-quality voice and is common in wireless DECT applications, but support outside DECT devices is limited.

G.728. G.728 uses low-delay code-excited linear prediction (LD-CELP) at 16 kbps and 8,000 samples per second.[8.] It is rated toll quality, and its low delay suits satellite, cellular and video-conferencing systems, but few bits are available for error protection.

G.729. G.729 uses CS-ACELP coding to encode voice into 8 kbps streams; the G.729 and G.729 Annex A variants differ mainly in computational complexity, and both give speech quality comparable to that of 32 kbps ADPCM.[9] Its low delay and low bandwidth requirement make it popular in VoIP, although speech quality is marginally reduced and patent licensing has historically been required.

3.3 Other speech coders

iLBC. The Internet Low Bitrate Codec (iLBC) is a free narrowband speech codec designed for robust voice communication over IP. Its payload bit rate is 13.33 kbps with 30 ms frames and 15.20 kbps with 20 ms frames.[9,11] Speech quality degrades gracefully when frames are lost or delayed, as happens with lost or late IP packets; its basic quality is higher than that of G.729A, and it is highly robust to packet loss.[9] iLBC has been adopted by the PacketCable consortium and is used in PC-to-phone applications such as Skype, Google Talk and Yahoo! Messenger with Voice.[9] Its limitations are weaker support in traditional VoIP equipment and the lack of a wideband option.

iSAC. The Internet Speech Audio Codec (iSAC) is a wideband speech codec developed by Global IP Solutions (GIPS), which Google acquired in 2011, for VoIP and streaming audio; its encoded blocks must be encapsulated in a transport protocol such as RTP.[12] It has been used in applications such as AIM Triton, Gizmo5, QQ and Google Talk, and, having been proprietary, became

part of the open-source WebRTC project in June 2011 with a royalty-free license.[12]

Opus. Opus is a lossy audio coding format designed to code both speech and general audio efficiently in a single format while keeping latency low enough for real-time interactive communication.[13] Several listening tests have ranked it above other standard formats, including MP3 and AAC, at any given bit rate up to transparency, and it is used as the VoIP codec in applications such as Discord, WhatsApp and PlayStation.[13] The format accepts input sampling rates of 8, 12, 16, 24 and 48 kHz,[14] and the 8, 16, 24 and 48 kHz settings are examined here. Its main strength is flexibility; its main limitation is weaker support in traditional VoIP equipment, although this is changing.

Table 1 compares the main characteristics of the codecs reviewed.

Table 1: Characteristics of the most common VoIP codecs.

Codec	Developer	Bit rate (kbps)	Sampling (kHz)	Frame (ms)	Coding algorithm	MOS	Remarks
G.711	ITU-T	64	8	10	PCM (companding)	4.1	A-law and μ-law
G.722	ITU-T	64	16	10	SB-ADPCM	4.13	Wideband (50–7,000 Hz)
G.723.1	ITU-T	5.3 / 6.3	8	30	MP-MLQ / ACELP	3.9	Part of H.324
G.726	ITU-T	16 / 24 / 32 / 40	8	5	ADPCM	3.85	Supersedes G.721 and G.723 (ADPCM)
G.728	ITU-T	16	8	5	LD-CELP	3.61	Very low delay
G.729	ITU-T	8	8	10	CS-ACELP	3.92	Algorithmic delay 15 ms
iLBC	GIPS	13.33 / 15.2	8	30 / 20	Frame-independent LPC	4.14	Free; robust to packet loss
iSAC	GIPS (now Google)	10–32	16 / 32	30–60	Channel-adaptive transform coding	3.5	Open source in WebRTC
Opus	IETF (Xiph.Org)	6–510	8–48	2.5–60	SILK + CELT	4–5	Up to 255 channels

G.711 applies only logarithmic companding; the other codecs use lossy compression. MOS values are nominal values compiled from the literature.

3.4 Quality and error measures

Let x(n) be the original signal and y(n) the coded signal, each of N samples. The mean squared error (MSE) is the average of the squared differences between the two signals:

$$\mathrm{MSE} = \frac{1}{N}\sum_{n=1}^{N}[x(n) - y(n)]^2 \qquad (1)$$

The peak signal-to-noise ratio (PSNR) relates the maximum possible signal power to the power of the corrupting noise and is widely used to quantify reconstruction quality after lossy compression:

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{\mathrm{MAX}^2}{\mathrm{MSE}}\right) \qquad (2)$$

where MAX is the peak signal value. The signal-to-noise ratio (SNR) compares the level of the desired signal with that of the noise and is usually expressed in decibels:

$$\mathrm{SNR} = 10\log_{10}\left(\frac{P_{\mathrm{signal}}}{P_{\mathrm{noise}}}\right) \qquad (3)$$

As implemented by snr(x,y) with the coded signal as the second argument (Section 2.3), the SNR index is instead.

$$R = 10\log_{10}\left(\frac{\sum_n x^2(n)}{\sum_n y^2(n)}\right) \qquad (4)$$

which is 0 dB for identical signals and is positive when the coded signal carries less power than the original.

4. RESULTS AND DISCUSSION

4.1 Time-domain analysis

Figure 1 compares the original waveform with the coded waveforms. Every codec preserves the temporal structure of the original: both speech segments, the pause between them and the total duration of 11.33 s are reproduced without visible timing distortion. The differences lie in amplitude.

As Table 2 shows, the coded signals have a peak-to-peak range of 0.4361, about one third smaller than the 0.6511 of the original, with a maximum of 0.2355 and a minimum of −0.2006

against 0.3022 and −0.3489. These extremes are identical for all nine coded signals at the precision reported, so peak amplitude does not discriminate between codecs; the statistic is probably dominated by a few isolated high-amplitude samples that are attenuated by coding. The median amplitude was zero for all signals, as expected for a zero-mean speech waveform, and all signals span the same 0–11.33 s interval.

Table 2: Amplitude statistics of the original and coded signals in the time domain.

Signal	Minimum	Maximum	Range
Original	−0.3489	0.3022	0.6511
G.711 A-law	−0.2006	0.2355	0.4361
G.711 μ-law	−0.2006	0.2355	0.4361
G.729	−0.2006	0.2355	0.4361
iLBC 8 kHz	−0.2006	0.2355	0.4361
iSAC	−0.2006	0.2355	0.4361
Opus 48 kHz	−0.2006	0.2355	0.4361
Opus 24 kHz	−0.2006	0.2355	0.4361
Opus 16 kHz	−0.2006	0.2355	0.4361
Opus 8 kHz	−0.2006	0.2355	0.4361

Median amplitude was 0 for all signals.

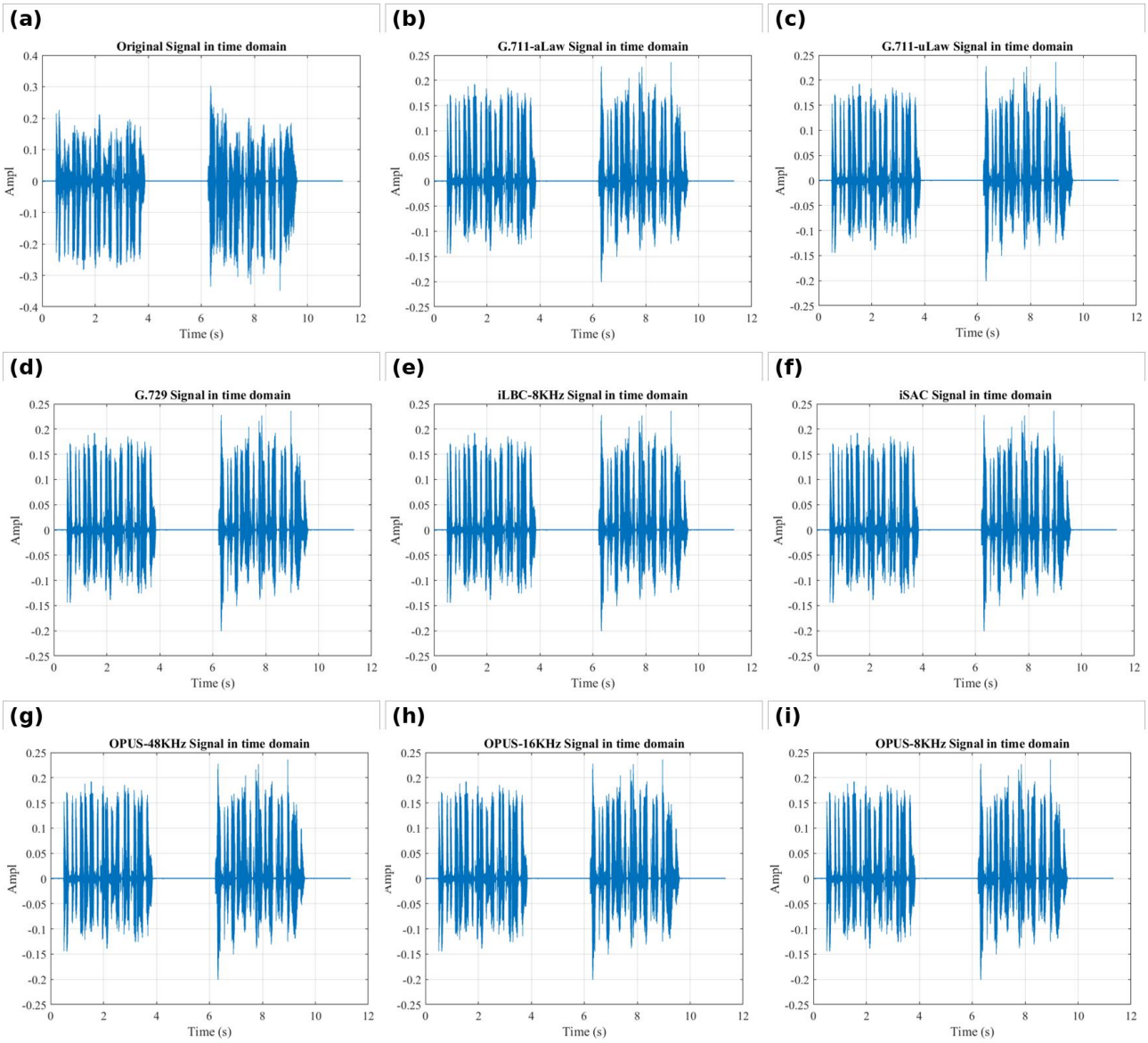


Fig. 1: Time-domain waveforms of (a) the original signal and the signals coded with (b) G.711 A-law, (c) G.711 μ-law, (d) G.729, (e) iLBC 8 kHz, (f) iSAC, (g) Opus 48 kHz, (h) Opus 16 kHz and (i) Opus 8 kHz. Opus 24 kHz is omitted for brevity; its statistics are in Table 2.

4.2 Frequency-domain analysis

Fig. 2 shows the magnitude spectra. The energy of the original is concentrated at low frequencies, with a secondary lobe at about 0.5×10^5 on the plotted frequency axis. This lobe remains clearly visible in iSAC and in Opus at 48 and 16

kHz, but is much smaller or absent in G.711, G.729, iLBC and Opus at 8 kHz, which is consistent with the wider audio bandwidth of the former group.

Table 3 gives the spectral statistics. All spectra span the same frequency-axis range (1 to about 4.998×10^5, and 5.013×10^5 for G.711 μ-law), so the statistics of interest are the magnitudes. Every codec lowers the peak magnitude of the original (1177) to between 505.4 and 923.1, that is, to 43–78% of the original. Opus at 48 kHz retained the most (78%) and G.711 μ-law the least (43%); G.711 A-law and μ-law differ clearly in spectral peak (791.8 against 505.4), although they are identical in Tables 2 and 4. Opus at 48 kHz also has a markedly higher median magnitude (0.277) than the other signals (0.011–0.029), which indicates a higher level of low-magnitude spectral content.

Table 3: Magnitude statistics of the original and coded signals in the frequency domain.

Signal	Minimum	Maximum	Median	Peak (%)
Original	1.339×10^{-5}	1177	0.0101	100
G.711 A-law	1.733×10^{-6}	791.8	0.01588	67
G.711 μ-law	1.387×10^{-5}	505.4	0.0107	43
G.729	6.961×10^{-5}	649.6	0.01312	55
iLBC 8 kHz	9.911×10^{-6}	782.3	0.01604	66
iSAC	1.003×10^{-5}	667.2	0.01066	57
Opus 48 kHz	1.736×10^{-5}	923.1	0.277	78
Opus 24 kHz	1.641×10^{-5}	696.7	0.02864	59
Opus 16 kHz	1.036×10^{-5}	739.7	0.01549	63
Opus 8 kHz	9.097×10^{-5}	747.2	0.01585	63

Peak (%) is the maximum magnitude as a percentage of that of the original signal.

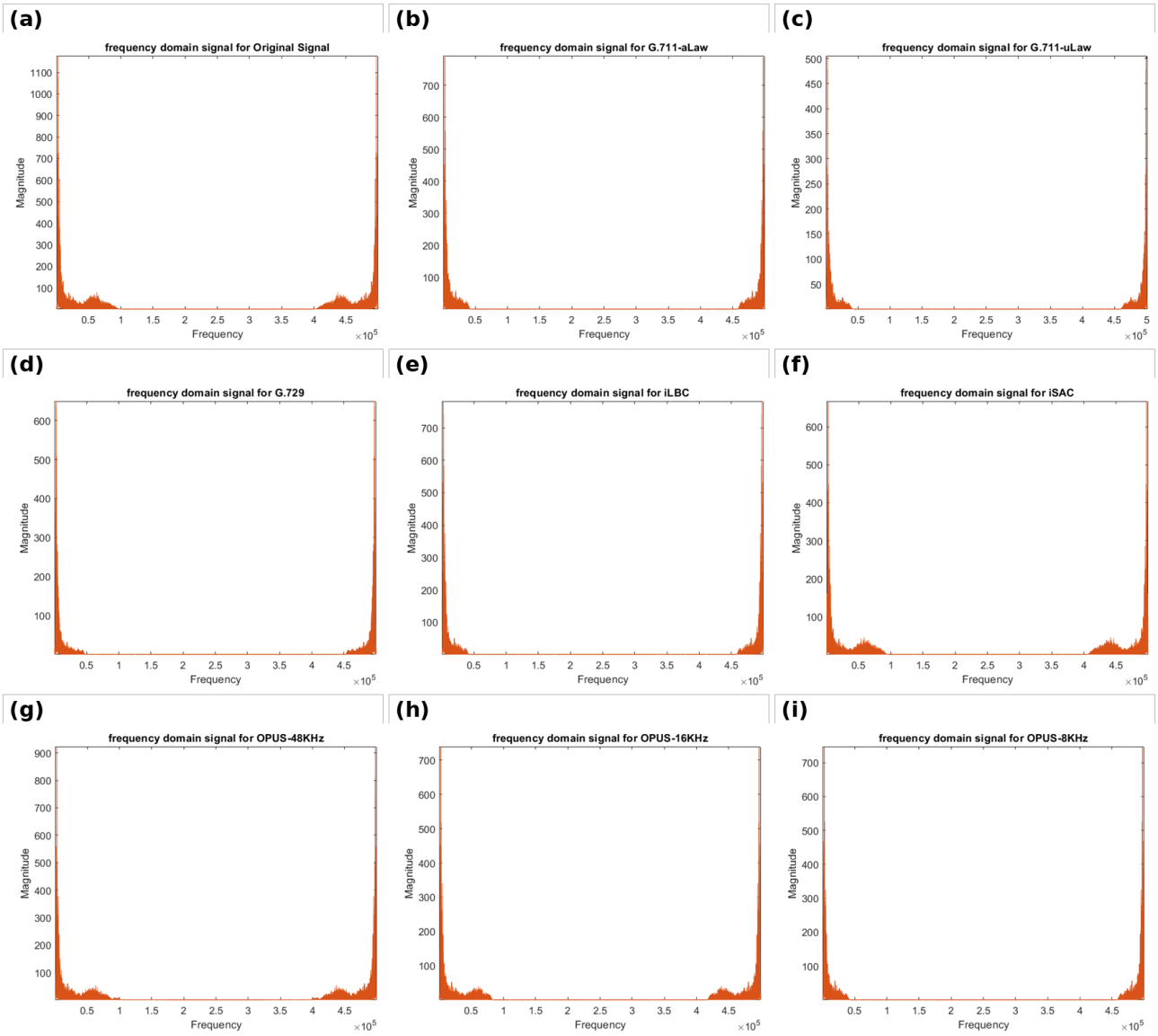


Fig. 2: Frequency-domain magnitude spectra of (a) the original signal and the signals coded with (b) G.711 A-law, (c) G.711 μ-law, (d) G.729, (e) iLBC 8 kHz, (f) iSAC, (g) Opus 48 kHz, (h) Opus 16 kHz and (i) Opus 8 kHz. Opus 24 kHz is omitted for brevity; its statistics are in Table 3.

4.3 Error measures

Table 4 lists the SNR index, PSNR and MSE between each coded signal and the original, which is the reference (SNR index 0 dB, infinite PSNR, zero MSE). Across all codecs, PSNR lies between 23.44 and 23.98 dB and MSE between 0.0040 and 0.0045. The lowest MSE and highest PSNR were obtained with iSAC (23.98 dB) and G.729 (23.96 dB), and the highest MSE was obtained with Opus at 8 kHz, iLBC and Opus at 48 kHz. The spread between best and worst PSNR is only 0.54 dB, and G.711 A-law and μ-law are identical on all three measures, so these measures resolve the codecs only weakly. For Opus, there is no consistent trend with sampling rate.

Table 4: SNR index, PSNR and MSE between the original and coded signals.

Signal	SNR index (dB)	PSNR (dB)	MSE
Original	0.0000	∞	0.0000
G.711 A-law	3.6743	23.5422	0.0044
G.711 μ-law	3.6743	23.5422	0.0044
G.729	5.0085	23.9633	0.0040
iLBC 8 kHz	3.1058	23.4997	0.0045
iSAC	4.5406	23.9823	0.0040
Opus 48 kHz	2.1283	23.4467	0.0045
Opus 24 kHz	3.8941	23.8657	0.0041
Opus 16 kHz	3.6922	23.5501	0.0044
Opus 8 kHz	3.6755	23.4387	0.0045

SNR index: original-to-coded power ratio from snr(x,y); PSNR assumes a peak value of 1.

The SNR index ranges from 2.13 dB (Opus at 48 kHz) to 5.01 dB (G.729). Since $10^{-SNR/10}$ gives the ratio of coded to original power, the coded signals carry roughly 32–61% of the original power, which agrees with the lower amplitude range and spectral peaks in Tables 2 and 3. Opus at 48 kHz loses the least power and G.729 the most.

These figures should not be read as a measure of perceived quality. Lossy speech codecs aim to preserve intelligibility and perceived quality at a low bit rate rather than to reproduce the waveform sample by sample, and MSE-based measures are also sensitive to level and time-alignment differences between signals. The narrow spread of the results therefore shows that the tested codecs reproduce the original to a comparable degree; it does not make any codec better or worse. The codecs differ mainly in how well they meet the requirements of a VoIP network.

4.4 Choosing a codec

When a call is set up, the caller and the called phone negotiate the codec, each using a prioritised list of codecs it supports.[15] The choice should reflect the real-world bandwidth and call volume of the network. A deployment that prefers call quality can place G.722 first and G.711 second, whereas one limited by bandwidth can place G.729 above G.711; because almost all VoIP phones and providers accept G.711, G.722 is the more limited of the two, yet IT professionals favour it for high-grade conversations that do not burden the local network.[15] Table 1 shows the underlying trade-offs: G.729 needs 8 kbps, one-eighth of the 64 kbps of G.711; G.722 doubles the audio bandwidth at the same 64 kbps; iLBC accepts a higher rate of 13.33–15.2 kbps in exchange for robustness to packet loss; and Opus spans 6–510 kbps and 8–48 kHz.

4.5 Limitations

Only one speech recording was analysed; the coding settings of each codec were not controlled in this study, and no subjective listening test, packet-loss or jitter condition was included. The results are therefore indicative and should not be generalised to other speakers, languages or network conditions.

5. CONCLUSION

Capacity and quality of service are the main challenges in VoIP, and the choice of codec strongly affects both. This study reviewed the main ITU-T and other speech codecs used in VoIP and compared a representative subset experimentally using one speech recording. All coded signals retained the temporal structure and spectral envelope of the original, with PSNR values of 23.44–23.98 dB and MSE values of 0.0040–0.0045. Opus at 48 kHz preserved the spectral peak best, while G.729 and iSAC gave the lowest MSE. Because these differences are small, signal fidelity alone does not identify a superior codec, and the codec should be selected to match network requirements such as bandwidth, delay, robustness to packet loss, licensing and interoperability. The findings are limited by the use of a single sample and by the sensitivity of the error measures to level

and alignment, so future work should use several speakers and languages, perceptual quality measures or listening tests, and degraded network conditions.

ACKNOWLEDGEMENT

This work was prepared within the Multimedia module of the Information Technology Pathway (Fall 2022/2023), Department of Electrical Engineering and Computing, The Libyan Academy (Zintan Branch).

AUTHOR CONTRIBUTIONS

All authors contributed equally for: Conceptualization, methodology, writing – original draft, investigation, data curation, validation, manuscript preparation, critical review, writing – review & editing, and final approval of the manuscript.

CONFLICT OF INTEREST

The authors declare no conflicts of interest concerning the research, authorship, and publication of this article.

FUNDING

This research received no external funding.

REFERENCES

1. Chao HC, Chu YM, Tsuei TG. Codec schemes selection for wireless Voice over IP (VoIP). In: Proceedings of the Pacific-Rim Conference on Multimedia. Berlin, Heidelberg: Springer; 2001.
2. Cisco. IP Telephony/Voice over IP (VoIP) [Internet]. 2007 Jun 29. Available from: https://www.cisco.com/c/en/us/tech/voice/ip-telephony-voice-over-IP-voip/index.html
3. Gupta N, Kumar N, Kumar H. Comparative analysis of voice codecs over different environment scenarios in VoIP. In: Proceedings of the 2nd International Conference on Intelligent Computing and Control Systems (ICICCS). IEEE; 2018.
4. Ortega MO, Chafla Altamirano G, Falconi Abad M. Evaluation of the voice quality and QoS in real calls using different voice over IP codecs. In: Proceedings of the 2018 IEEE Colombian Conference on Communications and Computing (COLCOM); 2018 May 16–18; Medellín, Colombia. IEEE; 2018. doi:10.1109/ColComCon.2018.8466727
5. ITU-T. Recommendation P.800: Methods for subjective determination of transmission quality. Geneva, Switzerland: International Telecommunication Union; 1996.
6. Cisco. Cisco Unified Border Element configuration guide: codec basics [Internet]. Available from: https://www.cisco.com/c/en/us/td/docs/ios-xml/ios/voice/cube/configuration/cube-book/cube-codec-basic.pdf
7. 8x8 Staff. What is a codec in VoIP? [Internet]. 8x8; 2022 Jan 8. Available from: https://www.8x8.com/blog/what-is-a-codec-in-voip
8. Speech compression | speech coder [Internet]. Vocal Technologies. Available from: https://vocal.com/speech-coders (accessed 9 Jan 2023).
9. Davidson J, Peters J, Bhatia M, Kalidindi S, Mukherjee S. Voice over IP fundamentals. 2nd ed. Indianapolis, USA: Cisco Press; 2006.
10. MathWorks. snr: Signal-to-noise ratio [Internet]. Available from: https://www.mathworks.com/help/signal/ref/snr.html (accessed 21 Sep 2026).
11. Andersen S, Duric A, Astrom H, Hagen R, Kleijn W, Linden J. Internet Low Bit Rate Codec (iLBC). RFC 3951. Internet Engineering Task Force; 2004 Dec. Available from: https://www.rfc-editor.org/rfc/rfc3951
12. Internet Speech Audio Codec [Internet]. Wikipedia; 2016 Apr 30. Available from: https://en.wikipedia.org/wiki/Internet_Speech_Audio_Codec
13. Opus (audio format) [Internet]. Wikipedia; 2012 Sep 11. Available from: https://en.wikipedia.org/wiki/Opus_(audio_format)
14. Valin JM, Vos K, Terriberry T. Definition of the Opus audio codec. RFC 6716. Internet Engineering Task Force; 2012 Sep. Available from: https://www.rfc-editor.org/rfc/rfc6716
15. Zerby J. What are VoIP codecs and how do they affect call sound quality? [Internet]. Nextiva Blog; 2020 Nov 16. Available from: https://www.nextiva.com/blog/voip-codecs.html