

Bandwidth Extension of Speech Using Perceptual Criteria

Visar Berisha, Steven Sandoval, and Julie Liss
Arizona State University

*SYNTHESIS LECTURES ON ALGORITHMS AND SOFTWARE IN
ENGINEERING #13*



MORGAN & CLAYPOOL PUBLISHERS

ABSTRACT

Bandwidth extension of speech is used in the International Telecommunication Union G.729.1 standard in which the narrowband bitstream is combined with quantized high-band parameters. Although this system produces high-quality wideband speech, the additional bits used to represent the high band can be further reduced. In addition to the algorithm used in the G.729.1 standard, bandwidth extension methods based on spectrum prediction have also been proposed. Although these algorithms do not require additional bits, they perform poorly when the correlation between the low and the high band is weak. In this book, two wideband speech coding algorithms that rely on bandwidth extension are developed. The algorithms operate as wrappers around existing narrowband compression schemes. More specifically, in these algorithms, the low band is encoded using an existing toll-quality narrowband system, whereas the high band is generated using the proposed extension techniques. The first method relies only on transmitted high-band information to generate the wideband speech. The second algorithm uses a constrained minimum mean square error estimator that combines transmitted high-band envelope information with a predictive scheme driven by narrowband features. Both algorithms make use of novel perceptual models based on loudness that determine optimum quantization strategies for wideband recovery and synthesis. Objective and subjective evaluations reveal that the proposed system performs at a lower average bit rate while improving speech quality when compared to other similar algorithms.

KEYWORDS

Speech compression, bandwidth extension, loudness, psychoacoustics

Bibliography

- [1] A. Spanias, "Speech coding: A tutorial review," in *Proc. of IEEE*, vol. 82, no. 10, October 1994. DOI: [10.1109/5.326413](https://doi.org/10.1109/5.326413). 1, 49
- [2] N. Benvenuto, "The 32 kbit/s adpcm coding standard," *ATT Tech. J.*, pp. 12–21, 1986. 1
- [3] G. Frantz and R. Wiggins, "The development of "solid state speech" technology at texas instruments," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Processing*, 1981. DOI: [10.1109/MSP.1981.28418](https://doi.org/10.1109/MSP.1981.28418). 1
- [4] GSM, *GSM Digital Cellular Communication Standards: Enhanced Full-Rate Transcoding*, ETSI/GSM Std. GSM 06.60, 1996. 1
- [5] M. Schroeder and B. Atal, "Code-excited linear prediction(CELP): High-quality speech at very low bit rates," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Processing*, vol. 10, April 1985, pp. 937– 940. DOI: [10.1109/ICASSP.1985.1168147](https://doi.org/10.1109/ICASSP.1985.1168147). 1
- [6] K. Hellwig, P. Vary, D. Massaloux, J. Petit, C. Galand, and M. Rosso, "Speech codec for the european mobile radio system," in *Proc. of Global Telecom. Conf.*, vol. 2, November 1989, pp. 1065–1069. DOI: [10.1109/GLOCOM.1989.64121](https://doi.org/10.1109/GLOCOM.1989.64121). 1
- [7] G. D. Hair and T. W. Rekieta, "Automatic speaker verification using phoneme spectra," *J. Acoust. Soc. Amer.*, vol. 51, no. 1A, pp. 131–131, 1972. DOI: [10.1121/1.1981413](https://doi.org/10.1121/1.1981413). 2
- [8] V. Berisha and A. Spanias, "Wideband speech recovery using psychoacoustic criteria," *EURASIP Journal on Audio, Speech, and Music Processing*, vol. 2007, 2007. DOI: [10.1155/2007/16816](https://doi.org/10.1155/2007/16816). 2, 5
- [9] A. McCree, T. Unno, A. Anandakumar, A. Bernard, and E. Paksoy, "An embedded adaptive multi-rate wideband speech coder," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Processing*, vol. 2, May 2001, pp. 761–764. DOI: [10.1109/ICASSP.2001.941026](https://doi.org/10.1109/ICASSP.2001.941026). 2, 24
- [10] T. Unno and A. McCree, "A robust narrowband to wideband extension system featuring enhanced codebook mapping," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Processing*, Philadelphia, PA, March 2005. DOI: [10.1109/ICASSP.2005.1415236](https://doi.org/10.1109/ICASSP.2005.1415236). 2, 3, 4, 6, 13, 26, 32, 49, 52
- [11] P. Jax and P. Vary, "Enhancement of band-limited speech signals," in *Proc. of Aachen Symposium on Signal Theory*, September 2001, pp. 331–336. DOI: [10.1109/ICASSP.1998.674450](https://doi.org/10.1109/ICASSP.1998.674450). 24, 30, 52

62 BIBLIOGRAPHY

- [12] P. Jax and P. Vary, "Artificial bandwidth extension of speech signals using MMSE estimation based on a hidden Markov model," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Processing*, vol. 1, April 2003, pp. 680–683. DOI: [10.1109/ICASSP.2003.1198872](https://doi.org/10.1109/ICASSP.2003.1198872). 3, 4, 24, 49, 52
- [13] M. Nilsson and W. Kleijn, "Avoiding over-estimation in bandwidth extension of telephony speech," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Processing*, vol. 2, May 2001, pp. 869–872. DOI: [10.1109/ICASSP.2001.941053](https://doi.org/10.1109/ICASSP.2001.941053). 3, 4, 13, 49, 52
- [14] G. Chen and V. Parsa, "HMM-based frequency bandwidth extension for speech enhancement using line spectral frequencies," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Processing*, vol. 1, May 2004, pp. 709–712. DOI: [10.1109/ICASSP.2004.1326084](https://doi.org/10.1109/ICASSP.2004.1326084). 2
- [15] Siyue Chen and Henry Leung, "Speech bandwidth extension by data hiding and phonetic classification," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Processing*, vol. 4, April 2007, pp. 593–596. DOI: [10.1109/ICASSP.2007.366982](https://doi.org/10.1109/ICASSP.2007.366982). 2
- [16] Chen, S. and Leung, H., "Artificial bandwidth extension of telephony speech by data hiding," in *Proc. IEEE Int. Symp. on Circuits and Systems*, May 2005, pp. 3151–3154. DOI: [10.1109/ISCAS.2005.1465296](https://doi.org/10.1109/ISCAS.2005.1465296). 2
- [17] V. Berisha and A. Spanias, "A scalable bandwidth extension algorithm," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Processing*, vol. 4, April 2007, pp. 601–604. DOI: [10.1109/ICASSP.2007.366984](https://doi.org/10.1109/ICASSP.2007.366984). 2
- [18] G. Geiser and P. Vary, "Backwards compatible wideband telephony in mobile networks: CELP watermarking and bandwidth extension," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Processing*, vol. 4, April 2007, pp. 533–536. DOI: [10.1109/ICASSP.2007.366967](https://doi.org/10.1109/ICASSP.2007.366967). 6
- [19] *An 8–32 kbit/s scalable wideband coder bitstream interoperable with G.729*, ITU-T Recommendation G.729.1, 2006. 6, 8, 24, 26, 32, 50, 53
- [20] A. McCree, "A 14 kb/s wideband speech coder with a parametric highband model," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Processing*, vol. 2, 2000. DOI: [10.1109/ICASSP.2000.859169](https://doi.org/10.1109/ICASSP.2000.859169). 2, 24
- [21] P. Jax and P. Vary, "An upper bound on the quality of artificial bandwidth extension of narrowband speech signals," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Processing*, vol. 1, May 2002, pp. 237–240. DOI: [10.1109/ICASSP.2002.5743698](https://doi.org/10.1109/ICASSP.2002.5743698). 2, 4, 24, 49, 53
- [22] M. Nilsson, and S.V. Anderson, and W.B. Kleijn, "On the mutual information between frequency bands in speech," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Processing*, vol. 3, May 2000, pp. 1327–1330. DOI: [10.1109/ICASSP.2000.861823](https://doi.org/10.1109/ICASSP.2000.861823).

- [23] M. Nilsson, and M. Gustafsson, and S.V. Anderson, and W.B. Kleijn, "Gaussian mixture model based mutual information estimation between frequency bands in speech," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Processing*, vol. 1, May 2002, pp. 525–528. DOI: [10.1109/ICASSP.2002.5743770](https://doi.org/10.1109/ICASSP.2002.5743770). 2, 4, 49, 53
- [24] C. F. Chan and W. K. Hui, "Wideband re-synthesis of narrowband CELP coded speech using multiband excitation model," in *Proc. Int. Conf. on Spoken Language Processing*, vol. 1, Philadelphia, PA, October 1996, pp. 322–325. DOI: [10.1109/ICSLP.1996.607118](https://doi.org/10.1109/ICSLP.1996.607118). 3, 49
- [25] V. Berisha and A. Spanias, "Enhancing the quality of coded audio using perceptual criteria," in *Proc. IEEE Wkshp. on Multimedia Signal Processing*, October 2005. DOI: [10.1109/MMSP.2005.248613](https://doi.org/10.1109/MMSP.2005.248613). 3, 35, 36, 49
- [26] V. Berisha and A. Spanias, "Enhancing vocoder performance for music signals," in *Proc. IEEE Int. Symp. on Circuits and Systems*, vol. 4, May 2005, pp. 4050–4053. DOI: [10.1109/ISCAS.2005.1465520](https://doi.org/10.1109/ISCAS.2005.1465520).
- [27] V. Berisha and A. Spanias, "Bandwidth extension of audio based on partial loudness criteria," in *Proc. IEEE Wkshp. on Multimedia Signal Processing*, October 2006. DOI: [10.1109/MMSP.2006.285286](https://doi.org/10.1109/MMSP.2006.285286). 30, 49
- [28] V. Berisha and A. Spanias, SPLIT-BAND Speech Compression Base on Loudness Estimation, ASU, Tempe, Arizona, US 8,392198 B1, 2012. 3
- [29] H. Yasukawa, "Enhancement of telephone speech quality by simple spectrum extrapolation method," in *Proc. of EUROSPEECH*, September 1995, pp. 1545–1548. 3, 11, 13, 51, 52
- [30] H. Yasukawa, "Signal restoration of broad band speech using nonlinear processing," in *Proc. of EUSIPCO*, September 1996, pp. 987–990. 3, 12, 23, 51, 52
- [31] H. Yasukawa, "Wideband Speech Recovery From Bandlimited Speech in Telephone Communications," in *Proc. IEEE Int. Symp. on Circuits and Systems*, no. 4, 1999, pp. 202–205. DOI: [10.1109/ISCAS.1998.698794](https://doi.org/10.1109/ISCAS.1998.698794). 3, 13, 51, 52
- [32] P. Jax and P. Vary, *Audio Bandwidth Extension*. West Sussex, England: Wiley, 2005, ch. Bandwidth Extension for Speech, pp. 171–235. 4, 5, 13, 14, 15, 16, 17, 19, 28, 52
- [33] H. Carl and U. Heute, "Bandwidth enhancement of narrow-band speech signals," in *Proc. of EUSIPCO*, vol. 2, September 1994, pp. 1178–1181. 4, 14, 16, 19, 30, 52
- [34] Y. Yoshida and M. Abe, "An algorithm to reconstruct wideband speech from narrowband speech based on codebook mapping," in *Proc. Int. Conf. on Spoken Language Processing*, 1994, pp. 1591–1594. 4, 16, 52

64 BIBLIOGRAPHY

- [35] Y.M. Cheng and D. O'Shaughnessy and P. Mermelstein, "Statistical recovery of wideband speech from narrowband speech," *IEEE Transactions on Speech and Audio Processing*, vol. 2, no. 4, pp. 544–548, 1994. DOI: [10.1109/89.326637](https://doi.org/10.1109/89.326637). 4, 23, 24, 52
- [36] S. Yao and C. F. Chan, "Block-based bandwidth extension of narrowband speech signals by using CDHMM," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Processing*, Philadelphia, PA, March 2005. DOI: [10.1109/ICASSP.2005.1415233](https://doi.org/10.1109/ICASSP.2005.1415233). 4, 24, 52
- [37] Y. Nakatoh, M. Tsushima, and T. Norimatsu, "Generation of broadband speech from narrowband speech using piecewise linear mapping," in *Proc. of EUROSPEECH*, vol. 3, 1997, pp. 1643–1646. 4, 21, 52
- [38] C. Avendano, and H. Hermansky, and E.A. Wan, "Beyond nyquist: towards the recovery of broad-bandwidth speech from narrow-bandwidth speech," in *Proc. of EUROSPEECH*, vol. 1, September 1995, pp. 165–168. 4, 20, 52
- [39] J. Epps, "Wideband extension of narrowband speech for enhancement and coding," Ph.D. dissertation, The University of New South Wales, 2000. 6, 19, 21, 23, 53
- [40] M. Dietz, and L. Liljeryd, and K. Kjørting, and O. Kunz, "Spectral band replication, a novel approach on audio coding," in *IEEE Aerospace and Electronic Systems*, 2002. 6, 8, 53
- [41] J. S. Garofolo, and L. F. Lamel, and W. M. Fisher, and J. G. Fiscus, and D. S. Pallett, and N. L. Dahlgren, "The DARPA TIMIT acoustic-phonetic continuous speech corpus CD ROM," NTIS order number PB91-100354, Tech. Rep., February 1993. 6
- [42] I. Maddieson, *Patterns of sounds*. Cambridge University Press, 1984. DOI: [10.1017/CBO9780511753459](https://doi.org/10.1017/CBO9780511753459). 7
- [43] P. Kroon and W.B. Kleijn, "Linear prediction-based analysis synthesis coding," in *Speech Coding and Synthesis*, 1995. DOI: [10.1007/978-1-4615-2281-2_3](https://doi.org/10.1007/978-1-4615-2281-2_3). 8, 35, 54
- [44] H. Hermansky, "Perceptual linear predictive (PLP) analysis of speech," *J. Acoust. Soc. Amer.*, vol. 87, no. 4, pp. 1738–1752, April 1990. DOI: [10.1121/1.399423](https://doi.org/10.1121/1.399423). 8, 54
- [45] H. Strube, "Linear prediction on a warped frequency scale," *J. Acoust. Soc. Amer.*, vol. 68, pp. 1071–1076, 1980. DOI: [10.1121/1.384992](https://doi.org/10.1121/1.384992). 8, 35, 54
- [46] *Information Technology-Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbit/sec, IS11172-3: Audio*, ISO/IEC JTC1/SC29/WG11, 1992. 8, 35, 40, 42, 54, 55
- [47] B. C. Moore, *An Introduction to the Psychology of Hearing*, 5th ed. New York: Academic Press, 2003. 8, 35, 43, 44, 54

- [48] B. Moore, B. R. Glasberg, and T. Baer, "A model for the prediction of thresholds, loudness, and partial loudness," *J. Audio Eng. Soc.*, vol. 45, no. 4, 1997. 8, 9, 35, 38, 43, 44, 47, 50, 54, 55
- [49] "The digital theater systems (dts)," web-page: www.dtsonline.com. 8, 35, 42, 55
- [50] G. Davidson, *The Digital Signal Processing Handbook*. New York: CRC Press, 1998, ch. Digital Audio Coding: Dolby AC-3, pp. 41.1–41.21. 8, 35, 42, 55
- [51] T. Painter and A. Spanias, "Perceptual segmentation and component selection for sinusoidal representations of audio," *IEEE Transactions on Speech and Audio Processing*, vol. 13, pp. 139–162, 2005. DOI: [10.1109/TSA.2004.841050](https://doi.org/10.1109/TSA.2004.841050). 9, 42, 55
- [52] V. Atti and A. Spanias, "Speech analysis by estimating perceptually relevant pole locations," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Processing*, vol. 1, March 2005, pp. 217–220. DOI: [10.1109/ICASSP.2005.1415089](https://doi.org/10.1109/ICASSP.2005.1415089). 9, 35, 55
- [53] H. Purnhagen, N. Meine, and B. Edler, "Sinusoidal coding using loudness-based component selection," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Processing*, vol. 2, May 2002, pp. 1817–1820. DOI: [10.1109/ICASSP.2002.5744977](https://doi.org/10.1109/ICASSP.2002.5744977). 9, 55
- [54] L. Laaksonen and J. Kontio and P. Alku, "Artificial bandwidth expansion method to improve intelligibility and quality of AMR-coded narrowband speech," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Processing*, vol. 1, March 2005, pp. 809–812. DOI: [10.1109/ICASSP.2005.1415237](https://doi.org/10.1109/ICASSP.2005.1415237). 12, 23
- [55] J. Paulus, "Variable rate wideband speech coding using perceptually motivated thresholds," in *IEEE Speech Coding Workshop*, 1995, pp. 35–36. DOI: [10.1109/SCFT.1995.658114](https://doi.org/10.1109/SCFT.1995.658114). 14
- [56] C. Erdmann, P. Vary, K. Fischer, W. Xu, M. Marke, T. Fingscheidt, I. Varga, M. Kaindl, C. Quinquis, B. Kovesi, and D. Massaloux, "A candidate proposal for a 3GPP adaptive multi-rate wideband speech codec," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Processing*, vol. 2, May 2001, pp. 757–760. DOI: [10.1109/ICASSP.2001.941025](https://doi.org/10.1109/ICASSP.2001.941025). 27
- [57] *AMR-WB; Transcoding Functions*, 3GPP TS 26.190, 2003. 14, 21, 23, 26, 27, 51
- [58] Y. Linde, A. Buzo, and R. Gray, "An algorithm for vector quantizer design," *IEEE Trans. on Comm.*, vol. 28, pp. 84–95, 1980. DOI: [10.1109/TCOM.1980.1094577](https://doi.org/10.1109/TCOM.1980.1094577). 17, 18
- [59] ITU, "Coding of speech at 8kb/s using conjugate-structure algebraic-code-excited linear-prediction (CS-ACELP)," no. Draft Recommendation G.729, 1995. 18
- [60] J. Conway and N. Sloane, "Voronoi regions of lattices, second moments of polytopes, and quantization," *IEEE Trans. on Inf. Theory*, vol. 28, 1982. DOI: [10.1109/TIT.1982.1056483](https://doi.org/10.1109/TIT.1982.1056483). 19

- [61] S. Chennoukh, A. Gerrits, G. Miet, and R. Sluijter, "Speech enhancement via frequency bandwidth extension using line spectral frequencies," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Processing*, vol. 1, May 2001, pp. 665–668. DOI: [10.1109/ICASSP.2001.940919](https://doi.org/10.1109/ICASSP.2001.940919). 20, 21
- [62] J. Cabral and L. Oliveira, "Pitch-synchronous time-scaling for high-frequency excitation regeneration," in *Proc. of Interspeech*, 2005. 26, 31
- [63] D. Griffin and J. Lim, "Multiband excitation vocoder," *IEEE Transactions on Speech and Audio Processing*, vol. 37, no. 4, pp. 246–275, 1988. DOI: [10.1109/29.1651](https://doi.org/10.1109/29.1651). 27
- [64] R.J. McAulay and T.F. Quatieri, *Speech Coding and Synthesis*. Elsevier, 1995, ch. Sinusoidal Coding, pp. 121–173. 27
- [65] R. McAulay and T. Quatieri, *Speech Coding and Synthesis*. Elsevier, 1995, ch. Time-domain and frequency-domain techniques for prosodic modification of speech, pp. 519–555. 28
- [66] W. Verhelst and M. Roelands, "An overlap-add technique based on waveform similarity (WSOLA) for high quality time-scale modification of speech," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Processing*, vol. 2, April 1993, pp. 554–557. DOI: [10.1109/ICASSP.1993.319366](https://doi.org/10.1109/ICASSP.1993.319366). 28
- [67] U. Kornagel, "Spectral widening of the excitation signal for telephone-band speech enhancement," in *Proc. Int. Wkshp. on Acoust. Echo and Noise Control*, 2001, pp. 215–218. 30
- [68] Y. Qian and P. Kabal, "Combining equalization and estimation for bandwidth extension of narrowband speech," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Processing*, vol. 1, May 2004, pp. 713–716. DOI: [10.1109/ICASSP.2004.1326085](https://doi.org/10.1109/ICASSP.2004.1326085). 32, 34
- [69] V. Atti and A. Spanias, "Rate determination based on perceptual loudness," in *Proc. IEEE Int. Symp. on Circuits and Systems*, vol. 2, May 2005, pp. 848–851. DOI: [10.1109/IS-CAS.2005.1464721](https://doi.org/10.1109/IS-CAS.2005.1464721). 35
- [70] A. Spanias, T. Painter, and V. Atti, *Audio Signal Processing and Coding*. New York: Wiley-Interscience, 2006. 35, 39, 41
- [71] B.C.J. Moore and B.R. Glasberg, "Audibility of time-varying signals in time-varying backgrounds: Model and data," *J. Acoust. Soc. Amer.*, vol. 115, pp. 2603–2603, May 2001. DOI: [10.1121/1.1738839](https://doi.org/10.1121/1.1738839). 35
- [72] B.C.J. Moore and B.R. Glasberg, "A model of loudness applicable to time-varying sounds," *J. Audio Eng. Soc.*, vol. 50, pp. 331–342, May 2002. 35
- [73] E. Zwicker and H. Fastl, *Psychoacoustics*. Springer, New York, 1990. 35, 38, 42, 43

- [74] ITU, “Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs,” no. Recommendation 862, 1995. 36
- [75] B. Paillard, P. Mabillean, S. Morissette, and J. Soumagne, “PERCEVAL: perceptual evaluation of the quality of audio signals,” *J. Acoust. Soc. Amer.*, vol. 40, no. 1-2, pp. 21–31, 1992. 36
- [76] A. Tewfik, M. Swanson, and B. Zhu, “Data embedding in audio: Where do we stand?” in *Proc. IEEE Int. Conf. Acoust., Speech Signal Processing*, March 1999, p. 2075. DOI: [10.1109/ICASSP.1999.758340](https://doi.org/10.1109/ICASSP.1999.758340). 36
- [77] E. Vickers, “Automatic long-term loudness and dynamics matching,” in *Proc. of Audio Eng. Soc. Conv.*, September 2001. 36
- [78] H. Fletcher, “Perceptual segmentation and component selection for sinusoidal representations of audio,” *Revs of Modern Physics*, vol. 12, pp. 47–65, 1940. DOI: [10.1103/RevModPhys.12.47](https://doi.org/10.1103/RevModPhys.12.47). 36, 37, 38
- [79] T. Painter and A. Spanias, “Perceptual coding of digital audio,” in *Proc. of the IEEE*, vol. 88, April 2000, pp. 451–515. DOI: [10.1109/5.842996](https://doi.org/10.1109/5.842996). 38, 39, 40, 41
- [80] B. Moore and B. R. Glasberg, “Formulae describing frequency selectivity as a function of frequency and level, and their use in calculating excitation patterns,” *Hearing Research*, pp. 209–225, 1987. DOI: [10.1016/0378-5955\(87\)90050-5](https://doi.org/10.1016/0378-5955(87)90050-5). 42, 43
- [81] B. Edler and G. Schuller, “Audio coding using a psychoacoustic pre- and post-filter,” in *Proc. IEEE Int. Conf. Acoust., Speech Signal Processing*, 2000, pp. 881–884. DOI: [10.1109/ICASSP.2000.859101](https://doi.org/10.1109/ICASSP.2000.859101). 49
- [82] *AMR-NB; Transcoding Functions*, 3GPP TS 26.090, 2001. 51
- [83] F. Nagel and S. Disch, “A harmonic bandwidth extension method for audio codecs,” in *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*, 2009, pp. 145–148. DOI: [10.1109/ICASSP.2009.4959541](https://doi.org/10.1109/ICASSP.2009.4959541). 57
- [84] F. Nagel, S. Disch, and S. Wilde, “A continuous modulated single sideband bandwidth extension,” in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, 2010, pp. 357–360. DOI: [10.1109/ICASSP.2010.5495843](https://doi.org/10.1109/ICASSP.2010.5495843). 57
- [85] G. Alipoor and M. Savoji, “Wide-band speech coding based on bandwidth extension and sparse linear prediction,” in *Telecommunications and Signal Processing (TSP), 2012 35th International Conference on*, 2012, pp. 454–459. DOI: [10.1109/TSP.2012.6256335](https://doi.org/10.1109/TSP.2012.6256335). 57

- [86] N. Park, Y. Lee, and H. Kim, "Artificial bandwidth extension of narrowband speech signals for the improvement of perceptual speech communication quality," in *Communication and Networking*, ser. Communications in Computer and Information Science, T.-h. Kim, H. Adeli, W.-c. Fang, T. Vasilakos, A. Stoica, C. Paterikakis, G. Zhao, J. Villalba, and Y. Xiao, Eds. Springer Berlin Heidelberg, 2011, vol. 266, pp. 143–153. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-27201-1_17 DOI: 10.1007/978-3-642-27201-1_17. 57
- [87] C. Yağlı, M.A. Tuğtekin Turan, and Engin Erzin, "Artificial bandwidth extension of spectral envelope along a Viterbi path," *Speech Communication*, vol. 55, no. 1, pp. 111 – 118, 2013. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0167639312000957> DOI: 10.1016/j.specom.2012.07.003. 57
- [88] L. Laaksonen and J. Virolainen, "Binaural artificial bandwidth extension (b-abe) for speech," in *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*, 2009, pp. 4009–4012. DOI: 10.1109/ICASSP.2009.4960507. 57
- [89] J. Li, D. Yu, J.-T. Huang, and Y. Gong, "Improving wideband speech recognition using mixed-bandwidth training data in cd-dnn-hmm," in *Spoken Language Technology Workshop (SLT), 2012 IEEE*, 2012, pp. 131–136. DOI: 10.1109/SLT.2012.6424210. 57
- [90] H. Pulakka and P. Alku, "Bandwidth extension of telephone speech using a neural network and a filter bank implementation for highband mel spectrum," *Audio, Speech, and Language Processing, IEEE Transactions*, vol. 19, no. 7, pp. 2170–2183, 2011. DOI: 10.1109/TASL.2011.2118206. 57
- [91] H. Pulakka, U. Remes, S. Yrttiaho, K. Palomaki, M. Kurimo, and P. Alku, "Bandwidth extension of telephone speech to low frequencies using sinusoidal synthesis and a gaussian mixture model," *Audio, Speech, and Language Processing, IEEE Transactions*, vol. 20, no. 8, pp. 2219–2231, 2012. DOI: 10.1109/TASL.2012.2199110. 57
- [92] I. Katsir, D. Malah, and I. Cohen, "Evaluation of a speech bandwidth extension algorithm based on vocal tract shape estimation," in *Acoustic Signal Enhancement; Proceedings of IWAENC 2012; International Workshop on*, 2012, pp. 1–4. 57
- [93] Y. Kwon, Y. Li, and S. Kang, "Bandwidth extension of g.729 speech coder using search-free codebook mapping," in *Telecommunications and Signal Processing (TSP), 2012 35th International Conference on*, 2012, pp. 437–440. DOI: 10.1109/TSP.2012.6256331. 57
- [94] P. Heewan, K. Sangwon, and A. Spanias, "Search-free codebook mapping for artificial bandwidth extension," *IEICE Transactions on Communications*, vol. 95, no. 4, pp. 1479–1482, 2012. DOI: 10.1587/transcom.E95.B.1479. 57

- [95] K.N. Ramamurthy and A.S. Spanias, *MATALB Software for the Code Excited Linear Prediction Algorithm: The Federal Standard-1016*, Morgan and Claypool Publishers, Vol. 2, No. 1, ISBN 1608453847, Jan. 2010 [1](#)
- [96] J.J. Thiagarajan and Andreas Spanias, *Analysis of MPEG-1 Layer III (MP3) Algorithm Using MATLAB*, Morgan and Claypool Publishers, Vol. 3, No. 3, ISBN 10:1608458016, ISBN 13:1608458011, November 2011. [39](#)