

Title (en)
AUDIO ENCODER, AUDIO DECODER, AUDIO ENCODING METHOD AND AUDIO DECODING METHOD FOR DITHERED QUANTIZATION FOR FREQUENCY-DOMAIN SPEECH AND AUDIO CODING

Title (de)
AUDIODECODIERER, AUDIODECODIERER, AUDIODECODIERUNGSVERFAHREN UND AUDIODECODIERUNGSVERFAHREN ZUR GEDITHERTEN QUANTISIERUNG FÜR FREQUENZEBENENSPRACH- UND -AUDIODECODIERUNG

Title (fr)
ENCODEUR AUDIO, DÉCODEUR AUDIO, PROCÉDÉ DE CODAGE AUDIO ET PROCÉDÉ DE DÉCODAGE AUDIO DE QUANTIFICATION JUXTAPOSÉE POUR CODAGE VOCAL ET AUDIO DANS LE DOMAINE FRÉQUENTIEL

Publication
EP 3544005 A1 20190925 (EN)

Application
EP 18187597 A 20180806

Priority
EP 18163459 A 20180322

Abstract (en)
An audio encoder for encoding an audio signal, wherein the audio signal is represented in a spectral domain, is provided. The audio encoder comprises a spectral envelope encoder (110) configured for determining a spectral envelope of the audio signal and for encoding the spectral envelope. Moreover, the audio encoder comprises a spectral sample encoder (120) configured for encoding a plurality of spectral samples of the audio signal. The spectral sample encoder (120) is configured to estimate an estimated bitrate needed for encoding for each spectral sample of one or more spectral samples of the plurality of spectral samples depending on the spectral envelope. Moreover, the spectral sample encoder (120) is configured to encode each spectral sample of the plurality of spectral samples, depending on the estimated bitrate needed for encoding for the one or more spectral samples, according to a first coding rule or according to a second coding rule being different from the first coding rule.

IPC 8 full level
G10L 19/22 (2013.01); **G10L 19/00** (2013.01); **G10L 19/032** (2013.01)

CPC (source: EP)
G10L 19/22 (2013.01); **G10L 19/0017** (2013.01); **G10L 19/032** (2013.01)

Citation (applicant)

- US 7447631 B2 20081104 - TRUMAN MICHAEL MEAD [US], et al
- T. BACKSTROM: "Speech Coding with Code-Excited Linear Prediction", 2017, SPRINGER
- TS 26.445, EVS CODEC DETAILED ALGORITHMIC DESCRIPTION; 3GPP TECHNICAL SPECIFICATION (RELEASE 12), 3GPP, 2014
- TS 26.190, ADAPTIVE MULTI-RATE (AMR-WB) SPEECH CODEC, 3GPP, 2007
- "Unified Speech and Audio Coding", MPEG-D (MPEG AUDIO TECHNOLOGIES, 2012
- M. BOSI ET AL.: "ISO/IEC MPEG-2 advanced audio coding", J. AUDIO ENG. SOC., vol. 45, no. 10, 1997, pages 789 - 814
- J. BENESTY; M. SONDLI; Y. HUANG: "Springer Handbook of Speech Processing", 2008, SPRINGER
- A. ZAHEDI; J. ØSTERGAARD; S. H. JENSEN; S. BECH; P. NAYLOR: "Audio coding in wireless acoustic sensor networks", SIGNAL PROCESS., vol. 107, 2015, pages 141 - 152
- T. BACKSTROM; F. GHIDO; J. FISCHER: "Blind recovery of perceptual models in distributed speech and audio coding", PROC. INTERSPEECH, 2016, pages 2483 - 2487, XP055369017, DOI: doi:10.21437/Interspeech.2016-27
- T. BACKSTROM; J. FISCHER: "Coding of parametric models with randomized quantization in a distributed speech and audio codec", PROC. ITG FACHTAGUNG SPRACHKOMMUNIKATION, 2016, pages 1 - 5
- P. T. BOUFONOS; R. G. BARANIUK: "1-bit compressive sensing", PROC. IEEE INF. SCI. SYST. 42ND ANN. CONF., 2008, pages 16 - 21, XP031282831
- A. MAGNANI; A. GHOSH; R. M. GRAY: "Optimal one-bit quantization", PROC. IEEE DATA COMPRESSION CONF., 2005, pages 270 - 278, XP010782778, DOI: doi:10.1109/DCC.2005.66
- S. VAUDENAY: "Decorrelation: a theory for block cipher security", J. CRYPTOL., vol. 16, no. 4, 2003, pages 249 - 286
- S. SAEEDNIA: "How to make the Hill cipher secure", CRYPTOLOGIA, vol. 24, no. 4, 2000, pages 353 - 360, XP000975779
- C.-C. KUO; W. THONG: "Reduction of quantization error with adaptive Wiener filter in low bit rate coding", PROC. 11TH EUR. IEEE SIGNAL PROCESS. CONF., 2002, pages 1 - 4, XP032754048
- G. E. OIEN; T. A. RAMSTAD: "On the role of Wiener filtering in quantization and DPCM", PROC. IEEE NORWEGIAN SIGNAL PROCESS. SYMP. WORKSHOP, 2001
- J. RISSANEN; G. G. LANGDON: "Arithmetic coding", IBM J. RES.DEVELOP., vol. 23, no. 2, 1979, pages 149 - 162, XP000938669
- J. D. GIBSON; K. SAYOOD: "Lattice quantization", ADV. ELECTRON. ELECTRON PHYS., vol. 72, 1988, pages 259 - 330
- A. GERSHO; R. M. GRAY: "Vector Quantization and Signal Compression", 1992, SPRINGER
- Z. XIONG; A. D. LIVERIS; S. CHENG: "Distributed source coding for sensor networks", IEEE SIGNAL PROCESS. MAG., vol. 21, no. 5, September 2004 (2004-09-01), pages 80 - 94, XP011118155, DOI: doi:10.1109/MSP.2004.1328091
- Z. XIONG; A. D. LIVERIS; Y. YANG: "Handbook on Array Processing and Sensor Networks", 2009, WILEY-IEEE PRESS, article "Distributed source coding", pages: 609 - 643
- M. BOSI; R. E. GOLDBERG: "Introduction to Digital Audio Coding and Standards", 2003, KLUWER
- A. BERTRAND: "Applications and trends in wireless acoustic sensor networks: A signal processing perspective", PROC. 18TH IEEE SYMP. COMMUN. VEH. TECHNOL. BENELUX, 2011, pages 1 - 6, XP032073747, DOI: doi:10.1109/SCVT.2011.6101302
- I. F. AKYILDIZ; T. MELODIA; K. R. CHOWDURY: "Wireless multimedia sensor networks: A survey", IEEE WIRELESS COMMUN., vol. 14, no. 6, December 2007 (2007-12-01), pages 32 - 39, XP011199018
- B. GIROD; A. M. AARON; S. RANE; D. REBOLLO-MONEDERO: "Distributed video coding", PROC. IEEE, vol. 93, no. 1, January 2005 (2005-01-01), pages 71 - 83, XP011123854, DOI: doi:10.1109/JPROC.2004.839619
- F. DE LA HUCHA ARCE; M. MOONEN; M. VERHELST; A. BERTRAND: "Adaptive quantization for multi-channel Wiener filter-based speech enhancement in wireless acoustic sensor networks", WIRELESS COMMUN. MOBILE COMPUT., vol. 2017, 2017
- A. ZAHEDI; J. ØSTERGAARD; S. H. JENSEN; P. NAYLOR; S. BECH: "Coding and enhancement in wireless acoustic sensor networks", PROC. IEEE DATA COMPRESSION CONF., 2015, pages 293 - 302, XP032963977, DOI: doi:10.1109/DCC.2015.20
- A. MAJUMDAR; K. RAMCHANDRAN; L. KOZINTSEV: "Distributed coding for wireless audio sensors", PROC. IEEE WORKSHOP APPL. SIGNAL PROCESS. AUDIO ACOUST., pages 209 - 212
- H. DONG; J. LU; Y. SUN: "Distributed audio coding in wireless sensor networks", PROC. IEEE INT. CONF. COMPUT. INTELL. SECUR., vol. 2, 2006, pages 1695 - 1699, XP031013101

- G. BARRIAC; R. MUDUMBAI; U. MADHOW: "Distributed beamforming for information transfer in sensor networks", PROC. 3RD INT. SYMP. INF. PROCESS. SENS. NETW., 2004, pages 81 - 88, XP058250798, DOI: doi:10.1145/984622.984635
- R. LIENHART; I. KOZINTSEV; S. WEHR; M. YEUNG: "On the importance of exact synchronization for distributed audio signal processing", PROC. IEEE INT. CONF. ACOUST., SPEECH, SIGNAL PROCESS., vol. 4, 2003, pages IV-840 - 3
- O. ROY; M. VETTERLI: "Rate-constrained collaborative noise reduction for wireless hearing aids", IEEE TRANS. SIGNAL PROCESS., vol. 57, no. 2, February 2009 (2009-02-01), pages 645 - 657, XP011238019
- S. BRAY; G. TZANETAKIS: "Distributed audio feature extraction for music", PROC. INT. CONF. MUSIC INF. RETRIEVAL, 2005, pages 434 - 437
- N. RAJPUT; A. A. NANAVALI: "Speech in Mobile and Pervasive Environments", 2012, WILEY, article "Distributed speech recognition", pages: 99 - 114
- D. PEARCE: "Automatic Speech Recognition on Mobile Devices and Over Communication Networks", 2008, SPRINGER, article "Distributed speech recognition standards", pages: 87 - 106
- S. KORSE; T. JAHNEL; T. BACKSTROM: "Entropy coding of spectral envelopes for speech and audio coding using distribution quantization", PROC. INTERSPEECH, 2016, pages 2543 - 2547
- S. DAS; A. CRACIUN; T. JAHNEL; T. BACKSTROM: "Spectral envelope statistics for source modelling in speech enhancement", PROC. ITG FACHTAGUNG SPRACHKOMMUNIKATION, 2016, pages 1 - 5
- G. H. GOLUB; C. F. VAN LOAN: "Matrix Computations", 2004, JOHN HOPKINS UNIV. PRESS
- V. PULKKI: "Spatial sound reproduction with directional audio coding", J. AUDIO ENG. SOC., vol. 55, no. 6, 2007, pages 503 - 516
- A. EDELMAN; N. R. RAO: "Random matrix theory", ACTA NUMERICA, vol. 14, 2005, pages 233 - 297
- D. KNUTH: "The Art of Computer Programming", 1998, ADDISON-WESLEY
- D. E. KNUTH: "The Art of Computer Programming", vol. 2, 2007, ADDISON-WESLEY, article "Seminumerical algorithms"
- P. DIACONIS; M. SHAHSHAHANI: "The subgroup algorithm for generating uniform random variables", PROBAB. ENG. INF. SCI., vol. 1, no. 1, 1987, pages 15 - 32
- G. W. STEWART: "The efficient generation of random orthogonal matrices with an application to condition estimators", SIAM J. NUMER. ANAL., vol. 17, no. 3, 1980, pages 403 - 409
- P. L'ECUYER: "Encyclopedia of Quantitative Finance", 2010, WILEY, article "Pseudorandom number generators"
- M. MATSUMOTO; T. NISHIMURA: "Mersenne twister: A623-dimensionally equidistributed uniform pseudo-random number generator", ACM TRANS. MODEL. COMPUT. SIMUL., vol. 8, no. 1, 1998, pages 3 - 30, XP055244849, DOI: doi:10.1145/272991.272995
- A. JAGANNATHAN: "Mersenne twister-A pseudo random number generator and its variants", 2008, DEPT. ELECT. COMPUT. ENG., GEORGE MASON UNIV.
- T. BACKSTROM; C. R. HELMRICH: "Arithmetic coding of speech and audio spectra using TCX based on linear predictive spectral envelopes", PROC. IEEE INT. CONF. ACOUST., SPEECH, SIGNAL PROCESS., April 2015 (2015-04-01), pages 5127 - 5131, XP033064629, DOI: doi:10.1109/ICASSP.2015.7178948
- J. FISCHER; T. BACKSTROM: "Wiener filtering in distributed speech and audio coding", IEEE SIGNAL PROCESS. LETT., 2017
- T. BACKSTROM: "Estimation of the probability distribution of spectral fine structure in the speech source", PROC. INTERSPEECH, 2017, pages 344 - 348
- J. S. GAROFOLO ET AL.: "TIMIT: Acoustic-Phonetic Continuous Speech Corpus", LINGUISTIC DATA CONSORTIUM, 1993
- "Method for the Subjective Assessment of Intermediate Quality Levels of Coding Systems", ITU-R RECOMMENDATION BS.1534, 2003
- S. NADARAJAH: "A generalized normal distribution", J. APPL. STATIST., vol. 32, no. 7, 2005, pages 685 - 694, XP009156797, DOI: doi:10.1080/02664760500079464
- C. WALCK: "Handbook on Statistical Distributions for Experimentalists", 2007, UNIV. STOCKHOLM
- A. BELA; A. FRIGYIK; M. GUPTA: "Tech. Rep. UWEETR-2010-0006", 2010, DEPT. ELECT. ENG., UNIV. WASHINGTON, article "Introduction to the Dirichlet distribution and related processes"
- M. SCHOEFFLER; F.-R. STOTER; B. EDLER; J. HERRE: "Towards the next generation of web-based experiments: A case study assessing basic audio quality following the ITU-R recommendation BS. 1534 (MUSHRA)", PROC. 1ST WEB AUDIO CONF., 2015
- M. NEUENDORF; M. MULTRUS; N. RETTELBACH; G. FUCHS; J. ROBILLIARD; J. LECOMTE; S. WILDE; S. BAYER; S. DISCH; C. HELMRICH: "The ISO/MPEG unified speech and audio coding standard - consistent high quality for all content types and at all bit rates", JOURNAL OF THE AES, vol. 61, no. 12, 2013, pages 956 - 977, XP040636948
- "Frame error robust narrow-band and wideband embedded variable bit-rate coding of speech and audio from 8-32 kbit/s", ITU-T G.718, 2008
- J. MAKINEN; B. BESSETTE; S. BRUHN; P. OJALA; R. SALAMI; A. TALEB: "AMR-WB+: a new audio coding standard for 3rd generation mobile audio services", PROC. ICASSP, vol. 2, 2005, pages 1109 - 1112, XP010790838, DOI: doi:10.1109/ICASSP.2005.1415603
- M. BOSI; K. BRANDENBURG; S. QUACKENBUSH; L. FIELDER; K. AKAGIRI; H. FUCHS; M. DIETZ; J. HERRE; G. DAVIDSON; Y. OIKAWA: "ISO/IEC MPEG-2 Advanced audio coding", AES CONVENTION, vol. 101, 2012
- G. FUCHS; M. MULTRUS; M. NEUENDORF; R. GEIGER: "Mdct-based coder for highly adaptive speech and audio coding", EUROPEAN SIGNAL PROCESSING CONFERENCE (EUSIPCO 2009, 2009, pages 24 - 28
- I. H. WITTEN; R. M. NEAL; J. G. CLEARY: "Arithmetic coding for data compression", COMMUNICATIONS OF THE ACM, vol. 30, no. 6, 1987, pages 520 - 540, XP000615171, DOI: doi:10.1145/214762.214771
- "Methods for subjective determination of transmission quality", ITU-T RECOMMENDATION P.800, 1996
- G. FUCHS; C. R. HELMRICH; G. MARKOVIC; M. NEUSINGER; E. RAVELLI; T. MORIYA: "Proc. ICASSP", 2015, IEEE, article "Low delay LPC and MDCT-based audio coding in the EVS codec", pages: 5723 - 5727
- S. DISCH; A. NIEDERMEIER; C. R. HELMRICH; C. NEUKAM; K. SCHMIDT; R. GEIGER; J. LECOMTE; F. GHIDO; F. NAGEL; B. EDLER: "Audio Engineering Society Convention", vol. 141, 2016, AUDIO ENGINEERING SOCIETY, article "Intelligent gap filling in perceptual transform coding of audio"
- J. VANDERKOOY; S. P. LIPSHITZ: "Dither in digital audio", JOURNAL OF THE AUDIO ENGINEERING SOCIETY, vol. 35, no. 12, 1987, pages 966 - 975
- R. W. FLOYD; L. STEINBERG: "An adaptive algorithm for spatial gray-scale", PROC. SOC. INF. DISP., vol. 17, 1976, pages 75 - 77
- M. LI; J. KLEJSA; W. B. KLEIJN: "Distribution preserving quantization with dithering and transformation", IEEE SIGNAL PROCESSING LETTERS, vol. 17, no. 12, 2010, pages 1014 - 1017, XP011320162
- T. BACKSTROM: "Enumerative algebraic coding for ACELP", PROC. INTERSPEECH, 2012
- T. BACKSTROM; J. FISCHER: "Fast randomization for distributed low-bitrate coding of speech and audio", IEEE/ACM TRANS. AUDIO, SPEECH, LANG. PROCESS., vol. 26, no. 1, January 2018 (2018-01-01), XP058381855, DOI: doi:10.1109/TASLP.2017.2757601
- J.-M. VALIN; G. MAXWELL; T. B. TERRIBERRY; K. VOS: "Audio Engineering Society Convention", vol. 135, 2013, AUDIO ENGINEERING SOCIETY, article "High-quality, low-delay music coding in the OPUS codec"

Citation (search report)

- [X1] WO 2014161994 A2 20141009 - DOLBY INT AB [NL]
- [X] BACKSTROM TOM ET AL: "Arithmetic coding of speech and audio spectra using tcx based on linear predictive spectral envelopes", 2015 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING (ICASSP), IEEE, 19 April 2015 (2015-04-19), pages 5127 - 5131, XP033064629, DOI: 10.1109/ICASSP.2015.7178948
- [A] TOM BACKSTROM ET AL: "Fast Randomization for Distributed Low-Bitrate Coding of Speech and Audio", IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, IEEE, USA, vol. 26, no. 1, 1 January 2018 (2018-01-01), pages 19 - 30, XP058381855, ISSN: 2329-9290, DOI: 10.1109/TASLP.2017.2757601
- [A] BOUFONOS P T ET AL: "1-Bit compressive sensing", INFORMATION SCIENCES AND SYSTEMS, 2008. CISS 2008. 42ND ANNUAL CONFERENCE ON, IEEE, PISCATAWAY, NJ, USA, 19 March 2008 (2008-03-19), pages 16 - 21, XP031282831, ISBN: 978-1-4244-2246-3

- [T] TOM BÄCKSTRÖM ET AL: "Dithered Quantization for Frequency-Domain Speech and Audio Coding", INTERSPEECH 2018, 1 January 2018 (2018-01-01), ISCA, pages 3533 - 3537, XP055579878, DOI: 10.21437/Interspeech.2018-46

Cited by

CN111681639A; CN116959459A; US2022076365A1; US11727520B2

Designated contracting state (EPC)

AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR

Designated extension state (EPC)

BA ME

DOCDB simple family (publication)

EP 3544005 A1 20190925; EP 3544005 B1 20211215

DOCDB simple family (application)

EP 18187597 A 20180806