



- (51) **International Patent Classification:**
G10L 19/00 (2013.01) *G10L 25/30* (2013.01)
- (21) **International Application Number:**
PCT/EP2023/057107
- (22) **International Filing Date:**
20 March 2023 (20.03.2023)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
22163062.7 18 March 2022 (18.03.2022) EP
22182048.3 29 June 2022 (29.06.2022) EP
- (71) **Applicant:** **FRAUNHOFER-GESELLSCHAFT ZUR FÖRDERUNG DER ANGEWANDTEN FORSCHUNG E.V.** [DE/DE]; Hansastraße 27c, 80686 München (DE).
- (72) **Inventors:** **PIA, Nicola**; c/o Fraunhofer-Institut für Integrierte Schaltungen IIS, Am Wolfsmantel 33, 91058 Erlangen (DE). **GUPTA, Kishan**; c/o Fraunhofer-Institut für Integrierte Schaltungen IIS, Am Wolfsmantel 33, 91058 Erlangen (DE). **KORSE, Srikanth**; c/o Fraunhofer-Institut für Integrierte Schaltungen IIS, Am Wolfsmantel 33, 91058 Erlangen (DE). **MULTRUS, Markus**; c/o Fraunhofer-Institut für Integrierte Schaltungen IIS, Am Wolfsmantel 33, 91058 Erlangen (DE). **FUCHS, Guillaume**; c/o Fraunhofer-Institut für Integrierte Schaltungen IIS, Am Wolfsmantel 33, 91058 Erlangen (DE).
- (74) **Agent:** **BURGER, Markus** et al.; Schoppe, Zimmermann, Stöckeler, Zinkler, Schenk & Partner mbB, Radlkoferstr. 2, 81373 München (DE).
- (81) **Designated States** (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ,

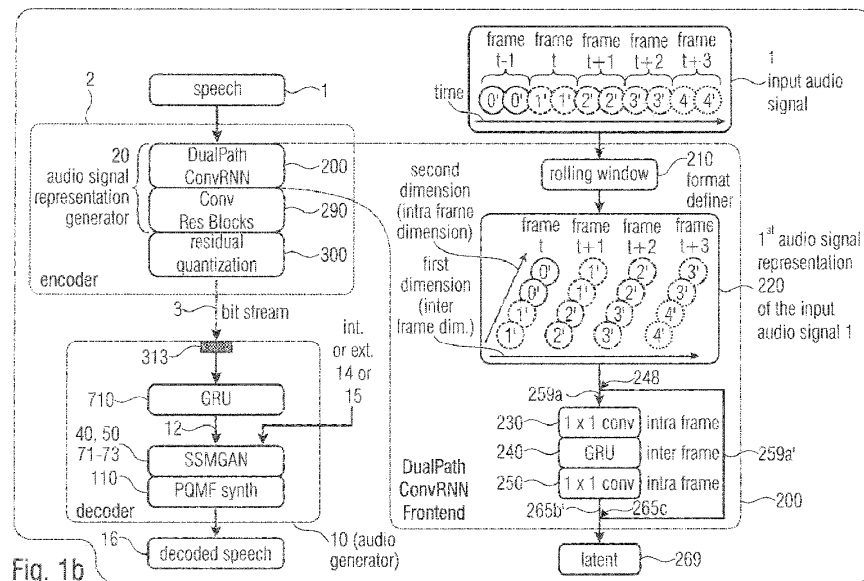
(54) **Title:** VOCODER TECHNIQUES

Fig. 1b

(57) **Abstract:** There is disclosed and audio generator (10), configured to generate an audio signal (16) from a bitstream (3), the bitstream (3) representing the audio signal (16), the audio signal being subdivided in a sequence of frames, the audio generator (10) comprising: a first data provisioner (702) configured to provide, for a given frame, first data (15) derived from an input signal (14); a first processing block (40, 50, 50a-50h), configured, for the given frame, to receive the first data (15) and to output first output data (69) in the given frame, wherein the first processing block (50) comprises: at least one preconditioning learnable layer (710) configured to receive the bitstream (3), or a processed version (112) thereof, and, for the given frame, output target data (12) representing the audio signal (16) in the given frame; at least one conditioning learnable layer (71, 72, 73) configured, for the given frame, to process the target data (12) to obtain conditioning feature parameters (74, 75) for the given frame; and a styling element (77), configured to

[Continued on next page]

CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))

apply the conditioning feature parameters (74, 75) to the first data (15, 59a) or normalized first data (59, 76'), wherein the at least one preconditioning learnable layer (710) includes at least one recurrent learnable layer.