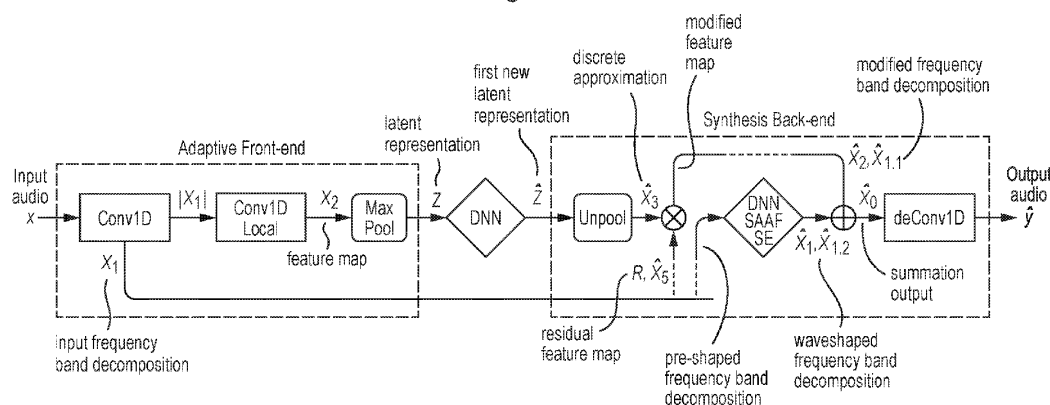




- | | |
|--|--|
| <p>(51) International Patent Classification:
 <i>G10H 1/00</i> (2006.01) <i>G10H 3/18</i> (2006.01)
 <i>G10H 1/12</i> (2006.01) <i>G06N 3/04</i> (2006.01)
 <i>G10L 25/30</i> (2013.01) <i>G06N 3/08</i> (2006.01)</p> | <p>don, Greater London E1 4NS (GB). REISS, Joshua Daniel; Queen Mary University of London, Mile End Road, London, Greater London E1 4NS (GB). BENETOS, Emmanouil; School of EECS, Queen Mary University of London, Mile End Road, London Greater London E1 4NS (GB).</p> |
| <p>(21) International Application Number:
 PCT/GB2020/051150</p> | <p>(74) Agent: J A KEMP LLP; 80 Turnmill Street, London EC1M 3QU (GB).</p> |
| <p>(22) International Filing Date:
 12 May 2020 (12.05.2020)</p> | <p>(81) Designated States (<i>unless otherwise indicated, for every kind of national protection available</i>): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA,</p> |
| <p>(25) Filing Language:
 English</p> | |
| <p>(26) Publication Language:
 English</p> | |
| <p>(71) Applicant: QUEEN MARY UNIVERSITY OF LONDON [GB/GB]; Mile End Road, London, Greater London E1 4NS (GB).</p> | |
| <p>(72) Inventors: MARTINEZ RAMIREZ, Marco Antonio; School of Electronic Engineering and Computer Science, Queen Mary University of London, Mile End Road, London, Greater London E1 4NS (GB).</p> | |

- (54)
- Title:**
- TIME-VARYING AND NONLINEAR AUDIO PROCESSING USING DEEP NEURAL NETWORKS

Fig. 2.0



(S7) Abstract: A computer-implemented method of processing audio data, the method comprising receiving input audio data (x) comprising a time-series of amplitude values; transforming the input audio data (x) into an input frequency band decomposition (X1) of the input audio data (x); transforming the input frequency band decomposition (X1) into a first latent representation (Z); processing the first latent representation (Z) by a first deep neural network to obtain a second latent representation (Z[^], Z1[^]); transforming the second latent representation (Z[^], Z1[^]) to obtain a discrete approximation (X3[^]); element-wise multiplying the discrete approximation (X3[^]) and a residual feature map (R, X5[^]) to obtain a modified feature map, wherein the residual feature map (R, X5[^]) is derived from the input frequency band decomposition (X1); processing a pre-shaped frequency band decomposition by a waveshaping unit to obtain a waveshaped frequency band decomposition (X1[^], X1.2[^]), wherein the pre-shaped frequency band decomposition is derived from the input frequency band decomposition (X1), wherein the waveshaping unit comprises a second deep neural network; summing the waveshaped frequency band decomposition (X1[^], X1.2[^]) and a modified frequency band decomposition (X2[^], X1.1[^]) to obtain a summation output (X0[^]), wherein the modified frequency band decomposition (X2[^], X1.1[^]) is derived from the modified feature map; and transforming the summation output (X0[^]) to obtain target audio data (v[^]).

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, GH, GM, KE, LR, LS, MW, MZ, NA, RW, 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, 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))