US011416742B2

(12) **United States Patent**
Sung et al.

(10) **Patent No.: US 11,416,742 B2**
(45) **Date of Patent: Aug. 16, 2022**

(54) **AUDIO SIGNAL ENCODING METHOD AND APPARATUS AND AUDIO SIGNAL DECODING METHOD AND APPARATUS USING PSYCHOACOUSTIC-BASED WEIGHTED ERROR FUNCTION**

(71) Applicants: **Electronics and Telecommunications Research Institute**, Daejeon (KR); **THE TRUSTEES OF INDIANA UNIVERSITY**, Bloomington, IN (US)

(72) Inventors: **Jongmo Sung**, Daejeon (KR); **Minje Kim**, Bloomington, IN (US); **Aswin Sivaraman**, Bloomington, IN (US); **Kai Zhen**, Bloomington, IN (US)

(73) Assignees: **Electronics and Telecommunications Research Institute**, Daejeon (KR); **THE TRUSTEES OF INDIANA UNIVERSITY**, Bloomington, IN (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 1016 days.

(21) Appl. No.: **16/122,708**

(22) Filed: **Sep. 5, 2018**

(65) **Prior Publication Data**

US 2019/0164052 A1 May 30, 2019

Related U.S. Application Data

(60) Provisional application No. 62/590,488, filed on Nov. 24, 2017.

(30) **Foreign Application Priority Data**

Dec. 15, 2017 (KR) 10-2017-0173405

(51) **Int. Cl.**
G06N 3/08 (2006.01)
G10L 19/008 (2013.01)
(Continued)

(52) **U.S. Cl.**
CPC ***G06N 3/08*** (2013.01); ***G10L 19/008*** (2013.01); ***G10L 19/032*** (2013.01); *G10L 25/30* (2013.01); *G10L 25/69* (2013.01)

(58) **Field of Classification Search**
CPC G06N 3/08
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

6,035,177 A * 3/2000 Moses H04H 20/31 380/252
7,921,007 B2 * 4/2011 Van de Par G10L 21/038 704/229

(Continued)

OTHER PUBLICATIONS

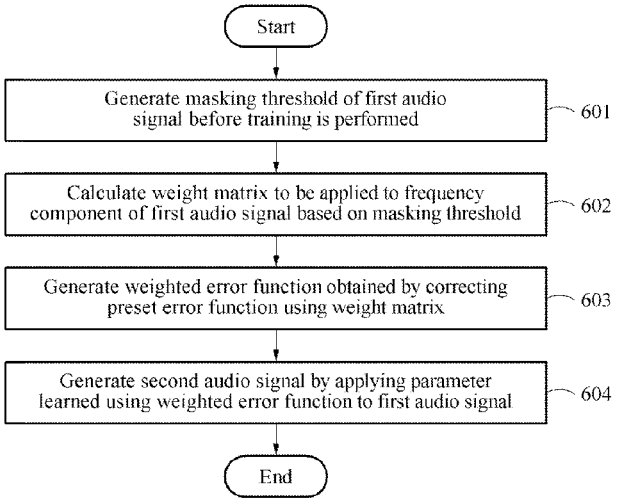
Liu et al., A Perceptually-Weighted Deep Neural Network for Monaural Speech Enhancement in Various Background Noise Conditions, 2017 25th European Signal Processing Conference (EUSIPCO), Aug. 28, 2017, pp. 1-5, IEEE, Kos, Greece.

Primary Examiner — Olisa Anwah
(74) *Attorney, Agent, or Firm* — William Park & Associates Ltd.

(57) **ABSTRACT**

Provided is a training method of a neural network that is applied to an audio signal encoding method using an audio signal encoding apparatus, the training method including generating a masking threshold of a first audio signal before training is performed, calculating a weight matrix to be applied to a frequency component of the first audio signal based on the masking threshold, generating a weighted error function obtained by correcting a preset error function using the weight matrix, and generating a second audio signal by applying a parameter learned using the weighted error function to the first audio signal.

7 Claims, 11 Drawing Sheets



(51) **Int. Cl.**

G10L 19/032	(2013.01)
G10L 25/30	(2013.01)
G10L 25/69	(2013.01)

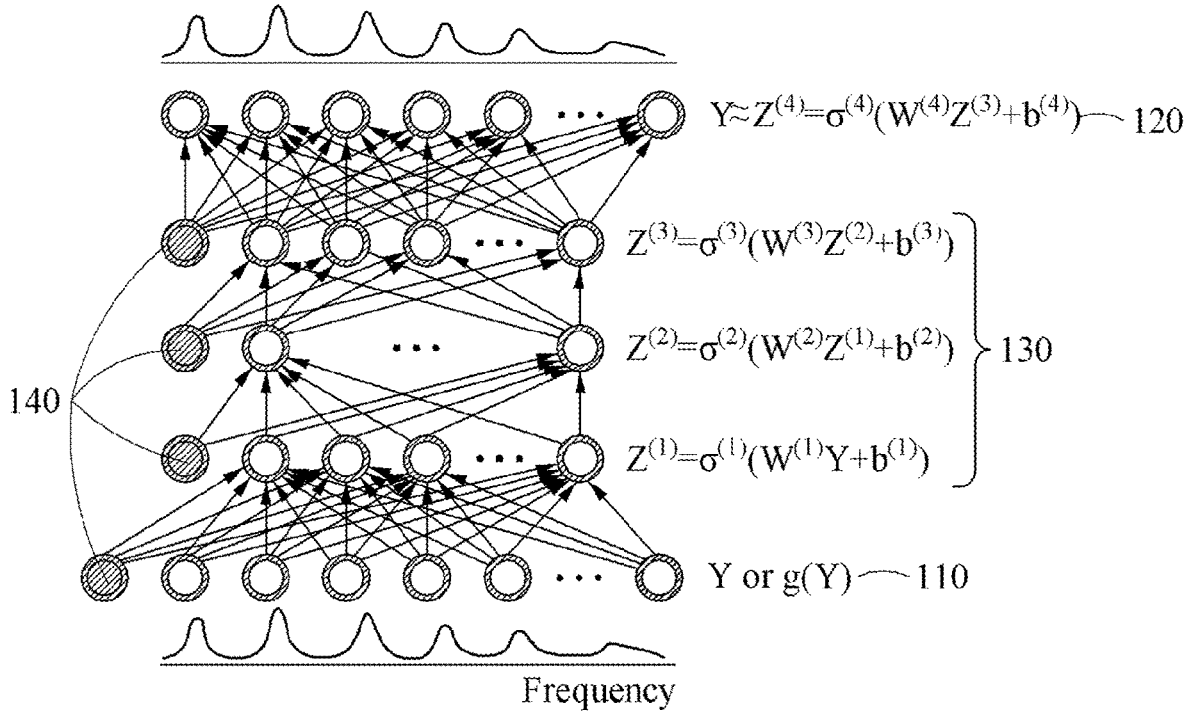
(56) **References Cited**

U.S. PATENT DOCUMENTS

8,588,427	B2 *	11/2013	Uhle	H04R 5/04 381/17
10,043,527	B1 *	8/2018	Gurijala	G10L 19/018
2003/0115042	A1 *	6/2003	Chen	G10L 25/69 704/E19.002
2004/0002859	A1 *	1/2004	Liu	G10L 19/035 704/229
2008/0082328	A1	4/2008	Lee	
2009/0089049	A1 *	4/2009	Moon	G10L 19/035 704/E19.001
2010/0076769	A1	3/2010	Yu	
2011/0026724	A1 *	2/2011	Doclo	G10K 11/17854 381/71.8
2011/0119039	A1 *	5/2011	Grancharov	G10L 25/69 703/2
2012/0020415	A1 *	1/2012	Yang	H04N 17/004 375/E7.279
2012/0023051	A1 *	1/2012	Pishehvar	G06N 3/049 706/21
2012/0093341	A1	4/2012	Kim et al.	
2015/0003625	A1	1/2015	Uhle et al.	
2016/0078880	A1	3/2016	Avendano et al.	
2016/0247502	A1	8/2016	Park et al.	
2020/0234720	A1 *	7/2020	Sung	G10L 19/0017
2020/0329330	A1 *	10/2020	Mitchell	H04R 29/007

* cited by examiner

FIG. 1


Y≈Z(4)=σ(4)(W(4)Z(3)+b(4))
120
Z(3)=σ(3)(W(3)Z(2)+b(3))
Z(2)=σ(2)(W(2)Z(1)+b(2))
130
Z(1)=σ(1)(W(1)Y+b(1))
140
Y or g(Y)
110
Frequency