(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) **EP 4 002 361 A1**

(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:
25.05.2022 Bulletin 2022/21

(21) Application number: **21190382.8**

(22) Date of filing: **09.08.2021**

(51) International Patent Classification (IPC):
G10L 21/0208 *(2013.01)* ***G10L 25/30*** *(2013.01)*
G10L 21/0232 *(2013.01)*

(52) Cooperative Patent Classification (CPC):
G10L 25/30; G10L 21/0208; G10L 21/0232

(84) Designated Contracting States:
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 States:
BA ME
Designated Validation States:
KH MA MD TN

(30) Priority: **23.11.2020 GB 202018375**

(71) Applicants:
- **Rossi, Lisa**
  Cambridgeshire CB3 9AF (GB)
- **Leputa, Mateusz**
  Aberdeenshire AB24 4LY (GB)

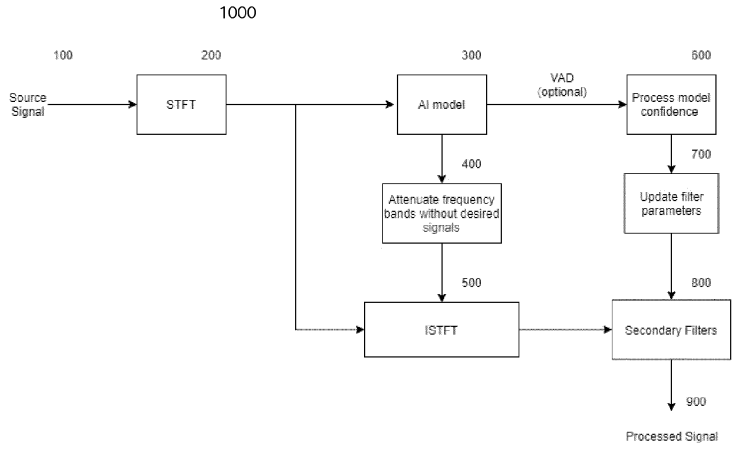
(72) Inventors:
- **Rossi, Lisa**
  Cambridgeshire CB3 9AF (GB)
- **Leputa, Mateusz**
  Aberdeenshire AB24 4LY (GB)

(74) Representative: **ip21 Ltd**
Central Formalities Department
Suite 2
The Old Dairy
Elm Farm Business Park
Wymondham
Norwich, Norfolk NR18 0SW (GB)

(54) AUDIO SIGNAL PROCESSING SYSTEMS AND METHODS

(57) An audio signal detection system comprising a signal processing unit, the system comprising a receiver for receiving an audio signal input having a plurality of audio frequency bands; the signal processing unit comprising at least one signal processing module configured to perform the steps of: processing the audio signal input to provide a frequency domain converted signal input; providing a first signal path comprising the step of employing a neural network for running a machine-learned model to receive the converted signal input and provide a first output representing a primary filter; providing a second signal path to reconstruct the audio signal input; and applying the primary filter to the frequency-domain representation of the audio signal input to provide a signal output.

Figure 1A



EP 4 002 361 A1

(Cont. next page)

Figure 1B

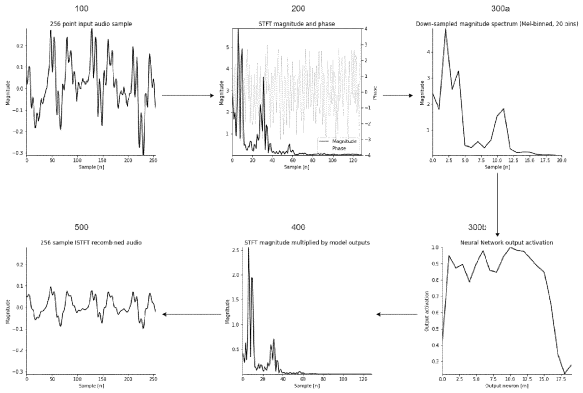


Figure 1C

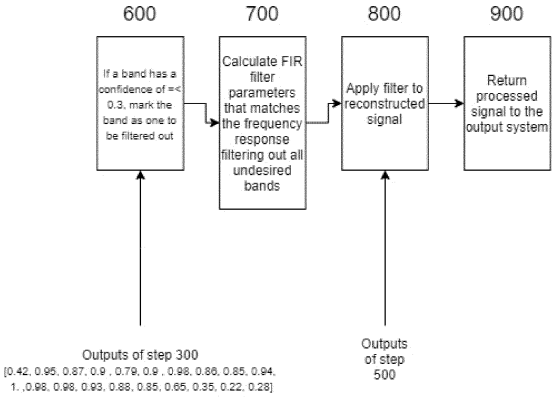


Figure 1D

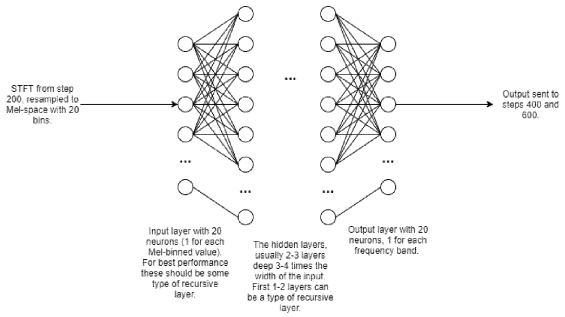
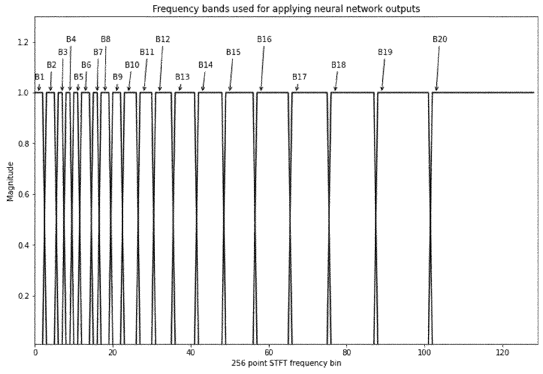


Figure 1E



Description

Technical Field

[0001] Embodiments of the present invention generally relate to audio processing systems and methods, in particular in relation to hearing protection devices.

Background

[0002] According to the World Health Organisation, noise pollution is the second biggest environmental problem affecting health. Prolonged exposure to noise pollution can have detrimental effects on health, such as cardiovascular disease, cognitive impairment, tinnitus and hearing loss. Noise pollution is particularly evident in mining, manufacturing and construction industries.

[0003] Noise protective devices exist and include earplugs, earmuffs, radio-integrated headsets and noise proof panels. However, existing devices offer limited communication and no selective control. For example, in the manufacturing industry, noise levels are particularly high and require the use of common ear protective devices which may hinder the ability for someone to hear another co-worker shouting or asking for help or a safety alarm.

[0004] Aspects of the present invention aim to overcome problems with the existing devices.

Summary

[0005] According to a first, independent aspect of the invention, there is provided an audio signal detection system comprising a signal processing unit, the system comprising a receiver for receiving an audio signal input having a plurality of audio frequency bands;
the signal processing unit comprising at least one signal processing module configured to perform the steps of:

processing the audio signal input to provide a frequency domain converted signal input;
providing a first signal path comprising the step of employing a neural network for running a machine-learned model to receive the converted signal input and provide a first output representing a primary filter;
applying the primary filter to the frequency domain converted signal output; and
providing a second signal path to reconstruct the audio signal input into time domain and to provide a signal output.

[0006] In a dependent aspect, the converted signal input comprises temporal information and the machine-learned model is configured to detect and process temporal information.

[0007] In a dependent aspect, employing a neural network comprises using at least one recursive neural layer.

[0008] In a dependent aspect, processing the audio signal input to provide a frequency domain converted signal input comprises calculating, for each one of the plurality of audio frequency bands, respective magnitude and phase values of a plurality of short-time Fourier transforms, STFT, for each audio frequency band.

[0009] In a further dependent aspect, the step of processing the audio signal input to provide a frequency domain converted signal input further comprises sampling the converted signal input, for example using Mel-frequency binning. This represents a pre-processing step where the STFT magnitude values are resampled into Mel space.

[0010] In a dependent aspect, the first output comprises at least one frequency band identified by the machine-learned model and at least one attenuation value.

[0011] In a further dependent aspect, the magnitude values of a plurality of short-time Fourier transforms are multiplied by a value of the first output.

[0012] In a dependent aspect, reconstructing the audio signal input comprises applying an inverse short-time Fourier transform calculation.

[0013] In a dependent aspect, running a machine-learned model to receive the converted signal input further provides a second output comprising a confidence value.

[0014] In a further dependent aspect, the signal processing module is configured to use the confidence value to identify undesirable audio frequency bands and provide a plurality of filter parameters representing a secondary filter.

[0015] In a further dependent aspect, the step of applying the primary filter to the frequency-domain audio signal input comprises applying the secondary filter to the reconstructed audio signal input.

[0016] In a dependent aspect, the signal processing unit comprises a filter bank module comprising a plurality of filters for receiving the output of the machine—learned model, wherein the at least one audio frequency band is identified based on the stacked output of a power of each one of the plurality of filters.

[0017] In a dependent aspect, the system further comprises an output device configured to receive the audio signal output.

[0018] Advantageously, the system can recognise and identify sounds (e.g. machinery noise) received from two or more microphones for example. This represents a solution for detecting frequency bands of an input audio signal which contain particular audio information.

[0019] Accordingly, a smart selective noise control solution is enabled. In particular, when employed as part of a noise control or selective control device, the system allows users to select the sounds they want to hear or remove. This advantageously leads to more enjoyable user experience as well as improves safety in the working environment.

[0020] For example, devices according to embodiments of the invention can guarantee the user is never

exposed to noise levels above the safety limits, thus, they can automatically lower any sound higher than the prescribed level.

[0021] Advantageously, the system may be implemented on computational devices including mobile devices such as mobile phones. It will be appreciated that the system may be deployed using any existing framework that either already supports a processing unit such as an embedded/lightweight microcontroller or employing any suitable AI method that can be ported to the chosen device.

[0022] Advantageously, all input audio signal may be attenuated to a safe threshold when returned to the users of the system.

[0023] In a dependent aspect, there is provided an audio processing system for real-time noise control and selection for use on portable or wearable devices, the system comprising an audio signal detection system described above.

[0024] In a further dependent aspect, the audio processing system comprises a controller. The controller may be a processor of the type known in the art at hardware level.

[0025] In a further dependent aspect, the audio processing system comprises a user interface for receiving a user input. The input may include an audio signal threshold level. In preferred embodiments, the user input represents a selection to attenuate or/and amplify noise detected in the audio signal input. The user interface may be a hardware or a software interface which allows the user of the system to select the desired functionality of the system. Users can selectively attenuate or/and amplify any noise identified by the software.

[0026] In further dependent aspects, audio processing system further comprises one or more of : at least one microphone, a sound pressure level measurement device, one or more speakers, and a data storage device. In preferred embodiments, mass storage facilitates the collection of data which can then be uploaded to a server location and used to train and improve the algorithms. The mass storage may take a form of a Micro SD peripheral (i.e. memory storage card) included within the device.

[0027] In a dependent aspect, the audio processing system comprises at least one indicator device; for example the indicator devices may comprise a plurality of light-emitting diodes. For example, the hardware version of the user interface including a LED indicator can indicate the functionality of the system currently enabled and/or battery status.

[0028] In a dependent aspect, there is provided a noise protective device comprising an audio processing system as described above. The noise protective device may be integrated in a hearing protection or communication device e.g. headset or earphone etc.

[0029] According to a second, independent aspect of the invention, there is provided an audio signal processing method comprising the steps of:

providing a signal processing unit, the system comprising a at least one signal processing module and a receiver for receiving an audio signal input having a plurality of audio frequency bands;

processing the audio signal input to provide a frequency domain converted signal input;

providing a first signal path comprising the step of employing a neural network for running a machine-learned model to receive the converted signal input and provide a first output representing a primary filter;

providing a second signal path to reconstruct the audio signal input; and

applying the primary filter to the frequency-domain representation of the audio signal input to provide a signal output.

[0030] Aspects of the present invention are now described with reference to the examples shown in the accompanying Figures.

Brief Description of the Drawings

[0031]

Figure 1A is a block diagram of a signal filtering process;

Figure 1B shows an example of signal values corresponding to the process of Figure 1A;

Figure 1C is a block diagram of a secondary filtering method example;

Figure 1D is an example model used at step 300 of Figure 1A;

Figure 1E shows an example of frequency bands used for applying neural network outputs;

Figure 2 illustrates an audio input storage and data collection process;

Figure 3 illustrates hardware components of a system according to a preferred embodiment.

Detailed Description

[0032] With reference to Figure 1A, an audio signal filtering process 1000 is described.

[0033] In a preferred embodiment, an input audio signal 100 representing a raw audio signal comprising a plurality of frequency bands is processed at step 200 using a series of short-time Fourier transforms (STFT) for each frequency band. For example, the frequency bands of an input audio signal 100 can be identified based on a stacked output of the power of each filter bank or the stacked magnitude of a series of short-time Fourier transforms (STFT) for each frequency band.

[0034] The system is provided with the input audio signal 100 from a microphone 1, an example of which is shown in Figure 3. At step 200, the signal is converted to the frequency domain using a STFT and sent via two signal paths to processing steps 300 and 500.

[0035] Figure 1B shows an example of signal values corresponding to the process of Figure 1A. In this example, the input audio signal is sampled with 256 audio signal input values corresponding to 256 frequency bands. The STFT magnitude and phase value are also shown on the signal value plot corresponding to step 200 shown in Figure 1B.

[0036] Referring back to Figure 1A, at step 300 a neural network is employed to run an AI model for detecting frequency bands of the input audio signal which contain desired information. It will be appreciated that a number of AI models are suitable. In a preferred embodiment, the AI model has a capability to detect and maintain temporal information about a series of inputs and as long as these inputs can be provided and processed by the AI model fast enough to provide the output in real-time.

[0037] Accordingly, the AI model may receive as input a down-sampled frequency domain input. The input may be down sampled using known methods such as Mel-frequency binning to a fixed number of bins. The AI model outputs the attenuation value (i.e. how loud) each frequency band should be in the reconstructed signal.

[0038] As shown in Figure 1B, in an example, step 300 is split into two steps, 300a and 300b. At step 300a, which represents a pre-processing step, the STFT magnitude is resampled into Mel-space, using 20 bins. The result of the down-sampled spectrum after step 300a is shown in the plot of Figure 1B. At step 300b, the neural network acts to provide the AI model output In this example the phase is not processed by the AI model, but it is used unaltered at step 500 (via the second signal path) to reconstruct the audio signal e.g. via using an inverse short-time Fourier transform (ISTFT).

[0039] Figure 1D illustrates an example model used in step 300 to process the Mel-binned (resampled) STFT signal and produce the confidence values for 20 frequency bands. Each output corresponds to a frequency band on the STFT magnitude shown in Figure 1D.

[0040] In this example, an input layer with 20 neurons, one neuron for each Mel-binned value or frequency band, is provided. For optimum performance, the layer may be a recursive layer. Hidden layers, usually 2-3 layers deep are provided with 3-4 times the width of the input. Typically, the first 1-2 layers are a type of recursive layer. The output layer also has 20 neurons, one neuron for each Mel-binned value or frequency band. The output of step 300 is then sent to steps 400 and 600, respectively.

[0041] Figure 1E shows example of frequency bands that each neurons output is applied to. For example, all values within the B1 frequency band envelope are multiplied by the first output neuron, all within B2 are multiplied by second neuron output and so on. These envelopes may be used for down sampling the signal from the 129 magnitude values that result form 256-point STFT down to the 20 bins required by the neural network input.

[0042] At step 400, the frequency gains are applied on the frequency domain representation of the original signal via simple multiplication of each respective frequency band by its respective gain. The plot in Figure 1B corresponding to step 400 shows an example of STFT magnitude values multiplied by the AI model outputs received from step 300 (or 300a and 300b).

[0043] It will be appreciated that gain may correspond to a range of frequencies. The frequency ranges are chosen at system implementation time depending on the requirements and usually match up with the down-sampling at step 200.

[0044] At step 500, the signal is reconstructed into time domain in this example using an inverse short-time Fourier transform (ISTFT).

[0045] Accordingly, the AI model output is used to identify desired signal components in the output signal. At step 600, it is decided if such activity has been detected with a high confidence by measuring the magnitudes of the outputs (for example, if all outputs are close to or above 1.0, it is decided that the model detects useful information in all bands). At steps 700 and 800, a finite or infinite input response (FIR/IIR) filter is provided , the filter being dynamically updated to remove any remaining noise in the signal to enhance the overall output. The processed signal (output) is provided to an output system at step 900.

[0046] Figure 1 C shows an example implementation of secondary filtering methods, spanning steps 600 to 900. It will be appreciated that there are several possible implementations of the processing path 600-900. In this example, the neural network outputs and a simple thresholding mechanism are used to determine whether each band contains useful signals or not (at step 600). In this example, at step 600 it is determined if a frequency band has a confidence level of less than or equal to 0.3, if so, the frequency band is identified as one to be filtered out (undesired frequency bands). At step 700, a filter matching the desired frequency response is provided, using any desired method such as Wiener filter design, lest square filter etc. For example FIR filter parameters are calculated that match the frequency response to filter out all of the frequency bands identified to be filtered out.

[0047] The constructed filter from step 700 is then applied at step 800 to the reconstructed signal from step 500 and the filtered signal is output at step 900.

[0048] The process 1000 of Figure 1A can advantageously identify bands in the frequency spectrum of the input signal 100 which contain useful audio information such as voice or any desired audio signal. The identified frequency bands which are deemed to be useful are then kept whilst the frequency bands containing information that is deemed to be noise are discarded via the means of attenuating their respective frequency bands in the frequency domain representation.

[0049] This approach is advantageous because it allows for live filtering of the input audio signal with very little knowledge of the exact spectral densities of the desired signal. The approach is also lightweight enough to work on embedded hardware such as microcontrollers

of Field Programmable Gate Arrays (FPGAs). The process allows for attenuation of sporadic noises while most common modern approaches can only reliably cancel out continuous noises.

[0050] The output of the neural network 300 can also be used as a reliable voice activity detection (VAD) which can then be used in conjunction with more traditional filtering such as using a Weiner filter to enhance the quality of the processed sound.

[0051] Advantages of the filtering process 1000 of Figure 1A over the prior art include but are not limited to:

- Explicit detection of frequency bands containing voice or desired signal.
- Voice activity detection and/or desired signal detection.
- Ability to employ any filtering method in tandem with the AI application.
- Small network capable of running on mobile devices in real-time.
- Ability to attenuate sporadic noises.
- Separation of the desired signals from the source signal.

[0052] Figure 2 illustrates an audio input storage and data collection process 2000 for the device. A storage unit 21 is present on the device that may be in communication with a computer when the device is charging. While the device is charging, the encrypted audio data is uploaded to a server while software updates are downloaded to the device. The recorded audio data can be accessed for monitoring purposes and to improve and retrain the models.

[0053] For example, audio samples are recorded onto the device storage unit 21 during daily operation. When not in use the device may be in communication with an on-site server 22 and upload the audio data to the server. The on-site server 22 is configured in this example to select most distinct audio samples and upload them to an off-site server or cloud server 23. The off-site/cloud server 23 integrates the new data into a training dataset and may employ additional AI models to train with the new dataset 24.. The retrained model is then sent to be dispatched back to the devices deployed (i.e. via devices 23 , 22, and 21 in sequence).

[0054] Accordingly, mass storage facilitates the collection of data which can be uploaded to a server location and used to train and improve the algorithms.

[0055] Figure 6 illustrates an example embodiment of a system 3000 representing a hearing protection headset. In this example, the system comprises two microphones 1 for audio sampling, although it would be appreciated that the number of microphones may vary. Preferably, the system can further comprise a sound pressure level measurement device (not shown). Speakers 2 are provided for replaying the processed audio in real-time. The number of speakers may vary.

[0056] The system further comprises a controller unit 3 comprising a controller (i.e. a processor) and other electronic peripherals required to operate the system such as memory storage unit. For example, the mass storage may take a form of a Micro SD card included within the device. A LED indicator 4 is also provided to indicate for example when the headset 3000 is gathering audio samples.

Claims

1. An audio signal detection system comprising a signal processing unit, the system comprising a receiver for receiving an audio signal input having a plurality of audio frequency bands;
   the signal processing unit comprising at least one signal processing module configured to perform the steps of:

   processing the audio signal input to provide a frequency domain converted signal input;
   providing a first signal path comprising the step of employing a neural network for running a machine-learned model to receive the converted signal input and provide a first output representing a primary filter;
   providing a second signal path to reconstruct the audio signal input; and
   applying the primary filter to the frequency-domain representation of the audio signal input to provide a signal output.

2. A system according to claim 1, wherein the converted signal input comprises temporal information and the machine-learned model is configured to detect and process temporal information.

3. A system according to claim 1 or claim 2, wherein employing a neural network comprises using at least one recursive neural layer.

4. A system according to any one of the preceding claims, wherein processing the audio signal input to provide a frequency domain converted signal input comprises calculating, for each one of the plurality of audio frequency bands, respective magnitude and phase values of a plurality of short-time Fourier transforms, STFT, for each audio frequency band.

5. A system according to any one of the preceding claims, wherein the step of processing the audio signal input to provide a frequency domain converted signal input further comprises sampling the converted signal input.

6. A system according to any one of the preceding claims, wherein the first output comprises at least one frequency band identified by the machine-

learned model and at least one corresponding attenuation value.

7. A system according to claim 6, wherein the magnitude values of a plurality of short-time Fourier transforms are multiplied by the respective attenuation values.

8. A system according to any one of the preceding claims, wherein reconstructing the audio signal input comprises making an inverse short-time Fourier transform calculation.

9. A system according to any one of the preceding claims, wherein running the machine-learned model further provides a second output comprising a confidence value.

10. A system according to claim 9, wherein the signal processing module is configured to use the confidence value to identify undesirable audio frequency bands and provide a plurality of filter parameters representing a secondary filter.

11. A system according to any one of the preceding claims, wherein the step of applying the primary filter to the frequency-domain representation of the audio signal input comprises applying the secondary filter to the reconstructed audio signal input.

12. A system according to any one of the preceding claims, wherein the signal processing unit comprises a filter bank module comprising a plurality of filters for receiving the output of the machine—learned model, wherein the at least one audio frequency band is identified based on the stacked output of a power of each one of the plurality of filters.

13. An audio processing system for real-time noise control and selection for use on portable or wearable devices, the system comprising an audio signal detection system according to any one of the preceding claims.

14. A noise protective device comprising an audio processing system according to any one of claims 1 to 12.

15. An audio signal processing method comprising the steps of:

providing a signal processing unit, the system comprising a at least one signal processing module and a receiver for receiving an audio signal input having a plurality of audio frequency bands;
processing the audio signal input to provide a frequency domain converted signal input;
providing a first signal path comprising the step of employing a neural network for running a machine-learned model to receive the converted signal input and provide a first output representing a primary filter;
providing a second signal path to reconstruct the audio signal input; and
applying the primary filter to the frequency-domain representation of the audio signal input to provide a signal output.

Figure 1A

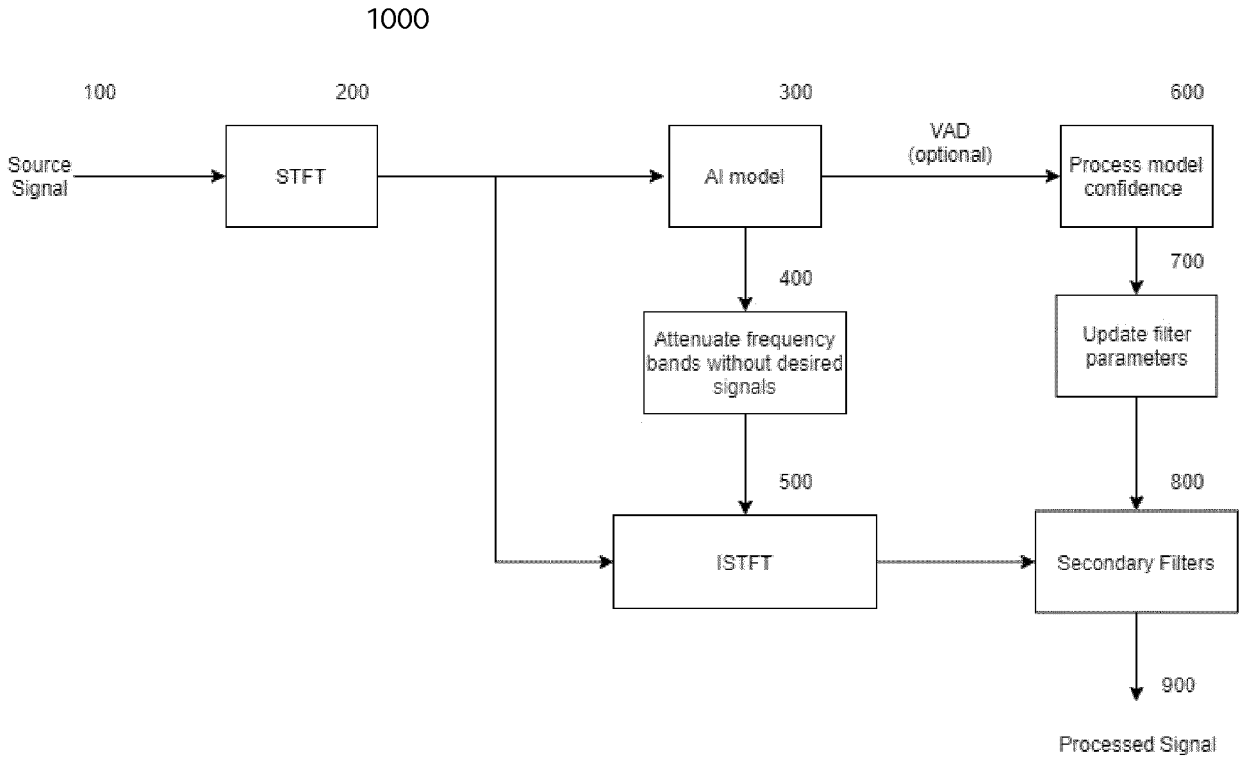

1000
100
200
300
600
Source Signal
STFT
AI model
VAD (optional)
Process model confidence
400
700
Attenuate frequency bands without desired signals
Update filter parameters
500
800
ISTFT
Secondary Filters
900
Processed Signal

Figure 1B

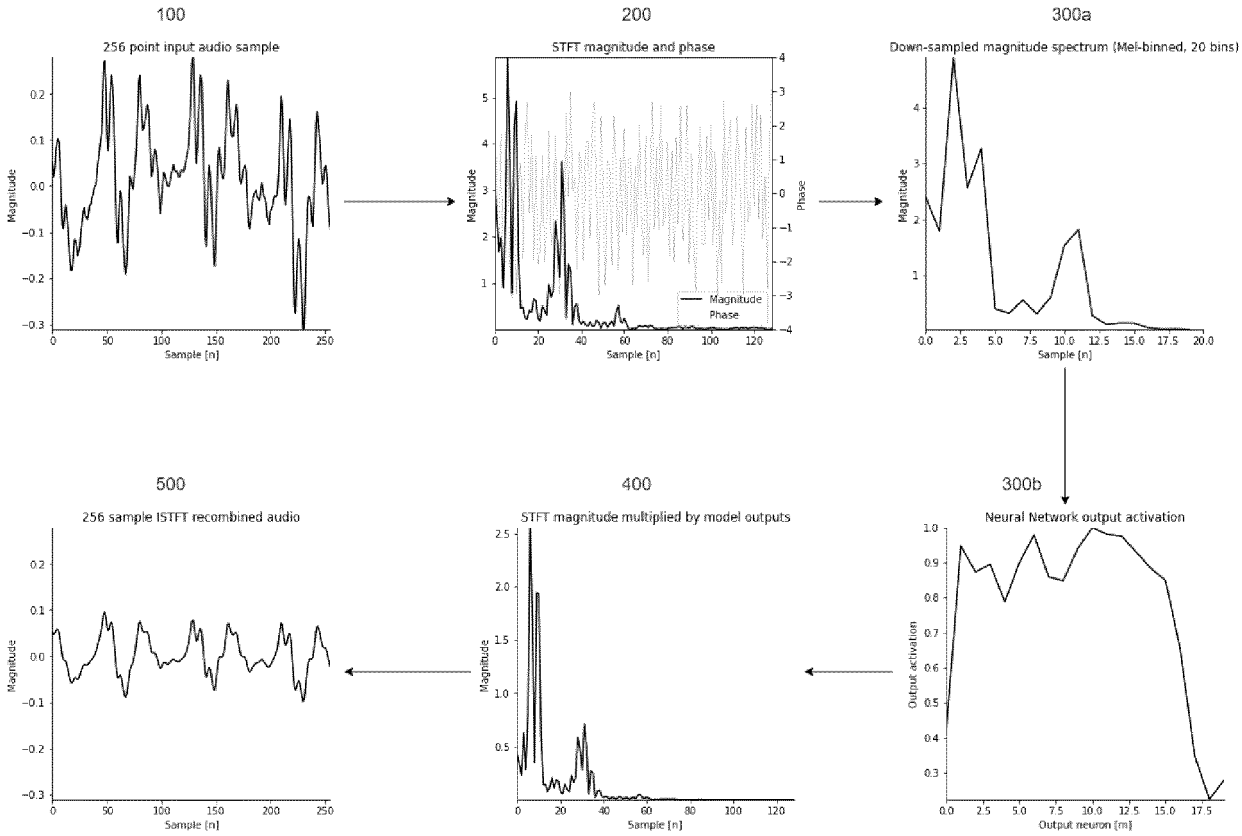

100
256 point input audio sample
Magnitude
Sample [n]
200
STFT magnitude and phase
Magnitude
Phase
Sample [n]
300a
Down-sampled magnitude spectrum (Mel-binned, 20 bins)
Magnitude
Sample [n]
300b
Neural Network output activation
Output activation
Output neuron [m]
400
STFT magnitude multiplied by model outputs
Magnitude
Sample [n]
500
256 sample ISTFT recombined audio
Magnitude
Sample [n]

Figure 1C

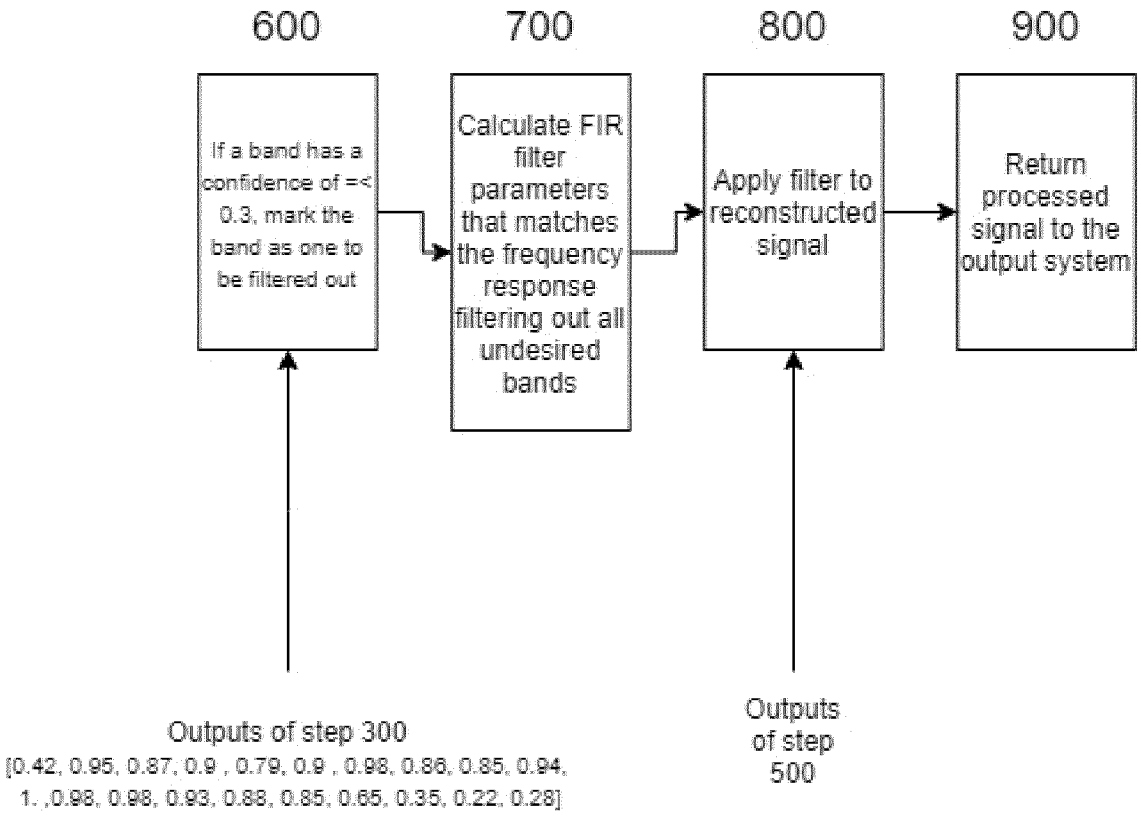

600
700
800
900
If a band has a confidence of =< 0.3, mark the band as one to be filtered out
Calculate FIR filter parameters that matches the frequency response filtering out all undesired bands
Apply filter to reconstructed signal
Return processed signal to the output system
Outputs of step 300
[0.42, 0.95, 0.87, 0.9 , 0.79, 0.9 , 0.98, 0.86, 0.85, 0.94, 1. ,0.98, 0.98, 0.93, 0.88, 0.85, 0.65, 0.35, 0.22, 0.28]
Outputs of step 500

Figure 1D

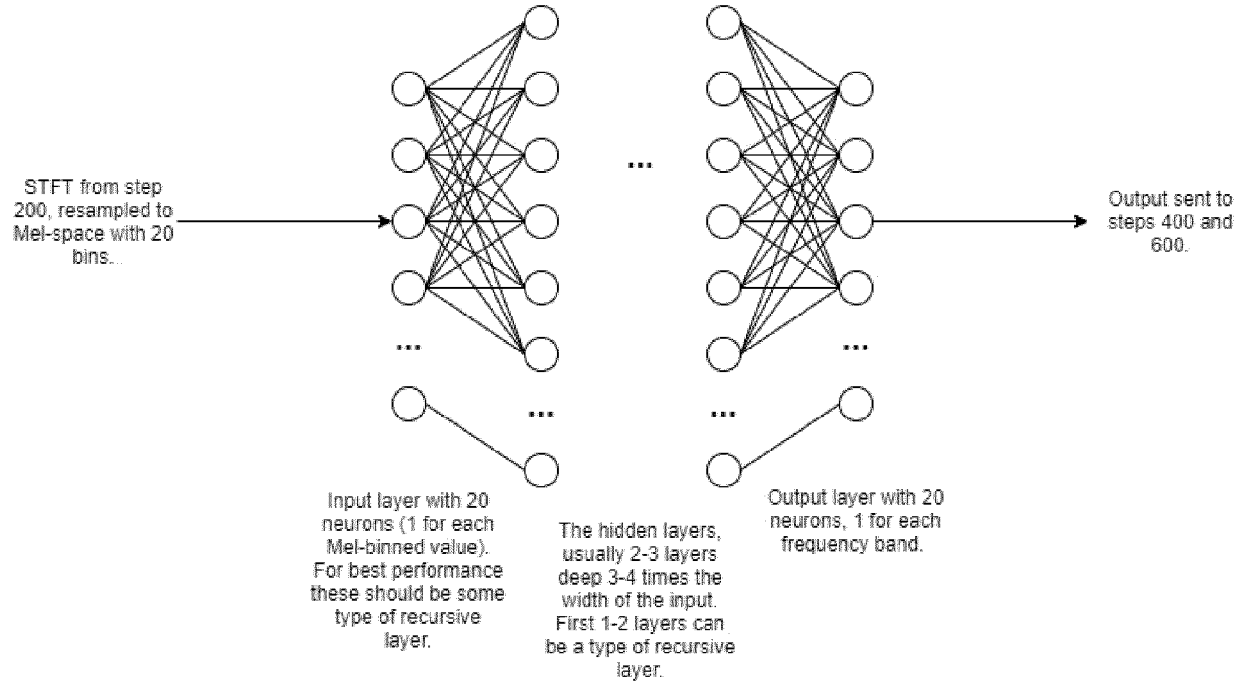


Figure 1E

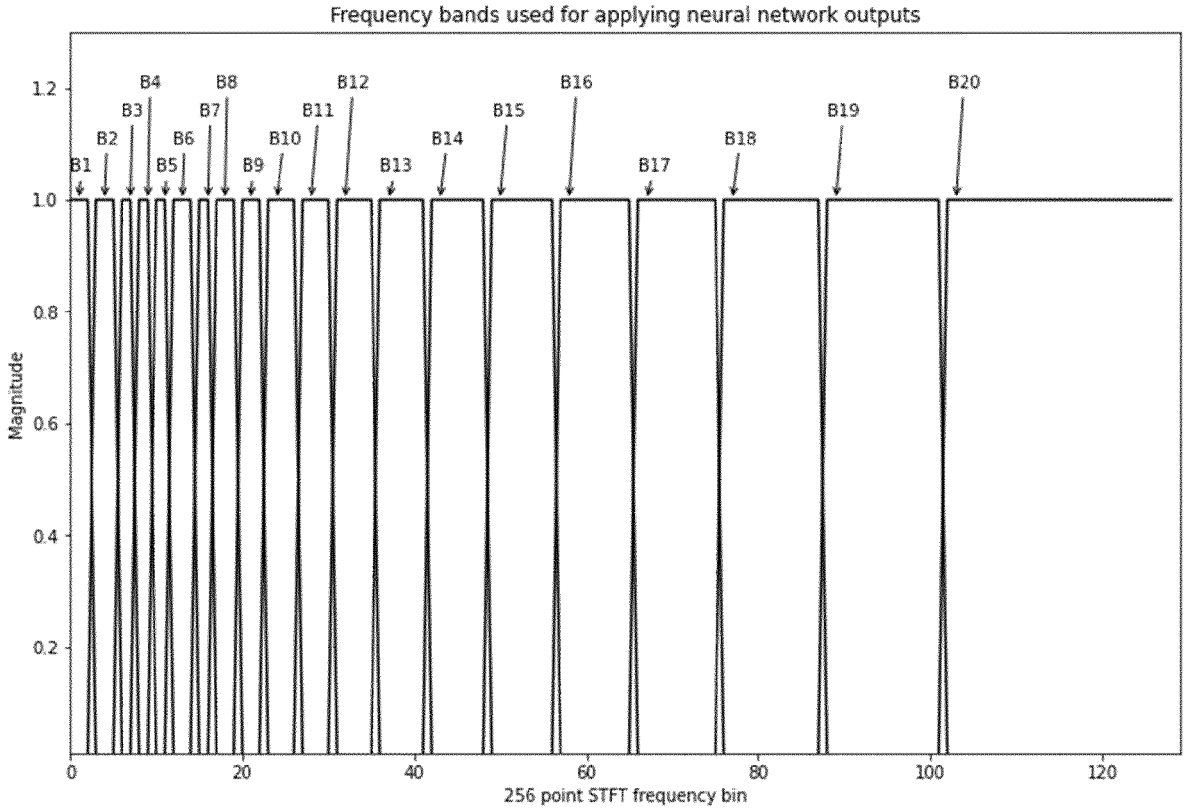


Figure 2

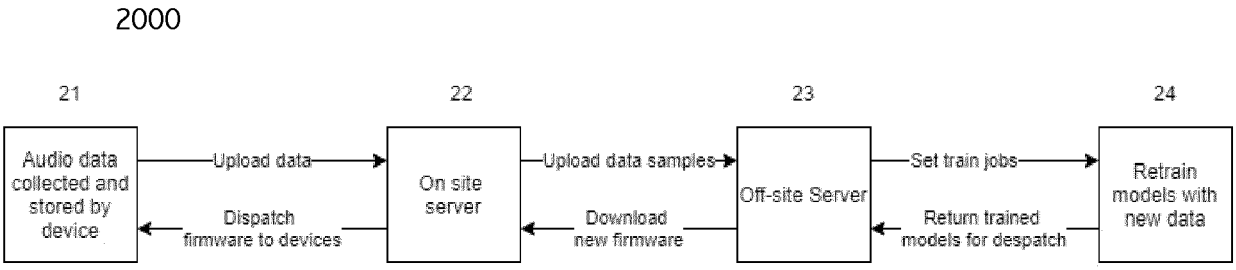
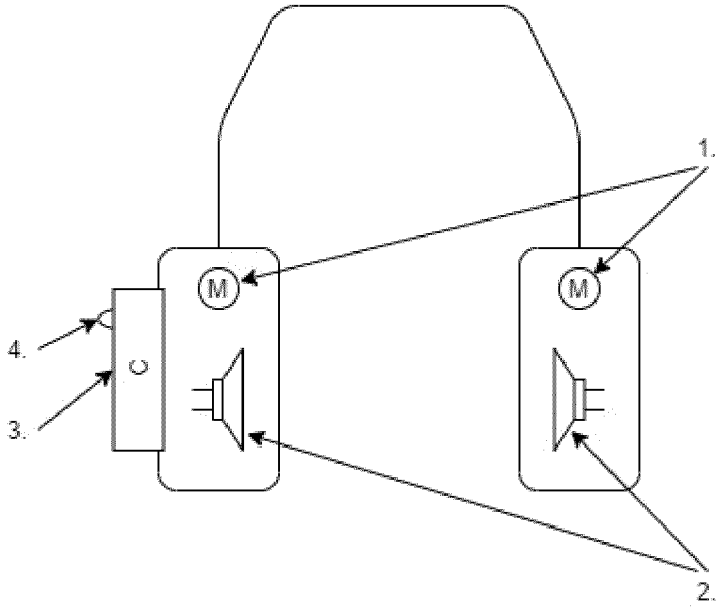

2000
21
22
23
24
Audio data collected and stored by device
Upload data
Dispatch firmware to devices
On site server
Upload data samples
Download new firmware
Off-site Server
Set train jobs
Return trained models for despatch
Retrain models with new data

Figure 3

3000



Europäisches Patentamt
European Patent Office
Office européen des brevets

EUROPEAN SEARCH REPORT

Application Number
EP 21 19 0382

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
X	VALIN JEAN-MARC: "A Hybrid DSP/Deep Learning Approach to Real-Time Full-Band Speech Enhancement", 2018 IEEE 20TH INTERNATIONAL WORKSHOP ON MULTIMEDIA SIGNAL PROCESSING (MMSP), 31 May 2018 (2018-05-31), pages 1-5, XP055783657, DOI: 10.1109/MMSP.2018.8547084 ISBN: 978-1-5386-6070-6 Retrieved from the Internet: URL:https://arxiv.org/pdf/1709.08243.pdf>	1-9, 13-15	INV. G10L21/0208 G10L25/30 G10L21/0232
Y	* figure 2 * * paragraph [II.C]; figure 3 * * paragraph [0III] *	14	
X	EP 2 151 822 A1 (FRAUNHOFER GES FORSCHUNG [DE]) 10 February 2010 (2010-02-10) * paragraphs [0029], [0031], [0033], [0044], [0079]; figures 4,8,10 * * paragraphs [0052], [0068], [0073] - [0079] * * paragraph [0002] *	1-8, 12-15	
A	BORGSTROM BENGT J ET AL: "Improving Statistical Model-Based Speech Enhancement with Deep Neural Networks", 2018 16TH INTERNATIONAL WORKSHOP ON ACOUSTIC SIGNAL ENHANCEMENT (IWAENC), IEEE, 17 September 2018 (2018-09-17), pages 471-475, XP033439100, DOI: 10.1109/IWAENC.2018.8521382 [retrieved on 2018-11-02] * paragraph [02.1]; figure 1 *	9	

-/--

TECHNICAL FIELDS SEARCHED (IPC)
G10L

1 The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
Munich	18 January 2022	Krembel, Luc

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

Europäisches Patentamt
European Patent Office
Office européen des brevets

EUROPEAN SEARCH REPORT

Application Number
EP 21 19 0382

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
A	SALEEM N ET AL: "Spectral Phase Estimation Based on Deep Neural Networks for Single Channel Speech Enhancement", JOURNAL OF COMMUNICATIONS TECHNOLOGY AND ELECTRONICS, NAUKA/INTERPERIODICA PUBLISHING, MOSCOW, RU, vol. 64, no. 12, 1 December 2019 (2019-12-01), pages 1372-1382, XP037029872, ISSN: 1064-2269, DOI: 10.1134/S1064226919120155 [retrieved on 2020-02-21] * paragraph [0003]; figure 1 * * page 1375 "Post processing" - 1376 *	10-12	
A	JAMAL NOREZMI ET AL: "A Hybrid Approach for Single Channel Speech Enhancement using Deep Neural Network and Harmonic Regeneration Noise Reduction", IJACSA) INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS, vol. 11, 1 January 2020 (2020-01-01), XP055880063, Retrieved from the Internet: URL:https://thesai.org/Downloads/Volume11No10/Paper_33-A_Hybrid_Approach_for_Single_Channel_Speech_Enhancement.pdf> * paragraphs [III.C], [III.D], [III.E] *	10-12	TECHNICAL FIELDS SEARCHED (IPC)
Y	US 2019/080710 A1 (ZHANG MI [US] ET AL) 14 March 2019 (2019-03-14) * paragraph [0056] *	14	

1 The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
Munich	18 January 2022	Krembel, Luc

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 21 19 0382

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

18-01-2022

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 2151822 A1	10-02-2010	AU 2009278263 A1	11-02-2010
		CA 2732723 A1	11-02-2010
		CN 102124518 A	13-07-2011
		EP 2151822 A1	10-02-2010
		ES 2678415 T3	10-08-2018
		HK 1159300 A1	27-07-2012
		JP 5666444 B2	12-02-2015
		JP 2011530091 A	15-12-2011
		KR 20110044990 A	03-05-2011
		RU 2011105976 A	27-08-2012
		TR 201810466 T4	27-08-2018
		US 2011191101 A1	04-08-2011
		WO 2010015371 A1	11-02-2010
US 2019080710 A1	14-03-2019	US 2019080710 A1	14-03-2019
		US 2021050030 A1	18-02-2021

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82