

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7188589号
(P7188589)

(45)発行日 令和4年12月13日(2022.12.13)

(24)登録日 令和4年12月5日(2022.12.5)

(51)国際特許分類 F I
G 0 6 N 3/04 (2006.01) G 0 6 N 3/04
G 0 6 N 3/08 (2006.01) G 0 6 N 3/08

請求項の数 7 (全13頁)

(21)出願番号	特願2021-528089(P2021-528089)	(73)特許権者	000004226 日本電信電話株式会社 東京都千代田区大手町一丁目5番1号
(86)(22)出願日	令和1年6月18日(2019.6.18)	(74)代理人	100121706 弁理士 中尾 直樹
(86)国際出願番号	PCT/JP2019/024058	(74)代理人	100128705 弁理士 中村 幸雄
(87)国際公開番号	WO2020/255242	(74)代理人	100147773 弁理士 義村 宗洋
(87)国際公開日	令和2年12月24日(2020.12.24)	(72)発明者	江村 暁 東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内
審査請求日	令和3年10月11日(2021.10.11)	審査官	三坂 敏夫

最終頁に続く

(54)【発明の名称】 復元装置、復元方法、およびプログラム

(57)【特許請求の範囲】

【請求項1】

クリップ後信号と当該クリップ後信号のうちクリップされた部分を表すクリップ情報とを含む入力データから信号復元ニューラルネットワークを用いて当該クリップ後信号に対応するクリップ前信号を推定する復元部を含み、

上記信号復元ニューラルネットワークは、クリップ前信号と当該クリップ前信号に対応するクリップ後信号と当該クリップ後信号に関するクリップ情報とを学習データとして、上記入力データを入力とし、上記クリップ前信号の推定値を出力するよう学習されたものであり、

上記クリップ情報は、上限値でクリップされた部分を表す上限クリップ情報と下限値でクリップされた部分を表す下限クリップ情報とからなる、

復元装置。

【請求項2】

クリップ後信号と当該クリップ後信号のうちクリップされた部分を表すクリップ情報とを含む入力データから信号復元ニューラルネットワークを用いて当該クリップ後信号に対応するクリップ前信号を推定する復元部を含み、

上記信号復元ニューラルネットワークは、クリップ前信号と当該クリップ前信号に対応するクリップ後信号と当該クリップ後信号に関するクリップ情報とを学習データとして、上記入力データを入力とし、上記クリップ前信号の推定値を出力するよう学習されたものであり、

上記信号復元ニューラルネットワークは、ゲート付き畳み込みニューラルネットワークであり、活性化関数が正負の値を出力する関数である、
復元装置。

【請求項 3】

クリップ後信号と当該クリップ後信号のうちクリップされた部分を表すクリップ情報とを含む入力データから信号復元ニューラルネットワークを用いて当該クリップ後信号に対応するクリップ前信号を推定する復元部を含み、

上記信号復元ニューラルネットワークは、クリップ前信号と当該クリップ前信号に対応するクリップ後信号と当該クリップ後信号に関するクリップ情報とを学習データとして、上記入力データを入力とし、上記クリップ前信号の推定値を出力するよう学習されたものであり、

10

上記復元部は、上記信号復元ニューラルネットワークを二段直列に接続し、一段目の信号復元ニューラルネットワークの出力と上記クリップ情報とからなる入力データを二段目の信号復元ニューラルネットワークへ入力して二段目の信号復元ニューラルネットワークの出力を上記クリップ前信号の推定値とする、

復元装置。

【請求項 4】

復元部が、クリップ後信号と当該クリップ後信号のうちクリップされた部分を表すクリップ情報とを含む入力データから信号復元ニューラルネットワークを用いて当該クリップ後信号に対応するクリップ前信号を推定し、

20

上記信号復元ニューラルネットワークは、クリップ前信号と当該クリップ前信号に対応するクリップ後信号と当該クリップ後信号に関するクリップ情報とを学習データとして、上記入力データを入力とし、上記クリップ前信号の推定値を出力するよう学習されたものであり、

上記クリップ情報は、上限値でクリップされた部分を表す上限クリップ情報と下限値でクリップされた部分を表す下限クリップ情報とからなる、

復元方法。

【請求項 5】

復元部が、クリップ後信号と当該クリップ後信号のうちクリップされた部分を表すクリップ情報とを含む入力データから信号復元ニューラルネットワークを用いて当該クリップ後信号に対応するクリップ前信号を推定し、

30

上記信号復元ニューラルネットワークは、クリップ前信号と当該クリップ前信号に対応するクリップ後信号と当該クリップ後信号に関するクリップ情報とを学習データとして、上記入力データを入力とし、上記クリップ前信号の推定値を出力するよう学習されたものであり、

上記信号復元ニューラルネットワークは、ゲート付き畳み込みニューラルネットワークであり、活性化関数が正負の値を出力する関数である、

復元方法。

【請求項 6】

復元部が、クリップ後信号と当該クリップ後信号のうちクリップされた部分を表すクリップ情報とを含む入力データから信号復元ニューラルネットワークを用いて当該クリップ後信号に対応するクリップ前信号を推定し、

40

上記信号復元ニューラルネットワークは、クリップ前信号と当該クリップ前信号に対応するクリップ後信号と当該クリップ後信号に関するクリップ情報とを学習データとして、上記入力データを入力とし、上記クリップ前信号の推定値を出力するよう学習されたものであり、

上記復元部は、上記信号復元ニューラルネットワークを二段直列に接続し、一段目の信号復元ニューラルネットワークの出力と上記クリップ情報とからなる入力データを二段目の信号復元ニューラルネットワークへ入力して二段目の信号復元ニューラルネットワークの出力を上記クリップ前信号の推定値とする、

50

復元方法。

【請求項 7】

請求項 1 から 3 のいずれかに記載の復元装置としてコンピュータを機能させるためのプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

この発明は、クリップ後の信号からクリップ前の信号を復元する技術に関する。

【背景技術】

【0002】

装置間で信号を入出力する際に、信号振幅が装置の入出力レンジより大きい部分は一定値にクリッピングされる。クリッピングは、例えば、センサーから信号を得るとき、信号を何らかの機器に出力するとき、アナログ信号をA/Dコンバータに入力してデジタル化するとき等、広範囲な状況で生じる可能性がある。そこで、クリップされた信号からクリップ前の信号波形を復元する研究がこれまでなされてきた。

【0003】

そのような方法として、SPADE (SParse Audio DEclipper) と呼ばれる方法が提案されている (非特許文献 1)。以下、SPADEについて説明する。

【0004】

なお、文中で使用する記号「 $\bar{\cdot}$ 」「 \wedge 」は、本来直前の文字の真上に記載されるべきものであるが、テキスト記法の制限により、当該文字の直後に記載する。数式中においてはこれらの記号は本来の位置、すなわち文字の真上に記述している。例えば、「 z^{\wedge} 」は数式中では次式で表される。

【0005】

【数 1】

$$\bar{z}$$

【0006】

また、例えば、「 z^{\wedge} 」は数式中では次式で表される。

【0007】

【数 2】

$$\hat{z}$$

【0008】

本来の信号 (クリップ前の信号) を信号ベクトル $x=[x_1, \dots, x_N]$ 、クリップされた信号を信号ベクトル $y=[y_1, \dots, y_N]$ で表現する。クリッピング前後の信号の各サンプルは、式 (1) の関係がある。

【0009】

【数 3】

$$y_i = \begin{cases} \theta & \text{for } x_i > \theta, \\ x_i & \text{for } |x_i| \leq \theta, \\ -\theta & \text{for } x_i < -\theta \end{cases} \quad \dots(1)$$

【0010】

ここで θ はクリッピングレベルである。クリッピング後の信号サンプルは、上限でクリップされる信号サンプル S_+ 、クリップされない信号サンプル S_r 、下限でクリップされる信

10

20

30

40

50

号サンプル s のいずれかに属する。

【 0 0 1 1 】

SPADEでは、まず辞書行列 D を定める。そして、辞書行列 D の逆行列 D^{-1} を信号ベクトル x にかけて得られる信号表現ベクトル z に注目して、信号の複雑さを z 中の非ゼロ要素の数、すなわち z の L_0 ノルム $\|z\|_0$ はかる。辞書行列 D としては、DFT行列 (Discrete Fourier Transform Matrix) やDCT行列 (Discrete Cosine Transform Matrix) 等が用いられる。

【 0 0 1 2 】

SPADEは、クリップ前の信号の複雑さを k とし、所定の更新量 s を複雑さ k の初期値として想定する。まず、入力信号すなわちクリップ後の信号 y を、 D^{-1} で信号表現ベクトル z^{\sim} へ変換する。 z^{\sim} 中で絶対値の大きい方から k 個の要素を残し、それ以外の値を0にすることで、複雑さ k の信号表現ベクトル z^{\wedge} へと変換する。この操作はhard thresholdingと呼ばれ、数式で $z^{\wedge} = H_k(z)$ と表現される (下記表1のステップ2に対応)。次に、この信号表現ベクトル z^{\wedge} に D をかけて、推定信号ベクトル $x^{\sim} = Dz^{\wedge}$ へ変換する。推定信号ベクトル x^{\sim} は、この段階でのクリップ前の信号ベクトル x の推定結果になる。通常、この推定信号ベクトル x^{\sim} と入力信号ベクトル y では非クリップ部分でも乖離がある。そこで下記の二つの条件を満たす信号表現ベクトル z^{\wedge} を求める (下記表1のステップ3に対応)。

【 0 0 1 3 】

条件1 . クリップされた Dz^{\wedge} が y と一致する。

【 0 0 1 4 】

条件2 . z^{\wedge} と z^{\sim} の距離が最小になる。

【 0 0 1 5 】

z^{\wedge} と z^{\sim} の距離があらかじめ決めた値より大きい場合には、「想定する信号の複雑さ k が不足しているためにターゲット信号を表現できない」と判定して複雑さ k を更新量 s ずつ増やし、上記の処理を繰り返す。

【 0 0 1 6 】

以上の処理を、最適化手法ADMM (非特許文献2) を用いて実装すると、表1のアルゴリズムが得られる。

【 0 0 1 7 】

【表1】

```

1:  $\hat{z}^{(0)} = D^*y, u^{(0)} = 0, i = 1, k = s$ 
2:  $\bar{z}^{(i)} = \mathcal{H}_k(\hat{z}^{(i-1)} + u^{(i-1)})$ 
3:  $\hat{z}^{(i)} = \arg \min \|z - (\bar{z}^{(i)} - u^{(i-1)})\|_2^2$ 
   subject to  $Dz$  in  $\Gamma(y)$ 
4: if  $\|\hat{z}^{(i)} - \bar{z}^{(i)}\|_2 \leq \epsilon$  or  $i > max\_iter$  then
5:   terminate
6: else
7:    $u^{(i)} = u^{(i-1)} + \hat{z}^{(i)} - \bar{z}^{(i)}$ 
8:    $i \leftarrow i + 1$ 
9:   if  $i \bmod r = 0$  then
10:     $k \leftarrow k + s$ 
11:   end if
12:   go to 2
13: end if
14: return  $x = D\hat{z}^{(i)}$ 

```

【 0 0 1 8 】

SPADEは、通常のフレーム信号処理と組み合わせて使用される。つまり、入力されたク

リップ後の信号は、重なりをもつ一定長のフレームに分割され、各フレームに窓かけ処理をした後、上記SPADE処理が適用される。そして、その処理結果にフレーム合成処理が適用され、復元したクリップ前の信号が得られる。

【先行技術文献】

【非特許文献】

【0019】

【文献】S. Kitic, N. Bertin, and R. Gribnoval, "Sparsity and cosparsity for audio declipping: a flexible non-convex approach", The 12th International conference on Latent Variable Analysis and Signal Separation (LVA/ICA2015), 2015. S. Boyd, N. Parkikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers", Foundation and Trend in Machine Learning, vol. 3, no. 1, 2011.

10

【発明の概要】

【発明が解決しようとする課題】

【0020】

しかしながら、SPADEには、リアルタイムにセンサー信号の波形を復元する必要がある場合には、演算量が変動する、という問題がある。これは、SPADEが複雑さの想定を順次増やしながら繰り返し型で波形復元処理を進めるためであり、かつ、そもそも入力信号の複雑さが未知であり絶えず変動しているためである。また、クリップされる部分が増えるほど、元の信号の特性が復元後の信号に反映されにくくなる、という問題もある。

20

【0021】

この発明の目的は、上記のような技術的課題に鑑みて、一定の演算量で精度よくクリップされた信号を復元することができる技術を実現することである。

【課題を解決するための手段】

【0022】

上記の課題を解決するために、この発明の一態様の復元装置は、クリップ後信号と当該クリップ後信号のうちクリップされた部分を表すクリップ情報とを含む入力データから信号復元ニューラルネットワークを用いて当該クリップ後信号に対応するクリップ前信号を推定する復元部を含み、信号復元ニューラルネットワークは、クリップ前信号と当該クリップ前信号に対応するクリップ後信号と当該クリップ後信号に関するクリップ情報とを学習データとして、入力データを入力とし、クリップ前信号の推定値を出力するよう学習されたものである。

30

【発明の効果】

【0023】

この発明の復元技術によれば、一定の演算量で精度よくクリップされた信号を復元することができる。

【図面の簡単な説明】

【0024】

【図1】図1は、波形復元装置の機能構成を例示する図である。

【図2】図2は、波形復元部の構成を例示する図である。

40

【図3】図3は、波形復元方法の処理手順を例示する図である。

【図4】図4は、第二実施形態の波形復元部の機能構成を例示する図である。

【図5】図5は、コンピュータの機能構成を例示する図である。

【発明を実施するための形態】

【0025】

以下、この発明の実施の形態について詳細に説明する。なお、図面中において同じ機能を有する構成部には同じ番号を付し、重複説明を省略する。

【0026】

[第一実施形態]

第一実施形態の信号復元装置(以下、「復元装置」とも呼ぶ)は、ゲート付き畳み込み

50

ニューラルネットワーク (Gated Convolutional Neural Network) (例えば、参考文献 1, 2 参照) から構成される信号復元ニューラルネットワークにより、クリップ後の信号からクリップ前の信号を復元する信号処理装置である。ニューラルネットワークの演算は固定であるため、信号復元ニューラルネットワークによる信号復元処理全体の演算量は一定となる。また、十分な学習データを用いて事前に十分に信号復元ニューラルネットワークを学習させることで、復元後の信号にクリップ前の信号の特性がより良く反映されることが期待できる。

【0027】

[参考文献 1] Y. N. Dauphin, A. Fan, M. Auli, and D. Grangier, "Language Modeling with Gated Convolutional Networks," arXiv:1612.08083, Submitted on 23 Dec 2016 (v1).

10

[参考文献 2] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. S. Huang, "Free-Form Image Inpainting with Gated Convolution," arXiv:1806.03589, Submitted on 10 Jun 2018.

第一実施形態の波形復元装置 1 は、図 1 に例示するように、フレーム分割部 11、波形復元部 12 (以下、「復元部」とも呼ぶ)、およびフレーム合成部 13 を備える。波形復元部 12 は、図 2 に例示するように、信号復元ニューラルネットワーク 121 および置換部 122 を備える。この波形復元装置 1 が、図 3 に例示する各ステップの処理を行うことにより第一実施形態の波形復元方法が実現される。

【0028】

20

波形復元装置 1 は、例えば、中央演算処理装置 (CPU: Central Processing Unit)、主記憶装置 (RAM: Random Access Memory) などをも有する公知又は専用のコンピュータに特別なプログラムが読み込まれて構成された特別な装置である。波形復元装置 1 は、例えば、中央演算処理装置の制御のもとで各処理を実行する。波形復元装置 1 に入力されたデータや各処理で得られたデータは、例えば、主記憶装置に格納され、主記憶装置に格納されたデータは必要に応じて中央演算処理装置へ読み出されて他の処理に利用される。波形復元装置 1 の各処理部は、少なくとも一部が集積回路等のハードウェアによって構成されていてもよい。

【0029】

図 2 を参照して、信号復元ニューラルネットワーク 121 内で、入力データがどのように中間データに変換されていき、最終的に出力されるかを説明する。

30

【0030】

はじめに、信号復元ニューラルネットワーク 121 の前段階 (例えば、フレーム分割部 11 等) において、波形復元装置 1 に入力されたクリップ後信号のベクトルと上限クリップ情報のベクトルと下限クリップ情報のベクトルとから入力データを構成する。クリップ後信号のベクトルは、L サンプルのクリップ後信号からなる L 次元ベクトルである。上限クリップ情報は、上限値以上の信号サンプルがあった場所に 1、それ以外の場所に 0 が設定されている L 次元ベクトルである。下限クリップ情報は、下限値以下の信号サンプルがあった場所に 1、それ以外の場所に 0 が設定されている L 次元ベクトルである。すなわち、図 2 に示すように、クリップ後信号のベクトルを上限クリップ情報のベクトルと下限クリップ情報のベクトルに挟んで構成した $L \times 3$ 行列が入力データになる。

40

【0031】

信号復元ニューラルネットワークを学習する際には、上記の入力データとクリップ前信号が学習データとして与えられる。学習済みの信号復元ニューラルネットワークで推定する際には、復元したいクリップ後信号に関する入力データを入力し、その出力をクリップ前信号の推定値とする。最後に、置換部 122 が、クリップ後信号のベクトル中で上限クリップもしくは下限クリップされている部分を、信号復元ニューラルネットワークで推定された値に置き換えて、復元されたクリップ前信号として出力する。

【0032】

信号復元ニューラルネットワークは、多層のゲート付き畳み込みニューラルネットワー

50

クで構成される。畳み込みニューラルネットワークは、入力データ（信号）を時間方向に複数切り出してフィルタをかけ、活性化関数に通すことで特徴ベクトルを出力する。信号長 $L = 1024$ のとき、フィルタ長には、例えば $3 \sim 20$ タップを用いる。フィルタの種類を増やすことで特徴ベクトルの数、すなわちチャンネル数を増やす。図 2 において、四角形で示したデータ ($L_1 \sim L_5$) は各中間データであり、その縦幅が時間方向サンプル数に、その横幅がチャンネル数に対応する。通常畳み込みニューラルネットワークの 1 層分の変換は、 Y を入力ベクトルとして、次式で表される。

【0033】

【数 4】

$$h(Y) = \tanh(Y * W + b)$$

10

【0034】

一方、ゲート付き畳み込みニューラルネットワークでは、この変換が次式になる。

【0035】

【数 5】

$$h(Y) = \sigma(Y * W + b) \otimes \sigma(Y * V + c)$$

【0036】

ただし、

【0037】

【数 6】

⊗

20

【0038】

は要素ごとの積 (element-wise product) であり、 σ は活性化関数であり、 W, b, V, c は学習済みのパラメータである。本実施形態では、入力される信号および出力する信号は共に正負の値を取るため、活性化関数には正負の値を出力する関数 (例えば、 \tanh など) を用いる。

30

【0039】

信号復元ニューラルネットワークは、クリップ後信号を高次の特徴量へエンコードしていく過程と、高次の特徴量を復元信号へデコードする過程とからなり、最終的にデコード過程から L 次元ベクトルが出力される。エンコードの過程では、フィルタの種類を増やすことでチャンネル数を増やすと同時にマックスプーリングを用いて時間方向サンプル数を減らしていく。デコードの過程では、逆にフィルタの種類を減らすことでチャンネル数を減らすと同時にアップサンプリングを用いて時間方向サンプル数を増やしていく。図 2 では隠れ層が 5 層の構成を示したが、本発明の層数はこれに限定されない。これより層数の少ない構成も多い構成も考えられる。

【0040】

なお、入力データから中間データ、中間データから中間データ、中間データから出力への各変換 ($G_1 \sim G_6$) には、図 2 に示すように、ゲート付き畳み込みニューラルネットワークとマックスプーリング、バッチノーマライゼーションが使われる。また、信号復元ニューラルネットワーク全体を学習するときのコスト関数には、参考文献 2 と同様に、クリップ前の信号と復元後の信号との差信号の L_1 ノルムを用いる。

40

【0041】

以下、図 3 を参照して、第一実施形態の波形復元装置 1 が実行する波形復元方法の処理手続きを説明する。

【0042】

波形復元装置 1 には、クリップ後信号と、クリップ後信号に関する上限クリップ情報と

50

、クリップ後信号に関する下限クリップ情報とが入力される。

【 0 0 4 3 】

ステップ S 1 1 において、フレーム分割部 1 1 は、入力されたクリップ後信号と上限クリップ情報と下限クリップ情報とをそれぞれ L サンプルずつに分割して入力データを生成する。すなわち、入力データは、L サンプルのクリップ後信号を表す L 次元ベクトルと、クリップ後信号の各サンプルに対応する上限クリップ情報を表す L 次元ベクトルと、クリップ後信号の各サンプルに対応する下限クリップ情報を表す L 次元ベクトルとを組にしたデータである。より具体的には、L 次元のクリップ後信号のベクトルが L 次元の上限クリップ情報のベクトルと L 次元の下限クリップ情報のベクトルとに挟まれた $L \times 3$ 行列が入力データとなる。フレーム分割部 1 1 は、生成した入力データを波形復元部 1 2 へ送る。

10

【 0 0 4 4 】

ステップ S 1 2 において、波形復元部 1 2 は、信号復元ニューラルネットワーク 1 2 1 を用いて入力データからクリップ前信号を推定する。すなわち、波形復元部 1 2 は、フレーム分割部 1 1 から受け取った入力データを信号復元ニューラルネットワーク 1 2 1 へ入力し、置換部 1 2 2 で入力データのクリップ後信号のベクトルのうち上限値でクリップまたは下限値でクリップされている部分を信号復元ニューラルネットワーク 1 2 1 で推定された値に置換することで、クリップ前信号のベクトルを生成する。波形復元部 1 2 は、推定したクリップ前信号のベクトルをフレーム合成部 1 3 へ送る。

【 0 0 4 5 】

ステップ S 1 3 において、フレーム合成部 1 3 は、推定されたクリップ前信号のベクトルにフレーム合成処理を適用し、クリップ前信号を復元する。

20

【 0 0 4 6 】

[第二実施形態]

第一実施形態の信号復元ニューラルネットワークでは、信号のおおまかな形状は復元されるが、細部の形状は復元されにくい傾向がある。そこで、第二実施形態の波形復元部では、細部形状の復元精度を上げるために、図 4 に示すように、信号復元ニューラルネットワークを二段直列に接続する。すなわち、第一実施形態の信号復元ニューラルネットワーク 1 2 1 - 1 により復元された信号ベクトルに対して、さらに二段目の信号復元ニューラルネットワーク 1 2 1 - 2 を適用し、クリップ前信号のベクトルを推定するように構成する。

30

【 0 0 4 7 】

第一実施形態と同様に、入力データは、クリップ後信号のベクトルを上限クリップ情報のベクトルと下限クリップ情報のベクトルとで挟んで構成する。信号長が L のとき、入力データは $L \times 3$ 行列になる。二段目の信号復元ニューラルネットワーク 1 2 1 - 2 を学習する際は、この入力データとクリップ前信号とを学習データとして与える。学習後、信号復元ニューラルネットワークで推定する際は、入力データを一段目の信号復元ニューラルネットワーク 1 2 1 - 1 へ入力し、二段目の信号復元ニューラルネットワーク 1 2 1 - 2 の出力をクリップ前信号の推定値とする。

【 0 0 4 8 】

二段目の信号復元ニューラルネットワーク 1 2 1 - 2 の内部構成は、図 2 に示した第一実施形態の信号復元ニューラルネットワークと同様である。すなわち、信号復元ニューラルネットワーク 1 2 1 - 2 は、クリップ後の信号を高次の特徴量へエンコードしていく過程と、高次の特徴量を復元信号へデコードする過程とからなり、最後にデコード過程から L 次元ベクトルが出力される。各中間データの時間方向サンプル数およびチャンネル数は、一段目の信号復元ニューラルネットワーク 1 2 1 - 1 と同一であってもよいし、違っていてもよい。また層数についても一段目の信号復元ニューラルネットワーク 1 2 1 - 1 と同じであってもよいし、違っていてもよい。

40

【 0 0 4 9 】

[第二実施形態の変形例]

クリップ後の信号ではなく欠落のある信号を対象として元の信号を復元する場合にも、

50

クリップ後の信号と同様に元の信号の細部の情報が復元された信号から失われやすい。そのため、欠落のある信号を復元する場合にも、第二実施形態の構成を適用できる。この場合、入力データは、欠落有り信号ベクトルと欠落情報ベクトルからなる $L \times 2$ 行列になる。図 4 に示すように、二段目の信号復元ニューラルネットワークを用いることで、一段目の推定信号から、より復元精度の高い復元信号を推定することができる。

【 0 0 5 0 】

[発明のポイント]

本発明のポイントは以下の 3 点である。

【 0 0 5 1 】

1. クリップ後信号をゲートつき畳み込みニューラルネットワークを用いて復元する信号復元ニューラルネットワークにおいて、上限クリップ情報のベクトルと下限クリップ情報のベクトルとでクリップ後信号のベクトルを挟んで、入力データを構成する。

10

【 0 0 5 2 】

2. ゲートつき畳み込みニューラルネットワークの内部では、正負の値を出力する関数 (\tanh) を活性化関数として使う。

【 0 0 5 3 】

3. 信号の復元精度を高めるために、信号復元ニューラルネットワークを二段構成にする。はじめに一段目の信号復元ニューラルネットワークを学習させる。その学習後の推定結果を用いて二段目の信号復元ニューラルネットワークを学習させる。

【 0 0 5 4 】

20

以上、この発明の実施の形態について説明したが、具体的な構成は、これらの実施の形態に限られるものではなく、この発明の趣旨を逸脱しない範囲で適宜設計の変更等があっても、この発明に含まれることはいうまでもない。実施の形態において説明した各種の処理は、記載の順に従って時系列に実行されるのみならず、処理を実行する装置の処理能力あるいは必要に応じて並列的あるいは個別に実行されてもよい。

【 0 0 5 5 】

[プログラム、記録媒体]

上記実施形態で説明した各装置における各種の処理機能をコンピュータによって実現する場合、各装置が有すべき機能の処理内容はプログラムによって記述される。そして、このプログラムを図 5 に示すコンピュータの記憶部 1020 に読み込ませ、制御部 1010、入力部 1030、出力部 1040 などに動作させることにより、上記各装置における各種の処理機能がコンピュータ上で実現される。

30

【 0 0 5 6 】

この処理内容を記述したプログラムは、コンピュータで読み取り可能な記録媒体に記録しておくことができる。コンピュータで読み取り可能な記録媒体としては、例えば、磁気記録装置、光ディスク、光磁気記録媒体、半導体メモリ等のようなものでもよい。

【 0 0 5 7 】

また、このプログラムの流通は、例えば、そのプログラムを記録した DVD、CD-ROM 等の可搬型記録媒体を販売、譲渡、貸与等することによって行う。さらに、このプログラムをサーバコンピュータの記憶装置に格納しておき、ネットワークを介して、サーバコンピュータから他のコンピュータにそのプログラムを転送することにより、このプログラムを流通させる構成としてもよい。

40

【 0 0 5 8 】

このようなプログラムを実行するコンピュータは、例えば、まず、可搬型記録媒体に記録されたプログラムもしくはサーバコンピュータから転送されたプログラムを、一旦、自己の記憶装置に格納する。そして、処理の実行時、このコンピュータは、自己の記憶装置に格納されたプログラムを読み取り、読み取ったプログラムに従った処理を実行する。また、このプログラムの別の実行形態として、コンピュータが可搬型記録媒体から直接プログラムを読み取り、そのプログラムに従った処理を実行することとしてもよく、さらに、このコンピュータにサーバコンピュータからプログラムが転送されるたびに、逐次、受け

50

取ったプログラムに従った処理を実行することとしてもよい。また、サーバコンピュータから、このコンピュータへのプログラムの転送は行わず、その実行指示と結果取得のみによって処理機能を実現する、いわゆるASP (Application Service Provider) 型のサービスによって、上述の処理を実行する構成としてもよい。なお、本形態におけるプログラムには、電子計算機による処理の用に供する情報であってプログラムに準ずるもの(コンピュータに対する直接の指令ではないがコンピュータの処理を規定する性質を有するデータ等)を含むものとする。

【 0 0 5 9 】

また、この形態では、コンピュータ上で所定のプログラムを実行させることにより、本装置を構成することとしたが、これらの処理内容の少なくとも一部をハードウェア的に実現することとしてもよい。

10

20

30

40

50

【 図 面 】

【 図 1 】

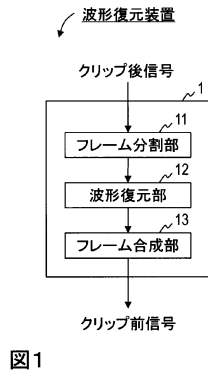


図 1

【 図 2 】

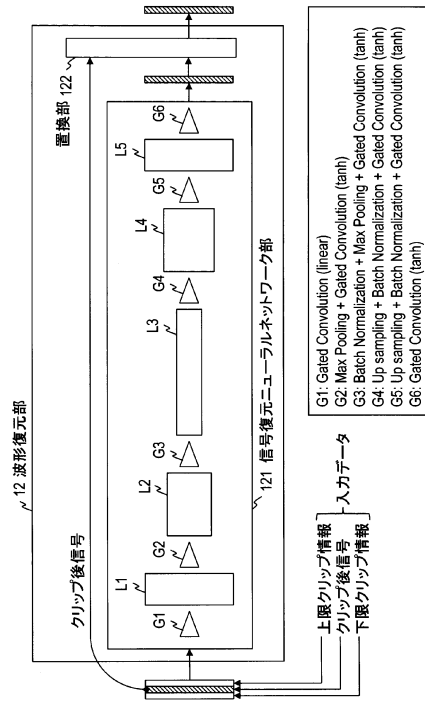


図 2

【 図 3 】

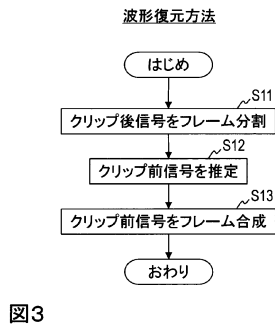


図 3

【 図 4 】

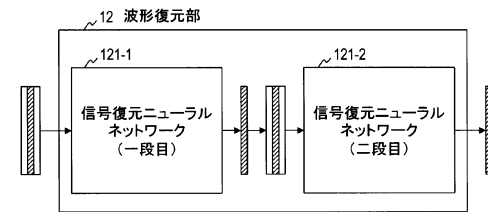


図 4

10

20

30

40

50

【 図 5 】

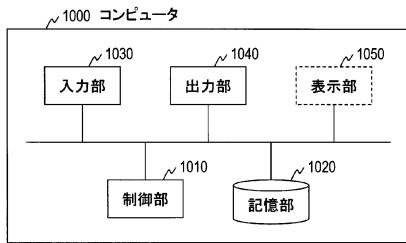


図5

10

20

30

40

50

フロントページの続き

- (56)参考文献 特開 2 0 0 5 - 2 7 5 4 1 0 (J P , A)
特開 2 0 1 3 - 1 6 2 3 4 7 (J P , A)
IIZUKA,Satoshi et al. , Globally and Locally Consistent Image Completion , [online] , 米国
 , ACM , 2017年07月 , pages:1-14 , [Retrieved from the Internet] URL:<https://dl.acm.org/doi/pdf/10.1145/3072959.3073659>
YU,Jiahui et al. , Free-Form Image Inpainting with Gated Convolution , [online] , 米国 , arXiv , 2018年06月10日 , pages:1-12 , [retrieved on 2022.8.24] , Retrieved from the Internet
:< U R L : <https://arxiv.org/pdf/1806.03589v1.pdf> >
- (58)調査した分野 (Int.Cl. , D B 名)
G 0 6 N 3 / 0 0 - 3 / 1 2
G 0 6 N 1 0 / 0 0 - 2 0 / 2 0