

(19) 日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2020-9444

(P2020-9444A)

(43) 公開日 令和2年1月16日 (2020.1.16)

(51) Int.Cl.

G06N 3/08 (2006.01)

F I

G06N 3/08

テーマコード (参考)

審査請求 未請求 請求項の数 17 O L (全 33 頁)

(21) 出願番号	特願2019-124747 (P2019-124747)	(71) 出願人	390019839
(22) 出願日	令和1年7月3日 (2019.7.3)		三星電子株式会社
(31) 優先権主張番号	10-2018-0077894		Samsung Electronics Co., Ltd.
(32) 優先日	平成30年7月4日 (2018.7.4)		大韓民国京畿道水原市靈通区三星路129
(33) 優先権主張国・地域又は機関	韓国 (KR)		129, Samsung-ro, Yeongtong-gu, Suwon-si, Gyeonggi-do, Republic of Korea
		(74) 代理人	100107766
			弁理士 伊東 忠重
		(74) 代理人	100070150
			弁理士 伊東 忠彦
		(74) 代理人	100091214
			弁理士 大貫 進介

最終頁に続く

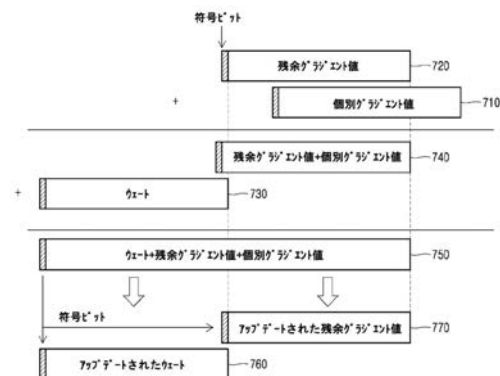
(54) 【発明の名称】 ニューラルネットワークにおいてパラメータを処理する方法及び装置

(57) 【要約】

【課題】 ニューラルネットワークにおいて、パラメータを処理する方法及びその装置を提供する。

【解決手段】 低精度ナンバーシステムを利用するニューラルネットワークにおいて、パラメータを処理する方法及びその装置に係り、ニューラルネットワークのウェイトをアップデートするための個別グラジエント値と残余グラジエント値とを計算し、加算器を利用し、個別グラジエント値、残余グラジエント値及びウェイトを合算し、合算結果に基づいて、ウェイト及び残余グラジエント値をアップデートする方法及びその装置である。

【選択図】 図7



【特許請求の範囲】

【請求項 1】

低精度ナンバーシステムを利用するニューラルネットワークにおいて、パラメータを処理する方法において、

前記ニューラルネットワークのウェートをアップデートするための個別グラジエント値を計算する段階と、

前記個別グラジエント値を累積した累積グラジエント値、及び前記ウェートのビット桁に基づいて、残余グラジエント値を計算する段階と、

前記残余グラジエント値のビット桁に対応するように、前記個別グラジエント値を調整する段階と、

加算器を利用し、前記調整された個別グラジエント値、前記残余グラジエント値及び前記ウェートを合算する段階と、

前記合算結果に基づいて、前記ウェート及び前記残余グラジエント値をアップデートする段階と、を含む方法。

【請求項 2】

前記残余グラジエント値を計算する段階は、

前記累積グラジエント値のうち、前記ウェートのビット桁に合算可能な値を有効グラジエント値と決定する段階と、

前記累積グラジエント値から前記有効グラジエント値を減算することにより、前記残余グラジエント値を計算する段階と、を含むことを特徴とする請求項 1 に記載の方法。

【請求項 3】

前記個別グラジエント値を調整する段階は、

前記個別グラジエント値のうち、前記残余グラジエント値の最下位ビット桁未満の値が省略されるように、前記個別グラジエント値を量子化する段階と、

前記残余グラジエント値の最上位ビット桁に対応するビット桁まで値が存在するように、前記量子化された個別グラジエント値をパディングする段階と、を含むことを特徴とする請求項 1 又は 2 に記載の方法。

【請求項 4】

前記合算する段階は、

前記加算器のビット数に基づいて、前記調整された個別グラジエント値及び前記残余グラジエント値を前記加算器にマッピングし、前記加算器を利用して中間合算値を計算する段階と、

前記加算器のビット数に基づいて、前記ウェートを前記加算器にマッピングし、前記加算器を利用し、前記中間合算値及び前記ウェートを合算する段階と、を含むことを特徴とする請求項 1－3 のうち何れか一項に記載の方法。

【請求項 5】

前記合算する段階は、

前記加算器の全てのビット桁に値がマッピングされるように、前記調整された個別グラジエント値、前記残余グラジエント値及び前記ウェートをパディングする段階と、

前記加算器を利用し、前記パディングされた個別グラジエント値、前記パディングされた中間合算値、及び前記パディングされたウェートを合算する段階と、を含むことを特徴とする請求項 4 に記載の方法。

【請求項 6】

前記アップデートする段階は、

前記ウェートのビット桁に対応する前記合算結果のビット桁の値を前記ウェートにアップデートし、前記ウェートのビット桁に対応しない前記合算結果のビット桁の値を前記残余グラジエント値にアップデートする段階を含むことを特徴とする請求項 1－5 のうち何れか一項に記載の方法。

【請求項 7】

前記方法は、

前記合算結果の最上位ビット（MSB）である符号ビットを獲得する段階と、

前記符号ビットが前記アップデートされたウェイト、及び前記アップデートされた残余グラジエント値のうち少なくともいずれか1つの最上位ビットになるように追加する段階と、をさらに含むことを特徴とする請求項1－6のうち何れか一項に記載の方法。

【請求項8】

低精度ナンバーシステムを利用するニューラルネットワークにおいて、パラメータを処理する方法において、

前記ニューラルネットワークのウェイトをアップデートするための個別グラジエント値を計算する段階と、

前記個別グラジエント値を累積した累積グラジエント値、及び前記ウェイトのビット桁に基づいて、残余グラジエント値を計算する段階と、

前記残余グラジエント値のビット桁に対応するように、前記個別グラジエント値を調整する段階と、

前記残余グラジエント値から符号ビットを除外した残りの値を前記ウェイトに連結し、中間連結値を計算する段階と、

加算器を利用し、前記調整された個別グラジエント値、及び前記中間連結値を合算する段階と、

前記合算結果に基づいて、前記ウェイト及び前記残余グラジエント値をアップデートする段階と、を含む方法。

【請求項9】

前記残余グラジエント値を計算する段階は、

前記累積グラジエント値のうち、前記ウェイトのビット桁に合算可能な値を有効グラジエント値と決定する段階と、

前記累積グラジエント値から前記有効グラジエント値を減算することにより、前記残余グラジエント値を計算する段階と、を含むことを特徴とする請求項8に記載の方法。

【請求項10】

前記個別グラジエント値を調整する段階は、

前記個別グラジエント値のうち、前記残余グラジエント値の最下位ビット桁未満の値が省略されるように、前記個別グラジエント値を量子化する段階と、

前記残余グラジエント値の最上位ビット桁に対応するビット桁まで値が存在するように、前記量子化された個別グラジエント値をパディングする段階と、を含むことを特徴とする請求項8又は9に記載の方法。

【請求項11】

前記合算する段階は、

前記加算器のビット数に基づいて、前記調整された個別グラジエント値、及び前記中間連結値を前記加算器にマッピングし、前記加算器を利用し、前記調整された個別グラジエント値、及び前記中間連結値を合算する段階を含むことを特徴とする請求項8－10のうち何れか一項に記載の方法。

【請求項12】

前記合算する段階は、

前記加算器の全てのビット桁に値がマッピングされるように、前記調整された個別グラジエント値、及び前記中間連結値をパディングする段階と、

前記加算器を利用し、前記パディングされた個別グラジエント値、及び前記パディングされた中間連結値を合算する段階と、を含むことを特徴とする請求項11に記載の方法。

【請求項13】

前記アップデートする段階は、

前記ウェイトのビット桁に対応する前記合算結果のビット桁の値を前記ウェイトにアップデートし、前記ウェイトのビット桁に対応しない前記合算結果のビット桁の値を前記残余グラジエント値にアップデートする段階と、を含むことを特徴とする請求項8－12の

うち何れか一項に記載の方法。

【請求項 14】

前記方法は、

前記合算結果の最上位ビット（MSB）である符号ビットを獲得する段階と、

前記符号ビットが前記アップデートされたウェート、及び前記アップデートされた残余グラジエント値のうち少なくともいずれか 1 つの最上位ビットになるように追加する段階と、をさらに含むことを特徴とする請求項 8－13 のうち何れか一項に記載の方法。

【請求項 15】

低精度ナンバーシステムを利用するニューラルネットワークを実施する装置において、少なくとも 1 つのプログラムが保存されたメモリと、

前記少なくとも 1 つのプログラムを実行することにより、前記ニューラルネットワークにおいて、パラメータを処理する少なくとも 1 つのプロセッサと、を含み、

前記少なくとも 1 つのプロセッサは、

前記ニューラルネットワークのウェートをアップデートするための個別グラジエント値を計算し、前記個別グラジエント値を累積した累積グラジエント値、及び前記ウェートのビット桁に基づいて、残余グラジエント値を計算し、前記残余グラジエント値のビット桁に対応するように、前記個別グラジエント値を調整し、加算器を利用し、前記調整された個別グラジエント値、前記残余グラジエント値及び前記ウェートを合算し、前記合算結果に基づいて、前記ウェート及び前記残余グラジエント値をアップデートする装置。

【請求項 16】

低精度ナンバーシステムを利用するニューラルネットワークを実施する装置において、少なくとも 1 つのプログラムが保存されたメモリと、

前記少なくとも 1 つのプログラムを実行することにより、前記ニューラルネットワークにおいて、パラメータを処理する少なくとも 1 つのプロセッサと、を含み、

前記少なくとも 1 つのプロセッサは、

前記ニューラルネットワークのウェートをアップデートするための個別グラジエント値を計算し、前記個別グラジエント値を累積した累積グラジエント値、及び前記ウェートのビット桁に基づいて、残余グラジエント値を計算し、前記残余グラジエント値のビット桁に対応するように、前記個別グラジエント値を調整し、前記残余グラジエント値から符号ビットを除外した残りの値を前記ウェートに連結し、中間連結値を計算し、加算器を利用し、前記調整された個別グラジエント値、及び前記中間連結値を合算し、前記合算結果に基づいて、前記ウェート及び前記残余グラジエント値をアップデートする装置。

【請求項 17】

請求項 1 ないし 8 のうちいずれか 1 項に記載の方法をコンピュータで実行するためのプログラムを記録したコンピュータで読み取り可能な記録媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ニューラルネットワークにおいてパラメータを処理する方法及び装置に関する。

【背景技術】

【0002】

ニューラルネットワーク（neural network）は、生物学的脳をモデリングしたコンピュータ科学的アーキテクチャ（computational architecture）を参照する。最近、ニューラルネットワーク技術の発展により、多種の電子システムにおいて、ニューラルネットワーク装置を使用して入力データを分析し、有効な情報を抽出する研究が活発に進められている。

【0003】

ニューラルネットワーク装置は、複雑な入力データに係わる多量の演算を必要とする。該ニューラルネットワーク装置が入力をリアルタイムで分析し、情報抽出のために、

10

20

30

40

50

ニューラルネットワーク演算を効率的に処理することができる技術が要求される。例えば、該ニューラルネットワーク装置の複雑な入力データの処理に必要な演算量を低減させながらも、精度損失を最小化させることができる技術が要求される。

【先行技術文献】

【特許文献】

【0004】

【特許文献1】特開2012-198804号公報

【特許文献2】特開2012-208843号公報

【発明の概要】

【発明が解決しようとする課題】

10

【0005】

本発明が解決しようとする課題は、ニューラルネットワークにおいて、パラメータを処理する方法及びその装置を提供するところにある。本実施形態がなす技術的課題は、前述のような技術的課題に限定されるものではなく、以下の実施形態から、他の技術的課題も類推される。

【課題を解決するための手段】

【0006】

前述の技術的課題を達成するための技術的手段として、本開示の第1側面は、低精度ナンバーシステム (low-precision number system) を利用するニューラルネットワークにおいて、パラメータを処理する方法において、前記ニューラルネットワークのウェートをアップデートするための個別グラジエント値を計算する段階と、前記個別グラジエント値を累積した累積グラジエント値、及び前記ウェートのビット桁に基づいて、残余グラジエント値を計算する段階と、前記残余グラジエント値のビット桁に対応するように、前記個別グラジエント値を調整する段階と、加算器を利用し、前記調整された個別グラジエント値、前記残余グラジエント値及び前記ウェートを合算する段階と、前記合算結果に基づいて、前記ウェート及び前記残余グラジエント値をアップデートする段階と、を含む方法を提供することができる。

20

【0007】

また、前記累積グラジエント値のうち、前記ウェートのビット桁に合算可能な値を有効グラジエント値と決定する段階と、前記累積グラジエント値から前記有効グラジエント値を減算することにより、前記残余グラジエント値を計算する段階と、を含む方法を提供することができる。

30

【0008】

また、前記個別グラジエント値のうち、前記残余グラジエント値の最下位ビット桁未満の値が省略されるように、前記個別グラジエント値を量子化する段階と、前記残余グラジエント値の最上位ビット (MSB: most significant bit) 桁に対応するビット桁まで値が存在するように、前記量子化された個別グラジエント値をパディング (padding) する段階と、を含む、方法を提供することができる。

【0009】

また、前記加算器のビット数に基づいて、前記調整された個別グラジエント値及び前記残余グラジエント値を前記加算器にマッピングし、前記加算器を利用して中間合算値を計算する段階と、前記加算器のビット数に基づいて、前記ウェートを前記加算器にマッピングし、前記加算器を利用し、前記中間合算値及び前記ウェートを合算する段階と、を含む方法を提供することができる。

40

【0010】

また、前記加算器の全てのビット桁に値がマッピングされるように、前記調整された個別グラジエント値、前記残余グラジエント値及び前記ウェートをパディングする段階と、前記加算器を利用し、前記パディングされた個別グラジエント値、前記パディングされた中間合算値、及び前記パディングされたウェートを合算する段階と、を含む方法を提供することができる。

50

【0011】

また、前記ウェートのビット桁に対応する前記合算結果のビット桁の値を前記ウェートにアップデートし、前記ウェートのビット桁に対応しない前記合算結果のビット桁の値を前記残余グラジエント値にアップデートする段階と、を含む方法を提供することができる。

【0012】

また、前記合算結果の最上位ビットである符号ビットを獲得する段階と、前記符号ビットが前記アップデートされたウェート、及び前記アップデートされた残余グラジエント値のうち少なくともいずれか1つの最上位ビットになるように追加する段階と、をさらに含む方法を提供することができる。

【0013】

本開示の第2側面は、低精度ナンバーシステムを利用するニューラルネットワークにおいて、パラメータを処理する方法において、前記ニューラルネットワークのウェートをアップデートするための個別グラジエント値を計算する段階と、前記個別グラジエント値を累積した累積グラジエント値、及び前記ウェートのビット桁に基づいて、残余グラジエント値を計算する段階と、前記残余グラジエント値のビット桁に対応するように、前記個別グラジエント値を調整する段階と、前記残余グラジエント値から符号ビットを除外した残りの値を前記ウェートに連結（concatenation）して中間連結値を計算する段階と、加算器を利用し、前記調整された個別グラジエント値、及び前記中間連結値を合算する段階と、前記合算結果に基づいて、前記ウェート及び前記残余グラジエント値をアップデートする段階と、を含む、方法を提供することができる。

【0014】

本開示の第3側面は、低精度ナンバーシステムを利用するニューラルネットワークを実施（implement）する装置において、少なくとも1つのプログラムが保存されたメモリと、及び前記少なくとも1つのプログラムを実行することにより、前記ニューラルネットワークにおいて、パラメータを処理する少なくとも1つのプロセッサを含み、前記少なくとも1つのプロセッサは、前記ニューラルネットワークのウェートをアップデートするための個別グラジエント値を計算し、前記個別グラジエント値を累積した累積グラジエント値、及び前記ウェートのビット桁に基づいて、残余グラジエント値を計算し、前記残余グラジエント値のビット桁に対応するように、前記個別グラジエント値を調整し、加算器を利用し、前記調整された個別グラジエント値、前記残余グラジエント値及び前記ウェートを合算し、前記合算結果に基づいて、前記ウェート及び前記残余グラジエント値をアップデートする装置を提供することができる。

【0015】

本開示の第4側面は、低精度ナンバーシステムを利用するニューラルネットワークを実施する装置において、少なくとも1つのプログラムが保存されたメモリ；及び前記少なくとも1つのプログラムを実行することにより、前記ニューラルネットワークにおいて、パラメータを処理する少なくとも1つのプロセッサを含み、前記少なくとも1つのプロセッサは、前記ニューラルネットワークのウェートをアップデートするための個別グラジエント値を計算し、前記個別グラジエント値を累積した累積グラジエント値、及び前記ウェートのビット桁に基づいて、残余グラジエント値を計算し、前記残余グラジエント値のビット桁に対応するように、前記個別グラジエント値を調整し、前記残余グラジエント値から符号ビットを除外した残りの値を前記ウェートに連結し、中間連結値を計算し、加算器を利用し、前記調整された個別グラジエント値、及び前記中間連結値を合算し、前記合算結果に基づいて、前記ウェート及び前記残余グラジエント値をアップデートする装置を提供することができる。

【0016】

本開示の第5側面は、第1側面の方法をコンピュータで実行させるためのプログラムを記録したコンピュータで読み取り可能な記録媒体を提供することができる。

【図面の簡単な説明】

【0017】

【図1】一実施形態によるニューラルネットワークにおいて、浮動小数点フォーマットのパラメータを、固定小数点フォーマットのパラメータに量子化する例示について説明する図面である。

【図2】一実施形態によるニューラルネットワークで行われる演算について説明するための図面である。

【図3A】一実施形態による浮動小数点値の例示を図示した図面である。

【図3B】一実施形態による固定小数点値の一例を図示した図面である。

【図3C】一実施形態による固定小数点値の他の例を図示した図面である。

【図4】一実施形態による、ニューラルネットワークのパラメータがアップデートされる過程について説明するための図面である。

【図5】一実施形態によるニューラルネットワークにおいて、パラメータを処理する方法の一例を示すフローチャートである。

【図6】一実施形態による、固定小数点パラメータ形式のウェイトがアップデートされる例示について説明するための図面である。

【図7】一実施形態による、ウェイト及び残余グラジエント値をアップデートする第1例示を示す図面である。

【図8】一実施形態による、ウェイト及び残余グラジエント値をアップデートする第2例示を示す図面である。

【図9】一実施形態による、加算器を利用し、ウェイト及び残余グラジエント値をアップデートする第1例示を示す図面である。

【図10】一実施形態による、加算器を利用し、ウェイト及び残余グラジエント値をアップデートする第2例示を示す図面である。

【図11】一実施形態による、加算器を利用し、ウェイト及び残余グラジエント値をアップデートする第3例示を示す図面である。

【図12】一実施形態による、加算器を利用し、ウェイト及び残余グラジエント値をアップデートする第4例示を示す図面である。

【図13】一実施形態による、ウェイト及び残余グラジエント値をアップデートする方法について説明するためのフローチャートである。

【図14】一実施形態によるニューラルネットワーク装置のハードウェア構成を図示したブロック図である。

【発明を実施するための形態】

【0018】

本明細書において、多様なところに記載される「一部実施形態において」または「一実施形態において」というような語句は、必ずしもいずれも同一実施形態を示すものではない。

【0019】

本開示の一部実施形態は、機能的なブロック構成、及び多様な処理段階によっても示される。そのような機能ブロックの一部または全部は、特定機能を実行する多様な個数のハードウェア構成及び／またはソフトウェア構成によっても具現される。例えば、本開示の機能ブロックは、1以上のマイクロプロセッサによっても具現されるか、あるいは所定機能のための回路構成によっても具現される。また、例えば、本開示の機能ブロックは、多様なプログラミング言語またはスクリプティング言語によっても具現される。該機能ブロックは、1以上のプロセッサで実行されるアルゴリズムによっても具現される。また、本開示は、電子的な環境設定、信号処理、及び／またはデータ処理などのために、従来技術を採用することができる。「メカニズム」、「要素」、「手段」及び「構成」のような用語は、汎用され、機械的であって物理的な構成として限定されるものではない。

【0020】

また、図面に図示された構成要素間の連結線または連結部材は、機能的な連結及び／または物理的または回路的な連結を例示的に示したものに過ぎない。実際の装置においてはは

10

20

30

40

50

、代替可能であったり、追加されたりする多様な機能的な連結、物理的な連結または回路連結によっても、構成要素間の連結が示される。

【0021】

以下、図面を参照し、本発明の実施形態について詳細に説明する。

【0022】

図1は、一実施形態によるニューラルネットワークにおいて、浮動小数点フォーマットのパラメータを、固定小数点フォーマットのパラメータに量子化する例示について説明する図面である。

【0023】

図1を参照すれば、ニューラルネットワーク学習器10は、ニューラルネットワークを生成するか、あるいはニューラルネットワークを訓練する(train) (または、学習させる(learn))か、浮動小数点フォーマットのニューラルネットワークを、固定小数点フォーマットのニューラルネットワークに量子化するか、あるいはニューラルネットワークを再訓練(retrain)する機能のような多様なプロセッシング機能を有するコンピュータデバイスに該当する。例えば、ニューラルネットワーク学習器10は、PC (personal computer)、サーバデバイス、モバイルデバイスのような多種のデバイスによっても具現される。

【0024】

ニューラルネットワーク学習器10は、与えられた初期ニューラルネットワークを反復的に訓練する(学習させる)ことにより、訓練されたニューラルネットワーク11を生成することができる。このとき、該初期ニューラルネットワークは、ニューラルネットワークの処理精度を確保する観点から、浮動小数点フォーマットのパラメータ、例えば、32ビット浮動小数点精度(32bit floating point precision)のパラメータを有することができる。ここで、該パラメータは、例えば、ニューラルネットワークの入出力アクティベーション、ウェイト、バイアスのような、ニューラルネットワークに入出力される多種のデータを含んでもよい。ニューラルネットワークの反復的な訓練が進められることにより、ニューラルネットワークの浮動小数点パラメータは、与えられた入力に対して、さらに正確な出力を演算するためにも調整される(tuned)。

【0025】

ニューラルネットワーク学習器10は、訓練されたニューラルネットワーク11を、ニューラルネットワーク推論器20のようなハードウェアアクセラレータに伝達することができる。ニューラルネットワーク推論器20は、モバイルデバイス、埋め込み(embedded)デバイスなどにも含まれる。ニューラルネットワーク推論器20は、ニューラルネットワーク21の駆動のための専用ハードウェアであり、比較的、低電力または低性能に具現されるために、浮動小数点演算よりは、固定小数点演算において、さらに適して具現される。一方、ニューラルネットワーク推論器20は、ニューラルネットワーク駆動のための専用モジュールであるTPU (tensor processing unit)、ニューラルエンジン(Neural Engine)などに該当するが、それらに制限されるものではない。

【0026】

量子化されたニューラルネットワーク21を駆動するニューラルネットワーク推論器20は、ニューラルネットワーク学習器10とは別途の独立したデバイスによっても具現され得る。しかし、それに制限されるものではなく、ニューラルネットワーク推論器20は、ニューラルネットワーク学習器10と同一デバイス内にも具現される。

【0027】

一方、該浮動小数点は、該固定小数点に比べ、相対的に多くの演算量と、多くのメモリアクセス頻度とが要求される。特に、ニューラルネットワーク処理に必要な演算量のほとんどは、多様なパラメータの演算を行うMAC (multiplication and accumulation) 演算であることが知られている。従って、比較的処理性能が低いスマートフォン、タブレット、ウェアラブルデバイスなどに含まれたニューラルネットワーク推論器20においては、浮動小数点フォーマットのパラメータを使用するニューラルネットワークの処理

は必ずしも円滑なものではない。結局、そのようなデバイスにおいて、演算量を十分に低減させながら、許容可能な精度損失内において、ニューラルネットワークを駆動させるためには、ニューラルネットワークで処理される浮動小数点フォーマットのパラメータは、量子化されることが望ましい。ここで、該パラメータ量子化は、固定精度の浮動小数点フォーマットのパラメータを、低精度の固定小数点フォーマットのパラメータに変換することを意味する。

【0028】

該ニューラルネットワーク量子化装置は、訓練されたニューラルネットワーク11のパラメータを、所定ビットの固定小数点フォーマットに変換する量子化を行う。該ニューラルネットワーク量子化装置は、ニューラルネットワーク学習器10、ニューラルネットワーク推論器20に含まれて動作するか、あるいは独立した第3のハードウェアアクセラレータとして動作することができる。該ニューラルネットワーク量子化装置は、ニューラルネットワークが採用される（deployed）ニューラルネットワーク推論器20のデバイスの処理性能を考慮し、訓練されたニューラルネットワーク11のパラメータを、所定ビットの固定小数点フォーマットに変換する量子化を行い、量子化されたニューラルネットワーク21を採用するニューラルネットワーク推論器20に伝達することができる。

【0029】

量子化されたニューラルネットワーク21が採用されるニューラルネットワーク推論器20は、具体的な例示として、ニューラルネットワークを利用した音声認識、映像認識などを行う自律走行自動車、ロボティクス、スマートフォン、タブレットデバイス、AR（augmented reality）デバイス、IoT（internet of things）デバイスなどにも含まれるが、それらに制限されるものではない。

【0030】

図2は、一実施形態によるニューラルネットワークで行われる演算について説明するための図面である。図2を参照すれば、ニューラルネットワーク2は、入力レイヤ、隠れ層及び出力レイヤを含む構造を有し、受信される入力データ（例えば、 I_1 及び I_2 ）を基に演算を行い、遂行結果を基に、出力データ（例えば、 O_1 及び O_2 ）を生成することができる。

【0031】

ニューラルネットワーク2は、前述のように、2層以上の隠れ層を含むDNN（deep neural network）またはn階層ニューラルネットワークでもある。例えば、図2に図示されているように、ニューラルネットワーク2は、入力レイヤ（Layer 1）、2層の隠れ層（Layer 2 及び Layer 3）、及び出力レイヤ（Layer 4）を含むDNNであってもよい。ニューラルネットワーク2がDNNアーキテクチャによって具現された場合、有効な情報を処理することができるさらに多いレイヤを含むので、ニューラルネットワーク2は、シングルレイヤを有するニューラルネットワークより複雑なデータ集合を処理することができる。一方、ニューラルネットワーク2は、4層のレイヤを含むように図示されているが、それは、例示に過ぎず、ニューラルネットワーク2は、さらに少なかったり多かたりするレイヤを含むか、あるいはさらに少なかったり多かたりするノードを含んでもよい。すなわち、ニューラルネットワーク2は、図2に図示されたところとは異なる、多様な構造のレイヤを含んでもよい。

【0032】

ニューラルネットワーク2に含まれたレイヤそれぞれは、複数のノードを含んでもよい。該ノードは、ニューロン（neuron）、プロセッシングエレメント（PE：processing element）、ユニット（unit）、またはそれらと類似した用語として知られた、複数の人工ノード（artificial node）にも該当する。例えば、図2に図示されているように、Layer 1は、2個のノード、Layer 2 及び Layer 3 それぞれは、3個のノードを含んでもよい。ただし、それらは、例示に過ぎず、ニューラルネットワーク2に含まれたレイヤそれぞれは、多様な個数のノードを含んでもよい。

【0033】

10

20

30

40

50

ニューラルネットワーク 2 のレイヤそれぞれに含まれたノードは、互いに連結されてデータを処理することができる。例えば、1つのノードは、他のノードからデータを受信して演算することができ、該演算結果をさらに他のノードに出力することができる。

【0034】

ノードそれぞれの入力及び出力のそれぞれは、入力アクティベーション及び出力アクティベーションともされる。すなわち、該アクティベーションは、1ノードの出力であると共に、次のレイヤに含まれたノードの入力に該当するパラメータでもある。一方、ノードそれぞれの、以前レイヤに含まれたノードから受信されたアクティベーション、ウェイト及びバイアスに基づいて、自体のアクティベーションを決定することができる。該ウェイトは、各ノードでの出力アクティベーションを計算するために利用されるパラメータであり、ノード間の連結関係に割り当てられる値でもある。

10

【0035】

該ノードそれぞれの、入力を受信して出力アクティベーションを出力する演算ユニット (computational unit) またはプロセッシングエレメント (processing element) によっても処理され、ノードそれぞれの入力・出力は、マッピングされる。例えば、 σ は、アクティベーション関数 (activation function) であり、 w_{jk}^i は、 $(i-1)$ 番目レイヤに含まれた k 番目ノードから、 i 番目レイヤに含まれた j 番目ノードへのウェイトであり、 b_j^i は、 i 番目レイヤに含まれた j 番目ノードのバイアスであり、 a_j^i は、 i 番目レイヤの j 番目ノードのアクティベーションであるとするとき、該アクティベーション a_j^i は、次のような数式 (1) を利用しても計算される。

20

【0036】

【数 1】

$$a_j^i = \sigma \left(\sum_k (w_{jk}^i \times a_k^{i-1}) + b_j^i \right) \quad (1)$$

図 2 に図示されているように、2 番目レイヤ (Layer 2) の最初ノード CH1 のアクティベーションは、 a_{11}^2 とも表現される。また、 a_{11}^2 は、数式 (1) により、 $a_{11}^2 = \sigma (w_{11}^2 \times a_{11}^1 + w_{12}^2 \times a_{12}^1 + b_{11}^2)$ の値を有することができる。ただし、前述の数式 (1) は、ニューラルネットワーク 2 において、データを処理するために利用されるアクティベーション、ウェイト及びバイアスについて説明するための例示であるのみ、それに制限されるものではない。該アクティベーションは、以前レイヤから受信されたアクティベーションの加重値和 (weighted sum) を sigmoid 関数や ReLU (Rectified Linear Unit) 関数などのアクティベーション関数に通過させることによって獲得された値でもある。

30

【0037】

一方、ニューラルネットワーク 2 は、低精度ナンバーシステム (low-precision number system) を利用することができる。低精度ナンバーシステムは、従来のニューラルネットワークより相対的に少ないビット個数を有する値をパラメータとして利用することにより、精度 (precision) を低くしながら、演算量を低減させるシステムを意味する。例えば、32ビットまたは64ビットの浮動小数点または固定小数点値を利用する従来のニューラルネットワークとは異なり、低精度ナンバーシステムを利用するニューラルネットワーク 2 は、8ビットまたは16ビットの浮動小数点または固定小数点値をパラメータとして利用することができる。

40

【0038】

ニューラルネットワーク 2 が、8ビットまたは16ビットの浮動小数点または固定小数点値を利用するということは、ニューラルネットワーク 2 のアクティベーション値、ウェイト及びノードのバイアス値などのパラメータが、8ビットまたは16ビットの浮動小数点または固定小数点形式を有するということを意味する。ただし、低精度ナンバーシステ

50

ムを利用するニューラルネットワーク 2 は、さらに少ないビット個数を有する浮動小数点または固定小数点データをパラメータとして利用することもでき、前述の例示に制限されるものではない。

【0039】

ニューラルネットワーク 2 においては、多くのデータ集合が相互連結された複数のノード間で交換され、レイヤを過ぎながら、多くの演算過程を経る。低精度ナンバーシステムを利用するニューラルネットワーク 2 は、多くの演算過程において、従来のニューラルネットワークより少ないビット個数を有するデータを利用するので、従来のニューラルネットワークと比較し、演算量が低減される。ただし、該低精度ナンバーシステムを利用するニューラルネットワーク 2 は、演算量を低減させながらも、精度損失を最小化させることができる技術が要求される。

10

【0040】

図 3 A は、一実施形態による浮動小数点値の例示を図示した図面であり、図 3 B は、一実施形態による固定小数点値の一例を図示した図面であり、図 3 C は、一実施形態による固定小数点値の他の例を図示した図面である。

【0041】

図 3 A を参照すれば、浮動小数点値 30 は、符号ビット 310、指数部 320、仮数部 330 及びバイアス 340 から構成される。該浮動小数点は、1 つの数を、小数点部分を示す部分と小数点位置を示す部分とに分けて表現する表記法である。

【0042】

20

仮数部 330 は、小数点部分を示す部分に対応し、指数部 320 は、小数点位置を示す部分に対応する。符号ビット 310 は、浮動小数点値 30 の符号を決定することができる。バイアス 340 は、指数部 320 に加減される値であり、負の指数を表現するために決定される値でもある。浮動小数点値 30 は、符号ビット 310、指数部 320 に対応するビット、及び仮数部 330 に対応するビットを含んでもよい。バイアス 340 は、浮動小数点値 30 に対して事前に決定され、別途に保存される。

【0043】

一方、浮動小数点値 30 は、符号ビット 310 を sign、指数部 320 を exponent、仮数部 330 を mantissa、バイアス 340 を bias とするとき、次のような数式 (2) による値を有することができる。

30

【0044】

【数 2】

$$\text{Floating point value} = (-1)^{\text{sign}} \cdot 2^{\text{exponent} - \text{bias}} \cdot \text{mantissa} \quad (2)$$

図 3 B を参照すれば、固定小数点値 35 は、符号ビット 315、整数部 325、小数部 335 及び小数点 345 から構成される。該固定小数点は、小数点を使用し、固定された桁数の小数を示す表記法を意味する。

40

【0045】

符号ビット 315 は、固定小数点値 35 の符号を決定し、整数部 325 は、固定小数点値 35 の整数を示す部分に対応し、小数部 335 は、固定小数点値 35 の小数を示す部分に対応する。小数点 345 は、固定小数点値 35 の整数部 325 及び小数部 335 を区分する基準になる点を意味する。

【0046】

一方、固定小数点値 35 が示す値は、図 3 C を参照しても説明される。図 3 C を参照すれば、固定小数点値 35 は、8 ビット固定小数点値でもある。また、固定小数点値 35 は、整数部 325、小数部 335 及び小数点 345 から構成される。

50

【0047】

整数部325及び小数部335を示すビットそれぞれは、1または0の値を有することができる。また、図3Cに図示されているように、整数部325及び小数部335を示すビットは、順に−8、+4、+2、+1、+0.5、+0.25、+0.125及び+0.0625の値を示すことができる。整数部325の最上位ビット(MSB:most significant bit)が1である場合、整数部325の最上位ビットが示す値は、−8であるので、整数部325及び小数部335に含まれる他のビットがいかなる値を有するにしても、固定小数点値35が示す値は、負数である。整数部325の最上位ビットは、固定小数点値35の符号を決定する符号ビット315に対応する。

【0048】

ただし、図3Cは例示に過ぎず、固定小数点値35は、16ビット固定小数点値でもあり、任意の適切な数のビット個数を有する固定小数点値でもある。また、固定小数点値35は、符号化表現法、1の補数法及び2の補数法のうちいずれか1つの表現法によって負数を示すことができる。

【0049】

図4は、一実施形態による、ニューラルネットワークのパラメータがアップデートされる過程について説明するための図面である。

【0050】

ニューラルネットワークのパラメータは、ニューラルネットワークに含まれるノード間の連結関係に割り当てられるウェイトでもある。以下において、該ニューラルネットワークで処理されるパラメータをウェイトとして説明するが、該パラメータは、必ずしもウェイトに限定されるものではない。該パラメータは、ニューラルネットワークで処理される任意のデータを意味する。

【0051】

個別グラジエント値(m)は、ウェイト(θ)をアップデートするために計算される値でもある。該ニューラルネットワークにおいて学習を行う対象は、ウェイト(θ)であり、該ニューラルネットワークがだんだんにさらに良好な結果を出力するためには、ウェイト(θ)が学習過程において、反復的にアップデートされる必要がある。

【0052】

一方、ウェイト(θ)を効果的にアップデートするためには、ネットワークで出す結果値と、実際の結果値との差を定義する損失関数(J(θ))の値を最小化させる必要があり、損失関数(J(θ))の値を最小化させるために、多様なOptimizerが利用される。

【0053】

一実施形態においてOptimizerは、グラジエントディセント(gradient descent)方式を利用することができる。該グラジエントディセント方式は、ウェイト(θ)に対して、グラジエントの反対方向に、一定サイズほど移動することを反復し、損失関数(J(θ))の値を最小化させるθ値を求める方式である。該グラジエントディセント方式は、下記数式(3)のように表現することができる。数式(3)において、 $\nabla_{\theta} J(\theta)$ は、損失(J(θ))の勾配であり、ηは、学習速度である。

【0054】

【数3】

$$\theta \leftarrow \theta - \eta \nabla_{\theta} J(\theta) \quad (3)$$

特に、損失関数(J(θ))を計算するとき、全体学習データを使用することを、batch gradient descentと言うが、その場合、過度に多くの計算量が必要である。それを防止するために、全体データ(batch)の代わりに、一部小データの集まり(mini-batch)

10

20

30

40

50

についてのみ損失関数（ $J(\theta)$ ）を計算するSGD（stochastic gradient descent）方法が使用される。

【0055】

図4を参照すれば、オプティマイザ（Optimizer）がグラジエントディセント方式を利用する場合、個別グラジエント値（ m ）は、 $\eta \nabla_{\theta} J(\theta)$ になる。

【0056】

他の実施形態において、Optimizerは、モーメンタム（momentum）方式を利用することができる。該モーメンタム方式は、グラジエントディセントを介して移動する過程において、「慣性」を与えるものである。現在グラジエントを介して、移動する方向とは別個に、過去に移動した方式を記憶しながら、その方向に一定程度追加して移動する方式である。該モーメンタム方式は、下記数式（4）のように表現することができる。数式（4）において、 v_t は、時間ステップ t での移動ベクトルであり、 γ は、いかほどモーメンタムを与えるかということに係わるモーメンタム項（momentum term）である。

【0057】

【数4】

$$\begin{aligned} v_t &= \gamma v_{t-1} + \eta \nabla_{\theta} J(\theta) \\ \theta &\leftarrow \theta - v_t \end{aligned} \quad (4)$$

図4を参照すれば、Optimizerがモーメンタム方式を利用する場合、個別グラジエント値（ m ）は、 $\gamma v_{t-1} + \eta \nabla_{\theta} J(\theta)$ になる。

【0058】

Optimizerは、前述のSGDグラジエントディセント方式及びモーメンタム方式の以外にも、Adagrad（Adaptive Gradient）方式、RMSProp方式、AdaDelta（Adaptive Delta）方式、Adam（Adaptive Moment Estimation）方式などを利用することができる。

【0059】

一方、一般的なニューラルネットワークにおいて、個別グラジエント値（ m ）は、ウェイト（ θ ）よりはるかに小さい値を有する。該ニューラルネットワークが低精度ナンバーシステム（例えば、8ビットまたは16ビット）を利用する場合、個別グラジエント値（ m ）は、ウェイト（ θ ）が示すことができる値の範囲に含まれない。それにより、ウェイト（ θ ）が多様なアップデート方式（SGDグラジエントディセント方式、モーメンタム方式など）によってアップデートされるとき、個別グラジエント値（ m ）は、無視され、ウェイト（ θ ）は、アップデートされない。該ニューラルネットワークがだんだんとさらに良好な結果を出力するためには、ウェイト（ θ ）が反復してアップデートされなければならないにもかかわらず、個別グラジエント値（ m ）の大きさにより、ウェイトがアップデートされない。

【0060】

以下では、説明の便宜上、Optimizerがグラジエントディセント方式を利用することに限定して説明する。

【0061】

図5は、一実施形態によるニューラルネットワークにおいて、パラメータを処理する方法の一例を示すフローチャートである。

【0062】

該ニューラルネットワークは、低精度ナンバーシステムを利用するニューラルネットワークでもある。パラメータは、ニューラルネットワークに含まれるノード間の連結関係に割り当てられるウェイトでもある。以下において、ニューラルネットワークで処理される

パラメータをウェイトとして説明するが、該パラメータは、必ずしもウェイトに限定されるものではない。該パラメータは、ニューラルネットワークで処理される任意のデータを意味する。

【0063】

図5を参照すれば、段階510において、ニューラルネットワーク装置は、ニューラルネットワークに含まれるノード間の連結関係に割り当てられるウェイトのための個別グラジエント (individual gradient) 値を計算することができる。該ウェイトは、浮動小数点値または固定小数点値でもある。

【0064】

該個別グラジエント値は、ウェイトをアップデートするために計算される値でもある。該ニューラルネットワークで学習を行う対象は、ウェイトであり、該ニューラルネットワークがだんだんとさらに良好な結果を出力するためには、該ウェイトが、学習過程において反復してアップデートされる必要がある。該ウェイトをアップデートするための方法のうち一つとして、例えば、傾斜下降法 (gradient descent algorithm) が利用される。該傾斜下降法は、損失関数または誤差関数の最低点を求めるために、損失関数または誤差関数の勾配方向にパラメータを調整することにより、誤差を最小化させるアルゴリズムを意味する。例えば、損失関数が L であり、該パラメータが θ であり、学習速度 (learning rate) が η であるとするとき、該パラメータは、次のような数式 (5) によっても調整される。

【0065】

【数5】

$$\theta \leftarrow \theta - \eta \frac{\partial L}{\partial \theta} \quad (5)$$

パラメータをアップデートするために計算される $\eta (\partial L / \partial \theta)$ は、損失関数の勾配方向にパラメータを調整するためのグラジエントを意味する。例えば、パラメータがウェイトであるとき、 $\eta (\partial L / \partial \theta)$ は、ウェイトをアップデートするために計算される個別グラジエント値でもある。

【0066】

一方、一般的なニューラルネットワークにおいて、個別グラジエント値 $\eta (\partial L / \partial \theta)$ は、ウェイトよりはるかに小さい値を有する。ニューラルネットワークが、低精度ナバーシシステム (例えば、8ビットまたは16ビット) を利用する場合、個別グラジエント値 $\eta (\partial L / \partial \theta)$ は、ウェイトが示すことができる値の範囲に含まれないかもしれない。それにより、該ウェイトが、数式 (5) によってアップデートされるとき、個別グラジエント値 $\eta (\partial L / \partial \theta)$ は、無視され、該ウェイトは、アップデートされないことが懸念される。該ニューラルネットワークがだんだんとさらに良好な結果を出力するためには、ウェイトが反復してアップデートされなければならないにもかかわらず、個別グラジエント値 $\eta (\partial L / \partial \theta)$ の大きさに起因して、該ウェイトがアップデートされない場合がある。

【0067】

本開示によるニューラルネットワーク装置は、段階520及び段階530を遂行することにより、個別グラジエント値は無視されず、ウェイトをアップデートする過程において反映される。段階520において、該ニューラルネットワーク装置は、個別グラジエント値を累積バッファ (accumulation buffer) に累積することができる。該累積バッファは、個別グラジエント値を累積するための専用ハードウェアでもあり、該個別グラジエント値を累積するために割り当てられたメモリでもある。該個別グラジエント値は、累積バッ

ファに累積し、累積グラジエント値にもなる。該累積バッファは、累積グラジエント値を保存することができる。

【0068】

低精度ナンバーシステムを利用するニューラルネットワークにおいて、個別グラジエント値がウェートをアップデートする過程において無視されないように、該ニューラルネットワーク装置は、個別アップデート値を、ウェートに影響を与えることができる大きさまで累積することができる。

【0069】

段階530において、該ニューラルネットワーク装置は、累積バッファの累積グラジエント値が臨界値以上である場合、累積グラジエント値を利用し、ウェートをアップデートすることができる。該ニューラルネットワーク装置は、累積バッファの累積グラジエント値が臨界値以上であるか否かということ判断することができる。該臨界値は、累積グラジエント値がウェートに影響を与えることができる大きさまで累積したか否かということ決定するための値でもある。例えば、該臨界値は、ウェートの最下位ビット桁に対応する値でもある。

【0070】

該ウェートが固定小数点値である場合、ウェートの最下位ビット桁は、ウェートを示すビットのうち最下位ビットに対応する桁でもある。例えば、ウェートの最下位ビットが 2^{-4} 桁を示す場合、ウェートの最下位ビット桁は、 2^{-4} 桁でもあり、臨界値は、その最下位ビット桁 2^{-4} 桁に対応する値である 2^{-4} でもある。

【0071】

該ウェートが浮動小数点値である場合、ウェートの最下位ビット桁は、ウェートの指数部を示すビットのうち最下位ビットに対応する桁でもある。ただし、該ウェートが浮動小数点値である場合には、最下位ビット桁が決定されるために、バイアスが考慮される。例えば、ウェートの指数部を示すビットのうち最下位ビットが 2^0 桁を示したとしても、バイアスが8である場合、該ウェートの最下位ビット桁は、2の (2^0) 乗 $=2^1$ 桁ではなく、 $2^1 - 8 = 2^{-7}$ 桁であり、該臨界値は、その最下位ビット桁 2^{-7} 桁に対応する値である 2^{-7} でもある。

【0072】

ただし、それは、例示に過ぎず、該臨界値は、ウェートの最下位ビット桁に対応する値と、累積グラジエント値の最上位ビット桁に対応する値との間の任意の適切な値でもある。

【0073】

該累積グラジエント値が臨界値以上である場合、該累積グラジエント値は、ウェートに影響を与えることができる大きさを有するので、該ニューラルネットワーク装置は、該累積グラジエント値を利用し、ウェートをアップデートすることができる。

【0074】

具体的には、該ニューラルネットワーク装置は、該累積グラジエント値のうちウェートのビット桁に合算可能な部分を有効グラジエント (effective gradient) 値と決定することができる。有効グラジエント値とは、該累積グラジエント値において、ウェートをアップデートするために、実際に利用される値を意味する。該ウェートが示すことができる値の範囲は、該ウェートに割り当てられたビット桁によって限界が存在する。従って、該累積グラジエント値のうちウェートの最下位ビット桁以上の値だけがウェートに合算される。

【0075】

例えば、ウェートが1ビットの符号ビット、3ビットの整数部、及び4ビットの小数部を含む8ビット固定小数点値であるとするとき、ウェートのビット桁は、 2^2 、 2^1 、 2^0 、 2^{-1} 、 2^{-2} 、 2^{-3} 及び 2^{-4} 桁でもある。該ウェートが示すことができる値の範囲の限界により、該ウェートにある値が合算されるためには、その値が 2^2 、 2^1 、 2^0 、 2^{-1} 、 2^{-2} 、 2^{-3} 及び 2^{-4} のうちいずれか1つの桁に対応する値を少なくと

も一つ含まなければならない。例えば、該ウェートに、1. 2 6 5 6 2 5（十進数）という値が合算されるべきであると仮定する。1. 2 6 5 6 2 5は、 $1 \times 2^0 + 1 \times 2^{-2} + 1 \times 2^{-6}$ であるので、該ウェートのビット桁に対応する 1×2^0 及び 1×2^{-2} は、ウェートに合算されるが、 1×2^{-6} は、ウェートが示することができる値の範囲の限界に起因して、ウェートに合算されない。

【0076】

一方、ウェートのビット桁に合算可能な部分は、ウェートの最下位ビット桁が示す値の倍数を含んでもよい。前述の例示において、該ウェートの最下位ビット桁に対応する値は、 2^{-4} であり、累積アップデート値のうち該ウェートのビット桁に合算可能な部分は、 2^2 , 2^1 , 2^0 , 2^{-1} , 2^{-2} , 2^{-3} 及び 2^{-4} 桁に対応する値である 2^2 , 2^1 , 2^0 , 2^{-1} , 2^{-2} , 2^{-3} 及び 2^{-4} である。 2^2 , 2^1 , 2^0 , 2^{-1} , 2^{-2} , 2^{-3} 及び 2^{-4} は、いずれも最下位ビット桁に対応する値である 2^{-4} の倍数である。

10

【0077】

該臨界値は、ウェートの最下位ビット桁に対応する値でもあるので、有効グラジエント値は、臨界値の倍数によっても構成される。例えば、累積グラジエント値を、Accumulated gradient、臨界値をthresholdとすると、該有効グラジエント値は、次のような数式（6）による。

【0078】

【数6】

20

$$\text{effective gradient} = \left\lfloor \frac{\text{Accumulated gradient}}{\text{threshold}} \right\rfloor \times \text{threshold} \quad (6)$$

数式（6）で、「 $\lfloor \text{Accumulated gradient} / \text{threshold} \rfloor$ 」は、Accumulated gradientをthresholdで除した商を意味する。該有効グラジエント値は、累積グラジエント値を臨界値で除した商に臨界値を乗じた値を意味する。また、該有効グラジエント値は、該累積グラジエント値より小さい値を有する、臨界値の倍数のうち最も大きい値でもある。例えば、Accumulated gradientが2. 2（十進数）であり、thresholdが1（十進数）である場合、該有効グラジエント値は、累積グラジエント値2. 2より小さい値を有する、臨界値1の倍数である1、2のうちより大きい値である2である。

30

【0079】

該ニューラルネットワーク装置は、ウェートと有効グラジエント値とを合算することができる。該有効グラジエント値は、ウェートをアップデートするために、実際に利用される値であるので、該ニューラルネットワーク装置は、該ウェートに該有効グラジエント値を合算することにより、ウェートをアップデートすることができる。

【0080】

40

また、該ニューラルネットワーク装置は、累積バッファの累積グラジエント値から有効グラジエント値を減算することができる。該累積バッファの該累積グラジエント値のうち該有効グラジエント値がウェートに反映されるので、該ニューラルネットワーク装置は、該累積バッファの該累積グラジエント値から該有効グラジエント値を減算することができる。例えば、ウェートがアップデートされることにより、該累積グラジエント値は、次のような数式（7）によっても減算される。

【0081】

【数 7】

$$\text{Accumulated gradient} \leftarrow \text{Accumulated gradient} - \left\lfloor \frac{\text{Accumulated gradient}}{\text{threshold}} \right\rfloor \times \text{threshold} \quad (7)$$

数式(7)で、 $\lfloor \text{Accumulated gradient} / \text{threshold} \rfloor \times \text{threshold}$ は、有効グラジエント値であるので、累積バッファの累積グラジエント値は、以前累積グラジエント値から有効グラジエント値が減算された値でもある。

10

【0082】

該ニューラルネットワーク装置は、ウェートをアップデートした後、累積バッファの累積グラジエント値に、個別グラジエント値を続けて累積することができ、前述の過程を反復することができる。本開示によるニューラルネットワークにおいて、パラメータの処理方法は、一般的にウェートよりはるかに小さい値を有する個別グラジエント値を無視したり捨てたりする代わりに、ウェートに影響を与えることができる大きさまで累積してウェートに適用することにより、低精度ナンバーシステムを利用するニューラルネットワークでの四捨五入誤差(rounding error)を最小化させることができる。また、本開示によるニューラルネットワークにおいて、パラメータ処理方法は、累積グラジエント値のうちウェートにアップデートされて残った部分も、続けて累積バッファに維持することにより、精度損失を最小化させることができる。

20

【0083】

一方、該ニューラルネットワーク装置は、累積アップデート値が臨界値以上であるか否かということを、既設定アップデート周期ごとに判断することができる。例えば、アップデート周期 $t=4$ 、すなわち、個別グラジエント値4個が計算される時間を意味するが、それに制限されるものではない。

【0084】

図6は、一実施形態による、固定小数点パラメータ形式のウェートがアップデートされる例示について説明するための図面である。

30

【0085】

図6を参照すれば、固定小数点パラメータを利用するニューラルネットワークにおいて、ウェート60は、第1符号ビット610、第1整数部620、第1小数部630及び第1小数点640から構成された固定小数点値であり、累積グラジエント値65は、第2符号ビット615、第2小数部635及び第2小数点(小数点)から構成された固定小数点値でもある。一実施形態によれば、ウェート60は、第1小数点640が移動可能な動的固定小数点(dynamic fixed point)値でもある。第1小数点640が移動されることにより、第1小数部630のビット個数(長さ)が調整される。

【0086】

累積グラジエント値65は、整数部を含まない。例えば、第1小数点640及び第2小数点645が同一である場合、累積グラジエント値65は、ウェート60に比べ、相当に小さい値のみを示すことができるので、整数部を含まず、小数部(すなわち、第2小数部635)のみを含んでもよい。一方、第2小数点は、第1小数点640と同一であってもよく、異なってもよい。第2小数点は、状況によっても調整される。

40

【0087】

ウェート60は、 m ビット(m は、任意の自然数)の固定小数点値でもあり、累積グラジエント値65は、 n ビット(n は、任意の自然数)の固定小数点値でもある。 m 及び n は、同一値でもあり、異なる値でもある。低精度ナンバーシステムを利用するニューラルネットワークにおいて、 m 及び n それぞれは、8または16の値を有することができるが、それらに制限されるものではない。

50

【0088】

累積グラジエント値65は、ウェート60と比較し、相対的に小さい値を有することができる。従って、ウェート60を示すビットそれぞれに対応する値は、累積グラジエント値65を示すビットそれぞれに対応する値より概して大きくなる。ただし、ウェート60を示すビットのうちいずれか1つのビットに対応する値は、累積グラジエント値65を示すビットのうちいずれか1つのビットに対応する値と同一でもある。ウェート60がアップデートされるためには、累積グラジエント値65の一部が、ウェート60に合算されなければならないので、ウェート60の下位ビットのうち一部は、累積グラジエント値65の上位ビットのうち一部とオーバーラップする。

【0089】

ウェート60と累積グラジエント値65との間でオーバーラップされるビット650がkビット（kは、任意の自然数）であるとするとき、該ニューラルネットワークは、実質的に（ $m+n-2-k$ ）ビットの固定小数点データを利用するところと同一効果を得ることができる。

【0090】

ウェート60と累積グラジエント値65との間でオーバーラップされるビット650は、臨界値660を基準に決定されてもよい。臨界値660は、累積グラジエント値65がウェート60に影響を与えることができる大きさまで累積されているか否かということを決断するための値でもある。例えば、臨界値660は、ウェート60の最下位ビット桁に対応する値でもある。オーバーラップされるビット650は、累積グラジエント値65を示すビットのうち、臨界値660以上の値を示すビットでもある。オーバーラップされるビット650は、ウェート60の下位ビットのうち一部とオーバーラップされるので、ウェート60とも合算される。一方、ウェート60の下位ビットは、第1小数部630を示すことができる。

【0091】

個別グラジエント値が累積され、累積グラジエント値65がウェート60に影響を与えることができる大きさまで大きくなれば、該ニューラルネットワーク装置は、累積グラジエント値65を利用し、ウェート60をアップデートすることができる。例えば、該ニューラルネットワーク装置は、累積グラジエント値65の第2小数部635を示すビットのうちウェート60の第1小数部630とオーバーラップされる少なくとも1つのビット650が示す値をウェート60と合算することができる。オーバーラップされる少なくとも1つのビット650が示す値は、有効グラジエント値に対応する。

【0092】

一方、累積グラジエント値65の第2小数部635を示すビットのうちウェート60の第1小数部630とオーバーラップされる少なくとも1つのビット650が示す値をウェート60と合算する過程において、パディング（padding）方式が適用されてもよい。

【0093】

図7は、一実施形態による、ウェート及び残余グラジエント値をアップデートする例示を示す図面である。

【0094】

図7を参照すれば、ニューラルネットワーク装置は、個別グラジエント値710、残余グラジエント値720及びウェート730を利用し、ウェート730及び残余グラジエント値720をアップデートすることができる。一実施形態において、該ニューラルネットワーク装置は、低精度ナンバーシステムを利用し、個別グラジエント値710、残余グラジエント値720及びウェート730を示すことができる。例えば、個別グラジエント値710、残余グラジエント値720及びウェート730は、8ビットまたは16ビットの浮動小数点または固定小数点値でもあるが、それらに制限されるものではない。

【0095】

該ニューラルネットワーク装置は、ニューラルネットワークに含まれるノード間の連結関係に割り当てられるウェートのための個別グラジエント値710を計算することができ

10

20

30

40

50

る。個別グラジエント値 710 は、ウェイト 730 をアップデートするために計算される値でもある。該ニューラルネットワークで学習を行う対象は、ウェイトであり、該ニューラルネットワークがだんだんとさらに良好な結果を出力するためには、ウェイト 730 が学習過程において、反復してアップデートされる必要がある。図 4 で説明したように、該ニューラルネットワーク装置の Optimizer は、グラジエントディセント方式、モーメンタム方式などを利用し、個別グラジエント値 710 を計算することができる。

【0096】

該ニューラルネットワーク装置は、個別グラジエント値 710 を累積バッファに累積することができる。低精度ナンバーシステムを利用するニューラルネットワークにおいて、個別グラジエント値 710 がウェイトをアップデートする過程で無視されないように、該ニューラルネットワーク装置は、個別グラジエント値 710 をウェイトに影響を与えることができる大きさまで累積することができる。

10

【0097】

該ニューラルネットワーク装置は、累積バッファの累積グラジエント値が臨界値以上である場合、該累積グラジエント値のうち、ウェイト 730 のビット桁に合算可能な部分を有効グラジエント値と決定することができる。該ニューラルネットワーク装置は、累積グラジエント値から有効グラジエント値を減算することにより、残余グラジエント値 720 を計算することができる。

【0098】

図 7 を参照すれば、ニューラルネットワーク装置は、個別グラジエント値 710 と残余グラジエント値 720 とを合算し、中間合算値 740 を計算することができる。図 7 においては、個別グラジエント値 710 及び残余グラジエント値 720 のいずれも最上位ビット (MSB) に符号ビットを含んでおり、ニューラルネットワーク装置は、2 つの値の符号ビットを考慮し、個別グラジエント値 710 と残余グラジエント値 720 との中間合算値 740 を計算することができる。

20

【0099】

そのとき、個別グラジエント値 710 と残余グラジエント値 720 とのビット数及びビット桁が対応しない場合、該ニューラルネットワーク装置は、残余グラジエント値 720 のビット数及びビット桁に対応するように、個別グラジエント値 710 を調整することができる。一実施形態において、該ニューラルネットワーク装置は、個別グラジエント値 710 を量子化してパディングすることにより、残余グラジエント値 720 のビット数及びビット桁に対応するように、個別グラジエント値 710 を調整することができる。

30

【0100】

ニューラルネットワーク装置は、中間合算値 740 とウェイト 730 とを合算し、最終合算値 750 を計算することができる。該ニューラルネットワーク装置は、最終合算値 750 から、アップデートされたウェイト 760 及びアップデートされた残余グラジエント値 770 を獲得することができる。一実施形態において、該ニューラルネットワーク装置は、ウェイト 730 のビット桁に基づいて、最終合算値 750 からアップデートされたウェイト 760、及びアップデートされた残余グラジエント値 770 を獲得することができる。例えば、該ニューラルネットワーク装置は、ウェイト 730 のビット桁に対応する最終合算値 750 のビット桁の値を、アップデートされたウェイト 760 として獲得することができる。また、ニューラルネットワーク装置は、ウェイト 730 のビット桁に対応しない最終合算値 750 のビット桁の値を、アップデートされた残余グラジエント値 770 として獲得することができる。

40

【0101】

そのとき、アップデートされたウェイト 760 の最上位ビットは、最終合算値 750 の最上位ビットである符号ビットでもある。また、該ニューラルネットワーク装置は、最終合算値 750 の符号ビットを獲得し、獲得された符号ビットがアップデートされた残余グラジエント値 770 の最上位ビットになるように、アップデートされた残余グラジエント値 770 に追加される。

50

【0102】

一方、アップデートされた残余グラジエント値770に新たな個別グラジエント値が合算されることにより、アップデートされた残余グラジエント値770と、新たな個別グラジエント値との合算結果である累積グラジエント値が臨界値以上になった場合、前述の過程が反復され、ウェート及び残余グラジエント値が再アップデートされる。

【0103】

図8は、一実施形態による、ウェート及び残余グラジエント値をアップデートする例示を示す図面である。

【0104】

図8を参照すれば、ニューラルネットワーク装置は、個別グラジエント値810と残余グラジエント値820とのビット数及びビット桁が対応しないので、該ニューラルネットワーク装置は、残余グラジエント値820のビット数及びビット桁に対応するように、個別グラジエント値810を調整することができる。一実施形態において、該ニューラルネットワーク装置は、個別グラジエント値810を量子化してパディングすることにより、残余グラジエント値820のビット数及びビット桁に対応するように、個別グラジエント値810を調整することができる。

【0105】

以下、便宜上、図7と重複する説明は、省略する。

【0106】

該ニューラルネットワーク装置は、符号ビットが除かれた残余グラジエント値820をウェート830に連結(concatenation)し、中間連結値840を計算することができる。そのとき、中間連結値840の最上位ビットは、ウェート830の符号ビットに該当する。

【0107】

該ニューラルネットワーク装置は、中間連結値840と個別グラジエント値810とを合算し、最終合算値850を計算することができる。該ニューラルネットワーク装置は、最終合算値850から、アップデートされたウェート860及びアップデートされた残余グラジエント値870を獲得することができる。一実施形態において、ニューラルネットワーク装置は、ウェート830のビット桁に基づいて、最終合算値850から、アップデートされたウェート860、及びアップデートされた残余グラジエント値870を獲得することができる。例えば、該ニューラルネットワーク装置は、最終合算値850において、ウェート830のビット桁に対応する最終合算値850のビット桁の値を、アップデートされたウェート860として獲得することができる。また、ニューラルネットワーク装置は、ウェート830のビット桁に対応しない最終合算値850のビット桁の値を、アップデートされた残余グラジエント値870として獲得することができる。

【0108】

そのとき、アップデートされたウェート860の最上位ビットは、最終合算値850の最上位ビットである符号ビットでもある。しかし、アップデート以前の残余グラジエント値820と同様に、アップデートされた残余グラジエント値870は、符号ビットを含まない。すなわち、図7とは異なり、該ニューラルネットワーク装置は、最終合算値850の符号ビットを、アップデートされた残余グラジエント値870に追加しない。

【0109】

一方、アップデートされた残余グラジエント値870に、新たな個別グラジエント値が合算されることにより、アップデートされた残余グラジエント値870と、新たな個別グラジエント値との合算結果である累積グラジエント値が、臨界値以上になった場合、前述の過程が反復され、ウェート及び残余グラジエント値が再アップデートされる。

【0110】

図9は、一実施形態による、加算器を利用し、ウェート及び残余グラジエント値をアップデートする例示を示す図面である。

【0111】

10

20

30

40

50

ニューラルネットワーク装置は、加算器 900 を利用し、個別グラジエント値、残余グラジエント値及びウェートを合算することができる。加算器 900 は、算術論理演算装置（ALU：arithmetic logic unit）の一種であり、所定ビット数を有することができる。加算器 900 のビット数により、処理能が決定される。すなわち、加算器 900 のビット数により、加算器 900 で合算することができるデータのビット数が決定される。

【0112】

該ニューラルネットワーク装置は、個別グラジエント値 910、残余グラジエント値 920 及びウェート 930 を利用し、ウェート 930 及び残余グラジエント値 920 をアップデートすることができる。個別グラジエント値 910、残余グラジエント値 920 及びウェート 930 は、低精度ナンバーシステムで利用される固定小数点または浮動小数点値でもある。

10

【0113】

図 9 においては、32 ビット加算器 900 を利用し、16 ビットの個別グラジエント値 910、残余グラジエント値 920 及びウェート 930 を合算する過程が図示される。

【0114】

図 9 を参照すれば、該ニューラルネットワーク装置は、個別グラジエント値 910 のビット数及びビット桁が、残余グラジエント値 920 のビット数及びビット桁に対応するように、個別グラジエント値 910 を調整することができる。一実施形態において、ニューラルネットワーク装置は、個別グラジエント値 910 を量子化してパディングすることにより、残余グラジエント値 920 のビット数及びビット桁に対応するように、個別グラジエント値 910 を調整することができる。

20

【0115】

具体的には、ニューラルネットワーク装置は、個別グラジエント値 910 において、残余グラジエント値 920 の最下位ビット桁未満の値が省略されるように、個別グラジエント値 910 を量子化させることができる。また、該ニューラルネットワーク装置は、残余グラジエント値 920 の最上位ビット桁に対応するビット桁まで値が存在するように、量子化された個別グラジエント値をパディングすることができる。すなわち、該ニューラルネットワーク装置は、個別グラジエント値 910 に、量子化及びパディングを適用することによって調整された個別グラジエント値 912 を計算することができる。調整された個別グラジエント値 912 のビット数及びビット桁は、残余グラジエント値 920 のビット数及びビット桁に対応する。

30

【0116】

一方、個別グラジエント値 910 のビット数及びビット桁のうち少なくともいずれかが残余グラジエント値 920 のビット数及びビット桁に対応する場合、個別グラジエント値 910 を量子化したりパディングしたりする過程は、省略される。

【0117】

該ニューラルネットワーク装置は、加算器 900 のビット数に基づいて調整された個別グラジエント値 912 及び残余グラジエント値 920 を、加算器 900 にマッピングすることができる。すなわち、32 ビット加算器 900 を利用して演算することができるように、ニューラルネットワーク装置は、16 ビットの調整された個別グラジエント値 912 及び残余グラジエント値 920 を、32 ビットにマッピングすることができる。一実施形態において、ニューラルネットワーク装置は、1 の補数（complement）方式または 2 の補数方式が適用された補数器を利用し、16 ビットの調整された個別グラジエント値 912 及び残余グラジエント値 920 を 32 ビットにマッピングすることができる。しかし、マッピング方式は、それに制限されるものではない。

40

【0118】

例えば、ニューラルネットワーク装置は、16 ビットの調整された個別グラジエント値 912 及び残余グラジエント値 920 それぞれの符号ビットを、加算器 900 の最上位ビット桁にマッピングし、残り 15 ビットを、加算器 900 の最下位ビットを除いた残り下位 15 ビット桁にマッピングすることができる。その場合、加算器 900 の最上位ビット

50

を除いた残り上位 15 ビット桁及び最下位ビット桁は、パディングされる。

【0119】

ニューラルネットワーク装置は、加算器 900 を利用し、加算器 900 にマッピングされた個別グラジエント値 914 及び残余グラジエント値 924 を合算することにより、中間合算値 940 を計算することができる。図 9 においては、個別グラジエント値 914 及び残余グラジエント値 924 のいずれも、符号ビットを最上位ビットとして含んでおり、ニューラルネットワーク装置は、2 つの値の符号ビットを考慮し、個別グラジエント値 914 と残余グラジエント値 924 とを合算することにより、中間合算値 940 を計算することができる。

【0120】

該ニューラルネットワーク装置は、加算器 900 のビット数に基づいて、ウェイト 930 を、加算器にマッピングすることができる。前述のように、32 ビット加算器 900 を利用して演算することができるように、ニューラルネットワーク装置は、16 ビットのウェイト 930 を 32 ビットにマッピングすることができる。該ニューラルネットワーク装置は、加算器 900 の最上位ビット桁方向から下位ビット桁方向に、ウェイト 935 をマッピングすることができる。すなわち、該ニューラルネットワーク装置は、加算器 900 の上位 16 ビット桁に、16 ビットのウェイト 935 をマッピングすることができる。また、対応するウェイト 930 が存在しない加算器 900 の下位 16 ビット桁は、パディングされる。

【0121】

ニューラルネットワーク装置は、加算器 900 にマッピングされたウェイト 935 及び中間合算値 940 を合算し、最終合算値 950 を計算することができる。該ニューラルネットワーク装置は、最終合算値 950 から、アップデートされたウェイト 960 及びアップデートされた残余グラジエント値 970 を獲得することができる。

【0122】

一実施形態において、ニューラルネットワーク装置は、ウェイト 930 のビット桁に基づいて、最終合算値 950 からアップデートされたウェイト 960 及びアップデートされた残余グラジエント値 970 を獲得することができる。該ニューラルネットワーク装置は、ウェイト 930 のビット桁に対応する最終合算値 950 のビット桁の値を、アップデートされたウェイト 960 として獲得することができる。また、該ニューラルネットワーク装置は、ウェイト 930 のビット桁に対応しない最終合算値 950 のビット桁の値を、アップデートされた残余グラジエント値 970 として獲得することができる。

【0123】

例えば、ニューラルネットワーク装置は、32 ビット加算器 900 の演算結果である 32 ビットの最終合算値 950 において、ウェイト 930 のビット桁に対応する最終合算値 950 の上位 16 ビット桁の値を、アップデートされたウェイト 960 として獲得し、下位 16 ビット桁の値は、アップデートされた残余グラジエント値 970 として獲得することができる。

【0124】

また、該ニューラルネットワーク装置は、最終合算値 950 の最上位ビットである符号ビットを獲得し、符号ビットがアップデートされた残余グラジエント値 970 の最上位ビットになるように追加することができる。

【0125】

図 10 は、一実施形態による、加算器を利用し、ウェイト及び残余グラジエント値をアップデートする例示を示す図面である。

【0126】

以下、図 9 と重複する内容は、便宜上、省略する。

【0127】

図 10 においては、32 ビット加算器 1000 を利用し、16 ビットの個別グラジエント値 1010、残余グラジエント値 1020 及び 8 ビットのウェイト 1030 を合算する

10

20

30

40

50

過程が図示される。

【0128】

該ニューラルネットワーク装置は、加算器1000のビット数に基づいて、ウェート1030を加算器にマッピングすることができる。すなわち、32ビット加算器1000を利用して演算することができるように、8ビットのウェート1030を32ビットにマッピングすることができる。一方、図9のウェート930は、16ビットであった一方、図10においては、ウェート1030が8ビットである。その場合、該ニューラルネットワーク装置は、加算器1000の最上位桁まで値が存在するように、ウェート1030をパディングすることができる。該ニューラルネットワーク装置は、加算器1000の最上位ビット桁方向から下位ビット桁方向にパディングされたウェートをマッピングすることができる。すなわち、該ニューラルネットワーク装置は、加算器1000の上位16ビット桁に、16ビットのパディングされたウェートをマッピングすることができる。また、対応するウェート1030が存在しない加算器1000の下位16ビット桁も、パディングされることができる。

10

【0129】

該ニューラルネットワーク装置は、加算器1000にマッピングされたウェート1035及び中間合算値1040を合算し、最終合算値1050を計算することができる。該ニューラルネットワーク装置は、最終合算値1050から、アップデートされたウェート1060、及びアップデートされた残余グラジエント値1070を獲得することができる。

20

【0130】

一実施形態において、ニューラルネットワーク装置は、ウェート1030のビット桁に基づいて、最終合算値1050からアップデートされたウェート1060、及びアップデートされた残余グラジエント値1070を獲得することができる。該ニューラルネットワーク装置は、ウェート1030のビット桁に対応する最終合算値1050のビット桁の値を、アップデートされたウェート1060として獲得することができる。また、該ニューラルネットワーク装置は、ウェート1030のビット桁に対応しない最終合算値1050のビット桁の値を、アップデートされた残余グラジエント値1070として獲得することができる。

【0131】

例えば、ニューラルネットワーク装置は、32ビット加算器1000の演算結果である32ビットの最終合算値1050において、ウェート1030のビット桁に対応しない下位16ビット桁の値を、アップデートされた残余グラジエント値1070として獲得することができる。一方、ニューラルネットワーク装置は、32ビットの最終合算値1050の上位16ビット桁の値のうち、パディング部（上位8ビット桁）を除いた残りビット桁の値を、アップデートされたウェート1060として獲得することができる。

30

【0132】

また、ニューラルネットワーク装置は、最終合算値1050の最上位ビットである符号ビットを獲得した後、アップデートされたウェート1060の最上位ビットを、最終合算値1050の符号ビットに置き換え、最終合算値1050の符号ビットを、アップデートされた残余グラジエント値1070の最上位ビットになるように追加することができる。

40

【0133】

図11は、一実施形態による、加算器を利用し、ウェート及び残余グラジエント値をアップデートする例示を示す図面である。

【0134】

以下、図9と重複する内容は、便宜上、省略する。

【0135】

図11においては、32ビット加算器1100を利用し、16ビットの個別グラジエント値1110、残余グラジエント値1120及びウェート1130を合算する過程が図示される。

【0136】

50

図 1 1 を参照すれば、該ニューラルネットワーク装置は、個別グラジエント値 1 1 1 0 のビット数及びビット桁が、残余グラジエント値 1 1 2 0 のビット数及びビット桁に対応するように、個別グラジエント値 1 1 1 0 を調整することができる。一実施形態において、該ニューラルネットワーク装置は、個別グラジエント値 1 1 1 0 に、量子化及びパディングを適用することによって調整された個別グラジエント値 1 1 1 2 を計算することができる。

【0 1 3 7】

該ニューラルネットワーク装置は、符号ビットが除かれた残余グラジエント値 1 1 2 0 をウェート 1 1 3 0 に連結し、中間連結値 1 1 4 0 を計算することができる。中間連結値 1 1 4 0 の最上位ビットは、ウェート 1 0 3 0 の符号ビットに該当し、1 6 ビットのウェ

10

ート 1 1 3 0 及び符号ビットが除かれた 1 5 ビットの残余グラジエント値 1 1 2 0 が連結になって 3 1 ビットの中間連結値 1 1 4 0 が計算されることができる。ニューラルネットワーク装置は、中間連結値 1 1 4 0 及び調整された個別グラジエント値 1 1 1 2 それぞれを加算器 1 1 0 0 にマッピングすることができる。例えば、ニューラルネットワーク装置は、加算器 1 1 0 0 の上位 3 1 ビット桁に、中間連結値 1 1 4 0 をマッピングすることができる。その場合、加算器 1 1 0 0 の最下位ビット桁も、パディングされる。

【0 1 3 8】

また、ニューラルネットワーク装置は、1 6 ビットの調整された個別グラジエント値 1 1 1 2 の符号ビットを、加算器 1 1 0 0 の最上位ビット桁にマッピングし、残り 1 5 ビットを、加算器 1 1 0 0 の最下位ビットを除いた残り下位 1 5 ビット桁にマッピングすることができる。その場合、加算器 1 1 0 0 の最上位ビットを除いた残り上位 1 5 ビット桁及び最下位ビット桁は、パディングされる。

20

【0 1 3 9】

該ニューラルネットワーク装置は、中間連結値 1 1 4 0 と、マッピングされた個別グラジエント値 1 1 1 4 とを合算し、最終合算値 1 1 5 0 を計算することができる。該ニューラルネットワーク装置は、最終合算値 1 1 5 0 から、アップデートされたウェート 1 1 6 0、及びアップデートされた残余グラジエント値 1 1 7 0 を獲得することができる。

【0 1 4 0】

一実施形態において、ニューラルネットワーク装置は、ウェート 1 1 3 0 のビット桁に基づいて、最終合算値 1 1 5 0 から、アップデートされたウェート 1 1 6 0 及びアップデートされた残余グラジエント値 1 1 7 0 を獲得することができる。例えば、該ニューラルネットワーク装置は、最終合算値 1 1 5 0 において、ウェート 1 1 3 0 のビット桁に対応する最終合算値 1 1 5 0 のビット桁の値を、アップデートされたウェート 1 1 6 0 として獲得することができる。また、ニューラルネットワーク装置は、ウェート 1 1 3 0 のビット桁に対応しない最終合算値 1 1 5 0 のビット桁の値を、アップデートされた残余グラジエント値 1 1 7 0 として獲得することができる。

30

【0 1 4 1】

そのとき、アップデートされたウェート 1 1 6 0 の最上位ビットは、最終合算値 1 1 5 0 の最上位ビットである符号ビットでもある。しかし、アップデート以前の残余グラジエント値 1 0 2 0 と同様に、アップデートされた残余グラジエント値 1 0 7 0 は、符号ビットを含まない。

40

【0 1 4 2】

図 1 2 は、一実施形態による、加算器を利用し、ウェート及び残余グラジエント値をアップデートする例示を示す図面である。

【0 1 4 3】

以下、図 1 0 と重複する内容は、便宜上、省略する。

【0 1 4 4】

図 1 2 においては、3 2 ビット加算器 1 2 0 0 を利用し、1 6 ビットの個別グラジエント値 1 2 1 0、残余グラジエント値 1 2 2 0 及び 8 ビットのウェート 1 2 3 0 を合算する

50

過程が図示される。

【0145】

図12を参照すれば、該ニューラルネットワーク装置は、個別グラジエント値1210のビット数及びビット桁が、残余グラジエント値1220のビット数及びビット桁に対応するように、個別グラジエント値1210を調整することができる。一実施形態において、該ニューラルネットワーク装置は、個別グラジエント値1210に、量子化及びパディングを適用することによって調整された個別グラジエント値1212を計算することができる。

【0146】

該ニューラルネットワーク装置は、符号ビットが除かれた残余グラジエント値1220をウェイト1230に連結し、中間連結値1240を計算することができる。また、該ニューラルネットワーク装置は、中間連結値1240を加算器1200にマッピングすることができる。

【0147】

一実施形態において、ウェイト1230の最上位ビットである符号ビットは、加算器1200の最上位ビット桁にもマッピングされる。一方、図11のウェイト1130は、16ビットであった一方、図12では、ウェイト1230が8ビットである。その場合、ニューラルネットワーク装置は、加算器1200の最上位桁まで値が存在するように、中間連結値1240をパディングすることができる。また、図11と同様に、加算器1200の最下位ビット桁も、パディングされる。

【0148】

また、該ニューラルネットワーク装置は、16ビットの調整された個別グラジエント値1212の符号ビットを、加算器1200の最上位ビット桁にマッピングし、残り15ビットを、加算器1200の最下位ビットを除いた残り下位15ビット桁にマッピングすることができる。その場合、加算器1200の最上位ビットを除いた残り上位15ビット桁及び最下位ビット桁は、パディングされる。

【0149】

該ニューラルネットワーク装置は、中間連結値1240と、マッピングされた個別グラジエント値1214とを合算し、最終合算値1250を計算することができる。該ニューラルネットワーク装置は、最終合算値1250から、アップデートされたウェイト1260、及びアップデートされた残余グラジエント値1270を獲得することができる。

【0150】

一実施形態において、該ニューラルネットワーク装置は、ウェイト1230のビット桁に基づいて、最終合算値1250からアップデートされたウェイト1260、及びアップデートされた残余グラジエント値1270を獲得することができる。例えば、該ニューラルネットワーク装置は、最終合算値1250において、ウェイト1230のビット桁に対応する最終合算値1250のビット桁の値を、アップデートされたウェイト1260として獲得することができる。また、ニューラルネットワーク装置は、ウェイト1230のビット桁に対応しない最終合算値1250のビット桁の値を、アップデートされた残余グラジエント値1270として獲得することができる。

【0151】

例えば、ニューラルネットワーク装置は、32ビット加算器1200の演算結果である32ビットの最終合算値1250において、ウェイト1230のビット桁に対応しない下位16ビット桁の値を、アップデートされた残余グラジエント値1270として獲得することができる。一方、該ニューラルネットワーク装置は、32ビットの最終合算値1250の上位16ビット桁の値のうち、パディング部（上位8ビット桁）を除いた残りビット桁の値を、アップデートされたウェイト1260として獲得することができる。

【0152】

また、ニューラルネットワーク装置は、最終合算値1250の最上位ビットである符号ビットを獲得した後、アップデートされたウェイト1260の最上位ビットを、最終合算

10

20

30

40

50

値 1 2 5 0 の符号ビットに置き換えることができる。しかし、アップデート以前の残余グラジエント値 1 2 2 0 と同様に、アップデートされた残余グラジエント値 1 2 7 0 は、符号ビットを含まない。

【0153】

図 1 3 は、一実施形態による低精度ナンバーシステムを利用するニューラルネットワークにおいて、パラメータを処理する方法を図示したフローチャートである。

【0154】

図 1 3 を参照すれば、段階 1 3 1 0 において、該ニューラルネットワーク装置は、ニューラルネットワークのウェートをアップデートするための個別グラジエント値を計算することができる。

10

【0155】

該ニューラルネットワーク装置は、ニューラルネットワークに含まれるノード間の連結関係に割り当てられるウェートのための個別グラジエント値を計算することができる。該個別グラジエント値は、ウェートをアップデートするために計算される値でもある。図 4 で説明したように、ニューラルネットワーク装置の optimizer は、グラジエントディセンスト方式、モーメンタム方式などを利用し、個別グラジエント値を計算することができる。

【0156】

段階 1 3 2 0 において、該ニューラルネットワーク装置は、個別グラジエント値を累積した累積グラジエント値及びウェートのビット桁に基づいて、残余グラジエント値を計算することができる。

20

【0157】

該ニューラルネットワーク装置は、個別グラジエント値を累積バッファに累積することができる。低精度ナンバーシステムを利用するニューラルネットワークにおいて、個別グラジエント値がウェートをアップデートする過程で無視されないように、該ニューラルネットワーク装置は、個別グラジエント値をウェートに影響を与えることができる大きさまで累積することができる。

【0158】

該ニューラルネットワーク装置は、累積バッファの累積グラジエント値が臨界値以上である場合、累積グラジエント値のうちウェートのビット桁に合算可能な部分を、有効グラジエント値と決定することができる。該ニューラルネットワーク装置は、累積グラジエント値から有効グラジエント値を減算することにより、残余グラジエント値を計算することができる。

30

【0159】

段階 1 3 3 0 において、該ニューラルネットワーク装置は、残余グラジエント値のビット桁に対応するように、個別グラジエント値を調整することができる。

【0160】

一実施形態において、該ニューラルネットワーク装置は、個別グラジエント値を量子化してパディングすることにより、残余グラジエント値のビット数及びビット桁に対応するように、個別グラジエント値を調整することができる。

【0161】

具体的には、該ニューラルネットワーク装置は、個別グラジエント値のうち、残余グラジエント値の最下位ビット桁未満の値が省略されるように、個別グラジエント値を量子化することができる。また、該ニューラルネットワーク装置は、残余グラジエント値の最上位ビット桁に対応するビット桁まで値が存在するように、量子化された個別グラジエント値をパディングすることができる。

40

【0162】

一実施形態において、段階 1 3 3 0 後、段階 1 3 4 0 及び段階 1 3 5 0 に進められる。

【0163】

段階 1 3 4 0 において、該ニューラルネットワーク装置は、加算器を利用して調整された個別グラジエント値、残余グラジエント値及びウェートを合算することができる。

50

【0164】

該ニューラルネットワーク装置は、調整された個別グラジエント値と残余グラジエント値とを合算し、中間合算値を計算することができる。調整された個別グラジエント値及び残余グラジエント値のいずれも、最上位ビット（MSB）として符号ビットを含むので、該ニューラルネットワーク装置は、2つの値の符号ビットを考慮し、個別グラジエント値と残余グラジエント値との中間合算値を計算することができる。

【0165】

また、該ニューラルネットワーク装置は、中間合算値とウェートとを合算し、最終合算値を計算することができる。

【0166】

段階1350において、該ニューラルネットワーク装置は、合算結果に基づいて、ウェート及び残余グラジエント値をアップデートすることができる。

【0167】

一実施形態において、該ニューラルネットワーク装置は、ウェートのビット桁に基づいて、最終合算値からアップデートされたウェート、及びアップデートされた残余グラジエント値を獲得することができる。例えば、該ニューラルネットワーク装置は、ウェートのビット桁に対応する最終合算値のビット桁の値を、アップデートされたウェートとして獲得することができる。また、該ニューラルネットワーク装置は、ウェートのビット桁に対応しない最終合算値のビット桁の値を、アップデートされた残余グラジエント値として獲得することができる。

【0168】

そのとき、アップデートされたウェートの最上位ビットは、最終合算値の最上位ビットである符号ビットでもある。また、該ニューラルネットワーク装置は、最終合算値の符号ビットを獲得し、獲得された符号ビットが、アップデートされた残余グラジエント値の最上位ビットになるように、アップデートされた残余グラジエント値に追加される。

【0169】

他の実施形態において、段階1330後、段階1345、段階1355及び段階1365に進められる。

【0170】

段階1345において、該ニューラルネットワーク装置は、残余グラジエント値から符号ビットを除外した残りの値をウェートに連結し、中間連結値を計算することができる。中間連結値840の最上位ビットは、ウェートの符号ビットに該当する。

【0171】

段階1355において、該ニューラルネットワーク装置は、加算器を利用して調整された個別グラジエント値及び中間連結値を合算することができる。

【0172】

調整された個別グラジエント値及び中間連結値のいずれも、最上位ビット（MSB）として符号ビットを含むので、該ニューラルネットワーク装置は、2値の符号ビットを考慮し、個別グラジエント値と中間連結値との中間合算値を合算することができる。

【0173】

段階1365において、該ニューラルネットワーク装置は、合算結果に基づいて、ウェート及び残余グラジエント値を、アップデート合算結果に基づいて、ウェート及び残余グラジエント値をアップデートすることができる。

【0174】

一実施形態において、該ニューラルネットワーク装置は、ウェートのビット桁に基づいて、最終合算値からアップデートされたウェート、及びアップデートされた残余グラジエント値を獲得することができる。例えば、該ニューラルネットワーク装置は、ウェートのビット桁に対応する最終合算値のビット桁の値を、アップデートされたウェートとして獲得することができる。また、該ニューラルネットワーク装置は、ウェートのビット桁に対応しない最終合算値のビット桁の値を、アップデートされた残余グラジエント値として獲

10

20

30

40

50

得することができる。

【0175】

そのとき、アップデートされたウェートの最上位ビットは、最終合算値の最上位ビットである符号ビットでもある。しかし、アップデート以前の残余グラジエント値と同様に、アップデートされた残余グラジエント値は、符号ビットを含まない。すなわち、該ニューラルネットワーク装置は、最終合算値の符号ビットを、アップデートされた残余グラジエント値に追加しない。

【0176】

図14は、一実施形態によるニューラルネットワーク装置のハードウェア構成を図示したブロック図である。

10

【0177】

図14を参照すれば、ニューラルネットワーク装置1400は、少なくとも1つのプロセッサ1410及びメモリ1420を含んでもよい。図14に図示されたニューラルネットワーク装置1400には、本実施形態と係わる構成要素だけが図示されている。従って、ニューラルネットワーク装置1400に、図14に図示された構成要素以外に、他の汎用的な構成要素がさらに含まれてもよいということは、当業者に自明である。例えば、ニューラルネットワーク装置1400は、センサモジュール（小数点）、通信モジュール（小数点）及びインターフェース（小数点）などをさらに含んでもよい。

【0178】

ニューラルネットワーク装置1400は、家電製品、モバイルコンピュータデバイス及びサーバのうち少なくとも一つに含まれるか、あるいは家電製品、モバイルコンピュータデバイス及びサーバのうち少なくとも一つに、有無線で連結されるようにも具現される。また、ニューラルネットワーク装置1400は、デバイスとサーバとの結合によっても具現される。ユーザの入力を受信するデバイスと、ニューラルネットワークを学習させるサーバとが別途に存在し、デバイスとサーバとの通信を介して、推論が行われる。ここで、単一のデバイスで推論が行われもするが、前記例示に制限されるものではないということはあるまでもない。

20

【0179】

プロセッサ1410は、1個または複数個のプロセッサによっても具現される。例えば、プロセッサ1410は、多数の論理ゲートのアレイによっても具現され、汎用的なマイクロプロセッサと、マイクロプロセッサで実行されるプログラムが保存されたメモリとの組み合わせによっても具現される。例えば、プロセッサ1410は、CPU（central processing unit）、GPU（graphics processing unit）、ハードウェアアクセラレータなどでもある。

30

【0180】

プロセッサ1410は、ニューラルネットワーク装置1400を制御するための全般的な役割を行うことができる。例えば、プロセッサ1410は、ニューラルネットワーク装置1400内のメモリ1420に保存されたプログラムを実行することにより、ニューラルネットワーク装置1400を全般的に制御することができる。また、少なくとも1つのプロセッサ1410は、メモリ1420に保存されたプログラムを実行することにより、ニューラルネットワーク装置1400の機能を並列的に遂行することができる。

40

【0181】

少なくとも1つのプロセッサ1410は、該ニューラルネットワークのレイヤに含まれたノードそれぞれの機能を遂行することができる。例えば、少なくとも1つのプロセッサ1410は、アクティベーション値及びウェートを利用し、ニューラルネットワーク演算またはマトリックス演算を行うことができる。

【0182】

プロセッサ1410は、図1ないし図13で説明した低精度ナンバーシステムを利用するニューラルネットワークにおいて、パラメータを処理するための一連のプロセスを制御することができる。

50

【0183】

メモリ1420は、ニューラルネットワーク装置1400内で処理される各種データを保存するハードウェアであり、例えば、メモリ1420は、ニューラルネットワーク装置1400で処理されたデータ及び処理されるデータを保存することができる。また、メモリ1420は、ニューラルネットワーク装置1400によって駆動されるアプリケーション、ドライバなどを保存することができる。メモリ1420は、DRAM (dynamic random access memory)・SRAM (static random access memory) のようなRAM (random access memory)、ROM (read-only memory)、EEPROM (electrically erasable programmable read-only memory)、CD-ROM (compact disc read only memory)、ブルーレイ (登録商標 (Blu-ray))、または他の光学ディスクストレージ、HDD (hard disk drive)、SSD (solid state drive)、またはフラッシュメモリを含んでもよい。

10

【0184】

また、メモリ1420は、ニューラルネットワーク装置1400に含まれた少なくとも1つのプロセッサ1410が共有する内蔵メモリ、または少なくとも1つのプロセッサ1410を支援する複数の内蔵メモリをさらに含んでもよい。

【0185】

本実施形態は、コンピュータによって実行されるプログラムモジュールのようなコンピュータによって実行可能な命令語を含む記録媒体の形態にも具現される。コンピュータ可読媒体は、コンピュータによってアクセスされる任意の可用媒体でもあり、揮発性及び不揮発性の媒体、分離型及び非分離型の媒体をいずれも含む。また、該コンピュータ可読媒体は、コンピュータ記録媒体及び通信媒体をいずれも含んでもよい。該コンピュータ記録媒体は、コンピュータ可読命令語、データ構造、プログラムモジュールまたはその他データのような情報保存のための任意の方法または技術によって具現された揮発性及び不揮発性、分離型及び非分離型の媒体をいずれも含む。該通信媒体は、典型的に、コンピュータ可読命令語、データ構造、プログラムモジュールのような変調されたデータ信号のその他データ、またはその他伝送メカニズムを含み、任意の情報伝達媒体を含む。

20

【0186】

また、本明細書において、「部」は、プロセッサまたは回路のようなハードウェア構成 (hardware component)、及び／またはプロセッサのようなハードウェア構成によって実行されるソフトウェア構成 (software component) でもある。

30

【0187】

前述の本明細書の説明は、例示のためのものであり、本明細書の内容が属する技術分野の当業者であるならば、本発明の技術的思想や、必須な特徴を変更せずにも、他の具体的な形態に容易に変形が可能であるということを理解することができるであろう。従って、以上で記述した実施形態は、全ての面において、例示的なものであり、限定的なものではないと理解しなければならない。例えば、単一型と説明されている各構成要素は、分散されても実施され、同様に、分散されていると説明されている構成要素も、結合された形態でも実施される。

【0188】

本実施形態の範囲は、前述の詳細な説明よりは、特許請求の範囲によって示され特許請求の範囲の意味、範囲そしてその均等概念から導き出される全ての変更、または変形された形態が含まれるものであると解釈されなければならない。

40

【産業上の利用可能性】

【0189】

本発明の、ニューラルネットワークにおいて、パラメータを処理する方法及びその装置は、例えば、情報処理関連の技術分野に効果的に適用可能である。

【符号の説明】

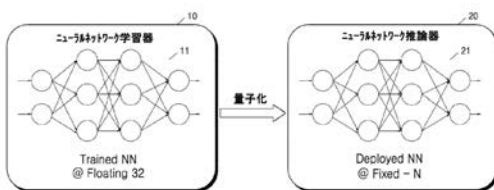
【0190】

10 ニューラルネットワーク学習器

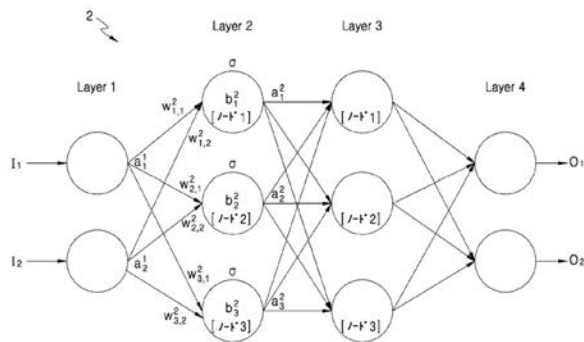
50

1 1, 2 1 ニューラルネットワーク
 2 0 ニューラルネットワーク推論器
 7 1 0 個別グラジエント値
 7 2 0 残余グラジエント値
 7 6 0 アップデートされたウェイト
 7 7 0 アップデートされたグラジエント値
 1 4 0 0 ニューラルネットワーク装置
 1 4 1 0 プロセッサ
 1 4 2 0 メモリ

【図 1】



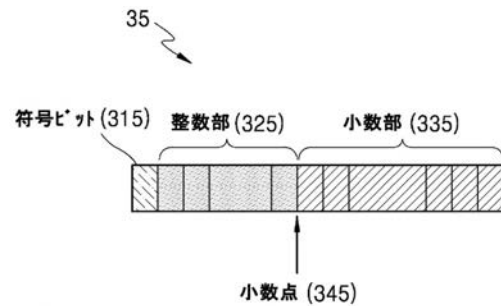
【図 2】



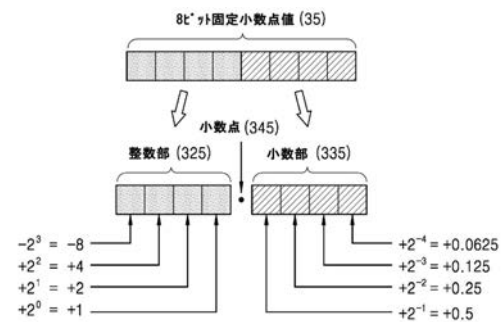
【図 3 A】



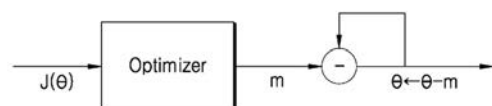
【図 3 B】



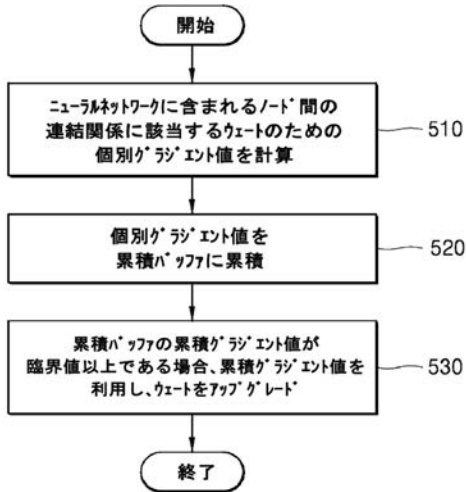
【図 3 C】



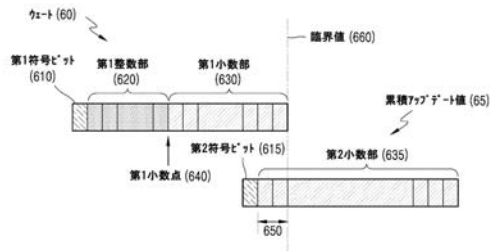
【図 4】



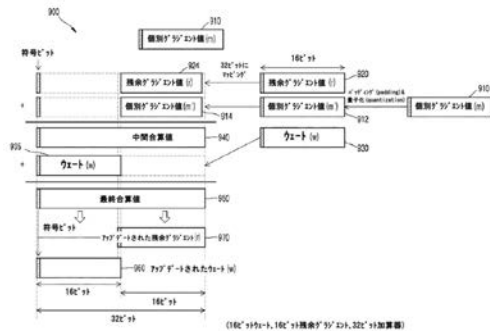
【図 5】



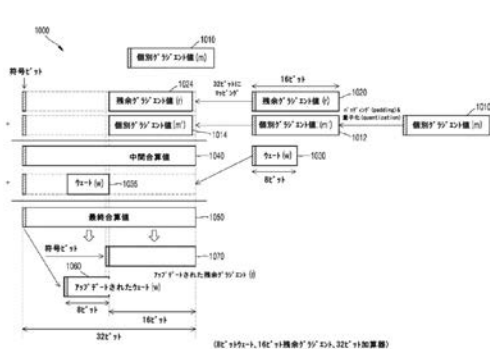
【図 6】



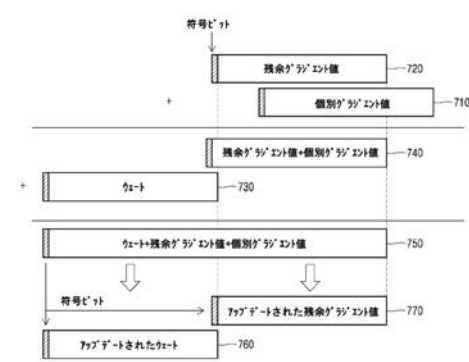
【図 9】



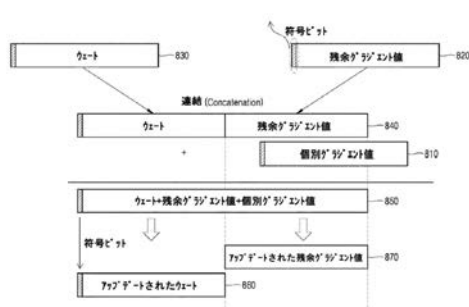
【図 10】



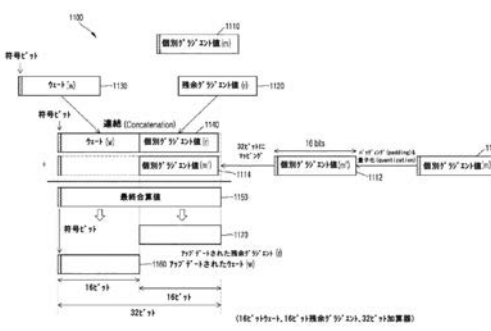
【図 7】



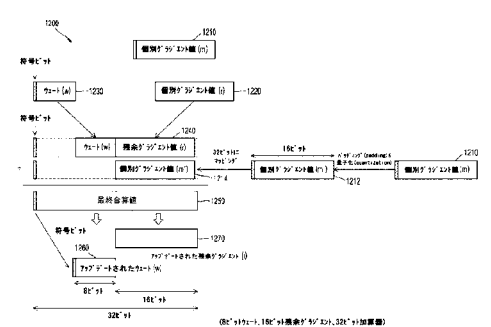
【図 8】



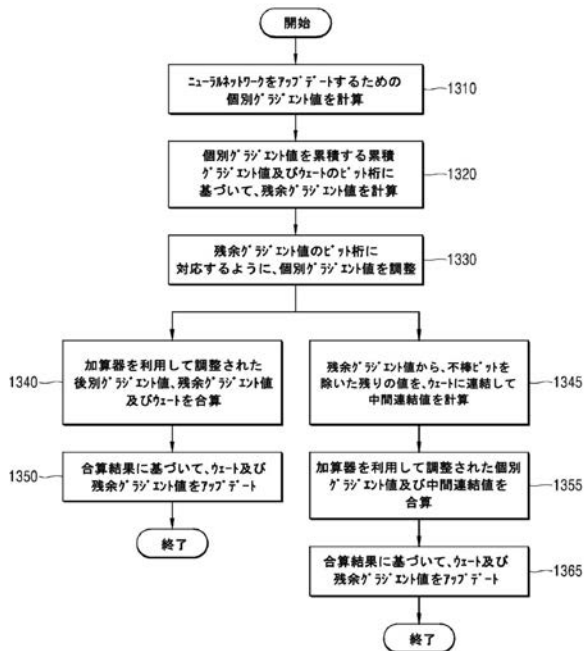
【図 11】



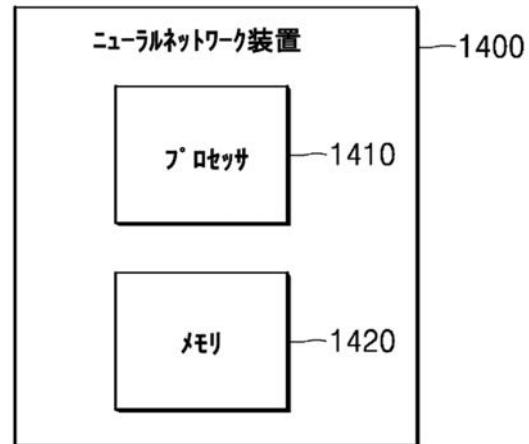
【図 12】



【図 1 3】



【図 1 4】



【手続補正書】

【提出日】令和1年7月3日(2019.7.3)

【手続補正1】

【補正対象書類名】明細書

【補正対象項目名】0004

【補正方法】変更

【補正の内容】

【0004】

【特許文献1】特許第3345478号公報

【特許文献2】中国特許出願公開第107480770号明細書

フロントページの続き

- (72)発明者 李 俊行
大韓民国京畿道水原市靈通区三星路1 3 0 三星綜合技術院内
- (72)発明者 朴 ▲ひょん▼宣
大韓民国京畿道水原市靈通区三星路1 3 0 三星綜合技術院内
- (72)発明者 宋 準鎬
大韓民国京畿道水原市靈通区三祐路1 3 0 三星綜合技術院内