

【公報種別】特許法第 17 条の 2 の規定による補正の掲載
【部門区分】第 6 部門第 3 区分
【発行日】令和 5 年 10 月 11 日(2023.10.11)

【公開番号】特開 2022-188301(P2022-188301A)
【公開日】令和 4 年 12 月 20 日(2022.12.20)
【年通号数】公開公報(特許)2022-234
【出願番号】特願 2022-169395(P2022-169395)
【国際特許分類】

G 0 6 F 17/16(2006.01)

10

G 0 6 N 3/063(2023.01)

【F I】

G 0 6 F 17/16 M

G 0 6 N 3/063

【手続補正書】

【提出日】令和 5 年 9 月 29 日(2023.9.29)

【手続補正 1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

20

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項 1】

複数の層を有するネットワークにおける層の複数のフィルタ係数と複数のデータとの乗算演算を実行する複数の乗算手段と、

前記複数の乗算手段による複数の乗算演算結果を、それぞれの乗算演算結果をシフト量に基づいて 2 のべき乗で乗算することによりシフトする複数のシフト演算手段と、
入力データの桁数および該入力データのビット幅に基づいて、前記複数のシフト演算手段に対応する前記シフト量を、前記ネットワークのそれぞれの層について設定する制御手段と、

30

前記複数のシフト演算手段によりシフトされた複数の乗算演算結果の総和を計算する加算手段と、

を備え、

前記複数の乗算手段の数は、前記複数のシフト演算手段の数に対応することを特徴とする情報処理装置。

【請求項 2】

前記複数の乗算手段は、前記複数の層のそれぞれについての乗算演算を同一の回路で実行することを特徴とする請求項 1 に記載の情報処理装置。

【請求項 3】

40

前記複数の乗算手段は、前記複数の層のそれぞれについての乗算演算を同一の回路で順次実行することを特徴とする請求項 2 に記載の情報処理装置。

【請求項 4】

前記制御手段は、前記層への入力データのビット幅に応じてフィルタ係数の転送回数を切り替えることを特徴とする請求項 1 に記載の情報処理装置。

【請求項 5】

前記乗算手段は、前記乗算演算とともに加算演算を実行することを特徴とする請求項 1 に記載の情報処理装置。

【請求項 6】

複数の層を有するネットワークにおける層の複数のデータを、該複数のデータのそれぞ

50

れをシフト量に基づいて 2 のべき乗で乗算することによりシフトする複数のシフト演算手段と、

入力データの桁数および該入力データのビット幅に基づいて、前記複数のシフト演算手段に対応するシフト量を、前記ネットワークのそれぞれの層について設定する制御手段と、

前記複数のシフト演算手段によりシフトされた複数のデータのデータと複数のフィルタ係数との乗算演算を実行する複数の乗算手段と、

前記複数の乗算手段により計算された複数の乗算演算結果の総和を計算する加算手段と

を備え、

前記複数の乗算手段の数は、前記複数のシフト演算手段の数に対応することを特徴とする情報処理装置。

10

【請求項 7】

前記複数の乗算手段と前記複数のシフト演算手段とは並列に動作することを特徴とする請求項 1 または 6 に記載の情報処理装置。

【請求項 8】

前記加算手段は、ある層に対する前記総和を該層の次の層のデータとしてメモリに格納することを特徴とする請求項 1 または 6 に記載の情報処理装置。

【請求項 9】

更に、

前記総和に対して活性化処理を行う活性化手段を備えることを特徴とする請求項 1 または 6 に記載の情報処理装置。

20

【請求項 10】

更に、

前記活性化処理の結果に対してプーリング処理を行うプーリング手段を備えることを特徴とする請求項 9 に記載の情報処理装置。

【請求項 11】

前記プーリング手段は、ある層に対する前記プーリング処理を行った結果を、該層の次の層のデータとしてメモリに格納することを特徴とする請求項 10 に記載の情報処理装置。

【請求項 12】

30

更に、

前記メモリに格納されたデータに基づいて、動画像における各フレームに対する画像処理および / または画像認識を行う手段を備えることを特徴とする請求項 8 に記載の情報処理装置。

【請求項 13】

前記入力データのビット幅は、2 ビット、4 ビット、または 8 ビットであることを特徴とする請求項 1 または 6 に記載の情報処理装置。

【請求項 14】

前記複数の乗算手段で演算されるデータのビット幅は、前記入力データのビット幅より短いことを特徴とする請求項 1 または 6 に記載の情報処理装置。

40

【請求項 15】

前記乗算手段では、1 ビットの演算を行うことを特徴とする請求項 6 に記載の情報処理装置。

【請求項 16】

情報処理装置であって、

複数の演算手段であって、複数の層を含むネットワークの層に入力されるデータのビット幅が前記演算手段のビット幅と同じである場合には、該ネットワークの層の複数のフィルタ係数と複数のデータとの乗算を、該複数の演算手段の演算結果の和が計算される前に該複数のデータをビットシフトすることなく行い、前記層に入力されるデータのビット幅が前記演算手段のビット幅の倍数である場合には、前記複数のフィルタ係数と前記複数のデ

50

ータとの乗算を、前記複数の演算手段の演算結果の和が計算される前にシフト量に基づいて該複数のデータをビットシフトすることで行い、ここで、前記層に入力されるデータは、前記複数のデータに分割される、該複数の演算手段と、
前記ネットワークのそれぞれの層について、該層に入力されるデータの桁に基づいて、前記シフト量を設定する制御手段と
を備えることを特徴とする情報処理装置。

【請求項 17】

情報処理装置が行う情報処理方法であって、
前記情報処理装置の複数の乗算手段が、複数の層を有するネットワークにおける層の複数のフィルタ係数と複数のデータとの乗算演算を実行する工程と、
前記情報処理装置の複数のシフト演算手段が、前記複数の乗算手段による複数の乗算演算結果を、それぞれの乗算演算結果をシフト量に基づいて2のべき乗で乗算することによりシフトする工程と、
前記情報処理装置の制御手段が、入力データの桁数および該入力データのビット幅に基づいて、前記複数のシフト演算手段に対応する前記シフト量を、前記ネットワークのそれぞれの層について設定する工程と、
前記情報処理装置の加算手段が、前記複数のシフト演算手段によりシフトされた複数の乗算演算結果の総和を計算する工程と、
を備え、

10

前記複数の乗算手段の数は、前記複数のシフト演算手段の数に対応することを特徴とする情報処理方法。

20

【請求項 18】

情報処理装置が行う情報処理方法であって、
前記情報処理装置の複数のシフト演算手段が、複数の層を有するネットワークにおける層の複数のデータを、該複数のデータのそれぞれをシフト量に基づいて2のべき乗で乗算することによりシフトする工程と、
前記情報処理装置の制御手段が、入力データの桁数および該入力データのビット幅に基づいて、前記複数のシフト演算手段に対応するシフト量を、前記ネットワークのそれぞれの層について設定する工程と、
前記情報処理装置の複数の乗算手段が、前記複数のシフト演算手段によりシフトされた複数のデータと複数のフィルタ係数との乗算演算を実行する工程と、
前記情報処理装置の加算手段が、前記複数の乗算手段により計算された複数の乗算演算結果の総和を計算する工程と、
を備え、

30

前記複数の乗算手段の数は、前記複数のシフト演算手段の数に対応することを特徴とする情報処理方法。

【請求項 19】

情報処理装置が行う情報処理方法であって、
前記情報処理装置の複数の演算手段が、複数の層を含むネットワークの層に入力されるデータのビット幅が前記演算手段のビット幅と同じである場合には、該ネットワークの層の複数のフィルタ係数と複数のデータとの乗算を、該複数の演算手段の演算結果の和が計算される前に該複数のデータをビットシフトすることなく行い、前記層に入力されるデータのビット幅が前記演算手段のビット幅の倍数である場合には、前記複数のフィルタ係数と前記複数のデータとの乗算を、前記複数の演算手段の演算結果の和が計算される前にシフト量に基づいて該複数のデータをビットシフトすることで行い、ここで、前記層に入力されるデータは、前記複数のデータに分割される、工程と、
前記情報処理装置の制御手段が、前記ネットワークのそれぞれの層について、該層に入力されるデータの桁に基づいて、前記シフト量を設定する工程と
を備えることを特徴とする情報処理方法。

40

【請求項 20】

50

コンピュータを、請求項 1 乃至 1 6 の何れか 1 項に記載の情報処理装置の各手段として機能させるためのコンピュータプログラム。

【手続補正 2】

【補正対象書類名】明細書

【補正対象項目名】0 0 1 1

【補正方法】変更

【補正の内容】

【0 0 1 1】

本発明の一様態は、複数の層を有するネットワークにおける層の複数のフィルタ係数と複数のデータとの乗算演算を実行する複数の乗算手段と、

10

前記複数の乗算手段による複数の乗算演算結果を、それぞれの乗算演算結果をシフト量に基づいて 2 のべき乗で乗算することによりシフトする複数のシフト演算手段と、入力データの桁数および該入力データのビット幅に基づいて、前記複数のシフト演算手段に対応する前記シフト量を、前記ネットワークのそれぞれの層について設定する制御手段と、

前記複数のシフト演算手段によりシフトされた複数の乗算演算結果の総和を計算する加算手段と、

を備え、

前記複数の乗算手段の数は、前記複数のシフト演算手段の数に対応することを特徴とする。

20

30

40

50