

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6778010号
(P6778010)

(45) 発行日 令和2年10月28日(2020.10.28)

(24) 登録日 令和2年10月13日(2020.10.13)

(51) Int.Cl.

G06T 7/00 (2017.01)

F 1

G06T 7/00 350C

請求項の数 10 (全 15 頁)

(21) 出願番号 特願2016-80476 (P2016-80476)
 (22) 出願日 平成28年4月13日 (2016.4.13)
 (65) 公開番号 特開2017-191458 (P2017-191458A)
 (43) 公開日 平成29年10月19日 (2017.10.19)
 審査請求日 平成31年4月11日 (2019.4.11)

(73) 特許権者 000001007
 キヤノン株式会社
 東京都大田区下丸子3丁目30番2号
 (74) 代理人 100076428
 弁理士 大塚 康徳
 (74) 代理人 100115071
 弁理士 大塚 康弘
 (74) 代理人 100112508
 弁理士 高柳 司郎
 (74) 代理人 100116894
 弁理士 木村 秀二
 (74) 代理人 100130409
 弁理士 下山 治
 (74) 代理人 100134175
 弁理士 永川 行光

最終頁に続く

(54) 【発明の名称】 画像識別システム、画像識別方法

(57) 【特許請求の範囲】

【請求項 1】

複数の階層を有するニューラルネットワークを用いて、入力画像から識別対象を識別するための識別処理を行う画像識別システムであって、

撮像した画像を前記入力画像として入力し、該入力画像から特徴を抽出するための、第1パラメータを用いる畳み込みフィルタ演算を行う撮像装置である第1の演算装置と、前記第1の演算装置と異なる第2の演算装置であって、

前記第1の演算装置から前記畳み込みフィルタ演算の結果である第1演算結果を取得する取得手段と、

前記第1演算結果に基づく前記特徴と、重み係数パラメータである第2パラメータから構成される行列と、の行列積を求める行列積演算であって、該第2パラメータを記憶するためのメモリ容量が前記第1パラメータを記憶するためのメモリ容量と比べて多い前記行列積演算を行う演算手段と

を有する前記第2の演算装置と、

前記行列積演算の結果である第2演算結果を用いて、前記識別処理において前記識別対象を識別する識別手段と

を備えることを特徴とする画像識別システム。

【請求項 2】

前記第1の演算装置は、前記入力画像のそれぞれの部分画像に対して同一の値である前記第1パラメータを適用する前記畳み込みフィルタ演算を行い、前記第2の演算装置は、

同一のデータに対して前記行列におけるそれぞれの列ベクトルを適用する前記行列積演算を行うことを特徴とする請求項1に記載の画像識別システム。

【請求項3】

前記第1の演算装置は、前記入力画像と、前記第1パラメータから構成されたフィルタカーネルと、を用いて前記畳み込みフィルタ演算を行うことを特徴とする請求項1または2に記載の画像識別システム。

【請求項4】

前記演算手段は、前記第1演算結果である前記特徴と、前記行列と、の行列積を求める行列積演算を行うことを特徴とする請求項3に記載の画像識別システム。

【請求項5】

前記識別手段は、前記入力画像中の人物を識別することを特徴とする請求項1乃至4の何れか1項に記載の画像識別システム。

【請求項6】

前記第1の演算装置は、前記識別手段を有し、

前記第2の演算装置は、前記第2演算結果を前記第1の演算装置に対して出力し、

前記識別手段は、前記第2の演算装置から出力された前記第2演算結果に基づいて前記第1の演算装置のユーザの認証を行うことを特徴とする請求項1乃至5の何れか1項に記載の画像識別システム。

【請求項7】

前記第2の演算装置は、前記特徴より高次の特徴である前記第2演算結果を前記第1の演算装置に対して出力し、

前記識別手段は、前記第2の演算装置から出力された前記第2演算結果に基づいて前記ユーザの認証を行うことを特徴とする請求項6に記載の画像識別システム。

【請求項8】

前記画像識別システムは複数の前記第1の演算装置を有し、

前記演算手段は、複数の前記第1の演算装置による前記第1演算結果を連結した結果を用いて演算を行うことを特徴とする請求項1乃至7の何れか1項に記載の画像識別システム。

【請求項9】

前記演算手段は、複数の前記第1の演算装置による前記第1演算結果を連結した結果と前記第2パラメータとを用いて前記行列積演算を行うことを特徴とする請求項8に記載の画像識別システム。

【請求項10】

複数の階層を有するニューラルネットワークを用いて、入力画像から識別対象を識別するための識別処理を行う画像識別システムによる画像識別方法であって、

撮像装置である第1の演算装置が、撮像した画像を前記入力画像として入力し、該入力画像から特徴を抽出するための、第1パラメータを用いる畳み込みフィルタ演算を行い、前記第1の演算装置と異なる第2の演算装置が、

前記第1の演算装置から前記畳み込みフィルタ演算の結果である第1演算結果を取得し、

前記第1演算結果に基づく前記特徴と、重み係数パラメータである第2パラメータから構成される行列と、の行列積を求める行列積演算であって、該第2パラメータを記憶するためのメモリ容量が前記第1パラメータを記憶するためのメモリ容量と比べて多い前記行列積演算を行い、

前記識別処理において前記識別対象が、前記行列積演算の結果である第2演算結果を用いて識別される

ことを特徴とする画像識別方法。

【発明の詳細な説明】

【技術分野】

【0001】

10

20

30

40

50

本発明は、画像識別技術に関するものである。

【背景技術】

【0002】

ディープネット（或いはディープニューラルネット、ディープラーニングとも称される）と呼ばれる多階層のニューラルネットワークが、近年非常に大きな注目を集めている。ディープネットは、特定の演算手法を指すものではないが、一般的には、入力データ（例えば画像データ）に対して、階層的な処理（ある階層の処理結果を、その後段の階層の処理の入力とする）を行うものを指す。

【0003】

特に画像識別の分野では、畳み込みフィルタ演算を行う畳み込み層と、全結合演算を行う全結合層と、から構成されるディープネットが主流になりつつある。このようなディープネットでは、処理の前半に複数の畳み込み層を配置し、処理の後半に複数の全結合層を配置することが一般的である（非特許文献1）。

【0004】

畳み込みフィルタ演算の一例を図4を用いて説明する。図4において、符号401は処理対象画像を示しており、符号402はフィルタカーネルを示している。図4では、カーネルサイズが 3×3 のフィルタ演算を行う場合を示している。このような場合、次式に示す積和演算処理により、畳み込みフィルタ演算結果が算出される。

【0005】

【数1】

$$f_{i,j} = \sum_{s=1}^{\text{rowSize}} \sum_{t=1}^{\text{columnSize}} (d_{i+s-1,j+t-1} \times w_{s,t}) \quad \dots (1)$$

【0006】

ここで、 $d_{i,j}$ は、処理対象画像401上の画素位置(i, j)における画素値を表し、 $f_{i,j}$ は、画素位置(i, j)におけるフィルタ演算結果を表している。また、 $w_{s,t}$ は、画素位置($i+s-1, j+t-1$)における画素値に適用するフィルタカーネル402の値（フィルタ係数パラメータ）を表している。また、「columnSize」及び「rowSize」はフィルタカーネル402のサイズ（それぞれ列数、行数）を表している。フィルタカーネル402を処理対象画像401中で移動させつつ、上記の演算を行うことで、畳み込みフィルタ演算の出力結果を得ることができる。

【0007】

この畳み込みフィルタ演算、及びシグモイド変換に代表される非線形変換処理から、畳み込み層が構成される。入力データに対してこの畳み込み層の演算を階層的に繰り返し行うことで、画像の特徴を表現する特徴量が得られる。

【0008】

ディープネットにおいて複数の畳み込み層に続いて配置されている全結合層では、最後の畳み込み層の出力結果（特徴量）に対して、次式に示されるような行列積演算を行う。

【0009】

【数2】

$$C = A \times B = [a_1 \ \dots \ a_m] \begin{bmatrix} b_{1,1} & \dots & b_{1,n} \\ \vdots & \ddots & \vdots \\ b_{m,1} & \dots & b_{m,n} \end{bmatrix} \quad \dots (2)$$

【0010】

ここで、 m 次元のベクトルAは、最後の畳み込み層からの出力である特徴量のベクトルであり、 $m \times n$ 行列Bは、全結合層の重みパラメータである。演算結果である n 次元ベクトルCが、ベクトルAと行列Bとの行列積演算の結果となる。

10

20

30

40

50

【 0 0 1 1 】

この行列積演算及び、シグモイド変換に代表される非線形変換処理から、全結合層が構成される。畳み込み層から出力される特徴量に対してこの行列積演算を階層的に繰り返しを行うことで、最終識別結果が得られる。

【 0 0 1 2 】

ここで、上述した畳み込みフィルタ演算と行列積演算とでは、その演算を実行するプラットフォームに要求される要件がかなり異なる。以下では、これについてより詳細に説明する。

【 0 0 1 3 】

畳み込みフィルタ演算も行列積演算も、入力データとパラメータとの内積演算という意味では同種の演算とみなすことができる。畳み込みフィルタ演算の場合、入力データは入力画像或いは前階層の畳み込み層の出力結果であり、パラメータはフィルタ係数パラメータとなる。同様に、行列積演算の場合、入力データは最後の畳み込み層から出力された特徴量或いは前階層の全結合層の出力結果であり、パラメータは全結合層の重みパラメータとなる。このように、入力データとパラメータとの内積演算という意味では両演算は同種類の演算ではあるが、両演算の性質は大きく異なる。

10

【 0 0 1 4 】

畳み込み層で行われる畳み込みフィルタ演算では、前述のように、画像中をフィルタカーネルを移動させつつ演算を行う。つまり、フィルタカーネルの位置（スキャン位置）ごとに入力画像から部分データ（スキャンウインドウで切り出された部分画像）を切り出し、該部分データとフィルタカーネルとを用いた上記の演算を行うことにより、各位置での演算結果を得ることができる。

20

【 0 0 1 5 】

それに対し、全結合層で行われる行列積演算は、ベクトル状に並べられた入力データ（特徴量）に対して、重みパラメータで構成される行列を掛ける演算を行う。つまり、重みパラメータの行列の列ベクトルを切り出し、入力データと該切り出された列ベクトルとの演算を行うことにより、演算結果の各ベクトル要素を得ることができる。

【 0 0 1 6 】

以上をまとめると、畳み込み層の畳み込みフィルタ演算と全結合層の行列積演算とでは、入力データ量とパラメータ量とにより規定される演算特性に下記のような違いがある。つまり、畳み込みフィルタ演算では、入力データの複数の部分集合データのそれぞれに対して、同一のフィルタカーネルを適用することで、畳み込みフィルタ演算結果を得るものである。従って、入力データ量に比してフィルタカーネル（フィルタ係数パラメータ）の量は少ないものとなる。

30

【 0 0 1 7 】

それに対し、行列積演算では、同一入力データに対して、重み係数パラメータ（行列）の複数の部分集合（列ベクトル）のそれぞれを適用することで、行列積演算結果を得るものである。従って、入力データ量に比して重み係数パラメータの量は多いものとなる。

【 0 0 1 8 】

また、畳み込みフィルタ演算も行列積演算も、演算量は入力データ量に比例する。畳み込みフィルタ演算では、入力データ量（入力画像のサイズ）にフィルタカーネルのサイズを掛けたものが演算量と言える。従って、畳み込みフィルタ演算の演算量は入力データ量に比例する（入力画像の端部の処理は無視する）。同様に、行列積演算では、入力データ量に重み係数パラメータの行列の列数（列ベクトルの数）を掛けたものが演算量と言える。従って、行列積演算の演算量も入力データ量に比例する。

40

【 0 0 1 9 】

これらのことから、畳み込み層の畳み込みフィルタ演算と全結合層の行列積演算とでは、演算の特性に関して以下のようにも言える。つまり、畳み込みフィルタ演算は、演算量に比してフィルタカーネル（フィルタ係数パラメータ）の量は少ないものと言え、行列積演算は、演算量に比して重み係数パラメータの量は多いものと言える。

50

【0020】

以上説明したように、ディープネットにおける演算処理には、入力データ量とパラメータ量により規定される演算特性がそれぞれ異なる2つの演算（畳み込み層における畳み込みフィルタ演算と全結合層における全結合演算）が含まれていることがわかる。

【先行技術文献】**【特許文献】****【0021】**

【特許文献1】特開平10-171910号公報

【特許文献2】WO2013/102972

【非特許文献】

10

【0022】

【非特許文献1】Krizhevsky, A., Sutskever, I. and Hinton, G. E. "ImageNet Classification with Deep Convolutional Neural Networks" NIPS 2012

【発明の概要】**【発明が解決しようとする課題】****【0023】**

畳み込み層における畳み込みフィルタ演算や、全結合層における行列積演算では、多数の積和演算を行う必要があるため処理量が多く、処理時間のかかる処理となる。また、畳み込みフィルタ演算に必要となるフィルタカーネルや行列積演算に必要となる重みパラメータを格納しておくメモリについても、ディープネットの階層（畳み込み層の数や全結合層の数）が多数になると、より大容量のメモリが要求される。

20

【0024】

従って一般的に、ディープネットを処理するためには潤沢な演算リソースが必要とされ、PC (Personal Computer)、サーバ、クラウド等に対し、演算リソースが貧弱である組み込み機器で処理することはこれまで考慮されて来なかった。特に、パラメータ量が多い全結合層の行列積演算も含めた一連のディープネット演算を組み込み機器で行うことは、組み込み機器に許容されるメモリ量の観点から現実的でなかった。また、同様に演算量の多い畳み込み層の畳み込みフィルタ演算も含めた一連のディープネット演算をPCやサーバやクラウド等で行うとすると、それらの演算資源を逼迫させる恐れがあった

特許文献1では、2次元ニューラルネットワークを、2つの1次元ニューラルネットワークに分解して演算を実行することで結合数（パラメータ数）の削減を行っている。しかしながら、特許文献1で開示されている手法では、複数の演算特性を持つ演算から構成されている一連の演算を、それぞれの演算特性を考慮して分割し、それぞれの演算を好ましい処理プラットフォームにおいて処理するということは考慮されていない。つまり、これまで詳細に説明したように、畳み込みフィルタ演算と行列積演算とでは演算特性に違いがあるが、この演算特性に応じて処理プラットフォームを変えるようなことは考慮されてこなかった。

30

【0025】

また、一連のディープネットの演算全てをサーバやクラウド等で行うとすると、画像を撮影した撮影機器から該画像を、ディープネット演算を行うサーバやクラウド等に送信する必要がある。伝送路を効率的に利用するという観点からすると、送信する画像のデータ量を削減した方が望ましい。しかしながらこれまで、ディープネットの演算を行うことと、送信する画像のデータ量を削減することは別のこととして扱われ、全体として効率の良い手法が検討されていなかった。

40

【0026】

特許文献2では、プライバシー保護を目的として、画像の撮像端末において画像から特徴量の抽出を行い、抽出した特徴量をサーバに送信して画像中的人物位置を特定する手法を開示している。しかしながら、この手法は撮像端末とサーバで行う処理に関し、それぞれの演算特性を考慮して割り振ったものではない。従って、特許文献2の手法では、演算資源を効率的に使用することや、アプリケーション（特許文献2では人物位置の特定を想

50

定するアプリケーション)を変更したときの柔軟性等は考慮されていなかった。

【0027】

本発明はこのような問題に鑑みてなされたものであり、入力データ量とパラメータ量により規定される演算特性の異なるそれぞれの演算を好ましい処理プラットフォームにおいて処理するための技術を提供する。

【課題を解決するための手段】

【0028】

本発明の一様態は、複数の階層を有するニューラルネットワークを用いて、入力画像から識別対象を識別するための識別処理を行う画像識別システムであって、

撮像した画像を前記入力画像として入力し、該入力画像から特徴を抽出するための、第1パラメータを用いる畳み込みフィルタ演算を行う撮像装置である第1の演算装置と、

前記第1の演算装置と異なる第2の演算装置であって、

前記第1の演算装置から前記畳み込みフィルタ演算の結果である第1演算結果を取得する取得手段と、

前記第1演算結果に基づく前記特徴と、重み係数パラメータである第2パラメータから構成される行列と、の行列積を求める行列積演算であって、該第2パラメータを記憶するためのメモリ容量が前記第1パラメータを記憶するためのメモリ容量と比べて多い前記行列積演算を行う演算手段と

を有する前記第2の演算装置と、

前記行列積演算の結果である第2演算結果を用いて、前記識別処理において前記識別対象を識別する識別手段と

を備えることを特徴とする。

【発明の効果】

【0029】

本発明の構成によれば、入力データ量とパラメータ量により規定される演算特性の異なるそれぞれの演算を好ましい処理プラットフォームにおいて処理することができる。

【図面の簡単な説明】

【0030】

【図1】画像識別システムの構成例を示すブロック図。

【図2】ディープネットの演算例を示す図。

【図3】画像識別システムの構成例を示すブロック図。

【図4】畳み込みフィルタ演算の一例を示す図。

【図5】画像識別システムの構成例を示すブロック図。

【発明を実施するための形態】

【0031】

以下、添付図面を参照し、本発明の実施形態について説明する。なお、以下に説明する実施形態は、本発明を具体的に実施した場合の一例を示すもので、特許請求の範囲に記載した構成の具体的な実施例の1つである。

【0032】

[第1の実施形態]

本実施形態では、演算量もパラメータ量も多大であるディープネットの処理を、低コスト且つ柔軟に実現するための画像識別システムの一例について説明する。また本実施形態では、一連のディープネットの処理を(上記非線形変換処理を除いて)、入力データの量(或いは入力データの量と比例関係にある演算量)とパラメータの量とで規定される演算特性の違いによって2種類の演算(第一、第二の演算)に分けている。そして、該2種類の演算を、それぞれの演算の演算特性(第一の演算特性、第二の演算特性)に応じた処理プラットフォームにて実行するようにしている。

【0033】

本実施形態では、第一の演算として、入力データの量に比してパラメータの量が少ない演算を考え、第二の演算として、入力データの量に比してパラメータの量が多い演算を考

10

20

30

40

50

える。然るに、第一の演算特性とは、「入力データの量に比してパラメータの量が少ない」という演算特性であり、第二の演算特性とは、「入力データの量に比してパラメータの量が多い」という演算特性である。

【0034】

「背景技術」の項で詳細に説明したように、一連のディープネット処理における演算のうち、畳み込み層における畳み込みフィルタ演算は、第一の演算に相当する。これは、畳み込みフィルタ演算が、スキャン位置ごとに、入力画像から部分データ（部分画像）を切り出し、該切り出された部分データとフィルタカーネルとの上記演算を行うことにより、各スキャン位置での演算結果を得るような演算であるからである。つまり、この場合の第一の演算は、同一のフィルタカーネルと、切り出された複数の部分データそれぞれとの演算となっている。10

【0035】

また、全結合層における行列積演算は、第二の演算に相当する。これは、行列積演算が、重みパラメータ行列の列ベクトルを切り出して、入力データと切り出された重みパラメータとの上記演算を行うことにより、演算結果の各ベクトル要素を得るような演算であるからである。

【0036】

本実施形態では上記のように、畳み込み層における畳み込みフィルタ演算を、第一の演算特性を有する第一の演算とし、全結合層における行列積演算を、第二の演算特性を有する第二の演算とする場合の例について説明する。さらに本実施形態では、第一の演算を組み込みデバイスで行い、第二の演算をPC（パーソナルコンピュータ）やサーバ等のコンピュータ装置（少なくとも組み込みデバイスよりも潤沢なメモリ容量のメモリを使用可能な装置）で行う場合の例について説明する。組み込みデバイスとしては、画像撮像装置（例えばカメラ）内の演算専用ハードウェアを想定する。20

【0037】

一般に組み込みデバイスで想定されるハードウェアは、特定の演算を高速に処理するよう設計される。従って、畳み込みフィルタ演算を効率よく処理するようなハードウェアを作成することは、公知の技術（例えば特許5184824号公報や特許5171118号公報等）を使って可能である。

【0038】

しかしながら組み込みデバイス内に大量のパラメータを格納することは困難である。大量のパラメータを格納するためには、大容量のメモリが必要になる。しかし、一般に回路面積や実装面積が制限される組み込みデバイス内にそのような大容量メモリを用意することは困難である。また、コストの面からも、カメラ等の画像撮像装置の内部に大容量メモリを用意することは、現実的でない。つまり、組み込みデバイスでの演算は、その演算に必要なパラメータの量が少ないような演算であることが望まれる。逆にパラメータの量が多い演算を組み込みデバイスで行うのは非現実的であるといえる。30

【0039】

これに対し、サーバに代表されるような汎用の計算機（PCやクラウド等）は、大容量メモリを装備している若しくは使用可能であることが一般的である。従って、パラメータの量が多い演算をサーバで行うのは理に適っているといえる。40

【0040】

本実施形態では、演算の演算特性（パラメータ量の大小等）と、演算プラットフォームの特性（大容量メモリ装備の現実性）と、を考慮して、一連のディープネットの処理におけるそれぞれの演算の演算プラットフォームへの割り当てを行う。これにより、低コストでディープネットの処理を実現している。

【0041】

本実施形態では、画像からの特徴量抽出処理に畳み込みフィルタ演算を用い、抽出した特徴量を用いた識別処理にパーセプトロンに代表される行列積演算を用いる構成になっているものを典型的なディープネットとしている。この特徴量抽出処理は畳み込みフィルタ50

演算を何度も繰り返す多階層処理であることが多く、また識別処理にも全結合の多階層のパーセプトロンが用いられることがある。この構成は、近年盛んに研究されているディープネットとしては非常に一般的な構成である。

【0042】

ここで、図2を用いてディープネットの演算例について説明する。図2では、入力層に入力された入力画像1101に対して、畳み込みフィルタ演算により特徴抽出を行って特徴量1107を取得し、該取得した特徴量1107に対して識別処理を行って識別結果1114を得るような処理を示している。入力画像1101から特徴量1107を得るまでに畳み込みフィルタ演算を何度も繰り返している。また、特徴量1107に対して全結合のパーセプトロン処理を複数回行い、最終的な識別結果1114を得ている。

10

【0043】

まず、前半の畳み込みフィルタ演算について説明する。特徴面1103a～1103cは、第1段目の階層1108の特徴面である。特徴面とは、所定の特徴抽出フィルタ（畳み込みフィルタ演算及び非線形処理）の検出結果を示すデータ面である。特徴面1103a～1103cは、入力画像1101に対する畳み込みフィルタ演算及び上記非線形処理により生成される。例えば、特徴面1103aは、フィルタカーネル11021aを用いた畳み込みフィルタ演算及び該演算の結果の非線形変換により得られる。なお、図2中のフィルタカーネル11021b及び11021cはそれぞれ、特徴面1103b及び1103cを生成する際に使用されるフィルタカーネルである。

【0044】

次に、第2段目の階層1109の特徴面1105aを生成する演算について説明する。特徴面1105aは前段の階層1108の3つの特徴面1103a～1103cと結合している。従って、特徴面1105aのデータを算出する場合、特徴面1103aに対してはフィルタカーネル11041aで示すカーネルを用いた畳み込みフィルタ演算を行い、この結果を保持する。同様に、特徴面1103b及び1103cに対しては、それぞれフィルタカーネル11042a及び11043aの畳み込みフィルタ演算を行い、これらの結果を保持する。これらの3種類のフィルタ演算の終了後、それぞれのフィルタ演算結果を加算して非線形変換処理を行う。以上の処理を画像全体に対して処理することにより、特徴面1105aを生成する。特徴面1105bの生成も同様に、階層1108の特徴面1103a～1103cに対するフィルタカーネル11041b、11042b、11043bによる3つの畳み込みフィルタ演算を行い、それぞれのフィルタ演算結果を加算して非線形変換処理を行う。

20

【0045】

また、第3段目の階層1110の特徴量1107の生成の際には、前段の階層1109の特徴面1105a～1105bに対するフィルタカーネル11061及び11062による2つの畳み込みフィルタ演算を行う。

30

【0046】

続いて後半のパーセプトロン処理を説明する。図2では2階層のパーセプトロンになっている。パーセプトロンは、入力特徴量のそれぞれの要素に対する重み付き和を非線形変換したものである。従って、特徴量1107に対して行列積演算を行い、その結果に対して非線形変換を行えば、中間結果1113を得ることができる。さらに同様の処理を繰り返せば、最終的な識別結果1114を得ることができる。

40

【0047】

次に、このような図2のディープネットを用いて画像識別を行う画像識別システムの構成例について、図1のブロック図を用いて説明する。図1に示す如く、本実施形態に係る画像識別システム101は、カメラなどの画像撮像装置102と、サーバやPCなどの演算装置106と、を有する。そして画像撮像装置102と演算装置106とは有線若しくは無線でもって互いにデータ通信が可能なように接続されている。

【0048】

画像識別システム101は、画像撮像装置102が撮像した撮像画像に対してディープ

50

ネットを用いた演算を行い、その結果として該撮像画像中に何が写っているのか（例えば、人、飛行機等）を識別するものである。

【0049】

先ず、画像撮像装置102について説明する。画像撮像装置102は、画像を撮像し、該画像に対し、上記のディープネットにより実現される画像識別処理のうち前半の処理、即ち畳み込みフィルタ演算及び非線形変換の結果を、後段の演算装置106に対して出力する。

【0050】

画像取得部103は、光学系やCCD、画像処理回路などにより構成されており、外界の光を映像信号に変換し、該変換した映像信号に基づく画像を撮像画像として生成し、該生成した撮像画像を入力画像として後段の第一の演算部104に対して送出する。10

【0051】

第一の演算部104は、画像撮像装置102内に設けられている組み込みデバイス（例えば専用ハードウェア）により構成されており、画像取得部103から受けた入力画像に対し、畳み込みフィルタ演算及び非線形変換を行って、特徴量を抽出する。これにより、現実的な処理リソースでの処理を可能にしている。第一の演算部104は上記の通り、周知の組み込みデバイスであり、その具体的な構成については公知の技術（例えば特許5184824や特許5171118）で実現できる。

【0052】

第一のパラメータ格納部105には、第一の演算部104が畳み込みフィルタ演算で用いるパラメータ（フィルタカーネル）が格納されている。これまで何度も述べたように、畳み込みフィルタ演算は入力データ（或いはそれに比例する演算量）に比してパラメータ量が少ないという演算特性を有するので、組み込みデバイス内のメモリでもフィルタカーネルを格納することが可能である。20

【0053】

第一の演算部104は、入力画像と第一のパラメータ格納部105に格納されているフィルタカーネルとを用いて畳み込みフィルタ演算を何度も行うことで、該入力画像から特徴量を算出する。つまり図2の特徴量1107を算出するまでの畳み込みフィルタ演算を第一の演算部104で行う。第一の演算部104は、算出した特徴量1107を第一の演算結果として、演算装置106に対して送信する。30

【0054】

次に、演算装置106について説明する。演算装置106は、画像撮像装置102から送信された第一の演算結果に対し、上記のディープネットにより実現されている画像識別処理のうち後半の処理、すなわち、全結合演算及び非線形変換の結果を出力する。

【0055】

第二の演算部107は、演算装置106内に設けられている汎用演算器により実現されている。第二のパラメータ格納部108には、第二の演算部107が全結合演算で用いるパラメータ、すなわち、行列積演算で必要となるパラメータ（重み係数パラメータ）が格納されている。前述のように、演算装置106に大容量メモリを搭載することは一般的であるので、パラメータ量が多いという第二の演算特性を持つ演算（行列積演算）を、演算装置106側で（第二の演算部107が）行うことは非常に合理的である。40

【0056】

第二の演算部107は、画像撮像装置102から送信された第一の演算結果と第二のパラメータ格納部108に格納されている重み係数パラメータとを用いて行列積演算を何度も行うことで最終的な識別結果を算出する。つまり図2の特徴量1107から、最終的な識別結果1114を算出するまでの行列積演算を第二の演算部107で行う。本実施形態では、入力画像中に何が写っているのかを識別するディープネットの処理を行っているので、最終的な識別結果としては、人物や飛行機等の識別クラスラベルが出力されることになる。

【0057】

10

20

30

40

50

なお、第二の演算部 107 による識別結果の出力先や出力形態については特定の出力先、出力形態に限るものではない。例えば、ディスプレイなどの表示装置に識別結果を画像や文字などでもって表示しても良いし、外部の機器に対して識別結果を送信しても良いし、識別結果をメモリに格納しても良い。

【0058】

このように本実施形態によれば、それぞれ異なる演算特性を持つ複数の演算を含むディープネット処理を、それぞれの演算特性に適した演算プラットフォームで処理するように分割することで、低コストで画像識別システムを構成することができる。

【0059】

またディープネットにおける畳み込み層では、サブサンプリング（畳み込みフィルタ演算のスキャンウインドウを動かすストライドを大きくする）や、ブーリング（周辺画素との統合）により、階層が進むにつれて特徴面サイズが小さくなることが一般的である。このため、図2の入力画像1101のサイズよりも、特徴量1107のサイズが小さくなることもある（例えば非特許文献1記載のディープネット）。このため、画像撮像装置102から演算装置106に対して入力画像そのものを送るよりも、画像撮像装置102にて入力画像から特徴量を抽出し、該抽出した特徴量を演算装置106に送った方が、送信するデータ量が小さくなる。つまり、通信路の効率的な利用という観点からも本実施形態は有効であるといえる。10

【0060】

また、一般にディープネットの前半部で行われる畳み込み層の演算は特徴量抽出処理と呼ばれる。特徴量抽出処理は、アプリケーション（ディープネットを用いて実現しようとしている画像識別タスク）によらず、共通にできることが多い。実際、非特許文献1で示されたディープネットの特徴量抽出処理部分（畳み込み層部分）を、各種タスクで使いまわすことがよく行われている（Ali Sharif Razavian, Hossein Azizpour, Josephine Sullivan, Stefan Carlsson, "CNN Features off-the-shelf: an Astounding Baseline for Recognition"）。つまり、畳み込み層の構成（フィルタカーネル、ネットワーク構成）はそのままにして、全結合層の構成（重み係数パラメータ、ネットワーク構成）を変えるだけで、アプリケーションの切り替えを実現できる。20

【0061】

従って、本実施形態のように、畳み込み層の演算を行う演算プラットフォームと全結合層の演算を行う演算プラットフォームとに分離していれば、次のような効果がある。つまり、全結合層の演算プラットフォームの設定（重み係数パラメータ、ネットワーク構成）さえ変更すれば、各種アプリケーションを実現することができる。30

【0062】

また、本実施形態のように、画像撮像装置102と演算装置106とを有する画像識別システムでは、演算装置106側の変更だけで、各種アプリケーションの切り替えや追加を実現できる。一般に、画像撮像装置102の設定を変更することは極めて煩わしい。然るに、その手間なく、アプリケーションを切り替えたり、新たなアプリケーションを追加できたりすることは、画像識別システムを維持、拡張していくうえで非常に有用な利点であり、高い柔軟性を有していると言える。40

【0063】

[第2の実施形態]

本実施形態では、演算装置106に対して複数台の画像撮像装置102が通信可能に接続されており、該複数台の画像撮像装置102の各々から特徴量を演算装置106に対して送信する画像識別システムについて説明する。本実施形態を含め、以下の実施形態では、第1の実施形態との差分について重点的に説明し、以下で特に触れない限りは第1の実施形態と同様であるものとする。

【0064】

カメラを複数台用意し、複数台のカメラで撮影されたそれぞれの画像を元に、画像中に何が写っているのかを特定するようなアプリケーションは、監視カメラでは一般的である50

。例えば、入退室管理のアプリケーションでは、入退室の許可を求める人物を複数台のカメラで撮影し、その画像から、対象人物のIDを特定することが行われている。

【0065】

本実施形態に係る画像識別システムの構成例について、図3のブロック図を用いて説明する。図3に示す如く、本実施形態に係る画像識別システム301は、演算装置306に複数台の画像撮像装置102a～102cが通信可能に接続されている。画像撮像装置の参照番号102に付したa、b、cはそれぞれ、各々の画像撮像装置を識別するために付けたものであり、画像撮像装置102a～102cは何れも図1の画像撮像装置102と同様の構成を有し、同様の動作を行う。なお、図3では画像撮像装置の台数を3としているが、この数に限るものではない。

10

【0066】

次に、演算装置306について説明する。第二の演算部307は、演算装置306内に設けられている汎用演算器により実現されている。第二の演算部307は、画像撮像装置102a～102cの各々から第一の演算結果を受けると行列積演算及び非線形変換を行つて、画像撮像装置102a～102cの各々が撮像した画像から対象人物の識別情報（例えばID）を特定して出力する。本実施形態では、画像撮像装置102a～102cの各々から第一の演算結果を受けるので、第二の演算部307は、それらを連結して新たな特徴量を作成し、その特徴量に対して行列積演算を行う。第二の演算部307は、演算装置306内にある汎用演算器により実現されている。

20

【0067】

第二のパラメータ格納部308には、第二の演算部307が行う行列積演算で必要となるパラメータ（重み係数パラメータ）が格納されている。本実施形態では、前述のように、3つの第一の演算結果を連結した特徴量に対して行列積演算を行うので、第二のパラメータ格納部308に格納されている重み係数パラメータの量もその分大きくなる。

【0068】

第二の演算部307では、複数の第一の演算結果と第二のパラメータ格納部308に格納されている重み係数パラメータと用いて行列積演算を何度も行うことで最終的な識別結果を算出する。本実施形態では、画像中に写っている人物の識別情報（名前等）を特定する処理を行っているので、人物を特定する識別情報が最終的な識別結果として出力される。

30

【0069】

本実施形態では、ディープネットにおける畳み込み層の演算を行う演算プラットフォームと全結合層の演算を行う演算プラットフォームとを分離している。このように構成することは、それぞれの演算特性に適した演算プラットフォームを選択できることのみならず、本実施形態で説明したように、画像撮像装置が複数に増えたとしても、柔軟に対応できる画像識別システムを実現することにも繋がっている。例えば、画像撮像装置内で全てのディープネットの処理を行うような画像識別システムでは、画像撮像装置が1台の場合には、画像撮像装置だけで全ての処理が完結するが、画像撮像装置が複数台になると、複数の処理結果を統合するものが必要になってしまふ。これでは、柔軟なシステムとは言い難い。

40

【0070】

[第3の実施形態]

第1、第2の実施形態では、第二の演算部で最終的な識別結果を算出していたが、第二の演算部で算出した結果を、再度第一の演算部に戻し、第一の演算部にて最終的な識別結果を算出してもよい。このように構成すると、個々の画像撮像装置固有の設定や、その画像撮像装置で画像を撮像した時の情報、或いは個々の画像撮像装置を操作するユーザの嗜好等を考慮に入れて最終的な識別結果を推定することが可能となる。また、ディープネットを用いた画像識別アプリケーションの幅が広がる。

【0071】

例えば、スマートフォン等でディープネットによる顔画像を用いたログイン認証を行う

50

のようなアプリケーションを実現する場合を考える。そのような場合、スマートフォンに内蔵された画像撮像装置でユーザの顔画像を撮影し、顔画像に対して畳み込み層の演算を行って特徴量（第一の演算結果）を算出し、演算装置に送る。演算装置では全結合層の演算を行って、さらに高次の特徴量（第二の演算結果）を算出し、再度画像撮像装置に送り返す。画像撮像装置では、予め登録されている高次特徴量と、今回演算装置から送り返されてきた高次特徴量とを比較し、ログインの可否を判断する。

【0072】

このような画像識別システムの構成例について、図5のブロック図を用いて説明する。図5に示す如く、本実施形態に係る画像識別システム501は、画像撮像装置502と、演算装置106と、を有し、それぞれが互いにデータ通信が可能なように接続されている。第二の演算部107は、第二の演算結果を算出すると、該第二の演算結果を画像撮像装置502に対して送信する。
10

【0073】

次に、画像撮像装置502について説明する。第一の演算部504は、画像撮像装置502内に設けられている組み込みデバイス（例えば専用ハードウエア）により構成されており、第一のパラメータ格納部105に加えて、第三のパラメータ格納部509を有する。第一の演算部504は第1の実施形態と同様に、画像取得部103からの入力画像と第一のパラメータ格納部105に格納されているパラメータとを用いて畳み込みフィルタ演算を行い、その演算結果を非線形変換したものを第二の演算部107に送信する。また、第一の演算部504は、第二の演算部107から第二の演算結果を受けると、第三のパラメータ格納部509に格納されているパラメータを用いた演算を行い、最終的な識別結果（第三の演算結果）を得る。
20

【0074】

第三のパラメータ格納部509には、画像撮像装置502固有の情報が格納されている。例えば、前述のログイン可否のアプリケーションを実施する場合、第三のパラメータ格納部509には、正式ユーザの登録情報が格納されている。正式ユーザの登録情報としては、予めユーザ登録を行う際に、自身の顔画像に対して、第二の演算結果を得るまでの処理を行って得られた第二の演算結果を用いればよい。このようにしておけば、ユーザ登録時に算出された第二の演算結果と、ログイン認証時に算出された第二の演算結果とを比較することで、ログインの可否を判断できる。前述のログイン可否のアプリケーションを実施する場合、このようなログイン可否判断処理は、第一の演算部504によって行われる。
30

【0075】

ここで第一の演算結果を登録情報としないのは以下のようない理由からである。第一の演算結果は畳み込み層の演算に基づく情報であるため、局所的な特徴量の集まりであるといえる。従って、第一の演算結果を用いただけでは、表情、照明、顔向き等の変動にロバストに認証することは難しい。そこで、より大域的な特徴量の抽出が期待できる第二の演算結果を登録情報として用いる方が、認証精度が向上することが予想されるためである。

【0076】

このように構成することで、画像撮像装置固有の情報（本実施形態では、予め登録された正式ユーザの情報）を用いた画像識別アプリケーションの実現が可能となる。画像撮像デバイス固有の情報（例えば正式ユーザの情報）も含めて演算装置に送れば、同様のことが実現可能ではあるが、その場合には、セキュリティの確立やプライバシー保護等、システムを構成する上の要件が増えることにつながる。また、そもそも、個人情報につながるような情報を演算装置に送信することに抵抗や不安を感じるユーザも存在するため、本実施形態のように構成した方が、アプリケーションを利用するユーザの心理的抵抗感が低減することが期待できる。
40

【0077】

なお、上記の各実施形態の構成は、その一部若しくは全部を適宜組み合わせて新たな構成の画像識別システムを構築することも可能である。また、第一の演算部や第二の演算部
50

は、その全てを専用ハードウェア（CPU等のプロセッサやRAMやROM等のメモリなどが配置された回路）で構成しても良いが、その一部の構成をソフトウェアで構成しても良い。その場合、そのソフトウェアは、対応する演算部のプロセッサが実行することで、対応する機能を実現する。また、上記の各実施形態で説明した画像識別システムは何れも、以下のような要件を満たす画像識別システムの一例として説明したものである。

【 0 0 7 8 】

- ・ 入力画像に対する識別処理における複数の演算処理のうち、使用するパラメータの量が、該パラメータを適用するデータの量に比して小さい演算処理を行う第1の演算装置
 - ・ 入力画像に対する識別処理における複数の演算処理のうち、使用するパラメータの量が、該パラメータを適用するデータの量に比して大きい演算処理を行う第2の演算装置
 - ・ 第2の演算装置は、第1の演算装置よりもメモリ容量が大きいメモリが使用可能である

10

(その他の実施例)

本発明は、上述の実施形態の1以上の機能を実現するプログラムを、ネットワーク又は記憶媒体を介してシステム又は装置に供給し、そのシステム又は装置のコンピュータにおける1つ以上のプロセッサーがプログラムを読み出し実行する処理でも実現可能である。また、1以上の機能を実現する回路(例えば、ASIC)によっても実現可能である。

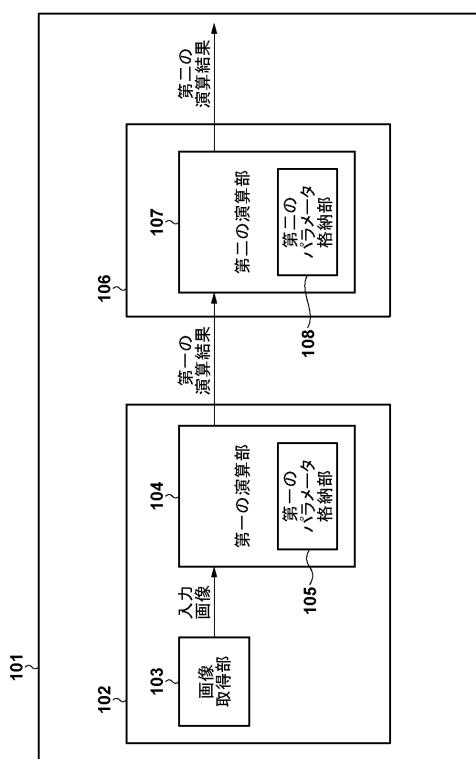
【符号の説明】

(0 0 7 9)

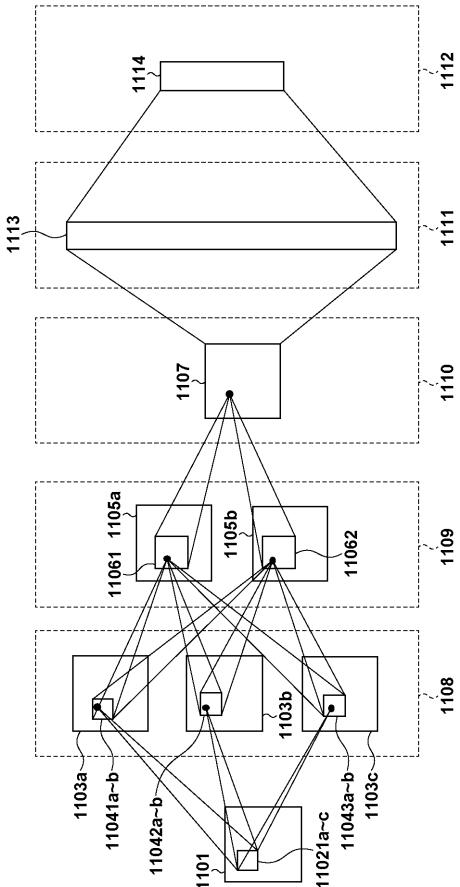
102：画像撮像装置 104：第一の演算部 106：演算装置 107：第二の演算部

20

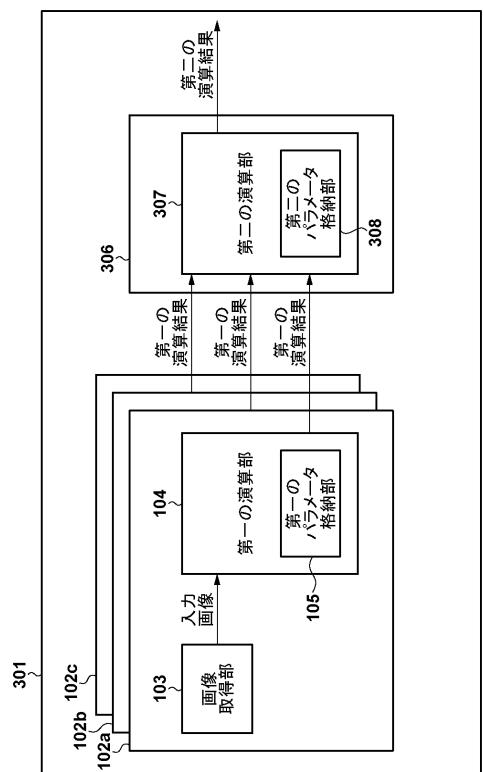
(1)



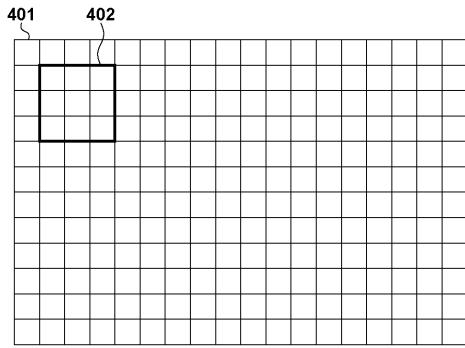
(2)



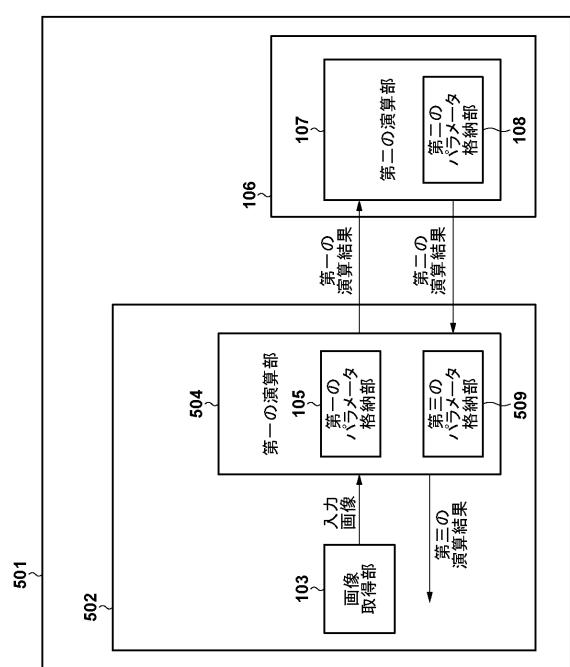
【図3】



【図4】



【図5】



フロントページの続き

(72)発明者 山本 貴久
東京都大田区下丸子3丁目30番2号 キヤノン株式会社内

(72)発明者 加藤 政美
東京都大田区下丸子3丁目30番2号 キヤノン株式会社内

(72)発明者 森 克彦
東京都大田区下丸子3丁目30番2号 キヤノン株式会社内

(72)発明者 伊藤 嘉則
東京都大田区下丸子3丁目30番2号 キヤノン株式会社内

(72)発明者 野村 修
東京都大田区下丸子3丁目30番2号 キヤノン株式会社内

審査官 武田 広太郎

(56)参考文献 國際公開第2015/066628 (WO, A1)
特開2015-082195 (JP, A)

(58)調査した分野(Int.Cl., DB名)
G 06 T 7 / 00