

(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2019-75003**  
(P2019-75003A)

(43) 公開日 **令和1年5月16日(2019.5.16)**

(51) Int. Cl. F I テーマコード (参考)  
**G06F 7/552 (2006.01)** G O 6 F 7/552 A  
**G06F 7/523 (2006.01)** G O 6 F 7/523

審査請求 未請求 請求項の数 8 O L (全 25 頁)

(21) 出願番号	特願2017-202006 (P2017-202006)	(71) 出願人	504182255 国立大学法人横浜国立大学 神奈川県横浜市保土ヶ谷区常盤台79番1号
(22) 出願日	平成29年10月18日 (2017.10.18)	(74) 代理人	100149548 弁理士 松沼 泰史
		(74) 代理人	100188558 弁理士 飯田 雅人
		(74) 代理人	100196058 弁理士 佐藤 彰雄
		(72) 発明者	島 圭介 神奈川県横浜市保土ヶ谷区常盤台79番1号 国立大学法人横浜国立大学内
		(72) 発明者	関島 律 神奈川県横浜市都筑区東山田4-5-5-402

(54) 【発明の名称】 近似計算装置、近似計算方法及びプログラム

(57) 【要約】

【課題】 べき乗の計算を比較的簡単な計算に置き換えられるようにする。

【解決手段】 近似計算装置が、2の実数乗の指数を整数部分と0以上1未満の小数部分とに分離する指数分離部と、前記小数部分について2のべき乗を直線近似で計算する小数部分計算部と、前記整数部分についての2のべき乗の計算結果と、前記小数部分計算部による前記小数部分についての2のべき乗の計算結果とを乗算する乗算部と、を備える。

【選択図】 図1

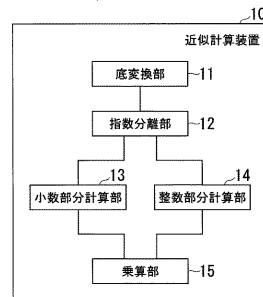


図1

**【特許請求の範囲】****【請求項 1】**

2の実数乗の指数を整数部分と0以上1未満の小数部分とに分離する指数分離部と、前記小数部分について2のべき乗を直線近似で計算する小数部分計算部と、前記整数部分についての2のべき乗の計算結果と、前記小数部分計算部による前記小数部分についての2のべき乗の計算結果とを乗算する乗算部と、  
を備える近似計算装置。

**【請求項 2】**

前記整数部分について2のべき乗をシフト演算で計算する整数部分計算部を備える、請求項1に記載の近似計算装置。

10

**【請求項 3】**

前記小数部分計算部は、前記0以上1未満の区間を分割した区間毎の直線近似にて前記小数部分についての2のべき乗を計算する、請求項1または請求項2に記載の近似計算装置。

**【請求項 4】**

前記小数部分計算部は、前記0以上1未満の区間を分割した区間における2のべき乗の両端を結んだ線分にて前記直線近似を行う、請求項3に記載の近似計算装置。

**【請求項 5】**

ネイピア数の実数乗を2の実数乗に換算する底変換部を備える、請求項1から4のいずれか一項に記載の近似計算装置。

20

**【請求項 6】**

前記底変換部は、混合ガウスモデルに基づくベイズ識別におけるネイピア数の実数乗を2の実数乗に換算する、請求項5に記載の近似計算装置。

**【請求項 7】**

2の実数乗の指数を整数部分と0以上1未満の小数部分とに分離する指数分離ステップと、  
前記小数部分について2のべき乗を直線近似で計算する小数部分計算ステップと、  
前記整数部分についての2のべき乗の計算結果と、前記小数部分計算ステップでの前記小数部分についての2のべき乗の計算結果とを乗算する乗算ステップと、  
を含む近似計算方法。

30

**【請求項 8】**

プログラミング可能な装置に

2の実数乗の指数を整数部分と0以上1未満の小数部分とに分離する指数分離ステップと、  
前記小数部分について2のべき乗を直線近似で計算する小数部分計算ステップと、  
前記整数部分についての2のべき乗の計算結果と、前記小数部分計算ステップでの前記小数部分についての2のべき乗の計算結果とを乗算する乗算ステップと、  
を実行させるためのプログラム。

**【発明の詳細な説明】****【技術分野】**

40

**【0001】**

本発明は、近似計算装置、近似計算方法及びプログラム、特に混合ガウスモデルを用いたベイズ識別に好適な近似計算装置、近似計算方法及びプログラムに関する。

**【背景技術】****【0002】**

筋電義手やブレインマシンインターフェース等において生体信号のパターン識別が用いられるが、これには混合ガウスモデル(Gaussian Mixture Model; GMM)に基づくベイズ識別が有効であることが知られているが、GMMの演算をリアルタイムで処理しようとすると実装コストが膨大となるという問題がある。

**【0003】**

50

一方、リアルタイム処理においてべき乗の計算を行う場合がある。

例えば、特許文献1には、混合ガウスモデルを適用したニューラルネットワークに生体信号を入力し、電動義手を操作することが記載されている。混合ガウスモデルで事後確率の計算を行う際、指数関数を含む計算を行う。

【先行技術文献】

【特許文献】

【0004】

【特許文献1】特許第3816762号公報

【発明の概要】

【発明が解決しようとする課題】

【0005】

リアルタイム処理においてべき乗の計算を行う場合、べき乗の計算を高速に行う必要がある。べき乗の計算を比較的簡単な計算に置き換えることができれば、計算の高速化を図ることができ、また、比較的容易に計算のハードウェア実装を行える。

【0006】

本発明は、べき乗の計算を比較的簡単な計算に置き換えることができる近似計算装置、近似計算方法およびプログラムを提供する。

【課題を解決するための手段】

【0007】

本発明の第1の態様によれば、近似計算装置は、2の実数乗の指数を整数部分と0以上1未満の小数部分とに分離する指数分離部と、前記小数部分について2のべき乗を直線近似で計算する小数部分計算部と、前記整数部分についての2のべき乗の計算結果と、前記小数部分計算部による前記小数部分についての2のべき乗の計算結果とを乗算する乗算部と、を備える。

【0008】

前記整数部分について2のべき乗をシフト演算で計算する整数部分計算部を備えるようにしてもよい。

【0009】

前記小数部分計算部は、前記0以上1未満の区間を分割した区間毎の直線近似にて前記小数部分についての2のべき乗を計算するようにしてもよい。

【0010】

前記小数部分計算部は、前記0以上1未満の区間を分割した区間における2のべき乗の両端を結んだ線分にて前記直線近似を行うようにしてもよい。

【0011】

ネイピア数の実数乗を2の実数乗に換算する底変換部を備えるようにしてもよい。

【0012】

前記底変換部は、混合ガウスモデルに基づくベイズ識別におけるネイピア数の実数乗を2の実数乗に換算するようにしてもよい。

【0013】

本発明の第2の態様によれば、近似計算方法は、2の実数乗の指数を整数部分と0以上1未満の小数部分とに分離する指数分離ステップと、前記小数部分について2のべき乗を直線近似で計算する小数部分計算ステップと、前記整数部分についての2のべき乗の計算結果と、前記小数部分計算ステップでの前記小数部分についての2のべき乗の計算結果とを乗算する乗算ステップと、を含む。

【0014】

本発明の第3の態様によれば、プログラムは、プログラミング可能な装置に2の実数乗の指数を整数部分と0以上1未満の小数部分とに分離する指数分離ステップと、前記小数部分について2のべき乗を直線近似で計算する小数部分計算ステップと、前記整数部分についての2のべき乗の計算結果と、前記小数部分計算ステップでの前記小数部分についての2のべき乗の計算結果とを乗算する乗算ステップと、を実行させるためのプログラムで

10

20

30

40

50

ある。

【発明の効果】

【0015】

本発明によれば、べき乗の計算を比較的簡単な計算に置き換えることができる。

【図面の簡単な説明】

【0016】

【図1】本発明の実施形態に係る近似計算装置の機能構成を示す概略ブロック図である。

【図2】同実施形態における2の近似の例を示す図である。

【図3】同実施形態の近似GMMに基づく確率ニューラルネットワークの構成例を示す図である。

10

【図4】同実施形態に係る近似関数 $g_1(z)$ を用いた場合の、指数関数の近似誤差率を示す図である。

【図5】同実施形態に係る近似関数 $g'_1(z)$ を用いた場合の、指数関数の近似誤差率を示す図である。

【図6】同実施形態に係る近似関数 $g_1(z)$ 、 $g'_1(z)$ それぞれの近似誤差率を示す図である。

【図7】同実施形態に係る近似関数 $g'_1(z)$ による指数関数の近似の例を示す図である。

【図8】同実施形態に係る近似関数 $g_2(z)$ を用いた場合の、対数関数の近似誤差率を示す図である。

20

【図9】同実施形態に係る近似関数 $g_2((3)Oc)$ による対数関数の近似の例を示す図である。

【発明を実施するための形態】

【0017】

以下、本発明の実施形態を説明するが、以下の実施形態は請求の範囲にかかる発明を限定するものではない。また、実施形態の中で説明されている特徴の組み合わせの全てが発明の解決手段に必須であるとは限らない。

図1は、本発明の実施形態に係る近似計算装置の機能構成を示す概略ブロック図である。図1に示すように、近似計算装置10は、底変換部11と、指数分離部12と、整数部分計算部13と、小数部分計算部14と、乗算部15とを備える。

30

【0018】

近似計算装置10は、べき乗の近似計算を行う。

底変換部11は、ネイピア数( $e = 2.71828 \dots$ )の実数乗を2の実数乗に換算する。例えば、底変換部11は、混合ガウスモデル(Gaussian Mixture Model; GMM)に基づくベイズ識別におけるネイピア数の実数乗を2の実数乗に換算する。

但し、近似計算装置10が近似計算を行うべき乗は、底がネイピア数のものに限定されない。例えば、近似計算装置10が、底が10のべき乗の入力を受け、底変換部11が、底が10のべき乗を底が2のべき乗に換算するようにしてもよい。あるいは、近似計算装置10が、底が元々2であるべき乗の入力を受けて近似計算を行うようにしてもよい。

【0019】

40

指数分離部12は、2の実数乗の指数を整数部分と0以上1未満の小数部分とに分離する。

整数部分計算部13は、指数の整数部分について2のべき乗をシフト演算(ビットシフト)で計算する。

小数部分計算部14は、指数の小数部分について2のべき乗を直線近似で計算する。

小数部分計算部14が、0以上1未満の区間を分割した区間毎の直線近似にて、指数の小数部分についての2のべき乗を計算するようにしてもよい。さらには、小数部分計算部14が、0以上1未満の区間を分割した区間における2のべき乗の両端を結んだ線分にて直線近似を行うようにしてもよい。

【0020】

50

乗算部 15 は、整数部分計算部 13 による指数の整数部分についての 2 のべき乗の計算結果と、小数部分計算部 14 による小数部分についての 2 のべき乗の計算結果とを乗算する。

これにより、近似計算装置 10 は、べき乗の計算を四則演算およびシフト演算といった比較的簡単な計算に置き換えることができる。

近似計算装置 10 が行うべき乗の近似計算について、式を参照して説明する。

【0021】

混合ガウスモデルにおいて、特徴ベクトル  $x$  を  $D$  次元の実数ベクトルとすると、特徴ベクトル  $x$  の生起確率密度関数  $f(x)$  は、式 (1) のように示される。

【0022】

【数 1】

$$f(x) = \sum_{c=1}^C \sum_{m=1}^{M_c} \gamma_{c,m} M(x; \mu_{c,m}, \Sigma_{c,m}) \quad \dots (1)$$

【0023】

ここで、 $c$  ( $c$  は、 $1 \leq c \leq C$  の正整数) は、各クラスを識別する識別番号であり、 $C$  はクラス数を示す。 $m$  ( $m$  は、 $1 \leq m \leq M_c$  の正整数) は、クラス  $c$  を構成するコンポーネントを識別する識別番号であり、 $M_c$  は、クラス  $c$  を構成するコンポーネントの数を示す。混合ガウスモデルでは、正規分布 (ガウス分布) の重ね合わせによってクラスが構成されており、ここでいうコンポーネントは、クラスを構成する正規分布である。

また、 $\mu_{c,m}$ 、 $\Sigma_{c,m}$ 、 $\gamma_{c,m}$  ( $0 \leq \gamma_{c,m} \leq 1$ ) は、それぞれ、クラス  $c$  のコンポーネント  $m$  の平均ベクトル、分散共分散行列、他のコンポーネントとの混合比を示す。全てのクラス  $c$  の全てのコンポーネント  $m$  の混合比  $\gamma_{c,m}$  の合計は 1、すなわち、式 (2) を満たすものとする。

【0024】

【数 2】

$$\sum_{c=1}^C \sum_{m=1}^{M_c} \gamma_{c,m} = 1 \quad \dots (2)$$

【0025】

また、正規分布  $M(x; \mu_{c,m}, \Sigma_{c,m})$  は、式 (3) のように示される。

【0026】

【数 3】

$$M(x; \mu_{c,m}, \Sigma_{c,m}) = \frac{\exp(z_{c,m})}{\sqrt{(2\pi)^D |\Sigma_{c,m}|}} \quad \dots (3)$$

【0027】

ここで、指数関数  $\exp$  は、ネイピア数  $e$  のべき乗を示す。 $D$  は、上記のように特徴ベクトルの次元数を示す。

指数  $z_{c,m}$  は、式 (4) のように示される。

【0028】

【数 4】

$$z_{c,m} = -\frac{1}{2} (x - \mu_{c,m})^T \Sigma_{c,m}^{-1} (x - \mu_{c,m}) \quad \dots (4)$$

【0029】

ここでは、上付きの「T」は、行列またはベクトルの転置を示す。上付きの「-1」は逆行列を示す。

特徴ベクトル  $x$  をクラスに分類する際の、クラス  $c$  のコンポーネント  $m$  に対する事後確率は式(5)のように示される。

10

【0030】

【数 5】

$$p(c, m|x) = \frac{\mu_{c,m} M(x; \mu_{c,m}, \Sigma_{c,m})}{\sum_{c'=1}^C \sum_{m'=1}^{M_c} \gamma_{c',m'} M(x; \mu_{c',m'}, \Sigma_{c',m'})} \quad \dots (5)$$

【0031】

正規分布を指数関数を用いて表記すると、式(5)は式(6)のように示される。

【0032】

【数 6】

20

$$p(c, m|x) = \frac{\mu_{c,m} \exp\left\{-\frac{1}{2}(x - \mu_{c,m})^T \Sigma_{c,m}^{-1} (x - \mu_{c,m})\right\}}{\sum_{c'=1}^C \sum_{m'=1}^{M_c} \gamma_{c',m'} \exp\left\{-\frac{1}{2}(x - \mu_{c',m'})^T \Sigma_{c',m'}^{-1} (x - \mu_{c',m'})\right\}} \quad \dots (6)$$

【0033】

式(6)における指数関数  $\exp$  の定義域は  $[-, 0]$  と広く、これら指数関数の演算に時間を要することが考えられる。

そこで、指数関数を高速かつ高精度に演算できる近似式を構築する。特に、ハードウェア実装を考慮して近似式を構築する。

30

まず、指数関数  $\exp$  を 2 のべき乗で表現すると式(7)のように示される。

【0034】

【数 7】

$$\exp(z) = 2^{z \log_2 e} = 2^{zk} \quad \dots (7)$$

【0035】

ここで、 $k$  は式 8 のように示される定数である。

【0036】

【数 8】

40

$$k = \log_2 e \quad \dots (8)$$

【0037】

式(7)を用いた指数関数  $\exp$  から 2 のべき乗への換算は、底変換部 11 が行う処理の例に該当する。

式(7)の  $zk$  を整数部分と小数部分とに分けると、式(9)のように示される。

【0038】

【数 9】

$$\exp(z) = 2^{zk} = 2^\alpha 2^\beta \quad \dots (9)$$

【0039】

ここで、 $\alpha$  は  $zk$  の整数部分であり、式 (10) のように示される。

【0040】

【数 10】

$$\alpha = \text{floor}(zk) \quad \dots (10)$$

【0041】

関数  $\text{floor}$  は、整数への切り捨てを示す。具体的には関数  $\text{floor}$  は、引数を超えない最大の整数を示す。

式 (9) の  $\beta$  は  $zk$  の小数部分であり、式 (11) のように示される。

【0042】

【数 11】

$$\beta = zk - \alpha \quad \dots (11)$$

【0043】

の定義域は  $[0, 1)$ 、すなわち  $0 \leq \beta < 1$  である。指数  $zk$  を整数部分  $\alpha$  と小数部分  $\beta$  とに分離する処理は、指数分離部 12 の処理の例に該当する。

ここで、 $\beta$  の定義域を等間隔に  $n$  分割して、区間毎に  $2^\beta$  をテイラー展開に基づいて線形近似する。2 次以上の項を無視できると仮定すると、 $t$  の周りにおける  $2^\beta$  のテイラー展開は、式 (12) のように示される。

【0044】

【数 12】

$$2^\beta \approx 2^t + 2^t \ln 2 (\beta - t) \quad \dots (12)$$

【0045】

また、 $\beta$  の定義域  $[0, 1)$  を等間隔に  $n$  分割した各区間は、式 (13) のように示される。

【0046】

【数 13】

$$\left[ \frac{l-1}{n}, \frac{l}{n} \right) \quad l = 1, \dots, n \quad \dots (13)$$

【0047】

すなわち、 $(l-1)/n \leq \beta < l/n$  であり、 $l$  は、 $1 \leq l \leq n$  の正整数である。

各区間の中間点を  $t_{n,l}$  とすると、式 (14) のように示される。

【0048】

【数 14】

$$t_{n,l} = \frac{2l-1}{2n} \quad \dots (14)$$

【0049】

中間点  $t_{n,l}$  の周りにおける  $2^\beta$  のテイラー展開は、式 (15) のように示される。

【 0 0 5 0 】

【 数 1 5 】

$$2^\beta \doteq 2^{t_{n,l}} + 2^{t_{n,l}} \ln 2 (\beta - t_{n,l}) \quad \dots (15)$$

【 0 0 5 1 】

式 ( 1 5 ) は、式 ( 1 6 ) のように示すことができる。

【 0 0 5 2 】

【 数 1 6 】

$$2^\beta \doteq a_{n,l} \beta + b_{n,l} \quad \dots (16)$$

【 0 0 5 3 】

ここで、 $a_{n,l}$  は、式 ( 1 7 ) のように示される。

【 0 0 5 4 】

【 数 1 7 】

$$a_{n,l} = 2^{t_{n,l}} \ln 2 \quad \dots (17)$$

【 0 0 5 5 】

また、 $b_{n,l}$  は、式 ( 1 8 ) のように示される。式 ( 1 6 ) は の一次関数であり、  
を直線近似する。

【 0 0 5 6 】

【 数 1 8 】

$$b_{n,l} = 2^{t_{n,l}} (1 - t_{n,l} \ln 2) \quad \dots (18)$$

【 0 0 5 7 】

式 ( 1 6 ) による  $2$  の近似を用いて、 $\exp(z)$  を式 ( 1 9 ) のように近似できる

30

【 0 0 5 8 】

【 数 1 9 】

$$\exp(z) = 2^\alpha 2^\beta \doteq 2^\alpha (a_{n,l} \beta + b_{n,l}) \quad \dots (19)$$

【 0 0 5 9 】

さらに、係数  $a_{n,l}$  および  $b_{n,l}$  を簡単化する。式 ( 1 7 ) に式 ( 1 4 ) を代入すると、係数  $a_{n,l}$  は式 ( 2 0 ) のように示される。

【 0 0 6 0 】

【 数 2 0 】

$$a_{n,l} = 2^{\frac{2l-1}{2n}} \ln 2 \quad \dots (20)$$

【 0 0 6 1 】

また、式 ( 1 8 ) に式 ( 1 4 ) を代入すると、係数  $b_{n,l}$  は式 ( 2 1 ) のように示される。

【 0 0 6 2 】

40



【数 2 1】

$$b_{n,l} = 2^{\frac{2l-1}{2n}} \left( 1 - \frac{2l-1}{2n} \ln 2 \right) \cdots (21)$$

【0063】

式(20)および式(21)の「 $(2l-1)/2n$ 」( $=t_{n,l}$ )、および、2の自然対数「 $\ln 2$ 」を近似式で置き換える。

$n$ の値を大きくしていくと、 $(2l-1)/2n$ は、 $1/n$ に収束する。これは、式(22)のように示される。

【0064】

【数 2 2】

$$\lim_{n \rightarrow \infty} \frac{2l-1}{2n} = \frac{l}{n} \cdots (22)$$

【0065】

また、 $n$ の値を大きくしていくと、 $(1-2^{-1/n})n$ は、 $\ln 2$ に収束する。これは式(23)のように示される。

【0066】

【数 2 3】

$$\lim_{n \rightarrow \infty} \left( 1 - 2^{-\frac{1}{n}} \right) n = \ln 2 \cdots (23)$$

【0067】

式(22)に基づき、係数 $a_{n,l}$ および $b_{n,l}$ の $(2l-1)/2n$ を $1/n$ に置き換える。また、式(23)に基づき、係数 $a_{n,l}$ および $b_{n,l}$ の $\ln 2$ を $(1-2^{-1/n})n$ に置き換える。

係数 $a_{n,l}$ から置き換え後の係数 $a'_{n,l}$ は、式(24)のように示される。

【0068】

【数 2 4】

$$a'_{nl} = 2^{\frac{l}{n}} \left( 1 - 2^{-\frac{1}{n}} \right) n \cdots (24)$$

【0069】

係数 $b_{n,l}$ から置き換え後の係数 $b'_{n,l}$ は、式(25)のように示される。

【0070】

【数 2 5】

$$b'_{nl} = 2^{\frac{l}{n}} \left\{ 1 - l \left( 1 - 2^{-\frac{1}{n}} \right) \right\} \cdots (25)$$

【0071】

式(16)の係数 $a_{n,l}$ および $b_{n,l}$ を係数 $a'_{n,l}$ および $b'_{n,l}$ で置き換えて、2の近似は式(26)のように示される。

【0072】

【数 2 6】

$$2^\beta \doteq a'_{n,l} \beta + b'_{n,l} \cdots (26)$$

10

20

30

## 【 0 0 7 3 】

式 ( 2 6 ) は、 $2^x$  の示す曲線を  $n$  分割した各区間の曲線を、両端を結んだ直線で近似している。

図 2 は、式 ( 2 6 ) による  $2^x$  の近似の例を示す図である。図 2 は、 $2^x$  の定義域を 2 分割する場合の例を示している。図 2 のグラフの横軸は  $x$  の値を示し、縦軸は  $2^x$  の値を示す。

## 【 0 0 7 4 】

線  $L_{11}$  は、 $2^x$  を示す。点  $P_{11}$ 、 $P_{12}$ 、 $P_{13}$  は、それぞれ、 $2^x$  の定義域を分割した区間の端部と線  $L_{11}$  との交点を示す。点  $P_{11}$  は、 $x = 0$  の直線と線  $L_{11}$  との交点を示す。点  $P_{12}$  は、 $x = 0.5$  の直線と線  $L_{11}$  との交点を示す。点  $P_{13}$  は、 $x = 1$  の直線と線  $L_{11}$  との交点を示す。 10

線  $L_{12a}$ 、線  $L_{12b}$  は、それぞれ  $x$  の区間  $[0, 0.5)$ 、 $[0.5, 1)$  にて線  $L_{11}$  を近似する。線  $L_{12a}$  は、点  $P_{11}$  と点  $P_{12}$  とを結んだ直線 (線分) である。線  $L_{12b}$  は、点  $P_{12}$  と点  $P_{13}$  とを結んだ直線 (線分) である。

このように、式 ( 2 6 ) によれば、直感的に把握しやすい態様で  $2^x$  を近似することができる。

## 【 0 0 7 5 】

式 ( 2 6 ) による近似を指数関数  $e^{x \cdot p}$  の近似に適用する。具体的には、式 ( 1 9 ) の係数  $a_{n,l}$  および  $b_{n,l}$  を係数  $a'_{n,l}$  および  $b'_{n,l}$  で置き換えて、指数関数  $e^{x \cdot p}$  の近似は式 ( 2 7 ) のように示される。 20

## 【 0 0 7 6 】

## 【 数 2 7 】

$$\exp(z) = 2^\alpha 2^\beta \cong 2^\alpha (a'_{n,l} \beta + b'_{n,l}) \quad \dots (27)$$

## 【 0 0 7 7 】

式 ( 2 7 ) における  $2^\alpha$  は、二進数表現におけるシフト演算 (ビットシフト) で演算できる。また、 $a'_{n,l} \beta + b'_{n,l}$  は、 $\beta$  の一次関数である。このように、式 ( 2 7 ) によれば、指数関数  $e^{x \cdot p}$  をシフト演算と線形方程式で近似することができる。簡単な処理で演算を行える点で、式 ( 2 7 ) によれば、指数関数  $e^{x \cdot p}$  の演算を高速化することができる。特に式 ( 2 7 ) は、シフト演算を用いる点、および、一次関数といった簡単な関数を用いる点でハードウェア実装が容易である。式 ( 2 7 ) をハードウェア実装することで、指数関数  $e^{x \cdot p}$  の演算を高速化することができる。 30

式 ( 2 7 ) の「 $2^\alpha$ 」の計算は、整数部分計算部 1 3 の処理の例に該当する。式 ( 2 7 ) の「 $a'_{n,l} \beta + b'_{n,l}$ 」の計算は、小数部分計算部 1 4 の処理の例に該当する。これらの計算結果の乗算は、乗算部 1 5 の処理の例に該当する。

## 【 0 0 7 8 】

次に、指数関数  $e^{x \cdot p}$  の近似を混合ガウスモデルに適用する。

正規分布を示す式 ( 3 ) に式 ( 2 7 ) の近似を導入して、正規分布は式 ( 2 8 ) のように近似される。 40

## 【 0 0 7 9 】

## 【 数 2 8 】

$$M_{app}(x; \mu_{c,m}, \Sigma_{c,m}) = A_n 2^{\alpha_{c,m}} (a'_{n,l} \beta_{c,m} + b'_{n,l}) \quad \dots (28)$$

## 【 0 0 8 0 】

ここで、 $A_n$  は実数定数であり、式 ( 2 9 ) のように示される。

## 【 0 0 8 1 】

【数 2 9】

$$A_n = \frac{1}{\int_{-\infty}^{\infty} \{2^\alpha (a'_{n,l} \beta + b'_{n,l})\} dx} \quad \dots (29)$$

【0082】

式(28)に示す正規分布の近似  $M_{app}(x; \mu_{c,m}, \Sigma_{c,m})$  を近似ガウス分布(または近似正規分布)と称する。

近似ガウス分布  $M_{app}(x; \mu_{c,m}, \Sigma_{c,m})$  について、式(30)が成り立つ。

10

【0083】

【数 3 0】

$$\int_{-\infty}^{\infty} M_{app}(x; \mu_{c,m}, \Sigma_{c,m}) dx = 1 \quad \dots (30)$$

【0084】

また、近似ガウス分布  $M_{app}(x; \mu_{c,m}, \Sigma_{c,m})$  について、式(31)が成り立つ。

【0085】

【数 3 1】

20

$$M_{app}(x; \mu_{c,m}, \Sigma_{c,m}) \geq 0 \quad \dots (31)$$

【0086】

式(30)および式(31)より、近似ガウス分布  $M_{app}(x; \mu_{c,m}, \Sigma_{c,m})$  は確率密度関数である。

近似ガウス分布  $M_{app}(x; \mu_{c,m}, \Sigma_{c,m})$  を式(1)に適用すると式(32)のようになる。

【0087】

【数 3 2】

30

$$f_{app}(x) = \sum_{c=1}^C \sum_{m=1}^{M_c} \gamma_{c,m} M_{app}(x; \mu_{c,m}, \Sigma_{c,m}) \quad \dots (32)$$

【0088】

$f_{app}(x)$  は、混合ガウスモデルの生起確率密度関数  $f(x)$  を模擬する。 $f_{app}(x)$  を用いての混合ガウスモデルの模擬を、近似混合ガウスモデル(近似GMM)と称する。

式(26)の近似を式(5)に適用して、近似GMMに基づいたベイズ推定におけるクラス  $c$  のコンポーネント  $m$  に対する事後確率  $p_{app}(c, m | x)$  は、式(33)のよう示される。

40

【0089】

【数 3 3】

$$p_{app}(c, m | x) = \frac{\mu_{c,m} M(x; \mu_{c,m}, \Sigma_{c,m})}{\sum_{c'=1}^C \sum_{m'=1}^{M_{c'}} \gamma_{c',m'} 2^{\alpha_{c',m'}} (a'_{n,l} \beta_{c',m'} + b'_{n,l})} \quad \dots (33)$$

【 0 0 9 0 】

、 のいずれもクラス  $c$  およびコンポーネント  $m$  毎に値をとるため、添え字  $c$  および  $m$ 、あるいは、 $c'$  および  $m'$  を付している。

このように、近似 GMM によれば、指数関数を用いずに GMM に基づくベイズ推定を近似可能であり、この点で演算を高速化することができる。

【 0 0 9 1 】

次に、近似 GMM のニューラルネットワークへの適用について説明する。以下、近似 GMM を適用したニューラルネットワークを、近似 GMM に基づく確率ニューラルネットワークと称する。

式 (5) を、平均ベクトル  $\mu_{c,m}$  および共分散逆行列  $\Sigma_{c,m}^{-1}$  を用いて展開すると、式 (34) のように、新たな入力ベクトル  $X$  と、 $\Sigma_{c,m}^{-1}$  とを用いて指数関数の引数を見かけ上線形化できる。

【 0 0 9 2 】

【 数 3 4 】

$$p(c, m|x) = \frac{\exp(\eta^{(c,m)T} X)}{\sum_{c'=1}^C \sum_{m'=1}^{M_c} \exp(\eta^{(c',m')T} X)} \quad \dots (34)$$

【 0 0 9 3 】

ここで、新たな入力ベクトル  $X$  は、式 (35) のように示される。

【 0 0 9 4 】

【 数 3 5 】

$$X = [1, x^T, x_1^2, x_1 x_2, \dots, x_1 x_D, x_2^2, x_2 x_3, \dots, x_2 x_D, \dots, x_D^2]^T \quad \dots (35)$$

【 0 0 9 5 】

すなわち、新たな入力ベクトル  $X$  は、元の入力ベクトル  $x$  の要素  $x_1, \dots, x_D$  のうち 2 つ以下の要素の積を要素としている。

また、 $\Sigma_{c,m}^{-1}$  は、式 (36) のように示される。

【 0 0 9 6 】

【 数 3 6 】

$$\eta^{(c,m)} = \left[ \eta_0^{(c,m)}, \sum_{i=1}^D s_{i,1}^{(c,m)} \mu_i^{(c,m)}, \dots, \sum_{i=1}^D s_{i,D}^{(c,m)} \mu_i^{(c,m)}, \right. \\ \left. -\frac{1}{2} s_{1,1}^{(c,m)}, s_{1,2}^{(c,m)}, \dots, s_{1,D}^{(c,m)}, \dots, \right. \\ \left. -\frac{1}{2} (2 - \delta_{ij}) s_{i,j}^{(c,m)}, \dots, -s_{D,1}^{(c,m)}, \dots, -\frac{1}{2} s_{D,D}^{(c,m)} \right]^T \quad \dots (36)$$

【 0 0 9 7 】

ここで、 $s_{i,j}^{(c,m)}$  は、共分散逆行列  $\Sigma_{c,m}^{-1}$  の要素であり、式 (37) のように示される。

【 0 0 9 8 】

10

20

30

【数 3 7】

$$\Sigma_{c,m}^{-1} = [S_{i,j}^{(c,m)}] \quad \dots (37)$$

【0 0 9 9】

また、 $\eta_0^{(c,m)}$  は、式 (38) のように示される。

【0 1 0 0】

【数 3 8】

$$\begin{aligned} \eta_0^{(c,m)} = & -\frac{1}{2} \sum_{i=1}^D \sum_{j=1}^D S_{i,j}^{(c,m)} \mu_i^{(c,m)} \mu_j^{(c,m)} \\ & - \frac{D}{2} \log 2\pi - \frac{1}{2} \log |\Sigma_{c,m}| + \log \gamma_{c,m} \quad \dots (38) \end{aligned}$$

【0 1 0 1】

$\delta_{ij}$  はクロネッカーのデルタ (Kronecker Delta) であり、式 (39) のように示される。

【0 1 0 2】

【数 3 9】

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \quad \dots (39)$$

【0 1 0 3】

式 (34) に式 (24)、式 (25) および式 (27) を適用すると、式 (40) のように示される。

【0 1 0 4】

【数 4 0】

$$p(c, m | x) = \frac{2^{\alpha_{c,m}} (a'_{n,l} \beta_{c,m} + b'_{n,l})}{\sum_{c'=1}^C \sum_{m'=1}^{M_c} 2^{\alpha_{c',m'}} (a'_{n,l} \beta_{c',m'} + b'_{n,l})} \quad \dots (40)$$

【0 1 0 5】

式 (40) の  $\alpha_{c,m}$  は式 (41) のように示すことができる。

【0 1 0 6】

【数 4 1】

$$\alpha_{c,m} = \text{floor}(w^{(c,m)T} X) \quad \dots (41)$$

【0 1 0 7】

式 (40) の  $\beta_{c,m}$  は式 (42) のように示すことができる。

【0 1 0 8】

【数 4 2】

$$\beta_{c,m} = w^{(c,m)T} X \quad \dots (42)$$

【0 1 0 9】

式 (6) に示されるように  $k = \log_2 e$  であり、この  $k$  を用いて新たなパラメータの

ベクトル  $w^{(c,m)}$  を式 (43) のようにおくことができる。

【0110】

【数43】

$$w^{(c,m)} = k\eta^{(c,m)} \quad \dots (43)$$

【0111】

式 (40) を式 (34) と比較すると、GMM のパラメータ  $\mu_{c,m}$ 、 $\sigma_{c,m}$  および  $\sigma_{c,m}$  を、無制約のベクトル  $w^{(c,m)}$  で表現可能である。

ベクトル  $w^{(c,m)}$  の値を適切に獲得することができれば、入力ベクトル  $x$  を非線形変換した新たな入力ベクトル  $X$  とベクトル  $w^{(c,m)}$  との積を求め、係数を用いた線形方程式およびビットシフトにてベイズ推定の演算を実行できる。そこで、ベクトル  $w^{(c,m)}$  をニューラルネットワークの重みとして、 $w^{(c,m)}$  の値を機械学習によって決定する。

10

【0112】

図3は、近似GMMに基づく確率ニューラルネットワークの構成例を示す図である。図3に示す構成で、ニューラルネットワーク20は、非線形変換部21と、第1層22と、第2層23と、第3層24とを備える。

非線形変換部21は、式(35)に示されるように、入力ベクトル  $x = [x_1, x_2, \dots, x_D]^T$  を新たな入力ベクトル  $X$  に変換する。ここで、新たな入力ベクトル  $X$  の要素数を  $H$  とし、 $X = [X_1, X_2, \dots, X_H]^T$  と表記する。 $H = 1 + (D + 3) / 2$  である。

20

【0113】

第1層22は、新たな入力ベクトル  $X$  の要素数  $H$  と同数のノード  $22-1 \sim 22-H$  を備える。第1層のノード  $22-h$  ( $h$  は、 $1 \leq h \leq H$  の整数) は、入力される  $X_h$  を、第2層23の各ノードへ出力する。ノード  $22-h$  への入力を  $(1)I_h$  とし、ノード  $22-h$  からの出力を  $(1)O_h$  とすると、式(44)のように示される。

【0114】

【数44】

$$(1)O_h = (1)I_h = X_h \quad \dots (44)$$

【0115】

第2層23は、近似GMMにおける全クラス的全コンポーネントと同数のノード  $23-1, 1, \dots, 23-1, M1, \dots, 23-c, m, \dots, 23-C, M_c$  を備える。

第1層22の全ノードから第2層23の全ノードへ経路(Path)が設けられ、経路毎に重みが設定される。第1層22のノードからの出力に重みが乗算されて第2層23のノードに入力され、第2層のノードは、第1層からの出力の重み付け合計の入力を受ける。第1層22のノード  $22-h$  から第2層23のノード  $23-c, m$  への経路の重みを  $w_h^{(c,m)}$  とすると、ノード  $23-c, m$  への入力  $(2)I_{c,m}$  は、式(45)のように示される。

40

【0116】

【数45】

$$(2)I_{c,m} = \sum_{h=1}^H (1)O_h w_h^{c,m} = \sum_{h=1}^H X_h w_h^{c,m} \quad \dots (45)$$

【0117】

50

第2層23の各ノードは、式(40)の演算を行って、クラス毎かつコンポーネント毎の事後確率を算出し出力する。ノード23-c, mからの出力<sup>(2)</sup>O<sub>c,m</sub>は、式(46)のように示される。

【0118】

【数46】

$$^{(2)}O_{c,m} = \frac{2^{\alpha_{c,m}}(a'_{n,l}\beta_{c,m}+b'_{n,l})}{\sum_{c'=1}^C \sum_{m'=1}^{M_c} 2^{\alpha_{c',m'}}(a'_{n,l}\beta_{c',m'}+b'_{n,l})} \dots (46)$$

【0119】

式(46)に示されるように、第2層23の各ノードが行う計算は、近似計算装置10が行うべき乗の近似計算を含む。この点で、第2層23の各ノードは近似計算装置10の例に該当する。

第3層24は、クラスの数Cと同数のノード24-1~24-Cを備える。

第2層23の各ノードは、自らが対応するクラスに応じて、第3層24の各ノードのうち当該クラスに対応するノードヘデータ(値)を出力する。第2層のノード23-c, mは、第3層24のノード24-cへ出力を行う。第3層24の各ノードは、第2層のノードからの出力の合計の入力を受け、当該合計を出力する。

第3層24のノード24-cへの入力<sup>(3)</sup>I<sub>c</sub>とし、ノード24-cからの出力<sup>(3)</sup>O<sub>c</sub>とすると、式(47)のように示される。

20

【0120】

【数47】

$$^{(3)}I_c = ^{(3)}O_c = \sum_{m=1}^{M_c} ^{(2)}O_{c,m} \dots (47)$$

【0121】

次に、ニューラルネットワーク20における学習則について説明する。ニューラルネットワーク20は、重みw<sub>h<sup>c,m</sup></sub>の値を学習する。

30

学習用データとしてN個の入力ベクトルx<sup>(n)</sup>(nは、1 n Nの正整数)と、式(48)に示される教師信号ベクトルT<sup>(n)</sup>とが与えられた場合を考える。

【0122】

【数48】

$$T^{(n)} = [T_1^{(n)}, \dots, T_c^{(n)}, \dots, T_C^{(n)}], \quad T_c^{(n)} \in [0,1],$$

$$\sum_{c=1}^C T_c^{(n)} = 1 \quad \dots (48)$$

【0123】

カルバック・ライブラー情報量(Kullback-Leibler Divergence)を最小化することでGMMのパラメータを学習的に獲得できるとの前提に基づき、ニューラルネットワーク20の評価関数を式(49)のように定義する。

【0124】

【数 4 9】

$$J = \log 2 \sum_{n=1}^N J_n = \log 2 \sum_{n=1}^N \sum_{c=1}^C T_c^{(n)} \log_2 {}^{(3)}O_c \quad \dots (49)$$

【0 1 2 5】

ニューラルネットワーク 20 は、式 (49) の J を小さくするように学習を行う。

重み  $w_h^{c,m}$  の修正量を  $\Delta w_h^{c,m}$  とし、学習率を  $\mu$  とすると、一括学習法によって修正量  $\Delta w_h^{c,m}$  は、式 (50) のように示される。

10

【0 1 2 6】

【数 5 0】

$$\Delta w_h^{c,m} = -\mu \sum_{n=1}^N \frac{\partial J_n}{\partial w_h^{c,m}} \quad \dots (50)$$

【0 1 2 7】

式 (50) の  $(\partial J_n) / (\partial w_h^{c,m})$  は、式 (51) のように表される。

【0 1 2 8】

【数 5 1】

20

$$\frac{\partial J_n}{\partial w_h^{c,m}} = \left( {}^{(2)}O_{c,m} - \frac{{}^{(2)}O_{c,m}}{{}^{(3)}O_c} T_k^{(n)} \right) X_h^{(n)} \quad \dots (51)$$

【0 1 2 9】

ここで、 $X_h^{(n)}$  は、入力ベクトル  $x^{(n)}$  から得られる新たな入力ベクトルである。

式 (50) および式 (51) は、単純な四則演算で計算可能であり、この点で高速に演算可能である。

30

また、式 (49) の  $\log_2 {}^{(3)}O_c$  を高速に演算できるよう近似を行う。 $\log_2 {}^{(3)}O_c$  は、式 (52) のように変形できる。

【0 1 3 0】

【数 5 2】

$$\log_2 {}^{(3)}O_c = \log_2 (m_c \times 2^{n_c}) = \log_2 m_c + n_c \quad \dots (52)$$

【0 1 3 1】

ここで、 $m_c$ 、 $n_c$  はいずれも整数であり、 $1 < m_c < 2$  である。

40

指数関数  $e^x$  の近似の場合と同様、 $m_c$  の定義域  $[1, 2)$  を  $n$  分割し、 $\log_2 m_c$  を近似する。 $m_c$  の定義域を等間隔に  $n$  分割した区間は、式 (53) のように示される。

【0 1 3 2】

【数 5 3】

$$\left[ \frac{n+l-1}{n}, \frac{n+l}{n} \right) \quad l = 1, \dots, n \quad \dots (53)$$



【 0 1 3 3 】

この区間における  $\log_2 m_c$  上の両端の点は、式 (54) のように示される。

【 0 1 3 4 】

【数 5 4】

$$\left(\frac{n+l-1}{n}, \log_2 \frac{n+l-1}{n}\right), \left(\frac{n+l}{n}, \log_2 \frac{n+l}{n}\right) \quad \dots (54)$$

【 0 1 3 5 】

式 (54) に示される両端の点を結ぶ直線 (線分)  $g'_{n,l}$  によって、 $\log_2 m_c$  は式 (55) のように近似される。

10

【 0 1 3 6 】

【数 5 5】

$$\log_2 m_c \cong c_{n,l} m_c + d_{n,l} \quad \dots (55)$$

【 0 1 3 7 】

ここで、 $c_{n,l}$  は、式 (56) のように示される。

【 0 1 3 8 】

【数 5 6】

20

$$c_{n,l} = n \log_2 \frac{n+l}{n+l-1} \quad \dots (56)$$

【 0 1 3 9 】

 $d_{n,l}$  は、式 (57) のように示される。

【 0 1 4 0 】

【数 5 7】

$$d_{n,l} = \log_2 \frac{(n+l-1)(n+l)}{n(n+l)(n+l-1)} \quad \dots (57)$$

【 0 1 4 1 】

さらに、 $n = 1$ 、 $l = 1$  とした場合、 $c_{n,l} = 1$ 、 $d_{n,l} = -1$  となり、 $\log_2 m_c$  を式 (58) のように線形近似可能である。

【 0 1 4 2 】

【数 5 8】

$$\log_2 m_c \cong m_c - 1 \quad \dots (58)$$

【 0 1 4 3 】

40

この場合、式 (52) は、式 (59) のように示される。

【 0 1 4 4 】

【数 5 9】

$$\log_2 {}^{(3)}O_c \cong m_c + n_c - 1 \quad \dots (59)$$

【 0 1 4 5 】

また、式 (49) は、式 (60) のように示される。

【 0 1 4 6 】

50

【数 6 0】

$$J \doteq \log 2 \sum_{n=1}^N \sum_{c=1}^C T_c^{(n)} (m_c + n_c - 1) \quad \dots (60)$$

【0 1 4 7】

式(50)の学習率 $\mu$ は、式(61)に基づいて学習サイクル毎に更新することが考えられる。

【0 1 4 8】

10

【数 6 1】

$$\mu = \frac{\mu_k J^\lambda}{\sum_{h=1}^H \sum_{c=1}^C \sum_{m=1}^{M_c} \left( \sum_{n=1}^N \frac{\partial J_n}{\partial w_{h,c,m}} \right)} \quad \dots (61)$$

【0 1 4 9】

ここで、学習の収束時間の上限設定値を $t_f$ とすると、 $\mu_k$ は式(62)のように示される。

【0 1 5 0】

20

【数 6 2】

$$\mu_k = \frac{J_0^{1-\lambda}}{t_f(1-\lambda)} \quad \dots (62)$$

【0 1 5 1】

は $0 < \lambda < 1$ を満たす任意の定数である。ここで、 $\lambda = 0.5$ とすると、 $\mu_k$ は式(63)のように示される。

【0 1 5 2】

【数 6 3】

30

$$\mu_k = \frac{2\sqrt{J_0}}{t_f} \quad \dots (63)$$

【0 1 5 3】

関数 $h(J) = \sqrt{J}$ を $t$ の周りでテイラー展開して第2項まで線形近似を行うと、式64のように示される。

【0 1 5 4】

【数 6 4】

$$h(J) = \sqrt{J} \doteq \sqrt{t} + \frac{1}{2} t^{-\frac{1}{2}} (J - t) \quad \dots (64)$$

【0 1 5 5】

ここで、 $t = 1$ とすると式(65)のように示される。

【0 1 5 6】

【数 6 5】

$$\sqrt{J} \doteq \frac{J+1}{2} \quad \dots (65)$$

【 0 1 5 7 】

式 ( 6 5 ) は、 $J = 1$  における接線であり、 $J = 1$  の近傍でのみ近似が成り立つ。近似の精度をより高めるために、式 ( 6 5 ) を式 ( 6 6 ) のように変形する。

【 0 1 5 8 】

【 数 6 6 】

$$\sqrt{J} = \sqrt{J'} \times 2^q \doteq \frac{J'+1}{2} \times 2^q \quad \dots (66)$$

【 0 1 5 9 】

ここで、 $0.5 \leq J' \leq 2$  である。また、 $q$  は自然数のベクトルである。

このように、比較的簡単な計算にて学習を実現可能である。

次に、上述した近似の精度について検討する。

式 ( 1 9 ) に示される、指数関数の近似関数  $g_1(z) = 2^{(a_{n,1} + b_{n,1})}$  の、指数関数  $e^x p = f_1(z)$  に対する誤差率  $E_1$  を式 ( 6 7 ) のように定義する。

【 0 1 6 0 】

【 数 6 7 】

$$E_1 = \left| \frac{f_1(z) - g_1(z)}{f_1(z)} \right| \quad \dots (67)$$

【 0 1 6 1 】

また、式 ( 1 9 ) に示される、指数関数  $e^x p$  の近似関数  $g'_1(z) = 2^{(a_{n,1} + b_{n,1})}$  の、指数関数  $e^x p = f_1(z)$  に対する誤差率  $E'_1$  を式 ( 6 8 ) のように定義する。

【 0 1 6 2 】

【 数 6 8 】

$$E'_1 = \left| \frac{f_1(z) - g'_1(z)}{f_1(z)} \right| \quad \dots (68)$$

【 0 1 6 3 】

図 4 は、近似関数  $g_1(z)$  を用いた場合の、指数関数の近似誤差率を示す図である。

図 4 のグラフの横軸は  $z$  の値を示し、縦軸は近似誤差率を示す。

線 L 2 1 は、分割数  $n = 1$  の場合の近似誤差率を示す。線 L 2 2 は、分割数  $n = 2$  の場合の近似誤差率を示す。線 L 2 3 は、分割数  $n = 4$  の場合の近似誤差率を示す。

図 4 に示すように、分割数  $n$  が増加すると近似誤差率は小さくなる。近似誤差率が最も大きい分割数  $n = 1$  の場合でも、近似誤差率の最大値は 7.59% であり、比較的高精度に近似できている。

【 0 1 6 4 】

図 5 は、近似関数  $g'_1(z)$  を用いた場合の、指数関数の近似誤差率を示す図である。図 5 のグラフの横軸は  $z$  の値を示し、縦軸は近似誤差率を示す。

線 L 3 1 は、分割数  $n = 1$  の場合の近似誤差率を示す。線 L 3 2 は、分割数  $n = 2$  の場合の近似誤差率を示す。線 L 3 3 は、分割数  $n = 4$  の場合の近似誤差率を示す。

図 5 に示すように、分割数  $n$  が増加すると近似誤差率は小さくなる。近似誤差率が最も大きい分割数  $n = 1$  の場合でも、近似誤差率の最大値は 6.15% であり、比較的高精度に近似できている。

【 0 1 6 5 】

図 6 は、近似関数  $g_1(z)$ 、 $g'_1(z)$  それぞれの近似誤差率を示す図である。図

6では、近似関数  $g_1(z)$ 、 $g'_1(z)$  それぞれについて、分割数  $n = 1, 2, 4$  それぞれの場合の近似誤差率の平均値および最大値を示している。図6を参照すると、近似関数  $g'_1(z)$  では、近似関数  $g_1(z)$  よりも近似誤差率の平均値が大きくなっているものの、最大値は小さくなっている。もっとも、近似関数  $g_1(z)$  と  $g'_1(z)$  との近似誤差率の相違は1%程度と比較的小さい値である。

【0166】

図7は、近似関数  $g'_1(z)$  による指数関数の近似の例を示す図である。図7は、分割数  $n = 1$  の場合の例を示している。図7のグラフの横軸は  $z$  を示し、縦軸は  $\exp(z)$  を示す。

線L41は、指数関数  $\exp(z)$  の実際値を示す。線L42は、近似関数  $g'_1(z)$  による近似値を示す。

10

図7に示されるように、近似関数  $g'_1(z)$  によって指数関数  $\exp(z)$  を近似できている。

【0167】

次に、対数関数の近似精度について検討する。

式(59)に示される対数関数の近似関数  $g_2({}^{(3)}O_c) = m_c + n_c - 1$  の、対数関数  $f_2({}^{(3)}O_c) = \log_2({}^{(3)}O_c)$  に対する近似誤差率を式(69)のように定義する。

【0168】

【数69】

20

$$E_2 = \left| \frac{f_1({}^{(3)}O_c) - g'_1({}^{(3)}O_c)}{f_2({}^{(3)}O_c)} \right| \dots (69)$$

【0169】

図8は、近似関数  $g_2(z)$  を用いた場合の、対数関数の近似誤差率を示す図である。図8のグラフの横軸は  ${}^{(3)}O_c$  の値を示し、縦軸は近似誤差率を示す。

線L51は、分割数  $n = 1$  の場合の近似誤差率を示す。線L52は、分割数  $n = 2$  の場合の近似誤差率を示す。線L53は、分割数  $n = 4$  の場合の近似誤差率を示す。

30

指数関数の近似の場合と同様、対数関数の近似の場合も分割数  $n$  が増加すると近似誤差率は小さくなる。

${}^{(3)}O_c$  が1に近付くと  $\log_2({}^{(3)}O_c)$  が0に近付くため、近似誤差率は次第に大きくなる。見かけ上の近似誤差率が大きくなるものの、近似誤差  $f_2({}^{(3)}O_c) - g_2({}^{(3)}O_c)$  が最も大きい分割数  $n = 1$  の場合でも、近似誤差は平均で  $0.058 \pm 0.026$  であり、比較的高精度に近似できている。

【0170】

図9は、近似関数  $g_2({}^{(3)}O_c)$  による対数関数の近似の例を示す図である。図9は、分割数  $n = 1$  の場合の例を示している。図9のグラフの横軸は  ${}^{(3)}O_c$  を示し、縦軸は  $\log_2({}^{(3)}O_c)$  を示す。

40

線L61は、対数関数  $\log_2({}^{(3)}O_c)$  の実際値を示す。線L62は、近似関数  $g_2({}^{(3)}O_c)$  による近似値を示す。

図9に示されるように、近似関数  $g_2({}^{(3)}O_c)$  によって対数関数  $\exp(\log_2({}^{(3)}O_c))$  を近似できている。

【0171】

次に、近似混合ガウスモデルに基づくクラス識別の実験結果について説明する。図3に示すニューラルネットワークをFPGA(Field-programmable Gate Array)に実装してクラス識別および学習の実験を行った。

ハードウェア上では、以下の流れで指数関数の演算を行った。

1 - 1.  ${}^{(2)}I_{c,m}$  の整数値を  $c_{i,m}$  とし、小数値を  $c_{f,m}$  とする。

50

1 - 2 .  $c, m$ に1を加え、 $c, m$ ビット左シフトする。

【0172】

また、対数関数の演算は、以下のように行った。

2 - 1 .  $(^3)O_c = 1$ のとき2 - 2 .へ進む。 $(^3)O_c > 1$ のとき2 - 3 .へ進む。

2 - 2 .  $(^3)O_c$ を2倍(1ビット左シフト)し、2 - 1 .へ戻る。

2 - 3 . 2 - 2 .で2倍した回数(1ビット左シフトした回数)を $n_c$ とし、式(59)の演算を行う。

【0173】

平方根関数は、以下のように計算可能である。

3 - 1 .  $J > 2$ のとき3 - 2 .に進む。 $J = 2$ のとき、3 - 3 .に進む。

3 - 2 .  $J$ を1/4倍(2ビット右シフト)し、3 - 1 .へ戻る。

3 - 3 . 3 - 2 .で1/4倍した回数(2ビット右シフトした回数)を $q$ として、式(66)の演算を行う。

以上より、ハードウェア上で指数関数、対数関数、平方根関数の各演算をシフト演算および加減算にて実行可能である。

【0174】

実験の比較対象として、上記の近似を行わない混合ガウスモデルによるニューラルネットワークをソフトウェアで構築した。ソフトウェアによるニューラルネットワークの構築は、CPUクロック周波数3.40ギガヘルツ(GHz)、RAM(Random Access Memory)16.0ギガバイト(GB)のコンピュータおよびC言語を用いて行った。

以下では、近似混合ガウスモデルによるニューラルネットワークのFPGAを用いた構成を近似GMMハード構成と称する。また、近似を行わない混合ガウスモデルによるニューラルネットワークのソフトウェアによる構成をGMMソフト構成と称する。

【0175】

クラスの識別率については、GMMソフト構成、近似GMMハード構成共に約99と高い識別率を示し、有意差は認められなかった。

学習時間については、GMMソフト構成が6.43秒であったのに対し、近似GMMハード構成では8.47ミリ秒と、およそ800分の1の時間に短縮された。

上記のように、指数関数、対数関数、平方根関数の各演算をシフト演算および加減算にて実行可能であり、この点で、ハードウェア実装が容易である。ハードウェア実装および計算の単純化により、上記のように学習時間が大幅に短縮された。

【0176】

以上のように、指数分離部12は、2の実数乗の指数を整数部分と0以上1未満の小数部分とに分離する。小数部分計算部14は、指数の小数部分について2のべき乗を直線近似で計算する。乗算部15は、整数部分についての2のべき乗の計算結果と、小数部分計算部14による小数部分についての2のべき乗の計算結果とを乗算する。

これにより、近似計算装置10では、指数の小数部分についての2のべき乗を簡単な線形演算に置き換えることができる。この点で、近似計算装置10によれば、ハードウェア実装が容易であり、ハードウェア実装および演算の単純化による演算時間の短縮が見込まれる。

【0177】

また、整数部分計算部13は、指数の整数部分について2のべき乗をシフト演算で計算する。

近似計算装置10によれば、指数の整数部分について2のべき乗をシフト演算で行う点で、ハードウェア実装が容易であり、ハードウェア実装および演算の単純化による演算時間の短縮が見込まれる。

【0178】

また、小数部分計算部14は、0以上1未満の区間を分割した区間毎の直線近似にて指数の小数部分についての2のべき乗を計算する。

10

20

30

40

50

これにより、近似計算装置 10 では、特に高い演算精度を要求される場合は分割数を多くして近似の精度を高めることができる。かつ、区間毎の直線近似にて、指数の小数部分についての 2 のべき乗を簡単な線形演算に置き換えることができる。この点で、近似計算装置 10 によれば、ハードウェア実装が容易であり、ハードウェア実装および演算の簡単化による演算時間の短縮が見込まれる。

【0179】

また、小数部分計算部 14 は、0 以上 1 未満の区間を分割した区間における 2 のべき乗の両端を結んだ線分にて直線近似を行う。

これにより、近似計算装置 10 では、指数の小数部分についての 2 のべき乗を、より簡単な線形演算に置き換えることができる。この点で、近似計算装置 10 によれば、ハードウェア実装が容易であり、ハードウェア実装および演算の簡単化による演算時間の短縮が見込まれる。

10

【0180】

また、底変換部 11 は、ネイピア数の実数乗を 2 の実数乗に換算する。

これにより、近似計算装置 10 では、ネイピア数を底とする指数関数を、より簡単な演算に置き換えることができる。

【0181】

また、底変換部 11 は、混合ガウスモデルに基づくベイズ識別におけるネイピア数の実数乗を 2 の実数乗に換算する。

近似計算装置 10 によれば、混合ガウスモデルを近似する近似混合ガウスモデルを用いることができる。近似混合ガウスモデルは、混合ガウスモデルと比較して、より簡単な演算で実装可能である。この点で、近似計算装置 10 によれば近似混合ガウスモデルのハードウェア実装が容易であり、ハードウェア実装および演算の簡単化による演算時間の短縮が見込まれる。

20

【0182】

なお、近似計算装置 10 が行う演算及び制御の全部または一部の機能を実現するためのプログラムを、コンピュータまたはハードウェアが読み取り可能な記録媒体に記録して、この記録媒体に記録されたプログラムをコンピュータシステムまたはハードウェアに読み込ませ、実行することで各部の処理を行ってもよい。なお、ここでいう「コンピュータシステム」とは、OS や周辺機器等のハードウェアを含むものとする。

30

また、「コンピュータシステム」は、WWWシステムを利用している場合であれば、ホームページ提供環境（あるいは表示環境）も含むものとする。

また、「コンピュータ読み取り可能な記録媒体」とは、フレキシブルディスク、光磁気ディスク、ROM、CD-ROM等の可搬媒体、コンピュータシステムに内蔵されるハードディスク等の記憶装置のことをいう。また上記プログラムは、前述した機能の一部を実現するためのものであっても良く、さらに前述した機能をコンピュータシステムにすでに記録されているプログラムとの組み合わせで実現できるものであっても良い。

【0183】

以上、本発明の実施形態を図面を参照して詳述してきたが、具体的な構成はこの実施形態に限られるものではなく、この発明の要旨を逸脱しない範囲の設計変更等も含まれる。

40

【符号の説明】

【0184】

- 10 近似計算装置
- 11 底変換部
- 12 指数分離部
- 13 整数部分計算部
- 14 小数部分計算部
- 15 乗算部
- 20 ニューラルネットワーク
- 21 非線形変換部

50

- 2 2 第 1 層
- 2 3 第 2 層
- 2 4 第 3 層

【 図 1 】

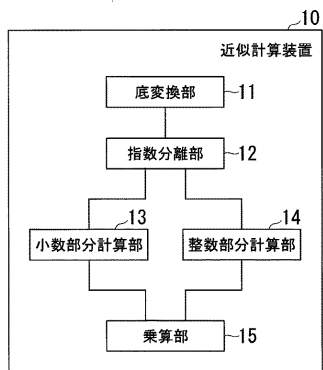


図 1

【 図 2 】

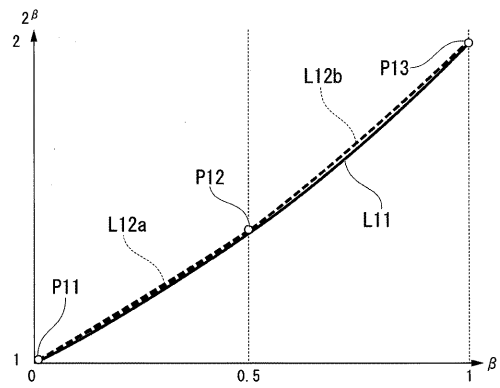


図 2

【 図 3 】

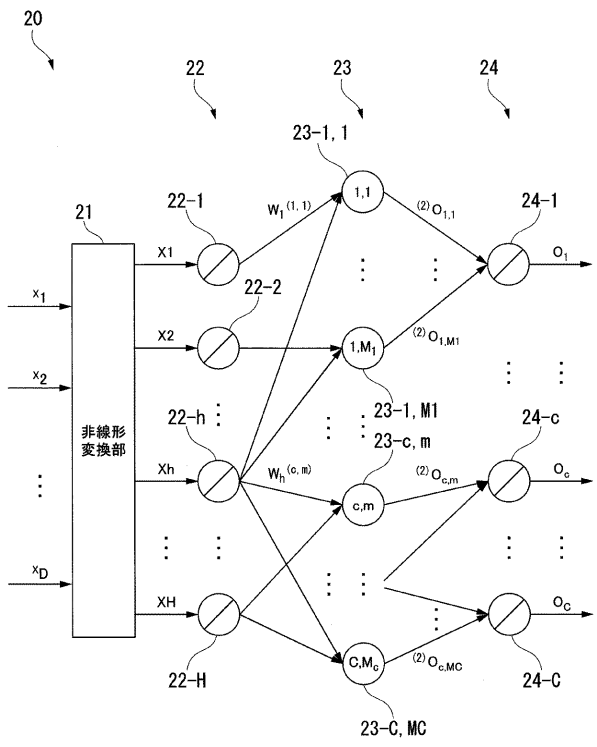


図 3

【 図 4 】

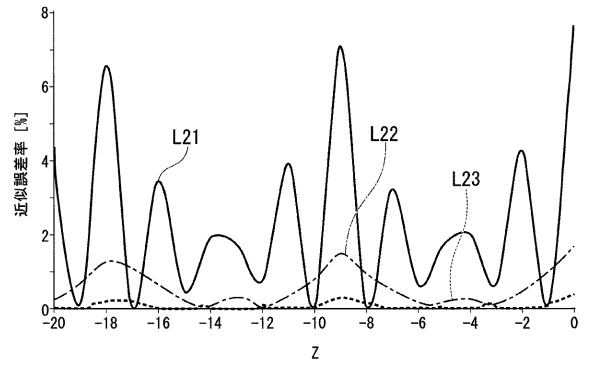


図 4

【 図 5 】

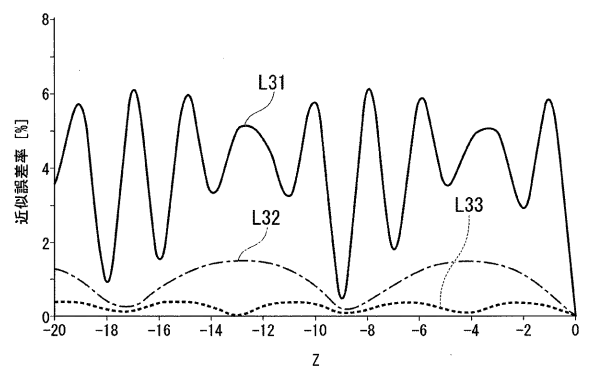


図 5

【 図 6 】

近似関数	n	近似誤差率の平均値 [%]	近似誤差率の最大値 [%]
g1	1	2.09±2.00	7.59
	2	0.51±0.47	1.69
	4	0.12±0.12	0.40
g'1	1	4.05±1.86	6.15
	2	1.00±0.46	1.51
	4	0.25±0.11	0.37

図 6

【 図 7 】

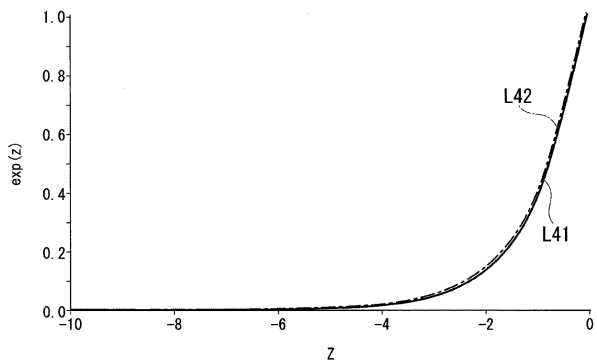


図 7

【 図 8 】

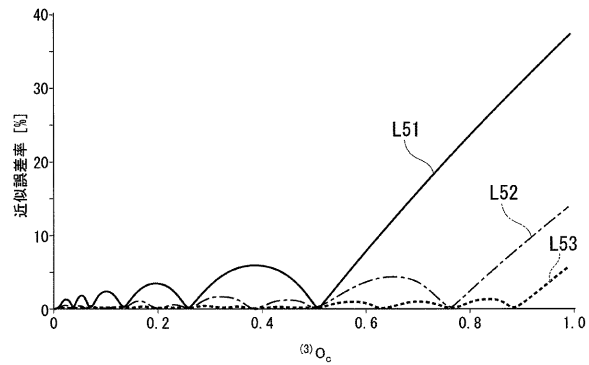


図 8



【 図 9 】

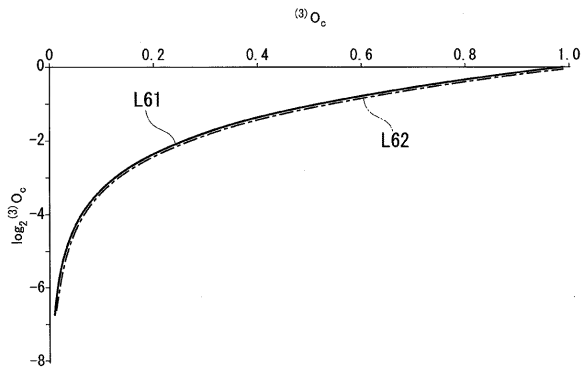


図 9