

(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2004-258748
(P2004-258748A)

(43) 公開日 平成16年9月16日(2004.9.16)

(51) Int. Cl. ⁷	F I	テーマコード (参考)
G06F 17/10	G06F 17/10	5B056
G06N 3/00	G06N 3/00 550A	

審査請求 未請求 請求項の数 15 O L (全 43 頁)

<p>(21) 出願番号 特願2003-46003 (P2003-46003)</p> <p>(22) 出願日 平成15年2月24日 (2003.2.24)</p> <p>特許法第30条第1項適用申請有り 2002年11月18日 Orchid Country Club発行の「Proceedings of the 9th International Conference on Neural Information Processing」に発表</p> <p>(出願人による申告) 国等の委託研究の成果に係る特許出願(通信・放送機構、平成14年4月1日付け委託契約研究テーマ「自律分散型無線ネットワークの研究開発」、産業再生法第30条の適用を受けるもの)</p>	<p>(71) 出願人 393031586 株式会社国際電気通信基礎技術研究所 京都府相楽郡精華町光台二丁目2番地2</p> <p>(74) 代理人 100064746 弁理士 深見 久郎</p> <p>(74) 代理人 100085132 弁理士 森田 俊雄</p> <p>(74) 代理人 100083703 弁理士 仲村 義平</p> <p>(74) 代理人 100096781 弁理士 堀井 豊</p> <p>(74) 代理人 100098316 弁理士 野田 久登</p> <p>(74) 代理人 100109162 弁理士 酒井 将行</p>
---	---

最終頁に続く

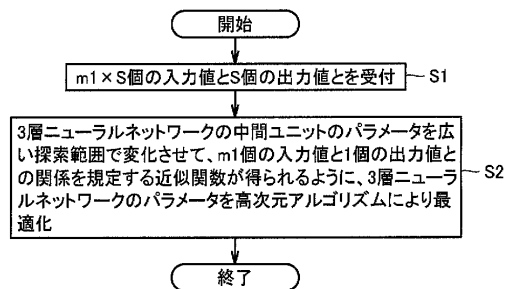
(54) 【発明の名称】 近似関数を求める演算をコンピュータに実行させるためのプログラムおよびそのプログラムを記録したコンピュータ読取り可能な記録媒体

(57) 【要約】

【課題】 入力と出力との関係を規定する近似関数を求める演算をコンピュータに実行させるためのプログラムを提供する。

【解決手段】 プログラムは、次のステップS1, S2を順次実行して近似関数を演算する。ステップS1は、サンプルデータに含まれるm1×S個の入力値及びS個の出力値を受付ける。ステップS2は、中間ユニットの個数を1個ずつ増加させながら中間ユニットのパラメータ w_{ij}, j, j を広い探索範囲で変化させて、m1個の入力値と1個の出力値との関係を規定する近似関数が得られるように、3層ニューラルネットワークのパラメータ $w_{ij}, j, j, W_j,$ を高次元アルゴリズムを用いて最適化する。

【選択図】 図17



【特許請求の範囲】**【請求項 1】**

各々が m (m は自然数) 個の入力値と n (n は自然数) 個の出力値とから成る S (S は自然数) 個のサンプルデータを用いて m 個の入力と n 個の出力との関係を規定する近似関数を求める演算をコンピュータに実行させるためのプログラムであって、

前記 $m \times S$ 個の入力値と前記 $n \times S$ 個の出力値とを受付ける第 1 のステップと、

前記 m 個の入力値に対して n 個の出力演算値を演算する超球面識別タイプの 3 層ニューラルネットワークの全パラメータのうち、識別超球面のパラメータの値を通常の探索範囲よりも広い探索範囲で変化させて、前記超球面識別タイプの演算により前記 $m \times S$ 個の入力値に対する $n \times S$ 個の出力演算値を演算し、その演算した $n \times S$ 個の出力演算値を用いて前記近似関数が得られるように前記全パラメータの値を最適化する第 2 のステップとをコンピュータに実行させ、

前記第 2 のステップは、前記全パラメータの数で定義される次元数よりも高い高次元空間を設定し、その設定した高次元空間において前記全パラメータの値が最適値以外である領域を素速く通過し、前記全パラメータの値が前記最適値である領域に容易に入ることが期待される高次元アルゴリズムにより前記全パラメータの最適化を行なう、コンピュータに実行させるためのプログラム。

【請求項 2】

前記第 2 のステップは、前記 3 層ニューラルネットワークに含まれ、かつ、前記超球面識別タイプの演算を行なう中間ユニットの個数を初期値に設定して前記 $n \times S$ 個の出力演算値を演算し、前記全パラメータの最適化を行なう、請求項 1 に記載のコンピュータに実行させるためのプログラム。

【請求項 3】

前記第 2 のステップは、

前記全パラメータを初期値に設定して前記超球面識別タイプの演算により前記 $n \times S$ 個の出力演算値を演算する第 1 のサブステップと、

前記演算された $n \times S$ 個の出力演算値を評価するコスト関数値を演算し、その演算したコスト関数値を所定値と比較する第 2 のサブステップと、

前記コスト関数値が前記所定値以下のとき、前記コスト関数値が得られるときの全パラメータの値を最適値とする第 3 のサブステップと、

前記コスト関数値が前記所定値よりも大きいとき、前記コスト関数値を低減させるための全パラメータを前記高次元アルゴリズムにより前記広い探索範囲で演算する第 4 のサブステップと、

前記第 4 のサブステップにより演算された全パラメータを用いて前記第 1 のサブステップを実行し、その後、前記第 2 から第 4 のサブステップを実行する第 5 のサブステップと、前記第 1 から第 5 のサブステップを規定回数まで繰返し実行したときのコスト関数値が前記所定値よりも大きいとき、前記中間ユニットの個数を増加して前記第 1 から第 5 のサブステップを実行する第 6 のサブステップとを含む、請求項 2 に記載のコンピュータに実行させるためのプログラム。

【請求項 4】

前記中間ユニットの個数は、1 個ずつ増加される、請求項 3 に記載のコンピュータに実行させるためのプログラム。

【請求項 5】

前記中間ユニットの個数の初期値は、1 である、請求項 2 から請求項 4 のいずれか 1 項に記載のコンピュータに実行させるためのプログラム。

【請求項 6】

前記第 2 のステップは、前記全パラメータの数を初期値に設定して前記 $n \times S$ 個の出力演算値を演算し、前記全パラメータの最適化を行なう、請求項 1 に記載のコンピュータに実行させるためのプログラム。

【請求項 7】

前記第 2 のステップは、

前記全パラメータを初期値に設定して前記超球面識別タイプの演算により前記 $n \times S$ 個の出力演算値を演算する第 1 のサブステップと、

前記演算された $n \times S$ 個の出力演算値を評価するコスト関数値を演算し、その演算したコスト関数値を所定値と比較する第 2 のサブステップと、

前記コスト関数値が前記所定値以下のとき、前記コスト関数値が得られるときの全パラメータの値を最適値とする第 3 のサブステップと、

前記コスト関数値が前記所定値よりも大きいとき、前記コスト関数値を低減させるための全パラメータを前記高次元アルゴリズムにより前記広い探索範囲で演算する第 4 のサブステップと、

10

前記第 4 のサブステップにより演算された全パラメータを用いて前記第 1 のサブステップを実行し、その後、前記第 2 から第 4 のサブステップを実行する第 5 のサブステップと、前記第 1 から第 5 のサブステップを規定回数まで繰返し実行したときのコスト関数値が前記所定値よりも大きいとき、前記全パラメータの数を増加して前記第 1 から第 5 のサブステップを実行する第 6 のサブステップとを含む、請求項 6 に記載のコンピュータに実行させるためのプログラム。

【請求項 8】

前記全パラメータは、所定数ずつ増加され、

前記第 1 から第 5 のサブステップは、前記全パラメータの数が増加されたとき、前記全パラメータの数が増加される前の所定数のパラメータの値を固定して実行される、請求項 7

20

に記載のコンピュータに実行させるためのプログラム。

【請求項 9】

前記全パラメータの数が増加される前の所定数のパラメータを第 1 のパラメータとし、前記増加された所定数のパラメータを第 2 のパラメータとしたとき、

前記第 4 のサブステップは、前記第 1 のパラメータを固定し、前記第 2 のパラメータを前記広い探索範囲で変化させて前記高次元アルゴリズムにより前記コスト関数値を低減させるための全パラメータを演算する、請求項 8 に記載のコンピュータに実行させるためのプログラム。

【請求項 10】

前記第 2 のサブステップは、前記受付けた $n \times S$ 個の出力値と前記演算された $n \times S$ 個の出力演算値との二乗誤差の和の平均を前記コスト関数値として演算する、請求項 3 ~ 5 及び請求項 7 ~ 9 のいずれか 1 項に記載のコンピュータに実行させるためのプログラム。

30

【請求項 11】

前記 $n \times S$ 個の出力値は、ガウシアン様の分布の結合により近似される、請求項 1 から請求項 10 のいずれか 1 項に記載のコンピュータに実行させるためのプログラム。

【請求項 12】

前記 $m \times S$ 個の入力値及び前記 $n \times S$ 個の出力値は、コンピュータにより演算されたデータである、請求項 1 から請求項 11 のいずれか 1 項に記載のコンピュータに実行させるためのプログラム。

【請求項 13】

40

前記 $m \times S$ 個の入力値及び前記 $n \times S$ 個の出力値は、微小構造中に閉じ込められた粒子の量子準位を演算する量子準位演算プログラムによって演算されたデータであり、

前記量子準位演算プログラムは、

線形のシュレディンガー方程式に基づいて初期の波動関数を演算し、その演算された初期の波動関数を複数の離散化された成分から成る数値列として与えるステップ A と、

前記離散化された複数の成分を持つ第 1 の波動関数と粒子の相互作用を考慮した非線形項を含むハミルトニアンとを用いて前記微小構造中に存在する粒子数で規格化され、かつ、全系のエネルギーを示すコスト関数を演算するステップ B と、

前記演算されたコスト関数を用いて、系の全体エネルギーが最小となる最終的な波動関数を演算するステップ C と、

50

前記最終的な波動関数と前記ハミルトニアンとを用いて前記最終的な波動関数で表わされる状態のエネルギーを演算するステップDとを含む、請求項12に記載のコンピュータに実行させるためのプログラム。

【請求項14】

前記高次元アルゴリズムは、

解くべき問題に現われ、かつ、前記最適化すべき全パラメータの空間を意味空間と定義するステップと、

前記全パラメータと共役な共役パラメータによって新しい空間を定義するステップと、

前記意味空間に前記新しい空間を加えて高次元空間を定義するステップと、

前記高次元空間において問題を設定するステップと、

前記全パラメータの値が最適値以外の値である領域を素速く通過し、前記全パラメータの値が最適値である領域に容易に入ることが期待される自律的運動を前記高次元空間において行なって前記全パラメータの最適値を検出するステップとから成る、請求項1から請求項13のいずれか1項に記載のコンピュータに実行させるためのプログラム。

10

【請求項15】

請求項1から請求項14のいずれか1項に記載されたプログラムを記録したコンピュータ読取り可能な記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

この発明は、入力と出力との関係を規定する近似関数を求める演算をコンピュータに実行させるためのプログラム及びそのプログラムを記録したコンピュータ読取り可能な記録媒体に関するものである。

20

【0002】

【従来の技術】

自己相互作用を考慮した精度の高い量子状態計算法がいくつか提案されている。いずれの方法も、量子井戸の構造パラメータ及び量子井戸に印加される外部電場等を入力変数の組とし、その入力変数の組からある1つの値の組が抽出されると、その抽出された1つの値の組を入力として微分方程式の積分及び各種パラメータの最適化等の計算量の多い演算を実行する。そして、与えられた入力に対して、物理的に許容されるエネルギー準位又は波動関数等が出力として計算される。

30

【0003】

なお、以上、本発明についての従来の技術を、出願人の知得した一般的技術情報に基づいて説明したが、出願人の記憶する範囲において、出願前までに先行技術文献情報として開示すべき情報を出願人は有していない。

【0004】

【発明が解決しようとする課題】

しかし、従来の方法では、個々の入力値に対して出力値を求めるために、計算のための専用のプログラム及び十分な計算リソース（メモリ及びCPU（Central Processing Unit）等）が必要である。また、実際のナノデバイスの設計においては、具体的な要求条件は、殆どの場合、エネルギー準位及び波動関数等、上述した計算手法の出力に対して課される。即ち、量子井戸におけるエネルギー準位又は波動関数が設計値として与えられ、その与えられたエネルギー準位又は波動関数を実現するために、井戸層の幅、井戸層のバンドギャップ、バリア層の高さ、バリア層の幅、バリア層のバンドギャップ及び井戸層又はバリア層におけるドーパントのドーピング量等のパラメータをどのように設定すればよいかを解く必要がある。

40

【0005】

この場合、上述した従来の方法を用いれば、試行錯誤的に入力値を変えながら条件を満たす系を探さなければならないが、そのためには、膨大な計算量が必要になり、ナノデバイスの設計の度に演算を行なうのは効率的でない。

50

【0006】

そこで、この発明は、かかる問題を解決するためになされたものであり、その目的は、入力と出力との関係を規定する近似関数を求める演算をコンピュータに実行させるためのプログラムを提供することである。

【0007】

また、この発明の別の目的は、入力と出力との関係を規定する近似関数を求める演算をコンピュータに実行させるためのプログラムを記録したコンピュータ読取り可能な記録媒体を提供することである。

【0008】

【課題を解決するための手段および発明の効果】

この発明によれば、各々が m (m は自然数)個の入力値と n (n は自然数)個の出力値とから成る S (S は自然数)個のサンプルデータを用いて m 個の入力と n 個の出力との関係を規定する近似関数を求める演算をコンピュータに実行させるためのプログラムは、 $m \times S$ 個の入力値と $n \times S$ 個の出力値とを受付ける第1のステップと、 m 個の入力値に対して n 個の出力演算値を演算する超球面識別タイプの3層ニューラルネットワークの全パラメータのうち、識別超球面のパラメータの値を通常の探索範囲よりも広い探索範囲で変化させて、超球面識別タイプの演算により $m \times S$ 個の入力値に対する $n \times S$ 個の出力演算値を演算し、その演算した $n \times S$ 個の出力演算値を用いて近似関数が得られるように全パラメータの値を最適化する第2のステップとをコンピュータに実行させ、第2のステップは、全パラメータの数で定義される次元数よりも高い高次元空間を設定し、その設定した高次元空間において全パラメータの値が最適値以外である領域を素速く通過し、全パラメータの値が最適値である領域に容易に入ることが期待される高次元アルゴリズムにより全パラメータの最適化を行なう、コンピュータに実行させるためのプログラムである。

10

20

【0009】

好ましくは、第2のステップは、3層ニューラルネットワークに含まれ、かつ、超球面識別タイプの演算を行なう中間ユニットの個数を初期値に設定して $n \times S$ 個の出力演算値を演算し、全パラメータの最適化を行なう。

【0010】

好ましくは、第2のステップは、全パラメータを初期値に設定して超球面識別タイプの演算により $n \times S$ 個の出力演算値を演算する第1のサブステップと、演算された $n \times S$ 個の出力演算値を評価するコスト関数値を演算し、その演算したコスト関数値を所定値と比較する第2のサブステップと、コスト関数値が所定値以下のとき、コスト関数値が得られるときの全パラメータの値を最適値とする第3のサブステップと、コスト関数値が所定値よりも大きいとき、コスト関数値を低減させるための全パラメータを高次元アルゴリズムにより広い探索範囲で演算する第4のサブステップと、第4のサブステップにより演算された全パラメータを用いて第1のサブステップを実行し、その後、第2から第4のサブステップを実行する第5のサブステップと、第1から第5のサブステップを規定回数まで繰返し実行したときのコスト関数値が所定値よりも大きいとき、中間ユニットの個数を増加して第1から第5のサブステップを実行する第6のサブステップとを含む。

30

【0011】

好ましくは、中間ユニットの個数は、1個ずつ増加される。
好ましくは、中間ユニットの個数の初期値は、1である。

40

【0012】

好ましくは、第2のステップは、全パラメータの数を初期値に設定して $n \times S$ 個の出力演算値を演算し、全パラメータの最適化を行なう。

【0013】

好ましくは、第2のステップは、全パラメータを初期値に設定して超球面識別タイプの演算により $n \times S$ 個の出力演算値を演算する第1のサブステップと、演算された $n \times S$ 個の出力演算値を評価するコスト関数値を演算し、その演算したコスト関数値を所定値と比較する第2のサブステップと、コスト関数値が所定値以下のとき、コスト関数値が得られる

50

ときの全パラメータの値を最適値とする第3のサブステップと、コスト関数値が所定値よりも大きいとき、コスト関数値を低減させるための全パラメータを高次元アルゴリズムにより広い探索範囲で演算する第4のサブステップと、第4のサブステップにより演算された全パラメータを用いて第1のサブステップを実行し、その後、第2から第4のサブステップを実行する第5のサブステップと、第1から第5のサブステップを規定回数まで繰返し実行したときのコスト関数値が所定値よりも大きいとき、全パラメータの数を増加して第1から第5のサブステップを実行する第6のサブステップとを含む。

【0014】

好ましくは、全パラメータは、所定数ずつ増加される。そして、プログラムの第1から第5のサブステップは、全パラメータの数が増加されたとき、全パラメータの数が増加される前の所定数のパラメータの値を固定して実行される。

10

【0015】

好ましくは、全パラメータの数が増加される前の所定数のパラメータを第1のパラメータとし、増加された所定数のパラメータを第2のパラメータとしたとき、第4のサブステップは、第1のパラメータを固定し、第2のパラメータを広い探索範囲で変化させて高次元アルゴリズムによりコスト関数値を低減させるための全パラメータを演算する。

【0016】

好ましくは、第2のサブステップは、受付けた $n \times S$ 個の出力値と演算された $n \times S$ 個の出力演算値との二乗誤差の和の平均をコスト関数値として演算する。

【0017】

好ましくは、 $n \times S$ 個の出力値は、ガウシアン様の分布の結合により近似される。

20

【0018】

好ましくは、 $m \times S$ 個の入力値及び $n \times S$ 個の出力値は、コンピュータにより演算されたデータである。

【0019】

好ましくは、 $m \times S$ 個の入力値及び $n \times S$ 個の出力値は、微小構造中に閉じ込められた粒子の量子準位を演算する量子準位演算プログラムによって演算されたデータである。そして、量子準位演算プログラムは、線形のシュレディンガー方程式に基づいて初期の波動関数を演算し、その演算された初期の波動関数を複数の離散化された成分から成る数値列として与えるステップAと、離散化された複数の成分を持つ第1の波動関数と粒子の相互作用を考慮した非線形項を含むハミルトニアンとを用いて微小構造中に存在する粒子数で規格化され、かつ、全系のエネルギーを示すコスト関数を演算するステップBと、演算されたコスト関数を用いて、系の全体エネルギーが最小となる最終的な波動関数を演算するステップCと、最終的な波動関数とハミルトニアンとを用いて最終的な波動関数で表わされる状態のエネルギーを演算するステップDとを含む。

30

【0020】

好ましくは、高次元アルゴリズムは、解くべき問題に現われ、かつ、最適化すべき全パラメータの空間を意味空間と定義するステップと、全パラメータと共役な共役パラメータによって新しい空間を定義するステップと、意味空間に新しい空間を加えて高次元空間を定義するステップと、高次元空間において問題を設定するステップと、全パラメータの値が最適値以外である領域を素速く通過し、全パラメータの値が最適値である領域に容易に入ることが期待される自律的運動を高次元空間において行なって全パラメータの最適値を検出するステップとから成る。

40

また、この発明によれば、近似関数を求める演算をコンピュータに実行させるためのプログラムを記録したコンピュータ読取り可能な記録媒体は、請求項1から請求項14のいずれか1項に記載されたプログラムを記録したコンピュータ読取り可能な記録媒体である。

【0021】

この発明によるプログラムは、超球面識別タイプの演算を行なうときの超球面パラメータを通常の探索範囲より広い範囲で変化させて3層ニューラルネットワークにより入力値に対する出力演算値を演算する。そして、この発明によるプログラムは、その演算した出力

50

演算値を評価するコスト関数が所定値以下になるように高次元アルゴリズムを用いてパラメータを最適化し、入力値と出力値との関係を規定する近似関数を演算する。

【0022】

従って、この発明によれば、入力値に対する出力値を容易に得ることが可能な近似関数を求めることができる。

【0023】

また、この発明においては、プログラムは、出力演算値を演算する際、パラメータの探索範囲を通常の探索範囲よりも広い範囲で変化させ、超球面識別タイプの演算を行なう。

【0024】

従って、この発明によれば、局所的特徴及び大局的特徴の両方を効率良く出力演算値に反映させることができる。 10

【0025】

更に、この発明においては、プログラムは、3層ニューラルネットワークのパラメータの数により設定される次元数よりも高い高次元空間において自律的運動を行なう高次元アルゴリズムを用いてパラメータの最適化を行なう。そして、自律的運動とは、パラメータの最適値以外の値が存在する領域を素速く通過し、最適値が存在する領域に容易に入る運動を言う。

【0026】

従って、この発明によれば、出力演算値を評価するコスト関数の局所解に捉まりにくく、かつ、コスト関数の平坦領域を素速く通過してパラメータを最適化できる。 20

【0027】

特に、超球面識別タイプの演算を行なう際にパラメータの探索範囲を広くすると、コスト関数の平坦領域が増加するが、高次元アルゴリズムは、平坦領域を素速く通過して最適解に到達する特徴を有するので、局所的特徴及び大局的特徴の両方を出力演算値に反映させ、かつ、早くパラメータを最適化できる。

【0028】

【発明の実施の形態】

本発明の実施の形態について図面を参照しながら詳細に説明する。なお、図中同一または相当部分には同一符号を付してその説明は繰返さない。

【0029】

図1は、この発明によるプログラムが近似関数を求める演算に用いる入力値と出力値とを示す。集合10は、入力値の組を示し、 $(x_1(1), \dots, x_{m_1}(1))$ 、 $(x_1(2), \dots, x_{m_1}(2))$ 、 \dots 、 $(x_1(S), \dots, x_{m_1}(S))$ 、 \dots 、 $(x_1(M), \dots, x_{m_1}(M))$ を含む。また、集合20は、出力値の集合を示し、 $z(1)$ 、 $z(2)$ 、 \dots 、 $z(S)$ 、 \dots 、 $z(M)$ を含む(S, M :自然数)。 30

【0030】

そして、出力値 $z(1)$ は、入力値の組 $(x_1(1), \dots, x_{m_1}(1))$ に対して得られ、出力値 $z(2)$ は、入力値の組 $(x_1(2), \dots, x_{m_1}(2))$ に対して得られ、以下、同様にして出力値 $z(M)$ は、入力値の組 $(x_1(M), \dots, x_{m_1}(M))$ に対して得られる。また、集合10に含まれる入力値の組 $(x_1(1), \dots, x_{m_1}(1))$ 、 $(x_1(2), \dots, x_{m_1}(2))$ 、 \dots 、 $(x_1(S), \dots, x_{m_1}(S))$ 、 \dots 、 $(x_1(M), \dots, x_{m_1}(M))$ 及び集合20に含まれる出力値 $z(1)$ 、 $z(2)$ 、 \dots 、 $z(S)$ 、 \dots 、 $z(M)$ は、コンピュータによって予め正確に演算された値である。 40

【0031】

なお、これらの入力値の組及び出力値を求める方法については後述する。

集合10及び20から近似関数を求める演算に用いるためのサンプルデータ30が生成される。即ち、サンプルデータ30は、サンプル1~サンプルSから成る。そして、サンプル1は、入力値の組 $(x_1(1), \dots, x_{m_1}(1))$ と出力値 $z(1)$ とから成り 50

、サンプル 2 は、入力値の組 $(x_1(2), \dots, x_{m-1}(2))$ と出力値 $z(2)$ とから成り、以下、同様にしてサンプル S は、入力値の組 $(x_1(S), \dots, x_{m-1}(S))$ と出力値 $z(S)$ とから成る。つまり、サンプルデータ 30 は、集合 10 及び 20 に含まれる入力値の組及び出力値から抽出された一部の入力値の組及び出力値によって構成される。

【0032】

このようにして、近似関数を求める演算に用いるべきサンプルデータ 30 が準備される。

【0033】

この発明によるプログラムは、サンプルデータ 30 を用いて入力値 $(x_1(1), \dots, x_{m-1}(1))$ 、 $(x_1(2), \dots, x_{m-1}(2))$ 、 \dots 、 $(x_1(S), \dots, x_{m-1}(S))$ と出力値 $(z(1), z(2), \dots, z(S))$ との関係 10
を規定する近似関数 $(z(n) = f(x_1(n), \dots, x_{m-1}(n)))$ 、 $(n = 1, \dots, S)$ を満たす関数 f を 3 層ニューラルネットワークに基づく関数モデル及び高次元アルゴリズム (新上和正、「高次元アルゴリズム」、Bit, Vol. 31, No. 7, pp. 2 - 8 (1999)、新上和正、「高次元アルゴリズム：最適化問題を解く 1 つの方法」、日本ファジィ学会誌、Vol. 11, No. 3, pp. 382 - 396 (1999) 参照) を用いて演算する。なお、以下においては、3 層ニューラルネットワークに基づく関数モデルによる演算を単に「3 層ニューラルネットワークによる演算」と言う。

【0034】

図 2 は、入力変数が 2 個の場合における入力値と出力値との関係を示す。図 2 の (a) は 20
、2 個の入力値の組 $x_1(n), x_2(n)$ と、2 個の入力値の組 $x_1(n), x_2(n)$ のプロット点及び関数 f の等高線とを示す。図 2 の (a) において、黒丸は、2 個の入力値の組 $x_1(n), x_2(n)$ の各々が $[0, 1]$ の範囲で変化した場合における 2 個の入力値の組 $x_1(n), x_2(n)$ のプロット点を示し、実線は、関数 f の等高線を示す。また、図 2 の (b) は、関数 f の 3 次元表現である。

【0035】

この場合、サンプルデータ 30 は、2 個の入力値の組 $x_1(n), x_2(n)$ を含むので、2 個の入力値の組 $x_1(n), x_2(n)$ は、プロット点 $(1), (2), \dots, (S)$ によって表わされる。そして、プロット点 $(1), (2), \dots, (S)$ の各々が 30
入力値として入力された場合、等高線 1 ~ 6 によって表わされる出力値が得られる。この等高線 1 ~ 6 によって表わされた曲面を 3 次元表現したものが図 2 の (b) に示す曲面 7 である。従って、プロット点 $(1), (2), \dots, (S)$ を入力値とした場合の出力値は、曲面 7 上に存在することになり、この発明によるプログラムは、プロット点 $(1), (2), \dots, (S)$ に対して曲面 7 を表わす関数 f を 3 層ニューラルネットワーク及び高次元アルゴリズムを用いて演算する。

【0036】

曲面 7 は、なだらかな曲面から成る。つまり、フーリエ級数又は三角関数の級数和によって表わした方が適切な激しい振動を多く含む曲面ではなく、ガウシアン様の分布の結合により表わされる。従って、この発明によるプログラムを用いて関数 f を求める演算を行なう場合、好ましくは、出力値 $(z(1), z(2), \dots, z(S))$ は、ガウシアン 40
様の分布の結合により表わされるような「なだらかな曲面」を構成する。

【0037】

なお、この発明において、「なだらかな曲面」とは、ガウシアン様の分布の結合により表わされる曲面を言う。

【0038】

図 3 は、この発明によるプログラムが入力値に対する出力演算値 $Z(n)$ を演算する 3 層ニューラルネットワークの概念図を示す。3 層ニューラルネットワーク 40 は、入力層 41 と、中間層 42 と、出力層 43 とを含む。

【0039】

入力層 41 は、 $m-1$ 個の入力ユニット $41_i (i = 1, \dots, m-1)$ から成る。中間層 50

4 2 は、 $m 2$ 個の中間ユニット 4 2 j ($j = 1, \dots, m 2$) から成る。出力層 4 3 は、1 個の出力ユニット 4 3 1 から成る。

【0040】

入力ユニット 4 1 i は、入力層 4 1 の i 番目のユニットに入力される入力値 $x_i(n)$ を受け、その受けた入力値 $x_i(n)$ を $m 2$ 個の中間ユニット 4 2 j の各々に伝達する。

【0041】

中間層 4 2 は、入力値と出力値との間の特徴抽出の主要な役割を担う層である。中間ユニット 4 2 j の内部状態、出力及び閾値をそれぞれ y_j, Y_j, τ_j とし、 i 番目の入力ユニット 4 1 i と j 番目の中間ユニット 4 2 j との間の結合のパラメータを w_{ij} としたとき、中間ユニット 4 2 j は、式 (1) により内部状態 y_j を演算する。

【0042】

【数 1】

$$y_j(n) = \theta_j^2 - \sum_{i=1}^{m1} (x_i(n) - w_{ij})^2 \quad \dots(1)$$

【0043】

即ち、中間ユニット 4 2 j は、超球面識別タイプの演算を行なう。中間ユニット 4 2 j が超球面識別タイプの演算を行なう理由については後述する。

【0044】

そして、中間ユニット 4 2 j は、演算した内部状態 y_j を式 (2) に代入して出力 Y_j を演算する。

【0045】

【数 2】

$$Y_j(n) = \frac{1}{1 + \exp(-y_j(n)/T_j)} \quad \dots(2)$$

【0046】

即ち、中間ユニット 4 2 j は、シグモイド関数により出力 Y_j を演算する。式 (2) において、 T_j は、シグモイド関数の遷移領域のスロープを調整するパラメータである。この実施の形態においては、式 (2) の右辺の分母に含まれる指数関数が数値的に発散するのを防止するために T_j は、式 (3) により定義される。

【0047】

【数 3】

$$T_j = \tau_j^2 + 0.01 \quad \dots(3)$$

【0048】

式 (3) において、 τ_j は、中間ユニット 4 2 j の出力関数の傾きを表わす。中間ユニット 4 2 j は、出力 Y_j を演算すると、その演算した出力 Y_j を出力層 4 3 の出力ユニット 4 3 1 へ出力する。

【0049】

出力層 4 3 は、 $m 2$ 個の中間ユニット 4 2 j の各々の出力結果を適切な重み付けにより最終出力の調整を行なう。従って、出力ユニット 4 3 1 は、式 (4) により出力演算値 $Z_1(n)$ を演算する。

【0050】

【数 4】

10

20

30

40

$$Z_1(n) = \sum_{j=1}^{m_2} W_j Y_j(n) - \Theta \quad \dots(4)$$

【0051】

式(4)において、 W_j は、 j 番目の中間ユニット42jとの結合重みであり、 Θ は、出力ユニット431の閾値である。

【0052】

なお、この発明においては、中間ユニットの個数は、最初、1個に設定され、その設定された1個の中間ユニットを用いた近似関数 f を求める演算結果に応じて、1個ずつ増加される。従って、出力ユニット431は、最初、中間ユニット421からの出力 Y_1 と、結合重み W_1 とを式(4)に代入して出力演算値 $Z_1(n)$ を演算し、中間ユニットの個数が増加されれば、その増加された中間ユニットからの出力及び増加された中間ユニットとの結合重みを用いて出力演算値 $Z_1(n)$ を演算する。

10

【0053】

このように、3層ニューラルネットワーク40は、 m_1 個の入力値($x_1(n), \dots, x_{m_1}(n)$)に対して、 w_{ij}, W_j をパラメータとして1個の出力演算値 $Z_1(n)$ を演算する。そして、 S 個のサンプル1~ S の各々は、 m_1 個のデータから成る入力値($x_1(n), \dots, x_{m_1}(n)$)を含むので、3層ニューラルネットワーク40は、 $m_1 \times S$ 個の入力値に対して S 個の出力演算値 $Z_1(1), \dots, Z_1(S)$ を演算する。

20

【0054】

中間ユニット42jが内部状態 y_j を求めるために式(1)により超球面識別タイプの演算を行なう理由について説明する。一般に、内部状態 y_j を求めるために式(5)により表わされる入力値 $x_i(n)$ と結合のパラメータ w_{ij} との積和演算がよく使用される。

【0055】

【数5】

$$y_j(n) = -\theta_j + \sum_{i=1}^{m_j} x_i \cdot w_{ij} \quad \dots(5)$$

30

【0056】

この場合、識別曲面の方程式 $y_j = 0$ は、1つの超平面を指定する。その例を図4及び図5に示す。図4は、2個の入力値の組 $x_1(n), x_2(n)$ を用いた場合に式(5)を用いて演算された内部状態 y_j を示す。また、図5は、2個の入力値の組 $x_1(n), x_2(n)$ を用いた場合の中間ユニットからの出力 Y_j を示す。

【0057】

図4に示すように、識別曲面の方程式 $y_j = 0$ は、1つの超平面50を指定する。そして、図5に示すように、中間ユニットの出力 Y_j は、超平面50を境界にして変化する曲面によって表わされる。超平面50の両側の領域51及び52は、ほぼ平坦であり、領域51及び52の各々において、更に狭い領域における特徴を表現することはできない。即ち、2個の入力値の組 $x_1(n), x_2(n)$ に対して演算された中間ユニットの出力 Y_j は、超平面50を境界にして変化するという大局的特徴を表現できるが、領域51及び52の更に狭い領域における局所的特徴を表現することができない。

40

【0058】

式(5)を用いて内部状態 y_j を演算することを超平面識別タイプの演算をすと言う。そして、超平面識別タイプの演算を行なった場合、上述したように、大局的特徴を表現できるが、1つの中間ユニットで局所的特徴を表現することはできない。局所的特徴を表現するには多数の中間ユニットを要する。

【0059】

50

一方、式(1)を用いて内部状態 y_j を演算した場合、識別曲面の方程式 $y_j = 0$ は、1つの超球面を指定する。その例を図6及び図7に示す。図6は、2個の入力値の組 $x_1(n)$ 、 $x_2(n)$ を用いた場合に式(1)を用いて演算された内部状態 y_j を示す。また、図7は、2個の入力値の組 $x_1(n)$ 、 $x_2(n)$ を用いた場合の中間ユニット42jからの出力 Y_j を示す。

【0060】

図6に示すように、識別球面の方程式 $y_j = 0$ は、1つの超球面60を指定する。そして、図7に示すように、中間ユニット42jの出力 Y_j は、超球面60の外側では平面であり、超球面60の内側では凸曲面になる。また、超球面60は、一般的には狭い領域に形成される。従って、中間ユニット42jの内部状態 y_j を超球面識別タイプの演算により求めることにより局所的特徴を表現することができる。

10

【0061】

上述したように、超平面識別タイプの演算により内部状態 y_j を求めれば、大局的特徴を表現できるが、局所的特徴を表現できない。一方、超球面識別タイプの演算により内部状態 y_j を求めれば、局所的特徴を表現できるが、大局的特徴を表現することができない。

【0062】

従って、理想的には、局所的特徴の表現に有利な中間ユニットと、大局的特徴の表現に有利な中間ユニットとを揃えればよいが、それぞれどれだけの個数を揃えればよいかが未知であるため超平面識別タイプの演算及び超球面識別タイプの演算を混在して行なえば、計算量が増加する。

20

【0063】

そこで、この発明においては、局所的特徴の表現に有利な超球面識別タイプの演算を採用し、中間ユニット42jが内部状態 y_j 及び出力 Y_j を演算する際のパラメータ (c_j , w_{ij} , t_j) を通常の探索範囲よりも広い範囲で変化させることにより大局的特徴を表現することにした。

【0064】

2個の入力値の組 $x_1(n)$ 、 $x_2(n)$ を用いた場合、識別曲面の方程式 $y_j = 0$ は、式(1)より、パラメータ w_{ij} を中心とし、パラメータ c_j を半径とする円を指定する。パラメータ c_j 及び w_{ij} の範囲を変えた場合の関数 f の変化について説明する。

【0065】

図8は、パラメータ c_j 及び w_{ij} の範囲を局所的特徴に相応する小さな半径にした場合における識別超球面の取り得る相対的位置関係と中間ユニット42jの出力例とを示す。図9は、パラメータ c_j 及び w_{ij} の範囲を広くした場合における識別超球面の取り得る相対的位置関係と中間ユニット42jの出力例とを示す。図10は、パラメータ w_{ij} (中心) が入力変数 x_1 、 x_2 の定義域外に存在する小さな超球である場合における識別超球面の取り得る相対的位置関係と中間ユニット42jの出力例とを示す。

30

【0066】

パラメータ (w_{1j} , w_{2j}) (中心) が入力変数 x_1 、 x_2 の定義域内 ($[0, 1]^2$ の範囲) にあって、パラメータ c_j (半径) の取り得る範囲が $[0, 0.5]$ 程度と小さい場合 (図8の(a)参照)、中間ユニット42jは、局所的特徴を反映した曲面61又は62によって表わされる出力 Y_j を出力する (図8の(b)参照)。

40

【0067】

曲面61は、パラメータ (w_{1j} , w_{2j}) (中心) が (0.5, 0.5) であり、パラメータ c_j (半径) が 0.1 であり、 T_j が 0.03 である場合に得られる。また、曲面62は、パラメータ (w_{1j} , w_{2j}) (中心) が (0.5, 0.5) であり、パラメータ c_j (半径) が 0.1 であり、 T_j が 0.05 である場合に得られる。

【0068】

このように、パラメータ (w_{1j} , w_{2j}) (中心) を入力変数 x_1 、 x_2 の定義域内に設定し、パラメータ c_j (半径) を小さい値にした場合、中間ユニット42jは、局所的特徴を反映した曲面61及び62によって表わされる出力 Y_j を出力する。

50

【0069】

パラメータ (w_{1j} , w_{2j}) (中心) が入力変数 x_1 , x_2 の定義域外 ($[-0.2, 1.2]^2$ の範囲) にあって、パラメータ r_j (半径) の取り得る範囲が $[0, 1.5]$ と大きい場合 (図9の(a)参照)、中間ユニット42jは、曲面63又は64によって表わされる出力 Y_j を出力する (図9の(b)参照)。

【0070】

曲面63は、パラメータ (w_{1j} , w_{2j}) (中心) が (1.2, 1.2) であり、パラメータ r_j (半径) が 1.0 であり、 T_j が 0.1 である場合に得られる。また、曲面64は、パラメータ (w_{1j} , w_{2j}) (中心) が (1.2, 1.2) であり、パラメータ r_j (半径) が 1.0 であり、 T_j が 0.5 である場合に得られる。

10

【0071】

曲面63及び64は、図5に示す曲面に似た曲面である。従って、パラメータ (w_{1j} , w_{2j}) (中心) を入力変数 x_1 , x_2 の定義域外に設定し、パラメータ r_j (半径) を大きい値にした場合、中間ユニット42jは、大局的特徴を反映した曲面63及び64によって表わされる出力 Y_j を出力する。つまり、中間ユニット42jは、近似的に超平面識別タイプの演算を行なう。このことは、超球面識別タイプの演算においてパラメータ r_j 及び w_{ij} の範囲を局所的特徴を反映する範囲よりも広い範囲に設定することによって、中間ユニット42jは、大局的特徴を反映した出力 Y_j を出力できることを意味する。

【0072】

パラメータ (w_{1j} , w_{2j}) (中心) が入力変数 x_1 , x_2 の定義域外にあって、パラメータ r_j (半径) が小さい場合 (図10の(a)参照)、中間ユニット42jは、曲面65によって表わされる出力 Y_j を出力する (図10の(b)参照)。

20

【0073】

この場合、中間ユニット42jは、定義域全域において、ほぼゼロの値から成る曲面65を出力し、最終出力である出力演算値 $Z_1(n)$ に殆ど寄与しない。

【0074】

上述したように、パラメータ r_j , w_{ij} 及び T_j (つまり r_j) の範囲を変化させることにより、中間ユニット42jは、超球面識別タイプの演算において局所的特徴を反映した出力 Y_j (図8参照) 及び大局的特徴を反映した出力 Y_j (図9参照) の両方を出力することが可能である。そして、超球面識別タイプの演算により局所的特徴及び大局的特徴を反映するようにしても中間ユニット42jのパラメータの数が増加することはない。従って、この発明においては、中間ユニット42jは、パラメータ r_j , w_{ij} 及び T_j (つまり r_j) の範囲を通常の超球面識別タイプの演算を行なう範囲よりも広い範囲まで変化させて超球面識別タイプの演算を行ない、出力 Y_j を出力することにした。これが、中間ユニット42jにおいて超球面識別タイプの演算を行なうことにした理由である。

30

【0075】

なお、「通常の探索範囲」とは、超球面識別タイプの演算により局所的特徴のみを出力 Y_j に反映させるためにパラメータ r_j , w_{ij} 及び T_j (つまり r_j) を変化させる範囲を言う。

【0076】

中間ユニット42jからの出力 Y_j は、出力ユニット431に入力され、出力ユニット431は、出力 Y_j 及びパラメータ w_j , を式(4)へ代入して出力演算値 $Z_1(n)$ を演算する。

40

【0077】

出力演算値 $Z_1(n)$ が演算されると、サンプルデータ30に含まれる出力値 $z_1(n)$ と出力演算値 $Z_1(n)$ との二乗誤差の和の平均 $V(w_{ij}, r_j, T_j)$ が式(6)により演算される。

【0078】

【数6】

$$V(w_{ij}, \theta_j, \tau_j, W_j, \theta) = \frac{1}{S} \sum_{n=1}^S \{z(n) - Z(x_1(n), \dots, x_{m1}(n))\}^2 \quad \dots (6)$$

【0079】

二乗誤差の和の平均 $V(w_{ij}, \theta_j, \tau_j, W_j, \theta)$ は、3層ニューラルネットワーク40によって演算された出力演算値 $Z_1(n)$ が実際の出力値 $z_1(n)$ に近い度合いを示す指標であり、出力演算値 $Z_1(n)$ を評価するコスト関数である。

【0080】

この発明によるプログラムは、パラメータ $w_{ij}, \theta_j, \tau_j, W_j, \theta$ を変化させて(中間ユニット42jのパラメータ w_{ij}, θ_j, τ_j については、局所的特徴を反映する範囲よりも広い範囲で変化させる)、出力演算値 $Z_1(n)$ を演算し、出力演算値 $Z_1(n)$ を式(6)に代入して演算したコスト関数 $V(w_{ij}, \theta_j, \tau_j, W_j, \theta)$ の関数値が所定値以下になるようにパラメータ $w_{ij}, \theta_j, \tau_j, W_j, \theta$ を最適化する。この場合、この発明によるプログラムは、最初、中間ユニット42jの個数を1個に設定し($j=1$)、その設定した1個の中間ユニット421を用いて最終的に演算したコスト関数 $V(w_{ij}, \theta_j, \tau_j, W_j, \theta)$ の関数値が所定値よりも大きいとき、関数値が所定値以下になるまで中間ユニット42jの個数を1個ずつ増加し、関数値が所定値以下になるようにパラメータ $w_{ij}, \theta_j, \tau_j, W_j, \theta$ を最適化する。

【0081】

このパラメータ $w_{ij}, \theta_j, \tau_j, W_j, \theta$ が最適化されれば、入力値 $x_1(n), \dots, x_{m1}(n)$ と出力値 $z_1(n)$ ($n=1 \sim S$) との関係の規定する近似関数 f が決定されるので、パラメータ $w_{ij}, \theta_j, \tau_j, W_j, \theta$ を最適化することは、近似関数 f を求める演算を行なうことに相当する。

【0082】

上述したように、中間ユニット42jは、パラメータ w_{ij}, θ_j, τ_j を局所的特徴を反映する範囲よりも広い範囲で変化させて超球面識別タイプの演算を行なうが、パラメータ w_{ij}, θ_j, τ_j を広い範囲で変化させると、コスト関数 $V(w_{ij}, \theta_j, \tau_j, W_j, \theta)$ において広い平坦領域を増加させることになる。図10の(a)に示すように、中心 w_{ij} が入力値の定義域外にあり、半径 τ_j が小さい場合、中間ユニット42jの出力 Y_j は、定義域全域においてほぼゼロとなり、最終出力である出力演算値 $Z_1(n)$ に殆ど寄与しない。従って、パラメータ空間における、このような状況に対応する領域ではコスト関数 $V(w_{ij}, \theta_j, \tau_j, W_j, \theta)$ が平坦になってしまう。

【0083】

$w_{ij}, \theta_j, \tau_j, W_j, \theta$ のようなパラメータを最適化する場合によく用いられる誤差逆伝播法(BP: Error Back-Propagation Algorithm)又は焼きなまし法(SA: Simulated Annealing)のようなアルゴリズムは、広い平坦領域を含むコスト関数に対して非常に効率が悪い。広い平坦領域の出現を避けるためにコスト関数にペナルティ項を追加する方法も考えられるが、この方法では、本来のコスト関数をゆがめたり、コスト関数の評価に要する計算量を増加させることになる。

【0084】

そこで、この発明においては、広い平坦領域、及びニューラルネットワークの学習において常に問題となる複数の局所解を含むコスト関数に対して有効な最適化手法である高次元アルゴリズムを採用することにした。

【0085】

高次元アルゴリズムによるパラメータの最適化について説明する。図11は、高次元アルゴリズムによる最適化のフローチャートを示す。解くべき問題に現われる全ての最適化すべき変数 q の空間を意味空間と定義する(ステップA)。そして、変数 q と共役な変数 p

10

20

30

40

50

を人為的に導入し、変数 p の新しい空間を定義する（ステップ B）。その後、変数 p の空間を意味空間に加え、意味空間を高次元化した空間を高次元空間と定義する（ステップ C）。

【0086】

そして、変数 q , p の高次元空間において、問題を設定する（ステップ D）。最後に、高次元空間において、自律的運動によって最適解を探索し、コスト関数を最小とする最適解を検出する（ステップ E）。ここで、自律的運動とは、解の存在しない領域を素速く通過し、解の存在する領域に容易に入る運動を言う。

【0087】

このように、高次元アルゴリズムは、最適化すべき変数 q の意味空間を高次元化し、その高次元化した高次元空間において自律的運動を行なうことによって最適解を検出する。 10

【0088】

図 12 は、高次元アルゴリズムによる解の探索方法を示す概念図である。また、図 13 は、パラメータが 2 個 (k_1 , k_2) の場合におけるコスト関数のランドスケープを示す。図 13 に示すランドスケープには、多くの山及び谷が存在し、解は全ての谷の最も低い点（極小値）に対応する。従って、解を求めるためには、複雑に入り組んだ多くの山及び谷を通過して最も低い谷に到達する必要がある。

【0089】

つまり、図 12 に示すように、最適化すべき変数 q の意味空間 70 において解 71 に到達するには、矢印 72 で示される経路を移動して多くの山及び谷を通過する必要がある。しかし、大きな意味空間 70 において小さな解 71 に到達するのは困難である。 20

【0090】

そこで、変数 q と共役な変数 p を追加して変数 q , p の高次元空間 80 を定義する。この高次元空間 80 においては、意味空間 70 において解でない領域は小さく、解に相当する領域は拡大される。従って、高次元空間 80 においては、意味空間 70 における解 71 が解の存在する領域 81 に拡大され、矢印 82 によって示される経路を運動して領域 81 に容易に到達する。そして、高次元空間 80 では解でない領域は小さくなり、解に相当する領域は拡大されるので、矢印 82 によって示される経路を運動する場合、解でない領域を素速く通過し、解の存在する領域 81 に容易に入る。つまり、高次元空間 80 においては、自律的運動が行なわれて解が検出される。 30

【0091】

このように、探索すべき空間を高次元化すれば目的物を探索し易いことは、次の例によって明確に理解できる。例えば、長い切り口の六角形の鉛筆を探索する場合を考える。この鉛筆を正面から見ると、小さい六角形の断面が見えるだけであるが、視線をシフトすれば、その切り口の長さも見える。この「長さ」を次元と考えれば、上述した意味空間 70 から高次元空間 80 へ次元を高次元化することにより、解を探索し易くなることを容易に理解できる。

【0092】

高次元アルゴリズムは、このように探索すべき空間を高次元化し、その高次元化した高次元空間において解を探索する結果、コスト関数の局所解に捉まりにくく、かつ、コスト関数の平坦領域を素速く通過でき、大局解を容易に求めることができる。 40

【0093】

3 層ニューラルネットワーク 40 により演算された出力演算値 $Z_1(n)$ が実際の出力値 $z_1(n)$ に近くなるように高次元アルゴリズムを用いてパラメータ w_{ij} , j , j , w_j , を最適化する場合、この発明によるプログラムは、高次元空間 80 における変数 q , p の関数であるハミルトニアン $H(q, p)$ を用いる。このハミルトニアン $H(q, p)$ は、運動する力学系を表わす具体的な道具であり、この具体的な道具（ハミルトニアン $H(q, p)$ ）によって表わされた空間においては、意味空間 70 を成す変数 q と、意味空間 70 を高次元空間 80 へ高次元化し、かつ、変数 q と共役である変数 p とを導入し易いからである。実際の力学系の場合、変数 q は、運動する物体の位置を表わし、変数 50

p は、運動する物体の速度を表わす。高次元アルゴリズムは、この力学系のアナロジーを取って最適パラメータを探索することを特徴とする最適化手法である。

【0094】

この発明の近似関数を求める問題の場合、高次元アルゴリズムによって最適化すべき変数 q は $q = w_{i j}, j, j, W_j,$ となる (図 11 のステップ A 参照)。そして、この変数 $q = w_{i j}, j, j, W_j,$ は、運動する物体の位置に対応する。ポテンシャルエネルギー $V(q)$ は、変数 $q = w_{i j}, j, j, W_j,$ の関数である。

【0095】

次に、変数 q に共役な変数 p が人為的に導入され、変数 p の成す新しい空間が定義される (図 11 のステップ B 参照)。この新しい空間を意味空間に加えたものを高次元空間と定義する (図 11 のステップ C 参照)。変数 p は、上述したように運動する物体の速度に対応し、変数 p の関数である運動エネルギー $T(p)$ に対応する関数を運動する力学系と同様に定義する (図 11 のステップ D 参照)。

10

【0096】

その後、意味空間における関数 $V(q)$ に新しい空間における関数 $T(p)$ を加えて意味空間を高次元化した高次元空間における関数 $H(q, p)$ が変数 q, p によって定義される (図 11 のステップ D 参照)。

【0097】

ハミルトニアン力学系においては、運動する物体の任意の時間 t における位置は、ハミルトニアン $H(q, p)$ から導かれる運動方程式により決定されるため、高次元アルゴリズムにおいても同様にする。

20

【0098】

最後に、高次元空間において、自律的運動によって最適解を検出することは、ハミルトニアン $H(q, p)$ から導かれる運動方程式を解いて最適化された変数 q を見つけることに相当する (図 11 のステップ E 参照)。

【0099】

ポテンシャルエネルギー $V(q)$ は、式 (7) によって表わされる。

【0100】

【数 7】

$$V(q) = V_C(q) + V_L(q) \quad \dots(7)$$

30

【0101】

即ち、ポテンシャルエネルギー $V(q)$ は、コスト関数 $V_C(q)$ と拘束ポテンシャル $V_L(q)$ との和とする。この拘束ポテンシャル $V_L(q)$ は、パラメータ $q = w_{i j}, j, j, W_j,$ の探索範囲を制限するものである。そして、拘束ポテンシャル $V_L(q)$ は、式 (8) によって表わされる。

【0102】

【数 8】

$$V_L(q) = c \sum_{n=1}^N v_n(q_n) \quad \dots(8)$$

40

【0103】

式 (8) において、c は、拘束の強さを制御するパラメータであり、この実施の形態においては $c = 1$ に設定される。また、式 (8) における $v_n(q_n)$ は、式 (9) によって表わされる。

【0104】

【数 9】

$$v_n(q_n) = (a_n - q_n)^2 \theta(a_n - q_n) + (q_n - b_n)^2 \theta(q_n - b_n) \quad \dots(9)$$

【 0 1 0 5 】

式(9)において、 $\theta(u)$ は、階段関数を表わし、 $u = a_n - q_n$ 又は $q_n - b_n > 0$ のとき、 $\theta(u) = 1$ であり、 $u = a_n - q_n$ 又は $q_n - b_n < 0$ のとき、 $\theta(u) = 0$ である。

【 0 1 0 6 】

即ち、 $v_n(q_n)$ は、図14に示す曲線89によって表わされ、パラメータ $q = w_{ij}$ 、 j 、 j 、 W_j の探索範囲の限界を規定する。

【 0 1 0 7 】

また、運動エネルギー $T(p)$ は、式(10)によって表わされ、ハミルトニアン $H(q, p)$ は、式(11)によって表わされる。

【 0 1 0 8 】

【 数 1 0 】

$$T(p) = \frac{1}{2} p^2 \quad \dots(10)$$

【 0 1 0 9 】

【 数 1 1 】

$$H(q, p) = T(p) + V(q) \quad \dots(11)$$

【 0 1 1 0 】

その結果、系の運動を記述する運動方程式は、式(12)によって表わされる。

【 0 1 1 1 】

【 数 1 2 】

$$\begin{cases} \frac{dq_i}{dt} = \frac{\partial H(q, p)}{\partial p_i} = p_i \\ \frac{dp_i}{dt} = -\frac{\partial H(q, p)}{\partial q_i} = -\frac{\partial V_C(q)}{\partial q_i} - \frac{\partial V_L(q)}{\partial q_i} \end{cases} \quad \dots(12)$$

(i = 1, \dots, N)

【 0 1 1 2 】

また、式(12)の下側の式の右辺に現われるコスト関数 $V_C(q)$ 及び拘束ポテンシャルを示す関数 $V_L(q)$ のパラメータ $q = w_{ij}$ 、 j 、 j 、 W_j による微分形をそれぞれ式(13)及び式(14)に示す。

【 0 1 1 3 】

【 数 1 3 】

10

20

30

40

$$\left. \begin{aligned}
 \frac{\partial V_C}{\partial w_{ij}} &= \frac{4}{ST_j} \sum_{n=1}^S (Z(n) - z(n)) W_j (x_i(n) - w_{ij}) Y_j(n) (1 - Y_j(n)) \\
 \frac{\partial V_C}{\partial \theta_j} &= \frac{4\theta_j}{ST_j} \sum_{n=1}^S (Z(n) - z(n)) W_j Y_j(n) (1 - Y_j(n)) \\
 \frac{\partial V_C}{\partial \tau_j} &= \frac{4\tau_j}{ST_j^2} \sum_{n=1}^S (Z(n) - z(n)) W_j y_j(n) Y_j(n) (1 - Y_j(n)) \\
 \frac{\partial V_C}{\partial W_j} &= \frac{2}{S} \sum_{n=1}^S (Z(n) - z(n)) Y_j(n) \\
 \frac{\partial V_C}{\partial \Theta} &= -\frac{2}{S} \sum_{n=1}^S (Z(n) - z(n))
 \end{aligned} \right\} \dots (13)$$

10

【0114】

【数14】

20

$$\frac{\partial V_L}{\partial q_i} = 2c \{ -(a_n - q_n) \theta(a_n - q_n) + (q_n - b_n) \theta(q_n - b_n) \} \quad \dots (14)$$

但し、 $i=1, \dots, N$

【0115】

ランダムに選択した初期値 ($q(0)$, $p(0)$) を出発点とし、式(12)の $2N$ 個の一階微分方程式の組をベルレー (Verlet) 法又はルンゲクッタ (Runge-Kutta) 法を用いて数値的に解くことにより、系の軌道 ($q(t)$, $p(t)$) が得られる。そして、十分な時間、軌道 $q(t)$ に沿ったコスト関数 $V_C(q(t))$ を監視することにより、パラメータ $q = w_{ij}$, θ_j , τ_j , W_j の最適値を見つけることができる。

30

【0116】

式(12)に示す運動方程式により記述される運動系は、混合性を持ち、かつ、全エネルギー E が一定である力学系である。この場合、系は、位相空間の $H(q, p) = E$ を満たす等エネルギー曲面上を等しい確率で到る所を動き回ることが期待される (等重率の原理)。この等重率の原理に基づけば、系が位置 q の近傍の微小体積 $q + dq$ 内に滞在する時間の期待値 ($\tau(q)$) は、式(15)によって表わされる (ケイ、シンジヨー (K. Shinjo) 及びティー、ササダ (T. Sasada) 著、「ハミルトニアン システムズ ウィズ メニィ ディグリー オブ フリーダム: アシンメトリック モーション アンド インテンシティ オブ モーション イン フェーズ スペース (Hamiltonian systems with many degrees of freedom: asymmetric motion and intensity of motion in phase space)」、フィジカル レビュー (Physical Review) E54, pp4685-4700, 1996)。

40

【0117】

【数15】

$$\delta(q) \propto [E - V(q)]^{N/2-1} dq \quad \dots(15)$$

【0118】

式(15)より、自由度Nが3以上のとき、高いポテンシャル値を持つ領域では滞在時間の期待値が低く、低いポテンシャル値を持つ領域では滞在時間の期待値が高くなり、更に、この傾向は、自由度Nが大きければ大きいほど顕著になる。

【0119】

従って、高次元アルゴリズムは、このような特徴を有するハミルトニアン $H(q, p)$ の力学系のアナロジーを取って最適解を探索することにより、上述したコスト関数の局所解に捉まりにくく、かつ、コスト関数の平坦な領域を素速く通過するという特徴を有する。

【0120】

この高次元アルゴリズムの特徴を概念的に説明する。

図15は、コスト関数の局所解に捉まる度合いを高次元アルゴリズム(HA)と焼きなまし法(SA)について比較して示す。図15の(a)は、高次元アルゴリズム(HA)の場合を示し、図15の(b)は、焼きなまし法(SA)の場合を示す。図15の(a)及び(b)において、横軸は最適化すべき変数 q を表わし、縦軸はコスト関数 $V(q)$ を表わす。

【0121】

また、図16は、コスト関数の平坦領域を通過する速さを高次元アルゴリズム(HA)と焼きなまし法(SA)について比較して示す。図16の(a)は、高次元アルゴリズム(HA)の場合を示し、図16の(b)は、焼きなまし法(SA)の場合を示す。図16の(a)及び(b)において、横軸は最適化すべき変数 q を表わし、縦軸はコスト関数 $V(q)$ を表わす。

【0122】

図15に示すように、高次元アルゴリズム(HA)の場合、コスト関数 $V(q)$ の谷(局所解)から抜け出す役割を運動エネルギー E が担うが、全エネルギー一定の条件より、運動エネルギー E は、コスト関数の関数値が小さい位置では大きくなる。つまり、局所解に入り込んだら運動エネルギー E が大きくなる。従って、高次元アルゴリズム(HA)は、局所解に捉まりにくい(図15の(a)参照)。一方、焼きなまし法(SA)の場合、運動は、正の絶対温度を持つことにより局所解から抜け出すことが可能であるが、この絶対温度は位置には依存しない。従って、絶対温度がコスト関数の山よりも低ければ、局所解を抜け出すことが困難である。その結果、焼きなまし法(SA)は、コスト関数の局所解に捉まり易い(図15の(b)参照)。

【0123】

また、図16に示すように、高次元アルゴリズム(HA)の場合、等速直線運動により一方の方向に運動するため、コスト関数 $V(q)$ の平坦領域を素速く通過する(図16の(a)参照)。一方、焼きなまし法(SA)の場合、ランダムウォーク(紙面の右方向及び左方向に各ステップごとにランダムに運動する)によって運動するため、コスト関数 $V(q)$ の平坦領域をなかなか通過できない(図16の(b)参照)。

【0124】

このように、高次元アルゴリズムは、コスト関数の局所解に捉まりにくく、かつ、平坦領域を素速く通過できるという特徴を有する。その結果、少ない計算量によって最適解に到達できる。

【0125】

図17は、この発明によるプログラムが入力値の組 $x_1(n), \dots, x_{m-1}(n)$ と出力値 $z_1(n)$ との関係を規定する近似関数 f を求める演算を行なうフローチャートを示す。近似関数 f を求める演算が開始されると、サンプルデータ30に含まれる $m-1 \times S$ 個の入力値及び S 個の出力値が受け付けられる(ステップS1)。そして、1個に設定された中間ユニット $42_j (j=1)$ を用いて3層ニューラルネットワーク40のパラメータ

10

20

30

40

50

$w_{ij}, j, j, W_j,$ が最適化される。

【0126】

即ち、3層ニューラルネットワーク40の中間ユニット42jのパラメータ w_{ij}, j, j を広い探索範囲で変化させて、m1個の入力値と1個の出力値との関係を模倣する近似関数fが得られるように、3層ニューラルネットワーク40のパラメータ $w_{ij}, j, j, W_j,$ を高次元アルゴリズムを用いて最適化する(ステップS2)。これにより、パラメータ $w_{ij}, j, j, W_j,$ が最適化されれば、近似関数fが決定されるので、一連の動作は終了する。

【0127】

図18は、図17に示すステップS2の詳細な動作を説明するためのフローチャートを示す。図17に示すステップS1の後、3層ニューラルネットワーク40のパラメータ $w_{ij}, j, j, W_j,$ を初期値に設定し、超球面識別タイプの演算によりS個の出力演算値 $Z_1(1), \dots, Z_1(S)$ が演算される(ステップS21)。

【0128】

そして、S個の出力演算値 $Z_1(1), \dots, Z_1(S)$ を評価するコスト関数値が演算される(ステップS22)。その後、コスト関数値が所定値以下であるか否かが判定され(ステップS23)、コスト関数値が所定値以下であるとき一連の動作は終了する。

【0129】

一方、ステップS23において、コスト関数値が所定値以下でないと判定されたとき、計算回数が規定回数以下であるか否かが判定される(ステップS24)。そして、計算回数が規定回数以下であるとき、コスト関数値を低減させるためのパラメータが高次元アルゴリズムにより広い探索範囲で演算される(ステップS25)。

【0130】

即ち、式(13)及び(14)により、ステップS22において演算されたコスト関数値及び拘束条件をパラメータ $w_{ij}, j, j, W_j,$ の各々で偏微分した値を求め、式(12)により次に取るべきパラメータ $w_{ij}, j, j, W_j,$ の値を演算する。そして、その演算した値を、ステップS21における演算に用いるパラメータ $w_{ij}, j, j, W_j,$ として用いる。

【0131】

このようにして、ステップS25の後、ステップS21～ステップS24が繰返し実行される。

【0132】

一方、ステップS24において、計算回数が規定回数以下でないと判定されたとき、一連の繰返しのうち、最もコスト関数値が小さくなるパラメータを、その中間ユニットの最適値と固定し、3層ニューラルネットワーク40の中間ユニット42jの個数が1個増加される(ステップS26)。そして、新しく追加された中間ユニットに対してステップS21～ステップS25が繰返し実行される。

【0133】

このように、この発明によるプログラムは、中間ユニット42jの個数が、最初、1個に設定され、その設定された1個の中間ユニット421を用いて、コスト関数値を小さくするパラメータを見つけるため、高次元アルゴリズムによりパラメータ $w_{ij}, j, j, W_j,$ を広い探索範囲で変化させて、対応するコスト関数値を次々と演算する。そして、中間ユニット42jの個数を1個にして規定回数の演算を行なっても、最も小さなコスト関数値が所定値以下にならないとき、中間ユニット42jの個数が1個増加され、同じ演算が繰返し実行される。中間ユニット42jの個数を2個に設定して規定回数の演算を行なっても、コスト関数値が所定値以下にならないとき、中間ユニット42jの個数が更に1個増加される。そして、図18に示すステップS23においてコスト関数値が所定値以下であると判定されるまで、高次元アルゴリズムによる新たなパラメータの探索と中間ユニット42jの個数の増加とが繰返し実行される。

【0134】

図17に示すステップS2が最初に実行される場合、中間層42の中間ユニット42jの個数は1個に設定されているので、中間層42は、中間ユニット421のみによって出力 Y_1 を演算する。図19は、中間ユニットの個数を1個に設定した場合の3層ニューラルネットワーク40の概念図を示す。従って、3層ニューラルネットワーク40は、図19に示す入力層41、中間層42A及び出力層43によって出力演算値 $Z_1(n)$ を演算する。

【0135】

この場合、中間ユニット421は、入力層41の入力ユニット41i ($i = 1, \dots, m_1$) からそれぞれ入力値 $x_{1,1}(n), \dots, x_{m_1,1}(n)$ ($n = 1 \sim S$)を受け、その受けた入力値 $x_{1,1}(n), \dots, x_{m_1,1}(n)$ ($n = 1 \sim S$)と結合のパラメータ $w_{1,1}, w_{2,1}, w_{m_1,1}$ と閾値 θ_1 とを式(1)に代入して内部状態 y_1 を演算する。そして、中間ユニット421は、演算した内部状態 y_1 及びパラメータ T_1 を式(2)に代入して出力 Y_1 を演算し、その演算した出力 Y_1 を出力層43の出力ユニット431へ出力する。

【0136】

出力ユニット431は、中間ユニット421から受けた出力 Y_1 と結合重み W_1 と閾値とを式(4)に代入して出力演算値 $Z_{1,1}(n)$ を演算する。つまり、パラメータ($w_{i_1,1}, \theta_1, W_1$) $_{1,1}$ を用いて出力演算値 $Z_{1,1}(n)$ が演算される。なお、パラメータ($w_{i_1,1}, \theta_1, W_1$) $_{1,1}$ 及び出力演算値 $Z_{1,1}(n)$ の添字{11}のうち、前者の{1}は中間ユニット42jの個数を表わし、後者の{1}は1回目に設定されたパラメータであることを表わす。

【0137】

コスト関数 $V_c(q)$ は、入力値 $x_{1,1}(n), \dots, x_{m_1,1}(n)$ ($n = 1 \sim S$)と出力値 $z_{1,1}(n)$ との関係を規定する近似関数 f を求める演算においては式(6)によって表わされる。従って、演算された出力演算値 $Z_{1,1}(n)$ 及び実際の出力値 $z_{1,1}(n)$ が式(6)に代入されてコスト関数 $V((w_{i_1,1}, \theta_1, W_1)_{1,1})$ のコスト関数値 $V_{1,1}$ が演算される(図18のステップS22参照)。なお、コスト関数値 $V_{1,1}$ の添字{11}の意味は、パラメータ($w_{i_1,1}, \theta_1, W_1$) $_{1,1}$ 及び出力演算値 $Z_{1,1}(n)$ の添字の意味と同じである。

【0138】

そして、コスト関数値 $V_{1,1}$ が所定値以下でないとき、コスト関数値 $V_{1,1}$ よりも小さいコスト関数値 $V_{1,2}$ を求めるためのパラメータ($w_{i_1,1}, \theta_1, W_1$) $_{1,2}$ を高次元アルゴリズムにより求める(図18のステップS25参照)。その後、パラメータ($w_{i_1,1}, \theta_1, W_1$) $_{1,2}$ を用いてS個の出力演算値 $Z_{1,2}(n)$ が演算され(図18のステップS21参照)、式(6)によりコスト関数値 $V_{1,2}$ が演算される(図18のステップS22参照)。

【0139】

そして、コスト関数値 $V_{1,2}$ が所定値以下であるか否かが判定され(図18のステップS23)、コスト関数値 $V_{1,2}$ が所定値以下でないとき、既に演算された最小のコスト関数値よりも更に小さいコスト関数値 $V_{1,h}$ ($h: h < k$ の自然数、 k : 規定回数)を求めるためのパラメータ($w_{i_1,1}, \theta_1, W_1$) $_{1,h}$ が演算される(図18のステップS25参照)。

【0140】

このように、計算回数が規定回数 k 以下であるとき、高次元アルゴリズムによりなるべく小さいコスト関数値 $V_{1,h}$ (h は、 k 以下の自然数)を見つけるためのパラメータ($w_{i_1,1}, \theta_1, W_1$) $_{1,h}$ の値が演算され、その演算されたパラメータ($w_{i_1,1}, \theta_1, W_1$) $_{1,h}$ を用いて新たなコスト関数値 $V_{1,h}$ が演算され、これが次々と繰返される。つまり、高次元アルゴリズムによりパラメータ($w_{i_1,1}, \theta_1, W_1$) $_{1,h}$ を変化させて対応するコスト関数値を監視しながら最も小さいコスト関数値と、

10

20

30

40

50

最も小さいコスト関数値に対応するパラメータとを探す。

【0141】

そして、計算回数が規定回数に達すると、中間ユニット42jの個数は、1個増加され、2個に設定される(図18のステップS26参照)。

【0142】

図20は、中間ユニットの個数を2個に設定した場合の3層ニューラルネットワーク40の概念図を示す。従って、3層ニューラルネットワーク40は、図20に示す入力層41、中間層42B及び出力層43によって出力演算値 $Z_{21}(n)$ を演算する。

【0143】

この場合、中間ユニット421は、入力層41の入力ユニット41i ($i = 1, \dots, m_1$)からそれぞれ入力値 $x_1(n), \dots, x_{m_1}(n)$ ($n = 1 \sim S$)を受け、その受けた入力値 $x_1(n), \dots, x_{m_1}(n)$ ($n = 1 \sim S$)と結合のパラメータ $w_{1,1}, w_{2,1}, w_{m_1,1}$ と閾値 θ_1 とを式(1)に代入して内部状態 y_1 を演算する。そして、中間ユニット421は、演算した内部状態 y_1 及びパラメータ T_1 を式(2)に代入して出力 Y_1 を演算し、その演算した出力 Y_1 を出力層43の出力ユニット431へ出力する。なお、パラメータ $w_{1,1}, w_{2,1}, w_{m_1,1}, \theta_1$ 及び T_j は、最も小さいコスト関数値を記録したときのパラメータ(w_{i1}, θ_1, T_1) $_{1k}$ に固定される。

【0144】

また、中間ユニット422は、入力層41の入力ユニット41i ($i = 1, \dots, m_1$)からそれぞれ入力値 $x_1(n), \dots, x_{m_1}(n)$ ($n = 1 \sim S$)を受け、その受けた入力値 $x_1(n), \dots, x_{m_1}(n)$ ($n = 1 \sim S$)と結合のパラメータ $w_{1,2}, w_{2,2}, w_{m_1,2}$ と閾値 θ_2 とを式(1)に代入して内部状態 y_2 を演算する。そして、中間ユニット422は、演算した内部状態 y_2 及びパラメータ T_2 を式(2)に代入して出力 Y_2 を演算し、その演算した出力 Y_2 を出力層43の出力ユニット431へ出力する。

【0145】

出力ユニット431は、中間ユニット421から受けた出力 Y_1 と中間ユニット422から受けた出力 Y_2 と結合重み W_1, W_2 と閾値 θ とを式(4)に代入して出力演算値 $Z_{21}(n)$ を演算する。

【0146】

そして、出力演算値 $Z_{21}(n)$ を用いてコスト関数値 V_{21} が演算され、中間ユニット421のみを用いた場合と同様の演算が繰返し実行される。中間ユニット42jの個数を2個に設定し、高次元アルゴリズムによりパラメータ(w_{i1}, θ_1, T_1) $_{2h}$ を変化させて小さいコスト関数値 V_{2h} を見つける計算を規定回数kまで繰返し実行してもコスト関数値 V_{2h} が所定値以下にならないとき、中間ユニット42jの個数が更に1個増加され、中間ユニット42jの個数を2個に設定した場合と同様の演算が繰返し実行される。

【0147】

このように、中間ユニット42jの個数がある値に設定されると、その設定された個数の中間ユニット42jを用いて演算された出力演算値 $Z(n)$ を評価するコスト関数値を小さくする最適値を求めるために、新たなパラメータ(w_{ij}, θ_j, T_j) $_{2h}$ が高次元アルゴリズムにより次々と演算される。そして、新たなパラメータの演算を規定回数まで行なってもコスト関数値が所定値以下にならないとき、中間ユニット42jの個数を1個増加してコスト関数値を小さくするように新たなパラメータ(w_{ij}, θ_j, T_j) $_{2h}$ が高次元アルゴリズムにより次々と演算される。

【0148】

従って、この発明によるプログラムは、3層ニューラルネットワーク40の中間ユニット42jの個数を1個ずつ増加させながらパラメータ(w_{ij}, θ_j, T_j) $_{2h}$ を最適化することを特徴とする。

10

20

30

40

50

【0149】

なお、上記においては、中間ユニット42jの個数は、最初、「1」個に設定されると説明したが、この発明においては、これに限らず、最初、複数に設定されてもよい。つまり、この発明においては、中間ユニット42jの個数は、最初、1個以上の初期値に設定されればよい。

【0150】

3層ニューラルネットワーク40のパラメータ(w_{ij} , j , j , W_j ,)は、中間ユニット42jの個数が増加するごとに($m1+3$)個ずつ増加する。例えば、中間ユニット42jの個数が1個から2個に増加したとき、全体のパラメータは、(w_{i_1} , 1 , 1 , W_1 ,)から(w_{i_1} , 1 , 1 , W_1 ; w_{i_2} , 2 , 2 , W_2 ,)へ($m1+3$)個増加する。従って、中間ユニット42jの個数を1個ずつ増加してパラメータ(w_{ij} , j , j , W_j ,)を最適化することは、パラメータの数を所定数ずつ増加してパラメータ(w_{ij} , j , j , W_j ,)を最適化することに相当する。そして、パラメータの数が増加される前の所定数のパラメータのうち、パラメータ(w_{i_1} , 1 , 1 , W_1)の値を固定し、増加されたパラメータ(w_{i_2} , 2 , 2 , W_2)とパラメータのうち、超球面識別タイプの演算を行なう中間ユニット422のパラメータ(w_{i_2} , 2 , 2)を広い探索範囲で変化させて出力演算値 $Z_1(n)$ を演算する。この場合、図17に示すステップS2では、パラメータの個数を所定数の初期値に設定してパラメータの最適化を行なう。

10

【0151】

従って、この発明によるプログラムは、3層ニューラルネットワークのパラメータの数を所定数ずつ増加してパラメータを最適化することを特徴とする。

20

【0152】

近似関数fを求める演算を行なう場合における3層ニューラルネットワーク40の効果について説明する。即ち、3層ニューラルネットワーク40の中間ユニット42jにおいて超球面識別タイプの演算を行なう場合の効果について説明する。

【0153】

図21は、二入力-出力の3つのテスト関数の等高線を示す。図21の(a)は、テスト関数zを $z = -0.76x_1 + 0.19x_2 + 0.78$ としたときの等高線を示す。図21の(b)は、テスト関数zを $z = \sin(x_1) \cdot \sin(x_2)$ としたときの等高線を示す。図21の(c)は、テスト関数zを $z = 0.5 \exp\{-5(x_1 - 0.2)^2 - 5(x_2 - 0.2)^2\} + 0.9 \exp\{-5(x_1 - 0.8)^2 - 10(x_2 - 0.6)^2\}$ としたときの等高線を示す。図21の(a), (b), (c)において、横軸及び縦軸は入力変数 x_1 , x_2 を表わす。

30

【0154】

また、各学習条件は、次に示すとおりである。

$$(i) \quad z = -0.76x_1 + 0.19x_2 + 0.78$$

サンプル数 $N=50$

所定値 $\varepsilon \leq 0.003$

$$(ii) \quad z = \sin(\pi x_1) \cdot \sin(\pi x_2)$$

サンプル数 $N=100$

所定値 $\varepsilon \leq 0.05$

$$(iii) \quad z = 0.5 \exp\{-5(x_1 - 0.2)^2 - 5(x_2 - 0.2)^2\} \\ + 0.9 \exp\{-5(x_1 - 0.8)^2 - 10(x_2 - 0.6)^2\}$$

10

サンプル数 $N=100$

所定値 $\varepsilon \leq 0.03$

問題 (i)、(ii) 及び (iii) の各々において、超球面識別タイプの演算及び超平面識別タイプの演算を行ない、その演算結果の比較を行なった。なお、両方の演算において、パラメータの最適化は高次元アルゴリズムを用いて行なわれた。

【0155】

問題 (i) の場合、次の結果が得られた。

20

< 1 > 超平面識別ネットワーク

獲得した中間ユニットの個数 = 1

平均学習回数 11, 507回

< 2 > 超球面識別ネットワーク

獲得した中間ユニットの個数 = 1

平均学習回数 64, 103回

図 21 の (a) に示すように、このテスト関数の等高線は線型であり、超平面識別ネットワークに明らかに有利な問題である。しかし、超球面識別ネットワークにおいても、1個の中間ユニットで比較的厳しい要求精度 (0.003) を達成していることから、パラメータの探索範囲を広くすることにより、局所的特徴のみならず、大局的特徴も反映させることができ、写像能力が強化されることが解る。

30

【0156】

一方、超球面識別ネットワークは、平均学習回数が超平面識別ネットワークの6倍程度多い。これは、パラメータの探索範囲を広くしたことが原因しているが、明らかに不利なケースにも拘わらず、高々6倍程度の増加に留まっている。

【0157】

次に、問題 (ii) の結果を示す。

< 1 > 超平面識別ネットワーク

獲得した中間ユニットの個数 = 8

平均学習回数 393, 829回

< 2 > 超球面識別ネットワーク

獲得した中間ユニットの個数 = 1

平均学習回数 12, 045回

図 21 の (b) に示すように、このテスト関数の等高線は、閉じた曲線であり、超球面識別ネットワークに明らかに有利な場合である。超平面識別ネットワークについては、要求精度を 0.03 に上げると学習の収束が急速に困難になったため、 0.05 とした。必要な中間ユニットの個数は、超平面識別ネットワークが8個、超球面識別ネットワークが1個という結果であり、これから、写像能力は、超球面識別ネットワークの方が各段に優れていることが解る。

40

【0158】

50

平均学習回数に関しても、8倍の中間ユニット数(パラメータ数)を要する超平面識別ネットワークでは約33倍多く、超平面識別ネットワークは、局所的特徴の表現には効率が悪いことが解る。

【0159】

最後に、問題(iii)の結果を示す。

<1>超平面識別ネットワーク

獲得した中間ユニットの個数 = 6

平均学習回数 264, 706回

<2>超球面識別ネットワーク

獲得した中間ユニットの個数 = 2

平均学習回数 2, 665回

図21の(c)に示すように、このテスト関数は、非対称な閉じた等高線を持ち、現実の応用問題にありそうな一例である。超平面識別ネットワークは、要求精度 0.01を達成するのが困難であったため、0.03とした。

【0160】

中間ユニットの個数は、超球面識別ネットワークが2個であり、超平面識別ネットワークが6個である。その結果、この場合も超球面識別ネットワークの方が写像能力が高いことが解る。平均学習回数も、超平面識別ネットワークは、超球面識別ネットワークの100倍近く多い。

【0161】

このように、パラメータの探索範囲を適切に拡大した超球面識別ネットワークを高次元アルゴリズムで学習する方法は、一般に、超平面識別ネットワークの場合よりも少ないパラメータ数で要求精度を達成できる高い写像能力を持ち、かつ、学習性能も良好であることが確認された。

【0162】

図17に示すステップS1, S2及び図18に示すステップS21~S26を備えるプログラムは、図22に示すパーソナルコンピュータによって実行される。図22は、パーソナルコンピュータの概略ブロック図である。パーソナルコンピュータ90は、データバスBSと、CPU91と、RAM(Random Access Memory)92と、ROM(Read Only Memory)93と、シリアルインタフェース94と、端子95と、CD-ROMドライブ96と、ディスプレイ97と、キーボード98とを備える。

【0163】

CPU91は、ROM93に格納されたプログラムをデータバスBSを介して読出す。また、CPU91は、シリアルインタフェース94、端子95及びインターネット網を介して取得したプログラム、またはCD(Compact Disk)99からCD-ROMドライブ96を介して読出したプログラムをROM93に格納する。更に、CPU91は、キーボード98から入力されたユーザからの指示を受付ける。

【0164】

RAM92は、CPU91が上述した近似関数fを求める演算を行なう際のワークメモリである。ROM93は、プログラム等を格納する。シリアルインタフェース94は、データバスBSと端子95との間でデータのやり取りを行なう。

【0165】

端子95は、ケーブルによってパーソナルコンピュータ90をインターネットに接続するためのインタフェース(図示せず)に接続するための端子である。CD-ROMドライブ96は、CD99に記録されたプログラムを読出す。ディスプレイ97は、各種の情報を視覚情報としてユーザに与える。キーボード98は、ユーザからの指示を受付ける。

【0166】

CPU91は、キーボード98を介して入力されたユーザの指示に応じて、ROM93に格納されたプログラムを読出し、その読出したプログラムを実行する。そして、CPU9

10

20

30

40

50

1 は、図 17 及び図 18 に示すフローチャートに従って近似関数 f を求める演算を行ない、最適化されたパラメータをディスプレイ 97 に表示する。

【0167】

この発明によるプログラムは、CD 99 から CD-ROM ドライブ 96 を介して CPU 91 によって読み込まれて ROM 93 に格納され、またはシリアルインタフェース 94、端子 95 及びインターネットを介して取得されて ROM 93 に格納される。

【0168】

このように、ユーザは、この発明によるプログラムをパーソナルコンピュータ 90 により実行して近似関数 f を求める演算を行なうことができる。

【0169】

上述したように、この発明によるプログラムが入力値 $x_1(n), \dots, x_{m_1}(n)$ ($n = 1 \sim S$) と出力値 $z_1(n)$ との関係を規定する近似関数 f を求める演算を行なう場合、サンプルデータ 30 の元になる集合 10 に含まれる入力値 ($x_1(1), \dots, x_{m_1}(1)$)、($x_1(2), \dots, x_{m_1}(2)$)、 \dots 、($x_1(S), \dots, x_{m_1}(S)$)、 \dots 、($x_1(M), \dots, x_{m_1}(M)$) 及び集合 20 に含まれる出力値 ($z(1), z(2), \dots, z(S), \dots, z(M)$) は、例えば、量子井戸構造における井戸層の幅 d_{w_1} 、井戸層の電子密度 n_{w_1} 、バリア層の高さ W_{d_2} 及びバリア層の幅 d_b 等から成る入力値と、量子井戸構造における粒子（電子及び正孔）のエネルギー準位 E_{w_1} から成る出力値として取得される。

【0170】

そこで、井戸層の幅 d_{w_1} 、井戸層の電子密度 n_{w_1} 、バリア層の高さ W_{d_2} 及びバリア層の幅 d_b 等から成る入力値 ($x_1(1), \dots, x_{m_1}(1)$)、($x_1(2), \dots, x_{m_1}(2)$)、 \dots 、($x_1(S), \dots, x_{m_1}(S)$)、 \dots 、($x_1(M), \dots, x_{m_1}(M)$) と、エネルギー準位 E_{w_1} から成る出力値 ($z(1), z(2), \dots, z(S), \dots, z(M)$) とを取得する方法について説明する。

【0171】

[サンプルデータの取得方法]

エネルギー準位 E_{w_1} は、量子井戸構造における粒子の自己相互作用を考慮して演算される。

【0172】

図 23 は、量子準位演算プログラムが演算の対象とする量子井戸の概念図である。縦軸はエネルギーを示し、横軸は位置を示す。量子井戸 100 は、バリア層 101、102 と井戸層 103 とから成る。井戸層 103 に閉じ込められた電子はエネルギー準位 104 を形成する。

【0173】

この発明においては、粒子の相互作用を含まないシュレディンガー方程式を解いて 1 つの量子井戸 100 における波動関数 ψ を求める。そして、求めた波動関数 ψ から出発して、変分法の原理に基づき系全体のエネルギーを最小化する波動関数 ψ を求める。この場合、系全体のエネルギーは、粒子の相互作用を非線形項として含むハミルトニアン H の期待値 $\langle H \rangle$ を、与えられた波動関数 ψ に対して計算したものと定義される。その後、量子井戸 100 の一方のバリア層 101 の端から他方のバリア層 102 の端までを複数のポイント $x_1 \sim x_N$ (N は自然数) に分割し、その分割した各ポイント x_i ($1 \leq i \leq N$) に対応して波動関数 ψ を N 個の波動関数 $\psi_1, \dots, \psi_i, \dots, \psi_N$ に離散化する。この場合、ポイント $x_1 \sim x_N$ の各々と隣接するポイントとの間の距離は全て等しいように位置 $x_1 \sim x_N$ が決定される。

【0174】

波動関数 ψ を波動関数 $\psi_1, \dots, \psi_i, \dots, \psi_N$ へ離散化すると、量子井戸 100 の系全体のエネルギーが最小となるように、波動関数 $\psi_1, \dots, \psi_i, \dots, \psi_N$ の各々を演算する。そして、演算した波動関数 $\psi_1, \dots, \psi_i, \dots, \psi_N$ を用

10

20

30

40

50

いて系の全体エネルギーを求める。

【0175】

このように、粒子の相互作用を含まないシュレディンガー方程式を解いて求めた波動関数を N 個に離散化し、系全体のエネルギーが最小となるように、離散化した波動関数 $\psi_1, \dots, \psi_i, \dots, \psi_N$ の各々を演算する。

【0176】

以下、粒子の相互作用を含まないシュレディンガー方程式を解いて求めた波動関数を変分法の原理に適用して系全体のエネルギーを演算する際に、ハミルトニアンに取り入れられる非線形項として粒子のクーロン相互作用を用いた場合について説明する。

【0177】

図24は、量子準位演算プログラムを構成する各ステップを示すフローチャートである。量子準位演算プログラムが実行されると、量子井戸100に閉じ込められた電子の自己相互作用のない場合の波動関数が演算される(ステップS100)。即ち、運動エネルギーの項と外部電界によるポテンシャル項とから成るハミルトニアンを用いて電子の初期の波動関数が演算される。この演算は、転送行列法、S行列法、及び狙い撃ち法のいずれかを用いて行われる。

【0178】

そして、最小化される系全体のエネルギーを示す式は式(16)により与えられる。

【0179】

【数16】

$$S[\psi] = \int dx \psi^\dagger(x) \left[-\frac{\hbar^2}{2m} \Delta + V(x) \right] \psi(x) + \frac{e^2}{2} \int dx \int dy \frac{[|\psi(x)|^2 - d(x)][|\psi(y)|^2 - d(y)]}{\epsilon(x)|x-y|} \quad \dots(16)$$

【0180】

式(16)の右辺の第2項は、自己相互作用を含むクーロン相互作用項である。そして、 $\epsilon(x)$ は、位置 x における誘電定数であり、 $-e$ 及び m は、それぞれ、電子の電荷及び質量である。また、 $d(y)$ は、位置 y における固定ドナーの体積密度であり、固定ドナーは $+e$ の電荷を有するものと仮定している。即ち、コンピュータによる計算のために、イオン化されたドナーは、自由電子により $+e$ の電荷を運ぶものと仮定した。

【0181】

式(16)は、量子井戸100の系全体のエネルギーを表わすが、式(16)を用いて演算することは困難であるので、空間を離散化するとともに系全体に存在する粒子数で規格化することとした。即ち、ステップS100で求めた波動関数を N 個の波動関数 $\psi_1, \dots, \psi_i, \dots, \psi_N$ に離散化するとともに系全体の粒子数で規格化することにより、式(16)から式(17)が得られる。

【0182】

【数17】

10

20

30

40

$$\begin{aligned}
S[\{\psi_i\}] &= \frac{\psi^\dagger (H_0 + H') \psi}{N_t} \Delta x - \frac{e}{2} \sum_{i=1}^N V_i' \frac{d_i}{N_c} \Delta x \\
&= -\frac{\hbar^2}{2m(\Delta x)^2} \sum_{i=1}^N \frac{2\psi_i^2 - \psi_{i-1}\psi_i - \psi_i\psi_{i+1}}{N_t} \Delta x \\
&\quad + e \sum_{i=1}^N V_i \frac{\psi_i^2}{N_t} \Delta x + \frac{e}{2} \sum_{i=1}^N V_i' \left(\frac{\psi_i^2}{N_t} - \frac{d_i}{N_c} \right) \Delta x \quad \dots(17)
\end{aligned}$$

10

【0183】

式(17)において、 N_t は規格化因子であり、 N_e は電子数である。また、 V_i' は、クーロンポテンシャルである。

【0184】

式(17)は、系に存在する1つの電子あたりのエネルギーを表わし、以下では、最適化問題の用語に合わせて「コスト関数」と呼ばれる。そして、式(17)を離散化したN個の波動関数 $\psi_1, \dots, \psi_i, \dots, \psi_N$ の各々によって偏微分してN個の導関数を演算する。即ち、式(18)が得られる。

20

【0185】

【数18】

$$\frac{\partial \langle H' \rangle}{\partial \psi_t} = 2V_i' \frac{\psi_i}{N} \Delta v - 2E_{\text{int}} \frac{\psi_i}{N} \Delta v. \quad \dots(18)$$

$$\begin{aligned}
\text{但し、} V_i' &= -\frac{q}{\epsilon} \sum_{j=1}^i \sum_{k=1}^N \left(N_e \frac{|\psi_k|^2}{N_t} - d_k \right) \\
&\quad \times \text{sgn}(j-k)(\Delta x)^2,
\end{aligned}$$

30

$$E_{\text{int}} = \frac{q}{2} \sum_{i=1}^N V_i' \left(\frac{|\psi_i|^2}{N_t} - \frac{d_i}{N_c} \right) \Delta x$$

【0186】

式(18)において、 E_{int} は1粒子当たりの非線形の相互エネルギーを意味する。そして、式(16)から式(17)及び式(18)を求めることは、コスト関数及びコスト関数の導関数を計算すること(ステップS102)に相当する。

40

【0187】

なお、式(18)は、式(17)の両辺を波動関数 $\psi_1, \dots, \psi_i, \dots, \psi_N$ の各々によって偏微分したものではなく、式(17)の自己相互作用を考慮したハミルトニアン H' を波動関数 $\psi_1, \dots, \psi_i, \dots, \psi_N$ の各々によって偏微分したものになっているが、その理由は次の理由による。

【0188】

量子準位演算プログラムは、自己相互作用を考慮して量子井戸100の井戸層103に閉じ込められた電子の系全体のエネルギーが最小になるように波動関数を計算することを目的とするため、式(17)の全体を波動関数 $\psi_1, \dots, \psi_i, \dots, \psi_N$ の各々に

50

よって偏微分するのではなく、自己相互作用を考慮したハミルトニアン H' を波動関数 $\psi_1, \dots, \psi_i, \dots, \psi_N$ の各々によって偏微分することにより自己相互作用の影響を最大限に反映して演算することにしたものである。

【0189】

従って、ステップ S102 においては、コスト関数（式（17））に離散化した波動関数 $\psi_1, \dots, \psi_i, \dots, \psi_N$ を代入して系全体のエネルギー $S\{\psi_i\}$ を演算し、コスト関数の導関数が自己相互作用を考慮したハミルトニアン H' を波動関数 $\psi_1, \dots, \psi_i, \dots, \psi_N$ の各々によって偏微分することにより演算される（式（18））。

【0190】

その後、ステップ S102 において演算した N 個の導関数及び次式（19）を用いて新しい波動関数を演算する（ステップ S104）。

【0191】

【数19】

$$\psi_i^{\text{new}} = \psi_i^{\text{old}} + \eta \left. \frac{\partial S[\{\psi_i\}]}{\partial \psi_i} \right|_{\psi = \psi^{\text{old}}} \quad \dots(19)$$

【0192】

式（19）において、 η は、量子準位演算プログラムによる最小エネルギーの演算が収束するようにするためのスケールリングファクターである。また、式（19）の右辺第2項は、自己相互作用を考慮したハミルトニアン H' を波動関数 ψ_i の各成分によって偏微分して式（18）を演算し、その演算した式（18）に波動関数 ψ_i^{old} を代入して演算される。

【0193】

式（19）による新しい波動関数 ψ_i^{new} は、既に演算された波動関数 ψ_i^{old} に自己相互作用による変化分（式（19）の右辺第2項）を加算したものである。従って、新しい波動関数 ψ_i^{new} は、自己相互作用の変化分を反映して演算される。そして、新しい波動関数 ψ_i^{new} は、離散化した N 個の波動関数 $\psi_1, \dots, \psi_i, \dots, \psi_N$ の各々に対して演算される。

【0194】

式（19）により新しい波動関数 ψ_i^{new} が演算されると、その新しい波動関数 ψ_i^{new} を式（17）の波動関数 ψ_i に代入して、新しい波動関数によるコスト関数（式（17））及びその導関数（式（18））が演算される（ステップ S106）。

【0195】

そして、コスト関数が増加するか否かにより、またはステップ S106 において演算した N 個の導関数の全てが零であるか否かを判定することにより、系全体のエネルギーが収束するか否かが判定される（ステップ S108）。

【0196】

ステップ S108 において、コスト関数が増加しないとき、又は N 個の導関数の全てが零でないとき、系全体のエネルギーは収束しないと判定され、ステップ S104、S106、S108 が繰返し実行される。これは、コスト関数が増加しないとき、コスト関数の増加分が零かコスト関数が減少していることを示し、系全体のエネルギーが更に減少する可能性があるからであり、コスト関数の導関数が零でないとき、コスト関数が増加していることを示し、この場合も系全体のエネルギーが更に減少する可能性があるからである。

【0197】

一方、ステップ S108 において、コスト関数が増加したとき、または N 個の導関数の全てが零であるとき、系全体のエネルギーが収束したと判定される。そして、1 ステップ前の波動関数を出力する（ステップ S110）。

【0198】

10

20

30

40

50

例えば、ステップ S 1 0 4 , S 1 0 6 , S 1 0 8 を 5 回繰返して実行し、5 回目にステップ S 1 0 8 においてコスト関数が増加、又は N 個の導関数の全てが零になったとすると、4 回目にステップ S 1 0 4 において式 (1 9) を用いて演算した N 個の成分をステップ S 1 1 0 において出力する。5 回目にステップ S 1 0 8 においてコスト関数が増加、又は N 個の導関数が全て零になったということは、ステップ S 1 0 4 において 4 回目に演算した波動関数を用いて、ステップ S 1 0 6 において演算したコスト関数が最小になったことを意味するからである。

【 0 1 9 9 】

ステップ S 1 1 0 において、系全体のエネルギーを最小にする N 個の成分が決定されると、その決定された N 個の成分から成る波動関数 (最終的な波動関数) と、自己相互作用を考慮したハミルトニアンである式 (2 0) を用いて系全体のエネルギーを演算し (ステップ S 1 1 2)、全体の演算動作が終了する。

10

【 0 2 0 0 】

【 数 2 0 】

$$\hat{H} = -\frac{\hbar^2}{2m}\Delta + V(x) + \frac{e^2}{\varepsilon(x)} \int dy \frac{|\psi(y)|^2 - d(y)}{|x-y|} \quad \dots (20)$$

【 0 2 0 1 】

上述した各ステップのうち、ステップ S 1 0 2 , S 1 0 4 , S 1 0 6 , S 1 0 8 , S 1 1 0 は 1 つの粒子に対する動作を表わす。従って、量子準位演算プログラムは、系全体のエネルギーを最小とするように N 個の成分を決定するとき、1 つの粒子に着目し、その 1 つの粒子に対する波動関数において自己相互作用による影響が最も小さくなるように N 個の成分を決定することを第 1 の特徴とする。そして、1 つの粒子に対する波動関数が決定されると、その決定された波動関数を系全体の粒子に適用して系全体のエネルギーを演算することを第 2 の特徴とする。

20

【 0 2 0 2 】

その結果、量子井戸 1 0 0 のバリア層 1 0 1 , 1 0 2 又は井戸層 1 0 3 へのドーピング量の増加に起因して井戸層 1 0 3 に閉じ込められる電子数が増加し、電子の相互作用による影響が大きくなって系全体のエネルギーが収束するように演算できる。

30

【 0 2 0 3 】

以下、量子準位演算プログラムを用いて演算した例について具体的に説明する。

【 0 2 0 4 】

図 2 5 は、量子準位演算プログラムが演算の対象とする量子井戸の具体例である。縦軸はエネルギーを示し、横軸はバリア層 1 0 1 , 1 0 2 及び井戸層 1 0 3 の厚み方向の距離 z を示す。2 つのバリア層 1 0 1 , 1 0 2 の各々は、3 5 モノレイヤー (= 1 0 n m) の A l _{0.2} G a _{0.8} A s から成り、井戸層 1 0 3 は、3 5 モノレイヤー (= 約 1 0 n m) の G a A s から成る。そして、伝導帯側のバンドの不連続値 E c (= E c 1 - E c 2) は 1 6 7 m e V である。

40

【 0 2 0 5 】

図 2 5 に示す系において井戸層 1 0 3 である G a A s のボトムエッジから 2 8 . 2 5 7 9 5 m e V の位置に第 1 準位 E 1 が形成され、その第 1 準位 E 1 を占める電子の初期の波動関数が波動関数 である。

【 0 2 0 6 】

以下、井戸層 1 0 3 にドナーをドーピングした場合 (「井戸ドーピング」と言う。)、及びバリア層 1 0 1 , 1 0 2 にドナーをドーピングした場合 (「バリアドーピング」と言う。) について説明する。そして、ドーピングされたドナーは全て活性化され、自由電子の総数は、ドナーの総数に等しいと仮定する。

【 0 2 0 7 】

50

また、バリア層 101, 102 及び井戸層 103 の各モノレイヤーを 10 ポイントに分割する。即ち、波動関数 $\psi_1, \dots, \psi_i, \dots, \psi_{105}$ に離散化する。

【0208】

更に、バリア層 101, 102 及び井戸層 103 から成る量子井戸 100 の両端では波動関数は零であると近似している。

【0209】

更に、式 (19) のスケーリングファクター α を -1.15×10^7 に固定して計算した。このスケーリングファクター α の値は、あくまで 1 モノレイヤー当たり 10 ポイントに分割した場合の値であり、他の分割数の場合には、他の値が用いられる。例えば、1 モノレイヤー当たり 20 ポイントに分割したのであれば、 $\alpha = -2.30 \times 10^7$ が用いられる。また、スケーリングファクター α の値としては、これらの値以外の値も想定される。

10

【0210】

図 26 は、井戸ドーピングを行なった場合の計算結果をドーピング量に対して示したものである。図 26 の (a) ~ (d) の各々において、縦軸はエネルギーを示し、横軸はバリア層 101, 102 及び井戸層 103 の厚み方向の距離 z を示す。図 26 の (a) は、ドーピング量が $1.0 \times 10^{18} \text{ cm}^{-3}$ の場合を示し、図 26 の (b) は、ドーピング量が $5.0 \times 10^{18} \text{ cm}^{-3}$ の場合を示し、図 26 の (c) は、ドーピング量が $8.0 \times 10^{18} \text{ cm}^{-3}$ の場合を示し、図 26 の (d) は、ドーピング量が $1.0 \times 10^{19} \text{ cm}^{-3}$ の場合を示す。また、図 26 の (a) ~ (d) において波動関数 ψ_0 は初期の波動関数であり、記号 $E_{w1} \sim E_{w4}$ は基底状態のエネルギー値を示し、記号 $w_1 \sim w_4$ は上述した量子準位演算プログラムを用いて計算された波動関数を示す。

20

【0211】

量子準位演算プログラムを用いて 4 つのドーピング量に対する計算を行なった結果、その計算時間は 30 秒以下と非常に短かった。

【0212】

井戸ドーピングにおいては、電子及びドナーは、両方とも井戸層 103 に存在するため、電子とドナーとの間で電荷の打消しが起こり、これが量子井戸の外側におけるバンドの曲がりを抑制する。そして、これは、高いドーピング量においても当てはまる。

【0213】

また、計算された波動関数のひずみは小さいことが解かった。更に、基底状態のエネルギー $E_{w1} \sim E_{w4}$ の変化は、イオン化された不純物 (ドナー) 及び電子間の相互作用による電子のポテンシャルに起因するが、この変化は、そう大きくないのに対し、井戸層 103 におけるバンドの曲がりが増加に伴い大きくなる。

30

【0214】

表 1 は、井戸ドーピングにおいて、量子準位演算プログラムを用いて計算した基底状態のエネルギー値を従来の S - P 法を用いた計算結果と比較して示す。

【0215】

【表 1】

密度 ($\times 10^{18}/\text{cm}^3$)	S-P method (meV)	present method (meV)
1.0	30.32965	30.32967
1.2	30.69431	30.69435
1.4	31.04355	31.04359
1.6	31.37784	31.37788
1.7	31.53953	31.53956
2.0	-	32.00348
4.0	-	34.38380
6.0	-	35.78042
8.0	-	36.46603
10.0	-	36.63213

10

【0216】

表1から明らかなように、ドーピング量が $1.0 \times 10^{18} \sim 1.7 \times 10^{18} \text{ cm}^{-3}$ の範囲においては、量子準位演算プログラムを用いた計算結果は、従来のS-P法による計算結果と良い一致を示し、その差は、殆ど無視できる程度の $1.3 \times 10^{-4} \%$ 以下である。

20

【0217】

また、従来のS-P法は、 $1.8 \times 10^{18} \text{ cm}^{-3}$ 以上のドーピング量に対して発散するのに対し、量子準位演算プログラムを用いた計算結果は、少なくとも $1.0 \times 10^{19} \text{ cm}^{-3}$ のドーピング量までは確実に収束することがわかった(数値が得られていることは収束していることを示す。以下同じ。)

【0218】

このように、量子準位演算プログラムを用いた計算方法は、井戸層103へドーピングした場合において、ドーピング量の低い領域では従来のS-P法による計算結果と良い一致を示し、高いドーピング量の範囲まで収束した計算結果を示す。その結果、量子準位演算プログラムを用いることにより、高いドーピング量の範囲まで基底状態のエネルギー値を得ることができる。

30

【0219】

図27は、バリアドーピングを行なった場合の計算結果をドーピング量に対して示したものである。図27の(a)~(d)において、縦軸はエネルギーを示し、横軸はバリア層101, 102及び井戸層103の厚み方向の距離zを示す。また、図27の(a)~(d)において、記号 $0, w1 \sim w4, Ew1 \sim Ew4$ は、図26の(a)~(d)における意味と同じである。更に、図27の(a)は、ドーピング量が $1.0 \times 10^{17} \text{ cm}^{-3}$ の場合を示し、図27の(b)は、ドーピング量が $3.0 \times 10^{17} \text{ cm}^{-3}$ の場合を示し、図27の(c)は、ドーピング量が $5.0 \times 10^{17} \text{ cm}^{-3}$ の場合を示し、図27の(d)は、ドーピング量が $7.0 \times 10^{17} \text{ cm}^{-3}$ の場合を示す。

40

【0220】

この場合も、計算時間は30秒以下であった。バリア層101, 102へのドーピング量が増加するに伴い波動関数 $w1 \sim w4$ のひずみが増加する。これは、次の理由による。バリアドーピングにおいては、閉じ込められた電子は井戸層103に存在し、ドナーはバリア層101, 102に存在するため、井戸層103の電子は、電子間の相互作用による反発力とバリア層101, 102に存在するドナーからの引力とにより両側に存在するバリア層101, 102の方へ拡がる。そして、バリア層101, 102へのドーピング量が増加するに従ってバリア層101, 102に存在するドナーからの引力が増加するので、井戸層103における電子の拡がりが増加する。その結果、バリア層101

50

、102へのドーピング量が増加するに従って波動関数 $w_1 \sim w_4$ のひずみが増加する。

【0221】

また、バリアドーピングの場合、井戸ドーピングの場合に比べてバンドの曲がりが大きくなる。これは、次の理由による。イオン化されたドナーによる電子のポテンシャルは、井戸層103のボトムにおけるバンド端のポテンシャルを増加させる。一方、電子とドナーとは空間的に分離されているので、イオン化されたドナーからの電気的な影響を除去しようとする電子の働きは低下する。その結果、井戸層103からバリア層101、102へ電界が及び、バンドの曲がりが大きくなる。

【0222】

表2は、バリアドーピングにおいて、量子準位演算プログラムを用いて計算した基底状態のエネルギー値を従来のS-P法を用いた計算結果と比較して示す。

【0223】

【表2】

密度 ($\times 10^{17}/\text{cm}^3$)	S-P method (meV)	present method (meV)
1.0	46.85800	46.85793
2.0	65.20937	65.20934
3.0	83.28149	83.28160
4.0	101.0339	101.0337
5.0	118.4102	118.4102
6.0	-	135.3280
7.0	-	151.6623

【0224】

表2から明らかなように、ドーピング量が $1.0 \times 10^{17} \sim 5.0 \times 10^{17} \text{ cm}^{-3}$ の範囲においては、量子準位演算プログラムを用いた計算結果は、従来のS-P法による計算結果と良い一致を示す。そして、量子準位演算プログラムを用いた計算方法は、従来のS-P法では収束しない $6.0 \times 10^{17} \text{ cm}^{-3}$ 、 $7.0 \times 10^{17} \text{ cm}^{-3}$ のドーピング量において収束する。

【0225】

このように、量子準位演算プログラムを用いた計算方法は、バリア層101、102へドーピングした場合において、ドーピング量の低い領域では従来のS-P法による計算結果と良い一致を示し、高いドーピング量の範囲まで収束した計算結果を示す。その結果、量子準位演算プログラムを用いることにより、高いドーピング量の範囲まで基底状態のエネルギー値を得ることができる。

【0226】

図28は、ドーピング量が $2.0 \times 10^{18} \text{ cm}^{-3}$ の井戸ドーピングを行なった場合において、量子準位演算プログラムを用いた計算結果を従来のS-P法による計算結果と比較して示す。図28の(a)、(b)において、縦軸はエネルギーであり、横軸はバリア層101、102及び井戸層103の厚み方向の距離 z である。図28の(a)は、従来のS-P法による計算結果を示し、図28の(b)は、量子準位演算プログラムを用いた計算結果を示す。

【0227】

記号 c 、 $i v t$ は、波動関数を示し、記号 $V c$ 、 $V i v t$ はバンド端のポテンシャルを示す。

【0228】

従来のS-P法は、シュレディンガー方程式の解法とポアソン方程式の解法とを繰返す過

10

20

30

40

50

程において、小さなバランスのずれや計算のエラーが増幅され、その結果、計算ステップの進行に対して波動関数 c が振動する。

【0229】

これに対して、量子準位演算プログラムを用いた計算方法は、図24に示すステップS104, S106, S108の繰返しにおいて、波動関数が本当の解に近づくに従って式(19)による補正が小さくなり、その結果、波動関数 $i v t$ は収束する。

【0230】

上記においては、1モノレイヤー当たりの分割数が10ポイントの場合について説明したが、この分割数を変化させた場合について図29に示す。縦軸は基底状態のエネルギーを表わし、横軸は1モノレイヤー当たりの分割数を表わす。また、曲線105(実線で示される)は、量子準位演算プログラムを用いた計算結果であり、曲線106(点線で示される)は、従来のS-P法による計算結果である。なお、ドーピングは、井戸層103へ行なわれ、ドーピング量は $1.0 \times 10^{18} \text{ cm}^{-3}$ である。

10

【0231】

両方の方法において、分割数が増加するに従って基底状態のエネルギー値が小さくなり、空間をより小さく分割した方が正確な波動関数が得られることが解った。また、空間の分割数に関しては、量子準位演算プログラムを用いた計算方法は従来のS-P法による計算方法と大きな差がないことが解かった。

【0232】

上述した井戸ドーピング及びバリアドーピングにおける計算においては、初期状態を狙い撃ち法(Shooting Method)により求めた初期状態を用いた。

20

【0233】

表3は、井戸ドーピング及びバリアドーピングにおいて、数学的に厳密な波動関数を用いた場合の、量子準位演算プログラムを用いた計算結果を示す。初期状態として採用した厳密な波動関数に対応するエネルギー値は、 28.27683 meV である。

【0234】

【表3】

井戸層ドーピング		バリア層ドーピング	
密度 ($\times 10^{18}/\text{cm}^3$)	エネルギー (meV)	密度 ($\times 10^{17}/\text{cm}^3$)	エネルギー (meV)
1.0	30.41680	1.0	46.91780
2.0	32.14000	2.0	65.34559
3.0	33.51818	3.0	83.54844
4.0	34.60840	4.0	101.5173
6.0	36.09270	5.0	119.2431
8.0	36.85680	6.0	136.7163
10.0	37.08482	7.0	153.9277

30

40

【0235】

その結果、初期状態として採用した厳密な波動関数に対応するエネルギー値として 28.27683 meV のエネルギー値を用いた場合の方が、各ドーピング量に対するエネルギー値は大きくなることが解かった(表1及び表2参照)。しかし、その差は、殆ど1.5%であり、実際の半導体材料における量子準位の見積もりにおいては許容される範囲である。従って、量子準位演算プログラムにおいては、狙い撃ち法(Shooting Method)により初期状態のエネルギー値を演算しても、特に、問題はないと考えられる。

【0236】

50

このようにして、量子準位演算プログラムを用いて量子井戸構造における井戸層の幅 d_w 、井戸層の電子密度 n_w 、バリア層の高さ W_d 及びバリア層の幅 d_b 等から成る入力値と、エネルギー準位 E_w から成る出力値とが取得される。

【0237】

そして、集合10に含まれる入力値 $(x_1(1), \dots, x_{m_1}(1))$ 、 $(x_1(2), \dots, x_{m_1}(2))$ 、 \dots 、 $(x_1(S), \dots, x_{m_1}(S))$ 、 \dots 、 $(x_1(M), \dots, x_{m_1}(M))$ 及び集合20に含まれる出力値 $(z(1), z(2), \dots, z(S), \dots, z(M))$ が取得されると、出力値 $(z(1), z(2), \dots, z(S), \dots, z(M))$ が上述したなだらかな曲面上に存在するか否かを判定し、出力値 $(z(1), z(2), \dots, z(S), \dots, z(M))$ がなだらかな曲面上に存在する場合、集合10に含まれる入力値 $(x_1(1), \dots, x_{m_1}(1))$ 、 $(x_1(2), \dots, x_{m_1}(2))$ 、 \dots 、 $(x_1(S), \dots, x_{m_1}(S))$ 、 \dots 、 $(x_1(M), \dots, x_{m_1}(M))$ から入力値 $(x_1(1), \dots, x_{m_1}(1))$ 、 $(x_1(2), \dots, x_{m_1}(2))$ 、 \dots 、 $(x_1(S), \dots, x_{m_1}(S))$ を抽出し、集合20に含まれる出力値 $(z(1), z(2), \dots, z(S), \dots, z(M))$ から出力値 $(z(1), z(2), \dots, z(S), \dots, z(M))$ を抽出してサンプルデータ30を準備する。

【0238】

図22に示すパーソナルコンピュータ90が量子準位演算プログラムを用いてエネルギー準位 E_w から成る出力値を演算する場合、CPU91は、量子井戸構造における井戸層の幅 d_w 、井戸層の電子密度 n_w 、バリア層の高さ W_d 及びバリア層の幅 d_b 等から成る入力値と、エネルギー準位 E_w から成る出力値とをRAM92に記憶する。そして、CPU91は、近似関数 f を求める演算を行なうプログラムの実行をキーボード98から指示されると、その指示に応じてRAM92に記憶した入力値及び出力値を読み出してディスプレイ97に表示する。

【0239】

ユーザは、ディスプレイ97に表示された入力値及び出力値を見て、出力値がなだらかな曲面上に存在するか否かを判定し、出力値がなだらかな曲面上に存在する場合、サンプルデータ30を構成する入力値 $x_1(n)$ 、 \dots 、 $x_{m_1}(n)$ 及び出力値 $z_1(n)$ ($n = 1 \sim S$) をキーボード98から指定する。

【0240】

そして、CPU91は、指定された入力値 $x_1(n)$ 、 \dots 、 $x_{m_1}(n)$ 及び出力値 $z_1(n)$ からサンプルデータ30を構成し、この発明によるプログラムをROM93から読み出し、その読み出したプログラムを実行して入力値 $x_1(n)$ 、 \dots 、 $x_{m_1}(n)$ と出力値 $z_1(n)$ との関係を規定する近似関数 f を求める演算を行なう。

【0241】

また、出力値 $z_1(n)$ ($n = 1 \sim S$) がなだらかな曲面上に存在するか否かの判定基準を予めCPU91に与えておき、入力値 $x_1(n)$ 、 \dots 、 $x_{m_1}(n)$ 及び出力値 $z_1(n)$ の抽出をCPU91が自動的に行なうようにしてもよい。

【0242】

従って、この発明によるプログラムは、より具体的には、量子井戸構造を決定するバリア層101、102及び井戸層103のパラメータを入力値とし、量子井戸における電子及び正孔のエネルギー準位を出力値とするサンプルデータを用いて、入力値と出力値との関係を規定する近似関数 f の演算を行なうことを特徴とする。

【0243】

上記においては、入力値 $x_1(n)$ 、 \dots 、 $x_{m_1}(n)$ 及び出力値 $z_1(n)$ は、パーソナルコンピュータ90により演算される場合を説明したが、この発明は、これに限らず、入力値 $x_1(n)$ 、 \dots 、 $x_{m_1}(n)$ 及び出力値 $z_1(n)$ を実験結果として取得し、その取得した実験結果をパーソナルコンピュータ90に入力してサンプルデータ3

10

20

30

40

50

0を構成するようにしてもよい。

【0244】

また、入力値 $x_1(n), \dots, x_{m_1}(n)$ 及び出力値 $z_1(n)$ の一方が実験結果であり、入力値 $x_1(n), \dots, x_{m_1}(n)$ 及び出力値 $z_1(n)$ の他方をパーソナルコンピュータ90による計算結果としてサンプルデータ30を構成するようにしてもよい。

【0245】

更に、上記においては、出力層43が1個の出力ユニットから成る3層ニューラルネットワーク40を用いて近似関数 f を求める演算を行なう場合について説明したが、この発明によるプログラムは、図30に示す3層ニューラルネットワーク40Aを用いて近似関数 f を求める演算を行なうようにしてもよい。 10

【0246】

図30は、この発明によるプログラムが入力値に対する出力演算値を演算する3層ニューラルネットワークの他の概念図を示す。3層ニューラルネットワーク40Aは、3層ニューラルネットワーク40の出力層43を出力層43Aに代えたものであり、その他は、3層ニューラルネットワーク40と同じである。

【0247】

出力層43Aは、出力ユニット43k ($k = 1 \sim m_3$) から成る。そして、出力ユニット43kは、中間ユニット42j ($j = 1 \sim m_2$) からそれぞれ出力 Y_j ($j = 1 \sim m_2$) を受け、その受けた出力 Y_j 、結合重み W_{jk} 及び閾値 θ_k を式(4)に代入して出力演算値 $Z_1(n), \dots, Z_{m_3}(n)$ ($n = 1 \sim S$) を演算する。 20

【0248】

3層ニューラルネットワーク40Aが用いられる場合も、この発明によるプログラムは、図17及び図18に示すフローチャートに従って近似関数 f を求める演算を実行する。そして、3層ニューラルネットワーク40Aを用いて近似関数 f を求める演算が実行される場合、 $n \times S$ 個の出力値 $z_1(n), \dots, z_{m_1}(n)$ が準備される。

【0249】

その他は、上述したとおりである。

更に、近似関数 f を求める演算に用いる入力値及び出力値は、量子井戸に関するデータに限らず、出力値がなだらかな曲面上に存在するデータであれば、どのような種類の入力値及び出力値であってもよい。 30

【0250】

今回開示された実施の形態はすべての点で例示であって制限的なものではないと考えられるべきである。本発明の範囲は、上記した実施の形態の説明ではなくて特許請求の範囲によって示され、特許請求の範囲と均等の意味および範囲内でのすべての変更が含まれることが意図される。

【図面の簡単な説明】

【図1】この発明によるプログラムが近似関数を求める演算に用いる入力値と出力値とを示す図である。

【図2】入力値が2個の場合における入力値と出力値との関係を示す図である。 40

【図3】この発明によるプログラムが入力値に対する出力演算値を演算する3層ニューラルネットワークの概念図である。

【図4】2個の入力値を用いた場合に式(5)を用いて演算された内部状態を示す図である。

【図5】2個の入力値を用いた場合の図3に示す中間ユニットからの出力を示す図である。

【図6】2個の入力値を用いた場合に式(1)を用いて演算された内部状態を示す図である。

【図7】2個の入力値を用いた場合の図3に示す中間ユニットからの出力を示す図である。

【図 8】パラメータ w_{ij} 及び w_{ij} の範囲を局所的特徴に相応する小さな半径にした場合における識別超球面の取り得る相対的位置関係と中間ユニット $42j$ の出力例とを示す図である。

【図 9】パラメータ w_{ij} 及び w_{ij} の範囲を広くした場合における識別超球面の取り得る相対的位置関係と中間ユニット $42j$ の出力例とを示す図である。

【図 10】パラメータ w_{ij} (中心) が入力変数の定義域外にある小さな超球である場合における識別超球面の取り得る相対的位置関係と中間ユニット $42j$ の出力例とを示す図である。

【図 11】高次元アルゴリズムによる最適化のフローチャートである。

【図 12】高次元アルゴリズムによる解の探索方法を示す概念図である。

【図 13】パラメータが 2 個の場合におけるコスト関数のランドスケープを示す図である。

【図 14】式 (9) に示される q_n (q_n) の概略図である。

【図 15】コスト関数の局所解に捉まる度合いを高次元アルゴリズム (HA) と焼きなまし法 (SA) について比較して示す図である。

【図 16】コスト関数の平坦領域を通過する速さを高次元アルゴリズム (HA) と焼きなまし法 (SA) について比較して示す図である。

【図 17】この発明によるプログラムが入力値 $x_1(n), \dots, x_{m_1}(n)$ と出力値 $z_1(n)$ との関係の規定する近似関数 f を求める演算を行なうフローチャートである。

【図 18】図 17 に示すステップ S2 の詳細な動作を説明するためのフローチャートである。

【図 19】中間ユニットの個数を 1 個に設定した場合の 3 層ニューラルネットワークの概念図である。

【図 20】中間ユニットの個数を 2 個に設定した場合の 3 層ニューラルネットワークの概念図である。

【図 21】二入力 - 出力の 3 つのテスト関数の等高線を示す図である。

【図 22】パーソナルコンピュータの概略ブロック図である。

【図 23】量子準位演算プログラムが演算の対象とする量子井戸の概念図である。

【図 24】量子準位演算プログラムを構成する各ステップを示すフローチャートである。

【図 25】量子準位演算プログラムが演算の対象とする量子井戸の具体例である。

【図 26】井戸ドーピングにおいて、ドーピング量を変化させたときの計算結果である。

【図 27】バリアドーピングにおいて、ドーピング量を変化させたときの計算結果である。

【図 28】量子準位演算プログラムを用いた計算結果と従来法による計算結果との比較を示す図である。

【図 29】基底状態のエネルギー値の分割数依存性を示す図である。

【図 30】この発明によるプログラムが入力値に対する出力演算値を演算する 3 層ニューラルネットワークの他の概念図である。

【符号の説明】

1 ~ 6 等高線、7 曲面、10, 20 集合、30 サンプルデータ、40, 40A 3 層ニューラルネットワーク、41 入力層、42, 42A, 42B 中間層、43, 43A 出力層、50 超平面、51, 52, 81 領域、60 超球面、61 ~ 65 曲面、70 意味空間、71 解、72, 82 矢印、80 高次元空間、89, 105, 106 曲線、90 パーソナルコンピュータ、91 CPU、92 RAM、93 ROM、94 シリアルインタフェース、95 端子、96 CD-ROM ドライブ、97 ディスプレイ、98 キーボード、99 CD、100 量子井戸、101, 102 バリア層、103 井戸層、104 エネルギー準位、411 ~ 41m1 入力ユニット、421 ~ 42m2 中間ユニット、431 ~ 43m3 出力ユニット。

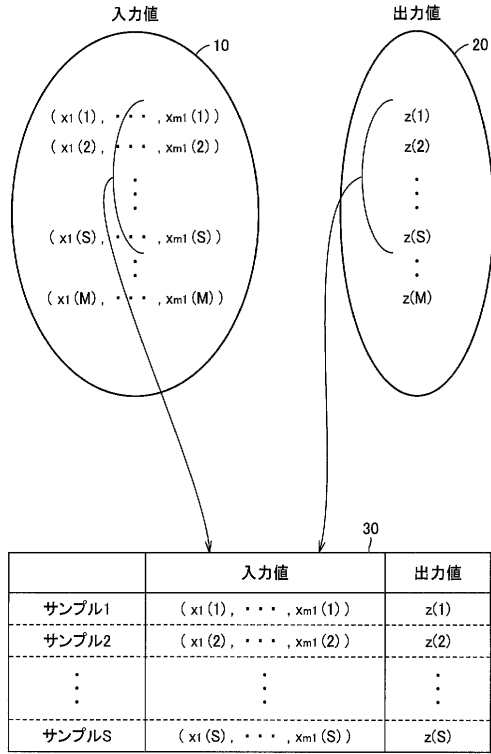
10

20

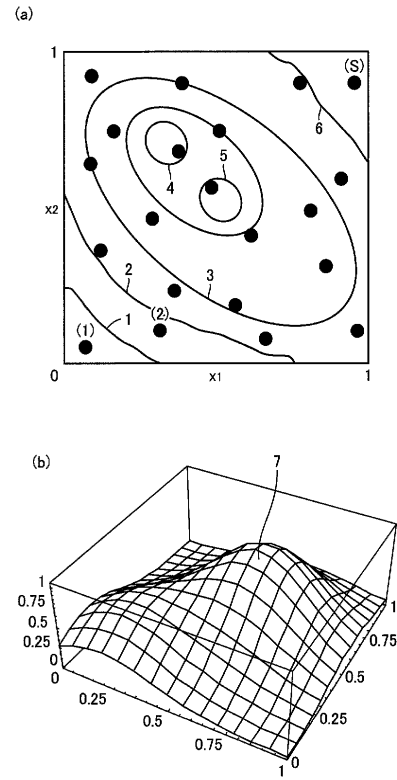
30

40

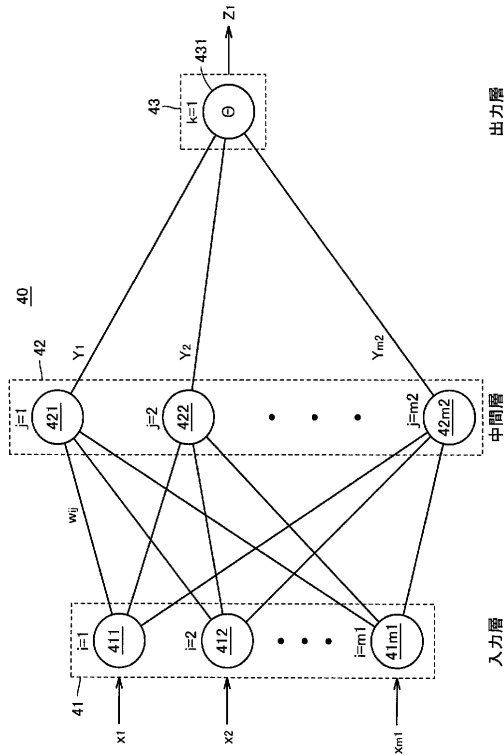
【 図 1 】



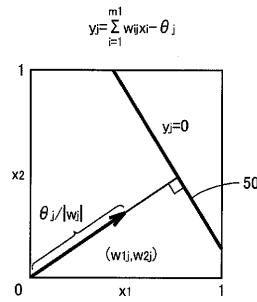
【 図 2 】



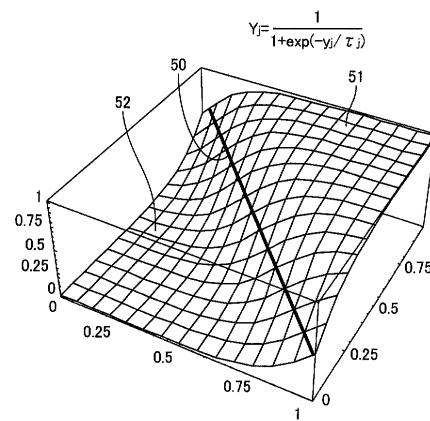
【 図 3 】



【 図 4 】

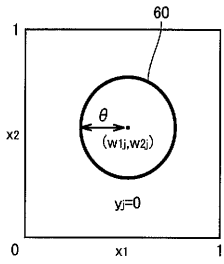


【 図 5 】



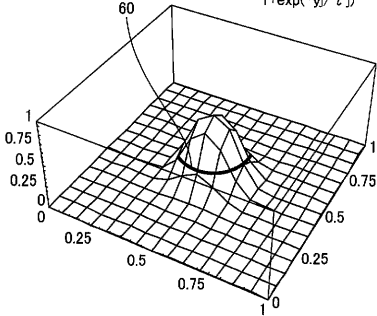
【 図 6 】

$$y_j = \theta_j^2 - \sum_{i=1}^{m1} (x_i - w_{ij})^2$$

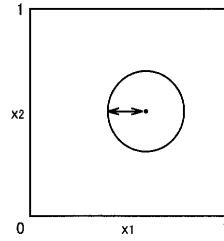


【 図 7 】

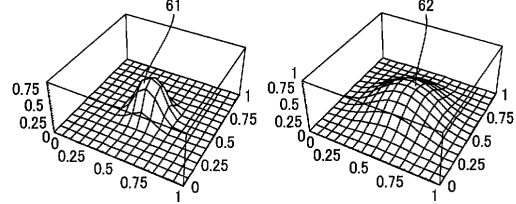
$$Y_j = \frac{1}{1 + \exp(-y_j / \tau_j)}$$



【 図 8 】

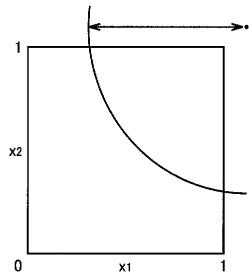


(a)

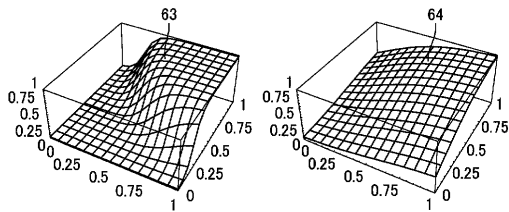


(b)

【 図 9 】

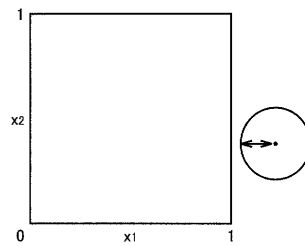


(a)

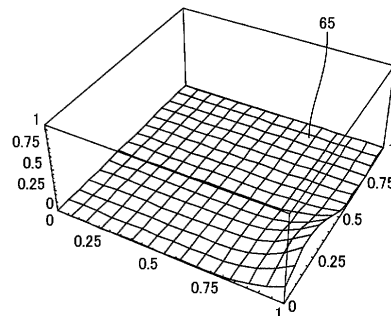


(b)

【 図 10 】

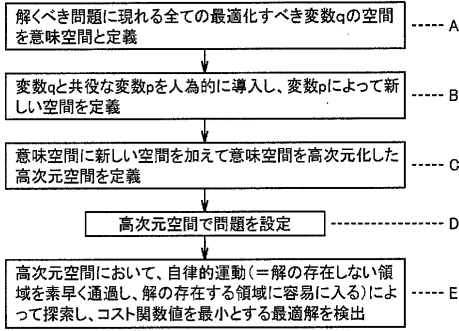


(a)

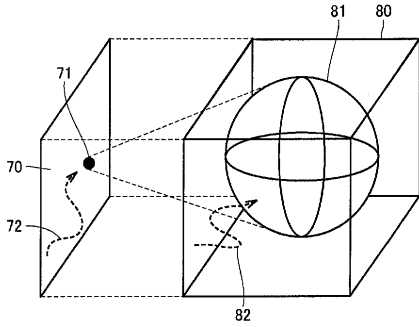


(b)

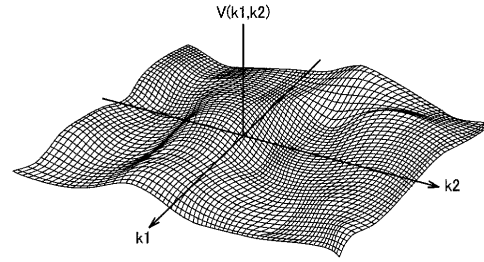
【 図 1 1 】



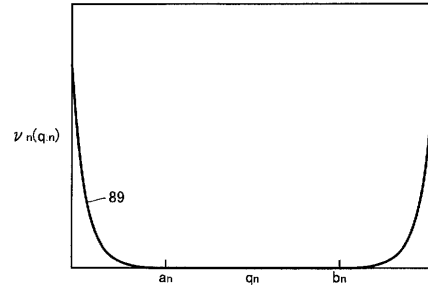
【 図 1 2 】



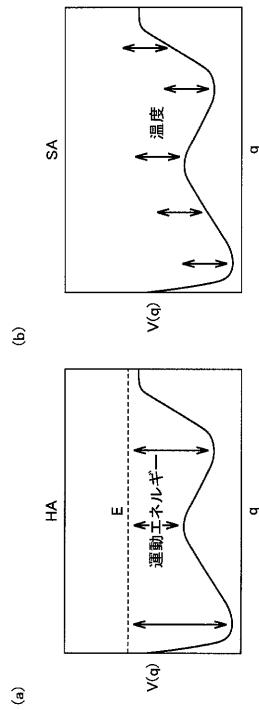
【 図 1 3 】



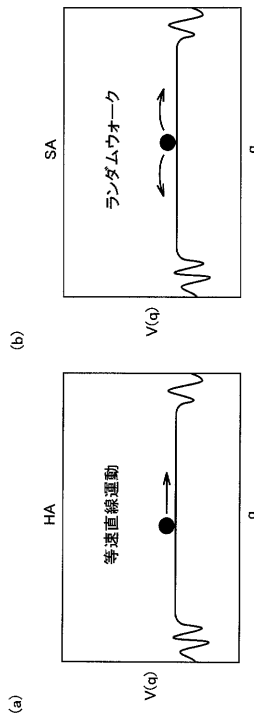
【 図 1 4 】



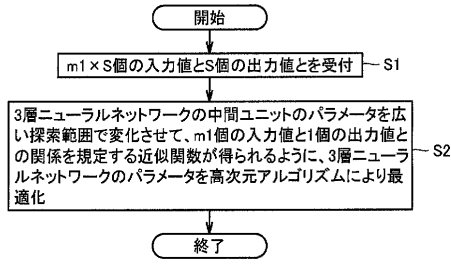
【 図 1 5 】



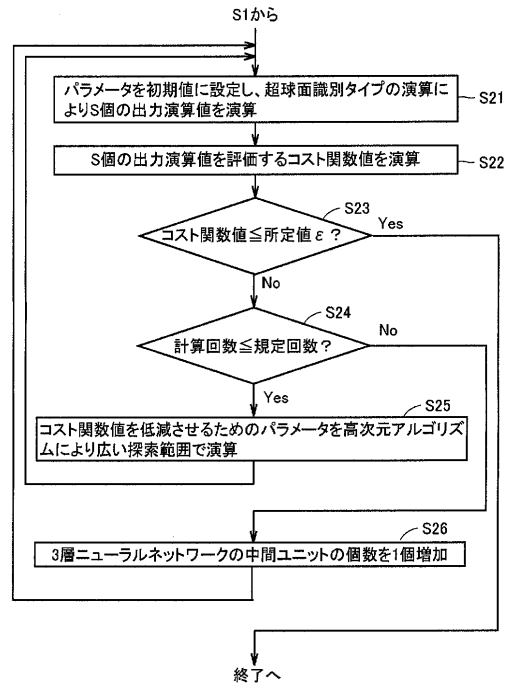
【 図 1 6 】



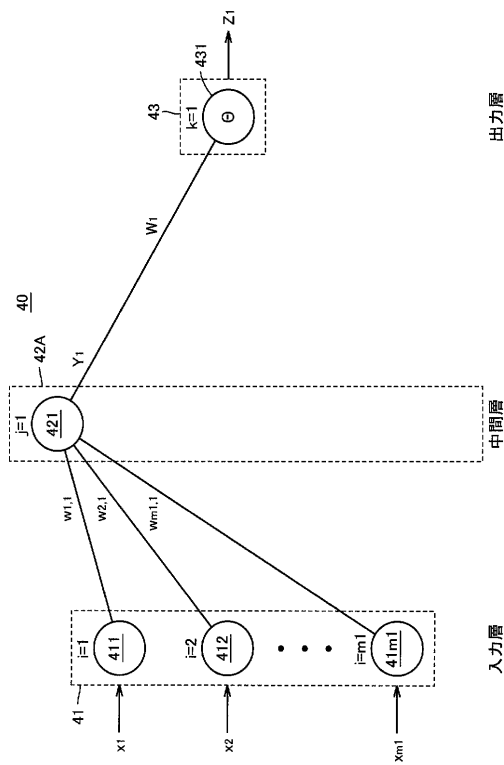
【図 17】



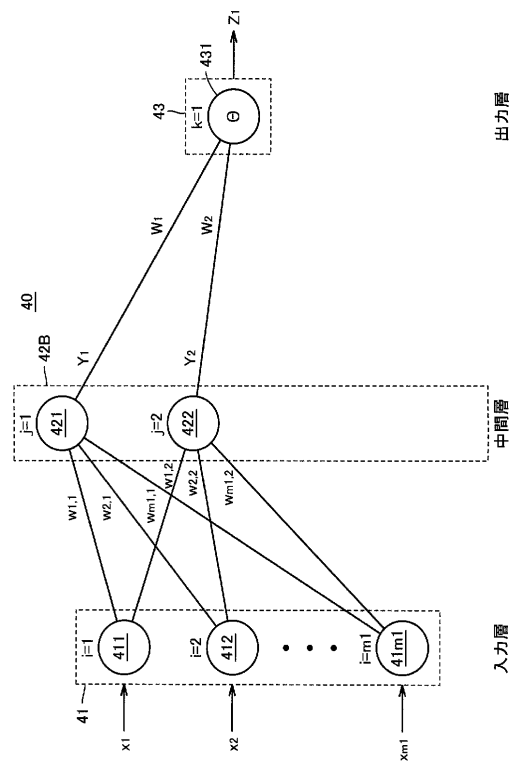
【図 18】



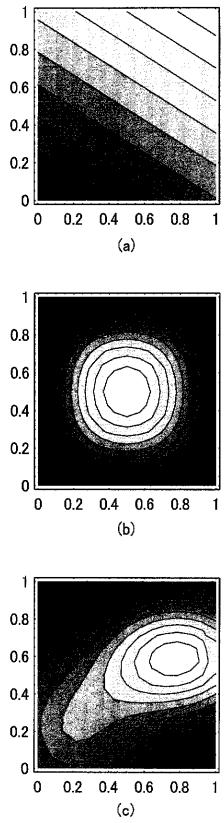
【図 19】



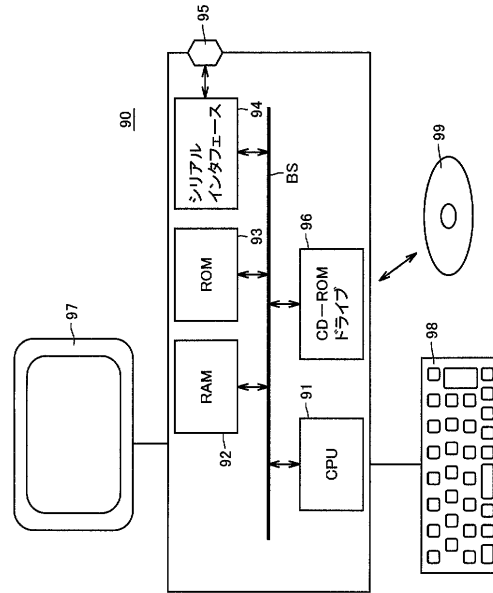
【図 20】



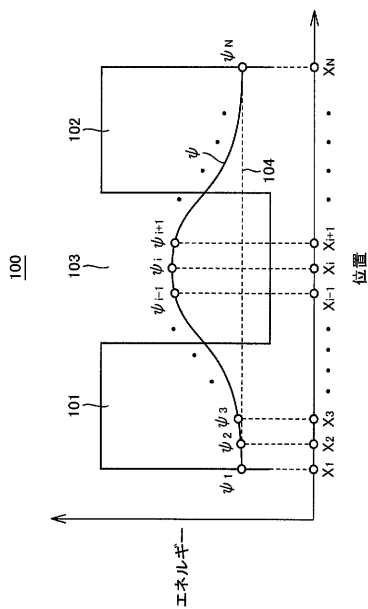
【図 2 1】



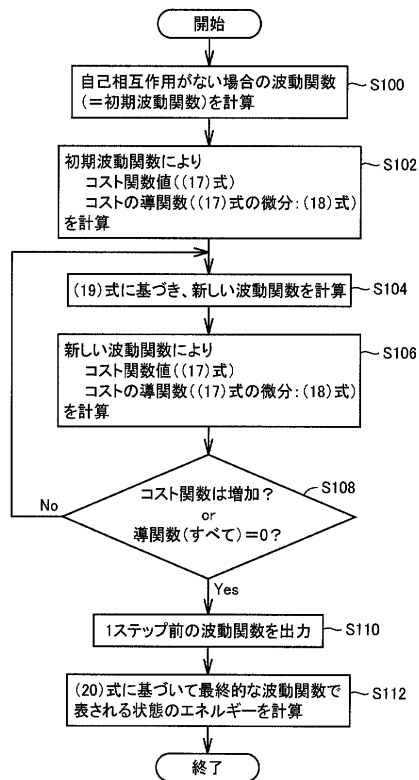
【図 2 2】



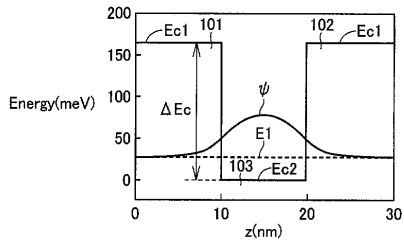
【図 2 3】



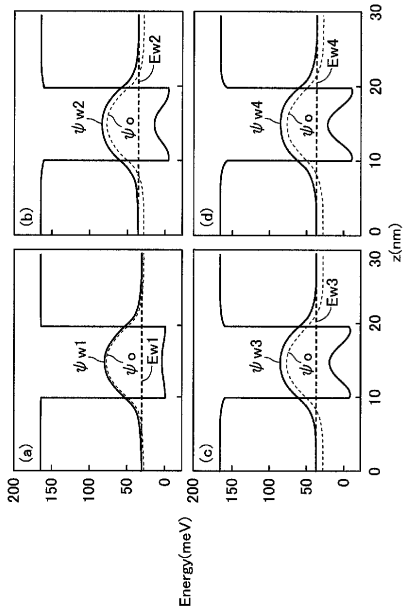
【図 2 4】



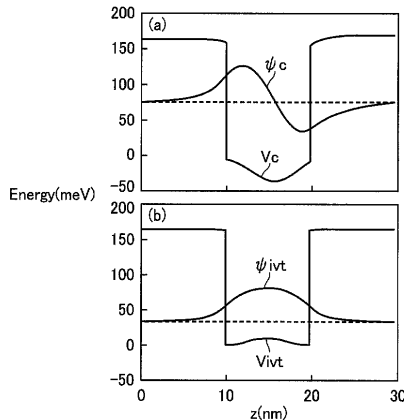
【 2 5 】



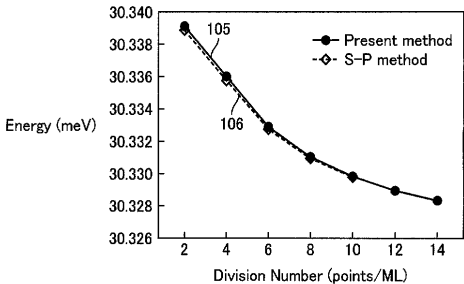
【 2 6 】



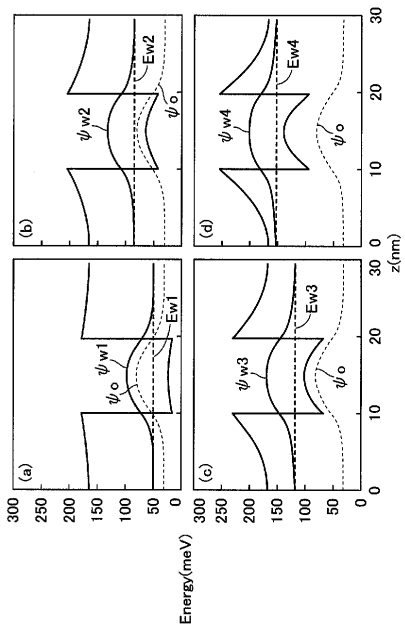
【 2 8 】



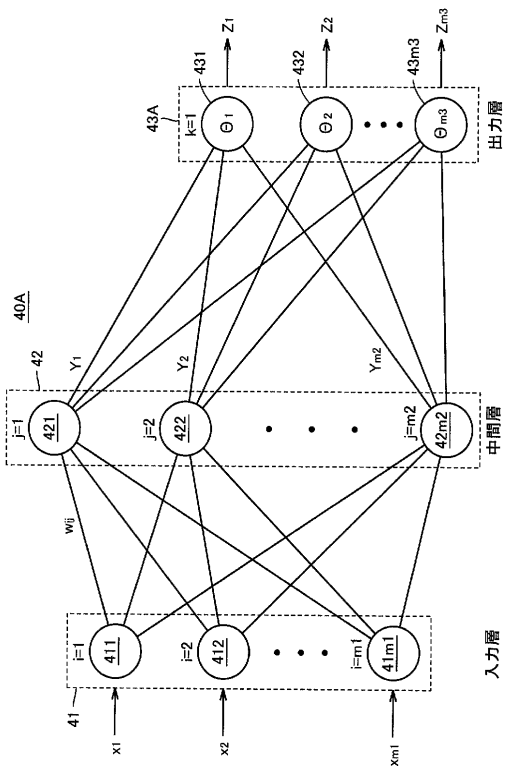
【 2 9 】



【 2 7 】



【 3 0 】



フロントページの続き

- (72)発明者 河野 芳江
京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内
- (72)発明者 安藤 太郎
京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内
- (72)発明者 斎藤 茂
京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内
- Fターム(参考) 5B056 AA06 BB22 BB51