

文章编号: 1001-0920(2011)06-0898-04

基于受控旋转门的量子神经网络模型算法及应用

李盼池^{a,b}, 宋考平^a, 杨二龙^a

(东北石油大学 a. 石油与天然气工程博士后科研流动站, b. 计算机与信息技术学院, 黑龙江大庆 163318)

摘要: 提出一种量子神经网络模型及算法. 首先借鉴受控非门的含义提出一种受控量子旋转门, 基于该门的物理意义, 提出一种量子神经元模型, 该模型包含对输入量子比特相位的旋转角度和对旋转角度的控制量两种设计参数; 然后基于上述量子神经元提出一种量子神经网络模型, 基于梯度下降法详细设计了该模型的学习算法; 最后通过模式识别和时间序列预测两个仿真验证了该模型及算法在收敛能力和鲁棒性方面优于普通的 BP 网络.

关键词: 量子计算; 受控量子旋转门; 量子神经元; 量子神经网络

中图分类号: TP18

文献标识码: A

Controlled-rotating gate-based quantum neural networks model algorithm with application

LI Pan-chi^{a,b}, SONG Kao-ping^a, YANG Er-long^a

(a. Post-doctoral Research Center of Oil and Gas Engineering, b. School of Computer and Information Technology, Northeast Petroleum University, Daqing 163318, China. Correspondent: LI Pan-chi, E-mail: lipanchi@vip.sina.com)

Abstract: A quantum neural networks model and its learning algorithm are presented. Firstly, a quantum controlled-rotating gate is proposed by analyzing the meaning of the controlled-NOT gate. Then a quantum neuron model is constructed from the physical meaning of the controlled rotating gate. The model includes two kinds of design parameters, rotation angle and its control range. Secondly, a quantum neural networks model based on quantum neuron is proposed. By using gradient descent algorithm, a learning algorithm of this model is designed in detail. With applications of pattern recognition and time series prediction, the simulation results show that the proposed algorithm is superior to the common BP neural networks in both convergence capability and robustness.

Key words: quantum computation; quantum controlled-rotation gate; quantum neuron; quantum neural networks

1 引言

被誉为 20 世纪伟大科学革命之一的量子力学尽管创立已有百年的历史, 然而它在神经计算中的应用仅仅始于 20 世纪 90 年代. Kak^[1]于 1995 年首次提出了量子神经计算的概念. Gopathy 等人^[2]于 1997 年借用量子理论中量子态叠加的思想, 提出了基于多级激励函数的量子神经网络(QNN)模型, 在 3 层网络结构中, 隐层量子神经元的激励函数采用多个传统激励函数的叠加, 使网络有了一种固有的模糊性. Ventura 等人^[3]于 1998 年研究了量子计算的学习问题, 提出了量子联想存储的算法, 相对于传统存储量子联想存储具有指数级的存储容量. Ajit 等人^[4]于 2000 年研究了量子神经网络的结构和模型, 从多宇宙的量子理

论观点提出了构建叠加的多宇宙量子神经网络模型的思想. 解光军等人^[5]于 2001 年首次在国内撰文阐述了量子神经计算的概念, 对国内量子神经网络研究起到了一定的先导作用. 并于 2004 年深入研究了以通用量子逻辑门组(即相移门和受控非门)作为计算基函数, 构造量子神经网络模型的新方法, 研究结果表明该模型的性能优于传统神经网络^[6-7]. 李盼池等人^[8]于 2007 年提出一种基于量子输入和量子权值的量子自组织特征映射网络模型及聚类算法, 该算法可明显提高普通自组织网络的聚类性能; 2008 年, 基于量子旋转门和量子受控非门的物理意义, 提出一种基于通用量子门组演化的量子 BP 网络模型^[9]; 2009 年, 又提出一种基于量子权值及量子活性值的量子神经

收稿日期: 2010-03-18; 修回日期: 2010-05-07.

基金项目: 中国博士后科学基金项目(20090460864, 201003405); 黑龙江省博士后科学基金项目(LBH-Z09289); 黑龙江省教育厅科学基金项目(11551015).

作者简介: 李盼池(1969—), 男, 博士, 副教授, 从事量子搜索、量子智能优化等研究; 宋考平(1962—), 男, 教授, 博士生导师, 从事油气田开发工程研究.

网络模型, 基于梯度下降法构造了该模型的超线性收敛学习算法^[10]. 目前量子计算与神经计算的融合正逐步成为一个崭新的研究方向.

本文在分析量子受控运算的基础上, 提出一种基于受控量子旋转门的量子神经元模型和量子神经网络模型, 基于量子计算原理构造了该模型的学习算法. 仿真结果表明, 该模型及算法在收敛能力和鲁棒性两个方面明显优于普通BP(CBP)网络.

2 量子比特及量子门的概念

2.1 量子比特

在量子计算中, 量子比特是作为传统计算中“比特”的对应物而引入的, 目的是描述量子线路的状态. 量子比特有两个可能的状态 $|0\rangle$ 和 $|1\rangle$, 记号“ $| \cdot \rangle$ ”称为Dirac记号. 比特和量子比特的区别在于, 量子比特的状态可以落在 $|0\rangle$ 和 $|1\rangle$ 之外, 可以是状态的线性组合, 常称为叠加态, 即

$$|\phi\rangle = \alpha|0\rangle + \beta|1\rangle, \quad (1)$$

其中 α 和 β 是一对复数, 称为量子态的概率幅, 即量子态 $|\phi\rangle$ 以 $|\alpha|^2$ 的概率坍缩到 $|0\rangle$, 或者以 $|\beta|^2$ 的概率坍缩到 $|1\rangle$, 且

$$|\alpha|^2 + |\beta|^2 = 1. \quad (2)$$

因此, 量子态也可由概率幅表示为 $|\phi\rangle = [\alpha, \beta]^T$.

2.2 受控量子旋转门及数学描述

在量子计算中, 通过对量子位的状态进行一系列的酉变换, 可以实现某些逻辑功能. 这些变换所起的作用相当于逻辑门所起的作用, 因此称在一定时间间隔内实现逻辑变换的量子装置为量子门. 量子门是物理实现量子计算的基础, 它包含了量子计算的特点.

1) 量子旋转门. 量子旋转门定义为

$$R(\varphi) = \begin{bmatrix} e^{i\varphi} & 0 \\ 0 & e^{i\varphi} \end{bmatrix}. \quad (3)$$

若令 $|\phi\rangle = \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix}$, 则有 $R(\varphi)|\phi\rangle = \begin{bmatrix} \cos(\theta + \varphi) \\ \sin(\theta + \varphi) \end{bmatrix}$, 显然 $R(\varphi)$ 起到了相位旋转的作用.

2) 受控量子旋转门. 受控非门的意义为, 若控制比特置0, 则目标比特不变; 若控制比特置1, 则目标比特翻转. 由此提出一种受控量子旋转门, 该门有两个输入比特, 分别是控制比特和目标比特, 若控制比特置0, 则目标比特保持不变; 若控制比特置1, 则目标比特将产生相位旋转. 因此, 受控旋转门描述为

$$C(k, \varphi) = \begin{bmatrix} e^{ik\varphi} & 0 \\ 0 & e^{ik\varphi} \end{bmatrix}. \quad (4)$$

其受控作用根据参数 k 的取值, 可分为如下3种情况:

a) 当 $k = 1$ 时, $C(k, \varphi) = [\cos(\theta + \varphi) \ \sin(\theta + \varphi)]^T$, 使 $|\phi\rangle$ 的相位旋转 φ ;

b) 当 $k = 0$ 时, $C(k, \varphi) = [\cos \theta \ \sin \theta]^T$, 此时 $|\phi\rangle$ 的相位不变;

c) 当 $0 < k < 1$ 时, $C(k, \varphi) = [\cos(\theta + k\varphi) \ \sin(\theta + k\varphi)]^T$, 使 $|\phi\rangle$ 的相位旋转 $k\varphi$.

3 量子神经网络模型

3.1 量子神经元模型

本文提出的基于受控量子旋转门的量子神经元模型包括输入、移相、聚合、旋转、输出5个部分, 其中输入和输出用量子位表示, 移相、旋转分别由量子旋转门 $R(\theta_i)$ 和受控旋转门 $U(\lambda, \varphi)$ 实现. 模型见图1. 图1中 $R(\theta_i)$ 的定义同式(3), $U(\lambda, \varphi)$ 的定义为

$$U(\lambda, \varphi) = C(f(\lambda), \varphi). \quad (5)$$

其中: $f(\cdot)$ 为Sigmoid函数, $C(\cdot, \cdot)$ 的定义同式(4).

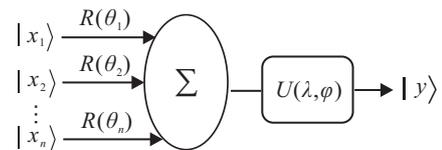


图1 量子神经元模型

记 $|x_i\rangle = [\cos t_i \ \sin t_i]^T$, 相移和聚合的结果可定义为

$$\sum_{i=1}^n R(\theta_i)|x_i\rangle = [\cos \theta \ \sin \theta]^T, \quad (6)$$

其中

$$\theta = \arg \left(\sum_{i=1}^n R(\theta_i)|x_i\rangle \right) = \arctan \left(\frac{\sum_{i=1}^n \sin(t_i + \theta_i)}{\sum_{i=1}^n \cos(t_i + \theta_i)} \right).$$

经 $U(\lambda, \varphi)$ 作用后, 量子神经元的输出为

$$|y\rangle = U(\lambda, \varphi) \sum_{i=1}^n R(\theta_i)|x_i\rangle = [\cos(\theta + f(\lambda)\varphi) \ \sin(\theta + f(\lambda)\varphi)]^T. \quad (7)$$

3.2 量子神经网络模型

量子神经网络由若干个量子神经元按一定的拓扑结构和连接规则组成. 本节提出的3层前馈QNN模型如图2所示. 其中输入层、隐含层和输出层分别有 n, p, m 个量子神经元, 网络输出为输出层量子神经元处于状态 $|1\rangle$ 的概率幅.

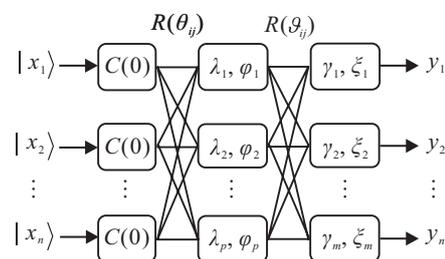


图2 3层前馈QNN模型

记

$$\alpha_j = \arg \left(\sum_{i=1}^n R(\theta_{ij}) |x_i\rangle \right) + f(\lambda_j) \varphi_j,$$

$$\beta_k = \arg \left(\sum_{j=1}^p R(\vartheta_{jk}) |h_j\rangle \right) + f(\gamma_k) \xi_k.$$

各层输入输出关系可描述为

$$|h_j\rangle = [\cos \alpha_j \quad \sin \alpha_j]^T, \quad (8)$$

$$y_k = \sin \beta_k. \quad (9)$$

其中: $i = 1, 2, \dots, n, j = 1, 2, \dots, p, k = 1, 2, \dots, m$.

4 量子神经网络算法

4.1 训练样本的量子态描述

为了实现训练样本的量子态描述, 对于 n 维欧氏空间以实值向量描述的训练样本 $\bar{X} = [\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n]^T$, 定义转换公式如下:

$$|X\rangle = [|x_1\rangle, |x_2\rangle, \dots, |x_n\rangle]^T, \quad (10)$$

其中 $|x_i\rangle = \left[\cos \left(\frac{2\pi}{1 + e^{-x_i}} \right), \sin \left(\frac{2\pi}{1 + e^{-x_i}} \right) \right]^T$.

4.2 网络参数的调整规则

令 \hat{y}_i 和 y_i 为经归一化后的期望输出和实际输出, 定义误差函数为

$$E = \frac{1}{2} \sum_{i=1}^m (\hat{y}_i - y_i)^2. \quad (11)$$

记

$$|x_i\rangle = [\cos \phi_i \quad \sin \phi_i]^T, \quad A_j = \sum_{i=1}^n \sin(\phi_i + \theta_{ij}),$$

$$B_j = \sum_{i=1}^n \cos(\phi_i + \theta_{ij}), \quad \alpha_j = \arctan(A_j/B_j) + f(\lambda_j) \varphi_j,$$

$$C_k = \sum_{j=1}^p \sin(\alpha_j + \vartheta_{jk}), \quad D_k = \sum_{j=1}^p \cos(\alpha_j + \vartheta_{jk}),$$

$$\beta_k = \arctan(C_k/D_k) + f(\gamma_k) \xi_k,$$

$$S_{jk} = (\cos(\alpha_j + \vartheta_{jk}) D_k + \sin(\alpha_j + \vartheta_{jk})) / (C_k^2 + D_k^2),$$

$$T_{ij} = (\cos(\phi_i + \theta_{ij}) B_j + \sin(\phi_i + \theta_{ij})) / (A_j^2 + B_j^2).$$

根据梯度下降法, 有

$$\Delta \gamma_k = -\partial E / \partial \gamma_k = (\hat{y}_k - y_k) \cos \beta_k \xi_k f(\gamma_k) (1 - f(\gamma_k)), \quad (12)$$

$$\Delta \xi_k = -\frac{\partial E}{\partial \xi_k} = (\hat{y}_k - y_k) \cos \beta_k f(\gamma_k), \quad (13)$$

$$\Delta \vartheta_{jk} = -\frac{\partial E}{\partial \vartheta_{jk}} = (\hat{y}_k - y_k) \cos \beta_k S_{jk}, \quad (14)$$

$$\Delta \lambda_j = -\partial E / \partial \lambda_j = (\hat{y}_k - y_k) \cos \beta_k S_{jk} \varphi_j f(\lambda_j) (1 - f(\lambda_j)), \quad (15)$$

$$\Delta \varphi_j = -\frac{\partial E}{\partial \varphi_j} = (\hat{y}_k - y_k) \cos \beta_k S_{jk} f(\lambda_j), \quad (16)$$

$$\Delta \theta_{ij} = -\frac{\partial E}{\partial \theta_{ij}} = (\hat{y}_k - y_k) \cos \beta_k S_{jk} T_{ij}. \quad (17)$$

其中: $i = 1, 2, \dots, n, j = 1, 2, \dots, p, k = 1, 2, \dots, m$.

参数调整规则可统一描述为 $\chi(t+1) = \chi(t) + \eta \Delta \chi(t+1)$. 其中: χ 分别等于 $\theta_{ij}, \vartheta_{jk}, \varphi_j, \xi_k, \lambda_j$ 和 γ_k ; η 为学习速率.

由网络输出式(9)可以看出, 4个迭代序列 $\{\theta_{ij}(t)\}, \{\vartheta_{jk}(t)\}, \{\varphi_j(t)\}, \{\xi_k(t)\}$ 的全局最优解并不唯一. 事实上, 存在如下定理.

定理 1 设 $\bar{\theta}_{ij}, \bar{\vartheta}_{jk}, \bar{\varphi}_j, \bar{\xi}_k$ 分别是迭代序列 $\{\theta_{ij}(t)\}, \{\vartheta_{jk}(t)\}, \{\varphi_j(t)\}, \{\xi_k(t)\}$ 的全局最优解, 则当 n_1, n_2, n_3, n_4 为任意整数时, 下式确定的 $\hat{\theta}_{ij}, \hat{\vartheta}_{jk}, \hat{\varphi}_j, \hat{\xi}_k$ 也是全局最优解:

$$\begin{aligned} \hat{\theta}_{ij} &= 2n_1\pi + \bar{\theta}_{ij}, \quad \hat{\vartheta}_{jk} = 2n_2\pi + \bar{\vartheta}_{jk}, \\ \hat{\varphi}_j &= 2n_3\pi/f(\lambda_j) + \bar{\varphi}_j, \quad \hat{\xi}_k = 2n_4\pi/f(\gamma_k) + \bar{\xi}_k. \end{aligned} \quad (18)$$

证明 设 \hat{y}_k 为与实际输出 y_k 对应的期望输出, $|\hat{h}_j\rangle$ 为与 $|h_j\rangle$ 对应的隐层期望输出. 由

$$\begin{aligned} \alpha_j &= \arg \left(\sum_{i=1}^n R(2n_1\pi + \bar{\theta}_{ij}) |x_i\rangle \right) + \\ & f(\lambda_j) (2n_3\pi/f(\lambda_j) + \bar{\varphi}_j), \end{aligned}$$

得到 $|h_j\rangle = [\cos \alpha_j, \sin \alpha_j]^T = |\hat{h}_j\rangle$. 由

$$\begin{aligned} \beta_k &= \arg \left(\sum_{j=1}^p R(2n_2\pi + \bar{\vartheta}_{jk}) |\hat{h}_j\rangle \right) + \\ & f(\gamma_k) (2n_4\pi/f(\gamma_k) + \bar{\xi}_k), \end{aligned}$$

得到 $y_k = \sin(\beta_k) = \hat{y}_k$. 故由式(18)确定的 $\hat{\theta}_{ij}, \hat{\vartheta}_{jk}, \hat{\varphi}_j, \hat{\xi}_k$ 也是全局最优解. \square

5 实际应用

5.1 平面点集分类

给定平面上 9 个点共两类, 模式分布如图 3 所示, 样本数据见表 1.

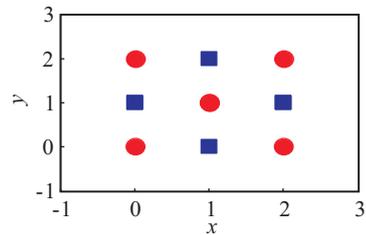


图 3 9点模式分布

表 1 9点模式信息

坐标	(0,0)	(1,0)	(2,0)	(0,1)	(1,1)	(2,1)	(0,2)	(1,2)	(2,2)
类别	1	2	1	2	1	2	1	2	1

分别用 QNN 和 CBP 作为分类器, 网络结构取 2-10-1 型, 限定迭代步数为 15 000, 限定误差精度为 0.1, 学习速率从集合 $\{0.1, 0.2, \dots, 1.0\}$ 中取值. 对于集合中的每种学习速率, 分别用 QNN 和 CBP 进行 100 次仿真, 并记录收敛次数、最小迭代步数、平均迭代步

数等评价指标. 当学习速率改变时, 两种模型的收敛率对比见图 4, 迭代步数对比见图 5.

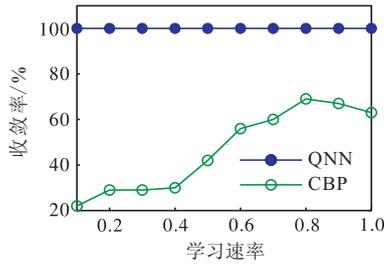


图 4 QNN 和 CBP 的收敛率对比

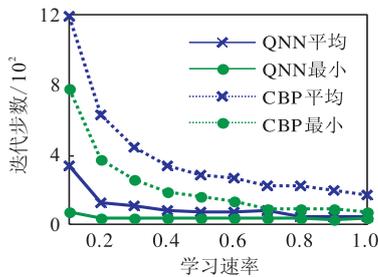


图 5 QNN 和 CBP 的迭代步数对比

根据图 4, 当学习速率变化时, QNN 的收敛率均为 100%; 而 CBP 的收敛率最小为 22%(当 $\eta = 0.1$ 时), 最大仅为 69%(当 $\eta = 0.8$ 时). 根据图 5, 当学习速率变化时, 以平均迭代步数为例, QNN 最大为 3 987.70 (当 $\eta = 0.1$ 时), 最小为 275.01(当 $\eta = 0.9$ 时), 波动范围仅为 3 712.69; 然而, CBP 最大为 12 335.45(当 $\eta = 0.1$ 时), 约为 QNN 的 3.1 倍, 最小为 1 697.24(当 $\eta = 1.0$ 时), 约为 QNN 的 6.2 倍, 波动范围高达 10 638.21, 约为 QNN 的 2.9 倍. 仿真结果表明, QNN 不仅迭代步数少、收敛率高, 而且当参数变化时具有较好的鲁棒性.

5.2 太阳黑子数预测

采用太阳黑子的年平均数序列进行仿真, 自 1700 年~2007 年, 共 308 个数据, 分布特征见图 6.

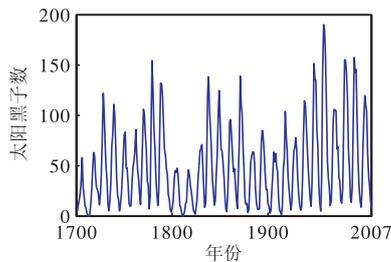


图 6 1700~2007 年的太阳黑子数据

多次仿真结果最终确定的预测方案是: 用连续 10 年的数据预测其后 1 年的太阳黑子数. 因此, 可以构造出 298 个样本数据. 用 1700 年~1957 年共 258 年的数据构造训练样本集, 完成网络训练. 用余下的 50 个样本作为预测集, 以检验网络的预测能力. 由样本特征可知, 输入层为 10 个节点, 输出层为一个节点. 隐层节点取为 30, 限定迭代步数取为 2 000, 学习速率

取为 0.5, 惯性因子取为 0.1. QNN 和 CBP 的训练结果见表 2, 对预测集的预测结果见表 3, 预测结果与真实结果对比见图 7.

表 2 QNN 和 CBP 的训练结果对比

样本个数	最大误差		最小误差		均方误差	
	QNN	CBP	QNN	CBP	QNN	CBP
248	51.387	106.498 2	0.002 6	0.018 0	0.835 5	1.562 2

表 3 QNN 和 CBP 的预测结果对比

样本个数	最大误差		最小误差		均方误差	
	QNN	CBP	QNN	CBP	QNN	CBP
50	10.582 0	44.447 6	0.041 8	0.561 5	0.389 8	2.585 7

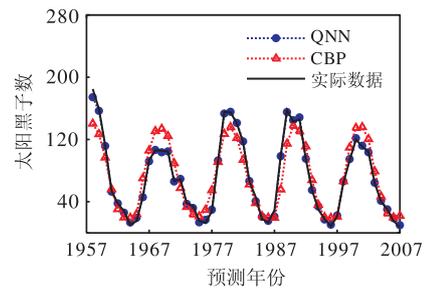


图 7 QNN 和 CBP 的预测结果对比

根据以上仿真结果可以看出, QNN 模型预测结果的最大误差、最小误差、均方误差均明显小于 CBPNN 模型, 这是由于在预测模型中引入了量子机制的缘故. 在 QNN 中, 通过量子旋转门和受控量子旋转门的协同作用, 增强了网络的学习能力、模式信息概括能力、函数逼近能力以及泛化推广能力, 从而使 QNN 呈现出良好的适应性.

6 结 论

量子神经网络是量子计算与神经计算相融合的产物. 本文基于量子旋转门和受控量子旋转门提出了一种新的量子神经网络模型, 详细设计了该模型的算法. 理论分析证明, 该模型具有周期性的多重吸引子, 可以提高收敛的概率. 实验结果表明, 对于简单模式识别问题, 该模型表现出良好的性能, 不仅具有较快的收敛速度、较高的收敛率, 而且具有较好的鲁棒性. 对于复杂时间序列预测问题, 该模型同样表现出较好的非线性逼近能力, 从而验证了量子机制的引入可以改善网络的性能, 为量子计算与神经计算的融合开辟了一条新途径.

参考文献(References)

[1] Kak S. On quantum neural computing[J]. Information Sciences, 1995, 83(3): 143-160.