基于神经网络的一类大系统动态递阶优化方法[1)]

侯增广　高慧琪　吴沧浦
(北京理工大学自动控制系　北京　100081)

摘　要　提出了一种用于求解离散时间大系统动态递阶优化问题的神经网络模型(LHONN)，该网络以全集成化为特征：1)各子系统的动态方程嵌入相应的局部优化网络中，使得网络结构具有较低的维数，易于硬件实现；2)其上级协调网络和局部优化网络的求解过程同时进行，优化求解速度高，适宜于实时系统优化。

关键词　大系统，动态系统，递阶优化，神经网络

1　引言

由于神经网络具有大规模分布存储和并行处理信息的能力，引起了诸多领域研究人员的极大兴趣，并在理论和实践研究中取得了长足的进展。其中，自动控制领域中神经网络的研究已经成为探索神经网络的最为活跃和最有前途的领域之一，在用以解决优化和优化控制问题上尤其显现了突出的良好效果。Hopfield 首先把神经网络用来解决"旅行商"问题，这是用神经网络求解优化问题的一项具有奠基性的工作，由于神经网络方法极大地改善了传统的基于迭代的优化方法所困扰的收敛速度和计算量等问题，人们开始把神经网络作为一种有效的快速求解优化问题的一个突破口，做了大量成功的研究。

由于大系统往往具有较高的维数，这使得传统优化计算中的"维数灾"问题变得非常突出；并且传统方法在其数值迭代求解时，工作效率低，难以满足实时优化的需要。本文给出了求解一类大系统动态递阶优化问题的神经网络(LHONN)(Neural Networks for Large-Scale Dynamical Hierarchical Optimization)，该网络具有如下特点：1) 将各子系统的动态方程及其初始条件约束嵌入于相应的局部优化网络中，不但克服了该项约束所带来的求解困难，同时显著地降低了神经网络的状态维数；2) 该网络模型特别适合于并行处理；3) 该神经网络的协调网络和局部优化网络的求解计算同步进行，具有很高的求解效率；4) 该网络具有全集成化的特点，适宜于硬件实现，而且该网络稳定性好，参数选择容易，便于实际大系统优化的应用。

2　离散时间大系统动态递阶优化问题的描述

考虑如下的由 N 个相互关联子系统构成的离散时间大系统

1)　国家自然科学基金资助项目.
收稿日期　1995-12-20

$$x_i(k+1) = A_i x_i(k) + B_i u_i(k) + C_i z_i(k), \qquad x_i(0) = x_{i0}, \tag{1}$$

子系统间的关联方程为 $$z_i(k) = \sum_{j=1}^{N} G_{ij} x_j(k). \tag{2}$$

其中 $x_i \in R^{n_i}, u_i \in R^{m_i}, z_i \in R^{r_i}, i=1,2,\cdots,N, k=0,1,\cdots,K-1, A_i, B_i, C_i, G_{ij}$为具有相应维数的矩阵,且 $\mathrm{rank}(A_i)=n_i, \mathrm{rank}(B_i)=m_i, \mathrm{rank}(C_i)=r_i$.

一个大系统优化问题(P)可以表述为求 N 个控制向量序列 $u_i(k)$在满足式(1),(2)的约束下使得下列目标函数极小化

$$J = \sum_{i=1}^{N} J_i$$
$$= \sum_{i=1}^{N}\left[\frac{1}{2}\|x_i(k)\|^2_{Q_i(K)} + \sum_{k=0}^{K-1}\frac{1}{2}\left(\|x_i(k)\|^2_{Q_i(k)} + \|z_i(k)\|^2_{S_i(k)} + \|u_i(k)\|^2_{R_i(k)}\right)\right], \tag{3}$$

其中 $Q_i(k) \in R^{n_i \times n_i}$为半正定矩阵,$R_i(k) \in R^{m_i \times m_i}, S_i(k) \in R^{r_i \times r_i}$为正定矩阵.

递阶优化的基本思想是分级分布式优化,由下级的 N 个局部优化单元和上级的协调器构成。传统的递阶优化方法基本分两步:一是各局部决策单元按照协调器的干预信号进行优化计算;二是协调器根据各子系统的反馈信息调节干预信号。值得注意的是,在数值计算时以上两步只能反复交替进行,优化效率低,不能满足实时优化的要求.

3　求解一类大系统动态递阶优化问题的神经网络设计

3.1　原理

该网络方法的原理是目标协调法,定义原问题的对偶函数为:

$$\Phi(\lambda) = \min_{x,u,z}\{L(x,u,z,\lambda)\} \quad \text{s.t.}\ (1) \quad , \tag{4}$$

式中 $$L(x,u,z,\lambda) = \sum_{i=1}^{N} L_i(\cdot), \tag{5}$$

$$L_i(\cdot) = \frac{1}{2}\|x_i(K)\|^2_{Q_i(K)} + \sum_{k=0}^{K-1}\left[\frac{1}{2}\|x_i(k)\|^2_{Q_i(k)} + \frac{1}{2}\|u_i(k)\|^2_{R_i(k)} + \frac{1}{2}\|z_i(k)\|^2_{S_i(k)}\right.$$
$$\left. + \lambda_i^T(k) z_i(k) - \sum_{j=1}^{N}\lambda_j^T(k) G_{ji} x_i(k)\right]. \tag{6}$$

其中 Lagrange 乘子 $\lambda_i \in R^{r_i}$. 这里原问题(P)的极小化问题等价于对偶函数(4)的极大化问题.

本文所要设计的神经网络是要具有如下的任务求解功能:1)局部优化网络(LONN),在每一时刻,对协调器给予的 $\lambda_i(k)$,求出相应子系统的最优决策 $u_i(k)$和最优状态 $x_i(k)$;2)协调网络(CNN),在每一时刻,对于由局部优化网络反馈来的信息,产生新的协调信号 $\lambda_i(k)$. 尤其重要的是,局部优化网络和协调网络同时工作,给出最优解.

3.2　设计过程

在大系统递阶优化过程中协调器的任务是使 N 个子系统处于关联平衡状态. 所以,在各子系统优化所要求满足的约束(1)式中,可以把控制项 $u_i(k)$和关联项 $z_i(k)$均当作输入项处理,则一个局部优化网络所要完成的任务是

$$\min_{u_i(x_i(k)), z_i(x_i(k))} L_i(\cdot) \quad \text{s.t.}\ (1), \quad i = 1,2,\cdots,N, k = 0,1,\cdots,K-1. \tag{7}$$

因此,大系统优化问题(P)的一个局部优化网络(LONN)为:

$$\begin{cases} C_{u_i}^k \dfrac{\mathrm{d}u_i(k)}{\mathrm{d}t} = -\dfrac{\partial L_i(\cdot)}{\partial u_i(k)}, & (8) \\ C_{z_i}^k \dfrac{\mathrm{d}z_i(k)}{\mathrm{d}t} = -\dfrac{\partial L_i(\cdot)}{\partial z_i(k)}, \quad i=1,2,\cdots,N, \quad k=0,1,\cdots,K-1. & (9) \end{cases}$$

其中 $C_{u_i}^k \in R^{m_i \times m_i}, C_{z_i}^k \in R^{r_i \times r_i}$是正定对角阵,为收敛速率控制因子.

令 $$M_i(K)=\frac{1}{2}\|x_i(K)\|_{Q_i(K)}^2, \tag{10}$$

及 $$M_i(k)=\frac{1}{2}\|x_i(k)\|_{Q_i(k)}^2+\frac{1}{2}\|u_i(k)\|_{R_i(k)}^2+\frac{1}{2}\|z_i(k)\|_{S_i(k)}^2+\lambda_i^T(k)z_i(k) \\ -\sum_{j=1}^{N}\lambda_j^T(k)G_{ji}x_i(k), \quad i=1,2,\cdots,N, \quad k=0,1,\cdots,K-1, \tag{11}$$

则 $$\frac{\partial L_i(\cdot)}{\partial u_i(k)}=\frac{\partial M_i(k)}{\partial u_i(k)}+\sum_{j=k+1}^{K}\frac{\partial x_i^T(j)}{\partial u_i(k)}\cdot\frac{\partial M_i(j)}{\partial x_i(j)}. \tag{12}$$

又令 $$p_i(K)=\frac{\partial M_i(K)}{\partial x_i(K)}, \tag{13}$$

及 $$p_i(k)=\frac{\partial M_i(k)}{\partial x_i(k)}+\frac{\partial x_i^T(k+1)}{\partial x_i(k)}p_i(k+1), i=1,2,\cdots,N, \quad k=1,2,\cdots,K-1. \tag{14}$$

整理后得

$$C_{u_i}^k\frac{du_i(k)}{dt}=-\left[R_i(k)u_i(k)+B_i^Tp_i(k+1)\right], i=1,2,\cdots,N, \quad k=0,1,\cdots,K-1. \tag{15}$$

同理可得

$$C_{z_i}^k\frac{dz_i(k)}{dt}=-\left[S_i(k)z_i(k)+\lambda_i(k)+C_i^Tp_i(k+1)\right], \\ i=1,2,\cdots,N, \quad k=0,1,\cdots,K-1. \tag{16}$$

而协调网络(CNN)由下面的微分方程组决定

$$C_{\lambda_i}^k\frac{d\lambda_i(k)}{dt}=z_i(k)-\sum_{j=1}^{N}G_{ij}x_j(k), \quad i=1,2,\cdots,N, \quad k=0,1,\cdots,K-1, \tag{17}$$

其中 $C_{\lambda_i}^k \in R^{r_i\times r_i}$是正定对角矩阵,为收敛速率控制因子.

这样,微分方程组(15—17)一起组成了一个完全连接的大系统动态递阶优化神经网络(LHONN).

3.3 网络的稳定性及其与原问题的等价性

命题 1. 若$U=U^T\in R^{n\times n}$是正定(半正定)的,$W\in R^{n\times m}$,则 $W^TUW=H\in R^{m\times m}$正定(半正定),当且仅当 $\mathrm{rank}(W)=m$.

定理 1. 在满足本文的设计条件时,神经网络(LHONN)是完全稳定的,即对于网络的任意初始状态,该神经网络均能收敛于对应大系统某个初始状态的一个稳定的平衡点.

证明. 对于递阶优化神经网络(LHONN),定义其 Lyapunov 函数为

$$V(\cdot)=\sum_{i=1}^{N}\sum_{k=0}^{K-1}V_i(\cdot), \tag{18}$$

式中 $$V_i(\cdot)=\frac{1}{2}\|\nabla_{u_i(k)}L(\cdot)\|^2+\frac{1}{2}\|\nabla_{z_i(k)}L(\cdot)\|^2+\frac{1}{2}\|\nabla_{\lambda_i(k)}L(\cdot)\|^2. \tag{19}$$

就 Lyapunov 函数(18)对时间 t 求导得

$$\frac{\mathrm{d}V(\cdot)}{\mathrm{d}t}=\sum_{i=1}^{N}\sum_{k=0}^{K-1}\frac{\mathrm{d}V_i(\cdot)}{\mathrm{d}t},\tag{20}$$

不失一般性,令式(15—17)中的速率控制因子为单位阵,同时注意到

$$\frac{\partial L(\cdot)}{\partial \boldsymbol{u}_i(k)}=\frac{\partial L_i(\cdot)}{\partial \boldsymbol{u}_i(k)},\frac{\partial L(\cdot)}{\partial \boldsymbol{z}_i(k)}=\frac{\partial L_i(\cdot)}{\partial \boldsymbol{z}_i(k)},\left(\frac{\partial L^2(\cdot)}{\partial \boldsymbol{\lambda}_i(k)\partial \boldsymbol{z}_i(k)}\right)^T=\frac{\partial L^2(\cdot)}{\partial \boldsymbol{z}_i(k)\partial \boldsymbol{\lambda}_i(k)},$$

并且 $\dfrac{\partial L^2(\cdot)}{\partial \boldsymbol{\lambda}_i(k)\partial \boldsymbol{u}_i(k)},\dfrac{\partial L^2(\cdot)}{\partial \boldsymbol{u}_i(k)\partial \boldsymbol{\lambda}_i(k)},\dfrac{\partial L^2(\cdot)}{\partial \boldsymbol{\lambda}_i(k)\partial \boldsymbol{\lambda}_i(k)}$均为零矩阵.

由于 $\dfrac{\partial L_i^2(\cdot)}{\partial \boldsymbol{u}_i(k)\partial \boldsymbol{u}_i(k)}=R_i(k)+B_i^T\dfrac{\partial \boldsymbol{p}_i(k+1)}{\partial \boldsymbol{u}_i^T(k)},\dfrac{\partial L_i^2(\cdot)}{\partial \boldsymbol{z}_i(k)\partial \boldsymbol{u}_i(k)}=C_i^T\dfrac{\partial \boldsymbol{p}_i(k+1)}{\partial \boldsymbol{u}_i^T(k)},$

$$\frac{\partial L_i^2(\cdot)}{\partial \boldsymbol{u}_i(k)\partial \boldsymbol{z}_i(k)}=B_i^T\frac{\partial \boldsymbol{p}_i(k+1)}{\partial \boldsymbol{z}_i^T(k)},\frac{\partial L_i^2(\cdot)}{\partial \boldsymbol{z}_i(k)\partial \boldsymbol{z}_i(k)}=S_i(k)+C_i^T\frac{\partial \boldsymbol{p}_i(k+1)}{\partial \boldsymbol{z}_i^T(k)}.$$

又 $\dfrac{\partial \boldsymbol{p}_i(k+1)}{\partial \boldsymbol{u}_i^T(k)}=\Lambda_iB_i,\dfrac{\partial \boldsymbol{p}_i(k+1)}{\partial \boldsymbol{z}_i^T(k)}=\Lambda_iC_i$,这里 $\Lambda_i=\sum\limits_{l=0}^{K-k-1}(A_i^l)^TQ_i(k+l+1)A_i^l$.

整理后可得

$$\frac{\mathrm{d}V_i(\cdot)}{\mathrm{d}t}=-\left[\left(\frac{\partial L_i(\cdot)}{\partial \boldsymbol{u}_i(k)}\right)^T\left(\frac{\partial L_i(\cdot)}{\partial \boldsymbol{z}_i(k)}\right)^T\right]\begin{bmatrix}R_i(k)+B_i^T\Lambda_iB_i & B_i^T\Lambda_iC_i\\ C_i^T\Lambda_iB_i & S_i(k)+C_i^T\Lambda_iC_i\end{bmatrix}\begin{bmatrix}\dfrac{\partial L_i(\cdot)}{\partial \boldsymbol{u}_i(k)}\\ \dfrac{\partial L_i(\cdot)}{\partial \boldsymbol{z}_i(k)}\end{bmatrix}.\tag{21}$$

注意到 $\begin{bmatrix}R_i(k)+B_i^T\Lambda_iB_i & B_i^T\Lambda_iC_i\\ C_i^T\Lambda_iB_i & S_i(k)+C_i^T\Lambda_iC_i\end{bmatrix}=\begin{bmatrix}R_i(k) & \Theta\\ \Theta & S_i(k)\end{bmatrix}+\begin{bmatrix}B_i & \Theta\\ \Theta & C_i\end{bmatrix}^T\begin{bmatrix}\Lambda_i & \Lambda_i\\ \Lambda_i & \Lambda_i\end{bmatrix}\begin{bmatrix}B_i & \Theta\\ \Theta & C_i\end{bmatrix},$

式中 Θ 表示零矩阵(不一定为方阵). 因 $\begin{bmatrix}\Lambda_i & \Lambda_i\\ \Lambda_i & \Lambda_i\end{bmatrix}\in R^{2n_i\times 2n_i}$为半正定,又 $\mathrm{rank}\begin{bmatrix}B_i & \Theta\\ \Theta & C_i\end{bmatrix}=m_i+r_i$,由命题 1 可得 $\begin{bmatrix}B_i^T\Lambda_iB_i & B_i^T\Lambda_iC_i\\ C_i^T\Lambda_iB_i & C_i^T\Lambda_iC_i\end{bmatrix}$为半正定,又 $\begin{bmatrix}R_i & \Theta\\ \Theta & S_i\end{bmatrix}$为正定,则 $\begin{bmatrix}R_i(k)+B_i^T\Lambda_iB_i & B_i^T\Lambda_iC_i\\ C_i^T\Lambda_iB_i & S_i(k)+C_i^T\Lambda_iC_i\end{bmatrix}$为正定,综之可得

$$\mathrm{d}V_i(\cdot)/\mathrm{d}t\leqslant 0,\quad i=1,2,\cdots,N,\quad k=0,1,\cdots,K-1,\tag{22}$$

即 $$\mathrm{d}V(\cdot)/\mathrm{d}t\leqslant 0.\tag{23}$$

若(23)式中等式成立,当且仅当

$$\frac{\partial L_i(\cdot)}{\partial \boldsymbol{u}_i(k)}=0,\frac{\partial L_i(\cdot)}{\partial \boldsymbol{z}_i(k)}=0,\frac{\partial L(\cdot)}{\partial \boldsymbol{\lambda}_i(k)}=0,i=1,2,\cdots,N,\quad k=0,1,\cdots,K-1.\tag{24}$$

所以,神经网络(LHONN)是完全稳定的,式(24)对应于网络唯一稳定的平衡点. □

根据网络的设计与推理过程可知,该网络的稳定点与原问题的最优解相对应.

3.4 网络实现

本文所设计的神经网络(LHONN)的特点之一是易于硬件集成化实现,其协调网络(CNN)和局部优化网络(LONN)是一有机结合的整体,可以方便地做成一个基于某大系统

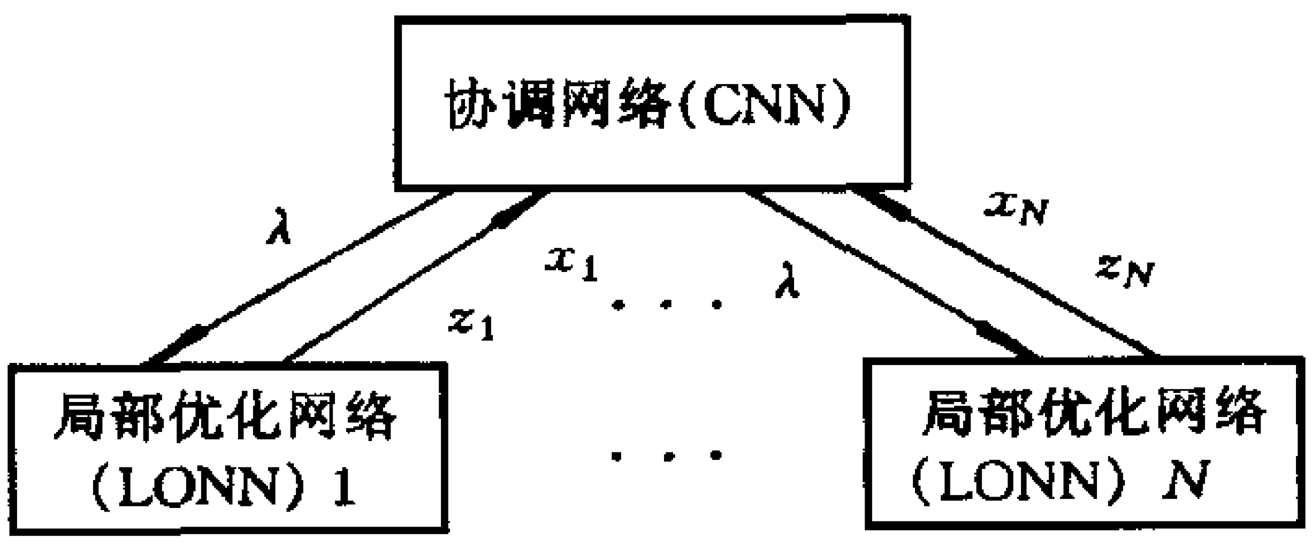
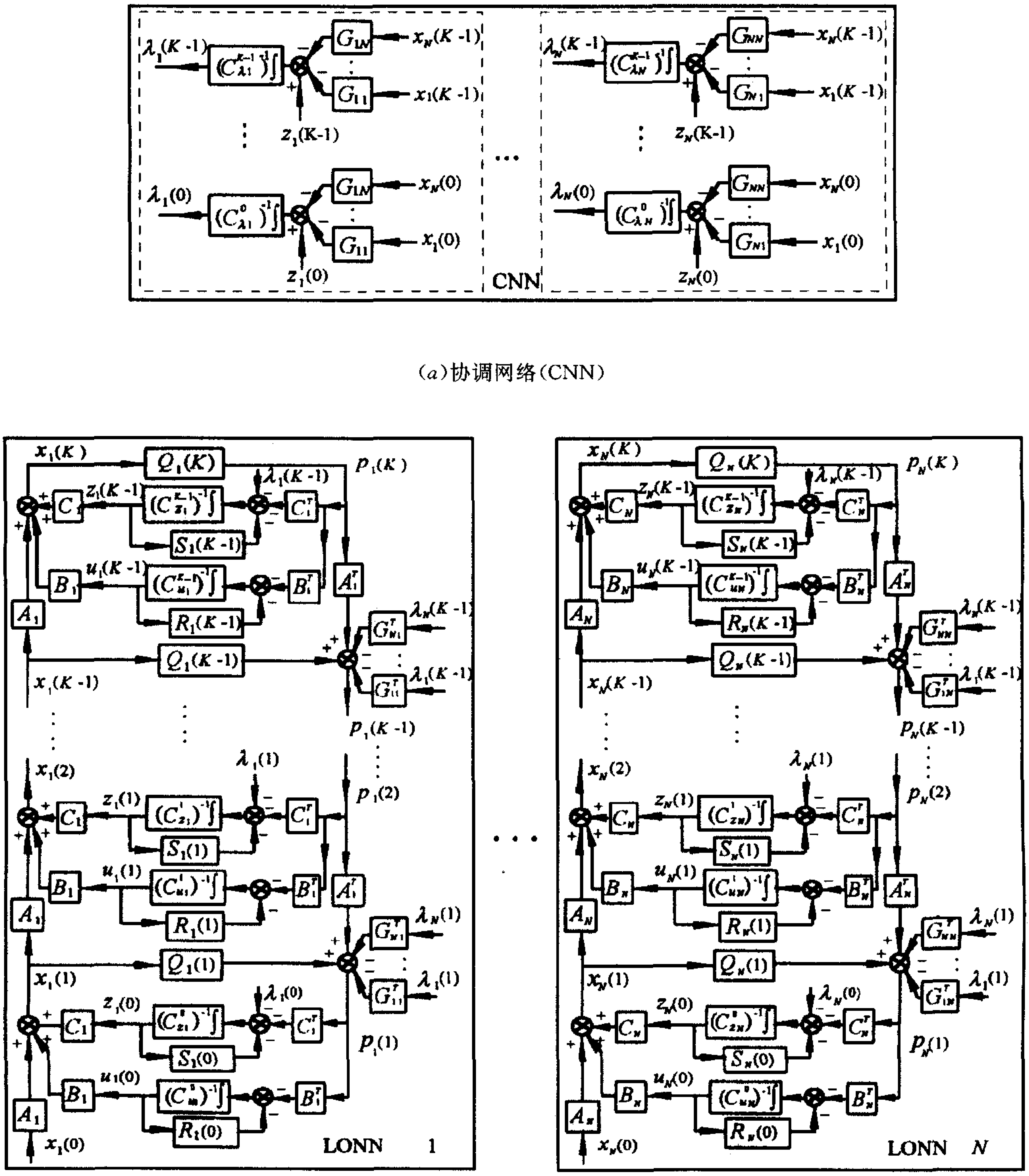


图 1　优化神经网络(LHONN)结构

模型的优化求解器电路.该网络的子网络之间连接的基本结构如图 1 所示，基于问题(P)的神经网络的电路实现原理如图 2 所示，从图 1 和图 2 不难建立起该神经网络的细化的完全连接图.



(a)协调网络(CNN)

(b)局部优化网络(LONN)

图 2 神经网络(LHONN)电路实现原理图

4 结语

由于神经网络具有并行处理功能，以具有递阶结构的神经网络求解大系统的优化问题，避免了传统优化方法在计算时所困扰的"维数灾"及迭代等待等问题，极大地提高了求

解效率,尤其是本文所给出的优化神经网络,具有全集成化的特点,适宜于硬件实现. 一个电路实现的神经网络,其优化计算可以瞬间完成,因此,该神经网络对工业大系统的实际应用极富有意义.

参　考　文　献

1 Hopfield J J, Tank D W. "Neural"computation of decisions in optimization problems. *Biol. Cybern.*, 1985, **52**: 141—152

2 Jamshidi M. Large-scale systems modeling and control. NewYork: North-Holland, 1983

3 李人厚,邵福庆. 大系统的递阶与分散控制. 西安:西安交通大学出版社, 1986

4 Huang X, Wu C P. Neural networks for optimal control problems. In: *Proc. Int. Joint Conf. on Neural Networks*, Beijing. 1992, **3**: 123—126

5 Moore K L, Naidu S. Linear quadratic regulation using neural networks. In: *Proc. Int. Joint Conf. on Neural Networks*, Seattle. 1991, 735—739

6 黄琳. 系统与控制理论中的线性代数. 北京:科学出版社,1986

A NEURAL NETWORK BASED METHOD FOR DYNAMICAL HIERARCHICAL OPTIMIZATION OF A CLASS OF LARGE-SCALE SYSTEMS

HOU ZENGGUANG　GAO HUIQI　WU CANGPU

(*Dept. of Automatic Control, Beijing Institute of Technology, Beijing* 100081)

Abstract A neural network for dynamical hierarchical optimization of discrete-time large-scale systems, i. e., LHONN, is presented in the paper. It is a fully integrated network with the following principal characteristics: 1) the dynamic equations of the subsystems are imbedded into the local optimization networks, which results in the lower dimension of the neural network, so it is easy for implementation; 2) the coordination neural network (CNN) and the local optimization neural networks (LONN) work simultaneously to seek for the optimal solution of the system, which leads to high speed of problem solving. Thus the LHONN is more suitable for real-time optimization problems.

Key words Large-scale systems, dynamical systems, hierarchical optimization, neural networks

侯增广 1969 年生. 1991 年、1993 年在燕山大学自动化系分别获学士、硕士学位,现为北京理工大学自动控制系博士生. 目前的研究兴趣是系统优化的理论与方法,神经网络技术,机器人控制.

吴沧浦 照片及简介见本刊第 20 卷第 6 期.