

大规模光电混合神经网络中的 可编程拓扑重构*

张延焯 常胜江 申金媛 张文伟

(南开大学现代光学研究所, 光学信息技术科学开放研究实验室, 天津 300071)

摘 要 提出一种在大规模光电混合神经网络系统中实现可编程拓扑重构的有效方法。文中介绍了这种方法的原理、运用技术以及采用这种方法在已有 NP1024 光学/数字神经网络处理器中运行单层反馈型和多层前馈型等不同网络拓扑结构的实验结果。

关键词 光学神经网络, 网络拓扑结构, 可编程拓扑重构。

1 引 言

目前, 人工神经网络的硬件实现主要有三种途径: 虚拟神经网络、微电子神经网络芯片和光神经网络。三种途径都存在各自的长处和不足。光神经网络在网络规模以及相乘与求和运算的速率方面已被公认是有竞争力的, 但是灵活性差是它的主要的弱点之一, 其中包括可学习性和网络拓扑的重构性等, 这些都是实现通用型人工神经网络或神经网络计算机的必要条件。

为了满足神经网络硬件实现的要求, 将不同方法结合起来, 构成优势互补的混合系统, 是当前神经网络硬件实现的主要研究趋势。例如可编程光电混合或光学/数学神经网络系统就是将光学和电子电路(或计算机)相结合而构成的。在光电混合神经网络系统中, 光的作用主要是大规模的加权互连(包括相乘与求和), 而电的作用则主要是非线性运算和编程控制。过去, 对光电混合神经网络系统的研究工作很多, 主要是关于系统的体系结构^[1~3]、规模极限^[4]、双极性光互连的编码方法^[5~7]以及关键器件^[8, 9]等方面。

本文提出一种在大规模光电混合神经网络系统中实现可编程拓扑重构的有效方法, 即: 利用光互连的空间可交叉性和计算机处理的可编程性, 不要求高分辨率和高速驱动的空间光调制器(完全符合要求的这种器件目前尚不存在)便可根据应用的要求, 在同一系统中实现单层反馈型、多层前馈和前馈与反馈混合型等不同拓扑结构的神经网络。虽然这种方法是针对特定的 NP1024 光学数字神经网络处理器提出的, 但方法的基本原理是广泛适用的。

* 国家教委博士点基金和国家自然科学基金资助的课题 69587005。

收稿日期: 1997-08-08

2 NP1024 光电混合神经网络处理器

NP1024 神经网络处理器是一个二维光学/数字神经网络系统。图 1 是它的原理示意图。其中的光学部分除光源 S 外, 主要功能为: 光学张量-矩阵相乘和双极性互连张量模板 (IWM) 的光束方向编码。微计算机的功能除系统控制和液晶开头阵列(LCSA) 的驱动外, 还有正负加权求和量的相减与神经元的非线性函数输出。光电之间的接口由 CCD 或光探测器阵列(PDA) 和图像接口板(A/D) 来完成。1024 个神经元的状态排成 32×32 的方阵 X_{ij} , 由液晶开关阵列来表示(开为 1 态, 关为 0 态)。神经元之间的互连权重张量(W_{ij}) $_{kl}$ 也按对每一神经元的互连权重, 排成 32×32 的子矩阵 W_{ij} , 然后再按对应神经元的位置 kl 排成 32×32 的互连权重模板(见图 2), 此互连权重模板由电子束扫描在透明片上制成。因此, 所有神经元之间的加权互连运算便可通过矩阵 X_{ij} 与张量(W_{ij}) $_{kl}$ 的内乘来实现。图 1 中的正互连模板 IWM⁺ 和负互连模板 IWM⁻ 通过光束方向编码合成为双极互连张量模板的光路, 将两者重迭成像于小透镜阵列(LA) 的前焦面处。透镜阵列与透镜 L₅ 和 L₆ 所构成的光学成像系统, 可在 CCD 或光探测器阵列的光敏面上形成透镜阵列后焦面上因两种光束方向而形成的所有双光点的像。它们的光强给出正互连模板(IWM⁺) 和负互连模板(IWM⁻) 中所有对应子矩阵 W_{ij}

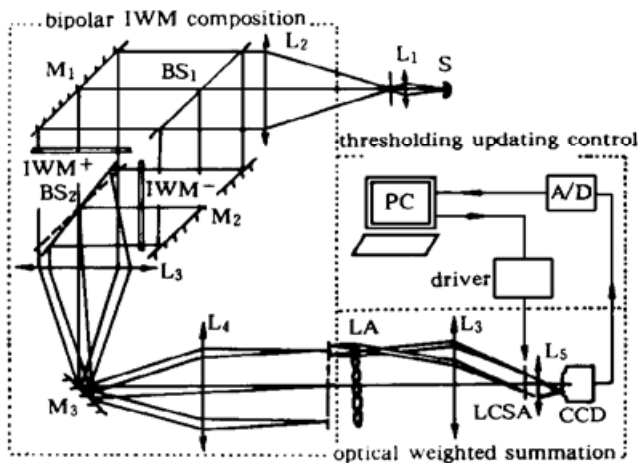


Fig. 1 The principle diagram of NP1024 neuro-processor system

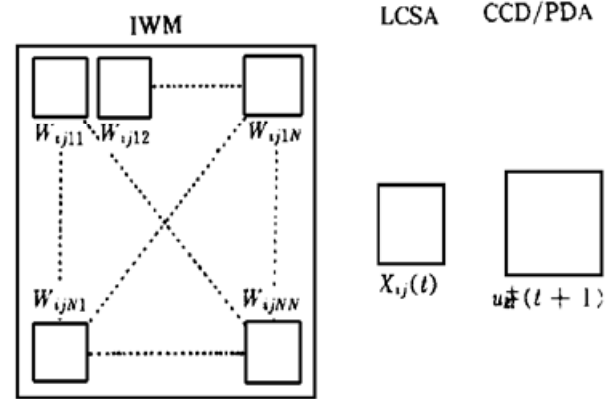


Fig. 2 The arrangement of interconnection weights (W_{ij}) $_{kl}$ in the IWM mask

与矩阵 X_{ij} 相内乘的结果。即:

$$u_{kl}^{\pm}(t+1) = \sum_i \sum_j (W_{ij}^{\pm})_{kl} X_{ij}(t), \quad (k, l = 1, 2, \dots, 32).$$

这些光点的光强被 CCD 接收后, 经 A/D 转换在计算机中实现相减和非线性运算, 得到神经元在下一个时刻 ($t + 1$) 的状态, 即:

$$X_{kl}(t+1) = f[u_{kl}^+(t+1) - u_{kl}^-(t+1)], \quad (k, l = 1, 2, \dots, 32).$$

式中 $f(x)$ 为所选用的非线性函数。应当说明, 这里所谓正互连模板 IWM⁺ 是指其中的每一子矩阵元 W_{ij}^+ 只取相应 W_{ij} 中为正的, 其余置零, 即:

$$W_{ij}^+ = W_{ij}, \text{ 若 } W_{ij} > 0; \text{ 或 } W_{ij}^+ = 0, \text{ 若 } W_{ij} \leq 0.$$

同样, 对负互连模板 IWM⁻ 则有: $W_{ij}^- = |W_{ij}|$, 若 $W_{ij} < 0$; 或 $W_{ij}^- = 0$, 若 $W_{ij} \geq 0$ 。

3 拓扑重构与互连模板

拓扑重构方法的关键是互连模板的设计和相应的计算机软件。分别讨论如下:

1) 任何单层反馈型网络, 只要其神经元总数 $N \leq 1024$, 便可编制如图 2 所示的互连张

量模板直接在系统中实现;

2) 多个单层反馈型网络。只要各网络神经元数之和小于等于 1024, 都可以同时在系统中运行。这时, 互连张量模板中各个网络的神经元在液晶开关阵列中所占的空间位置要事先给定, 与此同时, 互连张量模板中相应的子矩阵元的排布也要与该网络在液晶开关阵列中所占的空间位置严格对应, 如图 3 所示。为了简便, 图中仅以 5×5 的系统规模为例来说明模板设计的原理, 图中三个单层反馈型网络的规模分别为: $N_1 = 13, N_2 = 8, N_3 = 4$ 。当规模为 32×32 时, 方法一样。

3) 多层前馈型网络。只要网络中属于光层的神经元总数小于等于 1024, 也可在此系统中实现。所谓光层, 是指该多层前馈型网络中这样的神经元层, 其中的神经元数大于某一临界值 N_T , 从这时起, 完全在计算机中完成一个子网程(即一层神经元的加权求和与非线性输出和馈送)的运行时间, 反而比在光电混合系统中运算得更快。每一个光电混合神经网络系统, 都存在其特定的 N_T 值, 它是由光电混合系统完成一个子网程运算所需要的时间(通常此时间是神经元数 N 的线性函数, 见图 4 曲线 A)与计算机完成同样运算所需要的时间(通常此时间是神经元数 N 的二次幂函数, 见图 4 曲线 B)相比较而定, 如图 4 所示。

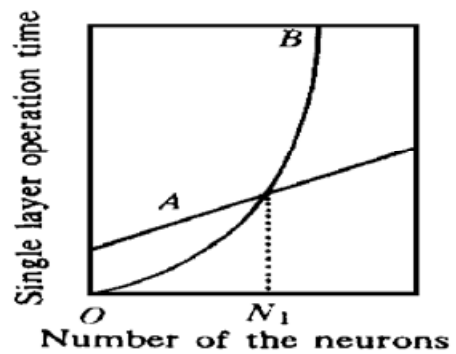
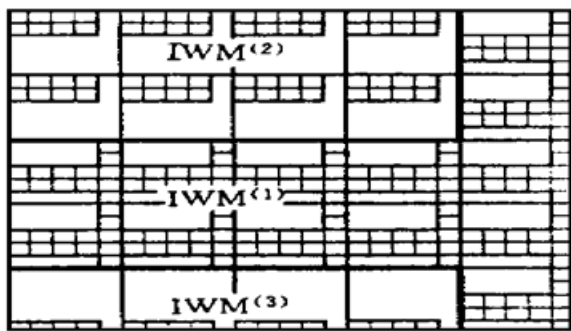


Fig. 3 The arrangement of IWM for simultaneous running multiple single layer networks

Fig. 4 The cycle time of hybrid and virtual neural layers vary with neuron numbers

为实现一个多层前馈型网络, 首先由网络结构和系统的 N_T 值确定该网络中的光层(其他神经元层可称电层), 只要属于光层的神经元总数小于等于 1024, 便可在液晶开关阵列中设定各层神经元的位置。这时, 互连张量模板中相应的互连权重子矩阵的排布和各子矩阵中矩阵元的排布, 也要与各层在液晶开关阵列中设定的位置相对应。与图 3 中的排布规则不同的是, 每一层神经元互连权重子矩阵的排布要与该层神经元在液晶开关阵列中的位置相对应; 而各子矩阵中矩阵元的排布却要与上一层神经元在液晶开关阵列中的位置相对应, 这是由多

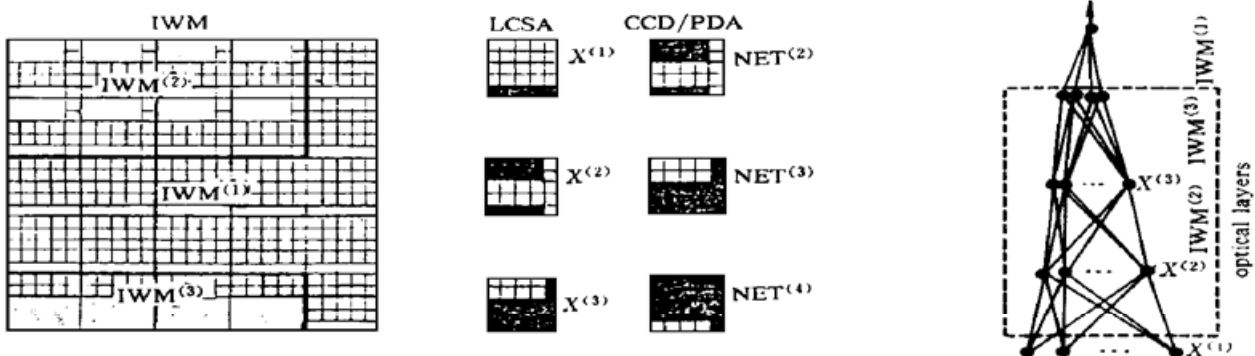


Fig. 5 The arrangement of IWM for a multi-layered feedforward network

层前馈型网络仅在层间连接的特点所决定的。图 5 给出一个结构为 20- 13- 8- 4- 1 的多层前馈网络的互连张量模板排布的示意图。同样为了简便, 图中仅以 5×5 的系统规模为例来说明模板设计的原理, 并假定系统的 N_T 值等于 3。显然, 此网络中光层神经元的总数为 $13 + 8 + 4 = 25$, 恰好等于系统的规模, 可以实现。这里, 输入层神经元数为 20, 但这些输入神经元不要求相减和非线性运算, 因此不占用系统的非线性神经元数。至于输出层, 由于其神经元数为 1, 小于 N_T 值, 因而完全在计算机中运行即可, 也不占用光电混合系统的神经元。与图 3 不同的是, 图 5 中每一层互连权重子矩阵的排布与各子矩阵中矩阵元的排布是不同的。

4 初步实验结果

采用前述拓扑重构方法, 在 NP1024 中先后实现了 Hopfield 网络联想存储器(单层反馈型)和神经网络目标分类系统(前馈与反馈混合网络)。前者的神经元数为 1024, 输入和再现的模式如图 6 所示。后者的网络结构为 400- 8- 8- 4, 目前该系统中的液晶开关阵列器件速度还不够快, 在计算机为 486/66 时, 系统的实测 N_T 值为 360。因此, 该网络仅第一层为光层, 系统的互连张量模板有较大的富余空间。该目标分类系统的目标模式为平面任意旋转的四种飞行器(见图 7)。

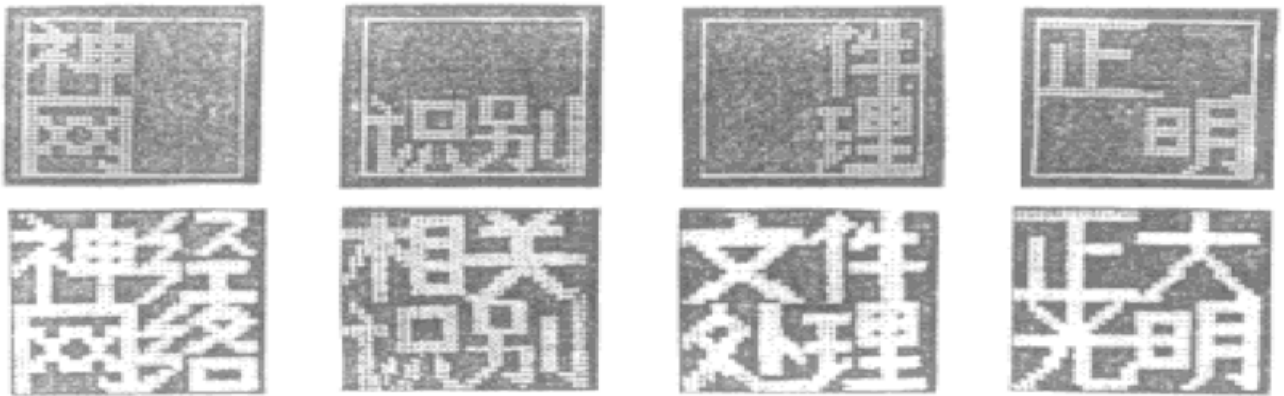


Fig. 6 The input and retrieval patterns of the associative neural memory



Fig. 7 Experimental examples of the targets and hidden neurons of the multilayered neural classification system

结 论 本文提出一种通过互连模板排布和在线计算机编程处理, 在大规模光电混合神经网络系统中实现可编程拓扑重构的方法。同时, 文中还讨论了运用这些方法的一些原则和技巧。该方法不要求有高分辨率和高速驱动的空间光调制器(完全符合要求的这种器件目前尚不存在)便可根据应用的要求, 在同一系统中实现单层反馈型、多层前馈型或前馈与反馈混

合型等不同拓扑结构的神经网络。当然,为了实现自学习和操作上的方便,高性能的空间光调制器是不可缺少的。不过,即使今后有了合乎要求的空间光调制器,本方法中关于互连模板排布和利用的原则和技巧仍然是有意义的。

本文所提出的方法揭示出神经网络系统中可编程完全互连的神经元数(即网络规模)的重要意义除了通常所理解的并行度高之外,还可以被利用来实现网络的拓扑重构。同时,作者在系统中利用了光和电在神经网络硬件实现中的多种互补性,如大规模互连与运行灵活性,加权求和与非线性运算,神经元数在 N_T 值上下时光与电神经网络实际运行速度的比较等等。考虑到实用神经网络需求的多样性,光神经网络和虚拟神经网络以及神经网络芯片的优势都是不可缺少的。由此可见,光电混合方法在神经网络硬件实现中并非只是一种权宜之计。

参 考 文 献

- [1] B. K. Jenkins, A. R. Tanguay, Jr., Optical architectures for neural network implementations. *The Handbook of Brain Theory and Neural Networks*. M. A. Arbib (Ed.), The MIT Press, 1995, 673~677
- [2] S. Q. Gao, Y. X. Zhang, J. W. Yang *et al.*, Coaxial architecture of an optical neural network with lenslet array. *Opt. Lett.*, 1994, **19**(24) : 2155~ 2157
- [3] Y. X. Zhang, S. J. Chang, Z. B. Feng *et al.*, Multimode operation in an optical/digital neuroprocessor with 1024 neurons. *Opt. Engng.*, 1996, **35**(8) : 2132~ 2135
- [4] T. Sakano, K. Noguchi, T. Matsumoto, Optical limits for spatial inter-connection networks using 2-D optical array devices. *Appl. Opt.*, 1990, **29**(8) : 1094~ 1100
- [5] J. W. Yang, S. Q. Gao, Y. X. Zhang *et al.*, 2D optical fully bipolar neural network using colour encoding. *IEE. Proc. Optoelectron*, 1995, **142**(2) : 121~ 124
- [6] M. Kranzendorf, B. J. Bigner, L. Zhang *et al.*, Optical connectionist machine with polarization-based bipolar weight values. *Opt. Engng.*, 1989, **28**(8) : 844~ 848
- [7] F T. S. Yu, X. Yang, T. Lu, Space-time sharing optical neural network. *Opt. Lett.*, 1991, **16**(4) : 247~ 249
- [8] A. R. Tanguay, Jr., B. K. Jenkins, Optical components for neural network implementations. *The Handbook of Brain Theory and Neural Networks*, M. A. Arbib (Ed.), The MIT Press, 1995, 677~682
- [9] D. C. Burns, I. Underwood, A. F. Murray *et al.*, Design and Characterisation of an optoelectronic neural chip. *Opt. Commun.*, 1996, **127**(3) : 313~ 324

Reconfiguration of Topology in an Opto-Electronic Neural System

Zhang Yanxin Chang Shengjiang Shen Jinyuan Zhang Wenwei

(Nankai University, Institute of Modern Optics, Optical Information Science Laboratory, Tianjin 300071)

(Received 8 August 1997)

Abstract An efficient method of programmable reconfiguration of the topology in an optical/digital neural system with large scale is proposed. The principle and application technique of the method are described. And the experimental results of operation of a single layer feedback network as well as a multi-layer feedforward network in a NP1024 optical/digital neuroprocessor by using the proposed method are reported.

Key words optical neural networks, topology of neural network structures, reconfiguration of the topology.