

基于正交化算法的三值互连神经网络模型*

常胜江 杨建文 高胜泉 申金媛 张延焯
(南开大学现代光学研究所, 天津 300071)

摘 要 提出了以信息损失最少为原则的三值(±1, 0)互连权重编码方法, 这种编码方法比以前的三值权重编码方法显著地提高了神经网络的性能。由于互连权重只有三值, 恰恰弥补了光互连精度不高的缺点, 易于光学实现。

关键词 神经网络, 三值编码。

1 引 言

神经网络联想存储器中的信息是以互连矩阵的形式分布地存储在网络之中, 因此, 网络的信息编码过程也就是求解互连权重的过程。互连权重的求解方法是一种学习算法, 其目标是寻找适当的网络结构和连接权重, 使尽可能多的记忆模式成为网络的稳定吸引子, 每一吸引子皆有尽可能大的吸引域, 并且便于硬件实现。

Hopfield 首次提出用 Hebb 规则给出单层反馈网络的互连权重的求解公式:

$$W_{ij} = \sum_{m=1}^M S_i^{(m)} S_j^{(m)} - gM\delta_{ij}$$

($i, j = 1, 2, \dots, N$)

式中 $S^{(m)}$ 为 M 个 N 维记忆矢量 ($m = 1, 2, \dots, M$), 其分量为 $S_1^{(m)}, S_2^{(m)}, \dots, S_N^{(m)}$, $\delta_{ij} = 1$ (当 $i = j$) 或 0 (当 $i \neq j$), g 为网络参数, $0 \leq g \leq 1$ 。这种互连权重的求解方法, 显示了神经网络的分布存储和联想、容错, 内容寻址等特点, 但是存储容量较低且与记忆模式的具体形式有关, 并且连接权重的动态范围与记忆模式的数目成正比, 当 M 较大时, 难以用光学互连技术实现。

多年来, 对 Hopfield 模型的改进一直围绕着信息编码问题, 出现了多种互连权重的编码方法, 其中尤以互连权重值限定为 ±1 和 0 的编码方案最为光学界重视。这种实现联想存储的最简单互连形式恰好弥补了光互连精度不高的缺点。目前, 已经出现了许多种互连权重值限定为 ±1 或 0 的编码方案及其优化方法, 如 Clipped 模型^[1], IPA 模型^[2], Monte Carlo 优化模型^[3], 神经网络优化模型^[4, 5]等等。虽然这些模型易于光学实现, 但存储容量及内容寻址的容错能力较低。因此, 如何得到存储容量大、寻址能力强的三值(±1, 0)互连矩阵一直是光学神经网络研究的重要内容。

* 本工作得到国家自然科学基金及天津市 21 世纪青年基金资助。

收稿日期: 1995 年 4 月 24 日; 收到修改稿日期: 1995 年 11 月 27 日

本文针对 Hopfield 模型的存储容量及寻址能力较低的缺点，引入正交化算法，并对其高动态范围的互连权重进行三级截值，实现了互连权重的三值(±1, 0)编码。在理论上指出了截值点的寻找方法。计算机模拟结果表明，这种三值互连权重的编码方法能比以前的三值权重编码方法显著地提高神经网络的性能。

2 正交化算法

对于 Hopfield 神经网络模型，其存储容量在神经元数 $N \rightarrow \infty$ 时的渐近值为 $a_c = M/N \approx 0.138$ ，提高其存储容量及寻址能力的方法很多，正交化算法就是其中的一种方法。给定一组双极矢量 $V^{(m)}$ ($i = 1, 2, \dots, N, m = 1, 2, \dots, M$)，通过 Gram-Schmidt 正交化处理，可以得到正交的矢量^[6]：

$$V^{*(m)} = V^{(m)} - \sum_{n_0=1}^{m-1} \frac{[V^{(m)} \cdot V^{*(n_0)}]}{\|V^{*(n_0)}\|} V^{*(n_0)} \tag{2.1}$$

互连矩阵 W 可以表示成如下形式：

$$W_{ij} = \sum_{m=1}^M \frac{V_i^{*(m)} V_j^{*(m)}}{\|V^{*(m)}\|} \tag{2.2}$$

Gram-Schmidt 方法是一种提取模式特征的方法，提取的特征通过权重编码的形式存储在网络中，这可以直接通过互连模板看出，图 1 所示的是存储 N 、 K 、 U (单极) 三个模式时正交化 Hopfield 模型的互连权重模板，从正交化模板(正部分)中可以清楚地看出 N 、 K 、 U 三个英文字母。也就是说正交化方法在存储信息时并没有去掉模式间的重叠部分，从信息编码的角度来讲，Gram-Schmidt 方法是一种既能提取模式特征信息，又能保存模式整体信息的一种算法。在文献[7]中，对正交化 Hopfield 模型进行了理论分析，结果表明正交化算法可以大大提高 Hopfield 网络的存储容量，理论极限值 $a_c = M/N = 1$ ，同非正交的 Hopfield 网络比较，寻址能力也有很大的提高。

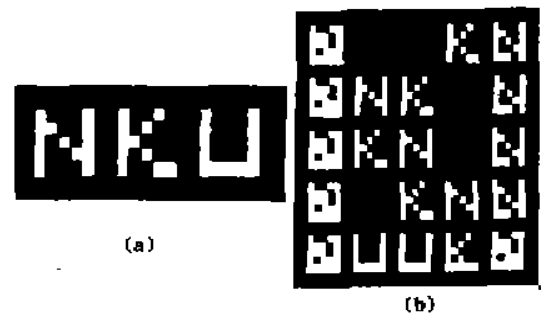


Fig. 1 Interconnection weight mask for storing N 、 K 、 U
(a) Stored patterns, (b) Interconnection Weight mask

3 基于正交化算法的三值权重编码

设由式(2.1)、(2.2)求得的正交化 Hopfield 神经网络模型的互连矩阵元为 W_{ij} ，截值后得到的互连矩阵元为 W'_{ij} ，截断点 m_0 为一正常数。令互连矩阵元 W'_{ij} 满足下式：

$$W'_{ij} = \begin{cases} m_0 & W_{ij} > m_0 \\ 0 & |W_{ij}| \leq m_0 \\ -m_0 & W_{ij} < -m_0 \end{cases} \tag{3.1}$$

采用误差分析的方法来确定 m_0 的值，互连矩阵元 W_{ij} 经(3.1)式处理后引起的误差 ΔW_{ij} 为：

$$\Delta W'_{ij} = W'_{ij} - W'_{ij} = \begin{cases} W'_{ij} - m_0 & W'_{ij} > m_0 \\ W'_{ij} & |W'_{ij}| \leq m_0 \\ -W'_{ij} - m_0 & W'_{ij} < -m_0 \end{cases}$$

对所有互连矩阵元进行截断后总的误差 E 为:

$$E = \sum_{i,j} |\Delta W'_{ij}| \quad (3.2)$$

最佳的截取点 m_0 就是使(3.2)式最小的那个点。由(2.1)、(2.2)两式求得 W'_{ij} 后代入(3.2)式,利用插值法,很容易求得 m_0 的值。(3.2)式的意义是让截值后的互连矩阵逼近正交化互连矩阵,由于网络记忆的信息是通过正交化互连权重分布地存储在神经网络中的,因此上面互连权重的截取方法是信息损失最少的方法。

通过下面的分析可以看到 W' 就是所要求的三值($\pm 1, 0$)互连矩阵。从(3.1)式可知,截值后的互连矩阵 W' 写为:

$$W' = [W'_{ij}] = [\{\pm m_0, 0\}] \quad (3.3)$$

$[\{\pm m_0, 0\}]$ 表示互连矩阵 W' 的所有元素为 $\pm m_0$ 或 0。

当以矢量 $V^{(n)}$ 进行内容寻址时,迭代过程可写为:

$$V^{(n)}(t+1) = \text{sgn} [W' V^{(n)}(t)] = \text{sgn} \{[\{\pm m_0, 0\}] V^{(n)}(t)\} \quad (3.4)$$

由于神经元激活函数取符号函数,所以从互连矩阵中提出 m_0 不会影响迭代的结果,即:

$$V^{(n)}(t+1) = \text{sgn} \{[\{\pm m_0, 0\}] V^{(n)}(t)\} = \text{sgn} \{[\{\pm 1, 0\}] V^{(n)}(t)\} \quad (3.5)$$

此时互连矩阵相当于三值矩阵,也即:

$$W' = [\{\pm 1, 0\}]$$

因此(3.1)式就是对存储模式的 $\pm 1, 0$ 三值互连权重编码。

4 模拟结果

本文采用计算机模拟对上述编码方法的关联存储模型进行了分析,结果如下:

4.1 存储容量 a_c^*

一般来说,网络的联想效果由输出矢量的出错率表示,出错率定义为:

$$q = 1 - \frac{\sum_{i=1}^N V_i U_i}{N} \quad (4.1)$$

U 与 V 分别为存储矢量与输出矢量,取一模式 U (U 本身或含有噪声的 V) 作为初始模式,经状态叠代达到稳定后,即可按(4.1)式计算稳定模式 V 与 U 之间的出错率 q 。定义当统计平均出错率 $\langle q \rangle < 0.015$ 时, $a_c = M/N$ 的值作为该神经网络的存储容量,记为 a_c^* ,最大的存储模式数为 $M = a_c^* N$ 。由于受微机速度及容量的限制,本文只对神经元数 $N = 25$ 时的 a_c^* 值进行了仿真测量。图 2 为正交化 Clipped 模型(记作 OC Model,对应于 $m_0 = 0$)和三级截值模型(记作 TC Model)对于 500 个随机模式的统计平均出错率 $\langle q \rangle$ 与参数 $a_c = M/N$ 的关系曲线。从图 2 可知,OC 模型的 $a_c^* = 0.16$,而 TC 模型可达 $a_c^* \approx 0.52$,是 OC 模型存储容量的三倍多,同 Hopfield 模型的存储容量 $a_c = 0.16$ ^[3] 相比存储容量提高了 3.15 倍。

4.2 寻址能力

a) 理论分析: 神经网络的容错性是有别于其它信息处理系统最主要的特点,因此网络优劣的判别标准之一就是它的寻址能力。路明哲等人^[5]分析了网络迭代时模拟输出量与存储能力、容错能力的关系,结果表明,内容寻址的模拟输出量与阈值差别越大,网络的存储能

力和寻址能力就越强。

从(3.5)式可知网络迭代时，互连矩阵 $[\{\pm m_0, 0\}]$ 可等效于三值互连权重矩阵 $[\{\pm 1, 0\}]$ ，等效过程是从矩阵中提出 m_0 ，此时模拟输出量与阈值(存储矢量为双极时通常为 0)之差的变化为：

$$[\{\pm m_0, 0\}]V^{(m)}(t) \rightarrow [\{\pm 1, 0\}]V^{(m)}(t) \tag{4.2}$$

可见从互连矩阵 $[\{\pm m_0, 0\}]$ 中提出 m_0 后，模拟输出量扩大了 $1/m_0$ 倍。在计算中可发现 m_0 的值一般在 $10^{-1} \sim 10^{-3}$ ，也就是说在三值权重编码中把模拟输出量扩大了 $10 \sim 1000$ 倍。因此，从理论上说经三值权重编码后网络的寻址能力会有所提高。

b) 模拟结果： 为了考察 TC 模型在有噪声情况下的寻址能力，在每个存储模式中用 N' 个随机数 ± 1 来代替该模式中的相同数目的分量，并用它作为与存储模式 Hamming 距离为 $d_H = N - N'$ 的初始模式，按 4.1 中的算法，求在不同的 d_H 下的 a_c^* 值，得到 $a_c^* \sim d_H$ 曲线。图 3 为 TC 模型、OC 模型、HP 模型(Hopfield 模型)的 $a_c^* \sim d_H$ 曲线。显然，图 3 曲线左边的区域系统总能对初始模式正确识别($\langle q \rangle < 0.015$)，而在曲线的右边系统不能正确识别。由图 3 可知，TC 模型比 OC 模型及 HP 模型有更大的寻址能力。

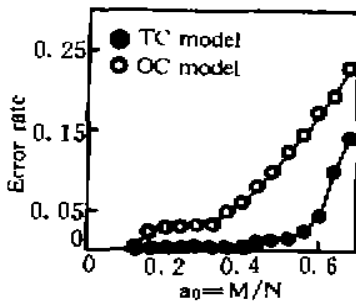


Fig. 2 The dependence of error rate $\langle q \rangle$ on a_c for OC and TC models

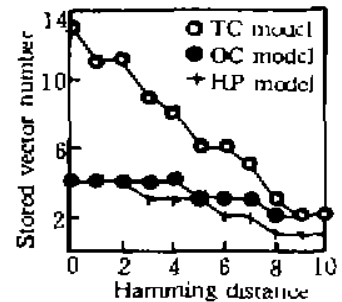


Fig. 3 Storage capacity of OC model, TC model and HP model vs. Hamming distance of addressing vector to the corresponding stored vector for $N=25$

4.3 迭代速度

为了研究 TC 模型的内容寻址收敛速度，采用 1-9 个阿拉伯数字作为存储模式[图 4]，并以接近这些存储模式的矢量作为输入模式，对 TC 模型、OC 模型、HP 模型的收敛速度进行统计分析。表 1 是三种模型收敛速度的统计平均值。由表 1 可知 TC 模型的收敛速度较快。

Table 1. Convergent speed of 3 models

model	TC	OC	HP
the average number of iteration	1.8	2.9	2.8

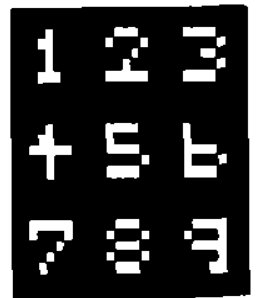


Fig. 4 Stored patterns (5 * 5 pixels)

结 论 计算机模拟结果表明，由本文提出的基于正交化算法的三值权重编码方法所建立的三级截值模型比目前的三值互连神经网络模型的存储容量提高了 3~4 倍，寻址能力及迭代速度都有明显的改善。因互连权重有只三值，极易于光学技术实现，是一种较理想的光学神经网络模型。另外，本文所提出的对互连权重的截值方法也完全适用于非正交化的 Hopfield 模型。

参 考 文 献

- [1] D. Psaltis, N. Farhat, Optical implementation of the Hopfield model. *Appl. Opt.*, 1985, 24(2): 1469~1475
- [2] T. L. u, X. Xu, S. Wu *et al.*, Neural network model using interpattern association. *Appl. Opt.*, 1990, 29(2): 284~289
- [3] 黄五群, 高成群, 陈天伦等, 一种适于光学实现的非线性神经网络模型及其蒙特卡洛学习算法. *光学学报*, 1990, 10(5): 446~451
- [4] 路明哲等, 利用神经网络优化关联存储模型. *红外与毫米波学报*, 1993, 12(1): 42~46
- [5] 张延忻, 常胜江, 光学神经网络中的信息编码及其优化. *激光与光电子学进展*(增刊, 中国光学会'95论文集), 1995, 7: 180
- [6] T. L. u, X. Xu, S. Wu *et al.*, Two-dimensional programmable optical neural network. *Appl. Opt.*, 1989, 28(22): 284~289
- [7] 黄五群等, A study on orthogonalized Hopfield model of neural network. *CCAST-WL Workshop Series*; 1994, 30: 46~57
- [8] Lu MZ, ZhanYL, Mu GG, Bipolar optical neural network with adaptive threshold. *Optik*, 1992, 81(4): 178~182

A Neural Networks Model with Ternary Interconnection Weights Based on Orthogonalized Algorithm

Chang Shengjiang Yang Jianweng Gao Shengquan
Shen Jinyuan Zhang Yanxin

(*Institute of Modern Optics, Nankai University, Tianjin 300071*)

(Received 24 April 1995; revised 27 November 1995)

Abstract An encoding method for constructing ternary interconnection weights based on the rule of least-lost of information is proposed in this paper. The encoding method improves the limitation on the accuracy of optical interconnection and is potential for optical implementation. The performance of ONN is evidently improved by this encoding method compared to the other methods proposed before. A detailed description of this method is presented.

Key words optical neural network (ONN), ternary encoding.

端面附近缺陷在掺铒光纤中引起的振荡

庞 勇 叶 勇 蒋佩璇 徐大雄

(北京邮电大学无线电工程系, 北京 100088)

摘 要 实验研究了由于掺铒光纤(EDF)端面附近存在缺陷引起光反射而导致的振荡, 这种振荡发生在掺铒光纤放大的自发辐射(ASE)谱的峰值波长处, 由于较长掺铒光纤的放大的自发辐射谱近似于其小信号增益谱, 因此也可以认为发生在掺铒光纤放大器(EDFA)的最大增益处, 对其增益产生不良影响。

关键词 掺铒光纤, 缺陷, 振荡。

1 引 言

对掺铒光纤激光器(EDFL)而言, 需要利用诸如法布里-珀罗腔^[1]、环形腔^[2]及光纤光栅^[3]等方法为其提供所需的光反馈, 获得稳定的激光, 而掺铒光纤放大器则不然, 为了防止系统产生的振荡影响放大器的性能, 就要设法抑制光反射^[4]。

一般情况下, 掺铒光纤中的光反射由瑞利散射以及熔接点反射引起, 这些因素可以通过技术措施得到改善。当光纤中存在缺陷时, 也能够产生相当大的反射, 从而可能导致自发的振荡, 对掺铒光纤放大器的性能产生不良影响, 这些缺陷可能是在掺铒光纤制作过程中造成的, 也可能是在使用过程中对掺铒光纤的处理不当造成的。

本文报道掺铒光纤端面附近存在缺陷时掺铒光纤放大器的实验研究结果, 该缺陷是由于对掺铒光纤放大器的出射端光纤端面处理不当引起的, 因为系统产生了振荡, 放大器的增益被降低了。

2 实验系统

图 1 是实验系统的示意图。该实验系统可以用来测量掺铒光纤放大器的增益特性, 也可以观察掺铒光纤的放大的自发辐射谱特性, 如果在方向耦合器与光谱分析仪的连接端插入适当的光滤波器, 还可以对掺铒光纤放大器的噪声特性进行粗略的测量。

在本实验中, 所用信号源是输出波长为 1537.25 nm 的分布反馈半导体激光器, 泵浦源是 1480 nm 波段大功率半导体激光器组件, 其输出功率与驱动电流的关系如图 2 所示(以上两者均由中国科学院半导体研究所海特光电技术公司提供)。使用的 1:1 方向耦合器是上海电缆传输研究所研制的, 其附加损耗小于 0.15 dB, 上海大学(嘉定校区)光纤通讯研究所

提供了 1480/1550 波分复用器, 其插入损耗约为 1.5dB, 掺铒光纤长度约为 4 m, 图 1 中左侧三个接头损耗约为 0.01 dB, 右侧接头损耗约为 0.2 dB。

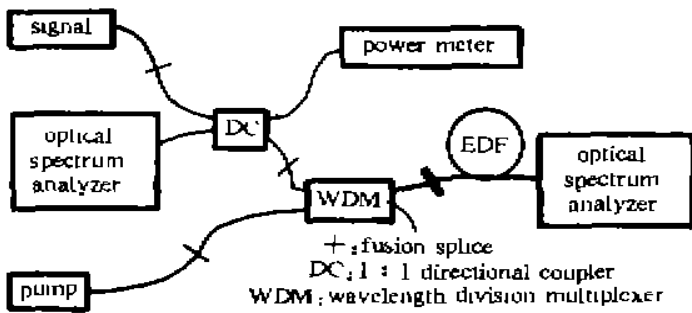


Fig. 1 Schematic view of experimental setup

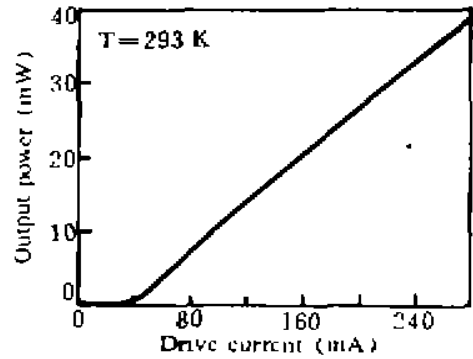


Fig. 2 Variation of the pump output power vs. drive current

3 实验结果及分析

利用上面的实验系统, 对掺铒光纤的放大的自发辐射谱以及增益进行了测量, 实验中, 所使用的光谱分析仪的分辨率为 0.1 nm。

为了简明和节省篇幅, 对光谱仪扫描出的放大的自发辐射功率谱的实验曲线进行了数据采样, 并且分别将正向和反向放大的自发辐射功率谱集中描绘在图 3 和图 4 中。数据采样的步长为 1 nm。

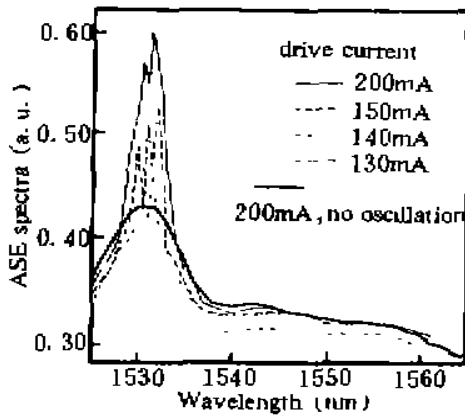


Fig. 3 Forward ASE spectra

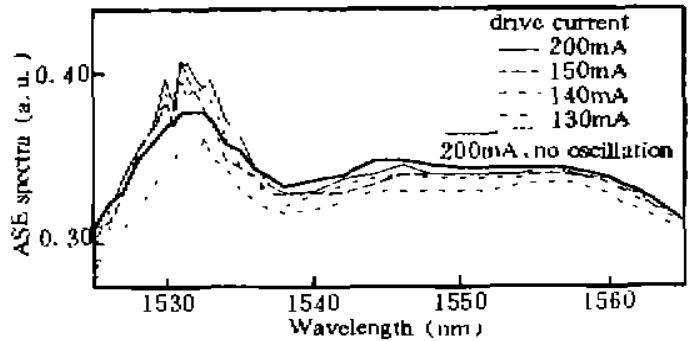


Fig. 4 Backward ASE spectra

图 3 为一系列的正向放大的自发辐射功率谱曲线, 各曲线对应的泵浦源驱动电流示于图例中, 同时给出了端面经过处理去掉缺陷后的放大的自发辐射功率谱曲线(泵浦源驱动电流为 200 mA, 见图例), 图中纵坐标为相对强度。当驱动电流为 130 mA 时, 系统尚未产生振荡, 当电流升高到 140 mA 时, 与没有端面缺陷时电流为 200 mA 的放大的自发辐射功率谱比较, 可以看出系统已在 1531 nm 附近出现振荡的迹象, 就象三能级激光系统达到了泵浦阈值, 随后可以看到电流的增加导致此处放大的自发辐射功率的迅速增强, 而其它波长处的放大的自发辐射功率则由于受到抑制而没有明显的增加, 显然系统已产生了振荡。在 1531 nm 附近出现振荡是因为此处的放大的自发辐射功率明显强于其它波长处, 其由于反射而引起的光反馈也最强, 又因为本实验系统没有波长选择手段, 所以此处具有形成振荡的优势。从分辨率为 0.1 nm 的光谱仪上可以观察到振荡包含了数个纵模(电流不同时数目不同), 由于数

据采样的步长为 1 nm, 又因为这些纵模的间隔很小, 所以图 3 描绘出的仅是一个包络, 而掩盖了振荡的精细结构。

对反向放大的自发辐射而言, 则未观察到显著的振荡(见图 4, 图中未补偿其通过波分复用器及定向耦合器而带来的约 4.5 dB 的损耗), 但是从电流为 130 mA 开始, 从光谱仪上就可以看到在 1531 nm 附近出现许多毛刺(由于上述同样的原因, 图 4 中看不出这些结构), 不过随着电流的增加, 振荡并不明显增强, 可见反向放大的自发辐射比正向更易产生振荡, 而其振荡强度却始终很弱。由于泵浦功率低时反向放大的自发辐射要强于正向放大的自发辐射, 所以反向放大的自发辐射比正向放大的自发辐射更早产生振荡, 但是从实验曲线上可以观察到反向放大的自发辐射具有较正向放大的自发辐射平坦得多的分布, 因而在 1531 nm 附近出现振荡的优势不是很明显, 所以反向放大的自发辐射中的振荡始终不是很强。从图 4 中还可以看到, 电流为 150 mA 后, 反向放大的自发辐射的增加不大, 这说明掺铒光纤中已出现饱和倾向。

反向放大的自发辐射功率谱的平坦特性尚不能很好解释, 理论上说^[5], 其与正向放大的自发辐射应有相似的形状, 但是实验结果也不容否认。由于与正向放大的自发辐射相比, 反向放大的自发辐射在测量之前经过了方向耦合器及波分复用器, 这两个器件的频率响应可能会带来一定的影响, 但不会很严重。

通过对掺铒光纤出射端端面进行重新处理, 观察到振荡现象消失了, 因此认为振荡是由于光纤出射端端面缺陷引起的, 而此缺陷可能是端面处理不慎产生的, 比如挤压产生的微裂纹等。

掺铒光纤中振荡的出现将会影响放大器的性能, 以增益特性为例, 通过比较有振荡和无振荡两种情况下系统获得的增益(见图 5, 信号功率约为 1 μ W, 同向泵浦结构), 可以看出振荡对放大器增益的抑制作用。从图 5 中也可以观察到增益趋势在电流为 150 mA 左右发生弯曲, 同样说明放大器出现了饱和现象。

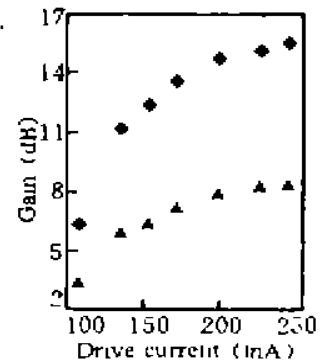


Fig. 5 Variation of gain vs. pump drive current

结 论 本文通过实验研究了由于端面附近缺陷在掺铒光纤中引起的振荡, 实验发现在正向放大的自发辐射中振荡较强, 而在反向放大的自发辐射中, 因为其功率谱较为平坦, 振荡比较弱。实验还比较了有振荡和无振荡两种情况下系统的增益, 说明了振荡对系统增益的抑制作用。

从实验结果也可以看出, 这种端面缺陷引起的反射要比光纤中的瑞利散射及熔接点反射严重得多, 能够对掺铒光纤放大器的性能造成更大的影响, 因此要特别注意防止这种人为造成的缺陷。

因为振荡对掺铒光纤放大器系统的性能有不良影响, 所以在掺铒光纤放大器的实验中要尽量消除导致振荡产生的各种因素。

参 考 文 献

- [1] Y. Kimura, M. Nakazawa, Lasing characteristics of Er^{3+} -doped fibers from 1553 up to 1603 nm. *J. Appl. Phys.*, 1988, 64(2): 516~520

- [2] G. J. Cowle, D. N. Payne, D. Reid, Single-frequency travelling-wave erbium-doped fibre loop laser, *Electron. Lett.*, 1991, 27(3): 229~230
- [3] J. L. Zyskind, V. Mizrahi, D. J. DiGiovanni *et al.*, Short single frequency erbium-doped fibre laser. *Electron. Lett.*, 1992, 28(15): 1385~1387
- [4] Y. Sato, Y. Yamabayashi, K. Aida, Noise figure monitoring of optical amplifiers via backward amplified spontaneous emission. *IEEE Photonics Technol. Lett.*, 1994, 8(2): 202~204
- [5] E. Desurvire, J. R. Simpson, Amplification of spontaneous emission in erbium-doped single-mode fibers. *J. Lightwave Technology*, 1989, 7(5): 835~845

Oscillation in Erbium-Doped Fiber Due to Near-End Defects

Pang Yong Ye Yong Jiang Peixuan Xu Daxiong

(Department of Radio Engineering, Beijing University of Posts & Telecommunications, Beijing 100088)

(Received 10 April 1995; revised 1 September 1995)

Abstract In this paper, we report the experimental results of oscillation in erbium-doped fiber (EDF), due to the considerable optical reflection caused by the near-end defects. The oscillation occurs at the peak ASE wavelength, which might be equal to the peak gain wavelength in long erbium-doped fibers, and is harmful to the gain characteristics of EDF amplifiers.

Key words erbium-doped fiber, defect, oscillation