

用于多个三维目标快速识别的 非相干光电混合神经网络系统*

孙颖 李清 张延焯 母国光

(南开大学现代光学研究所, 天津, 300071)

摘要 提出一种快速光电混合三层神经网络系统, 以实现三维目标的容错性模式识别. 系统的第一级网络采用正负互连并行处理, 并用白光光源代替体积较大的 He-Ne 激光器, 用 TFT 液晶显示屏做为实时输入器件, 既消除了相干噪声, 又实现了系统的小型化与快速识别化.

关键词 级联模型, 线性组合算法, 液晶显示屏(LCD), 非相干系统.

引言

三维目标的模式识别是光学信息处理的重要分支之一, 它涉及到从不同视角观察的投影图像的识别, 以及分离背景等问题. 目前已提出多种识别三维目标的算法, 如综合鉴别函数(SDF)^[1]、锁栓滤波器(LAT)^[2]、串码滤波器(SCFs)^[3]及级联模型等^[4], 使该问题的理论研究逐步完善和深入. 但由于这些算法极其复杂, 计算机不可能在短时间内快速完成识别. 故需有一套性能稳定的硬件实现系统. 我们曾提出一种光电混合系统实现三维目标识别^[5], 它虽具备光学系统的快速和并行性特点, 但由于采用相干光源, 用透明片输入图像, 且正负互连分两次进行, 从而给系统的实用化造成一定障碍.

本文提出一种实时非相干系统, 用小型白光光源代替体积庞大的 He-Ne 激光器, 并用 TFT 液晶显示屏(LCD)做为实时输入器件, 代替原系统的透明片输入, 且第一级网络的正负互连采用双通道并行处理方式, 大大缩短了识别时间. 由于白光光源体积很小, 故系统不但可进行实时快速识别、抑制了相干噪声, 还实现了小型化. 实验表明, 此非相干快速系统更适于实际应用.

1 基本原理及结构参数

1.1 结构及其参数

采用级联模型, 输入层神经元数为 H ; 中间层是目标的旋转不变性编码, 神经元数为

*国家自然科学基金资助项目

本文 1994 年 8 月 29 日收到, 最后修改稿 1994 年 11 月 28 日收到

N ; 输出层神经元数也为 N , 是正交化 Hopfield 模型的稳定输出状态. 待识别目标种类数为 M , 每一目标对应的特征码为 $C_m = [C_{m1}, C_{m2}, \dots, C_{mN}] (m=1, 2, \dots, M)$.

第一级网络将同一目标的不同投影图像映射为同一组特征码, 完成第一步的聚类 and 三维旋转不变性要求. 第二级网络以目标特征码 C_m 为存储模式, 通过选择正交化特征码, 及网络的反馈迭代, 达到容错性要求.

1.2 学习算法

综合串码滤波器和综合鉴别函数原理, 在第一级网络中采用线性组合学习算法.

设有 M 种三维目标, 每一种取 L 个不同的投影图像, 用 $H = I \times J$ 维二值矩阵来表示, 即 $\{f_{ij}(m, l)\} (m=1, 2, \dots, M; l=1, 2, \dots, L; i=1, 2, \dots, I; j=1, 2, \dots, J)$. 令互连子矩阵 $\|W_{ij}\|_k (k=1, 2, \dots, N)$ 表示输入层到中间层第 k 个神经元的互连权重, 取值为所有训练样本的线性组合, 即

$$W_{ijk} = \sum_{m=1}^M \sum_{l=1}^L a_{mlk} f_{ij}(m, l), \quad k=1, 2, \dots, N \quad (1)$$

按照对第一级网络的要求, 应有

$$\sum_{i=1}^I \sum_{j=1}^J f_{ij}(m, l) W_{ijk} = C_{mk}, \quad m=1, 2, \dots, M; k=1, 2, \dots, N \quad (2)$$

将式(1)代入式(2), 则有

$$\sum_{m=1}^M \sum_{l=1}^L a_{mlk} \left[\sum_{i=1}^I \sum_{j=1}^J f_{ij}(m, n) f_{ij}(m', l') \right] = C_{m'k}. \quad (3)$$

式(3)中 $k=1, 2, \dots, N; m'=1, 2, \dots, M$. 选定正交编码 $C_{m'k}$ 的值后, 由式(3)可解出展开系数 a_{mlk} , 从而确定互连权重矩阵 $\|W_{ij}\|_k$.

第二级网络采用正交化的 Hopfield 模型, 互连权重 T_{ij} 可直接由正交化编码 Hebb 规则求出. 当任输入一幅投影图像 $X = \{X_{ij}\}$ 时, 可在第一级网络中用光学方法实现 X 与互连权重 $\|W_{ij}\|$ 的内积, 从而得到一组码值 $\{C_i\} (j=1, 2, \dots, N)$. 它做为第二级网络的输入模式, 经过状态迭代, 收敛为与其最接近的特征码

$$C_i(t+1) = \text{sgn} \left[\sum_{j=1}^N T_{ij} C_j(t) \right]. \quad (4)$$

由此可知待识别目标的类别.

2 快速非相干光电混合系统

光电混合白光系统如图 1 所示. 系统采用白光光源做为照明光源, 在光源前加一小孔滤波, 故光源 S 近似为白光点源, 透镜 L_1 将点源 S 的成像在透镜 L_2 的中心, 这样既可保证输入面 P_1 上照明较均匀, 同时对成像透镜 L_2 的口径要求也较低. 将实时器件 TFT 液晶显示屏(LCD)放置在输入平面 P_1 处, 在微机 COM1 上按要求生成 4×4 个等大等间距的输入图像, 通过接口设备送至显示屏 LCD, 就可显示出 16 个等大的像做为系统的输入图像. 我们

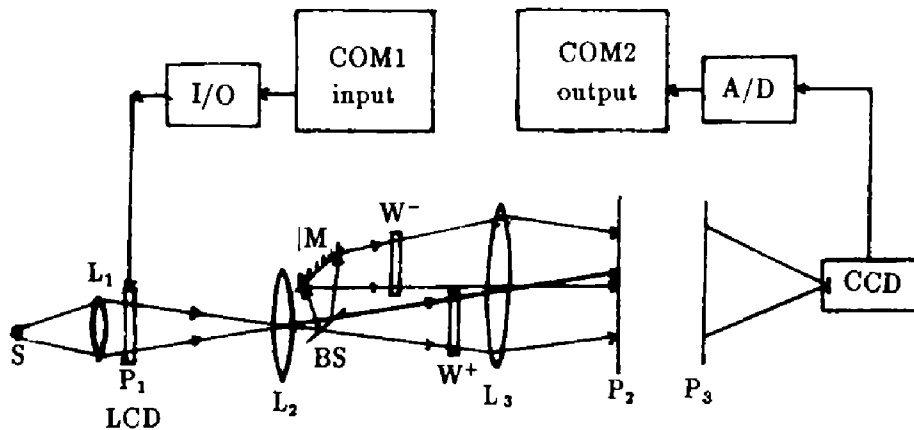


图 1 实时非相干神经网络系统

Fig. 1 Incoherent real-time neural network system

将双极互连权重分为正负两部分,即 W^+ 和 W^- . 在建立系统时保证 W^+ 和 W^- 都处在输入平面 P_1 的像面位置. BS 为分束器件,当一束载有 16 个图像的成像光束通过透镜 L_2 而入射到 BS 上以后,分束器 BS 可将其分成透射和反射两部分,经适当调节,可保证反射与透射图像强度相同. 于是载有图像的透射光照射在正权重片 W^+ 上;同时 BS 分出的反射光经反射镜 M 反射后,也照射在负权重片 W^- 上. 由以上可知,透射与反射光载有相同的图像信息,这样就实现了正负互连的并行处理. 我们通过调节各输入图像彼此的间距,并调节所成像的大小,使生成在权重片 W^+ 和 W^- 上的 32 个像分别与 32 个相应的互连权重子矩阵完全重叠,以便实现输入图像与互连权重的多通道乘积运算. L_3 为会聚透镜,它可将输入图像与正、负互连权重之间的乘积分布进一步会聚到 P_2 平面的小透镜阵列上,即透镜阵列上每一个小透镜对应一个物像与权重子矩阵的乘积分布. 于是在透镜阵列的焦平面 P_3 上便可得到 32 个会聚光点,光点处的光强值便是输入图像与每个权重子矩阵的内积值. 在 P_3 平面后适当位置处放一个 CCD 摄像机,采集光点后进行取阈处理,然后将正负互连通道所得的 32 个内积值在微机 COM2 中分别对应相减,得到输入图像与双级互连权重的 16 个内积值. 这就完成了第一级网络的映射过程,即将一幅二值投影图像映射为一组特征码 $\{C_j\}$ ($j=1,2,\dots,16$). 在微机 COM2 中, $\{C_j\}$ 做为第二级网络的输入模式,以第二级网络的互连权重进行状态迭代,最后收敛为 4 种模式编码中与其最接近的一种,由此可断定输入物属何种目标.

3 实验结果

3.1 样本的选择及互连模板的制作

我们选取了客机、战斗机、轰炸机和火箭 4 种目标,如图 2 所示,即 $M=4$. 选取方式为:在坐标系 (x_1, x_2, x_3) 中,令 $(\theta_1, \theta_2, \theta_3)$ 分别表示目标绕 x_1, x_2, x_3 轴的旋转角度, $\theta_1 = \theta_2 = \theta_3 = 0^\circ$ 表示目标处于零方位. 然后, (1) 从零方位开始,保持 $\theta_1 = \theta_2 = 0^\circ$, 每隔 10° 改变 θ_3 , 则每种目标得到 36 个投影样本; (2) 令 $\theta_2 = 30^\circ$, 对 $\theta_1 = 30^\circ, 45^\circ, 60^\circ$ 分别重复 (1) 的操作, 于是每种目标得到 108 个投影样本; (3) 令 $\theta_2 = 45^\circ$, 重复 (2) 的操作, 每种目标又得到 108 个样本.



图 2 4 种飞行器

Fig. 2 Four kinds of aircrafts

由以上选择可知,每种目标共选取 $L=36+108+108=252$ 个样本,故 4 种目标共有 1008 个样本模式. 每个样本像元数为 $H=I \times J=100 \times 100$,互连权重个数为 $N=16$. 4 组目标特征码彼此正交,如表 1 所示.

表 1 四种目标特征码
Table 1 The feature codes of four kinds of targets

目标	C_{m1}	C_{m2}	C_{m3}	C_{m4}	C_{m5}	C_{m6}	C_{m7}	C_{m8}	C_{m9}	C_{m10}	C_{m11}	C_{m12}	C_{m13}	C_{m14}	C_{m15}	C_{m16}
班机	-1	-1	+1	+1	+1	-1	-1	-1	-1	+1	+1	+1	+1	-1	+1	-1
战斗机	+1	+1	-1	-1	+1	-1	-1	+1	-1	-1	+1	-1	+1	+1	+1	-1
轰炸机	+1	-1	+1	-1	-1	+1	-1	-1	+1	-1	+1	+1	-1	+1	+1	-1
火箭	-1	+1	-1	+1	-1	+1	+1	-1	-1	-1	-1	+1	+1	+1	+1	-1

由 1008 个样本模式,根据前述的线性组合算法,设计出 16 个互连权重矩阵. 由于权重是通过线性方程组求解的,故灰阶数很高,达数千个. 我们用统计学方法将每个权重的灰阶压缩至 36 个,以便于硬件实现. 然后利用计算全息方法对其进行编码,并采用微电子学技术制备了正负两片互连模板. 将两块模板分别放在图 1 系统中的 W^+ 和 W^- 所在的平面处,做为第一级网络的互连权重.

3.2 计算机仿真及实验结果

对于压缩灰阶后的互连权重,我们首先在微机上对网络性能进行了检验,模拟结果为:(1)能正确识别所有样本模式;(2)能正确识别 90% 以上的非样本模式;(3)输入局部遮挡(或含噪声)的样本模式,当遮挡部分(或含噪声部分)小于 20% 时,网络亦能正确识别.

我们还对图 1 所示的白光系统进行了识别性能的检验. 对 4 种目标,我们共选取了 100 余幅投影样本进行检验. 其中 80% 以上是三维任意旋转的投影样本. 在图 1 的非相干光电混合系统中,照明光的均匀性、液晶显示屏的对比度、及 CCD 的热噪声等均给实验带来不利影响. 我们采取适当的补偿措施,使实验误差尽可能降低.

因每幅检测样本在系统中都有 16 个输出码,将其与 4 种目标的特征码进行比较,即可确定待测样本的类别. 我们定义码的正确率 = $\frac{\text{正确输出码的个数}}{\text{输出码总个数}}$,对所选取的 100 余幅检测样本得到实验结果如下.

在第一级光学网络中,由于各种噪声的影响,码的正确识别率为 79.4%,进入第二级网络后,由于该网络有较强的容错性,经过反复迭代,最后整个系统输出码的正确识别率可达到 93.8%,将检测样本输出码与目标特征码进行比较,得出检测样本的正确识别率达 80% 以上. 因检测样本中包含未经训练的非样本模式,故此非相干实时系统的性能是良好的. 且

该系统的识别时间达到 0.1s 数量级,实现了快速识别的目的.

4 结论

本文提出的快速非相干光电混合系统采用白光光源照明、正负互连权重双通道并行处理,并用实时输入器件 TFT 液晶显示屏来输入待识别图像.与过去的相干光学系统相比,白光实时系统不但发挥了光的快速及并行性特点,还有效地消除了相干噪声,实现了系统的小型化、实时化,因此更能适于实际应用.

参考文献

- 1 Casasent D. *Appl. Opt.*, 1984, **23**(10):1620~1627
- 2 Schils G F, Sweeney D W. *J. Opt. Soc. Am.*, 1988, **A5**(8):1309~1321
- 3 Yin Shizhuo, Cheng Lu, Mu Guoguang. *Optik*, 1989, **82**(4):129~131
- 4 Zhang Yanxin, Gao Chengqun, Huang Wuqun, et al. *Optical Computing and Processing*, 1992, **2**(1):3~8
- 5 孙颖,张延焯,杨翔鹏,等. *光子学报*, 1992, **21**(5):307~311

AN INCOHERENT OPTOELECTRONIC HYBRID NEURAL NETWORK FOR FAST RECOGNITION OF MULTIPLEX 3-D TARGETS*

Sun Ying Li Qing Zhang Yanxin Mu Guoguang

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

Abstract A fast incoherent optoelectronic hybrid neural network system is reported for pattern recognition of multiplex 3-D objects with error tolerance. In the first network, the positive and negative interconnections and parallel processing are used. A white-light source is used to replace He-Ne laser and a TFT liquid crystal display (LCD) panel is used as a real-time input device. In this system, coherent noise is eliminated, moreover, fast recognition and smaller size are realized.

Key words cascaded model, linear combination algorithm, liquid crystal display (LCD) panel, incoherent system

*The project supported by the National Natural Science Foundation of China