

# 视觉模式识别网络及光电实现\*

王 宁 刘立人

(上海光学精密机械研究所, 信息光学实验室, 上海 201800)

**摘 要** 在形态学网络及改进的 Hamming 网络基础上, 提出视觉模式识别网络及训练算法。形态学网络作为特征提取网络, 具有相对于图像平移和尺度不变的特性, 改进的 Hamming 网络加速完成特征矢量的分类操作。经过学习的视觉模式识别网络, 可完成图像特征的提取, 实现相对于图像平移和尺度不变的模式识别。在此基础上, 建立了光电视觉模式识别体系, 从而实现实时模式识别。

**关键词** 视觉模式识别网络, 训练算法, 光电实现结构。

## 1 引 言

Hubel 和 Wiesel<sup>[1]</sup>将视觉皮层细胞分为三组: 简单细胞, 复杂细胞和超复杂细胞。简单细胞仅在小范围的感受野中, 对图像的线和边界响应; 复杂细胞在较大范围的感受野中, 响应图像的曲线和边界; 超复杂细胞在大范围的感受野中, 感应图像的角度和界标。根据上述模型, Fukushima<sup>[2]</sup>建立了多层视觉识别网络, 该网络经过学习实现图像的特征识别, 但是网络结构和训练复杂。Wilson<sup>[3]</sup>提出以形态学网络实现相同比例字符的分类。本文建立的视觉模式识别网络和训练算法, 以形态学网络实现图像的特征提取, 改进的 Hamming<sup>[4]</sup>网络实现特征矢量的分类操作。经过训练的视觉网络, 能够实现相对于图像平移和尺度不变的模式识别。视觉网络训练包括作为探针的结构核优化和改进的 Hamming 网络互连权重的确定。在此基础上, 建立了光电视觉模式识别网络, 实现网络的实时操作。

## 2 视觉模式识别网络

图 1 给出视觉模式识别网络的结构图, 它是由单层的形态学网络和双层的 Hamming 网络组成。形态学网络是基于数学形态学 Rank order 变换, Rank order 变换  $\Xi$  的集合定义为:

$$A \Xi_t B = \{a: |A \cap B_a| \geq t\} \quad (1)$$

其中  $A$  表示离散的空间图像。 $B$  代表结构核。 $\cap$  代表逻辑“与”操作。 $|A|$  表示  $A$  的坐标集合,  $t$  是阈值。当  $t = |B|$  时, Rank order 变换退化为形态学的基本操作—腐蚀变换。当  $t = 1$  时, Rank order 变换成为形态学的另一基本操作—膨胀变换。根据 Rank order 变换的阈值卷积

\* 本课题得到国家高技术局和上海光机所所长基金的资助。

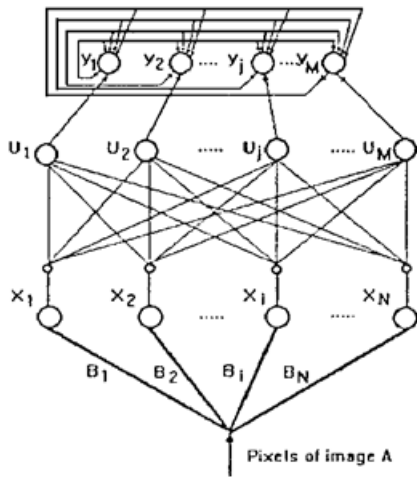


Fig. 1 The schematic diagram of the multilayer visual pattern recognition network

表示<sup>[4]</sup>，定义形态学网络如下：

$$X_i = \bigcup_{k,l} f_{t_i} \left[ \sum_{\zeta,\eta} A(k-\zeta, l-\eta) B_i(\zeta, \eta) \right]$$

$$f_{t_i}[x] = \begin{cases} 1 & x \geq t_i \\ 0 & \text{otherwise} \end{cases}$$

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

其中  $(k, l)$  和  $(\zeta, \eta)$  分别表示图像  $A$  和  $B$  的坐标， $f_{t_i}$  表示阈值函数， $\cup$  代表逻辑“或”操作。 $X_i$  表示形态学网络第  $i$  个神经元的输出， $B_i$  表示形态学网络第  $i$  个结构核。形态学网络利用结构核提取图像特征，网络的每一神经元对应一个或一组结构核。任何图像的角度代表图像的特征，该角度特征不随图像大小改变，因而以图像角度作为结构核提取图像特征，可得到随图像大小不变的特征矢量。又因为数学形态学变换具有随图像平移不变性，所以形态学网络是平移不变网络。形态学网络的

神经元是由图像角度构成的结构核组成，网络的互连表示图像中是否包含似同结构核的特征，如果含有似同结构核的特征，网络互连权重为 1，否则为 0。通过以上讨论，说明形态学网络能够实现相对输入图像平移和尺度变化的不变特征矢量提取。将结构核按照复杂程度排列，线状结构核在前，曲线和角状结构核排后，这样当形态学网络对一幅图像进行特征提取时，似同于视觉皮层细胞对图像特征的响应。

由于 Hamming 网络<sup>[5]</sup>少于 Hopfield 网络的互连且其训练算法简单，成为模式识别中的重要分类网络。通过测量输入与存储矢量的 Hamming 距离，确定图像与存储模式的最佳匹配。为了进一步增加非匹配模式与输入图像的 Hamming 距离，同时进一步减小最佳匹配模式与输入图像的 Hamming 距离，使 Hamming 网络收敛速度加快，特改进 Hamming 网络的定义如下：

$$U_j(0) = \begin{cases} \sum_{i=1}^N X_i^{(j)} X_i - D_j & \text{if } \sum_{i=1}^N X_i^{(j)} X_i > D_j \\ 0 & \text{otherwise} \end{cases}$$

$$D_j = \sum_{i=1}^N |X_i - X_i^{(j)}| \quad (3)$$

其中  $D_j$  表示输入与第  $j$  个存储矢量的距离。Hamming 网络的第二层是输出最大网络 (Maxnet)，定义如下：

$$U_k(s+1) = g[U_k(s) - \epsilon \sum_{j \neq k} U_j(s)]$$

$$(j \geq 1, k \leq M, \epsilon \leq \frac{1}{M}) \quad (4)$$

其中  $s$  表示循环次数， $\epsilon$  表示抑制参数。经过连续循环操作，网络将只有一个输出，它代表最佳匹配模式。

### 3 视觉模式识别网络的训练算法

图像的边界作为视觉模式识别网络的输入图像。视觉模式识别网络的训练算法分为二

- 步:
- 1) 根据训练模式选取和优化探针结构核;
  - 2) 由训练模式相对于优化探针结构核的形态学网络输出作为训练模式的特征矢量, 构成改进 Hamming 网络的互连权重矢量, 它表示 Hamming 网络记忆存储的模式样板, 从而完成该视觉模式识别网络的学习。

图 2 给出训练算法的流程图。训练过程如下:

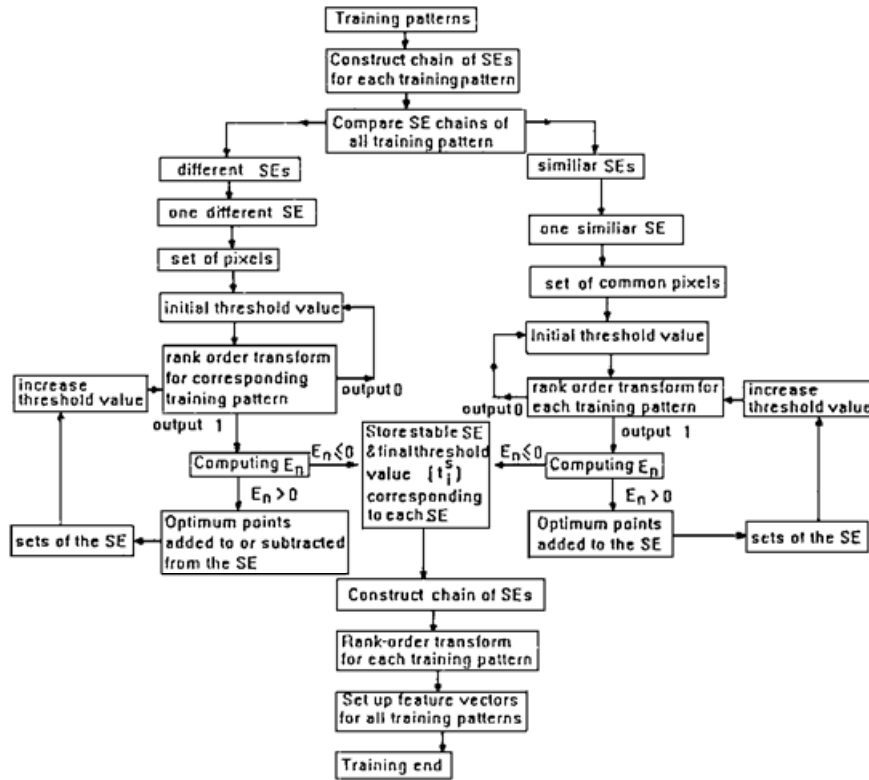


Fig. 2 The flow diagram of the visual network training algorithm

分解用于训练的模式, 以每一训练图像的边界角度构成网络初始结构核。将相似结构核和不相似结构核分开为二组。对于相似结构核, 取其相同坐标构成相似结构核的训练初始核, 把每一不相似结构核作为不同结构核的训练初始核, 采用图 2 流程图编程训练。训练中以  $E_n$  平方差判据判定网络训练程度,  $E_n$  的定义为:

$$E_n = \{ |B_{n+1}| - |B_n| \}^2 - \delta \quad (5)$$

$|B_n|$  表示结构核  $B$  经过  $n$  次优化操作的坐标集合,  $\delta$  是该平方差函数的阈值。  $E_n$  描述结构核优化过程中的误差。当  $E_n \leq 0$  时, 结构核的优化达到稳定状态, 记录此时的结构核以及 Rank order 阈值。经过优化的结构核按照从简单到复杂的序列排列组成结构核链。以此结构核链和相应的 Rank order 阈值对训练模

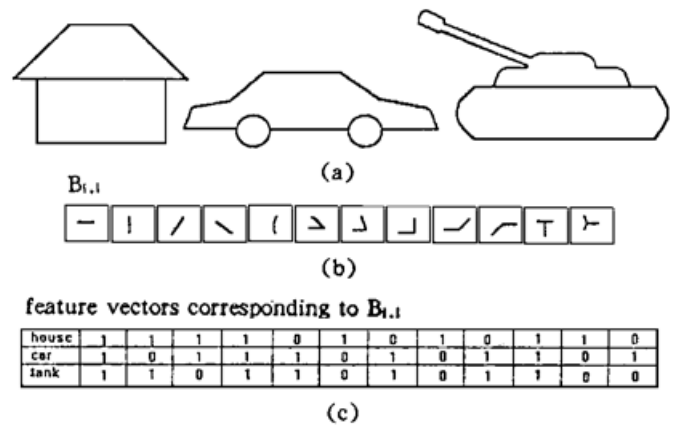


Fig. 3 (a) Training patterns and each has  $64 \times 64$  pixels, (b) chain of SEs (each SE is composed of  $16 \times 16$  pixels) after supervised training, (c) the feature vectors of the three training patterns corresponding to the optimized chain of SEs

式进行 Rank order 变换, 操作结果为形态学网络经过学习的输出, 组成改进的 Hamming 网络互连权重, 它表示存储模式的特征矢量, 从而完成视觉模式识别网络的训练。图 3 中: (a) 展示三个训练模式(occupies a window of  $64 \times 64$ ); (b) 给出优化的结构核链, (c) 给出三个训练模式相对于优化结构核链的特征矢量。

### 4 视觉模式识别网络的光电实现

光学以其并行输入和输出, 高密度的互连, 成为大规模互连网络和并行处理的研究热

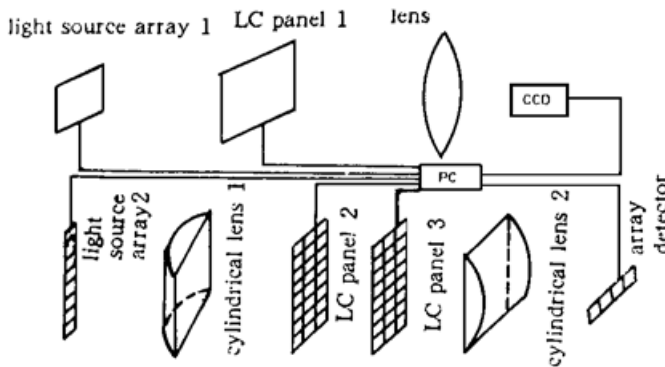


Fig. 4 An optoelectronic architecture of the visual network

点。利用光学高速并行特性, 发展光学模式识别体系, 是未来模式识别的任务之一。为此作者在建立了视觉模式识别网络的同时, 提出该网络的光电实现结构。图 4 是光电视觉模式识别网络的结构图。图的上半部是光学非相干相关器, 它用于实现形态学网络的邻域互连: LC panel 1 显示输入图像; 列阵光源表示经过训练的结构核; 透镜实现输入图像(LC panel 1)与探针结构核(列阵光源)的非相干相关。CCD 接收非相干相关图; PC 作为阈值和

存储器件, 完成形态学网络对于图像特征矢量的提取, 并将图像特征矢量进行面积编码。形态学网络输出为 1, 面积编码图对应白色, 否则网络输出为 0 对应面积编码图全黑。图的下半部是光学矢量矩阵乘法器(Crossbar), 用于实现改进 Hamming 网络的第一层互连操作: LC panel 2 显示 Hamming 网络的互连权重, 它是由训练模式的形态学网络输出经过面积编码组成(形态学网络输出是 1, 对应面积编码图为白色, 即 LC panel 2 透光, 反之该区域面积编码图为黑色, 即网络输出是 0, 对应 LC panel 2 不透光); LC panel 3 显示输入图像的特征矢量与训练存储图像的特征矢量面积编码图之间的逻辑与操作(以面积编码图光学逻辑与操作表示二矢量之间的距离, 即对应方程(3)中的  $D_j$ ) 的补。改进 Hamming 网络第一层光学实现的逻辑表达式为:

$$U_j(0) = \begin{cases} \sum_{i=1}^N \{X_i^{(j)} X_i \cap [X_i^{(j)} \cup X_i^c]\} & \text{if } \sum_{i=1}^N X_i^{(j)} X_i > D_j \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

其中  $X_i^c$  表示输入图像特征矢量  $X_i$  面积编码图的补,  $X_i^{(j)}$  表示改进 Hamming 网络的互连权重矢量的面积编码图, 代表训练模式特征矢量的面积编码图。图 5 中(a)展示一幅图像相对于探针结构核的非相干相关图; (b) 表示输入图像的形态学网络输出的面积编码图; (c) 和(d) 分别是 Hamming 网络互连权重的面积编码图(对应 LC panel 2 光强分布)和输入图像的特征矢量与训练存储图像的特征矢量面积编码图之间的逻辑与操作的补(对应 LC panel 3 光强分布); (e) 给出该视觉网络的输出。

**结 论** 本文提出视觉模式识别网络及其训练算法, 并在此基础上, 建立了该网络的光电体系。视觉模式识别网络是由一层形态学网络和二层改进 Hamming 网络组成, 形态学网络是特征提取网络, 改进 Hamming 网络实现特征矢量的分类操作。它能提高网络的收敛速度。形

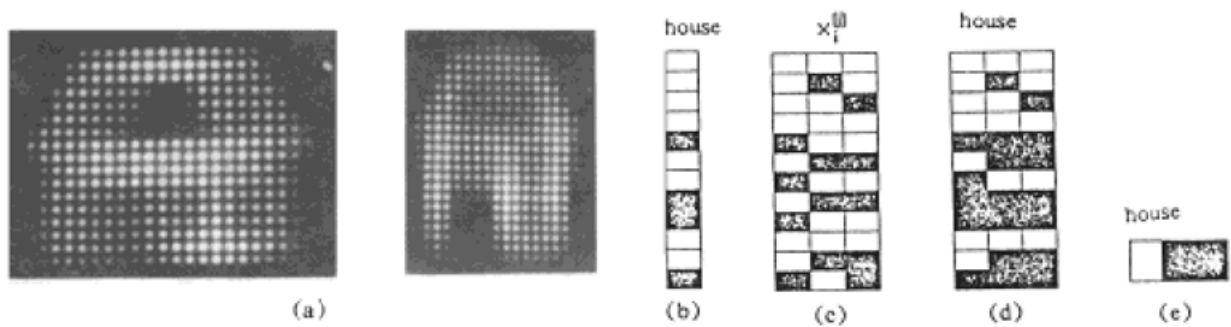


Fig. 5 (a) an image through neighborhood interconnection, (b) the first layer output of the visual network with the threshold value equal to  $4/5 t_i^0$ , (c) the pattern of LC panel 2, (d) the pattern of LC panel 3, (e) the output of the visual net

态学网络利用结构核提取图像的角度特征, 该特征不随图像大小改变, 因而可以得到随图像大小不变的特征矢量。由于数学形态学变换具有随图像平移不变的特性, 所以形态学网络是随图像平移不变网络。利用该模式识别网络能够实现相对于图像平移和尺度不变的特征提取, 完成随图像平移和尺度不变的模式识别。

### 参 考 文 献

- [1] D. H. Hubel, T. N. Wiesel, Receptive fields and functional architecture in two nonstriated visual areas (18 and 19) of cat. *J. Neurophysiology*, 1965, **28**: 229~ 289
- [2] K. Fukushima, Neural networks for visual pattern recognition. *Inst. Electron. Inf. Commun. Eng. Trans.*, (JPN) 1991, **74**: 179~ 190
- [3] S. S. Wilson, Training of object classes using mathematical morphology. *Proc. SPIE*, Nonlinear Image Processing, 1992, **1658**: 267~ 275
- [4] R. P. Lippmann, An introduction to computing with neural nets. *IEEE*, ASSP Magazine, 1987, **4**: 4 ~ 21

## A Visual Pattern Recognition Network and Optoelectronic Implementation

Wang Ning      Liu Liren

(Shanghai Institute of Optics and Fine Mechanics, The Chinese Academy of Sciences, Shanghai 201800)

(Received 30 October 1995; revised 4 March 1996)

**Abstract** A visual pattern recognition network and its training algorithm are proposed. The network is constructed of a one-layer morphology network and a two-layer modified Hamming net, which can implement invariant pattern recognition corresponding to image translation and size projection. After supervised learning, the visual network can extract image features and classify patterns. More over an optoelectronic architecture for real-time pattern recognition is proposed.

**Keywords** visual pattern recognition network, training algorithm, optoelectronic architecture.