

运用四元数神经网络模型 识别 16 色彩色字符*

陈振湘 帅建伟 刘瑞堂 吴伯僖

(厦门大学 物理系 361005)

used in PR
赵仰光

摘 要

本文把四元数离散神经网络模型应用于十六色彩色中英文字符的识别中, 讨论了彩色文字识别中的一些特点, 如由于颜色的改变, 导致信噪比发生变化, 所以网络的存储容量和纠错能力也随之而改变.

关键词 四元数, 神经网络, 字符识别.

1 引 言

随着以计算机和通讯为典型代表的微电子技术的迅速深入发展, 人类正在稳步迈入高度信息化的社会. 人们预言: 信息产业将是 21 世纪最重要最盈利的市场. 彩色数值图像的使用日益广泛, 计算机图像处理已成为信息学领域的一门重要分支学科. 对于人眼来说, 它所大量识别的图像是彩色图像, 建立自动识别彩色图像的计算机理论模型, 无论从应用还是从理论上来看, 都具有很重要的意义. 人工神经网络^[1]由于模拟了人脑的神经元的基本功能和网络的基本结构特征, 所以具有了人脑的分布式记忆、并行处理、联想记忆、分类与误差自动校正等智能功能. 常见的神经网络模型, 其神经元具有两态, 适用于黑白图像的识别. 建立多态神经网络模型以用于多值灰度或彩色图像的识别, 已提出了一些模型^[2-6]. 文[2]把四元数^[7]引入神经网络中, 建立了 16 态四元数神经网络模型. 该模型由于把四元数引入到神经元中, 所以可以用一个神经元表示一个具有 16 态取值的像点, 应用于 16 色数值彩色图像的识别中, 本文通过几个中英文彩色字符识别的计算机数值识别模拟实例, 较详细地讨论了其彩色文字识别中的一些特点.

2 四元数神经网络模型

四元数, 也称为 Hamilton 代数, 实质上是从实数轴 R , 复平面 C 扩张到四维空间的一个高维数, 它具有一个实部和三个虚部. 令四元数系

$$Q(R) = \{a : a = a + bi + cj + dk, a, b, c, d \in R\}$$

其中 i, j, k 表示虚部的三个基矢元, 用自然的方式定义它的元素的加法, 以及元素与实数的乘法, 元素间的乘法规定为用分配法去展开下式:

$$(a_0 + a_1i + a_2j + a_3k)(b_0 + b_1i + b_2j + b_3k) \quad (1)$$

* 本课题为国家自然科学基金重点项目资助课题; 本文于 1994 年 11 月 22 日收到

而其中各基矢元的乘法关系为:

$$\begin{aligned} i^2 = j^2 = k^2 = -1, & \quad ij = -ji = k, \\ jk = -kj = i, & \quad ki = -ik = j \end{aligned} \quad (2)$$

由上可知, 四元数虽不满足乘法交换律, 但满足结合律, 即有 $(\alpha \beta)\gamma = \alpha(\beta \gamma)$.

对任一四元数 $a = a + bi + cj + dk$, 定义其四元数共轭 $a^* = a - bi - cj - dk$, 故有 $aa^* = a^2 + b^2 + c^2 + d^2 = |a|^2$, 其中 $|a|$ 表示为 a 的模.

设四元数神经网络模型有 N 个神经元, 每个神经元的状态为如下 16 种状态中的任一种: $(\pm 1 \pm i \pm j \pm k)$. 存储 M 个记忆模式 $T^\mu = \{\pm 1 \pm i \pm j \pm k\}^N (\mu = 1, 2, \dots, M)$, 存储的样本由推广的 Hebb 学习规则构成连接矩阵, 即:

$$J_{mn} = \sum_{\mu=1}^M T_m^\mu (T_n^\mu)^* \quad (3)$$

网络的四元数动力学方程为

$$T_m(t+1) = \Theta \left\{ \sum_{n=1}^M J_{mn} T_n(t) \right\} \quad (4)$$

四元数阈值函数 $\Theta\{x\}$ 的取值规则如下: 当 x 的实部以及某一虚部不小于 0 时, $\Theta\{x\}$ 的对应部分取为正单位元, 当 x 的实部以及某一虚部小于 0 时, $\Theta\{x\}$ 的对应部分取为负单位元.

文 [2] 利用信噪比理论得当存储 M 个随机样本, 且 $M \ll N$ 时, 存储样本为四元数神经网络的稳定吸引子, 且其存储容量为 $1/2 \ln N$, 与文献 [8] 中所得的 Hopfield 模型的存储容量相同. 从几何上来看, 二值 Hopfield 模型把一维输入空间分为正负两半映射到输出空间的两个点上; 四元数离散神经网络模型表示在四维空间中, 在每一轴上取双极离散值: 正负单位元, 这样把输入空间扩展为四维. 其映射实质是: 以坐标轴及各 0 值平面为界, 把四维输入空间均分为 2^4 个象限, 各象限对应于输出空间的一个点. 这样网络对噪声的敏感性与二值 Hopfield 模型相当. 由于该网络的信噪比与二值 Hopfield 模型一样, 因此该模型的存储容量与 Hopfield 模型基本相同.

3 彩色字符的识别

由于计算机使用的是 0,1 数据文件方式, 其彩色图像是由三基色: 红 R, 兰 B, 绿 G 的数值灰度图像相加混合组成的数值图像来近似表示. 计算机中已得到了大量应用的是由三基色构成十六种色彩来近似表示一幅彩色图像, 一个像点由 4bit 表示, 其中有三个 bit 分别表示三基色, 余下的一个 bit 表示 2 级色饱和度, 这样构成了 16 色图形显示方式. 考虑到四元数具有一个实部和三个虚部, 在四元数神经网络模型中, 每个神经元为四位二值形式, 与计算机的 16 色表示方式很相似, 可以把它和 16 色彩色图像显示方式对应起来: 实数部分可对应色饱和度的两个级别, 三虚部对应三基色: 红, 兰, 绿, 则它可以存储十六色数值彩色图像, 应用于 16 色彩色数值图像的识别中. 在表 1 中, 列出了 16 种颜色的计算机编码和神经网络编码:

对于离散四元数神经网络模型, 为表征输入图像 $T = \{s_m = a_{m0} + a_{m1}i + a_{m2}j + a_{m3}k, m = 1, \dots, N\}$ 与存储样本 $T^n = \{s_m^v = b_{m0} + b_{m1}i + b_{m2}j + b_{m3}k, m = 1, \dots, N\}$ 的差别大小, 须先定义一个距离函数 D :

$$D = \sum_{m=1}^N \sum_{l=0}^3 |a_{ml} - b_{ml}| \quad (5)$$

距离函数 D 表示两图像的 bit 差别, 所以称 D 为广义汉明距离.

表 1 由三基色构成的十六色的神经元编码和计算机编码

颜色	神经元	计算机	颜色	神经元	计算机
0 黑	$(-1-i-j-k)$	0000	8 深灰	$(+1-i-j-k)$	1000
1 兰	$(-1-i-j+k)$	0001	9 淡兰	$(+1-i-j+k)$	1001
2 绿	$(-1-i+j-k)$	0010	10 淡绿	$(+1-i+j-k)$	1010
3 青	$(-1-i+j+k)$	0011	11 淡青	$(+1-i+j+k)$	1011
4 红	$(-1+i-j-k)$	0100	12 淡红	$(+1+i-j-k)$	1100
5 紫	$(-1+i-j+k)$	0101	13 淡洋红	$(+1+i-j+k)$	1101
6 棕	$(-1+i+j-k)$	0110	14 黄	$(+1+i+j-k)$	1110
7 淡灰	$(-1+i+j+k)$	0111	15 白	$(+1+i+j+k)$	1111

我们对四元数离散神经网络模型的 16 色彩色图像识别进行了初步的计算机模拟, 随机产生 10,000 个与任一记忆样本有某一固定汉明距离的变形字符, 包括色彩改变或出现多余色彩斑点等, 统计其纠错能力. 识别时设其背景颜色为黑色. 如对 7×10 点阵中字母的识别, 网络中存贮不同的三个汉字: “口” “中” “王”, 如图 1 所示, 其颜色分别为兰, 红, 白. 模拟表明, 这三个具有不同颜色的不同汉字均为网络的稳定存贮图像. 而若输入一些 bit 差别达 15, 即汉明距离等于 15 的稍有变形的汉字, 该模型正确识别率为 100%. 若输入一些 bit 差别达 20 的稍有变形的汉字, 正确识别率为 93%. 图 2 列出了三个差别达 20 的稍有变形的汉字, 网络均能正确识别. 若在网络中存贮四个不同颜色的同一汉字“王”, 其颜色为淡红, 淡兰, 淡绿, 淡灰. 模拟表明, 这四个具有不同颜色的“王”字均为该网络的稳定存贮图像. 当输入一些稍有变形的汉字“王”, 且其 bit 差别为 10 时, 模型正确识别率为 88%.

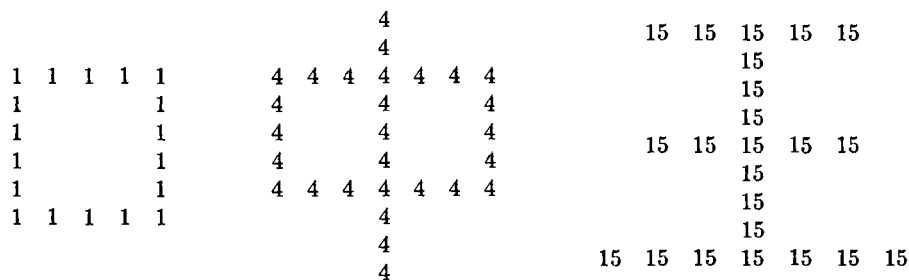


图 1 网络存贮不同的三个 7×10 点阵的汉字: 兰色的“口”, 红色的“中”和白色的“王”, 背景为黑色. 各数字代表的颜色见表 1.

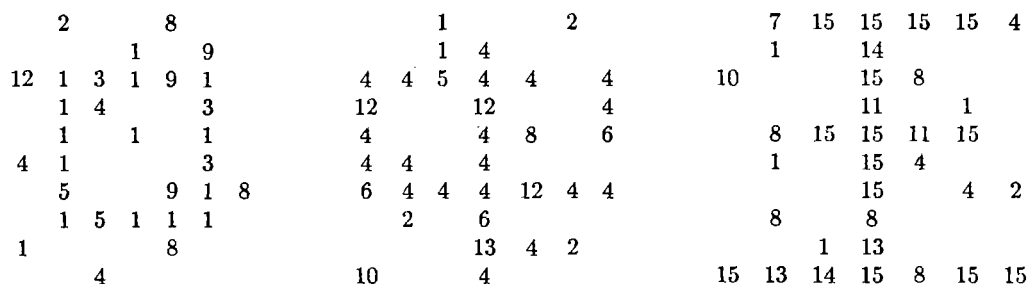


图 2 随机产生的具有汉明距离达 20 的变化彩色汉字

汉明距离达 30 的变形字母 A, 模型正确识别率降为 87%; 对汉明距离达 50 的变形字母, 其正确识别率却又为 76%, 比 58% 较高。

不难知, 由于颜色的改变, 也即其存贮记忆态的编码改变, 导致记忆态的伪正交性不同, 也即实际具体的信噪比不同, 从而决定了网络的存贮容量和纠错能力随不同的存贮字符而改变。

应该看到, 在四元数神经网络的彩色图像识别应用中, 对四元数的利用仅是由于我们取的四元数神经元的各个状态 $(\pm 1 \pm i \pm j \pm k)$ 与计算机 16 色数值图像的 4 位 0,1 的二进制编码表示每一像点有着数学上的一一对映关系, 但四元数间的运算关系与像素间的变换关系并不存在物理上的对应特征。所以该模型实际上也可用于对十六值灰度图像的识别中, 而本文中的上述各例也可以视为十六值灰度图像, 所得的结论对灰度图像的识别也同样适用。

4 结 论

本文运用四元数 (Hamilton 量) 离散神经网络模型对彩色中文英文字符进行了识别模拟, 讨论了彩色文字识别中的一些特点以及该模型的一些性质。四元数离散神经网络模型对彩色图像的识别, 实质上是用一个高维数来表示一个多值像点的状态。因此, 该模型不仅适用于 16 色数值图像的识别应用中, 也可用于 16 值灰度数值图像的识别应用中。

参 考 文 献

- [1] Hopfield J J. Neural Networks and Physical Systems with Emergent Collective Computational Abilities. Proc. Natl Acad Sci U.S.A, 1982, 79: 2554—2558.
- [2] Shuai J W, Chen Z X, Liu R T, Wu B X. The Hamilton Neural Network Model. Physica A, 1995, 216: 20—31.
- [3] 帅建伟, 陈振湘, 刘瑞堂, 吴伯倬. 八元数神经网络模型: 256 色图像的识别. 光学学报, 1995, 16: 556—560.
- [4] Rieger H. Storing an Extensive Number of Grey-toned Patterns in a Neural Network Using Multistate Neurons. J Phys A, 1990, 23(23): 1273—1279.
- [5] Kanter I. Potts-glass Models of Neural Networks. Phys Rev A, 1988, 37: 2739—2742.
- [6] Noest A J. Discrete-statephasor Neural Networks. Phys Rev A, 1988, 38: 2196—2199.
- [7] Condon E U et al. Handbook of Physics. New York: Second Edition, 1967, 1—22.
- [8] Mclellan R J, Posner E C, Rodemich E R, Venkatesh S S. The Capacity of the Hopfield Associative Memory. IEEE Tran Info Theory, 1987, IT-33(4): 461—482.

USING THE HAMILTON NEURAL NETWORK MODEL TO RECOGNIZE THE 16-LEVEL COLOR CHARACTERS

Chen Zhenxiang Shuai Jianwei Liu Ruitang Wu Boxi

(Department of Physics, Xiamen University, 361005)

ABSTRACT

We apply the Hamilton discrete neural network to recognize 16-level color characters, and discuss some features of the model. If we change the color of the characters, the signal-noise-ratio is changed, so the storage capacity and the error tolerance capacity are also changed.

Key Words Hamilton Number, Neural Network, Character Recognition.