

# 用振子神经网络实现二值图象中的多目标分割\*

王磊<sup>1</sup> 莫玉龙<sup>2</sup> 戚飞虎<sup>1</sup>

(1. 上海交通大学计算机科学与工程系, 上海 200030)(2. 上海大学通信与信息工程学院, 上海 201800)

**摘要** 利用基于视觉模型建立起来的局部激励、全局抑制振子神经网络(OSNN)实现二值图象中数字及字母的目标/背景分割;从理论上阐述了该网络实现目标/背景分割的机理,并分析了如何合理选择参数以利于成功分割。

**关键词** 目标分割,振子神经网络。

目标分割的机理并详细分析了如何合理选择参数。

## 1 引言

## 2 网络模型

众所周知,人类视觉感知的一个基本能力是能够对视觉感知场的元素很轻易地进行相关聚类,即将目标从背景中分离出来并识别它们。但是,如何用机器模拟人的视觉过程来实现这些功能尚未得到彻底的解决。为此,有关这个课题的研究一直在进行之中<sup>[1~6]</sup>。对于二值图象的多目标分割,一种基于视觉模型中相关神经激活理论设计的振子神经网络引人关注<sup>[5,6]</sup>。Ch. von der Maisburg 等人从脑函数和感知结构出发推导出可以用瞬态相关理论代表视觉模型<sup>[3,7]</sup>,该理论表明根据各个分散细胞激活性能的瞬态相关性可以记录各个目标的不同特性,从而分出各目标<sup>[3]</sup>。该理论已从猫的视觉皮层细胞的记录数据中得到了直接的实验验证<sup>[8,9]</sup>。若用振子神经元代表目标的某一特征(比如象素)就可以实现这种瞬态相关:一个目标由一组振子同步振动而显现,不同目标由不同组振子在不同时刻共振而区分。根据这一思想, D. L. Wang 等人提出一种四邻域局部激励、全局抑制神经网络<sup>[5,6]</sup>分割二值图象中的目标(字母)和背景,其分割结果仅模拟定出了目标的位置,而没有定出目标的准确形状。在此基础上,本文将局部激励扩展为八邻域,以实现有斜线的数字和字母混合二值图的分割,并实现了目标准确形状的定位。此外,本文介绍了多

在图1所示的网络模型中,每一个神经元振子(用空心圆表示)对应图象的一个象素,它只与其近邻的神经元有联系。全局抑制子(用黑方块表示)则与所有神经元都连接。每个振子的结构如图2所示,激励单元  $X_i$  和抑制单元  $Y_i$  之间反馈循环,它们构成动力学方程组

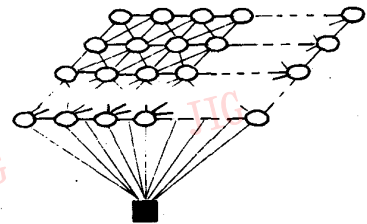


图1 网络模型

Fig. 1 Model of the network



图2 单个神经元振子模型

Fig. 2 Model of the single neuron oscillator

\* 本课题受国家自然科学基金资助

收稿日期:1996-10-15;收到修改稿日期:1997-04-24

$$dX_i/dt = 3X_i - X_i^3 + 2 - Y_i + \rho + I_i + S_i \quad (1)$$

$$dY_i/dt = \epsilon(\gamma(1 + \tanh(X_i/\beta)) - Y_i) \quad (2)$$

全局抑制子

$$dZ/dt = \phi(\sigma_\infty - Z) \quad (3)$$

式(1)中的  $\rho$  代表加在第  $i$  个振子上的高斯噪声, 噪声的加入有两个原因: 一是为了去同步, 二是为了增强系统的鲁棒性;  $I_i$  代表该振子受到的外部激励;  $S_i$  表示近邻振子及全局抑制子对该振子的影响

$$S_i = \sum_{k \in N(i)} W_{ik} S_\infty(X_k, \theta_x) - W_z S_\infty(Z, \theta_{zx}) \quad (4)$$

$$S_\infty(X, \theta) = 1 / \{1 + \exp[-K(X - \theta)]\} \quad (5)$$

$N(i)$  表示第  $i$  个振子的近邻, 除了在边界上它可能为 3 或 5, 其余都为 8。  $W_{ik}$  是第  $i$  个振子与邻域第  $k$  个振子之间的连接权值;  $W_z$  为全局抑制子与各个振子之间的连接权值。式(5)中的  $K$  是个大数, 因此当  $X > \theta$  时,  $S_\infty(X, \theta) \rightarrow 1$ ; 当  $X < \theta$  时,  $S_\infty(X, \theta) \rightarrow 0$ , 由此可见,  $\theta_x$  和  $\theta_{zx}$  都是门限。它们分别决定邻域点和全局抑制子是否影响该振子。式(2)中的  $\epsilon$  是一很小的数<sup>[6]</sup>; 参数  $\gamma$  控制单个振子停留在激活状态(如图 3 中的峰顶)和非激活状态(如图 3 中的平坦谷底)的相对时间。  $\gamma$  越大, 消耗在激活状态的时间越少。参数  $\beta$  用于控制  $dY_i/dt = 0$  这条 sigmoid 曲线(见文献[5]图 1)的陡峭程度。在式(3)中, 当每一个振子的  $X_i < \theta_{zx}$  时,  $\sigma_\infty = 0$ ; 当至少有一个振子的  $X_i \geq \theta_{zx}$  时,  $\sigma_\infty = 1$ 。故门限  $\theta_{zx}$  决定全局抑制子是否起作用: 在前一种情况下,  $Z \rightarrow 0$ , 所有的振子都不受到抑制; 在后一种情况下,  $Z \rightarrow 1$ , 抑制子起全局抑制作用。参数  $\phi$  决定全局抑制子响应  $Z \rightarrow 0$  或  $Z \rightarrow 1$  的速度。

### 3 目标分割机理及参数选择

为了使目标能从背景中分出来, 必须激活目标振子而抑制背景振子, 因此,  $I_i$  的取值方式为: 目标振子的  $I_i > 0$ , 比如取  $I_i = 0.2$ , 这些振子被激活后振动起来,  $X_i$  发生周期性的激活和抑制, 如图 3 虚线所示; 背景振子的  $I_i < 0$ , 可取一较大的负值如  $I_i = -3.0$  以使背景振子受到很强的抑制, 如图 3 实线所示。由于这些不活跃的背景振子对于分割目标没有影响, 因此在整个分割过程中不必理会, 这样还可节省处理时间。

在  $t = 0$  时刻, 所有神经元的状态都是随机的, 但是经过一段很小的时间变化后, 所有的目标振子都因为很强的邻域耦合而同时同步振动, 接着由于全局抑制子的抑制作用, 它们又几乎同时到达不活跃阶段。

当所有的  $X_i$  都小于  $\theta_{zx}$  和  $\theta_x$  时,  $S_i$  变为 0。这时, 目标振子要经过一段较长的时间才能再次到达激活状态(如图 3 点线所示)。如果各目标振子接收到的外部输入相同, 那么这些目标永无法去同步, 所以必须引入噪声  $\rho$ , 它使不同目标中振动状态相近的振子拉开距离, 一旦某个振子的  $X_i$  超过  $\theta_{zx}$ ,  $Z$  就立刻趋近 1, 这使其它不包含该振子的目标的  $S_i$  成为负值而被抑制, 而该激活点的近邻点的  $S_i$  则为正(若选择合适的  $W_{ik}$  和  $W_z$ ), 这样近邻点被激活了, 它们又去激活它们的近邻点, 如此传播开去, 直至整个目标的振子同步振动起来。一段时间后, 该目标被抑制, 它所有的  $X_i$  都比其余目标振子的  $X_i$  小, 因此, 再过一段时间后, 其余目标中的某一个又按与上述同样的方式单独同步振动起来。最后, 各个目标的振子周期性地分别在不同时刻同步振动。总之, 目标的同步依靠的是局部激励, 目标之间的去同步主要靠全局抑制子的作用, 另外还依靠噪声。因此, 合理选择连结权值  $W_{ik}$  和  $W_z$ , 门限  $\theta_{zx}$  和  $\theta_z$  很重要。

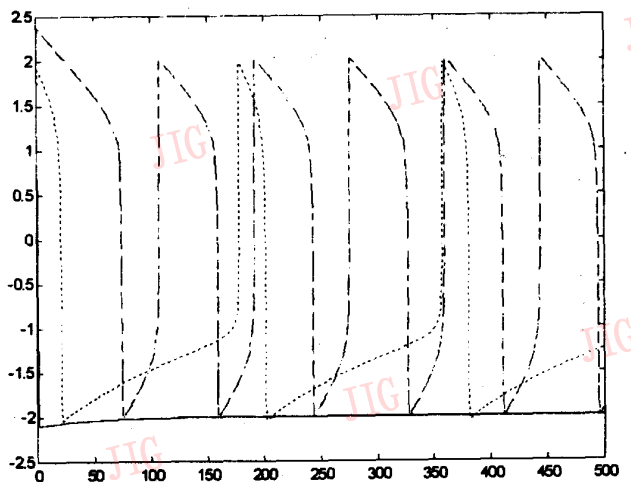


图 3 单个振子激励单元  $X_i$  在不同外部输入下随时间变化的曲线( $\rho = 0.05$ )。实线对应  $I = -3.0, S = -1$ ; 虚线对应  $I = 0.2, S = 5.0$ ; 点线对应  $I = 0.2, S = 0$ 。其他参数同第四部分实验数据。

Fig. 3 Effects of varying external input on single oscillator's activity part versus time ( $\rho = 0.05$ ). Solid line,  $I = -3.0, S = -1$ ; dashed line,  $I = 0.2, S = 5.0$ ; dotted line,  $I = 0.2, S = 0$ . Other parameters are the same as that in the experimental part.

连结权值  $W_{ik}$  是否取有效值(即  $W_{ik} \neq 0$ ) 的原则是: 当且仅当两个振子是近邻且都对对应目标点(即都受到外部激励)时才取值。这就保证了同一目标振子能同步而不同目标之间能去同步。其次, 为了满足同

步条件,目标振子的  $S_i$  不能相差很大,因此  $W_{ik}$ , 应根据第  $i$  个振子有多少个满足以上条件的邻域点动态取值,从而使  $\sum_{k \in N(i)} W_{ik}$  为一常数  $C$ 。比如,某振子若有 3 个满足条件的邻域点,则  $W_{ik} = C/3$ 。这样取值很重要,否则目标的同步性就会很差。比如,字母“6”左侧中间的横、竖线交叉点及其邻近 3 个点都有四邻域点,其余点大都只有二邻域点,若  $W_{ik}$  取固定值,则这几个点的  $S_i$  比其余点的  $S_i$  大将近一倍,使这 4 点组成的区域比其它部分激活时间长,严重影响了“6”的同步性。至于  $W_z$ , 它应比最小的  $W_{ik}$  略小一些。这是因为,若目标中具有最小  $W_{ik}$  (如字母“6”中右上端点)的振子最先被激活,一旦其  $X_i$  值大于门限  $\theta_{zx}$ , 全局抑制子立刻起作用,若  $W_z$  大于该  $W_{ik}$ ,  $S_i$  就出现负值,至使该点无法耦合其邻域点,进而使该目标无法被激活。此外,取值  $C$  和  $W_z$  还要考虑到  $S_i = C - W_z$  的大小。 $S_i$  太大,目标被激活后同步振动的时间太长,浪费时间; $S_i$  若太小则会影响同步:最先振动起来的振子还未等及同目标的其它振子都振动起来就开始回到了不活跃阶段(参看图 3 虚线和点线不同的脉冲宽度)。至于  $\theta_{zx}$  和  $\theta_x$ , 两值以相近为宜, $\theta_x$  略小一点,这有助于初始去同步。在本实验的取值前提下,若取  $\theta_x = 0.1$ , 只能分出“C”;若取  $\theta_x = 0.12$ , 只能分出“C”和“F”;当取  $\theta_x = 0.14$  时,实验就成功了。

在文献[5]中,定位目标的方法是先每个振子的  $X$  值归一:  $(X - X_{min}) / (X_{max} - X_{min})$ , 然后以归一化后的  $X_i$  为半径在每个象素点位置作圆。当一目标同步振动时,它各个振子的圆最大,由此显示提取的目标。但此方法未能重现目标。我们直接将归一化后的以强度图象(intensity image)方式显示,这样处于激活状态的目标的  $X_i$  最大,趋近 1, 而其它受抑振子的  $X_i$  都很小,因此目标就得以重现。这种方法既简单又省时。

### 4 模拟实验

实验用的是一  $14 \times 22$  包含 4 个目标的图象。有关数据为: 方差为 0.02 的高斯噪声, 常数  $C = 6.0$ ,  $W_z = 1.0$ ,  $\epsilon = 0.02$ ,  $\gamma = 6.0$ ,  $\beta = 0.1$ ,  $K = 50$ ,  $\theta_x = 0.14$ ,  $\theta_{zx} = \theta_{xz} = 0.15$ ,  $\phi = 3.0$ 。初始的  $X_i$  和  $Y_i$  在  $0 \sim 1$  范围随机取值。时间步长  $\Delta t = 0.3$ , 此值不宜过大, 以免  $X_i$  突变太快而影响初始去同步。实验结果如图 4 所示: 图 4(a) 为全体目标同步显现时的结果, 图 4(b)、图 4(c)、图 4(d) 和图 4(e) 分别为目标 C、6、2、F

在不同时刻单独显现出来的结果, 图 4(f) 上面 4 幅图表示各目标所有的  $X_i$  随时间变化的情况, 底下则是全局抑制子的变化情况。很明显, 经过一段时间的调整后, 4 个目标周期性单独显出, 分割成功。

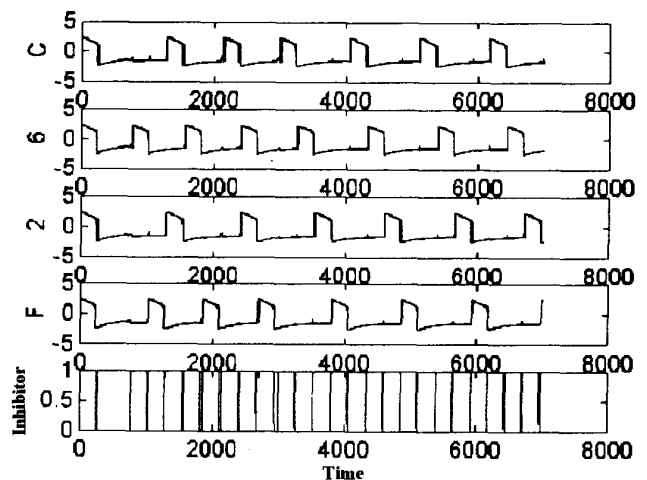
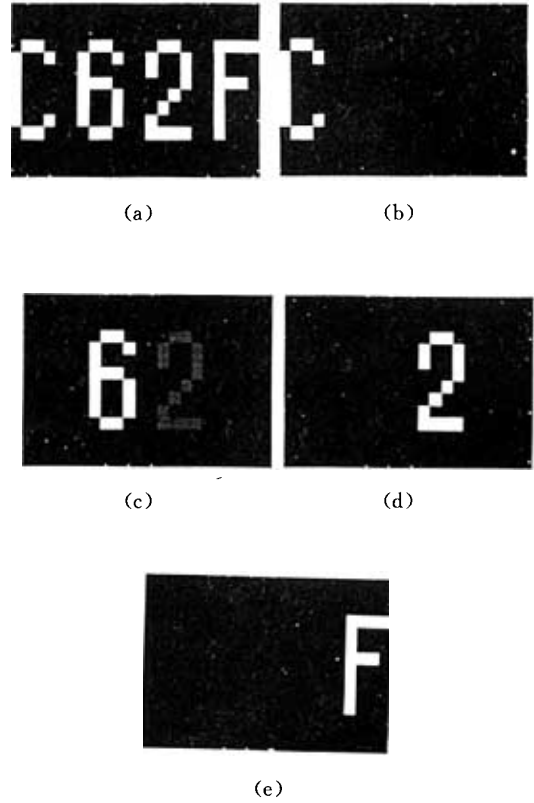


图 4 模拟实验结果。总共积分 7000 步

Fig. 4 Experimental results.

Took 7000 integration steps.

## 5 讨 论

本方法对于独立连通的目标都能实现提取,而与目标的位置分布无关。本模拟实验中,在每个积分步长都要用龙格—库塔法解  $N$ (总象素)个动力学方程组,因此虽然与某些传统方法如区域生长、投影方法等相比,运算时间相对较长,但是,因为该方法的理论依据是视觉函数和感知模型,所以具有特殊的理论研究意义。如果在专用硬件环境下,它可用于诸如车牌或车厢号码等的识别,具有一定的实用价值。此外,该方法如何用于灰度图象的分割,是一个值得研究的课题。

### 参 考 文 献

- 1 Herbert J. Reitboeck, Michael Stoecker, Christoph hahn. object separation in dynamic neural networks. IEEE Inter. Conf. on Neural networks, 1993, 2: 638~641.
- 2 Ch. von der Malsburg, J. Buhmann. Sensory segmentation with coupled neural oscillators. Biol. Cybern., 1992, 67: 233~242.
- 3 Ch. von der Malsburg. The correlation theory of brain functions. Internal Report 81-2, Max-Planck-Institut for Biophysical Chemistry, Goettingen, FRG (1981).
- 4 Paul Sajda, Finkel L H. A neural network model of object segmentation and feature binding in visual cortex, 1992, 4, 43~48.
- 5 Wang D L, Terman D. Locally excitatory globally inhibitory oscillators networks; theory and application to pattern segmentation. IEEE Inter. Conf. Neural Network, 1994, 2: 945~950.
- 6 Terman D, Wang D L. Global competition and local cooperation in a network of neural oscillators. Physica D 81, 1995, 148~176.
- 7 Ch. von der Malsburg, Schneider W. A neural cocktail-party processor. Biol. Cybern., 1986, 54: 29~40.
- 8 Eckhorn R. et al. Coherent oscillations; A mechanism of feature linking in the visual cortex. Biol. cybern. 1988, 60: 121~130.
- 9 Gray C M. et al. Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. Nature, 1989, 338: 334~337.



王 磊,毕业于苏州大学物理系并获理学学士学位,1994年于该系毕业并获光学专业硕士学位。现在上海交通大学电子信息学院攻读通信与电子专业博士学位,系上海交大与上海大学(嘉定分区)联合培养。目前的主要研究领域是神经网络、图象处理及模式识别等。

# Multiple Object Segmentation in Binary Images by Oscillator Neural Networks(OSNN)

Wang Lei<sup>1</sup>, Mo Yulong<sup>2</sup>, Qi Feihu<sup>3</sup>

(1. Dep. of Computer science and Technology, Shanghai Jiao Tong University, Shanghai 200030)

(2. Institute of Communication & Information Engineering, Shanghai University, Shanghai 201800)

**Abstract** This paper proposes a method of segmenting numbers and characters from the background in the binary images based on locally excitatory and globally inhibitory oscillator neural network (OSNN) which was built according to the visual model. It also discusses the mechanism of object/background segmentation made by the neural net and the problem of how to selecting parameters resonably in detail.

**Keywords** Object segmentation, Oscillator neural network