

# 脉冲神经网络图像分割的编码方法

崔文博, 蔺想红, 徐满意

(西北师范大学计算机科学与工程学院, 兰州 730070)

**摘 要:** 针对脉冲神经网络图像分割中的脉冲编码问题, 基于 Time-to-First-Spike 编码策略提出 2 种编码方式: 线性编码和非线性编码。线性编码方法采用从图像像素值到神经元脉冲发放时间的线性函数对应关系, 而非线性编码方法采用 Sigmoid 函数的对应关系。应用 2 种方法对图像进行分割, 实验结果表明, 非线性编码方法的分割结果优于线性编码方法, 分割图像具有更大的香农熵值, 并且非线性编码方法在图像分割时具有更大的取值区间, 更容易对参数进行选择, 取得最佳的图像分割结果。

**关键词:** Time-to-First-Spike 编码; Sigmoid 函数; 脉冲神经网络; 图像分割; 感受野; 最大香农熵

## Coding Method of Image Segmentation in Spiking Neural Network

CUI Wen-bo, LIN Xiang-hong, XU Man-yi

(College of Computer Science and Engineering, Northwest Normal University, Lanzhou 730070, China)

**【Abstract】** According to the difficulties of coding strategy of Spiking Neural Network(SNN) for image segmentation, two types of Time-to-First-Spike coding methods are proposed: linear coding and non-linear coding. Linear coding uses a linear function that realizes the corresponding relationship from pixel values to the spike times of the neurons, while non-linear coding uses corresponding relationship of Sigmoid function. Experimental results of image segmentation show that, the segmentation result using the non-linear coding is better than the result using the linear coding, and the segmentation image of non-linear coding has greater Shannon entropy. The method of non-linear coding is easier to select optimal parameters, and acquires the best segmentation result of image.

**【Key words】** Time-to-First-Spike coding; Sigmoid function; Spiking Neural Network(SNN); image segmentation; receptive field; maximum Shannon entropy

DOI: 10.3969/j.issn.1000-3428.2012.24.046

### 1 概述

近年来, 由脉冲神经元模型为基本单元构成的脉冲神经网络(Spiking Neural Network, SNN)越来越受到研究者的关注。应用脉冲神经网络对实际问题进行求解, 首要的问题是要理解信息是如何编码的, 以及被编码的信息又怎样用于一系列神经计算中, 即解码过程。在对神经元放电编码规律的认识过程中, 形成了各种不同的编码方法, 例如: 频率编码, Time-to-First-Spike 编码, 延时编码, 群体编码, 相位编码等<sup>[1-3]</sup>。传统的人工神经网络是对生物神经元的脉冲发放频率进行模拟, 而脉冲神经网络以神经元的脉冲发放时间信息作为网络的输入和输出, 因而更接近于真实的生物神经系统。与传统的应用脉冲频率编码的人工神经网络相比, 基于脉冲精确定时特性的脉冲神经网络拥有更强大的计算能力, 可模拟各种神经元信号和任意的连续函数, 非常适合于实现大脑神经信号的处理问题<sup>[4]</sup>。

图像分割是图像分析、图像理解和计算机视觉的基

础, 是图像处理中的难点。越来越多的研究者将脉冲神经网络应用于图像分割中, 取得了丰硕的成果<sup>[5-6]</sup>。应用脉冲神经网络进行图像分割, 首先要面对的一个问题是: 如何将图像的像素值转换成脉冲神经元可以处理的脉冲序列。文献[7]采用了脉冲响应模型的脉冲神经网络进行图像分割和边缘检测。这种方法的特点是将输入层的图像像素值通过群体编码策略转换为脉冲序列, 并应用无监督或监督学习算法调整网络的连接权值, 在输出层得到图像的分割结果或边缘检测结果。文献[8]应用脉冲频率编码, 提出了 Integrate-and-Fire 神经元网络和 BP(Back-Propagation)神经网络的混合模型对彩色图像进行分割。这种方法的特点是将图像像素值通过特定结构的脉冲神经网络转换为脉冲序列, 然后将脉冲序列的频率统计结果输入 BP 神经网络, 其实质仍然是基于频率编码的传统神经网络对图像进行分割。

本文提出基于 Time-to-First-Spike 编码策略的线性编

**基金项目:** 国家自然科学基金资助项目(61165002); 甘肃省自然科学基金资助项目(1010RJZA019); 西北师范大学科研基金资助项目(NWNU-LKQN-10-3)

**作者简介:** 崔文博(1986—), 男, 硕士研究生, 主研方向: 神经网络, 图像处理; 蔺想红, 副教授、博士; 徐满意, 硕士

**收稿日期:** 2012-05-25 **修回日期:** 2012-07-20 **E-mail:** cbowen@126.com

码和非线性编码 2 种方法, 实现从图像像素灰度值到神经元脉冲的转换过程, 并重点分析了 2 种编码方法对图像分割的影响。

### 2 图像分割中像素值的编码策略

Time-to-First-Spike 编码是一种比较简单地将模拟量转变为脉冲时间的编码方法。它假设神经元只产生一个脉冲, 产生脉冲的时刻与模拟量的值成比例。这种思想来自于生物神经元处理信息的一些现象, 文献[9]通过生物实验说明在大脑中的每个处理步骤中, 各神经元最多只能发放一个脉冲。因此, 发放的第 1 个脉冲中将包含大部分的相关信息。由于简单易用, Time-to-First-Spike 编码策略不管是在研究生物神经元的的信息处理, 还是脉冲神经网络的数据编码方面都得到了广泛应用<sup>[10-11]</sup>。将 Time-to-First-Spike 编码策略应用于图像分割领域, 本文给出下面 2 类实现方式: 线性编码方法和非线性编码方法。

#### 2.1 线性编码方法

线性编码方法采用了从模拟量到脉冲时间的线性对应方式。假设图像中像素值为 0 的像素点所对应的神经元将首先发放脉冲, 并标记该脉冲发放时间为 0 ms。而图像中像素值为 255 的像素点所对应的神经元最后发放脉冲, 并标记该脉冲发放时间为  $T_{max}$  ms, 可根据下面的线性函数实现图像像素灰度值到神经元脉冲发放时间的转换:

$$t = \frac{T_{max}}{255} p \tag{1}$$

其中,  $p$  表示图像的像素灰度值;  $t$  表示神经元脉冲的发放时间;  $T_{max}$  表示脉冲的最大发放时间, 本文实验中  $T_{max}$  取值为 10 ms。图 1 给出了线性编码方法的示意图。图 1(a) 表示像素灰度值和脉冲发放时间的线性对应关系, 图 1(b) 显示了 3 个神经元在像素灰度值信号 25、128、250 输入后产生的脉冲时间, 其中在刺激开始后, 3 个神经元发放脉冲的时间分别为 0.98 ms、5.02 ms 和 9.8 ms。

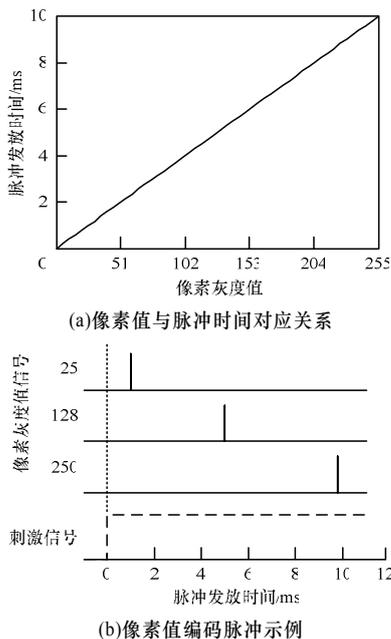


图 1 线性编码方法示意图

#### 2.2 非线性编码方法

非线性编码方法采用了从模拟量到脉冲时间的非线性对应方式。同样假设图像中像素值为 0 的像素点所对应的神经元最早发放脉冲, 而图像中像素值为 255 的像素点所对应的神经元最后发放脉冲, 具体从图像像素灰度值到神经元脉冲发放时间的转换关系采用如下式所示的 Sigmoid 函数:

$$t = \frac{T_{max}}{1 + \exp(\sigma(128 - p))} \tag{2}$$

其中,  $p$  表示图像的像素灰度值;  $t$  表示神经元脉冲的发放时间;  $\sigma$  表示非线性编码参数;  $T_{max}$  表示脉冲的最大发放时间, 本文实验中对  $T_{max}$  取值为 10 ms。

在非线性编码方法中, 参数  $\sigma$  的不同取值会让式(2)所对应曲线的形状发生比较明显的变化。当非线性编码参数  $\sigma$  接近于 1 时, 对应的曲线接近于阶跃函数, 而当非线性编码参数  $\sigma$  比较小的时候, 对应的曲线会接近于线性函数。图 2 给出了当非线性编码参数  $\sigma$  取值为 0.05 时, 非线性编码方法从图像的像素灰度值到脉冲发放时间的示意图。图 2(a) 表示像素灰度值和脉冲发放时间的非线性对应关系, 图 2(b) 显示了 3 个神经元在像素灰度值信号 25、128、250 输入后产生的脉冲时间, 其中在刺激开始后, 3 个神经元发放脉冲的时间分别为 0.06 ms、5 ms 和 9.98 ms。

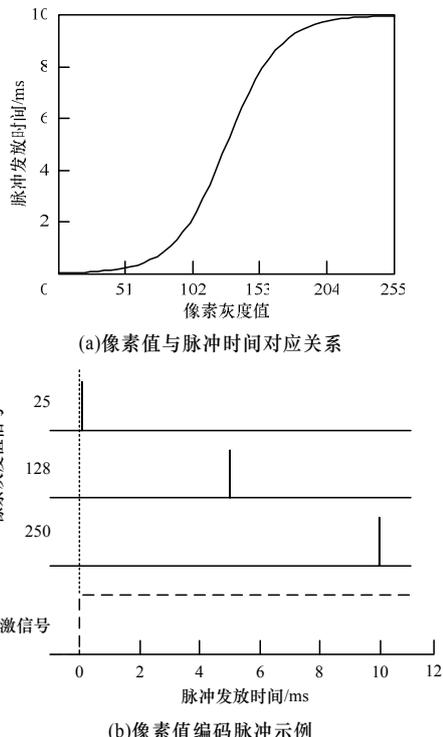


图 2 非线性编码方法示意图

### 3 脉冲神经网络图像分割的原理

人类的视觉系统可以高效地进行图像的分割与识别, 神经学科学家们已经发现从简单细胞到那些视网膜和外侧膝状体细胞有各种各样的神经细胞的感受野。在感受野和 Integrate-and-Fire 神经元模型<sup>[12]</sup>的基础上, 用于图像分割的 3 层脉冲神经网络模型如图 3 所示。该脉冲神经网络

络模型分为3层：输入层，中间层和输出层。

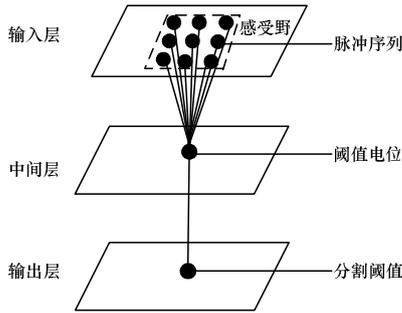


图3 脉冲神经网络图像分割模型

首先原始图像要经过输入层把图像的像素值转换为脉冲序列，转换方法则依据 Time-to-First-Spike 编码策略的线性编码或非线形编码。假设原始图像的大小为  $M \times N$ ，则输入层中就有  $M \times N$  个神经元与之对应，然后以该层中的每一个神经元  $A_i (i=1, 2, \dots, M \times N)$  为中心构成  $M \times N$  个  $3 \times 3$  大小的感受野。其中如果神经元位于输入层的边缘位置，则构成不了一个完整的  $3 \times 3$  大小的感受野，那么就在编码的过程中对这些空余的位置标记成不发放脉冲。

经过输入层以后，图像的像素值将以感受野为单位转换成脉冲序列，然后每一个感受野中的脉冲序列将以脉冲发放的先后顺序送到中间层中与输入层神经元  $A_i$  相对应的中间层神经元  $B_i (i=1, 2, \dots, M \times N)$ 。并且输入层感受野中的每一个神经元与中间层神经元  $B_i$  之间都将分别赋予一个连接权值  $\omega_{uv}$ ，权值  $\omega_{uv}$  的计算公式如下：

$$\omega_{uv} = W_{\max} \exp \left( - \left( \frac{\|F(u) - F(v)\|^2}{d} + \frac{\|X(u) - X(v)\|^2}{d} \right) \right) \quad (3)$$

其中， $u$  表示神经元  $A_i$  感受野中神经元对应的像素； $v$  表示与  $A_i$  对应的中间层神经元  $B_i$  所对应的像素； $\|F(u) - F(v)\|$  表示代表神经元的像素  $u$  和  $v$  的灰度差值； $\|X(u) - X(v)\|$  表示像素  $u$  和  $v$  之间的欧氏距离； $d$  为常数，取值为 3； $W_{\max}$  表示神经元之间的交互强度或权值尺度，本文中取值为 10 mV。

已知脉冲的发放时间和神经元之间的连接权值之后就可以求得神经元的膜电位，当中间层神经元的膜电位超过阈值电位时发放脉冲，并记录下脉冲的发放时间。当  $3 \times 3$  的感受野中的脉冲序列按照从小到大的顺序依次送到神经元  $B_i$  以后，神经元  $B_i$  的膜电位仍然没有超过阈值电位，则表示该神经元没有发放脉冲，该神经元脉冲发放时间用  $\infty$  表示。

第 3 层为输出层，同样该层的每一个神经元  $C_i (i=1, 2, \dots, M \times N)$  对应中间层的每一个神经元  $B_i$ 。在该层中设置一个分割阈值  $T_{th}$ ，当中间层神经元的脉冲发放时间超过分割阈值  $T_{th}$  时，不发放脉冲，并标记神经元  $C_i$  所对应的像素值为 255。当中间层神经元的脉冲发放时间没有超过分割阈值  $T_{th}$  时，则输出一个脉冲，并标记神经元  $C_i$  对应的像素值为 0。这样，在输出层根据分割阈值  $T_{th}$  就可以得到分割的二值图像。

### 4 实验分析

#### 4.1 图像分割结果的评价准则

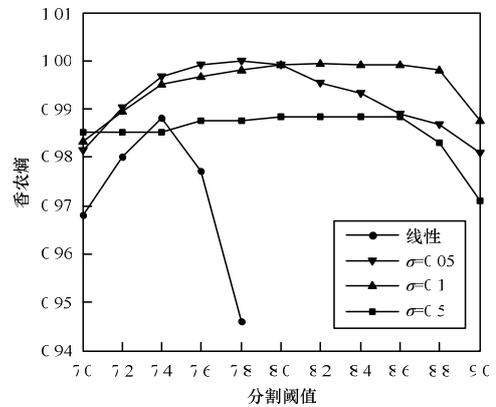
为了检验线性编码和非线性编码，以及非线性编码中参数  $\sigma$  的不同取值情况下图像分割结果的优劣，需要一个判别准则。目前许多文献提出了不同的评价准则，如最大香农熵准则、最大交叉熵准则、最大边缘数准则等。本文采用了常用的最大香农熵<sup>[13]</sup>的方法确定最佳分割结果，分割图像香农熵的计算公式如下：

$$H(P) = -P_1 \lg P_1 - P_0 \lg P_0 \quad (4)$$

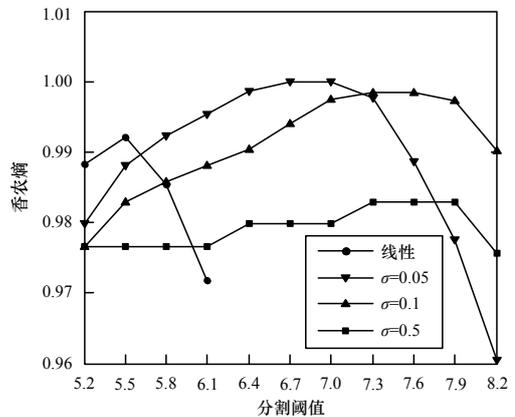
其中， $P_0$  和  $P_1$  分别表示分割后的二值图像中像素灰度值为 255 和 0 的概率。对大多数图像而言，最大香农熵代表了图像的信息量，如果分割后的图像的香农熵越大，分割图像从原始图像中得到的信息量越大，其分割图像的细节越丰富，且总体分割效果越好。所以，利用该准则进行判优，当  $H(P)$  取最大值时所对应的分割结果最佳。

#### 4.2 线性与非线性编码的比较

在本文实验中，Integrate-and-Fire 神经元模型的静息电位取值为 -70 mV，阈值电位取值为 -55 mV，膜时间常量取值为 20 ms。选取了  $256 \times 256$  大小的 Lena 图像和 Camera man 图像作为实验图像，分别对这 2 幅图像在不同的编码方法和非线性编码参数  $\sigma$  的不同取值下进行分割实验，并通过图像分割结果比较与分析，最后总结出 2 种编码方法的差异及其特点。2 种图像在不同编码方法下对应的分割结果曲线如图 4 所示。



(a) Lena 图像对应的分割结果曲线



(b) Camera man 图像对应的分割结果曲线

图4 2种图像在不同编码方法下对应的分割结果曲线

首先,从图4的2幅图像可以看出,当非线性编码参数 $\sigma=0.05$ 和 $0.1$ 时求得的最大香农熵值要大于在线性编码条件下求得的最大香农熵值。当Lena图像采用线性编码方法时取的最大香农熵为 $0.988\ 202$ ,采用非线性编码方法时,在非线性编码参数 $\sigma=0.05$ 条件下取得最大香农熵值 $0.999\ 92$ 。而文献[13]中应用脉冲耦合神经网络方法求得Lena图像的最大香农熵为 $0.999\ 8$ 。当Camera man图像采用线性编码方法时取得的最大香农熵为 $0.998\ 148$ ,采用非线性编码方法时,同样在非线性编码参数 $\sigma=0.05$ 条件下取得最大香农熵 $0.999\ 894$ 。而文献[13]中应用脉冲耦合神经网络方法求得的最大香农熵为 $0.999\ 5$ 。所以,当非线性的编码参数取得某一合适值时求得的最大香农熵大于线性编码条件下求得的最大香农熵,并且该条件下求得的最大香农熵也大于文献[13]中应用脉冲耦合神经网络方法求得的最大香农熵。从而也验证了非线性编码方法具有更好的生物可解释性。

其次,从图4(a)的Lena图像分割结果的变化趋势可以发现,对于线性编码方法只有当分割阈值 $T_{th}=7.4\ ms$ 时才会取得较大的香农熵。而对于非线性编码方法当参数 $\sigma=0.05$ 时,分割阈值 $T_{th}$ 从 $7.4\ ms\sim 8.6\ ms$ 之间都能取得较大的香农熵。当参数 $\sigma=0.1$ 时,分割阈值 $T_{th}$ 从 $7.4\ ms$ 到 $8.8\ ms$ 之间都能取得较大的香农熵。当参数 $\sigma=0.5$ 时,取值区间也较大但香农熵值较小。从图4(b)的Camera man图像同样可以发现类似的规律。对于线性编码方法只有当分割阈值 $T_{th}=5.5\ ms$ 时才会取得相对较大的香农熵。而对于非线性编码方法当参数 $\sigma=0.05$ 时,分割阈值 $T_{th}$ 从 $5.8\ ms\sim 7.3\ ms$ 之间都能取到较大的香农熵值。当非线性编码方法参数 $\sigma=0.1$ 时,分割阈值 $T_{th}$ 从 $6.7\ ms\sim 7.9\ ms$ 之间都能取到较好的香农熵值。当非线性编码方法参数 $\sigma=0.5$ 时,取值区间也较大但是香农熵值较小。所以在非线性编码方法中,当非线性编码参数 $\sigma=0.1$ 和 $0.05$ 时都可以在较大的取值区间取得分割图像较大的香农熵值。

2种图像在不同编码方法下对应的最佳分割结果如图5所示。

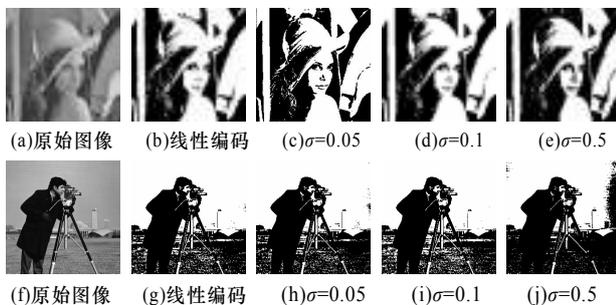


图5 2种图像在不同编码方法下对应的最佳分割结果

## 5 结束语

本文应用基于Time-to-First-Spike编码策略的脉冲神经网络进行图像分割,提出了Time-to-First-Spike编码策

略的线性编码和非线性编码2种方法,并重点分析了2种编码方法对图像分割结果的影响。对利用脉冲神经网络的图像分割模型得出的实验数据进行分析发现,非线性编码方法对图像的分割结果优于线性编码方法,并且最优分割结果所对应的香农熵大于传统神经网络所取得的香农熵。

对于今后的研究将主要着眼于2个方面:(1)对基于脉冲神经网络的图像分割,寻找更加具有生物可解释性的编码方法。(2)进一步优化脉冲神经网络图像分割的网络结构及其相应参数。

## 参考文献

- [1] Quiroga R Q, Panzeri S. Extracting Information from Neuronal Populations: Information Theory and Decoding Approaches[J]. Nature Reviews Neuroscience, 2009, 10(3): 173-185.
- [2] Johnson D H. Information Theory and Neural Information Processing[J]. IEEE Transactions on Information Theory, 2010, 56(2): 653-666.
- [3] Cho M W, Choi M Y. Theory of Neural Communication Based on Spatio-temporal Coding[J]. BMC Neuroscience, 2011, 12(S1): 38.
- [4] Ghosh-Dastidar S, Adeli H. Spiking Neural Networks[J]. International Journal of Neural Systems, 2009, 19(4): 295-308.
- [5] Buhmann J, Lange T, Ramacher U. Image Segmentation by Networks of Spiking Neurons[J]. Neural Computation, 2005, 17(5): 1010-1031.
- [6] 蔡荣太, 吴庆祥. 基于脉冲神经网络的红外目标提取[J]. 计算机应用, 2010, 30(12): 3327-3330.
- [7] Meftah B, Lezoray O, Benyettou A. Segmentation and Edge Detection Based on Spiking Neural Network Model[J]. Neural Processing Letters, 2010, 32(2): 131-146.
- [8] Wu Qingxiang, McGinnity T M, Maguire L, et al. Colour Image Segmentation Based on a Spiking Neural Network Model Inspired by the Visual System[C]/Proc. of the 6th International Conference on Intelligent Computing. Heidelberg, Germany: [s. n.], 2010: 49-57.
- [9] Thorpe S, Fize D, Marlot C. Speed of Processing in the Human Visual System[J]. Nature, 1996, 381(6582): 520-522.
- [10] Thorpe S J, Delorme A, Van Rullen R. Spike-based Strategies for Rapid Processing[J]. Neural Networks, 2001, 14(6-7): 715-725.
- [11] Bohte S M, Kok J N, La Poutré J A. Error-backpropagation in Temporally Encoded Networks of Spiking Neurons[J]. Neurocomputing, 2002, 48(1-4): 17-37.
- [12] Izhikevich E M. Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting[M]. Cambridge, USA: MIT Press, 2007.
- [13] 马义德, 戴若兰, 李廉. 一种基于脉冲耦合神经网络和图像熵的自动图像分割方法[J]. 通信学报, 2002, 23(1): 46-51.

编辑 顾逸斐