

文章编号 1004-924X(2007)03-0401-06

基于 SOFM 神经网络的图像融合二值化方法

潘梅森, 荣秋生

(湖南文理学院 计算机科学与技术系, 湖南 常德, 415000)

摘要:提出了一种基于自组织特征映射(SOFM)神经网络的图像融合二值化方法。介绍了 SOFM 神经网络的特点及学习算法,根据 SOFM 的聚类确定图像第一阈值作为循环迭代的初始值,对整幅图像进行循环迭代得到第二阈值,使用第二阈值对原始图像进行二值化,得到第一幅待融合图像;通过改进的 Bernsen 方法对原始图像进行二值化,得到第二幅待融合图像;最后根据图像灰度值选小的原则作为图像融合方法,得到最终的二值化图像。该方法既能有效地消除伪影,又能较好地分离字符和文字。模拟实验结果表明,该方法的二值化效果明显优于 Bernsen 方法和 Ostu 方法,且具有良好的适应性。

关键词:图像融合;二值化;阈值;SOFM 神经网络;像素

中图分类号:TP391;TP183 **文献标识码:**A

Image fusion binarization method based on SOFM neural network

PAN Mei-sen, RONG Qiu-sheng

(*Department of Computer Science and Technology, Hunan University of Arts and Science, Changde 415000, China*)

Abstract: An image fusion binarization method based on Self-organization Feature Map(SOFM) neural network is presented. The characteristic and the learning algorithm of SOFM neural network are introduced. The first threshold of the image is derived from the clustering characteristic of SOFM, and it is treated as a initialization of the circulated iteration method to obtain the second threshold of the image. The binarization for the original image is carried on using the second threshold to obtain the first fusing image. Then, the binarization for the original image is performed more again using the improved Bernsen method to obtain the second fusing image. Finally, two fusing images are merged based on the minimum principle of image gray values to get the final binarization image. This method can effectively eliminate the ghost and can also separate the characters very well. Experimental results show that the method is effective, and its binarization effect surpasses Bernsen method and Ostu method obviously, moreover it has good compatibility.

Key words: image fusion; binarization; threshold; SOFM neural network; pixel

1 引言

图像二值化是图像预处理中的一项重要技术,在模式识别、光学字符识别、医学成像等方面都有重要应用。由于图像二值化过程中将会损失原图像的许多有用信息,因此在进行二值化处理时,能否保留原图像上的主要特征十分关键。

在图像二值化过程中,阈值的选择是非常重要的。选择阈值的方法有整体阈值法和局部阈值法。整体阈值法是指在二值化过程中只使用一个全局阈值,它将图像的每个像素的灰度值与阈值进行比较,若大于阈值,则取为白色;否则,取为黑色。局部阈值法由当前像素灰度值与该像素邻域灰度值来确定像素的阈值^[1]。Ostu 方法^[2]是典型的整体阈值法,它是使用穷尽的方法来得到最优阈值,不需要对直方图做预处理而直接通过计算求得图像阈值的方法,它的判断标准是使分开的两部分之间方差最大。由于整幅图像使用一个阈值,因此对输入图像量化噪声和不均匀光照等情况抵抗能力较差,字符粘连现象较严重。Bernsen 方法^[2]是典型的局部阈值法,该方法虽然能够根据局部灰度特性来自适应地选择阈值,有较大的灵活性,但是在字符图像二值化中存在出现严重伪影现象的问题。

本文提出了基于自组织特征映射神经网络(Self-Organization Feature Map, SOFM)的图像融合二值化方法,其基本思想是:首先利用 SOFM 把整个图像灰度空间分类,计算各类的平均灰度值,并和相应权值相乘求得图像的第一阈值,根据第一阈值采用循环迭代方法求得第二阈值,使用第二阈值对原始图像进行二值化处理,得到待融合的第一幅图像;然后使用改进的 Bernsen 方法对原始图像进行二值化处理,得到待融合的第二幅图像;最后根据两幅图像像素灰度值选小方法进行图像融合,得到最终的二值化图像。基于 SOFM 神经网络的图像融合的二值化方法在很大程度上克服了字符图像出现伪影、字符不能很好分离等现象,具有较好的适应性,为图像二值化提供了一种新的途径。论文第二部分详细介绍了这一方法的实现过程,第三部分对该方法进行模拟实验及结果比较。

2 基于 SOFM 神经网络的图像融合二值化方法

基于 SOFM 神经网络的图像融合二值化方法大体上包括三个步骤:①使用 SOFM 神经网络对原始图像二值化处理;②使用改进的 Bernsen 方法对原始图像进行二值化处理;③对①步和②步产生的二值化图像进行融合。

2.1 使用 SOFM 神经网络对原始图像二值化处理

作为竞争神经网络的一种应用,许多学者已经把 Kohonen 的 SOFM 用于数据聚类。该神经网络具备这样的特点:输入节点和输出神经元之间通过权值互连;在输出神经元之间进行竞争选择,输出神经元之间存在侧抑制^[3-4]。因此, SOFM 是一种具有侧向联想能力的二维网格结构网络,如图 1 所示。神经元分布在二维网格上,而 SOFM 正是通过引入网格,在神经元之间建立拓扑邻域关系,使侧抑制的模拟得以实现,拓扑邻域关系如图 2 所示。

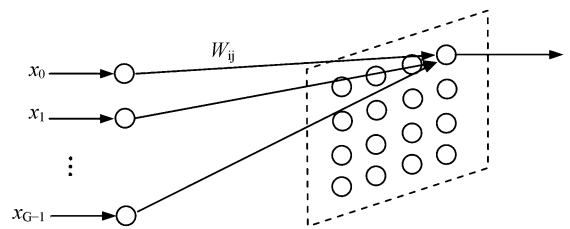


图 1 自组织特征映射神经网络模型

Fig. 1 SOFM model

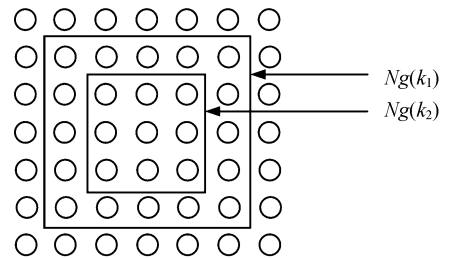


图 2 自组织特征映射神经网络拓扑邻域

Fig. 2 SOFM neighborhood of topology

设训练矢量空间为 \mathbf{X} , 个数为 I , x_q 是 \mathbf{X} 的取值, $q \in [0, I-1]$, 网络输入节点数为 k (等于训练矢量维数), 输出神经元个数为 G (即分类数),

各输入节点到各输出神经元的权值为 w_{ij} , $i \in [0, G-1]$, $j \in [0, k-1]$, w_{ij} 也是类集 $\mathbf{Y} = \{y_i; i=0, 1, 2, \dots, G-1\}$ 中类 y_i 的第 j 个分量。SOFM 学习算法包含竞争、合作和更新三个过程,整个学习算法描述如下:

(1) 设置变量:迭代总次数为 S , 并且有:

$\mathbf{x}_q(n) = [x_{q0}(n), x_{q1}(n), x_{q2}(n), \dots, x_{qk-1}(n)]^T$
为训练矢量, $q=0, 1, 2, \dots, I-1$;

$\mathbf{W}_i(n) = [w_{i0}(n), w_{i1}(n), w_{i2}(n), \dots, w_{ik-1}(n)]^T$
为权值矢量, $i=0, 1, 2, \dots, G-1$;

(2) 初始化:将权值矢量 \mathbf{W}_i 用较小随机矢量初始化,设置初始学习速率 $\eta(0)$ 。

(3) 抽样:从训练矢量空间中选取一个训练矢量 \mathbf{x}_q 做为输入。

(4) 选取获胜神经元 C :通过欧氏距离最小的原则来确定:

$$\|\mathbf{x}_q - \mathbf{W}_C\| = \min_i \|\mathbf{x}_q - \mathbf{W}_i\|, i=0, 1, 2, \dots, G-1, \quad (1)$$

从而实现了神经元的竞争过程。

(5) 对获胜神经元拓扑邻域 $N_g(n)$ 内的兴奋神经元,以 Hebb 学习规则:

$$\mathbf{W}_i(n+1) = \mathbf{W}_i(n) + \eta(n)(\mathbf{x}_q - \mathbf{W}_i(n)), \quad (2)$$

以更新神经元的权值矢量,从而实现了神经元的合作和更新过程。

(6) 更新学习速率 $\eta(n)$ 和拓扑邻域:

$$\eta(n+1) = \eta(0) \left(1 - \frac{n+1}{S}\right), \quad (3)$$

$$N_g(n+1) = INT \left[N_g(0) \left(1 - \frac{n+1}{S}\right) \right], \quad (4)$$

(7) 判断迭代次数 n 是否超过 S , 如果 $n \leq S$ 就转到③, 否则结束迭代过程。

SOFM 学习算法属于无导师学习算法,网络经过以上学习训练后,如果训练充分且算法收敛,则 SOFM 具有聚类能力,每一个输出神经元权值矢量表示一个分类,全部输出神经元权值矢量就表示了类集 \mathbf{Y} 。

假设数字图像 \mathbf{P} 大小为 $M \times N$, 左上角像素点位置为 $(1, 1)$ 。在进行图像灰度值分类时,首先把图像分成 $k = l \times l$ 尺寸的子图像,按行或按列转换为 $l^2 \times 1$ 维列矢量,这样就构成了分类器的输入矢量空间 \mathbf{X} ; 然后按失真度最小原则把灰度值分成 G 类,构成分类器的类空间 \mathbf{Y} 。每个类是 k 维矢量,在得到类 y_i 后,分别把其分量相加求

其平均值,即:

$$\text{mean}_i = \frac{1}{k} \sum_{j=0}^{k-1} y_{i,j} \quad i = 0, 1, 2, \dots, G-1, \quad (5)$$

其中 $y_{i,j}$ 表示类 y_i 的第 j 维分量。统计出类 y_i 在输入矢量空间 \mathbf{X} 中的个数 num_i , 然后计算第一阈值 T_1 :

$$T_1 = \sum_{i=0}^{G-1} \frac{\text{num}_i}{I} \cdot \text{mean}_i, \quad (6)$$

在求得第一阈值 T_1 后,再采用循环迭代方法求第二阈值 T_2 。给定一个 ξ , ξ 表示前后两次迭代阈值差值的临界,是迭代结束条件,也表示了在 ξ 误差范围内 T_2 和 T_1 相等, ξ 选择过大, T_2 结果不精确;选择过小,循环迭代次数增加,一般说来,选择 $\xi \in [0.5, 2.5]$ 。循环迭代方法基本步骤描述如下:

(1) 以 T_1 为阈值,把整个图像像素灰度空间划分为两个子空间 π_1 和 π_2 , 并分别统计这两个空间像素个数为 n_1 和 n_2 , 空间划分规则如下:

$$\begin{aligned} \pi_1 &= \{f(i, j) \geq T_1 \mid i=1, 2, \dots, M; j=1, 2, \dots, N\} \\ \pi_2 &= \{f(i, j) < T_1 \mid i=1, 2, \dots, M; j=1, 2, \dots, N\}, \end{aligned} \quad (7)$$

(2) 分别把 π_1 和 π_2 空间像素灰度值相加求其平均值为 average_{π_1} 和 average_{π_2} , 即

$$\begin{aligned} \text{average}_{\pi_1} &= \frac{1}{n_1} \sum_{f(i, j) \in \pi_1} f(i, j) \\ \text{average}_{\pi_2} &= \frac{1}{n_2} \sum_{f(i, j) \in \pi_2} f(i, j), \end{aligned} \quad (8)$$

(3) 计算阈值 T_2 :

$$T_2 = \frac{\text{average}_{\pi_1} + \text{average}_{\pi_2}}{2}, \quad (9)$$

(4) 如果 $|T_2 - T_1| > \xi$, 则 $T_1 = T_2$, 转①继续循环迭代; 否则,退出循环, T_2 即为所求。

在求得第二阈值 T_2 后,即可对原始图像进行二值化操作, $g_1(i, j)$ 表示 $f(i, j)$ 二值化后的值。其规则如下:

$$g_1(i, j) = \begin{cases} 255 & f(i, j) \geq T_2 \\ 0 & f(i, j) < T_2, \end{cases} \quad (10)$$

2.2 使用改进的 Bernsen 方法对图像进行二值化处理

Bernsen 方法^[5]是一种局部阈值法,它是根据局部灰度特点来自适应地选取阈值。首先找出整个图像 \mathbf{P} 中灰度值的最大值和最小值,分别记

为 $\max(\mathbf{P}_{M \times N})$ 和 $\min(\mathbf{P}_{M \times N})$ 。采用 3×3 窗口在图像上滑动,该窗口中心像素灰度值为 $f(i, j)$, 则该窗口内所有像素值构成如下集合:

$$\mathbf{W}_{i,j} = \{f(i+k, j+r) | k, r = -1, 0, 1\}, \quad (11)$$

$g_2(i, j)$ 表示 $f(i, j)$ 二值化后的值,则改进的 Bernsen 方法描述如下:

(1) 找出该窗口中像素的最大值和最小值为 $\max(\mathbf{W}_{i,j})$ 和 $\min(\mathbf{W}_{i,j})$;

(2) 计算该窗口中心像素的二值化阈值 $T(i, j)$:

$$T(i, j) = \frac{\max(\mathbf{W}_{i,j}) + \min(\mathbf{W}_{i,j})}{2}, \quad (12)$$

(3) 对该窗口中心像素进行二值化:

① 如果该像素灰度值 $f(i, j) = \max(\mathbf{P}_{M \times N})$, 则该像素二值化为 $g_2(i, j) = 255$;

② 如果该像素灰度值 $f(i, j) = \min(\mathbf{P}_{M \times N})$, 则该像素二值化为 $g_2(i, j) = 0$;

③ 如果该像素灰度值 $\min(\mathbf{P}_{M \times N}) < f(i, j) < \max(\mathbf{P}_{M \times N})$, 则该像素二值化为:

$$g_2(i, j) = \begin{cases} 255 & f(i, j) \geq T(i, j) \\ 0 & f(i, j) < T(i, j) \end{cases}, \quad (13)$$

2.3 二值图像进行图像融合

图像融合^[6]是一种对多幅图像进行提取与综合处理的技术,其目的是融合多幅图像中的冗余信息和互补信息,强化图像中的信息,增加图像的可靠性以使融合后的图像具有良好的鲁棒性和可视性,以利于对图像的进一步分析,理解及目标的检测、识别或跟踪。

图像融合方法很多,由于本文方法只涉及到二值图像,因此选择了像素灰度值选小方法进行图像融合。假设对 $\mathbf{G}_1, \mathbf{G}_2, \dots, \mathbf{G}_T$ 个大小为 $M_1 \times N_1$ 的源图像进行融合,融合后图像为 \mathbf{G} , $g_1(i, j), g_2(i, j), \dots, g_T(i, j)$ 分别表示 $\mathbf{G}_1, \mathbf{G}_2, \dots, \mathbf{G}_T$ 在点 (i, j) 的像素值。则像素灰度值选小的融合过程可表示为:

$$g(i, j) = \min\{g_1(i, j), g_2(i, j), \dots, g_T(i, j)\} \quad i = 1, 2, \dots, M_1; j = 1, 2, \dots, N_1, \quad (14)$$

即在融合处理时,比较各幅图像中对应位置 (i, j) 处像素灰度值的大小,以其中灰度值小的像素作为融合后图像 \mathbf{G} 在位置 (i, j) 处的像素值。

根据上述图像融合方法,在 $g_1(i, j)$ 和 $g_2(i, j)$ 确定后,接下来对图像进行融合, $g(i, j)$ 为

$g_1(i, j)$ 和 $g_2(i, j)$ 融合后的值。根据上述融合规则,本文的融合过程如下:

$$g(i, j) = \min\{g_1(i, j), g_2(i, j)\} \quad i = 1, 2, \dots, M; j = 1, 2, \dots, N. \quad (15)$$

3 模拟实验和结果比较

在模拟实验中,运行环境为:CPU 为 AMD Athlon 3000+, 操作系统为 Windows2000, 软件环境为 MATLAB6.5。以汽车牌照图像为例,对 Ostu 方法、Bernsen 方法和本文方法进行了比较。图 3 是车牌号码正常拍摄时的试验结果;图 4 是车牌号码倾斜时的试验结果;图 5 是车牌号码上



图 3 车牌号码正常拍摄时三种方法比较
Fig. 3 Three methods comparison under normal photography for vehicle license plate

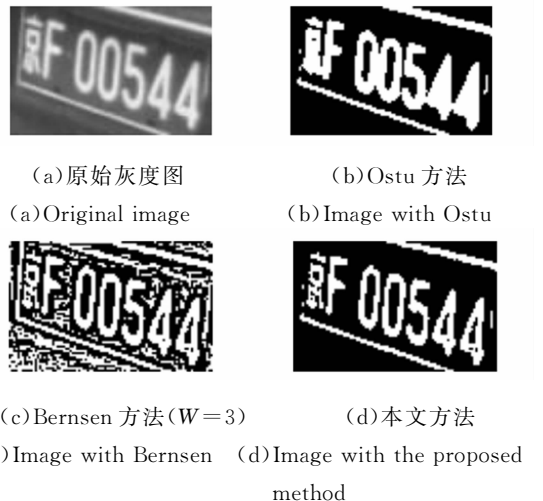


图 4 车牌号码倾斜时三种方法比较
Fig. 4 Three methods comparison under tilting photography for vehicle license plate

方有阴影时的试验结果。在实验时,本文方法把整个实验图像灰度值分成4类,每类的维数为9。



图5 车牌号码上方有阴影时三种方法比较
Fig. 5 Three methods comparison under shadow photography for vehicle license plate

从图3可以看到,当车牌号码正常拍摄时,图像用Ostu方法二值化后图像清晰,但“鲁”字上方的“鱼”字头和下方的“日”字笔画没有正常分离出来;图像用Bernsen方法二值化后文字轮廓清晰,但有明显的伪影;用本文方法二值化的结果和Ostu方法基本相同,但“鲁”字笔画已经完全分离,该字清晰可辨,非常利于识别。

从图4可以看到,当车牌号码倾斜时,图像用Ostu方法二值化后图像较模糊,不利于字符的分

割,特别是“京”字和“4”字;图像用Bernsen方法二值化后虽然文字轮廓比较清晰,但有明显的伪影,不利于字符的识别;而本文的方法二值化后,每个字符是分离的,伪影基本消除,但也存在不足,如“4”字中出现了空洞。

从图5可以看到,当车牌号码上方有阴影时,图像用Ostu方法二值化后比较清晰,但阴影部分的文字笔画较细,文字下方没有完全分离;图像用Bernsen方法二值化后文字轮廓较清晰,文字下方已经分离,但有明显的伪影;用本文方法二值化的结果和Ostu方法大致相同,但字符笔画更细腻,文字下方已经完全分离出来,便于图像的后继处理。

4 结束语

基于SOFM神经网络的图像融合的二值化方法采用SOFM神经网络分类方法确定图像第一阈值,将其做为循环迭代方法的初始值,并求出第二阈值对原始图像进行二值化,得到待融合的第一幅图像;采用改进的Bernsen方法对原始图像二值化,得到待融合的第二幅图像;最后对两幅二值化图像进行融合,得到最后的二值图像。实验结果表明该方法的二值化效果明显优于Bernsen方法,在车牌号码倾斜时大大优于Ostu方法,而且该方法具有良好的适应性,是一种较好的图像二值化方法。

参考文献:

- [1] 何永义,夏俊昇.工程图像的矢量化技术[J].光学精密工程,2002,10(8):374-378.
HE Y Y, XIA J SH. Vectorization technology for engineering image[J]. *Opt. Precision Eng.*, 2002, 10(8): 374-378. (in Chinese)
- [2] 陈丹,张峰,贺贵明.一种改进的文本图像二值化算法[J].计算机工程,2003,29(13):85-86.
CHEN D, ZHANG F, HE G M. An improved binarization algorithm for document image [J]. *Comput. Eng.*, 2003, 29(13): 85-86. (in Chinese)
- [3] 高隽.人工神经网络原理及仿真实例[M].北京:机械工业出版社,2003.
GAO J. *Artificial Neural Networks Principle and Simulation Example*[M]. Beijing: Mechanical Industry Press, 2003. (in Chinese)
- [4] 黄勇,陈琳.一种基于SOFM聚类的星图识别算法[J].光学精密工程,2004,12(6):346-351.
HUANG Y, CHEN L. Star pattern recognition algorithm based SOFM clustering function[J]. *Opt. Precision Eng.*, 2004, 12(6): 346-351. (in Chinese)

- [5] 黄志斌. 面向车辆牌照的 L 快速二值化算法[J]. 华侨大学学报(自然科学版), 2002, 23(4): 427-430.
HUANG ZH B. Algorithm of L quick binarization catering to license plate of vehicle[J]. *J. Huaqiao Univ. (Natural Science)*, 2002, 23(4): 427-430. (in Chinese)
- [6] 程英蕾, 赵荣椿, 李卫华, 等. 基于像素级的图像融合方法研究[J]. 计算机应用研究, 2004, 23(5): 169-172.
CHENG Y L, ZHAO R CH, LI W H, *et al.*. Overview of methods of data fusion for images based on pixel-level[J]. *Appl. Res. Comput.*, 2004, 23(5): 169-172. (in Chinese)

作者简介:潘梅森(1972—), 男, 湖南临澧人, 高级工程师, 研究方向为图像处理、人工神经网络。E-mail: pmsjj@126.com;
荣秋生(1973—), 男, 湖南临澧县人, 讲师, 研究方向为多媒体技术、数据挖掘。