

# 基于细胞神经网络和网格特征的碑刻文字识别<sup>①</sup>

国方媛<sup>1</sup> 李国东<sup>2</sup>

(1.华北电力大学 经济与管理学院 北京 102206; 2.华北电力大学 数理学院 北京 102206)

**摘要:** 碑刻作为文字的载体, 具有极其重要的历史价值和艺术价值。但是, 因为时间的久远, 其拓本大量的噪声对碑刻文字的辨识产生一定的影响。针对这种情况, 提出了一种新方法, 首先利用细胞神经网络(CNN)技术, 去除单个碑刻文字的噪声, 然后用网格特征提取方法对其进行识别。本文在MATLAB平台上进行了实例模拟, 取得了较好的效果。

**关键词:** 细胞神经网络; 网格特征; 二维中值滤波; 碑刻; 文字识别

## Table Recognition Based on CNN and Grid Feature

GUO Fang-Yuan<sup>1</sup>, LI Guo-Dong<sup>2</sup>

(1.Department of Economics and Management, North China Electric Power University, Beijing 102206, China; 2.Department of Mathematics and Physics, North China Electric Power University, Beijing 102206, China)

**Abstract:** As a carrier of the text, the table has very important historical and artistic value. However, because of a long time, a lot of noises on the rubbings have impact on recognizing text. This paper proposes a new method for this case. Firstly, remove the noises of a single table using CNN technology. And then recognize the table using the grid feature extraction method. The simulation example is based on MATLAB software, and we achieve good results.

**Keywords:** CNN; grid feature; 2D median filter; table; text recognition

## 1 引言

碑刻作为一种以石头为载体的特殊文献, 向来为不同领域的研究者所重视, 具有重要的参考价值<sup>[1]</sup>:

(1) 碑刻一般除少数伪造者外, 多是第一手资料, 不像史书经过传抄刊刻而多有脱讹, 故而被历史学家当作十分重要的史料, 应用于历史写作。

(2) 文物学家把碑刻作为一种文物来搜集、记录和保护, 进行碑文的抄录和研究。碑刻文字是他们研究、观摩和学习书法艺术的重要资料。

(3) 碑刻文字既反映区域社会和民间生活, 也反映国家或地方性制度。它记载的是普遍承认的行为规范, 可以反映特定时代特定人群的心态。碑刻对研究历史上普通人的生活, 也是一宗宝贵的财富。

随着时间的推移, 年代越早的碑刻拓本越模糊不清, 含有大量的噪声, 这对碑刻文化无疑是一种严重的损失。针对这种情况, 本文提出了去除噪声识别文字的新方法。

本文构成如下: 第一部分为引言, 介绍了碑刻的重要价值; 第二部分介绍了利用细胞神经网络技术去除噪声的方法; 第三部分应用网格特征提取法对去除噪声的碑刻文字进行识别; 最后是结论部分。

## 2 碑刻文字去噪处理

### 2.1 二维中值滤波

中值滤波器是一种非线性平滑滤波器, 在1971年由J. W. Juke首先提出。它对滤除脉冲干扰及图像

<sup>①</sup> 收稿时间:2010-03-06;收到修改稿时间:2010-03-30

扫描噪声最为有效,可以既去除噪声又保持图像的细节。中值滤波器的主要功能是让与周围像素灰度值的差比较大的像素改取与周围像素值相近的值,从而可以去掉孤立的噪声点。由于它不是简单地取均值,所以产生的模糊比较少。

二维中值滤波可以利用某种形式的二维窗口来表示。设  $\{x_{i,j}, (i,j) \in N^2\}$  表示数字图像各点的灰度值,滤波窗口为  $A$  的二维中值滤波可定义为

$$y_{i,j} = \underset{A}{\text{Med}}\{x_{i,j}\} \quad (1)$$

$$= \text{Med}\{x_{(i+r),(j+s)}, (r,s) \in A, (i,j) \in N^2\}$$

其中,  $y_{i,j}$  表示二维中值滤波后数字图像各点的灰度值。

假设窗口内各点的像素值为  $\begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix}$ , 将它

们按由小到大的顺序排列  $x_{12}, x_{33}, x_{31}, x_{13}, x_{32}, x_{11}, x_{23}, x_{22}, x_{21}$ , 那么  $x_{32}$  就是二维中值滤波的输出,中心点的像素值由  $x_{22}$  变为  $x_{32}$  [2]。二维中值滤波的窗口形状和尺寸对滤波效果影响很大,根据不同的图像内容和不同的应用要求,往往采用不同的窗口形状和尺寸。一般对于有缓慢变化的较长的轮廓线物体的图像,采用方形或圆形窗口为宜。

## 2.2 CNN 及其去噪模板

细胞神经网络(Cellular Neural Network, CNN)是由“局部耦合的细胞”所组成的空间阵列,其中每个细胞均为一个非线性动力系统——具有输入、输出和根据某些动力学定律而演化的态变量。CNN 定义其网络中的任一细胞仅与其邻域内的细胞相连。这些相连的细胞通过权重直接相互作用,非直接相连的细胞则通过网络的动力学传输效应间接相互作用。一个  $M \times N$  的细胞神经网络安排在  $M$  行  $N$  列的二维阵列空间中。CNN 最初作为电路上可实现的模型自 1988 年由 Chua 等人提出以来[3],其形式不断得到推广[4,5],现已成为图像处理的新工具,可以用来去除图像噪声、探测图像边缘等。标准 CNN 的状态方程为:

$$\begin{aligned} \dot{x}_{i,j} &= -x_{i,j} + \sum_{C(i+k),(j+l) \in S_r(i,j)} a_{k,l} y_{(i+k),(j+l)} + \sum_{C(i+k),(j+l) \in S_r(i,j)} b_{k,l} u_{(i+k),(j+l)} + z_{i,j} \\ &= -x_{i,j} + \sum_{k=-r}^r \sum_{l=-r}^r a_{k,l} y_{(i+k),(j+l)} + \sum_{k=-r}^r \sum_{l=-r}^r b_{k,l} u_{(i+k),(j+l)} + z_{i,j} \quad (2) \\ \Delta & \\ &= -x_{i,j} + A(y_{i,j}) + B(u_{i,j}) + z_{i,j} \end{aligned}$$

标准 CNN 的输出方程为:

$$y_{i,j} = \frac{1}{2}(|x_{i,j} + 1| - |x_{i,j} - 1|), \quad i = 1, 2, \dots, M; j = 1, 2, \dots, N \quad (3)$$

其中,  $x_{i,j}$  表示数字图像各点的灰度值,  $y_{i,j}$  表示经过细胞神经网络处理后数字图像各点的灰度值,  $u_{i,j}$  表示输入的数字图像的灰度值,  $z_{i,j}$  表示细胞的阈值。参数  $r$  称为影响球半径。参数  $a_{k,l}$ ,  $b_{k,l}$  对应于电路中相应的  $A$  模板和  $B$  模板。这是细胞神经网络研究及其特征表现的核心。适当地选择 CNN 参数,可以产生奇妙的图像处理功能。

下面的 CNN 模板可以实现去除噪声的功能:

大范围任务:

- ① 原始图像: 一幅静态的灰度图像  $P$ , 其各点的灰度值为  $x_{i,j}$ ;
  - ② 输入:  $U(t) = P$ , 即初始输入为灰度图像  $P$ ;
  - ③ 输出:  $Y(t) \Rightarrow Y(\infty)$  = 没有孤立黑色像素的图像。
- 局部规则:

$$u_{i,j} \quad y_{i,j}(\infty)$$

- ① 白色像素  $\rightarrow$  白色像素;
- ② 黑色像素  $\rightarrow$  白色像素, 如果周围都是白色像素;
- ③ 黑色像素  $\rightarrow$  黑色像素, 如果周围至少有一个黑色像素;
- ④ 灰色像素  $\rightarrow$  黑色像素, 如果  $\nabla^2 > 1$ ;
- ⑤ 灰色像素  $\rightarrow$  白色像素, 如果  $\nabla^2 < 1$ ;
- ⑥ 灰色像素  $\rightarrow$  白色、黑色或灰色像素, 如果  $\nabla^2 = 1$ 。其中,  $\nabla^2 = \sum_{k=-r}^r \sum_{l=-r}^r b_{k,l} u_{(i+k),(j+l)}$ 。

定理. 如果给定如下的模板  $A$ 、模板  $B$  和阈值

$$Z, \quad A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 8 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad Z = -1, \quad \text{那么细}$$

胞神经网络能够实现局部规则。

## 2.3 实例模拟

根据以上所述理论,提出去除碑刻文字图像噪声的新算法,其步骤如下:

- ① 对原始图像进行中值滤波;
- ② 利用上述 CNN 模板处理步骤 1 中值滤波后的图像;

③ 对步骤 2 处理后的图像进行第二次中值滤波。图(a)是碑刻拓片上分割下来的一个单字，可以看到，图像伴有大量的噪声，含有很多无效的像素点，这将极大地影响后续处理。为了进一步识别文字，必须去除噪声，加强图像中的有用信息。

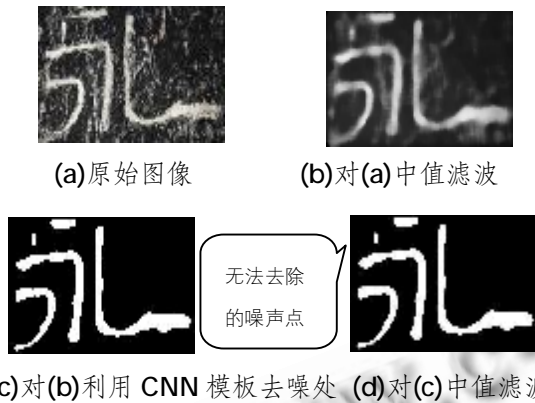


图 1 去除碑刻文字图像噪声的过程

分析：首先，采用 MATLAB 中的 medfilt2 二维中值滤波函数，以 9×9 的方形窗口对图(a)进行中值滤波，处理后的图像如图(b)所示。中值滤波后的图像去除了一些细小的噪声，文字的大体轮廓显现出来，而且文字笔画边界清晰可见。因为中值滤波作为一种局部平滑技术，能够很好地突出整个文字的边界特征，这也是本文采用它的一个主要原因。但是中值滤波后的图像(b)比较模糊，为此，利用上述的 CNN 模板

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 8 & 1 \\ 1 & 1 & 1 \end{bmatrix}, Z = -1$$

继续进行去噪处理，处理后的图像如图(c)所示。经过上面的处理，绝大部分噪声已经去除，为了使图像更为平滑并且最大程度地减少噪声，对图(c)以 5×5 的方形窗口进行第二次中值滤波，处理后的图像如图(d)所示。图(d)中的文字边界明显平滑许多，清除了不少的毛刺，图像噪声基本去除，为碑刻文字的识别工作做了良好的铺垫。但图像中存在一点无法去除的噪声点，这点可能是由于长期的风化作用或人为因素的干扰，笔画的力度与字体笔画的力度大致相同，以致无法彻底清除。文字还有待下面进一步的识别。

### 3 碑刻文字识别的实现

去除噪声后的碑刻文字还不能直接进行识别，要

通过归一化、二值化、网格特征提取和相似度判别等步骤。

#### 3.1 归一化

归一化包括位置归一化和尺寸归一化。

##### (1) 位置归一化

为了消除文字位置上的偏差，即将整个文字移到规定的位置上，称为位置归一化。本文采用最大梯度差检测法来实现文字的位置归一化，即确定文字的外接矩形。

对于数字图像，梯度(Gradient)是一阶导数的二维等效式，因此通过求取梯度可以判断边缘点。图像在位置的梯度向量定义为

$$\nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix}^T \quad (4)$$

式中  $G_x$  和  $G_y$  分别是梯度在  $x$ ， $y$  方向上的分量<sup>[6]</sup>。

由于文字与背景的色彩对比度强烈，而且文字笔画具有正负梯度交替变化的特征，利用最大梯度差进行检测文字，检测后的图像行经合并后即能形成文字图像区域。

定义从背景进入文字笔画处的梯度值为  $G_{in}$ ，从笔画进入背景的梯度值  $G_{out}$ ， $G_{in}$  与  $G_{out}$  符号相反，梯度的绝对值接近相等。在文字笔画的附近选取某一区域，求得区域内的最大梯度值  $G_{max}$  和最小梯度值  $G_{min}$ ，则该区域内的最大梯度差 MGD (Maximum Gradient Different) 为

$$MGD = G_{max} - G_{min} \quad (5)$$

设定阈值，最大梯度差 MGD 大于阈值的即为文字图像行<sup>[7]</sup>。

##### (2) 尺寸归一化

不同尺寸大小的文字，归一化后成为同一尺寸大小的文字，称为尺寸归一化。本文采用 MATLAB 中的 imresize 函数，用双三次插值法调整文字大小。

#### 3.2 二值化

二值图像是一种所有像素值只能在两种可能的离散值中取一的图像，也称为黑白图像。在 MATLAB 中，二值图像用一个由 0(黑)和 1(白)组成的二值矩阵表示，1 表示该像素处于前景，0 表示该像素处于背景。这种方式来操作图像可以更容易识别出图像的结构特征<sup>[8]</sup>。MATLAB 中提供 im2bw 函数把图像转换为二值图像。

### 3.3 网格特征提取

文字图像被均匀地或非均匀地划分为若干区域，称之为“网格”。在每一个网格内寻找各种特征，如笔画点与背景点的比例，交叉点的个数，网格部分的笔画密度等，称为网格特征提取。这种方法对文字特征的统计以网格为单位，即使个别点的统计有误差也不会造成大的影响，增强了特征的抗干扰性。

文本针对碑刻文字相对规整，对于同一个字变化幅度不大的特点，将文字图像分为个网格，然后统计每个网格内白色像素面积，计算其占整幅图像白色像素面积的百分比，这样就形成维特征向量。

### 3.4 相似度判别准则

从图像中提取文字的特征后，需要把去除噪声的碑刻文字与在文字库里的标准字进行比较，即相似度对比识别。根据与待识别文字一样的特征提取方法，提取标准字的特征，这样就得到 2 个  $n^2$  维特征向量，识别方法采用欧式距离与相似系数相结合的方法。

欧式距离的公式如下：

$$D(X,G) = \left[ \sum_{i=1}^n (x_i - g_i)^2 \right]^{\frac{1}{2}} \quad (6)$$

其中， $x_i$ ， $g_i$  为 2 个  $n$  维特征向量，欧式距离越小，说明相似度越高。

相似系数的公式如下：

$$R(X,G) = \frac{\sum_{i=1}^n x_i \cdot g_i}{\left[ \sum_{i=1}^n x_i^2 \cdot \sum_{i=1}^n g_i^2 \right]^{\frac{1}{2}}} \quad (7)$$

其几何意义是 2 个  $n$  维特征向量  $x_i$ ， $g_i$  的夹角余弦，相似系数越大，说明相似度越高。

本文采用的相似度判别准则为：相似度

$$E = \frac{[1 - D(X,G)] + R(X,G)}{2} \quad (8)$$

### 3.5 实例模拟

综上所述，碑刻文字识别的步骤如下：(1)归一化；(2)二值化；(3)提取网格特征，形成维特征向量；(4)相似度对比识别。

考虑到碑刻文字尚属于手写体范畴，根据多次实验的结果对比，本文中实例模拟取  $n=5$ 。

在实例模拟中，提取“永”、“水”、“司”三个字与去除噪声后的碑刻文字(图 2)比较相似的程度。



图 2 待识别的文字



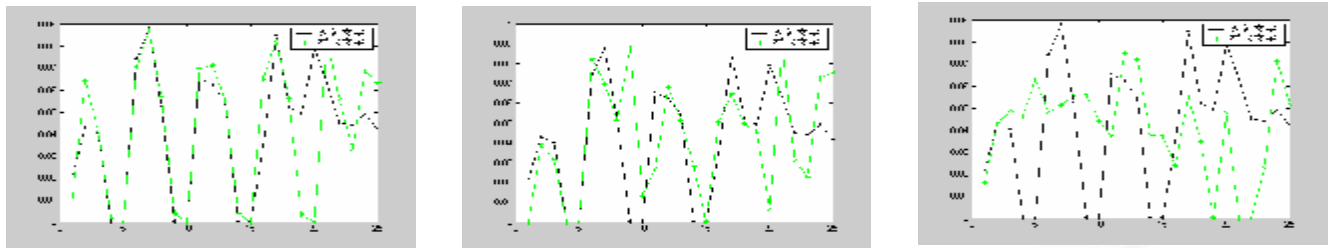
图 3 去除噪声后的碑刻文字识别过程

识别过程分析：首先，利用最大梯度差检测法将图 3(a)中的文字位置归一化，结果如图 3(b)所示。再将 3(b)尺寸归一化为  $200 \times 200$  像素，结果如图 3(c)。然后将 3(c)二值化，如图 3(d)。为了提取文字特征，利用 MATLAB 中的 `imcrop` 函数将 3(d)分为 25 等份，形成网格，如图 3(e)。

对图 2 中其他的文字做同样的处理，然后统计每个网格内白色像素面积占其整幅图像白色像素面积的百分比，各自形成 25 维特征向量。MATLAB 图像处理工具箱提供了 `bwarea` 函数来计算二进制图像的面积，面积粗略地说就是图像中前景的像素的个数。本文采用 `bwarea` 函数，是因为它并不是仅仅简单地计

算非 0 像素的数目,而是在计算面积的过程中,对不同的像素赋予不同的权值,这个加权的目的是为了补

偿用离散图像代表连续图像误差的过程。



(a) 碑刻文字与“永”字比较的特征曲线

(b) 碑刻文字与“水”字比较的特征曲线

(c) 碑刻文字与“司”字比较的特征曲线

图 4 特征曲线的比较

从特征曲线来看,去除噪声的碑刻文字与“永”字的相似度最高。经过计算,得到表 1。

表 1 相似度对比

	欧式距离 $D$	相似系数 $R$
与“永”字	0.1072	0.9137
与“水”字	0.1421	0.8428
与“司”字	0.1704	0.7534

由此得到与“永”字的相似度  $E_1 = 90.3\%$ ,与“水”字的相似度  $E_2 = 85.0\%$ ,与“司”字的相似度  $E_3 = 79.2\%$ 。

#### 4 结语

本文提出了一种识别碑刻文字的新方法,利用细胞神经网络技术结合二维中值滤波去除碑刻文字的噪声,然后使用网格特征提取识别文字。本文在 MATLAB 平台上进行了实例模拟,取得了较好的效果。本文研究的后续工作是对整幅碑刻的识别及不同字体的区分。

#### 参考文献

- 1 何如月. 碑刻文献在历史研究中的价值. 考古与文物, 2005,(4):89—89.
- 2 何骥鸣,李强,明艳. 改进型中值滤波器的实现和性能测试. 计算机系统应用, 2009,18(8):172—174.
- 3 Chua LO, Yang L. Cellular neural networks: Theory. IEEE Trans, Circuits Syst, 1988,35:1257—1272,1273—1290.
- 4 Chua LO, Roska T. The paradigm. IEEE Trans Circuits Syst, 1993,40(3):147—156.
- 5 Dogaru R, Chua LO. Universal CNN cells. Int. J. Bifur. Chaos, 1999,9(1):1—48.
- 6 余松煜,周源华,张瑞. 数字图像处理. 上海:上海交通大学出版社, 2007.275—275.
- 7 胡小锋,赵辉. Visual C++/MATLAB 图像处理与识别实用案例精选. 北京:人民邮电出版社, 2004.231—232.
- 8 高成. Matlab 图像处理与应用. 北京:国防工业出版社, 2007.198—198.