Chinese Character Thinning System Based on LabVIEW and BP Neural Network

Zhi WANG\(^1\), Zheng-lin LI\(^{1,*}\), Lei GUAN\(^1\), Dian-fang HUANG\(^1\) and Yu-wei ZHANG\(^1\)

\(^1\)Guangxi University of Science and Technology, Liuzhou

*Corresponding author

Keywords: Thinning, LabVIEW, BP neural network, Combinatorial templates.

Abstract. In order to effectively realize the thinning of Chinese character, a Chinese character thinning system based on LabVIEW platform and BP neural network optimized parallel thinning algorithm is proposed. The BP neural network is trained by the optimized combinatorial templates to obtain the optimized training parameters. Then the optimized parameters are input into the thinning system of Chinese characters for quick thinning. Experiments show that the system not only has the advantages of parallel algorithm based on combinatorial templates, but also simplifies the operation, and its processing speed is about 20% higher than that of combinatorial templates, which is an ideal thinning system. It is widely used in fingerprints, circuit boards and medical image recognition.

Introduction

In Chinese character recognition, feature extraction is a more important part of the identification process, and thinning is the key technology of feature extraction. In the process of thinning, some thinning principles must be followed to ensure the connectivity and topology of the original graphics. Both [1] and [2] give some thinning principle, which can be summed up in two points:

1. Try to ensure that the quality of thinning, including connectivity, topology and burr less;
2. Under the premise of ensuring (1), try to ensure that the processing speed is faster.

At present, the image thinning gradually began to be divided into two parts: one is through the logical relationship to thinning the binary image, and the second is through the neural network model to thinning the binary image. The first is to use the target point of the 8-neighborhood or 16-neighborhood pixels to achieve iterative loop judgment, making the image completely thinning, such as Zhang-Suen thinning algorithm [2], OPTA thinning algorithm [3], and so on. The second is to optimize the input samples, expected output, activation functions, weights, learning steps and other parameters, and build an accurate neural network model, such as BP neural network [5], pulse coupled neural network [6], multiple convolution neural network [7], depth learning [8], and so on. A new thinning algorithm is proposed by combining the effect of combination optimization template and the nonlinear mapping ability of BP neural network, the thinning algorithm system is implemented by using LabVIEW platform [9,10,11], and the fast thinning effect of Chinese characters can be obtained by simple operation.

Comparison of Thinning Algorithms for Existing Templates

The Chinese character image is preprocessed to obtain the binary image, the current pixel is P, and the 16-neighborhood template is shown in Fig. 1. The Zhang-Suen algorithm is a classic 8-neighborhood parallel algorithm, which is divided into two sub-processes. The first sub-process is to remove the useless points by logical judgment. The second sub-process is to traverse all the pixels. Among them, the specific logic judgment condition of two sub-processes is shown in [2]. The Chinese character image shown in Fig. 5(a) is thinning by Zhang-Suen algorithm, the thinning effect and the partial magnification are shown in Fig. 5(b), 5(f).
The OPTA thinning algorithm is also a parallel algorithm, slightly different is that it has two sets of templates, respectively, 8 deletion templates and 2 reserved templates, see [3]. Image thinning is achieved by comparing the logical relationship between the deleted template and the reserved template. The algorithm steps are as follows:

1. The pixels and delete template matching, if the pixel meets the delete template, enter step (2), otherwise go to step (3);
2. It is matched with the reserved template, and if the template is satisfied, the pixel is retained, go to step (1), otherwise, delete the pixel, proceed to step (1);
3. Repeat steps (1), (2) until all the pixels are traversed, and after traversing, go to step (4);
4. To determine whether there is a pixel in the previous thinning is deleted, if so, then go to step (1), otherwise, the end of the thinning.

The algorithm can maintain the original structure of the topology, but there are many burrs, not smooth and other shortcomings. The OPAT thinning algorithm is used to thinning the Chinese character images shown in Fig. 5(a), and the thinning effect and the local magnification are shown in Fig. 5(c), 5(g).

![Figure 1. 16-neighborhood template.](image1)

![Figure 2. Delete template in [1].](image2)

![Figure 3. Combination template in [1].](image3)

In order to solve the shortcomings of the above algorithm, [1] proposed a set of improved combination optimization template, including the delete template and the combination template, as shown in Fig. 2, Fig. 3. The algorithm steps, see [1], are in agreement with the OPTA algorithm, except that the reserved template is optimized to obtain the combination template as shown in Fig. 3. Similarly, the Chinese character image through the MEI algorithm [1] for thinning, the resulting thinning effect and local magnification are shown in Fig. 5(d), 5(h).

From Fig. 5(f)-4(h), it can be seen that the Zhang-Suen algorithm and the OPTA algorithm have the disadvantages that the pixel width is not 1, breakpoint, not smooth enough, and the thinning effect of Chinese characters by MEI algorithm is the best, it can effectively maintain the image connectivity and topology.
Proposed a New Thinning Algorithm and Experimental Simulation

New Thinning Algorithm Based on BP Neural Network

A new thinning algorithm is proposed based on the better thinning quality of MEI algorithm, which is called M-BP algorithm (MEI-Back–Propagation, M-BP). The algorithm not only has the advantages of the above algorithm, but also faster than the thinning algorithm in [1]. For BP neural networks, the preprocessing of sample data is important for the establishment of network model and the accuracy of pattern recognition [11].

![M-BP neural network model and training process.](image)

After a large number of experiments, the use of four layers of BP neural network to achieve better results, respectively, the output layer, the first hidden layer, the second hidden layer, the output layer. If the target point and the P16 point are not taken into account, the input node of the BP neural network has 14, and there are 16384 input samples. According to the empirical formula, the number of nodes in the first hidden layer is 8, and the number of the second hidden layer nodes is 3, the expected output is the result of whether the point is deleted, so the output node is 1. The BP neural network model is shown in Fig. 4(a).

![The original Chinese characters, and thinning results obtained by (b) Zhang-Suen algorithm, (c) OPTA, (d) MEI algorithm, (e) M-BP algorithm. (f) Partial magnification of (b). (g) Partial magnification of (c). (h) Partial magnification of (d). (i) Partial magnification of (e).](image)

The output formula of neurons in BP neural network is:

\[
y = f\left(\sum_{i=1}^{n} w_{ij} x_{j} - \theta\right)
\]  
(1)

Where \( \theta \) is the neuron threshold, \( f(\cdot) \) is the activation function, which mainly includes Sigmoid function, Tanh function, linear function, where the hidden layer uses Tanh function:

\[
f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}
\]  
(2)

The global error function is:

\[
E = \frac{1}{2} \sum_{p=1}^{P} \sum_{j=1}^{m_j} (t_{pj} - y_{pj})^2
\]  
(3)

The output layer weight formula is:
\[ \Delta w_{jk} = \eta \sum_{p=1}^{P} \frac{\partial E_p}{\partial \text{net}_{pk}} \frac{\partial \text{net}_{pk}}{\partial w_{jk}} = \eta \sum_{p=1}^{P} (t_{pj} - y_{pj})y_{pj}(1-y_{pj})o_{pk} \]  

(4)

The hidden layer weights formula is:

\[ \Delta v_{ki} = \eta \sum_{p=1}^{P} \frac{\partial E_p}{\partial \text{net}_{pk}} \frac{\partial \text{net}_{pk}}{\partial v_{ki}} = \eta \sum_{p=1}^{P} \left( \sum_{j=1}^{N} \delta_{pj} w_{jk} \right) o_{pk} (1-o_{pk})X_{pi} \]  

(5)

Generalized error formula:

\[ \delta_{pj} = (t_{pj} - y_{pj})y_{pj}(1-y_{pj}) \]  

(6)

Where \( t_{pj} \) represents the expected output, \( o_{pk} \) represents the output of the hidden layer, \( y_{pj} \) represents the output of the output layer, \( \eta \) represents the learning step, and setting it to 0.1.

The Chinese character image shown in Figure 5. (a) is skeletonized by the M-BP network model constructed in Figure 4. (a), and the number of training steps is 29 steps, as shown in Figure 4. (b), the thinning effect as shown in Figure 5. (e), the partial magnification of Chinese characters by thinning is shown in Figure 5. (i).

Table 1. shows the comparison of the four thinning algorithms in terms of execution time, implementation complexity, and thinning effect. The execution time is measured based on the different thinning algorithm for the original image of Figure 5. (a).

<table>
<thead>
<tr>
<th>Algorithms</th>
<th>Execution Time (ms)</th>
<th>Implementation complexity</th>
<th>Thinning Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhang-Suen</td>
<td>33.7</td>
<td>Complex</td>
<td>Medium</td>
</tr>
<tr>
<td>OPTA</td>
<td>23.6</td>
<td>Medium complex</td>
<td>Medium</td>
</tr>
<tr>
<td>MEI Yuan</td>
<td>21.3</td>
<td>Medium complex</td>
<td>good</td>
</tr>
<tr>
<td>M-BP</td>
<td>17.5</td>
<td>Simple</td>
<td>good</td>
</tr>
</tbody>
</table>

M-BP algorithm is equivalent to the MEI algorithm in terms of thinning quality, but it is about 20% higher than the MEI algorithms at the thinning speed. It should also be noted that the model selection of the M-BP algorithm is related not only to the 8-neighborhood template or 16-neighborhood template, but also to the number of elements in the template.
perform M-BP network training. The resulting actual output needs to be rounded up and then displayed through the VISION module to refine the image. The output data of M-BP network training needs to be rounded up and then transfer to the VISION module to show the final thinning images. The experimental results show that the thinning effect of the system is ideal.

Summary

We proposed a new thinning algorithm using BP neural network. A large number of experiments have proved that the thinning quality of the algorithm is not only ideal, but the thinning speed can increase about 20% of its original speed. Combined with LabVIEW to achieve the thinning algorithm can be simpler to achieve the thinning of Chinese characters. At the same time, the system can not only have a simple and effective thinning of Chinese characters, but also has a better effect on fingerprints, circuit boards, and medical CT maps [12]. It will be applied in the fields of character recognition, industrial process control, and machine vision.

Acknowledgement

This work was supported in part by the National Natural Science Foundation of China (NSFC) under grants No. 61464001, the Fund of Key Laboratory of Industrial Process Intelligent Control Technology of Guangxi Higher Education Institutes under grants No. IPICT-2016-08, Innovation Project of Guangxi University of Science and Technology Graduate Education under grants No. GKYC201622, GKYC201625.

References
