TELKOMNIKA, Vol.15, No.1, March 2017, pp. 306~313
ISSN: 1693-6930, accredited **A** by DIKTI, Decree No: 58/DIKTI/Kep/2013
DOI: 10.12928/TELKOMNIKA.v15i1.4063

Recognition Technology for Four Arithmetic Operations

Wang Meng[1], Bian Guangrong*[2], Zheng Kai[3], He Ninghui[4], Zhang Yanlei[5]
[1,3]School of Mechanical and Electronic Engineering, Wuhan University of Technology, Wuhan, Hubei, 430070, P.R.China
[2,5]Department of Aviation Ammunition, Air Force Service College, Xu Zhou, Jiangsu, 221000, P.R.China
[4]Power Science Research Institute of State Grid Ningxia Power Co, Yingchuan, Ningxia, 750011, P.R.China
Corresponding author, e-mail: woaipapamamajiejie@126.com, 44973901@qq.com*, 232464433@qq.com

Abstract

Numeral recognition is an important research direction in field of pattern recognition, and it has broad application prospects. Aiming at four arithmetic operations of general printed formats, this article adopts a multiple hybrid recognition method and is applied to automatically calculating. This method mainly uses BP neural network and template matching method to distinguish the numerals and operators, in order to increase the operation speed and recognition accuracy. Sample images of four arithmetic operations are extracted from printed books, and they are used for testing the performance of proposed recognition method. The experiments show that the method provides correct recognition rate of 96% and correct calculation rate of 89%.

Keywords: *four arithmetic operations, image preprocessing, numeral recognition, BP neural network*

1. Introduction

Arabic numerals, in its handwritten or printed form, are the only numeral system which is accepted virtually in the entire civilized world [1]. Four arithmetic operations are some elementary arithmetic based on the Arabic numerals. They are composed of Arabic numerals (0~9) and operator signs (plus, minus, multiplication, and division). The invention of calculator is very useful, and helps people solve many complex four arithmetic operations. Traditional calculators give outcome replying on the numeral calculating expressions input by users. In recent years, some products equipped with photo function can also work like a calculator. These products finish the calculating not relying on the users' manual input, but relying on the photographing unit that has the function of automatic recognition. Technology of numbers recognition has been mainly applied in these products. By taking photo, the calculating expression is processed and divided into several numeric characters, and then according to the basic mathematical operation rules, the result can be calculated. Therefore, the recognition accuracy of the operation expression is the core problem, and directly affects the reliability of the result.

Many researchers are working on the research of effective algorithms for Optical Character Recognition (OCR) Arabic texts. Paper [2] made recognition decision according to structure characteristics of numerals, and put forward to feature extraction method and recognition algorithm based on characteristics of closed and semi-closed hole. This method achieved high correct rate for standard and no noise numerals. Paper [3] proposed a novel method of direction feature extraction approach combined with principal component analysis (PCA) and support vector machine (SVM) for recognizing handwritten Bangla numerals, and got an average recognition rate of 95.05%. Paper [4] used improved BP neural network, and had an application on off-line handwritten recognition. The method could not only raise precision of BP artificial neural network, but also improve the convergent speed. Paper [5] also use BP Network for handwritten alpha numeral identification. The basic characteristics of artificial neural networks are nonlinear mapping, classification of study and real-time optimization, so it has opened up new avenues for pattern recognition [4]. According to actual situations, different kinds of neural networks are adopted for example BP network [6] and RBF network [7]. Most of design of numeral recognition systems is divided into two modules: image preprocessing

Received June 15, 2016; Revised August 11, 2016; Accepted October 26, 2016

module and recognition module. Image preprocessing module conducts image by a series of transformations, and then send the sample feature vector, finally extracted to digital identification system, then identify them and give the results [8].

The main research of this paper is to explore a new method for automatic calculation based on recognition technology. Because arithmetic operation is composed of Arabic numerals and operator signs, numbers recognition algorithm must be extended in order to have the recognition function of operation signs. Therefore, this paper proposed a recognition algorithm that realizes recognizing four arithmetic operations, which can be used for intelligent calculating based on recognition technology. Applying this recognition technology for four arithmetic operations is to complete two parts of recognition task, those are: numeral recognition and operator recognition. In our work, we have two phases for four arithmetic operations recognition. Firstly, according to the structural characters, we extract and classify the operator signs from the operation. Then we identify the rest numerals through numeral recognition algorism. The work of numeral recognition is mainly implemented by using BP neural network and template matching method. Training of BP neural network occurs before the trained system is implemented in numerals recognition. When the recognition of network is failed, program will enter into template matching procedures. By this, we both make use of neural network recognition algorithm to reduce computing time, and take advantage of template matching algorithm to improve the recognition rate.

The structure of this paper is as follows: section 2 presents some steps of image preprocessing, which will be used in this paper. In Section 3, details of BP neural network recognition and template matching method are presented. Their respective advantages are also analyzed and compared. Section 4 presents the process of our proposed recognition system, and provided the experiments. We discuss our results and comment on their comparisons. Section 5 concludes this paper.

2. Image Preprocessing

Image preprocessing is the transform processing of image and the preparation of image analysis. The purpose is to make the image more concise, more unique. The text, remain to be recognized, is always from scanned printed or handwritten image files. Generally, the image will be affected by the illumination, paper and so on. These factors will lead to distortion and stains. So we need to exclude the interference, and extract useful information.

2.1. Filtering

Filtering process is used to clean up the image from any redundant feature points, which mainly take the form of either a curve joining an end point with a branch or a cross point; or a small loop joining two branch points [9]. This paper adopts the median filtering method for image smoothly filtering. In this way, it not only eliminates the noise in the image, but also preserves image edge. At the same time, the signal-to-noise ratio of image is improved.

Median filtering is used similar to an averaging filter, in that each output pixel is set to an average of the pixel values in the neighborhood of the corresponding input pixel [10]. First of all, gray value of i-th pixel within the template (window) is arranged from small to large like as:

$$x_{i1} < x_{i2} < \cdots < x_{ik} < \cdots < x_{in} \tag{1}$$

Then the intermediate value can be calculated as:

$$f(i) = Med\{x_{i1}, x_{i2}, \cdots, x_{i(n-1)}, x_{in}\}. \tag{2}$$

So, $f(i)$ is chosen as the mean value, and this value is regarded as the output of the filter value. In the strong disturbing of pepper type (or pulse), median filtering can remove interference, because the mean value forces the gray value of interfering points to become close to gray value of the adjacent pixels points adjacent values.

In this paper, we used every pixel as the center to constitute the window of 3×3, and the filtering process is as shown in Figure 1.

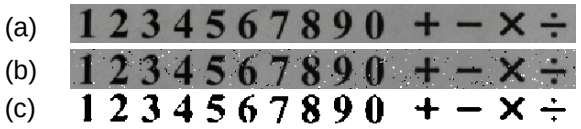

Figure 1. Filtering process, (a) Original image, (b) Adding salt and pepper noise, (c) Filtering and thresholding

2.2. Dividing

Character segmentation is an operation that seeks to decompose an image of a sequence of characters into sub images of individual symbols [11]. Dividing is to make it convenient for further processing of single characters separately. Papers [12-14] present several dividing methods according to different situations.

Considering the characters are printed form with relatively fixed way of writing, there are at least one row or column background pixels between neighboring characters. According to this characteristic, we can adopt the dividing method based on projection.

2.3. Thinning

The stroke width of original scanned characters is generally several pixels. Thinning is to retain important pixel of a character, remove irrelevant pixel, and get the stroke skeleton character. The stroke skeleton is usually a single pixel width, which is enough to describe the structure of the character. We call this image of a single pixel stroke with skeleton vividly, and the process of obtaining skeleton is called thinning [15]. This paper uses the classic Hilditch thinning algorithm.

Hilditch algorithm, different from the method of look-up table, under certain judgment, directly operates in the program. According to the operation result, the algorithm will determine whether delete this point or not. The thinning result is shown in Figure 2.

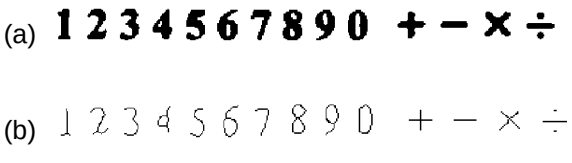

Figure 2. Thinning process
(a) Before thinning, (b) After thinning

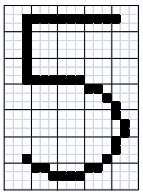

Figure 3. Coarse Grid Character

2.4. Feature Extraction

Feature extraction plays a very important role in all pattern recognition systems. In the past years, many feature extraction techniques have been proposed for handwritten number recognition [16]. The recognition rate is decided by the reliability of the extracted features.

This paper extracts coarse grid characters as character's feature. Through above preprocessing steps, we normalize each character to a 21×15 lattice. The 21×15 lattices will be separated into 35 parts, and each part contains 9 pixels as shown in Figure 3. For each character, the algorithm will scan from left to right, from top to bottom line by line, and count the statistics of each section's highlights points.

Through grid division, we can extract 35 feature values which embody the characteristics of preprocessed character. These values construct feature vector, and will be put into recognition program.

3. Basic Theories of BP Neural Network Recognition and Template Matching Method

Neural network recognition and template matching method are two kinds of recognition methods. They are frequently used. They both have their own advantages. For neural network, trained network used for recognition saves computing time, while for template matching method, it always have higher accuracy.

3.1. BP Neural Network Recognition

BP Ann usually has learning and recognizing process. There are two phases of positive transmitting processing and error reverse transmitting processing in the learn processing of BP Ann. In the recognizing process, the signal only propagates forward, and recognition result is given from the output layer.

For a practical neural network, if its structure is relatively simple, the ability of network may be not enough to apply to the actual solution. If its structure is complex, corresponding network computation quantity increases largely. At the same time, the generalization capability of the network may not get corresponding improvement.

In this paper, neural network is realized by using a 3-layer neural network. The structure is shown in Figure 4.

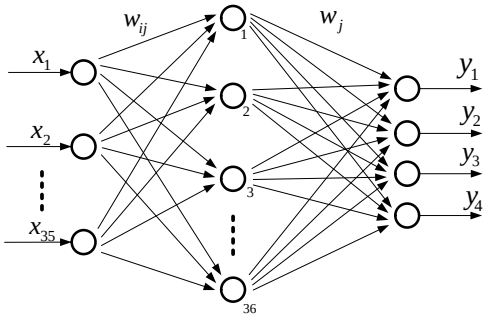

Figure 4. Neural Network Structure

The input layer has 35 neurons, and they respectively correspond to the values from feature vector $p=[x_1,x_2,\cdots x_{35}]$. The output layer has 4 neurons whose output values construct a matrix vector $t=[y_1,y_2,y_3,y_4]$. y_1, y_2, y_3 and y_4 can form a four bit binary code, which has 16 kinds of combinations. Here we use 14 kinds of them. These 14 combinations respectively represent numerals (0-9) and operator signs (+, −, ×, and ÷).

Original images scanned from books printed by different fonts are collected. Each character has 20 kinds of samples. In order to eliminate the samples of the same or similar circumstances, we choose different fonts as training samples as far as possible. The image preprocessing method mentioned in last chapter is adopted to get training samples. We extract the feature vector to represent each character, and finally normalize the vector. Table 1 shows the target vector of each character.

In the learning stage, input sample enters the network through the input layer, and then will participate in calculating process from hidden units to output units. During this process, neurons' state of each layer only influences the neurons' state of next layer. After the forward process, current output and expected output are compared. If they are not equal, the process will get into the back propagation procedure and correspondingly modify the connection weights between the neurons. After several adjustments and iterative approach, error will approaches to a minimum value.

In the recognizing stage, we extract the feature vector of character that remains to be recognized, and input the data to trained network. According to the output of network, we can judge which class the character belongs to.

Table 1. Characters and Corresponding Target Vector

character	Target vector	character	Target vector
1	[-0.95,-0.95,-0.95,0.95]	8	[0.95,-0.95,-0.95,-0.95]
2	[-0.95,-0.95,0.95,-0.95]	9	[0.95,-0.95,-0.95,0.95]
3	[-0.95,-0.95,0.95,0.95]	0	[0.95,-0.95,0.95,-0.95]
4	[-0.95,0.95,-0.95,-0.95]	+	[0.95,-0.95,0.95,0.95]
5	[-0.95,0.95,-0.95,0.95]	−	[0.95,0.95,-0.95,-0.95]
6	[-0.95,0.95,0.95,-0.95]	×	[0.95,0.95,-0.95,0.95]
7	[-0.95,0.95,0.95,0.95]	÷	[0.95,0.95,0.95,-0.95]

3.2. Template Matching Method

Among various pattern recognitions, template matching classification is one kind of classical methods. Its mathematical model is easy to set up. This method treats all samples in the training set as the standard samples. The test sample to be recognized will be compared with the standard samples one by one. According to the classification discriminate function, the category of most similar and closest standard sample is chosen as category of the test sample.

In the statistical approach, a pattern is represented by a set of N-dimensional feature vector and the decision making process is determined by a similarity measure such as a distance metric or a discriminate function [17]. In this paper, we will use shortest distance between the feature vectors as the discriminate function.

For one test sample $X=[x_1,x_2,\cdots x_N]$ and standard sample $X_i=[x_{i1},x_{i2},\cdots x_{iN}]$, where $1<i<14$, the discriminate function can be express as:

$$D_i=(\sum_{k=1}^{N}(x_k-x_{ik})^2)^{1/2} \quad (3)$$

By comparison of all the distance values, the category corresponding to the minus value is the final decision.

The template matching method is simple to realize and have higher accuracy than neural network. However, it requires strong storage capacity and costs large amount of computing time. This is because the stored templates may be a lot, so the storage amount is quite large. Moreover, each test sample is required to calculate a distance value with every standard template, so large amount of calculation should be done.

For BP neural network recognition, after finishing training the network, the recognition result can be obtained according to the output layer directly. The information of training samples has been recorded in the connection weights between neurons. This kind of network, used for recognition, costs less computing time.

Considering the advantages and disadvantages of the two recognition methods respectively, we combine the two methods to achieve better recognition performance and higher efficiency. Firstly, BP neural network is used to recognize the test four arithmetic operations. The recognition results of characters can directly made by the trained network, if the output of network meets the following formula:

$$\sum_{i=1}^{4}\left|\left|y_i\right|-0.95\right|<\delta \quad (4)$$

Where δ is a set small threshold. Otherwise, the characters are defined as unrecognized or error identification character. Then the unrecognized characters will be reidentified by the template matching recognition program.

4. Application to the Four Arithmetic Operations

Our work aims at applying the research to automatically calculating of four arithmetic operations. To take into account the requirements of recognition reliability and the response time, we employ a multiple hybrid recognition method.

4.1. Recognition Process

After preprocessing, each character in arithmetic expression is normalized to a unified form, which is convenient for the feature extraction and the following process. Then they will be treated by a set of processes. The recognition process is shown in Figure 5.

The first step is to distinguish the operator signs from the arithmetic expression. The structure characteristics of operator signs are obviously different from numerals, so we can extract the operator signs first. Moreover, the characteristics of each operator sign vary so greatly that we can use simple feature matching method to recognize them.

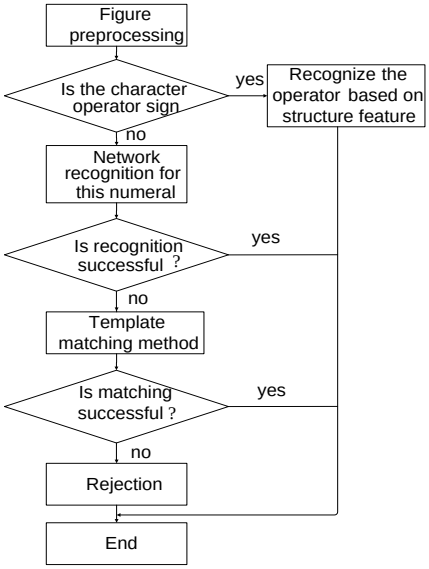

Figure 5. Recognition Process of Proposed Method

The second step is to use trained BP neural network for the recognition of numerals. Feature vector of a numeral is fed into network. The output vector of network is $[y_1, y_2, y_3, y_4]$, where $y_i = \pm 0.95$, $i = 1,2,3,4$. Through experiments, we found that once $|y_i|$ is not approximately equal to 0.95, network will often give wrong recognition result. Furthermore, the lager the error signal $\big||y_i| - 0.95\big|$ is, the higher the incorrect rate is. In order to decrease the incorrect rate, we consider the numeral whose output vector does not meet the inequality $\sum_{i=1}^{4} \big||y_i| - 0.95\big| < \delta$ as unrecognized character. At the same time, we think network recognition algorithm is not successful.

In the third step, template matching recognition program will reidentify the unrecognized character. Each standard template is compared with the sample to be recognized, and the most similar is chosen as the result of recognition. This method needs large storage capacity, and costs large amount of calculating time, so only when network recognition method is failed, it will be used.

4.2. Experiment Results and Comparative Analysis

In order to evaluate the performance of the proposed method, we carry out four groups of experiments. In the first group, we use conventional BP neural network recognition method that was mentioned in papers [5-6]. Test character image is fed into recognition system. The system makes the decision according to the output vector of network. The recognition result is show in Table 2. Then we conduct the experiment using the method of combining template matching [17] and neural network, which is one of steps in our proposed method. If the result of

network is not reasonable, process will get into template matching program. The recognition result is show in table 3. The third group is to use the proposed method. First step is to exact the operator characters from arithmetic expression. Then we use the method of combining neural network and template matching method to identify 10 numerals. The recognition result is show in Table 4. These three groups of experiment are aiming at identifying a single character. Each group had 100 tests.

Table 2. Result of traditional BP neural network method

Character	Correct	Error	Rejection	Character	Correct	Error	Rejection
1	76	24	0	8	63	37	0
2	66	34	0	9	82	18	0
3	71	29	0	0	60	40	0
4	55	45	0	+	41	59	0
5	59	41	0	−	35	65	0
6	89	11	0	×	82	18	0
7	75	25	0	÷	70	30	0

According to Table 2, by using traditional neural network method, correct recognition rate of characters is lower than 70% on average. Comparing tables 3-4 with Table 2, we can find that using the method of multiple hybrid recognition obviously improve accuracy rate of recognition. Comparing Table 4 with Table 3, we can see the method of separately identifying operator characters and numerals improve the recognition rate for operation characters. The proposed recognition method has the recognition rate of 96% on average for single character.

Considering the proposed recognition method has better performance, we conduct the fourth group experiment that is applied to recognize four arithmetic operations and automatically calculate it. We first take prepared identification program to identify the numerals and operators from the photos, and then according to the rules of four arithmetic operations, the program automatically calculates the result. We extract the test figures from book of different paper conditions. In each figure, the number of characters is between 3 and 5 randomly. One hundred tests are conducted for each condition. The result is shown in table 5. It shows that much of the recognition accuracy has to do with the paper condition. But for common books, it has correct rate of 89%.

Table 5. Test on different paper

Paper condition	Correct	Error	Rejection
clean books	93	5	2
regular books	89	7	4
old books	83	10	7

Comparing to a single character's correct recognition rate of 96%, the recognition system applied to four arithmetic operations automatically calculating performs relatively poor. That is because the prerequisite of correctly calculating is to correctly recognize all the characters in four arithmetic operations. So the longer the operation expression is, the lower the correct rate is.

5. Conclusion

In this paper, we develop a multiple hybrid recognition method, and apply it to four arithmetic operations automatically calculating. We first introduce the image preprocessing steps, and discuss two kinds of classical recognition method. Then we combine these two methods in the system in order to improve recognition performance. In the proposed recognition system, operator signs and numerals are identified by using different recognition methods. The recognition result achieved by the proposed hybrid method is more reliable than that by traditional neural network method alone. Experiments also demonstrate that the proposed recognition method still keeps high accuracy, when it is applied to four arithmetic operations calculating.

It is obvious that different paper condition will affect the recognition accuracy. In our future work, we will adopt some more effective image preprocessing approaches in order to adapt to different conditions. We will try to employ the experience of other advanced recognition algorithm to optimize our proposed recognition method.

References

[1] L Yang, HL Sun, W Liang, HJ Yuan, JX Li. *Numeral Recognition Based on Projective Invariants*. The 6th International Forum on Strategic Technology. 2011: 1078-1081.

[2] ZJ Zhang, GH Chen, JW Li, YL Ma, N Ju. *The Research on Digit Recognition Algorithm for Automatic Meter Reading System*. Proceedings of the 8th World Congress on Intelligent Control and Automation. 2010: 5399-5403.

[3] Y Wen, Y Lu, PF Shi. Handwritten Bangla Numeral Recognition System and Its Application to Postal Automation. *Pattern Recognition*. 2007; 40(1): 99-107.

[4] CD Zhang, XN Huang. *Off-line Handwritten Digit Recognition Based on Improved BP Artificial Neural Network*. IEEE International Conference on Service Operations and Logistics, and Informatics. 2008; 1: 626-629.

[5] G Sarker, M Besra, S Dhua. *A Malsburg Learning Back Propagation Combination for Handwritten Alpha Numeral Recognition*. Computer Engineering and Applications, IEEE. 2015: 493-498.

[6] Z Liu, L Li, M Yu. *An Algorithm of Handwritten Digital Recognition Based on BP-Bagging*. International Conference on Information Technology and Management Innovation. 2015: 1164-1168.

[7] D Mellouli, TM Hamdani, AM Alimi. *Deep Neural Network with RBF and Sparse Auto-encoders for Numeral Recognition*. International Conference on Intelligent Systems Design and Applications, IEEE. 2015: 468-472.

[8] B Lu, YL Wang. *Overview of Handwritten Numeral Recognition Based on BP Neural Network*. International Conference on Computer Science and Network Technology. 2011: 1502-1505.

[9] A Elnagar, F Al-Kharousi, S Harous. *Recognition of Handwritten Hindi Numerals using Structural Descriptors*. IEEE International Conference on Computational Cybernetics and Simulation. 1997; 2: 983-988.

[10] V Radha, M Krishnaveni. *Threshold Based Segmentation Using Median Filter for Sign Language Recognition System*. World Congress on Nature & Biologically Inspired Computing. 2009: 1394-1399.

[11] RG Casey, E Lecolinet. *A survey of Methods and Strategies in Character Segmentation*. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1996; 18(7): 690-706.

[12] A Rahman, B Verma. Effect of Ensemble Classifier Composition on Offline Cursive Character Recognition. *Information Processing and Management*. 2013; 49: 852-864.

[13] GH Zhang, ZY Xiong, K Li, CY Xing, SY Xia. A Novel Image Segmentation Algorithm Based on Graph Cut Optimization Problem. *TELKOMNIKA Telecommunication, Computing, Electronics and Control*. 2015; 13(4): 1337-1342.

[14] J Zhang, CJ Xie, LT Song, R Li, HB Chen. Robust Image Segmentation Using LBP Embedded Region Merging. *TELKOMNIKA Telecommunication, Computing, Electronics and Control*. 2016; 14(1): 368-377.

[15] YR Du, B Xie, LP Wang. Thinning Algorithm of Binary Image Based on Improved Templates. *Applied Mechanics and Materials*. 2014; 543(54): 2547-2550.

[16] WL Shen, CW Hsien. *Effective Multiple-features Extraction for Off-line SVM-Based Handwritten Numeral Recognition*. International Conference on Information Security and Intelligence Control. 2012; 194-197.

[17] L Sukhan, JC Pan. *Unconstrained Handwritten Numeral Recognition Based on Radial Basis Competitive and Cooperative Networks with Spatio-Temporal Feature Representation*. IEEE Transactions on Neural Networks. 1996; 7(2): 455-474.