# Handwritten Character Recognition Using CNN Gabor-Type Filters

Ertugrul Saatci* and Vedat Tavsanoglu*

***Abstract* — This paper proposes an approach for handwritten character recognition using nonlinear normalisation, a CNN Gabor-Type filter, a Location Based Dominant Orientation Map and cross correlation. Based on a test set of 26 test characters acting as template and a set consisting of 4 sets of 26 unknown handwritten test characters, max. 92 % correct recognition is provided. Recognition rate is studied for different values of the filter parameters. The results presented correspond to optimal parameter values.**

## 1 Introduction

Gabor-Type Filters (GTF) have recently been introduced by Shi [1] by allowing modulating functions other than the Gaussian in Gabor filters. Shi has also shown that such filters can be implemented by CNNs and called them CNN Gabor-Type Filters.

In [2] we have presented a Feature Extraction System that made use of a filter bank of CNN-GTFs to generate an Orientation Map (OM), which is defined as the plot of the Total Orientation in the image versus the Orientation Angle.

One of the drawbacks of using this system was that the Location Information (LI) disappeared in the process. In fact, LI is essential in the distinction of characters such as E and F; M and W; V and W; b, d, p and q etc. In order to keep LI we propose to modify the previous system such that the system stores LI along with the orientation angle. This will lead to a modified OM which will henceforth be called Location Based Dominant Orientation Map (LBDOM).

Our ultimate goal in this paper will be that of producing a handwritten character recognition system. In this system the feature extraction stage described above will be followed by a decision stage which consists of a correlator that correlates the LBDOM of an unknown character with that of a template. As Gabor-Type filters can be implemented using CNN VLSI chips [3], this system is expected to be developed into a fast character recognition system.

Variations in handwritten characters degrade performance of a character recognition system. In order to reduce the effects of such variations, normalisation of characters is carried out prior to feature extraction. Normalisation is a crucial preprocessing tool in the development of a robust handwritten character recognition system and it is well known that [4] it helps to increase the recognition rate.

In the literature there are two types of normalisation techniques: linear and nonlinear normalisation. Conventionally, linear normalisation methods which convert the image into one having constant position, size, rotation and inclination has been used. However, these did not prove to be adequate for handwritten characters because of their irregularities and partial distortion in shape. As a remedy, nonlinear normalisation methods [5] have been proposed which are subsequently verified [6] to be more effective than linear methods. Therefore as the preprocessing stage in our recognition system, we use a Nonlinear Normalisation Algorithm [5].

## 2 Gabor Filters

A 2-D Gabor filter is described by the impulse response:

$$h(x,y) = g(x,y)e^{j(\omega_{xo}x+\omega_{yo}y)}$$

where $g(x,y)$ is the Gaussian function given by:

$$g(x,y) = \frac{1}{2\pi\sigma^2}\cdot e^{\frac{-(x^2+y^2)}{2\sigma^2}} \quad ,$$

$(w_{xo}, w_{yo})$ is the spatial frequency and $\sigma^2$ is the standard deviation of the Gaussian. The output $v(x,y)$ of the filter $h(x,y)$ to an image $u(x,y)$ is obtained through the convolution sum:

$$v(x,y,\omega_{xo},\omega_{yo}) = \frac{1}{2\pi\sigma^2}\sum_{x_1,y_1} u(x_1,y_1)$$
$$\cdot e^{\frac{-(x-x_1)^2-(y-y_1)^2}{2\sigma^2}} e^{j(\omega_{xo}(x-x_1)+\omega_{yo}(y-y_1))}. \qquad (1)$$

### 2.1 CNN Gabor filters

For the filtering of an $M \times N$ pixel 2-D image, $u(m,n) \in \mathbf{R}$ where $m \in \{0,1,...,M-1\}$ and $n \in \{0,1,...,N-1\}$, we use a 2-D CNN array of

*South Bank University, School of Electrical, Electronic and Information Engineering, Borough Road, London SE1 OAA, UK. E-mail:saatcie@sbu.ac.uk, Tel: +44 20 7815 7526, Fax: +44 20 7815 7599.

$M \times N$ cells where the state at the $(m,n)th$ cell $\dot{v}(m,n) \in \mathcal{C}$ satisfies [1]:

$$\dot{v}(m,n) = \sum_{k,l=-\rho}^{\rho} a_{k,l} v(m+k, n+l) + bu(m,n) \quad (2)$$

where dot denotes differentiation with respect to time. The $A = [a_{k,l}]_{k,l=-\rho}^{\rho}$ and $b$ are complex coefficients called the feedback and feedforward cloning templates and $\rho$ is defined to be the connection radius.

In the case of 2-D band-pass CNN Gabor filter tuned to the centre frequency $(w_{xo}, w_{yo})$, feedback cloning template $A$ is given by Shi [1] as:

$$A = \begin{bmatrix} 0 & e^{-jw_{y0}} & 0 \\ e^{jw_{x0}} & -(4+\lambda^2) & e^{-jw_{x0}} \\ 0 & e^{jw_{y0}} & 0 \end{bmatrix} \quad (3)$$

In order to obtain the frequency response of the cells we use (2). If the filter is stable, it does not oscillate and eventually settles to a stable equilibrium point for which (2) takes the form [7]:

$$\sum_{k,l=-1}^{1} a_{k,l} v(m+k, n+l) + bu(m,n) = 0$$

Under this condition, using the feedback template in (3) and choosing the feedforward cloning template as $b = \lambda^2$, we obtain

$$v(m,n) = \frac{1}{4+\lambda^2}[e^{jw_{x0}} v(m-1,n) + e^{-jw_{y0}} v(m,n+1) + e^{-jw_{x0}} v(m+1,n) + e^{jw_{y0}} v(m,n-1) + bu(m,n)] \quad (4)$$

The frequency response of the CNN Gabor-Type filter is obtained from (4) as:

$$H(e^{jw_x}, e^{jw_y}) = \frac{V(e^{jw_x}, e^{jw_y})}{U(e^{jw_x}, e^{jw_y})} = \frac{\lambda^2}{4+\lambda^2 - 2cos(w_x - w_{x0}) - 2cos(w_y - w_{y0})} \quad (5)$$

## 3 Location Based Dominant Orientation Map

Gabor filters are orientation selective and respond maximally to edges which are oriented at an angle $\theta = atan(w_{y0}/w_{x0})$ where $\theta$ is defined to be the angle between the horizontal axis and the line perpendicular to the edge. In order to detect the angle $\theta$ of a particular orientation in an image, we use a filter bank of $n_p$ Gabor filters whose spatial frequencies are:

$$\{(w_{x0}^k = rcos\theta_k, w_{y0}^k = rsin\theta_k) \mid \theta_k = \frac{k\pi}{n_p}, k = 0, \cdots, (n_p - 1)\} \quad (6)$$

where $r$ is the radius of spatial frequency. The angle $\theta_k$ associated with the filter of the maximum output is taken as the orientation of the particular edge in the image.

We intend to exploit the orientation selectivity property of CNN Gabor-Type filters to construct a feature extraction tool to be used in the recognition of handwritten characters. In [2], an OM was introduced and a system was developed that generated the OM. The output of this system gave the OM as a plot of the Total Orientation in the image versus the Orientation Angle. The disadvantage of using this system was that its output did not contain any LI. In actual fact, LI is essential in the distinction of characters such as E and F; M and W; V and W; b, d, p and q etc. In order to be able to make use of LI, we have developed a system that detects the orientation with respect to the location in the image at all possible angles. This system is illustrated in Fig.1. and can be described as follows: $n_p$ matrices of dimensions $64 \times 64$ are allocated to the $n_p$ outputs of the filter bank consisting of $n_p$ filters. The orientation angle values for each pixel in $n_p$ matrices are then compared and the matrix entry which has the maximum value, which is the dominant orientation for that pixel, is replaced by a "1". The corresponding entry location in all other $n_p - 1$ matrices are set to zero. Once this operation has been completed we obtain $n_p$ new matrices which constitute what will be henceforth called the "Location Based Dominant Orientation Map (LBDOM)".

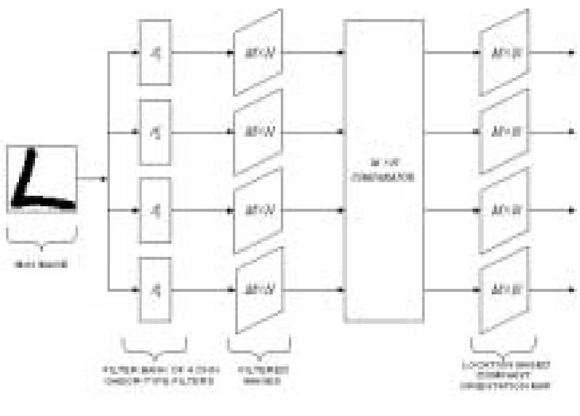

Figure 1: LBDOM.

## 4 Recognition of handwritten characters

For the reasons explained in Section 1 a nonlinear normalisation technique, namely line-density equalisation, is applied before inputting the image to the filter bank of Gabor-Type filters.

Once the unknown character has gone through the nonlinear normalisation stage, it is than fed to a filter bank of CNN-GTF. Filtered images are compared in a comparator to construct the LBDOM as explained in the previous section. At the final stage the LBDOM of the unknown handwritten character is compared with the LBDOMs of all characters in the template. The comparison is carried out by computing the cross correlation of each of the $n_p$ matrices in the LBDOM of the unknown character with the corresponding matrix of the LBDOM of each character in the template set. A single value is obtained by adding $n_p$ cross correlation values. The decision is made by assuming that the unknown character is the same as that from the template set which yields the maximum value for the sum of all $n_p$ cross correlations. The recognition rate of the system is defined as the percentage of the number of correctly recognised characters in a set.

## 5 Selection of $r$ and $\lambda$

Considering (5) and (6) reveals that $r$ is the radius of spatial frequency that controls the location of CNN-GTF centre frequency $(w_{x0}, w_{y0})$ in the spatial frequency domain. On the other hand, the parameter $\lambda$ determines the spread of the CNN-GTF frequency response. Large selection of $\lambda$ makes the filter wide-band in the spatial frequency domain which yields better results. Hence the appropriate choice for the parameters $r$ and $\lambda$ is crucial in CNN Gabor-Type filtering. The values for these parameters should be chosen such that most of the energy is captured by the filter. Only in this case steering the filter by changing $\theta$ results in significant variations of the filter output. For handwritten characters it is shown in [2] that most of the energy is localised at lower frequencies. Therefore values of $r$ should be chosen small enough to capture most of the energy on the frequency plane. When the spectrum of the CNN-GTF matches most of the frequency spectrum of the character maximum response from the filter will be obtained.

## 6 Example

In this paper we use a filter bank of $n_p = 4$ filters to detect the dominant orientation. 5 different sets of handwritten characters are used. One of these sets was taken as the template set. The optimal parameter values are found to be $r = 1$ and $\lambda = 3$. Table 1. shows the recognition results obtained. Figure 2. shows the template set and the four sets of unknown characters before and after nonlinear normalisation.

## 7 Conclusions

In this study handwritten character recognition is carried out using Gabor-Type filters implemented by CNNs. A newly developed Location Based Dominant Orientation Map is used which converts the filter output to a suitable form of extracted features. Alongside the advantages, there are also drawbacks of using the using LBDOM. The fact that LBDOM keeps local features may also lead to incorrect decisions if nonlinear normalisation is not used. This problem should also be given a thorough investigation. Although filtering is studied using different parameter values, the results presented in this paper correspond to optimal parameter values. Further research into filtering with different parameter values may prove useful. Future research will also concentrate on using filter banks with various numbers of filters and a system with a neural network acting as the decision making tool.

| set 1 | result | set 2 | result | set 3 | result | set 4 | result |
|---|---|---|---|---|---|---|---|
| A-01 | A | A-02 | A | A-03 | A | A-04 | A |
| B-01 | B | B-02 | H | B-03 | B | B-04 | B |
| C-01 | C | C-02 | C | C-03 | E | C-04 | C |
| D-01 | D | D-02 | B | D-03 | D | D-04 | D |
| E-01 | E | E-02 | E | E-03 | E | E-04 | E |
| F-01 | F | F-02 | F | F-03 | F | F-04 | F |
| G-01 | G | G-02 | G | G-03 | G | G-04 | G |
| H-01 | H | H-02 | H | H-03 | H | H-04 | H |
| I-01 | I | I-02 | I | I-03 | I | I-04 | I |
| J-01 | J | J-02 | J | J-03 | J | J-04 | J |
| K-01 | K | K-02 | K | K-03 | K | K-04 | H |
| L-01 | E | L-02 | L | L-03 | L | L-04 | L |
| M-01 | M | M-02 | M | M-03 | M | M-04 | M |
| N-01 | N | N-02 | N | N-03 | N | N-04 | N |
| O-01 | O | O-02 | D | O-03 | Q | O-04 | O |
| P-01 | P | P-02 | P | P-03 | P | P-04 | P |
| Q-01 | Q | Q-02 | D | Q-03 | Q | Q-04 | G |
| R-01 | R | R-02 | R | R-03 | R | R-04 | P |
| S-01 | S | S-02 | S | S-03 | S | S-04 | S |
| T-01 | T | T-02 | T | T-03 | T | T-04 | T |
| U-01 | U | U-02 | U | U-03 | U | U-04 | U |
| V-01 | X | V-02 | Y | V-03 | V | V-04 | V |
| W-01 | W | W-02 | U | W-03 | U | W-04 | U |
| X-01 | X | X-02 | X | X-03 | X | X-04 | X |
| Y-01 | Y | Y-02 | P | Y-03 | Y | Y-04 | Y |
| Z-01 | Z | Z-02 | I | Z-03 | Z | Z-04 | J |
| correct | percent | correct | percent | correct | percent | correct | percent |
| 24 | 92.3% | 18 | 69.2% | 23 | 88.4% | 21 | 80.7% |

Table 1: Result of the recognition for $r = 1$, $\lambda = 3$ and number of orientations are 4.

Figure 2: Template and four character set before (left) and after (right) nonlinear normalisation.

## References


[1] B.E. Shi "Gabor-Type Filtering in Space and Time with Cellular Neural Networs", IEEE Trans. on CAS-I,vol45, 2, pp.121-132, February 1998.

[2] V. Tavsanoglu, E. Saatci, "Feture Extraction for Character Recognition using Gabor-Type Filters Implmented by Cellular Neural Networks", Proceedings of the 6th IEEE International Workshop on Cellular Neural Networks and Their Applications 2000 (CNNA 2000), pp.63-68, 2000.

[3] B. E Shi, "2D Focal Plane Steerable and Scalable Cotical Filters", Proceedings of the Seventh International Conference on Microelectronics for Neural Fuzzy and Bio-Inspired Systems, MicroNeuro'99, pp.232-239, 1999.

[4] J. Tsukumo and H. Tanaka, "Classification of handprinted Chinese characters using nonlinear normalization and correlation methods", Proc. of 9th Int. Conf. Pattern on Recognition, pp.168-171, 1988.

[5] H. Yamada, K. Yamamoto and T. Saito, "A nonlinear normalization method for handprinted Kanji character recognition -Line Density Equalization-", Proc. of 9th Int. Conf. on Pattern Recognition, pp.172-175, 1988.

[6] S.W. Lee, J.S. Park, Y.Y. Tang, "Performance evaluation of nonlinear shape normalization methods for the recognition of large-set handwritten characters", Proceedings of the Second International Conference on Document Analysis and Recognition, pp. 402-407, 1993.

[7] B.E. Shi and L.O. Chua, "Resistive grid image filtering: input/output analysis via the CNN framework", IEEE Trans. Circuits Syst. I, Vol.39, pp.531-548, July 1992.