CNN CLONING TEMPLATE: CONNECTED COMPONENT DETECTOR

by

T. Matsumoto, L. O. Chua, and H. Suzuki

Memorandum No. UCB/ERL M89/65

25 May 1989

COVER PAGE

CNN CLONING TEMPLATE: CONNECTED COMPONENT DETECTOR

by

T. Matsumoto, L. O. Chua, and H. Suzuki

Memorandum No. UCB/ERL M89/65

25 May 1989

TITLE PAGE

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

CNN CLONING TEMPLATE: CONNECTED COMPONENT DETECTOR

by

T. Matsumoto, L. O. Chua, and H. Suzuki

Memorandum No. UCB/ERL M89/65

25 May 1989

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

CNN CLONING TEMPLATE: CONNECTED COMPONENT DETECTOR†

T.Matsumoto, L.O.Chua and H.Suzuki ††

† This research is supported in part by the Japanese Ministry of Education, the Institute of Science and Engineering, Waseda University and by the US office of Naval Research under Contract N000 14-89-J1402.

†† T.Matsumoto and H.Suzuki are with the Department of Electrical Engineering, Waseda University, Tokyo 160, Japan.

L.O.Chua is with the Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA 94720

Abstract A CNN template for ***connected component detection*** is reported. Using this template a handwritten character recognition system is proposed. An initial test result already shows 94% ~ 100% recognition rates for numerals.

I. INTRODUCTION

This note reports an extremely simple *cloning template* for CNN (***Cellular Neural Network***)[1] which is capable of detecting the ***number of connected components*** of a vector† in $\{+1,-1\}^N$. By exploiting this unique capability, an architecture for *a handwritten character recognition* system is proposed. A preliminary test result (100 handwritten numbers 0 ~ 9) already shows a 94% ~ 100% correct recognition rate.

II. CLONING TEMPLATE $\boxed{1 \mid 2 \mid -1}$

Consider first a 1-dimensional CNN defined by

$$C\frac{dv_{x_i}}{dt}=-\frac{1}{R_x}V_{x_i}+\sum_{C(k)\in N_r(i)}A(i;j)V_{y_k}+\sum_{C(k)\in N_r(i)}B(i;j)V_{u_k}+I \tag{1}$$

$$V_{y_i}=\frac{1}{2}\left(|V_{x_i}+1|-|V_{x_i}-1|\right),\ i=1....N \tag{2}$$

where C(k) denotes the k-th cell and $N_r(i)$ denotes the r-neighborhood of the i-th cell. Note that *the only nonlinearity* involved is the dependency of V_{yi} on V_{xi} as defined by (2). For simplicity, let us recast (1) into the following equivalent compact form

$$C\frac{dV_{x_i}}{dt}=-\frac{1}{R_x}V_{x_i}+A*V_{y_i}+B*V_{u_i}+I \tag{3}$$

where * denotes two - dimensional "convolution operator". Our cloning template is given by

$$\boxed{A=\boxed{1 \mid 2 \mid -1}\ ,\quad B=0\quad ,\quad I=0} \tag{4}$$

This astonishingly simple template is endowed with a remarkable property. In order to illustrate this property, consider the 1-d vector in $\{1,-1\}^{16}$ defined by Fig.1(a) where ■ = +1,and □ = -1. Let this vector be the *initial condition* for (3). Then (3) *converges* to the vector shown in Fig.2(a). Similarly, if one choses Fig.1(b) as the initial condition, then (3) converges to the vector shown in Fig.2(b). Namely, starting from the right most cell, the final state of the pixels *alternates* between

† Throughout this note, the N pixels along each row of a bipolar pattern will be represented by a row vector of 1's (for black) and -1's (for white). Hence, any *bipolar* pattern can be represented by a string of vectors in $\{+1,-1\}^N$.

■ and □, or vice - versa. Moreover,

$$\boxed{\begin{aligned} &\#(\,\blacksquare\text{'s in the final state}\,) \\ &= \#(\text{connected components of } \blacksquare\text{'s in the initial state}) \end{aligned}} \tag{5}$$

It should be noted that even though the initial and the final states of each pixel are discrete value (± 1), the "transient" values are continuous because of the dynamics (3). Fig.3 shows the time waveforms of $V_{xi}(t)$ corresponding to the initial state defined by Fig.1(a), $i = 1....16$, where

$$\begin{aligned} A(j;j-1) &= 1.0 \times 10^{-3}\Omega^{-1} \\ A(j;j\quad) &= 2.0 \times 10^{-3}\Omega^{-1} \\ A(j;j+1) &= -1.0 \times 10^{-3}\Omega^{-1} \\ R_x = 1k\Omega &, \; C = 1nF \end{aligned}$$

Note that each "positive pulse" ■ or connected ■'s is propagated toward right and settles down to a ■ next to a □. In fact, one can think of this circuit as a special purpose "shift register". For consistency in our interpretation, we *assume* that initial condition: $V_{x1}(0) = -1$, i.e., the *left most* cell must be □. If, $V_{x1}(0) = +1$, then, ■ and □ change their value, counting the number of connected components of □'s (See Fig.1(c) and Fig.2(c)), instead of the ■'s.

III. HANDWRITTEN CHARACTER RECOGNITION

Consider the 2-dimensional CNN defined by

$$C\frac{dV_{x_{ij}}}{dt} = -\frac{1}{R_x}V_{x_{ij}} + A * V_{y_{ij}} \tag{6}$$

where

$$A = \begin{array}{|c|c|c|} \hline 0 & 0 & 0 \\ \hline 1 & 2 & -1 \\ \hline 0 & 0 & 0 \\ \hline \end{array}$$

In order to illustrate how this CNN can be used for handwritten character recognition, consider the initial condition given by Fig.4(a), which is a handwritten character "5" represented as a bipolar(± 1) pattern. Fig.4(b) shows the final state. Observe that connected components along the "horizontal" direction are correctly detected. In a similar manner one can consider the connected components in the "vertical" and the "diagonal" directions using the related templates

$$A = \begin{array}{|c|c|c|}\hline 0 & 1 & 0 \\ \hline 0 & 2 & 0 \\ \hline 0 & -1 & 0 \\ \hline \end{array}\ , \quad \begin{array}{|c|c|c|}\hline 0 & 0 & -1 \\ \hline 0 & 2 & 0 \\ \hline 1 & 0 & 0 \\ \hline \end{array} \quad \text{and} \quad \begin{array}{|c|c|c|}\hline 1 & 0 & 0 \\ \hline 0 & 2 & 0 \\ \hline 0 & 0 & -1 \\ \hline \end{array} \tag{7}$$

respectively. Fig.4(c), (d) and (e) show the final states associated with the above templates. Note that there is a great amount of "data compression" already. Note also that the "compressed data" Fig.4(b) and (c) are *completely shift invariant*.Furthermore, they have certain amount of robustness against rotation as well as scale variations. The compressed data Fig.4(d) and (e) also acquire a reasonable amount of robustness against shift, rotation and scale variations.

Based upon these observations, the system given by Fig.5 was tested. The raw data consist of 100 for each of the handwritten numerals 0 ~ 9 supplied by a standard data base ETL3[2]. Since the raw data are ragged (see Fig.6, for instance), a preprocessor was used to remove noise. The preprocessed data was then fed in parallel into four CNN with the templates given by (6) and (7). The compressed data was further reduced into an 83 dimensional vector (details will be reported elsewhere) and then fed into the decision making network. The Back Prop Paradigm[3] was used with 83 input units, 10 hidden units, and 10 output units. Out of 100 data for each numeral, 33% was used for training the Back Prop. The network has never seen the remaining 67%. The recognition rates are given in Table 1.

Further research problem include:

1.Rigorous proof of why (3) with (4) gives (5).

2.Hardware implementations.

3.Applications to more complex characters including Japanese and Chinese characters.

REFERENCES

[1]Leon O.Chua and Lin Yang,"Cellular Neural Network",IEEE Trans CAS vol.35,No.10,pp1257 - 1272 and 1273 - 1290,1988

[2]ETL-3,Electro Technical Laboratory,Tsukuba,Ibaraki,Japan 1974

[3]D.E.Rumelhart, J.L.McClelland and the PDP Research Group, Parallel Distributed Processing, vol.1,2, Cambridge MA:Bradford, 1986

Figure Caption

Fig.1 Three different initial states for CNN.

Fig.2 The final states of CNN starting from the initial states of Fig.1.

Fig.3 Transient behaviors of the 16 cells.

Fig.4 Initial state (a) and final states (b) - (e) with four templates given by (6) and (7).

Fig.5 A character recognition system

Fig.6 Typical raw data supplied by [2].

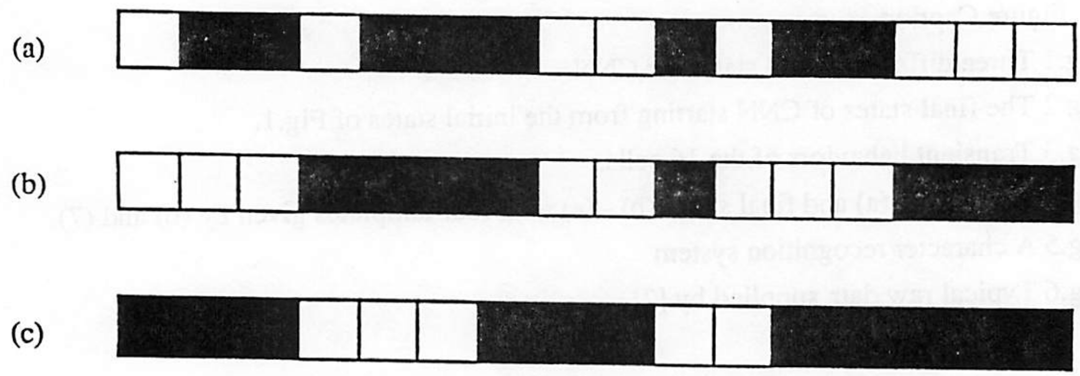

Fig.1

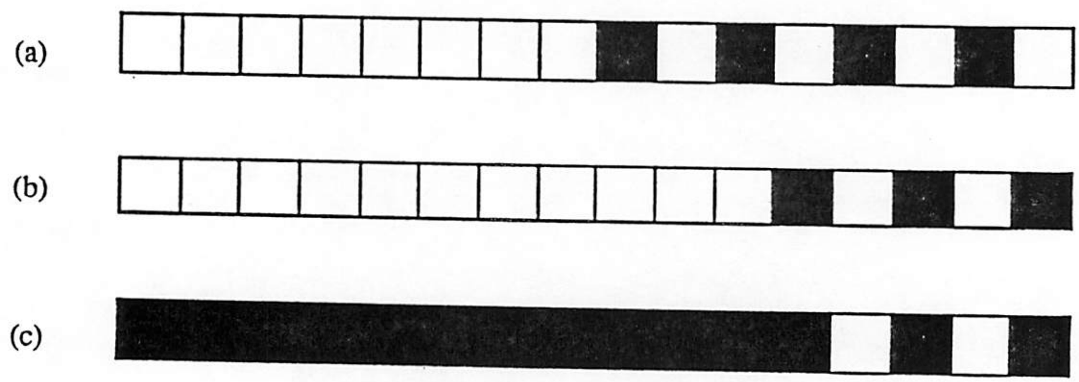

Fig.2

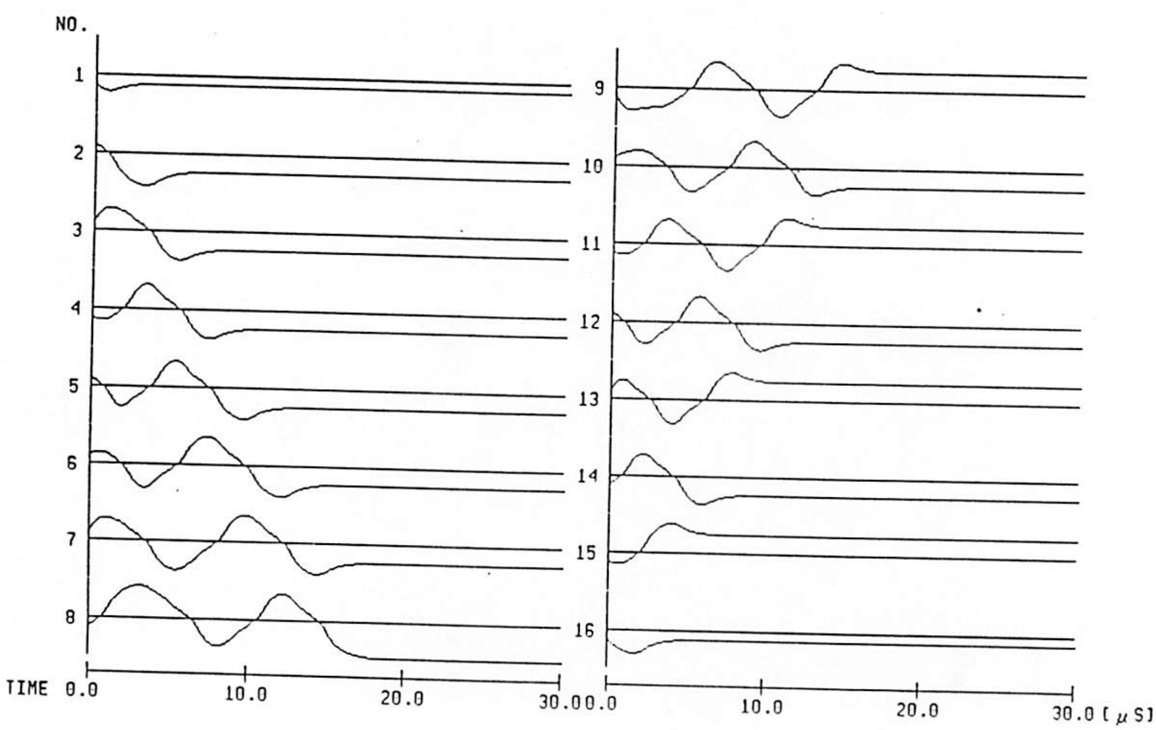

Fig.3

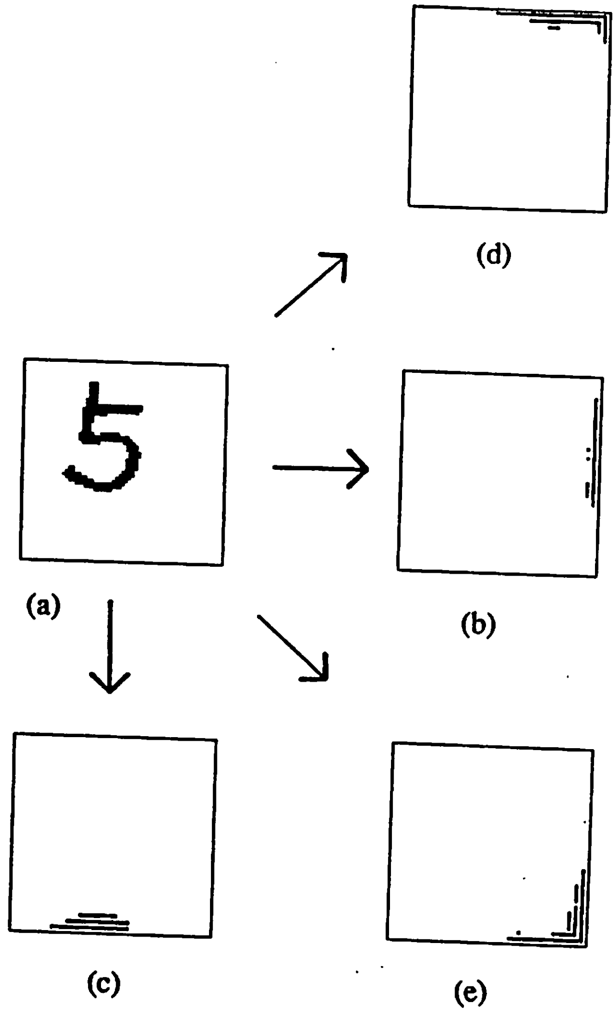

Fig.4

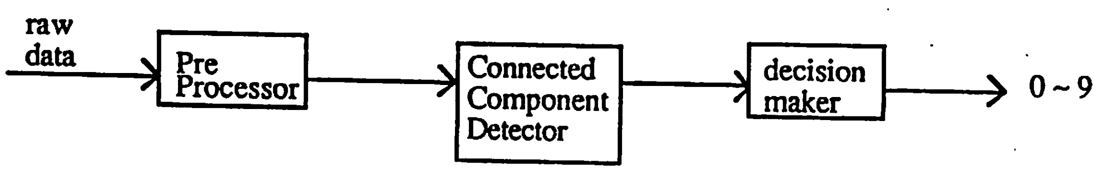

Fig.5

No.	recognition rate (%)
0	100.0
1	98.0
2	98.0
3	95.0
4	99.0
5	94.0
6	99.0
7	100.0
8	97.0
9	94.0

Table 1

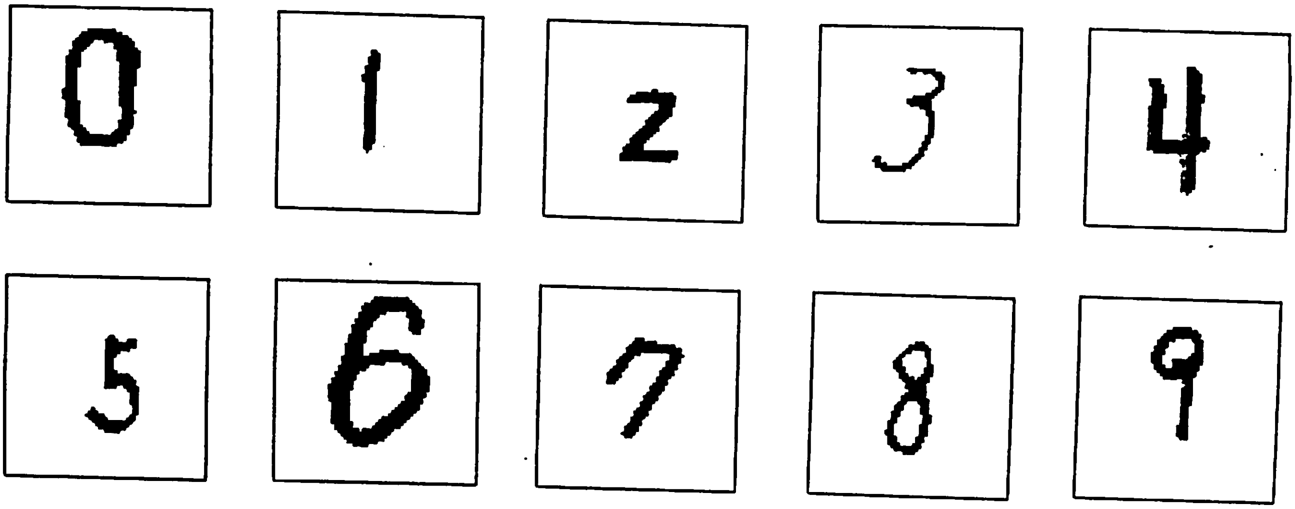

Fig.6