Object Recognition using Temporal Pattern Recognition Networks with Adaptive Segmentation of Quantizer Neuron Architecture (TASQA)

Susumu MARUNO* Chow Yuet Dih*** Yasuharu SHIMEKI* Takashi ANEZAKI** Yoshikazu Okahashi**
maruno@crl.mei.co.jp chow@crl.mei.co.jp shimeki@crl.mei.co.jp

*Central Research Laboratories
Matsushita Electric Ind. Co., Ltd.
Moriguchi, Osaka 570, Japan

**FA Engineering Laboratory
Matsushita Electric Ind. Co., Ltd.
Kadoma, Osaka 571, Japan

***AV/Information Research Center
Asia Matsushita Electric (S) Pte. Ltd.
BLK 1022, Hougang Avenue 1, #04-3526
Tai Seng Industrial Estate, Singapore 195

ABSTRACT

One of the biggest issues of an object recognition is the recognition with rotation invariance under a fluctuating noisy environment. We proposed a Temporal Pattern Recognition Network with Adaptive Segmentation of Quantizer Neuron Architecture (TASQA) and a ϕ-s transformation of shapes and applied them to object recognition.

The shape of the object is converted to a series of angles as a function of the circumference of the shape (ϕ-s data) and can be treated as a series of temporal patterns. TASQA consists of networks (ASQA network) with quantizer neurons, which can proliferate themselves and form the the networks automatically during training, and a layer of neurons with self feedback (self feedback layer). The self feedback layer unifies the temporal recognition results of ASQA networks during a certain period defined by the time constant of self feedback and this function can realize the function of selective attention to certain areas of a series of temporal patterns. As a result, TASQA realizes rotation invariance in recognition and we obtained 100% recognition accuracy of 5 images with fluctuating noise taken by CCD camera .

ϕ-s TRANSFORMATION

Figure 1 shows a schematic diagram of ϕ-s transformation. Points on a boundary of the object are transformed into a series of counter clock wise angles ϕ as a function of the circumference of the boundary **s**. As a result, the shape of the object can be treated as temporal wave data and the rotation δ of the object is expressed as the vertical shift of the wave data. The ϕ-s data of the object is input to the ASQA networks in the object recognition system using TASQA.

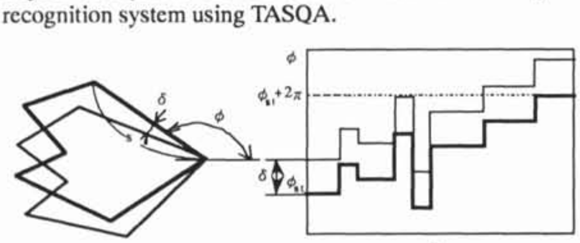

Figure 1. Schematic diagram of ϕ-s transformation

NETWORK STRUCTURE OF TASQA

Figure 2.1 shows the structure of TASQA. TASQA consists of ASQA networks and a self feedback layer. ASQA networks with quantizer neurons recognize a temporal pattern of ϕ-s data and send temporal results to the self feedback layer. The self feedback layer unifies the temporal results during a certain period defined by the time constant of self feedback.

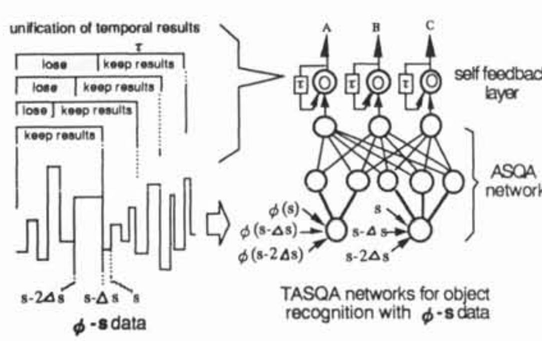

Figure 2.1. Structure of TASQA

STRUCTURE OF ASQA NETWORKS

One of the biggest issues of neural networks is how to design the structure. ASQA networks consists of quantizer neurons which can proliferate themselves and form the optimum network structure for the recognition automatically during training. As a result, there is no need to design the structure of the networks.

Principle of quantizer neuron

Figure 2.2 shows the basic structure of the quantizer neuron (QN). It has a quantizer, a quantizer data input terminal R, a selection signal input terminal S and output terminals O_j. The quantizer of the quantizer neuron maps the feature data input Q from terminal R into the discrete- valued d which indicates one of O_j. The weights τ_j of O_j (j=1 to j_{max}) are set according to the discrete-valued d using combination function given by equation (1).

As a result, weights are set dynamically according to the feature data input from quantizer input terminal as shown in figure 2.3.

The selection signal T is input from the terminal S and the output signals $T\tau_j$ are sent to the neurons connected to O_j.

$$\tau_j = g\,(j,\,\delta) \qquad (1)$$

$$(\text{ example: } g\,(j,\,\delta) = 1\text{-}b\,(j - \delta)^2 \;)$$

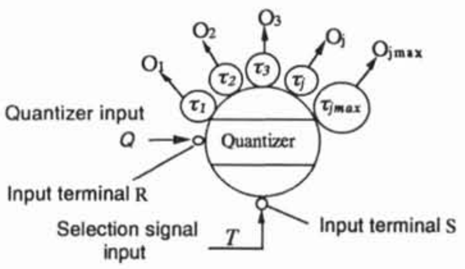

Figure 2.2. Structure of the quantizer neuron

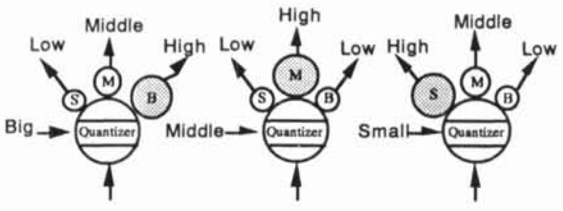

Figure 2.3. Schematic diagram of the function of quantizer neuron

Structure of quantizer for ASQA

The quantizer of the quantizer neuron for ASQA consists of quantizer units shown in figure 2.4. Each quantizer unit has an upper limit X_{Ui} and a lower limit X_{Li} for a range of quantization, a variable A_i for storing the average of input data, a variable σ_i for storing the variance of input data and a variable E_i for storing internal energy of each unit. Weight τ_i is connected to each quantizer unit and when the input data x is within the range of quantization of unit c, all weights are set according to equation (2) and equation (3).

$$\tau_c = 1 \quad when\ X_{Lc} \leqq x < X_{Uc} \qquad (2)$$

$$\tau_i = g(i,\,c) = 1\text{-}\,\beta\,(i - c\,)^2 \qquad (3)$$

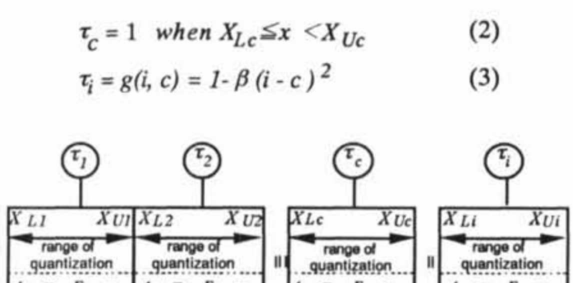

Figure 2.4. Structure of quantizer units

Adaptive segmentation of quantizer units

Quantizer neurons have only one quantizer unit and the upper limit X_{Ui} and the lower limit X_{Li} for the range of quantization of the unit are set to cover the range of input data at the initial condition.

When the data x is inputted, the quantizer unit calculates the average A_i of input data, the variance σ_i of input data, number of training iteration n and the internal energy E_i using each previous value according to equations (4) through (7).

$$A_i \cong \frac{n_{io} \bullet A_{io} + x}{n_{io}+1} \qquad (4)$$

$$\sigma_i \cong \frac{n_{io} \bullet \sigma_{io} + |\,x\text{-}A_i\,|}{n_{io}+1} \qquad (5)$$

$$n_i = n_{io} + 1 \qquad (6)$$

$$E_i = n_i \bullet \sigma_i \qquad (7)$$

When the internal energy E_i exceeds the threshold α, the quantizer unit U_i splits into two units U_i and U_{i+1}, as shown in figure 2.5. The upper limit X_{Ui} of the unit U_i is set to the average of input data A_i . The lower limit X_{Ui+1} and the upper limit X_{Ui+1} of the new unit U_{i+1} are set to the average of input data A_i and the previous upper limit X_{Ui} of the unit U_i according to equations (9) through (12). The average A, the variance σ_i, the internal energy E_i and the training iteration n of each units are also initialized according to equations (13) through (17).

<u>Condition of segmentation</u> :

$$E_i > \alpha \quad (\alpha : \text{threshold}) \qquad (8)$$

$$X_{Li'} = X_{Li} \quad (9)\,, \qquad X_{Ui'} = A_i \quad (10)$$

$$X_{Li+1} = A_i \quad (11)\,, \qquad X_{Ui+1} = X_{Ui} \quad (12)$$

$$A_i = \frac{X_{Li'} + X_{Ui'}}{2} \quad (13) \qquad A_{i+1} = \frac{X_{Li+1} + X_{Ui+1}}{2} \quad (14)$$

$$\sigma_{i'} = \sigma_{i+1} = 1 \quad (15) \qquad n_{i'} = n_{i+1} = 1 \quad (16)$$

$$\tau_{i+1} = \tau_i \quad (17)$$

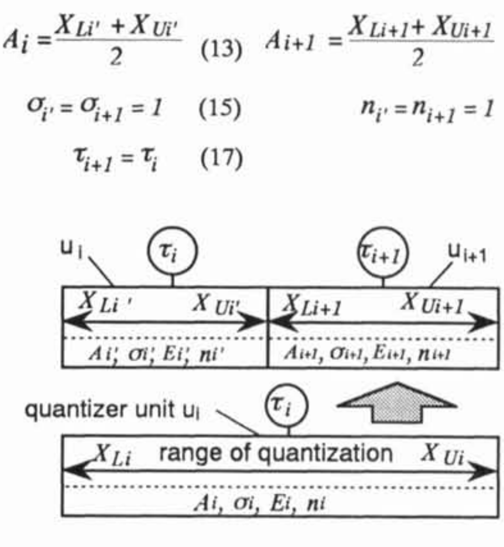

Figure 2.5. Segmentation of quantizer units

Proliferation of quantizer neurons

After the segmentation of quantizer units, a new neuron QN_2' which is the reproduction of QN_2 is made as shown in figure 2.6; thus, ASQA networks are constructed automatically as shown in figure 2.7. The output of the output neuron p of ASQA networks is given by equation (18).

$$y_p = f\left(\sum_{i=1}^{N} \sum_{j=1}^{max} t \cdot \tau_{ij} \cdot \tau_{ijp} \right) \tag{18}$$

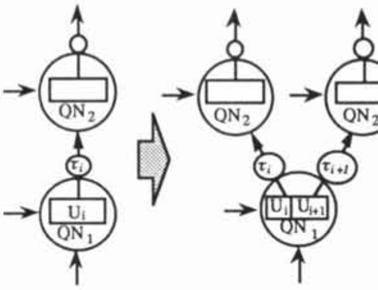

Figure 2.6. Proliferation of quantizer neurons

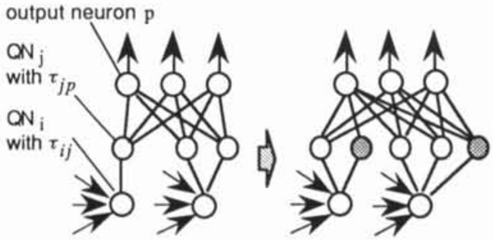

Figure 2.7. Self generation of ASQA networks

Structure of self feedback layer

The self feedback layer consists of neurons with self feedback. Each neuron is labeled as a target object. Each neuron p unifies the temporal recognition results $y_p(T)$ of ASQA networks during a certain period defined by the time constant Γ of self feedback and the output $u_p(T)$ at time T is given by equations (19).

$$u_p(T) = h\left(\int_0^T y_p(t) \cdot e^{-t/\Gamma} dt \right) \tag{19}$$

(h : threshold function)

Training of TASQA

During the training process, weights between the neurons of the output layer and the hidden layer of ASQA network are also changed so that the final output of self feedback layer becomes correct while the neurons proliferate. If the answer p is incorrect, supervisor datum p^* is input to the neurons of the hidden layer and the weight τ_{ijp^*} is increased by $\delta\tau_{ijp^*}$ according to the covariance rule.

EXPERIMENT OF OBJECTS RECOGNITION WITH TASQA

We took 5 different pictures(from image 1 to image 5) of the same sets of mechanical components including 10 objects by CCD camera. Each image has the same set of mechanical components; however, the shape of components are different image by image because of the difference of lighting and noises. We extracted the boundaries of objects as shown in figure 3.1 and converted those into ϕ-s data sets. We prepared TASQA networks as already shown in figure 2.1. The input data set to TASQA networks are $\phi(s)$, $\phi(s-\Delta s)$, $\phi(s-2\Delta s)$, s, s-Δs and s-2Δs as also shown in figure 2.1. Three neurons of self feedback layer are labeled as the target A, the target B and the other categories and we taught TASQA with 2 object (the object 11 in image 2 and the object 19 in image 4 shown in figure 3.1) as the recognition targets A and B. We performed a recognition experiment using data sets not used for training with various time constants of self feedback.

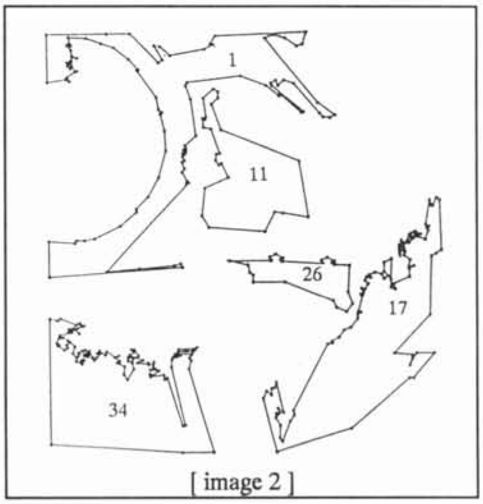

Figure 3.1. Example of the boundary shape

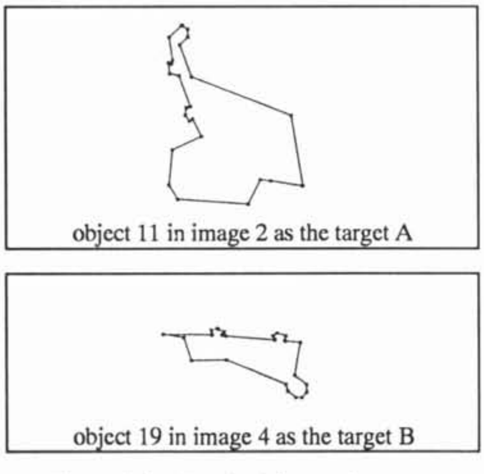

Figure 3.2. Taught object as the targets
The object 11 in image 2 is taught as the target A and the object19 in image 4 is taught as the target B.

Recognition results

[Recognition output of TASQA]

Figure 3.2 shows the output of each neuron in the self feedback layer as a function of circumference of the boundary (data s of ϕ-s data sets) of the objects when the ϕ-s data of object 7 in image 4 shown in figure 3.3 is input to TASQA and the time constant Γ is set to infinity.

Right hand side of object 7 was different from the trained object 11 in image 2 as the target A because of shadow and as a result, the output of the neuron labeled as the target A started to decrease and the output of the neuron labeled as the other categories started to increase from the point sx. This result shows that optimization of the time constant will give the most accurate answer.

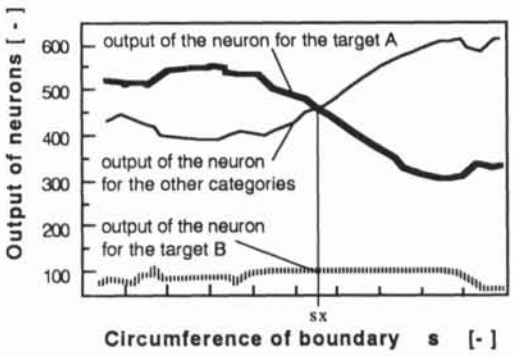

Figure 3.2. Recognition output vs. circumference of boundary of object

The output of the neuron for the target A started to decrease and the output of the neuron for the other categories started to increase from the point sx.

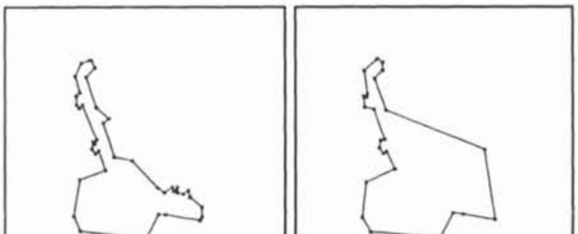

Figure 3.3. Boundary shape of input object

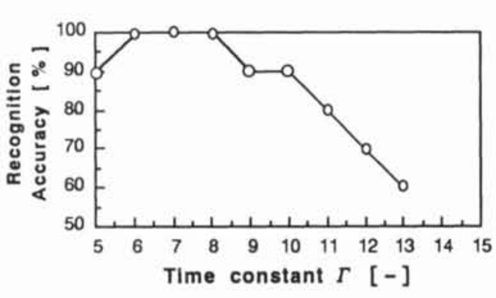

Figure 3.4. Average accuracy of object recognition vs. time constant Γ of self feedback

The recognition accuracy depended on the time constant Γ and the time constant of about half length of the target ϕ-s data sets used for training gave us the average accurate answer rate of 100%.

[Accuracy of the objects recognition]

Figure 3.4 shows the average recognition rate of the object recognition of 5 images as a function of the time constant Γ of self feedback. The recognition accuracy depends on the time constant Γ; and a time constant of about half the length of target ϕ-s data sets used for training gave us the best recognition rate of 100%.

DISCUSSIONS AND CONCLUSIONS

We proposed a Temporal Pattern Recognition Network with Adaptive Segmentation of Quantizer Neuron Architecture (TASQA) and a ϕ-s transformation of shapes and applied those to object recognition.

One of the biggest issues in object recognition is rotation invariance under an environment of fluctuating noise. We performed object recognition experiments using 5 images including a set of mechanical components taken by CCD camera with ϕ-s transformation and TASQA. The recognition accuracy was found to depend on the time constant Γ of self feedback. A time constant equal to about half the length of target ϕ-s data sets used for training gave us an average recognition accuracy of 100% for noisy rotated images.

TASQA consists of ASQA network with quantizer neurons and a self feedback layer. Quantizer neurons can proliferate themselves and form the the networks automatically during training; therefore there is no need to design the network. The self feedback layer unifies the temporal recognition results of ASQA networks during a period defined by the time constant of self feedback, and this function can direct attention selectively to certain areas of a series of temporal patterns. The selective attention areas can be easily adjusted by changing the time constant of self feedback according to the length of ϕ-s training data used as target objects to TASQA. Thus, TASQA can combat the biggest issues of recognition, rotation invariance under fluctuating noisy environment, and wider application of pattern recognition is expected.

Future tasks are (1)Optimization of the input data set of ϕ-s data and (2)Study on recognition of user specified points on the boundary of the shapes as grasp points for robot vision.

References

[1] S. Maruno, et al, "Shape recognition using temporal pattern recognition network with ASQA", Proc. 1992 IEICE Fall Conference, Part6, pp.6-38, September 1992.

[2] S. Maruno, A. Schoenberg, "Voice Command Recognition System for Handicapped Persons using Multi-Layered Networks", Proc. Annual Int. Conf. IEEE Eng. in Medicine & Biology Society Vol.10, Part4/4, pp.1590-1591, November 1988.

[3] S. Maruno, "Multi Functional Layered Network using Quantizer Neurons",Computer World '90,pp.202-209, November 1990.

[4] S. Maruno, et al, "Multi Functional Layered Network using Quantizer Neurons",Proc. IJCNN '91 Singapore, pp.541-546, November 1991.

[5] W. J. Freeman, "Lessons from Neurobiology for Feedback Architectures and Oscillatory Dynamics in Pattern Recognition with Artificial Neural Networks", Proc. Int. Conf. on Fuzzy Logic & Neural Networks.IIZUKA'90, Vol.1, pp.27-30, July 1990.

[6] D. Reilly, L. Cooper, C. Elbaum, "A Neural Model For Category Learning", Biol. Cybern., Vol.45,pp.35-41,1982.