Target Recognition in Mobile Robot Vision Systems via Neural Networks with Local Interconnections

LEONARDA CARNIMEO*
Dipartimento di Elettrotecnica ed Elettronica
Politecnico di Bari
via Orabona, 4, 70125 Bari
ITALY

Abstract: - In the field of robotics real-time image processing can provide the information necessary for mobile robots to execute a task in indoor environment. In this paper the application of a simple neural network-based system to translation and scale-invariant object recognition in the artificial vision structure of a mobile robot is proposed. The suggested bio-inspired vision system is mainly constituted by an encoder and a neural associative memory. Bipolar images constitute the input to the neural associative memory, which performs the recognizing stage of the bio-inspired vision system. The capability of the proposed system to detect and recognize scaled and translated targets is investigated on suitable test situations.

Key-Words: - Neural Networks, Cellular Associative Memories, Artificial Vision Systems

1 Introduction

The problem of recognizing objects and detecting targets in artificial vision systems independently of their position and size has been extensively investigated in recent years [1, 2]. In particular, high order neural networks have been already used for translation, rotation and scale-invariant object recognition [3-5]. However, most of these methods are too computationally expensive and cannot guarantee a real-time image processing, which is essential in many applications. As an example, by considering the features of artificial vision systems for mobile robots, the capability of real-time image processing reveals fundamental as it can provide the information necessary to execute a task in indoor environments. For this purpose, in this paper a vision architecture for object recognition and target detection based on a Cellular Neural Network (CNN) that achieves a satisfactory invariance under translation and scaling is developed. The choice of these Neural Systems with Local Interconnections is motivated by their well established implementability in VLSI and high speed operation [6, 7]. In fact, due to the local features of both Cellular Neural Networks and image segmentation techniques [8,9], a natural implementation of associative memories able to guarantee an extremely fast learning and recall is provided by these neural systems, being suitable encoding procedures already been adopted to map original images into binary segmented ones, adequate for the storage into a cellular associative memory [10, 11]. In this paper, the cellular neural-based architecture is described in Section II. For comprehension purposes, in Section III the synthesis procedure of Cellular Associative Memories developed in [9] is summarized. In Section IV the application of a simple cellular neural based system to translation and scale-invariant object recognition in the artificial vision structure of a mobile robot is developed. The class of object/background gray images is dealt with. The ability of the proposed architecture to detect and recognize scaled and translated targets is investigated and discussed on suitable test situations.

2 CNN-based Architecture of the Vision System

In this section, a bio-inspired robot vision system has been considered, which involves a camera, which captures images, and an image processing neural system, constituted by an encoder and a cellular associative memory with the aim of making a robot recognize correctly reference targets when moving in indoor environments. In particular, the encoder realizes the image thresholding of the original 256-level images captured by the camera [11]. A discrete-time CNN is then synthesized to behave as a neural associative memory by storing the reference bipolar images. Then, the images captured by the camera when the robot moves are segmented and compared with the reference ones stored into the neural associative memory. Therefore, the task of the bio-inspired cellular memory consists in realizing the real-time translation and scale-invariant pattern matching by comparing segmented images and reference memorized ones.

3 Neural Associative Memories via Discrete-Time CNNs

Taking into account the above reported considerations, in this section a Discrete-Time Cellular Neural Network (DTCNN) is designed to behave as an associative memory using the synthesis procedure developed in [9], which is hereby summarized for the aim of a good comprehension. The model of an (NxN)-cell rectangular DTCNN can be expressed in vector form as [9]:

$$\boldsymbol{u}(k+1) = \boldsymbol{T}\,\boldsymbol{v}(k) + \boldsymbol{I} \qquad (1a)$$

$$\boldsymbol{v}(k) = \boldsymbol{f}(\boldsymbol{u}(k)) \qquad (1b)$$

where $\boldsymbol{u} = [u_1,...,u_n]^{\mathrm{T}} \in R^n$ and $\boldsymbol{v} = [v_1,...,v_n]^{\mathrm{T}} \in R^n$ are the state vector and output one, respectively, with n= NxN, $\boldsymbol{I} = [I_1,...,I_n]^{\mathrm{T}} \in R^n$ contains the current sources values and $\boldsymbol{f} = [f,...,f]^{\mathrm{T}} \in R^n$, where the function $f: R \rightarrow R$ is a continuous, piecewise linear output function in the form

$$f(u) = (|u+1| - |u-1|)/2 \qquad (2)$$

The sparse matrix $\boldsymbol{T} = [T_{ij}] \in R^{n \times n}$ is the interconnection matrix, which takes into account the local connection property of the cellular neural network architecture. Any point $\boldsymbol{u}^i \in R^n$ is said to be an equilibrium point of (1) if

$$\boldsymbol{u}^i = \boldsymbol{T}\,\boldsymbol{v}^i + \boldsymbol{I} \qquad (3)$$

where $\boldsymbol{u}^i = [u_1^i, u_2^i, \ldots, u_n^i]^{\mathrm{T}} \in R^n$ and $\boldsymbol{v}^i = [v_1^i, v_2^i, \ldots, v_n^i]^{\mathrm{T}} \in R^n$. It can be proved that the suggested model assures the asymptotic stability of any equilibrium point of system (3), which is a necessary condition to generate an associative memory. In particular, each of the m images, which constitute the set of memories $\boldsymbol{v}^i$, i = 1,...,m, to be stored, has to correspond to an equilibrium point $\boldsymbol{u}^i$ of the DTCNN. Equation (3) can therefore be rewritten in compact form as:

$$\boldsymbol{U} = \boldsymbol{T}\,\boldsymbol{V} + \boldsymbol{I}' \qquad (4)$$

where $\boldsymbol{V} = [\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots, \boldsymbol{v}^m] \in R^{n \times m}$, $\boldsymbol{I}' = [\boldsymbol{I}, ..., \boldsymbol{I}] \in R^{n \times m}$ and $\boldsymbol{U} = [\boldsymbol{u}^1, \boldsymbol{u}^2, \ldots, \boldsymbol{u}^m] \in R^{n \times m}$.

Our objective consists in determining the matrices $\boldsymbol{T}$ and $\boldsymbol{I}'$ so that the constraint (4) is satisfied.

Equation (4) can be written as:

$$\boldsymbol{R}\,\boldsymbol{w}_k^{\mathrm{T}} = \boldsymbol{U}_k^{\mathrm{T}} \qquad k = 1, ..., n \qquad (5)$$

where $\boldsymbol{R} = [\boldsymbol{V}^{\mathrm{T}} |\ \boldsymbol{J}\ ] \in R^{m \times (n+1)}$; $\boldsymbol{w}_k = [T_{k1}, T_{k2}, \ldots, T_{kn} \mid I_k] \in R^{1 \times (n+1)}$; $\boldsymbol{U}_k = [u_k^1, u_k^2, \ldots, u_k^m] \in R^{1 \times m}$, k= 1,.., n and $\boldsymbol{J} = [1, 1, ..., 1]^{\mathrm{T}} \in R^{m \times 1}$.

Equation (5) has to be solved taking into account the constraints dictated by the DTCNN structure in the synthesis procedure and defining a matrix $\boldsymbol{S} = [S_{jk}] \in R^{n \times n}$ as follows:

$S_{jk} = 1$ if the k-th cell belongs to the same r-neighbourhood of the j-th cell;

$S_{jk} = 0$ otherwise (j = 1, ..., n; k = 1, ..., n).

Now, a reduced matrix $\boldsymbol{R}_{rk}$ can be obtained from the matrix $\boldsymbol{R}$ by eliminating those columns the indices of which correspond to the zero elements in the k-th row of $\boldsymbol{S}$. Moreover, a vector $\boldsymbol{w}_{rk}$ can be defined as the vector obtained from $\boldsymbol{w}_k$ by eliminating its zero elements. Thus, from (5) it results:

$$\boldsymbol{R}_{rk}\,\boldsymbol{w}_{rk}^{\mathrm{T}} = \boldsymbol{U}_k^{\mathrm{T}} \qquad k = 1, ..., n \qquad (6)$$

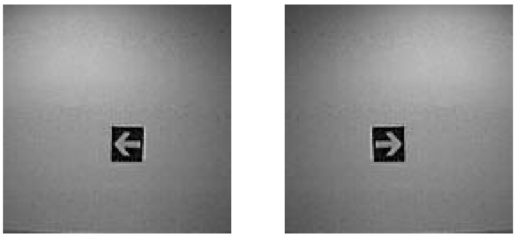
From (9) it follows:

$$\boldsymbol{w}_{rk}^{\mathrm{T}} = \boldsymbol{R}_{rk}^{+}\,\boldsymbol{U}_k^{\mathrm{T}} \qquad k = 1, ..., n \qquad (7)$$

where $\boldsymbol{R}_{rk}^{+}$ denotes the pseudo-inverse of $\boldsymbol{R}_{rk}$ [9]. The synthesis procedure concludes by expanding the vector $\boldsymbol{w}_{rk}{}^{T}$ with zero elements until the vector $\boldsymbol{w}_k{}^{T}$ is obtained.

4 Simulation Results

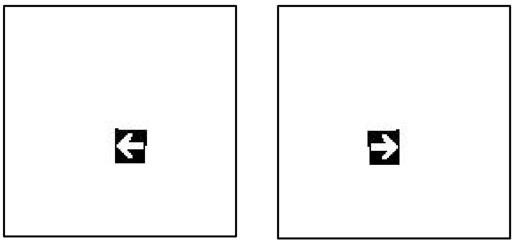
The performance of the proposed CNN-based vision system is investigated by considering the reference gray-level images reported in Fig.2.



(a) (b)

Figure 1: Gray-level (128x128)-captured images

In this figure two (128x128)-patterns, as seen by the artificial vision system of a mobile robot, are visualized. The target scene consists of an object shaped as an arrow positioned on a background. Since the aim of the suggested robot vision system consists in recognizing a pattern independently of its size and translated position via a DTCNN architecture, a preprocessing thresholding stage has been applied to the original pattern to obtain the bipolar reference images to be memorized in the cellular associative memories (-1=black; 1=white). These reference images, segmented with a threshold T=59, are shown in Fig. 2.



(a) (b)

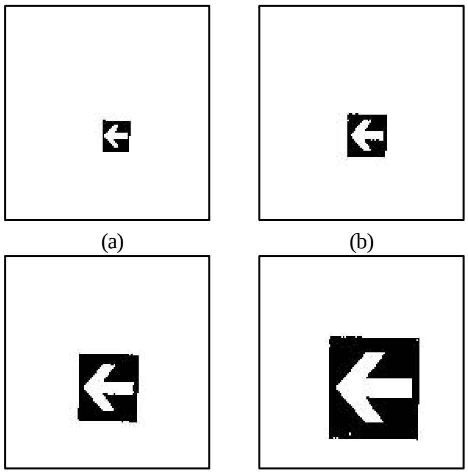
Figure 2: Bipolar reference images to be stored in the cellular associative memory

The cellular memory has been designed by considering a (128x128)-cell DTCNN with the neighbourhood reported in [9] (r = 1).

After designing the CNN-based memory, its recovery capabilities have been tested by considering sequences of images corresponding to the progressively closer vision of a mobile robot moving in an indoor environment. Interesting results have been obtained.

As an example, in this work a sequence of 18 images, representing the target pattern of Fig.2a, in a scaled view has been considered.

In Fig.3 four selected images of the sequence are reported.



(a) (b)

(c) (d)

Figure 3: Selected images of the sequence: image 1; (b) image 7; (c) image 14; (d) image 18.

The recovery capabilities of the designed architecture have been tested by assuming each image of a scaled pattern of the sequence as a noisy image of the corresponding stored reference one to be submitted to the designed memory.

Object matching results were encouraging, since quite significant values of the recall rate defined as:

$$\text{Recall rate} = \frac{\text{Number of correct pixels}}{\text{Total amount of pixels}} \cdot 100$$

were achieved just after few steps.

To better illustrate the memory retrieving capabilities, two selected noisy patterns and the related dynamic evolutions have been reported in Figs.4 and 5, respectively.

In Fig.4 the original pattern has been completely recovered after 30 steps, even though satisfactory results have been reached after 12 steps (Fig.6).

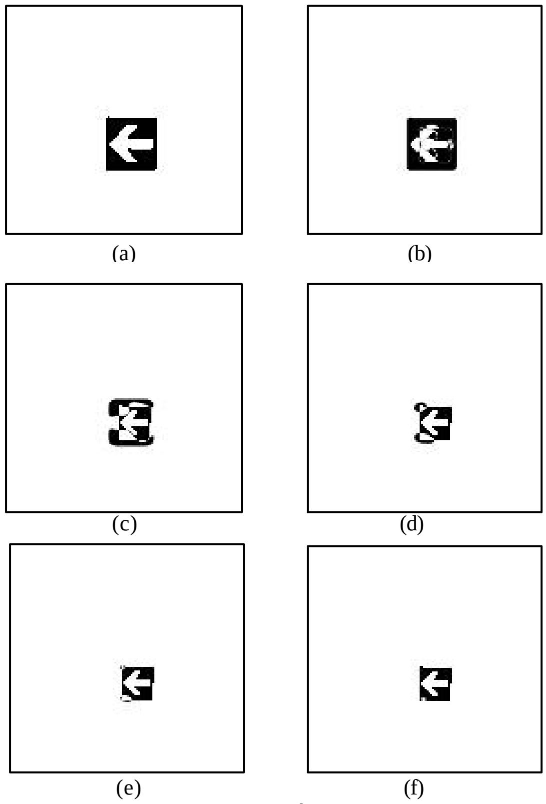


Figure 4: Recovery evolution of the designed DTCNN: (a) submitted 9-th image of the sequence; (b) output at step 1; (c) output at step 10; (d) output at step 21; (e) output at step 26; (f) final output at step 30.

The reference pattern reported in Fig.5 has been completely recovered after 61 steps.

From Fig.7 it can be noticed that significant results can be already reached after 47 steps (Fig.7).

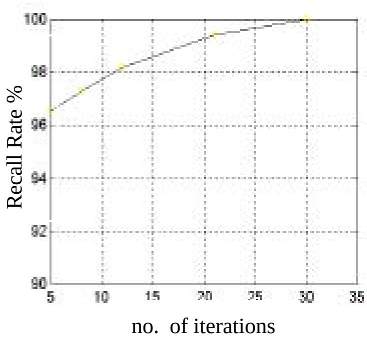


Figure 6: Diagram of the recall rate vs the no. of iterations for the recovering of image 9

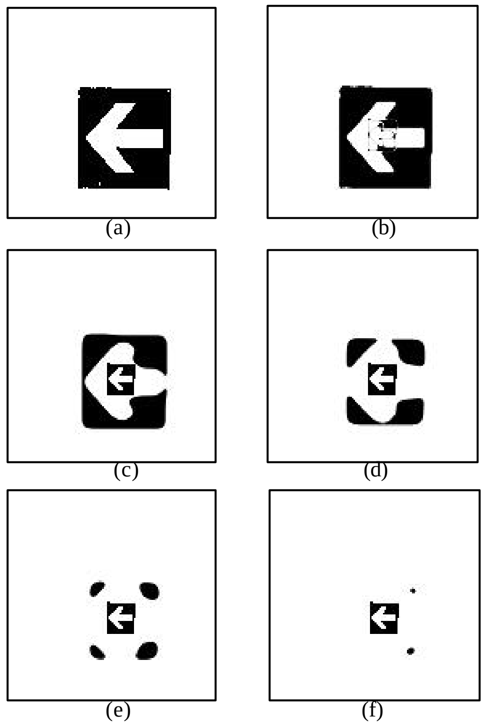


Figure 5: Recovery evolution of the designed DTCNN: (a) submitted 18-th image of the sequence; (b) output at step 1; (c) output at step 15; (d) output at step 30; (e) output at step 45; (f) output at step 58.

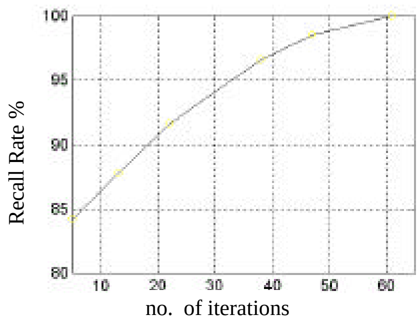


Figure 7: Diagram of the recall rate vs the no. of iterations for the recovering of image 18

Finally, several tests for translation-invariant recognition have been carried out.

For this purpose, selected translated images of the reference patterns have been reported in Fig.8. Also in this case the recovery capabilities of the system have been analyzed by assuming each translated input image as a noisy image of the corresponding stored reference pattern. In the considered cases the original patterns were recovered in about 15 steps.

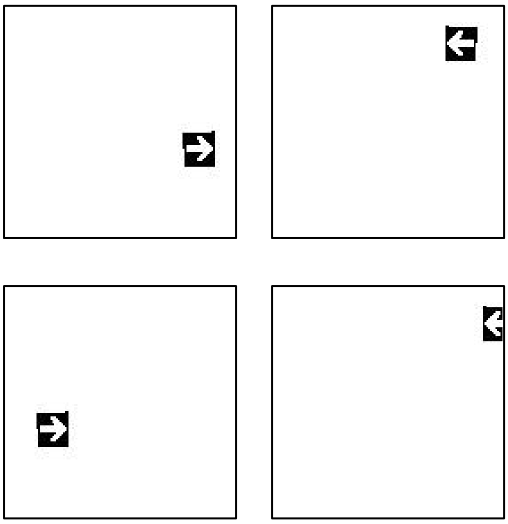


Figure 8: Translated images of the original patterns

5 Conclusions

In this paper the application of a neural network-based system with local interconnections to translation and scale-invariant target recognition in the artificial vision structure of a mobile robot has been developed. The capability of the proposed system to detect and recognize scaled and translated targets has been investigated on suitable test situations. Simulation results have shown satisfactory recovery capabilities both for scaled and translated patterns.

Acknowledgements

This work was partly supported by the Ministero dell'Università e della Ricerca Scientifica and partly by the Politecnico di Bari.

References

[1] L. A. Torres-Mendez, J. C. Ruiz-Suarez, Luis E. Sucar, and G.Gomez, Translation, Rotation and Scale-Invariant Object Recognition, *IEEE Trans. SMC Part C*, vol.30, no. 1, 2000, pp.125-130.

[2] T. Tsuji, K. Ito and P. G. Morasso, Neural Network Learning of Robot Arm Impedance in Operational Space, *IEEE Trans. SMC, Part B*, vol. 26, no.2, April 1996, p. 290.

[3] S.G. Perantonis, and P.J. G. Lisboa, Translation, Rotation and Scale-invariant Pattern Recognition by High-order Neural Networks and Moment Classifiers, *IEEE Trans. Neural Networks*, vol.3, no. 3, 1992, pp.241-251.

[4] L. Spirkovska, and M.B. Reid, Coarse-code higher-order neural networks for PSRI object recognition, *IEEE Trans. Neural Networks*, vol.4, no. 7, 1993, pp.276-283.

[5] M. Brucoli, L. Carnimeo, and G. Grassi, 'Associative Memory Design using Discrete-Time Second-Order Neural Networks with Local Interconnections, *IEEE Trans. CAS-I*, vol.44, no. 1, 1997, pp.1-6.

[6] K. R. Crounse, T.Roska and L. O. Chua, Image Halftoning with Cellular Neural Networks, *IEEE Trans. CAS-II.*, vol.40, no.4, 1993, pp.267-283.

[7] P. L.Venetianer, F.Werblin and T.Roska, Analogic CNN Algorithms for some Image Compression and Restoration Tasks, *IEEE Trans. CAS-I.*, vol.42, no.5, 1995, pp.278-284.

[8] H. Harrer, J. A. Nossek, and R. Stelzl, 'An Analog Implementation of Discrete-Time Cellular Neural Networks,*IEEE Trans. Neur. Net.*, vol.3, 1992, p.466.

[9] M. Brucoli, L. Carnimeo, and G. Grassi, 'A Global Approach to the Design of Discrete-Time Cellular Neural Networks for Associative Memories, *Int. J. Circuit Theory & Appl.*, vol.24, no.4, 1996, p.489.

[10] M. Brucoli, D. Cafagna, and L. Carnimeo, Design of Cellular Associative Memories for Robot Vision via a Fuzzy Image Segmentation, *Proc. of ECCTD'99*, Aug. 29 – Sept.2, Stresa, Italy, 1999.

[11] P. K. Sahoo, S. Soltani, A. K. C. Wong, Y.C.Chen, A survey of thresholding techniques, *Comp. Vision, Graphics, & Image Proc.*, vol.41x, 1988, pp.233-260.