Abstract—A visual target tracking identification by employing using a “Winner-takes-all” artificial neural network is proposed in this paper. In this approach a modified Kohonen Neural Network is the mechanism used both to determine the position as to represent the target trajectory given a sequence of images. Some of the advantages employing this technique is that the initial condition are supplied randomly and that the performance of the algorithm is independent of the initial condition as well as of the number of them. Besides, this algorithm converge for the center of mass of the target. This methodology is useful in remote and local systems when information is given by images be it related to aerospace applications, robotics, radar systems, or industrial applications. The proposed algorithm is here used in the identification of airplane trajectory by using digital images.

I. INTRODUCTION

The position identification and trajectory tracking of fixed and moving targets is focus of interest in diverse areas of application. Identification and analysis of movement, surveillance visual systems, visual interface of communication and behavior analysis, only to mention few, are examples of such an application [1]. Besides that, this approach can be employed, for instance, in aerospace segment – launcher trajectory tracking, flights, images of satellite with objects in movement –, in robotics, and industrial applications.

The identification of static and dynamic patterns by employing visual perception is a complex task that computers should be able to carry out in order to emulate the human behavior. In the visual perception tasks, human beings are able to easily recognize regularities when extracting information of environment or when interacting with other objects or agents. These patterns are detected and identified through input data by biological neural networks [2] concurrently that establish specific brain behavior for this activity. Through the biological visual system it is possible to obtain both the position as the properties of the objects and its relationship with itself and with the environment in which it is inserted.

The field of computer vision is composed of systems that simulate the human visual mechanisms. An alternative for emulation the human visual perception system is to employ techniques based on the human brain model. In doing so, there are artificial neural networks which presents similarities and characteristics that incite this use for practical problems such as identification, classification, digital image and signal processing and control. The main characteristics that motivate the use of neural network are its ability to learn from experience, adaptability when facing adverse conditions, and noise and disturb tolerance [3], [4], [5].

The moving target tracking problem may be solved by techniques from many areas. In particular, the field of artificial intelligence and computational intelligence are natural source of solution since they aim to emulate human behavior in machines. Techniques of artificial intelligence applied to digital signal and image processing have been used to solve problems of recognizing and tracking of moving targets. Among the existing methodologies, the use of fuzzy systems is employed for moving-target tracking problem [5]. In order to find the global optimal fuzzy tracker this approach consider the local concept approach as a method for tracking a moving target in an approximate field. Tracking moving target using fuzzy neural network is presented in [6]. The position of the target is estimated by angle and distance estimators assuming that moving target radiates narrow band waves. A neural network directed conditional probability generator and sequential classifier based on Bayes decision rule is found in [7] for target dynamic behavior and target classification. The dynamics is determined by using velocity/acceleration and curvature sequences from each track. A new biological neural network inspired in the Hodgkin and Huxleys biological membrane model is presented in [8] to dynamic collision-free trajectory generation in a non stationary environment.

When dealing with visual information, the approach in [9] uses an artificial neural network that receives as input the image for supplying the direction, distance of the scenario and the position of the moving target. Inspired by the fly visual system of Diptera males for conducting fast-moving aerial pursuits and interceptions of females a novel network for determining the motion of objects is proposed in [10] that uses as input the light intensity variation. Also inspired in animals, a neural network structure based on bat is used to face recognition and velocity of a dynamic target moving toward the camera [11]. In [12] a multiple elastic modules model based on Self-Organizing neural network incorporating a mechanisms analogous to thermodynamic temperature value – i.e., related to the nonhomogeneous adaptive temperature field – is described to escape of poor local minima when used for passive tracking problem. A feedforward neural network is used in [13] for static and moving identification for robot control. This approach also predicts the position of the target and of the manipulator by using time derivatives of the position of the object.

Ernesto Araujo is both with Computer Science / Health Informatics Department (DIS), Universidade Federal de Sao Paulo (UNIFESP), R. Botucatu, 862, 04023-062, Sao Paulo, SP, Brazil (email: ernesto.araujo@unifesp.br) and with Integration and Testing Laboratory (LIT), Instituto Nacional de Pesquisas Espaciais (INPE), Av. Astronautas, 1758, 12.227-010, Sao Jose dos Campos, SP, Brazil (email: ernesto.araujo@lit.inpe.br).
[14], an non-supervised neural network was employed in the development of a data composition system in a radar data multichannel mechanism for target identification. The characteristics of these data from three distinct radar channels are extracted from digital signal processing. Finally, a review of the algorithms used for detection and tracking in optical weak small targets environments is described in [15]. A tutorial for target tracking is described in [16] emphasizing the multisensor management and the fusion algorithms. A survey on neural-network-based target tracking and the state-of-the-art is presented in [17].

Following the successful application of artificial neural network in dealing with visual target tracking this paper addresses the problem of position identification and moving tracking by employing a modified Kohonen Neural Network. This approach can be employed as off-line as on-line according to the problem to be solved and is based on the KNN described in [18], [19], [20], [21] for pattern recognition in visual analysis.

II. FINDING AND FOLLOWING THE TARGET BY USING COMPETITIVE NEURAL NETWORKS

The Kohonen Neural Network (KNN) is an approach able to recognize patterns and clusterings to group similarities of input data with the advantage of not requiring previous training or learning. The network updates its parameters while in use [3], [6], [7]. Contrary to supervised training, which is accomplished when a desired output corresponding to a given input is furnished, the learning in the KNN is obtained just presenting the input to the network. In so doing, the network adapts their weights in such a way to follow data distributed in groups, i.e., clusters. This strategy is also known as the “winner-takes-all” and may be characterized as a competitive learning network, as well. During the competitive learning, the neurons fight for training their own weights. For each input presented to the KNN only one neuron is activate (Fig. 1). This neuron is labeled winner and assumes this condition due to the fact that it is the closest of the pattern computed and grouped from previous inputs. The learning mechanism detects groups by geometric closeness of similar characteristics of input data, according to eq. 1:

$$
\|x - w_j\| = \|x - w_i\| \forall i \neq j ,
$$

where $x$ corresponds to an output $y$ in such a way that only one neuron is activate, $y = +1$ while the other neurons are considered to be in inhibition.

The winner neuron is the best to be trained, i.e., to step forward to a set recognized and grouped according to previous inputs. At each input data presented to the algorithm the winner neuron tends to approach the group.

These characteristics motivate the use of KNN in the attempt of identifying the position and trajectory of fixed and moving targets, since there is no exigencies of an anticipate training and could be carried out in a fast manner.

When applied to target tracking and position position, the proximity of the neurons and the target is determined by the Euclidean distance between the position vector, $x$, and the synaptic weights, $w_i$, in all the dimensions that characterize the target:

$$
D = \|x - w_j\| = \left[\sum_{i=1}^{n}(x_i - w_{ij})^2\right]^{1/2}
$$

where “$j$” is the number of neurons, and “$i$” is the number of elements in each vector $x$ and $w_i$. The proposed approach may be used with any dimensional problem and the states of the system of interest are represented as $x^k = [x_1^k, x_2^k, \ldots, x_n^k]$ that represent the $n$ clusters, in such a way that $k$ is related to the amount of elements of the input pattern. The number of clusters or classes, $n$, is assumed to be previously known [20].

The winner neuron, $x_j$, is the one that is closest of the vector $x$. Thus, the best neuron is the one who presents the small Euclidean norm:

$$
x_j = \min_j \|x - w_j\|
$$

The Kohonen learning (clustering) rule is, then, described by the similarity matching:

$$
\|x - \hat{w}_i\| = \min_{1 \leq i \leq n} \{\|x - \hat{w}_i\|\}
$$

followed by the updating stage in which the KNN teaches the winner neuron by driven it closer to the inputs as possible:

$$
\hat{w}_{j(new)} = \hat{w}_{j(old)} + \beta (x - \hat{w}_{j(old)})
$$

$$
\hat{w}_{i(new)} = \hat{w}_{i(old)}, \forall i = 1, 2, \ldots, n, i \neq j,
$$

Fig. 1. General description of Kohonen Neural Network learning.

Fig. 2. New learning curve for the modified KNN.
where \( \beta \) is the learning rate, i.e., the step of training of the neural network and \( \hat{w}_j \) is the normalized vector:

\[
\hat{w}_j = \frac{w_j}{\|w_j\|}.
\]  \( (6) \)

The approach presented here is different from the classical KNN due to the learning curve that guides each neuron of the network. Usually, the function employed in the classical KNN is the Mexican hat. The proposed approach employs a modified KNN in order to guarantee that the winner neuron achieves the target and the looser neurons go toward to contrary directions. The learning curve used for training that differentiate the proposed approach from the classical one is shown in Fig. 2.

Consider, for example, the use of KNN in the position identification of static target. It is similar in dynamic analysis as the target being in equilibrium. The neuron that was initially considered winner usually approximates to the target step-by-step and, in this sense, its weight achieve a closer target distance in each update movement, i.e., iteration of the algorithm.

The convergence of the winner neuron toward to a static target presents, in general, a decaying behavior as the one depicted in Fig. 3(a). The convergence to the target is determined by the number of elements, \( n \), of the input patterns for the same class:

\[
\lim_{k \to n} \|X - X_{eq}\| \to 0
\]  \( (7) \)

for each image supplied in a certain sample. When the target is in equilibrium the winner neuron of the KNN shows a convergence approximately as that depicted in Fig. 3(b).

This convergence of the neuron is not limited to the particular case of the target in equilibrium, i.e., not moving. It is also valid for the task of linear trajectory tracking but with a decaying slower than when the target is fixed. When applied to nonlinear trajectories there is a convergence except that it is not exponential as it is for the linear moving target or fixed target problem. It is true both for fixed or variable learning. If the step size is constant, it may be considered as an additional parameter for the convergence of the neural network. In these conditions, the convergence is not purely exponential – oscillations may occur while there is target tracking.

When taking into account the step size it is determined as the learning rate value and the distance of the target and the neuron as in eq. (5). Although varying, the step size may assumes a high or small value mostly in function of the learning rate value. When high, the algorithm fast approximate to the target, yet may never really converge to
(a) From initial condition to target position 1.  
(b) From target position 1 to position 2.  
(c) From target position 2 to position 3.  
(d) From target position 3 to position 4.  
(e) From target position 4 to position 5.  
(f) From target position 5 to position 6.  
(g) From target position 6 to position 7.  
(h) From target position 7 to position 8.  
(i) From target position 8 to position 9.

Fig. 5. Neuron departs from randomly initial position and identification the target at the position $i$.

it since most of time the neuron will jump the central point of the target. In the other hand, if the step size is small the algorithm may never reach the target when changing from a certain image to another one, since the number of inputs reflecting the target may be not enough when compared to the step size. Alternatives exist to deal with this sort of problem. One is to employ a variable learning rate where the step size to up-to-date is also computed in function of the error obtained from the difference of the target and the current neuron. The other one is to determine an medium step size value. initially, in this paper, the later alternative is employed.

III. ILLUSTRATIVE EXAMPLE

The following example uses virtual images of airplanes for validate the proposed algorithm both in static and dynamic applications. There are single as multiple targets being used with distinct number of initial conditions. In all examples, targets are randomly disposed and so are the initial conditions. The learning step assumes the value $\beta = 0.03$ for all simulations.

A. Single and Multiple Fixed Target Identification

Initially, the algorithm is employed to a image with two targets in order to verify its efficiency. The line represents the movement of the neuron determined by the weight updated image-by-image as shown in Fig. 4(a). The neural initial positions for each iteration are represented by the asterisks and are automatically and randomly generate. In a glance, it is possible to notice that the algorithm performance is independent of the position of the initial condition. The neurons converge each one for distinct targets. While one neuron is attracted to one target the other converge to the second one.

Another advantage of this method is the possibility for generating as many as initial condition as desired without interfere in the algorithm performance while neurons converge to the target. In this case, the number of neurons (initial conditions) are activate as there is targets. In order to show this characteristic two targets being pursuit when presented four initial conditions is illustrated in Fig 4(b). While two of them converge to the target, the others stay out of the target.
The initial conditions that are geometrically not distant to the target are those that approximate to the target. It is worth mentioning that the KNN converge to the center of mass of the target.

B. Multiple Moving Target Identification

The proposed approach is, then, extended to a sequence of images that contain a target moving according to a random trajectory.

The results obtained when using the modified KNN to track the target is given in Fig. 5 and Fig. 6. The first one shows the result when tracking an isolated target step-by-step for each frame. In this example, the number of initial conditions is equivalent to the target. The second example illustrates the case where the number of initial conditions are larger than the number of targets. As occurred in the examples before mentioned, the same commentaries for static target are still valid for moving target tracking. The number of initial conditions, or their position does not interfere in the performance of the modified KNN as proposed here.

Additionally, for both conditions the center of mass of the neuron to reach the target. The oscillations are function of the learning rate (step size) used to train the KNN during the neuron weight update that is function of distance from neuron to target. The learning rate value is chosen for initial simulation was chosen to be constant, as carried out in these examples.

IV. Conclusions

The tracking problem has been treated by diverse tools. In this paper, a modified Kohonen neural network is used to identify positions and the tracking of trajectories based on in available visual data. This methodology may be employed both to remote or local systems when the information is related to image.

The proposed approach has shown efficiency both to position identification as to moving target tracking. When the target is static the Kohonen neural network was successful in identifying the target position. This result is extended to a sequence of images when simulating an airplane in movement. The proposed modified Kohonen neural network achieves the target trajectory.

The neural network was employed in various conditions to simulate diverse possible target behaviors. The simulation encompasses static and dynamic targets, be it with single or multiple targets, or yet be it with equal or different number of target and initial conditions.

Advantages of using the proposed technique is that computing the stochastic and statistic characteristic of this sort of behavior is not necessary. Furthermore, the initial conditions are supplied randomly, and the performance of the modified KNN is not dependent of the initial position and the amount of them.

REFERENCES


![Fig. 6. Moving target trajectory identification.](image-url)


