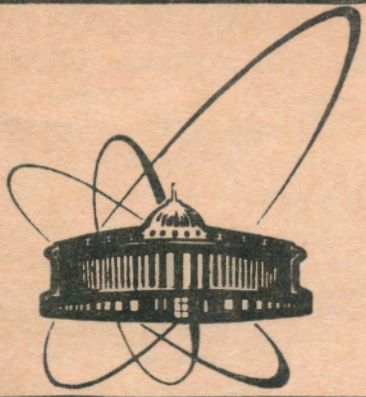

сообщения
объединенного
института
ядерных
исследований
дубна

E10-92-352

A.A.Glasov, I.V.Kisel, E.V.Konotopskaya,
V.N.Neskoromnyi, G.A.Ososkov

TRACK RECONSTRUCTION IN DISCRETE DETECTORS BY NEURAL NETWORKS

1992

1. Introduction

In recent years, in the high energy physics the interest increases in the neural networks [1, 2]. The basic application of neural networks is the search for local (or global) minima of given functions. These functions themselves are chosen in such a way that their minima correspond either to some fixed patterns (pattern recognition problem) or to optimal value of some variable (optimization problem). The efficiency of a network is determined by neuron nonlinear response and by strong interconnection between neurons. A neural network allows a fast solution of a wide class of problems with the great combinatorics complexity. Therefore, this gives much hope of solving the pattern recognition problem and constructing associative memory. Satisfactory solution of the problem can be reached in time of several iterations of a network.

In the present paper, on the basis of famous approaches of applying neural networks to the track finding problem [3] the investigations are made according to the specific properties of such discrete detectors in high energy physics as multiwire proportional chambers. These investigations result in the modification of the so-called rotor model in a neural network [2]. The energy function of a network in this modification contains only one "cost" term without any "punishment" terms which are determined usually by the constrains of a problem. This speeds up calculations considerably. The reduction of the energy function is done by the neuron selection with the help of simple geometrical and energetical criteria. Besides, the cellular automata were applied to preliminary selection of data [4] that made it possible to create an initial network configuration with the energy closer to its global minimum. Thus, the obtained computational algorithm was tested on 10^4 real three-prong events obtained from the ARES-spectrometer [5]. The results are satisfactory including the noise robustness and good resolution of nearby going tracks.

2. Neural networks

Artificial neural networks came from biology bearing the following features of biological neural networks with their ability to solve complicated problems of recognition and optimization:

- An element of the neural network is called *neuron* that can be connected with any other neurons of the network. All of them work in parallel.
- The synaptic strength (links) between neurons can vary in time.

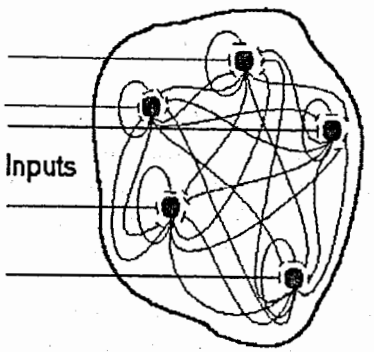

Figure 1: Neural network structure.

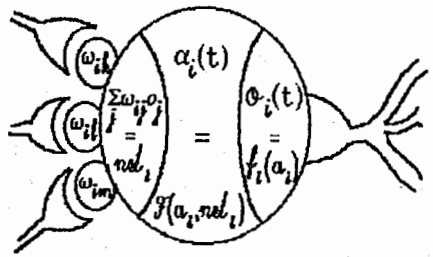

Figure 2: A neuron as a basic element of the network.

- The response function of a given neuron depends on its input and has a specific shape independently of the problem solved.

The information carried by a network is contained in the numerical values of these links. Data processing in such networks is performed in parallel by a great number of elements, which leads to fast calculations and high reliability of results. Fig. 1 shows a neural network as a structure of discrete computational elements — neurons, denoted by circles. In each moment every neuron u_i (Fig. 2) has an activation level $a_i(t)$ which by the function f_i can be transformed to the output level $o_i(t)$. The latter via directed links transmits to other neurons with different weights w_{ij}. These weights are the measure of the influence between two elements. The updating rule for the given neuron is the following: all inputs of the neutron are summarized and by the application of some function F_i the new neuron activation is defined. This leads to the iterative evolution of the system.

Let us consider one of the simplest examples of the neural network, namely the Hopfield model of statistical mechanics [6]. It consists of N binary neurons, described by the discrete variables s_i with two possible states (usually $\{-1, 1\}$). Synaptic strength between neurons is defined by symmetrical functions $w_{ij} = w_{ji}$, $w_{ii} = 0$. In pattern recognition problems, for example, the relation

$$w_{ij} \sim \sum_{m=1}^{M} \xi_i^{(m)} \xi_j^{(m)}$$

determines M patterns $\{\xi_i^{(m)}\}_{i=1}^{N}$ stored in our network[1].

To define the updating rule F it is necessary to assign to i-th neuron an activation threshold B_i so that the new value of the variable s_i is calculated from

$$s_i = \text{sign} \left(\sum_{i \neq j} w_{ij} s_j - B_i \right).$$

When all $B_i \equiv 0$, we have Glauber dynamics at zero temperature, which can be reformulated in terms of the local fields

$$H_i = \sum_{j=1}^{N} w_{ij} s_j \tag{1}$$

with the updating rule

$$s_i = \text{sign}\ (H_i). \tag{2}$$

As has been shown by Hopfield [6], the final state corresponds to one of the local minima of the energy function

$$E(s) = -\frac{1}{2} \sum_{ij} w_{ij} s_i s_j, \tag{3}$$

which is equivalent to the energy function for N interacting Izing spins s_i. The updating rule (1, 2) means that every spin s_i is to be oriented along the local field H_i that decreases the total energy (3).

From the point of view of thermodynamics this model is similar to the spin-glass model, in which the couplings between spins s_i are random. It is well-known [7] that the energy function of spin-glass in the configuration space has a lot of minima separated by very high and narrow potential barriers. This fact makes it difficult to reach the global minimum of the energy function $E(s)$, which is very essential for optimization problems. This obstacle can be avoided by the invention of some "temperature" T in the dynamics (1, 2) that permits subbarrier configurations and tunnelling to the global minimum. The original spins s_i are substituted by their thermal averages $v_i = \langle s_i \rangle_T$, the latter being determined only by the average value of the local field $U_i = \langle H_i \rangle_T$. Thus, modified dynamics is described by the mean-field-theory (MFT) new updating equation:

$$v_i = \tanh(U_i/T) = \tanh \left(\sum_{j=1}^{N} w_{ij} v_j / T \right). \tag{4}$$

[1] Here $s_i \equiv o_i = a_i$, $f(a_i) \equiv 1$.

A nontrivial solution of eq.(4) is possible only if $T < T_c \approx 1$ for $w_{ij} \sim 1$. The question arises about the possibility of tunnelling with such low temperatures. At the same time, a sufficiently high temperature can shift the position of global minimum, so that it is necessary to lower the temperature. Fortunately, the penetration into the potential barrier is proportional to the temperature when the latter is low enough, and the width of the barrier has the order of N^{-1}. In the survey [8], the result independence of the temperature in the region $T \sim 1$ is confirmed by numerical experiments.

3. Conventional track finding approach

Another example of neural network applications, more typical of high energy physics, is the problem of track finding. The results of measurements in a discrete track detector are the set of N points placed on n coordinate planes. Our task is to draw the minimum amount of smooth curves which connect points on different planes satisfied given restrictions. Let us consider the 2-dimensional case for the sake of simplicity. The possible restrictions might be: all curves converge at one point, the vertex; there are no bifurcations of tracks; all tracks must begin on the first plane and finish on the last one.

The conventional approach (see, e.g. [2]) introduces binary neurons s_{ij}. If points i and j are connected by a track, than $s_{ij} = 1$, and $s_{ij} = 0$ otherwise. The energy function consists of two terms:

$$E = E^{(cost)} + E^{(constr)}, \tag{5}$$

where the first term guarantees the closeness of directions for adjacent segments:

$$E^{(cost)} = -\frac{1}{2}\sum_{ijkl}\delta_{jk}\frac{\cos^m\theta_{ijl}}{r_{ij}r_{jl}}. \tag{6}$$

The notation for segments r_{ij} and angles θ_{ijl} is clear from Fig. 3, m is odd integer. The constraint term in (5) consists of two contributions:

$$E^{(constr)} = -\frac{\alpha}{2}\left[\sum_{ik}s_{ik}s_{kl} + \sum_{jl}s_{ij}s_{jl}\right] + \frac{\beta}{2}\left[\sum_{ij}s_{ij} - N\right]^2. \tag{7}$$

Here, the first term forbids bifurcating tracks, the second one enforces the closeness of N and the number of binary neurons set to 1, α and β are the balancing weight multipliers.

The conventional MFT-dynamics for the energy function (5) allows one to obtain the steady solution in 40–50 iterations. Despite the possibility of the

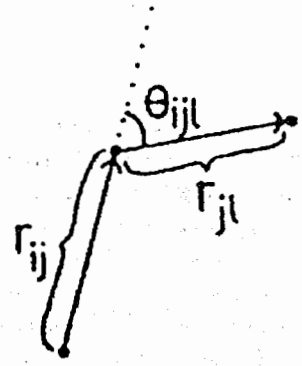

Figure 3: Definition of segment lengths r_{ij} and angles θ_{ijl} between segments.

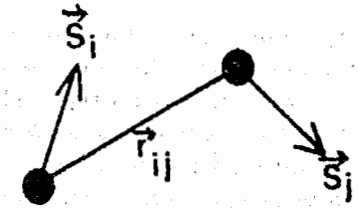

Figure 4: Rotors $\vec{s}_i$ and line segment vectors $\vec{r}_{ij}$ [2].

parallel implementation, the convergence of this scheme is so slow that the total number of computer operations is still larger than for the old tedious histogramming approach [2].

It was, probably, a reason for C. Peterson to propose a new "rotor model" approach [2]. A rotor $\vec{s}_i$, i.e. a unit vector, is put at each measured point. As it can be seen from Fig.4, one can achieve in some vicinity an accordance of a track, its segments (chords) $\vec{r}_{ij}$ connecting two adjacent points, and corresponding rotors by the following energy function:

$$E = -\frac{1}{2}\sum_{ij}\frac{1}{|\vec{r}_{ij}|^m}\vec{s}_i\vec{s}_j - \frac{\alpha}{2}\sum_{ij}\frac{1}{|\vec{r}_{ij}|^m}(\vec{s}_i\vec{r}_{ij})^2. \tag{8}$$

Here, the first term forces neighboring rotors to be close to each other. The second term is in charge of the same between rotors and track segments.

The MFT-dynamics is as follows:

$$\vec{v}_i = \frac{\vec{H}_i}{|\vec{H}_i|}\,\frac{I_1(|\vec{H}_i|/T)}{I_0(|\vec{H}_i|/T)}, \tag{9}$$

where, as before, $\vec{v}_i = \langle\vec{s}_i\rangle_T$.

Strictly speaking, MFT can be applied only to the long range interacting systems [9]. Despite that, applying of the MFT allows one to obtain an appropriate solution for some choice of T_c and α [8], but the number of iterations is

still large enough. In our opinion, the reason of that is the random choice of the initial distribution of rotors $\vec{s}_i$. Another disadvantage of this model is its tendency to straight tracks even if they have large curvatures.

4. Modified rotor track finding

The aim of this paper is the application of the neural networks to the track reconstruction in such discrete detectors as multiwire proportional chambers (MWPC) of experimental setup ARES. Tracks in these MWPCs are the circular arcs (maximum radius is about 60 cm). Explored experimental data consist of three-prong events only. Together with these simplifying features there exists enough noise counts which were not considered elsewhere before. Moreover, we have no such a parallel computing power as CRAY-computer. Therefore, some faster global algorithm is more preferable for us. Carrying out this, we have to take into account the following specific features of MWPC. At each experimental point there is an information not only about its coordinate, but also about the angle with which a particle intersects the chamber. Spectrometer ARES MWPCs are cylindrical but our approach is valid for both cylindrical and parallel chambers. Considering the projections of tracks onto the plane perpendicular to the magnetic field we can treat all tracks as arcs of circles.

Each chamber can be imaginary splitted into cells surrounding signal wires [4, 10]. When a charged particle hits the cell, the wire inside produces a signal. If the track crosses few adjacent cells in the chamber, then all they work. Such a cell group forms a cluster. If one knows that the cluster in a chamber is on, it means for straight tracks that one can detect a possible range of track slopes for this cluster. Physical constraints on tracks (e.g., it should begin in the target and leave the outer plane) also determine the region of the plane where the track may lie (Fig. 5).

This information about the possible track direction can be naturally realized in the rotor model. However, it should be modified to surmount its both disadvantages previously noted: locality and noise sensitivity due to arbitrary initial rotor distribution.

We propose below a "tensor" model of the spin interaction which is free from these disadvantages and can give a global algorithm. Let us choose the two-spin interaction energy in such a way that the neural network has to place vectors $\vec{v}_i = \langle s_i \rangle_T$ tangentially to track and effectively diminish $\vec{v}_j$ for all j that lie apart from the true trajectory. Let two points A_i and A_j lie on the circle (see Fig. 6). Then, tangential vectors are connected by the expression

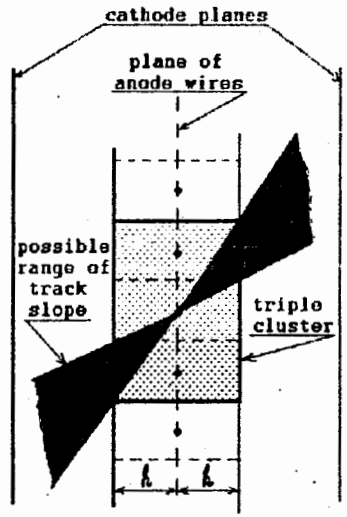

Figure 5: Schematic view of MPWC fragment.

$\beta_i = -\beta_j$, where β_i, β_j are the angles between these vectors and the chord r_{ij}. The energy function for the pair (i, j) can be expressed in a much simpler form than each component of (5):

$$E_{ij} = -\vec{s}_i \vec{s'}_j,$$

where $\vec{s'}_j$ is the mirror reflection of $\vec{s}_j$ relative to r_{ij}, has a minimum for all points placed at one track. Analytically, vector $\vec{s'}_j$ can be obtained by the following transformation:

$$\vec{s'}_j = \begin{pmatrix} \cos 2\varphi_{ij} & \sin 2\varphi_{ij} \\ \sin 2\varphi_{ij} & -\cos 2\varphi_{ij} \end{pmatrix} \cdot \vec{s}_j = w_{ij} \vec{s}_j. \qquad (10)$$

Here φ_{ij} is the angle between the chord r_{ij} and the coordinate axis Ox. The globality of calculating the total energy function is due to the possibility of summarizing all pair energies E_{ij} instead of the nearest points only

$$E = -\frac{1}{2} \sum_{ij} \vec{s}_i w_{ij} \vec{s}_j \qquad (11)$$

without any additional constraint terms like in (7, 8).

MFT-dynamics appears here more naturally than before because of the long range interaction. The local field $\vec{H}_i$ introduced in (1) is created by

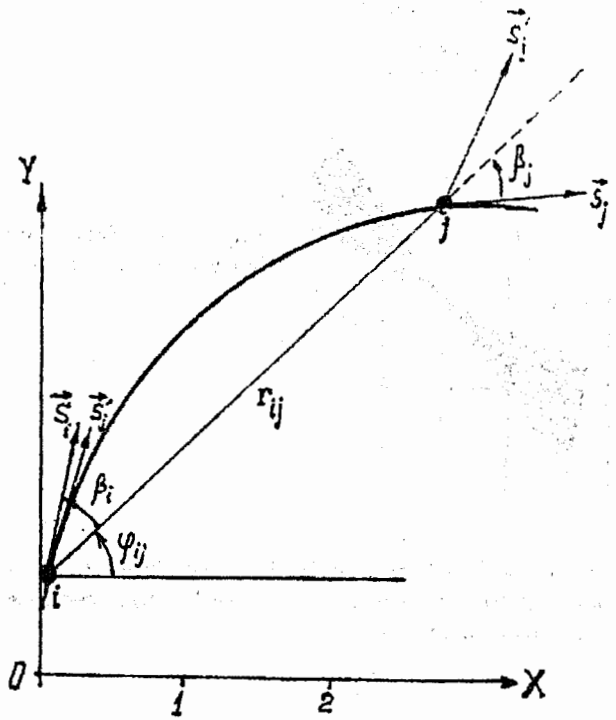

Figure 6: Spin interaction for circular tracks.

all other spins. The equation of dynamics (9) is common for the rotor model approach with our energy function (11). This equation minimizes energy function

$$E = \sum_i \vec{s}_i \vec{H}_i \to \min. \qquad (12)$$

For the better convergency of the MFT-dynamics due to its global character, it is necessary to invent the special robust-type restrictions that are not present explicitly in the conventional MFT-algorithm. During the calculations of mean field H_i in the sum (1) neurons lying "too far" were not included. The "local" radius of the track R_{ij} for any pair (i,j) can be calculated. Then, two parameters were calculated: the average curvature κ and the mean square

deviation ΔR of R_{ij}. Both parameters were controlled to satisfy the cut-off criteria. In particular, for the ARES setup final ΔR should be less than 1 cm. An additional restriction was the prohibition for the neurons from the same chamber to have the mutual interaction. Neurons from different chambers were interacting only one by such one that have maximum contribution to minimizing of the energy function (11).

Besides the variation of the angle β we have one more degree of freedom, namely the coordinate of a spin inside the central cell of the cluster. To control this parameter the simplest dynamics was chosen: MFT-dynamics with zero temperature[2]. According to this dynamics, the coordinate of a spin was assigned to the point where the local field H_i is maximum (if this point will appear outside the central cell, it is moved to its nearest outer border).

The final algorithm reads as follows:

m-th iteration: For each neuron $\vec{v}_i$ we calculate, from the interaction with all other neurons, the mean field $H_i^{(m)}$, the average curvature $\kappa^{(m)}$ and $\Delta R^{(m)}$ with all above restrictions to kill "far" neurons.

$(m+1)$-th iteration: Renew the values of $\vec{v}_i$ according to the rule

$$\vec{v}_i^{(m+1)} = \frac{\vec{H}_i^{(m)}}{|\vec{H}_i^{(m)}|} \frac{I_1(|\vec{H}_i^{(m)}|/T)}{I_0(|\vec{H}_i^{(m)}|/T)},$$

and the coordinate of a neuron inside the center cluster cell.

To define the stopping rule, the explorations were made to calculate the distribution of the iteration number for our neural network. From Fig. 7 it is clear that after 5-th iteration we have, in fact, the neuron grouping along different tracks. So, in our algorithm we stop the general loop for all points after the 5-th iteration and continue looping with neurons interacting only inside the same track to clean tracks from noisy counts due to δ-electrons until $\Delta R < 1$ cm.

5 The results and discussion

Neural network constructed according to the above algorithm was realized on IBM PC/AT-386 and tested on the set of 10^4 real three-prong events from ARES-spectrometer. Data were obtained in experiments during the search for

[2]It is possible due to the small range of the parameter, therefore only one minimum can exist.

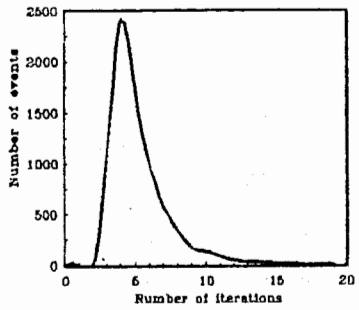

Figure 7: The distribution of the number of iterations.

the forbidden decay $\mu^+ \to e^+e^+e^-$ [11] and the study of the rare decay $\pi^+ \to e^+\nu_e e^+e^-$ [12]. Only 10 chambers were used in this experiments. That means that every track contains on the average 10 experimental points (clusters).

Fig. 8 shows one typical example of initial configurations of the neural network, and the final result of its evolution. It is necessary to point out that the construction of the proper initial rotor configuration was possible due to applying the cellular automaton [4] on the stage of preliminary processing. This automaton removed almost all noisy points (denoted by crosses) and fulfilled grouping of experimental points according to different tracks[3]. As one can see, in the case of nearby going tracks, our neural network performed their good resolution.

Fig. 9 shows the distribution of the initial number of points in an event (dotted line) and the final distribution after the neural network evolution. Since each track consists of not more than 10 points, it easy to note that after the full noise reduction we found one-track events (less than 10 points), two-track (10 – 20 points), three-track (20 – 30) and even four-track (more than 30 points) events.

The distribution of the number of points on the track is shown on Fig. 10. Dotted line shows the distribution of rejected points in a group recognized as a track and filled line shows the number of the remaining points. One can see that the average point number per track is 8 – 9 points.

In conclusion, we can say that the global character of the applied algo-

[3]However, in this analysis, having a goal to study the reliability of our neural network, we did not use these previous results of neuron grouping.

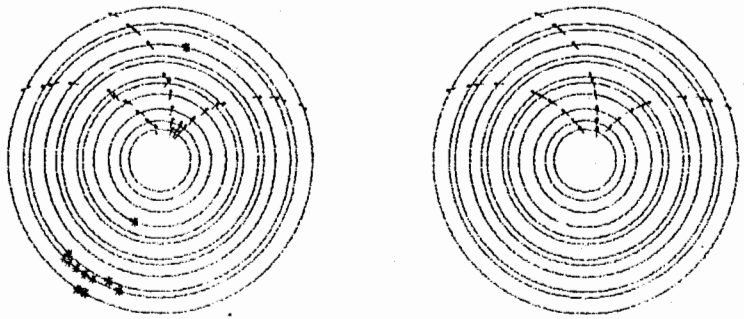

Figure 8: An example of the initial and final state of the neural network.

rithm together with certain robust restrictions made it possible to achieve the fast convergency of the constructed neural network and good track resolution. Moreover, this algorithm can work in the case of many-prong events without any special complications. We are convinced that our algorithm can also be applied to the case of an arbitrary magnetic field configuration. Only expression (10) should be modified to satisfy the equation of the particle motion in this field.

References

[1] B. Denby, Comp. Phys. Comm., **49**, 429 (1988).

[2] C. Peterson, Preprint LU TP 90 – 6, Lund (1990).

[3] C. Peterson, Nucl. Instr. & Meth., **A 279**, 537 (1989).
G. Stimphl-Abele, L. Garrido, Comp. Phys. Comm., **64**, 46 (1991).

[4] A. Glazov et. al., JINR Commun. E10-91-507, Dubna (1991).

[5] V. Baranov et. al., Nucl. Instr. & Meth., **B 17**, 438 (1986).

[6] G. Hopfield, D. Tank, Biol. Cyb. **52**, 141 (1985).

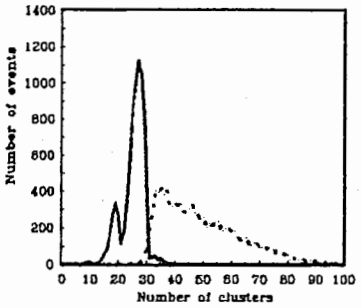

Figure 9: The distribution of the point number in an event before (dotted line) and after the network evolution.

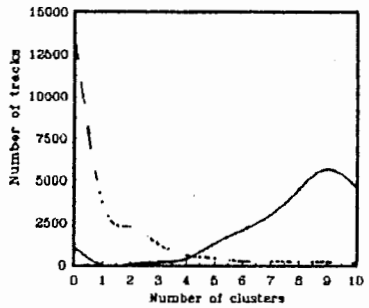

Figure 10: The distribution of rejected points in a group recognized as a track (dotted line) and that of remaining points (filled line).

[7] S. Kirkpatrik, D. Sherrington, Phys. Rev. **B 17**, 4384 (1978).

[8] C. Peterson, B. Södberg, Int. J. of Neural Syst. 1, 3 (1989).

[9] N. Bogolubov, Jr., A Method for Studying Model Hamiltonians. Oxford — New York, Pergamon Press (1972).

[10] N. Chernov et. al., JINR Commun. E10-91-361, Dubna (1991).

[11] V. Baranov et. al., J. Phys. G: Nucl. Part. Phys. **17**, 57 (1991).

[12] V. Baranov et. al., Jad. Phys. (to be published).

Received by Publishing Department
on August 17, 1992.

Глазов А.А. и др. E10-92-352

Восстановление треков в дискретных детекторах
при помощи нейронных сетей

На базе применения нейронных сетей к распознаванию треков проведен учет специфики таких дискретных детекторов, как многопроволочные пропорциональные камеры, приводящий к модификации так называемой роторной модели нейронной сети. Энергетическая функция сети в этой модификации содержит только один стоимостной член, что значительно ускоряет вычисления. Обеспечивается это введением селекции нейронов с помощью простых геометрических и энергетических критериев. Кроме того, применение клеточного автомата для предварительного отбора данных позволило создать начальную конфигурацию сети с энергией, близкой к своему глобальному минимуму. Алгоритм был протестирован на 10 тыс. реальных трековых событий, полученных на спектрометре АРЕС, и показал хорошие результаты, в том числе устойчивость к шумам и надежное разрешение близко лежащих треков.

Работа выполнена в Лаборатории вычислительной техники и автоматизации ОИЯИ.

Сообщение Объединенного института ядерных исследований. Дубна 1992

Glasov A.A. et al. E10-92-352

Track Reconstruction in Discrete Detectors
by Neural Networks

On the basis of applying neural networks to the track finding problem the investigations are made according to the specific properties of such discrete detectors as multiwire proportional chambers. These investigations result in the modification of the so-called rotor model in a neural network. The energy function of a network in this modification contains only one "cost" term. This speeds up calculations considerably. The reduction of the energy function is done by the neuron selection with the help of simple geometrical and energetical criteria. Besides, the cellular automata were applied to preliminary selection of data that made it possible to create an initial network configuration with the energy closer to its global minimum. The algorithm was tested on 10^4 real three-prong events obtained from the ARES-spectrometer. The results are satisfactory including the noise robustness and good resolution of nearby going tracks.

The investigation has been performed at the Laboratory of Computing Techniques and Automation, JINR.

Communication of the Joint Institute for Nuclear Research. Dubna 1992