E10-2005-216

S. G. Dmitrievsky, Yu. A. Gornushkin, G. A. Ososkov

NEURAL NETWORKS, CELLULAR AUTOMATA, AND ROBUST APPROACH APPLICATIONS FOR VERTEX LOCALIZATION IN THE OPERA TARGET TRACKER DETECTOR

Submitted to «Нейроинформатика»

Дмитриевский С. Г., Горнушкин Ю. А., Ососков Г. А. E10-2005-216
Применение нейронных сетей, клеточных автоматов и робастного подхода для локализации вершины взаимодействия в трековой системе целеуказания детектора OPERA

Описан нейросетевой подход для восстановления вершины нейтринного взаимодействия с помощью мишенного трекера в эксперименте OPERA. Использована прямоточная нейронная сеть (NN) с алгоритмом обратного распространения ошибки. Для минимизации энергетического функционала NN использован метод сопряженных градиентов. Предварительная обработка данных проводилась с помощью клеточного автомата. Для определения мюонного трека применялось преобразование Хафа, а для восстановления оси адронного ливня использовался робастный метод подгонки. Сравнительный анализ предлагаемого метода с методами, основанными на использовании нейросетевого пакета SNNS, показал их одинаковую эффективность. Предлагаемый подход находится в стадии дальнейшего развития.

Работа выполнена в Лаборатории ядерных проблем им. В. П. Джелепова ОИЯИ.

Препринт Объединенного института ядерных исследований. Дубна, 2005

Dmitrievsky S. G., Gornushkin Yu. A., Ososkov G. A. E10-2005-216
Neural Networks, Cellular Automata, and Robust Approach Applications for Vertex Localization in the OPERA Target Tracker Detector

A neural-network (NN) approach for neutrino interaction vertex reconstruction in the OPERA experiment with the help of the Target Tracker (TT) detector is described. A feed-forward NN with the standard back propagation option is used. The energy functional minimization of the network is performed by the method of conjugate gradients. Data preprocessing by means of cellular automaton algorithm is performed. The Hough transform is applied for muon track determination and the robust fitting method is used for shower axis reconstruction. A comparison of the proposed approach with earlier studies, based on use of the neural network package SNNS, shows their similar performance. The further development of the approach is underway.

The investigation has been performed at the Dzhelepov Laboratory of Nuclear Problems, JINR.

Preprint of the Joint Institute for Nuclear Research. Dubna, 2005

INTRODUCTION

The OPERA experiment [1] is designed for a direct observation of ν_τ appearance in the CNGS long baseline beam (from CERN to Gran Sasso Laboratory) as a result of $\nu_\mu \to \nu_\tau$ oscillation.

OPERA exploits nuclear emulsions as very high resolution tracking devices for the direct detection of τ leptons produced in the charge current (CC) interaction of the ν_τ with matter of the detector. The target part of the detector (see Fig. 1) consists of 62 target brick walls (photoemulsion layers, interlaced with led layers,

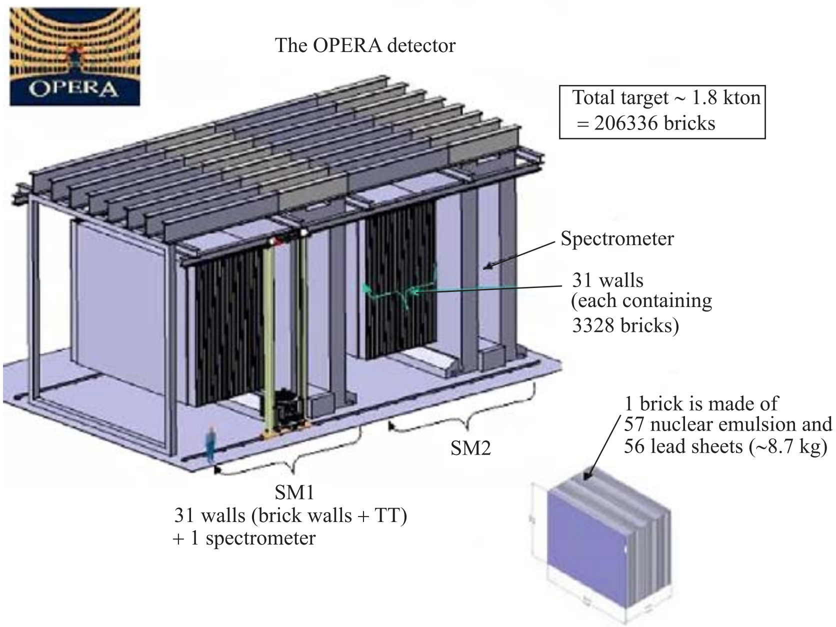

Fig. 1. General view of the OPERA setup

are organized in bricks). Each wall consists of $\sim$ 3300 bricks. It is accompanied by two planes (X–Y) of electronic Target Tracker (TT) detectors. The TT planes consist of 256 scintillator strips, each is $\sim$ 7 m long. The strips are read out by WLS fibers optically coupled to photodetectors on both ends of a strip. The TT mainly serves for a location of the event vertex's position, i. e. for a Brick Finding (BF) — a procedure of a target brick identifying where a neutrino interaction occurred.

A neural network (NN) approach was successfully employed in [2, 3] for the purpose of the BF. The NN used there had been generated by the Stuttgart NN Simulator (SNNS) [10] with feed-forward, standard back-propagation option. Input data were previously filtered by a kind of Hough transform method. The NN was trained separately on different classes of events. The BF strategy proposed in [2, 3, 4] was also different for muonic and non-muonic events, although in both cases it uses the idea of intersection of the previously found brick wall with either muon track or hadronic shower axis. In case of τ decay the total BF efficiency achieved in [2, 3] was 72.4% for muonic, 68.1% for hadronic, and 72.5% for electronic events.

The goal of this work was to develop a BF program based on our NN code and on some other data handling algorithms developed by the JINR's OPERA group. As a first step of validation of our BF approach we tried to reproduce the results of the previous studies [2, 3].

In our study we also split the BF task in two steps: we identify, first, the vertex wall, i. e. look for Z position of the vertex, and then, use it to locate a brick with the vertex of the given event.

1. EVENT CLEANING WITH THE HELP OF A CELLULAR AUTOMATON

If one would try to identify the brick wall, containing neutrino interaction vertex, straightforwardly, as a wall in front of the first TT plane, containing some hits, it would often give a wrong result because of a presence of back-scattering (BS) particles. BS is particularly dangerous since it is correlated with the vertex. Since for the vertex reconstruction the tracking information is most useful, isolated hits (the product of neutron or gamma interactions with the detectors, natural radioactivity background, PMT noise, etc.) are misleading in most of the cases.

So, to facilitate the vertex location some preliminary event cleaning off isolated hits is needed. We propose the procedure, based on the cellular automaton (CA) approach [5].

A cellular automaton can be regarded as a simplified local form of neural networks. It is a dynamical system that evolves in discrete, usually two-dimensional, space consisting of binary cells.

The evolution rules are local, i. e. the system dynamics is determined by an unchanged set of rules, for example, a table, in accordance with which the new state of a cell is calculated on the basis of the states of the nearest neighbors surrounding it. It is important that this change of states is made simultaneously and in parallel, and the time proceeds discretely. An example of CA realization is the famous «Life» game.

We tested a set of different survival rules in order to eliminate more efficiently or, at least, reduce the number of disconnected hits which can distort the topology of the events. The results can be seen in Table 1. Here BS $= 0$ means that there

Table 1. Tuning of CA rules

			BS = 0	**BS = 1**	**BS = 2**	**CL**
Without cleaning			45.8 %	43.1%	11.1 %	0.0 %
With cleaning						
NAP	**DR**	**NN**				
2	5	1	64.4 %	32.6 %	3.0 %	0.2 %
2	4	1	65.5	31.9	2.6	0.2
2	3	1	67.1	30.8	2.1	0.3
2	2	1	69.8	28.8	1.4	0.5
2	5	2	64.0	34.1	1.9	0.3
2	4	2	65.6	32.9	1.5	0.4
2	3	2	67.7	31.2	1.0	0.5
2	2	2	71.0	28.4	0.6	1.5
4	5	1	59.8	34.7	5.5	0.1
4	4	1	60.6	34.3	5.1	0.1
4	3	1	61.7	33.8	4.6	0.2
4	2	1	63.6	32.6	3.8	0.4
4	1	1	67.2	30.2	2.6	2.1
4	5	2	60.9	34.3	4.8	0.2
4	4	2	62.1	33.9	4.4	0.2
4	3	2	63.8	32.6	3.7	0.3
4	2	2	66.6	30.6	2.7	0.6
4	1	2	72.2	26.3	1.5	0.3

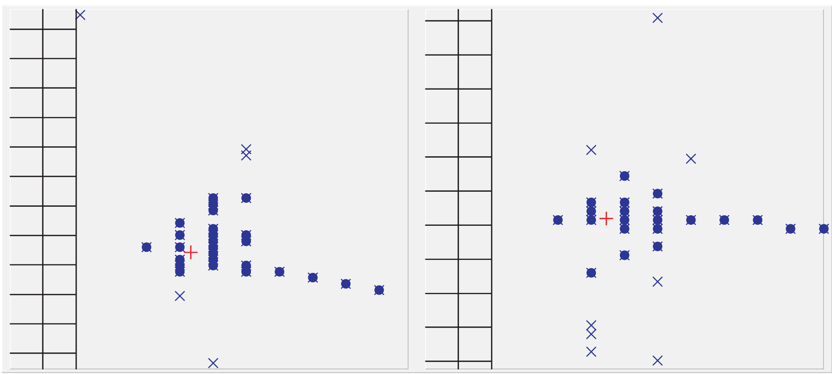

Fig. 2. An example of CA selected strips for a ν_μ CC event. Cross «+» marks the vertex real position

is no BS particles, BS $= 1$ — there is one hitted TT wall before the vertex wall, BS $= 2$ — there are two hitted walls before the vertex wall. NAP — number of adjacent walls in which neighbors are being found, DR — distance along an adjacent wall measured in strip widths. NN — number of neighbors needed for a hit survival. CL — cleaning losses.

Finally, the cleaning method consists in removing of each hit that has none of 14 nearest neighbors in two adjacent walls. In Fig. 2 the selected CA hits are shown with respect to all hits for one of the ν_μCC events.

2. CLASSIFICATION OF NEUTRINO EVENTS BY THEIR TOPOLOGY IN TT

The CNGS ν_μ beam is optimized for ν_τ appearance search. The number of neutrino interactions expected for 5 years of operation is more than 32000. The number of interacting ν_τ is expected to be 240. Thus, overwhelming majority of neutrino events will be the ν_μ interactions. For each neutrino interaction the brick, containing the vertex, has to be identified, extracted and the emulsions have to be analyzed. A brick removed from the brick wall for emulsion analysis is not coming back again. So, to preserve a mass of the OPERA detector and to decrease the scanning load it is important to identify a vertex brick of any neutrino event with the highest efficiency. Only after comprehensive scanning analysis a conclusion on a type of the neutrino event can be made.

The neutrino interaction events can be separated in a few classes according to their topology in the TT detector. The BF procedure can be optimized differently for those classes depending on presence of muon track and hadronic showers.

CC muon neutrino interactions can be either quasi-elastic (QE) or deep inelastic scattering (DIS) events. The first event kind is characterized by a long muon track while in the second case the muon track is accompanied by a hadronic shower.

Neutrino events were separated in three categories depending on the number of the hitted walls and on the mean number of hits per TT plane. This separation is motivated by a fact that the NN performance used to locate the vertex wall is improved by separating events with small shower development from those of important one. The first kind of the events more likely corresponds to the QE interactions and the second one to DIS, where the BS is quite frequent. The three classes of the events are defined as follows:

1) Events with one or two hitted TT walls;

2) Events with more than two hitted walls and a mean number of hits/plane less than 2.5;

3) Events with more than two hitted walls and a mean number of hits/plane more than 2.5.

The procedure of BF is described below only for the second and the third classes of the events because for the first class some other procedure should be applied.

3. MUON TRACK DETERMINATION BY MEANS OF THE HOUGH TRANSFORM

In ν_μ CC and $\tau \to \mu$ events a muon is produced which can be recognized by its long track. So we can try to determine the neutrino interaction vertex using a muon track information.

In case of the muonic events after CA filtering we use the method of variable slope histograms (VSH) for a muon track recognition.

The VSH method [6] is a particular case of the Hough transform [7] for straight lines revealing. The idea of the method consists in fragmentation of an inspected region by narrow parallel bands in every of which the number of hits is calculated (see Fig. 3). A slope of the bands is gradually changed from $\alpha_{\min}$ to $\alpha_{\max}$. When one of such slopes coincides with some of tracks, it would produce a maximum in the histogram corresponding to this slope.

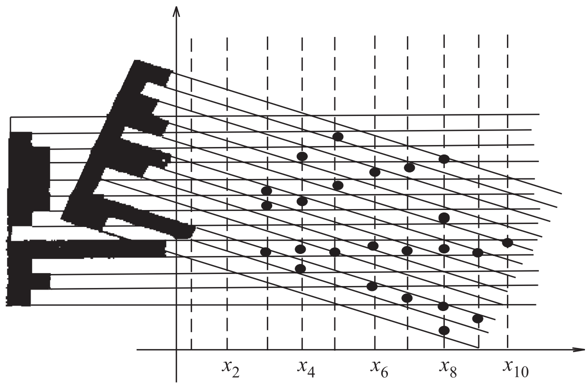

Fig. 3. Graphic example of the VSH method principle

The method of projection is used for bins infilling. All hits are projected onto Y axis on a fixed angle by rule $Y_{\rm pr} = y - x\alpha$ (see Fig. 4) and values of corresponding histogram bins are increased. Starting from $\alpha_{\min}$ we increase gradually the projection angle in the following way: for the kth direction α_k is calculated as $\alpha_k = \alpha_{\min} + k\Delta y / X_{\max}$.

The criterion of the muon track definition is that a bin of histogram with the maximum value must contain at least 10 counts. When the maximal bin is found, the directions corresponding to 95% of maximum are fixed. If the number of

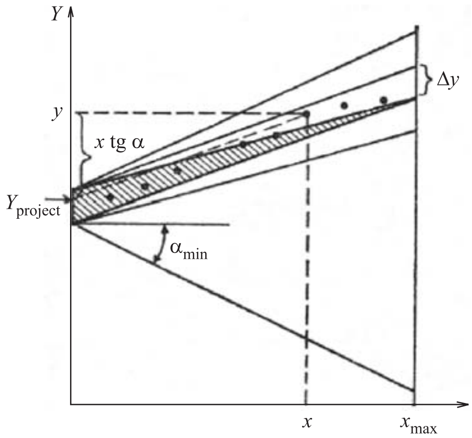

Fig. 4. The method of projection

such directions is less than 3 their parameters are used to determine then a vertex brick. Otherwise only a shower axis is reconstructed. In Fig. 5 the muon track found by the VSH method is shown.

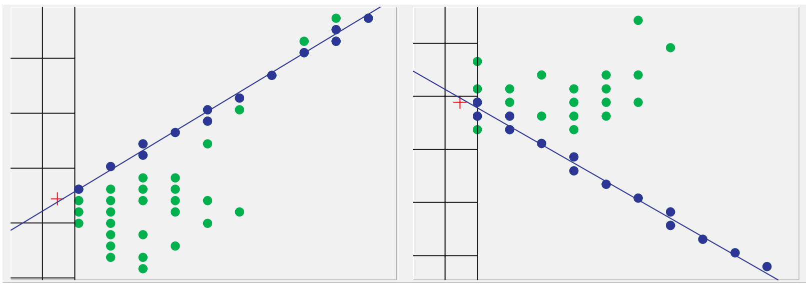

Fig. 5. Muon track found by the VSH method

4. PRINCIPAL SHOWER AXIS RECONSTRUCTION USING A ROBUST LINE-FITTING METHOD

The TT detector has a pitch of 26 mm and it is difficult to single out distinct tracks near the vertex of the event. In that case the reconstruction of a shower axis may be more useful for finding a general direction to the vertex.

For this purpose it is necessary to employ some line-fitting method. However, usual Least Square Method (LSM) may fail in case of our events because of presence of points-outliers or/and strong contamination. In both cases the crucial LSM assumption of residual normality is violated. So we propose to replace the Least Square functional

$$L(\mathbf{p}) = \sum_i \epsilon_i^2 \tag{1}$$

by a functional

$$L(\mathbf{p}, \sigma) = \sum_i \rho(\epsilon_i), \tag{2}$$

where $\mathbf{p}$ is the vector of parameters, ϵ_i are residuals and $\rho(\epsilon)$ is a compact contribution function.

The functional (2) minimization with respect to its parameters can be performed by solving the equation $\dfrac{\partial L(p)}{\partial p} = \sum_{i=1}^{n} \dfrac{\partial \rho(\epsilon_i)}{\partial \epsilon_i} \dfrac{\partial \epsilon_i}{\partial p} = 0$. As it can be demonstrated [8], that equation by denoting $w(\epsilon) = \dfrac{1}{\epsilon} \dfrac{\partial \rho(\epsilon)}{\partial \epsilon}$ can be modified to the form $\dfrac{\partial L(p)}{\partial p} = \sum_{i=1}^{n} w(\epsilon_i) \dfrac{\partial \epsilon_i}{\partial p} \epsilon_i = 0$, which is similar to the normal LSM equations, but with the replacement of numerical weight coefficients by the weight function $w(\epsilon)$ to be recalculated at each step of the iterative procedure.

It gives a weight to each measured point depending on its distance ϵ from the fitted curve. The weight function must quickly decrease with growing the residual ϵ.

To determine a principal shower axis we use the robust fit of a straight line to all hits excluding those of the found muon track, but including hits that belong to the first 3 walls with the energy deposition exceeding the average hit amplitude (to take into account double hits in some strips close to the vertex). We also take into account amplitude information (see Fig. 6) in such a way that a robust weight of each point changes accordingly to the energy deposited in the corresponding hit. Equation (3) describes a 2D weight function which is shown in Fig. 7.

$$w(E, d, \sigma) = A \exp\left(-\frac{d^2}{2\sigma^2}\right), \tag{3}$$

where $A = \sqrt[4]{\dfrac{E}{E_{\max}}}$ and σ decreases gradually during an iteration process.

The robust fit starts from an initial approximation, which is a line passing through the center of gravity of all hits and parallel to Z axis.

In Fig. 8 the shower axis reconstructed by a robust fitting method is shown for one of the $\tau \to h$ events.

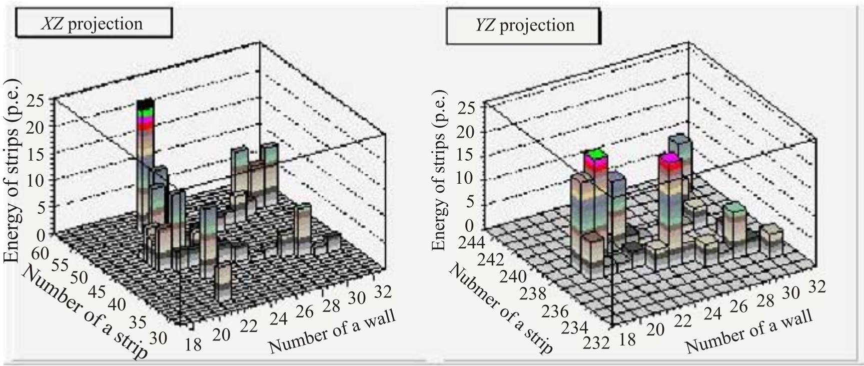

Fig. 6. Hit amplitudes

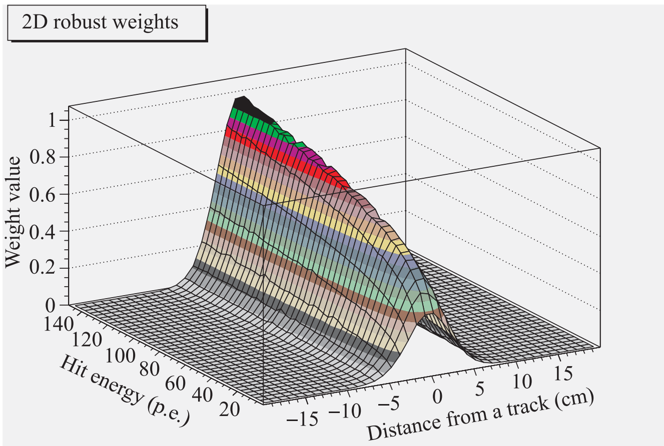

Fig. 7. 2D weight function

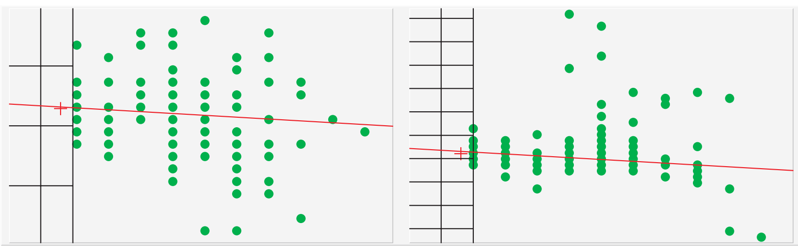

Fig. 8. Axes for hadronic shower reconstructed by the robust fit

5. VERTEX WALL DETERMINATION

For the purpose of the vertex wall determination we use a neural network approach based on a multilayer perceptron (MLP).

A typical MLP network consists of a set of source nodes forming the input layer, one or more hidden layers of computation nodes (artificial neurons), and an output layer of nodes (see Fig. 9). Each layer is fully connected to the next

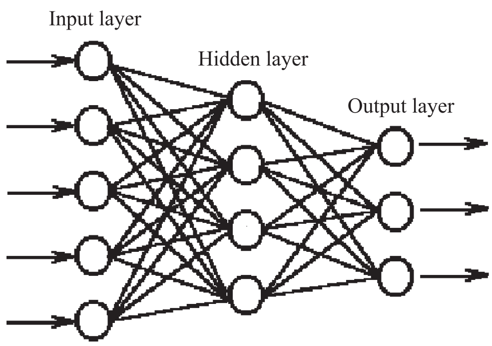

Fig. 9. MLP structure

one. The multiple real-valued input signal propagates through the network layer-by-layer according to a set of weights of neurons putting each subsequent output through some nonlinear activation function $g(x)$ (Fig. 10).

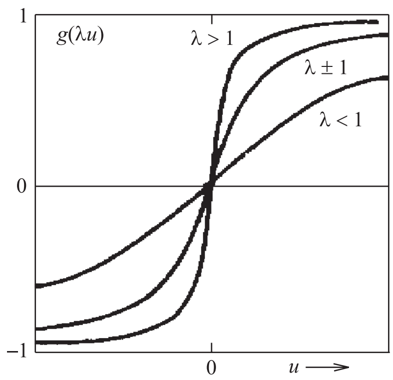

Fig. 10. Activation function

MLPs are still the most usable NN type in physics due to possibilities of their training by Monte-Carlo data and nonexpensive hardware implementation.

We use MLP with standard back propagation training algorithm [9]. The energy functional minimization of the network is performed by the method of conjugate gradients (CG). Our first goal was to reproduce the results of [2, 3] making use of our own NN code and we started with about the same input parameters. The numbers of neurons in the input and hidden layers are equal to 14. Input variables estimated for the first 3 hitted TT walls are as follows:

- Total amplitude of hits in a particular wall;
- Number of the hits in the wall;
- Dispersion of the hits in the wall;
- The mean distance of the hits in the wall with respect to the event shower axis.

In addition to those variables, ratios of energy in the next wall with respect to the previous one, E2/E1 and E3/E2, are also included to the input parameters to train the NN. Their distributions for different kinds of BS are shown in Fig. 11.

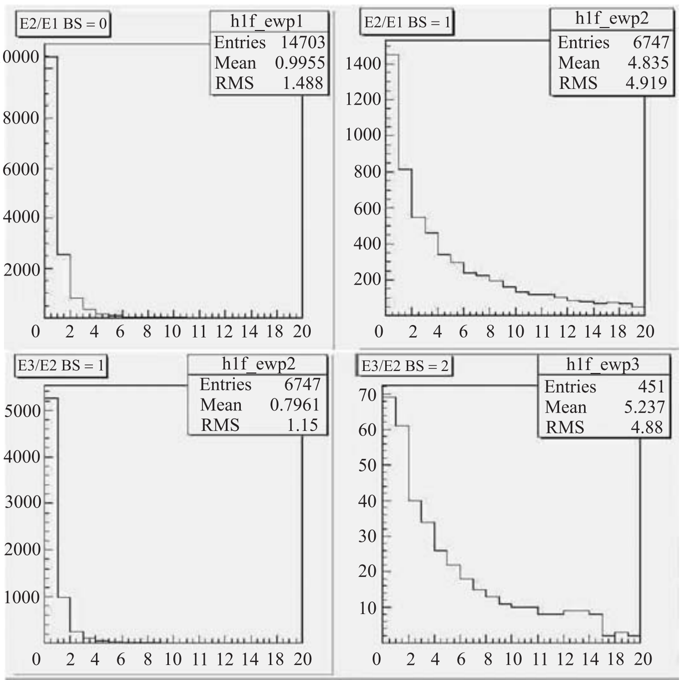

Fig. 11. E2/E1 and E3/E2 ratio distributions

As can be seen in Fig. 12 the most significant input parameters, for example, for $\tau \to \mu$ type of event are mean distances of hits in the first three walls with respect to the event shower axis.

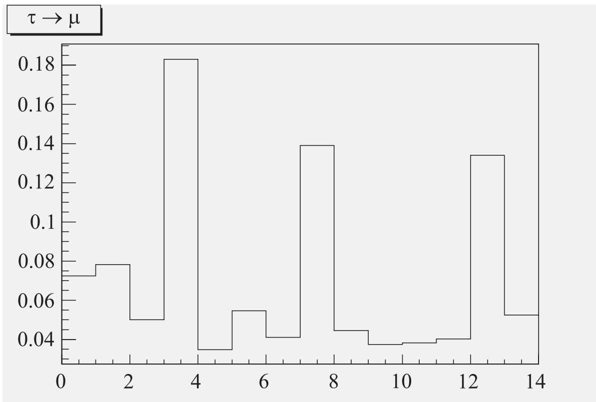

Fig. 12. Relative significance of the input parameters

6. MLP(CG) vs. SNNS COMPARISON

Stuttgart Neural Network Simulator (SNNS) package [10] includes vast spectrum of different NN architectures and training algorithms and it is intended for solving wide range of problems. Our net (or the MLP(CG)) was elaborated especially for the OPERA data analysis, therefore it is simple, compact and can be directly built in OPERA software. We found it to be more flexible and easy to use than SNNS with similar performance.

The comparative study of MLP(CG) and SNNS with analogous training and minimization algorithm shows that corresponding mean square errors (MSE) after 8000 epochs are approximately equal to each other. Comparison results are shown in Table 2.

Table 2. MLP(CG) vs. SNNS

	MLP(CG)		**SNNS**	
	Efficiency	**MSE**	**Efficiency**	**MSE**
$\tau \to \mu$ 2cl.	84.4 ± 2.9	0.249	84.5 ± 3.3	0.247
$\tau \to \mu$ 3cl.	84.1 ± 3.4	0.199	84.2 ± 3.2	0.198
ν_μ CC 2 cl.	85.2 ± 3.6	0.224	85.1 ± 3.1	0.228
ν_μ CC 3 cl.	84.9 ± 2.6	0.207	84.9 ± 2.8	0.209
ν_μ NC 2 cl.	79.9 ± 4.6	0.269	78.8 ± 3.8	0.285
ν_μ NC 2 cl.	83.2 ± 3.5	0.226	83.2 ± 3.4	0.225
$\tau \to e$ 3 cl.	89.5 ± 3.6	0.178	89.8 ± 3.8	0.174
$\tau \to h$ 3 cl.	84.4 ± 3.1	0.197	85.3 ± 2.7	0.187

7. VERTEX BRICK IDENTIFICATION

After the vertex wall is selected by the NN, we use its position and a muon track or a shower axis parameters to determine x–y coordinates of the vertex brick in the wall.

The distributions of vertex residual for $\tau \to \mu$ and ν_μCC events (approximated by the sum of two Gaussians for comparison) can be seen in Fig. 13 and Fig. 14, correspondingly. The sigma values are comparable with that given in [4].

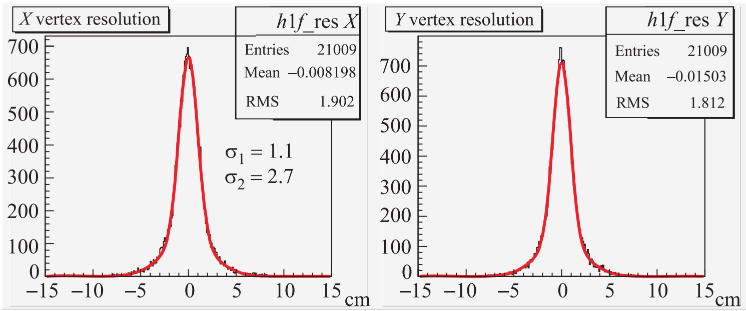

Fig. 13. $\tau \to \mu$ vertex resolution

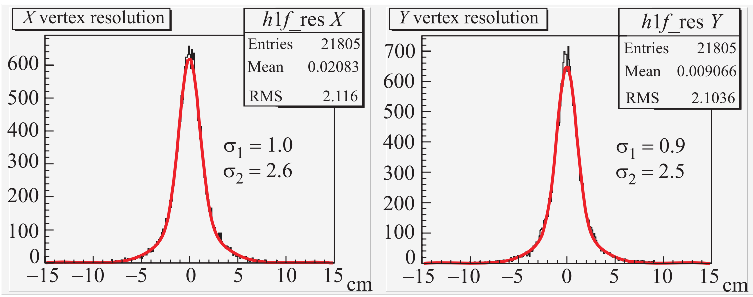

Fig. 14. ν_μ CC vertex resolution

8. PRELIMINARY RESULTS

In our studies we were restricted by the officially MC statistics which is not probably sufficient in some cases.

Wall finding (WF) and brick identification (BI) efficiencies calculated for available simulated data sets are presented in Table 3. Total BF efficiency is a multiplication of those two. The uncertainty of these results is about 5%. For a comparison purpose efficiency for various types of ν_τ events obtained in [2, 3] are shown in Table 4.

Table 3. Wall finding (neural network), brick identification, and total BF efficiency

	WF	BI	BF
$\tau \to \mu$	84.3	78.2	66.0
$\tau \to e$	89.5	75.8	67.8
$\tau \to h$	84.4	75.4	63.6
ν_μ CC	85.1	80.9	68.8
ν_μ NC	82.1	67.1	55.1

Table 4. Corresponding efficiencies obtained earlier

	WF	BI	BF
$\tau \to \mu$	88.6	81.7	72.4
$\tau \to e$	86.1	84.2	72.5
$\tau \to h$	83.8	81.3	68.1

9. CONCLUSION AND OUTLOOK

We have created the BF program, based on some algorithms usually applied for similar purposes.

The wall finding efficiency for the studied neutrino reactions is found to be the same as in the previous works, however, the current BF efficiency is still lower. It may be because of lack of MC statistics, of some difference in MC samples, etc., however, it is clear that a further optimization of the algorithms is necessary.

As the next steps, we plan

• to work on a further algorithm optimization;

• to implement the BF strategy with identification not just one but few most probable vertex bricks;

• to make better separation of neutrino events to classes;

• to generate sufficient statistics of realistic MC events making use of the information of calibration and commissioning of the TT.

Acknowledgements. The authors express their gratitude to A. Krasnoperov for MC simulations, to A. Stadnik for his valuable help with developing MLP software, and to Marcos Dracos for valuable discussions and support of this work.

REFERENCES

1. *Guler M. et al.* CERN/SPSC 200-028, SPSC/P318, LNGS P25/2000, July 10, 2000.
2. *Laktineh I.* Brick Finding Efficiency in Muonic Decay Tau Neutrino Events. OPERA Internal Note, 21 January, 2002.
3. *Laktineh I.* Brick Finding Efficiency of No-Muon Tau Neutrino Events in OPERA. OPERA Internal Note, 18 November, 2002.
4. *Heritier C., Autiero D.* Status of Brickfinding. OPERA Meeting, Napoli, October 23–25, 2003.
5. *Glazov A. et al.* Filtering Tracks in Discrete Detectors Using a Cellular Automaton // Nucl. Instr. Meth. A. 1993. V. 329. P. 262–268.
6. *Nikitin V. A., Ososkov G. A.* Automatization of Physical Experiment Measurement and Data Processing. M.: MSU, 1986.
7. *Hough P. V. C.* A Method and Means for Recognizing Complex Patterns. US Patent: 3, 069, 654, Dec. 1962.
8. *Ososkov G. A.* Elastic Arm Methods of Data Analysis as a Robust Approach // Tatra Mount. Math. Publ. 2003. V. 26. P. 291–306.
9. *Khanna T.* Foundation of Neural Networks. N.Y.: Addison–Wesley, 1989.
10. *Zell A. et al.* SNNS User Manual, Version 4.1. Report N 6/95. SNNS is (Copyright) 1990-95 SNNS Group, Institute of Parallel and Distributed High-Performance Systems (IPVR), University of Stuttgart, Germany.

Received on December 26, 2005.

Корректор *Т. Е. Попеко*

Подписано в печать 24.04.2006.
Формат $60 \times 90/16$. Бумага офсетная. Печать офсетная.
Усл. печ. л. 1,18. Уч.-изд. л. 1,67. Тираж 290 экз. Заказ № 55317.

Издательский отдел Объединенного института ядерных исследований
141980, г. Дубна, Московская обл., ул. Жолио-Кюри, 6.
E-mail: publish@pds.jinr.ru
www.jinr.ru/publish/