Available online at www.sciencedirect.com
SCIENCE DIRECT®

Nuclear Instruments and Methods in Physics Research A 559 (2006) 103–105

NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A

www.elsevier.com/locate/nima

The FEDRA—Framework for emulsion data reconstruction and analysis in the OPERA experiment

V. Tioukov[a,*], I. Kreslo[b], Y. Petukhov[c], G. Sirri[d]

[a]*Istituto Nazionale di Fisica Nucleare, Sezione di Napoli, Italy*
[b]*LHEP, University of Bern, Switzerland*
[c]*IHEP, Protvino, Russia*
[d]*INFN, Bologna, Italy*

Available online 22 December 2005

Abstract

OPERA is a massive lead/emulsion target for a long-baseline neutrino oscillation search. More than 90% of the useful experimental data in OPERA will be produced by the scanning of emulsion plates with the automatic microscopes. The main goal of the data processing in OPERA will be the search, analysis and identification of primary and secondary vertices produced by neutrino in lead-emulsion target.

The volume of middle- and high-level data to be analysed and stored is expected to be of the order of several Gb per event. The storage, calibration, reconstruction, analysis and visualization of this data is the task of FEDRA system written in C^{++} and based on ROOT framework. The system is now actively used for processing of test beams and simulation data. Several interesting algorithmic solutions permits us to make effective code for fast pattern recognition in heavy signal/noise conditions. The system consists of the storage part, intercalibration and segments linking part, track finding and fitting, vertex finding and fitting and kinematical analysis parts. Kalman Filtering technique is used for tracks & vertex fitting. ROOT-based event display is used for interactive analysis of the special events.

PACS: 29.40.Rg; 29.40.Gx; 29.85.+c

Keywords: Nuclear emulsion; Pattern recognition; Analysis framework

1. Emulsion data flow and data analysis

The basic unit of the OPERA [1] detector is so-called *emulsion brick*—based on the concept of the emulsion cloud chamber (ECC) sandwich consisting of 56 lead and 57 emulsion plates packed together. Charged particles passed through the brick leave the *latent image* in the emulsion media. After the emulsion development dark silver grains of about 0.6 μm in diameter are remaining on the path of the particles. The density of the grains along the *M.I.P.* path is about of $32/100\,\mu\mathrm{m}^{-1}$. It means that in the emulsion layer of 45 μm thick we have in average about 14 consecutive grains permitting us to reconstruct the three-dimensional track vector "microtrack" per each emulsion side (Fig. 1). Grains location and microtracks reconstruction is performed by the online scanning systems [2,3] with the usage of computers-driven microscopes. This information (microtracks) is the input raw level data for the off-line analysis (Fig. 2). The final goal of the off-line processing is the full reconstruction of three-dimensional event topology, definition of the kinematical parameters and the event analysis. Typically one brick contains the most of information necessary for one event analysis and can be considered as a quasi-independent detector.

2. Framework for emulsion data reconstruction and analysis (FEDRA) structure and functionality

FEDRA is a program dedicated for off-line processing of emulsion data acquired by the automatic scanning

*Corresponding author.
E-mail address: valeri@na.infn.it (V. Tioukov).

0168-9002/$ - see front matter
doi:10.1016/j.nima.2005.11.214

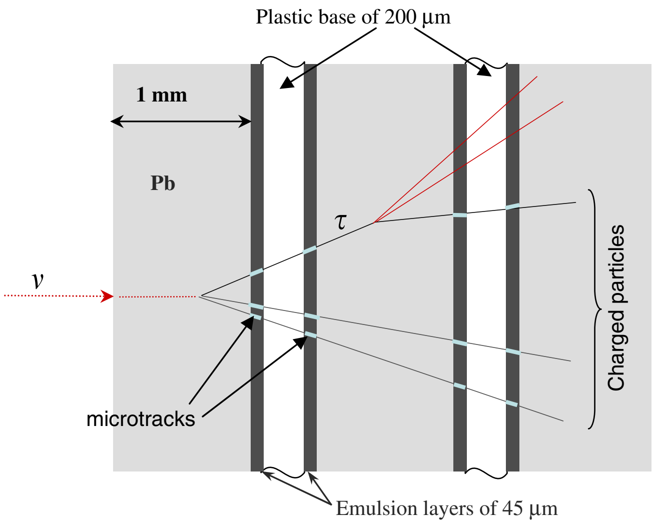

Fig. 1. Fragment of the OPERA brick illustrating the ECC structure and the event topology. What we really can see after scanning the microtracks.

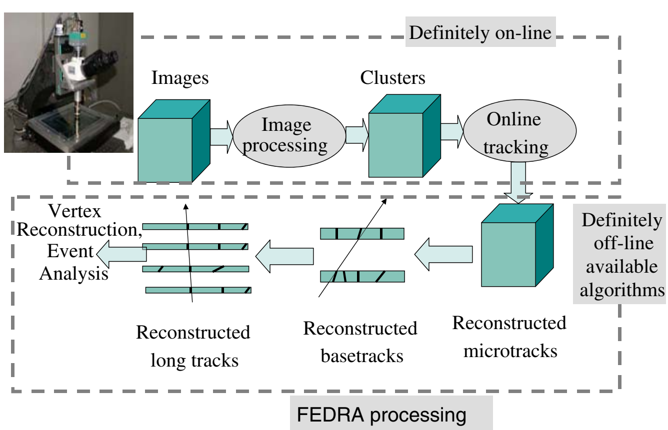

Fig. 2. Scheme of the emulsion data flow with the main processing steps from the image taking up-to the final event analysis.

system. Designed in a very modular way it allows to make off-line tests reconstruction algorithms in the emulsion data and perform the complete analysis of scanned or simulated into ECC events starting from the microtrack level.

It is organized as a set of shared libraries of C^{++} classes and routines and makes an intensive use of ROOT tools and structures. The input/output part is very flexible and allows the use of input data in ROOT or ORACLE or ASCII format: these data may come from the different scanning systems or from the Monte-Carlo simulation.

The main program modules cover the following tasks:

- Raw data storage (libEdb): set of classes for the storage of raw data coming from microscope in the format of ROOT trees. This include the possibility to store microtracks, clusters (grains), raw images and scanning parameters. Most commonly we store only microtracks and attached clusters, but for the test purposes and for the analysis of special events the possibility for storing all clusters and images is useful.
- Input–output and flow control module (libEIO) contains the set of classes for processing parameters handling and for the intermediate and final input–output of the each reconstruction step, currently available in the format of the ROOT trees. This approach makes easy the users control and analysis after each level of processing as well as the attachment of additional libraries or routines dedicated for the specific purposes. Interface to the ORACLE database for the storage of the processed data is under development.
- Reconstruction modules (libEdr, libVt + +) keep the most of algorithms necessary for data reconstruction and pattern recognition
  - ○ microtrack linking to form the basetracks,
  - ○ plate-to-plate alignment using the basetrack information,
  - ○ track finding: recognition of the long volume tracks formed by the basetracks,
  - ○ track fitting using the Kalman Filter (KF) technique,
  - ○ vertex finding and fitting with KF,
  - ○ event analysis using the three-dimensional Event Display and user customized scripts.
- Internal Monte-Carlo library (libEMC) for the fast algorithms testing. Multiple scattering, energy loss effects in the material of ECC, n-body decay kinematics for the testing of vertex and tracks reconstruction, apparatus resolution and smearing effects are taken into account. The possibility to have the exact knowledge and control on the input data is essential for the tuning of reconstruction algorithms.
- Interactive event display (libEdd). Traditional for ROOT-based applications event display with the possibility of three-dimensional animation, zooming–unzooming, rotation and picking vertices, tracks, segments and hits objects for the control and inspection is provided.
- Image and grains analysis library (libEGA) developed mainly for the test purpose. While these tasks are normally performed on the hardware level by the scanning system, in some special cases it is convenient the developing and testing offline the dedicated algorithms for example for the high-precision measurements of for the study of low-level resolution components.
- Additional and service libraries (libEmath, libEphys)—contain some specifications for emulsion processing mathematical and physical functions and classes.

3. Review of the data reconstruction algorithms

3.1. Microtracks linking

After the scanning of one emulsion plate the output is two arrays of microtracks—one per each side and the first task is to find the correct matching of 2 microtracks corresponding to the same charged particle. This is done by the selection of the best couples with minimal χ^2 criteria. Microtracks coordinates, angles and number of clusters (puls height) participate in χ^2 calculation. Angular dependency of the microtrack resolution is taken into account. The resulting track vector called *basetrack* connect the most inner (base) points of the corresponding microtracks. Linking is usually done by iterations where the first ones are used for emulsion shrinkage correction and the data quality check. This permit to improve significantly signal/noise separation and base-tracks resolution.

3.2. Plates (patterns) alignment

Plate to plate patterns matching is one of the most important reconstruction operations necessary to pass from the reference systems (RS) of the individual plates to the global one. In case of emulsions intercalibration this is the routine operation has to be done for each scanned plate. This procedure calculate the affine transformations between the adjacent patterns (shift, rotation and expansion) and apply it in a way to have all (56) patterns in the same RS.

Based on the hypothesis that tracks passed through the assembled ECC are nearly straight, the corresponding patterns may be found. This permit us to pass from the mechanical accuracy of plate positioning (about 50 μm) to the intrinsic accuracy of the emulsion (0.1–0.5 μm) This operation is formalized in FEDRA and could be applied to any kind of patterns in any-dimensional phase-space (for segments we use the four-dimensional one: x, y, tx, ty). Operation is speed-optimized for combinatorial reduction.

3.3. Tracks finding and fitting

Track finding procedure consists of the three main steps. On the first step all couples of adjacent basetracks are found. Then from the couples we form long chains of segments without the interrupts (missing plates). These chains serve as a triggers to start the KF procedure for track fitting and following. The last step is the track propagation taking into account the possibility to loose segments in one or more plates (normally we permit a gap of 3 consecutive plates). The final result of the procedure are the long tracks consisting of the array of segments accompanied by the “fit function” provided by the KF procedure. The main criteria for tracks/segment acceptance is the probability given by KF. We use KF with smoothing to improve the tracking accuracy. Apparatus resolution provided as the input parameters and multiple scattering effects are taken into account for the probability and fit calculation.

3.4. Vertex finding and fitting

The preliminary triggering operation for the vertex finding is the track-to-track couples search using the minimal distance criteria. Some topological cuts are used for the reduction of combinatorics. The final criteria for the track couples selection is the vertex probability limit calculated with the full covariance matrix of the participating tracks for the 5 parameters (x, y, tx, ty, p). Starting from couples the n-tracks vertices are constructed using the Kalman Filtering technique. For the mathematical part relevant for the KF algorithms for tracks and vertices fit we use libVt + + [5] library. This library was developed originally for Hera-B and has no any emulsion specifics.

3.5. Physics data analysis

A number of analysis tools for data check and quality control are provided by the system. For the physics and final events analysis a few high-level algorithms are already available as the library functions: momentum measurement by multiple scattering, kinematics analysis, etc. A number of new algorithms including ones for shower analysis with use of Neural Network, π/μ, π/e separation, were recently developed in the different scanning laboratories based on the FEDRA framework and may be smoothly integrated into the system as the standard libraries.

4. Conclusion

We developed the system for the off-line reconstruction and analysis of data coming from nuclear emulsion scanning. The system already covers most of general tasks necessary for the pattern recognition and event analysis. The code is self-consistent and depends only on the ROOT library, it means that it could be compiled and used on any OS with ROOT installed. It is fully tested and exploited now on Linux and on MS Windows (with VC + +7.0 and higher). The present algorithms have been mostly checked using Monte-Carlo data and are being validated and used intensively for the processing of the OPERA test beam data and OPERA simulation data. The FEDRA package is available through the OPERA CVS server and supported by the INFN Napoli OPERA group [4].

References

[1] M. Guler, et al., The OPERA proposal, CERN-SPSC-2000-028; M. Guler, et al., CERN-SPSC-2001- 025; Y. Declais, et al., CERN-SPSC-2002-029.

[2] T. Nakano, in: Proceedings of International Europhysics Conference HEP2001.

[3] N. Armenise, et al., Nucl. Instr. and Meth. A 551 (2005) 261–270.

[4] http://www.ntslab01.na.infn.it/fedra

[5] Th. Glebe Hera-B Note 00-175 (2000).