Neutrino Event Reconstruction in a Liquid Argon TPC

Gary Barker
Department of Physics, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL
E-mail: G.J.Barker@Warwick.ac.uk

Abstract. We present some preliminary findings and results from activities in Europe and the USA working towards an automated, algorithmic, reconstruction of particle interactions in liquid argon time projection chambers.

1. Introduction
The capability of Liquid argon (LAr) to record particle interactions with a density of detail comparable to the bubble chamber experiments is well recognised. The challenge of reconstructing these interactions to extract physics measurements has been partially met over the years but a fully automated, algorithmic, approach has yet to develop. The reasons for this are partly historical e.g.

- Bubble-chamber event rates/volumes were low by modern standards and hence there was no strong need for an automated solution.
- Available computing power limited what was possible.
- Collider experiments with electronic readout began to dominate the subject.

Tracking algorithms have therefore developed, optimised for reconstructing the sparse, but high precision, space points measured by typical collider experiments (see Fig.1).

Other reasons why automated reconstruction in liquid argon is still developing relate to the fact that it is not such a simple task. Reasons for this include:

- Events are topologically rather complicated (see Fig.2). Tracks and showers develop side-by-side in the same detector volume - compare this to the collider experiment where showers develop outside of the inner tracker volumes (see Fig.1).
- Unlike in the collider format, interactions in LAr can be initiated anywhere throughout the volume making knowledge of the start-point of events uncertain. This is a hurdle for any algorithmic approach.
- The high density of energy deposits (or ‘hits’) in LAr reveal a tremendous level of detail (vertices, kinks, delta rays etc) presenting challenges for efficient reconstruction.
- Multiple scattering occurs continuously throughout the volume and not just at well defined boundaries between detector elements.
2. Previous Work

Early attempts to automate reconstruction in bubble chamber data were led by the development of the Hough Transform technique [3] which later found wide application as a feature extraction method for image analysis. More recently, the full reconstruction of interaction events in a LAr TPC has been demonstrated by the ICARUS programme. The full chain of hit reconstruction, hit-clustering to form track objects and track trajectory fitting in 2D and 3D have been demonstrated (see e.g. [4]). However these results have relied on some degree of visual scanning and selection before applying algorithms. Recently, with the imminent start-up of ICARUS T600 in the CNGS beam and data taking with LAr TPC prototypes in Europe and the USA ramping up (e.g. ArDM and ArgoNeut respectively), interest has again turned to address the issue of a fully automated event reconstruction.

3. Current Status

The physics programme for any future large LAr TPC demands that: (a) hits are classified into ‘track-like’ and ‘shower-like’ objects; (b) particle tracks (e.g. muons, pions, protons) and showers (e.g. electrons and photons) are identified. Reconstructed quantities such as particle trajectories, $dE/dx$, event topologies, kinematics etc will enable these objectives to be met but their accurate reconstruction depends crucially on the correct assignment of hits to a common event feature i.e. hit-clustering. Recognising this, recent strategies for automatic event reconstruction have centred on the development of clustering algorithms to define groups of hits likely to comprise either a particle track or shower from which a track fitting algorithm or some higher-level shower algorithm (to extract shower energy, shape etc) can be applied as appropriate. As alluded to in the introduction, this task can be significantly simplified if additional information about interaction vertex locations, kinks etc is available. Such ‘key-point’ detection in event records is an area under active study for use in conjunction with clustering and fitting methods in order to boost performance. In what follows, we briefly summarise some of the studies underway in these areas.
3.1. Density-Based Clustering: DBSCAN

The aim of a hit clustering algorithm is to group together all hits which originate from a common parent while keeping to a minimum, contamination of the group by hits that originate from elsewhere. The DBSCAN algorithm [5] has recently been applied to ArgoNeuT data. This is an example of density-based clustering where clusters are identified as regions of high hit density sitting in an environment of lower density which work by demanding that the region in space around each hit in the cluster (defined by a radius $\epsilon$) contains at least $N_{\text{min}}$ other hits. Figure 3 shows two views of a neutrino interaction from the two readout planes of the ArgoNeuT detector. The colours depict the preliminary results [6] of applying DBSCAN – hits of the same colour have been grouped by the algorithm into the same cluster. The performance is promising although close inspection of the event displays shows that sometimes hit distributions that belong to the same cluster can be split into smaller clusters. In order to help optimise the performance studies of the package, Ordering Points To Identify the Clustering Structure (OPTICS) [7], have begun which stores hits in a convenient ordering with which to investigate optimal $\epsilon$-clustering scales [8].

![Figure 3. DBSCAN applied to an interaction in the ArgoNeuT detector. Drift time is the $y$-axis, readout wire number the $x$-axis and the grey-scale represents the size of energy deposit recorded.](image)

3.2. Key-Point Detection

The reconstruction task in LAr should be simplified if the location of certain ‘key-points’ are known e.g. interaction vertices, start/end points of tracks, delta-electron kinks etc. This is illustrated in Fig. 4 which shows that the DBSCAN algorithm can have difficulty grouping hits into separate track objects around an interaction vertex. If the location of the vertex point was known a-priori, the hits in the immediate location of the vertex could be removed from the clustering step allowing the algorithm to find correctly two, separate, clusters. In addition, Fig. 5 highlights that a straight line Hough-transform track fitting algorithm returns no information on the start and end-points of tracks and so these must be estimated by other means which are subject to errors.

These conclusions have motivated a study based on the Harris-Stephens operator [10] which numerically highlights features by computing the differential of pixel response in the event image and noting that these are large in all directions for something like a vertex, while large mainly in one direction for something like a track end-point. Some preliminary results are shown in Fig 6 and a quantitative analysis of the performance finds that in simulated quasi-elastic
Figure 4. DBSCAN clustering around a vertex point.

Figure 5. Result of a Hough-transform based track fitting algorithm around the track end-point [9]

Figure 6. The result of a Harris-Stephens key-point algorithm applied to a simulated charged current neutrino interaction in LAr.

$\nu_\mu$ interactions in LAr at $E_{\nu_\mu} = 0.77$ GeV, vertex and both the start and end-points of tracks can be identified in 3D with $\sim 85\%$ efficiency [8].

4. Conclusion

Activity towards the goal of automated reconstruction in LAr TPC’s has begun in Europe and the USA and preliminary results are beginning to emerge from this effort. A central issue is that of hit clustering, for which the DBSCAN algorithm has returned some promising results and early work has identified the advantages of using this in combination with event key-point reconstruction – for which the Harris-Stephens method has been successfully applied. The work reported on here is still very much in the preliminary stages. Plans for future development include investigating ‘trained algorithms’ such as neural networks and to continue investigating existing techniques from the disciplines of Machine Vision and Image Processing.
Acknowledgements
I wish to thank the following colleagues for useful discussions and access to results: Andrew Bennieston, Ben Morgan, Yorck Ramachers and Dan Roythorne in Warwick and members of the LArSoft collaboration, especially Brian Rebel at Fermilab.

References
[1] Event display pictures available from: http://aleph.web.cern.ch/aleph/
[2] A. Ankowski et al, ICARUS Collab. 2008 hep-ex:0812.2373
[3] Hough 1959 Proc. Conf High Energy Acc. Instr., CERN
[4] ICARUS Collab. 2006 Phys. Rev. D 74 112001
[5] Sander et al 1998 Data Mining and Knowledge Discovery 2 169-194
[6] Results courtesy of Kinga Partyka, Brian Rebel and the LarSoft Collab.
[7] M. Ankerst et al 1999 ACM SIGMOID Int. Conf, on Management of Data 49-60
[8] Results courtesy of Ben Morgan, Dan Roythorne and the University of Warwick group
[9] Results courtesy of Joshua Spitz, Brian Rebel and the LarSoft Collab.
[10] C. Harris, M. Stephens 1988 Proc. 4th Alvey Vision Conf. 147-151