DETECTING PARTICLE TRACKS
WITH A
MODIFIED LOCAL HOUGH TRANSFORM

by

AVRONIL BHATTACHARJEE

THESIS

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

1995

MAJOR: COMPUTER SCIENCE

Approved by:

Advisor Date

Advisor Date

Date

Dedication

To my family and friends

and then to the cause of science.

Acknowledgments

Through the development of this thesis many people have given support and encouragement. My gratitude to my advisors : **Dr. I. K. Sethi**, who has been very patient during the first months of the project, and especially to **Dr. Claude Pruneau**, who also as my friend and mentor has made all this possible.

I would like to thank everyone in the RHIC group at WSU for making the High Energy Physics computation laboratory, where most of the work was done, a friendly and very productive environment, especially, Radim Cernej, who spent many hours with me to get PAW to do just what I wanted, Joseph Bielecki for his tremendous assistance in getting this thesis ready and Dr. W. K. Wilson, who helped with the generation of immense amounts of data, spending many long nights with GEANT.

Hebrews 11:1-3

Now faith is the substance of things hoped for, the evidence of things not seen. For by it the elders obtained a good report. Through faith we understand that the worlds were framed by the word of God, so that things which are seen were not made of things which do appear.

Contents

List of Tables

List of Figures

Chapter 1

Introduction to our problem

1.1 Problem definition

The Hough Transform, is a widely known and very effective tool for pattern recognition where data points are "binned" or histogrammed according to their match to a given parametric model, for extraction of the most likely underlying structure. It has seen numerous applications in Machine Vision, shape & object recognition, robotics and many allied areas. In this work we present the use of this familiar and powerful method for a "new" application in the area of High Energy Physics.

The basic problem is one of detecting & reconstructing the 3-dimensional tracks of the decay products of a relativistic heavy ion collision inside a detector. The paths are sparsely sampled in the form of electric pulses, as the ions cross the detector, and from these hit-point coordinates in space, the few thousand odd decay paths have to be reconstructed in as short a time as possible.

The problem is compounded by the extremely high density of the hits in space, the presence of noise, spurious points and a finite error in the measurement of the hit positions. In fact if treated with brute force it becomes a very high order combinatorial optimization problem with little scope for direct solution.

1.2 Problem constraints

As we shall outline next (in chapter two), for the STAR detector one essential recognition problem to be solved is that of track reconstruction for Au (Gold) onAu collisions. Current theoretical models predict that the number of particle produced per central Au + Au collision should be in the neighborhood of 2000 particle tracks per event, in the acceptance range of the detector. Needless to say, numerous computer simulations have already been performed and have demonstrated that the technology chosen for STAR, a large time projection chamber (TPC), can handle such a large multiplicity of particles properly and allow reconstruction of a reasonable fraction (in excess of 90%) of all tracks.

Up till now, STAR TPC computer simulations have used a track reconstruction algorithm based on a conventional road finding technique as by Sakreda etal [14]. Altough this technique is quite satisfactory from the point of view of yield (reconstruction efficiency better than 95%) and accuracy (good purity and good momentum resolution), it is found to be quite CPU intensive and time consuming. Currently it takes over a quarter of an hour on a single CPU workstation to analyse a full TPC event. Also by its very nature the road finding algorithm is difficult to exploit for parallel or distributed implementations. It is unlikely to be easily mapped to multi processor systems and hence provides a lower ceiling on achievable speeds.

Given that STAR expects to acquire data at a rate of one central Au + Au collision per second, the analysis of the data acquired in the first year alone would take a few hundred CPU-years if one is limited to the current algorithm and computer speeds. Although one foresees increases in the CPU speed by up to a factor of 10 till RHIC turns on, and one plans to parallelize the analysis on large CPU farms, the analysis of the TPC data still appears somewhat unbearable unless one can find faster tracking algorithms.

1.3 Outline of solution

In this context it appears critical to explore other alternatives. These new techniques may not be optimal for the low track density environments encountered in the early heavy ion collision studies, but could yield comparitively superior performance in very high track density events. We shall therefore briefly review various other approaches as reported in the literature (chapter 3).

Amongst the new techniques that have been published the Hough transform is especially attractive in view of its intrinsic simplicity and robustness. A number of possible improvements to the published works, especially to speeden up, and achieve good results on very high density events are fortunately possible, and form the main object of this thesis. After taking a look at the physical characteristics of our data and the TPC (chapter 4) we present our new approach based on the modified Hough transform coupled with the pitch angle decomposition for making the proposed solution data-parallel and achieving high speed track recovery.

We have tested our algorithm in detail on fairly realistic simulations of heavy ion collisions and these results of computer simulations are presented in chapter 6 and lastly chapter 7 summarizes the key features of our approach and provides a few pointers to further improvements along our present line of work.

The next chapter introduces us to the goals and layout of the **STAR** experiment and briefly points out its major physical components of the detector and its various sub assemblies which in turn impose the hard (physical) constraints on our track reconstruction algorithm.

Chapter 2

The STAR experiment

2.1 The Scientific Motivation for STAR

The purpose of **STAR** (Solenoidal Tracker At RHIC), a heavy ion Au-Au high energy relativistic collision experiment, is primarily the detection of signatures of quark-gluon plasma (QGP) formation and the investigation of the phenomena of strongly interacting matter at high energy density. The correlation of many observables on an event-by-event basis will be crucial. and these correlations will be used to identify possible QGP signatures and special events. Thus the detection system for STAR must be capable of measuring many experimental observables simultaneously. Even if signatures for the QGP are not explicitly determined, the detector system should be versatile enough for other special events. The motivation and goals of the STAR experiment will provide a foundation for the physics and purpose beyond the TPC (time projection chamber) detector and the data analysis tools developed in our study which comprises this thesis.

2.2 The Physics Goals of STAR

At RHIC, two fundamentally new aspects of hadron production will be studied: correlations between global observables on an event-by-event basis and the use of hard scattering of partons as a probe of the properties of high density nuclear matter. Very high charged particle densities, $dn_{ch}/d\eta \approx 700$ expected in nucleus-nucleus collisions, provide a favorable environment for event-by-event measurement of global observables. The experiment is designed to detect event-by-event fluctuations expected near phase transitions by monitoring threshold behavior of the experimental observables as a function of energy density.

Full azimuthal coverage of particle identification and continuous tracking with a seamless detector will provide maximum particle yields needed for successful measurements. Hard QCD processes can be investigated due to jet yields at RHIC. This can be reached by tracking combined with segmented electromagnetic calorimetry and high p_t single particle measurements. Possible QGP characteristics can be observed by measuring hadronic cascades of hard-scattered partons. This measurement will also give new information on the nucleon structure function and parton shadowing in nuclei. Inclusive and event-by-event observables for proton-proton, proton-nucleus, and nucleus-nucleus interactions are measurement requirements for the STAR experiment. To conclusively identify signatures of QGP formation all observables will be studied as a function of impact parameter and incident energy [15].

2.3 Design Concept and Criteria

STAR is designed as a large acceptance device with simultaneous measuring capacity for both soft and hard physics. The relevant design criteria are:

- large rapidity coverage
- detection of as many charged particles as possible with high efficiency, to provide high statistics for event-by-event observables and fluctuation studies
- 2π continuous azimuthal coverage for reliable event characterization
- high tracking efficiency as close to the vertex as possible, to contain the size and cost of the experiment
- adequate track length for tracking, momentum measurement and particle identification for a majority of the particles
- high kaon detection efficiency for event-by-event observables and for inclusive measurement of the ϕ-meson
- momentum resolution $\Delta p/p \leq\ 0.02$ for $p_t\ <\ 1.5$ GeV/c
- two-track resolution providing a momentum difference resolution of a few MeV/c for HBT correlation studies.
- accurate determination of the primary vertex for high momentum resolution and to be able to identify particles from the primary interaction
- accurate determination of secondary vertices for detecting strange particle decays such as Λ, $\bar{\Lambda}$, K^o_s, Ξ^+, Ξ^-, and Ω^-
- large & uniform acceptance to maximize rates and minimize edge effects in jet reconstruction
- accurate determination of the primary vertex in order to achieve a momentum resolution $\Delta p/p\ <\ 0.05$ for $p_t\ =\ 10$ GeV/c for high p_t particle spectra and jets
- electromagnetic calorimetry combined with tracking and good momentum resolution up to $p_t\ =\ 10$ GeV/c to measure and trigger on high E_t jets
- 2π azimuthal coverage of electromagnetic calorimetry for event characterization

The design of the detection system of the experiment conforms to the physics goals. Most of the mid-rapidity charged particles need to be detected, and their tracking, momentum analysis, and particle identification performed. Track detection must be obtained in higher than expected maximum charged particle multiplicities for central collisions ($n_{ch} = 3000$ in the range $|\eta| < 1$ at mid-rapidity). Pions/kaons can be distinguished for $p < 0.7$ GeV/c and kaons/protons can be distinguished for $p < 1$ GeV/c via dE/dx. Decay particles will be measured and their second vertices will be reconstructed. Resolution at a 2 m radius from the vertex is expected to be 2 cm. Momentum resolutions ($\Delta p/p$) of 0.02 at 0.1 GeV/c are needed for confirmation of the physics expected, and $\Delta p/p$ in the order of a few percent is necessary for $p = 10$ GeV/c to measure rapidly falling spectra at high p_t and particles from jets. Limiting momentum resolution at low p_t is due to multiple scattering, which will be minimized by keeping detector material minimal through all of the detection system. Figure 2.1 shows the STAR experiment detectors.

The momentum will be measured for a large pseudo-rapidity range ($|\eta| < 2$) at mid-rapidity with complete azimuthal coverage ($\Delta\phi = 2\pi$). In the range $|\eta| < 1$, particle identification will be obtained. The information that the detection system will provide are firstly tracking and momentum analysis through a time projection chamber (TPC) and a silicon vertex tracker (SVT) encompassed by a solenoidal magnet, and secondly particle identification via dE/dx, and position of primary and secondary vertices. A collision geometry trigger will be provided by the following specific detector components:

- a central trigger scintillator barrel around the TPC
- vertex position detectors near the beam line but outside the magnet.

- calorimeters located in the area of the beam insertion magnets to reject events dictated by the number of spectators

Furthermore, an electromagnetic calorimeter might reside inside the solenoidal magnet for triggering on transverse energy and jet cross sections.

The trigger detectors will be used to select events on the basis of collision geometry, event topology, and particle species content. Track information will be accessed from the silicon vertex tracker (SVT), which will be able to fit primary vertex tracks and determine secondary vertex tracks of decaying particles below a 6 cm radius constrained by the geometry of the detector, and a time projection chamber (TPC) inside a solenoidal magnet producing a helical path for continuous tracking, accurate momentum resolution and particle identification.

The physics of the STAR Experiment demand acquiring information of projectiles from protons to Au, available beam energies. Impact parameters of soft physics processes will be studied from minimum bias to centralized collision geometry trigger range, and later on more refined methods of triggering will be one of the essential points of the future of STAR. Rare events and jet production studies are hoped to be obtained through longer data acquisition periods [15].

Next we discuss the uniform magnetic field and the layout of the Time Projection Chamber which is the focus of the track recognition algorithm described in this work.

2.4 Magnetic Field

The solenoidal magnet around the TPC and SVT will produce a uniform field along the beam direction. Field strength will be a maximum of 0.5 T. Shaping the magnetic field will be done through an iron return yoke and shaped pole pieces. This design is to facilitate high tracking accuracy and azimuthal symmetry. Other smaller coils will be used to enhance uniformity of the field. The maximum strength was chosen to allow resolution of tracking particles with momentum as high as 10 GeV/c while retaining low p_t particles (down to 60 Mev/c) also with reasonable tracks. The uniformity of field enables a number of simplifying assumptions in the track reconstruction algorithm, as the charged particle tracks can be considered to be nearly perfect helices in space.

2.5 Time Projection Chamber (TPC)

The TPC will have two longitudinal drift sections, each 2.1 m long. The particle carriers created from the ionization will drift in the longitudinal direction, due to the electric field in the TPC, to its end-caps. 70,000 pads make up each end-cap. Two pad sizes, one for 50 cm<radius<125 cm (2.85 mm x 11.5 mm) and one for 125 cm≤radius<200 cm (6.2 mm x 19.5 mm) will be implemented. The pads will be read out into 512 time samples. Ignoring energy losses, singly-charged particles with $p_t <$ 40 MeV/c will spiral inside the inner radius of 50 cm of the TPC. Only those singly-charged particles with $p_t >$ 150 MeV/c will spiral completely through to the outside radius of 200 cm of the TPC. The acceptance hence is limited to the longitudinal length and inner and outer radius of the TPC, giving a value of $|\eta| <$ 1 for outer radius and $|\eta| < 2$ for all particles which reach the inner radius, but not necessarily traverse completely through to the outer. In this work we will mainly focus on the reconstruction of the tracks inside the TPC.

2.6 Silicon Vertex Tracker (SVT)

The SVT is to extend the resolution of the TPC closer to the vertex to position primary interaction vertices, improve momentum and dE/dx resolutions, and determine secondary vertices to an accuracy less than 100 μm. Positioned closer to the vertex collision, the SVT will extend the acceptance of the TPC for particle momenta lower than 150 MeV/c. The SVT will provide this information by storing three dimensional space points and resulting tracking vectors with high resolution. From this track information, the SVT can be linked to the TPC to continue tracking of projectiles. Although we do not use the SVT information in this work it is a natural and trivial extension to use them by merging SVT tracks to the TPC templates.

2.7 Electromagnetic Calorimeter (EMC)

The EMC will measure and trigger on the total E_t and local $d^2E_t/d\eta d\phi$ transverse energy deposition in p-p, p-A, and A-A collisions. From the soft nuclear processes, the temperature and energy density of the reaction can be determined by analyzing the transverse energy at mid-rapidity. Triggering on jets, high energy direct photons, and high p_t particles can be accomplished with the EMC. Properties of interactions of hard scattered partons in the hot nuclear medium may be inferred from this information. Spin structure of polarized pp collisions can be investigated as well as single spin transverse asymmetries. The barrel of the EMC will be made of lead-scintillator and be located inside the superconducting solenoid coil and inside the iron flux return. Its coverage is $|\eta| \leq 1.05$ and 2π in azimuth. Energy resolution for the EM calorimeter is expected to be near $0.20/\sqrt{E}$.

2.8 Triggering and Trigger Detectors

A particular signature for the quark-gluon plasma is not known yet, so the triggering must be versatile in that it is able to explore various paths and flexible to upgrades and adoption when understanding grows. High dead times will be used in the start of STAR until enough information is collected to selectively trigger. The goals of the triggering detectors are:

- detect Au-Au interactions with $>$ 95% efficiency
- detect p-p interactions with $>$ 50% efficiency
- locate Au-Au vertex to within ± 6 cm within < 200 ns
- detect multiple events occurring within 40 μs with 90% efficiency
- select highest 2.5% of multiplicity or E_t distribution in < 200 ns
- select high transverse electromagnetic energy
- select based on geometry of hit patterns
- allow simple integration of upgrades to trigger system

The interaction rates for Au-Au and p-p are expected to be 1 kHz and 1 MHz with the event recording rate expected to be 1 Hz and 30 Hz, respectively. The trigger

Table 2.1: *The detectors contributing to the trigger process. Note, ϕ from 0 to 2π.*

Detector	$\lvert\eta\rvert$	*# of Channels*
Central Trigger Barrel (CTB)	0 - 1	200
TPC Endcaps (MWPC)	1 - 2	2000
Vertex Position Detectors (VPD)	3.3 - 3.8	50
Veto Calorimeter (VC)	> 5	2
EM Calorimeter (EMC)	0 - 1	1200

system will prioritize the events and select those of higher priority from the sample of events. Events can be selected out during the TPC readout/digitization, but not during the drift time-window.

The scale for trigger decision time is based on three natural time scales for STAR:

1. the RHIC beam's crossing time of 110 ns
2. the maximum drift time in the TPC of 40 μs
3. the readout/digitization time for the TPC of 10 ms

The STAR trigger comprises three distinct levels.The first level is based on vertex location and centrality and is obtained in about 110 ns. The second is based on the topology of hits and is bounded by about 40 μs. The third level is on software decisions of tracking information and takes roughly 70 ms. Each process of the trigger decision time is expected to have a keep / flush decision capability.

Figure 2.1: *The STAR detector geometry setup.*

Chapter 3

Present Approaches to tracking

3.1 Review of neural methods

Recently, there have been a number of new techniques studied for track recognition and reconstruction in time projection chambers [7, 9, 11]. Basically, some of these new techniques use the a-priory knowledge that the tracks are to a very good approximation helices in space, while others like the Artificial Neural Net (ANN) approach use only certain global energy minimization methods.

Denby [2] and several others have suggested tracking approaches based on a particular neural architecture, namely the Hopfield Network [13] which can be set up for minimizing energy functions specified in terms of the neurons' inter-connectivity. The general idea is to set up neurons to correspond to small track segments and then changing the connection strength of any two neurons, according to the probability of the corresponding two track segment's belonging to the same track.

We consider a set of N space points measured in a TPC. From these points we define a set of *legal segments* defined as short tracklets joining two points in our hit set such that the separation of these two points is less than some cutoff distance R_c, where this R_c is also of the order of a few times the average distance between hit points. Thus each point is considered to interact with a local neighborhood within a sphere of this radius in space.

Each legal segment is now represented by a neuron in an array. The circle in figure 3.1 for example represents the directed edgelet $3 \rightarrow 2$. The final solution then consists of an array in which only legal segments are "ON". Each track then may be read off by a non bifurcating, unbroken chain of segments by following the links from point to point on the table. On a valid track no point should have more than one track segment entering or leaving it and conversely no point should appear more than once.

From this we can set up inhibitory connections between each "ON" neuron and the others in its row or column. The reinforcing connection weights are set up so that short end to end connections are favored. For example the (3,2) neuron will provide support to all outgoing links from neuron 2 (that is all tracklets originating in row 2). This chaining together of similarly directed segments is similar to the minimal spanning tree approach.

The interconnection weights can be made both inhibitory and excitory and are generally a function of the angles between the tracklets as well, as smoothness in the joining of the small segments is an important criterion. For reasons of convergence one chooses to make the connection strength symmetric between a pair of neurons and this connection strength between neurons i and j denoted T_{ij} is given by :

$$T_{ij} \propto K \times [cos^n \theta_{ij}] \div R_{ij} \tag{3.1}$$

Careful tuning of the network parameters are essential and many features such as the rate, magnitude and sequence of weight updates strongly affect the quality of results. In general, resolution between parallel, closeby tracks is poor and a long set of iterations may be necessary for convergence. Further, techniques like **simulated annealing** have to be used to prevent capture of the net in local energy minimae instead of finding a globally optimal solution.

Although neural methods as reported by Denby [2] or Stimpfl & Garrido [4] appear promising to disentangle tracks in high density areas or in merging tracks for HBT analysis, these methods suffer from the same basic problem, they rely only on a concept of global energy minimization and do not use local structural properties (that is templates), which is essential for reconstruction speed.

An alternative approach to tracking consists in seeking tracks with a simple and fast (although possibly not quite as accurate) algorithm. If the first pass on the data succeeds in finding a significant fraction of "plausible" tracks, one may then envision performing additional passes, if needed, with slower but possibly more accurate algorithms. Additionally, in the light of recent developments in massively parallel computing, it also appears sensible to search for approaches that can be mapped easily onto distributed architectures. Although such algorithms may not perform their best on current platforms, one might envision easily porting them on to the new parallel hardware.

3.2 Model or template based tracking

In the context of a TPC, the particle tracking problem can be reduced to finding helices. An unknown subset of the data points correspond to detector noise and should be left unmatched. The bulk of the data however is expected to roughly follow helicoidal trajectories or simply circles in the transverse plane of the detector. Our strategy should thus be to find such circles in the transverse plane of the TPC. Moreover, as one expects that in excess of 95 % of all particles produced, originate from or very near the main vertex. It is quite reasonable to assume, in the first pass, that all circles should pass through the origin. This assumption is strictly speaking valid only for "primary" tracks, i.e. tracks produced exactly at the main vertex.

Many **secondary** tracks, resulting from the decay of short lived particles produced at the main vertex, are produced at finite distances from the main vertex, and should not, in principle, be found by an algorithm searching for circles passing it exactly through the origin. In practice however many of these short lived particles decay just a few centimeters away from the main vertex and lead to daughter particles having trajectories passing very close to the primary or main vertex. One thus expects the circle finder should be able to find a large fraction of the so-called secondary tracks also.

A generic approach to such template based matching is the robust vote based technique of Hough transforms [8], which have been used by a number of authors. Ohlsson etal [7], in particular, use it to provide a fairly fast search technique and generate track templates. In their approach most of the initial work is concentrated along the generation of elastic arms [7] (or track templates) in terms of radius of curvature, and position of the helices, to which the hits are then assigned using a Gaussian distribution function of the allowable distance of a hit from the template. The process is theoretically analogous to an energy minimization process where the elastic arms are attracted to the best fit points as the system's "temperature" is lowered.

3.3 Local versus global tracking : pros & cons

An interesting comparison of these methods is in terms of **local** versus **global** tracking. For the neural methods which are based on the "smoothness" concept alone, the tracks do not need to come from the origin, or even be helices. Some arbitrary energy function is minimized for the hit assignments, and the geometrical properties of the underlying tracks are largely ignored. The network merely settles to the most reasonably shaped curves in the volume being examined. This "model free" approach is however of great use in cases like **secondary vertex** tracking where *a-priory* information is not available.

In comparison the binning (Hough) techniques choose **long range** features and use geometrical constraints for fits to the track model. In practice, such a histogramming approach works faster whenever feasible. It needs to be emphasized that this applies only where tracks are approximated by analytical models such as circles or straight lines and not arbitrary curves. An added advantage is that rejecting noise and fitting hits is much easier and more robust too in any *model based* reconstruction as opposed to *model free* approaches.

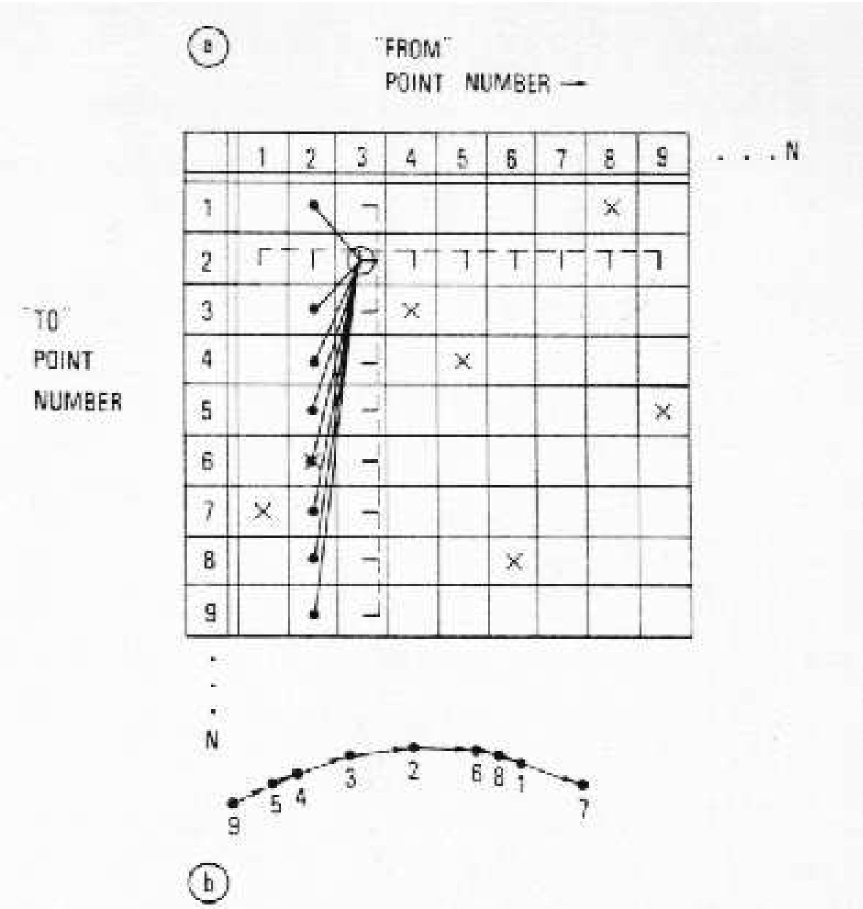


Figure 3.1: *The array of Neurons*

Chapter 4

The TPC data model

4.1 Detector characteristics

The STAR baseline detection system consists of a large Time Projection Chamber (TPC) inside a solenoidal magnet with 0.5 T magnetic field to enable tracking, momentum analysis, particle and vertex identification etc. The chamber has a cylindrical geometry. It is to be operated with P10 or Helium - Ethane gas mixture at atmospheric pressure. The length is about 5 meters and the inner and outer radius of the chamber are about 50cm and 200cm respectively.

The particle beams enter the TPC from either side along the central axis and collide in the middle. Free electrons produced by the ionization resulting from the passage of charged particles through the gas, drift towards the two end caps under the influence of the uniform longitudinal magnetic field. Each end cap contains 45 rows of detector pads. These pads measure the amount of charge arriving locally as a function of time. The pad position determines the (x,y) location of the track in the transverse plane while the drift time determines the position along the beam axis.

The charge read out is transferred to switched capacitor array pipelines, multiplexed and digitized. This digitized data is stored for analysis. The primary stage is thus extracting the position of the charge ionization cluster resulting from the passage of the charged particles. The set of these "space points" forms the input to the second stage which deals with track reconstruction and forms the object of discussion in our work.

4.2 Mathematical modelling

The input data therefore consists of space points (x_i, y_i, z_i) belonging to tracks to be detected and reconstructed. In addition to these, a random (additional) number of points which are spurious may also be present. The task at hand involves two stages. The first stage consists of recognition of tracks and the assignment of the hits to the tracks. In the second stage one fits these points to a track model to extract the momentum vector of the corresponding trajectory.

The magnetic field within the TPC is very uniform and parallel to the beam axis as described in chapter 2. Neglecting the effects of energy loss, the radius of the trajectories in the transverse plane of the field should be roughly constant along the trajectory. A simple model based on the helicoidal trajectories is thus fully justified. The parameters of the model are illustrated in figure 4.1 and the governing equations of the model are given below :

$$x(s) = x_0 + R_H[cos(\Phi_0 + h \times s \times cos(\lambda)/R_H) - cos(\Phi_0)] \tag{4.1}$$

$$y(s) = y_0 + R_H[sin(\Phi_0 + h \times s \times cos(\lambda)/R_H) - sin(\Phi_0)] \tag{4.2}$$

$$z(s) = z_0 + s \times sin(\lambda) \tag{4.3}$$

where s is the path traveled along the track and $s = s_0 = 0$ gives the first hit point (x_0, y_0, z_0) as in Figure 4.1. R_H is the radius of the helix and Φ_0 is the azimuthal angle giving the "direction" of the particle at the first point on the track. The factor h is +1 for positive particles and -1 for negative particles.

The helicoidal model thus offers much simplicity, in that the longitudinal and the transverse motion are decoupled and can be treated separately.Figures 4.2(a & b) show the cross sectional views of a typical TPC event, at less than half density of a typical Au on Au event for the (x-y) & (y-z) projections. It is immediately apparent that at such great track densities conventional pattern finding is extremely difficult. The human eye can barely distinguish any patterns at this level.

The corresponding radius of curvature and momentum spectra for such an event are also shown next in figures 4.3(a & b) which provide us valuable measures of the real extent of radius of curvature of tracks which is to be used later in setting constraints on our Hough Transform's generated templates.

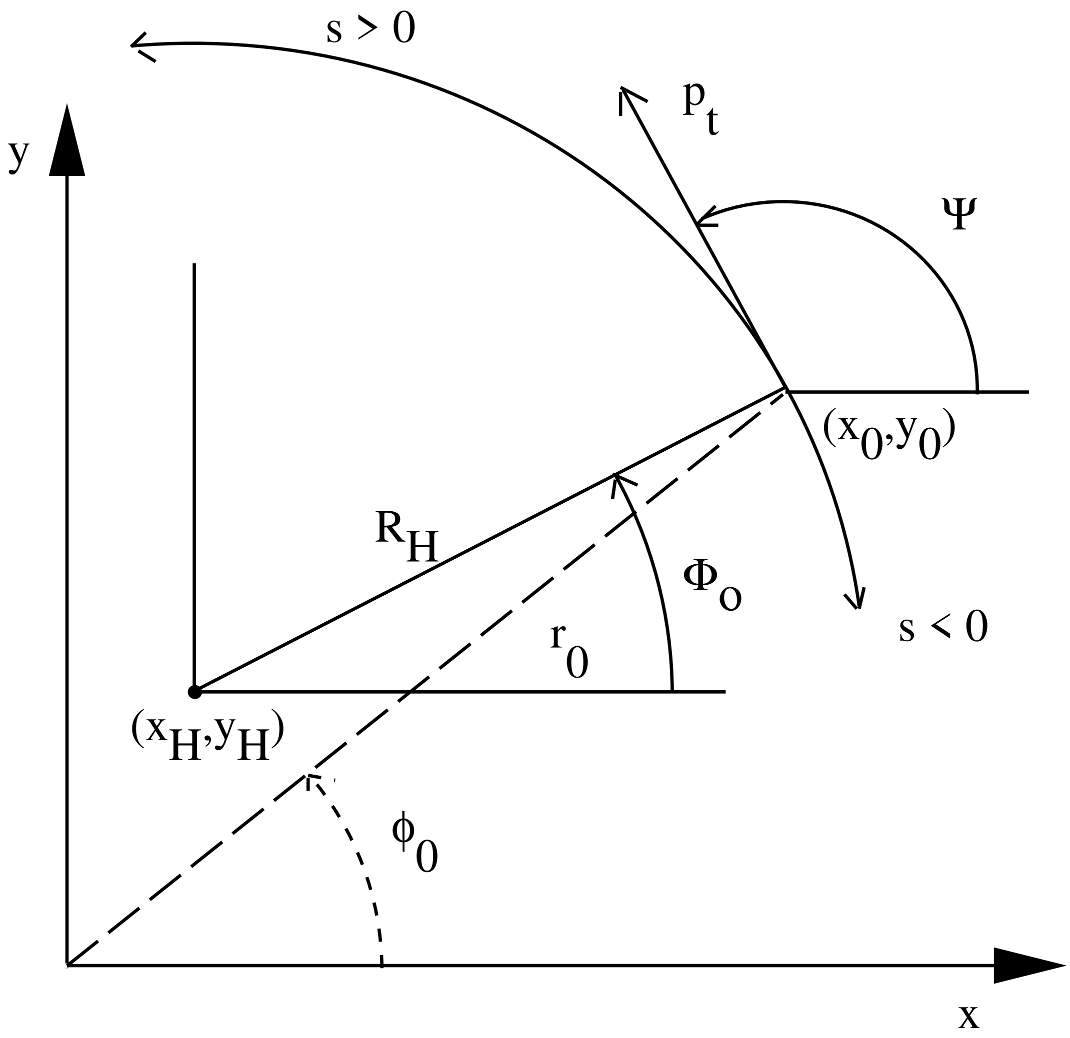


Figure 4.1: *The helicoidal track model for the TPC*

In specific, Figure 4.2a shows a cross sectional view of roughly one fourth of the hits in the TPC and Figure 4.2b shows the same event from a top view. The figures illustrate the tremendous density of the full events and the real complexity of any kind of traditional pattern recognition in such environments. The next two figures show the momentum and radial spectrum of the event which gives a good estimate of the expected radius of the track templates in the range 2 to 16 meters.

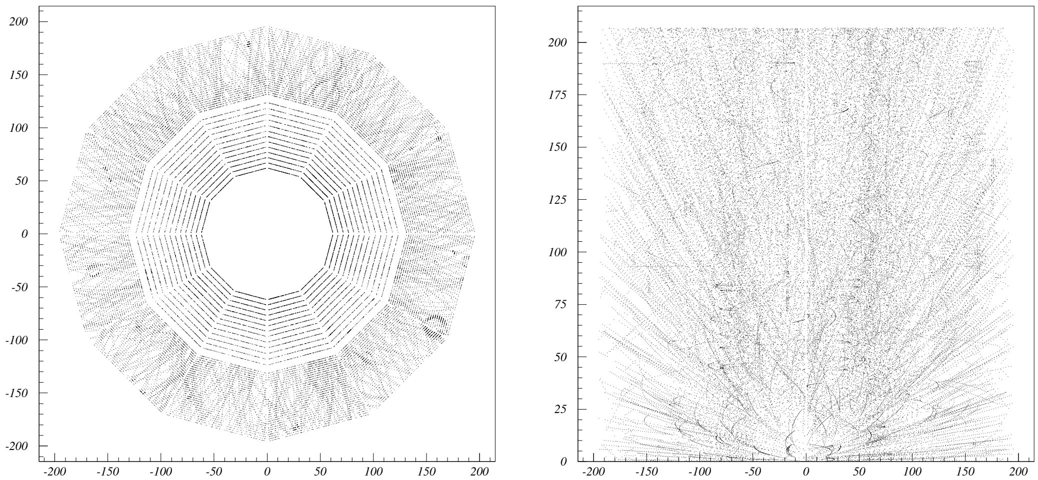


Figure 4.2: *(a) X-Y projection of TPC tracks (b) Y-Z projection of TPC tracks*

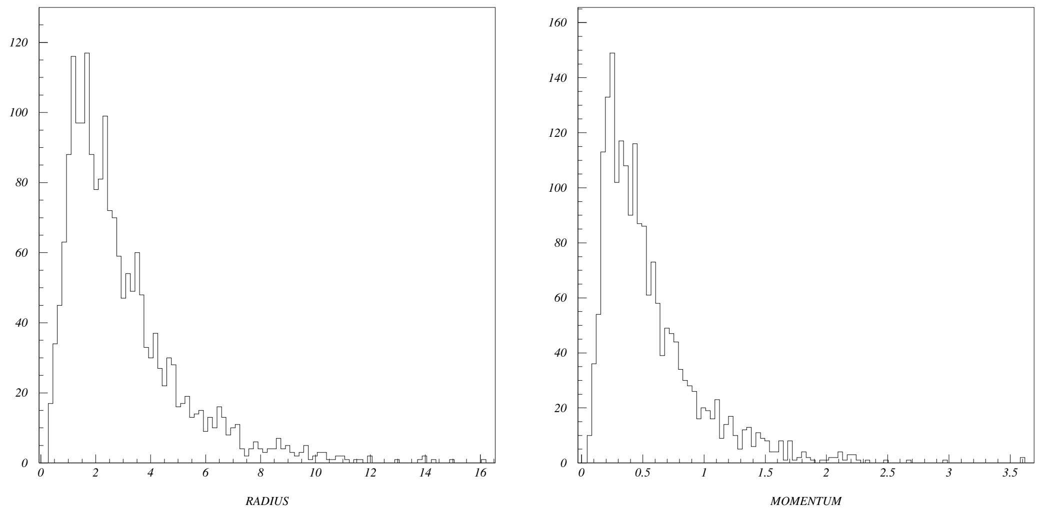


Figure 4.3: *(a)Transverse radii in meters (b) Track energy spectrum in Gev*

Chapter 5

The Tracking Algorithm

5.1 The Modified Hough Transform

The tracking tool we adopt is the Hough Transform [8] which is extremely popular in Computer Vision and other related pattern recognition areas. It has wide applications and is basically a histogramming or binning technique where the data points are used to gather votes for the existence of likely tracks and such vote peaks are then identified in the (track's) parameter space.

The technique therefore involves two parts. We first *transform* or map the data points to some suitably chosen "parameter-space" such that all points having identical values of the parameters of our interest are grouped or clustered together. In the second part this parameter (or feature) space is scanned and thresholded to identify the *peaks*, or high density clusters. For a good choice of parameters and bin size such tall peaks can clearly show the different "likely" tracks that best fit our original data. For a good choice of transform space binning, tracks should map to a single point, collect many votes and should thus be easily separable.

For perfect analytical curves the transform is trivial and is a mere change in the curve's representation. However for noisy and imperfect models the best fit parameters can be identified by looking in the parameter space, for the most likely values of the model being voted on. In particular we develop a **local** Hough transform, that is one which uses local features for voting based on some model rather than a single global entity like energy.

We proceed along a broadly similar line to Ohlsson's approach with a concentration on the local aspect of the Hough transform (for good signal clarity) and then also decompose the problem into independent sets for parallel analysis which can then followed (if needed) by a global merging routine.

This method is based mainly on extremely simple (nearly trivial) calculations which are performed on every pair of space points. Since we assume a helix through the origin, any two hit points and the fixed vertex (origin) always yield exactly one center for the circle through these three points Figure[5.1]. We then add a vote to the corresponding bin in a vote map. As all pairs are covered, likely circles have their centers voted on by many hit points that lie along it. Such peaks in the voting (parameter) space are then isolated by thresholding to identify the exact radius and positions of the likely circles in this sub event.

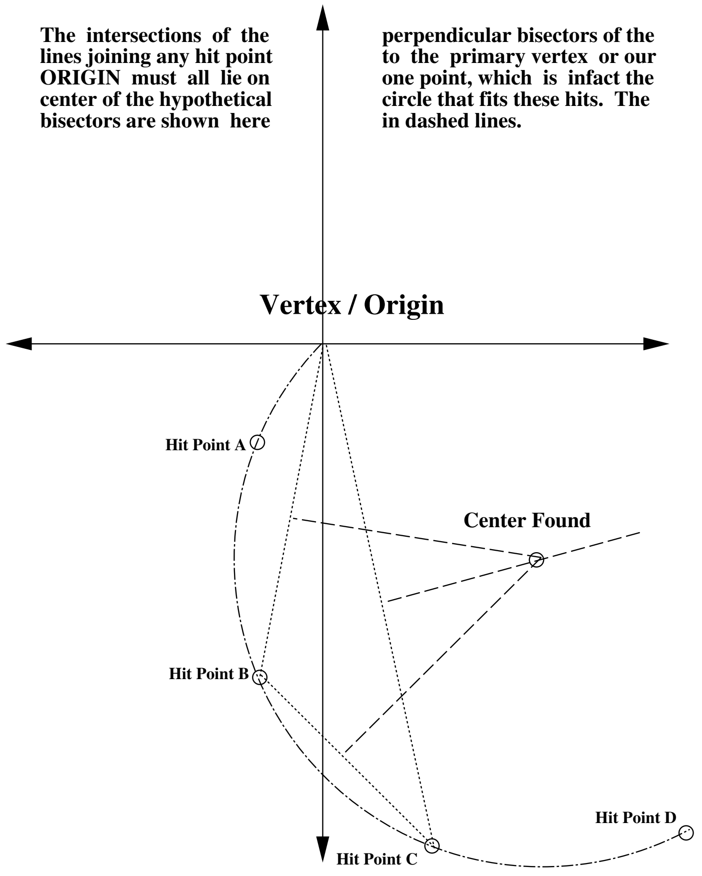


Figure 5.1: *The local Hough Transform : Center finding in X-Y space*

5.2 Implementational caveats

The key features here are the very high speed of performing these set of very trivial calculations against their possibly very large number. The cost actually grows as ${}^{N}C_2$ which warns of a combinatorial explosion. However this can be easily reduced by the following few precautions:

(i) Intersect only a pair of points that are not on the same pad layer. Tracks generally move *across* the layers and not *along* them. Hence points on the same pad-row need not be considered together.

(ii) Intersect only those points that have small angular separation (15 degrees say). This is based on the logic that points very far off from each other are not likely to be on the same track. A track does not generally bend over a large angle for any two hit points. So "close" neighbors are primarily targeted.

(iii) Intersect only those points that do not form a straight line with respect to the primary vertex. Such points can be easily eliminated by a collinearity test and cannot lie along a circle passing through the origin (vertex).

These measures lead to about halving the number of possible calculations with simple checks, before locating the center of the hypothetical circle through any pair of hit points. One hidden benefit of the combinatorial nature of the algorithm however is its extremely high robustness. Even for about 20 hit points along a track this method then picks up 120 votes (about 50% of the possible maximum of ${}^{20}C_2$) instead of the just 20 votes (maximum) for a straight forward (non local) Hough Transform. The signal to noise ratio is shown as the peak separation in figure 6.2a, and is extremely clear.

The distinction stems from the fact that in an "ordinary" Hough transform every point would vote for ALL possible circles that it (and the origin) could lie on. In theory this is actually an infinite number of circles. In the local transform we use a **pair** of points and the origin to vote for exactly **one** circle that fits these three points. The combinatorics involved in the pair-wise mapping of points is what causes the growth to be ${}^{N}C_2$. The gain is in terms of robustness in picking up tracks and the quality of peak separation that is achieved.

5.3 Pitch Angle Decomposition

We see in figure 4.2a that the density of the event in the (x-y) projection is so high that normal pattern recognition cannot be directly performed. In fact even the human eye cannot detect tracks here. In low density traditional tracking the search for tracks could thus be accomplished using the (x,y) $\rightarrow$ (r,ϕ) mapping in the transverse plane alone. On the other hand the Au + Au event size for RHIC precludes such simple 2D analysis. We must therefore use other information (or some function of the 3rd dimension) for suitably splitting the problem. The second objective is high speed and for this too we would like a "cut" that produces nearly identical data sets that can then be all simultaneously handled by the same track extraction algorithm, probably running in parallel, on multiple CPUs.

One such feature that naturally arises is the "dip-angle" of a track, or its *lambda*. This dip or pitch angle is defined as :

$$tan(\lambda) = P_t \div P_z \tag{5.1}$$

where P_t and P_z are the transverse and axial momentum of the track being considered and $(x,\ y,\ z)$ are the space coordinates of a hit. P_t here is the vector sum of the momentum in X and Y directions :

$$P_t = \sqrt{P_x^2 + P_y^2} \tag{5.2}$$

The track's λ (pitch angle) can also be however approximated by λ':

$$tan(\lambda') = z \div \sqrt{x^2 + y^2} \tag{5.3}$$

Although this is only an approximation to the actual value of *lambda*, is generally quite close to the real quantity. In fact the "walk" in "*lambda*" ($\lambda - \lambda$')is less than two and half degrees for most tracks as seen from figure 5.2. The *lambda* spectrum and the corresponding change in *lambda* or $\Delta\lambda$ values are plotted in figures 5.2a & 5.2b respectively.This implies that a typical event having a range of (about) ± 45 degrees could be split into roughly 40 *"lambda windows"* each of which could then be handled in parallel.

5.4 On choosing the angular windows

This partition in λ-angle space is critical to the algorithm's performance and effectively determines the degree of parallelization and hence the speed. The smaller the value of $\Delta\lambda$ (or window width) the greater the number of parallel windows and speed increases roughly linearly. With a bigger value of $\Delta\lambda$ each window has a greater number of data points, and since the transform is combinatorial, halving the total number of windows quadruples the time per window but overall leads to exactly doubling the total time since the number of windows to be processed have been reduced too. Wide windows therefore slow down the process considerably.

Too narrow a window causes tracks to "bend" out of the window range and appear across the neighboring windows. This splitting of a track may cause it not to have enough hit points in any one slice to be detected or may alternatively lead to duplication as more than a minimal threshold of hit points may appear in two windows across which it lies. The choice of $\Delta\lambda$ must be therefore made judiciously. Figures 5.3a & 5.3b show the accuracy of the λ split for typical TPC tracks.

From figure 5.3b we see that windows wider than 4 degrees are too wide, as no significant gains are made in recovery beyond this point. This is based on the logic that a high percentage of tracks need to be recovered (typically 85%) while minimizing execution time. We see therefore that wide windows have too many data points in them leading to much greater overall time at the advantage of slightly higher recovery. Minimum levels of recovery have already been described so these help fix the window size at about 2.5 degrees below which recovery fraction becomes questionable.

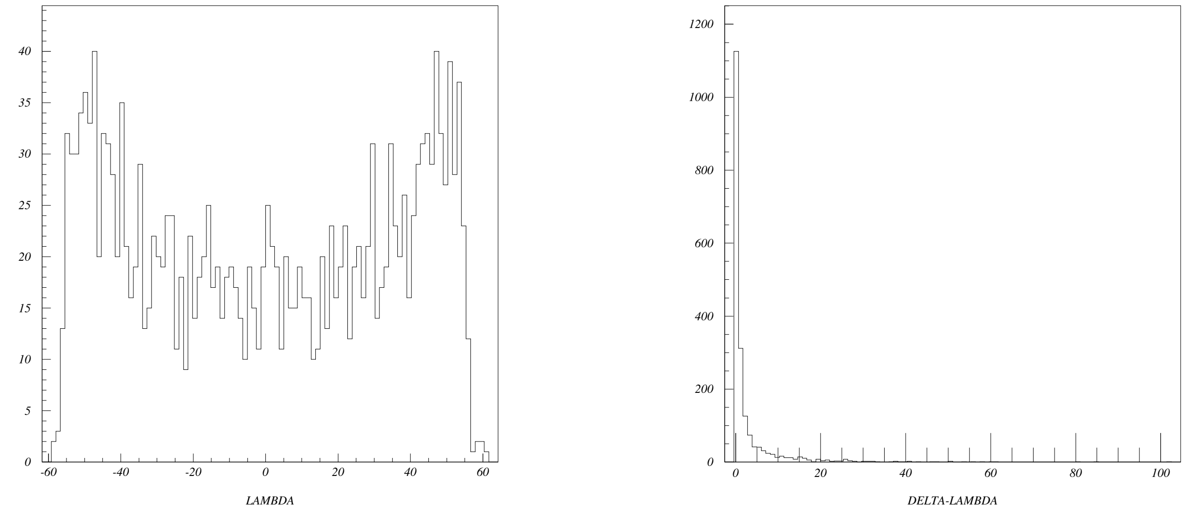


Figure 5.2: *(a) Lambda values for all tracks (b) Walk in lambda for tracks*

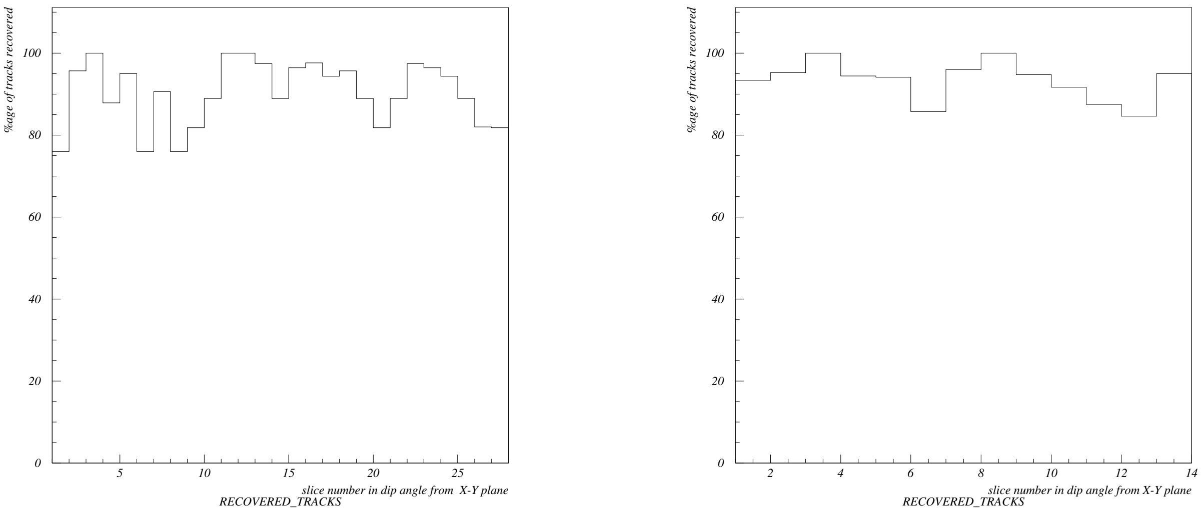


Figure 5.3: *Percentage tracks found (a) 2*deg *windows (b) 4*deg *windows*

5.5 The Overall Structure of the Software

Our track finding algorithm operates on a variety of data sources including output of the GEANT simulation software, and our own slightly simplified models. The key to the approach is to read in the hit coordinates alone (in a simple standard format) and after calculating the local approximation to λ, use this value to put the hit point in one of several possible "windows".

At this stage we also build some calibration files using the additional data on the grouping of points by track number as provided by GEANT. This grouping is of course not visible to the track finder. The calibration table stores the parameters of the two thousand odd tracks that exist in an event and is used later for accuracy calculations.

Having partitioned the data we apply a local Hough Transform on each window in a distributed (over several CPUs) manner. Each Transform calculates probable candidate centers for possible tracks that lie in its window. These are written out to a vote file. A thresholding routine then selects the most probable tracks (centers with high votes) and after an optional merging of closeby low height peaks we have as an output a list of likely track templates. The choice of threshold can be guessed from sample event snapshots such as Figure 6.5b, which shows typical peak heights and peak separation. The efficiency of recovery is also graphed in Figure 6.4 for various thresholds and is discussed in the next chapter.

The raw hit points are then each assigned to one best fit template (closest distance to hit) or are rejected as noise if the fit is not good enough to any of the templates generated, that is if the hit does not lie within some cutoff distance from any track. The final piece of code checks, using the calibration table, on what fraction of tracks in a window were recovered and for more detailed calculations, also computes the fraction of raw hits correctly assigned to the same track.

Presented next is a broad outline the overall structure in Figures 5.4 to 5.6. The flow of data across the modules of code is shown & one thread of the distributed version is illustrated as an example.

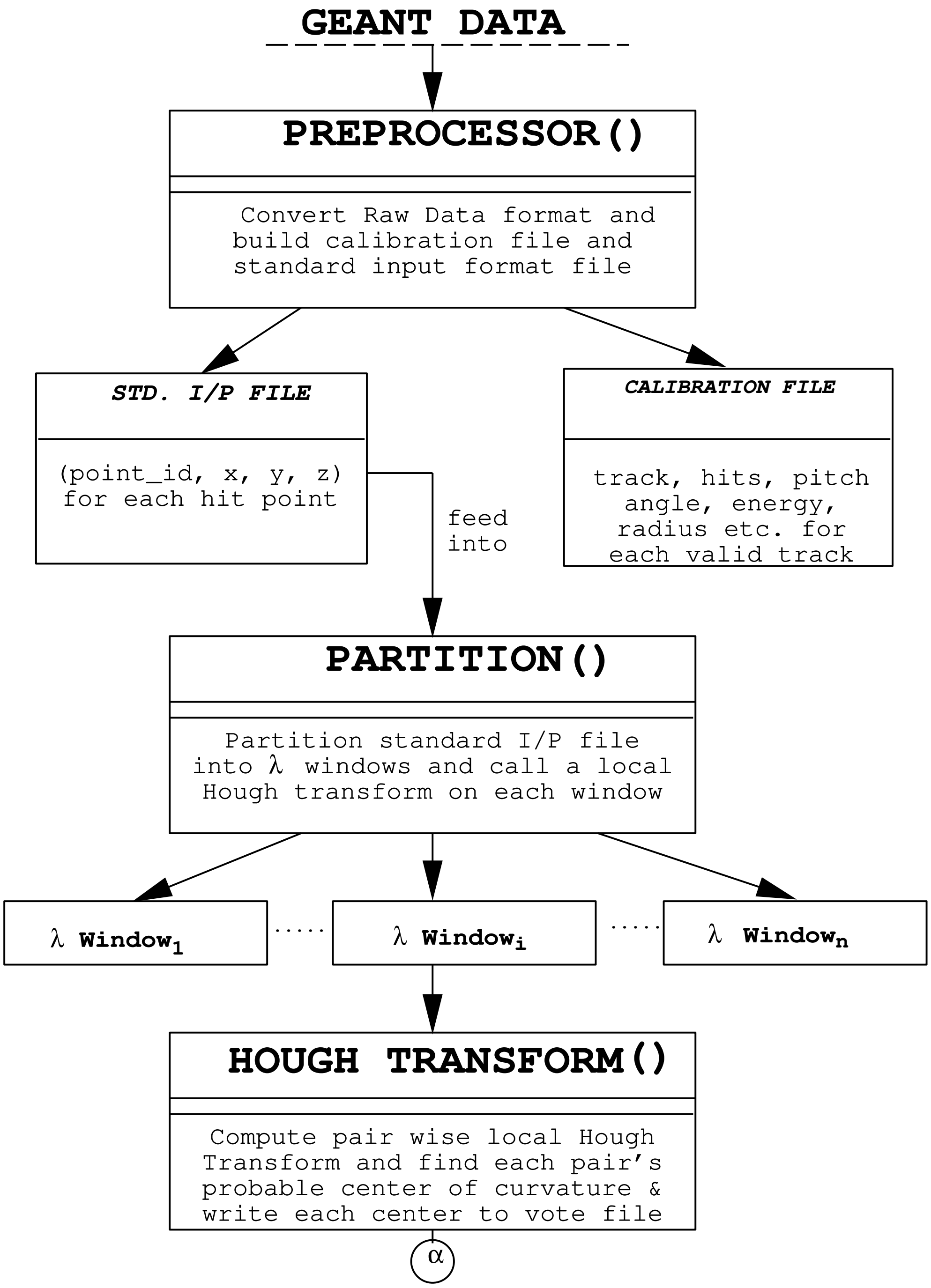


Figure 5.4: *Flowchart Of Overall Structure Of Track Finder : I*

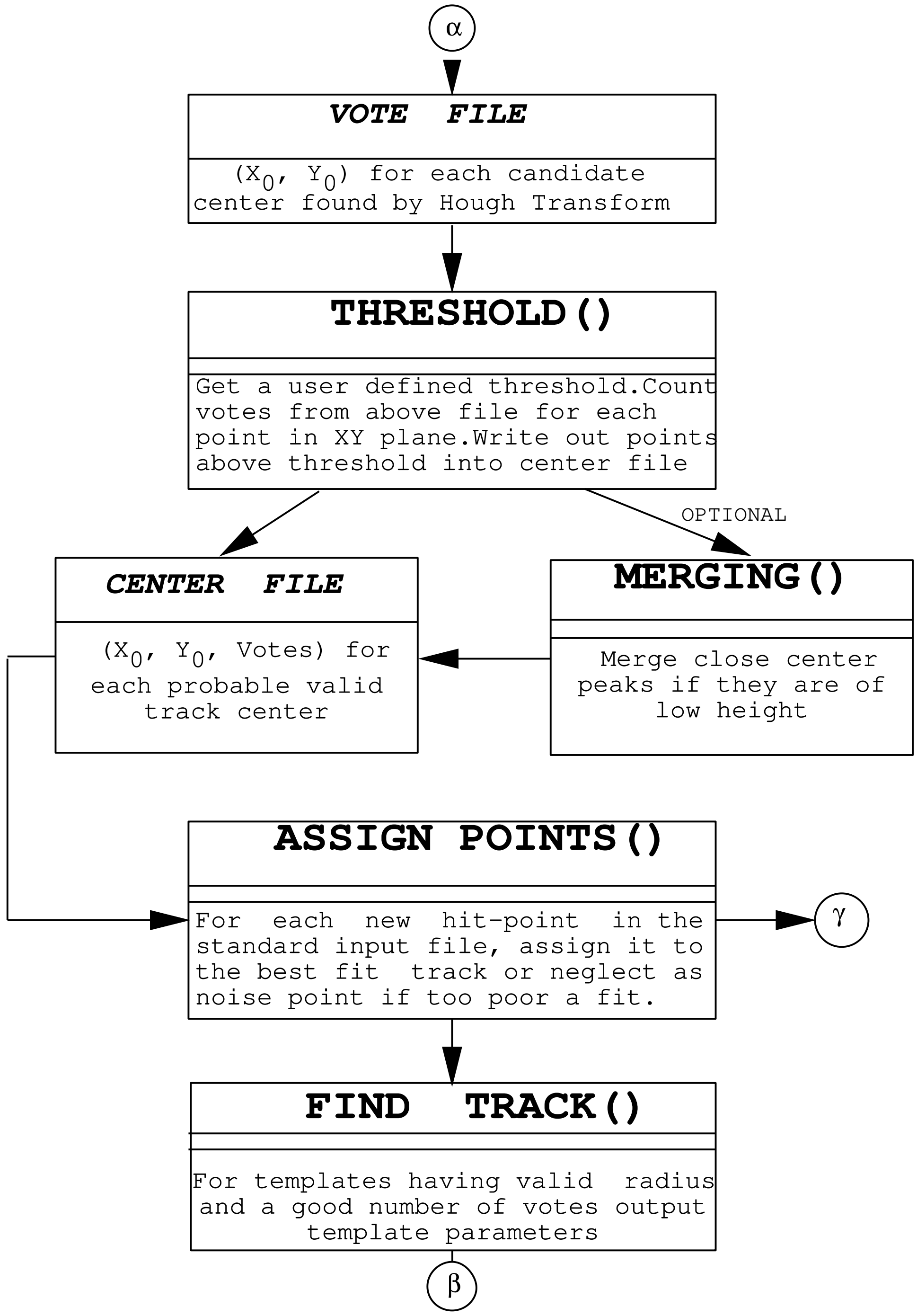


Figure 5.5: *Flowchart Of Overall Structure Of Track Finder : II*

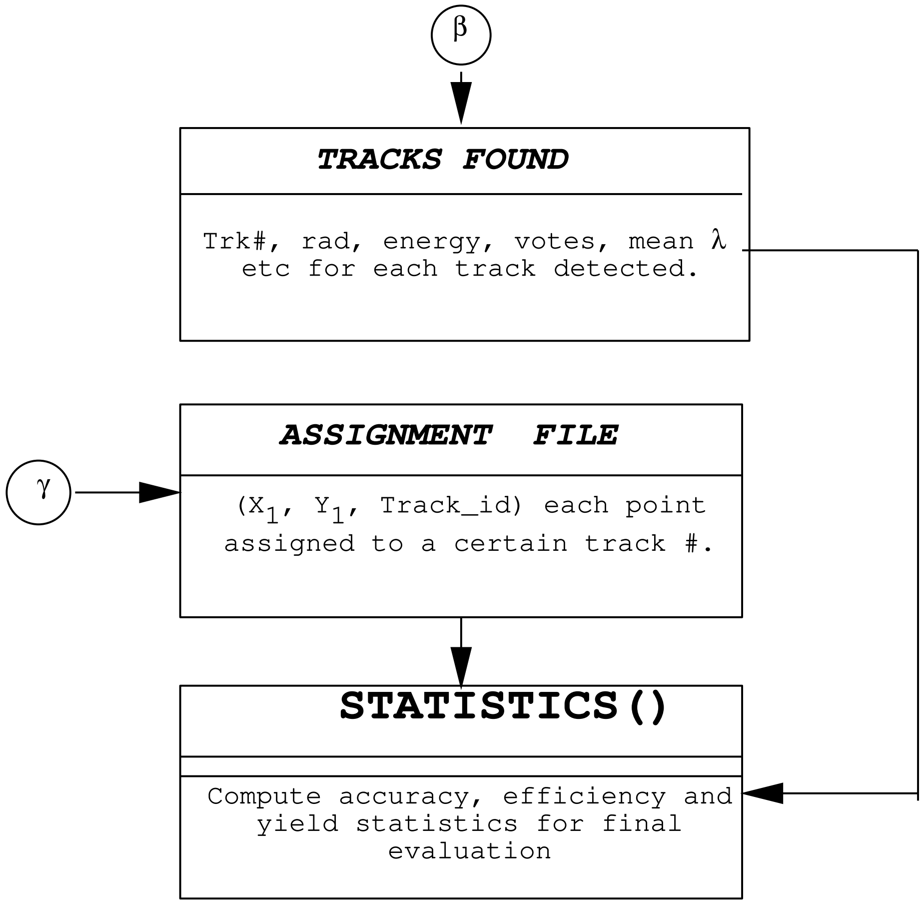


Figure 5.6: *Flowchart Of Overall Structure Of Track Finder : III*

Chapter 6

Performance evaluation & results

6.1 Numerical Simulations

We studied the performance of the algorithm outlined in the previous sections with two data models. For the initial testing and development phase of the transform a simple model with a selectable number of tracks with the required curvature, number of hits and "smear" of the point's measurement was used. A full characterization of the potential of the algorithm with the real TPC data however requires a more realistic simulation of the detector output.

Such simulations were performed with the STAR implementation of GEANT ref[23]. GEANT is used to propagate the particles produced in Au + Au collisions modeled with FRITIOF. It takes into account the decays, scattering and interaction of particles within the detector. A TPC response simulator was also used to include the effects of charge drift and dispersion in the TPC. We used over twenty large events simulated by GEANT and various slices generated by the toy simulator for our study. Various factors such as track recovery, peak separation, radial efficiency etc. were evaluated and typical results are reproduced in the next few sections.

6.2 Track extraction efficiency

Figure 6.1a shows the simplified tracks of our model (in projection). There are 30 tracks ranging in radius from 1 to 10 meters (or about 0.5 Mev/c to 3 Gev/c) with 30 hits per track. A smearing of $\pm 5mm$ has been applied to each hit point's coordinates. Figure [6.1b] shows the actual centers of curvature used to generate these semicircular tracks. This is the typical toy data on which a transform is done.

Figures 6.2a & 6.2b show the peaks in the Hough space as detected by the Hough transform and the corresponding centers identified from these. This set of figures (6.2a & 6.2b) are presented to demonstrate the quality of our Hough transform and its extremely high efficiency in locating tracks in a slice of the event as 100% centers are correctly reconstructed. The positional accuracy of the centers calculated (against their fraction recovered) is also fairly high. There is an error in the range of $\pm 4cm$ which is equivalent to an accuracy of about 96% assuming the average track radius to be 100 cms. The next two figures 6.3a & 6.3b show the quality of "point-assignment" to the track parameters calculated. The first Figure shows the percentage of hits correctly assigned to tracks at a bin size of $\pm 2cm$. The next Figure (6.3b) shows the percentage of correct assignments for the same sample at $\pm 4cm$. We see that using the simplest distance based assignment criterion most points are mapped directly to the correct track. Average accuracy exceeds 95% and any kind of χ^2 fitting can easily refine these nearly complete assignments.

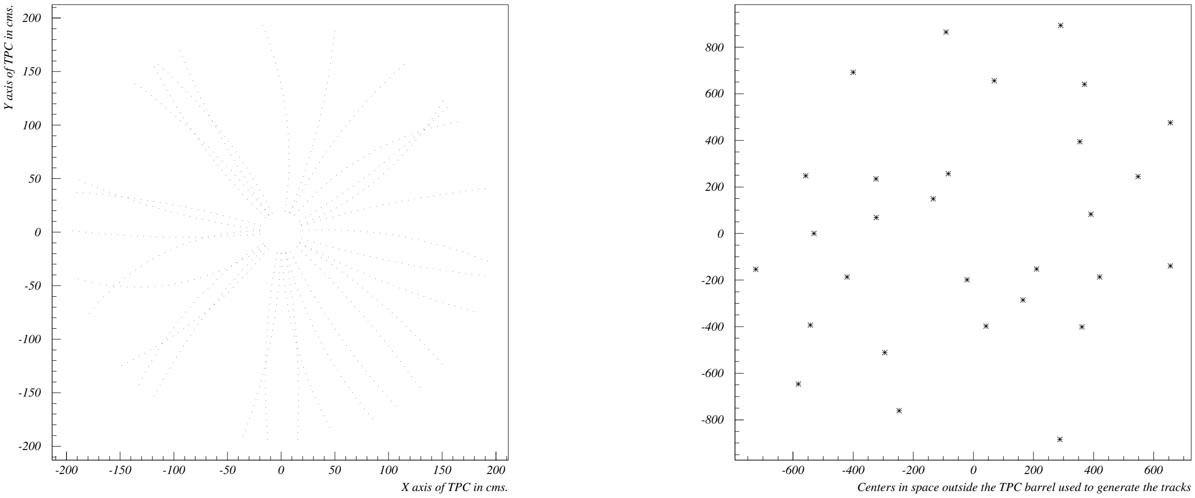
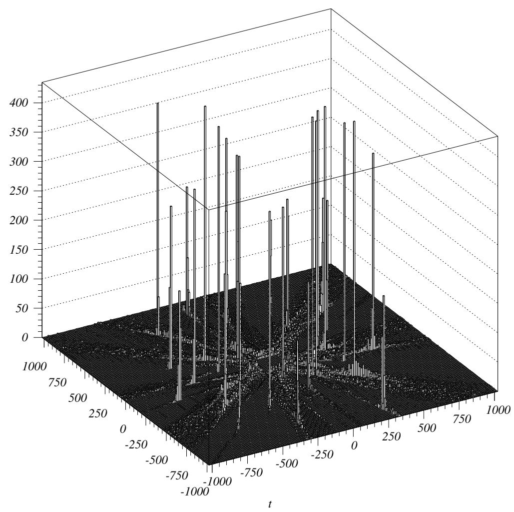


Figure 6.1: *(a) The simplified model 30 tracks (b) The 30 centers for the helices*

The positional accuracy of the centers calculated (against their fraction recovered) is also fairly high. There is an error in the range of $\pm 4cm$ which is equivalent to an accuracy of about 96% assuming the average track radius to be 100 cms. The next two figures 6.3a & 6.3b show the quality of "point-assignment" to the track parameters calculated. The first Figure shows the percentage of hits correctly assigned to tracks at a bin size of $\pm 2cm$. The next Figure (6.3b) shows the percentage of correct assignments for the same sample at $\pm 4cm$. We see that using the simplest distance based assignment criterion most points are mapped directly to the correct track. Average accuracy exceeds 95% and any kind of χ^2 fitting can easily refine these nearly complete assignments.



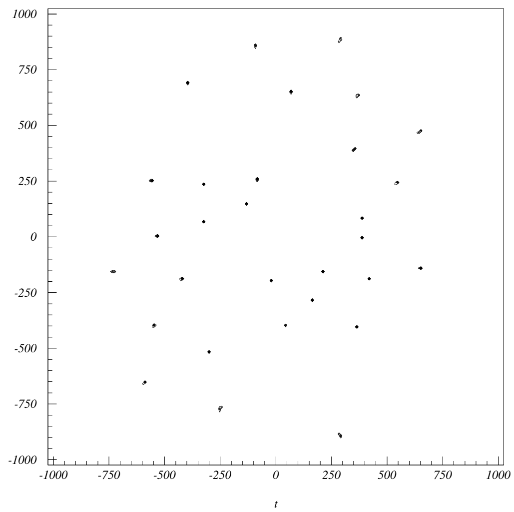


Figure 6.2: *(a) The Hough Transform of 30 tracks (b) All 30 centers recovered*

The other criterion of interest to us is the choice of threshold in identifying peaks in the Hough space. This is a tricky issue. Too low a choice brings in spurious tracks and too high a choice loses some tracks which may have fewer hits and hence fewer votes. In our simulations we expect 20 to 45 (maximum) hits. Based on this, the range of votes should be somewhere between 120 to 150. The recovery rate for various thresholds is plotted in figure 6.4. A choice of threshold of 120 votes seems to work well in the typical events that were studied. Increasing the threshold beyond 150 causes “real” tracks to be lost while decreasing the threshold below 90 finds all tracks, but generates too many spurious templates. The time taken for the transform on every slice is also a function of the number of points per slice, hence window sizes have to be minimised.

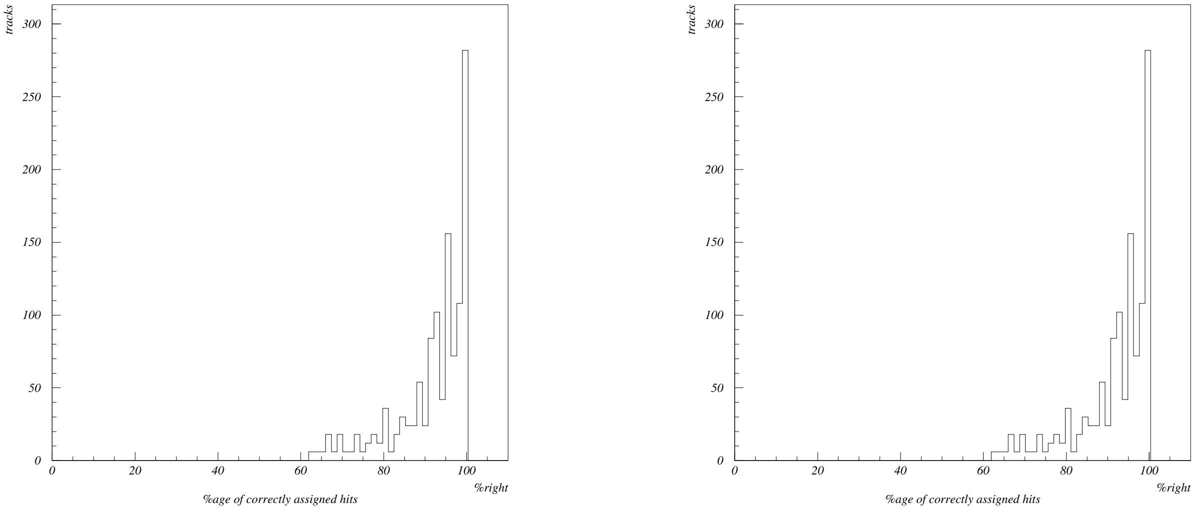


Figure 6.3: *(a) Hits assigned correctly at* $\pm 4cm$ *(b) Hits assigned correctly at* $\pm 2cm$

For typical GEANT events track finding takes 6 to 9 seconds per slice. Hence on a single CPU SGI Indy, 45 slices are done in about 6 minutes . With a multiprocessor system however we can approach the 6 second ceiling for the whole event. In principle, each window can have its own CPU for the transform and hence it takes (in theory) no longer to do the whole event than it does for any individual slice.

The next set of figures 6.5a & 6.5b show a typical 2 degree "slice" of a detailed GEANT simulated event's cut and the lego plot of the peak separation with respect to background after applying the Hough Transform to this sub-slice. In Figure 6.5a we see that after a typical Λ cut a sub slice has recognizable tracks and a scattering of "noise" points to be left unassigned in this slice. The peaks in Figure 6.5b show that all tracks in this slice are found and that their vote (height) separation with respect to the background is very clear.

Figure 6.6a demonstrates the fairly uniform track recovery performance as a function of the track's momentum (or radius). It can be seen that recovery of tracks of various momenta (radius of curvature) is uniformly excellent and is nearly independent of the momentum except for the lower cutoff of about 1m radius, below which the track loops inside the TPC causing it to be missed by the detection scheme. Lastly the execution (CPU) time's dependency on the width of the λ window which is directly proportional to the number of data points in a slice.

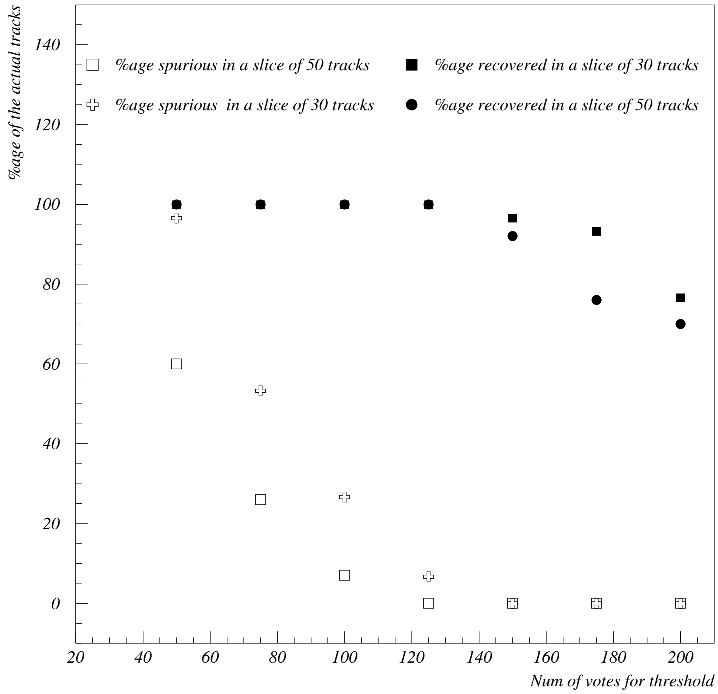
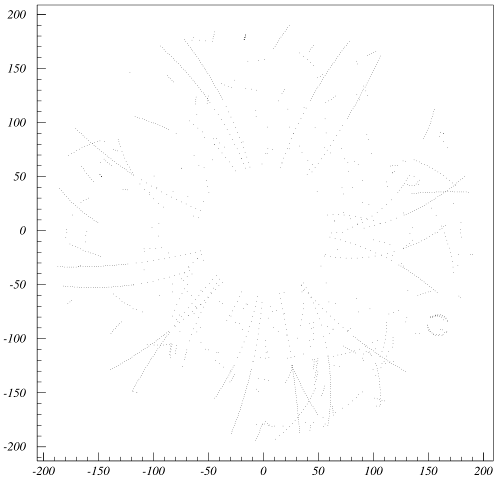


Figure 6.4: *Choosing a threshold for the Hough peaks*



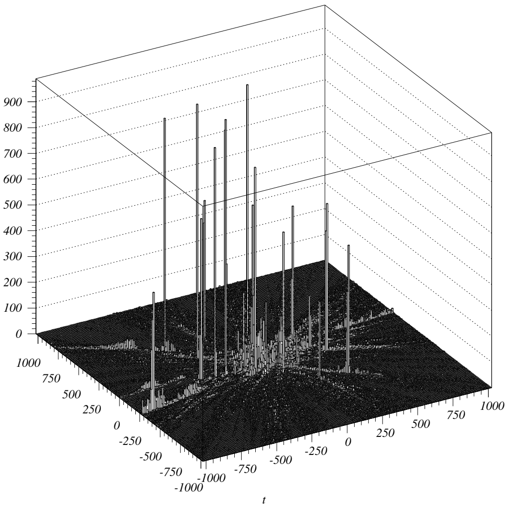
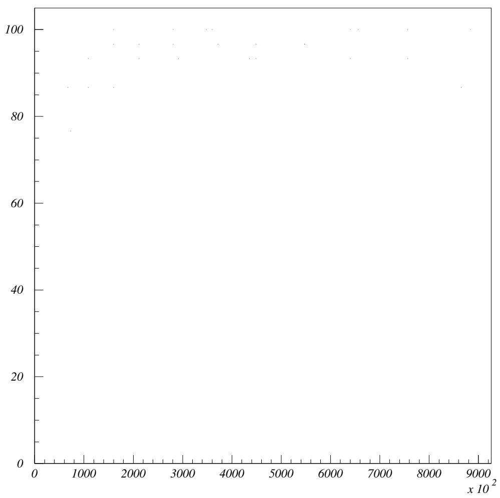


Figure 6.5: *(a) 2* deg *slice of a GEANT event (b) Hough Transform track yield*



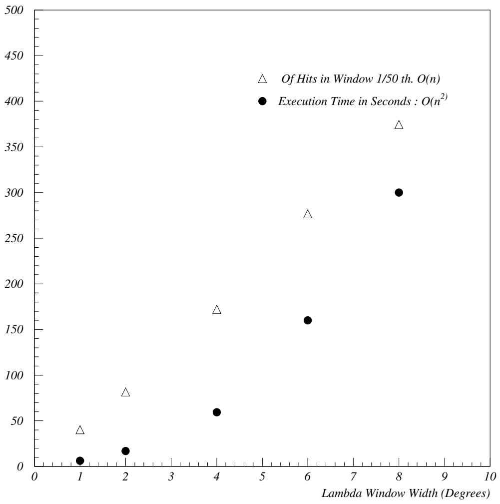


Figure 6.6: *(a) Yield versus radius (b) Speed versus window Size*

Chapter 7

Conclusions and suggestions

7.1 Performance

We have presented in this work a new technique for track reconstruction that is distinct, in as much it is highly distributed and offers a path to very high speed on line triggering for large events. The performance of the technique has been studied extensively on fairly detailed simulations of real events and excellent performance characteristics have been noted in general.

The new algorithm is about one order of magnitude faster than conventional methods and while not as accurate, has a very good yield, in excess of 90% of “findable” tracks which is quite acceptable at the trigger level. There exists a speed versus accuracy tradeoff which has been presented in the last chapter on performance analysis and this can be exploited to suit the aintended application’s needs.

The method’s strength and weakness are tied together in the *lambda* (pitch angle) decomposition, where some tracks are lost, but also where the raw hit data is split into small chunks, all of which can then be parallely processed. We sacrifice some accuracy due to tracks bending across windows and thereby breaking but gain immensely in speed by this tradeoff.

This very technique may in fact be logically extended to include cuts in Φ (transverse plane angle) too, after having made the cuts in λ and can then yield another order of magnitude speedup. The quality of such recovery however remains to be investigated.

Overall good recovery and fair accuracy is achieved in a very reasonable time scale. The algorithm also is distributed and can be automatically scaled to utilize the (serial or distributed) hardware that may be available.

7.2 Scope for future

The method presented in this work thus finds tracks of high energy easily but performance below a few hundred Mevs, equivalent to a rdius of curvature of less than 100 cms may not be good due to such tracks being split across λ windows due to their sharper curvature.

Also post processing and duplicate removal is necessary, should this approach be used for complete tracking as a stand alone application. Nearly complete assignments can however be directly made at this stage too, substantially lowering the burden on any post processor or conventional fitting routine that may then be called.

Such distributed and parallelized techniques we believe therefore, are going to be all the more useful, at the trigger and even higher levels, as their efficiency rapidly increases with the growth in event size and the dropping costs of modern multi processor systems.

Real time requirements will therefore make such approaches increasingly relevant in the near future and we hope to see more effort and developments along the lines of distributed data parallel solutions over the coming years.

References

[1] C.Peterson, "Parallel distributed approaches to combinatorial optimization problems - benchmark studies on TSP",*Neural Comput.* 2 (1990) pp. 261.

[2] B. Denby, "Neural networks and cellular automata in experimental high energy physics",*Computer Physics Communication* 64 (1991) 46.

[3] C. Peterson, "Track finding with neural networks", *Nuclear Instrumentation Methods* A 279 (1989) pp. 537.

[4] G. Stimpfl-Abele and L. Garrido, "Fast track finding with neural networks", *Computer Physics Communication* 64 (1991) 46.

[5] M. Gyulassy and H. Harlander, "Elastic tracking and neural networks algorithms for complex pattern recognition", *Computer Physics Communications* vol. 66 (1991) 31.

[6] A. Yuille, K. Honda and C. Peterson, "Particle tracking by deformable templates", in: Proc. 1991 *IEEE INNS International Joint Conference on Neural Networks*, Seattle, WA (July 1991), vol. 1, pp. 7.

[7] M. Ohlsson and C. Peterson, "Track finding with deformable templates : the elastic arms approach",*Comput. Phys. Commun.* 71 (1992) pp. 77 - 98.

[8] H. Grote, "Pattern recognition in high-energy physics", *Rep. Prog. Phys.*(1987) pp. 473.

[9] M. Gyulassy and M. Harlander, "High Resolution Multiparticle Tracking without Preprocessing via Elastic Tracking",*Nuc. Sci. Div., LBL publ.* 31276.

[10] G. L. Dempsey, "A Hough Transform System Based on Neural Networks", *IEEE Trans. on Indus. Elect.*, vol. 39, no. 6, dec. 1992.

[11] M. Gyulassy and M. Harlander, Elastic Tracking and Neural Network Algorithms for Complex Pattern Recognition, *Nuc. Sci. Div., LBL publ.* 29654.

[12] R.O. Duda and P.E. Hart, Pattern Classification and Scene Analysis (*Wiley*, New York, 1973).

[13] J. Hopfield and D.W. Tank, *Science* Vol 233 (1986) pp. 625.

[14] I. Sakreda and J.T.Mitchell,*Star Note 190.*

[15] *Conceptual Design Report for the Solenoidal Tracker At RHIC.* Berkeley, Ca.: The STAR Collaboration, (June 15, 1992).

[16] John Stachel, "Systematics of Global Observables in Relativistic Heavy Ion Collisions." talk presented at the VIIIth International Conference on Ultra-relativistic Nucleus-Nucleus Collisions. Menton, France, (1990).

[17] Philip R. Bevington,*Data Reduction and Error Analysis for the Physical Sciences.* New York: McGraw-Hill Book Co., 1969.

[18] P.Rehak and E. Gatti, "State of the Art in Semiconductor Detectors." *Nuclear Instruments and Methods in Physics Research.* A289 (1990) 410-417.

[19] B. Jacak, "Transverse momentum distributions of hadrons." *Nuclear Physics.* A525 (1991) 77c-89c.

[20] L. van Hove, *Ann. Phys.* 192 (1989) 66.

[21] E.V. Shuryak, *Physics Review Letters.* 68 (1992) 3270.

[22] R. Bellwied, "Search for Rare Quark Configurations at RHIC." *STAR Note.* 108 (1993) pp. 11.

[23] R.Brun et al., *GEANT User Guide* CERN, (Sept 1987).

[24] *The LUND Monte Carlo Programs.* CERN, (Nov. 1, 1989).

[25] F. James, "Monte Carlo Theory and Practice." *Reports on Progress in Physics.* 43 (1980) 1145-1189.

[26] N.I. Chernov and G.A. Oskov. "Effective Algorithms for Circle Fitting." *Computer Physics Communications.* 33 (1984) 329-333.

[27] C. Boymann, "SVT-TPC Matching Program." *STAR Note.* 83 (1993) pp. 4.

[28] D. Liko, "Track Fitting in the STAR Detector Using the Kalman Filter Method." *STAR Note.* 87 (1992) pp. 15.

[29] T Ferbel ed., *Experimental Techniques in High-Energy Nuclear and Particle Physics.* New Jersey: World Scientific, 1991.

Abstract

DETECTING PARTICLE TRACKS
WITH A
MODIFIED LOCAL HOUGH TRANSFORM

by

AVRONIL BHATTACHARJEE
December 1995

Advisors: Dr. Ishwar K. Sethi & Dr. Claude Pruneau

Major: Computer Science

Degree: Master of Science

A fast track reconstruction algorithm based on a modified Hough transform has been developed and studied for the STAR experiment at RHIC. It is based on a distributed implementation of a local transform of the space hit points to generate track templates. This is done on a set of slices of the event parallely and the results of the individual cuts are merged back again to provide a fairly fast and quite accurate reconstruction of the original tracks. The thesis discusses this and the present day tracking methods and compares the key features of our approach. Performance metrics in terms of efficiency and speed of reconstruction for our algorithm are also provided.

Autobiographical Statement

Avronil Bhattacharjee

Education:

Wayne State University, Detroit, Michigan M.S.(Computer Science): 1995

Punjab University, Chandigarh, India B.E. *Hons* (Computer Science & Engg.): 1994

Experience:

Program Manager : Office Products Division, Microsoft Corporation. (Jan '96 ...)

WSU Graduate Research Assistant: High Energy Nuclear Science Laboratory (Fall 1995)

WSU Graduate Research Assistant: Computer Vision & Neural Networks Laboratory (Fall 1994 - Summer 1995)

Publications and Presentations:

Binary Robot Vision For Tool Recognition : Seventh Indian Engineering Congress, Bangalore, India (1992).

Design and Development of 2-D Vision Software for Robot Guidance : International Conference on Computerization and Automation in Factories, Indian Institute of Technology Delhi (1994).

Finding Particle Tracks With a Modified Hough Transform : Submitted to I.E.E.E. conference on Computer Vision & Pattern Recognition, San Francisco (1996).

Scholarships & Awards:

WSU Graduate Research Assistantship (1994-1995).

Bachelor's Degree Silver Medal, Punjab University(1994).

Punjab University Merit Scholarship(1994).

National School Scholarship(1988).

Silver Medal for Technical Paper Presentation at the Seventh Indian Engineering Congress (student's section), Indian Institute of Science, Bangalore (1992).