DØ Note 3778
September 23, 2000

HTF: histogramming method for finding tracks. The algorithm description.

Alexander Khanov
Kansas State University, Manhattan, KS

Abstract

A histogramming track finder for the D0 tracker is described. The histogramming preselection seems to be an important step of the pattern recognition before applying conventional technique (Kalman filter). The algorithm details are explained and basic implementation features are discussed.

1 Introduction.

The conventional pattern recognition technique used in many high energy physics experiments is Kalman filter based local road finders [1]. While commonly used and very powerful, this technique is not very well suited for complicated cases with $10^4 - 10^6$ hits and high local detector occupancies, because the number of mathematical operations grows very quickly with the number of hits. Another problem is that the pattern recognition becomes involved for detectors without distinct layer structure (like the Silicon Microstrip Tracker of D0), and it is sensitive to the single hit inefficiency. Therefore, a global preselection step that reduces the combinatorics before applying the CPU consuming calculations is desired. As such step, we propose the histogramming approach based on the Hough transform – the method originally invented in 50's for bubble chamber experiments [2].

The method was applied for the pattern recognition in the ATLAS inner tracker [3] where it is the basis of the Level 2 trigger. It was considered for the CMS inner tracker [4]. Similar methods are used for the secondary vertex finding [5].

2 The basic algorithm elements.

Let us consider a particle in homogeneous magnetic field in absence of any material. The trajectory of this particle in the plane perpendicular to the direction of the field is a circle (or, in degenerate case, a straight line) that can be characterized by three parameters, e.g. (ρ, d_0, ϕ), where $\rho = qB/p_t$ is curvature (q is charge, B is magnetic field, and $p_t = \sqrt{p_x^2 + p_y^2}$ is transverse momentum), d_0 is the distance of closest approach to $(0, 0)$ (*impact parameter*), and ϕ is the direction of the track at the point of closest approach to $(0, 0)$.

We will be interested in trajectories with small impact parameters $d_0 \sim 0$. For those trajectories, every pair of points in *coordinate space* (x, y) corresponds to a single point in *parameter space* (ρ, ϕ). If the coordinates are not known precisely, the point in the parameter space becomes a compact area which size can be estimated by means of the error propagation.

The most trivial idea of the pattern recognition comes as follows. Given minimal p_t of the tracks to be found, we can divide the parameter space $-\rho_0 < \rho < \rho_0$, $0 < \phi < 2\pi$ ($\rho_0 = qB/p_t^{min}$) into cells thus creating a 2D histogram. Then, for every pair of hits we increment the contents of cells intersecting with the compact area corresponding to this pair. Since all hits from the same track refer to the same point in the parameter space, the track of n hits will produce a peak of a height of $n(n-1)/2$. Pairs of points from different tracks will give rise to the randomly distributed background. Naively applied, this procedure requires $\sim N_h^2$ operations, where N_h is the total number of hits in the event. This number can be further reduced by organizing detectors in ϕ sectors and considering only hits from nearby sectors.

There is, however, a modification of the above approach known as the Hough transform which improves the situation. We notice that each hit in the coordinate space corresponds to a line in the parameter space (a band, if the errors are taken into account). All those lines intersect at one point which corresponds to the parameters of the trajectory in question. Again, we divide the parameter space in cells, and for every *hit* increment the contents of cells covered by the band to which the hit coordinates correspond. In practice, this means that we make a loop over ρ divisions, for each ρ calculate the covered ϕ region and fill the histogram. In this case the number of required operations is $\sim N_h \times N_\rho$, where N_ρ is the number of divisions in ρ.

The Hough transform can be illustrated by the following example (Fig.1). Let us consider a single 1.5 GeV muon track leaving five hits in the SMT. For a given hit at radius $r = 20\ cm$ there will be a family of circles passing through this hit and the beginning of the (x, y) coordinate system (a). This family corresponds to the line in (ρ, ϕ) variables (b). All five hits from the track give rise to five lines crossing in one point $(\rho, \phi) \approx (13^0, -1/1.5\ GeV^{-1})$ (c). Finally, quantizing the parameter space, one gets a histogram with the clearly pronounced peak (d).

After the histogram is filled one can perform two useful operations:

1. discard cells with less than N_h^{min} hits,
2. discard cells that have all their hits contained in one of the nearby cells.

The conventional next step is to apply a peak finder and find the histogram maxima which would correspond to the tracks. In our case, when N_h^{min} can be as low as 4 and the background level is huge, this step does not seem feasible. Instead, all remaining cells are considered to be track candidates and undergo further filtering. For that purpose they are converted to *templates*. A template is a set of hits organized into measuring planes plus a set of (approximately) known trajectory parameters. A template is typically a candidate for one track. The principal difference between a track and a template is that the latter may have several hits in the same measuring plane, at most one of them supposed to survive after final filtering.

A powerful cleaning procedure which can be done still without bringing the 3D information is the 2D Kalman filter. During this procedure all three parameters of the (xy) trajectory are fitted and the material effects (multiple scattering and energy loss) are taken into account. After the filtering, at most one $r\phi$ hit survives in each layer. The step can produce more than one template.

So far no z information was used. Now we can bring in this information in order to further reduce the number of fake templates. At this moment, the $r\phi$ clusters maybe associated to the stereo ones, therefore giving many combinations per layer. [1] To get rid of fakes we apply another histogramming, now considering the coordinate space (r, z) and the parameter space (z_0, C), where z_0 is the position of the track origin along the z axis, and $C = dz/dr$ is the track inclination in (r, z). This works as soon as trajectories with modest impact parameters are considered.

Another procedure applied is the η-splitting technique which works as follows: we consider two hypotheses ($\eta > 0$ and $\eta < 0$) and for each case leave only hits from the detectors which can be crossed by tracks under given hypothesis. That means, for a fixed r order (increasing r) hits may come from detectors with same or increasing z (for $\eta > 0$) or with same or decreasing z ($\eta < 0$) but may not be mixed.

Having done this second preselection, we are left with templates consisting of hits that are well aligned.

3 The Kalman filter.

The Kalman filter starts from two hits taken from the planes with the smallest number of points and the beam line $(0, 0)$. However, the propagation equations do not impose the beam constraint. The procedure is capable of skipping a hit (accepting a "dummy" measurement). The number of possible continuations is restricted according to the scheme that is a tunable parameter of the procedure. Other parameters include: the maximal χ^2 increment

[1] That is what happens in CFT; in SMT for the moment the 3D hits are used from the very beginning.

due to a measurement (varies); the cut on the maximal number of lost hits (1); the cut on the minimal number of found hits (4 in SMT and 7 in CFT). The procedure takes into account the material effects.

4 The strategies.

Presently, two strategies are implemented. The tracks found by both methods are put together and duplicated tracks are discarded.

Strategy A begins with hits in SMT:

- take SMT hits, perform the $(x, y) \to (\rho, \phi)$ Hough transform;
- clean the template sample, keep templates with at least $N_h^{min}(SMT) = 4$ hits;
- for every template, perform the $(r, z) \to (z_0, C)$ Hough transform;
- clean the template sample;
- apply the η-splitter: for each template consider two hypotheses ($\eta > 0$ and $\eta < 0$) and make templates out of the hit subsets compatible with either of those hypotheses;
- using the 3D Kalman filter, build the SMT tracks;
- if applicable, extrapolate them to CFT and build complete tracks.

Strategy B begins with hits in CFT:

- take 2D CFT hits, perform the $(x, y) \to (\rho, \phi)$ Hough transform;
- clean the template sample, keep templates with at least $N_h^{min}(CFT) = 7$ hits (out of 8 maximum possible);
- apply the (x, y) Kalman filter and build 2D CFT tracks;
- for every track, make all possible 3D hits out of 2D hits contributing to the track;
- perform the $(r, z) \to (z_0, C)$ Hough transform;
- clean the template sample;

- apply the (r, z) Kalman filter and further clean the sample;
- using the 3D Kalman filter, build the CFT tracks;
- extrapolate them to SMT and build complete tracks.

More strategies seem promising but have not been implemented. One interesting approach would be to begin with 2D hits from both SMT and CFT systems, and having found $r\phi$ tracks, continue with 3D. It is not clear if this will work well for the forward area, where no CFT support is available.

For the moment, strategies A and B are run independently and duplicates are removed. This presumes a double calculation which could be avoided by dividing the whole parameter space into "regions of interest".

5 Fake tracks and ghosts.

As long as monte carlo events are considered, the reconstructed tracks can be associated to the ideal (simulated) tracks. This association can be done in different ways, e.g. comparing parameters of the reconstructed and simulated tracks. It turns out that two or more reconstructed tracks can be associated to the same monte carlo track; we will call these tracks *ghosts*. On the other hand, some tracks may not be associated to any monte carlo tracks at all; those tracks are called *fakes*. In real experiment, the actual number of fakes and ghosts cannot be calculated.

One of natural sources of the fake and ghost tracks is the ambiguities in stereo matched hits in the SMT. E.g., let us consider tracks of four SMT hits, two of them are coming from single sided silicon sensors (without stereo information) and two are coming from stereo detectors. If occupancy is large then one $r\phi$ cluster can be ambiguously matched to several stereo clusters. Any of these stereo matches gives rise to a perfectly valid track.

In order to reduce the number of fake and ghost tracks, some precautions have been taken. First, no tracks with less than two SMT hits are considered. Those tracks, even if real, typically have quite poor impact parameter resolution and are of no interest for applications like b-tagging. Second, if a track found in the SMT is extrapolated to the CFT and no CFT hits are attached although it does point to the CFT layers, such track is rejected. The same selection is applied for tracks found in the CFT and extrapolated to the SMT.

6 Conclusions.

An alternative preselection technique, employing Hough transform and histogramming, has been developed for the pattern recognition in DØ. At the moment, both SMT to CFT and CFT to SMT strategies are used. Performance studies will follow.

References

[1] R.Frühwirth, NIM A262 (1987) 444

[2] P.V.C. Hough, Machine Analysis of Bubble Chamber Pictures, Int. Conf. on High Energy Accelerators and Instrumentation, CERN, 1959

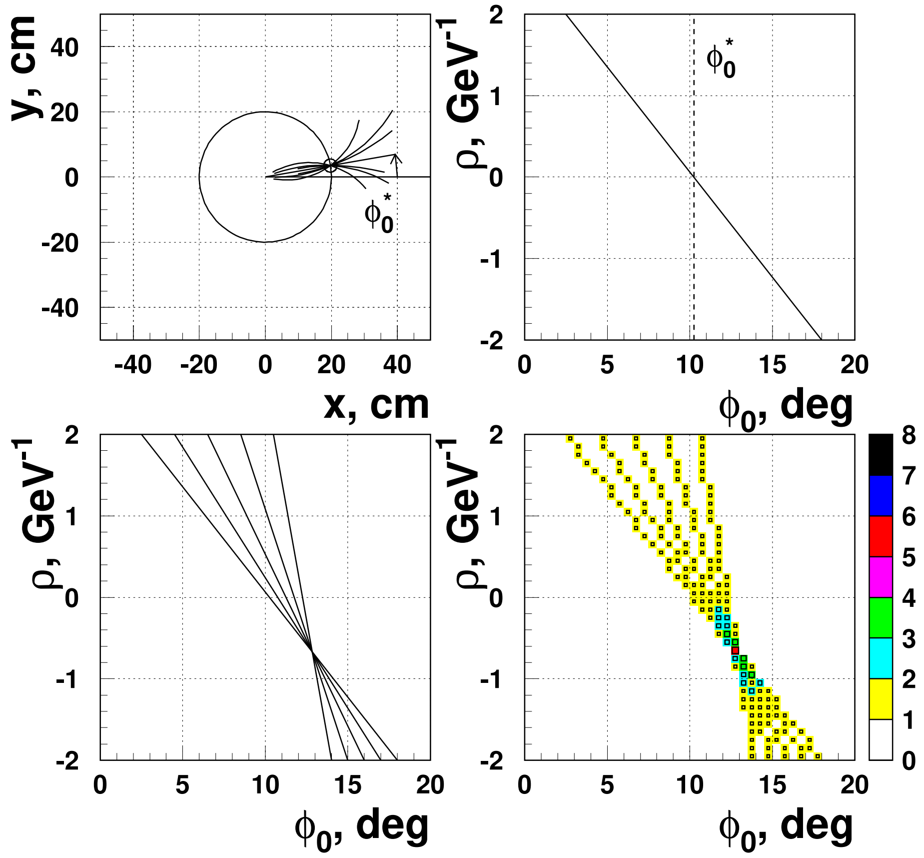
[3] J.Baines, R.Dankers, S.Sivoklokov, ATL-DAQ-99-013

[4] S.Banerjee, F.Bruyant, S.Piperov, CMS TN-1995/128

[5] G.Brooijmans, D0 note 003669

Figure 1: The histogramming technique shown for an example of a single 1.5 GeV muon track of 5 hits. (a) The family of trajectories containing a given hit. (b) The geometric place of all trajectories containing a given hit in parameter space. (c) Curves from different hits intersect at one point corresponding to the track parameters. (d) The point of intersection can be seen as a peak in the (ρ, ϕ) histogram.

(a) (b)



(c) (d)

A The algorithm parameters.

The principle tunable parameters of the algorithm are listed below. As the code development continues, most of them will become the run control parameters (*rcp*). Shown in parentheses are the default values plugged in for the moment.

- P_t^{min}: minimal transverse momentum of tracks to be found, GeV (0.5)
- η^{max}: maximum pseudo rapidity of tracks to be found (4)
- N_{SMT}: the minimum number of SMT hits in a track for the $SMT \to CFT$ strategy (4)
- N_{CFT}: the minimum number of CFT hits in a track for the $CFT \to SMT$ strategy (7)
- $\Delta\chi^2/NDF$: the maximum χ^2 increment due to a measurement (50.)
- N_{lost}: the maximum number of hits lost during the propagation (1)
- N_{br}: the maximum number of possible continuations (5)

One should take into account that 1) these values are context dependent and may vary during the procedure; 2) the cuts on track parameters, like P_t^{min}, do not imply that tracks below P_t^{min} will not be found; 3) there are many more less important parameters not described here.

Another possibility to control the code is to choose between the operation modes. Below are listed main controls of this type.

- The strategies ($SMT \to CFT$ and/or $CFT \to SMT$).
- The *trf++* refit.