New developments in track and vertex reconstruction

Talk given at DESY/Universität Hamburg, 24.01.05
Are Strandlie,
Gjøvik University College, Norway
Outline

- Introduction
- Review of traditional track reconstruction algorithms
- New developments
  - Gaussian-sum filter for electron reconstruction
  - Gaussian-sum filter for vertex reconstruction
  - Deterministic Annealing Filter for combined track finding and fitting (very preliminary – work in progress)
- Summary and conclusions
Introduction

Recent developments in track and vertex reconstruction

R. Frühwirth, CMS

Institute for High Energy Physics
of the Austrian Academy of Sciences
Vienna, Austria

DESY, Hamburg, January 13, 2003

A. Strandlie, DESY, 24.01.2005
Introduction

- I will – after a brief review – mainly focus on what has happened after Rudi’s talk
- The amount of topics will therefore be limited
- Hopefully at least some of the topics can be covered in enough detail to be understandable
- The main focus will be on track reconstruction, except one example of a vertex reconstruction algorithm
Track reconstruction is traditionally divided into two separate parts:

- **track finding** or **pattern recognition** (division of full set of hits in a detector into several sub-sets – each subset containing hits believed to come from one particle)
- **track fitting** (**estimation** of track parameters from the track candidates or sub-sets of hits produced by the track finder)

Methods used for these tasks as well as the boundaries between them have continuously changed over the years.
Review

- Track finding in early bubble chamber experiments was done in a purely manual way:
  - events were inspected on a scanning table
  - machines measuring bubble positions on the event image were guided by an operator

- One of the popular, early methods for automatic track finding was the so-called Hough transform
  - finding straight lines by transforming each measurement to a line in a parameter space
  - measurements lying along straight lines in original picture create lines crossing at one point in parameter space

- Tracks were fitted by a global least-squares estimator
**Review**

CERN

2m bubble chamber

*Fig. 1* The CERN 2 m hydrogen bubble chamber. The beam enters from the left. The central feature is the massive iron magnet yoke, the chamber body is inside the magnet. The film magazines for the cameras (then four) can be seen behind the two persons on the lower platform.

A. Strandlie, DESY, 24.01.2005
Review

The famous "scanning ladies"
Review

bubble chamber picture and subdivision into segments

A. Strandlie, DESY, 24.01.2005
Review

Fig. 3 A framelet giving a reasonably complex bubble pattern. The electronically-drawn transform appears at the bottom.

one segment

projection of transform
Review

- New challenges had to be faced by least-squares fitting algorithms when electronic experiments came on stage in early 70’s

- Prime example: Split Field Magnet (SFM) detector at the CERN Intersecting Storage Rings (ISR)
  - precise treatment of material effects in wire chambers was needed
  - also important effects from crossing beam pipe at shallow angles
  - correct approach pioneered by M. Regler (CERN 73-2 1973)

- Track finding was still mainly considered as a separate procedure preceding the track fit
Review

SFM detector at CERN ISR

A. Strandlie, DESY, 24.01.2005
procedure had to be as selective as possible. Besides the scattering on the foam-frame of the proportional chambers, a large contribution to multiple scattering was introduced by the walls of the beam tube. The momentum resolution was about ±5% because of the small sagitta at such high momenta. Therefore, full use had to be made of the constraints provided by the equations of motion in the strongly inhomogeneous magnetic field, and the least-squares sum had to be carefully weighted, considering the chamber resolution as well as the multiple scattering [25].

excerpt from paper by Nagy et al. (Nucl. Phys. B 1979)
In WA13 experiment at CERN, a new formulation of the least-squares method was developed. This method – the so-called progressive method due to P. Billoir – included measurements recursively into the fit (NIM 1984). Estimates of track parameters were updated each time information from new measurement was included.
Review

layout of WA13 experiment

Fig. 6. Layout of the WA13 experiment in the Omega Spectrometer.

Fig. 7. Comparison between the standard fitting method of ROMEO and the recursive implementation of the optimal estimator: distance, at the point of closest approach, between two tracks extrapolated backwards to the vertex.

A. Strandlie, DESY, 24.01.2005
During development of reconstruction software for DELPHI experiment at LEP collider at CERN, it was realized by R. Frühwirth that progressive method was equivalent to Kalman filter (NIM A 1987).

Immediate implications:

- existence of Kalman smoother
- easy to obtain optimal estimates anywhere along the track
- enables efficient outlier rejection, as track parameter predictions are calculated from all other measurements in a track.
Due to the recursive nature of the Kalman filter, it was soon realized that it could be used for track finding and fitting concurrently.

Given an estimate of the track parameters of a segment of the track, hits inside a search window compatible with the prediction into the next layer are considered for being included in the track candidate.

No more strict separation between these procedures.

Application of Kalman filter as track finder was pioneered by P. Billoir in DELPHI experiment at LEP (Billoir, Comp. Phys. Comm. (CPC) 1989).
Review

- It was tried out on real data for the first time in ZEUS experiment at DESY (Billoir and Qian, NIM A 1990)
- The method was generalized by R. Mankel to possibly propagate several track candidates concurrently in HERA-B experiment at DESY (Mankel, NIM A 1997)
- Is today by far the most widely used method for track finding and fitting
Review

principle of method

result from event in DELPHI TPC

Fig. 1. Principle of progressive track recognition.

Fig. 3. Reconstruction of tracks in a simulated event: (a) space points; (b) tracks found.
Review

reconstructed event from ZEUS

Fig. 7. Graphic display for a reconstructed event. (a) Input points: dark dots = real points; light dots = image points; crosses = backgrounds; pluses = beam-pipe centre and beam positions. (b) Reconstructed event: solid lines = reconstructed tracks; dotted lines = cell walls.
Main Tracker in HERA-B

principle of concurrent track evolution

Fig. 3. Illustration of the concurrent track evolution strategy for an ambiguous situation caused by three nearby tracks in a superlayer of the HERA-B outer tracker. The propagation proceeds upstream from the right to the left. We assume that the propagation started with a seed of hits from track T1.
Review

- Track finding and track fitting was traditionally considered as separate tasks.
- In bubble chamber and early electronic experiments, many types of track finding algorithms were used.
- Tracks were in all (or at least most) cases fitted by a global least-squares method.
- Since the invention of the Kalman filter, strict boundaries between track finding and fitting do no longer exist.
Review

- Kalman filter has been the dominating method in track reconstruction for many years

- Most important developments after the Kalman filter:
  - adaptive methods (Deterministic Annealing Filter (DAF), Multi-track Filter (MTF))
  - Gaussian-sum filters (GSF)

- Both of these classes of methods can be regarded as generalizations of the Kalman filter
  - DAF is an iteratively reweighted Kalman filter
  - GSF takes the form of several Kalman filters running in parallel
GSF works by modelling non-Gaussian effects of various kinds as Gaussian mixtures

- measurement errors
- bremsstrahlung energy loss
- multiple scattering
- non-Gaussian tails in track parameter distributions used for vertex fitting
Review

- Adaptive methods work by **downweighting non-Gaussian effects**
  - measurement errors
  - resolution of left-right ambiguities
  - distorted hits
  - wrongly associated tracks in a vertex fit
  - immature convergence of alignment parameters in sequential procedure (with tracks)

- I will in the following concentrate on a few, recent applications of Gaussian-sum filters and adaptive methods
New developments

- An important effect needing treatment during track reconstruction is **energy loss**
- Ionization loss is often regarded as **deterministic correction to track model**
- Electrons need special treatment:
  - suffer from **bremsstrahlung** energy loss
  - **dominant component** of energy loss above ~100 MeV/c
New developments

distribution of relative energy loss

strongly peaked with long tail
New developments

- Kalman filter implicitly assumes that bremsstrahlung distribution is Gaussian
  - since it is known to be optimal only when errors are Gaussian
  - such an approximation is quite crude for the bremsstrahlung distribution
- Plausible that approach which better takes the actual shape of the distribution into account can do better
- GSF algorithm for electron reconstruction uses an approximation of the bremsstrahlung distribution by Gaussian mixtures instead of a single Gaussian
New developments

- In the GSF, track states are described by Gaussian mixtures instead of single Gaussians.
- As with the Kalman filter, the GSF proceeds by alternating propagation and update steps.
- Adding material effects in a detector layer amounts to a convolution of the density of the predicted state with the density describing the disturbance of the track in the material.
- This leads to a potential combinatorial explosion in the number of components in the state mixture.
  - Number of components has to be limited to a tolerable value.
New developments

GSF for electrons

M Gaussians

N Gaussians

N * M Gaussians

Possible combinatorial explosion!

Predicted state

Updated state

Surface with material

Are Strandlie, CERN

b-tau meeting, 25.02.03

A. Strandlie, DESY, 24.01.2005
New developments

- **Gaussian-mixture approximations of the Bethe-Heitler distribution** (well-known, analytical approximation of the bremsstrahlung distribution) have been calculated recently.

- Expressions exist in the published literature (Frühwirth, CPC 2003).

- These have been used in a **GSF implementation** in the CMS tracker at CERN (W. Adam, R. Frühwirth, A. Strandlie, T. Todorov, Proc. CHEP’03, 2003).
New developments

Rz-projection of one quadrant of CMS tracker

A. Strandlie, DESY, 24.01.2005
New developments

The mixture parameters (weights, mean values and variances) have been obtained by minimizing the following distances with respect to the mixture parameters:

- Kullback-Leibler divergence:
  \[ D_{KL} = \int_{-\infty}^{\infty} \ln[f(z)/g(z)] f(z) \, dz \]

- Cumulative distribution function distance:
  \[ D_{CDF} = \int_{-\infty}^{\infty} |F(z) - G(z)| \, dz \]

A. Strandlie, DESY, 24.01.2005
New developments

Distances as a function of radiation length

A. Strandlie, DESY, 24.01.2005
New developments

Distance between polynomial of degree 5 and original fit

Weights, means and variances calculated on the fly in ORCA implementation of GSF. First and second moments are forced to those of the Bethe-Heitler distribution.
New developments

- In order to avoid a combinatorial explosion of components, several strategies were tried out for limiting the number:
  - pair with smallest distance (Kullback-Leibler) repeatedly combined into single component
  - component with largest weight combined with the closest one. Repeated with remaining components
  - components with smallest weights dropped

- First approach turned out to be superior, slightly better than the second

- The third is fast but not precise and does not preserve the first two moments of the mixture, leading to a biased estimate
New developments

estimated distribution of charged, inverse momentum for single track

GSF estimate clearly non-Gaussian!

A. Strandlie, DESY, 24.01.2005
New developments

residuals of estimated with respect to true parameters in a simplified simulation of the CMS tracker.

A. Strandlie, DESY, 24.01.2005
New developments

pull quantities
- not Gaussian either!
New developments

half-widths of intervals covering 68% of normalised momentum residuals as function of maximum number of components kept during reconstruction
New developments

same quantities as previous slide as a function of momentum
New developments

probability transform
(generalization of
chisquare probability)
of charged, inverse
momentum
New developments

FULL SIMULATION

intervals covering 68 % and 95 % of residual momentum distributions as a function of momentum
New developments

same quantities as previous slide but with Gaussian smeared measurements
New developments

Probability transform for charged, inverse momentum
New developments

resolution improves by 20 % for KF and 35 % for GSF
New developments

GSF for vertex reconstruction

A. Strandlie, DESY, 24.01.2005
New developments

- **Kalman filter for vertex reconstruction:**
  - all tracks believed to come from the same vertex are fitted together in a least-squares fashion in order to yield the vertex position
  - the parameter vector for a track is considered as a measurement
  - information from these are introduced recursively, each time yielding an improved vertex position
  - as for the Kalman filter for track reconstruction, the procedure is mathematically equivalent to a global least-squares fit
The Gaussian-sum Filter for vertex reconstruction

- Track parameter error distributions modeled by a mixture of Gaussians
- Vertex State vector \( x \), is also distributed according to a mixture of Gaussians
- Iterative procedure: estimate of the vertex is updated with one track at the time
- Add new track to vertex, each component of the Vertex State is updated with each component of the track (Combinatorial combination of all track components)
- The filter is a weighted sum of several Kalman Filters
  - GSF is implemented as a number of Kalman filters run in parallel
  - The weights of the components are calculated separately
  - Non-linear estimator: weights depend on the measurements
- The new Vertex State \( x_k \) is therefore distributed according to a mixtures of \( N_k \)
  \[
  ( = N_{\text{track} - k} \times N_{\text{vertex} - k-1} ) \text{ Gaussians}
  \]
  - The number of components increases exponentially (combinatorial explosion)
  - The GSF vertex filter shows little sensitivity to the number of components kept
New developments

Simulation

Simplified simulation in a fully controlled environment:

- Tracks generated at a common vertex
- No track reconstruction
- Track parameters are smeared according to known distributions:
  - 2 component Gaussian mixture:
    - Narrow component: 90 % Relative weight
      (Standard deviation of Impact parameter = 100μm)
    - Wide component: 10 % Relative weight
      Std dev. 10x larger (Impact parameter = 1000μm)
      Ratios of Standard deviation = 10
- For the Kalman Filter:
  - tracks smeared according to two-component mixture
  - single component used in the fit:
    - track parameter variance of dominating component
    - estimated position independent of scaling of variance (but not position uncertainty or $\chi^2$)

A. Strandlie, DESY, 24.01.2005
New developments

Kalman Filter fit

Four track-vertex fit with the Kalman Filter:

- Residuals – y coord. (cm)
- Pull – y coord.
- $P(\chi^2) < 0.01$

A. Strandlie, DESY, 24.01.2005
New developments

Gaussian-sum Filter fit

Four track-vertex fit with the GSF (using the full Gaussian mixture)

- Residuals – y coord. (cm)
- Pull – y coord.
- P(χ²)

A. Strandlie, DESY, 24.01.2005
New developments

Robust filters: Adaptive Vertex fitter (AVF)

- Iterative, re-weighted LS fit
- More robust by down-weighting outliers:
  - Weight of a track depends on its distance to the vertex
- Adaptive VF very stable with high break-down point

\[ w_i = \frac{1}{1 + \exp \left( \frac{X^2 - X^2_0}{2T} \right)} \]

- Default implementation: Adaptive Vertex fitter (AVF)
  - LS (Kalman) updater: sensitive only to the core of the distributions
- Adaptive Gaussian Sum Filter (A-GSF):
  - AVF with GSF updater: Gaussian mixture correctly taken into account!
New developments

Robustness tests

- LS estimators:
  - Optimal when no outliers are present
  - Very sensitive to outliers
- What about the GSF?

- Type 1 outliers = mismeasured tracks:
  - $\sigma_{\text{outliers}}/\sigma_{\text{inliers}} \neq 1$
- Type 2 outliers = track from another vertex:
  - 4 tracks from main vertex
  - 1 track from second vertex

Tracks smeared with same mixture of 2 Gaussians
New developments

Type I outliers

4 tracks from main vertex
3 inliers (Filter sees mixture with correct covariance matrices)
1 outlier with $\sigma_o/\sigma_i = 3$

Residuals – y coordinate [cm]

GSF
Without outlier

GSF

Kalman

Adaptive

A-GSF

A. Strandlie, DESY, 24.01.2005
New developments

<table>
<thead>
<tr>
<th>Filter</th>
<th>mean $P(\chi^2)$</th>
<th>Resolution</th>
<th>Pull</th>
<th>C(50%)</th>
<th>C(90%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kalman</td>
<td>0.18</td>
<td>115</td>
<td>1.61</td>
<td>297</td>
<td>326</td>
</tr>
<tr>
<td>GSF</td>
<td>0.37</td>
<td>83</td>
<td>1.11</td>
<td>54</td>
<td>167</td>
</tr>
<tr>
<td>Adaptive</td>
<td>0.18</td>
<td>92</td>
<td>1.34</td>
<td>59</td>
<td>180</td>
</tr>
<tr>
<td>A-GSF</td>
<td>0.24</td>
<td>84</td>
<td>1</td>
<td>55</td>
<td>168</td>
</tr>
</tbody>
</table>

Type 1 outliers

A. Strandlie, DESY, 24.01.2005
New developments

Type 2 outliers

4 tracks from main vertex
1 track from second vertex,
\( \Delta y = 3 \text{mm} \)

GSF
Without outlier

GSF

Residuals – y coordinate [cm]

Kalman

Adaptive

A-GSF

CHEP04 - 30th September 2004 - p. 17
New developments

<table>
<thead>
<tr>
<th>Filter</th>
<th>mean $P(\chi^2)$</th>
<th>Resolution</th>
<th>Pull</th>
<th>C(50%)</th>
<th>C(90%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kalman</td>
<td>0.06</td>
<td>224 bias!</td>
<td>1.24</td>
<td>573</td>
<td>853</td>
</tr>
<tr>
<td>GSF</td>
<td>0.12</td>
<td>63</td>
<td>0.97</td>
<td>42</td>
<td>113</td>
</tr>
<tr>
<td>Adaptive</td>
<td>0.41</td>
<td>69</td>
<td>0.99</td>
<td>46</td>
<td>155</td>
</tr>
<tr>
<td>A-GSF</td>
<td>0.43</td>
<td>63</td>
<td>0.86</td>
<td>42</td>
<td>112</td>
</tr>
</tbody>
</table>

Type 2 outliers

A. Strandlie, DESY, 24.01.2005
New developments

DAF for combined track finding and fitting
New developments

- Deterministic Annealing Filter (DAF) was originally invented for resolving left-right ambiguities in the ATLAS TRT (Frühwirth and Strandlie, CPC 1999)
  - turned out to be very precise and robust for this purpose
  - later successfully applied to vertex fitting in CMS
  - implemented for alignment with tracks in CMS
  - reconstruction of tracks in narrow jets in CMS
New developments

- DAF has been implemented in standard reconstruction program in CMS tracker (Winkler, PhD Thesis 2003)
- Systematic comparisons to standard, combinatorial Kalman filter (CKF) have been made
  - track finding performed by CKF in both cases, fitting procedure different
- Clear improvements in resolution of track parameters seen in ”difficult” situations, such as reconstruction of high-energy narrow jets
New developments

impact parameter resolution

probability distributions

\[ E_T = 200 \text{ GeV} \]

\[ \chi^2 \text{ probability} \]

A. Strandlie, DESY, 24.01.2005
New developments

b-tagging efficiencies for DAF almost independent of jet energy

A. Strandlie, DESY, 24.01.2005
New developments

- Very recent idea: evaluate the viability of the DAF as a combined track finder and fitter in scenarios with potentially very large amounts of hit distortions and additional noise
  - as with the Kalman filter, form a track seed as the starting point of the track finding
- Kalman filter builds up a combinatorial tree of track candidates starting from the seed
  - nowadays therefore often called a combinatorial Kalman filter (CKF)
New developments

- Basic idea of this application of the DAF:
  - collect hits around the extrapolation of the track candidate from the seed
  - run the DAF on this set of hits
  - monitor the chisquare of the track in order to decide whether the track candidate is a good track or not
New developments

- Potential advantages with respect to Kalman filter:
  - no combinatorial explosion in a dense track environment or with presence of large amounts of noise
  - DAF known to be very robust when fitting tracks with distorted hits and additional noise
New developments

- Potential disadvantages with respect to Kalman filter:
  - naive hit collection performed by blind extrapolation from seed can yield track candidate with many hits due to limited precision of seed and potentially long extrapolation distances
  - not obvious how fast the DAF will converge and what it will converge to
New developments

Current study (R. Frühwirth, A. Strandlie):

- simulation experiment of tracker setup similar to barrel part of CMS and ATLAS inner trackers
- three innermost layers have resolution typical to that of a pixel detector
- ten outermost layers resemble a silicon strip detector
- tracks propagated inwards-out and hits simulated with knowledge of the measurement resolution and the amount of material in the detector
New developments

- additional, nearby noise hits generated
- fraction of good hits distorted
- seeds formed by three 3D hits in layers closest to the beam
- track candidates followed outwards from seed
- DAF and combinatorial Kalman filter (CKF) implemented and compared
New developments

First scenario: assuming that collection of hits in a band around the tracks has been performed by separate procedure

- therefore simulating additional noise hits inside this band
- also distorting a fraction of the true track hits
- running CKF in the standard way outwards from the seed
- for the DAF, using direct propagation from seed towards end of tracker as initial track
New developments

average number of track candidates for CKF

A. Strandlie, DESY, 24.01.2005
New developments

CKF
1/p_T
resolution

DAF
1/p_T
resolution
New developments

- With no distorted hits (but with additional noise):
  - CKF about as accurate as DAF
  - fraction of track candidates with contamination about 2 % for both
- With 25 % distorted hits:
  - contamination fraction 22 % for CKF and 14 % for DAF
- With 50 % distorted hits:
  - contamination fraction 42 % for CKF and 35 % for DAF
  - DAF significantly more precise than CKF
- CKF about 15 times slower than DAF (keeping up to 64 candidates per seed), but CKF does not perform much worse with a significantly smaller amount of candidates
New developments

- Second (more realistic) scenario: noise hits simulated in entire tracker and hit collection performed as part of the track finding
  - density of noise hits decreasing with second or third power of radius
  - largest occupancy in the innermost layers
  - CKF proceeds "standard" way by building up combinatorial tree of track candidates from each seed
- DAF uses two-stage procedure
  - extrapolate three layers after seed, pick up hits in search window and run to convergence
  - extrapolate towards end of tracker in more narrow search window, pick up hits and run to convergence

A. Strandlie, DESY, 24.01.2005
New developments

density of noise hits

A. Strandlie, DESY, 24.01.2005
New developments

average number of track candidates for CKF

A. Strandlie, DESY, 24.01.2005
New developments

average number of hits in search window for DAF
New developments

number of DAF iterations

A. Strandlie, DESY, 24.01.2005
New developments

CKF and DAF resolutions without distorted hits
New developments

CKF and DAF resolutions with 10% distorted hits
New developments

- DAF performs very well when efficient hit collection procedure can be assumed
  - superior to CKF in particular for significant fraction of distorted track hits
- Some outliers (fraction increasing with noise density) when using the DAF also for the hit collection
  - due to convergence to wrong set of hits in first pass (i.e. first three layers after the seed)
  - for the remaining (non-outlier) set of tracks, the DAF is more precise
  - main challenge will be to overcome the difficult first phase of track finding when noise level is high
  - could form seeds at the outside where noise level is smaller
  - work still very much in progress…..
Summary and conclusions

- A brief review of traditional algorithms has been given
- Some recent developments have been presented in some detail:
  - GSF for electron reconstruction
  - GSF for vertex reconstruction
  - DAF for combined track finding and fitting
Summary and conclusions

- The Kalman filter is still very much alive and is the default choice for many applications.
- New developments such as Gaussian-sum filters and adaptive methods have been shown to be superior for some applications and are still gaining popularity.
Summary and conclusions

Main drawbacks with advanced methods:

- in general more time consuming than Kalman filter
- in my opinion approaching the limit of complexity:
  - have to be used with some intelligence
  - requires substantial degree of expertise by the implementer
Thanks to Rainer Mankel for inviting me to this seminar!