Advances in Tracking and Trigger Concepts

Prof. Dr. Ivan Kisel Uni-Frankfurt, FIAS, GSI

HEP Research Centers

| Research Center | Accelerator (GeV) | Experiment | Physics |
|----------------------|---|------------|--------------------|
| SLAC, USA | PEP-II, <mark>e</mark> ⁻x e⁺ (9 x 3.1) | BaBar | B-Physics |
| Fermilab, | Toyatron $p \times p (1000 \times 1000)$ | D0 | Universal |
| USA | (1000×1000) | CDF | Universal |
| BNL, USA | | PHENIX | Quark-Gluon-Plasma |
| | RHIC, Heavy lons | STAR | Quark-Gluon-Plasma |
| KEK, Japan | KEK-B, e⁻ x e⁺ (8 x 3.5) | BELLE | B-Physics |
| CERN. | LHC, p x p (7000 x 7000) | ATLAS | Universal |
| | | CMS | Universal |
| Switzerland | | ALICE | Quark-Gluon-Plasma |
| | | LHCb | B-Physics |
| | HERA, <mark>e^{+/-} x</mark> p (27.5 x 920) | ZEUS | Proton-Physics |
| DESY, Germany | | H1 | Proton-Physics |
| | | HERMES | Spin-Physics |
| | | HERA-B | B-Physics |
| FAIR/GSI, Germany | | PANDA | Quark-Physics |
| | ו פוט וועו פוט, p, Heavy IONS | CBM | Quark-Gluon-Plasma |

Ivan Kisel, Uni-Frankfurt, FIAS, GSI

ACAT-2013, Beijing, 19.05.2013 2/40

кек

HEP Experiments: Collider and Fixed-Target



HEP Experiments: select interesting physics on-line

From Raw Data to Physics



Stages of Event Reconstruction



- Hough Transformation
- Elastic Neural Net



• Kalman Filter



Global Methods: Conformal Mapping + Histogramming

Global methods are especially suitable for fast tracking in projections **Example:** Collider experiment with a solenoid, where tracks are circular trajectories

> Conformal Mapping: Transform circles into straight lines $u = x/(x^2+y^2)$ $v = -y/(x^2+y^2)$

Histogram:

Collect a histogram of azimuth angles φ Find peaks in the histogram Collect hits into tracks



Global Methods: Conformal Mapping + Histogramming



Histogram:

Collect a histogram of azimuth angles ϕ Find peaks in the histogram Collect hits into tracks



Features:

- Impressive visual simplification of the problem
- Each step is easy to implement in hardware

• ...

Weak points:

- Needs to know an exact position of the interaction point
- Do not finds tracks, but only approximate track parameters
- No grouping of hits into track candidates
- Finds only primary tracks
- Problems with non-uniform magnetic field
- ...

Useful implemented in hardware and for very simple event topologies only

Global Methods: Hough Transformation



Global Methods: Hough Transformation



Features:

- Generalization of the histogramming method
- Easy to implement in hardware
- ...

Weak points:

- Needs a global track model
- Appropriate only for uniform magnetic fields
- Does not include multiple scattering
- Provides only track parameters etc.
- No competition between track candidates
- Needs a lot of memory (x, y, tx, ty, q/p -> 5D histogramming)
- ...

Useful implemented in hardware and for simple event and trigger topologies

Local Methods: Kalman Filter for Track Following



Local Methods: Kalman Filter for Track Following



Features:

- Psychologically easy to accept hit by hit track finding
- Combined track finder and fitter based on KF
- Development of a new experiment starts with an ideal MC track finder and a realistic KF track fitter, then it is easy to implement the track finder as track following
- ...

Weak points:

- Based on a single track approach
- Needs seeding (starting short track segments)
- Efficiency is limited by the seeding efficiency (and detectors!)
- Works at the hit level, searching for hits within a region
- Repeats the same calculations, after discarding track candidates
- No global competition between track candidates
- ...

Useful for relatively simple event topologies and as a second after the ideal track finder

Local Methods: Cellular Automaton as Track Finder



Local Methods: Cellular Automaton as Track Finder



Features:

- Local relations -> parallel algorithm
- Staged implementation: hits -> segments -> tracks
- Polynomial (2nd order) combinatorics
- Track competition at the global level
- Includes the KF fitter, if necessary, for high track densities
- ...

Weak points:

- Not easy to understand a parallel algorithm (Game of Life)
- Parallel hardware is coming now
- ...

Useful for complicated event topologies with large combinatorics and for parallel hardware

Many-core HPC: Cores, Threads and Vectors

HEP experiments work with high data rates, therefore need High Performance Computing (HPC) !



Fundamental redesign of traditional approaches to data processing is necessary

Ivan Kisel, Uni-Frankfurt, FIAS, GSI

ACAT-2013, Beijing, 19.05.2013 14/40

Many-Core CPU/GPU Architectures



CPU/GPU Programming Frameworks



- Intel ArBB (Array Building Blocks)
 - Extension to the C language
 - Intel CPU/GPU specific
 - SIMD exploitation for automatic parallelism
- NVIDIA CUDA (Compute Unified Device Architecture)
 - Defines hardware platform
 - Generic programming
 - Extension to the C language
 - Explicit memory management
 - Programming on thread level
- OpenCL (Open Computing Language)
 - Open standard for generic programming
 - Extension to the C language
 - Supposed to work on any hardware
 - Usage of specific hardware capabilities by extensions

• Vector classes (Vc)

- Overload of C operators with SIMD/SIMT instructions
- Uniform approach to all CPU/GPU families
- Uni-Frankfurt/FIAS/GSI

Choice of CPU/GPU/Programming is a practical question

Vectorization: SIMD (Single Instruction Multiple Data)



Use headers to overload +, -, *, / operators --> the source code is unchanged !

Header (Intel's SSE), later Vector Classes (Vc)

```
SIMD instructions
typedef F32vec4 Fvec t;
/* Arithmetic Operators */
friend F32vec4 operator +( const F32vec4 &a, const F32vec4 &b ) { return _mm_add_ps(a, b); }
friend F32vec4 operator –( const F32vec4 &a, const F32vec4 &b ) { return _mm_sub_ps(a, b); }
friend F32vec4 operator *( const F32vec4 &a, const F32vec4 &b ) { return mm mul ps(a, b); }
friend F32vec4 operator /( const F32vec4 &a, const F32vec4 &b ) { return mm div ps(a, b); }
/* Functions */
friend F32vec4 min(const F32vec4 &a, const F32vec4 &b){ return mm min ps(a, b); }
friend F32vec4 max( const F32vec4 &a, const F32vec4 &b){ return mm max ps(a, b); }
/* Square Root */
friend F32vec4 sqrt( const F32vec4 &a ){ return _mm_sqrt_ps (a); }
/* Absolute value */
friend F32vec4 fabs( const F32vec4 &a ){ return _mm_and_ps(a, _f32vec4_abs_mask); }
/* Logical */
friend F32vec4 operator&( const F32vec4 &a, const F32vec4 &b ){ // mask returned
 return _mm_and_ps(a, b);
friend F32vec4 operator ( const F32vec4 &a, const F32vec4 &b ){ // mask returned
 return _mm_or_ps(a, b);
friend F32vec4 operator<sup>(</sup> const F32vec4 &a, const F32vec4 &b){ // mask returned
 return _mm_xor_ps(a, b);
}
friend F32vec4 operator!( const F32vec4 &a ){ // mask returned
 return mm xor ps(a, f32vec4 true);
friend F32vec4 operator ( const F32vec4 &a, const F32vec4 &b ){ // mask returned
 return _mm_or_ps(a, b);
}
/* Comparison */
friend F32vec4 operator < ( const F32vec4 &a, const F32vec4 &b ){ // mask returned
 return mm cmplt ps(a, b);
```

Code (Part of the Kalman Filter)

```
inline void AddMaterial( TrackV &track, Station &st, Fvec_t &qp0 )
{
   cnst mass2 = 0.1396*0.1396;
```

```
Fvec_t tx = track.T[2];
Fvec_t ty = track.T[3];
Fvec_t txtx = tx*tx;
Fvec_t tyty = ty*ty;
Fvec_t txtx1 = txtx + ONE;
Fvec_t h = txtx + tyty;
Fvec_t t = sqrt(txtx1 + tyty);
Fvec_t h2 = h*h;
Fvec_t qp0t = qp0*t;
```



vc = vec_add(va, vb)

cnst c1=0.0136, c2=c1*0.038, c3=c2*0.5, c4=-c3/2.0, c5=c3/3.0, c6=-c3/4.0;

 $Fvec_t s0 = (c1+c2*st.logRadThick + c3*h + h2*(c4 + c5*h + c6*h2))*qp0t;$

Fvec_t a = (ONE+mass2*qp0*qp0t)*st.RadThick*s0*s0;

```
CovV &C = track.C;
```

```
C.C22 += txtx1*a;
C.C32 += tx*ty*a; C.C33 += (ONE+tyty)*a;
}
```

How to Parallelize Reconstruction in Running (!) Experiments



S. Borkar et al. (Intel), "Platform 2015: Intel Platform Evolution for the Next Decade", 2005.



| 1 | Modify the existing code | ATLAS |
|---|---|-------|
| 2 | Implement a new algorithm in the combinatorial part | CMS |
| 3 | Merge on-line and off-line codes | ALICE |
| 4 | Use an advanced existing algorithm as seed finder | STAR |
| 5 | Design and develop a new code from scratch | CBM |

Provide stable and reproducible/better results

ATLAS



Marcus Elsing, 4th International Workshop for Future Challenges in Tracking and Trigger Concepts, CERN, 28-30.11.2012

LS1: Modify the existing code

CMS



Current tracking limitations

Iterative tracking is performing well in early steps; early steps are able to provide fast the bulk of tracking efficiency and, timing wise, they are pretty stable with respect to PU.

Problems are in later steps designed (loosen cuts) to recover efficiency for more difficult (e.g. displaced) tracks.

Late steps result in too many fake seeds with respect to early steps; each seed needs to be propagated in the attempt of building a track. This is time consuming and time per final good track is not favorable for these later steps.

The long story short: combinatorics...

Track building (e.g. full track reconstruction) is eventually able to get rid of these fake proto-tracks; similarly, more information should be used to clean up not useful seeds in advance. But this requires time.

New developments should address this paradox.

24.10.2012 Cerati / Sguazzoni / Stenson 41

Giacomo Sguazzoni, 4th International Workshop for Future Challenges in Tracking and Trigger Concepts, CERN, 28-30.11.2012

LS1: Implement a new algorithm in the combinatorial part

ALICE



ALICE HLT: Event of the first HI run with the GPU CA tracker



Merge off-line and on-line codes

STAR





Use an advanced existing algorithm as seed finder

STAR TPC CA Track Finder



lxir075.gsi.de: 4 Intel Xeon Westmere CPU E7-4860, 10 cores per CPU, HT, 2.27 GHz, 24 MB L3 cache, 64 GB RAM

CA is stable w.r.t. track multiplicity, 10 (40 HLT) times faster than the TF track finder and has strong scalability up to 80 cores.

Ivan Kisel, Uni-Frankfurt, FIAS, GSI

CBM



CBM: Kalman Filter Track Fit on Cell

| | Stage | Description | Time/track | Speedup | |
|------------|-------|-------------------------------------|------------------|---------|-----------|
| Cell Intel | | Initial scalar version | 12 ms | _ | |
| | 1 | Approximation of the magnetic field | $240~\mu { m s}$ | 50 | 10000 6 1 |
| | 2 | Optimization of the algorithm | $7.2~\mu{ m s}$ | 35 > | on any PC |
| | 3 | Vectorization | $1.6~\mu{ m s}$ | 4.5 J | |
| | 4 | Porting to SPE | $1.1~\mu{ m s}$ | 1.5 | |
| | 5 | Parallelization on 16 SPEs | $0.1~\mu{ m s}$ | 10 | |
| | | Final simulized version | $0.1~\mu{ m s}$ | 120000 | |

Comp. Phys. Comm. 178 (2008) 374-383



blade11bc4 @IBM, Böblingen: 2 Cell Broadband Engines, 256 kB LS, 2.4 GHz The KF speed was increased by 5 orders of magnitude



Motivated by, but not restricted to Cell !

CBM: Kalman Filter Track Fit Library

Kalman Filter Methods

Kalman Filter Tools:

- KF Track Fitter
- KF Track Smoother
- Deterministic Annealing Filter

Kalman Filter Approaches:

- Conventional DP KF
- Conventional SP KF
- Square-Root SP KF
- UD-Filter SP
- · Gaussian Sum Filter

Track Propagation:

- Runge-Kutta
- Analytic Formula





Implementations

Vectorization (SIMD):

- Header Files
- Vector Classes Vc
- Array Building Blocks ArBB
- OpenCL

Parallelization (many-cores):

- Open MP
- ITBB
- ArBB
- OpenCL

Precision:

- single
- double





Strong many-core scalability of the Kalman filter library

CBM Track Finding Methods: from Pixels to Strip Detectors



| | Developer | Tracking Method | < 2005 | > 2005 |
|---|---------------------------|----------------------|--------------|--------|
| 1 | LHEP JINR, Dubna | Conformal Mapping | \checkmark | × |
| 2 | ZITI, Mannheim | Hough Transformation | \checkmark | × |
| 3 | LIT JINR, Dubna | Track Following | \checkmark | × |
| 4 | Uni-Heidelberg, GSI, FIAS | Cellular Automaton | \checkmark | 1 |

Cellular Automaton is appropriate for complicated event topologies with large combinatorics

Ivan Kisel, Uni-Frankfurt, FIAS, GSI

CBM CA Track Finder: Efficiency



Efficient and stable event reconstruction

CBM CA Track Finder: Reliability and Scalability



Stable algorithm down to 80% detector efficiency and strong many-core scalability

ACAT-2013, Beijing, 19.05.2013 31/40

CBM CA Track Finder at High Track Multiplicity

A number of minimum bias events is gathered into a group, which is then treated by the track finder as one event



Au+Au mbias events at 25 AGeV, 8 STS, 0 x 7,5 strip angles

Towards 4D (space+time) event reconstruction

CBM CA Track Finder: Efficiency and Time vs. Track Multiplicity



Stable reconstruction efficiency and time as a second order polynomial up to 100 minimum bias events in a group

CBM KF Particle: Vertices and Decayed Particles





- Mother and daughter particles have the same state vector and are treated in the same way
- Geometry independent
- Kalman filter based



KFParticle provides uncomplicated approach to physics analysis (used in CBM, ALICE and STAR)

Ivan Kisel, Uni-Frankfurt, FIAS, GSI

ACAT-2013, Beijing, 19.05.2013 34/40

CBM KF Particle Finder for Physics Analysis and Selection



Ivan Kisel, Uni-Frankfurt, FIAS, GSI

ACAT-2013, Beijing, 19.05.2013 35/40

CBM Standalone First Level Event Selection (FLES) Package



The first version of the FLES package is vectorized, parallelized, portable and scalable

Ivan Kisel, Uni-Frankfurt, FIAS, GSI

ACAT-2013, Beijing, 19.05.2013 36/40

CBM: Parallelization in the Event Reconstruction

| Algorithm | Vector SIMD | MultiThreading | CUDA | OpenCL CPU/GPU |
|-----------------------------|--------------|----------------|--------------|----------------|
| Hit Producers | | | | |
| STS KF Track Fit | \checkmark | \checkmark | \checkmark | $\sqrt{1}$ |
| STS CA Track Finder | \checkmark | \checkmark | | |
| MuCh Track Finder | \checkmark | \checkmark | \checkmark | |
| TRD Track Finder | \checkmark | \checkmark | \checkmark | |
| RICH Ring Finder | \checkmark | \checkmark | | (√/√) |
| Vertexing (KFParticle) | \checkmark | \checkmark | | |
| Off-line Physics Analysis | \checkmark | | | |
| FLES Analysis and Selection | √ | \checkmark | | |

| Andrzej Nowak (OpenLab, CERN) by Hans von der Schmitt (ATLAS) at GPU Workshop, DESY, 15-16 April 2013 | | | | | | | |
|---|------|-----------------------------|---------------|-------|---------|--------|------------|
| | SIMD | Instr. Level Parallelism | HW Threads | Cores | Sockets | Factor | Efficiency |
| MAX | 4 | 4 | 1.35 | 8 | 4 | 691.2 | 100.0% |
| Typical | 2.5 | 1.43 | 1.25 | 8 | 2 | 71.5 | 10.3% |
| HEP | 1 | 0.80 | 1 | 6 | 2 | 9.6 | 1.4% |
| CBM@FAIR | 4 | 3 | 1.3 | 8 | 4 | 499.2 | 72.2% |

x Algorithm x Memory

Parallelization becomes a standard in the CBM experiment

Software Evolution: Many-Core Barrier



Consolidate Efforts: Common Reconstruction Package



Ivan Kisel, Uni-Frankfurt, FIAS, GSI

Consolidate Efforts: International Workshops

International Workshop for Future Challenges in Tracking and Trigger Concepts

- 1st GSI, Darmstadt, Germany, 07-11.06.2010;
- 2nd CERN, Geneva, Switzerland, 07-08.07.2011;
- 3rd FIAS, Frankfurt, Germany, 27-29.02.2012;
- 4th CERN, Geneva, Switzerland, 28-30.11.2012;
- 5th BNL, Brookhaven or LBNL, Berkeley, USA, this year.



Contact: Sverre Jarp, Jerome Lauret, Ivan Kisel

ACAT-2013, Beijing, 19.05.2013 40/40

devanded Studies 🔒 BROOKKAAVEN HIC FAIR 65 55 🗴 🧱

© A.Senger