

Parallelization of the SIMD Ensemble Kalman Filter for Track Fitting Using Ct

Thursday, 25 February 2010 17:00 (25 minutes)

A great portion of data mining in a high-energy detector experiment is spent in the complementary tasks of track finding and track fitting. These problems correspond, respectively, to associating a set of measurements to a single particle, and to determining the parameters of the track given a candidate path [Avery 1992]. These parameters usually correspond to the 5-tuple state of the model of motion of a charged particle in a magnetic field.

Global algorithms for track fitting have been superseded by recursive least-square estimation algorithms that, assuming the measurement noise is Gaussian, result in an optimal estimate [Frühwirth and Widl 1992]. However, this assumption hardly ever holds due to energy loss and multiple scattering. Extensions to the Kalman filter have been proposed and implemented [Frühwirth 1997] that use sums of Gaussians to model non-Gaussian distributions. In addition, non-linear filtering is necessary to eliminate the effects of outliers including misclassified measurements. Track fitting based on sums of Gaussians can be implemented through an ensemble of single Gaussian fits, each of which is the result of a single Kalman filter [Gorbunov et al. 2008].

Efficient parallel implementations of these algorithms are also crucial due to the large amount of data that needs to be processed. Since many separate tracks need to be fitted, the simplest way to parallelize the problem is to fit many independent tracks at once. However, with the ensemble algorithm it is also possible to parallelize across the ensemble within a single track fit. Parallelism mechanisms available in the hardware include multiple nodes and multiple cores.

Prior implementations can also make use of the SIMD vector units present in most modern CPUs [Gorbunov et al. 2008]. We have performed an implementation of the ensemble approach using Ct. Ct is a generalized data-parallel programming platform that can efficiently target both multiple cores and SIMD vector units from a single high-level specification in C++. Compared to the earlier SIMD implementation, Ct allows for hardware and instruction set portability.

We have obtained scalable speedup with respect to a scalar baseline implemented using both single and double precision. In the prior work, a speedup of 1.6x for scalar vs. vectorized versions of the algorithm were reported for double precision. Our results are comparable to these.

In addition, we have explored the implementation of sequential Monte Carlo in Ct using particle filtering to model arbitrary probability distributions. However, one challenge in this context is the computation of the likelihood function, which requires estimates of the measurement error for every measurement and their correlation.

References

- Avery, P. 1992. Applied fitting theory V: Track fitting using the Kalman filter. Tech. rep.
Frühwirth, R., and Widl, E. 1992. Track-based alignment using a Kalman filter technique. Communications in Computational Physics .
Frühwirth, R. 1997. Track-fitting with non-Gaussian noise. Communications in Computational Physics (Jan.).
Gorbunov, S., Kebschull, U., Kisel, I., Lindenstruth, V., and Müller, W. F. J. 2008. Fast SIMDized Kalman filter-based track fit. Computer Physics Communications .

Primary authors: Dr GHULOUM, Anwar (Intel); Dr MCCOOL, Michael D. (Intel/University of Waterloo); ESTEVES, R. Gabriel (Intel/University of Waterloo)

Presenter: Dr MCCOOL, Michael D. (Intel/University of Waterloo)

Session Classification: Thursday, 25 February - Data Analysis - Algorithms and Tools

Track Classification: Data Analysis - Algorithms and Tools