The Challenges of Integrating CUDA Engines Into an Existing Package

By Eri Rubin, OptiTex Ltd.

The Company

- OptiTex is a leading developer of 2D/3D CAD solutions for the fashion and textile industries.
- OptiTex has over 20 years of experience and more than 30,000 active installations worldwide

OptiTex
Next Generation 2D/3D CAD

Talk overview

- Background
- Goals of the Project
- Overview of our Garments simulation Engine
- Evolution of our CUDA development
- Our current development methodology
- Our achievements
- Future goals
- Questions

OptiTex
Next Generation 2D/3D CAD

Background

- Why do we want to accelerate ?
  - Achieve an interactive virtual try-on experience for the end customer
- Why use CUDA to accelerate ?
  - After exhaustive testing of other hardware and software solutions, only CUDA seemed promising in performance, cost, and development ease.

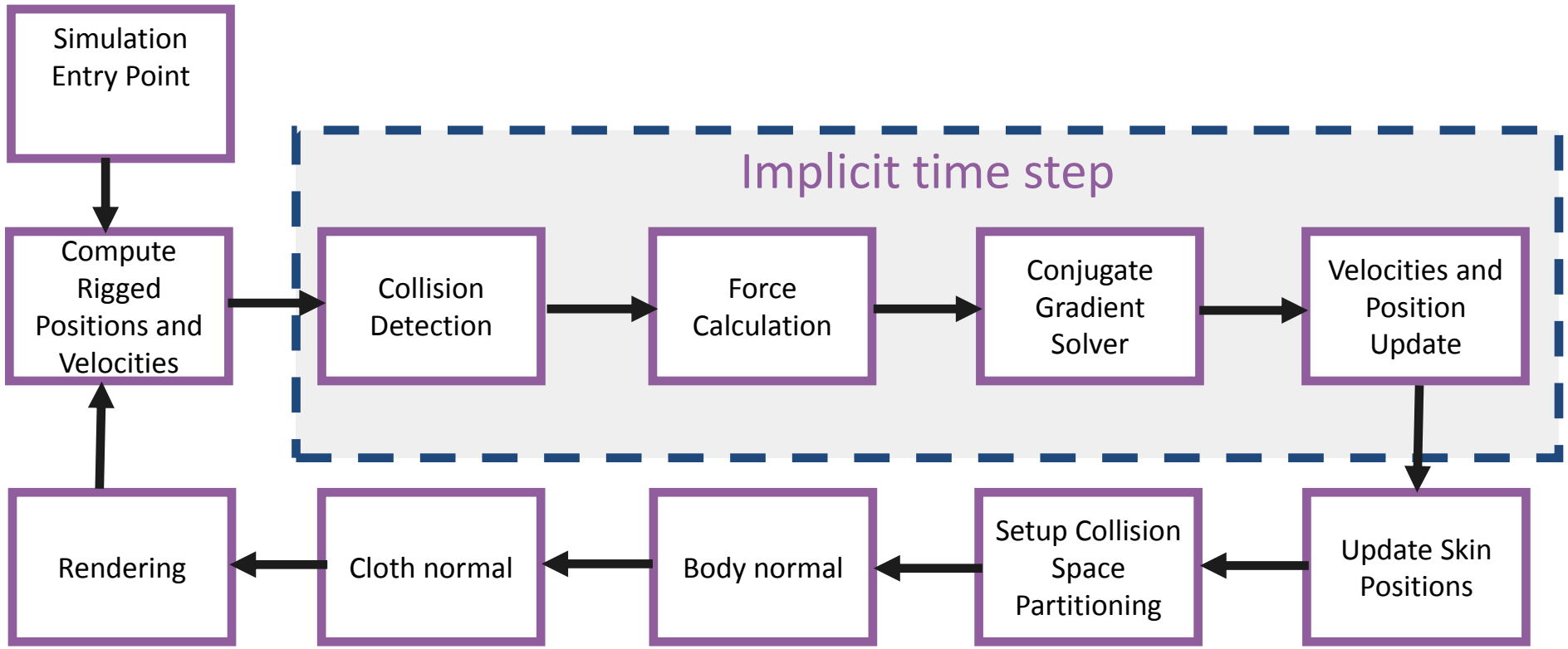
OptiTex
Next Generation 2D/3D CAD

Project Goals

- Achieve real time performance for reasonable sized garments.
- Maintain the same quality of simulation as the CPU version.
- Keep current functionality active, all work must be an optional add on.

OptiTex
Next Generation 2D/3D CAD

Garment simulation engine



Size of the project

- Over 200,000 lines of CPU code
- 25 kernels with a few thousand lines of code
- 10 separate modules
- Full preconditioned constraint conjugate gradient
- Full continues collision detection system
- Very easy to get lost in the mayhem
- Can be very hard to debug

OptiTex
Next Generation 2D/3D CAD

From the beginning

- Initial testing phase
- Check if we expect to hit our goals
- What changes will we need to make to the current system
- What will be the best way to integrate the CUDA engine into our existing system

OptiTex
Next Generation 2D/3D CAD

Project Elements

- New platform integration
  - CUDA
- Algorithm adaptation
  - Move to parallel algorithms, change the order.
- Performance assessment
  - To identify the bottlenecks
- Result validation
  - Compare GPU results to CPU results

OptiTex
Next Generation 2D/3D CAD

Identify the bottle necks

- In allot of cases people tend to start “optimizing” without having a clear picture of what are the real bottle necks in the system.
- This is the first stage, and comes back later aswell
- Premature optimization is the root of all evil !!!
- After finding the locations of the bottle neck drill down to find out if you are memory bound or compute bound.

The plug-in system

- A modular system meant to enable extensibility of the software.
- Such a system is useful beyond CUDA integration, for such things such as file formats and so forth.
- To compliment the system you need to expose an API that will allow the plug-ins access to the needed functions and data in the system.

Data Structures

- Modify existing array of structures to structure of arrays
- Each component of the mesh dose not hold its own data members, instead has a function to access the relevant array in its mesh with its member index.
- On top of that, index locality must be maintained to minimize random access
- Complex data structures can be used on the gpu.

Parallel algorithms

- Wrap your head around the data parallel paradigm
- Separate the data transfer from the computation
- Not all algorithms map intuitively to this model, for example reduce, or hashing. Yet efficient implementations have been coded for such algorithms
- Certain computations may change time and place

Comparing results

- Add a mode were both code paths occur with the same input.
- How do you compare a vector of floats ?
- The difference in precision between CPU and GPU
- Different execution order of the operations.
- Create a set of utility functions that automatically copies the data back from the GPU to the CPU and find the difference between the elements based on a certain threshold.

Optimizing CUDA

- Create timers to take measurements of real running time
- Find out if you are memory bound or computation bound
- Find out how close you are to the theoretical limits of the hardware
- Use CUDA profiling tools to understand where there might be room for optimizations

OptiTex
Next Generation 2D/3D CAD

Optimization Cycle

- Try and identify locations with potential room for improvement.
- Implement your optimization
- See if it really improves the performance.
- Verify that you are still getting the correct results.
- Rinse and repeat …

Achievements to date

- Created a solid framework for extensions to our system without compromising the stability and development of our existing
- Developed a full Garment simulation engine that runs entirely on the GPU.
- Achieved real time performance (30 fps on a 8k cloth mesh)
- Same quality as the CPU simulation.

OptiTex
Next Generation 2D/3D CAD

Future goals

- Achieve higher frame rates for larger meshes
- Unify code bases
- Enhance cuda implementation to support all complex cases

OptiTex
Next Generation 2D/3D CAD

Questions ?

Eri Rubin, CUDA Project Manager
eri.rubin@optitex.com
www.optitex.com