Teraflops for the Masses: Killer Apps of Tomorrow

Pradeep K. Dubey
Senior Principal Engineer
Corporate Technology Group
EDGE
UNC, Raleigh, May 23, 2006

Evolution continues ...

Media Evolution
- Modality-specific streaming
- Upcoming Transition
- Next Transition
- Modality-aware transformation
- Scene complexity: moderate
- Local processing dominated
- Scene complexity: large
- Global processing dominated

Graphics Evolution
- Modality-aware transformation
- Next Transition
- Multimodal recognition
- Scene complexity: real-world
- Physical simulation dominated

Mining Evolution
- Dataset: static/structured
- Response: offline
- Dataset: dynamic, multimodal
- Response: interactive
- Dataset: massive+streaming
- Response: real-time

Workload convergence: multimodal recognition and synthesis over complex datasets
Evolving towards model-based computing

Recognition

Mining

Synthesis

Real-world animation

Ray tracing

Global illumination

Behavioral Synthesis

Physical simulation

Kinematics

Emotion synthesis

Audio synthesis

Video/image synthesis

Document synthesis

Multimodal event/object Recognition

Statistical Computing

Clustering

Bayesian networks

Neural networks

LP/IP/QP/STO optimization

Large dataset mining

Semantic Web/Grid Mining

Streaming Data Mining

Text-based Data Mining

Based Retrieval

Active Filters

Personal indexing

Identity Reduction

Dynamic Environments

Efficient access to large, unstructured datasets

Stream processing

Evolving towards model-based computing

What is a tumor?

Is there a tumor here?

What if the tumor progresses?

It is all about dealing efficiently with complex multimodal datasets

Images courtesy: http://sphweb.bwh.harvard.edu/fmo00/pages/images_movies.html
RMS: Recognition Mining Synthesis

Today
- Model-based multimodal recognition
- Real-time analytics on dynamic, unstructured, multimodal datasets
- Photo-realism and physics-based animation

Model-less
- Real-time streaming and transactions on static – structured datasets
- Very limited realism

Tomorrow
- Model-less
- Real-time streaming and transactions on static – structured datasets
- Very limited realism

Most RMS apps are about enabling interactive (real-time) RMS Loop or iRMS
Next-Generation Entertainment

RMS Primitives:

- Model the ball
- Find the ball
- Replace the ball
- Shade/Bounce the ball

Going beyond media-stream encode-decode-transcode!

Real-time DCC Loop

What if ... what if ...

- Model the ball
- Shade/Bounce the ball

RMS Closed Loop
Rendering+Physics+Vision

Going beyond ‘red-eye removal’
Where are we headed...

Machine learning
Neural networks
Probabilistic reasoning
Fuzzy logic
Belief networks
Evolutionary computing
Chaos theory

Physics
Dynamics

Soft Computing
Constraints

Soft Physics?
Constraint Dynamics

RMS Computing Care

Unstructured Information Management
Analytics
Vision/Tracking
Gaming

Benefiting Applications
Real-time asset management, text mining
camera stream mining, 3D graphics ...

Pool of Mathematical Techniques
Conic optimization, Subspace projection ...

Pool of RMS Functions
Interior-point, Spectral Bundle, SVM ...
Workload Convergence ➔ RMS Primitives

Hollywood to Wall Street

Common Core Computing Kernel

- Stochastic Optimization (Sim Annealing, Genetic Alg., Bayes Learning)
- Numerical Integration (Monte Carlo, Quasi-MC, Gaussian)
- Convex Optimization (LP, QP, SOCP, SDP, Network, SVM)
- Combinatorial Optimization (Integer Prog, Dynamic Prog)
- Differential Equations Solvers (Parabolic, Elliptic, Hyperbolic, Finite Element Method, Stochastic)
- Iterative Solvers (Conjugate Gradients, Gauss-Seidel, Jacobi, GMRES)

Entertainment Tomorrow
- Object Recognition
- Object Tracking
- Computer Vision: Depth from Stereo
- Mesh Refinement
- Rendering: Path Tracing
- Computational Fluid Dynamics
- AI for Games: Path planning

Business/Finance Today
- Portfolio Selection
- Asset Allocation
- Multi-Look Option Pricing
- Asset-Liability Management
- Risk Management
- Interest-Rate Derivative Pricing
- Multi-Party Auctions
RMS Computing Core: Scaling to Next Generation Needs

- Map-based shading
- SIMPLEX based linear optimization
- Mass-Spring based deformation
- Marker-based explicit surface tracking
- Linear manifold based recognition/modeling
- Linear Complementarity problem
- Low dimension classifiers
- Global Illumination based IPM based LP/QP/NLP optimization
- FD/FE/FV based deformation
- Level Set based implicit surface tracking
- Non-linear manifolds computer vision
- Non-linear Complementarity
- High dimension classifiers
- Non-Linear
- And
- Generative

Summary

There are mass applications that require significant increase in compute density

- There is nothing as general-purpose as physics!
- Visual computing is a proxy of this much larger class (RMS)

These applications are not linear extensions of existing usage

- Optimal platform for such apps should not be linear extension either

There is a significant performance difference between a brute-force CMP Vs. a smart CMP targeted for this class

- There is significant opportunity for silicon differentiation

These apps will likely be the driver for most future technology vectors

- Programming to processor to memory technology