Research Thrust R3:
Overview of Image and Data Information Management

David Kaeli

NSF Year 9 Site Visit
April 22, 2009
CenSSIS Image and Data Information Management

- Addressing key research barriers in computational efficiency, embedded computing and image/sensor data management
- Exploiting hardware/software acceleration/parallelization to enable new discovery in SSI applications
- Producing an image/data repository and software-engineered Subsurface Sensing and Imaging Toolsets
- Developing enabling tools targeting system-level projects
  - Near real-time reconstruction and visualization
System Level Impact of R3 Activities

R1

Fundamental Science

S1 S2 S3

Validating TestBEDs

Bio-Med

Enviro-Civil

S4 S5

R2

R3 Image and Data Information Management
Year 9 R3 Researchers

**Faculty/Staff**
- Dana Brooks – NU
- George Chen – MGH
- Charles Dimarzio - NU
- Synho Do - MGH
- Burak Erem – NU
- James Goodman - UPRM
- James Hale – Spelman
- Shawn Hunt – UPRM
- Luis Jiménez - UPRM
- David Kaeli - NU
- Miriam Leeser – NU
- Waleed Meleis – NU
- Homer Pien - MGH
- Wilson Rivera - UPRM
- Samuel Rosario – UPRM
- Nayda Santiago – UPRM
- Greg Sharp - MGH
- Charles Stewart – RPI
- Miguel Velez - UPRM
- Furong Yang – NU

**Graduate students**
- Malak Alshawakeh - NU
- Emmanuel Arzuaga - NU
- Fatemeh Azmandian – NU
- Maria Constanza-Torres - UPRM
- Burak Erem – NU
- Byunghyun Jang – NU
- Jorge Manrique - UPRM
- Perhaad Mistry - NU
- Rafal Norton – NU
- Dana Schaa – NU
- Blas Trigueros - UPRM
- Gehua Yang – RPI
- Ayse Yilmazer - NU
- Juemin Zhang - NU
Year 9 R3 Researchers

Undergraduates

- Christine Cortes - UPRM
- Suzette Gomez – UPRM
- Yajaira Gonzalez - UPRM
- Stephen Hobson – NU
- Chawandia Mack – Spelman
- Gabriel Martinez - UPRM
- Martin Ramirez – UPRM
- Ruben Rios - UPRM
- Christian Sanchez – UPRM
- Joralis Sanchez - UPRM
Overview of Year 9

- **Focus on sustainability** - Won sustaining proposals, developed new university and industrial ties, submitted new proposals
  - NSF STTR, NSF MRI, DARPA, Industry, NSF IGERT, NIH, and NIEHS

- **Technical leadership in GPGPU acceleration**
  - GPGPU-1 Workshop held with the Year 8 RICC – 250 attendees!
    - Special issue of the Journal of Parallel and Distributed Processing in 2008
  - GPGPU-2 Workshop held with ACM ASPLOS in 2009 – 50+ attendees
    - Proceedings published by ACM International Conference Series in 2009
    - [http://www.cis.udel.edu/chips-mentoring-workshop](http://www.cis.udel.edu/chips-mentoring-workshop)
  - Special session at IEEE International Symposium on Biomedical Imaging – “Biomedical Computing on Many Cores” – June 2009
Overview of Year 9

- Continued support of existing software toolboxes and new toolbox development
  - The Coastal Image Analysis Toolbox (HyCIAT)
    - A toolset for the study the water optical properties, bathymetry and fractional abundances using hyperspectral image analysis
  - The Citation Indexing System
    - A database system similar in nature to Google Scholar for CenSSIS publications

- Apply CenSSIS Database to address new problems
  - NIEHS PRoTECT proposal
    - Environmental, biological and demographic data in a single repository
    - Providing data mining and GIS tools for enhanced data management/analysis
GPGPU

- Graphics Processing Units – NVIDIA, ATI, Intel
  - Provides massive parallelism at low cost
- Being applied to the following CenSSIS applications
  - Abundance Estimation - UPRM
  - Hyperspectral Image Reconstruction – UPRM/NU
  - Breast Cancer Tomosynthesis Reconstruction – MGH/NU
  - Cardiac CT Reconstruction – MGH/NU
  - Phase Unwrapping - NU
- Received joint NU-MGH grant from NVIDIA (1024 CPUs)
- Received an AMD/ATI Fellowship
  - Highly competitive - 4 awarded nationally
- Received grant to accelerate hyperspectral imaging on Larrabee
Phase Unwrapping and Affine Transformations for Optical Quadrature Microscopy Using a GPU

Percentage Distribution of time among Different Components

15.55 sec | 2.97 sec | 2.16 sec

7.2x speedup
GPGPU – Reduce the cost of image reconstruction!

$/sec of Breast Tomosynthesis Reconstruction on a NVIDIA GTX8800 GPU

* Collaboration with Richard Moore (MGH) and Daniel Kopans (MGH)

- GPUs will begin to replace the need for large compute clusters
- GPUs can enable low-cost real-time image-guided biopsies
GPGPU – Cardiac CT Reconstruction results with 128 cores

- Forward projection for Cardiac CT Reconstruction @ MGH
- All times in seconds
- Naive – Memory - Caching – explore different GPU optimizations
- 2 and 4 GPUs – multi-GPU configurations considered

35.5x speedup on 1 GPU

*Collaboration with Homer Pien (MGH), Clem Karl (BU) and Synho Do (MGH)
CenSSIS Toolboxes

- 3 completed CenSSIS toolboxes
  - http://www.censsis.neu.edu/
  - More than 400 downloads annually
  - Continued enhancements and support

- 3 ongoing CenSSIS toolbox development efforts
  - Coastal imaging and analysis
  - Capturing CenSSIS publications/datasets
  - Remote control for the 3D Fusion Microscope
CenSSIS Toolboxes

- Multiview Tomography Toolbox (MVT)
  - C++ fddlib: January 2003 (v. 1.0)
    - July 2003 (v. 1.1)
  - MATLAB mvt: October 2004
  - Used in the SSI Textbook and other funded projects

- Rensselaer Generalized Registration Library (RGRL)
  - Released September 2004
  - Effort has spawned “Dual-Align” startup company
  - http://www.dualalign.com

- Hyperspectral Image Analysis Toolbox (HIAT)
  - 1225 downloads in past 4 years
  - Version 2.2 available for download
  - The range of algorithms included in the toolbox have been funded by NASA, DoD, ARMY, and NSF
A tool to estimate shallow water optical properties, bathymetry and fractional abundances of bottom composition using hyperspectral image.

Builds upon the existing HIAT framework.

See the live demo!
Provides similar capability as Google Scholar

Specific to CenSSIS publications/information

Can store both data and pointers to data stored in other repositories

See the live demo!
CenSSIS Image Database System

- Over 18,000 image files online
- Database support continues through the end of the NSF support
- The CenSSIS database is proposed as the enabling technology in the NIEHS PRoTECT Program
R3 Contributions over 9 years

- Acceleration of Image Reconstruction
  - Application of FPGA, Grid Computing and GPU devices to a wide range of imaging applications and visualization
  - Contributions to GRID-based QoS
- CenSSIS Toolboxes
  - Heavy use of 3 existing toolboxes by government, industry and academia (and CenSSIS S-level projects)
  - Startup company effort (www.dualalign.com)
  - Continued development on new tools targeted at enhancing current platforms
- CenSSIS Database
  - Provides a web-based repository for documenting CenSSIS work
  - Provides a basis for future database projects
R3A Posters

- R3A p3 - “Performance Prediction Targeting Multi-GPU Execution,” D. Schaa and D. Kaeli
- R3A p4 – “Implementing a Highly Parameterized Digital PIV System on Reconfigurable Hardware,” A. Bennis and M. Leeser
- R3A p6 – “Analyzing the Use of GPUs for Hyperspectral Image Analysis,” Y. Gonzalez-Gonzalez and N. Santiago
R3B Posters

- R3B p1 – “Speeding up the Hyperspectral Image Analysis Toolbox,” S. Rosario and M. Velez-Rosario

- R3B p2 – “Accelerating Interactive 4D Motion Analysis with GPUs,” C. Mack, J. Hale and D. Kaeli

- R3B p3 - “The Gordon CenSSIS Citation Index System,” F. Yang, E. Arzuaga and The Spelman Team