Immersive Exploration of Big Data Arie Kaufman **Distinguished Professor and Chair CS** Chief Scientist, CEWIT **Co-Director IUCRC CDDA** Stony Brook University, NY

NSF Big Data, January 2015

Stony Brook University





MS in Data Science & Engineering

- MS in CS with Specialization in Data Science and Engineering
- 4 required courses:
 - Intro to Data Science
 - Statistics for Data Science
 - Algorithms
 - Al or Machine Learning
- 6-7 electives



Center for Dynamic Data Analytics (CDDA)

A National Science Foundation Industry-University Cooperative Research Center (IUCRC)

Jointly with Rutgers University

CDDA Mission

- Conduct integrative, multi-disciplinary research to manage, analyze, and visualize massive, complex, multidimensional and multi-scale dynamic data
- Turn chaos into knowledge
- Our Content of United Action of Content o
- Wide range of application domains (IT, healthcare, pharmaceutical, biotechnology, commerce, retail, finance, insurance, media, entertainment, transportation, logistics, manufacturing, defense, security, education)

What Does the CDDA Do?

 Create a partnership between academia, industry and government to advance dynamic data analytics and to address big data challenges in the application domains

Companies joint the CDDA for membership fee

share IP

support own interest (IP not shared)

Our Industrial Partners

- O CA Technologies Mobile data & security
- Northrop Grumman Big data analytics
- Samsung Mobile health
- Softheon Obama care
- Omni Scient Creations Social VR
- VJ Technologies Industrial imaging
- MedCAS Medical Imaging

Pending

- Motorola
- IBM
- Johnson & Johnson
- Verizon
- MediData
- Canon

Our Strategic Focus Areas

Managing Big Data

- Distributed & scalable storage systems/Cloud computing
- Cybersecurity
- Networking

Analyzing Big Data

- Machine learning and logic programming
- Natural language processing and text mining
- Image and video processing
- Biomedical imaging and informatics
- Statistical/stochastic analysis

• Visualizing Big Data

- Visual analytics
- Human-computer interaction
- Virtual reality facilities

Immersive Exploration of Big Data

• Visual Immersive Exploration

- Visualization of big data
- User interaction with virtual environment
- Analysis/exploration Visual Analytics
- Immersive Cabin
 - 5-wall immersive stereo virtual reality

Reality Deck

- Unique Gigapixel immersive visualization facility
- Very large-scale visualization
- Multi-modal interaction; multi-user

The Immersive Cabin

The Immersive Cabin

Visualization

- 5-wall enclosed Virtual Reality room/cave
 - 2 projectors per wall, active stereo
- GPU Cluster: 5 workstations with NVIDIA Quadro GPUs
- Tracking system
 - 8 camera IR tracking system
 - User wears passive reflective markers
- Interaction tools
 - Application specific navigation devices
 - Wireless gamepad, 3D mouse, tablet, etc.
 - Gesture-based user interface
- Software
 - Visualization, distributed rendering, synchronization, physics

The Immersive Cabin



Procedurally generated urban scene Realtime raytracing with global illumination

New CS building at SBU

Volumetric smoke dispersion in NYC

11 - H S 11 - H =



IC Applications – Architectural Design



Advanced Energy Research and Technology Center (AERTC), SBU Simmons Center for Geometry and Physics, SBU

IC Applications – Microtomography



Collaboration with Brookhaven National Laboratory





IC Applications – Medical Visualization



Immersive Virtual Colonoscopy

Top View 6-sided CAVE



Front View 6-sided CAVE Front View 5-sided CAVE with Conformal Mapping



Immersive Data Structure Visualization

Top View

Top View 6-sided CAVE

Front View 6-sided CAVE Front View - Conformally Mapped

Front View 5-sided CAVE with conformal mapping

Immersive Cabin - Limitations

Low resolution

- 1400x1050 per wall, approx. 12dpi
 - Highest resolution CAVE is 100 Megapixel
- Fine details in the data are lost

Image quality issues

- Limited contrast, dynamic range and color accuracy
- Not suitable for critical applications
 - Medical Visualization
 - Surveillance

• Limited physical space

- Floor surface is approx. 3m x 3m
- Limited opportunities for collaboration

Reality Deck addresses these limitations







The Reality Deck

The Reality Deck



Device for Big Data and Data Science

- NSF MRI grant
- Range of big data applications
- O 31 active applications
 - architecture, medicine, homeland security,, reconnaissance, scientific visualization, super-computing simulation, astronomy, climate modeling, electrical smart grid control center, G-pixel cameras, drug design, mapping, education, entertainment, etc.

The Reality Deck – The Facility

One-of-a-kind Exploration Facility

- 1.5 billion pixels
 - For visualizing big data from research / industry
 - Approaching the human visual acuity
- Immersive visualization
 - Data surrounds the user
 - Novel interaction techniques

Output States of Unique assembly of

- High-resolution LCD displays
- Large-scale GPU cluster
- Sensors, communication/networking devices
- Computer vision and human-computer interaction technologies

The Reality Deck - Hardware

• Visualization

- 30'x40'x11' environment
- 416 UQXGA LCD/LED Displays
- Samsung 27" S27A850D 2560x1440
 - Fed by 50ft-100ft DisplayPort cables
 - Fast response time, wide viewing angles, good dynamic range

GPU-Cluster

• 20-node GPU cluster, each node equipped with:

- 2x Six-core CPUs, 48 GB RAM
- 4x AMD FirePro V9800, 4GB RAM
- Each AMD drives 6 screens
- AMD S400 hardware video synchronization card
- 40Gb Infiniband adapter
- 1TB storage





The Reality Deck - Interaction

- 24 camera optical tracking system
- 24.4 immersive sound system
- Various input devices
 - tablets, Kinect, Wii remote, gamepad, SpaceNavigator, PixelSense table, ...

Gestural User Interface Prototype



The Reality Deck – in Numbers

- 1,533,542,400 pixels (1.5 Gigapixel)
- Over 6 miles of DisplayPort cables
- 240 CPU cores: 2.3 TFLOPs, 1.2 TB distributed memory
- 80 GPUs: 220 TFLOPs, 320 GB distributed memory

Applications in the Reality Deck

Wide-Angle G-pixel Video Cameras

- G-pixel camera system for military and civilian government applications
 - Can monitor entire city at once from a single aerial platform
 - Current technology allows visualization only of a small portion of the video (ArguSight: 76M pixel frame)
 - Need to zoom & pan
- The RealityDeck can display the full data without panning and zooming



Duke AWARE Camera Project

- Duke Imaging and Spectroscopy Program (DISP)
- I: David Brady
- 98 microcameras through a single lens
- Simultaneous image capture for temporal coherency
- Gigapixel resolution in color



[Image Credit: DISP Website]







Gigapixel Image Visualization

Dubai dataset 45 Gigapixels, 180° horizontal field of view





Gigapixel Telescope

6 Gigapixel View of the Milky Way (NASA's GLIMPSE and MIPSGAL projects)



Close-up of a single display







VISTA Telescope @ ESO – Milky Way 9Gp



Infinite Canvas

- Map very large 2D datasets to the Reality Deck
 - Take advantage of the immersion
- A very natural exploration interface
 - Based on head-tracking only
 - Walking and turning
- Data summarization with spiral navigation interface
 - Intuitive mapping back to the original visualization

Multi-user interaction





Art Gallery -10,000 Paintings

- I700s to modern art
 - Sorted chronologically
 - Visualization uses the *Infinite Canvas* for exploration



NY City Gigapixel Aerial Photograph 3m/10m resolution elevation data 0.5m resolution aerial photograph











Climate Visualization

Mesh Visualization

Future NYC scene

- 40M polygons
- Interactive rendering speed









Protein Visualization

- Data from the Protein Data Bank (PDB)
 - Other databases supported too
- Solution State State
- E-Coli bacteria protein



Reality Deck White House Video



Reality Deck

Immersive Gigapixel Display for Big Data Analytics

Thank You !

Supported by multiple NSF (incl. NSF MRI), NIH and ONR awards,







Stony Brook University