Feature Extraction & Tracking: An Overview for Ultra-ScaleVisualization

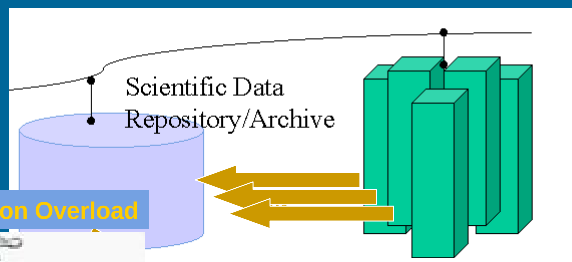
Deborah Silver
Prof, Department of Electrical and Computer Engineering
Rutgers, The State University of New Jersey
http://www.caip.rutgers.edu/vizlab.html

SC 06 Workshop on Ultra-Scale Visualization
Tampa, FL November 2006

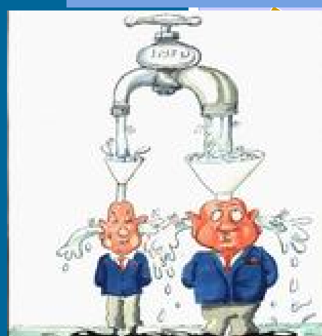
Outline – Retrospective on Feature Tracking

- Overview & Motivation
- Background in Feature Extraction & Tracking
- Review of Different algorithms & recent research
- Future

Post processing visualization workflow



Information Overload

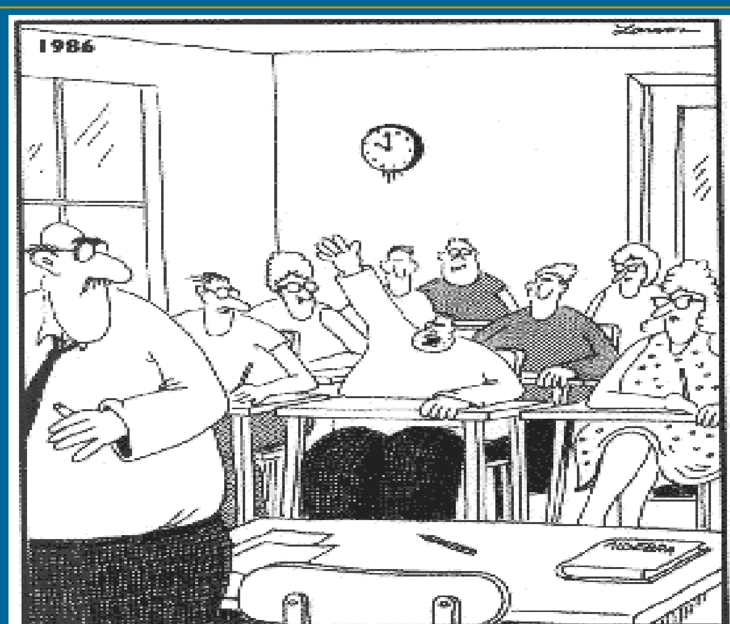


Supercomputing simulation

More Data

Even More Data

visualization

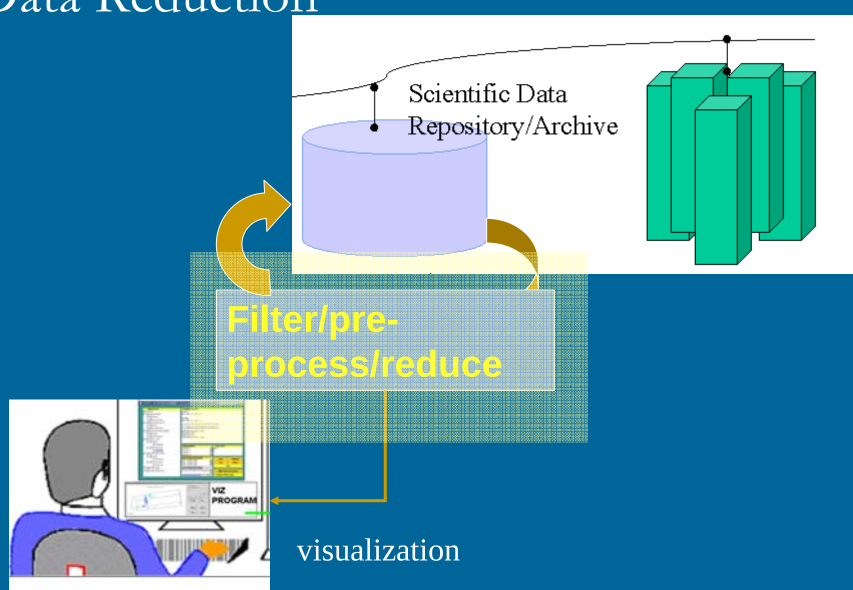


"Mr. Osborne, may I be excused? My brain is full."

Space/Time reduction

As the datasets become larger and larger, it becomes physically impossible to do in-depth discovery of all of the data. Filtering techniques are necessary to help the scientist focus on regions of interest in the ***space/time*** domain.

Data Reduction



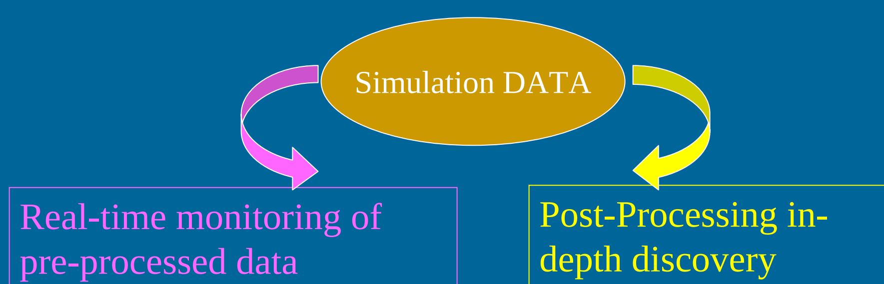
Effective Filtering

To Create *Effective Filtering + Visualizations* of Massive $3D^{+}$ Time Varying Simulation Data

➔

Feature Extraction & Tracking

Tracking can be done as part of a pre- or post-processing of the data.



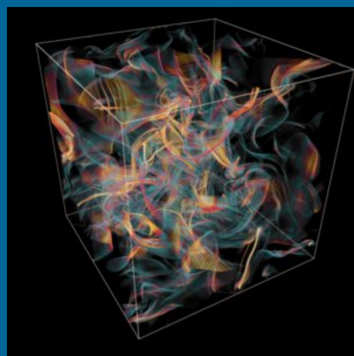
Pre-processing vs. Post-processing

- Automatic
- During simulation
- Produces still images and plots only
- Visualization parameters must be preset
- Requires separate data files for each time frame

- User-specified
- After simulation
- Allows user-interactive, in-depth analysis
- Visualization can be adjusted or optimized
- Allows data files to be whole or separate

What is tracking?

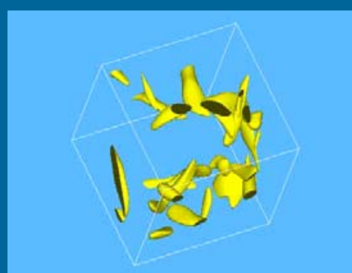
- Following "features" over time
- "Features" can be anything --- defined as coherent blobs/objects meeting certain conditions



Numerical Simulation of Compressible Turbulence:
D. Porter, P. Woodward, S. Anderson, K. Winkler and S. Hodson Oct 1997

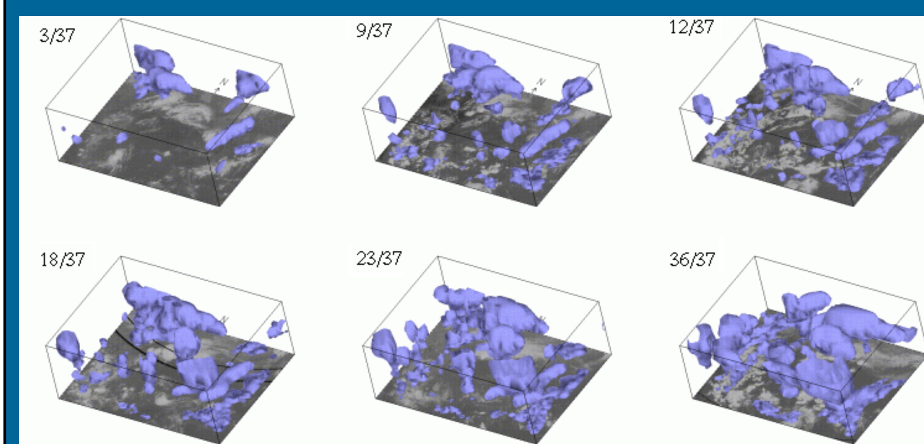
http://www.lcse.umn.edu

10GB Dataset (billion cell run, 512^3), vorticity scalar, volume rendering



Pseudo-spectral Simulation.
Isosurface of vorticity magnitude at 48% of maximum
128^3, 100 timestepsn (shown)

Weather Simulation

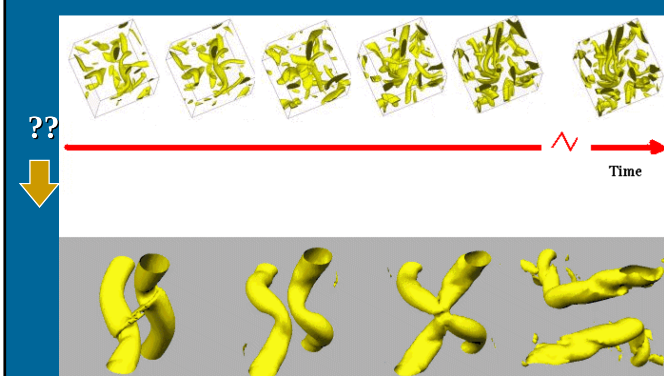


Data courtesy of Dr. YanChing Zhang at EPA

Difficulties:

- Size – too much data
- Clutter (Visual)
- Quantification: measurements
- Querying capabilities:
  - How many regions are there?
  - Where are the big regions?
  - Is “XXXX” present?
  - How does this compare to a different simulation?
- Classification: store in a database

Example-3D Event searching



Vortex Reconnection

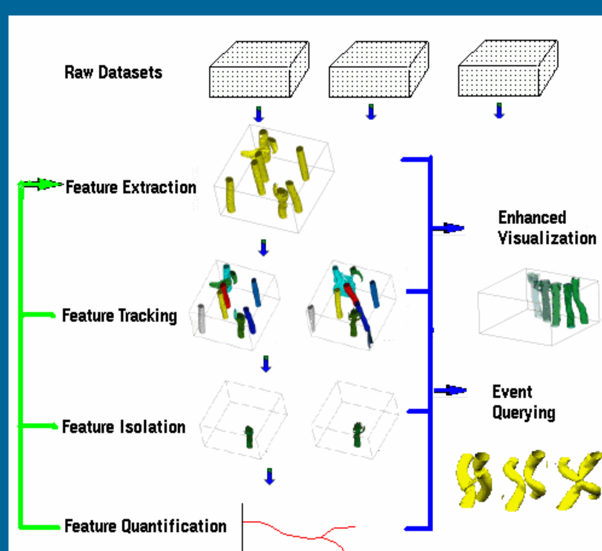
- Automatically find interesting events
- Follow Topological changes
- Classify events
- Search events

Feature & Event classification for Fusion-

Feature Based Techniques to characterize and catalogue interesting phenomena

Objects/Features	**Move**	**Interact ----Come Together/Apart**
(Plasma specific) Blobs Filaments Avaloids Striations Bursts Radial streamers IPO (Intermittent plasma Objects) Holes (opposite of blobs, density rarefactions) Chaotic Field line regions Wiggles Flux tubes Loss Cone *(CFD-general)* bubble hole blast wave packet blobcloud patch critical pt. point eddy ring Favor roll filament separatrix finger spike gyres spiral hairpin striation helix vortex	*(Plasma specific)* Spin/wake (blobs) Zonal flows Flow shears Rotation *(CFD general)* advect swirl entangle transport disperse wind flow hop migrate stream	*(Plasma specific)* Coalesce (blobs) Breakup (blobs) *(CFD General)* accrete aggregate align bind bifurcate burst collapse condense disassemble disrupt finger fission focus fold fuse pair roll-up plow reflect scatter spike split striate strip wind about

Feature-based Process Model



Major Components

- Feature Extraction
  - Define the features of interest. Domain dependent. Pre-defined or interactive.
- Feature Tracking
  - Automatically correlate extracted regions from one dataset to the next
- Quantification / Measurements for extraction & tracking.

--> BETTER VISUALIZATION

Features

- Basic definition
  - Regions of interest consisting of connected nodes satisfying some criteria (e.g. threshold interval)
- Each domain has its own definition
  - Volume intervals [G95]
  - Segmentation [S95]
  - Selective Visualization [W95]
  - Domain specific: Shock waves, Vortex Cores, Eddies, Medical ...

Thresholding Criteria

- Threshold criteria defining subsets of interest in a dataset are domain dependent
  - Volume intervals

    $$S(a,b) = (x, f(x) \mid a \le g(x) \le b)$$

  - Multi-Range Thresholding
    I.e. $(2<x<3, 6<x<8)$
  - Multi-Valued Thresholding
    Thresholding on multiple variables
  - Other more complex criteria
    e.g. combing vector data, etc.

3D Segmentation

- Region Growing
  - Select an initial seed(either interactively or automatically as extrema values)
  - Recursively test its neighbors depending on connectivity for inclusion, until no more points can be included.
  - New seeds are chosen, and repeat the above step until all the nodes that satisfy thresholding criteria are visited.

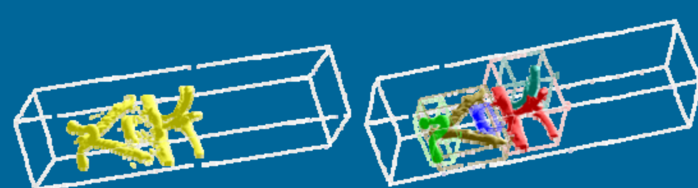
Object Information: Quantification

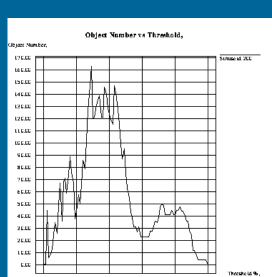
- Mass
- Centroid
- Local Extrema Values
- Volume
- Moments
- Bounding Surface (isosurface)
- Surface Area
- Skeletons
- Ellipsoids

Global Quantifications:

- Total number of objects
- Global extrema value
- Largest object

Example - Feature Extraction





Number of Objects per Threshold value

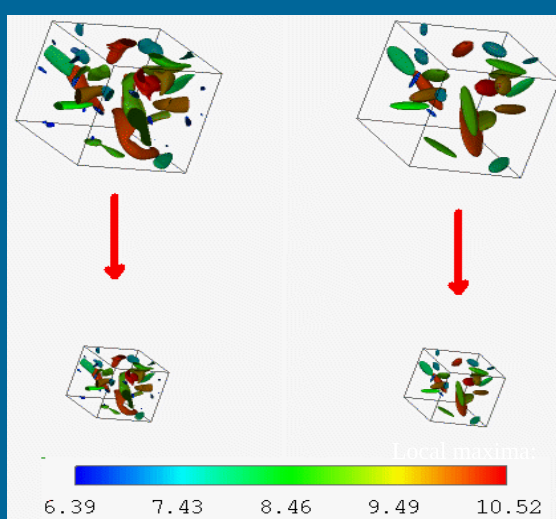
Object 1: Volume = 610, mass=1.2, ...

Object 2: Volume = 823, mass=1.8, ...

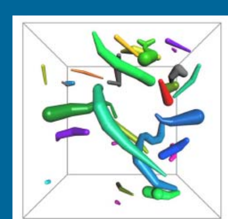
Feature Abstraction

- Abstract feature using a “reduced” representation object
- Compression of geometry
- Encapsulation of idea --- non-photorealistic rendering
- Reduced modeling

Abstraction: Data Reduction



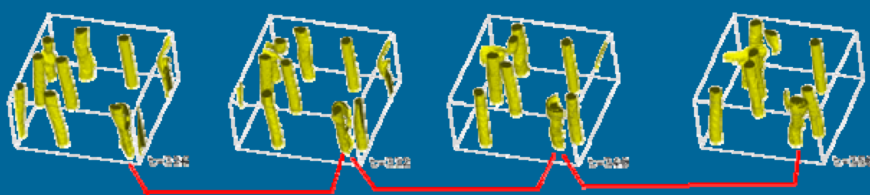
Abstraction using ellipsoids



Abstraction using skeletons

Feature Tracking

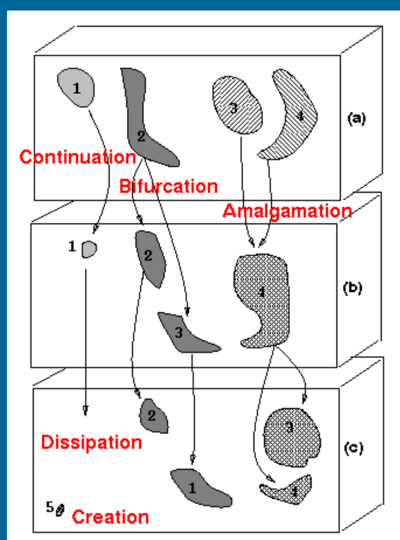
- Automatically correlate extracted regions from one dataset to the next



Assumption:
Sufficient Sampling Frequency
such that corresponding
features overlap in space.

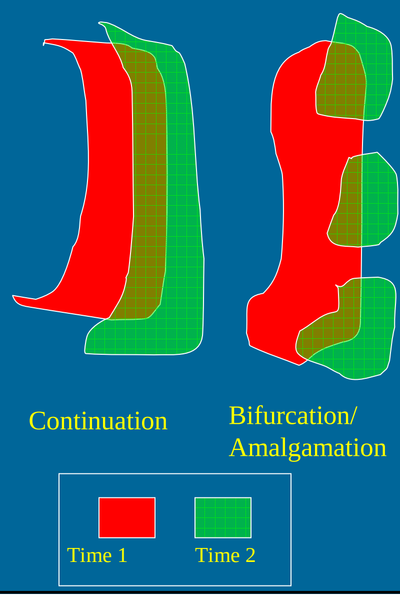
Tracking

- Continuation
- Bifurcation
- Amalgamation
- Dissipation
- Creation



Observations

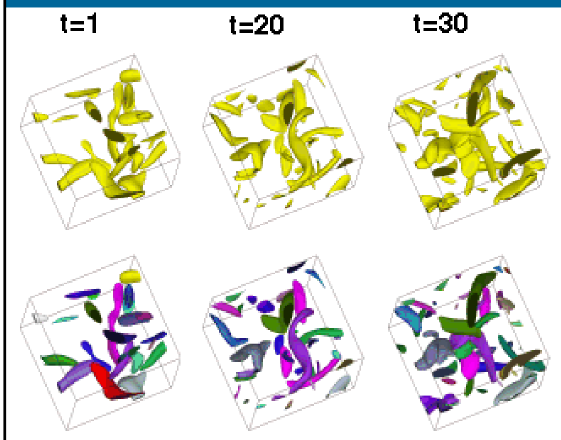
- **Continuation**: if feature O^{i}_{A} corresponds to O^{i+1}_{B}, then O^{i}_{A} overlaps with O^{i+1}_{B}.
- **Bifurcation** or **Amalgamation**: if a feature splits into a group of N objects, then all O^{i}_{A} in N overlap with O^{i+1}_{B}

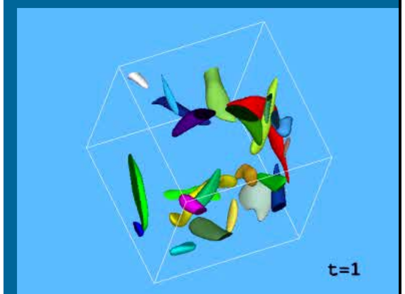


Visualization Paradigm

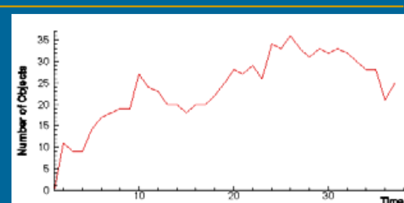
- DAG
- Enhanced surface rendering
- Enhanced volume rendering
- Feature isolation
- Trace and trajectory
- Feature Juxtaposition

Enhanced Surface Rendering

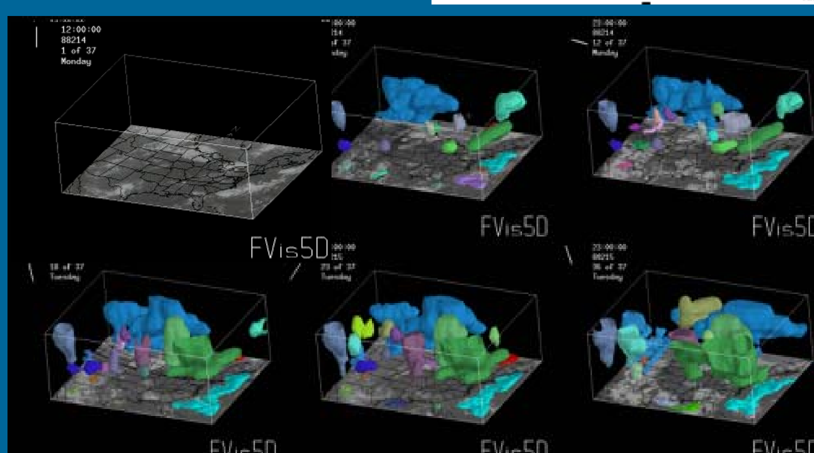




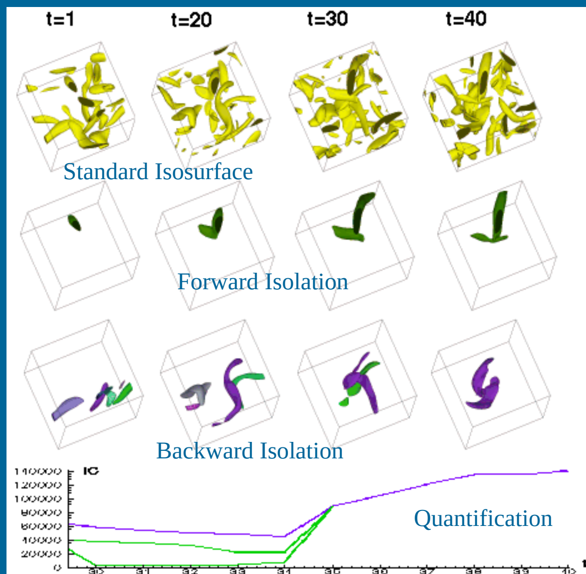
Dr. Zhang, EPA
Weather Simulation

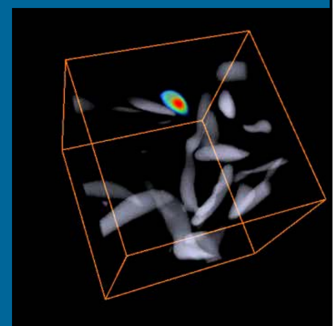


Graphs for Real-time monitoring

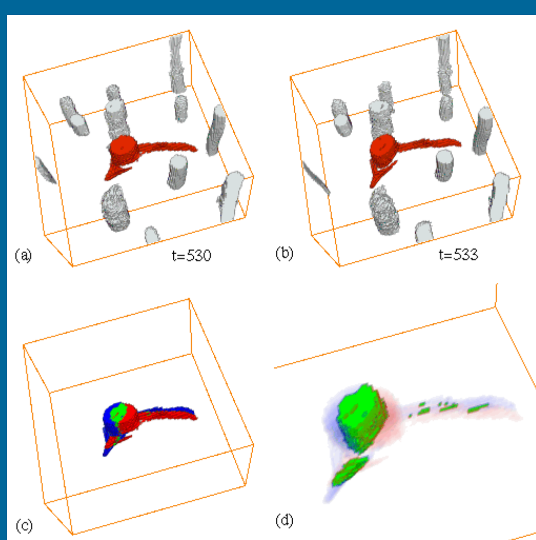


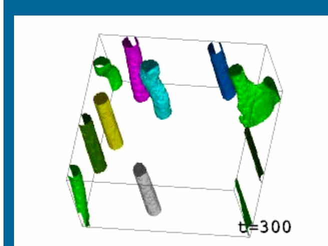
Feature Isolation/Quantification



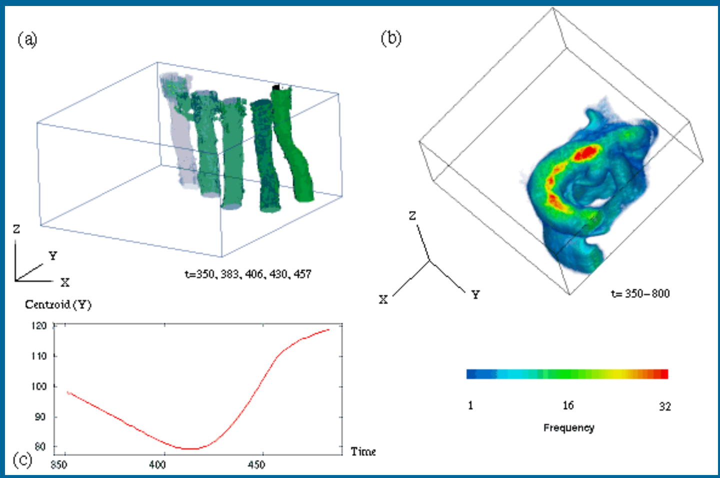


Juxtaposition





Trace and Trajectory



Tracking Methodology Review

- Define Feature
- Feature (vector) matching, i.e., centroid, area, skeletons, moments, etc.
- Overlapping Test

  Minimize the cases for best matching test.
  - Best Matching Test

    Object O_A and Object O_B are considered matched where

    $$R = \frac{Volume(O_A \cap O_B)}{\sqrt{Volume(O_A)Volume(O_B)}}$$
- Isosurface Tracking – 3D & 4D

Overlap

- D. Silver and X. Wang, Tracking and Visualizing Turbulent 3D Features. *IEEE Transactions on Visualization and Computer Graphics*, Volume 3, Number 2, June 1997.
- D. Silver and X. Wang, Tracking Features in Unstructured Datasets, Proceedings of IEEE Visualization '98 Conference, October 1998, Research Triangle Park, NC.
- J. Liang, H. Liu, D. Silver and M. Parashar, Rule-based Visualization in a Computational Steering Collaboratory, *International Conference on Computational Science (ICCS '04*), June 2004.
- J. Chen, D. Silver and J. Liang, The Feature Tree: Visualizing Feature Tracking in Distributed AMR Datasets, *IEEE Parallel Visualization and Graphics Symposium*, October 2003.
- J. Chen, D. Silver and M. Parashar, Real-time Feature Extraction and Tracking in a Computational Steering Environment, *Proceedings of Advanced Simulations Technologies Conference (ASTC '03)* March 2003.
- J. Chen, D. Silver and Y. Kusurkar, Distributed Feature Extraction and Tracking, SPIE IS&T Electronic Imaging Symposium, Visualization and Data Analysis Conference, January 2002, San Jose, CA.
- D. Silver and X. Wang, Volume Tracking. *IEEE Visualization '96 Conference Proceeding*. October 1996.
- T. Walsum, F. Post, D. Silver and F. Post Feature Extraction and Iconic Visualization. *IEEE Transactions on Visualization and Graphics*, July 1996.
- P. Rona , D. Jackson, K. Bemis, C. Jones, K. Mitsuzawa, D. Palmer, and D. Silver, *Acoustic Advances Study of Sea Floor Hydrothermal Flow*, EOS, Transactions, AGU, Vol 83, Number 44, October 2002.
- D. Silver and N. Zabusky, Quantifying Visualizations for Reduced Modeling in Nonlinear Science: Extracting Structures from Data Sets. *Journal of Visual Communication and Image Representation*, Volume 4, Number 1, March 1993.
- N. Zabusky, O. Boratov, R. Pelz, M. Gao, D. Silver and S. Cooper, Emergence of Coherent Patterns of Vortex Stretching During Reconnection: A Scattering Paradigm, *Phys. Rev. Letters* 67(18), 28, pp.2469-2472, October 1991.
- X. Wang and D. Silver, Visualizing Evolving Scalar Phenomena, *Journal of Future Generation Computer Systems (FGCS)*, Elsevier, February 1999 Volume 15, Number 1, p 99-108.
- D. Silver, Object Oriented Visualization. *IEEE Computer Graphics and Applications*, Volume 15, Number 3, May 1995.

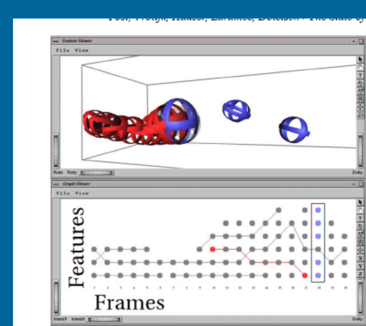
Attribute Based Feature Tracking,

F. Reinders, F. H. Post, H. J. W. Spoelder; VisSym 99

Visualization of Time-Dependent Data using Feature Tracking and Event Detection

The Visual Computer, 2001

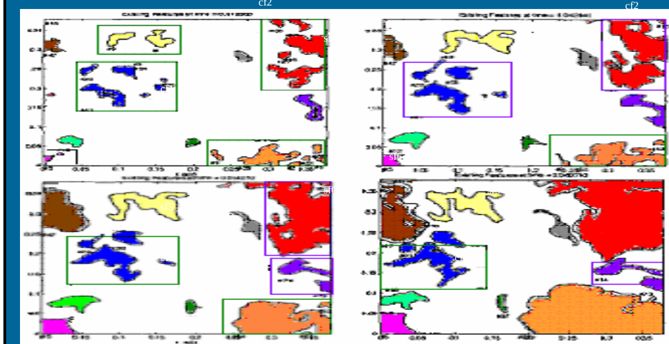
- Fits skeletons to regions of interest
- Compares attributes such as position, size, and orientation to determine correspondence
- Uses “event graph” to visualize feature development and interactions



Case Study: Application of Feature Tracking to Analysis of Autoignition Simulation Data

W. S. Koegler; Vis. 2001

- *Box overlap, matched left over for small fast-moving regions*

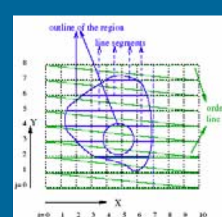


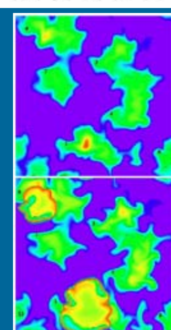
Four time steps are shown in an autoignition (HO_2) simulation, t= 6, 12, 15, 18. Time tracking is used to correlate and color-code features from one time step to the next. Large features ignite.

Using Bitmap Index for Interactive Exploration of Large Datasets

K. Wu, W. Koegler, J. Chen, A. Shoshani; SSDBM 2003

- **Preprocessing outline:**
  - Choose a bin arrangement for each parameter to partition the data
  - Generate one bitmap for each bin of each parameter. Each bit is calculated as
    ((paramValue >= binThreshold) ? 1 : 0)
  - Compress bitmaps as they are written, taking advantage of the connectedness of regions of 1's and 0's
- **Process outline:**
  - Define features by a user-composed query (e.g., x > a AND y > b) that is satisfied by performing boolean operations on the bitmaps, resulting in another bitmap
  - Use a form of region growing to locate individual connected regions (features), and write a compressed bitmap for each
  - Track features by using AND operations to detect overlap
- **Key aspects:**
  - Bitmap preprocessing:
    - accelerates later operations
    - requires significant disk I/O
    - May take several queries (~20) to justify the processing time
    - Well suited to parallelization
  - Feature extraction can be based on more than one criterion

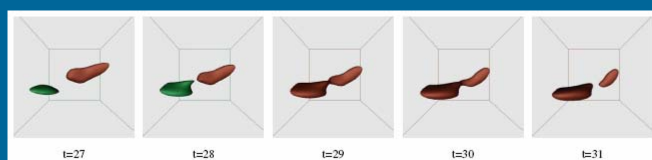




Volume Tracking Using Higher Dimensional Isosurfacing,

G. Ji, H.-W. Shen, R. Wenger; Vis. 2003

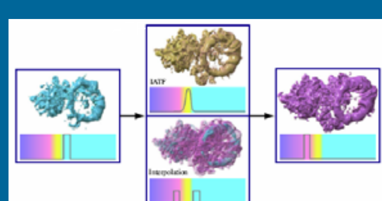
- Constructs an artificial fifth dimension to aid calculation of 4D interval volumes
- Uses an overlap threshold to determine feature correspondence
- **Key aspects:**
  - Uses the higher dimensional isosurfaces to lower computation cost and accelerate tracking



Intelligent Feature Extraction and Tracking for Visualizing Large-Scale 4D Flow Simulations

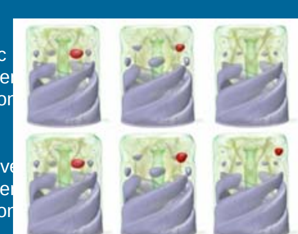
F.-Y. Tzeng, K.-L. Ma; SC 2005

- **Process outline:**
  - User input is combined with data and histograms and fed to a neural network to train the process
    - An overall interactive adaptive transfer function (IATF) is created once the user has chosen transfer functions for a few keyframes
    - Features of interest are tracked by position and internal values, influenced by the transfer function, once the user has chosen examples from a few keyframes
  - Feature correspondence is determined by 4D region growing, after applying the adaptive transfer function to each step
- **Key aspects:**
  - Uses machine learning through a neural network to produce intelligent results
  - Tracked features are delimited by an adaptive transfer function, allowing tracking of features over changing data values



static
transfe
functio

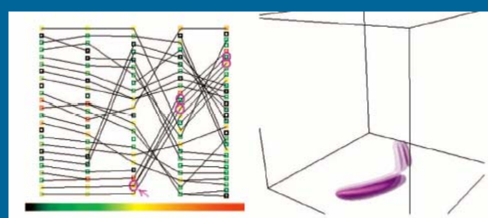
adaptive
transfe
functio



Time-Varying Contour Topology

B.-S. Sohn, C. Bajaj; *IEEE Trans. on Vis. and CG,* 2006

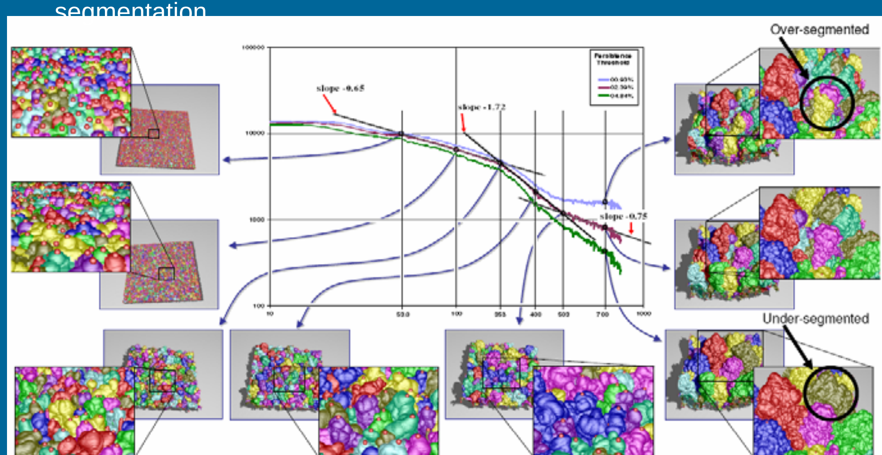
- **Process outline:**
  - Calculate isosurfaces
  - Build contour trees that describe topology changes of the time-varying isosurfaces
  - Calculate feature attributes (e.g., surface area, volume)
  - Through GUI, user can select a node of the contour tree to highlight the related feature
  - Surface extraction at runtime can be accelerated by growing the surface from a seed cell stored in the related node of the contour tree
- **Key aspects:**
  - Allows interactive isosurface tracking of a tetrahedral dataset
  - Provides the graph of topology changes as a way for the user to detect significant developments



Understanding the Structure of the Turbulent Mixing Layer in Hydrodynamic Instabilities

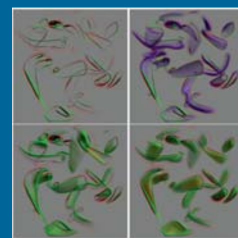
D. Laney, P.-T. Bremer, P. Miller, V. Pascucci; Vis 2006

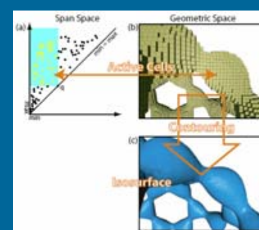
- Application, using the critical points of an surface to automate internal segmentation



Also of recent interest…

- “Multi-variate, Time-varying, and Comparative Visualization with Contextual Cues”, J. Woodring, H.-W. Shen; Vis 2006
- “Using Difference Intervals for Time-Varying Isosurface Visualization”, K. W. Waters, C. S. Co, K. I. Joy; Vis 2006





“Feature Tracking Using Earth Mover's Distance and Global Optimization”,

Guangfeng Ji, Han-Wei Shen, PG 2006

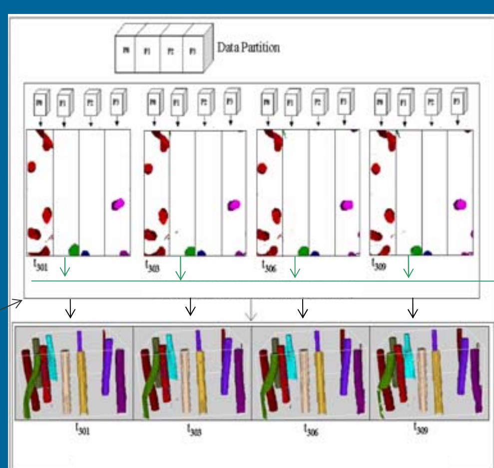
The Earth Movers Distance (EMD) computes the distance between two distributions, which are represented by signatures. The signatures are sets of weighted features that capture the distributions. The EMD is defined as the minimum amount of work needed to change one signature into the other. The notion of "work" is based on the user-defined ground distance which is the distance between two features. The size of the two signatures can be different. Also, the sum of weights of one signature can be different than the sum of weights of the other (partial match). Because of this, the EMD is normalized by the smaller sum.‘

Intuitively, given two distributions, one can be seen as a mass of earth properly spread in space, the other as a collection of holes in that same space. Then, the EMD measures the least amount of work needed to fill the holes with earth. Here, a unit of work corresponds to transporting a unit of earth by a unit of *ground distance*

Tracking issues for Ultra-scale Visualization

- Post-processing tracking too slow, data too large → tracking while simulation is progressing
  - Feature extraction must be preset
  - Mechanism to change thresholds etc..
  - Quantities extracted
- Real time steering
- Can be used for simulation (feedback to simulation)
- Multiresolution data
- Database classification for discovery

Challenges- robust code, distributed code, standard in vis packages

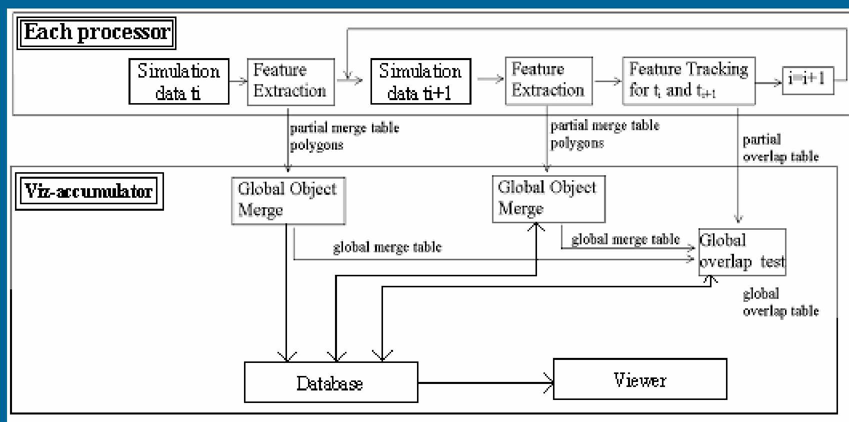


Partial merge across processor boundaries

Feature tracking from t_i to t_{i+1}

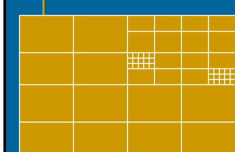
Distributed feature extraction and tracking. Each processor computes a partial extraction and tracking using ghost communications. The full solution is then merged.

Distributed Feature Tracking



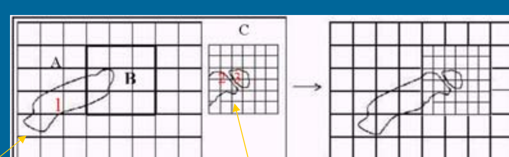
Challenges…

AMR datasets: Adaptive Mesh Refinement



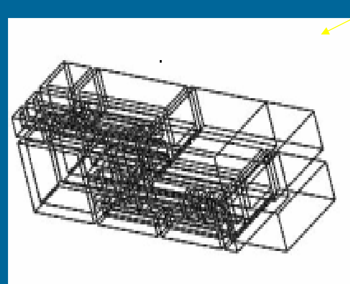
2D AMR

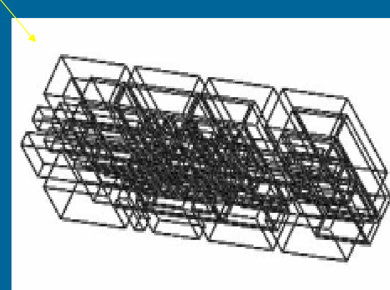
Level 1 Level 2



A splits into two pieces in Level 2: C2 and C3

Regridding from ti to ti+1





Future Research: Event Classification

- For massive simulations
- Database querying functions (pre-processed)
- Event tracking

Characterization of events for combustion

Objects	Move	Interact ----Come Together/Apart
Combustion specific	*Combustion specific*	*Combustion specific*
Kernal Flame (premixed, part, diffusion,edge) Shockwave Fuel Jet Stochiometric Line Region of chemical reaction Region of Flow	burn curve convect diffuse ignite quench percolate propagate reignite react strain	flame-wall interactions merge annihilate upstream annihilate downstream
CFD-general	*CFD-general*	*CFD General*
bubble blast wave blobcloud critical pt. eddy hole packet patch pint ring favor filament finger gyre hairpin helix roll separatrix spike spiral striation vortex	advect entangle disperse flow hop migrate stream swirlt transport wind	accrete aggregate align bind bifurcate breakup burst collapse condense disassemble disrupt finger fission focus fold fuse pair roll-up plow reflect scatter spike split striate strip wind about

Feature & Event classification for Fusion-

Feature Based Techniques to characterize and catalogue interesting phenomena

Objects/Features	**Move**	**Interact ----Come Together/Apart**
(Plasma specific) Blobs Filaments Avaloids Striations Bursts Radial streamers IPO (Intermittent plasma Objects) Holes (opposite of blobs, density rarefactions) Chaotic Field line regions Wiggles Flux tubes Loss Cone *(CFD-general)* bubble hole blast wave packet blobcloud patch critical pt. point eddy ring Favor roll filament separatrix finger spike gyres spiral hairpin striation helix vortex	*(Plasma specific)* Spin/wake (blobs) Zonal flows Flow shears Rotation *(CFD general)* advect swirl entangle transport disperse wind flow hop migrate stream	*(Plasma specific)* Coalesce (blobs) Breakup (blobs) *(CFD General)* accrete aggregate align bind bifurcate burst collapse condense disassemble disrupt finger fission focus fold fuse pair roll-up plow reflect scatter spike split striate strip wind about

Why Feature Tracking

- Reduce the Size of Data
- Reduce Complexity
- Provide Quantification
- Enhance Visualization
- Facilitate Event Searching

Thank You!!

http://www.caip.rutgers.edu/vizlab.html

VIZLAB

Visualization and Graphics

Funding acknowledgements: DOE, NSF