Readily Regenerable Microstructure Representations

2007 ASM/TMS Symposium

Padma Raghavan
Keita Teranishi, Long-Qing Chen and Zi-Kui Liu

The Pennsylvania State University

CCMD@GaTech+PSU

MATCASE@PSU
Outline

- Introduction
  - Matcase: An E-laboratory for materials design
  - Role of Microstructures
- Readily Regenerable Representations
  - Framework
  - Compression schemes
  - Results: compression and quality
- Conclusions
Introduction
Matcase: E-Lab for Materials Design

- An extensible interactive system for automated design space exploration
- Users (clients) connect to initiate materials design via web-portal
- Web-portal creates a service to the user and executes remote tasks
- Remote tasks are managed by Globus-enabled services
User View

- Users (clients) connect to initiate materials design via web-portal
- Web-portal *creates a service to the user and executes remote tasks*
- Remote tasks are managed by OGSA-enabled services suitable for wide-area parallel computing
System Architecture

A reconfigurable system with 3 main services

- Interaction handler
  - Gets input from clients and provides intermediate/final results

- Analyzer
  - Creates instances of interaction and simulation handlers
  - Manage “rules” for meaningful composition
  - Bridges between interaction handler and simulation handler for each client

- Simulation handler
  - Executes remote tasks using Globus grid-services
  - Creates instances of local “services” to process input/output between steps
  - Transfers input/output for client between the server and remote computers
Server is not developed directly

Server-generator produces server from specifications

Remote resource

Credentia regression (Myproxy)
Example: Automating Phase Stability Map Generation
Method: Adaptive Recursive Parameter Space Partitioning

- Create a regular mesh on parameter
  - Compute phase at each mesh point
  - If adjacent mesh points are in different phases, create finer mesh and check the phase of each new mesh point
- Repeat until
  Small domains
  All adjacent domains have same phase
Sample Results:

- **Initial Results:** More accurate diagram with smaller number of simulations
- **Ongoing:** more evaluations
  - Scheduling to maximize throughput
  - Allowing asynchronous parallelism
Microstructure Repositories For Fast Analysis

Problem: microstructure data is large
- 40 Mbytes for 1024x1024 mesh, over 1 TByte for one binary system for main time-steps
- PBytes over many systems

Large data size slows computation
- Large storage system scaling costs
- Slower computation ~ disk speed factor 10,000 slower than peak rate of CPU
Role of Microstructures

- Microstructures = results of each design experiment
- Microstructure repository
  - Comparisons for verification/validation
  - Reduces computational time, re-direct/steer focus of the experiment for meaningful science
- Building block for reverse mode analysis -- determining composition, processing parameters for desired properties
Readily Regenerable Representations
Background

- Earlier work
  - 2-Point Statistical representation
    - Statistically accurate
    - Does not preserve morphology of the original data
  - Reduced Order Modeling
    - Needs a large data set of “basis” microstructures

- Our work:
  - Multilevel and sparse skeletal techniques for 2D/3D microstructure
  - Can be used for data mining, reduced order modeling
Readily Regenerable Representations

Our approach:
- Store a reduced representation that is very compact
- Regenerate by applying a few steps of phase-field simulation or other recovery/regeneration schemes

Data reduction techniques
- Image compression, e.g. JPEG
- New schemes:
  - Recursive formulations
  - Sparse skeletal representations
Readily Regenerable Compression Algorithms- 2D
Multilevel Method

- Resize image using bi-cubic interpolation
  - (1024x1024-512x512-256x256-……)
  - Store the smallest image

- Regenerate by enlarging image
  - Apply corrections for smaller images (if needed)
  - Numeric interpolation
**Sparse Skeletal Representation**

- Observation: Values at interface change sharply
- Microstructure = collection of detailed grain skeletons + aggregate values of matrix, grain interior
- Large factors of compression from area to perimeter relationship
- Order-of-magnitude savings: $\sim N \sqrt{N}$ of the total $N \times N$ points in an image
Experiments

- Ni-Al, Al-Cu microstructures
- Results = average for 10 microstructures
- Compression algorithms
  - Existing: **SVD**: Singular Value Decomposition (dimension reduction); **JPEG** and **Segmentation** (image processing)
  - New: Multilevel JPEG (**ML-JPEG**) and Sparse Skeletal Representations (**SSR**)
- Refinement: 10-50 time steps of phase field simulation
Ni-Al (gamma, gamma prime)

- 100,000 time steps@ 10,000 time steps
- 512x512 mesh, 2 nm
- Temperature: 1073K
- Al-concentration: 13.8%, 17.42%, 21.47%
- Each mesh point includes:
  - 1 value for concentration
  - 4 values for the order (orientation)
Al-Cu (alpha and theta-prime)

- 100,000 @ 10,000 time steps
- 512x512 mesh, 2 nm
- Temperature: 500K
- Cu-concentration: 8.82%
  - 8 and 16 particles in initial microstructure
- Each mesh point includes:
  - 1 value for concentration
  - 2 values for the order of orientation
JPEG vs SSR

Stress Field: Original

Observe Error Scale range between JPEG and SSR

SSR has no error in boundaries
Metrics

- **$M$ original**  $\hat{M}$ regenerated
- **Compression factor**  
  
  \[ F = \frac{\text{size}(M)}{\text{size}(\hat{M})} \]
- **Quality**

\[ \frac{\|\hat{M} - M\|}{\|M\|} \]
- **Error in order parameters**

\[ \max_{i,j} |\hat{M}_{ij} - M_{ij}| \]
- **Error in stress field**
Compression Factors

Height of bar: Compression factor
Segmentation: highest
SVD: lowest
Quality: Error In Stress Field

- Segment: Worst; SSR: Best
Quality: Error In Composition

- Darker shade: after refinement
- Segment: Worst; SSR: Best
Quality: Error In Order Parameter

- Darker shade: after refinement
- Segment: Worst; SSR: Best
Error Definition

- Thus far: $M(t)$ compared with $\hat{M}(t) + x$
  - $x$ steps of refinement
- Should we compared $M(t+x)$ with $\hat{M}(t) + x$?

Next: results of Ni-Al, for $x= 10, 30, 50$
## Ni-Al: Effect of Evolution

<table>
<thead>
<tr>
<th>Method</th>
<th>$x = 0$</th>
<th>$x = 10$</th>
<th>$x = 30$</th>
<th>$x = 50$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M_t$</td>
<td>$M_{t+x}$</td>
<td>$M_t$</td>
<td>$M_{t+x}$</td>
</tr>
<tr>
<td>Ni-Al system</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SVD</td>
<td>0.0726</td>
<td>0.0609</td>
<td>0.0600</td>
<td>0.0485</td>
</tr>
<tr>
<td>JPEG</td>
<td>0.0174</td>
<td>0.0159</td>
<td>0.0149</td>
<td>0.0169</td>
</tr>
<tr>
<td>ML-JPEG</td>
<td>0.0243</td>
<td>0.0274</td>
<td>0.0264</td>
<td>0.0312</td>
</tr>
<tr>
<td>SSR</td>
<td>0.0177</td>
<td>0.0150</td>
<td>0.0141</td>
<td>0.0140</td>
</tr>
</tbody>
</table>

- Refinement with $x$ time-steps represents further evolution
- Errors better defined with respect to $M(t+x)$
Readily Regenerable Compression Algorithms - 3D

80M bytes for 128x128x128 Ni-Al
3D Data Compression Methods

- There are NO standard data compression schemes
- Extend standard 2D JPEG
  - Partition data into 8x8 blocks
  - Discrete Cosine Transform (DCT) for 8x8 block to coefficients of 64 wave components
  - Quantize (truncate) insignificant coefficients
SSR for 3D

- Extends naturally—store detail at grain surfaces
- SSR should lead to O(N) savings
  - From volume to surface correspondence
  - NxNxN vs NxN
New Scheme: Multilevel Discrete Wavelet Transform (ML-DWT)

- Discrete Wavelet Transform
- Wavelets decompose data using a single wave with different scaling applied to different spatial domains
  - Global approximation by rescaling wavelength to large size (applied to large subdomain)
  - Local approximation through rescaling and shifting (applied to small subdomains)
ML-DWT: continued

- 3D DWT executes, 1D DWT for X, Y, and Z directions
  - Decompose to 8 subblocks
  - Subblock for lowest frequency mode is recursively decomposed
  - Transformed coefficients have a small range of variation
    - Good for lossless compression
Experiments

- 3D JPEG, 3D Wavelet, and 3D SSR
- 256x256x64 Ni-Al system
  - Composition, 4 Order Parameters
  - 160 Mbytes
  - More than a week to simulate
  - Captured at 3 different stages
Ni-Al Microstructures

- 2D cross section (at z=20) of 256x256x64 mesh

30min

180min

390min
• SSR has low error and high compression factor
• DWT is better than 2D JPEG (DCT)
Conclusions

- Framework and new data reduction schemes to enable microstructure repositories for fast analysis

2D results:
  - JPEG achieves good compression quality
  - SSR has best compression factor/quality tradeoff
    - Very accurate stress field values and order parameters

Extensions to 3D: SSR, ML-DWT better than JPEG

Cost of compression and regeneration is under 1% of phase field simulation time

2D results to appear Comp. Mat. Sci., ’07; 3D results under preparation.