

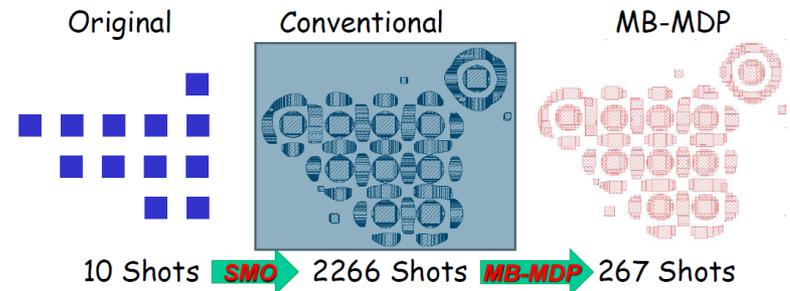
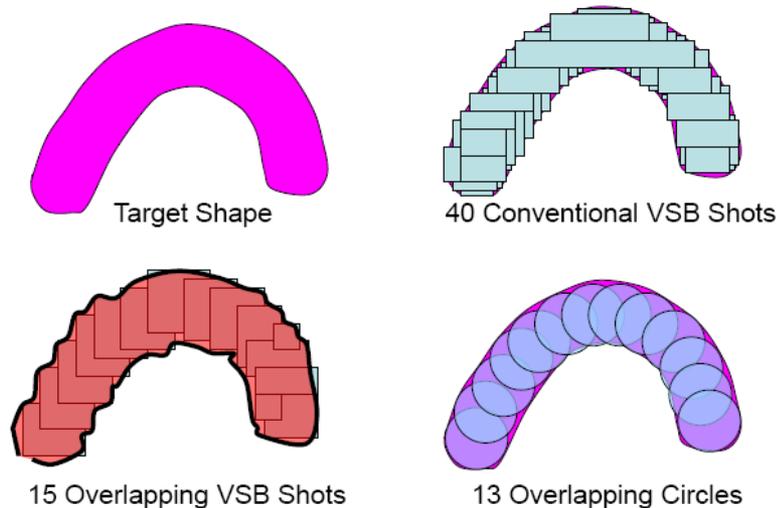
GPU-based MRC Methods for Overlapping eBeam Shots

Kokoro Kato, Ph.D.
Hitachi High-Tech Science Corporation

1. Model Based MDP (MB-MDP)
2. GPU computing and its benefits
3. Key points of GPU computing
4. Experimental results
5. Conclusion

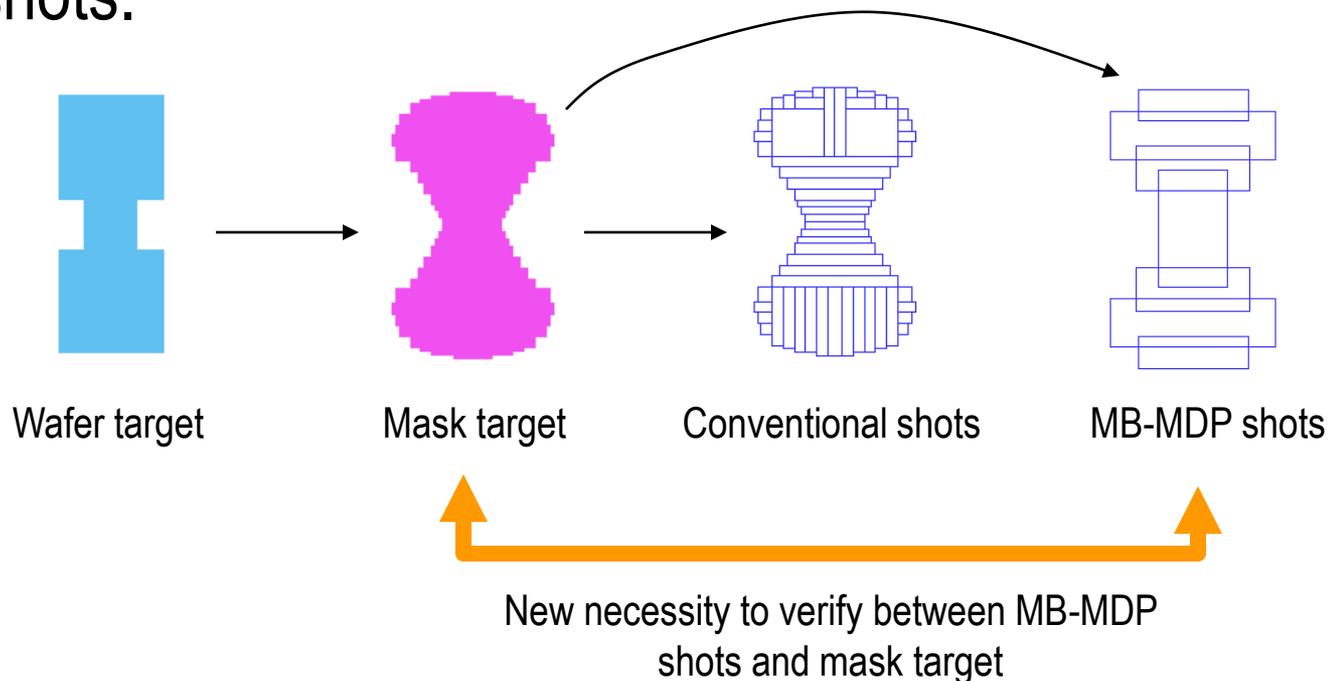
1. Model Based MDP (MB-MDP)

- Model Based MDP (MB-MDP) has been used for EB shot reduction and speed-up of mask writing time.
- By allowing overlapping shots, EB shot count can be significantly reduced.

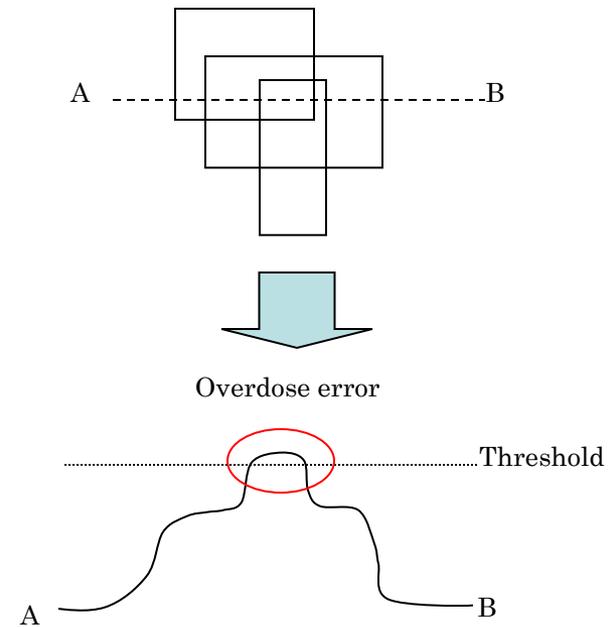
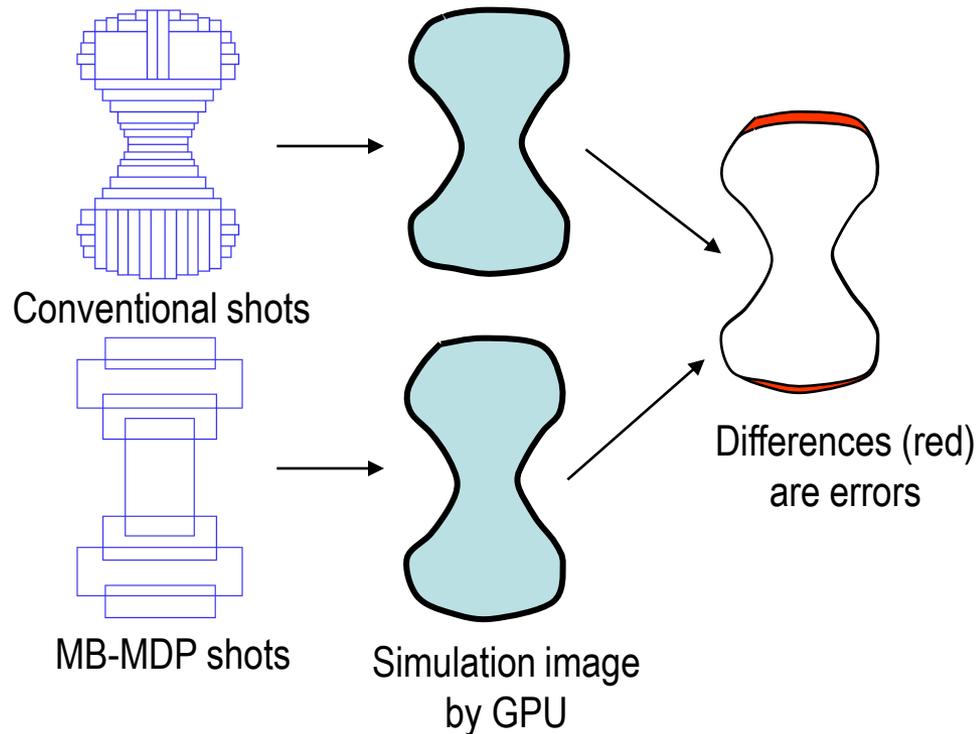


From <http://www.ebeam.org/home>

- Usually, EB shot data verification is done before mask writing
 - This process is called Mask Rule Check (MRC)
- However, MB-MDP shots are not identical to mask target. So, new methods are needed to verify the MB-MDP shots.

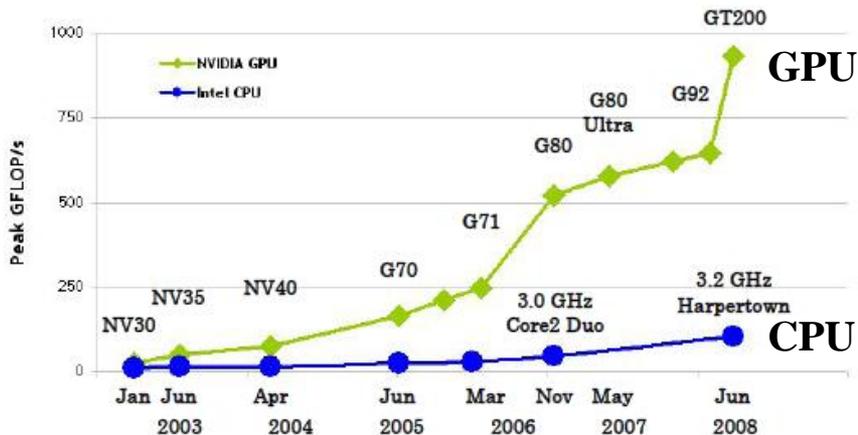


- In this presentation, we introduce two methods for verification of overlapping EB shots using GPGPU.
 - Comparison of simulated contour images between conventional shots and MB-MDP shots
 - Overdose detection



2. GPU computing and its benefits

- One of the greatest merits of GPU computing is its speed.
 - Currently many technologies of high speed computation based on GPUs have been reported.
 - Especially, an excellent program development environment called CUDA enables to build such high performance programs easily.



GT200 = GeForce GTX 280	G71 = GeForce 7900 GTX	NV35 = GeForce FX 5950 Ultra
G92 = GeForce 9800 GTX	G70 = GeForce 7800 GTX	NV30 = GeForce FX 5800
G80 = GeForce 8800 GTX	NV40 = GeForce 6800 Ultra	

GFLOP comparison (CPU vs GPU)



Architecture of CPU and GPU

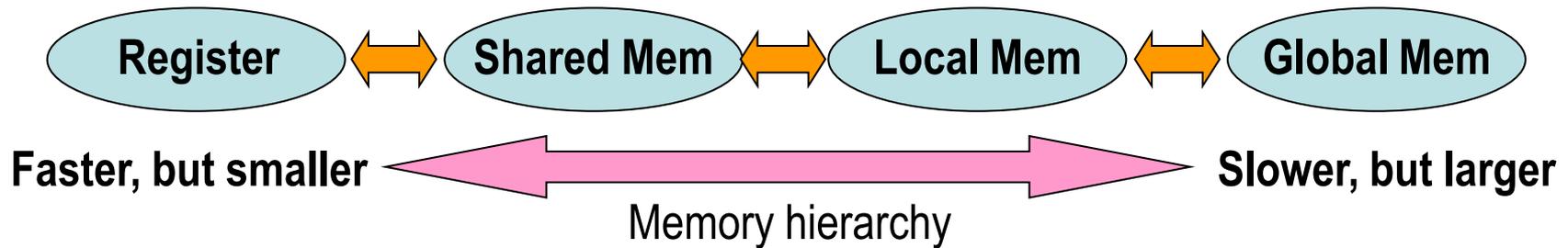
CPU has a single cache memory and controller and no more than 4 or 6 cores. On the other hand, GPU can have multiple cache memory and controllers, and more than one hundred cores.

- Another major merit of GPU is low power consumption.
 - Comparison of GFLOP per Watt
 - GPU is much better than CPU
- Computational lithography technologies need great amount of CPU powers.
 - Total reduction of power consumption is needed.
- In this study, we develop GPU computing techniques for faster and greener MRC

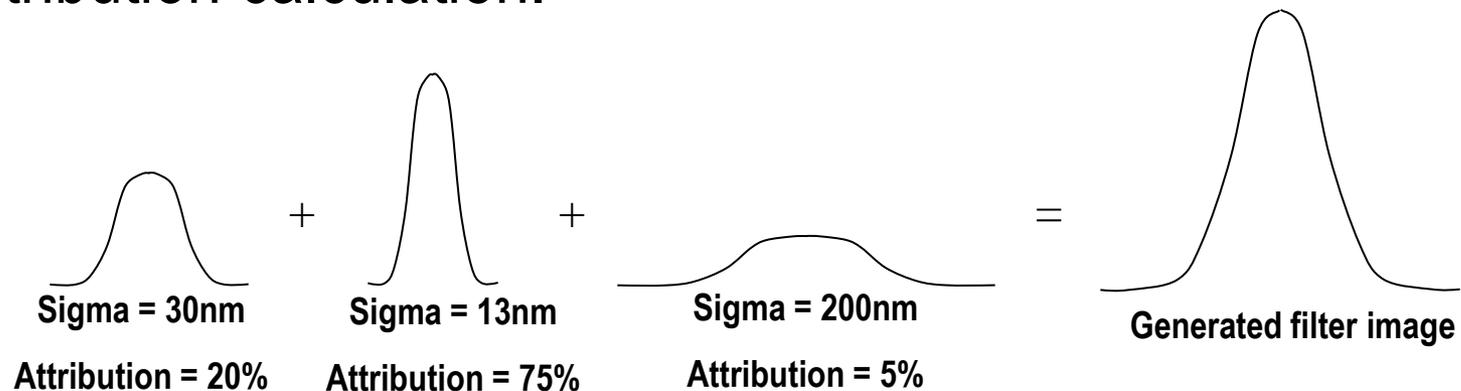
3. Key points of GPU computing

■ Need to pay high attention to optimal memory usage.

- Shared memory – small (16k), but very fast
- Global memory – large (1G), but very slow

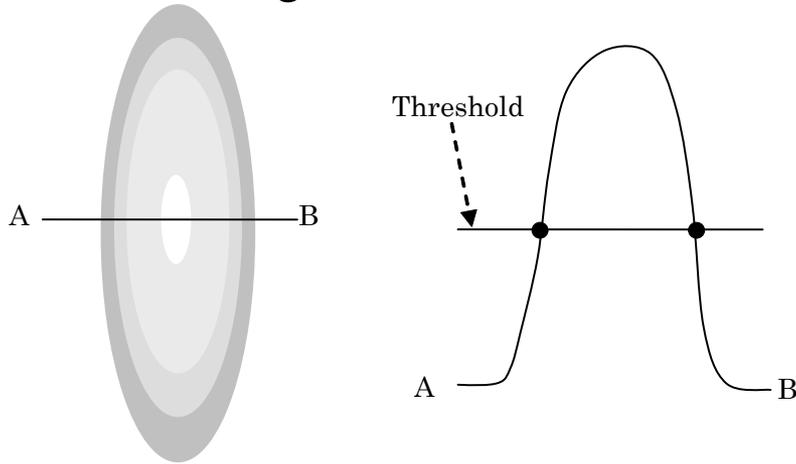


■ We applied multiple Gaussian filtering method for energy distribution calculation.



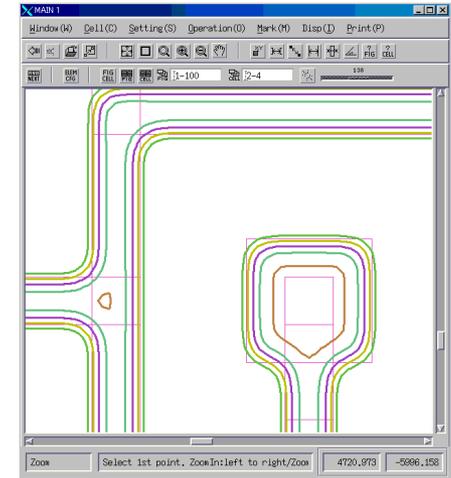
Contour generation

Contour generation flow



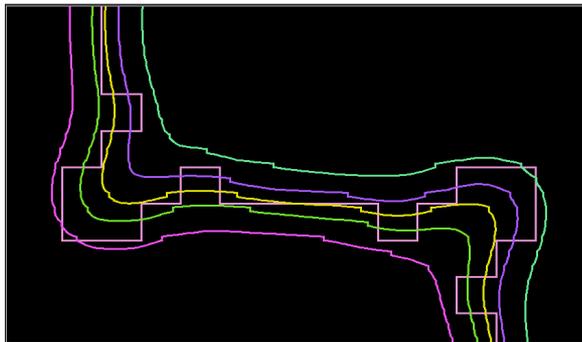
(a) Energy distribution and cut line (b) Intensity at cut line (c) Generated contour lines

Contour generation flow

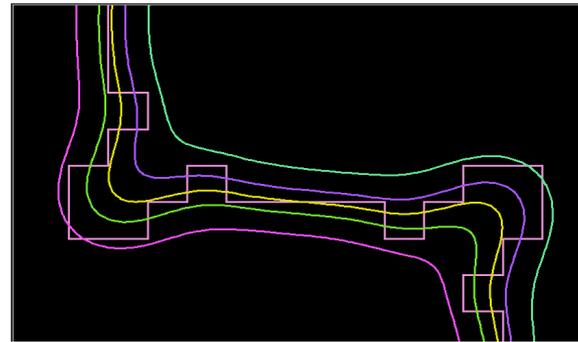


Generated contours (multi-threshold)

Smoothing is needed



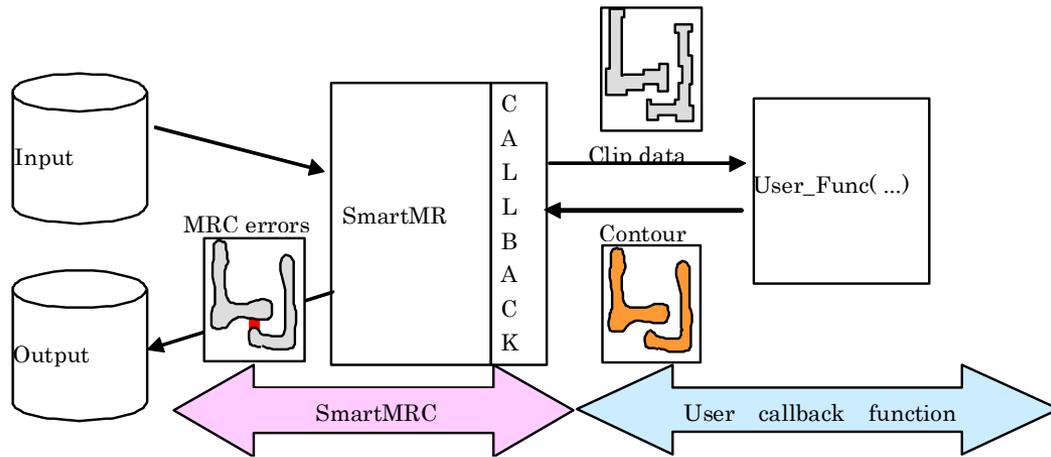
Before smoothing



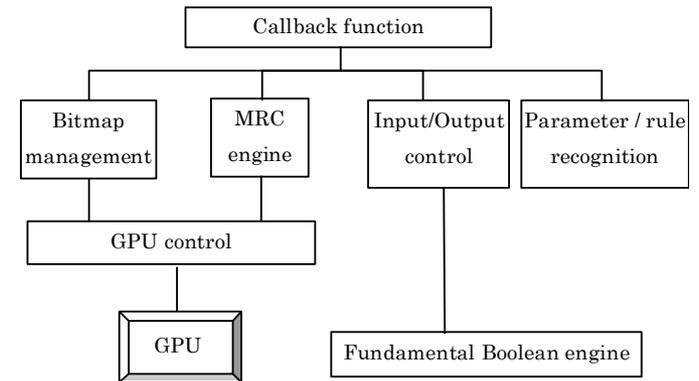
After smoothing

Program Architecture

- We use the API (Application Program Interface) function of SmartMRC as the fundamental framework of the software development and evaluation.
 - By using the API function of SmartMRC, programmers can access to the mask layout data easily without knowing the details of the mask data format.



API mechanism of SmartMRC



Program architecture

4. Experimental results

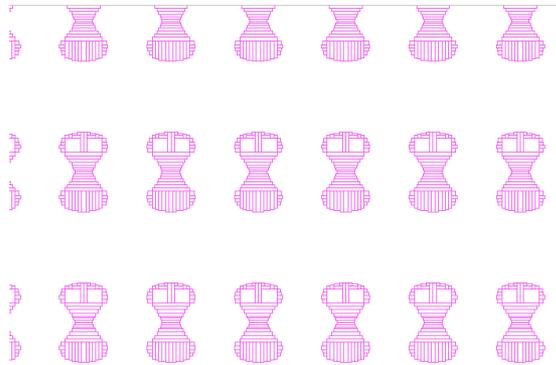
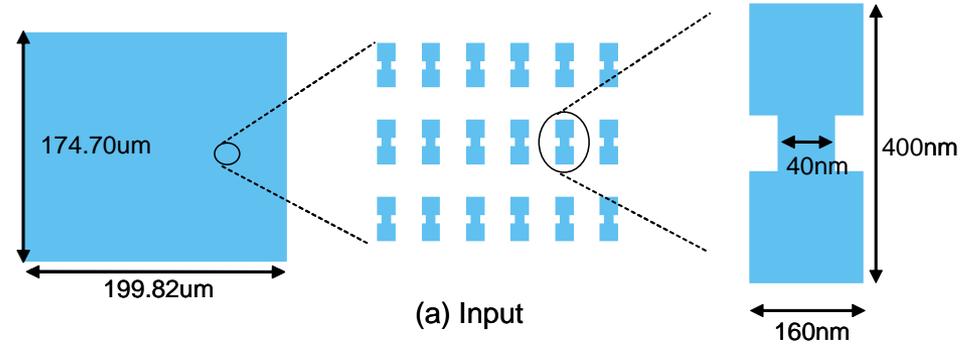
- We evaluated new MRC system with 2 types of data.
- The specs of the system is as follows:

CPU	: Intel® Xeon® CPU E5645 @ 2.40GHz X 2
Core number	: 24 logical cores
GPU	: Tesla C2050 / C2070
Core number	: 448 CUDA Cores
Global-memory size	: 2.64 Gbyte
GPU Clock-speed	: 1.15 GHz
Constant-memory size	: 64 Kbyte
Shared-memory per block	: 48 Kbyte
Register count per block	: 32,768

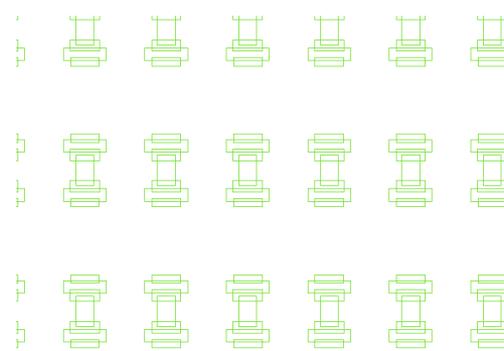
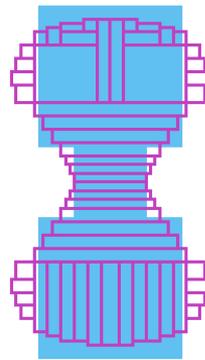
Dumbbell type data (DB)

- Dumbbell type data – array of dumbbell shape with height=400nm and width= 40/160nm.

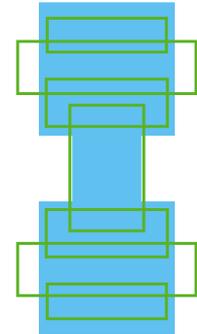
Type	Figure count
Input	125,000
Overlapping	875,000
Conventional	6,875,000



(b) Conventional shots



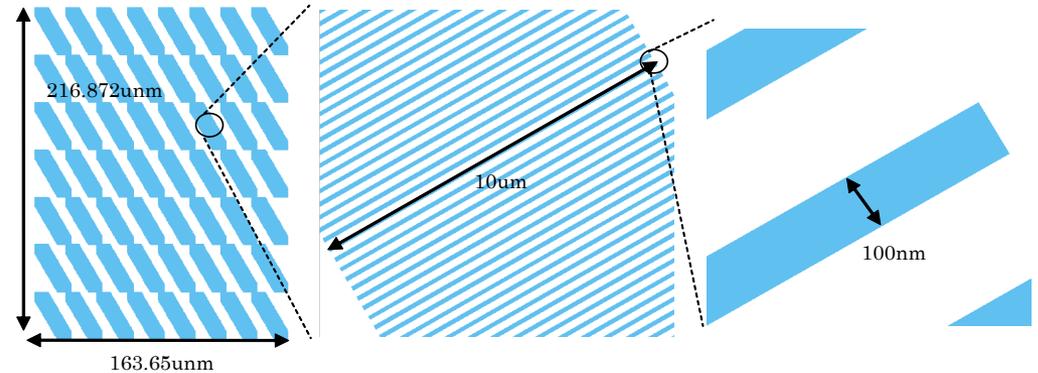
(c) Overlapping shots



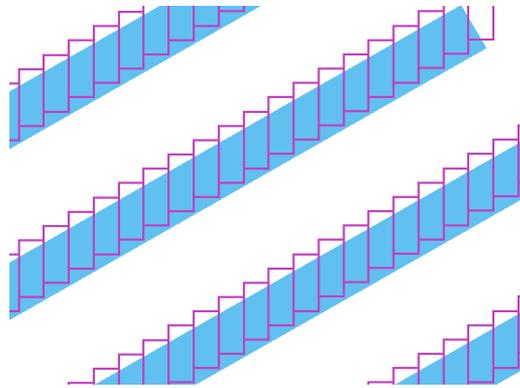
Angled-line type data (AL)

- Dumbbell type data – array of lines with angle=30degree and width= 100nm.

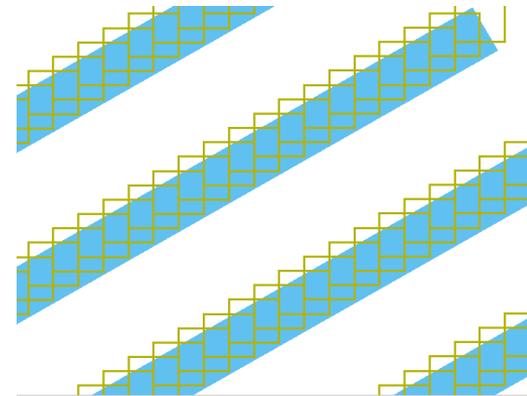
Type	Figure count
Input	5,600
Overlapping	417,600
Conventional	974,400



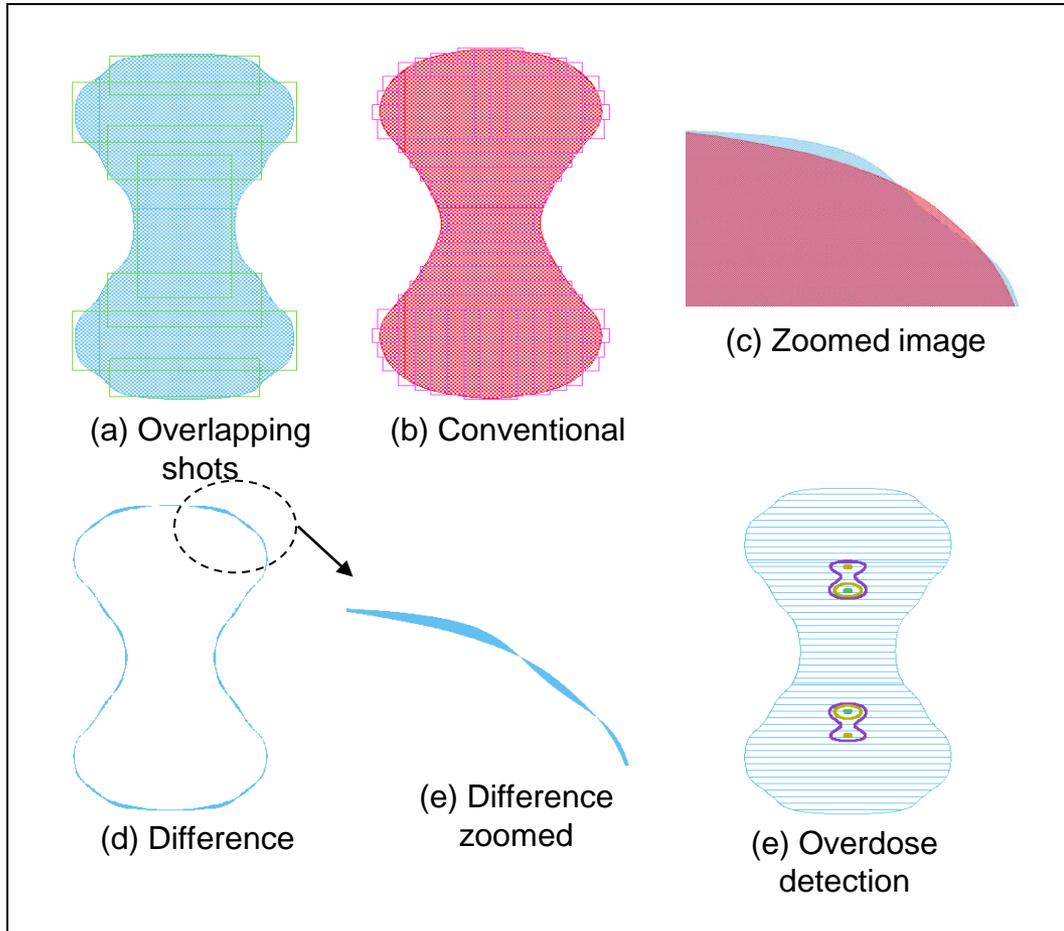
(a) Input



(b) Conventional shots



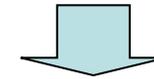
(c) Overlapping shots



Result images

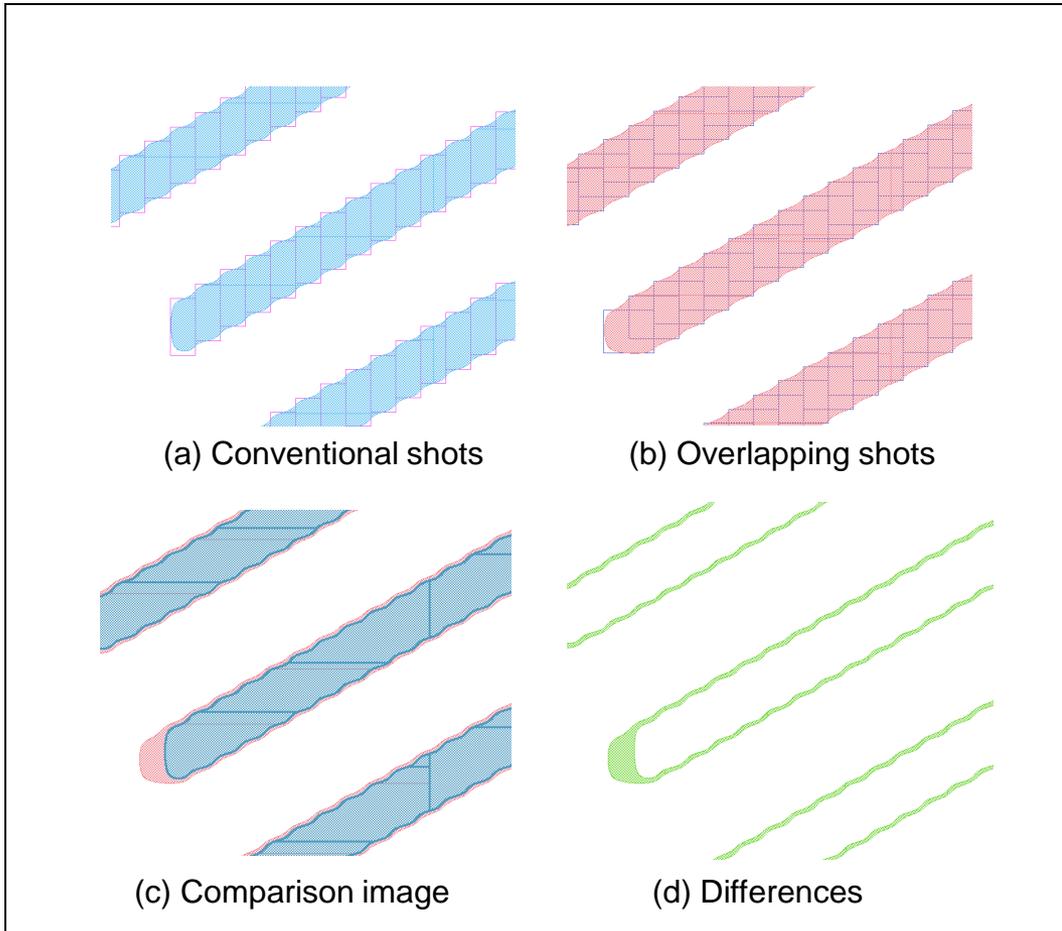
Results

Type	Processing time	Power consumption
GPU	1,258 s	100 Wh
CPU	6,130 s	340 Wh



-5 times faster

- 3.4 times smaller power consumption



Result images

Results

Type	Processing time	Power consumption
GPU	1,258 s	130 Wh
CPU	5,460 s	300 Wh



-4 times faster

- 2.3 times smaller
power consumption

5. Conclusion

- We have proposed novel verification methods for MB-MDP shots using GPU computing techniques.
- We have achieved X4-5 faster speed and X2.3-3.4 smaller power consumption compared to conventional CPU methods.
- Future plans
 - System development for actual mask production
 - Further improvement of basic algorithm for faster calculation

This study is a joint work of AIST (Advanced Industrial Science and Technology) and Hitachi High-Tech Science Corporation.