3D Image Deconvolution on Heterogeneous Compute Stations

Seeing clearly using personal super computers

Luke Domanski, Pascal Vallotton, John Taylor – CSIRO Mathematical and Information Sciences

Introduction

Deconvolution is an important operation in many areas of science, including astronomy[7], microscopy[3], and medical imaging. It reduces the effects of blurring introduced during image capture, revealing objects and details that may not have been visible in the raw image. Iterative methods for deconvolution can provide superior results compared to simpler approaches, but can be computationally expensive. We present our results on the use of GPUs and heterogeneous computing to increase the performance of iterative deconvolution algorithms.

Imaging Model

The image of a point light source (point spread function) in an image device characterizes the blurring discussed above. Each point of an observed object produces one of these point spread functions (PSF), hence, the resulting image is a convolution of the true object with the PSF (Fig 1). In some systems the PSF can change shape based on position and is called *spatially variant*.

The Richardson-Lucy (RL) algorithm[2,4] can be used to find the true object in the presence of Poisson image noise by iteratively re-applying the imaging model to an improving estimate. At the heart of the RL algorithm are two convolutions that can be implemented efficiently with Fast Fourier Transforms (FFT). It has been shown that using NVIDIA CUFFT library and Tesla hardware can improve the performance of the RL algorithm[5] compared to optimized CPU FFT libraries, such as Intel's MKL.

Parallel Deconvolution

Algorithm

Deconvolution can be performed in sections[6] by breaking the image into overlapping tiles and processing each tile independently (Fig 1). This approach can be used to decompose the image for parallel processing and to account for a spatially variant PSF by using a different PSFs for each tile[1,8].

The work performed on each tile is independent and does not rely on the processor being used. We implement two RL functions: one for the CPU using MKL, and one for the GPU using CUFFT 2.2 and custom CUDA kernels. We use MPI to run these functions in parallel across multiple processors.

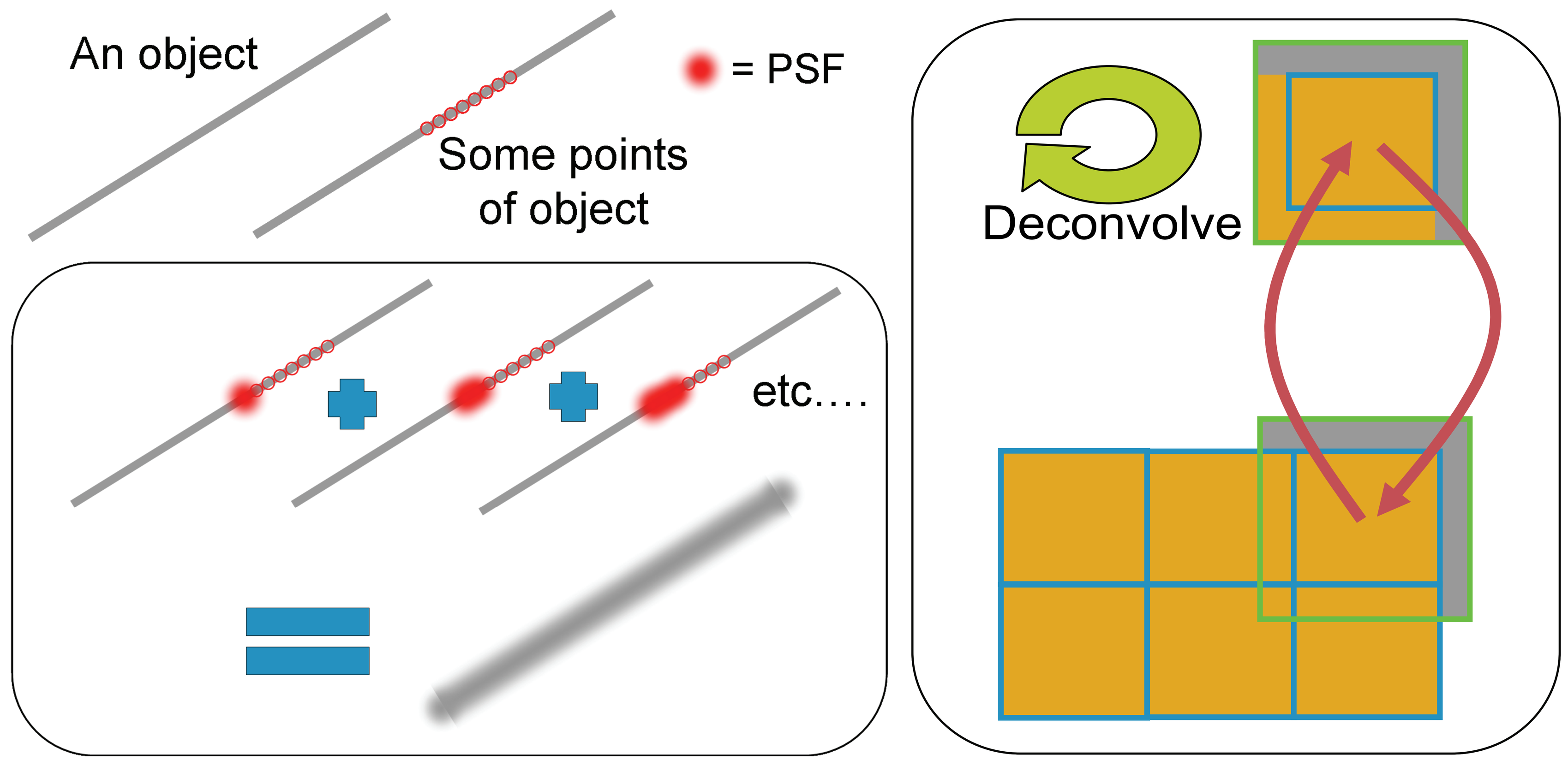

Figure 1: Left) Blurring in imaging systems. Right) Sectioned deconvolution for parallel processing and spatially variant point spread functions

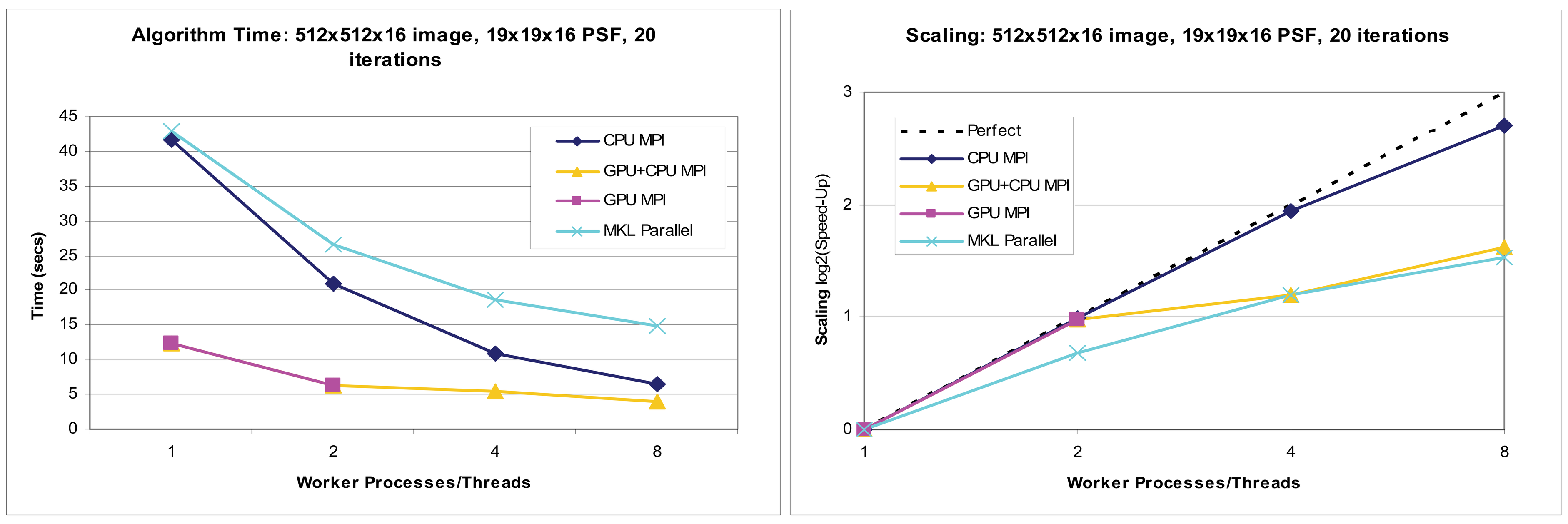

Figure 2: Left) Algorithm time as a function of number of workers (for GPU+CPU 2 GPUs are utilised first, then the CPUs). Right) Scaling of algorithm with the number of workers.

Compute resource management

Mechanisms are used to ensure workers take full advantage of available compute resources, and avoid resource contention. Counting semaphores are used to lock available resources. Workers try to claim the processors for the most efficient RL modules first, progressing to slower implementations if resources are unavailable. This approach ensures that resources are used in a combination that leads to the best performance. The semaphore approach allows migration to a compute cluster in the future. Because semaphores are publically accessible locks at the scope of a workstation, workers need not be aware of what workstation they run on or what other nodes are doing. They only try to get the best processor available based on the semaphores on their workstation. Note that a GPU worker requires both a GPU and CPU.

Results

Tests were performed on a Dual-socket Quad-core Xeon workstation equipped with two Tesla C1060s (Fig 2). While MKL provides parallel FFT using threads, it was found that using *N* single thread MPI workers was advantageous as opposed to 1 worker running *N* MKL threads. Both the GPU-only and CPU-only algorithms performed well and scaled with the number of physical processors, with 2 GPU workers achieving similar performance to 8 CPU workers.

In the GPU+CPU algorithm the GPU workers dominate processing, so introducing CPU workers increases performance only slightly over the dual GPU results and narrowly beats 8 CPU-only workers. This results in poor scaling of the GPU+CPU algorithm with the number of workers on a single workstation. Running the application on a cluster of workstations leveraging a high-speed network could see performance advantages of the GPU+CPU algorithm increase, as more dominant workers will be introduced. Better scaling is expected from the algorithm with the number of workstations as opposed to workers.

Conclusion

We have shown results for our implementation of parallel deconvolution on a heterogeneous workstation consisting of both GPUs and CPUs. Our approach takes full advantage of all computing resources available, and provides better performance than using the available GPUs or CPUs alone.

For spatially variant deconvolution this translates to a speed-up of over 10x compared to a single thread CPU algorithm using Intel's MKL library. Future work aims to test the application on our compute cluster consisting of 200 GPUs, 800 CPU cores, and Inifiband interconnect.

References

1. Boden et. al. Massively parallel spatially variant maximum-likelihood restoration of Hubble Space Telescope imagery, *J. Opt. Soc. Am. A, OSA,* 1996, *13*, 1537-1545 / **2.** Lucy, L. B. An iterative technique for the rectification of observed distributions, *The Astronomical Journal,* 1974, *79*, 745-754 / **3.** McNally et. al. Three-Dimensional Imaging by Deconvolution, Microscopy *Methods: A Companion to Methods in Enzymology,* 1999, *19*, 373-385(13) / **4.** Richardson, W. H. Bayesian-Based Iterative Method of Image Restoration, *J. Opt. Soc. Am., OSA,* 1972, *62*, 55 / **5.** Domanski et. al. Two and Three-Dimensional Image Deconvolution on Graphics Hardware. In *Proceedings of the 18th World IMACS/MODSIM Congress,*1010-1016, 13-17 2009 / **6.** Trussell, H. & Hunt, B. Sectioned methods for image restoration, *IEEE Transactions on Acoustics, Speech, and Signal Processing,* 1978, *26*, 157-164 / **7.** The Restoration of HST Images and Spectra – II, Hanisch, R. J. & White, R. L. eds. November 18-19, 1993 / **8.** Shearer et. al. Parallel image restoration with spatially variant point spread function: description and first clinical results, *Medical Imaging 2001: Image Processing, SPIE,* 2001, *4322*, 787-79

Further information

contact: Luke Domanski
phone: +61 (2) 9325 3427
email: Luke.Domanski@csiro.au
web: www.cmis.csiro.au

www.csiro.au

CSIRO