

Online real-time reconstruction of adaptive TSENSE with commodity CPU / GPU hardware

S. Roujol¹, B. D. De Senneville¹, E. Vahala², T. S. Sørensen³, C. Moonen¹, and M. Ries¹

¹UMR 5231, CNRS/Université Bordeaux 2, Laboratory for Molecular and Functional Imaging, Bordeaux, France, ²Philips Medical Systems, Vantaa, Finland, ³University of Aarhus, Department of Computer Science and Institute of Clinical Medicine, Aarhus, Denmark

Purpose/Introduction

Adaptive TSENSE [1] has been suggested as a robust parallel imaging method suitable for MR-guidance of interventional procedures. However, in practice, the reconstruction of adaptive TSENSE images obtained with large coil arrays leads to long reconstruction times and image latencies and thus hampers its use for applications such as MR-guided thermotherapy or cardiovascular catheterization. Here, we demonstrate a real-time reconstruction pipeline for adaptive TSENSE optimized for low image latencies and high frame-rates on affordable commodity PC hardware.

Material and Methods

Reconstruction pipeline: A multi-threaded reconstruction system has been implemented as shown in figure 1 with the aim of parallelizing data transport, a preparative reconstruction in read-out direction, the principal reconstruction in phase-encoding direction and the TSENSE reconstruction. This allows the data acquisition of the $n+1$ th image can be started as soon as the n th acquisition cycle is terminated. To further reduce the image latency, the TSENSE reconstruction, including a dynamic update of the coil sensitivities, is offloaded to a Graphics Processing Unit (GPU). As Hansen et al. have shown, the highly linear nature of the reconstruction steps required for SENSE are well-suited for GPU offloading [2]. The employed GPU can be seen as a massively parallel co-processor. Our implementation has been designed minimizing memory exchange between RAM and dynamic random access memory (DRAM) of the GPU. All data have been reorganized in memory so that all memory access from individual threads result in contiguous access by the memory controller. The two main time consuming tasks of the TSENSE reconstruction are the SENSE unfolding matrix recalculation (requiring a matrix inversion for each pixel using a LU-decomposition) and the temporal filtering.

Reconstructor Hardware: the reconstructor was a dual processor (INTEL 3.1 GHz Penryn, four cores) workstation with 8 GB of RAM and dual 1 Gb/s network interface cards. The GPU was a NVIDIA 8800GTX card with 768 MB of DRAM connected over a PCIe x16 link.

Reconstructor Software: for the data transport from the MR-acquisition system to the reconstructor and from the reconstructor to the viewing station(s) a real-time implementation of the common object request broker architecture (CORBA) known as The Ace Orb (TAO) [3] was used. The GPU implementation of the adaptive TSENSE method was realized using CUDA [4].

Results

As shown on figure 2, data transport varies between 17 ms (four-fold accelerated, four channels, ~135 kB per image) to 76 ms (two-fold accelerated, 16 channels, ~1 MB per image) depending on the data size. The reconstruction itself has a theoretical peak performance between 75 images/s (SENSE 4, 16-channels) to 330 images/s (SENSE 2, 4-channels). However, in practice the achievable peak data-throughput was found to be limited by the I/O subsystem of the acquisition system to ~2.1 MB/s, which corresponds to an image frame-rate of 20 images/s for a two-fold TSENSE accelerated data set (128x128, two-fold read-out oversampling, 16 receiver channels) with a overall image latency of 90 ms, or 40 images/s and an image latency of 60 ms for a four-fold acceleration.

Discussion

The presented work shows that the CPU/GPU reconstruction time for TSENSE is shorter than the MR-acquisition time even in cases when large coil arrays are used. In practice, this means that the speed for real-time imaging is limited by the boundaries imposed by the MR-sequence, such as sampling time and SNR considerations. The proposed reconstruction achieves image latencies from 20 to 90 ms for all coil configurations and acceleration factors, and is thus well-suited for manual feedback required for applications such as MR-guided surgical interventions and can be implemented on affordable commodity hardware.

References

- [1] Guttman M. et al [1995], Magn Reson Med, 34:814-823. [2] Hansen M. S. et al [2008], Magn Reson Med, ;59(3):463-8. [3] Schmidt D. C. et al, [1998] Computer Communications, 21:294-324. [4] CUDA Programming Guide, 2.0 ed 2008, pp. 1-107.

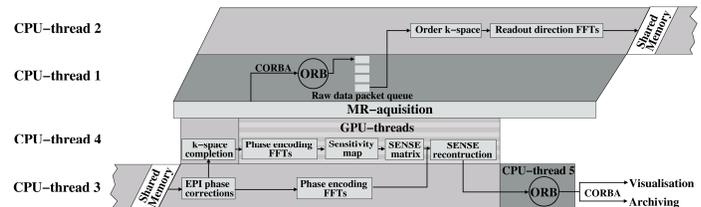


Figure 1. Overview of the thread architecture of the reconstruction pipeline. In order to achieve high throughput and short image latencies, as many independent data handling/reconstruction steps as possible are carried out in parallel: CPU-threads #1 and #2 handle/reconstruct k -space data from the dynamic image $n+1$, while CPU-threads #3 and #4 finalize simultaneously the reconstruction of dynamic image n . The time-consuming processing steps for the adaptive TSENSE reconstruction are offloaded to GPU-hardware which in itself uses up to 128 threads in parallel. A separate thread for dispatching the data to a visualization and an archiving system is used.

TSENSE factor	2				3				4			
MR-acquisition time (ms)	44											
TE (ms)	22											
Channel number	4	6	8	16	4	6	8	16	4	6	8	16
Data transport time (ms)	31	31	76	76	19	24	44	59	17	22	27	47
CPU only Reconstruction (ms)	26.2	33.7	44.1	84.4	26.3	35.8	45.6	88.8	26.9	35.5	47.4	90.3
CPU only Latency (ms)	57.2	64.7	120.1	160.4	45.3	59.8	89.6	147.8	43.9	57.5	74.4	137.3
CPU/GPU Reconstruction (ms)	3.0	4.2	5.7	11.3	3.1	4.5	5.9	11.5	3.5	5.0	6.4	12.2
CPU/GPU Latency (ms)	34.0	35.2	81.7	87.3	22.1	28.5	49.9	70.5	20.5	27.0	33.4	59.2

Figure 2. Reconstruction time and image latency for a single slice reconstruction of resolution 128x128 with TSENSE factor 2, 3, 4 for 4, 6, 8, 16 coil channels. With the proposed CPU/GPU implementation, total computation times are far below the TR on all tests demonstrating that real-time reconstruction is feasible. Since computation times for multi-slice acquisitions are almost linear with the slice number, only results measured with a single slice are reported.

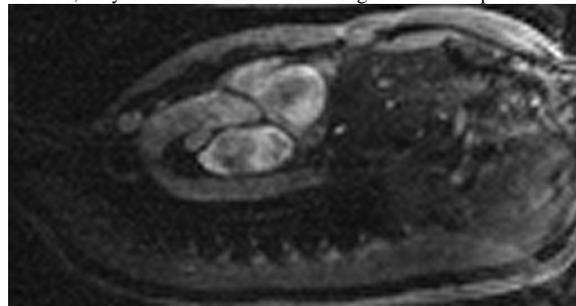


Figure 3. Real-time reconstructed MR-image using a combined CPU/GPU reconstruction (left-right: foot-head, up-down: anterior-posterior direction): Right ventricular outflow tract and the aorta, with a TSENSE acceleration factor of 4.