EFFICIENT ALGORITHM AND VLSI ARCHITECTURES FOR COMPRESSED SENSING SIGNAL RECOVERY

A thesis submitted
in partial fulfilment for the degree of

Doctor of Philosophy

By

Thomas James Thomas
SC16D019

Department of Avionics
Indian Institute of Space Science and Technology
Thiruvananthapuram, India

December 2022

Abstract

Compressed Sensing (CS) empowers the signal processing landscape to acquire sub-Nyquist measurements of real-world signals by exploiting their underlying sparse nature in suitable domains. Sparse recovery algorithms are integral to CS as they facilitate the reconstruction of higher dimensional signals from such compressed measurements. The recovery speed and hardware complexity has been a major focus of research in the design of such sparse recovery algorithms, with the computational effort for successful signal recovery remaining high, even for moderate sized problems. Dedicated hardware-driven sparse recovery has been pursued in the literature to offer swifter and cost-effective solutions for signal recovery compared to software implementations, but are restricted by the slow convergence or hardware complexity of the underlying algorithm. To solve this problem, the research work described in this thesis focuses on the development of a novel reconstruction algorithm to swiftly and accurately decode CS measurements. Further, this work presents the design of two distinct hardware architectures related to the goals of processing speed improvement and resource minimization respectively by employing suitable hardware-friendly optimizations on the proposed algorithm.

The preliminary work is concentrated on the design of a novel sparsity independent CS reconstruction algorithm that employs parallel index selection and regularization to curtail the number of iterations required to reconstruct the signal and thereby enhance the reconstruction speed at the algorithmic level. A restricted isometry property (RIP) based analysis is provided to guarantee the exact recovery of k-sparse signals. A rigorous experimental evaluation of the proposed algorithm is carried out in high-dimensional, sparsity blind and noisy scenarios. The proposed algorithm is found to achieve a significant speed-up with respect to the state-of-art while maintaining the reconstruction accuracy.

Subsequently, our next focus of research has been to exploit the achieved algorithmic speed-up on dedicated hardware. Complexity reduction techniques are adopted to opti-

mize the proposed algorithm with the specific goal of maximizing the reconstruction speed. The reformulated algorithm incorporates a cheaper regularization strategy and a modified Gram-Schmidt (MGS) based incremental QR decomposition (QRD) approach to augment the support set. The proposed design incorporates an iterative QRD architecture with feedback circuitry to exploit the parallelism of the multi-atom support augmentation step at increased hardware costs. Additionally, a fast inverse square root block circumvents the need for parallel divider blocks giving considerable hardware and latency savings. The design reuses the iterative QRD block to implement the interdependent computations of the algorithm by sophisticated scheduling techniques. The proposed design is found to accelerate sparse reconstructions respective to the state-of-art for similar problem sizes by a factor of 2 at least. Based on the achieved reconstruction speed for 36-sparse vectors, the proposed hardware is able to attain a processing throughput of roughly 13100 Vectors/second or equivalently 13.4 million samples per second (MSPS). The proposed implementation does not require prior knowledge of k for termination or signal estimation such that the reconstruction process can remain unmonitored for signals of varying k. It is observed that for lower sparsity levels, reconstruction speed is greatly improved without significant change in RSNR.

In order to significantly curtail hardware consumption by allowing a relaxation in reconstruction time, we subsequently focus on developing an improved version of the proposed algorithm where a gradient descent inspired least-mean-squares (LMS) approach is proposed to replace the complex least squares (LS). The proposed architecture bypasses matrix transpose requirement for parallel access to each row of the measurement sub-matrix by interleaving row write operations with the pipelined LMS process. The proposed implementation is found to curtail DSP and logic slices by $3\times$ and $2\times$ respective to the implementation in the state-of-art possessing comparable reconstruction time. An alternate row based LMS update scheme is proposed to reduce the latency cost and leads to a 28% improvement in the recovery time with a 12% degradation in the RSNR with respect to the

full LMS update.