Preface

This book was originally written to serve as the material for an advanced one-semester (fourteen 2-hour lectures) graduate course for engineering students at the Technion, Israel. It is originally based on a review paper that I wrote together with Alfred M. Bruckstein and David L. Donoho, which appeared in SIAM Review on February 2009. The content of this paper has been massively modified and extended to become an appropriate material for an advanced graduate course. In this book I introduce the topic of sparse and redundant representations, show the questions posed in this field, and the answers they get, present the flavor of the research in this arena, the people involved, and more. It is my desire to expose the readers of this book (and the students in the course) to the beauty of this field, and the potential it encompasses for various disciplines.

So, what is it all about? Mathematicians active in this field (and there are many of those) would mention the deep relation between this emerging field and harmonic analysis, approximation theory, and wavelets, in their answer. However, this is not my answer. As an engineer, my interest is more in the practical sides of this field, and thus my entire motivation for studying sparse and redundant representations comes from the applications they serve in signal and image processing. This is not to say that I find little interest in the theoretical sides of the activity in sparse and redundant representations. However, these theoretical results should be considered in the context of the wider picture.

As I see it, this field is all about one specific mathematical model for signal sources. Modeling sources is key in signal and image processing. Armed with a proper model, one can handle various tasks such as denoising, restoration, separation, interpolation and extrapolation, compression, sampling, analysis and synthesis, detection, recognition, and more. Indeed, a careful study of the signal and image processing literature reveals that there is an evolution of such models and their use in applications. This course is about one such model, which I call \textit{Sparse-Land} for brevity. This specific model is intriguing and fascinating because of the beauty of its theoretical foundations, the superior performance it leads to in various applications, its universality and flexibility in serving various data sources, and its unified view, which makes all the above signal processing tasks clear and simple.
At the heart of this model lies a simple linear system of equations, the kind of which seems long studied in linear algebra. A full-rank matrix $A \in \mathbb{R}^{n \times m}$ with $n < m$ generates an underdetermined system of linear equations $Ax = b$ having infinitely many solutions. We are interested in seeking its sparsest solution, i.e., the one with the fewest nonzero entries. Can such a solution ever be unique? If so, when? How can such a solution be found in reasonable time? It is hard to believe, but these questions and more of the like form the core engine to this model and the wide field that has grown around it. Positive and constructive answers to these questions in recent years expose a wide variety of surprising phenomena, and those set the stage for *Sparse-Land* to serve in practice. Clearly, research in this field calls for an intensive use of knowledge from linear algebra, optimization, scientific computing, and more.

This field is relatively young. Early signs of its core ideas appeared in a pioneering work by Stephane Mallat and Zhifeng Zhang in 1993, with the introduction of the concept of dictionaries, replacing the more traditional and critically-sampled wavelet transform. Their work put forward some of the core ideas that later became central in this field, such as a greedy pursuit technique that approximates a sparse solution to an underdetermined linear system of equations, characterization of dictionaries by their coherence measure, and more.

A second key contribution in 1995 was by Scott Shaobing Chen, David Donoho, and Michael Saunders, who introduced another pursuit technique that uses the $\ell_1$-norm for evaluating sparsity. Surprisingly, they have shown that the quest for the sparsest solution could be tackled as a convex programming task, often leading to the proper solution.

With these two contributions, the stage was set for a deeper analysis of these algorithms and their deployment in applications. A crucial step towards this goal was made in 2001, with the publication of the work by Donoho and Huo. In their daring paper, Donoho and Huo defined and (partly) answered what later became a key question in this field: Can one guarantee the success of a pursuit technique? Under what conditions? This line of analysis later became the skeleton of this field, providing the necessary theoretical backbone for the *Sparse-Land* model. It is amazing to see how fast and how vast this area of research has grown in recent years, with hundreds of interested researchers, various workshops, sessions, and conferences, and an exponentially growing number of papers.

The activity in this field is spread over all major universities and research organizations around the world, with leading scientists from various disciplines. As this field of research lies at the intersection between signal processing and applied mathematics, active in this field are mathematicians interested in approximation theory, applied mathematicians interested in harmonic analysis, statisticians, and engineers from various fields (computer science, electrical engineering, geophysics, and more).

A little bit about myself: I started my activity in this area after a short visit by David Donoho to the Technion, where he gave a talk describing the results of the above-mentioned paper. I am ashamed to admit that I did not attend the talk! Nevertheless, my good friend and mentor, Freddy Bruckstein, was there to appreciate
the opportunity. Freddy’s sixth sense for important research directions is already famous at the Technion, and did not fail him this time too. Freddy insisted that we work on improving the results that Donoho presented, and few months later we had our first result in this field. This result led to a joint work with Donoho during a post-doc period at Stanford, and I have found myself deeply involved in research on these topics ever since.

My origins as an engineer, and the fact that I worked with a leading mathematician, give me a unique perspective which I bring to this book. I present a coherent, well-structured, and flowing story that includes some of the theoretical foundations to sparse and redundant representations, numerical tools and algorithm for actual use, and applications in signal and image processing that benefit from these. I should stress that I do not pretend to give a well-balanced view of this entire field, and instead, I am giving my own point-of-view. In particular, I am not covering all the accumulated knowledge in this book because: (i) it is impossible to mention everything; (ii) not all details are important to the story I have to tell; (iii) new results in this field are emerging on a daily basis, making it impossible to give a completely updated view; and (iv) I have to admit that I do not understand all the results in this field.

Compressed-Sensing is a recent branch that separated from sparse and redundant representations, becoming a center of interest of its own. Exploiting sparse representation of signals, their sampling can be made far more effective compared to the classical Nyquist-Shannon sampling. In a recent work that emerged in 2006 by Emmanuel Candes, Justin Romberg, Terence Tao, David Donoho, and others that followed, the theory and practice of this field were beautifully formed, sweeping many researchers and practitioners in excitement. The impact this field has is immense, strengthened by the warm hug by information-theorists, leading mathematicians, and others. So popular has this field become that many confuse it with being the complete story of sparse representation modeling. In this book I discuss the branch of activity on Compressed-Sensing very briefly, and mostly so as to tie it to the more general results known in sparse representation theory. I believe that the accumulated knowledge on compressed-sensing could easily fill a separate book.

The CS Department, the Technion, Israel
September 2009

Michael Elad