

Chapter 6

Conclusions and recommendations

In this study, I investigated the use of compressive sensing for various applications, and as a viable alternative sampling theorem. In the compressive sampling of signals, there are similar workflows that can be followed separately for image-based and audio-based signals. Both follow a common workflow in the encoding stage, but differ slightly in the decoding stage. Depending on the size of the signal, both types can be processed in one go, but more often than not, the signal has to be processed in several manageable slices. Image-based signals can be sliced into patches with no overlap, while audio-based signals can be divided into subbands with significant overlap in order to suppress reconstruction artifacts.

Due to the sparsity and incoherence requirements of CS, the ideal starting point is by using partial DCT matrices encoded by randomly distributed samples. As for the random distribution, uniform distribution achieves the lowest reconstruction error, while triangular distribution achieves more consistent

results throughout a wide range of compression ratios. For this reason, as well as for simplicity in its generation, uniform distribution was used throughout this study. Comparison of three common reconstruction algorithms—namely, LASSO, OMP, and SLO—show that LASSO strikes the balance between run time and reconstruction quality. Thus, this algorithm was used in most cases, except where otherwise stated.

CS is most easily visualized using visual signals. Entire images can be compressively sampled at once if they are small enough. For larger images, they need first be sliced into patches, and each patch can be compressively sampled separately, then stitched back together after reconstruction. Additionally, simultaneous compression and encryption can be achieved using the same algorithms by encoding the sensing matrices with deterministic chaotic functions instead of random distributions. Using a logistic map, the CS compression-encryption system was shown to achieve correlation coefficients of as low as 0.02, and no higher than 0.42. Key sensitivity analysis showed that the system was sensitive to initial value perturbations on the order of 10^{-15} , making it robust against brute force attacks. The structural similarity index (SSIM) was determined to be a perceptually accurate metric for this type of signal, and a value of 0.8 and above is deemed acceptable. The experiments performed in the study achieved reconstruction errors of no lower than 0.82 SSIM.

Denser signals, such as audio, benefit largely from CS. Undersampling artifacts are more easily apparent in these signals. If the signal is simple enough, i.e., those with relatively unchanging frequencies, then it can be processed all at once. The same procedure with image-based CS can be followed, but in addition, the slicing the signal into subbands (analogous to image patches) require

some overlap between adjacent subbands in order to suppress undersampling artifacts. Experiments with recordings of single guitar notes showed that even while undersampling just enough to capture the base frequency of a note, up to sixth-order harmonics can still be recovered using CS. For larger and complex signals such as speech, it may be necessary to transform the signal to a higher dimension where it can be represented more sparsely, i.e. from the temporal domain to the modulation domain. A speech recording of a complete sentence was compressively sampled successfully with a 40% compression ratio, and 75% overlap between adjacent subbands. Error space mapping in terms of PESQ and the more widely used average segmental SNR (SNR_{seg}) shows that the reconstruction quality of a signal in terms of PESQ can solely be described by the compression ratio, while in terms of SNR_{seg} , by the number of subbands used to divide the signal. The perceptual evaluation of speech quality (PESQ) was determined to be a perceptually accurate metric for this type of signal, and a value of 3.0 and above is deemed acceptable. The experiments performed in the study achieved reconstruction errors of no lower than 2.5, which is quite lower than the acceptable threshold.

Future studies may look into further standardizing the CS workflow by introducing overlap between adjacent image patches—similar to audio—to improve reconstruction quality. For audio-based signals, one may explore varying parameters such as compression ratio, percent overlap, and number of subbands to also improve reconstruction quality. Finally, all these can be combined to compressively sample signals which may contain any combination of audio and images, such as full HD color videos and hyperspectral images.