Research on Image and Video Coding Algorithms for Compressive Imaging

PEETAKUL, Jirayu

(開始ページ / Start Page)
1
(終了ページ / End Page)
138
(発行年 / Year)
2022-03-24
(学位授与番号 / Degree Number)
32675甲第541号
(学位授与年月日 / Date of Granted)
2022-03-24
(学位名 / Degree Name)
博士(工学)
(学位授与機関 / Degree Grantor)
法政大学 (Hosei University)
(URL)
https://doi.org/10.15002/00025234

Research on Image and Video Coding Algorithms

for Compressive Imaging

JIRAYU PEETAKUL

Doctoral Dissertation Reviewed by Hosei University

Research on Image and Video Coding Algorithms

for Compressive Imaging

JIRAYU PEETAKUL

Abstract

The traditional camera based on a hundred-year-old sampling theorem developed by Whittaker–Nyquist–Kotelnikov–Shannon has resulted in a massive problem of redundant data in image and video applications, which oversamples signal twice higher than information rate. It necessitates the use of complex lossy coding algorithms to reduce redundancy. However, the most recent coding algorithms are going far beyond coding efficiency; for instance, improving coding performance by 20% would cost roughly 50% more complexity and resources, which is still a significant issue today. A new camera architecture based on block-based compressed sensing (CS) has recently gained popularity because it offers lower sampling costs and produces far less amount of raw data. Meanwhile, it is sufficient to represent the original content accurately. CS is based on the Johnson–Lindenstrauss lemma, which deals with low-distortion embedding of points from high to low dimensions via random projection, resulting in a compressed vector. It theoretically eliminates the need for coding algorithm. However, the recent studies found that raw data from the CS camera is still redundant in the form of linear combination, potentially necessitating additional coding to reduce redundancy. This thesis presents a new sensing matrix that outperforms existing sensing matrices in data acquisition performance and speed at low sampling rates while dramatically improving image quality. Furthermore, a newly developed data structure of a block-based CS camera called data cube is introduced, making coding raw CS data easier. Simplified image and video coding algorithms for compressive imaging, both vector-based and data cube-based, are introduced in software and hardware, including intra-prediction, inter-prediction with quantization, and entropy coding to improve bitrate reduction performance.

Acknowledgements

This work would not have been possible without the financial support of Japan Science and Technology Agency (JST) or graduate school of science and engineering at Hosei Unversity. I am especially grateful to Professor Jinjia Zhou and my colleague at institute of Integrated Science and Technology (IIST) at Hosei University, who have been supportive of my research goals and actively worked to provide me with protected academic time to pursue those goals. I am grateful to all of those with whom I have had the pleasure to work during this and other related projects. Each of the members of my dissertation committee has provided me extensive personal and professional guidance and taught me a great deal about both scientific research and life in general. Nobody has been more important to me in the pursuit of this project than members of my family. I would like to thank my parents, whose love and guidance are with me in whatever I do. They are the ultimate role models. Finally, I would like to thank myself for thinking and acting differently throughout my career in order to achieve academic goals and provide a unique perspective on emerging science and technology, which you did incredibly well in terms of people starting to adopt your ideas and algorithms on far land across the open-source ocean. While others are succumbing to some study trends, you remain devoted to your chosen field. You have demonstrated that you are a true man and inventor in this field. Your ideas may be too advanced for this decade based on research trends; do not be surprised if your efforts have had no immediate impact on academia and industry; the CS camera is now a very niche application and research area, with only a few laboratories or companies owning the device. In this regard, you are doing well. Remember that great inventions always getting notice after graduation; 5-10 years later, you could be awarded the field medal in applied mathematics for your efforts, I know you want one.

Contents

List of Figures

List of Tables

1

Introduction

1.1 Background

The Nyquist-Shannon sampling theorem (also known as the $Whittaker - Nyquist - Kotelnikov - Shannon$ theorem) is a widely used method for converting an analog signal to a digital signal ([81]). It is a signal processing theorem that serves as a fundamental link between continuous-time and discrete-time signals. It creates a sufficient condition for a sample rate that allows a discrete sequence of samples to capture all of the information from a continuous-time signal with finite bandwidth. This sampling theorem introduces the concept of a sample rate sufficient for perfect fidelity for the class of functions that are band-limited to a given bandwidth so that no information is lost during the sampling process. It expresses the necessary sample rate in terms of bandwidth for the class of functions. The theorem also leads to a formula for reconstructing the original continuous-time function from samples perfectly.

Considering popular consumer devices that strongly embraced the Nyquist-Shannon sampling theorem, complementary metal-oxide-semiconductor (CMOS) image sensors, which primarily used a sample-and-hold (S/H) circuit for analog to digital conversion (A/D). Commonly, CMOS image sensors capture light using a high-precision bit of ADC (often 12 to 16 bits for a single ADC unit), quantize to a manageable size of 8 bits, and output as raw pixel data. In general, we will not directly measure band-limited signals at preferred bits because it may miss some information, but will instead perform oversampling to band-limited signals and then quantize to a manageable size to ensure that we do not miss any information. Furthermore, losing some signal fractional coefficients due to quantization does not affect signal reconstruction ability.

Nowadays, as resolution requirements increase, throughput can significantly increase by a large number of pixels, such as 4K, 8K, and 16K. In this case, image and video coding algorithms are required to reduce data size. We would prefer lossless compression, which reduces file size by 15 to 25% of the original size while preserving all meaningful information, However, the file size cannot be reduced further due to limited redundant sequences. This causes issues with real-world memory space. Recalling Fourier coefficients, an image can be represented in another transform basis of sparsity. All data compression algorithms that rely on sparsity to reduce spatial and temporal redundancy, whereby a signal is compressed more efficiently in terms of the sparse vector of coefficients using a generic transform basis such as discrete cosine transform (DCT)—a lossy variant of Fourier transforms or discrete wavelet transform (DWT). This is the fundamental concept underlying all data compression methods, including JPEG/JPEG2000, JPEG-XR, advanced video coding (AVC), high-efficiency video coding (HEVC), and versatile video coding (VVC). The complexity of algorithms, on the other hand, can be tenfold increased with each generation and screen resolution. Furthermore, there is a significant problem with power consumption. Because of their inter-frame and group of pictures (GOPs)-based schemes, they require a lot of memory. As a result, because they are highly complex and consume a lot of power on their own, they would never be used to reduce power consumption and interfaces within an electronic device. Furthermore, they are typically not executable in terms of software implementation, but rather necessitate the purchase of more expensive hardware, which will become more expensive with each generation as future resolution increases. This means that we are currently focusing on statistical numerical compression performance via bit-per-pixels (bpps) while trading complexity. For example, some algorithms compress 5% better than previous generations while increasing algorithm complexity by 20%. In this case, the latency between encoder and decoder is also high because a large amount of memory is required as a frame buffer.

JPEG-XS, later known as TICO CODEC, is a new standard compression algorithm from the JPEG group for 4K/8K/16K IP-based transmission where they implemented lightweight, simple encoder and decoder, low FPGA/ASIC footprint, low latency, highly parallelizable in CPU and GPU, primarily focusing on devices with limited power and resources such as wireless cameras and handheld live streaming. In video production, the concept of this device is less expensive than AVC, HEVC, and VVC, which trade lossless quality for a lower compression ratio (resulting in higher bpps). This CODEC is now widely used in virtual reality (VR), augmented reality (AR), mixed reality (MR), automotive, ADAS systems, and machine vision (cameras, frame grabbers, extenders), wireless video systems, and drone applications where the bandwidth is set as high as possible to unlimited.

As a result, we are all wondering if the data compression algorithm has reached an efficiency limit that is generally not worth the additional effort ([87]), while the rest of the world is focusing on unlimited bandwidth to solve the bottleneck in transmission.

1.2 Motivation

Now, it is reasonable to assume that the root problems of high complexity of image and video coding algorithms were all based on an inefficient signal acquisition model. Regarding the problem challenges presented previously mentioned, image and video coding algorithms have reached their efficiency limit, and further implementation is only required to maintain an increased screen resolution.

The compressed sensing (CS) sampling theorem was recently developed to replace the Nyquist-Shannon sampling theorem in dozen applications, most notably wireless imaging for surveillance systems, particularly when operating continuously throughout the day and night. The CS theorem is based on an underdetermined linear system and random projection regarding the Johnson-Lindenstrauss Lemma (JL-Lemma), which produces lower-dimensional data (also known as compressed data) from high-dimensional data. This method, which does not require data compression in theory, has the potential to revolutionize image and video acquisition models by simplifying sensor architecture and reducing oversampled data. Block-based CS cameras have been proposed, which is groundbreaking in the CMOS image sensor industry, recently. Each work claims that CS solves problems such as power consumption, internal heat, frame rate, and readout rate ([20]) ([112]) ([50]) ([67]) ([53]) ([52]), which is known to be superior to traditional architecture.

When the mechanism on each iteration of the JL-Lemma random sampling process is closely examined, each low-dimensional sample still represents weak redundant information that is compressible. As a result, image and video coding algorithms are still required for the CS camera. However, unlike traditional image and video data structures, which are held in two coordinates of width and height, with a third optional coordinate to represent color-space. The CS data structure, on the other hand, operates in the form of a single column vector, rendering conventional data compression algorithms ineffective and necessitating the development of a new coding algorithm.

1.3 Contribution

There is concern that the lack of a coding algorithm for CS data structures will lead to bandwidth and data storage issues, both of which are limited when the CS theorem is widely used in many applications. Hence, a novel coding algorithm for CS data is becoming increasingly important in order to ensure that the new sampling technique serves the internet of things (IoT) and centralizes the artificial intelligence (AI) era smoothly. The contribution of this study is to investigate unusual CS data structures and develop novel data compression algorithms for the CS camera in both software and hardware implementation to serve future technology.

The study is divided into two sections: novel sensing matrix for CS, which allows image and video visualization to be improved over the classical sensing matrices stated in the theorem, and convex optimization, which allows convex optimization to reach convergence faster because each data point is well placed in orderliness, and reduces the state in non-deterministic computation, and multiple compression algorithms, which operate in the following styles: vector-based (which is commonly found in state-of-the-art works), and newly proposed data cube concept (also known as cube) and its coding algorithm. First and foremost, the coding algorithm is composed of various components such as intra-prediction, inter-prediction, and moving detection. Each component is then broken down and explained in detail, as follows: Intra prediction employs spatial correlation between blocks within a single frame to generate residual data by extrapolating from previously coded blocks and subtracting them from the current block. Inter-prediction is the inverse of intra-prediction in that it uses temporal correlation between multiple blocks in successive frames to generate prediction candidates, which are then subtracted from the current block to generate residual. However, intra-prediction and inter-prediction require a large number of candidates to compute, which is a computational burden. The static or low-change block will be skipped by the moving detection algorithm. It can reduce intra-prediction and inter-prediction workload by handling only the blocks that frequently change. Furthermore, by compressing low-dimensional CS data, the proposed encoder will be simpler, less expensive, and more efficient than traditional image and video coding algorithms. To be clear where is this effort standing in the ocean of real-wold implementation realm, this section has included a comparison statistic graph as shown in Figure 1.1.

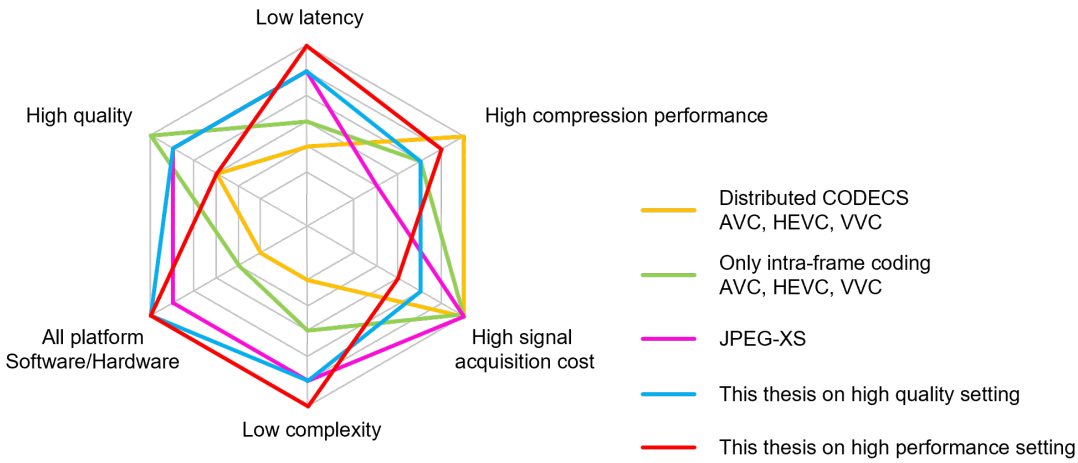

Figure 1.1: Statistic graph comparison of existing standard image and video coding algorithms, where this work is clearly stated in terms of lower signal acquisition cost, platform flexibility, low complexity, and low latency.

1.3.1 Structured Sensing Matrix By Scrambling Orthogonal Walsh Matrix for Compressed Sensing

First and foremost, a sensing matrix is important in CS because it limits the quality of acquired data, which directly affects image quality. Several sensing matrices have been proposed to determine the best basis sensing matrices. However, existing sensing matrices may fail in some sparse transforms used in image and video acquisition, such as the Fourier and Canonical transforms. The challenge is that novel sensing matrices must provide high image quality even when sampling with low SR close to the transform basis limited bound, whereas existing sensing matrices all fail to provide high image quality. An inverse problem and an undetermined linear system are formulated over the consistency of signal acquisition, according to the JL-Lemma, in which it should provide a close relationship and a close distance between each data point with a high probability. Arranging the number of zero-crossing in sensing matrices, for example, should provide a good relationship between each data point in the case of binary sensing matrices. However, as a result, it will lose randomness, making it unreconstructed under certain conditions.

In this work, a novel structured sensing matrix called Continuously ordering Walsh matrix (CoW) is proposed by scrambling orthogonal Walsh construction. Furthermore, the lowest bound condition of this novel sensing matrix is provided, which can be used to determine signal recoverability prior to the sample. According to spare optimization and the JL-Lemma, it enables CS camera to acquire signals more efficiently while reducing the number of observations and the computational overhead associated with sparse optimization. This sensing matrix allows for faster convergence due to the short distance of data clusters in the non-polynomial hard problem of non-deterministic computation when it comes to optimization. This sensing matrix transforms computation constraints into a nearly deterministic one. In another context, the problem of non-deterministic computation is simplified into nearly deterministic computation, where accepted and rejected branches are reduced, resulting in faster sparse recovery.

1.3.2 Vector-based coding algorithm for CS camera

1.3.2.1 A Measurement Coding System for Block-based Compressive Sensing Images by Using Pixel-Domain Features

This paper proposes a method for reducing redundancy in compressively sensed images from a CS camera via Natural ordered Hadamard (NoH). This work employs an intra-prediction algorithm inspired by the traditional image and video coding. Because of the differences between the CS and conventional data structures, it is unable to perform neighboring pixel-to-pixel coding; instead, it suggests an alternative intra-prediction method that performs neighboring vector-to-vector coding. To begin, sensing patterns, to the best of our knowledge, can be used to determine which pixel has been read out. As a result of these properties, it is possible to make a prediction candidate by utilizing single magnitude data from a neighboring block such as the left and upper sides of the current block. These single magnitudes will be multiplied by the sensing matrix to generate prediction candidates, which is a novel method in comparison to previous work. After obtaining several prediction candidates, the residual can be calculated by using the sum-of-absolute difference to calculate the minimum error between the current block and each of the prediction candidates (SAD). The residual will then be transferred to SQ to further reduce magnitude. Finally, Huffman coding will be used to compress the data for efficient transmission over communication channels.

1.3.2.2 Intra Prediction Based Measurement Coding Algorithm for Block-Based Compressive Sensing Images

This work is an extension of previous work aimed at improving intra-prediction performance by switching the sensing matrix from NoH to sequency ordered Walsh-Hadamard (SoWH), which, when combined with intra-prediction, provides better compression performance with higher image quality than previous work. Following that, the software algorithm was adapted into a hardware algorithm that was tested on a field-programmable gate array (FPGA). When compared to software implementation, performance has increased by a factor of ten. Surprisingly, CS data structures allow for lower circuit complexity than conventional data structures. Furthermore, when processing large image/video files, such as 4K, it can operate at a lower operating frequency while providing similar to higher throughput.

1.3.2.3 Temporal Redundancy Reduction in Compressive Video Sensing by using Moving Detection and Inter-Coding

Moving detection and inter-coding are proposed in this work to reduce temporal redundancy in compressive video sensing. To begin, moving detection is performed using coding area extraction with a local adaptive threshold to classify the measurement with an association of error distinction. However, false-positive detection may occur at random, increasing the transmission cost and increasing uncertainty. The adaptive quantization parameters are adjusted based on how frequently the area is detected to reduce transmission costs. The detected area is further compressed by encoding the difference between the current measurement and the best-matched measurement in neighboring frames.

1.3.2.4 Multiple Candidates Based Hybrid Hierarchical Search for Compacting Compressively Sensed Video

In general, finding the best match motion vector required a long search time when using inter prediction with moving estimation and moving compensation (MEMC). Coarse-grained search MEMC has been proposed as a way to cut search time in half. However, it also sacrificed accuracy and reduced bit-rates by half. In this paper, an efficient computational and higher accuracy MEMC with moveable search vectors is proposed. It provides the option to stop searching in a vector that potentially does not meet the criteria and redirects computational resources to another vector, implying that this algorithm performs a partially fine-grained search to some part of the area of interest compared to traditional MEMC and coarse-grained search MEMC. Furthermore, this work improved on previous work in terms of lower bit-rate while maintaining visual quality.

1.3.2.5 Measurement Coding Framework for High-Resolution Compressive Imaging

This paper proposed a completed measurement coding framework for the 4K CS camera. This framework consists of intra and inter-prediction, as well as a precise scalar quantization design. This work provides universal coding in which performance is not limited by measurement matrices, as commonly found in state-of-the-art works. By comparing to existing work, this work offers better coding performance with nearly no distortion targeting streaming content. It outperformed state-of-the-art works.

1.3.3 Cube-based coding algorithm for CS camera

1.3.3.1 Cube-based Video Coding Algorithm for Compressive imaging

First and foremost, a new perspective on vector data structures in the form of a data cube is proposed. In this case, the data cube comprises multiple downsampled images that can be obtained from an undetermined linear equation via multiple sensing matrix patterns. It enables us to create a more adaptable data compression algorithm for the CS camera. Because each layer in a data cube represents low-resolution image, for example, when sampling 4K frame (3840×2160 pixels) using a CS camera with a sub-block size of 16×16, we instead get 240 × 135 pixels, where the amount of sub-image is equal to the number of observations. In contrast to previous works that limit algorithm functionality to a few specific sensing matrices, this paper proposes a new compression algorithm for CS cameras to achieve universality for data sampled via a wide range distribution of sensing matrices. A normalized sub-image is created by averaging each layer in a cube data structure. Rather than performing intra and inter-prediction on each layer individually, an algorithm performs intra and inter-prediction on the generated sub-image, yielding the predicted template. The predicted template is then subtracted from each layer, resulting in faster coding performance and more reliable results, allowing residual data to be processed more effectively by quantization and entropy coding than in previous works.

1.4 Thesis Outline

This thesis carefully describes each necessary fundamental subject in this thesis to assist the reader in understanding the future data acquisition concept, which is vastly different from the conventional approach. Furthermore, it reflects the significant difference in data compression algorithms. First and foremost, the fundamentals of CS, which will revolutionize data acquisition methodology in general, are described. Furthermore, in section 2, an early work that propagated these technologies to become reality is summarized. Because the data acquisition model in this work differs greatly from conventional methods, a data structure in vector space is described in subsection 2.1, which helps the reader better understand the proposed algorithms throughout the thesis. The relationship between sparsity and compression geometries is then discussed, which is essential knowledge throughout this thesis. In subsection 2.2, compression geometries from a sparsity standpoint are provided, where the conventional approach samples a large number of coefficients but eventually keeps some fractional coefficients for better storage management. This concept will be explained in a casual manner using well-known transformation function such as Fourier. It will assist the reader in better understanding the theory behind any compression algorithm. The most important fundamental foundation, which is related to an undetermined system and dictionary creation, is discussed (also known as sensing matrices and measurement matrices) in subsection 2.3. This fundamental enables us to measure sparse signals in sufficient quantities to represent data in other space domains such as image and video without discarding any over sample coefficients. The following subsection 2.4 explains how to recover undetermined data and restore the original length of a sparse signal. It should be noted that, in general, only small minimization problems were considered, with only convex optimization achieving excellent results with no difficulty in implementation except for computation intensity. Iterative methods based on greedy or expectation–maximization algorithms are not considered in this work because they require prior information that is impossible to determine in real-world situations other than brute-force search, such as sparsity levels and transform basis. Using an iterative method, this has been identified as a flaw in some applications that are directly related to consumer imaging systems. Finally, the current technological transition of CMOS imaging architecture that uses CS with current representation and compression models is summarized, which helps the reader understand where the proposals stand.

2

Compressed Sensing

The inherent structure observed in natural data implies that the data can be represented sparsely in an appropriate coordinate system. In other words, if natural data is expressed on a well-chosen basis, only a few parameters are required to characterize the modes that are active and in what proportion. All data compression is based on sparsity, in which a signal is more efficiently represented in terms of a sparse vector of coefficients in a generic transform basis, such as Fourier or wavelet bases. Recent fundamental advances in mathematics have turned this paradigm upside down. Instead of collecting a high-dimensional measurement and then compressing, it is now possible to acquire compressed measurements and solve for the sparsest high-dimensional signal that is consistent with the measurements. This so-called CS is a valuable new perspective that is also relevant for complex systems in engineering, with the potential to revolutionize data acquisition and processing.

CS, which leverages the concept of transform coding, has emerged as a new framework for signal acquisition and sensor design that enables a potentially large reduction in sampling and computation costs for sensing signals with a sparse or compressible representation. While the Nyquist–Shannon sampling theorem states that a certain minimum number of samples are required to perfectly capture an arbitrary bandlimited signal, when the signal is sparse in a known basis, we can greatly reduce the number of measurements that must be stored. As a result, when sensing sparse signals, we may be able to perform better than classical results suggest. This is the fundamental idea behind CS: rather than first sampling at a high rate and then compressing the sampled data, we would like to find ways to directly sense the data in a compressed form —

i.e., at a lower sampling rate. Candès, Romberg, and Tao, as well as Donoho, established the field of CS by demonstrating that a finite-dimensional signal with a sparse or compressible representation can be recovered from a small set of linear, non-adaptive measurements (6) (16) (13). The development of these measurement schemes, as well as their application to practical data models and acquisition systems, are main challenges in the field of CS. To the rest, this chapter gently introduces the fundamental principles of sparsity as well as the mathematical theory that enables CS.

Early Works

While the concept of CS has recently gained popularity in the signal processing community, hints in this direction have been present since the eighteenth century. Prony's proposed an algorithm for estimating the parameters of a small number of complex exponential sampled in the presence of noise in 1795 (it has been translated from French to English and stored in a modern database (85)) In the early 1900s, Carathéodory demonstrated that a positive linear combination of any k sinusoids is uniquely determined by its value at $t = 0$ and any other $2k$ points in time. When k is small and the range of possible frequencies is wide, this represents far fewer samples than the number of Nyquist-rate samples. This work was generalized in the 1990s by George, Gorodnitsky, and Rao, who investigated sparsity in biomagnetic imaging and other contexts (40) (38) (39) (77). Concurrently, Bresler, Feng, and Venkataramani proposed a sampling scheme for acquiring certain classes of signals consisting of k components with nonzero bandwidth (as opposed to pure sinusoids) under constraints on the possible spectral supports, though exact recovery was not guaranteed in general (11) (34) (107). In the early 2000s Blu, Marziliano, and Vetterli developed sampling methods for certain classes of parametric signals that are governed by only k parameters, showing that these signals can be sampled and recovered from just $2k$ samples (108).

A related problem is the recovery of a signal from a partial observation of its Fourier transform. Beurling proposed a method for extrapolating these observations to determine the entire Fourier transform. One can demonstrate that if the signal consists of a finite number of impulses, then Beurling's approach will correctly recover the entire Fourier transform (of this non-bandlimited signal) from any sufficiently large piece of its Fourier transform. His approach is to find the signal with the smallest ℓ_1 norm among all signals that agree with the acquired Fourier measurements — is remarkably similar to some of the algorithms used in CS.

Candès, Romberg, Tao, and Donoho's works demonstrated that a sparsely represented signal can be recovered exactly from a small set of linear, non-adaptive measurements. This finding implies that it may be possible to sense sparse signals with far fewer measurements, hence the term *compressed sensing*. However, there are three significant differences between CS and classical sampling. To begin, sampling theory is typically applied to infinite-length, continuous-time signals. CS, on the other hand, is a mathematical theory that focuses on measuring finite-dimensional vectors in $\mathbb{R}^n$. Second, rather than sampling the signal at specific points in time, CS systems typically acquire measurements as inner products of the signal and more general test functions. This is in keeping with modern sampling methods, which also acquire signals through more general linear measurements (103). Third, the two frameworks differ in the manner in which they deal with signal recovery, i.e., the problem of recovering the original signal from the compressive measurements. In the Nyquist–Shannon framework, signal recovery is achieved through sinc interpolation – a linear process that requires little computation and has a simple interpretation. Compressed sensing has already had a significant impact on several applications. One application is in medical imaging (59), where it has enabled speedups by a factor of seven in pediatric MRI while preserving diagnostic quality (105). Moreover, the broad applicability of this framework has inspired research that extends the CS framework by proposing practical implementations for numerous applications, including sub-Nyquist sampling systems (24) (100) (97) (101), compressive imaging architectures (30) (2) (79), and compressive sensor networks (29) (23) (45).

2.1 Vector Space

For much of its history, signal processing has concentrated on signals generated by physical systems. Many natural and man-made systems can be modeled as linear. As a result, it is natural to consider signal models that complement this type of linear structure. This concept has been incorporated into modern signal processing by modeling signals as vectors in an appropriate vector space. This captures the linear structure that we often seek, namely that when we add two signals together, we get a new, physically meaningful signal. Moreover, vector spaces allow us to apply intuitions and tools from geometry in $\mathbb{R}^3$, such as lengths, distances, and angles, to describe and compare signals of interest. This is useful even when our signals are in high-dimensional or infinite-dimensional spaces.

In the case of a discrete, finite domain, we can view our signals as vectors in an n-dimensional Euclidean space, denoted by $\mathbb{R}^n$. When dealing with vectors in $\mathbb{R}^n$, we will frequently dealing with make frequent use of the ℓ_p norms, which are defined for $p \in [1, \infty)$ as

$$\|x\|_p = \begin{cases} (\sum_{i=1}^{n} |x_i|^p)^{\frac{1}{p}}, & p = [1, \infty) \\ \begin{matrix} max\,|x_i| \\ i = 1, 2, ..., n \end{matrix}, & p = \infty \end{cases} \tag{2.1}$$

In Euclidean space we can also consider the standard inner product in $\mathbb{R}^n$, which we denote $\langle x, z \rangle = z^T x = \sum_{i=1}^{n} x_i z_i$. This inner product leads to the ℓ_2 norm: $\|x\|_2 = \sqrt{\langle x, x \rangle}$ In some cases, it is useful to extend the concept of ℓ_p norms to the case where $p < 1$. In this case, the "norm" defined in 2.1 fails to satisfy the triangle inequality, so it is a quasi-norm.

We will also make frequent use of the notation $\|x\|_0 := |supp(x)|$, where $supp(x) = \{i : x_i \neq 0\}$ denotes the support of x and $|supp(x)|$ denotes the cardinality of $supp(x)$. Note that $\|\cdot\|_0$ is not even a quasi-norm, but one can easily show that

$$\lim_{p \to 0} \|x\|_p^p = |supp(x)|\,, \tag{2.2}$$

justifying this choice of notation. The ℓ_p (quasi-) norms have notably different properties for different values of p. To illustrate this, in Figure 2.1 we show the unit sphere, i.e., $\left\{x : \|x\|_p = 1\right\}$ induced by each of these norms in $\mathbb{R}^2$. We typically use norms as a measure of the strength of a signal, or the size of an error. For example, suppose we are given a signal $x \in \mathbb{R}^2$ and wish to approximate it using point in a one-dimensional affine space A. If we measure the approximation error using an ℓ_p norm, then our task is to find the $\hat{x} \in A$ that minimizes $\|x - \hat{x}\|_p$. The choice of p will have a significant effect on the properties of the resulting approximation error.

An example is illustrated in Figure 2.2. To compute the closet point in A to x using each ℓ_p norm, we can imagine growing an ℓ_p sphere centered on x until it intersects with A. This will be the point $\hat{x} \in A$ that is close to x in the corresponding ℓ_p norm. We observe that larger p tends to spread out the error that is more unevenly distributed

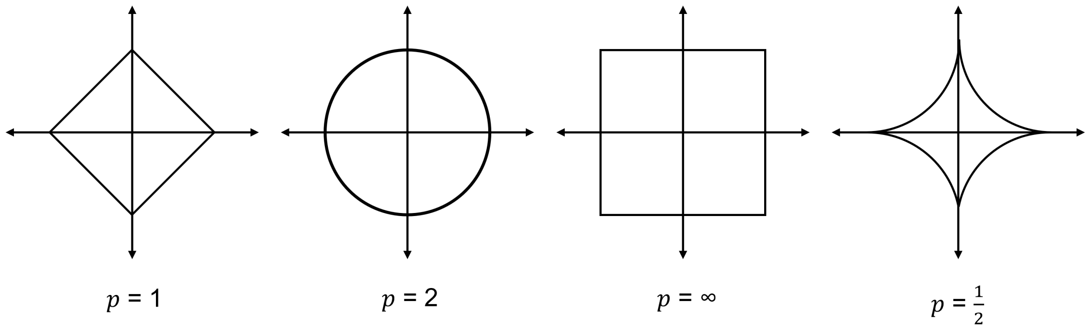

Figure 2.1: Unit spheres in $\mathbb{R}^2$ for the ell_p norms with $p = 1, 2, \infty$, and for the ℓ_p quasinorm with $p = \frac{1}{2}$

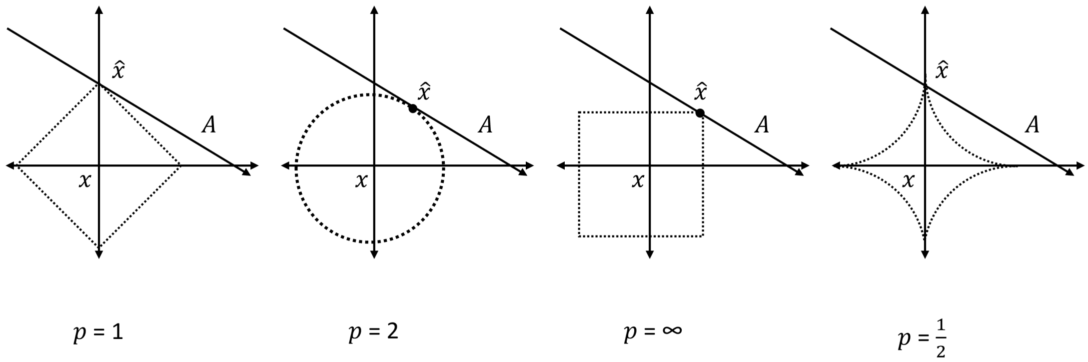

Figure 2.2: Best approximation of a point in $\mathbb{R}^2$ by a one-dimensional subspace using the ℓ_p norms for $p = 1, 2, \infty$, and the ℓ_p quasi-norm with $p = \frac{1}{2}$.

and tends to be sparse. This intuition generalizes to higher dimensions and playing an important role in the development of CS theory.

2.2 Sparsity and Compression Geometries

Most natural signals, such as images and audio, are highly compressible. This compressibility means that when the signal is written on an appropriate basis only a few modes are active, thus reducing the number of values that must be stored for an accurate representation. Said another way, a compressible signal $x \in \mathbb{R}^n$ may be written as a sparse vector $s \in \mathbb{R}^n$ (containing mostly zeros) in a transform basis $\Psi \in \mathbb{R}^{n \times n}$. It can be represented mathematically as follows:

$$x = \Psi s \tag{2.3}$$

In particular, the vector s is called K-sparse in Ψ if it contains exactly K nonzero elements. Sparse signals are frequently well approximated by a linear combination of only a few elements from a known basis or dictionary. When this representation is exact, the signal is said to be sparse. If the basis Ψ is generic, such as the Fourier or wavelet basis as shown in Figure 2.3, then only a few active terms in s are required to reconstruct the original signal x, reducing the amount of data required to store or transmit the signal.

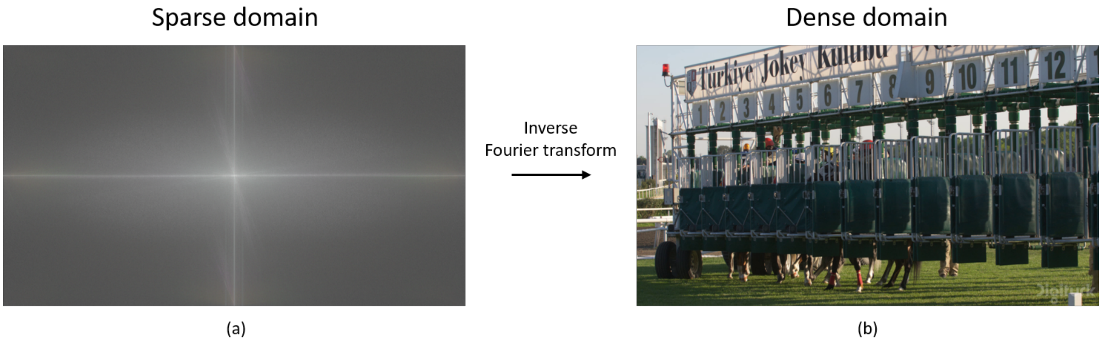

Figure 2.3: Illustration of a sparse signal captured on a Fourier basis and transformed to an image (in the dense domain) using the inverse Fourier transform.

Furthermore, high-dimensional signals often contain little information when compared to their ambient dimension. Talking mathematically, generally, we would say that signal x is k-sparse when it has a most k non-zeroes, i.e., $\|x\|_0 \leq k$. We let

$$\sum_k = \{x : \|x\|_0 \leq k\} \tag{2.4}$$

denote the set of all k-sparse signals. In most cases, we will be dealing with signals that are not sparse in and of themselves, but do admit a sparse representation in some basis Φ. In this case, we will still refer to x as being k-sparse, with the understanding that we can express x as $x = \Phi c$ where $\|c\|_0 \leq k$. In signal processing and approximation theory, sparsity has long been used for compression and other tasks (25) (66) (91) and denoising (26), and in statistics and learning theory as a method for avoiding overfitting (104). Sparsity also figures prominently in the theory of statistical estimation and model selection (89) (92), in the study of the human visual system (68), and has been exploited heavily in image processing tasks, since the multiscale wavelet transform (60) provides nearly sparse representations for natural images.

Consider image compression and denoising problems as a traditional application of sparse models. It is important to note that image compressibility is related to the overwhelming dimensionality of image space. Figure 2.4 presents a basic paradigm of data compression via fast Fourier coefficients by keeping only 5% of coefficients, where we can only notice a change in file size but feel the same in overall image details. So

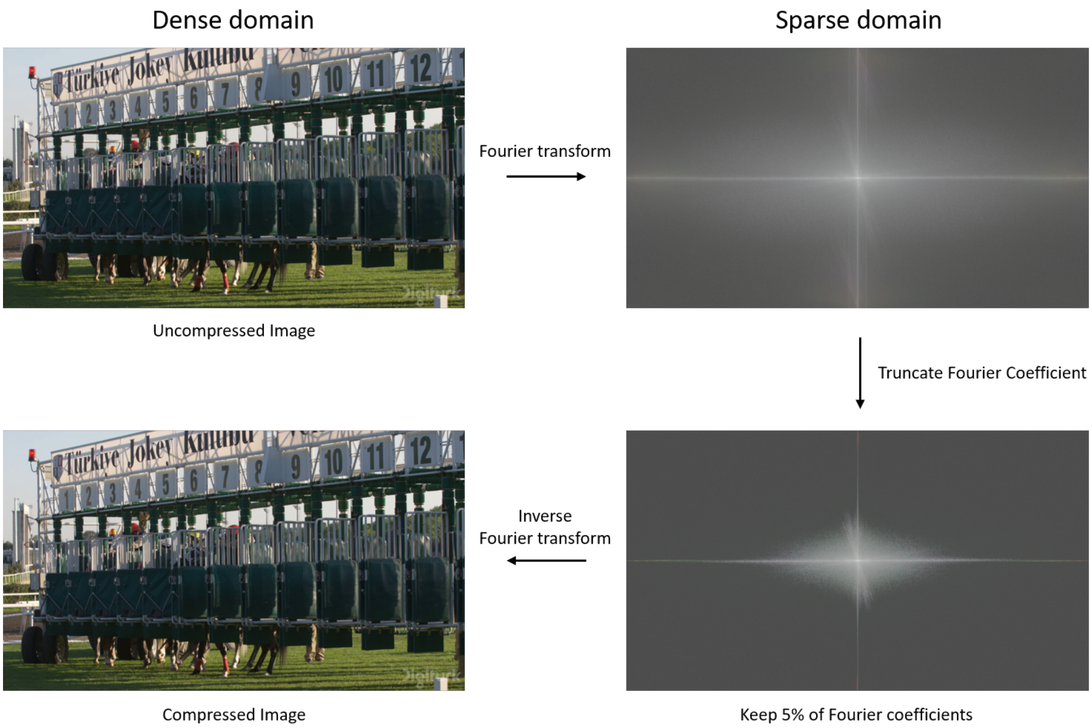

Figure 2.4: Illustration of conventional compression scheme via fast Fourier transform

that, when we know in advance that drawing 5% of Fourier coefficients from Fourier space is enough to represent an image, there is no need to acquire 100% of Fourier coefficients in the first place and then discard 95% of them. Can we pick those 5% signals in the first place? Then we do not need to truncate them out.

2.3 Undetermined System and Sensing matrices

If there are fewer equations than unknowns in a system of linear equations or a system of polynomial equations, the system is said to be underdetermined, in contrast to an overdetermined system, where there are more equations than unknowns. The concept of constraint counting can be used to clarify the terminology. Each unknown can be thought of as a degree of freedom that is available. Each equation introduced into the system can be thought of as a constraint that limits one degree of freedom. As a result, when the number of equations and the number of free variables are equal, the critical case (between overdetermined and underdetermined) occurs. There is a corresponding constraint that removes a degree of freedom for every variable that gives a degree of freedom. In contrast, the underdetermined case occurs when the system is underconstrained—that is, when the unknowns outnumber the equations.

An underdetermined linear system has either no solution or infinitely many solutions, as shown in the following example:

$$x + y + z = 1 \tag{2.5}$$

$$x + y + z = 0 \tag{2.6}$$

is an underdetermined system with no solution; any equation system with no solution is said to be inconsistent. On the other hand, the system

$$x + y + z = 1 \tag{2.7}$$

$$x + y + 2z = 3 \tag{2.8}$$

is consistent and has an infinite number of solutions, for example, $(x,\, y,\, z) = (1,\, 2,\, 2)$, $(2,\, 3,\, 2)$, and $(3,\, 4,\, 2)$ All of these solutions can be characterized by subtracting the first equation from the second, which demonstrates that all solutions obey $z = 2$; using this in either equation demonstrates that any value of y is possible, with $x = 1y$. The Rouché–Capelli theorem states that any system of linear equations (underdetermined or not) is inconsistent if the rank of the augmented matrix is greater than the rank of the coefficient matrix. If the ranks of these two matrices are equal, the system must have at least one solution; because in an underdetermined system, this rank must be less than the number of unknowns, there are indeed an infinite number of solutions, with the general solution having k free parameters, where k is the difference between the number of variables and the rank. There are algorithms for determining whether an underdetermined system has solutions and, if so, for expressing all solutions as linear functions of k variables, where the simplest method is Gaussian elimination.

This system is the foundation of CS, and it requires hints to recover sparse signals successfully. These hints, referred to as dictionaries, enable the linear combination of spare signals to meet existing solutions. In terms of CS foundation, sensing matrices should be composed of non-uniform distributions such as Gaussian or Bernoulli and uniform distributions such as Discrete Cosine transform or Hadamard transform, all of which must satisfy a specific linear algebra condition known as the restricted isometry property (RIP). It is used to characterize matrices that are nearly orthonormal, at least when operating on sparse vectors. The concept was introduced by Emmanuel Candès and Terence Tao (14) and is used to prove many theorems in the field of CS. Although there are no known large matrices with bounded restricted isometry constants (computing these constants is strongly non-deterministic polynomial also known as NP-hard (93) and difficult to approximate), many random sensing matrices have been shown to remain bounded. It has been demonstrated, in particular, that random Gaussian, Bernoulli, and partial Fourier matrices satisfy the RIP with a nearly linear number of measurements in the sparsity level with exponentially high probability (111).

It defines RIP to sensing matrices given by: let Φ be an $m \times n$ matrix and let $1 \leq s \leq n$ be an integer. Suppose that there exists a constant $\delta_s \in (0,1)$ such that, for every $m \times s$ submatrix Φ_s of Φ and for every s-dimensional vector y.

$$(1 - \delta_s) \|y\|_2^2 \leq \|\Phi_s y\|_2^2 \leq (1 + \delta_s) \|y\|_2^2 \tag{2.9}$$

Then, the matrix Φ is said to satisfy the s-restricted isometry property with restricted isometry constant δ_s. This is equivalent to

$$\|\Phi_s^* \Phi_s - I\|_{2\to 2} \leq \delta_s \tag{2.10}$$

where I is the identity matrix and $\|X\|_{2\to 2}$ is operator norm. Finally, this is equivalent to stating that all eigenvalues of $\Phi_s^* \Phi_s$ are in the interval $[1 - \delta_s, 1 + \delta_s]$. Further, an eigenvalue is used to describe for any matrix that satisfies RIP property with restricted isometric constant (RIC) which is used to define the infimum of all possible δ for a given $\Phi \in \mathbb{R}^{n\times m}$. It can be stated by the following condition:

$$inf\left[\delta : (1 - \delta)\right] \|y\|_2^2 \leq \|\Phi_s y\|_2^2 \leq (1 + \delta) \|y\|_2^2, \forall |s| \leq K, \forall y \in \mathbb{R}^{|s|} \tag{2.11}$$

which can be denoted shortly as δ_K Back to eigenvalues, the following condition holds:

$$1 - \delta_K \leq \lambda_{min}\left(\Phi_\tau^* \Phi_\tau\right) \leq \lambda_{max}\left(\Phi_\tau^* \Phi_\tau\right) \leq 1 + \delta_K \tag{2.12}$$

For Gaussian matrices, the tightest upper bound on the RIC can be computed. This can be accomplished by computing the exact probability that all Wishart matrices' eigenvalues fall within a given interval. Regarding the theoretical foundation above, CS acquisition model, in which high-dimensional signal length n can be mapped into lower-dimensional signal length m via random sampling method stated by JL-Lemma can be given by

$$y = \Phi x \tag{2.13}$$

where y is compressed signal with length of m and x is compressible signal with length of n. Clearer explain, it can represent with the following procedure:

$$\begin{array}{ccccc} y & & \Phi & & x \\ \begin{bmatrix} y_1 = (x_1\Phi_{1,1}) + (x_2\Phi_{1,2}) + \cdots + (x_n\Phi_{1,n}) \\ y_2 = (x_1\Phi_{2,1}) + (x_2\Phi_{2,2}) + \cdots + (x_n\Phi_{2,n}) \\ \vdots \\ y_m = (x_1\Phi_{m,1}) + (x_2\Phi_{m,2}) + \cdots + (x_n\Phi_{m,n}) \end{bmatrix} & = & \begin{bmatrix} \Phi_{1,1} & \Phi_{1,2} & \cdots & \Phi_{1,n} \\ \Phi_{2,1} & \Phi_{2,2} & \dots & \Phi_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_{m,1} & \Phi_{m,2} & \dots & \Phi_{m,n} \end{bmatrix} & \times & \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ \vdots \\ x_n \end{bmatrix} \end{array} \tag{2.14}$$

2.4 Sparse Optimization

In this subsection, spare optimization is discussed, which is related to curve fitting to obtain a curve, or a function, from given data, and how sparse optimization effectively work for CS problem ranging from a small problem to a large problem. There are key ideas to implement a sparse solver for CS as the following:

- The curve fitting is well-known formulated as an optimization problem to choose one solution among candidates.
- Regularization is used to avoid over-fitting.
- Sparse optimization is reduced to ℓ_1 optimization, which is convex and efficiently solved by numerical optimization.

2.4.1 Underdetermined System and Minimum ℓ_2-norm Solution

To begin with least squares and regularization with a simple example, consider the linear equation in CS given by

$$y = \Phi x \tag{2.15}$$

where $y \in \mathbb{R}^m$ is given vector, $\Phi in \mathbb{R}^{m \times n}$ is given matrix, and $x \in \mathbb{R}^n$ is an unknown vector, in which $m < n$, and Φ has full row rank, which can be described in equation as the following:

$$rank(\Phi) = m \tag{2.16}$$

Here, under this assumption, there exist infinitely many solutions. First, let find the smallest ℓ_1-norm solution among them. This is formulated as an optimization problem, given by

$$\underset{x \in \mathbb{R}^n}{minimize} \ \frac{1}{2} \|x\|_2^2 \ s.t. \ y = \Phi x \tag{2.17}$$

This problem is called as ℓ_2 optimization problem, and the solution is the minimum ℓ_2-norm solution. The method of Lagrange multipliers is used to solve this problem. First, the Lagrange function, or simply Lagrangian, of the optimization problem is defined by

$$L(x, \lambda) = \frac{1}{2} x^T x + \lambda^T (\Phi x - y) \tag{2.18}$$

where, the variable $\lambda \ in \mathbb{R}^m$ is called Lagrange multiplier. Then, we can obtain the optimal solution of 2.17 by finding the stationary point $(\hat{x}, \hat{\lambda})$ of the Lagrange function L. By differentiating L by the variable x,

$$\frac{\partial L}{\partial x} = \frac{\partial}{\partial x} \left(\frac{1}{2} x^T x + \lambda^T \Phi x \right) = x + \Phi^T \lambda \tag{2.19}$$

It follows that the stationary point $(\hat{x}, \hat{\lambda})$ satisfies

$$\hat{x} + \Phi^T \hat{\lambda} = 0 \tag{2.20}$$

Then differentiating L by $lambda$ gives

$$\frac{\partial L}{\partial \lambda} = \Phi x - y \tag{2.21}$$

and hence

$$\Phi \hat{x} - y = 0 \tag{2.22}$$

From this and 2.20, we have

$$-\Phi\Phi^T\hat{\lambda} = y \tag{2.23}$$

Since Φ has full row rank, the matrix $\Phi\Phi^T$ is non-singular and has its inverse. Therefore, from 2.23 we have

$$\hat{\lambda} = -(\Phi\Phi^T)^{-1}y \tag{2.24}$$

Assigning this to 2.23 gives the minimum ℓ_2-norm solution $\hat{x}$ as

$$\hat{x} = -(\Phi\Phi^T)^{-1}y \tag{2.25}$$

Finally, in summary, if we are given a full-row-rank matrix Φ and a vector y, we can compute the minimum ℓ_2-norm solution by the formula 2.25

2.4.2 Regression and Least Squares

Now, given 2 dimensional data $(t_1, y_1), (t_2, y_2), ..., (t_m, y_m)$. Next, consider a polynomial of order $n-1$,

$$y = f(t) = a_{n-1}t^{n-1} + a_{n-2}t^{n-2} + ... + a_1 t + a_0 \tag{2.26}$$

The curve fitting is to find coefficients $a_0, a_1, ..., a_{n-1}$ with which the polynomial curve has the best fit to the m-point data. For example, $t_1, t_2, ..., t_m$ are sampling instants, and $y_1, y_2, ..., y_m$ are temperature data from a sensor at a portion. From these data, we often want to know the curve behind the data. We call such data analysis the regression analysis or polynomial curve fitting. First, we consider the polynomial curve 2.26 goes through the data points $(t_1, y_1), (t_2, y_2), ..., (t_m, y_m)$, and hence we have m linear equations with unknowns $a_{n-1}, a_{n-2}, ..., a_1, a_0$:

$$\begin{aligned}
&a_{n-1}t_1^{n-1} + a_{n-2}t_1^{n-2} + ... + a_1 t_1 + a_0 = y_1, \\
&a_{n-1}t_2^{n-1} + a_{n-2}t_2^{n-2} + ... + a_1 t_2 + a_0 = y_2, \\
&... \\
&a_{n-1}t_m^{n-1} + a_{n-2}t_m^{n-2} + ... + a_1 t_m + a_0 = y_m,
\end{aligned} \tag{2.27}$$

Define a matrix

$$\Phi \triangleq \begin{bmatrix} t_1^{n-1} & t_1^{n-2} & \cdots & t_1 & 1 \\ t_2^{n-1} & t_2^{n-2} & \cdots & t_2 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ t_m^{n-1} & t_m^{n-2} & \cdots & t_m & 1 \end{bmatrix} \in \mathbb{R}^{m\times n} \tag{2.28}$$

and vectors

$$x \triangleq \begin{bmatrix} a_{n-1} \\ a_{n-2} \\ \vdots \\ a_1 \\ a_0 \end{bmatrix} \in \mathbb{R}^n, x \triangleq \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} \in \mathbb{R}^m \tag{2.29}$$

The system of linear equation can be represented in a matrix from: $\Phi x = y$. The matrix Φ is known as a Vandermonde matrix, and if $m = n$, then Φ is a square matrix and its determinant is given by

$$det(\Phi) = \prod_{1 \leq i < j \leq m} (t_i - t_j) = (t_1 - t_2)(t_1 - t_3)...(t_{m-1} - t_m) \tag{2.30}$$

It will be followed that if

$$t_i \neq t_j, \forall i, j \ s.t. \ i \neq j \tag{2.31}$$

then Φ is non-singular and has its inverse. Hence, the solution $\hat{x}$ of 2.27 is given by using Φ^{-1} as

$$\hat{x} = \Phi^{-1} y \tag{2.32}$$

2.4.3 ℓ_1-Minimization With Equality Constraints

Since the original polynomial is sparse and non-zero coefficients $\|x\|_0$ is assumed to be unknown. By borrowing the idea of the optimization priory mentioned, ℓ_0-norm as the cost function is used, and consider the following optimization problem:

$$\underset{x \in \mathbb{R}^n}{minimize} \ \|x\|_0 \ s.t. \ y = \Phi x \tag{2.33}$$

However, this is quite hard to solve using the exhaustive search method when the problem size is large. Therefore, the key idea of sparse optimization is to use ℓ_1-norm

$$\|x\|_1 = \sum_{i=1}^{n} |x_i| \tag{2.34}$$

instead of ℓ_0-norm. Consider the following optimization problem as relaxation of the ℓ_0 optimization 2.35:

$$\underset{x \in \mathbb{R}^n}{minimize} \ \|x\|_1 \ s.t. \ y = \Phi x \tag{2.35}$$

We call this optimization the ℓ_1 optimization. The method to obtain a sparse vector by the ℓ_1 optimization is known as the basis pursuit, which is recovery guarantee on any sensing matrices. However, using basis pursuit is a brute-force method that consumes a lot of computation resources and time. It will not meet the real-world requirement, which is primarily real-time unless solver optimization is performed. As a result, this thesis proposes using primal-dual interior-point optimization to optimize basis pursuit, which benefits optimization time the most.

2.4.4 ℓ_1-Minimization With Primal-Dual Interior-Point Method

Advances in interior-point methods for convex optimization over the past 20 years, led by the work in ([83]), have made large-scale solvers for the inverse problems feasible. In the past, Boyd and Vandenberghe outline a relatively simple primal-dual algorithm for linear programming, which we have followed very closely for the implementation of (P_1), (P_A), and (P_D). Their algorithm will be briefly reviewed here for completeness and to establish the notation. First, the standard form of linear program is given by

$$\underset{z}{minimize} \ \langle c_0, z\rangle \ s.t. \ A_0 z = b, \ f_i(z) \leq 0 \tag{2.36}$$

where the search vector $z \in \mathbb{R}^n$, $b \in \mathbb{R}^k$, A_0 is $k\ timesn$ matrix, and each of the f_i, $i = 1, ..., m$ is a linear functional:

$$f_i(z) = \langle c_i, z\rangle + d_i, \tag{2.37}$$

for some $c_i \in \mathbb{R}^n$, $d_i \in \mathbb{R}$. At the optimial point $\hat{z}$, there will exist dual vectors $\hat{v} \in \mathbb{R}^k$, $\hat{\lambda} \in \mathbb{R}^m$, $\hat{\lambda} \geq 0$ such that the Karush-Kuhn-Tucker (KKT) conditions are satisfied:

$$(KKT) \ \ c_0 + A_0^T \hat{v} + \sum_i \hat{\lambda}_i c_i = 0 \tag{2.38}$$

$$\hat{\lambda}_i f_i(\hat{z}) = 0, i = 1, ..., m, \tag{2.39}$$

$$A_0 \hat{z} = b \tag{2.40}$$

$$fi(\hat{z}) \leq 0, i = 1, ..., m \tag{2.41}$$

In a nutshell, the primal-dual algorithm finds the optimal $\hat{z}$ (along with optimal dual vectors $\hat{v}$ and $\hat{\lambda}$) by solving this system of nonlinear equations. The solution procedure is the classical Newton method: at an interior point (z^k, v^k, λ^k) (by which it mean $f_i(z^k < 0, \lambda^k > 0))$, the system is linearized and solved. However, the step to new point $(z^{k+1}, v^{k+1}, \lambda^{k+1})$ must be modified so that it will remain in the interior. In practice, we alway relax the complementary slackness condition $\lambda_i f_i = 0$ to

$$\lambda_i^k f_i(z^k) = \frac{-1}{\tau^k} \tag{2.42}$$

where we judiciously increase the parameter τ^k as we progress through the Newton iterations. This biases the solution of the linearized equations towards the interior, allowing a smooth, well defined central path from an interior point to the solution on the boundary. The primal, dual, and central residuals quantify how close a point $(z, , \lambda)$ is to satisfying (KKT) with 2.42 in place of the slackness condition:

$$\tau_{dual} = c_0 + A_0^T v + \sum_i \lambda_i ci \tag{2.43}$$

$$\tau_{cent} = -\Lambda f - (\frac{1}{\tau})1 \tag{2.44}$$

$$\tau_{primal} = A_0 z - b \tag{2.45}$$

where Λ is a diagonal matrix with $(\Lambda_{ii}) = \lambda_i$, and $f = (f_i(z)...f_m(z))^T$. From a point (z, v, λ), we want to find a step $(\Delta z, \Delta v, \Delta \lambda)$ such that

$$\tau_r(z + \Delta z, v + \Delta v, \lambda + \Delta \lambda) = 0 \tag{2.46}$$

by linearizing 2.46 with Taylor expansion around (z, v, λ) to be

$$\tau_r(z + \Delta z, v + \Delta v, \lambda + \Delta \lambda) \approx \tau_\tau(z, v, \lambda) + J_{\tau_\tau}(z, v\lambda) \begin{pmatrix} \Delta z \\ \Delta v \\ \Delta \lambda \end{pmatrix} \tag{2.47}$$

where $J_{\tau_\tau(z,v\lambda)}$ is the Jacobian of τ_τ, so then we have the new system as the following:

$$\begin{pmatrix} 0 & A_0^T & C^\tau \\ -\Lambda C & 0 & -F \\ A_0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \Delta z \\ \Delta v \\ \Delta \lambda \end{pmatrix} = - \begin{pmatrix} c_0 + A_0^T v + \sum_i \lambda_i c_i \\ -\Lambda f - \frac{1}{\tau} 1 \\ A_0 z - b \end{pmatrix} \tag{2.48}$$

, where $m \times n$ matrix C has the c_i^T as rows, and F is diagonal with $(F)_{ii} = f_i(z)$. Luckily, it is possible to eliminate $\Delta \lambda$ term using:

$$\Delta \tau = -\Lambda F^{-1} C \Delta z - \lambda - (\frac{1}{\tau}) f^{-1} \tag{2.49}$$

leaving us with the new core system:

$$\begin{pmatrix} -C^{\tau}F^{-1}\Lambda C & A_0^T \\ A_0 & 0 \end{pmatrix} \begin{pmatrix} \Delta z \\ \Delta v \end{pmatrix} = \begin{pmatrix} -c_0 + (\frac{1}{\tau}C^{\tau})f^{-1} - A_0^T v \\ b - A_0 z \end{pmatrix} \tag{2.50}$$

With the $(\Delta z, \Delta v, \Delta\lambda)$ we have a step direction. To choose the step length $0 < s \leq 1$, we ask that it satisfy two criteria: First, $z + s\Delta z$ and $\lambda + s\Delta\lambda$ are in the interior. Second, the norm of the residuals has decreased sufficiently following the condition of $\|\tau_\tau(z + s\Delta z, v + s\Delta v, \lambda + s\Delta\lambda)\|_2 \leq (1 - \alpha s) \cdot \|\tau_\tau(z, v, \lambda)\|_2$, where α is a user-specified parameter. Since the f_i are linear functional. We can choose the maximum step size that keeps the problem in the interior by

$$I_f^+ = \{i : \langle c_i, \Delta z\rangle > 0\}, I_\lambda^- \{i : \Delta\lambda < 0\}, \tag{2.51}$$

and set

$$s_{max} = 0.99 \cdot min\left\{1, \left\{-f_i(z))/\langle c_i, \Delta z\rangle, i \in I_f^+\right\}, \left\{-\lambda_i/\Delta\lambda_i, i \in I_f^-\right\}\right\} \tag{2.52}$$

Then starting with $s = s_{max}$, we check if two items above is satisfied; if not, we set $s^{'} = \beta \cdot s$ and try again to get accurate recovery results. Furthermore, it was discovered during thesis writing that the alternating direction method of multipliers (ADMM) is an algorithm that solves convex optimization problems by breaking them down into smaller pieces, each of which is then easier to handle. It provided interesting optimization as well as the primal-dual interior-point method. However, ADMM was not used in the research and implementation of this thesis (9).

2.5 Imaging Architecture based on Compressed Sensing

The conventional image sensor will currently have a very high resolution. A 4K image, defined as 3840×2160 pixels, produces an 8.29 megapixel image, whereas an 8K image produces a 33.17 megapixel image. This means that 8.29 million and 33.17 million conversions are required to convert analog to digital. Furthermore, it generates internal heat, reducing the life of the image sensor. The work in (30) introduced the first trial to revolutionize image sensor by utilizing a single-pixel with a digital micromirror device (DMD) to achieve random sampling behavior, which is a good fit with 3D microscopy camera (70). However, there is a significant disadvantage in other applications where a single-pixel camera requires a long exposure time that is incompatible with consumer use.

Later, the major image sensor companies, such as Sony and Hamamatsu, introduced new methods to capture images using CS by still fabricating multiple pixels sensor, but partitioning an entire sensor into individual blocks, giving much sensing performance boost while also reducing exposure time when compared to signal-pixel method. Furthermore, redesigning an imaging system saves significant energy by reducing not only RAW data size but also transistor count per pixel, pixel size, fill factor, and simplifying analog-to-digital circuits as in (74) (33) (**?**) stated. Various data gathering techniques for these novel image sensor have been proposed such as voltage summation (21) and current summation (47). In which the most successful implementation, Oike (Sony) and El Gamal (Standford University) proposed programmable block-based CS camera with per-column $\Sigma\Delta$ analog-to-digital (ADC) converters to reduce sensing time and increase frame-rate up to 480 fps and 1920 fps when the sampling rate is equal to 1/4 and 1/16, respectively (67), which plan to be used this technique in future Sony image sensor. Figure 2.5 shows a comparison of image sensor architecture. In terms of resources and power consumption, the CS camera outperforms conventional image sensors. The recent applications for CS cameras range from scientific cameras to examine gas leakage to wireless surveillance cameras (49) (1) and optical-based on-board missions (44), where RAW measurements will only stream from image sensor nodes to receiver (80). It requires less storage and bandwidth consumption than conventional approach.

Surprisingly, because CS cameras are implemented using a block-based approach, it allows spatial and temporal relationships between adjacent blocks and previous frames to be discovered. The same pixel will be measured multiple times according to the JL-lemma, resulting in each RAW measurement element containing a comparable amount of information about the signal being acquired. As a result, CS data is said to be more compressible than before. This investigation has been confirmed by the most recent and influential article (41). As a result, there is an open question about how to reduce redundancy in CS data, particularly in video applications, which typically continuously acquire large amounts of data. A novel compression algorithm for CS data is urgently required to address future storage and bandwidth consumption issues.

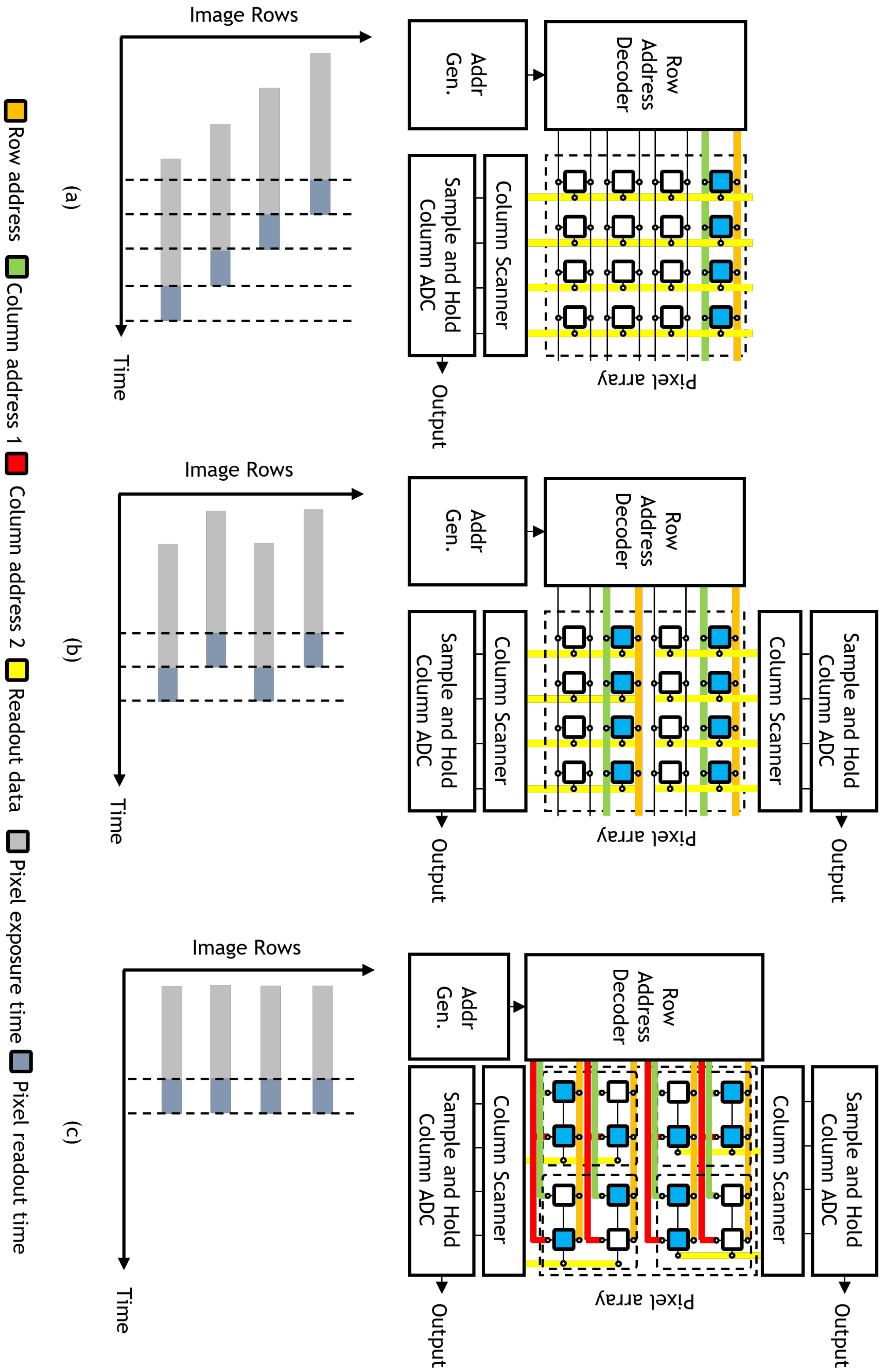

Figure 2.5: Illustration of conventional compression scheme via fast Fourier transform

3

Novel structured Sensing Matrix By Scrambling Orthogonal Walsh Matrix for Compressive Sensing

The theory of compressed sensing (CS) has the potential to revolutionize traditional signal acquisition methods. In CS, a sensing matrix has a remarkable meaning in that it can be used to determine the efficiency of signal acquisition. Several measurement matrices have been proposed to determine the best sensing matrix for CS. In this work, a novel structured sensing matrix called Continuously ordering Walsh matrix (CoW) is proposed by scrambling orthogonal Walsh construction. It enables the CS to acquire signals more efficiently while minimizing the number of observations and computational overhead associated with sparse optimization. We assess performance using reconstructed image quality when sampling at a rate of $5\% - 15\%$. According to simulation results, CoW provided significantly better image quality with significantly lower recovery error when compared to images sampled by existing structured measurement matrices.

3.1 Introduction

The band-limited model in Shannon's sampling theorem ([81]) is replaced by sparse model of compressed sensing (CS) ([27]). The model is based on solving underdetermined linear systems of equations: $y = \Phi x$, where $x \in \mathbb{R}^{n\times 1}$ is unknown solution, $\Phi \in$

$\mathbb{R}^{m \times n}$ is sensing matrix with $m < n$ where $\frac{m}{n}$ is sampling rate (SR), and $y \in \mathbb{R}^{m \times 1}$ is a measurement vector (also known as compressed data), with assumption that a signal can be exploited to recover it from far fewer samples than Shannon's sampling theorem requires. By applying this scheme to sensor design, it also implies that sensing costs, power consumption, and the complexity of data compression algorithms can be drastically reduced.

CS-camera has recently successfully gained popularity due to their possibility to reduce read-out timing and power consumption in 4K and 8K resolutions. They mostly adopted a linear-feedback shift register (LFSR) circuit to generate the sensing matrices. However, there is widespread concern that uncontrollable sensing sequences from LFSR have a significant impact on image quality with fairly poor compression performance (96). As the result, it has been a major drawback. Structured sensing matrices based on orthogonal transforms have recently been adopted in CS-camera. They allowed for the control of image quality and compression performance which have been published recently (37) (55) (56) (106) (57). There are many kinds of structured sensing matrices available that aim to reduce the number of observations close to $5\% - 15\%$ of SR (3). First and foremost, Hadamard and Walsh have been extensively researched and successfully applied to CS-camera because it provides good sensing performance, fast reconstruction, and is hardware friendly. However, it provides fairly poor image quality when further reduce SR lower than 60% (31) (10) (122).

Several studies have been carried out to investigate the effect of Hadamard and Walsh projection order selection on image reconstruction quality by simply reordering orthogonal matrices. In (88) the Russian-doll ordering Hadamard was proposed, in which the projection patterns are sorted by increasing the number of zero-crossing components. Cake-cutting ordering Hadamard was proposed in (113) by rearranging zero-crossing components over sensing matrix. It outperformed Russian-doll ordering Hadamard in image reconstruction quality. Later, Origami ordering Hadamard was proposed to improve the image quality of Cake-cutting ordering Hadamard (114). All of the above sensing matrices were tested on the CS camera, which provided significantly better image quality and compression performance than the LSFR. None of these works, however, explain why reordering such matrices can improve image quality at low SR.

3.2 Preliminaries

This thesis proposes a hypothesis that randomness is essential for recovering from the restricted isometry property (RIP) condition ([17]), whereas orderliness of zero-crossing from the smallest to the largest number of zero-crossings is important for image quality. If any sensing matrices possess those properties, they should be able to provide high compression performance while preserving image quality when sampling at low SR. The theoretical foundation of random projection and sensing matrices in CS is related to Johnson-Lindenstrauss-Schechtman (JL-lemma), 1986 as the following: given any $0 < \varepsilon < 1$ with any integer of n, and $\Phi \in \mathbb{R}^{m \times n}$ be a random orthonormal matrix. Then, for any set x of n point in $\mathbb{R}^m$, the following the following inequality about pairwise distance between any two data points a_i and a_j in x holds true with high probability as the following:

$$(1-\epsilon)\left\|a_i - a_j\right\|_2 \leq \left\|\Phi^T a_i - \Phi^T a_i\right\|_2 \leq (1+\epsilon)\left\|a_i - a_j\right\|_2 \tag{3.1}$$

This lemma obviously demonstrates that in Euclidean space, high-dimensional data can be randomly projected into lower-dimensional space while the pairwise distances between all data points are well preserved in close distance with a high probability, and the lower-dimensional signal can be fully recovered back to the original-dimensional without any problem. It has a lower sampling cost than the Nyquist-Shannon sampling theorem, which requires a high-dimensional signal all of the time. There are several different ways to construct the sensing matrices such as permutation, row-by-row reordering, and subsampled randomized. However, this work proposes reordering in conjunction with scrambling. It will not change the distance between each data point, but it will result in a more compact arrangement, which will allow the sparse solver to work faster and reach the convergent faster compares to existing sensing matrices.

According to the promising hypothesis, this can be explained by NP-hard, NP-completed, and NP problems regarding non-deterministic Turing machines, a theoretical model of computation whose governing rules specify more than one possible action when in some given situations. Unlike a deterministic Turing machine, which operates on a sequential state machine, the next state of a non-deterministic Turing machine is not entirely determined by its action and the current symbol it sees. In general, traditional ℓ_1-minimization via convex optimization is NP-hard problem where it cannot be solved in polynomial time constraint. Many convex optimization problems, such as linear programming (LP), second-order cone programming (SOCP), and semi-definite programming (SDP), can be solved in polynomial time using interior-point primal-dual methods, which reduce NP-hard to NP-completed and NP but never to P, in short, there is a limit. By carefully examining the non-deterministic Turing machine and the JL-lemma, it is possible to conclude that an effective prospect sensing matrix must reduce an irrelevant number of branches that eventually reach to reject as much as possible. Non-deterministic Turing machines, on the other hand, can be reduced to nearly deterministic Turing machines, resulting in faster convergent when solving ℓ_1-minimization problem.

3.3 New construction of structured sensing matrices

A new sensing matrix construction by scrambling orthogonal Walsh sensing matrix is proposed. First and foremost, generate Hadamard matrix $H\left(2^n\right) = H\left(2\right) \bigotimes H\left(2^{n-1}\right)$, where n is even number ($n = 2, 4, 6, ...$), then, apply Gray code permutation and bit-reverse to $H\left(2^n\right)$, resulting in Walsh matrix $W\left(2^n\right)$ as shown in Fig 3.1. Subsequently,

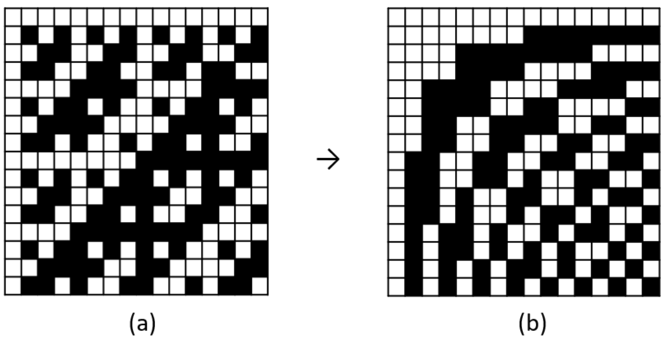

Figure 3.1: An example of matrices transformation via gray code permutations and bit-reversal where (a) Hadamard sensing matrix and (b) Walsh sensing matrix.

Walsh matrix size of $b \times b$ is partition into multiple masks size of $\sqrt{b} \times \sqrt{b}$ as shown in Fig 3.2. Rather than reordering rows by rows as is common in existing methods, this work

linewidthlinewidth

Figure 3.2: An example of partitioning of Walsh matrix size of $b \times b$ into multiple masks size of $\sqrt{b} \times \sqrt{b}$, where $b = 16$.

scrambles them by applying a zigzag pattern to partitioned masks. The procedures are summarized as follows:

Step 1. Partition $W(2^n)$ into multiple pieces, where the size of each piece equal to $\sqrt{b} \times \sqrt{b}$.

Step 2. Apply zigzag scramble to the multiple pieces of $W(2^n)$.

Step 3. Lastly, vectorize each piece from $\sqrt{b} \times \sqrt{b}$ and stack back into $b \times b$ matrix.

This method produces a new construction of Walsh matrix that gradually arranges measurement patterns from lowest number of zero-crossing to highest number of zero-crossing as shown in Fig 3.3. This order is named as Continuously ordering Walsh sensing matrix (CoW). For recovery guarantee according to foundation theorem, in

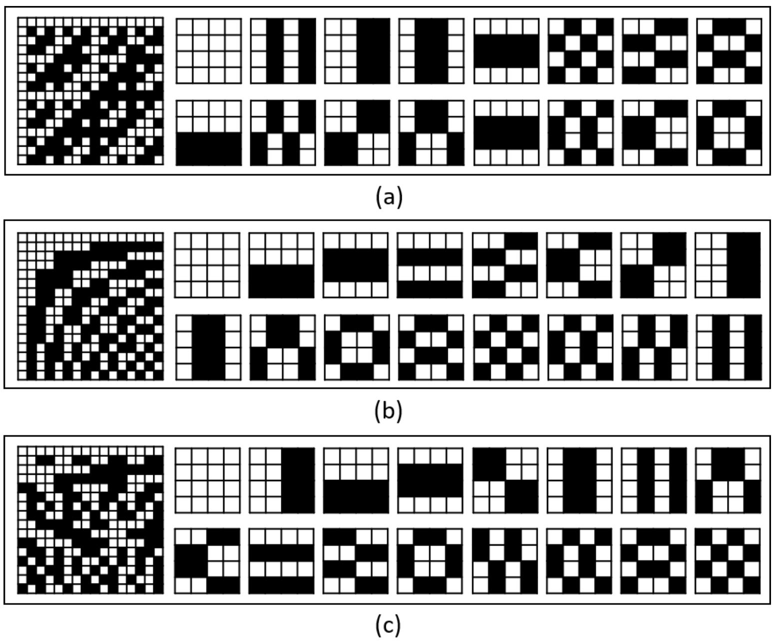

Figure 3.3: Structured measurement matrices and projection patterns comparison of (a) Hadamard, (b) Walsh, and (c) Continuously ordering Walsh (this work), where $b = 4$ and $n = m = 16$

general, CoW is not satisfied classical RIP condition given by

$$(1-\delta)\|x\|_2^2 \leq \|\Phi\|_2^2 \leq (1-\delta)\|x\|_2^2 \tag{3.2}$$

where $\delta \in \{1, 0\}$ It can be used to judge only non-uniform random such as Gaussian distribution. This condition is rigorous. In the actual situation, it is not easy to use this metric to judge if the sensing matrix is suitable for CS camera.

In ([5]) proposed alternative RIP conditions that use to guarantee recoverability of structured measurement matrices. Let $c_0(\varepsilon) > 0$ be positive constant, $\varepsilon \in (0, 1)$ if $\Phi_{m \times n}$ satisfy probability of an event of

$$\left| \|\Phi\|_2^2 - \|x\|_2^2 \geq \varepsilon \|x\|_2^2 \right| \tag{3.3}$$

but must be less than

$$2exp\left\lceil -nc_0\varepsilon \right\rceil, 0 < \varepsilon < 1 \tag{3.4}$$

then for a given δ, there exists a constant c_0, which enable the probability Φ of structured measurement matrices satisfy alternative RIP condition is above

$$1 - 2\left(12/\delta^k exp\left(-nc_0\left(\delta/2\right)\right)\right) \tag{3.5}$$

For instance, CoW is constructed by uniform distribution that produces property of $\{-1, 1\}$ with the probability of 1/2. Then,

$$c_0(\varepsilon) \tag{3.6}$$

for CoW will be equal to

$$\varepsilon^2/4 - \varepsilon^3/6 \tag{3.7}$$

Hence, CoW satisfy alternative RIP by

$$1 - 2\left(12/\delta^k exp\left(-n\varepsilon^2/4 - \varepsilon^3/6\left(\delta/2\right)\right)\right) \tag{3.8}$$

Commonly, CoW composes of two possible values consisting of $\{-1, 1\}$ which is possibly shifted to $\{0, 1\}$ for hardware compatibility. Moreover, each measurement mask can generate on-the-fly, which does not need to be stored as in some existing measurement matrices.

3.4 Experimental results

In this section, the proposed sensing matrix performance is demonstrated. The simulation results were delivered by a standard solver called ℓ_1-minimization via primal-dual interior-point method. To compare performance to existing measurement matrices used in CS cameras, such as linear-feedback shift register (LFSR), Hadamard, Walsh, and Parlay ordering Walsh (PoW). It should be noted that the Hadamard and Walsh ordering algorithm matrix generator codes were not available online at the time of writing the dissertation. They will only be made available upon reasonable request, to which every effort was made but no response was received. As shown in experimental results, CoW provided less recovery error via RMSE when sampled at target SR $\in \{5, 10, 15\}$ compared to existing measurement matrices on each test image in Table 3.1. However, at these target SR, statistical results are actually far beyond the point in which using in order to evaluate sensing matrices performance. Our human visual perception can use to judge better than statistical results. As a result, Figure 3.4 shows a comprehensive comparison of reconstructed images. It can be seen that CoW gave clearer image with fewer artifacts at high-frequency components. Further, it can significantly improve PSNR and SSIM averagely by 10% on each SR. Lastly, the evaluation results are provided by SSIM and RMSE comparison graphs with a different number of measurements per block as shown in Fig. 3.5, which CoW performed more extraordinary performance on both matrices than existing structured measurement matrices.

3.5 Summary

This work proposed a novel sensing matrix for CS camera by scrambling orthogonal Walsh matrix. When compared to existing measurement matrices, it allowed the CS-camera to capture images with fewer samples while providing better image quality with fewer artifacts.

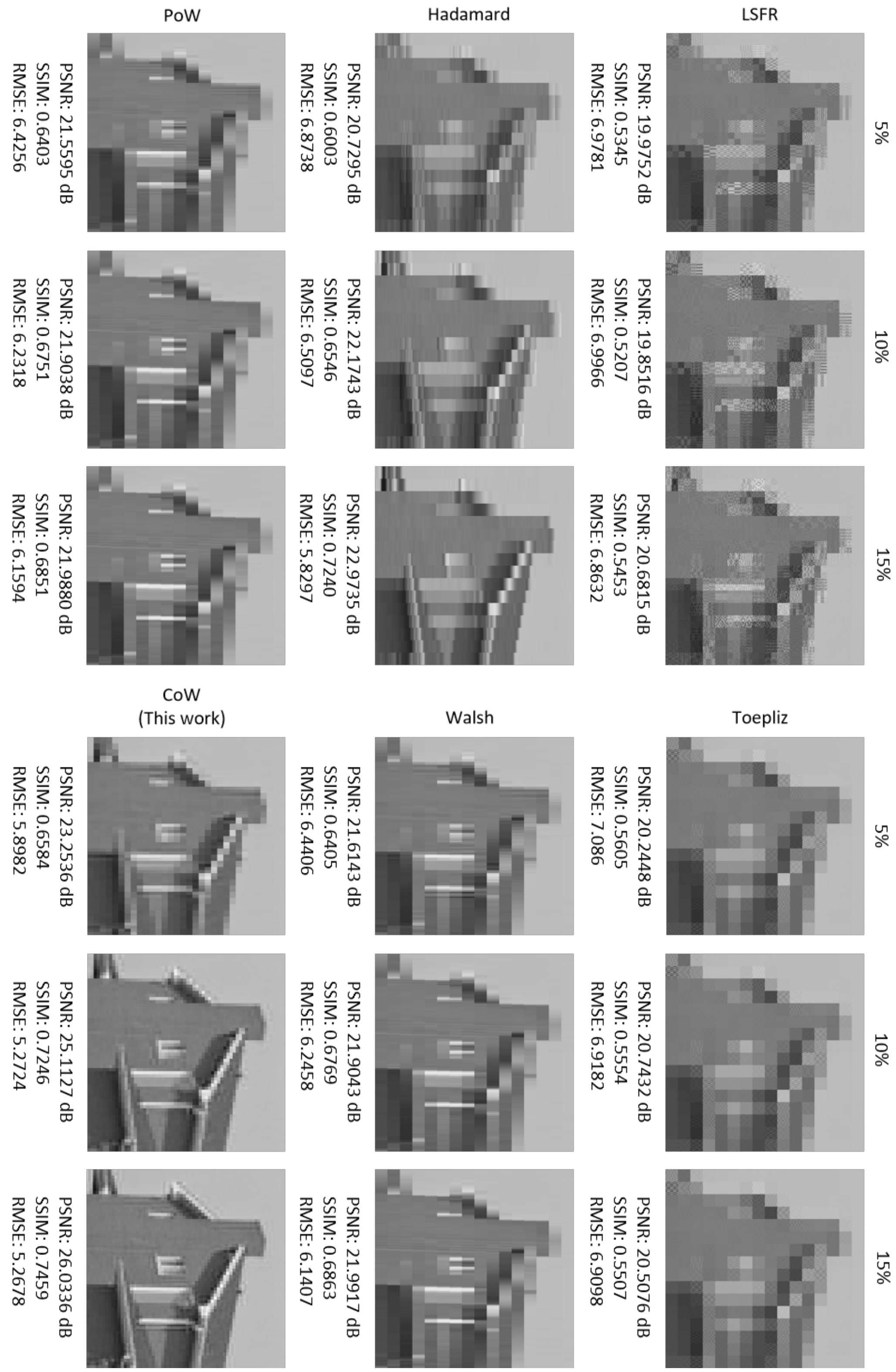

Figure 3.4: An example of reconstructed images of LSFR, Toeplitz, Hadamard, Walsh, PoW, and CoW (this work), where $b = 16$, $n = b \times b = 256$, and SR $\in \{5, 10, 15\}$.

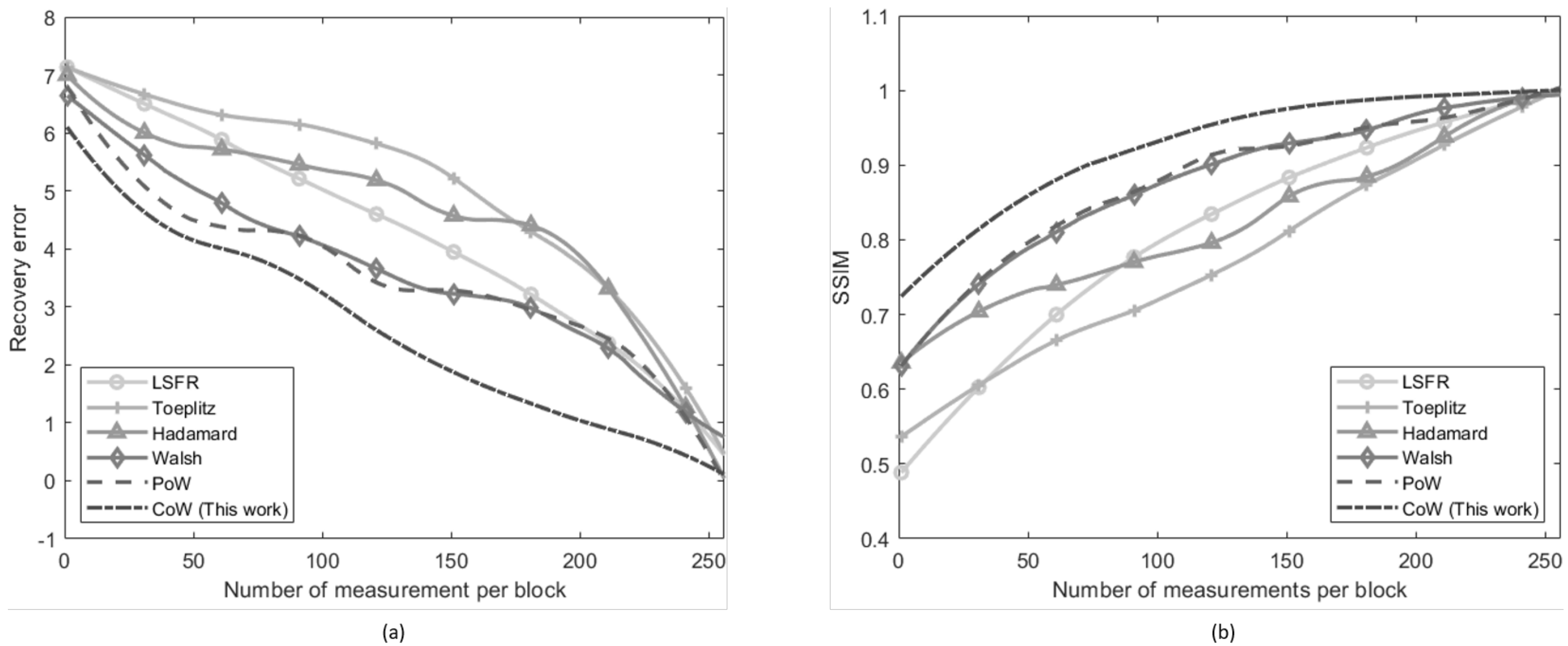

Figure 3.5: Comparison of (a) recovery error via RMSE and (b) SSIM of LSFR, Toeplitz, Hadamard, Walsh, PoW, and CoW (this work), where $b = 16$, $n = b \times b = 256$, and $m \in \{1, 2, 3, ..., n\}$.

Table 3.1: Recovery error via RMSE comparisons of (a) LSFR, (b) Toeplitz, (c) Hadamard, (d) Walsh, (e) PoW, (f) CoW, where $b = 16$ and $\text{SR} \in \{5, 10, 15\}$ (lower is better)

Images	SR = 5%					
	(A)	(B)	(C)	(D)	(E)	(F)
Man	8.38	8.39	8.29	7.59	7.60	**7.45**
Boat	7.25	7.23	7.09	6.88	6.86	**6.55**
House	6.97	7.08	6.87	6.44	6.42	**5.89**
Leopard	7.04	7.03	6.92	6.58	6.58	**6.36**
Lena	7.37	7.43	7.30	6.09	**6.08**	6.23
Average	7.40	7.43	7.29	6.71	6.70	**6.49**

Images	SR = 10%					
	(A)	(B)	(C)	(D)	(E)	(F)
Man	8.39	8.31	8.02	7.38	7.38	**6.84**
Boat	7.33	7.11	6.69	6.73	6.72	**6.04**
House	6.99	6.91	6.50	6.24	6.23	**5.27**
Leopard	7.08	6.97	6.65	6.51	6.51	**5.96**
Lena	7.36	7.38	7.12	5.80	5.78	**5.51**
Average	7.43	7.33	6.99	6.53	6.52	**5.92**

Images	SR = 15%					
	(A)	(B)	(C)	(D)	(E)	(F)
Man	8.31	8.34	7.79	7.31	7.31	**6.62**
Boat	7.23	7.17	6.65	6.63	6.25	**5.92**
House	6.89	6.90	5.82	6.14	6.15	**5.26**
Leopard	7.04	7.01	6.53	6.48	6.48	**5.77**
Lena	7.23	7.40	7.00	5.73	5.70	**5.23**
Average	7.34	7.36	6.75	6.45	6.37	**5.76**

4

A Measurement Coding System for Block-based Compressive Sensing Images by Using Pixel-Domain Feature

This work proposes a method for reducing redundancy in compressively sensed images from a CS camera via Natural ordered Hadamard (NoH). This work employs an intra-prediction algorithm inspired by the traditional image and video coding. Because of the differences between the CS and conventional data structures, it is unable to perform neighboring pixel-to-pixel coding; instead, it suggests an alternative intra-prediction method that performs neighboring vector-to-vector coding. To begin, sensing patterns, to the best of our knowledge, can be used to determine which pixel has been read out. As a result of these properties, it is possible to make a prediction candidate by utilizing single magnitude data from a neighboring block such as the left and upper sides of the current block. These single magnitudes will be multiplied by the sensing matrix to generate prediction candidates, which is a novel method in comparison to previous work. After obtaining several prediction candidates, the residual can be calculated by using the sum-of-absolute difference to calculate the minimum error between the current block and each of the prediction candidates (SAD). The residual will then be transferred to SQ to further reduce magnitude. Finally, Huffman coding will be used to compress the data for efficient transmission over communication channels.

4.1 Introduction

The concept of compressed sensing (CS) is a new paradigm in signal processing for efficient data acquisition and feature extraction. The sparse signal can be recovered from very few samples, which maps the signal from m-dimensional into n-dimensional, where $(m < n)$. Thus, CS camera captures the signals in compressed form, rather than sampling at the Nyquist-Shannon rate and compressing. The robustness of CS by taking the linear projections and computational complexity of the recovery algorithm. The m-dimensional can be recovered to n-dimensional using efficient recovery algorithm (e.g. basis pursuit, OMP, and AMP). As a result, new imaging architectures for CMOS image sensors (CIS) have been proposed in order to sense and compress simultaneously, resulting in a faster image acquisition system. In a wireless cameras network for video surveillance, a massive amount of data is produced. In addition, massive data needs to be transmitted efficiently and secured over the network to monitoring sites where the video streams can be processed and analyzed. However, there is still a significant amount of redundant data in the measurement domain that must be compressed before transmission. In order to address this issue, the CS data coding algorithm is proposed in this paper in order to further reduce data redundancy and improve measurement process efficiency. To begin, instead of a random measurement matrix, Hadamard is used to sensing, compress, and generate predictive candidates. Next, using pixel domain features, a new intra-prediction architecture is proposed to reduce spatial redundancy in the measurement domain. Lastly, image quantization and Huffman coding are used to further compress the data. The reconstruction is carried out using a single iteration basis pursuit with an inverse fast Walsh Hadamard transform. When compared to previous works, the experimental results show that the proposed system can significantly improve coding efficiency, increasing PSNR by 1.94dB - 2.3dB and lowering bitrate by 42% - 65% in terms of bpp.

4.2 Propose measurement coding system

An input images is divided into sub-block size $B \times B$ where $X \in (x_1, x_2, x_3, ..., x_N)$ is vectorized signal of sub-block which will be used through underdetermined system with sensing matrix Φ as shown in Figure 4.1. Furthermore, the intra-preduction algorithm from H.264 is used to inspire this work, and four directional prediction modes are promised, including left mode, upper mode, mean mode, and constant mode, which obtain information from neighboring blocks as shown in Figure 4.2.

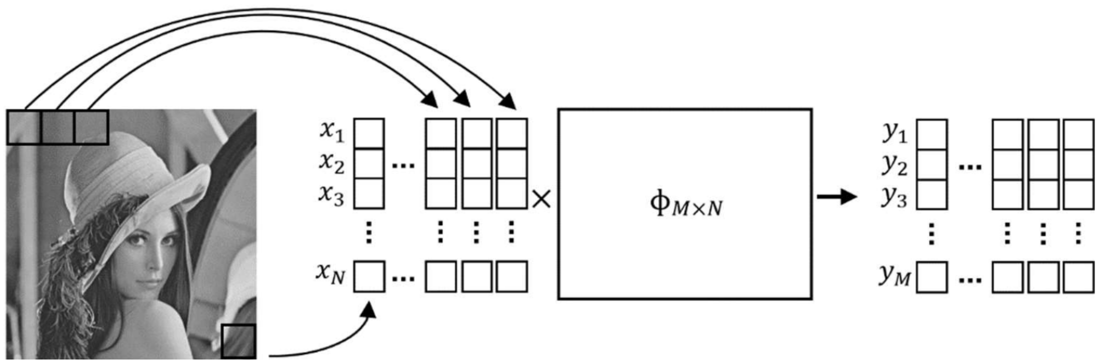

Figure 4.1: An illustrate of block-based compressed sensing sampling process

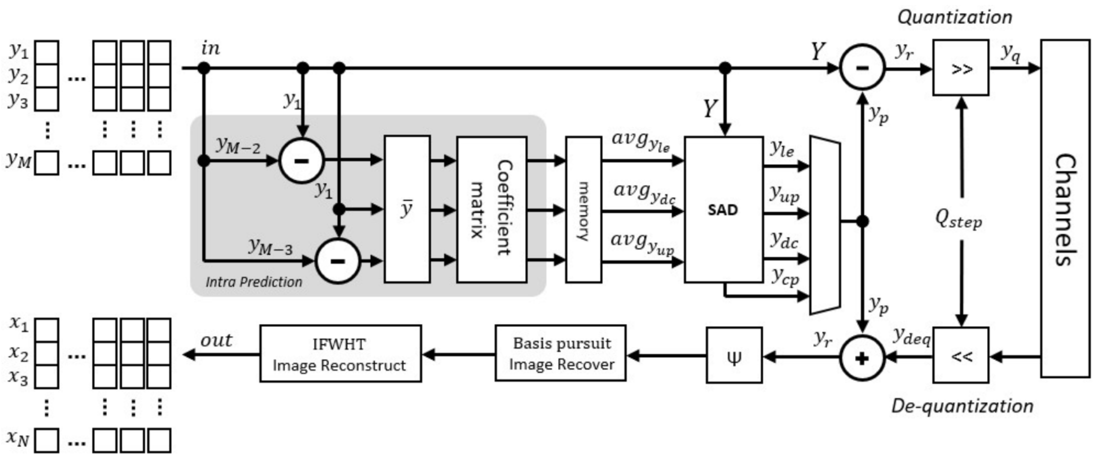

Figure 4.2: An illustrate of propose measurement coding system

Hadamard Transform Based Measurement Matrix

Many researchers have concentrated on random matrices, which are generated by identical or independent probability distributions such as Gaussian, Bernoulli, or Uniform. Random matrices, on the other hand, are simple to build and have a high probability of satisfying the restricted isometry property (RIP). However, there are some drawbacks to using a random matrix. The use of random matrices, for example, slows the recovery process and makes large-scale problems impossible to solve. This work used hadamard as sensing matrix with complementary entries. The columns of this matrix are orthogonal. Given a matrix H of order N, H is said to be a hadamard matrix if the transpose of the matrix H is closely related to its inverse. This can be expressed by

$$HH^T = NI_{N\times N} \tag{4.1}$$

where $I_{N\times N}$ is identity matrix and H^T is transposed matrix of H. The random hadamard matrix consists of taking random rows from the hadamard matrix. This measurement matrix satisfies the RIP with probability at least $1-5/N-e^{-\beta}$ providing $M \geq C_0(1+\beta)KlogN$ where β and C_0 contain positive constants, K is sparsity level of signal. After that, the matrix is binarized to increases measurement processing time and decrease the hardware cost. As the hyphothesis that random matrices has influential to image degradation, the comparison between random and the proposed matrix are demonstrated by increased 6% in PSNR and decrease 39% in bpp.

Spatial Redundancy Reduction using Pixel-Domain Features

To reduce spatial redundancy, intra-prediction has been applied to this research. However, the concept of traditional intra-prediction cannot be applied to CS directly due to no similarity in measurement domain. Therefore, the new intra-prediction architecture has been proposed for CS. Since we can obtaine pixel values only in compressed form $Y \in \{y_1, y_2, y_3, ..., y_M\}$, where y_1 is the summation of all pixel. Meanwhile, Hadamard measurement matrix introduces a patterns to extract two kind of boundary information by y_{M-2} and y_{M-3} for upper and left respectively. Then, we can approximated the boundary information by subtract y_1 with y_{M-2} and y_{M-3}, respectively. After the subtraction the summation of the pixel will be represent in compressed form. Next, divide the summation by the total number of pixels to get an average value. Furthermore, a special prediction mode called mean mode y_{dc} is added as shown in Figure 4.3. To select a predictive candidate, sum of absolute differences (SAD) is required to compare between Y and predictive candidates. In case there is no predictive candidate, the prediction candidate willl be zero for transmitting without prediction. The simple quantization is needed for entropy coding before transferred to Huffman coding.

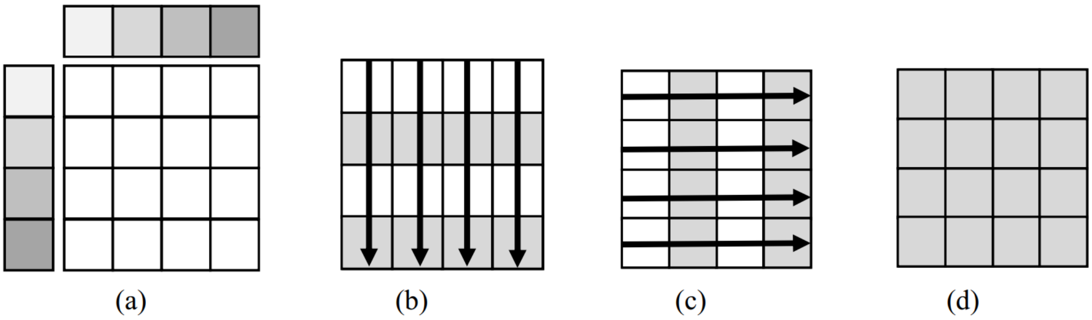

Figure 4.3: Illustration of (a) the block to be predicted surrounded by nearest neighbor pixels and directional prediction modes available for predictive candidates: (b) vertical for upper, (f) horizontal for left, and (d) mean for dc mode

4.3 Experimental results

The proposed measurement coding for BCS was implemented using MATLAB. All of experiment use 512 × 512 grayscale images of Lenna, Mandril, Goldhill, and Pentagon from "The USC-SIPI image database". The image was divided into non-overlapping by B = 4, 8, and 16, SR = 0.75, 0.5, and 0.25, and quantization parameter $Q_{step} \in \{0, 1, 2, 3, 4, 5, 6\}$. The image is reconstructed using basis pursuit with the inverse discrete fast Walsh-Hadamard transform (IFWHT). RD-distortion and performance is measured in terms of PSNR, bitrate in bits per pixel (bpp) using entropy quantize to estimate actual bitrate, and structural similarity index (SSIM) for perceptual metric that quantifies image quality degradation. The comparison of random and hadamard matrices with same condition by intra prediction and quantization step is demonstrated. As shown in Table 1. both matrix can achieve high PSNR. However, random matrix has higher bitrate than hadamard matrix. Thus, the construction of measurement matrix is greatly effects coding efficiency by increased 6% in PSNR and decrease 39% in bpp. The BD-rate curve between PSNR and bpp of four test image are plotted, which our framework can greatly achieve almost same PSNR with lower bpp as shown in Figure 4.4 The comparison of visual qualities and compression artifacts are demonstrated in Figure 4.5 and Figure 4.6. Since we applied intra-prediction and quantization, our proposal can greatly increase 1.94dB - 2.3dB in PSNR and reduced 42% - 65% in bpp when compared to BCSSPL, SQ, and DPCM respectively as shown in Table 4.1.

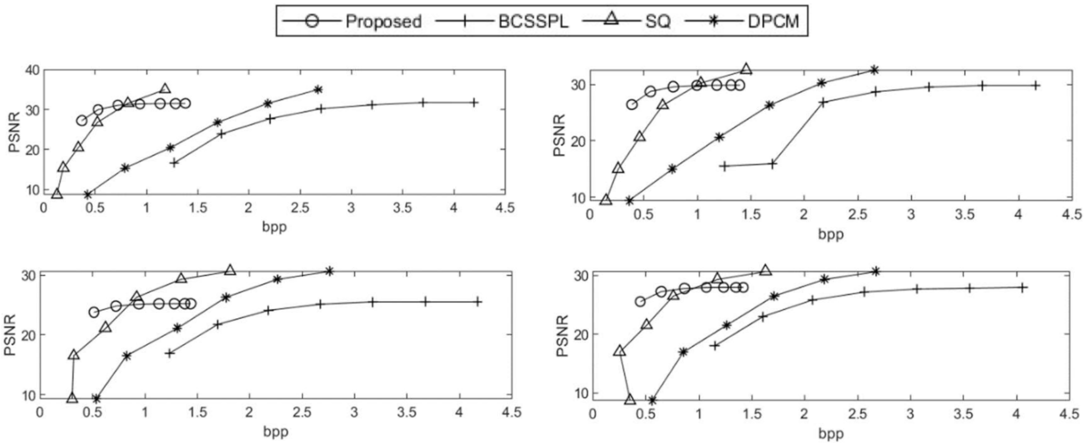

Figure 4.4: BD-rate curve of four test images: (a) Lenna, (b) Goldhill, (c) Mandrill, and (d) Pentagon among Proposed, BCSSPL ([35]), SQ, and DPCM ([62]) with $B = 4$, SR = 0.5 with $Q_{step} \in \{0, 1, 2, 3, 4, 5, 6\}$

Table 4.1: PSNR comparison among BCSSPL ([35]), SQ, and DPCM ([62]), and Proposed ($Q_{step} = 4, B = 4, SR = 0.5$).

Image	BCSSPL ([35])		SQ		DPCM ([62])		Proposed	
	PSNR	bpp	PSNR	bpp	PSNR	bpp	PSNR	bpp
Lenna	27.62	2.20	26.74	1.69	26.74	0.52	31.05	0.84
Goldhill	26.77	2.17	26.29	1.67	26.29	0.67	29.56	0.89
Mandrill	24.12	2.17	26.67	1.77	26.27	0.91	25.13	1.65
Pentagon	25.83	2.07	26.46	1.70	26.46	0.75	27.78	0.98
Avg.	26.08	2.15	26.44	1.70	26.44	0.71	28.38	1.09

4.4 Summary

CS has been considered as innovative technology in signal sampling and compressing. In this paper, we present our architecture and simulation results of BCS framework using the measurement-domain to generate pixel-domain features, which can reduce the redundant information with low computational complexity and realize high compression ratio via hadamard matrix. The pixel-domain features is obtaining by subtracting with intra-prediction for upper, left and dc mode respectively and divided the summation by total active pixel. The simulation results shown that this framework can increase 1.94dB - 2.3dB in PSNR and reduced 42% - 65% of bitrate in terms of bit-per-pixel when

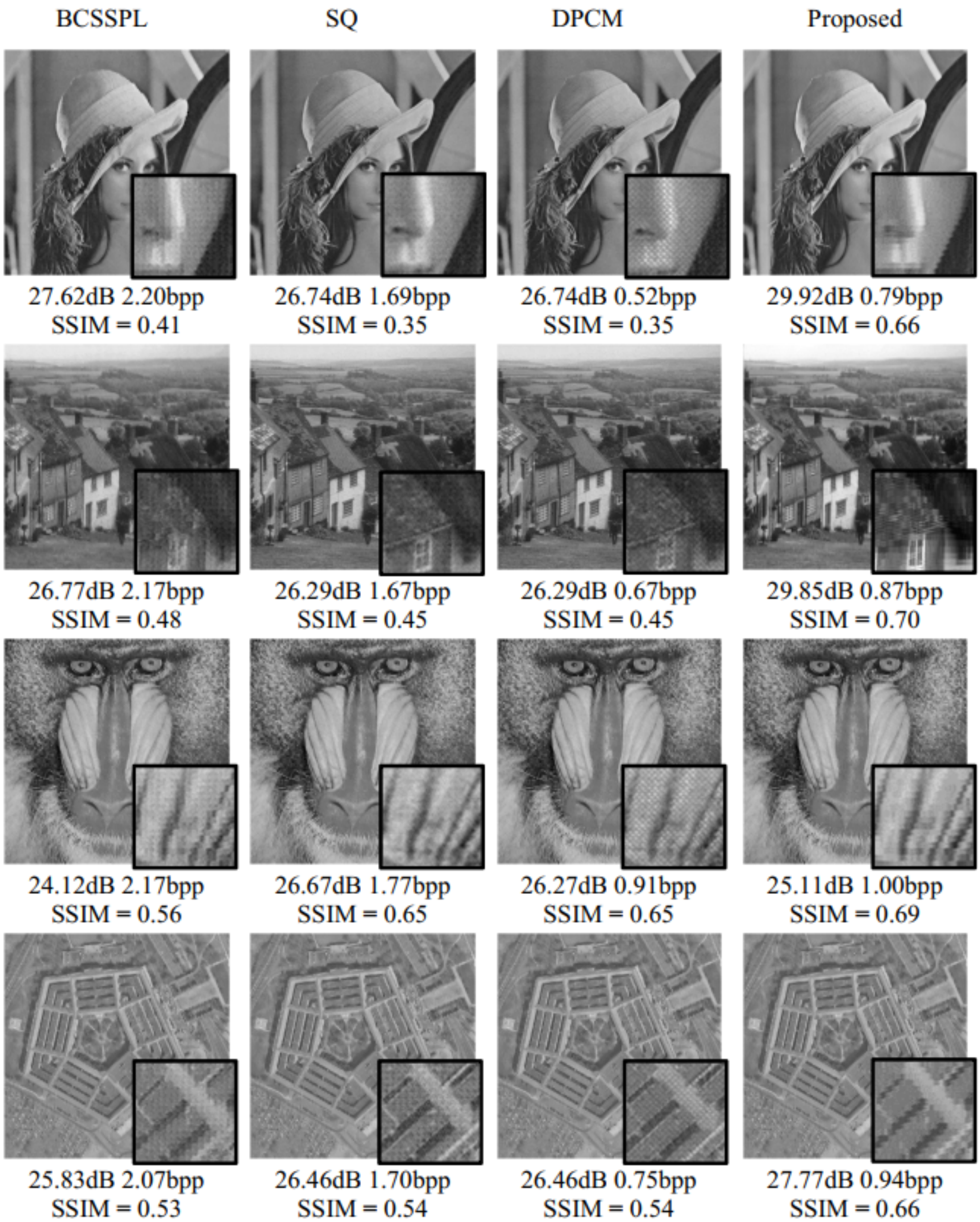

Figure 4.5: Comparison of visual qualities and compression artifacts of four test images: Lenna, Goldhill, Mandrill, and Pentagon among our proposal, BCSSPL ([35]), SQ, and DPCM ([62]) with $B = 4$, SR$= 0.5$ with $Q_{step} = 4$

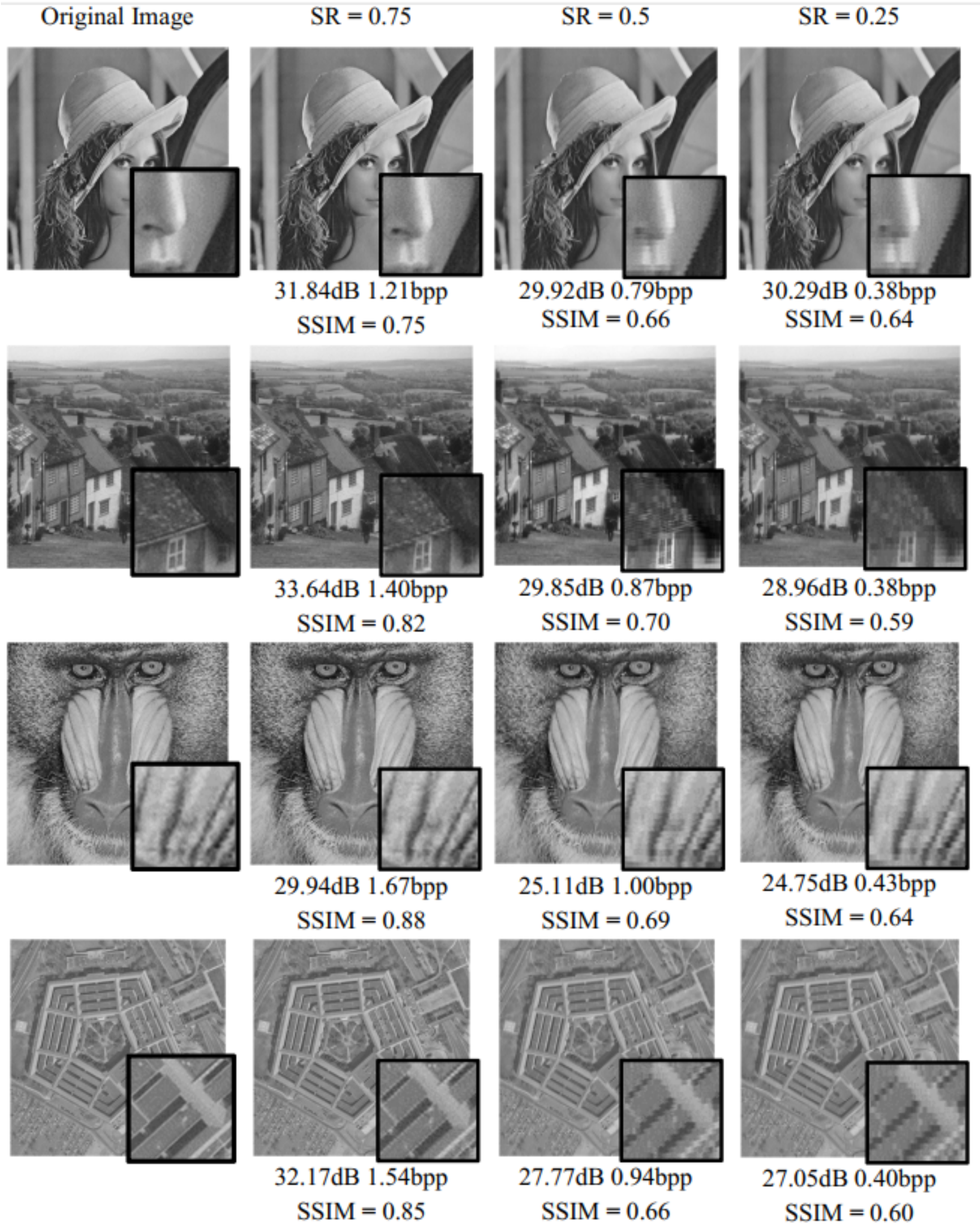

Figure 4.6: Comparison of visual qualities of four test images: Lenna, Goldhill, Mandrill, and Pentagon with $B = 4$, SR = 0.75, 0.5, and 0.25 with $Q_{step} = 4$.

compared to previous works. The design of measurement matrix is greatly importance for image and video compression using compressed sensing.

5

Intra Prediction Based Measurement Coding Algorithm for Block-Based Compressive Sensing Images

Block-based compressive image sensor (BCIS) captures light and represents them as compressed data called measurement. It has potential to revolutionize conventional image and video acquisition system that builds upon high complexity and redundant process. However, by comparing the compression performance between these two systems, BCIS cannot reduce bitrate to similar factor as the compressed media by pixel-based compression algorithms. It still requires enormous amounts of bit to store and transmit data. In this work, we introduce intra prediction based measurement coding (IPMC) algorithm for giving an extra compression performance to measurement. Moreover, importantly, there is a requirement that sensing matrix for BCIS must not be derived from non-uniform distribution in order to control prediction accuracy and quality. Therefore, we use structural sensing matrix made of sequency-ordered Walsh-Hadamard. Furthermore, it allows boundary pixels of adjacent blocks to be accessible through measurement, which helps intra prediction to generate its candidates accurately. The algorithm encodes prediction error between target measurement and multiple prediction candidates, resulting in smaller data size. This work can significantly reduce bpp by 10.90% and simultaneously increase 3.95 dB in PSNR compared to the state-of-the-art works. Moreover, we implemented the proposal on FPGA. It gave 10 times higher throughput than software. The core power consumption is at 50 mW and working at 88 MHz when processing 3840 $\times$ 2160 pixels with the sampling rate of 1/4.

5.1 Introduction

Over the past few years, block-based compressive image sensor (BCIS) has gained significant interest in imaging technology. It can solve analog-to-digital converter (ADC) problems in conventional image sensor such as slow pixel readout time and power consumption. A single-pixel camera was successfully developed by using digital micromirror device (DMD) array (32). It is useful for microscopy and microanalysis applications (70). Nevertheless, it required long sensing time when the resolution is relatively high. Robucci et al. proposed separable-transform BCIS (79). It could capture image faster than single-pixel camera. Nevertheless, it had limited in frame-per-second (fps). Later, Oike and El Gamal proposed programmable BCIS with per-column Sigma-Delta ADC to reduce sensing time and increase frame-rate up to 1920 fps (67), which overcame the problems of (32) and (79). In (75) and (49), they gave an opinion that BCIS can revolutionize image and video capturing and compressing scheme, where bitrate could be varied depending on preferred quality. However, at this stage, it is not suitable for consumer devices because they do not have enough computational resources for decoding the measurement. In the meantime, there is a suitable application for BCIS such as wireless surveillance system because the measurement can be decoded at monitoring sites with unlimited resources(48) (63). To transmit measurement wirelessly, it must be compressed into more compact format, resulting in lower transmission costs (80) (42). Therefore, in this work, we propose four modes intra prediction based measurement coding (IPMC) algorithm including upper, left, average, and no prediction. However, there is a requirement that sensing matrix for BCIS must not be derived from non-uniform distribution in order to control prediction accuracy and quality. Therefore, we use structural sensing matrix (SSM) made of sequency-ordered Walsh-Hadamard (SoWH). It allows boundary pixels of adjacent blocks to be accessible through measurement, which helps intra prediction to generate its candidates accurately. The algorithm encodes prediction error between target measurement and multiple prediction candidates, resulting in smaller data size.

To demonstrate the applicability and versatility, we evaluate the proposal using numerous video datasets in 4K resolution using peak-to-noise-ratio (PSNR), structural similarity index measure (SSIM), and bits-per-pixel (bpp). Moreover, we increase compression throughput by extending the proposal from software to hardware using FPGA. We establish notation and provide a brief background of compressive sensing theory in chapter A. and related works on measurement coding in chapter B. Section II provides the proposed IPMC algorithm for BCIS and hardware architecture. Section III provides extensive simulation results and compares them with state-of-the-art works. Further, we also report hardware implementation summary. Section IV provides conclusions.

5.2 Compressive sensing theory

Comressive sensing (CS) is built upon two major fundamental conditions consisted of sparsity and incoherent (98) (19). There are essential conditions in order to apply CS. First, a signal characteristic of x must be sparse when expressing in a specific orthonormal transform basis. Next, sensing matrix $\Phi \in \mathbb{R}^{m \times n}$, it can be made of random distribution called random sensing matrix (RSM), where the number of dimensional m must be less than n, known as sampling rate (SR).

The measurement $y \in \mathbb{R}^{m \times 1}$ can be obtained by projecting Φ to x. Nevertheless, traditional CS is not suitable for a large scale problem because it requires long sensing time. In (36) proposed partitioning approach to traditional CS by dividing an entire frame into multiple non-overlapping blocks. Instead that n equal to frame size, now n will be equal to $b \times b$, where b is block size. Hence, x will be sampled with smaller Φ, resulting in faster projection. It can be expressed mathematically as:

$$y_i = \Phi x_i \tag{5.1}$$

where $y_i \in \{\dot{y_1}, \dot{y_2}, \dot{y_3}, ..., \dot{y_m}\}$ is measurement of compressible signal of $x_i \in \{\dot{x_1}, \dot{x_2}, \dot{x_3}, ..., \dot{x_n}\}$ and i is the block order through raster scanning as shown in Figure 8.2. To guarantee a good reconstructed image, the sensing matrix should satisfy restricted isometry property. In this work, we recover y by using a classical method via convex optimization called ℓ_1-minimization.

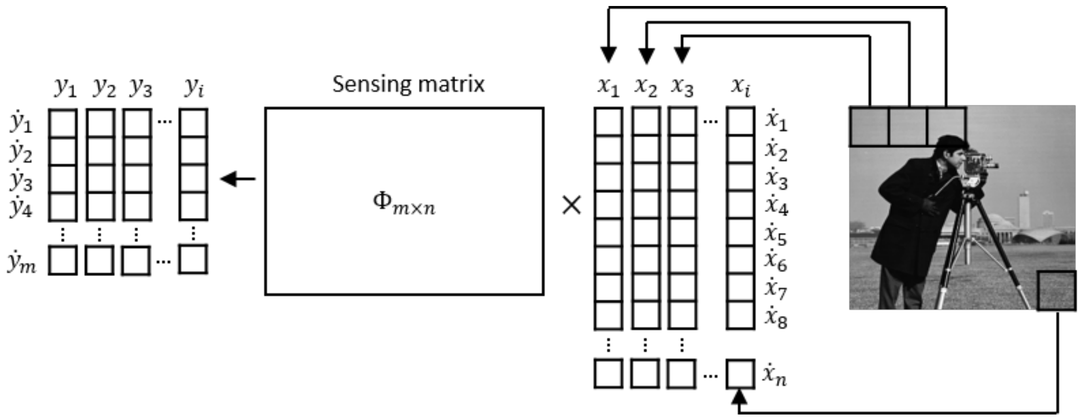

Figure 5.1: Non-overlapping blocks compressive sensing diagram.

5.3 Related works on measurement coding

Up to the present time, most of the CS literature has been devoted to study the recovery of sparse signals from a small number of measurement, but less in measurement coding algorithm. By referring to the legacy vector compression algorithms, introduced for lossless coding (84), and its extension for lossy coding (109). Although, it is possible to use these legacy approaches to encode measurement. Nevertheless, it requires a precise design for each system specifically, which is not convenient.

Scalar quantization (SQ) provided a straightforward approach to compress measurement. By comparing to vector compression algorithms, SQ gave higher performance and versatility than vector compression. Nevertheless, it has been established that SQ is highly inefficient in terms of information-theoretic rate-distortion (RD) performance (58) (65) (115) (123) (22). Additionally, it require an iterative recovery algorithm to predict corrupted quantized measurement such as quantized iterative hard thresholding (QIHT) (46), quantized compressed sampling matching pursuit (QCoSaMP), and adaptive outlier pursuit for quantized iterative hard thresholding (AOP-QIHT) (82).

Next, differential pulse-code modulation (DPCM) was introduced to reduce bitrate (62). DPCM used a single prediction candidate to predict target measurement. However, the disadvantage is that the single prediction candidate may contain irrelevant information to target measurement, resulting in unstable bitrate reduction. Afterward, spatially directional predictive coding (SDPC) was introduced in (116). This work was implemented based on DPCM. It gave higher compression performance than SQ and

DPCM, while improved image quality. However, this work used RSM as sensing matrix, resulting in unstable quality and unstable bitrate when coding the same image. Importantly, they could not embed this kind of sensing matrix into hardware device, where it can only handle binary signal sources.

Later, intra prediction based measurement coding with modification of RSM was introduced in (121). This work was inspired by intra prediction concept from conventional pixel-based compression algorithms that uses boundary pixels of adjacent blocks to predict target block. By imitating the conventional approach, they modified sensing patterns of RSM corresponding to obtain boundary pixels of adjacent blocks called hybrid sensing matrix (HSM). They used that boundary pixels information to generate intra prediction candidates. This work significantly reduced bitrate lower than SQ, DPCM, and SDPC. Nevertheless, it produced sampling artifact to the image due to the modification of sensing matrix.

In our previous work (73), we adopted SSM made of Natural ordered Hadamard (NoH) to obtain boundary information. We proposed four modes intra prediction included upper, left, average, and no prediction. This work significantly reduced bitrate and improved image quality compared to other works in this literature.

5.4 Proposed intra prediction based measurement coding for BCIS

This work adopted SSM made of Natural ordered Hadamard (NoH) to obtain boundary information. We proposed four modes intra prediction included upper, left, average, and no prediction. This work significantly reduced bitrate and improved image quality compared to other works in this literature. In this work, we improve coding performance and image quality based on the previous work in 4. The overall architecture including BCIS and IPMC algorithm can be seen in Figure 5.2.

There are three primary signals control BCIS including column selector, row selector, and pixel selector where each pixel will be selected according to the sensing matrix. Subsequently, we adopt SoWH as SSM. It can gather information more efficiently than RSM, HSM, and SSM made of NoH due to higher orderliness, resulting in better image quality. Furthermore, it allows pixels boundary information of neighboring blocks to be accessible through measurement without modifying the sensing matrix. After

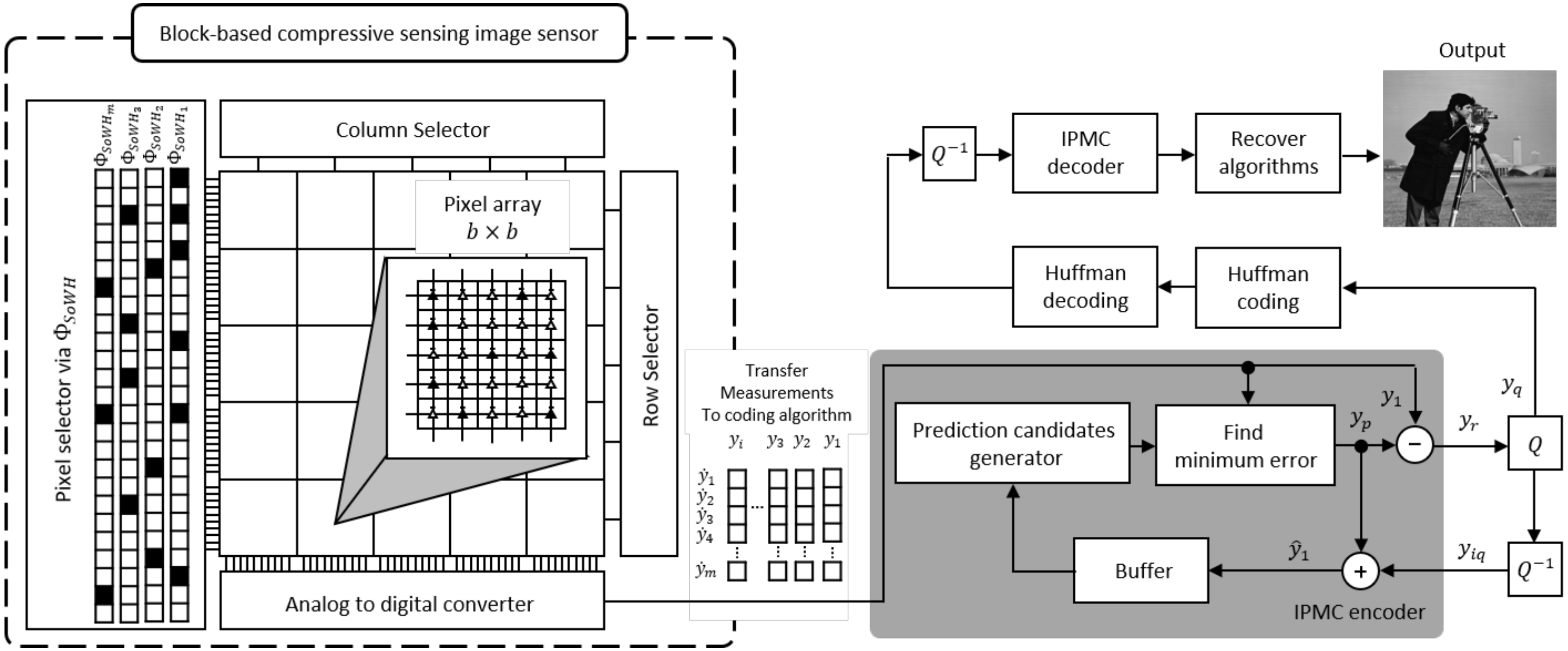

Figure 5.2: The overall architecture of BCIS and IPMC algorithm, where each measurement of block denoted by y_1, y_2, y_3, and y_m, respectively. Subsequently, each element in measurement denoted by $\dot{y_1}$, $\dot{y_2}$, $\dot{y_3}$, and $\dot{y_m}$ were sampled by Φ_{SoWH_1}, Φ_{SoWH_2}, Φ_{SoWH_3}, and Φ_{SoWH_m}, respectively. To demonstrate the data flow of the coding process, we illustrated an example of the process by using y_1, which refers to the first block.

obtaining the measurement, it will be compressed using the IPMC algorithm. By utilizing the boundary pixels of adjacent blocks, this work provides four modes of intra prediction. The target measurement encodes by minimizing distortion with multiple prediction candidates, resulting in a smaller data set. For Huffman coding, quantization is used to reduce the probability symbol. We also include inverse quantization within the transmitter to estimate prediction candidate loss by quantization at the receiver. The estimated prediction candidates are then used to predict the next target measurement. Otherwise, the decoder at receiver will act as error accumulator, in which a single corrupted measurement can initiate recovery error to the whole image. Moreover, to increase coding performance, throughput, and to realize IPMC algorithm in real-world applications, we implement the proposal in hardware level and evaluate it on FPGA.

Sequency-ordered Walsh-Hadamard sensing matrix

To use BCIS to capture the light, there is an implementation constraint that sensing matrix Φ must be $\{0,1\}$ because the pixel selector can handle only digital signal (i.e., low (0) and high (1)). Let Φ_{NoH} can be obtained by order n, it can be said to be Φ_{NoH} if the transpose of the matrix Φ_{NoH}^T is closely related to its inverse. It can be expressed as given below:

$$\Phi_{NoH}\Phi_{NoH}^T = nI_{n\times n} \tag{5.2}$$

where $nI_{n\times n}$ is the identity matrix and Φ_{NoH}^T is the transpose of matrix. By applying Sylvester's construction to Φ_{NoH}, resulting in Walsh-Hadamard matrix denoted by Φ_{WH} as the following:

$$\begin{aligned}\Phi_{WH}\left(2^k\right) &= \begin{bmatrix}\Phi_{NoH}\left(2^{k-1}\right) & \Phi_{NoH}\left(2^{k-1}\right)\\ \Phi_{NoH}\left(2^{k-1}\right) & -\Phi_{NoH}\left(2^{k-1}\right)\end{bmatrix}\\ &= \Phi_{NoH}\left(2\right)\bigotimes\Phi_{NoH}\left(2^{k-1}\right)\end{aligned} \tag{5.3}$$

for $2 \leq k \in n$, where $\bigotimes$ denotes the Kronecker product. Subsequently, we applying bit-reversal and gray-code permutation, resulting in sequency order of Φ_{WH} denoted by Φ_{SoWH} this sensing matrix satisfies the RIP with a probability of at least $1-5/n-e^{-\beta}$ providing $m \geq c(1+\beta))\kappa\, log\, n$ where β is a positive constant and κ is sparsity level. In general, Hadamard transform will return matrix in $\{-1,1\}$. Therefore, we binarize them from $\{-1,1\}$ to $\{0,1\}$ to possibly embed the sensing matrix into BCIS.

Intra prediction based measurement coding algorithm

In general, boundary pixels of adjacent blocks have information that closely related to target block. Hence, we use that boundary information to deliver four modes intra prediction including upper, left, average, and no prediction.

Firstly, prediction parameter preparation, different sensing patterns can refer to each order of Φ_{SoWH}, which use to obtain each element of y. It offers several features that allow boundary information to be accessible from measurement. Hence, we can trace back to which pixels in the block had read. In this work, there are three significant sensing patterns as shown in Figure 5.3. where white and black squares indicate the

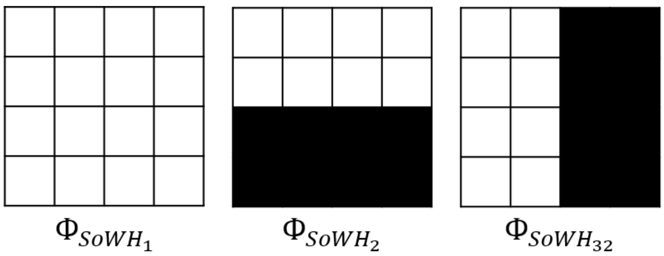

Figure 5.3: An example of significant sensing patterns in Φ_{SoWH}.

pixel that is being read and skip, respectively.

For instance, by multiplying x with Φ_{SoWH_1}, the first element denoted by $\dot{y}_1$ is the summation of 4×4 pixels; The second element $\dot{y}_2$ can be obtained by multiplying x with Φ_{SoWH_2}, which is the summation of upper-half 2×4 pixels; and the third element $\dot{y}_{32}$ can be obtained by multiplying x with $\Phi_{SoWH_{32}}$, which is the summation of half-left 4×2 pixels. However, the parameters that necessary for generating prediction candidates are located in black squares, which are opposite-side of Φ_{SoWH_2} and $\Phi_{SoWH_{32}}$. To retrieve them, since $\dot{y}_1$ is a summation of all pixels in the block. Therefore, the data in black squares can be obtained by subtracting $\dot{y}_1$ with $\dot{y}_2$ and $\dot{y}_1$ with $\dot{y}_{32}$, resulting in sum of bottom-half 2×4 pixels and sum of the right-half 4×2 pixels, respectively. To understand the concept clearer, we present the subtraction process by referring to sensing patterns subtraction as shown in Figure 5.4. This method delivers the same results as modifying the sensing matrix to obtain boundary pixels of adjacent blocks. Besides, the image quality will not be disturbed as the work in . Further, the group of pixels after subtraction over image is illustrated in Figure 5.5. At this stage, the

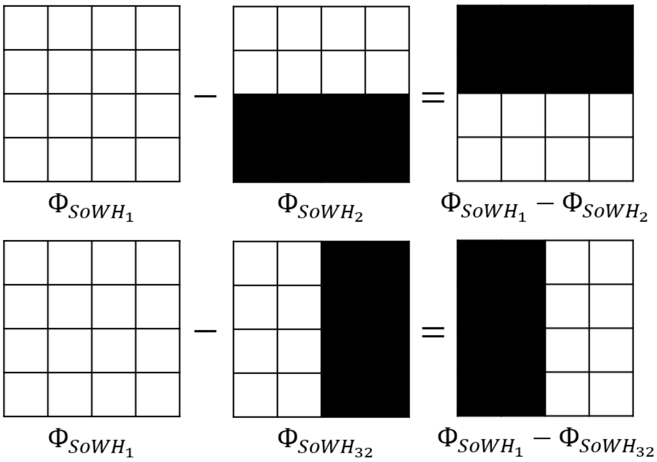

Figure 5.4: An example of effective pixels after element subtraction.

Figure 5.5: An example of group of pixels over image, where the yellow area indicates the pixel group for target measurement; the blue area indicates the pixel group for left mode; the green area indicates the pixel group for up mode.

parameters are the representation of multiple pixels. It is necessary to average them by dividing by the number of active pixels (i.e., in the case of 2 × 4 pixels and 4 × 2 pixels, the number of active pixels equals 8). Afterward, we multiply the averaged parameters with Φ_{SoWH} to generate vector known as intra prediction candidate. To sum up, the candidate generation procedure of each mode can be explained by the following equations: Up mode:

$$y_u = \frac{(\dot{y}_1 - \dot{y}_2)}{\sum_{j=1}^{n}(\Phi_{SoWH_{1,j}} - \Phi_{SoWH_{2,j}})} \times \Phi_{SoWH} \tag{5.4}$$

Left mode:

$$y_l = \frac{(\dot{y_1} - \dot{y_{32}})}{\sum_{j=1}^{n}(\Phi_{SoWH_{1,j}} - \Phi_{SoWH_{32,j}})} \times \Phi_{SoWH} \tag{5.5}$$

Average mode

$$y_{avg} = \frac{(y_u + y_l)}{2} \tag{5.6}$$

The final prediction candidate y_p can be estimated by finding minimum error between target measurement y with prediction candidates $y_c \in \{y_u, y_l, y_{avg}\}$. It can be expressed as the following:

$$y_p = \underset{y_c \in \{y_h, y_v, y_{avg}\}}{\arg\min} \left\| y_i - y_c \right\|_{l_1} \tag{5.7}$$

In addition, in case there is no prediction candidate selected from y_c, y_p will be equal to zero, which means no prediction. The residual measurement y_r can be calculated by subtracting y with y_p. It can be expressed by

$$y_r = y - y_p \tag{5.8}$$

Scalar Quantization

We further reduce bitrate and probability symbols of y_r using SQ. It maps residual measurement y_r into a finite sequence of codewords with quantization step equal to Δ. It can be expressed by:

$$\Delta = \left\lfloor \frac{\max(y_r) - \min(y_r)}{2^{Q_b}} \right\rfloor, \tag{5.9}$$

$$y_q = \left\lfloor \frac{y_r}{\Delta} \right\rfloor \tag{5.10}$$

where Q_b is quantization bit and quantized measurement denoted by y_q. Subsequently, inverse quantization maps y_q into y_{iq} that is an approximation of y_r. It can be expressed by:

$$y_{iq} = \Delta \cdot y_q \tag{5.11}$$

In this work, we fixed Q_b at 4 bits, which is sufficient to reduce bitrate and probability symbols. Furthermore, we include inverse quantization inside the transmitter to estimate prediction candidates loss by quantization at the receiver. Subsequently, we use that estimated prediction candidates to predict the next target measurement. If both sides do not have the same prediction candidates information, the decoder will act as error accumulator, in which a single corrupted measurement can ruin the whole image. Note that, y_q needs to transmit along with 2 bits side information of prediction mode and Δ of each block to the receiver.

Hardware implementation of proposed IPMC algorithm for BCIS

In this section, we extend IPMC algorithm from software to hardware for increasing throughput. The hardware architecture can be seen in Figure 5.6. It can be placed next to BCIS. Hence, the target measurement y can be encoded and transmitted immediately. The hardware procedure of measurement obtaining and coding can be described as the following step:

Step 1: Send block coordinate denoted by $rows\_addr$ and $columns\_addr$ to BCIS.

Step 2: Obtain y from BCIS according to $rows\_addr$ and $columns\_addr$.

Step 3: Fetch prediction parameters from registers. We note that when block coordinate $rows\_addr = 1$ and $columns\_addr = 1$, y will go straight to quantization without prediction. This is a special case in IPMC algorithm because there are no prediction parameters available for the first block.

Step 4: Average prediction parameters and multiply them with Φ_{SoWH} to generate y_u and y_l.

Step 5: Find minimal error of y among $y_c \in \{y_u, y_l, y_{avg}\}$, resulting in final prediction candidate y_p.

Step 6: Subtract y with y_p, resulting in y_r.

Step 7: Apply quantization to y_r, resulting in y_q.

Nevertheless, the data structure of y is vector. Without optimization, it requires at least $m-1$ clock cycles to encode y. Therefore, we optimize vector summation module using a tree-like pipeline technique as shown in Figure 5.7a. and non-pipeline in Figure 5.7b, in which clock cycle can be shorten from $m-1$ to $log_2(m)$. Consequently, it requires slightly higher resources than non-pipeline. Subsequently, it is necessary to

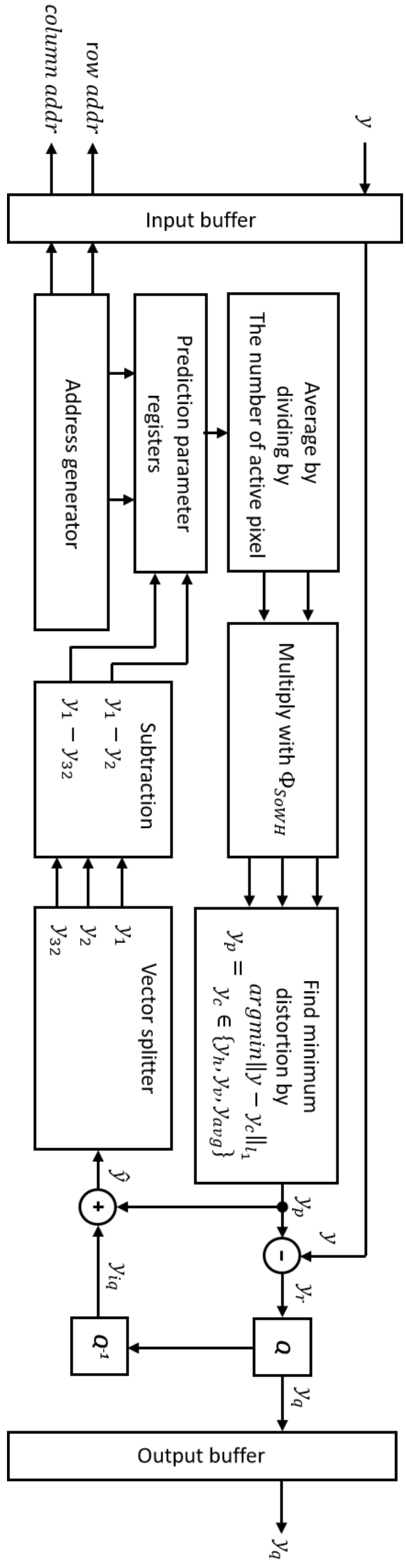

Figure 5.6: Hardware architecture of the proposed IPMC algorithm.

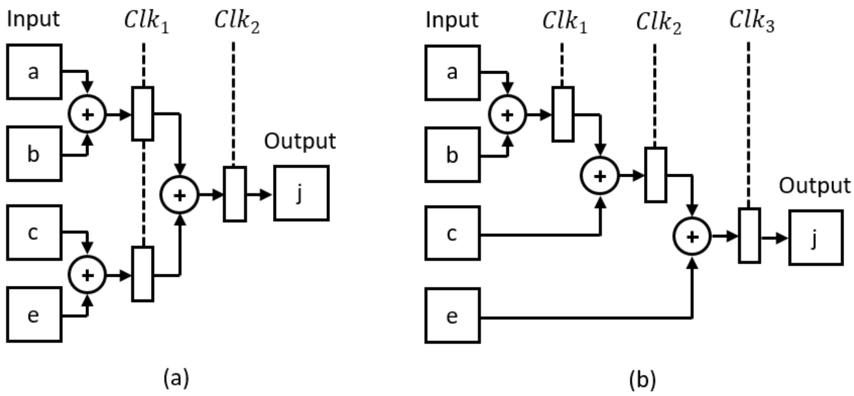

Figure 5.7: A diagram of vector summation for $m = 4$, where (a) tree-like vector pipeline which the summation can be done within $log_2(m)$ clock cycles and (b) non-pipeline which the summation can be done within $m - 1$ clock cycles.

prepare the prediction parameters for the next target measurement. The procedure of prediction parameters preparation can be described as the following step:

Step 1. Apply inverse quantization to y_q, resulting in y_{iq}.

Step 2. Decode y_{iq} by adding y_p, resulting in $\hat{y}$.

Step 3. Extract $\hat{y}$ using vector splitter to obtain prediction parameters (i.e., $\dot{\hat{y}}_1, \dot{\hat{y}}_2$, and $\dot{\hat{y}}_{32}$).

Step 4. Subtract $\dot{\hat{y}}_1$ with $\dot{\hat{y}}_2$ and $\dot{\hat{y}}_1$ with $\dot{\hat{y}}_{32}$.

Step 5. Store the results in registers for the next prediction.

5.5 Experimental results

The performance of the IPMC algorithm is evaluated in this section using PSNR, SSIM, and bpp. The simulation results delivered by MATLAB using l_1-minimization via primal-dual interior-point method. We used multiple 4K datasets consisted of Beauty, ReadySetGo, Bosphorus, and HoneyBee. Lastly, we reported hardware implementation results in terms of device specification and throughput.

Table 5.1: Overall performance comparison with the state-of-the-art on various datasets, where $b = 16$, SR=1/4, and $Q_b = 4$.

Methods	Beauty dataset		ReadySetGo dataset	
	PSNR (dB)/SSIM	bpp	PSNR (dB)/SSIM	bpp
Bernoulli + SQ	34.45 / 0.86	2.71	30.83 / 0.67	2.72
NoH + SQ	32.72 / 0.77	**0.39**	29.69 / 0.48	**0.51**
Gaussian + SQ	33.27 / 0.75	2.62	30.79 / 0.50	2.64
DPCM + SQ	36.15 / 0.88	2.71	31.23 / 0.72	2.71
SDPC + SQ	36.27 / 0.88	2.71	31.31 / 0.73	2.70
Intra Pred. + HSM + SQ	35.57 / 0.87	2.30	31.04 / 0.70	2.18
This work	**38.90 / 0.92**	2.21	**34.36 / 0.87**	2.05

Methods	Bosphorus dataset		HoneyBee dataset	
	PSNR (dB)/SSIM	bpp	PSNR (dB)/SSIM	bpp
Bernoulli + SQ	31.07 / 0.71	2.70	31.84 / 0.85	2.72
NoH + SQ	30.02 / 0.51	**0.25**	30.00 / 0.67	**0.35**
Gaussian + SQ	29.83 / 0.36	2.59	30.79 / 0.67	2.63
DPCM + SQ	34.95 / 0.92	2.74	32.85 / 0.89	2.74
SDPC + SQ	34.99 / 0.92	2.69	32.96 / 0.90	2.73
Intra Pred. + HSM + SQ	34.70 / 0.91	2.20	31.72 / 0.83	1.90
This work	**36.47 / 0.94**	1.92	**35.64 / 0.94**	1.67

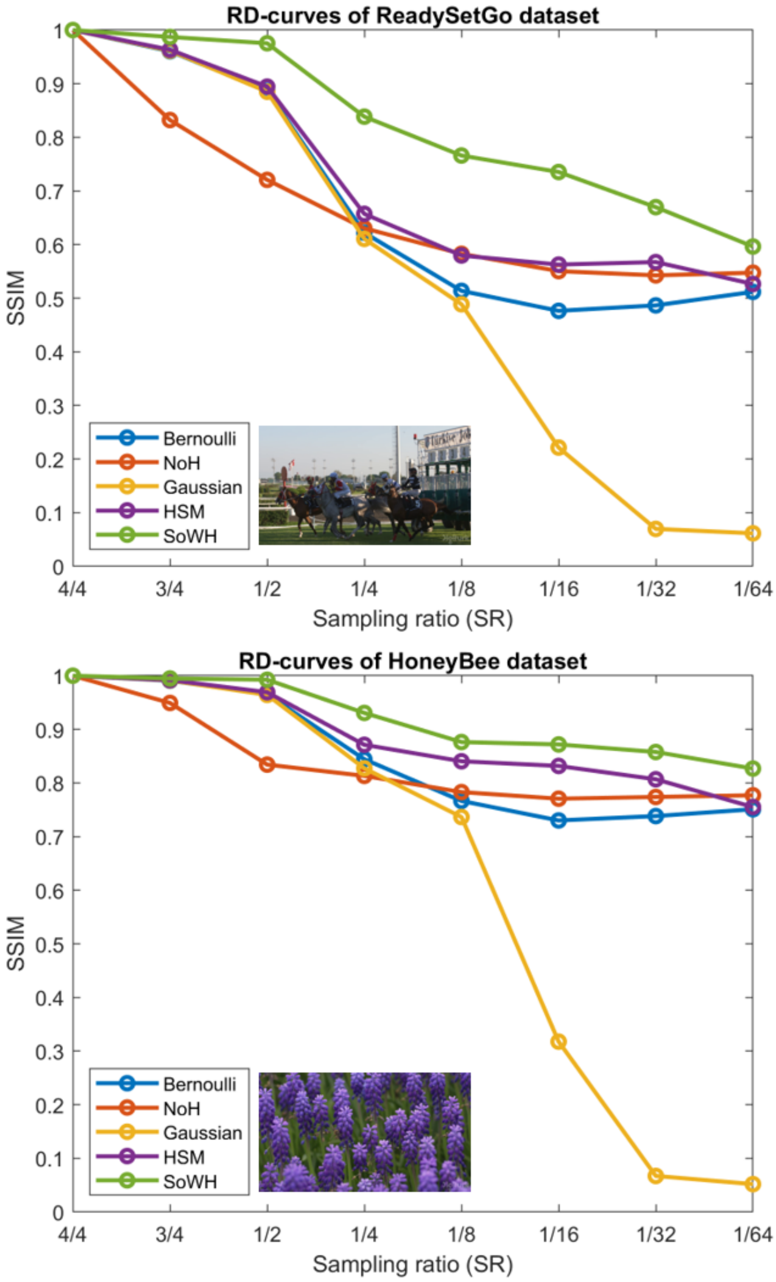

Figure 5.8: Comparing the performance of sensing matrices including Bernoulli , NoH, Gaussian, HSM, and SoWH, where SR $\in \{4/4, 3/4, 1/2, 1/4, 1/8, 1/16, 1/32, 1/64\}$

Overall quality comparison of various sensing matrices:

As the results shown in Figure 5.8, we compared SoWH with various existing sensing matrices such as Bernoulli, NoH, Gaussian, and HSM using ReadySetGo and HoneyBee datasets. It can be seen that SoWH gave the best quality in terms of SSIM than other sensing matrices at the same SR, reflecting a higher ability to gather compressible signals.

Overall performance comparison with state-of-the-art works:

As the results shown in Table 5.1, firstly, we compared our proposal with the works that used SQ to code measurement such as Bernoulli + SQ, NoH + SQ, and Gaussian + SQ . Our proposal overcame them in terms of higher PSNR and SSIM, and lower bpp. Nevertheless, NoH + SQ gave incredible results in bpp reduction because the data structure of measurement has highly uncorrelated. By using the equation (9), it returns a large parameter of Δ. Thus, SQ will give a huge image degradation as can be noticed by artifacts. Based on the uncontrollable performance of SQ, where the performance will be varied depending on sensing matrix. We assume that SQ is an inefficient coding method, which correspond to the opinion stated in the most recent literature. Next, we compared our proposal with state-of-the-art works that utilized measurement coding and SQ such as DPCM + SQ, SDPC + SQ, and Intra Pred. + HSM . This work significantly outperformed by reduced 10.90% of bpp, increased in PSNR and SSIM by 3.95 dB and 10.17%, respectively.

These results emphasized our opinion that compression performance can be increased by designing a good pair of measurement coding algorithm and sensing matrix. The measurement sampled by SSM has higher data structure consistency, which enabled coding algorithm to perform better, resulting in higher compression performance. Hence, the most important element in measurement will be encoded and will not be ruined by quantization, resulting in an improvement of PSNR and SSIM. Lastly, we provided visual quality comparison of reconstructed images in Figure 5.9. This work provided better image quality than state-of-the-art works without compression artifacts at the edge of object. Subsequently, we reported RD-curve performance in various setting of Q_b as shown in Figure 5.10. It can be seen that this work gave a remarkable

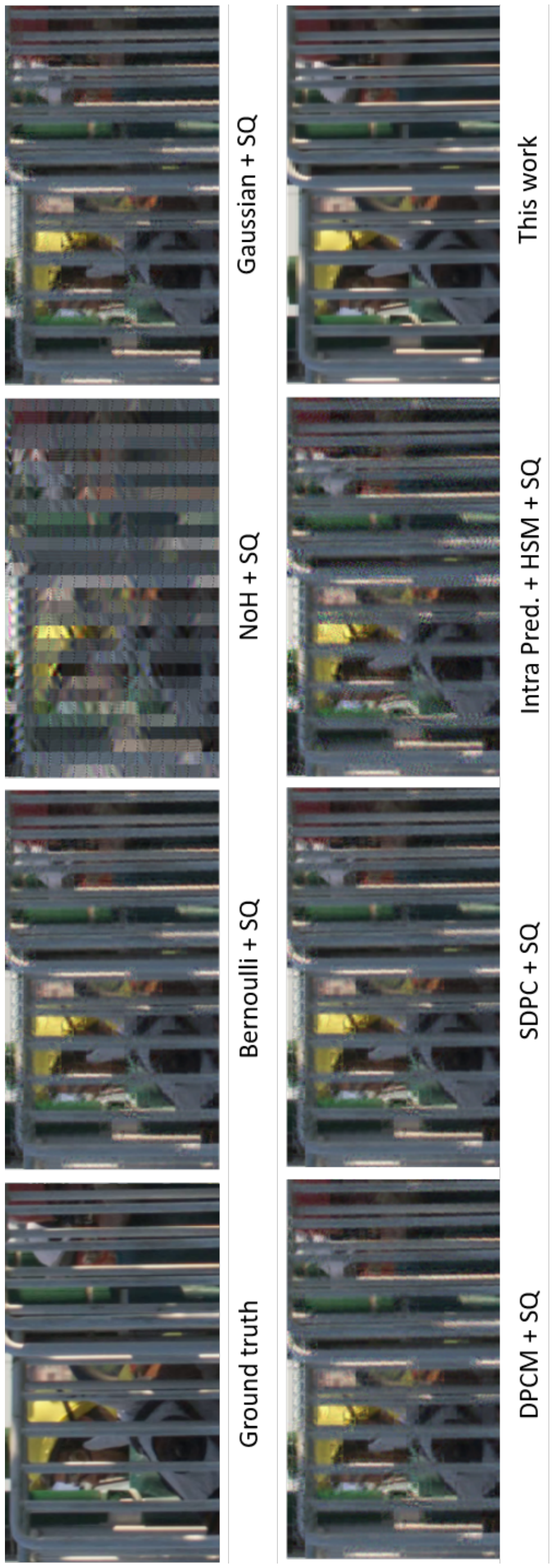

Figure 5.9: An example of 400% cropped quality comparison of ReadySetGo dataset consisted of Bernoulli + SQ, NoH + SQ, Gaussian + SQ, DPCM + SQ, SDPC + SQ, Intra Pred. + HSM + SQ , and this work

coding performance, where the data were encoded and not ruined by quantization even at shallow Q_b.

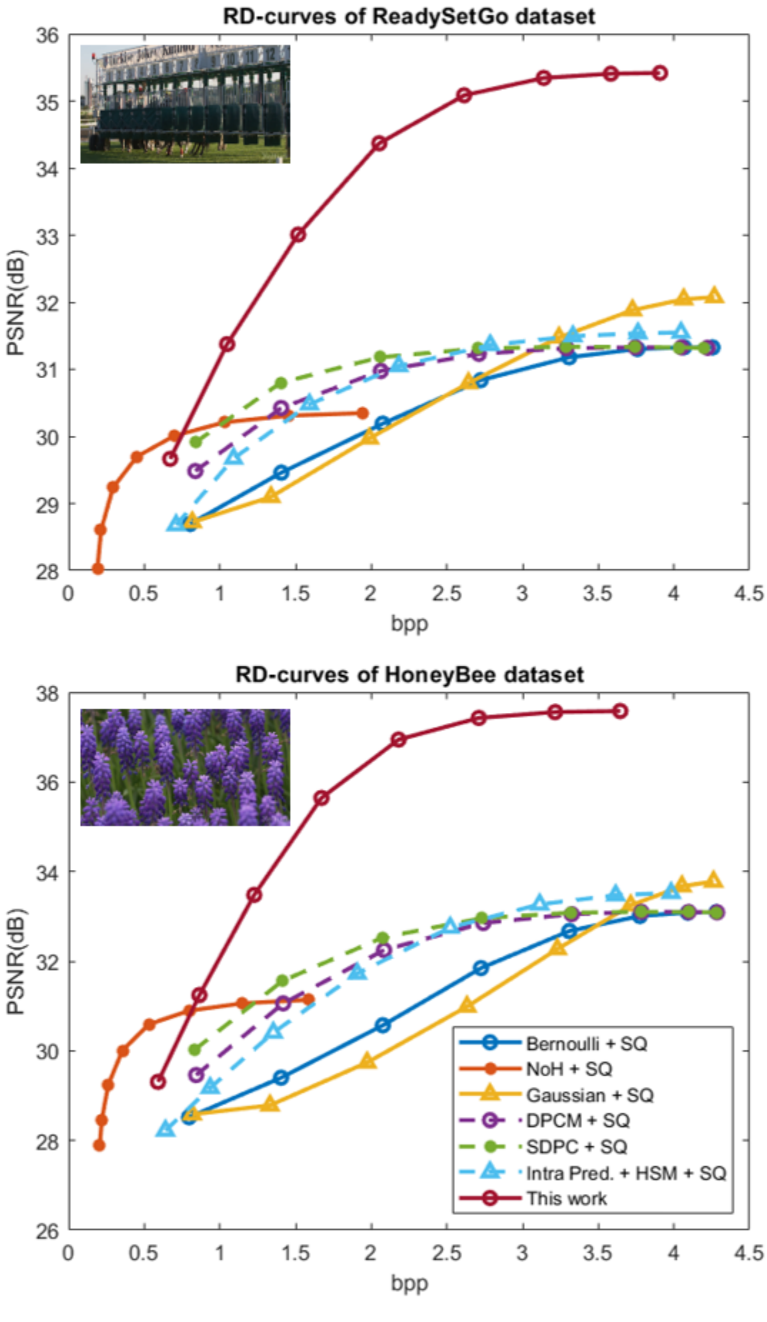

Figure 5.10: RD-curve comparison of Bernoulli + SQ, NoH + SQ, Gaussian + SQ, DPCM + SQ, SDPC + SQ, Intra Pred. + HSM + SQ, and this work using ReadySetGo and HoneyBee datasets, where each point represent to $Q_b \in \{1, 2, 3, 4, 5, 6, 7, 8\}$

Results of hardware implementation:

we reported hardware specification of IPMC algorithm in Table 5.2. The full block diagram and schematic in the Altera Quartus tools is located in Figure 6.5. The IPMC algorithm consumed total logic utilization by 5,948/41,910 logic elements and total registers by 2,138. The throughput of this algorithm is 5 Gpixels/s and operated at 88 MHz. This architecture cost the power of 50 mW for encoding 3840 $\times$ 2160 pixels, where the SR is fixed at 1/4. Lastly, we provided top-level timing diagram of BCIS and IPMC algorithm in Figure 5.12.

Table 5.2: Hardware implementation summary

Evaluation device	Altera Cyclone V
Technology	TSMC 28nm low power
ADC resolution	16 bits
Resolution	3840 $\times$2160
Maximum frequency	88 MHz
Logic utilization/ Total logic gates	5,948/41,910
Total registers	2,138
Sampling ratio	1/4
Streaming package length	2.0736 Mbits/frame
Throughput	5 Gpixels/s
Core power consumption	50 mW

5.6 Summary

BCIS is an innovative approach, in which turned conventional image and video system upside down. In this work, we closed the gap of compression performance between these BCIS and conventional systems. Our proposal capitalizes on a good pair of sensing matrix and the IPMC algorithm, which gave an extra compression performance to the traditional CS paradigm. Further, our proposal gave the highest compression performance compared to state-of-the-art works, which gradually closing the possibility gap to replace image and video acquisition system with BCIS and novel coding algorithm.

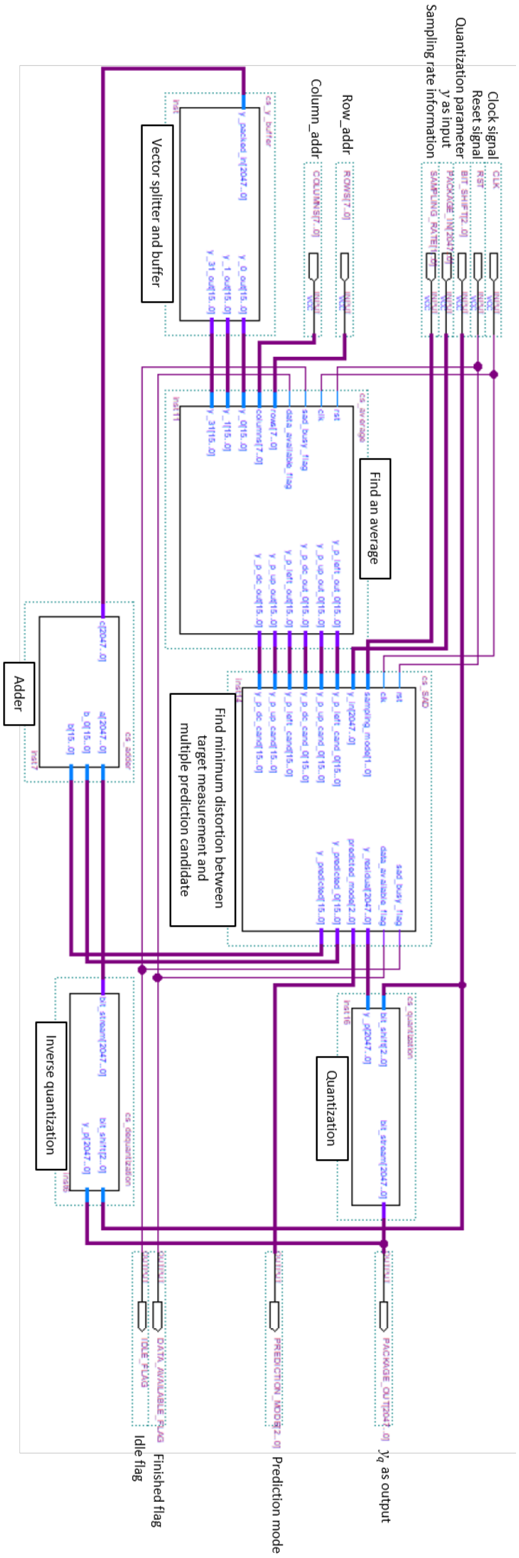

Figure 5.11: Block Diagram/Schematic of the hardware implementation in the Altera Quartus tools.

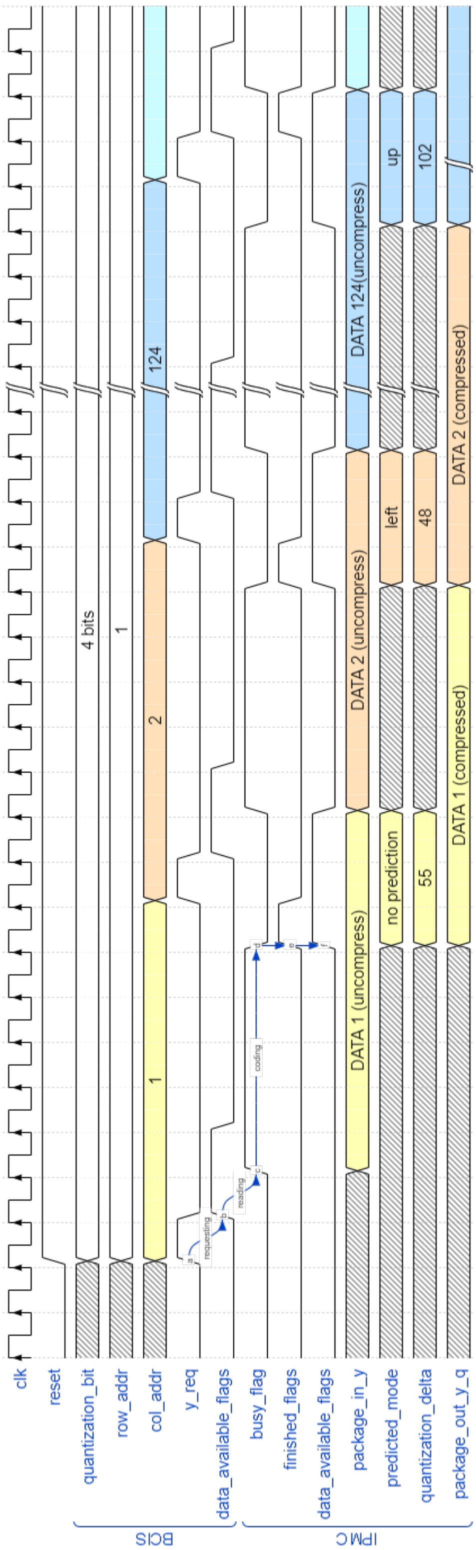

Figure 5.12: The timing diagram of hardware implementation between IPMC algorithm and BCIS including communication signals.

6

Temporal Redundancy Reduction in Compressive Video Sensing by using Moving Detection and Inter-Coding

In wireless surveillance camera (WSC) application, the scene could be unchanging for hours. Thus, wasting power and bandwidth consumption. Temporal redundancy has become primary concern to reduce power and bandwidth consumption. However, existing temporal redundancy reduction algorithm has high computational complexity. By changing the sampling method on CMOS image sensor (CIS) could reduce computation complexity, power, and bandwidth consumption using compressive sampling (CS). It is signal acquisition technique for simultaneous data sampling and compressing. The theory asserts that CS promised to reduce A/D sampling rates without adversely affecting signal recoverability. Furthermore, amount of measurements is reduced during acquisition that is no additional compression algorithm. In this paper, we proposed block-based measurement coding framework by reducing temporal redundancy for compressive video sensing. The transmitter can minimize bandwidth usage by finding the temporal redundancy on measurement based on three consecutive frames differencing method. Moreover, we used local adaptive threshold to extract useful measurement while skipping redundancy on the frames. Hence, we reduce bandwidth usage by transferring only information about the change in the scene. The receiver will compensate for the skipped measurement with previous measurement on the same location of the frame. We recover a signal using 1-minimization via primal-dual (PD) interior point algorithm and transform by using inverse fast Walsh-Hadamard transforms (IFWHT) with overlapping reconstruction to reduce staircase artifact while keeping the details. We demonstrated results on six surveillance video sequences. our proposal yields trade-off results in PSNR at 38.49dB, SSIM at 0.92, and BPP at 0.31 bpp, where sampling rate equal to 1/2.

6.1 Introduction

Compressive sensing is a signal acquisition and compression technique that allows for the efficient acquisition and reconstruction of a signal from a small number of measurements obtained by linear projections onto a sparse signal. Many works used intra prediction-based measurement coding to further compress the measurements. In this paper, we proposed using moving detection and inter-coding to reduce temporal redundancy in compressive video sensing. Firstly, the moving detection is performed coding area extraction using local adaptive threshold to classify the measurement with an association of error distinction. However, false positive detection could be occurred randomly, which transmission cost can be increasing uncertainty. In order to reduce transmission cost, the adaptive quantization parameters are adjusted by how frequently the area is detected. Moreover, we further compress the detected area by encoding the difference of current measurement and the best matching measurement in neighboring frames. Finally, an efficient recovery algorithm of sparse signal is performed by using ℓ_1-minimization via primal-dual interior-point algorithm and reconstructed by inverse fast Walsh-Hadamard transform with horizontal kernel filter to prevent staircase artifacts simultaneously. The experimental results show that our proposal can greatly reduce bandwidth usage in terms of BPP by 63.15%, improve in PSNR by 1.56dB, and SSIM by 14.81% on average when compared to the state-of-the-art.

6.2 Propose temporal redundancy reduction by using moving detection and inter-coding algorithm

Consecutive frames' information is usually similar, resulting in temporal redundancy. To further reduce the complexity, background subtraction is the most straightforward method for analyzing temporal variability and extracting the moving object of a pixel. However, because data is represented as a compressed vector of pixels domain, it is impossible to directly apply existing classical techniques.

System overview

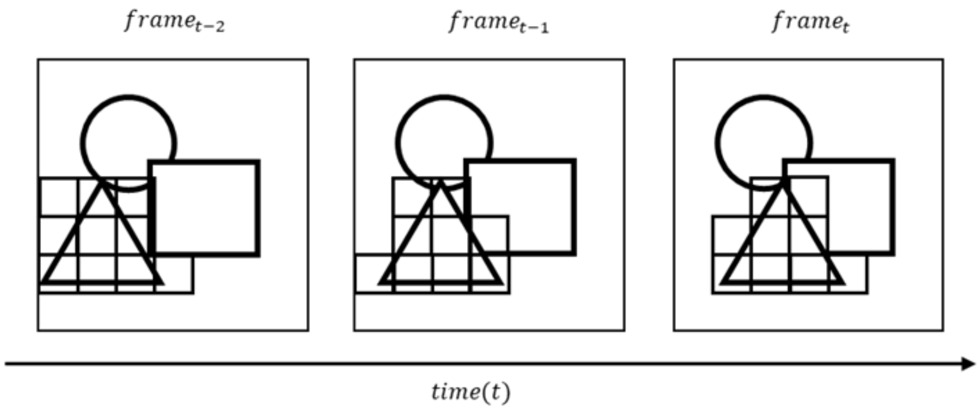

Figure 6.1: Three consecutive frames with static and region of moving components

We divided the measurement type into two parts: static measurement, which is non-moving, and dynamic measurement, which is moving on the pixel domain. To begin, we must calculate the mean square error between the current measurement Y_t and the predicted measurement Y predicted (MSE). We used local adaptive thresholding to classify the measurement data with the association of error distinction independently, which resulted in a more precise classification than global thresholding. If the difference is not greater than or equal to the threshold, we can conclude that the current Y_t is a static measurement. Otherwise, dynamic measurement is used. As a result, changing Y_t causes objects to move in the pixel domain (e.g. illumination change or moving parts). Lower than the threshold will result in a background or redundant object that must be skipped. To notify the receiver that the transmitter omitted some measurement, we must include SkipFlag in the transmission packet by setting SkipFlag to "1." The receiver will then compensate for the missed measurement using non-moving parts. Otherwise, SkipFlag will be set to "0," which means no skip. To reduce bandwidth usage and packet length even further, we used 8-bit quantization to perform left and right shift on Y_t bit patterns, resulting in Y_q. Furthermore, we reduced the number of bits on data streaming required to stream and store a packet before transferring over a communication channel by using Huffman coding.

Measurement modeling

Because the image has varying environmental conditions from time to time and no ground truth is available, it is necessary to model statistical measurements on each frame. We cannot use current measurements to create a statistical measurement model for the next frame. Because measurement is a sum of pixels in compressed form, it may not always contain the exact non-moving part of the pixel domain. Because every position on a measurement is generally self-related, using an existing denoising algorithm will affect detection and reconstruction results. We used three consecutive frames to model statistical measurement in this work. By locating the smallest difference between predicted and actual measurements $Y_{predicted}$. To begin, we stacked the dequantized measurements of three consecutive frames denoted by $Y_{dequantized}$, $Y_{dequantized-1}$, and $Y_{dequantized-2}$. Second, as $Y_{candidate}$, we chose the candidate with the lowest difference between measurement in stack and $Y_{dequantized}$ via MSE. Finally, the new Y predicted can be computed by averaging the $Y_{dequantized}$ and the lowest different $Y_{candidate}$, see Algorithm 1.

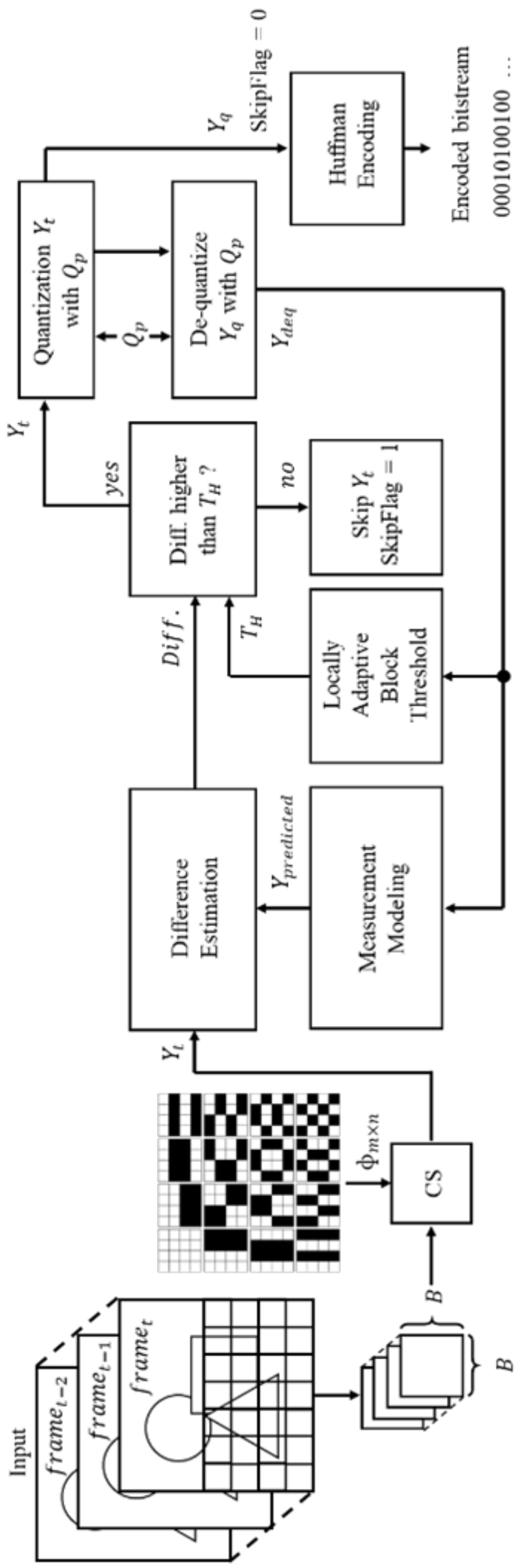

Figure 6.2: Block diagram of temporal redundancy reduction and residual decomposition

Algorithm 1: Temporal Redundancy Reduction for Residual Extraction

```
Input: rows, columns, colors, Q_p, Y_q, SkipFlag;
Output: Y_quantized and SkipFlag;
for k ← 0 to colors do
    for k ← 0 to colors do
        for i ← 0 to rows do
            for i ← 0 to rows do
                finding the difference between Y_estimate(i,j,k) and Y_t(i,j,k) via MSE
                if difference ¿ Threshold_i,j,k then
                    Y_quantized ← quantizeY_t(i,j,k) withQ_p
                    Transfering Y_quantized where SkipFlag = 0
                else
                    Break and go to the next Y_T
                end
            end
            Y_dequantized-2(i,j,k) = Y_dequantized-1(i,j,k) = Y_dequantized(i,j,k) //Stack
              swapping
            Y_dequantized(i,j,k) = dequantize Y_quantized with Q_p
            Finding lowest different between measurement in stack and Y_predicted
              through MSE as Y_candidate
            Y_predicted(i,j,k) ← average of Y_predicted(i,j,k) and Y_candidate
            Threshold(i,j,k) = mean Y_estimated(i,j,k)
        end
    end
end
```

where rows are the height of the block after it has been divided; columns are the width of the block after it has been divided; colors are color planes (default=1, grayscale); The quantization parameters are denoted by Q_p. Y t denotes current measurement, while $Y_{quantized}$ denotes Y_T after quantization. $SkipFlag$ equal to 0 is non-skip (the default), and $SkipFlag$ equal to 1 is skip.

Local Adaptive Thresholding

In order to account for variations in measurement, local adaptive thresholding has become primary method to classify the measurement. Since all of measurement in one frame contained different level of illumination and object in pixel domain. Global thresholding would be failed on adaptive environment such as climate change or camera shake caused by unstable camera mount, which is always happen in urban environment. Moreover, it also sensitive to noise and difficult to initialize threshold value. Unlike local adaptive thresholding, where threshold values are chosen independently. It is allowed BCS can operated automatically with nonuniform illumination where a global thresholding technique will not work satisfactorily. We can calculate threshold values of each block by finding an average mean of $Y_{predicted}$.

Skipped Measurement Compensation and Sparse Recovery

When $SkipFlag$ is equal to 1 and the sum of Y_q is equal to 0, the receiver will compensate for the missing block with information from the previous measurement on the same frame location. On the other hand, if $SkipFlag$ is equal to 0 and the sum of the elements of Y_t is non-zero, we can directly proceed to the optimization and reconstruction algorithm while storing patch measurement into buffer for an incoming frame. The k-sparse signal recovery is performed by a well-known solver known as l1-minimization via PD interior point algorithm. The recovery process can be very efficient in terms of computational speed per frame by using block-based CS. To avoid staircase artifacts, IFWHT has been chosen to transform the signal back into the image using overlapping reconstruction, as shown in Algorithm 2.

Algorithm 2: Missing Measurement Compensation and Sparse Recovery

Input: rows, columns, colors, Q_p, Y_q, $SkipFlag$;
Output: Image;
for $k \leftarrow 0$ **to** *colors* **do**
  for $k \leftarrow 0$ **to** *colors* **do**
    for $i \leftarrow 0$ **to** *rows* **do**
      if $SkipFlag == 1$ **then**
        $Y'_{(i,j,k)} = \hat{Y}'_{(i,j,k)}$
      else
        $\hat{Y}'_{(i,j,k)} = Y_{quantized_{(i,j,k)}}$ dequantization by bitwise shift with Q_p
        $\hat{Y}'_{(i,j,k)} = Y'_{(i,j,k)}$
      end
    end
  end
end

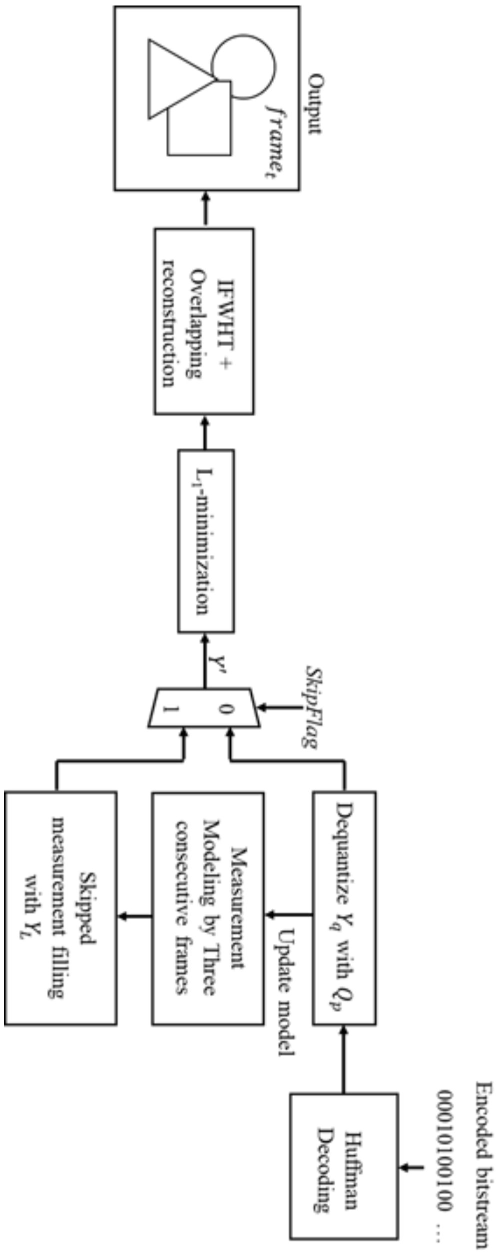

Figure 6.3: Block diagram of temporal redundancy reduction and residual decomposition

6.3 Experimental results

We demonstrated a block-based measurement coding framework for compressive video sensing by reducing temporal redundancy using 8-bit six surveillance video sequences with three crowded and uncrowded people. We sampled an entire scene by dividing it into non-overlapping blocks and sampling with the same operator. To classify the change in measurement, we used local adaptive thresholding. We used 8-bit quantization on bit patterns of measurement to further reduce bandwidth usage and packet length. Furthermore, we reduced the number of bits on data streaming required to stream and store a packet before transferring to the receiver using Huffman coding. Spare recovery is performed using ℓ_1-minimization and IFWHT transforms with overlapping reconstruction to reduce staircase artifacts while preserving details.

6.4 Summary

We demonstrated that using moving detection and inter-coding for compressive video sensing can further reduce temporal redundancy. Our proposal is fast in restoration, has good visual qualities, and significantly lowers BPP. According to the experimental results, the encoder can detect moving objects and recover the test scenes accurately. The initial frame and environmental noise had no effect on the performance of moving detection. Furthermore, by compressing the detected area further, we can significantly reduce bandwidth usage. For a perceptual metric that quantifies image quality degradation, the coding efficiency and performance are measured in terms of PSNR, BPP, and SSIM. This proposed method can reduce sampling costs and reduce communication and storage burdens while achieving comparable estimation performance, resulting in a simple approach against bandwidth usages.

Figure 6.4: Simulation results of proposed temporal redundancy reduction with crowded of people by using VIRAT, AVL-Town Centre, and WILDTRACK sequences, respectively. The first row is the ground truth. The second row is residual results of moving detection in pixel-domain. The third row is residual motion in pixel-domain. The fourth row is fully reconstruction results. The fifth row is crop results to show the remaining details, where $b \times b = 16 \times 16$, $Q_p \in \{2, 4, 6, 8\}$, and SR = 1/2

Figure 6.5: The visual comparison among three methods with our proposal using WILDTRACK sequence. The top row is the original scene. The second row is residual motion in pixel-domain. The third row is fully reconstruction results in pixel-domain. The fourth row is cropped and zoomed results for comparing the remaining details, where $b \times b = 16 \times 16$, $Q_p \in \{\ 2,\ 4,\ 6,\ 8\}$, and SR = 1/2

7

Multiple Candidates Based Hybrid Hierarchical Search for Compacting Compressively Sensed Video

Compressive sensing is a simultaneously signal acquisition and compression technique for efficiently acquiring and reconstructing a signal from a small number of measurements, which can be obtained by linear projections onto sparse signal. In recent years, spatial and temporal redundancy in measurement has become a primary concern. In this paper, we proposed hybrid three levels hierarchical search block matching algorithm with multiple vector candidates to reduce bit-rate consumption and forbidden computational cost. The multiple vector candidates technique can increase match percentage and reduce erroneous match at highest level. Moreover, we introduce termination mechanism to stop further search in vector as soon as it might be led to erroneous in lower levels. The experimental results show that our proposal can greatly reduce bandwidth usage in terms of BPP by 64.15%, improve in PSNR by 0.48dB, and SSIM by 1.16% on average when compared to the state-of-the-art works.

7.1 Introduction

The conventional method that we used nowadays for signal acquisition called Nyquist Shannon sampling theorem is inefficient because the signals of interest contain only a small number of significant frequencies relative to the bandlimited. Therefore, we need to oversample underdetermined signals in a large portion in the first place, reduce redundancy of minimal process data and represent in more compact format such as JPEG, resulting in high redundancy of compression algorithm, implementation complexity, and memory consumption. Compressed sensing or compressive sensing (CS) is a signal acquisition technique for efficiently acquiring and reconstructing a signal in mathematically sophisticated approach, by finding solutions to underdetermined linear systems ([27]). Noted that, CS is not a certain data compression methodology in information theory perspective, but it is dimensionality reduction in the first place we sample target signal. CS-based CMOS image sensor (CS-CIS) has gained significant interest in the past few years. It can greatly reduce analog-to-digital conversion circuits (ADCs) sampling rate, on-chip processing unit complexity, storage requirement, and ready to transmit right from sensors compared to conventional signal acquisition method using Nyquist-Shannon rate. CS-CIS performs acquisition and compression simultaneously on focal plane and transfers all heavy computation burden components to decoder, where the measurement streams can be processed and analyzed with unlimited resources, resulting in a low-complexity encoder ([18]). In recent years, spatial and temporal redundancy in measurement has become a primary concern, which needs to be further compressed. However, it is impossible to apply an existing multimedia compression algorithm into CS since the measurement is represented in compressed vector form, which its length can be varying. Therefore, there is a demand for compression algorithm in the measurement domain to reduce spatial and temporal redundancy, transmission cost, and power consumption. Several works introduced a coding method to further reduce spatial redundancy in measurements before transferring over communication channels while still maintain visual quality. J. Zhang, et al., have proposed a novel coding strategy for block-based compressive sensing, called spatially directional predictive coding (SDPC), which efficiently utilizes the intrinsic spatial correlation of natural images. For each block of compressive sensing measurements, its optimal prediction is selected from a set of multiple prediction candidates that are generated by four designed directional predictive modes (49). Jian bin, et al., also proposed four modes of intra-prediction with specific rows modification of random measurement matrix to gain boundary information from pixel-domain. It is allowed intra-prediction to generated predictive candidates accurately. This proposed can greatly reduce bandwidth usages when compared to other methods. However, it has degradation in image quality due to the modification of measurement matrix (118). In our previous works, we proposed hardware friendly measurement coding algorithm to further improve coding and transmission efficiency by using intra-prediction and partial Walsh-Hadamard to reduce the spatial redundancy in measurement with neighboring features. Our previous experimental results can reduce bit-rate in terms of bpp by half and improved image quality when compared to above methods (73). However, most of the existing works only focus on reducing the spatial redundancy while there is higher temporal redundancy in video transmission, in which mandatory reduction in frame-to-frame rather than being neighbor-to-neighbor manner. In this work, we proposed hybrid three levels hierarchical search block matching algorithm with multiple vector candidates (H3L-HSBM) for motion estimation and compensation to reduce an erroneous match in higher level, bitrate, and increase match percentage. Moreover, we introduce termination mechanism to stop further search in vectors as soon as it might be led to erroneous in lower levels. In such a case that some vectors have been terminated, the algorithm will dedicate computational resources to the remaining vectors. Thus, the computational resources would not be wasted and increase match percentage in the remaining vectors, resulting in hybrid algorithm compared to existing algorithms such as exhaustive search block matching (ESBM) and three levels hierarchical search block matching algorithm (3L-HSBM).

7.2 Analyze of extending the existing pixel-domain motion estimation algorithm to measurement domain

In pixel-domain, the procedure of a block-matching algorithm is to find the best matched displaced block from previous reference frame, within a designated area, for each $b \times b$ pixels in current frame. In CS camera, however, the pixel information cannot be represented individually. Moreover, measurement is sampled in block-based manner and represented in k-sparse vector length of m, which is ambitious to determine the displacement. To determine the difference of k-sparse, we adopted a straightforward method called Manhattan distance (MD) to find minimal distortion in designated area for each candidate location (u, v) area as following:

$$MD(u,v) = \sum_{i=0}^{b}\sum_{j=0}^{b} |Y_t(x+i, y+i) - Y_{t-1}(x+i+u, y+j+v)| \tag{7.1}$$

where (x, y) is the coordinate of the current block Y_t. The values of u and v are limited between $-d$ to d, which $(x + u, y + v)$ is valid candidate position. To the best of our knowledge, ESBM yields the best performance in terms of best-matched percentage and bit-rate reduction. The searching mechanism has been demonstrated in Figure 7.1.a.

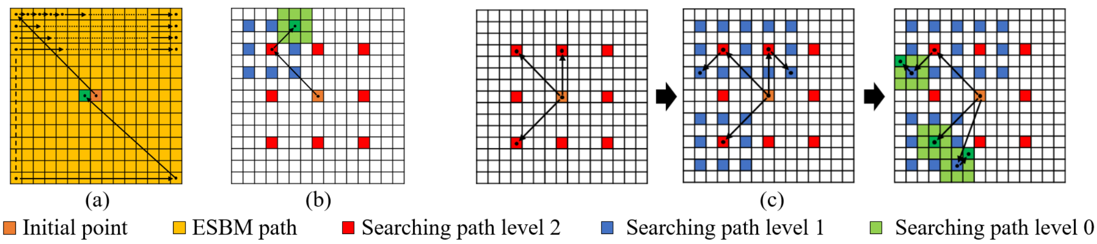

Figure 7.1: Illusions of block-matching searching path of (a) ESBM, (b) HSBM, and (c) MCH-HSBM

ESBM offered the best match candidate at 225 boxes but slow in searching process when the search range $-d, d$ is equal to 7; HSBM offered faster searching process at 25 boxes but worst in matching performance when the search range $-d, d = 7$ (highest), 4 (middle), and 1 (lowest) and searching location of each levels equal to 9 as show in 7.1.b. HSBM seems to be faster than ESBM in searching time but lower in bit-rate reduction performance. In this work, since Walsh-Hadamard sequency-order offered low frequency sampling pattern at y_1 called measurement matrix features, which is a summation of all pixels in range of $b \times b$. By using only y_1 divined by total pixel in $b \times b$ is enough to determine the displacement. Hence, we can reduce average filter in hardware level implementation. Noted that, other measurement matrices rather than Walsh-Hadamard sequency-order can use equation 7.1 with additional average filter. Therefore, our proposal hierarchical search block-matching can offer lower computational complexity and a large degree of flexibility compared to other algorithms.

7.3 Propose multiple candidates based hybrid hierarchical search for compacting compressively sensed video

Since both ESBM and HSBM have their own advantages and disadvantages. In this work, we introduced MCH-HSBM as optimal solution for motion estimation and compensation to reduce an erroneous match at higher level, bitrate, and increase match percentage simultaneously. The early termination mechanism is proposed to stop searching in branches as soon as it might be led to erroneous in lower levels. In such a case that some branches have been terminated, an algorithm will dedicate computational resources of terminated branch to help the remaining branches as shown in 7.1.c. The MCH-HSBM algorithm is proceeding as four following steps: Firstly, we assigned the common search range for three levels searching, where $-d, d = 7$ (highest), 4 (middle), and 1 (lowest) and searching location of each levels equal to 9; Secondly, we find the minimal distortion block between present frame and previous frame in designated area. The algorithm would perform multiple vectors searching to reduce erroneous of block matching, which the number of vector candidates denoted by V_c; Thirdly, we used sorting algorithm and selected the minimum V_c values to implement early termination mechanism, in which to stop searching in the path that might propagate the cause of erroneous to lower levels; Finally, the algorithm will return all possible candidate values and locations. The residual given by Y_r will be calculated by subtracting Y_t with final candidate $Y_{t_{low}}$ (u, v) multiplied with Φ to generate predictive candidate as following:

$$Y_r = Y_t(x, y) - (Y_{t_{low}}(u, v) \times \Phi) \tag{7.2}$$

However, since we implemented MCH-HSBM with the measurement that sampled in non-overlapped block-based manner. Thus, the best match candidate might be located between two candidate blocks. The first candidate by $Y_{t_{mid}}$ u_{mid}, v_{mid} and the second candidate is obtained by lowest levels given by $Y_{t_{low}}$ u_{low}, v_{low}. The final result will be judge by normalizing two candidates and find the minimum measurements as residual given by Y_r. If $Y_{t_{low}}$ is less than $Y_{t_{mid}}$, Y_r will be equal to $Y_{t_{low}}$. Otherwise, Y_r will be equal to the average of $Y_{t_{mid}}$ and $Y_{t_{low}}$ multiplied with Φ.

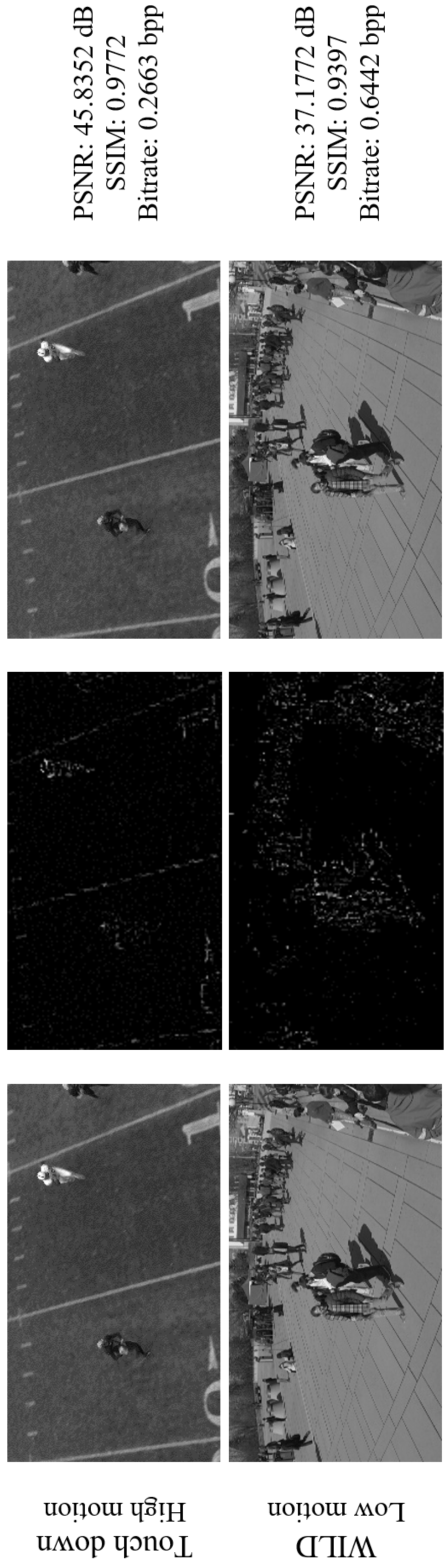

Figure 7.2: Visual comparison of proposed MCH-HSBM (a) Ground truth, (b) Residual frame, and (c) Reconstruct using Touch down and WILD sequences.

7.4 Experimental results

We demonstrated and compared our proposal with state-of-the-art works by using two video sequences consisted of Touch down for high motion test and WILD for low motion test. The video resolution is 1280×720. We adopted uniform scalar quantization (SQ) to quantize the measurement before being transfer to communication channel, in which each sample value is rounded to the nearest value from a finite set of possible quantization levels given by 2^{bits}. We fixed quantization parameter at 4-bit to serve ordinary media information. The entropy encoding can be performed by using Huffman coding. The algorithm required 16-bits overhead for each block depends on designated searching area and hierarchical level (for instance, if block size is set to be 16×16, searching area is set to be 7, and three levels of hierarchical then the overhead is equal to $16/256 = 0.0625$ bpp). Since we used y_1, which is measurement matrix feature to find the displacement and calculate predictive candidate. Therefore, it is unnecessary to store an entire frame of measurements, resulting in low memory utilization. our proposal is outperformed in both high and low motion scene.

The multiple vector candidates with early termination mechanism yield trade-off performance in total search boxes, memory consumption, and bitrate reduction. The simulation results in Table 7.1 and Table 7.2 show that MCH-HSBM is outperformed when V_c equal to 3, which it can improve PSNR by 0.48 dB, SSIM by 1.16%, and bit-rate reduction by 88.49% on average compared to state-of-the-art works.

7.5 Summary

Multiple candidates based hybrid hierarchical search (MCH-HSBM) is proposed for compressively sensed video compacting, which significantly reduces bitrate while maintaining good visual quality. Furthermore, using the measurement matrix feature can reduce hardware resources and implementation complexity.

Table 7.1: An average PSNR (dB), SSIM, and bitrate (bpp) comparisons of 100 frames using Touch down sequence

Block size	16×16	16×16	16×16
Sampling rate	50%	50%	50%
Sensing matrix	Modified DCT	Modified DCT	Walsh-Hadamard Sequency-order
Coding method	Intra	Intra	Intra
Candidate vector	-	-	-
Search vector	-	-	-
Search area	-	-	-
Search boxes	-	-	-
Quantization method	Scalar	Bit-shift	Bit-shift
Quantization bit	4	4	4
Transformation basis	IDCT	IDCT	IFWHT
Post-processing	Median filter	Horizontal kernel filter	Horizontal kernel filter
Avg. PSNR (dB)	33.74	37.76	36.94
Avg. SSIM	0.83	0.93	0.93
Avg. Bitrate (bpp)	2.84	1.99	0.98

Table 7.2: An average PSNR (dB), SSIM, and bitrate (bpp) comparisons of 100 frames using Touch down sequence (continue)

		This work	
Block size	16×16	16×16	16×16
Sampling rate	50%	50%	50%
Sensing matrix	Walsh-Hadamard Sequency-order	Walsh-Hadamard Sequency-order	Walsh-Hadamard Sequency-order
Coding method	MEMC + Inter	MEMC + Inter	MEMC + Inter
Search vector	Exhastive	1	3
Search area	16×16	16×16	16×16
Search boxes	255	25	54
Quantization method	Scalar	Scalar	Scalar
Quantization bit	4	4	4
Transformation basis	IFWHT	IFWHT	IFWHT
Post-processing	Horizontal kernel filter	Horizontal kernel filter	Horizontal kernel filter
Avg. PSNR (dB)	37.22	37.24	37.24
Avg. SSIM	0.94	0.94	0.94
Avg. Bitrate (bpp)	0.36	0.41	0.36

8

Measurement Coding Framework for High-Resolution Compressive Imaging

Compressive imaging system employs a simultaneous sensing and compressing scheme to provide a novel imaging system for multimedia. It enables high-resolution image sensors to perform better in terms of lower read-out latency and lower power consumption of analog-to-digital converters, which dominate the power of sensing systems. This system digital output is not pixels as in traditional imaging systems, but rather a compressed vector called measurements. There are two major challenges in designing a compressive imaging system. To begin, the sampling rate has an effect on the reconstructed image quality, and most existing sensing matrices cannot produce good image quality at very low sampling rates. Second, there is redundant information in RAW measurements because the sensor measures the same pixel several times to generate output. In this paper, we propose a measurement coding framework comprised of intra-inter coding, quantization, and entropy coding to improve RAW measurement compression performance. We test the performance on a variety of 4K datasets. We improved image quality in PNSR and SSIM by 8.5 and 5.5 %, respectively, in experimental results. Furthermore, we reduced bpp by 40% when compared to previous works. This proposal provided an efficient sensing matrix and high compression performance, allowing us to revolutionize the next generation of image/video acquisition systems.

8.1 Introduction

Modern image sensors capture the light and output as RAW pixel data. With higher and higher resolution requirement, the throughput can significantly increase by a large number of pixels. To reduce the data size, image and video coding algorithms are mandatory such as JPEG/JPEG2000, JPEG-XR, JPEG-XS, advanced video coding (AVC), high-efficiency video coding (HEVC), and versatile video coding (VVC). However, the complexity can be reached up to 10 times higher on each generation in software and hardware implementation. All data compression algorithms relying on sparsity by reducing spatial and temporal redundancy, whereby a signal is compressed more efficiently in terms of the sparse vector of coefficient using a generic transform basis such as discrete cosine transform (DCT)—a variant of Fourier transforms or discrete wavelet transform (DWT), which is lossy. Thus, we all question if we will perform lossy compression plainly to satisfy low bit-rate.

8.2 Compressive sensing theorem

Recently, the fundamental advances in mathematics have turned the conventional paradigm upside down. Instead of collecting high redundant pixel data and then compressing, it is now possible to acquire lower dimensional pixel data and then solve the sparest signal consistent, resulting in higher dimensional pixel data. This paradigm called “compressed sensing” or ”compressive sensing” (CS) ([18]).

CS is built upon two major fundamental conditions consisted of sparsity and incoherent ([98]) ([19]). There are essential conditions in order to apply CS to specific applications. First, a signal characteristic must be sparse when expressed in a specific orthonormal transform basis. This characteristic implies that only a few coefficients would contain the majority of the signal information. It can be expressed by:

$$x = \Psi\theta \tag{8.1}$$

where $x \in \mathbb{R}^{n\times 1}$ is the vectorized signal and $\theta \in \mathbb{R}^{n\times 1}$ is the sparse vector that contains the projection of x in the basis $\Psi \in \mathbb{R}^{n\times n}$. The theory asserts that incoherence is key associated to the quality of measurements and how the signal is sampled. Therefore, the sensing matrix $\Phi \in \mathbb{R}^{m\times n}$ used to sample x must have low coherence with Ψ. The theory asserts that Φ can be obtained by various random distribution such as

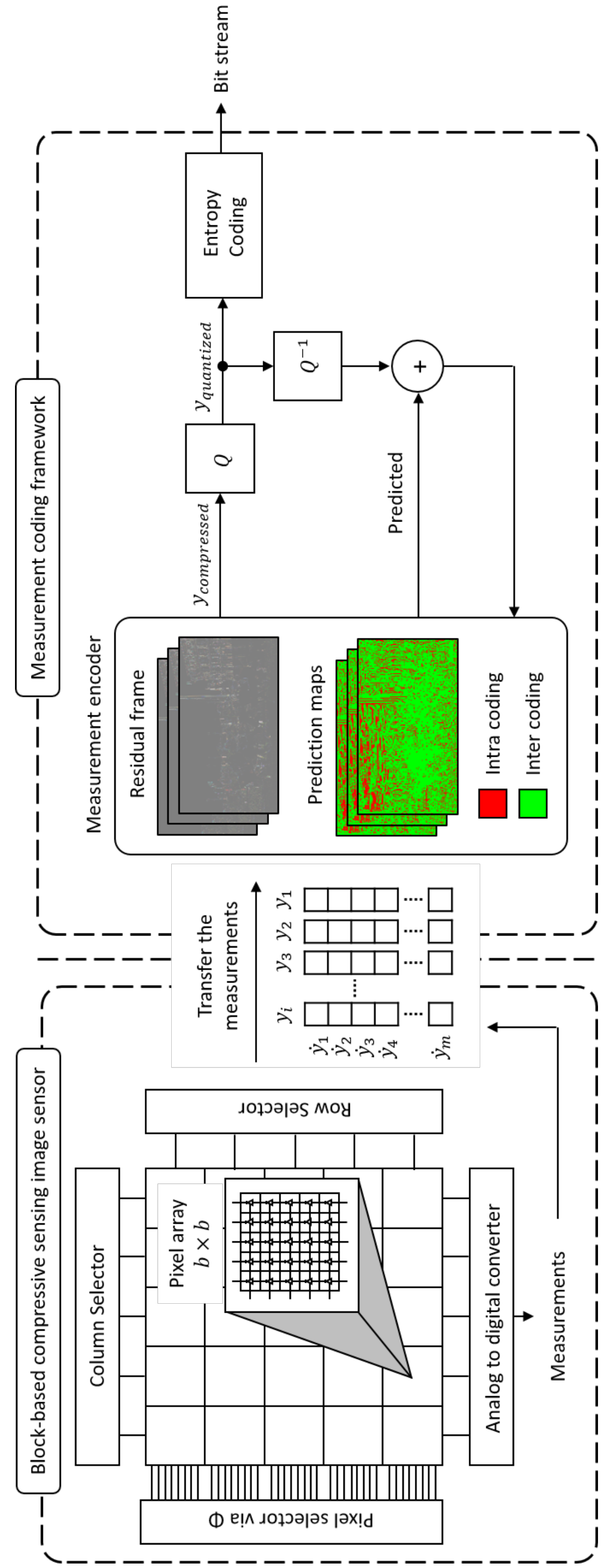

Figure 8.1: Block diagram of propose measurement coding framework for BCI, where Q and Q^{-1} are quantization and inverse quantization, respectively.

Bernoulli, Gaussian called random sensing matrix (RSM). However, the idea of using RSM in multimedia has been drawbacks because it generates a whole new sensing matrix for each frame, resulting in uncontrollable quality.

We obtain measurements $y \in \mathbb{R}^{m \times 1}$ by using underdetermined system project to x, where the dimensional of m must be less than n, known as sampling rate (SR). However, traditional CS is not suitable for a large scale problem due to long sensing time and optimizing time. In (36) proposed alternative projection technique to traditional CS by dividing an entire data frame into multiple non-overlapping blocks, where now $n = b \times b$ and b is block size. Hence, each block will sample with smaller Φ. It can be represented mathematically as:

$$y_i = \Phi x_i \tag{8.2}$$

The measurement $y_i \in \{\dot{y}_1, \dot{y}_2, \dot{y}_3, ..., \dot{y}_m\}$ is measurement of compressible signal of $x_i \in \{\dot{x}_1, \dot{x}_2, \dot{x}_3, ..., \dot{x}_n\}$, where i is the block order through raster scanning as shown in Figure 8.2.

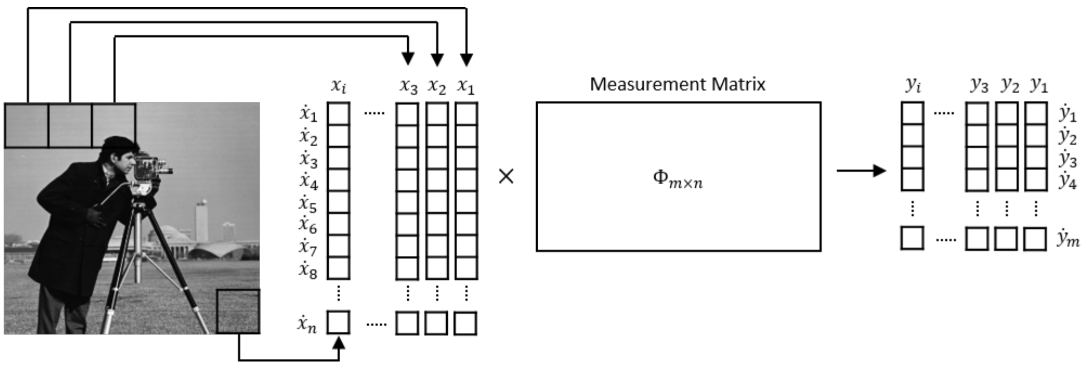

Figure 8.2: Block-based compressive sensing diagram.

To guarantee a good reconstruction, there is a hinge on a characterization of sensing matrix called restricted isometry property (RIP) (14). We can determine the lower bound dimensional of m for non-uniform distributed sensing matrix using the following equation:

$$(1 - \delta_s) \|x\|_2^2 \leq \|\Phi\|_2^2 \leq (1 - \delta_s) \|x\|_2^2 \tag{8.3}$$

where $\delta_s \in \{1, 0\}$. However, theoretically, RIP condition is a rigorous metric, which does not sound practical in some conditions of sensor design. Hence, it is challenging to use RIP condition to judge if the sensing matrix is suitable for BCI system. We found that the simplest solution is to look at reconstructed images directly. By solving ill-posed linear inverse problems via convex optimization to recover the signal, CS states that if the signal x is compressible by sparse transform Φ and Ψ is highly incoherent to Φ. Therefore, it can accurately recover from dimensional m of incomplete measurements in the coefficient domain as:

$$\hat{x} = \Phi\Psi^{-1}\Psi x \tag{8.4}$$

where Ψ^{-1} is inverse transform. y can recover by using a classical method via convex optimization called ℓ_1-minimization (12) (15) (17) as follows:

$$\hat{x} = \underset{\|\Phi x = y\| \leq \varepsilon}{\arg\min} \|x\|_{l_1}, \; where \begin{cases} y = \Phi x \\ \|w\| \leq \varepsilon \end{cases} \tag{8.5}$$

which can simplify by

$$\widehat{x} = \arg\min \frac{1}{2}\|\Phi x - y\|^2 + \lambda\|x\|_1 \tag{8.6}$$

Further, there are greedy-based recovery algorithms have been proposed for CS such as orthogonal matching pursuit (OMP) (99) and its extension stage-wise OMP (28), A*OMP (51), CoSaMP (64), and TwIST2 (8). In comparison with ℓ_1-minimization, greedy-based methods are generally faster because they take advantage of sparsity structure via minimizing a sequence of subspace problem. However, it requires higher prior-knowledge and a decent amount of measurements to make good reconstructed results, in which some applications are impracticable to obtain such information if they need to reduce SR at the sensing device frequently.

8.3 Compressive sensing based imaging technology

Several works investigated the possibility of CS-based imaging technology for multimedia and wireless surveillance system called compressive imaging (CI). By redesigning the conventional imaging system, CI provides significant energy savings as it not only cuts the RAW data size but also reduces transistor count per pixel, decreases pixel size, increases fill factor, and simplifies ADC (74) (33) (43). Further, it gives a reasonable bit-rate regarding image quality. Thus, theoretically, it does not require an additional compression algorithm. A single-pixel camera is the first image sensor that successfully developed using a digital micromirror device (DMD) (32). However, DMD-based CI required long sensing time when the resolution is relatively large. Hence, it is only suitable for microscopy and microanalysis applications ([70]). Various signal gathering technique for CI devices have been proposed such as voltage summation ([21]) and current summation ([47]) in block-based manner, which are practical for multimedia compare to single-pixel camera. Afterward, Oike and El Gamal proposed programmable block-based CI (BCI) with per-column $\Sigma\Delta$ ADCs to reduce sensing time and increase frame-rate up to 480 fps and 1920 fps when SR is equal to 1/4 and 1/16, respectively ([67]).

BCI can outperform conventional image sensors (CIS) in lower resources and lower power consumption. Nevertheless, the existing algorithms require relatively high computational resources to recover RAW measurements. Up to the present time, it may not be suitable for commercial multimedia. On the other hand, there are practical applications for BCI such as wireless surveillance system ([48]) ([1]) and optical-based on-board mission ([44]), where RAW measurements will only stream from image sensor nodes to receiver ([80]). The monitoring site (i.e., police station or ground station) is a good example that can recover RAW measurements with unlimited computational resources and power.

Moreover, BCI allows spatial and temporal relationships among adjacent and previous frame to be exploited. Each element in RAW measurement carries a similar amount of information about the signal being acquired. Hence, RAW measurements are redundant ([41]). There is an open problem on BCI for multimedia that how to store or transmit them in an efficient way?

8.4 Related works on measurement coding

Regarding CS theory, RAW measurements are reasonable to store and transmit as they sampled, where the amount of bit-rate will correspond to SR. However, in some applications, there is a retraction in dimensionality reduction. For instance, BCI for 4K resolution has a special microarchitecture with SR reduction limit of 50% to preserve high visual quality, resulting in approximately 13.27MB per frame (while conventional CIS produces 26.54MB per frame). However, the amount of bit-rate is still higher than the compressed media by pixel-based compression algorithms, which makes BCI difficult to replace traditional approaches.

To save storage space and reduce bit-rate in transmission, we state an opinion that RAW measurements require a compression algorithm, which can produce a similar factor of bit-rate reduction as using pixel-based compression algorithms.

Up to the present time, most of the CS literature has been devoted to study the recovery of sparse signals from a small number of measurements, but less in compression algorithms. By referring to the legacy vector compression algorithms, introduced for lossless compression (84), and its extension for lossy compression (109). Although, it is possible to use these legacy approaches to code RAW measurements. However, to deploy in conjunction with BCI, it requires a precise design for each system specifically, which is not convenient.

Scalar quantization (SQ) provided a straightforward approach. However, it has been established that SQ is highly inefficient in terms of information-theoretic rate-distortion (RD) performance (123) (22). By comparing to traditional source coding, SQ gave low bit-rate utilize but high recovery failure rate. Additionally, it required an iterative recovery algorithm to predict corrupted quantized measurements such as quantized iterative hard thresholding (QIHT) (46), quantized compressed sampling matching pursuit (QCoSaMP), and adaptive outlier pursuit for quantized iterative hard thresholding (AOP-QIHT) (82). The quality degradation would emphatically appear in the reconstructed image, which can be easily detected by human visual perception as artifacts.

Differential pulse-code modulation (DPCM) was introduced to reduce bit-rate over transmission channels (62). DPCM used single previously compressed measurement to predict RAW measurement. However, the previously compressed measurement may contain irrelevant information with no relation to RAW measurement, in circumstance, resulting in unstable bit-rate reduction. Nevertheless, DPCM provided a practical method with low recovery error rate. The work in (119) used lossy discrete cosine transform (DCT) to encode RAW measurements. This approach achieved better coding efficiency than SQ and DPCM. Consequently, it had quality degradation higher than SQ and DPCM. The work from (120) proposed an alternative approach from most state-of-the-art works. They investigated in dead-zone of quantizer. They stated that quality degradation mainly came from the quantizer and SR. Hence, in their studies, The system was designed strictly included an opinion on how to choose sensing matrix, the trade-off between quantization and SR, and provided reconstruction algorithm using deep learning. They controlled the quantization step and SR to achieve near-optimal quality at any given bit-rate. Hence, their approach can gain a better balance between reconstruction quality and bpp. The directional compression scheme introduced in (118). Spatially directional predictive coding (SDPC) is an efficient measurement compression scheme that accomplished SQ and DPCM with a higher successful recovery rate and low bit-rate. However, this work used Gaussian distributed as sensing matrix, resulting in unstable quality and bit-rate utilization. Importantly, they could not embed this kind of sensing matrix to hardware that can handle only binary sensing matrix. Later, four directional intra coding with partial modification of RSM was proposed (121). They were inspired by intra coding concept of H.264/AVC in order to propose their algorithm. Subsequently, the partial modification of RSM was used to obtained specific pixels for generating prediction candidates. They embedded specific sensing patterns into RSM in order to gain boundary information while sensing called hybrid sensing matrix (HSM). Afterward, they used gained information and then multiplied it with HSM to generate prediction candidates. This work can significantly reduce bit-rate compared to state-of-the-art works. Nevertheless, it cost tremendous degradation to image quality due to the modification of sensing matrix that was not satisfied by RIP conditions and inverse transformation.

In the past few years, in (31) (78) (90) proposed an idea that RAW measurements were more compressible than those sampled by RSM and HSM, if sensing matrix made of diagonal-constant matrix such as toeplitz, vandermonde polynomial, and alternant. It was called a structural sensing matrix (SSM). Subsequently, this matrix construction gave stable image quality all the time because each block will always sample by the same sensing matrix. The work in (73) motivated by pixel-based compression algorithms, we introduced intra coding for BCI to reduce bit-rate utilization. We used SSM made of Natural ordered Hadamard (NoH) to obtain boundary information from neighboring blocks partially, producing three prediction candidates. This work could overcome bit-rate lower than state-of-the-art works with reliable image quality. Moreover, in (72), we proposed an extension by coding only a high transition vector in conjunction with inter coding. With similar technique as in (73), in this work, we used sequency ordered Walsh-Hadamard (SoWH) properties to predict redundant motion from the successive frame. We could significantly overcome bit-rate utilization lower than state-of-the-art works. However, it introduced numerous image clipping artifacts due to non-overlapping block-based sampling approach. The previous works were mainly focused on bit-rate reduction while preserving visual quality. However, we could not further reduce bit-rate and SR lower than 50% due to the unsatisfactory image quality. Although, we added five more prediction candidates into intra coding in (73). However, reported in (94), it could not provide significant improvement in compression performances compared to (73) (72), which averagely better by 1-2%. Therefore, our previous methods reached the limit of bit-rate reduction.

8.5 Propose measurement coding for high-resolution compressive imaging

In this work, we study the connected components among adjacent sensing pattern, in which those RAW measurements sampled by high connected components sensing matrix should provide higher image quality. We hypothesize that if each element in RAW measurement carries information that associates each other, it should help recovery algorithm to estimate the original signal precisely, even sampling at lowest bound of SR. Therefore, we introduce a new sensing matrix that improves the reconstructed image accordingly. Subsequently, we introduce a compression algorithm to compress RAW measurements, resulting in lower bit-rate than SR control. The overall architecture of the proposal can be seen in Fig 8.1. There are three primary signals that control BCI, including column selector, row selector, and pixel selector. Multiple pixels are sum and read-out by analog-to-digital converter several times regarding SR, resulting in each element of RAW measurement. Afterward, BCI transfers RAW measurements to measurement coding framework for further compressing. We compress RAW measurements using intra-inter coding with multiple candidates from neighboring blocks. Furthermore, we reduce the probability symbol by further quantizing compressed measurements. Besides, we include inverse quantization inside the encoder to estimate information for decoding at the receiver. Without estimating information between both sides, the decoder will act as error accumulator, in which a single corrupt decoded measurement can initiate recovery error to the whole image. Nevertheless, this framework allows users to skip quantization for lossless storing and transmitting. Moreover, before storing or transmitting, compressed measurements can further reduce bit-rate using entropy coding such as Huffman Coding or Arithmetic Coding.

The measurement encoder receives RAW measurement from the image sensor. RAW measurements sampled by SSM show an identical significant correlation of data spatially and temporally, which are highly compressible than RAW measurements sampled by RSM. Therefore, we benefit from those identical amounts of information to deliver intra-inter coding by estimating from the neighboring blocks. We note that our scheme will not conduct in pixel-to-pixel but vector-to-vector due to the data structure of RAW measurements. The prediction candidate of intra coding can select by finding minimum difference among the surrounding neighbor measurements on the same frame. On the other hand, the prediction candidate of inter coding can obtain by finding minimum difference with multiple candidates from previous frame, as shown in Fig 8.3. An

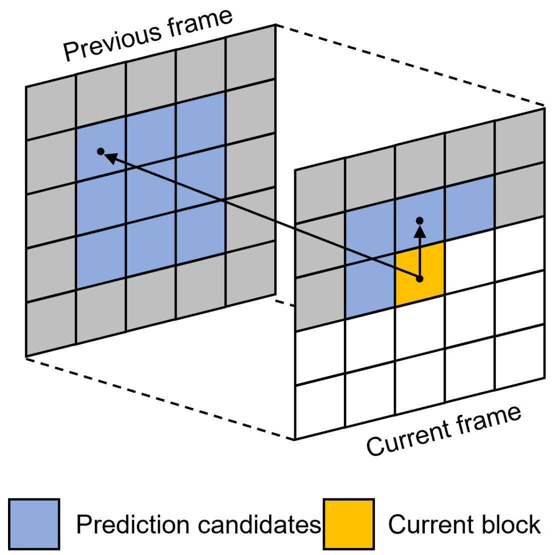

Figure 8.3: Diagram of prediction candidate positions in current frame and previous frame for intra-inter coding, respectively.

intra coding candidates are locate in 4 directions as shown in Fig 8.4a, in which a set of candidates can obtain by $y_{intra} \in \{\, y_a,\, y_b,\, y_c,\, y_d\}$. The inter coding candidates are locate in previous frame, as shown in Fig 8.4b, in which a set of prediction candidate can obtain from 9 directions consists of $y_{inter} \in \{\, \hat{y}_{\hat{a}},\, \hat{y}_{\hat{b}},\, \hat{y}_{\hat{c}},\, \hat{y}_{\hat{d}},\, \hat{y}_{\hat{e}},\, \hat{y}_{\hat{f}},\, \hat{y}_{\hat{g}},\, \hat{y}_{\hat{h}},\, \hat{y}_{\hat{i}}\}$. In our study, we note that when the block size $b \times b$ is relatively large, by increasing prediction candidates, it does not make any difference due to non-overlapped block partitioning.

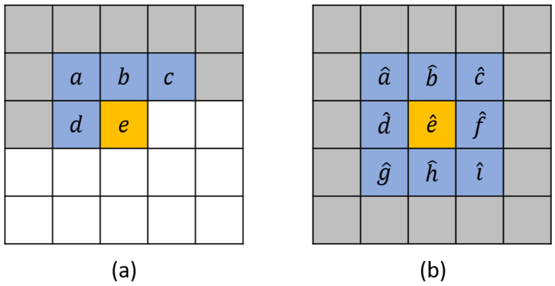

Figure 8.4: The neighbor prediction candidates that being used for (a) intra coding and (b) inter coding.

To determine prediction candidate for each mode, it can be expressed by:

$$y_{intra_{compressed}} = \|y_{intra} - y\|_{l_1} \tag{8.7}$$

$$y_{inter_{compressed}} = \|y_{inter} - y\|_{l_1} \tag{8.8}$$

where $\|*\|_{l_1}$ is normalization by adding all absolute value of entire vector and finding minimum difference among prediction candidates and current RAW measurement y. The compressed measurement of each mode can denote by $y_{intra_{compressed}}$ and $y_{inter_{compressed}}$. Afterward, we normalize $y_{intra_{compressed}}$ and $y_{inter_{compressed}}$ to find the smallest magnitude, and return as final compressed measurement $y_{compressed}$. It can be expressed by:

$$y_{compressed} = \begin{cases} y_{intra_{compressed}}, & \text{if } \left\|y_{intra_{compressed}}\right\|_1 < \\ & \quad \left\|y_{inter_{compressed}}\right\|_1 \\ y_{inter_{compressed}}, & otherwise \end{cases} \tag{8.9}$$

Nevertheless, if there is no previous frame available, or the current frame is key-frame, the encoder will only perform intra coding. Note that,$y_{compressed}$ need to store or transmit along with side information compose of 1 bit for coding mode and 4 bits for direction, which decoder will use this information to decode $y_{compressed}$. Moreover, we introduce a memory model in order to store prediction candidates for the next y. The memory model can calculate mathematically by

$$memory_{intra} = m \times \left(\frac{w}{b}\right) \tag{8.10}$$

$$memory_{inter} = m \times \frac{(w \times h)}{b^2} \tag{8.11}$$

where w and h is width and height of frame, respectively. For instance, let the frame size equal to 3840×2160, each measurement element required 16 bits. By using block-based approach where $b = 16$, for each block $b \times b$, then n will equal to 256. Let $m = 16$ for 93.75% reduction in sampling cost. For intra coding, the memory required is 7.68 KB per color channel. On the other hand, for inter coding, the memory required is 1,036.8 KB per color space. Finally, for RGB, we can estimate the total memory to prediction candidates are approximately 4 MB.

Quantization

After obtained $y_{compressed}$, applies SQ further to reduce the probability set input symbol for entropy coding, resulting in higher compression performance. SQ maps $y_{compressed}$ into a finite sequence of codewords with a quantization step size equal to Δ. It can be expressed by

$$\Delta = \left\lfloor \frac{\max\left(y_{compressed}\right) - \min\left(y_{compressed}\right)}{2^{Q_b}} \right\rfloor, \tag{8.12}$$

$$y_{quantized} = \left\lfloor \frac{y_{compressed}}{\Delta} \right\rfloor \tag{8.13}$$

where Q_b is quantization bit. The quantized measurement can be denoted by $y_{quantized}$. Subsequently, inverse SQ maps $y_{quantized}$ into $y_{dequantized}$ that is an approximation of $y_{compressed}$. It can be expressed by

$$y_{dequantized} = \Delta \cdot y_{quantized} \tag{8.14}$$

However, in general, it is difficult to determine how many bit $y_{compressed}$ need for each block since Q_b can be vary depending on parameters from $y_{compressed}$. To determine Q_b, this work used SSIM to analyze quantization effect over $y_{compressed}$ against $Q_b \in \{1, 2, 3, 4, 5, 6, 7, 8\}$ on multiple datasets. We reported an average results of each coding mode separately, where SR $= \{3/4, 1/2, 1/4, 1/8, 1/16, 1/32, 1/64\}$, $b = 16$, as shown in Fig 8.5 and Fig 8.6, respectively. As the results show in Fig 8.7, Q_b can choose flexibly depending on the acceptable loss, in which high loss also corresponds to bpp reduction. Further, quantization will give artifacts at the edge of the object. However, it is not clear

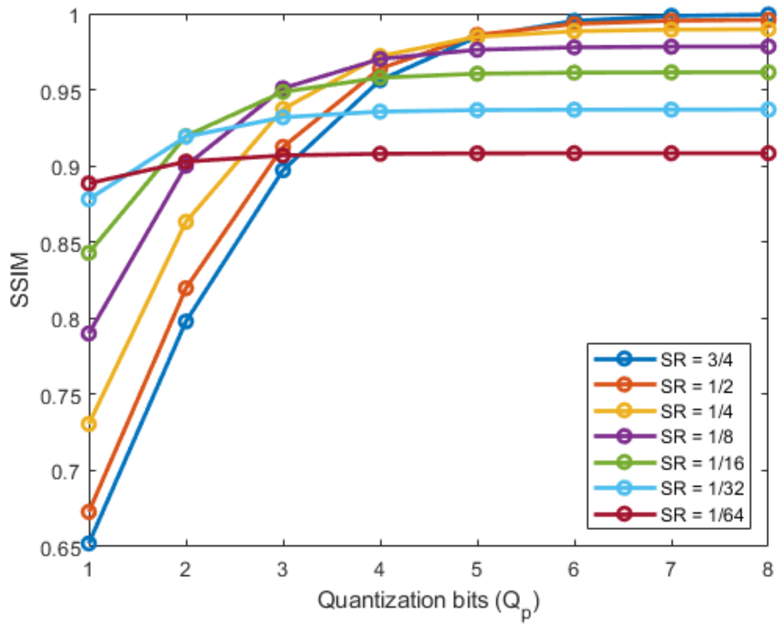

Figure 8.5: An average results of quantization effect to reconstructed image against multiple Q_b on intra coding.

how to choose the best Q_b automatically without seeing the reconstructed images. This problem has been raised recently in which the most recent developed theorems stated opinion in (4), (54), and (86) that it should be about $log_2(m)$. Lastly, this proposal allows users to skip SQ for lossless storing and transmitting. To further reduce bit-rate, $y_{quantized}$ can further compress by Huffman Coding or Arithmetic Coding.

8.6 Experimental results

In this section, we demonstrated the proposal applicability and versatility using PSNR, SSIM, and bpp as quantitative metrics. The simulation results delivered by MATLAB using l_1-minimization via primal-dual interior-point methods. We used multiple 4K datasets (**?**), including Beauty, ReadySetGo, Bosphorus, and HoneyBee. First of all, we provided extensive simulation results of quality degradation characteristics of Φ_{CoWH} varying by different SR in Fig 8.8, where $b = 16$ and SR $\in \{4/4, 1/8, 1/16, 1/32, 1/64\}$.

It can be seen that when SR equal to 1/16, Φ_{CoWH} provided the last acceptable

Table 8.1: SQ performance with various sensing matrices as baseline, where SR= 1/16, $b = 16$, and $Q_b = 4$.

Methods	Beauty dataset		ReadySetGo dataset	
	PSNR (dB)/SSIM	bpp	PSNR (dB)/SSIM	bpp
Bernoulli-SQ	34.09 / 0.78	3.9956	31.58 / 0.66	3.9958
NoH-SQ	34.66 / 0.80	3.9690	32.18 / 0.69	3.9691
SoWH-SQ	34.79 / 0.81	3.9689	32.96 / 0.75	3.9687
CoWH-SQ	34.21 / 0.85	3.9679	33.02 / 0.76	3.9680
Average as baseline	34.43 / 0.81	3.9753	32.43 / 0.71	3.9754

Methods	Bosphorus dataset		HoneyBee dataset	
	PSNR (dB)/SSIM	bpp	PSNR (dB)/SSIM	bpp
Bernoulli-SQ	32.59 / 0.80	3.9959	33.25 / 0.88	3.9959
NoH-SQ	33.03 / 0.71	3.9698	33.52 / 0.86	3.9696
SoWH-SQ	32.62 / 0.69	3.9698	33.86 / 0.87	3.9695
CoWH-SQ	30.67 / 0.66	3.9698	33.44 / 0.90	3.9692
Average as baseline	32.22 / 0.71	3.9763	33.51 / 0.87	3.9760

Table 8.2: Intra coding performance comparison with state-of-the-art works consist of DPCM-SQ ([62]), SDPC-SQ ([116]), In.4M.Pred.-HSM-SQ , In.3M.Pred.-NoH-SQ ([73]), and In.9M.Pred.-SoWH-SQ ([94]), where SR= 1/16, $b = 16$, and $Q_b = 4$.

Methods	Beauty dataset		ReadySetGo dataset	
	PSNR (dB)/SSIM	bpp	PSNR (dB)/SSIM	bpp
DPCM-SQ	34.77 / 0.79	1.4653	34.35 / 0.85	1.8434
SDPC-SQ	34.53 / 0.79	1.2386	34.06 / 0.82	1.5244
In.4M.Pred.-HSM-SQ	34.77 / 0.79	0.9863	36.04 / 0.92	1.2159
In.3M.Pred.-NoH-SQ	35.85 / 0.84	0.9202	36.04 / 0.92	1.1497
In.9M.Pred.-SoWH-SQ	37.13 / 0.86	0.6591	35.79 / 0.90	0.9104
This work-SQ	38.41 / 0.89	0.3781	35.54 / 0.89	0.6711

Methods	Bosphorus dataset		HoneyBee dataset	
	PSNR (dB)/SSIM	bpp	PSNR (dB)/SSIM	bpp
DPCM-SQ	35.24 / 0.93	1.7356	34.80 / 0.92	1.9188
SDPC-SQ	34.37 / 0.90	1.5224	34.44 / 0.92	1.7584
In.4M.Pred.-HSM-SQ	35.42 / 0.93	1.2813	34.83 / 0.92	1.5004
In.3M.Pred.-NoH-SQ	39.13 / 0.97	1.2995	35.75 / 0.94	1.4952
In.9M.Pred.-SoWH-SQ	40.37 / 0.94	0.9498	37.94 / 0.96	0.9644
This work-SQ	41.61 / 0.92	0.6001	40.14 / 0.98	0.4336

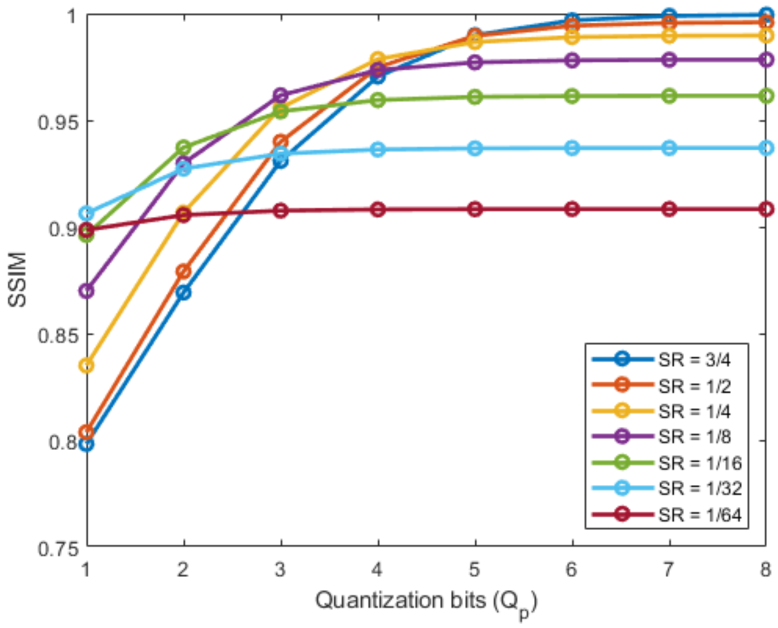

Figure 8.6: An average results of quantization effect to reconstructed image against multiple Q_b on inter coding.

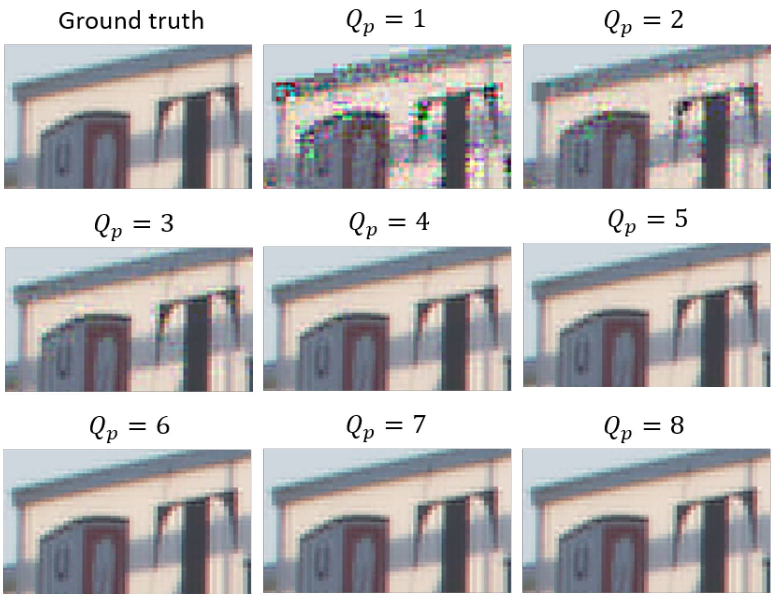

Figure 8.7: An example of quantization effect to reconstructed image against Q_b using ReadySetGo dataset.

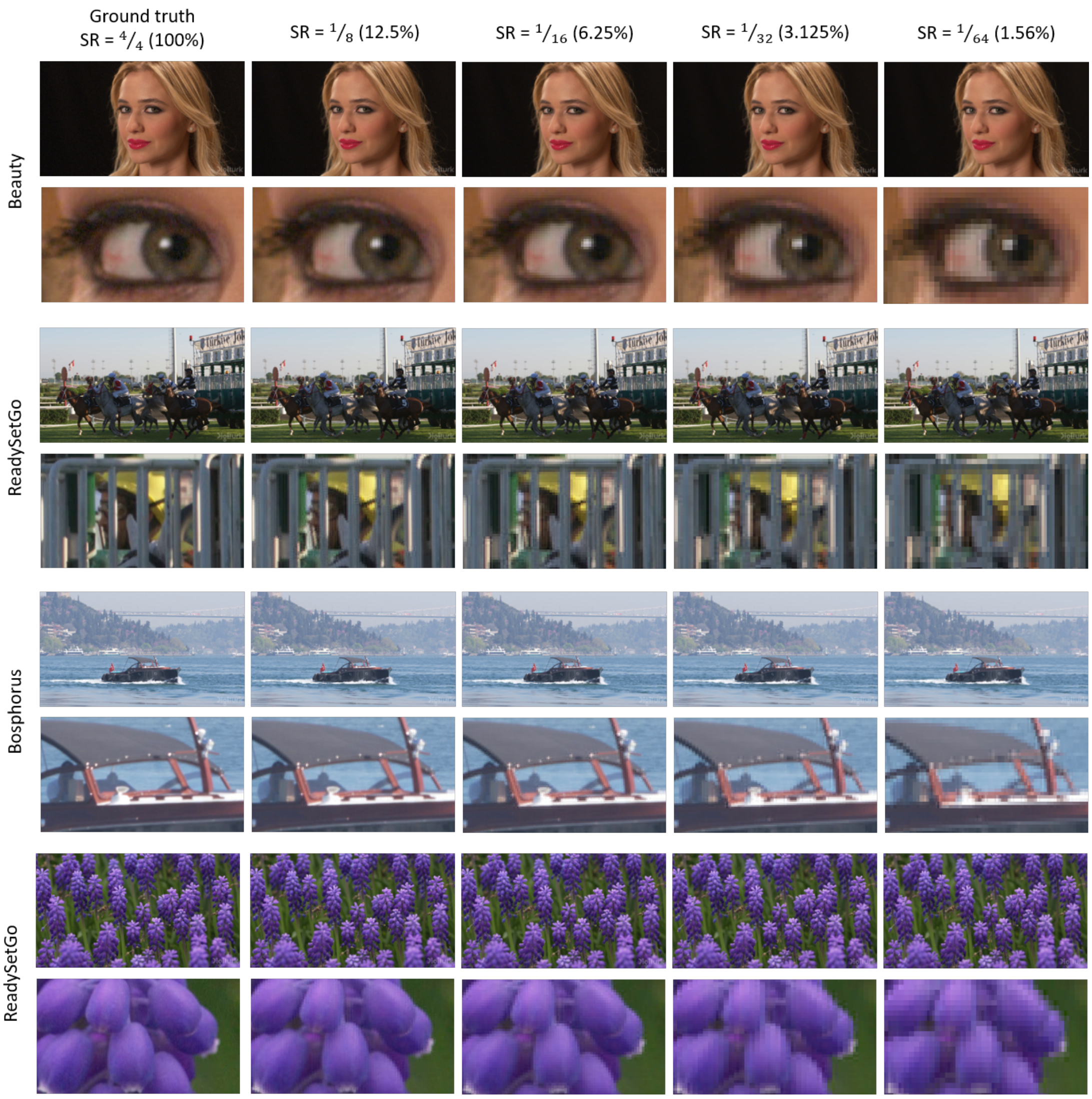

Figure 8.8: An example of quality degradation characteristics of Φ_{CoWH} with 400% zoomed results.

quality with less undersampling artifacts compared to 1/32 and 1/64. Therefore, This proposal fixed SR at 1/16 to evaluate our proposal and compared it with state-of-the-art works. This work used reproducible state-of-the-art codes. However, this proposal adapted their quantization method to be SQ with a fixed Q_b at 4 bits for general quantization.

Defining baseline with SQ

As the results show in Table 8.1, Φ_{CoWH} overcame other sensing matrices such as $\Phi_{Bernoulli}$, Φ_{NoH}, and Φ_{SoWH}. Φ_{CoWH} gave better PSNR than $\Phi_{Bernoulli}$. However, by comparing to Φ_{NoH} and Φ_{SoWH}, Φ_{CoWH} produced slightly lower PSNR but higher in SSIM. Besides, Φ_{CoWH} gave the lowest bpp compare to the other sensing matrices. Nevertheless, it seems that only SQ could not satisfy bpp reduction, which corresponds to our stated opinion at the beginning that RAW measurements need other methods to compress RAW measurement. At this stage, SQ is undoubtedly inefficient approach for bpp reduction. Therefore, we used average SQ performance among various sensing matrices as the baseline for each dataset. In which any measurement coding algorithm must perform better than baseline.

Overall intra coding comparison with state-of-the-art works.

By comparing with baseline in Table 8.2, our proposal gave remarkable bpp reduction averagely by 10 times lower with higher quality, which can notice by PSNR and SSIM. By comparing to exiting DPCM-SQ and SDPC-SQ, which applied directional candidates to encode the measurement. However, both DPCM-SQ and SDPC-SQ implemented based on $\Phi_{bernoulli}$. They could not control the consistency of the information contained inside the measurement, resulting in less ability to reduce bpp. By statistical comparison in bpp reduction, our proposal accomplished DPCM-SQ and SDPC-SQ averagely by 70% and 65% in bpp reduction, respectively. Additionally, by comparing to the advanced scheme in sensing matrix and compression scheme, we accomplished In.4M.Pred.-HSM-SQ averagely 58% of bpp reduction Further, by comparing with our previous works such as In.3M.Pred.-NoH-SQ and In.9M.Pred.-SoWH-SQ, this work accomplished them in averagely 57% and 40% of bpp reduction, respectively. It can be said that modifying the structure of sensing matrix in order to obtain specific information to generate prediction candidates is not efficient. Moreover, adding higher prediction candidates did not reflect to higher compression performance. Finally, we provide extensive performance comparison curves in SSIM of our proposal among SQ, baseline, and state-of-the-art works against multiple $Q_b = \{1, 2, 3, 4, 5, 6, 7, 8\}$ as shown in Fig. 8.9 and Fig. 8.10, this proposal presented highest visual quality compared to state-of-the-art works.

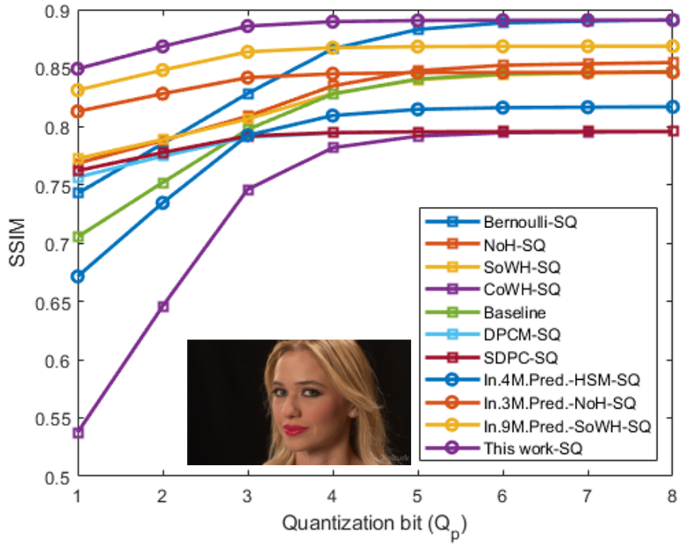

Figure 8.9: Example of quality comparison of propose measurement coding framework among state-of-the-art works using Beauty dataset.

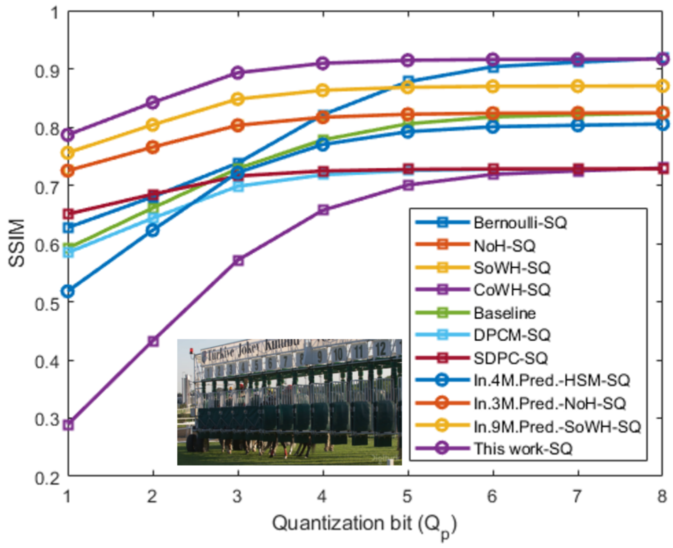

Figure 8.10: Example of quality comparison of propose measurement coding framework among state-of-the-art works using ReadySetGo dataset.

Coding efficiency of intra and inter-prediction coding

Most exiting state-of-the-art works compressed RAW measurements using intra coding and quantize with small Q_b. However, this study found that RAW measurements can

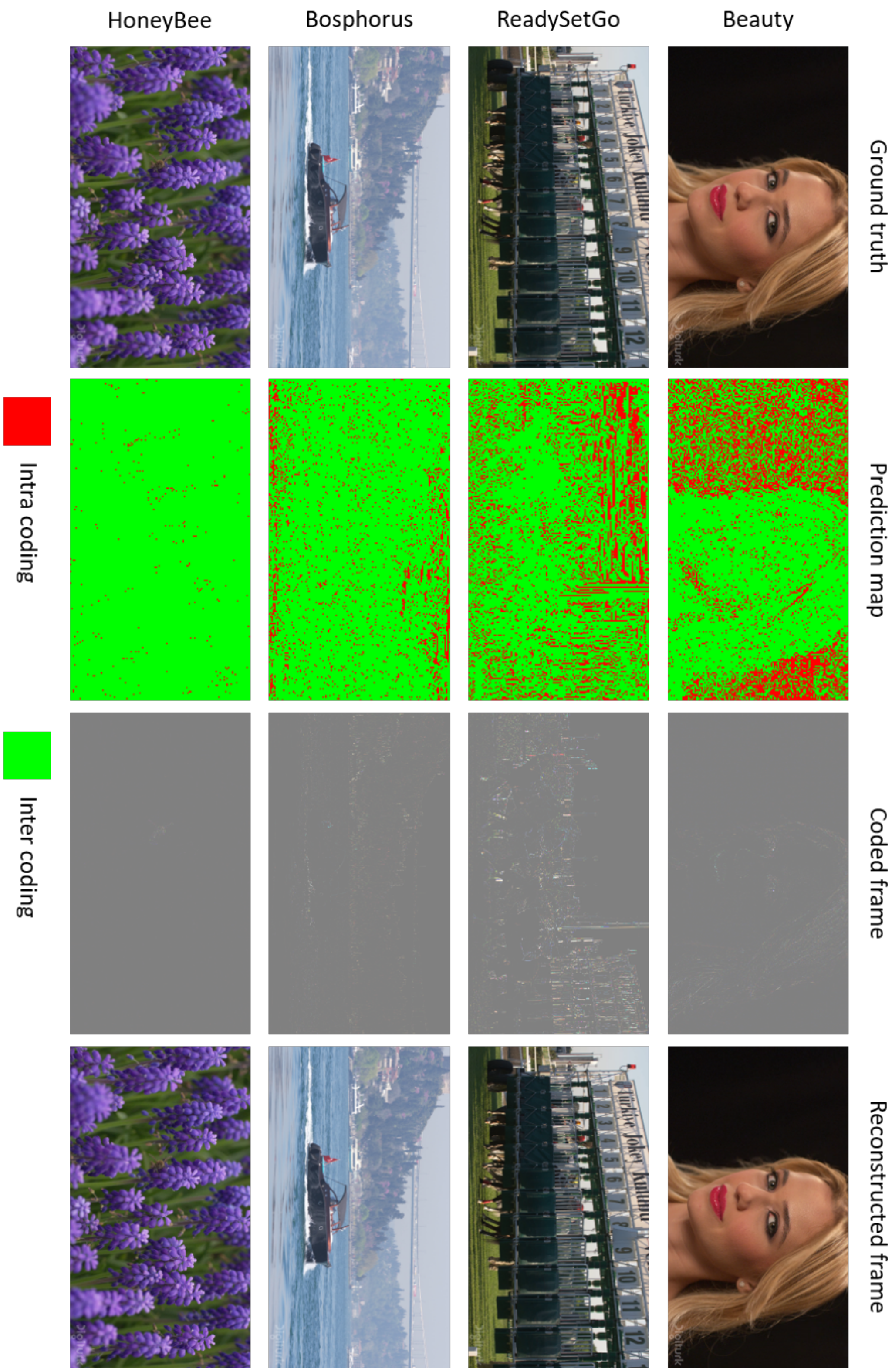

Figure 8.11: An example of prediction maps and coded frame, where red indicates intra coding and green indicates inter coding.

compress more efficiently using inter coding. Hence, by using intra-inter coding, it should give higher bpp reduction while still maintaining image quality. In this section, an experimental results of intra-inter coding for efficiently transmitting compressed measurements to the receiver is provided. According to the demonstration in Fig 8.11, the prediction map is present, which prediction mode used to compress RAW measurement. We demonstrate intra-inter coding in the long run by 100 frames using high motion (ReadySetGo) and low motion (Beauty) datasets. Nevertheless, we set up the experimental to transmit key-frame every 10 frames. In which measurement coding framework will encode RAW measurement using intra coding. According to the

Table 8.3: Intra-inter coding performance of 100 frames in terms of PSNR (dB), SSIM, and bpp.

Dataset	PSNR (dB)	SSIM	bpp
Beauty	41.16	0.94	0.2159
ReadySetGo	38.63	0.94	0.4293
Bosphorus	40.77	0.97	0.3043
HoneyBee	40.56	0.98	0.2448

results shown in Table 8.3, by applying intra-inter coding to RAW measurements, it can greatly reduce bpp by 40% compared the results in Table 8.2. Moreover, we can increase PSNR and SSIM by 10% because RAW measurements that compressed by inter-coding give relatively small residuals, which do not much disturb by quantization as intra-coding. Lastly, we provided bpp comparison between intra-inter coding and state-of-the-art works in Fig. 8.12 and Fig. 8.13. We note that only intra-coding cannot deliver an efficient approach in bpp reduction. On the other hand, inter-coding takes substantial benefit from a non-overlapping block-based partitioning to deliver more efficient compression performance. Subsequently, we note that spiking bpp usually happens when applying intra-inter prediction in conjunction with a non-overlapping block-based partitioning. It depends on how best matched the candidate that encoder can find on RAW measurement at that time. In general, intra coding will cause these spiking behavior.

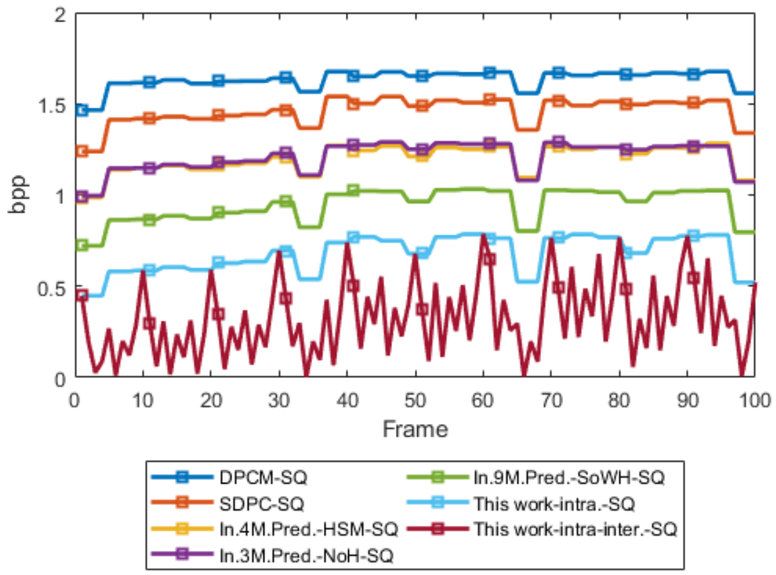

Figure 8.12: Result of bpp produced during 100 frames among state-of-the-art works using Beauty dataset, which consider as low motion sequence.

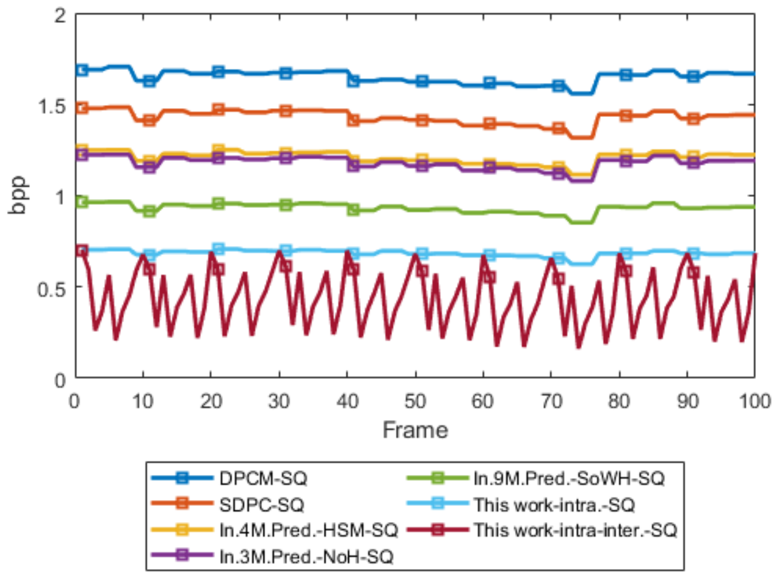

Figure 8.13: Result of bpp produced during 100 frames among state-of-the-art works using ReadySetGo dataset, which consider as high motion sequence.

8.7 Summary

Our proposal capitalizes on future high-resolution of faster imaging and efficient multimedia system. The proposed sensing matrix gave a better sensing scheme, allowing BCI to sample at lower SR than other sensing matrices. It allowed an image to be captured faster without ADC bottleneck. Subsequently, the proposed measurement coding framework gave an extra compression performance. Hence, RAW measurements can be stored and transmitted more efficiently than plainly using underdetermined systems regarding CS theory. The proposed measurement coding framework has low-complexity and outstanding performance compared to state-of-the-art works. In summary, our proposal can solve an occurring problem of conventional CIS design and high complexity multimedia compression algorithms. It allows the next generation of image/video acquisition systems to be revolutionized.

9

Cube-based Video Coding Framework for Block-based Compressive Imaging

Block-based compressive imaging enables new video acquisition methodology while reducing raw data size, theoretically eliminating the need for complex coding algorithms. When transmitting raw data, however, the redundancy associated with random projection remains. This paper takes a fresh look at raw data structure by viewing it as cube made up of multiple downsampled images rather than a vector. As a result, we can view each individual data point as a pixel, allowing us to code more flexibly and versatility than state-of-the-art works. Following that, we propose a tailored video coding algorithm for this structure that includes directional 9 modes intra and inter prediction with block-matching motion estimation, transformation, and quantization. We evaluated coding performance using various 4K datasets, resulting in 60-65 % lower bit-per-pixels while improving visual quality compared to state-of-the-art works.

9.1 Introduction

The traditional camera based on a hundred-year-old sampling theorem developed by Whittaker–Nyquist–Kotelnikov–Shannon has resulted in a massive problem of redundant data in image and video applications, which oversamples signal twice higher than information rate. It necessitates the use of complex lossy coding algorithms to reduce redundancy. However, the most recent coding algorithms are going far beyond coding efficiency; for instance, improving coding performance by 20% would cost roughly 50% more complexity and resources, which is still a significant issue today.

Recently, a new camera architecture based on block-based compressed sensing (CS) has gained popularity because it offers lower sampling costs, resulting in far fewer raw data generated but sufficient to represent the original content (69) (7). CS is based on the Johnson–Lindenstrauss lemma, which deals with low-distortion embedding of points from high to low dimensions via random projection, resulting in a compressed vector. It theoretically eliminates the need for coding algorithm. However, the recent studies found that raw data from CS camera is still redundant in the form of linear combination, potentially necessitating additional coding to reduce redundancy (110) (102) (76).

This paper takes a fresh look at raw data by viewing it as cube made up of multiple downsampled images rather than a vector. As a result, we can view each data point as a pixel. Following that, we propose a tailored video coding algorithm for cube structure that includes directional 9 modes intra and inter prediction with block-matching motion estimation, transformation, quantization, and entropy coding. When compared to state-of-the-art works, this proposal provides a significant improvement in coding performance and flexibility.

9.2 Early Works

At the first glance, scalar quantization (SQ) provided a straightforward and fast approach. By comparing to traditional source coding, SQ gave low bit-rate utilize but high recovery failure rate. Differential pulse-code modulation (DPCM) was introduced to reduce bit-rate over transmission channels for block-based CS camera (62). It computes a difference relative to two inputs using the values of two consecutive samples of previously coded CS data and raw CS data. The difference can be quantized further in this scenario, allowing for an excellent way to incorporate a controlled loss in the encoding. However, the previously compressed data may contain irrelevant information about the raw CS data that follows. In this case, it will result in an unstable bit-rate reduction, which can be reduced and increased at random. Later, spatially directional predictive coding (SDPC) was introduced (117), which effectively used the inherent spatial correlation of raw CS data. The best prediction is chosen from a set of prediction candidates generated by 4 designed directional predictive modes. This work outperformed DPCM in terms of bit-rate reduction while maintaining visual quality. Nonetheless, this work still remain recovery errors caused SQ. Furthermore, due to the limitations of data structure of vector, this work was unable to add more prediction modes. Directional 4 modes intra-prediction was proposed in (121). This work was inspired by intra-prediction concept in conventional video coding algorithms. This work embedded specific sensing patterns into the binary random sensing matrix to gain boundary information while sensing called semi-structured sensing matrix. Afterward, they used gained information and then multiplied it with sensing matrix to generate prediction candidates. This work significantly reduced bit-rate. Nevertheless, it cost tremendous degradation in image quality due to the modification of sensing matrix.

The works mentioned above have become an important fundamental foundation of the most recent video coding algorithm for CS camera. The work in (73) proposed directional 4 modes intra-prediction by generating prediction candidates using the stable structure of hadamard matrix, which partially points to the block boundary through sampled data. This method was used for the first time without modification to sensing matrix, and it performed previous works in terms of coding performance and provided excellent video quality. Furthermore, this work was extended from the software to hardware level. It turned out that it required far fewer resources and consumed far less power than the traditional video coding algorithm when performed at 4K resolution (71). The work in (95) then proposed directional 7 modes intra-prediction to reduce bit-rate using a stable structure of sequence ordered walsh-hadamard as sensing matrix. It significantly outperformed previous works by 15% lower bit-rate, while the coding algorithm did not produce much compression artifacts to video.

9.3 Proposed Video Coding Algorithm

The majority of early works appear to have limited the functionality of coding algorithms to specific sensing matrices. In this paper, we propose alternative data structure to observe raw CS data, which we believe will provide the coding algorithm with greater flexibility and universality. To begin, CS is a method for reducing the dimensional of signal x with a length of n into linear combination signal of y with length of m, where

m must be less than n. The process can be done through sensing matrix Φ with dimensional of $m \times n$ as shown in figure 9.1. However, look closely at the block-based

$$\begin{array}{ccccc} y & & \Phi & & x \\ \begin{bmatrix} y_1 = (x_1\Phi_{1,1}) + (x_2\Phi_{1,2}) + \cdots + (x_n\Phi_{1,n}) \\ y_2 = (x_1\Phi_{2,1}) + (x_2\Phi_{2,2}) + \cdots + (x_n\Phi_{2,n}) \\ \vdots \\ y_m = (x_1\Phi_{m,1}) + (x_2\Phi_{m,2}) + \cdots + (x_n\Phi_{m,n}) \end{bmatrix} & = & \begin{bmatrix} \Phi_{1,1} & \Phi_{1,2} & \cdots & \Phi_{1,n} \\ \Phi_{2,1} & \Phi_{2,2} & \cdots & \Phi_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_{m,1} & \Phi_{m,2} & \cdots & \Phi_{m,n} \end{bmatrix} & \times & \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \end{array}$$

Figure 9.1: An illustrate of undetermined system.

CS procedure; an entire frame is partitioned into several blocks with size of $b \times b$. An elements in the block will be vectorized into x and iteratively linearized into element of y via each 1D kernel order of Φ. It expects to obtain several downsampled images corresponding to m as output by grouping the same element order from all blocks as shown in figure 9.2. We called this data structure as cube. To be more clarify, we demonstrate the procedure using 4K frame size of 2160 $\times$ 3840 pixels in figure 9.3. As

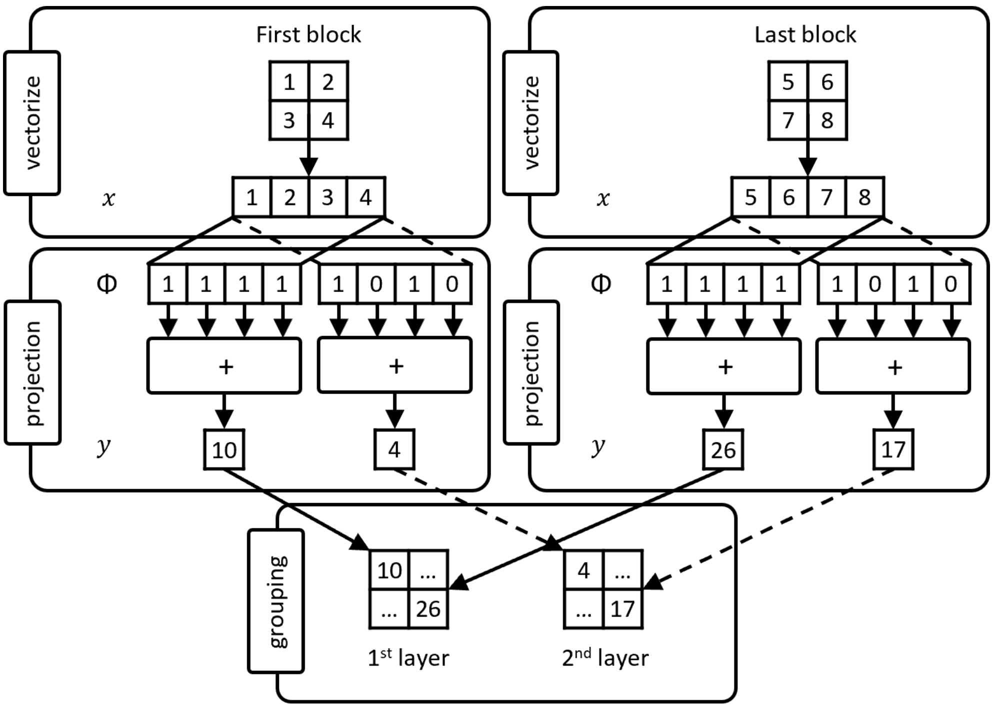

Figure 9.2: An example of block-based CS procedure, where the block size is set to be 2×2, $n = 4$, $m = 2$, which generate 2 downsampled images.

can be seen, the original frame has been downsampled into $135 \times 240 \times 8$ pixels. It achieves such a remarkable shift in perspective that we can now code raw CS data more efficiently using techniques similar to conventional coding algorithms.

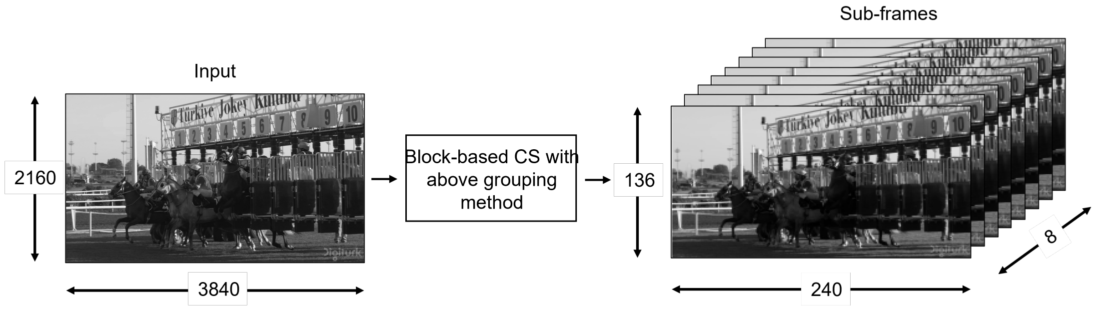

Figure 9.3: An example of a block-based CS downsampling procedure into a data cube over a 4K frame, where block size equal to 16×16 and $m = 8$.

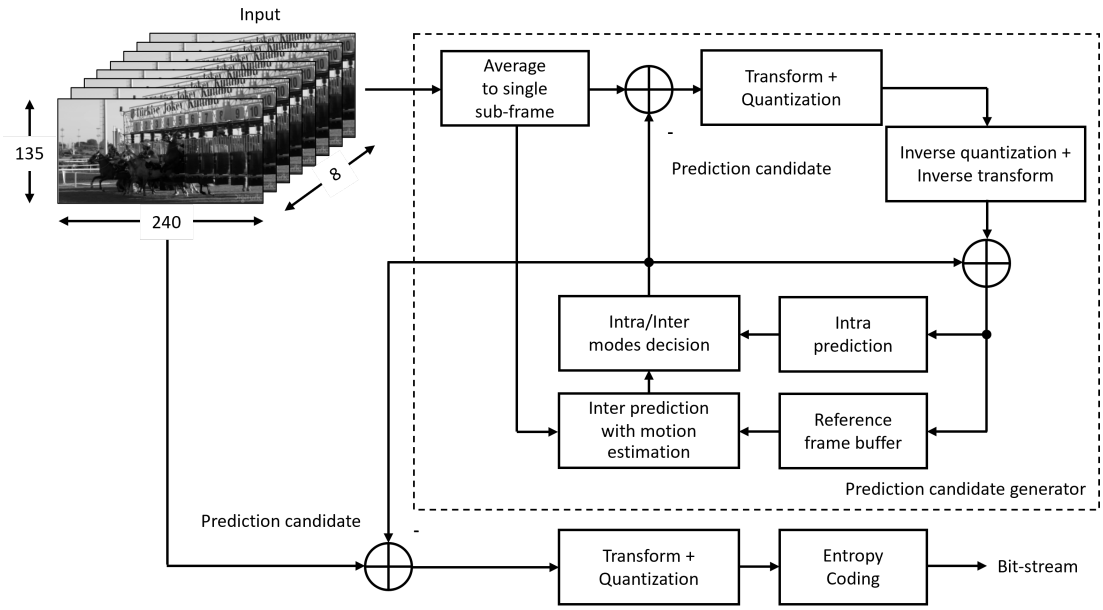

Figure 9.4: A proposed framework architecture tailored to a data cube for compressive imaging.

Following that, we propose a video coding framework tailored to these data structures, in which overall architecture can be seen in figure 9.4. Because the majority of the information in each layer of cube is roughly the same, coding each layer separately would result in redundant computation. As a result, we decided to average them into a single image and apply a coding algorithm to it. This method returns a prediction candidate from the prediction candidate generator, which can be used to code the entire cube indirectly.

Inside prediction candidate generator, we divide an image into multiple tiles with size of 4×4. Then, we apply directional 9 modes intra-prediction in order to code image spatially, which utilizes neighboring pixels to make a prediction candidate consisting of Mode 0: vertical mode, Mode 1: horizontal mode, Mode 3: DC mode, Mode 4: diagonal/left mode, Mode 5: diagonal/right mode, Mode 6: vertical/right mode, Mode 7: horizontal/down mode, Mode 8: vertical/left mode, and Mode 9: horizontal/up mode, as shown in figure 9.5. The sum-of-absolute-differences (SAD) is used to find the

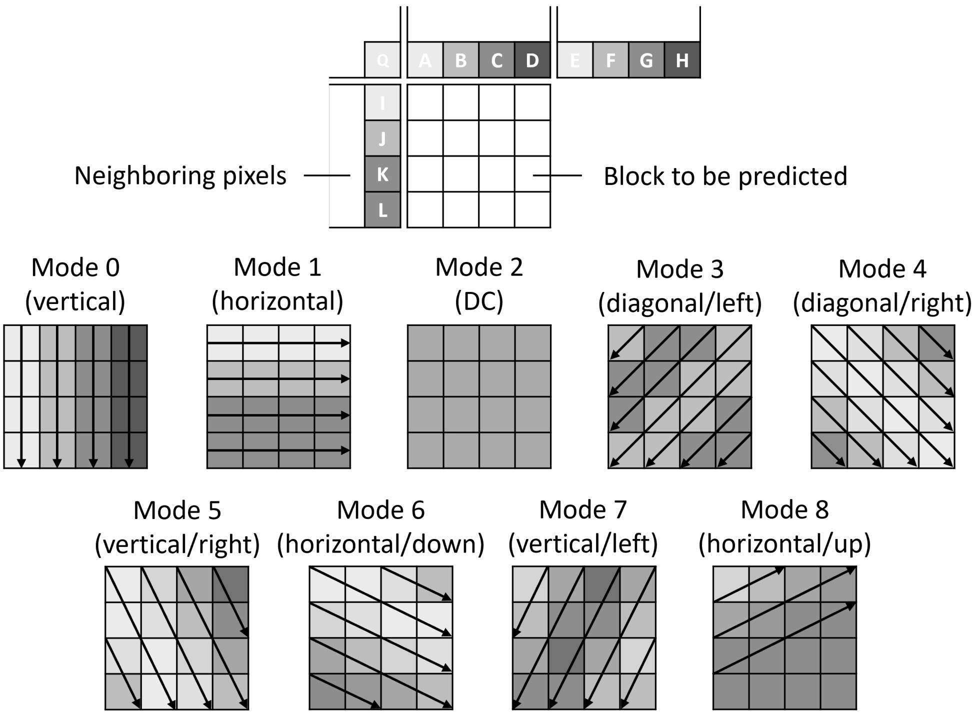

Figure 9.5: Standard 4×4 block directional 9 modes intra-prediction.

minimum distortion by comparing the current block (CB) that is being processed with reference blocks (RB) from each mode, which can describe by the following equation:

$$SAD = \sum \sum |CB - RB| \tag{9.1}$$

In addition, we apply a block-matching motion estimation algorithm with inter-prediction to code image temporally. The motion estimation comes with two modes:

slow and fast mode. The slow mode performs exhaustive search motion estimation in a specified area, whereas the fast mode performs multiple candidate three-level hierarchical search motion estimation, where v denotes the number of candidates. Same as intra-prediction, this coding mode uses SAD to find minimum distortion between CB and RB. Unlike the traditional approach, we modify the classical algorithm by employ-

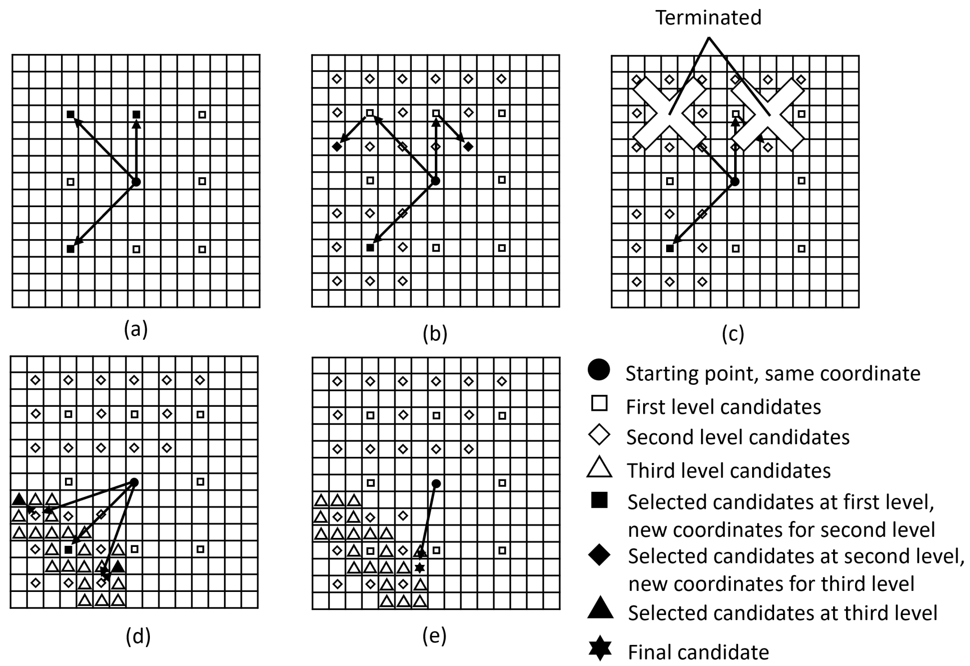

Figure 9.6: Fast three-level hierarchical search motion estimation capable of stopping motion vectors with v equal to 3, where (a) the top-level of hierarchical motion estimation, (b) the middle-level of hierarchical motion estimation, (c) two motion vectors in the middle-level of hierarchical motion estimation are stopped, (d) the lowest-level of hierarchical motion estimation received computational resources to active motion vector, and (e) the matched candidate is returned.

ing sorting algorithm, which allows algorithm to select new coordinate on each level by looking at v first minimum distortion candidates between current block and references blocks. The procedure is illustrated in figure 9.6.

The final candidate is determined by subtracting the prediction candidate from each coding mode with the current block. The mode with the smallest residual magnitude will be chosen as the final prediction candidate. The residual block is calculated by subtracting the final prediction candidate from the current block. Furthermore, we apply the DCT transform to the residual block, allowing quantization to perform admirably and simulating quantization loss when transmitting the coded block to the receiver. In this work, we use a custom 4×4 quantization table as shown in figure 9.7, where each coding mode uses a different table. The quantization table is adjustable by multiplying quantization parameter (Q_p) for greater symbols reduction performance.

$$\begin{bmatrix} 0.5 & 1 & 2 & 3 \\ 1 & 1 & 2 & 3 \\ 2 & 2 & 2 & 3 \\ 3 & 3 & 3 & 3 \end{bmatrix} \qquad \begin{bmatrix} 2 & 4 & 8 & 12 \\ 4 & 4 & 8 & 12 \\ 8 & 8 & 8 & 12 \\ 12 & 12 & 12 & 12 \end{bmatrix}$$

(a) (b)

Figure 9.7: Custom 4×4 quantization table, where (a) quantization table for intra-prediction and (b) quantization table for inter-prediction.

We embed decoder within a prediction candidate generator that also includes inverse quantization, inverse transformation, and a coded block restoration adder. The output will be used to predict neighboring blocks. Following the prediction candidate generator, the final prediction candidate will be used to generate a transmittable coded cube by subtracting with the entire cube, followed by DCT transformation and quantization using the same settings as in the prediction candidate generator.

9.4 Experimental Results

In this section, we evaluate the proposed algorithm on 4K resolution datasets such as Beauty, ReadySetGo, and Bosphorous ([61]). We use binary random sensing matrix, where $n = b \times b = 16 \times 16$ and $m = 64$. We compare coding performance using a variety of quantitative matrices, including PSNR and SSIM and bit-per-pixel (bpp). To compare existing works, we run an experiment in two parts: intra-prediction only and intra-prediction with inter-prediction. In Table 9.1, we compared our proposal to state-of-the-art coding methods using 3 quantization setups via Q_p. We set the quantization bit to 4 bits for state-of-the-art works, as stated in reports.

Overall, the work of (a) performs insufficiently in terms of bpp reduction while maintaining PSNR and SSIM because they only used a single prediction candidate from previously coded CS data. The work of (b)-(e) uses an advanced coding algorithm

Table 9.1: Intra-prediction performance comparison with the state-of-the-art on various datasets, where $b \times b = n = 16 \times 16$, $m = 64$, (a) binary random matrix with DPCM-SQ ([62]), (b) binary random matrix with SDPC-SQ ([117]), (c) binary random matrix with modified-binary random matrix with 3 modes intra prediction-SQ ([121]), (d) hadamard with 4 modes intra prediction-SQ ([71]), (e) sequency-ordered walsh-hadamard with 7 modes intra prediction-SQ ([95]), (f) this work only intra-prediction with $Q_p = 3$, (g) this work only intra-prediction with $Q_p = 6$, and (h) this work only intra-prediction with $Q_p = 9$.

Coding methods	Beauty dataset	ReadySetGo dataset	Bosphorus dataset
	PSNR / SSIM / bpp	PSNR / SSIM / bpp	PSNR / SSIM / bpp
(a)	34.77 / 0.79 / 2.46	34.35 / 0.85 / 2.84	35.24 / 0.93 / 2.73
(b)	34.53 / 0.79 / 2.23	34.06 / 0.82 / 2.52	34.37 / 0.90 / 2.52
(c)	34.77 / 0.79 / 1.98	36.04 / 0.92 / 2.21	35.42 / 0.93 / 2.28
(d)	35.85 / 0.84 / 1.92	36.04 / 0.92 / 2.14	39.13 / 0.97 / 2.29
(e)	37.13 / 0.86 / 1.65	35.79 / 0.90 / 1.91	40.37 / 0.94 / 1.94
(f)	47.92 / 0.98 / 0.83	39.11 / 0.97 / 0.99	43.22 / 0.98 / 0.95
(g)	47.65 / 0.98 / 0.64	39.07 / 0.97 / 0.84	43.13 / 0.98 / 0.78
(h)	47.24 / 0.98 / 0.53	39.01 / 0.97 / 0.75	42.98 / 0.97 / 0.68

with intra-prediction to improve the performance of bpp reduction on each proposal by increasing prediction modes and changing sensing matrices. PSNR and SSIM, on the other hand, have not improved significantly. In contrast, our propose in (f)-(h) with $Q_p = 3$, 6, and 9 outperforms state-of-the-art works in PSNR and SSIM, resulting in richer visual quality. Further, we can reduce bpp average by 60-65% across all datasets. Furthermore, as shown in figure 9.8, we show the performance curve of recovery error via mean square error (MSE) against multiple Qp ranging from 2 to 50, with a step size of 2. We show the quality drop caused by quantization in 9.9, where the quality drops and quantization artifacts begin to appear as Q_p increases.

We present the experimental results of proposal, including intra-prediction and inter-prediction in Table 9.2, where inter-prediction of (a)-(c) performed in slow mode and (d)-(f) performed in fast mode with the same condition of Q_p. It gave better bpp reduction averagely by 48-50% compared to only intra-prediction approach in Table 9.1. We will note that the performance of each searching mode in inter-prediction can be determined by downsampled image size, which means that if the frame size is not large enough, it will be difficult to see the performance.

Table 9.2: Intra-prediction with 2 modes inter-prediction performance on 100 frames of various datasets, where $b \times b = n = 16 \times 16$, $m = 64$. The results in (a), (b), (c), represent intra-prediction with slow mode inter-prediction; (d), (e), (f), represent intra-prediction with fast mode inter-prediction, where $Q_p = 3, 6$, and 9, respectively.

Coding methods	Beauty dataset	ReadySetGo dataset	Bosphorus dataset
	PSNR / SSIM / bpp	PSNR / SSIM / bpp	PSNR / SSIM / bpp
(a)	48.79 / 0.98 / 0.4496	39.27 / 0.97 / 0.7151	43.17 / 0.98 / 0.6237
(b)	48.66 / 0.98 / 0.3078	39.26 / 0.97 / 0.5581	43.68 / 0.98 / 0.4632
(c)	48.50 / 0.98 / 0.2387	39.24 / 0.97 / 0.4719	43.62 / 0.98 / 0.3794
(d)	48.79 / 0.98 / 0.4495	39.27 / 0.97 / 0.7151	43.71 / 0.98 / 0.6245
(e)	48.67 / 0.98 / 0.3084	39.26 / 0.97 / 0.5581	43.68 / 0.98 / 0.4634
(f)	48.05 / 0.98 / 0.2387	39.24 / 0.97 / 0.4719	43.64 / 0.98 / 0.3794

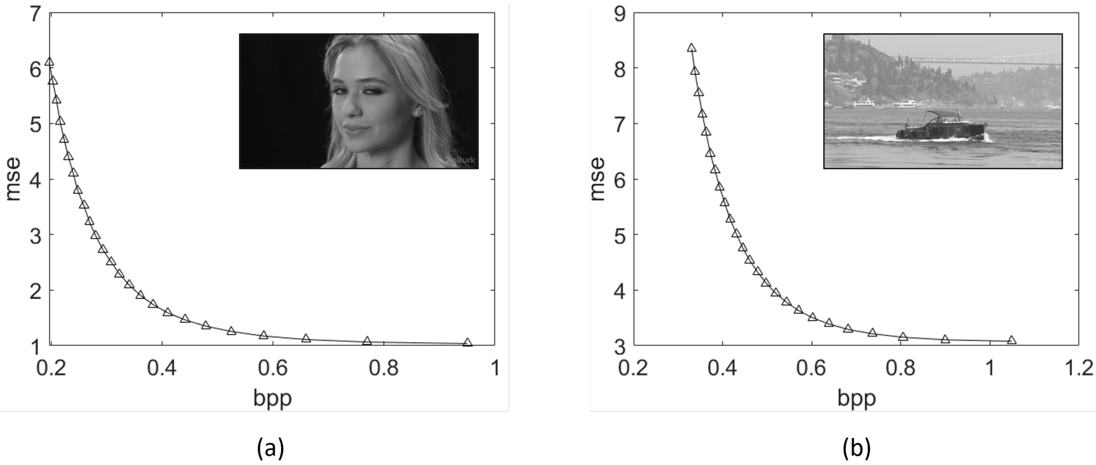

Figure 9.8: Comparison graph of recovery error via mse and bpp reduction using Beauty dataset in (a) and Bosphorous dataset in (b), where Q_p ranges from 2 to 50 and step size is equal to 2.

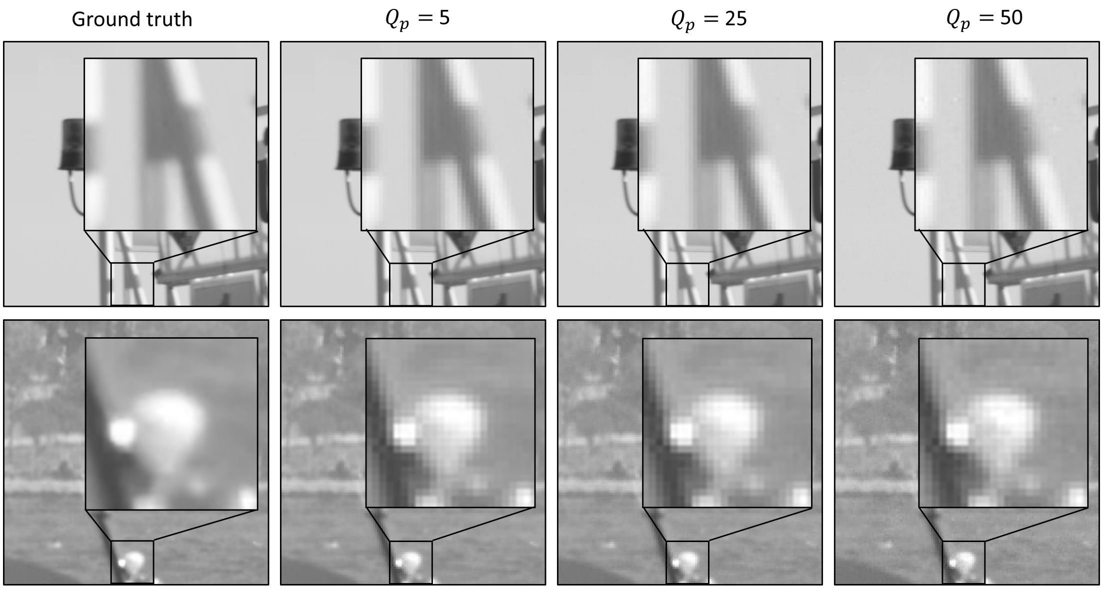

Figure 9.9: Remaining quality and quantization artifacts comparison of Readysetgo and Bosphorous datasets under different setup of $Q_p \in 5, 25, 50$.

9.5 Summary

This paper presents a video coding framework for compressive video sensing while also changing the perspective from which we view raw CS data as a cube made up of multiple downsampled images rather than a vector. The proposed video coding algorithm for the CS camera significantly improved coding performance and provided more flexibility for future implementation. Notably, the video coding algorithm functionality will not be constrained by the structure of sensing matrices, as is commonly found in state-of-the-art works.

10

Discussion

This section discusses CS thoughts on the future CMOS image sensor, as well as image and video coding algorithms. First and foremost, we are well aware that the traditional approach has reached its limits in terms of data acquisition and compression efficiency, so we rarely use the entire RAW data as sampled. Each generation of algorithms was designed to maintain an increasing screen resolution, such as HEVC, which was designed for 1080p and 2K resolutions. Of course, it performed well at that resolution but poorly at 4K and 8K, where VVC performs better, as we are all expected. However, following AVC, which promised to be the last known software-based codec, all next-generation codecs require specific encoder and decoder chips, primarily found as a co-processor in CPU and GPU. They are expensive, and if we did not have programs that required co-processor operation, we would never turn them on to say hello in practice. Currently, we are all still using AVC, a video format that has been around for 20 years. The major companies such as Microsoft, VLC, NHK, Sony, Apple, Netflix, and Google are also aware of coding efficiency in both software and hardware, resulting in software-based codec for both image and video such as JPEG-XS for XR/VR, VP9/10, and AV1, where they do care less about compression ratio but lower latency and cheaper in computational resources.

Nowadays, deep learning-based data compression algorithms have been successfully implemented; however, the amount of data required to generate specific models under strict conditions makes real-world applications difficult. For example, some deep learning methods require at least one week to train a prospect model without knowing whether it will be successful or not; at this point, a massive amount of data has been sampled using the Nyquist-Shannon sampling theorem. We can say that tons of data are required further to discard a small fraction of the target data. Further, it requires high-end hardware to do this codec, such as GPU, which the bottleneck problem can occur if we reduce the computational resources at the decoder. Hence, the deep learning-based approach appears to be more inefficient and troublesome than the conventional codec.

Later, the data acquisition model was changed from Nyquist-Shannon to CS, which was recently and successfully implemented to make researchers aware of the root cause problem. In other words, compression algorithms are proposed to solve an inefficient data acquisition model in which we cannot store the RAW data. Nowadays, any data can be acquired efficiently with less redundancy and without the need for additional data compression. However, it should be noted that some codec is still required for image and video applications that use CS camera because the amount of data is still enormous when frame-rate is high; however, an algorithm complexity will be straightforward and possibly determined by the quality requirement.

This thesis was finally proven by applying CS to image and video applications to solve problems in CMOS image sensors and alternate image and video coding algorithms. Furthermore, a new perspective on CS data has been addressed in order to effect a regime shift in which we should not have expected CS cameras to produce vector data but volumetric data with depth proportional to the number of observations. Under various 4K datasets, the proposed algorithms can compress CS data in real-time by software without the use of an expensive chip. Finally, the future direction of this research field is provided. The high-level image and video coding framework are written in MATLAB and Python3, but some back-end algorithms should be re-written in C/C++ or FORTRAN for faster numerical computation. The sparse solver is then written in Python3 and modified to accept infinity (Inf) and Not-a-Number (NaN) conditions. It is generally unacceptable in computer programming, but it is required in applied mathematics and convex optimization via linear programming. Moreover, FPGA/ASIC implementation is greatly needed; however, it would be difficult to implement floating-point following with Inf and NaN conditions so that the result may be different. Here, I would suggest doing hardware/software co-design where we will not transfer all of the algorithms to hardware, but reasoning divides the computational section.

11

Dissertation Summary

This section contains a summary of my works. In chapter 3, A novel sensing matrix called Continuously ordered Walsh was introduced. It transforms binary sensing matrix research by focusing to enhance worst-case sensing conditions, resulting in shorter sensing times and huge improve in image quality. Apart from the breakthrough in binary sensing matrix, which served to improve the foundation theory of the compressive sensing theorem. I discovered that some redundancy remained in the measurement data across all sampling rates. As a result, when it comes to image and video applications, measurement data must be further compressed and placed in a suitable container before being sent to the receiver. In chapter 4, 5, and 8, a vector-based image and video coding algorithms were proposed that addressed both spatial and temporal redundancy reduction by using intra-prediction and inter-prediction, respectively. In chapter 6, I proposed a special module that only codes moving components in scenes. Later, the modules were combined into single video coding framework for compressive imaging. It achieved the highest novelty in video coding framework for compressive imaging, which is still progressing as of the writing of this dissertation and has become a standard for future improvement. In addition, I look into the possibility of hardware implementation in order to compare hardware complexity and implementation cost to traditional hardware. In which the result turnout to be straightforward and cheaper in development even demonstrated at the same resolution such as 4K and 8K. In chapter 7, inter-prediction module was further improved in complexity and coding performance by adding multiple candidates based hybrid hierarchical block-matching search algorithm. This algorithm searches for the best match vector among neighboring blocks to the target block, resulting in better compression performance than co-location-based coding. Furthermore, the algorithm can halt search vectors that are potentially unable to find suitable candidates and redirect computation resources to other searching vectors. It produced results similar to an exhaustive search but with a shorter coding time. In chapter 9, the last chapter, takes a fresh look at raw data by viewing it as cube made up of multiple downsampled images rather than a vector. As a result, we can view each data point as a pixel. Following that, we propose a tailored video coding algorithm for cube structure that includes directional 9 modes intra and inter prediction with block-matching motion estimation, transformation, quantization, and entropy coding. When compared to state-of-the-art works, this proposal provides a significant improvement in coding performance and flexibility.

References

[1] S. Aasha Nandhini, R. Sankararajan, and K. Rajendiran. Video compressed sensing framework for wireless multimedia sensor networks using a combination of multiple matrices. *Computers Electrical Engineering*, 44:51–66, 2015. 27, 96

[2] G. R. Arce, D. J. Brady, L. Carin, H. Arguello, and D. S. Kittle. Compressive coded aperture spectral imaging: An introduction. *IEEE Signal Processing Magazine*, 31(1):105–115, 2014. 13

[3] Y. Arjoune, N. Kaabouch, H. El Ghazi, and A. Tamtaoui. A performance comparison of measurement matrices in compressive sensing. *International Journal of Communication Systems*, 31(10):e3576, 2018. e3576 IJCS-17-0570.R1. 30

[4] Y. Baig, E. M.-K. Lai, and J. P. Lewis. Quantization effects on compressed sensing video. In *2010 17th International Conference on Telecommunications*, pages 935–940, 2010. 103

[5] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin. A simple proof of the restricted isometry property for random matrices. *Constructive Approximation*, 28(3):253–263, Dec 2008. 34

[6] R. G. Baraniuk. Compressive sensing [lecture notes]. *IEEE Signal Processing Magazine*, 24(4):118–121, 2007. 12

[7] W. Benjilali, W. Guicquero, L. Jacques, and G. Sicard. Hardware-compliant compressive image sensor architecture based on random modulations and permutations for embedded inference. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 67(4):1218–1231, 2020. 116

[8] J. M. Bioucas-Dias and M. A. T. Figueiredo. A new TwIST: Two-step iterative shrinkage/thresholding algorithms for image restoration. 16(12):2992–3004, Dec. 2007. 95

[9] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.*, 3(1):1–122, Jan. 2011. 26

[10] C. Boyer, J. Bigot, and P. Weiss. Compressed sensing with structured sparsity and structured acquisition. *Applied and Computational Harmonic Analysis*, 46(2):312–350, 2019. 30

[11] Y. Bresler and P. Feng. Spectrum-blind minimum-rate sampling and reconstruction of 2-d multiband signals. In *Proceedings of 3rd IEEE International Conference on Image Processing*, volume 1, pages 701–704 vol.1, 1996. 12

[12] E. Candes and J. Romberg. "l1-magic: Recovery of Sparse Signals via Convex Programming". Oct. 2005. 95

[13] E. Candes, J. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489–509, 2006. 12

[14] E. Candes and T. Tao. Decoding by linear programming. *IEEE Transactions on Information Theory*, 51(12):4203–4215, 2005. 19, 94

[15] E. J. Candes. J. *Romberg, Tao, T. Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information. IEEE Trans. Inf. Theory 2006*, 52, 2006. 95

[16] E. J. Candes and J. Romberg. Quantitative robust uncertainty principles and optimally sparse decompositions. *Foundations of Computational Mathematics*, 6(2):227–254, Apr 2006. 12

[17] E. J. Candes, J. K. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 59(8):1207–1223, 2006. 31, 95

[18] E. J. Candes and T. Tao. Near-optimal signal recovery from random projections: Universal encoding strategies. Nov. 2006. 83, 92

[19] E. Candès and J. Romberg. (2006). *Sparsity and Incoherence in Compressive Sampling. Inverse Problems. 23. 10*, 23.(10.). 51, 92

[20] M. Dadkhah, M. J. Deen, and S. Shirani. Cmos image sensor with area-efficient block-based compressive sensing. *IEEE Sensors Journal*, 15(7):3699–3710, 2015. 3

[21] M. Dadkhah, M. Jamal Deen, and S. Shirani. Block-based cs in a cmos image sensor. *IEEE Sensors Journal*, 14(8):2897–2909, 2014. 27, 96

[22] W. Dai and O. Milenkovic. Information theoretical and algorithmic approaches to quantized compressive sensing. 2011. 52, 97

[23] M. A. Davenport, C. Hegde, M. F. Duarte, and R. G. Baraniuk. Joint manifolds for data fusion. *IEEE Transactions on Image Processing*, 19(10):2580–2594, 2010. 13

[24] M. A. Davenport, J. N. Laska, J. R. Treichler, and R. G. Baraniuk. The pros and cons of compressive sensing for wideband signal acquisition: Noise folding versus dynamic range. *IEEE Transactions on Signal Processing*, 60(9):4628–4642, 2012. 13

[25] R. DeVore. Nonlinear approximation. *Acta Numer*, 7:51–150, 1998. 16

[26] D. Donoho. De-noising by soft-thresholding. *IEEE Transactions on Information Theory*, 41(3):613–627, 1995. 16

[27] D. L. Donoho. Compressed sensing. In *IEEE Transactions on Information Theory*, volume 52, pages 1289–1306. Apr. 2006. 29, 83

[28] D. L. Donoho, Y. Tsaig, I. Drori, and J. Starck. Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit. 58(2):1094–1121, Feb. 2012. 95

[29] M. Duarte, S. Sarvotham, D. Baron, M. Wakin, and R. Baraniuk. Distributed compressed sensing of jointly sparse signals. In *Conference Record of the Thirty-Ninth Asilomar Conference onSignals, Systems and Computers, 2005.*, pages 1537–1541, 2005. 13

[30] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk. Single-pixel imaging via compressive sampling. *IEEE Signal Processing Magazine*, 25(2):83–91, March 2008. 13, 26

[31] M. F. Duarte and Y. C. Eldar. Structured compressed sensing: From theory to applications. In *IEEE Transactions on Signal Processing*, volume 59, pages 4053–4085. 10, doi, 2011. 30, 98

[32] M. F. Duarte et al. Single-pixel imaging via compressive sampling. 25(2):83–91, Mar. 2008. 50, 95

[33] A. Farina, A. Candeo, A. D. Mora, A. Bassi, R. Lussana, F. Villa, G. Valentini, S. Arridge, and C. D'Andrea. Novel time-resolved camera based on compressed sensing. *Opt. Express*, 27(22):31889–31899, Oct 2019. 27, 95

[34] P. Feng and Y. Bresler. Spectrum-blind minimum-rate sampling and reconstruction of multiband signals. In *1996 IEEE International Conference on Acoustics, Speech, and Signal Processing Conference Proceedings*, volume 3, pages 1688–1691 vol. 3, 1996. 12

[35] J. E. Fowler, S. Mun, and E. W. Tramel. *"Multiscale Block Compressed Sensing with Smoothed Projected Landweber Reconstruction, " in Proceedings of the European Sig-nal Processing Conference, Barcelona, Spain.* Aug. 2011. viii, xiii, 44, 45

[36] L. Gan. Block compressed sensing of natural images. In *2007 15th International Conference on Digital Signal Processing*, pages 403–406, 2007. 51, 94

[37] L. Gan, T. T. Do, and T. D. Tran. Fast compressive imaging using scrambled block hadamard ensemble. In *2008 16th European Signal Processing Conference*, pages 1–5, 2008. 30

[38] I. Gorodnitsky and B. Rao. Sparse signal reconstruction from limited data using focuss: a re-weighted minimum norm algorithm. *IEEE Transactions on Signal Processing*, 45(3):600–616, 1997. 12

[39] I. Gorodnitsky, B. Rao, and J. George. Source localization in magnetoencephalography using an iterative weighted minimum norm algorithm. In *[1992] Conference Record of the Twenty-Sixth Asilomar Conference on Signals, Systems Computers*, pages 167–171 vol.1, 1992. 12

[40] I. F. Gorodnitsky, J. S. George, and B. D. Rao. Neuromagnetic source imaging with focuss: a recursive weighted minimum norm algorithm. *Electroencephalography and Clinical Neurophysiology*, 95(4):231–251, 1995. 12

[41] M. Goukhshtein, P. T. Boufounos, T. Koike-Akino, and S. C. Draper. Distributed coding of quantized random projections. *IEEE Transactions on Signal Processing*, 68:5924–5939, 2020. 27, 96

[42] V. K. Goyal, A. K. Fletcher, and S. Rangan. “compressive sampling and lossy compression, ” IEEE Signal Processing Magazine. 25(2):48–56, Mar. 2008. 50

[43] P. S. Gupta and G. S. Choi. Image acquisition system using on sensor compressed sampling technique. *Journal of Electronic Imaging*, 27(1):1 – 16, 2018. 95

[44] D. Guzzi, G. Coluccia, D. Labate, C. Lastri, E. Magli, V. Nardino, L. Palombi, I. Pippi, D. Coltuc, A. Z. Marchi, and V. Raimondi. Optical compressive sensing technologies for space applications: instrumental concepts and performance analysis. 11180:2266 – 2275, 2019. 27, 96

[45] J. Haupt, W. U. Bajwa, M. Rabbat, and R. Nowak. Compressed sensing for networked data. *IEEE Signal Processing Magazine*, 25(2):92–101, 2008. 13

[46] L. Jacques, K. Degraux, and C. D. Vleeschouwer. Quantized iterative hard thresholding: Bridging 1-bit and high-resolution quantized compressed sensing. 2013. 52, 97

[47] L. Jacques, P. Vandergheynst, A. Bibet, V. Majidzadeh, A. Schmid, and Y. Leblebici. Cmos compressed imaging by random convolution. In *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 1113–1116, 2009. 27, 96

[48] H. Jiang, W. Deng, and Z. Shen. Surveillance video processing using compressive sensing. *Inverse Problems Imaging,6,2,201*, 214, 1930. 50, 96

[49] H. Jiang, S. Zhao, Z. Shen, W. Deng, P. A. Wilford, and R. Haimi-Cohen. Surveillance video analysis using compressive sensing with low latency. *Bell Labs Technical Journal*, 18(4):63–74, 2014. 27, 50, 84

[50] B. Kaliannan and V. S. R. Pasupureddi. A low power cmos imager based on distributed compressed sensing. In *2014 27th International Conference on VLSI Design and 2014 13th International Conference on Embedded Systems*, pages 534–538, 2014. 3

[51] N. B. Karahanoglu and H. Erdogan. A* orthogonal matching pursuit: Best-first search for compressed sensing signal recovery. *Digital Signal Processing*, 22, 2012. 95

[52] H. Lee, D. Seo, W.-T. Kim, and B.-G. Lee. A compressive sensing-based cmos image sensor with second-order $\sigma\delta$ adcs. *IEEE Sensors Journal*, 18(6):2404–2410, 2018. 3

[53] S. Leitner, H. Wang, and S. Tragoudas. Design of scalable hardware-efficient compressive sensing image sensors. *IEEE Sensors Journal*, 18(2):641–651, 2018. 3

[54] C. Li, H. Jiang, P. Wilford, Y. Zhang, and M. Scheutzow. A new compressive video sensing framework for mobile broadcast. *IEEE Transactions on Broadcasting*, 59(1):197–205, 2013. 103

[55] K. Li and S. Cong. State of the art and prospects of structured sensing matrices in compressed sensing. *Frontiers of Computer Science*, 9(5):665–677, Oct 2015. 30

[56] M. Lotfi and M. Vidyasagar. Compressed sensing using binary matrices of nearly optimal dimensions. *IEEE Transactions on Signal Processing*, 68:3008–3021, 2020. 30

[57] W. Lu, T. Dai, and S.-T. Xia. Binary matrices for compressed sensing. *IEEE Transactions on Signal Processing*, 66(1):77–85, 2018. 30

[58] W. Lu, W. Li, K. Kpalma, and J. Ronsin. Compressed sensing performance of random bernoulli matrices with high compression ratio. 22(8):1074–1078, Aug. 2015. 52

[59] M. Lustig, D. Donoho, and J. M. Pauly. Sparse mri: The application of compressed sensing for rapid mr imaging. *Magnetic Resonance in Medicine*, 58(6):1182–1195, 2007. 13

[60] S. Mallat. A wavelet tour of signal processing. academic press. 1999. 16

[61] A. Mercat, M. Viitanen, and J. Vanne. Uvg dataset: 50/120fps 4k sequences for video codec analysis and development. In *Proceedings of the 11th ACM Multimedia Systems Conference*, MMSys '20, page 297–302, New York, NY, USA, 2020. Association for Computing Machinery. 122

[62] S. Mun and J. E. Fowler. Dpcm for quantized block-based compressed sensing of images. In *2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO)*, pages 1424–1428, 2012. viii, xiii, xiv, 44, 45, 52, 97, 105, 116, 123

[63] S. A. Nandhini, R. Sankararajan, and K. Rajendiran. Video Compressed Sensing framework for Wireless Multimedia Sensor Networks using a combination of multiple matrices. *Computers Electrical Engineering*, 44, 2015. 50

[64] D. Needell and J. A. Tropp. Cosamp: Iterative signal recovery from incomplete and inaccurate samples. *Applied and Computational Harmonic Analysis*, 26, 2008. 95

[65] S. Nirmalraj and G. Nagarajan. Biomedical image compression using fuzzy transform and deterministic binary compressive sensing matrix. 2020. 52

[66] H. Nyquist. Certain topics in telegraph transmission theory. *Trans AIEE*, 47:617–644, 1928. 16

[67] Y. Oike and A. El Gamal. Cmos image sensor with per-column adc and programmable compressed sensing. *IEEE Journal of Solid-State Circuits*, 48(1):318–328, 2013. 3, 27, 50, 96

[68] B. A. Olshausen and D. J. Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381(6583):607–609, Jun 1996. 16

[69] C. Park, W. Zhao, I. Park, N. Sun, and Y. Chae. A 51-pj/pixel 33.7-db psnr 4× compressive cmos image sensor with column-parallel single-shot compressive sensing. *IEEE Journal of Solid-State Circuits*, 56(8):2503–2515, 2021. 116

[70] M. Pascucci, S. Ganesan, A. Tripathi, O. Katz, V. Emiliani, and M. Guillon. Compressive three-dimensional super-resolution microscopy with speckle-saturated fluorescence excitation. *Nature Communications*, 10(1):1327, Mar 2019. 26, 50, 96

[71] J. Peetakul, Y. Fan, and J. Zhou. Intra prediction-based measurement coding algorithm for block-based compressive sensing images. *IEEE Access*, 9:56031–56040, 2021. xiv, 117, 123

[72] J. Peetakul and J. Zhou. Temporal Redundancy Reduction in Compressive Video Sensing by using Moving Detection and Inter-Coding. *2020 Data Compression Conference (DCC), Snowbird, UT, USA*, 10.:387–387, 2020. 98, 99

[73] J. Peetakul, J. Zhou, and K. Wada. A measurement coding system for block-based compressive sensing images by using pixel-domain features. In *2019 Data Compression Conference (DCC)*, pages 599–599, 2019. xiii, 53, 84, 98, 99, 105, 117

[74] H. Podmore, A. Scott, and R. Lee. On-chip compressed sensing fourier-transform visible spectrometer. *IEEE Photonics Journal*, 10(6):1–10, 2018. 27, 95

[75] S. Pudlewski, T. Melodia, and A. Prasanna. Compressed-sensing enabled video streaming for wireless multimedia sensor networks. *IEEE T. Mobile Comp*, 11(6), 2011. 50

[76] F. Qian, Y. Gong, and L. Wang. A memristor-based compressive sampling encoder with dynamic rate control for low-power video streaming. *J. Emerg. Technol. Comput. Syst.*, 16(2), Jan. 2020. 116

[77] B. Rao. Signal processing with the sparseness constraint. In *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP '98 (Cat. No.98CH36181)*, volume 3, pages 1861–1864 vol.3, 1998. 12

[78] A. Ravelomanantsoa, H. Rabah, and A. Rouane. Compressed sensing: A simple deterministic measurement matrix and a fast recovery algorithm. In *IEEE Transactions on Instrumentation and Measurement*, volume 64, pages 3405–3413. 10, doi, Dec. 2015. 98

[79] R. Robucci, J. D. Gray, L. K. Chiu, J. Romberg, and P. Hasler. Compressive sensing on a cmos separable-transform image sensor. *Proceedings of the IEEE*, 98(6):1089–1101, 2010. 13, 50

[80] R. Saab, R. Wang, and Yılmaz. From compressed sensing to compressed bit-streams: Practical encoders, tractable decoders. *IEEE Transactions on Information Theory*, 64(9):6098–6114, 2018. 27, 50, 96

[81] C. Shannon. Communication in the presence of noise. *Proceedings of the IRE*, 37(1):10–21, 1949. 1, 29

[82] H. J. M. Shi, M. Case, X. Gu, S. Tu, and D. Needell. Methods for quantized compressed sensing. Jan. 2016. 52, 97

[83] S. Shindoh. Yu. nesterov and a. nemirovskii, interior point polynomial algorithms in convex programming, siam, 1994, 405pp. *Bulletin of the Japan Society for Industrial and Applied Mathematics*, 6(4):344–345, 1996. 24

[84] D. Slepian and J. Wolf. "noiseless coding of correlated information sources. " *IEEE Trans. Inf. Theory, vol. 19, no*, 19(4):471–480, Sept. 1973. 52, 97

[85] V. Slivinskas and V. Šimonyte. On the foundation of prony's method. *IFAC Proceedings Volumes*, 19(5):121–126, 1986. 2nd IFAC Symposium on Stochastic Control, Vilnius, Lithuanian SSR, USSR, 19-23 May 1986. 12

[86] X. Song, X. Peng, J. Xu, G. Shi, and F. Wu. Distributed compressive sensing for cloud-based wireless image transmission. *IEEE Transactions on Multimedia*, 19(6):1351–1364, 2017. 103

[87] G. J. Sullivan, J.-r. Ohm, A. Ortega, E. Delp, A. Vetro, and M. Barni. dsp forum - future of video coding and transmission. *IEEE Signal Processing Magazine*, 23(6):76–82, 2006. 3

[88] M.-J. Sun, L.-T. Meng, M. P. Edgar, M. J. Padgett, and N. Radwell. A russian dolls ordering of the hadamard basis for compressive single-pixel imaging. *Scientific Reports*, 7(1):3464, Jun 2017. 30

[89] R. T. T. Hastie and J. Friedman. The elements of statistical learning. springer. 2001. 16

[90] T. Tanaka. Performance analysis of l1-norm minimization for compressed sensing with non-zero-mean matrix elements. *IEEE International Symposium on Information Theory (ISIT), Vail, CO*, 10.:401–405, 2018. 98

[91] D. Taubman and M. Marcellin. Jpeg 2000: Image compression fundamentals, standards and practice. 2001. 16

[92] R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):267–288, 1996. 16

[93] A. M. Tillmann and M. E. Pfetsch. The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing. *IEEE Transactions on Information Theory*, 60(2):1248–1259, 2014. 19

[94] T. T. T. Tran, J. Peetakul, C. D. K. Pham, and J. Zhou. Bi-directional intra prediction based measurement coding for compressive sensing images. *2020 IEEE 22nd International Workshop on Multimedia Signal Processing (MMSP), Tampere*, 10.:1–6, 2020. xiii, 99, 105

[95] T. T. T. Tran, J. Peetakul, C. D. K. Pham, and J. Zhou. Bi-directional intra prediction based measurement coding for compressive sensing images. In *2020 IEEE 22nd International Workshop on Multimedia Signal Processing (MMSP)*, pages 1–6, 2020. xiv, 117, 123

[96] M. Trevisi, A. Akbari, M. Trocan, and Á. Rodríguez-vázquez and r. *Carmona-Galán,*. 30

[97] J. Tropp, M. Wakin, M. Duarte, D. Baron, and R. Baraniuk. Random filters for compressive sampling and reconstruction. In *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, volume 3, pages III–III, 2006. 13

[98] J. A. Tropp. The sparsity gap: Uncertainty principles proportional to dimension. *44th Annual Conference on Information Sciences and Systems (CISS), Princeton, NJ*, 10.:1–6, 2010. 51, 92

[99] J. A. Tropp and A. C. Gilbert. Signal recovery from random measurements via orthogonal matching pursuit. 53(12):4655–4666, Dec. 2007. 95

[100] J. A. Tropp, J. N. Laska, M. F. Duarte, J. K. Romberg, and R. G. Baraniuk. Beyond nyquist: Efficient sampling of sparse bandlimited signals. *IEEE Transactions on Information Theory*, 56(1):520–544, 2010. 13

[101] R. Tur, Y. C. Eldar, and Z. Friedman. Innovation rate sampling of pulse streams with application to ultrasound imaging. *IEEE Transactions on Signal Processing*, 59(4):1827–1842, 2011. 13

[102] A. S. Unde and D. P.P. Rate–distortion analysis of structured sensing matrices for block compressive sensing of images. *Signal Processing: Image Communication*, 65:115–127, 2018. 116

[103] M. Unser. Sampling-50 years after shannon. *Proceedings of the IEEE*, 88(4):569–587, 2000. 13

[104] V. Vapnik. The nature of statistical learning theory. 1999. 16

[105] S. S. Vasanawala, M. T. Alley, B. A. Hargreaves, R. A. Barth, J. M. Pauly, and M. Lustig. Improved pediatric mr imaging with compressed sensing. *Radiology*, 256(2):607–616, Aug 2010. 20529991[pmid]. 13

[106] P. G. Vaz, D. Amaral, L. F. R. Ferreira, M. Morgado, and J. ao Cardoso. Image quality of compressive single-pixel imaging using different hadamard orderings. *Opt. Express*, 28(8):11666–11681, Apr 2020. 30

[107] R. Venkataramani and Y. Bresler. Further results on spectrum blind sampling of 2d signals. In *Proceedings 1998 International Conference on Image Processing. ICIP98 (Cat. No.98CB36269)*, volume 2, pages 752–756 vol.2, 1998. 12

[108] M. Vetterli, P. Marziliano, and T. Blu. Sampling signals with finite rate of innovation. *IEEE Transactions on Signal Processing*, 50(6):1417–1428, 2002. 12

[109] A. Wyner and J. Ziv. "the rate-distortion function for source coding with side information at the decoder. " *IEEE Trans. Inf. Theory, vol. 22, no*, 22(1):1–10, Jan. 1976. 52, 97

[110] Z. Xu, L. Zhang, J. Shen, H. Zhou, X. Liu, J. Cao, and K. Xing. Mrcs: matrix recovery-based communication-efficient compressive sampling on temporal-spatial data of dynamic-scale sparsity in large-scale environmental iot networks. *EURASIP Journal on Wireless Communications and Networking*, 2019(1):18, Jan 2019. 116

[111] F. Yang, S. Wang, and C. Deng. Compressive sensing of image reconstruction using multi-wavelet transforms. In *2010 IEEE International Conference on Intelligent Computing and Intelligent Systems*, volume 1, pages 702–705, 2010. 19

[112] H. Ye, L. Tian, Q. Zhang, H. Wang, and S. Feng. Cmos image sensor with programmable compressed sensing. In *2015 IEEE 11th International Conference on ASIC (ASICON)*, pages 1–4, 2015. 3

[113] W.-K. Yu. Super sub-nyquist single-pixel imaging by means of cake-cutting hadamard basis sort. *Sensors*, 19(19), 2019. 30

[114] W.-K. Yu and Y.-M. Liu. Single-pixel imaging with origami pattern construction. *Sensors*, 19(23), 2019. 30

[115] X. Yuan and R. Haimi-Cohen. Image compression based on compressive sensing: End-to-end comparison with JPEG. 22(11):2889–2904, Nov. 2020. 52