

NEW WAVELET COEFFICIENT RASTER SCANNINGS FOR COMPRESSIVE IMAGING

Marco F. Duarte,¹ Sina Jafarpour,² and Robert Calderbank¹

¹Departments of Computer Science and Electrical and Computer Engineering

Duke University, Durham, NC 27708

² Department of Computer Science

Princeton University, Princeton, NJ 08544

E-mails: marco.duarte@duke.edu, sina@princeton.edu, robert.calderbank@duke.edu

ABSTRACT

The Delsarte-Goethals frame has been proposed for deterministic compressive sensing of sparse and compressible signals. Its performance in compressive imaging applications falls short of that obtained for arbitrary sparse vectors. Prior work has proposed specially tailored signal recovery algorithms that partition the recovery of the input vector into clustered and unclustered portions. In this paper we present a formal analysis of the Delsarte-Goethals frame that shows that its hampered performance in compressive imaging is due to the presence of clustered sparse vectors in its null space. Such a clustered structure is characteristic in commonly-used raster scanning of 2-D wavelet representations. We also show that a simple randomized raster scanning of the wavelet coefficient vector can remedy these shortcomings, allowing the use of standard recovery algorithms. Additionally, we propose an alternative deterministic raster scanning that achieves similar recovery performance. Experimental results confirm the improvements in recovery performance for both the noiseless and noisy measurement regimes.

Keywords— Compressive sensing, compressive imaging, Delsarte-Goethals frame, raster scanning, sparsity, wavelets

1. INTRODUCTION

In compressed sensing (CS) [5, 3], we wish to acquire a signal $f \in \mathbb{R}^C$ by taking its product with a matrix $\Phi \in \mathbb{R}^{N \times C}$, obtaining a measurement vector $y \in \mathbb{R}^N$. When $N \ll C$, this acquisition scheme effectively compresses the signal f . Since in this case the signal recover problem is ill posed, one must exploit prior information on the signal such as sparsity or compressibility. CS relies on the use of specially tailored signal recovery algorithms based on sparsity to recover the signal f from the measurements y and the CS matrix Φ . Most work in CS relies on random constructions on the matrix Φ ; that is,

The work is supported in part by NSF under grant DMS-0914892, by ONR under grant N00014-08-1-1110, and by AFOSR under grants FA9550-09-1-0643 and FA9550-09-1-0551. MFD is also supported by NSF Supplemental Funding DMS-0439872 to UCLA-IPAM, P.I. R. Caflisch. We thank Kangyu Ni for many useful discussions and for providing us with the implementation of the two-step recovery algorithm from [7, 8].

the entries of the matrix are drawn independently from a suitable probability distribution such as Gaussian or Rademacher. Such matrices have been shown to provide enough information about a K -sparse signal f through the measurements y when $N = \mathcal{O}(K \log C)$.

The Delsarte-Goethals Frame [2] (DGF) was proposed as a deterministic CS matrix construction that enables efficient recovery of almost all sparse signals without the use of randomness in the design of the measurement matrix. The DGF stands out because of its deterministic nature and a set of accompanying fast signal recovery algorithms [2]. The DGF has been shown to succeed for recovery of an overwhelming majority of sparse signals. It is also robust to noise, measurement loss, and other practical considerations. However, the DGF must be applied directly on a sparse or compressible vector. That is, the vector f measured using the CS matrix φ must be sparse or compressible in the canonical domain.

Initial results on the performance of the DGF for deterministic compressive imaging were presented by Ni et al. [7, 8]. While most natural images are not themselves sparse or compressible, transforms such as the the 2-D discrete wavelet transform are suitable to obtain sparse or compressible image representations. The DGF can then be applied directly to the wavelet coefficient vector of the image being acquired. It was hinted in [7] that the DGF may potentially have problems sensing images due to the characteristic clustered structure that appears under standard raster scanning of 2-D wavelet coefficients. Ni et al. proposed a two-stage recovery algorithm that addresses these concerns by exploiting the concentration of nonzero coefficients in the coarsest scales of the wavelet transform.

In this paper, we study the behavior of the DGF for such signals with clustered coefficient vectors in additional depth. Specifically, we show that the DGF performs suboptimally for recovery of sparse signals whose largest coefficients are concentrated (or clustered) within the coefficient vector.¹ We prove formally that there exist sparse vectors exhibiting such clustering within the null space of the DGF, which implies that there are clustered coefficient vectors that are provably unrecoverable from DGF measurements. Additionally, this behavior of the

¹It is natural to note that the fraction of sparse signals that are clustered is vanishingly small, which is part of the reason we are so interested in this matrix.

DGF severely hampers the recovery of coefficient vectors having their energy clustered in small groups. The 2-D wavelet coefficients of natural images are a prominent example of clustered vectors due to the large magnitude of coarse-scale wavelet coefficients and their concentration within a small portion of the coefficient vector.

To overcome the recovery issues in compressive imaging caused by this behavior, we show experimentally an improvement in performance afforded by altering the raster scanning of the coefficient vector so that the clustered structure of the nonzero coefficients is diluted. While a random raster scanning is the easiest approach, it would negate many of the attractive features due to the deterministic nature of the DGF implementation. As an alternative, we propose a deterministic alternative raster scanning that is effective at dissipating the clusters that appear in standard raster scanning of 2-D wavelet coefficients. This modification provides substantially improved performance over the standard use of the DGF [7, 8]. We verify this improvement experimentally for recovery of natural images from both noiseless and noisy measurements.

This paper is organized as follows. Section 2 provides background and briefly reviews related work. Section 3 provides our theoretical results on the performance of the DGF for deterministic compressive imaging, and Section 4 details our modified raster scanning schemes that provide improved performance. Section 5 ends the paper with validating experimental results.

2. BACKGROUND AND RELATED WORK

The Delsarte-Goethals Frame: The Delsarte-Goethals set $DG(m, r)$, $m, r \in \mathbb{Z}$, is a vector space containing $2^{(r+1)m}$ binary symmetric matrices of size $m \times m$ with the property that the difference of any two distinct matrices has rank at least $m - 2r$.

The Delsarte-Goethals frame [2] (DGF) φ is a measurement matrix of size $N = 2^m$ and $C = 2^m R$, with $R \in [1, 2^{m(r+1)}]$ an integer. Its rows can be indexed using elements $x \in \mathbb{F}_2^m$, represented as binary vectors of length m . Similarly, its columns can be indexed using ordered pairs (P, b) , where $P \in DG(m, r)$ and $b \in \mathbb{F}_2^m$. In this way, we label and define the entry of φ in row x and column (P, b) as $\varphi_{P,b}(x) = i^{xPx^\top + 2bx^\top}$. Here x^\top denotes the transpose of x . Note that all the arithmetic in the expressions $xPx^\top + 2bx^\top$ takes place in the ring of integers modulo 4, since the expression appears as an exponent for $i = \sqrt{-1}$. Given P, b , the vector $xPx^\top + 2bx^\top$ is a codeword in the Delsarte-Goethals code (defined over the ring of integers modulo 4). For a fixed matrix P , the 2^m columns $\{\varphi_{P,b}\}_{b \in \mathbb{F}_2^m}$ form an orthonormal basis Γ_P that can also be obtained by post-multiplying the Walsh-Hadamard basis by the unitary transform $\text{diag}[i^{xPx^\top}]$. In this way, the DGF can be written as a concatenation of bases $\varphi = [\Gamma_{P_1} \Gamma_{P_2} \dots]$, with the matrix aspect ratio $R = C/N$ determining the number of bases contained in φ .

The DGF is well suited for compressive sensing. In particular, when normalized to obtain unit-norm columns, its worst-case coherence value is $\mu(\varphi) = 1/\sqrt{N}$ and its spectral norm is $\|\varphi\|_2 = C/N$ (since it is a tight frame). These two proper-

ties guarantee that the matrix enables successful recovery of an overwhelming majority of sparse signals [2, 9]. In the sequel, we also write $DGF(m, r)$ for the DGF when the dependence on m, r must be made explicit.

The Haar Wavelet Basis: The Haar wavelet basis is arguably the simplest construction of a discrete wavelet system. The basis is completely defined by the scaling function, which has a constant value throughout its support, and its wavelet function, which has only two distinct nonzero values of the same magnitude, each covering half of the function's support. For C a power of two, the length- C Haar scaling function is given by

$$\phi(n) = \sqrt{1/C}, 0 \leq n \leq C - 1.$$

We also define multiscale Haar wavelet functions $\psi_{s,j}$; each function is labeled by its scale $s = 0, \dots, \log_2 C - 1$, and its offset $j = 0, \dots, 2^s - 1$:

$$\psi_{s,j}(n) = \begin{cases} \sqrt{2^s/C} & 2^{-s}Cj \leq n < 2^{-s}C(j+1/2), \\ -\sqrt{2^s/C} & 2^{-s}C(j+1/2) \leq n < 2^{-s}C(j+1), \\ 0 & \text{otherwise.} \end{cases}$$

This structure for the support of the wavelet (i.e., the location of its nonzero values) is known as a dyadic structure: the wavelet's support is of size $2^{-s}C$, and the offset is a multiple of its size. For simplicity, we denote by $\mathcal{D}_{s,j}$ the set of indices in the dyadic interval at scale s and offset j :

$$\mathcal{D}_{s,j} = \{2^{-s}Cj, 2^{-s}Cj+1, \dots, 2^{-s}C(j+1) - 1\}. \quad (1)$$

Raster scanning of wavelet representations: Wavelet coefficient vectors of natural images exhibit significant structure on the location of the largest-magnitude entries within the coefficient vector. The wavelet coefficients are usually raster scanned first by scale, and then by offset:

$$f_w = [s_{0,0} \ w_{0,0} \ w_{1,0} \ w_{1,1} \ w_{2,0} \ \dots \ w_{3,0} \ \dots].$$

Here $s_{0,0}$ denotes the scaling coefficient, and $w_{s,j}$ denotes the wavelet coefficient at scale s and offset j . When a partial-level wavelet transform is used, the raster scan begins with the multiple scaling coefficients, rasterized into a vector form, followed by the wavelet coefficients ordered as in (1).

Wavelet transforms of natural images have certain properties that affect the configuration of their raster-scanned representations. First, the coefficients decay in magnitude as the scale increases, resulting in a concentration of the largest coefficients in the beginning of the vector. Furthermore, the number of 2-D wavelet coefficients of an image at scale s is equal to $3 \cdot 4^s$, implying that the largest coefficients of the signal are concentrated within a very small portion of the vector.

Related Work: Leveraging the properties of wavelet representation vectors, Ni et al. proposed a two-stage recovery algorithm [7]. The first stage estimates the coefficients at coarsest scales using a submatrix projection, which due to their location entails a simple projection into the basis Γ_{P_1} . The second stage estimates the remainder of the coefficient vector using a greedy sparse signal recovery algorithm. The incremental nature of the

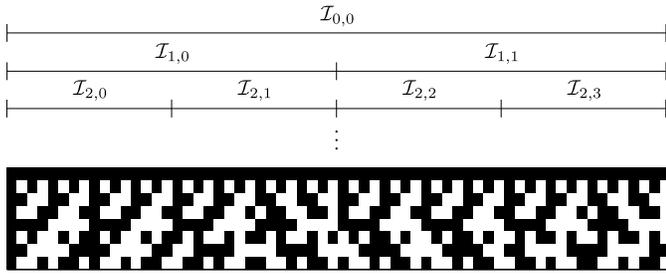


Fig. 1. Example dyadic column intervals for a DGF of size $C = 64$, $N = 8$ ($m = 3$).

proposed estimation algorithm reduces the number of nonzeros of the signal being estimated, resulting in an improvement over standard CS recovery of the full coefficient vector.

3. THE NULL SPACE OF THE DGF

In this section, we study the behavior of the DGF when applied to clustered coefficient vectors. More specifically, we focus on linear combinations of Haar wavelets; these functions are piecewise constants and have dyadic supports, which line up with subgroups of the index sets (P, b) assigned to columns of the DGF. The measurement vectors for these functions correspond to averages of the columns in subgroups of the index set which, as we will show, vanish under specific conditions. Thus, these functions are included in the null space of the DGF.

We begin by defining notation that will simplify the exposition. For each dyadic wavelet scale-offset pair (s, j) , we let the set $\mathcal{I}_{s,j} \subseteq DG(m, r) \times \mathbb{F}_2^m$ denote the pairs (P, b) for the columns at the positions contained in the dyadic interval $\mathcal{D}_{s,j}$; we call $\mathcal{I}_{s,j}$ a *dyadic column interval*. Some examples are shown in Figure 1. We provide some properties of the column indices (P, b) contained in the subsets $\mathcal{I}_{s,j}$.

Proposition 1. *If $s \geq \log_2 R$, then $\mathcal{I}_{s,j} = \{P_{s,j}\} \times \mathcal{F}_{s,j}$, where $P_{s,j}$ is a fixed matrix in $DG(m, r)$, and $\mathcal{F}_{s,j} = a_{s,j} \oplus \mathbb{F}_2^{\log_2 C - s}$, with $a_{s,j} \in \mathbb{F}_2^{s - \log_2 R}$.*

In words, $\mathcal{F}_{s,j}$ is a subset of \mathbb{F}_2^m whose elements share the $s - \log_2 R$ most significant bits; fluctuations on later bits span the subset. Thus, the subset $\mathcal{I}_{s,j}$ is defined by the matrix $P_{s,j}$ and the “header” $a_{s,j}$ containing the fixed most significant bits of b over $\mathcal{I}_{s,j}$.

We now consider the linear combination of two Haar wavelets projected by the DGF at the same scale $\eta_{s,j_1,j_2,\pm}(x) = \varphi(\psi_{s,j_1} \pm \psi_{s,j_2})$, and look to determine which combinations of wavelets belong in the null space of φ , $\mathcal{N}(\varphi)$. The following theorem is proven in [6].

Theorem 1. *Let φ denote the DGF(m, r) and write $x_{h,l}$ and $x_{f,l}$ for the first and last l entries of the vector $x \in \mathbb{F}_2^m$, $1 \leq l \leq m$, respectively. For $0 \leq s < \log_2 C$ and $0 \leq j_1, j_2 < 2^s$, $j_1 \neq j_2$, we have $\eta_{s,j_1,j_2,+} = 0$ (i.e., $\psi_{s,j_1} \pm \psi_{s,j_2} \in \mathcal{N}(\varphi)$) if*

- $s < \log_2 R$ (for all j_1, j_2);
- $\log_2 R \leq s \leq \log_2 R + 2r$, $a_{s,j_1} = a_{s,j_2}$ (that is, the two dyadic wavelet intervals are at the same position within the domain of the corresponding bases $\Gamma_{P_{s,j_1}}, \Gamma_{P_{s,j_2}}$) and $x(P_{s,j_1} - P_{s,j_2})x^\top = 0 \pmod{4}$ for all $x \in \mathbb{F}_2^m$ such that $x_{f, \log_2 C - s} = 0$; or
- $s \geq \log_2 R$ and $x(P_{s,j_1} - P_{s,j_2})x^\top = 2(a_{s,j_2} - a_{s,j_1})x_{h, s - \log_2 R}^\top \pmod{4}$ for all $x \in \mathbb{F}_2^m$ such that $x_{f, \log_2 C - s} = 0$.

The conditions also hold for $\eta_{s,j_1,j_2,-} = 0$ by adding a term of 2 in the equalities $\pmod{4}$.

The theorem shows that there exist vectors with $2N$ nonzeros, composed of the sum or difference of two Haar wavelets, that belong in the null space of the DGF. We have also found experimentally vectors with significantly fewer nonzeros that are in $\mathcal{N}(\varphi)$. As an example, for the DGF(15, 0) and the DGF(13, 0), we have found vectors of sparsity $\|f\|_0 = 1024$ in their null spaces, composed of the difference of two Haar wavelets at scales $s = 9$ and $s = 7$, respectively.

The proof of Theorem 1 can be extended to real-valued adaptation using (a) zero-diagonal DG matrices P , as used in [7], and (b) by generating a larger matrix from the DGF via the Gray map; see [6] for details. We note in passing that the two-step approximation of [7] separates the recovery of the clustered portion from the coefficient vector, covering its first N entries, from the recovery of the unclustered portion. Thus, the approach alleviates the issues described in the theorem by recovering the clustered portion without using sparsity-based tools.

4. ALTERNATIVE RASTER SCANNINGS

A simple approach to alleviate the issues of the DGF with clustered sparse and compressible coefficient vectors is to modify the raster scanning of the coefficients so that the clustered structure is dissipated. It is easy to show that a random raster scanning will dissipate the clustering with high probability. However, a random raster scanning is not completely compatible with the DGF, as it nullifies several useful properties of its deterministic nature, such as scalability in computation to the signal length and implementation simplicity.

As an alternative, we introduce a deterministic raster scanning that evenly distributes the content of coefficient clusters and scales across the bases Γ_{P_i} composing the DGF. We start from the standard raster scanning for wavelet coefficients and apply a permutation described by the mapping

$$p(n) = (n \bmod R)N + \lfloor n/R \rfloor,$$

which provides a bijective mapping of the set $\{0, \dots, C - 1\}$ onto itself. In words, the mapping takes each set of consecutive coefficients of size R and spreads it across the R bases $\{\Gamma_{P_i}\}$; the subsequent blocks are partitioned similarly and ordered lexicographically. The resulting raster scanning separates the 2-D wavelet coefficients at a given scale and adjacent offsets by a

Algorithm	SNR (dB)	Time (s)
TSA	21.74	1008
BP + Random Raster	23.60	820
IHT + Random Raster	22.25	804
BP + Deterministic Raster	23.52	822
IHT + Deterministic Raster	20.98	813

Table 1. Performance and computational cost of signal recovery algorithms. The use of the proposed raster scanings significantly improves image recovery performance.

Algorithm	$\sigma = 0.01$	$\sigma = 0.05$	$\sigma = 0.1$
TSA	21.64	19.51	16.37
BP + Random Raster	23.41	20.71	18.18
IHT + Random Raster	22.24	20.82	17.20

Table 2. Performance of signal recovery algorithms for noisy measurements. The use of random raster scanning significantly improves the performance under the noisy measurement regime.

distance of N entries within the coefficient vector. Furthermore, neighboring coefficients in this raster scan either correspond to different scales or to coefficients for wavelets with offsets whose difference is equal to N .

5. EXPERIMENTAL RESULTS

In this section, we present some experimental results that validate the improvements in performance of deterministic compressive imaging afforded by the raster scanings proposed in Section 4. Our experiments use an MRI image of size $\mathcal{C} = 512 \times 512$. We use the Daubechies-8 discrete wavelet transform to obtain compressible coefficients for the image, and we set the number of measurements to $N = 65536$, providing a compression ratio of 25%. We test the two-step approximation (TSA) algorithm from [7], basis pursuit (BP) [4], and iterative hard thresholding (IHT) [1] using the standard $DGF(16, 0)$. We also test BP and IHT under the proposed raster scanings described in Section 4. First, we compare the quality and computational cost of the different approaches; the results, shown in Table 1, show that under similar computational cost, the proposed matrix designs provide significant improvement in the quality of recovery. Representative example recovered images are shown in Figure 2. We also compare the quality of recovery for the different raster scanning approaches when independent and identically distributed Gaussian noise of variance σ^2 is added to the measurements for several values of σ . The resulting signal-to-noise ratios (SNRs), shown in Table 2, show once again that the proposed raster scanings significantly improve the performance of image recovery.

6. REFERENCES

[1] T. Blumensath and M. E. Davies. Iterative hard thresholding for compressed sensing. *Appl. Comput. Harmonic Analysis*, 27(3):265–274, Nov. 2008.

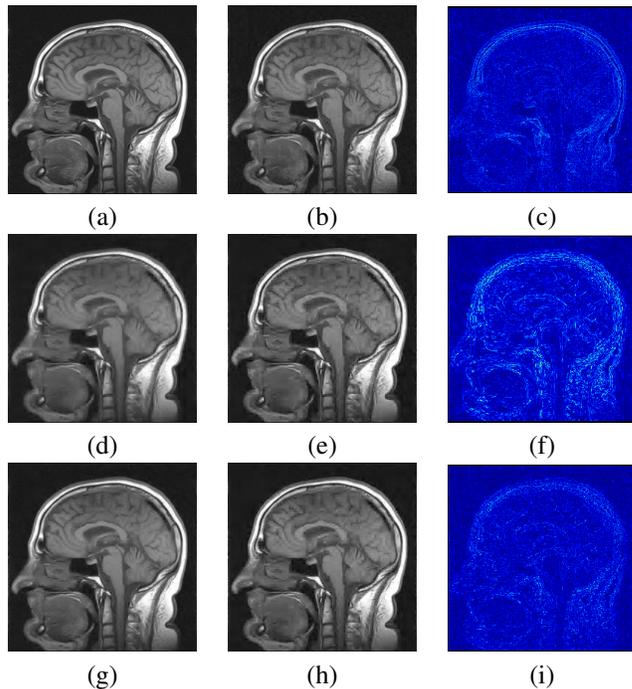


Fig. 2. Comparison of CS recovery performance. (a) Original image. (b) Two-step approximation [7]. (c) Difference image of (a), (b). (d-e) BP and IHT with standard DGF. (f) Difference image of (a), (d). (g-h) BP and IHT using random raster scanning. (i) Difference image of (a), (g).

[2] R. Calderbank, S. Howard, and S. Jafarpour. Construction of a large class of deterministic sensing matrices that satisfy a statistical isometry property. *IEEE J. Selected Topics in Signal Processing*, 4(2):358–374, Apr. 2010.

[3] E. J. Candès. Compressive sampling. In *Proc. International Congress of Mathematicians*, volume 3, pages 1433–1452, Madrid, Spain, 2006.

[4] S. Chen, D. L. Donoho, and M. Saunders. Atomic decomposition by basis pursuit. *SIAM J. Scientific Computing*, 20(1), 1998.

[5] D. L. Donoho. Compressed sensing. *IEEE Trans. Info. Theory*, 52(4):1289–1306, September 2006.

[6] M. F. Duarte and R. Calderbank. The null space of the Delsarte-Goethals frame. Technical Report TR-2010-09, Duke University, Department of Computer Science, Durham, NC, Feb. 2011.

[7] K. Ni, S. Datta, P. Mahanti, S. Roudenko, and D. Cochran. Efficient deterministic compressed sensing for images with chirps and Reed-Muller sequences. *Preprint*, 2010.

[8] K. Ni, P. Mahanti, and S. Roudenko. Stability of efficient deterministic compressed sensing for images with chirps and Reed-Muller sequences. *Preprint*, 2010.

[9] J. A. Tropp. On the conditioning of random subdictionaries. *Appl. Comp. Harmonic Analysis*, 25(1):1–24, July 2008.