Statistically Matched Wavelet based texture synthesis in a Compressive Sensing Framework

Mithilesh Kumar Jha[1,*], Brejesh Lall[1], Sumantra Dutta Roy[1]

Department of Electrical Engineering, Indian Institute of Technology Delhi

Abstract

This paper proposes an Statistically Matched Wavelet based textured image coding scheme for efficient representation of texture data in a compressive sensing (CS) framework. Statistically Matched Wavelet based data representation causes most of the captured energy to be concentrated in the approximation sub-space, while very little information remains in the detail sub-space. We encode not the full-resolution Statistically Matched Wavelet subband coefficients, but only the approximation subband coefficients (LL) using standard image compression scheme like JPEG2000. The detail subband coefficients i.e HL, LH and HH are jointly encoded in a compressive sensing framework. Compressive sensing technique has proved that it is possible to achieve a sampling rate lower than the Nyquist rate with acceptable reconstruction quality. The experimental results demonstrate that the proposed scheme can provide better PSNR and MOS with a similar compression ratio than the conventional DWT-based image compression schemes in a CS framework and other Wavelet based texture synthesis scheme like HMT-3S.

Keywords: Compressive Sensing(CS), Statistically Matched Wavelet, Texture Representation, Texture Synthesis, Noiselet Transform

1. Introduction

Texture data contain spatial, temporal, statistical and perceptual redundancies. Representing texture data using standard compression schemes like MPEG-2 [1] and H.264 [2] is not efficient, as they are based on Shannon-Nyquist sampling [3] and do not account for perceptual redundancies. They are often resource consuming (as they acquires too many sample) due to its fine details in textured image and high frequency content. Variety of applications in computer vision, graphics and image processing (such as robotics, defence, medicine and geo-sciences) demand better compression with good perceptual reconstruction quality, instead of bit-accurate (high PSNR) reconstruction. This is because

*Corresponding Author

Email addresses: jham73@gmail.com (Mithilesh Kumar Jha), brejesh@ee.iitd.ac.in (Brejesh Lall), sumantra@ee.iitd.ac.in (Sumantra Dutta Roy)

[1]Contact: Department of Electrical Engineering, Indian Institute of Technology Delhi, Hauz Khas, New Delhi - 110016, INDIA.

the human brain is able to decipher important variations in data at scales smaller than those of the viewed objects. Ndjiki-Nya et al. [4, 5, 6, 7, 8], Bosch et al. [9, 10], Bryne et al. [11, 12], and Zhang and Bull [13, 14] have proposed techniques to re-construct visually similar texture from sample data. Statistically Matched Wavelet [15] is aimed at designing a filter bank that matches a given pattern in the image and can better represent the corresponding image as compared to other wavelet family.

Compressive Sensing (CS) technique [16] has proved that it is possible to achieve a sampling rate lower than the Nyquist rate [3] with acceptable reconstruction quality. Leveraging the concept of transform coding, compressive sensing enables a potentially large reduction in sampling and computation costs for sensing signals that have sparse or compressible representation (by a sparse representation, we mean that for a signal of length N, we can represent it with $K << N$ nonzero coefficients). Compressive sensing framework opens a new research dimension in which most of the sparse signals can be reconstructed from a small number of measurements (M), using algorithms like convex optimization, greedy methods and iterative thresholding [17, 18]. Significant theoretical contributions have been published on the compressive sensing in recent years [19, 16, 17] for image processing applications [20, 21, 22, 23, 24]. Compressive Sensing framework mainly consists of three stages, i.e. sparsification by transformation, measurement (projection) and optimization (reconstruction). Designing a good measurement matrix with large compression effects and designing a good signal recovery algorithm, are the two major challenges for applying CS technique in image compression.

1.1. The prior works and motivation

Existing methods of textured image compression scheme can be broadly classified as parametric or non-parametric. Non-parametric approaches can be applied to a wide variety of textures (with irregular texture patterns) and provide better perceptual results. However, these schemes are often computationally more complex. Parametric approaches can achieve very high compression at low computational cost. However, these techniques are not effective for structured textures such as those with primarily non-stationary data content.

Non-parametric approaches are pixel-based or patch-based. Efros and Leung [25] proposed pixel-based non-parametric sampling to synthesize texture. Wei and Levoy [26] further improve the above using a multi-resolution image pyramid based on a hierarchial statistical method. A limitation of the above pixel-based methods is an incorrect synthesis owing to incorrect matches in searching for similar statistics. Patch-based methods overcome this limitation by considering features matching patch-boundaries with multiple pixel statistics. People generally use Markov Random Field (MRF) models for texture analysis [26, 7]. The popular choice for texture synthesis is a patch-based graph cut [27]. Region based texture representation and synthesis algorithms [28, 29, 30] have been explored recently to address the limitations of block based representation in handling the homogenous texture and blocking artifacts. Byrne et al. [11, 12] have demonstrated region-based synthesis structure using a morphological and spectral image representation technique [31]. Region based representation has also been explored using multi-resolution wavelet based decomposition as reported in [32, 33, 34]. However that work is limited to document class of images.

Recent work in the parametric approach is typically based on Auto Regressive (AR) [35] or Auto Regressive Moving Average (ARMA)-based modelling [36, 37, 38]. AR- and

ARMA-based models in texture synthesis enable blocks to be selectively removed at the encoding stage, and reconstructed at the decoder with acceptable perceptual quality. AR- and ARMA-based approach are suitable for the textures with stationary data, like steady water, grass and sky, however they are not suitable for structured texture with non-stationary data as blocks with non-stationary data are not amenable to AR modelling. Further, they are block-based approaches, and blocking artifacts can appear in the synthesised image. Portilla and Simoncelli [39] propose a statistical model for texture images based on joint statistical constraint on the wavelet coefficients. In this author proposed an algorithm for synthesizing textures using sequential projections on to the constraints surfaces, however the model is the choice of the statistical constraints and not suitable for texture with structural pattern. Wavelet-domain hidden Markov tree (HMT-3S) [40] were recently used for texture analysis and synthesis, where it was assumed that three subbands of the two-dimensional DWT, i.e. HL, LH and HH, are independent. The HMT-3S adds dependencies across sub-bands by treating jointly the three hidden elements across the three orientations, however HMT-3S is established in the nonredundant DWT and is inferior to redundant DWT for statistical modelling. Also, for structural textures with regular patterns, both HMT and HMT-3S fail to reproduce the regular patterns.

In this paper, we propose an Statistically Matched Wavelet based texture data representation and a Compressive Sensing based texture synthesis scheme (*henceforth we refer the proposed scheme in this paper as SMWT-CS*). Statistically Matched Wavelet based representation [41] causes most of the captured energy to be concentrated in the approximation sub-space, while very little information is retained in the detail sub-space. We encode not the full-resolution Statistically Matched Wavelet subband coefficients (as normally done in a standard wavelet-based image compression), but only the approximation subband coefficients (LL) using standard image compression scheme like JPEG2000 [42] (which accounts for 1/4th of the total coefficients and can be represented using fewer bits). The detail subband coefficients i.e HL, LH and HH (which accounts for 3/4th of the total coefficients) are jointly encoded in a compressive sensing framework, and can therefore be represented with fewer measurements. Quality matrix is an essential component for assessing the performance of an image compression system. We have computed Mean Opinion Score (MOS) for subjective assessment as it captures the visual perception of human subjects. We have also computed PSNR for objective assessment of the quality for comparative study.

Table 1. provide a comparison of existing texture analysis and synthesis scheme. We have included our proposed scheme to put the work in perspective. As can be seen the proposed scheme is more efficient than the existing schemes. The main difference between a conventional wavelet based compression scheme in a CS framework (DWT-CS) [43, 21, 44] and our approach is the use of Statistically Matched Wavelet based sparsification [15, 41], as opposed to generic DWT used in previous works. The existing schemes use regular DWT that is not able to fully exploit the signal properties (orthogonal/nonreduandant DWT cannot well preserve the regularity and/or periodicity [45]) and hence sparsity that results is sub-optimal. CS optimization is done over detail subspace allows scalability in choosing measurement vector size providing a trade-off between compression and perceptual reconstruction quality. The other novelty is that we apply standard coding on the low resolution part (LL subband) of the information, because due to the nature of wavelet decomposition (this is true especially for multi-level wavelet

Table 1: Comparative study of texture analysis and synthesis schemes

Proposed Work	**Texture Analysis and Synthesis technique**	**Texture type**	**Quality assessment**	**Limitation**	**Complexity**
Wang and Adelson [**29**]	Affine warping	Rigid textures	None	Not suitable for structural texture	Medium
Dumitras and Haskell [**30**]	Steerable pyramids	Rigid textures	None	Not suitable for structural texture	Medium
Ndjiki-Nya and Wiegand [**8**]	Perspective warping	Rigid and non-rigid textures	Yes	Prone to propagation error	High
Zhang and Bull [**14**]	ARMA	Rigid and non-rigid textures	Yes	Not suitable for structural texture	Medium
Portilla and Simoncelli [**39**]	Wavelet-based (Joint-statistics)	Rigid and non-rigid textures	None	Not suitable for structural texture	Medium
Fan and Xia **[40]**	Wavelet-based (HMT-3S)	Rigid and non-rigid textures	None	Not suitable for structural texture	Medium
Our scheme (SMWT-CS)	Compressive Sensing and Statistically Matched Wavelet	Rigid and non-rigid textures	Yes (PSNR and MOS)	CS optimization still evolving	Medium

decomposition); the low frequency information is not sparse and hence not amenable for CS based encoding. Our proposed scheme can be easily integrated into an existing encoding framework [35] with texture analysis and synthesis being performed by the proposed scheme. So far, multi-resolution based representation has been used in the domain of documentation class of image, however no work has reported using statistically matched wavelet and compressive sensing based analysis and synthesis technique for a generic texture class images for compression with subjective reconstruction quality, to the best of our knowledge.

1.2. Overview of the scheme

Fig. 1 gives an overall view of the proposed scheme. In this figure, the texture analyzer block decomposes the input textured image into approximation (LL) and detail (HL, LH, HH) subbands using Statistically Matched Wavelet based image representation [15]. The basic idea is to design a Statistically Matched Wavelet filter bank using source data and decompose the input textured image into approximation and detail sub-space [41]. Standard image encoder like JPEG2000 [42] is used to encode approximation subband coefficients (LL). The detail subband coefficients i.e HL, LH and HH are jointly encoded in a compressive sensing framework. The proposed SMWT-CS encoder block performs compressive measurement using Noiselet transform [46] over detail sub-space coefficients (joint representation for HL, LH and HH sub-band coefficients) along with quantization and entropy coding. For a K-sparse signal x, we compute the measurement vector Mx1, which is much smaller in length than the length of signal (Nx1) and therefore the compression is guaranteed. Texture synthesis block does the compressive sensing optimization to synthesize the samples from detail sub-space measurements, using convex optimization [19, 16], i.e. l_1-norm minimization with equality constraint. Combining the

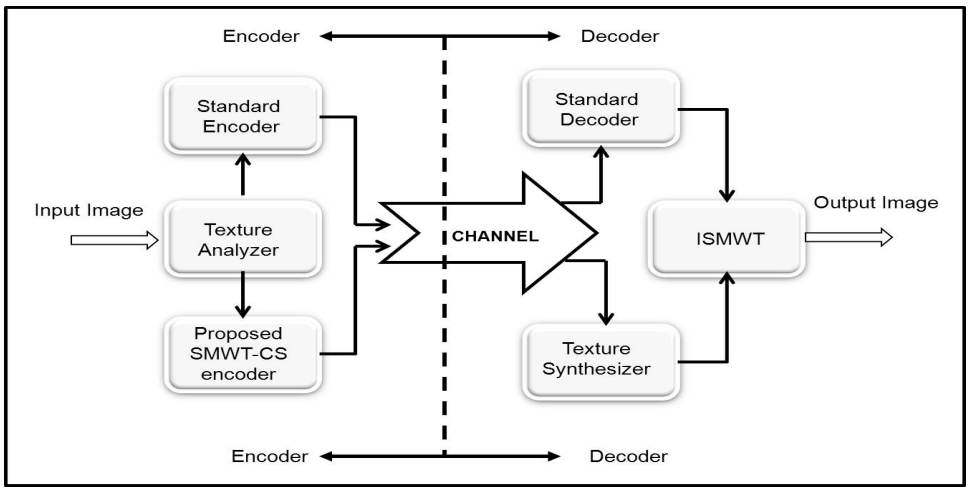


Figure 1: Overview of analysis and synthesis based image compression

decoded samples from the standard decoder and the texture synthesizer and doing the inverse of Statistically Matched Wavelet transform, a synthesized image is reconstructed.

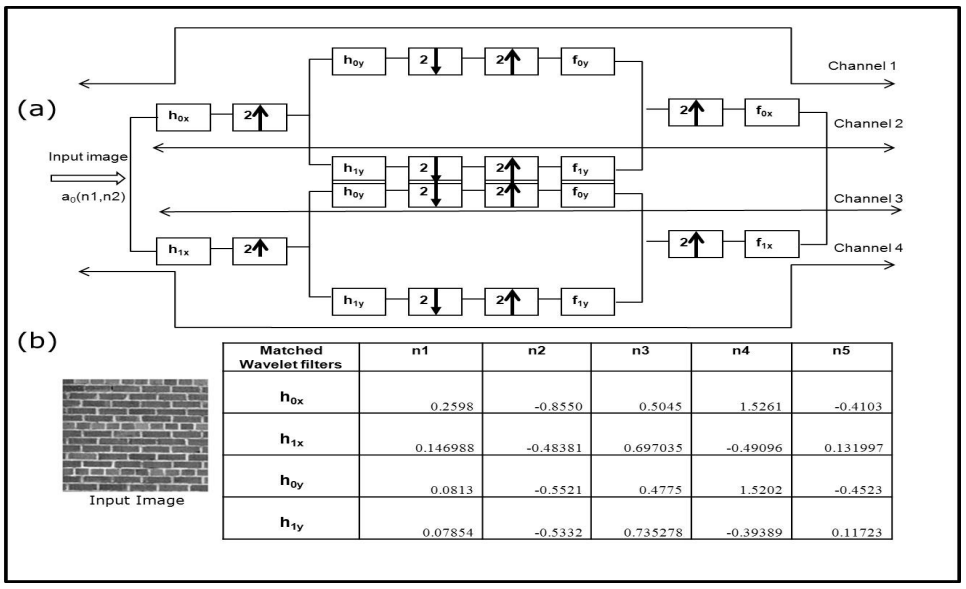
The rest of the paper is organized as follows: Section 2. provides an overview of statistically matched wavelet based texture representation scheme. Section 3. presents the design and implementation of the compressive measurements and encoding framework while Section 4. provides a detailed view of synthesis framework. The experimental results are discussed in Section 5. followed by conclusion in Section 6.

2. Statistically Matched Wavelet based Texture Representation

Gupta et al. [15] propose Statistically Matched Wavelet for the estimation of wavelets that is matched to a given signal in the statistical sense. This concept is further extended for image data and Fig. 2 shows a 2-D two-band separable kernel wavelet system (with an example of estimated matched wavelet filters for brickwall image). In this figure, x and y represents the horizontal and vertical directions respectively. The scaling filter is represented by f_{0x} and f_{0y} while the wavelet filter is represented by f_{1x} and f_{1y} corresponding to the horizontal and vertical direction. The dual of them are represented by h_{0x}, h_{0y}, h_{1x} and h_{1y}. The input to this system is a 2-D signal (an image in our framework). The output of Channel-1 in Fig. 2 is called approximation sub-space or scaling sub-space while the outputs of other three channels (Channel-2,3,4 in Fig. 2) are called detail sub-space. This system is designed as biorthogonal wavelet system so that it satisfies the conditions (Eq. 1, 2, 3 and 4) for perfect reconstruction of the two-band filter bank.

$$h_{1i}(n) = (-1)^n f_{0i}(d-n) \tag{1}$$

$$f_{1i}(n) = (-1)^n h_{0i}(d-n) \tag{2}$$



Matched Wavelet filters	n1	n2	n3	n4	n5
h_{0x}	0.2598	-0.8550	0.5045	1.5261	-0.4103
h_{1x}	0.146988	-0.48381	0.697035	-0.49096	0.131997
h_{0y}	0.0813	-0.5521	0.4775	1.5202	-0.4523
h_{1y}	0.07854	-0.5332	0.735278	-0.39389	0.11723

Figure 2: Statistically Matched Wavelet Estimation (a) Matched Wavelet filter bank (output of Channel 1 represents approximation sub-space while output of Channel-2, 3, 4 are detailed sub-space) (b) Estimated analysis filter coefficients for brickwall test image.

$$\sum_n h_{0i}(n-2m1)f_{0i}(n-2m2) = \delta(m1-m2), \qquad \forall m1, m2 \in Z \tag{3}$$

$$\sum_n h_{0i}(n)h_{1i}(n) = 0. \tag{4}$$

where i can be x or y, respective to horizontal or vertical direction and d is any odd delay. f_{0i}, is the scaling filter while f_{1i}, is the wavelet filter. h_{0i} and h_{1i} are the dual of scaling and wavelet filters respectively.

The optimization criterion used for matched wavelet is the minimization of the energy in the detail sub-space [41]. If $a(x,y)$ is a 2-D image signal and $\hat{a}(x,y)$ represents the 2-D image reconstructed using only the output of channel 2, 3 and 4 (detail sub-space) in Fig. 2, then the error function $e(x,y)$ can be defined as:

$$e(x,y) = a(x,y) - \hat{a}(x,y) \tag{5}$$

To ensure that the maximum input signal energy moves to approximation sub-space, the energy E in the difference signal $e(x,y)$ should be maximized with respect to both x and y direction filters. This leads to the following set of equations [41] :

$$\sum_k h_{1_x}[\sum_m a_{0_x}(2m+k)a_{0_x}(2m+r)] = 0$$
$$For \quad r = 0, 1, 2, ..., j-1, j+1, ...N-1 \tag{6}$$

$$\sum_{k} h_{1_y}[\sum_{m} a_{0_y}(2m+k)a_{0_y}(2m+r)] = 0$$
$$for \quad r = 0, 1, 2, ..., j-1, j+1, ...N-1 \tag{7}$$

Here the $j^t h$ filter weight is kept constant leading to a close form expression. These are a set of N-1 linear equations in filter weight and can be solved simultaneously. All rows of image are placed adjacent to each other to form a 1-D signal having variations in horizontal direction only and represented as a_{0_x} in Eq. 6. Similarly all the columns of image are placed together to form 1-D signal having variations in vertical direction only and represented as a_{0_y} in Eq. 7. The solution of Eq. (6) and Eq. (7) gives the analysis high pass filters (wavelet filters) h_{1_x} and h_{1_y}. From these other filters (analysis low pass, synthesis high pass and synthesis low pass) are computed using finite impulse response perfect reconstruction bi-orthogonal filter bank design as presented with Eq. 1, Eq. 2, Eq. 3 and Eq. 4.

The estimation of scaling and wavelet function is done separately for the variations in the horizontal and vertical directions of the input image. Passing the input image through the 2-D statistically matched wavelet filter bank, we get the output as sub-sampled image corresponding to approximation and detail sub-space. Fig. 3, demonstrates the decomposition of one of the input tests sequence in approximation and detail sub-space and Fig. 4 gives the distribution of the matched wavelet coefficients in approximation and detail sub-space (horizontal, vertical and diagonal subspace). As one can observe from Fig. 3 and Fig. 4, most of the information is represented by approximation sub-space alone and very little information is present in the detail sub-space. In the rest of the paper we refer to the approximation or scaling sub-space (output of channel 1 in Fig.(2)) as ***the candidate data for host encoding*** and the detail sub-space (sum of output of channel 2, 3 and 4 in Fig.(2)) as ***the candidate data for CS measurements and synthesis***

3. Compressive Measurement and Encoding

Compressive sensing exploits the fact that most natural or artificial signals are sparse in some domain and hence compressible. A real valued signal $x \in \mathbf{R}^N$, can be represented as a function of basis vectors as follows [16, 17] :

$$\mathbf{x} = \sum_{i}^{N} s_i \psi_i \qquad or \qquad \mathbf{x} = \Psi \mathbf{s} \tag{8}$$

Where $\mathbf{x}$ and $\mathbf{s}$ are Nx1 column vectors, Ψ is an NxN sparsifying basis matrix. The signal x is called K-sparse if it can be represented as a linear combination of only K-basis vectors, i.e. only K elements of the vector $\mathbf{s}$ are non-zero. The signal x can be treated as compressible if it can be well approximated by a signal with only K ($K << N$) nonzero coefficients. Fig. 5 shows sparse representation of subbands coefficients corresponding to brickwall and escalator test sequences. Significant wavelet coefficients in the figure are represented by blue pixels while all other non-significant coefficients are shown in white. One can observe that most of the detail subbands coefficients are close to zero and

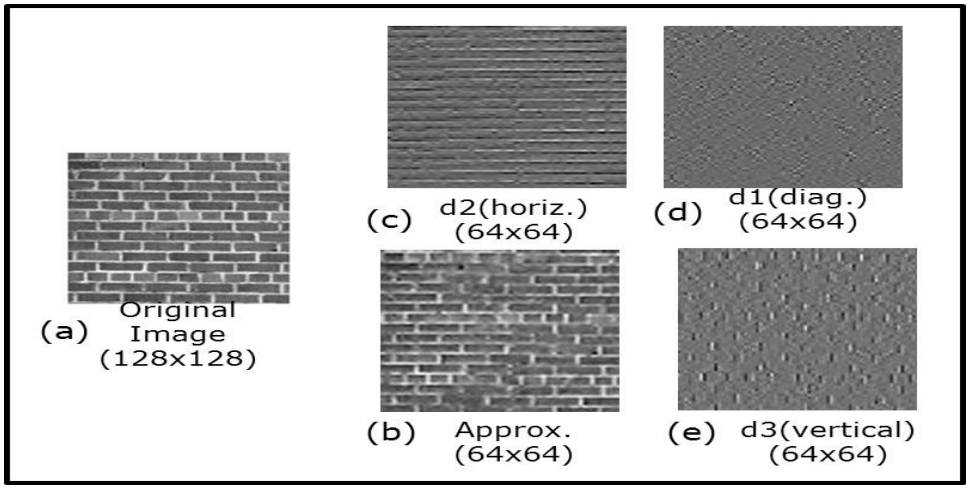


Figure 3: Illustration of statistically matched wavelet decomposition (a)Input image,(b) approximation sub-space image,(c to e) detailed sub-space image respective to horizontal, diagonal and vertical sub-bands.

therefore CS can be best utilised to exploit this sparsity for encoding detailed subbands coefficients. Compressive measurement is computed through linear projection as shown in Eq. 9.

$$\mathbf{y} = \Phi\mathbf{x} = \Phi\Psi\mathbf{s} = \Theta\mathbf{s} \tag{9}$$

Here $\mathbf{y}$ is an Mx1 measurement vector where $M < N$. Θ is an MxN measurement matrix. The matrix Φ represents a dimensionality reduction matrix, i.e. it maps $\mathbf{R}^N$ to $\mathbf{R}^M$, where M is typically much smaller than N. The main challenges with CS theory is that how should we design the sensing matrix Φ so that it preserves the information in the signal x. Several theories have been proposed in the literature for reconstructing x from y, if Θ satisfies a Restricted Isometry Property (RIP) [16]. We have used Statistically Matched Wavelet for the sparsifying matrix (Ψ) and Noiselet [46] for sensing matrix (Φ) for compressive measurements according to Eq. 9. Measurements from Noiselet have been chosen because they are highly incoherent with the considered sparse domain and RIP tends to hold for reasonable values of M. In addition, noiselet comes with very fast algorithms and just like the Fourier transform, the noiselet matrix does not need to be stored to be applied to a vector. This is of crucial importance for efficient numerical computations without which applying CS can be very complex. The CS measurement matrix created is orthogonal and self-adjoint, thus being easy to manipulate. Scalar quantization over the measurement vector is proposed to achieve better compression ratio.

It is important to note that Eq. 9 is ill-conditioned since there are more unknowns than the number of equations as $M<N$, however it has been shown that if the signal x is K-sparse and the locations of the K non-zero elements are known then the problem can be solved provided $M \geq K$ through a simplified linear equation by deleting all those columns and elements corresponding to zero or non-significant elements. Fig. 6

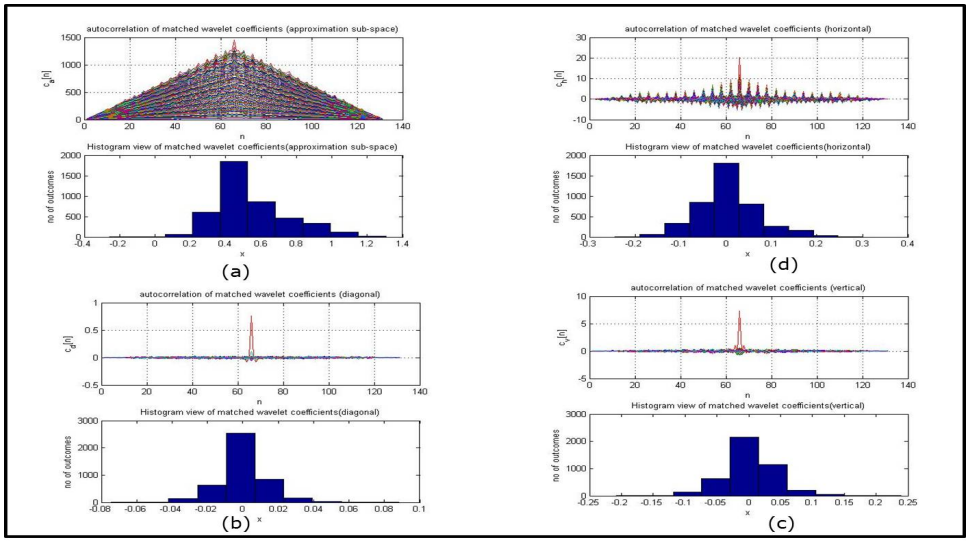


Figure 4: Illustration of statistically matched wavelet coefficients distribution in approximation and detail sub-space, (a)Histogram and auto co-relation plot of Matched wavelet coefficients in approximation sub-space, (b to d)Histogram and auto co-relation plot of Matched wavelet coefficients in detail subspace respective to diagonal, vertical and horizontal sub-bands.

gives an overall block diagram of the proposed encoder. Bit mapper block is responsible for generating a final bit stream encompassing encoded data from standard codecs for approximation sub-space and CS measurements for detail sub-space. The algorithm at the end of this section provides the implementation summary of the proposed encoder framework.

$$y = y/Q; \quad y = round(y); \quad y = y * Q, \ where \ Q, \ is \ the \ quantization \ factor; \tag{10}$$

Algorithm : CS measurement and encoding

1. Design a 2-D separable kernel filter bank using Statistically Matched Wavelet [Detailed in Sec. 2]
   (a) Decompose input image into approximation and detail subbands coefficients
      i. Use Wavelet function from MATLAB "wavecut" to represent approximation and detail coefficients
2. Design a CS measurement matrix Φ using Noiselet Transform [46] and detail subbands coefficients
3. Do quantization of the CS measurements using Eq. 10
4. Do Entropy coding of the quantized CS measurements
5. Combine standard and CS encoded bit streams

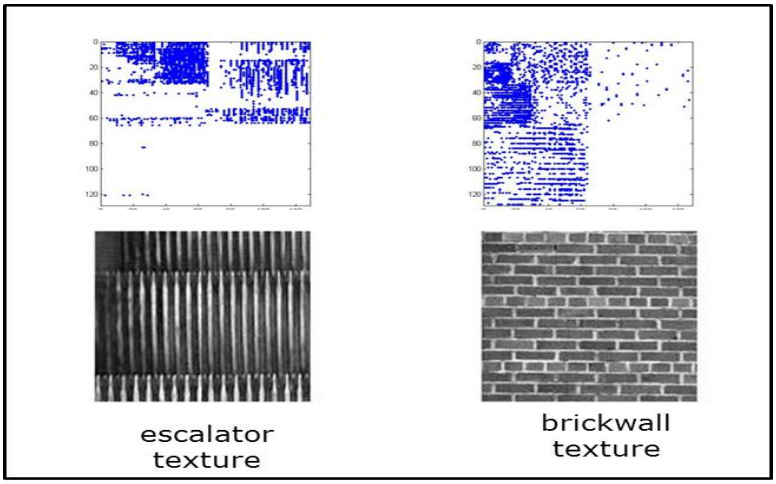


Figure 5: Sparse representation of Statistically Matched Wavelets sub-bands coefficients for *brickwall* and *escalator* texture.

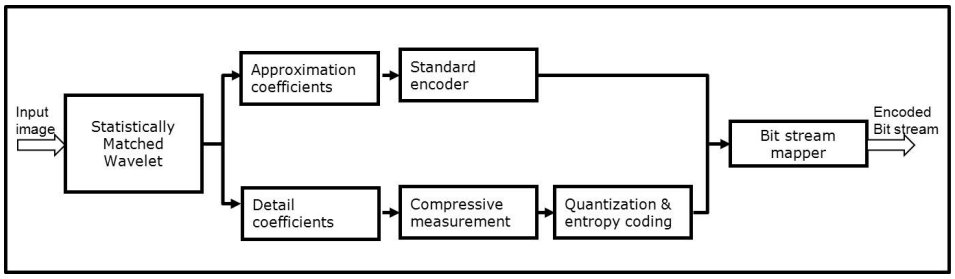


Figure 6: Encoder Framework

4. Texture Synthesis Framework

In this section, we present the overall texture synthesis framework. Fig. 7 gives an overall block diagram of the proposed decoder.

At the decoder, the detail subbands coefficients are synthesized from the CS measurements using the compressive optimization, while the approximation coefficients are decoded using standard JPEG2000. In all our experiments, we have used convex optimization (l_1-norm minimization with equality constraint) for good recovery in a CS framework. This matches the l_0 norm as RIP is ensured due to noiselet measurements. Because of non-differentiability of the l_1-norm, this optimization principle leads to sparser decompositions [17] and ensures fast and stable resolution better than other methods from the class of greedy algorithms. The l_1 norm also provides a computationally viable approach to sparse signal recovery. In all our experiments we have used standard basis pursuit using a primal-dual algorithm [47] which finds the vector with smallest l_1-norm (Eq. 11). The algorithm at the end of this section provides the implementation summary of the convex optimization and proposed decoder framework. Combining the decoded

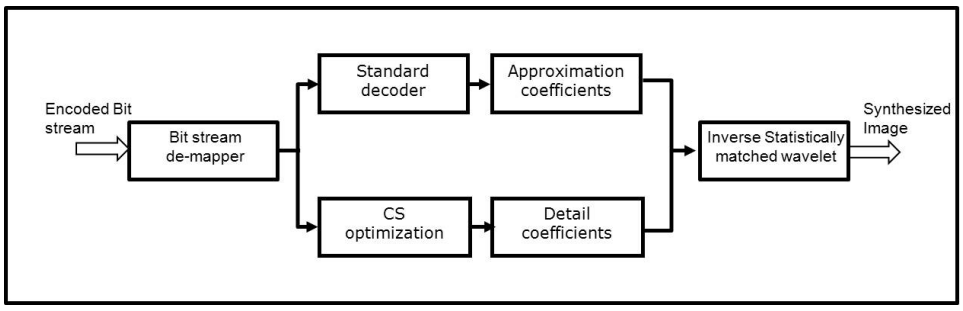


Figure 7: Decoder Framework

samples from the standard decoder and the texture synthesizer (CS optimisation) and doing the inverse of Statistically Matched Wavelet transform, a synthesized image is reconstructed.

$$min\|x\|_1 \quad subject\ \ to \quad \Theta x = s_a, \quad \|x\|_1 = \sum_i |x_i| \tag{11}$$

Algorithm: Texture Synthesis in a CS framework

Inputs: Initial Reference Point,Observation Vector, Optimization Parameter

Initialize: s_a=Nx1 vector(guess point), y=kx1 measurement vector,
T(tolerance for primal-dual algorithm)=5, I(Maximum primal dual iterations)=20,
Tg(Tolerance for conjugate gradients)=1e-8, Ig(Maximum conjugate gradients)=300

If (Valid Starting Reference, s_a)
Minimize f(x) subject to Θ x=s_a (Where, f-convex, rankΘ=M<N)
(Use Primal Dual Interior Point Methods with Equality Constraint)

While (Surrogate duality gap<T OR Iterations>I) do
1) Optimality Condition : ∇ f(x^*)+Θ^T V^*=0, $\Theta(x^*)$=s_a
2) Compute the Newton step and decrement Δx_{nt}, $\lambda(x)$
3) Solve KxK Positive definite system of equations from Newton step
3a) Use Conjugate Gradients Method, MATLAB command, cgsolve
3b) Stop Conjugate Gradient algorithm, If (norm(Θ x-s_a)/norm(s_a) < Tg)
4) Do line Search. Choose step size **t** by backtracking line search
5) Update. x=x+$t\Delta x_{nt}$
6) Update Central and Dual residuals
end while
end If
Output: Sparse representation Wavelet Coefficients $\hat{\mathbf{s}}$

5. Results and Discussion

In this section, we present the experimental results. For our experiments, we have used texture database from Broadtz album [48] and Portilla et al [39] website to select different class of structural textured images (periodic, pseudo-periodic, aperiodic) and complex structured photographic textures. All the test sequences are 128x128, 8-bit, gray scale texture. PSNR and MOS have been used as the quality metrics. Mean Opinion Score (MOS) computation was done by collecting responses of various students and staff working in the lab and averaging them. All the experiments are carried out in MATLAB, running on windows XP PC with P4 CPU and 1GB RAM.

5.1. Statistically Matched Wavelet filter estimation

In this section we present the simulation results of Statistically Matched Wavelet-based 2-D separable kernel wavelet filters used to input data decomposition and sub-bands representation in the proposed framework. Table. 2 gives the result of statistically matched analysis wavelet filters estimated using input data with the filter length set to 5. Using the analysis filters we have computed all other wavelet filters as described in section 2, to construct the filter bank. Fig. 8, shows the reconstruction results of one of our test sequence (*escalator*), using different set of sub-bands coefficients for approximation and detail subbands. As one can see that by increasing the number of approximation sub-band coefficients (a=1000 to 4000) while keeping the number of detail subbands coefficients constant (d=1000), improves the reconstruction quality and PSNR significantly. As opposed to this if we keep the approximation subband coefficients at constant (a=1000) and increase the number of detail subbands coefficients (d=1000 to 4000), the reconstruction quality and PSNR remains almost unchanged. This experiment demonstrates the correctness of our theoretical claim that statistically matched wavelet based texture representation ensures that maximum energy is captured in the approximation sub-space with very little information left in detail sub-space. One can also observe from Fig. 8 that we need smaller number of coefficients from detail subbands (we select 1000 coefficients out of total 13872 coefficients (3*68*68)) as compared to those from approximation subband. This experiment was performed to illustrate the significance of approximation and detail subbands coefficients for good quality reconstruction.

5.2. Texture synthesis results

In this section we present the texture synthesis results of our proposed scheme (SMWT-CS) and conventional DWT-based texture synthesis scheme in a CS framework (DWT-CS) [43, 21]. We have compared our simulation results with the conventional DWT based image synthesis scheme in a CS framework as presented in [43, 21]. In addition, we have done a comparative study of our synthesis results with joint-statistics based statistical model for texture synthesis schemes as presented in [39, 40] and standard JPEG2000. The texture synthesis performance are measured using varying CS measurement samples such as, $M = 2000$ and $M = 4000$.

- Fig. 9 and Fig. 10 presents the synthesis results for structural texture such as *brickwall*, *escalator* and statistical texture such as $floorbox$ and *balckhole*, using conventional DWT-CS and our proposed scheme. Table. 3 gives the PSNR values and MOS scores of the synthesized texture. As one can observe from Table. 3, the

proposed scheme outperforms the conventional DWT-CS scheme and can provide significantly better PSNR (5 to 10 dB gain) for the same compressive measurements (M=2000 or M=4000) or better compression at the same PSNR. It is important to note that SMWT-CS is able to reconstruct the textural pattern smoother and sharper as compared to conventional DWT-CS based scheme for same CS measurements (Fig. 9). In addition, we can observe that the texture synthesis quality can be improved by increasing the CS measurement. This can be observed both subjectively (MOS assessment) as well as objectively (PSNR data), hinting at the scalability in the proposed framework.

- Fig. 11, Fig. 12 and Fig. 13 shows the synthesis results for structural textures with regular patterns such as $D20$, $D36$, $D75$ and structural textures with irregular patterns such as $D68$, $D76$, $D87$ from Broadtz album [48]. As one can observe, the proposed scheme can synthesize the regular and irregular structural patterns with better PSNR and perceptual quality as compared to conventional DWT-CS schemes for the same CS measurements. When compared with the synthesis results for the same textures with other wavelet based texture synthesis scheme such as HMT-3S [40], the proposed scheme outperforms them interms of perceptual synthesis quality. In fact, such joint-statistics based statistical texture model cannot handle the regular and periodic structures as reported by the authors in [40].

- Fig. 14, shows the synthesis results for complex structured photographic texture such as *stone-wall* and $fish\text{-}fabric$ from Portilla website [39]. As one can observe, the proposed scheme results in better PSNR and perceptual quality as compared to conventional DWT-CS schemes for the same CS measurements. Also, when compared with the synthesis results based on joint-statistical models for texture synthesis as suggested by Portilla et al. [39] for the same textures, the proposed scheme provides much better synthesis quality and the capability to synthesise all the patterns types.

When compared with JPEG2000, we obtain approx 4 dB PSNR gain for statistical texture such as *blackhole* and $floorbox$(Table. 3). For structural texture such as *brickwall* and *escalator* (Fig. 9), JPEG2000 performs better at the same bit rate. The reason for this is the presence of significant energy in the high frequency region for structured textures. The coding efficiency loss in a scene with very high frequency transition is due to the fact that compressive sensing recovery of sparse signals typically requires a number of measurements to be larger than the number of nonzero samples.

Table 2: Estimate of statistically matched Wavelet Filters

Test Sequence	Matched Wavelet Filters	n_1	n_2	n_3	n_4	n_5
BrickWall	h_1x	0.1441	-0.4869	0.6997	-0.4855	0.1300
	h_1y	0.0862	-0.5230	0.7329	-0.4099	0.1180
Escalator	h_1x	0.1564	-0.4931	0.6894	-0.4827	0.1553
	h_1y	0.1248	-0.3860	0.7252	-0.5496	0.0860
Floor Box	h_1x	0.1256	-0.4708	0.7171	-0.4852	0.1134
	h_1y	0.1018	-0.3761	0.7148	-0.5668	0.1264
Black hole	h_1x	0.1276	-0.4873	0.7047	-0.4808	0.1360
	h_1y	0.1058	-0.4367	0.7065	-0.5257	0.1503
D20	h_1x	0.1197	-0.4864	0.7371	-0.4459	0.0832
	h_1y	0.1080	-0.4851	0.7307	-0.4558	0.1062
D36	h_1x	0.1012	-0.4673	0.7495	-0.4516	0.0753
	h_1y	0.0661	-0.4249	0.7732	-0.4633	0.0517
D75	h_1x	0.1301	-0.4776	0.7051	-0.4896	0.1342
	h_1y	0.1366	-0.5195	0.7060	-0.4446	0.1242
D68	h_1x	0.1341	-0.4849	0.7026	-0.4832	0.1405
	h_1y	0.1171	-0.4321	0.7241	-0.5139	0.1057
D87	h_1x	0.1354	-0.4858	0.7052	-0.4811	0.1303
	h_1y	0.1529	-0.4996	0.6894	-0.4813	0.1418
Fish-fabric	h_1x	0.0942	-0.4654	0.7448	-0.4607	0.0870
	h_1y	0.0687	-0.4580	0.7686	-0.4373	0.0583

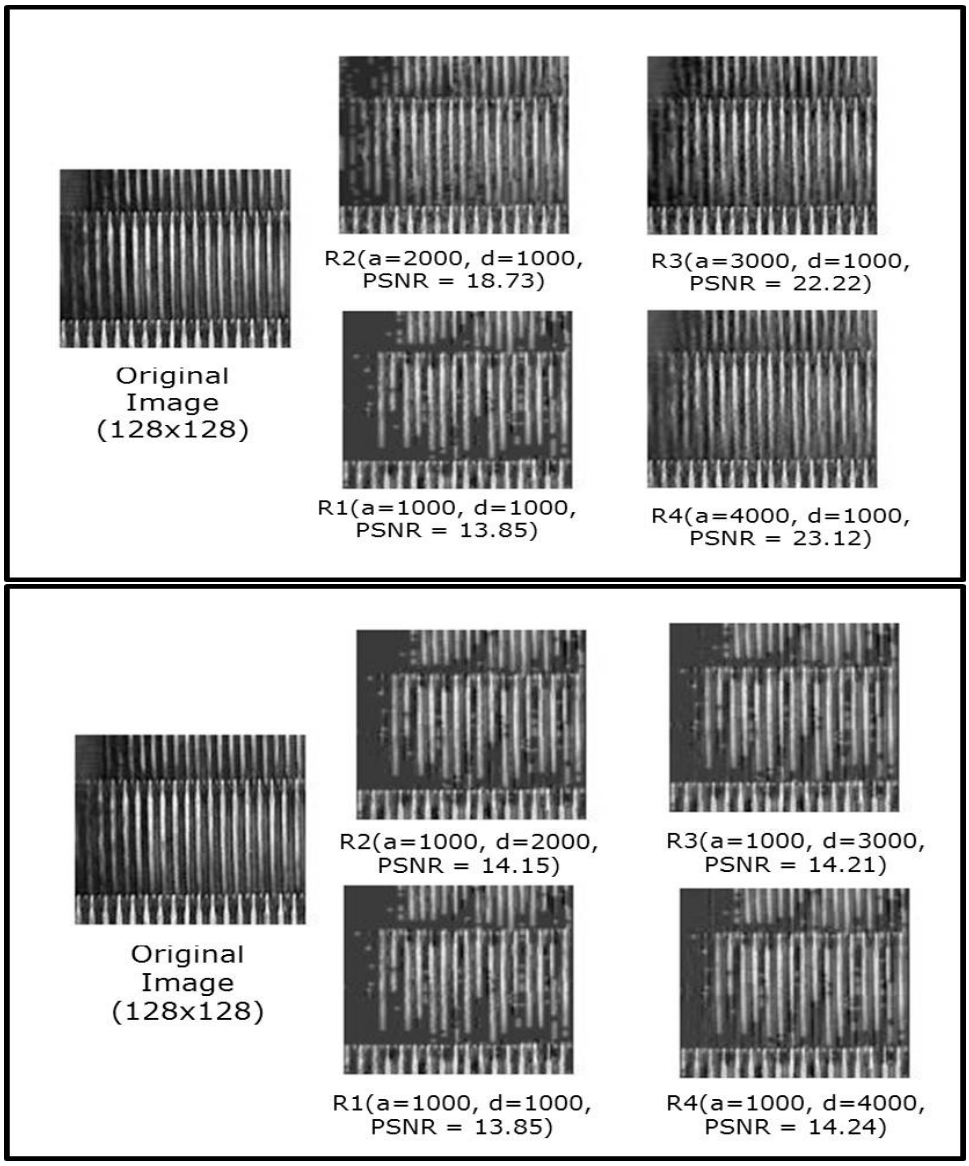


Figure 8: Illustrating the significance of approximation and detail subbands coefficients in image reconstruction. R1, R2, R3 and R4 are the reconstructed image using different approximation (indicated as a) and detail subbands coefficients (indicated as d in the figure). The upper part shows reconstruction for varrying a with constant d, while lower part shows reconstruction for constant a with varrying d.

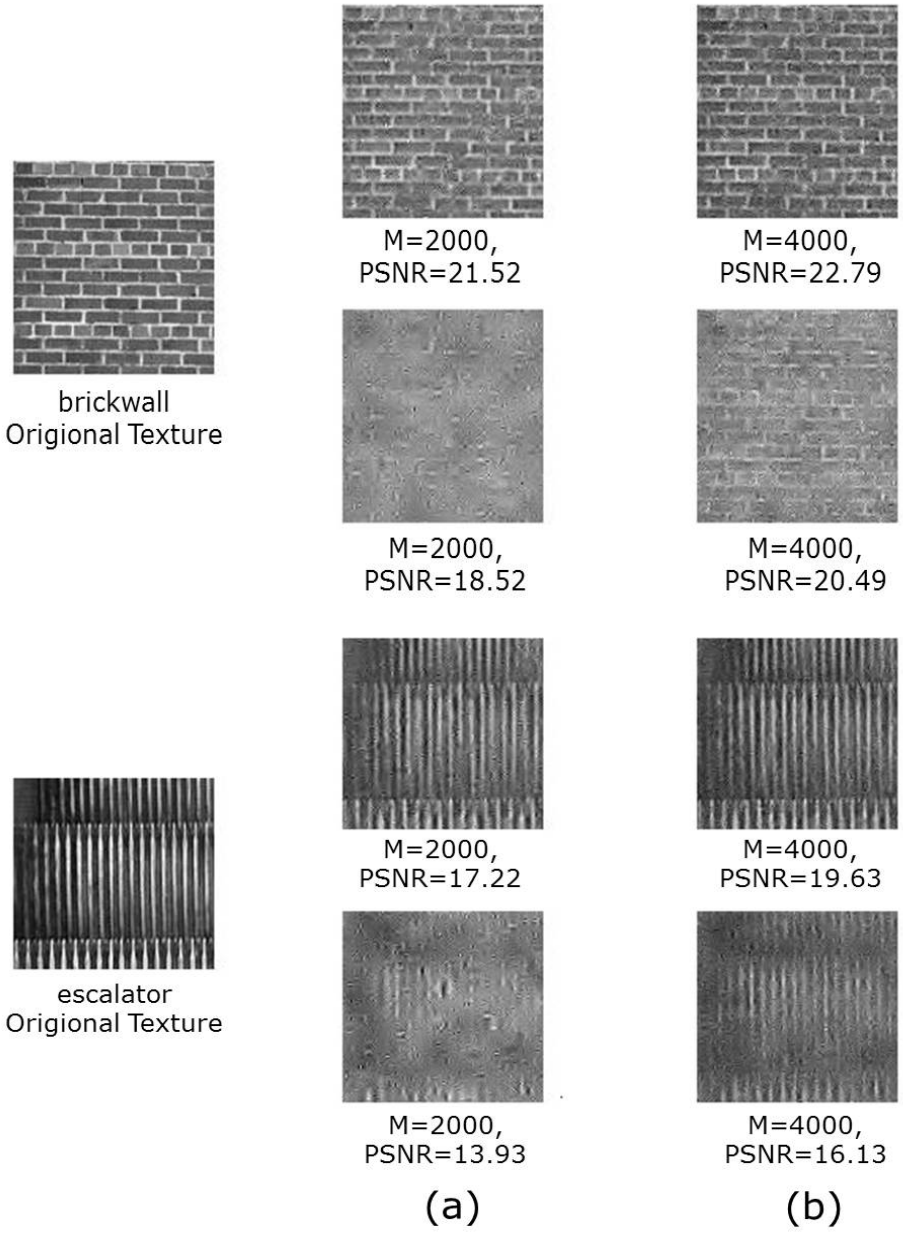


Figure 9: Synthesis results of structural texture with pseudo-periodic patterns (brickwall and escalator texture). For each origional texture, the lower image pair is the synthesized texture using conventional DWT based CS scheme for image synthesis [43, 21], and the upper image pair is the synthesized texture using our proposed scheme. (a) Synthesis results for CS measurements (M)=2000 (b) Synthesis results for CS measurements (M)=4000.

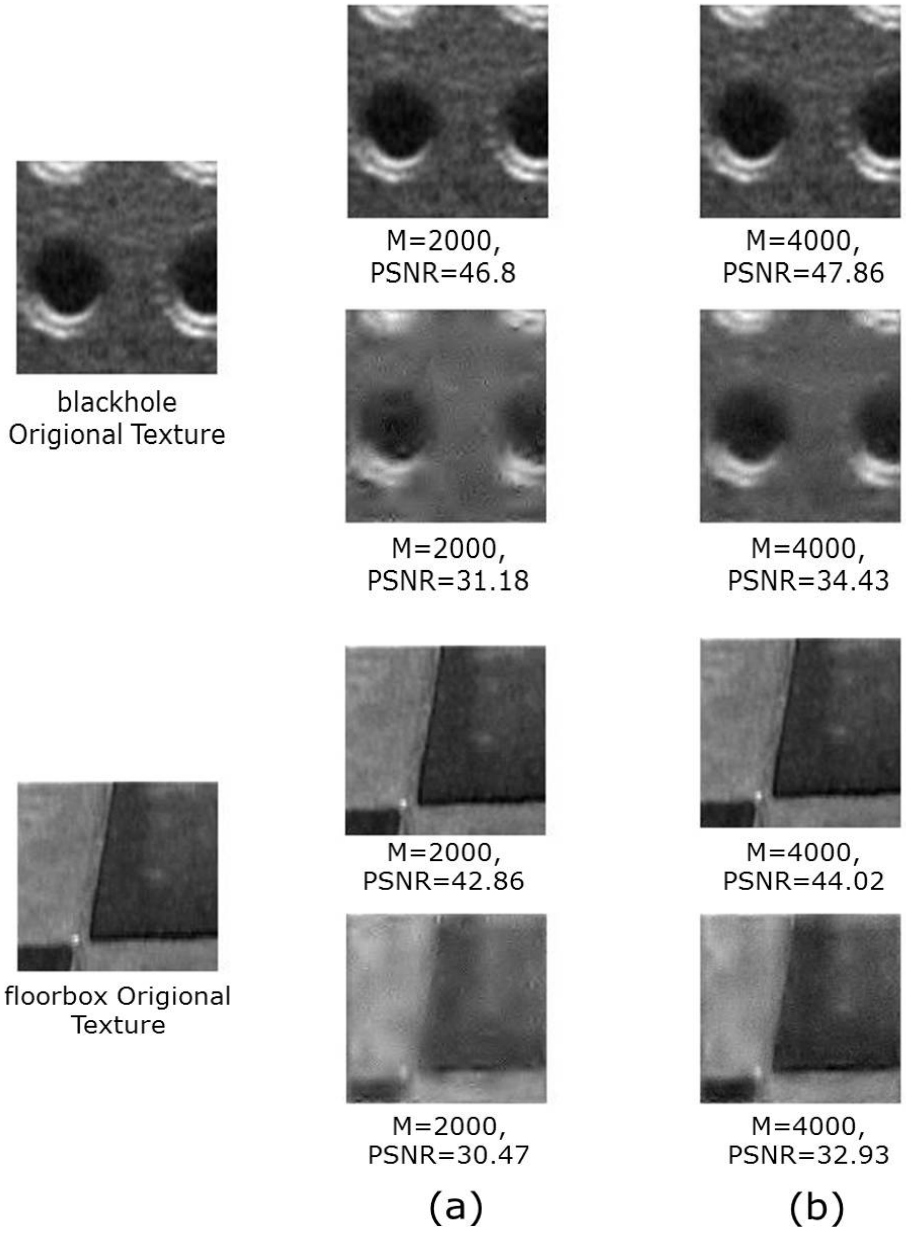


Figure 10: Synthesis results of statistical texture (blackhole and floorbox texture). For each origional texture, the lower image pair is the synthesized texture using conventional DWT based CS scheme for image synthesis [43, 21], and the upper image pair is the synthesized texture using our proposed scheme. (a) Synthesis results for CS measurements (M)=2000 (b) Synthesis results for CS measurements (M)=4000.

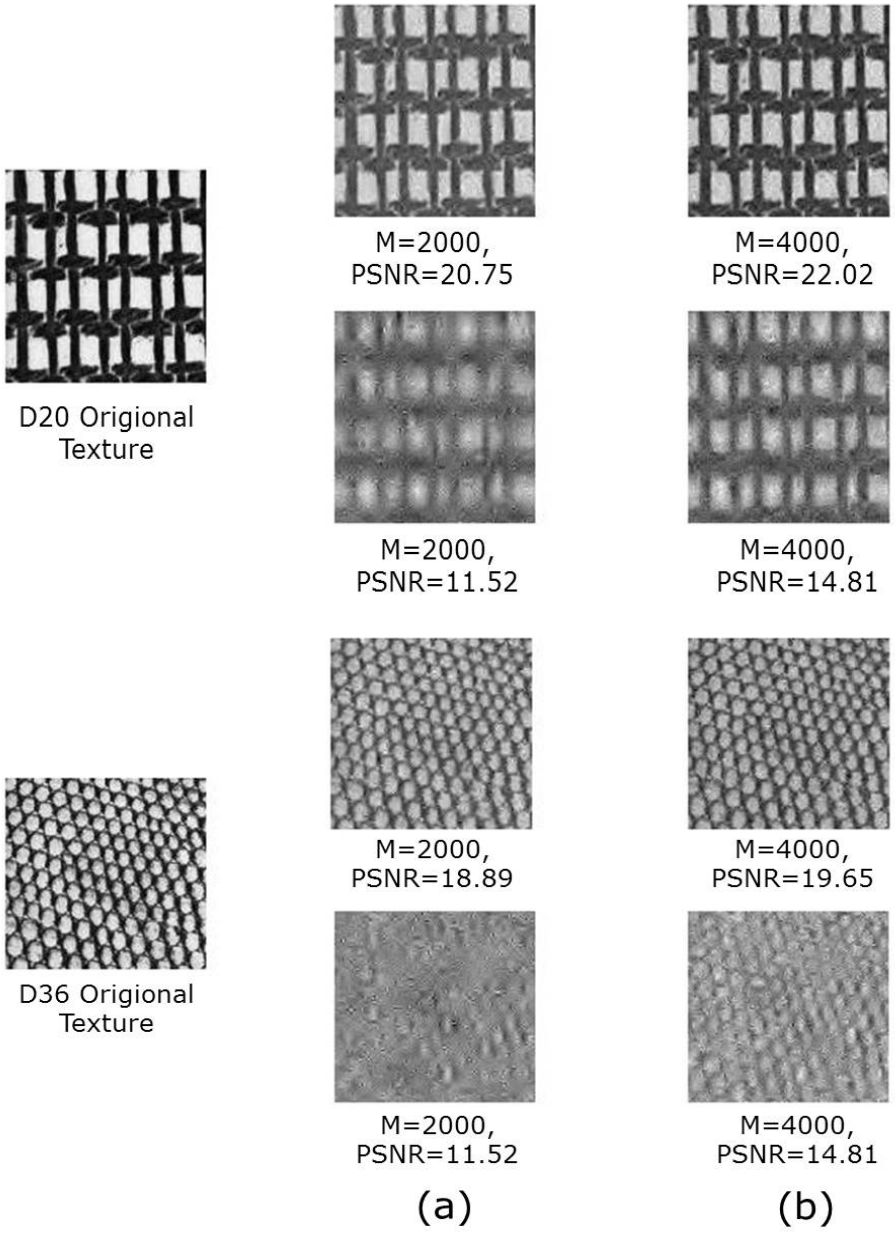


Figure 11: Synthesis results of structural texture with periodic patterns (D20 and D36 texture from broadtz album [48]). For each origional texture, the lower image pair is the synthesized texture using conventional DWT based CS scheme for image synthesis [43, 21], and the upper image pair is the synthesized texture using our proposed scheme. (a) Synthesis results for CS measurements (M)=2000 (b) Synthesis results for CS measurements (M)=4000.

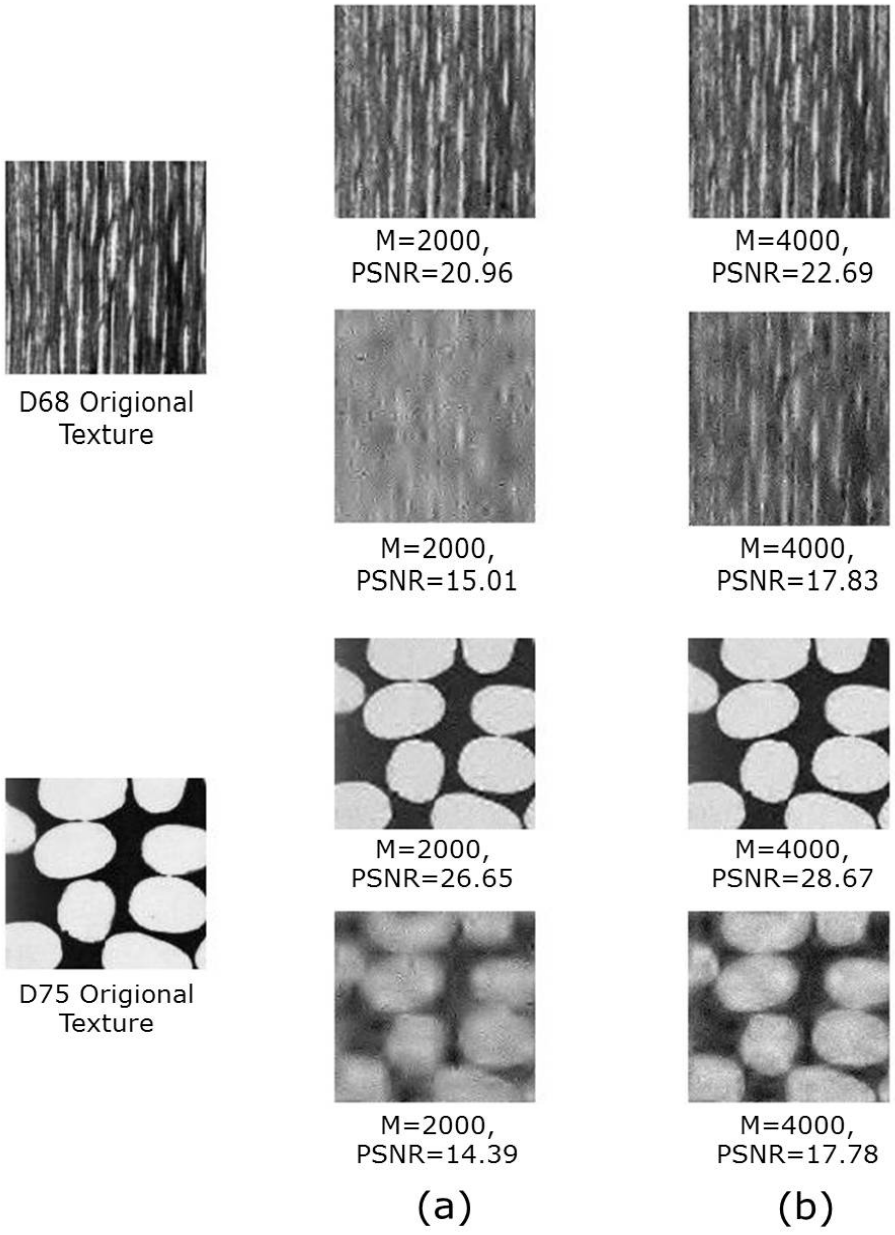


Figure 12: Synthesis results of structural texture with irregular and regular patterns (D68 and D75 texture from broadtz album [48]). For each origional texture, the lower image pair is the synthesized texture using conventional DWT based CS scheme for image synthesis [43, 21], and the upper image pair is the synthesized texture using our proposed scheme. (a) Synthesis results for CS measurements (M)=2000, (b) Synthesis results for CS measurements (M)=4000.

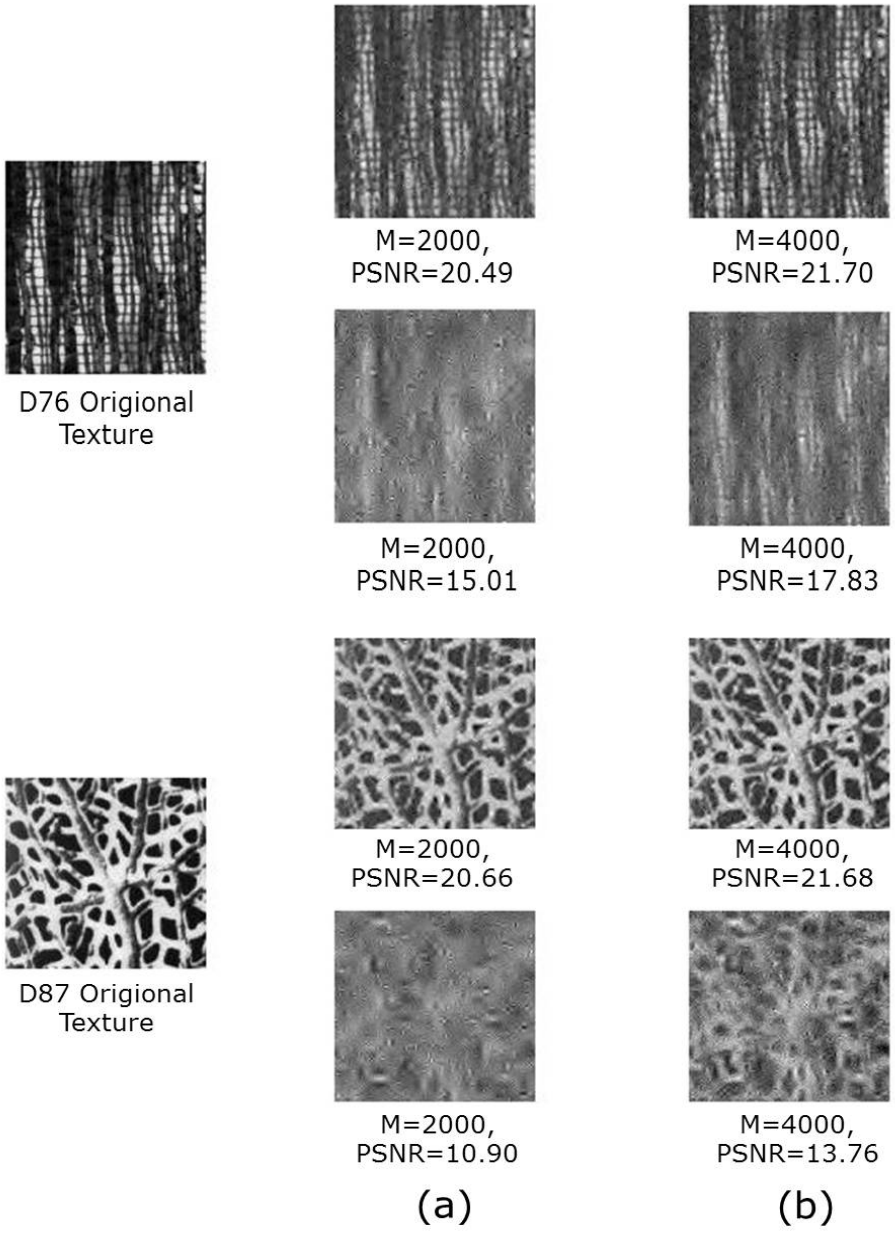


Figure 13: Synthesis results of structural texture with irregular patterns (D76 and D87 texture from broadtz album [48]). For each origional texture, the lower image pair is the synthesized texture using conventional DWT based CS scheme for image synthesis [43, 21], and the upper image pair is the synthesized texture using our proposed scheme. (a) Synthesis results for CS measurements (M)=2000, (b) Synthesis results for CS measurements (M)=4000.

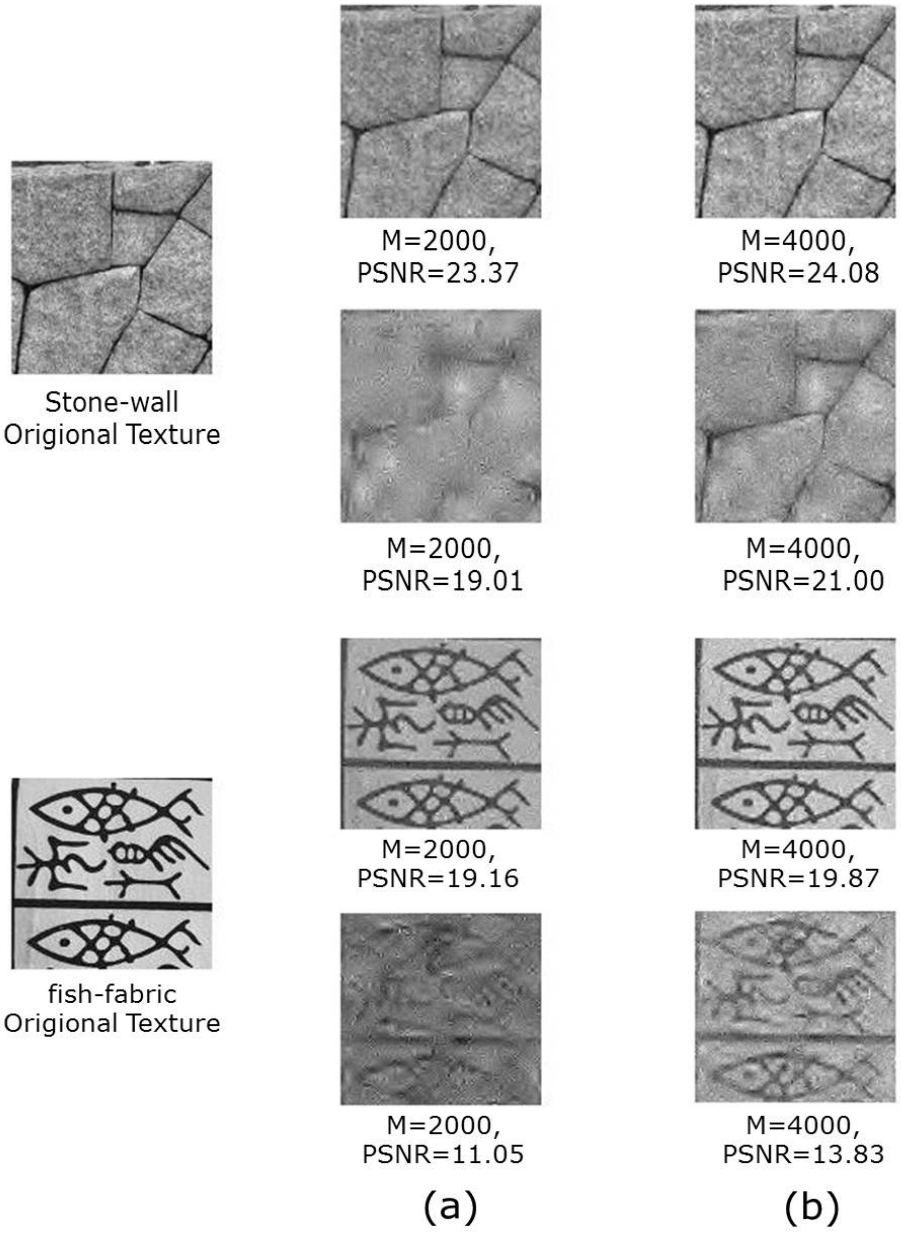


Figure 14: Synthesis results of complex structural texture with irregular patterns (*stone-wall* and $fish\text{-}fabric$ texture from portilla website [39]). For each origional texture, the lower image pair is the synthesized texture using conventional DWT based CS scheme for image synthesis [43, 21], and the upper image pair is the synthesized texture using our proposed scheme. (a) Synthesis results for CS measurements (M)=2000, (b) Synthesis results for CS measurements (M)=4000.

Table 3: Illustrating the PSNR and MOS data of the proposed model((a)floorbox, (b)blackhole,(c)escalator and (d)brickwall) using JPEG2000, conventional wavelet based CS scheme (DWT-CS) [43, 21] and our proposed CS scheme (SMWT-CS).

(a) Sequence : Floor Box (128x128)				
CS Measurements (No. of Coefficients for JPEG2000)	(PSNR in db) (JPEG2000)	(PSNR in db) (DWT-CS)	(PSNR in db) (Our Scheme, SMWT-CS)	(MOS) (Our Scheme, SMWT-CS)
2000	39.19	30.47	42.86	5.0
4000	41.30	32.93	44.02	5.0
6000	41.85	36.39	45.40	5.0
8000	41.90	38.19	46.96	5.0
10000	41.90	42.27	48.98	5.0
12000	41.90	45.01	51.16	5.0

(b) Sequence : Black hole (128x128)				
CS Measurements (No. of Coefficients for JPEG2000)	(PSNR in db) (JPEG2000)	(PSNR in db) (DWT-CS)	(PSNR in db) (Our Scheme, SMWT-CS)	(MOS) (Our Scheme, SMWT-CS)
2000	38.88	31.18	46.8	5.0
4000	41.51	34.43	47.86	5.0
6000	41.97	36.72	49.06	5.0
8000	41.98	39.63	50.58	5.0
10000	41.98	42.86	52.24	5.0
12000	41.98	46.30	54.16	5.0

(c) Sequence : Escalator (128x128)				
CS Measurement (No. of Coefficients for JPEG2000)	(PSNR in db) (JPEG2000)	(PSNR in db) (DWT-CS)	(PSNR in db) (Our Scheme, SMWT-CS)	(MOS) (Our Scheme, SMWT-CS)
2000	30.86	13.93	17.22	3.0
4000	31.08	16.13	19.63	3.5
6000	34.76	18.26	22.13	4.0
8000	38.39	20.88	24.75	4.2
10000	40.07	23.75	27.84	4.4
12000	40.56	27.48	31.76	4.8

(d) Sequence : Brickwall (128x128)				
CS Measurements (No. of Coefficients for JPEG2000)	(PSNR in db) (JPEG2000)	(PSNR in db) (DWT-CS)	(PSNR in db) (Our Scheme, SMWT-CS)	(MOS) (Our Scheme, SMWT-CS)
2000	30.47	18.52	21.52	3.5
4000	31.41	20.49	22.79	3.5
6000	33.25	22.34	24.50	4.1
8000	35.19	24.50	26.17	4.1
10000	37.48	27.14	28.35	4.4
12000	38.95	30.50	31.04	4.8

6. Conclusion

In this paper, we propose Statistically Matched Wavelet based texture data representation and synthesis in a compressive sensing framework (SMWT-CS). Statistically Matched Wavelet based representation causes most of the captured energy to be concentrated in the approximation sub-space, while very little information is retained in the detail sub-space. We encode not the full-resolution Statistically Matched Wavelet subband coefficients, but only the approximation subband coefficients (LL) using standard image compression scheme like JPEG2000. The detail subband coefficients i.e HL, LH and HH are jointly encoded in a compressive sensing framework using compressive projection and measurements. The experimental results demonstrate that the proposed scheme can provide significantly better PSNR for the same compressive measurements or better compression at the same PSNR as compared to conventional DWT-based image compression scheme in a CS framework. It can also be observed that performing linear compression over the approximation sub-space provides better reconstruction quality for the same number of samples as compared to CS measurements over approximation sub-space. This indicates that performing linear compression over approximation sub-space and CS measurements over detail sub-space provides optimal compression and reconstruction quality as against using only standard linear compression or using only CS measurements over both approximation and detail sub-space.

References

[1] ITU-T Rec. H.262 and ISO/IEC 13818-2 MPEG-2, Generic Coding of Moving Pictures and Associated Audio Information - Part 2 Video.

[2] ITU-T Rec. H.264 and ISO/IEC 14496-10 (MPEG-4/AVC), Advanced video coding for generic audio visual services, Standard version 7, ITU-T and ISO/IEC JTC 1.

[3] H. Nyquist, Certain topics in telegraph transmission theory, Trans. AIEE (1928) 617 – 644.

[4] P. Ndjiki-Nya, T. Hinz, C. Stuber, T. Wiegand, A content-based video coding approach for rigid and non-rigid texture, Proc. IEEE International Conference on Image Processing (ICIP) (2006) 3169 – 3172.

[5] P. Ndjiki-Nya, C. Stuber, T. Wiegand, Texture Synthesis method for generic video sequences, Proc. IEEE International Conference on Image Processing (ICIP) (2007) 397 – 400.

[6] P. Ndjiki-Nya, T. Hinz, T. Wiegand, Generic and robust video coding with texture analysis and synthesis, IEEE International Conference on Multimedia and Expo (2007) 1447 – 1450.

[7] P. Ndjiki-Nya, M. Koppel, D. Doshkov, T. Wiegand, Automatic structure-aware inpainting for complex image content, International Symposium on Visual Computing (2008) 1144 – 1156.

[8] P. Ndjiki, T. Wiegand, Video coding using closed-loop texture analysis and synthesis, JSTSP Special issue on Emerging technologies for video compression.

[9] M. Bosch, F. Zhu, E. J. Delp, Spatial texture models for video compression, Proc. IEEE International Conference on Image Processing (ICIP) (2003) 93 – 96.

[10] M. Bosch, F. Zhu, E. J. Delp, Video coding using motion classification, Proc. IEEE International Conference on Image Processing (ICIP) (2008) 1588 – 1591.

[11] J. Byrne, S. Ierodiaconou, D. R. Bull, D. Redmill, P. Hill, Unsupervised image compression-by-synthesis within a jpeg framework, Proc. IEEE International Conference on Image Processing (ICIP) (2008) 2892 – 2895.

[12] S. Ierodiaconou, J. Byrne, D. R. Bull, D. Redmill, P. Hill, Unsupervised image compression using graphcut texture synthesis, Proc. IEEE International Conference on Image Processing (ICIP) (2009) 2289 – 2292.

[13] F. Zhang, D. R. Bull, N. Canagarajah, Region based texture modelling for next generation video codecs, Proc. IEEE International Conference on Image Processing (ICIP) (2010) 2593 – 2596.

[14] F. Zhang, D. R. Bull, Enhanced video compression with region based texture models, Picture Coding Symposium(PCS) (2010) 54 – 57.

[15] A. Gupta, S. D. Joshi, S. Prasad, A new approach for estimation of statistically matched wavelet, IEEE Transactions on Signal Processing (2005) 1778 – 1793.
[16] D. Donoho, Compressed sensing, IEEE Transactions on Information Theory (2006) 1289 – 1306.
[17] R. G. Baraniuk, Compressed sensing[lecture notes], IEEE Signal Processing Magazine (2007) 118 – 121.
[18] S. Chen, D. Donoho, M. Saunders, Atomic decomposition by basis pursuit, SIAM J. on Scientific Computing Vol. 20 (1998) 33 – 61.
[19] E. Candes, J. Romberg, T. Tao, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information, IEEE Transactions on Information Theory (2006) 489 – 509.
[20] Y. Zhang, S. Mei, Q. Chen, Z. Chen, A novel image/video coding method based on compressive sensing theory, Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (2008) 1361 – 1364.
[21] D. Venkatraman, A. Makur, A compressive sensing approach to object-based surveillance video coding, Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (2009) 3513 – 3516.
[22] J. Prades-Nebot, Y. Ma, T. Huang, Distributed video coding using compressive sampling, Picture Coding Symposium (PCS) (2009) 1 – 4.
[23] M. B. Wakin, J. N. Laska, M. F. Duarte, D. Baron, S. Sarvotham, D. Takhar, K. F. Kelly, R. G. Baraniuk, Compressed imaging for video representation and coding, Picture Coding Symposium (2006) 1289 – 1306.
[24] Y. Yang, O. C. Au, L. Fang, X. Wen, W. Tang, Perceptual compressive sensing for image signals, IEEE International Conference on Multimedia and Expo (2009) 89 – 92.
[25] A. Efros, T. Leung, Texture synthesis by non-parametric sampling, International conference on computer vision (1999) 1033 – 1038.
[26] L. Y. Wei, M. Levoy, Fast Texture Synthesis using tree-structured vector quantization, in Proc. SIGGRAPH, New Orleans, Louisiana, USA (2000) 479 – 488.
[27] V. Kwatra, A.Schodl, I. Essa, G. Turk, A. Bobick, Graphcut textures : image and video synthesis using Graph Cuts, in Proc. SIGGRAPH, San Diego, CA, USA (2003) 277 – 286.
[28] P. Hill, Wavelet based texture analysis and segmentation for imgae retrieval and fusion, PhD,University of Bristol.
[29] J. Y. A. Wang, E. H. Adelson, Representing moving images with layers, IEEE Transaction on Image Processing (1994) 625 – 638.
[30] A. Dumitras, B. G. Haskell, An Encoder-decoder Texture Replacement Method with Application to Content-based Movie Coding, IEEE Transactions on Circuits and Sytems for Video Technology 14 (2004) 825 – 840.
[31] R. O'Callaghan, D. Bull, Combined morphological-spectral unsupervised image segmentation, IEEE Transaction on Image Processing (2005) 49 – 62.
[32] H. Choi, R. G. Baraniuk, Multiscale image segmentation using wavelet domain hidden markov model, IEEE Transaction on Image Processing (2001) 1309 – 1321.
[33] J. Li, R. M. Gray, Context-based multiscale classification of document images using wavelet coefficient distribution, IEEE Transaction on Image Processing (2000) 1604 – 1616.
[34] M. Acharya, M. K. Kundu, An adaptive approach to unsupervised texture segmentation using m-band wavelet, Signal Processing (2001) 1337 – 1356.
[35] A. Khandelia, S.Gorecha, B. Lall, S. Chaudhury, M. Mathur, Parametric Video Compression Scheme using AR based Texture Synthesis, Proc. IAPR- and ACM-sponsored Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP) (2008) 219 – 225.
[36] A. Stojanovic, M. Wien, J. R. Ohm, Dynamic texture synthesis for H.264/AVC inter coding, Proc. IEEE International Conference on Image Processing (ICIP) (2008) 1608 – 1612.
[37] A. Stojanovic, M. Wien, T. K. Tan, Synthesis-in-the-loop for video texture coding, Proc. IEEE International Conference on Image Processing (ICIP) (2009) 2293 – 2296.
[38] H. Chen, R. Hu, D. Mao, R. Thong, Z. Wang, Video coding using dynamic texture synthesis, IEEE International Conference on Multimedia and Expo (2010) 203 – 208.
[39] J. Portilla, E. P. Simoncelli, A Parametric Texture Model based on Joint Statistics of Complex Wavelet Coefficients, IJCV (2000) 49 – 71.
[40] G. Fan, X. Xia, Wavelet-based texture analysis and synthesis using hidden Markov Model, IEEE Transactions on Circuits and Sytems for Video Technology (2003) 106 – 120.
[41] S. Kumar, R. Gupta, N. Khanna, S. Chaudhury, S. D. Joshi, Text extraction and document image segmentation using matched wavelets and mrf model, IEEE Transaction on Image Processing (2007)

2117 – 2128.
[42] I.-T. R. T.800, Jpeg-2000:core coding system, International Telecommunication Union, Tech. Rep.
[43] A. Schulz, L. Velho, Eduardo, A. B. da Silva, On the empirical rate-distortion performance of compressive sensing, Proc. IEEE International Conference on Image Processing (ICIP) (2009) 3049 – 3052.
[44] C. Deng, W. Lin, B. Lee, C. T. Lau, Robust image compression based on compressive sensing, IEEE International Conference on Multimedia and Expo (2010) 462 – 467.
[45] A. Mojsilovic, M. V. Popovic, D. M. Rackov, On the selection of an optimal wavelet basis for texture characterization, ITIP (2000) 2043 – 2050.
[46] R. Coifman, F. Geshwind, Y. Meyer, Noiselets, Appl. Comp. Harman. Anal. vol. 10, no. 1.
[47] S. Boyd, L. Vandenberghe, Convex optimization, Cambridge University Press.
[48] P. Brodatz, Textures - A Photographic Album for Artists and Designers, New York : Dover.