Compressed Sensing Based Image Encoding Technique for Wireless Sensor Networks

A. Loganathan
Assistant Professor
Department of ECE
Velammal Engineering College
Chennai, India

S.Harish
PG Student
Department of ECE
SSN College of Engineering
Chennai, India

Dr. R.Kanthavel
Professor & Head
Department of ECE
Velammal Engineering College
Chennai, India

***Abstract*— The Wireless Sensor Network (WSN) is the one, which generally consists of cameras themselves, which have some local image processing, communication and storage capabilities, and one or more central computers, where image data from multiple cameras is further processed and fused. Here the computation resource is extremely limited. Because of these limitations, a new sampling method is introduced in the Image/video encoder of the WSN called Compressed Sensing (CS), which is the process of acquiring and reconstructing a signal that is supposed to be sparse or compressible, thus reducing the computational complexity. The image is divided into dense and sparse components [1], where the dense component uses the standard encoding procedure such as JPEG and the sparse measurements from the sparse components are encoded by the Arithmetic encoding. The correlation between the dense and sparse components is studied using the autoregressive model, by which the sparse components are predicted from the dense component at the receiver side. With the measurements (used in CS) and the predicted sparse components as the initial values, the projection onto convex set (POCS) recovery algorithm [2] is used to get back the original sparse components and hence the original image by applying the inverse of transform to the dense and recovered sparse components.**

Keywords-Compressed Sensing, JPEG, Arithmetic encoding, Image interpolation;PAR model, Variable adaptive interpolation, Projection onto Convex set.

I. INTRODUCTION

A wireless sensor network (WSN) is a collection of low-cost, low-power sensor nodes that communicate in short distance and collaborate together to reach the objective of a WSN application. Some applications of WSN are environmental monitoring, biomedical research, human imaging and tracking and military applications. In cases where the number of samples is large, compression must be done prior to transmission. In these applications there are limitations on computational power of sampling devices or communication channel bandwidth. Therefore reducing the complexity and power consumption of the sensors is desirable, which can be accomplished by incorporating a new sampling technique called compressed sensing, which is a signal processing technique that takes advantage of the signal's sparseness or compressibility in some domain, allowing the entire signal to be determined from relatively few measurements.

Emmanuel J. Candès and Michael B [3] explained that the use of CS theory can recover certain signals and images from far fewer samples or measurements than traditional methods of acquisition. This new sampling theory tries to combine procedures for sampling and compressing data simultaneously. Justin Romberg [5], whose article is an introduction to CS and recovery via convex programming, says about transforming the image into an appropriate basis and then coding only the important expansion coefficients.

Candès E, Romberg J [2] explained that it is possible to reconstruct certain types of signals accurately from limited measurements. In a nutshell, suppose that f is compressible in the sense that it is well-approximated by a linear combination of M vectors taken from a known basis Ψ. Then not knowing anything in advance about the signal, f can (very nearly) be recovered from only about $M \log N$ generic non-adaptive measurements. The recovery procedure is concrete and consists in solving a simple convex optimization program.

In this paper, the image is decomposed into two components: dense and sparse. The decomposition helps to generate a sparser signal which is more suitable for CS. Furthermore, there exists a strong correlation between these two components because the decomposition is usually not completely orthogonal [1]. Following the decomposition, the dense component is transmitted as it is by standard encoding technique such as JPEG, whereas for the sparse components, only very few measurements are transmitted using Arithmetic encoding [12]. At the receiver side, the dense component received, is interpolated to obtain the super-resolution image whose size will be equal to that of original image. Once the interpolation is done, the wavelet transform is applied to the interpolated image to obtain once again, the dense and sparse components. The sparse components are of only interest which will be used as the initialization for the POCS recovery algorithm along with the random measurements taken from the sparse components of the original image. The POCS recovery algorithm estimates the original sparse components in certain no of iterations [2]. The estimated sparse components along with the dense component are applied with inverse wavelet transform, to get back the original image.

II. COMPRESSED SENSING

As explained in the previous section, following the decomposition, the dense component is processed by traditional method and the sparse components by CS. If an N-dimensional signal could be represented sparsely in a certain transformed domain and if there are only K no of non-zero coefficients in the transformed domain (here wavelet transform is applied), then it would be enough to take a very few measurements say M which is much lesser compared to the no of measurements N, if there is no transformation applied [1]. The CS principle claims that the sparse components can be recovered back from such a small number of random linear measurements [1].

Suppose we have an N-dimensional signal

$$x \varepsilon R^N, \quad (1)$$

this could be represented sparsely in a certain domain by the transform matrix Ψ, if there are only K non-zero coefficients in the Ψ domain, that is, x is exactly K-sparse. Then there is no need to take N measurements when $N>>K$.

In order to take CS measurements, we first let ϕ denote an M by N measurement matrix with $M<<N$, which obeys the restricted isometric property. The measurements are obtained by a linear system

$$y = \phi x \quad (2)$$

the condition that M should satisfy is

$$K <= Const * M \log N \quad (3)$$

where *Const* is the over-measuring factor greater than 1. In order to take measurements of the sparse components, a Gaussian random ensemble ϕ is used in this paper. The other measurement matrices generally used are Bernoulli random matrix and Sub-Gaussian random matrices. The entries of a random Bernoulli matrix take the value *±(m)^(-1/2)* with equal probability, while the entries of a Gaussian matrix are independent and follow a normal distribution with expectation zero and variance *(1/m)*, where m is the mean [13]. Whatever the type, the measurement matrix must satisfy the Restricted Isometric property. In essence, a matrix satisfying RIP is such that the lengths of all sufficiently sparse vectors are approximately preserved under transformation by the matrix [13].

Since the dimension of the input image is very high, directly applying the 2D Gaussian random matrix ϕ to the sparse components is not practical. In order to apply the random matrix more efficiently, the sparse components are needed to be regrouped and a block based sampling strategy is followed. The sparse components are first divided into several groups by scaling and then reordering it into a number of vectors of the same dimension. In this way, random measurements to the vector with moderate size instead of the tremendous size can be taken. The encoding and decoding of these measurements are done at the transmitter and receiver side respectively.

III. TRANSMITTER SECTION

The transmitter section mainly involves domain transformation, JPEG encoding of the dense component and Arithmetic encoding of sparse measurements.

A. Domain Transformation

The image has to be converted into certain domain in which it could be represented sparsely, so as to take only few no of measurements. Sparseness is one of the reasons for the extensive use of popular transforms such as the discrete Fourier transform (DFT), the discrete wavelet transform (DWT) and the singular value decomposition (SVD). The aim of these transforms is to reveal certain structures of a signal and to represent these structures in a compact and sparse form. Sparse representations have therefore increasingly become recognized as providing extremely high performance for diverse applications such as: noise reduction, compression, feature extraction, pattern classification and blind source separation [15]. The fig. 1 represents the various levels of wavelet transform.

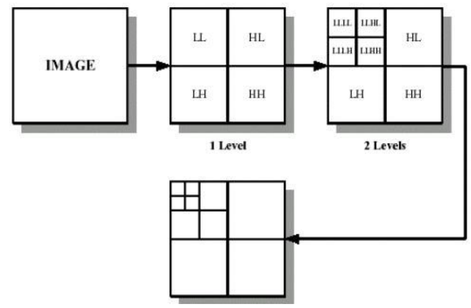

Figure 1. Levels of Wavelet transform

The wavelet transform is a more efficient approach whose coefficients are exactly represented by finite precision numbers. In discrete cosine transform, only spatial correlation of the pixels inside the single 2-D block is considered and the correlation from the pixels of the neighboring blocks is neglected. Also, it is impossible to completely decorrelate the blocks at the boundaries using DCT. Undesirable blocking artifacts affect the reconstructed images or video frames (high compression ratios or very low bit rates). The wavelet transform overcome the above mentioned limitations of applying DCT on images and hence used in [1].

B. Transmitter block diagram

After applying the 2D wavelet transform to the image, the resultant four bands LL, LH, HL and HH are to be encoded and transmitted. The transmitter block diagram of the proposed scheme is as shown in fig. 2.

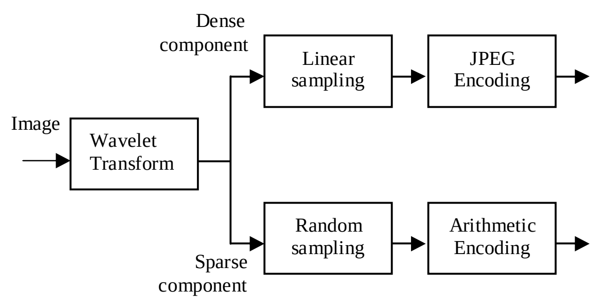

Figure 2. Transmitter block diagram

C. Encoding of dense component

The dense component of the input image, after transformation, is sampled by the normal linear sampling technique followed by quantization and the encoding scheme.

In JPEG encoding, the dense component is transformed into discrete cosine domain followed by the scalar quantization to reduce the number of bits to represent them, at the expense of quality. The output of quantization is a set of integer numbers which have to be encoded bit-by-bit. The quantization table for JPEG encoding is fixed and the quality matrix is formulated by manipulating the quantization table depending on the required quality of compression []. The JPEG is the lossy compression technique for images. The degree of compression can be adjusted, but with a tradeoff between storage size and quality. The JPEG encoding and decoding sections are as shown in the fig. 3.

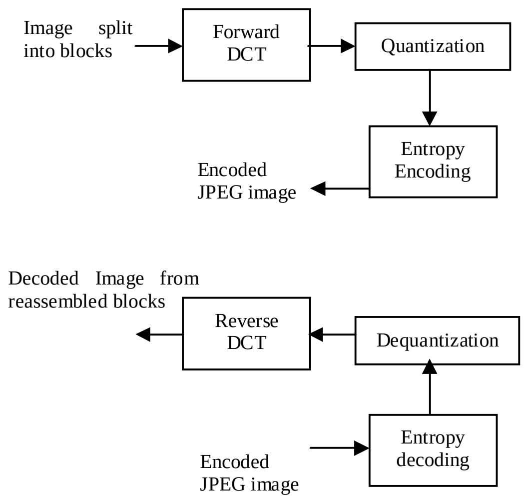

Figure 3. JPEG Encoding and Decoding

D. *Encoding of sparse components*

The sparse measurements taken by the technique of compressed sensing (ref. section 2) are to be encoded by arithmetic encoding. Arithmetic coding is a form of entropy coding used in lossless data compression. When a string is arithmetic encoded, frequently used characters will be stored with fewer bits and not-so-frequently occurring characters will be stored with more bits, resulting in fewer bits used in total. Arithmetic coding differs from other forms of entropy encoding such as Huffman coding in that rather than separating the input into component symbols and replacing each with a code, arithmetic coding encodes the entire message into a single number, usually a fraction *n* where ($0.0 \leq n < 1.0$), called a Tag value. The encoder considers only three important values : the next symbol that needs to be encoded, the current interval (at the very start of the encoding process, the interval is set to [0,1), but that will change), the probabilities the model assigns to each of the various symbols that are possible at this stage. The encoder divides the current interval into sub-intervals, each representing a fraction of the current interval proportional to the probability of that symbol in the current context. Whichever interval corresponds to the actual symbol that is next to be encoded becomes the interval used in the next step.

When all symbols have been encoded, the resulting interval unambiguously identifies the sequence of symbols that produced it. If the same final interval and model that is being used are available, the symbol sequence that must have entered the encoder to result in that final interval can be reconstructed. It is not necessary to transmit the final interval, however; it is only necessary to transmit one fraction that lies within that interval. In particular, it is only necessary to transmit enough digits of the fraction so that all fractions that begin with those digits fall into the final interval [12].

IV. RECEIVER SECTION

The receiver section comprises of two main sections: Interpolation and Signal recovery. The dense component which is transmitted as it is, using JPEG encoding technique, will be given as the input to the interpolation process, the output of which is applied again with wavelet transform. The resultant sparse components along with the arithmetic decoded sparse measurements are given as inputs to the POCS recovery algorithm to obtain the estimated sparse components. These recovered sparse components along with the dense component are applied with inverse wavelet transform to get back the original image. The fig. 4 describes the receiver section in detail.

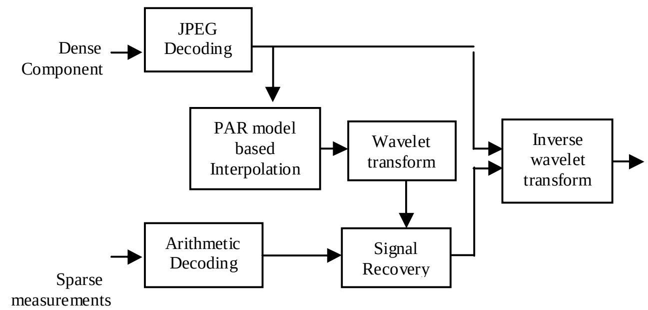

Figure 4. Receiver section

A. *PAR model based Interpolation*

Interpolation is a method of constructing new data points within the range of a discrete set of known data points. There are many types of interpolation namely bilinear interpolation, bi-cubic interpolation etc. The interpolation technique is dealt in this paper, combines a soft-decision interpolation technique that estimates missing pixels in the enlarged image, in groups rather than one at a time [7]. The soft decision estimation technique learns and adapts to varying scene structures at the edges using a 2-D piecewise autoregressive model (PAR). The model parameters are estimated in a moving window in the input low-resolution image. The pixel structure dictated by the learnt model is enforced by the soft-decision estimation process onto a block of pixels, including both observed and estimated. The result is equivalent to that of a high-order adaptive non separable 2-D interpolation filter. This new image interpolation approach preserves spatial coherence of interpolated images at the edges better than the existing methods, and produces best results so far over a wide range of scenes in terms of subjective visual quality [7].

Although the input image could be decomposed into the dense and sparse components, one can still observe that there exists a strong visual correlation between them. Therefore, it is possible to use the dense component to predict the original image, as well as the sparse component. The recent development on adaptive interpolation provides an effective tool to solve this problem. The adaptive interpolation describe the image as a 2D piecewise autoregressive (PAR) model, namely,

$$X_{i,j} = \sum_{(m,n)\in B_{i,j}} \alpha_{m,n} X_{i+m,j+n} + v_{i,j} \quad (4)$$

where (i, j) is the pixel to be interpolated,

$B_{i,j}$ is the window centered at pixel (i, j),

$v_{i,j}$ is the random perturbation independent of pixel and the image signal.

It is observed that the image is piecewise stationary. In other words, although the PAR model parameters $\alpha_{m,n}$ can vary significantly in different segments of a scene, they remain constant in a small local window. Let I_x be the image to be estimated by interpolating the decoded dense component I_y of the original image. Let $x_i \in I_x$ and $y_i \in I_y$ be the pixels of the estimated and the original image. The four connected neighbors of pixel location i in the original image are represented as x_{i-t} and y_{i-t} depending on whether they are in the original image or interpolated image with t values 0,1,2,3. With the introduced notations and the PAR image model, the problem of interpolation is estimated as an optimization problem.

$$\min_{a,b,x}\{\sum_{i\in W}\left\| y_i - \sum_{t=0}^{3}(a_t x_{i-t}^{(4)} + b_t y_{i-t}^{(8)})\right\| + \sum_{i\in W}\left\| x_i - \sum_{t=0}^{3}(a_t y_{i-t}^{(4)} + b_t x_{i-t}^{(8)})\right\|\} \quad (5)$$

To apply the PAR model in interpolating samples, the model parameters from an incomplete data set (half of the pixels are missing), has to be estimated. On one hand, the interpolation performance relies on a good model that fits the true data. On the other hand, the model parameters can be reliably estimated only if the missing pixels are known, thereby creating the dilemma. It can be overcome by treating model parameters a, b and the missing pixels both as variables in the proposed optimal estimation problem of (5). This allows estimating the model parameters and the missing pixels jointly under the constraint of the known quincunx image I_y. The optimization objective is to maintain a best statistical agreement between the estimated model and the interpolated pixels [7].

The optimization problem of (5) is non-linear and non-convex. Hence the problem is broken into three linear least-square sub-problems which develop an efficient solution. Under the assumption of piecewise local stationarity, the two images I_x and I_y have the same second-order statistics. Thus the model parameters $b= (b_0, b1, b2, b3)$ can be estimated by linear least-square fitting of the known quincunx image I_y, independent of I_x. The first subproblem to be solved is as follows,

$$\hat{b} = \arg\min_{b}\{\sum_{i\in W}\left\| y_i - \sum_{t=0}^{3} b_t y_{i-t}^{(8)}\right\|\} \quad (6)$$

$$\hat{a} = \arg\min_{a}\{\sum_{i\in W}\left\| y_i - \sum_{t=0}^{3} b_t y_{i-t}^{*}\right\|\} \quad (7)$$

where y_{i-t}^{*} are the north, south, east and west neighbours of the pixel location i in I_y.

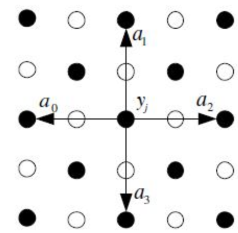

Figure 5. Sample relations in estimating model parameters

By solving the two linear least-square problems of estimating b and a, it reduces (5.2) to a linear least-square estimation problem.

$$\min_{a,b,x}\{\sum_{i\in W}\left\| y_i - \sum_{t=0}^{3}(\hat{a}_t x_{i-t}^{(4)} + \hat{b}_t y_{i-t}^{(8)})\right\| + \sum_{i\in W}\left\| x_i - \sum_{t=0}^{3}(\hat{a}_t y_{i-t}^{(4)} + \hat{b}_t x_{i-t}^{(8)})\right\|\} \quad (8)$$

Rather than making one estimate at a time, the objective function (8) accounts for the mutual influences among estimates of neighbouring missing pixels. These estimates are jointly optimized in a local window W so that the PAR model fits all pixels in W, regardless from I_x or I_y, in least squares sense [9].

By applying the above explained interpolation techniques on the edge and other regions respectively, the interpolated image which is equal in size to that of the original image is obtained. The wavelet transform is applied to the interpolated image to obtain again the dense and the sparse components.

B. Signal Recovery

The recovery algorithm used in this scheme is Projection onto Convex Set, which is a convex optimization problem. If a finite signal $f \in R^N$ has to be recovered from a set of K linear measurements which is the vastly underdetermined case $K<<N$, where there are many more unknowns than observations, the signal f cannot be completely recovered from the reduced measurements. However, if f is sparse, in the sense if it can be written as a superposition of a small number of vectors taken from a basis Ψ, then the exact recovery is possible and the 'true' signal f actually is the solution to a simple convex optimization problem.

The prediction of sparse components from the interpolated image, helps in two aspects: first, it could be used as the initialization of the iteration. As known, initialization is important to an iterative algorithm and the initial value need to be in a certain space for final convergence at local optimal. Secondly, the prediction can be used as a reference which helps for converging more rapidly and accurately. Suppose that f is sparse which means only a few of its entries are non-zero; that is, we can write f as a superposition of M spikes [2].

$$f(t) = \sum_{\tau\in T} \alpha_t \delta(t-\tau) \quad (10)$$

for some $T \subset \{0,1,...N-1\}, |T| = M$, where neither the locations nor the amplitudes of the spikes are not known. The central theorem states that for an overwhelming percentage of sets Ω with cardinality obeying the following condition,

$$|\Omega| = K \geq const.M\log N \quad (11)$$

f is the unique solution to the convex optimization problem.

$$\min_{g\in R^N} \|g\|_{l1} := \sum_t |g(t)| \quad \text{subject to} \quad F_\Omega g = y \quad (12)$$

that is, it is possible, with high probability, to recover f from the knowledge of its projection onto a randomly selected K-dimensional subspace [2].

Suppose a Gaussian ensemble is generated by choosing each $(M)_{k,n}$ independently from zero mean, normal distribution with unit variance

$$(M)_{k,n} \sim N(0,1) \quad k=0,1,\ldots,K-1 \quad n=0,1,\ldots,N-1 \tag{13}$$

and it is used to measure a sparse signal f, $y=Mf$. Again if $K \geq const.M \log N$, then f is the unique solution to

$$\min \|g\|_{l1} \text{ subject to } M * g = y \tag{14}$$

When the $K<<N$ measurement ensemble is constructed and used to measure f, we are essentially choosing a K dimensional subspace uniformly at random from the set of all K dimensional subspaces, and then projecting f onto it. The fact that f can be recovered means that although K can be much smaller than N, the projection retains enough critical information to specify f uniquely [2]. The Gaussian measurement ensemble easily allows to extend the results to signals that are sparse in any fixed orthonormal basis Ψ. To recover the signal, the equation (13) is modified to search over coefficient sequences in the Ψ domain.

$$\min \|\alpha\|_{l1} \text{ subject to } M\Psi\alpha = y \tag{15}$$

Because the subspace is chosen uniformly at random, it does not matter which set of axes the signal is aligned with. Mathematically speaking, if M has i.i.d. Gaussian entries and is orthonormal, then the distribution of the random matrix M is exactly the same as that of M; making measurements of f using M and solving (14) will recover signals with M-sparse representations in the domain when the condition $K <= Const* MlogN$ is satisfied [2]. This invariance property makes the Gaussian measurement ensemble especially attractive; we can recover sparse signals in any fixed basis from randomly sampled K measurement vectors with very high probability. To begin, suppose that the condition for exact reconstruction is satisfied; that is, the wavelet coefficients α are non-zero only on a small set T. Let $M'=M\Psi$ be the measurement matrix expressed in the wavelet domain. Since α is the unique solution to (14),

The $l1$ ball $\quad B = \{\beta : \|\beta\|_{l1} <= \|\alpha\|_{l1}\}$ (16)

and the hyperplane $\quad H = \{\beta : M'\beta = y\}$ (17)

meet at exactly one point $\quad B \cap H = \{\alpha\}$ (18)

There are two types of projections in this algorithm, the projection onto H and the projection onto ball B [2]. To find the closest vector $\hat{\beta}$ in H, the following formula is used,

$$\hat{\beta} = \beta + \varphi^{-1} * (y - \varphi * \hat{\beta}) \tag{19}$$

the sparse components which are obtained by applying wavelet transform β is projected onto the above mentioned hyperplane, which will result in an estimate $\hat{\beta}$. This estimate will be projected onto the ball.

To project the vector $\hat{\beta}$ onto l1 ball, a soft-thresholding operation is applied.

$$\begin{aligned} \hat{\beta} &= \beta(t) - \gamma ....\beta(t) > \gamma \\ \hat{\beta} &= 0 ....|\beta(t) \leq \gamma| \\ \hat{\beta} &= \beta(t) + \gamma ....\beta(t) < -\gamma \end{aligned} \tag{20}$$

To determine the threshold,

$$\|\hat{\beta}\|_{l1} \leq \|\hat{\alpha}\|_{l1}, \tag{21}$$

the coefficients are to be sorted out and a linear search has to be performed which will require $O(N \log N)$ operations [2].

V. OBSERVATIONS

TABLE I

IMAGE NAME	RMSE	PSNR WITH ENCODING
Coins	6.45	33.8574
Cameraman	5.65	32.9193

TABLE II

Q tab	CR	BPP	PSNR
10	21.93	0.3646	32.9236
20	18.81	0.4250	32.9241
40	15.58	0.5130	32.8833
60	11.34	0.6810	32.9687
80	10.54	0.7587	32.7670
90	08.25	0.4692	32.8507

CR – Compression Ratio
BPP – Bits Per Pixel
PSNR – Peak Signal to Noise Ratio

Figure 6.a Dense Component ,

Figure 6.b Interpolated dense Component

Figure 7.a Input image

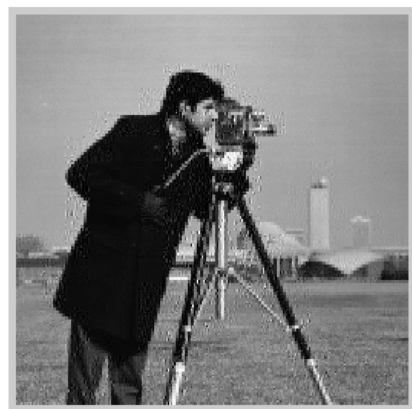

Figure 7.b Recovered image
(Q tab = 10, CR = 21.93, bpp = 0.36)

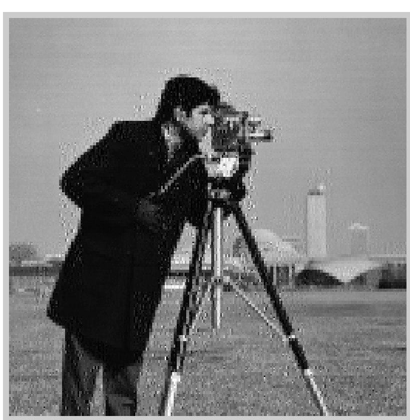

Figure 7.c Recovered image,
(Q tab = 50, CR = 12.58, bpp = 0.64)

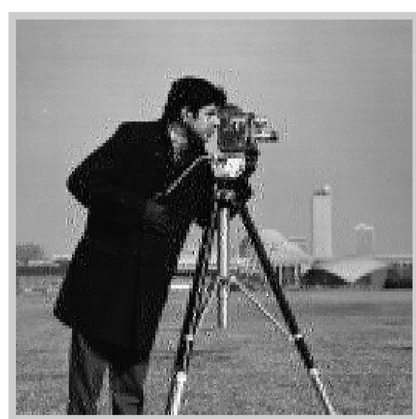

Figure 7.d Recovered image
(Q tab = 90, CR = 8.25, bpp = 0. 46)

VI. CONCLUSIONS

Since the sampling process of a compressive sensing device is simple and collected measurements are already compressed, the devices built based on compressive sensing need less computational power and have cheaper embedded hardware. The reason is that signal information is evenly distributed amongst the measurements and if some measurements are dropped during communication, it is still possible to recover the signal using received measurements [5]. In this paper, a new image representation scheme for the wireless sensor networks is proposed. The work shows that the decomposition of the input data before CS measuring is important and especially useful for reducing the number of measurements and recovery iterations. First, the decomposition removes the unnecessary component, which is not suitable for CS recovery. Second, the prediction by interpolation of the dense component helps the recovery procedure of the sparse component. When the quality of JPEG compression of the dense component is varied, the compression ratio as well as bits per pixel values change and tabulated in Table 2. The future work will be in the reduction of complexity in the arithmetic encoding and decoding of sparse measurements.

REFERENCES

[1] Bing Han a, Feng Wub, Dapeng, *Image representation by compressive sensing for wireless sensor networks*, J. Vis.Commun. Image R., pp.325-333, Elsevier, 2010.

[2] E. Candès, J. Romberg, Practical signal recovery from random projections, in: Wavelet Applications in Signal and Image Processing XI, Proc. SPIE Conf., 2004, pp. 5914-5931.

[3] Emmanuel J. Candès and Michael B. Wakin "An Introduction to Compressive Sampling" in *IEEE Signal processing mag.*, pp.1-20, March 2008.

[4] L. Gan, Block compressed sensing of natural images, in: Proc. Int. Conf. on Digital Signal Processing, Cardiff, UK, pp. 1-5, July 2007.

[5] Justin Romberg "Imaging via Compressive Sampling" in *IEEE Signal processing mag.*, pp.14-20, March 2008.

[6] Reza Pournaghi, McMaster University. (2009) 'Recovery of Compressive Sensed Images With Piecewise Autoregressive Modeling' in: Open Access Dissertations and Theses, pp. 1-88.

[7] Xiangjun Zhang and Xiaolin Wu, Senior Member, IEEE. (2008) 'Image Interpolation by Adaptive 2-D Autoregressive Modeling and Soft-Decision Estimation' in: IEEE Trans. Image Processing, Vol. 17, No. 6, pp. 887-896.

[8] Xin Li, Member, IEEE, and Michael T. Orchard, Fellow, IEEE. (2001) 'New Edge-Directed Interpolation' in: IEEE Trans. Image Processing, Vol. 10, No. 10, pp. 1521-1527.

[9] X. Zhang, X. Wu, F. Wu, *Image Coding on Quincunx Lattice with Adaptive Lifting and Interpolation*, in: Data Compression Conference, 2007, pp. 193–202.

[10] http://www.abmedia.dk/video/coding.htm. Accessed on 01.06.2012.

[11] Chaddha, N. , Agrawal, A. ; Gupta, A. ; Meng, T.H.Y, Variable compression using JPEG, in: Multimedia Computing and Systems, 1994., Proceedings of the International Conference on 15-19 May 1994, pp. 562-570.

[12] http://en.wikipedia.org/wiki/Arithmetic_coding.

[13] rauhut.ins.uni-bonn.de/LinzRauhut.pdf

[14] http://en.wikipedia.org/wiki/Bicubic_interpolation

[15]http://www.see.ed.ac.uk/~tblumens/Sparse/Sparse.html

[16] http://www.mathworks.in/matlabcentral/fileexchange/38518-jpeg-compression/content/jpegimplementation.m