

MRI IMAGE RECONSTRUCTION: USING NON LINEAR CONJUGATE GRADIENT SOLUTION

Jyothi Kiran. G.C

Scholar, Master in Engineering, Department of ECE, UVCE, Bangalore, India

ABSTRACT

The images extracted from Tomography are complex. Tomography images like Magnetic Resonance Images (MRI), angiograms, X-rays or CT scanned images extracted from scanner are further reconstructed to analyze smallest piece or portion of the images for investigation of diseases. The MR images when they extracted from scanner are under-sampled, added with artifacts called interference. L1-norm technique is used to minimize these interferences which are like random noise. A nonlinear conjugate gradient solution is adopted to minimize these noises. The natural images, they are real images, added with different noises like random / speckle / white noise / salt and pepper noise. If they are undersampled and incoherent are reconstructed by calculating the filter coefficients and denoising by conjugant gradient technique. A comparison of the PSNR, SSIM, and RMSE for natural (both color and gray) and medical images is tabulated. Matlab Algorithms are used in this work for reconstruction of images.

KEYWORDS --- conjugate gradient (CG), Compressed Sensing (CS), Magnetic Resonance Imaging (MRI), Sparse, Total Variance (TV), Wavelet Transforms (WT), Incoherence, Incoherent Sampling and Undersampling (US).

I. INTRODUCTION

An image is a two-dimensional array of pixel measurements on a uniform grid. MR images are implicitly sparse, Sparse representation means a matrix with few nonzero values. For example, angiograms are extremely sparse in the pixel representation. Complex medical images may not be sparse in the pixel representation, but they can be sparser in transform domain. To make them sparser Fourier transform or Wavelet Transform is used. WT is more precise transform. In this project, the efforts are made to remove these noises and aliasing interference. The incoherence which is introduced by under sample is removed by Non-linear Conjugate Gradient solution. Using L1- norm technique, the energy is minimized and trying to reduce noise in the data. Using Monte-Carlo Method, the incoherence is reduced based on number of iterations.

Compressed sensing technique used for efficiently acquiring and reconstructing an image, by finding solutions to underdetermined linear systems. This gives the image's sparseness or compressibility in transform domain, allowing the original image to be reconstructed by relatively few measurements.

Images are no more complex. Brightness along a line can be recorded as a set of values measured at equally spaced distances apart, or equivalently, at a set of spatial frequency values. Each of these frequency values is referred to as a frequency component. Image which changes rapidly with high frequency and slowly with low frequency.

In following sections a short description of MRI,K-space, Compressed Sensing, FFT, Incoherence and Wavelet Transforms. In wavelet transform section the explanation of one level and cascading filters, Haar and Daubechies wavelets, quadrature mirror filter. A complete explanation of Image restoration using L1-norm and nonlinear conjugate gradient solution and algorithms are included.

The results, conclusion and future work of the paper is included. And then acknowledge and latest references for different topics is included in this paper.

Block diagram of Wavelet Transformation, reconstruction of Medical images and natural images included.

A table is included in this paper it is comparing the PSNR, SSIM and RMSE of different images.

This paper also includes figures pertaining to output of the research.

II. DESCRIPTION

a) Magnetic Resonance Imaging (MRI)

Magnetic resonance imaging (MRI) or magnetic resonance tomography (MRT) is a medical imaging technique used in radiology to visualize internal structures of the body in detail. MRI makes use of the property of nuclear magnetic resonance (NMR) to image nuclei of atoms inside the body.

An MRI scanner [24] is a device in which the patient lies within a large, powerful magnet. MRI provides distinguish between the different tissues of the body, which is useful in imaging the brain, muscles, the heart and cancers compared with other medical imaging techniques such as computed tomography (CT) or X-rays.

The 2D image measured and acquired by MRI scanner[24] as K-space. The image in K-plane is said be is undersampled and sparsed. From this plane a part or piece of the image is taken and denoised. This is useful for the doctors to investigate diseases present in that part of body. The values measured using resonance frequency of the magnetic field.

b) K – Space

In MRI physics, k-space [23] is the 2D or 3D Fourier transform of the MR image measured. Its complex values are sampled during MR measurement, in a premeditated scheme controlled by a pulse sequence, i.e. an accurately timed sequence of radiofrequency and gradient pulses. K-space [23] refers to the temporary image space, usually a matrix, in which data from digitized MR signals are stored during data acquisition. When k-space is full (at the end of the scan) the data are mathematically processed to produce a final image.

K-space has the same number of rows and columns as the final image and is filled with raw data during the scan. A MR image is a complex-valued map of the spatial distribution of the transverse magnetization in the sample at a specific time point after an excitation. Fourier analysis asserts that low spatial frequencies (near the center of k-space) contain the signal to noise and contrast information of the image, whereas high spatial frequencies (outer peripheral regions of k-space) contain the information determining the image resolution. This is the basis for advanced scanning techniques, such as the keyhole acquisition, as shown in figure 1.

A symmetry property exists in k-space it's information is redundant and an image can be reconstructed using only one half of the k-space.

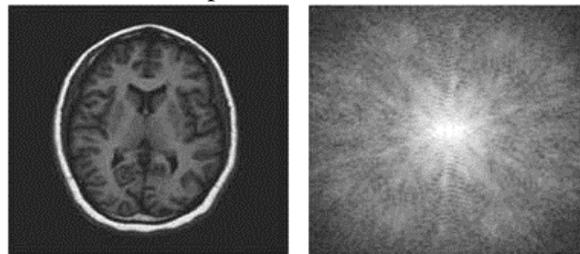


Figure 1. K-space

c) Sparse

The images obtained from the MR scanner are sparsed. That is, sparsity is implicit in MR images. To make more sparser for the manipulation of complicated images. The sparse[6] representation achieved either by Fast Fourier Transformation or Wavelet Transformation.

The sparse values in transform domain are called sparse coefficients or Fourier coefficients or wavelet coefficients.

Using these coefficients images are manipulated and reconstructed. The image is compressed without loss of data either by FFT or Wavelet Transform. This process reduces image processing time and reduces the memory space for storage of the image.

In MRI applications the imaging speed is very important. The speed at which data can be collected in MRI is limited by,

1. Physical (gradient amplitude and slew-rate) and
2. Physiological (nerve stimulation) constraints.

Therefore many researches are seeking for methods to reduce the amount of acquired data so that using these few data the image can be reconstructed.

d) Compressed Sensing

Compressed sensing (CS) also known as compressive sensing, compressive sampling [7], or sparse sampling is a signal processing technique for efficiently acquiring and reconstructing a signal, by finding solutions to underdetermined linear systems. This gives the signal's sparseness or compressibility in some domain, allowing the entire signal to be determined from relatively few measurements.

A prominent application is Tomography, such as MRI, CT scan, and X-ray.

Compressed sensing is redundant in many interesting signals—they are not pure noise. They contain many coefficients close to or equal to zero when represented in some domain. This method used in many forms of lossy compression.

Solutions to Compressed sensing [7] starts with taking a weighted linear combination of samples also called compressive measurements in a basis in which the signal is known to be sparse. Converting the image back into the intended domain involves solving an underdetermined matrix equation since the number of compressive measurements taken is smaller than the number of pixels in the full image.

The least-squares solution used to such problems is to minimize the L_2 -norm—that is, minimize the amount of energy in the system.

To enforce the sparsity constraint when solving for the underdetermined system of linear equations, one can minimize the number of nonzero components of the solution. The function counting the number of non-zero components of a vector was called the L_0 "norm" by David Donoho.

L_1 - norm, is used for compressed sensing which was introduced by Laplace, when the constructed images did not satisfy the Nyquist's–Shannon criterion. Around 2004 Emmanuel Candès, Terence Tao and David Donoho discovered that by minimum amount of data reconstruction of an image was possible even though the amount of data would be insufficient by the Nyquist's–Shannon criterion. This is the basis of compressed sensing.

e) Incoherence

Suppose we want our representation basis Ψ to be as "different" from our sampling basis Φ as possible. One measurement for the similarity of two bases is the coherence

$$\mu(\Phi, \Psi) = \sqrt{n} \cdot \max(\Phi_k, \Psi_j), \quad \text{where } 1 \leq k, j \leq n \quad \text{--Eq. 1}$$

which measures the maximum coherence between any two elements of Ψ and Φ . Linear algebra shows that the values of coherence obey $\mu(\Psi, \Phi) \in [1, \sqrt{n}]$. A lower value indicates a smaller coherence and therefore, if $\mu(\Psi, \Phi) = 1$ the bases are called maximally incoherent. A low coherence value indicates that a signal which is sparse in one of the bases will have a dense representation in the other one.

Undersampled [18] means the signal that does not satisfy the Nyquist's criteria. This produces aliasing which are called interference in the sparse transform domain as an essential ingredient for CS [5], [6].

The images from MRI scanner are in random k -space exhibits a very high degree of incoherence. Only a random subset of k -space [3] was chosen to simplify the mathematical proofs in CS. This interference is incoherent and they act like Additive White Gaussian Noise. This type of random subset of k -space undersampled data is referred as Incoherent Sampling.

f) Fast Fourier Transform (FFT)

FFT is discrete transform used in Fourier analysis transforms one function which is in the time domain into frequency domain representation. The input function is discrete and are created by sampling a continuous function, and they have limited duration. For example, the amplitude of a

person's voice over a time. FFT uses both sine's and cosines complex [21],[22] values and evaluates frequency components to reconstruct the finite data.

FFT used in signal processing, solving partial differential equations, for convolution operations, multiplying large numbers and in complex image processing.

$$F(u, v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) e^{-2\pi i(xu+yv)/N}, \quad \text{---Eq. 2}$$

FFT is given by,

The Inverse FFT is used to reproduce the original data and is given by,

$$f(x, y) = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F(u, v) e^{2\pi i(xu+yv)/N}. \quad \text{---Eq. 3}$$

g) Wavelet Transforms

A wavelet series is a representation of a square-integral (real- or complex-valued) function by a certain orthonormal series generated by a wavelet. Wavelet transformation is one of the most popular transform of the time-frequency-transformations.

A function $\Psi \in L^2(\mathbb{R})$ is called an orthonormal wavelet if it can be used to define a Hilbert basis, that is a complete orthonormal system, for the Hilbert space $L^2(\mathbb{R})$ of square integral functions. The Hilbert basis is constructed as the family of functions $\{\Psi_{jk}; j, k \in \mathbb{Z}\}$ by means of dyadic translations and dilations of Ψ ,

$\Psi_{jk}(x) = 2^{-j/2} \Psi(2^{-j}x - k)$ for integers $j, k \in \mathbb{Z}$. This family is an orthonormal system if it is orthonormal under the inner product $(\Psi_{jk}, \Psi_{lm}) = \delta_{jl} \delta_{km}$ where δ_{jl} is the Kronecker delta and (f, g) is the standard inner product

$$(f, g) = \int_{-\infty}^{\infty} f(x) \overline{g(x)} dx \quad \text{on } L^2(\mathbb{R}) \quad \text{---Eq. 4}$$

The every function $h \in L^2(\mathbb{R})$ may be expanded in the basis as

$$h(x) = \sum_{j,k=-\infty}^{\infty} c_{jk} \psi_{jk}(x) \quad \text{---Eq. 5}$$

With convergence of the series to be convergence in norm. Such a representation of a function f is known as a wavelet series. This implies that an orthonormal wavelet is self-dual. The integral wavelet transform is defined as

$$[W_{\psi} f](a, b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} \overline{\psi\left(\frac{x-b}{a}\right)} f(x) dx \quad \text{---Eq. 6}$$

The wavelet coefficients c_{jk} are then given by $c_{jk} = [W_{\psi} f](2^{-j}, k2^{-j})$. Here, $a = 2^{-j}$ is called the binary dilation or dyadic dilation, and $b = k2^{-j}$ is the binary or dyadic position.

Wavelet Transforms are used in image compression, Signal processing of accelerations for gait analysis, for fault detection to find spatial and temporal variability and to find Sea Surface Temperature.

There are different types of wavelet transforms for specific purposes. The common ones are : Mexican hat wavelet, Haar Wavelet, Daubechies wavelet, Triangular wavelet.

h) Discrete Wavelet Transforms

A discrete wavelet transform (DWT) [19] is any wavelet transform for which the wavelets are discretely sampled. A key advantage of wavelet transforms is captures both frequency and location. 2D discrete wavelet transform [7],[8] used in JPEG2000.

The original image is high-pass filtered, yielding the three large images, each describing local changes in brightness in the original image. It is then low-pass filtered and downsampled, yielding an approximation image; this image is high-pass filtered to produce the three smaller detail images, and low-pass filtered to produce the final approximation image in the upper-left.

Haar wavelets : The first DWT was invented by the Hungarian mathematician Alfréd Haar. For an input represented by a list of 2^n numbers, the Haar wavelet transform is simply pair up input values, storing the difference and passing the sum. This process is repeated recursively, pairing up the sums to provide the next scale: finally resulting in $2^n - 1$ differences and one final sum.

Daubechies wavelets: The most commonly used set of discrete wavelet transforms [7],[8] was formulated by the Belgian mathematician Ingrid Daubechies in 1988.

One level of the transform : The DWT of a signal x is calculated by passing it through a series of filters.

First the samples are passed through a low pass filter with impulse response g resulting in a convolution of the two:

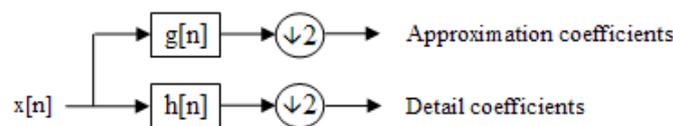
$$y[n] = (x * g)[n] = \sum_{k=-\infty}^{\infty} x[k]g[n - k]. \quad \text{---Eq.7}$$

The signal is also decomposed simultaneously using a high-pass filter h . The outputs giving the detail coefficients from the high-pass filter and approximation coefficients from the low-pass. The two filters are related to each other and they are known as a quadrature mirror filter [9],[10].

The filter outputs are then subsampled by 2 (Mallat's and the common notation is the opposite, g - high pass and h - low pass):

$$y_{low}[n] = \sum_{k=-\infty}^{\infty} x[k]g[2n - k] \quad ; \quad y_{high}[n] = \sum_{k=-\infty}^{\infty} x[k]h[2n - k] \quad ; \quad \text{---Eq.8}$$

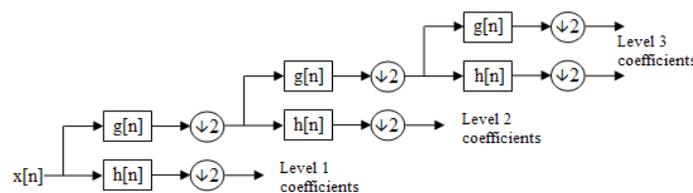
This decomposition has halved the time resolution since only half of each filter output characterizes the signal. However, each output has half the frequency band of the input so the frequency resolution has been doubled.



With the subsampling the subsampling operator \downarrow , $(y \downarrow k [n]) = y [kn]$

The above summation can be written as, $y_{low} = (x * g) \downarrow 2$; $y_{high} = (x * h) \downarrow 2$; ---Eq. 9

Cascading and Filter banks : The decomposition is repeated to further increases the frequency resolution and the approximation coefficients decomposed with high and low pass filters and then down-sampled. This is represented as a binary tree with nodes representing a sub-space with different time-frequency localization. The tree is known as Filter Bank.



At each level in the above diagram the signal is decomposed into low and high frequencies. Due to the decomposition process the input signal must be a multiple of 2^n where n is the number of levels. For example a signal with 32 samples, frequency range 0 to f_n and 3 levels of decomposition, 4 output scales are produced.

Quadrature mirror filter : A binomial QMF is also called an orthonormal binomial quadrature mirror filter – is an orthogonal wavelet with perfect reconstruction (PR) [1] was designed by Ali Akansu, et al. published in 1990[1],[2]. These binomial-QMF filters are identical to the wavelet filters designed independently by Ingrid Daubechies. It was shown that the magnitude square functions of low-pass and high-pass binomial-QMF filters are the unique maximally flat functions in a two-band PR-QMF design framework.

In digital signal processing, a quadrature mirror filter is a filter whose magnitude response is mirror image about $\pi/2$ of that of another filter. Together these filters are known as the Quadrature Mirror Filter pair.

The different types of orthonormal wavelets are Haar, Daubechies, Coiflets Mallat, are generated by scaling functions which satisfy a quadrature mirror filter relationships.

Daubechies wavelets : The Daubechies [5],[11]wavelets have the highest number of vanishing moments. They are orthogonal wavelets D2-D20[16] (even index numbers only). The index number refers to the number N of coefficients. Each wavelet has a number of zero moments or vanishing moments equal to half the number of coefficients. For example, D2 (the Haar wavelet) has one

vanishing moment, D4 has two, etc. A vanishing moment limits the wavelet's ability. D4 encodes polynomials with two coefficients, i.e. constant and linear signal components; and D6 encodes 3-polynomials, i.e. constant, linear and quadratic signal components.

The scaling functions for D2-20, The wavelet coefficients are derived by reversing the order of the scaling function coefficients and then reversing the sign of every second one, (i.e., D4 wavelet = {-0.1830127, -0.3169873, 1.1830127, -0.6830127}). Mathematically, this looks like $b_k = (-1)^k a_{N-1-k}$, where, k is the coefficient index, b is a coefficient of the wavelet sequence and a is a coefficient of the scaling sequence. N is the wavelet index, i.e., 2 for D2.

Forward DWT : One computes recursively, starting with the coefficient sequence $s^{(J)}$ and counting down from k=J-1 down to some M<J, single application of a wavelet filter bank, with filters $g=a^*$, $h=b^*$

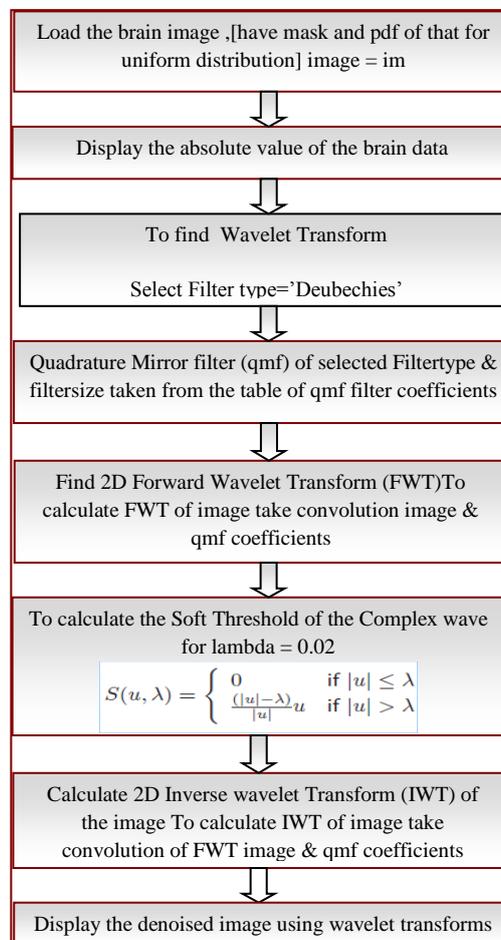
$$s_n^{(k)} := \frac{1}{2} \sum_{m=-N}^N a_m s_{2n+m}^{(k+1)} \text{ or } s^{(k)}(z) := (\downarrow 2)(a^*(z) \cdot s^{(k+1)}(z)) \text{ -- Eq. 10}$$

and

$$d_n^{(k)} := \frac{1}{2} \sum_{m=-N}^N b_m s_{2n+m}^{(k+1)} \text{ or } d^{(k)}(z) := (\downarrow 2)(b^*(z) \cdot s^{(k+1)}(z)) \text{ --- Eq.11}$$

For k=J-1,J-2,...,M and all $n \in \mathbb{Z}$. In the Z-transform notation: recursive application of the filter bank

III. BLOCK DIAGRAM OF WAVELET TRANSFORMS



The downsampling operator $(\downarrow 2)$ reduces an infinite sequence, given by its Z-transform, which is simply a Laurent series, to the sequence of the coefficients with even indices,

$$(\downarrow 2)(c(z)) = \sum_{k \in \mathbb{Z}} c_{2k} z^{-k} \text{ . ---Eq. 12}$$

The starred Laurent-polynomial $a^*(z)$ denotes the adjoint filter, it has time reversed adjoints coefficients

$$a^*(z) = \sum_{n=-N}^N a_{-n}^* z^{-n} \quad \text{---Eq.13}$$

(The adjoint of a real number being the number itself, of a complex number its conjugate, of a real matrix the transposed matrix, of a complex matrix its hermitian adjoint).

Multiplication is polynomial multiplication, which is equivalent to the convolution of the coefficient sequences. It follows that

$$P_k[f](x) := \sum_{n \in \mathbb{Z}} s_n^{(k)} \phi(2^k x - n) \quad \text{---Eq. 14}$$

is the orthogonal projection of the original signal f or at least of the first approximation $P_j[f](x)$ onto the subspace V_k , that is, with sampling rate of 2^k per unit interval. The difference to the first approximation is given by $P_j[f](x) = P_j[f](x) + D_k[f](x) + \dots + D_{j-1}[f](x)$ where, the difference or detail signals are computed from the detail coefficients as

$$D_k[f](x) := \sum_{n \in \mathbb{Z}} d_n^{(k)} \psi(2^k x - n) \quad \text{---Eq. 15}$$

where, Ψ denotes the mother wavelet of the wavelet transform.

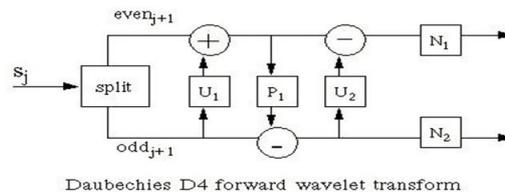
Inverse DWT : Given the coefficient sequence $s^{(M)}$ for some $M < J$ and all the difference sequences $d^{(k)}$, $k=M, \dots, J-1$, one computes recursively,

$$s_n^{(k+1)} := \sum_{k=-N}^N a_k s_{2n-k}^{(k)} + \sum_{k=-N}^N b_k d_{2n-k}^{(k)} \quad \text{or} \quad s^{(k+1)}(z) = a(z) \cdot (\uparrow 2)(s^{(k)}(z)) + b(z) \cdot (\uparrow 2)(d^{(k)}(z)) \quad \text{---Eq.16}$$

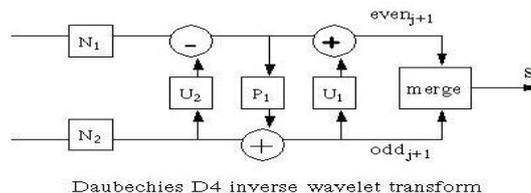
For, $k=J-1, J-2, \dots, M$ and all $n \in \mathbb{Z}$. In the Z-transform notation. The upsampling operator $(\uparrow 2)$ creates zero-filled holes inside a given sequence. That is, every second element of the resulting sequence is an element of the given sequence, every other second element is zero or

$$(\uparrow 2)(c(z)) := \sum_{n \in \mathbb{Z}} c_n z^{-2n} \quad \text{---Eq. 17}$$

This linear operator is, in the Hilbert space $\ell^2(\mathbb{Z}, \mathbb{R})$, the adjoint to the downsampling operator $(\downarrow 2)$. For eg. Fig 2.D4 Forward WT



And Fig 3. D4 Inverse WT



IV. IMAGES RESTORATION

The Image Restoration Technique involves two methods,

1. Constrained method
2. Unconstrained method.

In both the methods, the image restoration achieved by Least Square Minimization solution.

1. In constrained Method a linear algebraic method is used. This is the processes of nonlinear image reconstruction to the CS setting. Suppose, ' m ' is the image of interest in a vector ,

Let ‘ ψ ’ denote the linear operator that transforms from pixel representation into a sparse representation, and

Let ‘Fu’ be the undersampled Fourier transform, corresponding to one of the k-space

The reconstruction is obtained by solving the following constrained optimization problem is by:

minimize $\|\Psi\mathbf{m}\|_1$ and To Show that $\|\mathbf{F}_u\mathbf{m} - \mathbf{y}\|_2 < \epsilon$. ---Eq 18

Here m is the reconstructed image, where y is the measured k-space data from the scanner and ϵ controls the fidelity of the reconstruction to the measured data.

The threshold parameter ϵ is usually set below the expected noise level. The objective function is the L_1 norm, which is defined as $\|x\|_1 = \sum_i |x_i|$. Minimizing $\|\Psi\mathbf{m}\|_1$ promotes sparsity. The constraint $\|\mathbf{F}_u\mathbf{m} - \mathbf{y}\|_2 < \epsilon$ enforces data consistency. In words, among all solutions which are consistent with the acquired data, in the above equation finds a solution which is compressible by the transform ψ .

When finite-differences are used as a sparsifying transform, the objective is often referred to as Total-Variation(TV), since it is the sum of the absolute variations in the image. The objective then is usually written as TV(m). Even when using other sparsifying transforms in the objective, it is useful to include a TV penalty[24]. Total variation denoising, also known as total variation regularization. It is a process, used in digital image processing, that has applications in noise removal. It is based on the principle that signals with excessive and possibly spurious detail have high total variation, that is, the integral of the absolute gradient of the signal is high. According to this principle, reducing the total variation of the signal a close match to the original signal, removes unwanted detail while preserving important details such as edges. The concept was developed by Rudin et al. in 1992.

This noise removal technique has advantages over simple techniques such as linear smoothing or median filtering which reduce noise but at the same time smooth away edges to a greater or lesser degree. By contrast, total variation denoising is effective at simultaneously preserving edges and smoothing away noise in flat regions, even at low signal-to-noise ratios.

for a digital signal y_n , we can, define the total variation as:
$$V(y) = \sum_n |y_{n+1} - y_n|$$
 -Eq.19

Given an input signal x_n , the total variation denoising is used to find an approximation, let it be y_n , that has smaller total variation than x_n but is "close" to x_n . One measure of closeness is the sum of square errors:

$$E(x, y) = \frac{1}{2} \sum_n (x_n - y_n)^2 \quad \text{---Eq.20}$$

So the total variation denoising problem amounts to minimizing the following discrete functional over the signal y_n :

$$E(x, y) + \lambda V(y) \quad \text{---Eq.21}$$

By differentiating this functional with respect to y_n , we can derive a corresponding Euler-Lagrange equation, that can be numerically integrated with the original signal x_n as initial condition. Since this is a convex functional, techniques from convex optimization can be used to minimize it and find the solution of y_n .

λ is a called regularization parameter play a critical role in the denoising process.

When, $\lambda = 0$, there is no denoising and the result is the same as the input signal.

$\lambda \rightarrow \infty$, the total variation term plays a strong role, which forces the result to have smaller total variation, at the expense of being less like the input (noisy) signal. Thus, the choice of regularization parameter is critical to achieving just the right amount of noise removal. The total variation norm is,

$$V(y) = \sum_{i,j} \sqrt{|y_{i+1,j} - y_{i,j}|^2 + |y_{i,j+1} - y_{i,j}|^2} \quad \text{----Eq. 22}$$

A variation that is sometimes used, since it may sometimes be easier to minimize, is an anisotropic version

$$V_{\text{aniso}}(y) = \sum_{i,j} \sqrt{|y_{i+1,j} - y_{i,j}|^2 + |y_{i,j+1} - y_{i,j}|^2} = \sum_{i,j} |y_{i+1,j} - y_{i,j}| + |y_{i,j+1} - y_{i,j}| \quad \text{---Eq.23}$$

The standard total variation denoising problem is still of the form $\min_y E(x, y) + \lambda V(y)$ where E is the 2D L2-norm.

This can be considered as requiring the image to be sparse by both the specific transform and finite-differences at the same time.

In this case equation can be written as, minimize $\|\Psi\mathbf{m}\|_1 + \alpha \text{TV}(\mathbf{m})$ and to Show that $\|\mathbf{F}_u\mathbf{m} - \mathbf{y}\|_2 < \varepsilon$ ---Eq. 24

where, α trades ψ with finite-differences sparsity.

Minimizing the L1 - norm of an objective result in a sparse solution.

On the other hand, minimizing the L2 norm, which is defined as $\|\mathbf{x}\|_2 = (\sum_i |x_i|^2)^{1/2}$ ---Eq. 24

Special purpose methods for solving the above equation when CS was first introduced. Proposed methods include: interior point methods projections onto convex sets, iterative soft threshold ,and iteratively reweighted least squares. In the following chapter it is described, using non-linear conjugate gradients and backtracking line-search.

V. NON-LINEAR CONJUGATE GRADIENT SOLUTION OF CS OPTIMIZATION PROCEDURE

Gradient [19], [20] is the pixel intensity of the image which increases horizontally or vertically. The gradient (or gradient vector field) of a scalar function $f(x_1, x_2, x_3, \dots, x_n)$ is denoted ∇f or $\vec{\nabla} f$ where ∇ (the nabla symbol) denotes the vector differential operator, del. The notation "grad(f)" is also commonly used for the gradient. The gradient of f is defined as the unique vector field whose dot product with any vector \mathbf{v} at each point \mathbf{x} is the derivative of ∇f along \mathbf{v} . That is, $(\nabla f(\mathbf{x})) \cdot \mathbf{v} = D_{\mathbf{v}} f(\mathbf{x})$. $\|\mathbf{F}_u\mathbf{m} - \mathbf{y}\|_2$ is a constrained convex optimization problem [19].

2. Consider the unconstrained problem in so-called Lagrange form:

$$\text{argmin} \|\mathbf{F}_u\mathbf{m} - \mathbf{y}\|_2^2 + \lambda \|\Psi\mathbf{m}\|_1 \quad \text{---Eq.25}$$

Where, ' λ ' is a regularization parameter that determines the trade-off between the data consistency and the sparsity. The parameter ' λ ' can be selected appropriately such that the solution of above is exactly as same as constrained convex optimization. The value of ' λ ' can be determined by solving equation for different values, and then choosing ' λ ' so that

$$\|\mathbf{F}_u\mathbf{m} - \mathbf{y}\|_2 < \varepsilon \quad \text{---Eq.26}$$

For a quadratic function $f(\mathbf{x})$: $f(\mathbf{x}) = \|\mathbf{F}_u\mathbf{m} - \mathbf{y}\|^2$ ---Eq. 27

The minimum of f is obtained when the gradient is 0: $\nabla_x f = 2 \mathbf{F}_u' (\mathbf{F}_u \mathbf{m} - \mathbf{y}) = 0$ ---Eq. 28

Where as linear conjugate gradient seeks a solution to the linear equation $\mathbf{F}_u' \mathbf{F}_u \mathbf{m} = \mathbf{F}_u' \mathbf{y}$, the nonlinear conjugate gradient method is generally used to find the local minimum of a nonlinear function using its gradient $\nabla_x f$ alone. It works when the function is approximately quadratic near the minimum, which is the case when the function is twice differentiable at the minimum.

Given a function $f(\mathbf{x})$ of N variables to minimize, its gradient $\nabla_x f$ indicates the direction of maximum increase. One simply starts in the opposite (steepest descent) direction: $\Delta \mathbf{x}_0 = -\nabla_x f(\mathbf{x}_0)$ with an adjustable step length α and performs a line search in this direction until it reaches the minimum of f : $\alpha = \arg \min f(\mathbf{x}_0 + \alpha \Delta \mathbf{x}_0)$

$\mathbf{x}_1 = \mathbf{x}_0 + \alpha_0 \Delta \mathbf{x}_0$ After this first iteration in the steepest direction $\Delta \mathbf{x}_0$, the following steps constitute one iteration of moving along a subsequent conjugate direction \mathbf{S}_n , where $\mathbf{S}_0 = \Delta \mathbf{x}_0$:

1. Calculate the steepest direction: $\Delta \mathbf{x}_n = -\nabla_x f(\mathbf{x}_n)$,
2. Compute β_n according to one of the formulas below,
3. Update the conjugate direction: $\mathbf{S}_n = \Delta \mathbf{x}_n + \beta_n \mathbf{S}_{n-1}$
4. Perform a line search: optimize $\alpha_n = \text{argmin} f(\mathbf{x}_n + \alpha \mathbf{S}_n)$,
5. Update the position: $\mathbf{x}_{n+1} = \mathbf{x}_n + \alpha_n \mathbf{S}_n$,

With a pure quadratic function the minimum is reached within N iterations, but a non-quadratic function will make slower progress. Subsequent search directions lose conjugacy requiring the search direction to be reset to the steepest descent direction at least every N iterations, or sooner if progress stops. However, resetting every iteration turns the method into steepest descent. The algorithm stops when it finds the minimum, determined when no progress is made after a direction reset or when some tolerance criterion is reached.

Within a linear approximation, the parameters α and β are the same as in the linear conjugate gradient method but have been obtained with line searches. Three of the best known formulas for β_n are titled Fletcher-Reeves (FR), Polak-Ribière (PR), and Hestenes-Stiefel (HS) after their developers. They are given by the following formulas:

• Fletcher-Reeves:
$$\beta_n^{FR} = \frac{\Delta x_n^T \Delta x_n}{\Delta x_{n-1}^T \Delta x_{n-1}} \quad \text{---Eq.29}$$

• Polak-Ribière:
$$\beta_n^{PR} = \frac{\Delta x_n^T (\Delta x_n - \Delta x_{n-1})}{\Delta x_{n-1}^T \Delta x_{n-1}} \quad \text{---Eq. 30}$$

• Hestenes-Stiefel:
$$\beta_n^{HS} = -\frac{\Delta x_n^T (\Delta x_n - \Delta x_{n-1})}{s_{n-1}^T (\Delta x_n - \Delta x_{n-1})} \quad \text{---Eq. 31}$$

These formulas are equivalent for a quadratic function, but for nonlinear optimization the preferred formula is a matter of heuristics or taste. A popular choice is $\beta = \max\{0, \beta^{PR}\}$ which provides a direction reset automatically. But the α and β are fixed values taken as constant.

The RMS value of two data is calculated by the formula, :

$$RMS = \sqrt{\frac{1}{n} \sum x^2(t)} = \sqrt{\frac{1}{n^2} \sum |X(f)|^2} = \sqrt{\sum_n \left| \frac{X(f)}{n} \right|^2}. \quad \text{---Eq. 32}$$

Peak Signal to Noise Ratio(PSNR) is used to measure the quality of reconstruction of lossy compression codec's. The signal in this case is the original data, and the noise is the error introduced by compression. When comparing compression codecs, PSNR is defined as, $PSNR = 10 \cdot \log_{10} (MAX1^2/MSE)$; $PSNR = 20 \cdot \log_{10} (MAX1) - 10 \cdot \log_{10} (MSE)$ ---Eq. 33

MAX1 is the maximum possible pixel value of the image. And MSE is the root mean square error value. PSNR of the reconstructed image is between 30dB to 50dB approximately.

The structural similarity (SSIM) index is a method for measuring the similarity between two images. The SSIM index is a the measuring of image quality based on an initial uncompressed or distortion-free image as reference. SSIM is designed to improve peak signal-to-noise ratio (PSNR) and mean squared error(MSE).

The iterative algorithms for L1-norm can be written in following manner:

Input parameters,

y - k-space measurements

Fu - undersampled Fourier operator associated with the measurements - sparsifying transform operator

λ - a data consistency tuning constant

Optional Parameters:

TolGrad - stopping criteria by gradient magnitude (default 10⁻⁴)

MaxIter- stopping criteria by number of iterations (default 100)

α, β - line search parameters (defaults $\alpha = 0.05, \beta = 0.6$)

Output:

M – the numerical approximation.

Step 1- Iterations

Step 2 - Backtracking line search: means from highest value till the minimum value the process is iterated.

Step 3 - The conjugate gradient requires the computation of $f(m)$ which is, $\nabla f(m) = 2F_u^*(F_u m - y) + \lambda \nabla \|\Psi m\|_1$ --

Step 4 - The L¹ norm is the sum of absolute values. The absolute value function is not a smooth Function and as a result is not well defined for all values of 'm'. Instead, we approximate the absolute value with a smooth function by using the relation, $|x| \approx \sqrt{x^*x + \mu}$, where μ is a positive smoothing parameter.

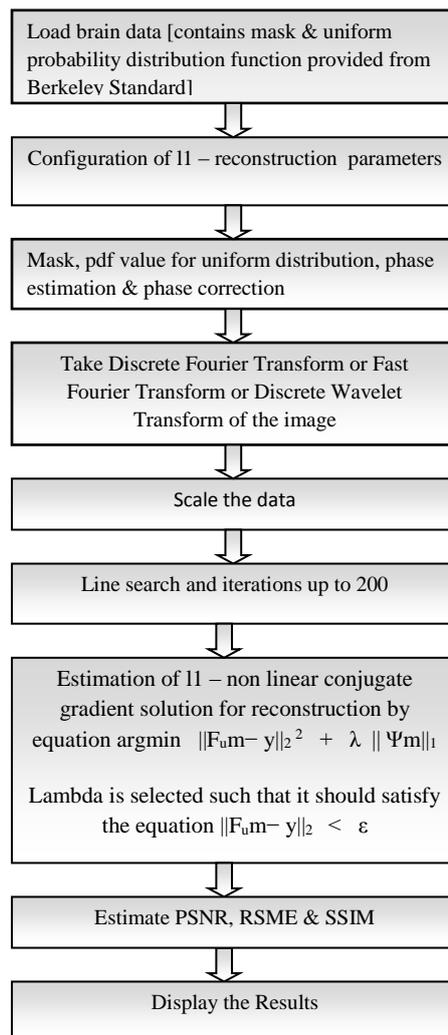
Step 5 - Differentiating above equation we get, $\frac{d|x|}{dx} \approx \frac{x}{\sqrt{x^*x + \mu}}$.

Step 6 – If W is the diagonal matrix the elements are, $w_i = \sqrt{(\Psi m)_i^* (\Psi m)_i + \mu}$.

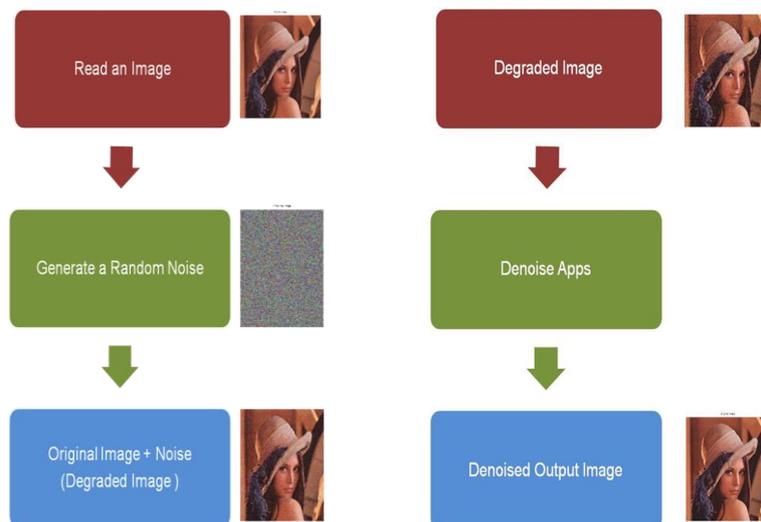
Step 7 – Then the conjugate gradient can be approximated as, $\nabla f(m) \approx 2F_u^*(F_u m - y) + \lambda \Psi^* W^{-1} \Psi m$

The smoothing factor for this equation is $\mu \in [10^{-15}, 10^{-6}]$. The number of CG iterations depends on images to images.

VI. BLOCK DIAGRAM OF MEDICAL IMAGES



VII. BLOCK DIAGRAM FOR NATURAL IMAGES



VIII. RESULTS

In this paper the scanned image assumed to be consists of interference by undersampling. To denoise and reconstruct it a piece or part of image is considered, masked and taken probability distribution function and reconstructed. PSNR, RMSE[4] and SSIM of it is calculated the result is as shown below. In fig.2 a original image of MRI brain. Fig.3.it is masked added with interference and fig.5 shows the piece of the same image and reconstructed we can see the denoised using L1- norm nonlinear conjugate gradient solution as explained in the previous session. Fig 6. Shows the Histogram of the original, noisy and denoised image it is observeded that a good denoising improvement.

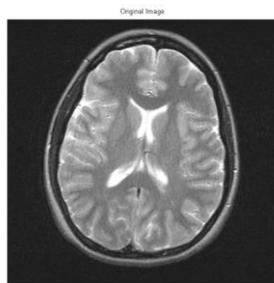


Fig.2. Original image

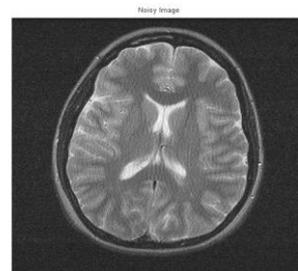


Fig.3. Noisy brain image

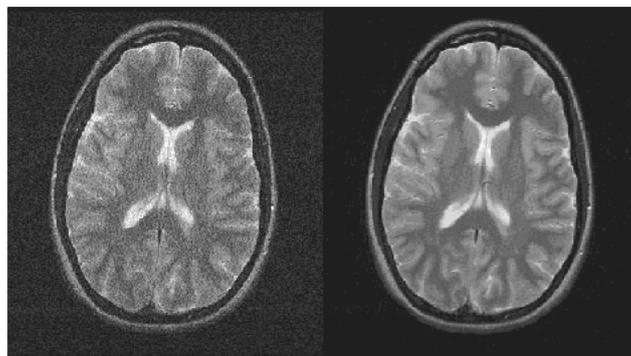


Fig.4. Comparing original and noisy image

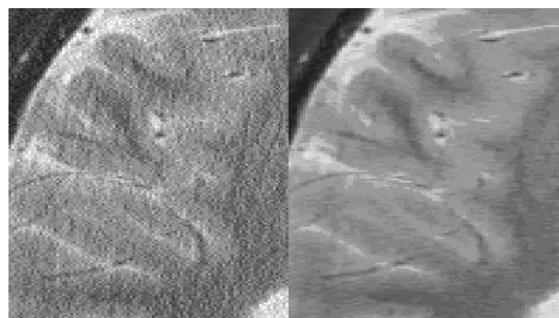


Fig .5.Brain MRI Image Part of the brain image extracted denoised and reconstructed

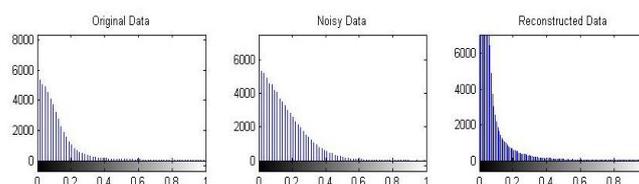


Fig 6. Histogram of the brain

Comparison of PSNR, MSE and SSIM for different images is tabulated in Table.1.

Similarly, the concept of nonlinear conjugate gradient solution is applied to natural images (for both color and gray) where they are undersampled and added with different types of noise. The different noises with different sigma values applied and output is observed with Histograms. In case of natural images logarithmic values are considered.

IX. CONCLUSION

The images scanned from MRI scanner are further reconstructed for the use of doctors for investigation of diseases. The images are compressed by Wavelet transform. By WT the edges of complex images are detected and phase encodes are calculated. By minimizing energy using L1 – norm the hidden objects are detected. By using Non-linear conjugate gradient the incoherence of the image is minimized. Using backtracking line-search the percentage of error is reduced with minimum error. Using Monte-Carlo Method, the interference in the undersampled image are minimized. Finally, PSNR, SSIM and RMSE are tabulated for different images. It is found that the RMSE is reduced. The PSNR and SSIM are improved approximately about 90%.

In the case of natural images 95% of the added noise is removed. Even though the percentage of noise (sigma value) is high the noise is removed with filter coefficients and many iterations and logarithm conjugate gradient method. It is tested for both color and grey images. Here also PSNR, SSIM and RMSE are improved.

Using this technique i.e. MRI Image Reconstruction using non linear conjugate gradient solution small piece of the complex image can be separated and the defects in that piece of image can be detected. It is very helpful to doctors to detect the defects in sensitive parts of the body.

X. FUTURE WORK

This concept can be adapted to T2 weighted multi-slice data of a brain. This concept is enhanced for video images as well as 3D images. This can be adapted to multi-slice fast spin-echo brain imaging also. Using the same CS reconstruction concept 5-fold acquisition data can be reconstructed.

Table 1. Comparison of psnr, rmse and ssim for different images

SL NO	IMAGE	PSNR(IN DB)		SSIM(IN DB)		RMSE	
		Noisy	Denoise	Noisy	Denoise	Noisy	Denoise
1	Parrot (color)	5.17	13.91	0.58	0.83	27.71	16.76
2	Elephant (color)	7.64	16.27	0.53	0.89	14.99	14.99
3	Lena (color)	6.89	13.25	0.55	0.72	31.84	31.84
4	Brain (medical)	27.78	40.56	0.99	1.00	0.026	0.00675
5	Phantom (medical)	21.89	43.98	0.48	0.99	5.74	0.113
6	Hero (color)	5.67	13.15	0.68	0.77	29.63	29.63
7	Lena (B&W)	7.30	15.20	0.72	0.87	23.84	23.84
8	Peppers (B&W)	6.85	16.29	0.80	0.87	39.87	8.3
9	Deer (color)	6.31	14.78	0.50	0.70	22.24	16..25
10	Barbara (B&W)	7.19	11.98	0.49	0.67	25.72	12.85

ACKNOWLEDGEMENT

I express my deep sense and gratitude to my guide Dr. Sreenivasa Murthy, Associative Professor, department of ECE, UVCE, Bangalore University, Bangalore, and Karnataka State, India.

I am very much thankful to my husband and my daughters Rithu and Riya for their moral support and inspiration. And I am very thankful to the only one God

REFERENCES

- [1] A.N. Akansu Multiplierless Suboptimal PR-QMF Design Proc. SPIE 1818, *Visual Communications and Image Processing*, p. 723, November, 1992
- [2] A.N. Akansu Multiplierless PR Quadrature Mirror Filters for Subband Image Coding *IEEE Trans. Image Processing*, p. 1359, September 1996
- [3] G. Beylkin, R. Coifman, V. Rokhlin, "Fast wavelet transforms and numerical algorithms" *Comm. Pure Appl. Math.*, 44 (1991) pp.141–183 http://www.bearcave.com/misl/misl_tech/wavelets/lifting/average.jpg
- [4] Cartwright, Kenneth V (Fall 2007). "Determining the Effective or RMS Voltage of Various Waveforms without Calculus". *Technology Interface* 8 (1): 20 pages
- [5] Daubechies Wavelet Filter
- [6] Donoho et. al (*Sparse Solution of Underdetermined Linear Equations by Stagewise Orthogonal Matching Pursuit*, 2006, Stanford University, Statistics Department) ; roth@cs.brown.edu
- [7] M. Davenport, "The Fundamentals of Compressive Sensing", IEEE Signal Processing Society Online Tutorial Library, April 12, 2013.
- [8] Duren, P. (1970), *Theory of H^p -Spaces*, New York: Academic Press
- [9] Greiser A, von Kienlin M. *Efficient k-space sampling by density-weighted phase-encoding. Magn Reson Med* 2003; 50:1266–1275.
- [10] Hardware implementation of wavelets
- [11] Hazewinkel, Michiel, ed. (2001), "Daubechies wavelets", *Encyclopedia of Mathematics*, Springer, ISBN 978-1-55608-010-4
- [12] I. Daubechies "Orthonormal bases of compactly supported wave lets", *IEEE Trans. Inform. Theory*, vol. 36, pp.961 -1005 1990
- [13] I. Kaplan, The Daubechies D4 Wavelet Transform.
- [14] Marseille GJ, de Beer R, Fuderer M, Mehlkopf AF, van Ormondt D. *Non uniform phase-encode distributions for MRI scan time reduction. J Magn Reson* 1996; 111:70–75.
- [15] Mohlenkamp, M. J, *A Tutorial on Wavelets and Their Applications*. University of Colorado, Boulder, Dept. of Applied Mathematics, 2004.
- [16] S.G. Mallat *A Wavelet Tour of Signal Processing* (1999 Academic Press) p. 255
- [17] N. Malmurugan, A. Shanmugam, S. Jayaraman and V. V. Dinesh Chander. "A New and Novel Image Compression Algorithm Using Wavelet Footprints"
- [18] Peters DC, Korosec FR, Grist TM, Block WF, Holden JE, Vigen KK, Mistretta CA. *Undersampled projection reconstruction applied to MR angiography. Magn Reson Med* 2000;43: 91–101.
- [19] Polikar, R, *Multiresolution Analysis: The Discrete Wavelet Transform*. , Rowan University, NJ, Dept. of Electrical and Computer Engineering
- [20] Rudin L, Osher S, Fatemi E. Non-linear total variation noise removal algorithm. *Phys. D*1992; 60:259–268.
- [21] Rudin, Walter (1980), *Real and Complex Analysis* (2rd ed.), New Delhi: Tata McGraw-Hill,
- [22] Rudin, Walter (1987), *Real and complex analysis* (3rd ed.), New York: McGraw-Hill, ISBN 978-0-07-054234-1, MR 924157
- [23] Tsai CM, Nishimura D. *Reduced aliasing artifacts using variable-density k-space sampling trajectories. Magn Reson Med* 2000; 43:452–458.
- [24] Wajer F. "Non-Cartesian MRI Scan Time Reduction through Sparse Sampling". *PhD thesis, Delft University of Technology*, 2001.

AUTHORS BIOGRAPHY

Jyothi Kiran Gowdru Chandrappa is working in GSS Institute of Technology as Assistant Professor. She has completed her ME in UVCE, Bangalore.

