Assignment 1: CS 754, Advanced Image Processing

Shreyas Narahari-203050037
Nitish Gangwar-203050069

February 9, 2021

Answer of Q1

(Part 1.1:)

we know that,

$$|\theta - \theta*| \leq C_1 s^{-1/2} ||\theta - \theta_s|| + C_2 \epsilon$$

Here we are dealing with the inverse problem so it is very intuitive to analyze the relative reconstruction error instead of absolute reconstruction error.So, as s increases the mean of the distribution also increases so does its variance.finally resulting into the increase in the mean square error but a decrease in relative mean squared error and hence we can say that the term $|\theta - \theta_s|_2$ starts decreasing when s increases.

(Part 1.2:)

We know that, ϵ is an upper bound on the size of the noisy contribution.i.e.

$$||\eta||_2 \leq \epsilon$$

and we also know through theorem 3's tube constraint that

$$||\Phi(x^* - x_0)||_2 \leq ||\Phi x^* - y||_2 + ||\Phi x_0 - y||_2 \leq 2\epsilon$$

and finally we ends up getting the final equation as shown below:

$$||h||_2 \leq \frac{1}{\sqrt{s}} \frac{1 + \delta_{2s}(\sqrt{2} - 1)}{1 - (1 + \sqrt{2})\delta_{2s}} ||x_0 - x_{0,T0}||_1 + \frac{4\sqrt{1 + \delta_2 s}}{1 - (1 + \sqrt{2})\delta_2 s} \epsilon$$

here we can see the terms in the RHS indicating the error caused by the recovery algorithm and that due to noise.So, error due to noise is dependent on ϵ. In case of quantization noise,considering the Uniform random noise.we have

$$\epsilon \geq r\sqrt{m}$$

(Part 1.3:)

Theorem 3A will be a better choice because we know from Theorem 3 that:

$$||h||_2 \leq \frac{4\sqrt{1+\delta_2 s}}{1-(1+\sqrt{2})\delta_2 s}\epsilon + \frac{1}{\sqrt{s}}\frac{1+\delta_2 s(\sqrt{2}-1)}{1-(1+\sqrt{2})\delta_2 s}||x_0 - x_{0,T_0}||_1$$

The two constant factors in the RHS of the above given equation are both increasing function of δ_{2s}.So, if the value of δ_{2s} decreases then from the above inequality the RHS will also decreases so we will get more tighter estimation of our solution.So, theorem 3A will be a better choice.

(Part 1.4:)

when ϵ is set to 0 then we know in basis Pursuit problem we have:

$$x^* = argmin_x ||x||_1$$

such that

$$||y - Ax||_2 \leq \epsilon$$

So, if ϵ is set to 0 then we get exactly what we started with

$$y = Ax$$

but since there was some non-zero noise magnitude so exact reconstruction is not possible.

Answer of Q2 :

(Part 2.b:)

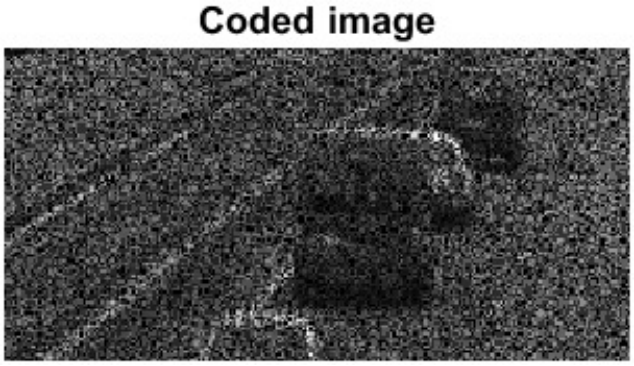


Figure 1: Coded snapshot for $N_f = 3$

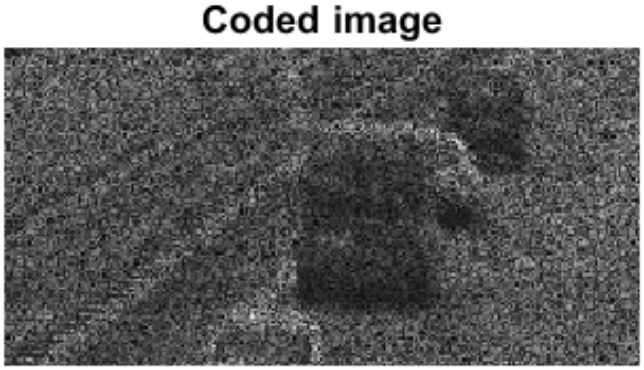


Figure 2: Coded snapshot for $N_f = 5$

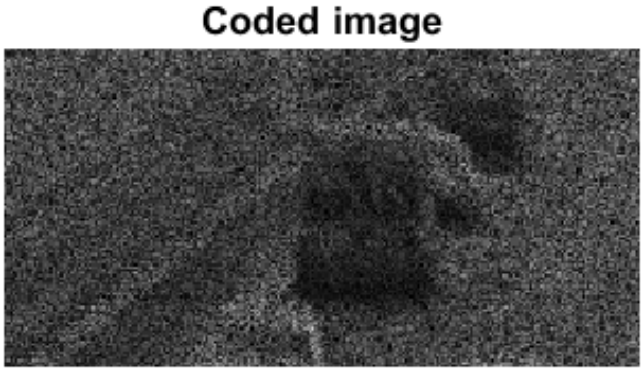


Figure 3: Coded snapshot for $N_f = 7$

(Part 2.c:)

we already have

$$E_u = \sum_{t=1}^{T} C_t.F_t$$

we can consider the above given equation similar to

$$y = \phi f$$

and, we know that

$$f = \psi\theta$$

hence finally we have,

$$y = \phi\psi\theta$$

Here, we have been asked to represent above given equation in the following format

$$Ax = b$$

Here, b is the vectorized format of coded snapshot. A is basically defined as the product of $\phi\psi$.

where ϕ is the diagonal matrix which contains element from the randomly coded mask denoted by C_t.

ψ is 2D DCT basis matrix which is calculated by taking Kronecker product of two 2D-DCT and x is basically θ which needs to be computed.

(Part 2.d:)

we know that OMP is an iterative greedy algorithm that at each step selects the column of A which is most correlated with the current residuals.This column is then added into the set of previously selected columns.The algorithm updates the residuals by projecting the observation y onto the linear subspace spanned by the columns that have already been selected and the algorithms then iterates.
Now for estimating the error,we consider the following model

$$y = Ax$$

and incase of patch by patch reconstuction then
A is basically the product ϕ and ψ, where ϕ is an diagonal matrix having the values from mask matrix and ψ is the 2D DCT basis matrix of size 8x8 ie $\psi_{8x8} \bigotimes \psi_{8x8}$.

from the Johnson-Linderstrauss lemma we have

$$0.75||a||_2^2 \leq ||Za||_2^2 \leq 1.25||a||_2^2$$

for all a $\in$ A.
here A is a collection of points in R^m and let Z be a N $\times$ m matrix whose entries are all i.i.d NORMAL(0,N^{-1}) except with the probability

$$2|A|e^{-cN}.$$

another lemma which shows that there is a subset A of the unit sphere in R^m such that

1. each vector on the sphere lies within l_2 distance 0.125 of some point in A.
2. the cardinality of A does not exceed 24^m.

On applying the Johnson–Lindenstrauss lemma . This procedure shows that Z cannot distort the points in the covering A much. A simple approximation argument shows that Z cannot distort any other point on the sphere. Working out the details leads to the following result.
From Baraniuk et al. paper.we got,

$$0.5||x||_2 \leq ||Zx||_2 \leq 1.5||x||_2$$

for all x $\in R^m$ with probability at least

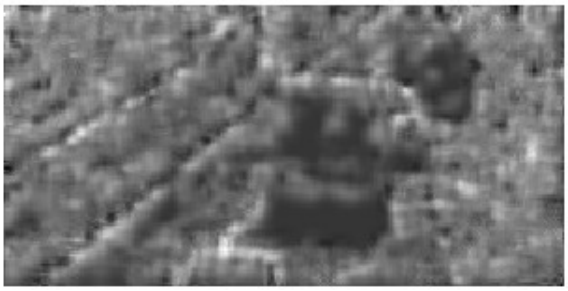
$$1 - 2.24^m.e^{-cN}.$$

For gaussian measurements they have obtained more precise estimates for the constant and was found to be at K=20.But as the number m of nonzero components approaches towards the infinity it is possible to take $K = 4 + \eta$ for any positive number η.and hence the following theorem has been stated as shown below:
Theorem: Fix $\delta \in$(0,0.36) and choose $N \geq Kmln(d/\delta)$.Suppose that s is an arbitrary m-sparse signal in R^d.Draw N measurement vectors $x_1, x_2, ..., x_N$ independently from the standard Guassian distribution on R^d.Given the data OMP can reconstruct the signal with probability exceeding $1-2\delta$.For this theoretical result, it suffices that K=20.when m is large ,it suffices to take K=4.

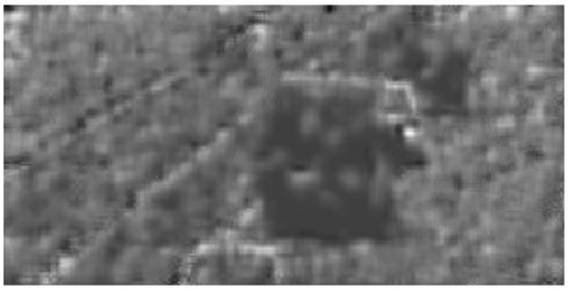
(Part 2.e:)

Reconstruction done when N_f =3 where N_f denotes number of frames

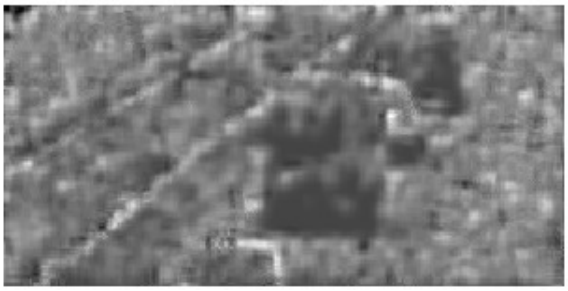
Reconstructed Frame No : 1 RMSE: 3.912173e-01



Reconstructed Frame No : 2 RMSE: 3.959315e-01



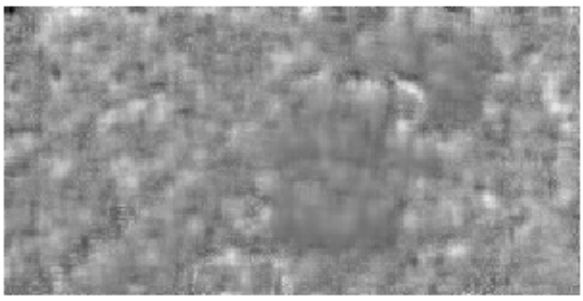
Reconstructed Frame No : 3 RMSE: 3.941431e-01



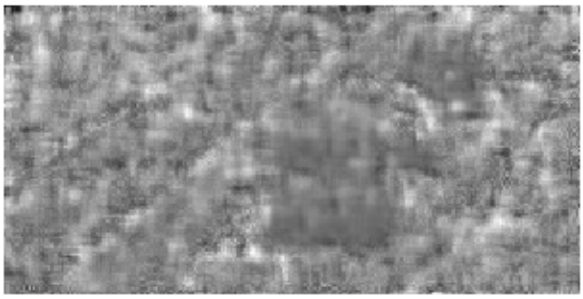
(Part 2.f:)

Reconstruction done when N_f =5 where N_f denotes number of frames:

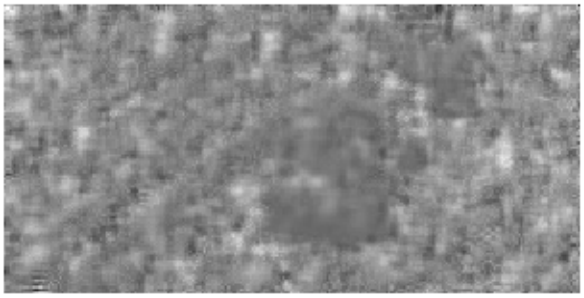
Reconstructed Frame No : 1 RMSE: 4.247001e-01



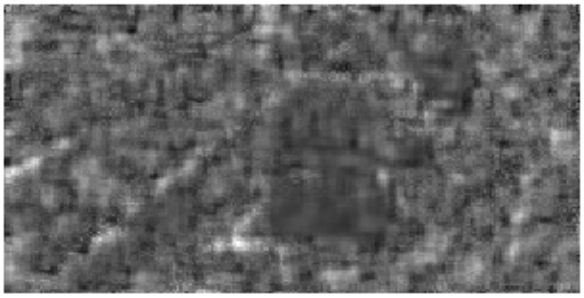
Reconstructed Frame No : 2 RMSE: 4.293451e-01



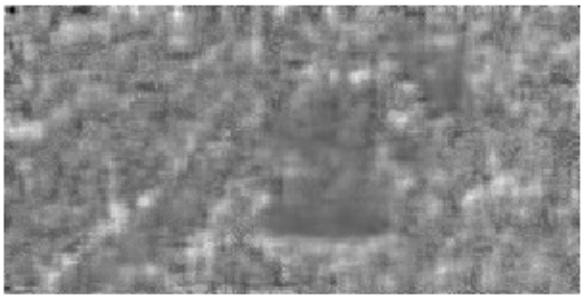
Reconstructed Frame No : 3 RMSE: 4.287452e-01



Reconstructed Frame No : 4 RMSE: 4.272792e-01

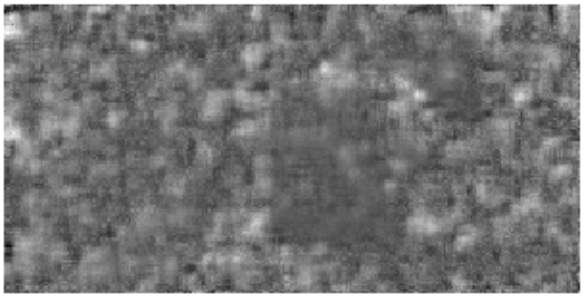


Reconstructed Frame No : 5 RMSE: 4.309228e-01

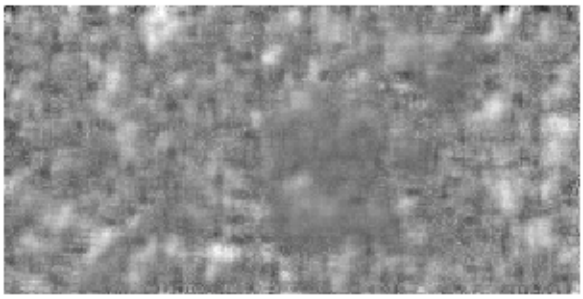


Reconstruction done when N_f =7 where N_f denotes number of frames:

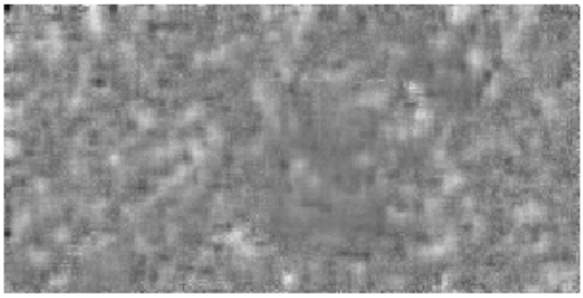
Reconstructed Frame No : 1 RMSE: 4.489795e-01



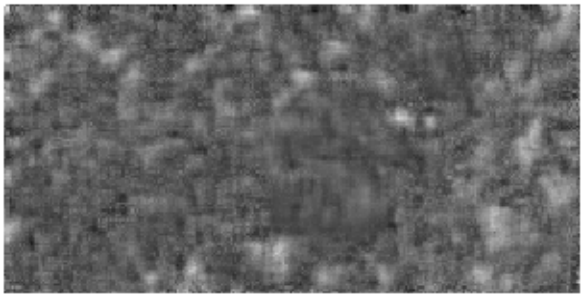
Reconstructed Frame No : 2 RMSE: 4.491971e-01



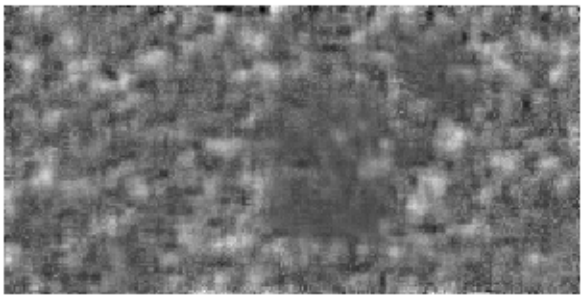
Reconstructed Frame No : 3 RMSE: 4.514804e-01



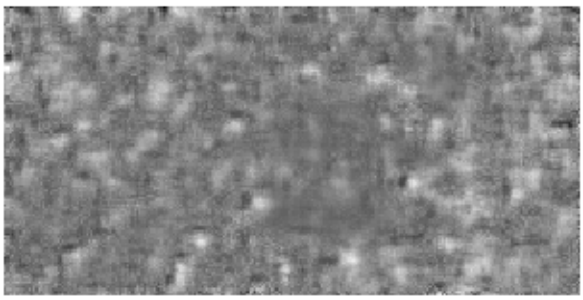
Reconstructed Frame No : 4 RMSE: 4.548403e-01



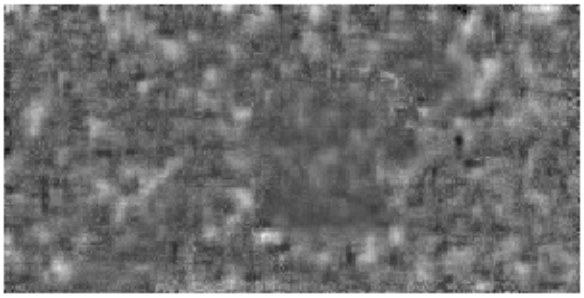
Reconstructed Frame No : 5 RMSE: 4.559046e-01



Reconstructed Frame No : 6 RMSE: 4.600024e-01



Reconstructed Frame No : 7 RMSE: 4.615071e-01



(Part 2.h:)

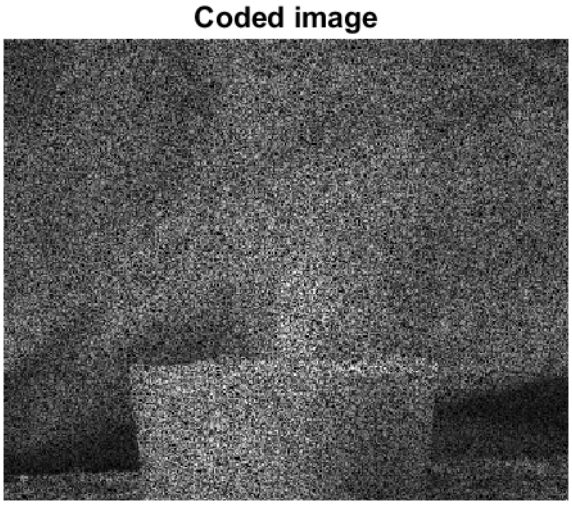
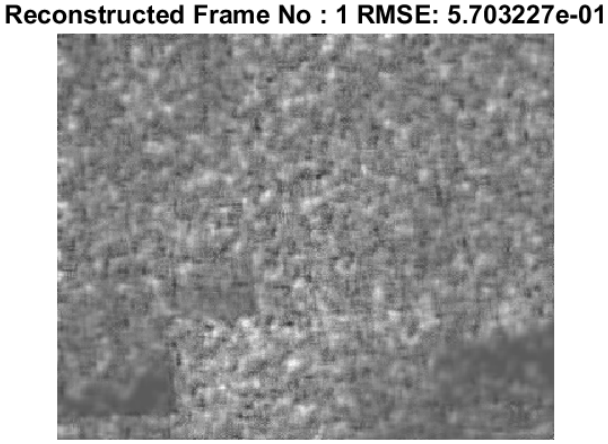
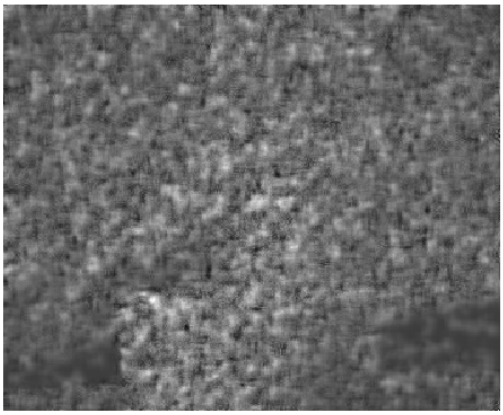


Figure 4: Coded snapshot of flame video for $N_f = 5$ frames

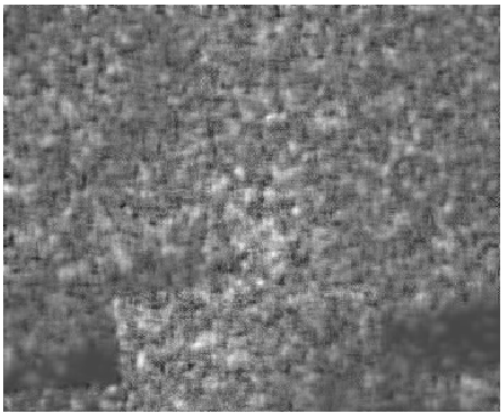
Reconstruction done when N_f =5 where N_f is number of frames:



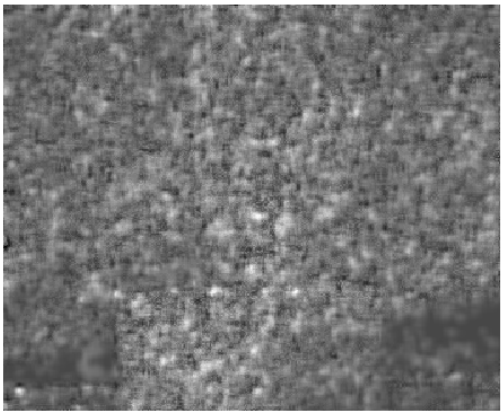
Reconstructed Frame No : 2 RMSE: 5.675376e-01



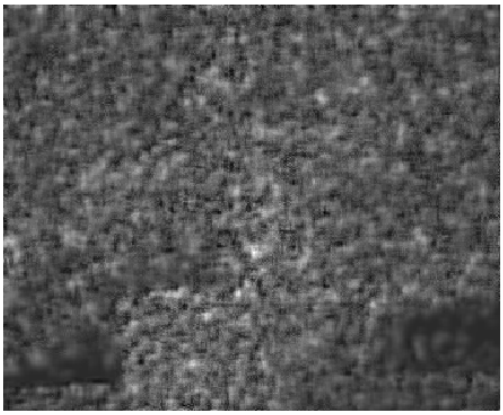
Reconstructed Frame No : 3 RMSE: 5.675781e-01



Reconstructed Frame No : 4 RMSE: 5.698350e-01



Reconstructed Frame No : 5 RMSE: 5.665471e-01



Answer of Q3 :

For proving the upper bound, the coherence is considered as:

$$\mu(\mathbf{\Phi}, \mathbf{\Psi}) = \sqrt{n}\text{max}_{i,j\in\{0,1,\ldots,n-1\}}|\mathbf{\Phi^{i}}^{t}\mathbf{\Psi_j}|$$

First we prove the upper bound for the coherence. From Cauchy–Schwarz inequality we get that:

$$|a||b| \geq |ab|$$

Considering the relationship between the measurement matrix ϕ and representation matrix ψ,

$$|\mathbf{\Phi^{i}}^{t}||\mathbf{\Psi_j}| \geq |\mathbf{\Phi^{i}}^{t}\mathbf{\Psi_j}|$$

As $\mathbf{\Phi^{i}}^{t}$ and $\mathbf{\Psi_j}$ are unit vectors we get,

$$\mathbf{1} \geq |\mathbf{\Phi^{i}}^{t}\mathbf{\Psi_j}|$$

$$max(|\mathbf{\Phi^{i}}^{t}\mathbf{\Psi_j}|) \leq \mathbf{1}$$

From the above statements we can infer that the upper bound of the coherence is,

$$\mu(\mathbf{\Phi}, \mathbf{\Psi}) \leq \sqrt{n}$$

For proving the lower bound we consider a unit vector $\boldsymbol{g} \in \mathbb{R}^n$ which can be expressed as $\boldsymbol{g} = \sum_{k=1}^{n} \alpha_k \mathbf{\Psi_k}$ where $\mathbf{\Psi}$ is an orthonormal basis.
And for orthonormal vectors it is known that,

$$\mathbf{\Psi_i}.\mathbf{\Psi_j} = \begin{cases} 0, & \text{if } i \neq j \\ 1, & \text{if } i = j \end{cases} \tag{1}$$

Using the above property we get the coherence as ,

$$\mu(\boldsymbol{g}, \boldsymbol{\Psi}) = \sqrt{n}\text{max}_{i\in\{0,1,\ldots,n-1\}}\frac{|\alpha_i|}{\sum_{j=1}^{n}\alpha_j^2}$$

As $\boldsymbol{\alpha_i}$ components of $\boldsymbol{g}$ with respect to $\boldsymbol{\psi_i}$ and $|g| = 1$ so that $\sum_{j=1}^{n}\alpha_j^2 = 1$. From this observation the coherence reduces to,

$$\mu(\boldsymbol{g}, \boldsymbol{\Psi}) = \sqrt{n}\text{max}_{i\in\{0,1,\ldots,n-1\}}|\alpha_i|$$

The minimum coherence can be attained when all the values of the coefficient α_i are equal. So as $\sum_{j=1}^{n}\alpha_j^2 = 1$ we can say that,

$$\alpha^2_{i\in\{0,1,\ldots,n-1\}=\frac{1}{n}}$$

$$\alpha_{i\in\{0,1,\ldots,n-1\}=\frac{1}{\sqrt{n}}}$$

Thus the minimum value that can be achieved is,

$$\mu(\boldsymbol{g}, \boldsymbol{\psi}) = \sqrt{n} \times \frac{1}{\sqrt{n}} = 1$$

$$\mu(\boldsymbol{g}, \boldsymbol{\psi}) \geq 1$$

This minimum value is attained when all the α_i values are equal to $\frac{1}{\sqrt{n}}$ so from the relationship between the unit vector $\boldsymbol{g}$ and orthonormal basis $\boldsymbol{\psi}$ we get,

$$\boldsymbol{g} = \sum_{k=1}^{n}\frac{1}{\sqrt{n}}\boldsymbol{\Psi_k}$$

$$\boldsymbol{g} = \frac{1}{\sqrt{n}}\sum_{k=1}^{n}\boldsymbol{\Psi_k}$$

Answer of Q4 :

Let us consider the i^{th} element of $\boldsymbol{x}$ as $\boldsymbol{x_i}$ and i^{th} column of $\boldsymbol{\phi}$ as $\boldsymbol{\phi_i}$.

(Part 4.a)

Here the vector y will be of dimension 1 and $y = \phi_i * x_i$ but here the main issue is that we do not know the value of ϕ_i or $_i$ as we cannot estimate the position i as there can be elements such that $y = \phi_{i1} * x_i = \phi_{i2} * x_i$ and we cannot detect those cases as two separate cases because any two vectors of size 1 are linearly dependent. But, If the position information is provided, then we can take the provided position to be i then we know the values of y and ϕ_i so x_i can be uniquely estimated as $x_i = \dfrac{y}{\phi_i}$.

(Part 4.b)

Here the vector y will have a size of 2 where we can consider the elements to be y_1 and y_2 where $y_1 = \phi_{i1} * x_i$ and $y_2 = \phi_{i2} * x_i$, we consider the elements ϕ_{i1} and ϕ_{i2} to be the elements of the column vector ϕ_i giving us $y = \phi_i * x_i$. In this case the condition that $y = \phi_i * x_i = \phi_j * x_i$ which can also be written as $(\phi_i - \phi_j)x_i = 0$ should not be true. This is possible as we can build the measurement matrix to be such that no two columns are linearly dependent.

(Part 4.c)

Here the vector y will have a size of 3 where y can be expressed as $y = \phi_i * x_i + \phi_j * x_j$ we cannot get a unique estimation as if we consider two other columns of the measurement matrix ϕ_p and ϕ_q we should be able to get a measurement matrix such that all combinations of four columns should not satisfy the condition that $y = \phi_i * x_i + \phi_j * x_j = \phi_p * x_i + \phi_q * x_j$ ie. $\phi_i * x_i + \phi_j * x_j - \phi_p * x_i - \phi_q * x_j = 0$ this should not be equal to the zero vector for uniqueness to exist. But from the given set of conditions it forms an under determined set of equations as the number of variables to estimate is 4 and number of equations is 3. So, we cannot uniquely determine x given these constraints as we can always find some combination of columns such that we get zero vector.

(Part 4.d)

Here the vector y will have a size of 4 which can be expressed as $y = \phi_i * x_i + \phi_j * x_j$ in which case we can get a unique estimation for x. To understand this better we will consider two more columns of the measurement matrix ϕ_p and ϕ_q such that the condition $y = \phi_i * x_i + \phi_j * x_j = \phi_p * x_i + \phi_q * x_j$ ie. $\phi_i * x_i + \phi_j * x_j - \phi_p * x_i - \phi_q * x_j = 0$ should not hold. But here as we have four equations and four variables to estimate we can have combinations of columns (ϕ_i) such that we will have a deterministic solution and can give guarantee that any four columns of the measurement matrix are linearly independent.

Answer of Q5:

Similarities in Architecture:

1. Both architecture uses the coded aperture and codes a single coded snapshot.

2. CASSI and CACTI share identical mathematical forward models.

Differences in architecture

1. In CACTI, modulation of the image data stream by harmonic oscillation of this mask requires no code transmission or operating power.

2. The passive coding scheme used in CACTI holds other advantages, such as compactness and polarization independence, over reflective LCoS-based modulation strategies.

3. CACTI's passive coding scheme facilitates scalability without requiring extra power; one may simply use a larger coded aperture to modulate larger values of N pixels.

Cost function used:

TwIST:

$$f_e = argmin||g - Hf||^2 + \lambda\Omega(f)$$

f_e = N_F frame estimate of the continuous motion f
$g =$ coded image of dimension R$^{\sqrt{N}\times\sqrt{N}}$
H = system's discrete forward matrix that accounts for sampling factors of dimension R $^{N\times NN_F}$ includes the optical impulse response,pixel sampling function, and time-varying transmission function.
$f =$ denotes the three dimensional scene and has a dimension of $R^{\sqrt{N}\times\sqrt{N}\times N_F}$
λ= regularization weights
Below given equation denote the Total variation regularizer:

$$\Omega(f) = \sum_{k}^{N_F} \sum_{i,j}^{N_F} \sqrt{(f_{i+1,j,k} - f_{i,j,k}))^2 + (f_{i,j+1,k} - f_{i,j,k})^2}$$

$\Omega(f)$= regularizer used to penalize the characteristics of of the estimated f.

Generalized Alternating Projection(GAP)

$$[P_\pi(\tilde{f})]_{i,j,k} = \tilde{f}_{i,j,k} + \frac{T_{i,j,k}}{\sum_{k'=1}^{N_f} T_{i,j,k'}^2}(g_{i,j} - \sum_{k'=1}^{N_f} T_{i,j,k'}\tilde{f}_{i,j,k'})$$

$[P_\pi(\tilde{f})]_{i,j,k}$= Euclidean projection of any $\tilde{f} \notin \pi$ onto π
T is a coded aperture's transmission function generates random values of size (m,n,p). where , m*n is random binary matrix shifted vertically by p pixels.

$$P_{\Lambda(C)}(f) = \psi^{-1}(argmin_{\theta:||\theta||_{G\beta} \leq C}||\theta - \psi(f)||_2)$$

where this C is defined below and $P_{\Lambda(C)}(f)$ is Euclidean projection of any f that does not belongs to Λ onto Λ GAP in terms of weighted $l_{2,1}$ that undergoes change in the size:
Projection on the linear manifold:

$$f^{(t)} = P_\pi(\tilde{f}^{(t-1)}), t \geq 1$$

Projection on the weighted $l_{2,1}$ ball of changing size.

$$\tilde{f}^{(t)} = P_{\Lambda(C^{(t)})}(f^{(t)}) = \psi^{-1}(\theta^{(t)}), t \geq 1,$$

The GAP algorithm starts with $\tilde{f}^{(0)} = 0$, iterates between the following two steps until $||f^t - \tilde{f}^t||_2$ converges in t. Where,

$$\theta_{i,j,k}^{(t)} = w_{i,j,k}^{(t)} max\left\{1 - \frac{\beta_l ||w_{G_{lm^*+1}}^{(t)}||_2}{\beta_{l_{m^*+1}}||w_{G_l}^{(t)}||_2}, 0\right\}, \forall (i,j,k) \in G_l, l = 1, ..., m,$$

where $w_{i,j,k}$ denotes the group of transform coefficients.These $w_{i,j,k}$ were grouped

together into m disjoint subsets and each such is assigned a weight that is denoted by β_l .

ball size C^t is given by:

$$C^{(t)} = \sum_{q=1}^{m^*} \beta_{l_q}^2 \left(\frac{||w_{G_{l_q}}^{(t)}||_2}{\beta_{l_q}} - \frac{||w_{G_{l_{m^*+1}}}^{(t)}||_2}{\beta_{l_{m^*+1}}} \right)$$

Answer of Q6:

Title: Sparse imaging for fast electron microscopy
Published in: SPIE Electronic Imaging, 2013, Burlingame, California, United States
Year: 2013
Link to paper: click here to access paper

Hardware Architecture

Scanning Electron Microscope(SEM) is used along with electronics to drive the beam location and sample the detector.A Zeiss GmbH column was used with a Schottky thermal field emission source and Gemini optics.An aperture of 10μ m was used for a nominal beam energy of 10 keV.The incident beam was deflected onto the sample using the standard scanning coils and current amplifiers in the column.However, some custom electronics was used to set the desired beam location using an standard column controls.Standard column controls were also used for magnification.Then, the pixels in the field of view can be seen by applying a voltage of -10V to +10V, which is converted to a current in the coil amplifiers.A digital to analog converter(DAC) was used to drive the desired voltages.The detector was sampled using an analog to digital converter(A/D) that was synchronized to the DAC.National Instruments PCI-6110 multi-function data acquisition system was used to implement the A/D and DAC.The system has a maximum frequency of 2.5MHz with two analog outputs,an output resolution of 16 bits per sample and an input resolution of 12 bits per sample.
They achieved variable dwell time by digitally averaging multiple samples at the same time location.A basic dwell time of 400ns using one sample per pixel results in low-SNR images, while a high-SNR dwell time of 6.4μs achieved by averaging 16 samples per pixel.For characterizing the dynamics of the amplifiers and scan coils ,they recommended a step wise jump in position from one extreme of the beam's scanning range to the other over a calibration sample and this is done using scan coil dynamics.

Reconstruction or Cost function

Authors have used Split Bregman Interpolation for fast recovery of sparsely sampled images.Measurements taken were denoted as :

$$y = \phi x + \eta$$

$\phi =$ row subset of identity matrix.
$\eta =$ noise with power σ^2.
Then, basis pursuit was used for image reconstruction.i.e.

$$min_x ||\psi^T x||_1 + ||\nabla x||_1$$
$$s.t. ||y - \phi x|| \leq \sigma^2$$

$\psi =$block-DCT basis with 32 $\times$ 32 pixel blocks.
Total variation regularizer is:

$$||\nabla x||_1 = \sum_i \sqrt{|(\nabla_h x)_i|^2 + |(\nabla_v x)_i|^2}$$

Basis Pursuit can be solved efficiently using the split Bregman method which recasts the constrained problem into an unconstrained problem of the form.

$$min_x ||\psi^T x||_1 + ||\nabla x||_1 + \frac{\mu}{2}||y - \phi x||.$$

Then, split Bregman formulation was used to solve the problem upto a arbitrary precision.At the kth iteration,following equation is solved

$$min_{x,u,v,w} ||w||_1 + ||(u,v)||_2 + \frac{\mu}{2}||\phi x - y||_2^2 + \frac{\lambda}{2}||u - \nabla_u x - b_u^k||_2^2 + \frac{\lambda}{2}||v - \nabla_v x - b_v^k||_2^2 + \frac{\gamma}{2}||w - \psi^T x - b_w^k||_2^2$$

and then the Bregman parameters b_u^k, b_v^k, b_w^k were updated

$$b_u^{k+1} = b_u^k + (\nabla_u x^{k+1} - u^{k+1})$$
$$b_v^{k+1} = b_v^k + (\nabla_v x^{k+1} - v^{k+1})$$
$$b_w^{k+1} = b_w^k + (\psi^T x^{k+1} - w^{k+1})$$

The elements which had l_1 norms were solved efficiently via element-wise shrinkage

$$u_i^{k+1} = \frac{max(s_i^k - \frac{1}{\lambda}, 0)}{s_i^k}((\nabla_u x^k)_i + b_{u,i}^k)$$

$$v_i^{k+1} = \frac{max(s_i^k - \frac{1}{\lambda}, 0)}{s_i^k}((\nabla_v x^k)_i + b_{v,i}^k)$$

$$w_i^{k+1} = shrink((\psi^T x^{k+1})_i + b_{w,i}^k, \frac{1}{\gamma})$$

where,

$$s_i^k = \sqrt{|(\nabla_u x^k)_i + u_i^k|^2 + |(\nabla_v x^k)_i + v_i^k|^2}$$

and

$$shrink(x, \rho) = sgn(x)max(|x| - \rho, 0)$$

and then the kth equation was solved for x

$$(\mu\phi^T\phi - \lambda\delta + \gamma I)x^{k+1} = \mu\phi^T y + \lambda\nabla_u^T(u^k - b_u) + \lambda\nabla_v^T(v^k - b_v) + \gamma\psi(w^k - b_w)$$

For typical images on the interval [0,1],authors have used $\mu = \lambda = 1$ and $\gamma = 10^{-2}$ were resonable values for reconstruction.

References

1. https://ieeexplore.ieee.org/document/5895106/metrics#metrics.
2. https://sci-hub.se/10.1109/tit.2007.909108
3. https://www.spiedigitallibrary.org/conference-proceedings-of-spie/8657/1/Sparse-imaging-for-fast-electron-microscopy/10.1117/12.2008313.short?SSO=1
4. http://www.math.lsa.umich.edu/ annacg/papers/TG05-signal-recovery-TR.pdf
5. https://en.wikipedia.org/wiki/Cauchy%E2%80%93Schwarz_inequality
6. https://math.stackexchange.com/questions/211370/linear-combination-of-vectors
7. https://statistics.stanford.edu/sites/g/files/sbiybj6031/f/1999-13.pdf