# Assignment 1
# Advanced Image Processing (CS754)

[1] Shashwat Pathak (213070010)
[2] Mohit Kumar Meena (213070021)

## Contents

# 1 Question 1

**(1)**. From the definition, we know that $\delta_{2S}$ is the increasing function of S (Sparsity of signal), Also the constants C0 and C1 increases with increasing $\delta_{2S}$ causing the overall error term in the bound to increase. The term $1/\sqrt{S}$ in the upper bound decreases but it does not cause a significant reduction. Hence the condition of discrepancy does not arise practically.

**(2)**. Intuitively, it looks like error bound is independent of m (number of measurements) but in real scenario this is not the case. Theorem 3 states that **A** (having dimension m*n) should obey RIP of order 2S, where $\delta_{2S} < 0.41$. And we know that error bound is dependent on $\delta_{2S} < 0.41$. Additionally, In theorem 1 it's been stated that we can get exact reconstruction of **f** with very high probability given $m \geq O(Clog(n/\delta))$. So indirectly the error bound depends on **m**.

**(3)**. Theorem 3 is better than Theorem 3A. This is because of the fact that that larger the bound (smaller $\delta_{2S}$) as shown in equation below, requires less sparser signal than the case of smaller bound (smaller $\delta_{2S}$).

$$(1-\delta_{2s})||\theta||^2 \leq ||\mathbf{A}\theta||^2 \leq \ (1-\delta_{2s})||\theta||^2$$

Also, every matrix **A** satisfying condition for theorem 3A ($\delta_{2S} < 0.1$) will also satisfy for theorem 3 ($\delta_{2S} < 0.41$).

**(4)**. We know that $\epsilon$ is the upper bound on noise vector ($\eta$) and it's been given that $\eta$ is non-zero, so considering $\epsilon = 0$ will be highly inappropriate. Practically, signal is corrupted i.e. $y = \phi\psi\theta + \eta$ so if we consider $\epsilon = 0$ i.e. $y = \phi\psi\theta$ will introduce high errors and may lead to wrong solution.

# 2 Question 2

**To prove:**
The coherence between $\phi$ and $\psi$ as given in equation 1 lies in the range $[1,\sqrt{n}]$.

$$\mu(\phi,\psi) = \sqrt{n} \max_{i\in\{0,1.....m-1\},j\in\{0,1......n-1\}} |\phi^{i^t}\psi| \tag{1}$$

**Proof:**
**a). Upper bound**
According to Cauchy-Schwartz inequality, the dot product of two vectors is always less than or equal to their individual products. i.e.

$$|\phi^i\psi j| \leq |\phi^i||\psi^j|$$

Given that $\phi$ and $\psi$ are unit normalized, so $|\phi^i| = |\psi^j = 1|$. Now using equation (1), we get,

$$\max_{i\in\{0,1.....m-1\},j\in\{0,1......n-1\}} |\phi^{i^t}\psi| \leq \sqrt{n} \tag{2}$$

Therefore maximum value of coherence ($\mu(\phi,\psi)$) is $\sqrt{n}$.

**b). Lower bound**
Utilizing the hint i.e.

$$g = \sum_{k=1}^{n} \alpha_k \psi k$$

Given that **g** is an orthonormal unit vector i.e. $g^T g = 1$. So,

$$g^T g = \left(\sum_{k=1}^{n} \alpha_k^T \psi_k^T\right)\left(\sum_{k=1}^{n} \alpha_k \psi_k\right) = 1$$

Now, exploiting the fact that $\psi$ is orthonormal and unit normalized along columns ($\psi^T$ is unit normalized along rows). we get,

$$g^T g = \sum_{k=1}^{n} \alpha_k^2 = 1 \tag{3}$$

Now we need to prove,

$$\mu(g, \psi) = \sqrt{n} \max_{i \in \{0,1,......n-1\}} \frac{|\alpha_i|}{\sum_{j=1}^{n} \alpha_j^2} \tag{4}$$

Substituting the results from equation 3 in equation 4 we get,

$$\mu(g, \psi) = \sqrt{n} \max_{i \in \{0,1,......n-1\}} \frac{|\alpha_i|}{1} = \sqrt{n} \max_{i \in 0,1,......n-1} |\alpha_i| \tag{5}$$

To solve further let's use proof by contradiction. i.e. let us assume that

$$\max_{i \in \{0,1,......n-1\}} |\alpha_i| < \frac{1}{\sqrt{n}}$$

this implies,

$$|\alpha_k|^2 < \frac{1}{n} \quad \forall i \in \{0, 1, .....n-1\} \tag{6}$$

$$\sum_{k=1}^{n} \alpha_k^2 < 1$$

This clearly contradicts the results obtained in the equation 3. So,

$$\max_{i \in 0,1,......n-1} |\alpha_i| \geq \frac{1}{\sqrt{n}} \tag{7}$$

Now, substituting the minimum value of above expression i.e. $\frac{1}{\sqrt{n}}$ in equation 5, we'll get the minimal value of coherence $\mu(g, u)$ to be 1.
Using equation 2 and equation 7, we concluded that the minimal value of coherence is attained when one of the **g** is $\frac{1}{\sqrt{n}}{}_{k=0}^{n-1}\psi_k$.

# 3 Question 3

**a).** We know that

$$y = \sum_{j=1}^{n} \phi_j x_j$$

Here, $\phi_j$ is the $j^{th}$ column of $\phi$ and $x_j$ is the element in the $j^{th}$ row. When **m**(number of measurements)=1, then there doesn't exist any unique solution. This can be explained as follows:
Using above equation,

$$x_j = \frac{y}{\phi_j}$$

So, when y=0 and $\phi_j \neq 0$ then $x_j = 0$ but this contradicts the fact that **x** has one non-zero element, so $y \neq 0$. Now, if $y \neq 0$ but $\phi_j = 0$, then there exists infinitely many solutions of **x**. Now, considering the last case i.e. When $y \neq 0$ and $\phi_j \neq 0$ then even in this case, there exist different value of **x** for different **j** value. Therefore, we cannot uniquely determine the value of x.
When index(j) of non-zero element in x is given then we can easily determine the unique value of x using $x_j = \frac{y}{\phi_j}$.

**b).** If m (number of measurements)=2, even in this case that **x** to have unique solution but there exist some cases when we can find unique **x**. Consider,

$$y_1 = \sum_{j=1}^{n} \phi_{1j} x_j$$

$$y_2 = \sum_{j=1}^{n} \phi_{2j} x_j$$

So dividing the above two equations, we get ratio $\frac{\phi_{1j}}{\phi_{2j}}$. Now depending upon this ratio we can identify which column of $\phi$ satisfy this condition. But as said, this doesn't guarantee the unique value of **x**, Since more than one column value can satisfy this ratio.

**c).** When m=3 and x is 2 sparse so to determine the unique values of x, any four 3D vectors have to be linearly independent but this is not the case since m=3. Consider,

$$y = \sum_{j=1}^{n} \phi_j x_j$$

Specifically,

$$y_1 = \phi_{11} x_1 + \phi_{12} x_2 \tag{8}$$

$$y_2 = \phi_{21} x_1 + \phi_{22} x_2 \tag{9}$$

This will result in either unique solution, infinite solutions or no solution. Accordingly, we can design an algorithm to find unique solution (if any).
**Algorithm:**
**- Unique solution:** If the solution of above equations is unique, then check whether it satisfies the equation

$$y_3 = \phi_{31} x_1 + \phi_{32} x_2 \tag{10}$$

If **Yes**, then store it and halt the execution, if **No** then proceed further.

**- No solution:** Simply proceed further.

**- Infinite solution:** Then again solve for either equation 1&3 or equation 2&3. If unique solution then check if it satisfies the third equation then it's the unique solution and store it, if it's no solution then simply iterate over and if the solution is infinite then simply **halt** the execution.

**d).** When **m**=4 and **x** is 2 sparse then we can uniquely determine the values of x. Since, every 4 columns of matrix $\phi$ are linearly independent and it's been proved in signal reconstruction uniqueness issues that " If any **2S** columns of any $m * n$ matrix $\phi$ are linearly independent then any S-sparse signal **x** can be uniquely reconstructed from measurements $y = \phi x$. Again the algorithm to find the unique values of x is illustrated as:

**Algorithm:**
-Pick any two measurements (out of 4). e.g.

$$y_1 = \phi_{11}x_1 + \phi_{12}x_2$$

$$y_2 = \phi_{21}x_1 + \phi_{22}x_2$$

**- Unique solution:** If the solution of above equations is unique, then check whether it satisfies the other two measurements.i.e.

$$y_3 = \phi_{31}x_1 + \phi_{32}x_2$$

$$y_4 = \phi_{41}x_1 + \phi_{42}x_2$$

If **Yes**, then store it and halt the execution, if **No** then proceed further.

**- No solution:** Simply proceed further.

**- Infinite solution:** Then again solve for other measurements. If they results in unique solution then check if it satisfies the remaining measurements if it is so then it's the unique solution and store it, if it's no solution then simply iterate over and if the solution is infinite then simply **halt** the execution.

# 4 Question 4

Given that,
P1: $\min(||x||_1)$ with respect to $x$ such that $||y - Ax||_2 \leq e$
Q1: $\min(||Ax - y||_2)$ with respect to $x$ such that $||x||_1 \leq t$

**To prove:**
If there exists a unique solution $x$ for P1 for some $e$, then $x$ uniquely satisfies Q1 as well.

**Proof:**

Suppose we have a vector $x = \alpha$ that satisfies P1, such that $min(||x||_1) = \beta$, thus $||x||_1 = \beta$. Also, $||y - A\alpha|| \leq e$.

Now if $t$ is such that $\beta \leq t \leq 0$ then $||x||_1 \leq t$ since $||x||_1 = \beta$. This satisfies the constraint on Q1.

Now, suppose for Q1, $\min(||A\alpha - y||_2) = \delta$, and we select an $e$ such that $\delta \leq e$. Since from constraint on P1, we know that $||y - A\alpha||_2 \leq e$. Thus, Q1 and P1 can both be uniquely satisfied by $x = \alpha$, if we select $t$ and $e$ such that $\beta \leq t \leq 0$ and $\delta \leq e$ are satisfied.

# 5 Question 5

**a) Title:** A Compressive Sensing Method for Noise reduction of Speech and Audio Signals
**Published in**: 2011 IEEE 54th International Midwest Symposium on Circuits and Systems
**Date of Publication**: 23rd September 2011
**Link**: https://ieeexplore.ieee.org/document/6026662

**b)** This paper is an algorithmic implementation. There is no hardware component involved in it. The paper proceeds on the argument that speech and audio signals are k-spare in some domain while the inherent noise factor in the signal is not. Thus, the two are theoretically separable by the compressive sensing method.

**c)** The speech noise reduction is formulated as a L1 norm minimization of the sparse representation of the estimated clean speech signal with respect to some constraints.

To understand this more clearly, let us assume that $\hat{x}$ is a noisy speech signal such that

$$\hat{x} = x + e$$

where $x$ is the clean audio signal and $e$ is the noise in the signal. The clean signal is obtained as a result of solving the following optimization problem:

$$Problem\ 1 : \min ||W.x||_1\ , such\ that\ ||F_u.x - y||_2 \leq ||F_u.e||_2 \leq ||F.e||_2 = ||E||_2$$

here, $W$ is the wavelet transform operator, $F_u$ is the random Fourier operator, $F$ is the Fourier transform matrix and $E$ is is the set of fourier coefficients for the noise signal. Now we modify the above constraint using the Parseval's theorem which states the energy equivalence of a signal in time and frequency domain. Thus, the above constraint equation changes to the following:

$$||F_u.x - y||_2 \leq ||F_u.e||_2 \leq ||F.e||_2 = \frac{2\pi}{\gamma}||x||_2$$

where, $\gamma$ is the signal-to-noise ratio, and is defined as $\gamma = \frac{\sum_t x^2(t)}{\sum_t e^2(t)}$.

In order to solve $Problem 1$, it is converted into a non-constrained quadratic optimization problem. The estimation of clean signal $x^*$ can be done by solving the following problem:

$$Problem\ 2 : x^* = \arg\min_x \lambda[\,||F_u.x - y||_2 + \frac{2\pi}{\gamma}.||x||_2]\ + (1-\lambda)||W.x||_1$$

A quadratic optimization problem can be solved by several standard methods, such as Gradient Linear Search (GDLS), and an Iteratively Re-weighted Least Squared algorithm (IRLS). The authors went with GDLS to solve this problem.

# 6 Question 6

**a)** The cars.avi video has been read using the 'mmread' function (implemented in q6.m). Since the MMRead module was externally provided, we had to add the folder to execution path in order to use MMRead.

**b)** The coded snapshot was generated by multiplying each frame with the binary coded pattern with the grayscale image and then summing up the products for each frame. The pattern was generated using the 'randi' function in MATLAB. Figure 6 shows the coded snapshot for $T = 3$ frames for 'cars.avi'

**c)** For the case of T= 3, the equation to obtain the coded snapshot $E_u$ is as follows:

$$E_u = \sum_{t=1}^{3} C_t.F_t$$

where, $C_t$ is the binary coded pattern and $F_t$ is the image at instant $t$ Now, in order to represent this in the form $\boldsymbol{Ax} = \boldsymbol{b}$, we will decompose the summation as follows:

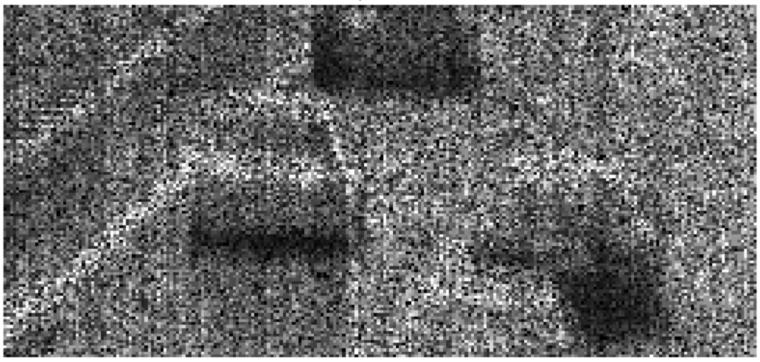


Figure 1: Figure showing the coded snapshot for T=3 for **cars.avi**

$$\boldsymbol{E} = ([\ diag(\boldsymbol{C1}) \mid diag(\boldsymbol{C2}) \mid diag(\boldsymbol{C3})\ ]) * \boldsymbol{F}$$

Assume the image $F$ is sparse in 2D-DCT space, thus $\boldsymbol{F} = \boldsymbol{\psi\theta}$ , here $\boldsymbol{\psi}$ is the diagonalized 2D-DCT matrix and $\boldsymbol{\theta}$ is the vectorized sparse representation of F in the 2D-DCT domain.

Now, $\boldsymbol{\psi} = \begin{bmatrix} \psi_1 & & \\ & \psi_2 & \\ & & \psi_3 \end{bmatrix}$ here, $\psi_i$ is the 2D-DCT matrix for the $i^{th}$ frame. For an $8 \times 8$ patch with three frames each, the $\psi_i$ has dimensions $64 \times 64$. Thus, $\boldsymbol{\psi}$ has the dimensions $192 \times 192$ for three frames. Thus, the equation of E changes to

$$\boldsymbol{E} = ([\ diag(\boldsymbol{C1}) \mid diag(\boldsymbol{C2}) \mid diag(\boldsymbol{C3})\ ]) * \boldsymbol{\psi} * \boldsymbol{\theta}$$

Lets, say $\boldsymbol{\phi} = [\ diag(\boldsymbol{C1}) \mid diag(\boldsymbol{C2}) \mid diag(\boldsymbol{C3})\ ]$, which is the essentially the sensing matrix for the compressive sensing method, we will get, $E = \boldsymbol{\phi} * \boldsymbol{\psi} * \boldsymbol{\theta}$.

On comparing this with $\boldsymbol{Ax} = \boldsymbol{b}$, we realize that the coded snapshot ($\boldsymbol{E}$) corresponds to $\boldsymbol{b}$, $\boldsymbol{x}$ corresponds to $\boldsymbol{\theta}$, and $\boldsymbol{A}$ corresponds to $\boldsymbol{\phi} * \boldsymbol{\psi}$.

**d)** We know from theorem three for compresssive sensing in noisy signals, that in case of gaussian noise, the noise term is a chi-squared random variable (refer slide 75 in $CS_theory.pdf$), hence the error term $\epsilon$ is likely to be within three standard deviations from the mean. Thus we set the error term to be higher to $9m\sigma^2$. for the case of T=3, patchwise implementation of OMP with a patch size of $8 \times 8$ we get the value of $m$ as 64. The value of $\sigma$ is given as 2. Thus, the error term ($\epsilon$) which is the threshold for convergence of OMP, is equal to 2304.

**e)** Average Relative Mean Squared Error between the reconstructed and the original image for T=3 in cars.avi is **28.233**

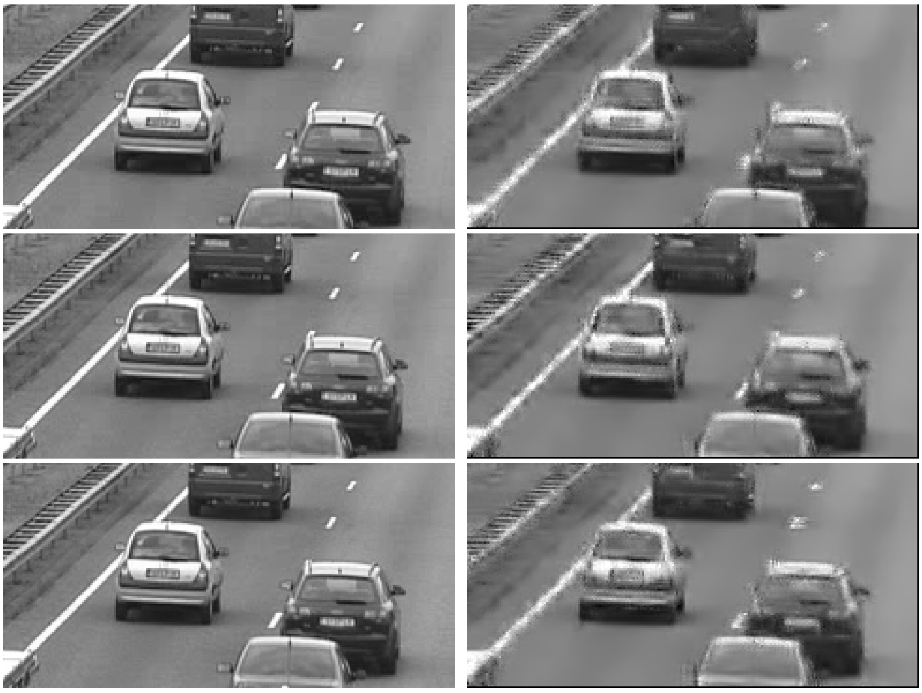

Figure 2: Figure showing the comparison between the original and reconstructed frames for T=3 in **cars.avi**

**f)** Average Relative mean error between the original and the reconstructed image for T=5 over all the frames is is **37.25**.

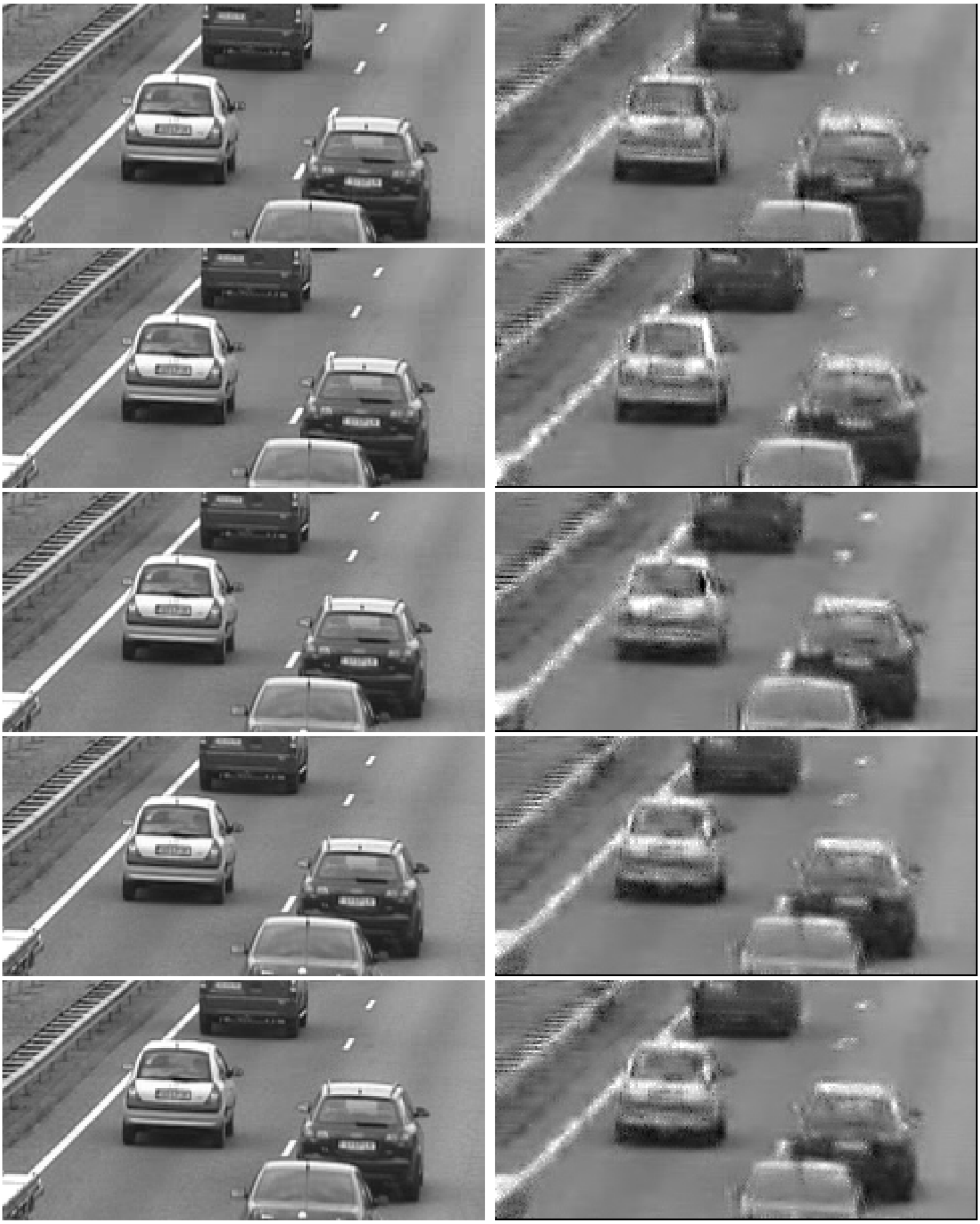

Figure 3: Figure showing the comparison between the original and reconstructed frames for T=5 in **cars.avi**

Average Relative mean error between the original and the reconstructed image for T=7 over all the frames is is **44.44**.

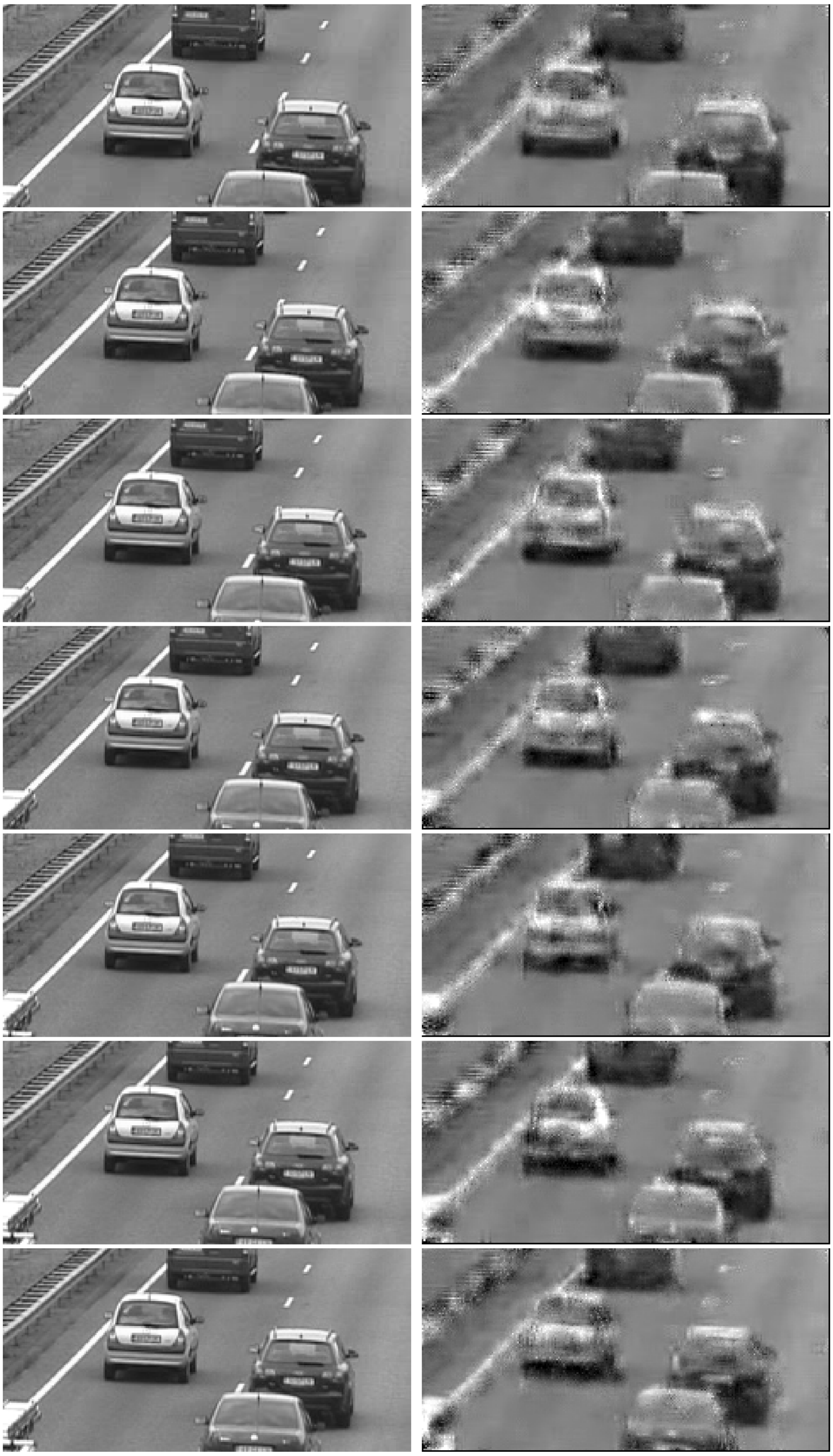

Figure 4: Figure showing the comparison between the original and reconstructed frames for T=7 in **cars.avi**

**g)** All the sections of this question has been attempted by considering a $120 \times 240$ section of the image in the lower right half region. This procedure has been followed for both 'cars.avi' and 'flame.avi'

**h)** Average Relative mean error between the original and the reconstructed image for T=5 over all the frames is is **10.45**.

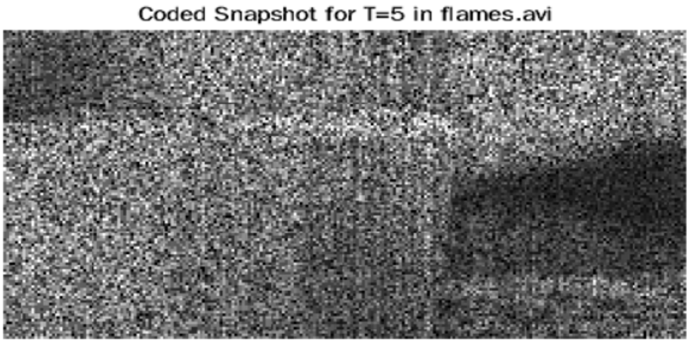


Figure 5: figure showing coded snapshot for T=5 in **flame.avi**

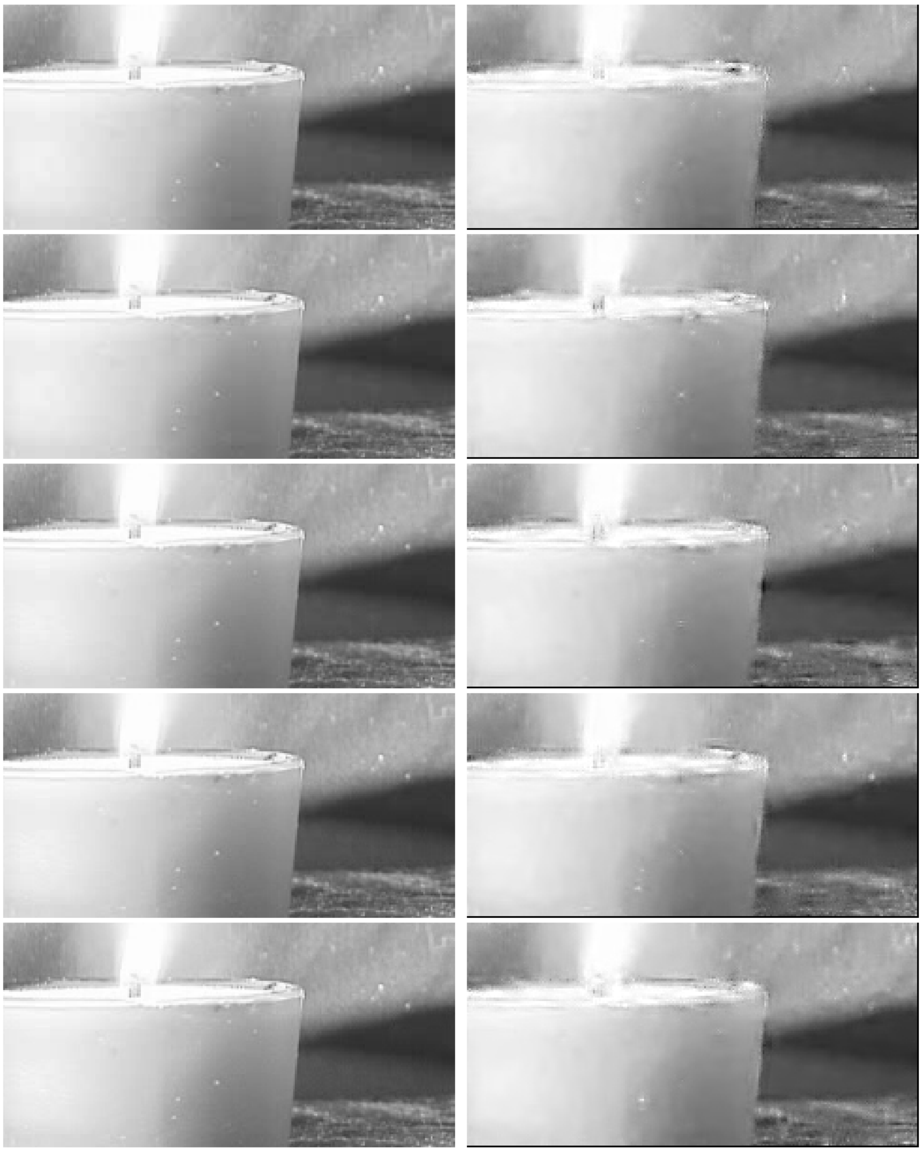

Figure 6: Figure showing the comparison between the original(left) and reconstructed frames(right) for T=5 in **flame.avi**