<u>Digital Video: Perception and Algorithms</u>

<u>Assignment 1</u>

By: - Pratyush Gauri.   Sr.No:- 20227

Lucas Kanade Optical Flow Algorithm

Lucas–Kanade method is a widely used differential method for optical flow estimation. It assumes that the flow is essentially constant in a local neighbourhood of the pixel under consideration and solves the basic optical flow equations for all the pixels in that neighbourhood, by the least squares criterion.
Thus, the optical flow equation can be assumed to hold for all pixels within a window centred at a particular point.
When we use closed form solution for linear case, we also introduce a constraint that the smallest eigenvalue of ATA is more than a particular value. So, because of this constraint, we restrict the noise. Namely, the local image flow (velocity) vector (Vx, Vy) must satisfy,

$$\mathrm{Ix(q1)Vx} + \mathrm{Iy(q1)Vy} = -\mathrm{It(q1)}$$
$$\mathrm{Ix(q2)Vx} + \mathrm{Iy(q1)Vy} = -\mathrm{It(q2)}$$
.
.
.
.
.
$$\mathrm{Ix(qn)Vx} + \mathrm{Iy(qn)Vy} = -\mathrm{It(qn)}$$

Where q1, q2, ..., qn are the pixels in the window and $I_x(qi)$, $I_y(qi)$, $I_t(qi)$ are the partial derivatives of the image I with respect to position x, y and time t, evaluated at the point q_i and at the current time.
These equations can be written in matrix form Av=b.
This system has more equations than unknowns and thus it is usually over-determined. The Lucas–Kanade method obtains a compromise solution by the least squares principle.

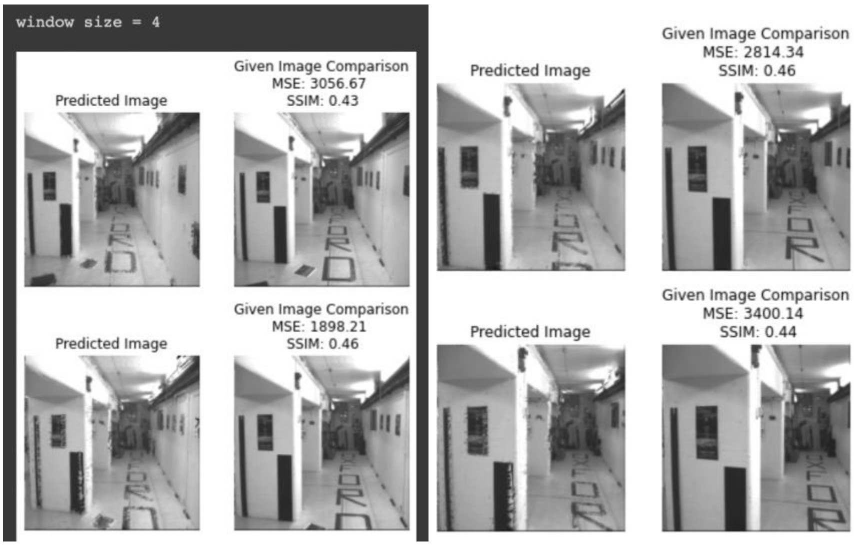
$$A^T A v = A^T b$$
$$v = (A^T A)^{-1} A^T b$$

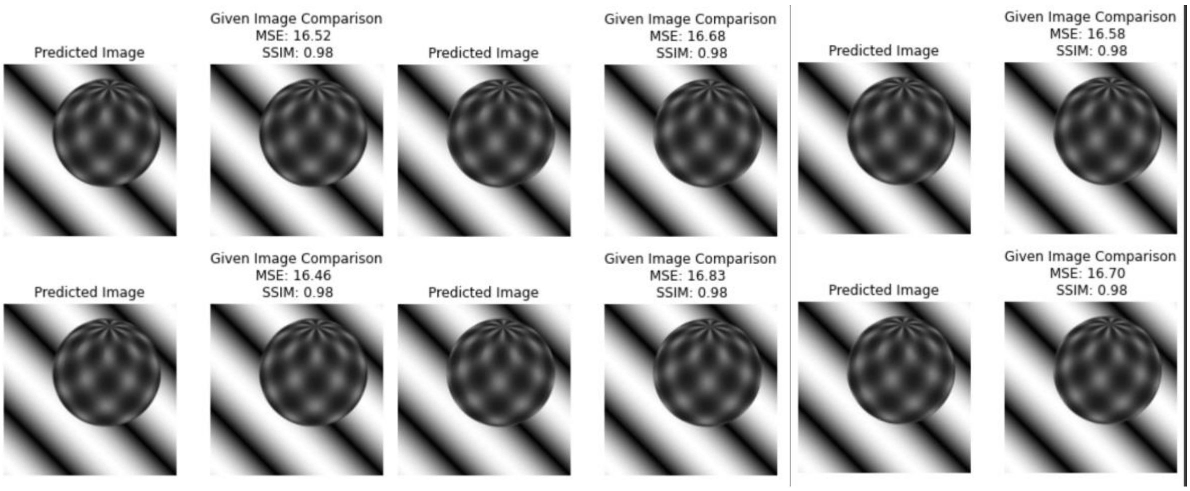
We make 2 Assumptions while deriving the optical constraint equation:
1. Brightness of image point remains constant over time.
2. Displacement and time step are small.

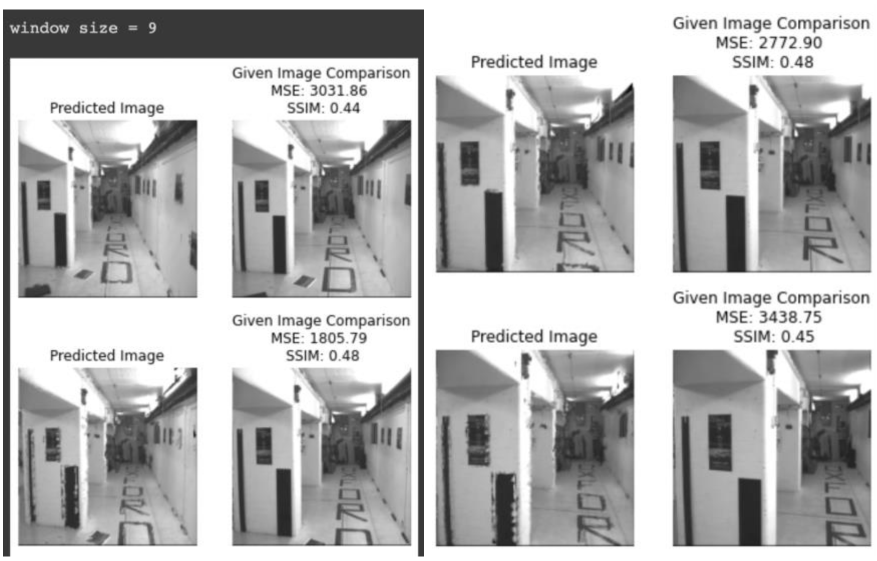
When does optical flow equation works?
1. $A_T A$ must be invertible, that is its determinant is not equal to 0.
2. $A_T A$ must be well -conditioned (change in input cause change in output)

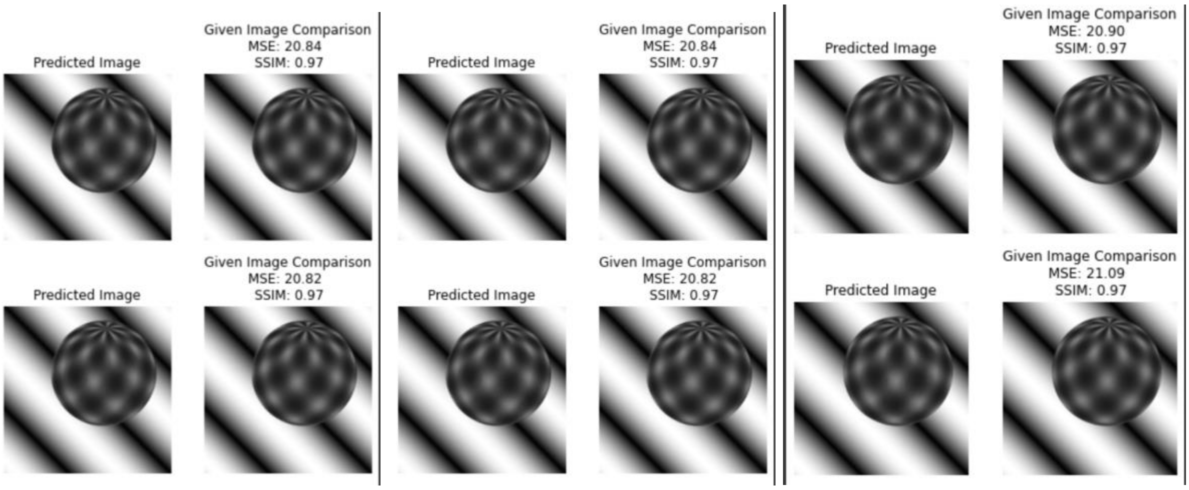
When we solve for the least squares solution for Av = b, we add an additional constraint that we find the solution to the linear systems only if the smallest eigenvalues of AT A is greater than a threshold τ. This was introduced so that there are no large values (noise) for the optical flow and showed better empirical performance.

Q1) We visualised the following relations which can be seen in the images. The mean squared error observed for the corridor prediction was much higher but was comparatively a lot lesser in sphere prediction. The window size was varied as 4,9,12 and 18. I have taken alternate frames prediction i.e., have predicted frame1,3,5 etc.

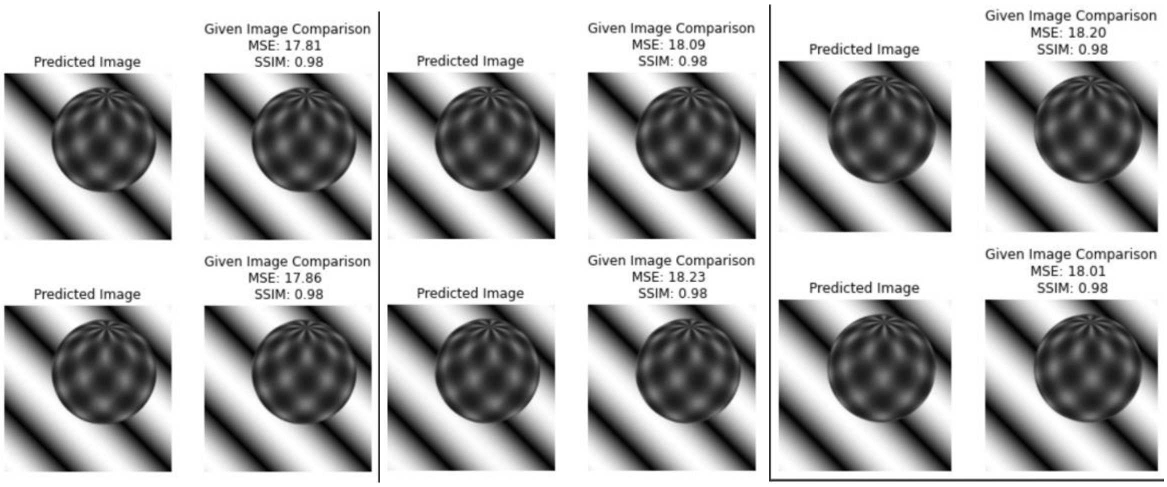


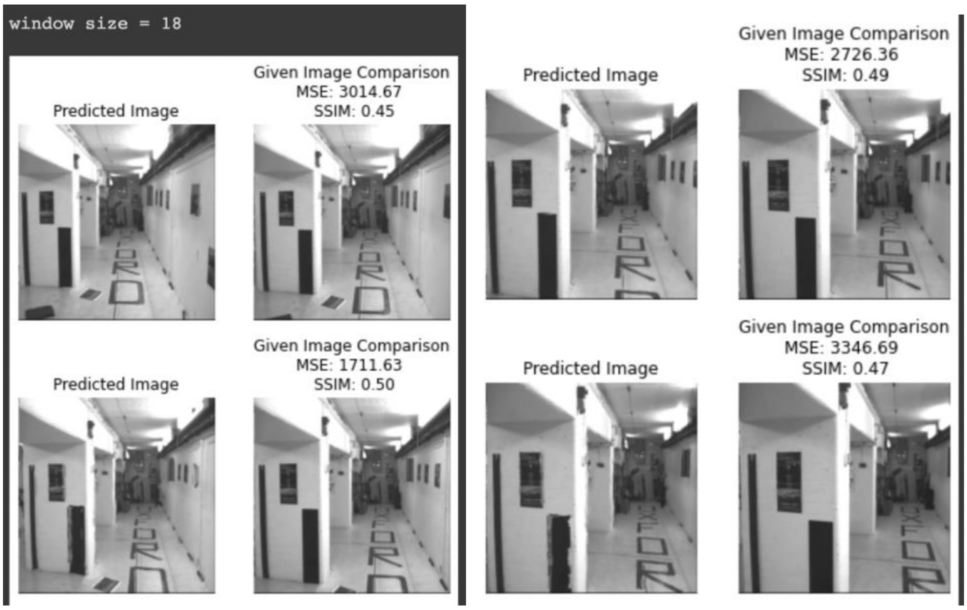


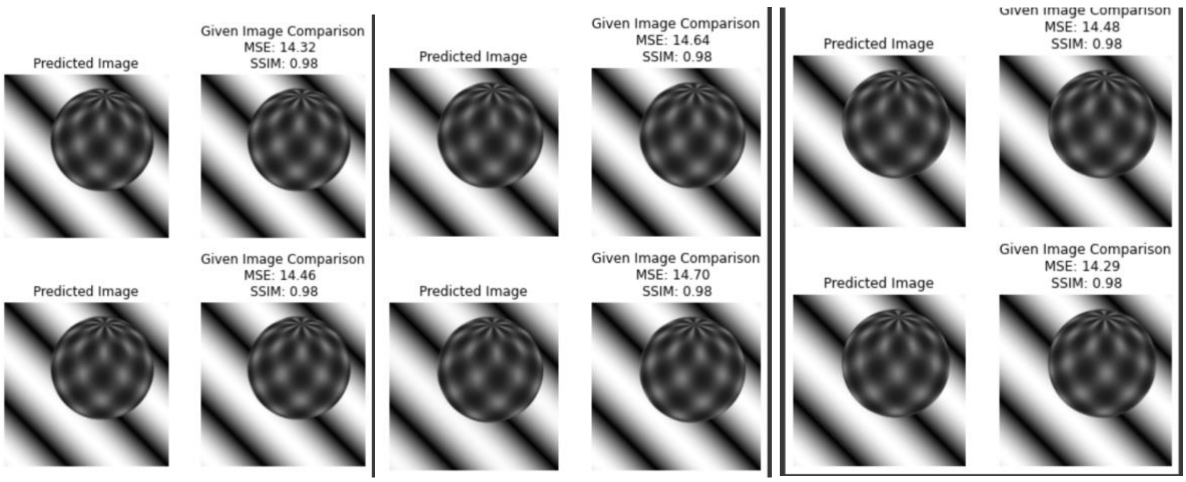



Predicted Image
Given Image Comparison
MSE: 20.84
SSIM: 0.97
Predicted Image
Given Image Comparison
MSE: 20.84
SSIM: 0.97
Predicted Image
Given Image Comparison
MSE: 20.90
SSIM: 0.97
Predicted Image
Given Image Comparison
MSE: 20.82
SSIM: 0.97
Predicted Image
Given Image Comparison
MSE: 20.82
SSIM: 0.97
Predicted Image
Given Image Comparison
MSE: 21.09
SSIM: 0.97


window size = 12
Predicted Image
Given Image Comparison
MSE: 3018.86
SSIM: 0.45
Predicted Image
Given Image Comparison
MSE: 2715.79
SSIM: 0.49
Predicted Image
Given Image Comparison
MSE: 1752.99
SSIM: 0.49
Predicted Image
Given Image Comparison
MSE: 3401.60
SSIM: 0.46


Predicted Image
Given Image Comparison
MSE: 17.81
SSIM: 0.98
Predicted Image
Given Image Comparison
MSE: 18.09
SSIM: 0.98
Predicted Image
Given Image Comparison
MSE: 18.20
SSIM: 0.98
Predicted Image
Given Image Comparison
MSE: 17.86
SSIM: 0.98
Predicted Image
Given Image Comparison
MSE: 18.23
SSIM: 0.98
Predicted Image
Given Image Comparison
MSE: 18.01
SSIM: 0.98





Q2) For numerical performance, I have used the mean square error and the Structural Similarity Index as the metrics to compare the images. We can clearly see in the images that on the corridor frames, the MSE reaches over 1000 and sometimes near 4000 as well which is very high and SSIM therefore observed is quite low for them. For the Spherical frames, MSE obtained is below 20 mostly and therefore SSIM is almost closer to 1.

Discrete Horn Schunck

The gradient of intensity is calculated same as the previous Lucas Kanade method.
The Horn-Schunck algorithm assumes smoothness in the flow over the whole image. Thus, it tries to minimize distortions in flow and prefers solutions which show more smoothness. Advantages of the Horn–Schunck algorithm include that it yields a high density of flow vectors, i.e., the flow information missing in inner parts of homogeneous objects is filled in from the motion boundaries. On the negative side, it is more sensitive to noise than local methods.
We assume our initial estimate of optical flow to be zero. Then for every pixel in the image apply the below iterative algorithm to calculate the original optical flow, for p=0,1, 2...p_{max}:

$$a^{(p+1)} = a_{avg}^{(p)} - \lambda f_x \frac{f_x a_{avg}^{(p)} + f_y b_{avg}^{(p)} + f_t}{1 + \lambda \parallel \nabla f \parallel^2}$$

$$b^{(p+1)} = b_{avg}^{(p)} - \lambda f_y \frac{f_x a_{avg}^{(p)} + f_y b_{avg}^{(p)} + f_t}{1 + \lambda \parallel \nabla f \parallel^2}$$

where, $\nabla f = [\, f_x f_y \,]^T$

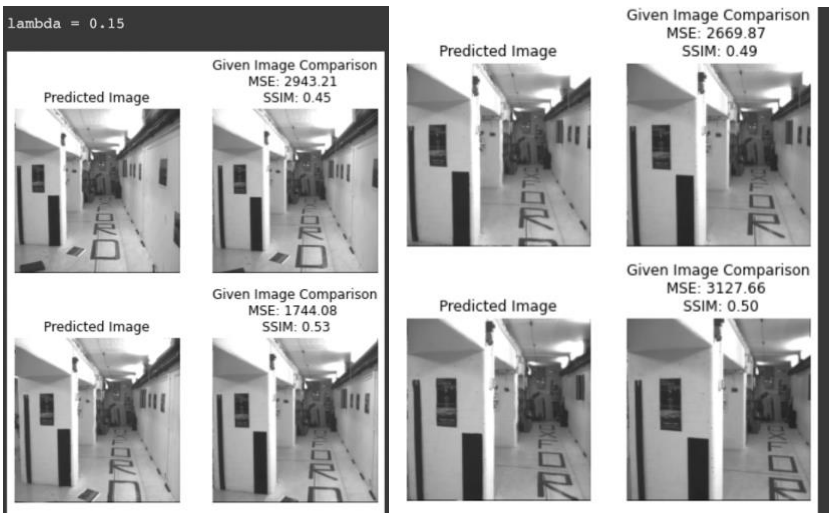
$$a_{avg} = \frac{1}{4}(a[m+1, n\,] + a[m-1\,, n] + a[m, n+1] + a[m, n-1])$$

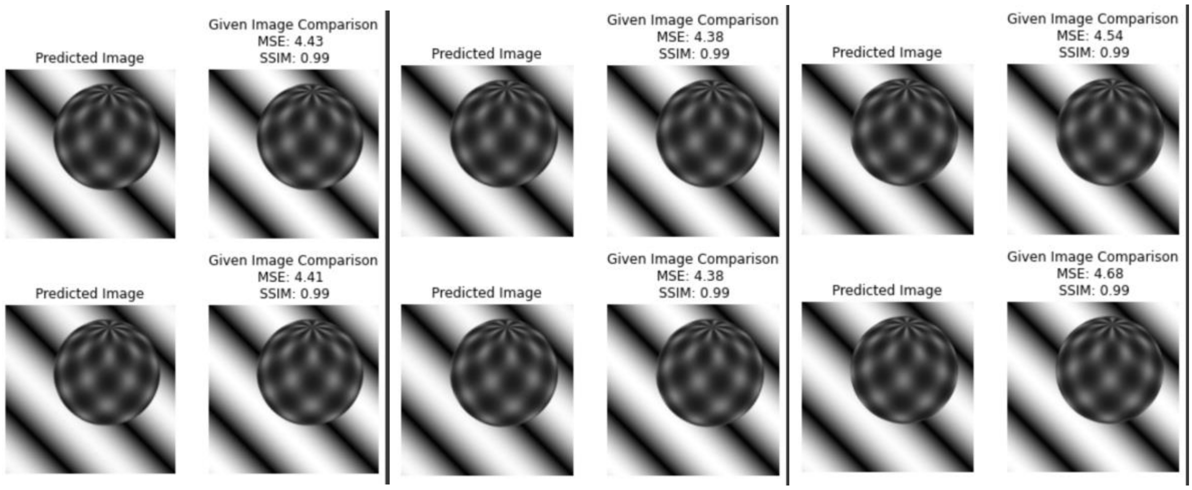
$$b_{avg} = \frac{1}{4}(b[m+1, n\,] + b[m-1\,, n] + b[m, n+1] + b[m, n-1])$$

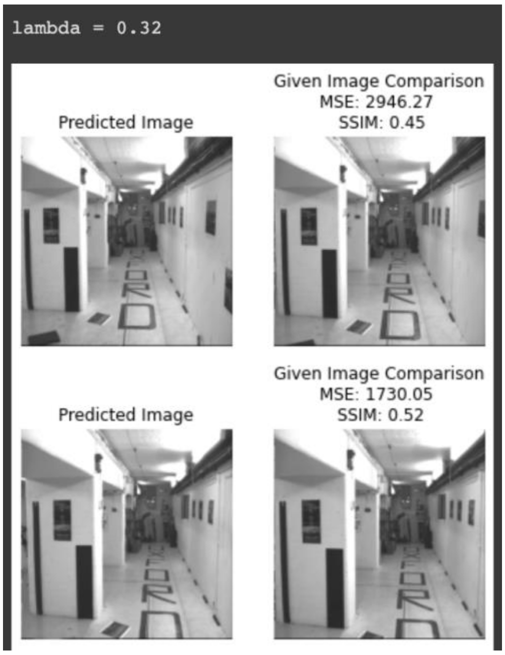
We can also choose the stopping criterion for stooping the algorithm like:

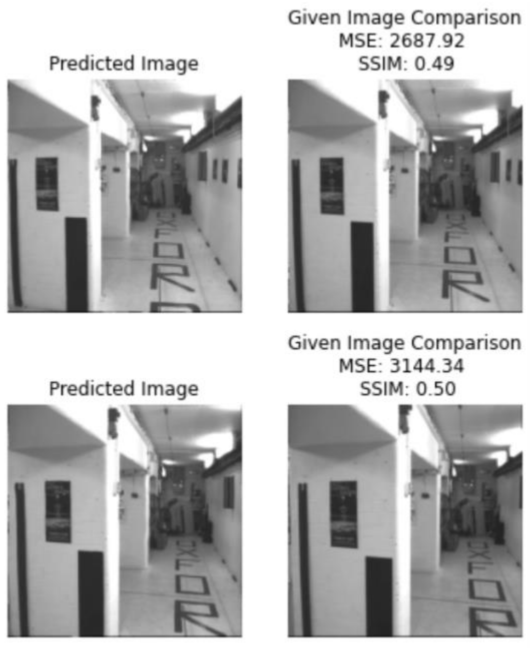
$$max_{(m,n)}\left|a_{(p+1)}(m,n) - a_{(p)}(m,n)\right| < \epsilon$$
$$max_{(m\,n)}\left|b_{(p+1)}(m,n) - b_{(p)}(m,n)\right| < \epsilon$$

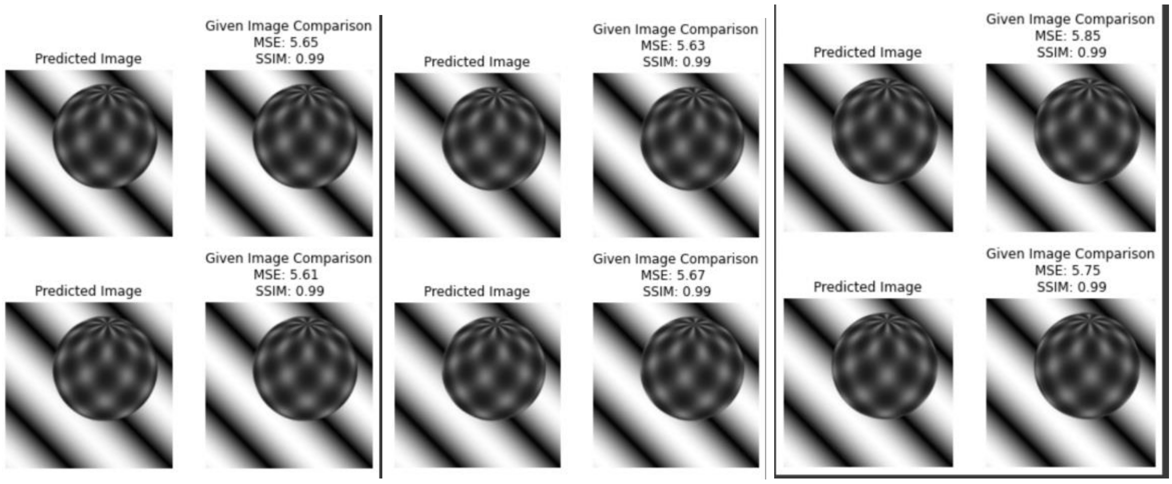
Q1) We visualised the following relations which can be seen in the images. The mean squared error observed for the corridor prediction was much higher comparable to that in Lucas Kanade and much lower in sphere prediction (even lower than LK method). The lambda parameter was varied as 0.15, 0.32, 0.65, 1, 5. I have taken alternate frames prediction i.e., have predicted frame1,3,5 etc.

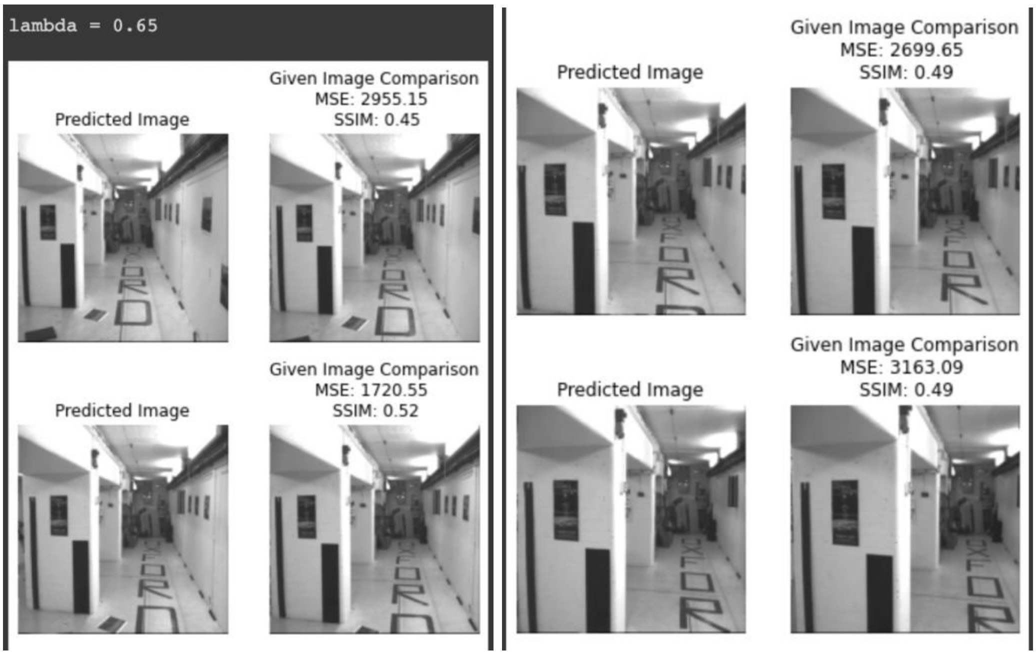


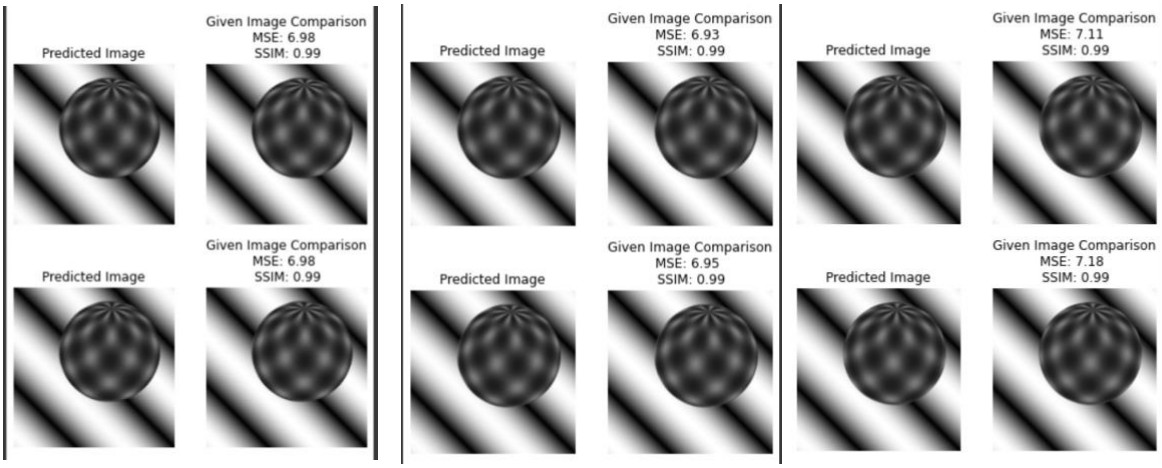

Predicted Image
Given Image Comparison
MSE: 4.43
SSIM: 0.99
Predicted Image
Given Image Comparison
MSE: 4.41
SSIM: 0.99
Predicted Image
Given Image Comparison
MSE: 4.38
SSIM: 0.99
Predicted Image
Given Image Comparison
MSE: 4.38
SSIM: 0.99
Predicted Image
Given Image Comparison
MSE: 4.54
SSIM: 0.99
Predicted Image
Given Image Comparison
MSE: 4.68
SSIM: 0.99

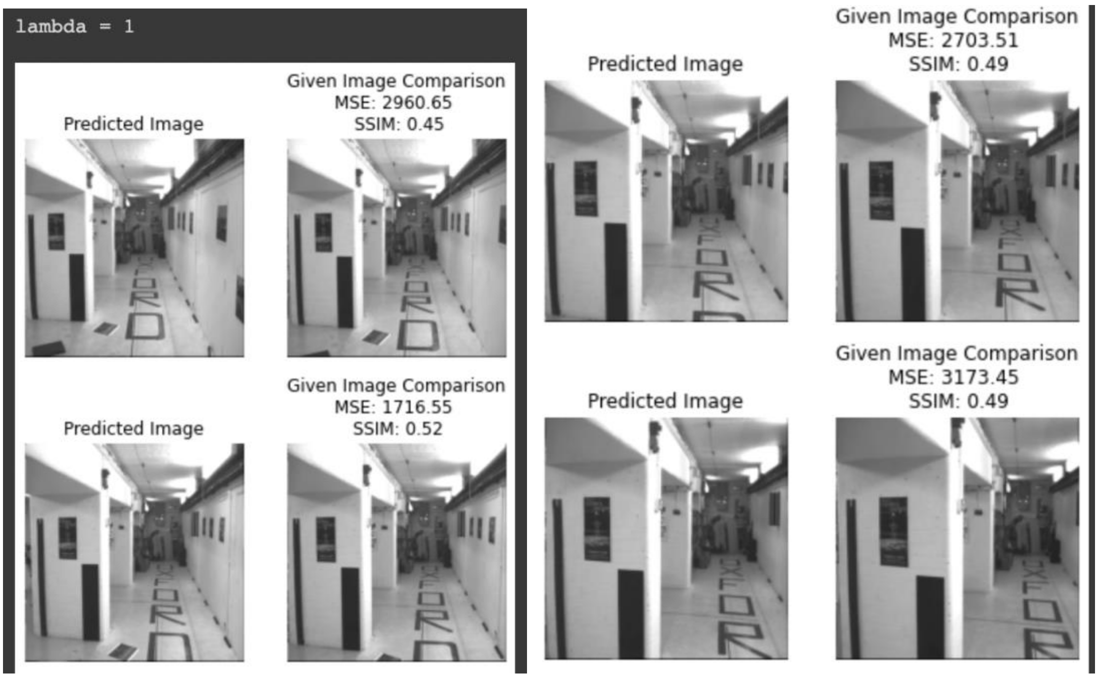

lambda = 0.32
Predicted Image
Given Image Comparison
MSE: 2946.27
SSIM: 0.45
Predicted Image
Given Image Comparison
MSE: 1730.05
SSIM: 0.52

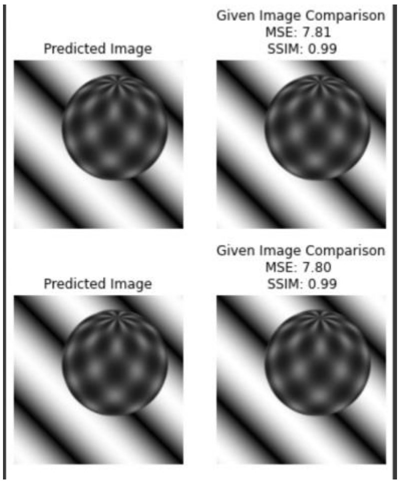
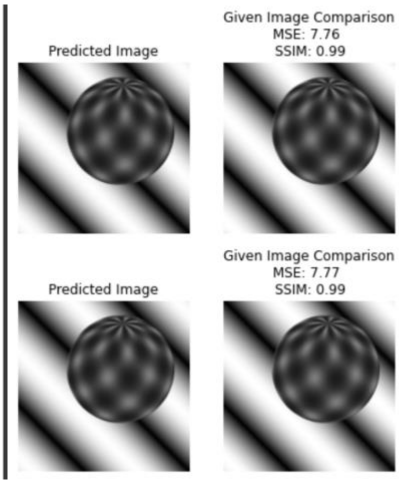
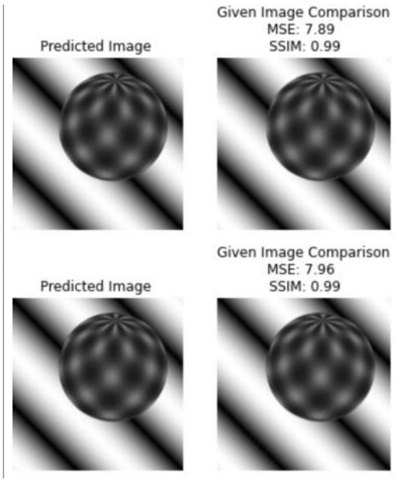
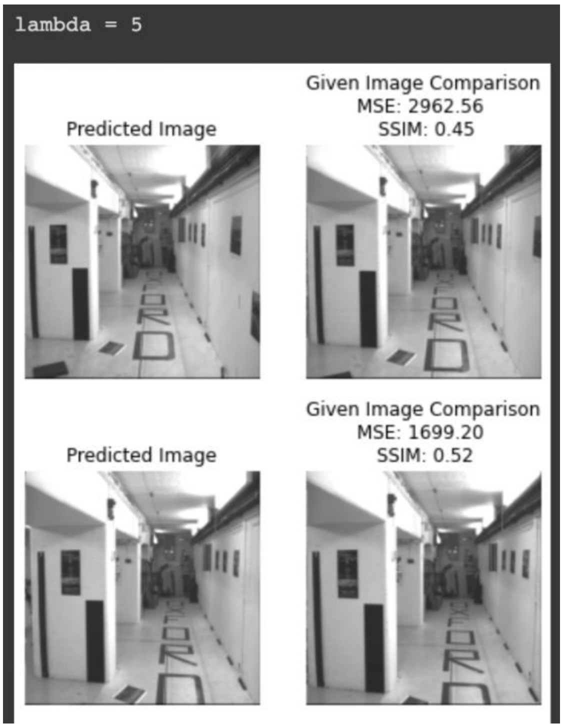
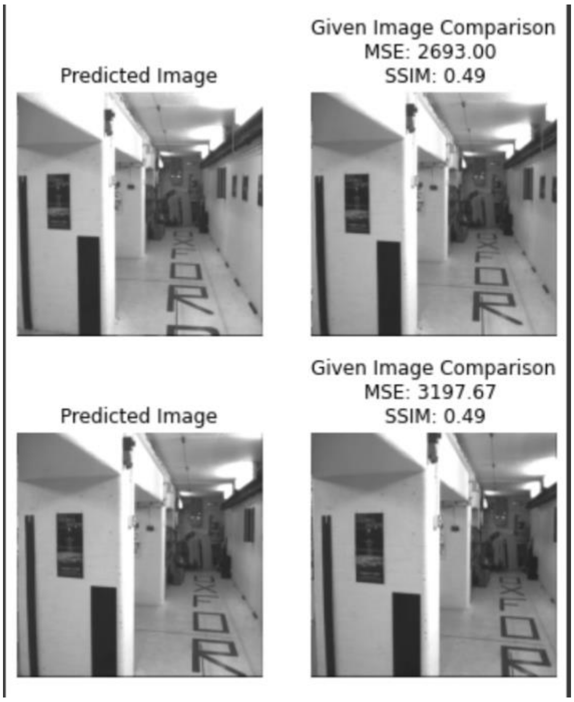
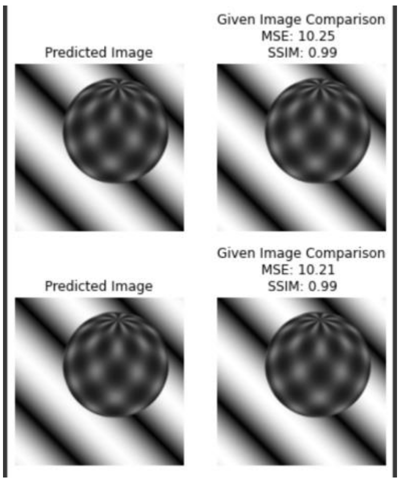
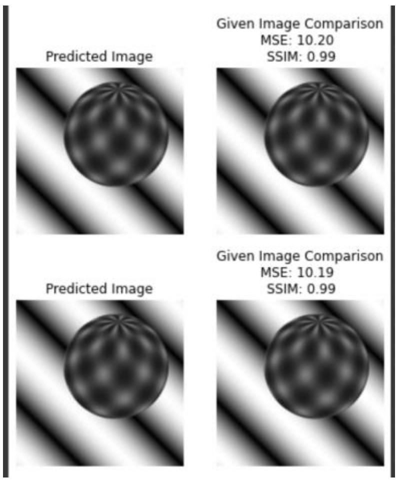
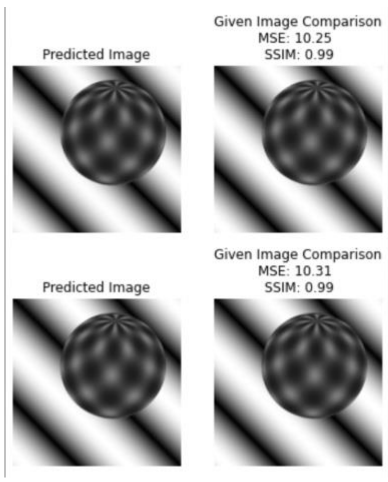

Predicted Image
Given Image Comparison
MSE: 2687.92
SSIM: 0.49
Predicted Image
Given Image Comparison
MSE: 3144.34
SSIM: 0.50


Predicted Image
Given Image Comparison
MSE: 5.65
SSIM: 0.99
Predicted Image
Given Image Comparison
MSE: 5.61
SSIM: 0.99
Predicted Image
Given Image Comparison
MSE: 5.63
SSIM: 0.99
Predicted Image
Given Image Comparison
MSE: 5.67
SSIM: 0.99
Predicted Image
Given Image Comparison
MSE: 5.85
SSIM: 0.99
Predicted Image
Given Image Comparison
MSE: 5.75
SSIM: 0.99


lambda = 0.65
Predicted Image
Given Image Comparison
MSE: 2955.15
SSIM: 0.45
Predicted Image
Given Image Comparison
MSE: 1720.55
SSIM: 0.52
Predicted Image
Given Image Comparison
MSE: 2699.65
SSIM: 0.49
Predicted Image
Given Image Comparison
MSE: 3163.09
SSIM: 0.49


Predicted Image
Given Image Comparison
MSE: 6.98
SSIM: 0.99
Predicted Image
Given Image Comparison
MSE: 6.98
SSIM: 0.99
Predicted Image
Given Image Comparison
MSE: 6.93
SSIM: 0.99
Predicted Image
Given Image Comparison
MSE: 6.95
SSIM: 0.99
Predicted Image
Given Image Comparison
MSE: 7.11
SSIM: 0.99
Predicted Image
Given Image Comparison
MSE: 7.18
SSIM: 0.99


lambda = 1
Predicted Image
Given Image Comparison
MSE: 2960.65
SSIM: 0.45
Predicted Image
Given Image Comparison
MSE: 1716.55
SSIM: 0.52
Predicted Image
Given Image Comparison
MSE: 2703.51
SSIM: 0.49
Predicted Image
Given Image Comparison
MSE: 3173.45
SSIM: 0.49


Predicted Image
Given Image Comparison
MSE: 7.81
SSIM: 0.99
Predicted Image
Given Image Comparison
MSE: 7.80
SSIM: 0.99

Predicted Image
Given Image Comparison
MSE: 7.76
SSIM: 0.99
Predicted Image
Given Image Comparison
MSE: 7.77
SSIM: 0.99

Predicted Image
Given Image Comparison
MSE: 7.89
SSIM: 0.99
Predicted Image
Given Image Comparison
MSE: 7.96
SSIM: 0.99

lambda = 5
Predicted Image
Given Image Comparison
MSE: 2962.56
SSIM: 0.45
Predicted Image
Given Image Comparison
MSE: 1699.20
SSIM: 0.52

Predicted Image
Given Image Comparison
MSE: 2693.00
SSIM: 0.49
Predicted Image
Given Image Comparison
MSE: 3197.67
SSIM: 0.49

Predicted Image
Given Image Comparison
MSE: 10.25
SSIM: 0.99
Predicted Image
Given Image Comparison
MSE: 10.21
SSIM: 0.99

Predicted Image
Given Image Comparison
MSE: 10.20
SSIM: 0.99
Predicted Image
Given Image Comparison
MSE: 10.19
SSIM: 0.99

Predicted Image
Given Image Comparison
MSE: 10.25
SSIM: 0.99
Predicted Image
Given Image Comparison
MSE: 10.31
SSIM: 0.99

Q2) For numerical performance, I have used the mean square error and the Structural Similarity Index as the metrics to compare the images. We can clearly see in the images that on the corridor frames, the MSE reaches over 1000 and sometimes near 4000 as well which is very high and SSIM therefore observed is quite low for them. For the Spherical frames, MSE obtained is below 10 mostly and therefore SSIM is almost closer to 1.
One observation here can be noted that as the parameter value lambda increases, the MSE is observed to increase slightly in the sphere frame set.

Q3) The places where optical flow is different is described below. We can visualize the flow obtained in different method as shown below for the sphere frames: -

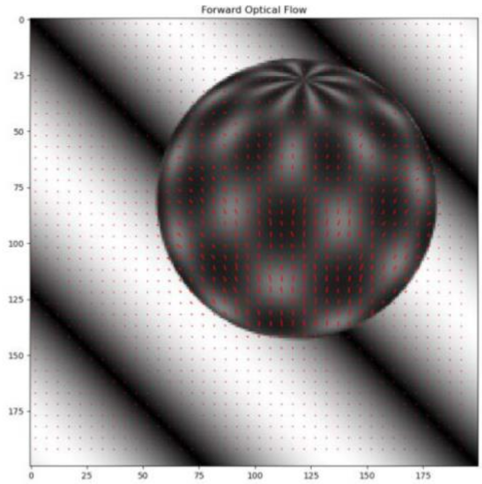


Figure 1: Lucas Kanade Flow

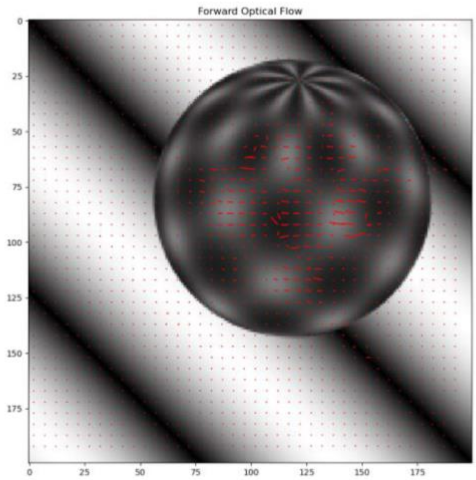


Figure 2: Horn Schunck Flow

In Horn Schunck, the smoothness constraints fail at edges and corners because of which we observe difference in optical flow but Lucas Kanade gives better flow at those locations since the ATA matrix is invertible at those points giving positive eigen values.
But Lucas-Kanade optical flow fails on smooth region as the matrix is not invertible there but instead Horn-Schnuck performs well.

Q4) Occlusion and disocclusion along with large optical flow is the reason behind we are not able to predict some of the intensities.
The sphere is moving, and the camera is fixed for the set of sphere frames which makes it independent of occlusions as even if some points are occluded on one side, the other side takes care of it but in corridor we are moving the camera instead which causes problems to occlusions. At those points, the intensity cannot be predicted. Also, some intensities are formed which are not there in previous frame (i.e., disocclusion) because of which we are not able to predict intensities on those points.
Large optical flow in corridor frames can also be a problem compared to much smaller ones in Horn-Schunck.