
DIGITAL VIDEO: PERCEPTION & ALGORITHMS

ASSIGNMENT-1

NAME: AVNISH KUMAR **S.No.: 20964**

LINK TO ALL VIDEOS WITH COMBINATIONS OF DIFFERENT SCALES AND K VALUES

Q1. Multi-scale Discrete Horn-Schunck optical flow

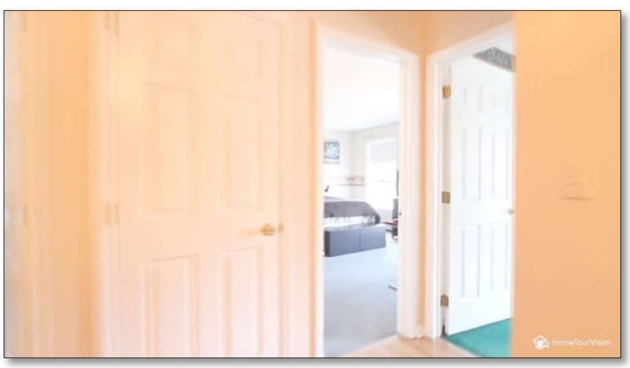
Studying the variations in PSNR and LPIPS:

FOLDER: 00001-

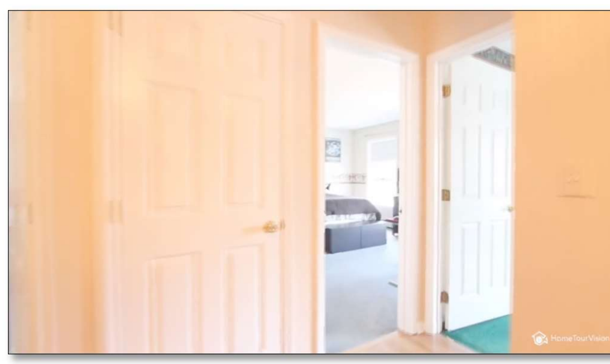
K=2	SCALE=1	SCALE=2	SCALE=4	SCALE=8
PSNR	83.18756809366	83.25580185580	83.21939229838	83.19897752446
LPIPS	0.077	0.076	0.080	0.082

For K=2

Example of interpolated frame-2, scale-1



Example of interpolated frame-2, scale-2



FOLDER: 00006-

K=2	SCALE=1	SCALE=2	SCALE=4	SCALE=8
PSNR	74.82265205094	74.71213188331	74.77341542692	74.72981111832
LPIPS	0.110	0.124	0.143	0.155

For K=2

Example of interpolated frame-2, scale-1

Example of interpolated frame-2, scale-2

Q2. Numerical performance comparisons across scales and for different K

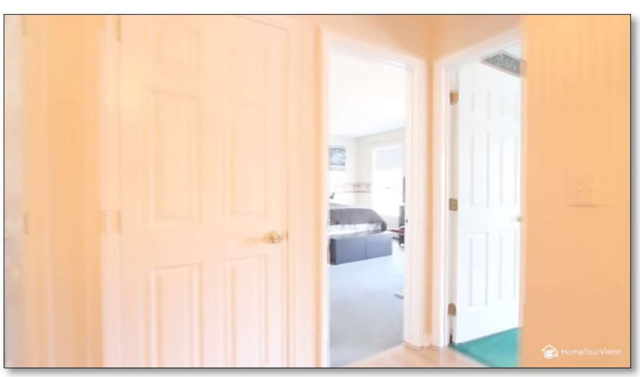
FOLDER: 00001-

K=2	SCALE=1	SCALE=2	SCALE=4	SCALE=8
PSNR	83.18756809366	83.25580185580	83.21939229838	83.19897752446
LPIPS	0.077	0.076	0.080	0.082

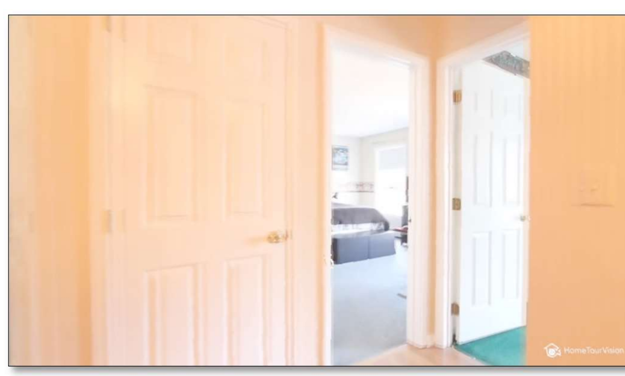
K=3	SCALE=1	SCALE=2	SCALE=4	SCALE=8
PSNR Frame-2	80.73725209085	80.76460682050	80.93983534880	81.0044484201
LPIPS Frame-2	0.073	0.073	0.074	0.078
PSNR Frame-3	80.7085033000	80.6869500804	80.56875831270	80.589445110
LPIPS Frame-3	0.074	0.075	0.083	0.087

For K=3:

Example of interpolated frame-2, scale-2



Example of interpolated frame-3, scale-2



K=4	SCALE=1	SCALE=2	SCALE=4
PSNR Frame-2	79.94872924323	79.9747309396	80.07224049823
LPIPS Frame-2	0.067	0.066	0.067
PSNR Frame-3	77.8854998726	77.90282912075	77.93908883249
LPIPS Frame-3	0.103	0.104	0.109
PSNR Frame-4	79.96222673096	79.87198852808	79.44397445606
LPIPS Frame-4	0.064	0.066	0.079

FOLDER: 00006-

For K=3

Example of interpolated frame-2, scale-2

Example of interpolated frame-3, scale-2

K=2	SCALE=1	SCALE=2	SCALE=4	SCALE=8
PSNR	74.82265205094	74.71213188331	74.77341542692	74.72981111832
LPIPS	0.110	0.124	0.143	0.155

For K=3:

K=3	SCALE=1	SCALE=2	SCALE=4
PSNR Frame-2	73.59696724748	73.65426718026	73.67623340487
LPIPS Frame-2	0.115	0.117	0.127
PSNR Frame-3	73.09500983724	72.72922112722	72.62148830241
LPIPS Frame-3	0.138	0.168	0.200

CONCLUSIONS:

Visual performance of video:

Report Section:

Image Quality Analysis

In the reconstructed videos using the Horn-Schunck optical flow method, noise is notably present along edges. This phenomenon arises from the violation of the luminance consistency condition, a key assumption of this optical flow technique.

1. Effect of Change in $'\mathcal{K}'$ Value

Furthermore, our investigation revealed that increasing the 'K' values in the method tends to result in less perceptually pleasing videos, particularly in regions prone to occlusions. As the 'K' value increases, the impact of occlusions becomes more pronounced. Notably, this issue is particularly evident in the images from folder 00001, where the area near the door exhibits less smooth transitions between frames. This degradation in visual quality is reflected in the decreasing PSNR values, indicating the introduction of more noise as the 'K' value increases.

Note: Similar results were observed for folder 00006 images.

2. Effect of Change in Scale

Another significant observation is the effect of changing the scale factor during Multiscale optical flow, for example, scaling to 2x. Initially, this change leads to an increase in PSNR (Peak Signal-to-Noise Ratio). However, as the scale factor is further increased, PSNR begins to decrease. Simultaneously, LPIPS (Learned Perceptual Image Patch Similarity) scores tend to increase. This behavior is due to the fact that at larger scales, the primary optical flow calculated can be significantly larger in the original frame, even when there are minimal actual changes (dx/dt) in the positions of objects within the image. Similarly, LPIPS initially decreases and then tends to increase.

3. Consideration of $\mathcal{LPIPS}$

Lastly, it is worth noting that LPIPS may not consistently align with human perception, particularly when assessing noise near edges. This discrepancy can be attributed to the model's inherent biases, which are shaped by specific perceptual features learned during training. Consequently, LPIPS might overlook important distinctions or similarities that are contextually significant for a given task.

These observations shed light on the nuances of image quality assessment when utilizing optical flow methods, highlighting the importance of considering both traditional metrics like PSNR and perceptual metrics like LPIPS to comprehensively evaluate visual content.

Click on the hyperlink below to access the interpolated and un-interpolated videos of Folder 00001 and 00006.

LINK: Videos made with combinations of K = 2,3,4 and Scales = 1,2,4

Simple heuristics

3. While estimating the interpolated frame, there could be pixels which do not map to any location to either the forward or the backward frame. So, we have to rely on one of the frame to get the pixel otherwise there will be holes present in the output image and will affect overall perception of the video.

So, When there is occlusion present in both the forward and backward frame. I simply use the pixel from the previous frame.

AVERAGE LPIPS CALCULATIONS FOR DIFFERENT K- VALUES AND SCALES:

Folder 00001 All frames average LPIPS:

SCALE:1

- **K=2 AVG: 0.10521 +/- 0.00835**
- **K=3 AVG: 0.13555 +/- 0.00983**
- **K=4 AVG: 0.11650 +/- 0.01115**

SCALE:2

- **K=2 AVG: 0.13489 +/- 0.00945**
- **K=3 AVG: 0.13289 +/- 0.01217**
- **K=4 AVG: 0.13462 +/- 0.01251**

Folder 00006 All frames average LPIPS:

SCALE:1

- **K=2 AVG: 0.15791 +/- 0.01259**
- **K=3 AVG: 0.15950 +/- 0.01405**
- **K=4 AVG: 0.20330 +/- 0.01543**

SCALE:2

- **K=2 AVG: 0.20009 +/- 0.01318**
- **K=3 AVG: 0.21276 +/- 0.01556**
- **K=4 AVG: 0.22257 +/- 0.01628**