Due: Mar. 22, 2024 @ 11:59 p.m.

General notes to keep in mind:

- *All deliverables for the assignment must be submitted as a* ***single ZIP file per group*** *via the Brightspace D2L course shell. Submissions containing multiple ZIP files per group or those with a file that is not in the ZIP format will NOT be graded.*
- *The code submitted must be written purely using the Python programming language and it should execute within the Python 3.12.1 interpreter running on the Windows operating system (version 10 or above). The submitted code should NOT require external python modules other than opencv-python 4.9.0.80, flowpy 0.6.0, tensorflow 2.15.0.post1, scikit-image 0.22.0, fastai 2.7.14, torch 2.2.0, scikit-learn 1.4.0, matlplotlib 3.8.2, pandas 2.2.0, imagecodecs 2024.1.1 and their dependencies.*
- *Read the "Assignment code submission requirements" carefully and prepare the code accordingly. It is your responsibility to ensure that the submitted code executes. If the grader is unable to execute your code and/or your code does NOT adhere to the submission requirements, your code may not be graded.*
- *The written responses required to the questions in the assignments must be compiled into* ***single PDF*** *file named as* `report.pdf`*. You are encouraged to use LaTeX for typesetting your written responses, but however, the use of Microsoft Word™ or any other such programs is also acceptable.*

Optical flow estimation

The Middlebury Optical Flow dataset has a rich history that traces back to the early 2000s when computer vision researchers recognized the need for standardized benchmarks to evaluate the performance of optical flow algorithms. Conceived at Middlebury College in Vermont, USA, the dataset made its debut in 2005 and has since undergone several updates and expansions. The initial release consisted of challenging image pairs captured in controlled environments, laying the foundation for benchmarking optical flow methods. Over the years, the dataset evolved to include diverse scenes, varying lighting conditions, and complex motion patterns, reflecting the growing demand for robust algorithms in real-world applications. Its pivotal role in fostering advancements in optical flow research is evident as the dataset became a staple reference in countless publications and competitions. With each iteration, the Middlebury Optical Flow dataset continued to shape the landscape of computer vision by providing a reliable and comprehensive platform for testing and refining optical flow algorithms. The dataset creators have held a longstanding challenge to independently evaluate the various submitted optical flow estimation algorithms on an set of benchmark image sequences (see Figure 1). After 15 years and over 700 submissions this challenge was finally retired in April last year. The final leaderboard for the Middlebury Optical Flow benchmark challenge can be seen here. The enduring legacy of the Middlebury Optical Flow benchmark challenge underscores the significance of benchmark datasets in driving progress within the field of computer vision. In this assignment, you will **gain familiarity with optical flow estimation using images from the Middlebury Optical Flow dataset.**

(a) Initial frame (b) Final frame

Figure 1: The "Basketball" image sequence from the Middlebury Optical Flow benchmark evaluation dataset [Szeliski et al. 2011].

Data

The data for this assignment can be downloaded from here. It consists of the "Grove3" and "RubberWhale" color image sequences from the Middlebury Optical Flow dataset. For each image sequence, the initial frame is given in `frame10.png` and the final frame is given in `frame11.png`. The ground truth optical flow from the initial to final frame is given in `flow10.flo`. The frame taken in the "middle" of the initial and final frames is given in `frame10i11.png`. Two other image sequences from the Middlebury Optical Flow dataset are gathered and are intentionally kept "blinded" from you and will be used to perform an independent validation of your submitted optical flow method in Question 5.

Question 1 [10 marks]

Generate visualizations of the ground truth optical flow similar to the ones shown in *Figure 2* in [Szeliski et al. 2011] for the "Grove3" and "RubberWhale" image sequences. You may find the flowpy 0.6.0 package useful for this purpose.

Question 2 [25 marks]

Using the Farnebäck algorithm (see Optical Flow tutorial OpenCV), compute the "forward" optical flow, i.e., flow from the initial to the final frame for both the "Grove3" and "RubberWhale" image sequences. Generate the optical flow maps using the Red, Green, Blue channels respectively. Also, generate the optical flow map using the grayscale image obtained using the "preferential weighting of the Green channel" method mentioned in *Lecture 4, Slide 16*. Generate visualizations of the estimated optical flow maps in each of the four cases and determine their accuracy using the average *endpoint error metric* (see *Equation 23* in [Szeliski et al. 2011]) calculated in reference to the given ground truth optical flow map in `flow10.flo`. Provide a short discussion comparing the accuracies of the estimated optical flow maps across the four cases.

Question 3 [15 marks]

Repeat experiment in the above Question 2, but now switching the initial and final frames around, i.e., compute the "backward" optical flow from the final frame to initial frame. Visually compare the "backward" optical flow maps with the "forward" optical flow maps estimated in Question 2. Briefly comment on the findings of your visual comparison.

Question 4 [25 marks]

Generate estimates of the "middle" frame for both the "Grove3" and "RubberWhale" image sequences using *backward warping* with *bilinear interpolation* and the "best" (in terms of the average endpoint error) "forward" and "backward" optical flow maps obtained in Question 2 and Question 3 respectively. Provide visualizations of these two different "middle" frame estimates obtained using these "best" "forward" and "backward" optical flow maps respectively, and determine their accuracy using the *root-mean-square (RMS) difference metric* (see *Equation 24* in [Szeliski et al. 2011]) calculated in reference to the given ground truth "middle frame" in `frame10i11.png`. Provide a short discussion comparing the accuracies of the estimated "middle frames"s obtained using the "best" "forward" and "backward" optical flow maps respectively.

Question 5 [25 marks]

Leveraging the knowledge you have gained so far in optical flow estimation and/or using the various optical flow algorithms available in OpenCV, come up with the "best" approach for computing the optical flow between two given color image frames. You may use any other optical flow map estimation strategy that we have not discussed in the course to design this "best" optical flow estimation approach. Submit this "best" optical flow estimation approach as the following method:

```
def computeOpticalFlow(initial_frame_path, final_frame_path, flow_map_path, model_dir_path):
"""Computes the optical flow map from the color image in "initial_frame_path" to the
color image in "final_frame_path" and saves the estimated optical flow map to "flow_map_path"

initial_frame_path          full path to a color image in .png format
```

```
final_frame_path            full path to a color image in .png format

flow_map_path               full path to the save the flow map in .flo format

model_dir_path (optional)   full path to the folder containing deep learning model related files
"""
    # your code goes here
    # ....
```

The above method will be evaluated on two different image sequences by the grader to determine its optical flow map estimation performance. The performance will be quantified using the average endpoint error across the two image sequences. Note that you are required to submit the deep learning model files in case your optical flow method is based on a deep neural network.

Note on grading

The grading for Question 1, Question 2, Question 3 and Question 4 will be based on the appropriateness of the submitted code and the written responses. The grading for Question 5 will be based on the relative performance of your submitted optical flow estimation method. The submission(s) with the best performing (referred below as 1^{st} ranked model) in terms of the average *endpoint error* (rounded to 2 decimal places) will receive full marks on Question 5 (i.e., 25 marks). All other submissions will receive marks that are proportional to the increase in the achieved average endpoint error with respect to the 1^{st} ranked model. For example, if the average endpoint error of the optical flow method of a given submission is 10% higher than the 1^{st} ranked one, then that submission will receive 22.5 marks for Question 5.