Answers to Theory Questions

Section 1

Q1,1,1) Describe how optical flow could be used to create slow motion video.

A) The optical flow equation helps us to find the distribution of apparent velocities of movement of brightness patterns of pixels in a given image. One naive way would be to reduce the frame rate of the video while keeping the number of frames constant. This would show the changes happening in the video slowly. But in such a naive solution, the changes between 2 consecutive frames will stay the same despite changing the gaps between which these 2 frames are shown.
And that is exactly what we call frame interpolation.
Let 2 consecutive captured frames be F1 and F2. Then, we can try to estimate what happened between these 2 frames by a technique called **frame interpolation** (synthesizing several frames in the middle of two adjacent frames of the original video.).
Prof Nayar mentions this in the 135th sec of his lecture (link).
If we are able to estimate the motion of a pixel between images F1 and F2, we can also estimate the position of the pixel in the scene in hypothetical intermediate frames between F1 and F2 and can then use this information to artificially render the intermediate frames.

Q1.1.2) Explain briefly how optical flow is used in the matrix's bullet-time scene.

A) As per Wikipedia, Bullet time is a visual effect or visual impression of detaching the time and space of a camera (or viewer) from those of its visible subject. In the video, there were 2 unconventional things:
   a) *Extreme transformation of time* (slow enough to show normally imperceptible and unfilmable events, such as flying bullets). **Obstacle**: Filming at such a high frame rate so as to catch a bullet moving is difficult
   b) *Extreme transformation of space* (by way of the ability of the camera angle—the audience's point-of-view—to move around the scene at a normal speed while events are slowed). **Obstacle:** Capturing at such a high frame rate and that too from different viewing angles would require a physical camera would have to move **implausibly fast.**

On reading about the filming of the scene in this detailed blog, I found that the "space transformation problem" was solved somewhat by *placing 120 still cameras and two film cameras in a 360 degree angle around the action spot.* Placing the cameras close helped create the illusion of motion, as each camera only captured a single still photo. These cameras were triggered at extremely close intervals, so the action would appear to unfold slowly as the viewpoint moved at a "normal" pace.

Once the images from all the cameras at the scene were scanned into the computer, **motion interpolation** (using algorithms to "animate" more frames in a sequence of stills) was used.
As mentioned in my previous answer, knowing how much a pixel moved in time "t" between 2 images I1 and I2 helps to estimate the position of the pixel in theoretical intermediate frames between I1 and I2. Placing these intermediate frames between the actual captured frames creates a **perceived increase in frame rate**. This makes the motion feel less jittery.

Q1.1.3) So breathtaking, heartbreaking and brimming with emotion, WDMC is a journey into the afterlife and deals with a dead man's attempt to reunite with his wife. Catch a glimpse of the "Painted World" here! Describe briefly on how optical flow is used to create this "painterly effect".

A) The scene had 2 peculiar things:
   a) The world appears to be made of painted brushstrokes

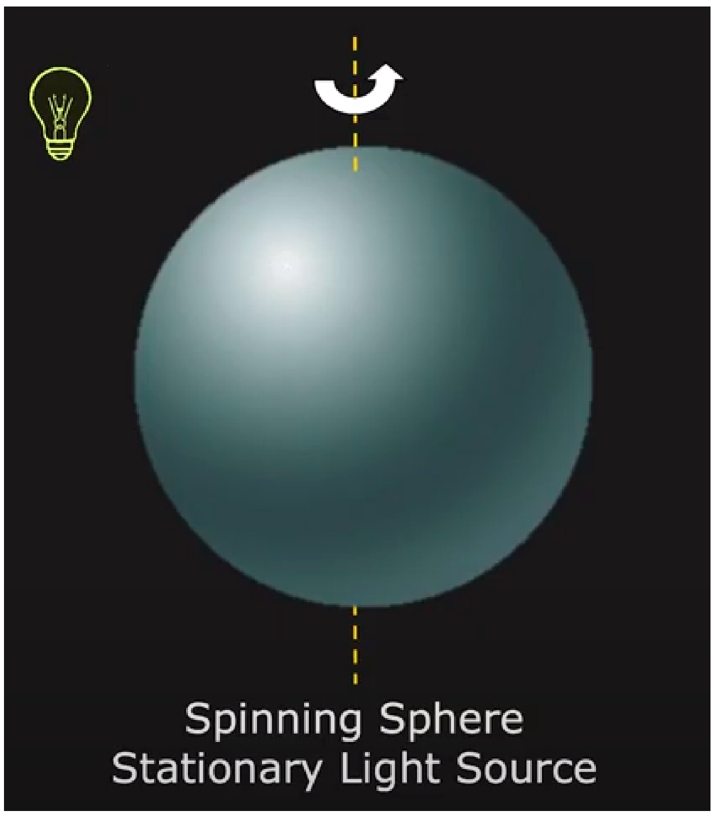
b) The background always contains a bright backlight no matter what position the camera is in. However, the foliage reacts realistically to the wind and the actor's movements.

As I read in this blog and this blog , the following were the steps followed:

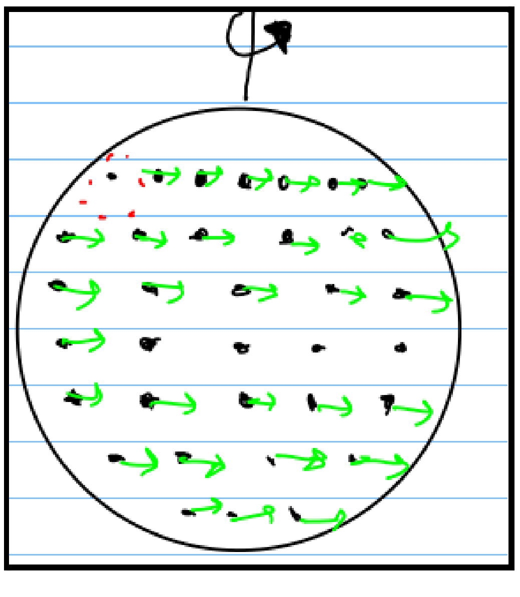
- William's scenes (motion) were shot naturally in the Glacier National Park with orange markers spread throughout the scene.
- The actor was then removed from the footage by rotoscope and the hole was filled by merging the surrounding optical flow map.
- The orange markers were used to recreate the basic movement of the camera in 3D. Optical flow vector maps were created to track each pixel throughout the frames.
- LIDAR was used to scan topography and create a cloud of points, and helped reconstruct the 3D information about the landscape where the scene was shot.
- Tracking the original footage shot gave the creators a complete vector map of the movement of each pixel in the original natural sequence.
- The painted objects/effects were then added to the desired pixels on the footage.
- Since using optical flow, the creators already knew the motion of the original pixels, they were able to apply the same optical flow vectors to give the impression of movement for the newly added objects.
- Finally, the actor was dropped back into the sequence.

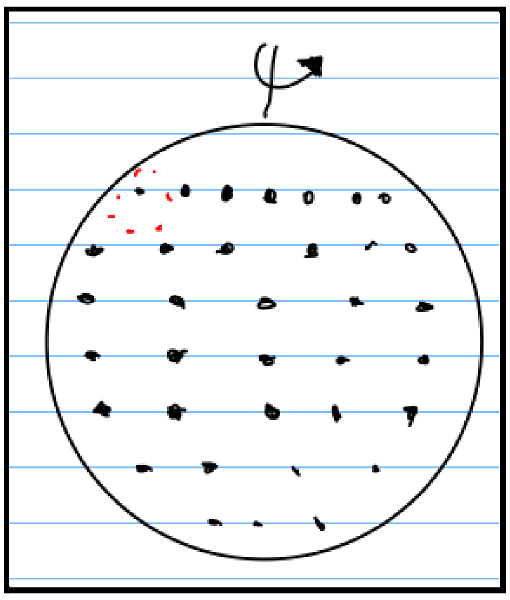
Q1.1.4)Consider a Lambertian ball that is: (i) rotating about its axis in 3D under constant illumination and (ii) stationary and a moving light source. What does the 2D motion field and optical flow look like in both cases.

In case 1,

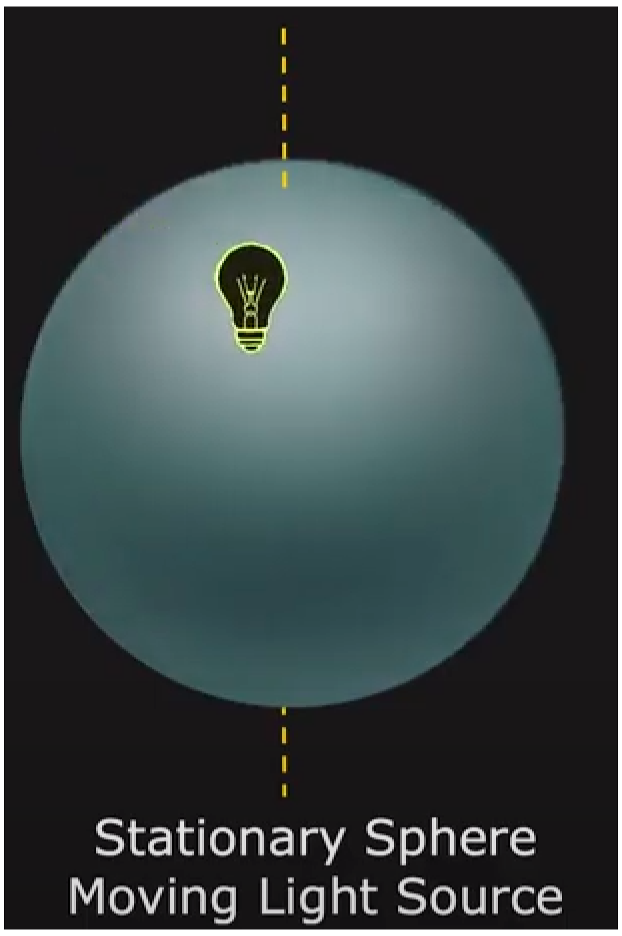


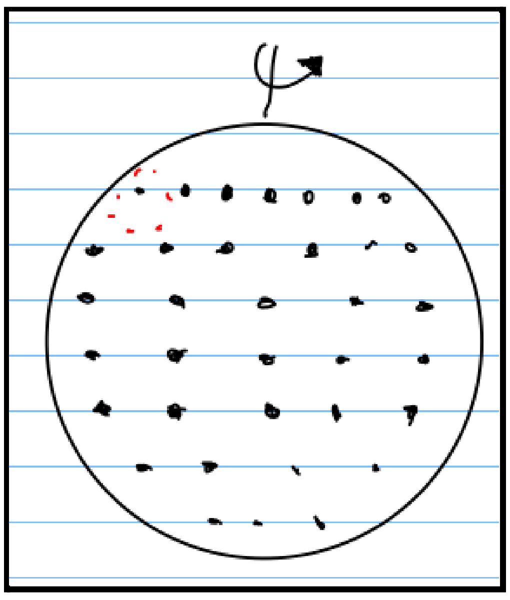
Above is a sphere rotation in the direction shown.

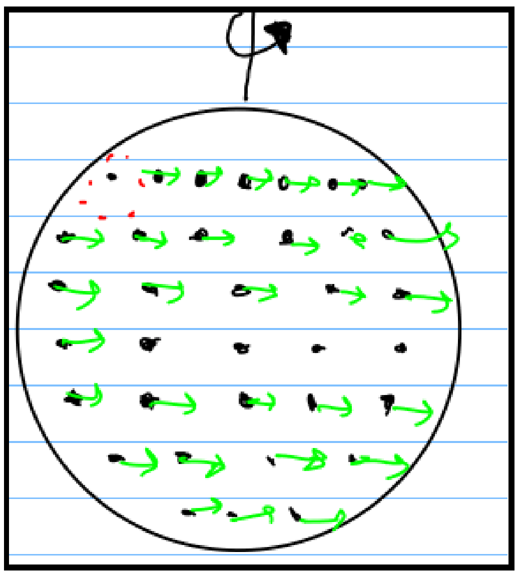
- Motion Field NOT EQUAL to zero as points are physically moving
    - 

- Optical flow = ZERO as no changes are observed in the brightness as all the points on the sphere have the same reflectance property. All images taken by a stationary camera will be identical to each other.
    - 

In case 2 ie moving light source.



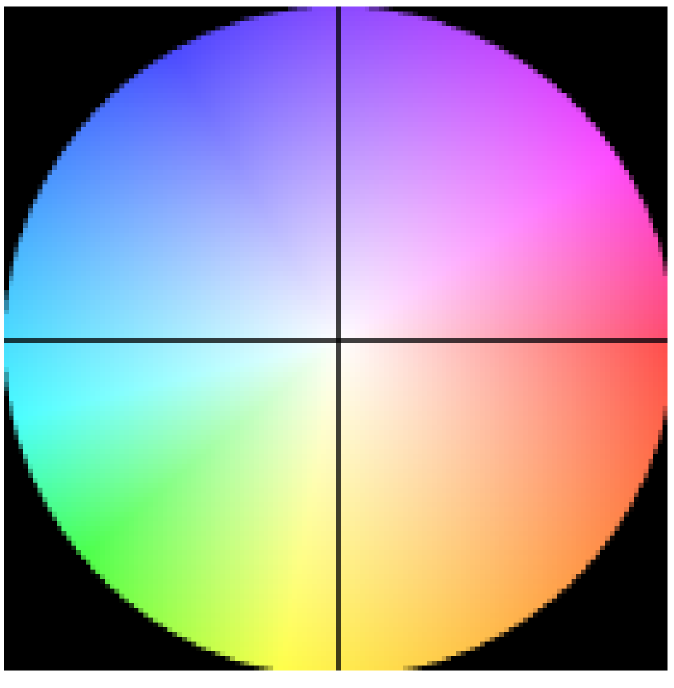
- Motion Field EQUAL to zero as points are NOT physically moving
    - 
- Optical flow NOT EQUAL TO zero brightness pattern change in consecutive frames due to the motion of the illuminating source.

○ 

Answer to some of the parts in Q2: (Other answers are there in the written version)

Q5) Did you observe that the ground truth visualizations are in HSV color space? Try to reason it.

I noticed the visualization color wheel was as follows:



Here, the direction is coded by hue and length is coded by saturation.

On reading more, I found that:

- Hue is the color portion of the model, expressed as a number from 0 to 360 degrees.
- Saturation describes the amount of gray in a particular color, from 0 to 100 percent. Reducing this component toward zero introduces more gray and produces a faded effect.
- Value works in conjunction with saturation and describes the brightness or intensity of the color, from 0 to 100 percent

In optical flow, we have two things to represent mainly:

- Direction of flow
- Magnitude of flow

It is not obvious and trivial to represent these 2 quantities in RGB space (arithmetic operations do not visualize well to the human eye in RGB space). For eg: if we correlate the magnitude of flow with the R-component, the human eye won't clearly be able to know the difference from (LOW RED, G, B) to (HIGH RED, G, B).

Due to the directional use case in optical flow visualization, HSV color space works well.

To the human eye, on viewing an image (where optical flow is plotted in HSV space), both direction and magnitude become easily evident. Eg:

- Directional vectors can be mapped to the H (vector angle) and hence, if I observe that the color at a pixel is red, I can easily interpret it to be moving right.
- Saturation value can be mapped to the magnitude of the vector, Hence, pixels having colors closer to white can be interpreted as NON-MOVING as per our estimated optical flow.