Event-based Robot Vision. TU Berlin SS 2023
Midterm 21.06.2023

Student name: _____ Student TU Berlin ID: _____

"Closed books" (no slides, no cheat-sheet, no code, no videos). No (smart-)phones. No Google search.

Honor code: I hereby certify that I have not given or received aid in the examination.
Sign your name and date: _____

Multiple choice questions: 5 pts each.

1.The Dynamic Vision Sensor (DVS) is inspired in:

- ☐ The sustained pathway of the human visual system.
- ☐ The camera invented by Eadweard Muybridge.
- ☐ The transient pathway of the human visual system.
- ☐ The V1 and V4 layers of the visual cortex.

2. An ATIS camera is one that acquires visual information and outputs it in the form of:

- ☐ Synchronous change detection events and synchronous grayscale frames.
- ☐ Synchronous change detection events and synchronous grayscale events.
- ☐ Asynchronous change detection events and synchronous grayscale frames.
- ☐ Asynchronous change detection events and asynchronous grayscale events.

3. The potential disadvantages of event cameras are:

- ☐ Noise is abundant and non-stationary.
- ☐ Data output is not compatible with computer vision algorithms for standard cameras.
- ☐ Spatial (pixel) resolution is not as high as in current standard cameras.
- ☐ All of the above.

4. With a stationary standard camera viewing a golf player, the player swinging and the people standing still watching the player are:

- ☐ Both undersampled.
- ☐ Oversampled and undersampled, respectively.
- ☐ Undersampled and oversampled, respectively.
- ☐ Both oversampled.

5. The original (non-linearized) event generation model…

- ☐ States that events are triggered by moving edges or flickering lights.
- ☐ Assumes brightness constancy.
- ☐ Assumes a Taylor probability distribution for the contrast threshold C.
- ☐ Is given by a formula in terms of the image gradient.

6. Early event-based tracking methods manually defined a "template", that is, a model of the appearance of the object to track. Then, as events are produced by the object, different data association strategies try to assign the event to the template. Which is true?

- ☐ ICP (Iterative Closest Point) performs soft data association, assigning each event to multiple template points, according to a probability density function.
- ☐ EKLT (Event-based Lucas Kanade tracking) does not have data association.
- ☐ ICP (Iterative Closest Point) performs hard data association, assigning each event to only one template point.
- ☐ PnP (perspective n-point problem) has a one-to-one data association, assigning each event to a previous event.

7. In Exercise 4, what is the state that needs to be updated with every incoming event to achieve image reconstruction from events?

- ☐ The state is the time map.
- ☐ The state is the brightness map.
- ☐ The state consists of the time and brightness maps.
- ☐ The state consists of one time map for each event polarity and a brightness map.

8. In Exercise 4, what is the effect of the cut-off frequency alpha?

- ☐ The larger the alpha, the longer the integration (fill-in homogeneous regions) effect.
- ☐ The smaller the alpha, the longer the integration effect.
- ☐ Alpha does not affect the integration result, which only depends on the number of events processed.
- ☐ In this low-pass filter, the smaller the alpha, the smaller the noise gets through it.

9. In the contrast maximization method (Exercise 6), each pixel of the image of warped events…

- ☐ Counts pairs of associated events along point trajectories.
- ☐ Counts the number of events triggered at that pixel.
- ☐ Counts the number of events along a point trajectory.
- ☐ Computes the average timestamp of the events at that pixel.

10. In the contrast maximization method (Exercise 6), what is the (user-supplied) additional knowledge used?

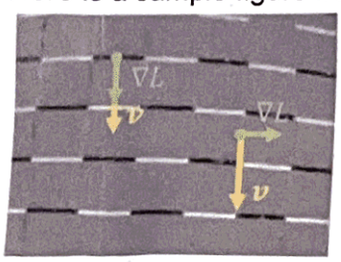
- ☐ Information from all past events.
- ☐ Appearance information of the scene (i.e., the scene map).
- ☐ The type of motion / type of warped used.
- ☐ There is no additional knowledge.

1. Name **three** potential advantages of event-based cameras. How can each of the advantages that you mentioned be explained by the pixel design of the DVS, or by the corresponding sampling model?

   Example: Advantage X is due to ... (15 pts)
   (Please answer in another sheet of paper)

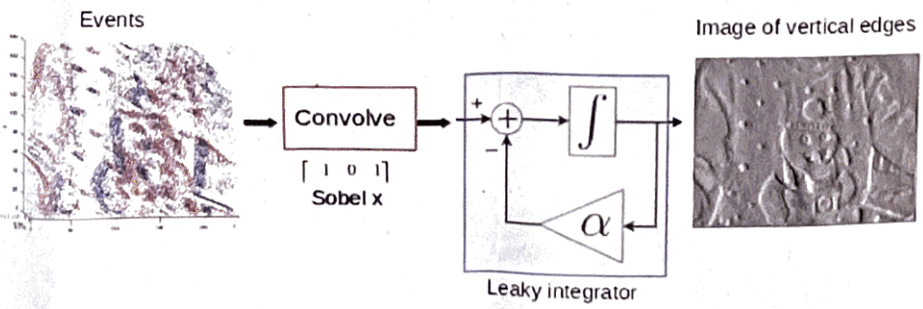
2. What is the event representation known as **Voxel grid**? What are its advantages and disadvantages? (10 pts)
   (Please answer in another sheet of paper)

3. Assuming constant illumination, please explain the relationship between the **event generation model(s)** and the optical flow. Please provide equations and explain how they are connected, mathematically. (15 pts).
   Here is a sample figure from one of the models, to trigger your memory:



   (Please answer in another sheet of paper)

4. **Coding. Exercise 5. dvs_integrator_conv**



   An event e_k = (x_k, y_k, t_k, pol_k) is produced by the event camera. Assuming that we use the simplified 3-element Sobel-x kernel shown in the figure, how do you implement the update equations of the state variables? That is, the equations that implement the convolution and the leaky integration with decaying parameter (cut-off frequency) alpha. Write (C++) code or pseudocode using the state variables and the event e_k, and please skip the zeros in the kernel (i.e., avoid wasteful multiplications by zero). (10 pts)

   (Please answer in another sheet of paper)