Indian Institute of Science Education and Research Bhopal

Computer Vision(DSE-312/EECS-320)

Assignment-3

Name: Aditya Mishra
Roll No.: 21013
Date of submission: November 18, 2024 Marks Obtained:

1. **All questions are mandatory. Plagiarism and copying from anywhere (similar submission) can debar you from this course and invite the academic dishonesty policy.**
2. **Implement all algorithms purely in Python without using specialized libraries like OpenCV or PIL for the processing. You may use libraries for basic operations (like loading an image), but the algorithms should be coded from scratch.**
3. **Display both the original and processed images to compare results.**
4. **Make a short 7-minute video and explain your code.**
5. **A report reflecting on what you have learned. Visualization of the output must be there along with other necessary details.**

1. Apply the filters mentioned below on the image attached and analyze their impact. Describe what you found after applying each filter and why certain phenomena occur. ***(Marks: 6)***
   - SIFT
   - Bag of Words
   - HOG

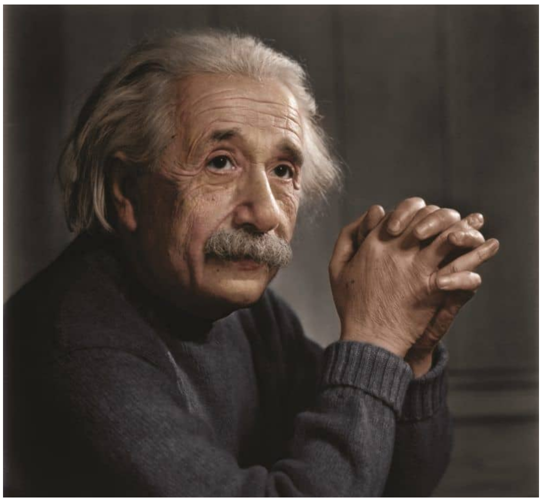

Figure 1: Image of Albert Einstein.

Answer: The codes can be accessed on the ***GitHub***. We first convert the image to gray scale and then apply the mentioned filters.

(i) SIFT (Scale-Invariant Feature Transform) algorithm detects and describes local features in an image, identifying key points (edges, corners, or distinctive areas) and their orientations. Each key point is assigned a descriptor vector to summarize its surrounding area.

Figure 2: SIFT Keypoints - The interest poin

Observations: The SIFT keypoints (illustrated in Figure 2) represent regions in the image with high variations in intensity, such as the contours of Einstein's face and hands.

The algorithm captures features invariant to scale, rotation, and slight illumination changes, making it robust for identifying unique structures. This happens because, SIFT identifies areas with sharp intensity gradients because these regions often contain the most information about the object's structure. The descriptors are normalized to handle changes in brightness, ensuring robustness across different lighting conditions.

(ii) The **Bag of Words** model clusters descriptors from SIFT (or other feature detectors) into a finite set of *visual words*. The output is a histogram showing the frequency of these visual words in the image.

Observations: The provided histogram (depicted in Figure 3) shows a relatively uniform distribution of visual word frequencies with some peaks. Peaks in the histogram correspond to the most frequently occurring patterns (e.g., repetitive textures or prominent features like wrinkles and hair in Einstein's image).

Lower-frequency words represent less common or unique features. This happens because the BoW histogram aggregates SIFT descriptors into clusters, reducing

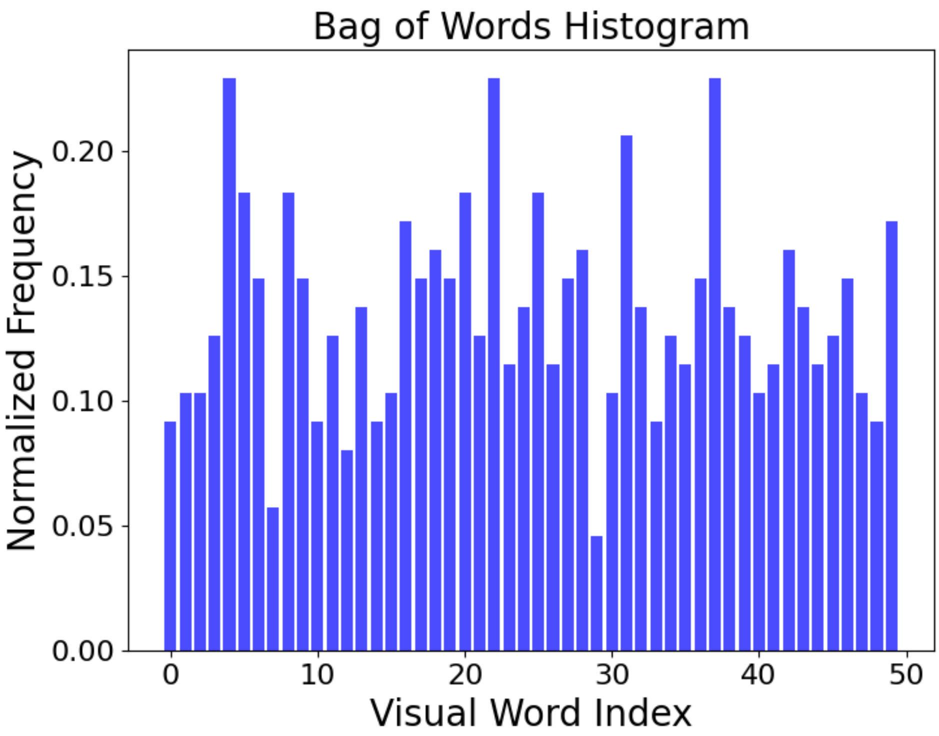

Figure 3: The figure illustrates output is a histogram showing the frequency of these visual words in the image as generated by the Bag of Words.

the image to a compact, structured representation. The diversity in visual word frequencies reflects the range of textures and details present in the image.

(iii) The **HOG (Histogram of Oriented Gradients)** filter focuses on capturing the gradient orientations and their magnitudes to emphasize edge structures in an image.

Observations: The HOG image (4) highlights strong edges and contours, especially around Einstein's face, hair, and hands. Fine textures (like wrinkles or sweater texture) appear less pronounced compared to dominant edges like facial contours. The HOG descriptor would provide numerical values summarizing the directional gradients in the image.

This happens because HOG captures structural patterns by computing gradients in small cells and normalizing them over larger blocks. This ensures consistency under lighting variations. Strong gradients correspond to regions with clear edges or transitions in brightness, making them the most visible in the HOG visualization.

2. Imagine you're monitoring pedestrian movement at a crosswalk. Your task is to track the direction and speed of pedestrians in a video using optical flow analysis. Use OpenCV's built-in video vtest.avi, which simulates real-world pedestrian movement. ***(Marks: 6)***

   - Using the Lucas-Kanade method, track specific points in the video to capture the movement of pedestrians.
   - Visualize the direction of movement using arrows to indicate the flow direction at each point.
   - Provide a brief summary: What patterns do you observe in pedestrian movement? Are there any areas where pedestrians tend to cluster or move faster?

Figure 4: The figure depicts the Histogram of Oriented Gradients (HOG) and shows the strong edges and contours.

Answer: Figure 5 illustrates optical flow with arrows illustrating the direction and magnitude of pedestrian movement at each point. The red arrows represent flow vectors, emphasizing the overall movement patterns throughout the scene.

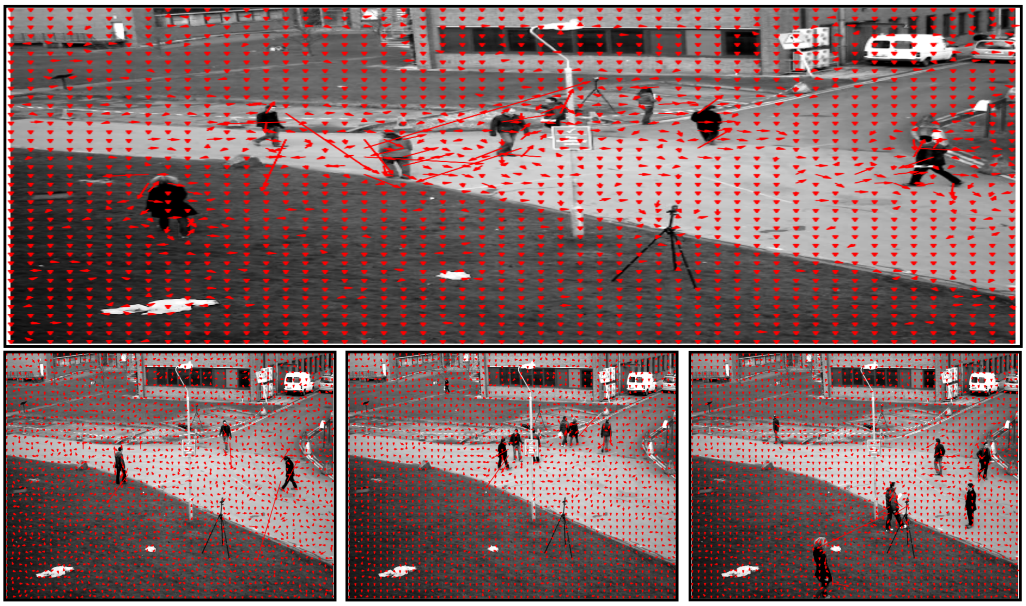

Figure 5: Visualization of optical flow with arrows indicating the direction and magnitude of pedestrian movement at each point. Red arrows represent the flow vectors, highlighting the overall movement patterns across the scene.

The following observations are observed in the pedestrian movement:

1. **Pedestrian Movement Patterns:** The optical flow vectors indicate the general direction and speed of pedestrian movement. Most individuals seem to be moving along defined pathways (sidewalks and open spaces) intended for walking. The directions appear consistent with expected pedestrian traffic, with little deviation in movement paths.
2. **Clustering Areas:** Certain areas, such as intersections and places with multiple pedestrians, show denser optical flow vectors. These likely represent regions where pedestrians naturally cluster or interact (intersection point in the given video). Near objects like streetlights or poles, movement is reduced, indicating pauses or changes in pace.
3. **Faster Movement:** Flow vectors with larger magnitudes suggest areas of quicker pedestrian movement. This is more prominent in open areas and straight paths, where pedestrians have fewer obstacles or interruptions. Slower movement is observed near corners and where clustering occurs.