Indian Institute of Technology Bhubaneswar

Security & Forensics Lab II

Laboratory Experiment No. 4

(Held on 21/01/2022 & 28/01/2022)

Submitted By:

Ayush Modi

21CS06005

M. Tech, CSE (1st year)

Aim

1) Implementation of RESIDUAL-BASED FORENSIC COMPARISON OF VIDEO SEQUENCES. The paper proposes a method to compare two video sequences. It finds out the matching of source cameras and content modification (if exists). P. Mullan, D. Cozzolino, L. Verdoliva and C. Riess, "Residual-based forensic comparison of video sequences," 2017 IEEE International Conference on Image Processing (ICIP), Beijing, 2017, pp. 1507-1511.

Go through the paper and implement the statistical descriptors as mentioned in Section 2.1. Using this statistical descriptor as the feature, find the Mahalanobis distance between the two feature vectors computed from two video sequences. Find out the similarity index as discussed in Section 2.2. Create a dataset as presented in Section 3.1 and Section 3.2 and produce the results shown in Figure 4.

Theory

Computation of Statistical Descriptor –

We use a 1-D high-pass filter h(u) with coefficients [1, −3, 3, −1] along rows, columns, and temporal direction to compute the residuals. To suppress strong edges from image content, residuals are truncated and quantized, such that the residual vector IR(c) at position c, is

$$I_R(c) = \min\left(t,\ \max\left(-t,\ \left\lfloor \frac{(I * h)(c)}{q} \right\rfloor\right)\right) \quad , \qquad (1)$$

where I denote a 1-D slice of the video, h the high-pass filter from above, t the truncation threshold and q the quantization factor. When the intensities are represented as values between 0 and 255, it is reasonable to select small positive integer values for t and q. In our experiments, we set t = 2 and q = 3. This leads to 2t + 1 values. Thus, for n neighboring residuals, there exist (2t + 1) n different combinations of values, which is also referred to as co-occurrence. We chose n = 4 in accordance with other recent forensic works that are based on co-occurrence features. Mirrored co-occurrences like [0, 2, 0, 1] and [1, 0, 2, 0] are considered identical. Thus, the number of individual co- occurrences reduces to 169. We compute histograms of co- occurrence along rows, columns, and frames of the video. These histograms are concatenated to form a 3 x 169 = 507-dimensional statistical block descriptor, on windows of size 128 × 128 pixels, with a step size of 8 pixels.

We use optical flow to compensate for such motion effects, and to increase the likelihood that all feature vectors are either computed on background or on the overlay. More specifically, we estimate per pixel a dense motion field that tracks motion from one frame z − 1 to the next frame z by computing a translation vector (vx, vy) T such that

$$I(x, y, z) = I(x + v_x, y + v_y, z - 1) \ . \qquad (2)$$

Comparison of Two video Sequences –

To compare descriptors from two video subsequences, we select all descriptors from one sequence, and compute their mean μ and co-variance Σ, after passing the feature vectors through a square-root non-linearity. Then, each descriptor of the other sequence is evaluated on these statistics by computing the Mahalanobis distance.

$$d_{\mathrm{Mahal}}(\vec{g}; \vec{\mu}, \Sigma) = \sqrt{(\vec{g} - \vec{\mu})^{\mathrm{T}} \Sigma^{-1} (\vec{g} - \vec{\mu})} \ . \qquad (3)$$

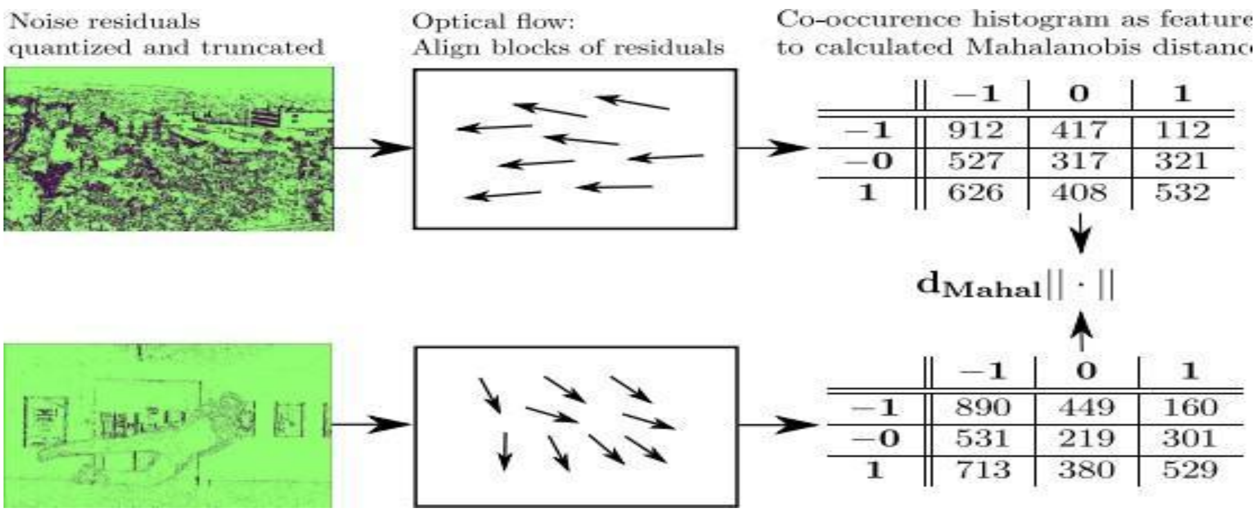

Fig-1: Mahalanobis Distance

Flow chart

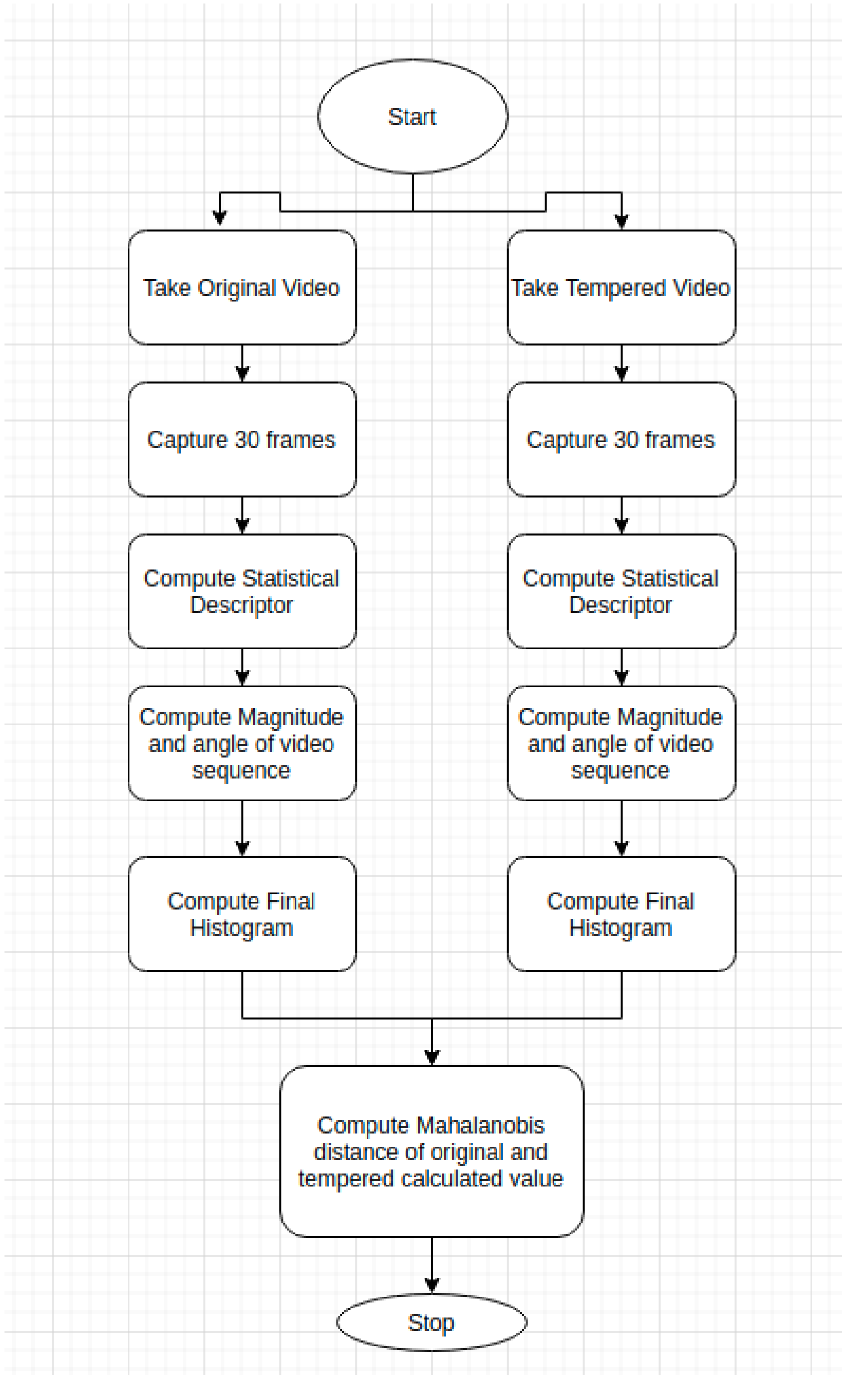

Fig-2: Flow Chart

Results

Datasets after running the program:

In [49]:
```
df=pd.DataFrame(data=np.array(fhist[0]))
df.head()
```

Out[49]:

	0	1	2	3	4	5	6	7	8	9	...	656	657	658	659	660	661	662	663	664	665
0	15996	2	2	0	0	0	0	0	0	0	...	0	80	272	0	0	57	1592	0	0	7
1	15995	3	2	0	0	0	0	0	0	0	...	0	80	201	0	0	57	1443	0	0	7
2	15992	4	4	0	0	0	0	0	0	0	...	0	80	161	0	0	57	1223	0	0	7
3	4	15992	4	0	0	0	0	0	0	0	...	0	80	146	0	0	57	972	0	0	7
4	15996	2	2	0	0	0	0	0	0	0	...	0	79	125	0	0	57	727	0	0	7

5 rows × 666 columns

In [50]:
```
df2=pd.DataFrame(data=np.array(tfhist[0]))
df2.head()
```

Out[50]:

	0	1	2	3	4	5	6	7	8	9	...	656	657	658	659	660	661	662	663	664	665
0	15984	8	8	0	0	0	0	0	0	0	...	0	54	49	0	0	0	0	0	0	0
1	15983	9	8	0	0	0	0	0	0	0	...	0	49	49	0	0	0	0	0	0	0
2	15980	10	10	0	0	0	0	0	0	0	...	0	24	44	0	0	0	0	0	0	0
3	10	15980	10	0	0	0	0	0	0	0	...	0	13	33	0	0	0	0	0	0	0
4	15984	8	8	0	0	0	0	0	0	0	...	0	13	33	0	0	0	0	0	0	0

5 rows × 666 columns

Fig-3: Dataset values

Difference between frame from Original and Tempered video:

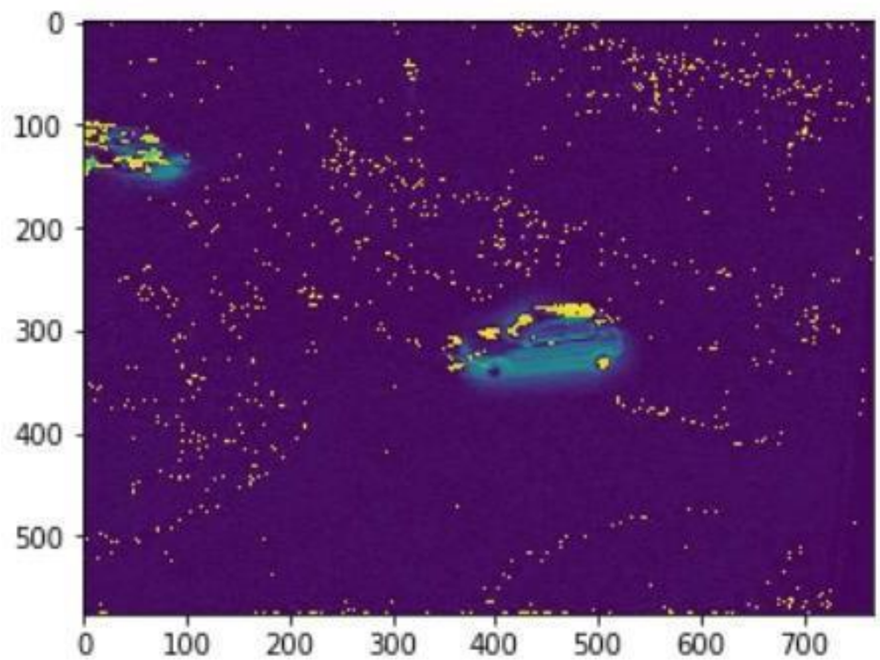

Fig-4: Difference between frame-5

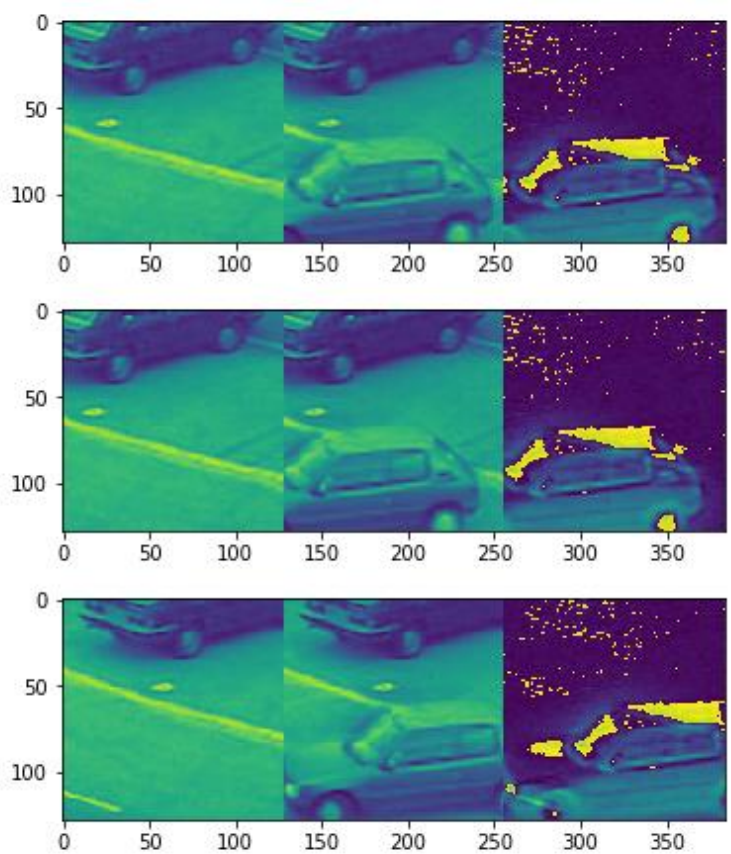

Fig-5: Difference between frames

Conclusion

1) The Optical flow and Histogram of original and tempered video was calculated.
2) Mahalanobis distance was calculated between the original and tempered videos.

Code

Google drive link -
https://drive.google.com/drive/folders/1XeLUCBr-v9Z6MnVbx_g8NDX6Ec1Gfk4E?usp=sharing

Github link to the file –
https://github.com/21CS06005/RESIDUAL-BASED-FORENSIC-COMPARISON-OF-VIDEO-SEQUENCES.git