Detecting video frame rate up conversion based on periodic properties of edge intensity

Vishwajeet Jadhav
IIT Roorkee
vishwajeet_pj@ece.iitr.ac.in
22116106

Vaibhav Prajapati
IIT Roorkee
vaibhav_p@ece.iitr.ac.in
22116102

Parth Badgujar
IIT Roorkee
parth_db@ece.iitr.ac.in
22116065

Pratham Agarwal
IIT Roorkee
pratham_a@ece.iitr.ac.in
22116071

I. Introduction

Frame Rate Up-Conversion (FRUC) is a widely used technique in video processing, aimed at enhancing the smoothness and visual appeal of video playback. By interpolating additional frames between existing ones, FRUC increases the frame rate of a video, allowing for a more fluid viewing experience, particularly on high-refresh-rate displays. While this technique offers significant advantages in terms of visual quality, it also introduces artifacts that may compromise the integrity of the video content.

The ability to detect whether a video has undergone FRUC is crucial in various domains, such as video forensics, copyright infringement detection, and content authenticity verification. For example, identifying FRUC-processed videos can help differentiate between original high-frame-rate content and upconverted material, ensuring compliance with industry standards and preventing deceptive practices

II. Methodology

Upconversion of frame rate involves interpolation which leads to loss of edge intensity. This can be measured using the number of edge-pixels in successive frames. The research paper[1] proposes a method of detecting FRUC with the help of edge intensities calculated using high-pass filtering. There proposed algorithm is as follows :

1) Convert video to consecutive frames I_k, where k is the frame index.
2) Extract luminesce component of each frame to form a new sequence Y_k
3) Extract edge information of luminance sequence Y_k by convolution with a kernel (Kirsch, Prewitt or Sobel)
4) Count the number of edge pixels N_k in each frame. Whether a pixel is an edge pixel or not is decided by a threshold operation on the high pass filter output

Therefore edge intensity of each frame can be given by

$$\rho_k = \frac{N_k}{S}$$

where S is the total number of pixels in the frame.

Original frame rate is then predicted using a Kauffman adaptive moving average (KAMA) based approach.

The algorithm has been tested using Kirsch, Prewitt and Sobel edge detectors as well as a proposed Wavelet transform based edge detection technique. Simulations have been performed on the Foreman sequence and a custom shot video for various frame rates in the range of 15 fps to 60 fps. The prediction results are towards the end of the report .

A. Kirsch Operator for edge intensity

Kirsch operator is a non-linear gradient matrix that takes the derivative of the image intensity level in 8 directions.

Kirsch operator consists of 8 different filters each for each direction :

$$g^{(1)} = \begin{bmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix} \qquad g^{(2)} = \begin{bmatrix} 5 & 5 & -3 \\ 5 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}$$

$$g^{(3)} = \begin{bmatrix} 5 & 0 & -3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{bmatrix} \qquad g^{(4)} = \begin{bmatrix} 5 & 5 & -3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{bmatrix}$$

$$g^{(5)} = \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{bmatrix} \qquad g^{(6)} = \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ 5 & 5 & 5 \end{bmatrix}$$

$$g^{(7)} = \begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix} \qquad g^{(8)} = \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{bmatrix}$$

For each filter $g^{(z)}$ the edge magnitude can be calculated by convoluting the filter with the image.

The final edge intensity can be calculated by :

$$d_A = \sqrt{d_1^2 + d_2^2 + d_3^2 + d_4^2 + d_5^2 + d_6^2 + d_7^2 + d_8^2}$$

The results obtained after calculating edge intensity values for Foreman sequence upconverted video (45 fps) can be seen in figure 1. A regular drop in intensity level can be seen in the case of upconverted video. This is detected using Kaufman adaptive moving average (KAMA) for the edge intensities of the upconverted video.

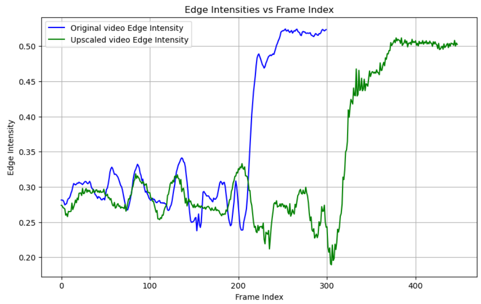

Fig. 1: The comparison of edge-intensity for the original sequence and up-converted sequence.

The output of Kirsch operator followed by binary classification of pixels as edge pixels can be seen in figure 2

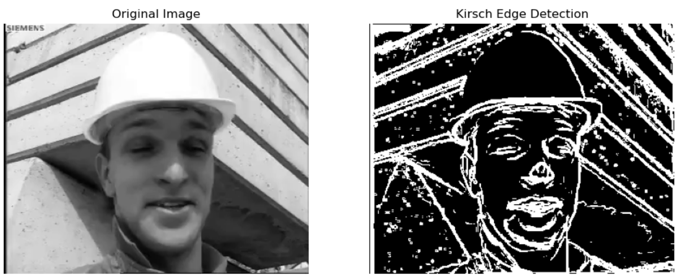

Fig. 2: Edge detection output of Kirsch edge detection

B. Prewitt Operator for Edge Intensity

The Prewitt operator is a first-order derivative-based method for edge detection. It estimates gradients in the horizontal and vertical directions.

The Prewitt operator uses two fixed convolution masks to compute gradients in the x and y directions:

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}, \quad G_y = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$$

The gradient in each direction can be calculated by convoluting these masks with the image intensity matrix.

The edge intensity can then be computed as the magnitude of the gradient:

$$d_A = \sqrt{d_x^2 + d_y^2}$$

Similar to what was done before in case of Kirsch operator, the results for Prewitt operator are as follows :

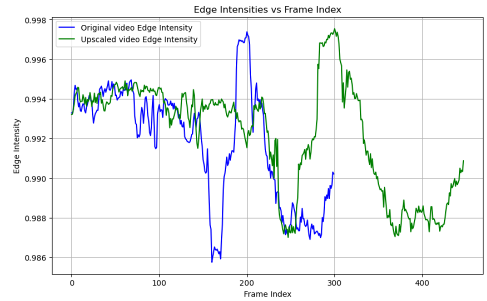

Fig. 3: The comparison of edge intensity for the original sequence and upconverted sequence using Prewitt operator.

The output of the Prewitt operator, showing binary edge classification, is presented in Figure 4. This highlights the detected edges in the upconverted video.

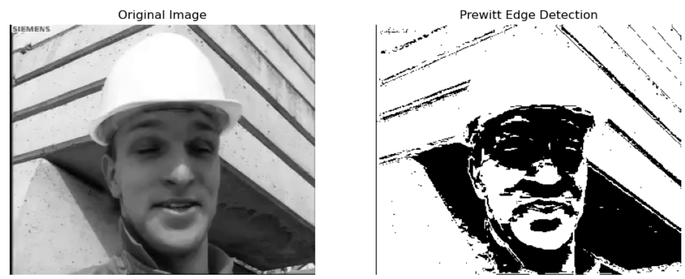

Fig. 4: Edge detection output using Prewitt edge detection.

C. Sobel Operator for Edge Intensity

The Sobel operator is another first-order derivative-based edge detection method, which improves upon the Prewitt operator by adding weights to emphasize pixels closer to the center of the filter. It estimates gradients in the horizontal and vertical directions with higher sensitivity.

The Sobel operator uses two convolution masks to compute gradients in the x and y directions:

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \quad G_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

The gradient in each direction can be calculated by convoluting these masks with the image intensity matrix.

The edge intensity can then be computed as the magnitude of the gradient:

$$d_A = \sqrt{d_x^2 + d_y^2}$$

Similar to the case of the Kirsch and Prewitt operators, the Sobel operator results are shown below:

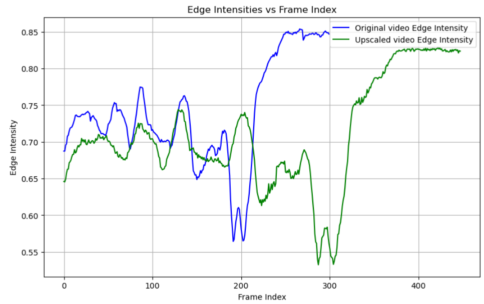

Fig. 5: The comparison of edge intensity for the original sequence and upconverted sequence using Sobel operator.

The output of the Sobel operator, showing binary edge classification, is presented in Figure 6. This demonstrates the edges detected in the upconverted video.

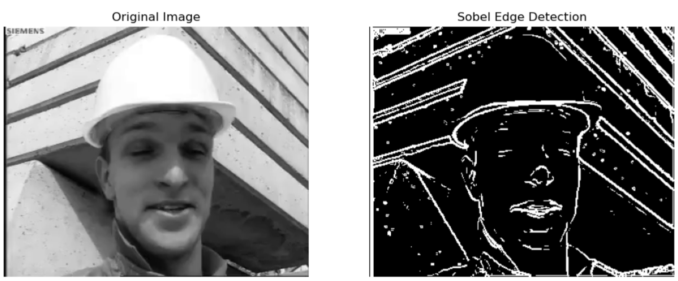

Fig. 6: Edge detection output using Sobel edge detection.

III. Proposed Wavelet Transform approach for edge intensity calculation

The Wavelet Transform is a multi-resolution analysis method that decomposes an image into multiple frequency bands. Unlike Fourier transform-based methods, wavelet transform captures both spatial and frequency information.

In wavelet decomposition, an image is divided into four sub-bands:

- LL (Low-Low): Represents the approximation or low-frequency components, capturing the smooth regions of the image.
- LH (Low-High): Captures horizontal high-frequency components, highlighting vertical edges.
- HL (High-Low): Captures vertical high-frequency components, highlighting horizontal edges.
- HH (High-High): Contains diagonal high-frequency components, capturing finer edge details and noise.

High-frequency components in the LH, HL, and HH sub-bands account for edge information. These are then used to extract edge-intensity level [2].

The edge intensity for an image can be computed by combining the high-frequency sub-bands as follows:

$$d_A = \sqrt{(d_{LH})^2 + (d_{HL})^2 + (d_{HH})^2}$$

The results obtained using the Wavelet Transform for the Foreman sequence upconverted video (45 fps) are shown below:

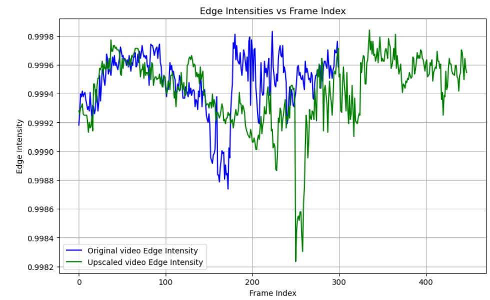

Fig. 7: The comparison of edge intensity for the original sequence and upconverted sequence using Wavelet Transform.

The edge detection output using the Wavelet Transform, showing binary classification of edge pixels, is presented in Figure 8. This highlights the detected edges in the upconverted video.

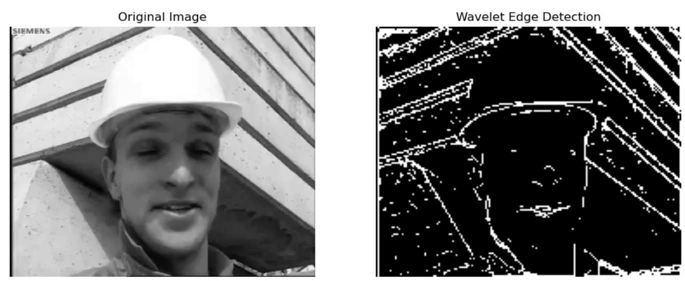

Fig. 8: Edge detection output using Wavelet Transform.

IV. Original FPS prediction results

The frame-rate predictions using all the aforementioned algorithms is summarised in the tables III below. All video frame rate conversions were done using FFMPEG. The expected frame-rate for the original video is 30 fps for both cases.

TABLE I: Predicted Original FPS for Different FRUC Rates Using Various Techniques for Foreman sequence

FRUC Rate	Prewitt (fps)	Kirsch (fps)	Sobel (fps)	Wavelet (fps)
60	27.24	32.75	37.06	31.65
55	25.44	30.15	32.85	31.85
50	26.05	27.55	30.76	29.95
45	21.45	24.95	26.56	23.05
40	19.14	22.76	24.76	23.06
35	16.14	19.85	21.67	21.26
25	11.40	15.00	15.10	14.60
20	8.60	13.00	12.50	11.90
15	6.90	9.50	9.80	7.40

TABLE II: Predicted Original FPS for Different FRUC Rates Using Various Techniques for Room video

FRUC Rate	Prewitt (fps)	Kirsch (fps)	Sobel (fps)	Wavelet (fps)
60	25.40	29.30	22.53	27.63
55	23.74	26.34	20.22	24.67
50	20.22	23.00	17.07	22.54
45	21.25	20.87	14.94	20.50
40	16.70	18.46	12.90	18.19
35	14.67	15.97	11.33	16.62
25	11.02	10.83	7.96	11.67
20	10.00	8.98	5.09	8.79
15	6.61	6.50	4.77	7.00

V. Conclusion

It may be observed that the proposed algorithm has worked successfully only in the cases where upconverted frame rate was sufficiently above the original frame rate. It has also failed in predicting the original frame rate in the cases where the frame rate was reduced. This may be attributed to the fact that downsampling of frames occurs in this process which has not been accounted for. Amongst the four techniques tested, Kirsch operator and wavelet transform have been able to predict the original frame rate more accurately than Prewitt and Sobel detectors. Additionally, camera panning, which was more prominant in the second video than the Foreman sequence poses as a hindrance to fps prediction.

VI. References

References

[1] Y. Yao, G. Yang, X. Sun, and L. Li, “Detecting video frame-rate up-conversion based on periodic properties of edge-intensity,” *Journal of Information Security and Applications*, vol. 26, pp. 39–50, 2016, ISSN: 2214-2126. DOI: 10.1016/j.jisa.2015.12.001. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2214212615000691.

[2] K. Kumar, N. Mustafa, J. .-. Li, R. A. Shaikh, S. A. Khan, and A. Khan, “Image edge detection scheme using wavelet transform,” in *2014 11th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, Chengdu, China, 2014, pp. 261–265. DOI: 10.1109/ICCWAMTIP.2014.7073404.