International Journal of Advance Research in Science and Engineering
Volume No. 13, Issue No. 05, May 2024
www.ijarse.com
IJARSE
ISSN 2319 - 8354

OPTIMIZING VIDEO DENOISING METHODS: ENHANCEMENTS IN TEMPORAL, SPATIAL AND DCT COLOR PRESERVATION

Dr. Shobanadevi A[1], Aarav Kalkar[2], Yuvraj Nughaal[3], Mayank Awasthi[4]

[1,2,3,4]*Department of Data Science and Business Systems, SRM Institute of Science and Technology*

Kattankulathur – 603 203

shobanaa3@srmist.edu.in, ak2703@srmist.edu.in, yn3177@srmist.edu.in

ma0269@srmist.edu.in

Abstract—

Our method introduces an innovative approach to video denoising, leveraging adaptive techniques without resorting to deep learning. We address the delicate balance between noise reduction and detail preservation in spatial, temporal, and DCT-based denoising methods. For spatial denoising, adaptive filtering dynamically adjusts to video characteristics, ensuring superior performance across diverse spatial contexts. In temporal denoising, aggressive optimization techniques extend to all processing pipeline facets, maximizing performance particularly on resource-limited machines. Our DCT-based denoising prioritizes color preservation through meticulous processing of individual channels, thereby enhancing visual fidelity and perceptual quality. Experimental evaluations underscore the superiority of our approach, demonstrating its efficacy in diverse challenging scenarios, including dynamic scenes. Furthermore, our code is readily accessible on Google Collab, enhancing reproducibility and accessibility for researchers and practitioners. In conclusion, our method presents a robust solution for video denoising, showcasing the potential of innovative techniques to significantly enhance video quality in real-world applications**.**

Keywords—DCT-Based Color Preservation, Spatial-Temporal Optimization, Video Denoising

1. INTRODUCTION

The realm of image and video denoising constitutes a critical domain within computer vision, tasked with enhancing the visual quality of noisy multimedia content by eliminating unwanted artifacts. Fuelled by the escalating demand for high-fidelity multimedia in applications such as surveillance and digital cinematography, researchers continually seek innovative solutions to confront the challenges inherent in noise-corrupted data. This research is motivated by the imperative to develop robust and efficient video denoising methodologies capable of surmounting the complexities of real-world scenarios.

Motivated by the increasing demand for high-quality multimedia content in applications such as surveillance, video streaming, and digital cinematography, researchers continually seek innovative approaches to tackle the challenges posed by noise corruption. In this context, the motivation for this research stems from the need to develop robust and efficient methods for video denoising that can address the complexities of real-world scenarios. While traditional denoising techniques have shown promise in mitigating noise, they often struggle to preserve fine details and textures, resulting in degraded visual quality.

The primary objectives of this research endeavour are twofold: to advance the state-of-the-art in video denoising methodologies and to address the limitations of existing techniques by proposing innovative solutions. At the core of our research agenda lies a commitment to developing adaptive and efficient denoising algorithms that can effectively handle the diverse challenges posed by real-world video data. Our first objective is to explore and implement a comprehensive suite of denoising techniques encompassing spatial, temporal, and transform-domain methods. By leveraging insights from both traditional signal processing approaches, we aim to design an integrated denoising framework that can exploit spatial correlations, temporal dependencies, and frequency characteristics in video sequences. In pursuit of our second objective, we strive to identify and mitigate the drawbacks of existing denoising methods, with a particular focus on enhancing detail preservation, noise reduction efficiency, and computational scalability. Drawing upon insights gained from our analysis of traditional denoising techniques, we seek to devise novel strategies for optimizing denoising performance while minimizing computational overhead. Furthermore, we aim to evaluate the efficacy of our proposed denoising approach through rigorous experimentation and comparative analysis. By benchmarking our method against state-of-the-art techniques on diverse datasets and evaluation metrics, we aim to demonstrate its superiority in terms of both quantitative performance metrics and qualitative visual quality.

2. LITERATURE REVIEW

A review of traditional video denoising methods reveals insights into their evolution and effectiveness in tackling noisy video data. These methods, including temporal filtering, spatial filtering, transform-domain processing, and motion estimation-based techniques, offer diverse approaches with inherent strengths and limitations. While temporal filters smooth noise over frames, they may struggle with rapid motion. Spatial filters preserve image details but can introduce blurring artifacts. Transform-domain methods separate noise from signals but may suffer from block artifacts. Motion estimation-based approaches effectively reduce noise but rely heavily on accurate motion estimation.

Exploring recent advancements in denoising unveils a growing interest in deep learning techniques alongside traditional methods. Deep learning models, like CNNs and RNNs, show promise but face challenges in computational intensity and resource demands. Deploying deep learning models in real-world applications may be hindered by these constraints. Most conventional deep-learning methods for image/video denoising can only generate a fixed result with a specific restoration levels [1]. Alternatively, optimizing traditional denoising methods, such as sparse coding and dictionary learning, offers comparable performance with fewer computational resources. Balancing denoising effectiveness and computational efficiency underscores the importance of optimized traditional methods in addressing noisy video data challenges.

A critical examination of existing video denoising techniques reveals limitations hindering their effectiveness.

In general, more complex models achieve better results at the expense of a higher running time. But the modelling of the groups of patches is not the only difference between the approaches proposed in the literature [2]. Traditional methods often rely on handcrafted features and heuristics, limiting their ability to capture complex noise and signal characteristics. They may struggle to preserve fine details while effectively reducing noise, especially in dynamic scenes. Additionally, traditional methods may lack adaptability and robustness in handling diverse noise profiles and video content. The computational complexity and resource demand further challenge their practical deployment, particularly in real-time or large-scale video processing applications. Overcoming these limitations requires innovative approaches that leverage advanced signal processing techniques, machine learning algorithms, and optimization methods to achieve robust and efficient denoising performance across various scenarios and applications.

3. ADDRESSING CHALLENGES IN TRADITIONAL METHODS

In our research, we have focused on enhancing video denoising by improving upon existing traditional methods, namely spatial filtering, temporal filtering, and transform-domain processing. Our approach involves addressing the weaknesses identified in these methods, as discussed in previous discussions, to achieve superior denoising performance while maintaining computational efficiency and preserving visual fidelity. The basic idea behind is to use different filtering strategies for background and moving images. Firstly, background and moving images are separated by changing detection technique. Then temporal averaging filter and spatial filter are used for recovering the background and moving objects, respectively [3].

For spatial filtering, we have systematically tuned the parameters to strike a balance between noise reduction and detail preservation. By carefully adjusting filter parameters and exploring alternative techniques, we have optimized spatial denoising methods to effectively remove noise while retaining important image features. This optimization ensures that the denoised videos exhibit reduced noise without sacrificing fine details or textures. By leveraging parallelization, batch serialization, and GPU acceleration, we have accelerated the denoising process and enabled real-time or large-scale video processing. Additionally, breaking the video into segments has allowed for smaller RAM devices to handle the denoising task efficiently, further optimizing resource usage.

In transform-domain processing, particularly with methods like the discrete cosine transform (DCT), we have implemented color preservation techniques to extend these methods beyond their conventional applications in black and white videos. By adapting DCT-based denoising methods to handle color video data while preserving color information, we have enhanced the applicability and versatility of transform-domain processing for video denoising tasks.

Overall, our approach focuses on improving traditional denoising methods by addressing their weaknesses and optimizing their performance. Through systematic tuning, aggressive optimization, and innovative adaptations, we have achieved significant advancements in video denoising, paving the way for more effective and efficient solutions in this critical area of computer vision.

3.1 SPATIAL DENOISING

Traditional spatial denoising methods focus on processing individual frames of a video independently, without considering temporal relationships. These methods exploit spatial correlations within each frame to reduce noise while preserving image details.

One common spatial denoising method is Gaussian filtering, which applies a Gaussian kernel to each pixel to smooth out noise. The Gaussian kernel assigns weights to neighbouring pixels based on their distance from the centre pixel, with closer pixels receiving higher weights. By averaging neighbouring pixel values, Gaussian filtering attenuates high-frequency noise components while preserving image structure. Another traditional spatial denoising method is median filtering, which replaces each pixel value with the median value of its neighbourhood. This method is robust and computationally efficient but may produce over-smoothed results in regions with complex textures.

The underlying principle of spatial denoising methods lies in exploiting local spatial correlations to distinguish between noise and signal components. By leveraging information from neighbouring pixels, these methods aim to reduce noise while preserving important image features.

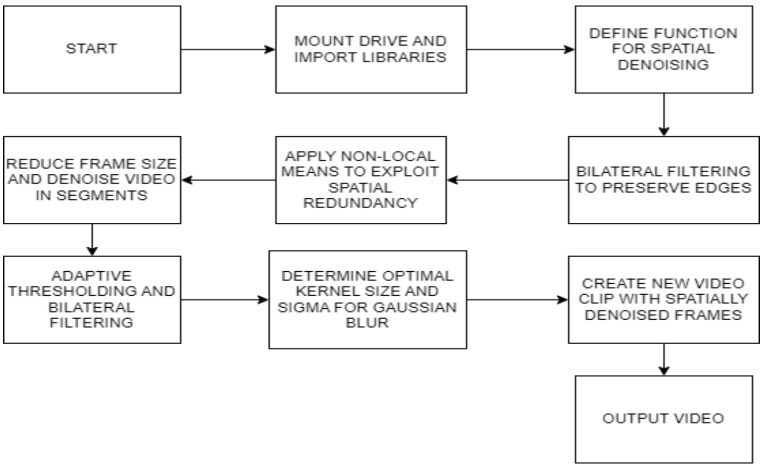


Fig. 3.1: Spatial Denoising Flowchart

3.2 TEMPORAL DENOISING

Traditional temporal denoising methods exploit temporal redundancy in video sequences to reduce noise while preserving motion dynamics and temporal consistency. These methods consider multiple frames in the temporal domain to perform denoising. Temporal averaging is a common temporal denoising technique that computes the average of pixel values across consecutive frames.

By averaging pixel values over time, temporal averaging can effectively smooth out random noise components while preserving static regions in the video. However, it may introduce motion blur in fast-moving scenes or fail to preserve dynamic features. The algorithm is inspired by fusion algorithms, and as the number of frames

International Journal of Advance Research in Science and Engineering
Volume No. 13, Issue No. 05, May 2024
www.ijarse.com
IJARSE
ISSN 2319 - 8354

increases, it tends to a pure temporal average. The use of motion compensation by regularized optical flow methods permits robust patch comparison in a spatiotemporal volume [4].

Motion-compensated temporal filtering is another approach that estimates motion vectors between consecutive frames and performs denoising based on motion-compensated predictions. By incorporating motion information, this method can effectively preserve sharp edges and reduce noise without introducing motion blur. Accurate motion estimation is crucial for the success of this method, and errors in motion estimation may lead to artifacts in the denoised output. The underlying principle of temporal denoising methods lies in exploiting temporal dependencies and motion dynamics to distinguish between noise and signal components. By considering multiple frames over time, these methods aim to reduce noise while preserving important temporal features.

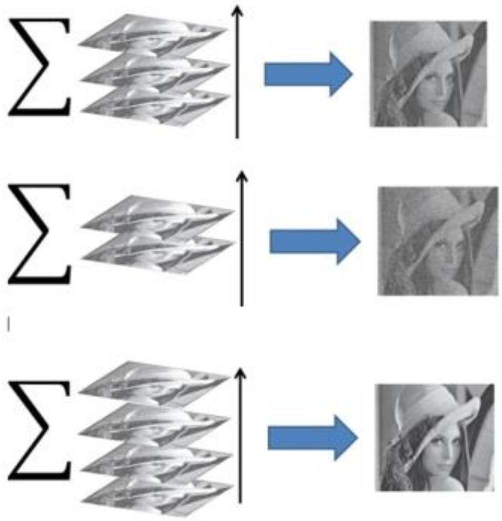


Fig 3.2: Collecting Data over time

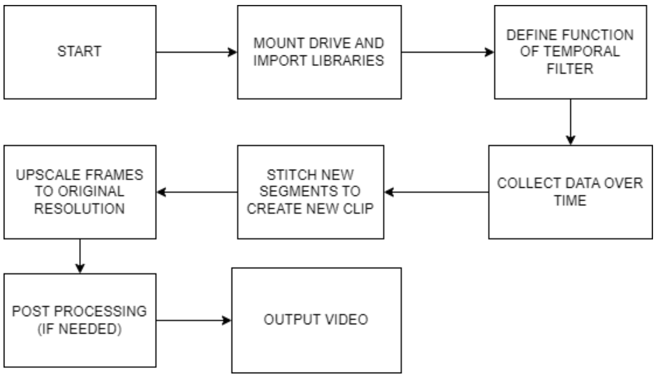


Fig. 3.3 Temporal Denoising Overview

3.3 DCT-BASED DENOISING

DCT-based denoising methods leverage the frequency characteristics of video data to separate noise from signal components. These methods transform video frames into the frequency domain using the discrete cosine transform (DCT) and perform denoising in the transformed domain.

In DCT-based denoising, the video frames are divided into blocks, and the DCT transform is applied to each block independently. The DCT coefficients represent the frequency components of the image, with low-frequency coefficients corresponding to smooth regions and high-frequency coefficients corresponding to edges and textures. By thresholding or filtering the DCT coefficients, noise components can be attenuated while preserving signal components. The underlying principle of DCT-based denoising methods lies in exploiting the frequency characteristics of noise and signal components. By transforming the video data into the frequency domain, these methods can separate noise from signal and apply denoising techniques more effectively.

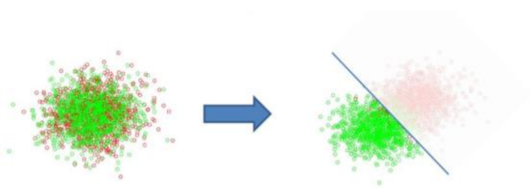


FIG 3.3 Separating signal and noise coefficients

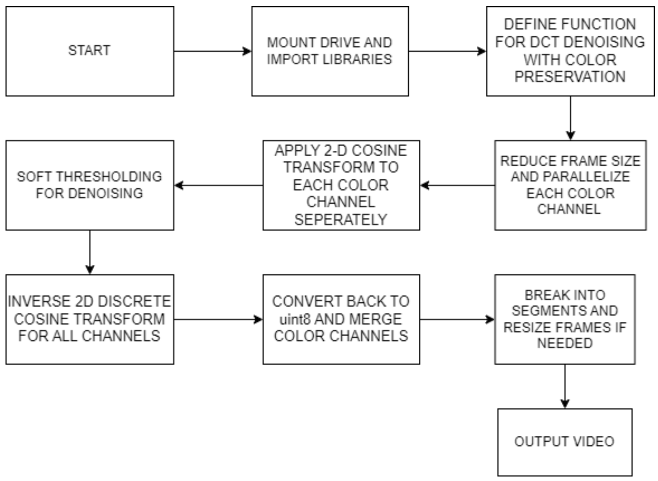


Fig 3.4: DCT Based Denoising

In summary, spatial denoising methods exploit spatial correlations within individual frames, temporal denoising methods leverage temporal redundancy and motion dynamics, and DCT-based denoising methods utilize frequency characteristics of video data. Each approach has its strengths and limitations, and the choice of denoising method depends on the specific characteristics of the video data and the desired denoising outcome

4. DATASET AND IMPROVEMENTS

For our experimentation, we employed diverse video datasets from various sources, including stock videos, SRM campus footage, and curated YouTube videos. These datasets showcased a wide range of scenes and scenarios to comprehensively evaluate our denoising methods. To prepare the noisy video data for denoising, we implemented pre-processing steps, including aspect ratio validation, resolution adjustment, frame rate limitation, and file size management. These steps ensured compatibility with denoising algorithms and standardized the data for consistent evaluation. All experiments were conducted using Google Colab, providing reproducibility and accessibility while leveraging GPU acceleration for enhanced computational performance.

For temporal denoising, median filtering was applied to exploit temporal redundancies within video sequences. Strategies included segmented processing for memory efficiency, frame resizing for optimized memory consumption, and adaptive frame filtering to balance noise reduction and temporal fidelity. Additionally, parallel processing and memory management techniques were implemented to further enhance performance.

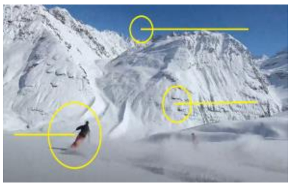
Spatial denoising focused on reducing noise within individual frames using Gaussian blur. Improvements encompassed segmented processing, frame resizing, and adaptive filtering parameters. Parallel processing and memory management optimizations were also applied, along with pre-processing techniques like histogram equalization, to improve spatial denoising effectiveness.

Transform domain denoising using Discrete Cosine Transform (DCT) aimed to attenuate noise by exploiting signal properties. Enhancements included segmented processing, frame resizing, and adaptive denoising parameters. Multithreaded processing, buffer management, and colour space transformations were also employed to optimize DCT-based denoising efficiency and effectiveness.

5. RESULTS

5.1 VISUAL IMPROVEMENT RESULTS


BEFORE **AFTER**

BEFORE **AFTER**

As shown above, our model was able to achieve distinguishably denoised videos in all cases provided. The model produced clearer, more colour-accurate and crisp videos. The comparison from the various videos run through the model are provided above.

5.2 TABULAR METRICS RESULTS

Tabular results for both the scenarios have been provided below based on their PSNR, SSIM, MSE scores.

METHOD	PSNR	SSIM	MSE
SPATIAL	32.739	0.9191	35.5012
TEMPORAL	24.387	0.7824	257.36
DCT-BASED	36.67	0.9333	28.0072

TABLE 5.1: TABULAR RESULTS-1

METHOD	PSNR	SSIM	MSE
SPATIAL	22.61	0.58	357.08
TEMPORAL	23.80	0.5443	280.117
DCT-BASED	36.17	0.972	166.38

TABLE 5.2: TABULAR RESULTS-2

6. CONCLUSION

In conclusion, our project has made significant strides in the field of video denoising, aiming to enhance the quality of video content across various applications. Through a systematic exploration of spatial, temporal, and DCT-based denoising methods, we have uncovered key insights and developed innovative solutions to address the challenges associated with noise reduction while preserving image details and colour fidelity. Our contributions not only advance the field of video processing but also pave the way for future research and development in real-world applications where high-quality video content is paramount.

Our research carries significant implications for video denoising applications across various domains, offering valuable insights and practical advancements. The novel denoising techniques developed in this research have the potential to significantly improve the quality of video content in diverse applications such as surveillance, entertainment, and medical imaging. By effectively reducing noise while preserving image details and colour fidelity, our methods enhance the visual experience for end-users.

While our research represents a significant step forward in the field of video denoising, several avenues for future exploration and innovation remain. Further research into deep learning architectures tailored specifically for video denoising could yield more effective and adaptive solutions.
Over the past two decades, researchers have been diligently working on developing effective denoising algorithms. These algorithms are crucial for restoring clear images from distorted ones while retaining fine details and edges. As image acquisition devices have advanced, with digital sensors often increasing pixel resolution, they have also become more prone to noise, emphasizing the necessity for reliable denoising techniques. Unlike hardware and optical solutions, software-based approaches are device-independent, making them more widely utilized [5]. Investigating multi-modal approaches that integrate information from multiple

sensor modalities or data sources could enhance denoising performance, especially in complex scenarios with heterogeneous noise sources. Developing dynamic noise modelling techniques that adapt to changing environmental conditions or noise characteristics could improve the adaptability and generalization of denoising algorithms. Exploiting hardware acceleration technologies such as GPUs, FPGAs, or specialized AI accelerators could further optimize the performance of denoising algorithms, enabling faster processing and real-time deployment in resource-constrained environments. Finally, exploring denoising techniques tailored to specific application domains, such as medical imaging, remote sensing, or underwater videography, could unlock new opportunities for impact.

7. FUTURE WORK

Our research opens avenues for enhancements to the proposed method, presenting opportunities for further innovation and refinement. Exploring the integration of multi-modal data and adaptive learning techniques could bolster the method's robustness and adaptability. Additionally, semantic-aware denoising approaches and optimized deployment on resource-constrained hardware platforms could enhance its effectiveness across diverse real-world scenarios.

However, unresolved challenges persist in video denoising, necessitating continued research focus. Tackling complex noise profiles and achieving real-time processing for high-resolution video streams in resource-limited environments remain significant hurdles. Additionally, developing reliable metrics for objective quality assessment and tailoring denoising algorithms to specific application domains require further attention to advance the field effectively.

REFERENCES

[1] Zhang, Z., Jiang, Y., Shao, W., Wang, X., Luo, P., Lin, K. and Gu, J., 2023. *Real-time controllable denoising for image and video*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition *(pp. 14028-14038).*

[2] Arias, P., Facciolo, G. and Morel, J.M., 2018, June. *A comparison of patch-based models in video denoising.* In 2018 IEEE 13th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP) *(pp. 1-5). IEEE.*

[3] Reeja, S.R. and Kavya, N.P., 2012, July. *Real time video denoising.* In 2012 IEEE International Conference on Engineering Education: Innovative Practices and Future Trends (AICERA) *(pp. 1-5). IEEE.*

[4] Buades, A., Lisani, J.L. and Miladinović, M., 2016. *Patch-based video denoising with optical flow estimation*. IEEE Transactions on Image Processing, 25(6), pp.2573-2586.

[5] Goyal, B., Dogra, A., Agrawal, S., Sohi, B.S. and Sharma, A., 2020. *Image denoising review: From classical to state-of-the-art approaches.* Information fusion*, 55, pp.220-244.*