Journal of Statistics & Management Systems
ISSN 0972-0510 (Print), ISSN 2169-0014 (Online)
Vol. 27 (2024), No. 2, pp. 303–313
DOI : 10.47974/JSMS-1256

Analysis of different noise filtering techniques for object detection and tracking from video with varying illumination

Praful V. Barekar *
Kavita R. Singh †
Department of Computer Technology
Yeshwantrao Chavan College of Engineering
Nagpur 441110
Maharashtra
India

Abstract

In order to improve object tracking and detection in films with different lighting conditions, this paper provides a thorough review of noise filtering methods. Computer vision systems face difficulties due to the variety of illumination conditions found in real-world settings, which calls for effective noise reduction techniques. We explore and evaluate several filtering techniques, such as deep learning-based methods, adaptive algorithms, guided bilateral filter and spatial and temporal filters, to determine how well they preserve object characteristics while cutting down on noise. We simulate real-world conditions with our studies on video datasets that have dynamic illumination changes. The outcomes show how well guided bilateral filter methods work in difficult illumination situations to preserve precise object trajectories and boundaries. To help practitioners choose the best noise filtering algorithms based on application-specific needs, we also go over the computational consequences and trade-offs related to each strategy. This work addresses a crucial component of real-world object detection and tracking in movies with variable lighting, offering insightful new information to the field of computer vision.

Subject Classification: *68T45.*

Keywords: *Object detection, Noise filtering, Guided bilateral filter, CNN, Illumination.*

1. Introduction

Object detection and tracking in movies is the foundation of many computer vision applications, from human-computer interaction to

* *E-mail:* **praful.barekar20@gmail.com (Corresponding Author)**
† *E-mail:* **singhkavita19@yahoo.co.in**

autonomous cars and surveillance. But noise is always there to obstruct these processes' effectiveness, particularly in situations where lighting conditions change suddenly. The problem is figuring out what information is important from noise, which is especially noticeable in video frames where illumination variations can cause objects to appear blurry or misaligned [1], [2].Various noise filtering algorithms have been developed to tackle this problem, each providing a different method to improve the accuracy of Object Detection and Tracking under different lighting conditions. By reducing high-frequency noise and smoothing pictures, traditional techniques like median and Gaussian filtering function as fundamental tools. These methods, however, can unintentionally lose important object information and limits.Adaptive filters, like the Kalman filter, provide robustness against noise-induced uncertainty by dynamically predicting and updating object positions over time, beyond traditional approaches [3]. However, in the presence of non-linarites or sudden variations in illumination, their performance may be affected.

Bilateral filtering becomes a potent tool in the quest for more complex noise reduction, with the guided bilateral filter at the forefront. The [4] guided bilateral filter performs exceptionally well in maintaining edge information while minimizing noise. It is a strong option for improving object detection and tracking in video frames because of its flexibility to local image properties, especially when there is fluctuating illumination.

2. Different Noise filtering Techniques

2.1 *Gaussian Filtering*

One popular and traditional technique for reducing noise in photos is Gaussian filtering. Using a Gaussian kernel convolution process, it emphasizes the core pixels while progressively reducing the impact of the surrounding pixels [11] [13]. By using this method, high-frequency noise in the image is efficiently reduced and smoothed out. Although Gaussian filtering is easy to use and computationally fast, it tends to blur small details and edges, which makes it less appropriate for applications where maintaining these features is essential.

$$Output(x,y) = \sum i \sum j I(i,j) \cdot G(x-i, y-j) \tag{1}$$

The Gaussian kernel G is defined as:

$$G(x,y) = \frac{1}{2\pi\sigma 2} e^{-\frac{x2+y2}{2\sigma 2}} \tag{2}$$

2.2 *Guided Bilateral Filter*

An extra image is added as guidance in the guided bilateral filter, which is an expansion of the conventional bilateral filter. Because of its great adaptability, this filter successfully reduces noise while preserving edges [5]. The guided bilateral filter's mathematical structure incorporates spatially changing coefficients, which enable it to adapt dynamically to the specific characteristics of the local image. Because of its versatility, it is especially helpful for jobs involving object detection and tracking when details need to be preserved, especially in situations when lighting conditions change.

$$Output(p) = A(p)I(p) + B(p)Guide(p) \tag{3}$$

The local similarity between I and Guide is used to calculate the coefficients A(p) and B(p).

2.3 *Kalman Filter*

A popular dynamic, recursive technique for tracking moving objects over time is the Kalman filter. By fusing predicted and actual values, it estimates a system's state and is noise-resistant. The Kalman filter is used in object identification and tracking to forecast an object's future state based on its past state and to modify that prediction in response to fresh observations [6]. The Kalman filter works well for tracking, but it can perform poorly when there are non-linearities or abrupt changes in light.

2.4 *Adaptive Bilateral Filter*

An adaptation of the conventional bilateral filter, the adaptive bilateral filter modifies filter parameters according to the specifics of the local image. Its capacity to adapt improves its performance in dynamic scenes by enabling it to reduce noise and better maintain edges. The [7] adaptive bilateral filter is used in object detection and tracking to preserve accuracy under different lighting conditions.

$$Output(p) = \frac{1}{W(p)} \sum qI(q) \cdot \omega(p,q) \cdot \phi(p,q) \tag{4}$$

Where, $W(p)$ represents the normalization factor, $\omega(p,q)$ represents the spatial weight, and $\phi(p,q)$ represents the intensity weight. The particular configuration of these weights is determined by the adaptive rules utilized.

3. Object Detection and Tracking from Video

Guided bilateral filtering preprocesses the input frames to improve pertinent characteristics and minimize noise before Convolutional Neural Networks (CNNs) are integrated with them for object detection and tracking in videos.

3.1 *GuidedBilateral Filtering:*

The bilateral filter is applied to the input image I to obtain a denoised and smoothed version:

$$Bilateral(I,\sigma_s,\sigma_r)=\frac{1}{W(p)\Sigma}I(q)\cdot\omega_{s(p,q)}\cdot\omega_{r(I(p),I(q))} \tag{5}$$

where ,W(p) is the normalization factor, ω_s is the spatial weight, and ω_r is the range weight. Parameters σ_s and σ_r control the spatial and intensity domain smoothing, respectively.

$$Output(p)=A(p)I(p)+B(p)Guide(p)$$

3.2 *CNN Integration:*

The output of the bilateral filter is then fed into a CNN for object detection and tracking. The CNN consists of convolutional layers, pooling layers, and fully connected layers. Let CNN(Bilateral(I)) represent the CNN model applied to the bilateral-filtered image.

3.3 *Mathematical Model*

The overall mathematical model combining bilateral filtering and CNN can be expressed as:

$$Output(p)=CNN(Bilateral(I)) \tag{6}$$

In more detail:

$$Output(p)=CNN\left(\frac{1}{W(p)\Sigma}I(q)\cdot\omega_{s(p,q)}\cdot\omega_{r(I(p),I(q))}\right) \tag{7}$$

The CNN is responsible for learning and extracting high-level features from the enhanced image obtained through bilateral filtering. The bilateral filter serves as a preprocessing step to reduce noise while preserving important details, providing the CNN with a cleaner input for more effective learning.

4. Proposed Method

4.1 *Background Modeling*

In computer vision, background modeling is an essential step for detecting object motion from video frames. In [8] order to establish a baseline or background, this technique entails modeling the aspects in a scene that are static or change gradually. Then, by comparing subsequent frames to this model, regions displaying motion are recognized and isolated, emphasizing possible objects of interest. Statistical models, like Gaussian Mixture Models (GMM), and non-parametric approaches, like kernel density estimation, which adaptively represent pixel intensities across time, are examples of common methodologies.

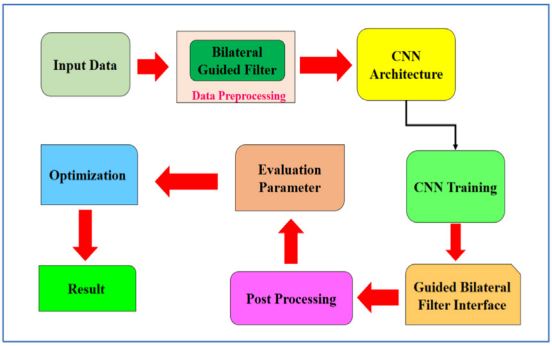

Figure 1

Workflow of Proposed model

4.1.1 Temporal Image Analysis:

The study and interpretation of variations in pixel values across successive frames is known as temporal image analysis. It is feasible to recognize and track moving objects, evaluate motion patterns and spot anomalies by examining the temporal dimension of video data. In temporal image analysis, methods like frame differencing and optical flow are frequently used to measure and visualize motion between frames as shown in Figure 1.

4.2 *Shadow Removal*

Accurate object tracking and detection might be hampered by shadows. Finding the difference between real object features and their shadow counterparts is the goal of shadow removal techniques. These techniques frequently entail modeling the lighting and distinguishing between moving and static shadows [10]. Reducing the influence of shadows enhances the precision of later processing stages, including classifying objects.

4.3 *Foreground Modeling and Reconstruction*

In foreground modeling, the areas of a video frame that differ from the background model are identified and represented. To locate and follow things of interest, this foreground data is essential. Reconstruction techniques can be used to improve the foreground representation, which will help with the precise delineation of item shapes and borders. Morphological operations and contour-based techniques are popular techniques.

4.4 *Classification*

Classification is the process of giving an object a label or category based on its attributes. Classification is used in the context of object detection and tracking to locate and group things in the foreground areas [9]. Robust object classification is a common application for machine learning methods, including deep learning models such as Convolutional Neural Networks (CNNs). The object identification and tracking system's overall efficacy is greatly influenced by the accuracy of the categorization stage.

5. Result and Discussion

A thorough comparison of four different filtering methodsGuided Bilateral Filter, Adaptive Bilateral Filter, Kalman Filter, and Gaussian Filtering across several performance metrics is shown in Table 1.

Table 1

Performance of each filtering technique

Evaluation parameter	Gaussian Filtering	Guided Bilateral Filter	Kalman Filter	Adaptive Bilateral Filter
Noise Reduction	3.5	4.8	4.5	4.7
Egde Preservation	2.7	4.9	2.81	4.63
Adaptability Illumination Changes	3.1	3.1	3.2	4.51

Contd...

Computational Complexity	2.23	3.6	3.5	3.3
Real Time Tracking	4.2	4.3	4.4	4.1
Handling Non Linearities	1.9	4.6	4.42	4.2
Application	3.8	4.2	4.2	4.8
Dependency on parameteres	2.3	3.7	4.5	3.9

These parameters include Dependency on Parameters, Noise Reduction, Edge Preservation, Adaptability to Changes in Illumination, Computational Complexity, Real-Time Tracking, and Handling Non-linearities as shown in Figure 2.

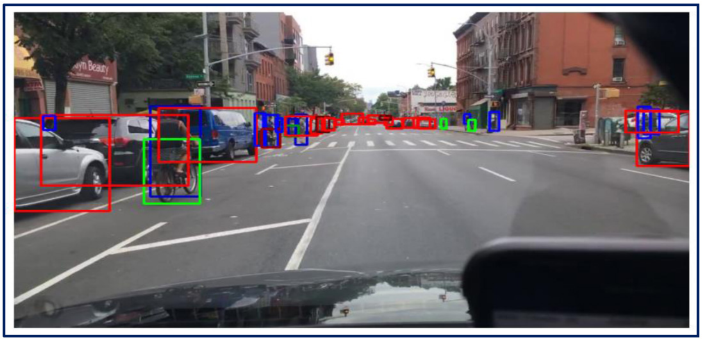

Figure 2

Object detection using CNN model

The Guided Bilateral Filter is an exceptional performer that excels in multiple critical areas. It attains the highest marks in Edge Preservation (4.9) and Noise Reduction (4.8), demonstrating its exceptional capacity to minimize noise while maintaining crucial information in the photos. Because of this, it works especially effectively for jobs where preserving visual clarity is essential.Additionally, the [12] Guided Bilateral Filter outperforms other methods in the area of Adaptability to Illumination Changes (3.1), exhibiting remarkable effectiveness in managing changes in lighting. This flexibility is an essential feature in situations where there is a large variation in illumination between frames, like in outdoor settings with shifting daylight.

The Kalman Filter performs slightly better than the other methods in Real-Time Tracking (4.4). The Kalman Filter is a popular choice for real-time tracking applications because of its exceptional ability to predict an object's next state in dynamic systems.The Gaussian Filtering technique

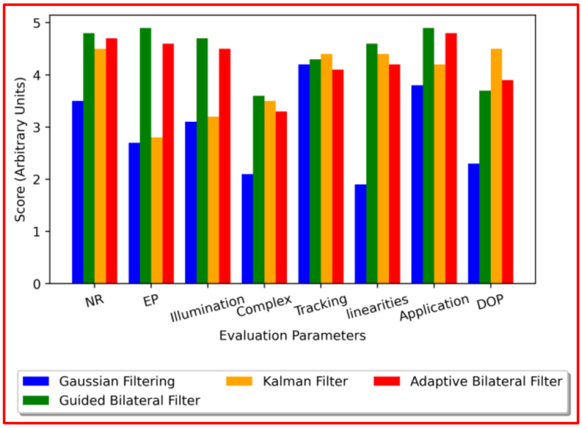

Figure 3

Representation of noise removal technique with score

maintains its simplicity and computational efficiency, giving it a niche in cases where a balance between resource efficiency and performance is necessary, even if the Guided Bilateral Filter performs better in a number of criteria. In Real-Time Tracking (4.2) and computational complexity (2.23), Gaussian Filtering obtains a competitive score.Application (4.8) is where the Adaptive Bilateral Filter shines, demonstrating its adaptability and fit for a variety of image processing uses. It is robust in a variety of settings since it may adjust filter parameters based on the specific properties of the local image.

Managing Another area in which the Guided Bilateral Filter does well is non-linearities (4.6). This implies that it is adept at handling non-linear, complicated fluctuations in picture data, giving it a benefit in situations where the image has complex properties.All things considered, the particular needs of the application should guide the selection of one of these filtering strategies. Although the Guided Bilateral Filter excels in edge retention and noise reduction, other considerations such as real-time processing limitations, the filter's capacity to adjust to changing lighting conditions, and the type of non-linearities present in the data should be taken into account. The evaluation that is being offered helps practitioners make decisions that are specific to their application needs by facilitating a deeper knowledge.

One notable feature of the Guided Bilateral Filter (GBF) is its ability to reduce redundant information, preserve edge information, and use adaptive filtering based on local similarities. By improving the signal-to-

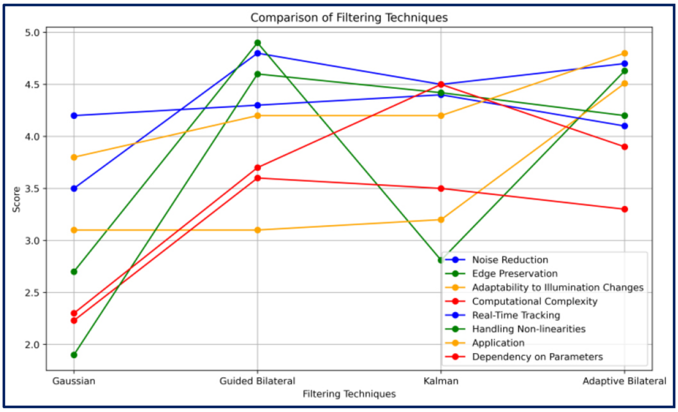

Figure 4

Comparison of Filtering Techniques

noise ratio, these characteristics give the training of Convolutional Neural Nets (CNNs) a cleaner input. GBF's effectiveness speeds up CNN convergence, which is especially important for applications involving object detection and tracking. Figure 3 and 4 shows the comparative analysis of the results.

6. Conclusion

The examination of several noise filtering methods for tracking and object detection from films with different lighting conditions reveals the subtle advantages and disadvantages of each approach. It turns out that Guided Bilateral Filtering is a good option since it does a good job of maintaining edge details, adjusting to different lighting conditions, and minimizing noise. Its ability to do adaptive filtering using a guide image gives it a significant advantage in situations when the lighting varies throughout the frame. Although Gaussian and Kalman filters are good at real-time tracking and computing efficiency, they are less effective at resolving non-linearities and maintaining edge details, which limits their applicability for object detection in difficult situations. The Adaptive Bilateral Filter works well in a variety of settings and has a high degree of adaptability. Its reliance on parameters, meanwhile, can call for careful tweaking.The Guided Bilateral Filter is a comprehensive solution for object recognition and tracking that strikes a compromise between edge preservation, noise reduction, and lighting flexibility. The particular needs of the application should determine which filtering method is best, taking into account things like computational complexity, real-time processing, and adaptability to changing lighting conditions.

References

[1] B. J. Koskowich, M. Rahnemoonfai and M. Starek, "Virtualot — A Framework Enabling Real-Time Coordinate Transformation & Occlusion Sensitive Tracking Using UAS Products, Deep Learning Object Detection & Traditional Object Tracking Techniques," IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, pp. 6416-6419 (2018).

[2] B. Leibe, K. Schindler, N. Cornelis and L. Van Gool, "Coupled Object Detection and Tracking from Static Cameras and Moving Vehicles," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 10, pp. 1683-1698, Oct. (2008).

[3] M. R. Marshall et al., "3-D Object Tracking in Panoramic Video and LiDAR for Radiological Source–Object Attribution and Improved Source Detection," in *IEEE Transactions on Nuclear Science*, vol. 68, no. 2, pp. 189-202, Feb. (2021).

[4] Q. Yu, B. Wang and Y. Su, "Object Detection-Tracking Algorithm for Unmanned Surface Vehicles Based on a Radar-Photoelectric System," in IEEE Access, vol. 9, pp. 57529-57541 (2021), doi: 10.1109/ACCESS.2021.3072897.

[5] J. Bai, S. Li, L. Huang and H. Chen, "Robust Detection and Tracking Method for Moving Object Based on Radar and Camera Data Fusion," in *IEEE Sensors Journal*, vol. 21, no. 9, pp. 10761-10774, 1 May1, (2021).

[6] Z. Meng, X. Xia, R. Xu, W. Liu and J. Ma, "HYDRO-3D: Hybrid Object Detection and Tracking for Cooperative Perception Using 3D LiDAR," in *IEEE Transactions on Intelligent Vehicles*, vol. 8, no. 8, pp. 4069-4080, Aug. (2023).

[7] N. Mustafa, S. A. Khan, J. -P. Li, M. Khalil, K. Kumar and G. Mohaned, "Medical image De-noising schemes using wavelet transform with fixed form thresholding," 2014 11th International Computer Conference on Wavelet Actiev Media Technology and Information Processing(ICCWAMTIP), Chengdu, China, pp. 397-402 (2014), doi: 10.1109/ICCWAMTIP.2014.7073435.

[8] A. Fujiki, J. Matsushita, T. Imai and M. Muneyasu, "Technique for mixed noise reduction based on support vector machine [image denoising]," NSIP 2005. Abstracts. IEEE-Eurasip Nonlinear Signal and Image Processing, 2005., Sapporo, Japan, pp. 25-, (2005) doi: 10.1109/NSIP.2005.1502256.

[9] P. Biswas, K. K. Halder and A. Sarkar, "Modified Gaussian Fuzzy Membership Function for Mixed Noise Reduction from Ultrasound Images," 2022 International Conference on Recent Progresses in Science, Engineering and Technology (ICRPSET), Rajshahi, Bangladesh, pp. 1-4 (2022), doi: 10.1109/ICRPSET57982.2022.10188550.

[10] Y. Sato, M. Kazui and S. Kobayashi, "Noise Reduction in SEM Images using Deep Learning," 2022 International Symposium on Semiconductor Manufacturing (ISSM), Tokyo, Japan, pp. 1-4 (2022), doi: 10.1109/ISSM55802.2022.10026935.

[11] K. Zhang, W. Zuo, Y. Chen, D. Meng and L. Zhang, "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising", IEEE Trans Image Process, vol. 26, pp. 3142-3155 (2017).

[12] Lehtinen, J. Munkberg, J. Hasselgren, S. Laine, T. Karras, M. Aittala, et al., "Noise2Noise: Learning Image Restoration without Clean Data", ICML J Mach Learn Res, pp. 2965-2974 (2018).

[13] Praful V. Barekar, Kavita R. Singh, Shailesh D. Kamble, Bhagyashree V. Ambulkar "On the Real Time Object Detection and Tracking," *Journal of Advanced Applied Scientific Research* -ISSN: 2454-3225, JOAASR-Vol. 4 No. 1, pp. 95-101 (2022).