A Comprehensive Review of Contour Detection and Background Subtraction Methods for Extracting Moving Objects from Videos

Mr. Prateek Oswal[1], Dr. Harsh Mathur[2]
[1]*Phd Scholar, RNTU, Bhopal*
[2]*Guide & Associate Professor, RNTU, Bhopal*

Abstract: Object extraction is a critical step in various computer vision applications such as surveillance, autonomous vehicles, and human-computer interaction. Contour detection and background subtraction are two fundamental techniques used to isolate and identify moving objects in video streams. This paper provides a comprehensive review of state-of-the-art methods for contour detection and background subtraction, emphasizing their strengths, limitations, and applicability to different real-world scenarios. The review also explores the integration of these methods to improve robustness and accuracy in object extraction. Finally, we discuss emerging trends and future research directions in this domain.

1. INTRODUCTION

The ability to detect and extract moving objects from video streams is central to the success of numerous applications in fields ranging from security surveillance to augmented reality. Two widely used approaches for achieving this goal are contour detection and background subtraction. Contour detection focuses on identifying the edges or boundaries of objects, while background subtraction involves isolating foreground objects by subtracting a background model from the current frame.

This paper aims to provide a detailed review of these methods, highlighting their evolution, theoretical underpinnings, and practical applications. Furthermore, we examine the synergy between contour detection and background subtraction to improve the accuracy and efficiency of object extraction.

2. CONTOUR DETECTION TECHNIQUES

2.1 Overview

Contour detection is a fundamental technique in computer vision that focuses on identifying the boundaries or edges of objects in an image. It involves detecting significant changes in pixel intensity, color, or texture that typically occur at the borders between different regions or objects. By delineating these boundaries, contour detection provides crucial structural information about the image content, which is essential for numerous higher-level tasks, such as object recognition, image segmentation, and scene understanding.

The process of contour detection can be broken down into several key steps:

1. Preprocessing: Images are often preprocessed to reduce noise and improve edge clarity. Techniques such as Gaussian blurring are commonly used to smooth the image while preserving important details.
2. Edge Detection: Edge detection algorithms like Sobel, Prewitt, and Canny play a vital role in identifying potential contours by highlighting regions of abrupt intensity changes.
3. Thresholding and Binarization: To isolate contours, thresholding is applied to the edge-detected image, converting it into a binary image where potential contours are clearly marked.
4. Contour Approximation: Advanced techniques, such as polygonal approximation using the Ramer-Douglas-Peucker algorithm, simplify the detected contours into a set of straight-line segments for easier interpretation.
5. Post-Processing: To refine results, post-processing steps may be applied to filter out noise, close gaps, or connect disjoint segments in the contours.

Contour detection is widely applied across various domains:

- Object Detection and Recognition: Extracted contours help identify and classify objects within an image.

- Image Segmentation: Contours are used to separate distinct regions, aiding in tasks such as medical image analysis or autonomous vehicle navigation.
- Shape Analysis: The contours of objects provide valuable information for shape matching and comparison.

2.2 Classical Methods

- Canny Edge Detector: One of the most popular methods, it uses gradient-based edge detection followed by non-maximum suppression and hysteresis thresholding.
- Sobel and Prewitt Operators: These techniques use convolution kernels to compute intensity gradients and identify edges.

2.3 Modern Approaches

- Structured Edge Detection: Exploits machine learning to learn edge structures from labeled data.
- Deep Learning-Based Methods: Neural networks, such as U-Net and HED (Holistically-Nested Edge Detection), have significantly improved the accuracy and robustness of contour detection.

2.4 Challenges

While contour detection is a powerful tool in image processing, it faces several challenges that can impact its accuracy and effectiveness. Some of the primary challenges include:

1. Noise and Artifacts
Images often contain noise due to various factors such as sensor limitations, compression artifacts, or environmental conditions. Noise can cause false edges or disrupt the continuity of true contours, leading to false positives (incorrectly detected edges) or false negatives (missed edges). Contour detection algorithms need to be robust enough to distinguish between noise and actual boundaries, which can be difficult in noisy environments.

- Solution: Preprocessing techniques like Gaussian smoothing, median filtering, and denoising algorithms can help reduce noise and preserve important edge information.

2. Illumination Changes
Changes in lighting conditions, such as shadows, reflections, or variations in brightness, can cause significant challenges in detecting contours accurately. These changes often lead to incomplete or broken contours as variations in intensity do not always correspond to true object boundaries.

- Solution: Adaptive methods that adjust for lighting conditions, such as histogram equalization or contrast enhancement, can help mitigate the impact of varying illumination.

3. Complex Backgrounds
In many images, especially in natural scenes or cluttered environments, the background can be complex and contain multiple objects with similar intensities or colors to the object of interest. This can confuse contour detection algorithms, leading to contour leakage where the detected contour includes background information or fails to isolate the object.

- Solution: Background subtraction or segmentation techniques can help isolate the foreground from the background, improving the accuracy of contour detection.

4. Multiple or Overlapping Objects
When there are multiple objects in close proximity or overlapping, contour detection may struggle to separate their individual boundaries. Contours may merge or become fragmented, making it difficult to correctly delineate objects.

- Solution: Techniques like watershed segmentation or advanced edge detection algorithms can help in such scenarios by better handling the boundaries between adjacent or overlapping objects.

5. Scale Variability
Objects in images can vary greatly in scale, which can make contour detection algorithms sensitive to objects of different sizes. A contour detection method that works well on large objects may fail to detect smaller objects, and vice versa.

- Solution: Multi-scale or pyramid-based approaches can help detect contours at different resolutions, allowing for better handling of objects across a wide range of sizes.

6. Curved and Complex Shapes

While detecting straight-line contours is relatively straightforward, curved or irregular shapes present additional challenges. Traditional edge detection methods often struggle to accurately trace complex or highly detailed boundaries.

- Solution: Advanced techniques, such as active contours (snakes) or deep learning-based methods, can better handle the detection of curved shapes and complex boundaries by learning from more diverse features.

7. Ambiguity in Edge Definition
In some cases, edges between two regions may not be well-defined, either due to similar textures or gradual transitions. This ambiguity can lead to either missed contours or incorrect contour identification.

- Solution: More sophisticated edge detection methods like the Canny edge detector or methods using gradients and local texture information can help in detecting these ambiguous edges.

8. Real-time Processing Constraints
In real-time applications, such as robotics or autonomous vehicles, contour detection must be performed quickly and efficiently. Complex algorithms that yield high accuracy may be computationally expensive and unsuitable for real-time systems.

- Solution: Optimized algorithms and hardware acceleration techniques (e.g., using GPUs) are necessary to achieve fast contour detection without sacrificing accuracy.

9. Texture and Color-Based Contours
While intensity-based contours are the most common, there are situations where texture or color variations are more reliable indicators of boundaries. However, detecting contours based solely on texture or color can be challenging due to variability in textures and color representations in different lighting conditions.

- Solution: Techniques that incorporate multi-channel or multi-feature information, such as texture gradients, color histograms, or deep learning-based models, can improve contour detection in these scenarios.

3. BACKGROUND SUBTRACTION TECHNIQUES

3.1 Overview

Background subtraction is a widely used technique in computer vision, particularly in video surveillance, object tracking, and motion detection. The primary goal of background subtraction is to detect moving objects by identifying differences between a background model and the current video frame. This technique is highly effective in scenarios where the scene's background remains relatively static, while objects in the foreground exhibit motion.

The process of background subtraction involves the following key steps:

1. Background Modeling: A model of the background is created based on the initial frames of the video or through continuous updates over time. This background model typically represents the static elements in the scene and is updated to account for slight variations due to lighting changes or environmental shifts.
2. Foreground Detection: Once the background model is established, each new frame is compared to it. The differences between the background model and the current frame are used to identify regions of the image where significant changes occur, which are presumed to be moving objects or foreground elements.
3. Thresholding: A thresholding step is applied to the difference image (the result of comparing the background model to the current frame) to classify the regions with significant changes as foreground. The areas with small changes are deemed part of the background and are discarded.
4. Post-Processing: The detected foreground regions are often refined using post-processing techniques, such as morphological operations, to remove noise or fill gaps in the detected foreground areas.

Background subtraction is essential for a variety of applications, including:

- Motion Detection: By highlighting moving objects, background subtraction allows for the detection of dynamic changes in a scene.
- Object Tracking: Once moving objects are detected, their movement can be tracked across successive frames, facilitating tasks like pedestrian or vehicle tracking.

- Surveillance Systems: Background subtraction plays a key role in security systems by isolating intruders or anomalous activities in monitored areas.
- Robotics: Robots use background subtraction for navigation and interaction with dynamic environments, where they need to identify obstacles or moving objects.

3.2 Classical Approaches

- Frame Differencing: Compares consecutive frames to detect motion.
- Gaussian Mixture Models (GMM): Models each pixel as a mixture of Gaussians to handle dynamic backgrounds.
- Median Filtering: Computes the median of pixel values over time to create a robust background model.

3.3 Advanced Techniques

- ViBe and PBAS: Non-parametric methods that adaptively update the background model for dynamic environments. These methods maintain a history of pixel values, making them robust to minor variations and noise.
- Deep Learning-Based Approaches: Networks like FgSegNet use convolutional architectures to learn complex spatio-temporal patterns in videos. These models are particularly effective in handling challenging scenarios such as dynamic backgrounds, shadows, and occlusions.

3.4 Challenges

Background subtraction methods often struggle with:

- Illumination Variations: Sudden changes in lighting can cause false detections.
- Dynamic Backgrounds: Elements such as waving trees or rippling water are difficult to differentiate from moving objects.
- Shadows and Reflections: These can lead to incorrect segmentation of objects.
- Initialization Requirements: Some methods require a period of "background learning," during which no moving objects should be present.

4. INTEGRATION OF CONTOUR DETECTION AND BACKGROUND SUBTRACTION

Combining contour detection and background subtraction can enhance the accuracy of moving object extraction. Background subtraction provides a coarse foreground mask, which can be refined using contour detection to delineate object boundaries more precisely.

4.1 Complementary Benefits

- Boundary Precision: Background subtraction effectively isolates foreground regions, while contour detection sharpens object boundaries.
- Noise Reduction: Integration reduces false positives by leveraging both pixel-level changes and edge information.

4.2 Case Studies

- Surveillance Systems: In security applications, background subtraction identifies regions of interest, while contour detection ensures accurate tracking by delineating objects clearly. For instance, tracking individuals in crowded environments benefits from this synergy.
- Autonomous Vehicles: Vehicles use these methods to detect pedestrians, other vehicles, and obstacles. Background subtraction identifies moving objects, and contour detection helps refine their shapes for precise localization.
- Human-Computer Interaction (HCI): Gesture recognition systems often rely on this combination to detect hand movements against complex backgrounds.

4.3 Implementation Considerations

- Computational Efficiency: Combining these methods increases computational requirements, making optimization crucial for real-time applications.
- Parameter Tuning: Properly tuning thresholds for both background subtraction and contour detection is essential to avoid over-segmentation or under-detection.

5. EMERGING TRENDS AND FUTURE DIRECTIONS

5.1 Real-Time Processing

The demand for real-time processing is driving advancements in hardware acceleration and algorithm optimization. Techniques such as GPUaccelerated computing and parallel processing frameworks like OpenCV and CUDA are enabling the deployment of computationally intensive models in time-critical applications.

5.2 Robustness to Environmental Variations

Research is increasingly focused on making these methods robust to:

- Weather Conditions: Techniques are being developed to handle challenges such as rain, snow, and fog, which can obscure object boundaries.
- Lighting Changes: Algorithms are incorporating adaptive thresholding and learning-based approaches to manage sudden illumination changes.
- Occlusions: Methods combining temporal and spatial cues are being explored to predict the presence of partially occluded objects.

5.3 Integration with High-Level Semantics

Combining contour detection and background subtraction with semantic segmentation and object recognition can provide a richer understanding of scenes. For instance, recognizing the type of moving object (e.g., pedestrian vs. vehicle) enhances decision-making in applications like autonomous driving.

5.4 Multi-Sensor Fusion

The fusion of data from multiple sensors, such as RGB cameras, thermal imaging, and LiDAR, is becoming increasingly important. Multi-sensor approaches provide complementary information, improving robustness in challenging environments. For example, LiDAR data can provide depth information to complement 2D contour detection.

5.5 Explainable AI (XAI)

As deep learning models dominate the field, there is a growing emphasis on explain ability. Researchers are working on methods to interpret and visualize the decisions made by models, fostering trust in critical applications such as surveillance and healthcare.

6. CONCLUSION

Contour detection and background subtraction are cornerstone techniques for extracting moving objects in video streams. While each method has its unique strengths and challenges, their integration offers a powerful approach to achieving accurate and robust object extraction. Continued research in this domain promises significant advancements, driven by innovations in deep learning, real-time processing, and multi-modal data fusion. These developments will pave the way for more sophisticated and reliable computer vision systems.

7. REFERENCES

[1] S. Mokssit, D. B. Licea, B. Guermah, and M. Ghogho, "A Spatial Transformation based next Frame Predictor," *IEEE Access*, doi: 10.1109/ACCESS.2024.3524997.

[2] H. Fang, G. Liao, Y. Liu, C. Zeng, X. He, and M. Xu, "A Joint Moving Target Detection Method in Video SAR Via Low-Rank Sparse Decomposition and Transformer," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 18, pp. 1007-1019, 2025, doi: 10.1109/JSTARS.2024.3503639.

[3] X. Sheng, Z. Shen, and G. Xiao, "Saliency-aware Spatio-temporal Modeling for Action Recognition on Unmanned Aerial Vehicles," *IEEE Latin America Transactions*, vol. 22, no. 12, pp. 1026-1033, Dec. 2024, doi: 10.1109/TLA.2024.10789633.

[4] 4.Y. Ueno and M. Seo, "Improving Facial Video Conversion Using Optical Flow and Edge Feature," in *2024 IEEE 13th Global Conference on Consumer Electronics (GCCE)*, Kitakyushu, Japan, 2024, pp. 726-728, doi: 10.1109/GCCE62371.2024.10760910.

[5] E. Ngharamike, L. -M. Ang, M. Wang, and K. P. Seng, "Analysis of the Impact of Light Sources on the Capturing and Extraction of ENF from Video Recordings," in *2024 9th International Conference on Communication, Image and Signal Processing (CCISP)*, Gold Coast, Australia, 2024, pp. 89-96, doi: 10.1109/CCISP63826.2024.10765509.

[6] Y. Zhuo, C. Gu, and S. Kuriyama, "Consistency-Preserving Text-Based One-Shot Video Tuning with Segmentation Mask Guidance," in *2024 11th International Conference on Advanced Informatics: Concept, Theory and Application (ICAICTA)*, Singapore, 2024, pp. 1-6, doi: 10.1109/ICAICTA63815.2024.10763203.

[7] "IEEE Standard Adoption of Moving Picture, Audio and Data Coding by Artificial Intelligence

(MPAI) Technical Specification Multimodal Conversation--Version 2," *IEEE Std 3300-2024*, vol., no., pp. 1-89, Nov. 14, 2024, doi: 10.1109/IEEESTD.2024.10756013.

[8] M. Li, S. Gong, W. Cao, X. Yin, H. Wang, and J. Wang, "Video Significance Object Detection Model Based on Spatiotemporal Progressive Learning," in *2024 IEEE 7th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, Chongqing, China, 2024, pp. 1410-1414, doi: 10.1109/ITNEC60942.2024.10733173.

[9] S. S. Mallick, I. Das, S. Das, and R. Karmakar, "An Intelligent Crowd Density and Motion Direction Estimation for Real-Time Crowd Dynamics," in *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Kamand, India, 2024, pp. 1-6, doi: 10.1109/ICCCNT61001.2024.10723903.

[10] H. Chen, "YOLO Algorithm in Analysis and Design of Athletes' Actions in College Physical Education," in *2024 International Conference on Interactive Intelligent Systems and Techniques (IIST)*, Bhubaneswar, India, 2024, pp. 764-768, doi: 10.1109/IIST62526.2024.00002.