

Temporal Gradient-Based Moving Object Detection Using Derivative Filters and Adaptive Thresholding

Eduard Aliaga Torrens

Department of Electrical Engineering or Computer Science
Northeastern University
Email: aliagatorrens.e@northeastern.upc.edu

Abstract—This paper presents a temporal gradient-based approach for detecting moving objects in image sequences captured by a stationary camera. The method relies on the observation that background pixels exhibit minimal temporal variation, while moving objects produce significant intensity changes over time. Motion is detected using temporal derivative operators, including a central finite-difference filter and a derivative-of-Gaussian filter at multiple temporal scales. The impact of spatial smoothing and different threshold selection strategies is systematically evaluated. Experimental results show that moderate Gaussian spatial smoothing combined with derivative-of-Gaussian temporal filtering and adaptive noise-model thresholding yields the most reliable motion segmentation, effectively suppressing background noise while preserving object boundaries.

I. DESCRIPTION OF ALGORITHMS

The proposed motion detection algorithm operates on an ordered sequence of grayscale image frames acquired from a stationary camera. Let $I(x, y, t)$ denote the pixel intensity at spatial location (x, y) and time index t . The objective is to detect moving objects by identifying large temporal intensity variations at each pixel.

A. Temporal Derivative Estimation

Motion is detected by computing the temporal derivative of pixel intensities. Two temporal derivative operators are considered:

1) *Finite Difference Operator*: A simple central difference approximation is used:

$$\frac{\partial I}{\partial t}(x, y, t) \approx \frac{1}{2} [I(x, y, t + 1) - I(x, y, t - 1)]. \quad (1)$$

This operator requires only three consecutive frames and provides a direct estimate of temporal intensity change.

2) *Derivative of Gaussian (DoG)*: To improve robustness to temporal noise, a 1-D Gaussian smoothing kernel $G_\sigma(t)$ is applied along the temporal axis before differentiation. The derivative is computed as:

$$\frac{\partial I}{\partial t}(x, y, t) = \left(\frac{d}{dt} G_{\sigma_t}(t) \right) * I(x, y, t), \quad (2)$$

where σ_t denotes the temporal standard deviation and $*$ represents convolution along the time dimension. Multiple values of σ_t are evaluated to analyze the effect of temporal smoothing.

B. Spatial Smoothing Before Temporal Differentiation

To reduce spatial noise prior to temporal differentiation, in some experiments spatial smoothing is applied independently to each frame. Three smoothing strategies are evaluated:

- 3×3 box filter,
- 5×5 box filter,
- 2-D Gaussian filter with standard deviation σ_s .

Let $H_s(x, y)$ denote the chosen spatial filter. The smoothed frame is obtained as:

$$\tilde{I}(x, y, t) = H_s(x, y) * I(x, y, t). \quad (3)$$

Temporal derivatives are then computed using $\tilde{I}(x, y, t)$.

C. Motion Mask Generation

After computing the temporal derivative $D(x, y, t)$, motion is detected by thresholding its magnitude:

$$M(x, y, t) = \begin{cases} 1, & |D(x, y, t)| > T, \\ 0, & \text{otherwise.} \end{cases} \quad (4)$$

Three threshold selection strategies are considered:

- Fixed threshold $T = c$,
- Percentile-based threshold using the p -th percentile of $|D|$,
- Adaptive noise-model threshold.

See the following sections for the explanation of the last two threshold algorithms mentioned.

D. Percentile-Based Threshold

An alternative data-driven thresholding strategy is based on the empirical distribution of the temporal derivative magnitude. Histogram-based threshold selection methods are widely used in image segmentation and global decision rules [1], [2]. Since moving objects typically occupy a relatively small portion of the image, large derivative values correspond to a small upper fraction of the distribution of $|D(x, y, t)|$.

Let $F_{|D|}(\cdot)$ denote the cumulative distribution function of the absolute derivative values over all pixels in a frame. The threshold is defined as the p -th percentile (quantile) of this distribution:

$$T = \text{Percentile}_p(|D(x, y, t)|), \quad (5)$$

where $p \in [0, 100]$ is a user-defined percentile. Quantile-based decision rules provide a non-parametric way to adapt the threshold to the observed data distribution without assuming a specific noise model [2].

Equivalently, T satisfies

$$P(|D(x, y, t)| \leq T) = \frac{p}{100}. \quad (6)$$

The motion mask is then obtained as

$$M(x, y, t) = \begin{cases} 1, & |D(x, y, t)| > T, \\ 0, & \text{otherwise.} \end{cases} \quad (7)$$

Larger values of p produce stricter thresholds and fewer detected motion pixels, while smaller values increase sensitivity at the cost of potentially higher false detections.

E. Adaptive Noise-Model Threshold

Since most pixels correspond to stationary background, their temporal derivatives are assumed to follow a zero-mean Gaussian noise model. Under this assumption, motion detection can be formulated as a statistical hypothesis test where large deviations from the background distribution are classified as foreground [3].

To obtain a robust estimate of the background noise variance, the Median Absolute Deviation (MAD) estimator is used. The MAD is widely employed in robust statistics because it is resilient to outliers and does not require prior knowledge of the underlying distribution [4], [5]. It is defined as:

$$\text{MAD} = \text{median}(|D| - \text{median}(|D|)|). \quad (8)$$

For a zero-mean Gaussian distribution, the MAD can be converted to an estimate of the standard deviation using the consistency constant 0.6745 [4]:

$$\sigma_{\text{noise}} = \frac{\text{MAD}}{0.6745}. \quad (9)$$

The detection threshold is then defined as

$$T = k \sigma_{\text{noise}}, \quad (10)$$

where k controls the detection sensitivity. Larger values of k reduce false detections by enforcing stricter deviation from the background noise model, while smaller values increase sensitivity to weak motion at the expense of potentially higher false positives. Under the Gaussian background assumption, the detection rule $T = k \sigma_{\text{noise}}$ can be interpreted as a statistical hypothesis test. If background derivatives follow a zero-mean normal distribution, the probability of observing a deviation larger than $k\sigma$ is given by the tail probability of the Gaussian distribution,

$$P(|D| > k\sigma) = 2\Phi(-k),$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution [3]. Therefore, increasing k corresponds to enforcing a stricter confidence level, reducing the false positive rate in a statistically controlled manner.

In practice, the noise standard deviation σ_{noise} is estimated robustly using the Median Absolute Deviation (MAD), which is resilient to outliers and does not require prior knowledge of the full distribution [4], [5]. As k increases, the adaptive threshold effectively selects progressively smaller upper-tail portions of the derivative distribution, making its behavior conceptually related to high-percentile thresholding, but grounded in an explicit statistical model of background noise.

F. Visualization

The binary motion mask is overlaid onto the original frame by setting detected motion pixels to maximum intensity for visualization purposes. Quantitative comparison between filter configurations is performed by reporting the percentage of detected motion pixels for each combination of spatial smoothing, temporal derivative, and thresholding strategy.

II. EXPERIMENTS AND VALUES OF PARAMETERS USED

The proposed motion detection algorithms were evaluated on image sequences captured by a stationary camera. All frames were converted to grayscale and processed as double-precision intensity arrays. Experiments were conducted on a representative frame located at the center of each sequence to ensure sufficient temporal context on both sides for derivative computation.

A. Temporal Derivative Parameters

Two temporal derivative operators were evaluated:

- Central finite difference filter $0.5[-1, 0, 1]$.
- Derivative of Gaussian (DoG) with temporal standard deviations $\sigma_t \in \{0.5, 1.5, 5.0\}$.

For the DoG operator, the temporal support was truncated at $\pm 3\sigma_t$ to ensure numerical stability and sufficient smoothing.

B. Spatial Smoothing Parameters

To analyze the effect of spatial noise reduction prior to temporal differentiation, the following spatial filters were applied:

- 3×3 box filter,
- 5×5 box filter,
- 2-D Gaussian filter with spatial standard deviations $\sigma_s \in \{0.5, 1.5, 5.0\}$.

Spatial smoothing was applied independently to each frame before computing temporal derivatives.

C. Thresholding Parameters

Motion masks were obtained by thresholding the magnitude of the temporal derivative. Three thresholding strategies were evaluated:

- Fixed thresholds $T \in \{2, 5, 10, 15, 20, 30, 50\}$,
- Percentile-based thresholds with percentiles $p \in \{85, 90, 95\}$,
- Adaptive noise-model thresholds with sensitivity parameters $k \in \{2.0, 3.0, 4.0, 5.0\}$.

For the percentile-based experiments, a default value of $p = 90$ was used for systematic comparison across filter configurations. In the adaptive approach, the background noise standard

deviation was estimated using the Median Absolute Deviation (MAD), and the threshold was set to $T = k\sigma_{\text{noise}}$.

D. Evaluation Criteria

To enable quantitative comparison between filtering configurations, three complementary metrics were computed for each combination of spatial smoothing, temporal derivative, and thresholding strategy:

- **Motion Percentage:** Let $M(x, y)$ denote the binary motion mask, where $M(x, y) = 1$ indicates motion and 0 otherwise, and let N be the total number of pixels in the image. The motion percentage is defined as

$$\text{Motion (\%)} = \frac{\sum_{x,y} M(x, y)}{N} \times 100.$$

This metric measures detection sensitivity (extent of detected motion) but does not directly assess segmentation quality.

- **Signal-to-Noise Ratio (SNR):** Let $A(x, y) = |D(x, y)|$ denote the absolute temporal derivative magnitude. Rather than using the thresholded motion mask, we compute a distribution-based SNR from A . Define

$$\begin{aligned}\Omega_{95} &= \{(x, y) : A(x, y) \geq P_{95}(A)\}, \\ \Omega_{50} &= \{(x, y) : A(x, y) \leq P_{50}(A)\},\end{aligned}$$

where $P_q(A)$ denotes the q -th percentile of the values of A . The derivative SNR is then defined as

$$\text{SNR} = \frac{1}{|\Omega_{95}|} \sum_{(x,y) \in \Omega_{95}} A(x, y) / \text{std}(\{A(x, y) : (x, y) \in \Omega_{50}\}) + \epsilon,$$

with a small ϵ for numerical stability. This metric measures how strongly the largest derivative responses stand out relative to the variability of low-response (background-like) pixels, and is therefore constant for a fixed derivative field regardless of the thresholding strategy.

- **Largest Connected Component Ratio (LargestCC):**

Let C_{\max} denote the number of pixels in the largest connected component of the binary mask M , and let $N_m = \sum_{x,y} M(x, y)$ be the total number of motion pixels. The LargestCC ratio is defined as

$$\text{LargestCC} = \frac{C_{\max}}{N_m}.$$

This metric measures spatial coherence: values close to 1 indicate that most detected motion pixels belong to a single compact region, while lower values reflect fragmentation and scattered detections.

While motion percentage alone may remain approximately constant under percentile-based thresholding, SNR and LargestCC provide insight into detection robustness and spatial consistency. Together with qualitative visualization of derivative magnitudes, binary masks, and overlays, these metrics allow systematic assessment of the trade-offs between temporal smoothing, spatial filtering strength, and threshold selection.

III. RESULTS AND DISCUSSION

For consistency across all configurations, the 90th percentile threshold is used as the reference operating point in the figures presented in the results and discussion section. This percentile provides a balanced trade-off between noise suppression and preservation of the motion silhouette, as reflected by stable SNR and LargestCC values across filtering strategies.

A. Temporal Derivative

Figure 1 compares the temporal derivative operators without spatial smoothing. The quantitative results in Table I reinforce the qualitative observations and reveal systematic trends across percentile thresholds.

The simple central-difference operator $0.5[-1, 0, 1]$ produces sharp motion boundaries but is highly sensitive to frame-to-frame intensity fluctuations. As shown in Table I, the simple central-difference operator exhibits an SNR of 31.15 for all tested percentile thresholds. Although increasing the percentile reduces the motion percentage (from 14.29% at $p = 80$ to 4.75% at $p = 95$), spatial coherence improves only moderately (LargestCC increases from 0.174 to 0.363). This indicates that stronger thresholding suppresses background noise but does not fully compensate for the intrinsic instability of the unsmoothed temporal derivative.

Introducing temporal Gaussian smoothing significantly improves robustness. For DoG with $t_\sigma = 0.5$, the SNR increases to 41.55, and LargestCC rises gradually from 0.147 ($p = 80$) to 0.378 ($p = 95$). However, spatial coherence remains limited at moderate percentiles (e.g., 0.243 at $p = 90$), showing that mild temporal smoothing provides only partial noise suppression.

When $t_\sigma = 1.5$, performance improves substantially. The SNR increases sharply to 112.85—more than three times that of the simple derivative. Spatial coherence also increases consistently with percentile, reaching 0.467 at the selected reference percentile $p = 90$ and 0.663 at $p = 95$. This configuration demonstrates strong separation between motion and background while preserving a well-defined silhouette.

For large temporal smoothing ($t_\sigma = 5.0$), SNR further increases to 267.51, and spatial coherence becomes very high at moderate percentiles (LargestCC = 0.781 at $p = 80$ and 0.731 at $p = 90$). These values indicate excellent suppression of background fluctuations and highly compact motion regions. However, at very selective thresholds ($p = 95$), LargestCC drops to 0.418 despite the high SNR, reflecting the loss of weaker motion components due to excessive smoothing combined with aggressive thresholding.

We can see a clear monotonic increase in SNR with larger temporal scales, while LargestCC generally improves with both σ_t and percentile up to a point. Moderate temporal smoothing ($t_\sigma \approx 1.5$) provides an excellent trade-off: it achieves SNR values above 100 and strong spatial coherence at $p = 90$, while preserving finer structural details compared to the extreme smoothing case ($t_\sigma = 5.0$).

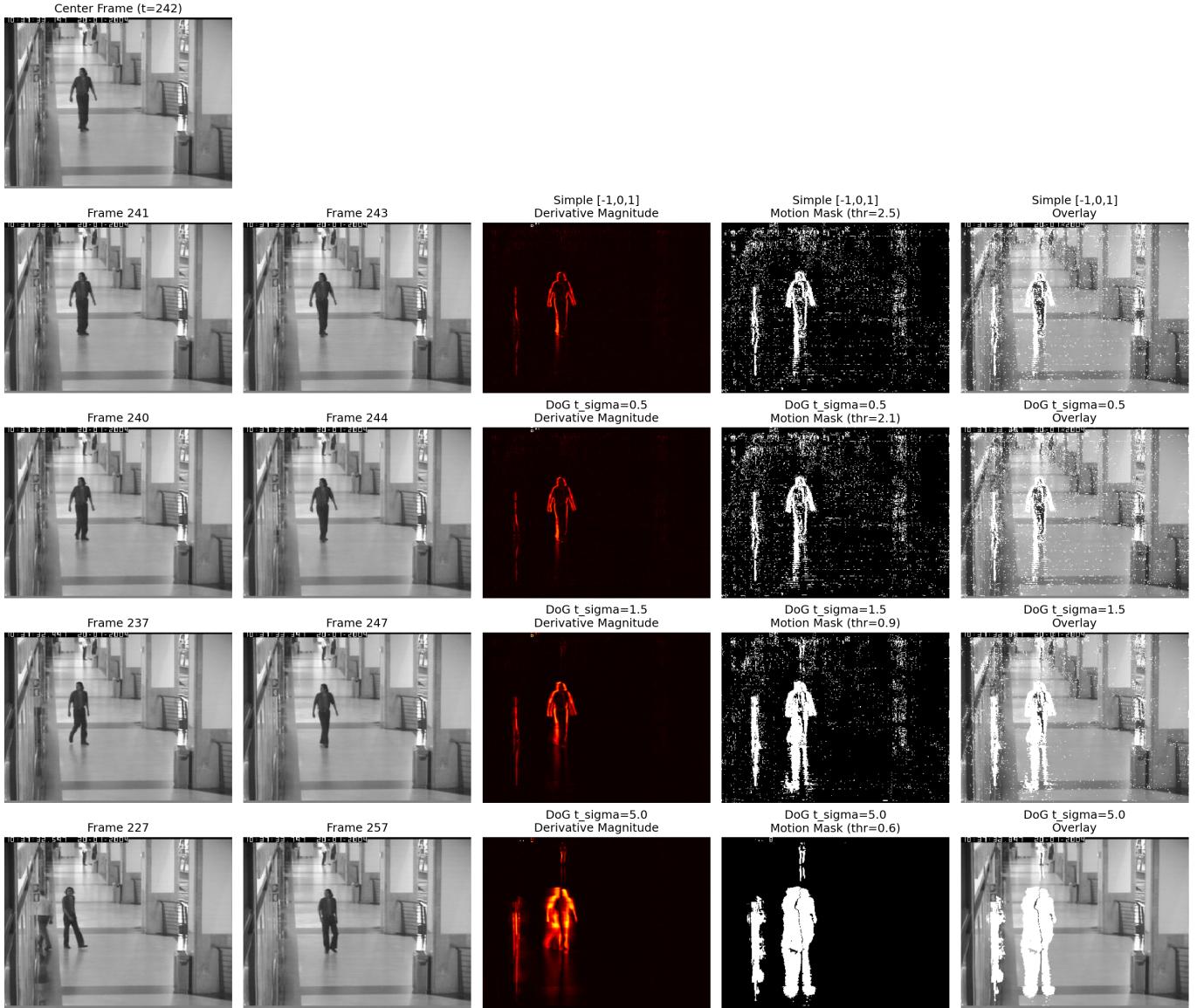


Fig. 1. Comparison of temporal derivative operators without spatial smoothing. The first row shows the center frame. Subsequent rows correspond to the simple central difference filter and derivative-of-Gaussian (DoG) filters with increasing temporal standard deviation $\sigma_t \in \{0.5, 1.5, 5.0\}$. For each method, the neighboring frames used for computation, the derivative magnitude, the resulting motion mask, and the overlay on the original frame are shown. Increasing σ_t reduces background noise and produces smoother motion masks, at the cost of reduced boundary sharpness.

B. Spatial Smoothing Before Temporal Differentiation

a) Effect of Spatial Smoothing on Noise Suppression: The quantitative results in Tables II–VII, together with the visual evidence in Figures 2 and 3, clearly demonstrate the impact of spatial smoothing prior to temporal differentiation, extending the baseline behavior observed in Table I.

Without spatial filtering (Table I), the simple central-difference operator achieves a low SNR of 31.15 and limited spatial coherence (LargestCC = 0.239 at the reference percentile $p = 90$). Visually, this corresponds to strong background contamination and highly fragmented masks.

Introducing a 3×3 box filter (Figure 2, Table II) substantially improves robustness. For the simple derivative at $p = 90$, SNR

increases from 31.15 to 81.43 and LargestCC from 0.239 to 0.331. The masks in Figure 2 show reduced pixel-level noise, although noticeable background artifacts remain. A similar improvement is observed for DoG with $t_\sigma = 1.5$, where SNR increases from 112.85 (no spatial smoothing) to 181.41 and LargestCC from 0.467 to 0.575 at $p = 90$. This confirms that even mild spatial averaging significantly improves motion coherence.

Increasing the box filter size to 5×5 (Figure 3, Table IV) further enhances separation. For the simple derivative at $p = 90$, SNR rises to 100.76 and LargestCC to 0.387. For DoG with $t_\sigma = 1.5$, SNR increases to 226.66 and LargestCC to 0.640, representing a substantial improvement over the

unsmoothed case. As seen in Figure 3, background fluctuations are strongly attenuated and motion regions become more compact. However, silhouettes appear thicker and boundaries less sharply defined due to increased spatial averaging.

Gaussian spatial filtering provides more controlled smoothing behavior. For $\sigma_s = 0.5$ (Table V), improvements are moderate: for the simple derivative at $p = 90$, SNR reaches 58.74 and LargestCC 0.254, visually comparable to mild box filtering. When $\sigma_s = 1.5$ (Table VI), performance improves significantly. With DoG $t_\sigma = 1.5$ at $p = 90$, SNR reaches 231.07 and LargestCC increases to 0.648, clearly outperforming both the unsmoothed and box-filtered cases. Visually, this configuration achieves strong background suppression while maintaining well-defined motion contours.

For strong spatial smoothing ($\sigma_s = 5.0$, Table VII), SNR continues to increase dramatically. For example, combining $\sigma_s = 5.0$ with $t_\sigma = 5.0$ yields SNR = 717.89. However, spatial coherence does not improve monotonically. At $p = 90$, LargestCC decreases to 0.445 for this most aggressive configuration, despite the extremely high SNR. This behavior reflects over-smoothing: although high-frequency noise is effectively eliminated, meaningful motion gradients are attenuated and object boundaries become broadened or distorted, reducing structural consistency.

b) Interaction Between Spatial and Temporal Filtering: The interaction between spatial and temporal filtering can be clearly observed by comparing the unsmoothed baseline in Table I with the spatially smoothed configurations in Tables II–VII, together with the qualitative behavior shown in Figures 2, 3, and 4–6.

For the simple central-difference operator, spatial smoothing alone substantially improves robustness but does not fully compensate for temporal instability. Without spatial filtering (Table I), the simple derivative yields SNR = 31.15 and LargestCC = 0.239 at the reference percentile $p = 90$, corresponding visually to strong background contamination and fragmented masks. Introducing a 3×3 box filter (Table II, Figure 2) increases SNR to 81.43 and LargestCC to 0.331, visibly reducing high-frequency noise while preserving the main motion structure. Expanding the support to a 5×5 box filter (Table IV, Figure 3) further raises SNR to 100.76 and LargestCC to 0.387, producing cleaner masks but noticeably thicker silhouettes.

Similar behavior is observed with Gaussian smoothing. For $\sigma_s = 0.5$ (Table V, Figure 4), improvements are moderate and comparable to the 3×3 box filter, with slight attenuation of background noise and preserved boundaries. When $\sigma_s = 1.5$ (Table VI, Figure 5), SNR reaches 105.36 for the simple derivative and spatial coherence improves to 0.402 at $p = 90$, producing visibly cleaner and more compact motion regions. However, even under strong spatial smoothing ($\sigma_s = 5.0$, Table VII, Figure 6), although SNR increases to 121.29, the simple operator remains less stable and more sensitive to temporal fluctuations than the DoG filter, and boundaries become increasingly blurred.

When spatial smoothing is combined with the DoG temporal

derivative, the improvement is more pronounced and visually evident. For moderate temporal scale ($t_\sigma = 1.5$), SNR increases from 112.85 without spatial smoothing (Table I) to 226.66 with a 5×5 box filter (Table IV, Figure 3) and to 231.07 with Gaussian smoothing at $\sigma_s = 1.5$ (Table VI, Figure 5). At the same percentile ($p = 90$), LargestCC rises from 0.467 (no spatial smoothing) to 0.640 (Box 5×5) and 0.648 (Gaussian $\sigma_s = 1.5$), indicating strong spatial coherence and compact motion regions. Visually, this configuration yields clean masks with well-defined silhouettes and minimal background artifacts. This demonstrates that moderate spatial and temporal smoothing reinforce each other, simultaneously improving noise suppression and connected-component consistency.

For large temporal scales combined with strong spatial smoothing, the trade-off becomes evident. For example, with $\sigma_s = 5.0$ and $t_\sigma = 5.0$ (Table VII, Figure 6), SNR reaches its maximum value of 717.89, indicating extremely strong background suppression. However, at $p = 90$, LargestCC decreases to 0.445 compared to 0.735 for $\sigma_s = 1.5$, $t_\sigma = 5.0$ (Table VI). As seen in Figure 6, silhouettes become excessively thick and fine structural details are lost. This confirms that excessive smoothing in both domains, while maximizing numerical separability, reduces spatial consistency and degrades boundary localization.

c) Best Performing Configuration: Considering the quantitative results at the reference percentile ($p = 90$), the most balanced configuration corresponds to Gaussian spatial smoothing with $\sigma_s = 1.5$ combined with a derivative-of-Gaussian temporal filter at $\sigma_t = 1.5$ (Table VI). This configuration achieves SNR = 231.07 and LargestCC = 0.648, indicating strong signal-to-noise separation together with high spatial coherence.

Importantly, although similar performance is obtained with a 5×5 box filter (SNR = 226.66, LargestCC = 0.640; Table IV), Gaussian smoothing provides slightly better connected-component consistency while preserving smoother spatial transitions. This suggests that controlled Gaussian averaging offers a more stable balance between noise suppression and boundary preservation.

More aggressive smoothing ($\sigma_s = 5.0$, $\sigma_t = 5.0$) yields the highest SNR overall (717.89; Table VII), but at the reference percentile the LargestCC drops to 0.445. This reduction reflects over-smoothing: while background fluctuations are heavily suppressed, meaningful motion gradients are attenuated and structural consistency deteriorates.

Conversely, weaker smoothing configurations (e.g., no spatial filtering or small σ_t values in Table I) produce substantially lower SNR and reduced spatial coherence.

Overall, moderate Gaussian spatial smoothing ($\sigma_s \approx 1.5$) combined with moderate temporal smoothing ($\sigma_t \approx 1.5$) provides the best trade-off between background suppression and preservation of meaningful motion structure.

C. Motion Mask Generation

a) Fixed Threshold Strategy: The quantitative behavior of the fixed threshold strategy is summarized in Table III. For

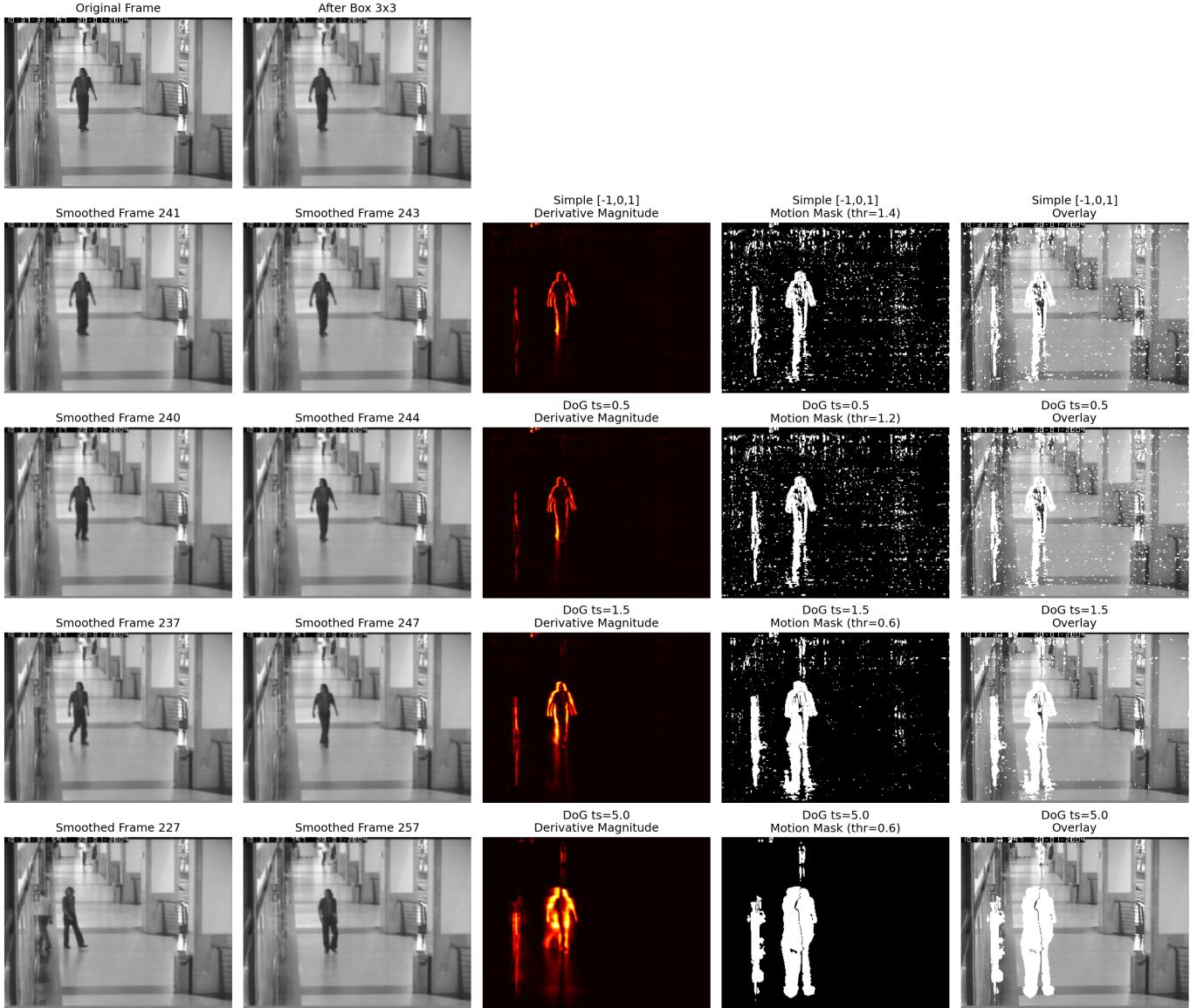


Fig. 2. Effect of applying a 3×3 box filter prior to temporal differentiation. The first row shows the original frame and its smoothed version. Subsequent rows display the neighboring smoothed frames used for derivative computation, the temporal derivative magnitude (for the simple central difference and DoG filters with increasing t_σ), the resulting motion masks, and the corresponding overlays. The 3×3 box filter slightly reduces high-frequency spatial noise; however, background artifacts remain visible in the motion masks, particularly when combined with the simple temporal derivative operator.

low threshold values (e.g., $T = 2$), a large proportion of pixels is classified as motion (14.29%), which is consistent with the fragmented and noisy masks observed in Figures 7 and 8. The LargestCC ratio is relatively low (0.174), indicating that the detected motion is dispersed across many small connected components rather than forming a coherent region. This confirms that small temporal fluctuations in the background are incorrectly interpreted as motion, leading to over-segmentation.

As the threshold increases to moderate values (e.g., $T = 5$ and $T = 10$), the percentage of detected motion pixels drops significantly (2.81% and 1.60%, respectively). At $T = 5$, the LargestCC ratio increases to 0.542, indicating that the

remaining motion pixels are more spatially coherent and primarily concentrated on the moving object. This corresponds to visibly cleaner masks in Figures 9. However, further increasing the threshold reduces the motion percentage below 2%, and although spatial coherence may remain relatively high (e.g., 0.538 at $T = 20$), relevant motion regions begin to disappear, reflecting under-segmentation.

For large threshold values ($T \geq 30$), the detected motion percentage falls below 1%, and the LargestCC ratio decreases again (0.303 at $T = 30$), indicating that only isolated high-gradient pixels remain. This thinning effect is also visible in Figures 9 and 10, where structural details of the moving object are progressively lost.

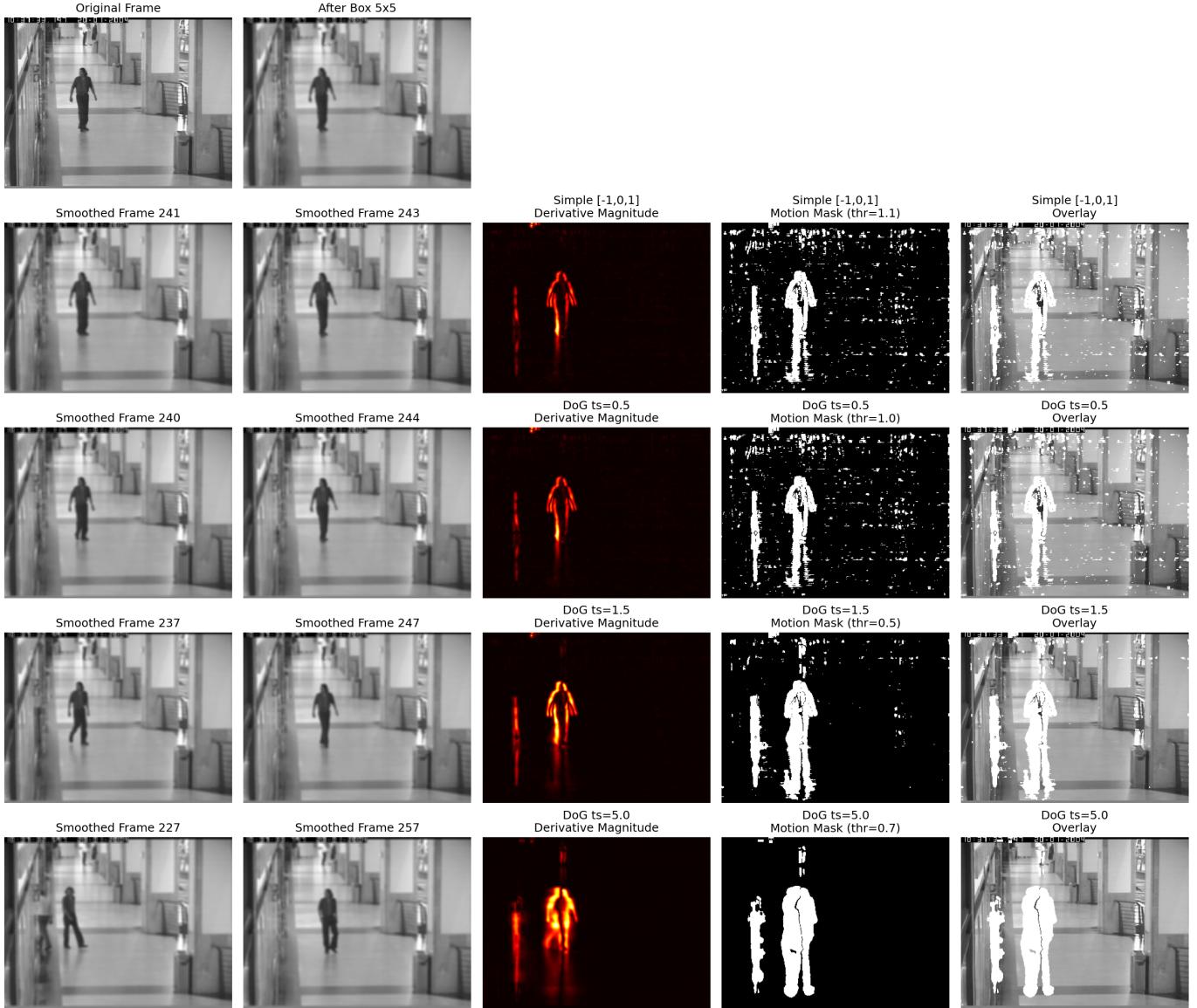


Fig. 3. Effect of applying a 5×5 box filter prior to temporal differentiation. The first row shows the original frame and its spatially smoothed version. Subsequent rows display the smoothed neighboring frames used for derivative computation, the temporal derivative magnitude (for the simple central difference and DoG filters with increasing t_σ), the resulting motion masks, and the corresponding overlays. Compared to the 3×3 filter, stronger spatial averaging reduces background noise more effectively but produces thicker and more blurred motion boundaries due to increased spatial smoothing.

b) Percentile-Based threshold:: As mentioned before, all figures used to compare filtering strategies were generated using the percentile-based thresholding strategy with $p = 90$. The quantitative results in Table I (no spatial smoothing) and Tables II–VII (with spatial smoothing) allow a systematic comparison of percentile values $p = 80, 85, 90$, and 95 across all temporal and spatial configurations.

For the representative case of the DoG temporal derivative with $t_\sigma = 1.5$ and no spatial smoothing (Table I), increasing the percentile reduces the detected motion area from 20% ($p = 80$) to 5% ($p = 95$), while LargestCC increases from 0.273 to 0.663. This indicates that stronger thresholds improve spatial coherence by removing scattered background responses, but

at the cost of progressively shrinking the detected silhouette. As illustrated in Fig. 11, lower percentiles ($p = 80, 85$) lead to over-segmentation and thick, noisy regions, whereas $p = 95$ produces thin masks that begin to erode weaker motion components. The intermediate value $p = 90$ provides a balanced compromise (Motion = 10%, LargestCC = 0.467), preserving object structure while suppressing most background noise.

A similar trend is observed when spatial smoothing is introduced. For example, with Gaussian smoothing $\sigma_s = 1.5$ and DoG $t_\sigma = 1.5$ (Table VI), LargestCC increases from 0.403 ($p = 80$) to 0.676 ($p = 95$), while motion extent decreases from 20% to 5%. As shown in Fig. 12, lower

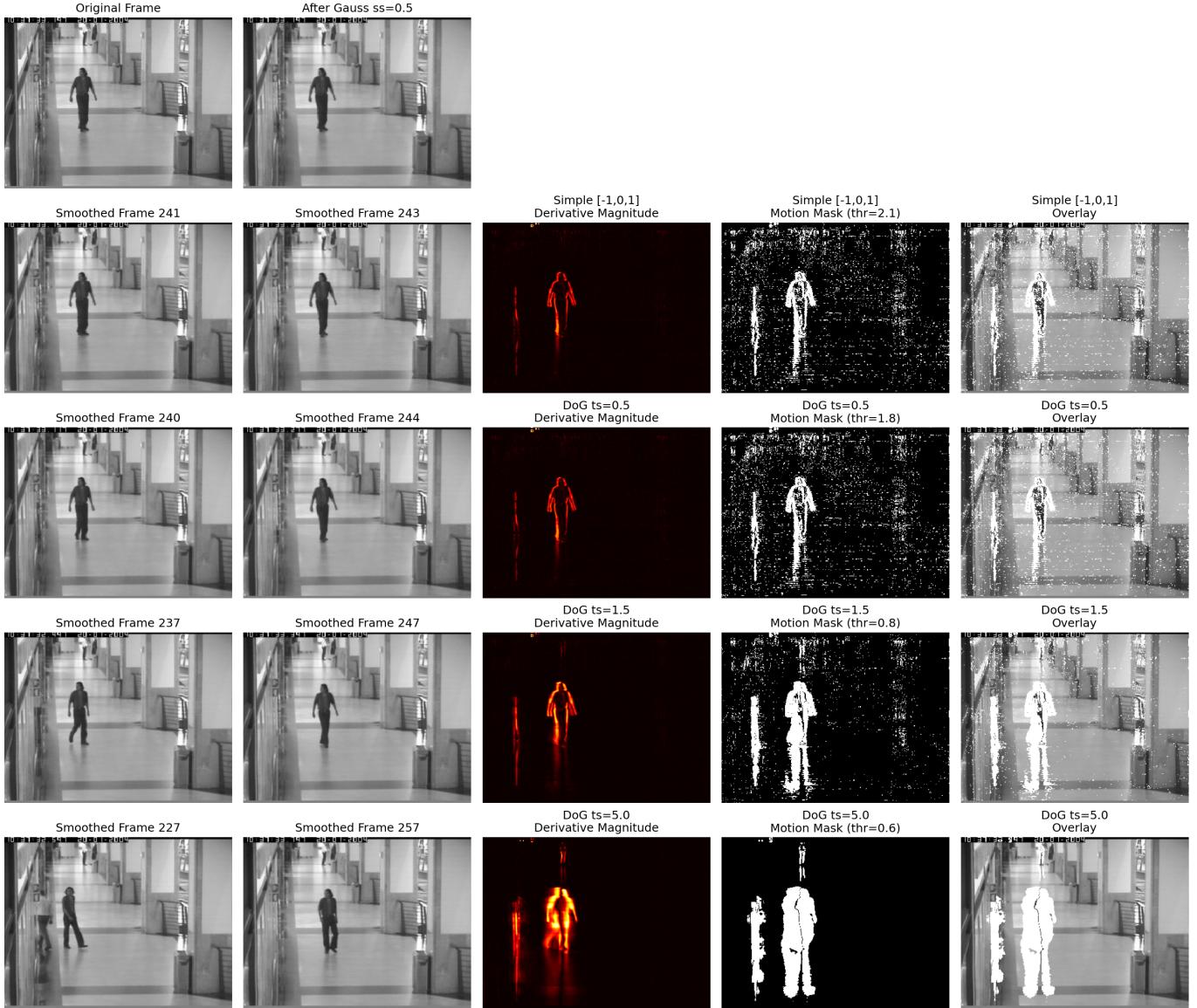


Fig. 4. Effect of applying Gaussian spatial smoothing with $s_\sigma = 0.5$ prior to temporal differentiation. The first row shows the original frame and its smoothed version. Subsequent rows display the neighboring smoothed frames used for derivative computation, the temporal derivative magnitude (for the simple central difference and DoG filters with increasing t_σ), the resulting motion masks, and the corresponding overlays. The small Gaussian smoothing slightly reduces high-frequency noise while preserving object boundaries, yielding results comparable to the 3×3 box filter but with smoother attenuation characteristics.

percentiles produce scattered background detections, whereas very high percentiles overly thin the silhouette. At $p = 90$, the configuration achieves Motion = 10% and LargestCC = 0.648, yielding a compact and spatially coherent mask without excessive erosion.

Across all tested configurations, the 90th percentile consistently provides a favorable trade-off between motion extent and spatial coherence. Lower percentiles tend to over-segment the scene and inflate the motion region, while higher percentiles, although numerically increasing LargestCC, remove weaker but meaningful motion gradients and reduce structural completeness. For this reason, $p = 90$ is adopted as the reference threshold in the subsequent analysis, ensuring consistent

comparison across spatial and temporal filtering strategies.

c) Adaptive Noise-Model Threshold:: The adaptive noise-model threshold defines the decision boundary as $T = k \sigma_{\text{noise}}$, where σ_{noise} is estimated from the background distribution of derivative magnitudes. Figures 13–16 illustrate how increasing k shifts the threshold toward larger derivative values, progressively reducing the number of detected motion pixels.

For the simple central-difference operator, the quantitative comparison in Table III highlights the effect of k . When $k = 2$, the detected motion percentage increases to 37.45%, far above the percentile-based baseline (14.29% at $p = 80$ in Table I), and LargestCC drops to 0.111, indicating severe over-segmentation and strong background contamination. This

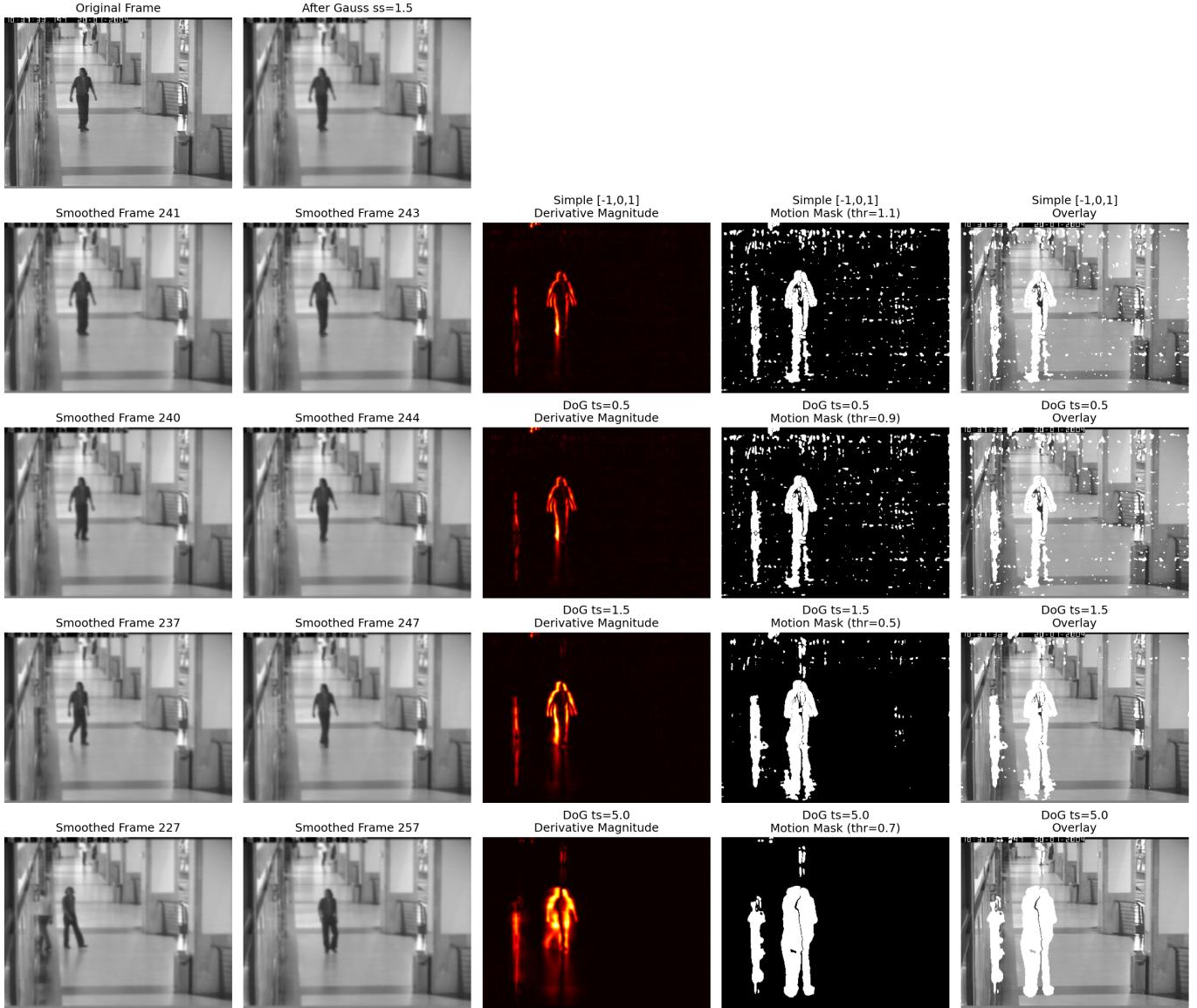


Fig. 5. Effect of applying Gaussian spatial smoothing with $s_\sigma = 1.5$ prior to temporal differentiation. The first row shows the original frame and its smoothed version. Subsequent rows display the neighboring smoothed frames used for derivative computation, the temporal derivative magnitude (for the simple central difference and DoG filters with increasing t_σ), the resulting motion masks, and the corresponding overlays. Compared to weaker spatial smoothing, $s_\sigma = 1.5$ provides substantial noise reduction while preserving the main object structure, yielding cleaner and more spatially coherent motion masks without excessive boundary thickening.

behavior is clearly visible in Fig. 13, where the mask contains large scattered background regions.

As k increases to 3 and 4, the detected motion percentage (14.29% and 9.11%) and LargestCC values (0.174 and 0.239) become numerically equivalent to the percentile-based thresholds at $p = 85$ and $p = 90$, respectively (Table I). This confirms that the adaptive strategy effectively reproduces the behavior of percentile thresholding when k is properly selected. Visually, intermediate values ($k \approx 3\text{--}4$) suppress background noise while preserving the main silhouette.

For $k = 5$, the motion percentage decreases to 4.75% and LargestCC increases to 0.363, matching the percentile result

at $p = 95$. As shown in Fig. 13, the mask becomes thinner and cleaner, but weaker motion gradients begin to disappear. Thus, increasing k improves spatial coherence but reduces structural completeness, mirroring the percentile trend.

The same qualitative behavior is observed for the DoG filters in Figs. 14–16. However, because DoG derivatives exhibit higher SNR values (Table I), the adaptive threshold produces more spatially coherent masks even for moderate k . For example, with $t_\sigma = 1.5$ or $t_\sigma = 5.0$, derivative responses are more concentrated around zero, and increasing k effectively suppresses background fluctuations while maintaining compact motion regions.

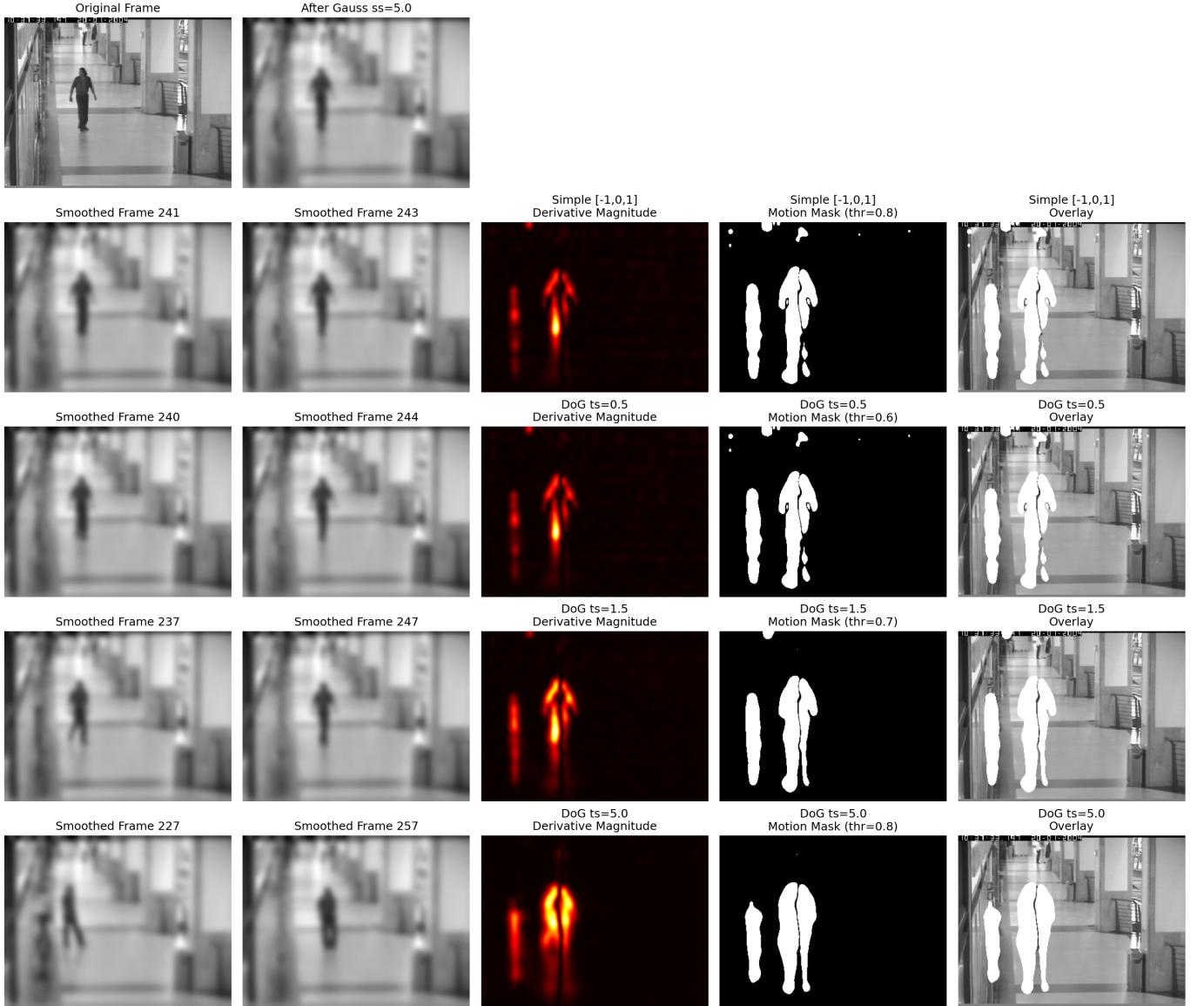


Fig. 6. Effect of applying Gaussian spatial smoothing with $s_\sigma = 5.0$ prior to temporal differentiation. The first row shows the original frame and its heavily smoothed version. Subsequent rows display the smoothed neighboring frames used for derivative computation, the temporal derivative magnitude (for the simple central difference and DoG filters with increasing t_σ), the resulting motion masks, and the corresponding overlays. Strong spatial smoothing significantly suppresses background noise; however, it introduces substantial blurring, enlarges object silhouettes, and removes fine structural details, leading to thicker and less precise motion boundaries.

d) Comparison:: Figure 17 illustrates the qualitative differences between the three thresholding strategies, while Table III summarizes their quantitative behavior.

Fixed thresholding is the least robust strategy. Its performance depends directly on the absolute scale of the derivative response, making it highly sensitive to parameter choice. As the threshold increases, motion detections rapidly shrink and spatial coherence varies unpredictably. Because the threshold is not normalized to the data distribution, the method does not generalize well across different temporal filters or smoothing configurations. This instability is clearly visible in Fig. 17, where small changes in the threshold significantly alter the

mask structure.

Percentile-based thresholding improves stability by adapting the decision boundary to the global distribution of derivative magnitudes. Increasing the percentile produces a smooth and monotonic transition from thicker, noisier masks to thinner and cleaner silhouettes. This behavior yields more consistent spatial coherence across configurations. However, the method enforces a predefined proportion of active pixels, which may not necessarily correspond to the true motion content of the scene.

The adaptive noise-model threshold provides a statistically grounded alternative. By estimating background variance and

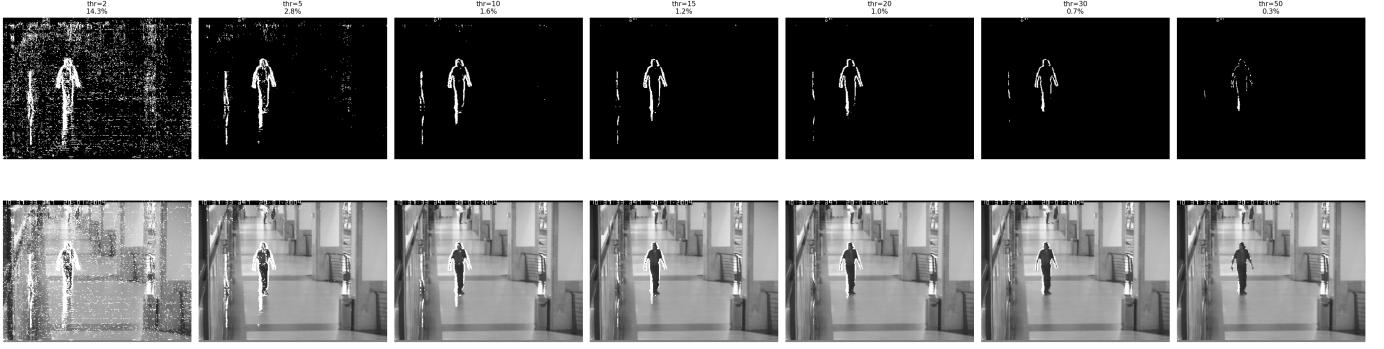


Fig. 7. Effect of varying fixed thresholds applied to the simple central difference derivative $[-1, 0, 1]$. The first row shows the resulting binary motion masks for thresholds $T = \{2, 5, 10, 15, 20, 30, 50\}$, with the percentage of detected motion pixels indicated above each image. The second row presents the corresponding overlays on the original frame. Low thresholds produce noisy masks with significant background detections, whereas high thresholds progressively suppress background responses but eventually remove relevant motion details, yielding thinner and partially fragmented silhouettes.

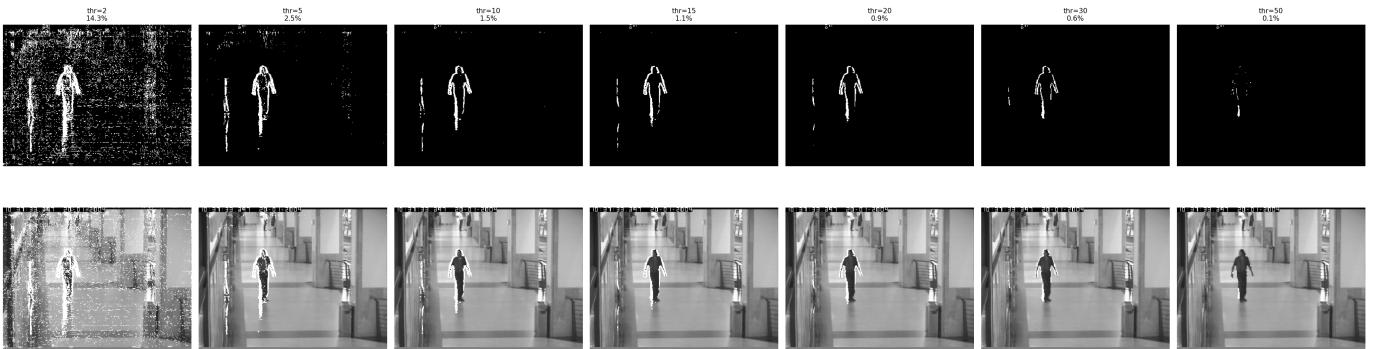


Fig. 8. Fixed-threshold analysis for the DoG temporal derivative with $t_\sigma = 0.5$. The top row shows the binary motion masks obtained for thresholds $T = \{2, 5, 10, 15, 20, 30, 50\}$, with the percentage of detected motion pixels indicated above each mask. The bottom row presents the corresponding overlays on the original frame. Small thresholds result in substantial background detections due to residual noise, while moderate thresholds ($T \approx 10–20$) produce compact and well-defined motion regions. Excessively large thresholds remove weaker gradient responses, leading to thinner silhouettes and partial loss of motion details.

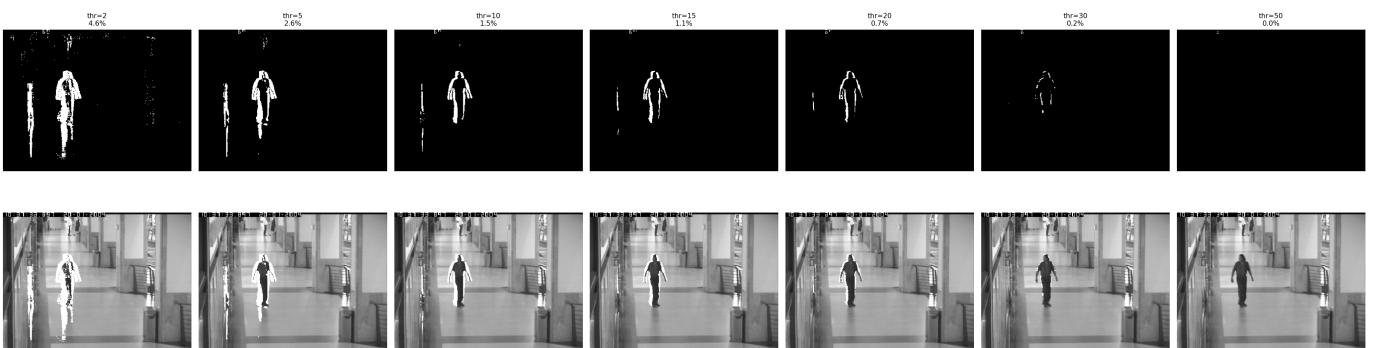


Fig. 9. Fixed-threshold sweep for the DoG temporal derivative with $t_\sigma = 1.5$. The top row shows the binary motion masks obtained for thresholds $T = \{2, 5, 10, 15, 20, 30, 50\}$, with the percentage of detected motion pixels indicated above each mask. The bottom row presents the corresponding overlays on the original frame. Compared to smaller temporal scales, the derivative response is more spatially coherent and less noisy. Moderate thresholds ($T \approx 5–15$) yield well-defined motion regions, whereas large thresholds progressively suppress weaker gradients and lead to thinning and partial loss of object boundaries.

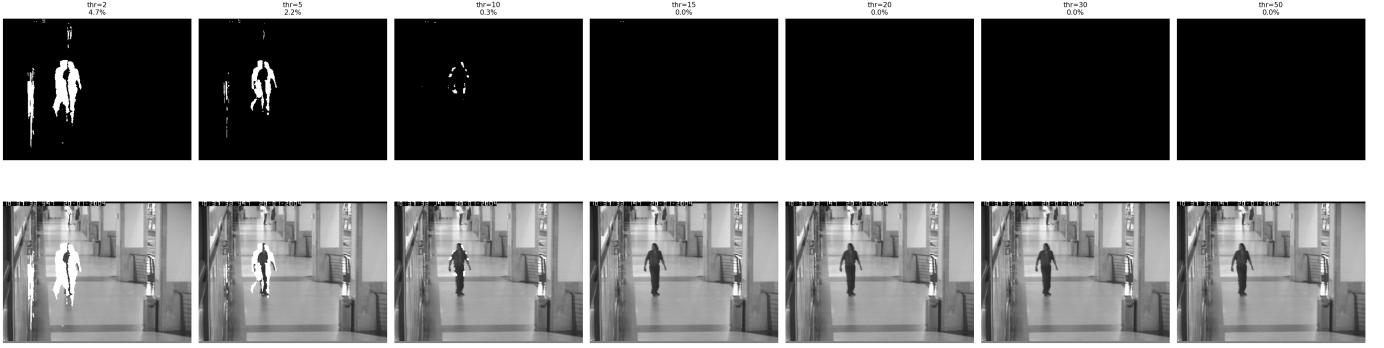


Fig. 10. Fixed-threshold sweep for the DoG temporal derivative with $t_\sigma = 5.0$. The top row shows the binary motion masks for thresholds $T = \{2, 5, 10, 15, 20, 30, 50\}$, with the percentage of detected motion pixels indicated above each mask. The bottom row presents the corresponding overlays. Due to the strong temporal smoothing, the derivative response is highly concentrated on the main motion region, resulting in compact masks even for low thresholds. However, increasing the threshold rapidly suppresses weaker gradient components, and for $T \geq 10$ the detected motion region becomes progressively reduced or disappears entirely.

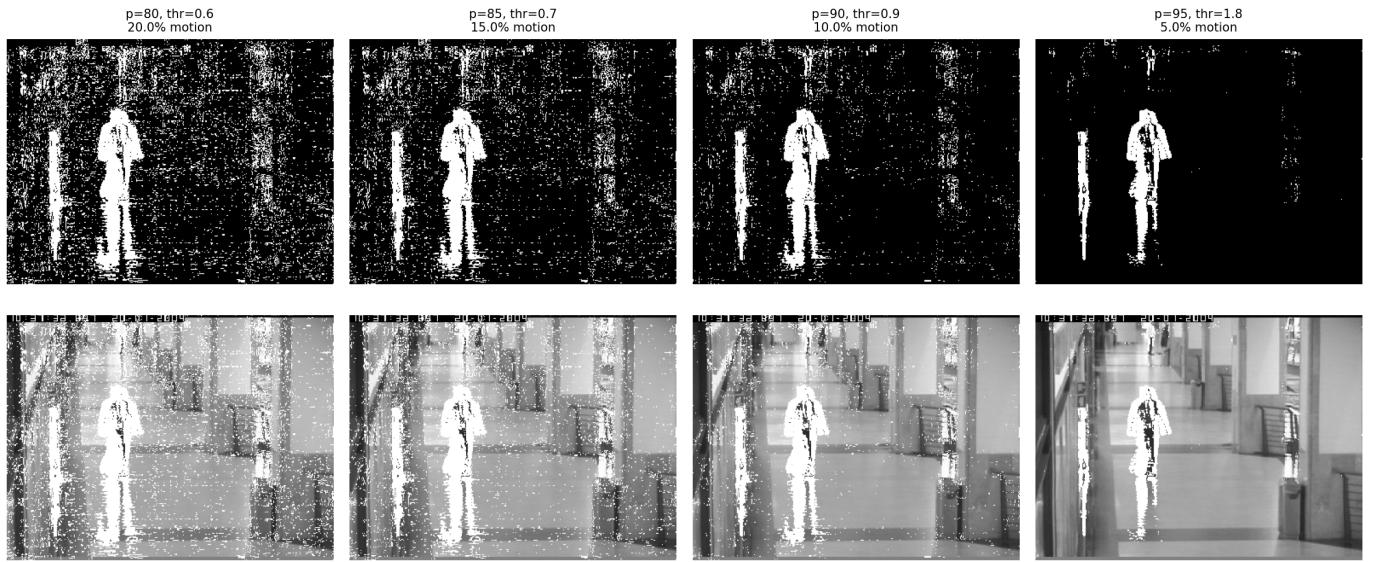


Fig. 11. Effect of varying the percentile threshold for motion detection using the derivative-of-Gaussian temporal filter ($\sigma_t = 1.5$) without spatial smoothing. Columns correspond to percentile values $p = 80, 85, 90$, and 95 . The top row shows the resulting binary motion masks, and the bottom row displays the corresponding overlays on the original frame. Lower percentiles (e.g., $p = 80$) produce dense and noisy detections, while higher percentiles (e.g., $p = 95$) yield cleaner but progressively thinner motion regions. The intermediate value $p = 90$ provides a balanced trade-off between background suppression and silhouette preservation.

defining the threshold as $T = k\sigma_{\text{noise}}$, it ties the decision boundary directly to noise statistics rather than to an arbitrary pixel proportion or fixed magnitude. Qualitatively, intermediate values of k yield compact and coherent motion regions comparable to well-chosen percentiles. The key distinction lies in interpretation: while percentile thresholding controls the fraction of detected pixels, the noise-model approach controls the confidence level of deviation from background noise.

These results show that distribution-aware methods (percentile and noise-model) produce more stable and consistent behavior than fixed thresholds. Among them, the adaptive noise-model offers greater interpretability and a principled statistical foundation, making it the most reliable strategy across varying temporal filtering scales.

IV. CONCLUSION

This work investigated a temporal gradient-based framework for moving object detection in image sequences acquired by a stationary camera. Motion was estimated through pixel-wise temporal derivatives using both a simple central finite-difference operator and derivative-of-Gaussian (DoG) filters at multiple temporal scales. The influence of spatial smoothing and threshold selection strategies was systematically evaluated using quantitative metrics (SNR and LargestCC) and qualitative visual inspection.

The results clearly show that temporal smoothing is essential for stable motion estimation. The simple central-difference operator produces sharp but highly unstable responses (SNR = 31.15 without spatial smoothing), leading to fragmented

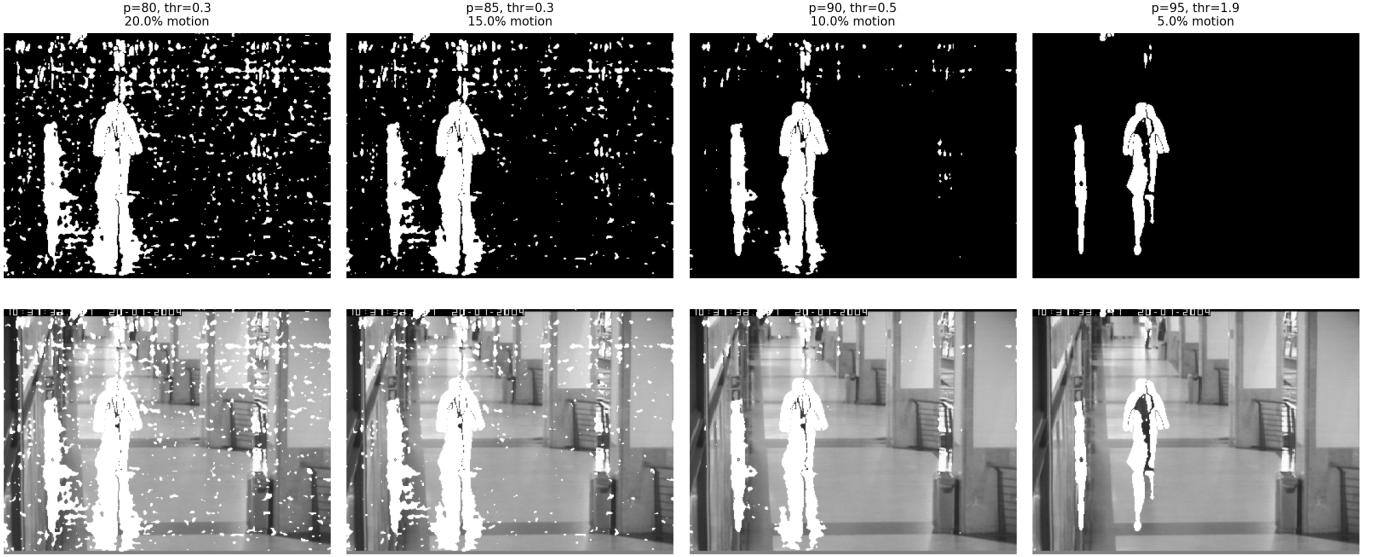


Fig. 12. Influence of the percentile threshold on motion detection using Gaussian spatial smoothing ($\sigma_s = 1.5$) followed by a derivative-of-Gaussian temporal filter ($\sigma_t = 1.5$). Columns correspond to percentile values $p = 80, 85, 90$, and 95 . The top row shows the binary motion masks, while the bottom row presents the corresponding overlays on the original frame. Lower percentiles ($p = 80, 85$) lead to over-segmentation with scattered background responses, whereas higher percentiles ($p = 95$) suppress most background noise but begin to erode weaker parts of the silhouette. The intermediate choice $p = 90$ provides the most balanced result, preserving object structure while maintaining strong background suppression.

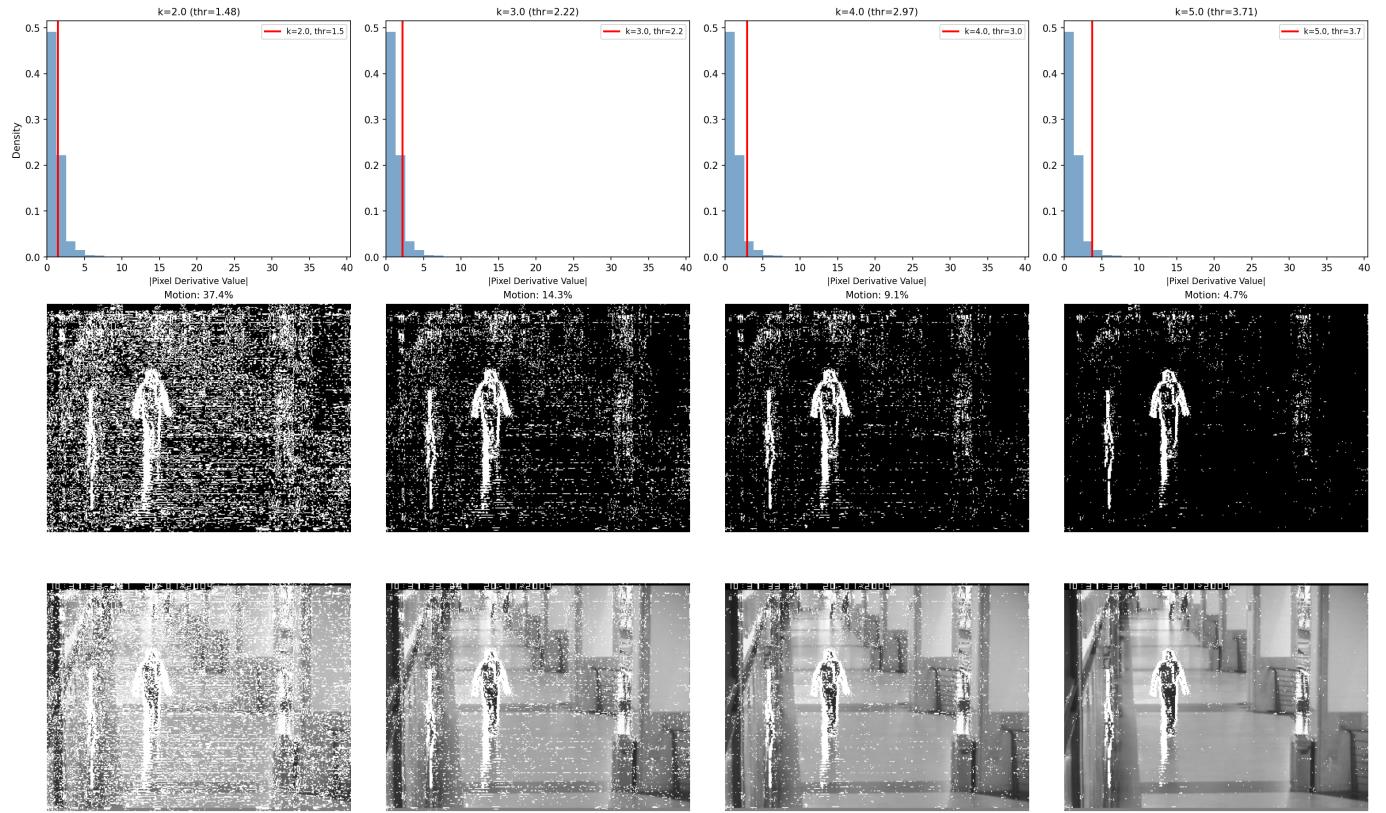


Fig. 13. Adaptive noise-model thresholding applied to the simple central-difference temporal derivative. The first row shows the histogram of the absolute derivative values, with the estimated threshold $T = k \sigma_{\text{noise}}$ indicated by the red vertical line for $k = \{2, 3, 4, 5\}$. The second row presents the corresponding binary motion masks, and the third row shows the overlays on the original frame. For small k values, a large portion of background pixels is incorrectly classified as motion. Increasing k progressively suppresses noise and reduces the detected motion percentage, yielding cleaner masks. Intermediate values ($k \approx 3\text{--}4$) provide a suitable compromise between noise reduction and preservation of the moving object's silhouette.

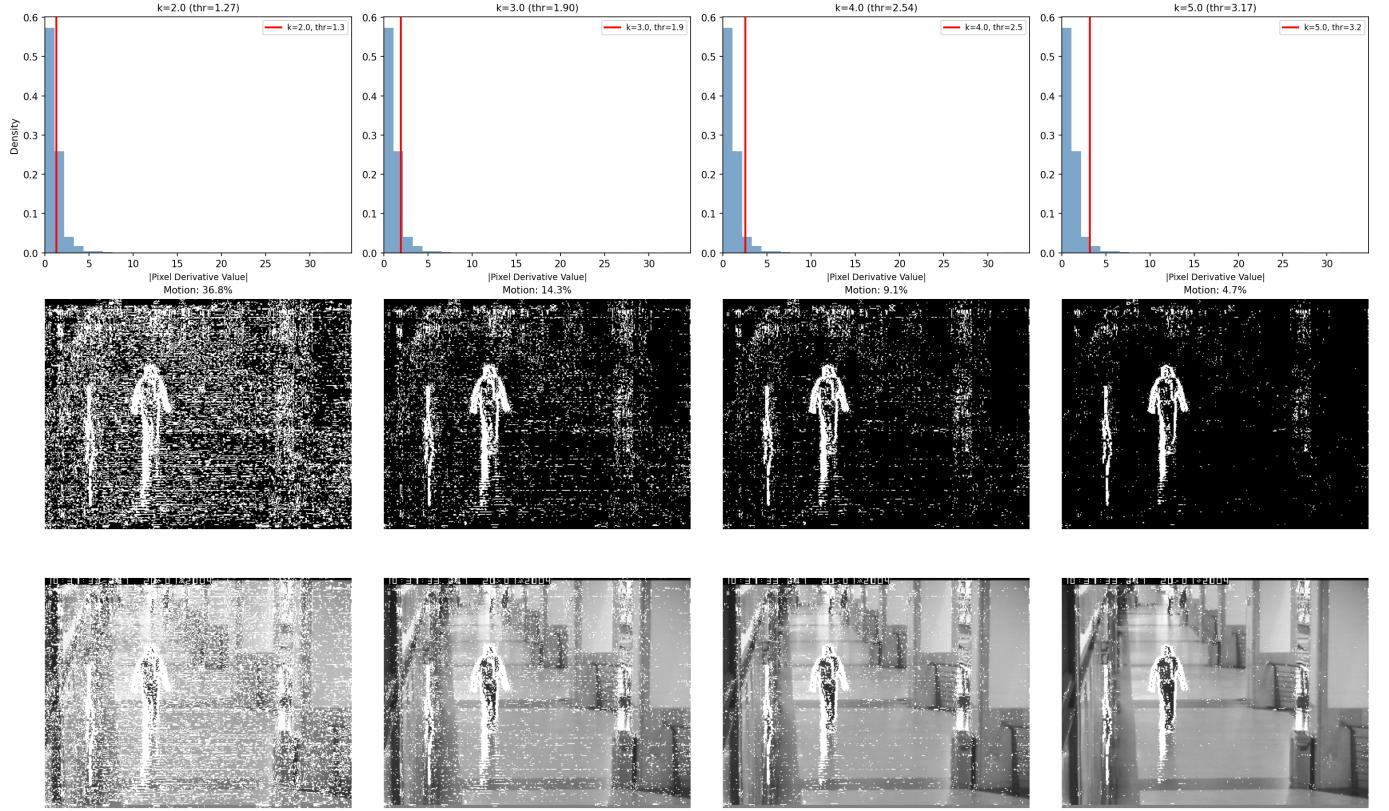


Fig. 14. Adaptive noise-model thresholding applied to the DoG temporal derivative with $t_\sigma = 0.5$. The first row shows the histogram of the absolute derivative values, where the threshold $T = k \sigma_{\text{noise}}$ is indicated by the red vertical line for $k = \{2, 3, 4, 5\}$. The second row presents the corresponding motion masks, and the third row shows the overlays on the original frame. Increasing k progressively suppresses background noise and reduces the percentage of detected motion pixels. Moderate values ($k \approx 3\text{--}4$) provide improved separation between the moving object and the background while preserving the main silhouette structure.

TABLE I
INFLUENCE OF PERCENTILE THRESHOLD ON TEMPORAL DERIVATIVES WITHOUT SPATIAL SMOOTHING.

Filter	Pctl	Thr	Motion (%)	SNR	LCC
Simple [-1,0,1]	80	2.00	14.29	31.15	0.174
Simple [-1,0,1]	85	2.00	14.29	31.15	0.174
Simple [-1,0,1]	90	2.50	9.11	31.15	0.239
Simple [-1,0,1]	95	3.50	4.75	31.15	0.363
DoG ts=0.5	80	1.70	19.38	41.55	0.147
DoG ts=0.5	85	1.71	14.95	41.55	0.182
DoG ts=0.5	90	2.14	9.89	41.55	0.243
DoG ts=0.5	95	2.99	5.00	41.55	0.378
DoG ts=1.5	80	0.61	20.00	112.85	0.273
DoG ts=1.5	85	0.73	15.00	112.85	0.341
DoG ts=1.5	90	0.94	10.00	112.85	0.467
DoG ts=1.5	95	1.78	5.00	112.85	0.663
DoG ts=5.0	80	0.18	20.00	267.51	0.781
DoG ts=5.0	85	0.27	15.00	267.51	0.651
DoG ts=5.0	90	0.62	10.00	267.51	0.731
DoG ts=5.0	95	1.83	5.00	267.51	0.418

TABLE II
QUANTITATIVE RESULTS FOR BOX 3×3 SPATIAL SMOOTHING WITH DIFFERENT TEMPORAL DERIVATIVES AND PERCENTILE THRESHOLDS. METRICS INCLUDE MOTION PERCENTAGE, DERIVATIVE SNR, AND LARGEST CONNECTED COMPONENT RATIO (LCC)

Temporal	Pctl	Thr	Motion (%)	SNR	LCC
Simple [-1,0,1]	80	1.06	20.00	81.43	0.191
Simple [-1,0,1]	85	1.22	14.99	81.43	0.240
Simple [-1,0,1]	90	1.44	10.00	81.43	0.331
Simple [-1,0,1]	95	2.22	5.00	81.43	0.520
DoG ts=1.5	80	0.36	20.00	181.41	0.343
DoG ts=1.5	85	0.43	15.00	181.41	0.436
DoG ts=1.5	90	0.58	10.00	181.41	0.575
DoG ts=1.5	95	1.70	5.00	181.41	0.673

TABLE III
THRESHOLD STRATEGY COMPARISON FOR THE SIMPLE [-1,0,1] TEMPORAL DERIVATIVE.

Strategy	Parameter	Thr	Motion (%)	SNR	LCC
Fixed	2	2.00	14.29	31.15	0.174
Fixed	5	5.00	2.81	31.15	0.542
Fixed	10	10.00	1.60	31.15	0.473
Fixed	20	20.00	1.02	31.15	0.538
Fixed	30	30.00	0.74	31.15	0.303
Percentile	80	2.00	14.29	31.15	0.174
Percentile	85	2.00	14.29	31.15	0.174
Percentile	90	2.50	9.11	31.15	0.239
Percentile	95	3.50	4.75	31.15	0.363
Noise model	$k = 2$	1.48	37.45	31.15	0.111
Noise model	$k = 3$	2.22	14.29	31.15	0.174
Noise model	$k = 4$	2.97	9.11	31.15	0.239
Noise model	$k = 5$	3.71	4.75	31.15	0.363

TABLE IV
QUANTITATIVE RESULTS FOR BOX 5×5 SPATIAL SMOOTHING WITH DIFFERENT TEMPORAL DERIVATIVES AND PERCENTILE THRESHOLDS. METRICS INCLUDE MOTION PERCENTAGE, DERIVATIVE SNR, AND LARGEST CONNECTED COMPONENT RATIO (LCC).

Temporal Filter	Percentile	Threshold	Motion (%)	SNR	LCC
Simple [-1,0,1]	80	0.78	19.99	100.76	0.217
Simple [-1,0,1]	85	0.90	14.99	100.76	0.276
Simple [-1,0,1]	90	1.12	9.97	100.76	0.387
Simple [-1,0,1]	95	1.92	5.00	100.76	0.619
DoG $t_\sigma = 0.5$	80	0.66	20.00	101.71	0.218
DoG $t_\sigma = 0.5$	85	0.77	15.00	101.71	0.277
DoG $t_\sigma = 0.5$	90	0.96	10.00	101.71	0.387
DoG $t_\sigma = 0.5$	95	1.65	5.00	101.71	0.620
DoG $t_\sigma = 1.5$	80	0.27	20.00	226.66	0.379
DoG $t_\sigma = 1.5$	85	0.33	15.00	226.66	0.480
DoG $t_\sigma = 1.5$	90	0.49	10.00	226.66	0.640
DoG $t_\sigma = 1.5$	95	1.86	5.00	226.66	0.668
DoG $t_\sigma = 5.0$	80	0.11	20.00	570.16	0.935
DoG $t_\sigma = 5.0$	85	0.24	15.00	570.16	0.675
DoG $t_\sigma = 5.0$	90	0.66	10.00	570.16	0.734
DoG $t_\sigma = 5.0$	95	1.83	5.00	570.16	0.419

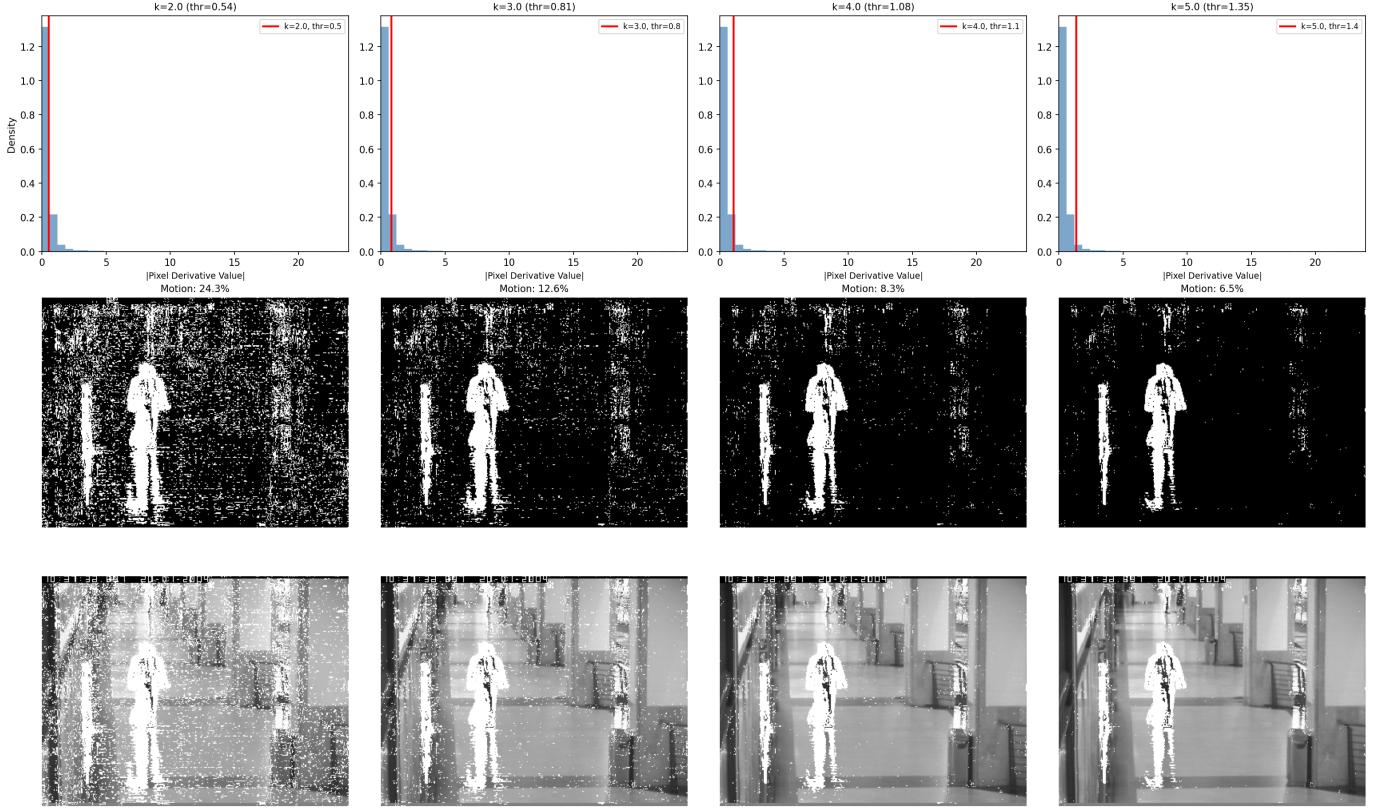


Fig. 15. Adaptive noise-model thresholding applied to the DoG temporal derivative with $t_\sigma = 1.5$. The first row shows the histogram of the absolute derivative values, with the threshold $T = k \sigma_{\text{noise}}$ indicated by the red vertical line for $k = \{2, 3, 4, 5\}$. The second row presents the resulting binary motion masks, and the third row shows the overlays on the original frame. As k increases, the threshold becomes more selective, reducing the percentage of detected motion pixels and suppressing background fluctuations. Compared to lower temporal scales, the derivative response is more spatially coherent, and intermediate values ($k \approx 3\text{--}4$) provide clean object delineation while preserving the main silhouette structure.

TABLE V
QUANTITATIVE RESULTS FOR GAUSSIAN SPATIAL SMOOTHING ($\sigma_s = 0.5$) WITH DIFFERENT TEMPORAL DERIVATIVES AND PERCENTILE THRESHOLDS.

Temporal Filter	Percentile	Threshold	Motion (%)	SNR	LCC
Simple [-1,0,1]	80	1.54	20.00	58.74	0.157
Simple [-1,0,1]	85	1.77	15.00	58.74	0.185
Simple [-1,0,1]	90	2.13	10.00	58.74	0.254
Simple [-1,0,1]	95	2.97	5.00	58.74	0.409
DoG $t_\sigma = 0.5$	80	1.31	20.00	59.12	0.158
DoG $t_\sigma = 0.5$	85	1.51	15.00	59.12	0.185
DoG $t_\sigma = 0.5$	90	1.82	10.00	59.12	0.255
DoG $t_\sigma = 0.5$	95	2.53	5.00	59.12	0.410
DoG $t_\sigma = 1.5$	80	0.50	20.00	133.53	0.302
DoG $t_\sigma = 1.5$	85	0.60	15.00	133.53	0.373
DoG $t_\sigma = 1.5$	90	0.78	10.00	133.53	0.504
DoG $t_\sigma = 1.5$	95	1.71	5.00	133.53	0.694
DoG $t_\sigma = 5.0$	80	0.15	20.00	321.06	0.829
DoG $t_\sigma = 5.0$	85	0.25	15.00	321.06	0.665
DoG $t_\sigma = 5.0$	90	0.63	10.00	321.06	0.731
DoG $t_\sigma = 5.0$	95	1.82	5.00	321.06	0.418

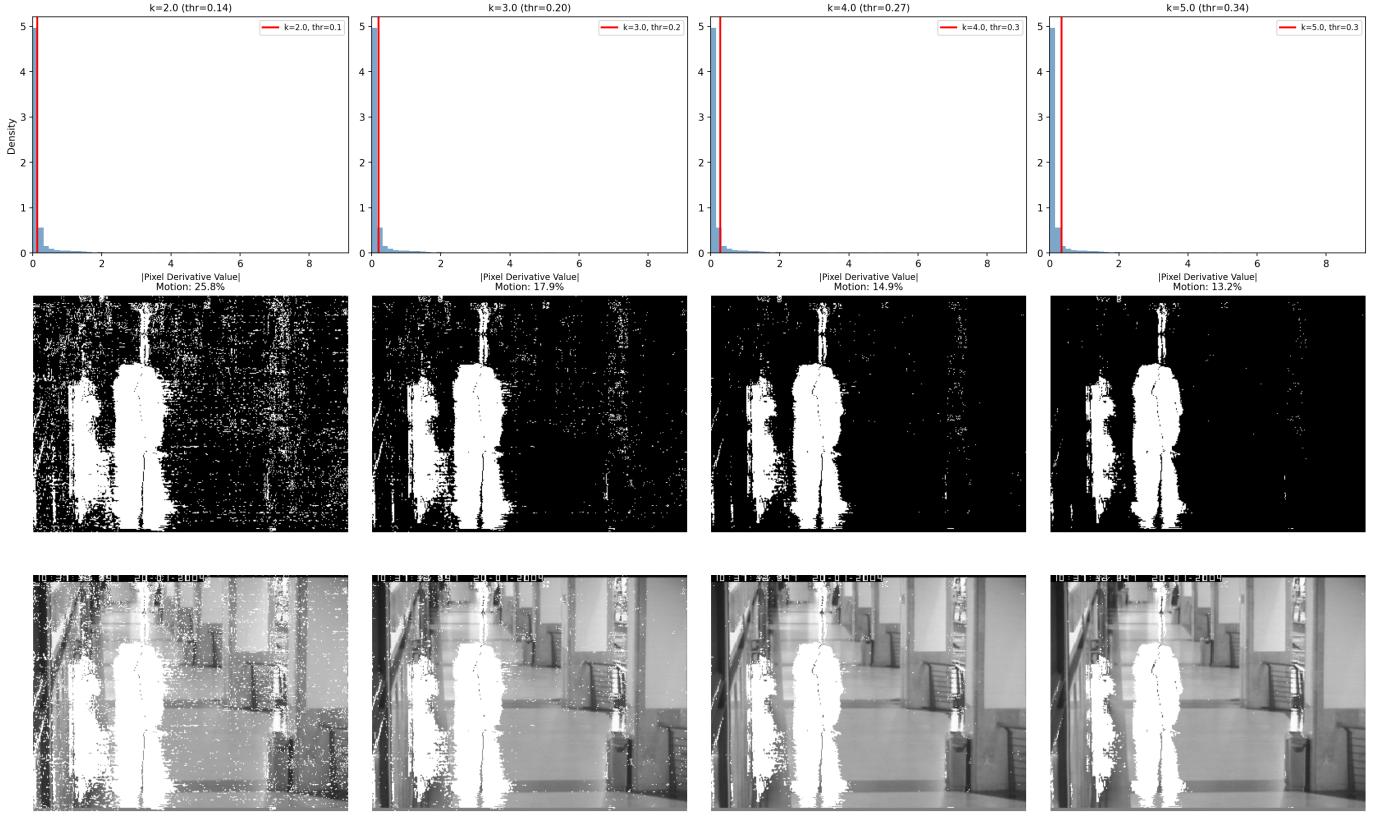


Fig. 16. Adaptive noise-model thresholding applied to the DoG temporal derivative with $t_\sigma = 5.0$. The first row shows the histogram of the absolute derivative values, with the threshold $T = k \sigma_{\text{noise}}$ indicated by the red vertical line for $k = \{2, 3, 4, 5\}$. The second row presents the corresponding binary motion masks, and the third row shows the overlays on the original frame. Due to the strong temporal smoothing, derivative values are more concentrated near zero, resulting in relatively low threshold magnitudes. As k increases, background responses are progressively suppressed; however, excessive selectivity may attenuate weaker motion components and slightly reduce the spatial extent of the detected object. Moderate values of k provide compact and coherent motion regions while maintaining the main silhouette structure.

masks and strong sensitivity to background fluctuations. In contrast, the DoG operator significantly improves robustness as the temporal scale increases. Moderate temporal smoothing ($t_\sigma \approx 1.5$) provides a strong improvement in signal-to-noise separation ($\text{SNR} \approx 112.85$ without spatial smoothing and above 230 when combined with moderate spatial filtering) while preserving object structure. Very large temporal scales ($t_\sigma = 5.0$) further increase SNR but may attenuate fine motion details and reduce structural consistency under aggressive thresholding.

Spatial smoothing consistently enhances robustness by suppressing high-frequency background noise. However, the quantitative analysis shows that the improvement is not strictly monotonic with increasing spatial scale. Moderate Gaussian smoothing ($s_\sigma \approx 1.5$) yields the most balanced behavior, achieving high SNR values (above 230 with $t_\sigma = 1.5$) and strong spatial coherence (LargestCC ≈ 0.65 at $p = 90$) without excessive boundary thickening. Strong spatial smoothing ($s_\sigma = 5.0$) produces very high SNR values (up to 717.89 when combined with $t_\sigma = 5.0$) but may reduce spatial consistency, reflecting over-smoothing effects.

Regarding threshold selection, fixed thresholds proved highly sensitive to parameter choice and did not generalize across

derivative magnitudes. Percentile-based thresholding provided consistent and predictable control of the detected motion percentage, producing monotonic changes in spatial coherence as the percentile increased. The adaptive noise-model thresholding strategy offered a statistically grounded alternative by linking the decision boundary to estimated background variance. Quantitatively, it reproduced the behavior of percentile-based thresholds when properly tuned, while offering improved interpretability through its explicit noise-based formulation.

Overall, the most reliable performance was obtained by combining moderate Gaussian spatial smoothing ($s_\sigma \approx 1.5$), DoG temporal differentiation at a moderate scale ($t_\sigma \approx 1.5$), and an adaptive threshold tied to background noise statistics. These findings demonstrate that effective gradient-based motion detection requires careful coordination between temporal filtering, spatial preprocessing, and statistically principled thresholding. The study highlights the importance of analyzing these components jointly rather than in isolation when designing robust motion detection systems.

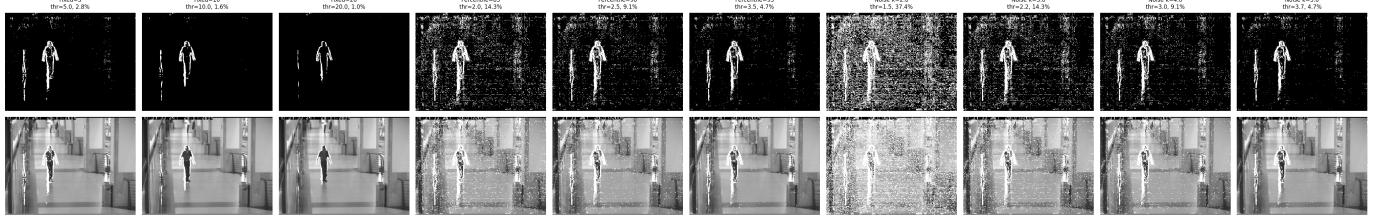


Fig. 17. Comparison of threshold selection strategies for the simple central-difference temporal derivative. The first row shows the binary motion masks obtained using fixed thresholds ($T = \{5, 10, 20\}$), percentile thresholds (85%, 90%, 95%), and noise-model thresholds ($k = \{2, 3, 4, 5\}$). The second row presents the corresponding overlays on the original frame. Fixed thresholding exhibits strong sensitivity to the selected value; percentile-based thresholding adapts to the derivative distribution but may retain background noise at lower percentiles; and the noise-model-based approach provides more statistically grounded and spatially coherent motion segmentation for intermediate k values.

TABLE VI

QUANTITATIVE RESULTS FOR GAUSSIAN SPATIAL SMOOTHING ($\sigma_s = 1.5$) WITH DIFFERENT TEMPORAL DERIVATIVES AND PERCENTILE THRESHOLDS.

Temporal Filter	Percentile	Threshold	Motion (%)	SNR	LCC
Simple [-1,0,1]	80	0.74	20.00	105.36	0.256
Simple [-1,0,1]	85	0.87	15.00	105.36	0.323
Simple [-1,0,1]	90	1.09	10.00	105.36	0.402
Simple [-1,0,1]	95	1.94	5.00	105.36	0.629
DoG $t_\sigma = 0.5$	80	0.63	20.00	106.11	0.257
DoG $t_\sigma = 0.5$	85	0.74	15.00	106.11	0.324
DoG $t_\sigma = 0.5$	90	0.93	10.00	106.11	0.404
DoG $t_\sigma = 0.5$	95	1.66	5.00	106.11	0.630
DoG $t_\sigma = 1.5$	80	0.27	20.00	231.07	0.403
DoG $t_\sigma = 1.5$	85	0.33	15.00	231.07	0.498
DoG $t_\sigma = 1.5$	90	0.49	10.00	231.07	0.648
DoG $t_\sigma = 1.5$	95	1.86	5.00	231.07	0.676
DoG $t_\sigma = 5.0$	80	0.11	20.00	576.39	0.935
DoG $t_\sigma = 5.0$	85	0.24	15.00	576.39	0.692
DoG $t_\sigma = 5.0$	90	0.66	10.00	576.39	0.735
DoG $t_\sigma = 5.0$	95	1.82	5.00	576.39	0.421

TABLE VII

QUANTITATIVE RESULTS FOR GAUSSIAN SPATIAL SMOOTHING ($\sigma_s = 5.0$) WITH DIFFERENT TEMPORAL DERIVATIVES AND PERCENTILE THRESHOLDS.

Temporal Filter	Percentile	Threshold	Motion (%)	SNR	LCC
Simple [-1,0,1]	80	0.39	20.00	121.29	0.371
Simple [-1,0,1]	85	0.47	15.00	121.29	0.458
Simple [-1,0,1]	90	0.75	10.00	121.29	0.597
Simple [-1,0,1]	95	2.57	5.00	121.29	0.486
DoG $t_\sigma = 0.5$	80	0.33	20.00	122.07	0.371
DoG $t_\sigma = 0.5$	85	0.40	15.00	122.07	0.459
DoG $t_\sigma = 0.5$	90	0.65	10.00	122.07	0.597
DoG $t_\sigma = 0.5$	95	2.21	5.00	122.07	0.486
DoG $t_\sigma = 1.5$	80	0.19	20.00	229.77	0.818
DoG $t_\sigma = 1.5$	85	0.27	15.00	229.77	0.649
DoG $t_\sigma = 1.5$	90	0.72	10.00	229.77	0.682
DoG $t_\sigma = 1.5$	95	2.53	5.00	229.77	0.480
DoG $t_\sigma = 5.0$	80	0.14	20.00	717.89	0.971
DoG $t_\sigma = 5.0$	85	0.29	15.00	717.89	0.702
DoG $t_\sigma = 5.0$	90	0.76	10.00	717.89	0.445
DoG $t_\sigma = 5.0$	95	1.89	5.00	717.89	0.451

REFERENCES

- [1] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.
- [2] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 4th ed. Pearson, 2018.
- [3] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.
- [4] P. J. Huber, *Robust Statistics*. John Wiley & Sons, 1981.
- [5] P. J. Rousseeuw and C. Croux, "Alternatives to the median absolute deviation," *Journal of the American Statistical Association*, vol. 88, no. 424, pp. 1273–1283, 1993.