Joint NLOS Object Detection Using Extraction of Moving Edges *

Aleksej MAKAROV

Laboratoire de Traitement des Signaux
Ecole Polytechnique Fédérale de Lausanne
1015 Ecublens, Suisse

RÉSUMÉ

Dans cet article nous présentons une méthode de détection d'intrus basée sur le traitement d'images. Les coûts de calcul et de mémoire impliqués par l'algorithme sont très bas. L'approche suggérée consiste à séparer les frontières de l'objet en mouvement de celles de l'arrière-plan. Le nombre de points de frontières est calculé pour chaque image. Un signal unidimensionnel est ainsi obtenu à partir d'une séquence d'images. Les changements significatifs sont détectés en analysant l'enveloppe du signal.

ABSTRACT

In this article we present an image processing based method for intrusion detection. The algorithm is characterized by a very low computational cost and limited memory requirements. The suggested approach consists in separating the moving object edges from the background. The number of edge-points is computed for each processed frame. A one-dimensional signal is thus obtained from a sequence of images. The changes of interest in this one-dimensional signal are detected by analyzing its envelope.

1 Introduction

Image processing based intrusion detection devices can store or transmit frames suspected to contain intrusions. The excessive computational and memory loads inherent to image processing make the standard digital signal processors (DSP) hardly appropriate for this application.

An intrusion detection algorithm should allow a sufficiently frequent checkup of the surveyed scene. If it is to be executed by a relatively cheap DSP, its computational cost should be very low.

The problem of detection is simplified when dealing with images provided by a fixed camera. Thus we can compare the current input image with a stored reference image of the empty scene (background). Lighting changes are the main cause of false alarms in image based intrusion detection. Edges are much more robust against lighting changes than luminance [1, 2]. An object can be detected if the current input image contains some coherent edges which were not present in the stored reference background. An important onedimensional parameter that describes the change in the frame is therefore the number of pixels assigned to the edges that have not been detected previously in the surveyed scene. A fixed background video sequence is thus converted into a onedimensional signal, which avoids large computational and memory costs. Low values of this signal correspond to the absence of intruding objects, while high values should signal an intrusion.

In section 2 we present a way to extract this signal and give a description of various kinds of noise appearing in this case. In section 3 we introduce the Joint Nonlinear Order Statistics as a tool to process this signal. The results on real sequences are shown in section 4. Finally, the conclusion is given in section 5.

*This work was supported by CERS grant No 2456.1

2 Intrusion signal and its ideal estimator

The image-processing part of this algorithm consists in a comparison of the stored edges of the fixed background with the edges extracted from the current input image. Since one frame consists of a large number of pixels, the simplest operations are used.

Edge-extraction is performed by subtracting consecutive rows (for vertical component) and columns (for the horizontal component of the gradient). The absolute values of the two gradient components are then added up and their sum is thresholded in order to obtain bilevel edge-images.

An "AND" logical operation is performed on the current bilevel gradient image and the complement of the reference gradient image input edges (see Figure 2). Hence, if an edge-point appears in the reference bacground image, and not in the current input image, it will be simply ignored. The background edge-points that coincide will annihilate each other. The resulting bilevel image will consist of the full gradient edges of the moving object (if any), the sparsed remains of the background gradient edges, and of more noise than before.

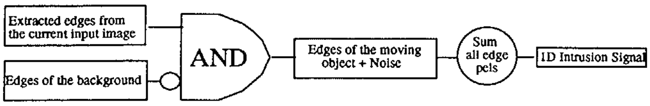

Figure 1: *Fast extraction of intrusion signals*

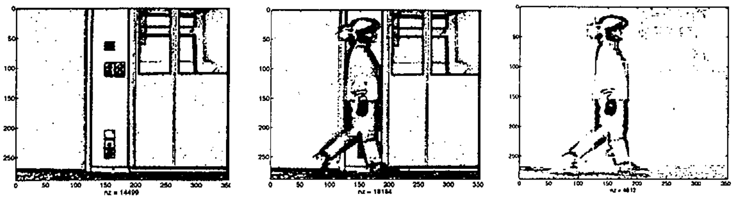

Figure 2: *Edges of the fixed background, edges of an input image with intruder, and resulting edges of the intruder, respectively.*

In the absence of intruding objects, due to the lack of relatively expensive cleaning operations (e.g., median), at the exit of the "AND" circuit in Figure 1 there will be a certain amount of noisy pels, especially the remnants of the background edges, resulting in a nonzero value of the intrusion signal. This value is hardly predictible and scene-dependent, so the comparison of the signal extracted according to Figure 1 with a fixed threshold is pointless. For convenience, the above onedimensional signal will be referred to as the intrusion signal.
A way to detect changes in intrusion signals which might signal a real intrusion would be to compare the increment of the central value of the signal with some appropriately chosen threshold. In this case, the positive increments whose magnitude is above the threshold should set off the alarm, while the corresponding negative increments should stop the alarm (under condition that the current estimate of the signal is close enough to the signal level prior to intrusion). For convenience, the aforementioned threshold will be referred to as the *incremental threshold.*
If we denote the intrusion signal by x, its increment at time t over the period Δt may be estimated as:

$$\Delta x = \mathcal{E}\left\{x\right\}_t - \mathcal{E}\left\{x\right\}_{t-\Delta t} \qquad (1)$$

where $\mathcal{E}$ holds for an estimator. Depending of the relation between the increment Δx and the incremental threshold, which will be denoted by η, we shall distinguish three events of interest:

- $-\eta \leq \Delta x < \eta$ no change
- $\Delta x \geq \eta$ OFF $\rightarrow$ ON
- $\Delta x < -\eta$ ON $\rightarrow$ OFF

where ON and OFF correspond to the alarm and the quiet state, respectively.
While the lighting changes may cause false alarms, there is a series of events closely related to the properties of the intruding object and the scene that may cause a prematurate cessation or interruptions of alarm. These large changes of the number of edge-points during an intrusion may occur when:

- flexible objects change their shapes (for example the changes in the shape of the human silhouette during the walk);
- objects change their distance from the camera;
- the local contrast between the object and the background changes;
- the objects are partially occluded.

According to the above considerations, we formulate the requirements an ideal estimator $\mathcal{E}$ should satisfy in order to detect accurately the intruding objects:

1. The quiet state (no intrusion) may be considered as a constant level corrupted with i. i. d. noise. The ideal estimator should flatten the changes in this state in order to avoid false alarms.
2. During the transitions from the quiet state to the intrusion, the ideal estimator should track the change as quickly as possible (early detection).
3. The end of the alarm should be signaled as soon as the intrusion signal returns to the quiet state.
4. The signal dependent noise during the detected intrusion (ON state) should not cause a prematurate cessation of alarm.

3 Joint Nonlinear Order Statistics

The order statistics of a set of data $\{x\}_t$

$$\{x\}_t = \{x_1, x_2, \ldots x_N\}$$

are defined as the set of the same data rearranged in ascending order:

$$x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(N)}$$

The indexes without parentheses correspond to the temporal order within the window, while the indexes between parentheses correspond to the magnitude order. The most popular NLOS estimator is the median, $x(\frac{N+1}{2})$ for odd N. The median smoothes out noisy signals and is robust to outliers. In the presence of large signal-dependent noise during an intrusion (see the signal areas under rectangle in Figure 3) and for small window-lengths (short delays) the increment Δx estimated by the median estimator can be large, causing thus interruptions or prematurate cessation of alarm. We shall show that this can be avoided if, instead of the median value, we observe the envelope of the intrusion signal.
The envelope of the intrusion signal may be formed by considering the signal extrema, $x_{(1)}$ and $x_{(N)}$, taken over a sliding window of length N. The main drawback of a single extremum is its high sensitivity to outliers. In the quiet state, for instance, a single outlier may give rise to a false alarm.
The robustness to outliers can be improved if both extrema are used simultaneously. If we denote by ΔM the difference between successive maxima:

$$\Delta M = \max\{x\}_{t+\Delta t} - \max\{x\}_t = \max_{S \leq i \leq N+S}(x_i) - \max_{1 \leq i \leq N}(x_i)$$

and by Δm the difference between successive minima:

$$\Delta m = \min\{x\}_{t+\Delta t} - \min\{x\}_t = \min_{S \leq i \leq N+S}(x_i) - \min_{1 \leq i \leq N}(x_i)$$

taken over a window of length N shifted by S samples with respect to its previous position, the Joint Extrema based Detector may be described by:

$$h(|\Delta M| - \eta) \cdot h(|\Delta m| - \eta) \cdot \delta_{\mathrm{sgn}(\Delta M),\mathrm{sgn}(\Delta m)} \cdot \mathrm{sgn}(\Delta M)$$
$$= \begin{cases} 1 & \text{OFF} \rightarrow \text{ON} \\ 0 & \text{no change} \\ -1 & \text{ON} \rightarrow \text{OFF} \end{cases} \qquad (2)$$

where h stands for the Heaviside unit step function, and δ for the Kronecker Delta symbol.

We shall show that the false alarm probability and the rate of false interruptions of alarm are very low for this detector. The most common case is when $S = 1$, i.e. when the sliding window is shifted to the following position by only one sample. If we denote by $\{x\}_t = \{x_1, x_2, \ldots x_N\}$ the data from the window at time t, at time $t + \Delta t$ x_1 will drop off and some new sample x_{N+1} will be included into the window. In order to simplify the problem, let us consider the (false) alarm probability when the intrusion signal is in the quiet state. From the definition of the Joint Extrema based Detector we see that both new minimum and new maximum should be larger than the respective extrema in the previous window. For a one-sample shift this is possible if and only if x_1 was the old minimum and x_{N+1} is the new maximum, or, pictorially:

$$\underbrace{\overbrace{x_1}^{\min\{x\}_t} \quad x_2 \quad \cdots\cdots\cdots \quad x_N}_{\text{current window} \quad \{x\}_t} \qquad \overbrace{x_{N+1}}^{\max\{x\}_{t+\Delta t}} \qquad (3)$$

This pattern is not very common in the quiet state. Under the hypothesis of i.i.d noise, the probability P_P of this event is:

$$P_P = \frac{1}{N+1} \cdot \frac{1}{N} \qquad (4)$$

This value will be referred to as the spatial arrangement factor. Once the above spatial arrangement is achieved, it is necessary that:

$$\begin{aligned} x_{(2)} - x_{(1)} &\geq \eta \quad \text{and} \\ x_{(N+1)} - x_{(N)} &\geq \eta \end{aligned} \qquad (5)$$

The joint probability density function (pdf) for $x(1), x(2), x(N)$ and $x(N+1)$ is given by [3]:

$$f_{1,2,N,N+1}(x, y, z, w) = (N+1)N(N-1)(N-2) \cdot$$
$$\cdot f(x)f(y)f(z)f(w)\left[F(z) - F(y)\right]^{N-3} \qquad (6)$$

where $f(x)$ and $F(x)$ are the pdf and the cumulative distribution function (cdf) of the signal x in the quiet (stationary) state. If we denote by ξ the difference between successive minima and by ζ the difference between successive maxima,

$$\begin{aligned} \xi &= x_{(2)} - x_{(1)} \\ \zeta &= x_{(N+1)} - x_{(N)} \end{aligned} \qquad (7)$$

and replace (x, y, z, w) in 6 by (x, ξ, z, ζ), by integrating over all possible values for x $(x_{(1)})$ and z $(x_{(N)})$, we obtain the joint pdf for extrema increments:

$$f(\xi, \zeta) = (N+1)N(N-1)(N-2) \cdot$$
$$\cdot \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x)f(x+\xi)f(z)f(z+\zeta)\left[F(z) - F(x+\xi)\right]^{N-3} \mathrm{d}x\mathrm{d}z \qquad (8)$$

The false alarm rate is equal to the product of the spatial arrangement factor P_P and the joint probability P_η that $\xi = x_{(2)} - x_{(1)} \geq \eta$ and $\zeta = x_{(N+1)} - x_{(N)} \geq \eta$:

$$P_D(\eta, f(x), N) = (N-1)(N-2) \int_{\xi=\eta}^{\infty}\int_{\zeta=\eta}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}$$
$$\left[F(z) - F(x+\xi)\right]^{N-3} f(x)f(x+\xi)f(z)f(z+\zeta)\, d\zeta\, d\xi\, dz\, dx \qquad (9)$$

During the transition from quiet state to intrusion , we need the Joint Extrema Detector to be sensitive enough in order to signal the alarm. If the window length is short enough with respect to the duration of transitions, and if the slope of the general trend of the intrusion signal is high enough with respect to the standard deviation of the noise, both factors P_P and P_η increase. The spatial arrangement factor P_P increases due to the self-ordering tendency: during a transition from a low to a high central value, the window minima are tend to be located at the left side, while the window maxima tend to be located at the right side of the window.. P_η increases with the slope of the general trend of intrusion signal, which should be close to zero in the quiet state.

The requirement that both extrema should change in the same direction for a sufficiently large value reduces the rate of false interruptions of alarm. For a sufficiently large window (so that the required spatial arrangement is difficult to achieve), this rate can be further reduced by increasing the threshold η during the alarm state.

The objections that can be made in what concerns the sensitivity of the detector is that it demands a special spatial arrangement as a necessary condition to detect intrusion. Increasing the shift improves the sensibility of the Joint Extrema Detector during transitions. Larger shifts should be used when the standard deviation of the noise is more important with respect to the slope during the transitions then in cases we tested. An ascending edge can be detected by using the sliding window with the shift S if the old minimum was among the the S dropped off samples and the new maximum among the newly inserted S samples. Or, pictorially, instead of pattern (3) we have

$$\underbrace{\overbrace{x_1\ x_2\ \ldots\ x_S}^{x_{(1)}} \quad x_{S+1} \quad \cdots\cdots\cdots \quad x_N}_{\text{current window} \quad \{x\}_t} \qquad \overbrace{x_{N+1}\ x_{N+2}\ \ldots\ x_{N+S}}^{x_{(N+S)}}$$

4 Results

The Joint Extrema based Detector was used to detect the changes corresponding to real intrusions in the intrusion signal obtained as in Figure 3. We tried several window sizes N in the range from 3 to 12 on a population of four sequences (288x352 pels, 25 frames per second) and in all cases the detection was successful. In the total of about 70 frames of these sequences in which the intruder was absent, no false alarm was detected. Several shifts (1,2,3,4,5) were tried with satisfying results. Since the window size was below the transition duration in our application, we could choose the highest precision, i.e. the 1-sample shift. The threshold in the quiet (stationary) state was set to 10 edge-pels per frame. During intrusion, it was set to 10 % of the current value of the intrusion signal. The cessation of alarm was disabled until the difference between the current value of the intrusion signal and its last value before the intrusion was detected became less than 30 % of the intrusion signal value before the alarm. In Figure 3, the first frame for which the alarm was raised, the middle frame and the last frame before the alarm is stopped are shown. The diagram below the frames represent the intrusion signal. The superposed rectangles cover the interval of the signal for which the intrusion is detected. The dotted impulses on the diagrams correspond to the detected ascending edges, while the full-line impulses mark the moments when the descending edges are detected.

5 Conclusion

In this article we presented a new intrusion detector whose computational load is very low (5 operations per pixel for pseudogradient, one logical operation per pixel in "AND" circuit, one cumulative addition in order to obtain the one-dimensional intrusion signal). Despite the simplicity of the image processing part, satisfactory results are obtained when the intrusion signal is tested by the Joint Extrema detector. The influence of signal-independent noise and slow illumination changes is discarded, while the sensitivity to the presence of physical objects is conserved.

References

[1] F.Bartollini and V.Cappellini. Automatic detection of intrusions by image processing. *Proceedings of the International Conference on Digital Signal Processing*, 2:468–471, July 1993.

[2] C. León J. Aranda and M. Frigola. A multitracking system for trajectory analysis of people in a restricted area. In V. Cappellini, editor, *Proceedings of the 4th International Workshop on Time-Varying Image Processing and Moving Object Recognition*, number 3, pages 118–124, Florence, Italy, June 10-11 1993. Elsevier Science B.V., Amsterdam, 1994.

[3] Herbert A. David. *Order Statistics*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, New York, 2nd edition, 1981.

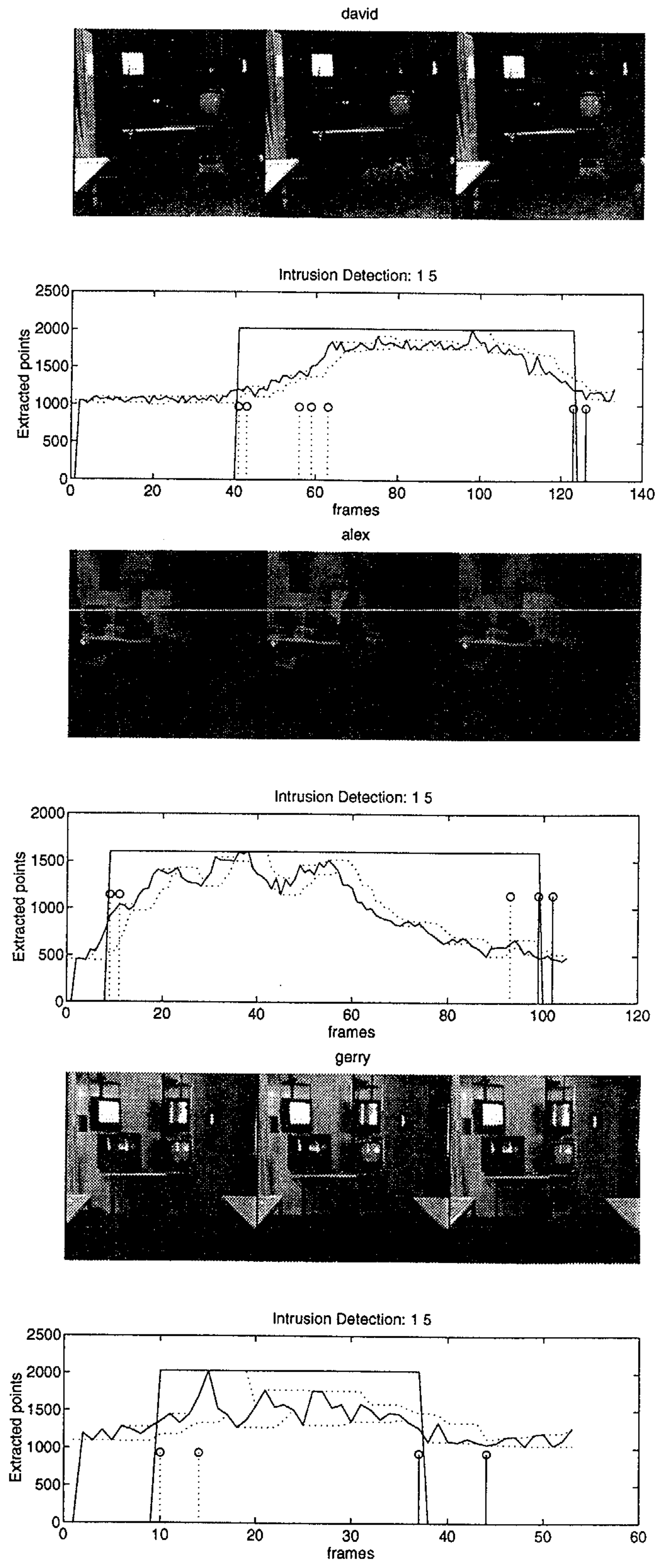

Figure 3: *The JNLOS detection applied to three different sequences, each of which contains an intrusion. The leftmost and the rightmost snaps correspond to the first and the last frames for which the intrusion is detected. On the diagrams, the horizontal axis represents the time (frames), while the corresponding values of the intrusion signal are given on the vertical axis. The dotted curves represent the envelope of the signal. The circles denote the samples for which the change of envelope is detected. The rectangles cover the portions of the sequences for which the intrusion is detected.*