Detection of Unusual Objects, Actions and Events in Streaming Video Surveillance Data

Bonny Banerjee (and students)
Institute for Intelligent Systems and Dept. of Electrical & Computer Engineering
The University of Memphis

Abstract

We present an online sparse coding based approach for abnormal event detection in videos. The approach is based on the intuition that abnormal events occur rarely. It extracts local spatiotemporal volumes and uses an online sparse dictionary learning algorithm to learn a set of atoms. While learning, the rarity of atoms is approximated online using Incremental Coding Length to measure the entropy gain of each atom. To speed up the process, a batch version of the well-known Orthogonal Matching Pursuit algorithm is used. The proposed approach operates in an unsupervised and online manner, hence applicable to real world streaming video. Experiments on three benchmark datasets (UCSD, UMN and Subway) and evaluations in comparison with a number of mainstream algorithms show that the proposed approach is comparable to the state-of-the-art.

Proposed Framework

Three steps:

1. Video Representation: Spatiotemporal interest point detector (Dollar et al. 2005)

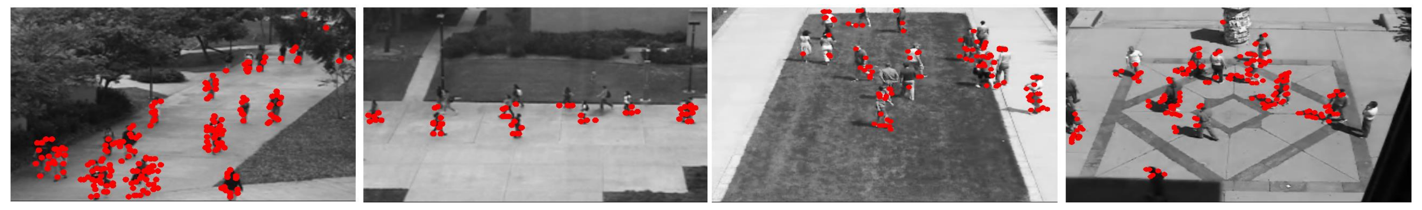

Figure 1: Examples of detected spatiotemporal interest points (best viewed in color) using the method in (Dollar et al. 2005).

2. Online Sparse dictionary learning:

$$\min_{\Gamma,D} \frac{1}{2}\sum_{i=1}^{N}\|x_i - D\gamma_i\|_2^2 \quad \text{subject to} \quad \forall i \ \|\gamma_i\|_0 \le \kappa$$

- Sparse coding: Orthogonal Matching Pursuit (OMP) (Rubinstein, Zibulevsky, and Elad 2008)

$$\min_{\gamma} \frac{1}{2}\|x - D\gamma\|_2^2 \quad \text{subject to} \quad \|\gamma\|_0 \le \kappa$$

- Dictionary update: Online dictionary update algorithm (Mairal et al. 2010)

$$\begin{aligned} D^t &= \underset{D}{\operatorname{argmin}} \frac{1}{t}\sum_{i=1}^{t}\left(\frac{1}{2}\|x_i - D\gamma_i\|_2^2\right) \\ &= \underset{D}{\operatorname{argmin}} \frac{1}{t}\left(\frac{1}{2}\mathrm{Tr}(D^T D A^t) - \mathrm{Tr}(D^T B^t)\right) \end{aligned}$$

3. Abnormal event detection: Incremental Coding Length (ICL) (Hou and Zhang 2008)

$$\mathrm{ICL}(p_j) = \frac{\partial H(\mathbf{p})}{\partial p_j} = -H(\mathbf{p}) - p_j - \log p_j - p_j \log p_j$$

where: $p_j = \frac{\sum_n |\Gamma_{j,n}|}{\sum_j \sum_n |\Gamma_{j,n}|}$

$$\theta_j = \begin{cases} \frac{\mathrm{ICL}(p_j)}{\sum_{i \in S} \mathrm{ICL}(p_i)}, & \text{if } j \in S \\ 0, & \text{otherwise} \end{cases}$$

$$\bar{\theta}^t = (1 - \alpha^t)\bar{\theta}^{t-1} + \alpha^t \theta^t$$

$$g(x|D) = \sum_{j=1}^{k} \bar{\theta}_j |\gamma_j|$$

Results

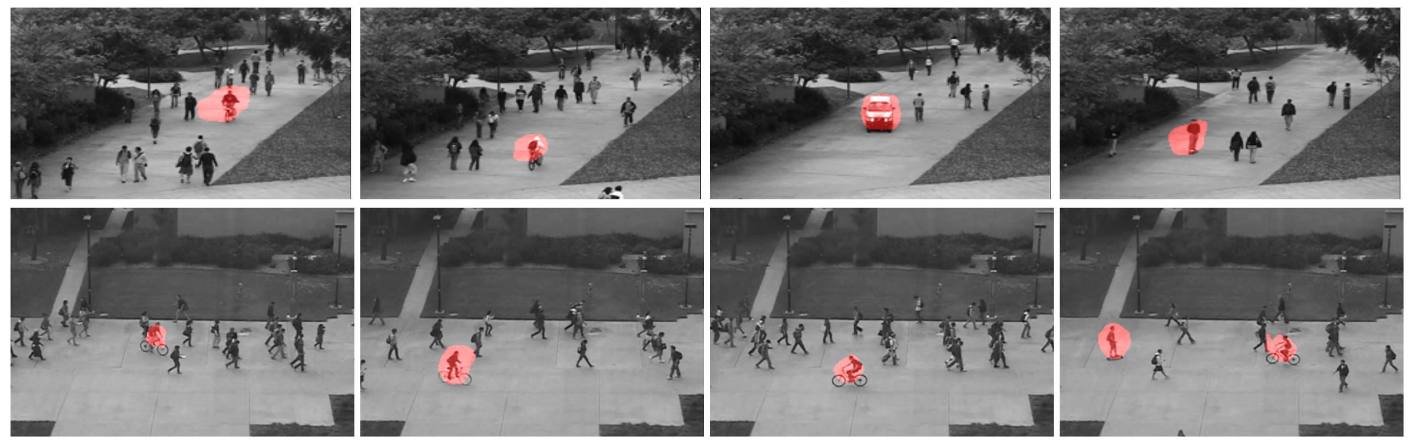

Figure 2: Abnormal frames and their detection result from UCSD Ped1 (top row) and UCSD Ped2 (bottom row) using our model. The bikers, skaters and cars were detected as anomalous patterns (highlighted in red, best viewed in color). The proposed method can detect multiple anomalous patterns within a single frame.

Method	EER (Ped1)	RD (Ped1)	EER (Ped2)	RD (Ped2)
Ours	19.8	69.5	22.3	67.5
H-MDT	17.8	75	18.5	70
Sparse	19	46	X	X
STC	15	73	13	74
MPPCA	35.6	23.2	35.8	22.4
Force flow	36.5	40.9	35	27.6
LMH	38.9	32.6	45.8	22.4

Table 1: Anomaly detection performance on USCD Ped1 and Ped2 datasets.

Figure 3: Abnormal frames from UMN dataset. Anomalous regions, as detected by our model, are highlighted in red (best viewed in color).

Method	AUC	EER
Ours	99.5	3.65
Chaotic invariant	99.4	5.3
Social force	94.9	12.6
Sparse	99.6	2.8
Local stat. aggr.	99.5	3.4
H-MDT	99.5	3.7

Table 2: Quantitative comparison between different methods on UMN dataset.

Figure 4: Anomaly detection in Subway dataset. Top row represents entrance gate and bottom row represents exit gate. Anomaly detection includes detection of wrong direction events and no payment events. Anomalies detected by our model are highlighted in red (best viewed in color).

Method	Dataset	Abnormal events	False alarm
Ours	Entrance Exit	60/66 19/19	5 2
STC	Entrance Exit	61/66 19/19	4 2
MPPCA	Entrance Exit	57/66 19/19	6 3
Dynamic SC	Entrance Exit	60/66 19/19	5 2
Sparse	Entrance Exit	27/31 9/9	4 0

Table 3: Performance of different methods on the Subway dataset.

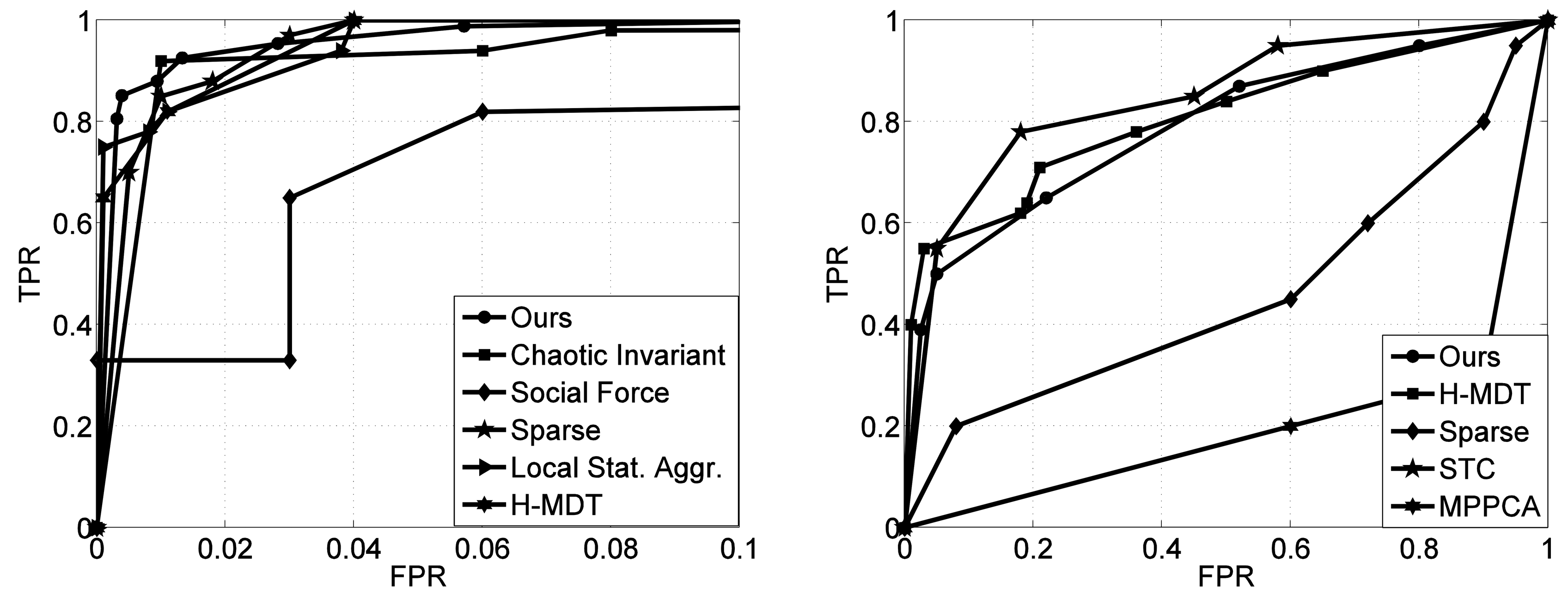

Figure 5: ROC curves for pixel-level criterion on UCSD Ped1 dataset (left) and ROC curves for frame-level criterion on UMN dataset (right).

Conclusion

- A rarity based approach for anomaly detection in streaming videos was proposed.
- A dictionary of atoms was learned from the data in an unsupervised manner using an online sparse coding framework.
- While learning, the rarity of atoms was approximated online using ICL and anomaly score for an input was computed as the sum, over all atoms, of the average energy multiplied by absolute coefficients.
- No prior assumption was made regarding the data or nature of anomaly and the online operation of the proposed method allows it to deal with varying data distribution and is useful to real-time applications.
- The proposed approach was extensively experimented with a number of benchmark datasets and the results are comparable to the state-of-the art.

Acknowledgement

We gratefully acknowledge support from the City of Memphis and Fedex Institute of Technology.

For enquiries, contact: Bonny Banerjee, bbnerjee@memphis.edu

References

- Dollar, P.; Rabaud, V.; Cottrell, G.; and Belongie, S. 2005. Behavior recognition via sparse spatio-temporal features. In 2nd Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance, 2005, 65–72. IEEE.
- Hou, X., and Zhang, L. 2008. Dynamic visual attention: Searching for coding length increments. In NIPS, 681–688.
- Mairal, J.; Bach, F.; Ponce, J.; and Sapiro, G. 2010. Online learning for matrix factorization and sparse coding. Journal of Machine Learning Research 11:19–60.
- Rubinstein, R.; Zibulevsky, M.; and Elad, M. 2008. Efficient implementation of the k-svd algorithm using batch orthogonal matching pursuit. CS Technion.