

Unsupervised Behavior-Specific Dictionary Learning for Abnormal Event Detection

Huamin Ren¹

hr@create.aau.dk

Weifeng Liu²

weifeng.liu@nbi.ku.dk

Søren Ingvor Olsen³

ingvor@di.ku.dk

Sergio Escalera⁴

sergio@maia.ub.es

Thomas B. Moeslund¹

tbm@create.aau.dk

¹ Department of Architecture, Design and Media Technology
Aalborg University
Aalborg, Denmark

² Niels Bohr Institute
University of Copenhagen
Copenhagen, Denmark

³ Department of Computer Science
University of Copenhagen
Copenhagen, Denmark

⁴ Dept. Applied Mathematics, University of Barcelona, Computer
Vision Center
Barcelona, Spain

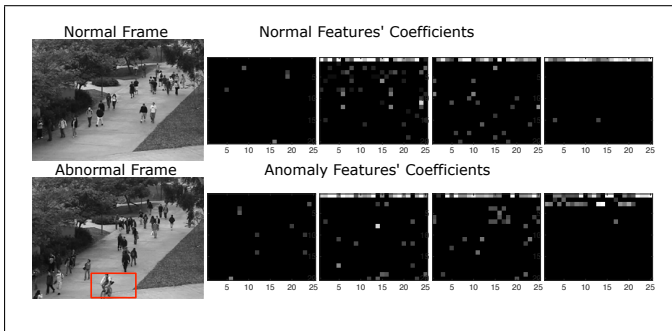


Figure 1: Non zero distribution in coefficients of features from normal and abnormal frames. From left to right: Lu13 [3] and our three frames: D_{basis} , $D_{feature}$ and $D_{missedAtoms}$. Accumulated non zeros in each dictionary are normalized in $[0,1]$; therefore, the whiter, the more non zeros.

Abnormal event detection is an important issue in video surveillance applications. The goal is to detect abnormal or suspicious behaviors while given training samples that contain only normal behaviors. Sparse representation has showed its effectiveness in abnormal event detection [2, 3, 4, 5], where a dictionary is commonly learned during training and anomalies are detected based on reconstruction error from the learned dictionary. Note that only a small proportion of the data is used for training - relatively to the huge amount of surveillance data, there is a high risk of incomplete normal patterns in the training data. Consequently, dictionary learning is crucial to the overall abnormality detection performance.

We propose a Behavior-Specific Dictionary (BSD) algorithm, which takes into consideration the relation of atoms in one dictionary without

prior knowledge. In our algorithm, atoms in the same dictionary are related to one type of behavior. Normal features could be sparsely represented by one dictionary with a similar pattern, while anomalies could not. To further reduce false alarms, features with compact non zero coefficients in each dictionary (called missed atoms) are added as new atoms in order to better distinguish between anomalies and infrequent behaviors. The non zero distribution of coefficients for normal and abnormal frames are demonstrated in Fig. 1, compared with Lu13 [3]. It can be seen that our frames could have a more concentrated non zero distribution for normal features, meanwhile obtaining spread out non zeros for abnormal features. Our proposed algorithm is shown in Fig.2.

Experiments are carried out on two datasets: UCSD Anomaly dataset and Anomaly Stairs dataset. The former dataset is used to get a comparable result with other state-of-art methods; while the latter one is used to test the effectiveness of our method when the training data is incomplete. On UCSD Anomaly dataset, our BSD obtains a AUC of 56.17% on pixel-level evaluation, which achieves the best result (promoting AUC by 10% compared to the best score (46.1%) published so far [2]), yet has a satisfactory result on frame-level evaluation (AUC is 70.69%). On Anomaly Stairs dataset, seen from Tab. 1, our BSD achieves the best detection with the least false alarms.

Pattern and training size	# Correctly detected abnormal videos (TP)			# False alarms (FA)		
	Lu13 [3]	Bao14 [1]	BSD (ours)	Lu13 [3]	Bao14 [1]	BSD (ours)
One group (1%)	11	12	15	13	16	14
One group (10%)	17	12	17	21	16	15
Two groups (1%)	14	13	17	15	19	17
Two groups (10%)	14	19	15	17	18	12
Three groups (1%)	14	18	18	15	19	16
Three groups (10%)	14	19	19	17	18	17
Average TP / FA	14.0	15.5	16.8	16.3	17.7	15.2
Average TPR / FPR	70.0%	77.5%	84.2%	77.8%	84.1%	72.2%

Table 1: Results on Anomaly Stairs dataset by varying the size of the training data.

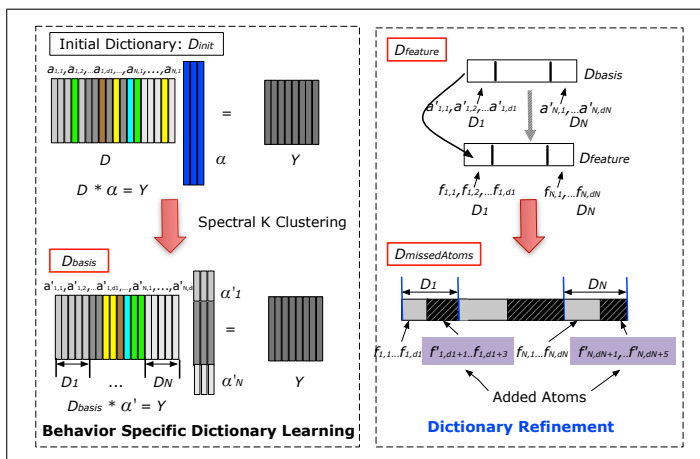


Figure 2: Behavior-specific dictionary learning (left) and refinement procedure (right). During BSD learning, multiple dictionaries are obtained, each of which corresponds to one particular normal behavior. Refinement is to find missed atoms to improve current dictionaries.

Our proposed algorithm learns behavior-specific dictionaries through unsupervised learning. The experiments proved the effectiveness of the proposed method, especially when the data available during the training is limited.

- [1] Chenglong Bao, Hui Ji, Yuhui Quan, and Zuowei Shen. L0 norm based dictionary learning by proximal methods with global convergence. In *CVPR*, pages 3858–3865, June 2014.
- [2] Y. Cong, J. Yuan, and J. Liu. Sparse reconstruction cost for abnormal event detection. In *CVPR*, pages 3449–3456, 2011.
- [3] C. Lu, J. Shi, and J. Jia. Abnormal event detection at 150 fps in matlab. In *ICCV*, pages 2720–2727, 2013.
- [4] H. Ren and T. B. Moeslund. Abnormal event detection using local sparse representation. In *AVSS*, pages 125–130, 2014.
- [5] B. Zhao, F. Li, and E. P. Xing. Online detection of unusual events in videos via dynamic sparse coding. In *CVPR*, pages 3313–3320, 2011.