How does structured sparsity work in abnormal event detection?
Ren, Huamin; Pan, Hong; Olsen, Søren Ingvor; Moeslund, Thomas B.

Publication date: 2015

Document Version
Early version, also known as pre-print

Citation for published version (APA):
How Does Structured Sparsity Work in Abnormal Event Detection?

Huamin Ren*, Hong Pan†, Søren Ingvor Olsen† and Thomas B. Moeslund*,
*Department of Architecture, Design and Media Technology
Aalborg University, Aalborg, Denmark
Email: hr@create.aau.dk,.tbm@create.aau.dk
†Department of Computer Science
University of Copenhagen, Copenhagen, Denmark
Email: hong.pan@di.ku.dk,ingvor@di.ku.dk

Abstract—In traditional sparse modeling, it is assumed that
a signal/feature/image can be either accurately or approxi-
mately represented by a sparse linear combination of atoms
from a learned dictionary. Structured sparsity, which is beyond
traditional sparse modeling, addresses collaborative structured
sparsity to add stability and prior information to the represen-
tation. Specifically, in structured sparse modeling, the atoms are
partitioned in groups, and a few groups are selected at a time
for the sparse coding. Supposing there are $n$ classes and $m_i$
training data for each class $i$, $B[i] = [b_{i1}\ldots b_{im_i}], i = 1...n$ and
each $b_{ij} \in R^D$, the dictionary of the training data has a block
structure where a few blocks of the dictionary correspond to
the training data in each class. Thus, a test example can be
represented as a linear combination of training data from a few
blocks of the dictionary corresponding to its class.

Structured sparsity has been found important in computer
vision such as face recognition, motion segmentation, and activity
recognition, since the data lie in multiple low-dimensional sub-
spaces of a high dimensional ambient space in these applications.
In fact, abnormal event detection can be another beneficiary -
given a testing frame, it should be identified as a normal frame if
all the features within the frame preserves a structured sparsity:
all features could be linearly represented by only a few atoms,
more importantly, these a few atoms come from the same or
similar behavior. Otherwise, it should be detected as an abnormal
frame. This structured sparsity is illustrated in Fig. 1.

It is straightforward to consider structured sparsity algo-
rithms to achieve the structured sparsity shown in Fig. 1.
However, it is infeasible to apply structured sparsity algorithms
directly in abnormal event detection, which are mainly due to two
reasons: 1) abnormal event detection has a highly biased training
data - only normal videos are used during the training, which
is the due to the fact that abnormal videos are limited or even
unavailable in advance in most video surveillance applications.
As a result, there could be only one label in the training data
which hampers supervised learning; 2) Even though there are
multiple types of normal behaviors, how many normal patterns
lie in the whole surveillance data is still unknown. This is because
there is huge amount of video surveillance data and only a
small proportion is used in algorithm learning, consequently,
the normal patterns in the training data could be incomplete. As
a result, any sparse structure learned from the training data
could have a high bias and ruin the precision of abnormal event
detection. Therefore, we in the paper propose an algorithm to
solve the abnormality detection problem by sparse representation,
in which local structured sparsity is preserved in coefficients. To
better meet the needs of practical video surveillance applications,
our method aims at dictionary learning preserving the structured
sparsity through a relatively small training data.

Our method contains three steps. Step 1: Initial dictionary
construction, which selects initial atoms to form multiple dic-
tionaries. These atoms are learned based on visual features;
Step 2: Transferring atoms in Step 1 into feature space, and
replace atoms in the initial dictionary with new feature atoms;
Step 3: Dictionary refinement, which preserves local structured
sparsity. We compare our method with state-of-the-art on the
public dataset: UCSD anomaly dataset; moreover, we carry our
experiments on the Anomaly Stairs dataset with a challenging
setting: an incomplete normal patterns or a small set is used for
training the model. Experimental results show the effectiveness
of our method.