ISSN: 2320 - 8791
www.ijreat.org

Sparse Coding Algorithm for Anomaly Detection

R.Kanimozhi[1], M.Jeeva[2]andT.karthick[3]

[1,2,3]Department of Computer Science and Engg., Anand Institute of Higher Technology, Chennai, Tamilnadu, India

Abstract

Surveillance is monitoring the behavior, activities, or other changing information, usually of people for the purpose of influencing, managing, directing, or protecting.The camera based surveillance system is used to detect the abnormal activities or unwanted activities. Such abnormal activities are infrequent when compared to regular activities. Now the present system of surveillance is done manually. Employing human operators in video monitoring is known to be ineffective, error prone, and expensive. Hence we propose a computer vision aided anomaly detection system which will enable the selection of video frames that contain an anomaly and those selected frames will be used for manual verification. Sparse coding algorithm has been developed to detect the panic in human crowds and alarm where there is a significant change in the behavior. Most of the existing algorithms are computationally expensive and too slow to be real time. Hence we propose Orthogonal Matching Pursuit (OMP) an inexpensive way to model a behavior using dictionary elements and OMP is used to design a panic detection algorithm.

Keyword:*abnormal behavior, orthogonal matching pursuit algorithm, spatial and temporal localization.*

1. Introduction

Recently, the surveillance of human activities has drawn a lot of attention from the research community and the camera based surveillance is being tried with the aid of computers. Surveillance is required to detect abnormal or unwanted activities. Such abnormal activities are very infrequent as compared to regular activities. At present, surveillance is done manually, where the job of operators is to watch a set of surveillance video screens to discover an abnormal event. Surveillance has become easier with the development of inexpensive digital cameras and fast networks. Many cameras have been installed and now almost every important place is under surveillance; however, manual surveillance does not seem to be an efficient solution. Humans are good at detecting anomalies; however, they are very poor at maintaining attention during mundane tasks. This is expensive and prone to error. There are a few problems of manual surveillance:

- Manual surveillance implies that a human is watching other human activities and this may be considered as a privacy breach. Such a surveillance system typically has to pass many rules and regulations checks.
- Humans are good detectors of an anomaly because of their strong analytical capability and great sense of context; however, their effectiveness drops sharply in un engaging jobs. Anomalies are extremely rare, so operators have to watch normal activities most of the time. It tends to make them bored and inattentive and thus the chances are high of missing anomalies.
- Humans are expensive, they need lots of space and resources; it makes manual surveillance unaffordable for many people in need.

Fig 1:Normalbehavior

Fig2: Abnormal behavior

A variety of methods are being invented to automatically detect anomalies in human crowds. With the use of such systems, only those frames will be displayed for manual verifications which are suspected by the anomaly detection algorithm; in the future, human verification might not be needed at all. A simple computer vision technique, which detects motion in a region of interest, can filter out many frames for anomaly detection in an empty place such as restricted areas and vacant properties. It shows the capacity to which computer vision can help in various surveillance operations.

Fig 3: A typical surveillance control room, where an operator is watching at a set of screens for surveillance purposes.

2. Related Work

A recently proposed technique, dictionary learning is gaining interest for panic detection. Using a set of training feature vectors, dictionary learning finds a few representative dictionary elements using an optimization process such that any training data can be reconstructed using those dictionary elements. As training samples correspond to normal behaviors, a panic is signaled if dictionary elements fail to reconstruct a feature vector within an acceptable error limit. Orthogonal matching pursuit selects a set of optimal dictionary elements and dictionary learning uses OMP or similar approaches in each iteration of optimization. So OMP is relatively inexpensive than dictionary learning.We propose a novel algorithm to model normal crowd behaviors using wavelet based orthogonal matching pursuit (OMP). Like dictionary learning, OMP based panic detection also uses dictionary elements to reconstruct test feature vectors, and a panic is alarmed if the reconstruction error goes beyond an acceptable limit. Each type of anomaly poses a different set of requirements for an anomaly detection system; hence, it is challenging to develop an anomaly detection algorithm to detect all types of anomalies. To simplify this difficult problem, algorithms are typically developed to address a particular subset of anomalies.

3. Panic Detection System

A human crowd panic detection system is composed of three main segments:

3.1 Crowd Behavior Representation

Any analysis of a crowd's behaviors requires expressing a crowd's characteristics in a tangible form. A set of features is required that show a significant change in the presence of irregularities. The type of features largely depends on the type of panic to be detected. A few of the features used in detecting a crowd panic are motion information, headcount, texture of the image, and the group-size

3.2 Crowd Modeling

Once features are decided to represent crowd behaviors, a model is needed to learn them. Panics are difficult to model because of high variations and rare occurrences; so discriminative classifiers such as support vector machine (SVM) are not effective to classify panic and normal behaviors. Panic detection is better suited to hypothesis testing.A hypothesis is an assumption about a given instance and hypothesis testing uses two types of hypothesis:

3.2.1 Null hypothesis

A null hypothesis represents a “no change" situationwith respect to the normal behavior.

3.2.2 Alternate hypothesis:

An alternate hypothesis appears when the null hypothesis is disproved. In a panic detection system, as the panic is defined as a deviation with respect to a normal behavior, a null hypothesis, it acts as an alternate hypothesis.

The main challenge of developing a panic detection algorithm is to allow natural variations of a normal behavior and also to detect anomalies with high accuracy, there are three types of modeling methods.

1. Parametric models:

In a parametric model, finding a perfect parametric model for a data set is difficult and an approximate parametric model is selected to represent them. These approximations sometimes cause a high inaccuracy.

2. Non-parametric models:

Unlike a parametric model, a non-parametric model does not assume a particular distribution about samples. It extracts distribution structure from the sample data and uses it to estimate the association of a given sample with the training class. As this approach does not make any assumption about the sample, it is more generic than the parametric model. The algorithm needs to calculate distance with each training sample, the space requirement of this method may be larger than a parametric model.

4. *OMP and dictionary learning*

4.1 Dictionary Learning

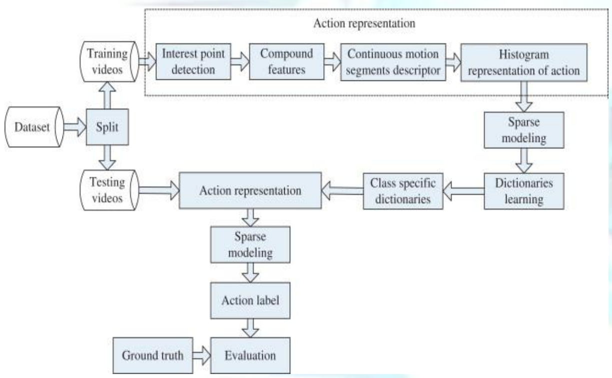

Fig 4: Dictionary Learning.

Dictionary learning is the task of learning or training a dictionary such that it is well adapted to its training data. Usually the objective is to give sparse representation of the training set, making the total error as small as possible, i.e. minimizing the sum of squared errors. Let the training data constitute the columns in a matrix Y and the sparse coefficients vectors are the columns in matrix X. The objective function of the dictionary learning can be stated formally as a minimization problem

$$\min_{\mathbf{D},\mathbf{X}} \|\mathbf{Y}-\mathbf{D}\mathbf{X}\|_F^2 \text{ subject to } \|\mathbf{X}\|_p \leq T,$$

4.2 Orthogonal Matching Pursuit Algorithm.

OMP is an extension of the matching pursuit (MP) algorithm.A sufficiently large dictionary is initialized D = [d1,…,d_{NM}]. MP provides a greedy way of finding an optimal set of dictionary elements $\hat{D} \subseteq D$ from D. The algorithm runs using multiple iterations and after each iteration, one dictionary element from D is determined. The $\hat{\mathbf{f}}_0 = \mathbf{f}$ vector $\hat{\mathbf{f}}_j$ represents the residue of **f** after the j^{th} iteration and
A DE is selected after the j^{th} iteration as

$$\hat{i} = \operatorname*{argmax}_{i \in [1, \ldots, N_M]} |\hat{\mathbf{f}}_{\mathbf{j-1}} \cdot \check{\mathbf{d}}_{\mathbf{i}}|$$

Where $\check{\mathbf{d}}_{\mathbf{i}}$ is a unit vector of the i^{th}DE.

5. Motion Estimation Methods

Motion estimation methods have certain properties of a moving point remains unchanged during the motion (constantness property) and motion vectors change smoothly across the neighboring pixels (smoothness constraint). Based on these assumptions, a cost function is formed and the motion is estimated by minimizing the cost function. Motion estimation methods differ in terms of the constantness assumptions and their formulations in the cost function. Broadly, there are two types of motion estimation methods, namely optical flow and SIFT flow.

5.1 Optical flow or optic flow

Is the pattern of apparent motionof objects, surfaces, and edges in a visual scene caused by the relative motion between an observer (an eye or a camera) and the scene. The optical flow methods try to calculate the motion between two image frames.Optical flow gives motion estimation for all pixels in a given image; however, it does not work well for large displacements.To solve this problem, proposed a warping based optical flow estimation.The main problem with the warping based approach is that as we warp an image to a smaller scale, some information is lost. For a small object, the loss can be so significant that it becomes indistinguishable. This leads to a high error in the initial estimate and the error continues until the finest scale.

5.2 SIFT flow

Optical flow gives poor results in the presence oflarge illumination changes. Also, it gets severely affected by large displacements. SIFT flow has been developed to overcome these limitations of the optical flow. There is no significant difference in the formulations of optical flow and SIFT flow, since both make constantness and smoothness assumptions. However SIFT flow uses scale invariant feature transform (SIFT) feature vectors to represent a pixel's property. SIFT uses information from neighboring pixels to provide robustness against illumination variations, scaling and rotations.

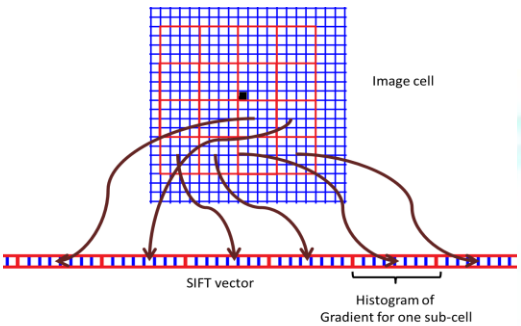

Fig5: SIFT feature formation

One big difference between optical flow and SIFT flow is that the search window size for SIFT flow is much larger since an object can move drastically from one image to another in scene alignment. Therefore, we need to design efficient algorithm to cope with the complexity.

6. Motion Representation

Expressing an optical flow in an image is a challenging task. A new technique of optical flow representation, called triple color flow presentation (TCFP), has been proposed to remove these limitations. In TCFP, three color channels (red, green, and blue) are used for three different purposes.

6.1 Triple Color Flow Presentation:

Red***:*** A motion estimate b = (u v) is unquantized and provides a vector from a pixel in the image I(t) to a point in the image I(t +T). This vector can be used to form the second image I’ as I’((x) + b; t + T) = I(x, t) where motion estimates b are quantized in b. For an ideal flow estimate, I0 must be same as I(t + T). However, a few pixels I in I’ do not correspond to a vector from any pixel in I(t). This can happen for the following reasons,

1. Regions corresponding to I were occluded or not available in the previous image I(t), so no flow was possible.
2. An error in the motion estimation.
3. A few unmapped pixels may appear because of the conversion of an unquantized motion estimates b to aquantized vector b. Pixels which are unmapped at frame t+T will have red channel values set to 1, where 1 is the largest value of the color dynamic range. A few red pixels are expected because of aforementioned occlusions and quantization; however a large number of red pixels gives valuable information about the quality of optical flow.

Green***:*** The estimated image I is converted to a gray scale image and shown using the green channel.

Blue: The original image I (t+T) is converted to a gray scale image and shown using the blue channel. Any existing error in motion estimation can be observed clearly using TCFP. If there are many unmapped pixels then there will be a large number of red pixels. For an idea optical flow the green and blue images should exactly match with each other. In the case of an inaccurate motion estimate, the image in the green channel shall not match with the image in the blue channel, which gives a distinct region in abnormally high blue or green intensity. Thus TCFP provides a framework to analyze and compare the accuracy of a motion estimation algorithm in the absence of the ground truth.

6.2Panic localization**:**

After modeling a normal behavior, a cost function is formed to test a sample against the model. For a given sample, if the cost function gives higher cost as compared to a threshold, the sample is classified as a panic. There are two methods for localization of errors in a coded video: spatial localization and temporal localization;

6.2.1 Temporal localization:

The main objective of a panic detection system is to detect frames which contain panic. With the aid of an accurate temporal localization, only suspected abnormal frames will be shown to operators. It can significantly increase the efficiency and accuracy of a surveillance system.

6.2.2 Spatial localization:

Once a panic is temporally localized, the spatial localization tells the exact location where the system has suspected an abnormal behavior. Though it is not the primary objective of a panic detection system, it can greatly improve a surveillance system. It can also help in the quick detection of false positive cases.

Fig 6: Original image

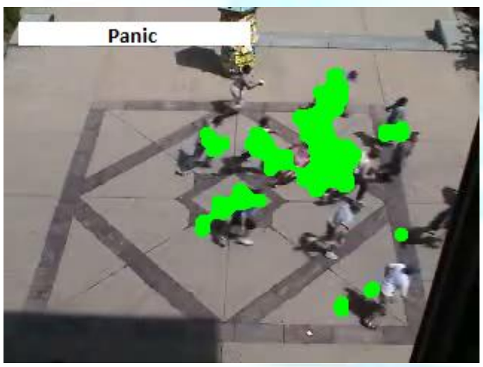

Fig 7: panic localized

7. Results

The algorithm will be tested with two types of flow, optical flow and SIFT flow. Optical flow will be computed using Brox et al.'s mex implementation and SIFT flow will be computed with Labview simulation software.
Two accuracy measurement schemes, namely AROC and F1-measure, are used to compare the accuracy. Most of the existing algorithms on the panic detection have reported and compared results in terms of AROC. So, the accuracy of the proposed algorithm will be compared with others in terms of AROC. AROC gives a good idea about the stability of a system; however, it does not give much information about the accuracy of a system. The F1-measure measures the accuracy of a classifier and hence it has also been reported for completeness of our analysis.

Accuracy measurement*:*

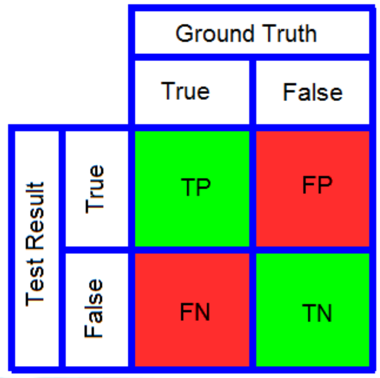

Fig 8: There can be two types of outcomes of a classification, namely correct classifications and incorrect classifications. True positive (TP) and true negative (TN) are correct classifications. False negative (FN) and false positive (FP) are wrong classifications. In the case of a panic detection algorithm, panic is considered as positive.

The accuracy of the proposed method will be compared with other existing methods and the proposed approach will give competitive accuracy with respect to the most accurate approaches. A significant accuracy will be obtained with respect to the dictionary learning based approach. As there is no standard method to count the number of positive and negative cases; it is difficult to make a direct inference about which one is the best. However, based on the competitive accuracy and the real-time computational complexity the proposed method will be very effective in real-time panic detection. The abnormal behavior of people detected in surveillance and intimated to the operator.

Fig 9: Motion along the camera axis

Fig 10: A running person.

8. Conclusion

The proposed OMP based panic detection method produces better accuracy than a state-of-the-art dictionary learning based method. The optical flow estimation is the major contributor to the time complexity of panic detection algorithms; it makes a panic detection algorithm non-real time. Optical flow estimates both the motion direction and the magnitude. The proposed approach uses only the magnitude of motion, we think that inexpensive motion estimation algorithms can be developed which provide only motion magnitude.

9. Future Work

One of the simplifications achieved in the proposed work is the use of only motion magnitude. The proposed algorithm uses optical flow to get a motion estimate. Optical flow estimation is computationally expensive and it gives both the magnitude and direction of the motion. A relatively inexpensive method such as can be developed which produces only motion magnitude and tested with the proposed algorithm. Without adaptation the proposed algorithm will detect even a smooth change as a panic; with adapted dictionary elements and coefficients, the system will take care of smooth transition and it will trigger an alarm only when there is a sudden change.

References

[1] A. Adam, E. Rivlin, I. Shimshoni, and D. Reinitz.Robust real-time unusual event detection using multiple fixed-location monitors. IEEE Trans. on Pattern Analysis and Machine Intelligence, March 2008.

[2] M. Aharon, M. Elad, and A. Bruckstein. K-SVD: An algorithm for designing over completedictionaries for sparse representation. IEEE Trans. on Signal Processing, Nov. 2006.

[3] S. Ali and M. Shah. A lagrangian particle dynamics approach for crowd flow segmentation and stability analysis.In IEEE Conf. on Computer Vision and Pattern Recognition, June 2007.

[4] E. L. Andrade, S. Blunsden, and R. B. Fisher.Modelling crowd scenes for event detection.In 18th international conference on Pattern Recognition, IEEE 2006.

[5] A. Basharat, A. Gritai, and M. Shah. Learning object motion patterns for anomaly detection and improved object detection. In IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2008.