BAYESIAN GENERATIVE MODEL BASED ON COLOR HISTOGRAM OF ORIENTED PHASE AND HISTOGRAM OF ORIENTED OPTICAL FLOW FOR RARE EVENT DETECTION IN CROWDED SCENES

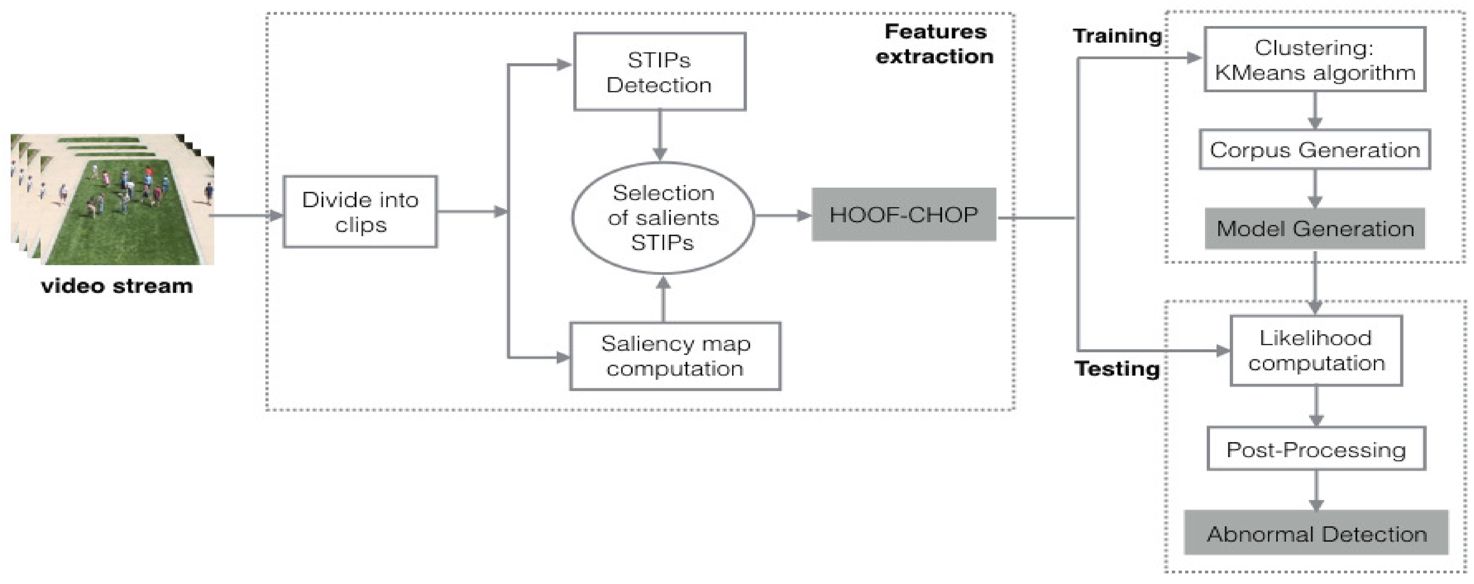
Dieudonné Fabrice ATREVI[1], Damien VIVET[2], Bruno EMILE[1]

[1] University Of Orleans, INSA Centre Val de Loire, PRISME EA 4229, F45072, Orleans, France
[2] University of Toulouse, DEOS/SCAN ISAE-Supaero, Toulouse, France

ABSTRACT

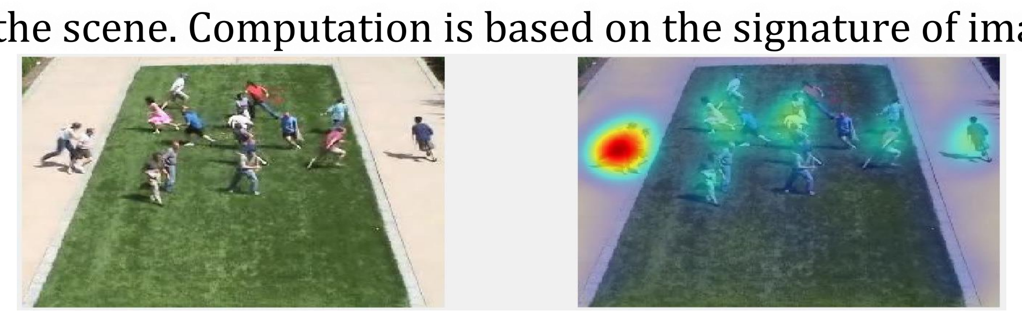
In this paper, we propose a new method for rares events detection in crowded scenes using a combination of Color Histogram of Oriented Phases (CHOP) and Histogram of Oriented Optical Flow (HOOF). We propose to detect and filter spatio-temporal interest points (STIP) based on the visual saliency information of the scene. Once salient STIPs are detected, the motion and appearance information of the surrounding scene are extracted. Finally, the extracted information from normal scenes are modelled by using a Bayesian generative model (Latent Dirichlet Allocation). The rare events are detected by processing the likelihood of the current scene in regard to the obtained model. The proposed method has been tested on the publicly available UMN dataset and compared with different state-of-the-art algorithms. We have shown that our method is very competitive and provides promising results.

PROPOSED APPROACH



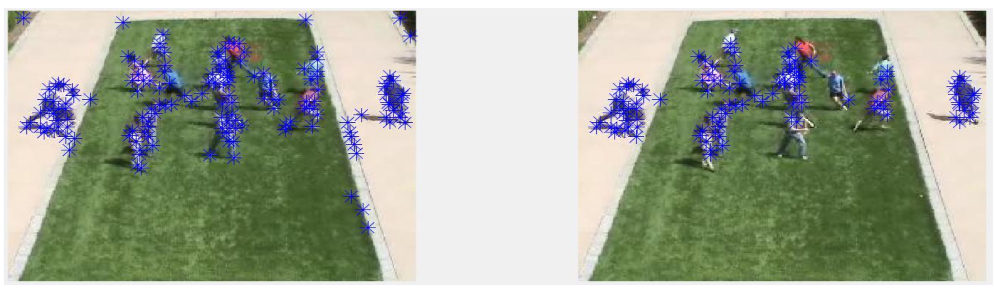
DETAILS

- Visual saliency map for the volume of image are used to highlight interesting region (foreground) in the scene. Computation is based on the signature of image [1];



- Detect and filter spatio-temporal interest point based on saliency score and optical flow magnitude in image volume:

SelectedPoint = DetectedPoint(mean(SaliencyScore > threshold) and mean(Optical Flow) and variance(Optical flow))

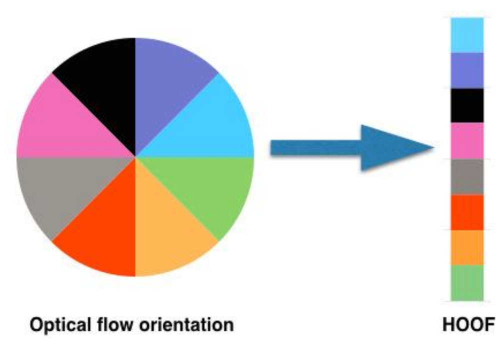


- Describe each STIP with CHOP (Color Histogram of Oriented Phase) for appearance information and HOOF (Histogram of Oriented Optical flow) for motion;

- Model frequents events with the «Latent Dirichlet Allocation» [2] algorithm using "Bag of Word" approach;

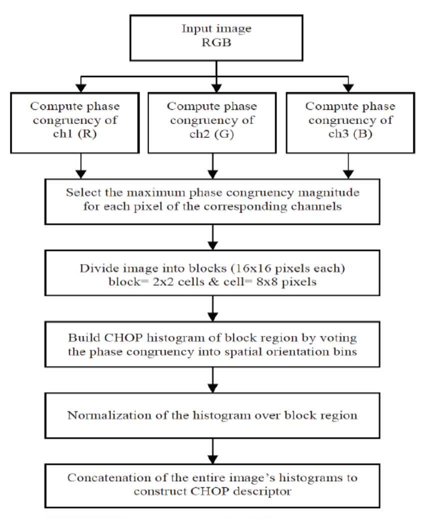
FEATURES

Histogram of Oriented Optical Flow[3]



- Horn-Shrunk algorithm for OF computation
- Bin vote are weighted with the magnitude

Color Histogram of Oriented Phase[4]



CHOP computation process

$$E(x) = \sqrt{F(x)^2 + F_H(x)^2}$$

Local Energy formulation

- F(x): Filtered signal
- $F_H(x)$: Hilbert Transform of F(x)

$$PC(x) = \frac{E(x)}{\epsilon + \sum_n A_n}$$

Phase congruency formulation

PARAMETERS

- Clip size: 15 frames
- Feature size/clip: 960-D;
- Number of topic: 40;
- Number of cluster: 40 visual words;
- Patches size: 16 x 16;

DATASET

- UMN dataset: 3 different scenes: lawn scene, indoor scene and plaza scene;
- The dataset presents escape movement as abnormality;
- Training set: Normal sequence of 5 first videos of indoor scene;
- Test set: Both normal and abnormal sequence of the remain video

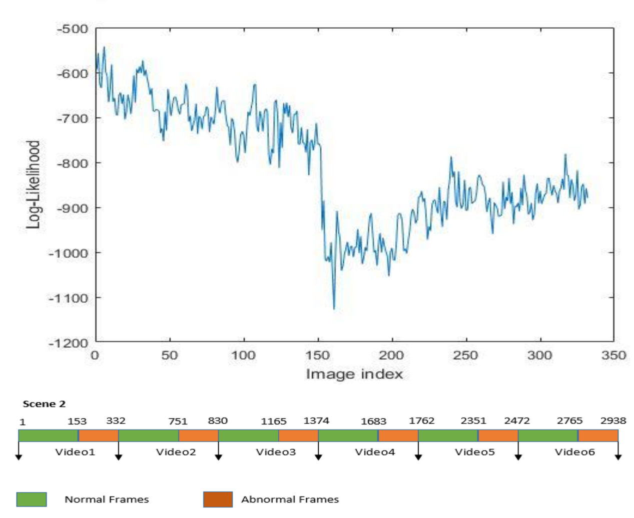
Scene 1: Lawn scene

Scene 2: Indoor scene

Scene 3: Plaza scene

RESULTS

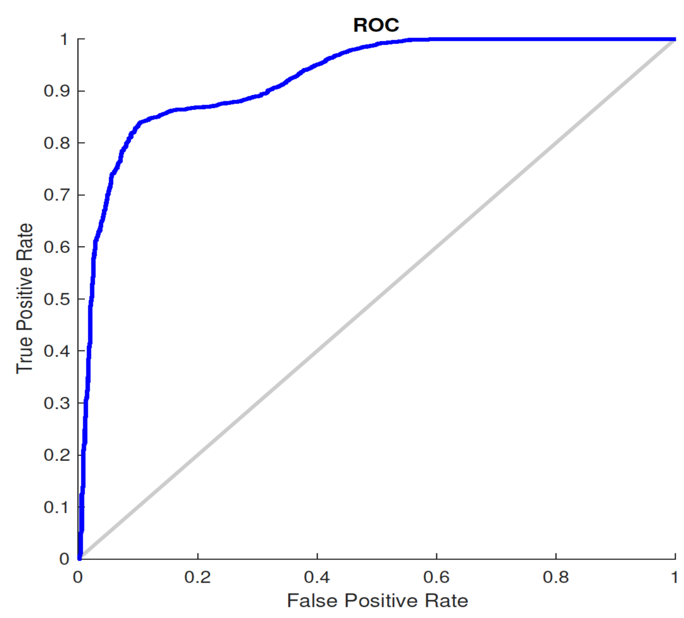
Log-Likelihood on video 1 of Indoor scene



- Abnormal detection based on the Likelihood of the clip;
- Find best parameters alpha and beta using the EM method over normal scene;
- Performance analysis based on the ROC curve and AUC

$$l(\alpha, \beta) = \log p(D|\alpha, \beta)$$

- D is the train Corpus
- Learned parameters



ROC curve of the proposed method

Method	SIFT	STIP	DT	Bag of Graph	**Ours**
AUC	0.85	0.85	0.81	0.95	**0.93**

Comparison with the bag-of-words existing methods

Method \ AUC	Lawn	Indoor	Plaza
Pure optical flow	0.84		
Social Force	0.96		
NN	0.93		
STCOG	0.9362	0.7759	0.9661
HMOFP	0.9976	0.9570	0.9869
Ours	**0.93**		

Comparison with others approach

CONCLUSION

In this paper, we proposed a new method that integrated the visual saliency for spatio-temporal interest point selection, a new feature descriptor based on the histogram of oriented optical flow and the color histogram of oriented phase for rare event detection in a controlled environment. The proposed approach successfully models frequent event in video using the Bayesian generative model LDA. We reached a competitive result with an accuracy of 93 % compared to the prior works on the abnormal UMN dataset.

ACKNOWLEDGE

This work is part of LUMINEUX project, supported by the Regional Centre-Val de Loire (France). The authors would like to acknowledge the Conseil Regional of Centre-Val de Loire for its support.

REFERENCES

[1] Xiaodi Hou, Jonathan Harel, and Christof Koch, "Image signature : Highlighting sparse salient regions", IEEE transactions on pattern analysis and machine intelligence, 34 (Jan) :194–201, 2012.

[2] David M Blei, Andrew Y Ng, and Michael I Jordan, "Latent Dirichlet Allocation", Journal of machine Learning research, 3 (Jan) : 993–1022, 2003.

[3] Rizwan Chaudhry, Avinash Ravichandran, Gregory Hager, and Rene Vidal, "Histograms of oriented optical flow and binet-cauchy kernels on nonlinear dynamical systems for the recognition of human actions", In computer vision and pattern recognition, CVPR 2009, pages 1932–1939, 2009.

[4] Hussin K Ragb and Vijayan K Asari, "Color and local phase based descriptor for human detection", In Aerospace and Electronics Conference (NAECON) and Ohio Innovation Summit (OIS), 2016 IEEE National, pages 68–73. IEEE, 2016.