Abnormal Behavior Detection in Arial Video Surveillance

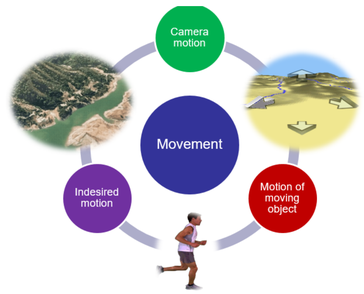
A. Walha, A. WALI, A. Alimi, Research Groups in Intelligent Machines Laboratory at Sfax University Tunisia

جامعة الملك عبدالله للعلوم والتقنية
King Abdullah University of Science and Technology

Introduction

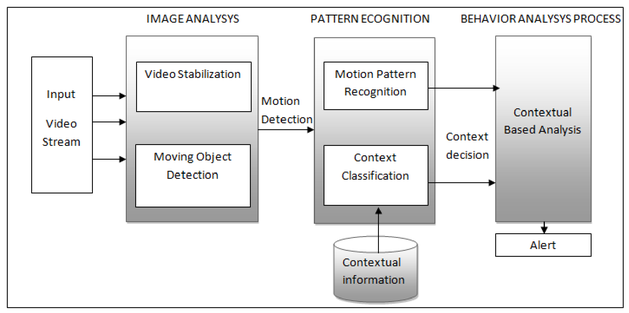
- Motion is a relative concept
  - Relative to objects in the image
  - Relative to image plane (static camera)
- Factors that make it difficult
  - Changing extrinsic and intrinsic camera parameters (Pan, tilt, translation, rotation, and zooming)
  - Illumination changes
  - Overlap of multiple motions with semi-transparency (dust, fog, mist, glasses, etc.)

Input: A real video captured from UAV
Goal: Abnormal behavior detection



Motion in video with moving camera.

Overview of the Proposed System



SIFT algorithm is used to detect and describe keypoints feature because of its invariability to scale changes, rotation changes and blur . Four steps:

- detection of extrema in scale-space
- localization and filtering of keypoints
- assignment of orientation
- and the generation of descriptors

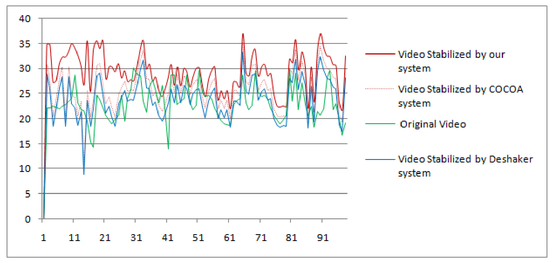
The function of DOG can be defined as follows:

$$D(x,y,\sigma) = (G(x,y,k\sigma) - G(x,y,\sigma)) * I(x,y) = L(x,y,k\sigma) - L(x,y,\sigma) \quad (1)$$

$G(x,y,\sigma)$ is a variable scale Gaussian, which is defined as

$$G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} exp^{-(x^2+y^2)/2\sigma^2} \quad (2)$$

Experimental Results



Graph of the PSNR of the original video and the stabilized video by Deshaker, COCOA System and our system.

PSNR compare the quality of the SV with Deshaker and COCOA. $PSNR$ between frame n and frame $n+1$ is defined as

$$PSNR(n) = 10\log_{10}\frac{I_{MAX}}{MSE(n)} \quad (3)$$

$$MSE(n) = \frac{1}{MN}\sum_{y=1}^{M}\sum_{x=1}^{N}[I_N(x,y) - I_{n+1}(x,y)]^2 \quad (4)$$

Sequence	Original ITF	Stabilized ITF		
		Proposed system	Deshaker system	COCOA System
Sequence 1	18.321	20.123	19.127	19.321
Sequence 2	17.876	19.542	19.210	19.452
Sequence 3	16.963	18.432	17.385	18.032
Sequence 4	17.856	19.874	18.985	19.432
Sequence 5	19.998	20.653	20.128	20.763

The measurement of Interframe Transformation Fidelity (ITF) is defined as

$$ITF = \frac{1}{N_{frame}-1}\sum_{k=1}^{N_{frame}-1} PSNR(k) \quad (5)$$

Detection Ratio(DR) and False Alarm Ratio (FAR) for performance evaluation.

$$DR = TP/(TP+FN) \quad (6)$$

$$FAR = FP/(TP+FP) \quad (7)$$

Video Stream	M O	Detection	DR	FAR
Video 1	2	4	1	0.2
Video 2	4	7	1	0.5
Video 3	3	7	1	0.6
Video 4	5	12	1	0.8
Video 5	7	14	1	0.34

Conclusion

1. Context modeling
2. For video stabilization, SIFT features are detected and matched for frame-to-frame to estimate the affinity matrix transformation.
3. Using a stabilized video, SIFT features are then used for moving object detection by compensation of camera motion and moving features to cluster independent motion.
4. Our result is improved by using filter kalman to estimate the motion of detected object.

References and Acknowledgements

[1] A. Walha, A. Wali, A. M. Alimi, *AVideo stabilization with moving object detecting and tracking for aerial video surveillance*, Multimedia Tools Appl, Vol. 74, Issue 1-2, pp 6745–6767, 2015.

[2] M. Ahmadi, A. Walha, A. Wali, A. M. Alimi, *A new Motion Estimation technique for Video Coding*, 15th International Conference on Intelligent Systems Design and Applications, ISDA 2015.

[3] A. Walha, A. Wali, A. M. Alimi, *Moving Object Detection System in Aerial Video Surveillance*, Advanced Concepts for Intelligent Vision Systems - 15th International Conference, ACIVS 2013.