Real Time Abnormal Motion Detection in Surveillance Video

Nahum Kiryati Tammy Riklin Raviv
Yan Ivanchenko Shay Rochel

Vision and Image Analysis Laboratory
School of Electrical Engineering
Tel Aviv University

Background

- Huge video surveillance system can apply hundreds and even thousands of cameras.
- To avoid communication bottlenecks, the acquired video is often compressed by a local processor within the camera, or at a nearby video-server.
- The compressed video is then transmitted to a central facility for storage and display.

Motivation

Difficulties

- Extensive human monitoring of the incoming video channels is impractical, expensive and ineffective
- Storing a large capacity of data is problematic

Solution

- Automatic systems that trigger recording or video transmission and attract the attention of a human observer to a particular video channel.
- Static scenes - Motion detection
- Dynamic scenes - Abnormal motion detection

Examples

- Airports
- Land transportation terminals
- Roads
- Office buildings
- Private usage

Constraints & Requirements

- Real time operation
- High reliability
- Low cost hardware
- Limited computation power

Related work

Motion Segmentation and abnormal motion behavior
Ermis et al ICIP 2008

Key Ideas

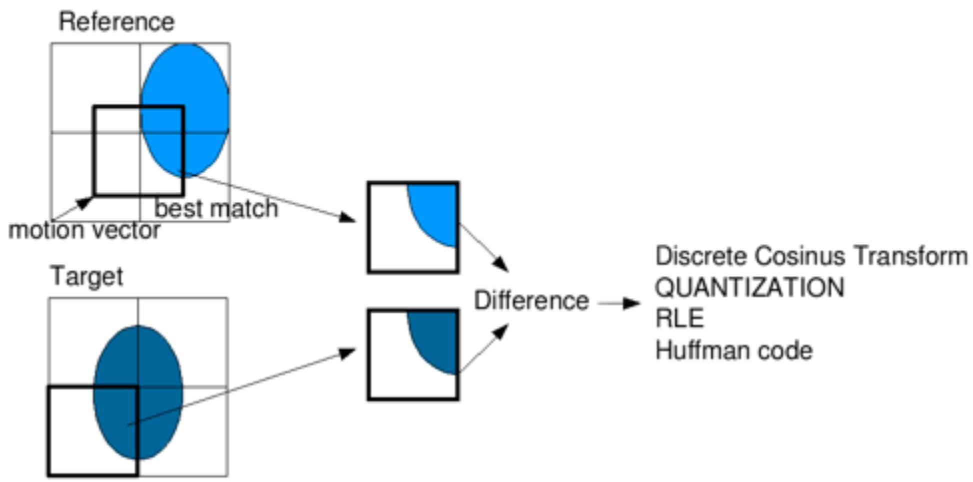
- Avoid segmentation or tracking
- Use the macro-block motion vectors that are generated anyway as part of standard video compression methods
- Derive motion features from the motion vectors
- Estimate the statistical distributions of the motion feature vectors that characterize normal activity during training
- Unlikely feature vectors during online operation indicate abnormal motion.

From Video to Motion Vectors

Video compression - Eliminating spatial & temporal redundancy

Intra-frame – compressed as a full still image

Inter-frame - represented by macro-block displacement vectors (motion vectors) relative to the reference frame and an error image.



http://en.wikipedia.org/wiki/Inter_frame

From Motion Vectors to Motion Features

An inter-frame contains N macroblocks.

Each macroblock (i, j) $i = 1 \ldots i_{\max}, j = 1 \ldots j_{\max}$

is associated with a motion vector $V_{i,j} = (V_{X_{i,j}}, V_{Y_{i,j}})$

Construct a small set f^l of $M << N$ features.

Define $\left|V_{i,j}^l\right| = \sqrt{(V_{X_{i,j}}^l)^2 + (V_{Y_{i,j}}^l)^2}$ and $\Phi_{i,j}^l = \arctan(\frac{\mathrm{V}_{\mathrm{Y}_{\mathrm{i,j}}}^{\mathrm{l}}}{\mathrm{V}_{\mathrm{X}_{\mathrm{i,j}}}^{\mathrm{l}}})$

as the magnitude and the direction of a motion vector $V_{i,j}^l$

From Motion Vectors to Motion Features

1. Total absolute motion $f_{TAM}^{l} = \sum_{i,j} \left| V_{i,j}^{l} \right|$

Divide the frame into k sub-regions A_k

2. Area of dominant motion $f_{ADM}^{l} = \hat{k} = \underset{k}{\operatorname{argmax}} \left(\sum_{i,j \in A_k} \left| V_{i,j}^{l} \right| \right)$

3. Motion homogeneity $f_{MH}^{l} = \dfrac{\max\limits_{k} \sum\limits_{i,j \in A_k} \left| V_{i,j}^{l} \right|}{f_{TAM}^{l} + \varepsilon}$

From Motion Vectors to Motion Features

The range of motion directions $\{-\pi,\pi\}$ is divided into R equal fractions of size $\nabla\varphi = \frac{2\pi}{R}$.

Let $r = 0 \ldots R-1$ be the angular fraction index

4. Principal motion direction

$$f^{l}_{PMD} = \hat{r} = \arg\max_{r} \sum_{i,j} \left(\left| \Phi^{l}_{i,j} - r\Delta\varphi \right| < \frac{\Delta\varphi}{2} \right)$$

5. Dominance of principal motion direction

$$f^{l}_{DPM} = \frac{\sum_{i,j \in D_r} \left| V^{l}_{i,j} \right|}{f^{l}_{TAM} + \varepsilon}, \quad D_r = \left\{ i, j \middle\| \Phi_{i,j} - \hat{r}\Delta\varphi \right| < \frac{\varphi}{2} \right\}$$

Manually designed features?

Why ?

Machine learning algorithms (such as Boosting or SVM)
learn set of discriminative features (or support vectors)
during training.
The `learned' set of features is adapted to the scene.

Why not?

Negative examples (in addition to the positive ones) are a must!
It is impossible to train the system on a comprehensive set of
abnormal motion patterns for every possible scene.

Probability Density Estimation

Construct an M dimensional histogram of the discrete feature vectors obtained during the training.

The normalized histogram defines the PDF of normal motion under the assumptions:

1) The training detects only normal motion
2) The training captures the variability of the possible normal motion patterns.

Histogram?

Why ?

Parametric methods for density estimation allow compact representation using statistical measures such as mean, variance and likelihood.
The data remains continuous.
Quantization (resolution) is not an issue.

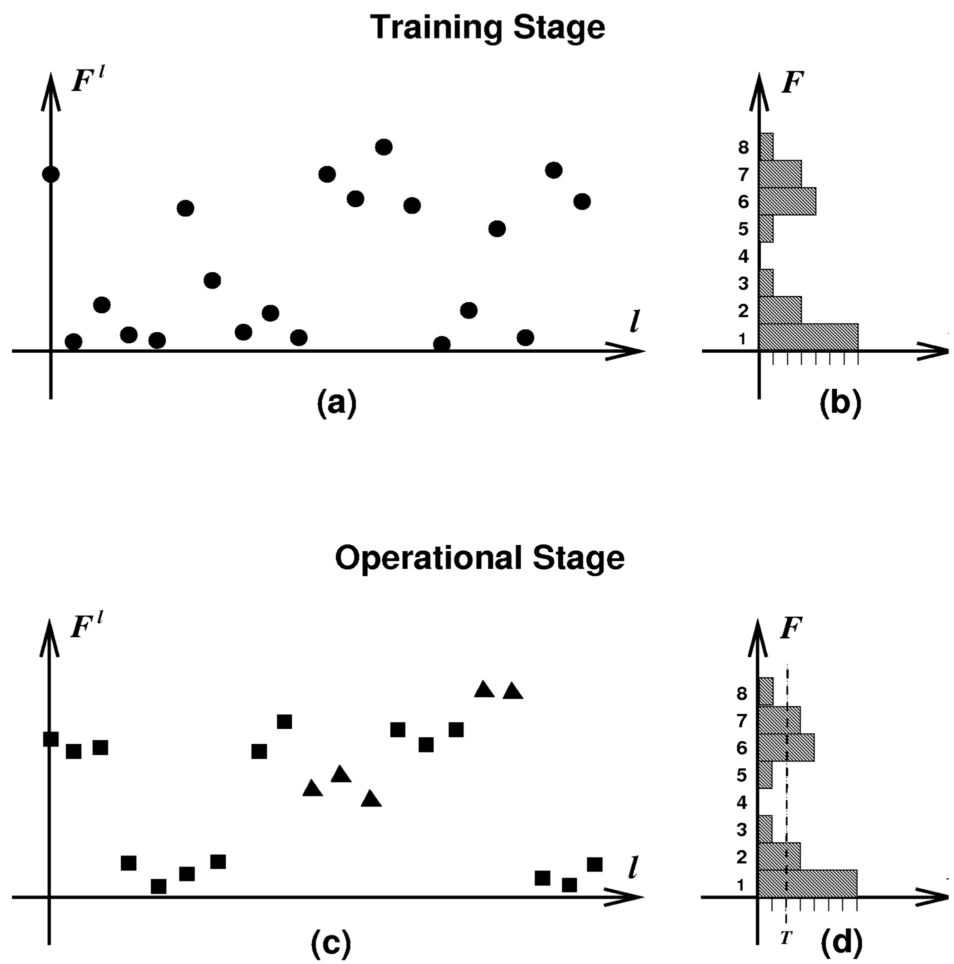
Why not

There is no reason to assume parametric distribution of the data
Histograms are the fastest and simplest non-parametric estimation methods.

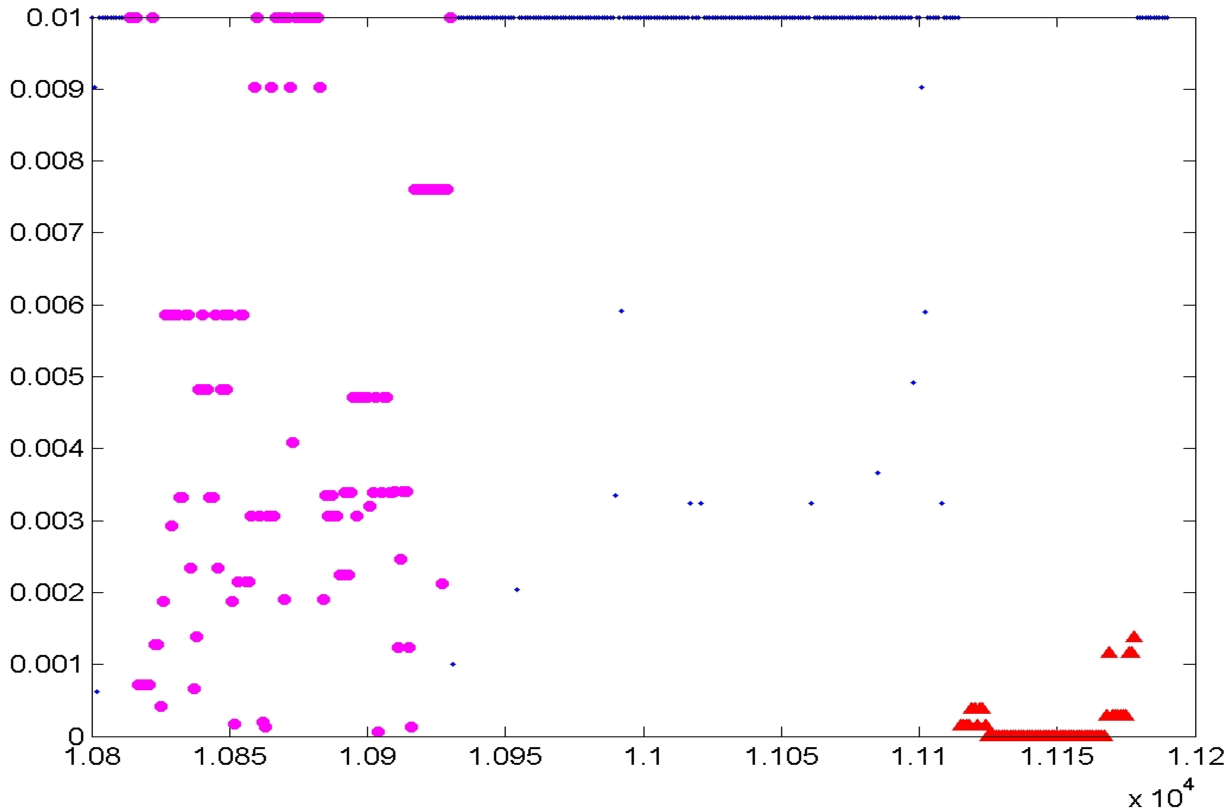
Abnormal Motion Detection

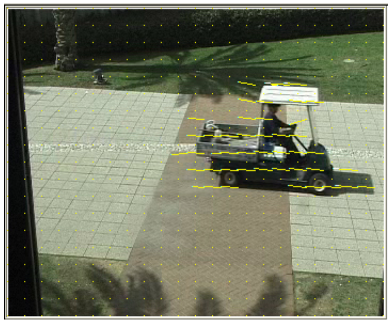
- Set a threshold T to define the unlikely feature vectors (histogram cells).
- Compute the feature vector associated with each of the incoming frames at the operational phase.
- The system detects abnormal motion event after L consecutive unlikely feature vectors.

Abnormal Motion Detection



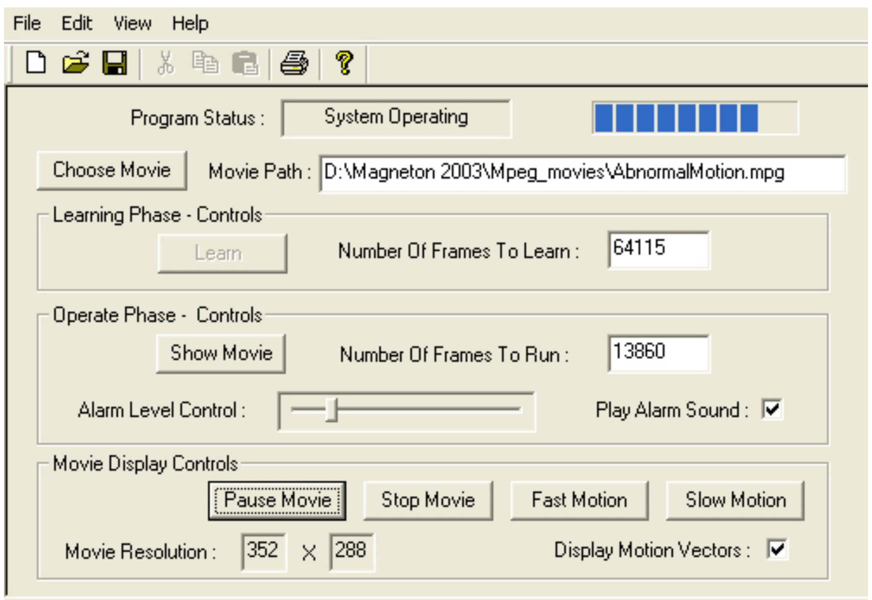
Event graph

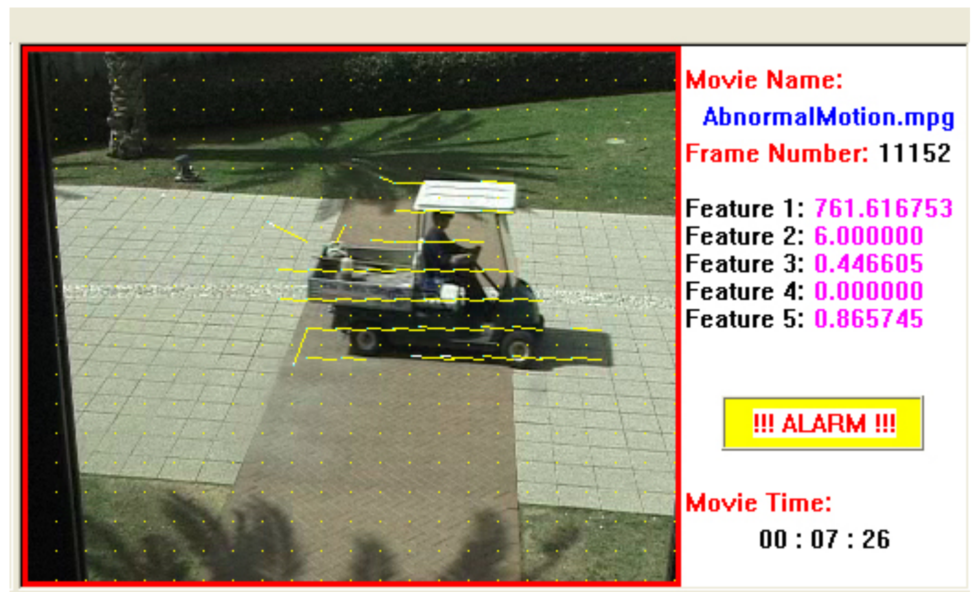

0.01
0.009
0.008
0.007
0.006
0.005
0.004
0.003
0.002
0.001
0
1.08
1.085
1.09
1.095
1.1
1.105
1.11
1.115
1.12
x 10^4



Experiments

Snapshot of the operational system





Experiments

Some technical details

- 50 minutes of video were acquired
- 41 minutes of normal traffic were used for training
- The 9 minutes test sequence contained normal and abnormal motion
- The movie was captured using SONY TRV900E PAL (25 pfs) digital video camera
- It was transformed to DV format an coded to MPEG-1 using generic
- MPEG-2 codec

Summary

The system presented is:

- Computationally efficient (75 CIF frames/sec on 2.8GHz PC)
- Reliable
- Operates on the compressed video stream
- Abnormal motion is not associated with a particular object in the scene
- Possible extensions – relate the concepts of normal and abnormal motion with time and causality

Questions?

The complete 9 minutes test movie can be found at:
http://abn-motion.axspace.com

We thank I. Dvir and D. Harari for the stimulating discussions.