

Abnormality Detection with Improved Histogram of Oriented Tracklets

Hossein Mousavi, Moin Nabi, Hamed Kiani, Alessandro Perina and Vittorio Murino

Pattern Analysis and Computer Vision (PAVIS)-Istituto Italiano di Tecnologia



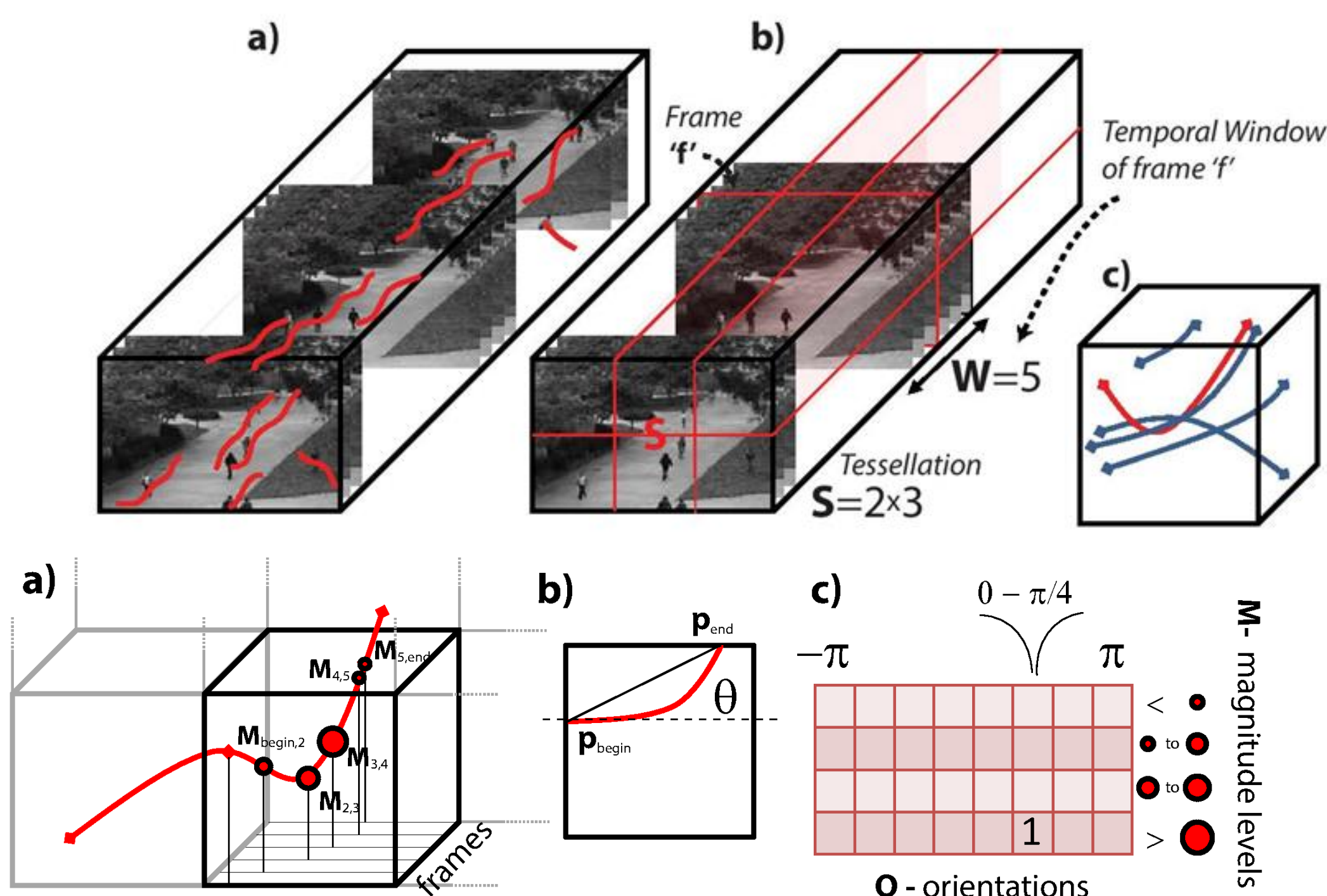
Abstract

Abnormality Detection: identifying abnormal behavioral patterns in videos.

Challenges:

- Difficult to detect and track individuals and objects due to clutters, low resolution, occlusion, etc.
- Not clear definition of abnormality, context dependent
- Scarcity of abnormal training samples

Histogram of Oriented Tracklets (WACV'15)

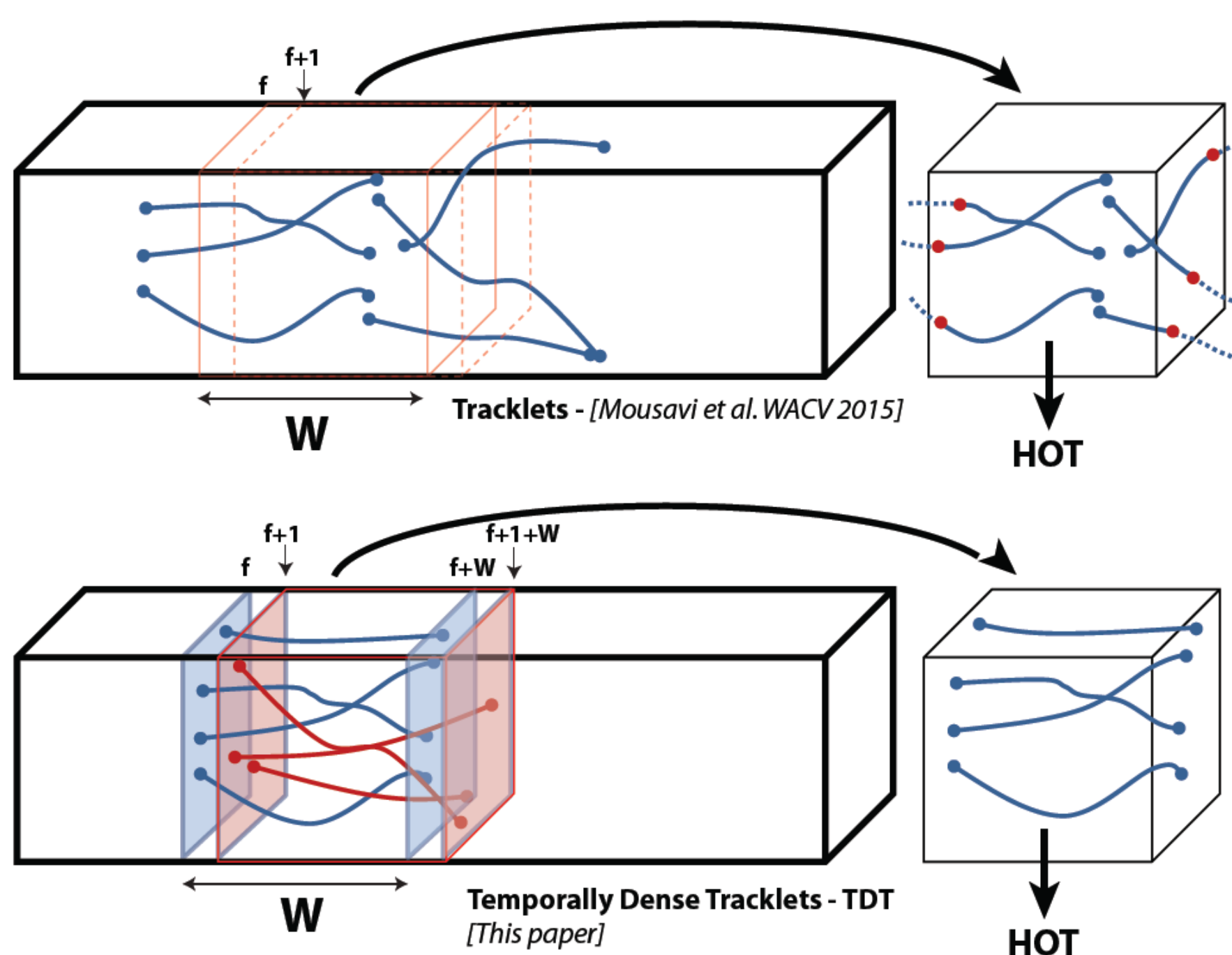


Drawbacks:

- Limited to salient points of the first frame,
- Tracking drift in the presence of occlusion,
- count-based histogramming: missing motion statistics

Improved HOT

Frame level salient point re-initialization



Motion statistics per each bin: mean and variance

$$mH_{o,m}^{s,f} = \frac{1}{J} \sum_{j=1}^J M^{j,s}$$

$$vH_{o,m}^{s,f} = \frac{1}{J} \sum_{j=1}^J (M^{j,s} - mH_{o,m}^{s,f})^2$$

Evaluation strategies

Fully bag of words (BW).

$$(mD)^f = \sum_s (mH)_{o,m}^{s,f} \quad \text{and} \quad (vD)^f = \sum_s (vH)_{o,m}^{s,f}$$

Per-frame, Per-sector (FS).

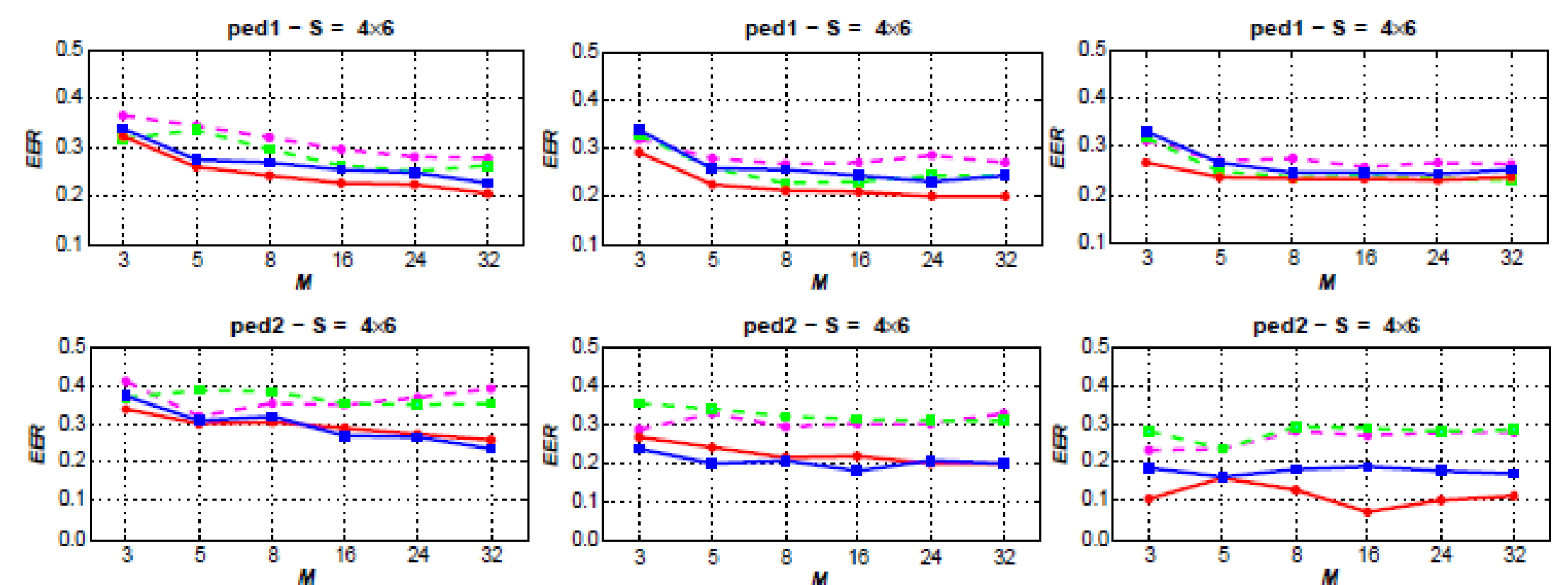
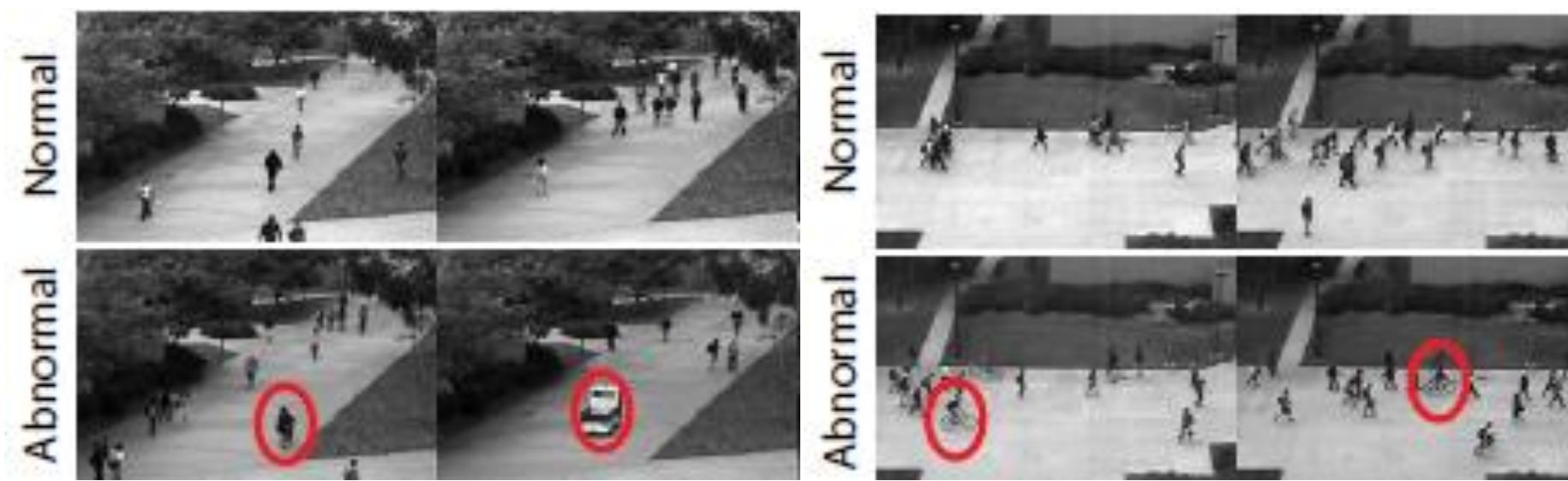
$$(mD)^f = \{(mH)_{o,m}^{1,f} | (mH)_{o,m}^{2,f} | \dots | (mH)_{o,m}^{S,f}\}$$

$$(vD)^f = \{(vH)_{o,m}^{1,f} | (vH)_{o,m}^{2,f} | \dots | (vH)_{o,m}^{S,f}\}$$

Experimental Results

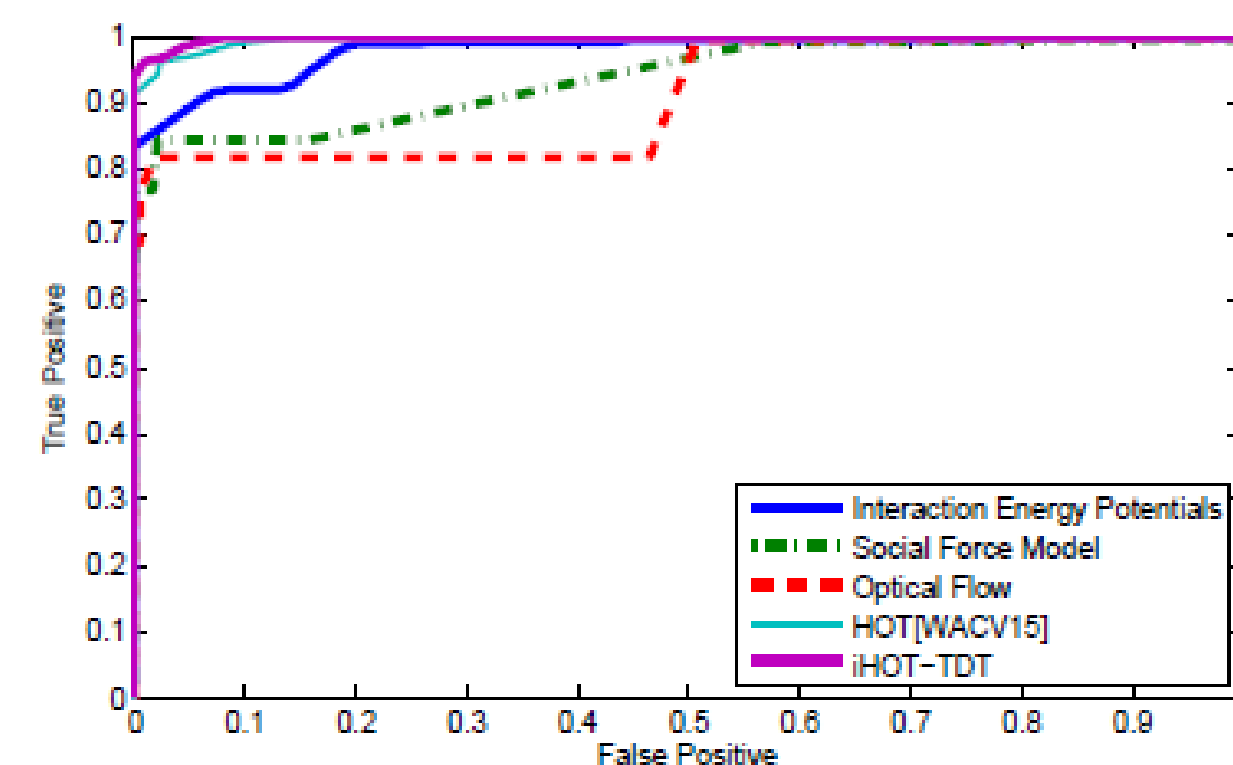
Three standard datasets: UCSD, BEHAVE and UMN, LDA + SVM

UCSD dataset



| ped1 | | ped2 | |
|--------------|--------|--------------|--------|
| Method | EER | Method | EER |
| MDT [19, 29] | 22.9% | MDT [19, 29] | 27.9% |
| SFM [9] | 36.5% | SFM [9] | 35.0% |
| LMH [24] | 38.9% | LMH [24] | 45.8% |
| HOT: BW [18] | 23.84% | HOT: BW [18] | 20.42% |
| HOT: FS [18] | 22.53% | HOT: FS [18] | 21.84% |
| iHOT: BW | 19.37% | iHOT: BW | 8.59% |
| iHOT: FS | 22.27% | iHOT: FS | 16.5% |

BEHAVE dataset



UMN dataset



| Dataset | iHOT-TDT | HOT [18] | SFM [9] | SR [26] | OF [9] | CI [27] |
|------------|----------|----------|---------|---------|--------|---------|
| scene-1 | 0.998 | 0.993 | 0.990 | 0.995 | 0.964 | n/a |
| scene-2 | 0.991 | 0.984 | 0.949 | 0.975 | 0.906 | n/a |
| scene-3 | 0.998 | 0.991 | 0.989 | 0.964 | 0.967 | n/a |
| all scenes | 0.994 | 0.991 | 0.960 | 0.978 | 0.840 | 0.990 |

Conclusions

Temporally Dense Tracklets:

- Initializing salient points in each frame
- Handling tracking drift, extracting salient points across entire video

Motion Statistics:

- Mean and variance of tracklet magnitudes per each bin

References

- Mousavi, H., Mohammadi, S., Perina, A., Chellali, R., Murino, V.: Analyzing tracklets for the detection of abnormal crowd behavior. WACV (2015)
- Li, W., Mahadevan, V., Vasconcelos, N.: Anomaly detection and localization in crowded scenes. PAMI (2014)
- Cui, X., Liu, Q., Gao, M., Metaxas, D.N.: Abnormal detection using interaction energy potentials. CVPR(2011)