

# MULTI-CHANNEL CORRELATION FILTERS FOR HUMAN ACTION RECOGNITION

HAMED KIANI<sup>1</sup>, TERENCE SIM<sup>1</sup> AND SIMON LUCEY<sup>2</sup>  
 {hkiani,tsim}@comp.nus.edu.sg, slucey@andrew.cmu.edu

<sup>1</sup>School of Computing, NUS, Singapore

<sup>2</sup>Robotics Institute, Carnegie Mellon University, USA

## ABSTRACT

In this work, we propose to employ multi-channel correlation filters for recognizing human actions (e.g. *walking, riding*) in videos. In our framework, each action sequence is represented as a multi-channel signal (frames) and the goal is to learn a multi-channel filter for each action class that produces a set of desired outputs when correlated with training examples. The experiments on the Weizmann and UCF sport datasets demonstrate superior computational cost (real-time), memory efficiency and very competitive performance of our approach compared to the state of the arts.

## CONTRIBUTIONS

- Extending canonical correlation filter theory to efficiently handle multi-channel signals
- A multi-channel detector whose training memory is independent of the number of training samples
- Superior performance to current state of the art correlation filters, and superior computational and memory efficiency in comparison to spatial detectors (e.g. linear SVM) with comparable detection performance

## QUALITATIVE RESULTS

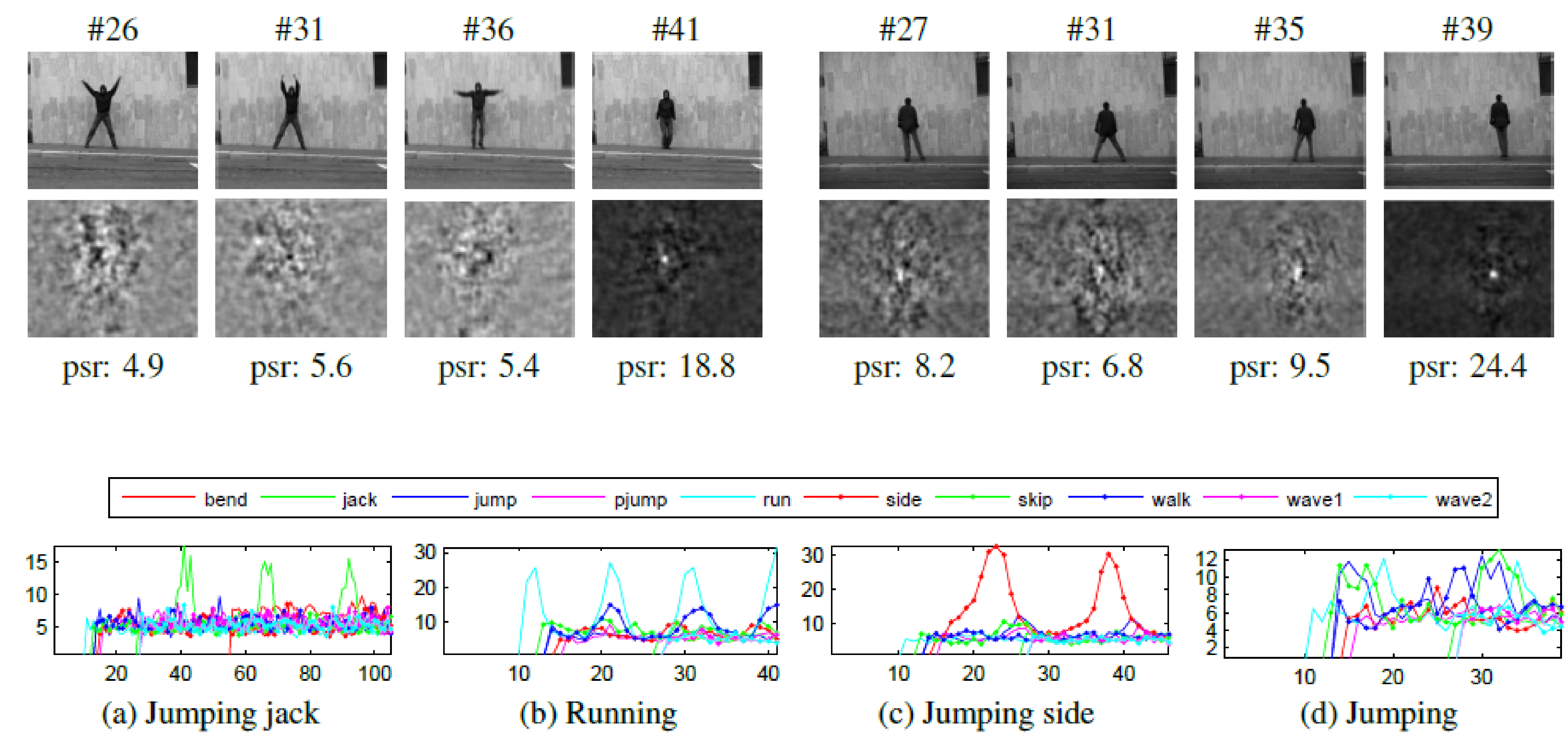
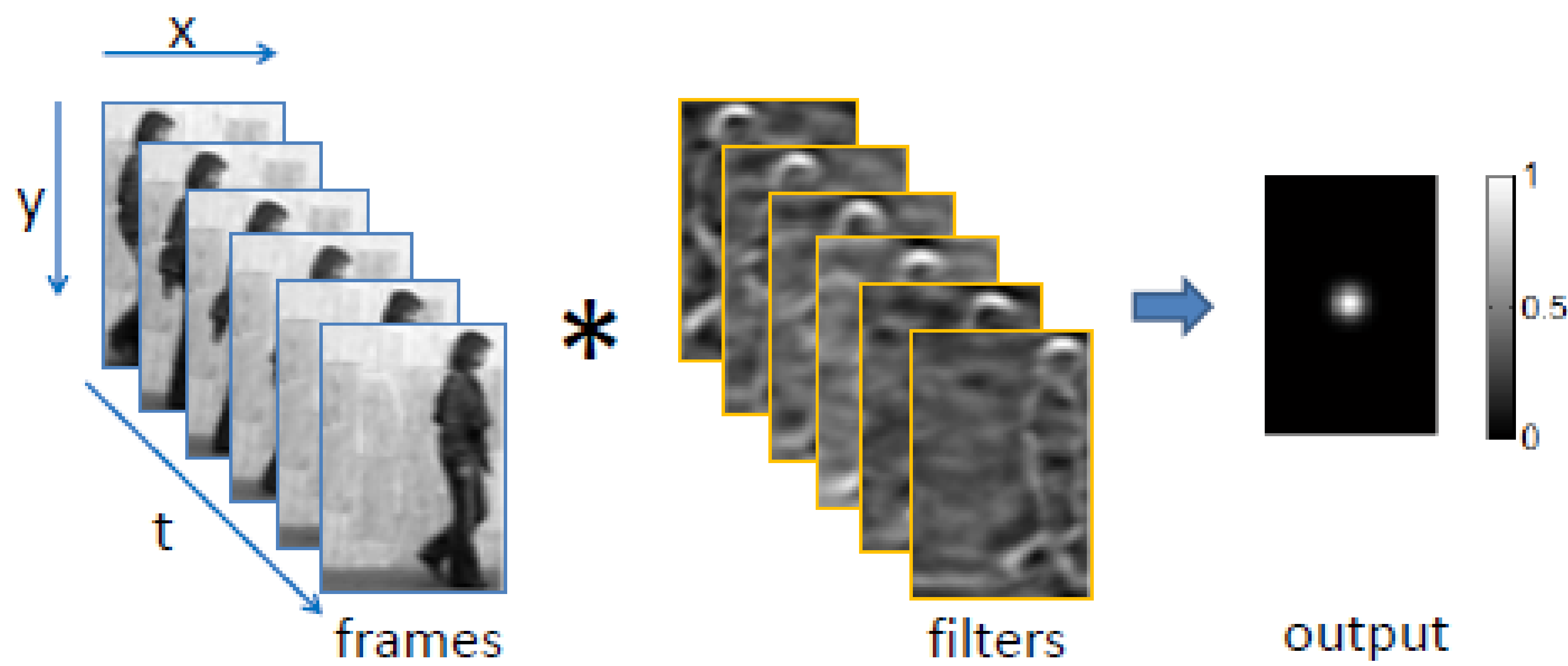


Figure 1. Comparing MCCF with SVM + HOG on the problem of pedestrian detection using Daimler dataset. Top: Memory usage (MB) of MCCF compared to SVM as a function of number of training images.

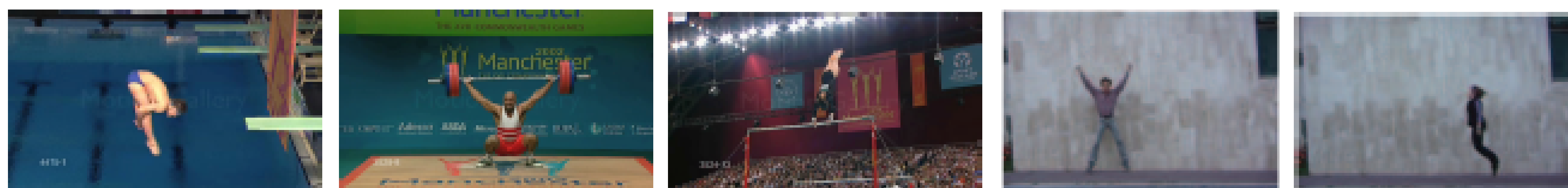
## PROPOSED FRAMEWORK



- MCCF Objective Function [1]:

$$\arg \min_{\mathbf{h}} \left\| \mathbf{y} - \sum_{k=1}^K \mathbf{h}^{(k)} * \mathbf{x}^{(k)} \right\|_2^2 + \lambda \sum_{k=1}^K \left\| \mathbf{h}^{(k)} \right\|_2^2$$

## THE WEIZMANN AND UCF SPORT DATASETS



## QUANTITATIVE RESULTS - 1

Bend	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Jack	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Jump	0.00	0.00	0.89	0.00	0.00	0.11	0.00	0.00	0.00
Pjump	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
Run	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00
Side	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00
Skip	0.00	0.00	0.00	0.11	0.00	0.89	0.00	0.00	0.00
Walk	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00
Wave1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00
Wave2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00

Diving	0.86	0.00	0.00	0.00	0.00	0.00	0.00	0.14	0.00
Golfing	0.00	0.80	0.00	0.00	0.00	0.00	0.20	0.00	0.00
Kicking	0.00	0.00	0.70	0.00	0.00	0.20	0.00	0.00	0.10
Lifting	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
Riding	0.00	0.00	0.00	0.00	0.92	0.08	0.00	0.00	0.00
Running	0.00	0.00	0.00	0.08	0.00	0.77	0.00	0.00	0.15
Skating	0.00	0.17	0.00	0.08	0.00	0.67	0.00	0.00	0.08
Swing bench	0.05	0.00	0.05	0.00	0.05	0.00	0.80	0.00	0.05
Swing bar	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.92	0.00
Walking	0.00	0.00	0.05	0.00	0.00	0.09	0.05	0.00	0.82

## QUANTITATIVE RESULTS - 2

Method	Weizmann	UCF sport
Huang et al.	100%	-
Cai et al.	98.7%	90.6%
Wang et al.	97.8 %	77.4%
Campos et al.	96.7 %	80.0%
Rodriguez et al.	86.6%	69.2%
Yeffet & Wolf	-	79.3%
Our method	97.8%	82.6%

Norm. intensity	Edge magnitude	Temp. derivative	HoG (5 bins)
89.4%	91.2%	92.3%	97.8%

## REFERENCES

- [1] H. Kiani, T. Sim and S. Lucey. Multi-Channel Correlation Filters. In ICCV '13.