Multi-Channel Correlation Filters For Human Action Recognition

Hamed Kiani[1], Terence Sim[1] and Simon Lucey[2]
{hkiani,tsim}@comp.nus.edu.sg, slucey@andrew.cmu.edu
[1]School of Computing, NUS, Singapore
[2]Robotics Institute, Carnegie Mellon University, USA

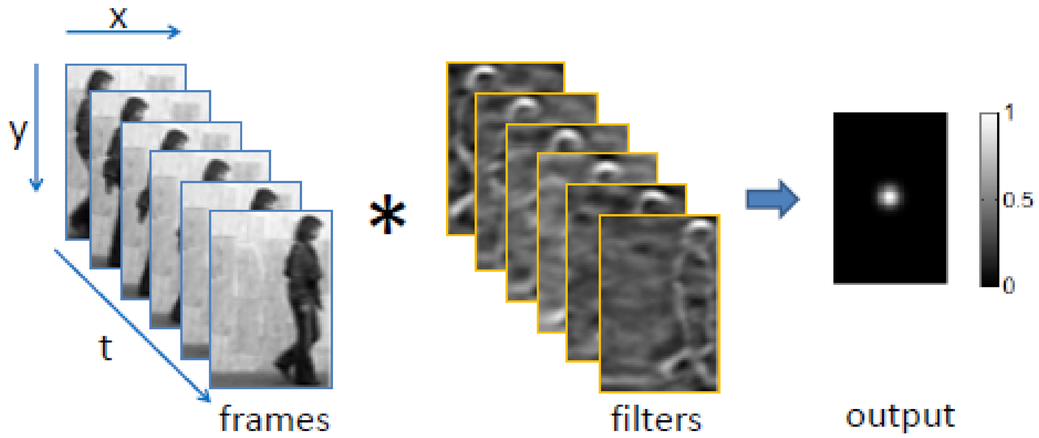
Abstract

In this work, we propose to employ multi-channel correlation filters for recognizing human actions (e.g. *waking*, *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

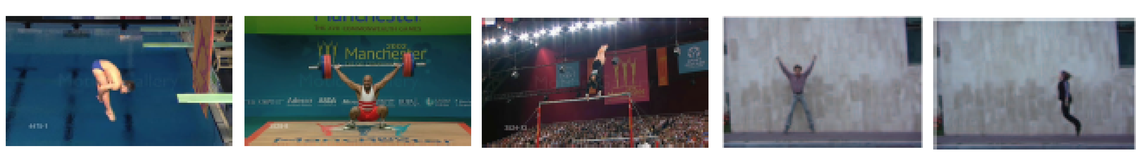
Proposed Framework



- 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



Qualitative results

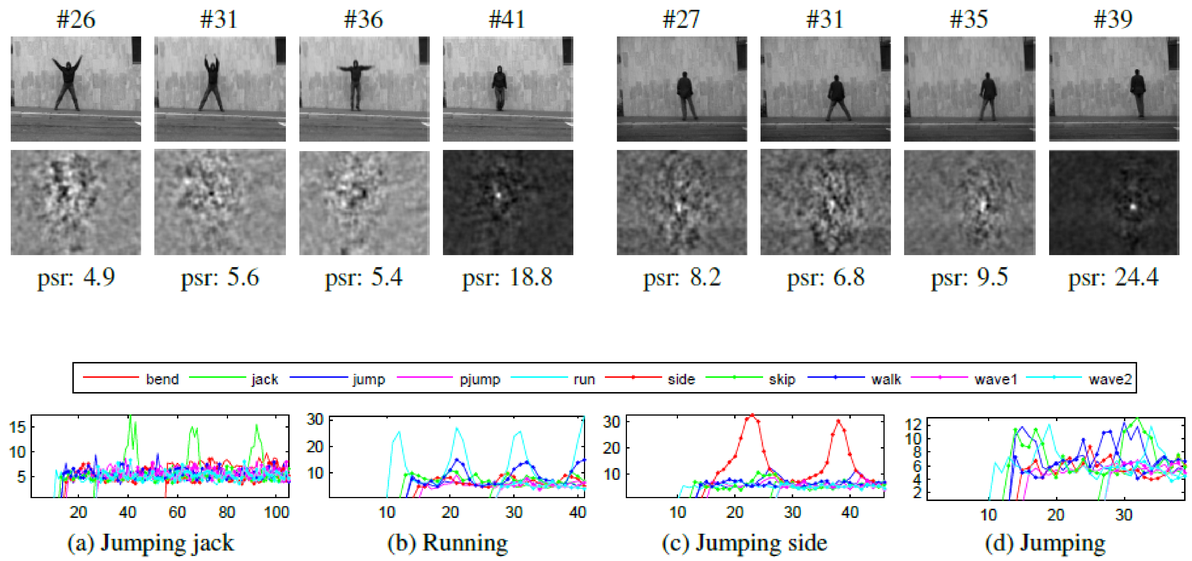
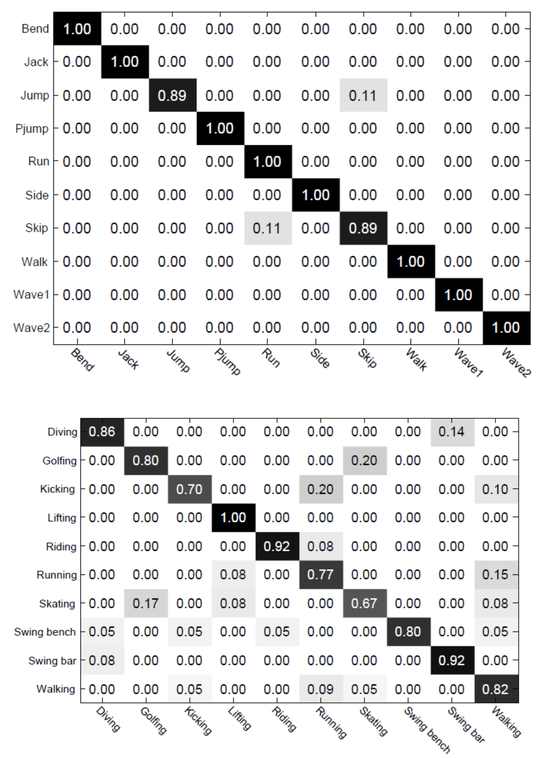


Figure 1. Comparing MCCF with SVM + HOG on the problem of pedestrian detection using Daimler dataset. Top: Memory usage (MB) of MCCF compared to SVM as a function of number of training images.

Quantitative Results - 1



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*.