



# 1. Problem: Human Action Recognition in Videos



a single frame/feature that semantically summarizes the sequence.

## **Our Contributions**

subspace that provides:

i) A low-rank approximation to X and

from the origin in  $U_X$ ).

**2.** We propose to use the subspace  $U_X$  as the action descriptor for X

## 2. Related Work

number of subspaces

properties of data

## 3. Video Pre-processing

# Generalized Rank Pooling for Activity Recognition Anoop Cherian<sup>1,2</sup> Basura Fernando<sup>1,2</sup> Mehrtash Harandi<sup>2,3</sup> Stephen Gould<sup>1,2</sup> <sup>1</sup>Australian Centre for Robotic Vision, <sup>2</sup>The Australian National University, <sup>3</sup>Data61/CSIRO For code and data: contact anoop.cherian@anu.edu.au





MPII Cooking Activities Dataset: ~5K videos, 65 classes

# **8.** Experiments: Results

Method/Dataset	FLOW	RGB
MPII	mAP (%)	mAP (%)
GRP (w/o constraints)	51	48.9
<b>GRP-Grassmann</b>	52.1	50.3
JHMDB	Avg.Acc.(%)	Avg.Acc.(4
GRP (w/o constraints)	59.4	41.8
<b>GRP-Grassmann</b>	64.2	42.5

Comparison between the influence of GRP on FLOW and RGB separately

Algorithm	Avg. Acc. (	
P-CNN	61.1	
P-CNN + IDT-FV	72.2	
Action Tubes	62.5	
Stacked Fisher Vectors	69.03	
IDT + FV	62.8	
Higher-order Pooling	73.3	
GRP (w/o constraints)	64.1	
GRP	70.6	
GRP + IDT-FV	73.7	

#### JHMDB Dataset

Algorithm	Avg. A	
Two stream	5	
Spatio-Temporal ResNet	7	
Temporal Segment Networ	ks 6	
TDD + IDT-FV	6	
Dynamic Image + IDT-F	V 6	
Hier. Rank Pooling + IDT-I	EV 6	
Dynamic Flow + IDT-FV	6	
GRP (w/o constraints	s) 6	
GRP	6	
GRP + IDT-FV	6	
GRP + IDT-FV (ResNet-	152) 7	
HMDB-51 Dataset		

### IEEE 2017 Conference on **Computer Vision and Pattern** Recognition



**HMDB Dataset** : ~6K videos, 51 classes **JHMDB Dataset :** ~1K videos, 21 classes **UCF101 Dataset:** ~13K videos,101 classes

