Exemplar-SVM for Action Recognition

Christina Peterson, Rui Hou
(clp8199@knights.ucf.edu, houray@gmail.com,)
University of Central Florida

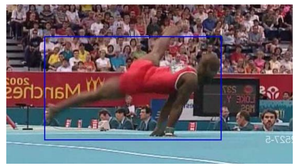
1. Problem

- **Objective:**
  - Action Recognition in Videos Using a Series of Exemplar-SVMs
- **Motivation**
  - Only one labeled video as a positive sample is required to train an Exemplar-SVM, reducing the amount of required labeled data
  - Allows for training with very little human supervision

Exemplars for Swing Bench Action Class



Exemplar 1



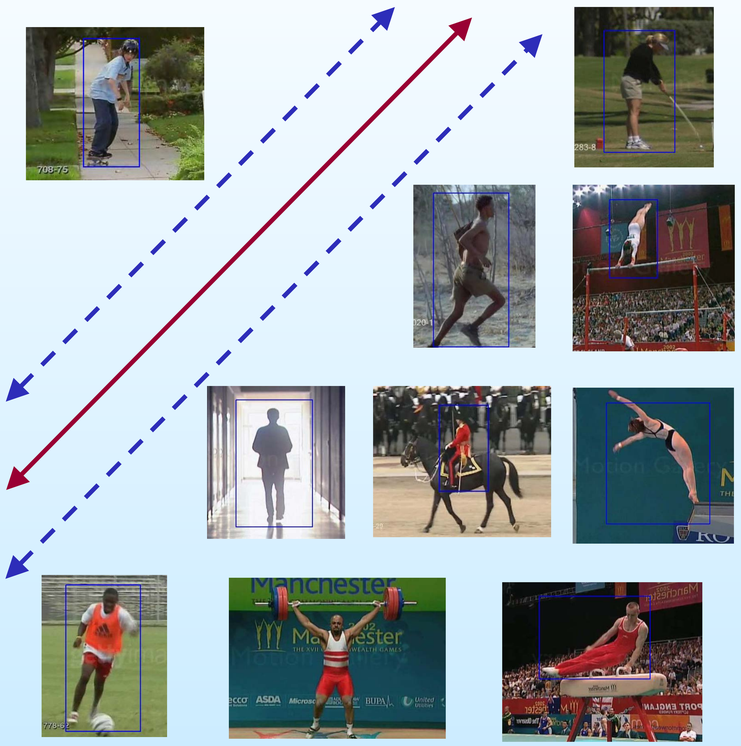
Exemplar 2



Exemplar 3

2. UCF Sports Actions Dataset

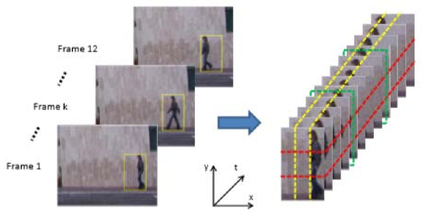
- **Set of ten actions collected from various sports which are typically featured on broadcast television channels**
- **Each exemplar is trained against the other actions in the training dataset**



3. Overall Approach

(1) Bag of Words

- **Collect Low Level Features within Bounding Box**
  - **STIP (Space-Time Interest Points)**
    - **Histogram of Oriented Gradient (HOG)**
    - **Histogram of Optical Flow (HOF)**
  - **Color**
  - **Texture**
- **Cluster by K-means**
  - **128 for Color, 256 for Texture, 1000 for STIP**
- **Construct Feature Vector**
  - **Create histogram of cluster centers per feature for each cell in bounding box**
  - **Normalize Histograms**
  - **Final Feature Vector is (128 + 256 + 1000) x 10 = 13,840 dimensions**



(2) Train Exemplar-SVMs

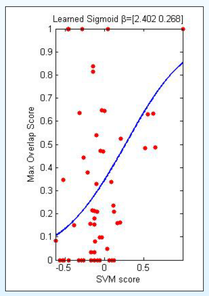
- **Select a minimal number of videos from each action class which serve as a good representation for the action**
- **Train a series of Linear-SVMs where the selected exemplar video is the positive sample and videos from other actions are the negative samples**
  - **Negative samples also include background volumes**

(3) Calibration

The calibration step is performed in order to make the raw scores of the Exemplar-SVMs compatible

- **Platt's Method**
  - **Fit a probability distribution to a validation set**

$$P(x) = \frac{1}{1 + e^{-\alpha(x-\beta)}}$$



(4) Calibrated Exemplar-SVM Scores as Mid-Level Features

All Exemplar-SVMs are tested on the validation set, and the calibrated scores of all exemplars are concatenated together to form a mid-level feature vector for the validation video

	Label	Action 1 Exem. 1	Action 1 Exem. 2	Action 1 Exem. 3					Action 10 Exem. 1	Action 10 Exem. 2	Action 10 Exem. 3
1	1	0.13	0.55	0.99	0.11	0.67	0.29	0.91	0.32	0.78	0.05
	1	0.44	0.82	0.99	0.36	0.18	0.84	0.22	0.53	0.62	0.86
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	10	0.38	0.52	0.02	0.99	0.55	0.74	0.81	0.94	0.26	0.12
10	10	0.99	0.02	0.34	0.28	0.52	0.96	0.83	0.20	0.58	0.32

(5) Train Multi-Class SVM based on Calibrated Exemplar-SVM Scores

A Multi-Class SVM is trained on the mid-level features vectors of the validation set

4. Experimental Evaluation

Testing was performed on the UCF Sports Actions Dataset

- **Exemplar-SVM set was held consistent**
- **The remaining videos from each action class were used to train and test the Multi-Class SVM**
- **The Multi-Class SVM was trained using a two-fold cross validation method**
  - **Videos split into two groups**
  - **First group used for training, second group used for testing**
  - **Second group used for training, first group used for testing**
  - **Results between test cases were averaged**

5. Results

- **Recognition Accuracies on UCF Sports Actions Dataset**

Method	Accuracy (%)	Diving	Golfing	Kicking	Lifting	Riding	Running	Skating	Swing-Bench	High-Swing	Walking
Rodriguez *et al.*	69.2	68	61	66	66	75	74	73	-	-	-
Yeffet and Wolf	79.3	100	61	65	67	67	69	92	-	-	86
Le et al.	86.5	100	77.8	80	100	66.7	69.2	83.3	100	100	90.9
Wu et al.	91.3	100	88	100	100	67	93	84	95	93	91
Action Bank	95.0	100	100	100	83	100	91	92	100	89	86
Standard Multiclass-SVM	90.2	100	50	71.4	100	100	100	100	100	100	50
Exemplar-SVMs (Ours)	94.4	96.6	87.5	91.7	100	83.3	100	98.1	97.9	94.4	82.5

- **Confusion Matrix for Exemplar-SVM**

	Di	Go	Ki	Li	Ho	Ru	Sk	Sb	Hs	Wa
Diving	96.6	1.9								1.5
Golf		87.5	8.3		4.2					
Kick		8.3	91.7							
Lift				100						
Horse-Ride					83.3		16.7			
Run						100				
Skateboard		1.9					98.1			
Swing-Bench								97.9	2.1	
High-Swing								2.8	94.4	2.8
Walk		3.3	3.3				10.9			82.5