Mining Representative Actions for Actor Identification

Wenlong Xie, Hongxun Yao,

Xiaoshuai Sun, Sicheng Zhao, Tingting Han, Cheng Pang

Harbin Institute of Technology, China

wlxie@hit.edu.cn

Outline

- Introduction
  - Actor identification
  - Motivation
- Method
  - Can we extract representative actions of each actor from a movie?
  - How can these actions help for actor identification?
- Experiments

Outline

Actor Identification

☐ Identification of characters in TV series or movies

☐ Important for many higher level multi-media analysis tasks

Sheldon

Penny

Motivation

☐ Previous methods mainly focused on static features

☐ Perform poorly when the appearances of actors are hard to detect or changes greatly over time

☐ Try to mine some dynamic features for actor identification

☐ Using our method as complementary to formal method, performance improved

What makes an actor impressive?

Faces
Clothes
Figures
• • •

Could it be vivid and dynamic?

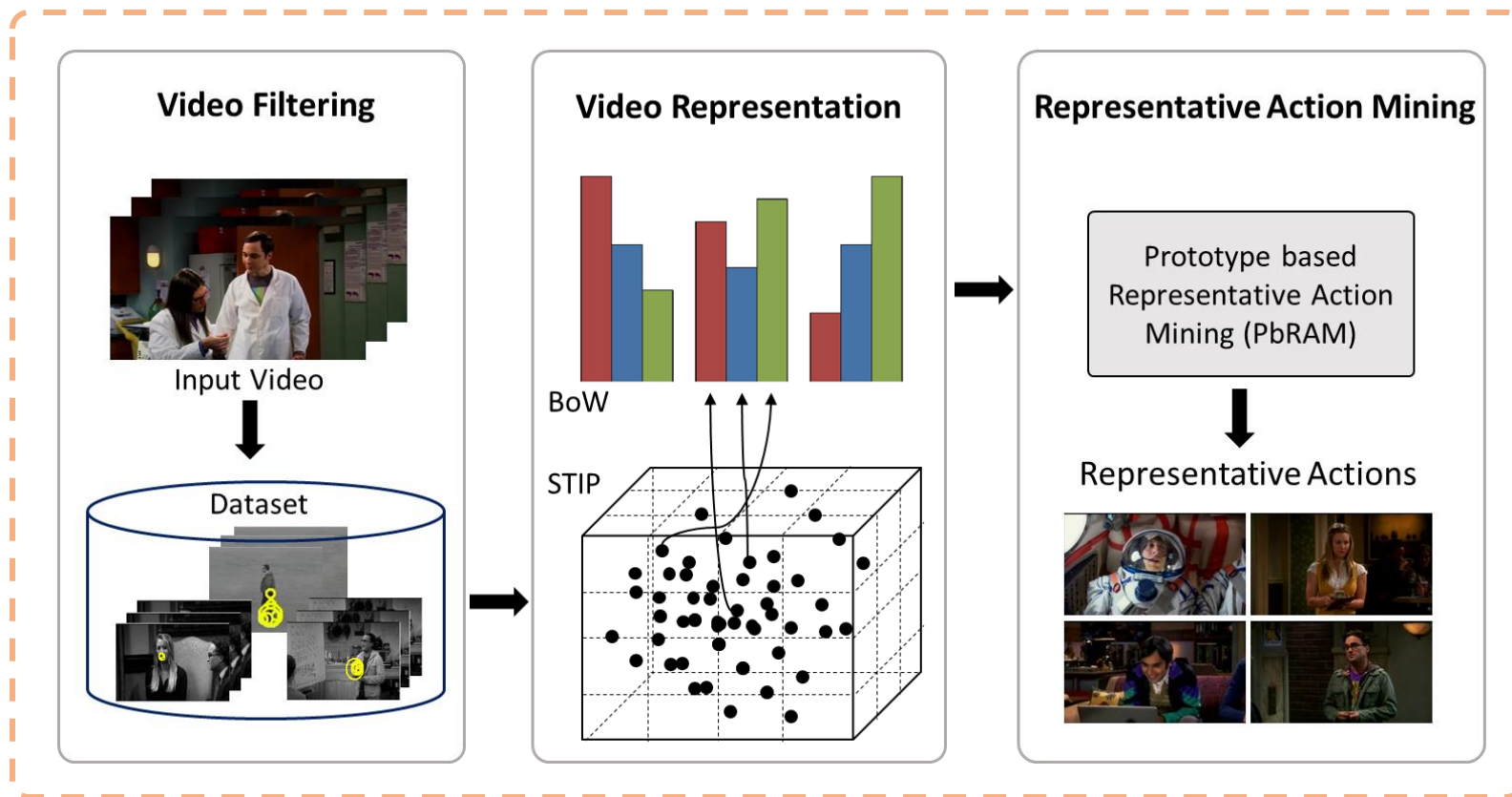# Representative actions of *Sheldon*

Outline

☐ Introduction
  ☐ Actor identification
  ☐ Motivation

☐ **Method**
  ☐ Can we extract representative actions of each actor from a movie?
  ☐ How can these actions help for actor identification?

☐ Experiments

Outline

Mining representative actions

Prototype based representative action mining

□ Representative actions:

— Occur **frequently** in **same** category

— **Distinguish** itself with others in **different** categories

□ Prototype [1] based model:

— **Most** representative of items **inside**

— **Least** representative of items **outside** a category

[1] Rosch, "Principles of categorization," Concepts: core readings, pp. 189–206, 1999.

Prototype based representative action mining

☐ Prototype: most represent inside, least represent outside

☐ Set k-means centers as the initial prototypes

☐ SVM to update the prototypes

positive exemplar ⟷ its prototypes

negative exemplar ⟷ other actors' prototypes

updated prototypes ⟷ actions with high test score

Algorithm 1: Prototype based representative action mining

Input: Actor number N_a, top N_r representative actions need to mine, BoW representation of each actor's actions $H_i = \{hist_j\}_{j=1}^{N_i}, i = 1, ..., N_a$, number of clustering centers k, threshold θ to split prototypes with other actions

Output: Representative actions index matrix $\mathbf{R}(N_a, N_r)$

1   **for** $i = 1, ... N_a$ **do**

2 | Run k-means on H_i to obtain cluster centers $C_i = \{c_j\}_{j=1}^{k}$;   → Prototypes initialize

3   **end**

4   Unite all the prototypes $C = \{C_i\}_{i=1}^{N_a}$;

5   **for** $i = 1, ... N_a$ **do**

6 | $Classifier_i \leftarrow$ trainSVM$(C_i, C - \{C_i\}, cosine)$;

7 | $score_i \leftarrow$ test all item in H_i with $Classifier_i$;   → Actions ranking

8 | Descending sort $score_i$;

9 | $N = \max\{j\}$, s.t. $score_i(j) > 0$;

10 | **if** $N \geqslant N_i * \theta$ **then**

11 | $\mathbf{R}(i, N_r) \leftarrow$ index of top N_r ranking actions of $score_i$;

12 | break;

13 | **end**

14 | $C_i \leftarrow$ corresponding top k score actions in H_i;   → Prototypes update

15   **end**

16   **return** $\mathbf{R}(N_a, N_r)$;

Representative actions mined by our algorithm

Representative actions by human judgement

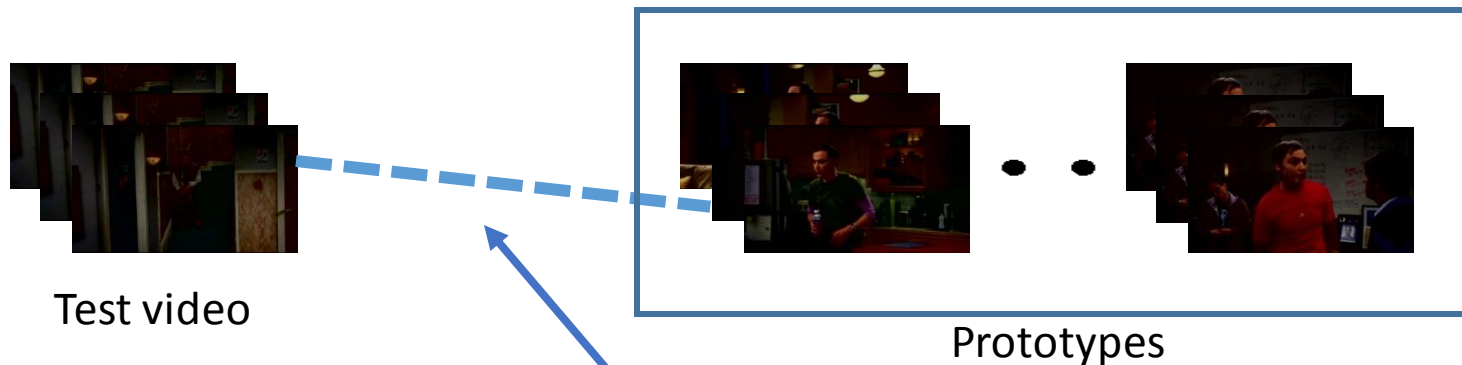# Representative actions of *Sheldon*

Chosen manually

Mined by algorithm

Outline

□ Introduction
  □ Actor identification
  □ Motivation
□ **Method**
  □ Can we extract representative actions of each actor from a movie?
  □ **How can these actions help for actor identification?**
□ Experiments

Actor identification with representative actions

Test video

Prototypes

$$score_i(\mathrm{x}^*, \mathrm{r}) = \exp\left\{-\lambda \boxed{D\left(\mathrm{x}^*, \mathrm{r}\right)}\right\}$$

λ : scaling constant

$$score\left(\mathrm{x}^*\right) = \underbrace{\omega_1 * score_1\left(\mathrm{x}^*, \mathrm{r}\right)}_{\text{face}^{[2]}} + \underbrace{\omega_2 * score_2\left(\mathrm{x}^*, \mathrm{r}\right)}_{\text{action}}$$

ω : weight

face [2]   action

[2] Josef Sivic, Mark Everingham, and Andrew Zisserman, "who are you?-learning person specific classifiers from video," *CVPR*, 2009, pp. 1145–1152.

Outline

- Introduction
  - Actor identification
  - Motivation
- Method
  - Can we extract representative actions of each actor from a movie?
  - How can these actions help for actor identification?
- **Experiments**

Experiments
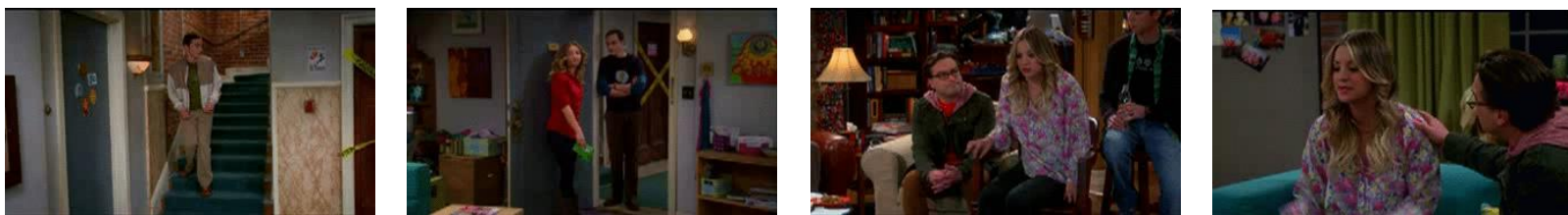

New dataset — **BigBangActions**

☐ Each clip corresponds to an actor's action

☐ Five actors, each has around 1000 clips

☐ Each clips contains 100 — 1000 frames

Actor identification examples

A success, **R** success

Frontal face & Temporally stable

A failed, **R** success

Hard to detect face or not temporally stable

A — Appearance matching method
R — Representative actions matching method

Experiments

- Consistent on different actors, average 75%

- Best performance on *Sheldon*

Confusion matrix of our proposed method for actor identification

Experiments

Method	Accuracy(%)
Sivic et al. [2]	59.78
Representative Actions	34.81
Combination	**75.32**

Performance comparison between the proposed method and the baseline based on actors' appearance

[2] Josef Sivic, Mark Everingham, and Andrew Zisserman, "who are you?-learning person specific classifiers from video," CVPR, 2009, pp. 1145–1152.

Conclusion

- Proposed a prototype based method to mine representative actions

- Using representative actions as complementary, our method greatly improves the performance of actor identification

- Constructed a new dataset BigBangActions and will be released soon

Thank you!

wlxie@hit.edu.cn