Human Detection in Digital Videos Using Motion Features Extractors

Rodrigo F. S. C. Oliveira & Carmelo J. A. Bastos-Filho
Universidade de Pernambuco (UPE), Recife, Brazil

ABSTRACT

Human detection in digital videos is a challenging task, because the human appearance may vary widely. Several algorithms to detect humans in digital images have been recently developed, such as the Aggregated Chanel Features (ACF) and most of them are based on features related to the human body shape. These algorithms give the best results regarding accuracy, but generate many false alarms. In this paper, we propose to introduce motion features to the ACF in order to accurately detect humans in digital videos. We demonstrate that our proposal returns more accurate results than the original ACF and presents a reduction in false alarms.

1 INTRODUCTION

The objective of human detection consists of determining the exact location and extent of people in images or video. It is an essential task for many applications, such as autonomous robots, pedestrian protection, and surveillance.

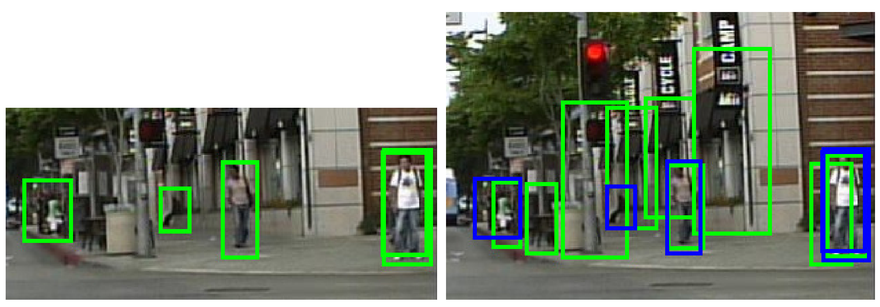

Figure 1: Left: Perfect detection. Right: Detection provided by the ACF.

Challenges:

- Diversity in human appearance such as height, body shape, color, and texture.
- Complexity and Dynamism of urban scenes: complex background, viewpoint and movement of cameras, uncontrolled and unbalanced illumination and occlusion.

Aggregated Channel Features (ACF):

- The ACF human detector was proposed by Dollar et al. [1] and is on of the most accurate human detector in the literature.
- Pipeline: Sliding Window, Fast Feature Pyramid and Channel Features Representation for color and shape (ex. HOG) and AdaBoost as classifier.
- The ACF can achieve good accuracy rates, but produces a reasonable amount of false positives (e.g. background pieces misclassified as human) as shown in the right image on Figure 1.

2 NEW ACF EXTENSION

We extend the ACF Human detector by applying the following motion feature extractors:

- Motion Boundary Histogram (MBH) [2]
- Internal Motion Histogram Central Difference (IMHcd) [2]
- Weak Stabilized Temporal Difference (WSTD) [3]
- Entropy of Histogram Oriented Gradient (EHOG)

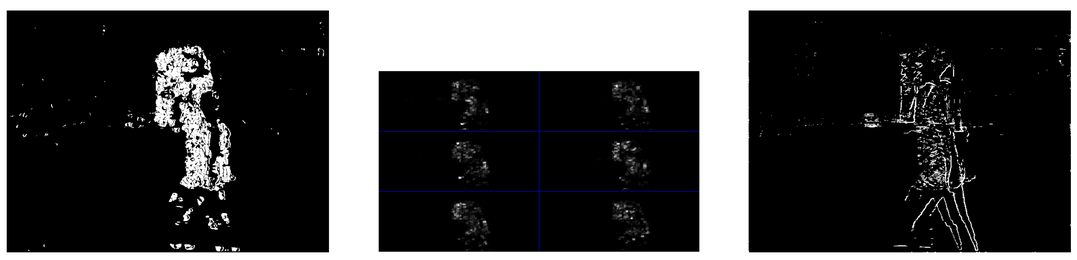

Figure 2: Visualization of features computed for two adjacent frames. First row: Adjacent frames. Second row from left to right: MBH, IMHcd, WSTD.

3 EXPERIMENTAL RESULTS

Motion Feature Extractors Evaluation: We observed the assessed motion feature extractors have presented similar performance for high detection threshold values, but IMHcd reaches the best performance regarding log average Miss-rate as shown in Figure 3.

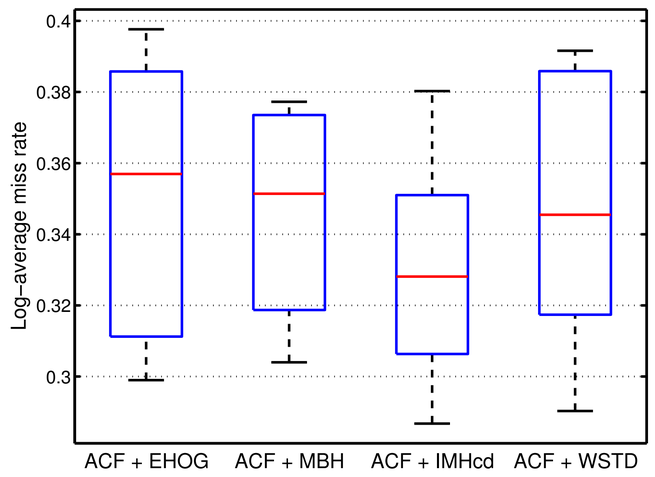

Figure 3: Box plot of log-average MR for human detectors generated by combining motion features extractors with the original ACF features.

New ACF Extension Performance Evaluation: The proposed ACF extension (ACF + IMHcd) is more accurate than the original ACF. This improvement is more clear at the end of the MR x FPPI curve (Figure 4) where the new extension reaches lower MR for a given FPPI value than the original ACF. For instance, the proposed extension presents a MR of 13% for FPPI equal to 1, while the original ACF presents 16%.

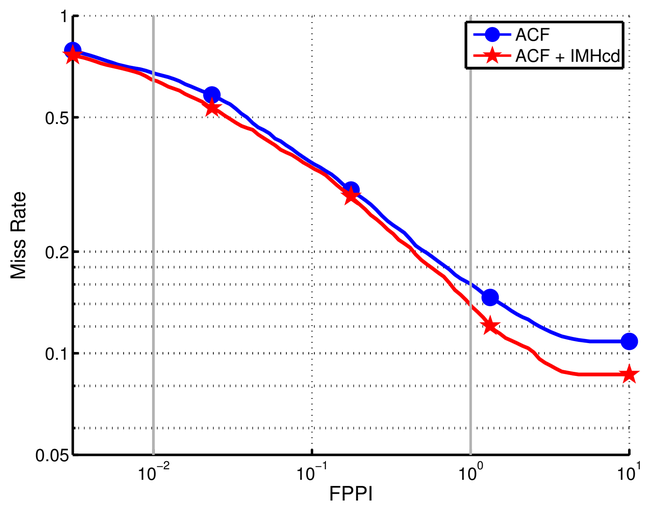

Figure 4: Curve MR x FPPI for the ACF and the ACF + IMHcd.

The proposed ACF extension can also reduce the false alarms emission. For instance, a high amount of false alarms is often emitted by the ACF in low detection threshold setting. However, we can notice the proposed ACF extension produces less false alarms than the original ACF in such setting as showed in Figure 5.

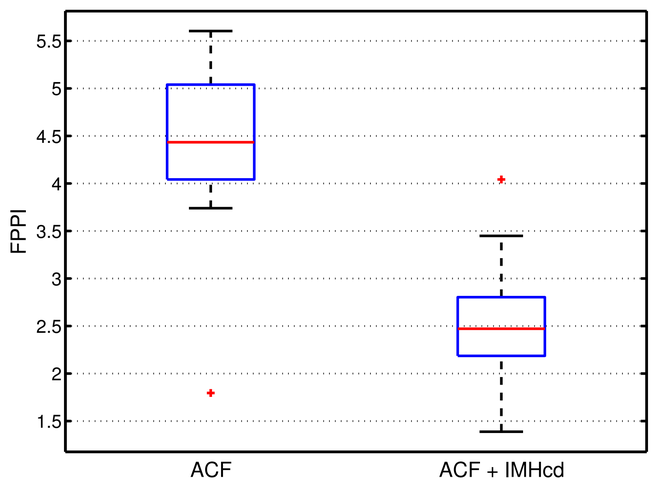

Figure 5: Box plots of FPPI values on MR reference level equal to 10 %.

4 CONCLUSIONS

- This paper extends the ACF human detector by adding motion information to its framework.
- We evaluated different motion feature extractors and conclude that the IMHcd outperforms MBH, WSTD and EHOG in the ACF framework.
- The proposed extension improves the ACF detection accuracy and mitigates its false alarms emission. These improvements are more evident in low detection threshold settings.
- Emission of false alarms is usually a critical point for human detection in most of real-world applications. Therefore, this paper presents a technique able to overcome this problem in order to make many critical application feasible.

References

[1] Ron Appel, Pietro Perona, and Serge Belongie. Fast feature pyramids for object detection. *PAMI*, 99:1, 2014.

[2] Navneet Dalal, Bill Triggs, and Cordelia Schmid. Human detection using oriented histograms of flow and appearance. In *ECCV 2006*, pages 428–441, 2006.

[3] Dennis Park, C. Lawrence Zitnick, Deva Ramanan, and Piotr Dollar. Exploring weak stabilization for motion feature extraction. In *CVPR 2013*, pages 2882–2889, 2013.