Multi-Cue Pedestrian Classification With Partial Occlusion Handling

Angela Eigenstetter

Computer Science Department, TU Darmstadt

Abstract. This paper presents a novel mixture-of-experts framework for pedestrian classification with partial occlusion handling. The framework involves a set of component-based expert classifiers trained on features derived from intensity, depth and motion. To handle partial occlusion, expert weights that are related to the degree of visibility of the associated component were computed. In experiments on extensive real-world data sets, with both partially occluded and non-occluded pedestrians, significant performance boosts were obtained over state-of-the-art approaches.

1 Introduction

The ability to visually recognize pedestrians is key for a number of application domains such as surveillance or intelligent vehicles. This task faces several difficulties like strongly varying pose and appearance as well as in case of a moving camera ever-changing backgrounds and partial occlusions. Most of the previous work focuses on the classification of pedestrians that are fully visible. However, in a real world environment significant amounts of occlusions can occur. Pedestrian classifiers designed for non-occluded pedestrians do typically not reach satisfying performance if some body parts of a pedestrian are occluded.

Component-based approaches which represent a pedestrian as an ensemble of parts, cf. [2], can only alleviate this problem to some extent without prior knowledge. The key to successful detection of partially occluded pedestrians is additional information about which body parts are occluded.

In this paper, a multi-cue component-based mixture-of-experts framework for pedestrian classification with partial occlusion handling is presented. At the core of the framework is a set of component-based expert classifiers trained on intensity, depth and motion features. Occlusions of individual body parts manifest in local depth- and motion-discontinuities. In the application phase, a segmentation algorithm is applied to extract areas of coherent depth and motion. Based on the segmentation result, occlusion-dependent weights are determined for the component-based expert classifiers to focus the combined decision on the visible parts of the pedestrian. See Figure 1.

Note that an extended version of this paper was accepted for CVPR 2010 [1].

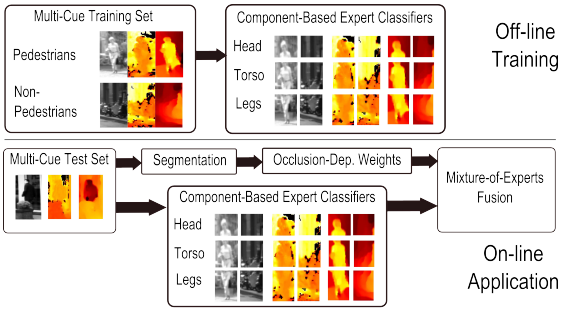

Fig. 1. Framework overview. Multi-cue component-based expert classifiers are trained off-line on features derived from intensity, depth and motion. On-line, multi-cue segmentation is applied to determine occlusion-dependent component weights for expert fusion. Data samples are shown in terms of intensity images, dense depth maps and dense optical flow (left to right).

2 Previous Work

Pedestrian classification has become an increasingly popular research topic recently. Most state-of-the art systems, cf. [3],[2],[4], derive a set of features from the available image data and apply pattern classification techniques.

Besides operating in the image intensity domain only, some authors have proposed multi-cue approaches combining information from different modalities, e.g. intensity, depth and motion [5],[6],[7]. See [2] for a current survey.

In view of detecting partially occluded pedestrians, component-based classification as suggested by [8],[9],[10],[11],[12],[13],[14],[15],[16] seems an obvious choice. Yet, only a few approaches [15],[16] explicitly incorporate a model of partial occlusion into their classification framework. However, both approaches make some restrictive assumptions.

The method of Wu and Nevatia, [16], requires a particular camera set-up, where the camera looks down on the ground-plane and consequently assumes that the head is always visible.

Wang et al., [15], use a monolithic (full-body) HOG/SVM classifier to determine occlusion maps from the responses of the underlying block-wise feature set. Based on the spatial configuration of the recovered occlusion maps, they either apply a full-body classifier or activate part-based classifiers in non-occluded regions or heuristically combine both full-body and part-based classifiers.

The main contribution of this work is a mixture-of-experts framework for pedestrian classification with partial occlusion handling. In contrast to [16], neither a particular camera set-up is required nor constant visibility of a certain body part is assumed. The suggested method is independent of the employed feature/classifier combination and the pedestrian component layout, unlike [15]. A secondary contribution involves the integration of intensity, depth and motion cues throughout the approach. Off-line, multi-cue component-based expert classifiers are trained and on-line multi-cue mean-shift segmentation is applied, see Figure 1.

3 Pedestrian Classification

Input to the framework is a training set $\mathcal{D}$ of pedestrian (ω_0) and non-pedestrian (ω_1) samples $\mathbf{x}_i \in \mathcal{D}$. Each sample $\mathbf{x}_i = [\mathbf{x}_i^i; \mathbf{x}_i^d; \mathbf{x}_i^f]$ consists of three different modalities, i.e. gray-level image intensity ($\mathbf{x}_i^i$), dense depth information via stereo vision ($\mathbf{x}_i^d$) [17] and dense optical flow ($\mathbf{x}_i^f$) [18]. $\mathbf{x}_i^d$ and $\mathbf{x}_i^f$ are treated similarly to gray-level intensity images $\mathbf{x}_i^i$, in that both depth and motion cues are represented as images, where pixel values encode distance from the camera and magnitude of optical flow vectors between two temporally aligned images, respectively. Note, that in case of optical flow only the horizontal component of flow vectors is considered and that no ego-motion compensation is applied. See Figure 5.

3.1 Component-Based Classification

The goal of pedestrian classification is to determine a class label ω_i for an unseen example $\mathbf{x}_i$. Since, a two-class problem with classes ω_0 (pedestrian) and ω_1 (non-pedestrian) is considered it is sufficient to compute the posterior probability $P(\omega_0|\mathbf{x}_i)$ that an unseen sample $\mathbf{x}_i$ is a pedestrian. The final decision then results from selecting the object class with the highest posterior probability :

$$\omega_i = \underset{\omega_j}{\operatorname{argmax}}\, P(\omega_j|\mathbf{x}_i) \tag{1}$$

The posterior probability $P(\omega_0|\mathbf{x}_i)$ is approximated using a component-based mixture-of-experts model. A sample $\mathbf{x}_i$ is composed out of K components. In the mixture-of-experts framework, [19], the final decision results from a weighted linear combination of so-called local expert classifiers which are specialized in a particular area of the feature space. With $\mathbf{F}_k(\mathbf{x}_i)$ representing a local expert classifier for the k-th component of $\mathbf{x}_i$ and $w_k(\mathbf{x}_i)$ denoting its weight, $P(\omega_0|\mathbf{x}_i)$ is approximated using:

$$P(\omega_0|\mathbf{x}_i) \approx \sum_{k=1}^{K} w_k(\mathbf{x}_i)\mathbf{F}_k(\mathbf{x}_i) \tag{2}$$

Note that the weight $w_k(\mathbf{x}_i)$ for each component expert classifier is not a fixed component prior, but depends on the sample $\mathbf{x}_i$ itself.

3.2 Multi-Cue Component Expert Classifiers

Given the component-based mixture-of-experts model, cf. Eq. (2), the component expert classifiers $\mathbf{F}_k(\mathbf{x}_i)$ are given by component-based classifiers for each cue (intensity, depth, flow) :

$$\mathbf{F}_k(\mathbf{x}_i) = \sum_{m\in(i,d,f)} \mathbf{f}_k^m(\mathbf{x}_i^m) \tag{3}$$

In this formulation, $\mathbf{f}_k^m(\mathbf{x}_i^m)$ denotes a local expert classifier for the k-th component of $\mathbf{x}_i$, which is represented in terms of the m-th cue.

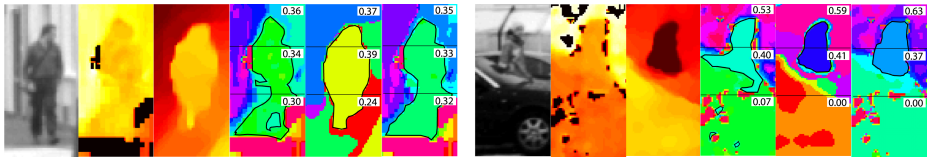

Fig. 2. Segmentation results for a non-occluded (left) and partially occluded pedestrian (right). From left to right, each sample shows: intensity image, stereo image, flow image, segmentation on stereo, segmentation on flow, combined segmentation on stereo and flow. The cluster chosen as pedestrian cluster $\boldsymbol{\phi}_{ped}$, cf. Eq. (7), is outlined in black. The computed occlusion-dependent component weights $w_k(\mathbf{x}_i)$, cf. Eq. (8), are also shown.

3.3 Occlusion-Dependent Component Weights

Weights $w_k(\mathbf{x}_i)$ introduced in Sec. 3.1 are derived from each example $\mathbf{x}_i$ to incorporate a measure of occlusion of certain pedestrian components into the model. Visibility information is extracted from each sample $\mathbf{x}_i$ by exploiting significant depth and motion discontinuities at the occlusion boundary, as shown in Figures 2 and 5.

The procedure to derive component weights $w_k(\mathbf{x}_i)$ is divided into three steps: First, a segmentation algorithm is applied, cf. [20], to the dense stereo and optical flow images of $\mathbf{x}_i$. Second, the segmented cluster which likely corresponds to the visible area of a pedestrian is selected. Third, the degree of visibility of each component given the selected cluster is estimated.

Mean-shift algorithm is chosen, [21], out of many possible choices because it provides a good balance between segmentation accuracy and processing efficiency [20]. The result of the mean-shift segmentation is a set of C clusters ϕ_c with $c = 1, \ldots, C$, as shown in Figure 2.

Let $\boldsymbol{\phi}_c$ and $\boldsymbol{\gamma}_k$ denote binary vectors defining the membership of pixel-locations of the sample $\mathbf{x}_i$ to the c-th cluster ϕ_c and k-th component γ_k, respectively. Further, a two-dimensional probability mass function $\mu_v(\mathbf{p})$ is utilized. It represents the probability that a given pixel $\mathbf{p} \in \mathbf{x}_i$ corresponds to a pedestrian, solely based on its location within $\mathbf{x}_i$.

To increase specificity, view-dependent probability masks $\mu_v(\mathbf{p})$ in terms of separate masks for front/back, left and right views are used. See Figure 3(a). Again, a vectorized representation of μ_v is denoted as $\boldsymbol{\mu}_v$.

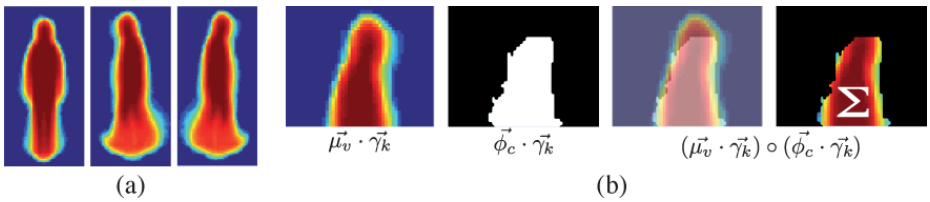

Fig. 3. (a) Probability masks for front/back, left and right view. The values of the probability masks are in the range of zero (dark blue) to one (dark red). (b) Visualization of the correlation-based similarity measure $\Psi_{in}(\boldsymbol{\phi}_c, \boldsymbol{\gamma}_k, \boldsymbol{\mu}_v)$ for the head component, see text.

To select the segmented cluster, which corresponds to the visible area of a pedestrian, a correlation-based similarity measure is used:

$$\Psi(\boldsymbol{\phi}_c, \boldsymbol{\gamma}_k, \boldsymbol{\mu}_v) = \Psi_{in}(\boldsymbol{\phi}_c, \boldsymbol{\gamma}_k, \boldsymbol{\mu}_v) + \Psi_{out}(\boldsymbol{\phi}_c, \boldsymbol{\gamma}_k, \boldsymbol{\mu}_v) \tag{4}$$

The first measure $\Psi_{in}(\boldsymbol{\phi}_c, \boldsymbol{\gamma}_k, \boldsymbol{\mu}_v)$ is designed to evaluate how well a cluster ϕ_c matches typical pedestrian geometry, represented by a view-dependent pedestrian probability mask μ_v, in a certain component γ_k. To compute $\Psi_{in}(\boldsymbol{\phi}_c, \boldsymbol{\gamma}_k, \boldsymbol{\mu}_v)$, the cluster $\boldsymbol{\phi}_c$ and the probability mask $\boldsymbol{\mu}_v$ are correlated and normalized within the component given by $\boldsymbol{\gamma}_k$:

$$\Psi_{in}(\boldsymbol{\phi}_c, \boldsymbol{\gamma}_k, \boldsymbol{\mu}_v) = \frac{(\boldsymbol{\mu}_v \cdot \boldsymbol{\gamma}_k) \circ (\boldsymbol{\phi}_c \cdot \boldsymbol{\gamma}_k)}{\boldsymbol{\mu}_v \circ \boldsymbol{\gamma}_k} \tag{5}$$

Here, $\cdot$ denotes point-wise multiplication of vectors, while $\circ$ denotes a dot product. Note that the main purpose of $\boldsymbol{\gamma}_k$ in this formulation is to restrict computation to a local body component γ_k. See Figure 3(b).

The second measure $\Psi_{out}(\boldsymbol{\phi}_c, \boldsymbol{\gamma}_k, \boldsymbol{\mu}_v)$ penalizes clusters which extend too far beyond a typical pedestrian shape. For that a similar correlation is performed using an "inverse" probability mask $\boldsymbol{\nu}_v = 1 - \boldsymbol{\mu}_v$:

$$\Psi_{out}(\boldsymbol{\phi}_c, \boldsymbol{\gamma}_k, \boldsymbol{\mu}_v) = 1 - \frac{(\boldsymbol{\nu}_v \cdot \boldsymbol{\gamma}_k) \circ (\boldsymbol{\phi}_c \cdot \boldsymbol{\gamma}_k)}{\boldsymbol{\nu}_v \circ \boldsymbol{\gamma}_k} \tag{6}$$

The cluster similarity measure $\Psi(\boldsymbol{\phi}_c, \boldsymbol{\gamma}_k, \boldsymbol{\mu}_v)$, see Eq. (4), is computed per cluster, component and view-dependent probability mask. To choose the cluster $\boldsymbol{\phi}_{ped}$ which most likely corresponds to visible parts of the pedestrian, a maximum operation is applied over components and views:

$$\boldsymbol{\phi}_{ped} = \underset{\boldsymbol{\phi}_c}{\operatorname{argmax}} \left(\max_{\boldsymbol{\gamma}_k \boldsymbol{\mu}_v} (\Psi(\boldsymbol{\phi}_c, \boldsymbol{\gamma}_k, \boldsymbol{\mu}_v)) \right) \tag{7}$$

Note, that only single clusters and pairs of clusters are merged together as possible candidates.

Once the cluster $\boldsymbol{\phi}_{ped}$, corresponding to visible parts of the pedestrian, is selected, the degree of visibility of each component is approximated. For each component $\boldsymbol{\gamma}_k$, the spatial extent of $\boldsymbol{\phi}_{ped}$ is related against clusters corresponding to occluding objects. The set of all clusters $\boldsymbol{\phi}_j$, which are possible occluders of $\boldsymbol{\phi}_{ped}$, is denoted by $\boldsymbol{\Upsilon}$. Possible occluders of $\boldsymbol{\phi}_{ped}$ are clusters which are closer to the camera than $\boldsymbol{\phi}_{ped}$. With $n(\boldsymbol{v})$ denoting the number of non-zero elements in an arbitrary vector $\boldsymbol{v}$, occlusion-dependent component weights $w_k(\mathbf{x}_i)$, with $\sum_k w_k(\mathbf{x}_i) = 1$, are then given by:

$$w_k(\mathbf{x}_i) \propto \frac{n(\boldsymbol{\phi}_{ped} \cdot \boldsymbol{\gamma}_k)}{\sum_{\boldsymbol{\phi}_j \in \boldsymbol{\Upsilon}} (n(\boldsymbol{\phi}_j \cdot \boldsymbol{\gamma}_k)) + n(\boldsymbol{\phi}_{ped} \cdot \boldsymbol{\gamma}_k)} \tag{8}$$

See Figure 2 for a visualization of the cluster $\boldsymbol{\phi}_{ped}$, corresponding to visible parts of the pedestrian, and the recovered occlusion-dependent component weights $w_k(\mathbf{x}_i)$.

Table 1. Training and test set statistics.

	Pedestrians (labeled)	Pedestrians (jittered)	Non-Pedestrians
Train Set	6514	52112	32465
Partially Occluded Test Set	620	11160	16235
Non-Occluded Test Set	3201	25608	16235

4 Experiments

4.1 Experimental Setup

The proposed multi-cue component-based mixture-of-experts framework was tested in experiments on pedestrian classification.

The training and test samples consist of manually labeled pedestrian and non-pedestrian bounding boxes in images captured from a vehicle-mounted calibrated stereo camera rig in an urban environment. For each manually labeled pedestrian, additional samples are created by geometric jittering.

Dense stereo is computed using the semi-global matching algorithm [17]. To compute dense optical flow, the method of [18] is used.

Training and test samples have a resolution of 36×84 pixels with a 6-pixel border around the pedestrians. In the experiments, $K = 3$ components γ_k were used, corresponding to head/shoulder (36×24 pixels), torso (36×36 pixels) and leg (36×48 pixels) regions, see Figure 4.

Regarding features for the component/cue expert classifiers $\mathbf{f}_k^m$, see Eq. (3), histograms of oriented gradients (HOG) are chosen out of many possible feature sets, cf. [22], [3], [2], [23]. The motivation for this choice is two-fold: First, HOG features are still among the best performing feature sets available; second, the framework is compared to the approach of Wang et al. [15] which explicitly requires and operates on the block-wise structure of HOG features. Linear support vector machines (SVMs) are employed for classification. In the implementation of [15], the occlusion handling of Wang et al. is used together with the same component layout (head, torso, legs), features (HOG) and classifiers (linear SVMs) as in the suggested framework, but only for the intensity cue.

To train the component classifiers, only non-occluded pedestrians (and non-pedestrian samples) are used. For testing, the performance is evaluated on two different test sets : one involving non-occluded pedestrians and one consisting of partially occluded pedestrians. The non-pedestrian samples are the same for both test sets. See Table 1 and Figure 5 for an overview of the dataset.

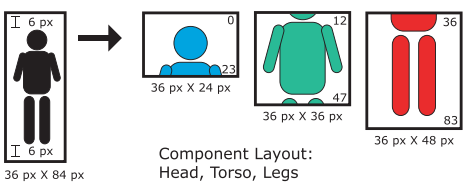

Fig. 4. Component layout as used in the experiments. Three overlapping components are used, corresponding to head, torso and leg regions, see text.

4.2 Performance on Partially Occluded Test Data

Partial Occlusion Handling The first experiment, evaluates the effect of different models of partial occlusion handling. All expert component classifiers are trained on intensity images only. Full-body HOG approach of [22] and the approach of [15] are used as baslines. The suggested framework is evaluated using four different strategies to compute occlusion-dependent component weights $w_k(\mathbf{x}_i)$ for $\mathbf{x}_i$, as defined in Sec. 3.3: weights resulting from mean-shift segmentation using depth only, flow only and a combination of both depth and flow. Additionally, uniform weights $w_k(\mathbf{x}_i)$ are considered, i.e. no segmentation. Results in terms of ROC performance are given in Figure 6(a).

All component-based approaches outperform the full-body HOG classifier (magenta *). The approach of Wang et al. [15] (cyan +) significantly improves performance over the full-body HOG classifier by a factor of two (reduction in false positives at constant detection rates). All variants of the framework in turn outperform the method of Wang et al. [15], with segmentation on combined depth and flow (green □) performing best. Compared to the use of uniform weights $w_k(\mathbf{x}_i)$ (black ×), the addition of multi-cue segmentation to compute component weights (green □) improves performance by approximately a factor of two.

Multi-Cue Classification The second experiment, evaluates the performance of multi-cue component classifiers, as presented in Sec. 3.2, compared to intensity-only component classifiers. Uniform component weights $w_k(\mathbf{x}_i)$, i.e. no segmentation, were used throughout all approaches. Results are given in Figure 6(b) (solid lines). A full-body intensity-only HOG classifier and a multi-cue full-body HOG classifier trained on intensity, stereo and flow data (dashed lines) are used as baseline. Multi-cue classification significantly improves performance both for the full-body and for the component-based approach. The best performance is

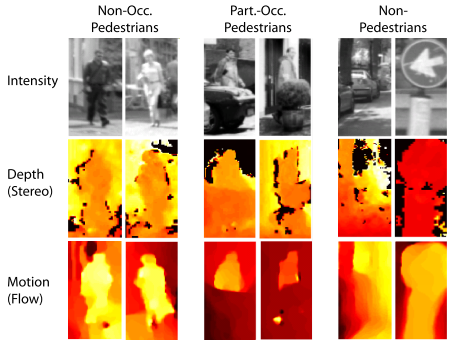

Fig. 5. Non-occluded pedestrians, partially occluded pedestrians and non-pedestrians samples in the data. In depth (stereo) images, darker colors denote closer distances. Note that the background (large depth values) has been faded out for visibility. Optical flow images depict the magnitude of the horizontal component of flow vectors, with lighter colors indicating stronger motion.

reached by the component-based approach involving intensity, stereo and flow (green □). The performance improvement over a corresponding component-based classifier using intensity-only (black ×) is up to a factor of two.

Multi-Cue Classification with Partial Occlusion Handling The next experiment, evaluates the proposed multi-cue framework involving occlusion-dependent component weights combined with multi-cue classification. The same cues were used for both segmentation and classification. Similar to the previous experiment, the baseline is given by full-body classifiers (cyan + and magenta *), as well as a component-based intensity-only classifier using uniform weights (black ×). See Figure 6(c).

The best performing system variant is the proposed component-based mixture-of-experts architecture using stereo and optical flow concurrently to determine occlusion-dependent weights $w_k(\mathbf{x}_i)$ and for multi-cue classification (green □). Compared to a corresponding multi-cue full-body classifier (magenta *), the performance boost is approximately a factor of four. A similar performance differences exists between the best approach (green □) and a component-based intensity-only classifier using uniform component weights (black ×).

4.3 Performance on Non-Occluded Test Data

In this section performance of the framework using non-occluded pedestrians is evaluated. The effect of partial occlusion handling is evaluated independently from the use of multiple cues for classification.

Figure 6(d) shows the effect of different models of partial occlusion handling combined with intensity-only component-based classifiers. The full-body HOG classifier (magenta *), as well as the approach of Wang et al. [15] (cyan +), serve as baselines. The best performance is reached by the full-body HOG classifier. All component-based approaches perform slightly worse. Of all component-based approaches, uniform component weights $w_k(\mathbf{x}_i)$, i.e. no occlusion handling, yields the best performance by a small margin. On non-occluded test samples, the best suggested approach with occlusion handling (green □) gives the same performance as Wang et al. [15] (cyan +).

Multi-cue classification, as shown in Figure 6(e), yields similar performance boosts compared to intensity-only classification as observed for the test on partially occluded data, cf. Sec. 4.2. Figure 6(f) depicts results of the integrated multi-cue mixture-of-experts framework with partial occlusion handling. Compared to a full-body classifier involving intensity, stereo and flow (magenta *), the best performing mixture-of-experts approach gives only slightly worse performance, particularly at low false positive rates. In relation to intensity-only full-body classification (cyan +), i.e. the approach of [22], the multi-cue framework improves performance by up to a factor of two.

5 Conclusion

This paper presented a multi-cue mixture-of-experts framework for component-based pedestrian classification with partial occlusion handling. For the partially

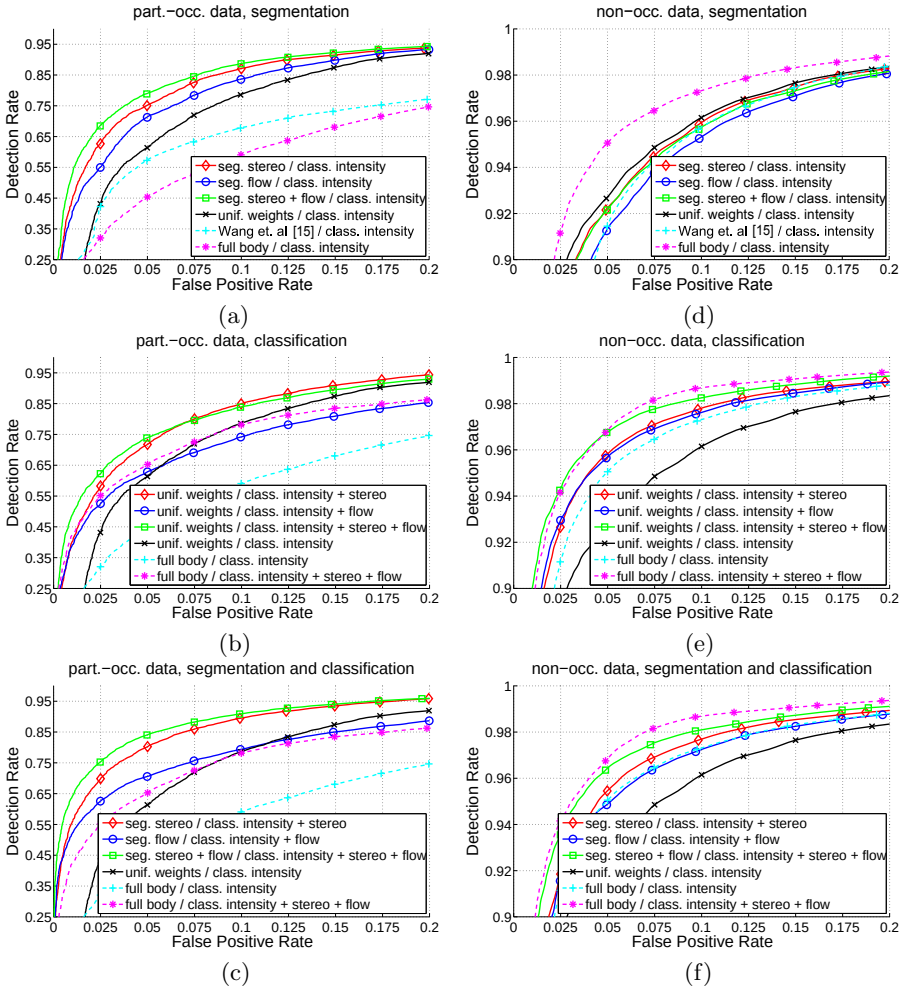

Fig. 6. Left column shows evaluation on partially occluded test set performing (a) partial occlusion handling (b) multi-cue classification in comparison to intensity-only classification (c) combined multi-cue partial occlusion handling and classification. Right column shows evaluation on non-occluded test set performing (d) partial occlusion handling strategies (e) multi-cue classification in comparison to intensity-only classification (c) combined multi-cue partial occlusion handling and classification.

occluded dataset, an improvement of more than a factor of two versus the baseline (component-based, no occlusion handling) and state-of-the-art [15] was obtained in the case of depth- and motion-based occlusion handling. In the case of multi-cue (intensity, depth, motion) classification an additional improvement of a factor of two was obtained versus the baseline (intensity only). The full-body classifiers performed worse than the beforementioned baselines. For the non-occluded dataset, occlusion handling does not appreciably deteriorate results, while multi-cue classification improves performance by a factor of two.

References

1. Enzweiler, M., Eigenstetter, A., Gavrila, D.M., Schiele, B.: Multi-cue pedestrian classification with partial occlusion handling. Proc. CVPR(2010)
2. Enzweiler, M., Gavrila, D.M.: Monocular pedestrian detection: Survey and experiments. IEEE PAMI**31**(12) (2009) 2179–2195
3. Dollar, P., Wojek, C., Schiele, B., Perona, P.: Pedestrian detection: A benchmark. Proc. CVPR(2009)
4. Hussein, M., Porikli, F., Davis, L.: A comprehensive evaluation framework and a comparative study for human detectors. IEEE ITS**10**(3) (2009) 417–427
5. Ess, A., Leibe, B., van Gool, L.: Depth and appearance for mobile scene analysis. In: Proc. ICCV. (2007)
6. Gavrila, D.M., Munder, S.: Multi-cue pedestrian detection and tracking from a moving vehicle. IJCV**73**(1) (2007) 41–59
7. Wojek, C., Walk, S., Schiele, B.: Multi-cue onboard pedestrian detection. In: Proc. CVPR. (2009)
8. Dollar et al., P.: Multiple component learning for object detection. Proc. ECCV(2008) 211–224
9. Felzenszwalb, P.F., Huttenlocher, D.P.: Pictorial structures for object recognition. IJCV**61**(1) (2005) 55–79
10. Leibe, B., Seemann, E., Schiele, B.: Pedestrian detection in crowded scenes. In: Proc. CVPR. (2005) 878–885
11. Micilotta, A.S., Ong, E.J., Bowden, R.: Detection and tracking of humans by probabilistic body part assembly. In: Proc. BMVC. (2005) 429–438
12. Mikolajczyk, K., Schmid, C., Zisserman, A.: Human detection based on a probabilistic assembly of robust part detectors. In: Proc. ECCV. (2004) 69–81
13. Mohan, A., Papageorgiou, C., Poggio, T.: Example-based object detection in images by components. IEEE PAMI**23**(4) (2001) 349–361
14. Seemann, E., Fritz, M., Schiele, B.: Towards robust pedestrian detection in crowded image sequences. In: Proc. CVPR. (2007)
15. Wang, X., Han, T., Yan, S.: A HOG-LBP human detector with partial occlusion handling. Proc. ICCV(2009)
16. Wu, B., Nevatia, R.: Detection and tracking of multiple, partially occluded humans by Bayesian combination of edgelet based part detectors. IJCV**75**(2) (2007) 247 – 266
17. Hirschmüller, H.: Stereo processing by semi-global matching and mutual information. IEEE PAMI**30**(2) (2008) 328–341
18. Wedel, A., Cremers, D., Pock, T., Bischof, H.: Structure- and motion-adaptive regularization for high accuracy optic flow. Proc. ICCV(2009)
19. Jacobs, R., Jordan, M.I., Nowlan, S.J., Hinton, G.E.: Adaptive mixtures of local experts. Neural Computation**3**(1) (1991) 79–87
20. Estrada, F.J., Jepson, A.D.: Benchmarking image segmentation algorithms. IJCV**85**(2) (2009) 167–181
21. Comaniciu, D., Meer, P.: Mean shift: A robust approach toward feature space analysis. IEEE PAMI**24**(5) (2002) 603–619
22. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: Proc. CVPR. (2005) 886–893
23. Munder, S., Gavrila, D.M.: An experimental study on pedestrian classification. IEEE PAMI**28**(11) (2006) 1863–1868