2016 IEEE Intelligent Vehicles Symposium (IV)
Gothenburg, Sweden, June 19-22, 2016

Parts Selective DPM for detection of pedestrians possessing an umbrella

Yuto Shimbo[1], Yasutomo Kawanishi[1], Daisuke Deguchi[2], Ichiro Ide[1], Hiroshi Murase[1]

***Abstract*— In recent years, pedestrian detection from an in-vehicle camera has been attracting attention. However, in the case of a raining situation, the detection accuracy decreases because the head of a pedestrian tends to be occluded by an umbrella. In oder to handle such cases, in this paper, as a variation of the Deformable Part Model (DPM) which is widely used in the field of object recognition, we propose "Parts Selective DPM (PS-DPM)" which selectively chooses the original part filters and additional part filters trained independently. In the detection of pedestrians possessing an umbrella, the selection of head and umbrella parts will make pedestrian detection more robust to the occlusion. We conducted experiments to evaluate the performance of the proposed method. As a result, pedestrian detection with the proposed PS-DPM achieved high detection accuracy in rainy weather, compared with the detection by the conventional DPM. Moreover, we confirmed that it did not decrease the pedestrian detection accuracy in fine weather.**

I. INTRODUCTION

In recent years, technology for autonomous vehicles is being developed. Recognition of the surrounding environment of a vehicle is becoming one of the most indispensable functions for developing an autonomous vehicle, where pedestrian detection plays an important role to avoid serious accidents. Rainy weather is a critical and difficult condition since the braking distance of a vehicle increases due to the wet road surface. In addition, it is sometimes difficult for pedestrians themselves to notice an approaching vehicle because their field of view is narrowed by an umbrella. Therefore, pedestrian detection in rainy weather is more important than that in fine weather. However, most researches focus on the pedestrian detection in fine weather, and rainy weather has not been considered much. As seen in Fig. 1, pedestrians will have different appearances in fine weather and in rainy weather. As shown in Fig. 1(b), since an umbrella occludes the head of a pedestrian, the detection becomes difficult. To overcome this problem, the proposed method in this paper tries to improve the accuracy of pedestrian detection in rainy weather by considering differences of pedestrian appearance due to personal possessions, especially an umbrella.

Recently, various object detection methods have been developed, and most of them make use of the statistical learning framework [1]–[9], [17]. Among them, the Deformable Part Model (DPM) is widely used since it can deal with various poses of the detection target and its code is publicly available [13]. However, a pedestrian in rainy weather is often occluded by an umbrella. Therefore, if detection is performed by DPM trained with pedestrians who do not possess an umbrella, the number of oversighted pedestrians will increase in rainy weather in comparison with fine weather. Some research groups have proposed methods that train the model of the detection target by training samples including occlusion. For example, Tang et al. proposed a double person detector that detects two persons at the same time by learning the overlap between people by DPM [11]. This method achieved accurate detection in crowded situations that conventional detectors could not cope with. However, the performance of this method degrades when it is applied for individual detection [15]. Pepik et al. proposed a method that can describe the relationship between "a hiding object (occluder)" and "a hidden object (occludee)" [15]. This method shows that accurate detection can be obtained by considering the occlusion which occurs frequently in certain circumstances. In this method, clustering is applied for training samples to the obtained occlusion patterns. Then, the model representing each occlusion pattern is generated. Although this method can represent various occlusion patterns, a huge number of object models needs to be constructed for representing various occlusion patterns.

(a) In fine weather (b) In rainy weather

Fig. 1. Appearances of pedestrians in fine weather and in rainy weather.

Therefore, this paper proposes a method to improve the detection accuracy by finding the optimal combination of parts that can deal with various occlusion patterns.

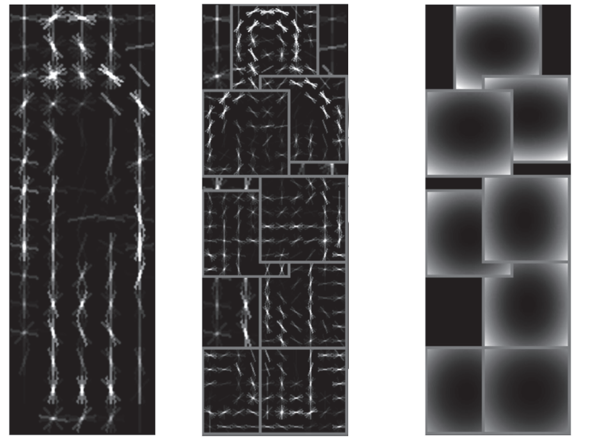
The contributions of this paper are as follows:

- Introduction of a parts selection mechanism into the DPM model structure to cope with parts-wise occlusion.
- Proposal of a flexible model structure that can add new parts trained independently.

In Section II, the idea of "Parts Selective DPM (PS-DPM)" is introduced. Then, Section III explains the detailed implementation of detection of pedestrians possessing an

[1]Graduate School of Information Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya-shi, Aichi, Japan shimboy@murase.m.is.nagoya-u.ac.jp, {kawanishi, ide, murase}@is.nagoya-u.ac.jp
[2]Information Strategy Office, Nagoya University, Furo-cho, Chikusa-ku, Nagoya-shi, Aichi, Japan ddeguchi@nagoya-u.jp

978-1-5090-1820-8/16/$31.00 ©2016 IEEE



(a) Root filter (b) Part filters (c) Deformation costs

Fig. 2. Model structure of DPM.

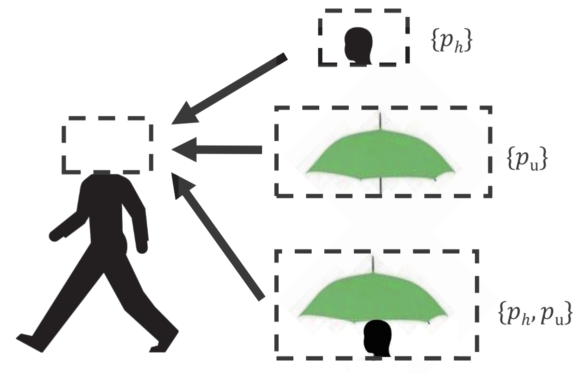


Fig. 3. Concept of the proposed Parts Selective DPM applied to a pedestrian.

umbrella using PS-DPM. Section IV gives experimental results and discussions. Finally, this paper is concluded in Section V.

II. Overview of Parts Selective DPM

The proposed Parts Selective DPM (PS-DPM) is designed as a model that can treat the detection of non-occluded target and occluded target in the same framework by finding the optimal combination of parts adaptively to the presence or absence of occluders. In this section, first, the object detection with the conventional DPM [1] is described. Next, we introduce the proposed PS-DPM. Then, the application of PS-DPM for pedestrian detection in rainy weather is described.

A. Object Detection with DPM

Figure 2 shows an example of the model structure of the conventional DPM. DPM consists of a set of filters $\mathcal{P} = \{p_0, p_1, ..., p_n\}$, where p_0 is a root filter representing the rough shape of an object, and p_i $(i = 1, ..., n)$ are part filters representing the local structures of the object. Each filter p_j $(j = 0, ..., n)$ is composed of a function ϕ_j that extracts a feature vector from the local image patch and a weight vector $\boldsymbol{F}_j$ for controlling the importance of the feature vector. Based on these notations, p_j is represented as $p_j = (\phi_j, \boldsymbol{F}_j)$. In addition, a part filter is composed of a part position $\boldsymbol{v}_i$ relative to the root filter and a deformation cost $\boldsymbol{d}_i$. Finally, the score of DPM at position $\boldsymbol{x}_0 = (x_0, y_0)$ is calculated by

$$\begin{aligned}\alpha(\mathcal{P}, \boldsymbol{x}_0) &= \boldsymbol{F}_0 \cdot \phi_0(\boldsymbol{x}_0) \\ &+ \sum_{p_i \in \mathcal{P} \setminus \{p_0\}} \max_{\Delta \boldsymbol{x}} \left\{ \boldsymbol{F}_i \cdot \phi_i(\boldsymbol{x}_0 + \boldsymbol{v}_i + \Delta \boldsymbol{x}) - \boldsymbol{d}_i \cdot \theta_d(\Delta \boldsymbol{x}) \right\} + b,\end{aligned} \quad (1)$$

where the first term is the response of the root filter, and the second term is the sum of the responses of part filters subtracted by the deformation cost. Here, " · " represents the inner product, b, a bias term, $\Delta \boldsymbol{x} = (\Delta x, \Delta y)$, and the deformation from the trained part position $\boldsymbol{v}_i$. Finally, the response of each part filter is obtained by finding $\Delta \boldsymbol{x}$ where the score takes the maximum value. $\theta_d(\Delta \boldsymbol{x})$ is represented as

$$\theta_d(\Delta \boldsymbol{x}) = (\Delta x, \Delta y, (\Delta x)^2, (\Delta y)^2)^T. \quad (2)$$

The deformation cost $\boldsymbol{d}_i$ is a weight vector for $\theta_d(\Delta \boldsymbol{x})$. According to $\boldsymbol{d}_i$ and the distance from $\boldsymbol{v}_i$, the detection score is penalized.

B. Parts Selective DPM

In Eq. (1), the score of DPM is calculated by summing the responses of all filters. However, when a detection target is partially occluded, poor responses are obtained from part filters corresponding to the occluded part. To avoid this problem, the proposed method selectively uses part filters that can represent the model appropriately, and tries to remove the effect from poor part filters. In the proposed method, part filters of $\mathcal{P}$ are divided into two groups as

$$\mathcal{P}_f = \{p_0, p_{f_1}, ..., p_{f_k}\},\ \mathcal{P}_s = \{p_{s_1}, ..., p_{s_l}\},$$
$$(|\mathcal{P}| = |\mathcal{P}_f| + |\mathcal{P}_s| = k + l + 1),$$

where $\mathcal{P}_f$ is a set of part filters that must be used for the score calculation, and $\mathcal{P}_s$ is a set of part filters that can be selectively used. Here, the root filter p_0 is always included in $\mathcal{P}_f$. Part filters can be arbitrary split into the two groups by the user. To introduce the proposed parts selection mechanism, Eq. (1) is rewritten as

$$\alpha'(\mathcal{P}, \boldsymbol{x}_0) = \alpha(\mathcal{P}_f, \boldsymbol{x}_0) + \max_{p' \in \mathfrak{P}(\mathcal{P}_s)} \gamma(p', \boldsymbol{x}_0), \quad (3)$$

$$\gamma(p', \boldsymbol{x}_0) = \sum_{p_i \in p'} \max_{\Delta \boldsymbol{x}} \{ \boldsymbol{F}_i \cdot \phi_i(\boldsymbol{x}_0 + \boldsymbol{v}_i + \Delta \boldsymbol{x}) - \boldsymbol{d}_i \cdot \theta_d(\Delta \boldsymbol{x}) \}, \quad (4)$$

where $\alpha(\mathcal{P}_f, \boldsymbol{x}_0)$ is calculated by Eq. (1), and $\max_{p' \in \mathfrak{P}(\mathcal{P}_s)} \gamma(p', \boldsymbol{x}_0)$ is the term for selecting the highest scoring combination of part filters from $\mathfrak{P}(\mathcal{P}_s)$. Note that $\mathfrak{P}(\mathcal{P}_s)$ is the power set of $\mathcal{P}_s$ excluding the empty set. For

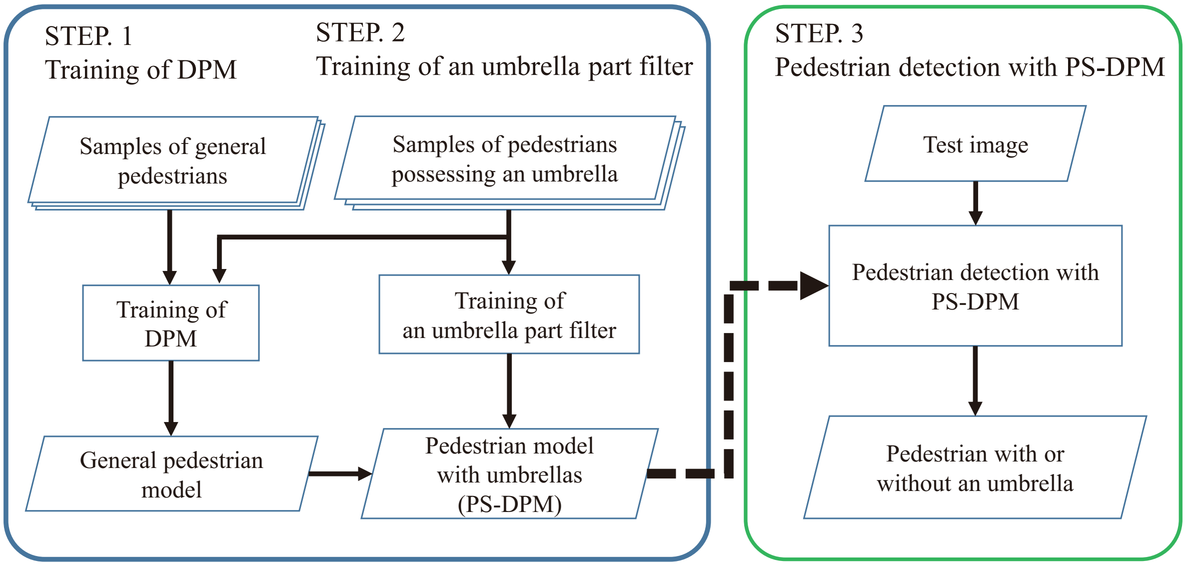


Fig. 4. Procedure of the proposed method.

instance, if $\mathcal{P}_s$ is a set consisting of three parts ($l = 3$), $\mathfrak{P}(\mathcal{P}_s)$ can be written as

$$\mathfrak{P}(\mathcal{P}_s) = \{\{p_{s_1}\}, \{p_{s_2}\}, \{p_{s_3}\}, \{p_{s_1}, p_{s_2}\}, \{p_{s_1}, p_{s_3}\}, \{p_{s_2}, p_{s_3}\}, \{p_{s_1}, p_{s_2}, p_{s_3}\}\}.$$

In the case of pedestrian detection in rainy weather, since pedestrians will possess an umbrella, their heads will have a high chance to be occluded by the umbrella. This will result in poor responses from part filters corresponding to the head. Therefore, the proposed method constructs an additional part filter p_u representing the umbrella, and puts it into the set of selectable parts set $\mathcal{P}_s$. Thus, $\mathcal{P}_s = \{p_h, p_u\}$. Parts selection in the proposed method is performed only for the head part filter p_h and the umbrella part filter p_u. Therefore, the proposed method calculates responses of p_h and p_u, and selects the best scoring combination of parts among $\mathfrak{P}(\mathcal{P}_s) = \{\{p_h\}, \{p_u\}, \{p_h, p_u\}\}$. All parts without the head and the umbrella are always used. The concept of the pedestrian model in the proposed method is shown in Fig. 3.

III. Application of the PS-DPM for Detecting Pedestrians Possessing an Umbrella

In this section, we will describe the application of the proposed PS-DPM for detecting pedestrians possessing an umbrella. First, a general pedestrian model is constructed by the conventional DPM. Next, a part filter that represents umbrellas is trained by training samples of pedestrian possessing an umbrella. Then, the umbrella part filter is added to the model. Finally, detection is performed on test images by using the umbrella filter and the head part filter, selectively. The procedure of the proposed method is shown in Fig. 4.

Fig. 5. Examples of the annotation for an additional part filter.

A. Training of a general pedestrian model

In this step, a general pedestrian model is trained by the conventional DPM framework [1], [12]. Note that this model is trained by all training samples no matter whether a pedestrian possesses an umbrella or not.

B. Training of a umbrella part filter

In this step, an umbrella part filter is trained and added to the general pedestrian model. To add a part filter, it is necessary to learn the following parameters: filter size (w, h), weight vector $\boldsymbol{F}_u$, part location $\boldsymbol{v}_u$, and deformation cost $\boldsymbol{d}_u$. By annotating the part location for the training samples, the shape of the additional part can be explicitly trained. Figure 5 shows examples of the training samples of pedestrians possessing an umbrella. The rectangle with a solid line represents the detection target (pedestrian), and that with a dashed line represents the additional part (umbrella, in this case).

The filter size (w, h) of the additional part filter is calculated by taking the average of all the annotated locations. Then, all the training samples are resized to the above filter

TABLE I

COMPARISON OF DATASETS

Dataset	Weather	Number of images	Image size [pixels]	Number of pedestrians	Number of pedestrians possessing an umbrella
Training dataset	Rainy	1,673	960 × 720	1,857	1,763 (95%)
Test dataset	Rainy	950	640 × 480	1,280	845 (66%)
Daimler dataset [16]	Sunny	7,656	640 × 480	2,452	0 (0%)

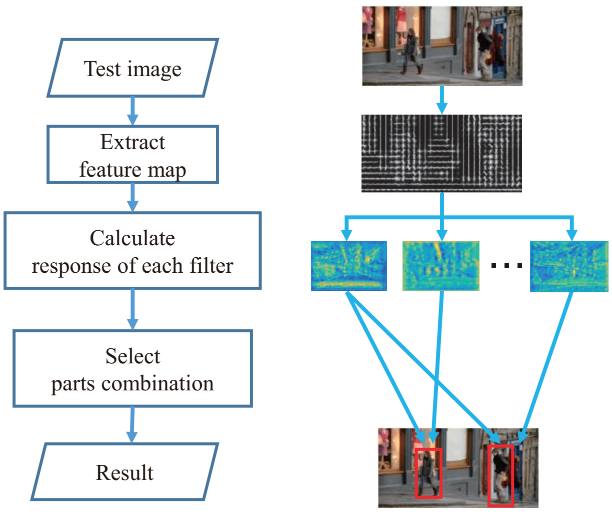


Fig. 6. Procedure of pedestrian detection with the proposed Parts Selective DPM.

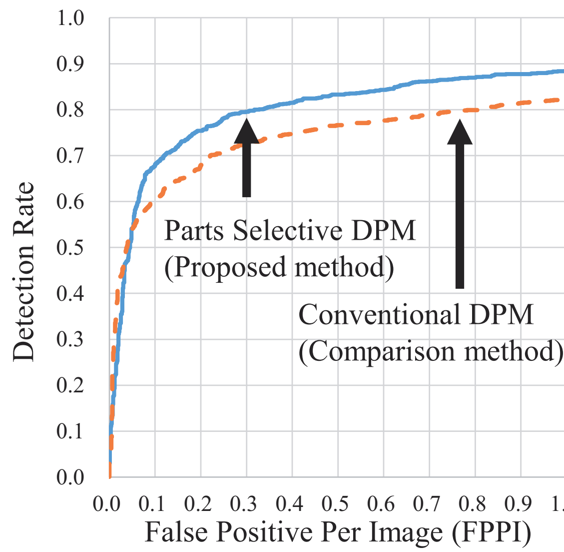


Fig. 7. Detection accuracy (Test dataset)

size, and the weight vector $\boldsymbol{F}_u$ is obtained through the training of linear SVM. The part location $\boldsymbol{v}_u = (\bar{x}, \bar{y})$ is calculated as the average of the umbrella positions (x, y) relative to the pedestrian in each training sample. As the deformation cost $\boldsymbol{d}_u$ for $\theta_d(\Delta\boldsymbol{x})$, the proposed method uses the variances σ_x, σ_y of the additional part position (x, y). In detail, $\boldsymbol{d}_u \cdot \theta_d(\Delta\boldsymbol{x})$ for the additional part filter is calculated by $N(\bar{x}, \sigma_x)N(\bar{y}, \sigma_y)$.

C. Pedestrian detection with PS-DPM

Figure 6 shows the detection flow of the proposed method. First, as in the conventional DPM, the proposed method computes HOG feature maps from a test image. Next, responses of all filters are calculated using feature maps. Finally, the score of each detection window is calculated by Eq. (3) which selects the optimal combination of part filters from $\mathfrak{P}(\mathcal{P}'_s) = \{\{p_h\}, \{p_u\}, \{p_h, p_u\}\}$.

IV. EXPERIMENTAL EVALUATIONS

To confirm the effectiveness of the proposed method for the detection of pedestrians possessing an umbrella, we conducted experiments using in-vehicle camera images taken in rainy weather as the test dataset. In addition, to confirm the effectiveness of the proposed method for the detection of general pedestrians, the proposed method was evaluated by using in-vehicle camera images taken in fine weather, too. In these experiments, Person Grammar Model [12] was used as the model structure of DPM. The proposed PS-DPM added the umbrella part filter p_u into this model. Here, umbrella part filter p_u was trained in three aspect ratios. As a comparison method, Person Grammar Model trained with the training samples of person class in the PASCAL VOC 2007 Dataset [10] and the training dataset described in Section III-A was prepared.

A. Datasets

In these experiments, we prepared two in-vehicle camera image sequences captured in rainy weather. One image sequence was used as the dataset for training the umbrella part filter, and the other one was used as the test dataset. As for the evaluation against fine weather, the proposed and the comparison methods were evaluated by Daimler Mono Pedestrian Detection Benchmark Dataset [16]. Table I summarizes the details of these datasets.

B. Results and Discussions

A detected bounding box b_d was considered as positive when it overlapped with the corresponding ground truth b_g. We evaluated each detected bounding box by the following equation:

$$\frac{|b_d \cap b_g|}{|b_d \cup b_g|} \geq 0.5, \qquad (5)$$

where $|\cdot|$ represents the number of pixels in an area.

Free-response Receiver Operating Characteristic (FROC) curves were drawn for the evaluation. Figure 7 shows FROC curves of the proposed and the comparison methods by changing the threshold of the detection score. Here, x and y axes of this graph are detection rate and False Positives Per Image (FPPI), respectively. Since in this graph, higher detection accuracy is obtained when the curve becomes closer to the upper left corner, we can see that the proposed method showed superiority over the comparison method. Figure 8 shows examples of detection results (FPPI $= 0.1$) by the proposed and the comparative methods. In these figures, the red rectangles indicate detected bounding boxes of pedestrians, and the blue rectangles indicate part filters selected for the detection. As seen here, the proposed method could detect pedestrians possessing an umbrella correctly even if their heads were occluded by an umbrella. From these results, we confirmed that the adaptive parts selection mechanism of the proposed PS-DPM contributed to improve the detection accuracy of pedestrians possessing an umbrella.

Meanwhile, Fig. 9 shows the detection accuracy when applying both methods to Daimler's fine weather dataset. We can see that the two curves are almost overlapped. Figure 10 shows examples of detection results (FPPI $= 0.1$) by the proposed and the comparative methods. From these figures, we confirmed that the detection accuracy does not decrease even if the proposed method was applied to pedestrians that do not possess an umbrella.

From this result, we can say that the proposed PS-DPM is useful for improving the detection accuracy in rainy weather while keeping its performance in fine weather.

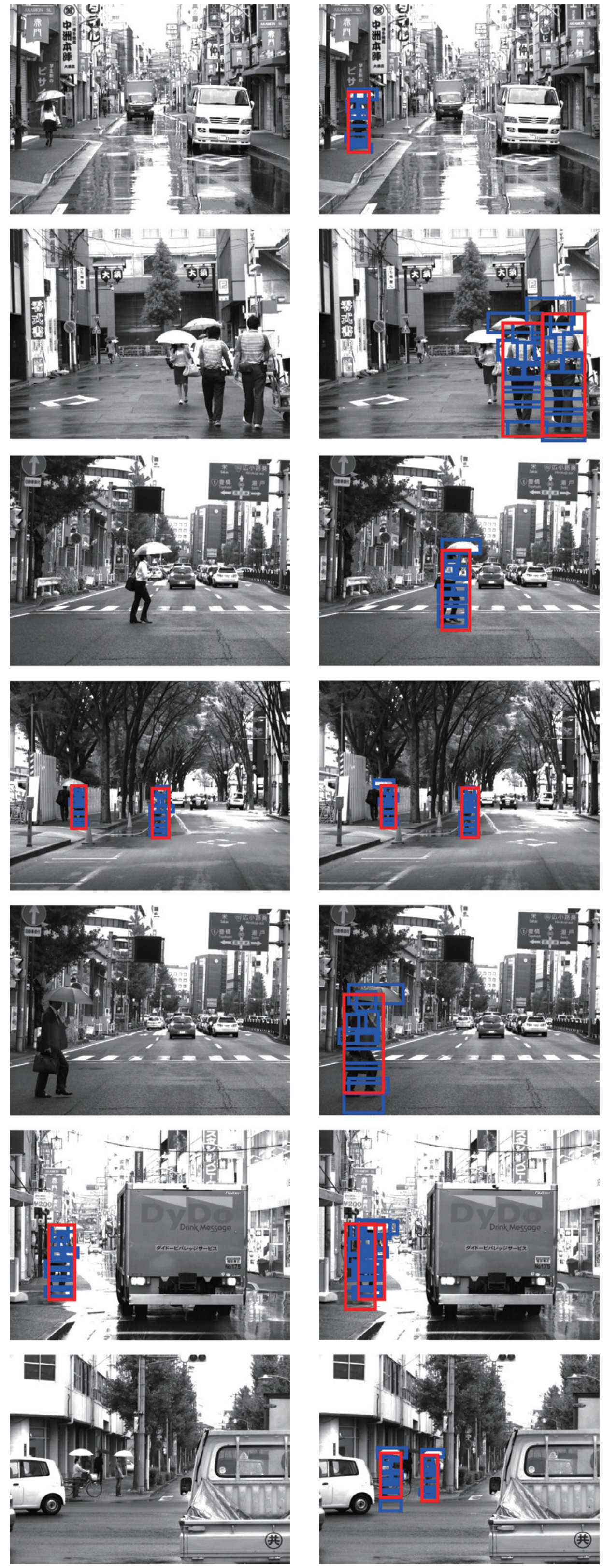
V. Conclusion

This paper proposed Parts Selective DPM (PS-DPM) for detecting pedestrians possessing an umbrella. By introducing the optimal parts selection mechanism into DPM, the proposed method could find the appropriate combination of parts for detecting pedestrians. From the experimental results, we confirmed the superiority of the proposed method over the conventional DPM, especially for detecting pedestrians in rainy weather, while keeping its performance in fine weather.

Future works include extension of the proposed PS-DPM to cope with scale change of parts, consideration of other personal possessions such as bags, experiments with larger variations of data, and real-time processing of the detection. In addition, we would like to introduce our ideas to Deep Learning based pedestrian detection techniques [8], [17].

Acknowledgment

Parts of this research were supported by JSPS Grant-in-Aid for Scientific Research.



(a) Conventional DPM (Comparison method) (b) Parts Selective DPM (Proposed method)

Fig. 8. Examples of pedestrian detection results (Test dataset).

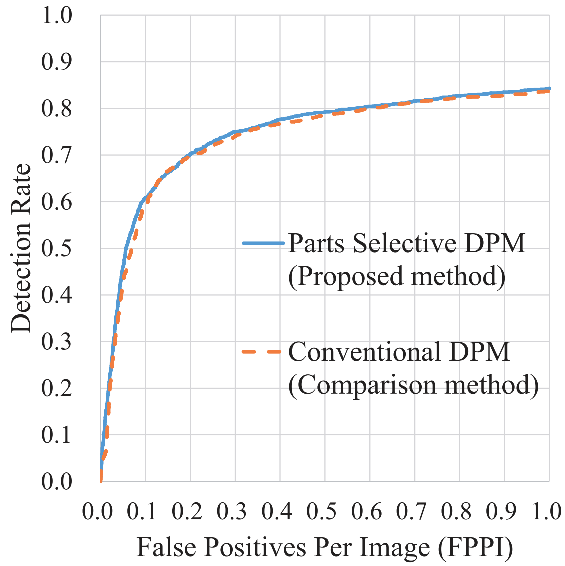
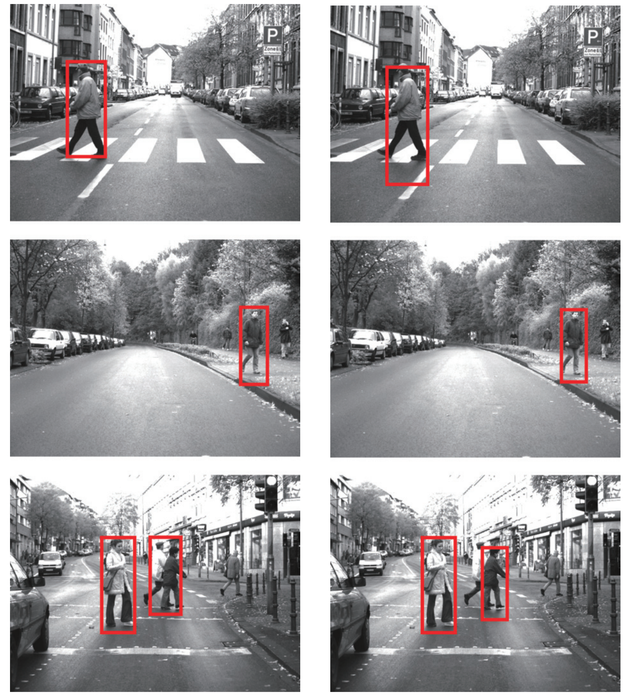


Fig. 9. Detection accuracy (Daimler dataset)



(a) Conventional DPM (Comparison method)　(b) Parts Selective DPM (Proposed method)

Fig. 10. Examples of pedestrian detection results (Daimler dataset [16]).

REFERENCES

[1] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part-based models," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.32, no.9, pp.1627–1645, Sept. 2010.

[2] R. Benenson, M. Omran, J. Hosang, and B. Schiele, "Ten years of pedestrian detection, what have we learned?" In Computer Vision—ECCV 2014 Workshops, Lecture Notes in Computer Science, vol.8626, pp.613–627, Springer, Mar. 2015.

[3] P. Viola and M. J. Jones, "Robust real-time face detection," Int. J. of Computer Vision, vol.57, no.2, pp.137–154, May 2004.

[4] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for human detection," In Proc. 2005 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, vol.1, pp.886–893, June 2005.

[5] D. Suzuo, D. Deguchi, I. Ide, H. Murase, H. Ishida, and Y. Kojima, "Environment adaptive pedestrian detection using in-vehicle camera and GPS," In Proc. 9th Int. Conf. on Computer Vision Theory and Applications, pp.354–361, Jan. 2014.

[6] C. Dubout and F. Fleuret, "Deformable Part Models with individual part scaling," In Proc. 24th British Machine Vision Conf., pp.1–10, Sept. 2013.

[7] B. Pepik, P. Gehler, M. Stark, and B. Schiele, "3D2PM — 3D Deformable Part Models," In Computer Vision—ECCV 2012, Lecture Notes in Computer Science, vol.7577, pp.356–370, Springer, Oct. 2015.

[8] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," In Proc. 2014 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, pp.580–587, June 2014.

[9] P. Dollár, R. Appel, S. Belongie, and P. Perona, "Fast feature pyramids for object detection", IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.36, no.8, pp.1532–1545, Aug. 2014.

[10] M. Everingham, L. V. Gool, C. K. Williams, J. Winn, and A. Zisserman, "The Pascal Visual Object Classes (VOC) challenge," Int. J. of Computer Vision, vol.88, no.2, pp.303–338, June 2010.

[11] S. Tang, M. Andriluka, and B. Schiele, "Detection and tracking of occluded people," Int. J. of Computer Vision, vol.110, no.1, pp.58–69, Oct. 2014.

[12] R. B. Girshck, P. Felzenszwalb, and D. McAllester, "Object detection with grammar models," In Proc. 25th Annual Conf. on Neural Information Processing Systems, Advances in Neural Information Processing Systems, vol.24, pp.442–450, Curran Associates, Inc., Dec. 2011.

[13] R. B. Girshck, P. Felzenszwalb, and D. McAllester, "Discriminatively trained Deformable Part Models, Release 5," <http://www.cs.berkeley.edu/~rbg/latent/>, Sept. 2012. [Accessed Jan. 18, 2016]

[14] C. N. J. Yu and T. Joachims, "Learning structural SVMs with latent variables," In Proc. 26th Annual ACM Int. Conf. on Machine Learning, pp.1169–1176, Sep. 2009.

[15] B. Pepik, M. Stark, P. Gehler, and B. Schiele, "Occlusion patterns for object class detection," Proc. 2013 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, pp.3286–3293, June 2013.

[16] M. Enzweiler and D. M. Gavrila, "Monocular pedestrian detection: Survey and experiments," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.31, no.12, pp.2179–2195, Dec. 2009.

[17] Y. Tian, P. Luo, X. Wang, and X. Tang, "Deep learning strong parts for pedestrian detection," In Proc. IEEE Int. Conf. on Computer Vision 2015, pp.1904–1912, Dec. 2015.