TABLE I

RECOGNITION ACCURACY COMPARISON ON TUD DATASET

	Classification criteria (%)	
	Strict	Soft
SVM (Normal)	47.58	75.40
SVM (Cost-relaxation)	50.40	78.63

TABLE II

CONFUSION MATRIX OF THE PROPOSED METHOD ON TUD DATASET

	S	SW	W	NW	N	NE	E	SE
S	0.41	0.08	0.08	0.00	0.26	0.03	0.05	0.10
SW	0.11	0.51	0.24	0.05	0.08	0.00	0.00	0.00
W	0.00	0.00	0.78	0.22	0.00	0.00	0.00	0.00
NW	0.11	0.08	0.21	0.42	0.18	0.00	0.00	0.00
N	0.17	0.05	0.00	0.05	0.61	0.02	0.02	0.07
NE	0.05	0.21	0.00	0.00	0.11	0.21	0.26	0.16
E	0.00	0.00	0.00	0.00	0.00	0.15	0.77	0.08
SE	0.12	0.00	0.00	0.08	0.04	0.04	0.44	0.28

TABLE III

CONFUSION MATRIX OF THE CONVENTIONAL METHOD ON TUD DATASET

	S	SW	W	NW	N	NE	E	SE
S	0.26	0.13	0.03	0.00	0.38	0.00	0.03	0.18
SW	0.14	0.54	0.16	0.08	0.08	0.00	0.00	0.00
W	0.00	0.09	0.57	0.30	0.00	0.00	0.00	0.04
NW	0.16	0.11	0.13	0.42	0.18	0.00	0.00	0.00
N	0.20	0.05	0.00	0.05	0.61	0.02	0.02	0.05
NE	0.11	0.16	0.00	0.00	0.16	0.37	0.16	0.05
E	0.00	0.08	0.00	0.00	0.00	0.19	0.65	0.08
SE	0.08	0.00	0.00	0.12	0.04	0.08	0.28	0.40

TABLE IV

RECOGNITION ACCURACY COMPARISON ON PDC DATASET

	Classification criteria (%)	
	Strict	Soft
SVM (Normal)	83.52	97.20
SVM (Cost-relaxation)	81.66	97.76

TABLE V

CONFUSION MATRIX OF THE PROPOSED METHOD ON PDC DATASET

	S	SW	W	NW	N	NE	E	SE
S	0.82	0.08	0.01	0.00	0.03	0.00	0.00	0.06
SW	0.08	0.66	0.23	0.00	0.01	0.01	0.00	0.00
W	0.01	0.13	0.79	0.05	0.01	0.00	0.01	0.00
NW	0.00	0.00	0.15	0.68	0.16	0.00	0.00	0.00
N	0.01	0.00	0.00	0.04	0.91	0.04	0.00	0.00
NE	0.00	0.00	0.00	0.00	0.20	0.62	0.16	0.01
E	0.00	0.00	0.01	0.00	0.00	0.04	0.90	0.05
SE	0.11	0.00	0.00	0.01	0.02	0.01	0.19	0.66

TABLE VI

CONFUSION MATRIX OF THE CONVENTIONAL METHOD ON PDC DATASET

	S	SW	W	NW	N	NE	E	SE
S	0.90	0.03	0.01	0.00	0.03	0.00	0.01	0.03
SW	0.16	0.48	0.31	0.01	0.02	0.01	0.01	0.00
W	0.01	0.09	0.83	0.04	0.01	0.00	0.01	0.00
NW	0.01	0.00	0.12	0.67	0.18	0.00	0.01	0.00
N	0.01	0.00	0.00	0.02	0.93	0.03	0.00	0.00
NE	0.00	0.00	0.00	0.00	0.20	0.64	0.14	0.00
E	0.00	0.00	0.01	0.00	0.01	0.04	0.91	0.03
SE	0.15	0.00	0.00	0.00	0.04	0.01	0.19	0.61

For both classifiers, we used Radial Basis Function (RBF) for the kernel function. We tuned the soft-margin parameter and a parameter of RBF kernel by cross validation suitable for the standard SVM. We set the cost-relaxation parameter $s = 0$ empirically for the following evaluations.

C. Feature

Various features to describe a pedestrian image have been proposed. Since we do not care about the features in the proposed method, we simply use Histograms of Oriented Gradients (HOG) feature proposed by Dalal et al. [19], which is the most popular feature for pedestrian description.

D. Evaluation Criteria

In the proposed method, since we aim to reduce fatal misclassifications by cost-relaxation of tolerable misclassifications, as evaluation criteria, we introduce not only the strict eight-orientations classification accuracy but also a "soft" classification criterion, which tolerates the misclassifications to the neighboring orientation classes. Thus, in the "soft" criterion, we do not count the misclassifications to the neighboring orientation classes as false.

The evaluation criteria are summarized as:

- Strict criterion: Strictly evaluates the misclassifications.
- Soft criterion: Tolerates the misclassifications to the neighboring orientation classes.

E. Single Dataset Evaluation Results

1) TUD dataset: First, we evaluated the methods on TUD dataset. The classification results are shown in TABLE I. We can see that the proposed method achieved higher classification accuracies in both criteria.

The confusion matrices are shown in TABLES II and III. In these tables, we can see that the proposed method reduced "fatal" misclassifications.

2) PDC dataset: Then, we evaluated the methods on PDC dataset. Since the dataset is not divided into training/testing samples, we performed five-fold cross-validation for the evaluation on this dataset.

The classification results are shown in TABLE IV. We can see that the proposed method achieved lower classification accuracy in strict criterion but higher classification accuracy in "soft" criterion.

The confusion matrices are shown in TABLES V and VI. In these tables, we can also see that the proposed method increased tolerable misclassifications but reduced "fatal" misclassifications.

F. Cross-Dataset Evaluation Results

To evaluate the generalization performance, we performed a cross-dataset evaluation. We trained the classifier using TUD dataset and evaluated on PDC dataset, and vice versa.

TABLE VII

RECOGNITION ACCURACY COMPARISON ON PDC DATASET TRAINED BY TUD DATASET

	Classification criteria (%)	
	Strict	Soft
SVM (Normal)	46.34	72.70
SVM (Cost-relaxation)	48.06	74.57

TABLE VIII

RECOGNITION ACCURACY COMPARISON ON TUD DATASET TRAINED BY PDC DATASET

	Classification criteria (%)	
	Strict	Soft
SVM (Normal)	47.58	75.40
SVM (Cost-relaxation)	50.40	78.63

The classification results are shown in TABLES VII and VIII. We can see that the proposed method outperformed the standard SVM in both criteria.

V. CONCLUSION

In this paper, we proposed a classification training method which reduces "fatal" misclassifications by cost-relaxation of "tolerable" misclassifications for pedestrian orientation classification, named misclassification tolerable learning. We introduced a new class group *conceptually similar classes* to the one-against-all classification scheme. The conceptually similar classes are defined as classes whose class labels are similar to the positive class. In the case of pedestrian orientation classification, the conceptually similar classes are defined as classes whose class labels are neighboring orientation classes to the positive orientation class. A pedestrian orientation classifier was implemented by relaxing the misclassification cost for the conceptually similar classes and tuning the classifier for separating the positive class and the non-similar classes in one-against-all SVM training. We confirmed the classification performance of the proposed method by evaluating on TUD and PDC datasets.

Future work includes extension of the proposed concept to classifiers other than SVM, and application on other classification problems.

ACKNOWLEDGEMENT

Parts of this research were supported by the Center of Innovation Program from Japan Science and Technology Agency, JST.

REFERENCES

[1] M. Enzweiler and D. M. Gavrila, "Integrated pedestrian classification and orientation estimation," in *Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2010, pp. 982–989.

[2] K. Goto, K. Kidono, Y. Kimura, and T. Naito, "Pedestrian detection and direction estimation by cascade detector with multi-classifiers utilizing feature interaction descriptor," in *Proceedings of the 2011 IEEE Intelligent Vehicles Symposium*, 2011, pp. 224–229.

[3] C. Weinrich, C. Vollmer, and H. M. Gross, "Estimation of human upper body orientation for mobile robotics using an SVM decision tree on monocular images," in *Proceedings of the 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012, pp. 2147–2152.

[4] D. Baltieri, R. Vezzani, and R. Cucchiara, "People orientation recognition by mixtures of wrapped distributions on random trees," in *Proceedings of the 12th European Conference on Computer Vision*, 2012, pp. 270–283.

[5] W. Liu, Y. Zhang, S. Tang, J. Tang, R. Hong, and J. Li, "Accurate estimation of human body orientation from RGB-D sensors," *IEEE Transactions on Cybernetics*, vol. 43, no. 5, pp. 1442–1452, 2013.

[6] J. Tao and R. Klette, "Part-based RDF for direction classification of pedestrians, and a benchmark," in *Proceedings of the 12th Asian Conference on Computer Vision Workshop*, 2014, pp. 418–432.

[7] F. Shinmura, D. Deguchi, I. Ide, H. Murase, and H. Fujiyoshi, "Estimation of human orientation using coaxial RGB-Depth images," in *Proceedings of the 2015 International Conference on Computer Vision Theory and Applications*, 2015, pp. 113–120.

[8] F. Shinmura, Y. Kawanishi, D. Deguchi, I. Ide, and H. Murase, "Pedestrian orientation classification utilizing single-chip coaxial RGB-ToF camera," in *Proceedings of the 2015 Workshop on Environment Perception for Automated On-road Vehicles*, 2015, pp. 7–11.

[9] T. Gandhi and M. M. Trivedi, "Image based estimation of pedestrian orientation for improving path prediction," in *Proceedings of the 2008 IEEE Intelligent Vehicles Symposium*, 2008, pp. 506–511.

[10] H. Shimizu and T. Poggio, "Direction estimation of pedestrian from multiple still images," in *Proceedings of the 2004 IEEE Intelligent Vehicles Symposium*, 2004, pp. 596–600.

[11] T. G. Dietterich and G. Bakiri, "Solving multiclass learning problems via error-correcting output codes," *Journal of Artificial Intelligence Research*, vol. 2, pp. 263–286, 1994.

[12] T. Hastie and R. Tibshirani, "Classification by pairwise coupling," *Annals of Statistics*, vol. 26, no. 2, pp. 451–471, 1998.

[13] J. C. Platt, N. Cristianini, and J. Shawe-Taylor, "Large margin DAGs for multiclass classification," in *Proceedings of Advances in Neural Information Processing Systems*, 2000, pp. 547–553.

[14] C.-W. Hsu and C.-J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE Transactions on Neural Networks*, vol. 13, no. 2, pp. 415–425, 2002.

[15] J. Milgram, M. Cheriet, and R. Sabourin, ""One Against One" or "One Against All": Which one is better for handwriting recognition with SVMs?" in *Proceedings of the 10th International Workshop on Frontiers in Handwriting Recognition*, 2006, pp. 1–6.

[16] M. Lapin, M. Hein, and B. Schiele, "Learning using privileged information: SVM+ and weighted SVM," *Neural Networks*, vol. 53, pp. 95–108, 2014.

[17] M. Andriluka, S. Roth, and B. Schiele, "Monocular 3D pose estimation and tracking by detection," in *Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE, 2010, pp. 623–630.

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

[19] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005, pp. 886–893.