A Boosted 3-D PCA Algorithm Based on Efficient Analysis Method

Kyung-Min Lee and Chi-Ho Lin *

School of Computer Science, Semyung University, Jecheon 27136, Korea; leegis0904@semyung.ac.kr

* Correspondence: ich410@semyung.ac.kr

Featured Application: The proposed method is useful in applications where it is not sufficient to apply a traffic safety technology of high reliability. We designed a new method because the existing method is inappropriate for traffic safety techniques due to its low reliability middle inference process in deep learning.

Abstract: In this paper, we propose a boosted 3-D PCA algorithm based on an efficient analysis method. The proposed method involves three steps that improve image detection. In the first step, the proposed method designs a new analysis method to solve the performance problem caused by data imbalance. In the second step, a parallel cross-validation structure is used to enhance the new analysis method further. We also design a modified AdaBoost algorithm to improve the detector accuracy performance of the new analysis method. In order to verify the performance of this system, we experimented with a benchmark dataset. The results show that the new analysis method is more efficient than other image detection methods.

Keywords: 3-D PCA; AdaBoost; image detection; performance; parallel

Citation: Lee, K.-M.; Lin, C.-H. A Boosted 3-D PCA Algorithm Based on Efficient Analysis Method. *Appl. Sci.* **2021**, *11*, 7609. <https://doi.org/10.3390/app11167609>

Academic Editor: Peng-Yeng Yin

Received: 31 May 2021

Accepted: 4 August 2021

Published: 19 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The rapid advancement of information and communications technology (ICT) has led to the active commercialization of the autonomous vehicle industry. Research into technologies that can detect the traffic environment is essential for establishing an autonomous driving infrastructure. One of the technologies for establishing infrastructure, the charge coupled device (CCD) camera, is limited in detecting objects in a complex traffic environment. In order to overcome this limitation, camera-based detection systems use various types of image detection technologies [1–3].

Image detection technology has trouble obtaining high accuracy performance due to environmental factors such as changes in the luminescence of the light that strikes an object, the size of the object, and the moving direction of the object. Recently, a convolutional neural network (CNN) [4] structure has been used to solve these problems. The CNN structure helps obtain a pattern to detect an image. It can detect an object with high performance by self-correcting the object region through the pattern data. However, it can require a vast amount of input data to obtain high detection accuracy performance. Furthermore, as the CNN model is like a black-box, there is no method to verify the correct detection process. Accordingly, CNN-based image detection cannot be easy to trust the judgment of autonomous vehicles controlled because it cannot be confirmed to trust a basis to calculate an estimated value for the detection result. The main cause for this problem is that the CNN structure consists of complex functions with many non-linear elements to calculate the detection result [5–14]. Therefore, a detection system that can reliably determine the judgment of an autonomous vehicle by analyzing the characteristics of an object along with high-precision result inference may be required.

The image detection techniques that require analysis depend on different types of linear element techniques. It is used in three types depending on the method: holistic,

feature-based, and hybrid. In the holistic object recognition method, the input of the object recognition system uses the entire object area. In general, the holistic object recognition method has the advantage of being implemented. However, it has been a disadvantage that leads to lower detection accuracy due to the high complexity of object recognition, such as overlapping, distortion, and deformation. The feature-based method detects spatial features such as the eyes, nose, and mouth extracted, and then the locations. Although it shows a method better than the holistic method, optimally extracting features is complicated by the large volume of spatial feature information. Finally, the hybrid method has the advantages of each technology as it is a mix of holistic and feature-based methods.

This paper will design how object detection can be improved in general images using a detection method based on the PCA [15] structure. Spatial feature extraction using PCA is a multiple variable analysis method and a usual dimensional reduction method. However, PCA-based detection methods obtain heavier covariance matrices by vectoring two-dimensional images in one dimension. To overcome this problem, we use methods to obtain effective computational complexity. One method uses the 2-D PCA (two-dimensional PCA) algorithm [16], which inputs the 2D features of the image as it is without changing the column vector of the image. However, 2-D PCA has the advantage of efficiently extracting the properties of objects, but it has been a disadvantage to solve that it is difficult to extract location features. Further, the more data, the more difficult it is to extract location features, the lower the detection performance. In order to solve these problems, a PCA-based feature extraction method is required considering object features and location features.

Therefore, in this paper, we propose a boost 3-D PCA algorithm on an efficient analysis method to increase object detection accuracy and cognitive analysis reliability due to black-boxes in deep learning. The proposed method has three improvement steps for image object detection results and analysis. The first is the design of an improved PCA-based feature extractor 3-D PCA extractor to consider feature and location data, and the second is a voting-based data set cross-validation partitioning method to solve the data balancing problem that occurs using a large amount of data. Finally, we design a 3-D PCA-based AdaBoost to detect object locations.

2. New Image Detection System for Pedestrian Traffic Safety

The proposed system in this paper uses a linear analysis-based object detection method as shown in Figure 1. The proposed detection system on the Boosted 3-D PCA extracts a high-performance estimated detection result from pedestrians in images obtained from CCTV.

Figure 1. Proposed detection system on the Boosted 3-D PCA.

2.1. Feature Extraction Using New Analysis-Based Method

A feature extractor capable of high-performance object extraction is needed to accurately detect pedestrians in images obtained from CCTV (closed-circuit television) in

intersections. In particular, the system proposed in this paper detects pedestrians in an intersection environment and uses PCA to construct an extractor considering the characteristics thereof. The structure of the proposed feature extractor is designed to reduce the amount of PCA computation, improve accuracy, and process large-capacity input images, like the 2-D PCA algorithm. Therefore, the proposed feature extraction structure shows high-performance feature extraction by enhancing the object detection performance for a large amount of input data after performing the PCA in one more dimension than 2-D PCA. The proposed feature extraction structure extracts three-dimensional features for the image dataset of $W \times H \times D$, as shown in Figure 2, and this is defined as the 3-D PCA method.

Figure 2. Dimension reduction in the 3-D PCA method.

The process of this structure mainly comprises the Y-Z axis PCA and X-Y axis PCA. At first, for the Y-Z axis PCA, when there are M input images of $x_i^{m_1 \times m_2}$ ($i = 1, 2, \dots, M$), which are sequentially entered like $m_1 \times m_2$ of Figure 3, the i -th image is expressed as $x_i^{m_1 \times m_2}$, and after conversion of each of the pixel indexes into 2D data on the X-Y axis, $z_j^{M \times m_2}$ ($j = 1, 2, \dots, m_1$), the PCA is conducted in the row direction.

Figure 3. Dimension reduction in Y-Z axis.

In this process, the PCA covariance matrix for the Y-Z axis is computed. The covariance matrix C_z can be expressed as Equation (1) with $z_j^{M \times m_2}$ and the average image $\bar{z}$.

$$C_z = \frac{1}{m_1} \sum_{j=1}^{m_1} (z_j - \bar{z})(z_j - \bar{z})^T \quad (1)$$

The covariance matrix C_z obtained in this method is projected to $Y_z = U_z^T (A - \bar{z})$ when a new recognition candidate A is given. In this process, the variance in the back-

ground is large compared to the X-Y axis. Therefore, when dimension reduction is performed for data on the Y-Z axis, the background is efficiently separated from the foreground. Then, the covariance matrix is calculated for the X-Y axis PCA from the average image $\bar{X}$ for the input image $x_i^{m_1 \times m_2}$. In order to reduce the complexity of the covariance matrix calculations, three steps are performed. First, for the row direction PCA, using the row directed covariance matrix C_x , the dimensions of the feature matrix U_x are reduced by $x^{m_1 \times dx}$.

The dx represents that the dimension is reduced by large feature matrices of dx in the U_x matrix, and the feature matrix in the row direction has the size of $m_1 \times dx$. The matrix U_x obtained in this manner is projected to $Y_x = U_x^T (A - \bar{X})$, where a new recognition Candidate A is given. Then, for the column direction PCA, using the covariance matrix C_y calculated in the column direction of the dimensional image, the feature matrix V is obtained by dimension reduction with $x^{dy \times m_2}$. The dy represents that the dimension is reduced by large feature matrices of dy in the V matrix, and the feature matrix in the column direction has the size of $dy \times m_2$. The V matrix obtained in this manner is projected to $Y_x = (A - \bar{X})V$. Finally, the results of PCA in the row and column directions are combined. Using the two feature matrices obtained in this manner, when new recognition candidates are entered, they can be projected to $Y_x = U_x^T (A - \bar{X})V$.

2.2. Voting Algorithm to Solve the Data Balance Problem

In image detection, the performance may depend on the composition of the data, such as the amount, size, and shape of the data. This problem is called the data balancing problem, and in order to solve it, a voting-based dataset segmentation algorithm is designed in this paper. The data division uses a cross-validation method of segmentation into K pieces, and this segmented dataset can address the difficulty of object detection with a large amount of data. In addition, as the detected model is generated in parallel by segmenting the data, the power required for computing is reduced. Each image feature map based on the proposed 3D PCA structure is extracted from the segmented datasets. Finally, the background and foreground are detected by applying an SVM (support vector machine) [17] to the feature map. The weight value for the object region is assigned to the segmented dataset by collecting these detected results. Thus, the efficiency of object detection is improved.

2.3. Problem Boosted 3-D PCA Algorithm Based on Modified Adaboost Algorithm

The proposed detection method is designed by modifying the AdaBoost [18,19] structure according to the 3-D PCA extractor structure. This is defined as a new boosted 3-D PCA method. The new boosted 3-D PCA method applies the AdaBoost algorithm to each feature map extracted in the X-Y axis PCA and Y-Z PCA processes for a 2D image. In the new boosted 3-D PCA method structure, the AdaBoost algorithm is applied to each dataset segmented into K -parts, as shown in Figure 4, to extract the low-performance detection result at the K -th level. Then, by combining the low-performance detection results of all stages, high-performance detection results are extracted. To this end, in the new boosted analysis algorithm, the calculation process is performed twice while considering the weight. The first result represents the sum of the transformed feature vectors using the weight $\omega_{k,p}$ of the low-performance detection result for the dataset segmented into K -parts and $v_{2,j}$ for the Y-Z axis. The second result represents the sum of the transformed feature vectors using the weight ω_k of the low-performance detection result on the X-Y axis, the transformed feature vector $v_{1,j}$ and the feature vectors for the first low-performance detection result. This calculation process considering the weights, is performed in K -parts. Here, K -parts are used so that the steps for accurate object detection can be defined. In this paper, K -parts for segmenting the dataset are defined. The generated high-performance detection model extracts an estimated detection value of β ($0 \leq \beta \leq 1$).

Figure 4. Cross-validation segmentation algorithm for solving data balancing problems.

When the value of β is 0, the object extraction attains an estimated detection value corresponding to a negative result. When the value of β is 1, the object extraction extracts an estimated detection value corresponding to a positive result.

3. Experiment and Analysis

In this section, the proposed detection system is comparatively experimented with the latest seven detectors and three benchmark datasets. To validate our proposed system, we have tested it on four publicly available sequences, which are PET'09 S3.MF (multiple flow) and PET'09 S0.CC (city center) from the PET benchmark [20], “Caltech Person” dataset from the California Institute of Technology dataset [21], and the INRIA Person dataset [22]. The first two datasets are consecutive frames captured by one fixed camera, while the sequences of the latter two datasets are chosen from different scenarios. For the urban center sequence, the first 100 frames are selected for testing, as the amount of this sequence is large, while all sequences of the other three datasets are adopted, during this experiment. The Caltech dataset consists of approximately 10 h of video of the vehicle operating in an urban traffic environment. For the INRIA dataset, parts of these images are resized that these pedestrians have a similar size to the training datasets. Pedestrian objects in this data set include a wide variety of scales.

Training datasets are benchmark datasets selected from the INRIA, Caltech, and PETS2000/2001 Pedestrian dataset, with an image size of 64×128 pixels. Datasets apply the k-fold cross-validation method for balancing data that occurs during learning. Furthermore, the data augmentation method is performed for input of the detection system as shown in Figure 5, and the experiment is performed by inputting it into the system.

Experiments on the INRIA dataset perform comparative evaluations with four detectors: CAP [23], HOG [22], GAB [24], CWPAC, CW2DPCA [25] and Y-PD [26]. As shown as in Table 1, the results were derived, and higher accuracy were derived than linear-based detection systems. In addition, the proposed method was improved by 1.3% over the Y-PD results with the YOLOv2 algorithm.

Figure 5. Perform data augmentation for input of detection system.

Table 1. Experimental results for benchmark datasets.

Datasets	Algorithm	Model	Accuracy
INRIA [22]	P. Viola and M. Jones [23]	CAP	71.84%
	N. Dalal and B. Triggs [22]	HOG	89.08%
	J. Friedman [24]	GAB	62.29%
	Z. Liu et al. [26]	Y-PD	90.90%
	Ours	Boosted 3-D PCA	92.24%
Caltech [21]	A. Alkandari and S. J. Aljaber [16]	PCA	88.42%
	Ours	Boosted 3-D PCA	90.26
PETS2000/2001 [20]	F. Alsaqre and O. Almathkour [25]	CWPCA	78.75/76.61%
		CW2DPCA	91.02/90.77%
		Boosted 3-D PCA	93.52/91.84%

In addition, the following experiments were conducted for comparative evaluation with PCA-based detectors: It was evaluated against three detectors, PCA, CWPCA, and CW2DPCA, and PCA was performed on a Caltech dataset basis, and CWPCA and CW2DPCA were performed on a PETS2000/2001 dataset basis. As shown in Table 1, the results were derived, and accuracy improvements such as Caltech: 1.84%; PETS2000/2001: 14.77/15.23%; and 2.5/1.07% were confirmed, respectively. PCA and CWPCA are detectors consisting of PCA with the same backbone, but there is a large difference in their respective comparative performance with the proposed algorithm. As shown in Table 1, we analyzed as the performance difference caused by different datasets. The Caltech dataset is a collection of images, and PETS dataset is a continuous collection of images, resulting in large performance differences due to its complexity of recognition and various object features.

4. Conclusions

In this paper, a new analysis method was designed and developed to enhance pedestrian traffic safety. A boosted 3-D PCA algorithm based on an efficient analysis method. The proposed method performs the following steps to improve detection performance. First, a new feature extraction structure according to the 3-D PCA method is designed. Second, a voting-based dataset cross-validation segmentation structure is designed to solve the data balancing problem. Finally, the detection accuracy performance is improved by designing a modified AdaBoost algorithm suitable for the 3-D PCA method structure. In order to verify the performance of the proposed system, a performance verification experiment using benchmark data was conducted. Experiments were conducted to verify the 3-D PCA method feature extractor's performance, verify the accuracy performance of the pedestrian domain detection, and verify the accuracy performance of pedestrian

and non-pedestrian detection considering the traffic environment. This experiment compares the proposed detection system with the latest seven detectors and three benchmark datasets. The results of the experiments revealed an increase in performance in comparison to an existing method. They showed that the detection accuracy performance of a large amount of data was improved. Therefore, it was confirmed that the new approach is an efficient method of pedestrian detection. The proposed method makes it possible to analyze black-box characteristics, which cannot be done with a CNN-based deep learning structure. In addition, the in-depth detection results produced using linear analysis established a foundation that will be useful for the promotion of pedestrian traffic safety.

Author Contributions: Conceptualization, K.-M.L. and C.-H.L.; methodology, K.-M.L. and C.-H.L.; software, K.-M.L.; validation, K.-M.L. and C.-H.L.; formal analysis, K.-M.L. and C.-H.L.; writing—original draft preparation, K.-M.L.; writing—review and editing, K.-M.L. and C.-H.L.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: This paper was supported by the Semyung University Research Grant of 2021.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lowe, D.G. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110. [\[CrossRef\]](#)
2. Bay, H.; Tuytelaars, T.; van Gool, L. Surf: Speeded up robust features. In Proceedings of the European Conference on Computer Vision, Graz, Austria, 7–13 May 2006; Springer: Berlin/Heidelberg, Germany, 2006; pp. 404–417.
3. Deselaers, T.; Keysers, D.; Ney, H. Features for images retrieval: An experimental comparison. *Inf. Retr.* **2008**, *11*, 77–107. [\[CrossRef\]](#)
4. Simard, P.Y.; Steinkraus, D.; Platt, J.C. Best practices for convolutional neural networks applied to visual document analysis. In Proceedings of the Seventh International Conference on Document Analysis and Recognition, Edinburgh, Scotland, 3–6 August 2003; Volume 2, pp. 958–963.
5. Byambajav, B.; Alikhanov, J.; Fang, Y.; Ko, S.; Jo, G.S. Transfer learning using multiple convnet layers activation features with principal component analysis for image classification. *Korea Intell. Inf. Syst. Soc.* **2018**, *24*, 205–225.
6. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Columbus, OH, USA, 23–28 June 2014; pp. 580–587.
7. Girshick, R. Fast R-CNN. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015; pp. 1440–1448.
8. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. *Commun. ACM* **2017**, *60*, 84–90. [\[CrossRef\]](#)
9. Redmon, J.; Divvala, S.; Girshick, R.; You, A.F. Only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.
10. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. Ssd: Single shot multibox detector. In Proceedings of the European Conference on Computer Vision (ECCV), Amsterdam, The Netherlands, 8–16 October 2016; pp. 21–37.
11. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Amsterdam, The Netherlands, 8–16 October 2016; pp. 770–778.
12. Park, J.-C.; Lee, S.-H.; Jung, J.-U.; Son, S.-B.; Oh, H.-S.; Jung, Y. Uncertainty-Based Deep Object Detection from Aerial Images. *Inst. Control. Robot. Syst.* **2020**, *26*, 891–899. (In Korean) [\[CrossRef\]](#)
13. Son, H.J.; Koh, E.H.; Sung, S.K. A Study on Integrated Navigation Algorithm using Deep learning based Lidar Odometry and Inertial Measurement. *Inst. Control. Robot. Syst.* **2020**, *26*, 715–723. (In Korean) [\[CrossRef\]](#)
14. Ma, G.; Kummert, A.; Park, S.B.; Schneiders, S.M.; Loffe, A. A symmetry search and filtering algorithm for vision based pedestrian detection system. *SAE Tech. Pap. Ser. World Congr. Detroit Mich.* **2008**. [\[CrossRef\]](#)
15. Alkandari, A.; Aljaber, S.J. Principal component analysis algorithm (PCA) for image recognition. In Proceedings of the Second International Conference on Computing Technology and Information Management (ICCTIM), Johor, Malaysia, 21–23 April 2015; pp. 21–23.

16. Daoqiang, Z.; Zhou, Z.-H. Two-Directional two-Dimensional PCA for efficient face representation and recognition. *Neurocomputing* **2005**, *69*, 224–231.
17. Hearst, M.A.; Dumais, S.T.; Osuna, E.; Platt, J.; Scholkopf, B. Support vector machines. *IEEE Intell. Syst. Appl.* **1998**, *13*, 18–28. [[CrossRef](#)]
18. Freund, Y.; Schapire, R.; Abe, N.J. A short introduction to boosting. *J.-Jpn. Soc. Artif. Intell.* **1999**, *14*, 771–780.
19. Cai, C.C.; Gao, J.; Minjie, B.; Zhang, P.; Gao, H. Fast pedestrian detection with Adaboost algorithm using GPU. *Int. J. Database Theory Appl.* **2015**, *8*, 125–132. [[CrossRef](#)]
20. Ferryman, J.; Shahrokni, A. PETS 2009: Dataset and challenge. In Proceedings of the Eleventh IEEE International Workshop on Performance Evaluation of Tracking and Surveillance, Snowbird, UT, USA, 7–9 December 2009. Available online: <http://pets2009.net/> (accessed on 15 February 2021).
21. Hariyono, J.; Jo, K.-H. Detection of pedestrian crossing road. In Proceedings of the IEEE International Conference on Image Processing (ICIP), Quebec City, QC, Canada, 27–30 September 2015; pp. 27–30.
22. Dalal, N. INRIA Person Dataset. Available online: <http://pascal.inrialpes.fr/data/human/> (accessed on 24 March 2021).
23. Viola, P.; Jones, M. Rapid object detection using a boosted cascade of simple features. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Kauai, HI, USA, 8–14 December 2001; pp. 1511–1518.
24. Friedman, J.; Hastie, T.; Tibshirani, R. Additive logistic regression: A statistical view of boosting. *Ann. Stat.* **2000**, *28*, 337–407. [[CrossRef](#)]
25. Alsaqre, F.; Almathkour, O. Moving Objects Classification via Category-Wise Two-Dimensional Principal Component Analysis. *Appl. Comput. Inform.* **2019**. ahead-of-print. [[CrossRef](#)]
26. Liu, Z.; Chen, Z.; Li, Z.; Hu, W. An Efficient Pedestrian Detection Method Based on YOLOv2. *Math. Probl. Eng.* **2018**, *2018*, 10. [[CrossRef](#)]