

REGION-DEPENDENT VEHICLE CLASSIFICATION USING PCA FEATURES

Jon Arróspide and Luis Salgado

ABSTRACT

Video-based vehicle detection is the focus of increasing interest due to its potential towards collision avoidance. In particular, vehicle verification is especially challenging due to the enormous variability of vehicles in size, color, pose, etc. In this paper, a new approach based on supervised learning using Principal Component Analysis (PCA) is proposed that addresses the main limitations of existing methods. Namely, in contrast to classical approaches which train a single classifier regardless of the relative position of the candidate (thus ignoring valuable pose information), a region-dependent analysis is performed by considering four different areas. In addition, a study on the evolution of the classification performance according to the dimensionality of the principal subspace is carried out using PCA features within a SVM-based classification scheme. Indeed, the experiments performed on a publicly available database prove that PCA dimensionality requirements are region-dependent. Hence, in this work, the optimal configuration is adapted to each of them, rendering very good vehicle verification results.

Index Terms— Intelligent vehicles, Hypothesis verification, Principal component analysis, Machine learning, Vehicle database

1. INTRODUCTION

Advanced Driver Assistance Systems (ADAS), and in particular pre-crash sensing, are receiving increasing attention in the last years with the aim of reducing the number of accidents and their severity. Road safety improvement has been the focus of many national and international projects, while manufacturers, universities, and research centers are also devoting great effort and resources to this end. According to traffic statistics, the main threats that a driver faces are usually caused by other drivers, therefore vehicle detection arises as the key challenge for ADAS. It also paves the way for further vehicle tracking and ultimately for collision avoidance.

Specifically, video-based approaches to vehicle tracking play an important role on account of their flexibility, low cost and intrinsic potential. To achieve vehicle detection, most of the video-based methods accounted in the literature proceed in a two-step approach: hypothesis generation, and hypothesis verification. In the former, potential locations of the objects in the image are hypothesized; in the latter, in turn, the presence of objects in the hypothesized regions is checked. The first stage, hypothesis generation, usually relies on knowledge-based methods, which search for specific cues in the image (such as color [1], or edge-density [2]), or motion-based methods (e.g., [3]). Verification of hypotheses, in turn, is commonly

accomplished through appearance-based techniques. These often entail supervised classification problems, in which a set of samples are trained in search of specific feature descriptors of vehicles. Principal Component Analysis, Histograms of Oriented Gradients (HOG) [4], Haar-like features [5], and Gabor filters [6], among others, have been widely applied for feature extraction.

In particular, in this study we address vehicle hypothesis verification through Principal Component Analysis. PCA has been extensively used for feature extraction of many objects, especially faces, and to a lesser extent vehicles [7][8][9][10]. However, existing approaches display a number of shortcomings that limit the applicability of PCA to vehicle detection. First, the importance of the vehicle pose is neglected: indeed, only rear views of vehicles are considered for training and testing the classifier. Hence, although good results are generally proven, they are only significant under the constraint that the vehicles are right in front of the observer, and do not represent a realistic driving situation in which the appearance of vehicles changes according to their relative position. For instance, in [9] the case of detecting passing vehicles from side perspectives is specifically excluded from the study. As a second limitation, approaches reported in the literature usually perform experiments in small (e.g., 90 vehicles for training in [8]) or unpublished datasets ([7]-[10]), which jeopardizes the significance of the experiments. Finally, the dimensionality of the PCA subspace is usually predefined (for instance, in [10] a 20-dimensional PCA vector is arbitrarily chosen) or, at the best, set according to the accumulated variance of the eigenvectors (80% in [9], 90 or 95% in [7]).

In this work, a new approach to PCA-based vehicle detection is proposed that tackles all the aforementioned shortcomings. First, the relative pose of vehicles is taken into account by using a large database of images. Namely, the database consists of four different regions according to the position of the vehicles with respect to the camera (front, left and right in the close/middle range, and far range), each comprising 1000 positive and 1000 negative images. Vehicle detection is thus addressed as a two-class classification problem, in which a different classifier is adapted to each of the specific image regions based on SVM. The database is significant both in size (8000 images) and in variability (different image quality, weather, and illumination conditions are considered). Finally, a thorough analysis is carried out on the performance of PCA as a function of the feature space dimensionality, and a different operation point is adapted for each image region.

2. PRINCIPAL COMPONENT ANALYSIS

It is often the case that in large data sets the vast majority of the points lie close to a subspace of much lower dimensionality than

This work was supported partially by the Ministerio de Ciencia e Innovación of the Spanish Government under projects TEC2010-20412 (Enhanced 3DTV) and TEC2007-67764 (SmartVision).

that of the original data space. Starting from a original space of dimensionality D , PCA aims at finding the space of dimensionality M , with $M < D$, that best grasps the intrinsic information within the data set. This subspace is known as principal subspace, and is formally defined so that the orthogonal projection of the data points onto this subspace maximizes the variance of the projected data [11]. PCA is extensively used for dimensionality reduction, lossy data compression, data visualization, and feature extraction.

A thorough explanation of PCA can be found in [11]. In summary, the principal subspace is given by the first M eigenvectors (i.e., those corresponding to the largest eigenvalues) of the covariance matrix given by

$$\mathbf{S} = \frac{1}{N} \sum_{n=1}^N (\mathbf{x}_n - \bar{\mathbf{x}})(\mathbf{x}_n - \bar{\mathbf{x}})^\top \quad (1)$$

where each sample of the data set is denoted \mathbf{x}_n , and $\bar{\mathbf{x}}$ is the sample set mean, $\bar{\mathbf{x}} = \frac{1}{N} \sum_{n=1}^N \mathbf{x}_n$. The directions of the eigenvectors are also known as the principal components.

As regards image representation through PCA, an image $I_n(x, y)$ of size $R \times C$ is represented as a vector:

$$\mathbf{z}_n = \{z_{n,k}\}_{k=1}^{R \times C}$$

$$z_{n,y+C(x-1)} = I_n(x, y), \quad 1 \leq x \leq R, \quad 1 \leq y \leq C$$

First the average of the image set must be computed:

$$\bar{\mathbf{z}}_n = \frac{1}{N} \sum_{n=1}^N \mathbf{z}_n \quad (2)$$

Then, the eigenvectors $\{\mathbf{u}_i\}_{i=1}^D$ and eigenvalues $\{\lambda_i\}_{i=1}^D$ of the covariance matrix are derived as explained above, and only the principal components are retained, $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_M\}$. Finally, the mean is subtracted from each image and the result is projected onto the principal subspace to provide each feature of the image representation: $\mathbf{f}_n^{(i)} = (\mathbf{z}_n^\top - \bar{\mathbf{z}}_n^\top) \mathbf{u}_i$. The final PCA-based representation of the data samples is composed of the projections onto all the principal components, $\mathbf{f}_n = (\mathbf{f}_n^{(1)}, \mathbf{f}_n^{(2)}, \dots, \mathbf{f}_n^{(M)})$.

3. VEHICLE IMAGE DATABASE

As stated in the introduction, a database of 8000 images has been used to train the classifiers. This is accessible in the Internet at <http://www.gti.ssr.upm.es/data>. In this database, images are divided according to their relative position into four regions: close/middle range in the front, close/middle range in the left, close/middle range in the right, and far range. Each of them comprises 1000 positive images and 1000 negative images. Images are selected from different highway video sequences (acquired with a forward-looking camera) in order to capture the high variability of samples under different acquisition conditions. Besides, images from different databases have been appended to further increase variability. In particular, 500 images are taken from the Caltech [12] and the TU Graz-02 [13] databases corresponding to close vehicles ahead. For the remaining regions (left, right, and far) few images have been found in these databases and their proportion in the used set is residual (only 2.5% of the images). The core is thus composed of images of our collection and is distributed as follows: 60% of the images are devoted to cloudy, sunny, and medium-illumination conditions (20% each), which correspond to the most frequent scenarios; 20% of the images are taken from sequences at down/dusk

Region	P90	P95
Front	114	246
Left	108	233
Right	89	206
Far	60	131
Mean	92.75	204
Single Region	110	267

Table 1. Number of principal components required to retain 90% (P90) and 95% (P95) of the variance when using a different classifier for each image region or a single classifier for all of them.

Region	Linear SVM		Quadratic SVM	
	P90	P95	P90	P95
Front	94.26	91.42	93.18	91.55
Left	91.92	89.54	91.80	90.18
Right	90.38	87.02	91.04	88.98
Far	91.56	89.82	92.58	91.32
Mean	92.03	89.45	92.15	90.51
Single Region	83.94	78.56	86.85	82.95

Table 2. Accuracy for P90 and P95 when using a different classifier for each image region or a single classifier for all of them.

involving lower illumination, 10% display soft rain, 5% are acquired with low resolution cameras, and 2.5% correspond to tunnels (i.e., with artificial light). The same proportions are enforced for the 500 images of our collection in the front close/middle region. Negative samples are extracted in an analogous manner. Each positive image contains a vehicle. Notwithstanding, in order to simulate the output of the hypothesis generation stage, which is often inaccurate, some samples overfit the vehicle (i.e., the image contain not only the vehicle but also some background), while others hold the vehicle only partially (all images in the database cover at least 50% of the vehicle rear). All images are scaled to 64×64 pixels.

4. PCA-BASED VEHICLE CLASSIFIER

The effectiveness of PCA for feature extraction and posterior vehicle recognition is evaluated according to the database presented above. Rather than blindly extracting features of the vehicle and non-vehicle classes and training a classifier, we propose to use a different classifier for each region of the image. Indeed, the larger the variability in the feature space (e.g. in terms of pose, illumination, etc.), the larger the number of principal components required. Therefore, by independently characterizing each of the four image regions, we expect that the dimensionality of the respective principal subspaces be reduced and that the performance be enhanced.

4.1. Experiments and Results

In order to verify this, let us first define a comparison framework. In particular, the sum of the eigenvalues of the discarded eigenvectors is a measure of distortion of the PCA approximation. In other words, the factor $\sum_{i=1}^M \lambda_i / \sum_{i=1}^D \lambda_i$ gives the idea of the percentage of the variance held after the approximation, that is, of the amount of information retained. Some approaches determine the dimensionality of the principal subspace by taking as many components needed as to keep a certain percentage of the variance, rather than taking a fixed number of components regardless of the data. In [7], for

Region	Principal subspace dimensionality										
	10	20	30	40	50	60	70	80	90	100	110
Front	93.92	95.80	95.76	96.22	95.38	95.86	95.62	93.70	93.78	93.96	93.46
Left	88.34	89.30	92.12	92.72	93.08	93.32	92.70	92.06	92.26	90.98	-
Right	84.30	89.04	90.20	90.72	90.92	91.04	90.88	90.70	-	-	-
Far	89.56	90.44	91.22	91.48	91.52	91.56	91.48	91.44	-	-	-

Table 3. Accuracy for different regions as a function of different principal subspace dimensionality.

instance, they perform experiments by retaining 90% and 95% of the variance. We will use the same criteria to establish a comparison framework between the two classifier configurations mentioned.

By using this comparison framework, the advantage of designing independent classifiers for each region instead of a common classifier becomes evident. Table 1 summarizes the number of principal components needed to keep 90 and 95% (referred to as P90 and P95, respectively) of the variance for both settings. As can be observed, the average number of components needed when using independent classifiers for each region is around 15% lower than that of using a single region for P90, and 25% lower for P95. What is more important, the classification performance is significantly enhanced, as the between-class variability is higher. Classification results for P90 and P95 are shown in Table 2. Experiments are performed using 50% holdout cross-validation 5-fold. Also, a linear SVM is used as a baseline classifier for feature comparison. Results are compared in terms of accuracy or correct classification rate, that is, the proportion of correctly classified samples.

Observe that the results are degraded for P95 with respect to P90. This suggests that using a too large set of features is counteractive for the classifier. This behavior is not connected to the linear nature of the classifier: experiments have been performed with a quadratic SVM (see Table 2), for which degradation is also observed when increasing the dimensionality of the principal subspace to keep 95% of the variance. Therefore, there must be an optimum number of features rendering maximum accuracy, which shall also depend on the image region. To determine the corresponding operation points, experiments have been conducted by varying the number of used principal components for each image region: multiples of 10 have been considered up to the dimensionality of the P90 principal subspace. Accuracy rates as a function of the principal subspace dimension are enclosed in Table 3, and graphically represented in Fig. 1. As expected, we observe that for all the image regions accuracy reaches a maximum in between 10 and the P90 subspace dimension and then starts to decrease. Naturally, the operation point is not the same for all regions. Namely, for the front close/middle region, high classification performance is achieved even for very low subspaces and a maximum accuracy of 96.22% is reached at $M = 40$. For the left and right close/middle regions and the far range more principal components are required, all the three having the operation point at $M = 60$, and respective accuracies of 93.32%, 91.04% and 91.56%. When adjusting the operation point according to the need of each region of the image, the average detection rate is thus of 93.04%.

Exact comparison with the methods reported in the literature is not possible since, as stated in the introduction, most of them use proprietary databases. However, rough comparison with the rates provided by these works unveil the superiority of the proposed strategy. For instance, in [10] a maximum detection rate of 96.09% is obtained using a polynomial kernel of order 5 for SVM, which is lower than the 96.22% obtained with our method for the front region. They use a fixed dimensionality of the feature space, i.e. 20,

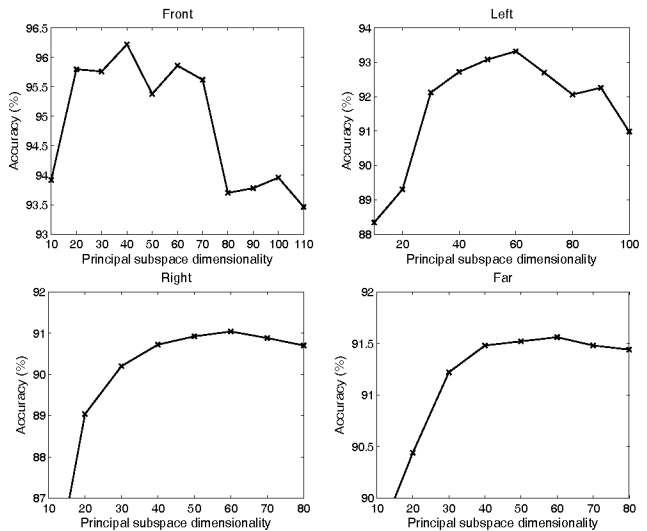


Fig. 1. Accuracy evolution for each image region as a function of the principal subspace dimensionality.

which is not optimal as has been shown in this study. Slightly worse classification performance is reported in [9]: a maximum recall of 95% and precision of 89.2% are obtained with a combined method using PCA and ICA (Independent Component Analysis). In this case the principal subspace is set to capture 80% of the variance. The PCA-SVM method proposed in [8] achieves maximum recall and precision rates of 75% and 86.5% for passenger cars, which is significantly worse than our method, although the used database seems also too small. These rates confirm the importance of the correct selection of the principal subspace dimensionality, as shown by the experiments performed above.

The proposed method does not only outperform existing methods in terms of accuracy. Indeed, the methods reported above only address classification of front rear-viewed vehicles, while the method proposed in this paper incorporate other viewpoints. Specifically, an extended database has been used involving also samples observed from left, right and far viewpoints, and independent training has been performed for each of them according to their different nature.

Although the obtained results are rather high, there seems to be a limit in the achievable rates through PCA. This suggests that the features extracted through PCA are intrinsically non-separable, that is, that there is some degree of overlapping between the vehicle and the non-vehicle classes. Since the database consists of instances of different nature according to weather and illumination conditions, resolution and source, let us graphically inspect the distribution of each of this image subgroups for both classes. In order to visualize it,

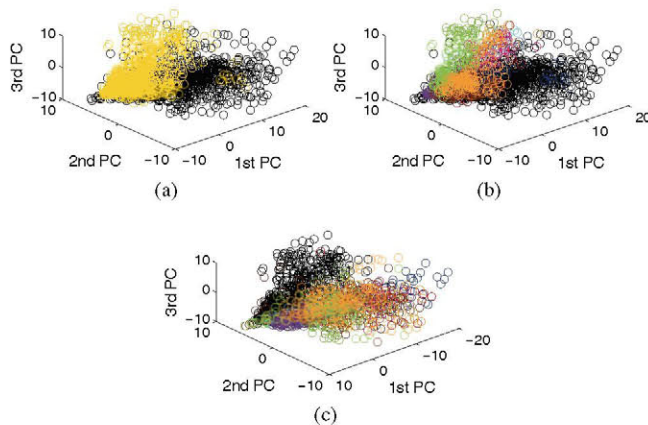


Fig. 2. Sample projection onto the 3-dimensional principal subspace for left close/middle region: (a) vehicle samples are shown in yellow and non-vehicle samples in black; (b) vehicle samples are broken down according to their nature and non-vehicle samples remain black; (c) non-vehicle samples are broken down and vehicle samples are shown in black.

only the first three principal components are considered. Although better separation is achieved when introducing more components, 3D representation eases interpretation of the classifier behavior.

To begin with, Fig. 2(a) shows the data distribution of both classes for the left region. This figure confirms that there is some class overlapping in the PCA feature space. Fig. 2(b) illustrates the vehicle sample distribution broken down in subgroups, according to the categorization explained in Section 3. The color code is the following: cloudy (orange), sunny (brown), medium-illumination (green), dusk (purple), rain (pink), low-resolution (blue), tunnel (red), and images from other databases (cyan). This figure hints that each of the vehicle class subgroups features different separability from the non-vehicle class. Analogously, the different subgroups of the non-vehicle class also cluster in different areas of the principal subspace (see Fig. 2(c)). This suggests that better classification results might be obtained through PCA analysis by designing a different classifier not only for each image region, but also according to the different nature (in terms of weather, illumination, resolution, etc.) of the images. Naturally, this approach involves a number of additional requirements. First, the image database needs to be amply enlarged: 1000 images of each subclass are needed to have the same representativity degree, which results in 64000 images (4 regions, 8 subgroups, 2 classes). On the other hand, an on-line system must be designed in order to detect the illumination and weather conditions of the image, and the camera resolution must be set in advance. The advantages and convenience of using this approach will be investigated in future work.

5. CONCLUSIONS AND FUTURE WORK

In this paper a new approach to PCA-based vehicle verification has been presented that addresses the shortcomings of the state-of-the-art methods. Specifically, the approach proposes region-dependent classification of candidates, in such a way that the relative pose of the vehicle is taken into account. Exhaustive experiments have been conducted, which confirm that region-specific classifiers outperform the monolithic classifier used in classical approaches. In addition,

in order to fill the gap of existing approaches, which typically rely on proprietary databases or use small datasets comprising only rear-viewed vehicles, a new database including 1000 vehicles and 1000 non-vehicle images for each of the regions has been used.

Future work will address the generation of a larger database comprising different categories within the vehicle and the non-vehicle classes according to weather and illumination conditions and image quality, which have been shown to have great potential towards the increase of inter-class separability.

6. REFERENCES

- [1] L.-W. Tsai, J.-W. Hsieh, and K.-C. Fan, "Vehicle detection using normalized color and edge map," *IEEE Trans. Image Processing*, vol. 6, no. 3, pp. 850–864, 2007.
- [2] J. Arróspide, L. Salgado, M. Nieto, and F. Jaureguizar, "On-board robust vehicle detection and tracking using adaptive quality evaluation," in *Proc. IEEE International Conference on Image Processing*, San Diego, CA, 2008, pp. 2008–2011.
- [3] J. Arróspide, L. Salgado, M. Nieto, and R. Mohedano, "Homography-based ground plane detection using a single on-board camera," *IET Intelligent Transport Systems*, vol. 4, no. 2, pp. 149–160, 2010.
- [4] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Diego, CA, 2005, vol. 1, pp. 886–893.
- [5] A. Haselhoff and A. Kummert, "A vehicle detection system based on haar and triangle features," in *Proc. IEEE Intelligent Vehicles Symposium*, Xi'an, China, 2009, pp. 261–266.
- [6] H. G. Jung, Y. H. Lee, P. J. Yoon, I. Y. Hwang, and J. Kim, "Sensor fusion based obstacle detection/classification for active pedestrian protection system," *Advances in Visual Computing, Lecture Notes in Computer Science*, vol. 4292, pp. 294–305, 2006.
- [7] Z. Sun, G. Bebis, and R. Miller, "Monocular precrash vehicle detection: features and classifiers," *IEEE Trans. Image Processing*, vol. 15, no. 7, pp. 2019–2034, 2006.
- [8] C. Zhang, X. Chen, and W. b. Chen, "A pca-based vehicle classification framework," in *Proc. International Conference on Data Engineering Workshops*, Atlanta, GA, 2006.
- [9] C.-C. R. Wang and J.-J. J. Lien, "Automatic vehicle detection using local features a statistical approach," *IEEE Trans. Intelligent Transportation Systems*, vol. 9, no. 1, pp. 83–96, 2008.
- [10] J. Zhou, D. Gao, and D. Zhang, "Moving vehicle detection for automatic traffic monitoring," *IEEE Trans. Vehicular Technology*, vol. 56, no. 1, pp. 51–59, 2007.
- [11] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer Science+Business Media, LLC, New York, NY, 2006.
- [12] Pasadena Computational Vision at California Institute of Technology, "The Caltech database," <http://www.vision.caltech.edu/html-files/archive.html>, accessed 28 Dec. 2011.
- [13] Graz University of Technology, "The TU Graz-02 database," http://www.emt.tugraz.at/~pinz/data/GRAZ_02/, accessed 28 Dec. 2011.