# Localised Contourlet Features in Vehicle Make and Model Recognition

I.Zafar, E.A.Edirisinghe, B.S.Acar
Department of Computer Sciences, Loughborough University, LE11 3TU, Loughborough U.K

## ABSTRACT

Automatic vehicle **M**ake and **M**odel **R**ecognition (MMR) systems provide useful performance enhancements to vehicle recognitions systems that are solely based on **A**utomatic **N**umber **P**late **R**ecognition (ANPR) systems. Several vehicle MMR systems have been proposed in literature. In parallel to this, the usefulness of multi-resolution based feature analysis techniques leading to efficient object classification algorithms have received close attention from the research community. To this effect, Contourlet transforms that can provide an efficient directional multi-resolution image representation has recently been introduced. Already an attempt has been made in literature to use Curvelet/Contourlet transforms in vehicle MMR. In this paper we propose a novel localized feature detection method in Contourlet transform domain that is capable of increasing the classification rates up to 4%, as compared to the previously proposed Contourlet based vehicle MMR approach in which the features are non-localized and thus results in sub-optimal classification. Further we show that the proposed algorithm can achieve the increased classification accuracy of 96% at significantly lower computational complexity due to the use of Two Dimensional Linear Discriminant Analysis (2DLDA) for dimensionality reduction by preserving the features with high between-class variance and low inter-class variance.

**Keywords**: Vehicle MMR, Contourlet Transform, Feature Matching, multi-resolution image analysis, 2DLDA.

## 1. INTRODUCTION

A significant amount of research has been carried out in the area of computer vision based vehicle classification. However these classification techniques have been limited mostly to algorithms distinguishing between different categories of vehicles i.e. car, bus, truck etc. In contrast, an effective vehicle recognizing system solicits the need of correctly identifying the make and model of vehicles within a given category. Several vehicle recognition systems based on correctly recognizing only the vehicle registration plates, are in widespread commercial use at present. However reports by police and media sources have indicated that number-plate cloning, i.e. using bogus registration plates, have been recently used to breach the security. This problem can be addressed by enhancing the reliability of access control systems by using a hybrid approach, i.e. the combined use of ANPR and vehicle MMR, automatic identification of a vehicle's visual description comprising of either one or more of properties such as, make, model and color techniques.

Vehicle MMR is a comparatively new research area. The basic idea is to extract suitable features from the images of a vehicle, which can help in recognizing its make and model. A relatively limited number of techniques that directly relate to vehicle MMR have been proposed in literature. Petrovic and Cootes [1, 2] proposed techniques for the recognition of cars by extracting gradient features from images. A number of feature extraction algorithms including direct and statistical mapping methods were applied to **R**egions-**O**f-**I**nterest (ROI) of frontal views of cars, to obtain sampled structures. These feature vectors were then extracted and classified using simple nearest neighbor classification methods. Munroe and Madden [3] investigated the use of machine learning classification techniques in vehicle MMR. Initially a Canny edge detector followed by a dilation process was used to extract feature vectors. Subsequently different machine learning classifiers were used to determine vehicle make and model associated with each feature vector. Dlagnekov [4] and Zafar, Edirisinghe and Acar [5] explored the problem of MMR by using **S**cale Invariant **F**eature **T**ransforms (SIFT) by Lowe [6]. It is used to identify the points of interest in car images which are subsequently utilized in matching. Zafar, Edirisinghe and Acar in [5] proposed a further improvement to this basic approach via restricting the SIFT key point detection to only the query image and using a SIFT descriptors belonging to all points within a maximum-likelihood area of the candidate images, for matching. Anthony [7] extended the work done by Dlagnekov [4] by replacing the SIFT features with features which characterize contour lines. In this approach, initially, edges are extracted from the rear views of car images. These are then extended to line segments by using a strip based line generator algorithm and are

Image Processing: Machine Vision Applications II, edited by Kurt S. Niel, David Fofi, Proc. of SPIE-IS&T Electronic Imaging,
SPIE Vol. 7251, 725105 · © 2009 SPIE-IS&T · CCC code: 0277-786X/09/$18 · doi: 10.1117/12.805878

subsequently used in matching. Zafar et.al [8] proposed the use of 2DLDA [12] in vehicle MMR, effectively utilizing the 2DLDA's ability to optimize the ratio between, between class scatter and within class scatter. The algorithm's recognition robustness as compared to previously proposed Principle Component Analysis (PCA) based approaches was demonstrated. Kazemi et.al in [9] investigated the use of Fast Fourier Transforms, Discrete Wavelet Transforms and Discrete Curvelet Transforms based image features in vehicle MMR.Rahati et.al in [10] proposed the direct replacement of Curvelet transforms in [9] with Contourlet transforms for vehicle MMR. As the work presented in this paper is based on the effective use of Curvelet transforms in vehicle MMR, a more detailed review of [9] and [10] is postponed to the next section. Negri et.al in [11] proposed an oriented-contour point based voting algorithm for multiclass vehicle type recognition, which is was particularly proved to be effective under occlusion

Although a number of different approaches have been published in literature for vehicle MMR, the search for a robust, efficient algorithm still remains an open research problem. In this paper we attempt to contribute to the current state-of-the-art in vehicle MMR. For clarity of presentation this paper is divided into five sections. Apart from this section which provided an insight to the problem domain and the existing state-of-the art solutions, section 2 provides a conceptual comparison between multi-resolution image analysis techniques, i.e., Wavelet, Curvelet and Contourlet transforms, leading to a discussion on research motivation. Section 3 presents the proposed Contourlet based approach to car MMR. Section 4 provides experimental results and a detailed analysis. Finally section 5 concludes with an insight to possible future improvements to the methodology proposed in this paper.

## 2. RESERACH MOTIVATION

All vehicle MMR approaches proposed in literature (see review in section 1) are based on an initial stage of feature detection, where the detected features are subsequently used in matching and classification. Majority of the methods proposed in literature use edge maps as features. However, the pixel domain edge extractors (e.g. Canny) are limited in their performance and therefore fail in accurately capturing the smooth curves/contours which are an important part of most images. Kazemi et.al's proposal of [9] which is a vehicle MMR method based on Discrete Curvelet Transforms [13] by Candes et.al, is the first attempt to address this problem. Curvelet transforms provide a multi-resolution, band pass and directional functional analysis method which is useful to represent the image edges and curved singularities more efficiently. Therefore they are more accurate in representing curved edges as compared to traditional wavelet transforms [13]. The multi-resolution representation of edges and curved singularities enables effective feature matching across the scales of the Curvelet transform, thus improving the robustness of the car MMR algorithm. However a major challenge in capturing the geometry and the directionality in images comes from the discrete nature of the data: the input is typically, sampled images defined on a rectangular grid. For example, directions other than horizontal and vertical look very different on a rectangular grid. In other words because of pixelization, the notion of smooth contours on sampled images are not obvious. For this reason unfortunately, a mathematical transform such as Curvelet transform that is initially developed in the continuous domain and then later discretized for sampled data [13] is not effective to be used with digital images.

Identifying the above shortcomings a new bread of transforms named Contourlet transforms was proposed by Do and Vetterli [14] in the discrete form as a simple directional extension for wavelets. Contourlet transforms starts with a discrete-domain construction and then studies its convergence to an expansion in the continuous domain and is thus more suitable for digital image processing. It provides improvements to 2-D separable wavelet transforms for representing images with smooth contours in all directions (see Figure 1).

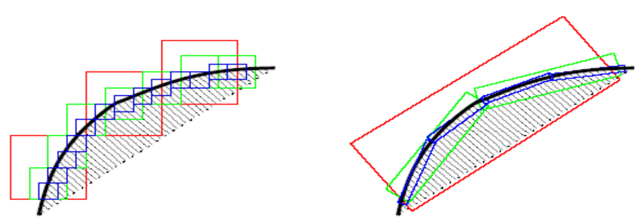


Fig. 1. Wavelet vs. Contourlet painting of the contours [14].

Identifying this advantage, Rahati et.al [10] proposed the direct replacement of Curvelet transforms in [9] with that of Contourlets. The idea was to pass each image through a Contourlet filter bank (see Figure 2) converting the images to an n-level multi-resolution decomposition (see Figure 3). The feature detection and classification was subsequently performed on the decomposed images, rather than on the pixel domain.

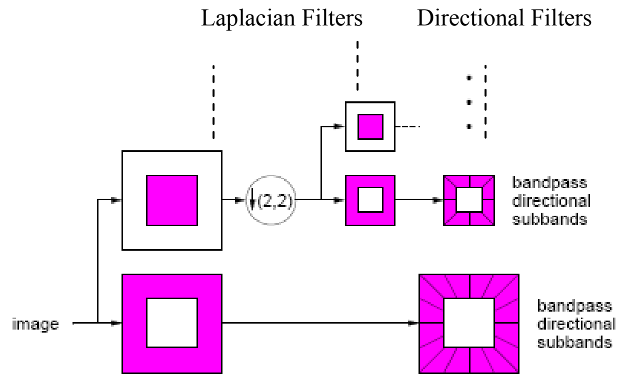



Fig. 2. The Contourlet filter bank [14].

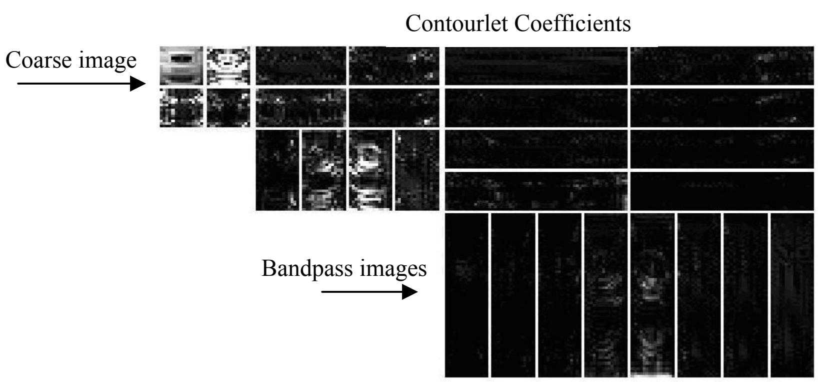



Fig. 3. Multi scale and multi directional decomposition of car image using contourlet transform.

In approaches [9, 10], features used are directly derived from the coefficients of sub-bands. Although an attempt has been made in exploiting the redundancy present in the Contourlet domain by using only pre selected sub-bands for image matching, the feature used in matching is the standard deviation of coefficient values of sub-bands. Unfortunately this measure ignores the localization of features that are an important property which often provides discrimination capability between classes in vehicle type recognition. Therefore we propose to use additional features of the Contourlet decomposition, namely, the local texture in the lowpass sub-band and the presence of oriented image edges across the scales of directional resolutions. Finally a Support Vector Machine (SVM) with a polynomial kernel is used as the classifier. [Note: A comparison of the performance of the proposed technique with that of [9] and [10] is presented in section 4]. The following section explains the proposed novel vehicle MMR technique in detail.

## 3. PROPOSED METHOD

Figure 4 illustrates a block diagram of the proposed approach to Contourlet based vehicle MMR. Details of the various modules are given below:

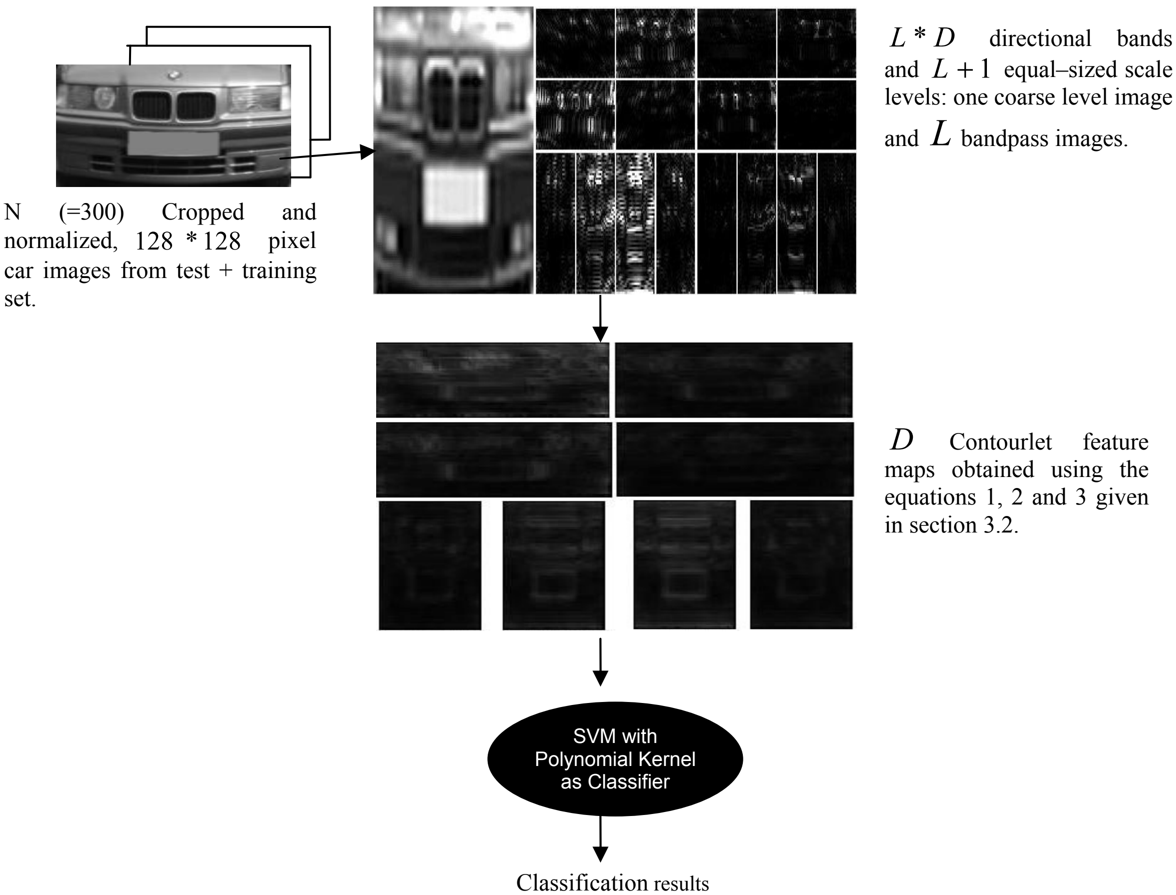



Fig.4. Proposed Recognition Algorithm.

### 3.1 Pre-Processing of test and training images

Initially all images are cropped using the algorithms proposed in [1], making sure that the cropped areas encapsulate all important components of the frontal views, such as the grill, lights, badge, bumper area, etc., of the cars but eliminates background clutter as much as possible. Subsequently all cropped images are normalized in size to $128 * 128$ pixels.

### 3.2 Contourlet decomposition

In this stage, cropped images of cars are decomposed by using contourlet transform. According to the theory of Contourlet transforms, Contourlet decomposition is constructed by applying two successive decomposition stages. The first stage transforms the original image in to a Laplacian Pyramid (LP) having $L+1$ scale levels. The second stage decomposes each LP scale level in to $D$ directional subbands by using a directional filter bank (see Figure 2). Inspired by the idea of extracting visually significant coefficients from [15, 16], we have discarded the down sampling operation in the Laplacian pyramid scheme. As a result we get $L+1$ equal–sized scale levels: one coarse level image and $L$ bandpass images. Same number of directions is obtained for each of $L$ scale levels, thus obtaining a total of $L * D$ equal size directional bands (see Figure 5). Note that in our experiments $L = 2$ and $D = 8$.

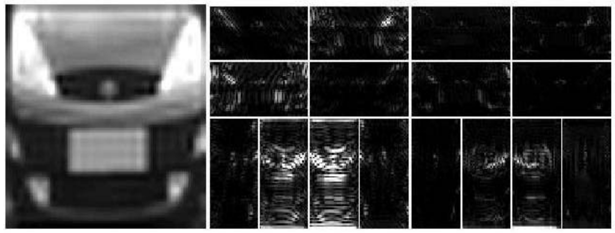



Fig. 5. Contourlet decomposition of an image of a car with equal sized scales and directions.

### 3.3 Feature selection

In section 2, the importance of using a localized feature extractor as against a measure such as the variance of coefficients values of the sub-bands, was discussed. As a result within the research context of the proposed work we have adopted a modified version of the localized feature extraction technique of Bouzidi, and Baaziz [15]. For each frequency direction *d* of the contourlet decomposition, localized Contourlet features are extracted through the calculation of a directional map $M_d$ given by following equation:

$$M_d(i, j) = T(i, j)^{\alpha} E(d, i, j)^{\beta} \tag{1}$$

Where

$$T(i, j) = \operatorname{var}\{ C_{L+1}(4(i-1)+1+y, 2j-1+x)\}_{x=0,2;\, y=0,3} \tag{2}$$

$$E(d, i, j) = \frac{1}{2}\sum_{l=1}^{L} \sum_{x=0,1} \sum_{y=0,1} [C_{l,d}(i+y, j+x)]^2 \tag{3}$$

In the above set of equations, $C_{l,d}$ represents the Contourlet sub-band at scale level *l* and frequency direction *d,* while $C_{L+1}$ is the coarse level sub-image. Therefore *T* and *E* are respectively the local texture from the original filtered image (i.e. the coarse image) and the oriented edges across the scales of directional resolutions respectively. As a result of the above feature detection we obtain a total of D Contourlet feature maps (with localized properties), i.e. one for each direction. Note that the directional feature map $M_d$ proposed in [15] consisted of a 'Brightness' term, but has been discarded in the above modified version. This is due to the reason that in vehicle MMR, matching between images should be invariant to changes in brightness. Note that in our experiments the values of $\alpha$ and $\beta$ were experimentally found out to be 0.1 and 0.2 respectively.

### 3.4 Dimensionality reduction

Though the feature matrix for each image, extracted as a result of proposed technique, is much lower in dimension $(128 * 128)$ than the traditional contourlet features obtained from multiple scales and directions, a further dimensionality reduction may optionally be applied by making use of 2DLDA approach [12]. According to the theory of 2DLDA [12], it provides distinction between classes by decreasing inter-class variance and increasing between-class variance. Thus by preserving the features with high between-class variance and low inter-class variance, feature matrix for each image is further reduced to enhance the recognition speed of classifier.

### 3.5 Support Vector Machine based classification

Finally the feature map obtained following the procedure described in section 3.3 and section 3.4 is used as input to a Support Vector Machine (SVM) classifier first proposed by Cortes and Vapnik in [17]. A SVM is an effective method for general purpose pattern recognition and is a powerful classification tool. The basic idea is to map input data into a high dimensional space and to find the separating hyper plane with maximal margin. A support vector (SV) kernel is used to map the data from the input space to the high-dimensional feature space which facilitates the problem to be processed in a linear form.

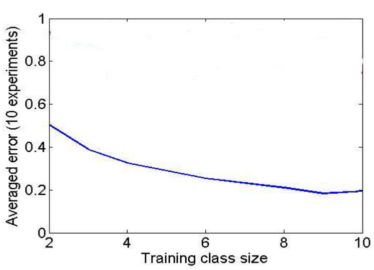
In literature it has been generally concluded that for most applications, a low degree polynomial kernel or a RBF kernel works quite well. In our case, a polynomial kernel with degree 1 and 2 provided the best results. Note that SVM 'one vs. all' has been used here. Theoretical details and mathematical derivations with regards to SVM can be found in [17].

## 4. EXPERIMENTAL RESULTS AND ANALYSIS

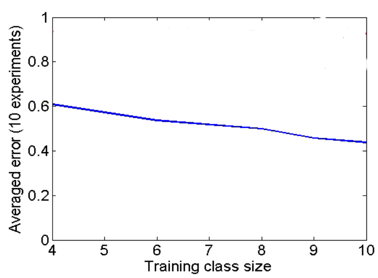
The experiments were conducted on a database of 300 images (frontal views) of cars belonging to 25 different classes [8]. Each class consisted of at least 11 images of different cars belonging to the same make and model. The images were cropped and normalized to $128 * 128$ pixels.

Initially an experiment was designed to determine whether increasing the size of training set improves the classification accuracy. It was observed that when the training set size increases, the accuracy values improve. For example in Figure 6(c) when the training set size increases from 4 to 10, the average classification error reduces from approximately 20% to 10%.

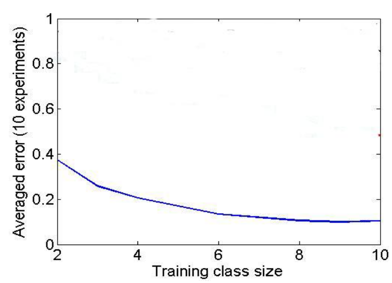
A second experiment was designed to determine the effect of using the novel Contourlet feature selection criteria presented in section 3.3 and to compare it with directly using Contourlet coefficients from different combinations of scales-levels, as features. The results are presented in Figure6.Comparing Figure 6 (g) with the rest of graphs illustrated in Figure 6 (i.e. (a) to (f)), we can clearly see that the proposed localized Contourlet feature selection criteria results in an improved classification accuracy when applied to vehicle MMR. An accuracy rate of around 94% is achieved as the training class size is improved to 10. The best results obtained when using the traditional approach is when all sub-bands are used (see figure 6(f)). The highest accuracy level obtained is 90%, when the training class size is 10.



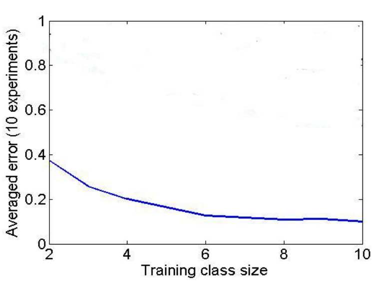

L2 (a)



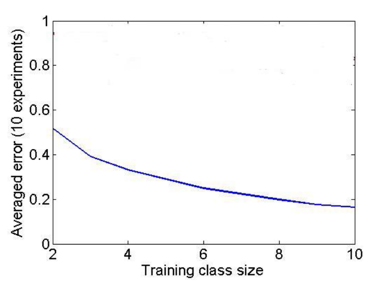

L3 (b)



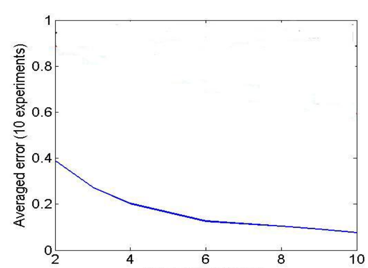

Coarse Level (c)



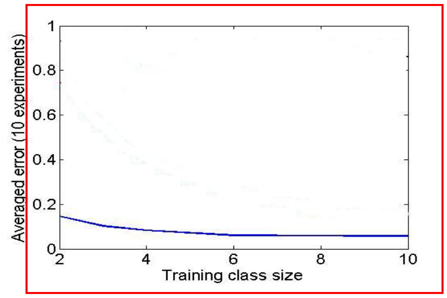

Coarse Level, L2 (d)




L2L3 (e)



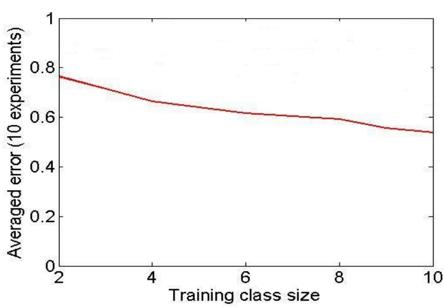

L2 L3 Coarse level (f)



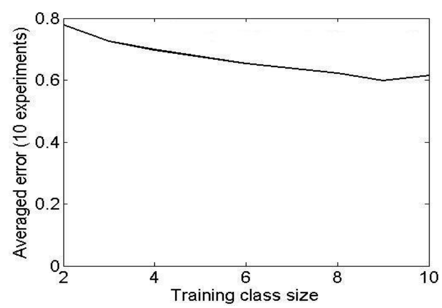

Proposed technique (g)

Fig. 6. SVM 'one against all' classification error vs. size of training class, (a)-(f): directly using Contourlet coefficients from different combinations of scales-levels, (g): using the proposed Contourlet feature selection criteria.

A further experiment was performed to evaluate the use of the feature representation and feature vector selection technique of Rahati et.al [10] on our database of car images for comparison purposes. According to Rahati et.al [10], best accuracy results are obtained when using scale level 3 and 4, divided into 8 and 16 directional sub-bands, respectively. Therefore we obtain a 24 element feature vector of standard deviations for each car image. The results are illustrated in Figure 7. Figure 7(a) shows the results when we use SVM 'one against all' [17], whereas Figure 7(b) uses SVM 'one against one and all'. Comparing results of graph in Figure 7(a) and the results of the proposed technique illustrated in Figure 6(g), it is clear that the proposed algorithm outperforms [10] by achieving an accuracy figure of 94%, as against an accuracy figure of approximately 50%. Our experimental results reveal that the use of SVM 'one against one and all' (see Fig. 7(b)) gives worse level of classification accuracy of approximately 40%.




(a)




(b)

Fig.7. (a): SVM 'one against all' classification error vs. training class for using standard deviation of scale-levels 3&4 of [10]. (b): SVM 'one against one and all' classification error vs. training class for using standard deviation of scale-levels 3&4 of [10].

A further experiment was performed to first reduce the dimensionality of extracted features (see section 3.3) using 2DLDA while enhancing the between-class variance and subsequently using the reduced feature set in classification (see section 3.4). In Figure 8, the classification accuracies obtained after dropping 10 and 100 small value eigenvectors (from a total of 128 eigenvectors) are compared with that obtained when 2DLDA is not used and used, but all eigenvectors are considered. The results illustrate that when a significant number of eigenvectors are dropped (for example 100) the classification accuracy increases, particularly when the training class size is small (for example 2). The reduction in dimensionality has resulted in the use of the optimum discriminant features in classification, hence improving the accuracy. By dropping 100 least significant projection axes, a feature matrix of $28 * 28$ dimensions is obtained for each car image. This is substantially lower in dimension than the feature matrix obtained as a result of section 3.3, which is $128 * 128$. Hence a further advantage of reducing the dimensionality of the feature space is a significant reduction of the computational cost of classification.

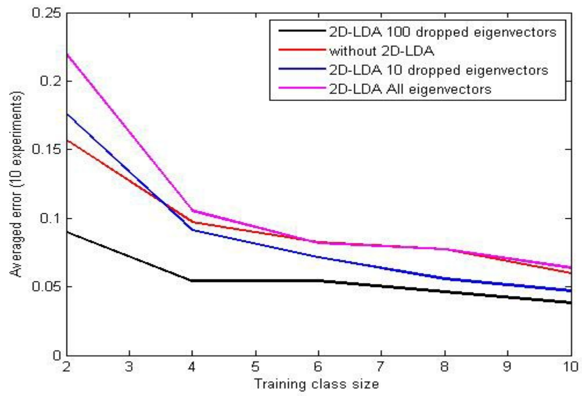
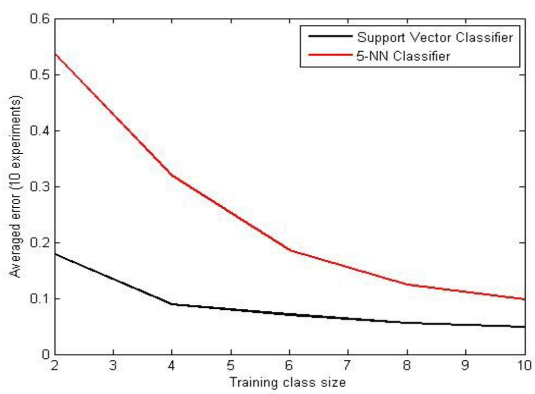
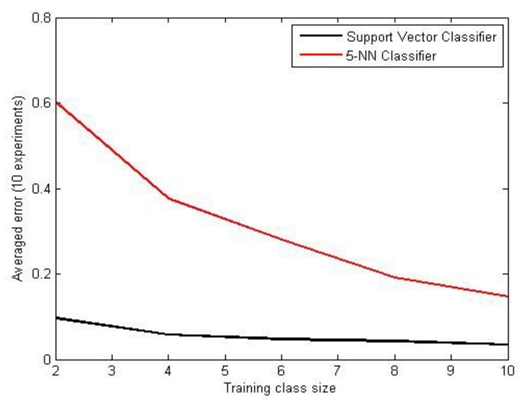



Fig. 8. Classification accuracy vs. training class size: using 2DLDA as dimensionality reduction tool.

We performed a last experiment to compare the performance of SVM against a commonly used classifier *k*-Nearest Neighbor (*k*-NN). Figure 9(a) and 9(b) illustrate the accuracy results of using SVM and *k*-NN classifiers on feature sets extracted following proposed procedures presented in sections 3.3 and 3.4, respectively. The results of Figure 9 illustrate that SVM outperforms *k*-NN.




(a)




(b)

Fig. 9 (a).SVM vs. *k*-NN classification results on extracted features from section 3.3. (b). SVM vs. *k*-NN classification results on reduced features after dropping 100 projection axes from section 3.4.

## 5. CONCLUSION

In this paper we have proposed a novel approach to automatic vehicle MMR, based on a localized directional feature selection criterion, on the Contourlet transform domain. We have shown that the use of localized features in Contourlet transform domain results in a substantial increase in classification accuracy as compared to using variance/standard deviation of Contourlet coefficient values. Further this increased accuracy is obtained at a significant reduction of computational cost. We have further concluded that 2DLDA can be used as a means for dimensionality reduction, without affecting the classification accuracy. This enables the reduction of the computational complexity of the classification stage substantially. Experimental results further concluded that SVM performs substantially better as a classifier in vehicle MMR as compared to k-NN ( $k = 5$ ) classifier.

## REFERENCES

1. Petrović, V., Cootes, T., "Analysis of features for rigid structure vehicle type recognition", Proc. of the BMVC,(2004).
2. Petrović, V., Cootes, T., "Vehicle type recognition with match refinement", International Conference on Pattern Recognition, 3, 95-98(2004).
3. Munroe, D.T., Madden, M.G., "Multi –class and single-class classification approaches to vehicle Model recognition from images", Proc.AICS, (2005).
4. Dlagnekov, L., "Video-based car surveillance: License plate make and model recognition", Masters Thesis, University of California at San Diego,(2005).
5. Zafar,I., Edirisinghe,E.A., Acar,B.S., "Vehicle Make & Model Identification using Scale Invariant Transforms ", Proceeding(583) Visualization Imaging and Image Processing(2007).
6. Lowe, D.G.,"Distinctive Image Features from Scale-Invariant Keypoints", International Journal of Computer Vision, (2004).
7. Anthony, D.," More local structure information for Make -Model Recognition".
8. Zafar,I., Edirisinghe,E.A., Acar,B.S., Bez,H.E., "Two Dimensional Statistical Linear Discriminant Analysis for Real-Time Robust Vehicle Type Recognition", Int. Conf. on Real-Time Image Proc., Proc. of the SPIE, vol 6496, pp. 649602 (2007).
9. Kazemi, F., Samadi, S., Pourreza, H., Akbarzadeh, M., "Vehicle Recognition Based on Fourier, Wavelet and Curvelet Transforms - a Comparative Study", IJCSNS Int. Jou. Of Comp. Sci. and Net. Secu. 7(2), (2007).
10. Rahati, S., Moravejian, R., Kazemi, E.M., Kazemi, F.M., "Vehicle recognition using contourlet transform and SVM", Proc. of the 5th IEEE Int. Conference on Information Technology: New Generations, 894-898(2008).
11. Negri, P., et.al, "An Oriented-Contour Point Based Voting Algorithm for Vehicle Type Classification", Proceedings of the 18th International Conference on Pattern Recognition, 1,574 – 577(2006).
12. Li, M., Yuan, B.,"2D-LDA: A statistical linear discriminant analysis for image matrix", Pattern Recognition letters, 26, 527-532(2005).
13. Candes,E., Demanet,L., Donoho,D., Ying,L., "Fast Discrete Curvelet Transforms", Int. Journal on Multiscale Modelling & Simulation (MMS), 5(3), 861-899(2006).
14. Do, M.N., and Vetterli, M., "The contourlet transform: an efficient directional multiresolution image representation", IEEE Transactions Image on Processing, 14(12), 2091-2106(2005).
15. Bouzidi, A., Baaziz, N.,"Contourlet Domain Feature Extraction for Image Content Authentication", IEEE 8th Workshop on Multimedia Signal Processing, 202 – 206(2006).
16. Barni, M., Bartolini, F., Piva, A.,"Improved wavelet-based watermarking through pixel-wise masking", IEEE Transactions on Image Processing, 10(5), 783 – 791(2001).
17. Cortes, C., and Vapnik, V., "Support-Vector Networks", Machine Learning, 20(3), 273-2979(1995).