Machine Vision and Applications
DOI 10.1007/s00138-008-0145-z

ORIGINAL PAPER

Traffic sign recognition system with β-correction

Sergio Escalera · Oriol Pujol · Petia Radeva

Received: 2 February 2007 / Accepted: 1 May 2008

Abstract Traffic sign classification represents a classical application of multi-object recognition processing in uncontrolled adverse environments. Lack of visibility, illumination changes, and partial occlusions are just a few problems. In this paper, we introduce a novel system for multi-class classification of traffic signs based on error correcting output codes (ECOC). ECOC is based on an ensemble of binary classifiers that are trained on bi-partition of classes. We classify a wide set of traffic signs types using robust error correcting codings. Moreover, we introduce the novel β-correction decoding strategy that outperforms the state-of-the-art decoding techniques, classifying a high number of classes with great success.

Keywords Multi-class classification · Error correcting output codes · Embedding of dichotomizers · Object recognition · Traffic sign classification · Adaboost

1 Introduction

Traffic sign classification in uncontrolled environments is a hard task in computer vision due to the high variability of symbol appearance caused by illumination changes, lack of visibility, or occlusions. In the last few years, several approaches to deal with the problem have been proposed. Usually, traffic sign recognition strategies are divided into two main groups: color-based and gray scale-based. Gray scale-based approaches focus on object geometry [1,2], whereas the color-based techniques have the advantage of preventing false positive detection [3,4]. Traffic sign recognition is studied for several purposes, like autonomous driving under certain simplified conditions or for assisted driving [5]. We focus on the goal of mobile mapping [6], as a technique used to compile cartographic information from a mobile system. One of the main difficulties that makes this problem hard is the great number of classes and the high resemblance among signs in poor resolution images. In order to deal with these hindrances, robust multi-class classifiers should be considered.

Multi-class classification is based on assigning labels to instances that belong to a finite set of classes N_c ($N_c > 2$). Designing a machine learning multi-classification technique is a difficult task. In this sense, it is common to conceive algorithms for distinguishing between just two classes, and combine them in some way to form a strong multi-class classifier. Pairwise (one-versus-one) [7] or one-versus-all [8] grouping techniques are the schemes most frequently used. Error correcting output codes (ECOC) were born as an alternative for handling multi-class problems using binary classifiers [9]. It is well-known that ECOC, when applied to multi-class learning problems, can improve the general performance [7]. One of the reasons for this improvement is its property to decompose the original problem into a set of complementary two-class problems—coded in the ECOC matrix—that allows the sharing of classifiers across the original classes.

Recently, there has been a renewed interest in the design of ECOC. The common predesigned coding strategies (one-versus-one and one-versus-all) have been improved with problem-dependent designs [10,11]. Problem-dependent

S. Escalera (✉) · P. Radeva
Department Ciències de la Computació, Computer Vision Center, UAB, 08193 Bellaterra, Spain
e-mail: sescalera@cvc.uab.es

P. Radeva
e-mail: petia@cvc.uab.es

O. Pujol
Department Matemàtica Aplicada i Anàlisi, UB, Gran Via 585, 08007 Barcelona, Spain
e-mail: oriol@cvc.uab.es

designs exploit the knowledge of each particular domain to focus on relevant classifiers for that problem. The selection of the most relevant classifiers allows to obtain a robust strong classifier that requires less dichotomies to split classes. However, a few studies related to the decoding step have been proposed [7].

In this paper, we deal with the problem of traffic sign classification by means of error correcting techniques. We use the information obtained from a mobile mapping system [6] to analyze road video sequences. We use Adaboost with the Haar-like features estimated over the integral image [12] to detect regions with high probability to contain a traffic sign. After applying a spatial normalization and model fitting, we classify the candidate signs in their respective categories. We compare the recently proposed coding strategies in the framework of ECOC, showing the improvement of these techniques when problem-dependent ECOC designs are combined with proper decoding strategies. In this way, two novel decoding strategies are presented to increase the ECOC performance. The proposed ECOC designs robustly classify several types of signs with high variability of appearance, outperforming traditional ECOC designs.

The paper is organized as follows: Sect. 2 overviews the ECOC coding strategies and presents the novel β-correction decoding approaches. Section 3 explains the system for traffic signs classification. Section 4 shows experimental results, and finally, Sect. 5 concludes the paper.

2 Error correcting output codes

The basis of the ECOC framework is to create a codeword for each of the N_c classes. Arranging the codewords as rows of a matrix, we define a "coding matrix" M, where $M \in \{-1, 0, 1\}^{N_c \times n}$ in the ternary case, n being the code length. From the point of view of learning, M is constructed by considering n binary problems (dichotomies), each corresponding to a matrix column. Joining classes in sets, each dichotomy defines a partition of classes (coded by +1, −1, according to their class set membership, or 0 if the class is not considered by the dichotomy).

A ternary coding design is shown in Fig. 1. The matrix is coded using seven dichotomies $\{h_1, \ldots, h_7\}$ for a 4-class problem of codewords $\{y^1, \ldots, y^4\}$. The white regions are coded by +1 (considered as positive for its respective dichotomy, h_i), the dark regions by −1 (considered as negative), and the gray regions correspond to the zero symbol (not considered classes for the current dichotomy). For example, the first classifier is trained to discriminate class 3 versus 1 and 2 without considering class 4, the second one classifies class 2 versus 1, 3, and 4, and so on. Applying the n trained binary classifiers, a code is obtained for each data point in the test set. This code is compared to the base codewords of each class defined in the matrix M, and the data point

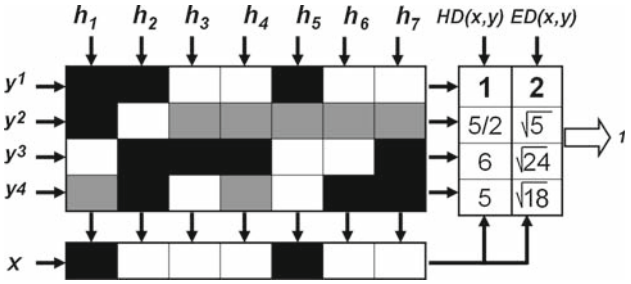

Fig. 1 Example of ternary matrix M for a 4-class problem. A new test codeword x is classified by class one when using the traditional Hamming and Euclidean decoding strategies

is assigned to the class with the "closest" codeword [7]. In the case of the figure, a new test input x is evaluated by all the classifiers and the system assigns the label (in this case, class 1) with the minimum Euclidean decoding distance $ED(x, y^i) = \sqrt{\sum_{j=1}^{n}(x_j - y_j^i)^2}$ and Hamming distance $HD(x, y_i) = \sum_{j=1}^{n}(1 - sign(x_j \cdot y_j^i))/2$, where y is a class codeword, and n is the total number of binary classifiers.

2.1 Coding designs

The traditional coding strategies are: one-versus-all [8], where each learner is trained to distinguish one class from the rest of the classes. Given N_c classes, this technique has a codeword length of N_c. One-versus-one [7] considers all pairs of classes. The codeword length, in this case, is $\frac{N_c(N_c-1)}{2}$. Dense random strategy [7] generates a random coding matrix M, where the values $\{+1, -1\}$ have a certain probability to appear. The sparse random strategy [7] is similar to the dense case, but includes the third symbol 0 with another appearance of probability value. The requirement of the random strategies is that the randomly generated matrix rows and columns should be as different as possible in terms of the Hamming distance. In this way, more classification errors can be corrected [10]. In the work of [7], the authors experimentally proposed the length of each technique: $10\log(N_c)$ for the dense case, and $15\log(N_c)$ for the sparse case.

Due to the high number of binary classifiers involved in the traditional coding strategies and the low robustness of the one-versus-all strategy in comparison with one-versus-one, new coding approaches have been proposed [10,11]. The new techniques are based on exploiting the problem domain by selecting the representative binary classification problems that increase the general performance while keeping the code length small.

2.1.1 Discriminant ECOC

The method in [10] is based on the embedding of discriminant tree structures derived from the problem domain. The binary

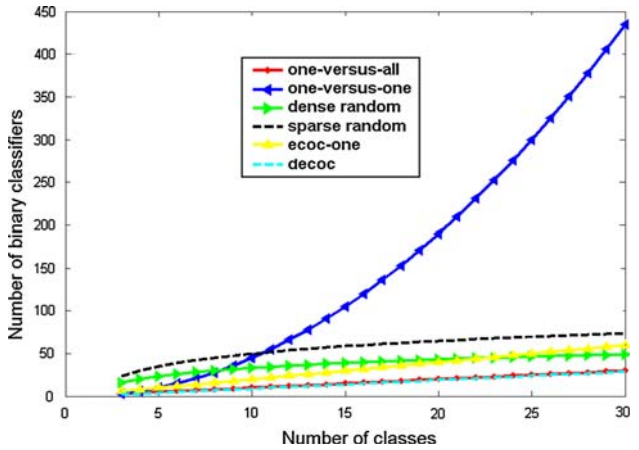

Fig. 2 Number of classifiers required for the coding strategies when the number of classes increases

trees are built by looking for the partition that maximizes the mutual information between the data and their respective class labels. Each node of the tree splits a subset of classes, and each internal node is embedded in the ECOC matrix as a column, coding by +1 the positions that correspond to the classes on the right sub-partitions of the tree, and by -1 the positions corresponding to the left tree sub-partitions of the classes. The length of the codeword is only $(N_c - 1)$.

2.1.2 ECOC-ONE

In our previous work [11], we proposed an extension of any initial ECOC configuration. The method uses a coding process that learns relevant binary problems guided by a validation subset. At each iteration of the algorithm (thus, each new binary classifier), the whole system is evaluated in the training and validation subsets, and the confusion matrix is used to search for the pair of classes with the highest classification error. The next step is to generate an optimal subset of classes containing the two conflictive classes in opposite subsets and update the system by embedding the new classifier weighted by its importance. It has been estimated that $2N_c$ bits are enough for a good performance improvement, since the first selected classifiers are the most discriminant.

Figure 2 shows the cost in terms of the number of binary classifiers required for each of the commented coding strategies. Observe the quadratic behavior of the one-versus-one strategy in contrast with the linear tendency of the rest of the methods when the number of classes increases.

2.2 Decoding designs

The ECOC matrix M uses three possible symbols $M \in \{-1, 0, +1\}$. The zero symbol allows to increase the number of bi-partitions of classes (thus, the number of possible binary classifiers), resulting in a higher number of binary problems to be learnt. However, the effect of the ternary symbol is still an open issue. Since a zero symbol means that the corresponding classifier is not trained on a certain class, to consider the "decision" of this classifier on those zero coded positions does not make sense. Moreover, the response of the classifier on a test sample will always be different from zero, so obligatorily an error will be registered. Let us return to Fig. 1, where an example about the effect of the zero symbol is shown. The classification result using the Hamming distance as well as the Euclidean distance is class 1. Note that class 2 has only the first two positions coded; thus, it is the only information provided about class 2. The first two coded locations of the codeword x correspond exactly to these positions. Thus, the correct classification should be class 2 instead of class 1. The use of standard decoding techniques that do not consider the effect of the third symbol (zero) frequently fail. In the figure, the HD and ED strategies accumulate an error value proportional to the number of zero symbols by row, and finally misclassify the test sample x. To deal with this problem, we propose two novel approaches that increase the performance of the ternary ECOC designs.

2.2.1 Laplacian decoding

We present the simple Laplacian approach to deal with the ternary ECOC decoding. This approach gives to each class a score according to the number of coincidences between the input codeword and the class codeword, normalized by the errors without considering the zero symbol. In this way, the coded positions of the codewords with more zero symbols attain more importance. The decoding score is estimated by:

$$d(x, y^i) = \frac{\alpha_i + 1}{\alpha_i + \beta_i + K} \tag{1}$$

where α_i is the number of coincidences from the test codeword and the codeword for class i, β_i, is the number of failures from the test codeword and the codeword for class i, and K is an integer value that codifies the number of classes considered by the classifier, in this case 2, due to the binary partitions of the base classifiers. The offset $1/K$ is the default value (bias) in case that the coincidences and failures tend to zero.

2.2.2 β-correction

Based on the present discrete Laplacian technique to decode, we define a method, called Pessimistic β-Density Distribution decoding. The method is based on estimating the probability density functions between two codewords. The main goal of this strategy is to model at the same time the accuracy and uncertainty based on a pessimistic score in order to obtain more reliable predictions. We use an extension of the

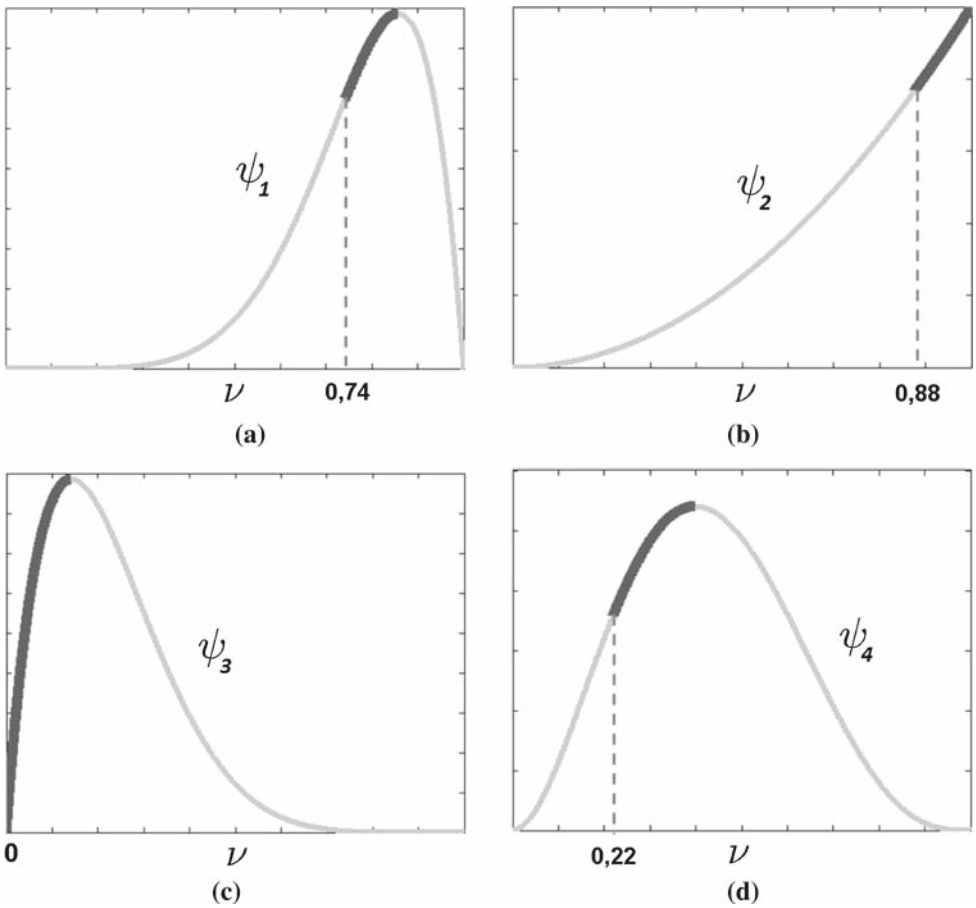

Fig. 3 Pessimistic score decoding for the test codeword x and the matrix M of Fig. 1. **a** Class 1, **b** class 2, **c** class 3, and **d** class 4. The probability for the second class allows a successful classification in this case

continuous binomial distribution, the β-distribution, defined as follows:

$$\psi_i(\nu, \alpha_i, \beta_i) = \frac{1}{K}\nu^{\alpha_i}(1-\nu)^{\beta_i} \tag{2}$$

where ψ_i is the β-Density Distribution between a codeword x and a class codeword y_i for class c_i, and $\nu \in [0, 1]$. The expectation of ψ_i is $\alpha_i/(\alpha_i + \beta_i)$. Note that it asymptotically tends to the Laplace corrected estimator without the prior K in Eq. (1).

Given a test codeword x and the set of functions $\psi(\nu, \alpha, \beta) = [\psi_1(\nu, \alpha_1, \beta_1), \ldots, \psi_N(\nu, \alpha_N, \beta_N)]$, the class c_i is assigned to x if it achieves the highest score s_i, defined as the pessimistic score satisfying the following equivalency:

$$s_i : \int_{v_i - s_i}^{v_i} \psi_i(\nu, \alpha_i, \beta_i)\mathrm{d}\nu = u \tag{3}$$

where u is a threshold parameter. After a preliminary set of experiments, we fixed $u = \frac{1}{3}$. Note that u governs the uncertainty influence in the final score. Figure 3 shows the estimated density functions $[\psi_1, \psi_2, \psi_3, \psi_4]$ for the design shown in Fig. 1. Observe that on the design of Fig. 1, the HD and the ED decoding strategies classify the test codeword x by class 1, although according to the present discussion, the decision should be class 2. In Fig. 3, one can see that the β-DEN decoding classifies the test data sample to its correct class 2, obtained by Fig. 3b. It can be shown that when a function ψ_i is estimated by a combination of values α_i and β_i, the sharpness is higher when it is generated by a majority of one of the two types. Moreover, this sharpness depends on the number of code positions different from zero and the balance between the number of matches and failures. In this way, the pessimistic score reflects the confidence in the expectation of the probability density function.

3 Traffic sign classification system

We focus on the goal of mobile mapping to compile cartographic information from a mobile system. In particular, we use the video sequences obtained from the mobile mapping system of [6]. In this system, the position and orientation of the different traffic signs are measured in movement with the car video cameras. The system has a stereo pair of calibrated cameras, which are synchronized with a GPS/INS system (see Fig. 6a). Therefore, the result of the acquisition step is a set of stereo-pairs of images with their position and orientation information.

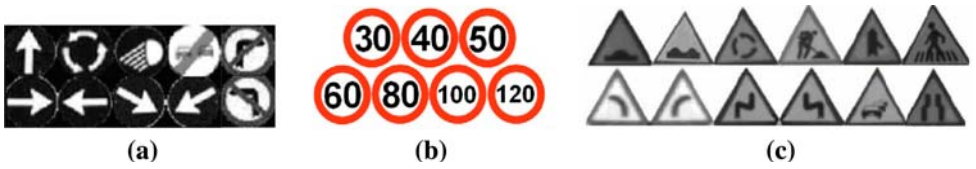

Fig. 4 **a** Circular classes, **b** speed classes, **c** triangular classes

Fig. 5 **a**, **b** Road conditions from the mobile mapping process video camera. **c** Speed sign samples under different conditions

From ten analyzed DVD video sequences, we obtained the classes of Fig. 4. The classes are divided into three main groups: speed, circular, and triangular, with a total of 27 different classes to recognize. The speed signs are treated as a special case due to their similarity and difficulty to discriminate in adverse conditions. Different road frames acquired from the video system where the illumination dramatically changes are shown in Fig. 5a,b.

The traffic sign recognition system used is divided into four main steps: object detection, model fitting, normalization, and classification. Each of these steps must be robust enough to minimize the propagation of errors in the system.

3.1 Detection

The detection process is based on the face detector presented by Viola and Jones in [12]. In particular, we use the Gentle version of Adaboost with decision stumps [13]. The weak classifiers are trained using the attentional cascade based on the extended set of Haar-like features (i.e., including the rotated ones) estimated on the integral image [12]. As a result of the detection process, we obtain robust and fast detection results, as shown in Fig. 6b [14].

The three attentional cascades (one for each group) were trained using a total of 1,500 positive samples divided into the three different groups.

Given an image where the Adaboost learning algorithm detected a road sign, a region of interest (ROI) that contains a sign is determined (circular or triangular). However, since we have missing information about sign scale and position, before the recognition process we apply a spatial normalization to improve final recognition.

3.2 Model fitting

The Hough transform [15] and fast radial symmetry [16] are applied in order to fit the model since they offer great robustness against noise.

3.2.1 Fast radial symmetry

The fast radial symmetry is calculated over a set of one or more ranges, depending on the scale of the features one is trying to detect. The value of the transform at a range indicates the contribution to radial symmetry of the gradients at a distance d away from each point. At each range d, we examine the gradient g at each point p, from which a corresponding positively-affected pixel, $p_{+ve}(p)$, and negatively-affected pixel, $p_{-ve}(p)$, are determined: $P_{\pm ve}(p) = p \pm \text{round}\frac{g(p)}{\|g(p)\|}d$, and accumulated in the orientation projection image O_d: $O_d(P_{\pm ve}(p)) = O_d(P_{\pm ve}(p)) + 1$. Now, to locate the center of radial symmetry, we search for the position (x, y) of maximal value in the accumulated orientations matrix $O^T = \sum_{i=1}^{d} O_d$. Locating that maximum, we determine the radius length. This procedure allows to obtain robust results fitting circular traffic signs.

3.2.2 Hough transform

The Hough transform has been shown to allow the detection of straight lines in a robust way. We apply this procedure in order to look for the three representative lines of the triangular sign and calculate their intersections to transform the image. However, we need to consider additional restrictions to obtain the three representative border lines of a triangular traffic sign. Each line has associated a position in relation to the others. Once we have the three detected lines, we calculate their intersection. Given the parameters a and b that define the equation $y = a \cdot x + b$ for each of the three lines, the intersection point (X, Y) for each pair of lines is defined as follows:

$$X_t = (b_2^i - b_1^i)/(a_1^i - a_2^i),$$
$$Y_t = a_1^i X_t + b_1^i \mid t, \quad i \in [1, \ldots, 3] \tag{4}$$

To assure that the lines are the expected ones, we complement the procedure searching for a corner at a circular region at

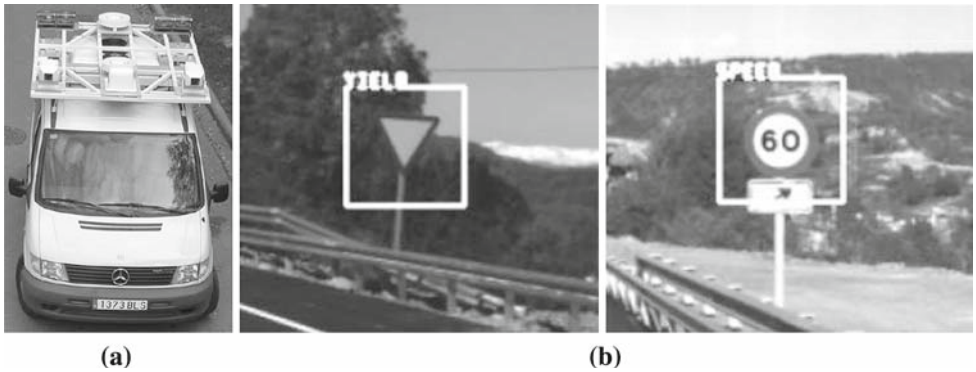

Fig. 6 **a** Geovan, **b** detected traffic signs

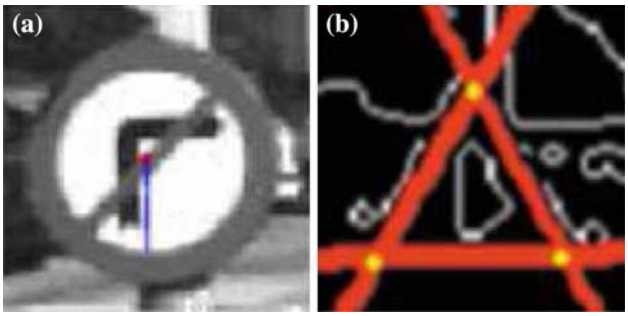

Fig. 7 **a** Fitting circular traffic signs, **b** fitting triangular traffic signs

each intersection neighborhood:

$$S = \{(x_i, y_i) \,|\, \exists p < ((x - x_i)^2 + (y - y_i)^2 - r^2)\} \,|\, i \in [1, \ldots, 3] \tag{5}$$

where S is the set of valid intersection points, and p corresponds to a corner point to be located in a neighborhood of the intersection point. In Fig. 7, an example of two detected sign regions and a circular and triangular traffic signs are fitted using the fast radial symmetry [16] and the Hough transform [15], respectively.

3.2.3 Normalization

Once the sign model is fitted using the commented methods, the next step is to normalize the fitted object before classification. The steps are: transform the image to make the recognition invariant to small affine deformations re-scaling to the signs database size (32×32 pixels), filter with the Weickert anisotropic filter [17], and mask the image to exclude background at the classification step. To prevent the effects of illumination changes, the histogram equalization improves image contrast and obtains a uniform histogram. An example of the normalization process applied to a detected circular sign is shown in Fig. 8. The image of Fig. 8a corresponds to a detected, fitted, and re-scaled circular sign. Note the poor resolution of the detected object. Figure 8b shows the histogram equalization from the previous image. In Fig. 8c, anisotropic Weickert filtering allows to homogenize regions affected by noise and slight illumination changes. Finally, Fig. 8d is obtained after applying a circular mask in order to reject background regions at the classification step. This process is repeated for each detected sign, applying a circular or triangular model mask depending on the group of the object.

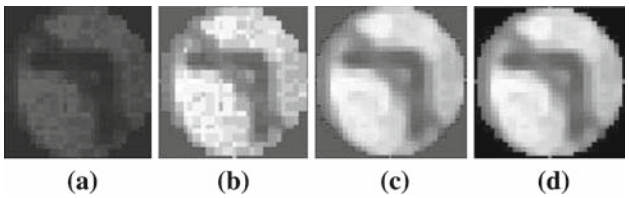

Fig. 8 **a** Fitted sign, **b** histogram equalization, **c** Weickert anisotropic filtering, and **d** masked region

3.3 Classification

Applying the three attentional cascades in the mobile mapping system video sequences, the detected and normalized regions are classified, depending on the type of the detected sign, using different classification strategies combining the coding and decoding strategies of ECOC presented in the previous chapter. A scheme of the whole system is shown in Fig. 9.

4 Experimental results

Before the results are presented, we discuss the data, comparatives, measurements, and experiments.

- *Data*: The data consists of 15,000 road frames at different conditions obtained by the mobile mapping system.
- *Comparatives*: The strategies used to validate the classification are 50 runs of Gentle Adaboost with decision stumps and Radial Basis Function SVM with the parameter gamma set to 1. These two classifiers generate the set of binary problems to embed in the set of ECOC configurations: one-versus-one, one-versus-all, dense-random, and the recently

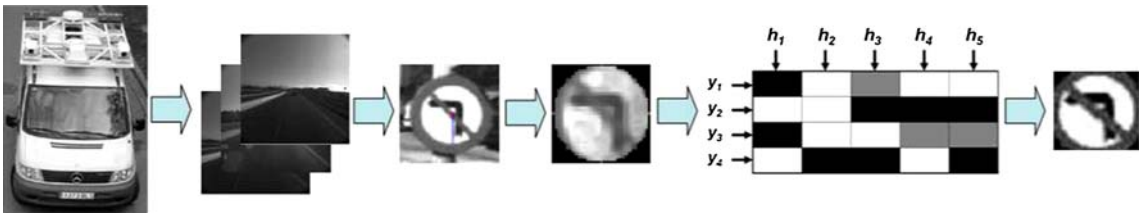

Fig. 9 Traffic sign recognition scheme. From *left* to *right* Geovan system, captured frames, detected sign, normalized sign, ECOC categorization, and final decision (recovered sign from the sign database)

proposed DECOC, and ECOC-ONE. Each of the ECOC strategies is evaluated using different decoding strategies: the traditional Euclidean distance, and the novel Laplacian and β-correction decoding. The number of classifiers used for each methodology are: $N_c(N_c - 1)/2$ for one-versus-one, N_c for one-versus-all, N_c for dense random [7], $N_c - 1$ for DECOC, and $2N_c$ for ECOC-ONE. In the dense random case, the coding matrix was selected from a set of 20,000 generated random matrix of N_c binary classifiers, which provides a fair comparison between one-versus-all and DECOC designs in terms of a similar number of binary problems.

- *Measurements*: The classification tests are performed using stratified tenfold cross-validation with a two-tailed t test at 95% confidence interval.
- *Experiments*: First, we evaluate the detection rate of the mobile mapping system. Then, based on the detected regions of interest, we perform the classification of traffic sign classes. Two feature sets are considered: first, we used the normalized pixel-values, and second, we apply the SIFT descriptor [18] over the set of normalized signs. These results are compared to the ones obtained by the multi-class built-in SVM. Moreover, public multi-class data sets from the UCI Machine Learning Repository [19] have been used to test the classification methodology.

4.1 System detection results

For each of the three groups of classes shown in Fig. 4, an attentional cascade using 500 positive samples and a random set of background images was trained. Each cascade was trained using 50 weak classifiers per level, rejecting 80% of negative samples per level, and learning a total of 12 levels per cascade [12,13].

Applying the attentional cascades over a test set of 15,000 road frames obtained from the mobile mapping system, we detected 1,104 regions that contain traffic signs from a set of 1,119 (thus, a mean detection accuracy of 98.66%). The detection rate of each particular group is shown in Table 1.

From the previously detected regions, where the minimum size of detected sign corresponds to 24×24 pixels resolution, we applied the model fitting and the spatial normalization explained in the previous sections. Then, two feature sets are designed to perform the final classification: first, all pixels are recovered as a 1,024 feature vector from the data set size (32×32 pixels resolution), and second, as a 128 SIFT feature vector.

Table 1 Detection rate using an attentional cascade

Problem	Detection rate (%)
Speed	98.20
Circular	97.98
Triangular	99.80
Mean performance	98.66

4.2 Error correcting classification using normalized pixel-based features

First, we generate three types of classification experiments using the normalized pixel-based features, each one considering 500 samples from each of the three different traffic signs groups. The experiment results for the circular and triangular groups are shown in Fig. 10a, b, respectively. The results in these cases are very similar due to the high discriminability of the classes. However, one-versus-one, DECOC, and ECOC-ONE coding strategies obtain the best performance. For the one-versus-one, the results are premised because of the high number of binary classifiers used. In the case of the speed signs, the classes are very similar and sensible to classification errors. This fact is magnified by the road conditions and the video camera resolution (see Fig. 5). In this case, the results of Fig. 10c show higher differences between the classification strategies. The best performance for SVM is obtained by one-versus-one and ECOC-ONE strategies using the Laplacian and β-density decoding techniques. For Gentle Adaboost, ECOC-ONE obtains the first position. This fact is related to the error correction capability and the problem-dependent designs of the ECOC-ONE and DECOC strategies, which exploit the speed domain to focus on difficult classes to split.

Concerning the decoding strategies, one can see that for a same base classifier and coding strategy, the Laplacian and β-density decodings improve the traditional Euclidean

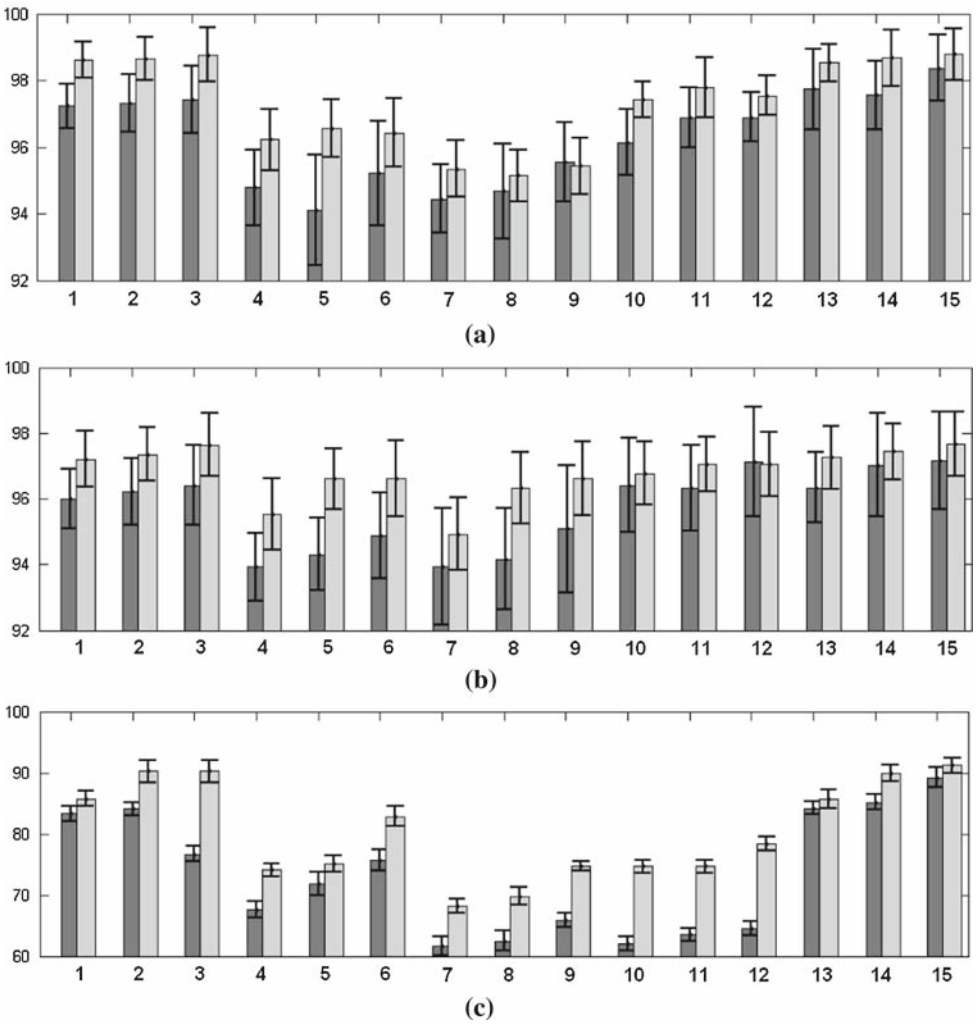

Fig. 10 Classification results using pixel-based features for circular **a**, triangular **b**, and speed **c** classes using Gentle Adaboost (*dark bars*) and $RBF\ SVM$ (*light bars*) in the coding and decoding strategies, respectively. From *left* to *right* *1* Euclidean one-versus-one, *2* Laplacian one-versus-one, *3* β-density one-versus-one, *4* Euclidean one-versus-all, *5* Laplacian one-versus-all, *6* β-density one-versus-all, *7* Euclidean dense random, *8* Laplacian dense random, *9* β-density dense random, *10* Euclidean DECOC, *11* Laplacian DECOC, *12* β-density DECOC, *13* Euclidean ECOC-ONE, *14* Laplacian ECOC-ONE, and *15* β-density ECOC-ONE

distance, and in general, they increase the classification performance of any ECOC design. In particular, β-density decoding attains the best positions, and the percentage of improvement is more relevant when applied to third symbol-based ECOC.

The mean rankings for each classification strategy using the results of the three presented experiments are shown in Fig. 11. The rankings are obtained by estimating each particular ranking r_i^j for each problem i and each ECOC configuration j, and computing the mean ranking R for each ECOC design as $R_j = \frac{1}{J}\sum_i r_i^j$, where J is the total number of problems (three experiments). One can observe that the best position is obtained by the ECOC-ONE strategy, followed by one-versus-one, DECOC, one-versus-all, and finally dense random strategy. Moreover, note that for each ECOC design, the Laplacian, and β-density in particular, increase the classification accuracy of Euclidean decoding for all the cases, as claimed.

4.3 Error correcting classification using SIFT-based features

Due to the similarity of shape appearance of the signs from the same class, we also used the SIFT descriptor [18] to compute the feature space of traffic sign classes. The SIFT descriptor has shown to be very useful to describe image regions in real applications. In our problem, we apply the SIFT descriptor on the normalized signs to project them into a feature vector of 128 features based on the inner orientations of the object. The same experiments as in the previous section have been performed with the new feature set. The experimental results for the circular, triangular, and speed groups are shown in Fig. 12a, b, c, respectively. The behavior of the different classification strategies is similar to using the normalized pixel-based features. The main difference is that the ECOC configurations that use Adaboost as the base classifier tend to slightly increase the classification performance,

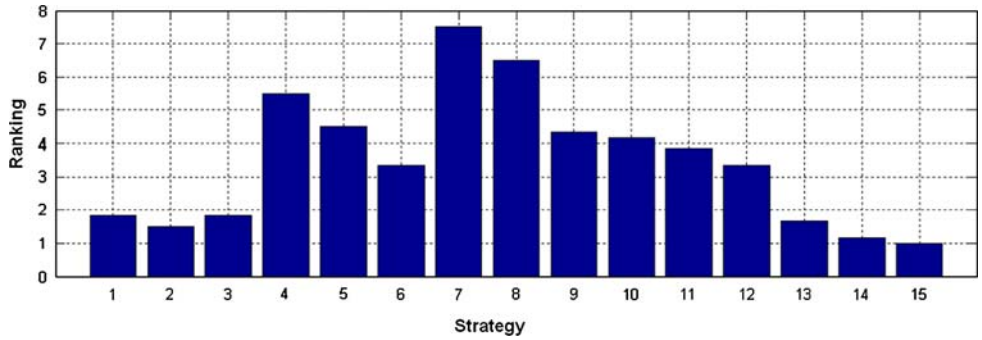

Fig. 11 Ranking position for each classification strategy using pixel-based features. From *left* to *right* *1* Euclidean one-versus-one, *2* Laplacian one-versus-one, *3* β-density one-versus-one, *4* Euclidean one-versus-all, *5* Laplacian one-versus-all, *6* β-density one-versus-all, *7* Euclidean dense random, *8* Laplacian dense random, *9* β-density dense random, *10* Euclidean DECOC, *11* Laplacian DECOC, *12* β-density DECOC, *13* Euclidean ECOC-ONE, *14* Laplacian ECOC-ONE, *15* β-density ECOC-ONE

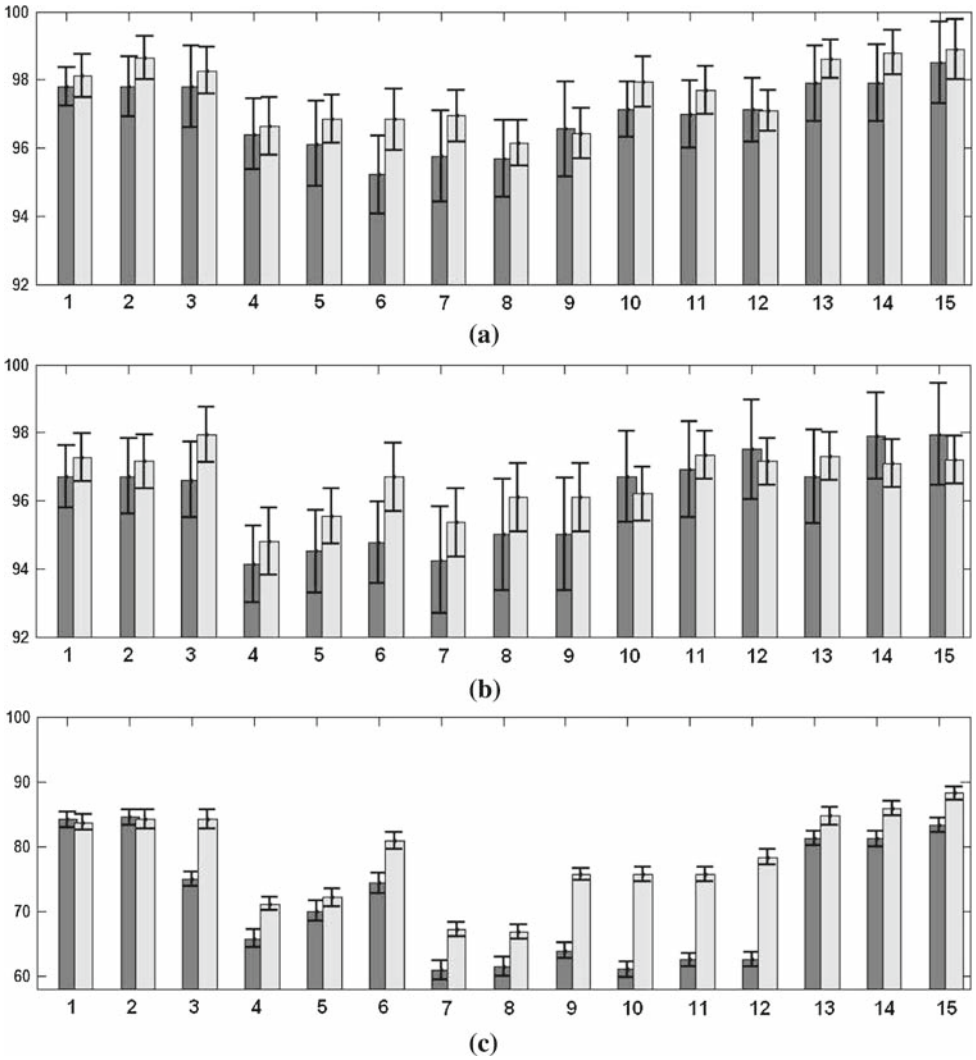

Fig. 12 Classification results using SIFT-based features for circular **a**, triangular **b**, and speed **c** classes using Gentle Adaboost (*dark bars*) and $RBF\ SVM$ (*light bars*) in the coding and decoding strategies, respectively. From *left* to *right* *1* Euclidean one-versus-one, *2* Laplacian one-versus-one, *3* β-density one-versus-one, *4* Euclidean one-versus-all, *5* Laplacian one-versus-all, *6* β-density one-versus-all, *7* Euclidean dense random, *8* Laplacian dense random, *9* β-density dense random, *10* Euclidean DECOC, *11* Laplacian DECOC, *12* β-density DECOC, *13* Euclidean ECOC-ONE, *14* Laplacian ECOC-ONE, and *15* β-density ECOC-ONE

whereas when using $RBF\ SVM$ as the base classifier, the obtained performance remains very similar.

Concerning the decoding strategies, one can see that for the same base classifier and coding strategy, the Laplacian and β-density decodings also improve the traditional Euclidean distance using SIFT-based features.

The mean rankings for each classification strategy using the results of the three presented experiments are shown in

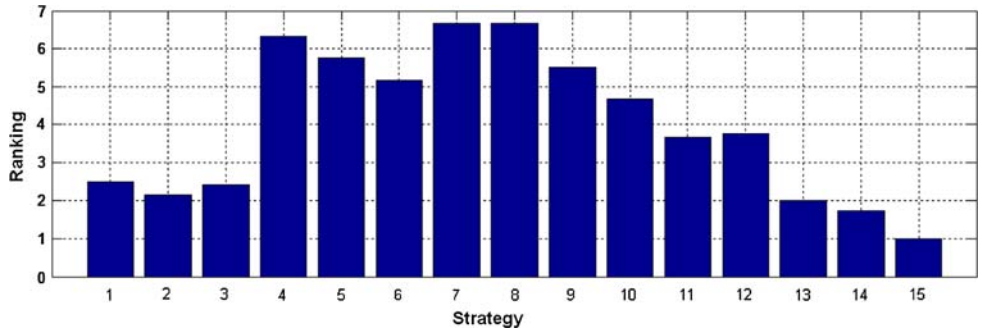

Fig. 13 Ranking position for each classification strategy using SIFT-based features. From *left* to *right* *1* Euclidean one-versus-one, *2* Laplacian one-versus-one, *3* β-density one-versus-one, *4* Euclidean one-versus-all, *5* Laplacian one-versus-all, *6* β-density one-versus-all, *7* Euclidean dense random, *8* Laplacian dense random, *9* β-density dense random, *10* Euclidean DECOC, *11* Laplacian DECOC, *12* β-density DECOC, *13* Euclidean ECOC-ONE, *14* Laplacian ECOC-ONE, *15* β-density ECOC-ONE

Table 2 UCI Repository data sets characteristics

Problem	#Train	#Attributes	#Classes	Problem	#Train	#Attributes	#Classes
Dermatology	366	34	6	Yeast	1,484	8	10
Vowel	990	10	11	Letter	20,000	16	26

Table 3 Dermatology performance using Gentle Adaboost

	One-versus-one	One-versus-all	Dense	DECOC	ECOC-ONE
ED	92.04 (2.32)	89.37 (1.89)	91.04 (2.37)	92.04 (2.19)	92.04 (2.32)
LAP	92.04 (2.32)	89.37 (1.89)	91.04 (2.37)	92.04 (2.20)	92.04 (2.20)
$\beta - DEN$	92.04 (2.32)	89.37 (1.89)	91.04 (2.37)	92.04 (2.04)	92.04 (2.11)

Table 4 Vowel performance using Gentle Adaboost

	One-versus-one	One-versus-all	Dense	DECOC	ECOC-ONE
ED	59.19 (2.83)	42.42 (2.28)	27.47 (2.07)	62.83 (2.62)	62.56 (2.87)
LAP	59.19 (2.83)	42.42 (2.28)	27.47 (2.07)	64.91 (2.68)	65.36 (2.17)
$\beta - DEN$	59.19 (2.83)	42.42 (2.28)	27.47 (2.07)	65.12 (2.62)	65.36 (2.17)

Fig. 13. One can observe that the rank positions are equivalent to those obtained using normalized pixel-based features. The best position is obtained by the ECOC-ONE strategy, followed by one-versus-one, DECOC, one-versus-all, and finally, dense random strategy. Moreover, note that for each ECOC design, the Laplacian and β-density, as in the previous experiment, increase the classification accuracy of Euclidean decoding.

4.4 Public UCI machine learning repository classification

In this experiment, we classify four multi-class data sets from the UCI Machine Learning Repository [19]. The details of the data sets are shown in Table 2. The results applying the different ECOC configurations with Gentle Adaboost and *RBF SVM* are shown in Tables 3, 4, 5, 6 and 7, 8, 9, 10, respectively. The best performance of each ECOC configuration is shown in bold. Note that the Laplacian and β-density in the worst case obtain the same results as applying the traditional Euclidean distance to decode.

4.5 Multi-class RBF comparison

To show the robustness of the presented classification framework, we compare the results obtained with the ECOC methods with a built-in multi-class SVM with RBF. The results are shown in Fig. 14a for pixel-based features and Fig. 14b for SIFT-based features. One can observe that the RBF multi-class SVM obtains inferior results to the ones obtained by one-versus-one, ECOC-ONE, and DECOC designs, and similar to one-versus-all and dense random strategies for the same types of features.

Table 5 Yeast performance using Gentle Adaboost

	One-versus-one	One-versus-all	Dense	DECOC	ECOC-ONE
ED	49.57 (1.38)	45.87 (1.12)	46.84 (1.34)	50.32 (1.44)	50.88 (2.16)
LAP	49.57 (1.38)	45.87 (1.12)	46.84 (1.34)	51.77 (1.35)	52.04 (1.38)
$\beta - DEN$	49.57 (1.38)	45.87 (1.12)	46.84 (1.34)	51.79 (1.37)	52.04 (1.55)

Table 6 Letter performance using Gentle Adaboost

	One-versus-one	One-versus-all	Dense	DECOC	ECOC-ONE
ED	88.96 (1.56)	84.34 (1.65)	84.53 (1.67)	87.56 (1.67)	88.85 (1.58)
LAP	88.96 (1.64)	86.89 (1.63)	85.73 (1.76)	88.76 (1.59)	90.12 (1.81)
$\beta - DEN$	88.96 (1.64)	87.12 (1.60)	88.26 (1.50)	89.01 (1.54)	90.32 (1.58)

Table 7 Dermatology performance using $RBF\ SVM$

	One-versus-one	One-versus-all	Dense	DECOC	ECOC-ONE
ED	95.59 (0.74)	94.54 (1.04)	80.86 (1.26)	95.07 (1.02)	95.52 (0.94)
LAP	95.59 (0.74)	94.54 (1.04)	80.86 (1.26)	96.10 (0.94)	95.59 (1.00)
$\beta - DEN$	95.59 (0.74)	94.54 (1.04)	80.86 (1.26)	96.10 (0.94)	96.10 (0.83)

Table 8 Vowel performance using $RBF\ SVM$

	One-versus-one	One-versus-all	Dense	DECOC	ECOC-ONE
ED	64.95 (3.71)	26.67 (2.11)	28.18 (3.16)	66.78 (2.67)	66.90 (2.73)
LAP	64.95 (3.71)	26.67 (2.11)	28.18 (3.16)	68.36 (3.02)	68.40 (2.94)
$\beta - DEN$	64.95 (3.71)	26.67 (2.11)	28.18 (3.16)	68.36 (3.08)	68.53 (3.13)

Table 9 Yeast performance using $RBF\ SVM$

	One-versus-one	One-versus-all	Dense	DECOC	ECOC-ONE
ED	50.79 (2.39)	35.52 (1.00)	27.82 (1.59)	50.79 (2.48)	51.04 (2.51)
LAP	50.79 (2.39)	35.52 (1.00)	27.82 (1.59)	50.79 (2.21)	52.20 (2.44)
$\beta - DEN$	50.79 (2.39)	35.52 (1.00)	27.82 (1.59)	50.79 (2.28)	52.34 (2.46)

Table 10 Letter performance using $RBF\ SVM$

	One-versus-one	One-versus-all	Dense	DECOC	ECOC-ONE
ED	86.11 (0.99)	36.38 (0.76)	67.28 (0.81)	82.09 (0.81)	86.25 (0.86)
LAP	86.22 (1.00)	36.38 (0.76)	68.73 (1.01)	85.15 (0.96)	88.89 (1.05)
$\beta - DEN$	86.47 (0.92)	36.38 (0.76)	70.37 (0.97)	85.62 (0.89)	89.03 (0.94)

4.6 Discussion

The multi-class variant of Adaboost that has demonstrated to be dominant to the other proposals in empirical studies is the Adaboost.MH [20]. The Adaboost.MH algorithm converts the N_c-class problem into that of estimating a two-class classifier on a training set N_c times as large, with an additional feature defined by the set of class labels [21]. It is the same as the one-versus-all scheme, representing a "Multi-label Hamming" to measure the separate classifiers, being essentially the one-versus-all ECOC with Hamming decoding, that, in comparison to ours, offers low performance [13]. For this reason, the comparison with multi-class Adaboost.MH has been omitted from the set of experiments.

The presented multi-classification system is robust, dealing with a high number of classes to distinguish. The combination of the recently proposed problem-dependent coding designs and the novel decoding strategies outperform the traditional schemes of error correction and also increase the performance of the traditional multi-class approaches, as the multi-class Adaboost and SVM.

We are currently performing the detection and object fitting steps, since the classification performance of the present application depends on the accuracy of these first steps of the system.

The presented multi-classification approach represents a powerful tool to be used on any application that requires distinguishing between a set of categories.

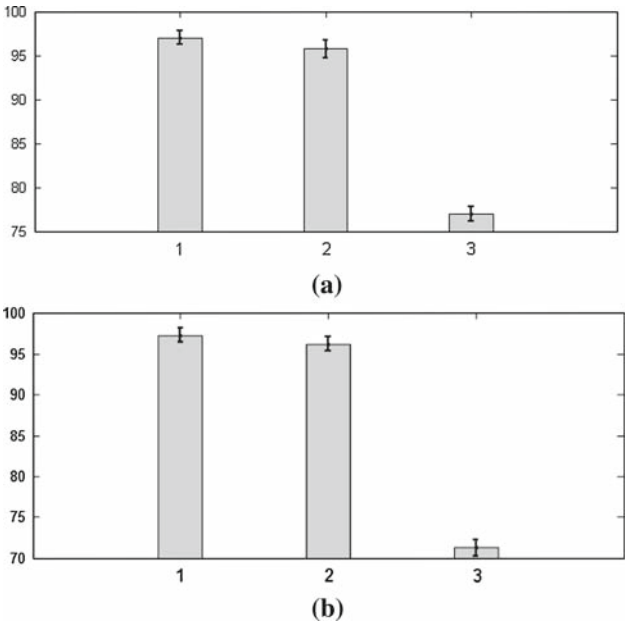

Fig. 14 Classification results using multi-class *SVM* for pixel-based features **a** and SIFT-based features **b**. From *left* to *right 1* Circular, *2* triangular, and *3* speed classification

5 Conclusions

In this paper, we presented a classification system that obtains a very high performance for the problem of traffic sign classification using error correcting techniques. The system has four main stages: traffic sign detection, model fitting, spatial normalization, and sign categorization. The multi-class classification techniques are evaluated on real video sequences obtained from a mobile mapping system. We compared the state-of-the-art and recently proposed designs for ECOC. Moreover, we presented two novel decoding techniques, which showed high robustness and better performance than traditional ECOC designs and the state-of-the-art multi-classifiers. The traffic sign recognition system obtains robust classification results in front, with a high number of classes and high variability of the objects' appearance.

Acknowledgments This work was supported in part by the projects, FIS-G03/1085, FIS-PI031488, MI-1509/2005, TIN2006-15308-C02-01, TIN2006-15308-C02-02, and the Institut Cartogràfic de Catalunya.

References

1. Piccioli, G., Micheli, E., Campani, M.: A robust method for road sign detection and recognition. ECCV **1**, 495–500 (1996)
2. Loy, G., Zelinsky, A.: Fast radial symmetry for detecting points of interest. IEEE Trans Pattern Analysis and Machine Intelligence, 8 (2003)
3. Shaposhnikov, D., Podladchikova, L., Golovan, A., Shevtsova, N., Hong, K., Gao, X.: Road sign recognition by single positioning of space-variant sensor (2002)
4. Hsu, S., Huang, C.: Road sign detection and recognition using matching pursuit method. Image Vis. Comput. **19**, 119–129 (2001)
5. Handmann, U., Kalinke, T., Tzomakas, C., Werner M., von Seelen, W.: An image processing system for driver assistance. IEEE International Conference on Intelligent Vehicles, pp. 481–486 (1998)
6. Casacuberta, J., Miranda, J., Pla, M., Sanchez, S., Serra, A., Talaya, J.: On the accuracy and performance of the geomobil system. Society for Photogrammetry and Remote Sensing (2004)
7. Allwein, E., Schapire, R., Singer, Y.: Reducing multiclass to binary: a unifying approach for margin classifiers. J. Mach. Learn. Res. **1**, 113–141 (2002)
8. Nilsson, N.: Learning Machines. McGraw-Hill, New York (1965)
9. Dietterich, T., Bakiri, G.: Solving multiclass learning problems via error-correcting output codes. J. Artif. Intell. Res. **2**, 263–286 (1995)
10. Pujol, O., Radeva, P., Vitrià, J.: Discriminant ECOC: a heuristic method for application dependent design of error correcting output codes. IEEE Trans. Pattern Anal. Mach. Intell. **28**, 1007–1012 (2006)
11. Escalera, S., Pujol, O., Radeva, P.: ECOC-ONE: a novel coding and decoding strategy. Int. Conf. Pattern Recogn. **3**, 578–581 (2006)
12. Viola, P., Jones, M.: Robust real-time object detection. Int. J. Comput. Vis. (2002)
13. Friedman, J., Hastie, T., Tibshirani, R.: Additive logistic regression: a statistical view of boosting. Technical Report (1998)
14. Baro, X., and Vitrià, J.: Traffic sign detection on greyscale images. CCIA, pp. 209–216 (2004)
15. Morse, B.: Segmentation (edge based, hough transform). Technical report (2000)
16. Loy, G., Zelinsky, A.: Fast radial symmetry for detecting points of interest. IEEE Trans. Pattern Anal. Mach. Intell. **25** (2003)
17. Weickert, J.: Anisotropic diffusion in image processing. European Consortium for Mathematics in Industry. B.G. Teubner, Stuttgart (1998)
18. Lowe, D.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. **20**, 91–110 (2003)
19. Asuncion, A., Newman, D.J.: UCI Machine Learning Repository [http://www.ics.uci.edu/~mlearn/MLRepository.html]. University of California, Department of Information and Computer Science, Irvine (2007)
20. Schapire, R., Singer, Y.: Improved boosting algorithms using confidence-rated prediction. Mach. Learn. **37**(3), 297–336 (1999)
21. Zhu, J., Rosset, S., Zou, H., Hastie, T.: Multi-class Adaboost. A multiclass generalization of the Adaboost algorithm, based on a generalization of the exponential loss (2005)
22. Rifkin, R., Klautau, A.: In defense of one-vs-all classification. J. Mach. Learn. Res. **5**, 101–141 (2004)

Author biographies

Sergio Escalera received the BS and MS degrees from Universitat Autonoma de Barcelona in 2003 and 2005, respectively. He is currently working toward the PhD degree in Computer Science. His research interests include machine learning and object recognition.

Oriol Pujol received the PhD degree in Computer Science from Universitat Autònoma de Barcelona in 2004. Currently, he is a lecturer at Universitat de Barcelona. His main research interest includes basic statistical machine learning techniques for object recognition and medical imaging analysis.

Petia Radeva has received her PhD at UAB on the development of physics-based models applied to image analysis. Currently, Petia Radeva is an associate professor in the Computer Science Department of the Universitat Autònoma de Barcelona. Her present research interest is concentrated on development of physics-based and statistical approaches for object recognition, medical image analysis and industrial vision.