Invited Paper

Strategies for detection of distorted road signs in background noise

Elisabet Pérez[*,a], María-Albertina Castro[**,b], Bahram Javidi[***,b]
[a]Optics and Optometry Dept., Universitat Politècnica de Catalunya, Spain;
[b]Electrical and Computing Engineering Dept., University of Connecticut, USA;

ABSTRACT

Design of an on-board processor that enables recognition of a given road sign affected by different distortions is presented. The road sign recognition system is based on a nonlinear processor. Analysis of different filtering methods allows us to select the best techniques to overcome a variety of distortions. The proposed recognition system has been tested in real still images as well as in video sequences. Scenes were captured in real environments, with cluttered backgrounds and contain many distortions simultaneously. Recognition results for various images show that the processor is able to properly detect a given road sign even if it is varying in scale, slightly tilted or viewed under different angles. Recognition is also achieved when dealing with partially occluded road signs. In addition, the system is robust to illumination fluctuations.

Keywords: Pattern Recognition, Distortion-Tolerant Processor, Road Sign Detection.

1. INTRODUCTION

A distortion-tolerant processor, which enables recognition of road signs, is obtained. A safety system to be installed in vehicles could be based on this processor in order to automatically detect and identify road signs. Afterwards, the recognition system could make an objective decision according to the information detected. One of the greatest difficulties on achieving this goal lies on the number of different distortions that may simultaneously modify the reference sign. Variations in scale, in-plane and out-of-plane rotations, background clutter, partially occluded signs, variable illumination, are some examples of distortions that can affect road signs.

Different contributions have been proposed for distortion-invariant systems in the field of pattern recognition. One of the most used techniques is the synthesis of composite filters, also called synthetic discriminant functions (SDF). They have been first introduced by Caulfield and Maloney[1,2] and developed by the group of Casasent.[3,4] Since then, other proposals of this category of filters have been made. For instance, minimum average correlation energy (MACE) filters are composite filters that improve the detection of the target by minimizing the average correlation plane energy, which implies that sidelobes are minimized as much as possible.[5] The aforementioned composite filters[3,5] have their equivalent for a nonlinear processor. A rotation-invariant system has been achieved by the use of these synthetic discriminant functions in a nonlinear joint transform correlator.[6,7] Nonlinear processors have also been studied in the case of an illuminant-invariant system.[8] Design for a recognition system tolerant to various distortions has also been proposed[9], and special attention has been paid to distortion-invariant systems under noisy environments.[10-13]

Pattern recognition techniques that deal with different types of distortions have also been introduced for a road sign recognition system. A monochromatic optical correlator for scale-invariant road sign detection was proposed[14] and its performance was analyzed and compared with the binary phase-only filter (POF).[15] White-light illumination has been shown as a useful tool to achieve scale-invariance within a limited range for road sign detection.[16,17] Partially invariant filters were applied in a multiple correlator to obtain road sign recognition with partial tolerance to in-plane rotations and scale-invariance.[18,19]

[*] eperez@oo.upc.es; phone 34 93 739 83 39; fax 34 93 739 83 01; Optics and Optometry Dept., Universitat Politècnica de Catalunya, Violinista Vellsolà, 37 - 08222 Terrassa (Spain).
[**] betina@engr.uconn.edu; phone 1 860 486 26 70; fax 1 860 486 24 47; Electrical and Computing Engineering. Dept., University of Connecticut, 260 Glenbrook Road, U-157, Storrs, CT 06269-2157 (USA).
[***] bahram@engr.uconn.edu; phone 1 860 486 28 67; fax 1 860 486 24 47; Electrical and Computing Engineering Dept., University of Connecticut, 260 Glenbrook Road, U-157, Storrs, CT 06269-2157 (USA).

In general, a given recognition technique is designed to provide satisfactory results when dealing with a particular distortion of the object. However, the same strategy usually gives poorer results if another type of distortion affects the object. Analysis and comparison of different techniques are carried out in this work. A recognition system simultaneously scale-invariant and tolerant to slight tilts or out-of-plane rotations due to different view angles of the acquisition system is obtained by combining various strategies. Tolerance to illumination fluctuations is needed in order to enable a recognition system to work under different illumination or weather conditions. Robustness to cluttered background is also important for a road sign recognition processor that analyzes images captured in real environments.

2. DISTORTION-TOLERANT RECOGNITION SYSTEM

In this section, a description of the proposed processor is provided. First of all, a brief introduction to nonlinear processors is given paying special attention to the k'th-law processor, which is used in the following experiments. Afterwards, two methods that allow obtaining distortion-tolerance are presented and compared: a bank of single nonlinear filters and composite nonlinear filters. Then, definition of peak-to-correlation energy (*PCE*) criterion is given in subsection 2.3 as it is used as a metric for performance evaluation. In subsection 2.4 we introduce a set of scenes that are used to train the recognition system and to test its performance. Finally, the strength of the best-suited nonlinearity for the nonlinear processing is determined.

2.1 Nonlinear processor

The proposed recognition system is based on a nonlinear processor.[20] In this work, nonlinear filtering is being used in the correlator due to its superior performance in comparison with linear filtering techniques in terms of discrimination capability, correlation peak sharpness, and noise robustness.[6-7,20] In a k'th-law processor,[20] the nonlinear operator applied symmetrically to the scene and to the reference Fourier transforms is defined by

$$g(\hat{f}) = \mathrm{sgn}(\hat{f})\left|\hat{f}\right|^{k} \exp\left[i\phi_{\hat{f}}\right] \qquad k \le 1, \tag{1}$$

where symbol ^ denotes Fourier transform, and parameter k controls the strength of the applied nonlinearity. For k=1 a linear filtering technique is obtained, whereas k=0 leads to a binarizing nonlinearity. Intermediate values of k permit to vary some features of the processor, such as its discrimination capability or its illumination-invariance. Precise index k needs to be determined to obtain a good performance of the processor depending on the application.

2.2 Bank of single nonlinear filters versus composite nonlinear filters

Different approaches to obtain distortion-tolerant recognition systems have been developed. They have in common the need of storing information of the reference target taking into account different distortions that can affect it. The most straightforward way to keep the information of the distorted versions of a target is to design a single filter for each type of distortion to be considered, which implies a filter bank. To determine if a target, distorted or not, is present in a given scene, it will be necessary to correlate or compare the scene each filter that belongs to the bank. This technique could be time-consuming and it is sometimes overcome by the use of composite filters.

In a general approach, the information included in a composite filter consists of several views of the target under different situations (different rotation angles, scale variations, changes in illumination, etc.). The synthesis of all the information in a unique composite filter is carried out taking into account different constraints. The constraint operations provide desirable features of the composite filter such as sharp correlations peaks, noise robustness, low output-noise floor, etc. The main advantage of a composite filter in front of a bank of single filters is the reduction of time in the processing step. Only a single correlation can be enough to compare a given image with the whole set of distorted versions of the target to be recognized. However, composite filters can lack noise robustness and discrimination capability. In a composite filter, the number of images (distorted versions) of the reference is limited in order to obtain a satisfactory performance.

In this work, a Fourier-plane nonlinear filter[6,7] is used as a composite nonlinear filter. Let s_i represent N training images and let P be the total number of pixels contained in each image. Instead of a matrix we use a vector notation to represent an image. Let consider a PxN training data matrix **S** that has the vector s_i as its i'th column. The expression for the k'th-law equal-correlation-peak synthetic-discriminant-function filter (ECP SDF) is then:

$$\hat{h}_{ECP}{}^{k} = \hat{S}^{k}\left[\left(\hat{S}^{k}\right)^{+}\hat{S}^{k}\right]^{-1} c*. \tag{2}$$

where the ^ symbol denotes Fourier domain, * stands for complex-conjugation, parameter k controls the strength of the applied nonlinearity and vector c contains the desired cross-correlation peak value for each training image.

The k'th-law ECP SDF filter is tested in the following simulation experiments to design the recognition system. The performance of the system using this composite nonlinear filter is compared to the performance of the processor when a bank of single nonlinear filters is considered.

2.3 Metrics for performance evaluation

There are several metrics to evaluate correlation filter performance. Some of them are described in summarizing works elsewhere.[21] To evaluate correlation results in our experiments we will use a criterion based on the peak-to-correlation energy (PCE) parameter. This parameter measures the ratio between the intensity value of the output peak at the target location and the total energy of the output plane, that is[21]

$$PCE = \frac{\left|c(0,0)\right|^2}{\iint \left|c(x,y)\right|^2 dx\,dy}. \tag{3}$$

In general, a high and sharp correlation peak is expected when there is an object in the scene that matches the reference target, thus leading to a high value for the PCE parameter. The better match between an object of the scene and the reference, the closer value to unity for the PCE parameter will be reached. For this reason, PCE parameter seems to be a reliable criterion to base the final recognition decision.

A threshold operation permits to accept a true target or to reject a false object. The threshold level is sometimes established arbitrarily. However, it can also be determined by means of a learning algorithm. A set of training images, containing true targets and false objects permits to measure the probability of error in the recognition process depending on the threshold value. A final threshold level for the recognition procedure is established by considering a null or a minimum probability of error in the identification of the training objects. In our case, based on the PCE criterion, objects that obtain PCE values above the threshold are considered as true targets; otherwise the system will reject the object in the recognition process.

Another assumption could be taken into account. Sometimes objects contained in the analyzed scene are compared, simultaneously or sequentially, to different reference targets. In such a situation, different correlation planes are computed for each scene. The final recognition result for the analyzed scene can be processed in different ways depending on arithmetic or logical operations applied to the correlation outputs. In this work, a winner-take-all model is used. The PCE parameter is computed for all the output planes and the output plane with the maximum PCE value is selected as the final response of the system. Only if the maximum PCE value is above the threshold, established in the learning process, an object contained in the scene will be recognized as similar to the target.

2.4 Training and non-training (testing) images

Several images have been captured in a real environment. Firstly, a stop sign is being used as the true target to be detected. Pictures containing a stop sign have been divided into two groups: the set of true targets images that trains the recognition system, and a different set of non-training stop signs for testing the system's performance. Another set of images containing a different road sign (false object) is used to train the system and to test its discrimination capability. A training target, a non-training target and a false object are shown in figure 1a, 1b and 1c, respectively. Information of training true targets centered on a zero background is used as reference to build nonlinear filters for image processing.

Each image is 128 x 128 pixels. They are normalized to have a maximum gray scale of unity and then zero padded to 256 x 256 pixels. The normalized images are Fourier transformed and k'th-law nonlinearity is applied to them. The nonlinear correlation output is obtained by taking the inverse Fourier transform of the product between the nonlinearly modified spectra of both the input signal and the reference target. As a reference target we will consider either a single sign to synthesize a single nonlinear filter, or multiple views of a sign to synthesize a composite nonlinear filter. A single nonlinear filter is obtained by applying the nonlinear operator of eq. (1) to the complex-conjugated Fourier transform of the reference. Composite nonlinear filters are obtained by applying eq. (2). Each composite filter contains a maximum number of 6 views of the target.

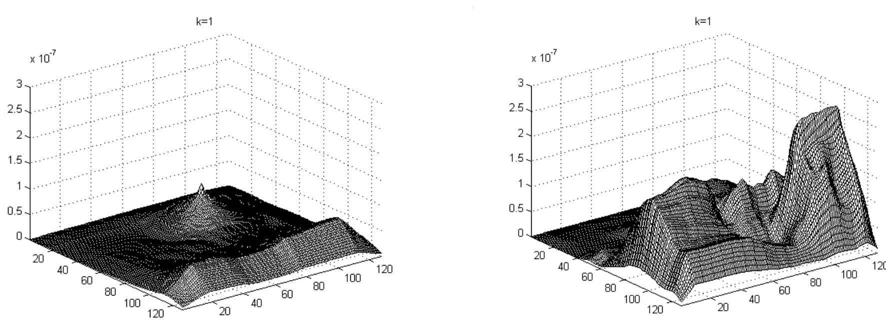
(a) (b) (c)

Figure 1. a) Training true target, b) Non-training true target, and c) False target.

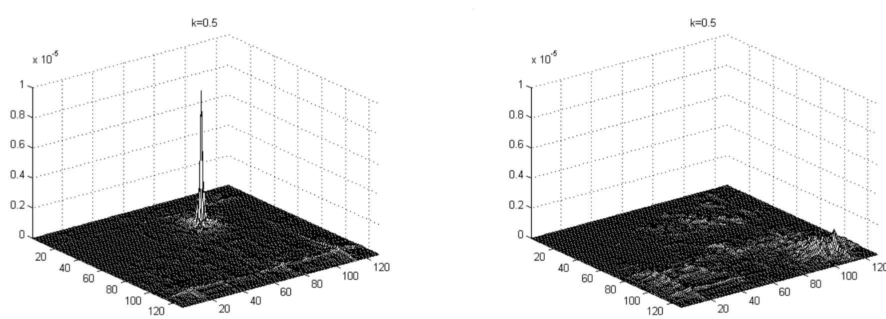
2.5 Determination of nonlinearity strength

We carry out a preliminary analysis in order to determine the nonlinearity that provides better recognition results. For this reason, we perform by simulation the correlation of all the training images with the reference patterns. We obtain a correlation output for each pair of images captured at the same distance. The training set consists of true target images and false targets. Figures 2 and 3 show the obtained results for different values of k. Analysis from the linear case (k=1) to a binarizing nonlinearity (k=0) is provided taking also into account intermediate values of parameter k.

Output planes for different values of k are shown in figure 2. For a given true target correlation planes are displayed on the left column. Output planes shown on the right column correspond to a given false sign. From figure 2 we observe that the highest peak on the lowest output-noise floor is achieved for k=0.1 (Figure 2g). A good discrimination for the false target is also obtained (Figure 2h).

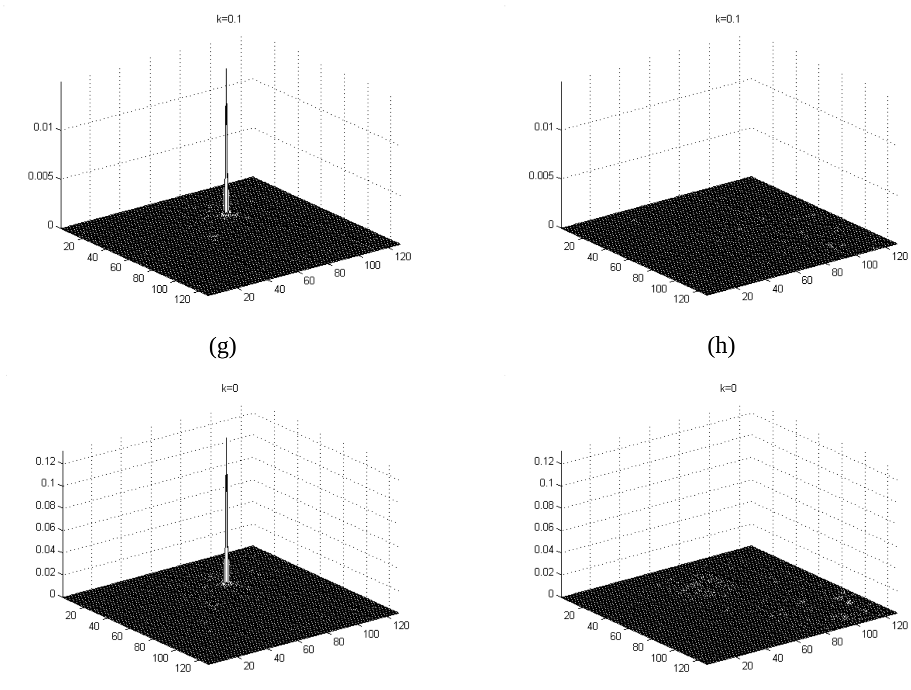


(a) (b)



(c) (d)

Figure 2. Correlation planes for different nonlinearities. Planes on the left column correspond to the analysis of a true target. Planes on the right correspond to a false target.

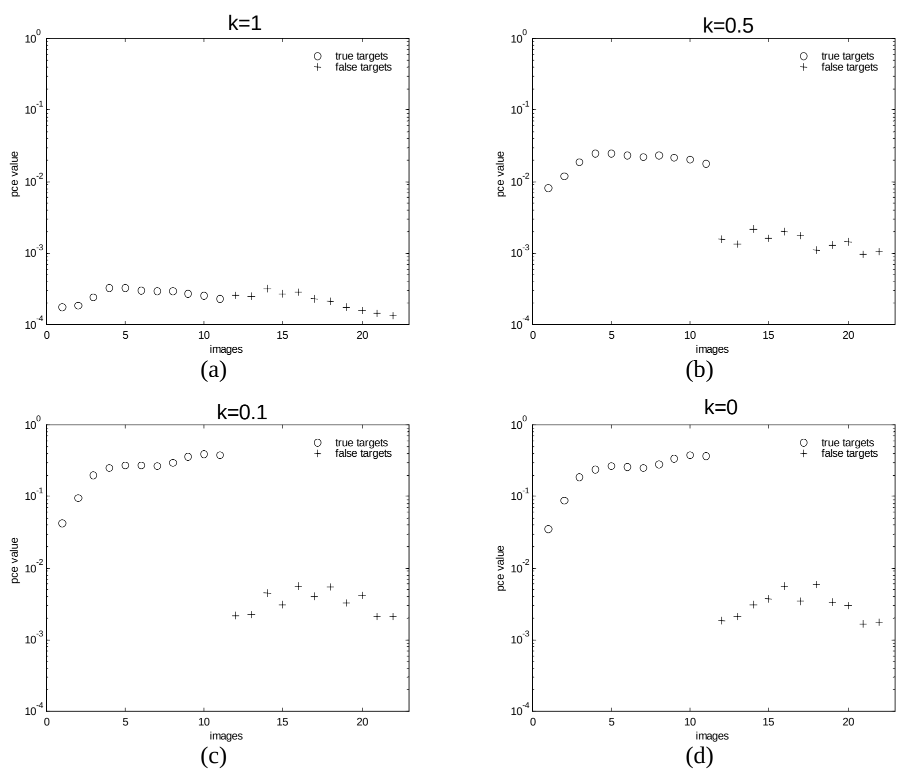


(g) (h)

(i) (j)

Figure 2. (Continue)

In figure 3, we represent PCE values obtained for each image in the training set. We observe that the largest differences between PCE values of true targets and PCE values of false targets are achieved for low values of parameter *k*.



(a) (b)

(c) (d)

Figure 3. PCE results for different nonlinearities. a) *k*=1 (linear case); b) *k*=0.5; c) *k*=0.1 and d) *k*=0 (binarizing nonlinearity).

In particular, our analysis shown that nonlinearity of k=0.1 improves correlation results in terms of peak sharpness and discrimination capability of the system. These results are in accordance with other results obtained for tolerance to target rotations.[6,7] Thus, value of k=0.1 is selected for all the simulations.

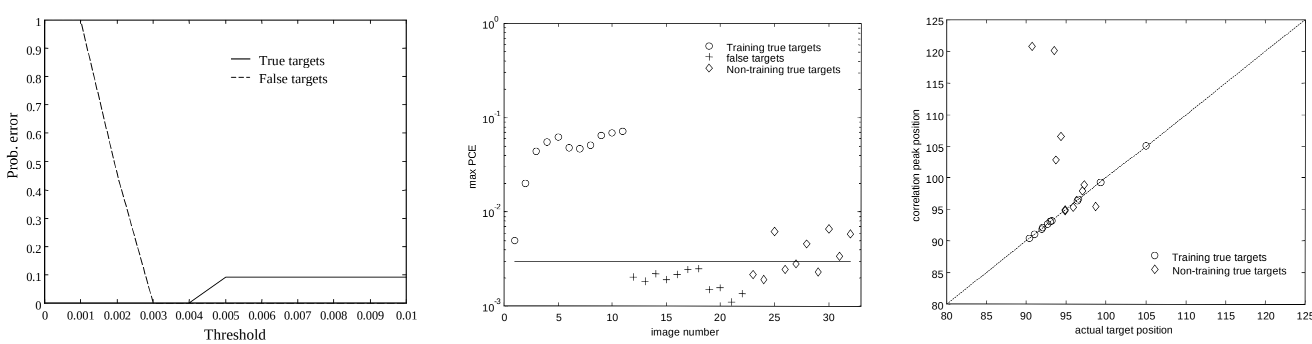
3. RECOGNITION RESULTS FOR DIFFERENT DISTORTIONS

3.1 Scale-invariance

Scale-invariance is required in a road sign recognition system to achieve detection of signs even when the acquisition system is in motion. The wider range of tolerance to variations in scale, the better capability for the recognition system to detect objects located at far distances. Performance of the bank of single nonlinear filters is compared to performance of the k'th-law ECP SDF filter.[22]

Recognition results are obtained for the entire training set, which is composed from true targets and false objects. The maximum PCE value is considered to classify signs as similar to the target or to discriminate them from the sign to be recognized. The learning algorithm allows selection of an appropriate threshold value for the output of the system. Output PCE values above the threshold correspond to objects considered as a true target, whereas, PCE values below the threshold imply the rejection of the object in the recognition process.

Figure 4a shows the probability of error in the recognition of road signs for the training set using a composite nonlinear filter. Solid line in the graphs indicates the probability of error in classification of training stop signs. A dashed line plots probability of error in rejection of false targets. A threshold value can be established when the probability of misclassification of false targets reaches the value of zero. Once the threshold is established, we use the entire set of images to test the performance of the system. Results for composite nonlinear filters are summarized in figure 4b. In this graph, output response of the system is displayed. The maximum PCE value achieved among the different output planes is plotted for all the images. A horizontal solid line shows the chosen threshold level. The maximum PCE value for non-training stop signs is above the threshold level in some cases. Thus, they are correctly recognized as the true class target. However, when the stop sign is located at a far distance from the acquisition system, the processor is not able to properly detect the road sign due to the small resolution and energy of the target. Figure 4c shows the position of the maximum correlation peak versus the actual position of the sign in the scene. The position is represented in this graph by means of the distance of the center of the sign to the origin of the image (pixel (0,0) is located in the left top corner). Figure 4c points out that some correlation peaks appear in a wrong position, thus, corresponding to false alarms.



(a) (b) (c)

Figure 4. Recognition results for the k'th-law ECP SDF filter. Scale-invariance of the system is tested. a) Probability of error in the classification of training images. b) Classification of true targets and false targets with respect to the established threshold value. c) Correlation peak position versus the actual target position in the scene.

For the case of a nonlinear filter bank, recognition results are shown in figure 5. Figure 5a plots the probability of error in the classification of training images. A larger region of null probability is obtained for the filter bank than for the composite nonlinear filters used in the previous experiment. Figure 5b plots the obtained PCE values along with the established threshold. No false alarm appears among the final output results and non-training stop signs are satisfactorily detected and located (Figure 5c). An advantage of using a bank of single nonlinear filters is the possibility of obtaining more information about the detected road sign for classification or parameter estimation. For instance, distance between the detected sign and the acquisition system will approximately coincide with distance between the camera and the training road sign that provides the best match. This information can be easily stored in the processor. Taking into account distances between training stop signs and the acquisition device, figure 5b can be represented in the equivalent graph shown in figure 5d. Figure 5d plots the maximum PCE value versus the distance between the detected road signs and the acquisition system. In this representation, it is possible to observe that non-training stop signs are recognized even for distances around 60 meters.

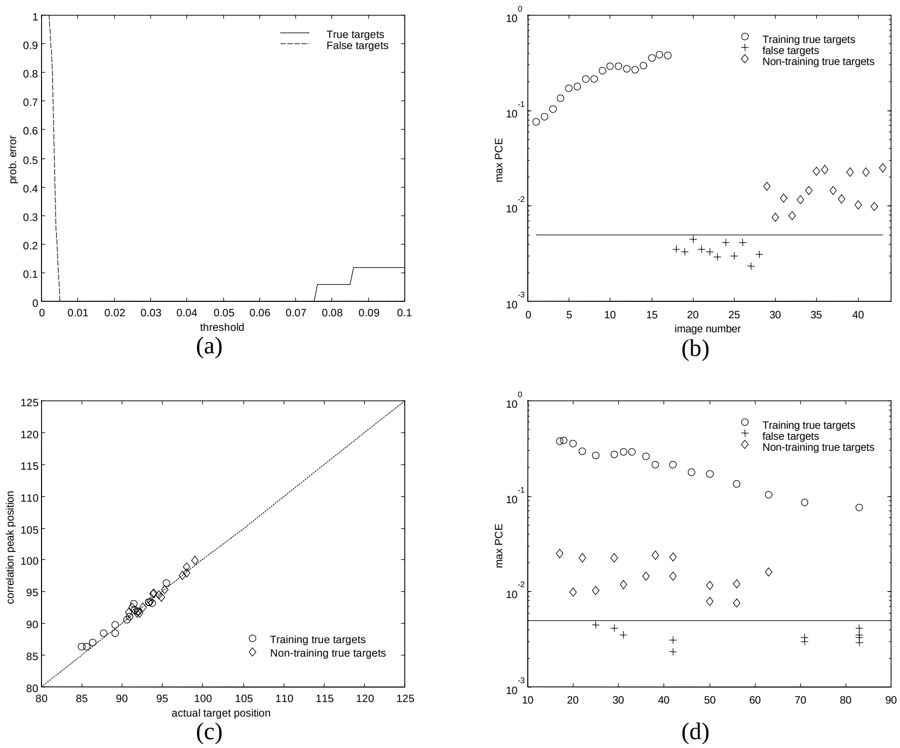


Figure 5. Recognition results for a bank of single nonlinear filters. Scale-invariance of the system is tested. a) Probability of error in the classification of training images. b) Classification of true targets and false targets with respect to the established threshold value. c) Correlation peak position versus the actual target position in the scene. d) PCE value versus the distance between the target and the acquisition system.

From the results shown in this experiment, composite nonlinear filters, which include scale distortions of the object to be recognized, do not satisfy requirements for discrimination capability in a scale-invariant system. These filters, however, have shown good performance when they were composed for in-plane and out-of-plane versions of the target.[6,7] One of the reasons of this different behavior could be explained by the largest changes in size and energy of objects when considering scaled versions than when considering rotated versions of the target. However, results obtained by the bank of single nonlinear filters are successful and encouraging. We have shown that a nonlinear processor based on a bank of nonlinear single filters makes feasible the detection of road signs varying in scale. Moreover, the proposed scale-invariant recognition system has been applied in the analysis of a video sequence. Detection and proper location of a non-training stop sign has been achieved for the different frames contained in the video sequence.[22]

3.2 Tolerance to in-plane rotations

In this subsection we examine the system's performance with respect to in-plane rotation of the objects. Two different methods can provide a recognition system with tolerance to in-plane rotations: synthesis of composite nonlinear filters by using in-plane rotated versions of the reference, or rotation of the input signal followed by its correlation with non-rotated versions of the target. In both cases, a digital algorithm to obtain rotated versions of the images is considered.

First, each training stop sign centered in a zero background is digitally rotated in increments of 3 degrees from -9 to 9 degrees around its vertical position. Rotated versions of the training sign are used to build a composite nonlinear filter by applying eq. (2) with k=0.1. We synthesize a composite nonlinear filter for each training stop sign captured at a different distance from the camera in order to maintain a scale-invariant system.

Figure 6a plots of the probability of error versus the threshold value in learning process. A k'th-law ECP SDF filter is used. Solid line in the graph indicates probability of error in the detection of true target and a dashed line plots the probability of error in the rejection of false objects, depending on the threshold value. A minimum threshold value can be established when the probability of misclassification of false objects reaches the value of zero. Afterwards, we test the performance of the recognition system by using a set of non-training stop signs captured with an in-plane rotation angle of 4 degrees. Results for the k'th-law ECP SDF filters are summarized in figure 6b. In this graph, the maximum PCE value achieved among the different output planes is plotted for all the images. A horizontal solid line is plotted at the value of the chosen threshold level. In general, stop signs obtain PCE values above the established threshold level. PCE values for false signs are below the threshold. However, we note that a PCE value obtained for a non-training stop sign is below the threshold. This implies that a false alarm appears in the recognition process. Furthermore, some of the correlation peaks for other testing images do not coincide with the actual target position in the scene as we can observe in figure 6c. This figure plots the position of the maximum correlation peak versus the actual position of the sign in the scene. In this graph, the incorrect position of some correlation peaks is noticed. They correspond also to false alarms.

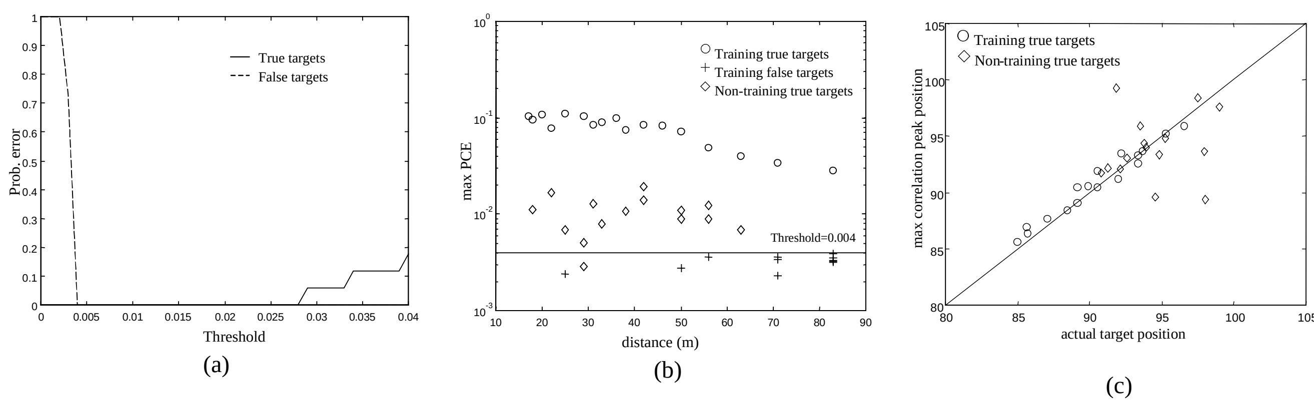


Figure 6. Tolerance to in-plane rotation of the target by using k'th-law ECP SDF filter. a) Probability of error in the classification of training images. b) Classification of true targets and false targets with respect to the established threshold value. c) Correlation peak position versus the actual target position in the scene.

In the second method, tolerance to in-plane rotations is achieved by rotating the input scene and it is compared to the single nonlinear filters belonging to the bank. The bank of filters contains information of the reference varying in scale to allow a scale-invariant recognition system. A digital algorithm for rotating the signal is used to obtain in-plane rotated versions of the scene to be analyzed. The input scene is rotated from -9 to 9 degrees in increments of 3 degrees. The output of the recognition system is related to the best match between the rotated versions of the input signal and the reference targets. Thus, the output coincides with the maximum PCE value.

Improvement of recognition results can be noticed in figure 7. In figure 7a we observe that the interval with null probability of error increases, whereas in figures 7b and 7c the successful recognition task is pointed out. That is, the training and non-training stop signs are correctly detected and located at the right position. They are also successfully distinguished from the other road signs used to test the discrimination capability of the system.

From comparison of the obtained results, we remark that if some tolerance to in-plane rotation is required in the recognition system, better results are achieved by using a single nonlinear filter and rotating the input image, rather than designing a composite nonlinear filter for rotation invariance.

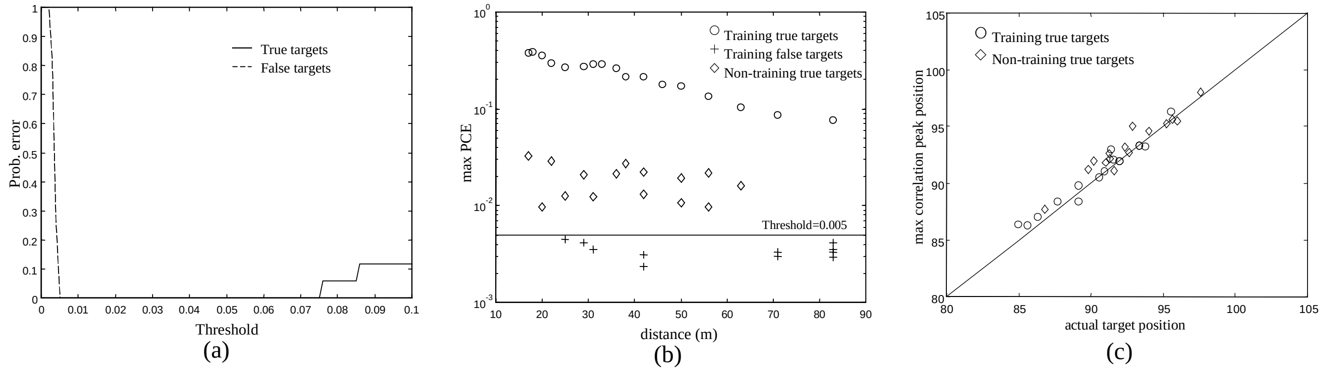


Figure 7. Tolerance to in-plane rotation of the target by rotating the input scene. a) Probability of error in the classification of training images. b) Classification of true targets and false targets with respect to the established threshold value. c) Correlation peak position versus the actual target position in the scene.

3.3 Tolerance to out-of-plane rotations

We focus our analysis on the system's tolerance to out-of-plane rotations. Due to the difficulty of generating digitally out-of-plane rotated versions of the images, we implement them optically. Thus, stop signs are captured out-of-plane rotated from -9 to 9 degrees in increments of 3 degrees during the acquisition process. They are used as training images. These signs centered in a zero background, are used to construct composite nonlinear filters. A k'th-law ECP SDF filter with k=0.1 is obtained for each distance between the sign and the acquisition system to maintain scale-invariance.

A learning algorithm allows establishing the threshold value for the output of the recognition system. The value of the threshold is determined based on the results of figure 8a. Several non-training images slightly out-of-plane rotated are captured and used to test the system's tolerance to this type of distortion. Pictures are taken from a view angle of 4 degrees. A wide range of distances between the road sign and the acquisition camera are also considered to keep scale-invariance.

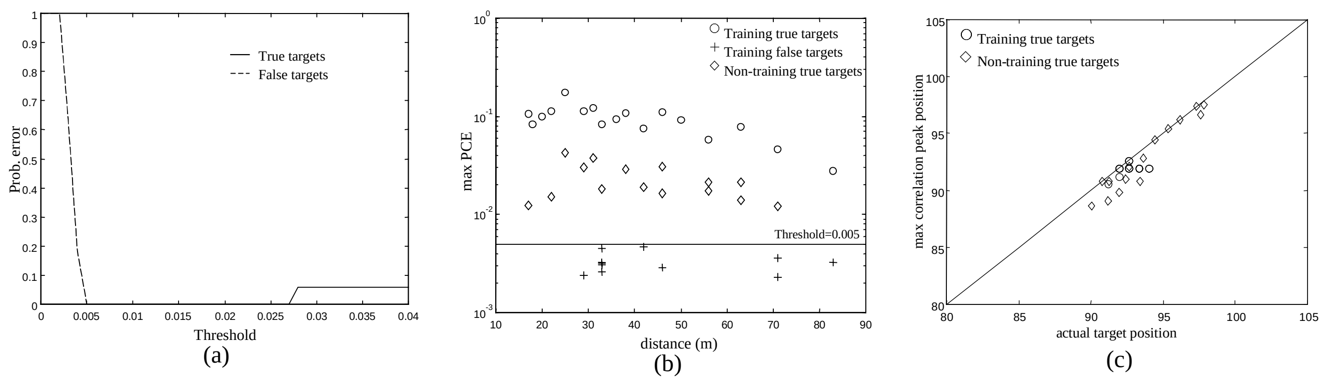


Figure 8. Tolerance to out-of-plane rotation of the target by using k'th-law ECP SDF filters. a) Probability of error in the classification of training images. b) Classification of true targets and false targets with respect to the established threshold value. c) Correlation peak position versus the actual target position in the scene.

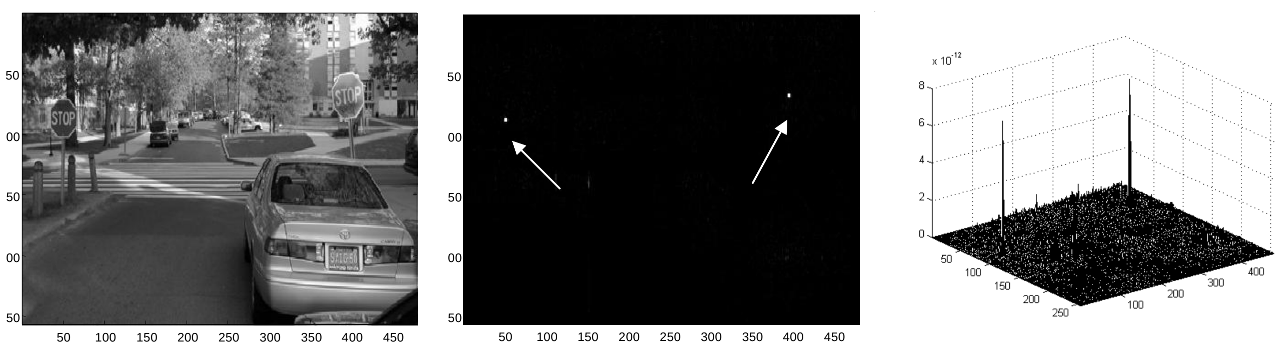
Recognition results, once the established threshold level is applied, are shown in figure 8b. Recognition of stop signs is always achieved by a PCE value larger than the threshold level. They are also discriminated from other signs. Correlation peaks corresponding to stop signs are located at the same position as the sign in the scenes (Figure 8c). Results contained in figures 8b and 8c show that the proposed recognition system is able to detect a partially out-of-plane rotated road sign at different distances from the acquisition system. This is due to the bank of composite nonlinear filters that are being used. Information of out-of-plane rotation is included in the k'th-law ECP SDF filters and allows detecting the sign even if it is slightly out-of-plane rotated or if it is captured from a different view angle by the acquisition system.

4. ANALYSIS OF REAL IMAGES IN NON-OVERLAPPING CLUTTERED BACKGROUND

The separate analyses of different distortions have shown an encouraging performance of the recognition system as a distortion-tolerant processor. We now apply the distortion-tolerant system to new captured images. They are selected as samples where it is difficult to recognize the road sign due to the amount of involved distortions. Selected images include stop signs modified by several distortions. They are captured under varying illumination due to shadows or different weather conditions, and in some cases the sign to be detected has been vandalized or appears partially occluded. In all the cases, a real cluttered background surrounds the stop signs.

The designed recognition system is based on a nonlinear processor that uses a bank of composite nonlinear filters. The bank of filters serves to achieve scale-invariance in a wide range of distances from the sign to the acquisition system. Composite nonlinear filters, in particular k'th-law ECP SDF filters, provide tolerance to out-of-plane rotation of targets. Finally, rotation of the input signal allows tolerance to in-plane rotations. A certain degree of tolerance to illumination fluctuations is achieved as a consequence of using a nonlinear processor with parameter *k* close to zero (*k*=0.1)

Figure 9a corresponds to an analyzed scene that includes two stop signs to be detected. These signs are located at both sides of the road, and they have different illumination. The stop sign on the left has a low average energy due to a shadow that completely covers it. This sign is partially in-plane and out-of-plane rotated. The stop sign on the right, however, has a non-uniform illumination due to shadows caused by the leaves and it has been vandalized. This sign is tilted, so that tolerance to in-plane rotation is needed to detect it correctly. It is also out-of-plane rotated. We observe that the background of the picture is quite cluttered and there are areas with larger energy than the energy of stop signs. Figure 9b shows a 2D representation of the output correlation plane where two high intensity peaks appear and coincide with the position of the two true targets. Correlation peaks are easily observed in the 3D representation of the correlation plane (figure 9c). We remark that the recognition of both stop signs is achieved under different illumination conditions of the signs. This is mainly due to the nonlinearity applied in the nonlinear process.

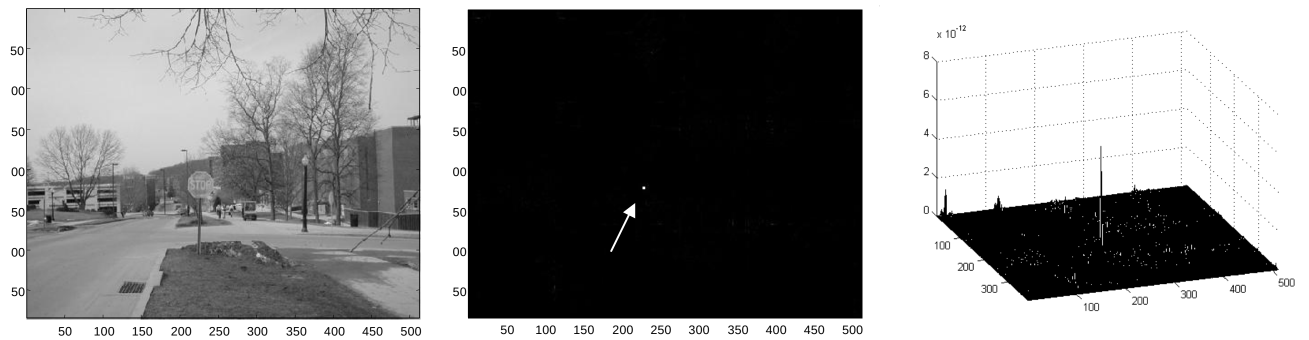


(a) (b) (c)

Figure 9. Recognition of stop signs by the proposed distorted-tolerant system. a) Input scene. b) 2D representation and c) 3D representation of the output plane.

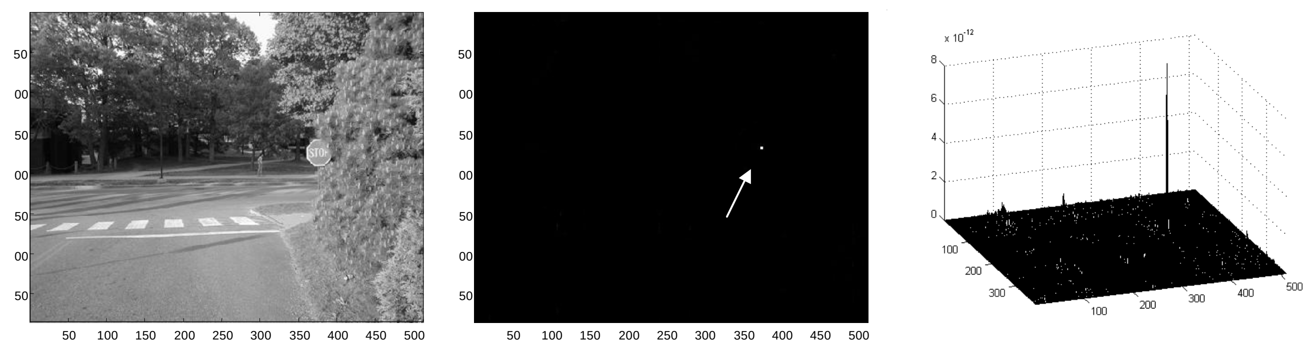
A second sample consists of a stop sign strongly faded (Figure 10a). The sign appears in a cluttered background and with an inverse contrast. However, a high and sharp correlation peak appears in the actual position of the road sign (figures 10b and 10c). This implies a satisfactory recognition of the sought sign.

Last example of stop sign recognition corresponds to the analysis of the image displayed in figure 11a. The stop sign that is contained in the scene appears partially occluded by a tree. Detection and location of this sign is also satisfactory as it can be seen from the 2D output graph of figure 11b or as a 3D representation in figure 11c. A high and sharp peak is obtained in a low output-noise floor.



(a) (b) (c)

Figure 10. Recognition of a faded stop sign by the proposed distorted-tolerant system. a) Input scene. b) 2D representation and c) 3D representation of the output plane.



(a) (b) (c)

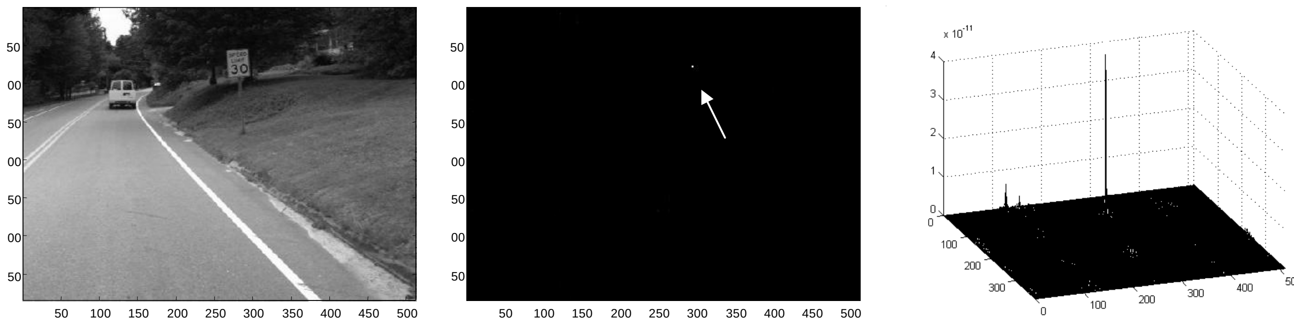
Figure 11. Recognition of a partially occluded stop sign by the proposed distorted-tolerant system. a) Input scene. b) 2D representation and c) 3D representation of the output plane.

To show the versatility of the proposed road sign recognition system, another road sign, which differs from a stop sign, is considered as the reference. Our second recognition task consists of detecting and locating every speed limit sign, regardless its speed limit number. To solve this task, a bank of composite nonlinear filters is built. For a given speed limit scale, a k'th-law ECP SDF filter includes the information of several speed limit numbers (from 25 mph to 65 mph). A bank of such composite filters permits a scale-invariant detection. Rotation of the input scene allows tolerance to in-plane rotation of the signs. In this case, tolerance to out-of-plane rotations has been neglected, however, it is possible to include the information of out-of-plane distorted views of the target in the composite nonlinear filters. Recognition results are presented in the following figures. Scenes shown in this work have been chosen as remarkable examples due to the difficulty on the detection of road signs.

Figure 12 shows the proper detection and location of a 30mph speed limit sign in the presence of an object of similar shape included in the background. A sharp and high peak is obtained in the target position, which a PCE value above the established threshold. No false alarms are obtained.

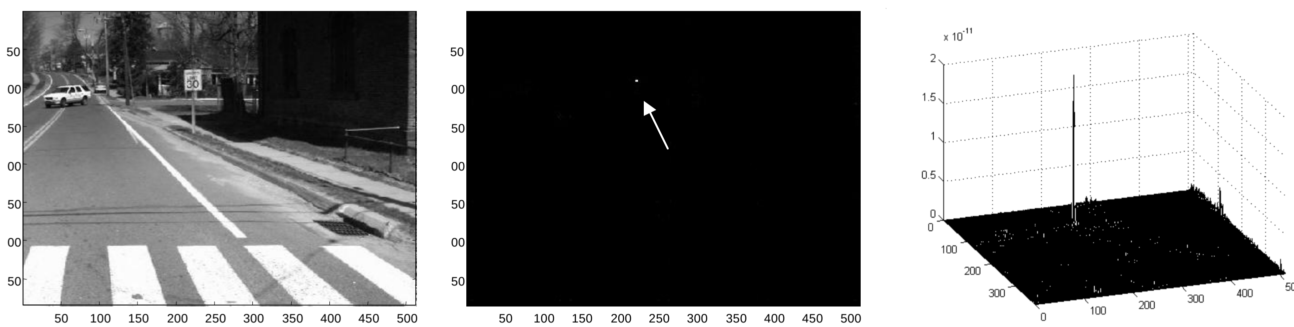
Image included in figure 13a corresponds to another example of recognition of a speed limit sign along with rejection of objects with similar energy. A high and sharp peak allows location of the target, whereas no false alarms appear (see Figures 13b and 13c).

Figure 14a shows a distant 50 mph speed limit sign. The correlation plane, which is displayed as a 2D image (Fig. 14b) and as a 3D graph (Fig. 14c), points out the successful recognition of this sign despite its low resolution and non-uniform illumination. The same sign, but closer to the acquisition system and partially occluded is also detected with accuracy as it is shown in figure 15. The recognition is achieved even though half of the sign information is occluded.



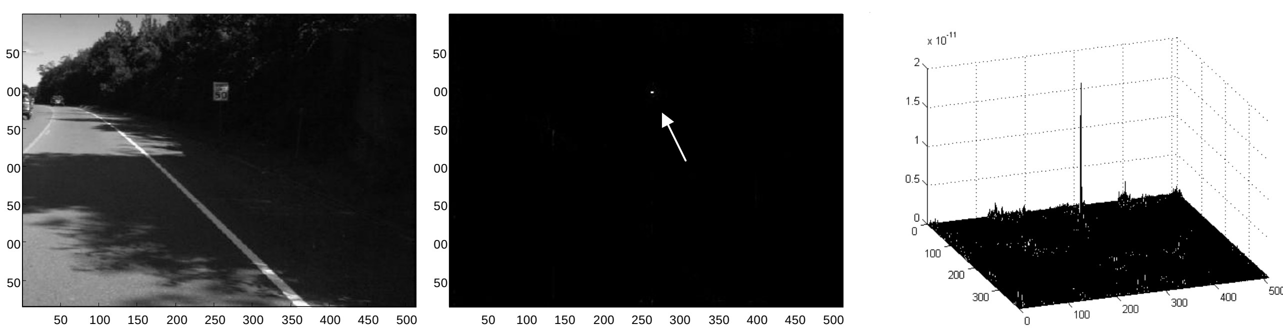
(a) (b) (c)

Figure 12. Recognition of a 30 mph speed limit sign with low energy by the proposed distorted-tolerant system. a) Input scene. b) 2D representation and c) 3D representation of the output plane.



(a) (b) (c)

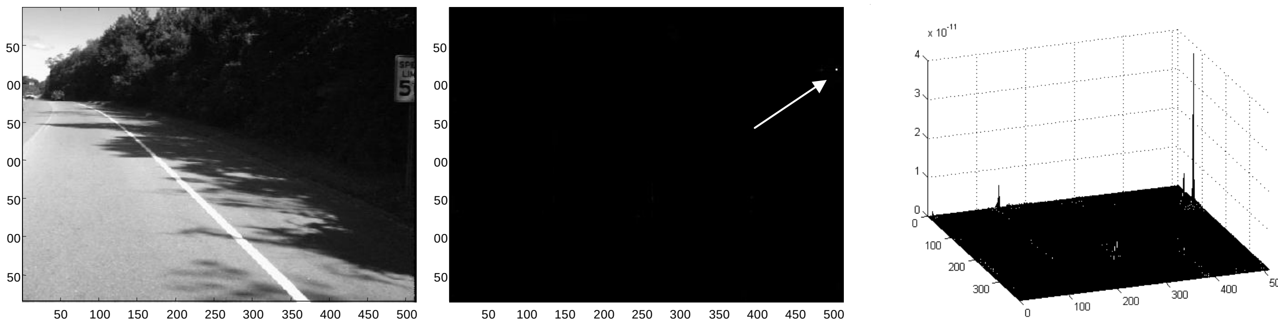
Figure 13. Recognition of a 30 mph speed limit sign on a cluttered background by the proposed distorted-tolerant system. a) Input scene. b) 2D representation and c) 3D representation of the output plane.



(a) (b) (c)

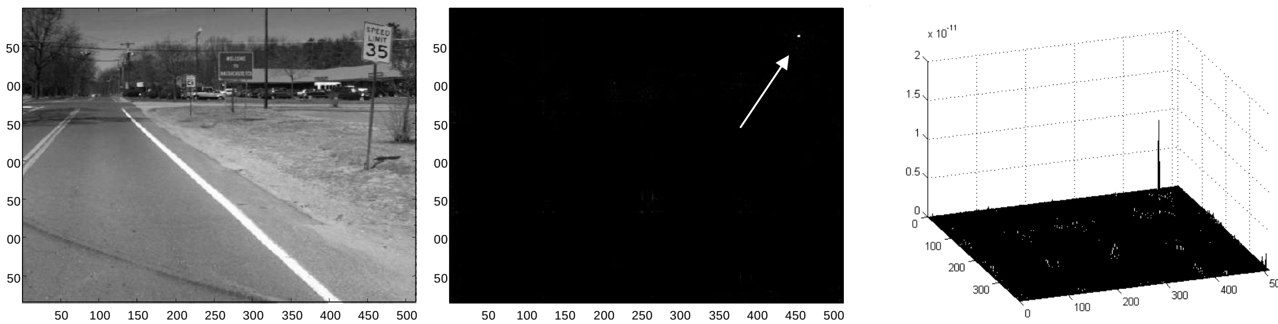
Figure 14. Recognition of a distant 50 mph speed limit sign by the proposed distorted-tolerant system. a) Input scene. b) 2D representation and c) 3D representation of the output plane.

Another example is the one shown in figure 16a, which corresponds to a rotated speed limit sign. It is satisfactorily recognized as the PCE value for its correlation peak is above the established threshold. Figures 16b and 16c show the obtained correlation plane in a 2D representation and in a 3D graph, respectively.



(a) (b) (c)

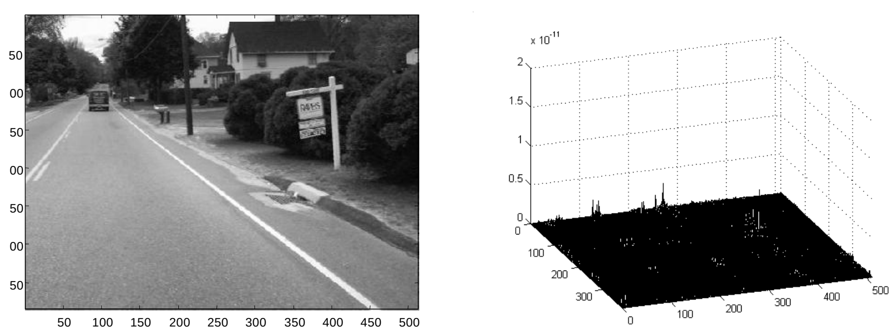
Figure 15. Recognition of a partially occluded 50 mph speed limit sign by the proposed distorted-tolerant system. a) Input scene. b) 2D representation and c) 3D representation of the output plane.



(a) (b) (c)

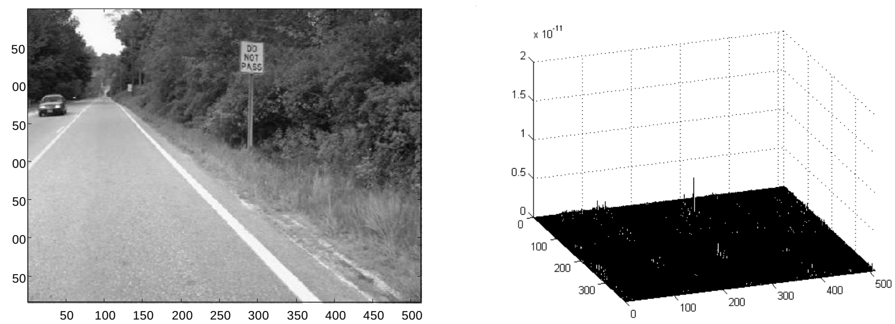
Figure 16. Recognition of a rotated 35 mph speed limit sign by the proposed distorted-tolerant system. a) Input scene. b) 2D representation and c) 3D representation of the output plane.

Finally, two examples of false signs are included in this work (see Figures 17 and 18). The analyzed images contain false targets that are perfectly discriminated from the speed limit targets, even though they have a high similarity in shape with respect to the object to be recognized. In both correlation planes (Figures 17b and 18b), the corresponding PCE value is below the established threshold, so that no target is detected.



(a) (b)

Figure 17. Recognition results for the proposed distorted-tolerant system when a false sign is analyzed. a) Input scene. b) 3D representation of the output plane.



(a) (b)

Figure 18. Recognition results for the proposed distorted-tolerant system when a false sign (do no enter sign) is analyzed. a) Input scene. b) 3D representation of the output plane.

From the results shown in this section, it is remarkable that the proposed recognition system is able to detect and locate road sign in real background images. The detection is successfully achieved even when the road sign is varying in scale, slightly rotated, illuminated under different conditions, faded or partially occluded. Detecting different road signs has shown versatility of the proposed processor.

5. CONCLUSIONS

A road sign recognition system has been proposed based on a nonlinear processor. Analysis of different filtering methods allows us to select the best techniques to overcome a variety of distortions. The most frequent distortions when dealing with road sign detection are scale variations, in-plane and out-of-plane rotation and illumination variations of the targets.

The entire processor performs several correlations between different input scenes and a set of reference targets. Multiple correlation results are then processed to give a single recognition output. A learning algorithm is carried out to establish a threshold value, which determines whether or not any object contained in an input scene is similar to the target.

Scale-invariance is provided to the recognition system by means of a bank of nonlinear filters. Filter bank recognition system shows a better performance than composite nonlinear filters. Images of a true target captured from different distances constitute the set of filters in the bank. Apart from locating a true sign, this method allows us to approximately determine the distance between the acquisition system and the road sign.

In-plane rotation invariance is achieved by rotating the input scene. Recognition results obtained by this method are compared to results obtained for composite nonlinear filters. Composite filters are constructed by using digital rotated versions of the reference target. In-plane rotation of the input scene allows better detection results than composite filters. Moreover, in the design of composite filters the maximum number of images included in a composite filter is limited, whereas range of the input scene rotation can be determined based on the application.

Using composite nonlinear filters rather than using individual filters in the filter bank can satisfy tolerance requirements for out-of-plane rotation of the targets, as we have shown by the detection of stop signs. By using the same procedure, tolerance to some information included in a given road sign (for instance, the number of a speed limit sign) is also achieved. In particular, k'th-law equal-correlation-peak synthetic-discriminant-function (ECP SDF) filters are used as composite nonlinear filters.

The entire recognition system has been tested in real still images as well as in video sequences. Scenes were captured in real environments, with cluttered backgrounds and contained many distortions simultaneously. Recognition results for various images show that, the proposed recognition system is able to properly detect a given road sign even if it is varying in scale, slightly tilted or viewed under different angles. In addition, the system is robust to changes in illumination due to shadows or weather conditions. It is also able to locate a faded or vandalized sign along with partially occluded road signs. Obviously, the processor can be designed for different varieties of road signs in noisy background scenes, as we have satisfactorily shown for stop and speed limit signs.

ACKNOWLEDGEMENTS

E. Pérez wishes to thank the *Ministerio de Ciencia y Tecnología* (Spain), project DPI2000-0991, and M.-A. Castro the *Consejo Nacional de Ciencia y Tecnología* (Mexico) for financial support.

REFERENCES

1. H. J. Caulfield, W. T. Maloney, "Improved discrimination in optical character recognition," *Appl. Opt.*, **8**, pp. 2354-2356, 1969.
2. H. J. Caulfield, "Linear combinations of filters for character recognition: a unified treatment," *Appl. Opt.*, **19**, pp. 3877-3879, 1980.
3. C. F. Hester, D. Casasent, "Multivariant technique for multiclass pattern recognition," *Appl. Opt.*, **19**, pp. 1758-1761, 1980.
4. D. Casasent, "Unified synthetic discriminant function computational formulation," *Appl. Opt.*, **23**, pp. 1620-1627, 1984.
5. A. Mahalanobis, B. V. K. Vijaya Kumar, D. Casasent, "Minimum average correlation energy filters," *Appl. Opt.*, **26**, pp. 3633-3640, 1987.
6. B. Javidi, D. Painchaud, "Distortion-invariant pattern recognition with Fourier-plane nonlinear filters," *Appl. Opt.*, **35**, pp. 318-331, 1996.
7. B. Javidi, W. Wang, G. Zhang, "Composite Fourier-plane nonlinear filter for distortion-invariant pattern recognition," *Opt. Eng.*, **36**, pp. 2690-2696, 1997.
8. B. Javidi, J. L. Horner, A. Fazlollahi, J. Li, "Illumination-invariant pattern recognition with a binary nonlinear joint transform correlator using spatial frequency dependent threshold function," *Proc. SPIE*, **2026**, pp. 100-106, 1993.
9. E. Silvera, T. Kotzer, J. Shamir, "Adaptive pattern recognition with rotation, scale, and shift invariance," *Appl. Opt.*, **34**, pp. 1891-1900, 1995.
10. Ph. Refregier, B. Javidi, G. Zhang, "Minimum mean-square-error filter for pattern recognition with spatially disjoint signal and scene noise," *Opt. Lett.*, **18**, pp. 1453-1455, 1993.
11. Ph. Refregier, V. Laude, B. Javidi, "Nonlinear joint-transform correlation: an optimal solution for adaptive image discrimination and input noise robustness," *Opt. Lett.*, **19**, pp. 405-407, 1994.
12. B. Javidi, J. Wang, "Optimum distortion-invariant filter for detecting a noisy distorted target in nonoverlapping background noise," *J.Opt. Soc. Am. A*, **12**, pp. 2604-2614, 1995.
13. B. Noharet, R. Hey, H. Sjoberg, "Comparison of the performance of correlation filters on images with real-world non-overlapping noise: influence of the target size, illumination and in-plane rotation distortion," *Proc. SPIE*, **3386**, pp. 212-221, 1998.
14. L. Guibert, G. Keryer, A. Servel, M. Attia, H. S. MacKenzie, P. Pellat-Finet, J. L. de Bougrenet de la Tocnaye, "On-board optical joint transform correlator for real-time road sign recognition," *Opt. Eng.*, **34**, pp. 135-143, 1995.
15. T. D. Wilkinson, Y. Perillot, R. J. Mears, J. L. de Bougrenet de la Tocnaye, "Scale-invariant optical correlators using ferroelectric liquid-crystal spatial-light modulators," *Appl. Opt.*, **34**, pp. 1885-1890, 1995.
16. J. J. Esteve-Taboada, Ph. Refregier, J. Garcia, C. Ferreira, "Target localization by wavelength multiplexing," abstracts of the *IV Euro-American Workshop on Optoelectronic Information Processing. Optics for Information Systems*, 2001.
17. J. J. Esteve-Taboada, J. Garcia, C. Ferreira, "Extended scale-invariant pattern recognition with white-light illumination," *Appl. Opt.*, **39**, pp. 1268-1271, 2000.
18. M. Taniguchi, K. Matsuoka, "Detection of road signs by using a multiple correlator and partial invariant filters," *Proc. SPIE*, **3490**, pp. 178-181, 1998.
19. M. Taniguchi, Y. Mokuno, K. Kintaba, K. Matsuoka, "Correlation filter design for classification of road sign by multiple optical correlators," *Proc. SPIE*, **3804**, pp. 140-147, 1999.
20. B. Javidi, "Nonlinear joint power spectrum based optical correlation," *Appl. Opt.*, **28**, pp.2358-2367, 1989.
21. J. L. Horner, "Metrics for assessing pattern-recognition performance," *Appl. Opt.*, **31**, pp. 165-166, 1992.
22. E. Pérez, B. Javidi, "Image processing for intelligent transportation systems: application to road sign recognition," in *Smart Imaging Systems*, edited by B. Javidi, **PM91**, pp. 207-232, SPIE Press, Bellingham 2001.