Object recognition within cluttered scenes employing a Hybrid Optical Neural Network (HONN) filter

Ioannis Kypraios, Rupert Young*, Philip Birch, Chris Chatwin

Laser and Photonics Systems Research Group

Department of Engineering and Design

University of Sussex

Brighton BN1 9QT

ABSTRACT

We have recently proposed a hybrid filter, which we call the Hybrid Optical Neural Network (HONN) filter. This filter combines the optical implementation and shift-invariance of correlator-type filters with the non-linear superposition capabilities of artificial neural network methods. The filter demonstrates good performance in maintaining high quality correlation responses and resistance to clutter to non-training in-class images at orientations intermediate to the training set poses. This paper presents the design and implementation of the HONN filter architecture and assesses its object recognition performance in clutter.

Keywords: correlation filter, artificial neural network, synthetic discriminant function, optical pattern recognition

* E-mail: R.C.D.Young@sussex.ac.uk; Tel: +44(0)1273 678908; Fax: +44(0)1273 690814

1 Introduction

There are three main considerations[1-3] in the design and performance assessment of a pattern recognition filter. Firstly, the filter must be able to detect the in-class object and demonstrate tolerance to in-plane and out-of-plane rotation. Secondly, it must possess good discrimination abilities between the in-class and the out-of-class objects. Usually, we are interested in designing the filter to give a wide distortion range in order to reduce the number of intermediate object poses in the training set images. Thirdly, the filter must exhibit good detection of the object in cluttered scenes, i.e. demonstrate good tolerance to noise and clutter in the input scene. Kumar and Hasserbrook[1] have presented several performance measures for correlation-type filters to enable their comparison, such as peak sharpness, peak location, light efficiency, discriminability, distortion invariance and target-to-clutter ratio.

In an effort to keep consistency between the different mathematical symbols of the artificial neural networks and optical correlators we need to unify their representation. We denote the variable names and functions by non-italic letters, the names of the vectors by italic lower case letters and the matrices by italic upper case. The frequency domain vectors, matrices, variable names and functions are represented by bold letters and the space domain vectors, matrices, variables and functions by plain letters.

The Synthetic Discriminant Function[4] (SDF) correlation filter belongs in the wider category of Linear Combination Filters[5-7] (LCF). The main idea of the SDF filter is the inclusion of the expected distortions in the filter design such that improved immunity to these distortions is achieved. Multi-class object recognition is also

possible by including the out-of-class objects (non-targets) in the filter design. The conventional SDF filter is based on using a weighted linear combination of distorted reference images to create a composite image, which should cross-correlate with input images to produce equal on-axis height correlation peaks for all the input images belonging to the same class. Let $H(m,n)$ denote the composite image in the space domain and $X_i(m,n)$ denote the training image set of size $[\mathrm{m}\times\mathrm{n}]$ in the space domain, where $\mathrm{i}=1,2,\cdots,\mathrm{N}$ and N is the number of the training images used in the synthesis of the SDF. The basic filter's transfer function, from the weighed linear combination of X_i, is given by:

$$\mathrm{H}(\mathrm{m},\mathrm{n})=\sum_{\mathrm{i}=1}^{\mathrm{N}} \mathrm{a}_{\mathrm{i}}\, X_i(m,n) \tag{1}$$

where the coefficients $\mathrm{a}_{\mathrm{i}}(\mathrm{i}=1,2,\cdots,\mathrm{N})$ are to set the constraints on the peak given by c**.** The a_{i} values are determined from:

$$\mathrm{a}_{\mathrm{i}} = R^{-1}\mathrm{c} \tag{2}$$

where R is the correlation matrix of X_i and c is the peak constraint vector. The elements of this are usually set to zeros for false class objects and to ones for true class objects.

Recent advances in the area of optical processing and in particular enabling the spatial light modulator (SLM) technology, has made possible the implementation of compact optical correlator systems. Chao et al.[8] have demonstrated experimental results taken from the realisation of the maximum average correlation height (MACH)[8–11] filter in their grayscale optical correlator. However, the modulation levels afforded by the SLM limit the number of the training set images in the

implementation of the optical filter. Several efforts[12, 13] have been made to overcome this limit by designing filters to produce acceptable uniformity of performance over the training set and being optimal to multiple performance criteria goals, such as those mentioned above.

There are several implementations presented in the literature of optical pattern-recognition neural networks (OPRNN), such as the vector-matrix-multiplier-based[14] OPRNN, the photorefractive-crystal-based holographic[15] OPRNN and the spatial light modulator[16, 17] (SLM) -based OPRNN. The SLM-based system and the vector-matrix-based systems are operated in the spatial domain. So, their main drawback is that shift invariance[18] can be achieved only at the expense of massive and redundant interconnections. The main drawback of the photorefractive-crystal-based holographic OPRNN[19,20], which is operated in the Fourier-transform plane, is the limited shift invariance achievable due to the narrow Bragg-angle restriction within a thick photorefractive medium. Chao and Stoner[21, 22] presented a technique for optical implementation of a feature-based optical neural network. It uses the neocognitron[22] paradigm as a development guideline and it is implemented by a multichannel correlator, forming a generic neural network layer to provide parallel and shift invariant feature correlation. Talukder and Casasent[23,24] presented a method called the maximum representation and discriminanating feature (MRDF) extraction technique to compute nonlinear features for simultaneous representation and classification. The method has been demonstrated to perform better than the most discriminating feature (MDF) techniques[25] and the Fisher linear discriminant technique and its variations. However, it is stated that the MRDF discrimination technique requires a large feature database in order to find the useful discrimination

information present [24]. For practical applications and for non-deformable objects with constant volume, usually only a limited number of training images is available.

Section 2 presents the overall structure of the hybrid optical neural network (HONN) filter [26] and gives details of the artificial neural network (NNET) block of the filter. It describes two possible custom NNETs realisable within the system. Section 3 presents our choice of the custom NNET and analyses its design. Section 4 presents the experiments carried out to evaluate the performance of the HONN filter; first the filter's peak sharpness and detectability, second its distortion range, then the filter's discrimination ability and finally the filter's tolerance-to-clutter performance. Section 5 concludes.

2 The Hybrid Optical Neural Network (HONN) Filter

The hybrid optical neural network filter combines the digital design of a filter by artificial neural network techniques with an optical correlator-type implementation of the resulting combinatorial correlator type filter (see Fig. 1). Thus in effect, there are two main design blocks in the hybrid optical neural network filter, the NNET and here we choose for the correlator type block to be an SDF-like filter [27,28]. The original images pass first through the NNET block. The output of the SDF block is a composite image of the hybrid optical neural network filter's output. To test the HONN filter, we correlate the filter with an input image.

2.1 *The Artificial Neural Network*

We have designed a custom artificial neural network architecture to fit our purposes. If we assume we have a training set consisting of N images, we train a specifically configured neural net with this set of images. The network has N neurons

at the hidden layer, i.e. equal to the number of training images. There is a single neuron at the output layer to separate two different object classes. (In a multi-class object recognition problem, the increase of the different classes of objects would require more than one neuron at the output layer to correctly separate all the training images.) The net input of each of the neurons in the hidden layer is given by:

$$net_1 = \sum_{j=1}^{m \times n} w_j \; x_1(j) \tag{3}$$

where *net* is the net input vector of each of the hidden neurons. w_j is the input vector from the input layer to the hidden neurons for the training image x_1 in vector form of size $[1 \times (m \times n)]$. Similarly, for the training image x_N of size ($[1 \times (m \times n)]$ in vector form) the net input, net_N, is given by:

$$net_N = \sum_{j=1}^{m \times n} w_N \; x_N(j) \tag{4}$$

From Eqns.(1) and (3) and (4) there is a direct analogy between the SDF filter synthesis procedure and the combination of all the layers' weighted input vectors.

Two possible and equivalent custom designs of artificial neural network architectures (NNET) are suggested to form the basis of the SDF filter synthesis. Assume there are three training images of a car, size $[100 \times 100]$ ($[1 \times (100 \times 100)]$ in vector form), of different angle of view, to pass through the NNET. The first design (see Fig. 2) assumes one input source used for all the training images. (The input source in Fig. 2 consists of 10,000 i.e. $[1 \times (100 \times 100)]$ input neurons equal to the size of each training image (in vector form). Each layer needs, by definition, to have the same input connections to each of its hidden neurons. However, Fig. 2 is referred to as four layer since there are three hidden layers (shown

here aligned under each other) and one output layer. The input layer does not contain neurons with activation functions and so is omitted in the numbering of the layers. Each of the hidden layers has only one hidden neuron. Though the network initially is fully connected to the input layer during the training stage, only one hidden layer is connected for each training image presented through the NNET. Fig. 2 is thus not a contiguous three layer network during training, which is why the distinction is made.

The second design assumes three separate input sources each of 10,000 or $[1 \times (100 \times 100)]$ input neurons for each training set image (see Fig. 3). Each layer consists of a single hidden neuron but all the layer weights are fully connected to the output layer. As before, there is a single output neuron. The weights from the input layer to the hidden layers, called the input weights, are partially connected with the neurons of the hidden layers. The input weights of the first input source are connected only with the hidden layer 1, the input weights of the second input source are connected only with the hidden layer 2 and the third input source are connected only with the hidden layer 3. Each layer has only a single hidden neuron. All the weights of the hidden layers are connected fully to the output layer. In this case there is a single output neuron, since we want to recognise two different classes of objects. In contrast to the first design, now both the input weights and the weights from the hidden layers to the output layer, i.e. the layer weights, stay connected during the training session.

In both of the designs each neuron of the hidden layer is trained only with one of the training set images. In effect, neuron_1 with the training image x_1, neuron_2 with the training image x_2 and so on, ending with neuron_n with the training image x_n. In the first design the number of the input sources is kept constant whereas in the

second design the number of the input sources is equal to the number of the training images. In both designs each hidden neuron learns one of the training images. In effect the number of the input weights for both designs increases proportionally to the size of the training set:

$$N_{iw} = N \times [m \times n] \tag{5}$$

where N_{iw} is the number of the input weights, N, is the size of the training set equal to the number of the training images and $[m \times n]$ is the size of the image of the training set. The latter design would allow parallel implementation, since all the training images could be input through the NNET in parallel due to the parallel input sources. However, to allow easier implementation, we chose the former design of the NNET.

3 Architecture and Implementation of the NNET

The hybrid optical neural network is implemented as a feedforward multi-layer architecture trained with a backpropagation algorithm. It has a single input source of input neurons equal to the size of the training image in vector form. In effect, for the training image $X_{i=1\ldots N}$ of size $[m \times n]$, there are $[m \times n]$ input neurons in the single input source. It is emphasised that the training set images of the NNET consist of several images of the object at different poses and at regular angular increments. Each image is passed through the NNET in gray-scale form, without an explicit amplitude scale normalisation. The actual size of the training set images is $[256 \times 256]$ in matrix form (i.e. m=n=256) or of size $[1 \times 256 \times 256] = [1 \times 65{,}536]$ in vector form. The input weights are initially fully-connected from the input layer to the hidden layers *but* do not stay fully connected during training. There are N_{iw}

input weights proportional to the size of the training set. The number of the hidden layers, N_l, is equal to the number of the images of the training set, N:

$$N = 1, 2, 3, \cdots, i \tag{6}$$

$$N_l = N \tag{7}$$

Each hidden layer consists of a single neuron. The layer weights are fully connected to the output layer. Since there is a single output neuron, the number of the layer weights, N_{lw}, equals the number of the training images, N:

$$N_{lw} = N \times N_{opn} \tag{8}$$

where $N_{opn} = 1$ is the number of the output neurons. There are bias connections to each one of the hidden layers:

$$N_b = N \tag{9}$$

where N_b is the number of the bias connections. There is one target connection from the single output neuron of the output layer.

The initial values of the input weights, the layer weights and the biases are based on the Nguyen-Widrow[29] initialisation algorithm[30]. The transfer function of the hidden layers is set as the Log-Sigmoidal function, whereas the transfer function of the output layer is linear. When a new training image is presented to the NNET we leave connected the input weights of only one of the hidden neurons. In order not to upset any previous learning of the rest of the hidden layer neurons we do not alter their weights when the new image is input to the NNET. It is emphasised that there is no separate feature extraction stage applied to the training set images. To achieve faster learning we used a modified steepest descent backpropagation algorithm based on heuristic techniques[31]. The adaptive training algorithm updates the weights and

bias values according to the gradient descent momentum and an adaptive learning rate:

$$\Delta w(t,t+1) = \mu \times \Delta w(t-1,t) + \alpha \times \mu \times \frac{\Delta P_f}{\Delta w(t+1,t)} \tag{10}$$

$$\Delta b(t,t+1) = \mu \times \Delta b(t-1,t) + \alpha \times \mu \times \frac{\Delta P_f}{\Delta b(t+1,t)} \tag{11}$$

$$\alpha = \left\{ \begin{array}{l} \alpha = \alpha + \varepsilon \quad \text{if } \Delta P_f < 0 \\ \alpha = \text{no change} \quad \text{if } 0 < \Delta P_f \ \&\& \ \Delta P_f > \max(P_f) \\ \alpha = \alpha - \varepsilon \quad \text{if } \Delta P_f > \max(P_f) \end{array} \right\} \tag{12}$$

where Δw is the update of the input and layer weights, Δb is the update of the biases of the layers, t is the iteration index of the NNET and μ is the momentum constant. The momentum allows the network to respond not only to the local gradient, but also to recent trends in the error surface. It acts like a low-pass filter by removing the small features in the error surface of the NNET. The momentum allows the network not to get stuck in a shallow local minimum, but to slide through such a minimum. P_f is the performance function, usually set as being the mean square error (*mse*) and ΔP_f is the derivative of the performance function[31]. The learning rate is indicated with the letter α. It adapts iteratively based on the derivative of the performance function ΔP_f. In effect, if there is a decrease in the ΔP_f, then the learning rate is increased by the constant ε. If ΔP_f increases but the derivative does not take a value higher than the maximum allowed value of the performance function, $\max(P_f)$, then the learning rate does not change. If ΔP_f increases more than $\max(P_f)$, then the learning rate decreases by the constant ε. The layer weights remain connected with all the hidden layers for all the training set and throughout all the training

session. The NNET typically converges for each training image in under one hundred iterations, requiring only few tens of milliseconds CPU time on a 2.5GHz Pentium 4 based PC. Also, due to the generalization properties of a NNET structure, the number of the training images decreases, in comparison to the typical number of images in the training set of linear combinatorial filters (such as the SDF filter).

The motivation to use the hybrid optical neural network filter comes from the ability of neural network methods to provide a non-linear interpolation of the training set poses of the test objects[32]. The enhancement of the non-linear properties of the NNET in the SDF filter synthesis procedure is most clearly seen from eqn. (14). We extract the layer and input weights and compute their dot product, $\Gamma_{i=1\cdots N}$. We then calculate the dot product of $\Gamma_{i=1\cdots N}$ with the corresponding training image, $X_{i=1\cdots N}$. Assume we have N training images of size $[\,m \times n\,]$ and we represent each training image with $x_{\,i=1\cdots N}(m, n)$, the weights from neuron ι to neuron κ with $w^{x_i}_{mn}$ and the layer weights with $l^{x_i}_{mn}$. Then:

$$
\begin{aligned}
\Gamma_{i=1\cdots N} &= W^{x_i} \cdot L^{x_i} \\
&= \begin{bmatrix} w_{11}^{x_i} & w_{12}^{x_i} & \cdots & w_{1n-1}^{x_i} & w_{1n}^{x_i} \\ w_{21}^{x_i} & w_{22}^{x_i} & \cdots & w_{2n-1}^{x_i} & w_{2n}^{x_i} \\ \vdots & & & & \\ w_{m1}^{x_i} & w_{m2}^{x_i} & \cdots & w_{mn-1}^{x_i} & w_{mn}^{x_i} \end{bmatrix} \cdot \begin{bmatrix} l_{11}^{x_i} & \cdots & l_{1q}^{x_i} \\ l_{21}^{x_i} & \cdots & l_{2q}^{x_i} \\ \vdots & & \\ l_{n1}^{x_i} & \cdots & l_{nq}^{x_i} \end{bmatrix} \\
&= \begin{bmatrix} w_{11}^{x_i} l_{11}^{x_i} + w_{12}^{x_i} l_{21}^{x_i} + \cdots + w_{1n}^{x_i} l_{n1}^{x_i} & \cdots & w_{11}^{x_i} l_{1q}^{x_i} + w_{12}^{x_i} l_{2q}^{x_i} + \cdots + w_{1n}^{x_i} l_{nq}^{x_i} \\ w_{21}^{x_i} l_{11}^{x_i} + w_{22}^{x_i} l_{21}^{x_i} + \cdots + w_{2n}^{x_i} l_{n1}^{x_i} & \cdots & w_{21}^{x_i} l_{1q}^{x_i} + w_{22}^{x_i} l_{2q}^{x_i} + \cdots + w_{2n}^{x_i} l_{nq}^{x_i} \\ \vdots & & \\ w_{m1}^{x_i} l_{11}^{x_i} + w_{m2}^{x_i} l_{21}^{x_i} + \cdots + w_{mn}^{x_i} l_{n1}^{x_i} & \cdots & w_{m1}^{x_i} l_{1q}^{x_i} + w_{m2}^{x_i} l_{2q}^{x_x} + \cdots + w_{mn}^{x_x} l_{nq}^{x_x} \end{bmatrix}
\end{aligned}
\tag{13}
$$

$$
S_{i=1\cdots N} = \Gamma_{i=1\cdots N} \cdot X_{i=1\cdots N}(m,n) \tag{14}
$$

where W^{x_i} and L^{x_i} are the matrices of the input and layer weights. $w_{mn}^{x_x}$ are the input and layer weights from the input neuron of the input vector element at row m and column n to the associated hidden layer for the training image $x_{i=1\cdots N}(m,n)$. $l_{nq}^{x_i}$ are the input and layer weights from the hidden neuron of the layer vector element at row m and column n to the associated output neuron. In our case, q = 1 since the output layer has only one neuron. $S_{i=1\cdots N}(m,n)$ is the transformed image calculated from the dot product of $\Gamma_{i=1\cdots N}$ with the corresponding training image $X_{i=1\cdots N}(m,n)$.

If we do not put a hard constraint on the correlation peak heights generated by the filter, and add the newly transformed images $S_{i=1\cdots N}$, then we can synthesise the *unconstrained* hybrid optical neural network (*U-HONN*) filter in the spatial domain:

$$U - HONN = \sum_{i=1\cdots N}^{N} S_i \left(x, y \right)$$

or in the frequency domain the same equation is rewritten as:

$$\boldsymbol{U - HONN = \sum_{i=1\cdots N}^{N} S_i(x,y)} \tag{14}$$

Based on a similar technique to SDF filter synthesis[10] we can constrain the correlation peak of the HONN filter:

$$C - HONN = \sum_{i=1\cdots N}^{N} a_i \cdot S_i(m,n)$$

$$= a_1 \cdot S_1(m,n) + a_2 \cdot S_2(m,n) + \cdots + a_N \cdot S_N(m,n)$$

$$= a_1 \cdot \left(\Gamma_1 \cdot X_1(m,n) \right) + a_2 \cdot \left(\Gamma_2 \cdot X_2(m,n) \right) + \cdots + a_N \cdot \left(\Gamma_N \cdot X_N(m,n) \right)$$

or the same equation in the frequency domain is rewritten as:

$$\boldsymbol{C - HONN = \sum_{i=1\cdots N}^{N} a_i \cdot S_i(m,n)} \tag{15}$$

This is the filter's transfer function. The HONN[26] filter is composed of a non-linear space domain superposition of the training set images; the multiplying coefficient, $a_i \cdot \Gamma_i$, now becomes a function of the input weights and the layer weights, rather than a simple linear multiplying constant as used in a conventional SDF filter. Thus, the non-linear HONN filter is inherently shift invariant and it may be employed in an optical correlator as would a linear superposition SDF type filter. It may be used as a space domain function in a joint transform correlator architecture or be Fourier

transformed and used as Fourier domain filter in a 4-f Vander Lugt type optical correlator[33].

4 Computer Simulation

4.1 *Test Data*

A data set of input images for the overall HONN filter was constructed of an S-type Jaguar car model at 10° increments of out-of-plane rotation at an elevation angle of approximately 45°. Another set of images was constructed for the Police car model Mazda Efini RX-7 at the same elevation angle to serve as the out-of-class data for discrimination tests (see Fig. 4. (a) and (b)). A third data set consisted of the background images of typical car parks (see Fig. 5) and the images of the S-type car model and the Mazda RX-7 car model added in the background scene. The size of all the images was $[\,256 \times 256\,]$ and the grey-scale bitmap format is used. All the images of the training set for the NNET are concatenated row by row into a vector of size $[\,1 \times (\,256 \times 256\,)\,]$ prior to input to the neural net. Normally this size of image is not possible, since to be implemented by enough input and layer weights:

$$\begin{aligned} N_{iw} &= 10 \times [\,256 \times 256\,] \\ &= 10 \times 65{,}536 \\ &= 655{,}360 \end{aligned} \tag{16}$$

This would be impossibly large since for a training set of $N = 10$ of individual vector size $[\,256 \times 256\,]$ there would be more than half-a-million input weight connections needed. Thus the selective weight connection architecture is employed to overcome this problem. Also, introducing into the NNET training session the heuristic training algorithm with momentum and an adaptive learning rate has

speeded up the learning phase and reduced the memory size needed to complete fully the training session.

It was found experimentally that by choosing different values of the classification levels for the true class Cl_T and false class Cl_F objects, one can control the HONN filter's behaviour to suit different application requirements. Thus we define:

$$\Delta Cl = \left| Cl_T - Cl_F \right| \tag{17}$$

where ΔCl is the absolute distance of the classification levels between the true class objects and the false class objects. For example, when $Cl_T = +80$ and $Cl_F = -80$, the resulting filter behaves more like a high-pass biased filter, which generally gives sharp peaks and good clutter suppression but is more sensitive to intra-class distortions. If $Cl_T = +10$ and $Cl_F = -10$, then the filter behaves more like a MVSDF filter[34] with relatively good intra-class distortion invariance but producing broad correlation peaks. In effect, when ΔCl increases, the HONN filter possesses better discriminatory properties but when ΔCl decreases the HONN filter has better generalising properties. Fig. 6. (a) shows the composite image of the SDF filter and Fig. 6. (b) shows the composite image generated by the SDF-MACH filter for the training images over the orientation range $[20\ 40\ 60\ 80]$ degrees. Fig. 6. (c) shows the composite image of the HONN filter for the in-class training images over the same orientation range with the setting $T_{true} = +80$ and $T_{false} = -80$ Fig. 6. (d) shows the composite image of the HONN filter for setting the in-class training images over the same range with $T_{true} = +280$ and $T_{false} = -280$.

Fig. 7. (a) and (b) show the correlation plane isometric of the HONN filter for the training set over the orientation range $[20\ 40\ 60\ 80]$ degrees and the test image of the 80° orientation view, when the targets were set to $T_{true} = +80$ and $T_{false} = -80$. Fig. 7. (c) and (d) show the correlation plane isometric and the grey-level image of the correlation plane of the HONN filter for the same training set and the test image of the 80° orientation view, but now the targets were set to $T_{true} = +280$ and $T_{false} = -280$. It is clear that the filter becomes more discriminative when we increase the absolute distance ΔCl of classification levels between the true class objects and the false class objects. Thus increasing ΔCl leads to an increased emphasis of the high spatial frequency content of the composite images comprising the filter, which in turn leads to a more localised correlation response.

4.2 *Results*

4.2.1 *Peak Sharpness and Detectability*

Several tests were carried out to test the performance of the hybrid optical neural network in terms of peak sharpness and general detectability[1–3] for the out-of-plane orientation considered. The aim of the tests was to assess the ability of the filter to detect non-training in-class images that are orientated at an intermediate angle of view between the training images. The training set consisted of images over a distortion range $[20\ 80]$ degrees at increments of 20°. To compare the filter with an SDF-type filter, we used the SDF- MACH filter, to constrain the correlation peaks generated by the MACH filter to achieve better class separation of the different objects. We tested the SDF-MACH filter and the HONN filter with the object's intermediate car poses over the same range at 10° increments. We added two randomly chosen intermediate car poses, at 130° and at 140°, in the training set of the

HONN filter to create a false class. The Target of the false class object is $T_{false} = -40$ and of the true class object is $T_{false} = +40$. Both filters have no information on the non-training, intermediate car images in the construction of their composite images. We constrained the correlation peaks in the constraint matrix of both filters to be $+1$ for the images of the true class object and 0 for the images of the false class object. We employed SDF-MACH filter parameters[9–11] set to: $\alpha = 0.002$, $\beta = 0.01$ and $\gamma = 0.1$.

Fig. 8. plots the correlation-peak height for each input image for both the HONN and the SDF-MACH filters. Both filters constrain the correlation peak height, using the constraint matrix, for the in-class training set images and the out-of-class training set images to $+1$ and 0, respectively. From the plot of Fig. 8. it is observed that the HONN filter and the SDF-MACH filter are invariant to the out-of-plane rotation, since they both have produced consistent correlation peaks for both in-class training and non-training images. However, the graph indicates better out-of-plane rotation invariance for the hybrid optical neural network filter. The consistency of the correlation peak values produced from the HONN filter demonstrate the filter's ability to interpolate well between the intermediate car poses at 10° increments. The SDF-MACH filter produced correlation peak values that fell to a greater degree around the intermediate non-training images as compared to the HONN filter'sresponses. The improvement in the HONN filter's interpolation ability must result from the non-linear function applied in the construction of the composite image comprising the SDF filter formed from the NNET output images.

A second set of tests was carried out to evaluate the deterioration of the filters' correlation response for the intermediate car poses. Fig. 9. shows the normalised

PCE values for both filters. The hybrid optical neural network filter and the SDF-MACH filter were normalised to the maximum value recorded for both filters, which was the 80° orientation test image. From the graph, it can be observed that the SDF-MACH filter and the HONN filter produced PCE values for the intermediate non-training images close to those produced by the training car images. In effect, both filters maintain correlation peak sharpness for the in-class training images. However, the SDF-MACH filter demonstrates the highest PCE values, producing sharper peaks and fewer sidelobes evident on the correlation plane. The values of the HONN filter's PCE values are, however quite acceptable.

From our experiments, we found that the change of the absolute distance of the classification levels, ΔCl, between the true class objects and the false class objects affects the PCE values produced from the HONN filter. When we increased the distance, the PCE values increased and, conversely, when we decreased it, the PCE values produced decreased. This is because as we increase the value of ΔCl, the hybrid optical filter increases its discrimination ability, but when we decreased it the filter increases its generalisation ability and so its discrimination ability deteriorates. The correlation peak broadens as ΔCl is decreased i.e. there is very similar trade-off in discrimination and generalisation ability to that found in SDF type correlation filters.

4.2.2 Distortion Range of the Filter

A second group of test sets was carried out to evaluate the distortion range of the hybrid optical neural network filter. The training set consisted of images for a distortion range over 0° to 90°. For comparison with an SDF-filter, we used the SDF-MACH filter, which we trained with the same training set. We used several

smaller test sets for both filters, which consisted of two in-class training images at a widely separated angle within the range $\left[10^\circ\ 20^\circ\ 30^\circ\ 40^\circ\ 50^\circ\ 60^\circ\ 70^\circ\ 80^\circ\ 90^\circ\right]$ and a third non-training in-class image lying on the bisector angle of the two in-class training images. Fig. 10. shows the reference angle of view image at Θ_0, the first training car pose image at Θ_1, the second training car pose image at Θ_2 and the intermediate non-training car pose image at Θ_3, where $\hat{\Theta}_3 \in \left[5^\circ, 45^\circ\right]$ i.e. $\Theta_3 = \left[5^\circ\ 10^\circ\ 15^\circ\ 20^\circ\ 25^\circ\ 30^\circ\ 35^\circ\ 40^\circ\ 45^\circ\right]$ We define $\Delta\Theta$ to be the absolute distance between the two training set images, and the angle Θ_3 to be on the bisector angle of Θ_1 and Θ_2. Thus:

$$\begin{aligned}
&\Delta\Theta = \left[10^\circ\ 20^\circ\ 30^\circ\ 40^\circ\ 50^\circ\ 60^\circ\ 70^\circ\ 80^\circ\ 90^\circ\right] \\
&\Delta\Theta = \left|\Theta_1 - \Theta_2\right| \\
&\Theta_3 = \frac{\Delta\Theta}{2} \\
&\Delta\Theta_1 = \left|\Theta_1 - \Theta_0\right| \\
&\Delta\Theta_2 = \left|\Theta_2 - \Theta_0\right| \\
&\Theta_0 = 0^\circ
\end{aligned} \tag{18}$$

As for the first group of tests used for evaluating the peak sharpness and the detectability of the filter, we added three randomly chosen training images, namely at orientations of $110^\circ, 130^\circ$ and 140°, in the training set of the hybrid optical neural network filter which fall inside the false class. The Targets of the true class objects and of the false class objects are kept the same as before, i.e. $T_{true} = +40$ and $T_{false} = -40$. None of the filters have any information built into them on the test images of the intermediate car poses. We kept the same constrained values in the constraint matrix of both filters for the in-class images and for the out-of-class images. The SDF-MACH filter was tuned to the same values as for the first group of

test sets, i.e. $\alpha = 0.002$, $\beta = 0.01$ and $\gamma = 0.1$, in order not to change the SDF-MACH filter's behaviour and to obtain, if possible, an overall comparison to the HONN filter.

Fig. 11. shows the correlation-peak height for each input image for both the HONN filter and the SDF-MACH filter. It is apparent that both filters have good performance in recognising all the intermediate car poses of the test set. The correlation-peak height of the in-class input images up to $\hat{\Theta}_3 = 15^\circ$, that are intermediate between two training images, lie within a band of greater than 90% of the pre-specified peak-height constant in the constraint matrix *C* for the HONN filter and more than 40% of the pre-specified peak-height constant in the constraint matrix *C* for the SDF-MACH filter. From the graph it can be observed that both filters demonstrated some degree of invariance to rotation over a range of $\hat{\Theta}_3 \in [\, 5^\circ, 45^\circ \,]$. The correlation peak-height of the in-class test set images which are intermediate between the two training images fall inside a band of greater than 70% of the specified peak-height constant in the constraint matrix *C* for almost all the test set images of the HONN filter, whereas there is a clear variation in the peak-heights from that specified for the training in-class images of the SDF-MACH filter. Thus it appears that the non-linear interpolation of the training set images of the intermediate car poses, inherent in the hybrid optical neural network filter, has resulted in less peak deterioration when Θ_3 gradually increased, and succeeded in keeping it almost a constant peak-height for the range $\Theta_3 = 40^\circ$.

4.2.3 Discrimination Ability

The third group of test sets was carried out to evaluate the discrimination ability of the HONN filter. The main aim of the tests was to discriminate between objects of

different classes while retaining invariance to in-class distortions. The training set consisted of images of the Jaguar S-type for a distortion range over 20° to 70° at 10° increments. Again, for comparison we used the SDF-MACH filter and we trained it with the same set of images. The test set, used in both filters, consisted of one training image at 40° orientation of the Jaguar S-type and a second image of the out-of-class Police RX-7 at the same angle of view. This time we only added two images at 130° and 140°, for the false class of the HONN filter and none for the SDF-MACH filter. We did not have to constrain the false class images of the objects to zero peak-height in the construction of both filters' composite images. The Target of the false class object is, $T_{false} = -40$, and the Target of the true class object is, $T_{true} = +40$. Both filters have no built-in information on the test images of the in-class object or the out-of-class object. The values of the parameters of the SDF-MACH filter were kept constant $(\alpha = 0.002, \beta = 0.01, \gamma = 0.1)$.

Table 1 was drawn from the tests we performed for the SDF-MACH and the HONN filters. The third column of the table contains the values taken for the in-class training image and the fourth column contains all the values taken for the out-of-class training image.

It can be observed from the Table 1 entries that both filters demonstrated good discrimination ability between the two objects, the Jaguar S-type car and the RX-7 Police patrol car. The SDF-MACH filter produced approximately 92% or more class separation, whereas the HONN filter produced 72% or higher class separation.

4.2.4 Clutter Tolerance

A fourth group of tests was carried out to evaluate the tolerance of the filter to background clutter in the input scene by the insertion of training images and non-

training out-of-class images into different car park scenes. The training set consisted of images of the Jaguar S-type for a distortion range over 20° to 70° degrees at 10° increments. As with the previous test sets, we used the SDF-MACH filter for comparison and we trained it with the same training set of images. We added three background images of typical empty car park scenes in the training set of the hybrid optical neural network to fall inside the false class. We had to constrain the false class images of the objects to zero peak-height constraint in the construction of both filters' composite images. The test set, used in both filters, consisted of one of the training in-class images inserted in a car park scene, a randomly chosen non-training in-class image of an intermediate car pose inserted in the same car park scene and one out-of-class non-training image of the RX-7 Police patrol car also inserted in the car park scene. The rest of the test set images were the same images of the Jaguar S-type and the RX-7 Police patrol car inserted in a second car park scene. The Targets of the true class objects were $T_{true} = +40$ and of the false class background images of the empty car park scenes were $T_{false\,1} = -40$. The false class image of the out-of-class non-training RX-7 Police car image was $T_{false\,2} = -40$. We kept the same values of the parameters for the Output Noise Variance (ONV), Average Similarity Measure (ASM) and the Average Correlation Energy (ACE) i.e. $\alpha = 0.002$, $\beta = 0.01$ and $\gamma = 0.1$, respectively.

Fig. 12. (a) and (b) show the in-class Jaguar S-type car and the out-of-class RX-7 Police patrol car, resized and inserted into the first background scene of a car park. Fig. 13. shows the isometric correlation planes of the in-class training view Jaguar S-type car image at 40° (resized and inserted into the first car park scene). Fig. 13. (a) shows the HONN filter response and (b) that of the SDF-MACH filter.

The correlation planes of the out-of-class non-training RX-7 Police patrol car at 40°, are shown in (c) for the HONN filter and (d) for the SDF-MACH filter. Fig. 14 show the isometric correlation planes of the in-class training view Jaguar S-type car image at 40° and of the in-class non-training Jaguar S-type intermediate car pose at 55°, respectively.

From the above results, it is apparent that the HONN filter was able to detect and classify correctly the Jaguar S-type car pose at 40° and suppress the background clutter scene. The HONN filter rejected the out-of-class RX-7 Police patrol car in the background scene. Both filters performed better with the in-class training set Jaguar car image and their performance deteriorated with the intermediate in-class in-training Jaguar S-type car image. The hybrid optical neural network filter as well as the SDF-MACH filter were able to detect more easily the object in the first background than in the second background scene. This is due to the highercomplexity of the second car park scene in comparison to the first. Table 2 shows the values of the Target-to-Clutter Ratio (TCR) of the two filters. It is clear that the HONN filter demonstrates higher TCR values for the two tested background scenes than the SDF-MACH filter. The SDF-MACH filter response deteriorates more when detecting the intermediate Jaguar S-type car image at 55° in both background scenes.

5 Conclusion

The combination of an artificial neural network (NNET) with the linear combination filter, such as the synthetic discriminant function (SDF) design method, results in the hybrid optical neural network (HONN) filter. The motivation for the HONN filter's creation was to combine the shift invariance inherent in the

combinatorial filter synthesis process and the non-linear superposition of the training set generated by neural network methods. The HONN filter so created is shift-invariant but incorporates a non-linear space domain superposition of the training set making it suitable for direct implementation in an optical correlator. For a filter to be of practical use it should demonstrate good detectability, good discrimination ability and invariance to distortions in the input scene.

In this paper several tests were carried out to investigate the performance of the HONN filter, giving emphasis to its performance within cluttered images. The filter has shown a good quality correlation response to the non-training in-class images at angles of view intermediate to the training set images. The HONN filter has kept almost constant, good sharpness, correlation peaks of over 93% of the correlation peak height value set in the filter's constraint matrix *C* for the intermediate car poses at 10° increments. Further tests have demonstrated the high distortion range of the HONN filter. It has produced good quality correlation peak-height values for all the in-class test set images which are intermediate between the two training images over the range of $\hat{\Theta}_3 \in [5^{\circ}, 45^{\circ}]$ falling inside a band of greater than 70% of the specified peak-height constant value set in the constraint matrix *C*. The non-linear interpolation of the training set images of the intermediate car poses achieved in the HONN filter resulted in less peak deterioration and high distortion range. The filter was found to have good discrimination ability. The HONN filter successfully rejected the out-of-class RX-7 Police patrol car and has detected and classified correctly the Jaguar S-type car pose within the background clutter scenes. It has kept overall high TCR values for all the tested cluttered images of the in-class training set and in-class intermediate car poses.

By changing the absolute distance of the classification levels, ΔCl, we have controlled the correlation plane peak and the filter's behaviour has varied from highly discriminating to more generalising properties. The results presented have been very promising for the realisation of the HONN filter in optical implementations for real-time applications. It should be possible, using the HONN filter, to incorporate fewer views in the training set for a given distortion range. This makes its optical implementation easier, since this is limited by the dynamic range of the input transducer.

Acknowledgement

The author would like to thank everyone in the 'Laser and Photonics Systems Research Group' to contribute for the completion of this research paper. The author wishes the work presented in here and any related material to benefit and not harm the human life: *"Science without virtue is immoral science (Plato)"*.

References

1. B. V. K. Vijaya Kumar and L. Hasserbrook, "Performance measures for correlation filters", *Applied Optics Vol. **29**, No. **20*** 2997-3001 (1990).
2. Ph. Refregier, "Filter design for optical pattern recognition: multicriteria optimization approach", *Optics Letters, Vol. **15**, No. **15,*** 854-856 (1990).
3. Ph. Refregier. "Optimal trade-off filters for noise robustness, sharpness of the correlation peak, and Horner efficiency", *Optics Letters, Vol, **16**, No. **11,*** 829-831 (1991).
4. D. Casasent, "Unified synthetic discriminant function computational formulation", *Applied Optics, Vol. **23**, No. **10**,* 1620-1627 (1984).
5. H. J. Caulfield, "Linear combinations of filters for character recognition: a unified treatment", *Applied Optics, **19 (23)**,*3877-3878 (1980).

6. H. J. Caulfield and W. T. Maloney, “Improved discrimination in optical character recognition”, *Applied Optics* ***08 (11)*** 2354-2356 (1969).

7. D. R. Selviah, J. E. Midwinter, A. W. Rivers and K. W. Lung, “Correlating matched filter model for analysis and optimisation of neural networks”, *IEE Proceedings,* ***136 (3)***, 143-148 (1989).

8. T.-H. Chao, G. Reyes and Y. Park, “Grayscale Optical Correlator”, *SPIE Vol.* ***3386*** 60-64 (1998).

9. H. Zhou and T.-H. Chao, “MACH filter synthesizing for detecting targets in cluttered environment for gray-scale optical correlator”, *SPIE Vol.* ***3715*** 394-398, Florida (1999).

10. A. Mahalanobis, B. V. K. Vijaya Kumar, S. Song, S. R. F. Sims, J. F. Epperson, “Unconstrained correlation filters”, *Applied Optics, Vol.* ***33***, 3751-3759 (1994).

11. A. Mahalanobis and B. V. K. Vijaya Kumar, “Optimality of the maximum average correlation height filter for detection of targets in noise”, *Optical Engineering* ***36 (10)***, 2642-2648 (1997).

12. B. V. K. Vijaya Kumar, “Tutorial survey of composite filter designs for optical correlators”, *Applied Optics, Vol.* ***31****, No.* ***33****,* 4773-4801 (1992).

13. D. L. Flannery, “Optimal trade-off distortion-tolerant constrained modulation correlation filters”, *JOSA A, Vol.* ***12****, No.* ***1*** 66-72 (1995).

14. N. Fahrat, D. Psaltis, A. Prata, and E. Paek, “Optical implementation of the Hopfield model”, *Applied Optics* ***24****,* 1469-1475 (1985).

15. B. H. Soffer, G. H. Dunning, Y. Owechiko, and E. Marom, “Associative holographic memory with feedback using phase conjugating mirrors”, *Optics Letters* ***11*****,** 118-120 (1986).

16. T.-H. Chao, “A spatial light modulator based optoelectronic associative memory”, in *Spatial Light Modulators and Applications, Vol.* ***14*** of 1990 OSA Technical Digest Series (Optical Society of America, Washington, D. C., 1990), 28-31.

17. F. T. S. Yu and T. Lu, “Self-learning optical neural network”, in *Spatial Light Modulators and Applications, Vol.* ***14*** of 1990 OSA Technical Digest Series (Optical Society of America, Washington, D. C., 1990), 24-27.

18. D.P. Casasent, L.M. Neiberg, “Classifier and shift-invariant automatic target recognition neural networks”, *Neural Networks, Vol.* ***8****,*.1117-1129, (1995).

19. M. A. Neifeld and D. Psaltis, “Optical implementations of radial basis classifiers”, *Applied Optics Vol.* ***32****, No.* ***08****,* 1370-1379 (1993).

20. H. Toyoda, N. Mukohzaka, Y. Suzuki and M. Ishikawa, “Adaptive optical processing system with

optical associative memory", *Applied Optics Vol. **32**, No. **08**,* 1354-1358 (1993).

21. T.-H. Chao and W. W. Stoner, "Optical implementation of a feature-based neural network with application to automatic target recognition", *Applied Optics, Vol. **32**, No. **8*** 1359-1369 (1993).

22. K. Fukushima, S. Miyake, and T. Ito, "Neocognitron: A neural network model for a mechanism of visual pattern recognition", *IEEE Trans. Syst. Man. Cybernetics, **SMC-13**,* 826-834 (1983).

23. A. Talukder and D. Casasent, "General methodology for simultaneous representation and discrimination of multiple object classes", *Opt. Eng. **37 (3)*** 904-913 (1998).

24. A. Talukder and D. Casasent, "Nonlinear features for Product Inspection", *SPIE Vol. **3715**,* Florida, 32-43 (1999).

25. D. L. Swets and J. J. Weng, "Using discriminant eigenfeatures for image retrieval", *IEE Trans. Pattern Anal. Mach. Intell. **18 (8)*** 831-836 (1996).

26. I. Kypraios, R. C. D. Young, P. Birch, C. Chatwin, "A non-linear training set superposition filter derived by neural network training methods for implementation in a shift invariant optical correlator", *SPIE Vol.**5106*** 84-95 (2003).

27. L. S. Jamal-Aldin, R. C. D. Yound, and C. R. Chatwin, "Synthetic discriminant function filter employing nonlinear space-domain preprocessing on bandpass-filtered images", *Applied Optics Vol. **37**, No. **11*** 2051-2062 (1998).

28. L. S. Jamal-Aldin, R. C. D. Yound, and C. R. Chatwin, "In-class distortion tolerance, out-of-class discrimination and clutter resistance of correlation fitlers that employ a space domain non-linearity applied to wavelet filtered input images", *SPIE Vol. **3386*** 111-122 (1998).

29. D. Nguyen and B. Widrow, "The truck backer-upper: An example of self-learning in neural networks, *Proceedings of the International Joint Conference on Neural Networks, Vol. **2**,* 357-363 (1989).

30. D. Nguyen and B. Widrow, "Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights", *Proceedings of the International Joint Conference on Neural Networks, Vol. **3**,* 21-26 (1990).

31. M. T. Hagan, H. B. Demuth and M. H. Beale, "Neural Network Design", *Boston, MA: PWS Publishing* (1996).

32. I. Kypraios, R. Young, C. Chatwin, "An investigation of the non-linear properties of correlation filter synthesis and neural network design", *Asian Journal of Physics, Vol. **11**, No. **3*** (2003).

33. A.Vander Lugt, “Signal detection by complex spatial filtering”. *IEEE Transactions on Information Theory*, ***10***, 139-145 (1964).

34. B. V. K. Kumar, “Minimum Variance synthetic discriminant function”, *Journal of Optical Society of America* ***A3,*** 1579-1584 (1986).

List of figures

Fig. 1 Block Diagram of the Hybrid Optical Neural Network (HONN) Filter.
Fig. 2 The architecture of Design A of the artificial neural network (NNET) Block of the HONN Filter.
Fig. 3 The architecture of Design B of the NNET Block of the HONN filter.
Fig. 4 (a) the Jaguar S-type car model and (b) the Mazda Efini RX-7
Fig. 5 (a), (b), (c) and (d) The Background Car Park Scenes part of the third data set used in the experiments.
Fig. 6 (a) The composite image of the basic SDF filter for the training set over the range [20 40 60 80] of the car images, (b) the composite image of the SDF-MACH filter for the training set over the same range of the car images, (c) the composite image of the HONN filter over the same range of the training images for $T_{true} = +80$ and $T_{false} = -80$ and (d) the composite image of the HONN filter over the same range of the training images for $T_{true} = +280$ and $T_{false} = -280$. All the images of the training set used are gray-scale of size $[256 \times 256]$.
Fig. 7 (a) and (b) The correlation plane isometric and grey-level image, respectively, of the HONN filter for the training set over the orientation range of $[20\ 40\ 60\ 80]$ degrees of the car images for the test image of 80° orientation. $T_{true} = +80$ and $T_{false} = -80$. (c) and (d) show the correlation plane isometric and the grey-level image of the HONN filter for the same training set and the same test image, but $T_{true} = +280$ and $T_{false} = -280$.
Fig. 8 The plot shows the correlation peak height, of the training set over the orientation range $[20\ 80]$ at increments of 20°, for the HONN and the SDF-MACH filters. The test set consisted of the car images over the same range, at increments of 10°. The correlation peak height of both filters for the training set in-class images and the out-of-class training images was constrained in the filters' constraint matrix to $+1$ and 0 respectively.
Fig. 9 The plot shows the peak correlation energy (PCE) over the orientation range $[20\ 80]$ at increments of 20° of the training set, for the HONN and the SDF-MACH filters. The test set consisted of the car images over the same range, at increments of 10°. The peak correlation energy values of both filters were normalised to the maximum value. The maximum value of both of the filters was recorded at 80° orientation.
Fig. 10 The diagram shows the reference angle, Θ_0, and the two in-class training images at the angles, Θ_1 and Θ_2. The test image is at the bisector angle at angle Θ_3.
Fig. 11 The graph plots the correlation peak height of the HONN filter and the SDF-MACH filter for the training set over the range angle Θ_3 from 5° to 45°. Any two in-class training images could be on any angle over the same range, $\hat{\Theta}_1, \hat{\Theta}_2 \in [0^\circ\ 90^\circ]$ and the test image was on the bisector (range) angle of the two in-class training images, i.e. $\Theta_3 = [5^\circ 10^\circ 15^\circ 20^\circ 25^\circ 30^\circ 35^\circ 40^\circ 45^\circ]$.
Fig. 12 (a) The in-class Jaguar S-type car inserted in the first background scene of a car park and (b) the out-of-class RX-7 Mazda Police patrol car inserted into the same background scene.
Fig. 13 The isometrics of the correlation planes of the in-class training view Jaguar S-type car image at 40°, resized and inserted into the first car park scene, (a) the HONN filter and (b) the SDF-MACH filter. The correlation planes of the out-of-class non-training RX-7 Police car at 40°, resized and inserted in the same car park scene, (c) the HONN filter and (d) the SDF-MACH filter.
Fig. 14 The isometrics of the correlation planes of the in-class non-training intermediate Jaguar S-type car pose at 55°, resized and inserted into the first car park scene, of (a) the HONN filter and (b) the SDF-MACH filter and resized and inserted into the second car park scene of (c) the HONN filter and (d) the SDF-MACH filter.

List of Tables

Table 1 Discrimination Ability (%) of the HONN and the SDF-MACH filter for the in-class training Jaguar S-type image at 40° and the out-of-class non-training Mazda RX-7 image at the same angle.
Table 2 Target-to-Clutter Ratio (TCR)

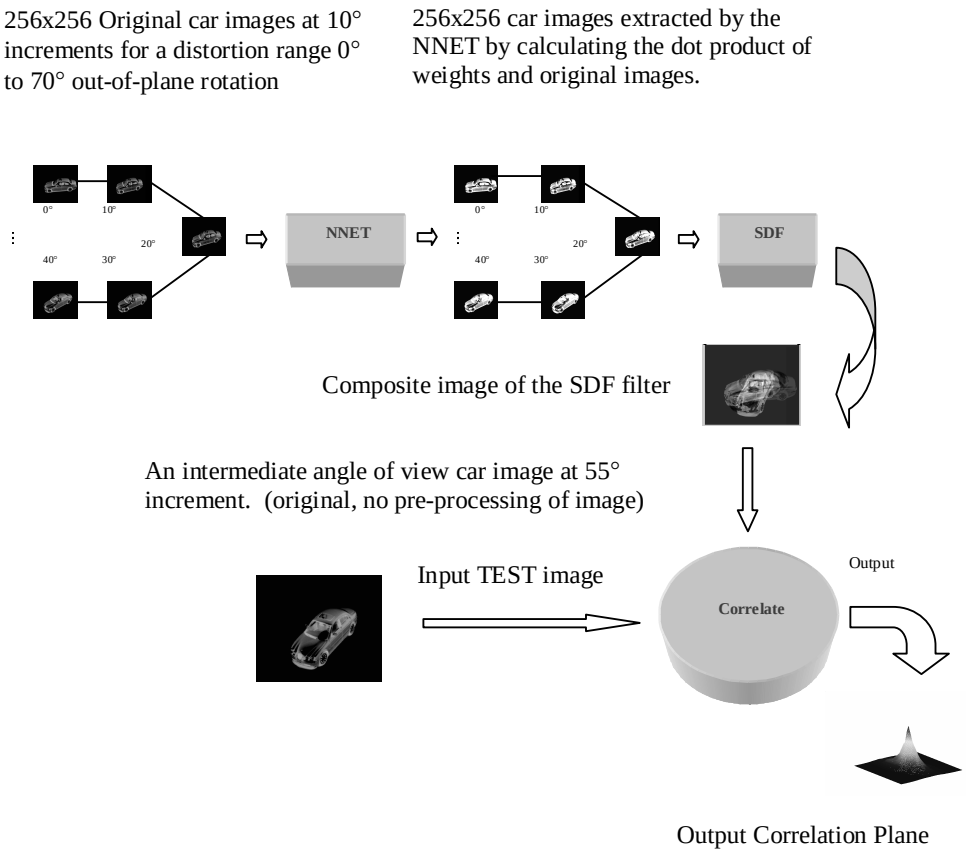


Fig. 1 Block Diagram of the Hybrid Optical Neural Network (HONN) Filter.

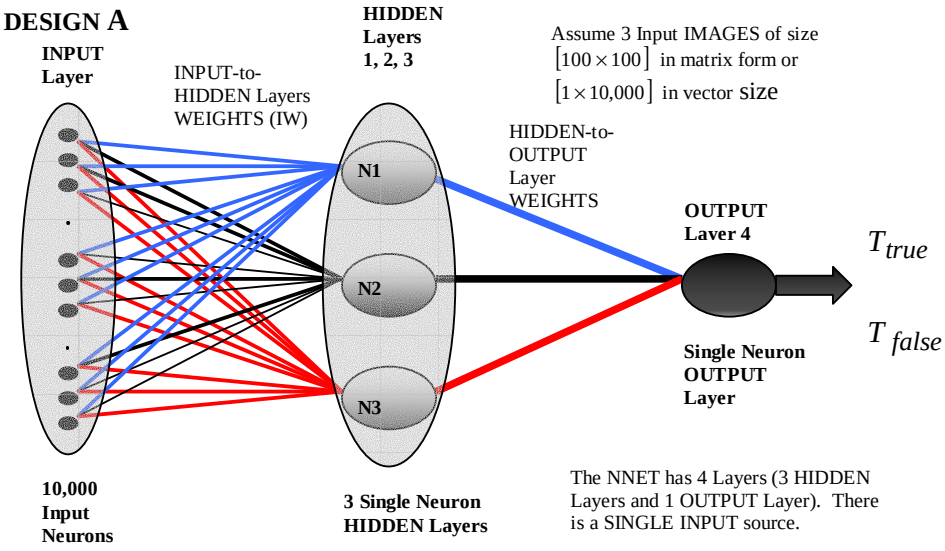


Fig. 2 The architecture of Design A of the artificial neural network (NNET) Block of the HONN Filter.

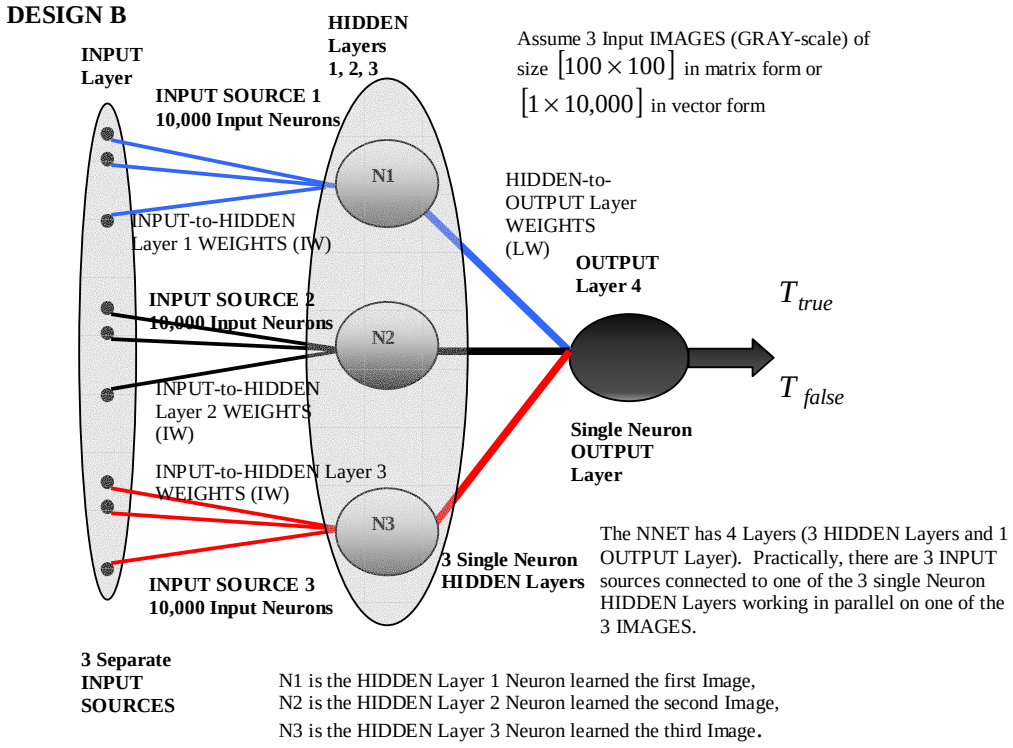


Fig. 3 The architecture of Design B of the NNET Block of the HONN filter.

(a) (b)

Fig. 4 (a) the Jaguar S-type car model and (b) the Mazda Efini RX-7

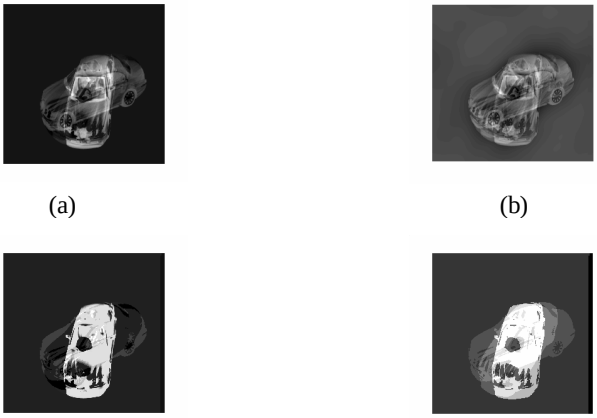
(a)

(b)

(c)

(d)

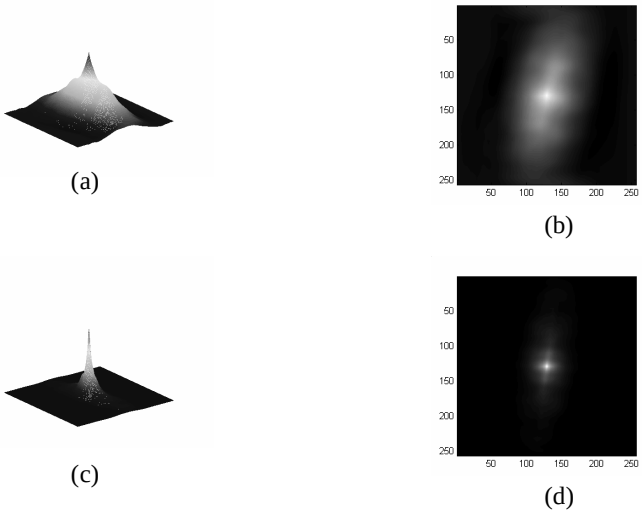
Fig. 5 (a), (b), (c) and (d) The Background Car Park Scenes part of the third data set used in the experiments.



(a) (b)

(c) (d)

Fig. 6 (a) The composite image of the basic SDF filter for the training set over the range [20 40 60 80] of the car images, (b) the composite image of the SDF-MACH filter for the training set over the same range of the car images, (c) the composite image of the HONN filter over the same range of the training images for $T_{true} = +80$ and $T_{false} = -80$ and (d) the composite image of the HONN filter over the same range of the training images for $T_{true} = +280$ and $T_{false} = -280$. All the images of the training set used are gray-scale of size $[256 \times 256]$.



(a) (b)

(c) (d)

Fig. 7 (a) and (b) The correlation plane isometric and grey-level image, respectively, of the HONN filter for the training set over the orientation range of $[20\ 40\ 60\ 80]$ degrees of the car images for the test image of 80° orientation. $T_{true} = +80$ and $T_{false} = -80$. (c) and (d) show the correlation plane isometric and the grey-level image of the HONN filter for the same training set and the same test image, but $T_{true} = +280$ and $T_{false} = -280$.

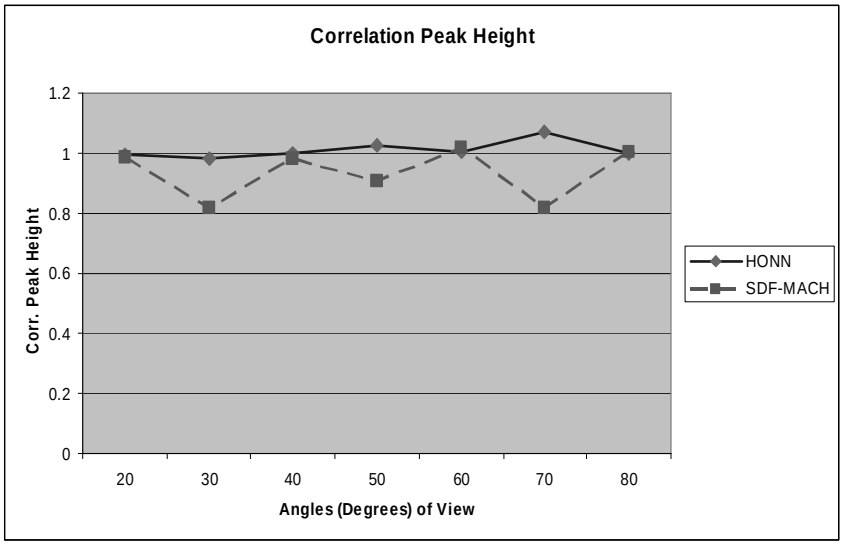


Fig. 8 The plot shows the correlation peak height, of the training set over the orientation range $[20\ 80]$ at increments of 20°, for the HONN and the SDF-MACH filters. The test set consisted of the car images over the same range, at increments of 10°. The correlation peak height of both filters for the training set in-class images and the out-of-class training images was constrained in the filters' constraint matrix to $+1$ and 0 respectively.

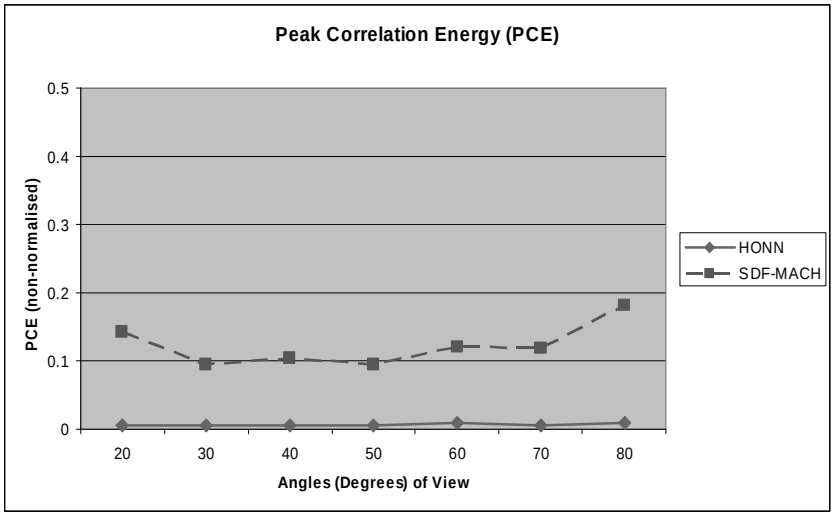


Fig. 9 The plot shows the peak correlation energy (PCE) over the orientation range $[20\ 80]$ at increments of 20° of the training set, for the HONN and the SDF-MACH filters. The test set consisted of the car images over the same range, at increments of 10°. The peak correlation energy values of both filters were normalised to the maximum value. The maximum value of both of the filters was recorded at 80° orientation.

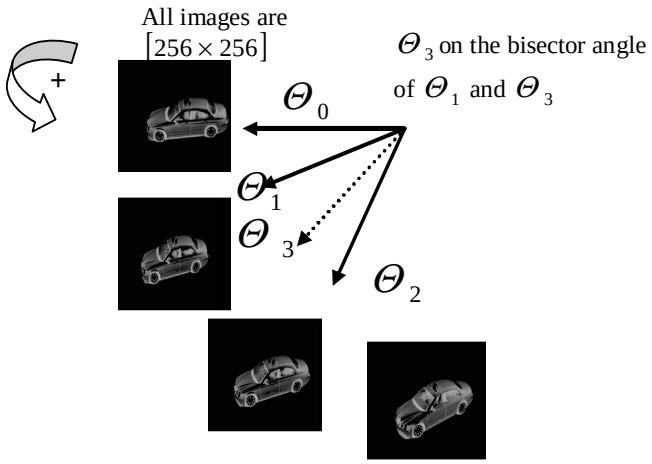


Fig. 10 The diagram shows the reference angle, Θ_0, and the two in-class training images at the angles, Θ_1 and Θ_2. The test image is at the bisector angle at angle Θ_3.

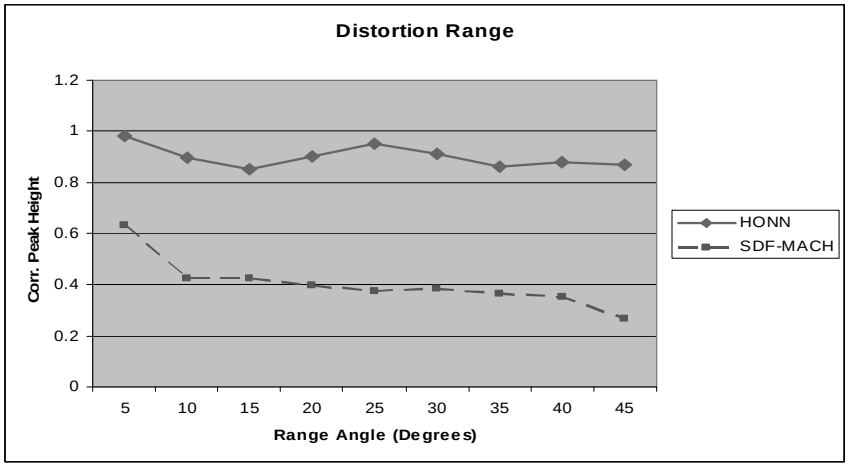


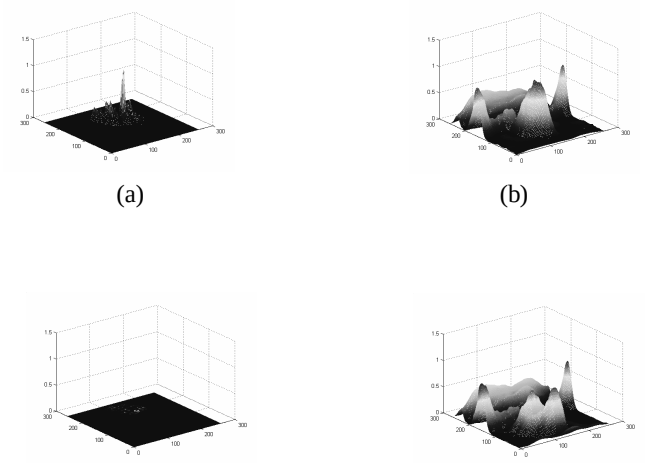
Fig. 11 The graph plots the correlation peak height of the HONN filter and the SDF-MACH filter for the training set over the range angle Θ_3 from 5° to 45°. Any two in-class training images could be on any angle over the same range, $\hat{\Theta}_1, \hat{\Theta}_2 \in \left[0^\circ \; 90^\circ\right]$ and the test image was on the bisector (range) angle of the two in-class training images, i.e. $\Theta_3 = \left[5^\circ 10^\circ 15^\circ 20^\circ 25^\circ 30^\circ 35^\circ 40^\circ 45^\circ\right]$.

Table 1 Discrimination Ability (%) of the HONN and the SDF-MACH filter for the in-class training Jaguar S-type image at 40° and the out-of-class non-training Mazda RX-7 image at the same angle.

			Correlation Peak Height	
Filter	**Parameter Values**	**Discr. Ability %**	**JAGUAR**	**MAZDA**
SDF-MACH	$\alpha = 0.002,\ \beta = 0.01,\ \gamma = 0.1$	92.07	0.9789	0.0776
HONN	ΔC_l	72.57	1.008	0.2765

(a) (b)

Fig. 12 (a) The in-class Jaguar S-type car inserted in the first background scene of a car park and (b) the out-of-class RX-7 Mazda Police patrol car inserted into the same background scene.



(a) (b)

(c) (d)

Fig. 13 The isometrics of the correlation planes of the in-class training view Jaguar S-type car image at 40°, resized and inserted into the first car park scene, (a) the HONN filter and (b) the SDF-MACH filter. The correlation planes of the out-of-class non-training RX-7 Police car at 40°, resized and inserted in the same car park scene, (c) the HONN filter and (d) the SDF-MACH filter.

(a) (b)

(c) (d)

Fig. 14 The isometrics of the correlation planes of the in-class non-training intermediate Jaguar S-type car pose at 55°, resized and inserted into the first car park scene, of (a) the HONN filter and (b) the SDF-MACH filter and resized and inserted into the second car park scene of (c) the HONN filter and (d) the SDF-MACH filter.

Table 2 Target-to-Clutter Ratio (TCR)

		Tolerance - to - Clutter (TCR)		
Filter	**BG Scenes**	**IN-CLASS**	**IN-CLASS**	**OUT-of-CLASS**
		Training	**Non-Training**	**Non-Training**
HONN	BG1	2.8445	2.0305	2.6557
HONN	BG2	4.9869	2.0584	3.7129
SDF-MACH	BG1	1.752	1.8801	1.5491
SDF-MACH	BG2	1.0208	1.0174	1.0322