We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS BOOK CITATION INDEX INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

15

Hybrid Optical Neural Network-Type Filters for Multiple Objects Recognition within Cluttered Scenes

Ioannis Kypraios
School of Engineering & Design, University of Sussex,
Falmer, Brighton, BN1 9QT
U.K.

1. Introduction

A robust invariant pattern detection and classification system needs to be able to recognise the object under any usual *a priori* defined distortions such as translation, scaling and in-plane and out-of-plane rotation (Wood, 1996) (see Fig. 1). Ideally, the system should be able to recognise (detect and classify) any complex scene of objects even within background clutter noise. This problem is a very complex and difficult one. Here, we will consider only non-deformable (*solid*) objects (Forsyth & Ponce, 2003). In effect, they maintain their form independent of any of the distortions just described. Early studies (Casasent & Psaltis, 1976) in trying to solve the invariant pattern recognition problem include the system based on a modified logarithmic Mellin transform (Grace & Spann, 1991; Sheng & Arsenault, 1986; Sheng & Lejeune, 1991). Other work (Mersereau & Morris, 1986) has focused on a system based on a circular harmonic filter (Hsu et al., 1982; Hsu & Arsenault, 1982, 1984) illuminated with white light illumination. (Jensen et al., 1987) have described an optical image pattern recognition system based on log-polar mapping (Bryngdahl, 1974; Cederquist & Tai, 1984) of a Fourier transformed input pattern to convert in-plane rotation and scale changes into shift properties. The system's implementation by a correlator has allowed translation invariance of the input pattern. Although, the real-time practical use of these systems has been superseded, useful concepts for future implementation of filters can be extracted from this work.

In literature, broadly, two main categories of pattern recognition systems exist. The first category consists of linear combinatorial-type filters (LCFs) (Stamos, 2001). Proper image analysis in the frequency domain is done with the help of Fourier Transformation (FT) (Lynn & Fuerst, 1998; Proakis & Manolakis, 1988). The second category consists of pure neural modelling approaches. (Wood, 1996) has given a brief but clear review of invariant pattern recognition methods. His survey has divided the methods into two main categories of solving the invariant pattern recognition problem. The first category has two distinct phases of separately calculating the features of the training set pattern to be invariant to certain distortions and then classifying the extracted features. The second one, instead of having two separate phases, has a single phase which parameterises the desired invariances and then adapts them. (Wood, 1996) has also described the integral transforms, which fall

under the first category of feature extractors. They are based on Fourier analysis, such as the multidimensional Fourier transform, Fourier-Mellin transform, triple correlation (Delopoulos et al., 1994) and others. Part of the first category is also the group of algebraic invariants, such as Zernike moments (Khotanzad & Hong, 1990; Perantonis & Lisboa, 1992), generalised moments (Shvedov et al., 1979) and others. Wood has given examples of the second category, the main representative of this category being based on artificial neural network (NNET) architectures. He has presented the weight-sharing neural networks (LeCun, 1989; LeCun et al. 1990), the high-order neural networks (Giles & Maxwell, 1987; Kanaoka et al. 1992; Perantonis & Lisboa, 1992; Spirkovska & Reid, 1992), the time-delay neural networks (TDNN) (Bottou et al., 1990; Simard & LeCun, 1992; Waibel et al., 1989) and others. Finally, he has included a third category or the miscellaneous group where it consists of the methods which cannot strictly be categorised under either the feature-extraction feature-classification approach or the parameterised approach. Such methods are image normalisation pre-processing (Yuceer & Oflazer, 1993) methods for achieving invariance to certain distortions. (Dobnikar et al., 1992) have compared the invariant pattern classification (IPC) neural network architecture versus the Fourier transform (FT) method. They used for their comparison black-and-white images. They have proven the generalisation properties and fault-tolerant abilities to input patterns of the artificial neural network architectures.

An alternative approach for the solution of the invariant pattern recognition problem has been well demonstrated previously with the Hybrid Optical Neural Network (HONN) filter. HONN 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. There are two main design blocks in the HONN filter, the NNET block and the optical correlator-type block. The input images pass first through the NNET block. The extracted images from the NNET block's output are used in the composite image synthesis of the correlator-type block where we have chosen to be of the combinatorial-type.

In order to keep consistency between the different mathematical symbols of NNET architectures and optical correlators we have applied similar notation rules throughout this chapter. 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.

Section 2 describes briefly the design and implementation of the general HONN filter. Section 3 describes how the design of the general HONN filter can be altered to accommodate multiple objects recognition of the same and of different class. Section 4 describes the design and implementations of the unconstrained-HONN (U-HONN), the constrained-HONN (C-HONN), and the modified-HONN (M-HONN) filters for multiple objects recognition. Section 5 consists of the comparative analysis of U-HONN, C-HONN and M-HONN filters for multiple objects recognition. Section 6 concludes.

2. General hybrid optical neural network filter

The main motivation for the initial design and implementation of the HONN filter was to achieve the performance advantages of both NNET architectures (Kypraios et al. 2004a) and the optically implemented correlators (Bahri et Kumar, 1988). NNET architectures exhibit non-linear superposition abilities (Kypraios et al., 2002) of the training set pattern images and generalisation abilities (Beale & Jackson, 1990) over the whole set of the whole set of the

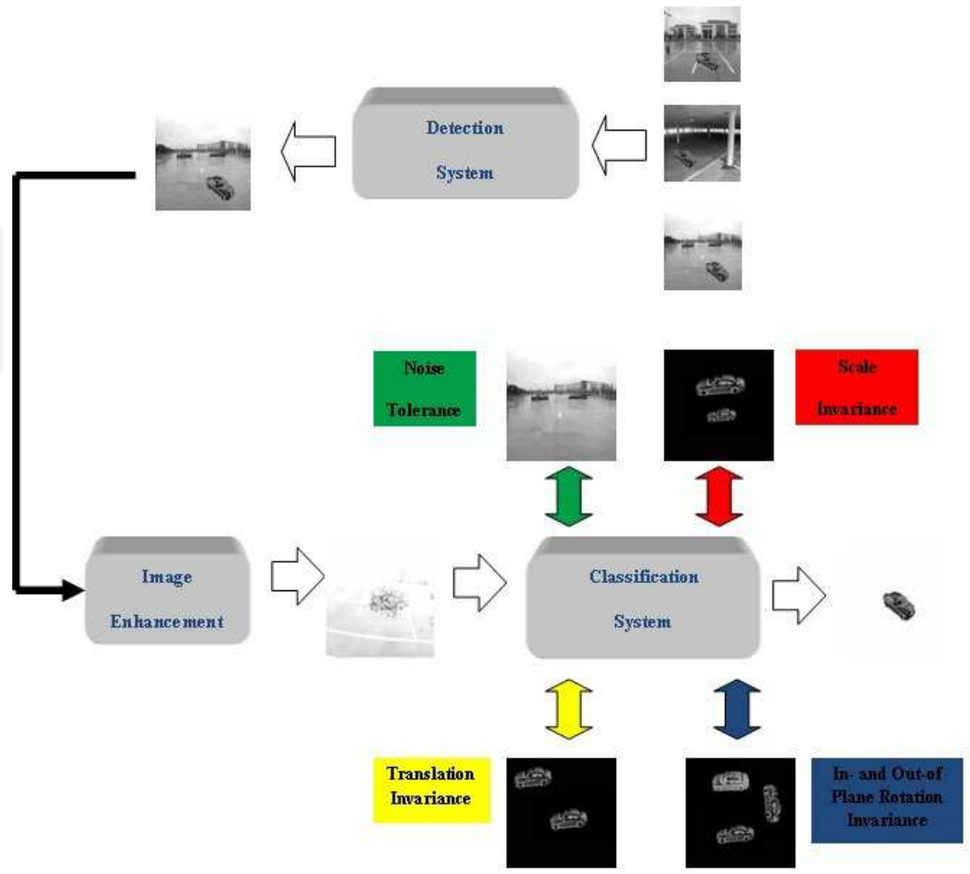


Fig. 1. A robust invariant pattern detection and classification system

input images. Optical correlators allow high speed implementation of the algorithms described. In effect, the HONN filter combines the digital design of a filter by NNET techniques with an optical correlator-type implementation of the resulting combinatorial correlator type filter (Kumar, 1992). Briefly, the original input images pass first through the NNET block and, then, the extracted images from the NNET block's output are used to form a combinatorial-type filter. Thus the output of the combinatorial-type correlator block is a composite image of the HONN filter's output. To test the HONN filter we correlate the filter with an input image.

Let $h(k,l)$ denote the composite image of the combinatorial-type correlator, such as synthetic discriminant function (SDF) filter, and $x_i(k,l)$ denote the training set images, where $i=1,2,\cdots,N$ and N is the number of the training images used in the synthesis of the combinatorial-type correlator. The basic filter's transfer function, from the weighed linear combination of x_i is given by:

$$h(k,l)=\sum_{i=1}^{N}\alpha_i\ x_i(k,l) \tag{1}$$

where the coefficients $\alpha_i(i=1,2,\cdots,N)$ are to set the constraints on the peak given by *c*. The α_i values are determined from:

$$\alpha = R^{-1}c \tag{2}$$

where α is the vector of the coefficients $\alpha_i(i=1,2,\cdots,N)$, $\boldsymbol{R}$ is the correlation matrix of t_i and $\boldsymbol{c}$ 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.

Let's assume that an image s is the input vector to an NNET's hidden neuron (node), $t_{p\kappa}$ represent the target output for a pattern p on node κ and $o_{p\kappa}$ represent the calculated output at that node. The weight from node ι to node κ is represented by $w_{\iota\kappa}$. The activation of each node κ, for pattern p, can be written as:

$$net_{p\kappa} = \sum\left(w_{\iota\kappa}o_{p\kappa} + b_{\iota}\right) \tag{3}$$

i.e. it is the weighted sum of the calculated output from the node ι to node κ. b_{ι} represents the bias vector of unit ι. We train a specifically configured NNET architecture with N training set images. The network has N neurons in 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. From eqn. (3) the net input of each of the neurons in the hidden layer is now given by:

$$net_{x_i} = \sum_{\iota=1}^{m\times n} w_{\iota}^{x_{\iota}} s_{\iota}^{x_i} \tag{4}$$

where *net* is the net input of each of the hidden neurons. $w_{\iota}^{x_i}$ are the input weights from the input layer to the hidden neurons for the training image x_i of size [mxn] in matrix form or of size [1x(mxn)] in vector form. Similarity, for the training image x_N of size [mxn] in matrix form ([1x(mxn)] in vector form) the net input, net_{x_N} is given by:

$$net_{x_N} = \sum_{i=1}^{mxn} w_{\iota}^{x_N} s_{\iota}^{x_N} \tag{5}$$

From eqns. (1), and (3) and (5) there is a direct analogy between the combinatorial-type filter synthesis procedure and the combination of all the layers' weighted input vectors.

Two possible and equivalent specially configured designs (Kypraios et al. 2004a) of NNET architectures can form the basis of the combinatorial-type filter synthesis. In both of the designs each neuron of the hidden layer is trained with only 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 effect the number of the input weights increases proportionally to the size of the training set:

$$N_{iw} = N\times[m\times n] \tag{6}$$

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 [mxn] is the size of the image of the training set. The latter design would allow parallel implementation, since all the training images could be

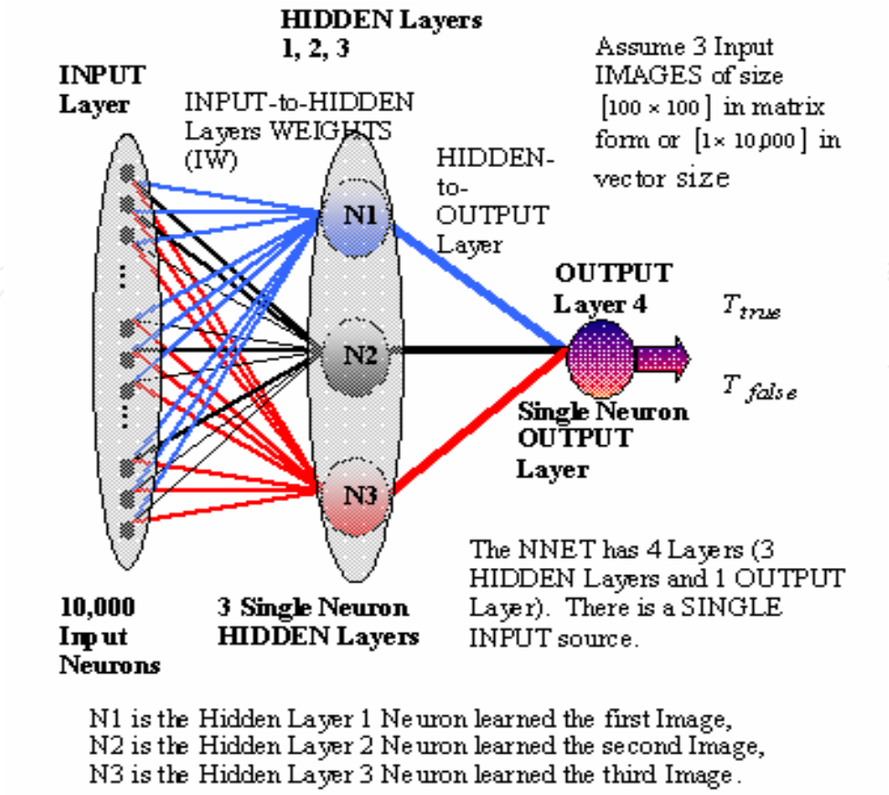


Fig. 2. Custom design of NNET block at general-HONN filter

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.

Hence, assume there are three training images of a car, size [100x100] ([1x(100x100)] in vector form), of different angle of view, to pass through the NNET. The chosen first design (see Fig. 2) uses one input source used for all the training images. Then, the input source consists of 10,000 i.e. [1x(100x100)] 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 of four-layered architecture 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 consist of 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 four-layered architecture during training which is why the distinction is made.

In the chosen configured NNET architecture design, the initial values of the input weights, the layer weights and the biases are based on the Nguyen-Widrow (Nguyen & Widrow, 1989, 1990) initialisation algorithm. The transfer function of the hidden layers is set as the Log-Sigmoidal function. 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 (Casasent et al., 1998; Talukder & Casasent, 1999) applied to the training set image. To achieve faster learning we used a modified steepest descent (Hagan et al., 1996) back propagation algorithm based on heuristic techniques. The adaptive training algorithm updates the weights and bias values according to the gradient descent momentum and an adaptive learning rate (Hagan et al., 1996):

$$\Delta w(i,i+1) = \mu \times \Delta w(i-1,i) + \lambda \times \mu \times \frac{\Delta P_f}{\Delta w(i+1,i)} \tag{7}$$

$$\Delta b(i,i+1) = \mu \times \Delta b(i-1,i) + \lambda \times \mu \times \frac{\Delta P_f}{\Delta b(i+1,i)} \tag{8}$$

$$\lambda = \left\{ \begin{array}{ll} \lambda = \lambda + \varepsilon & if \;\; \Delta P_f < 0 \\ \lambda = no \;\; change & if \;\; 0 < \Delta P_f \;\&\&\; \Delta P_f > \max(P_f) \\ \lambda = \lambda - \varepsilon & if \;\; \Delta P_f > \max(P_f) \end{array} \right\} \tag{9}$$

where now variable i is the iteration index of the network and is updated every time all the training set images pass through the NNET. Δw is the update function of the input and layer weights, Δb is the update function of the biases of the layers 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 functions like a low-pass filter by removing the small features in the error surface which allows NNET 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. The learning rate is indicated with the letter λ. 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.

3. Multiple objects recognition

All the HONN-type filters can accommodate multiple objects of the same class to be recognised within an input cluttered image due to the shift invariance properties inherited by its correlator unit. In the general HONN filter (see Fig. 3) all the input images first pass through the NNET unit. Each training image is multiplied (element wise) with the corresponding weight connections (mask). Then, the training set images after being transformed (masked) through the NNET unit by being multiplied with the mask, pass through the correlator unit where they are correlated with the masked test set images. The

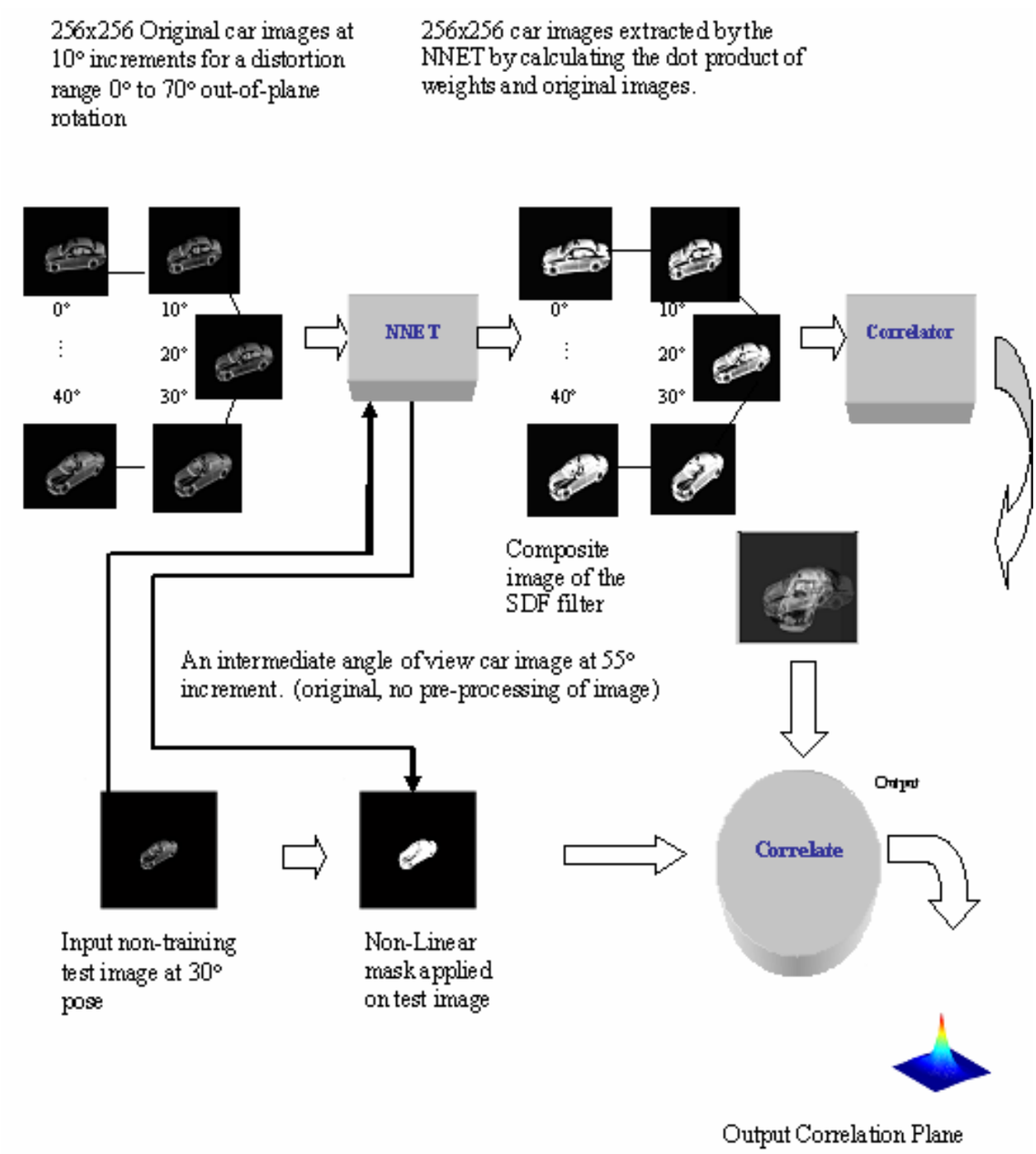


Fig. 3. Block diagram of the general-HONN filter

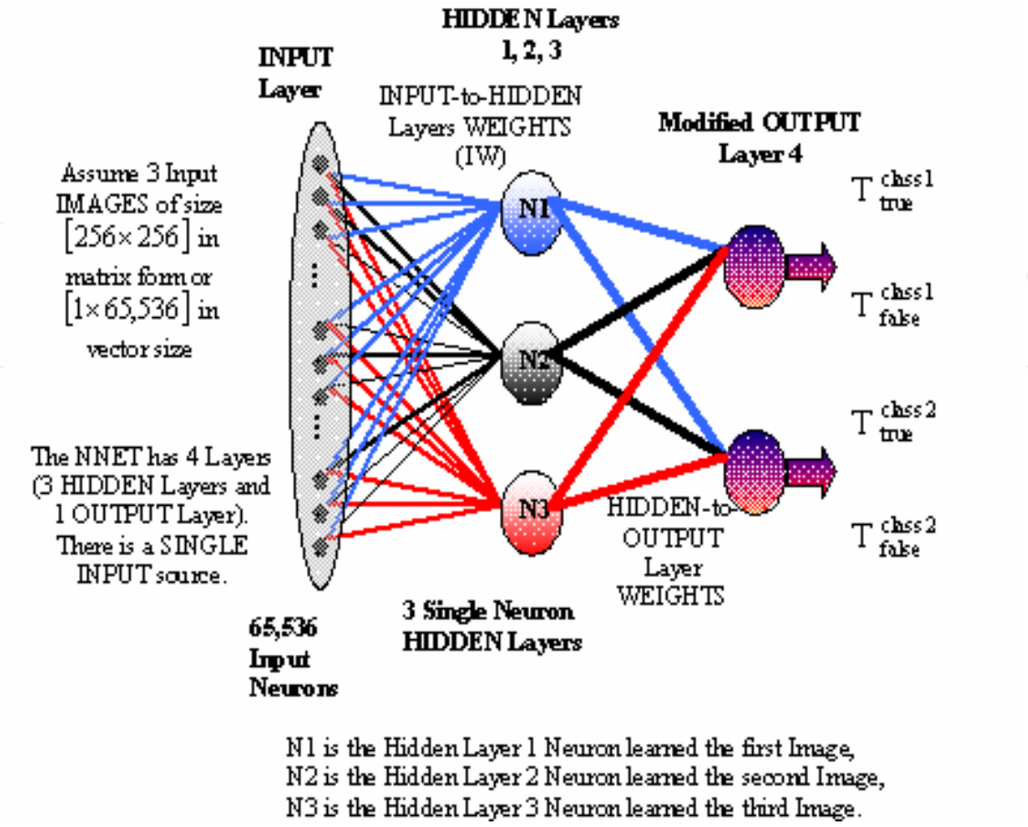


Fig. 4. Modified NNET block architecture for enabling multiple objects recognition of the same and of different classes

cross-correlation of each masked test set image with the transformed training set images (reference kernel) returns an output correlation plane peak value for each cross-correlation step. In effect, the maximum peak height values of the output correlation plane correspond to the recognised true-class objects.

3.1 Modified NNET block architecture for multiple objects of different classes recognition

Fig. 4 shows the modified NNET block architecture for accommodating multiple objects for more than one class recognition. In all the HONN-type filters presented here, i.e. Unconstrained-, Constrained-, and Modified-HONN filters, NNET is implemented as a feedforward multi-layer architecture trained with a backpropagation algorithm (Beale & Jackson, 1990). It has a single input source of input neurons equal to the size of the training image or video frame in vector form. In effect, for the training image or frame $x_{i=1...N}$ of size $[m \times n]$, there are $[m \times n]$ input neurons in the single input source. The input weights are fully connected from the input layer to the hidden layers. 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 or video frames of the training set N:

$$N = 1,2,3,\cdots,i \text{ and } N_l = N \tag{10}$$

We have set to each hidden layer to contain a single neuron. The layer weights are fully connected to the output layer. Now, the number of the layer weights N_{lw} is given by:

$$N_{lw} = N \times N_{opn} \text{ and } N_{opn} = N_{classes} \tag{11}$$

where N_{opn} is the number of the output neurons and $N_{classes}$ is the number of the different classes. In effect, we have augmented the output layer by adding more output neurons, one for each different class. On Fig. 4 we assume $N_{classes} = 2$. Thus:

$$N_{opn} = N_{classes} = 2 \text{ so, there are } N_{lw} = N \times 2 \tag{12}$$

and

$$N_{class1_{lw}} = N \text{ and } N_{class2_{lw}} = N \tag{13}$$

where $N_{class1_{lw}}$ and $N_{class2_{lw}}$ are the layer weights corresponding to object class 1 and object class 2, respectively. There are bias connections to each one of the hidden layers:

$$N_b = N \tag{14}$$

where N_b is the number of the bias connections. There are N_{target_w} target connections form the N_{opn} output neurons of the output layer:

$$N_{target_w} = N_{opn} \tag{15}$$

Thus, for $N_{classes} = 2$ there are N transformed images being created for class 1 and N transformed images being created for class 2. Then, both sets of transformed images are used for the synthesis of the filter's composite image.

4. Unconstrained-, Constrained-, and Modified-HONN filters for multiple objects recognition

Next, we describe the implementation of the unconstrained-hybrid optical neural network (U-HONN), the constrained-hybrid optical neural network (C-HONN), and the modified-hybrid optical neural network (M-HONN) filters for multiple objects recognition of the same and of different classes.

4.1 Unconstrained-HONN filter for multiple objects recognition

In general, unconstrained linear combinatorial-type filters (Mahalanobis et al., 1994; Mahalanobis & Kumar, 1997; Zhou & Chao, 1999) produce broader correlation peaks but offer better distortion tolerance. However, they are not explicitly optimised to provide good quality discrimination ability between classes. LCFs, such as the SDF filter (Sudharsanan et al., 1990) set hard constraints on the correlation peak height values of the training images. During the synthesis of the filter, peak height constraints are applied at the origin. In effect, all the training set images, when correlated with an LCF, are set to produce certain pre-specified peak values, but there is no information provided for the test image correlation peak height values of the training images. In the unconstrained correlation filter synthesis

(Mahalanobis et al., 1994; Mahalanobis & Kumar, 1997; Zhou & Chao, 1999) there are no hard constraints on the correlation peak heights. Thus, the assumption made is that by removing the hard constraints the number of possible solutions the filter can draw on increases by allowing the correlation peak height values to move freely to any value, so improving its performance.

Assume we have now $N_{classes}$ objects of different classes in the input image. Then, if do not set hard constraints on the correlation peak heights generated by the HONN filter, and add the transformed images $S_i^{class}(m,n)$, of size each $[m \times n]$, of each class, without any hard constraint weights at the origin, then we can synthesise the U-HONN filter (Kypraios et al., 2004b), which its transfer function is given by:

$$U-HONN = \sum_{i=1\ldots N_{classes}\times N}^{N_{classes}\times N} S_i^{class}(m,n) \tag{16}$$

or in the frequency domain eqn. (16) is re-written as:

$$U-HONN = \sum_{i=1\ldots N_{classes}\times N}^{N_{classes}\times N} S_i^{class}(u,v) \tag{17}$$

where $S_i^{class}(u,v)$ is the frequency domain transformed input image i of each class (with ***u, v)*** the frequency components of the image), N is the number of the input images, the image index $i=1\ldots(N_{classes}\times N)$, i.e. there are N transformed images of each of the $N_{classes}$ in the filter's synthesis, and index $N_{classes} = class1, class2, \ldots, classK$ (K any non-zero positive integer number, $K \in \Im^+$.

The non-linear U-HONN filter is inherently shift invariant and may be employed as an optical correlator-type filter. It may be used as a space domain function in a joint transform correlator architecture or be Fourier transformed and used as a Fourier domain filter in a 4-f Vander Lugt (Vander Lugt, 1964) type optical correlator.

4.2 Constrained-HONN filter for multiple objects recognition

Following a similar technique used for constraining an LCF we can constrain the correlation peaks at the centre of the correlation plane of the general HONN filter. The transfer function of the produced C-HONN filter (Kypraios et al., 2004a), now for multiple objects of the same and of different class, is given by:

$$C-HONN = \sum_{i=1\ldots N_{classes}\times N}^{N_{classes}\times N} \alpha_i \cdot S_i^{class}(m,n) \tag{18}$$

or in the frequency domain eqn. (18) is re-written as:

$$C-HONN = \sum_{i=1\ldots N_{classes}\times N}^{N_{classes}\times N} \alpha_i \cdot S_i^{class}(u,v) \tag{19}$$

where $S_i^{class}(u,v)$ is the frequency domain transformed input image i of each class (with ***u, v)*** the frequency components of the image), N is the number of the input images, the image index $i=1\ldots(N_{classes}\times N)$, i.e. there are N transformed images of each of the $N_{classes}$ in the

filter's synthesis, and index $N_{classes} = class1, class2, \ldots, classK$ (K any non-zero positive integer number, $K \in \Im^{+}$. Now, the transformed images $S_i^{class}(m,n)$, of size each $[m \times n]$, of each class, are multiplied with the hard constraint weights $\alpha_{i=1\ldots N_{classes} \times N}$ set on the correlation peak heights of the N input images at the centre of the correlation plane.
The C-HONN filter is composed of a non-linear space domain superposition of the training set images. The multiplying coefficient becomes a function of the input weights and the layer weights, rather than a simple linear multiplying constant as used in a conventional LCF synthesis. Thus, the non-linear C-HONN filter is inherently shift invariant and it may be employed in an optical correlator as would a linear superposition LCF, such as the SDF-type filters. As for the U-HONN 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 (Vander Lugt, 1964) type optical correlator.

4.3 Modified-HONN filter for multiple objects recognition

The following observations are made for the general HONN filter. Though the LCFs contain no information on non-reference objects in the training set used during their synthesis, the NNET includes information for reference and non-reference images of the true-class object. That is due to the NNET interpolating non-linearly between the reference images (Kypraios et al., 2002) included in the training set and forcing all the non-reference images to follow the activation graph. Moreover, the NNET generalises between all the reference and non-reference images. Motivated by these observations, we apply an optical mask to the filter's input (see Fig. 5). The mask is built by the weight connections of the reference images of the true-class object and is applied to all the tested images:

$$\Gamma_c = W^{x_c} \cdot L^{x_c}$$

$$= \begin{bmatrix} w_{11}^{x_c} & w_{12}^{x_c} & \cdots & w_{1n-1}^{x_c} & w_{1n}^{x_c} \\ w_{21}^{x_c} & w_{22}^{x_c} & \cdots & w_{2n-1}^{x_c} & w_{2n}^{x_c} \\ \vdots & & & & \\ w_{m1}^{x_c} & w_{m2}^{x_c} & \cdots & w_{mn-1}^{x_c} & w_{mn}^{x_c} \end{bmatrix} \cdot \begin{bmatrix} l_{11}^{x_c} & \cdots & l_{1q}^{x_c} \\ l_{21}^{x_c} & \cdots & l_{2q}^{x_c} \\ \vdots & & \\ l_{n1}^{x_c} & \cdots & l_{nq}^{x_c} \end{bmatrix}$$

$$= \begin{bmatrix} w_{11}^{x_c} \cdot l_{11}^{x_c} + w_{12}^{x_c} \cdot l_{21}^{x_c} + \cdots + w_{1n}^{x_c} \cdot l_{n1}^{x_c} & \cdots\cdots & w_{11}^{x_c} \cdot l_{1q}^{x_c} + w_{12}^{x_c} \cdot l_{2q}^{x_c} + \cdots + w_{1n}^{x_c} \cdot l_{nq}^{x_c} \\ w_{21}^{x_c} \cdot l_{11}^{x_c} + w_{22}^{x_c} \cdot l_{21}^{x_c} + \cdots + w_{2n}^{x_c} \cdot l_{n1}^{x_c} & \cdots\cdots & w_{21}^{x_c} \cdot l_{1q}^{x_c} + w_{22}^{x_c} \cdot l_{2q}^{x_c} + \cdots + w_{2n}^{x_c} \cdot l_{nq}^{x_c} \\ \vdots & & \\ w_{m1}^{x_c} \cdot l_{11}^{x_c} + w_{m2}^{x_c} \cdot l_{21}^{x_c} + \cdots + w_{mn}^{x_c} \cdot l_{n1}^{x_c} & \cdots\cdots & w_{m1}^{x_c} \cdot l_{1q}^{x_c} + w_{m2}^{x_c} \cdot l_{2q}^{x_c} + \cdots + w_{mn}^{x_c} \cdot l_{nq}^{x_c} \end{bmatrix} \quad (20)$$

where W^{x_c} and L^{x_c} are the matrices of the input and layer weights. $W_{mn}^{x_c}$ are the input 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 $x_c(m,n)$. $l_{nq}^{x_c}$ are the hidden layer weights

from the hidden neuron n to the associated output neuron q. Now, instead of multiplying each training image with the corresponding weight connections as for the C-HONN filter, we keep constant the weight connection values, setting them to be equal with a randomly chosen image included in the training set $x_c(m,n)$. The matrix Γ_c is used to build the optical mask for M-HONN.

The transfer function of the M-HONN filter (Kypraios et al., 2008, 2009) for multiple object recognition of different class objects is written as follows:

$$M-HONN = \sum_{i=1\ldots N_{classes}\times N}^{N_{classes}\times N} \alpha_i \cdot S_i^{class}(m,n)$$

$$= \alpha_1 \cdot \left(\Gamma_c^{class} \cdot X_1(m,n)\right) + \alpha_2 \cdot \left(\Gamma_c^{class} \cdot X_2(m,n)\right) + \cdots + \alpha_N \cdot \left(\Gamma_c^{class} \cdot X_N(m,n)\right) \quad (21)$$

or in the frequency domain eqn. (21) is re-written as:

$$M-HONN = \sum_{i=1\ldots N_{classes}\times N}^{N_{classes}\times N} \alpha_i \cdot S_i^{class}(u,v) \quad (22)$$

where $S_i^{class}(u,v)$ is the frequency domain transformed input image i of each class (with ***u, v)*** the frequency components of the image), N is the number of the input images, the image index $i = 1\ldots(N_{classes}\times N)$, i.e. there are N transformed images of each of the $N_{classes}$ in the filter's synthesis, and index $N_{classes} = class\,1, class\,2, \ldots, class\,K$ (K any non-zero positive integer number, $K \in \Im^+$. The transformed images $S_i^{class}(m,n)$ are calculated from the dot product of Γ_c^{class} for each class, which corresponds to an output neuron of the augmented NNET, with the corresponding training image $X_i(m,n)$.

Thus, the filter is composed of a non-linear space domain superposition of the training set images (similarly, it can be formed from video frames of the training set images). As for all the HONN-type filters, the multiplying coefficients now become a non-linear function of the input weights and the layer weights, rather than a simple linear multiplying constant as used in a constrained linear combinatorial-type filter synthesis procedure. The non-linear M-HONN filter is inherently shift invariant and it may be employed in an optical correlator as would a linear superposition LCF, such as the SDF-type filters. 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 (Vander Lugt, 1964) type optical correlator.

5. Comparative analysis of HONN-type filters

It was confirmed experimentally that by choosing different values of the target classification levels for the true-class T_{true}^{class} and false-class T_{true}^{false} objects i.e. the output layer's neuron's target output for the true-class object and the corresponding false-class object, and for each of the different object classes, respectively of the NNET (see Fig. 4), the U-HONN, C-HONN, and M-HONN filters', for multiple objects recognition, behaviour can be varied to suit different application requirements. Hence, by increasing the absolute distance of the target classification levels between the different object classes and between each object class and

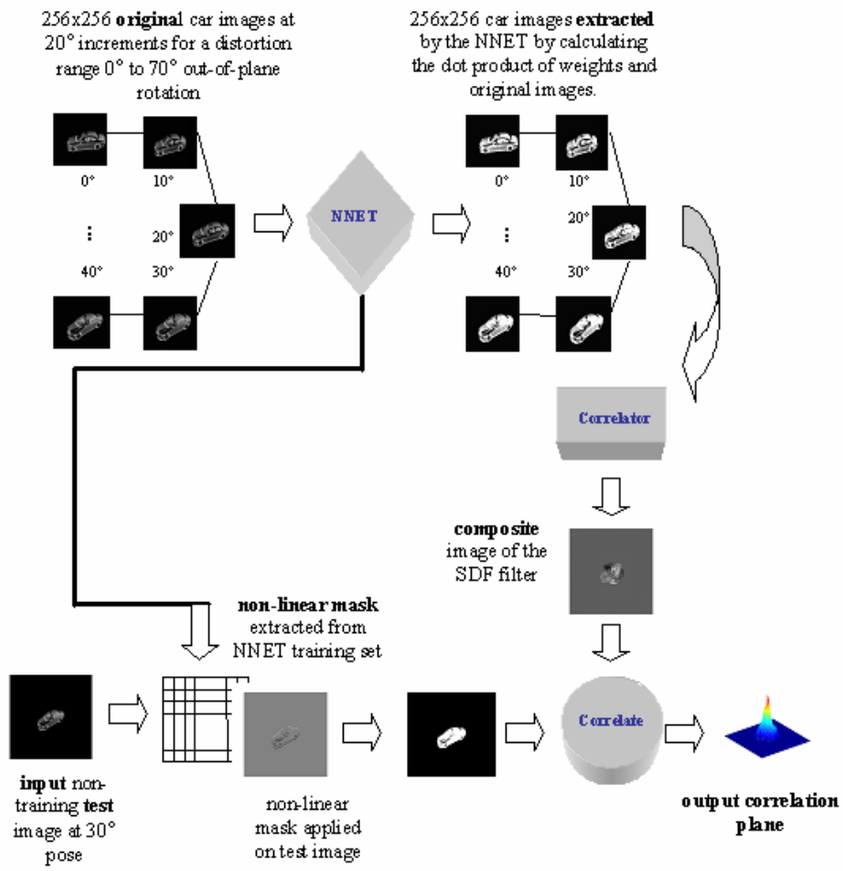


Fig. 5. Block diagram of the M-HONN filter

each corresponding false-class i.e. $\Delta T_{true}^{class} = \left| T_{false}^{class1} - T_{false}^{class2} \right|$ and $\Delta T_{true}^{class} = \left| T_{true}^{class1} - T_{true}^{class2} \right|$ the filters exhibited generally sharp peaks and good clutter suppression but are more sensitive to intra-class distortions i.e. they behave more like a high-pass biased filter, whereas by decreasing the absolute distance of the target classification levels between the different object classes and between each object class and each corresponding false-class the filters exhibited relatively good intra-class distortion invariance but producing broad correlation peaks i.e. they behave more like a minimum variance synthetic discriminant function (MVSDF) filter (Kumar, 1986). It is noted that though U-HONN, C-HONN and M-HONN filters behaviour can be varied by increasing or decreasing the absolute distance of the target classification levels between the different object classes and between each object class and each corresponding false-class, however even for the same training and test set images and for the same target classification levels the individual morphology of each filter's output correlation plane differs based on its individual transfer function characteristics. Therefore, next we will study comparatively the characteristics of its filter for multiple objects recognition within cluttered scenes. Thus, we have kept the same target classification levels for all the filters and all the conducted simulations to be able to extract useful comparison conclusions. It must be noted, again, the selected target classification levels are not optimised for each individual filter form, but rather we are aiming to set values that can allow reasonable performance for all of the filters, U-HONN, C-HONN and M-HONN.

(a) (b)

Fig. 6. (a) Test set image data, and (b) training set background car park scenes

Obviously, the optimised target classification levels achieve best performance when set individually for each of U-HONN, C-HONN and M-HONN filters (Kypraios, 2009; Kypraios et al. 2004a, 2004b, 2008, 2009).

5.1 Test data

For the conducted simulations for comparing the U-HONN, C-HONN and M-HONN filters' performance within cluttered scenes we used different image data sets. The first data set consists of an S-type Jaguar car model at 10° increments of out-of-plane rotation at an elevation angle of approximately 45°. A second image data set consists of images of a Mazda Efini RX-7 Police car model at 45° elevation angle. A third image data set consists of typical empty car park scenes (background scene images). A fourth image data set (cluttered scenes) consists of the background images of typical car parks and the images of the S-type car model and the MazdaRX-7 car model added in the background scene (see Fig. 6). All the image data sequences used in the training sets and test sets were used in grey-scale bitmap format. All the image data sequences used in the training sets and test sets were of size [256x256]. All the input video frames prior being processed by the NNET are concatenated row-by-row in to a vector form, i.e. [1x(256x256)]. Normally this size of input image data is practically impossible to be processed in real-time, since to be implemented by enough input and layer weights for 256x256 pixels input images would require in input weights:

$$N \times \left[1x\left(256 \times 256 \right) \right] \tag{23}$$

So, assume (see Fig. 4) N=10 images or video frames to be processed through any neural network architecture would require in input weights:

$$10 \times \left[1x\left(256 \times 256 \right) \right] = 655,360 \tag{24}$$

which is more than half-a-million input weight connections! Hence, it only becomes possible to overcome this problem by using the novel selective weight connection architecture.

Additionally, the heuristic training algorithm with momentum and an adaptive learning rate employed into the NNET training stage speeds up the learning phase and reduces the memory size needed to fully complete the training stage. Here, it worth mentioning that the video frame sequences for all the test series were processed by a Dual Core CPU at 2.4 GHz with 4GB RAM in few a msec.

5.2 Simulation results of C-HONN for multiple objects recognition

The training set consisted of true-class 1 object of the Jaguar S-type for a distortion range over 20° to 70° at 10° increments. At least one Mazda Efini RX-7 car image has been added in the training set to be of true-class 2 object. At least one background image of a typical car park scene has been included in the training set of the NNET block to fall inside the false-class. For the C-HONN filter we constrained in the correlator-type block the true-class 1 object images (Jaguar S-type) to unit peak-height output correlation value, the true-class 2 object images (Mazda Efini RX-7) to half-a-unit peak-height output correlation value, and all the false-class images (background scenes) to zero peak-height output correlation value. The test set consisted of a non-training in-class true-class 1 object at an intermediate car pose (non-training) and a non-training in-class true-class 2 object at an intermediate car pose (non-training) inserted in an unknown (non-training) background scene. During the true-classes objects' insertion additional Gaussian noise is added in the test set image, too. For our application purposes and for enabling us to extract useful comparison conclusions we have set in the NNET block the true-class 1 object target classification levels $T_{true}^{class1} = +40$ and the true-class 2 object target classification levels $T_{true}^{class2} = +20$. All the false-class images and all the background images with the car park scenes were set to $T_{false}^{class1} = -1$ for false-class 1 and to $T_{false}^{class2} = -1$ for false-class 2.

Several simulations with different test sets were conducted. For the purpose of this study indicatively we show one of the results recorded. Thus, Fig. 7 (a) shows the normalised, to the maximum correlation peak intensity value, isometric correlation plane for the used test set image. Fig. 7 (b) shows the position of a tracking box on top of the detected area at the output correlation plane of the true-class objects. Also, in Table 1, we have recorded the peak-to-correlation energy ratio (PCE) value, the peak-to-secondary-peak ratio (PSPR) value and the discrimination ability percentage (%) of recognising the different true-class objects for the shown output correlation plane. From the complete series of the recorded results, it is apparent that the C-HONN filter is able to detect and classify correctly both true-class objects, class 1 of Jaguar S-type and class 2 of Mazda Efini RX-7 at non-training intermediate out-of-plane rotation angles, and suppress background clutter scene. From Fig. 7 (a) and Table 1, C-HONN filter gave sharp correlation peaks with good correlation peak-height values. Note that the PSPR values should not be confused with the discrimination ability of the filter. PSPR values indicate the maximum peak-height value in comparison to the sidelobes and not to the overall output correlation plane for the test set image. Thus, the discrimination ability % value (also indicates the filter's inter-class separation ability) which C-HONN filter gave for the shown isometric correlation plane, and for separating class 1 and class 2 objects was 36.5616% (column 3 of Table 1). Though the LCFs offer no information for non-reference objects of the training set in their synthesis, the NNET block of all the HONN-type filters offers information for reference (trained) and non-reference

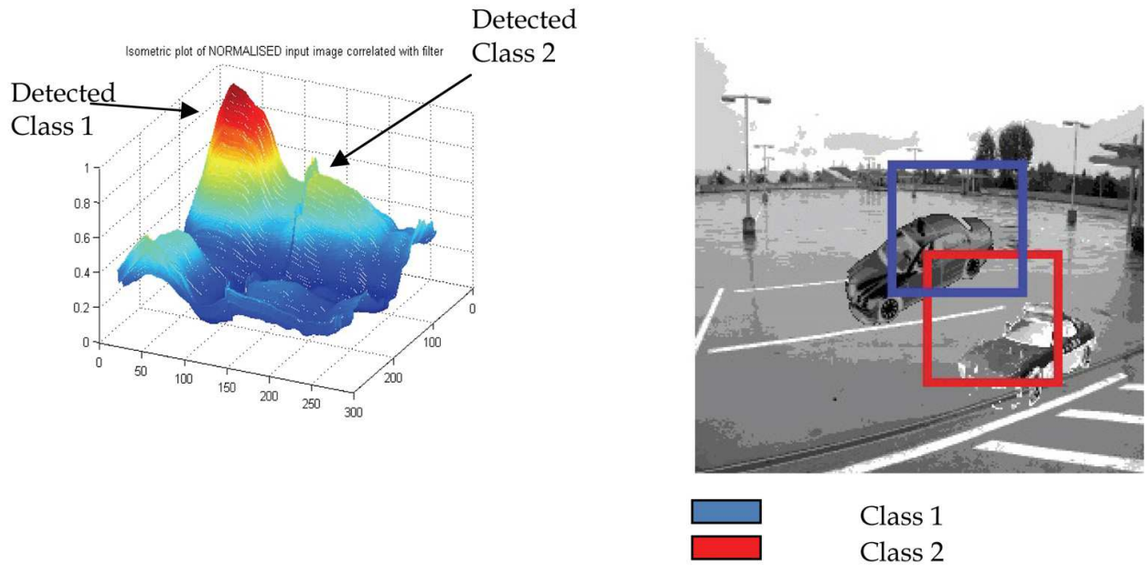


Fig. 7. (a) Normalised, to the maximum correlation peak intensity value, isometric output correlation plane of the C-HONN filter for a test set image, and (b) tracking boxes on top of the detected of the true-class objects areas at the output correlation plane; class 1of Jaguar S-type is shown with blue colour and class 2 of Mazda Efini RX-7 is shown with the red colour

(non-training) images of the true-class objects. Consequently, LCFs, such as SDF-type filters, depend solely on the information built inside the composite image formed from the reference images. However, the C-HONN filter for multiple objects recognition was able to generalise enough within the cluttered images and successfully recognise the true-class objects even at non-reference (non-training) out-of-plane rotation angles and within non-reference background car park scene.

5.3 Simulation results of U-HONN for multiple objects recognition

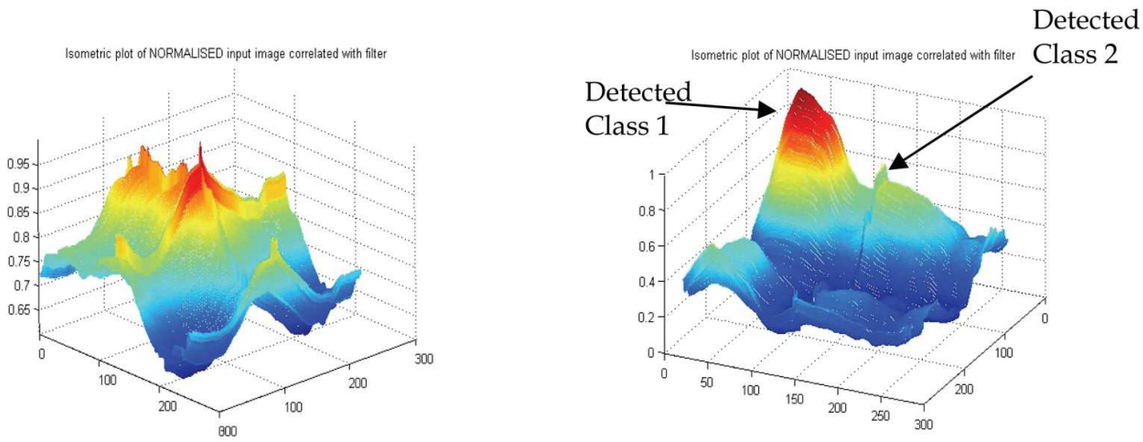
We used the same training and test data sets as for the C-HONN filter for multiple objects recognition. However, we created two slightly different training sets, one with at least one non-training background image included and a second one with no background images included. During the U-HONN filter for multiple objects recognition's synthesis of its composite image this time we set no hard constraints on the correlation peak-height values. For our application purposes and for enabling us to extract useful comparison conclusions

PCE	PSPR	Discrimination Ability %
0.0054	0.2458	36.5616

Table 1. C-HONN filter for multiple objects recognition within cluttered scenes performance assessment values

we have kept the same, as for the C-HONN filter for multiple objects recognition, target classification levels in the NNET block i.e. true-class 1 object target classification level $T_{true}^{class1} = +40$ and the true-class 2 object target classification level $T_{true}^{class2} = +20$. All the false-class images and all the background images with the car park scenes were set to $T_{false}^{class1} = -1$ for false-class 1 and to $T_{false}^{class2} = -1$ for false-class 2.

Several simulations with different test sets were conducted. For the purpose of this study indicatively we show one of the results recorded. Thus, Fig. 8 (a) shows the normalised, to the maximum correlation peak intensity value, isometric correlation plane for the used test set image with at least one background non-reference car park scene included in the training set, Fig. 8 (b) shows the normalised, to the maximum correlation peak intensity value, isometric correlation plane for the used test set image with no background car park scene included in the training set, and Fig. 9 shows the position of a tracking box on top of the detected area at the output correlation plane of the true-class objects for the isometric correlation plane shown in Fig. 8 (a). In Table 2 we have recorded the PCE values, the PSPR values and the discrimination ability % of recognising the different true-class objects for both shown output correlation planes in Fig. 8 (a) and Fig. 8 (b). From the complete series of the recorded results, it is apparent that the U-HONN filter is able to detect and classify correctly both true-class objects, class 1 of Jaguar S-type and class 2 of Mazda Efini RX-7 at non-training intermediate out-of-plane rotation angles, and suppress background clutter scene. But, when there was at least one non-reference background scene included in the U-HONN filter's synthesis then it increased the detected false-class areas at the output correlation plane (see Fig. 8 (b)). Thus, as it was expected from U-HONN filter for multiple objects recognition's design and transfer function (see eqns. (16) and (17)), the resulted solutions from correlating the test set image with the U-HONN filter's transfer function are increasing in comparison with the C-HONN filter since there are no hard constraints imposed on the correlation peak-heights for U-HONN filter. However, by including a false-class non-reference background image in the filter's synthesis, it produces more unwanted false-class peaks in comparison to having not included any background images. From Fig. 7 and Fig. 8 (a) (for background images included in the training set), and from Table 1 and Table 2 U-HONN filter for multiple objects recognition produces higher PSPR with smaller sidelobes values i.e. sharper correlation peaks, but C-HONN filter for multiple objects recognition produces higher correlation peak-height values with broader sidelobes (smaller PSPR values). The discrimination ability % value U-HONN filter gave for Fig. 8 (a) shown isometric correlation plane, and for separating class 1 and class 2 objects was 12.0242% and for Fig. 8 (b) was approximately 2% (column 3 of Table 2), which for both isometric plots it is less than the discrimination ability % value that C-HONN filter gave. In effect, U-HONN filter for multiple objects recognition it maximises the correlation peak-heights (including the false-class ones in the case of Fig. 8 (a) plot) at the output correlation plane in expense of broadening the sidelobes and, thus, decreasing its discrimination ability %, for the test set images to recognise the true-class objects in the cluttered scenes. Consequently, for U-HONN, broader sidelobes means better intra-class ability and better distortion range, i.e. it is able to maintain high correlation peak-heights for recognising intermediate non-reference out-of-plane rotation angles of the true-class objects with less correlation peak-height value decrease (drop) than C-HONN filter. Again, as for the C-HONN filter for multiple objects recognition, the U-HONN filter for multiple objects recognition was able to generalise



(a) (b)

Fig. 8. (a) Normalised, to the maximum correlation peak intensity value, isometric output correlation plane of the U-HONN filter for a test set image with at least one background non-reference car park scene included in the training set, and (b) normalised, to the maximum correlation peak intensity value, isometric correlation plane for the used test set image with no background car park scene included in the training set

enough within the cluttered images and successfully recognise the true-class objects even at non-reference (non-training) out-of-plane rotation angles and within non-reference background car park scene.

5.4 Simulation results of M-HONN for multiple objects recognition

We used the same training and test data sets as for the C-HONN filter for multiple objects recognition. As for the C-HONN filter for multiple objects recognition, we constrained in the correlator-type block of the M-HONN filter for multiple objects recognition the true-class 1 object images (Jaguar S-type) to unit peak-height output correlation value, the true-class 2 object images (Mazda Efini RX-7) to half-a-unit peak-height output correlation value, and all the false-class images (background scenes) to zero peak-height output correlation value. For our application purposes and for enabling us to extract useful comparison conclusions we have kept the same, as for the C-HONN filter for multiple objects recognition, target classification levels in the NNET block i.e. true-class 1 object target classification level $T_{true}^{class1} = +40$ and the true-class 2 object target classification level $T_{true}^{class2} = +20$. All the false-class images and all the background images with the car park scenes were set to $T_{false}^{class1} = -1$ for false-class 1 and to $T_{false}^{class2} = -1$ for false-class 2.

Several simulations with different test sets were conducted. For the purpose of this study indicatively we show one of the results recorded. Fig. 10 (a) shows the normalised, to the maximum correlation peak intensity value, isometric correlation plane for the used test set image and Fig. 10 (b) shows the position of a tracking box on top of the detected area at the output correlation plane of the true-class objects. In Table 3 the PSPR value is recorded for the test set image. From the complete series of the recorded results, it is apparent that the M-HONN filter is able to detect and classify correctly both true-class objects, class 1 of Jaguar

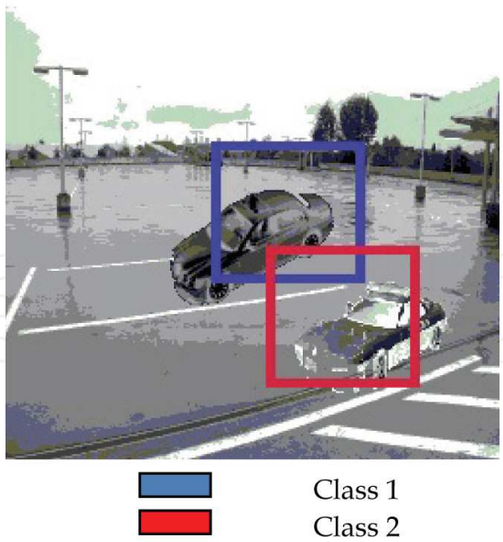


Fig. 9. Tracking boxes on top of the detected areas of the true-class objects at the output correlation plane of U-HONN filter for multiple objects recognition; class 1of Jaguar S-type is shown with blue colour and class 2 of Mazda Efini RX-7 is shown with the red colour

U-HONN for multiple objects	PCE	PSPR	Discrimination Ability %
Train Set with Backgorund Images included	0.0020	0.9522	12.0242
Train Set with NO Backgorund Images included	0.0044	0.0840	approx. 2.0

Table 2. U-HONN filter for multiple objects recognition within cluttered scenes performance assessment values

S-type and class 2 of Mazda Efini RX-7 at non-training intermediate out-of-plane rotation angles, and suppress background clutter scene. From the isometric plots shown in Fig. 7, Fig. 8 (b), Fig. 9 and Fig. 10, and from Table 1, Table 2 and Table 3 it is found that the M-HONN filter for multiple objects recognition produces higher PSPR values and i.e. sharper peaks than the C-HONN and U-HONN filters for multiple objects recognition. However, from the full series of the conducted tests it is recorded that the M-HONN filter for multiple objects recognition produces a higher drop for the non-training intermediate car poses within the background clutter than the C-HONN and U-HONN filters for multiple objects recognition (Kypraios, 2009; Kypraios et al. 2008). In effect, the M-HONN filter for multiple objects recognition confirms its design expectation of producing optimum performance (sharper peak-heights) within cluttered scenes than U-HONN and C-HONN filters for multiple objects recognition in expense of a drop in its intra-class (non-training images of intermediate out-of-plane rotation angles of true-class objects) distortion tolerance.

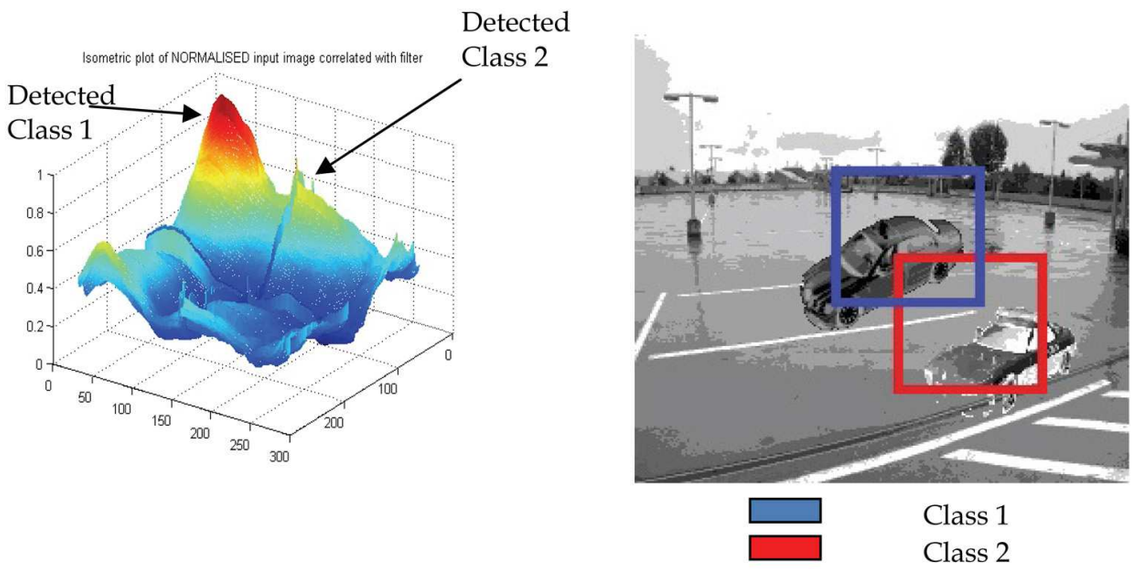


Fig. 10. (a) Normalised, to the maximum correlation peak intensity value, isometric output correlation plane of the M-HONN filter for a test set image, and (b) tracking boxes on op of the detected of the true-class objects areas at the output correlation plane; class 1of Jaguar S-type is shown with blue colour and class 2 of Mazda Efini RX-7 is shown with the red colour

PCE	PSPR	Discrimination Ability %
0.0050	0.3026	20.4320

Table 3. M-HONN filter for multiple objects recognition within cluttered scenes performance assessment values

M-HONN filter for multiple objects recognition gave for the shown isometric correlation plane in Fig. 10, and for separating class 1 and class 2, the discrimination ability % value of 20.4320% (column 3 of Table 3). Thus, C-HONN filter for multiple objects recognition exhibits higher discrimination ability than the M-HONN filter for multiple objects recognition. Still, the M-HONN filter is able to separate adequately the different classes of objects. Again, as for the U-HONN and C-HONN filters for multiple objects recognition, the M-HONN filter for multiple objects recognition was able to generalise enough within the cluttered images and successfully recognise the true-class objects even at non-reference (non-training) out-of-plane rotation angles and within non-reference (unknown) background car park scene.

6. Conclusion

We have compared with each other the performance of the U-HONN, C-HONN and M-HONN filters for multiple objects recognition. We have described how the U-HONN, C-

HONN and M-HONN filters can accommodate the recognition of multiple objects of the same or of different classes. Due to the shift invariance properties inherited by its correlator unit the filter can accommodate multiple objects of the same class to be detected within an input cluttered image. Also, the architecture of the NNET block of the general-HONN filter allows the recognition of multiple objects of different classes within the input cluttered image by augmenting the output layer of the unit. U-HONN, C-HONN and M-HONN filters for multiple objects recognition may be used as a space domain function in a joint transform correlator architecture or be Fourier transformed and used as a Fourier domain filter in a 4-f Vander Lugt-type optical correlator. It was confirmed experimentally that by increasing or decreasing the absolute distance of the target classification levels between the different object classes and between each object class and each corresponding false-class i.e. $\Delta T_{true}^{class} = \left| T_{false}^{class1} - T_{false}^{class2} \right|$ and $\Delta T_{true}^{class} = \left| T_{true}^{class1} - T_{true}^{class2} \right|$, the U-HONN, C-HONN, and M-HONN filters, for multiple objects recognition, behaviour can be varied to behave from more like a high-pass biased filter to more like a MVSDF filter for serving the different application requirements. However, even for the same training and test set images and for the same target classification levels the individual morphology of each filter's output correlation plane differs based on its individual transfer function characteristics. Therefore, for all the conducted tests the target classification levels for all the filters have been kept the same in order to be able to extract useful comparison conclusions. It must be noted the target classification levels are chosen to values which can allow good performance for all the filters and for keeping the same values. Obviously, best performance can be achieved by setting the values individually for each of U-HONN, C-HONN and M-HONN filters but not to the same target classification levels.

U-HONN, C-HONN and M-HONN filters for multiple objects recognition exhibit simultaneously shift and out-of-plane rotation invariances with a single pass over the data, i.e. there is not needed more than one filter to be trained for shift invariance and separately another one for out-of-plane rotation invariance. Additionally, they exhibit good tolerance-to-clutter performance without disturbing the other simultaneously exhibit properties of out-of-plane rotation and shift invariances. In general, the HONN-type filters are shown experimentally to be performing better than the LCFs. U-HONN, C-HONN and M-HONN filters are proven to recognize correctly the multiple true-class objects of the same or of different classes within non-reference (unknown i.e. not previously trained) background scenes. U-HONN filter for multiple objects recognition exhibits better distortion range, i.e. it maintains good correlation peak height for recognising intermediate non-reference out-of-plane rotation angles of the true-class objects, and higher peak-heights but in expense of broader sidelobes in recognising the true-class objects within the cluttered scene than the C-HONN and M-HONN filters. M-HONN filter design is optimised for producing best performance in recognising objects within cluttered scenes. Hence, it was found that it gave sharper peaks than the U-HONN and C-HONN filters for recognising the true-class objects of the different classes within the unknown car park scene. However, C-HONN filter for multiple objects recognition produces more controlled peak-heights and better discrimination ability between the true-class objects of different classes within the cluttered scenes than the U-HONN and M-HONN filters for multiple objects recognition. U-HONN, C-HONN and M-HONN filters for multiple objects recognition can be employed, amongst the many application areas, in image content-based Internet search engines. The simultaneous properties of U-HONN, C-HONN and M-HONN filters of shift and out-of-

plane rotation invariances, can reduce the number of stored images for each object class and, consequently reduce the time needed for an Internet image-to-image search engine (content-based search engine) to search the complete data set of matched images. Moreover, the accommodation of multiple objects recognition of the same and of different classes with the same single filter and with a single pass over the training and test data reduces the training times instead of training several filters for the different object classes.

7. References

Wood, J. (1996). Invariant Pattern Recognition: A Review. *Pattern Recognition,* Vol.29, No.1, pp. 1-17

Forsyth, D. A. & Ponce, J. (2003). *Computer Vision –A Modern Approach*, Prentice Hall International, Inc. (Pearson Education, Inc.),ISBN-10 0130851981, ISBN-13 978-0130851987

Sheng, Y., & Arsenault, H. H. (1986). Experiments on Pattern Recognition Using Invariant Fourier-Mellin Descriptors, *Journal of Optical Society of America A (Optics and Image Science),* Vol.3, pp. 771-776

Sheng, Y., & Lejeune, C. (1991). Invariant Pattern Recognition Using Fourier-Mellin Transforms and Neural Networks, *Journal of Optics,* Vol.22, pp. 223-228

Grace, A., & Spann, M. (1991). A Comparison Between Fourier-Mellin Descriptors and Moment Based Features For Invariant Object Recognition Using Neural Networks, *Pattern Recognition Letters,* Vol.12, pp. 635-643

Casasent, D. & Psaltis D. (1976). Position, Rotation and Scale Invariant Optical Correlation, *Applied Optics,* Vol.15, No.7, pp. 1795-1799

Mersereau, K. & Morris, G. (1986). Scale, Rotation, and Shift Invariant Image Recognition, *Applied Optics,* Vol.25, No.14, pp. 2338-2342

Hsu, Y. N., Arsenault, H. H. & April, G. (1982). Optical Pattern Recognition Using Circular Harmonic Expansion, *Applied Optics,* Vol. 21, pp. 4012

Hsu, Y. N. & Arsenault H. H. (1984). Pattern Discrimination By Multiple Circular Harmonic Components, *Applied Optics,* Vol.23, pp. 841

Jensen, A. S., Lindvold, L. & Rasmussen, E. (1987). Transformation of Image Positions, Rotations and Sizes Into Shift Parameters, *Applied Optics,* Vol.26, No.9, pp. 1775-1781

Bryngdahl, O. (1974). Geometrical Transformations In Optics, *Journal of Optical Society of America,* Vol.64, pp. 1092

Cederquist, J. & Tai, A. M. (1984). Computer-Generated Holograms For Geometric Transformations, *Applied Optics,* Vol.23, pp. 3099

Stamos, E. (2001). *Algorithms for Designing Filters for Optical Pattern Recognition,* D.Phil. Thesis, Department of Electronic and Electrical Engineering, University College London

Lynn, P. A. & Fuerst, W. (1998). *Introductory Digital Signal Processing- with Computer Applications,* John Wiley & Sons Ltd.

Proakis, J. G. & Manolakis, D. G. (1988). *Introduction to Digital Signal Processing,* Prentice Hall International Paperback Editions

Delopoulos, A., Tirakis, A. & Kollias, S. (1994). Invariant Image Classification Using Triple-Correlation-Based Neural Networks, *IEEE Transactions on Neural Networks*, Vol.5, No.3, pp. 392-408

Perantonis, S. & Lisboa, P. (1992). Translation, Rotation and Scale Invariant Pattern Recognition by High-Order Neural Networks and Moment Classifiers, *IEEE Transactions on Neural Networks,* Vol.3, No.2, pp. 241-251

Khotanzad, A. & Hong, H. (1990). Invariant Image Recognition by Zernike Moments, *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 12, pp. 489-497

Shvedov, A., Schmidt, A. & Yakubovich, V. (1979). Invariant Systems of Features in Pattern Recognition, *Automation Remote Control,* Vol. 40, pp. 131-142

LeCun, Y. (1989). Generalisation and Network Design Strategies, *Connectionism in Perspective,* Pfeirer, R., Schreter, Z., Fogelman-Soulié, F. & Steels, L., Elsevier Science, Amsterdam

LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., Jackel, L. D. (1990). Handwritten Digit Recognition with a Backpropagation Network, *Advances in Neural Information Processing Systems,* Touretzky, D., Morgan Kaufmann, Vol.2, pp. 396-404

Spirkovska, L. & Reid, M. (1992). Robust Position, Scale and Rotation Invariant Object Recognition Using Higher-Order Neural Networks, *Pattern Recognition,* Vol.25, pp. 975-985

Giles, C. L. & Maxwell, T. (1987). Learning, Invariance and Generalisation in Higher-Order Neural Networks, *Applied Optics,* Vol.26, pp. 4972-4978

Kanaoka, T., Chellapa, R., Yoshitaka, M. & Tomita, S. (1992). A Higher-Order Neural Network for Distortion Invariant Pattern Recognition, *Pattern Recognition Letters,* Vol.13, pp. 837-841

Waibel, A., Hanazawa, T., Hinton, G , Shikano, K. & Lang, K. (1989). Phoneme Recognition Using Time-Delay Neural Networks, *IEEE Transactions on Acoustics, Speech Signal Processing,* Vol.37, No.3, pp. 328-339

Bottou, L., Fogelman- Soulié, F., Blanchet, P. & Lienard, J. S. (1990). Speaker Independent Isolated Digit Recognition : Multilayer Perceptrons vs Dynamic Time Warping, *Neural Networks,* Vol.3, pp. 453-465

Simard, P. & LeCun, Y. (1992). Reverse TDNN : An Architecture for Trajectory Generation, *Advances in Neural Information Processing Systems,* Moody, J., Hanson, S. & Lipmann, R., Morgan Kauffmann, Vol.4, pp. 579-588

Yuceer, C. & Oflazer, K. (1993). A Rotation, Scaling and Translation Invariant Pattern Classification System, *Pattern Recognition,* Vol.26, No.5, pp. 687-710

Kypraios, I., Young, R. C. D., Birch, P. M. & Chatwin, C. R. (2004a). Object Recognition Within Cluttered Scenes Employing a Hybrid Optical Neural Network Filter, *Optical Engineering Special Issue on Trends in Pattern Recognition,* Vol.43, pp. 1839-1850

Bahri, Z. & Kumar, B. V. K. (1988). Generalized Synthetic Discriminant Functions, *Journal of Optical Society of America,* Vol.5, No.4, pp. 562-571

Kypraios, I., Young, R. C. D., Chatwin C. R. (2002). An Investigation of the Non-Linear Properties of Correlation Filter Synthesis and Neural Network Design, *Asian Journal of Physics,* Vol.11, No.3, pp. 313-344

Beale, R. & Jackson, T. (1990). *Neural Computing : An Introduction*, Institute of Physics Publishing, Hilger, ISBN 0852742622, 9780852742624, Bristol, Philadelphia

Kumar, B. V. K. (1992). Tutorial Survey of Composite Filter Designs for Optical Correlators, *Applied Optics,* Vol.31, No.23, 4773-4801

Nguyen, D. & Widrow, B. (1989). The Truck Backer-Upper: An Example of Self-Learning in Neural Networks, *Proceedings of the IEEE International Joint Conference on Neural Networks,* Vol.2, pp. 357-363

Nguyen, D. & Widrow, B. (1990). Improving the Learning Speed of 2-Layer Neural Networks by Choosing Initial Values of the Adaptive Weights, *Proceedings of the IEEE International Joint Conference on Neural Networks,* Vol.3, pp. 21-26

Casasent, D., Neiberg, L. M. & Sipe, M. A. (1998). Feature Space Trajectory Distorted Object Representation for Classification and Pose Estimation, *Optical Engineering,* Vol.37, No.3, pp. 914-923

Talukder, A. & Casasent, D. (1999). Non-Linear Features for Product Inspection, *Optical Pattern Recognition X, Proceedings of SPIE,* Vol.3715, pp. 32-43

Hagan, M. T., Demuth, H. B. & Beale, M. H. (1996). *Neural Network Design,* PWS Publishing, ISBN 0-9717321-0-8, Boston, MA

Sudharsanan, S. I., Mahalanobis, A. & Sundareshan, M. K. (1990). A Unified Framework for the Synthesis of Synthetic Discriminant Functions with Reduced Noise Variance and Sharp Correlation Structure, *Optical Engineering,* Vol.29, pp. 1021-1028

Mahalanobis, A., Kumar, B. V. K., Song, S., Sims, S. R. F. & Epperson, J. F. (1994). Unconstrained Correlation Filters, *Applied Optics,* Vol.33, No.17, pp. 3751-3759

Zhou, H. & Chao, T. H. (1999). MACH Filter Synthesising for Detecting Targets in Cluttered Environment for Gray-Scale Optical Correlator, *Optical Pattern Recognition X, Proceedings of SPIE,* Vol.3715, pp. 394-398

Mahalanobis, A. & Kumar, B. V. K. (1997). Optimality of the Maximum Average Correlation Height Filter for Detection of Targets in Noise, *Optical Engineering,* Vol.36, No.10, pp. 2642-2648

Kypraios, I., Young, R. C. D. & Chatwin, C. R. (2004b). Performance assessment of Unconstrained Hybrid Optical Neural Network filter for Object Recognition Tasks in Clutter, *Optical Pattern Recognition XV, Proceedings of SPIE,* Vol.5437, pp. 51-62

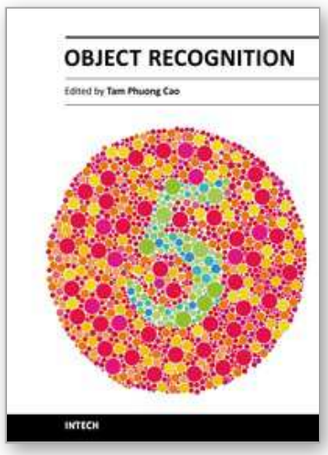
Vander Lugt, A. (1964). Signal Detection By Complex Spatial Filtering, *IEEE Transactions on Information Theory,* Vol.10, pp. 139-145

Kypraios, I., Lei, P. W., Birch, P. M., Young, R. C. D. & Chatwin C. R. (2008). Performance Assessment of the Modified-Hybrid Optical Neural Network Filter, *Applied Optics,* Vol.47, No.18, pp. 3378-3389

Kypraios, I., Young, R. C. D. & Chatwin, C. R. (2009). Modified-Hybrid Optical Neural Network Filter for Multiple Objects Recognition within Cluttered Scenes, *Optics and Photonics for Information Processing III, Proceedings SPIE,* Vol.7442, pp. 74420P-74420P-12

Kumar, B. V. K. (1986). Minimum Variance Synthetic Discriminant Functions, *Journal Optics Society America A,* Vol.3, pp. 1579-1584

Kypraios, I. (2009). A Comparative Analysis of the Hybrid Optical Neural Network-type Filters Performance within Cluttered Scenes, *51st International Symposium ELMAR, IEEE Region 8/IEEE Croatia/EURASIP,* Vol.1, pp. 71-77



Object Recognition

Edited by Dr. Tam Phuong Cao

ISBN 978-953-307-222-7
Hard cover, 350 pages
Publisher InTech
Published online 01, April, 2011
Published in print edition April, 2011

Vision-based object recognition tasks are very familiar in our everyday activities, such as driving our car in the correct lane. We do these tasks effortlessly in real-time. In the last decades, with the advancement of computer technology, researchers and application developers are trying to mimic the humanâ€™s capability of visually recognising. Such capability will allow machine to free human from boring or dangerous jobs.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Ioannis Kypraios (2011). Hybrid Optical Neural Network-Type Filters for Multiple Objects Recognition within Cluttered Scenes, Object Recognition, Dr. Tam Phuong Cao (Ed.), ISBN: 978-953-307-222-7, InTech, Available from: http://www.intechopen.com/books/object-recognition/hybrid-optical-neural-network-type-filters-for-multiple-objects-recognition-within-cluttered-scenes

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen