Similarly the ACF of the code c_2 is given by

$$R_{c22}(k) = \sum_{t=0}^{(M-1-|k|)/2} -a_t \cdot q_{[(|k|-1/2)+t]} + \sum_{t=1}^{(M-1-|k|)/2} -q_{t-1} \cdot a_{[|k|-1/2)+t]}$$

$$-(M-1) \leq k \leq (M-1) \qquad \text{with } k \text{ odd} \quad (12)$$

and

$$R_{c22}(k) = \sum_{t=0}^{(M-2-|k|)/2} a_t \cdot a_{[(|k|-2/2)+t+1]} + \sum_{t=0}^{(M-2-|k|)/2} q_t \cdot q_{[(|k|-2/2)+t+1]}$$

$$-(M-1) \leq k \leq (M-1) \qquad \text{with } k \text{ even} \quad (13)$$

Summing the ACFs results in zero for odd k. For even k

$$R_{c11}(k) + R_{c22}(k) = 4 \sum_{t=0}^{(M-2-|k|)/2} a_t \cdot a_{[(|k|-2/2)+t+1]} = 4 \sum_{t=0}^{N-1-(|k|/2)} a_t \cdot a_{[(|k|/2)+t]} \quad (14)$$

which is four times the ACF of the code $\boldsymbol{a}$ at even values of k.

Further transformations to a set of 2^n codes: If the code pair c_1 and c_2 are now transformed separately, a set of four codes is produced. Further transformations result in a set of 2^n codes which obey the following:

$$\sum_{j=1}^{j=2^n} \sum_{t=1}^{t=2^n} R_{B_iB_j}(k) = 0 \qquad \forall k,\ i \neq j \quad (15)$$

$$\frac{1}{2^{2n}} \max \left| \sum_{t=1}^{t=2^n} R_{B_jB_i}(k) \right| = \max |R_{aa}(k)| \qquad \forall k \neq 0 \quad (16)$$

B_i is the ith code in the set and R_{ab} is the CCF of $\boldsymbol{a}$ and $\boldsymbol{b}$.

Practical application: The beamwidth of linear arrays may be reduced considerably by transmitting two of more signals from a suitably disposed set of transmitting antennas and using appropriate matched filter processing.[2] At the receiving array the transmitted signals must be separated completely suggesting the need for orthogonal codes. If additive processing is employed, as in the dual transmission radar shown in Fig. 1, the signals need not be orthogonal if the CCFs of the waveforms sum to zero and the ACFs sum to provide a useful PSR. In such a scheme the waveform design task may be reduced to the search for a single phase code with the required range-velocity ambiguity diagram. By using the transformations described previously, a set of codes for a 2^n transmission radar may be obtained, offering a combined response that is a time-expanded version of the original code's response. Although the codes suffer in range-bearing, the problem may be eliminated by an appropriate choice of signal bandwidth, carrier frequency and array length.[3]

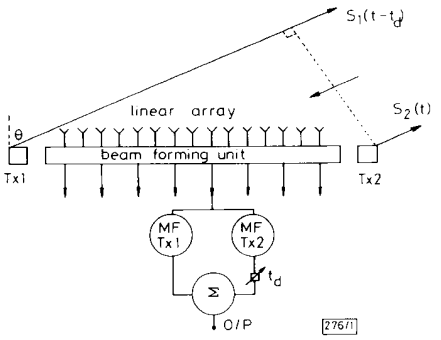

Fig. 1 *Dual transmission radar*

Acknowledgment: The helpful criticism of this paper provided by Dr. D. C. Cooper is gratefully acknowledged.

E. R. McKINNON *30th April 1990*

School of Electronic and Electrical Engineering
University of Birmingham
Birmingham B15 2TT, United Kingdom

References

1 TSENG, C. C., and LUI, C. L.: 'Complementary sets of sequences', *IEEE Trans.*, 1972, **IT-18**, (5)
2 COOPER, D. C.: 'The use of multiple transmitting antennas to provide beamwidth reduction for HF pulse radars with sector coverage'. IEE Conf. Publ., 195, Part 1, April 1981
3 MCKINNON, E. R.: 'Design of orthogonal phase shift coded pulses'. Birmingham University Final Year Project Report, 76, April 1989

DAUGMAN's GABOR TRANSFORM AS A SIMPLE GENERATIVE BACK PROPAGATION NETWORK

Indexing term: Networks and network theory

Much work has been performed on learning mechanisms for neural networks. A particular area of interest has been the use of neural networks for image processing problems. Two important pieces of work in this area are unified. An architecture and learning scheme for neural networks called generative back propagation has been previously developed and a system for image compression and filtering based on 2-D Gabor transformations which used a neural network type architecture described. Daugman's procedure is exactly replicated. A procedure which used a four layer neural network as a two-layer generative back propagation network with half of the units. The GBP update rule is shown to perform the same change as Daugman's rule, but more efficiently.

Introduction: Much work has been performed on learning mechanisms for neural networks. A particular area of interest has been the use of neural networks for image processing problems. This letter unifies two important pieces of work in this area.

Hinton[2] described an architecture and learning scheme for neural networks which he called generative back propagation. Daugman[1] described a system for image compression and filtering based on 2-D Gabor transformations which used a neural network type architecture.

The aim of this letter is to show that Daugman's procedure which used a four layer neural network can be implemented as a two-layer generative back propagation network with half the units and less computation.

Generative back propagation: Generative back propagation (GBP) is a method of approximating a set of descriptive parameters from an image. It is typically used to parameterise a pixel based scene in terms of a few real parameters.

The idea behind GBP is to use a feedforward neural network which generates scenes from a small set of parameters. The parameters can be taken as the strengths of the connections from an input layer with activations all equal to one to the units of the first layer. This network can be used to approximate the set of parameters corresponding to a novel scene using neural network learning algorithms.[3-6] The algorithms are used to adapt only the input parameter weights, leaving all other weights fixed. This operation is shown in the architecture of Fig. 1.

When the network has stabilised to a global minimum, the weights will represent the best fit parameterisation of the image.

Daugman Gabor transform: Daugman fixed a set of Gabor functions and tried to express a pixel based image as an

expansion in this basis of Gabor functions. The problem arises that, unlike Fourier transforms and pixel representations, Gabor functions do not form an orthogonal basis. This means

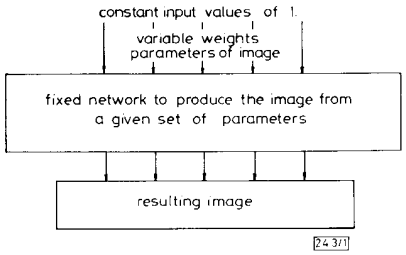

Fig. 1 *Typical network configuration for GBP*

that obtaining the expansion of the image with respect to the Gabor functions is a difficult problem.

The approach that Daugman uses is to employ a network of neuron-like units with a particular learning rule. The architecture is designed so that when the weights of the connections stabilise they are the best least-mean-squares aprpoximations to the Gabor parameters. This is achieved by finding the derivative of the error with respect to each of the Gabor parameters and using a gradient descent method to iteratively approximate the solution.

Suppose the Gabor functions are $G_i(x, y)$, $i = 1, \ldots, n$ and the picture has value $I(x, y)$ at pixel (x, y). The expansion in terms of the Gabor function is

$$I(x, y) = \sum_{i=1}^{n} a_i G_i(x, y)$$

where the a_i, $i = 1, \ldots, n$ are the sought parameter set.

Suppose that some value for the a_is is guessed. The derivative of the least-mean-square error E with respect to a_i is

$$D_{a_i} = \frac{\partial E}{\partial a_i} = 2 \sum_{x,y} [I(x, y) G_i(x, y)] - 2 \sum_{x,y} \left[\left(\sum_{k=1}^{n} a_k G_k(x, y) \right) G_i(x, y) \right]$$

The update rule for the parameters a_i is then given by

$$a_i' = a_i - \tfrac{1}{2} D_{a_i}$$

Daugman showed that this algorithm always converges to the correct solution, being the unique global minimum of the error function.

The network: Daugman uses the network shown in Fig. 2 which has no explicit teacher. The weights are fixed except for

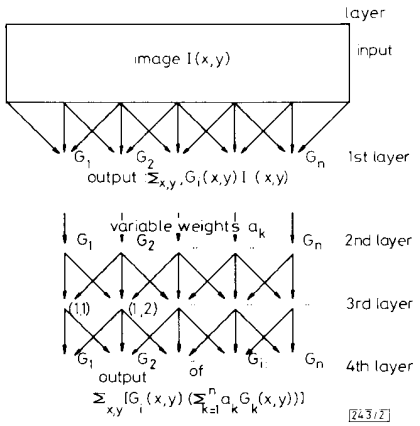

Fig. 2 *Neural network used to calculate Gabor parameters*

the variable weights a_k. These weights are updated using the difference between the output of the fourth layer and the output of the first layer. This is the value that was calculated for the gradient descent. Hence this network can be used to approximate the Gabor parameters.

To see why this is the case, we describe the precise architecture of the network. The network has four layers and all the nodes in the network have the identity function as their activation function. The nodes of the first, second and fourth layers correspond to the n Gabor functions G_i, $i = 1, \ldots, n$. The nodes of the input and the third layers are index by the points (x, y) in the image.

The inputs to the first layer are the values of the pixels $I(x, y)$. The weight connecting to the first layer node corresponding to G_i from pixel (x, y) is $G_i(x, y)$ so the output from this node is $\sum_{x,y} I(x, y) G_i(x, y)$. This is part of the value that will be required to update the approximation of the Gabor parameter. The output of the first layer are used to update the Gabor parameters.

Each node G_i in the second layer has one input which has constant value (1) connected through a variable weight whose strength is the approximated Gabor parameter for G_i. This means that nodes in the second layer produce the approximated Gabor parameter for G_i as outputs. The node for G_k is connected to the node corresponding to the pixel (x, y) in the third layer with strength $G_k(x, y)$. The output of node (x, y) in the third layer is then $\sum_{k=1}^{n} a_k G_k(x, y)$.

The strength of the connection from node (x, y) in the third layer to the node G_i of the fourth layer is $G_i(x, y)$. The output of node G_i in the final layer is then $\sum_{x,y} [(\sum_{k=1}^{n} a_k G_k(x, y)) G_i(x, y)]$ which is the other part of the expression required to update the approximated value of the Gabor parameters.

The input weights to the nodes in the second layer (Gabor parameter approximations) are then updated by the difference between the outputs of the first and the fourth layers. This corresponds exactly to gradient descent.

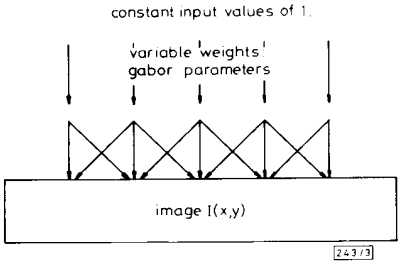

Fig. 3 *GBP network which mimics Daugman's network*

The equivalence: To demonstrate the equivalence it is sufficient to construct the GBP network that mimics Daugman's iterative approximations of the Gabor coefficients. This network is given in Fig. 3.

The network consist of two layers. All the nodes have the identity function as their activation function. The nodes in the first layer correspond to the set of Gabor functions. Each node has only one input which is the approximate Gabor parameter for that Gabor function. Each of the nodes in the first layer is then connected to the second layer whose nodes correspond to the pixels of the image. The strength of the connection between Gabor node G_i and pixel node (x, y) is fixed as $G_i(x, y)$.

The network is then trained in the standard GBP fashion. All of the inputs to the network are fixed at strength one. Ordinary back-propagation is then performed. But the only weights for which learning takes place are the weights to nodes in the first layer, which are the approximate Gabor parameters.

It is a simple matter to calculate the effect of this learning on the approximate Gabor parameter a_i for Gabor function G_i. The change in the value a_i is

$$\alpha \sum_{x,y} G_i(x, y) \left(\sum_{k=1}^{n} a_k G_k(x, y) - I(x, y) \right)$$

per iteration. Taking $\alpha = 1$ gives exactly the same update as the Daugman rule for the approximated Gabor parameters.

Conclusion: The Daugman specific architecture for calculating the Gabor parameters is a special case of the more general GBP architecture.

As the GBP architecture optimises some of the calculations involved in determining optimum Gabor parameters it is also more efficient than the corresponding Daugman network. Suppose there is an image of N pixels that are encoded using M Gabor parameters. Suppose further that K of the numbers $G_i(x, y)$ for all i and pixels (x, y) are nonzero. If all of these values are nonzero the $K = MN$. In any case there will be K connections from each Gabor layer to each pixel layer in both networks.

The GBP update now requires K multiplications and K additions for both the feed-forward pass and for the feedback pass and for the feedback pass. Hence $2K$ multiplications and $2K$ additions are required together with the N subtractions at the image layer.

For Daugman's method K multiplications and K additions are required for the computation of the output at the first layer. Similarly K multiplications and K additions are required for the computation of the outputs of both the third and fourth layers. An additional M subtractions are also required to compute the differences between the first and fourth layer outputs. So for Daugman's method $3K$ multiplications and $3K + M$ additions are required. This is K multiplications and $K + M - N$ additions more than the GBP.

D. COHEN
J. SHAWE-TAYLOR

Department of Computer Science
Royal Holloway and Bedford New College, Egham, United Kingdom

24th April 1990

References

1 DAUGMAN, J.: 'Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression', *IEEE Trans.*, 1988, **ASSP-36**, (7)
2 HINTON, G.: 'Generative back-propagation'. Abstracts 1st INNS, 1988
3 MCCLELLAND, J. L., and RUMELHART, D.: 'Parallel distributed processing' (MIT Press, 1986), Vols. 1 and 2
4 MCCLELLAND, J. L., and RUMELHART, D.: 'Learning representations by back propagating errors', *Nature*, 1986, **323**
5 MCCLELLAND, J. L., and RUMELHART, D.: 'Explorations in parallel distributed processing' (MIT Press, 1988)
6 SHAWE-TAYLOR, J., and COHEN, D.: 'The linear programming algorithm', to appear in *Neural Networks*

TWO-SECTION DISTRIBUTED FEEDBACK LASERS FOR INCOHERENT FREQUENCY-SHIFT-KEYING TRANSMISSION SYSTEMS

Indexing terms: Lasers and laser applications, Frequency shift keying, Optical fibres

Fabrication and performance characteristics of two-section distributed feedback lasers operating at 1·55 μm are presented. Using this type of laser, a 622 Mbit/s incoherent frequency-shift-keying system over a 40 km fibre has been demonstrated with a receiver sensitivity of −41 dBm at a BER of 10^{-9}.

Introduction: Two-section distributed feedback (2S-DFB) lasers have been developed to achieve large wavelength tunability and uniform frequency modulation (FM) response that are essential for frequency shift keying (FSK) transmission system applications.[1,2] A nonuniform FM response of a laser under FSK modulation will result in a signal pulse distortion at the receiver.[3,4] It has been shown that conventional single-electrode DFB lasers usually have an amplitude dip in their FM response caused by a thermal effect. The amplitude dip (also called thermal dip) is caused by the opposite phases of the change in refractive index caused by thermal modulation and carrier modulation.[5] The frequencies of the thermal dips are typically between a few kilohertz and a few megahertz. This frequency range is within the range of receiver bandwidth for the transmission systems operating at a few hundred Mbit/s to a few Gbit/s. Without appropriate equalisation, the signal pulse (especially in a nonreturn-to-zero system) will be distorted by the thermal dip of the laser under FSK modulation. 2S-DFB lasers are found to have more uniform FM response with thermal dips located at much lower frequencies (below one kiloHertz). The frequencies of the thermal dips are essentially outside the receiver band. The pseudorandom FSK signal generated from a two-section DFB laser can be detected at the receiver with a minimal distortion. We report the fabrication and performance characteristics of the InGaAsP 2S-DFB lasers operating at the 1·55 μm region and their applications in a 622 Mbit/s incoherent FSK system.

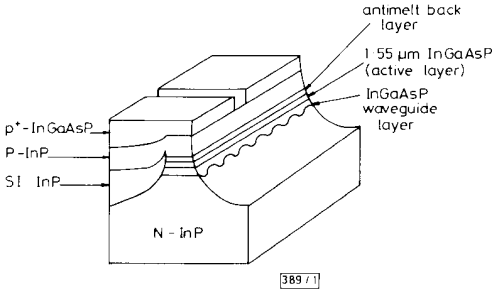

Fig. 1 *Two-section DFB laser structure*

Device fabrication and performance characteristics: The 2S-DFB lasers are based on the capped-mesa buried heterostructure (CMBH).[6] A schematic diagram of the 2S-DFB laser structure is shown in Fig. 1. The fabrication of the CMBH lasers is described in detail in Reference 6. The lasers are 508 μm long and the two sections are of equal length. The isolation resistance between the two sections is greater than 1 kΩ. The front facet and the back facet have antireflection coating (~1%) and high-reflectivity coating (~65%), respectively.

The light output from front facet against bias current characteristic of a laser at 20°C is shown in Fig. 2*a*. The front section current is kept at 80 mA, 100 mA, and 120 mA while the back section current is varied from 0 to 80 mA. The typical threshold currents of the lasers were in the range of 15 to 30 mA when both sections are equally biased. The lasers emitted at a single frequency with side mode suppression ratios greater than 35 dB up to power levels >20 mW. Fig. 2*b* shows the CW spectrum of the 2S-DFB laser used in the system experiment. The bias currents of the laser are 80 mA and 70 mA for the front and the back sections, respectively.

For FSK type system applications, it is essential that lasers exhibit reasonable FM sensitivities and uniform FM response extending to the GHz range. The measured FM response as a function of modulation frequency for a 2S-DFB laser is shown in Fig. 3. The modulation current is applied to the front section. Under this bias condition ($I_f = 80$ mA and $I_b = 70$ mA), the frequency at which the amplitude of the FM response dips is below 3 kHz and the phase of FM response is essentially flat for modulation frequencies in the 3 kHz to 500 MHz range. To obtain high enough FM sensitivities and thermal dips at low modulation frequencies (e.g., <10 kHz) for the lasers, the front section is biased high above threshold, and the modulation current should be applied to the front section.

CW spectral linewidths of several lasers were measured using a test set employing the delayed self-heterodyne principle. Fig. 4 shows the statistical data of the linewidth for several lasers under the bias conditions such that reasonable FM sensitivities (>200 MHz/mA) and low FM dip frequencies (<10 kHz) are obtained. The latter represents criteria for lightwave FSK system application. It is found that 80% of