

4

This Sheet Has Been Replaced

- 1 -



NEURAL NETWORKS AS FOR
VIDEO SIGNAL PROCESSING

5 This invention relates to the use of neural networks in
a television signal processing system.

Typically video signal processing circuitry is designed
by assuming a desired transfer function and then determining the
circuit elements necessary to realize the function. For example, in
designing a low pass filter certain "necessary" parameters are
10 assumed such as bandwidth, phase linearity ripple, etc. Then with
the aid of design tools the filter is synthesized and finally tested
to determine whether it provides the assumed response
characteristics. The problem with this approach is that the
preliminary assumptions may not comport with the desired
15 response.

A better approach is to assemble circuitry capable of a
variety of responses, and then train it to produce desired output
signals for a representative set of input signals. The training set
must truly represent the conditions found in actual operation, and
20 the system should be sufficiently flexible to generalize to input
signals not present during training. Once the training has been
accomplished, the circuit parameters may be utilized to fabricate
circuitry to emulate the "trained" system.

A class of circuitry termed "neural networks" is
25 capable of being trained to produce desired output responses to
sets of input signals. A generalized description of such networks
may be found in "Neural Nets for Adaptive Filtering and Adaptive
Pattern Recognition" by Bernard Widrow et al., Computer, March
1988, pp. 25-39, and "An Introduction to Computing with Neural
30 Nets" by Richard P. Lippmann, IEEE ASSP Magazine, April 1987,
pp. 4-22.

Neural networks are conducive to recognizing patterns
or particular signal signatures and have typically been utilized in
pattern recognition application. In general neural networks have
35 not been employed as linear signal processing elements, since the
component parts (perceptrons) are inherently non-linear.

The present inventors have found that neural
networks may be advantageously applied in processing of video
signals, in both linear and non-linear signal paths.



The present invention is a video signal processing circuit including at least one neural network in a signal processing path. In one embodiment a neural network is implemented to detect signal defects in a composite video signal. In a second
5 embodiment a neural network is arranged to provide interpolated signal values to be substituted for defective signal or as interstitial lines in an interlace-to-noninterlace scan converter.

Brief Description of the Drawings

10 FIGURE 1 is a block diagram of a television receiver including a neural network embodying the invention.

FIGURE 2 is a block diagram of a neural network for passing signal components having the appearance of signal defects.

15 FIGURE 3 is a block diagram of a neural network for performing a correlation function.

FIGURE 4 is a block diagram of a neural network for performing signal interpolation.

20 FIGURES 5-8 are tables of coefficients used in exemplary embodiments of the networks illustrated in FIGURES 2-4.

The invention will be described in the environment of an impulse noise or signal defect detection and correction system for processing composite video signals, however it will readily be
25 recognized that the concepts may be applied to more general video signal processing applications. The assumption is made that the video signals are in sampled data format, in binary or PCM form, and with the samples occurring at a rate f_s equal to four times the color subcarrier frequency.

30 As used herein a neural network is a multi-layered combination of perceptrons. Each perceptron includes a summing node for summing a plurality of weighted samples, and applying an offset Θ . The output of the summing node is then applied to a non-linear circuit with a sigmoidal response. The output of each
35 perceptron, Z_i , is of the form

$$Z_i = 1/(e^{-y_i(x)} + 1) \quad (1)$$

where



$$y_i(x) = \sum_j w_{ij} x_j - \Theta_i \quad (2)$$

where Θ_i are respective offset values, x_j are input sample values and w_{ij} are respective weighting values applied to respective sample inputs. Note that the range of output values Z_i is between zero and one. The function Z_i is generally "S" shaped and is designated herein as a sigmoidal function. An alternative sigmoidal function to equation (2) is the hyperbolic tangent, $\tanh(y_i(x)/2)$.

The perceptrons in a given layer all operate on the same set of input samples x_j . The perceptrons in successive layers operate on output values generated by a prior layer of perceptrons. The architecture of the network, that is the number of nodes, their interconnections and the external inputs are design choices as are the definitions of the desired output signals. The weights and offset values are determined in the laboratory in a training environment similar to the application in which the network will be implemented.

Training of neural networks is documented in the literature, however a brief explanation of network training is provided below.

For any degree of accuracy and any continuous function from R^N to vectors in R^M for which each component is between 0 and 1, there is a multi-layer perceptron which approximates the function to the given degree of accuracy, and in fact there is a method for constructing a two-layer perceptron to approximate a known function to a given degree of accuracy. However, part of the value of multi-layer perceptrons lies in the possibility of approximating unknown functions. This can be done if a set of examples is available, that is if one has a set of inputs from R^N and for each of those inputs, a desired output from R^M . This set of examples does not have to reflect an actual function. There may be a probabilistic relationship between the inputs and outputs, in which case the neural network will approximate a function which gives a statistical relationship between the inputs and outputs, such as the mean output given the input, the most probable output given the input, etc. The person training the network must choose a set of examples, the so-called "training set", appropriately.



The weights and thresholds of the network are adjusted to best approximate the training set according to some measure of the error of the approximation, most commonly the mean squared error on the training set. Methods for adjusting the weights and thresholds are currently a very active area of research, but almost all methods involve gradient descent. To do this, the gradient of the network's error for each example with respect to the weights and thresholds is computed. The errors and gradients for each example can then be averaged, giving the mean error and its gradient, and one can then apply classical optimization techniques, such as one of the conjugate gradient methods. Or one can perform a crude method of gradient descent called back-propagation. In back-propagation, each weight and threshold is adjusted by simply adding to it a small constant multiplied by the negative of the derivative of the network error with respect to the weight or threshold in question. The error can be either the error for a particular example, the error averaged over some subset of the training set, or the error averaged over the whole training set. A common modification of this procedure is to compute a decaying average for each derivative, in which the average is taken over the recently computed errors, whether they are for single examples, or sets of examples. These averages are then used to adjust the weights and thresholds instead of the derivatives for the most recently computed error.

The calculation of the gradient of the network's error with respect to the weights and thresholds is simplified by the structure of the network. The function computed by the network is a composition of several functions, each of which is realized by a layer in the network. Mathematically, one simply applies the chain rule to compute the derivative of the network's error with respect to the outputs of the nodes in a layer. These derivatives are commonly called the node's errors. Consider that these derivatives are known for some layer. The chain rule then gives the derivatives of the network's error with respect to this layer's weights and thresholds, but also it is useful for calculating the derivatives of the network's error with respect to the outputs of the nodes in the previous layer. This derivative for a node in the previous layer is the sum over the current layer's nodes of the



current layer node's error multiplied by the derivative of the current layer node's output with respect to the previous layer node's output.

An example of back-propagation training takes the form:

- 5 1. Compute the network's output for the current example.
2. Set the error of each node in the net to zero.
3. Set the error of each of the nodes in the output layer to the derivative of the network's error with respect to that node's output.
- 10 4. Set the current layer to the output layer.
5. For each node in the current layer, use its error to compute the derivative of the network's error with respect to this node's weights and threshold.
- 15 6. For each node in the current layer and each of its inputs, if the input is from a node in a previous layer, calculate the derivative of the current node's output with respect to the input, multiply by the current node's error, and add the result to the input node's error.
- 20 7. If there are no previous layers, the calculation is done. Otherwise, set the new current layer to the previous layer and go back to 5.

25 Thus the name back-propagation, since the error is propagated back from the output layer through each previous layer in order.

FIGURE 1 shows a portion of a television signal processing system including several neural networks for performing various functions. In FIGURE 1 a sampled data
30 composite video signal is provided by a source 10. In a television receiver the source 10 may include an antenna, tuner and IF circuitry and a signal sampling circuit (e.g., an analog-to-digital converter). Alternatively, in a video disc player or a video
cassette recorder, the source 10 may include a recorded medium
35 and pickup circuitry. The signal provided by the source 10 is presumed to be susceptible of including defects such as impulse noise or signal drop outs. The signal from source 10 is coupled to a defect detector 11 and a defect corrector 12. The detector 11



generates a control signal which is applied to control the defect corrector 12.

5 The defect corrector 12 includes a multiplexor and an interpolator. Signal from the source 10 is applied to one input of the multiplexor and to the interpolator. Interpolated samples generated by the interpolator are coupled to a second input of the multiplexor. The multiplexor is conditioned by the control signal to pass signal from the source 10 in the absence of signal defects and to pass interpolated samples on the occurrence of signal
10 defects. The signal provided by the multiplexor is the signal output by the defect corrector 12. Note, respective compensating delay elements may have to be included in the defect corrector to temporally align the interpolated and real signal, and the control signal with corresponding defects.

15 Output signal from the defect corrector 12 is applied to conventional video signal processing circuitry 13 which may include luminance/chrominance separating circuitry, chrominance demodulators, saturation-control, peaking, etc.

20 The luminance and chrominance signals from the processor 13 are coupled to an interlace-to-noninterlace scan converter 15 and to a memory device 14. The memory device 14 provides samples to a neural network interpolating filter 16 which generates interstitial lines of video signal. These interstitial lines of signal are also applied to the scan converter 15 wherein
25 real lines of video signal from the processor 13 are alternated with interpolated lines of video signal from the filter 16, which alternating lines are time compressed to form a twice rate non-interlaced signal. Output signal from the scan converter is provided to a display device (not shown).

30 FIGURES 2 and 3 illustrate portions of the neural network defect detector 11. The neural network of FIGURE 2 operates on successive groupings of samples within a video line and is trained to recognize occurrences of defect like signal occurrences. In general, the signal provided by this network is a
35 sequence of pulses, one pulse for each defect-like occurrence.

Certain images (e.g., vertical stripes or lines) generate horizontal signal components similar to signal defects of the impulse/dropout type, which will be recognized as defects by the



circuitry of FIGURE 2 and generate false detection pulses.

However, these horizontal signal components are normally redundant line-to-line, that is, they are spatially correlated.

Defects of the impulse/dropout type are random and spatially

5 non-correlated. The circuitry of FIGURE 3 operates on a two-dimensional matrix of samples provided by the circuitry of FIGURE 2, and is trained to recognize only spatially non-correlated signal defect occurrences. While the output signal from the FIGURE 2 circuitry may be used directly as a defect indicating
10 signal for controlling the defect corrector 12, inclusion of the FIGURE 3 circuitry greatly reduces the potential for false defect detection.

Referring to FIGURE 2, signal from the source 10 is applied to a tapped delay line 20 having output taps A-I, each tap
15 providing successive incremental delay of for example one signal sample period. Signals from the taps are applied to three similar perceptrons (PERCEPTRON 1, PERCEPTRON 2 AND PERCEPTRON 3). In the respective perceptrons the signals from the taps A-I are applied to respective weighting circuits W1-W9 wherein these
20 respective signals are scaled by different weighting values. The weighted signals are applied to a summing circuit 22. An offset value Θ is also applied to summing circuit 22 which generates successive samples $y_i(x)$ as defined by equation (2). The $y_i(x)$ samples are coupled to a nonlinear or sigmoidal element 24 which
25 provides perceptron output samples according to the function defined by equation (1). Element 24 will typically be realized using a read-only-memory (ROM) which is preprogrammed with the function defined by equation 1. The $y_i(x)$ values or sums are applied as address codes to the address input port of the ROM and
30 Z_i output values corresponding to the applied y_i values are available at the ROM's output port.

For the defect detection application only a relatively small number of values need be stored in the ROM. That is for n-bit words y_i , y_i can assume 2^n-1 different values and hence 2^n-1
35 different ROM addresses. When n is equal to, e.g., 6, there are 64 possible y_i values. However, the number of stored values or effective addresses may be limited to a much smaller number.



Samples are continuously loaded into the delay line 20 at the sample rate F_s and the three perceptrons generate output values from groupings of nine successive samples. It has been determined for the defect detection function that it is only
5 necessary to examine every other set of nine samples. That is, if the successive samples are numbered 1-15, the respective sets of nine samples processed by the three perceptrons are samples (1-9), (3-11), (5-13), (7-15), etc. As such the three perceptrons may be arranged to operate at one half the sample rate or $F_s/2$.

10 Output samples provided by PERCEPTRON 1, PERCEPTRON 2 AND PERCEPTRON 3 are coupled to a fourth perceptron 26 (PERCEPTRON 4) wherein they are respectively weighted, summed, offset and processed by the nonlinear sigmoidal element. For this application it has been found that it is
15 sufficient to examine only every other set of samples provided by the first layer of perceptrons (PERCEPTRON 1, PERCEPTRON 2 AND PERCEPTRON 3). Thus PERCEPTRON 4 may be arranged to operate at one quarter of the input sample rate or $F_s/4$. Output samples from the fourth perceptron are lowpass filtered (28) to remove
20 high frequency components and then latched (30) to temporally stabilize output samples. Exemplary weighting coefficients and offset values utilized in the perceptrons of FIGURE 2 are listed in TABLE I wherein the taps 1-9 correspond to signals A-I respectively. The taps 10, 11 and 12 correspond to the outputs of
25 PERCEPTRONS 1-3, respectively. The columns labeled 1-4 list the respective weights and offset values for PERCEPTRONS 1-4, respectively.

FIGURE 3 illustrates a neural network for detecting non-spatially correlated pulses provided by the neural net of
30 FIGURE 2. Samples from the latch 30 are coupled to a storage device 50 which concurrently provides five samples from each of three successive horizontal lines L_{n-1} , L_n and L_{n+1} . These 15 samples $S1-S15$ are applied to each of three perceptrons $P1$, $P2$ and $P3$. These perceptrons are constructed similar to the
35 PERCEPTRON 1 of FIGURE 2 except that there are 15 input signal paths with 15 weighting circuits rather than nine. It has been found that it is only necessary to examine alternate blocks of 15 samples, hence the perceptrons $P1$, $P2$ and $P3$ may be arranged to



operate at half the sample rate of latch 30 or $F_s/8$. Output samples from the perceptrons P1, P2 and P3 are applied to a three input perceptron P4 which provides the defect control signal applied to the defect corrector 12. Perceptron P4 is also similar to
5 the perceptron 26 of FIGURE 2.

Exemplary offset values and weighting coefficients applied to the perceptrons P1-P4 are listed in TABLE II. In TABLE II the taps labeled 1-15 correspond to samples S1-S15 respectively with respect to P1-P3. Regarding P4, the taps labeled
10 1-3 correspond to the signals provided by the perceptrons P1-P3 respectively.

FIGURE 4 illustrates an exemplary neural network which may be implemented as the interpolator in the defect corrector 12. This particular network is configured to provide
15 interpolated samples for a composite NTSC video signal including a chrominance component modulated in quadrature on a subcarrier which alternates in phase from horizontal line-to-horizontal line. Interpolated samples are generated at the sample rate F_s from five samples from each of two lines L_{n-2} , L_{n-1} preceding the
20 current line L_n and from five samples from each of two lines L_{n+1} , L_{n+2} , following the current line. In FIGURE 4 the spatial point for which a current sample is interpolated is indicated by the open circle. The twenty samples used in the calculation are designated X1-X20.

Input samples from the signal source 10 are applied to
25 a storage element 70 which concurrently provides the 20 samples X1-X20 from segments of four horizontal lines. These twenty samples are coupled to each of 10 perceptrons P1A-P10A. The ten output samples F1-F10 generated by the respective
30 perceptron P1A-P10A are applied to each of five perceptrons P1B-P5B. The five output samples F1A-F5A generated by the perceptrons P1B-P5B are coupled to a five input perceptron P6B which provides the interpolated sample. Each of the perceptrons in the FIGURE 4 embodiment are similar to PERCEPTRON 1 of
35 FIGURE 2 modified to accommodate the number of input signals applied respectively thereto. The offset values and weighting coefficients for the perceptrons P1A-P10A are listed in TABLE IIIA. The offset values and weighting coefficients for perceptrons



P1B-P6B are listed in TABLE IIIB. In TABLE IIIA the tap numbers 1-20 correspond respectively to the samples designated X1-X20. In TABLE IIIB regarding perceptrons P1B-P5B the taps numbered 1-10 correspond to the output signals provided by the perceptrons P1A-P10A respectively. Similarly, regarding perceptron P6B the taps numbered 1-5 correspond to output signals provided by the perceptrons P1B-P5B respectively.

The interpolating neural network illustrated in FIGURE 4 appears to be relatively complex, however it provides excellent substitution samples. The complexity is a consequence of the signal being composite video. Separate, simpler interpolators operating on component signals may be implemented subsequent to luminance/chrominance separation. However, the overall complexity will be about the same, and due to effects of the defects on the luminance/chrominance separation circuitry, performance may be compromised.

The FIGURE 4 circuitry is exemplary. Circuitry arranged to process a different matrix of samples are contemplated which include for example a neural network utilizing samples from lines L_{n-1} , L_n and L_{n+1} . Regardless of the choice of the sample matrix processed either vertical or horizontal resolution will be somewhat compromised and this compromise is typically determinative of the circuitry selected for a particular application,

The neural network of FIGURE 4 operates continuously and thus produces interpolated samples for every sample point in line L_n . As such circuitry of this type is amenable to providing interstitial lines of video signal for interlace-to-noninterlace scan conversion. For scan conversion however, the matrix illustrated in element 70 would typically represent samples of a component signal. The line L_n is an interstitial line being generated. Lines L_{n-1} and L_{n+1} represent lines from the current field of interlaced signal and lines L_{n-2} and L_{n+2} may represent lines from the previous field of interlaced signal. Nominally interlace-to-noninterlace scan converters use lines from the previous field for interstitial lines except where motion exists between fields/frames. When motion is present, interpolated values, generated from samples within the current field are substituted

for the interstitial line. An interpolating filter of the form shown in FIGURE 4 and operating on samples from the current and previous field may be arranged for generating interstitial lines. More typically however, the neural network interpolating filter
5 will be arranged to process samples from two adjacent lines (L_n-1 , L_n+1) from the current field of interlaced video signal. Thus a corresponding neural network interpolating filter for generating interstitial lines for scan conversion may take the form of the FIGURE 4 interpolator but with the number of inputs to the first
10 layer of perceptrons (P1A-P10A) reduced to, e.g., ten (X6-X15). The weighting coefficients and offset values for the entire network will be determined by a training exercise.

In the claims that follow the term television receiver is meant to include, for example such devices as conventional
15 television receivers, video cassette recorders, and video disc players.

This Sheet Has Been Replaced



5

CLAIMS:

1. A television receiver including a neural network for processing recovered video signal, characterized by said neural network including a first plurality of perceptrons (PERC 1,2,3) arranged to be operative on said video signal in one dimension (e.g. horizontal); means (26) for combining signals provided by said first plurality of perceptrons to provide an intermediate signal; a second plurality of perceptrons (PERC P1-10A, P1-5B) arranged to be operative on said intermediate signal in two dimensions (e.g. horizontal and vertical); and means (P6B) for combining signals provided by said second plurality of perceptrons to form a processed video signal, and utilization means responsive to said processed video signal.

2. The receiver set forth in claim 1 characterized in that said video signal occurs as samples at a predetermined sample rate; said first plurality of perceptrons provide processed signals at a rate proportional to said predetermined rate; and said neural network further includes a low pass filter for filtering said intermediate signal before application to said second plurality of perceptrons.

3. The receiver set forth in claim 2 characterized by means for conditioning said second plurality of perceptrons to provide processed signal at a lesser rate than said first plurality of perceptrons.

4. The receiver set forth in claim 2 characterized in that said first plurality of perceptrons are arranged to simultaneously operate on groups of N samples, and wherein successive groups of N samples are delayed P sample intervals, (N and P integers with $0 < P < N$).



This Sheet Has Been Replaced





5. The receiver set forth in claim 1 further characterized by a second neural network responsive to said video signal for providing interpolated video signal; and means
5 responsive to said processed video signal for substituting said interpolated video signal for defects in said video signal.

6. The receiver set forth in claim 1 characterized in that said neural network is an interpolating filter for generating
10 interpolated samples.

7. The receiver set forth in claim 1 characterized in that said neural network includes;

delay means, responsive to said received video signal
15 for concurrently providing a plurality of relatively delayed samples;

a first layer of N perceptrons (N an integer), each of said N perceptrons responsive to said plurality of relatively delayed samples;

20 at least one further perceptron responsive to signals provided by said first layer of N perceptrons, for providing said intermediate signal.

8. The receiver set forth in claim 1 characterized in
25 that each of said perceptrons includes means for weighting a plurality of delayed representations of said received video signal, means for providing sums of weighted signals provided by said weighting means, and means for processing said sums with a sigmoidal transfer function.

30 9. The receiver set forth in claim 8 characterized in that said sigmoidal transfer function, Z, is given by;

$$Z = 1/(e^{-y} + 1)$$

where y corresponds to said sums.





10. The receiver set forth in claim 9 characterized in that said means for processing said sums with a sigmoidal transfer function comprises a read only memory having address input
5 terminals coupled to receive said sums and programmed at respective memory locations with respective values Z.



* C L H S *