(19) **United States**

(12) **Patent Application Publication** (10) Pub. No.: **US 2006/0053086 A1**
    Fazzini   (43) **Pub. Date: Mar. 9, 2006**

(54) **METHOD AND APPARATUS FOR REMOVING VISIBLE ARTEFACTS IN VIDEO IMAGES**

(76) Inventor: **Paolo Giuseppe Fazzini**, Watford (GB)

Correspondence Address:
FLYNN THIEL BOUTELL & TANIS, P.C.
2026 RAMBLING ROAD
KALAMAZOO, MI 49008-1631 (US)

Publication Classification

(57)   **ABSTRACT**

A method and apparatus are provided for removing regularly occurring visible artefacts in decompressed video images. Firstly a decompressed video signal is received. This is filtered frame by frame to extract data related to the artefacts. The thus extracted data is then processed in a neural network processor which has been trained to identify the artefacts in order to produce data identifying their locations. The video signal is then corrected to reduce the effect of the thus identified artefacts.
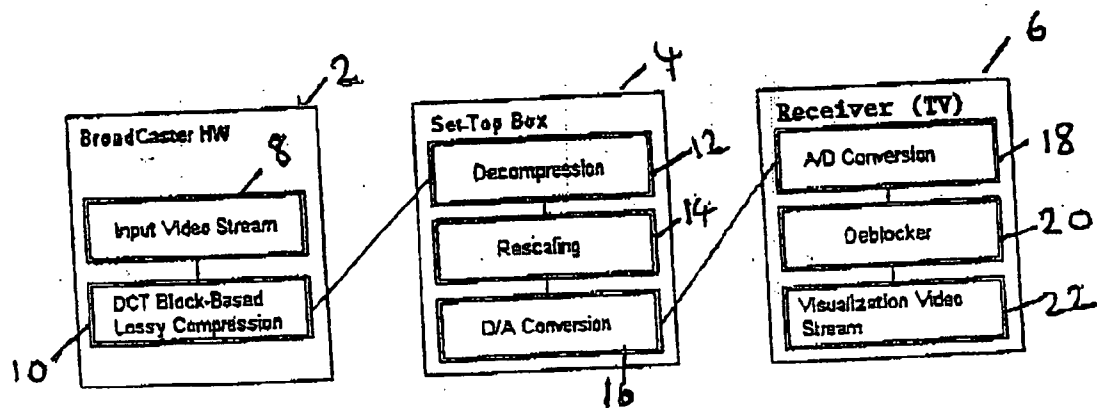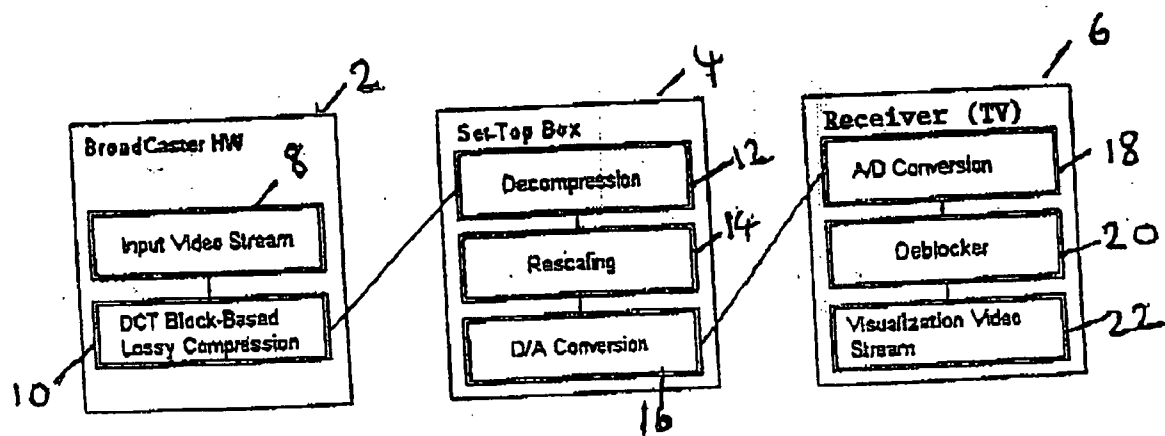
FIGURE 1

FIGURE 2

FIGURE 3

Input

Middle
Layer

Output
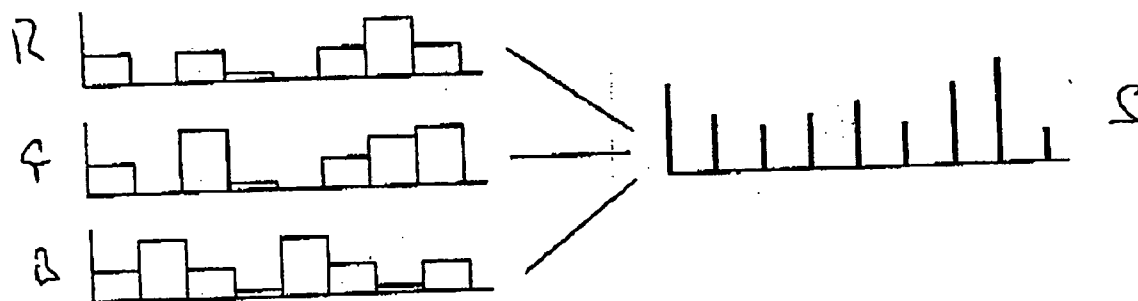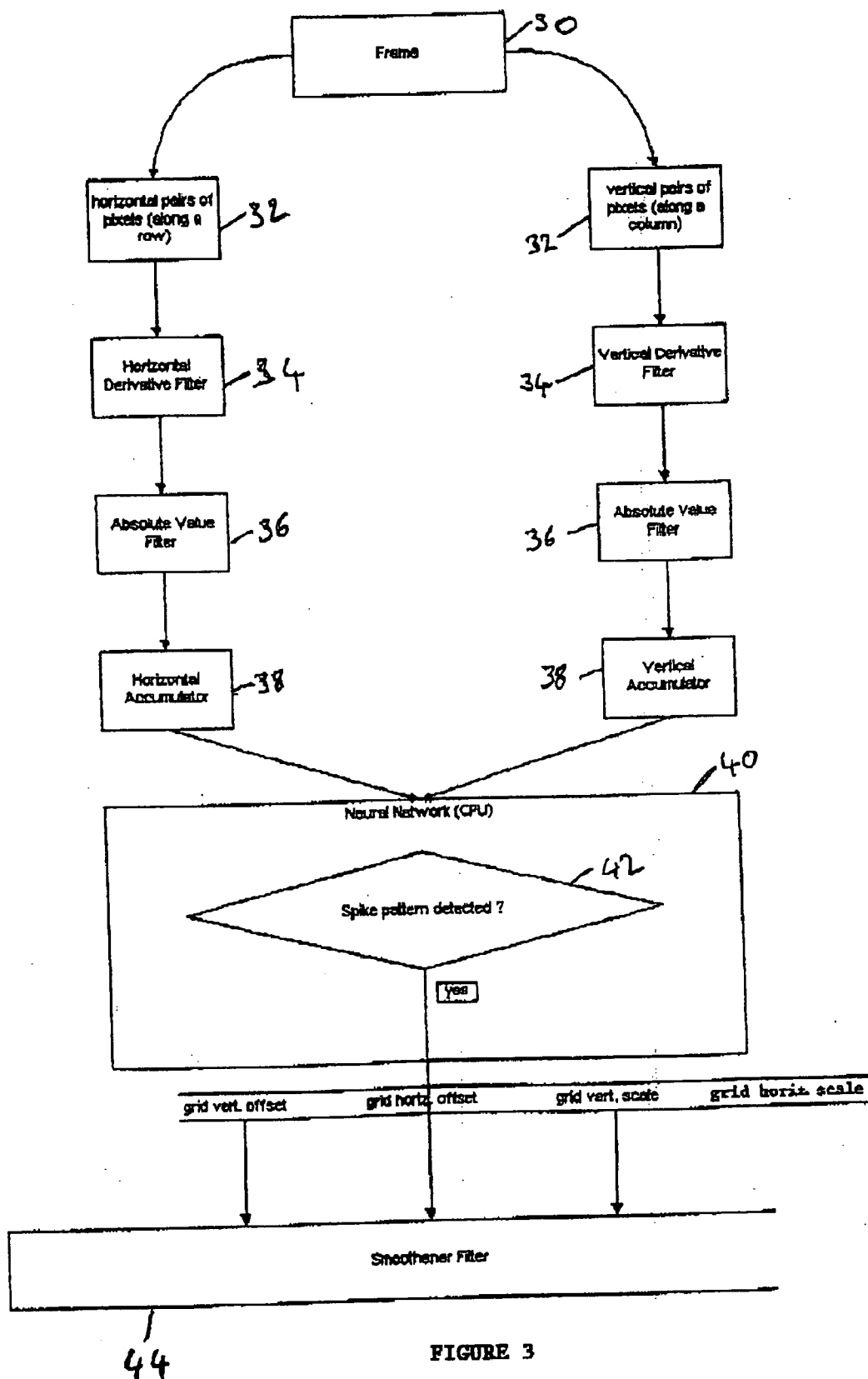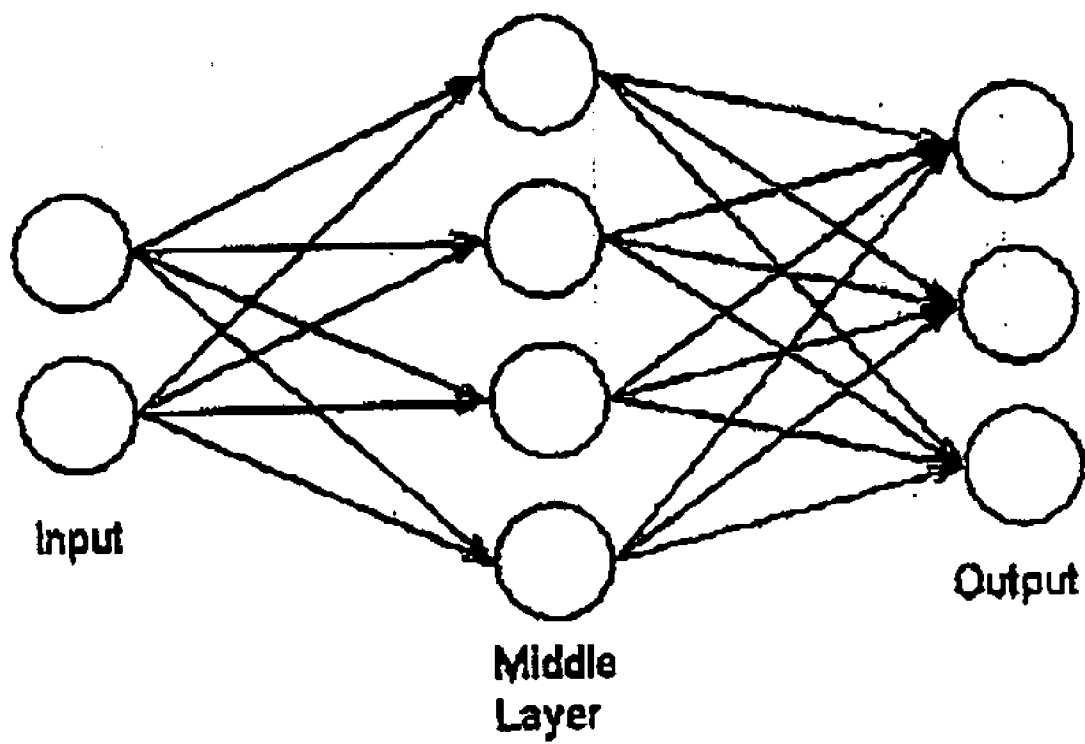
FIGURE 4

METHOD AND APPARATUS FOR REMOVING VISIBLE ARTEFACTS IN VIDEO IMAGES

[0001]   This invention relates to a method and apparatus for removing visible artefacts in video images, such as the blocky artefacts introduced by compression and subsequent decompression using MPEG which operates on a block by block basis.

[0002]   When compression and decompression schemes are used on video images there is inevitably some loss of data. Most compression/decompression schemes such as MPEG operate on a block by block basis. As the image data of adjacent blocks is different the elements of the data which may be lost during the compression/decompression process may differ from block to block. When the image is viewed after decompression this can lead to visible blocky artefacts in the image. This mismatch in adjustment of data between adjacent blocks can apply to any characteristic of the video image such as RGB values or other methods of defining the video data.

[0003]   If the position of the blocks in the image is known then a smoothing filter can be applied across the block transition for each characteristic of the image to help to eliminate the blockiness which arises from the compression/ decompression scheme. This will then reduce the degree to which the blocky artefacts are visible.

[0004]   Problems arise when trying to remove artefacts from a signal where some unknown scaling or shifting of the array of blocks has taken place during the compression/ decompression scheme. For example, satellite and digital television signals are typically decoded using some form of set top box. Many of these will apply a shifting or scaling to the array of blocks. The shifting or scaling will not be known to the receiver on which the image is to be viewed. Because of this shifting or scaling no final assumption can be made about the previous processing of the video stream.

[0005]   In addition to the blocky artefacts introduced by compression/decompression, other signals are present which may confuse any system designed to detect the blocky structure. These signals can be regarded as noise and are divided into three main groups.

[0006]   a) Noise from highly detailed scenes.

[0007]   b) Noise from blocks actually present in the real image such as backgrounds with bricks, metal grids etc; and

[0008]   c) Regular noise generated in the signal by the application of filters.

[0009]   The first noise group is of the random kind and is therefore potentially less of a problem. The second noise group is potentially a source of misleading blocky patterns in any analysis of the data. The third group may or may not be present because it is not known what previous video processing has been applied. The worst case scenario would be previous usage of low pass filters on a regular grid. This third kind of noise can be as significant as the blocks due to video compression.

[0010]   Preferred embodiments of the invention provide a method and apparatus for detecting regular occurring artefacts such as a blocky structure in a video image by analysing characteristics of the video signal over a sequence of frames and determining whether the data in these frames contains any regularly occurring structure such as a blocky structure. Once the structure has been detected, a correction may be made to the video image to reduce the effect of these blocky artefacts.

[0011]   Preferably the blocky structure is detected using a neural network system.

[0012]   Preferably the artefacts are detected over a number of frames of video data.

[0013]   Preferably the artefacts are repeatedly detected.

[0014]   A preferred embodiment of the invention will now be decribed in detail by way of example with reference to the accompanying drawings in which:

[0015]   FIG. 1 shows a block diagram of a system for compressing and decompressing video data in which regularly occurring artefacts may arise;

[0016]   FIG. 2 shows graphically the effect of the artefacts on three video components and the data to be derived from the artefacts;

[0017]   FIG. 3 shows a block diagram of a system for analysing each frame of a video signal over a sequence of frames in accordance with an embodiment of the invention; and

[0018]   FIG. 4 shows schematically a neural network system of the type which might be used in an embodiment of the present invention.

[0019]   In FIG. 1, there are three components, broadcaster hardware 2, a set top box or other decoder as would be provided to a user, and a receiver such as a television (6) which is used to display the signal decoded by the set top box.

[0020]   The broadcast hardware receives an input video stream 8 which is compressed using a DCT block based lossy compression system such as MPEG in a compressor 10. In the set top box, decompression of this compressed signal is performed by unit 12 and some resealing may be performed by unit 14 before a digital to analog conversion is performed to send the signal to a television receiver 6. The digital to analog conversion is performed in a D/A converter 16.

[0021]   In the receiver 6, in order to analyse more easily the image for any blocky structure, firstly an analog to digital conversion is performed by a deconverter 18. A deblocker 20 then detects a block-like structure in the image and applies appropriate correcting filtration to this in the visualisation of the video stream at 22.

[0022]   The A/D converter 18 of the receiver 6 could be eliminated if a set top box were provided with an output after the rescaling unit 14. There would then be two inputs to the television receiver, one being a digital input for use in determining the blocky structure in the system and the other being an analog signal to which correction for the blocky structure would be applied.

[0023]   The deblocker 20 analyses data along rows in the image and down columns to look for blocky artefacts. Before analysis, it is preferable to apply a 2D derivative filter over the entire frame. The object of this is to make more general the shape of data in the frame in an attempt to

extract only the features significant for further processing. Looking at this derivative filter in one dimension it is of the form:

[0024] Derivative filter=absolutevalue(data)(i+1)−data (i)).

[0025] Where 'i' is the pixel position and 'data' is the data at that pixel position for a given characteristic of the video image.

[0026] **FIG. 2** shows in general form the effect of this filter working along a line of the image for three video components R, G and B. The horizontal axis of the three graphs on the left hand side represent distance along a line and the vertical axis represents the level of the component. Thus it can be seen that if there is a blocky structure on the image there will be transitions in each of the R, G and B values at the block edges. Applying the derivative filter to these values and then summing the absolute values gives the output spikes S in the right hand graph. These spikes then represent the locations of block edges along the lines. Similar data will be available down the columns of the image.

[0027] Thus the grid structure of the blocks is transformed by the derivative filters into a set of data with spikes. If the grid is 8×8 then this will become a set of data with spikes every 8 values both vertically and horizontally.

[0028] Once this data is available for the whole image it can be processed to detect whether or not it represents a regular grid. As mentioned above, the presence of noise in the image may result in additional spikes being present. It is then necessary to use a system which is capable of determining which spikes relate to the blocky structure and which may relate to some other characteristic of the image. We have appreciated that using a neural network pattern recognition system is the best way to overcome the problems caused particularly by regularly reoccurring noise in the image.

[0029] For the set of spikes every 8 values the neural network will recognise this structure as belonging to a class of 8×8 grids. It will then output the grid step (the separation between the spikes) as a value of 8. Once the grid has been detected the position of the highest spikes gives the horizontal and vertical shift of the grid.

[0030] **FIG. 3** shows schematically the way in which an embodiment of an invention works. In this, each frame of data passes to a frame store **30**. Similar processes are then applied to the vertical columns and horizontal rows of pixels. Along each column or row, a pair of pixels is selected and stored at **32** before a derivative filter **34** is applied to this pair of pixels. This is a derivative filter of the form discussed above. An absolute value filter **36** is then applied to the output of each derivative filter and the outputs of these are applied to corresponding horizontal and vertical accumulators **38**. The purpose of these is to store the outputs of the derivative filters for pixel locations within the frame. Once the whole frame has been analysed in this manner and the horizontal and vertical accumulators store all the data for the frame, the data from them is supplied to a neural network CPU **40**.

[0031] The neural network CPU starts to analyse the received data after the first frame has been processed. It looks for spike patterns along lines and down columns of the

form shown in **FIG. 2**. Neural networks of course operate by learned behaviour. Thus, it may not initially be able to pick out the blocky structure caused by decompression. If this is the case it will not make any correction to the video signal being displayed. Over a number of frames, the spikes caused by decompression and scaling should become more apparent to the neural network CPU and will then enable it to provide a correction in the form of a smoothing filter to the video signal being displayed on the receiver.

[0032] When a spike pattern is detected at **42**, the neural network CPU outputs data relating to vertical and horizontal offsets of the grid it has detected, and vertical and horizontal scaling of the grid it has detected to a smoothning filter **44**. This filter is then applied to the video stream for display at the block intersections. The filter is a simple one in that is smooths across the video component values between adjacent pixels at the block transitions. More complex filters may be used in its most simple form this will be a filter which operates on two pixels only. In more complicated systems it could use more than one adjacent pixels.

[0033] In an MPEG system, the block sizes are typically 8×8 pixels.

[0034] It will be appreciated that when the receiver and deblocker are powered up, it may take a short interval of time before the neural network CPU starts to detect a blocky structure and therefore may take a short while before the smoothning filter is applied to the video signal being displayed. The grid will start to be detected once the system is stable.

[0035] In the presence of other grid-like structures in the image being displayed, the neural network CPU may detect these as being more significant than the blocky structure caused by decompression and resealing in the set top box **4**. In order to connect for such circumstances, a user input may be provided to the deblocker to ignore the blocky structure it has detected and to search for a further blocky structure.

[0036] The data from the frames of the video signal are repeatedly collected. The neural network circuitry includes data buffers. When these approach the storage limit, all the values are halved so this additional data can be read in. There is no waiting applied to recent or old data. All is given the same waiting.

[0037] The neural network is initially trained to recognise the particular blocky structure it is trying to identify. This is done by initially providing it with data patterns corresponding to those which might be provided with when trying to pick out the blocky structure. After a number of iterations in which patterns are shown to the neural network it is able to recognise the patterns on its own even though they may not be exactly the same as the ones shown to train it. Preferably this training procedure commences each time the system is powered up and the training patterns are stored in a separate memory coupled to the deblocking circuitry. This training procedure for the neural network leaves two optimal values being found for the free parameters which the neural network sets to drive its processing and thus its output. These parameters converge towards their optimal values during the training process. Once trained, the neural network is able to pick out the blocky structure so that correction can be made.

[0038] If the neural network does not behave as expected or something changes in the patterns it is trying to detect, the

training or tuning sequence has to be run again. If modified patterns are to be detected then these need to be included in the training set. Thus, if the deblocker has to detect blocky structure in a decompressed signal which uses a different block size then a training sequence for this has to be provided and run. A manual input can be provided to do this. It is not necessary to change the topology of the neural network itself in the circumstances.

[0039] In a preferred embodiment of the present invention, a two layer perception is used. This comprises 32 units in the input layer, 231 units in the middle layer, and 7 units in the output layer. These forms of neural networks are well known. The basic structure is shown in **FIG. 4**. Such neural networks are well known and have been widely studied. The exact operation of the neural network is not critical to the present invention.

[0040] Thus it will be seen that there are two main stages to the processor embodying this invention. Firstly, the neural network must be trained with a grid-like structure of the type it is intending to identify. Subsequently, it is supplied with grids on a frame by frame basis. It analyses this data and from it attempts to pick out the appropriate pattern before providing offset and scaling data. These can then be used to correct the image data, preferably by applying a smoothing filter across block transitions.

1. A method for removing regularly occurring visible artefacts in decompressed video images comprising the steps of:

receiving a decompressed video signal;

filtering the video signal frame by frame to extract data relating to the artefacts,

processing the thus extracted data in a neural network processor which has been trained to identify the artefacts to produce data identifying the locations of the artefacts in the video signal; and

correcting the video signal to reduce the effect of the thus identified artefacts.

2. A method according to claim 1 in which the artefacts are identified over a number of frames of the video signal.

3. A method according to claim 1 in which the artefacts are repeatedly identified.

4. A method according to claim 1 in which the correcting step comprises filtering the image to reduce the effect of the artefacts.

5. A method according to claim 4 in which the filtering step comprises applying a smoothing filter to the artefacts.

6. Apparatus for removing regularly occurring visible artefacts in a decompressed video signal comprising

means to receive the video signal;

means to filter the video signal frame by frame to extract data relating to the artefact;

a neural network processor trained to identify the artefacts and to produce date relating to the location of the artefacts in the video signal; and

means to correct the video signal to reduce the effect of the thus identified artefacts.

7. Apparatus according to claim 6 configured to identify the artefacts over a number of frames.

8. Apparatus according to claim 6 configured to repeatedly identify the artefacts.

9. Apparatus according to claims 7 in which the correcting means comprises means to filter the image to reduce the effect of the artefacts.

10. Apparatus according to claim 5 in which the means to filter comprises a smoothing filter.

11. A method for identifying regularly occurring visible artefacts in decompressed video images comprising the steps of:

receiving a decompressed video signal;

filtering the video signal frame-by-frame to extract data relating to the artefacts;

training a neural network processor to identify the artefacts;

using the neural network processor to identify locations of the artefacts; and

correcting the video signal to reduce the effect of the thus identified artefacts.

12. Apparatus for identifying regularly occurring visible artefacts in a decompressed video signal comprising:

means to receive a video signal;

means to filter the video signal frame-by-frame to extract data relating to the artefacts;

means for training a neural network processor to identify the artefacts;

using the neural network processor to produce data relating to the location of the artefacts in the video signal; and

means to correct the video signal to reduce the effect of the thus identified artefacts.

* * * * *