

5 **SYSTEMS AND METHODS FOR IMPROVING THE QUALITY OF
COMPRESSED VIDEO SIGNALS BY SMOOTHING BLOCK ARTIFACTS**

TECHNICAL FIELD

 This disclosure relates to digital video signals and more specifically to systems
and methods for improving the quality of compressed digital video signals by separating
10 the video signals into Deblock and Detail regions and by smoothing the Deblock region.

5

BACKGROUND OF THE INVENTION

It is well-known that video signals are represented by large amounts of digital data, relative to the amount of digital data required to represent text information or audio signals. Digital video signals consequently occupy relatively large bandwidths when transmitted at high bit rates and especially when these bit rates must correspond to the real-time digital video signals demanded by video display devices.

In particular, the simultaneous transmission and reception of a large number of distinct video signals, over such communications channels as cable or fiber, is often achieved by frequency-multiplexing or time-multiplexing these video signals in ways that share the available bandwidths in the various communication channels.

Digitized video data are typically embedded with the audio and other data in formatted media files according to internationally agreed formatting standards (*e.g.* MPEG2, MPEG4, H264). Such files are typically distributed and multiplexed over the Internet and stored separately in the digital memories of computers, cell phones, digital video recorders and on compact discs (CDs) and digital video discs (DVDs). Many of these devices are physically and indistinguishably merging into single devices.

In the process of creating formatted media files, the file data is subjected to various levels and types of digital compression in order to reduce the amount of digital data required for their representation, thereby reducing the memory storage requirement as well as the bandwidth required for their faithful simultaneous transmission when multiplexed with multiple other video files.

The Internet provides an especially complex example of the delivery of video data in which video files are multiplexed in many different ways and over many different channels (*i.e.* paths) during their downloaded transmission from the centralized server to the end user. However, in virtually all cases, it is desirable that, for a given original digital video source and a given quality of the end user's received and displayed video, the resultant video file be compressed to the smallest possible size.

Formatted video files might represent a complete digitized movie. Movie files may be downloaded 'on demand' for immediate display and viewing in real-time or for

- 5 storage in end-user recording devices, such as digital video recorders, for later viewing in real-time.

Compression of the video component of these video files therefore not only conserves bandwidth, for the purposes of transmission, but it also reduces the overall memory required to store such movie files.

- 10 At the receiver end of the abovementioned communication channels, single-user computing and storage devices are typically employed. Currently-distinct examples of such single-user devices are the personal computer and the digital set top box, either or both of which are typically output-connected to the end-user's video display device (*e.g.* TV) and input-connected, either directly or indirectly, to a wired copper distribution cable
- 15 line (*i.e.* Cable TV). Typically, this cable simultaneously carries hundreds of real-time multiplexed digital video signals and is often input-connected to an optical fiber cable that carries the terrestrial video signals from a local distributor of video programming. End-user satellite dishes are also used to receive broadcast video signals. Whether the end-user employs video signals that are delivered via terrestrial cable or satellite, end-user digital set
- 20 top boxes, or their equivalents, are typically used to receive digital video signals and to select the particular video signal that is to be viewed (*i.e.* the so-called TV Channel or TV Program). These transmitted digital video signals are often in compressed digital formats and therefore must be uncompressed in real-time after reception by the end-user.

- Most methods of video compression reduce the amount of digital video data by
- 25 retaining only a digital approximation of the original uncompressed video signal. Consequently, there exists a measurable difference between the original video signal prior to compression and the uncompressed video signal. This difference is defined as the video distortion. For a given method of video compression, the level of video distortion almost always becomes larger as the amount of data in the compressed video data is reduced by
- 30 choosing different parameters for those methods. That is, video distortion tends to increase with increasing levels of compression.

As the level of video compression is increased, the video distortion eventually becomes visible to the human vision system (HVS) and eventually this distortion becomes visibly-objectionable to the typical viewer of the real-time video on the chosen display

5 device. The video distortion is observed as a so-called artifact. An artifact is observed video content that is interpreted by the HVS as not belonging to the original uncompressed video scene.

Methods exist for significantly attenuating visibly-objectionable artifacts from compressed video, either during or after compression. Most of these methods apply only
10 to compression methods that employ the block-based Two-dimensional (2D) Discrete Cosine Transform (DCT) or approximations thereof. In the following, we refer to these methods as DCT-based. In such cases, by far the most visibly-objectionable artifact is the appearance of artifact blocks in the displayed video scene.

Methods exist for attenuating the artifact blocks typically either by searching
15 for the blocks or by requiring *a priori* knowledge of where they are located in each frame of the video.

The problem of attenuating the appearance of visibly-objectionable artifacts is especially difficult for the widely-occurring case where the video data has been previously compressed and decompressed, perhaps more than once, or where it has been previously
20 re-sized, re-formatted or color re-mixed. For example, video data may have been re-formatted from the NTSC to PAL format or converted from the RGB to the YCrCb format. In such cases, *a priori* knowledge of the locations of the artifact blocks is almost certainly unknown and therefore methods that depend on this knowledge do not work.

Methods for attenuating the appearance of video artifacts must not add
25 significantly to the overall amount of data required to represent the compressed video data. This constraint is a major design challenge. For example, each of the three colors of each pixel in each frame of the displayed video is typically represented by 8 bits, therefore amounting to 24 bits per colored pixel. For example, if pushed to the limits of compression where visibly-objectionable artifacts are evident, the H264 (DCT-based)
30 video compression standard is capable of achieving compression of video data corresponding at its low end to approximately 1/40th of a bit per pixel. This therefore corresponds to an average compression ratio of better than $40 \times 24 = 960$. Any method for attenuating the video artifacts, at this compression ratio, must therefore add an insignificant number of bits relative to 1/40th of a bit per pixel. Methods are required for

- 5 attenuating the appearance of block artifacts when the compression ratio is so high that the average number of bits per pixel is typically less than 1/40th of a bit.

For DCT-based and other block-based compression methods, the most serious visibly-objectionable artifacts are in the form of small rectangular blocks that typically vary with time, size and orientation in ways that depend on the local spatial-temporal characteristics of the video scene. In particular, the nature of the artifact blocks depends upon the local motions of objects in the video scene and on the amount of spatial detail that those objects contain. As the compression ratio is increased for a particular video, MPEG-based DCT-based video encoders allocate progressively fewer bits to the so-called quantized basis functions that represent the intensities of the pixels within each block. The number of bits that are allocated in each block is determined on the basis of extensive psycho-visual knowledge about the HVS. For example, the shapes and edges of video objects and the smooth-temporal trajectories of their motions are psycho-visually important and therefore bits must be allocated to ensure their fidelity, as in all MPEG DCT based methods.

20 As the level of compression increases, and in its goal to retain the above mentioned fidelity, the compression method (in the so-called encoder) eventually allocates a constant (or almost constant) intensity to each block and it is this block-artifact that is usually the most visually objectionable. It is estimated that if artifact blocks differ in relative uniform intensity by greater than 3% from that of their immediate neighboring blocks, then the spatial region containing these blocks is visibly-objectionable. In video scenes that have been heavily-compressed using block-based DCT-type methods, large regions of many frames contain such block artifacts.

BRIEF SUMMARY OF THE INVENTION

30 The present invention is directed to systems and methods in which, for a given amount of data required to represent a compressed video signal, the quality of the uncompressed displayed real-time video, as perceived by a typical human viewer, is improved. Systems and methods herein achieve this improvement by attenuating the appearance of blocks without necessarily having *a priori* knowledge of their locations. In some embodiments, the methods described herein attenuate the appearance of these blocks

- 5 such that the quality of the resultant real-time video, as perceived by the HVS, is improved.

In terms of the intensity difference between the compressed and uncompressed versions of a video, the blocky regions may not be the largest contributors to a mathematical metric of overall video distortion. There is typically significant
10 mathematical distortion in the detailed regions of a video but advantage is taken of the fact that the HVS does not perceive that distortion as readily as it perceives the distortion due to block artifacts.

In the embodiments discussed herein, the first step of the method separates the digital representations of each frame into two parts referred to as the Deblock region and
15 the Detail Region. The second step of the method operates on the Deblock region to attenuate the block artifacts resulting in a smoothed Deblock Region. The third step of the method recombines the smoothed Deblock region and the Detail Region.

In one embodiment, the identification of the Deblock region commences by selecting candidate regions and then comparing each candidate region against its
20 surrounding neighborhood region using a set of criteria, such as:

- a. Flatness-of-Intensity Criteria (F),
- b. Discontinuity Criteria (D) and
- c. Look-Ahead/Look-Behind Criteria (L).

5 The foregoing has outlined rather broadly the features and technical advantages of the present invention in order that the detailed description of the invention that follows may be better understood. Additional features and advantages of the invention will be described hereinafter which form the subject of the claims of the invention. It should be appreciated by those skilled in the art that the conception and specific embodiment
10 disclosed may be readily utilized as a basis for modifying or designing other structures for carrying out the same purposes of the present invention. It should also be realized by those skilled in the art that such equivalent constructions do not depart from the spirit and scope of the invention as set forth in the appended claims. The novel features which are believed to be characteristic of the invention, both as to its organization and method of operation,
15 together with further objects and advantages will be better understood from the following description when considered in connection with the accompanying figures. It is to be expressly understood, however, that each of the figures is provided for the purpose of illustration and description only and is not intended as a definition of the limits of the present invention.

5

BRIEF DESCRIPTION OF THE DRAWINGS

For a more complete understanding of the present invention, reference is now made to the following descriptions taken in conjunction with the accompanying drawing, in which:

FIGURE 1 shows a typical blocky image frame;

10

FIGURE 2 shows the Deblock region (shown in black) and Detail region (shown in white) corresponding to FIGURE 1;

[FIGURE 3 shows one example of the selection of isolated pixels in a frame;

FIGURE 4 illustrates a close up of Candidate Pixels C_i that are x pixels apart and belong to the Detail region *DET* because they do not satisfy the Deblock Criteria;

15

FIGURE 5 illustrates one embodiment of a method for assigning a block to the Deblock region by using a nine pixel crossed-mask;

FIGURE 6 shows an example of a nine pixel crossed-mask used at a particular location within an image frame;

20

FIGURE 7 shows one embodiment of a method for achieving improved video image quality; and

FIGURE 8 shows one embodiments of the use of the concepts discussed herein.

5 DETAILED DESCRIPTION OF THE INVENTION

One aspect of the disclosed embodiment is to attenuate the appearance of block artifacts in real-time video signals by identifying a region in each frame of the video signal for deblocking using flatness criteria and discontinuity criteria. Additional gradient criteria can be combined to further improve robustness. Using these concepts, the size of the video
10 file (or the number of bits required in a transmission of the video signals) can be reduced since the visual effects of artifacts associated with the reduced file size can be reduced.

One embodiment of a method to perform these concepts consists of three parts with respect to image frames of the video signal:

1. A process to identify a Deblock region (*DEB*) that distinguishes the
15 Deblock region from a so-called Detail region (*DET*);
2. An operation applied to the Deblock region *DEB* for the purposes of attenuating, by spatial smoothing, the appearance of block artifacts in the Deblock Region; and
3. A process to combine the now smoothed Deblock region obtained in
20 part 2 with the Detail Region.

In the method of this embodiment the spatial-smoothing operation does not operate outside of the Deblock Region: equivalently, it does not operate in the Detail Region. As will be discussed herein, methods are employed to determine that the spatial-smoothing operation has reached the boundaries of the Deblock region *DEB* so that
25 smoothing does not occur outside of the Deblock Region.

Video signals that have been previously subjected to block-based types of video compression (*e.g.* DCT-based compression) and decompression, and possibly to re-sizing and/or reformatting and/or color re-mixing, typically contain visibly-objectionable residues of block artifacts that first occurred during previous compression operations.
30 Therefore, the removal of block-induced artifacts cannot be completely achieved by attenuating the appearance of only those blocks that were created in the last or current compression operation.

In many cases, *a priori* information about the locations of these previously created blocks is unavailable and blocks at unknown locations often contribute to

- 5 objectionable artifacts. Embodiments of this method identify the region to be de-blocked by means of criteria that do not require *a priori* knowledge of the locations of the blocks.

In one embodiment, a flatness-of-intensity criteria method is employed and intensity-discontinuity criteria and/or intensity-gradient criteria is used to identify the Deblock region of each video frame which is to be de-blocked without specifically finding
10 or identifying the locations of individual blocks. The Deblock region typically consists, in each frame, of many unconnected sub-regions of various sizes and shapes. This method only depends on information within the image frame to identify the Deblock region in that image frame. The remaining region of the image frame, after this identification, is defined as the Detail region.

- 15 Video scenes consist of video objects. These objects are typically distinguished and recognized (by the HVS and the associated neural responses) in terms of the locations and motions of their intensity-edges and the texture of their interiors. For example, FIGURE 1 shows a typical image frame 10 that contains visibly-objectionable block artifacts that appear similarly in the corresponding video clip when displayed in real-
20 time. Typically within fractions of a second, the HVS perceives and recognizes the original objects in the corresponding video clip. For example, the face object 101 and its sub-objects, such as eyes 14 and nose 15, are quickly identified by the HVS along with the hat, which in turn contains sub-objects, such as ribbons 13 and brim 12. The HVS recognizes the large open interior of the face as skin texture having very little detail and
25 characterized by its color and smooth shading.

- While not clearly visible in the image frame of FIGURE 1, but clearly visible in the corresponding electronically displayed real-time video signal, the block artifacts have various sizes and their locations are not restricted to the locations of the blocks that were created during the last compression operation. Attenuating only the blocks that were
30 created during the last compression operation is often insufficient.

This method takes advantage of the psycho-visual property that the HVS is especially aware of, and sensitive to, those block artifacts (and their associated edge intensity-discontinuities) that are located in relatively large open areas of the image where there is almost constant intensity or smoothly-varying image intensity in the original

5 image. For example, in FIGURE 1, the HVS is relatively unaware of any block artifacts that are located between the stripes of the hat but is especially aware of, and sensitive to, the block artifacts that appear in the large open smoothly-shaded region of the skin on the face and also to block artifacts in the large open area of the left side (underneath of) the brim of the hat.

10 As another example of the sensitivity of the HVS to block artifacts, if the HVS perceives a video image of a uniformly-colored flat shaded surface, such as an illuminated wall, then block edge intensity-discontinuities of more than about 3% are visibly-objectionable whereas similar block edge intensity-discontinuities in a video image of a highly textured object, such as a highly textured field of blades of grass, are typically
15 invisible to the HVS. It is more important to attenuate blocks in large open smooth-intensity regions than in regions of high spatial detail. This method exploits this characteristic of the HVS.

However, if the above wall is occluded from view except in small isolated regions, the HVS is again relatively unaware of the block artifacts. That is, the HVS is less
20 sensitive to these blocks because, although located in regions of smooth-intensity, these regions are not sufficiently large. This method exploits this characteristic of the HVS.

As a result of applying this method to an image frame, the image is separated into at least two regions: the Deblock region and the remaining Detail region. The method can be applied in a hierarchy so that the above first-identified Detail region is then itself
25 separated into a second Deblock region and a second Detail region, and so on recursively.

FIGURE 2 shows the result 20 of identifying the Deblock region (shown in black) and the Detail region (shown in white). The eyes 14, nose 15 and mouth belong to the Detail region (white) of the face object, as does most of the right-side region of the hat having the detailed texture of stripes. However, much of the left side of the hat is a region
30 of approximately constant intensity and therefore belongs to the Deblock region while the edge of the brim 12 is a region of sharp discontinuity and corresponds to a thin line part of the Detail region.

As described in the following, criteria are employed to ensure that the Deblock region is the region in which the HVS is most aware of and sensitive to block artifacts and

5 is therefore the region that is to be de-blocked. The Detail region is then the region in which the HVS is not particularly sensitive to block artifacts. In this method, Deblocking of the Deblock region may be achieved by spatial intensity-smoothing. The process of spatial intensity-smoothing may be achieved by low pass filtering or by other means. Intensity-smoothing significantly attenuates the so-called high spatial frequencies of the
10 region to be smoothed and thereby significantly attenuates the edge-discontinuities of intensity that are associated with the edges of block artifacts.

One embodiment of this method employs spatially-invariant low pass filters to spatially-smooth the identified Deblock Region. Such filters may be Infinite Impulse Response (IIR) filters or Finite Impulse Response (FIR) filters or a combination of such
15 filters. These filters are typically low pass filters and are employed to attenuate the so-called high spatial frequencies of the Deblock region, thereby smoothing the intensities and attenuating the appearance of block artifacts.

The above definitions of the Deblock region *DEB* and the Detail region *DET* do not preclude further signal processing of either or both regions. In particular, using this
20 method, the *DET* region could be subjected to further separation into new regions *DET1* and *DEB1* where *DEB1* is a second region for Deblocking ($DEB1 \in DET$), possibly using a different Deblocking method or different filter than is used to Deblock *DEB*. *DEB1* and *DET1* are clearly sub-regions of *DET*.

Identifying the Deblock region (*DEB*) often requires an identifying algorithm
25 that has the capability to run video in real-time. For such applications, high levels of computational complexity (e.g., identifying algorithms that employ large numbers of multiply-accumulate operations (MACs) per second) tend to be less desirable than identifying algorithms that employ relatively few MACs/s and simple logic statements that operate on integers. Embodiments of this method use relatively few MACs/s. Similarly,
30 embodiments of this method ensure that the swapping of large amounts of data into and out of off-chip memory is minimized. In one embodiment of this method, the identifying algorithm for determining the region *DEB* (and thereby the region *DET*) exploits the fact that most visibly-objectionable blocks in heavily compressed video clips have almost-constant intensity throughout their interiors.

5 In one embodiment of this method, the identification of the Deblock region *DEB* commences by choosing Candidate Regions C_i in the frame. In one embodiment, these regions C_i are as small as one pixel in spatial size. Other embodiments may use candidate regions C_i that are larger than one pixel in size. Each Candidate region C_i is tested against its surrounding neighborhood region by means of a set of criteria that, if met,
10 cause C_i to be classified as belonging to the Deblock region *DEB* of the image frame. If C_i does not belong to the Deblock Region, it is set to belong to the Detail region *DET*.
Note, this does not imply that the collection of all C_i is equal to *DEB*, only that they form a sub-set of *DEB*.

 In one embodiment of this method, the set of criteria used to determine
15 whether C_i belongs to the Deblock region *DEB* may be categorized as follows:

- a. Flatness-of-Intensity Criteria (*F*),
- b. Discontinuity Criteria (*D*) and
- c. Look-Ahead/Look-Behind Criteria (*L*).

 If the above criteria (or any useful combination thereof) are satisfied, the
20 Candidate Regions C_i are assigned to the Deblock region (*i.e.*, $C_i \in DEB$). If not, then the Candidate Region C_i is assigned to the Detail Region *DET* ($C_i \in DET$). In a particular implementation, such as when Deblocking a particular video clip, all three types of criteria (*F*, *D* and *L*) may not be necessary. Further, these criteria may be adapted on the basis of the local properties of the image frame. Such local properties might be statistical or they
25 might be encoder/decoder-related properties, such as the quantization parameters or motion parameters used as part of the compression and decompression processes.

 In one embodiment of this method, the Candidate Regions C_i are chosen, for reasons of computational efficiency, such that they are sparsely-distributed in the image frame. This has the effect of significantly reducing the number of Candidate Regions C_i
30 in each frame, thereby reducing the algorithmic complexity and increasing the throughput (*i.e.*, speed) of the algorithm.

5 FIGURE 3 shows, for a small region of the frame, the selected sparsely-distributed pixels that can be employed to test the image frame of FIGURE 1 against the criteria. In FIGURE 3, the pixels 31-1 to 31-6 are 7 pixels apart from their neighbors in both the horizontal and vertical directions. These pixels occupy approximately $1/64^{\text{th}}$ of the number of pixels in the original image, implying that any pixel-based algorithm that is
10 used to identify the Deblock region only operates on $1/64^{\text{th}}$ of the number of pixels in each frame, thereby reducing the complexity and increasing the throughput relative to methods that test criteria at every pixel.

 In this illustrative example, applying the Deblocking criteria to FIGURE 1 to the sparsely-distributed Candidate region in FIGURE 3 results in the corresponding
15 sparsely-distributed $C_i \in DEB$ as illustrated in FIGURE 4.

 In one embodiment of this method, the entire Deblock region DEB is ‘grown’ from the abovementioned sparsely-distributed Candidate Regions $C_i \in DEB$ into surrounding regions.

 The identification of the Deblock region in FIGURE 2, for example, is ‘grown’
20 from the sparsely-distributed C_i in FIGURE 4 by setting N to 7 pixels, thereby ‘growing’ the sparse-distribution of Candidate region pixels C_i to the much larger Deblock region in FIGURE 2 which has the property that it is more contiguously connected.

 The above growing process spatially connects the sparsely-distributed $C_i \in DEB$ to form the entire Deblock region DEB .

25 In one embodiment of this method, the above growing process is performed on the basis of a suitable distance metric that is the horizontal or vertical distances of a pixel from the nearest Candidate region pixel C_i . For example, with Candidate region pixels C_i chosen at 7 pixels apart in the vertical and horizontal directions, the resultant Deblock region is as shown in FIGURE 2.

30 As one enhancement, the growing process is applied to the Detail region DET in order to extend the Detail region DET into the previously determined Deblock region DEB . This can be used to prevent the crossed-mask of spatially invariant low-pass

5 smoothing filters from protruding into the original Detail region and thereby avoid the possible creation of undesirable 'halo' effects. In doing so, the Detailed region may contain in its expanded boundaries unattenuated blocks, or portions thereof. This is not a practical problem because of the relative insensitivity of the HVS to such block artifacts that are proximate to Detailed Regions.

10 Alternate distance metrics may be employed. For example, a metric corresponding to all regions of the image frame within circles of a given radius centered on the Candidate Regions C_i may be employed.

The Deblock Region, that is obtained by the above or other growing processes has the property that it encompasses (*i.e.* spatially covers) the part of the image frame that
15 is to be Deblocked.

Formalizing the above growing process, the entire Deblock region DEB (or the entire Detail region DET) can be determined by surrounding each Candidate Region C_i (that meets the criteria $C_i \in DEB$ or $C_i \in DET$) by a Surrounding Grown region G_i whereupon the entire Deblock region DEB (or the entire Detail region DET) is the union of
20 all C_i and all G_i .

Equivalently, the *entire* Deblock region can be written logically as

$$DEB = \bigcup_i ((C_i \notin DET) \cup G_i) = \bigcup_i ((C_i \in DEB) \cup G_i)$$

where $\cup$ is the union of the regions and where again DET is simply the remaining parts of the image frame. Alternatively, the *entire* Detail region DET may be determined from the
25 qualifying Candidate Regions (using $C_i \notin DEB$) according to

$$DET = \bigcup_i ((C_i \notin DEB) \cup G_i) = \bigcup_i ((C_i \in DET) \cup G_i)$$

If the Grown Surrounding Regions G_i (32-1 to 32-N in FIGURE 3) are sufficiently large, they may be arranged to overlap or touch their neighbors in such a way as to create a Deblock region DEB that is contiguous over enlarged areas of the image
30 frame.

5 One embodiment of this method is illustrated in FIGURE 5 and employs a 9-pixel crossed-mask for identifying Candidate region pixels C_i to be assigned to the Deblock region or to the Detail region DET . In this embodiment, the Candidate Regions C_i are of size 1x1 pixels (*i.e.*, a single pixel). The center of the crossed-mask (pixel 51) is at pixel $x(r, c)$ where (r, c) points to the row and column location of the pixel where its
 10 intensity x is typically given by $x \in [0, 1, 2, 3, \dots, 255]$. Note that in this embodiment the crossed-mask consists of two single pixel-wide lines perpendicular to each other forming a + (cross).

Eight independent flatness criteria are labeled in FIGURE 5 as $ax, bx, cx, dx, ay, by, cy$ and dy and are applied at the 8 corresponding pixel locations. In the following,
 15 discontinuity (*i.e.*, intensity-gradient) criteria are applied inside crossed-mask 52 and optionally outside of crossed-mask 52.

FIGURE 6 shows an example of the nine pixel crossed-mask 52 used at a particular location within image frame 60. Crossed-mask 52 is illustrated for a particular location and, in general, is tested against criteria at a multiplicity of locations in the image
 20 frame. For a particular location, such as location 61 of image frame 60, the center of the crossed-mask 52 and the eight flatness-of-intensity criteria $ax, bx, cx, dx, ay, by, cy$ and dy are applied against the criteria.

The specific identification algorithms used for these eight flatness criteria can be among those known to one of ordinary skill in the art. The eight flatness criteria are
 25 satisfied by writing the logical notations $ax \in F, bx \in F, \dots, dy \in F$. If met, the corresponding region is 'sufficiently-flat' according to whatever flatness-of-intensity criterion has been employed.

The following example logical condition may be used to determine whether the overall flatness criterion for each Candidate Pixel $x(r,c)$ is satisfied:

30 if

$$(ax \in F \text{ and } bx \in F) \text{ or } (cx \in F \text{ and } dx \in F) \quad (1)$$

and

$$(ay \in F \text{ and } by \in F) \text{ or } (cy \in F \text{ and } dy \in F) \quad (2)$$

5 then

$$C_i \in Flat.$$

Equivalently, the above Boolean statement results in the truth of the statement $C_i \in Flat$ under at least one of the following three conditions:

10 a) Crossed-mask 52 lies over a 9-pixel region that is entirely of sufficiently-flat intensity, therefore including sufficiently-flat regions where 52 lies entirely in the interior of a block

OR

b) Crossed-mask 52 lies over a discontinuity at one of the four locations $(r+1, c)$ OR $(r+2, c)$ OR $(r-1, c)$ OR $(r-2, c)$ while satisfying the flatness criteria at the remaining three locations

OR

15 c) Crossed-mask 52 lies over a discontinuity at one of the four locations $(r, c+1)$ OR $(r, c+2)$ OR $(r, c-1)$ OR $(r, c-2)$ while satisfying the flatness criteria at the remaining three locations.

20 In the above-described process, as required for identifying Candidate pixels, crossed-mask 52 spatially covers the discontinuous boundaries of blocks, or parts of blocks, regardless of their locations, while maintaining the truth of the statement $C_i \in Flat$.

A more detailed explanation of the above logic is as follows. Condition a) is true when all the bracketed statements in (1) and (2) are true. Suppose there exists a
25 discontinuity at one of the locations given in b). Then statement (2) is true because one of the bracketed statements is true. Suppose there exists a discontinuity at one of the locations given in c). Then statement (1) is true because one of the bracketed statements is true.

30 Using the above Boolean logic, the flatness criterion is met when the crossed-mask 52 straddles the discontinuities that delineate the boundaries of a block, or part of a block, regardless of its location.

The employment of a specific algorithm for determining the Flatness Criteria F (that are applied to the Candidate Pixels C_i) is not crucial to the method. However, to

5 achieve high throughput capability, one example algorithm employs a simple mathematical flatness criterion for $ax, bx, cx, dx, ay, by, cy$ and dy that is, in words, ‘ the magnitude of the first-forward difference of the intensities between the horizontally adjacent and the vertically adjacent pixels’. The first-forward difference in the vertical direction, for example, of a 2D sequence $x(r, c)$ is simply $x(r+1, c) - x(r, c)$.

10 The above-discussed flatness criteria are sometimes insufficient to properly identify the region *DEB* in every region of every frame for every video signal. Assume now that the above flatness condition $C_i \in Flat$ is met for the Candidate Pixel at C_i . Then, in this method, a Magnitude-Discontinuity Criterion D may be employed to improve the discrimination between a discontinuity that is part of a boundary artifact of a block and
15 a non-artifact discontinuity that belongs to desired detail that exists in the original image, before and after its compression.

The Magnitude-Discontinuity Criterion method sets a simple threshold D below which the discontinuity is assumed to be an artifact of blocking. Writing the pixel $x(r, c)$ (61) at C_i in terms of its intensity x , the Magnitude Discontinuity Criterion is of the
20 form

$$dx < D$$

where dx is the magnitude of the discontinuity of intensity at the center (r, c) of crossed-mask 52.

The required value of D can be inferred from the intra-frame quantization step
25 size of the compression algorithm, which in turn can either be obtained from the decoder and encoder or estimated from the known compressed file size. In this way, transitions in the original image that are equal to or larger than D are not mistaken for the boundaries of blocking artifacts and thereby wrongly Deblocked. Combining this condition with the flatness condition gives the more stringent condition

30 Values for D ranging from 10% to 20% of the intensity range of $x(r, c)$ have been found to yield satisfactory attenuation of block artifacts over a wide range of different types of video scenes.

$$C_i \in Flat \text{ and } dx < D$$

5 There will almost certainly exist non-artifact discontinuities (that should therefore not be deblocked) because they were in the original uncompressed image frame. Such non-artifact discontinuities may satisfy $dx < D$ and may also reside where the surrounding region causes $C_i \in Flat$, according to the above criterion, which thereby leads to such discontinuities meeting the above criterion and thereby being wrongly
 10 classified for deblocking and therefore wrongly smoothed. However, such non-artifact discontinuities correspond to image details that are highly localized. Experiments have verified that such false deblocking is typically not objectionable to the HVS. However, to significantly reduce the probability of such rare instances of false deblocking, the following Look-Ahead (LA) and Look-Behind (LB) embodiment of the method may be
 15 employed.

 It has been found experimentally that, in particular video image frames, there may exist a set of special numerical conditions under which the required original detail in the original video frame meets both of the above local flatness and local discontinuity conditions and would therefore be falsely identified (*i.e.*, subjected to false deblocking and
 20 false smoothing). Equivalently, a small proportion of the C_i could be wrongly assigned to *DEB* instead of to *DET*. As an example of this, *a vertically-oriented* transition of intensity at the edge of an object (in the uncompressed original image frame) can meet both the flatness conditions and the discontinuity conditions for Deblocking. This can sometimes lead to visibly-objectionable artifacts in the displayed corresponding real-time video
 25 signal.

 The following LA and LB criteria are optional and address the above special numerical conditions. They do so by measuring the change in intensity of the image from crossed-mask 52 to locations suitably located outside of crossed-mask 52.

 If the above criteria $C_i \in Flat$ and $dx < D$ are met and also exceed a ‘looking
 30 ahead LA’ threshold criterion or a ‘looking back LB’ threshold criterion L , then the candidate C_i pixel is not assigned to the Deblock Region. In terms of the magnitudes of derivatives, one embodiment of the LA and LB criteria is:

if

$$(dxA \geq L) \text{ OR } (dxB \geq L) \text{ OR } (dxC \geq L) \text{ OR } (dxD \geq L)$$

5 then

$$C_i \notin DEB$$

In the above, terms such as $(dx_A \geq L)$ simply mean that the magnitude of the LA magnitude-gradient or change criterion dx as measured from the location (r, c) out to the location of pixel A in this case is greater than or equal to the threshold number L . The other three terms have similar meanings but with respect to pixels at locations B , C and D .

The effect of the above LA and LB criteria is to ensure that deblocking cannot occur within a certain distance of an intensity-magnitude change of L or greater.

These LA and LB constraints have the desired effect of reducing the probability of false deblocking. The LA and LB constraints are also sufficient to prevent undesirable deblocking in regions that are in the close neighborhoods of where the magnitude of the intensity gradient is high, regardless of the flatness and discontinuity criteria.

An embodiment of the combined criteria, obtained by combining the above three sets of criteria, for assigning a pixel at C_i to the Deblock region DEB , can be expressed as an example criterion as follows:

if

$$C_i \in Flat \text{ AND } x < D \text{ AND } ((dx_A < L \text{ AND } dx_B < L \text{ AND } dx_C < L \text{ AND } dx_D < L))$$

then

$$C_i \in DEB$$

As an embodiment of this method, the truth of the above may be determined in hardware using fast logical operations on short integers. Evaluation of the above criteria over many videos of different types has verified its robustness in properly identifying the Deblock Regions DEB (and thereby the complementary Detail Regions DET).

Many previously-processed videos have 'spread-out' block edge-discontinuities. While being visibly-objectionable, spread-out block edge-discontinuities straddle more than one pixel in the vertical and/or horizontal directions. This can cause

- 5 incorrect classification of block edge-discontinuities to the Deblock Region, as described by example in the following.

For example, consider a horizontal 1-pixel-wide discontinuity of magnitude 40 that separates flat-intensity regions that satisfy $C_i \in Flat$, occurring from say $x(r, c) = 100$ to $x(r, c + 1) = 140$ with the criterion discontinuity threshold $D=30$. The discontinuity is
 10 of magnitude 40 and this exceeds D , implying that the pixel $x(r, c)$ does not belong to the Deblock region DEB . Consider how this same discontinuity of magnitude 40 is classified if it is a spread-out discontinuity from say $x(r, c) = 100$ to $x(r, c + 1) = 120$ to $x(r, c + 2) = 140$. In this case, the discontinuities at (r, c) and $x(r, c + 1)$ are each of magnitude 20 and because they fail to exceed the value of D , this causes false Deblocking to occur: that is,
 15 both $x(r, c)$ and $x(r, c + 1)$ would be wrongly assigned to the Deblock region DEB .

Similar spread-out edge discontinuities may exist in the vertical direction.

Most commonly, such spread-out discontinuities straddle 2 pixels although the straddling of 3 pixels is also found in some heavily-compressed video signals.

One embodiment of this method for correctly classifying spread-out edge-
 20 discontinuities is to employ a dilated version of the above 9-pixel crossed-mask 52 which may be used to identify and thereby Deblock spread-out discontinuity boundaries. For example, all of the Candidate Regions identified in the 9-pixel crossed-mask 52 of FIGURE 5 are 1 pixel in size but there is no reason why the entire crossed-mask could not be spatially-dilated (*i.e.* stretched), employing similar logic. Thus, ax , bx , ...etc. are
 25 spaced 2 pixels apart, and surround a central region of 2x2 pixels. The above Combined Pixel-Level Deblock Condition remains in effect and is designed such that $C_i \in Flat$ under at least one of the following three conditions:

- d) Crossed-mask 52 (M) lies over a 20-pixel region that is entirely of sufficiently-flat intensity, therefore including sufficiently-flat regions where M lies entirely
 30 in the interior of a block

OR

- e) Crossed-mask 52 lies over a 2-pixel wide discontinuity at one of the four 1x2 pixel locations

$(r + 2 : r + 3, c)$ OR $(r + 4 : r + 5, c)$ OR $(r - 2 : r - 1, c)$ OR $(r - 4 : r -$

5 3, c)

while satisfying the flatness criteria at the remaining three locations

OR

f) Crossed-mask 52 lies over a 2-pixel wide discontinuity at one of the four
2x1 pixel locations

10 $(r, c + 2 : c + 3)$ OR $(r, c + 4 : c + 5)$ OR $(r, c - 2 : c - 1)$ OR $(r, c -$
 $4 : c - 3)$

while satisfying the flatness criteria at the remaining three locations.

In this way, as required, the crossed-mask M is capable of covering the 1-
pixel-wide boundaries as well as the spread-out 2-pixel-wide boundaries of blocks,
15 regardless of their locations, while maintaining the truth of the statement $C_i \in Flat$. The
minimum number of computations required for the 20-pixel crossed-mask is the same as
for the 9-pixel version.

There are many variations in the details by which the above flatness and
discontinuity criteria may be determined. For example, criteria for 'flatness' could involve
20 such statistical measures as variance, mean and standard deviation as well as the removal
of outlier values, typically at additional computational cost and slower throughput.
Similarly, qualifying discontinuities could involve fractional changes of intensity, rather
than absolute changes, and crossed-masks M can be dilated to allow the discontinuities to
spread over several pixels in both directions.

25 A particular variation of the above criteria relates to fractional changes of
intensity rather than absolute changes. This is important because it is well known that the
HVS responds in an approximately linear way to fractional changes of intensity. There are
a number of modifications of the above method for adapting to fractional changes and
thereby improving the perception of Deblocking, especially in dark regions of the image
30 frame. They include:

i. Instead of subjecting the image intensity $x(r,c)$ directly to the flatness and
discontinuity criteria as the Candidate Pixel C_i , the logarithm of intensity $C_i = \log_b$
($x(r,c)$) is used throughout, where the base b might be 10 or the natural exponent
 $e = 2.718...$

35 OR

- 5 ii. Instead of employing magnitudes of intensity differences directly, fractional differences are used directly as all or part of the criteria for flatness, discontinuities, look ahead and look back. For example, the flatness criteria may be modified from the absolute intensity threshold e in

$$|x(r+1, c) - x(r, c)| < e$$

- 10 to a threshold containing a relative intensity term, such as a relative threshold e_R of the form

$$e_R \equiv \left(e + \frac{x(r, c)}{I_{MAX}} \right)$$

where, in the example in the Appendix, we have used $e = 3$ and $I_{MAX} = 255$ which is the maximum intensity that can be assumed by $x(r, c)$.

- 15 The Candidate Regions C_i must sample the 2D space of the image frame sufficiently-densely that the boundaries of most of the block artifacts are not missed due to under-sampling. Given that block-based compression algorithms ensure that most boundaries of most blocks are separated by at least 4 pixels in both directions, it is possible with this method to sub-sample the image space at intervals of 4 pixels in each direction
- 20 without missing almost all block boundary discontinuities. Up to 8 pixels in each direction has also been found to work well in practice. This significantly reduces computational overhead. For example sub-sampling by 4 in each direction leads to a disconnected set of points that belong to the Deblock Region. An embodiment of this method employs such sub-sampling.

- 25 Suppose the Candidate Pixels are L pixels apart in both directions. Then the Deblock region may be defined, from the sparsely-distributed Candidate Pixels, as that region obtained by surrounding all Candidate Pixels by $L \times L$ squares blocks. This is easy to implement with an efficient algorithm.

- 30 Once the Deblock Regions are identified, there is a wide variety of Deblocking strategies that can be applied to the Deblock region in order to attenuate the visibly-objectionable perception of blockiness. One method is to apply a smoothing operation to the Deblock Region, for example by using Spatially-Invariant Low Pass IIR Filters or Spatially-Invariant Low Pass FIR Filters or FFT-based Low Pass Filters.

5 An embodiment of this method down samples the original image frames prior to the smoothing operation, followed by up sampling to the original resolution after smoothing. This embodiment achieves faster overall smoothing because the smoothing operation takes place over a smaller number of pixels.

 With the exception of certain filters such as the Recursive Moving Average
10 (*i.e.* the Box) 2D filter, 2D FIR filters have computational complexity that increases with the level of smoothing that they are required to perform. Such FIR smoothing filters require a number of MACs/s that is approximately proportional to the level of smoothing.

 Highly-compressed videos (*e.g.* having a quantization parameter $q > 40$) typically require FIR filters of order greater than 11 to achieve sufficient smoothing
15 effects, corresponding to at least 11 additions and up to 10 multiplications per pixel. A similar level of smoothing can be achieved with much lower order IIR filters, typically of order 2. One embodiment of this method employs IIR filters for smoothing the Deblock Region.

 Another method for smoothing is similar to that described above except that
20 the smoothing filters are spatially-varied (*i.e.*, spatially-adapted) in such a way that the crossed-mask of the filters is altered, as a function of spatial location, so as not to overlap the Detail Region. In this method, the order (and therefore the crossed-mask size) of the filter is adaptively reduced as it approaches the boundary of the Detail Region.

 The crossed-mask size may also be adapted on the basis of local statistics to
25 achieve a required level of smoothing, albeit at increased computational cost. This method employs spatially-variant levels of smoothing in such a way that the response of the filters cannot overwrite (and thereby distort) the Detail region or penetrate across small Detail Regions to produce an undesirable 'halo' effect around the edges of the Detail Region.

 A further improvement of this method applies a 'growing' process to the Detail
30 region *DET* in a) above for all Key Frames such that *DET* is expanded around its boundaries. The method used for growing, to expand the boundaries, such as that described herein may be used, or other methods known to one of ordinary skill in the art. The resultant Expanded Detail region *EXPDET* is used in this further improvement as the Detail region for the adjacent image frames where it overwrites the Canvas Images *CAN* of

5 those frames. This increases throughput and reduces computational complexity because it is only necessary to identify the Detail region *DET* (and its expansion *EXPDET*) in the Key Frames. The advantage of using *EXPDET* instead of *DET* is that *EXPDET* more effectively covers moving objects having high speeds than can be covered by *DET*. This allows the Key Frames to be spaced farther apart, for a given video signal, and thereby
 10 improves throughput and reduces complexity.

In this method, the Detailed region *DET* may be expanded at its boundaries to spatially cover and thereby make invisible any 'halo' effect that is produced by the smoothing operation used to Deblock the Deblock region.

In an embodiment of this method, a spatially-variant 2D Recursive Moving
 15 Average Filter (*i.e.* a so-called 2D Box Filter) is employed, having the 2D Z transform transfer functions

$$H(z_1, z_2) = \frac{(1 - z_1^{-L_1})(1 - z_2^{-L_2})}{(1 - z_1^{-1})(1 - z_2^{-1})} \frac{1}{L_1 L_2}$$

which facilitates fast recursive 2D FIR filtering of 2D order (L_1, L_2) . The corresponding 2D recursive FIR input-output difference equation is

20
$$y(r, c) = y(r-1, c) + y(r, c-1) - y(r-1, c-1) + \dots$$

$$\frac{1}{L_1 L_2} [x(r, c) + x(r-L_1, c) + x(r, c-L_2) + x(r-L_1, c-L_2)]$$

where y is the output and x is the input. This embodiment has the advantage that the arithmetic complexity is low and is independent of the level of smoothing.

In a specific example of the method, the order parameters (L_1, L_2) are spatially-
 25 varied (*i.e.*, spatiality of the above 2D FIR Moving Average filter is adapted to avoid overlap of the response of the smoothing filters with the Detail region *DET*).

FIGURE 7 shows one embodiment of a method, such as method 70, for achieving improved video image quality using the concepts discussed herein. One system for practicing this method can be, for example, by software, firmware, or an ASIC running
 30 in system 80 shown in FIGURE 8, perhaps under control of processor 82-1 and/or 84-1. Process 701 determines a Deblock region. When all Deblock regions are found, as

- 5 determined by process 702 process 703 then can identify all Deblock regions and by
implication all Detail regions.

Process 704 then can begin smoothing such that process 705 determines when
the boundary of the Nth Deblock region has been reached and process 706 determines
when smoothing of the Nth region has been completed. Process 708 indexes the regions
10 by adding 1 to the value N and processes 704 through 707 continue until process 707
determines that all Deblock regions have been smoothed. Then process 709 combines the
smoothed Deblock regions to the respective Detail regions to arrive at an improved image
frame. Note that it is not necessary to wait until all of the Deblock regions are smoothed
before beginning the combining process since these operations can be performed in
15 parallel if desired.

FIGURE 8 shows one embodiment 80 of the use of the concepts discussed
herein. In system 80 video (and audio is provided as an input 81. This can come from
local storage, not shown, or received from a video data stream(s) from another location.
This video can arrive in many forms, such as through a live broadcast stream, or video file
20 and may be pre-compressed prior to being received by encoder 82. Encoder 82, using the
processes discussed herein processes the video frames under control of processor 82-1.
The output of encoder 82 could be to a file storage device (not shown) or delivered as a
video stream, perhaps via network 83, to a decoder, such as decoder 84.

If more than one video stream is delivered to decoder 84 then the various
25 channels of the digital stream can be selected by tuner 84-2 for decoding according to the
processes discussed herein. Processor 84-1 controls the decoding and the output decode
video stream can be stored in storage 85 or displayed by one or more displays 86 or, if
desired, distributed (not shown) to other locations. Note that the various video channels
can be sent from a single location, such as from encoder 82, or from different locations,
30 not shown. Transmission from the decoder to the encoder can be performed in any well-
known manner using wireline or wireless transmission while conserving bandwidth on the
transmission medium.

Although the present invention and its advantages have been described in
detail, it should be understood that various changes, substitutions and alterations can be

- 5 made herein without departing from the spirit and scope of the invention as defined by the
appended claims. Moreover, the scope of the present application is not intended to be
limited to the particular embodiments of the process, machine, manufacture, composition
of matter, means, methods and steps described in the specification. As one of ordinary
skill in the art will readily appreciate from the disclosure of the present invention,
- 10 processes, machines, manufacture, compositions of matter, means, methods, or steps,
presently existing or later to be developed that perform substantially the same function or
achieve substantially the same result as the corresponding embodiments described herein
may be utilized according to the present invention. Accordingly, the appended claims are
intended to include within their scope such processes, machines, manufacture,
- 15 compositions of matter, means, methods, or steps.

5

CLAIMS

What is claimed is:

1. A method for removing artifacts from an image frame, said artifacts being visually disruptive to the HVS, said method comprising:
separating a digital representation of each image frame into a Deblock region that
10 is to be deblocked and a Detail region that is to remain essentially undeblocked.
2. The method of claim 1 further comprising:
smoothing said Deblock region of each said image frame; and
combining said smoothed Deblock region with said unblocked Detail region to
form an a new image frame having less visual disruption to the HVS than a pre-separated
15 image frame.
3. The method of claim 2 wherein said separating comprises at least one of the following criteria for determining said Deblock Region: intensity-flatness; discontinuity; look-head; look-behind.
4. The method of claim 3 wherein parameters of said criteria are chosen such
20 that attenuation occurs for compressed image frames in which locations of artifact blocks are *a priori* unknown.
5. The method of claim 4 wherein said artifact blocks occur in said compressed video frames due to one or more of the following: previously compressed multiple times; re-formatted image frames, color-mixed image frames; re-sized image
25 frames.
6. The method of claim 3 wherein said intensity-flatness criteria employs statistical measures comprising a local variance and a local mean of intensities.
7. The method of claim 3 wherein intensity change criteria are based on fractional changes of intensity.
- 30 8. The method of claim 2 wherein said smoothing comprises:
spatial smoothing to attenuate said Deblock Region.

- 5 9. The method of claim 2 wherein said smoothing comprises attenuation of blocks as well as other artifacts in said Deblock Region.
10. The method of claim 1 where said separating occurs within a DCT-based encoder.
11. The method of claim 2 wherein said smoothing comprises at least one of:
10 FIR filters, IIR filters.
12. The method of claim 11 wherein said filters can be either spatially-variant or spatially invariant.
13. The method of claim 11 wherein said smoothing comprises:
 at least one Moving Average FIR 2D Box filter.
- 15 14. The method of claim 2 wherein said smoothing comprises:
 means for ensuring that smoothing does not occur outside of the boundaries of said Deblock Region.
15. The method of claim 1 wherein said separating recursively separates said image frame into Deblock Regions and Detail Regions.
- 20 16. The method of claim 1 wherein said separating comprises:
 selecting candidate regions; and
 determining on a selected candidate by selected candidate region basis whether a selected candidate region belongs to said Deblock region according to certain criteria.
17. The method of claim 16 wherein said candidate regions are sparsely located
25 in each image frame.
18. The method of claim 17 wherein a separated Detail region is expanded to allow spatially invariant filtering of said Deblock region without causing a halo effect around said Detail Region.
19. The method of claim 18 wherein said expansion comprises:
30 growing each candidate pixel to a surrounding rectangle of pixels.

- 5 20. The method of claim 1 wherein a separated Detail region is expanded to allow spatially invariant filtering of said Deblock region without causing a halo effect around said Detail Region.
21. The method of claim 2 wherein said smoothing comprises:
 use of an N-pixel crossed-mask.
- 10 22. The method of claim 21 where N equals 9.
23. The method of claim 2 wherein said smoothing comprises:
 using dilated crossed-masks for Deblocking video signals having spread out edge discontinuities.
24. A system for presenting video, said system comprising:
15 an input for obtaining a first video frame having a certain number of bits per pixel; said certain number being such that when said video frame is presented to a display said display yields artifacts perceptible to a human visual system (HVS); and
 circuitry for producing a second video frame from said first video frame, said second video frame yielding artifacts less perceptible to said HVS when said second video
20 frame is presented to said display.
25. The system of claim 24 wherein said certain number extends to a low of 0.1 bits/pixel.
26. The system of claim 24 wherein said certain number is a number of bits/pixel provided by compression of said first video frame using an H.264 encoder.
- 25 27. The system of claim 25 wherein said certain number is at least $\frac{1}{2}$ the number of bits achieved by an H.264 encoder.
28. The system of claim 24 wherein said producing circuitry comprises:
 means for separating said video frame into a Detail region and a Deblock region;
 and
30 means for smoothing said Deblock region prior to combining said regions to form said second video frame.

5 29. The system of claim 28 further comprising:
a tuner for allowing a user to select one of a plurality of digital video streams, each
said video stream comprising a plurality of digital video frames.

 30. The system of claim 28 wherein said smoothing comprises:
a spatially invariant FIR filter having a certain crossed-mask size; and
10 a processor for preventing said spatially invariant filter from smoothing said Detail
regions.

 31. The system of claim 30 wherein said processor operates to expand said
Detail regions a distance approximately equal to $\frac{1}{2}$ said crossed-mask size.

 32. The system of claim 28 wherein said smoothing means comprises:
15 a spatially variant FIR filter.

 33. The system of claim 28 wherein said separating means comprises:
processing using at least one of the following criteria for determining said Deblock
Region: intensity-flatness; discontinuity; look-head; look-behind.

 34. The system of claim 33 wherein parameters of said criteria are chosen such
20 that artifact attenuation occurs for compressed image frames in which locations of artifact
blocks are *a priori* unknown.

 35. The system of claim 34 wherein said artifact blocks occur in said
compressed video frames due to one or more of the following: previously compressed
multiple times; re-formatted image frames, color-mixed image frames; re-sized image
25 frames.

 36. The system of claim 33 wherein said intensity-flatness criteria employs
statistical measures comprising a local variance and a local mean of intensities.

 37. The system of claim 33 wherein intensity change criteria are based on
fractional changes of intensity.

30 38. The system of claim 28 wherein said smoothing means comprises:
a processor operative for spatial smoothing to attenuate said Deblock Region.

- 5 39. The system of claim 28 wherein said smoothing means comprises:
a processor for attenuating blocks as well as other artifacts in said Deblock Region.
40. The system of claim 28 where said means for separating is a portion of a
DCT-based encoder.
41. The system of claim 28 wherein said smoothing means comprises at least
10 one of: FIR filters, IIR filters.
42. The system of claim 41 wherein said filters can be either spatially-variant or
spatially invariant.
43. The system of claim 28 wherein said smoothing means comprises:
at least one Moving Average FIR 2D Box filter.
- 15 44. The system of claim 28 wherein said separating means recursively separates
said image frame into Deblock Regions and Detail Regions.
45. The system of claim 28 wherein said separating means comprises:
means for selecting candidate regions; and
means for determining on a selected candidate by selected candidate region basis
20 whether a selected candidate region belongs to said Deblock region according to certain
criteria.
46. The system of claim 45 wherein said candidate regions are sparsely located
in each image frame.
47. A method of presenting video, said method comprising:
25 obtaining a first video frame having a certain number of bits per pixel; said certain
number being such that when said video frame is presented to a display said display yields
artifacts perceptible to a human visual system (HVS); and
producing a second video frame from said first video frame, said second video
frame yielding artifacts less perceptible to said HVS when said second video frame is
30 presented to said display.
48. The method of claim 47 wherein said certain number extends to a low of
0.1 bits/pixel.

5 49. The method of claim 47 wherein said producing comprises:
separating Detail and Deblock regions within each said frame; and
smoothing said Deblock region; and
combining said smoothed Deblock region with said separated Detail region.

10 50. The method of claim 49 wherein said smoothing comprises:
using a spatially invariant FIR filter having a certain crossed-mask size; and
expanding said Detail region a distance at least equal to $\frac{1}{2}$ said crossed-mask size
so as to avoid a halo effect at a border between said Deblock and Detail regions.

15 51. The method of claim 50 further comprising:
receiving at a device a plurality of digital video streams, each said stream having a
plurality of said digital video frames; and
wherein said obtaining comprises:
selecting one of said received digital video streams at said device.