Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2684110 A1 2008/10/23

(21) 2 684 110

(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION

(13) A1

(86) Date de dépôt PCT/PCT Filing Date: 2008/04/09
(87) Date publication PCT/PCT Publication Date: 2008/10/23
(85) Entrée phase nationale/National Entry: 2009/10/09
(86) N° demande PCT/PCT Application No.: IB 2008/051351
(87) N° publication PCT/PCT Publication No.: 2008/126019
(30) Priorité/Priority: 2007/04/12 (US60/911,480)

(51) Cl.Int./Int.Cl. *H04N 7/50* (2006.01)
(71) Demandeur/Applicant:
NOKIA CORPORATION, FI

(72) Inventeur/Inventor:
LAINEMA, JANI, FI

(74) Agent: SIM & MCBURNEY

(54) Titre : CODAGE PAR TRANSFORMATION AMELIORE DANS L'ESPACE
(54) Title: SPATIALLY ENHANCED TRANSFORM CODING

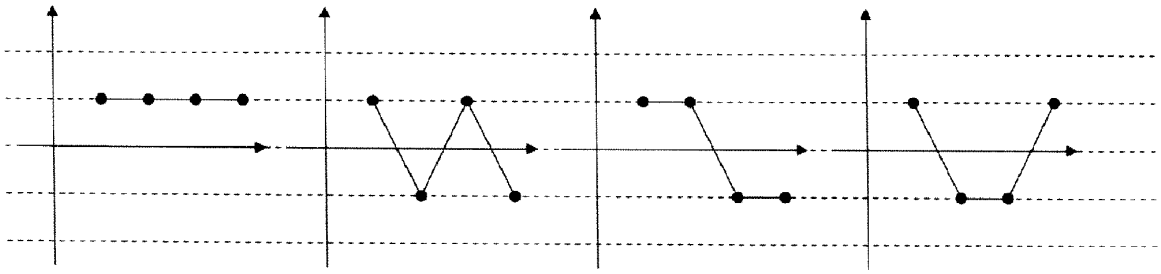

FIG. 1

(57) Abrégé/Abstract:
A system and method for improving prediction error coding performance of various video compression algorithms. Various embodiments combine the energy compaction features of transform coding with localization properties of spatial coding. In various embodiments, the effects of pixel "outliers" are removed from the transform and are coded separately as a spatially coded pixel prediction, thereby improving the coding gain of the transform.

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000001
IPR2024-00921 (Amazon.com, Inc., Amazon.com Services LLC v. Nokia Technologies Oy)

SPATIALLY ENHANCED TRANSFORM CODING

FIELD OF THE INVENTION

[0001] The present invention relates to the coding and decoding of digital video material. More particularly, the present invention relates to prediction error coding in both frequency and spatial domains.

BACKGROUND OF THE INVENTION

[0002] This section is intended to provide a background or context to the invention that is recited in the claims. The description herein may include concepts that could be pursued, but are not necessarily ones that have been previously conceived or pursued. Therefore, unless otherwise indicated herein, what is described in this section is not prior art to the description and claims in this application and is not admitted to be prior art by inclusion in this section.

[0003] A video codec comprises an encoder that transforms input video into a compressed representation suited for storage and/or transmission and a decoder that can uncompress the compressed video representation back into a viewable form. Typically, the encoder discards some information in the original video sequence in order to represent the video in a more compact form, i.e., at a lower bitrate.

[0004] Typical hybrid video codecs, for example ITU-T H.263 and H.264, encode video information in two phases. In the first phase, pixel values in a certain picture area or "block" are predicted. These pixel values can be predicted, for example, by motion compensation mechanisms, which involve finding and indicating an area in one of the previously coded video frames that corresponds closely to the block being coded. Additionally, pixel values can be predicted via by spatial mechanisms, which involve using the pixel values around the block to be coded in a specified manner. The second phase involves coding the prediction error, i.e. the difference between the predicted block of pixels and the original block of pixels. This is typically accomplished by transforming the difference in pixel values using a specified transform (e.g., a Discreet Cosine Transform (DCT) or a variant thereof), quantizing

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000002

the coefficients, and entropy coding the quantized coefficients. By varying the fidelity of the quantization process, the encoder can control the balance between the accuracy of the pixel representation (i.e., the picture quality) and the size of the resulting coded video representation (i.e., the file size or transmission bitrate).

[0005] The decoder reconstructs output video by applying prediction mechanisms that are similar to those used by the encoder in order to form a predicted representation of the pixel blocks (using motion or spatial information created by the encoder and stored in the compressed representation) and prediction error decoding (the inverse operation of the prediction error coding, recovering the quantized prediction error signal in the spatial pixel domain). After applying prediction and prediction error decoding processes, the decoder sums up the prediction and prediction error signals (i.e., the pixel values) to form the output video frame. The decoder (and encoder) can also apply additional filtering processes in order to improve the quality of the output video before passing it for display and/or storing it as a prediction reference for the forthcoming frames in the video sequence.

[0006] In typical video codecs, the motion information is indicated with motion vectors associated with each motion-compensated image block. Each of these motion vectors represents the displacement of the image block in the picture to be coded (in the encoder side) or decoded (in the decoder side) and the prediction source block in one of the previously coded or decoded pictures. In order to represent motion vectors efficiently, motion vectors are typically coded differentially with respect to block specific predicted motion vectors. In a typical video codec, the predicted motion vectors are created in a predefined way, for example by calculating the median of the encoded or decoded motion vectors of the adjacent blocks.

[0007] Typical video encoders utilize Lagrangian cost functions to find optimal coding modes, e.g., the desired macroblock mode and associated motion vectors. This kind of cost function uses a weighting factor λ to tie together the exact or estimated image distortion due to lossy coding methods and the exact or estimated amount of information that is required to represent the pixel values in an image area:

$$C = D + \lambda R \qquad (1)$$

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000003

[0008] In Eq. (1), C is the Lagrangian cost to be minimized, D is the image distortion (e.g., the mean squared error) with the mode and motion vectors considered, and R the number of bits needed to represent the required data to reconstruct the image block in the decoder (including the amount of data to represent the candidate motion vectors).

[0009] Transform coding of the prediction error signal in video or image compression system typically comprises DCT-based linear transform, quantization of the transformed DCT coefficients, and context based entropy coding of the quantized coefficients. However, the transform can efficiently pack energy of the prediction error signal only under certain statistics, and the coding performance deteriorates when the prediction error to be transformed becomes less correlated. This causes suboptimal performance, especially in modern video and image coding systems employing advanced motion compensation and spatial prediction processes in order to achieve good quality predictions for the image blocks to be coded (thus, minimizing and decorrelating the prediction error signal).

[0010] To address some of the above issues, a number of hybrid video coding schemes have been developed. These hybrid systems typically comprise a hybrid of two redundancy reduction techniques--prediction and transformation. Prediction can take the form of inter-picture prediction, which is used to remove temporal redundancies in the signal. Intra-picture prediction may also be used in the H.264/Advanced Video Coding (AVC) standard where spatial redundancies are removed by exploiting the similarities between neighboring regions within a picture frame. As a consequence of these inter-picture and intra-picture prediction techniques, a residual/error signal is formed by removing the predicated picture frame from the original. This prediction error signal is then typically block transform coded using an 8x8 DCT transform in order to reduce spatial redundancies in the signal.

SUMMARY OF THE INVENTION

[0011] Various embodiments of the present invention provide a system and method for representing the prediction error signal as a weighted sum of different basis functions of a selected transform and quantized spatial samples. The basis functions

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000004

of the selected transform may comprise an orthogonal set of basis vectors, or the basis functions may not comprise an orthogonal set. According to various embodiments, the prediction error signal for a single image block is constructed using both transform basis functions and spatial samples (i.e., pixel values), thereby combining the desired features of both the transform and spatial coding approaches discussed previously. This allows for the utilization of those selected transform basis functions that give good overall representation of the image block with minimal amount of transform coefficients (representing the component of prediction error signal that is well correlated with the basis functions). Additionally, various embodiments of the present invention allow for the efficient spatial representation of those components of the prediction error signal of the same image block that are not well correlated with the basis functions of the applied transform (such as certain types of sensor noise, high frequency texture and edge information).

[0012] According to various embodiments of the present invention, a system and method of encoding a prediction error signal for a block of data comprises calculating a difference signal representing differences between sample values of a predicted block of data and values for an original input block. Transform coding and spatial coding are both performed to the difference signal, thereby creating a first and second representations of first and second components of the difference signal. The first and second representations are then combined to provide the prediction error signal.

[0013] Various embodiments of the present invention also provide a system and method of decoding a prediction error signal for a block of data, comprising receiving a coded prediction error signal, the coded prediction error signal including a plurality of transform coefficients and a plurality of spatial samples. The plurality of transformed coefficients are decoded into decoded transform information, and the plurality of spatial samples are decoded into decoded spatial information. The decoded transform information, the decoded spatial information, and a reconstructed prediction of the block of data are then added, thereby forming a decoded representation of the block of data.

[0014] The implementations of various embodiments of the present invention serve to improve the compression efficiency of modern video and image codecs. Although

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000005

a certain amount of increased computational complexity of encoding may be needed, fast algorithms can be applied in order to lower the encoding complexity that approaches the complexity level for traditional transform based coding. Any effect on the complexity for a decoder is negligible when implementing various embodiments of the present invention.

[0015] These and other advantages and features of the invention, together with the organization and manner of operation thereof, will become apparent from the following detailed description when taken in conjunction with the accompanying drawings, wherein like elements have like numerals throughout the several drawings described below.

BRIEF DESCRIPTION OF THE DRAWINGS

[0016] FIG. 1 shows an example of basis functions of a Hadamard transform that can be used to encode image data;

[0017] Figure 2(a) is an example of four scalar values that are to be coded or decoded and representing the prediction error signal between the original image and the prediction image; Figure 2(b) shows the four weighted basis functions from Figure 1 being summed to the prediction error signal; and Figure 2(c) shows how the same prediction error signal can be reconstructed by summing up only the first basis function and a single spatial sample represented by an arrow;

[0018] Figure 3 is a block diagram of a video encoder constructed in accordance with one embodiment of the present invention;

[0019] Figure 4 is a flow chart showing one encoder algorithm which may be used in an embodiment of the present invention;

[0020] Figure 5 is a flow chart showing one decoder algorithm which may be used in various embodiments of the present invention;

[0021] Figure 6 is a block diagram of a video decoder constructed in accordance with one embodiment of the present invention;

[0022] Figure 7 is an overview diagram of a system within which various embodiments of the present invention may be implemented;

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000006

[0023] Figure 8 is a perspective view of an electronic device that can be used in conjunction with the implementation of various embodiments of the present invention; and

[0024] Figure 9 is a schematic representation of the circuitry which may be included in the electronic device of Figure 8.

DETAILED DESCRIPTION OF VARIOUS EMBODIMENTS

[0025] Various embodiments of the present invention provide a system and method for representing the prediction error signal as a weighted sum of different basis functions of a selected transform and quantized spatial samples. According to various embodiments, the prediction error signal for a single image block is constructed using both transform basis functions and spatial samples (i.e., pixel values), thereby combining the desired features of both the transform and spatial coding approaches discussed previously. This allows for the utilization of those selected transform basis functions that give good overall representation of the image block with minimal amount of transform coefficients (representing the component of prediction error signal that is well correlated with the basis functions). Additionally, various embodiments of the present invention allow for the efficient spatial representation of those components of the prediction error signal of the same image block that are not well correlated with the basis functions of the applied transform (such as certain types of sensor noise, high frequency texture and edge information).

[0026] Figures 1 and 2 show a simplified example of how various embodiments of the present invention are operable in one dimension. Figure 1 shows an example of basis functions of a one-dimensional linear transform that can be used to encode image data (namely a Hadamard transform). The weighted sum of these four basis functions can be used to represent any four samples. Figure 2(a) is an example of four scalar values that are to be coded or decoded and representing the prediction error signal between the original image and the prediction image. In Figure 2(b), the four weighted basis functions from Figure 1 are summed to the prediction error signal (1.5 times the first, 0.5 times the second, -0.5 times the third and -0.5 times the fourth basis function). In Figure 2(c) it is demonstrated how the same prediction error signal

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000007

can be reconstructed by summing up only the first basis function and a single spatial sample represented by an arrow. In this example, compression efficiency improvement can be expected, as the signal is represented by a single transform coefficient and a single spatial sample instead of four transform coefficients weighting the four Hadamard basis functions (as shown in Figure 2(b)).

[0027] Figure 3 is a block diagram of a video encoder constructed in accordance with one embodiment of the present invention. More particularly, Figure 3 shows how an image to be encoded 300 undergoes pixel prediction 302, prediction error coding 303 and prediction error decoding 304. For pixel prediction 302, the image 300 undergoes both inter-prediction 306 and intra-prediction 308 which, after mode selection 310, results in prediction representation of an image block 312. A preliminary reconstructed image 314 is also used for intra-prediction 308. Once all of the image blocks are processed, the preliminary reconstructed image 314 undergoes filtering at 316 to create a final reconstructed image 340, which is sent to a reference frame memory 318 and is also used for inter-prediction 306 of future frames.

[0028] The prediction representation of the image block 312, as well as the image to be encoded 300, are used together to define a prediction error signal 320 which is used for prediction error coding 303. In prediction error coding 303, the prediction error signal 320 undergoes feature selection 322 and spatial quantization 324, as well as transform 326 and quantization 328 (both after feature selection 322). The data describing prediction error and predicted representation of the image block 312 (e.g., motion vectors, mode information and quantized DCT+spatial samples) are passed to entropy coding 330. The prediction error decoding 304 is substantially the opposite of the prediction error coding 303, with the prediction error decoding including an inverse spatial quantizer 332, an inverse transform 334 and an inverse quantization 336. The result of the prediction error decoding 304 is a reconstructed prediction error signal 338, which is used in combination with the predicted representation of the image block 312 to create the preliminary reconstructed image 314.

[0029] Figure 4 is a flow chart showing one encoder algorithm which may be used in an embodiment of the present invention. However, as discussed below, a variety of different algorithms may be used according to principles of the present invention. At

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000008

400 in Figure 4, the difference between sample values of a predicted block and an original input block is calculated. It should be noted that "values" and "outlier values" as used herein are intended to refer to these difference values. At 410, outlier values are found. In this case, the amplitude depends on the expected accuracy of the coded representation; smaller amplitude representations qualify as outliers if one is targeting high bitrate good quality representation, and only high amplitude representations qualify if one is targeting a lower bitrate, lower quality representation. At 420, the modified prediction error signal is transform coded (involving transform, quantization and entropy coding), with the outlier values being substituted by interpolated representations of those values, e.g., by averaging the neighboring prediction error values. At 430, the modified prediction error signals are transform decoded back to the spatial domain. At 440, spatial coding is applied (involving quantization and entropy coding without transform) to the difference between the original outlier values and the transform decoded outlier values. At 450, the final coded representation of the prediction error signal is formed by joining the transform coded data and the spatial coded data. In an alternative embodiment, the spatial coding occurs before transform coding.

[0030] When considering the one-dimensional example depicted in Figures 2(a)-2(c), the algorithm outlined in Figure 4 would process the relevant data as follows. As discussed previously, Figure 2(a) represents the difference signal [1 1 3 1] to be coded. In this case, [3] is an outlier. Therefore, the prediction error signal is modified by substituting the [3] with an average of the second and fourth samples, making the signal [1 1 1 1]. The difference between the original outlier sample ([3]) and the transform coded outlier value ([1]) is then spatial coded, making the spatial coded signal [0 0 2 0]. The entropy coded representations of the transform coded and spatial coded signals are then written to the bitstream. Thereafter, the decoder can recover the coded signal by adding up the inverse transformed data [1 1 1 1] and the inverse spatial coded data [0 0 2 0] (together with the prediction samples of the image block).

[0031] When considering the above example, it is important to note that the choice of Hadamard basis vectors for the transform is only intended to be exemplary in nature. In fact, the various methods and techniques described herein can be applied to

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000009

any transform employing basis functions, and the basis functions do not necessarily need to be orthogonal.

[0032] A video or image codec, according to various embodiments the present invention, can be implemented in various ways. In the case of an encoder, the encoder can use different strategies to find the desired transform coefficients and spatial samples. For example, an encoder can first transform code the signal and apply spatial coding to the residual. The encoder can also first apply spatial coding to the signal, followed by transform coding of the residual. Additionally, some or all of the transform coefficients can be set to zero in order to improve the coding performance. Some or all spatial samples can also be set to zero in order to improve the coding performance. An encoder can also iteratively modify transform coefficients and/or spatial samples in order to improve the coding performance until a desired performance or a maximum defined number of iterations is achieved.

[0033] In terms of quantization and dequantization, the quantization and dequantization of transform coefficients and spatial samples can be tied together (e.g., a quantization step size for both transform coefficients and spatial samples can be derived from a single parameter). Alternatively, different quantizations and dequantizations can be applied to transform coefficients and spatial samples.

[0034] For the coding and decoding of spatial samples, such coding and decoding can depend on transform coefficients and vice versa. Alternatively, the coding and decoding of spatial samples, transform coefficients or both can depend on the prediction signal, on the other transform coefficients and spatial samples in the same image or in other images.

[0035] In addition to the above, it can be indicated that there are no transform coefficients or spatial samples for a specific image area or image. It can also be indicated that only spatial coding or only transform coding is used for a specific image area or image. The number of spatial samples can be coded and decoded as one unit, for example representing certain patterns of textures. Pre- and/or post-processing mechanisms can be applied to the prediction signal, prediction error signal reconstructed signal or any combination thereof. The method can be used to code and decode other information instead or in addition to the prediction error signals. The

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000010

codec can limit the usage of either transform coefficients (e.g., allow only low frequency transform coefficients to be present in the coded representation of the signal) or spatial samples.

[0036] Figure 5 shows a decoding process according to various embodiments of the present invention. At 500 in Figure 5, a video decoder receives a coded prediction error signal which comprises both transformed coefficients and spatial samples. At 510, both the transformed coefficients and the spatial samples are decoded. At 520, the decoded transform information, the decoded spatial information and reconstructed prediction of the image block are added together to form decoded representation of the image block.

[0037] Figure 6 is a block diagram of a video decoder constructed in accordance with one embodiment of the present invention. As shown in Figure 6, entropy decoding 600 is followed by both prediction error decoding 602 and pixel prediction 604. In prediction error decoding 602, in addition to an inverse transform 606 and inverse quantization 608, an inverse spatial quantizer 610 is used as discussed herein, ultimately resulting in a reconstructed prediction error signal 612. For pixel prediction 604, either intra-prediction or inter-prediction occurs at 614 to create a predicted representation of an image block 616. The predicted representation of the image block 616 is used in conjunction with the reconstructed prediction error signal 612 to create a preliminary reconstructed image 618, which in turn can be used both for prediction 614. Once all of the image blocks have been processed, the preliminary reconstructed image 618 is passed for filtering 620. The filtered image can also be stored in reference frame memory 624, making it usable for prediction 614 as well.

[0038] Figure 7 shows a system 10 in which various embodiments of the present invention can be utilized, comprising multiple communication devices that can communicate through one or more networks. The system 10 may comprise any combination of wired or wireless networks including, but not limited to, a mobile telephone network, a wireless Local Area Network (LAN), a Bluetooth personal area network, an Ethernet LAN, a token ring LAN, a wide area network, the Internet, etc. The system 10 may include both wired and wireless communication devices.

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000011

[0039] For exemplification, the system 10 shown in Figure 7 includes a mobile telephone network 11 and the Internet 28. Connectivity to the Internet 28 may include, but is not limited to, long range wireless connections, short range wireless connections, and various wired connections including, but not limited to, telephone lines, cable lines, power lines, and the like.

[0040] The exemplary communication devices of the system 10 may include, but are not limited to, an electronic device 50, a combination personal digital assistant (PDA0 and mobile telephone 14, a PDA 16, an integrated messaging device (IMD) 18, a desktop computer 20, a notebook computer 22, etc. The communication devices may be stationary or mobile as when carried by an individual who is moving. The communication devices may also be located in a mode of transportation including, but not limited to, an automobile, a truck, a taxi, a bus, a train, a boat, an airplane, a bicycle, a motorcycle, etc. Some or all of the communication devices may send and receive calls and messages and communicate with service providers through a wireless connection 25 to a base station 24. The base station 24 may be connected to a network server 26 that allows communication between the mobile telephone network 11 and the Internet 28. The system 10 may include additional communication devices and communication devices of different types.

[0041] The communication devices may communicate using various transmission technologies including, but not limited to, Code Division Multiple Access (CDMA), Global System for Mobile Communications (GSM), Universal Mobile Telecommunications System (UMTS), Time Division Multiple Access (TDMA), Frequency Division Multiple Access (FDMA), Transmission Control Protocol/Internet Protocol (TCP/IP), Short Messaging Service (SMS), Multimedia Messaging Service (MMS), e-mail, Instant Messaging Service (IMS), Bluetooth, IEEE 802.11, etc. A communication device involved in implementing various embodiments of the present invention may communicate using various media including, but not limited to, radio, infrared, laser, cable connection, and the like.

[0042] Figures 8 and 9 show one representative electronic device 50 within which the present invention may be implemented. It should be understood, however, that the present invention is not intended to be limited to one particular type of device. The

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000012

electronic device 50 of Figures 8 and 9 includes a housing 30, a display 32 in the form of a liquid crystal display, a keypad 34, a microphone 36, an ear-piece 38, a battery 40, an infrared port 42, an antenna 44, a smart card 46 in the form of a UICC according to one embodiment of the invention, a card reader 48, radio interface circuitry 52, codec circuitry 54, a controller 56 and a memory 58. Individual circuits and elements are all of a type well known in the art, for example in the Nokia range of mobile telephones.

[0043] The various embodiments of the present invention described herein is described in the general context of method steps or processes, which may be implemented in one embodiment by a computer program product, embodied in a computer-readable medium, including computer-executable instructions, such as program code, executed by computers in networked environments. Generally, program modules may include routines, programs, objects, components, data structures, etc. that perform particular tasks or implement particular abstract data types. Computer-executable instructions, associated data structures, and program modules represent examples of program code for executing steps of the methods disclosed herein. The particular sequence of such executable instructions or associated data structures represents examples of corresponding acts for implementing the functions described in such steps or processes.

[0044] Software and web implementations of various embodiments of the present invention can be accomplished with standard programming techniques with rule-based logic and other logic to accomplish various database searching steps or processes, correlation steps or processes, comparison steps or processes and decision steps or processes. It should be noted that the words "component" and "module," as used herein and in the following claims, is intended to encompass implementations using one or more lines of software code, and/or hardware implementations, and/or equipment for receiving manual inputs.

[0045] The foregoing description of embodiments of the present invention have been presented for purposes of illustration and description. The foregoing description is not intended to be exhaustive or to limit embodiments of the present invention to the precise form disclosed, and modifications and variations are possible in light of

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000013

the above teachings or may be acquired from practice of various embodiments of the present invention. The embodiments discussed herein were chosen and described in order to explain the principles and the nature of various embodiments of the present invention and its practical application to enable one skilled in the art to utilize the present invention in various embodiments and with various modifications as are suited to the particular use contemplated.

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000014

SPATIALLY ENHANCED TRANSFORM CODING

FIELD OF THE INVENTION

[0001] The present invention relates to the coding and decoding of digital video material. More particularly, the present invention relates to prediction error coding in both frequency and spatial domains.

BACKGROUND OF THE INVENTION

[0002] This section is intended to provide a background or context to the invention that is recited in the claims. The description herein may include concepts that could be pursued, but are not necessarily ones that have been previously conceived or pursued. Therefore, unless otherwise indicated herein, what is described in this section is not prior art to the description and claims in this application and is not admitted to be prior art by inclusion in this section.

[0003] A video codec comprises an encoder that transforms input video into a compressed representation suited for storage and/or transmission and a decoder that can uncompress the compressed video representation back into a viewable form. Typically, the encoder discards some information in the original video sequence in order to represent the video in a more compact form, i.e., at a lower bitrate.

[0004] Typical hybrid video codecs, for example ITU-T H.263 and H.264, encode video information in two phases. In the first phase, pixel values in a certain picture area or "block" are predicted. These pixel values can be predicted, for example, by motion compensation mechanisms, which involve finding and indicating an area in one of the previously coded video frames that corresponds closely to the block being coded. Additionally, pixel values can be predicted via by spatial mechanisms, which involve using the pixel values around the block to be coded in a specified manner. The second phase involves coding the prediction error, i.e. the difference between the predicted block of pixels and the original block of pixels. This is typically accomplished by transforming the difference in pixel values using a specified transform (e.g., a Discreet Cosine Transform (DCT) or a variant thereof), quantizing

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000015

the coefficients, and entropy coding the quantized coefficients. By varying the fidelity of the quantization process, the encoder can control the balance between the accuracy of the pixel representation (i.e., the picture quality) and the size of the resulting coded video representation (i.e., the file size or transmission bitrate).

[0005] The decoder reconstructs output video by applying prediction mechanisms that are similar to those used by the encoder in order to form a predicted representation of the pixel blocks (using motion or spatial information created by the encoder and stored in the compressed representation) and prediction error decoding (the inverse operation of the prediction error coding, recovering the quantized prediction error signal in the spatial pixel domain). After applying prediction and prediction error decoding processes, the decoder sums up the prediction and prediction error signals (i.e., the pixel values) to form the output video frame. The decoder (and encoder) can also apply additional filtering processes in order to improve the quality of the output video before passing it for display and/or storing it as a prediction reference for the forthcoming frames in the video sequence.

[0006] In typical video codecs, the motion information is indicated with motion vectors associated with each motion-compensated image block. Each of these motion vectors represents the displacement of the image block in the picture to be coded (in the encoder side) or decoded (in the decoder side) and the prediction source block in one of the previously coded or decoded pictures. In order to represent motion vectors efficiently, motion vectors are typically coded differentially with respect to block specific predicted motion vectors. In a typical video codec, the predicted motion vectors are created in a predefined way, for example by calculating the median of the encoded or decoded motion vectors of the adjacent blocks.

[0007] Typical video encoders utilize Lagrangian cost functions to find optimal coding modes, e.g., the desired macroblock mode and associated motion vectors. This kind of cost function uses a weighting factor λ to tie together the exact or estimated image distortion due to lossy coding methods and the exact or estimated amount of information that is required to represent the pixel values in an image area:

$$C = D + \lambda R \qquad (1)$$

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000016

[0008] In Eq. (1), C is the Lagrangian cost to be minimized, D is the image distortion (e.g., the mean squared error) with the mode and motion vectors considered, and R the number of bits needed to represent the required data to reconstruct the image block in the decoder (including the amount of data to represent the candidate motion vectors).

[0009] Transform coding of the prediction error signal in video or image compression system typically comprises DCT-based linear transform, quantization of the transformed DCT coefficients, and context based entropy coding of the quantized coefficients. However, the transform can efficiently pack energy of the prediction error signal only under certain statistics, and the coding performance deteriorates when the prediction error to be transformed becomes less correlated. This causes suboptimal performance, especially in modern video and image coding systems employing advanced motion compensation and spatial prediction processes in order to achieve good quality predictions for the image blocks to be coded (thus, minimizing and decorrelating the prediction error signal).

[0010] To address some of the above issues, a number of hybrid video coding schemes have been developed. These hybrid systems typically comprise a hybrid of two redundancy reduction techniques--prediction and transformation. Prediction can take the form of inter-picture prediction, which is used to remove temporal redundancies in the signal. Intra-picture prediction may also be used in the H.264/Advanced Video Coding (AVC) standard where spatial redundancies are removed by exploiting the similarities between neighboring regions within a picture frame. As a consequence of these inter-picture and intra-picture prediction techniques, a residual/error signal is formed by removing the predicated picture frame from the original. This prediction error signal is then typically block transform coded using an 8x8 DCT transform in order to reduce spatial redundancies in the signal.

SUMMARY OF THE INVENTION

[0011] Various embodiments of the present invention provide a system and method for representing the prediction error signal as a weighted sum of different basis functions of a selected transform and quantized spatial samples. The basis functions

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000017

of the selected transform may comprise an orthogonal set of basis vectors, or the basis functions may not comprise an orthogonal set. According to various embodiments, the prediction error signal for a single image block is constructed using both transform basis functions and spatial samples (i.e., pixel values), thereby combining the desired features of both transform and spatial coding approaches discussed previously. This allows for the utilization of those selected transform basis functions that give good overall representation of the image block with minimal amount of transform coefficients (representing the component of prediction error signal that is well correlated with the basis functions). Additionally, various embodiments of the present invention allow for the efficient spatial representation those components of the prediction error signal of the same image block that are not well correlated with the basis functions of the applied transform (such as certain types of sensor noise, high frequency texture and edge information).

[0012] Accordingly, in one aspect there is provided a method of encoding a prediction error signal for a block of data, comprising: calculating a difference signal representing differences between sample values of a predicted block of data and values for an original input block; creating a modified prediction error signal by substituting outlier values in the difference signal with non-outlier values; performing transform coding to the modified prediction error signal, thereby creating a first representation of a first component of the difference signal; performing spatial domain coding to the difference between the difference signal and the transform decoded modified error signal, thereby creating a second representation of the first component of the difference signal; and joining the first and second representations.

[0013] According to another aspect there is provided an apparatus, comprising: a processor; and a memory unit communicatively connected to a controller and comprising: computer code for calculating a difference signal representing differences between sample values of a predicted block of data and values for an original input block; computer code for creating a modified prediction error signal by substituting outlier values in the difference signal with non-outlier values; computer code for performing transform coding to the difference signal, thereby creating a first representation of a first component of the difference signal; computer code for performing spatial domain coding to the difference between the difference signal and the transform decoded modified error signal, thereby creating a second representation of the first component of the difference signal; and computer code for joining the first and second representations.

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000018

[0013a] According to yet another aspect there is provided a method of decoding a prediction error signal for a block of data, comprising: receiving a coded prediction error signal, the coded prediction error signal including a plurality of transform coefficients and a plurality of spatial domain samples; decoding the plurality of transformed coefficients into decoded transform information; decoding the plurality of spatial domain samples into decoded spatial domain information; adding the decoded transform information, the decoded spatial domain information, and a reconstructed prediction of the block of data, thereby forming a decoded representation of the block of data; and providing an indication that at least one of no transform coefficients exist in the block of data and no spatial domain samples exist in the block of data.

[0013b] According still yet another aspect there is provided an apparatus, comprising: a processor; and a memory unit communicatively connected to the processor and comprising: computer code for processing a received coded prediction error signal, the coded prediction error signal including a plurality of transformed coefficients and a plurality of spatial domain samples; computer code for decoding the plurality of transformed coefficients into decoded transform information; computer code for decoding the plurality of spatial domain samples into decoded spatial domain information; computer code for adding the decoded transform information, the decoded spatial domain information, and a reconstructed prediction of the block of data, thereby forming a decoded representation of the block of data; and computer code for providing an indication that at least one of no transform coefficients exist in the block of data and no spatial domain samples exist in the block of data.

[0013c] According to still yet another aspect there is provided a method of encoding a prediction error signal for a block of data, comprising: calculating a difference signal representing differences between sample values of a predicted block of data and values for an original input block; substituting outlier values in the difference signal with non-outlier values to create a modified prediction error signal; performing one of transform coding and spatial domain coding to the modified prediction error signal, creating one of a transform coded data and spatial domain coded data; performing an inverse of the one of transform coding and spatial domain coding; applying the other of transform coding and spatial domain coding to a residual from the inverse transform or spatial domain decoding, creating the other of transform coded data and spatial domain coded data; and joining the transform coded data and the spatial domain coded data.

[0014] The implementations of various embodiments of the present invention serve to improve the compression efficiency of modern video and image codecs. Although

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000019

a certain amount of increased computational complexity of encoding may be needed, fast algorithms can be applied in order to lower the encoding complexity that approaches the complexity level for traditional transform based coding. Any effect on the complexity for a decoder is negligible when implementing various embodiments of the present invention.

[0015] These and other advantages and features of the invention, together with the organization and manner of operation thereof, will become apparent from the following detailed description when taken in conjunction with the accompanying drawings, wherein like elements have like numerals throughout the several drawings described below.

BRIEF DESCRIPTION OF THE DRAWINGS

[0016] FIG. 1 shows an example of basis functions of a Hadamard transform that can be used to encode image data;

[0017] Figure 2(a) is an example of four scalar values that are to be coded or decoded and representing the prediction error signal between the original image and the prediction image; Figure 2(b) shows the four weighted basis functions from Figure 1 being summed to the prediction error signal; and Figure 2(c) shows how the same prediction error signal can be reconstructed by summing up only the first basis function and a single spatial sample represented by an arrow;

[0018] Figure 3 is a block diagram of a video encoder constructed in accordance with one embodiment of the present invention;

[0019] Figure 4 is a flow chart showing one encoder algorithm which may be used in an embodiment of the present invention;

[0020] Figure 5 is a flow chart showing one decoder algorithm which may be used in various embodiments of the present invention;

[0021] Figure 6 is a block diagram of a video decoder constructed in accordance with one embodiment of the present invention;

[0022] Figure 7 is an overview diagram of a system within which various embodiments of the present invention may be implemented;

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000020

[0023] Figure 8 is a perspective view of an electronic device that can be used in conjunction with the implementation of various embodiments of the present invention; and

[0024] Figure 9 is a schematic representation of the circuitry which may be included in the electronic device of Figure 8.

DETAILED DESCRIPTION OF VARIOUS EMBODIMENTS

[0025] Various embodiments of the present invention provide a system and method for representing the prediction error signal as a weighted sum of different basis functions of a selected transform and quantized spatial samples. According to various embodiments, the prediction error signal for a single image block is constructed using both transform basis functions and spatial samples (i.e., pixel values), thereby combining the desired features of both the transform and spatial coding approaches discussed previously. This allows for the utilization of those selected transform basis functions that give good overall representation of the image block with minimal amount of transform coefficients (representing the component of prediction error signal that is well correlated with the basis functions). Additionally, various embodiments of the present invention allow for the efficient spatial representation of those components of the prediction error signal of the same image block that are not well correlated with the basis functions of the applied transform (such as certain types of sensor noise, high frequency texture and edge information).

[0026] Figures 1 and 2 show a simplified example of how various embodiments of the present invention are operable in one dimension. Figure 1 shows an example of basis functions of a one-dimensional linear transform that can be used to encode image data (namely a Hadamard transform). The weighted sum of these four basis functions can be used to represent any four samples. Figure 2(a) is an example of four scalar values that are to be coded or decoded and representing the prediction error signal between the original image and the prediction image. In Figure 2(b), the four weighted basis functions from Figure 1 are summed to the prediction error signal (1.5 times the first, 0.5 times the second, -0.5 times the third and -0.5 times the fourth basis function). In Figure 2(c) it is demonstrated how the same prediction error signal

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000021

can be reconstructed by summing up only the first basis function and a single spatial sample represented by an arrow. In this example, compression efficiency improvement can be expected, as the signal is represented by a single transform coefficient and a single spatial sample instead of four transform coefficients weighting the four Hadamard basis functions (as shown in Figure 2(b)).

[0027] Figure 3 is a block diagram of a video encoder constructed in accordance with one embodiment of the present invention. More particularly, Figure 3 shows how an image to be encoded 300 undergoes pixel prediction 302, prediction error coding 303 and prediction error decoding 304. For pixel prediction 302, the image 300 undergoes both inter-prediction 306 and intra-prediction 308 which, after mode selection 310, results in prediction representation of an image block 312. A preliminary reconstructed image 314 is also used for intra-prediction 308. Once all of the image blocks are processed, the preliminary reconstructed image 314 undergoes filtering at 316 to create a final reconstructed image 340, which is sent to a reference frame memory 318 and is also used for inter-prediction 306 of future frames.

[0028] The prediction representation of the image block 312, as well as the image to be encoded 300, are used together to define a prediction error signal 320 which is used for prediction error coding 303. In prediction error coding 303, the prediction error signal 320 undergoes feature selection 322 and spatial quantization 324, as well as transform 326 and quantization 328 (both after feature selection 322). The data describing prediction error and predicted representation of the image block 312 (e.g., motion vectors, mode information and quantized DCT+spatial samples) are passed to entropy coding 330. The prediction error decoding 304 is substantially the opposite of the prediction error coding 303, with the prediction error decoding including an inverse spatial quantizer 332, an inverse transform 334 and an inverse quantization 336. The result of the prediction error decoding 304 is a reconstructed prediction error signal 338, which is used in combination with the predicted representation of the image block 312 to create the preliminary reconstructed image 314.

[0029] Figure 4 is a flow chart showing one encoder algorithm which may be used in an embodiment of the present invention. However, as discussed below, a variety of different algorithms may be used according to principles of the present invention. At

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000022

400 in Figure 4, the difference between sample values of a predicted block and an original input block is calculated. It should be noted that "values" and "outlier values" as used herein are intended to refer to these difference values. At 410, outlier values are found. In this case, the amplitude depends on the expected accuracy of the coded representation; smaller amplitude representations qualify as outliers if one is targeting high bitrate good quality representation, and only high amplitude representations qualify if one is targeting a lower bitrate, lower quality representation. At 420, the modified prediction error signal is transform coded (involving transform, quantization and entropy coding), with the outlier values being substituted by interpolated representations of those values, e.g., by averaging the neighboring prediction error values. At 430, the modified prediction error signals are transform decoded back to the spatial domain. At 440, spatial coding is applied (involving quantization and entropy coding without transform) to the difference between the original outlier values and the transform decoded outlier values. At 450, the final coded representation of the prediction error signal is formed by joining the transform coded data and the spatial coded data. In an alternative embodiment, the spatial coding occurs before transform coding.

[0030] When considering the one-dimensional example depicted in Figures 2(a)-2(c), the algorithm outlined in Figure 4 would process the relevant data as follows. As discussed previously, Figure 2(a) represents the difference signal [1 1 3 1] to be coded. In this case, [3] is an outlier. Therefore, the prediction error signal is modified by substituting the [3] with an average of the second and fourth samples, making the signal [1 1 1 1]. The difference between the original outlier sample ([3]) and the transform coded outlier value ([1]) is then spatial coded, making the spatial coded signal [0 0 2 0]. The entropy coded representations of the transform coded and spatial coded signals are then written to the bitstream. Thereafter, the decoder can recover the coded signal by adding up the inverse transformed data [1 1 1 1] and the inverse spatial coded data [0 0 2 0] (together with the prediction samples of the image block).

[0031] When considering the above example, it is important to note that the choice of Hadamard basis vectors for the transform is only intended to be exemplary in nature. In fact, the various methods and techniques described herein can be applied to

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000023

any transform employing basis functions, and the basis functions do not necessarily need to be orthogonal.

[0032] A video or image codec, according to various embodiments the present invention, can be implemented in various ways. In the case of an encoder, the encoder can use different strategies to find the desired transform coefficients and spatial samples. For example, an encoder can first transform code the signal and apply spatial coding to the residual. The encoder can also first apply spatial coding to the signal, followed by transform coding of the residual. Additionally, some or all of the transform coefficients can be set to zero in order to improve the coding performance. Some or all spatial samples can also be set to zero in order to improve the coding performance. An encoder can also iteratively modify transform coefficients and/or spatial samples in order to improve the coding performance until a desired performance or a maximum defined number of iterations is achieved.

[0033] In terms of quantization and dequantization, the quantization and dequantization of transform coefficients and spatial samples can be tied together (e.g., a quantization step size for both transform coefficients and spatial samples can be derived from a single parameter). Alternatively, different quantizations and dequantizations can be applied to transform coefficients and spatial samples.

[0034] For the coding and decoding of spatial samples, such coding and decoding can depend on transform coefficients and vice versa. Alternatively, the coding and decoding of spatial samples, transform coefficients or both can depend on the prediction signal, on the other transform coefficients and spatial samples in the same image or in other images.

[0035] In addition to the above, it can be indicated that there are no transform coefficients or spatial samples for a specific image area or image. It can also be indicated that only spatial coding or only transform coding is used for a specific image area or image. The number of spatial samples can be coded and decoded as one unit, for example representing certain patterns of textures. Pre- and/or post-processing mechanisms can be applied to the prediction signal, prediction error signal reconstructed signal or any combination thereof. The method can be used to code and decode other information instead or in addition to the prediction error signals. The

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000024

codec can limit the usage of either transform coefficients (e.g., allow only low frequency transform coefficients to be present in the coded representation of the signal) or spatial samples.

[0036] Figure 5 shows a decoding process according to various embodiments of the present invention. At 500 in Figure 5, a video decoder receives a coded prediction error signal which comprises both transformed coefficients and spatial samples. At 510, both the transformed coefficients and the spatial samples are decoded. At 520, the decoded transform information, the decoded spatial information and reconstructed prediction of the image block are added together to form decoded representation of the image block.

[0037] Figure 6 is a block diagram of a video decoder constructed in accordance with one embodiment of the present invention. As shown in Figure 6, entropy decoding 600 is followed by both prediction error decoding 602 and pixel prediction 604. In prediction error decoding 602, in addition to an inverse transform 606 and inverse quantization 608, an inverse spatial quantizer 610 is used as discussed herein, ultimately resulting in a reconstructed prediction error signal 612. For pixel prediction 604, either intra-prediction or inter-prediction occurs at 614 to create a predicted representation of an image block 616. The predicted representation of the image block 616 is used in conjunction with the reconstructed prediction error signal 612 to create a preliminary reconstructed image 618, which in turn can be used both for prediction 614. Once all of the image blocks have been processed, the preliminary reconstructed image 618 is passed for filtering 620. The filtered image can also be stored in reference frame memory 624, making it usable for prediction 614 as well.

[0038] Figure 7 shows a system 10 in which various embodiments of the present invention can be utilized, comprising multiple communication devices that can communicate through one or more networks. The system 10 may comprise any combination of wired or wireless networks including, but not limited to, a mobile telephone network, a wireless Local Area Network (LAN), a Bluetooth personal area network, an Ethernet LAN, a token ring LAN, a wide area network, the Internet, etc. The system 10 may include both wired and wireless communication devices.

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000025

[0039] For exemplification, the system 10 shown in Figure 7 includes a mobile telephone network 11 and the Internet 28. Connectivity to the Internet 28 may include, but is not limited to, long range wireless connections, short range wireless connections, and various wired connections including, but not limited to, telephone lines, cable lines, power lines, and the like.

[0040] The exemplary communication devices of the system 10 may include, but are not limited to, an electronic device 50, a combination personal digital assistant (PDA0 and mobile telephone 14, a PDA 16, an integrated messaging device (IMD) 18, a desktop computer 20, a notebook computer 22, etc. The communication devices may be stationary or mobile as when carried by an individual who is moving. The communication devices may also be located in a mode of transportation including, but not limited to, an automobile, a truck, a taxi, a bus, a train, a boat, an airplane, a bicycle, a motorcycle, etc. Some or all of the communication devices may send and receive calls and messages and communicate with service providers through a wireless connection 25 to a base station 24. The base station 24 may be connected to a network server 26 that allows communication between the mobile telephone network 11 and the Internet 28. The system 10 may include additional communication devices and communication devices of different types.

[0041] The communication devices may communicate using various transmission technologies including, but not limited to, Code Division Multiple Access (CDMA), Global System for Mobile Communications (GSM), Universal Mobile Telecommunications System (UMTS), Time Division Multiple Access (TDMA), Frequency Division Multiple Access (FDMA), Transmission Control Protocol/Internet Protocol (TCP/IP), Short Messaging Service (SMS), Multimedia Messaging Service (MMS), e-mail, Instant Messaging Service (IMS), Bluetooth, IEEE 802.11, etc. A communication device involved in implementing various embodiments of the present invention may communicate using various media including, but not limited to, radio, infrared, laser, cable connection, and the like.

[0042] Figures 8 and 9 show one representative electronic device 50 within which the present invention may be implemented. It should be understood, however, that the present invention is not intended to be limited to one particular type of device. The

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000026

electronic device 50 of Figures 8 and 9 includes a housing 30, a display 32 in the form of a liquid crystal display, a keypad 34, a microphone 36, an ear-piece 38, a battery 40, an infrared port 42, an antenna 44, a smart card 46 in the form of a UICC according to one embodiment of the invention, a card reader 48, radio interface circuitry 52, codec circuitry 54, a controller 56 and a memory 58. Individual circuits and elements are all of a type well known in the art, for example in the Nokia range of mobile telephones.

[0043] The various embodiments of the present invention described herein is described in the general context of method steps or processes, which may be implemented in one embodiment by a computer program product, embodied in a computer-readable medium, including computer-executable instructions, such as program code, executed by computers in networked environments. Generally, program modules may include routines, programs, objects, components, data structures, etc. that perform particular tasks or implement particular abstract data types. Computer-executable instructions, associated data structures, and program modules represent examples of program code for executing steps of the methods disclosed herein. The particular sequence of such executable instructions or associated data structures represents examples of corresponding acts for implementing the functions described in such steps or processes.

[0044] Software and web implementations of various embodiments of the present invention can be accomplished with standard programming techniques with rule-based logic and other logic to accomplish various database searching steps or processes, correlation steps or processes, comparison steps or processes and decision steps or processes. It should be noted that the words "component" and "module," as used herein and in the following claims, is intended to encompass implementations using one or more lines of software code, and/or hardware implementations, and/or equipment for receiving manual inputs.

[0045] The foregoing description of embodiments of the present invention have been presented for purposes of illustration and description. The foregoing description is not intended to be exhaustive or to limit embodiments of the present invention to the precise form disclosed, and modifications and variations are possible in light of

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000027

the above teachings or may be acquired from practice of various embodiments of the present invention. The embodiments discussed herein were chosen and described in order to explain the principles and the nature of various embodiments of the present invention and its practical application to enable one skilled in the art to utilize the present invention in various embodiments and with various modifications as are suited to the particular use contemplated.

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000028

WHAT IS CLAIMED IS:

1. A method of encoding a prediction error signal for a block of data, comprising:

calculating a difference signal representing differences between sample values of a predicted block of data and values for an original input block;

performing transform coding to the difference signal, thereby creating a first representation of a first component of the difference signal;

performing spatial coding to the difference signal, thereby creating a second representation of a first component of the difference signal; and

joining the first and second representations.

2. The method of claim 1, further comprising, before performing transform coding and spatial coding, substituting outlier values in the difference signal with non-outlier values to create a modified prediction error signal.

3. The method of claim 2, wherein transform coding is applied to the modified prediction error signal and spatial coding is applied to a residual thereof.

4. The method of claim 2, wherein spatial coding is applied to the modified prediction error signal and transform coding is applied to a residual thereof.

5. The method of claim 1, wherein, for transform coding, at least one transform coefficient is set to zero.

6. The method of claim 1, wherein, for spatial coding, at least one spatial sample is set to zero.

7. The method of claim 2, wherein the non-outlier values used to replace the outlier values comprise an average of neighboring prediction error values.

8. The method of claim 1, further comprising providing an indication that no transform coefficients exist in the block of data if appropriate.

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000029

9. The method of claim 1, further comprising providing an indication that no spatial samples exist in the block of data if appropriate.

10. The method of claim 1, wherein for the transform coding, a transform coefficient is formed, the transform coefficient comprising a weighting coefficient of a discrete transform basis vector.

11. The method of claim 10, wherein the discrete transform comprises a discrete orthogonal transform.

12. A computer program product, embodied in a computer-readable medium, comprising computer code configured to perform the implemented the processes of claim 1.

13. An apparatus, comprising:

a processor; and

a memory unit communicatively connected to a controller and comprising:

computer code for performing transform coding to the difference signal, thereby creating a first representation of a first component of the difference signal;

computer code for performing spatial coding to the difference signal, thereby creating a second representation of a first component of the difference signal; and

computer code for joining the first and second representations.

14. The apparatus of claim 13, wherein the memory unit further comprises computer code for, before performing transform coding and spatial coding, substituting outlier values in the difference signal with non-outlier values to create a modified prediction error signal.

15. The apparatus of claim 14, wherein transform coding is applied to the modified prediction error signal and spatial coding is applied to a residual thereof.

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000030

16. The apparatus of claim 14, wherein spatial coding is applied to the modified prediction error signal and transform coding is applied to a residual thereof.

17. The apparatus of claim 13, wherein, for transform coding, at least one transform coefficient is set to zero.

18. The apparatus of claim 13, wherein, for spatial coding, at least one spatial sample is set to zero.

19. The apparatus of claim 13, wherein the non-outlier values used to replace the outlier values comprise an average of neighboring prediction error values.

20. The apparatus of claim 13, further comprising providing an indication that no transform coefficients exist in the block of data if appropriate.

21. The apparatus of claim 13, further comprising providing an indication that no spatial samples exist in the block of data if appropriate.

22. A method of decoding a prediction error signal for a block of data, comprising:

receiving a coded prediction error signal, the coded prediction error signal including a plurality of transform coefficients and a plurality of spatial samples;

decoding the plurality of transformed coefficients into decoded transform information;

decoding the plurality of spatial samples into decoded spatial information; and

adding the decoded transform information, the decoded spatial information, and a reconstructed prediction of the block of data, thereby forming a decoded representation of the block of data.

23. The method of claim 22, wherein an indication is provided that no transform coefficients exist in the block of data if appropriate.

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000031

24. The method of claim 22, wherein an indication is provided that that no spatial samples exist in the block of data if appropriate.

25. The method of claim 22, wherein the decoding of the plurality of transformed coefficients and spatial samples includes dequantization of each, and wherein the dequantization of transformed coefficients and spatial samples are related to each other.

26. The method of claim 22, wherein the decoding of the plurality of transformed coefficients and spatial samples includes dequantization of each, and wherein the dequantization of transformed coefficients and spatial samples are different from each other.

27. The method of claim 22, wherein the decoding of the spatial samples is dependent upon the decoding of the transformed coefficients.

28. The method of claim 22, wherein the decoding of the transformed coefficients is dependent upon the decoding of the spatial samples.

29. The method of claim 22, wherein at least some of the plurality of spatial samples are decoded as a single unit.

30. A computer program product, embodied in a computer-readable medium, comprising computer code configured to perform the processes of claim 19.

31. An apparatus, comprising:

a processor; and

a memory unit communicatively connected to the processor and including:

computer code for processing a received coded prediction error signal, the coded prediction error signal including a plurality of transformed coefficients and a plurality of spatial samples;

computer code for decoding the plurality of transformed coefficients into decoded transform information;

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000032

computer code for decoding the plurality of spatial samples into decoded spatial information; and

computer code for adding the decoded transform information, the decoded spatial information, and a reconstructed prediction of the block of data, thereby forming a decoded representation of the block of data.

32. The apparatus of claim 31, wherein an indication is provided that no transform coefficients exist in the block of data if appropriate.

33. The apparatus of claim 31, wherein an indication is provided that that no spatial samples exist in the block of data if appropriate.

34. The apparatus of claim 31, wherein the decoding of the plurality of transformed coefficients and spatial samples includes dequantization of each, and wherein the dequantization of transformed coefficients and spatial samples are related to each other.

35. The apparatus of claim 31, wherein the decoding of the plurality of transformed coefficients and spatial samples includes dequantization of each, and wherein the dequantization of transformed coefficients and spatial samples are different from each other.

36. The apparatus of claim 31, wherein the decoding of the spatial samples is dependent upon the decoding of the transformed coefficients.

37. The apparatus of claim 31, wherein the decoding of the transformed coefficients is dependent upon the decoding of the spatial samples.

38. The apparatus of claim 31, wherein at least some of the plurality of spatial samples are decoded as a single unit.

39. A method of encoding a prediction error signal for a block of data, comprising:

calculating a difference signal representing differences between sample values of a predicted block of data and values for an original input block;

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000033

substituting outlier values in the difference signal with non-outlier values to create a modified prediction error signal;

performing one of transform coding and spatial coding to the modified prediction error signal, creating one of a transform coded data and spatial coded data

performing an inverse of the one of transform coding and spatial coding;

applying the other of transform coding and spatial coding to a residual from the inverse transform or spatial decoding, creating the other of transform coded data and spatial coded data; and

joining the transform coded data and the spatial coded data.

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000034

What is claimed is:

1. A method of encoding a prediction error signal for a block of data, comprising:

calculating a difference signal representing differences between sample values of a predicted block of data and values for an original input block;

creating a modified prediction error signal by substituting outlier values in the difference signal with non-outlier values;

performing transform coding to the modified prediction error signal, thereby creating a first representation of a first component of the difference signal;

performing spatial domain coding to the difference between the difference signal and the transform decoded modified error signal, thereby creating a second representation of the first component of the difference signal; and

joining the first and second representations.

2. The method of claim 1, wherein, for transform coding, at least one transform coefficient is set to zero.

3. The method of claim 1 or 2, wherein, for spatial domain coding, at least one spatial domain sample is set to zero.

4. The method of any one of claims 1 to 3, wherein the non-outlier values used to replace the outlier values comprise an average of neighboring prediction error values.

5. The method of any one of claims 1 to 4, further comprising providing an indication that no transform coefficients exist in the block of data if appropriate.

6. The method of any one of claims 1 to 4, further comprising providing an indication that no spatial domain samples exist in the block of data if appropriate.

7. The method of claim 1, wherein for the transform coding, a transform coefficient is formed, the transform coefficient comprising a weighting coefficient of a discrete transform basis vector.

8. The method of claim 7, wherein the discrete transform comprises a discrete orthogonal transform.

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000035

9. A computer-readable medium storing computer code configured to perform the method of any one of claims 1 to 8.

10. An apparatus, comprising:

a processor; and

a memory unit communicatively connected to a controller and comprising:

computer code for calculating a difference signal representing differences between sample values of a predicted block of data and values for an original input block;

computer code for creating a modified prediction error signal by substituting outlier values in the difference signal with non-outlier values;

computer code for performing transform coding to the difference signal, thereby creating a first representation of a first component of the difference signal;

computer code for performing spatial domain coding to the difference between the difference signal and the transform decoded modified error signal, thereby creating a second representation of the first component of the difference signal; and

computer code for joining the first and second representations.

11. The apparatus of claim 10, wherein, for transform coding, at least one transform coefficient is set to zero.

12. The apparatus of claim 10 or 11, wherein, for spatial domain coding, at least one spatial domain sample is set to zero.

13. The apparatus of any one of claims 10 to 12, wherein the non-outlier values used to replace the outlier values comprise an average of neighboring prediction error values.

14. The apparatus of any one of claims 10 to 13, further comprising providing an indication that no transform coefficients exist in the block of data if appropriate.

15. The apparatus of any one of claims 10 to 13, further comprising providing an indication that no spatial domain samples exist in the block of data if appropriate.

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000036

16. A method of decoding a prediction error signal for a block of data, comprising:

receiving a coded prediction error signal, the coded prediction error signal including a plurality of transform coefficients and a plurality of spatial domain samples;

decoding the plurality of transformed coefficients into decoded transform information;

decoding the plurality of spatial domain samples into decoded spatial domain information;

adding the decoded transform information, the decoded spatial domain information, and a reconstructed prediction of the block of data, thereby forming a decoded representation of the block of data; and

providing an indication that at least one of no transform coefficients exist in the block of data and no spatial domain samples exist in the block of data.

17. The method of claim 16, wherein the decoding of the plurality of transformed coefficients and spatial domain samples includes dequantization of each, and wherein the dequantization of transformed coefficients and spatial domain samples are related to each other.

18. The method of claim 16, wherein the decoding of the plurality of transformed coefficients and spatial domain samples includes dequantization of each, and wherein the dequantization of transformed coefficients and spatial domain samples are different from each other.

19. The method of claim 16, wherein the decoding of the spatial domain samples is dependent upon the decoding of the transformed coefficients.

20. The method of claim 16, wherein the decoding of the transformed coefficients is dependent upon the decoding of the spatial domain samples.

21. The method of claim 16, wherein at least some of the plurality of spatial domain samples are decoded as a single unit.

22. A computer-readable medium storing computer code configured to perform the method of any one of claims 16 to 21.

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000037

23. An apparatus, comprising:

a processor; and

a memory unit communicatively connected to the processor and comprising:

computer code for processing a received coded prediction error signal, the coded prediction error signal including a plurality of transformed coefficients and a plurality of spatial domain samples;

computer code for decoding the plurality of transformed coefficients into decoded transform information;

computer code for decoding the plurality of spatial domain samples into decoded spatial domain information;

computer code for adding the decoded transform information, the decoded spatial domain information, and a reconstructed prediction of the block of data, thereby forming a decoded representation of the block of data; and

computer code for providing an indication that at least one of no transform coefficients exist in the block of data and no spatial domain samples exist in the block of data.

24. The apparatus of claim 23, wherein the decoding of the plurality of transformed coefficients and spatial domain samples includes dequantization of each, and wherein the dequantization of transformed coefficients and spatial domain samples are related to each other.

25. The apparatus of claim 23, wherein the decoding of the plurality of transformed coefficients and spatial domain samples includes dequantization of each, and wherein the dequantization of transformed coefficients and spatial domain samples are different from each other.

26. The apparatus of claim 23, wherein the decoding of the spatial domain samples is dependent upon the decoding of the transformed coefficients.

27. The apparatus of claim 23, wherein the decoding of the transformed coefficients is dependent upon the decoding of the spatial domain samples.

28. The apparatus of any one of claims 23 to 27, wherein at least some of the plurality of spatial domain samples are decoded as a single unit.

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000038

29. A method of encoding a prediction error signal for a block of data, comprising:

calculating a difference signal representing differences between sample values of a predicted block of data and values for an original input block;

substituting outlier values in the difference signal with non-outlier values to create a modified prediction error signal;

performing one of transform coding and spatial domain coding to the modified prediction error signal, creating one of a transform coded data and spatial domain coded data;

performing an inverse of the one of transform coding and spatial domain coding;

applying the other of transform coding and spatial domain coding to a residual from the inverse transform or spatial domain decoding, creating the other of transform coded data and spatial domain coded data; and

joining the transform coded data and the spatial domain coded data.

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000039

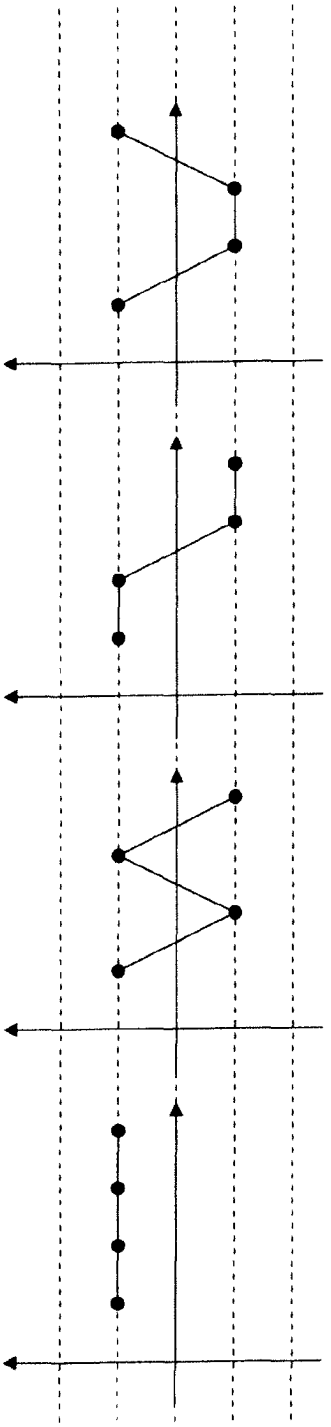

FIG. 1

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000040

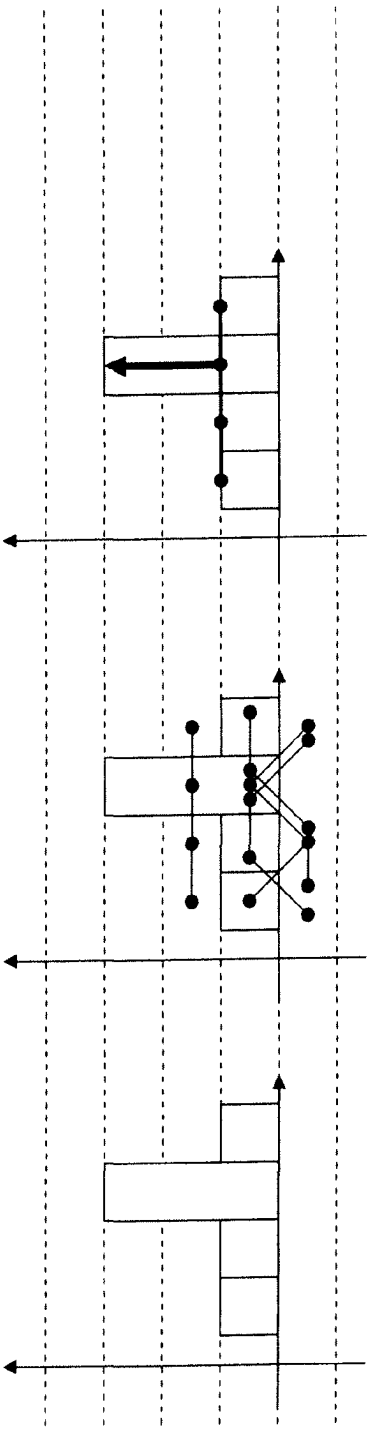

FIG. 2(a) FIG. 2(b) FIG. 2(c)

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000041

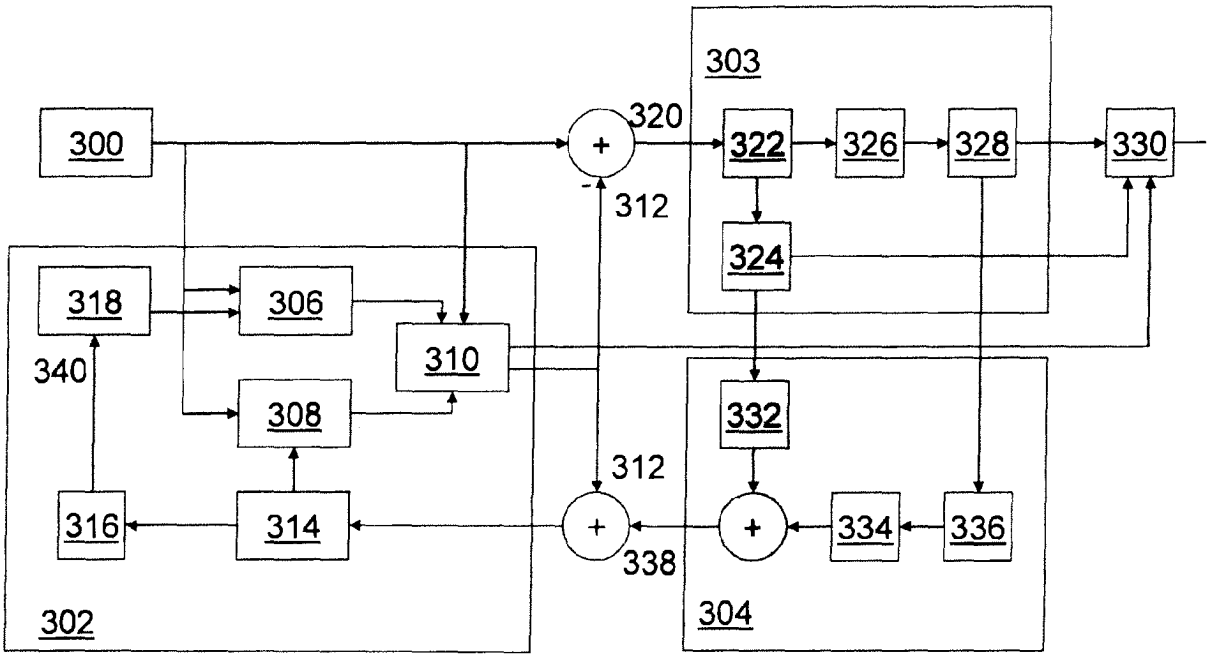

FIG. 3

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000042

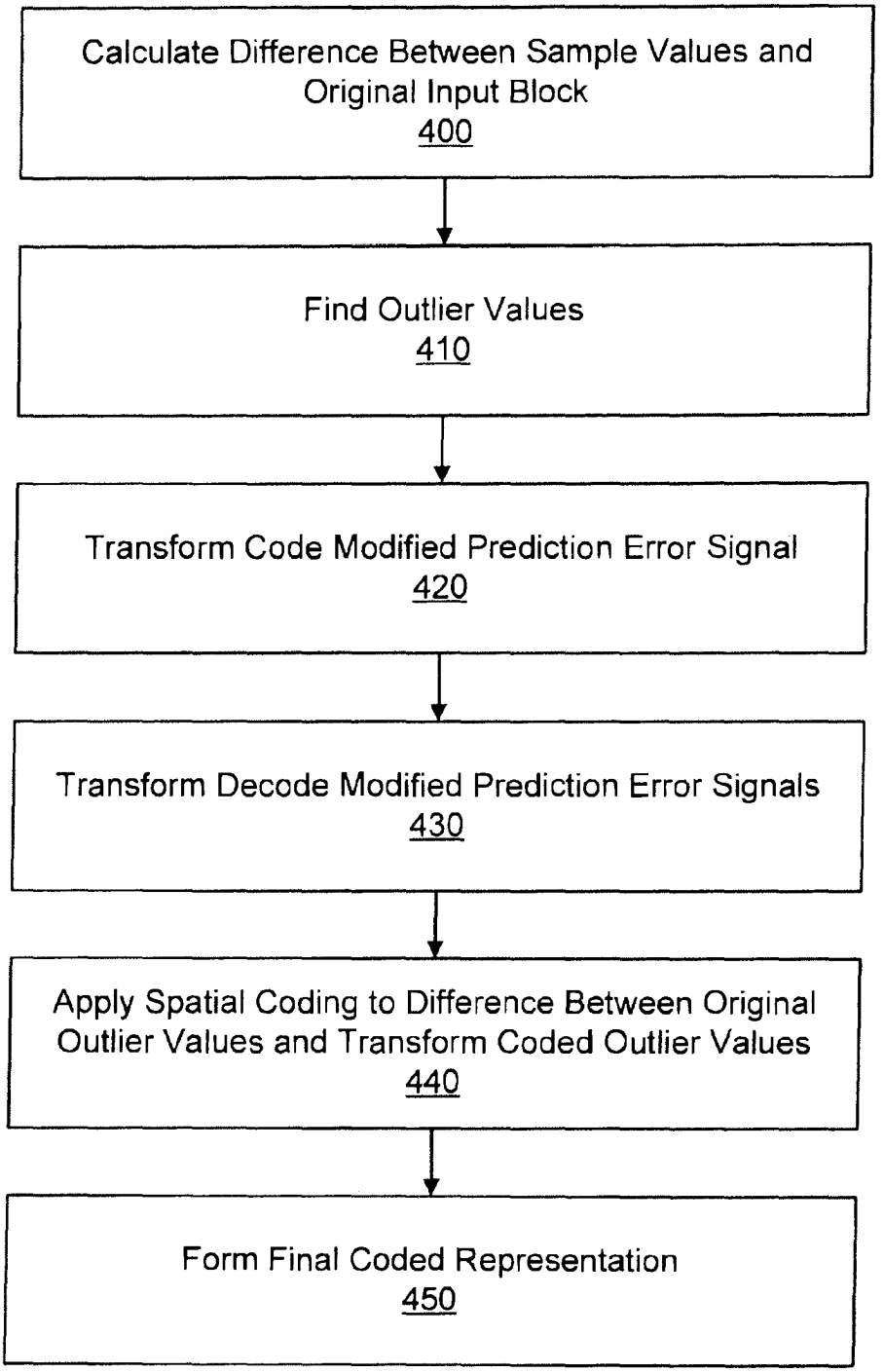

FIG. 4

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000043

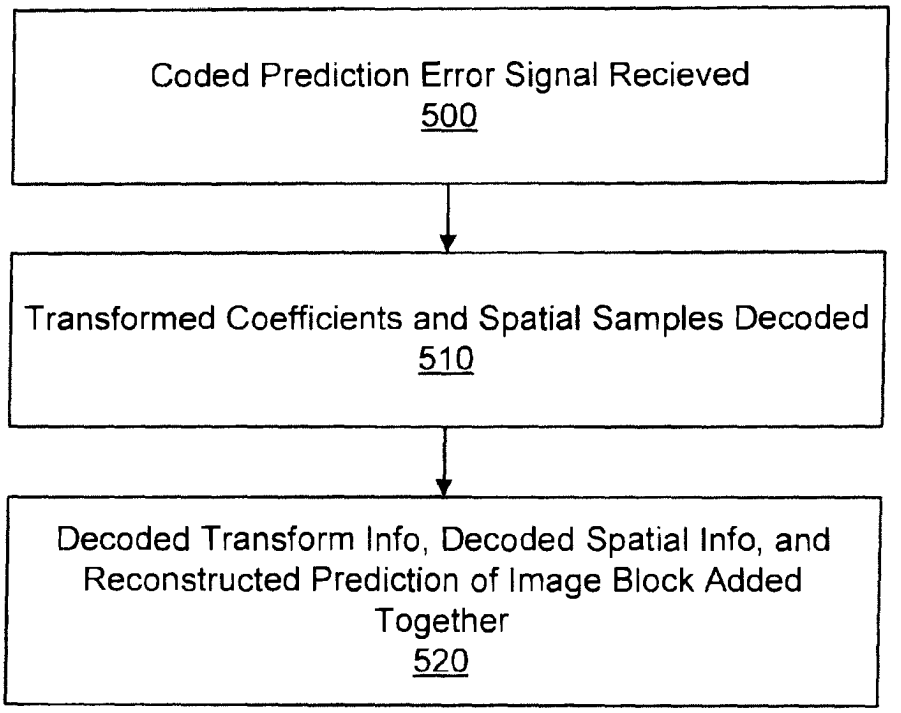

FIG. 5

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000044

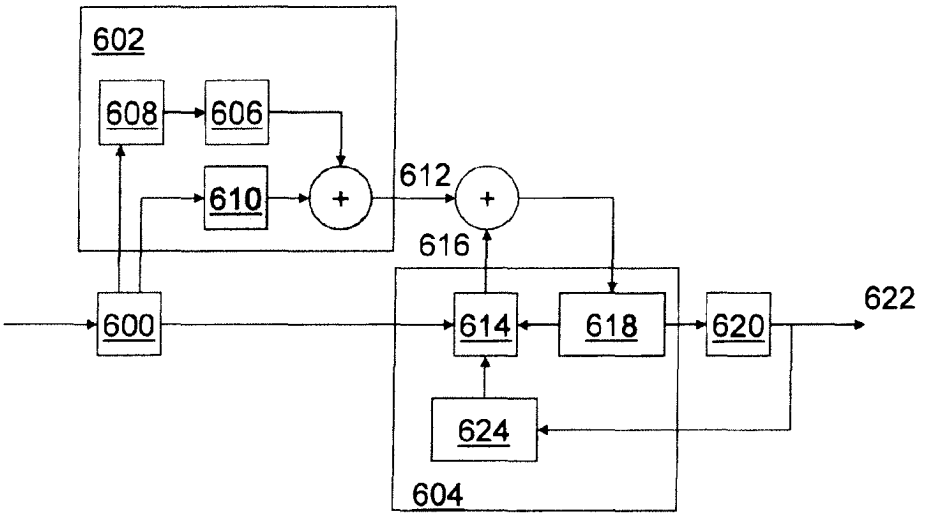

FIG. 6

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000045

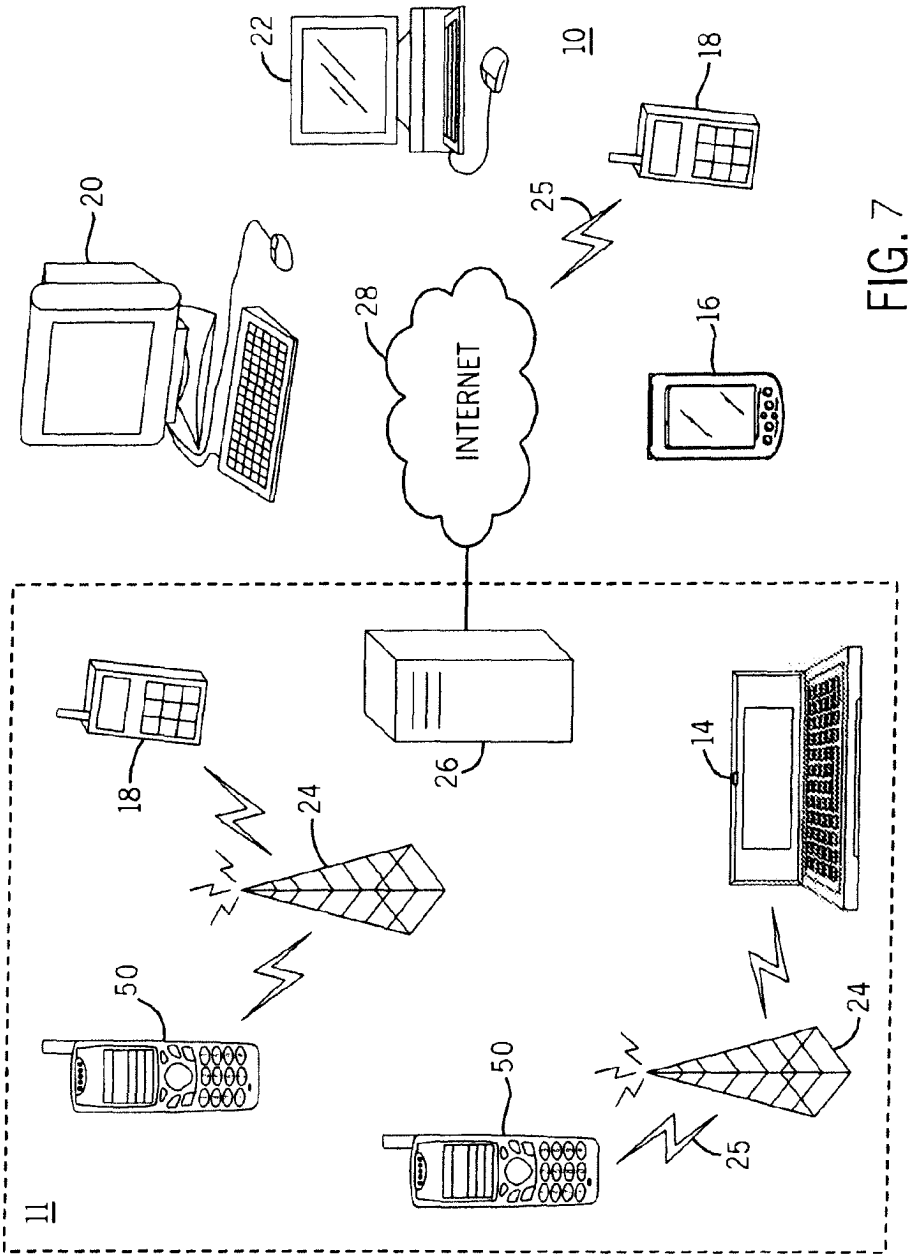

FIG. 7

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000046

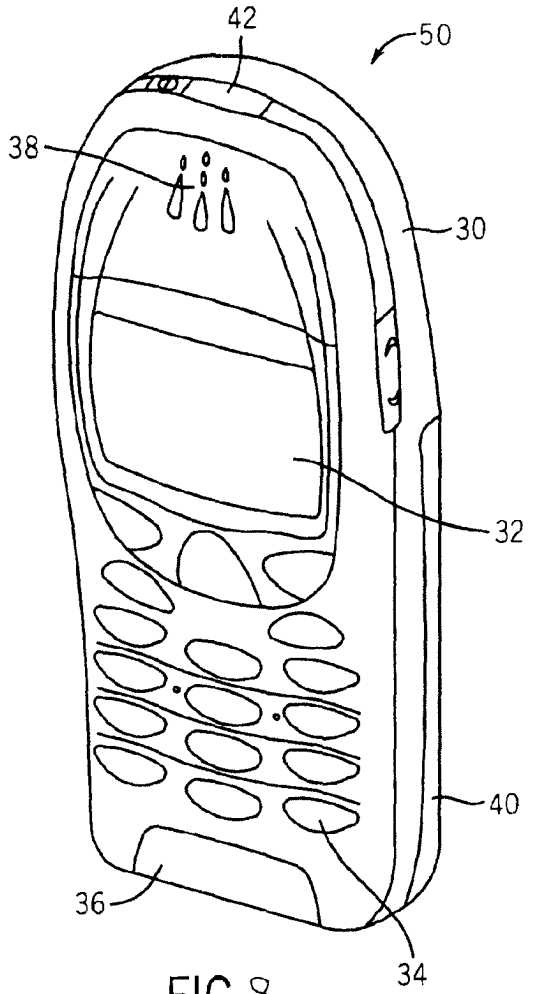

FIG.8

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000047

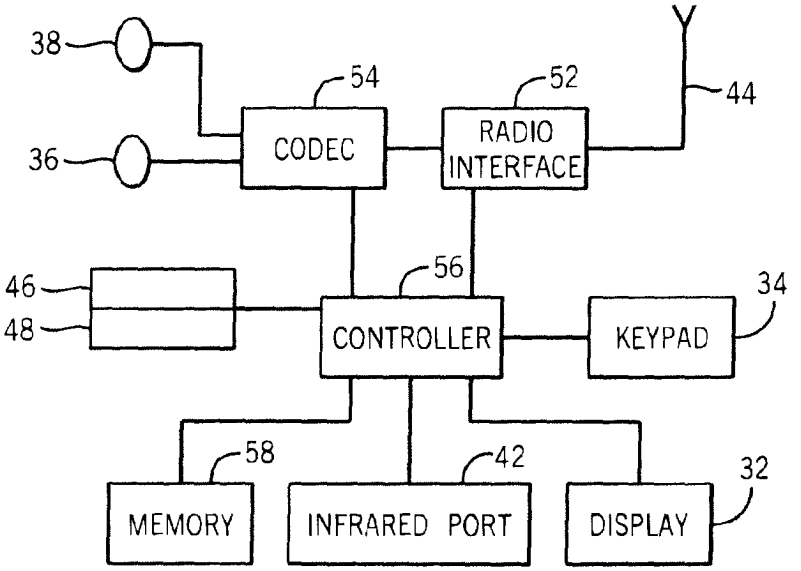

FIG. 9

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000048

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
23 October 2008 (23.10.2008)

PCT

(10) International Publication Number
WO 2008/126019 A3

(51) International Patent Classification:
H04N 7/50 (2006.01)

(21) International Application Number:
PCT/IB2008/051351

(22) International Filing Date: 9 April 2008 (09.04.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/911,480 12 April 2007 (12.04.2007) US

(71) Applicant *(for all designated States except US)*: **NOKIA CORPORATION** [FI/FI]; Keilalahdentie 4, FIN-02150 Espoo (FI).

(71) Applicant *(for LC only)*: **NOKIA, INC.** [US/US]; 6000 Connection Drive, Irving, TX 75039 (US).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **LAINEMA, Jani** [FI/FI]; Kisakentakatu 12 B 6, FIN-33230 Tampere (FI).

(74) Agent: **ALBERT, G.Peter**; Foley & Lardner LLP, 11250 El Camino Real, Suite 200, San Diego, CA 92130-2677 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

(88) Date of publication of the international search report:
4 December 2008

(54) Title: SPATIALLY ENHANCED TRANSFORM CODING

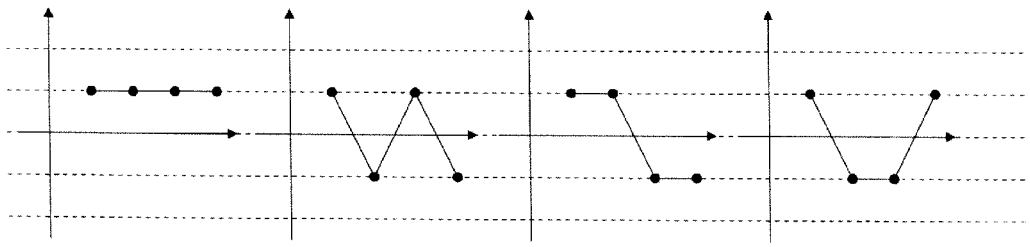

FIG. 1

(57) Abstract: A system and method for improving prediction error coding performance of various video compression algorithms. Various embodiments combine the energy compaction features of transform coding with localization properties of spatial coding. In various embodiments, the effects of pixel "outliers" are removed from the transform and are coded separately as a spatially coded pixel prediction, thereby improving the coding gain of the transform.

WO 2008/126019 A3

SIM & McBURNEY
Patent & Trade-Mark Agents

330 University Avenue
6th Floor
Toronto, Ontario
M5G 1R7
Tel.: (416) 595-1155
Fax (416) 595-1163
www.sim-mcburney.com

David A. Ruston
Direct Line: (416) 849-8390
ruston@sim-mcburney.com

October 9, 2009

Our Ref: 14010-27 DAR

BOX PCT

The Commissioner of Patents
Canadian Intellectual Property Office
Place du Portage I
50 Victoria Street, Room C-114
Gatineau, Quebec
Canada K1A 0C9

Industry Canada Industrie Canada
A/M/J Y/M/D
2009/10/09
286 - 09
CIPO OPIC B001409173

Dear Madam:

This is a request to enter the national phase in Canada for International Patent Application No. PCT/IB2008/051351, having a filing date of April 9, 2008, in the name of **NOKIA CORPORATION** entitled **SPATIALLY ENHANCED TRANSFORM CODING** comprising the following:

- International Application as Published
- Request for National Entry
- Claim to Convention Priority
- Declaration under PCT Rule 4.17
  [X] Enclosed
  [ ] To Follow

The Patent Office is respectfully requested to deduct from the Sim & McBurney Deposit Account No. 600000249 the total amount indicated below to be applied as follows:

Basic national fee - large entity	$ 400.00
Examination fee	$ 800.00
First maintenance fee due **April 9, 2010**	$ 100.00
Total	**$1,300.00**

The Requisite Fee Payment Form is enclosed herewith.

Michael I. Stewart, B.Sc. (Hons.) (Chemistry) | Toni P. Ashton, LL.B | Kenneth D. McKay, LL.B, CLP | Timothy M. Lowman, LL.B, CLP, MCIArb, Certified Specialist – Civil Litigation | Stephen M. Lane, B.A.Sc. (Chemical Engineering) | David A. Ruston, B.Sc. (Electrical Engineering) | Lola A. Bartoszewicz, Ph.D. (Clinical Biochemistry) | Andrew Jones, LL.B | Kimberly A. McManus, Ph.D. (Chemistry) | Matthew D. Powell, B.Sc. (Electrical Engineering) | Marlye Monfiston, LL.B | Rohit Parekh, M.Sc. (Molecular Biology) | A. SuMei Cheung, B.Eng. (Engineering Physics) |

Technical Consultants · Erica L. Lowthers, Ph.D. (Pharmacology) | Evangellos W. Vekris, Ph.D. (Materials Science & Engineering) | Patrick J. Cotter, LL.B | Sanjukta Tole, B.Sc. (Hons.) (Human Behavioural Biology)

Senior Consultants · Peter W. McBurney, B.Sc. (Hons.) (Electrical Engineering) (Electronics) | Brenda L. Boardman, B.A. | John H. Woodley (1947-2008) | Patricia A. Rae, Ph.D. (Biochemistry)

If payment of the examination fee is tendered herewith, the Applicant requests examination of the application.

Should the fees submitted with this letter be insufficient to cover all of the fees for which payment is explicitly or implicitly requested by this letter, the Commissioner is requested to charge the amount of the insufficiency to the Sim & McBurney deposit account already on file. In this eventuality, we would appreciate being advised by telephone or correspondence of the amount so deducted and our file reference as set out on this letter. In the event that there are insufficient funds in the Sim & McBurney deposit account to cover the indicated amount or any deficiency therein, it is requested that CIPO telephone Applicant's Agent or Janine Greenberg @ 416-849-8377, as soon as possible to notify of the same so that additional funds can be remitted promptly.

Respectfully submitted,

SIM & McBURNEY
Patent & Trade-Mark Agents

David A. Ruston/mgr
Encls.

Industry Canada Industrie Canada
A/M/J Y/M/D
2009/10/09
286- 09
CIPO OPIC
B001409175

PCT/CA

Our Reference: **14010-27 DAR**

FORM FOR REQUEST FOR ENTRY INTO NATIONAL PHASE UNDER ARTICLE 22 OR 39 OF THE PATENT COOPERATION TREATY

IN THE MATTER OF PCT International Application

 # PCT/IB2008/051351 (international application #) filed **April 9, 2008** (international filing date)

in the name(s) of **NOKIA CORPORATION** (full name of Applicant(s))

1. The above-identified Applicant(s), whose full post office address(es) is **Keilalahdentie 4, FIN-02150 Espoo, Finland**, own(s) the entire right in Canada to an invention entitled **SPATIALLY ENHANCED TRANSFORM CODING** made by

(a) **LAINEMA, Jani** of **Kisakentakatu 12 B 6, FIN-33230 Tampere, Finland** (inventor's name & address)

as described and claimed in the said international application and any amendments thereto.

2. The Applicant(s) claim(s) the protection afforded by Section 28(1) of the *Patent Act* in relation to the following prior filed application(s):

Country of Filing	Application No.	Filing Date
United States of America	60/911,480	April 12, 2007

3. The Applicant(s) appoint(s) **SIM & McBURNEY**, whose complete address in Canada is **6th Floor, 330 University Avenue, Toronto, Ontario, Canada M5G 1R7, Telephone (416) 595-1155, Fax No. (416) 595-1163**, as the representative of the Applicant(s) in Canada, pursuant to Section 29 of the *Patent Act*.

4. The Applicant(s) appoint(s) **SIM & McBURNEY**, whose complete address is as set out above, as the patent agent of the Applicant(s).

5. The Applicant(s) accordingly hereby request(s) commencement of national phase procedures consequent to the designation of Canada in such international application and pray(s) that a patent may be granted to the Applicant(s) for the said invention.

Sim & McBurney
Agents for the Applicant(s)

Re New Canadian Patent Application, National Phase of **PCT/IB2008/051351**

Title: **SPATIALLY ENHANCED TRANSFORM CODING**

Applicant: **NOKIA CORPORATION**

DECLARATION UNDER PCT RULE 4.17
ENTITLEMENT TO APPLY FOR AND BE GRANTED A PATENT

Declaration as to the Applicant's entitlement, as at the international filing date, to apply for and be granted a patent (Rules 4.17(ii) and 51 *bis*.1(a)(ii)), in a case where the Declaration under Rule 4.17(iv) is not appropriate:

in relation to International Application No. **PCT/IB2008/051351**, **NOKIA CORPORATION** was (were) entitled to apply for and be granted a patent by virtue of the following:

(i) **LAINEMA, Jani** of **Kisakentakatu 12 B 6, FIN-33230 Tampere, Finland**, is (are) the inventor(s) of the subject matter for which protection is sought by way of the international application,

(ii) **NOKIA CORPORATION** was entitled as employer of the inventor(s), **LAINEMA, Jani**,

(iii) an agreement between _____ (name) and _____ (name), dated _____,

(iv) an Assignment from _____(inventor(s)) to _____ (applicant), dated _____,

(v) consent from _____ (name) in favour of _____ (name), dated _____,

(vi) a court order issued by _____ (name of court), effecting a transfer from _____ (name) to _____ (name), dated _____,

(vii) transfer of entitlement from _____ (name) to _____ (name) by way of _____ (specify kind of transfer), dated _____,

(viii) the Applicant's name changed from _____ (name) to _____ (name) on _____ (date).

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
23 October 2008 (23.10.2008)

PCT

(10) International Publication Number
WO 2008/126019 A2

(51) International Patent Classification:
H04N 7/50 (2006.01)

(21) International Application Number:
PCT/IB2008/051351

(22) International Filing Date: 9 April 2008 (09.04.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/911,480 12 April 2007 (12.04.2007) US

(71) Applicant *(for all designated States except US)*: **NOKIA CORPORATION** [FI/FI]; Keilalahdentie 4, FIN-02150 Espoo (FI).

(71) Applicant *(for LC only)*: **NOKIA, INC.** [US/US]; 6000 Connection Drive, Irving, TX 75039 (US).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **LAINEMA, Jani** [FI/FI]; Kisakentakatu 12 B 6, FIN-33230 Tampere (FI).

(74) Agent: **ALBERT, G.Peter**; Foley & Lardner LLP, 11250 El Camino Real, Suite 200, San Diego, CA 92130-2677 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

(54) Title: SPATIALLY ENHANCED TRANSFORM CODING

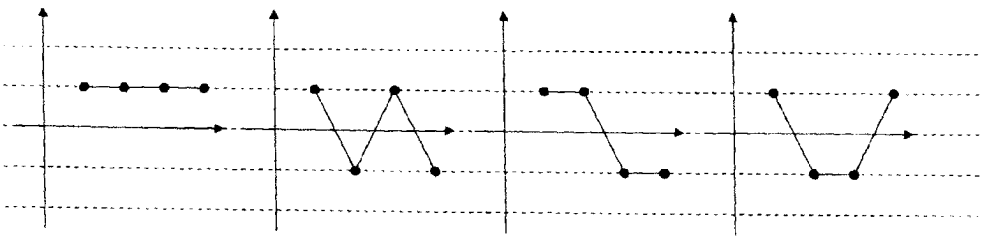

FIG. 1

(57) Abstract: A system and method for improving prediction error coding performance of various video compression algorithms. Various embodiments combine the energy compaction features of transform coding with localization properties of spatial coding. In various embodiments, the effects of pixel "outliers" are removed from the transform and are coded separately as a spatially coded pixel prediction, thereby improving the coding gain of the transform.

WO 2008/126019 A2

INTERNATIONAL SEARCH REPORT

International application No
PCT/IB2008/051351

A. CLASSIFICATION OF SUBJECT MATTER
INV. H04N7/50

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
H04N G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, INSPEC

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SANTAGO P ET AL: "USING CONVEX SET TECHNIQUES FOR COMBINED PIXEL AND FREQUENCY DOMAINCODING OF TIME-VARYING IMAGES" IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, IEEE SERVICE CENTER, PISCATAWAY, US, vol. SAC-05, no. 7, 1 August 1987 (1987-08-01), pages 1127-1139, XP001053635 ISSN: 0733-8716 paragraph [VII. DESCRIPTION OF THE IMPLEMENTED CODER] ----- -/--	1-39

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
29 September 2008	07/10/2008

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Wahrenberg, Annika

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/IB2008/051351

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	SMITH J R ET AL: "Joint adaptive space and frequency basis selection" IMAGE PROCESSING, 1997. PROCEEDINGS., INTERNATIONAL CONFERENCE ON SANTA BARBARA, CA, USA 26-29 OCT. 1997, LOS ALAMITOS, CA, USA,IEEE COMPUT. SOC, US, vol. 3, 26 October 1997 (1997-10-26), pages 702-705, XP010253684 ISBN: 978-0-8186-8183-7 the whole document	1-39
A	CLARKE R J: "ON TRANSFORM CODING OF MOTION-COMPENSATED DIFFERENCE IMAGES" IEE PROCEEDINGS I. SOLID- STATE & ELECTRON DEVICES, INSTITUTION OF ELECTRICAL ENGINEERS. STEVENAGE, GB, vol. 139, no. 3 PART 01, 1 June 1992 (1992-06-01), pages 372-376, XP000305103 ISSN: 0956-3776 the whole document	1-39
A	WIEGAND T ET AL: "Overview of the H.264/AVC video coding standard" IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, IEEE SERVICE CENTER, PISCATAWAY, NJ, US, vol. 13, no. 7, 1 July 2003 (2003-07-01), pages 560-576, XP011099249 ISSN: 1051-8215 the whole document	1-39
A	SEGALL C A ET AL: "Pre- and post-processing algorithms for compressed video enhancement" SIGNALS, SYSTEMS AND COMPUTERS, 2000. CONFERENCE RECORD OF THE THIRTY- FOURTH ASILOMAR CONFERENCE ON OCT. 29 - NOV. 1, 2000, PISCATAWAY, NJ, USA,IEEE, vol. 2, 29 October 2000 (2000-10-29), pages 1369-1373, XP010535224 ISBN: 978-0-7803-6514-8 the whole document	1-39

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000057

May 8, 2012

SIM & MCBURNEY
6th Floor
330 University Avenue
TORONTO Ontario
M5G 1R7

Application No.	:	**2,684,110**
Owner	:	NOKIA CORPORATION
Title	:	**SPATIALLY ENHANCED TRANSFORM CODING**
Classification	:	H04N-7/50 (2006.01)
Your File No.	:	**14010-27 DAR**
Examiner	:	Terry Cartile

YOU ARE HEREBY NOTIFIED OF A REQUISITION BY THE EXAMINER IN ACCORDANCE WITH SUBSECTION 30(2) OF THE PATENT RULES. IN ORDER TO AVOID ABANDONMENT UNDER PARAGRAPH 73(1)(A) OF THE PATENT ACT, A WRITTEN REPLY MUST BE RECEIVED WITHIN 6 MONTHS AFTER THE ABOVE DATE.

This application has been examined as filed at National Entry.

The number of claims in this application is 39.

Documents Cited:

D1: ¤ SANTAGO et al., "Using Convex Set Techniques for Combined Pixel and Frequency Domain Coding of Time-Varying Images", *IEEE Journal on Selected Areas in Communications, IEEE Service Center, Piscataway, USA*, Vol. SAC-05, No. 7, Pages 1127-1139, 1 August 1987 (01-08-1987), XP001053635

D2: ¤ SEGALL et al., "Pre- and Post-Processing Algorithms for Compressed Video Enhancement", *34th IEEE Asilomar Conference on Signals, Systems and Computers, Piscataway, USA*, Vol. 2, Pages 1369-1373, 29 October 2000 (29-10-2000), XP010535224

¤ document cited by the International Preliminary Report on Patentability

Lack of Novelty

Claims 1, 13, 22 and 31 do not comply with *Section 28.2 of the Patent Act.* The examiner considers that the analysis presented in the International Preliminary Report on Patentability (IPRP) in regard to *"Novelty"*, with respect to D1, is applicable under the Canadian standard for

anticipation. In view of the lack of amendment to the claims subsequent to the IPRP, the deficiencies cited in regard to *"Novelty"* during international examination are hereby identified, with respect to the claims identified in the IPRP, as defects for the purposes of national examination.

Obviousness

Claims 2-12, 14-21, 23-30 and 32-39 do not comply with *Section 28.3 of the Patent Act.* The examiner considers that the analysis presented in the International Preliminary Report on Patentability (IPRP) in regard to *"Inventive Step"* is applicable under the Canadian standard for obviousness. In view of the lack of amendment to the claims since the IPRP, the deficiencies cited in regard to *"Inventive Step"*, with respect to D1 and D2, during international examination are hereby identified, with respect to the claims identified in the IPRP, as defects for the purposes of national examination.

Indefiniteness

Claims 1-4, 6, 9, 13-16, 18, 21, 22, 24-29, 31 and 33-39 are indefinite and do not comply with *Subsection 27(4) of the Patent Act.* In each of these claims, the terms *"spatial coding"* and *"spatial samples"* are used. These terms are not clear, in that their meanings are broader than that used in the specification. Perhaps what was intended instead are the terms *"spatial domain coding"* and *"spatial domain samples"*.

In addition, **claim 19** is indefinite and does not comply with *Subsection 27(4) of the Patent Act.* The terms *"outlier values"* and *"non-outlier values"* have no antecedents.

In view of the foregoing defects, the applicant is requisitioned, under *Subsection 30(2) of the Patent Rules,* to amend the application in order to comply with the *Patent Act and the Patent Rules,* or to provide arguments as to why the application does comply.

Under *Section 34 of the Patent Rules,* any amendment made in response to this requisition must be accompanied by a statement explaining the nature thereof, and how it overcomes each of the above objections.

Terry Cartile
Senior Patent Examiner - Section E8 - Electrical Division
(819) 997-2951
2684110A.TCC

As per CIPO Client Service Standards, a response to a telephone enquiry or voice mail should be provided by the end of the next business day. In the event that attempts to reach the examiner are unsuccessful, the examiner's Section Head, *Nicholas Tsagaris*, can be reached at *(819) 953-0790.*

008

IN THE CANADIAN PATENT OFFICE
PCT NATIONAL PHASE

Applicant: Nokia Corporation

Serial No.: 2,684,110

Title: **SPATIALLY ENHANCED TRANSFORM CODING**

Filed: April 9, 2008

Our Ref.: 14010-27 DAR

Industry Canada Industrie Canada
A/M/J ... Y/M/D
2012/11/08
314- 12
CIPO OPIC A002085590

November 8, 2012

The Commissioner of Patents
CANADIAN INTELLECTUAL PROPERTY OFFICE
Place du Portage I
50 Victoria Street
Gatineau, Quebec
K1A 0C9

Dear Commissioner:

RESPONSE AND AMENDMENT

This is in response to the outstanding Official Action dated May 8, 2012 issued in respect of the above-identified application. Please amend the application in the following manner:

IN THE DISCLOSURE:

Please delete page 4 presently on file and insert therefor, replacement pages 4 and 4a submitted herewith.

IN THE CLAIMS:

Please delete claims 1 to 39 presently on file and insert therefor, replacement claims 1 to 29 submitted herewith.

REMARKS

The disclosure has been amended to provide statements of invention corresponding with the independent claims submitted herewith.

The claims have been amended to address the Examiner's objections under Section 27(4) of the *Patent Act*, and are believed to distinguish patentably over the cited prior art.

In the Official Action, the Examiner has referenced the International Preliminary Report on Patentability (IPRP) issued in respect of the parent PCT application. The Examiner has rejected former claims 1, 13, 22 and 31 under Section 28.2 of the *Patent Act* in view of the publication entitled "Using convex set techniques for combined pixel and frequency domain coding of time-varying images" authored by Santago et al. ("D1"). The Examiner has rejected former claims 1 to 39 under Section 28.3 of the *Patent Act* in view of D1 and the publication entitled "Pre- and post-processing algorithms for compressed video enhancement" authored by Segall et al. ("D2"). Applicant respectfully submits that the Examiner's rejections in view of the cited references are without merit at least for the reasons set forth below.

Independent claim 1 as amended now recites a method of encoding a prediction error signal for a block of data, comprising calculating a difference signal representing differences between sample values of a predicted block of data and values for an original input block; creating a modified prediction error signal by substituting outlier values in the difference signal with non-outlier values; performing transform coding to the modified prediction error signal, thereby creating a first representation of a first component of the difference signal; performing spatial domain coding to the difference between the difference signal and the transform decoded modified error signal, thereby creating a second representation of the first component of the difference signal; and joining the first and second representations.

D1 discloses an image coding technique, in which pixel and frequency domain information is combined in order to improve image quality over standard transform coding at given bit rates. The use of non-intersecting convex sets is shown to be optimum for two sets and a given error measure.

D2 provides a review of pre- and post-processing techniques for video compression together with compression enhancement algorithms.

Contrary to the Examiner's allegation, D1 does not anticipate the Applicant's invention as claimed. D1 does not show, disclose, teach or suggest a method of encoding a prediction error signal for a block of data, comprising, *inter alia*, creating a modified prediction error signal by substituting outlier values in the difference signal with non-outlier values, and performing spatial domain coding to the difference between the difference signal and the transform decoded modified error signal, thereby creating a second representation of the first component of the difference signal, as recited. Accordingly, Applicant respectfully submits that independent claim 1 distinguishes patentably over D1.

D2 does not remedy the above-noted deficiencies of D1. D2 does not show, teach or suggest a method of encoding a prediction error signal for a block of data, comprising, *inter alia*, performing spatial domain coding to the difference between the difference signal and the transform decoded modified error signal, thereby creating a second representation of the first component of the difference signal, as recited (emphasis added). Additionally, contrary to the allegation in paragraph 4 of the IPRP, Applicant finds no evidence that D2 discloses "creating a modified prediction error signal by substituting outlier values in the difference signal with non-outlier values". In the IPRP, it is alleged that "Pre-filtering the prediction error signal to remove outlier samples is well known in the art" (see paragraph 4.1). However, Applicant respectfully submits that even though paragraph 4.2 discloses that pre-filtering can be applied in order to change the values of the predicted residual, this does **NOT** correspond to the feature of "creating a modified prediction error signal by substituting outlier values in the difference signal with non-outlier values." Combining the teachings of D1 and D2 therefore would not lead one of ordinary skill in the art to the Applicant's invention as claimed. Accordingly, Applicant respectfully submits that independent claim 1, and the claims dependent therefrom, distinguish patentably over D1 and D2 and should be allowed.

Independent claims 10, 16, 23 and 29, and the claims dependent therefrom, are similarly believed to distinguish patentably over the cited references and should also be allowed.

In view of the above, it is believed the application is in order for allowance and action to that end is respectfully requested.

Respectfully submitted,
NOKIA CORPORATION

Sim & McBurney
Agents of Record

DAR/EWV/shd
Enclosures

(416) 849-8390

of the selected transform may comprise an orthogonal set of basis vectors, or the basis functions may not comprise an orthogonal set. According to various embodiments, the prediction error signal for a single image block is constructed using both transform basis functions and spatial samples (i.e., pixel values), thereby combining the desired features of both transform and spatial coding approaches discussed previously. This allows for the utilization of those selected transform basis functions that give good overall representation of the image block with minimal amount of transform coefficients (representing the component of prediction error signal that is well correlated with the basis functions). Additionally, various embodiments of the present invention allow for the efficient spatial representation those components of the prediction error signal of the same image block that are not well correlated with the basis functions of the applied transform (such as certain types of sensor noise, high frequency texture and edge information).

[0012] Accordingly, in one aspect there is provided a method of encoding a prediction error signal for a block of data, comprising: calculating a difference signal representing differences between sample values of a predicted block of data and values for an original input block; creating a modified prediction error signal by substituting outlier values in the difference signal with non-outlier values; performing transform coding to the modified prediction error signal, thereby creating a first representation of a first component of the difference signal; performing spatial domain coding to the difference between the difference signal and the transform decoded modified error signal, thereby creating a second representation of the first component of the difference signal; and joining the first and second representations.

[0013] According to another aspect there is provided an apparatus, comprising: a processor; and a memory unit communicatively connected to a controller and comprising: computer code for calculating a difference signal representing differences between sample values of a predicted block of data and values for an original input block; computer code for creating a modified prediction error signal by substituting outlier values in the difference signal with non-outlier values; computer code for performing transform coding to the difference signal, thereby creating a first representation of a first component of the difference signal; computer code for performing spatial domain coding to the difference between the difference signal and the transform decoded modified error signal, thereby creating a second representation of the first component of the difference signal; and computer code for joining the first and second representations.

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000064

[0013a] According to yet another aspect there is provided a method of decoding a prediction error signal for a block of data, comprising: receiving a coded prediction error signal, the coded prediction error signal including a plurality of transform coefficients and a plurality of spatial domain samples; decoding the plurality of transformed coefficients into decoded transform information; decoding the plurality of spatial domain samples into decoded spatial domain information; adding the decoded transform information, the decoded spatial domain information, and a reconstructed prediction of the block of data, thereby forming a decoded representation of the block of data; and providing an indication that at least one of no transform coefficients exist in the block of data and no spatial domain samples exist in the block of data.

[0013b] According still yet another aspect there is provided an apparatus, comprising: a processor; and a memory unit communicatively connected to the processor and comprising: computer code for processing a received coded prediction error signal, the coded prediction error signal including a plurality of transformed coefficients and a plurality of spatial domain samples; computer code for decoding the plurality of transformed coefficients into decoded transform information; computer code for decoding the plurality of spatial domain samples into decoded spatial domain information; computer code for adding the decoded transform information, the decoded spatial domain information, and a reconstructed prediction of the block of data, thereby forming a decoded representation of the block of data; and computer code for providing an indication that at least one of no transform coefficients exist in the block of data and no spatial domain samples exist in the block of data.

[0013c] According to still yet another aspect there is provided a method of encoding a prediction error signal for a block of data, comprising: calculating a difference signal representing differences between sample values of a predicted block of data and values for an original input block; substituting outlier values in the difference signal with non-outlier values to create a modified prediction error signal; performing one of transform coding and spatial domain coding to the modified prediction error signal, creating one of a transform coded data and spatial domain coded data; performing an inverse of the one of transform coding and spatial domain coding; applying the other of transform coding and spatial domain coding to a residual from the inverse transform or spatial domain decoding, creating the other of transform coded data and spatial domain coded data; and joining the transform coded data and the spatial domain coded data.

[0014] The implementations of various embodiments of the present invention serve to improve the compression efficiency of modern video and image codecs. Although

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000065

What is claimed is:

1. A method of encoding a prediction error signal for a block of data, comprising:

calculating a difference signal representing differences between sample values of a predicted block of data and values for an original input block;

creating a modified prediction error signal by substituting outlier values in the difference signal with non-outlier values;

performing transform coding to the modified prediction error signal, thereby creating a first representation of a first component of the difference signal;

performing spatial domain coding to the difference between the difference signal and the transform decoded modified error signal, thereby creating a second representation of the first component of the difference signal; and

joining the first and second representations.

2. The method of claim 1, wherein, for transform coding, at least one transform coefficient is set to zero.

3. The method of claim 1 or 2, wherein, for spatial domain coding, at least one spatial domain sample is set to zero.

4. The method of any one of claims 1 to 3, wherein the non-outlier values used to replace the outlier values comprise an average of neighboring prediction error values.

5. The method of any one of claims 1 to 4, further comprising providing an indication that no transform coefficients exist in the block of data if appropriate.

6. The method of any one of claims 1 to 4, further comprising providing an indication that no spatial domain samples exist in the block of data if appropriate.

7. The method of claim 1, wherein for the transform coding, a transform coefficient is formed, the transform coefficient comprising a weighting coefficient of a discrete transform basis vector.

8. The method of claim 7, wherein the discrete transform comprises a discrete orthogonal transform.

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000066

9. A computer-readable medium storing computer code configured to perform the method of any one of claims 1 to 8.

10. An apparatus, comprising:

a processor; and

a memory unit communicatively connected to a controller and comprising:

computer code for calculating a difference signal representing differences between sample values of a predicted block of data and values for an original input block;

computer code for creating a modified prediction error signal by substituting outlier values in the difference signal with non-outlier values;

computer code for performing transform coding to the difference signal, thereby creating a first representation of a first component of the difference signal;

computer code for performing spatial domain coding to the difference between the difference signal and the transform decoded modified error signal, thereby creating a second representation of the first component of the difference signal; and

computer code for joining the first and second representations.

11. The apparatus of claim 10, wherein, for transform coding, at least one transform coefficient is set to zero.

12. The apparatus of claim 10 or 11, wherein, for spatial domain coding, at least one spatial domain sample is set to zero.

13. The apparatus of any one of claims 10 to 12, wherein the non-outlier values used to replace the outlier values comprise an average of neighboring prediction error values.

14. The apparatus of any one of claims 10 to 13, further comprising providing an indication that no transform coefficients exist in the block of data if appropriate.

15. The apparatus of any one of claims 10 to 13, further comprising providing an indication that no spatial domain samples exist in the block of data if appropriate.

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000067

16. A method of decoding a prediction error signal for a block of data, comprising:

receiving a coded prediction error signal, the coded prediction error signal including a plurality of transform coefficients and a plurality of spatial domain samples;

decoding the plurality of transformed coefficients into decoded transform information;

decoding the plurality of spatial domain samples into decoded spatial domain information;

adding the decoded transform information, the decoded spatial domain information, and a reconstructed prediction of the block of data, thereby forming a decoded representation of the block of data; and

providing an indication that at least one of no transform coefficients exist in the block of data and no spatial domain samples exist in the block of data.

17. The method of claim 16, wherein the decoding of the plurality of transformed coefficients and spatial domain samples includes dequantization of each, and wherein the dequantization of transformed coefficients and spatial domain samples are related to each other.

18. The method of claim 16, wherein the decoding of the plurality of transformed coefficients and spatial domain samples includes dequantization of each, and wherein the dequantization of transformed coefficients and spatial domain samples are different from each other.

19. The method of claim 16, wherein the decoding of the spatial domain samples is dependent upon the decoding of the transformed coefficients.

20. The method of claim 16, wherein the decoding of the transformed coefficients is dependent upon the decoding of the spatial domain samples.

21. The method of claim 16, wherein at least some of the plurality of spatial domain samples are decoded as a single unit.

22. A computer-readable medium storing computer code configured to perform the method of any one of claims 16 to 21.

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000068

23. An apparatus, comprising:

a processor; and

a memory unit communicatively connected to the processor and comprising:

computer code for processing a received coded prediction error signal, the coded prediction error signal including a plurality of transformed coefficients and a plurality of spatial domain samples;

computer code for decoding the plurality of transformed coefficients into decoded transform information;

computer code for decoding the plurality of spatial domain samples into decoded spatial domain information;

computer code for adding the decoded transform information, the decoded spatial domain information, and a reconstructed prediction of the block of data, thereby forming a decoded representation of the block of data; and

computer code for providing an indication that at least one of no transform coefficients exist in the block of data and no spatial domain samples exist in the block of data.

24. The apparatus of claim 23, wherein the decoding of the plurality of transformed coefficients and spatial domain samples includes dequantization of each, and wherein the dequantization of transformed coefficients and spatial domain samples are related to each other.

25. The apparatus of claim 23, wherein the decoding of the plurality of transformed coefficients and spatial domain samples includes dequantization of each, and wherein the dequantization of transformed coefficients and spatial domain samples are different from each other.

26. The apparatus of claim 23, wherein the decoding of the spatial domain samples is dependent upon the decoding of the transformed coefficients.

27. The apparatus of claim 23, wherein the decoding of the transformed coefficients is dependent upon the decoding of the spatial domain samples.

28. The apparatus of any one of claims 23 to 27, wherein at least some of the plurality of spatial domain samples are decoded as a single unit.

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000069

29. A method of encoding a prediction error signal for a block of data, comprising:

calculating a difference signal representing differences between sample values of a predicted block of data and values for an original input block;

substituting outlier values in the difference signal with non-outlier values to create a modified prediction error signal;

performing one of transform coding and spatial domain coding to the modified prediction error signal, creating one of a transform coded data and spatial domain coded data;

performing an inverse of the one of transform coding and spatial domain coding;

applying the other of transform coding and spatial domain coding to a residual from the inverse transform or spatial domain decoding, creating the other of transform coded data and spatial domain coded data; and

joining the transform coded data and the spatial domain coded data.

Amazon.com, Inc., Amazon.com Services LLC - Ex. 1009, Page 000070

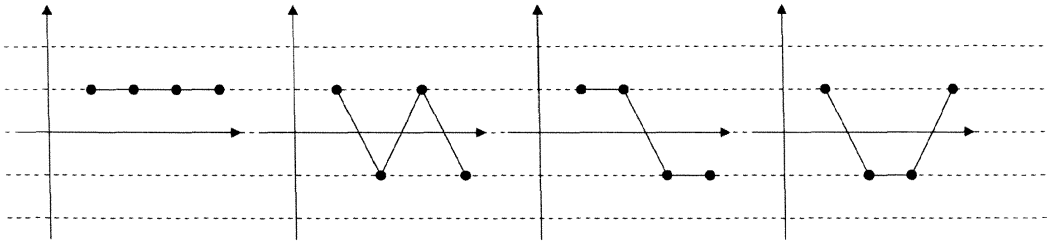

FIG. 1