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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
3 January 2013 (03.01.2013)

WIPO | PCT

(10) International Publication Number
WO 2013/001278 A1

(51) International Patent Classification:
H04N 7/26 (2006.01) *H04N 7/50* (2006.01)

(21) International Application Number:
PCT/GB2012/051412

(22) International Filing Date:
19 June 2012 (19.06.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
1110873.5 27 June 2011 (27.06.2011) GB

(71) Applicant *(for all designated States except US)*: **BRITISH BROADCASTING CORPORATION** [GB/GB]; Broadcasting House, London W1A 1AA (GB).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **MRAK, Marta** [HR/GB]; 42 Templewood, London W13 8DX (GB). **GABRIELLINI, Andrea** [IT/GB]; 14 St. Johns Road, Guilford GU2 7UH (GB). **SPRLJAN, Nikola** [HR/GB]; 42 Templewood, London W13 8DX (GB). **FLYNN, David** [GB/GB]; 48 Leslie Park Road, Croydon Surrey CR0 6TP (GB).

(74) Agent: **GARRATT, Peter Douglas**; Mathys & Squire LLP, 120 Holborn, London EC1N 2SQ (GB).

(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, CL, 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, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, 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, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

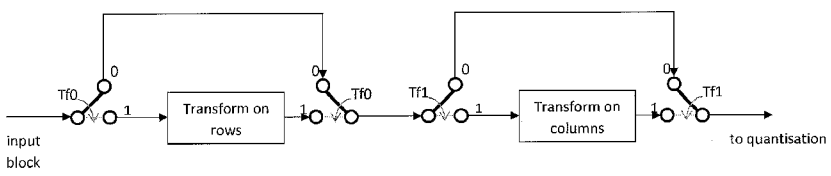
Published:

— *with international search report (Art. 21(3))*

WO 2013/001278 A1

(54) **Title**: VIDEO ENCODING AND DECODING USING TRANSFORMS

Figure 1: Encoder



(57) **Abstract**: Video encoding or decoding utilising a spatial transform operating on rows and columns of a block, with a set of transform skip modes including: transform on rows and columns; transform on rows only;transform on columns only; no transform. An indication of the selected mode is provided to the decoder.

VIDEO ENCODING AND DECODING USING TRANSFORMS

FIELD OF THE INVENTION

This invention is related to video compression and decompression systems, and in particular to a framework to adaptively model signal representation between prediction and entropy coding, by the adaptive use of transform functions and related tools, including scaling, quantisation, scanning, and signalling.

BACKGROUND OF THE INVENTION

Transmission and storage of video sequences are employed in several applications like e.g. TV broadcasts, internet video streaming services and video conferencing.

Video sequences in a raw format require a very large amount of data to be represented, as each second of a sequence may consist of tens of individual frames and each frame is represented by typically at least 8 bit per pixel, with each frame requiring several hundreds or thousands of pixels. In order to minimise the transmission and storage costs video compression is used on the raw video data. The aim is to represent the original information with as little capacity as possible, i.e., with as few bits as possible. The reduction of the capacity needed to represent a video sequence will affect the video quality of the compressed sequence, i.e. its similarity to the original uncompressed video sequence.

State-of-the-art video encoders, such as AVC/H.264, utilise four main processes to achieve the maximum level of video compression while achieving a desired level of video quality for the compressed video sequence: prediction, transformation, quantisation and entropy coding. The prediction process exploits the temporal and spatial redundancy found in video sequences to greatly reduce the capacity required to represent the data. The mechanism used to predict data is known to both encoder and decoder, thus only an error signal, or residual, must be sent to the decoder to reconstruct the original signal. This process is typically performed on blocks of data (e.g. 8x8 pixels) rather than entire frames. The prediction is typically performed against already reconstructed frames or blocks of reconstructed pixels belonging to the same frame.

The transformation process aims to exploit the correlation present in the residual signals. It does so by concentrating the energy of the signal into few coefficients. Thus the transform coefficients typically require fewer bits to be represented than the pixels of the residual. H.264 uses 4x4 and 8x8 integer type transforms based on the Discrete Cosine Transform (DCT).

The capacity required to represent the data in output of the transformation process may still be too high for many applications. Moreover, it is not possible to modify the transformation process in order to achieve the desired level of capacity for the compressed signal. The quantisation process takes care of that, by allowing a further reduction of the capacity needed to represent the signal. It should be noted that this process is destructive, i.e. the reconstructed sequence will look different to the original

The entropy coding process takes all the non-zero quantised transform coefficients and processes them to be efficiently represented into a stream of bits. This requires reading, or scanning, the transform coefficients in a certain order to minimise the capacity required to represent the compressed video sequence.

The above description applies to a video encoder; a video decoder will perform all of the above processes in roughly reverse order. In particular, the transformation process on the decoder side will require the use of the inverse of the transform being used on the encoder. Similarly, entropy coding becomes entropy decoding and the quantisation process becomes inverse scaling. The prediction process is typically performed in the same exact fashion on both encoder and decoder.

The present invention relates to the transformation part of the coding, thus a more thorough review of the transform process is presented here.

The statistical properties of the residual affect the ability of the transform (i.e. DCT) to compress the energy of the input signal in a small number of coefficients. The residual shows very different statistical properties depending on the quality of the prediction and whether the prediction exploits spatial or temporal redundancy. Other factors affecting the quality of the prediction are the size of the blocks being used and the spatial / temporal characteristics of the sequence being processed.

It is well known that the DCT approaches maximum energy compaction performance for highly correlated Markov-I signals. DCT's energy compaction performance starts dropping as the signal correlation becomes weaker. For

instance, it is possible to show how the Discrete Sine Transform (DST) can outperform the DCT for input signals with lower adjacent correlation characteristics.

The DCT and DST in image and video coding are normally used on blocks, i.e. 2D signals; this means that a one dimensional transform is first performed in one direction (e.g., horizontal) followed by a one dimensional transform performed in the other direction. As already mentioned the energy compaction ability of a transform is dependent on the statistics of the input signal. It is possible, and indeed it is also common under some circumstances, for the two-dimensional signal input to the transform to display different statistics along the two vertical and horizontal axes. In this case it would be desirable to choose the best performing transform on each axis. A similar approach has already been attempted within the new ISO and ITU video coding standard under development, High Efficiency Video Coding (HEVC). In particular, a combination of two one dimensional separable transforms such as a DCT-like [2] and DST [3] has been used in HEVC standard under development.

While previous coding standards based on DCT use a two-dimensional transform (2D DCT), newer solutions apply a combination of DCT and DST to intra predicted blocks, i.e. on blocks that are spatially predicted. It has been shown that DST is a better choice than DCT for transformation of rows, when the directional prediction is from a direction that is closer to horizontal then vertical, and, similarly, is a better choice for transformation of columns when the directional prediction is closer to vertical. In the remaining direction (e.g. on rows, when DST is applied on columns), DCT is used.

For implementation purposes, in video coding it is common to use integer approximations of DCT and DST, which will in rest of this text be simply referred to as DCT and DST. One of solutions for integer DCT-like transform uses 16-bit intermediate data representation and is known as partial butterfly. Its main properties are same (anti)symmetry properties as of DCT, almost orthogonal basis vectors, 16 bit data representation before and after each transform stage, 16 bit multipliers for all internal multiplications and no need for correction of different norms of basis vectors during (de)quantisation.

SUMMARY OF THE INVENTION

The present invention consists in, in one aspect, a method of video encoding utilising a spatial transform operating on rows and columns of a block, comprising the steps of establishing a set of transform skip modes including:

transform on rows and columns;
transform on rows only;
transform on columns only;
no transform;

selecting one of the said modes; and providing an indication of the selected mode for a decoder.

There are described in the following a transform mode and a system to apply a combination of transforms minimising the capacity required to represent a signal for a given output signal quality target. Moreover, a system to signal the selected combination of transform modes is presented.

DETAILED DESCRIPTION OF THE PRESENT INVENTION

The present invention will now be described by way of example with reference to the accompanying drawings, in which:

Figure 1 is a block diagram illustrating a feature on an encoder according to an embodiment of the invention;

Figure 2 is a block diagram illustrating the feature on a decoder according to the embodiment;

Figure 3 is a diagram illustrating an alternative to the known zig-zag scanning approach;

Figure 4 is a diagram illustrating a further alternative scanning approach;

Figure 5 is a block diagram illustrating a feature on an encoder according to a further embodiment of the invention;

Figure 6 is a block diagram illustrating the feature on a decoder according to the embodiment;

Figure 7 is a block diagram illustrating a feature on a decoder according to a further embodiment of the invention

This invention presents a mode to perform the transformation process - Transform Skip Mode (TSM). As described above, the most common transform used in video coding is the DCT. Its energy compacting performance depends on the correlation of the residual. It has also been described how the residual can be highly decorrelated, or correlated in one direction only, making the 2D DCT less efficient. It is proposed to skip the transformation process when the encoder makes such decision in a rate-distortion sense. The selected transform mode must be signalled to the decoder, which then performs a combination of transform/ skip transform as defined in signalling.

Four transform modes are defined as shown in Table 1.

TSM	Transform on rows	Transform on columns	Note
TS0	+	+	2D transform
TS1	+	-	1D transform
TS2	-	+	1D transform
TS3	-	-	no transform

Table 1 - Transform Skip Mode options

TS0 mode corresponds to 2D transform, i.e. 2D DCT. TS1 mode defines application of one dimensional horizontal DCT followed by a transform skip in the orthogonal direction, i.e. transform of columns is skipped. TS2 defines skipping of horizontal transform, while only columns are transformed. Finally, TS3 mode completely skips transforms in both axes, i.e., no transform is applied to the input signal.

Figures 1 and 2 show core transform skip mode block diagrams, for encoder and decoder, respectively. Each transform skip mode is selected with corresponding (Tf0, Tf1) pair of flags, such that TS0: (1, 1), TS1: (1, 0), TS2: (0, 1) and TS3: (0, 0).

As for any other additional bits from a compressed bit-stream that enable adaptive option, signalling of the transform skip mode can be costly. Therefore several strategies are devised to maximise the coding efficiency.

Four TSM options can be signalled using carefully designed code words. Those code words do not need to be transmitted for each block, but some other methods can be used to save necessary bit-rate.

Some of possibilities for reducing the signalling cost are listed in the following; each option influencing transform-related parts of the encoder and decoder:

1. The same transform mode used on all components (luminance - Y and chrominance - U and V) of a YUV block; therefore, for Y, U and V collocated blocks only one TSM choice is transmitted.
2. TSM not signalled when all quantised blocks (Y, U and V) have only coefficients with zero values.
3. TSM not signalled for blocks when Y block has only zero-value coefficients, and then 2D DCT is used on U and V components.
4. TSM signalled only for blocks with certain other modes (e.g. bidirectional predicted); otherwise 2D-DCT is applied.
5. Application of TSM signalled on a set of blocks (if "on" then TS modes signalled for each block from the set).
6. TSM signalled on a set of blocks (e.g. all sub-blocks share the same TSM).
7. TSM signalled if certain other block characteristics are present; e.g. TSM not signalled when Y block has only one non-zero value, and that value is in top-left corner of the block (DC component); in that case 2D-DCT is used for all components.

Four TSM modes (2D transform, two 1D block transforms and skipped transform on a block) can be defined with various code words, e.g. with simple 2-bit words, or with more bits (i.e. with unary codes):

TSM	2 bit signalling	Unary code
TS0	11	1
TS1	10	01
TS2	01	001
TS3	00	000

If arithmetic coding is used, each bin of the code word can be encoded with different probability models (i.e. initial context states for each slice), depending on the current block size and on QP value.

On the other hand, if variable length coding is used, TSM code words can be encoded independently of or merged with other syntax elements, to reduce the signalling overhead.

In some approaches, a block is not always transformed at once, but rather options for its partitioning into smaller sub-units are applied, and transforms are applied on each sub-units. Representative of such transform structure is Residual QuadTree (RQT) method. While application of TSM on blocks that are not further divided into smaller unit has been assumed so far, TSM can also be applied on such multi-split transform structures. Several options are identified:

1. TSM is decided on a block level, and the same transform choice is applied on each sub-unit.
2. TSM is enabled only at the root level of transformation structure, i.e. when a block is not further partitioned into smaller units when a multi-split structure is enabled; if a block is split into smaller units, each unit is transformed using 2D transform.
3. TSM is decided and signalled for each sub-unit, independently of its depth.
4. TSM is decided and signalled for sub-units, up to specific depth (size) of units; for lower sub-units, when TSB is not signalled, 2D transform is used.

Coefficients within a block can have different characteristics when the transform is not performed in one or both directions. Therefore different coding strategies can be applied, depending on the transform skip mode, to better compress given coefficients.

When a 2D transform is applied on a block, the resulting coefficients are often grouped towards top-left corner of a block, that is to say they are low-frequency components. Conventional scanning, e.g. zig-zag scanning, is therefore a good choice for coding of such signals.

If only 1D transform is applied (TS1 or TS2), adaptive scanning can be used. For example, a row-by-row, or a column-by-column scanning can be used for TS2 and TS1 cases respectively, since one can expect that applied transform concentrates the coefficients towards lower frequencies.

For the TS3 case, where a transform is not applied in any direction, a conventional scan (used for 2D transformed block) scan may be used. Alternatively, a different scanning pattern may be employed which takes into account the probability (implicit in the decision to conduct no transform) that non-zero coefficients are not uniformly distributed. For example, coefficients may be grouped in “islands” surrounded by “seas” of zero coefficients.

Thus, in one new arrangement, positions of the first and the last significant coefficients within a block can be transmitted in the bit-stream, and a conventional scanning of coefficients within a block can then be performed. This is shown in Figure 3 where white squares represent coefficients that are not encoded and have zero value, gray squares represent coefficients that will be encoded, i.e. include significant (non-zero) coefficients), where the first coded coefficient is labelled with "F" and the last encoded coefficient is labelled with "L". Scanning is performed only on rows and columns that belong to area defined by the first and the last coefficient. In this scanning method, x and y coordinates of the first coefficient must be the same or smaller than the x and y coordinates of the last significant coefficient.

This arrangement should lead to highly efficient coding in the case where non-zero coefficients are clustered, but requires the additional complexity in the encoder of determining the positions of the first and the last significant coefficients within a block, together with the need to signal those positions to the decoder.

In an alternative, a double zig-zag scan is used, as depicted in Figure 4, where a block of transform coefficients is represented with sub-blocks of coefficients. Each sub-block is visited in sub-block level zig-zag scan, and inside each block a zig-zag scan (or any other scan) is used. This enables better grouping of non-zero coefficients, which tend to be spatially close.

It will be desirable, where a decision is taken to skip either or both 1D transforms, to minimise or remove the need to change other elements of the process to accommodate the skipped transform or transforms.

Here, two implementation strategies for the adaptive transform stage are identified:

1) skipping selected transform of rows / columns, and modifying quantisation stage.

2) replacing selected transform of rows / columns by suitable scaling step and adapting the quantisation step if required.

While the first strategy is suitably presented with Figures 1 and 2, the second strategy that employs scaling is depicted in Figures 5 and 6. One of the main reasons why scaling is performed is to maintain levels of signal, with the highest supported precision, between transform blocks. This is indicated using dashed line in Figures 5 and 6.

Scaling is performed by scaling each input pixel value by a factor that is derived from norm-2 of corresponding transform vectors (which would be used to obtain a transform coefficient value, at the same position in a row/column, if the transform was selected). Some transforms have close to orthonormal properties of each vector and this property can further simplify the scaling design since a single value can be used to suitably scale whole row/column on which the transform is skipped.

In the following, scaling strategies are discussed in the context of integer DCT transform with 16 bit intermediate data representation. It will be recognised, however, that this is only an example.

Transforms used in HEVC have the norms (TN_N), where N is size of the transform, close to the following numbers:

- 4-point transform: $TN_4 = 128 = 2^7$; $TNS_4 = 7$;
- 8-point transform: $TN_8 = 181 = 2^{7.5}$; $TNS_8 = 7.5$;
- 16-point transform: $TN_{16} = 256 = 2^8$; $TNS_{16} = 8$;
- 32-point transform: $TN_{32} = 362 = 2^{8.5}$. $TNS_{32} = 8.5$;

where TNS is corresponding Transform Norm Shift parameter (power of 2 represented by left bit-shifting). Note that in HEVC each transform vector may have slightly different norm, but these numbers are good approximations for practical implementations. This fact is also reflected in the designs of quantisation and in the transform level adjustment to preserve 16-bit intermediate data representation. For example, in HEVC decoder design, 16-bit value enters inverse transform. In order to reach 16-bit precision between column (1st stage

inverse) and row (2nd stage inverse) transforms, and 9+DB precision after the row transform, the following signal level bit-shifts occur (considering N x N block size):

$$SHIFT = TNS_N - SHIFT_INV_1ST + TNS_N - (SHIFT_INV_2ND - DB),$$

where, by the standard, SHIFT_INV_1ST = 7 and SHIFT_INV_2ND = 12, and DB is the bit-depth increment for processing (e.g. 0 or 2). Internal processing bit-depth is B = 8 + DB. Therefore, SHIFT equals:

$$SHIFT = 2 \cdot TNS_N - 19 + DB = 2 \cdot TNS_N - 27 + B\ .$$

This corresponds to the parameter transform shift used in the HEVC quantisation. This leads to, for the example where 4 x 4 block is considered ($TNS_4 = 7$), to

$$-SHIFT_4 = 13 - B,$$

i.e. right shift by 13 - B.

While this example may be used to address signal level adjustment for TS3, some additional considerations have to be taken into account when the transform is applied in one direction only. That is because TNS_N are not always integer numbers, thus bit-shifting is not the only option for level adjustment. Other options for addressing unified designs for such combinations are addressed in the following text.

Where a transform is replaced with scaling, the adaptive transform stage is designed in a way that it can be interleaved within the integer DCT transform with 16-bit intermediate data representation, i.e. with the goal to replace some of its parts and to be compatible with the rest of the codec that supports original 2D transform. For example, not applying transform can be used on rows in a way which is still compatible with the part of 2D transform that is applied on columns. This means that quantisation applied for 2D transform can also be used with adaptive transform choice.

The forward transform skip is defined for rows and columns separately.

On samples x of rows the transform skip is applied as:

$$y = (x \cdot scale + offset) \text{ right shifted by } S \text{ bits} \qquad (a)$$

where:

$$S = M - 1 + DB$$

offset = 1 left shifted by (S - 1) bits

DB = B - 8 bit-depth increment for processing

M = log2(N), where N is row/column size in the number of pixels, and scale is an unsigned integer multiplier.

On columns, the transform skip is applied as in (a) where x are samples of columns, but with:

S = M + 6

offset = 1 left shifted by (S - 1) bits

In this way a bit-width of 16 after each transform stage is ensured, as in the 2D transform.

Again, scale factors are designed in a way to be near the norm-2 of related transform vectors ($scale_N^2 = TN_N^2 = N \cdot 64^2$) and to be an integer number. On samples x of columns the inverse transform skip is applied as

y = (x · scale + offset) right shifted by S bits

where:

S = 7

offset = 1 left shifted by (S - 1) bits

and scale is the same as in the forward skip.

On rows the same transform skip operation is applied, but with:

S = 12 - DB, where DB is the same as in the forward transform skip.

In order to save unnecessary processing of pixels, where one or both 1D transforms are skipped, scaling can be moved to quantisation. Moreover (for example), if only the vertical transform is kept, it can be adapted, to ensure maximal 16-bit representation of pixels. This enables full use to be made of the available bit width. Therefore, scaling in quantisation has to be adapted not only because of the scaling related to skipped transform but also related to new scaling within a transform.

<u>TSM = TS0 (2D transform)</u>

Regular 2D transform and corresponding quantisation is used.

<u>TSM - TS1 (1D transform on rows) and TS2 (1D transform on columns)</u>

In both cases the forward transform corresponds to the original transform of rows

y = (x + offset) right shifted by S bits, (b)

where:

x is the original value of residual block,

S = M - 1 + DB,

offset = 1 left shifted by (S - 1) bits

and M and DB are the same as in (a).

This ensures 16-bit intermediate data precision.

Quantisation is adapted and takes into account the level at which signal is now.

TSM = TS3 (no transform)

Residual pixels are directly quantised using the flat matrix so that the level of signal corresponds to the levels of quantised coefficients that are 2D transformed and quantised.

Another example of how the level of the signal can be adjusted when a transform is skipped is presented in the following, with reference to Figure 7. In this example the aim is to reduce a number of operations required to achieve desired performance. In that context, where a transform or its parts can be skipped or replaced, this technique uses a combination of one or more basic operations:

1. Changes to bit-shifting within transform stages;
2. Adjustment of quantisation that correspond to the scaling a signal by a factor smaller than 2;
3. Replacement of the transform or its parts by a scalar outside the quantisation.

Each scaling of the signal can be represented by scaling by a factor of 2^N (where N is a positive integer) and by scaling by a factor M that is smaller than 2. Note that in this case N is the transform size as in the previous example). In this invention, Operation 1 enables signal scaling by a factor of 2^N (bit-shifting) and Operation 2 enables scaling by M. The choice of M is typically limited and depends on the quantisation design. A typical component of a 1D transform in video coding is bit-shifting. Therefore Operation 1 applied here readily enables adjustment of a signal level by a factor of 2^N. In the case where both transforms are skipped, adjustment of the level of the signal can be performed in the

"Scaling" block from Figure 7, which corresponds to Operation 3. In any case, adjustment of the signal by a factor smaller than 2, a quantisation parameter offset, or quantisation scaling factor, can be suitably chosen to perform required signal level adjustment. For example, in High Efficiency Video Coding (HEVC), adding an offset of 3 to a quantisation parameter is equivalent to adjusting the level of the signal by sqrt(2) (root 2).

It will be understood that the invention has been described by way of example only and that a wide variety of modifications are possible without departing from the scope of the invention as set forth in the appended claims. Features which are here described in certain combinations may find useful application in other combinations beyond those specifically mentioned and may in certain cases be used alone. For example, the scanning approaches in video coding or decoding where:

positions of the first and the last coefficients to be encoded /decoded within a block are signalled to the decoder and a scanning of coefficients is performed between said first and the last coefficients; or

a double scan is performed, where a block of transform coefficients is represented with sub-blocks of coefficients; each sub-block is visited in sub-block level zig-zag scan, and inside each sub-block additional scan pattern is used;

may be useful beyond the case of transform skip mode.

CLAIMS

1. A method of video encoding utilising a spatial transform operating on rows and columns of a block, comprising the steps of establishing a set of transform skip modes including:

transform on rows and columns;
transform on rows only;
transform on columns only;
no transform;

selecting one of the said modes; and providing an indication of the selected mode for a decoder.

2. A method of decoding video which has been encoded utilising a spatial transform operating on rows and columns of a block with transform skip modes including:

transform on rows and columns;
transform on rows only;
transform on columns only;
no transform;

comprising the steps of providing an indication of the transform skip mode and applying inverse transforms in accordance with the mode.

3. A method according to Claim 1 or Claim 2, wherein mode selection is signalled to the decoder with each mode assigned a codeword.

4. A method according to any one of the preceding claims, where the order in which coefficients within a block are scanned in the entropy coding stage is adapted in accordance with the transform skip mode.

5. A method according to Claim 4, wherein row-by-row scanning is employed where the row transform is skipped and transform of columns is kept, and column-by-column scanning is employed where the column transform is skipped and transform on rows is kept.

6. A method according to any one of the preceding claims, wherein in the entropy coding stage positions of the first and the last coefficients to be encoded /decoded within a block are signalled to the decoder and a scanning of coefficients is performed between said first and the last coefficients.

7. A method according to any one of the preceding claims, wherein, a double scan is performed, where a block of transform coefficients is represented with sub-blocks of coefficients; each sub-block is visited in sub-block level scan, and inside each sub-block a different scan is used.

8. A method according to any one of the preceding claims, wherein the same transform skip mode is used on all components (luminance - Y and chrominance - U and V) of a YUV block.

9. A method according to any one of the preceding claims, wherein the transform skip mode is not signalled for blocks having only zero-value coefficients.

10. A method according to Claim 9, wherein the transform skip mode is not signalled when the luminance component has only zero values; in this case 2D transform is used on chroma components.

11. A method according to Claim 9, wherein the transform skip mode is not signalled when the only non-zero-value coefficient of the luminance component is the top-left corner of the block (DC component) in this case 2D transform is used on chroma components.

12. A method according to any one of the preceding claims, wherein the transform skip mode is signalled only for blocks with predefined other modes (e.g. predicted from other frames only).

13. A method according to any one of the preceding claims, wherein the transform skip mode is signalled on a set of blocks.

14. A method according to any one of the preceding claims, where the transform provides options for its partitioning into smaller sub-units and transforms are applied on each sub-units (for example the Residual QuadTree (RQT) method) and wherein:

the transform skip mode is enabled on a block level, and the same transform mode is applied on each sub-unit; or

transform skip mode is enabled only on the root level of transformation structure; for lower sub-units, when the transform skip mode is disabled, 2D transform is used; or

the transform skip mode is enabled for each sub-unit, independently of its depth; or

the transform skip mode is enabled for sub-units up to a specific depth of units; for lower sub-units, when the transform skip mode is disabled, 2D transform is used; or

transform skip mode is enabled for sub-units under a specific depth of units; for higher sub-units, when the transform skip mode is disabled, 2D transform is used.

15. A method according to any one of the preceding claims, wherein a quantisation stage is adapted according to the selected transform skip mode.

16. A method according to any one of the preceding claims, wherein a quantisation matrix that has the same values in each column is applied when vertical transform is skipped, a quantisation matrix that has the same values in each row is applied when horizontal transform is skipped.

17. A method according to any one of the preceding claims, comprising the step of scaling of coefficients that are not transformed, where the scaling factors are dependent upon the norms of corresponding transform vectors to bring the untransformed coefficients to the same level as transformed coefficients.

18. A method according to Claim 17, wherein the same scaling factors are used for all coefficients in scaled row or column.

19. A method according to any one of the preceding claims, wherein the row transform differs in dependence upon whether or not the column transfer is skipped and wherein the column transform differs in dependence upon whether or not the row transfer is skipped.

20. A method according to Claim 19 where shifting in the remaining 1D transform is adjusted so that the level of signal is adjusted by $D=2^N$, where N is an integer.

21. A method according to Claim 20 where any remaining scaling of $M < 2$ is approximated in quantisation.

.

22. A computer program product containing instructions causing programmable means to implement a method according to any one of the preceding claims.

23. A video encoder adapted and configured to operate in accordance with any one of Claim 1 and any claim when dependent on Claim 1.

24. A video decoder adapted and configured to operate in accordance with any one of Claim 2 and any claim when dependent on Claim 2.

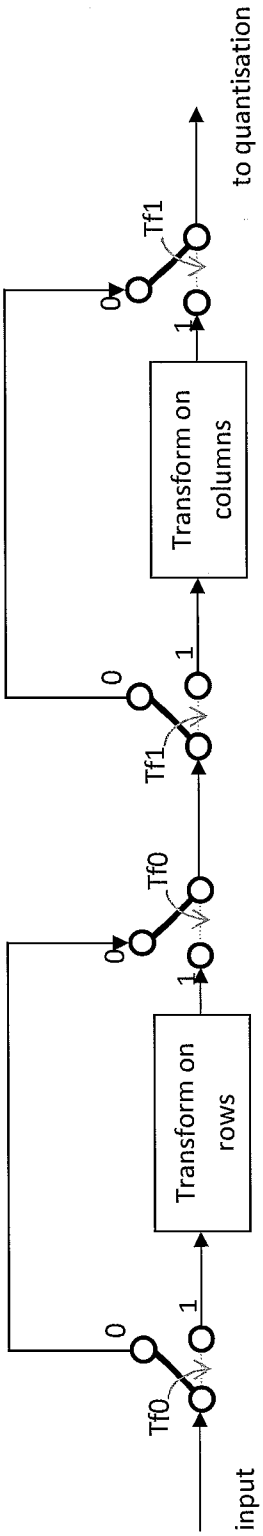


Figure 1: Encoder

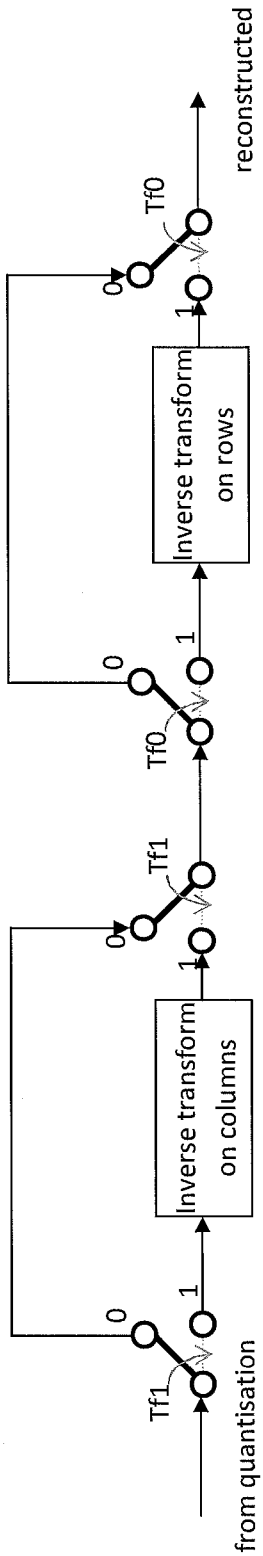


Figure 2: Decoder

Figure 3

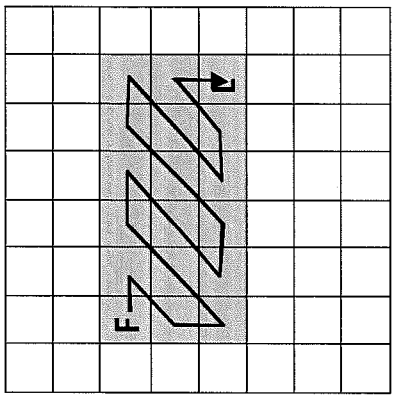
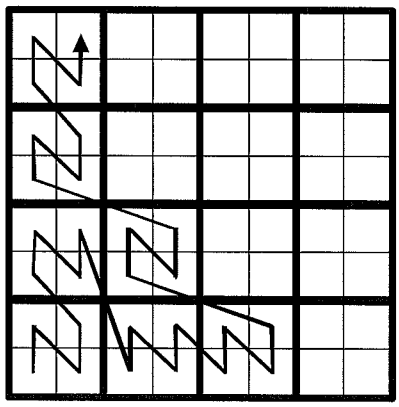


Figure 4



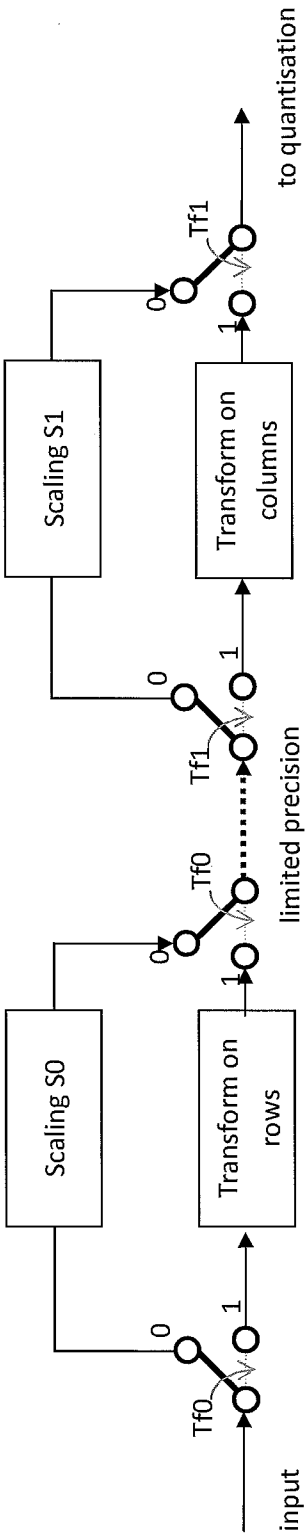


Figure 5: Encoder

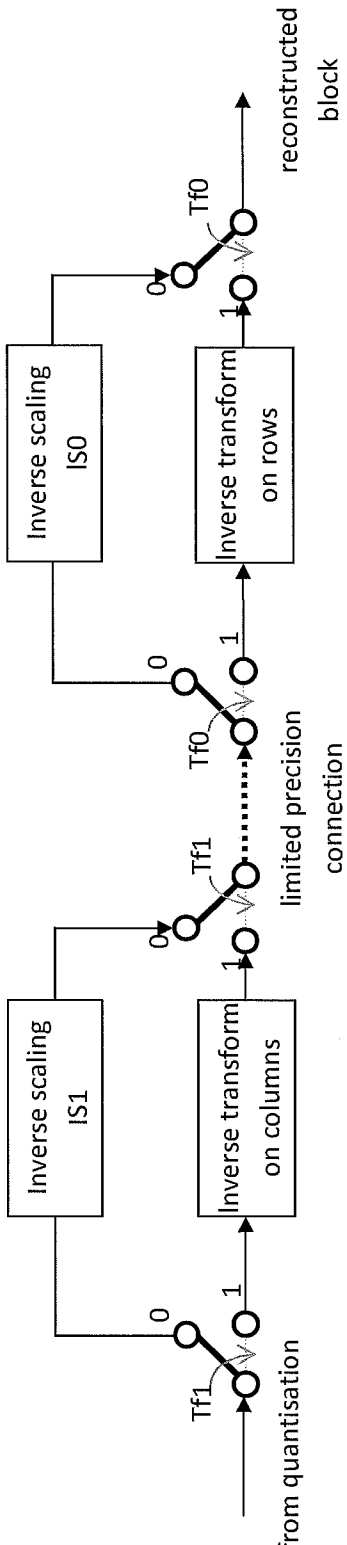


Figure 6: Decoder

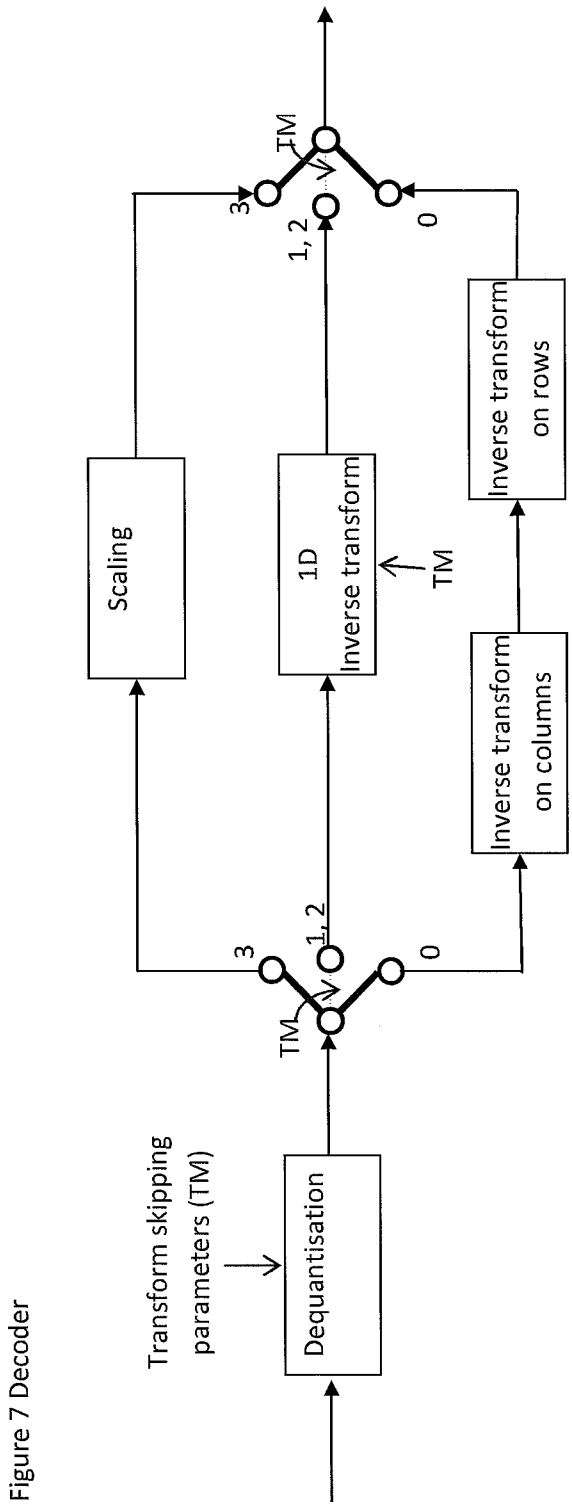


Figure 7 Decoder

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2012/051412

A. CLASSIFICATION OF SUBJECT MATTER
INV. H04N7/26 H04N7/50
ADD.

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

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 practicable, search terms used)

EPO-Internal, COMPENDEX, INSPEC, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	YUMI SOHN ET AL: "One Dimensional Transform For H.264 Based Intra Coding (Abstract)", 26. PICTURE CODING SYMPOSIUM;7-11-2007 - 9-11-2007; LISBON,, 7 November 2007 (2007-11-07), XP030080458, abstract; figures 2, 6 sections 1, 2.1, 2.3, 3 ----- -/--	1-6, 8-14, 17-24

[X] Further documents are listed in the continuation of Box C.

[X] 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 application or patent 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
30 July 2012	06/08/2012

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, Fax: (+31-70) 340-3016	Streich, Sebastian

2

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

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	FATIH KAMISLI ET AL: "Video compression with 1-D directional transforms in H.264/AVC", IEEE INTERNATIONAL CONFERENCE ONACOUSTICS SPEECH AND SIGNAL PROCESSING (ICASSP), 2010 , IEEE, PISCATAWAY, NJ, USA, 14 March 2010 (2010-03-14), pages 738-741, XP031697009, ISBN: 978-1-4244-4295-9 abstract; figures 1, 2 sect. 1-3, 5	1-11, 13-16, 19-24
X,P	MARTA MRAK ET AL: "Transform skip mode", 7. JCT-VC MEETING; 98. MPEG MEETING; GENEVA; (JOINT COLLABORATIVE TEAM ON VIDEO CODING OF ISO/IEC JTC1/SC29/WG11 AND ITU-T SG.16), no. JCTVC-G575, 8 November 2011 (2011-11-08), XP030110559, the whole document	1-24
A	EP 2 192 783 A1 (NARROSCHKE MATTHIAS [DE]; MUSMANN HANS-GEORG [DE]) 2 June 2010 (2010-06-02) abstract; figures 1,2,4-6 paragraphs [0004], [0018] paragraph [0024] - paragraph [0027] paragraph [0035]	1-24

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/GB2012/051412

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 2192783 A1	02-06-2010	AT 480104 T	15-09-2010
		CN 101356823 A	28-01-2009
		EP 1982524 A1	22-10-2008
		EP 2192783 A1	02-06-2010
		ES 2347912 T3	25-11-2010
		JP 2009522968 A	11-06-2009
		JP 2012124965 A	28-06-2012
		KR 20080087147 A	30-09-2008
		US 2011038410 A1	17-02-2011
		WO 2007079964 A1	19-07-2007