PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :		(11) International Publication Number:
H04N 7/26, 17/00, 7/50	**A1**	**WO 00/22831**
		(43) International Publication Date: 20 April 2000 (20.04.00)

(21) International Application Number: PCT/GB99/03359

(22) International Filing Date: 11 October 1999 (11.10.99)

(30) Priority Data:
9822092.4 9 October 1998 (09.10.98) GB

(71) Applicant *(for all designated States except US)*: SNELL & WILCOX LIMITED [GB/GB]; 6 Old Lodge Place, St Margaret's, Twickenham TW1 1RQ (GB).

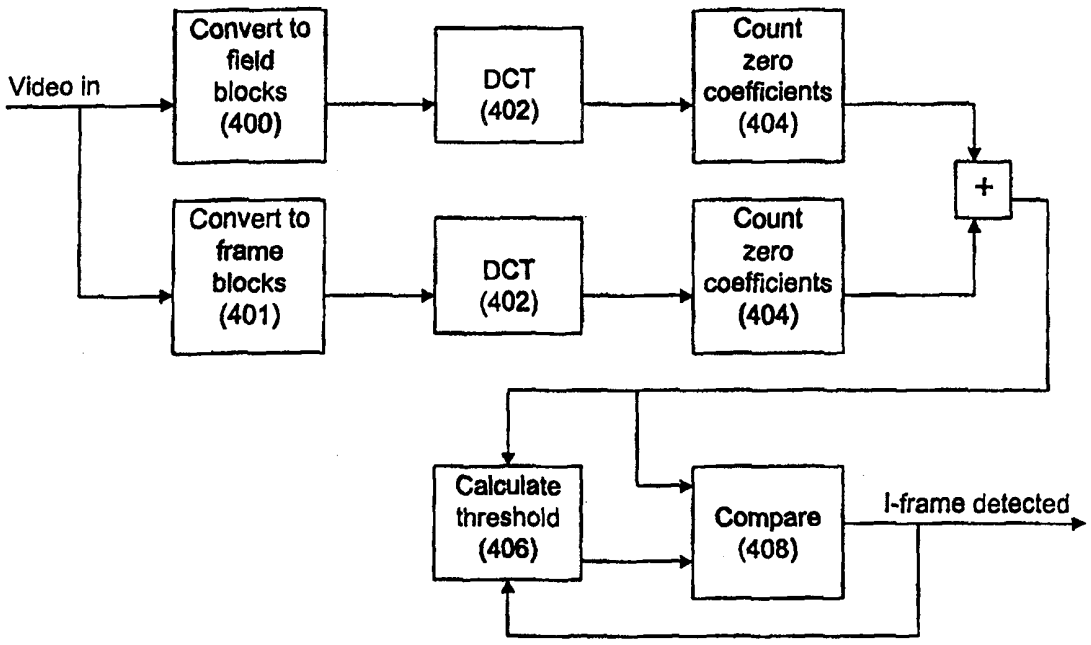
(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: KNEE, Michael, James [GB/GB]; 6 Woodbury Avenue, Petersfield, Hampshire GU32 2EE (GB). POOLE, Ian, James [GB/GB]; 3 Elmdale, Madeline Road, Petersfield, Hampshire GU31 4AJ (GB).

(74) Agents: GARRATT, Peter, Douglas et al.; Mathys & Squire, 100 Gray's Inn Road, London WC1X 8AL (GB).

(81) Designated States: AU, CA, JP, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: ANALYSIS OF COMPRESSION DECODED VIDEO IMAGE SEQUENCES



(57) Abstract

A method of detecting I–frames in a video signal which has previously been MPEG coded, involves taking a DCT and analysing the frequency of zero value coefficients. An I–frame, which does not utilise prediction coding is expected to have a higher number of zero coefficients than a predicted P– or B– frame.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

ANALYSIS OF COMPRESSION DECODED VIDEO IMAGE SEQUENCES

This invention relates to the analysis of compression decoded sequences and in the most important example to the analysis of video signals that have been decoded from an MPEG bitstream.

It is now understood that processes such as re-encoding of a video signal can be significantly improved with knowledge of at least some of the coding decisions used in the original encoding.

Proposals have been made for making some or all of these coding decisions available explicitly. Examples of these proposals can be seen in previous patent applications [see EP 0 765 576 and EP 0 913 058]. These methods involve the use of an Information Bus which passes MPEG coding parameters from a decoder to a subsequent re-coder.

However, in certain situations, no such explicit information is available, and the only available information is that contained in the decoded video signal.

It will be well understood that in the MPEG-2 video compression standard, there are different categories of frames which differ in the degree to which their frames are coded using prediction, and that these categories are denoted by I-, P- and B-frames respectively. An important coding decision to be taken into consideration in a re-encoding process is accordingly the frame structure in terms of the I-, P- and B-frames.

Accordingly, it is an object of the invention to determine by analysis of the video signal, information concerning the upstream coding and decoding process that is useful in minimizing degradation of picture quality in a subsequent, downstream coding and decoding process.

A further object of the present invention is to derive information from a decoded video signal concerning the categories of frames employed in the encoding process. It is a further object of this invention to assist in maintaining picture quality when cascading compression decoding and coding processes.

Accordingly the present invention consists in a method of analysing a signal derived in coding and decoding processes which utilise a quantisation process having a set of quantisation values in which the coded signal contains

categories of frames which categories differ in the degree to which their frames are coded using prediction, the method comprising the steps of measuring the occurrence of values in the signal corresponding with the set of possible quantisation values, and inferring the category of a specific frame by testing the occurrence of said values against a threshold.

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

Figure 1 is a block diagram illustrating the use of an Information Bus as in the prior art;

Figure 2 is a block diagram showing one embodiment of the present invention;

Figure 3 is a graph illustrating the operation of one embodiment of the present invention; and

Figure 4 is a block diagram illustrating in more detail a specific part of the apparatus of the Figure 2 embodiment.

Referring initially to Figure 1, the MPEG decoder (100), adapted as shown in the above prior references, receives an MPEG bitstream. In addition to the output of a standard MPEG decoder, this adapted decoder then produces an Information Bus output conveying the coding decisions taken in the upstream encoder, which are of course inherent in the MPEG bitstream. The Information Bus is then passed to the dumb coder (102) along with the video signal. This dumb coder then follows the coding decisions made by the upstream coder (not shown) which are conveyed by the Information Bus.

This invention is related to the situation in which the adapted MPEG decoder cannot be used to produce the Information Bus because the MPEG bitstream has already been decoded by a standard decoder and the input bitstream is no longer available. The aim is to try to estimate as many as possible of the MPEG coding parameters by analysing the decoded video signal.

When the Information Bus is used, transparent cascading is only possible when all the relevant parameters at sequence, GOP, picture, slice and macroblock rate are carried. Clearly, it is not possible to estimate all these parameters from a decoded picture alone. However, a proportion of the benefit

of the Information Bus can be obtained even if the only information carried from decoder to coder relates to the picture type (I-, P- or B-frame). In this case, the 'dumb' coder of Figure 1 becomes a full MPEG coder except that it sets the picture type to that received in the Information Bus.

The purpose of the present invention is to estimate the picture type by analysing the decoded video signal. Such an estimate would be used as shown in Figure 2. In this embodiment, the standard MPEG decoder (200) receives the bitstream input, and decodes it, producing the usual video signal output. This is then passed to the MPEG coder (206), and also to the picture type detector (202). The information from the picture type detector is then passed to an Information Bus generator (204). The Information Bus then generated, relating to picture type only, is then passed to the MPEG coder. This then follows the coding information available in the Information Bus in coding the video signal, producing an output bitstream.

The following description concentrates on one particular aspect of the picture type detector, the detection of I-frames. In this embodiment, the coding processes referred to are those of MPEG-2, in which the different categories of frames which differ in the degree to which they are coded using prediction are the picture types I, P, and B. The I-frames are coded with no prediction; the P-frames with only forward prediction and the B-frames with both forward and backward prediction.

The invention relies on the observation that intra coded blocks in MPEG-2 are the direct output of an inverse DCT function, whereas predicted macroblocks are the result of an inverse DCT function added to a prediction. If we take the forward DCT of a picture that has been decoded from an MPEG-2 bitstream, then we would expect the DCT coefficients of intra blocks to take only values that were in the set of quantizer reconstruction levels specified in the MPEG-2 standard. The DCT coefficients of predicted blocks might occasionally exhibit this property, but this would only be fortuitous. Unfortunately, without knowledge of the quantizer step size and weighting matrix used in the original encoder, the only quantizer reconstruction level that we know to exist is zero. However, because the distribution of DCT coefficients (even for intra blocks) is

highly peaked around zero, we can still expect a large number of DCT coefficients of intra blocks to be equal to zero.

If we count the number of zero DCT coefficients in each frame, we would expect a high number to indicate either that the frame is an I-frame and contains only intra blocks, or is a predicted frame in which a very high proportion of the blocks are intra coded. In either case, it would be acceptable to judge such a frame as an I-frame for the purposes of optimizing the performance of the re-coding step. One slight complication is that, for luminance blocks, there are two options for every macroblock for the input to the DCT process; **dct_type** can be either frame based or field based. This problem can be avoided by simply taking both kinds of DCT in parallel and including both in the count, accepting the fact that there will be a 2:1 'dilution' of the result. While chrominance blocks do not have this problem in Main Profile MPEG coding, they are not used in the preferred form of the invention because of the ambiguities involved in transcoding between the 4:2:0 and 4:2:2 formats.

Figure 3 illustrates aspects of this particular embodiment of the invention. Graph B in Figure 3 shows the number of zero coefficients encountered for each frame in a typical sequence in which the first and every 12th subsequent frame is an I-frame.

Clear peaks can be seen which indeed correspond to the existence of I-frames. In principle, we should be able to apply a threshold to the curve in order to detect I-frames. However, as the graph illustrates, a simple fixed threshold will not always work. The sequence has a change in scene content around frame 190, and this leads to a drop both in the counts of zero coefficients for I-frames and in the 'background count' for other frames. Some kind of adaptive threshold filter is therefore required. There now follows an example showing how the threshold could be adapted.

The filter sets an initial threshold which is then modified once per frame. If the zero-coefficent count (ZCC) exceeds the threshold, then an I-frame is deemed to have been detected and the threshold is reset to the ZCC value; otherwise the threshold is decreased by a factor, known here as the Threshold Modifier Factor (TMF):

$$T_n = \frac{(T_{n-1} * TMF + V_n)}{TMF + 1}$$

where T_n is the threshold at frame n, and V_n is the ZCC value at time n.

It can be seen that if the incoming ZCC data is constant, then the threshold will tend to that constant.

The TMF is determined using a confidence test. The confidence test looks at the position of the last I-frame found and the number of frames between the last two detected I-frames, and assumes that the I-frames are occurring at regular intervals. If there were no frames between the last two detected I-frames, then we assume that a false I-frame was detected, so the confidence is set to 0.5 until another I-frame is found; otherwise the confidence is determined as:

$$Confidence = \exp\left(\frac{(CurrentFramePosition - ExpectedIFramePosition)}{5}\right)$$

The confidence is clipped to the range 0 (no peak expected) to 1 (peak expected).

The TMF is determined as :

$$TMF = Max - Confidence * (Max - Min)$$

where Min and Max are the minimum and maximum values the TMF is allowed to take.

It can be seen that the greater the confidence, the smaller the TMF, hence the more rapidly the threshold decreases until an I-frame is detected.

The above description is simply an example; other systems of adaptive thresholding could equally be used.

A refinement to the detection method could be to make an explicit and separate detection of scene changes and incorporate this into the confidence measure. This could be augmented by *a priori* knowledge of the strategy

adopted by the original encoder to modify or reset the GOP structure at scene changes.

The performance of the complete detection method is illustrated in Figure 3 , where the four curves show: the threshold, graph A, the zero-coefficient count (ZCC) value, graph B, the confidence value, graph C, and the result - a downward pointing spike indicating that an I-frame has been detected – graph D.

A further particular aspect of the picture type detector involves the detection of all the different categories of frames which make up the frame structure in the video signal and not merely the I-frames. The method here is similar to that employed in detecting the I-frames. In general, a particular P- frame is not necessarily noticeably more intra-coded than a particular B-frame, but if the analysis is performed as for the detection of I-frames, and an average of the results over a large number of frames is taken, it is found that the P-frames are, on average, slightly more intra-coded than the B-frames. This result can be used to determine the particular frame structure which was used in the original encoding.

Given a specific number of non-I-frames between two I-frames, there are only a few different types of frame structure which are used in conventional encoders. For example, given an I-frame every sixth frame, the structure will typically be I, B, B, P, B, B, I or I, B, P, B, P, B, I. The changes, if at all, between frames structures in a given signal are also reasonably infrequent, allowing a large possible sample from which an average can be taken. If an average of the results of an analysis as above are taken, the subtle trends in the numbers of zero coefficients found for the different types of predicted frame can be used, along with the knowledge that there are only a certain known number of different frame structures available, to deduce the frame structure which was used in encoding the given signal. This information can be used alongside information regarding the position of the I-frames to give more detailed information to the MPEG coder (as in Figure 2, (206)) via the Information Bus.

Figure 4 gives a block diagram of an I-frame detector using the method described above. The video signal is first converted to field blocks (400) and

frame blocks (401). The signal from each of these is then passed first through a DCT (402), and the zero coefficients for the signal are then counted (404). The two signals are then added and the combined zero coefficient count is compared at 408 with a threshold. An I-frame is detected if the threshold is exceeded. The threshold is calculated at 406 utilising, as described above, information from the combined zero coefficient count and the location of the last detected I-frame.

It will be understood that this invention has been described by way of example only, and a wide variety of modification is possible without departing from the scope of the invention.

Thus, whilst the described counting of zero coefficients has the advantage of not requiring prior knowledge of the full set of possible quantisation values, there will be circumstances in which it will be appropriate to measure the occurrence of other values. Also, although the example of MPEG-2 is of course very important, the invention is also applicable to other coding schemes which utilise categories of frames which differ in the degree to which frames are coded using prediction.

CLAIMS

1. A method of analysing a signal derived in coding and decoding processes which utilise a quantisation process having a set of quantisation values and in which the coded signal contains categories of frames which categories differ in the degree to which frames are coded using prediction, the method comprising the steps of measuring the occurrence of values in the signal corresponding with the set of possible quantisation values, and inferring the category of a specific frame by testing the occurrence of said values against a threshold.

2. A method according to Claim 1, in which said threshold is varied in accordance with the expected pattern of said categories of frames in the coded signal.

3. A method according to Claim 2, in which said threshold varies with the number of frames since detection in the decoded signal of a particular category of frame.

4. A method according to Claim 3, in which said particular category of frame contains those frames which are coded with no prediction.

5. A method according to any one of the preceding claims, in which the occurrence of zero values is measured.

6. A method according to any one of the preceding claims, in which said coding and decoding processes are those employed in MPEG-2, and in which said categories of frames are denoted I-, P- and B-frames respectively.

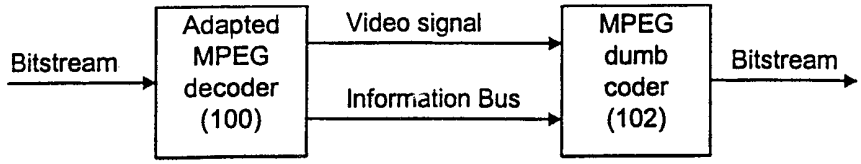


Figure 1 PRIOR ART

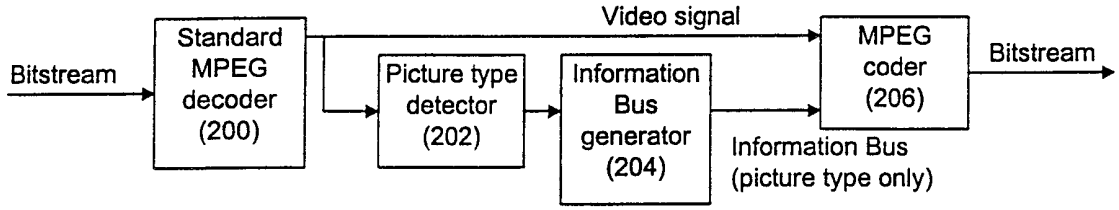


Figure 2

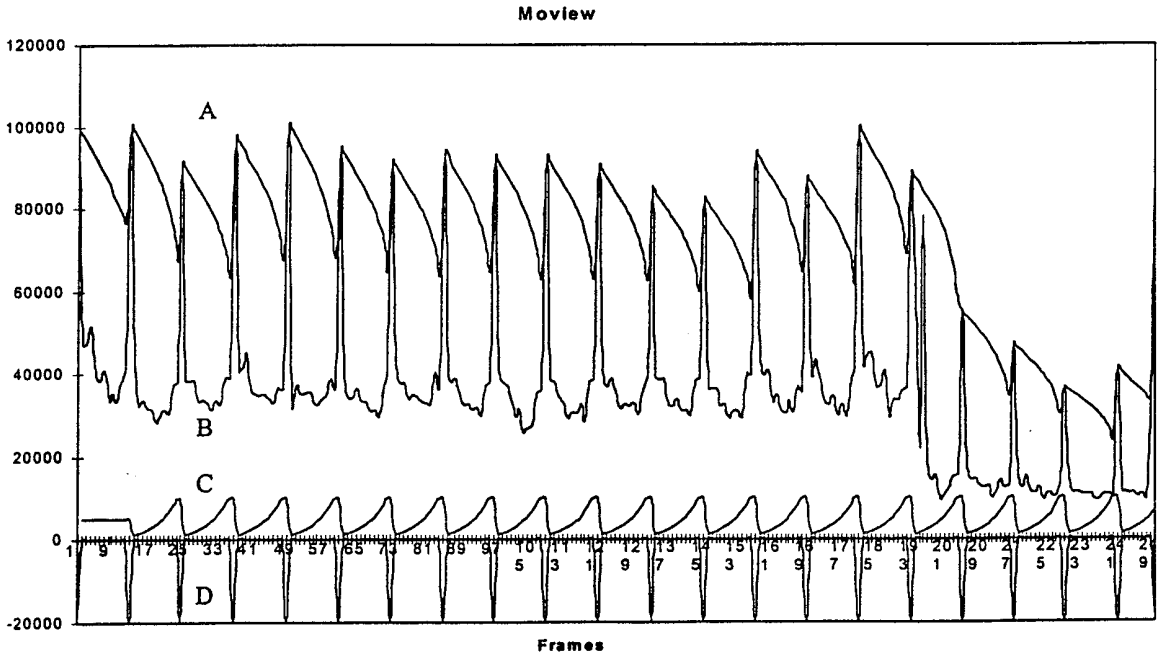


Figure 3

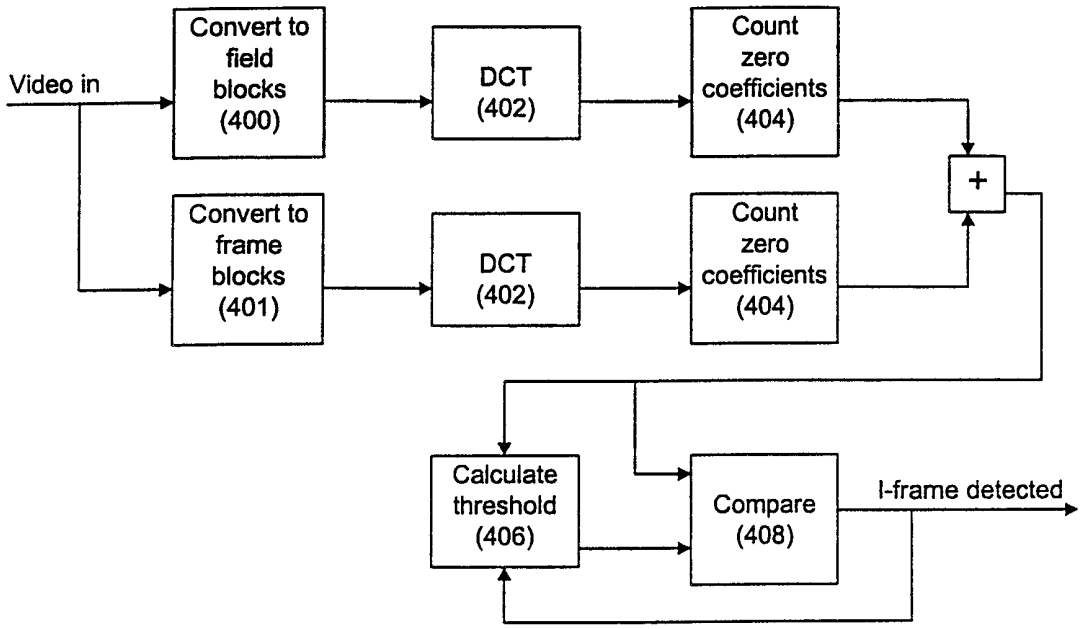

Video in
Convert to field blocks (400)
DCT (402)
Count zero coefficients (404)
+
Convert to frame blocks (401)
DCT (402)
Count zero coefficients (404)
Calculate threshold (406)
Compare (408)
I-frame detected

Figure 4

INTERNATIONAL SEARCH REPORT

Inte onal Application No
PCT/GB 99/03359

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 H04N7/26 H04N17/00 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)
IPC 7 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 practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 98 26602 A (SONY CORP ;MIHARA KANJI (JP)) 18 June 1998 (1998-06-18) abstract; figures	1-6
Y	MURATA ET AL: "Fast 2D IDCT Implementation with Multimedia Instructions for a Software MPEG2 Decoder" PROCEEDINGS OF THE 1998 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING, vol. 5, 12 May 1998 (1998-05-12), pages 3105-3108, XP002130031 abstract page 3106, paragraph 2.3	1-6
A	EP 0 710 030 A (MITSUBISHI ELECTRIC CORP) 1 May 1996 (1996-05-01) claims 13-18	1

-/--

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

[X] Patent family members are listed in 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
8 February 2000	21/02/2000

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	Foglia, P

2

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Inte onal Application No
PCT/GB 99/03359

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	EP 0 509 576 A (AMPEX) 21 October 1992 (1992-10-21) page 5, line 3 - line 14 page 12, line 11 - line 23	1

2

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inte onal Application No
PCT/GB 99/03359

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9826602 A	18-06-1998	JP 10174106 A	26-06-1998
		JP 10174098 A	26-06-1998
EP 0710030 A	01-05-1996	JP 8130743 A	21-05-1996
		AU 691268 B	14-05-1998
		AU 2052995 A	09-05-1996
		CA 2151085 A	01-05-1996
		CN 1126409 A	10-07-1996
		NO 952221 A	02-05-1996
		SG 33377 A	18-10-1996
		US 5831688 A	03-11-1998
EP 0509576 A	21-10-1992	AT 162918 T	15-02-1998
		CA 2063972 A	19-10-1992
		DE 69224227 D	05-03-1998
		DE 69224227 T	09-07-1998
		JP 5153406 A	18-06-1993
		MX 9201788 A	01-02-1993
		US 5930398 A	27-07-1999