

8457761



THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

February 09, 2024

**THIS IS TO CERTIFY THAT ANNEXED IS A TRUE COPY FROM THE
RECORDS OF THIS OFFICE OF THE FILE WRAPPER AND CONTENTS
OF:**

APPLICATION NUMBER: 13/666,680
FILING DATE: November 01, 2012
PATENT NUMBER: 9571833
ISSUE DATE: February 14, 2017



Certified by

Kathi

Performing the Functions and Duties of the
Under Secretary of Commerce
for Intellectual Property
and Director of the United States
Patent and Trademark Office

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	NC77198US-NP
		Application Number	
Title of Invention	METHOD FOR CODING AND AN APPARATUS		
The application data sheet is part of the provisional or nonprovisional application for which it is being submitted. The following form contains the bibliographic data arranged in a format specified by the United States Patent and Trademark Office as outlined in 37 CFR 1.76. This document may be completed electronically and submitted to the Office in electronic format using the Electronic Filing System (EFS) or the document may be printed and included in a paper filed application.			

Secrecy Order 37 CFR 5.2

☐ Portions or all of the application associated with this Application Data Sheet may fall under a Secrecy Order pursuant to 37 CFR 5.2 (Paper filers only. Applications that fall under Secrecy Order may not be filed electronically.)

Applicant Information:

Applicant 1					Remove
Applicant Authority		☒ Inventor		☐ Legal Representative under 35 U.S.C. 117	☐ Party of Interest under 35 U.S.C. 118
Prefix	Given Name	Middle Name	Family Name	Suffix	
	Mehmet	Oguz	BICI		
Residence Information (Select One) ☐ US Residency ☒ Non US Residency ☐ Active US Military Service					
City	Tampere	Country Of Residenceⁱ	TR		
Citizenship under 37 CFR 1.41(b)ⁱ		FI			
Mailing Address of Applicant:					
Address 1		Tammelan puistokatu 1-3 D 46			
Address 2					
City	Tampere	State/Province			
Postal Code	33500	Countryⁱ	FI		
Applicant 2					Remove
Applicant Authority		☒ Inventor		☐ Legal Representative under 35 U.S.C. 117	☐ Party of Interest under 35 U.S.C. 118
Prefix	Given Name	Middle Name	Family Name	Suffix	
	Jani		LAINEMA		
Residence Information (Select One) ☐ US Residency ☒ Non US Residency ☐ Active US Military Service					
City	Tampere	Country Of Residenceⁱ	FI		
Citizenship under 37 CFR 1.41(b)ⁱ		FI			
Mailing Address of Applicant:					
Address 1		Kisakentankatu 12 B 6			
Address 2					
City	Tampere	State/Province			
Postal Code	33230	Countryⁱ	FI		
Applicant 3					Remove
Applicant Authority		☒ Inventor		☐ Legal Representative under 35 U.S.C. 117	☐ Party of Interest under 35 U.S.C. 118
Prefix	Given Name	Middle Name	Family Name	Suffix	
	Kemal		UGUR		
Residence Information (Select One) ☐ US Residency ☒ Non US Residency ☐ Active US Military Service					
City	Tampere	Country Of Residenceⁱ	FI		

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	NC77198US-NP	
		Application Number		
Title of Invention	METHOD FOR CODING AND AN APPARATUS			
Citizenship under 37 CFR 1.41(b) i		TR		
Mailing Address of Applicant:				
Address 1	Lapintie 6D 25			
Address 2				
City	Tampere	State/Province		
Postal Code	33100	Country	FI	
All Inventors Must Be Listed - Additional Inventor Information blocks may be generated within this form by selecting the Add button. Add				

Correspondence Information:

Enter either Customer Number or complete the Correspondence Information section below. For further information see 37 CFR 1.33(a).			
☐ An Address is being provided for the correspondence Information of this application.			
Customer Number	73658		
Email Address	ipr.inhouse@nokia.com	Add Email	Remove Email

Application Information:

Title of the Invention	METHOD FOR CODING AND AN APPARATUS		
Attorney Docket Number	NC77198US-NP	Small Entity Status Claimed ☐	
Application Type	Nonprovisional		
Subject Matter	Utility		
Suggested Class (if any)		Sub Class (if any)	
Suggested Technology Center (if any)			
Total Number of Drawing Sheets (if any)	13	Suggested Figure for Publication (if any)	

Publication Information:

☐ Request Early Publication (Fee required at time of Request 37 CFR 1.219)
☐ Request Not to Publish. I hereby request that the attached application not be published under 35 U.S.C. 122(b) and certify that the invention disclosed in the attached application has not and will not be the subject of an application filed in another country, or under a multilateral international agreement, that requires publication at eighteen months after filing.

Representative Information:

Representative information should be provided for all practitioners having a power of attorney in the application. Providing this information in the Application Data Sheet does not constitute a power of attorney in the application (see 37 CFR 1.32). Enter either Customer Number or complete the Representative Name section below. If both sections are completed the Customer Number will be used for the Representative Information during processing.			
Please Select One:	☒ Customer Number	☐ US Patent Practitioner	☐ Limited Recognition (37 CFR 11.9)

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	NC77198US-NP
		Application Number	
Title of Invention	METHOD FOR CODING AND AN APPARATUS		
Customer Number	73658		

Domestic Benefit/National Stage Information:

This section allows for the applicant to either claim benefit under 35 U.S.C. 119(e), 120, 121, or 365(c) or indicate National Stage entry from a PCT application. Providing this information in the application data sheet constitutes the specific reference required by 35 U.S.C. 119(e) or 120, and 37 CFR 1.78(a)(2) or CFR 1.78(a)(4), and need not otherwise be made part of the specification.			
Prior Application Status		Remove	
Application Number	Continuity Type	Prior Application Number	Filing Date (YYYY-MM-DD)
	non provisional of	61/555703	2011-11-04
Additional Domestic Benefit/National Stage Data may be generated within this form by selecting the Add button.			Add

Foreign Priority Information:

This section allows for the applicant to claim benefit of foreign priority and to identify any prior foreign application for which priority is not claimed. Providing this information in the application data sheet constitutes the claim for priority as required by 35 U.S.C. 119(b) and 37 CFR 1.55(a).			
			Remove
Application Number	Country ⁱ	Parent Filing Date (YYYY-MM-DD)	Priority Claimed
			☒ Yes ☐ No
Additional Foreign Priority Data may be generated within this form by selecting the Add button.			Add

Assignee Information:

Providing this information in the application data sheet does not substitute for compliance with any requirement of part 3 of Title 37 of the CFR to have an assignment recorded in the Office.			
Assignee ¹			Remove
If the Assignee is an Organization check here. ☒			
Organization Name	Nokia Corporation		
Mailing Address Information:			
Address 1	Keilalahdentie 4		
Address 2			
City	Espoo	State/Province	
Country ⁱ	FI	Postal Code	02150
Phone Number		Fax Number	
Email Address			
Additional Assignee Data may be generated within this form by selecting the Add button.			Add

Signature:

A signature of the applicant or representative is required in accordance with 37 CFR 1.33 and 10.18. Please see 37 CFR 1.4(d) for the form of the signature.
--

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number		NC77198US-NP	
		Application Number			
Title of Invention		METHOD FOR CODING AND AN APPARATUS			
Signature		/Thomas J. Arria/		Date (YYYY-MM-DD)	2012-11-01
First Name	Thomas J.	Last Name	Arria	Registration Number	60223

This collection of information is required by 37 CFR 1.76. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 23 minutes to complete, including gathering, preparing, and submitting the completed application data sheet form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these records.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

METHOD FOR CODING AND AN APPARATUS

TECHNICAL FIELD

[0001] There is provided a method for encoding, a method for decoding, an apparatus, computer program products, an encoder and a decoder.

BACKGROUND INFORMATION

[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 may comprise an encoder which transforms input video into a compressed representation suitable for storage and/or transmission and a decoder that can uncompress the compressed video representation back into a viewable form, or either one of them. The encoder may discard some information in the original video sequence in order to represent the video in a more compact form, for example at a lower bit rate.

[0004] Many hybrid video codecs, operating for example according to the International Telecommunication Union's ITU-T H.263 and H.264 coding standards, 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 encoded video frames (or a later coded video frame) that corresponds closely to the block being coded. Additionally, pixel values can be predicted by spatial mechanisms which involve finding and indicating a spatial region relationship, for example by using pixel values around the block to be coded in a specified manner.

[0005] Prediction approaches using image information from a previous (or a later) image can also be called as Inter prediction methods, and prediction approaches using image information within the same image can also be called as Intra prediction methods.

[0006] The second phase is one of coding the error between the predicted block of pixels and the original block of pixels. This may be accomplished by transforming the difference in

pixel values using a specified transform. This transform may be e.g. a Discrete Cosine Transform (DCT) or a variant thereof. After transforming the difference, the transformed difference may be quantized and entropy encoded.

[0007] By varying the fidelity of the quantization process, the encoder can control the balance between the accuracy of the pixel representation, (in other words, the quality of the picture) and the size of the resulting encoded video representation (in other words, the file size or transmission bit rate).

[0008] The decoder reconstructs the output video by applying a prediction mechanism similar to that used by the encoder in order to form a predicted representation of the pixel blocks (using the motion or spatial information created by the encoder and stored in the compressed representation of the image) and prediction error decoding (the inverse operation of the prediction error coding to recover the quantized prediction error signal in the spatial domain).

[0009] After applying pixel prediction and error decoding processes the decoder combines the prediction and the prediction error signals (the pixel values) to form the output video frame.

[0010] The decoder (and encoder) may also apply additional filtering processes in order to improve the quality of the output video before passing it for display and/or storing as a prediction reference for the forthcoming frames in the video sequence.

[0011] In some video codecs, such as High Efficiency Video Coding Working Draft 4, video pictures may be divided into coding units (CU) covering the area of a picture. A coding unit consists of one or more prediction units (PU) defining the prediction process for the samples within the coding unit and one or more transform units (TU) defining the prediction error coding process for the samples in the coding unit. A coding unit may consist of a square block of samples with a size selectable from a predefined set of possible coding unit sizes. A coding unit with the maximum allowed size can be named as a largest coding unit (LCU) and the video picture may be divided into non-overlapping largest coding units. A largest coding unit can further be split into a combination of smaller coding units, e.g. by recursively splitting the largest coding unit and resultant coding units. Each resulting coding unit may have at least one prediction unit and at least one transform unit associated with it. Each prediction unit and transform unit can further be split into smaller prediction units and transform units in order to increase granularity of the prediction and prediction error coding processes, respectively. Each

prediction unit may have prediction information associated with it defining what kind of a prediction is to be applied for the pixels within that prediction unit (e.g. motion vector information for inter predicted prediction units and intra prediction directionality information for intra predicted prediction units). Similarly, each transform unit may be associated with information describing the prediction error decoding process for samples within the transform unit (including e.g. discrete cosine transform (DCT) coefficient information). It may be signalled at coding unit level whether prediction error coding is applied or not for each coding unit. In the case there is no prediction error residual associated with the coding unit, it can be considered there are no transform units for the coding unit. The division of the image into coding units, and division of coding units into prediction units and transform units may be signalled in the bitstream allowing the decoder to reproduce the intended structure of these units.

[0012] In some video codecs, motion information is indicated by motion vectors associated with each motion compensated image block. These motion vectors represent the displacement of the image block in the picture to be coded (in the encoder) or decoded (at the decoder) and the prediction source block in one of the previously coded or decoded images (or pictures). In order to represent motion vectors efficiently, motion vectors may be coded differentially with respect to block specific predicted motion vector. In some video codecs, 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.

[0013] Another way to create motion vector predictions is to generate a list or a set of candidate predictions from blocks in the current frame and/or co-located or other blocks in temporal reference pictures and signalling the chosen candidate as the motion vector prediction. A spatial motion vector prediction is a prediction obtained only on the basis of information of one or more blocks of the same frame than the current frame whereas temporal motion vector prediction is a prediction obtained on the basis of information of one or more blocks of a frame different from the current frame. It may also be possible to obtain motion vector predictions by combining both spatial and temporal prediction information of one or more encoded blocks. These kinds of motion vector predictions are called as spatio-temporal motion vector predictions.

[0014] In addition to predicting the motion vector values, the reference index in the reference picture list can be predicted. The reference index may be predicted from blocks in the current frame and/or co-located or other blocks in a temporal reference picture. Moreover, some

high efficiency video codecs employ an additional motion information coding/decoding mechanism, often called merging/merge mode, where all the motion field information, which includes motion vector and corresponding reference picture index for each available reference picture list, may be predicted and used without any modification or correction. Similarly, predicting the motion field information may be carried out using the motion field information of blocks in the current frame and/or co-located or other blocks in temporal reference pictures and the used motion field information is signalled among a list of motion field candidate list filled with motion field information of available blocks in the current frame and/or co-located or other blocks in temporal reference pictures.

[0015] In some video codecs the prediction residual after motion compensation is first transformed with a transform kernel (like DCT) and then coded. The reason for this is that often there still exists some correlation among the residual and transform can in many cases help reduce this correlation and provide more efficient coding.

[0016] Some 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:

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

where C is the Lagrangian cost to be minimized, D is the image distortion (e.g. 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).

[0017] Some video codecs such as hybrid video codecs may generate a list of motion vector predictions (MVP) consisting of motion vectors of spatial adjacent blocks (spatial MVP) and/or motion vectors of blocks in a previously decoded frame (temporal MVP). One of the candidate motion vectors in the list is signalled to be used as the motion vector prediction of the current block. After the list is generated, some of the motion vector prediction candidates may have the same motion information. In this case, the identical motion vector prediction candidates

may be removed to reduce redundancy. During the decoding, if the temporal motion vector prediction information is unavailable due to e.g. loss of reference frame, the decoder may not know if the temporal motion vector prediction candidate in the list is to be removed. This may lead to uncertainty for mapping the decoded candidate index to the candidates whose removal decision is based on comparing motion information with the temporal motion vector prediction. As a result, false assignment of motion vector prediction candidates may occur which may lead to degradation in the picture quality and drift of false motion information throughout the decoding process.

SUMMARY

[0018] The present invention introduces a method for generating a motion vector prediction list for an image block. In some embodiments video codecs employ in a motion prediction candidate list construction a way to reduce the complexity of the implementation. This can be achieved by performing a limited number of motion information comparisons between candidate pairs to remove the redundant candidates rather than comparing every available candidate pair. The decision of whether comparing two candidates may depend on the order of the candidates to be considered for the list and/or coding/prediction mode and/or location of the blocks associated with the candidates. In some embodiments a video codec employs a merge process for motion information coding and creates a list of motion prediction candidates from which one of the candidates is to be signalled as the motion information for the current coding or prediction unit. The motion prediction candidates may consist of several spatial motion predictions and a temporal motion prediction. The spatial candidates are obtained from the motion information of e.g. spatial neighbour blocks.

[0019] According to a first aspect of the present invention there is provided a method comprising:

receiving a block of pixels including a prediction unit; determining a set of spatial motion vector prediction candidates for the block of pixels; the spatial motion vector prediction candidates being provided with motion information;

selecting a first spatial motion vector prediction candidate from the set of spatial motion vector prediction candidates as a potential spatial motion vector prediction candidate to be included in a merge list for the prediction unit;

determining a subset of spatial motion vector predictions based on the location of the block associated with the first spatial motion vector prediction candidate;

comparing motion information of the first spatial motion vector prediction candidate with motion information of the spatial motion vector prediction candidate in the determined subset of spatial motion vector prediction candidates;

if at least one of the comparisons indicates that the motion vector information of the spatial motion vector prediction candidates correspond with each other, excluding the first spatial motion vector prediction candidate from the merge list.

[0020] According to a second aspect of the present invention there is provided a method comprising:

receiving an encoded block of pixels including a prediction unit;

determining a set of spatial motion vector prediction candidates for the encoded block of pixels; the spatial motion vector prediction candidates being provided with motion information;

selecting a first spatial motion vector prediction candidate from the set of spatial motion vector prediction candidates as a potential spatial motion vector prediction candidate to be included in a merge list for the prediction unit;

determining a subset of spatial motion vector predictions based on the location of the block associated with the first spatial motion vector prediction candidate;

comparing motion information of the first spatial motion vector prediction candidate with motion information of another spatial motion vector prediction candidate of the set of spatial motion vector prediction candidates;

if at least one of the comparisons indicates that the motion vector information of the spatial motion vector prediction candidates correspond with each other, excluding the first spatial motion vector prediction candidate from the merge list.

[0021] According to a third aspect of the present invention there is provided an apparatus comprising a processor and a memory including computer program code, the memory and the computer program code configured to, with the processor, cause the apparatus to:

receive a block of pixels

including a prediction unit;

determining a set of spatial motion vector prediction candidates for the block of pixels;
the spatial motion vector prediction candidates being provided with motion information;

selecting a first spatial motion vector prediction candidate from the set of spatial motion
vector prediction candidates as a potential spatial motion vector prediction candidate to be
included in a merge list for the prediction unit;

determining a subset of spatial motion vector predictions based on the location of the
block associated with the first spatial motion vector prediction candidate;

comparing motion information of the first spatial motion vector prediction candidate
with motion information of the spatial motion vector prediction candidate in the determined
subset of spatial motion vector prediction candidates;

if at least one the comparisons indicates that the motion vector information of the spatial
motion vector prediction candidates correspond with each other, excluding the first spatial
motion vector prediction candidate from the merge list.

[0022] According to a fourth aspect of the present invention there is provided an
apparatus comprising a processor and a memory including computer program code, the memory
and the computer program code configured to, with the processor, cause the apparatus to:

receive an encoded block of pixels

including a prediction unit;

determining a set of spatial motion vector prediction candidates for the encoded block
of pixels; the spatial motion vector prediction candidates being provided with motion
information;

selecting a first spatial motion vector prediction candidate from the set of spatial motion
vector prediction candidates as a potential spatial motion vector prediction candidate to be
included in a merge list for the prediction unit;

determining a subset of spatial motion vector predictions based on the location of the
block associated with the first spatial motion vector prediction candidate;

comparing motion information of the first spatial motion vector prediction candidate
with motion information of another spatial motion vector prediction candidate of the set of
spatial motion vector prediction candidates;

if at least one of the comparisons indicates that the motion vector information of the spatial motion vector prediction candidates correspond with each other, excluding the first spatial motion vector prediction candidate from the merge list.

[0023] According to a fifth aspect of the present invention there is provided a storage medium having stored thereon a computer executable program code for use by an encoder, said program code comprises instructions for:

receiving a block of pixels including a prediction unit;

determining a set of spatial motion vector prediction candidates for the block of pixels; the spatial motion vector prediction candidates being provided with motion information;

select a first spatial motion vector prediction candidate from the set of spatial motion vector prediction candidates as a potential spatial motion vector prediction candidate to be included in a merge list for the prediction unit;

determine a subset of spatial motion vector predictions based on the location of the block associated with the first spatial motion vector prediction candidate;

compare motion information of the first spatial motion vector prediction candidate with motion information of the spatial motion vector prediction candidate in the determined subset of spatial motion vector prediction candidates;

exclude the first spatial motion vector prediction candidate from the merge list, if at least one of the comparisons indicates that the motion vector information of the spatial motion vector prediction candidates correspond with each other

[0024] According to a sixth aspect of the present invention there is provided a storage medium having stored thereon a computer executable program code for use by a decoder, said program code comprises instructions for:

receiving an encoded block of pixels including a prediction unit;

determining a set of spatial motion vector prediction candidates for the encoded block of pixels; the spatial motion vector prediction candidates being provided with motion information;

selecting a first spatial motion vector prediction candidate from the set of spatial motion vector prediction candidates as a potential spatial motion vector prediction candidate to be included in a merge list for the prediction unit;

determining a subset of spatial motion vector predictions based on the location of the block associated with the first spatial motion vector prediction candidate;

comparing motion information of the first spatial motion vector prediction candidate with motion information of the spatial motion vector prediction candidate in the determined subset of spatial motion vector prediction candidates;

if at least one of the comparisons indicates that the motion vector information of the spatial motion vector prediction candidates correspond with each other, excluding the first spatial motion vector prediction candidate from the merge list.

[0025] According to a seventh aspect of the present invention there is provided an apparatus comprising:

means for receiving a block of pixels including a prediction unit;

means for determining a set of spatial motion vector prediction candidates for the block of pixels; the spatial motion vector prediction candidates being provided with motion information;

selecting a first spatial motion vector prediction candidate from the set of spatial motion vector prediction candidates as a potential spatial motion vector prediction candidate to be included in a merge list for the prediction unit;

determining a subset of spatial motion vector predictions based on the location of the block associated with the first spatial motion vector prediction candidate;

comparing motion information of the first spatial motion vector prediction candidate with motion information of the spatial motion vector prediction candidate in the determined subset of spatial motion vector prediction candidates;

if at least one of the comparisons indicates that the motion vector information of the spatial motion vector prediction candidates correspond with each other, excluding the first spatial motion vector prediction candidate from the merge list.

[0026] According to an eighth aspect of the present invention there is provided an apparatus comprising:

means for receiving an encoded block of pixels including a prediction unit;

means for determining a set of spatial motion vector prediction candidates for the encoded block of pixels; the spatial motion vector prediction candidates being provided with motion information;

means for selecting a first spatial motion vector prediction candidate from the set of spatial motion vector prediction candidates as a potential spatial motion vector prediction candidate to be included in a merge list for the prediction unit;

means for determining a subset of spatial motion vector predictions based on the location of the block associated with the first spatial motion vector prediction candidate;

means for comparing motion information of the first spatial motion vector prediction candidate with motion information of the spatial motion vector prediction candidate in the determined subset of spatial motion vector prediction candidates;

means for excluding the first spatial motion vector prediction candidate from the merge list, if at least one of the comparisons indicates that the motion vector information of the spatial motion vector prediction candidates correspond with each other.

DESCRIPTION OF THE DRAWINGS

[0027] For better understanding of the present invention, reference will now be made by way of example to the accompanying drawings in which:

[0028] Figure 1 shows schematically an electronic device employing some embodiments of the invention;

[0029] Figure 2 shows schematically a user equipment suitable for employing some embodiments of the invention;

[0030] Figure 3 further shows schematically electronic devices employing embodiments of the invention connected using wireless and wired network connections;

[0031] Figure 4a shows schematically an embodiment of the invention as incorporated within an encoder;

[0032] Figure 4b shows schematically an embodiment of a prediction reference list generation and modification according to some embodiments of the invention;

[0033] Figures 5a and 5b show a flow diagram showing the operation of an embodiment of the invention with respect to the encoder as shown in figure 4a;

[0034] Figure 6a illustrates an example of spatial and temporal prediction of a prediction unit;

[0035] Figure 6b illustrates another example of spatial and temporal prediction of a prediction unit;

[0036] Figure 7 shows schematically an embodiment of the invention as incorporated within a decoder;

[0037] Figures 8a and 8b show a flow diagram of showing the operation of an embodiment of the invention with respect to the decoder shown in figure 7;

[0038] Figure 9 illustrates an example of a coding unit and some neighbour blocks of the coding unit;

[0039] Figure 10a illustrates an example of a horizontal division of the coding unit;

[0040] Figure 10b illustrates an example of a vertical division of the coding unit;

[0041] Figure 11a illustrates locations of five spatial neighbours A0, A1, B0, B1, B2 for a prediction unit generated as the second prediction unit of a horizontally divided coding unit;

[0042] Figure 11b illustrates locations of five spatial neighbours for a prediction unit generated as the second prediction unit of a vertically divided coding unit; and

[0043] Figure 12 illustrates an example of blocks between some spatial neighbours of a coding unit.

DETAILED DESCRIPTION OF SOME EXAMPLE EMBODIMENTS

[0044] The following describes in further detail suitable apparatus and possible mechanisms for the provision of improving the prediction accuracy and hence possibly reducing information to be transmitted in video coding systems. In this regard reference is first made to Figure 1 which shows a schematic block diagram of an exemplary apparatus or electronic device 50, which may incorporate a codec according to an embodiment of the invention.

[0045] The electronic device 50 may for example be a mobile terminal or user equipment of a wireless communication system. However, it would be appreciated that embodiments of the invention may be implemented within any electronic device or apparatus which may require encoding and decoding or encoding or decoding video images.

[0046] The apparatus 50 may comprise a housing 30 for incorporating and protecting the device. The apparatus 50 further may comprise a display 32 in the form of a liquid crystal display. In other embodiments of the invention the display may be any suitable display technology suitable to display an image or video. The apparatus 50 may further comprise a keypad 34. In other embodiments of the invention any suitable data or user interface mechanism may be employed. For example the user interface may be implemented as a virtual keyboard or data entry system as part of a touch-sensitive display. The apparatus may comprise a microphone 36 or any suitable audio input which may be a digital or analogue signal input. The apparatus 50 may further comprise an audio output device which in embodiments of the invention may be any one of: an earpiece 38, speaker, or an analogue audio or digital audio output connection. The apparatus 50 may also comprise a battery 40 (or in other embodiments of the invention the device may be powered by any suitable mobile energy device such as solar cell, fuel cell or clockwork generator). The apparatus may further comprise an infrared port 42 for short range line of sight communication to other devices. In other embodiments the apparatus 50 may further comprise any suitable short range communication solution such as for example a Bluetooth wireless connection or a USB/firewire wired connection.

[0047] The apparatus 50 may comprise a controller 56 or processor for controlling the apparatus 50. The controller 56 may be connected to memory 58 which in embodiments of the invention may store both data in the form of image and audio data and/or may also store instructions for implementation on the controller 56. The controller 56 may further be connected

to codec circuitry 54 suitable for carrying out coding and decoding of audio and/or video data or assisting in coding and decoding carried out by the controller 56.

[0048] The apparatus 50 may further comprise a card reader 48 and a smart card 46, for example a UICC and UICC reader for providing user information and being suitable for providing authentication information for authentication and authorization of the user at a network.

[0049] The apparatus 50 may comprise radio interface circuitry 52 connected to the controller and suitable for generating wireless communication signals for example for communication with a cellular communications network, a wireless communications system or a wireless local area network. The apparatus 50 may further comprise an antenna 44 connected to the radio interface circuitry 52 for transmitting radio frequency signals generated at the radio interface circuitry 52 to other apparatus(es) and for receiving radio frequency signals from other apparatus(es).

[0050] In some embodiments of the invention, the apparatus 50 comprises a camera capable of recording or detecting individual frames which are then passed to the codec 54 or controller for processing. In some embodiments of the invention, the apparatus may receive the video image data for processing from another device prior to transmission and/or storage. In some embodiments of the invention, the apparatus 50 may receive either wirelessly or by a wired connection the image for coding/decoding.

[0051] With respect to Figure 3, an example of a system within which embodiments of the present invention can be utilized is shown. The system 10 comprises multiple communication devices which 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 wireless cellular telephone network (such as a GSM, UMTS, CDMA network etc), a wireless local area network (WLAN) such as defined by any of the IEEE 802.x standards, a Bluetooth personal area network, an Ethernet local area network, a token ring local area network, a wide area network, and the Internet.

[0052] The system 10 may include both wired and wireless communication devices or apparatus 50 suitable for implementing embodiments of the invention.

[0053] For example, the system shown in Figure 3 shows a mobile telephone network 11 and a representation of 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 similar communication pathways.

[0054] The example communication devices shown in the system 10 may include, but are not limited to, an electronic device or apparatus 50, a combination of a personal digital assistant (PDA) and a mobile telephone 14, a PDA 16, an integrated messaging device (IMD) 18, a desktop computer 20, a notebook computer 22. The apparatus 50 may be stationary or mobile when carried by an individual who is moving. The apparatus 50 may also be located in a mode of transport including, but not limited to, a car, a truck, a taxi, a bus, a train, a boat, an airplane, a bicycle, a motorcycle or any similar suitable mode of transport.

[0055] Some or further apparatuses 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 may include additional communication devices and communication devices of various types.

[0056] The communication devices may communicate using various transmission technologies including, but not limited to, code division multiple access (CDMA), global systems for mobile communications (GSM), universal mobile telecommunications system (UMTS), time divisional multiple access (TDMA), frequency division multiple access (FDMA), transmission control protocol-internet protocol (TCP-IP), short messaging service (SMS), multimedia messaging service (MMS), email, instant messaging service (IMS), Bluetooth, IEEE 802.11 and any similar wireless communication technology. A communications device involved in implementing various embodiments of the present invention may communicate using various media including, but not limited to, radio, infrared, laser, cable connections, and any suitable connection.

[0057] With respect to Figure 4a, a block diagram of a video encoder suitable for carrying out embodiments of the invention is shown. Furthermore, with respect to Figures 5a and 5b, the operation of the encoder exemplifying embodiments of the invention specifically with respect to construction of the list of candidate predictions is shown as a flow diagram.

[0058] Figure 4a shows the encoder as comprising a pixel predictor 302, prediction error encoder 303 and prediction error decoder 304. Figure 4a also shows an embodiment of the pixel

predictor 302 as comprising an inter-predictor 306, an intra-predictor 308, a mode selector 310, a filter 316, and a reference frame memory 318. In this embodiment the mode selector 310 comprises a block processor 381 and a cost evaluator 382. The encoder may further comprise an entropy encoder 330 for entropy encoding the bit stream.

[0059] Figure 4b depicts an embodiment of the inter predictor 306. The inter predictor 306 comprises a reference frame selector 360 for selecting reference frame or frames, a motion vector definer 361, a prediction list modifier 363 and a motion vector selector 364. These elements or some of them may be part of a prediction processor 362 or they may be implemented by using other means.

[0060] The pixel predictor 302 receives the image 300 to be encoded at both the inter-predictor 306 (which determines the difference between the image and a motion compensated reference frame 318) and the intra-predictor 308 (which determines a prediction for an image block based only on the already processed parts of the current frame or picture). The output of both the inter-predictor and the intra-predictor may be passed to the mode selector 310. The intra-predictor 308 may have more than one intra-prediction modes. Hence, each mode may perform the intra-prediction and provide the predicted signal to the mode selector 310. The mode selector 310 also receives a copy of the image 300.

[0061] The mode selector 310 determines which encoding mode to use to encode the current block. If the mode selector 310 decides to use an inter-prediction mode it will pass the output of the inter-predictor 306 to the output of the mode selector 310. If the mode selector 310 decides to use an intra-prediction mode it will pass the output of one of the intra-predictor modes to the output of the mode selector 310.

[0062] The output of the mode selector is passed to a first summing device 321. The first summing device may subtract the pixel predictor 302 output from the image 300 to produce a first prediction error signal 320 which is input to the prediction error encoder 303.

[0063] The pixel predictor 302 further receives from a preliminary reconstructor 339 the combination of the prediction representation of the image block 312 and the output 338 of the prediction error decoder 304. The preliminary reconstructed image 314 may be passed to the intra-predictor 308 and to a filter 316. The filter 316 receiving the preliminary representation may filter the preliminary representation and output a final reconstructed image 340 which may be saved in a reference frame memory 318. The reference frame memory 318 may be connected

to the inter-predictor 306 to be used as the reference image against which the future image 300 is compared in inter-prediction operations.

[0064] The operation of the pixel predictor 302 may be configured to carry out any known pixel prediction algorithm known in the art.

[0065] The pixel predictor 302 may also comprise a filter 385 to filter the predicted values before outputting them from the pixel predictor 302.

[0066] The operation of the prediction error encoder 302 and prediction error decoder 304 will be described hereafter in further detail. In the following examples the encoder generates images in terms of 16x16 pixel macroblocks which go to form the full image or picture. Thus, for the following examples the pixel predictor 302 outputs a series of predicted macroblocks of size 16x16 pixels and the first summing device 321 outputs a series of 16x16 pixel residual data macroblocks which may represent the difference between a first macro-block in the image 300 against a predicted macro-block (output of pixel predictor 302). It would be appreciated that other size macro blocks may be used.

[0067] The prediction error encoder 303 comprises a transform block 342 and a quantizer 344. The transform block 342 transforms the first prediction error signal 320 to a transform domain. The transform is, for example, the DCT transform. The quantizer 344 quantizes the transform domain signal, e.g. the DCT coefficients, to form quantized coefficients.

[0068] The prediction error decoder 304 receives the output from the prediction error encoder 303 and performs the opposite processes of the prediction error encoder 303 to produce a decoded prediction error signal 338 which when combined with the prediction representation of the image block 312 at the second summing device 339 produces the preliminary reconstructed image 314. The prediction error decoder may be considered to comprise a dequantizer 346, which dequantizes the quantized coefficient values, e.g. DCT coefficients, to reconstruct the transform signal and an inverse transformation block 348, which performs the inverse transformation to the reconstructed transform signal wherein the output of the inverse transformation block 348 contains reconstructed block(s). The prediction error decoder may also comprise a macroblock filter (not shown) which may filter the reconstructed macroblock according to further decoded information and filter parameters.

[0069] In the following the operation of an example embodiment of the inter predictor 306 will be described in more detail. The inter predictor 306 receives the current block for inter

prediction. It is assumed that for the current block there already exists one or more neighbouring blocks which have been encoded and motion vectors have been defined for them. For example, the block on the left side and/or the block above the current block may be such blocks. Spatial motion vector predictions for the current block can be formed e.g. by using the motion vectors of the encoded neighbouring blocks and/or of non-neighbour blocks in the same slice or frame, using linear or non-linear functions of spatial motion vector predictions, using a combination of various spatial motion vector predictors with linear or non-linear operations, or by any other appropriate means that do not make use of temporal reference information. It may also be possible to obtain motion vector predictors by combining both spatial and temporal prediction information of one or more encoded blocks. These kinds of motion vector predictors may also be called as spatio-temporal motion vector predictors.

[0070] Reference frames used in encoding the neighbouring blocks have been stored to the reference frame memory 404. The reference frames may be short term references or long term references and each reference frame may have a unique index indicative of the location of the reference frame in the reference frame memory. When a reference frame is no longer used as a reference frame it may be removed from the reference frame memory or marked as a non-reference frame wherein the storage location of that reference frame may be occupied for a new reference frame. In addition to the reference frames of the neighbouring blocks the reference frame selector 360 may also select one or more other frames as potential reference frames and store them to the reference frame memory.

[0071] Motion vector information of encoded blocks is also stored into the memory so that the inter predictor 306 is able to retrieve the motion vector information when processing motion vector candidates for the current block.

[0072] In some embodiments the motion vectors are stored into one or more lists. For example, motion vectors of uni-directionally predicted frames (e.g. P-frames) may be stored to a list called as list 0. For bi-directionally predicted frames (e.g. B-frames) there may be two lists (list 0 and list 1) and for multi-predicted frames there may be more than two lists. Reference frame indices possibly associated with the motion vectors may also be stored in one or more lists.

[0073] In some embodiments there may be two or more motion vector prediction procedures and each procedure may have its own candidate set creation process. In one

procedure, only the motion vector values are used. In another procedure, which may be called as a Merge Mode, each candidate element may comprise 1) The information whether 'block was uni-predicted using only list0' or 'block was uni-predicted using only list1' or 'block was bi-predicted using list0 and list1' 2) motion vector value for list0 3) Reference picture index in list0 4) motion vector value for list1 5) Reference picture index list1. Therefore, whenever two prediction candidates are to be compared, not only the motion vector values are compared, but also the five values mentioned above may be compared to determine whether they correspond with each other or not. On the other hand, if any of the comparisons indicate that the prediction candidates do not have equal motion information, no further comparisons need be performed.

[0074] The motion vector definer 361 defines candidate motion vectors for the current frame by using one or more of the motion vectors of one or more neighbour blocks and/or other blocks of the current block in the same frame and/or co-located blocks and/or other blocks of the current block in one or more other frames. These candidate motion vectors can be called as a set of candidate predictors or a predictor set. Each candidate predictor thus represents the motion vector of one or more already encoded block. In some embodiments the motion vector of the candidate predictor is set equal to the motion vector of a neighbour block for the same list if the current block and the neighbour block refer to the same reference frames for that list. Also for temporal prediction there may be one or more previously encoded frames wherein motion vectors of a co-located block or other blocks in a previously encoded frame can be selected as candidate predictors for the current block. The temporal motion vector predictor candidate can be generated by any means that make use of the frames other than the current frame.

[0075] The candidate motion vectors can also be obtained by using more than one motion vector of one or more other blocks such as neighbour blocks of the current block and/or co-located blocks in one or more other frames. As an example, any combination of the motion vector of the block to the left of the current block, the motion vector of the block above the current block, and the motion vector of the block at the up-right corner of the current block may be used (i.e. the block to the right of the block above the current block). The combination may be a median of the motion vectors or calculated by using other formulas. For example, one or more of the motion vectors to be used in the combination may be scaled by a scaling factor, an offset may be added, and/or a constant motion vector may be added. In some embodiments the combined motion vector is based on both temporal and spatial motion vectors, e.g. the motion

vector of one or more of the neighbour block or other block of the current block and the motion vector of a co-located block or other block in another frame.

[0076] If a neighbour block does not have any motion vector information a default motion vector such as a zero motion vector may be used instead.

[0077] Figure 9 illustrates an example of a coding unit 900 and some neighbour blocks 901—905 of the coding unit. As can be seen from Figure 9, if the coding unit 900 represents the current block, the neighbouring blocks 901—905 labelled A0, A1, B0, B1 and B2 could be such neighbour blocks which may be used when obtaining the candidate motion vectors.

[0078] Creating additional or extra motion vector predictions based on previously added predictors may be needed when the current number of candidates is limited or insufficient. This kind of creating additional candidates can be performed by combining previous two predictions and/or processing one previous candidate by scaling or adding offset and/or adding a zero motion vector with various reference indices. Hence, the motion vector definer 361 may examine how many motion vector candidates can be defined and how many potential candidate motion vectors exist for the current block. If the number of potential motion vector candidates is smaller than a threshold, the motion vector definer 361 may create additional motion vector predictions.

[0079] In some embodiments the combined motion vector can be based on motion vectors in different lists. For example, one motion vector may be defined by combining one motion vector from the list 0 and one motion vector from the list 1 e.g. when the neighbouring or co-located block is a bi-directionally predicted block and there exists one motion vector in the list 0 and one motion vector in the list 1 for the bi-directionally predicted block.

[0080] To distinguish the current block from the encoded/decoded blocks the motion vectors of which are used as candidate motion vectors, those encoded/decoded blocks are also called as reference blocks in this application.

[0081] In some embodiments not only the motion vector information of the reference block(s) is obtained (e.g. by copying) but also a reference index of the reference block in the reference picture list may be copied to the candidate list. The information whether the block was uni-predicted using only list0 or the block was uni-predicted using only list1 or the block was bi-predicted using list0 and list1 may also be copied. The candidate list may also be called as a candidate set or a set of motion vector prediction candidates.

[0082] Figure 6a illustrates an example of spatial and temporal prediction of a prediction unit. There is depicted the current block 601 in the frame 600 and a neighbour block 602 which already has been encoded. The motion vector definer 361 has defined a motion vector 603 for the neighbour block 602 which points to a block 604 in the previous frame 605. This motion vector can be used as a potential spatial motion vector prediction 610 for the current block. Figure 6a depicts that a co-located block 606 in the previous frame 605, i.e. the block at the same location than the current block but in the previous frame, has a motion vector 607 pointing to a block 609 in another frame 608. This motion vector 607 can be used as a potential temporal motion vector prediction-611 for the current frame.

[0083] Figure 6b illustrates another example of spatial and temporal prediction of a prediction unit. In this example the block 606 of the previous frame 605 uses bi-directional prediction based on the block 609 of the frame preceding the frame 605 and on the block 612 succeeding the current frame 600. The temporal motion vector prediction for the current block 601 may be formed by using both the motion vectors 607, 614 or either of them.

[0084] The operation of the prediction list modifier 363 will now be described in more detail with reference to the flow diagram of Figures 5a and 5b. The prediction list modifier 363 initializes a motion vector prediction list to default values in block 500 of Figure 5a. The prediction list modifier 363 may also initialize a list index to an initial value such as zero. Then, in block 501 the prediction list modifier checks whether there are any motion vector candidates to process. If there is at least one motion vector candidate in the predictor set for processing, the prediction list modifier 363 generates the next motion vector candidate which may be a temporal motion vector or a spatial motion vector. The comparison can be an identity/equivalence check or comparing the (absolute) difference against a threshold or any other similarity metric.

[0085] In the following, a merge process for motion information coding according to an example embodiment will be described in more detail. The encoder creates a list of motion prediction candidates from which one of the candidates is to be signalled as the motion information for the current coding unit or prediction unit. The motion prediction candidates may consist of several spatial motion predictions and a temporal motion prediction. The spatial candidates can be obtained from the motion information of e.g. the spatial neighbour blocks A0, A1, B0, B1, B2, whose motion information is used as spatial candidate motion predictions. The temporal motion prediction candidate may be obtained by processing the motion of a block in a

frame other than the current frame. In this example embodiment, the encoder operations to construct the merge list for the spatial candidates may include the following. The operations may be carried out by the prediction list modifier 363, for example.

[0086] A maximum number of spatial motion prediction candidates to be included in the merge list may be defined. This maximum number may have been stored, for example, to the memory 58 of the apparatus 50, or to another appropriate place. It is also possible to determine the maximum number by using other means, or it may be determined in the software of the encoder of the apparatus 50.

[0087] In some embodiments the maximum number of spatial motion prediction candidates to be included in the merge list is four but in some embodiments the maximum number may be less than four or greater than four.

[0088] In this example the spatial motion prediction candidates are the spatial neighbour blocks A0, A1, B0, B1, B2. The spatial motion vector prediction candidate A1 is located on the left side of the prediction unit when the encoding/decoding order is from left to right and from top to bottom of the frame, slice or another entity to be encoded/decoded. Respectively, the spatial motion vector prediction candidate B1 is located above the prediction unit. third; the spatial motion vector prediction candidate B0 is on the right side of the spatial motion vector prediction candidate B1; the spatial motion vector prediction candidate A0 is below the spatial motion vector prediction candidate A1; and the spatial motion vector prediction candidate B2 is located on the same column than spatial motion vector prediction candidate A1 and on the same row than the spatial motion vector prediction candidate B1. In other words, the spatial motion vector prediction candidate B2 is cornerwise neighbouring the prediction unit as can be seen e.g. from Figure 9.

[0089] These spatial motion prediction candidates can be processed in a predetermined order, for example, A1, B1, B0, A0 and B2. The first spatial motion prediction candidate to be selected for further examination is thus A1. Before further examination is performed for the selected spatial motion prediction candidate, it may be determined whether the merge list already contains a maximum number of spatial motion prediction candidates. Hence, the prediction list modifier 363 compares 502 the number of spatial motion prediction candidates in the merge list with the maximum number, and if the number of spatial motion prediction candidates in the merge list is not less than the maximum number, the selected spatial motion prediction candidate

is not included in the merge list and the process of constructing the merge list can be stopped 526. On the other hand, if the number of spatial motion prediction candidates in the merge list is less than the maximum number, a further analyses of the selected spatial motion prediction candidate is performed (blocks 504-522).

[0090] For all the spatial motion prediction candidates for which the further analyses is to be performed, some or all of the following conditions below may be tested for determining whether to include the spatial motion prediction candidate in the merge list.

[0091] The prediction list modifier 363 examines 504 if the prediction unit or block covering the spatial motion prediction candidate block is not available for motion prediction. If so, the candidate is not included in the merge list. The reason that the block is not available may be that the block is either coded in intra mode or resides in a different slice or outside of the picture area.

[0092] In addition to the common conditions above, for each spatial motion prediction candidate, if any of the following conditions holds, then the candidate is not included in the merge list, otherwise, it is included.

[0093] The prediction list modifier 363 determines 506 which spatial motion prediction candidate of the set of spatial motion prediction candidates is in question. If the spatial motion prediction candidate is the block A1, one or more of the following conditions may be examined 508, 510 to determine whether to include this spatial motion prediction candidate in the merge list or not. If the current coding unit 100 is vertically split into two rectangle prediction units 103, 104 as depicted in Figure 10b and the current prediction unit is the second prediction unit 104 in the coding/decoding order (508), this spatial motion prediction candidate is not included in the merge list. If the current coding unit 100 is not vertically split into two rectangle prediction units but it is horizontally split into two rectangle prediction units 101, 102 as depicted in Figure 10a and the current prediction unit is the second prediction unit in the coding/decoding order and the block A1 has the same motion information as the block B1 (510), this spatial motion prediction candidate (block A1) is not included in the merge list. In the example of Figure 10a the second prediction unit is the lower prediction unit 102 of the coding unit 100 and in the example of Figure 10b the second prediction unit is the rightmost prediction unit 104 of the coding unit 100. If none of the conditions above is fulfilled the block A1 is included in the merge list as a spatial motion prediction candidate (524).

[0094] If the spatial motion prediction candidate is the block B1, one or more of the following conditions may be examined 512, 514 to determine whether to include this spatial motion prediction candidate in the merge list or not. If the current coding unit 100 is horizontally split into two rectangle prediction units 101, 102 as depicted in Figure 10a and the current prediction unit is the second prediction unit 104 in the coding/decoding order (512), this spatial motion prediction candidate is not included in the merge list. If the current coding unit 100 is not horizontally split into two rectangle prediction units and if the block B1 has the same motion information than the block A1 (514), this spatial motion prediction candidate (block B1) is not included in the merge list. If none of the conditions above is fulfilled the block B1 is included in the merge list as a spatial motion prediction candidate (524).

[0095] If the spatial motion prediction candidate is the block B0, this spatial motion prediction candidate is not included in the merge list if the block B0 has the same motion information than the block B1 (516). Otherwise, if the number of spatial motion prediction candidates in the merge list is less than the maximum number of spatial motion prediction candidates, this spatial motion prediction candidate (block B0) is included in the merge list (524).

[0096] If the spatial motion prediction candidate is the block A0, this spatial motion prediction candidate is not included in the merge list if the block A0 has the same motion information than the block A1 (518). Otherwise, if the number of spatial motion prediction candidates in the merge list is less than the maximum number of spatial motion prediction candidates, this spatial motion prediction candidate (block A0) is included in the merge list (524).

[0097] If the spatial motion prediction candidate is the block B2, this spatial motion prediction candidate is not included in the merge list if the maximum number of spatial motion prediction candidates is four and the other blocks A0, A1, B0, and B1 are all decided to be included in the merge list (520). Otherwise, if the number of spatial motion prediction candidates in the merge list is less than the maximum number of spatial motion prediction candidates, the block B2 is not included in the merge list if the block B2 has the same motion information than the block B1 or the block A1 (522).

[0098] Then, after processing the blocks A1, B1, B0, A0 and B2 and including a subset of them in the merge list based on the above described conditions, no more redundancy check

between these candidates are performed and remaining temporal motion prediction candidate and/or other possible additional candidates may be processed.

[0099] Comparing two blocks whether they have the same motion may be performed by comparing all the elements of the motion information, namely 1) The information whether 'the prediction unit is uni-predicted using only reference picture list0' or 'the prediction unit is uni-predicted using only reference picture list1' or 'the prediction unit is bi-predicted using both reference picture list0 and list1' 2) Motion vector value corresponding to the reference picture list0 3) Reference picture index in the reference picture list0 4) Motion vector value corresponding to the reference picture list1 5) Reference picture index in the reference picture list1.

[00100] In some embodiments similar restrictions for comparing candidate pairs can be applied if the current coding unit is coded/decoded by splitting into four or any number of prediction units.

[00101] The maximum number of merge list candidates can be any non-zero value. In the example above the merger list candidates were the spatial neighbour blocks A0, A1, B0, B1, B2 and the temporal motion prediction candidate, but there may be more than one temporal motion prediction candidate and also other spatial motion prediction candidates than the spatial neighbour blocks. In some embodiments there may also be other spatial neighbour blocks than the blocks A0, A1, B0, B1, B2.

[00102] It is also possible that the maximum number of spatial motion prediction candidates included in the list can be different than four.

[00103] In some embodiments the maximum number of merge list candidates and maximum number of spatial motion prediction candidates included in the list can depend on whether a temporal motion vector candidate is included in the list or not.

[00104] A different number of spatial motion prediction candidates located at various locations in the current frame can be processed. The locations can be the same as or different than A1, B1, B0, A0 and B2.

[00105] The decision of including which spatial motion prediction candidates in the list can be realized in two steps. In the first step, some of the candidates are eliminated by checking whether the candidate block is available and/or the candidate block's prediction mode is intra and/or whether the current block is a second prediction unit of a coding unit coded with two

prediction units and the candidate has the same motion with the first prediction unit. In the second step, remaining candidates are examined and some or all of them are included in the merge list. The examination in the second step does not include comparing motion information of each possible candidate pair but includes a subset of the possible comparison combinations.

[00106] The decisions for the candidates can be taken in any order of A1, B1, B0, A0 and B2 or independently in parallel.

[00107] For each candidate and/or a subset of the candidates, the following conditions may also be checked: Whether the candidate block has the same motion as the first prediction unit of the current coding unit when the current coding unit is split into two rectangle prediction units and the current prediction unit is the second prediction unit in the coding/decoding order.

[00108] Additional conditions related to various properties of current and/or previous slices and/or current and/or neighbour blocks can be utilized for determining whether to include a candidate in the list.

[00109] Motion comparison can be realized by comparing a subset of the whole motion information. For example, only the motion vector values for some or all reference picture lists and/or reference indices for some or all reference picture lists and/or an identifier value assigned to each block to represent its motion information can be compared. The comparison can be an identity or an equivalence check or comparing the (absolute) difference against a threshold or any other similarity metric.

[00110] Conditions for deciding whether a candidate is to be included in the list can include motion information comparison with any subset of the candidates as long as not all possible candidate pairs are compared eventually.

[00111] Deciding whether a temporal motion vector candidate is to be included in the list can be based on comparing its motion information with motion information of a subset of the spatial motion vector prediction candidates.

[00112] When comparing motion information of two blocks, motion information of additional blocks can be considered too. For example, when comparing the block B2 and the block A1, all the blocks between the block B2 and the block A1 (illustrated in Figure 12) are checked whether they have the same motion; and when comparing the block B2 and the block B1, all the blocks between the block B2 and the block B1 (illustrated in Figure 12) are checked whether they have the same motion. This embodiment can be implemented so that the right-most

block of each prediction unit or all blocks of each prediction unit may store the information of how many consecutive blocks to the above have the same motion information. Also the bottom-most block of each prediction unit or all blocks of each prediction unit may store the information of how many consecutive blocks to the left have the same motion information. Using this information the condition for not including B0 in the list can be realized by checking if the number of consecutive blocks with the same motion to the left of B0 is greater than 0. The condition for not including A0 in the list can be realized by checking if the number of consecutive blocks with same motion to the above of A0 is greater than 0. The conditions for not including B2 can be modified as follows:

[00113] It is not examined whether the block B2 has same motion as the block B1 or whether the block B2 has same motion as the block A1, but how many consecutive blocks exists to the left of the block B1 with the same motion than the block B1 and/or how many consecutive blocks exist above the block A1 with the same motion. If the number of consecutive blocks with the same motion to the left of the block B1 is greater than the number of blocks between B2 and B1, or if the number of consecutive blocks with the same motion above the block A1 is greater than the number of blocks between the block B2 and the block A1, the block B2 is not included in the merge list.

[00114] If the above implementation is used, the value of how many consecutive blocks to the left/above have the same motion information can be determined by direct comparison of motion information or checking the prediction mode and/or the merge index if the block employs a merge process.

[00115] When coding/decoding the selected merge index, the information whether the merge process is employed for coding/decoding a Skip mode coding unit or an Inter Merge mode prediction unit can be taken into account. For example, if a context adaptive binary arithmetic coder (CABAC) is used for entropy coding/decoding, different contexts can be used for the bins depending on the coding mode (Skip mode or inter merge mode) of the current block. Furthermore, assigning two contexts depending on whether the merge process is employed in a Skip mode coding unit or an inter Merge mode prediction unit can be applied for only the most significant bin of the merge index.

[00116] During the process of removal of redundant candidates, comparison between motion vector predictor candidates can also be based on any other information than the motion

vector values. For example, it can be based on linear or non-linear functions of motion vector values, coding or prediction types of the blocks used to obtain the motion information, block size, the spatial location in the frame/(largest) coding unit/macroblock, the information whether blocks share the same motion with a block, the information whether blocks are in the same coding/prediction unit, etc.

[00117] The following pseudo code illustrates an example embodiment of the invention for constructing the merging list.

[00118] Inputs to this process are

- a luma location (xP, yP) specifying the top-left luma sample of the current prediction unit relative to the top-left sample of the current picture;
- variables specifying the width and the height of the prediction unit for luma, nPSW and nPSH; and
- a variable PartIdx specifying the index of the current prediction unit within the current coding unit.

[00119] Outputs of this process are (with N being replaced by A₀, A₁, B₀, B₁ or B₂ and with X being replaced by 0 or 1)

- the availability flags availableFlagN of the neighbouring prediction units,
- the reference indices refIdxLXN of the neighbouring prediction units,
- the prediction list utilization flags predFlagLXN of the neighbouring prediction units,
- the motion vectors mvLXN of the neighbouring prediction units.

[00120] For the derivation of availableFlagN, with N being A₀, A₁, B₀, B₁ or B₂ and (xN, yN) being (xP-1, yP + nPSH), (xP-1, yP + nPSH-1), (xP + nPSW, yP-1), (xP+nPSW-1, yP-1) or (xP-1, yP-1), the following applies.

- If one of the following conditions is true, the availableFlagN is set equal to 0, both components mvLXN are set equal to 0, refIdxLXN and predFlagLX[xN, yN] of the prediction unit covering luma location (xN, yN) are assigned respectively to mvLXN, refIdxLXN and predFlagLXN.
 - N is equal to B₂ and availableFlagA₀ + availableFlagA₁ + availableFlagB₀ + availableFlagB₁ is equal to 4.
 - The prediction unit covering luma location (xN, yN) is not available or PredMode is MODE_INTRA.

- N is equal to A1 and PartMode of the current prediction unit is PART_Nx2N or PART_nLx2N or PART_nRx2N and PartIdx is equal to 1.
- N is equal to A1 and PartMode of the current prediction unit is PART_2NxN or PART_2NxN_U or PART_2NxN_D and PartIdx is equal to 1 and the prediction units covering luma location (xP+nPSW-1, yP-1) (N = B1) and luma location (xN, yN) (Cand. N) have identical motion parameters:
 - mvLX[xP+nPSW-1, yP-1] == mvLX[xN, yN]
 - refIdxLX[xP+nPSW-1, yP-1] == refIdxLX[xN, yN]
 - predFlagLX[xP+nPSW-1, yP-1] == predFlagLX[xN, yN]
- N is equal to B1 and PartMode of the current prediction unit is 2NxN or PART_2NxN_U or PART_2NxN_D and PartIdx is equal to 1.
- N is equal to B1 and the prediction units covering luma location (xP-1, yP+nPSH-1) (N = A1) and luma location (xN, yN) (Cand. N) have identical motion parameters:
 - mvLX[xP-1, yP+nPSH-1] == mvLX[xN, yN]
 - refIdxLX[xP-1, yP+nPSH-1] == refIdxLX[xN, yN]
 - predFlagLX[xP-1, yP+nPSH-1] == predFlagLX[xN, yN]
- N is equal to B0 and the prediction units covering luma location (xP+nPSW-1, yP-1) (N = B1) and luma location (xN, yN) (Cand. N) have identical motion parameters:
 - mvLX[xP+nPSW-1, yP-1] == mvLX[xN, yN]
 - refIdxLX[xP+nPSW-1, yP-1] == refIdxLX[xN, yN]
 - predFlagLX[xP+nPSW-1, yP-1] == predFlagLX[xN, yN]
- N is equal to A0 and the prediction units covering luma location (xP-1, yP+nPSH-1) (N = A1) and luma location (xN, yN) (Cand. N) have identical motion parameters:
 - mvLX[xP-1, yP+nPSH-1] == mvLX[xN, yN]
 - refIdxLX[xP-1, yP+nPSH-1] == refIdxLX[xN, yN]
 - predFlagLX[xP-1, yP+nPSH-1] == predFlagLX[xN, yN]
- N is equal to B2 and the prediction units covering luma location (xP+nPSW-1, yP-1) (N = B1) and luma location (xN, yN) (Cand. N) have identical motion parameters:
 - mvLX[xP+nPSW-1, yP-1] == mvLX[xN, yN]
 - refIdxLX[xP+nPSW-1, yP-1] == refIdxLX[xN, yN]
 - predFlagLX[xP+nPSW-1, yP-1] == predFlagLX[xN, yN]

- N is equal to B2 and the prediction units covering luma location (xP-1, yP+nPSH-1) (N = A1) and luma location (xN, yN) (Cand. N) have identical motion parameters:
 - mvLX[xP-1, yP+nPSH-1] == mvLX[xN, yN]
 - refIdxLX[xP-1, yP+nPSH-1] == refIdxLX[xN, yN]
 - predFlagLX[xP-1, yP+nPSH-1] == predFlagLX[xN, yN]
- PartMode of the current prediction unit is PART_NxN and PartIdx is equal to 3 and the prediction units covering luma location (xP-1, yP) (PartIdx = 2) and luma location (xP-1, yP-1) (PartIdx = 0) have identical motion parameters:
 - mvLX[xP-1, yP] == mvLX[xP-1, yP-1]
 - refIdxLX[xP-1, yP] == refIdxLX[xP-1, yP-1]
 - predFlagLX[xP-1, yP] == predFlagLX[xP-1, yP-1]and the prediction units covering luma location (xP, yP-1) (PartIdx = 1) and luma location (xN, yN) (Cand. N) have identical motion parameters:
 - mvLX[xP, yP-1] == mvLX[xN, yN]
 - refIdxLX[xP, yP-1] == refIdxLX[xN, yN]
 - predFlagLX[xP, yP-1] == predFlagLX[xN, yN]
- PartMode of the current prediction unit is PART_NxN and PartIdx is equal to 3 and the prediction units covering luma location (xP, yP-1) (PartIdx = 1) and luma location (xP-1, yP-1) (PartIdx = 0) have identical motion parameters:
 - mvLX[xP, yP-1] == mvLX[xP-1, yP-1]
 - refIdxLX[xP, yP-1] == refIdxLX[xP-1, yP-1]
 - predFlagLX[xP, yP-1] == predFlagLX[xP-1, yP-1]and the prediction units covering luma location (xP-1, yP) (PartIdx = 2) and luma location (xN, yN) (Cand. N) have identical motion parameters:
 - mvLX[xP-1, yP] == mvLX[xN, yN]
 - refIdxLX[xP-1, yP] == refIdxLX[xN, yN]
 - predFlagLX[xP-1, yP] == predFlagLX[xN, yN]
- Otherwise, availableFlagN is set equal to 1 and the variables mvLX[xN, yN], refIdxLX[xN, yN] and predFlagLX[xN, yN] of the prediction unit covering luma location (xN, yN) are assigned respectively to mvLXN, refIdxLXN and predFlagLXN.

[00121] For the motion vector predictor candidate list generation process, each list candidate can include more information than the motion vector value, such as the reference lists used, the reference frames used in each list and motion vector for each list.

[00122] When all motion vector candidates have been examined, one motion vector is selected to be used as the motion vector for the current block. The motion vector selector 364 may examine different motion vectors in the list and determine which motion vector provides the most efficient encoding result, or the selection of the motion vector may be based on to other criteria as well. Information of the selected motion vector is provided for the mode selector for encoding and transmission to the decoder or for storage when the mode selector determines to use inter prediction for the current block. The information may include the index of the motion vector in the list, and/or motion vector parameters or other appropriate information.

[00123] The selected motion vector and the block relating to the motion vector is used to generate the prediction representation of the image block 312 which is provided as the output of the mode selector. The output may be used by the first summing device 321 to produce the first prediction error signal 320, as was described above.

[00124] The selected motion vector predictor candidate can be modified by adding a motion vector difference or can be used directly as the motion vector of the block. Moreover, after the motion compensation is performed by using the selected motion vector predictor candidate, the residual signal of the block can be transform coded or skipped to be coded.

[00125] Although the embodiments above have been described with respect to the size of the macroblock being 16x16 pixels, it would be appreciated that the methods and apparatus described may be configured to handle macroblocks of different pixel sizes.

[00126] In the following the operation of an example embodiment of the decoder 600 is depicted in more detail with reference to Figure 7.

[00127] At the decoder side similar operations are performed to reconstruct the image blocks. Figure 7 shows a block diagram of a video decoder 700 suitable for employing embodiments of the invention and Figures 8a and 8b show a flow diagram of an example of a method in the video decoder. The bitstream to be decoded may be received from the encoder, from a network element, from a storage medium or from another source. The decoder is aware of the structure of the bitstream so that it can determine the meaning of the entropy coded codewords and may decode the bitstream by an entropy decoder 701 which performs entropy

decoding on the received signal. The entropy decoder thus performs the inverse operation to the entropy encoder 330 of the encoder described above. The entropy decoder 701 outputs the results of the entropy decoding to a prediction error decoder 702 and a pixel predictor 704.

[00128] In some embodiments the entropy coding may not be used but another channel encoding may be in use, or the encoded bitstream may be provided to the decoder 700 without channel encoding. The decoder 700 may comprise a corresponding channel decoder to obtain the encoded codewords from the received signal.

[00129] The pixel predictor 704 receives the output of the entropy decoder 701. The output of the entropy decoder 701 may include an indication on the prediction mode used in encoding the current block. A predictor selector 714 within the pixel predictor 704 determines that an intra-prediction or an inter-prediction is to be carried out. The predictor selector 714 may furthermore output a predicted representation of an image block 716 to a first combiner 713. The predicted representation of the image block 716 is used in conjunction with the reconstructed prediction error signal 712 to generate a preliminary reconstructed image 718. The preliminary reconstructed image 718 may be used in the predictor 714 or may be passed to a filter 720. The filter 720, if used, applies a filtering which outputs a final reconstructed signal 722. The final reconstructed signal 722 may be stored in a reference frame memory 724, the reference frame memory 724 further being connected to the predictor 714 for prediction operations.

[00130] Also the prediction error decoder 702 receives the output of the entropy decoder 701. A dequantizer 792 of the prediction error decoder 702 may dequantize the output of the entropy decoder 701 and the inverse transform block 793 may perform an inverse transform operation to the dequantized signal output by the dequantizer 792. The output of the entropy decoder 701 may also indicate that prediction error signal is not to be applied and in this case the prediction error decoder produces an all zero output signal.

[00131] The decoder selects the 16x16 pixel residual macroblock to reconstruct. This residual macroblock is also called as a current block.

[00132] The decoder may receive information on the encoding mode used in encoding of the current block. The indication is decoded, when necessary, and provided to the reconstruction processor 791 of the predictor selector 714. The reconstruction processor 791 examines the indication and selects one of the intra-prediction mode(s), if the indication indicates that the

block has been encoded using intra-prediction, or the inter-prediction mode, if the indication indicates that the block has been encoded using inter-prediction.

[00133] For inter-prediction mode the reconstruction processor 791 may comprise one or more elements corresponding to the prediction processor 362 of the encoder, such as a motion vector definer, a prediction list modifier and/or a motion vector selector.

[00134] The reconstruction processor 791 initializes a motion vector prediction list to default values in block 800. As was the case in the encoding part, in this example the spatial motion prediction candidates are the spatial neighbour blocks A0, A1, B0, B1, B2 and these spatial motion prediction candidates are processed in the same predetermined order than in the encoder: A1, B1, B0, A0 and B2. The first spatial motion prediction candidate to be selected for further examination is thus A1. Before further examination is performed for the selected spatial motion prediction candidate, it is examined whether the merge list already contains a maximum number of spatial motion prediction candidates. If the number of spatial motion prediction candidates in the merge list is not less than the maximum number, the selected spatial motion prediction candidate is not included in the merge list and the process of constructing the merge list can be stopped 826. On the other hand, if the number of spatial motion prediction candidates in the merge list is less than the maximum number, a further analyses of the selected spatial motion prediction candidate is performed (blocks 804-822).

[00135] The decoder examines 804 if the prediction unit or block covering the spatial motion prediction candidate block is not available for motion prediction. If so, the candidate is not included in the merge list. The reason that the block is not available may be that the block is either coded in intra mode or resides in a different slice or outside of the picture area.

[00136] In addition to the common conditions above, for each spatial motion prediction candidate, if any of the following conditions holds, then the candidate is not included in the merge list, otherwise, it is included.

[00137] The decoder determines 806 which spatial motion prediction candidate of the set of spatial motion prediction candidates is in question. If the spatial motion prediction candidate is the block A1, one or more of the following conditions may be examined 808, 810 to determine whether to include this spatial motion prediction candidate in the merge list or not. If the current coding unit 100 is vertically split into two rectangle prediction units 103, 104 as depicted in Figure 10b and the current prediction unit is the second prediction unit 104 in the

coding/decoding order (808), this spatial motion prediction candidate is not included in the merge list. If the current coding unit 100 is not vertically split into two rectangle prediction units but it is horizontally split into two rectangle prediction units 101, 102 as depicted in Figure 10a and the current prediction unit is the second prediction unit in the coding/decoding order and the block A1 has the same motion information as the block B1 (810), this spatial motion prediction candidate (block A1) is not included in the merge list. In the example of Figure 10a the second prediction unit is the lower prediction unit 102 of the coding unit 100 and in the example of Figure 10b the second prediction unit is the rightmost prediction unit 104 of the coding unit 100. If none of the conditions above is fulfilled the block A1 is included in the merge list as a spatial motion prediction candidate (824).

[00138] If the spatial motion prediction candidate is the block B1, one or more of the following conditions may be examined 812, 814 to determine whether to include this spatial motion prediction candidate in the merge list or not. If the current coding unit 100 is horizontally split into two rectangle prediction units 101, 102 as depicted in Figure 10a and the current prediction unit is the second prediction unit 104 in the coding/decoding order (812), this spatial motion prediction candidate is not included in the merge list. If the current coding unit 100 is not horizontally split into two rectangle prediction units and if the block B1 has the same motion information than the block A1 (814), this spatial motion prediction candidate (block B1) is not included in the merge list. If none of the conditions above is fulfilled the block B1 is included in the merge list as a spatial motion prediction candidate (824).

[00139] If the spatial motion prediction candidate is the block B0, this spatial motion prediction candidate is not included in the merge list if the block B0 has the same motion information than the block B1 (816). Otherwise, if the number of spatial motion prediction candidates in the merge list is less than the maximum number of spatial motion prediction candidates, this spatial motion prediction candidate (block B0) is included in the merge list (824).

[00140] If the spatial motion prediction candidate is the block A0, this spatial motion prediction candidate is not included in the merge list if the block A0 has the same motion information than the block A1 (818). Otherwise, if the number of spatial motion prediction candidates in the merge list is less than the maximum number of spatial motion prediction

candidates, this spatial motion prediction candidate (block A0) is included in the merge list (824).

[00141] If the spatial motion prediction candidate is the block B2, this spatial motion prediction candidate is not included in the merge list if the maximum number of spatial motion prediction candidates is four and the other blocks A0, A1, B0, and B1 are all decided to be included in the merge list (820). Otherwise, if the number of spatial motion prediction candidates in the merge list is less than the maximum number of spatial motion prediction candidates, the block B2 is not included in the merge list if the block B2 has the same motion information than the block B1 or the block A1 (822).

[00142] Then, after processing the blocks A1, B1, B0, A0 and B2 and including a subset of them in the merge list based on the above described conditions, no more redundancy check between these candidates are performed and remaining temporal motion prediction candidate and/or other possible additional candidates may be processed.

[00143] When the merge list has been constructed the decoder may use 828 the indication of the motion vector received from the encoder to select the motion vector for decoding the current block. The indication may be, for example, an index to the merge list.

[00144] Basically, after the reconstruction processor 791 has constructed the merge list, it would correspond with the merge list constructed by the encoder if the reconstruction processor 791 has the same information available than the encoder had. If some information has been lost during transmission the information from the encoder to the decoder, it may affect the generation of the merge list in the decoder 700.

[00145] The above examples describe the operation mainly in the merge mode but the encoder and decoder may also operate in other modes.

[00146] The embodiments of the invention described above describe the codec in terms of separate encoder and decoder apparatus in order to assist the understanding of the processes involved. However, it would be appreciated that the apparatus, structures and operations may be implemented as a single encoder-decoder apparatus/structure/operation. Furthermore in some embodiments of the invention the coder and decoder may share some or all common elements.

[00147] Although the above examples describe embodiments of the invention operating within a codec within an electronic device, it would be appreciated that the invention as described below may be implemented as part of any video codec. Thus, for example,

embodiments of the invention may be implemented in a video codec which may implement video coding over fixed or wired communication paths.

[00148] Thus, user equipment may comprise a video codec such as those described in embodiments of the invention above.

[00149] It shall be appreciated that the term user equipment is intended to cover any suitable type of wireless user equipment, such as mobile telephones, portable data processing devices or portable web browsers.

[00150] Furthermore elements of a public land mobile network (PLMN) may also comprise video codecs as described above.

[00151] In general, the various embodiments of the invention may be implemented in hardware or special purpose circuits, software, logic or any combination thereof. For example, some aspects may be implemented in hardware, while other aspects may be implemented in firmware or software which may be executed by a controller, microprocessor or other computing device, although the invention is not limited thereto. While various aspects of the invention may be illustrated and described as block diagrams, flow charts, or using some other pictorial representation, it is well understood that these blocks, apparatus, systems, techniques or methods described herein may be implemented in, as non-limiting examples, hardware, software, firmware, special purpose circuits or logic, general purpose hardware or controller or other computing devices, or some combination thereof.

[00152] The embodiments of this invention may be implemented by computer software executable by a data processor of the mobile device, such as in the processor entity, or by hardware, or by a combination of software and hardware. Further in this regard it should be noted that any blocks of the logic flow as in the Figures may represent program steps, or interconnected logic circuits, blocks and functions, or a combination of program steps and logic circuits, blocks and functions. The software may be stored on such physical media as memory chips, or memory blocks implemented within the processor, magnetic media such as hard disk or floppy disks, and optical media such as for example DVD and the data variants thereof, CD.

[00153] The memory may be of any type suitable to the local technical environment and may be implemented using any suitable data storage technology, such as semiconductor based memory devices, magnetic memory devices and systems, optical memory devices and systems, fixed memory and removable memory. The data processors may be of any type suitable to the

local technical environment, and may include one or more of general purpose computers, special purpose computers, microprocessors, digital signal processors (DSPs) and processors based on multi core processor architecture, as non limiting examples.

[00154] Embodiments of the inventions may be practiced in various components such as integrated circuit modules. The design of integrated circuits is by and large a highly automated process. Complex and powerful software tools are available for converting a logic level design into a semiconductor circuit design ready to be etched and formed on a semiconductor substrate.

[00155] Programs, such as those provided by Synopsys, Inc. of Mountain View, California and Cadence Design, of San Jose, California automatically route conductors and locate components on a semiconductor chip using well established rules of design as well as libraries of pre stored design modules. Once the design for a semiconductor circuit has been completed, the resultant design, in a standardized electronic format (e.g., Opus, GDSII, or the like) may be transmitted to a semiconductor fabrication facility or "fab" for fabrication.

[00156] The foregoing description has provided by way of exemplary and non-limiting examples a full and informative description of the exemplary embodiment of this invention. However, various modifications and adaptations may become apparent to those skilled in the relevant arts in view of the foregoing description, when read in conjunction with the accompanying drawings and the appended claims. However, all such and similar modifications of the teachings of this invention will still fall within the scope of this invention.

[00157] In the following some examples will be provided.

[00158] In some embodiments a method comprises:

receiving a block of pixels including a prediction unit; determining a set of spatial motion vector prediction candidates for the block of pixels; the spatial motion vector prediction candidates being provided with motion information;

selecting a first spatial motion vector prediction candidate from the set of spatial motion vector prediction candidates as a potential spatial motion vector prediction candidate to be included in a merge list for the prediction unit;

determining a subset of spatial motion vector predictions based on the location of the block associated with the first spatial motion vector prediction candidate;

comparing motion information of the first spatial motion vector prediction candidate with motion information of the spatial motion vector prediction candidate in the determined subset of spatial motion vector prediction candidates;

if at least one of the comparisons indicates that the motion vector information of the spatial motion vector prediction candidates correspond with each other, excluding the first spatial motion vector prediction candidate from the merge list.

[00159] In some embodiments the method comprises including neighbouring blocks of the received block of pixels in the set of spatial motion vector prediction candidates.

[00160] In some embodiments the method comprises constructing the set of spatial motion vector predictions by using motion vectors of one or more encoded blocks in a same frame than the block of pixels.

[00161] In some embodiments the method comprises selecting spatial motion vector prediction candidates from the set of spatial motion vector prediction candidates as the potential spatial motion vector prediction candidate in a predetermined order.

[00162] In some embodiments the method comprises comparing motion information of the potential spatial motion vector prediction candidate with motion information of at most one other spatial motion vector prediction candidate of the set of spatial motion vector prediction candidates.

[00163] In some embodiments the method comprises prediction unit and a second prediction unit; and if so, excluding the potential spatial motion vector prediction candidate from the merge list if the prediction unit is the second prediction unit.

[00164] In some embodiments the method comprises
determining a maximum number of spatial motion vector prediction candidates to be included in a merge list; and

limiting the number of spatial motion vector prediction candidates in the merge list smaller or equal to the maximum number .

[00165] In some embodiments the method comprises
examining, if the number of spatial motion vector prediction candidates in the merge list smaller than the maximum number;

if so, examining whether a prediction unit to which the potential spatial motion vector prediction candidate belongs is available for motion prediction;

if so, performing at least one of the following:

for the potential spatial motion vector prediction candidate on the left side of the prediction unit, excluding the potential spatial motion vector prediction candidate from the merge list if any of the following conditions are fulfilled:

- the received block of pixels is vertically divided into a first prediction unit and a second prediction unit, and the prediction unit is the second prediction unit;
- the received block of pixels is horizontally divided into a first prediction unit and a second prediction unit, and if the prediction unit is the second prediction unit, and the potential spatial motion vector prediction candidate has essentially similar motion information than the spatial motion vector prediction candidate above the prediction unit;

for the potential spatial motion vector prediction candidate above the prediction unit, excluding the potential spatial motion vector prediction candidate from the merge list if any of the following conditions are fulfilled:

- the received block of pixels is horizontally divided into a first prediction unit and a second prediction unit, and the prediction unit is the second prediction unit;
- the potential spatial motion vector prediction candidate has essentially similar motion information than the spatial motion vector prediction candidate on the left side of the prediction unit;

for the potential spatial motion vector prediction candidate, which is on the right side of the potential spatial motion vector prediction candidate above the prediction unit, excluding the potential spatial motion vector prediction candidate from the merge list if the potential spatial motion vector prediction candidate has essentially similar motion information than the spatial motion vector prediction candidate above the prediction unit;

for the potential spatial motion vector prediction candidate, which is below the potential spatial motion vector prediction candidate on the left side of the prediction unit, excluding the potential spatial motion vector prediction candidate from the merge list if the potential spatial motion vector prediction candidate has essentially similar motion information than the spatial motion vector prediction candidate on the left side of the prediction unit;

for the potential spatial motion vector prediction candidate cornerwise neighbouring the prediction unit, excluding the potential spatial motion vector prediction candidate from the merge list if any of the following conditions are fulfilled:

- all the other potential spatial motion vector prediction candidates have been included in the merge list;
- the potential spatial motion vector prediction candidate has essentially similar motion information than the spatial motion vector prediction candidate above the prediction unit;
- the potential spatial motion vector prediction candidate has essentially similar motion information than the spatial motion vector prediction candidate on the left side of the prediction unit.

[00166] In some embodiments the method comprises including a temporal motion prediction candidate into the merge list.

[00167] In some embodiments the method comprises selecting one motion vector prediction candidate from the merge list to represent a motion vector prediction for the block of pixels.

[00168] In some embodiments a method according to the second aspect comprises:
receiving an encoded block of pixels including a prediction unit;
determining a set of spatial motion vector prediction candidates for the encoded block of pixels; the spatial motion vector prediction candidates being provided with motion information;

selecting a first spatial motion vector prediction candidate from the set of spatial motion vector prediction candidates as a potential spatial motion vector prediction candidate to be included in a merge list for the prediction unit;

determining a subset of spatial motion vector predictions based on the location of the block associated with the first spatial motion vector prediction candidate;

comparing motion information of the first spatial motion vector prediction candidate with motion information of another spatial motion vector prediction candidate of the set of spatial motion vector prediction candidates;

if at least one of the comparisons indicates that the motion vector information of the spatial motion vector prediction candidates correspond with each other, excluding the first spatial motion vector prediction candidate from the merge list.

[00169] In some embodiments the method comprises including neighbouring blocks of the received encoded block of pixels in the set of spatial motion vector prediction candidates.

[00170] In some embodiments the method comprises constructing the set of spatial motion vector predictions by using motion vectors of one or more decoded blocks in a same frame than the received encoded block of pixels.

[00171] In some embodiments the method comprises selecting spatial motion vector prediction candidates from the set of spatial motion vector prediction candidates as the potential spatial motion vector prediction candidate in a predetermined order.

[00172] In some embodiments the method comprises comparing motion information of the potential spatial motion vector prediction candidate with motion information of at most one other spatial motion vector prediction candidate of the set of spatial motion vector prediction candidates.

[00173] In some embodiments the method comprises examining whether the received encoded block of pixels is divided into a first prediction unit and a second prediction unit; and if so, excluding the potential spatial motion vector prediction candidate from the merge list if the prediction unit is the second prediction unit.

[00174] In some embodiments the method comprises
determining a maximum number of spatial motion vector prediction candidates to be included in a merge list; and
limiting the number of spatial motion vector prediction candidates in the merge list smaller or equal to the maximum number.

[00175] In some embodiments the method comprises
examining, if the number of spatial motion vector prediction candidates in the merge list smaller than the maximum number;
if so, examining whether a prediction unit to which the potential spatial motion vector prediction candidate belongs is available for motion prediction;
if so, performing at least one of the following:

for the potential spatial motion vector prediction candidate on the left side of the prediction unit, excluding the potential spatial motion vector prediction candidate from the merge list if any of the following conditions are fulfilled:

- the received encoded block of pixels is vertically divided into a first prediction unit and a second prediction unit, and the prediction unit is the second prediction unit;
- the received encoded block of pixels is horizontally divided into a first prediction unit and a second prediction unit, and if the prediction unit is the second prediction unit, and the potential spatial motion vector prediction candidate has essentially similar motion information than the spatial motion vector prediction candidate above the prediction unit;

for the potential spatial motion vector prediction candidate above the prediction unit, excluding the potential spatial motion vector prediction candidate from the merge list if any of the following conditions are fulfilled:

- the received encoded block of pixels is horizontally divided into a first prediction unit and a second prediction unit, and the prediction unit is the second prediction unit;
- the potential spatial motion vector prediction candidate has essentially similar motion information than the spatial motion vector prediction candidate on the left side of the prediction unit;

for the potential spatial motion vector prediction candidate, which is on the right side of the potential spatial motion vector prediction candidate above the prediction unit, excluding the potential spatial motion vector prediction candidate from the merge list if the potential spatial motion vector prediction candidate has essentially similar motion information than the spatial motion vector prediction candidate above the prediction unit;

for the potential spatial motion vector prediction candidate, which is below the potential spatial motion vector prediction candidate on the left side of the prediction unit, excluding the potential spatial motion vector prediction candidate from the merge list if the potential spatial motion vector prediction candidate has essentially similar motion information than the spatial motion vector prediction candidate on the left side of the prediction unit;

for the potential spatial motion vector prediction candidate cornerwise neighbouring the prediction unit, excluding the potential spatial motion vector prediction candidate from the merge list if any of the following conditions are fulfilled:

- all the other potential spatial motion vector prediction candidates have been included in the merge list;
- the potential spatial motion vector prediction candidate has essentially similar motion information than the spatial motion vector prediction candidate above the prediction unit;
- the potential spatial motion vector prediction candidate has essentially similar motion information than the spatial motion vector prediction candidate on the left side of the prediction unit.

[00176] In some embodiments the method comprises including a temporal motion prediction candidate into the merge list.

[00177] In some embodiments the method comprises selecting one motion vector prediction candidate from the merge list to represent a motion vector prediction for the received encoded block of pixels.

[00178] In some embodiments an apparatus according to the third aspect comprises a processor and a memory including computer program code, the memory and the computer program code configured to, with the processor, cause the apparatus to:

- receive a block of pixels
- including a prediction unit;
- determining a set of spatial motion vector prediction candidates for the block of pixels;
- the spatial motion vector prediction candidates being provided with motion information;
- selecting a first spatial motion vector prediction candidate from the set of spatial motion vector prediction candidates as a potential spatial motion vector prediction candidate to be included in a merge list for the prediction unit;
- determining a subset of spatial motion vector predictions based on the location of the block associated with the first spatial motion vector prediction candidate;

comparing motion information of the first spatial motion vector prediction candidate with motion information of the spatial motion vector prediction candidate in the determined subset of spatial motion vector prediction candidates;

if at least one the comparisons indicates that the motion vector information of the spatial motion vector prediction candidates correspond with each other, excluding the first spatial motion vector prediction candidate from the merge list.

[00179] In some embodiments an apparatus according to the fourth aspect comprises a processor and a memory including computer program code, the memory and the computer program code configured to, with the processor, cause the apparatus to:

receive an encoded block of pixels including a prediction unit;

determine a set of spatial motion vector prediction candidates for the encoded block of pixels; the spatial motion vector prediction candidates being provided with motion information;

select a first spatial motion vector prediction candidate from the set of spatial motion vector prediction candidates as a potential spatial motion vector prediction candidate to be included in a merge list for the prediction unit;

determine a subset of spatial motion vector predictions based on the location of the block associated with the first spatial motion vector prediction candidate;

compare motion information of the first spatial motion vector prediction candidate with motion information of the spatial motion vector prediction candidate in the determined subset of spatial motion vector prediction candidates;

exclude the first spatial motion vector prediction candidate from the merge list, if at least one of the comparisons indicates that the motion vector information of the spatial motion vector prediction candidates correspond with each other

[00180] In some embodiments a storage medium having stored thereon a computer program code a computer executable program code for use by an encoder, said program codes comprise instructions for use by an encoder, said program code comprises instructions for:

receiving a block of pixels including a prediction unit;

determining a set of spatial motion vector prediction candidates for the block of pixels; the spatial motion vector prediction candidates being provided with motion information;

selecting a first spatial motion vector prediction candidate from the set of spatial motion vector prediction candidates as a potential spatial motion vector prediction candidate to be included in a merge list for the prediction unit;

determining a subset of spatial motion vector predictions based on the location of the block associated with the first spatial motion vector prediction candidate;

comparing motion information of the first spatial motion vector prediction candidate with motion information of the spatial motion vector prediction candidate in the determined subset of spatial motion vector prediction candidates;

if at least one of the comparisons indicates that the motion vector information of the spatial motion vector prediction candidates correspond with each other, excluding the first spatial motion vector prediction candidate from the merge list.

[00181] In some embodiments a storage medium having stored thereon a computer program code a computer executable program code for use by an encoder, said program codes comprise instructions for use by an encoder, said program code comprises instructions for:

receiving an encoded block of pixels including a prediction unit;

determining a set of spatial motion vector prediction candidates for the encoded block of pixels; the spatial motion vector prediction candidates being provided with motion information;

selecting a first spatial motion vector prediction candidate from the set of spatial motion vector prediction candidates as a potential spatial motion vector prediction candidate to be included in a merge list for the prediction unit;

determining a subset of spatial motion vector predictions based on the location of the block associated with the first spatial motion vector prediction candidate;

comparing motion information of the first spatial motion vector prediction candidate with motion information of the spatial motion vector prediction candidate in the determined subset of spatial motion vector prediction candidates;

if at least one of the comparisons indicates that the motion vector information of the spatial motion vector prediction candidates correspond with each other, excluding the first spatial motion vector prediction candidate from the merge list.

[00182] In some embodiments an apparatus comprises:

means for receiving a block of pixels including a prediction unit;

means for selecting a first spatial motion vector prediction candidate from the set of spatial motion vector prediction candidates as a potential spatial motion vector prediction candidate to be included in a merge list for the prediction unit;

means for determining a subset of spatial motion vector predictions based on the location of the block associated with the first spatial motion vector prediction candidate;

means for comparing motion information of the first spatial motion vector prediction candidate with motion information of the spatial motion vector prediction candidate in the determined subset of spatial motion vector prediction candidates;

means for excluding the first spatial motion vector prediction candidate from the merge list, if at least one of the comparisons indicates that the motion vector information of the spatial motion vector prediction candidates correspond with each other.

[00183] In some embodiments an apparatus comprises:

means for receiving an encoded block of pixels including a prediction unit;

means for determining a set of spatial motion vector prediction candidates for the encoded block of pixels; the spatial motion vector prediction candidates being provided with motion information;

means for selecting a first spatial motion vector prediction candidate from the set of spatial motion vector prediction candidates as a potential spatial motion vector prediction candidate to be included in a merge list for the prediction unit;

means for determining a subset of spatial motion vector predictions based on the location of the block associated with the first spatial motion vector prediction candidate;

means for comparing motion information of the first spatial motion vector prediction candidate with motion information of the spatial motion vector prediction candidate in the determined subset of spatial motion vector prediction candidates;

means for excluding the first spatial motion vector prediction candidate from the merge list, if at least one of the comparisons indicates that the motion vector information of the spatial motion vector prediction candidates correspond with each other.

Claims:

1. A method comprising:
receiving a block of pixels including a prediction unit;
determining a set of spatial motion vector prediction candidates for the block of pixels;
the spatial motion vector prediction candidates being provided with motion information;
selecting a first spatial motion vector prediction candidate from the set of spatial motion vector prediction candidates as a potential spatial motion vector prediction candidate to be included in a merge list for the prediction unit;
determining a subset of spatial motion vector predictions based on the location of the block associated with the first spatial motion vector prediction candidate;
comparing motion information of the first spatial motion vector prediction candidate with motion information of the spatial motion vector prediction candidate in the determined subset of spatial motion vector prediction candidates;
if at least one of the comparisons indicates that the motion vector information of the spatial motion vector prediction candidates correspond with each other, excluding the first spatial motion vector prediction candidate from the merge list.
2. The method according to claim 1 comprising selecting spatial motion vector prediction candidates from the set of spatial motion vector prediction candidates as the potential spatial motion vector prediction candidate in a predetermined order.
3. The method according to claim 1, comprising comparing motion information of the potential spatial motion vector prediction candidate with motion information of at most one other spatial motion vector prediction candidate of the set of spatial motion vector prediction candidates.
4. The method according to claim 1 comprising examining whether the received block of pixels is divided into a first prediction unit and a second prediction unit; and if so, excluding the potential spatial motion vector prediction candidate from the merge list if the prediction unit is the second prediction unit.

5. The method according to claim 1, further comprising
determining a maximum number of spatial motion vector prediction candidates to be
included in a merge list; and
limiting the number of spatial motion vector prediction candidates in the merge list
smaller or equal to the maximum number .

6. The method according to claim 1 comprising:
examining, if the number of spatial motion vector prediction candidates in the merge list
smaller than the maximum number;
if so, examining whether a prediction unit to which the potential spatial motion vector
prediction candidate belongs is available for motion prediction;
if so, performing at least one of the following:
for the potential spatial motion vector prediction candidate on the left side of the
prediction unit, excluding the potential spatial motion vector prediction candidate from the
merge list if any of the following conditions are fulfilled:

- the received block of pixels is vertically divided into a first
prediction unit and a second prediction unit, and the prediction unit is the second
prediction unit;
- the received block of pixels is horizontally divided into a first
prediction unit and a second prediction unit, and if the prediction unit is the
second prediction unit, and the potential spatial motion vector prediction
candidate has essentially similar motion information than the spatial motion
vector prediction candidate above the prediction unit;

for the potential spatial motion vector prediction candidate above the prediction unit,
excluding the potential spatial motion vector prediction candidate from the merge list if any of
the following conditions are fulfilled:

- the received block of pixels is horizontally divided into a first
prediction unit and a second prediction unit, and the prediction unit is the second
prediction unit;

- the potential spatial motion vector prediction candidate has essentially similar motion information than the spatial motion vector prediction candidate on the left side of the prediction unit;

for the potential spatial motion vector prediction candidate, which is on the right side of the potential spatial motion vector prediction candidate above the prediction unit, excluding the potential spatial motion vector prediction candidate from the merge list if the potential spatial motion vector prediction candidate has essentially similar motion information than the spatial motion vector prediction candidate above the prediction unit;

for the potential spatial motion vector prediction candidate, which is below the potential spatial motion vector prediction candidate on the left side of the prediction unit, excluding the potential spatial motion vector prediction candidate from the merge list if the potential spatial motion vector prediction candidate has essentially similar motion information than the spatial motion vector prediction candidate on the left side of the prediction unit;

for the potential spatial motion vector prediction candidate cornerwise neighbouring the prediction unit, excluding the potential spatial motion vector prediction candidate from the merge list if any of the following conditions are fulfilled:

- all the other potential spatial motion vector prediction candidates have been included in the merge list;
- the potential spatial motion vector prediction candidate has essentially similar motion information than the spatial motion vector prediction candidate above the prediction unit;
- the potential spatial motion vector prediction candidate has essentially similar motion information than the spatial motion vector prediction candidate on the left side of the prediction unit.

7. The method according to claim 1 further comprising including a temporal motion prediction candidate into the merge list.

8. The method according to claim 1 comprising selecting one motion vector prediction candidate from the merge list to represent a motion vector prediction for the block of pixels.

9. A method comprising:
receiving an encoded block of pixels including a prediction unit;
determining a set of spatial motion vector prediction candidates for the encoded block of pixels; the spatial motion vector prediction candidates being provided with motion information;
selecting a first spatial motion vector prediction candidate from the set of spatial motion vector prediction candidates as a potential spatial motion vector prediction candidate to be included in a merge list for the prediction unit;
determining a subset of spatial motion vector predictions based on the location of the block associated with the first spatial motion vector prediction candidate;
comparing motion information of the first spatial motion vector prediction candidate with motion information of another spatial motion vector prediction candidate of the set of spatial motion vector prediction candidates;
if at least one of the comparisons indicates that the motion vector information of the spatial motion vector prediction candidates correspond with each other, excluding the first spatial motion vector prediction candidate from the merge list.

10. The method according to claim 9 comprising comparing motion information of the potential spatial motion vector prediction candidate with motion information of at most one other spatial motion vector prediction candidate of the set of spatial motion vector prediction candidates.

11. The method according to claim 9 comprising examining whether the received encoded block of pixels is divided into a first prediction unit and a second prediction unit; and if so, excluding the potential spatial motion vector prediction candidate from the merge list if the prediction unit is the second prediction unit.

12. The method according to claim 9 further comprising
determining a maximum number of spatial motion vector prediction candidates to be included in a merge list; and

limiting the number of spatial motion vector prediction candidates in the merge list smaller or equal to the maximum number.

13. The method according to claim 9 comprising:

examining, if the number of spatial motion vector prediction candidates in the merge list smaller than the maximum number;

if so, examining whether a prediction unit to which the potential spatial motion vector prediction candidate belongs is available for motion prediction;

if so, performing at least one of the following:

for the potential spatial motion vector prediction candidate on the left side of the prediction unit, excluding the potential spatial motion vector prediction candidate from the merge list if any of the following conditions are fulfilled:

- the received encoded block of pixels is vertically divided into a first prediction unit and a second prediction unit, and the prediction unit is the second prediction unit;

- the received encoded block of pixels is horizontally divided into a first prediction unit and a second prediction unit, and if the prediction unit is the second prediction unit, and the potential spatial motion vector prediction candidate has essentially similar motion information than the spatial motion vector prediction candidate above the prediction unit;

for the potential spatial motion vector prediction candidate above the prediction unit, excluding the potential spatial motion vector prediction candidate from the merge list if any of the following conditions are fulfilled:

- the received encoded block of pixels is horizontally divided into a first prediction unit and a second prediction unit, and the prediction unit is the second prediction unit;

- the potential spatial motion vector prediction candidate has essentially similar motion information than the spatial motion vector prediction candidate on the left side of the prediction unit;

for the potential spatial motion vector prediction candidate, which is on the right side of the potential spatial motion vector prediction candidate above the prediction unit, excluding the

potential spatial motion vector prediction candidate from the merge list if the potential spatial motion vector prediction candidate has essentially similar motion information than the spatial motion vector prediction candidate above the prediction unit;

for the potential spatial motion vector prediction candidate, which is below the potential spatial motion vector prediction candidate on the left side of the prediction unit, excluding the potential spatial motion vector prediction candidate from the merge list if the potential spatial motion vector prediction candidate has essentially similar motion information than the spatial motion vector prediction candidate on the left side of the prediction unit;

for the potential spatial motion vector prediction candidate cornerwise neighbouring the prediction unit, excluding the potential spatial motion vector prediction candidate from the merge list if any of the following conditions are fulfilled:

- all the other potential spatial motion vector prediction candidates have been included in the merge list;
- the potential spatial motion vector prediction candidate has essentially similar motion information than the spatial motion vector prediction candidate above the prediction unit;
- the potential spatial motion vector prediction candidate has essentially similar motion information than the spatial motion vector prediction candidate on the left side of the prediction unit.

14. The method according to claim 9 comprising selecting one motion vector prediction candidate from the merge list to represent a motion vector prediction for the received encoded block of pixels.

15. An apparatus comprising a processor and a memory including computer program code, the memory and the computer program code configured to, with the processor, cause the apparatus to:

- receiving a block of pixels including a prediction unit;
- determining a set of spatial motion vector prediction candidates for the block of pixels;
- the spatial motion vector prediction candidates being provided with motion information;

selecting a first spatial motion vector prediction candidate from the set of spatial motion vector prediction candidates as a potential spatial motion vector prediction candidate to be included in a merge list for the prediction unit;

determining a subset of spatial motion vector predictions based on the location of the block associated with the first spatial motion vector prediction candidate;

comparing motion information of the first spatial motion vector prediction candidate with motion information of the spatial motion vector prediction candidate in the determined subset of spatial motion vector prediction candidates;

if at least one the comparisons indicates that the motion vector information of the spatial motion vector prediction candidates correspond with each other, excluding the first spatial motion vector prediction candidate from the merge list.

16. An apparatus comprising a processor and a memory including computer program code, the memory and the computer program code configured to, with the processor, cause the apparatus to:

receive an encoded block of pixels including a prediction unit;

determine a set of spatial motion vector prediction candidates for the encoded block of pixels; the spatial motion vector prediction candidates being provided with motion information;

select a first spatial motion vector prediction candidate from the set of spatial motion vector prediction candidates as a potential spatial motion vector prediction candidate to be included in a merge list for the prediction unit;

determine a subset of spatial motion vector predictions based on the location of the block associated with the first spatial motion vector prediction candidate;

compare motion information of the first spatial motion vector prediction candidate with motion information of the spatial motion vector prediction candidate in the determined subset of spatial motion vector prediction candidates;

exclude the first spatial motion vector prediction candidate from the merge list, if at least one of the comparisons indicates that the motion vector information of the spatial motion vector prediction candidates correspond with each other

17. A storage medium having stored thereon a computer executable program code for use by an encoder, said program codes comprise instructions for use by an encoder, said program code comprises instructions for:

- receiving a block of pixels including a prediction unit;
- determining a set of spatial motion vector prediction candidates for the block of pixels; the spatial motion vector prediction candidates being provided with motion information;
- selecting a first spatial motion vector prediction candidate from the set of spatial motion vector prediction candidates as a potential spatial motion vector prediction candidate to be included in a merge list for the prediction unit;
- determining a subset of spatial motion vector predictions based on the location of the block associated with the first spatial motion vector prediction candidate;
- comparing motion information of the first spatial motion vector prediction candidate with motion information of the spatial motion vector prediction candidate in the determined subset of spatial motion vector prediction candidates;
- if at least one of the comparisons indicates that the motion vector information of the spatial motion vector prediction candidates correspond with each other, excluding the first spatial motion vector prediction candidate from the merge list.

18. A storage medium having stored thereon a computer executable program code for use by an encoder, said program codes comprise instructions for use by an encoder, said program code comprises instructions for:

- receiving an encoded block of pixels including a prediction unit;
- determining a set of spatial motion vector prediction candidates for the encoded block of pixels; the spatial motion vector prediction candidates being provided with motion information;
- selecting a first spatial motion vector prediction candidate from the set of spatial motion vector prediction candidates as a potential spatial motion vector prediction candidate to be included in a merge list for the prediction unit;
- determining a subset of spatial motion vector predictions based on the location of the block associated with the first spatial motion vector prediction candidate;

comparing motion information of the first spatial motion vector prediction candidate with motion information of the spatial motion vector prediction candidate in the determined subset of spatial motion vector prediction candidates;

if at least one of the comparisons indicates that the motion vector information of the spatial motion vector prediction candidates correspond with each other, excluding the first spatial motion vector prediction candidate from the merge list.

19. An apparatus comprising:

means for receiving a block of pixels including a prediction unit;

means for determining a set of spatial motion vector prediction candidates for the block of pixels; the spatial motion vector prediction candidates being provided with motion information;

means for selecting a first spatial motion vector prediction candidate from the set of spatial motion vector prediction candidates as a potential spatial motion vector prediction candidate to be included in a merge list for the prediction unit;

means for determining a subset of spatial motion vector predictions based on the location of the block associated with the first spatial motion vector prediction candidate;

means for comparing motion information of the first spatial motion vector prediction candidate with motion information of the spatial motion vector prediction candidate in the determined subset of spatial motion vector prediction candidates;

means for excluding the first spatial motion vector prediction candidate from the merge list, if at least one of the comparisons indicates that the motion vector information of the spatial motion vector prediction candidates correspond with each other.

20. An apparatus comprising:

means for receiving an encoded block of pixels including a prediction unit;

means for determining a set of spatial motion vector prediction candidates for the encoded block of pixels; the spatial motion vector prediction candidates being provided with motion information;

means for selecting a first spatial motion vector prediction candidate from the set of spatial motion vector prediction candidates as a potential spatial motion vector prediction candidate to be included in a merge list for the prediction unit;

means for determining a subset of spatial motion vector predictions based on the location of the block associated with the first spatial motion vector prediction candidate;

means for comparing motion information of the first spatial motion vector prediction candidate with motion information of the spatial motion vector prediction candidate in the determined subset of spatial motion vector prediction candidates;

means for excluding the first spatial motion vector prediction candidate from the merge list, if at least one of the comparisons indicates that the motion vector information of the spatial motion vector prediction candidates correspond with each other.

ABSTRACT

The invention relates to a method for encoding, a method for decoding, an apparatus, computer program products, an encoder and a decoder for video information. The motion vector for a block in a video image is predicted from a set of motion vector prediction candidates determined based on previously-coded motion vectors. A motion vector prediction candidate is included in the set based on the location of the block associated with the first spatial motion vector prediction candidate and in comparison with motion vector prediction candidates already in the set.

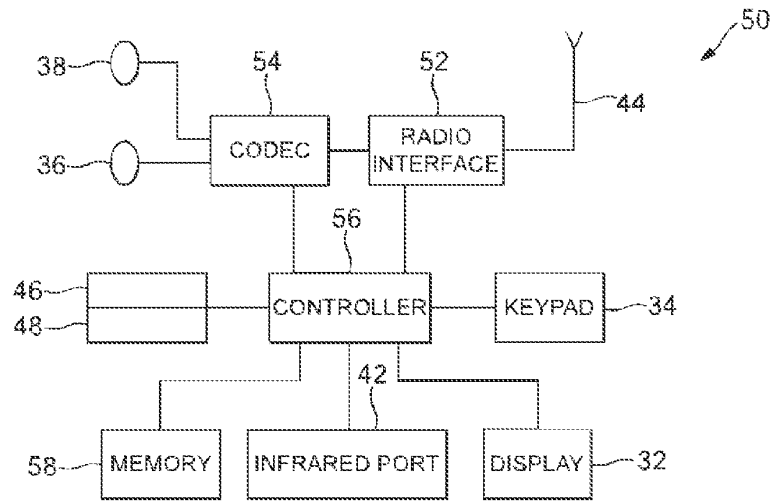


FIG. 1

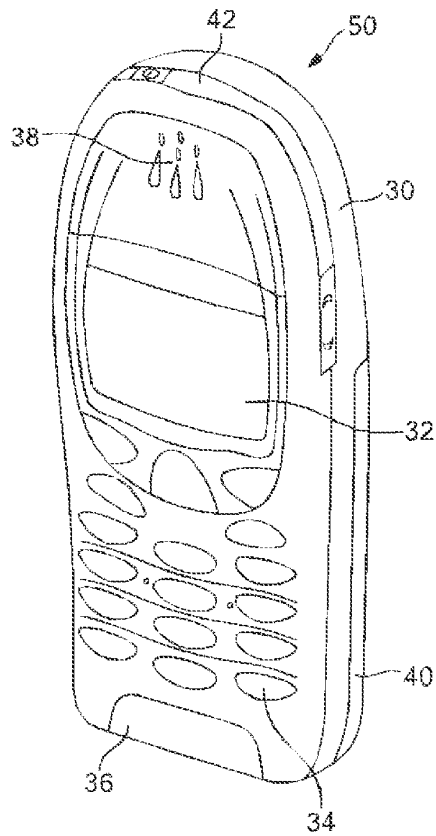


FIG. 2

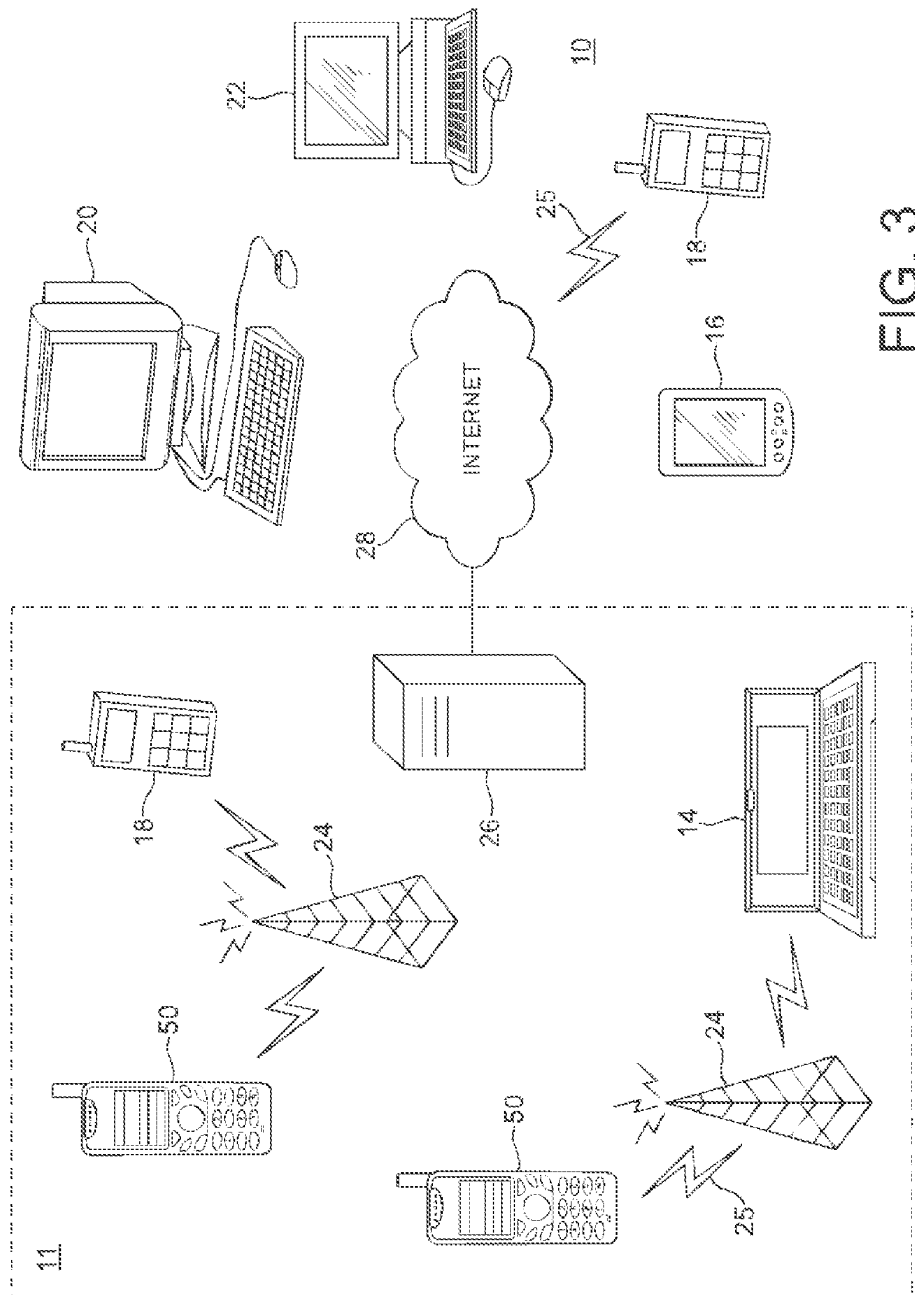


FIG. 3

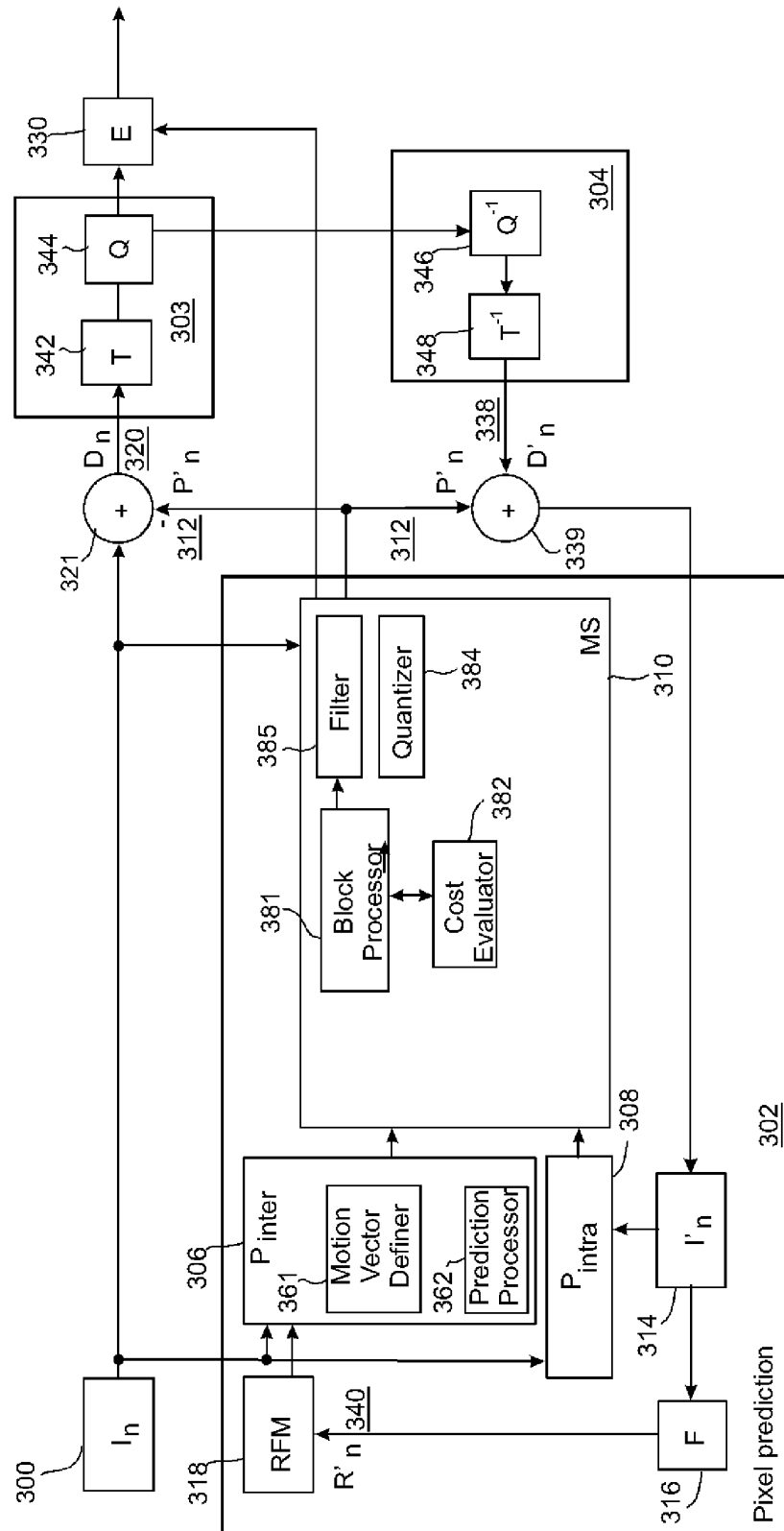


Fig. 4a

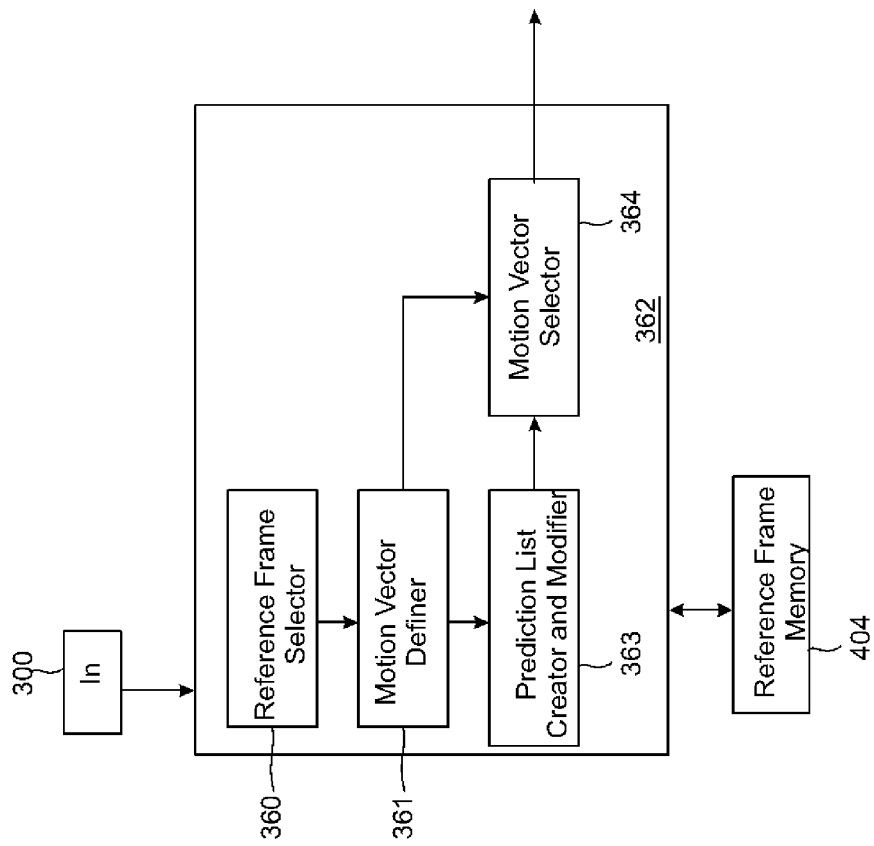


Fig. 4b

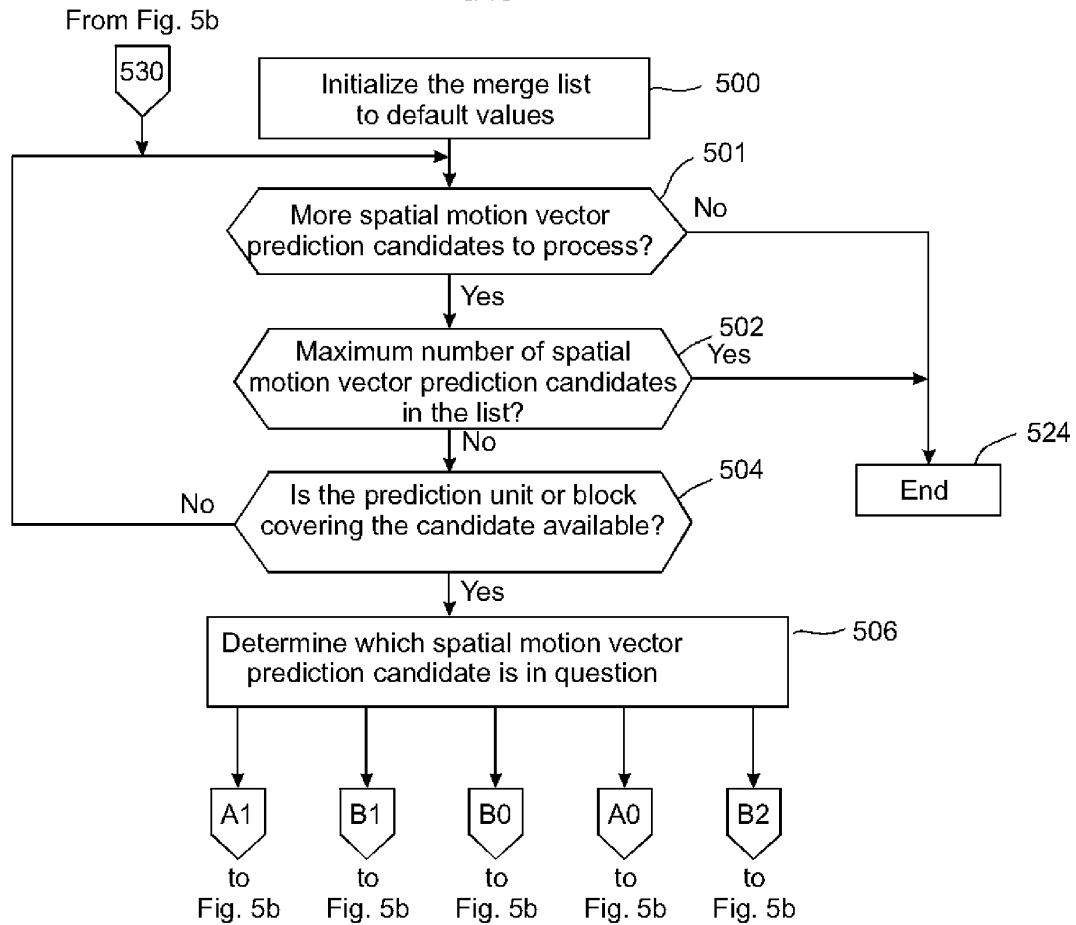


Fig. 5a

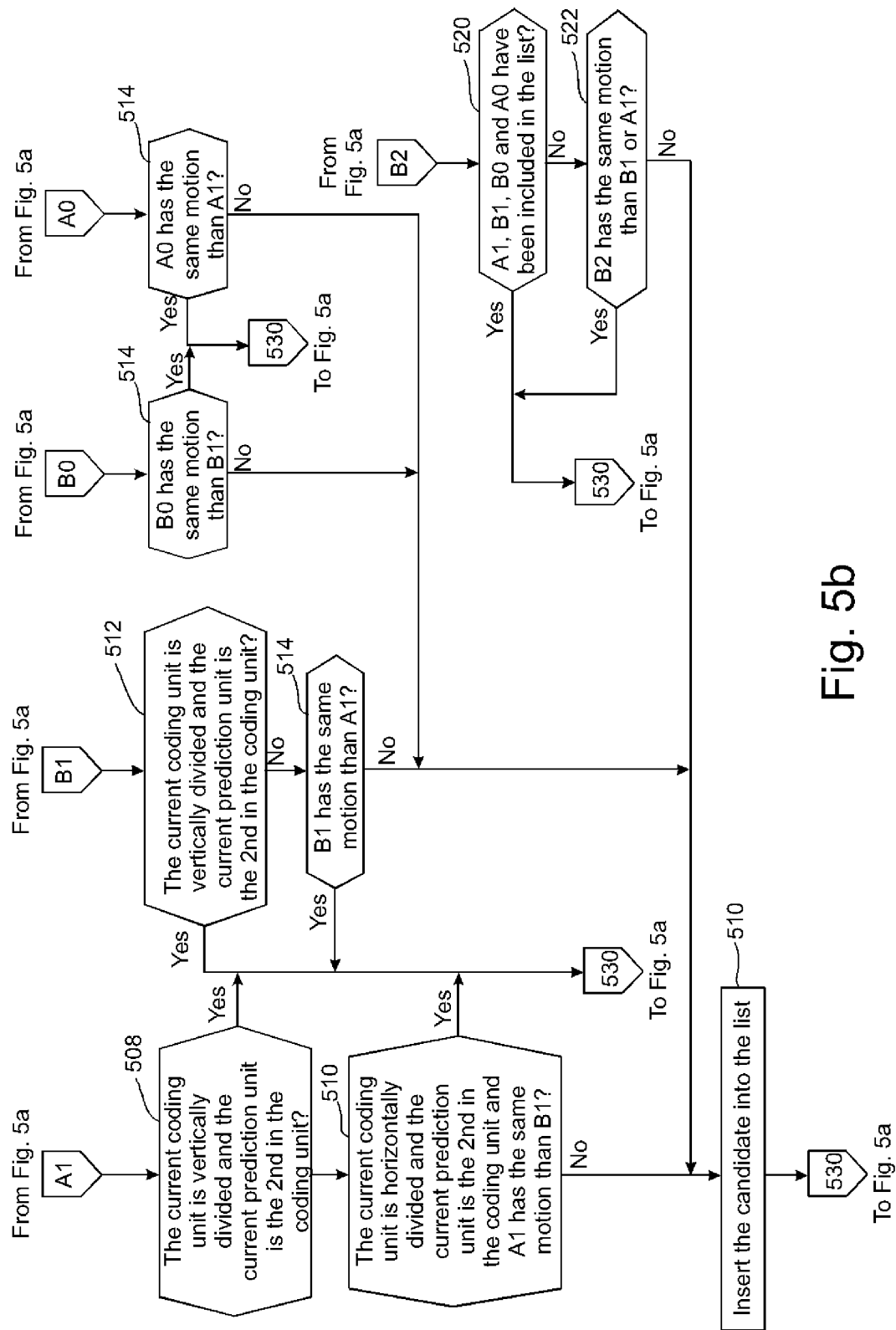


Fig. 5b

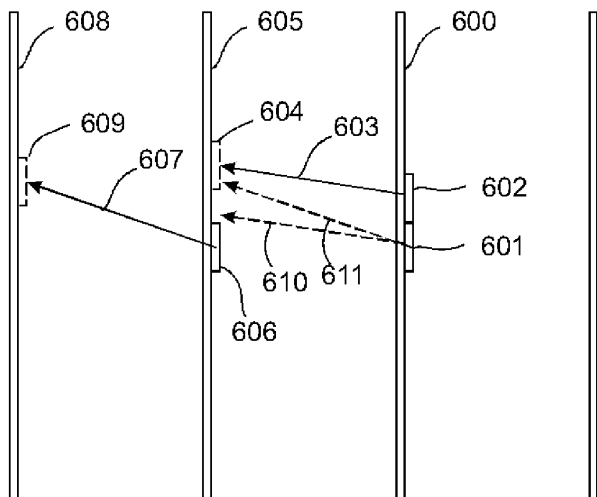


Fig. 6a

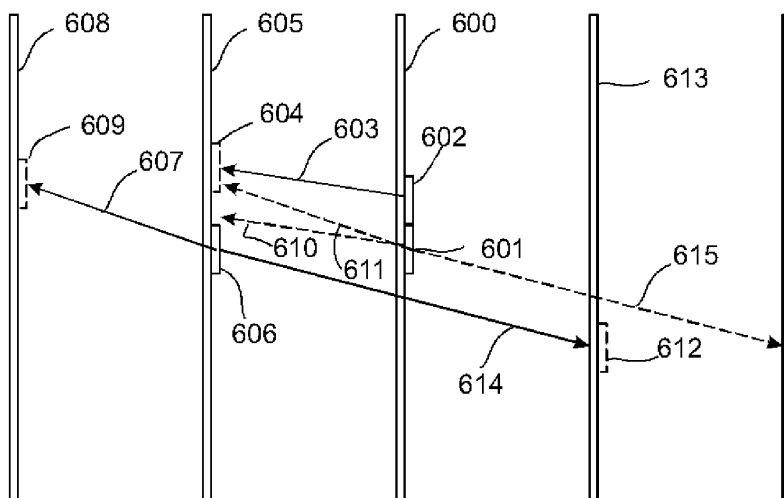


Fig. 6b

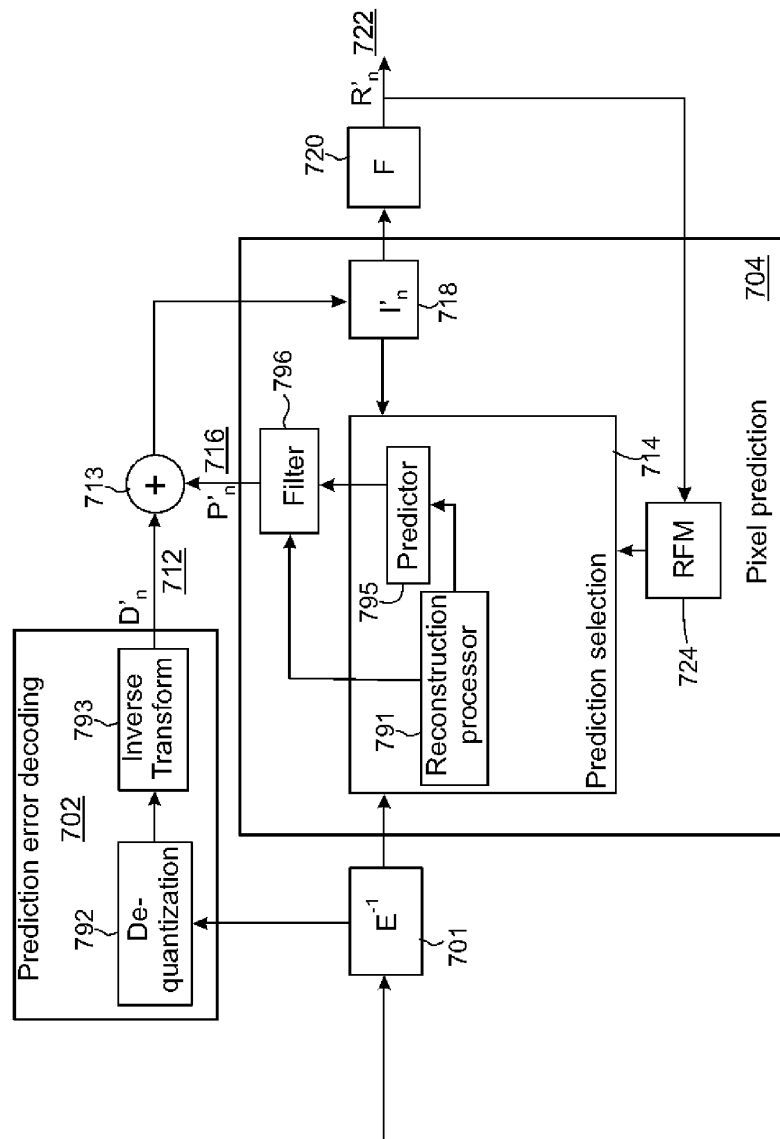


Fig. 7

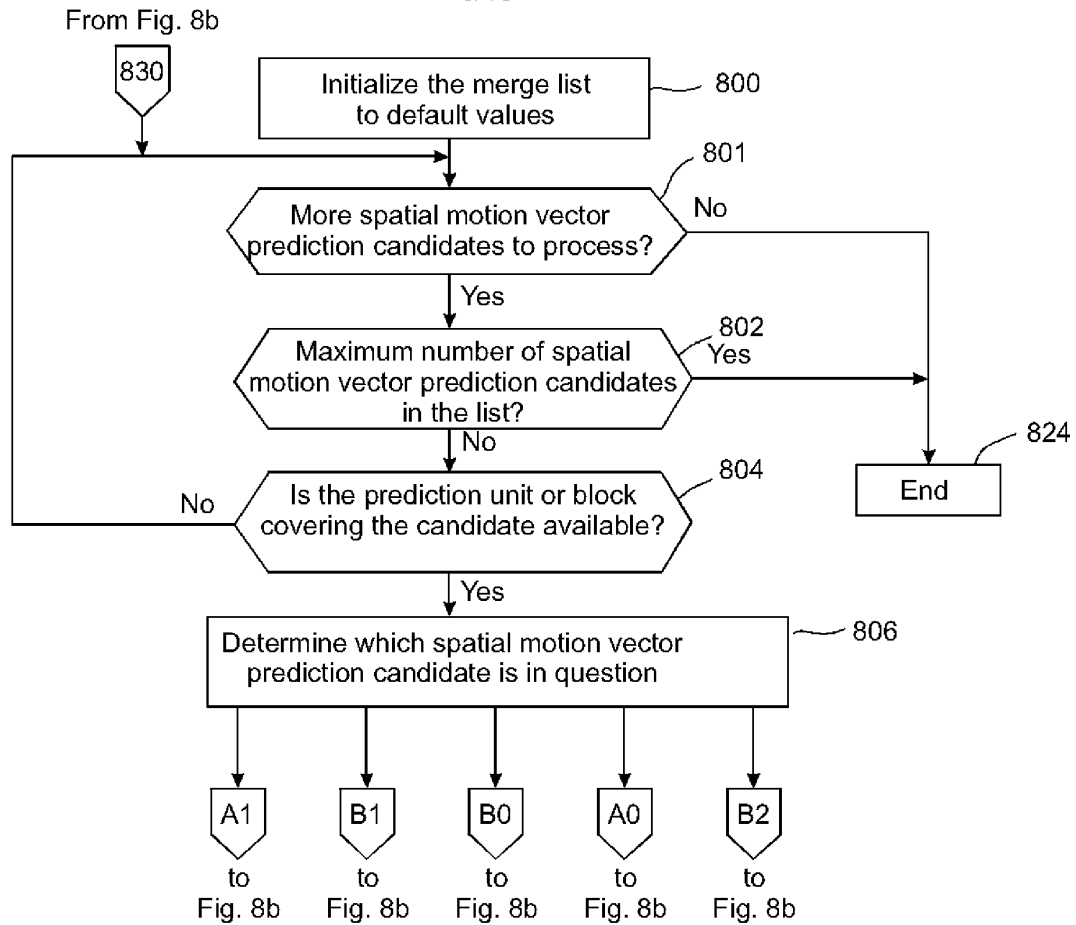


Fig. 8a

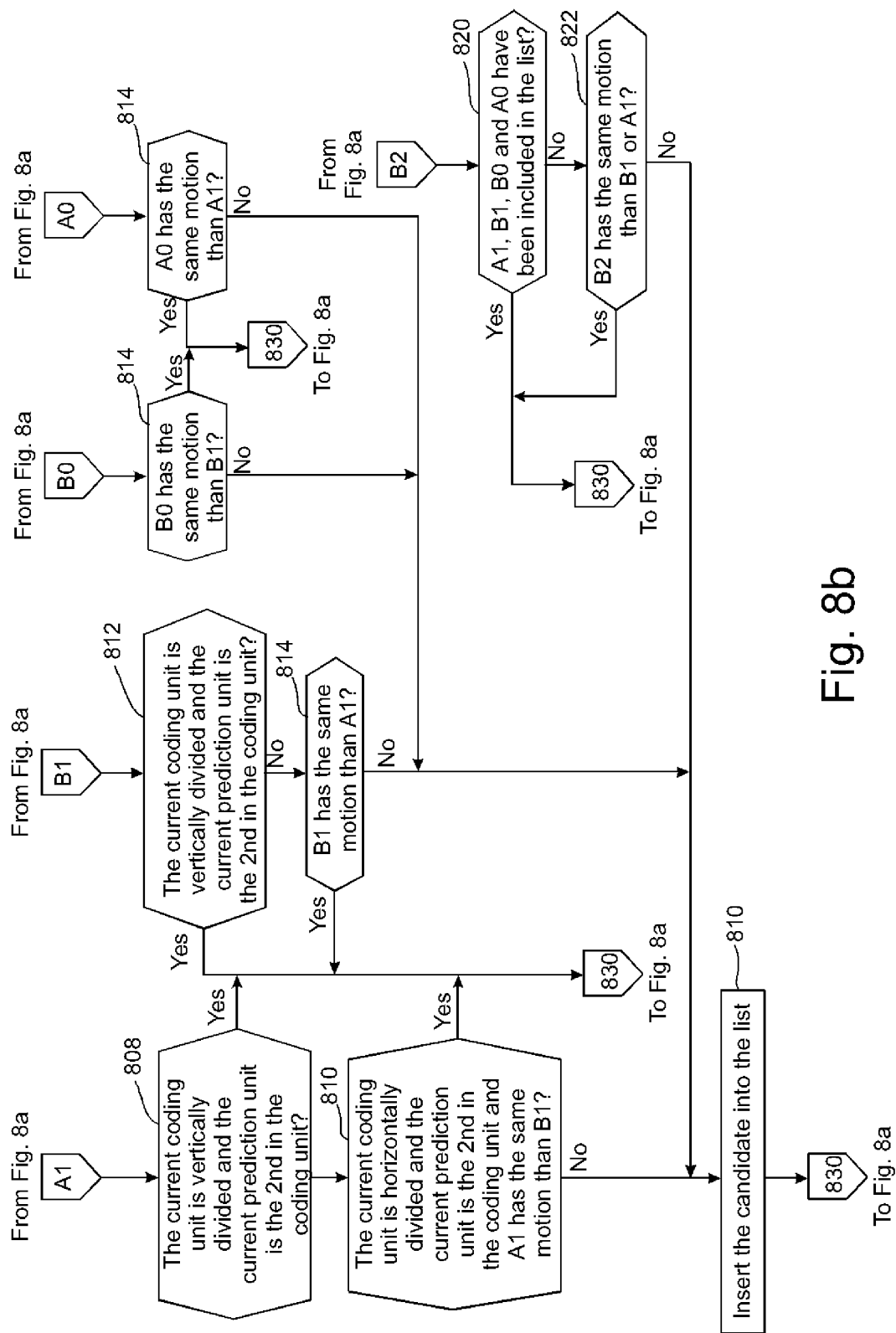


Fig. 8b

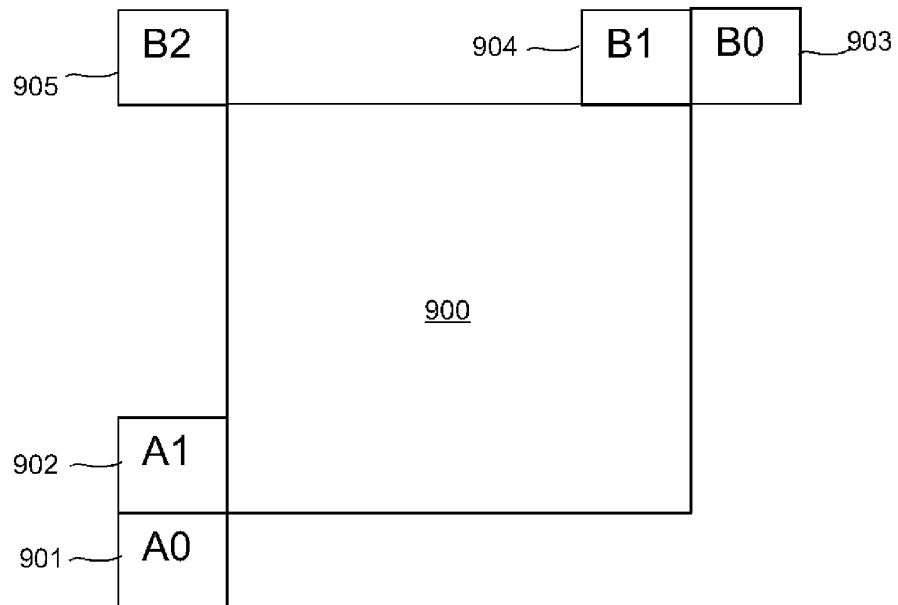


Fig. 9

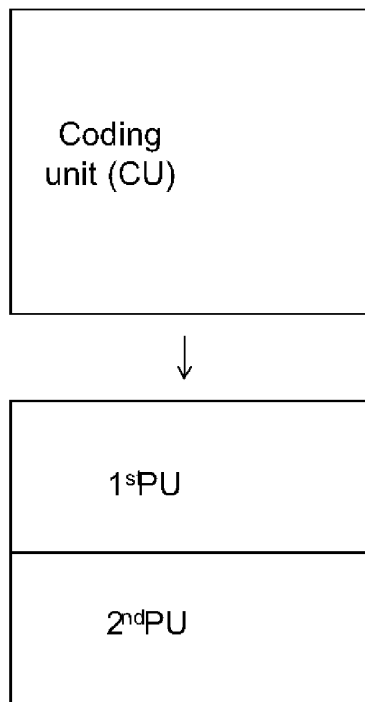


Fig. 10a

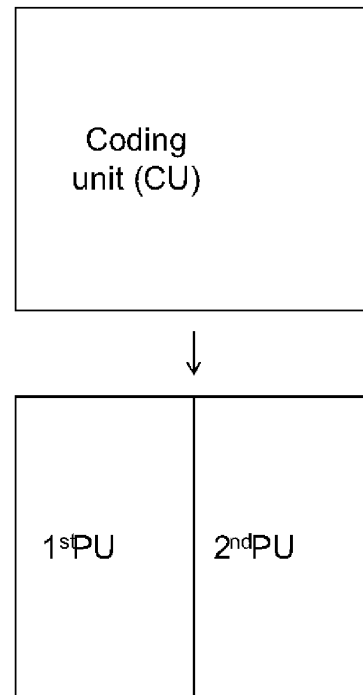


Fig. 10b

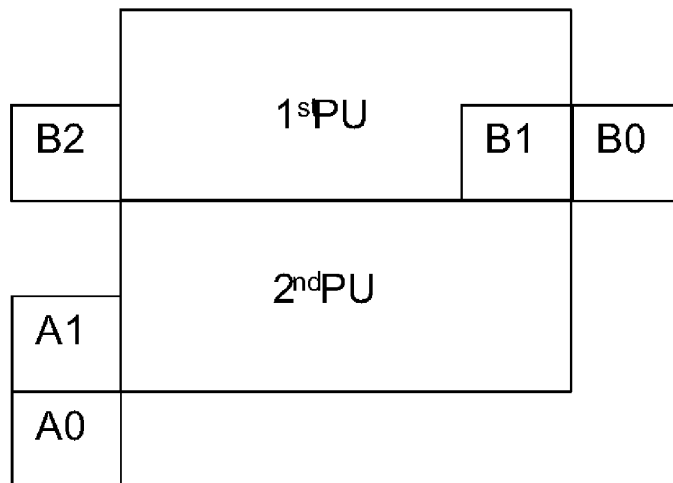


Fig. 11a

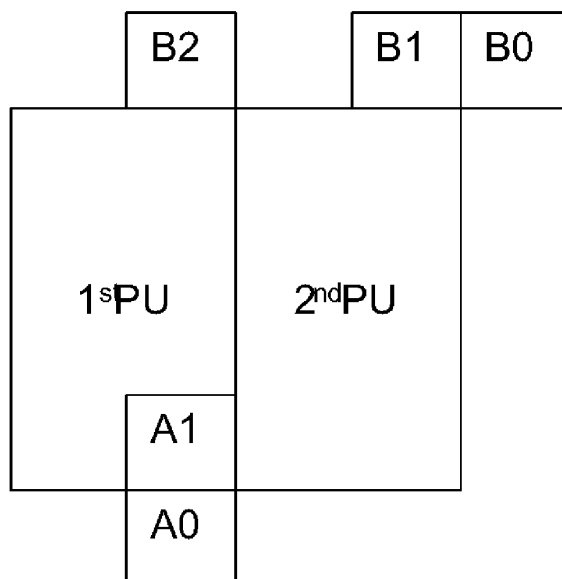


Fig. 11b

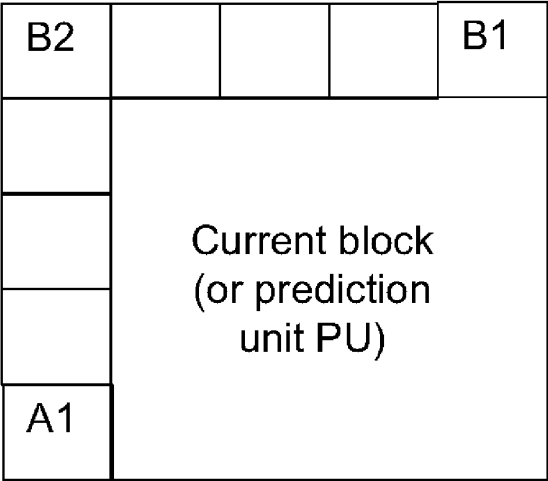


Fig. 12

Electronic Patent Application Fee Transmittal				
Application Number:				
Filing Date:				
Title of Invention:		METHOD FOR CODING AND AN APPARATUS		
First Named Inventor/Applicant Name:		Mehmet Oguz BICI		
Filer:		Thomas Joseph Arria/thao pham		
Attorney Docket Number:		NC77198US-NP		
Filed as Large Entity				
Utility under 35 USC 111(a) Filing Fees				
Description	Fee Code	Quantity	Amount	Sub-Total in USD(\$)
Basic Filing:				
Utility application filing	1011	1	390	390
Utility Search Fee	1111	1	620	620
Utility Examination Fee	1311	1	250	250
Pages:				
Claims:				
Independent claims in excess of 3	1201	5	250	1250
Miscellaneous-Filing:				
Petition:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD(\$)
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				
Miscellaneous:				
Total in USD (\$)				2510

Electronic Acknowledgement Receipt	
EFS ID:	14129936
Application Number:	13666680
International Application Number:	
Confirmation Number:	4782
Title of Invention:	METHOD FOR CODING AND AN APPARATUS
First Named Inventor/Applicant Name:	Mehmet Oguz BICI
Customer Number:	73658
Filer:	Thomas Joseph Arria/thao pham
Filer Authorized By:	Thomas Joseph Arria
Attorney Docket Number:	NC77198US-NP
Receipt Date:	01-NOV-2012
Filing Date:	
Time Stamp:	17:52:47
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	Deposit Account
Payment was successfully received in RAM	\$2510
RAM confirmation Number	4958
Deposit Account	500270
Authorized User	
The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows: Charge any Additional Fees required under 37 C.F.R. Section 1.17 (Patent application and reexamination processing fees) Charge any Additional Fees required under 37 C.F.R. Section 1.21 (Miscellaneous fees and charges)	

File Listing:					
Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Application Data Sheet	77198_ADS.pdf	837045	no	5
			cc4bf4c800fa962613531be16b5486bd271ba7		
Warnings:					
Information:					
2		77198US-NP_Specification.pdf	231397	yes	56
			6e610cea7fe41e52a4bdd06cb679afbceee123e		
	Multipart Description/PDF files in .zip description				
	Document Description		Start	End	
	Specification		1	45	
	Claims		46	55	
	Abstract		56	56	
Warnings:					
Information:					
3	Drawings-only black and white line drawings	77198_drawings.pdf	157966	no	13
			d0f31e08b0ec09fdccf729907fae7a17633ad799		
Warnings:					
Information:					
4	Fee Worksheet (SB06)	fee-info.pdf	36587	no	2
			5c191f1951b5032e2ac44f62defeeebc36a4f021		
Warnings:					
Information:					
Total Files Size (in bytes):			1262995		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PATENT APPLICATION FEE DETERMINATION RECORD Substitute for Form PTO-875						Application or Docket Number 13/666,680				
APPLICATION AS FILED - PART I										
(Column 1)		(Column 2)		SMALL ENTITY		OR OTHER THAN SMALL ENTITY				
FOR	NUMBER FILED	NUMBER EXTRA	RATE(\$)	FEE(\$)		RATE(\$)	FEE(\$)			
BASIC FEE (37 CFR 1.16(a), (b), or (c))	N/A	N/A	N/A			N/A	390			
SEARCH FEE (37 CFR 1.16(k), (l), or (m))	N/A	N/A	N/A			N/A	620			
EXAMINATION FEE (37 CFR 1.16(o), (p), or (q))	N/A	N/A	N/A			N/A	250			
TOTAL CLAIMS (37 CFR 1.16(j))	20	minus 20 = *				x 62 =	0.00			
INDEPENDENT CLAIMS (37 CFR 1.16(h))	8	minus 3 = *	5			x 250 =	1250			
APPLICATION SIZE FEE (37 CFR 1.16(s))	If the specification and drawings exceed 100 sheets of paper, the application size fee due is \$310 (\$155 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).						0.00			
MULTIPLE DEPENDENT CLAIM PRESENT (37 CFR 1.16(j))							0.00			
* If the difference in column 1 is less than zero, enter "0" in column 2.				TOTAL		TOTAL	2510			
APPLICATION AS AMENDED - PART II										
(Column 1)		(Column 2)		(Column 3)		SMALL ENTITY		OR OTHER THAN SMALL ENTITY		
AMENDMENT A		CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA	RATE(\$)	ADDITIONAL FEE(\$)		RATE(\$)	ADDITIONAL FEE(\$)
	Total (37 CFR 1.16(i))	*	Minus	**	=	x	=		x	=
	Independent (37 CFR 1.16(h))	*	Minus	***	=	x	=		x	=
	Application Size Fee (37 CFR 1.16(s))									
	FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))									
						TOTAL ADD'L FEE			TOTAL ADD'L FEE	
AMENDMENT B		CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA	RATE(\$)	ADDITIONAL FEE(\$)		RATE(\$)	ADDITIONAL FEE(\$)
	Total (37 CFR 1.16(i))	*	Minus	**	=	x	=		x	=
	Independent (37 CFR 1.16(h))	*	Minus	***	=	x	=		x	=
	Application Size Fee (37 CFR 1.16(s))									
	FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))									
						TOTAL ADD'L FEE			TOTAL ADD'L FEE	
* If the entry in column 1 is less than the entry in column 2, write "0" in column 3. ** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20". *** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3". The "Highest Number Previously Paid For" (Total or Independent) is the highest found in the appropriate box in column 1.										



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING OR 371(C) DATE	FIRST NAMED APPLICANT	ATTY. DOCKET NO./TITLE
13/666,680	11/01/2012	Mehmet Oguz BICI	NC77198US-NP

CONFIRMATION NO. 4782

FORMALITIES LETTER

73658

Nokia, Inc.

Attn: Intellectual Property Rights Docketing

200 South Mathilda Ave

Sunnyvale, CA 94086



OC000000057716268

Date Mailed: 11/26/2012

NOTICE TO FILE MISSING PARTS OF NONPROVISIONAL APPLICATION

FILED UNDER 37 CFR 1.53(b)

Filing Date Granted

Items Required To Avoid Abandonment:

An application number and filing date have been accorded to this application. The item(s) indicated below, however, are missing.

Applicant is given **TWO MONTHS** from the date of this Notice within which to file all required items below to avoid abandonment. Extensions of time may be obtained by filing a petition accompanied by the extension fee under the provisions of 37 CFR 1.136(a).

- A surcharge (for late submission of the basic filing fee, search fee, examination fee or inventor's oath or declaration) as set forth in 37 CFR 1.16(f) of \$ **130** for a non-small entity, must be submitted.

SUMMARY OF FEES DUE:

Total fee(s) required within **TWO MONTHS** from the date of this Notice is \$ **130** for a non-small entity

- \$ **130** Surcharge.

Items Required To Avoid Processing Delays:

Applicant is notified that the above-identified application contains the deficiencies noted below. No period for reply is set forth in this notice for correction of these deficiencies. However, if a deficiency relates to the inventor's oath or declaration, the applicant must file an oath or declaration in compliance with 37 CFR 1.63, or a substitute statement in compliance with 37 CFR 1.64, executed by or with respect to each actual inventor no later than the expiration of the time period set in the "Notice of Allowability" to avoid abandonment. See 37 CFR 1.53(f).

- A properly executed inventor's oath or declaration has not been received for the following inventor(s):
all

Applicant may submit the inventor's oath or declaration at any time before the Notice of Allowance and Fee(s) Due, PTOL-85, is mailed.

Replies must be received in the USPTO within the set time period or must include a proper Certificate of Mailing or Transmission under 37 CFR 1.8 with a mailing or transmission date within the set time period. For more information and a suggested format, see Form PTO/SB/92 and MPEP 512.

Replies should be mailed to:

Mail Stop Missing Parts
Commissioner for Patents
P.O. Box 1450
Alexandria VA 22313-1450

Registered users of EFS-Web may alternatively submit their reply to this notice via EFS-Web.

<https://sportal.uspto.gov/authenticate/AuthenticateUserLocalEPF.html>

For more information about EFS-Web please call the USPTO Electronic Business Center at **1-866-217-9197** or visit our website at <http://www.uspto.gov/ebc>.

If you are not using EFS-Web to submit your reply, you must include a copy of this notice.

/smunpanthovong/

Office of Data Management, Application Assistance Unit (571) 272-4000, or (571) 272-4200, or 1-888-786-0101



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING or 371(c) DATE	GRP ART UNIT	FIL FEE REC'D	ATTY. DOCKET NO.	TOT CLAIMS	IND CLAIMS
13/666,680	11/01/2012	2631	2510	NC77198US-NP	20	8

CONFIRMATION NO. 4782

73658

Nokia, Inc.

Attn: Intellectual Property Rights Docketing
200 South Mathilda Ave
Sunnyvale, CA 94086

FILING RECEIPT



OC000000057716267

Date Mailed: 11/26/2012

Receipt is acknowledged of this non-provisional patent application. The application will be taken up for examination in due course. Applicant will be notified as to the results of the examination. Any correspondence concerning the application must include the following identification information: the U.S. APPLICATION NUMBER, FILING DATE, NAME OF APPLICANT, and TITLE OF INVENTION. Fees transmitted by check or draft are subject to collection. Please verify the accuracy of the data presented on this receipt. **If an error is noted on this Filing Receipt, please submit a written request for a Filing Receipt Correction. Please provide a copy of this Filing Receipt with the changes noted thereon. If you received a "Notice to File Missing Parts" for this application, please submit any corrections to this Filing Receipt with your reply to the Notice. When the USPTO processes the reply to the Notice, the USPTO will generate another Filing Receipt incorporating the requested corrections**

Inventor(s)

Mehmet Oguz BICI, Tampere, TURKEY;
Jani LAINEMA, Tampere, FINLAND;
Kemal UGUR, Tampere, FINLAND;

Applicant(s)

Mehmet Oguz BICI, Tampere, TURKEY;
Jani LAINEMA, Tampere, FINLAND;
Kemal UGUR, Tampere, FINLAND;

Assignment For Published Patent Application

NOKIA CORPORATION, Espoo, FINLAND

Power of Attorney: None

Domestic Priority data as claimed by applicant

This appln claims benefit of 61/555,703 11/04/2011

Foreign Applications for which priority is claimed (You may be eligible to benefit from the **Patent Prosecution Highway** program at the USPTO. Please see <http://www.uspto.gov> for more information.) - None.

Foreign application information must be provided in an Application Data Sheet in order to constitute a claim to foreign priority. See 37 CFR 1.55 and 1.76.

If Required, Foreign Filing License Granted: 11/20/2012

The country code and number of your priority application, to be used for filing abroad under the Paris Convention, is **US 13/666,680**

Projected Publication Date: 05/09/2013

Non-Publication Request: No

Early Publication Request: No
Title

METHOD FOR CODING AND AN APPARATUS

Preliminary Class

375

PROTECTING YOUR INVENTION OUTSIDE THE UNITED STATES

Since the rights granted by a U.S. patent extend only throughout the territory of the United States and have no effect in a foreign country, an inventor who wishes patent protection in another country must apply for a patent in a specific country or in regional patent offices. Applicants may wish to consider the filing of an international application under the Patent Cooperation Treaty (PCT). An international (PCT) application generally has the same effect as a regular national patent application in each PCT-member country. The PCT process **simplifies** the filing of patent applications on the same invention in member countries, but **does not result** in a grant of "an international patent" and does not eliminate the need of applicants to file additional documents and fees in countries where patent protection is desired.

Almost every country has its own patent law, and a person desiring a patent in a particular country must make an application for patent in that country in accordance with its particular laws. Since the laws of many countries differ in various respects from the patent law of the United States, applicants are advised to seek guidance from specific foreign countries to ensure that patent rights are not lost prematurely.

Applicants also are advised that in the case of inventions made in the United States, the Director of the USPTO must issue a license before applicants can apply for a patent in a foreign country. The filing of a U.S. patent application serves as a request for a foreign filing license. The application's filing receipt contains further information and guidance as to the status of applicant's license for foreign filing.

Applicants may wish to consult the USPTO booklet, "General Information Concerning Patents" (specifically, the section entitled "Treaties and Foreign Patents") for more information on timeframes and deadlines for filing foreign patent applications. The guide is available either by contacting the USPTO Contact Center at 800-786-9199, or it can be viewed on the USPTO website at <http://www.uspto.gov/web/offices/pac/doc/general/index.html>.

For information on preventing theft of your intellectual property (patents, trademarks and copyrights), you may wish to consult the U.S. Government website, <http://www.stopfakes.gov>. Part of a Department of Commerce initiative, this website includes self-help "toolkits" giving innovators guidance on how to protect intellectual property in specific countries such as China, Korea and Mexico. For questions regarding patent enforcement issues, applicants may call the U.S. Government hotline at 1-866-999-HALT (1-866-999-4158).

LICENSE FOR FOREIGN FILING UNDER

Title 35, United States Code, Section 184

Title 37, Code of Federal Regulations, 5.11 & 5.15

GRANTED

The applicant has been granted a license under 35 U.S.C. 184, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" followed by a date appears on this form. Such licenses are issued in all applications where

the conditions for issuance of a license have been met, regardless of whether or not a license may be required as set forth in 37 CFR 5.15. The scope and limitations of this license are set forth in 37 CFR 5.15(a) unless an earlier license has been issued under 37 CFR 5.15(b). The license is subject to revocation upon written notification. The date indicated is the effective date of the license, unless an earlier license of similar scope has been granted under 37 CFR 5.13 or 5.14.

This license is to be retained by the licensee and may be used at any time on or after the effective date thereof unless it is revoked. This license is automatically transferred to any related applications(s) filed under 37 CFR 1.53(d). This license is not retroactive.

The grant of a license does not in any way lessen the responsibility of a licensee for the security of the subject matter as imposed by any Government contract or the provisions of existing laws relating to espionage and the national security or the export of technical data. Licensees should apprise themselves of current regulations especially with respect to certain countries, of other agencies, particularly the Office of Defense Trade Controls, Department of State (with respect to Arms, Munitions and Implements of War (22 CFR 121-128)); the Bureau of Industry and Security, Department of Commerce (15 CFR parts 730-774); the Office of Foreign Assets Control, Department of Treasury (31 CFR Parts 500+) and the Department of Energy.

NOT GRANTED

No license under 35 U.S.C. 184 has been granted at this time, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" DOES NOT appear on this form. Applicant may still petition for a license under 37 CFR 5.12, if a license is desired before the expiration of 6 months from the filing date of the application. If 6 months has lapsed from the filing date of this application and the licensee has not received any indication of a secrecy order under 35 U.S.C. 181, the licensee may foreign file the application pursuant to 37 CFR 5.15(b).

SelectUSA

The United States represents the largest, most dynamic marketplace in the world and is an unparalleled location for business investment, innovation and commercialization of new technologies. The USA offers tremendous resources and advantages for those who invest and manufacture goods here. Through SelectUSA, our nation works to encourage, facilitate, and accelerate business investment. To learn more about why the USA is the best country in the world to develop technology, manufacture products, and grow your business, visit SelectUSA.gov.

To: IPR.Inhouse@nokia.com,brian.odonoghue@nokia.com,thao.pham@nokia.com
From: PAIR_eOfficeAction@uspto.gov
Cc: PAIR_eOfficeAction@uspto.gov
Subject: Private PAIR Correspondence Notification for Customer Number 73658

Nov 26, 2012 05:21:34 AM

Dear PAIR Customer:

Nokia, Inc.
Attn: Intellectual Property Rights Docketing
200 South Mathilda Ave
Sunnyvale, CA 94086
UNITED STATES

The following USPTO patent application(s) associated with your Customer Number, 73658 , have new outgoing correspondence. This correspondence is now available for viewing in Private PAIR.

The official date of notification of the outgoing correspondence will be indicated on the form PTOL-90 accompanying the correspondence.

Disclaimer:

The list of documents shown below is provided as a courtesy and is not part of the official file wrapper. The content of the images shown in PAIR is the official record.

Application	Document	Mailroom Date	Attorney Docket No.
13666680	NTC.MISS.PRT	11/26/2012	NC77198US-NP
	APP.FILE.REC	11/26/2012	NC77198US-NP

To view your correspondence online or update your email addresses, please visit us anytime at <https://sportal.uspto.gov/secure/myportal/privatepair>.

If you have any questions, please email the Electronic Business Center (EBC) at EBC@uspto.gov with 'e-Office Action' on the subject line or call 1-866-217-9197 during the following hours:

Monday - Friday 6:00 a.m. to 12:00 a.m.

Thank you for prompt attention to this notice,

UNITED STATES PATENT AND TRADEMARK OFFICE
PATENT APPLICATION INFORMATION RETRIEVAL SYSTEM

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

POWER OF ATTORNEY and CORRESPONDENCE ADDRESS INDICATION FORM

Application Number	13/666680
Filing Date	November 1 2012
First Named Inventor	Mehmet Oguz BICI
Title	METHOD FOR CODING AND AN
Art Unit	TBD
Examiner Name	TBD
Attorney Docket Number	NC77198US-NP

I hereby revoke all previous powers of attorney given in the above-identified application.

I hereby appoint:

☒ Practitioners associated with the Customer Number:

73658

OR

☐ Practitioner(s) named below:

Name	Registration Number

as my/our attorney(s) or agent(s) to prosecute the application identified above, and to transact all business in the United States Patent and Trademark Office connected therewith.

Please recognize or change the correspondence address for the above-identified application to:

☒ The address associated with the above-mentioned Customer Number:

OR

☐ The address associated with Customer Number:

OR

☐ Firm or Individual Name

Address

City

State

Zip

Country

Telephone

Email

I am the:

☐ Applicant/Inventor.

☒ Assignee of record of the entire interest. See 37 CFR 3.71.
Statement under 37 CFR 3.73(b) is enclosed. (Form PTO/SB/96)

SIGNATURE of Applicant or Assignee of Record

Signature	/Thomas J. Arria/	Date	January 2, 2013
Name	Thomas J. Arria	Telephone	781-219-8760
Title and Company	Senior IPR Specialist, Nokia Corporation		

NOTE: Signatures of all the inventors or assignees of record of the entire interest or their representative(s) are required. Submit multiple forms if more than one signature is required, see below*.

☒ *Total of one forms are submitted.

This collection of information is required by 37 CFR 1.31, 1.32 and 1.33. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 3 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

LIMITED POWER OF ATTORNEY

that we, the undersigned, under the authority granted to each of us to sign jointly on behalf of

NOKIA CORPORATION
Keilalahdentie 4
02150 Espoo
Finland

hereby authorize the following person

Thomas Arria
Nokia Corporation
15 Wayside Road
01803 Burlington
U.S.A.

Nokia Corporation
P.O.Box 226,
FI-00045 NOKIA GROUP, Finland
(Keilalahdentie 4, 02150 ESPOO)

Tel. +358 7180 08000
Fax +358 7180 38226

Business Identity Code
0112038-9, Helsinki

in his capacity as Senior IP Specialist, Patenting, to sign alone in the name of Nokia Corporation all the documents that relate to the filing, prosecution and registration of patent applications, as well as post-grant operations relating to patents granted based on such patent applications. Such documents may include without limitation case-specific Powers of Attorney, Assignment Deeds for assigning the ownership of an invention to Nokia, Novelty Declarations and Inventorship Declarations.

The rights granted under this Power of Attorney may not be transferred further, other than as specified above.

The authority granted by this Power of Attorney shall expire six (6) months from the date of execution.

Notwithstanding the foregoing, the authority granted by this Power of Attorney shall expire when the authorized person's employment with Nokia ends.

This Power of Attorney is duly signed on this day of January 1, 2013.

NOKIA CORPORATION

By: 
Name: Harri Honkasalo
Title: Director,
Legal and Intellectual Property

By: 
Name: Paul Melin
Title: Vice President,
Intellectual Property

STATEMENT UNDER 37 CFR 3.73(b)

Applicant/Patent Owner: Mehmet Oguz BICI

Application No./Patent No.: 13/666680 Filed/Issue Date: November 1 2012

Entitled: METHOD FOR CODING AND AN APPARATUS

Nokia Corporation, a Corporation
(Name of Assignee) (Type of Assignee, e.g., corporation, partnership, university, government agency, etc.)

states that it is:

1. ☒ the assignee of the entire right, title, and interest; or
2. ☐ an assignee of less than the entire right, title and interest
(The extent (by percentage) of its ownership interest is _____ %)

in the patent application/patent identified above by virtue of either:

A. ☒ An assignment from the inventor(s) of the patent application/patent identified above. The assignment was recorded in the United States Patent and Trademark Office at Reel _____, Frame _____, or for which a copy thereof is attached.

OR

B. ☐ A chain of title from the inventor(s), of the patent application/patent identified above, to the current assignee as follows:

1. From: _____ To: _____
The document was recorded in the United States Patent and Trademark Office at
Reel _____, Frame _____, or for which a copy thereof is attached.
2. From: _____ To: _____
The document was recorded in the United States Patent and Trademark Office at
Reel _____, Frame _____, or for which a copy thereof is attached.
3. From: _____ To: _____
The document was recorded in the United States Patent and Trademark Office at
Reel _____, Frame _____, or for which a copy thereof is attached.

☐ Additional documents in the chain of title are listed on a supplemental sheet.

☒ As required by 37 CFR 3.73(b)(1)(i), the documentary evidence of the chain of title from the original owner to the assignee was, or concurrently is being, submitted for recordation pursuant to 37 CFR 3.11.

[NOTE: A separate copy (*i.e.*, a true copy of the original assignment document(s)) must be submitted to Assignment Division in accordance with 37 CFR Part 3, to record the assignment in the records of the USPTO. See MPEP 302.08]

The undersigned (whose title is supplied below) is authorized to act on behalf of the assignee.

/Thomas J. Arria/
Signature
Thomas J. Arria

Printed or Typed Name
Senior IPR Specialist, Nokia Corporation

Title

January 2, 2013

Date
781-219-8760

Telephone Number

This collection of information is required by 37 CFR 3.73(b). The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

ASSIGNMENT

Title of Invention	METHOD FOR CODING AND AN APPARATUS
This Assignment is directed to (check all that apply): ☒ U.S. Provisional No. <u>61/555703</u> filed on <u>November 4, 2011</u> ☒ U.S. Application No. <u>13/666680</u> filed on <u>November 1, 2012</u> ☐ which is a 371 National Stage of _____ filed on _____	

As the below named inventor(s), I/we have invented certain improvements in the patent application for which I/we have filed an application for Letters Patent of the United States of America as identified above;

WHEREAS, I/we authorize the attorney of record to update this document to include Patent Office information as deemed necessary (i.e., filing date, serial number, etc.);

WHEREAS, NOKIA CORPORATION, a corporation organized under the laws of Finland, having its principal office in Espoo, Finland (hereinafter referred to as "ASSIGNEE"), is desirous of acquiring the entire right, title and interest in and under the said invention and the said application, and in and to any and all Letters Patent which shall be granted therefore in the United States of America and in any and all foreign countries;

NOW, THEREFORE, for good and valuable consideration, the receipt and sufficiency of which is hereby acknowledged, I/we have sold and do hereby sell, assign, transfer and convey unto said ASSIGNEE, its successors, assigns and legal representatives, the entire right, title and interest in and to said invention and application, and in all divisionals, reissues, substitutions, continuations, continuation-in-part and, in any and all Letters Patents of the United States of America and all foreign countries or reissues, reexaminations, or extensions thereof which may be granted therefore or thereon, for the full end of the term for which said Letters Patent may be granted, together with the right to claim the priority of said application in all foreign countries in accordance with the International Convention, the same to be held and enjoyed by said ASSIGNEE, its successors and assigns, as fully and entirely as the same would have been held and enjoyed by me if this assignment and sale had not been made.

I/WE ALSO HEREBY authorize and request the Commissioner of Patents and Trademarks to issue all patents for said invention, or patents resulting therefrom to the said ASSIGNEE of my/our entire right, title and interest.

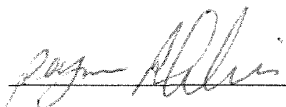
I/WE FURTHER HEREBY sell and assign to said ASSIGNEE, its successors, assigns and legal representatives the full and exclusive rights, title and interest to the invention disclosed in said application throughout the world, including, without limitation, the right to file applications and obtain patents, utility models, industrial models and designs for said invention in its own name throughout the world including all rights of priority, all rights to publish cautionary notices reserving ownership of said invention and all rights to register said invention in appropriate registries. I/we further agree to execute any and all powers of attorney, applications, assignments, declarations, affidavits, and any other papers in connection therewith necessary to perfect such rights, title and interest in ASSIGNEE, its successors, assigns and legal representatives.

I/WE HEREBY covenant that I/we have not and I/we will not execute any agreement in conflict herewith.

ASSIGNMENT

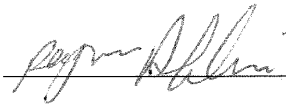
I/WE HEREBY further covenant, and agree to bind my/our heirs, legal representatives, and assigns, promptly to communicate to said ASSIGNEE or its representatives any facts known to me relating to said invention, to testify in any interference or legal proceedings involving said invention, to execute any additional papers which may be requested to confirm the right of the assignee, its representatives, successors, or assigns to secure patent or similar protection for the said invention in all countries and to vest in the assignee complete title to the said invention and Letters Patent, without further compensation, but at the expense of said ASSIGNEE, its successors, assigns, and other legal representatives.

IN WITNESS WHEREOF, I/we have hereunto set my/our hand and seal on the date indicated below.

Inventor: Mehmet Oguz BICI Date: NOV -05 - 2012
Signature:  Residence: Tampere, Finland
Witness Signature: 
Printed Name: PAYMAN AFLAKI Date: 5.11.2012

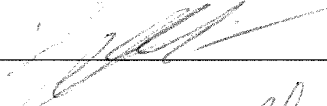
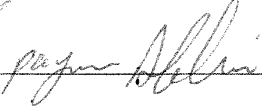
ASSIGNMENT

IN WITNESS WHEREOF, I/we have hereunto set my/our hand and seal on the date indicated below.

Inventor: Jani LAINEMA Date: NOV-02-2012
Signature:  Residence: Tampere, Finland
Witness Signature: 
Printed Name: PAYMAN AFLAKI Date: 2/11/2012

ASSIGNMENT

IN WITNESS WHEREOF, I/we have hereunto set my/our hand and seal on the date indicated below.

Inventor: Kemal UGUR Date: Nov - 02 - 2012
Signature:  Residence: Tampere, Finland
Witness Signature: 
Printed Name: PAYMAN AFLAKI Date: 2/11/2012

Electronic Patent Application Fee Transmittal				
Application Number:		13666680		
Filing Date:		01-Nov-2012		
Title of Invention:		METHOD FOR CODING AND AN APPARATUS		
First Named Inventor/Applicant Name:		Mehmet Oguz BICI		
Filer:		Thomas Joseph Arria/thao pham		
Attorney Docket Number:		NC77198US-NP		
Filed as Large Entity				
Utility under 35 USC 111(a) Filing Fees				
Description	Fee Code	Quantity	Amount	Sub-Total in USD(\$)
Basic Filing:				
Pages:				
Claims:				
Miscellaneous-Filing:				
Late filing fee for oath or declaration	1051	1	130	130
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD(\$)
Miscellaneous:				
Total in USD (\$)				130

Electronic Acknowledgement Receipt	
EFS ID:	14597236
Application Number:	13666680
International Application Number:	
Confirmation Number:	4782
Title of Invention:	METHOD FOR CODING AND AN APPARATUS
First Named Inventor/Applicant Name:	Mehmet Oguz BICI
Customer Number:	73658
Filer:	Thomas Joseph Arria/thao pham
Filer Authorized By:	Thomas Joseph Arria
Attorney Docket Number:	NC77198US-NP
Receipt Date:	02-JAN-2013
Filing Date:	01-NOV-2012
Time Stamp:	15:16:04
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	Deposit Account
Payment was successfully received in RAM	\$130
RAM confirmation Number	1609
Deposit Account	500270
Authorized User	
The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows: Charge any Additional Fees required under 37 C.F.R. Section 1.21 (Miscellaneous fees and charges)	

File Listing:					
Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Oath or Declaration filed	77198_executed_declaration.pdf	268093	no	3
			0df10ceb9f01d2e1b415cfb34f4031e3cf6351c8		
Warnings:					
Information:					
2		77198_POA_373b.pdf	378033	yes	7
			2a71d054ee4ec5f0e34ccb017d1ff096fe0b6141		
	Multipart Description/PDF files in .zip description				
	Document Description		Start	End	
	Power of Attorney		1	2	
	Assignee showing of ownership per 37 CFR 3.73.		3	7	
Warnings:					
Information:					
3	Fee Worksheet (SB06)	fee-info.pdf	30297	no	2
			4e38822359718c4cd9a4820faa6e6fd59c1598db		
Warnings:					
Information:					
Total Files Size (in bytes):			676423		
This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503. <u>New Applications Under 35 U.S.C. 111</u> If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application. <u>National Stage of an International Application under 35 U.S.C. 371</u> If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course. <u>New International Application Filed with the USPTO as a Receiving Office</u> If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.					

DECLARATION (37 CFR 1.63) FOR UTILITY OR DESIGN APPLICATION USING AN APPLICATION DATA SHEET (37 CFR 1.76)

Title of Invention	METHOD FOR CODING AND AN APPARATUS
As the below named inventor, I hereby declare that: This declaration is directed to: ☐ The attached application, or ☒ United States application or PCT international application number <u>13/666680</u> filed on <u>November 1, 2012</u>. The above-identified application was made or authorized to be made by me. I believe that I am the original inventor or an original joint inventor of a claimed invention in the application. I hereby acknowledge that any willful false statement made in this declaration is punishable under 18 U.S.C. 1001 by fine or imprisonment of not more than five (5) years, or both. WARNING: Petitioner/applicant is cautioned to avoid submitting personal information in documents filed in a patent application that may contribute to identity theft. Personal information such as social security numbers, bank account numbers, or credit card numbers (other than a check or credit card authorization form PTO-2038 submitted for payment purposes) is never required by the USPTO to support a petition or an application. If this type of personal information is included in documents submitted to the USPTO, petitioners/applicants should consider redacting such personal information from the documents before submitting them to the USPTO. Petitioner/applicant is advised that the record of a patent application is available to the public after publication of the application (unless a non-publication request in compliance with 37 CFR 1.213(a) is made in the application) or issuance of a patent. Furthermore, the record from an abandoned application may also be available to the public if the application is referenced in a published application or an issued patent (see 37 CFR 1.14). Checks and credit card authorization forms PTO-2038 submitted for payment purposes are not retained in the application file and therefore are not publicly available.	
LEGAL NAME OF INVENTOR Inventor: <u>Mehmet Oguz BICI</u> Date (Optional): <u>NOV - 05 - 2012</u> Signature: <u></u>	
Note: An application data sheet (PTO/SB/14 or equivalent), including naming the entire inventive entity, must accompany this form. Use an additional PTO/AIA/01 form for each additional inventor.	

This collection of information is required by 35 U.S.C. 115 and 37 CFR 1.63. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 1 minute to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

**DECLARATION (37 CFR 1.63) FOR UTILITY OR DESIGN APPLICATION USING AN
APPLICATION DATA SHEET (37 CFR 1.76)**

Title of Invention	METHOD FOR CODING AND AN APPARATUS
As the below named inventor, I hereby declare that: This declaration is directed to: ☐ The attached application, or ☒ United States application or PCT international application number <u>13/666680</u> filed on <u>November 1, 2012</u>. The above-identified application was made or authorized to be made by me. I believe that I am the original inventor or an original joint inventor of a claimed invention in the application. I hereby acknowledge that any willful false statement made in this declaration is punishable under 18 U.S.C. 1001 by fine or imprisonment of not more than five (5) years, or both. WARNING: Petitioner/applicant is cautioned to avoid submitting personal information in documents filed in a patent application that may contribute to identity theft. Personal information such as social security numbers, bank account numbers, or credit card numbers (other than a check or credit card authorization form PTO-2038 submitted for payment purposes) is never required by the USPTO to support a petition or an application. If this type of personal information is included in documents submitted to the USPTO, petitioners/applicants should consider redacting such personal information from the documents before submitting them to the USPTO. Petitioner/applicant is advised that the record of a patent application is available to the public after publication of the application (unless a non-publication request in compliance with 37 CFR 1.213(a) is made in the application) or issuance of a patent. Furthermore, the record from an abandoned application may also be available to the public if the application is referenced in a published application or an issued patent (see 37 CFR 1.14). Checks and credit card authorization forms PTO-2038 submitted for payment purposes are not retained in the application file and therefore are not publicly available. LEGAL NAME OF INVENTOR Inventor: <u>Jani LAINEMA</u> Date (Optional): <u>NOV-02-2012</u> Signature: <u></u> Note: An application data sheet (PTO/SB/14 or equivalent), including naming the entire inventive entity, must accompany this form. Use an additional PTO/AIA/01 form for each additional inventor.	

This collection of information is required by 35 U.S.C. 115 and 37 CFR 1.63. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 1 minute to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

DECLARATION (37 CFR 1.63) FOR UTILITY OR DESIGN APPLICATION USING AN APPLICATION DATA SHEET (37 CFR 1.76)

Title of Invention	METHOD FOR CODING AND AN APPARATUS
As the below named inventor, I hereby declare that: This declaration is directed to: ☐ The attached application, or ☒ United States application or PCT international application number <u>13/666680</u> filed on <u>November 1, 2012</u>. The above-identified application was made or authorized to be made by me. I believe that I am the original inventor or an original joint inventor of a claimed invention in the application. I hereby acknowledge that any willful false statement made in this declaration is punishable under 18 U.S.C. 1001 by fine or imprisonment of not more than five (5) years, or both. WARNING: Petitioner/applicant is cautioned to avoid submitting personal information in documents filed in a patent application that may contribute to identity theft. Personal information such as social security numbers, bank account numbers, or credit card numbers (other than a check or credit card authorization form PTO-2038 submitted for payment purposes) is never required by the USPTO to support a petition or an application. If this type of personal information is included in documents submitted to the USPTO, petitioners/applicants should consider redacting such personal information from the documents before submitting them to the USPTO. Petitioner/applicant is advised that the record of a patent application is available to the public after publication of the application (unless a non-publication request in compliance with 37 CFR 1.213(a) is made in the application) or issuance of a patent. Furthermore, the record from an abandoned application may also be available to the public if the application is referenced in a published application or an issued patent (see 37 CFR 1.14). Checks and credit card authorization forms PTO-2038 submitted for payment purposes are not retained in the application file and therefore are not publicly available.	
LEGAL NAME OF INVENTOR Inventor: <u>Kemal UGUR</u> Date (Optional): <u>Nov-02-2012</u> Signature: 	
Note: An application data sheet (PTO/SB/14 or equivalent), including naming the entire inventive entity, must accompany this form. Use an additional PTO/AIA/01 form for each additional inventor.	

This collection of information is required by 35 U.S.C. 115 and 37 CFR 1.63. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 1 minute to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

Doc code: IDS

Doc description: Information Disclosure Statement (IDS) Filed

PTO/SB/08a (01-10)

Approved for use through 07/31/2012. OMB 0651-0031

U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)	Application Number		13666680	
	Filing Date		2012-11-01	
	First Named Inventor	Mehmet Oguz BICI		
	Art Unit	2631		
	Examiner Name	TBD		
	Attorney Docket Number	NC77198US-NP		

U.S. PATENTS							Remove	
Examiner Initial*	Cite No	Patent Number	Kind Code ¹	Issue Date	Name of Patentee or Applicant of cited Document	Pages, Columns, Lines where Relevant Passages or Relevant Figures Appear		
	1							
If you wish to add additional U.S. Patent citation information please click the Add button.							Add	
U.S. PATENT APPLICATION PUBLICATIONS							Remove	
Examiner Initial*	Cite No	Publication Number	Kind Code ¹	Publication Date	Name of Patentee or Applicant of cited Document	Pages, Columns, Lines where Relevant Passages or Relevant Figures Appear		
	1	20120307905	A1	2012-12-06	Sunyeon et al.			
	2	20110170602	A1	2011-07-14	Tammy et al.			
If you wish to add additional U.S. Published Application citation information please click the Add button.							Add	
FOREIGN PATENT DOCUMENTS							Remove	
Examiner Initial*	Cite No	Foreign Document Number ³	Country Code ² j	Kind Code ⁴	Publication Date	Name of Patentee or Applicant of cited Document	Pages, Columns, Lines where Relevant Passages or Relevant Figures Appear	T ⁵
	1	2011062392	WO	A2	2011-05-26	SK Telecom Co. Ltd. et al.		☐
If you wish to add additional Foreign Patent Document citation information please click the Add button							Add	
NON-PATENT LITERATURE DOCUMENTS							Remove	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)	Application Number		13666680
	Filing Date		2012-11-01
	First Named Inventor	Mehmet Oguz BICI	
	Art Unit	2631	
	Examiner Name	TBD	
	Attorney Docket Number	NC77198US-NP	

Examiner Initials*	Cite No	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc), date, pages(s), volume-issue number(s), publisher, city and/or country where published.	T ⁵
	1	International Search Report and Written Opinion received for corresponding Patent Cooperation Treaty Application No. PCT/FI2012/051070, dated March 27, 2013, 12 pages.	☐
	2	Oudin, S. et al.; "Block Merging for Quadtree-Based Video Coding", IEEE Int. Conf. on Multimedia and Expo, July 11-15, 2011, 6 pages.	☐
	3	Sullivan, G. J.; "Overview of the High Efficiency Video Coding (HEVC) Standard", IEEE Trans. on Circuits and Systems for Video Technology, Vol. 22, No. 12, December 2012, pp. 1649-1668.	☐

If you wish to add additional non-patent literature document citation information please click the Add button **Add**

EXAMINER SIGNATURE

Examiner Signature		Date Considered	
--------------------	--	-----------------	--

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through a citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

¹ See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. ² Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). ³ For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. ⁴ Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. ⁵ Applicant is to place a check mark here if English language translation is attached.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)	Application Number	13666680
	Filing Date	2012-11-01
	First Named Inventor	Mehmet Oguz BICI
	Art Unit	2631
	Examiner Name	TBD
	Attorney Docket Number	NC77198US-NP

CERTIFICATION STATEMENT

Please see 37 CFR 1.97 and 1.98 to make the appropriate selection(s):

☐ That each item of information contained in the information disclosure statement was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(1).

OR

☐ That no item of information contained in the information disclosure statement was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the person signing the certification after making reasonable inquiry, no item of information contained in the information disclosure statement was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(2).

☐ See attached certification statement.

☐ Fee set forth in 37 CFR 1.17 (p) has been submitted herewith.

☒ None

SIGNATURE

A signature of the applicant or representative is required in accordance with CFR 1.33, 10.18. Please see CFR 1.4(d) for the form of the signature.

Signature	/Thomas J. Arria/	Date (YYYY-MM-DD)	2013-04-22
Name/Print	Thomas J. Arria	Registration Number	60223

This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 1 hour to complete, including gathering, preparing and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these records.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

PATENT COOPERATION TREATY

From the INTERNATIONAL SEARCHING AUTHORITY

PCT

NOTIFICATION OF TRANSMITTAL OF
THE INTERNATIONAL SEARCH REPORT AND
THE WRITTEN OPINION OF THE INTERNATIONAL
SEARCHING AUTHORITY, OR THE DECLARATION

(PCT Rule 44.1)

To: Nokia Corporation IPR Department Jussi Jaatinen Keilalahdentie 4 FI- 02150 Espoo FINLAND		Date of mailing (<i>day/month/year</i>) 27 March 2013 (27.03.2013)
Applicant's or agent's file reference NC77198WO		FOR FURTHER ACTION See paragraphs 1 and 4 below
International application No. PCT/FI2012/051070		International filing date (<i>day/month/year</i>) 02 November 2012 (02.11.2012)
Applicant NOKIA CORPORATION		

1. ☒ The applicant is hereby notified that the international search report and the written opinion of the International Searching Authority have been established and are transmitted herewith.
- Filing of amendments and statement under Article 19:**
The applicant is entitled, if he so wishes, to amend the claims of the international application (see Rule 46):
- When?** The time limit for filing such amendments is normally two months from the date of transmittal of the international search report.
- Where?** Directly to the International Bureau of WIPO, 34 chemin des Colombettes
1211 Geneva 20, Switzerland, Facsimile No.: +41 22 338 82 70
- For more detailed instructions, see *PCT Applicant's Guide*, International Phase, paragraphs 9.004 – 9.011.**
2. ☐ The applicant is hereby notified that no international search report will be established and that the declaration under Article 17(2)(a) to that effect and the written opinion of the International Searching Authority are transmitted herewith.
3. ☐ **With regard to any protest** against payment of (an) additional fee(s) under Rule 40.2, the applicant is notified that:
- ☐ the protest together with the decision thereon has been transmitted to the International Bureau together with the applicant's request to forward the texts of both the protest and the decision thereon to the designated Offices.
- ☐ no decision has been made yet on the protest; the applicant will be notified as soon as a decision is made.
4. **Reminders**
- The applicant may submit comments on an informal basis on the written opinion of the International Searching Authority to the International Bureau. The International Bureau will send a copy of such comments to all designated Offices unless an international preliminary examination report has been or is to be established. Following the expiration of 30 months from the priority date, these comments will also be made available to the public.
- Shortly after the expiration of **18 months** from the priority date, the international application will be published by the International Bureau. If the applicant wishes to avoid or postpone publication, a notice of withdrawal of the international application, or of the priority claim, must reach the International Bureau before the completion of the technical preparations for international publication (Rules 90*bis*.1 and 90*bis*.3).
- Within **19 months** from the priority date, but only in respect of some designated Offices, a demand for international preliminary examination must be filed if the applicant wishes to postpone the entry into the national phase **until 30 months** from the priority date (in some Offices even later); otherwise, the applicant must, **within 20 months** from the priority date, perform the prescribed acts for entry into the national phase before those designated Offices.
- In respect of other designated Offices, the time limit of **30 months** (or later) will apply even if no demand is filed within 19 months.
- For details about the applicable time limits, Office by Office, see www.wipo.int/pct/en/texts/time_limits.html and the *PCT Applicant's Guide*, National Chapters.

Name and mailing address of the ISA/FI National Board of Patents and Registration of Finland P.O. Box 1160, FI-00101 HELSINKI, Finland Facsimile No. +358 9 6939 5328	Authorized officer Timo Laakso Telephone No. +358 9 6939 500
--	--

Form PCT/ISA/220 (July 2010)

PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference NC77198WO	FOR FURTHER ACTION see Form PCT/ISA/220 as well as, where applicable, item 5 below.	
International application No. PCT/FI2012/051070	International filing date (<i>day/month/year</i>) 02 November 2012 (02.11.2012)	(Earliest) Priority Date (<i>day/month/year</i>) 04 November 2011 (04.11.2011)
Applicant NOKIA CORPORATION		

This international search report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This international search report consists of a total of 4 sheets.

☒ It is also accompanied by a copy of each prior art document cited in this report.

1. Basis of the report

a. With regard to the **language**, the international search was carried out on the basis of:

☒ the international application in the language in which it was filed.

☐ a translation of the international application into _____ which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b)).

b. ☐ This international search report has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43.6bis(a)).

c. ☐ With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, see Box No. I.

2. ☐ **Certain claims were found unsearchable** (see Box No. II).

3. ☐ **Unity of invention is lacking** (see Box No. III).

4. With regard to the **title**,

☐ the text is approved as submitted by the applicant.

☒ the text has been established by this Authority to read as follows:

METHOD FOR VIDEO CODING AND APPARATUS

5. With regard to the **abstract**,

☒ the text is approved as submitted by the applicant.

☐ the text has been established, according to Rule 38.2, by this Authority as it appears in Box No. IV. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

6. With regard to the **drawings**,

a. the figure of the **drawings** to be published with the abstract is Figure No. 6b

☐ as suggested by the applicant.

☒ as selected by this Authority, because the applicant failed to suggest a figure.

☐ as selected by this Authority, because this figure better characterizes the invention.

b. ☐ none of the figures is to be published with the abstract.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/FI2012/051070

A. CLASSIFICATION OF SUBJECT MATTER

See extra sheet

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: G06T, H04N

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

FI, SE, NO, DK

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

EPO-Internal, WPI, Google Scholar, IEEE Xplore

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2011062392 A2 (SK TELECOM CO LTD et al.) 26 May 2011 (26.05.2011) abstract, claim 1 & US 2012307905 A1 (KIM SUNYEON et al) 6 December 2012 (06.12.2012)	1-20
A	US 2011170602 A1 (LEE TAMMY et al.) 14 July 2011 (14.07.2011) the whole document	1-20
A	LOUDIN, S. et al.: "Block merging for quadtree-based video coding", IEEE Int. Conf. on Multimedia and Expo, 11-15 July 2011, 6p the whole document	1-20
T	SULLIVAN, G. J.: "Overview of the High Efficiency Video Coding (HEVC) Standard", IEEE Trans. on Circuits and Systems for Video technology, vol. 22, no. 12, Dec. 2012, pp. 1649-1668 the whole document	1-20

☐ 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 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

26 March 2013 (26.03.2013)

Date of mailing of the international search report

27 March 2013 (27.03.2013)

Name and mailing address of the ISA/FI
National Board of Patents and Registration of Finland
P.O. Box 1160, FI-00101 HELSINKI, Finland
Facsimile No. +358 9 6939 5328

Authorized officer
Timo Laakso
Telephone No. +358 9 6939 500

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/FI2012/051070

Patent document cited in search report	Publication date	Patent family members(s)	Publication date
WO 2011062392 A2	26/05/2011	US 2012307905 A1	06/12/2012
		CN 102714720 A	03/10/2012
		KR 20110054592 A	25/05/2011
.....			
US 2011170602 A1	14/07/2011	US 2013044815 A1	21/02/2013
		MX 2012008229 A	17/08/2012
		EP 2524507 A2	21/11/2012
		CN 102792697 A	21/11/2012
		US 2012155542 A1	21/06/2012
		SG 182491 A1	30/08/2012
		WO 2011087321 A2	21/07/2011
		CA 2787006 A1	21/07/2011
		AU 2011205896 A1	30/08/2012
		KR 20110083365 A	20/07/2011
.....			

INTERNATIONAL SEARCH REPORT

International application No.

PCT/FI2012/051070

CLASSIFICATION OF SUBJECT MATTER

Int. Cl.

H04N 7/26 (2006.01)

H04N 7/50 (2006.01)

H04N 7/34 (2006.01)

G06T 9/00 (2006.01)

PATENT COOPERATION TREATY

From the
INTERNATIONAL SEARCHING AUTHORITY

PCT

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

(PCT Rule 43bis.1)

To: Nokia Corporation IPR Department Jussi Jaatinen Keilalahdentie 4 FI- 02150 Espoo FINLAND		Date of mailing (<i>day/month/year</i>) 27 March 2013 (27.03.2013)
Applicant's or agent's file reference NC77198WO		FOR FURTHER ACTION See paragraph 2 below
International application No. PCT/FI2012/051070	International filing date (<i>day/month/year</i>) 02 November 2012 (02.11.2012)	Priority date (<i>day/month/year</i>) 04 November 2011 (04.11.2011)
International Patent Classification (IPC) or both national classification and IPC See supplemental box		
Applicant NOKIA CORPORATION		

1. This opinion contains indications relating to the following items: ☒ Box No. I Basis of the opinion ☐ Box No. II Priority ☐ Box No. III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability ☐ Box No. IV Lack of unity of invention ☒ Box No. V Reasoned statement under Rule 43bis.1(a)(i) with regard to novelty, inventive step and industrial applicability; citations and explanations supporting such statement ☐ Box No. VI Certain documents cited ☒ Box No. VII Certain defects in the international application ☐ Box No. VIII Certain observations on the international application 2. FURTHER ACTION If a demand for international preliminary examination is made, this opinion will be considered to be a written opinion of the International Preliminary Examining Authority ("IPEA") except that this does not apply where the applicant chooses an Authority other than this one to be the IPEA and the chosen IPEA has notified the International Bureau under Rule 66.1bis(b) that written opinions of this International Searching Authority will not be so considered. If this opinion is, as provided above, considered to be a written opinion of the IPEA, the applicant is invited to submit to the IPEA a written reply together, where appropriate, with amendments, before the expiration of 3 months from the date of mailing of Form PCT/ISA/220 or before the expiration of 22 months from the priority date, whichever expires later. For further options, see Form PCT/ISA/220.
--

Name and mailing address of the ISA/FI National Board of Patents and Registration of Finland P.O. Box 1160, FI-00101 HELSINKI, Finland Facsimile No. +358 9 6939 5328	Date of completion of this opinion 26 March 2013 (26.03.2013)	Authorized officer Timo Laakso Telephone No. +358 9 6939 500
--	--	--

Form PCT/ISA/237 (cover sheet) (July 2011)

**WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY**

International application No.
PCT/FI2012/051070

Box No. I Basis of this opinion

1. With regard to the **language**, this opinion has been established on the basis of:
 - ☒ the international application in the language in which it was filed.
 - ☐ a translation of the international application into _____ which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b)).
2. ☐ This opinion has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43bis.1(a))
3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, this opinion has been established on the basis of a sequence listing filed or furnished:
 - a. (means)
 - ☐ on paper
 - ☐ in electronic form
 - b. (time)
 - ☐ in the international application as filed
 - ☐ together with the international application in electronic form
 - ☐ subsequently to this Authority for the purposes of search
4. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.
5. Additional comments:

**WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY**

International application No.
PCT/FI2012/051070

**Box No. V Reasoned statement under Rule 43bis.1(a)(i) with regard to novelty, inventive step or industrial applicability;
citations and explanations supporting such statement**

1. Statement

Novelty (N)	Claims <u>1-20</u>	YES
	Claims _____	NO
Inventive step (IS)	Claims <u>1-20</u>	YES
	Claims _____	NO
Industrial applicability (IA)	Claims <u>1-20</u>	YES
	Claims _____	NO

2. Citations and explanations:

2.1 Documents cited in the International Search Report

D1: WO 2011062392 A2 (SK TELECOM CO LTD et al.) 26 May 2011 (26.05.2011)
D2: US 2011170602 A1 (LEE TAMMY et al.) 14 July 2011 (14.07.2011)
D3: OUDIN, S. et al.: "Block merging for quadtree-based video coding", IEEE Int. Conf. on Multimedia and Expo, 11-15 July 2011, 6p.
D4: SULLIVAN, G. J.: "Overview of the High Efficiency Video Coding (HEVC) Standard", IEEE Trans. on Circuits and Systems for Video technology, vol. 22, no. 12, Dec. 2012, pp. 1649-1668.

It is assumed that the contents of WO2011062392 correspond to those of US2012307905 which belongs to the same patent family.

2.2 Claimed invention

As described in claim 1, the invention relates to a method for video coding, comprising:

(F1) receiving a block of pixels including a prediction unit;
(F2) determining a set of spatial motion vector prediction candidates for the block of pixels; the spatial motion vector prediction candidates being provided with motion information;
(F3) selecting a first spatial motion vector prediction candidate from the set of spatial motion vector prediction candidates as a potential spatial motion vector prediction candidate to be included in a merge list for the prediction unit;
(F4) determining a subset of spatial motion vector predictions based on the location of the block associated with the first spatial motion vector prediction candidate;
(F5) comparing motion information of the first spatial motion vector prediction candidate with motion information of the spatial motion vector prediction candidate in the determined subset of spatial motion vector prediction candidates;
(F6) if at least one of the comparisons indicates that the motion vector information of the spatial motion vector prediction candidates correspond with each other, excluding the first spatial motion vector prediction candidate from the merge list.

(F1)...(F6) denote the technical features of the invention. Independent claims 9 and 15-20 define another method, two apparatuses, two storage mediums and two apparatuses with the corresponding technical features.

In addition, dependent claims 2-8 and 10-14 specify preferred embodiments of the invention.

The objective of the invention is to provide an improved arrangement for video coding for mobile communications systems.

Continued to next page

Supplemental Box

Continuation of: Box V (1 / 2)

2.3 Novelty under PCT Article 33(2)

D1 (abstract, claim 1) discloses a method of encoding a motion vector that includes: selecting one of a plurality of predicted candidate motion vector sets by using motion information of neighboring blocks of a current block; selecting one of predicted candidate motion vectors within a selected predicted candidate motion vector set, as a predicted motion vector; encoding a differential motion vector representing a difference between a current motion vector or motion vector of the current block and a selected predicted motion vector; and encoding a predicted motion vector index indicating the selected predicted motion vector. As a motion vector is encoded after selecting an efficient predicted candidate motion vector set, the size of a differential vector to be encoded can be reduced without necessarily encoding additional information to indicate which set of predicted candidate motion vectors has been selected, resulting in improved compression efficiency of motion vectors and in turn the improved video compression efficiency.

D2 introduces methods and apparatuses for encoding and decoding a motion vector. The method of encoding a motion vector includes: selecting a mode from among a first mode in which information indicating a motion vector predictor of at least one motion vector predictor is encoded and a second mode in which information indicating generation of a motion vector predictor based on pixels included in a previously encoded area adjacent to a current block is encoded; determining a motion vector predictor of the current block according to the selected mode and encoding information about the motion vector predictor of the current block; and encoding a difference vector between a motion vector of the current block and the motion vector predictor of the current block.

D3 presents an approach for block merging that removes redundancies by using a single parameter for a whole motion-compensated region of contiguous blocks. The proposed technique provides improvement over the specified H.264/AVC video coding standard and is included in the novel High Efficiency Video Coding (HEVC) standard.

D4 is a tutorial overview article of the novel HEVC video coding standard.

Documents D1-D3 represent the most relevant prior art. D4 is a later document not in conflict with the application that is cited to understand the theory underlying the invention.

2.3.1 Independent claims

Document D1 is regarded as the prior art closest to the subject matter of independent claim 1.

Document D1 (abstract, claim 1) discloses a methods for encoding a motion vector by selecting a set of predicted candidate motion vectors by using motion information of neighboring blocks of a current block. The subject matter of claim 1 differs from this in that merging of spatial motion vectors is not considered in D1 as in feature (F6) of claim 1. The subject matter of claim 1 is therefore novel. The same applies to independent claims 9 and 15-20 with the corresponding technical features.

2.3.2 Dependent claims

Because independent claims 1 and 9 are novel, also dependent claims 2-8 and 10-14 are novel.

Continued to next page

Supplemental Box

Continuation of: Box V (2 / 2)

2.4 Inventive step under PCT Article 33(3)

2.4.1 Independent claims

The solution presented in claim 1 is neither described in nor made obvious by the available prior art. Thus, claim 1 involves an inventive step. The same applies to independent claims 9 and 15-20 with the corresponding technical features.

2.4.2 Dependent claims

Because independent claims 1 and 9 involve an inventive step, also dependent claims 2-8 and 10-14 involve an inventive step.

2.5 Industrial applicability under PCT Article 33(4)

Claims 1-20 meet the requirement of industrial applicability because the claimed subject matter can be made or used in industry.

**WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY**

International application No.
PCT/FI2012/051070

Box No. VII Certain defects in the international application

The following defects in the form or contents of the international application have been noted:

Contrary to the requirements of PCT Rule 5.1(a)(ii), no relevant prior art is indicated in the description, nor are the documents reflecting such art identified therein.

Claims are not drafted in a two-part form in accordance with PCT Rule 6.3(b), which in the present case would be appropriate, with those features known from the prior art placed in the preamble (PCT Rule 6.3(b)(i)) and the remaining features included in the characterising part (PCT Rule 6.3(b)(ii)).

The features of the claims are not provided with reference signs placed in parentheses (PCT Rule 6.2(b)).

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/FI2012/051070

Supplemental Box

In case the space in any of the preceding boxes is not sufficient.

Continuation of: International Patent Classification (IPC)

Int. Cl.

H04N 7/26 (2006.01)

H04N 7/50 (2006.01)

H04N 7/34 (2006.01)

G06T 9/00 (2006.01)

BLOCK MERGING FOR QUADTREE-BASED VIDEO CODING

S. Oudin¹, P. Helle¹, J. Stegemann¹, C. Bartnik¹, B. Bross¹, D. Marpe¹, H. Schwarz¹, and T. Wiegand^{1,2}

¹Image Processing Department
Fraunhofer Institute for Telecommunications -
Heinrich Hertz Institute
Einsteinufer 37, 10587 Berlin, Germany

²Image Communication Chair
Department of Telecommunication Systems
Technical University of Berlin
Einsteinufer 17, 10587 Berlin, Germany

ABSTRACT

Quadtree-based block partitioning together with motion-compensated prediction has proven to be an efficient approach in video compression. However, when dealing with spatially neighboring blocks in uniformly displaced regions, quadtree-based partitioning may lead to redundant sets of transmitted motion parameters. This paper proposes and describes a simple but efficient block merging algorithm that aims at removing those redundancies by using only a single parameter set for a whole motion-compensated region of contiguous blocks. Simulation results show that our proposed merging technique works more efficiently than the conceptually similar direct mode as, e.g., specified in H.264/AVC. Due to its efficiency and simplicity, our proposed merging approach has been adopted into the first test model of the high efficiency video coding (HEVC) standardization project, as currently pursued by ITU-T VCEG and ISO/IEC MPEG.

Index Terms— motion compensation, quadtree partition, block merging, H.265, HEVC

1. INTRODUCTION

In image and video coding, pictures and their corresponding sets of sample arrays are usually decomposed into blocks such that each block is associated with a particular set of coding parameters. In H.264/AVC [1], e.g., each block is either spatially or temporally predicted and the resulting prediction residual is represented by using transform coding. Temporal or inter-picture prediction typically uses samples of one or two displaced blocks in previously decoded and reconstructed reference pictures for the samples of a given picture block. The motion vectors corresponding to these displacements are typically coded in a differential way by using an appropriately specified motion-vector predictor. Thus, without any further provision, one or two sets of motion parameters have to be signaled for each inter-predicted block as motion information with each set consisting of a motion-vector difference and an index specifying the corresponding reference picture.

In H.264/AVC, macroblocks or sub-macroblocks of bi-predictively coded areas can also be coded in so-called *direct*

mode without the need of transmitting any additional motion information [1, 2]. The direct mode infers the motion information for prediction from previously decoded blocks. In addition, if no prediction residual signal is transmitted for a direct-coded macroblock, it is also referred to as *skipped*, and skipped macroblocks can be indicated very efficiently as for the *skip mode* in H.264/AVC.

In the present paper, we introduce a coding tool called *block merging* which is conceptually similar to the direct mode of H.264/AVC. Block merging has been designed as part of a novel quadtree-based video coding approach [3, 4]. It exploits the spatial redundancy of motion information in neighboring blocks belonging to potentially different branches and levels in the quadtree hierarchy and it operates in a way that is conceptually similar to what has already been introduced in previous related publications [5, 6, 7].

The presented block merging algorithm has been first introduced as a feature of the Fraunhofer HHI response [3] to the joint Call for Proposals (CfP) on high efficiency video coding (HEVC) technology that was issued by ITU-T VCEG and ISO/IEC MPEG [8]. The intention of the CfP was to identify video coding technology with substantially higher compression capabilities than the existing H.264/AVC standard.

The rest of the paper is organized as follows. The next section briefly introduces the HHI proposal for HEVC, also called *HHI video codec* in the sequel. Section 3 deals with quadtree based partitioning and its drawbacks. The proposed block merging algorithm is described in detail in Section 4. Conducted experiments and their results are presented in Sections 5 and 6. Section 7 concludes this paper.

2. OVERVIEW OF THE HHI VIDEO CODEC

The HHI video codec is based on the hybrid video coding approach using temporal and spatial prediction followed by transform coding of the residual and entropy coding [3, 4]. In this video codec, the size of the prediction blocks can be adaptively chosen by using a quadtree-based partitioning. The block size used for DCT-based coding of the residual signal is also derived by using quadtree-based partitioning of the corresponding prediction block. The concept of a macroblock

as the basic processing unit in standardized video coding [1] is generalized to what is called a *coding tree block* (CTB). Fig. 1(a) shows an example of a coding tree block for luma samples (solid bold outline) and how it is subdivided into prediction blocks (solid lines) and transforms blocks (dashed lines). In Fig. 1(b), the corresponding quadtree structure for CTB partitioning is shown. Each node of the so-called prediction quadtree, as shown in solid lines in the example of Fig. 1(b), which corresponds to a prediction block and its related residual signal, can be further split into transform blocks of smaller size than the given prediction block size. In the example of Fig. 1, this is illustrated by the dashed blocks and the corresponding dashed nested quadtrees.

Block merging, as further described below, enables the creation of regions composed of neighboring prediction blocks with identical motion information. This motion information is composed of one or two sets of a motion-vector difference and an index specifying the corresponding reference picture. However, the dimensions of the transform blocks for residual coding are still constrained by corresponding residual quadtree partition, independent of the merging operation. The direct mode which is used in our simulation as a reference is similar to the spatial one specified in H.264/AVC [1]. This spatial direct mode can be applied to inter-partitions in pictures using multi-hypothesis prediction. For these inter-partitions, a flag is signaled indicating whether the partition is coded in direct mode or not. Interleaved Motion Vector Predictor (IMVP), as introduced in [3, 4], interleaves the prediction and coding for the components of a motion vector. More precisely, it uses the vertical component in order to predict its horizontal component. In the HHI video codec, no explicit skip mode is used. However, for each leaf in the residual quadtree, there is a flag which signals explicitly the presence of non-vanishing residual coefficients in the corresponding block. For more details of the HHI video codec, the reader is referred to [3, 4].

3. BACKGROUND AND PROBLEM FORMULATION

The quadtree data structure provides a suitable framework in video compression for optimizing the tradeoff between model accuracy (e.g., given by distortion D) and model coding cost, typically measured in bit rate R . The existence of simple tree pruning algorithms optimizing a Lagrangian functional

$$D + \lambda * R,$$

allows the development of compression algorithms which are based on quadtrees and related representations. The HHI video codec is a quadtree-structured video coding approach with syntax elements carrying the subdivision information for blocks of different types. Such syntax elements may be related to subdivision of blocks for the purpose of prediction or transform coding, as already outlined in Section 2 above.

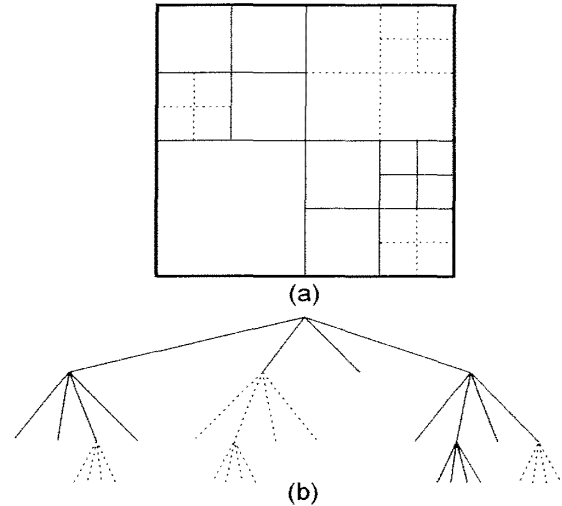


Fig. 1. (a) Example of coding tree block (solid bold) into prediction blocks (solid) and transform blocks (dashed) of variable size. (b) Corresponding nested quadtree structure.

However, the concept of quadtree-based image and video coding is intrinsically tied to certain drawbacks, as analyzed in [5]. For instance, the systematic decomposition of every block into four child blocks does not allow to jointly represent child blocks that belong to two different parent blocks. Also, if a given block is divided into four child blocks, all child blocks are typically coded separately, even if two or three of them may share the same coding parameters.

Thus, the sub-optimal rate-distortion properties of an initial quadtree-based subdivision can be substantially improved [5]. Block merging has been introduced in order to compensate the aforementioned drawbacks by enabling the merging of spatially neighboring blocks with identical coding parameters. As a consequence, block merging reduces the amount of redundant sets of transmitted coding parameters that are typically generated when using a quadtree-based block partitioning algorithm. In our proposed block merging algorithm, as further described below, specific syntax elements are introduced and coded together with the temporal prediction parameters for each leaf node of the prediction quadtree.

Merging is particularly well-suited for regions of homogeneous motion that can be sufficiently described by using a single motion parameter set. For example, a video signal often contains large spatially adjacent regions, each having homogeneous motion parameters that differ significantly among the regions. Inside each such region, the corresponding blocks can be merged, even if the maximum allowed block size for quadtree-based partitioning is reached.

Merging enables the creation of regions with arbitrary

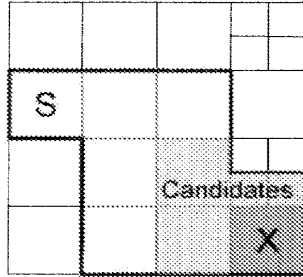


Fig. 2. A quadtree segmentation with multiple blocks merged into one region (bold line). Block S contains the prediction information for the region. Blocks in the gray shaded region are the candidates for block X to be merged with.

shapes. One illustrating example is shown in Fig. 2; other examples as outcome of our simulations are shown in Figs. 5–6. Ideally, the boundary of each region should coincide with the motion discontinuities in the given video signal. With quadtree-based block partitioning combined with merging, the picture can be subdivided into smaller and smaller blocks thereby approximating the motion boundary and minimizing the size of partitions containing the boundary. Subsequently, each created quadtree leaf block can be merged into a region on either side of the boundary.

4. THE PROPOSED BLOCK MERGING ALGORITHM

In order to illustrate the operation of our proposed block merging algorithm, Fig. 2 shows an example of a quadtree segmentation. The purpose of block merging, as already explained, is to compensate for the disadvantages of an initial quadtree-based subdivision by reducing the redundant sets of transmitted coding parameters. This is achieved by creating regions composed of neighboring prediction blocks with identical motion information. Each region must contain at least one block that is not merged, seeding the new motion information. In Fig. 2, block S (for seed) corresponds to such a particular block. Block X in Fig. 2 corresponds to the current block to be en-/decoded. All the blocks en-/decoded before the current block along the block scanning pattern form the set of causal blocks. Note that within each CTB, the subblocks are traversed in depth-first order. In any case, the block to be en-/decoded can be merged only with one of its causal neighboring blocks. In Fig. 2, these blocks are highlighted by the gray shaded region.

A list of candidate sets of motion information is created using the motion information collected from these neighboring blocks. If the block X is merged with one of these candidates, it means that block X copies and uses the same motion information of one of its neighbors. In order to identify

this motion information in the list, an index is signaled for the block X. The particular design of the block merging algorithm used to obtain simulation results for this paper only uses the motion parameter sets of the left and top neighboring blocks as candidates. Thus, the candidate list contains at most two entries.

The new syntax elements `merge_flag` and `merge_index` are introduced for signalization. The binary valued `merge_flag` is signaled for each motion-compensated partition block prior to any prediction parameter. The list of merge candidates may be empty because, e.g., no neighboring block is predicted by using motion compensation. In this case, block X cannot be merged and `merge_flag` is not transmitted. Otherwise, if `merge_flag` is equal to 0, the current block X is not merged with any of its neighboring candidate blocks and motion parameters for X are signaled explicitly. If the motion information of two different merge candidates are identical, the list of candidate sets of motion information can be reduced. In this way, the list of candidates of motion information is only composed of motion information which are different and thus, its size is as small as possible. If there is only one candidate in the list, no `merge_index` needs to be transmitted. Otherwise, the value of `merge_index` which is transmitted identifies the position of the candidate in the reduced list and its associated motion information which will be used for en-/decoding the current block X.

There is a conceptual similarity between the direct mode (of H.264/AVC) and our proposed block merging approach. Direct mode also infers the motion information for motion-compensated prediction (MCP) from previously decoded blocks. Thus, both algorithms are using motion information from neighboring blocks in order to reduce the bit rate required for transmitting motion parameters. However, for direct mode the motion information is derived independently for each reference picture. This means that for each reference picture (two in the HHI video codec), a motion predictor needs to be calculated from neighboring blocks for the current block to be en-/decoded. To derive a motion predictor in direct mode, up to four motion vectors of neighboring blocks are used, namely the above, left, above-left and above-right neighbors. The median of these vector components is used to calculate the motion predictor of the current block. For direct mode in H.264/AVC, also a flag is signaled for MCP partitions indicating whether the partition is coded in direct mode or not [1].

Despite this similarity, the two algorithms differ in the way they handle neighboring motion information. While the direct mode infers motion parameters from adjacent blocks, merging creates regions where all the blocks share the complete motion information. The creation of such regions is performed using only simple operations, such as comparing and copying the complete motion information from a neighbor. In comparison with this, the calculation of motion predictors in direct mode requires more computational complexity. Motion

Class	Sequence	BD rate CS1 [%]
B1 (1920x1080)	Kimono	-38.30
	ParkScene	-24.28
B2 (1920x1080)	Cactus	-29.22
	Basketball Drive	-35.97
	BQTerrace	-41.92
C (832x480)	Basketball Drill	-31.88
	BQMall	-29.71
	PartyScene	-28.09
	RaceHorses	-29.67
D (416x240)	Basketball Pass	-21.97
	BQSquare	-43.96
	BlowingBubbles	-23.60
	RaceHorses	-19.64
Average		-30.63

Table 1. Average BD rate savings (percentage) of the HHI codec relative to the H.264/AVC HP anchor at CS1 test conditions.

predictors are calculated for all reference pictures individually while using up to four neighboring blocks, where additional operations are required to calculate the mean/median values, in case there is more than one neighboring motion vector available.

5. SIMULATION SETTINGS AND RESULTS

The codec used for performing our experiments is the HHI video codec as presented in section 2. The experiments were conducted by encoding all test sequences, as defined in the CfP [8]. Two constraint sets (CS) restricting the temporal coding structure were defined in the CfP as follows:

- CS1: hierarchical B picture coding structure [9] with a structural temporal delay not larger than 8 pictures,
- CS2: hierarchical P picture coding structure without picture reordering in the decoder.

Furthermore, the coding tree block has a fixed size of 64×64 pixels (for luma) and a maximum quadtree depth of 4, resulting in a minimum block size of 4×4 . In order to obey CS1, inter-prediction with up to two hypotheses is employed within groups of 8 successive pictures in a hierarchical structure using 4 temporal layers. For CS2, inter prediction is restricted to one hypothesis from temporally preceding reference pictures. The intra period in CS1 is set depending on the number of frames per second for each sequence, whereas in CS2, only one intra picture at the beginning of each sequence has been utilized. Five quantization parameter (QP) values are chosen to obtain five RD points per sequence. The test sequences are separated into 4 classes according to their resolution and their content. For instance, the sequences of class E show typical videoconferencing scenes with static background. That is also why class E is only tested for CS2, the use case for low-delay applications. The other classes are characterized by different

Class	Sequences	BD rate CS1 [%]	BD rate CS2 [%]
B1 (1920x1080)	Kimono	-2.42	-2.37
	ParkScene	-2.78	-2.88
B2 (1920x1080)	Cactus	-2.65	-2.28
	Basketball Drive	-2.93	-2.56
	BQTerrace	-3.45	-2.62
C (832x480)	Basketball Drill	-2.79	-2.58
	BQMall	-4.80	-4.47
	PartyScene	-2.22	-3.46
	RaceHorses	-3.08	-3.00
D (416x240)	Basketball Pass	-2.82	-2.11
	BQSquare	-2.41	-2.14
	BlowingBubbles	-2.10	-1.98
	RaceHorses	-2.81	-2.51
Average		-2.87	-2.70
E (1280x720)	Vidyo1	n/a	-4.50
	Vidyo3	n/a	-5.79
	Vidyo4	n/a	-4.32
Total Average		-2.87	-3.10

Table 2. Averaged BD rate savings (in percentage) of the HHI codec with block merging relative to the same codec with direct mode (instead of block merging).

resolutions and a variety of content. With the exception of class E, all classes are tested for both CS1 and CS2. All presented bit rate savings are measured in terms of Bjontegaard Delta (BD) rate [10].

First, the HHI proposal for HEVC (which incorporates block merging) has been evaluated in comparison to H.264/AVC high profile (HP) as an anchor [8]. Table 1 shows the bit rate savings of the HHI proposal relative to this anchor. The average gain for CS1 is around 30%.

The efficiency of block merging has been evaluated by comparing two versions of the HHI codec. The first version is given by the original HHI proposal for HEVC (including block merging), while the other version incorporates direct mode instead of block merging, as described in Section 2. The results of this comparison at CS1 and CS2 test conditions with the direct mode version as a reference are listed in Table 2. It can be observed that the gains of block merging in terms of average BD rate savings are around 3%, and that, to a large extent, those gains are independent of resolution or sequence type. However, there are several sequences where coding efficiency lies significantly above the average such as, e.g., all class E sequences, as shown in Table 2. These particularly favorable results for class E are mostly due to the static background of these sequences which can be easily represented by using a single merged region with zero motion vector at the corresponding block merging seed.

In Fig. 3, the RD curves of the sequence BQMall for CS1 are presented. The curves show the RD performance of three competing schemes: the HHI codec with block merging (the HHI proposal for HEVC), the HHI codec with direct mode,

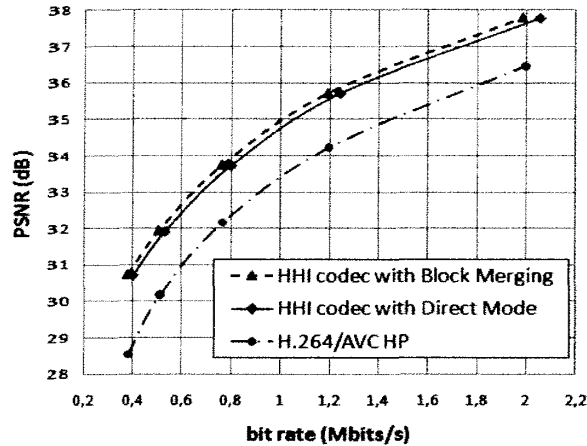


Fig. 3. RD performance plot comparing direct mode and block merging in the HHI video codec as well as H.264/AVC HP by using the test sequence BQMall at CS1 test conditions.

and the H.264/AVC HP anchor scheme. For BQMall, the average BD rate saving gain of the HHI proposal relative to H.264/AVC HP is 29.7%. The average gain relative to the HHI codec with direct mode is 4.8% BD rate savings, where the individual gains for the different rate points are distributed more or less uniformly, which has also been observed for all other sequences.

Fig. 4 compares the percentage of picture samples covered by MCP blocks using direct mode with that using block merging. The curves were obtained by encoding the BQMall sequence at different QP values using both versions of the HHI codec. As can be seen from Fig. 4, the percentage of coverage for direct mode in comparison with that for block merging is much smaller. This behavior was consistently observed for the entire test set of sequences, which is an indication for block merging being selected more often due to its more favorable RD cost. Also, it can be observed that with larger QP values, the coverage increases for either method, although more distinctively for the case of block merging.

6. ILLUSTRATION OF THE RESULTS OBTAINED WITH BLOCK MERGING

In the following, details of encoded frames are further analyzed to demonstrate the motion modeling capabilities of block merging. In Figs. 5 and 6, white lines separate merged regions and grey lines indicate MCP leaf blocks of CTBs.

Fig. 5 shows a detail from a picture belonging to the BQMall sequence. In this part of the sequence, the background is static and only people in the foreground are moving. As can be seen from Fig. 5, the boundaries of the merged regions are

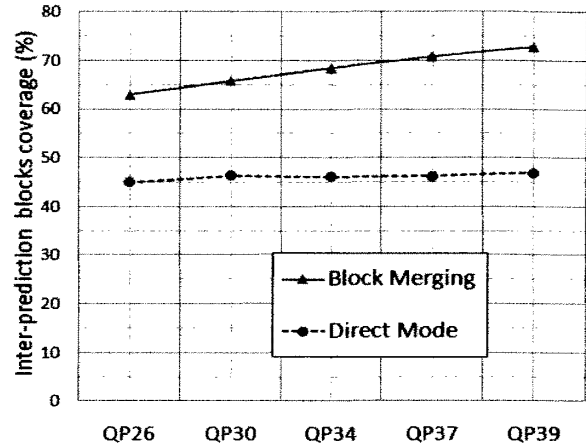


Fig. 4. Percentage coverage of MCP blocks using the direct mode and MCP blocks using block merging, both depending on the QP by using the sequence BQMall at CS1 test conditions.

clearly separating the moving persons from the static background.

Fig. 6 shows that block merging is effective also for more complex sequences. The corresponding Parkscene sequence was shot with a panning camera in front of a scene consisting of a park in the background and some tree branches in the foreground. The figure reveals distinct 'object planes' in the picture, all of which are moving differently such that each plane can be identified by a different set of block merging regions in Fig. 6. The tree branches correspond to the foreground, the colonnade belongs to a second plane, and the rest of the park can be attributed to the background.

As in the two examples examined above, natural video signals very often contain distinct moving objects, each with a different kind of motion. At the boundary of such objects, the field of motion parameters can be expected to exhibit very abrupt changes. These motion discontinuities can be modeled very efficiently by block merging, as it selects exactly one of the neighboring blocks to infer the motion information for the current block, leaving other candidates aside. In contrast to that, direct mode uses up to four spatially neighboring blocks to calculate an intermediate motion representation which, in the presence of distinct moving objects, may not be adequately representing the real motion in a given scene.

7. CONCLUSION AND FUTURE WORK

In this paper, block merging together with a quadtree-based block partitioning concept has been presented as an alternative to the conceptually similar direct mode for motion-compensated prediction. Simulation results show that the

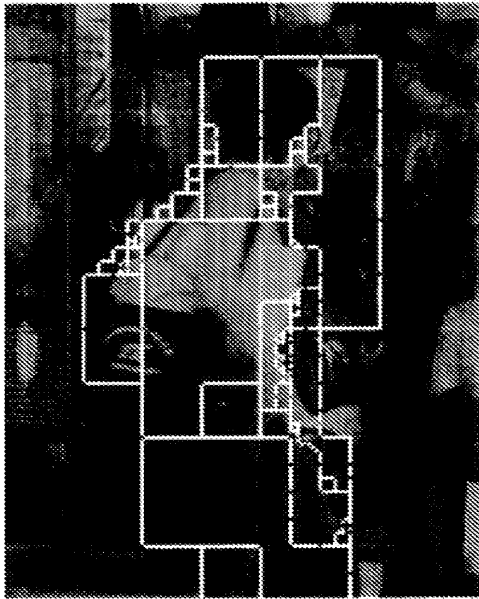


Fig. 5. Detail of sequence BQMall in Class C with a static background and moving persons in the foreground.

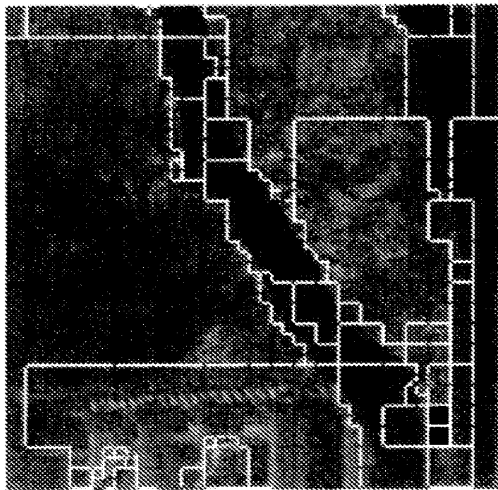


Fig. 6. Detail of sequence Parkscene in Class B with tree branches in the foreground and colonnades at the lower picture boundary, captured by a panning camera.

proposed merging algorithm better exploits spatiotemporal correlations in motion model parameters. Averaged BD rate savings of around 3% for a whole test set of video sequences have been achieved in comparison to the direct mode. Block merging not only yields better RD coding performance than the traditional direct mode, but also offers a computationally less demanding algorithm. These benefits were crucial for the proposed block merging algorithm to be adopted as part of the working draft of HEVC [11] which has been currently issued by the Joint Collaborative Team on Video Coding (JCT-VC). Further aspects of block merging within the HEVC test model are currently under investigation and further study [12].

8. REFERENCES

- [1] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra, "Overview of the H.264/AVC Video Coding Standard, *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, pp. 560–576, Jul. 2003.
- [2] A. M. Tourapis, F. Wu, and S. Li, "Direct Mode Coding for Bipedictive Slices in the H.264 Standard, *IEEE Trans. Circuits and Syst. Video Technol.*, vol. 15, no. 1, pp. 119–126, Jan. 2005.
- [3] M. Winken, S. Boße, B. Bross, P. Helle, T. Hinz, H. Kirchhoffer, H. Lakshman, D. Marpe, S. Oudin, M. Preiß, H. Schwarz, M. Siekmann, K. Sühring, T. Wiegand, "Description of video coding technology proposal by Fraunhofer HHI", Doc. *JCTVC-A116*, Dresden, Germany, Apr. 2010.
- [4] D. Marpe, H. Schwarz, S. Boße, B. Bross, P. Helle, T. Hinz, H. Kirchhoffer, H. Lakshman, T. Nguyen, S. Oudin, M. Siekmann, K. Sühring, M. Winken, and T. Wiegand, "Video Compression Using Quadrees, Leaf Merging and Novel Techniques for Motion Representation and Entropy Coding", *IEEE Trans. Circuits and Syst. Video Technol.*, vol. 20, no. 12, Dec. 2010.
- [5] R. Shukla, P. L. Dragotti, M. N. Do and M. Vetterli, "Rate-Distortion Optimized Tree-Structured Compression Algorithms for Piecewise Polynomial Images", *IEEE Trans. Image Process.*, vol. 14, no. 3, pp. 343–359, Mar. 2005.
- [6] R. De Forni and D. S. Taubman, "On the benefits of leaf merging in quadtree motion models", *Proc. IEEE Int. Conf. Image Process.*, vol. 2, pp. 858–861, Sep. 2005.
- [7] R. Mathew, D. S. Taubman, "Quad-Tree Motion Modeling with Leaf Merging", *IEEE Trans. Circuits and Syst. Video Technol.*, vol. 20, no. 10, pp. 1331–1345, Oct. 2010.
- [8] ISO/IEC JTC1/SC29/WG11 and ITU-T Q6/16, "Joint Call for Proposals on Video Compression Technology," *WG11 Doc. N11113 and ITU-T Q6/16 Doc. VCEG-AM91*, Kyoto, Jan. 2010.
- [9] H. Schwarz, D. Marpe, T. Wiegand, "Hierarchical B pictures," *Joint Video Team, Doc. JVT-P014*, Poznan, Poland, July 2005.
- [10] G. Bjøntegaard, "Calculation of Average PSNR Differences between RD Curves, *ITU-T Q6/16 Doc. VCEG-M33*, April 2001.
- [11] T. Wiegand, W.-J. Han, B. Bross, J. R. Ohm, G. J. Sullivan, "WD2: Working Draft 2 of High-Efficiency Video Coding" JCT-VC Meeting, Daegu, KR, Doc. *JCTVC-D503*, Output Document, Jan. 2011.
- [12] J. Jung, B. Bross, P. Chen, W.-J. Han, "Summary report of Core Experiment 9: MV Coding and Skip/Merge operations" JCT-VC Meeting, Geneva, Switzerland, Doc. *JCTVC-E029*, Input Document, Jan. 2011.

Overview of the High Efficiency Video Coding (HEVC) Standard

Gary J. Sullivan, *Fellow, IEEE*, Jens-Rainer Ohm, *Member, IEEE*, Woo-Jin Han, *Member, IEEE*, and Thomas Wiegand, *Fellow, IEEE*

Abstract—High Efficiency Video Coding (HEVC) is currently being prepared as the newest video coding standard of the ITU-T Video Coding Experts Group and the ISO/IEC Moving Picture Experts Group. The main goal of the HEVC standardization effort is to enable significantly improved compression performance relative to existing standards—in the range of 50% bit-rate reduction for equal perceptual video quality. This paper provides an overview of the technical features and characteristics of the HEVC standard.

Index Terms—Advanced video coding (AVC), H.264, High Efficiency Video Coding (HEVC), Joint Collaborative Team on Video Coding (JCT-VC), Moving Picture Experts Group (MPEG), MPEG-4, standards, Video Coding Experts Group (VCEG), video compression.

I. INTRODUCTION

THE High Efficiency Video Coding (HEVC) standard is the most recent joint video project of the ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC Moving Picture Experts Group (MPEG) standardization organizations, working together in a partnership known as the Joint Collaborative Team on Video Coding (JCT-VC) [1]. The first edition of the HEVC standard is expected to be finalized in January 2013, resulting in an aligned text that will be published by both ITU-T and ISO/IEC. Additional work is planned to extend the standard to support several additional application scenarios, including extended-range uses with enhanced precision and color format support, scalable video coding, and 3-D/stereo/multiview video coding. In ISO/IEC, the HEVC standard will become MPEG-H Part 2 (ISO/IEC 23008-2) and in ITU-T it is likely to become ITU-T Recommendation H.265.

Manuscript received May 25, 2012; revised August 22, 2012; accepted August 24, 2012. Date of publication October 2, 2012; date of current version January 8, 2013. This paper was recommended by Associate Editor H. Gharavi. (Corresponding author: W.-J. Han.)

G. J. Sullivan is with Microsoft Corporation, Redmond, WA 98052 USA (e-mail: garysull@microsoft.com).

J.-R. Ohm is with the Institute of Communication Engineering, RWTH Aachen University, Aachen 52056, Germany (e-mail: ohm@ient.rwth-aachen.de).

W.-J. Han is with the Department of Software Design and Management, Gachon University, Seongnam 461-701, Korea (e-mail: hurumi@gmail.com).

T. Wiegand is with the Fraunhofer Institute for Telecommunications, Heinrich Hertz Institute, Berlin 10587, Germany, and also with the Berlin Institute of Technology, Berlin 10587, Germany (e-mail: twiegand@ieee.org).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCSVT.2012.2221191

Video coding standards have evolved primarily through the development of the well-known ITU-T and ISO/IEC standards. The ITU-T produced H.261 [2] and H.263 [3], ISO/IEC produced MPEG-1 [4] and MPEG-4 Visual [5], and the two organizations jointly produced the H.262/MPEG-2 Video [6] and H.264/MPEG-4 Advanced Video Coding (AVC) [7] standards. The two standards that were jointly produced have had a particularly strong impact and have found their way into a wide variety of products that are increasingly prevalent in our daily lives. Throughout this evolution, continued efforts have been made to maximize compression capability and improve other characteristics such as data loss robustness, while considering the computational resources that were practical for use in products at the time of anticipated deployment of each standard.

The major video coding standard directly preceding the HEVC project was H.264/MPEG-4 AVC, which was initially developed in the period between 1999 and 2003, and then was extended in several important ways from 2003–2009. H.264/MPEG-4 AVC has been an enabling technology for digital video in almost every area that was not previously covered by H.262/MPEG-2 Video and has substantially displaced the older standard within its existing application domains. It is widely used for many applications, including broadcast of high definition (HD) TV signals over satellite, cable, and terrestrial transmission systems, video content acquisition and editing systems, camcorders, security applications, Internet and mobile network video, Blu-ray Discs, and real-time conversational applications such as video chat, video conferencing, and telepresence systems.

However, an increasing diversity of services, the growing popularity of HD video, and the emergence of beyond-HD formats (e.g., 4k×2k or 8k×4k resolution) are creating even stronger needs for coding efficiency superior to H.264/MPEG-4 AVC's capabilities. The need is even stronger when higher resolution is accompanied by stereo or multiview capture and display. Moreover, the traffic caused by video applications targeting mobile devices and tablet PCs, as well as the transmission needs for video-on-demand services, are imposing severe challenges on today's networks. An increased desire for higher quality and resolutions is also arising in mobile applications.

HEVC has been designed to address essentially all existing applications of H.264/MPEG-4 AVC and to particularly focus on two key issues: increased video resolution and increased use of parallel processing architectures. The syntax of HEVC

1051-8215/\$31.00 © 2012 IEEE

is generic and should also be generally suited for other applications that are not specifically mentioned above.

As has been the case for all past ITU-T and ISO/IEC video coding standards, in HEVC only the bitstream structure and syntax is standardized, as well as constraints on the bitstream and its mapping for the generation of decoded pictures. The mapping is given by defining the semantic meaning of syntax elements and a decoding process such that every decoder conforming to the standard will produce the same output when given a bitstream that conforms to the constraints of the standard. This limitation of the scope of the standard permits maximal freedom to optimize implementations in a manner appropriate to specific applications (balancing compression quality, implementation cost, time to market, and other considerations). However, it provides no guarantees of end-to-end reproduction quality, as it allows even crude encoding techniques to be considered conforming.

To assist the industry community in learning how to use the standard, the standardization effort not only includes the development of a text specification document, but also reference software source code as an example of how HEVC video can be encoded and decoded. The draft reference software has been used as a research tool for the internal work of the committee during the design of the standard, and can also be used as a general research tool and as the basis of products. A standard test data suite is also being developed for testing conformance to the standard.

This paper is organized as follows. Section II highlights some key features of the HEVC coding design. Section III explains the high-level syntax and the overall structure of HEVC coded data. The HEVC coding technology is then described in greater detail in Section IV. Section V explains the profile, tier, and level design of HEVC. Since writing an overview of a technology as substantial as HEVC involves a significant amount of summarization, the reader is referred to [1] for any omitted details. The history of the HEVC standardization effort is discussed in Section VI.

II. HEVC CODING DESIGN AND FEATURE HIGHLIGHTS

The HEVC standard is designed to achieve multiple goals, including coding efficiency, ease of transport system integration and data loss resilience, as well as implementability using parallel processing architectures. The following subsections briefly describe the key elements of the design by which these goals are achieved, and the typical encoder operation that would generate a valid bitstream. More details about the associated syntax and the decoding process of the different elements are provided in Sections III and IV.

A. Video Coding Layer

The video coding layer of HEVC employs the same hybrid approach (inter-/intrapicture prediction and 2-D transform coding) used in all video compression standards since H.261. Fig. 1 depicts the block diagram of a hybrid video encoder, which could create a bitstream conforming to the HEVC standard.

An encoding algorithm producing an HEVC compliant bitstream would typically proceed as follows. Each picture is split into block-shaped regions, with the exact block partitioning being conveyed to the decoder. The first picture of a video sequence (and the first picture at each clean random access point into a video sequence) is coded using only intrapicture prediction (that uses some prediction of data spatially from region-to-region within the same picture, but has no dependence on other pictures). For all remaining pictures of a sequence or between random access points, interpicture temporally predictive coding modes are typically used for most blocks. The encoding process for interpicture prediction consists of choosing motion data comprising the selected reference picture and motion vector (MV) to be applied for predicting the samples of each block. The encoder and decoder generate identical interpicture prediction signals by applying motion compensation (MC) using the MV and mode decision data, which are transmitted as side information.

The residual signal of the intra- or interpicture prediction, which is the difference between the original block and its prediction, is transformed by a linear spatial transform. The transform coefficients are then scaled, quantized, entropy coded, and transmitted together with the prediction information.

The encoder duplicates the decoder processing loop (see gray-shaded boxes in Fig. 1) such that both will generate identical predictions for subsequent data. Therefore, the quantized transform coefficients are constructed by inverse scaling and are then inverse transformed to duplicate the decoded approximation of the residual signal. The residual is then added to the prediction, and the result of that addition may then be fed into one or two loop filters to smooth out artifacts induced by block-wise processing and quantization. The final picture representation (that is a duplicate of the output of the decoder) is stored in a decoded picture buffer to be used for the prediction of subsequent pictures. In general, the order of encoding or decoding processing of pictures often differs from the order in which they arrive from the source; necessitating a distinction between the decoding order (i.e., bitstream order) and the output order (i.e., display order) for a decoder.

Video material to be encoded by HEVC is generally expected to be input as progressive scan imagery (either due to the source video originating in that format or resulting from deinterlacing prior to encoding). No explicit coding features are present in the HEVC design to support the use of interlaced scanning, as interlaced scanning is no longer used for displays and is becoming substantially less common for distribution. However, a metadata syntax has been provided in HEVC to allow an encoder to indicate that interlace-scanned video has been sent by coding each field (i.e., the even or odd numbered lines of each video frame) of interlaced video as a separate picture or that it has been sent by coding each interlaced frame as an HEVC coded picture. This provides an efficient method of coding interlaced video without burdening decoders with a need to support a special decoding process for it.

In the following, the various features involved in hybrid video coding using HEVC are highlighted as follows.

1) Coding tree units and coding tree block (CTB) structure:

The core of the coding layer in previous standards was

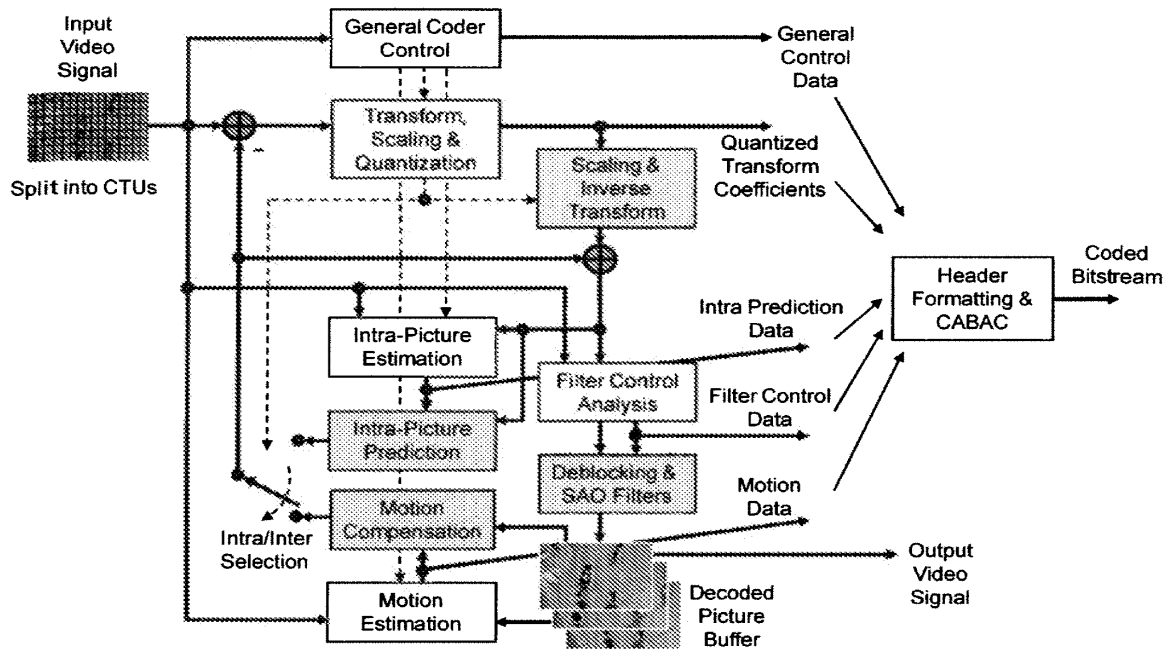


Fig. 1. Typical HEVC video encoder (with decoder modeling elements shaded in light gray).

the macroblock, containing a 16×16 block of luma samples and, in the usual case of 4:2:0 color sampling, two corresponding 8×8 blocks of chroma samples; whereas the analogous structure in HEVC is the coding tree unit (CTU), which has a size selected by the encoder and can be larger than a traditional macroblock. The CTU consists of a luma CTB and the corresponding chroma CTBs and syntax elements. The size $L \times L$ of a luma CTB can be chosen as $L = 16, 32$, or 64 samples, with the larger sizes typically enabling better compression. HEVC then supports a partitioning of the CTBs into smaller blocks using a tree structure and quadtree-like signaling [8].

- 2) *Coding units (CUs) and coding blocks (CBs)*: The quadtree syntax of the CTU specifies the size and positions of its luma and chroma CBs. The root of the quadtree is associated with the CTU. Hence, the size of the luma CTB is the largest supported size for a luma CB. The splitting of a CTU into luma and chroma CBs is signaled jointly. One luma CB and ordinarily two chroma CBs, together with associated syntax, form a coding unit (CU). A CTB may contain only one CU or may be split to form multiple CUs, and each CU has an associated partitioning into prediction units (PUs) and a tree of transform units (TUs).
- 3) *Prediction units and prediction blocks (PBs)*: The decision whether to code a picture area using interpicture or intrapicture prediction is made at the CU level. A PU partitioning structure has its root at the CU level.

Depending on the basic prediction-type decision, the luma and chroma CBs can then be further split in size and predicted from luma and chroma prediction blocks (PBs). HEVC supports variable PB sizes from 64×64 down to 4×4 samples.

- 4) *TUs and transform blocks*: The prediction residual is coded using block transforms. A TU tree structure has its root at the CU level. The luma CB residual may be identical to the luma transform block (TB) or may be further split into smaller luma TBs. The same applies to the chroma TBs. Integer basis functions similar to those of a discrete cosine transform (DCT) are defined for the square TB sizes 4×4 , 8×8 , 16×16 , and 32×32 . For the 4×4 transform of luma intrapicture prediction residuals, an integer transform derived from a form of discrete sine transform (DST) is alternatively specified.
- 5) *Motion vector signaling*: Advanced motion vector prediction (AMVP) is used, including derivation of several most probable candidates based on data from adjacent PBs and the reference picture. A merge mode for MV coding can also be used, allowing the inheritance of MVs from temporally or spatially neighboring PBs. Moreover, compared to H.264/MPEG-4 AVC, improved skipped and direct motion inference are also specified.
- 6) *Motion compensation*: Quarter-sample precision is used for the MVs, and 7-tap or 8-tap filters are used for interpolation of fractional-sample positions (compared to six-tap filtering of half-sample positions followed by linear interpolation for quarter-sample positions in

H.264/MPEG-4 AVC). Similar to H.264/MPEG-4 AVC, multiple reference pictures are used. For each PB, either one or two motion vectors can be transmitted, resulting either in unipredictive or bipredictive coding, respectively. As in H.264/MPEG-4 AVC, a scaling and offset operation may be applied to the prediction signal(s) in a manner known as weighted prediction.

- 7) *Intrapicture prediction*: The decoded boundary samples of adjacent blocks are used as reference data for spatial prediction in regions where interpicture prediction is not performed. Intrapicture prediction supports 33 directional modes (compared to eight such modes in H.264/MPEG-4 AVC), plus planar (surface fitting) and DC (flat) prediction modes. The selected intrapicture prediction modes are encoded by deriving most probable modes (e.g., prediction directions) based on those of previously decoded neighboring PBs.
- 8) *Quantization control*: As in H.264/MPEG-4 AVC, uniform reconstruction quantization (URQ) is used in HEVC, with quantization scaling matrices supported for the various transform block sizes.
- 9) *Entropy coding*: Context adaptive binary arithmetic coding (CABAC) is used for entropy coding. This is similar to the CABAC scheme in H.264/MPEG-4 AVC, but has undergone several improvements to improve its throughput speed (especially for parallel-processing architectures) and its compression performance, and to reduce its context memory requirements.
- 10) *In-loop deblocking filtering*: A deblocking filter similar to the one used in H.264/MPEG-4 AVC is operated within the interpicture prediction loop. However, the design is simplified in regard to its decision-making and filtering processes, and is made more friendly to parallel processing.
- 11) *Sample adaptive offset (SAO)*: A nonlinear amplitude mapping is introduced within the interpicture prediction loop after the deblocking filter. Its goal is to better reconstruct the original signal amplitudes by using a look-up table that is described by a few additional parameters that can be determined by histogram analysis at the encoder side.

B. High-Level Syntax Architecture

A number of design aspects new to the HEVC standard improve flexibility for operation over a variety of applications and network environments and improve robustness to data losses. However, the high-level syntax architecture used in the H.264/MPEG-4 AVC standard has generally been retained, including the following features.

- 1) *Parameter set structure*: Parameter sets contain information that can be shared for the decoding of several regions of the decoded video. The parameter set structure provides a robust mechanism for conveying data that are essential to the decoding process. The concepts of sequence and picture parameter sets from H.264/MPEG-4 AVC are augmented by a new video parameter set (VPS) structure.

- 2) *NAL unit syntax structure*: Each syntax structure is placed into a logical data packet called a network abstraction layer (NAL) unit. Using the content of a two-byte NAL unit header, it is possible to readily identify the purpose of the associated payload data.
- 3) *Slices*: A slice is a data structure that can be decoded independently from other slices of the same picture, in terms of entropy coding, signal prediction, and residual signal reconstruction. A slice can either be an entire picture or a region of a picture. One of the main purposes of slices is resynchronization in the event of data losses. In the case of packetized transmission, the maximum number of payload bits within a slice is typically restricted, and the number of CTUs in the slice is often varied to minimize the packetization overhead while keeping the size of each packet within this bound.
- 4) *Supplemental enhancement information (SEI) and video usability information (VUI) metadata*: The syntax includes support for various types of metadata known as SEI and VUI. Such data provide information about the timing of the video pictures, the proper interpretation of the color space used in the video signal, 3-D stereoscopic frame packing information, other display hint information, and so on.

C. Parallel Decoding Syntax and Modified Slice Structuring

Finally, four new features are introduced in the HEVC standard to enhance the parallel processing capability or modify the structuring of slice data for packetization purposes. Each of them may have benefits in particular application contexts, and it is generally up to the implementer of an encoder or decoder to determine whether and how to take advantage of these features.

- 1) *Tiles*: The option to partition a picture into rectangular regions called tiles has been specified. The main purpose of tiles is to increase the capability for parallel processing rather than provide error resilience. Tiles are independently decodable regions of a picture that are encoded with some shared header information. Tiles can additionally be used for the purpose of spatial random access to local regions of video pictures. A typical tile configuration of a picture consists of segmenting the picture into rectangular regions with approximately equal numbers of CTUs in each tile. Tiles provide parallelism at a more coarse level of granularity (picture/subpicture), and no sophisticated synchronization of threads is necessary for their use.
- 2) *Wavefront parallel processing*: When wavefront parallel processing (WPP) is enabled, a slice is divided into rows of CTUs. The first row is processed in an ordinary way, the second row can begin to be processed after only two CTUs have been processed in the first row, the third row can begin to be processed after only two CTUs have been processed in the second row, and so on. The context models of the entropy coder in each row are inferred from those in the preceding row with a two-CTU processing lag. WPP provides a form of processing parallelism at a rather fine level of

granularity, i.e., within a slice. WPP may often provide better compression performance than tiles (and avoid some visual artifacts that may be induced by using tiles).

- 3) *Dependent slice segments*: A structure called a dependent slice segment allows data associated with a particular wavefront entry point or tile to be carried in a separate NAL unit, and thus potentially makes that data available to a system for fragmented packetization with lower latency than if it were all coded together in one slice. A dependent slice segment for a wavefront entry point can only be decoded after at least part of the decoding process of another slice segment has been performed. Dependent slice segments are mainly useful in low-delay encoding, where other parallel tools might penalize compression performance.

In the following two sections, a more detailed description of the key features is given.

III. HIGH-LEVEL SYNTAX

The high-level syntax of HEVC contains numerous elements that have been inherited from the NAL of H.264/MPEG-4 AVC. The NAL provides the ability to map the video coding layer (VCL) data that represent the content of the pictures onto various transport layers, including RTP/IP, ISO MP4, and H.222.0/MPEG-2 Systems, and provides a framework for packet loss resilience. For general concepts of the NAL design such as NAL units, parameter sets, access units, the byte stream format, and packetized formatting, please refer to [9]–[11].

NAL units are classified into VCL and non-VCL NAL units according to whether they contain coded pictures or other associated data, respectively. In the HEVC standard, several VCL NAL unit types identifying categories of pictures for decoder initialization and random-access purposes are included. Table I lists the NAL unit types and their associated meanings and type classes in the HEVC standard.

The following subsections present a description of the new capabilities supported by the high-level syntax.

A. Random Access and Bitstream Splicing Features

The new design supports special features to enable random access and bitstream splicing. In H.264/MPEG-4 AVC, a bitstream must always start with an IDR access unit. An IDR access unit contains an independently coded picture—i.e., a coded picture that can be decoded without decoding any previous pictures in the NAL unit stream. The presence of an IDR access unit indicates that no subsequent picture in the bitstream will require reference to pictures prior to the picture that it contains in order to be decoded. The IDR picture is used within a coding structure known as a closed GOP (in which GOP stands for group of pictures).

The new clean random access (CRA) picture syntax specifies the use of an independently coded picture at the location of a random access point (RAP), i.e., a location in a bitstream at which a decoder can begin successfully decoding pictures without needing to decode any pictures that appeared earlier in the bitstream, which supports an efficient temporal coding

TABLE I
NAL UNIT TYPES, MEANINGS, AND TYPE CLASSES

Type	Meaning	Class
0, 1	Slice segment of ordinary trailing picture	VCL
2, 3	Slice segment of TSA picture	VCL
4, 5	Slice segment of STSA picture	VCL
6, 7	Slice segment of RADL picture	VCL
8, 9	Slice segment of RASL picture	VCL
10–15	Reserved for future use	VCL
16–18	Slice segment of BLA picture	VCL
19, 20	Slice segment of IDR picture	VCL
21	Slice segment of CRA picture	VCL
22–31	Reserved for future use	VCL
32	Video parameter set (VPS)	non-VCL
33	Sequence parameter set (SPS)	non-VCL
34	Picture parameter set (PPS)	non-VCL
35	Access unit delimiter	non-VCL
36	End of sequence	non-VCL
37	End of bitstream	non-VCL
38	Filler data	non-VCL
39, 40	SEI messages	non-VCL
41–47	Reserved for future use	non-VCL
48–63	Unspecified (available for system use)	non-VCL

order known as open GOP operation. Good support of random access is critical for enabling channel switching, seek operations, and dynamic streaming services. Some pictures that follow a CRA picture in decoding order and precede it in display order may contain interpicture prediction references to pictures that are not available at the decoder. These nondecodable pictures must therefore be discarded by a decoder that starts its decoding process at a CRA point. For this purpose, such nondecodable pictures are identified as random access skipped leading (RASL) pictures. The location of splice points from different original coded bitstreams can be indicated by broken link access (BLA) pictures. A bitstream splicing operation can be performed by simply changing the NAL unit type of a CRA picture in one bitstream to the value that indicates a BLA picture and concatenating the new bitstream at the position of a RAP picture in the other bitstream. A RAP picture may be an IDR, CRA, or BLA picture, and both CRA and BLA pictures may be followed by RASL pictures in the bitstream (depending on the particular value of the NAL unit type used for a BLA picture). Any RASL pictures associated with a BLA picture must always be discarded by the decoder, as they may contain references to pictures that are not actually present in the bitstream due to a splicing operation. The other type of picture that can follow a RAP picture in decoding order and precede it in output order is the random access decodable leading (RADL) picture, which cannot contain references to any pictures that precede the RAP picture in decoding order. RASL and RADL pictures are collectively referred to as leading pictures (LPs). Pictures that follow a RAP picture in both decoding order and output order, which are known as trailing pictures, cannot contain references to LPs for interpicture prediction.

B. Temporal Sublayering Support

Similar to the temporal scalability feature in the H.264/MPEG-4 AVC scalable video coding (SVC) extension [12],

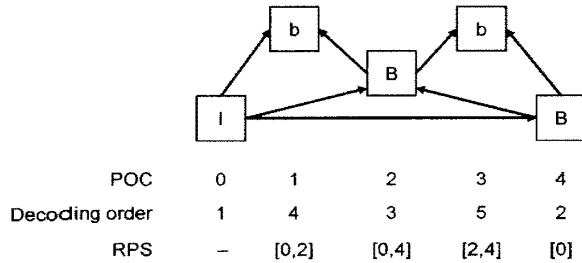


Fig. 2. Example of a temporal prediction structure and the POC values, decoding order, and RPS content for each picture.

HEVC specifies a temporal identifier in the NAL unit header, which indicates a level in a hierarchical temporal prediction structure. This was introduced to achieve temporal scalability without the need to parse parts of the bitstream other than the NAL unit header.

Under certain circumstances, the number of decoded temporal sublayers can be adjusted during the decoding process of one coded video sequence. The location of a point in the bitstream at which sublayer switching is possible to begin decoding some higher temporal layers can be indicated by the presence of temporal sublayer access (TSA) pictures and step-wise TSA (STSA) pictures. At the location of a TSA picture, it is possible to switch from decoding a lower temporal sublayer to decoding any higher temporal sublayer, and at the location of an STSA picture, it is possible to switch from decoding a lower temporal sublayer to decoding only one particular higher temporal sublayer (but not the further layers above that, unless they also contain STSA or TSA pictures).

C. Additional Parameter Sets

The VPS has been added as metadata to describe the overall characteristics of coded video sequences, including the dependences between temporal sublayers. The primary purpose of this is to enable the compatible extensibility of the standard in terms of signaling at the systems layer, e.g., when the base layer of a future extended scalable or multiview bitstream would need to be decodable by a legacy decoder, but for which additional information about the bitstream structure that is only relevant for the advanced decoder would be ignored.

D. Reference Picture Sets and Reference Picture Lists

For multiple-reference picture management, a particular set of previously decoded pictures needs to be present in the decoded picture buffer (DPB) for the decoding of the remainder of the pictures in the bitstream. To identify these pictures, a list of picture order count (POC) identifiers is transmitted in each slice header. The set of retained reference pictures is called the reference picture set (RPS). Fig. 2 shows POC values, decoding order, and RPSs for an example temporal prediction structure.

As in H.264/MPEG-4 AVC, there are two lists that are constructed as lists of pictures in the DPB, and these are called

reference picture list 0 and list 1. An index called a reference picture index is used to identify a particular picture in one of these lists. For uniprediction, a picture can be selected from either of these lists. For biprediction, two pictures are selected—one from each list. When a list contains only one picture, the reference picture index implicitly has the value 0 and does not need to be transmitted in the bitstream.

The high-level syntax for identifying the RPS and establishing the reference picture lists for interpicture prediction is more robust to data losses than in the prior H.264/MPEG-4 AVC design, and is more amenable to such operations as random access and trick mode operation (e.g., fast-forward, smooth rewind, seeking, and adaptive bitstream switching). A key aspect of this improvement is that the syntax is more explicit, rather than depending on inferences from the stored internal state of the decoding process as it decodes the bitstream picture by picture. Moreover, the associated syntax for these aspects of the design is actually simpler than it had been for H.264/MPEG-4 AVC.

IV. HEVC VIDEO CODING TECHNIQUES

As in all prior ITU-T and ISO/IEC JTC 1 video coding standards since H.261 [2], the HEVC design follows the classic block-based hybrid video coding approach (as depicted in Fig. 1). The basic source-coding algorithm is a hybrid of interpicture prediction to exploit temporal statistical dependences, intrapicture prediction to exploit spatial statistical dependences, and transform coding of the prediction residual signals to further exploit spatial statistical dependences. There is no single coding element in the HEVC design that provides the majority of its significant improvement in compression efficiency in relation to prior video coding standards. It is, rather, a plurality of smaller improvements that add up to the significant gain.

A. Sampled Representation of Pictures

For representing color video signals, HEVC typically uses a tristimulus YCbCr color space with 4:2:0 sampling (although extension to other sampling formats is straightforward, and is planned to be defined in a subsequent version). This separates a color representation into three components called Y, Cb, and Cr. The Y component is also called luma, and represents brightness. The two chroma components Cb and Cr represent the extent to which the color deviates from gray toward blue and red, respectively. Because the human visual system is more sensitive to luma than chroma, the 4:2:0 sampling structure is typically used, in which each chroma component has one fourth of the number of samples of the luma component (half the number of samples in both the horizontal and vertical dimensions). Each sample for each component is typically represented with 8 or 10 b of precision, and the 8-b case is the more typical one. In the remainder of this paper, we focus our attention on the typical use: YCbCr components with 4:2:0 sampling and 8 b per sample for the representation of the encoded input and decoded output video signal.

The video pictures are typically progressively sampled with rectangular picture sizes $W \times H$, where W is the width and

H is the height of the picture in terms of luma samples. Each chroma component array, with 4:2:0 sampling, is then $W/2 \times H/2$. Given such a video signal, the HEVC syntax partitions the pictures further as described follows.

B. Division of the Picture into Coding Tree Units

A picture is partitioned into coding tree units (CTUs), which each contain luma CTBs and chroma CTBs. A luma CTB covers a rectangular picture area of $L \times L$ samples of the luma component and the corresponding chroma CTBs cover each $L/2 \times L/2$ samples of each of the two chroma components. The value of L may be equal to 16, 32, or 64 as determined by an encoded syntax element specified in the SPS. Compared with the traditional macroblock using a fixed array size of 16×16 luma samples, as used by all previous ITU-T and ISO/IEC JTC 1 video coding standards since H.261 (that was standardized in 1990), HEVC supports variable-size CTBs selected according to needs of encoders in terms of memory and computational requirements. The support of larger CTBs than in previous standards is particularly beneficial when encoding high-resolution video content. The luma CTB and the two chroma CTBs together with the associated syntax form a CTU. The CTU is the basic processing unit used in the standard to specify the decoding process.

C. Division of the CTB into CBs

The blocks specified as luma and chroma CTBs can be directly used as CBs or can be further partitioned into multiple CBs. Partitioning is achieved using tree structures. The tree partitioning in HEVC is generally applied simultaneously to both luma and chroma, although exceptions apply when certain minimum sizes are reached for chroma.

The CTU contains a quadtree syntax that allows for splitting the CBs to a selected appropriate size based on the signal characteristics of the region that is covered by the CTB. The quadtree splitting process can be iterated until the size for a luma CB reaches a minimum allowed luma CB size that is selected by the encoder using syntax in the SPS and is always 8×8 or larger (in units of luma samples).

The boundaries of the picture are defined in units of the minimum allowed luma CB size. As a result, at the right and bottom edges of the picture, some CTUs may cover regions that are partly outside the boundaries of the picture. This condition is detected by the decoder, and the CTU quadtree is implicitly split as necessary to reduce the CB size to the point where the entire CB will fit into the picture.

D. PBs and PUs

The prediction mode for the CU is signaled as being intra or inter, according to whether it uses intrapicture (spatial) prediction or interpicture (temporal) prediction.

When the prediction mode is signaled as intra, the PB size, which is the block size at which the intrapicture prediction mode is established is the same as the CB size for all block sizes except for the smallest CB size that is allowed in the bitstream. For the latter case, a flag is present that indicates whether the CB is split into four PB quadrants that each

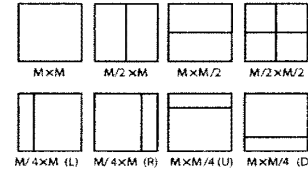


Fig. 3. Modes for splitting a CB into PBs, subject to certain size constraints. For intrapicture-predicted CBs, only $M \times M$ and $M/2 \times M/2$ are supported.

have their own intrapicture prediction mode. The reason for allowing this split is to enable distinct intrapicture prediction mode selections for blocks as small as 4×4 in size. When the luma intrapicture prediction operates with 4×4 blocks, the chroma intrapicture prediction also uses 4×4 blocks (each covering the same picture region as four 4×4 luma blocks). The actual region size at which the intrapicture prediction operates (which is distinct from the PB size, at which the intrapicture prediction mode is established) depends on the residual coding partitioning that is described as follows.

When the prediction mode is signaled as inter, it is specified whether the luma and chroma CBs are split into one, two, or four PBs. The splitting into four PBs is allowed only when the CB size is equal to the minimum allowed CB size, using an equivalent type of splitting as could otherwise be performed at the CB level of the design rather than at the PB level. When a CB is split into four PBs, each PB covers a quadrant of the CB. When a CB is split into two PBs, six types of this splitting are possible. The partitioning possibilities for interpicture-predicted CBs are depicted in Fig. 3. The upper partitions illustrate the cases of not splitting the CB of size $M \times M$, of splitting the CB into two PBs of size $M \times M/2$ or $M/2 \times M$, or splitting it into four PBs of size $M/2 \times M/2$. The lower four partition types in Fig. 3 are referred to as asymmetric motion partitioning (AMP), and are only allowed when M is 16 or larger for luma. One PB of the asymmetric partition has the height or width $M/4$ and width or height M , respectively, and the other PB fills the rest of the CB by having a height or width of $3M/4$ and width or height M . Each interpicture-predicted PB is assigned one or two motion vectors and reference picture indices. To minimize worst-case memory bandwidth, PBs of luma size 4×4 are not allowed for interpicture prediction, and PBs of luma sizes 4×8 and 8×4 are restricted to unipredictive coding. The interpicture prediction process is further described as follows.

The luma and chroma PBs, together with the associated prediction syntax, form the PU.

E. Tree-Structured Partitioning Into Transform Blocks and Units

For residual coding, a CB can be recursively partitioned into transform blocks (TBs). The partitioning is signaled by a residual quadtree.

Only square CB and TB partitioning is specified, where a block can be recursively split into quadrants, as illustrated in Fig. 4. For a given luma CB of size $M \times M$, a flag signals whether it is split into four blocks of size $M/2 \times M/2$. If

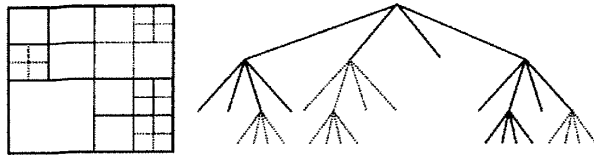


Fig. 4. Subdivision of a CTB into CBs [and transform block (TBs)]. Solid lines indicate CB boundaries and dotted lines indicate TB boundaries. (a) CTB with its partitioning. (b) Corresponding quadtree.

further splitting is possible, as signaled by a maximum depth of the residual quadtree indicated in the SPS, each quadrant is assigned a flag that indicates whether it is split into four quadrants. The leaf node blocks resulting from the residual quadtree are the transform blocks that are further processed by transform coding. The encoder indicates the maximum and minimum luma TB sizes that it will use. Splitting is implicit when the CB size is larger than the maximum TB size. Not splitting is implicit when splitting would result in a luma TB size smaller than the indicated minimum. The chroma TB size is half the luma TB size in each dimension, except when the luma TB size is 4×4 , in which case a single 4×4 chroma TB is used for the region covered by four 4×4 luma TBs. In the case of intrapicture-predicted CUs, the decoded samples of the nearest-neighboring TBs (within or outside the CB) are used as reference data for intrapicture prediction.

In contrast to previous standards, the HEVC design allows a TB to span across multiple PBs for interpicture-predicted CUs to maximize the potential coding efficiency benefits of the quadtree-structured TB partitioning.

F. Slices and Tiles

Slices are a sequence of CTUs that are processed in the order of a raster scan. A picture may be split into one or several slices as shown in Fig. 5(a) so that a picture is a collection of one or more slices. Slices are self-contained in the sense that, given the availability of the active sequence and picture parameter sets, their syntax elements can be parsed from the bitstream and the values of the samples in the area of the picture that the slice represents can be correctly decoded (except with regard to the effects of in-loop filtering near the edges of the slice) without the use of any data from other slices in the same picture. This means that prediction within the picture (e.g., intrapicture spatial signal prediction or prediction of motion vectors) is not performed across slice boundaries. Some information from other slices may, however, be needed to apply the in-loop filtering across slice boundaries. Each slice can be coded using different coding types as follows.

- 1) *I slice*: A slice in which all CUs of the slice are coded using only intrapicture prediction.
- 2) *P slice*: In addition to the coding types of an I slice, some CUs of a P slice can also be coded using interpicture prediction with at most one motion-compensated prediction signal per PB (i.e., uniprediction). P slices only use reference picture list 0.
- 3) *B slice*: In addition to the coding types available in a P slice, some CUs of the B slice can also be coded

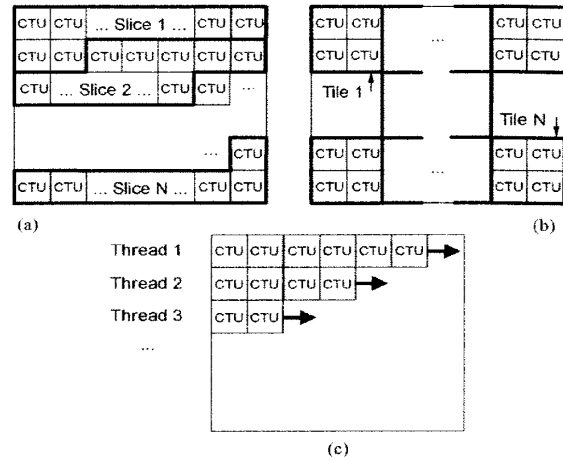


Fig. 5. Subdivision of a picture into (a) slices and (b) tiles. (c) Illustration of wavefront parallel processing.

using interpicture prediction with at most two motion-compensated prediction signals per PB (i.e., biprediction). B slices use both reference picture list 0 and list 1.

The main purpose of slices is resynchronization after data losses. Furthermore, slices are often restricted to use a maximum number of bits, e.g., for packetized transmission. Therefore, slices may often contain a highly varying number of CTUs per slice in a manner dependent on the activity in the video scene. In addition to slices, HEVC also defines tiles, which are self-contained and independently decodable rectangular regions of the picture. The main purpose of tiles is to enable the use of parallel processing architectures for encoding and decoding. Multiple tiles may share header information by being contained in the same slice. Alternatively, a single tile may contain multiple slices. A tile consists of a rectangular arranged group of CTUs (typically, but not necessarily, with all of them containing about the same number of CTUs), as shown in Fig. 5(b).

To assist with the granularity of data packetization, dependent slices are additionally defined. Finally, with WPP, a slice is divided into rows of CTUs. The decoding of each row can be begun as soon as a few decisions that are needed for prediction and adaptation of the entropy coder have been made in the preceding row. This supports parallel processing of rows of CTUs by using several processing threads in the encoder or decoder (or both). An example is shown in Fig. 5(c). For design simplicity, WPP is not allowed to be used in combination with tiles (although these features could, in principle, work properly together).

G. Intrapicture Prediction

Intrapicture prediction operates according to the TB size, and previously decoded boundary samples from spatially neighboring TBs are used to form the prediction signal. Directional prediction with 33 different directional orientations is defined for (square) TB sizes from 4×4 up to 32×32 . The

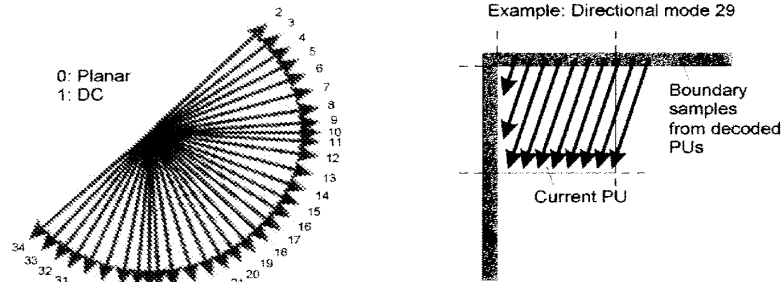


Fig. 6. Modes and directional orientations for intrapicture prediction.

possible prediction directions are shown in Fig. 6. Alternatively, planar prediction (assuming an amplitude surface with a horizontal and vertical slope derived from the boundaries) and DC prediction (a flat surface with a value matching the mean value of the boundary samples) can also be used. For chroma, the horizontal, vertical, planar, and DC prediction modes can be explicitly signaled, or the chroma prediction mode can be indicated to be the same as the luma prediction mode (and, as a special case to avoid redundant signaling, when one of the first four choices is indicated and is the same as the luma prediction mode, the Intra_Angular[34] mode is applied instead).

Each CB can be coded by one of several coding types, depending on the slice type. Similar to H.264/MPEG-4 AVC, intrapicture predictive coding is supported in all slice types. HEVC supports various intrapicture predictive coding methods referred to as Intra_Angular, Intra_Planar, and Intra_DC. The following subsections present a brief further explanation of these and several techniques to be applied in common.

1) *PB Partitioning*: An intrapicture-predicted CB of size $M \times M$ may have one of two types of PB partitions referred to as PART_2N×2N and PART_N×N, the first of which indicates that the CB is not split and the second indicates that the CB is split into four equal-sized PBs. (Conceptually, in this notation, $N = M/2$.) However, it is possible to represent the same regions that would be specified by four PBs by using four smaller CBs when the size of the current CB is larger than the minimum CU size. Thus, the HEVC design only allows the partitioning type PART_N×N to be used when the current CB size is equal to the minimum CU size. This means that the PB size is always equal to the CB size when the CB is coded using an intrapicture prediction mode and the CB size is not equal to the minimum CU size. Although the intrapicture prediction mode is established at the PB level, the actual prediction process operates separately for each TB.

2) *Intra_Angular Prediction*: Spatial-domain intrapicture prediction has previously been successfully used in H.264/MPEG-4 AVC. The intrapicture prediction of HEVC similarly operates in the spatial domain, but is extended significantly—mainly due to the increased size of the TB and an increased number of selectable prediction directions. Compared to the eight prediction directions of H.264/MPEG-4 AVC, HEVC supports a total of 33 prediction directions, denoted as Intra_Angular[k], where k is a mode number from

2 to 34. The angles are intentionally designed to provide denser coverage for near-horizontal and near-vertical angles and coarser coverage for near-diagonal angles to reflect the observed statistical prevalence of the angles and the effectiveness of the signal prediction processing.

When using an Intra_Angular mode, each TB is predicted directionally from spatially neighboring samples that are reconstructed (but not yet filtered by the in-loop filters) before being used for this prediction. For a TB of size $N \times N$, a total of $4N+1$ spatially neighboring samples may be used for the prediction, as shown in Fig. 6. When available from preceding decoding operations, samples from lower left TBs can be used for prediction in HEVC in addition to samples from TBs at the left, above, and above right of the current TB.

The prediction process of the Intra_Angular modes can involve extrapolating samples from the projected reference sample location according to a given directionality. To remove the need for sample-by-sample switching between reference row and column buffers, for Intra_Angular[k] with k in the range of 2–17, the samples located in the above row are projected as additional samples located in the left column; and with k in the range of 18–34, the samples located at the left column are projected as samples located in the above row.

To improve the intrapicture prediction accuracy, the projected reference sample location is computed with $1/32$ sample accuracy. Bilinear interpolation is used to obtain the value of the projected reference sample using two closest reference samples located at integer positions.

The prediction process of the Intra_Angular modes is consistent across all block sizes and prediction directions, whereas H.264/MPEG-4 AVC uses different methods for its supported block sizes of 4×4 , 8×8 , and 16×16 . This design consistency is especially desirable since HEVC supports a greater variety of TB sizes and a significantly increased number of prediction directions compared to H.264/MPEG-4 AVC.

3) *Intra_Planar and Intra_DC Prediction*: In addition to Intra_Angular prediction that targets regions with strong directional edges, HEVC supports two alternative prediction methods, Intra_Planar and Intra_DC, for which similar modes were specified in H.264/MPEG-4 AVC. While Intra_DC prediction uses an average value of reference samples for the prediction, average values of two linear predictions using four corner reference samples are used in Intra_Planar prediction

to prevent discontinuities along the block boundaries. The Intra_Planar prediction mode is supported at all block sizes in HEVC, while H.264/MPEG-4 AVC supports plane prediction only when the luma PB size is 16×16 , and its plane prediction operates somewhat differently from the planar prediction in HEVC.

4) *Reference Sample Smoothing*: In HEVC, the reference samples used for the intrapicture prediction are sometimes filtered by a three-tap $[1 \ 2 \ 1]/4$ smoothing filter in a manner similar to what was used for 8×8 intrapicture prediction in H.264/MPEG-4 AVC. HEVC applies smoothing operations more adaptively, according to the directionality, the amount of detected discontinuity, and the block size. As in H.264/MPEG-4 AVC, the smoothing filter is not applied for 4×4 blocks. For 8×8 blocks, only the diagonal directions, Intra_Angular[k] with $k = 2, 18$, or 34 , use the reference sample smoothing. For 16×16 blocks, the reference samples are filtered for most directions except the near-horizontal and near-vertical directions, k in the range of 9 – 11 and 25 – 27 . For 32×32 blocks, all directions except the exactly horizontal ($k = 10$) and exactly vertical ($k = 26$) directions use the smoothing filter, and when the amount of detected discontinuity exceeds a threshold, bilinear interpolation from three neighboring region samples is applied to form a smooth prediction.

The Intra_Planar mode also uses the smoothing filter when the block size is greater than or equal to 8×8 , and the smoothing is not used (or useful) for the Intra_DC case.

5) *Boundary Value Smoothing*: To remove discontinuities along block boundaries, in three modes, Intra_DC (mode 1) and Intra_Angular[k] with $k = 10$ or 26 (exactly horizontal or exactly vertical), the boundary samples inside the TB are replaced by filtered values when the TB size is smaller than 32×32 . For Intra_DC mode, both the first row and column of samples in the TB are replaced by the output of a two-tap $[3 \ 1]/4$ filter fed by their original value and the adjacent reference sample. In horizontal (Intra_Angular[10]) prediction, the boundary samples of the first column of the TB are modified such that half of the difference between their neighbored reference sample and the top-left reference sample is added. This makes the prediction signal more smooth when large variations in the vertical direction are present. In vertical (Intra_Angular[26]) prediction, the same is applied to the first row of samples.

6) *Reference Sample Substitution*: The neighboring reference samples are not available at the slice or tile boundaries. In addition, when a loss-resilience feature known as constrained intra prediction is enabled, the neighboring reference samples inside any interpicture-predicted PB are also considered not available in order to avoid letting potentially corrupted prior decoded picture data propagate errors into the prediction signal. While only Intra_DC prediction mode is allowed for such cases in H.264/MPEG-4 AVC, HEVC allows the use of other intrapicture prediction modes after substituting the nonavailable reference sample values with the neighboring available reference sample values.

7) *Mode Coding*: HEVC supports a total of 33 Intra_Angular prediction modes and Intra_Planar and Intra_DC prediction modes for luma prediction for all block

sizes. Due to the increased number of directions, HEVC considers three most probable modes (MPMs) when coding the luma intrapicture prediction mode predictively, rather than the one most probable mode considered in H.264/MPEG-4 AVC.

Among the three most probable modes, the first two are initialized by the luma intrapicture prediction modes of the above and left PBs if those PBs are available and are coded using an intrapicture prediction mode. Any unavailable prediction mode is considered to be Intra_DC. The PB above the luma CTB is always considered to be unavailable in order to avoid the need to store a line buffer of neighboring luma prediction modes.

When the first two most probable modes are not equal, the third most probable mode is set equal to Intra_Planar, Intra_DC, or Intra_Angular[26] (vertical), according to which of these modes, in this order, is not a duplicate of one of the first two modes. When the first two most probable modes are the same, if this first mode has the value Intra_Planar or Intra_DC, the second and third most probable modes are assigned as Intra_Planar, Intra_DC, or Intra_Angular[26], according to which of these modes, in this order, are not duplicates. When the first two most probable modes are the same and the first mode has an Intra_Angular value, the second and third most probable modes are chosen as the two angular prediction modes that are closest to the angle (i.e., the value of k) of the first.

In the case that the current luma prediction mode is one of three MPMs, only the MPM index is transmitted to the decoder. Otherwise, the index of the current luma prediction mode excluding the three MPMs is transmitted to the decoder by using a 5-b fixed length code.

For chroma intrapicture prediction, HEVC allows the encoder to select one of five modes: Intra_Planar, Intra_Angular[26] (vertical), Intra_Angular[10] (horizontal), Intra_DC, and Intra_Derived. The Intra_Derived mode specifies that the chroma prediction uses the same angular direction as the luma prediction. With this scheme, all angular modes specified for luma in HEVC can, in principle, also be used in the chroma prediction, and a good tradeoff is achieved between prediction accuracy and the signaling overhead. The selected chroma prediction mode is coded directly (without using an MPM prediction mechanism).

H. Interpicture Prediction

1) *PB Partitioning*: Compared to intrapicture-predicted CBs, HEVC supports more PB partition shapes for interpicture-predicted CBs. The partitioning modes of PART_2N $\times$ 2N, PART_2N $\times$ N, and PART_N $\times$ 2N indicate the cases when the CB is not split, split into two equal-size PBs horizontally, and split into two equal-size PBs vertically, respectively. PART_N $\times$ N specifies that the CB is split into four equal-size PBs, but this mode is only supported when the CB size is equal to the smallest allowed CB size. In addition, there are four partitioning types that support splitting the CB into two PBs having different sizes: PART_2N $\times$ nU, PART_2N $\times$ nD, PART_nL $\times$ 2N, and PART_nR $\times$ 2N. These types are known as asymmetric motion partitions.

2) *Fractional Sample Interpolation*: The samples of the PB for an intrapicture-predicted CB are obtained from those of a

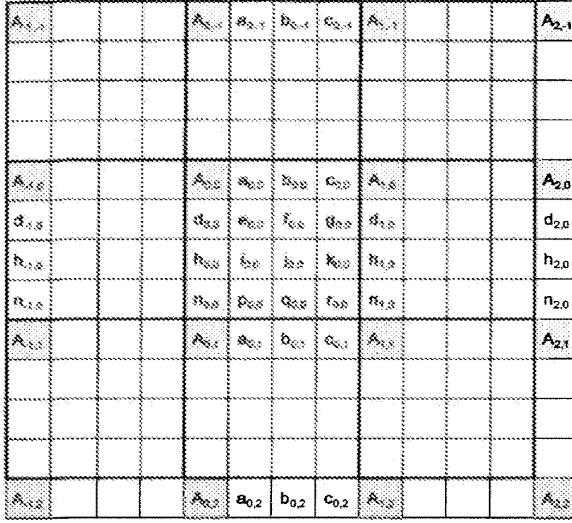


Fig. 7. Integer and fractional sample positions for luma interpolation.

corresponding block region in the reference picture identified by a reference picture index, which is at a position displaced by the horizontal and vertical components of the motion vector. Except for the case when the motion vector has an integer value, fractional sample interpolation is used to generate the prediction samples for noninteger sampling positions. As in H.264/MPEG-4 AVC, HEVC supports motion vectors with units of one quarter of the distance between luma samples. For chroma samples, the motion vector accuracy is determined according to the chroma sampling format, which for 4:2:0 sampling results in units of one eighth of the distance between chroma samples.

The fractional sample interpolation for luma samples in HEVC uses separable application of an eight-tap filter for the half-sample positions and a seven-tap filter for the quarter-sample positions. This is in contrast to the process used in H.264/MPEG-4 AVC, which applies a two-stage interpolation process by first generating the values of one or two neighboring samples at half-sample positions using six-tap filtering, rounding the intermediate results, and then averaging two values at integer or half-sample positions. HEVC instead uses a single consistent separable interpolation process for generating all fractional positions without intermediate rounding operations, which improves precision and simplifies the architecture of the fractional sample interpolation. The interpolation precision is also improved in HEVC by using longer filters, i.e., seven-tap or eight-tap filtering rather than the six-tap filtering used in H.264/MPEG-4 AVC. Using only seven taps rather than the eight used for half-sample positions was sufficient for the quarter-sample interpolation positions since the quarter-sample positions are relatively close to integer-sample positions, so the most distant sample in an eight-tap interpolator would effectively be farther away than in the half-sample case (where the relative distances of the integer-sample positions are symmetric). The actual filter tap values of the

TABLE II
FILTER COEFFICIENTS FOR LUMA FRACTIONAL SAMPLE INTERPOLATION

Index i	-3	-2	-1	0	1	2	3	4
$hfilter[i]$	-1	4	-11	40	40	-11	4	1
$qfilter[i]$	-1	4	-10	58	17	-5	1	

interpolation filtering kernel were partially derived from DCT basis function equations.

In Fig. 7, the positions labeled with upper-case letters, $A_{i,j}$, represent the available luma samples at integer sample locations, whereas the other positions labeled with lower-case letters represent samples at noninteger sample locations, which need to be generated by interpolation.

The samples labeled $a_{0,j}$, $b_{0,j}$, $c_{0,j}$, $d_{0,0}$, $h_{0,0}$, and $n_{0,0}$ are derived from the samples $A_{i,j}$ by applying the eight-tap filter for half-sample positions and the seven-tap filter for the quarter-sample positions as follows:

$$\begin{aligned}
 a_{0,j} &= (\sum_{i=-3,3} A_{i,j} qfilter[i]) >> (B-8) \\
 b_{0,j} &= (\sum_{i=-3,4} A_{i,j} hfilter[i]) >> (B-8) \\
 c_{0,j} &= (\sum_{i=-2,4} A_{i,j} qfilter[1-i]) >> (B-8) \\
 d_{0,0} &= (\sum_{i=-3,3} A_{0,j} qfilter[j]) >> (B-8) \\
 h_{0,0} &= (\sum_{i=-3,4} A_{0,j} hfilter[j]) >> (B-8) \\
 n_{0,0} &= (\sum_{j=-2,4} A_{0,j} qfilter[1-j]) >> (B-8)
 \end{aligned}$$

where the constant $B \geq 8$ is the bit depth of the reference samples (and typically $B = 8$ for most applications) and the filter coefficient values are given in Table II. In these formulas, $>>$ denotes an arithmetic right shift operation.

The samples labeled $e_{0,0}$, $f_{0,0}$, $g_{0,0}$, $i_{0,0}$, $j_{0,0}$, $k_{0,0}$, $p_{0,0}$, $q_{0,0}$, and $r_{0,0}$ can be derived by applying the corresponding filters to samples located at vertically adjacent $a_{0,j}$, $b_{0,j}$ and $c_{0,j}$ positions as follows:

$$\begin{aligned}
 e_{0,0} &= (\sum_{v=-3,3} a_{0,v} qfilter[v]) >> 6 \\
 f_{0,0} &= (\sum_{v=-3,3} b_{0,v} qfilter[v]) >> 6 \\
 g_{0,0} &= (\sum_{v=-3,3} c_{0,v} qfilter[v]) >> 6 \\
 i_{0,0} &= (\sum_{v=-3,4} a_{0,v} hfilter[v]) >> 6 \\
 j_{0,0} &= (\sum_{v=-3,4} b_{0,v} hfilter[v]) >> 6 \\
 k_{0,0} &= (\sum_{v=-3,4} c_{0,v} hfilter[v]) >> 6 \\
 p_{0,0} &= (\sum_{v=-2,4} a_{0,v} qfilter[1-v]) >> 6 \\
 q_{0,0} &= (\sum_{v=-2,4} b_{0,v} qfilter[1-v]) >> 6 \\
 r_{0,0} &= (\sum_{v=-2,4} c_{0,v} qfilter[1-v]) >> 6.
 \end{aligned}$$

The interpolation filtering is separable when B is equal to 8, so the same values could be computed in this case by applying the vertical filtering before the horizontal filtering. When implemented appropriately, the motion compensation process of HEVC can be performed using only 16-b storage elements (although care must be taken to do this correctly).

It is at this point in the process that weighted prediction is applied when selected by the encoder. Whereas H.264/MPEG-4 AVC supported both temporally implicit and explicit weighted prediction, in HEVC only explicit weighted prediction is applied, by scaling and offsetting the prediction with values sent explicitly by the encoder. The bit depth of the prediction is then adjusted to the original bit depth of the reference samples. In the case of uniprediction, the interpolated (and possibly weighted) prediction value is rounded,

TABLE III
FILTER COEFFICIENTS FOR CHROMA FRACTIONAL
SAMPLE INTERPOLATION

Index	-1	0	1	2
filter1[i]	-2	58	10	-2
filter2[i]	-4	54	16	-2
filter3[i]	-6	46	28	-4
filter4[i]	-4	36	36	-4

right-shifted, and clipped to have the original bit depth. In the case of biprediction, the interpolated (and possibly weighted) prediction values from two PBs are added first, and then rounded, right-shifted, and clipped.

In H.264/MPEG-4 AVC, up to three stages of rounding operations are required to obtain each prediction sample (for samples located at quarter-sample positions). If biprediction is used, the total number of rounding operations is then seven in the worst case. In HEVC, at most two rounding operations are needed to obtain each sample located at the quarter-sample positions, thus five rounding operations are sufficient in the worst case when biprediction is used. Moreover, in the most common usage, where the bit depth B is 8 b, the total number of rounding operations in the worst case is further reduced to 3. Due to the lower number of rounding operations, the accumulated rounding error is decreased and greater flexibility is enabled in regard to the manner of performing the necessary operations in the decoder.

The fractional sample interpolation process for the chroma components is similar to the one for the luma component, except that the number of filter taps is 4 and the fractional accuracy is 1/8 for the usual 4:2:0 chroma format case. HEVC defines a set of four-tap filters for eighth-sample positions, as given in Table III for the case of 4:2:0 chroma format (where, in H.264/MPEG-4 AVC, only two-tap bilinear filtering was applied).

Filter coefficient values denoted as filter1[i], filter2[i], filter3[i], and filter4[i] with $i = -1, \dots, 2$ are used for interpolating the 1/8th, 2/8th, 3/8th, and 4/8th fractional positions for the chroma samples, respectively. Using symmetry for the 5/8th, 6/8th, and 7/8th fractional positions, the mirrored values of filter3[$1-i$], filter2[$1-i$], and filter1[$1-i$] with $i = -1, \dots, 2$ are used, respectively.

3) *Merge Mode*: Motion information typically consists of the horizontal and vertical motion vector displacement values, one or two reference picture indices, and, in the case of prediction regions in B slices, an identification of which reference picture list is associated with each index. HEVC includes a merge mode to derive the motion information from spatially or temporally neighboring blocks. It is denoted as merge mode since it forms a merged region sharing all motion information.

The merge mode is conceptually similar to the direct and skip modes in H.264/MPEG-4 AVC. However, there are two important differences. First, it transmits index information to select one out of several available candidates, in a manner sometimes referred to as a motion vector competition scheme. It also explicitly identifies the reference picture list and reference picture index, whereas the direct mode assumes that these have some predefined values.

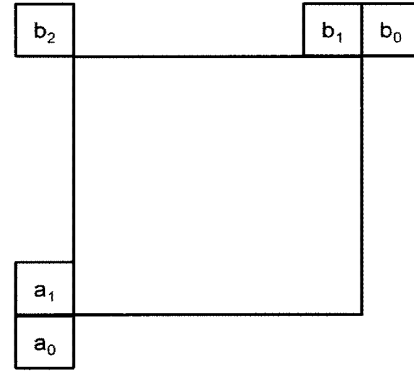


Fig. 8. Positions of spatial candidates of motion information.

The set of possible candidates in the merge mode consists of spatial neighbor candidates, a temporal candidate, and generated candidates. Fig. 8 shows the positions of five spatial candidates. For each candidate position, the availability is checked according to the order $\{a_1, b_1, b_0, a_0, b_2\}$. If the block located at the position is intrapicture predicted or the position is outside of the current slice or tile, it is considered as unavailable.

After validating the spatial candidates, two kinds of redundancy are removed. If the candidate position for the current PU would refer to the first PU within the same CU, the position is excluded, as the same merge could be achieved by a CU without splitting into prediction partitions. Furthermore, any redundant entries where candidates have exactly the same motion information are also excluded.

For the temporal candidate, the right bottom position just outside of the collocated PU of the reference picture is used if it is available. Otherwise, the center position is used instead. The way to choose the collocated PU is similar to that of prior standards, but HEVC allows more flexibility by transmitting an index to specify which reference picture list is used for the collocated reference picture.

One issue related to the use of the temporal candidate is the amount of the memory to store the motion information of the reference picture. This is addressed by restricting the granularity for storing the temporal motion candidates to only the resolution of a 16×16 luma grid, even when smaller PB structures are used at the corresponding location in the reference picture. In addition, a PPS-level flag allows the encoder to disable the use of the temporal candidate, which is useful for applications with error-prone transmission.

The maximum number of merge candidates C is specified in the slice header. If the number of merge candidates found (including the temporal candidate) is larger than C , only the first $C - 1$ spatial candidates and the temporal candidate are retained. Otherwise, if the number of merge candidates identified is less than C , additional candidates are generated until the number is equal to C . This simplifies the parsing and makes it more robust, as the ability to parse the coded data is not dependent on merge candidate availability.

For **B** slices, additional merge candidates are generated by choosing two existing candidates according to a predefined order for reference picture list 0 and list 1. For example, the first generated candidate uses the first merge candidate for list 0 and the second merge candidate for list 1. HEVC specifies a total of 12 predefined pairs of two in the following order in the already constructed merge candidate list as (0, 1), (1, 0), (0, 2), (2, 0), (1, 2), (2, 1), (0, 3), (3, 0), (1, 3), (3, 1), (2, 3), and (3, 2). Among them, up to five candidates can be included after removing redundant entries.

When the slice is a P slice or the number of merge candidates is still less than C , zero motion vectors associated with reference indices from zero to the number of reference pictures minus one are used to fill any remaining entries in the merge candidate list.

In HEVC, the skip mode is treated as a special case of the merge mode when all coded block flags are equal to zero. In this specific case, only a skip flag and the corresponding merge index are transmitted to the decoder. The B-direct mode of H.264/MPEG-4 AVC is also replaced by the merge mode, since the merge mode allows all motion information to be derived from the spatial and temporal motion information of the neighboring blocks with residual coding.

4) *Motion Vector Prediction for Nonmerge Mode*: When an interpicture-predicted CB is not coded in the skip or merge modes, the motion vector is differentially coded using a motion vector predictor. Similar to the merge mode, HEVC allows the encoder to choose the motion vector predictor among multiple predictor candidates. The difference between the predictor and the actual motion vector and the index of the candidate are transmitted to the decoder.

Only two spatial motion candidates are chosen according to the availability among five candidates in Fig. 8. The first spatial motion candidate is chosen from the set of left positions $\{a_0, a_1\}$ and the second one from the set of above positions $\{b_0, b_1, b_2\}$ according to their availabilities, while keeping the searching order as indicated in the two sets.

HEVC only allows a much lower number of candidates to be used in the motion vector prediction process for the nonmerge

case, since the encoder can send a coded difference to change the motion vector. Furthermore, the encoder needs to perform motion estimation, which is one of the most computationally expensive operations in the encoder, and complexity is reduced by allowing a small number of candidates.

When the reference index of the neighboring PU is not equal to that of the current PU, a scaled version of the motion vector is used. The neighboring motion vector is scaled according to the temporal distances between the current picture and the reference pictures indicated by the reference indices of the neighboring PU and the current PU, respectively. When two spatial candidates have the same motion vector components, one redundant spatial candidate is excluded.

When the number of motion vector predictors is not equal to two and the use of temporal MV prediction is not explicitly disabled, the temporal MV prediction candidate is included. This means that the temporal candidate is not used at all when two spatial candidates are available. Finally, a zero motion vector is included repeatedly until the number of motion vector prediction candidates is equal to two, which guarantees that the number of motion vector predictors is two. Thus, only a coded flag is necessary to identify which motion vector prediction is used in the case of nonmerge mode.

1. Transform, Scaling, and Quantization

HEVC uses transform coding of the prediction error residual in a similar manner as in prior standards. The residual block is partitioned into multiple square TBs, as described in Section IV-E. The supported transform block sizes are 4×4 , 8×8 , 16×16 , and 32×32 .

1) *Core Transform*: Two-dimensional transforms are computed by applying 1-D transforms in the horizontal and vertical directions. The elements of the core transform matrices were derived by approximating scaled DCT basis functions, under considerations such as limiting the necessary dynamic range for transform computation and maximizing the precision and closeness to orthogonality when the matrix entries are specified as integer values.

$$H = \begin{bmatrix} 64 & 64 & 64 & 64 & 64 & 64 & 64 & 64 & 64 & 64 & 64 & 64 & 64 & 64 & 64 & 64 \\ 90 & 87 & 80 & 70 & 57 & 43 & 25 & 9 & -9 & -25 & -43 & -57 & -70 & -80 & -87 & 90 \\ 89 & 75 & 50 & 18 & -18 & -50 & -75 & -89 & -89 & -75 & -50 & -18 & 18 & 50 & 75 & 89 \\ 87 & 57 & 9 & -43 & -80 & -90 & -70 & -25 & 25 & 70 & 90 & 80 & 43 & -9 & -57 & -87 \\ 83 & 36 & -36 & -83 & -83 & -36 & 36 & 83 & 83 & 36 & -36 & -83 & -83 & -36 & 36 & 83 \\ 80 & 9 & -70 & -87 & -25 & 57 & 90 & 43 & -43 & -90 & -57 & 25 & 87 & 70 & -9 & -80 \\ 75 & -18 & -89 & -50 & 50 & 89 & 18 & -75 & -75 & 18 & 89 & 50 & -50 & -89 & -18 & 75 \\ 70 & -43 & -87 & 9 & 90 & 25 & -80 & -57 & 57 & 80 & -25 & -90 & -9 & 87 & 43 & -70 \\ 64 & -64 & -64 & 64 & 64 & -64 & -64 & 64 & 64 & -64 & -64 & 64 & 64 & -64 & -64 & 64 \\ 57 & -80 & -25 & 90 & -9 & -87 & 43 & 70 & -70 & -43 & 87 & 9 & -90 & 25 & 80 & -57 \\ 50 & -89 & 18 & 75 & -75 & -18 & 89 & -50 & -50 & 89 & -18 & -75 & 75 & 18 & -89 & 50 \\ 43 & -90 & 57 & 25 & -87 & 70 & 9 & -80 & 80 & -9 & -70 & 87 & -25 & -57 & 90 & -43 \\ 36 & -83 & 83 & -36 & -36 & 83 & -83 & 36 & 36 & -83 & 83 & -36 & -36 & 83 & -83 & 36 \\ 25 & -70 & 90 & -80 & 43 & 9 & -57 & 87 & -87 & 57 & -9 & -43 & 80 & -90 & 70 & -25 \\ 18 & -50 & 75 & -89 & 89 & -75 & 50 & -18 & -18 & 50 & -75 & 89 & -89 & 75 & -50 & 18 \\ 9 & -25 & 43 & -57 & 70 & -80 & 87 & -90 & 90 & -87 & 80 & -70 & 57 & -43 & 25 & -9 \end{bmatrix}$$

For simplicity, only one integer matrix for the length of 32 points is specified, and subsampled versions are used for other sizes. For example, the matrix for the length-16 transform is as shown in the equation at the bottom of the previous page.

The matrices for the length-8 and length-4 transforms can be derived by using the first eight entries of rows 0, 2, 4, ..., and using the first four entries of rows 0, 4, 8, ..., respectively. Although the standard specifies the transform simply in terms of the value of a matrix, the values of the entries in the matrix were selected to have key symmetry properties that enable fast partially factored implementations with far fewer mathematical operations than an ordinary matrix multiplication, and the larger transforms can be constructed by using the smaller transforms as building blocks.

Due to the increased size of the supported transforms, limiting the dynamic range of the intermediate results from the first stage of the transformation is quite important. HEVC explicitly inserts a 7-b right shift and 16-b clipping operation after the first 1-D inverse transform stage of the transform (the vertical inverse transform stage) to ensure that all intermediate values can be stored in 16-b memory (for 8-b video decoding).

2) *Alternative 4 × 4 Transform*: For the transform block size of 4 × 4, an alternative integer transform derived from a DST is applied to the luma residual blocks for intrapicture prediction modes, with the transform matrix

$$H = \begin{bmatrix} 29 & 55 & 74 & 84 \\ 74 & 74 & 0 & -74 \\ 84 & -29 & -74 & 55 \\ 55 & -84 & 74 & -29 \end{bmatrix}.$$

The basis functions of the DST better fit the statistical property that the residual amplitudes tend to increase as the distance from the boundary samples that are used for prediction becomes larger. In terms of complexity, the 4 × 4 DST-style transform is not much more computationally demanding than the 4 × 4 DCT-style transform, and it provides approximately 1% bit-rate reduction in intrapicture predictive coding.

The usage of the DST type of transform is restricted to only 4 × 4 luma transform blocks, since for other cases the additional coding efficiency improvement for including the additional transform type was found to be marginal.

3) *Scaling and Quantization*: Since the rows of the transform matrix are close approximations of values of uniformly scaled basis functions of the orthonormal DCT, the prescaling operation that is incorporated in the dequantization of H.264/MPEG-4 AVC is not needed in HEVC. This avoidance of frequency-specific basis function scaling is useful in reducing the intermediate memory size, especially when considering that the size of the transform can be as large as 32 × 32.

For quantization, HEVC uses essentially the same URQ scheme controlled by a quantization parameter (QP) as in H.264/MPEG-4 AVC. The range of the QP values is defined from 0 to 51, and an increase by 6 doubles the quantization step size such that the mapping of QP values to step sizes is approximately logarithmic. Quantization scaling matrices are also supported.

To reduce the memory needed to store frequency-specific scaling values, only quantization matrices of sizes 4 × 4 and 8 × 8 are used. For the larger transformations of 16 × 16 and 32 × 32 sizes, an 8 × 8 scaling matrix is sent and is applied by sharing values within 2 × 2 and 4 × 4 coefficient groups in frequency subspaces—except for values at DC (zero-frequency) positions, for which distinct values are sent and applied.

J. Entropy Coding

HEVC specifies only one entropy coding method, CABAC [13] rather than two as in H.264/MPEG-4 AVC. The core algorithm of CABAC is unchanged, and the following subsections present several aspects of how it is used in the HEVC design.

1) *Context Modeling*: Appropriate selection of context modeling is known to be a key factor to improve the efficiency of CABAC coding. In HEVC, the splitting depth of the coding tree or transform tree is exploited to derive the context model indices of various syntax elements in addition to the spatially neighboring ones used in H.264/AVC. For example, the syntax element *skip_flag* specifying whether the CB is coded as interpicture predictively skipped and the syntax element *split_coding_unit_flag* specifying whether the CB is further split are coded by using context models based on the spatially neighboring information. The syntax element *split_transform_flag* specifying whether the TB is further split and three syntax elements specifying non-zero transform coefficients for each color component, *cbf_luma*, *cbf_cb* and *cbf_cr*, are coded based on the splitting depth of the transform tree. Although the number of contexts used in HEVC is substantially less than in H.264/MPEG-4 AVC, the entropy coding design actually provides better compression than would a straightforward extension of the H.264/MPEG-4 AVC scheme. Moreover, more extensive use is made in HEVC of the bypass mode of CABAC operation to increase throughput by reducing the amount of data that needs to be coded using CABAC contexts. Dependences between coded data are also carefully considered to enable further throughput maximization.

2) *Adaptive Coefficient Scanning*: Coefficient scanning is performed in 4 × 4 subblocks for all TB sizes (i.e., using only one coefficient region for the 4 × 4 TB size, and using multiple 4 × 4 coefficient regions within larger transform blocks). Three coefficient scanning methods, diagonal up-right, horizontal, and vertical scans as shown in Fig. 9, are selected implicitly for coding the transform coefficients of 4 × 4 and 8 × 8 TB sizes in intrapicture-predicted regions. The selection of the coefficient scanning order depends on the directionalities of the intrapicture prediction. The vertical scan is used when the prediction direction is close to horizontal and the horizontal scan is used when the prediction direction is close to vertical. For other prediction directions, the diagonal up-right scan is used.

For the transform coefficients in interpicture prediction modes of all block sizes and for the transform coefficients of 16 × 16 or 32 × 32 intrapicture prediction, the 4 × 4 diagonal up-right scan is exclusively applied to subblocks of transform coefficients.

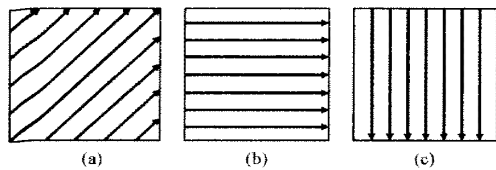


Fig. 9. Three coefficient scanning methods in HEVC. (a) Diagonal up-right scan. (b) Horizontal scan. (c) Vertical scan.

3) *Coefficient Coding*: Similar to H.264/MPEG-4 AVC, HEVC transmits the position of the last nonzero transform coefficient, a significance map, sign bits and levels for the transform coefficients. However, various changes for each part have been made, especially for better handling of the significantly increased size of the TBs.

First, the horizontal and vertical frequency coordinate positions of the last nonzero coefficient are coded for the TB before sending the significance maps of 4×4 subblocks that indicate which other transform coefficients have nonzero values, rather than sending a series of last-coefficient identification flags that are interleaved with the significance map as done in H.264/MPEG-4 AVC.

The significance map is derived for significance groups relating to the fixed size 4×4 subblocks. For all groups having at least one coefficient preceding the last coefficient position, a significant group flag specifying a nonzero coefficient group is transmitted, followed by coefficient significance flags for each coefficient prior to the indicated position of the last significant coefficient. The context models for the significant coefficient flags are dependent on the coefficient position as well as the values of the right and the bottom significant group flags.

A method known as sign data hiding is used for further compression improvement. The sign bits are coded conditionally based on the number and positions of coded coefficients. When sign data hiding is used and there are at least two nonzero coefficients in a 4×4 subblock and the difference between the scan positions of the first and the last nonzero coefficients is greater than 3, the sign bit of the first nonzero coefficient is inferred from the parity of the sum of the coefficient amplitudes. Otherwise, the sign bit is coded normally. At the encoder side, this can be implemented by selecting one coefficient with an amplitude close to the boundary of a quantization interval to be forced to use the adjacent quantization interval in cases where the parity would not otherwise indicate the correct sign of the first coefficient. This allows the sign bit to be encoded at a lower cost (in rate-distortion terms) than if it were coded separately—by giving the encoder the freedom to choose which transform coefficient amplitude can be altered with the lowest rate-distortion cost.

For each position where the corresponding significant coefficient flag is equal to one, two flags specifying whether the level value is greater than one or two are coded, and then the remaining level value is coded depending on those two values.

K. In-Loop Filters

In HEVC, two processing steps, namely a deblocking filter (DBF) followed by an SAO filter, are applied to the

reconstructed samples before writing them into the decoded picture buffer in the decoder loop. The DBF is intended to reduce the blocking artifacts due to block-based coding. The DBF is similar to the DBF of the H.264/MPEG-4 AVC standard, whereas SAO is newly introduced in HEVC. While the DBF is only applied to the samples located at block boundaries, the SAO filter is applied adaptively to all samples satisfying certain conditions, e.g., based on gradient. During the development of HEVC, it had also been considered to operate a third processing step called the adaptive loop filter (ALF) after the SAO filter; however, the ALF feature was not included in the final design.

1) *Deblocking Filter*: The deblocking filter is applied to all samples adjacent to a PU or TU boundary except the case when the boundary is also a picture boundary, or when deblocking is disabled across slice or tile boundaries (which is an option that can be signaled by the encoder). It should be noted that both PU and TU boundaries should be considered since PU boundaries are not always aligned with TU boundaries in some cases of interpicture-predicted CBs. Syntax elements in the SPS and slice headers control whether the deblocking filter is applied across the slice and tile boundaries.

Unlike H.264/MPEG-4 AVC, where the deblocking filter is applied on a 4×4 sample grid basis, HEVC only applies the deblocking filter to the edges that are aligned on an 8×8 sample grid, for both the luma and chroma samples. This restriction reduces the worst-case computational complexity without noticeable degradation of the visual quality. It also improves parallel-processing operation by preventing cascading interactions between nearby filtering operations.

The strength of the deblocking filter is controlled by the values of several syntax elements similar to the scheme in H.264/MPEG-4 AVC, but only three strengths are used rather than five. Given that P and Q are two adjacent blocks with a common 8×8 grid boundary, the filter strength of 2 is assigned when one of the blocks is intrapicture predicted. Otherwise, the filter strength of 1 is assigned if any of the following conditions is satisfied.

- 1) P or Q has at least one nonzero transform coefficient.
- 2) The reference indices of P and Q are not equal.
- 3) The motion vectors of P and Q are not equal.
- 4) The difference between a motion vector component of P and Q is greater than or equal to one integer sample.

If none of the above conditions is met, the filter strength of 0 is assigned, which means that the deblocking process is not applied.

According to the filter strength and the average quantization parameter of P and Q, two thresholds, t_c and β , are determined from predefined tables. For luma samples, one of three cases, no filtering, strong filtering, and weak filtering, is chosen based on β . Note that this decision is shared across four luma rows or columns using the first and the last rows or columns to reduce the computational complexity.

There are only two cases, no filtering and normal filtering, for chroma samples. Normal filtering is applied only when the filter strength is greater than one. The filtering process is then performed using the control variables t_c and β .

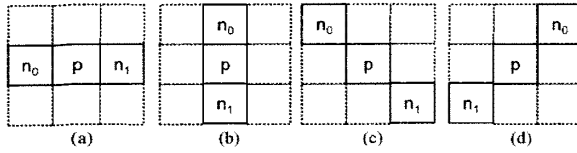


Fig. 10. Four gradient patterns used in SAO. Sample labeled “p” indicates a center sample to be considered. Two samples labeled “ n_0 ” and “ n_1 ” specify two neighboring samples along the (a) horizontal (sao_eo_class = 0), (b) vertical (sao_eo_class = 1), (c) 135° diagonal (sao_eo_class = 2), and (d) 45° (sao_eo_class = 3) gradient patterns.

In HEVC, the processing order of the deblocking filter is defined as horizontal filtering for vertical edges for the entire picture first, followed by vertical filtering for horizontal edges. This specific order enables either multiple horizontal filtering or vertical filtering processes to be applied in parallel threads, or can still be implemented on a CTB-by-CTB basis with only a small processing latency.

2) *SAO*: SAO is a process that modifies the decoded samples by conditionally adding an offset value to each sample after the application of the deblocking filter, based on values in look-up tables transmitted by the encoder. SAO filtering is performed on a region basis, based on a filtering type selected per CTB by a syntax element `sao_type_idx`. A value of 0 for `sao_type_idx` indicates that the SAO filter is not applied to the CTB, and the values 1 and 2 signal the use of the band offset and edge offset filtering types, respectively.

In the band offset mode specified by `sao_type_idx` equal to 1, the selected offset value directly depends on the sample amplitude. In this mode, the full sample amplitude range is uniformly split into 32 segments called bands, and the sample values belonging to four of these bands (which are consecutive within the 32 bands) are modified by adding transmitted values denoted as band offsets, which can be positive or negative. The main reason for using four consecutive bands is that in the smooth areas where banding artifacts can appear, the sample amplitudes in a CTB tend to be concentrated in only few of the bands. In addition, the design choice of using four offsets is unified with the edge offset mode of operation which also uses four offset values.

In the edge offset mode specified by `sao_type_idx` equal to 2, a syntax element `sao_eo_class` with values from 0 to 3 signals whether a horizontal, vertical or one of two diagonal gradient directions is used for the edge offset classification in the CTB. Fig. 10 depicts the four gradient patterns used for the respective `sao_eo_class` in this mode. Each sample in the CTB is classified into one of five `EdgeIdx` categories by comparing the sample value p located at some position with the values n_0 and n_1 of two samples located at neighboring positions as shown in Table IV. This classification is done for each sample based on decoded sample values, so no additional signaling is required for the `EdgeIdx` classification. Depending on the `EdgeIdx` category at the sample position, for `EdgeIdx` categories from 1 to 4, an offset value from a transmitted look-up table is added to the sample value. The offset values are always positive for categories 1 and 2 and negative for categories 3 and 4 - thus the filter generally has a smoothing effect in the edge offset mode.

TABLE IV
SAMPLE `EDGEIDX` CATEGORIES IN SAO `EDGE CLASSES`

EdgeIdx	Condition	Meaning
0	Cases not listed below	Monotonic area
1	$p < n_0$ and $p < n_1$	Local min
2	$p < n_0$ and $p = n_1$ or $p < n_1$ and $p = n_0$	Edge
3	$p > n_0$ and $p = n_1$ or $p > n_1$ and $p = n_0$	Edge
4	$p > n_0$ and $p > n_1$	Local max

Thus, for SAO types 1 and 2, a total of four amplitude offset values are transmitted to the decoder for each CTB. For type 1, the sign is also encoded. The offset values and related syntax elements such as `sao_type_idx` and `sao_eo_class` are determined by the encoder - typically using criteria that optimize rate-distortion performance. The SAO parameters can be indicated to be inherited from the left or above CTB using a merge flag to make the signaling efficient. In summary, SAO is a nonlinear filtering operation which allows additional refinement of the reconstructed signal, and it can enhance the signal representation in both smooth areas and around edges.

L. Special Coding Modes

HEVC defines three special coding modes, which can be invoked at the CU level or the TU level.

- 1) In *L-PCM* mode, the prediction, transform, quantization and entropy coding are bypassed, and the samples are directly represented by a pre-defined number of bits. Its main purpose is to avoid excessive consumption of bits when the signal characteristics are extremely unusual and cannot be properly handled by hybrid coding (e.g., noise-like signals).
- 2) In *lossless* mode, the transform, quantization, and other processing that affects the decoded picture (SAO and deblocking filters) are bypassed, and the residual signal from inter- or intrapicture prediction is directly fed into the entropy coder (using the same neighborhood contexts that would usually be applied to the quantized transform coefficients). This allows mathematically lossless reconstruction, which is achieved without defining any additional coding tools.
- 3) In *transform skipping* mode, only the transform is bypassed. This primarily improves compression for certain types of video content such as computer-generated images or graphics mixed with camera-view content (e.g., scrolling text). This mode can be applied to TBs of 4×4 size only.

SAO and deblocking filtering are not applied to lossless mode regions, and a flag controls whether they are applied to *L-PCM* regions.

V. PROFILES, TIERS, AND LEVELS

A. Profile, Level, and Tier Concepts

Profiles, tiers, and levels specify conformance points for implementing the standard in an interoperable way across various applications that have similar functional requirements. A profile defines a set of coding tools or algorithms that can be used in generating a conforming bitstream, whereas a level places

constraints on certain key parameters of the bitstream, corresponding to decoder processing load and memory capabilities. Level restrictions are established in terms of maximum sample rate, maximum picture size, maximum bit rate, minimum compression ratio and capacities of the DPB, and the coded picture buffer (CPB) that holds compressed data prior to its decoding for data flow management purposes. In the design of HEVC, it was determined that some applications existed that had requirements that differed only in terms of maximum bit rate and CPB capacities. To resolve this issue, two tiers were specified for some levels—a Main Tier for most applications and a High Tier for use in the most demanding applications.

A decoder conforming to a certain tier and level is required to be capable of decoding all bitstreams that conform to the same tier or the lower tier of that level or any level below it. Decoders conforming to a specific profile must support all features in that profile. Encoders are not required to make use of any particular set of features supported in a profile, but are required to produce conforming bitstreams, i.e., bitstreams that obey the specified constraints that enable them to be decoded by conforming decoders.

B. The HEVC Profile and Level Definitions

Only three profiles targeting different application requirements, called the Main, Main 10, and Main Still Picture profiles, are foreseen to be finalized by January 2013. Minimizing the number of profiles provides a maximum amount of interoperability between devices, and is further justified by the fact that traditionally separate services, such as broadcast, mobile, streaming, are converging to the point where most devices should become usable to support all of them. The three drafted profiles consist of the coding tools and high-layer syntax described in the earlier sections of this paper, while imposing the following restrictions.

- 1) Only 4:2:0 chroma sampling is supported.
- 2) When an encoder encodes a picture using multiple tiles, it cannot also use wavefront parallel processing, and each tile must be at least 256 luma samples wide and 64 luma samples tall.
- 3) In the Main and Main Still Picture profiles, only a video precision of 8 b per sample is supported, while the Main 10 profile supports up to 10 b per sample.
- 4) In the Main Still Picture profile, the entire bitstream must contain only one coded picture (and thus interpicture prediction is not supported).

Currently, the definition of 13 levels is planned to be included in the first version of the standard as shown in Table V, ranging from levels that support only relatively small picture sizes such as a luma picture size of 176×144 (sometimes called a quarter common intermediate format) to picture sizes as large as 7680×4320 (often called $8k \times 4k$). The picture width and height are each required to be less than or equal to $\sqrt{8 \cdot \text{MaxLumaPS}}$, where MaxLumaPS is the maximum luma picture size as shown in Table V (to avoid the problems for decoders that could be involved with extreme picture shapes).

There are two tiers supported for eight of these levels (level 4 and higher). The CPB capacity is equal to the

TABLE V
LEVEL LIMITS FOR THE MAIN PROFILE

Level	Max Luma Picture Size (samples)	Max Luma Sample Rate (samples/s)	Main Tier Max Bit Rate (1000 bits/s)	High Tier Max Bit Rate (1000 bits/s)	Min Comp. Ratio
1	36 864	552 960	128	—	2
2	122 880	3 686 400	1500	—	2
2.1	245 760	7 372 800	3000	—	2
3	552 960	16 588 800	6000	—	2
3.1	983 040	33 177 600	10 000	—	2
4	2 228 224	66 846 720	12 000	30 000	4
4.1	2 228 224	133 693 440	20 000	50 000	4
5	8 912 896	267 386 880	25 000	100 000	6
5.1	8 912 896	534 773 760	40 000	160 000	8
5.2	8 912 896	1 069 547 520	60 000	240 000	8
6	35 651 584	1 069 547 520	60 000	240 000	8
6.1	35 651 584	2 139 095 040	120 000	480 000	8
6.2	35 651 584	4 278 190 080	240 000	800 000	6

maximum bit rate times 1 s for all levels except level 1, which has a (higher) CPB capacity of 350 000 b. The specified maximum DPB capacity in each level is six pictures when operating at the maximum picture size supported by the level (including both the current picture and all other pictures that are retained in the decoder at any point in time for reference or output purposes). When operating with a smaller picture size than the maximum size supported by the level, the DPB picture storage capacity can increase to as many as 16 pictures (depending on the particular selected picture size). Level-specific constraints are also specified for the maximum number of tiles used horizontally and vertically within each picture and the maximum number of tiles used per second.

VI. HISTORY AND STANDARDIZATION PROCESS

After the finalization of the H.264/MPEG-4 AVC High Profile in mid-2004, both ITU-T VCEG and ISO/IEC MPEG have been trying to identify when the next major advances in coding efficiency would become ready for standardization. VCEG began studying potential advances in 2004, began identifying certain key technology areas (KTAs) for study in early 2005, and developed a common KTA software codebase for this paper [14]. Various technologies were proposed and verified using the KTA software codebase, which was developed from the H.264/MPEG-4 AVC reference software known as the joint model (JM).

From 2005 to 2008, MPEG began exploration activities toward significant coding efficiency improvements as well, organized several workshops, and issued a “call for evidence” [15] of such advances in April 2009. Expert viewing tests were conducted to evaluate submissions of responses to the call.

From their respective investigations, it was agreed that there were sufficient technologies with the potential to improve the coding efficiency significantly, compared to the existing video coding standards. The Joint Collaborative Team on Video Coding (JCT-VC) was planned to be established by both groups in January 2010, and a joint call for proposals (CfP) on video compression technology [16] was issued by the same

TABLE VI
STRUCTURE OF CODING TOOLS ASSOCIATED WITH HIGH EFFICIENCY
AND LOW COMPLEXITY CONFIGURATIONS OF HM 1

Functionality	High Efficiency	Low Complexity
CTB/CU structure	Tree structure from 8×8 to 64×64	
PU structure	Square and symmetric shapes	
TU structure	Three-level tree structure	Two-level tree structure
Transform	Integer transforms from four to 32 points	
Intra prediction	33 angular modes with DC mode	
Luma interpolation	12-tap separable	Six-tap directional
Chroma interpolation	Bilinear	
MV prediction	Spatial CU merging/AMVP	
Entropy coding	CABAC	Event-based VLC
Deblocking filter	Enabled	
Adaptive loop filter	Enabled	Disabled

time to identify the initial technologies that would serve as a basis of future standardization activities.

At its first meeting in April 2010, the JCT-VC established the HEVC project name, studied the proposals submitted in response to the CFP, and established the first version of a test model under consideration (TMuC) [17], which was produced collectively from elements of several promising proposals. A corresponding software codebase was implemented after this meeting. The technology submitted in several of the key proposal contributions was previously discussed in a special section of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY [18].

Although the TMuC showed significant coding efficiency improvements compared to prior standards, it had several redundant coding tools in each functional block of the video compression system, primarily due to the fact that the TMuC was a collective design from various contributions. During the second JCT-VC meeting in July 2010, the process began of selecting the minimal necessary set of coding tools for each functional block by thoroughly testing each component of the TMuC.

Based on the reported results of the exhaustive component testing [20], an HEVC test model version 1 (HM 1) [21] and the corresponding HEVC working draft specification version 1 (WD 1) [22] were produced as outputs of the third JCT-VC meeting in October 2010. Compared to the prior TMuC design, HM 1 was simplified greatly by removing coding tools that showed only marginal benefits relative to their computational complexity.

In several subsequent studies, the coding tools of the HM were classified into two categories called the high efficiency and low complexity configurations. Two corresponding test scenarios for verifying future contributions in the JCT-VC were also established. Table VI summarizes the HM 1 coding tools for the high efficiency and low complexity configurations.

From the fourth to the eleventh JCT-VC meetings, not just coding efficiency improvements, but many other aspects including computational complexity reduction, unification of various coding tools, and parallel friendly design were investigated, and the HEVC design was updated accordingly, until the current status of draft standard, as described in this

TABLE VII
SUMMARY OF CODING TOOLS OF HIGH EFFICIENCY
CONFIGURATION IN HM 1 AND HEVC

Functionality	HM 1 High Efficiency	HEVC (draft 9)
CTU structure	Tree structure from 8×8 to 64×64	Tree structure from 8×8 to 64×64
PU structure	Square and symmetric	Square, symmetric and asymmetric (only square for intra)
TU structure	Tree structure of square TUs	Tree structure of square TUs
Core transform	Integer transforms from 4 to 32 points (full factorization)	Integer transforms from 4 to 32 points (partially factorable)
Alternative transform	n/a	Integer DST type for 4×4
Intra prediction	33 angular modes with DC mode	33 angular modes with planar and DC modes
Luma interpolation	12-tap separable	8-tap/7-tap separable
Chroma interpolation	Bilinear	4-tap separable
MV prediction	AMVP	AMVP
MC region merging	Spatial CU merging	PU merging
Entropy coding	CABAC	CABAC
Deblocking filter	Nonparallel	Parallel
Sample adaptive offset	n/a	Enabled
Adaptive loop filter	Multiple shapes	n/a
Dedicated tools for parallel processing	Slices	Slices, tiles, wavefronts, and dependent slice segments

paper, was reached. In this context, it also turned out that the differentiation for low complexity and high efficiency was no longer necessary because possible to define the unified main profile. Table VII provides a summary of coding tools of the high efficiency configuration in HM 1 and the current specification of HEVC.

At the eighth JCT-VC meeting in February 2012, the draft version 6 of HEVC standard was produced, which was subsequently balloted as the ISO/IEC Committee Draft of the HEVC standard. The tenth JCT-VC Meeting in July 2012 released the draft version 8 for a Draft International Standard ballot, and the finalized text for Consent in ITU-T and Final Draft International Standard in ISO/IEC is expected to be produced in January 2013.

Future extensions of HEVC, which are already being explored and prepared by the JCT-VC's parent bodies, are likely to include extended-range formats with increased bit depth and enhanced color component sampling, scalable coding, and 3-D/stereo/multi-view video coding (the latter including the encoding of depth maps for use with advanced 3-D displays).

VII. CONCLUSION

The emerging HEVC standard has been developed and standardized collaboratively by both the ITU-T VCEG and ISO/IEC MPEG organizations. HEVC represents a number of advances in video coding technology. Its video coding layer design is based on conventional block-based motion-compensated hybrid video coding concepts, but with some important differences relative to prior standards.

When used well together, the features of the new design provide approximately a 50% bit-rate savings for equivalent perceptual quality relative to the performance of prior standards (especially for a high-resolution video). For more details on compression performance, please refer to [23]. Implementation complexity analysis is outside the scope of this paper; however, the decoder implementation complexity of HEVC overall is not a major burden (e.g., relative to H.264/MPEG-4 AVC) using modern processing technology, and encoder complexity is also manageable. For more details on implementation complexity, please refer to [24].

Further information on and documents from the project are available in the JCT-VC document management system (<http://phenix.int-evry.fr/jct/>).

ACKNOWLEDGMENT

The authors would like to thank the experts of ITU-T VCEG, ISO/IEC MPEG, and the ITU-T/ISO/IEC Joint Collaborative Team on Video Coding for their contributions.

REFERENCES

- [1] B. Bross, W.-J. Han, G. J. Sullivan, J.-R. Ohm, and T. Wiegand, *High Efficiency Video Coding (HEVC) Text Specification Draft 9*, document JCTVC-K1003, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC), Oct. 2012.
- [2] *Video Codec for Audiovisual Services at pxb4 kbit/s*, ITU-T Rec. H.261, version 1: Nov. 1990, version 2: Mar. 1993.
- [3] *Video Coding for Low Bit Rate Communication*, ITU-T Rec. H.263, Nov. 1995 (and subsequent editions).
- [4] *Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to About 1.5 Mbit/s—Part 2: Video*, ISO/IEC 11172-2 (MPEG-1), ISO/IEC JTC 1, 1993.
- [5] *Coding of Audio-Visual Objects—Part 2: Visual*, ISO/IEC 14496-2 (MPEG-4 Visual version 1), ISO/IEC JTC 1, Apr. 1999 (and subsequent editions).
- [6] *Generic Coding of Moving Pictures and Associated Audio Information—Part 2: Video*, ITU-T Rec. H.262 and ISO/IEC 13818-2 (MPEG 2 Video), ITU-T and ISO/IEC JTC 1, Nov. 1994.
- [7] *Advanced Video Coding for Generic Audio-Visual Services*, ITU-T Rec. H.264 and ISO/IEC 14496-10 (AVC), ITU-T and ISO/IEC JTC 1, May 2003 (and subsequent editions).
- [8] H. Samet, "The quadtree and related hierarchical data structures," *Comput. Survey*, vol. 16, no. 2, pp. 187–260, Jun. 1984.
- [9] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560–576, Jul. 2003.
- [10] S. Wenger, "H.264/AVC over IP," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 645–656, Jul. 2003.
- [11] T. Stockhammer, M. M. Hannuksela, and T. Wiegand, "H.264/AVC in wireless environments," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 657–673, Jul. 2003.
- [12] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 9, pp. 1103–1120, Sep. 2007.
- [13] D. Marpe, H. Schwarz, and T. Wiegand, "Context-adaptive binary arithmetic coding in the H.264/AVC video compression standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 620–636, Jul. 2003.
- [14] G. J. Sullivan, *Meeting Report for 26th VCEG Meeting*, ITU-T SG16/Q6 document VCEG-Z01, Apr. 2005.
- [15] *Call for Evidence on High-Performance Video Coding (HVC)*, MPEG document N10553, ISO/IEC JTC 1/SC 29/WG 11, Apr. 2009.
- [16] *Joint Call for Proposals on Video Compression Technology*, ITU-T SG16/Q6 document VCEG-AM91 and ISO/IEC MPEG document N11113, ITU-T and ISO/IEC JTC 1, Jan. 2010.
- [17] *Test Model Under Consideration*, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC) document JCTVC-A205, ITU-T and ISO/IEC JTC 1, Apr. 2010.
- [18] T. Wiegand, J.-R. Ohm, G. J. Sullivan, W.-J. Han, R. Joshi, T. K. Tan, and K. Ugur, "Special section on the joint call for proposals on High Efficiency Video Coding (HEVC) standardization," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 12, pp. 1661–1666, Dec. 2010.
- [19] K. McCann, *Tool Experiment 12: Evaluation of TMuC Tools*, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC) document JCTVC-B312, Jul. 2010.
- [20] K. McCann, *Summary of Evaluation of TMuC Tools in TE12*, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC) document JCTVC-C225, Oct. 2010.
- [21] K. McCann, B. Bross, S. Sekiguchi, and W.-J. Han, *Encoder-Side Description of HEVC Test Model (HM)*, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC) document JCTVC-C402, Oct. 2010.
- [22] T. Wiegand, W.-J. Han, B. Bross, J.-R. Ohm, and G. J. Sullivan, *Working Draft 1 of High Efficiency Video Coding*, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC) document JCTVC-C402, Oct. 2010.
- [23] J.-R. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan, and T. Wiegand, "Comparison of the coding efficiency of video coding standards—Including High Efficiency Video Coding (HEVC)," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1668–1683, Dec. 2012.
- [24] F. Bossen, B. Bross, K. Sühring, and D. Flynn, "HEVC complexity and implementation analysis," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1684–1695, Dec. 2012.



Gary J. Sullivan (S'83–M'91–SM'01–F'06) received the B.S. and M.Eng. degrees in electrical engineering from the University of Louisville, Louisville, KY, in 1982 and 1983, respectively, and the Ph.D. and Engineers degrees in electrical engineering from the University of California, Los Angeles, in 1991.

He has been a long-standing Chairman or Co-Chairman of various video and image coding standardization activities in ITU-T and ISO/IEC. He is best known for leading the development of the ITU-T H.264/ISO/IEC 14496-10 MPEG-4 Advanced Video Coding standard and its Scalable Video Coding and 3-D/Stereo/Multiview Video Coding extensions. He is a Video/Image Technology Architect with the Windows Division, Microsoft Corporation. With Microsoft Corporation, he has been the Originator and Lead Designer of the DirectX Video Acceleration video decoding feature of the Microsoft Windows operating system. His current research interests and areas of publication include image and video compression, rate-distortion optimization, motion estimation and compensation, scalar and vector quantization, and loss resilient video coding.

Dr. Sullivan has received the IEEE Masaru Ibuka Consumer Electronics Award, the IEEE Consumer Electronics Engineering Excellence Award, the INCITS Technical Excellence Award, the IMTC Leadership Award, and the University of Louisville J. B. Speed Professional Award in Engineering. The team efforts that he has led have been recognized by the ATAS Primetime Emmy Engineering Award and by a pair of NATAS Technology and Engineering Emmy Awards. He is a fellow of the SPIE.



Jens-Rainer Ohm (M'92) received the Dipl.-Ing., Dr.-Ing., and Habil. degrees from the Technical University of Berlin (TUB), Berlin, Germany, in 1985, 1990, and 1997, respectively.

From 1985 to 1995, he was a Research Associate with the Institute of Telecommunication, TUB. From 1992 to 2000, he has served as a Lecturer on digital image processing, coding and transmission at TUB. From 1996 to 2000, he was a Project Coordinator with the Heinrich Hertz Institute, Berlin. In 2000,

he became a Full Professor and since then has been the Chair of the Institute of Communication Engineering, RWTH Aachen University, Aachen, Germany. His current research and teaching activities cover the areas of motion-compensated, stereoscopic and 3-D image processing, multimedia signal coding, transmission and content description, audio signal analysis, as well as fundamental topics of signal processing and digital

communication systems. He has authored textbooks on multimedia signal processing, analysis and coding, on communication engineering and signal transmission, as well as numerous papers in the fields mentioned above.

Dr. Ohm has been participating in the work of the Moving Picture Experts Group (MPEG) since 1998. He has been the Chair and Co-Chair of various standardization activities in video coding, namely, the MPEG Video Subgroup since 2002, the Joint Video Team of MPEG and ITU-T SG 16 VCEG from 2005 to 2009, and currently, the Joint Collaborative Team on Video Coding as well as the Joint Collaborative Team on 3-D Video Coding Extensions. For T-CSV'T, he has served as an Associate Editor from 2001 to 2005. He is a member of various professional organizations, including VDE/ITG, EURASIP, and AES.



Woo-Jin Han (M'02) received the M.S. and Ph.D. degrees in computer science from the Korea Advanced Institute of Science and Technology, Daejeon, Korea, in 1997 and 2002, respectively.

He is currently a Professor with the Department of Software Design and Management, Gachon University, Seongnam, Korea. From 2003 to 2011, he was a Principal Engineer with the Multimedia Platform Laboratory, Digital Media and Communication Research and Development Center, Samsung Electronics, Suwon, Korea. Since 2003, he has contributed

successfully to the ISO/IEC Moving Pictures Experts Group, Joint Video Team, and Joint Collaborative Team on Video Coding standardization efforts. His current research interests include high-efficiency video compression techniques, scalable video coding, multiview synthesis, and visual contents understanding.

Dr. Han was an editor of the HEVC video coding standard in 2010.



Thomas Wiegand (M'05–SM'08–F'11) received the Dipl.-Ing. degree in electrical engineering from the Technical University of Hamburg-Harburg, Germany, in 1995, and the Dr.-Ing. degree from the University of Erlangen-Nuremberg, Germany, in 2000.

He is currently a Professor with the Department of Electrical Engineering and Computer Science, Berlin Institute of Technology, Berlin, Germany, chairing the Image Communication Laboratory, and is jointly heading the Image Processing Department, Fraunhofer Institute for Telecommunications–

Heinrich Hertz Institute, Berlin. He joined the Heinrich Hertz Institute in 2000 as the Head of the Image Communication Group, Image Processing Department. Since 1995, he has been an active participant in standardization for multimedia with successful submissions to ITU-T VCEG, ISO/IEC MPEG, 3GPP, DVB, and IETF. His current research interests include video processing and coding, multimedia transmission, and computer vision and graphics.

Dr. Wiegand was a recipient of the SPIE VCIP Best Student Paper Award in 1998, and the Fraunhofer Award and the ITG Award of the German Society for Information Technology in 2004. The projects that he co-chaired for development of the H.264/MPEG-4 AVC standard have been recognized by the 2008 ATAS Primetime Emmy Engineering Award and by a pair of NATAS Technology and Engineering Emmy Awards. In 2009, he received the Innovations Award of the Vodafone Foundation, the EURASIP Group Technical Achievement Award, and the Best Paper Award of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY. In 2010, he received the Eduard Rhein Technology Award. In 2011, he received the Best Paper Award of EURASIP and the Karl Heinz Beckurts Award. In October 2000, he was appointed as the Associated Rapporteur of ITU-T VCEG. In December 2001, he was appointed as the Associated Rapporteur/Co-Chair of the JVT. In February 2002, he was appointed as the editor of the H.264/MPEG-4 AVC video coding standard and its extensions (FRExt and SVC). From 2005 to 2009, he was the Co-Chair of MPEG Video. He is a recipient of the 2012 IEEE Masaru Ibuka Technical Field Award.

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국



(43) 국제공개일
2011년 5월 26일 (26.05.2011)

(10) 국제공개번호
WO 2011/062392 A2

- (51) 국제특허분류: H04N 7/32 (2006.01) H04N 7/24 (2011.01)
- (21) 국제출원번호: PCT/KR2010/007839
- (22) 국제출원일: 2010년 11월 8일 (08.11.2010)
- (25) 출원언어: 한국어
- (26) 공개언어: 한국어
- (30) 우선권정보: 10-2009-0111302 2009년 11월 18일 (18.11.2009) KR
- (71) 출원인 (US 을(를) 제외한 모든 지정국에 대하여): 에스케이텔레콤 주식회사 (SK TELECOM CO., LTD.) [KR/KR]; 서울 중구 을지로 2가 11번지, 100-999 Seoul (KR).
- (72) 발명자: 김
- (75) 발명자/출원인 (US 에 한하여): 김수년 (KIM, Sun-yeon) [KR/KR]; 서울 송파구 잠실동 잠실엘스아파트 158동 2603호, 138-220 Seoul (KR). 임정연 (LIM, Jeongyeon) [KR/KR]; 경기도 성남시 분당구 정자동 102 한솔마을아파트 415동 1902호, 463-010 Gyeonggi-do (KR). 이규민 (LEE, Gyumin) [KR/KR]; 경기도 성남시 분당구 금곡동 현대아리온오피스텔 308호, 463-853 Gyeonggi-do (KR). 최재운 (CHOI, Jaehoon) [KR/KR]; 경기도 성남시 분당구 수내동 로안팰리스 하우스빌 A동 2012호, 463-854 Gyeonggi-do (KR). 김용구 (KIM, Yonggoo) [KR/KR]; 서울시 서초구 반포본동 주공아파트 97동 507호, 137-813 Seoul (KR). 최윤식 (CHOE, Yoonsik) [KR/KR]; 경기도 고양시 일산서구 일산3동 후곡마을 테영아파트 1704동 501호, 411-313 Gyeonggi-do (KR). 최영호 (CHOI, Yunggho) [KR/KR]; 경기도 안양시 동안구 관양동 1595-1 동우베스티움 501호, 431-060 Gyeonggi-do (KR).
- (74) 대리인: 이철희 (LEE, Chulhee); 서울 강남구 역삼동 647-13 동강빌딩 5층, 135-911 Seoul (KR).
- (81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): 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, 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, RO, RS, RU, 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) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 역내 권리의 보호를 위하여): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), 유라시아 (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), 유럽 (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).
- 공개: — 국제조사보고서 없이 공개하며 보고서 접수 후 이를 별도 공개함 (규칙 48.2(g))

(54) Title: METHOD AND APPARATUS FOR ENCODING/DECODING A MOTION VECTOR BY SELECTING A SET OF PREDICTED CANDIDATE MOTION VECTORS, AND METHOD AND APPARATUS FOR IMAGE ENCODING/DECODING USING THE SAME

(54) 발명의 명칭: 후보 예측 움직임 벡터 집합 선택을 이용한 움직임 벡터 부호화/복호화 방법 및 장치와 그를 이용한 영상 부호화/복호화 방법 및 장치

(57) Abstract: The present invention relates to a method and apparatus for encoding/decoding a motion vector by selecting a set of predicted candidate motion vectors, and a method and apparatus for image encoding/decoding using the same. The method for encoding a motion vector according to the present invention comprises: selecting one set of predicted candidate motion vectors among plural sets of predicted candidate motion vectors by using information about motions of neighboring blocks of a current block; selecting one of the predicted candidate motion vectors in the selected set of predicted candidate motion vectors as a predicted motion vector; encoding a differential vector which represents the difference between the current motion vector as a motion vector of the current block and the predicted motion vector being selected; and encoding a predicted motion vector index that represents the predicted motion vector being selected. According to the present invention, as a motion vector is encoded by selecting an efficient set of predicted candidate motion vectors, the magnitude of a differential vector to be encoded can be reduced without necessarily encoding additional information to indicate which set of predicted candidate motion vectors has been selected. In result, the compression efficiency of a motion vector and further the compression efficiency of an image can be improved.

(57) 요약서: 본 발명은 후보 예측 움직임 벡터 집합 선택을 이용한 움직임 벡터 부호화/복호화 방법 및 장치와 영상 부호화/복호화 방법 및 장치에 관한 것이다. 본 발명은 현재 블록의 주변 블록의 움직임 정보를 이용하여 복수 개의 후보 예측 움직임 벡터 집합 중에서 하나의 후보 예측 움직임 벡터 집합을 선택하고, 선택되는 후보 예측 움직임 벡터 집합 내의 후보 예측 움직임 벡터 중에서 하나의 후보 예측 움직임 벡터를 예측 움직임 벡터로서 선택하며, 현재 블록의 움직임 벡터인 현재 움직임 벡터와 선택되는 예측 움직임 벡터의 차이인 차분 벡터를 부호화하며, 선택되는 예측 움직임 벡터를 나타내는 예측 움직임 벡터 인덱스를 부호화하는 단계를 포함하는 것을 특징으로 하는 움직임 벡터 부호화 방법을 제공한다. 본 발명에 의하면, 효율적인 후보 예측 움직임 벡터 집합을 선택하여 움직임 벡터를 부호화함에 따라 부호화할 차분 벡터의 크기를 줄이면서도 어떠한 후보 예측 움직임 벡터 집합이 선택되었는지를 나타내기 위한 부가 정보를 부호화하지 않아도 되므로, 움직임 벡터의 압축 효율을 향상시켜 결과적으로 영상의 압축 효율을 향상시킬 수 있다.

WO 2011/062392 A2

명세서

발명의 명칭: 후보 예측 움직임 벡터 집합 선택을 이용한 움직임 벡터 부호화/복호화 방법 및 장치와 그를 이용한 영상 부호화/복호화 방법 및 장치

기술분야

- [1] 본 발명은 후보 예측 움직임 벡터 집합 선택을 이용한 움직임 벡터 부호화/복호화 방법 및 장치와 그를 이용한 영상 부호화/복호화 방법 및 장치에 관한 것이다. 더욱 상세하게는, 영상을 예측 부호화하는 데 이용되는 움직임 벡터를 효율적으로 압축하여 영상의 압축 효율을 향상시키기 위한 방법 및 장치에 관한 것이다.

배경기술

- [2] H.264/AVC와 같은 통상적인 영상 압축 기술에서는 블록 기반으로 움직임을 추정하여 얻어지는 움직임 벡터를 예측 부호화하기 위해서, 부호화하고자 하는 블록의 주변 블록의 움직임 벡터들의 중앙값(Median)을 이용하여 예측 움직임 벡터를 결정하고, 부호화하고자 하는 움직임 벡터와 예측 움직임 벡터의 차분을 가변 길이 부호화하여 움직임 벡터를 압축한다.
- [3] 최근 ITU-T VCEG(Video Coding Expert Group)에서는 KTA(Key Technical Area)라는 이름으로 기존의 H.264/AVC 기반으로 더 높은 성능을 갖는 코덱을 연구해왔는데, 그 중 MVComp(Competition-based Motion Vector Coding)라는 기법을 통하여 기존의 H.264/AVC의 움직임 벡터 부호화 기법을 개선하였다.
- [4] MVComp는 여러 개의 후보 예측 움직임 벡터를 두고, 현재 움직임 벡터와 예측 후에 얻어지는 움직임 벡터와의 차분 값이 최소가 되는 후보 예측 움직임 벡터를 선택하여 복호화기에 선정된 후보 예측 움직임 벡터에 대한 정보를 전송해주는 방법으로, 기존의 H.264/AVC에 비해서 5% 정도의 부호화 압축 효율을 개선하였다. 하지만, 이 기법은 후보 예측 움직임 벡터의 개수가 늘어날수록 복호화기에 전송해야 할 인덱싱 부가 정보(Indexing Side Information)의 양이 증가하는 문제점이 있다.
- [5] 이에, 여러 개의 후보 예측 움직임 벡터 중에 현재 움직임 벡터와 가장 유사할 것과 같은 하나의 움직임 벡터를 부호화기에서 선택하여, 선택된 움직임 벡터가 최적의 움직임 벡터인지 아닌지만을 구별하기 위한 부가 정보를 보내는 기술 등이 제안되었다. 하지만, 제안된 방법은 현재 프레임의 비트를 에러 없이 전달 받았음에도 불구하고, 이전 프레임에서 발생한 에러 때문에 현재 프레임 및 다음 인트라 프레임 전까지의 프레임을 복원할 수 없게 되는 복호화기 충돌 문제(Decoder Crash Problem)와 복호화기에 많은 연산량 증가를 가져오는 한계를 가지는 문제점이 있다.
- [6] 이에, 현재 부호화할 블록 주위의 움직임 벡터를 이용하여 현재 블록의 움직임

백터를 결정하는 기법들이 제안되었다. 이 기법은 주변 블록의 움직임 백터를 이용하여 인덱싱 부가 정보를 효율적으로 전송하고 후보 예측 움직임 백터의 선택을 이전 프레임의 정보와 분리시켜 복호화기 충돌 문제를 해결하였지만, 한정된 개수의 후보 예측 움직임 백터를 사용하기 때문에 압축 성능에 있어서 제한적이라는 한계를 가지는 문제점이 있다.

발명의 상세한 설명

기술적 과제

- [7] 전술한 문제점을 해결하기 위해 본 발명은, 부호화할 움직임 백터와 더욱 유사한 예측 움직임 백터를 이용하여 움직임 백터를 부호화하면서도 그로 인해 발생하는 비트량을 줄여 움직임 백터를 효율적으로 압축하고 그에 따라 영상의 압축 효율을 향상시키는 데 주된 목적이 있다.

과제 해결 수단

- [8] 전술한 목적을 달성하기 위해 본 발명은, 움직임 백터를 부호화하는 장치에 있어서, 현재 블록의 주변 블록의 움직임 정보를 이용하여 복수 개의 후보 예측 움직임 백터 집합 중에서 하나의 후보 예측 움직임 백터 집합을 선택하는 후보 예측 움직임 백터 집합 선택기; 선택되는 후보 예측 움직임 백터 집합 내의 후보 예측 움직임 백터 중에서 하나의 후보 예측 움직임 백터를 예측 움직임 백터로서 선택하는 예측 움직임 백터 선택기; 현재 블록의 움직임 백터인 현재 움직임 백터와 선택되는 예측 움직임 백터의 차이인 차분 백터를 부호화하는 차분 백터 부호화기; 및 선택되는 예측 움직임 백터를 나타내는 예측 움직임 백터 인덱스를 부호화하는 예측 움직임 백터 부호화기를 포함하는 것을 특징으로 하는 움직임 백터 부호화 장치를 제공한다.
- [9] 또한, 본 발명의 다른 목적에 의하면, 영상을 부호화하는 장치에 있어서, 현재 블록의 현재 움직임 백터를 결정하고, 결정되는 현재 움직임 백터를 이용하여 현재 블록을 예측 부호화하는 영상 부호화기; 및 복수 개의 후보 예측 움직임 백터 집합 중 현재 블록의 주변 블록의 움직임 정보를 기초로 선택되는 후보 예측 움직임 백터 집합에서 예측 움직임 백터를 선택하고, 현재 움직임 백터와 선택되는 예측 움직임 백터의 차이인 차분 백터 및 선택되는 예측 움직임 백터를 나타내는 예측 움직임 백터 인덱스를 부호화하는 움직임 백터 부호화기를 포함하는 것을 특징으로 하는 영상 부호화 장치를 제공한다.
- [10] 또한, 본 발명의 또 다른 목적에 의하면, 움직임 백터를 복호화하는 장치에 있어서, 현재 블록의 주변 블록의 움직임 정보를 이용하여 복수 개의 후보 예측 움직임 백터 집합 중에서 하나의 후보 예측 움직임 백터 집합을 선택하는 후보 예측 움직임 백터 집합 선택기; 움직임 백터 데이터로부터 추출되는 인덱스 데이터를 복호화하여 예측 움직임 백터 인덱스를 복원하고, 선택되는 후보 예측 움직임 백터 집합에서 복원되는 예측 움직임 백터 인덱스에 의해 식별되는 후보 예측 움직임 백터를 예측 움직임 백터로서 복원하는 예측 움직임 백터 복원기;

및 움직임 벡터 데이터로부터 추출되는 차분 벡터 데이터를 복호화하여 차분 벡터를 복원하고, 복원되는 차분 벡터와 복원되는 예측 움직임 벡터를 가산하여 현재 블록의 현재 움직임 벡터를 복원하는 현재 움직임 벡터 복원기를 포함하는 것을 특징으로 하는 움직임 벡터 복호화 장치를 제공한다.

- [11] 또한, 본 발명의 또 다른 목적에 의하면, 영상을 복호화하는 장치에 있어서, 비트스트림으로부터 추출되는 움직임 벡터 데이터를 복호화하여 차분 벡터와 예측 움직임 벡터 인덱스를 복원하고, 현재 블록의 주변 블록의 움직임 정보를 이용하여 복수 개의 후보 예측 움직임 벡터 집합 중에서 하나의 후보 예측 움직임 벡터 집합을 선택하며, 선택되는 후보 예측 움직임 벡터 집합에서 복원되는 예측 움직임 벡터 인덱스에 의해 식별되는 후보 예측 움직임 벡터를 예측 움직임 벡터로서 선택하며, 복원되는 차분 벡터와 선택되는 예측 움직임 벡터를 가산하여 현재 블록의 현재 움직임 벡터를 복원하는 움직임 벡터 복호화기; 및 비트스트림으로부터 추출되는 영상 데이터를 복원되는 현재 움직임 벡터를 이용하여 예측 복호화하여 현재 블록을 복원하는 영상 복호화기를 포함하는 것을 특징으로 하는 영상 복호화 장치를 제공한다.
- [12] 또한, 본 발명의 또 다른 목적에 의하면, 움직임 벡터를 부호화하는 방법에 있어서, 현재 블록의 주변 블록의 움직임 정보를 이용하여 복수 개의 후보 예측 움직임 벡터 집합 중에서 하나의 후보 예측 움직임 벡터 집합을 선택하는 단계; 선택되는 후보 예측 움직임 벡터 집합 내의 후보 예측 움직임 벡터 중에서 하나의 후보 예측 움직임 벡터를 예측 움직임 벡터로서 선택하는 단계; 현재 블록의 움직임 벡터인 현재 움직임 벡터와 선택되는 예측 움직임 벡터의 차이인 차분 벡터를 부호화하는 단계; 및 선택되는 예측 움직임 벡터를 나타내는 예측 움직임 벡터 인덱스를 부호화하는 단계를 포함하는 것을 특징으로 하는 움직임 벡터 부호화 방법을 제공한다.
- [13] 또한, 본 발명의 또 다른 목적에 의하면, 영상을 부호화하는 방법에 있어서, 현재 블록의 현재 움직임 벡터를 결정하는 단계; 결정되는 현재 움직임 벡터를 이용하여 현재 블록을 예측 부호화하는 단계; 복수 개의 후보 예측 움직임 벡터 집합 중 현재 블록의 주변 블록의 움직임 정보를 기초로 선택되는 후보 예측 움직임 벡터 집합에서 예측 움직임 벡터를 선택하는 단계; 및 현재 움직임 벡터와 선택되는 예측 움직임 벡터의 차이인 차분 벡터와 선택되는 예측 움직임 벡터를 나타내는 예측 움직임 벡터 인덱스를 부호화하는 단계를 포함하는 것을 특징으로 하는 영상 부호화 방법을 제공한다.
- [14] 또한, 본 발명의 또 다른 목적에 의하면, 움직임 벡터를 복호화하는 방법에 있어서, 움직임 벡터 데이터로부터 추출되는 차분 벡터 데이터와 인덱스 데이터를 복호화하여 차분 벡터와 예측 움직임 벡터 인덱스를 복원하는 단계; 현재 블록의 주변 블록의 움직임 정보를 이용하여 복수 개의 후보 예측 움직임 벡터 집합 중에서 하나의 후보 예측 움직임 벡터 집합을 선택하는 단계; 선택되는 후보 예측 움직임 벡터 집합에서 복원되는 예측 움직임 벡터 인덱스에

- [125] 도 9는 본 발명의 일 실시예에 따른 동작인 벡터 복호화 방법을 설명하기 위한 나타낸 블록 구성도,
- [124] 도 8은 본 발명의 일 실시예에 따른 동작인 벡터 복호화 장치를 간략하게 구성도,
- [123] 도 7은 본 발명의 일 실시예에 따른 영상 부호화 장치를 간략하게 나타낸 블록
- [122] 도 6은 본 발명의 일 실시예에 따른 영상 부호화 방법을 설명하기 위한 순서도, 순서도,
- [121] 도 5는 본 발명의 일 실시예에 따른 동작인 벡터 부호화 방법을 설명하기 위한 방법을 구현한 예를 나타낸 예시도,
- [120] 도 4는 본 발명의 일 실시예에 따라 후보 예측 동작인 벡터 집합을 선택하는
- [119] 도 3은 본 발명의 일 실시예에 따른 주변 블록의 동작인 정보를 나타낸 예시도, 나타낸 블록 구성도,
- [118] 도 2는 본 발명의 일 실시예에 따른 동작인 벡터 부호화 장치를 간략하게 구성도,
- [117] 도 1은 본 발명의 일 실시예에 따른 영상 부호화 장치를 간략하게 나타낸 블록
- 도면의 간단한 설명**
- 영상시켜 결과적으로 영상의 압축 효율을 향상시킬 수 있다.
- 위한 부가 정보를 부호화하지 않아도 되므로, 동작인 벡터의 압축 효율을 줄이면서도 어떠한 후보 예측 동작인 벡터 집합이 선택되었는지를 나타내기
- 집합을 선택하여 동작인 벡터를 부호화함에 따라 부호화할 차분 벡터의 크기를
- [116] 이상에서 설명한 바와 같이 본 발명에 의하면, 효율적인 후보 예측 동작인 벡터
- 발명의 효과**
- 특징으로 하는 영상 부호화 방법을 제공한다.
- 이용하여 예측 부호화하여 현재 블록을 복원하는 단계를 포함하는 것을
- 및 스트림으로부터 추출되는 영상 데이터를 복원되는 현재 동작인 벡터를
- 예측 동작인 벡터를 가산하여 현재 블록의 현재 동작인 벡터를 복원하는 단계;
- 벡터를 예측 동작인 벡터로서 선택하는 단계; 복원되는 차분 벡터와 선택되는
- 집합에서 복원되는 예측 동작인 벡터 인덱스에 의해 선택되는 후보 예측 동작인
- 예측 동작인 벡터 집합을 선택하는 단계; 선택되는 후보 예측 동작인 벡터
- 정보를 이용하여 복수 개의 후보 예측 동작인 벡터 집합 중에서 하나의 후보
- 예측 동작인 벡터 인덱스를 복원하는 단계; 현재 블록의 주변 블록의 동작인
- 및 스트림으로부터 추출되는 동작인 벡터 데이터를 부호화하여 차분 벡터와
- [115] 또한, 본 발명의 또 다른 목적에 의하면, 영상을 부호화하는 방법에 있어서,
- 부호화 방법을 제공한다.
- 현재 동작인 벡터를 복원하는 단계를 포함하는 것을 특징으로 하는 동작인 벡터
- 및 복원되는 차분 벡터와 복원되는 예측 동작인 벡터를 가산하여 현재 블록의
- 의해 선택되는 후보 예측 동작인 벡터를 예측 동작인 벡터로서 복원하는 단계;

순서도,

- [26] 도 10은 본 발명의 일 실시예에 따른 영상 복호화 방법을 설명하기 위한 순서도이다.

발명의 실시를 위한 형태

- [27] 이하, 본 발명의 일부 실시예들을 예시적인 도면을 통해 상세하게 설명한다. 각 도면의 구성요소들에 참조부호를 부가함에 있어서, 동일한 구성요소들에 대해서는 비록 다른 도면상에 표시되더라도 가능한 한 동일한 부호를 가지도록 하고 있음에 유의해야 한다. 또한, 본 발명을 설명함에 있어, 관련된 공지 구성 또는 기능에 대한 구체적인 설명이 본 발명의 요지를 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명은 생략한다.
- [28] 또한, 본 발명의 구성 요소를 설명하는 데 있어서, 제 1, 제 2, A, B, (a), (b) 등의 용어를 사용할 수 있다. 이러한 용어는 그 구성 요소를 다른 구성 요소와 구별하기 위한 것일 뿐, 그 용어에 의해 해당 구성 요소의 본질이나 차례 또는 순서 등이 한정되지 않는다. 어떤 구성 요소가 다른 구성 요소에 "연결", "결합" 또는 "접속"된다고 기재된 경우, 그 구성 요소는 그 다른 구성 요소에 직접적으로 연결되거나 접속될 수 있지만, 각 구성 요소 사이에 또 다른 구성 요소가 "연결", "결합" 또는 "접속"될 수도 있다고 이해되어야 할 것이다.
- [29] 이하에서 후술할 움직임 벡터 부호화 장치(Motion Vector Encoding Apparatus), 움직임 벡터 복호화 장치(Motion Vector Decoding Apparatus), 영상 부호화 장치(Video Encoding Apparatus), 영상 복호화 장치(Video Decoding Apparatus)는 개인용 컴퓨터(PC: Personal Computer), 노트북 컴퓨터, 개인 휴대 단말기(PDA: Personal Digital Assistant), 휴대형 멀티미디어 플레이어(PMP: Portable Multimedia Player), 플레이스테이션 포터블(PSP: PlayStation Portable), 이동통신 단말기(Mobile Communication Terminal) 등과 같은 사용자 단말기이거나 응용 서버와 서비스 서버 등 서버 단말기일 수 있으며, 각종 기기 또는 유무선 통신망과 통신을 수행하기 위한 통신 모듈 등의 통신 장치, 움직임을 부호화하거나 복호화하거나 영상을 부호화하거나 복호화하기 위한 각종 프로그램과 데이터를 저장하기 위한 메모리, 프로그램을 실행하여 연산 및 제어하기 위한 마이크로프로세서 등을 구비하는 다양한 장치를 의미한다.
- [30] 또한, 움직임 벡터 부호화 장치 또는 영상 부호화 장치에 의해 비트스트림으로 부호화된 움직임 벡터 또는 영상은 실시간 또는 비실시간으로 인터넷, 근거리 무선 통신망, 무선랜망, 와이브로망, 이동통신망 등의 유무선 통신망 등을 통하여 케이블, 범용 직렬 버스(USB: Universal Serial Bus) 등과 같은 다양한 통신 인터페이스를 통해 움직임 벡터 복호화 장치 또는 영상 복호화 장치로 전송되어 움직임 벡터 복호화 장치에서 복호화되어 움직임 벡터로서 복원되거나 영상 복호화 장치에서 복호화되어 영상으로 복원되고 재생될 수 있다.

- [31] 통상적으로 동영상은 일련의 픽처(Picture)로 구성되어 있으며, 각 픽처들은 블록(Block)과 같은 소정의 영역으로 분할된다. 영상의 영역이 블록으로 분할되는 경우에는 분할된 블록은 부호화 방법에 따라 크게 인트라 블록(Intra Block), 인터 블록(Inter Block)으로 분류된다. 인트라 블록은 인트라 예측 부호화(Intra Prediction Coding) 방식을 사용하여 부호화되는 블록을 뜻하는데, 인트라 예측 부호화란 현재 부호화를 수행하는 현재 픽처 내에서 이전에 부호화되고 복호화되어 복원된 블록들의 화소를 이용하여 현재 부호화하고자 하는 블록인 현재 블록의 화소를 예측함으로써 예측 블록을 생성하고 현재 블록의 화소와의 차분값을 부호화하는 방식이다. 인터 블록은 인터 예측 부호화(Inter Prediction Coding)를 사용하여 부호화되는 블록을 뜻하는데, 인터 예측 부호화란 하나 이상의 과거 픽처 또는 미래 픽처를 참조하여 현재 픽처 내의 현재 블록을 예측함으로써 예측 블록을 생성하고 현재 블록과의 차분값을 부호화하는 방식이다. 여기서, 현재 픽처를 부호화하거나 복호화하는데 참조되는 픽처를 참조 픽처(Reference Picture)라고 한다.
- [32] 도 1은 본 발명의 일 실시예에 따른 영상 부호화 장치를 간략하게 나타낸 블록 구성도이다.
- [33] 본 발명의 일 실시예에 따른 영상 부호화 장치(100)는 영상을 부호화하는 장치로서, 움직임 벡터 부호화기(Motion Vector Encoder, 110)와 영상 부호화기(Video Encoder, 120)를 포함하여 구성될 수 있다.
- [34] 움직임 벡터 부호화기(110)는 복수 개의 후보 예측 움직임 벡터 집합(Candidate Predicted Motion Vector Set) 중 현재 블록의 주변 블록(Neighbor Block)의 움직임 정보(Motion Information)를 기초로 선택되는 후보 예측 움직임 벡터 집합에서 예측 움직임 벡터(Predicted Motion Vector)를 선택하고, 현재 움직임 벡터(Current Motion Vector)와 선택되는 예측 움직임 벡터의 차이인 차분 벡터(Differential Motion Vector) 및 선택되는 예측 움직임 벡터를 나타내는 예측 움직임 벡터 인덱스(Predicted Motion Vector Index)를 부호화한다. 차분 벡터는 부호화되어 차분 벡터 데이터로서 생성되고 예측 움직임 벡터 인덱스는 부호화되어 인덱스 데이터로서 생성된다. 따라서, 움직임 벡터 부호화기(110)는 차분 벡터 데이터와 그룹 인덱스 데이터를 포함하는 움직임 벡터 데이터를 생성한다.
- [35] 여기서, 움직임 벡터 부호화기(110)는 차분 벡터를 생성하는 데 있어서, 영상 부호화기(120)에서 현재 블록을 예측 부호화하기 위해 결정된 현재 움직임 벡터를 이용한다. 움직임 벡터 부호화기(110)에 대해서는 후술하는 과정에서 도 2를 통해 상세히 설명한다.
- [36] 영상 부호화기(120)는 현재 블록의 움직임 벡터인 현재 움직임 벡터를 결정하고 현재 움직임 벡터를 이용하여 현재 블록을 예측 부호화한다. 이와 같이 현재 블록이 예측 부호화되어 영상 데이터가 생성된다.
- [37] 이를 위해, 영상 부호화기(120)는 예측기(Predictor), 감산기(Subtractor), 변환기 및 양자화기(Transformer and Quantizer), 부호화기(Encoder)를 포함하여 구성될

수 있으며, 역 양자화기 및 역 변환기(Inverse Transformer and Inverse Quantizer), 가산기(Adder), 디블로킹 필터(Deblocking Filter), 메모리(Memory) 등을 추가로 포함할 수 있다. 여기서, 예측기는 현재 블록의 움직임 추정하여 현재 움직임 벡터를 결정하고 현재 움직임 벡터를 이용하여 현재 블록의 움직임을 보상하여 예측 블록(Predicted Block)을 생성하며, 감산기는 현재 블록과 예측 블록을 감산하여 잔여 블록(Residual Block)을 생성하며, 변환기 및 양자화기는 잔여 블록을 변환 및 양자화하여 양자화된 변환 계수(Quantized Transform Coefficient)를 생성하며, 부호화기는 양자화된 변환 계수를 부호화하여 영상 데이터를 생성한다. 또한, 역 양자화기 및 역 변환기는 양자화된 변환 계수를 역 양자화 및 역 변환하여 잔여 블록을 복원하고 가산기는 예측 블록과 복원되는 잔여 블록을 복원하여 현재 블록을 복원하며, 복원된 현재 블록은 디블로킹 필터(Deblocking Filter)에 의해 디블로킹 필터링되어 메모리에 픽처 단위로 누적되어 참조 픽처로서 저장되고 다음 블록 또는 다음 픽처를 예측하는 데 활용된다.

- [38] 도 2는 본 발명의 일 실시예에 따른 움직임 벡터 부호화 장치를 간략하게 나타낸 블록 구성도이다.
- [39] 본 발명의 일 실시예에 따른 움직임 벡터 부호화 장치는 도 1을 통해 기술한 영상 부호화 장치(100)에서는 움직임 벡터 부호화기(110)로 구현될 수 있다. 이하에서는 설명의 편의를 위해, 본 발명의 일 실시예에 따른 움직임 벡터 부호화 장치를 움직임 벡터 부호화기(110)라 칭한다.
- [40] 움직임 벡터 부호화기(110)는 후보 예측 움직임 벡터 집합 선택기(Candidate Predicted Motion Vector Set Selector, 210), 예측 움직임 벡터 선택기(Predicted Motion Vector Selector, 220), 차분 벡터 부호화기(Differential Motion Vector Encoder, 230) 및 예측 움직임 벡터 부호화기(Predicted Motion Vector Encoder, 240)를 포함하여 구성될 수 있다.
- [41] 후보 예측 움직임 벡터 집합 선택기(210)는 현재 블록의 주변 블록의 움직임 정보를 이용하여 복수 개의 후보 예측 움직임 벡터 집합 중에서 하나의 후보 예측 움직임 벡터 집합을 선택한다.
- [42] 본 발명의 일 실시예에서, 현재 블록의 주변 블록이란 현재 블록이 포함된 현재 픽처에서 현재 블록이 부호화되기 전에 이미 부호화되고 복호화되어 기 복원된 블록 중 현재 블록의 주변에 위치한 블록을 말한다. 주변 블록은 현재 블록과 인접한 인접 블록이 될 수도 있지만 반드시 인접한 블록만으로 한정되지는 않는다. 주변 블록의 움직임 정보란 기 부호화되고 복호화되어 복원된 주변 블록에 대해 움직임 벡터와 관련하여 획득된 정보들을 말하며, 주변 블록의 움직임 벡터와 예측 움직임 벡터 등이 될 수 있다.
- [43] 도 3은 본 발명의 일 실시예에 따른 주변 블록의 움직임 정보를 나타낸 예시도이다.
- [44] 3A는 현재 블록의 주변 블록과 주변 블록의 움직임 벡터를 예시적으로

나타내었다. 현재 블록의 주변에 위치한 블록들 중에서 현재 블록의 왼쪽에 인접한 블록을 블록 A, 위쪽에 인접한 블록을 블록 B, 위쪽 오른쪽에 인접한 블록을 블록 C라고 가정하면, 블록 A, 블록 B 및 블록 C가 현재 블록의 주변 블록이 될 수 있다. 이때, 블록 A, 블록 B 및 블록 C는 모두 현재 블록이 부호화되기 전에 이미 부호화되고 복호화되어 복원된 블록들이다. 3A에서는 블록 A, 블록 B 및 블록 C만이 현재 블록의 주변 블록인 것으로 도시하였지만, 이에 한정되지 않고 현재 블록의 위쪽 왼쪽에 위치한 블록 등과 같이 다른 블록들도 주변 블록으로 이용될 수 있다.

- [45] 3A에서, MV_a, MV_b, MV_c 는 블록 A, 블록 B 및 블록 C 각각의 움직임 벡터를 나타낸다. 블록 A, 블록 B 및 블록 C의 움직임 벡터는 각 블록을 부호화할 때 이미 결정되어 영상 부호화 장치(100) 또는 움직임 벡터 부호화기(110)의 버퍼 또는 메모리 등에 저장되어 있으므로, 영상 부호화 장치(100), 움직임 벡터 부호화기(110) 또는 후보 예측 움직임 벡터 집합 선택기(210)가 현재 블록을 부호화하거나 현재 블록의 움직임 벡터를 부호화할 때 언제든지 획득하여 이용할 수 있는 정보이다.
- [46] 3B는 주변 블록의 예측 움직임 벡터를 예시적으로 나타내었다. 3B에서, PMV_a, PMV_b, PMV_c 는 블록 A, 블록 B 및 블록 C 각각의 예측 움직임 벡터를 나타낸다. 블록 A, 블록 B 및 블록 C의 예측 움직임 벡터도 각 블록을 부호화할 때 이미 결정되어 영상 부호화 장치(100) 또는 움직임 벡터 부호화기(110)의 버퍼 또는 메모리 등에 저장되어 있으므로, 영상 부호화 장치(100), 움직임 벡터 부호화기(110) 또는 후보 예측 움직임 벡터 집합 선택기(210)가 현재 블록을 부호화하거나 현재 블록의 움직임 벡터를 부호화할 때 언제든지 획득하여 이용할 수 있는 정보이다.
- [47] 따라서, 3A와 3B에 나타낸 예의 경우, 주변 블록의 움직임 정보는 3C에 나타낸 바와 같이 $MV_a, MV_b, MV_c, PMV_a, PMV_b, PMV_c$ 가 될 수 있다.
- [48] 후보 예측 움직임 벡터 집합 선택기(210)는, 도 3에 도시한 바와 같이, 주변 블록의 움직임 벡터와 예측 움직임 벡터를 주변 블록의 움직임 정보로서 이용하여 복수 개의 후보 예측 움직임 벡터 집합 중에서 하나의 후보 예측 움직임 벡터 집합을 선택할 수 있다.
- [49] 일 예로, 후보 예측 움직임 벡터 집합 선택기(210)는 주변 블록의 움직임 정보를 이용하여 카메라의 움직임 여부를 판단하고 판단되는 카메라의 움직임 여부를 기초로 복수 개의 후보 예측 움직임 벡터 집합 중에서 하나의 후보 예측 움직임 벡터 집합을 선택할 수 있다. 이를 위해, 후보 예측 움직임 벡터 집합 선택기(210)는 주변 블록들 중에서 움직임 벡터가 제로 벡터(Zero Vector)인 주변 블록의 개수가 기 설정된 개수 이상인지 여부를 판단하여 카메라의 움직임 여부를 판단할 수 있다. 즉, 후보 예측 움직임 벡터 집합 선택기(210)는 움직임 벡터가 제로 벡터인 주변 블록의 개수가 기 설정된 개수 이상인 경우에는 카메라의 움직임이 없는 것으로 판단하고, 움직임 벡터가 제로 벡터인 주변

블록의 개수가 기 설정된 개수 미만인 경우에는 카메라의 움직임이 있는 것으로 판단할 수 있다.

- [50] 도 3에 도시한 예에서, 주변 블록들의 움직임 벡터 MV_a, MV_b, MV_c 가 각각 (0, 0), (0, 0), (0, 1)이고, 카메라의 움직임 여부를 판단하기 위한 개수로 2 개가 설정되었으며, 복수 개의 후보 예측 움직임 벡터 집합으로서 $\{MV_a, MV_{H.264}\}, \{MV_a, MV_{extspa}\}$ 두 개의 후보 예측 움직임 벡터 집합이 설정되었다고 가정하면, 움직임 벡터가 제로 벡터인 주변 블록의 개수가 2로서 기 설정된 개수인 2 이상이므로, 카메라의 움직임이 없는 것으로 판단되어 $MV_{H.264}$ 가 포함된 $\{MV_a, MV_{H.264}\}$ 가 후보 예측 움직임 벡터 집합으로서 선택될 수 있다. 이때, 주변 블록들의 움직임 벡터 MV_a, MV_b, MV_c 가 각각 (0, 0), (1, 0), (0, 1)이라고 가정하면, 움직임 벡터가 제로 벡터인 주변 블록의 개수가 1로서 기 설정된 개수인 2 미만이므로, 카메라의 움직임이 있는 것으로 판단되어 MV_{extspa} 가 포함된 $\{MV_a, MV_{extspa}\}$ 가 후보 예측 움직임 벡터 집합으로서 선택될 수 있다.
- [51] 이때, 카메라의 움직임이 없는 것으로 판단되면 $MV_{H.264}$ 가 포함된 집합을 후보 예측 움직임 벡터 집합으로서 선택하고 카메라의 움직임이 있는 것으로 판단되면 MV_{extspa} 가 포함된 집합을 후보 예측 움직임 벡터 집합으로서 선택하는 것은 예시적인 것일 뿐, 반드시 $MV_{H.264}$ 또는 MV_{extspa} 가 포함된 집합을 선택해야 하는 것은 아니며, 어떠한 후보 예측 움직임 벡터가 포함된 집합을 선택해야 하는지는 실험을 통해 경험적으로 결정될 수 있다. 즉, 카메라의 움직임 여부에 따라 선택되어 후보 예측 움직임 벡터 집합에 포함될 후보 예측 움직임 벡터는 실험을 통해 경험적으로 결정될 수 있으며, 이와 같이 경험적으로 결정되는 데이터는 영상 부호화 장치(110)와 영상 복호화 장치에 미리 설정하여 저장될 수 있다.
- [52] 다른 예로, 후보 예측 움직임 벡터 집합 선택기(210)는 주변 블록의 움직임 정보를 이용하여 주변 블록의 객체 움직임(Object Motion)을 판단하고, 판단되는 주변 블록의 객체 움직임을 기초로 복수 개의 후보 예측 움직임 벡터 집합 중에서 하나의 후보 예측 움직임 벡터 집합을 선택할 수 있다. 이를 위해, 후보 예측 움직임 벡터 집합 선택기(210)는 주변 블록의 예측 움직임 벡터를 이용하여 주변 블록의 객체 움직임을 판단할 수 있다. 즉, 후보 예측 움직임 벡터 집합 선택기(210)는 주변 블록의 객체 움직임을 판단하기 위해 주변 블록의 움직임 정보로서 주변 블록의 예측 움직임 벡터를 분석하고, 분석된 주변 블록의 예측 움직임 벡터의 분포에 따라 이용 가능한 후보 예측 움직임 벡터들 중 하나 이상의 후보 예측 움직임 벡터를 선택하고, 복수 개의 후보 예측 움직임 벡터 집합 중에서 선택된 하나 이상의 후보 예측 움직임 벡터를 포함하는 후보 예측 움직임 벡터 집합을 하나의 후보 예측 움직임 벡터 집합으로서 선택할 수 있다.
- [53] 도 3에 도시한 예에서, 설정된 복수 개의 후보 예측 움직임 벡터 집합으로서 $\{MV_a, MV_{H.264}\}, \{MV_b, MV_{H.264}\}, \{MV_c, MV_{H.264}\}$ 세 개의 후보 예측 움직임 벡터 집합이 설정되었다고 가정하면, 주변 블록들의 예측 움직임 벡터 $PMV_a, PMV_b,$

PMV_c 는 각각 MV_a, MV_a, MV_c 인데, 이 경우 현재 블록의 예측 움직임 벡터는 MV_a 가 될 확률이 MV_a 나 MV_c 가 될 확률보다 높다는 가정 하에서, $MV_a, MV_a, MV_c, MV_{H.264}, MV_{extspa}$ 등과 같은 이용 가능한 후보 예측 움직임 벡터들 중에서 MV_b 를 선택하여, 선택된 MV_b 를 포함하는 $\{MV_b, MV_{H.264}\}$ 가 후보 예측 움직임 벡터 집합으로서 선택될 수 있다. 이때, 주변 블록들의 예측 움직임 벡터 PMV_a, PMV_b, PMV_c 가 각각 MV_a, MV_a, MV_c 일 때, 현재 블록의 예측 움직임 벡터는 MV_b 가 될 확률이 MV_a 나 MV_c 가 될 확률보다 높다는 것은 예시적인 것일 뿐, 이는 실험을 통해 경험적으로 결정될 수 있다. 즉, 주변 블록의 예측 움직임 벡터의 분포에 따라 선택되어 후보 예측 움직임 벡터 집합에 포함될 후보 예측 움직임 벡터는 실험을 통해 경험적으로 결정될 수 있으며, 이와 같이 경험적으로 결정되는 데이터는 영상 부호화 장치(110)와 영상 복호화 장치에 미리 설정하여 저장될 수 있다.

- [54] 또 다른 예로, 도 4에 도시한 바와 같이, 선술한 카메라의 움직임 어부와 주변 블록의 객체 움직임 모두를 기초로 복수 개의 후보 예측 움직임 벡터 집합 중에서 하나의 후보 예측 움직임 벡터 집합을 선택할 수 있다.
- [55] 도 4는 본 발명의 일 실시예에 따라 후보 예측 움직임 벡터 집합을 선택하는 방법을 구현한 예를 나타낸 예시도이다.
- [56] 복수 개의 후보 예측 움직임 벡터 집합들이 $\{MV_a, MV_{H.264}\}, \{MV_b, MV_{H.264}\}, \{MV_c, MV_{H.264}\}, \{MV_a, MV_{extspa}\}, \{MV_b, MV_{extspa}\}, \{MV_c, MV_{extspa}\}$ 로 설정된 경우(S410), 후보 예측 움직임 벡터 집합 선택기(210)는 주변 블록의 움직임 정보를 분석하여(S420), 주변 블록의 객체 움직임을 기초로 선술한 예에서와 같이 이용 가능한 후보 예측 움직임 벡터 중에서 MV_a 를 선택하며(S430), 카메라 움직임이 없는지 여부를 판단하여(S410), 카메라 움직임이 없는 경우에는 $MV_{H.264}$ 를 선택하여(S450), $\{MV_a, MV_{H.264}\}$ 를 후보 예측 움직임 벡터 집합으로서 선택하며(S460), 카메라 움직임이 있는 경우에는 MV_{extspa} 를 선택하여(S470), $\{MV_a, MV_{extspa}\}$ 를 후보 예측 움직임 벡터 집합으로서 선택할 수 있다(S480).
- [57] 도 3 및 도 4를 통해서는 복수 개의 후보 예측 움직임 벡터 집합이 2 개, 3 개, 6 개인 경우에 대해서 예를 들었지만, 후보 예측 움직임 벡터 집합은 2 개 이상이라면 제한 없이 설정될 수 있다. 또한, 도 3 및 도 4를 통해서는 각 후보 예측 움직임 벡터 집합 내에 두 개의 후보 예측 움직임 벡터가 포함되는 것으로 예를 들었지만, 반드시 두 개에 한정되지 않고 두 개 이상 복수 개의 후보 예측 움직임 벡터가 포함될 수 있다.
- [58] 다시 도 2를 참조하면, 예측 움직임 벡터 선택기(220)는 선택되는 후보 예측 움직임 벡터 집합 내의 후보 예측 움직임 벡터 중에서 하나의 후보 예측 움직임 벡터를 예측 움직임 벡터로서 선택한다. 이를 위해, 예측 움직임 벡터 선택기(220)는 후보 예측 움직임 벡터 집합 선택기(210)에 의해 선택되는 후보 예측 움직임 벡터 집합 내에 포함되는 후보 예측 움직임 벡터들 중에서 윗-왜곡 측면에서 최적이 되는 후보 예측 움직임 벡터를 예측 움직임 벡터로서 선택할 수

있다. 여기서, 윌-왜곡 측면에서 최적이 되는 후보 예측 움직임 벡터란 해당 후보 예측 움직임 벡터들 각각을 이용하여 현재 움직임 벡터를 부호화하고 현재 블록을 부호화하여 생성되는 비트스트림의 윌-왜곡 비용이 가장 작은 후보 예측 움직임 벡터를 말한다.

- [59] 예를 들어, 예측 움직임 벡터 선택기(220)에 의해 선택되는 후보 예측 움직임 벡터 집합이 $\{MV_a, MV_{H.264}\}$ 라고 가정하면, 수학식 1을 이용하여 예측 움직임 벡터가 선택될 수 있다.

[60] [수학식 1]

$$[61] \quad BMV = \arg \min_l \left[\sum_i \sum_j |x_{i,j} - \hat{x}_{i,j}| + \lambda \cdot \{R(MV_c - MV_l) + R(l)\} \right]$$

- [62] 수학식 1에서, BMVI(Best Motion Vector Index)는 예측 움직임 벡터 선택기(220)에 의해 선택되는 예측 움직임 벡터를 식별하기 위한 예측 움직임 벡터 인덱스를 나타내며, $x_{i,j}$ 는 현재 화소의 좌표 (i,j) 에 대한 화소값을 나타내며, $\hat{x}_{i,j}$ 는 참조 화소의 좌표 (i,j) 에 대한 화소값을 나타내며, $R(MV - MV_l)$ 과 $R(l)$ 은 각각 현재 블록의 움직임 벡터와 예측 움직임 벡터와의 차이를 부호화하는 데 소요되는 비트량과 예측 움직임 벡터 인덱스를 부호화하는 데 소요되는 비트량이 계산된 값을 나타낸다.

- [63] 차분 벡터 부호화기(230)는 현재 블록의 움직임 벡터인 현재 움직임 벡터와 선택되는 예측 움직임 벡터의 차이인 차분 벡터를 부호화한다. 즉, 차분 벡터 부호화기(230)는 현재 움직임 벡터와 예측 움직임 벡터 선택기(220)에 의해 선택되는 예측 움직임 벡터를 감산하여 차분 벡터를 생성하고 차분 벡터를 부호화하여 차분 벡터 데이터를 생성한다. 예를 들어, 예측 움직임 벡터 선택기(220)에 의해 선택되는 예측 움직임 벡터가 MV_a 라고 가정하면, 차분 벡터는 수학식 2와 같이 계산될 수 있다. 수학식 2에서, DMV 는 차분 벡터를 나타내고, MV 는 현재 움직임 벡터를 나타낸다.

[64] [수학식 2]

$$[65] \quad DMV = MV_c - MV_a$$

- [66] 다만, 차분 벡터 부호화기(230)는 차분 벡터를 별도로 부호화하지 않고 예측 움직임 벡터 선택기(220)가 윌-왜곡 비용을 구하기 위해 차분 벡터를 구하고 부호화한 경우 예측 움직임 벡터 선택기(220)에 의해 부호화된 차분 벡터 데이터를 출력할 수 있다. 차분 벡터를 부호화하는 기법으로서는 고정 길이 부호화(Fixed Length Coding), 가변 길이 부호화(Variable Length Coding), 산술 부호화(Arithmetic Coding) 등과 같은 엔트로피 부호화(Entropy Coding) 기법이 이용될 수 있다.

- [67] 예측 움직임 벡터 부호화기(240)는 예측 움직임 벡터 선택기(220)에 의해 선택되는 예측 움직임 벡터를 나타내는 예측 움직임 벡터 인덱스를 부호화한다. 예측 움직임 벡터 인덱스를 부호화하는 기법으로서는 차분 벡터를 부호화할

때와 마찬가지로 고정 길이 부호화, 가변 길이 부호화, 산술 부호화 등과 같은 엔트로피 부호화 기법이 이용될 수 있다.

- [68] 예를 들어, 후보 예측 움직임 벡터 집합 선택기(210)에 의해 선택되는 후보 예측 움직임 벡터 집합이 $\{MV_a, MV_{H.264}\}$ 라고 가정하면, MV_a 의 예측 움직임 벡터 인덱스는 '0'이고 $MV_{H.264}$ 의 예측 움직임 벡터 인덱스는 '1'로 설정될 수 있다. 따라서, 이 경우, 후보 예측 움직임 벡터 집합 내에 두 개의 후보 예측 움직임 벡터만이 존재하므로 예측 움직임 벡터 인덱스가 부호화되어 생성되는 인덱스 데이터는 '0' 또는 '1'의 값을 가지는 1 비트로 생성될 수 있다. 여기서, 영상 복호화 장치 또는 움직임 벡터 복호화 장치에서 예측 움직임 벡터 인덱스를 올바르게 복원하기 위해서는 후보 예측 움직임 벡터 집합별로 각 집합 내에 포함되는 후보 예측 움직임 벡터에 대한 예측 움직임 벡터 인덱스가 영상 부호화 장치(100)와 영상 복호화 장치 또는 움직임 벡터 부호화 장치(110)와 움직임 벡터 복호화 장치에 미리 동일하게 설정되어 저장되어야 한다.
- [69] 도 5는 본 발명의 일 실시예에 따른 움직임 벡터 부호화 방법을 설명하기 위한 순서도이다.
- [70] 본 발명의 일 실시예에 따른 움직임 벡터 부호화 방법에 따르면, 움직임 벡터 부호화기(110)는 현재 블록의 주변 블록의 움직임 정보를 이용하여 복수 개의 후보 예측 움직임 벡터 집합 중에서 하나의 후보 예측 움직임 벡터 집합을 선택하고(S510), 선택되는 후보 예측 움직임 벡터 집합 내의 후보 예측 움직임 벡터 중에서 하나의 후보 예측 움직임 벡터를 예측 움직임 벡터로서 선택하며(S520), 현재 블록의 움직임 벡터인 현재 움직임 벡터와 선택되는 예측 움직임 벡터의 차이인 차분 벡터를 부호화하며(S530), 선택되는 예측 움직임 벡터를 나타내는 예측 움직임 벡터 인덱스를 부호화한다(S540).
- [71] 단계 S510에서, 움직임 벡터 부호화기(110)는 주변 블록의 움직임 정보를 이용하여 카메라의 움직임 여부를 판단하고 판단되는 카메라의 움직임 여부를 기초로 하나의 후보 예측 움직임 벡터 집합을 선택할 수 있다.
- [72] 또한, 단계 S510에서, 움직임 벡터 부호화기(110)는 주변 블록의 움직임 정보를 이용하여 주변 블록의 객체 움직임을 판단하고, 판단되는 주변 블록의 객체 움직임을 기초로 하나의 후보 예측 움직임 벡터 집합을 선택할 수 있다. 이를 위해, 움직임 벡터 부호화기(110)는 주변 블록의 예측 움직임 벡터를 이용하여 주변 블록의 객체 움직임을 판단할 수 있다.
- [73] 또한, 단계 S510에서, 움직임 벡터 부호화기(110)는 단계 S510에서, 주변 블록의 움직임 정보를 이용하여 주변 블록의 객체 움직임을 판단하고 판단되는 객체 움직임을 기초로 후보 예측 움직임 벡터를 선택하며, 주변 블록의 움직임 정보를 이용하여 카메라의 움직임 여부를 판단하고 판단되는 카메라의 움직임 여부를 기초로 후보 예측 움직임 벡터를 선택하며, 복수 개의 후보 예측 움직임 벡터 집합 중에서 객체 움직임을 기초로 선택되는 후보 예측 움직임 벡터 및 카메라의 움직임 여부를 기초로 선택되는 후보 예측 움직임 벡터를 포함하는

후보 예측 움직임 벡터 집합을 하나의 후보 예측 움직임 벡터 집합으로서 선택할 수 있다.

- [74] 도 6은 본 발명의 일 실시예에 따른 영상 부호화 방법을 설명하기 위한 순서도이다.
- [75] 본 발명의 일 실시예에 따른 영상 부호화 방법에 따르면, 영상 부호화 장치(100)는 현재 블록의 현재 움직임 벡터를 결정하고(S610), 결정되는 현재 움직임 벡터를 이용하여 현재 블록을 예측 부호화하며(S620), 복수 개의 후보 예측 움직임 벡터 집합 중 현재 블록의 주변 블록의 움직임 정보를 기초로 선택되는 후보 예측 움직임 벡터 집합에서 예측 움직임 벡터를 선택하며(S630), 현재 움직임 벡터와 선택되는 예측 움직임 벡터의 차이인 차분 벡터와 선택되는 예측 움직임 벡터를 나타내는 예측 움직임 벡터 인덱스를 부호화한다(S640).
- [76] 도 7은 본 발명의 일 실시예에 따른 영상 복호화 장치를 간략하게 나타낸 블록 구성도이다.
- [77] 본 발명의 일 실시예에 따른 영상 복호화 장치(700)는 움직임 벡터 복호화기(Motion Vector Decoder, 710) 및 영상 복호화기(Video Decoder, 720)를 포함하여 구성될 수 있다.
- [78] 움직임 벡터 복호화기(710)는 비트스트림으로부터 추출되는 움직임 벡터 데이터를 복호화하여 차분 벡터와 예측 움직임 벡터 인덱스를 복원하고, 현재 블록의 주변 블록의 움직임 정보를 이용하여 복수 개의 후보 예측 움직임 벡터 집합 중에서 하나의 후보 예측 움직임 벡터 집합을 선택하며, 선택되는 후보 예측 움직임 벡터 집합에서 복원되는 예측 움직임 벡터 인덱스에 의해 식별되는 후보 예측 움직임 벡터를 예측 움직임 벡터로서 선택하며, 복원되는 차분 벡터와 선택되는 예측 움직임 벡터를 가산하여 현재 블록의 현재 움직임 벡터를 복원한다. 움직임 벡터 복호화기(710)에 대해서는 후술하는 과정에서 도 8을 통해 상세히 설명한다.
- [79] 영상 복호화기(720)는 비트스트림으로부터 추출되는 영상 데이터를 복원되는 현재 움직임 벡터를 이용하여 예측 부호화하여 현재 블록을 복원한다. 이를 위해, 영상 복호화기(720)는 복호화기(Decoder), 역 양자화기 및 역 변환기, 예측기, 가산기, 디블로킹 필터, 메모리 등을 포함하여 구성될 수 있다. 여기서, 복호화기는 비트스트림으로부터 추출되는 영상 데이터를 복호화하여 양자화된 변환 계수를 복원하고, 역 양자화기 및 역 변환기는 복원되는 양자화된 변환 계수를 역 양자화 및 역 변환하여 잔여 블록을 복원하며, 예측기는 움직임 벡터 복호화기(710)에 의해 복원되는 현재 블록의 현재 움직임 벡터를 이용하여 현재 블록의 움직임을 보상함으로써 예측 블록을 생성하며, 가산기는 복원되는 잔여 블록과 예측 블록을 가산하여 현재 블록을 복원할 수 있다. 복원되는 현재 블록은 디블로킹 필터에 의해 디블로킹 필터링되고 픽처 단위로 누적되어 복원 영상으로서 출력되거나 메모리에 저장되어 예측기가 다음 블록 또는 다음 픽처를 예측하는데 활용된다.

- [80] 도 8은 본 발명의 일 실시예에 따른 움직임 벡터 복호화 장치를 간략하게 나타낸 블록 구성도이다.
- [81] 본 발명의 일 실시예에 따른 움직임 벡터 복호화 장치는 도 7을 통해 기술한 영상 복호화 장치(700)에서는 움직임 벡터 복호화기(710)로 구현될 수 있다. 이하에서는 설명의 편의를 위해, 본 발명의 일 실시예에 따른 움직임 벡터 복호화 장치를 움직임 벡터 복호화기(710)라 칭한다.
- [82] 움직임 벡터 복호화기(710)는 후보 예측 움직임 벡터 집합 선택기(810), 예측 움직임 벡터 복원기(Predicted Motion Vector Reconstructor, 820) 및 현재 움직임 벡터 복원기(Current Motion Vector Reconstructor, 830)를 포함하여 구성될 수 있다.
- [83] 후보 예측 움직임 벡터 집합 선택기(810)는 현재 블록의 주변 블록의 움직임 정보를 이용하여 복수 개의 후보 예측 움직임 벡터 집합 중에서 하나의 후보 예측 움직임 벡터 집합을 선택한다. 여기서, 후보 예측 움직임 벡터 집합 선택기(810)는 도 2를 통해 기술한 후보 예측 움직임 벡터 집합 선택기(210)와 동일 또는 유사하므로 상세한 설명은 생략한다.
- [84] 예측 움직임 벡터 복원기(820)는 움직임 벡터 데이터로부터 추출되는 인덱스 데이터를 복호화하여 예측 움직임 벡터 인덱스를 복원하고, 선택되는 후보 예측 움직임 벡터 집합에서 복원되는 예측 움직임 벡터 인덱스에 의해 식별되는 후보 예측 움직임 벡터를 예측 움직임 벡터로서 복원한다. 즉, 예측 움직임 벡터 복원기(820)는 움직임 벡터 데이터로부터 인덱스 데이터를 추출하고 복호화하여 예측 움직임 벡터 인덱스를 복원하고, 후보 예측 움직임 벡터 집합 선택기(810)에 의해 선택되는 후보 예측 움직임 벡터 집합 내의 후보 예측 움직임 벡터들 중에서 복원되는 예측 움직임 벡터 인덱스에 의해 식별되는 후보 예측 움직임 벡터를 예측 움직임 벡터로서 복원한다.
- [85] 예를 들어, 후보 예측 움직임 벡터 집합 선택기(810)에 의해 선택되는 후보 예측 움직임 벡터 집합이 $\{MV_a, MV_{H.264}\}$ 이고, MV_a 의 예측 움직임 벡터 인덱스는 '0'이고 $MV_{H.264}$ 의 예측 움직임 벡터 인덱스는 '1'로 설정되어 있으며, 움직임 벡터 데이터로부터 추출되는 인덱스 데이터가 '0'의 비트라고 가정하면, 복원되는 예측 움직임 벡터 인덱스에 의해 식별되는 후보 예측 움직임 벡터는 MV_a 가 되어, MV_a 가 예측 움직임 벡터로서 복원될 수 있다. 여기서, 후보 예측 움직임 벡터 집합별로 각 집합 내에 포함되는 후보 예측 움직임 벡터에 대한 예측 움직임 벡터 인덱스는 영상 부호화 장치(100)와 영상 복호화 장치(700) 또는 움직임 벡터 부호화기(110)와 움직임 벡터 복호화기(710)에 미리 동일하게 설정되어 저장되어야 한다.
- [86] 현재 움직임 벡터 복원기(830)는 움직임 벡터 데이터로부터 추출되는 차분 벡터 데이터를 복호화하여 차분 벡터를 복원하고, 복원되는 차분 벡터와 복원되는 예측 움직임 벡터를 가산하여 현재 블록의 현재 움직임 벡터를 복원한다. 즉, 현재 움직임 벡터 복원기(830)는 움직임 벡터 데이터로부터 차분

벡터 데이터를 추출하고 복호화하여 차분 벡터를 복원하고, 복원되는 차분 벡터와 예측 움직임 벡터 복원기(820)에 의해 복원되는 예측 움직임 벡터를 가산하여 현재 움직임 벡터를 복원한다.

- [87] 도 9는 본 발명의 일 실시예에 따른 움직임 벡터 복호화 방법을 설명하기 위한 순서도이다.
- [88] 본 발명의 일 실시예에 따른 움직임 벡터 복호화 방법에 따르면, 움직임 벡터 복호화기(710)는 움직임 벡터 데이터로부터 추출되는 차분 벡터 데이터와 인덱스 데이터를 복호화하여 차분 벡터와 예측 움직임 벡터 인덱스를 복원하고(S910), 현재 블록의 주변 블록의 움직임 정보를 이용하여 복수 개의 후보 예측 움직임 벡터 집합 중에서 하나의 후보 예측 움직임 벡터 집합을 선택하며(S920), 선택되는 후보 예측 움직임 벡터 집합에서 복원되는 예측 움직임 벡터 인덱스에 의해 식별되는 후보 예측 움직임 벡터를 예측 움직임 벡터로서 복원하며(S930), 복원되는 차분 벡터와 복원되는 예측 움직임 벡터를 가산하여 현재 블록의 현재 움직임 벡터를 복원한다(S940).
- [89] 단계 S920에서, 움직임 벡터 복호화기(710)는 주변 블록의 움직임 정보를 이용하여 카메라의 움직임 여부를 판단하고 판단되는 카메라의 움직임 여부를 기초로 하나의 후보 예측 움직임 벡터 집합을 선택할 수 있다.
- [90] 또한, 단계 S920에서, 움직임 벡터 복호화기(710)는 주변 블록의 움직임 정보를 이용하여 주변 블록의 객체 움직임을 판단하고, 판단되는 주변 블록의 객체 움직임을 기초로 하나의 후보 예측 움직임 벡터 집합을 선택할 수 있다. 이를 위해, 움직임 벡터 복호화기(710)는 주변 블록의 예측 움직임 벡터를 이용하여 주변 블록의 객체 움직임을 판단할 수 있다.
- [91] 또한, 단계 S920에서, 움직임 벡터 복호화기(710)는 주변 블록의 움직임 정보를 이용하여 주변 블록의 객체 움직임을 판단하고 판단되는 객체 움직임을 기초로 후보 예측 움직임 벡터를 선택하며, 주변 블록의 움직임 정보를 이용하여 카메라의 움직임 여부를 판단하고 판단되는 카메라의 움직임 여부를 기초로 후보 예측 움직임 벡터를 선택하며, 복수 개의 후보 예측 움직임 벡터 집합 중에서 객체 움직임을 기초로 선택되는 후보 예측 움직임 벡터 및 카메라의 움직임 여부를 기초로 선택되는 후보 예측 움직임 벡터를 포함하는 후보 예측 움직임 벡터 집합을 하나의 후보 예측 움직임 벡터 집합으로서 선택할 수 있다.
- [92] 도 10은 본 발명의 일 실시예에 따른 영상 복호화 방법을 설명하기 위한 순서도이다.
- [93] 본 발명의 일 실시예에 따른 영상 복호화 방법에 따르면, 영상 복호화 장치(700)는 비트스트림으로부터 추출되는 움직임 벡터 데이터를 복호화하여 차분 벡터와 예측 움직임 벡터 인덱스를 복원하고(S1010), 현재 블록의 주변 블록의 움직임 정보를 이용하여 복수 개의 후보 예측 움직임 벡터 집합 중에서 하나의 후보 예측 움직임 벡터 집합을 선택하며(S1020), 선택되는 후보 예측 움직임 벡터 집합에서 복원되는 예측 움직임 벡터 인덱스에 의해 식별되는 후보

예측 움직임 벡터를 예측 움직임 벡터로서 선택하며(S1030), 복원되는 차분 벡터와 선택되는 예측 움직임 벡터를 가산하여 현재 블록의 현재 움직임 벡터를 복원하며(S1040), 비트스트림으로부터 추출되는 영상 데이터를 복원되는 현재 움직임 벡터를 이용하여 예측 복호화하여 현재 블록을 복원할 수 있다.

- [94] 한편, 이상에서는 영상과 움직임 벡터가 블록 단위로 예측 부호화되고 복호화되는 것으로 예를 들어 설명했지만, 영상과 움직임 벡터가 반드시 블록 단위로 예측 부호화되고 복호화되어야 하는 것은 아니다. 예를 들어, 블록 단위가 아닌, 슬라이스, 픽처, 시퀀스 단위와 같은 소정의 부호화 단위로 예측 부호화될 수도 있고 블록의 형태가 아닌 비정형의 다양한 영역의 형태로 예측 부호화될 수도 있다.
- [95] 이상에서 전술한 바와 같이, 본 발명의 일 실시예에 따르면, 이미 부호화되고 복호화되어 복원된 주변 블록의 움직임 정보를 이용하여 블록 단위, 슬라이스 단위, 픽처 단위, 시퀀스 단위 등 소정의 부호화 단위마다 적응적으로 해당 부호화 단위에 적합한 후보 예측 움직임 벡터 집합을 선택함으로써, 부호화하고자 하는 움직임 벡터와 더욱 유사한 예측 움직임 벡터를 선택하여 차분 벡터의 크기를 줄이면서도 선택한 후보 예측 움직임 벡터 집합에 대한 정보를 부호화할 필요가 없으므로, 움직임 벡터를 부호화하는 데 소요되는 비트량을 줄일 수 있으며 그에 따라 영상의 압축 효율을 향상시킬 수 있다.
- [96] 이상에서, 본 발명의 실시예를 구성하는 모든 구성 요소들이 하나로 결합하거나 결합하여 동작하는 것으로 설명되었다고 해서, 본 발명이 반드시 이러한 실시예에 한정되는 것은 아니다. 즉, 본 발명의 목적 범위 안에서라면, 그 모든 구성 요소들이 하나 이상으로 선택적으로 결합하여 동작할 수도 있다. 또한, 그 모든 구성 요소들이 각각 하나의 독립적인 하드웨어로 구현될 수 있지만, 각 구성 요소들의 그 일부 또는 전부가 선택적으로 조합되어 하나 또는 복수 개의 하드웨어에서 조합된 일부 또는 전부의 기능을 수행하는 프로그램 모듈을 갖는 컴퓨터 프로그램으로서 구현될 수도 있다. 그 컴퓨터 프로그램을 구성하는 코드들 및 코드 세그먼트들은 본 발명의 기술 분야의 당업자에 의해 용이하게 추론될 수 있을 것이다. 이러한 컴퓨터 프로그램은 컴퓨터가 읽을 수 있는 저장매체(Computer Readable Media)에 저장되어 컴퓨터에 의하여 읽혀지고 실행됨으로써, 본 발명의 실시예를 구현할 수 있다. 컴퓨터 프로그램의 저장매체로서는 자기 기록매체, 광 기록매체, 캐리어 웨이브 매체 등이 포함될 수 있다.
- [97] 또한, 이상에서 기재된 "포함하다", "구성하다" 또는 "가지다" 등의 용어는, 특별히 반대되는 기재가 없는 한, 해당 구성 요소가 내재할 수 있음을 의미하는 것이므로, 다른 구성 요소를 제외하는 것이 아니라 다른 구성 요소를 더 포함할 수 있는 것으로 해석되어야 한다. 기술적이거나 과학적인 용어를 포함한 모든 용어들은, 다르게 정의되지 않는 한, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미가 있다. 사전에

정의된 용어와 같이 일반적으로 사용되는 용어들은 관련 기술의 문맥상의 의미와 일치하는 것으로 해석되어야 하며, 본 발명에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.

- [98] 이상의 설명은 본 발명의 기술 사상을 예시적으로 설명한 것에 불과한 것으로서, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 다양한 수정 및 변형이 가능할 것이다. 따라서, 본 발명에 개시된 실시예들은 본 발명의 기술 사상을 한정하기 위한 것이 아니라 설명하기 위한 것이고, 이러한 실시예에 의하여 본 발명의 기술 사상의 범위가 한정되는 것은 아니다. 본 발명의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 발명의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

산업상 이용가능성

- [99] 이상에서 설명한 바와 같이 본 발명은 영상을 부호화하고 복호화하는 영상 압축 처리 분야에 적용되어, 효율적인 후보 예측 움직임 벡터 집합을 선택하여 움직임 벡터를 부호화함에 따라 부호화할 차분 벡터의 크기를 줄이면서도 어떠한 후보 예측 움직임 벡터 집합이 선택되었는지를 나타내기 위한 부가 정보를 부호화하지 않아도 되므로, 움직임 벡터의 압축 효율을 향상시켜 결과적으로 영상의 압축 효율을 향상시킬 수 있는 효과를 발생하는 매우 유용한 발명이다.

[100]

[101] CROSS-REFERENCE TO RELATED APPLICATION

- [102] 본 특허출원은 2009년 11월 18일 한국에 출원한 특허출원번호 제 10-2009-0111302 호에 대해 미국 특허법 119(a)조(35 U.S.C § 119(a))에 따라 우선권을 주장하면, 그 모든 내용은 참고문헌으로 본 특허출원에 병합된다. 아울러, 본 특허출원은 미국 이외에 국가에 대해서도 위와 동일한 동일한 이유로 우선권을 주장하면 그 모든 내용은 참고문헌으로 본 특허출원에 병합된다.

청구범위

[청구항 1]

움직임 벡터를 부호화하는 장치에 있어서,
 현재 블록의 주변 블록의 움직임 정보를 이용하여 복수 개의 후보
 예측 움직임 벡터 집합 중에서 하나의 후보 예측 움직임 벡터
 집합을 선택하는 후보 예측 움직임 벡터 집합 선택기;
 상기 선택되는 후보 예측 움직임 벡터 집합 내의 후보 예측 움직임
 벡터 중에서 하나의 후보 예측 움직임 벡터를 예측 움직임
 벡터로서 선택하는 예측 움직임 벡터 선택기;
 상기 현재 블록의 움직임 벡터인 현재 움직임 벡터와 상기
 선택되는 예측 움직임 벡터의 차이인 차분 벡터를 부호화하는
 차분 벡터 부호화기; 및
 상기 선택되는 예측 움직임 벡터를 나타내는 예측 움직임 벡터
 인덱스를 부호화하는 예측 움직임 벡터 부호화기
 를 포함하는 것을 특징으로 하는 움직임 벡터 부호화 장치.

[청구항 2]

영상을 부호화하는 장치에 있어서,
 현재 블록의 현재 움직임 벡터를 결정하고, 상기 결정되는 현재
 움직임 벡터를 이용하여 상기 현재 블록을 예측 부호화하는 영상
 부호화기; 및
 복수 개의 후보 예측 움직임 벡터 집합 중 상기 현재 블록의 주변
 블록의 움직임 정보를 기초로 선택되는 후보 예측 움직임 벡터
 집합에서 예측 움직임 벡터를 선택하고, 현재 움직임 벡터와 상기
 선택되는 예측 움직임 벡터의 차이인 차분 벡터 및 상기 선택되는
 예측 움직임 벡터를 나타내는 예측 움직임 벡터 인덱스를
 부호화하는 움직임 벡터 부호화기
 를 포함하는 것을 특징으로 하는 영상 부호화 장치.

[청구항 3]

움직임 벡터를 복호화하는 장치에 있어서,
 현재 블록의 주변 블록의 움직임 정보를 이용하여 복수 개의 후보
 예측 움직임 벡터 집합 중에서 하나의 후보 예측 움직임 벡터
 집합을 선택하는 후보 예측 움직임 벡터 집합 선택기;
 움직임 벡터 데이터로부터 추출되는 인덱스 데이터를 복호화하여
 예측 움직임 벡터 인덱스를 복원하고, 상기 선택되는 후보 예측
 움직임 벡터 집합에서 상기 복원되는 예측 움직임 벡터 인덱스에
 의해 식별되는 후보 예측 움직임 벡터를 예측 움직임 벡터로서
 복원하는 예측 움직임 벡터 복원기; 및
 상기 움직임 벡터 데이터로부터 추출되는 차분 벡터 데이터를
 복호화하여 차분 벡터를 복원하고, 상기 복원되는 차분 벡터와
 상기 복원되는 예측 움직임 벡터를 가산하여 상기 현재 블록의

- 현재 움직임 벡터를 복원하는 현재 움직임 벡터 복원기를 포함하는 것을 특징으로 하는 움직임 벡터 복호화 장치.
- [청구항 4] 영상을 복호화하는 장치에 있어서,
비트스트림으로부터 추출되는 움직임 벡터 데이터를 복호화하여 차분 벡터와 예측 움직임 벡터 인덱스를 복원하고, 현재 블록의 주변 블록의 움직임 정보를 이용하여 복수 개의 후보 예측 움직임 벡터 집합 중에서 하나의 후보 예측 움직임 벡터 집합을 선택하며, 상기 선택되는 후보 예측 움직임 벡터 집합에서 상기 복원되는 예측 움직임 벡터 인덱스에 의해 식별되는 후보 예측 움직임 벡터를 예측 움직임 벡터로서 선택하며, 상기 복원되는 차분 벡터와 상기 선택되는 예측 움직임 벡터를 가산하여 상기 현재 블록의 현재 움직임 벡터를 복원하는 움직임 벡터 복호화기; 및 상기 비트스트림으로부터 추출되는 영상 데이터를 상기 복원되는 현재 움직임 벡터를 이용하여 예측 복호화하여 상기 현재 블록을 복원하는 영상 복호화기를 포함하는 것을 특징으로 하는 영상 복호화 장치.
- [청구항 5] 움직임 벡터를 부호화하는 방법에 있어서,
현재 블록의 주변 블록의 움직임 정보를 이용하여 복수 개의 후보 예측 움직임 벡터 집합 중에서 하나의 후보 예측 움직임 벡터 집합을 선택하는 단계;
상기 선택되는 후보 예측 움직임 벡터 집합 내의 후보 예측 움직임 벡터 중에서 하나의 후보 예측 움직임 벡터를 예측 움직임 벡터로서 선택하는 단계;
상기 현재 블록의 움직임 벡터인 현재 움직임 벡터와 상기 선택되는 예측 움직임 벡터의 차이인 차분 벡터를 부호화하는 단계; 및
상기 선택되는 예측 움직임 벡터를 나타내는 예측 움직임 벡터 인덱스를 부호화하는 단계
를 포함하는 것을 특징으로 하는 움직임 벡터 부호화 방법.
- [청구항 6] 제 5 항에 있어서,
상기 하나의 후보 예측 움직임 벡터 집합을 선택하는 단계는,
상기 주변 블록의 움직임 정보를 이용하여 카메라의 움직임 여부를 판단하고 상기 판단되는 카메라의 움직임 여부를 기초로 상기 하나의 후보 예측 움직임 벡터 집합을 선택하는 것을 특징으로 하는 움직임 벡터 부호화 방법.
- [청구항 7] 제 5 항에 있어서,
상기 하나의 후보 예측 움직임 벡터 집합을 선택하는 단계는,
상기 주변 블록의 움직임 정보를 이용하여 상기 주변 블록의 객체

움직임을 판단하고, 상기 판단되는 주변 블록의 객체 움직임을 기초로 상기 하나의 후보 예측 움직임 벡터 집합을 선택하는 것을 특징으로 하는 움직임 벡터 부호화 방법.

[청구항 8]

제 7 항에 있어서,
상기 하나의 후보 예측 움직임 벡터 집합을 선택하는 단계는,
상기 주변 블록의 예측 움직임 벡터를 이용하여 상기 주변 블록의 객체 움직임을 판단하는 것을 특징으로 하는 움직임 벡터 부호화 방법.

[청구항 9]

제 5 항에 있어서,
상기 하나의 후보 예측 움직임 벡터 집합을 선택하는 단계는,
상기 주변 블록의 움직임 정보를 이용하여 상기 주변 블록의 객체 움직임을 판단하고 상기 판단되는 객체 움직임을 기초로 후보 예측 움직임 벡터를 선택하는 단계;
상기 주변 블록의 움직임 정보를 이용하여 카메라의 움직임 여부를 판단하고 상기 판단되는 카메라의 움직임 여부를 기초로 후보 예측 움직임 벡터를 선택하는 단계; 및
상기 복수 개의 후보 예측 움직임 벡터 집합 중에서 상기 객체 움직임을 기초로 선택되는 후보 예측 움직임 벡터 및 상기 카메라의 움직임 여부를 기초로 선택되는 후보 예측 움직임 벡터를 포함하는 후보 예측 움직임 벡터 집합을 상기 하나의 후보 예측 움직임 벡터 집합으로서 선택하는 단계
를 포함하는 것을 특징으로 하는 움직임 벡터 부호화 방법.

[청구항 10]

영상을 부호화하는 방법에 있어서,
현재 블록의 현재 움직임 벡터를 결정하는 단계;
상기 결정되는 현재 움직임 벡터를 이용하여 상기 현재 블록을 예측 부호화하는 단계;
복수 개의 후보 예측 움직임 벡터 집합 중 상기 현재 블록의 주변 블록의 움직임 정보를 기초로 선택되는 후보 예측 움직임 벡터 집합에서 예측 움직임 벡터를 선택하는 단계; 및
상기 현재 움직임 벡터와 상기 선택되는 예측 움직임 벡터의 차이인 차분 벡터와 상기 선택되는 예측 움직임 벡터를 나타내는 예측 움직임 벡터 인덱스를 부호화하는 단계
를 포함하는 것을 특징으로 하는 영상 부호화 방법.

[청구항 11]

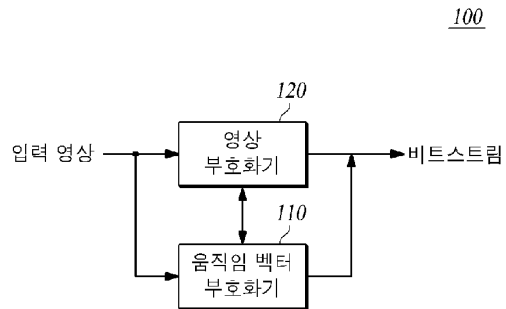
움직임 벡터를 복호화하는 방법에 있어서,
움직임 벡터 데이터로부터 추출되는 차분 벡터 데이터와 인덱스 데이터를 복호화하여 차분 벡터와 예측 움직임 벡터 인덱스를 복원하는 단계;
현재 블록의 주변 블록의 움직임 정보를 이용하여 복수 개의 후보

- 예측 움직임 벡터 집합 중에서 하나의 후보 예측 움직임 벡터 집합을 선택하는 단계;
 상기 선택되는 후보 예측 움직임 벡터 집합에서 상기 복원되는 예측 움직임 벡터 인덱스에 의해 식별되는 후보 예측 움직임 벡터를 예측 움직임 벡터로서 복원하는 단계; 및
 상기 복원되는 차분 벡터와 상기 복원되는 예측 움직임 벡터를 가산하여 상기 현재 블록의 현재 움직임 벡터를 복원하는 단계를 포함하는 것을 특징으로 하는 움직임 벡터 복호화 방법.
- [청구항 12] 제 11 항에 있어서,
 상기 하나의 후보 예측 움직임 벡터 집합을 선택하는 단계는,
 상기 주변 블록의 움직임 정보를 이용하여 카메라의 움직임 여부를 판단하고 상기 판단되는 카메라의 움직임 여부를 기초로 상기 하나의 후보 예측 움직임 벡터 집합을 선택하는 것을 특징으로 하는 움직임 벡터 복호화 방법.
- [청구항 13] 제 11 항에 있어서,
 상기 하나의 후보 예측 움직임 벡터 집합을 선택하는 단계는,
 상기 주변 블록의 움직임 정보를 이용하여 상기 주변 블록의 객체 움직임을 판단하고, 상기 판단되는 주변 블록의 객체 움직임을 기초로 상기 하나의 후보 예측 움직임 벡터 집합을 선택하는 것을 특징으로 하는 움직임 벡터 복호화 방법.
- [청구항 14] 제 13 항에 있어서,
 상기 하나의 후보 예측 움직임 벡터 집합을 선택하는 단계는,
 상기 주변 블록의 예측 움직임 벡터를 이용하여 상기 주변 블록의 객체 움직임을 판단하는 것을 특징으로 하는 움직임 벡터 복호화 방법.
- [청구항 15] 제 11 항에 있어서,
 상기 하나의 후보 예측 움직임 벡터 집합을 선택하는 단계는,
 상기 주변 블록의 움직임 정보를 이용하여 상기 주변 블록의 객체 움직임을 판단하고 상기 판단되는 객체 움직임을 기초로 후보 예측 움직임 벡터를 선택하는 단계;
 상기 주변 블록의 움직임 정보를 이용하여 카메라의 움직임 여부를 판단하고 상기 판단되는 카메라의 움직임 여부를 기초로 후보 예측 움직임 벡터를 선택하는 단계; 및
 상기 복수 개의 후보 예측 움직임 벡터 집합 중에서 상기 객체 움직임을 기초로 선택되는 후보 예측 움직임 벡터 및 상기 카메라의 움직임 여부를 기초로 선택되는 후보 예측 움직임 벡터를 포함하는 후보 예측 움직임 벡터 집합을 상기 하나의 후보 예측 움직임 벡터 집합으로서 선택하는 단계

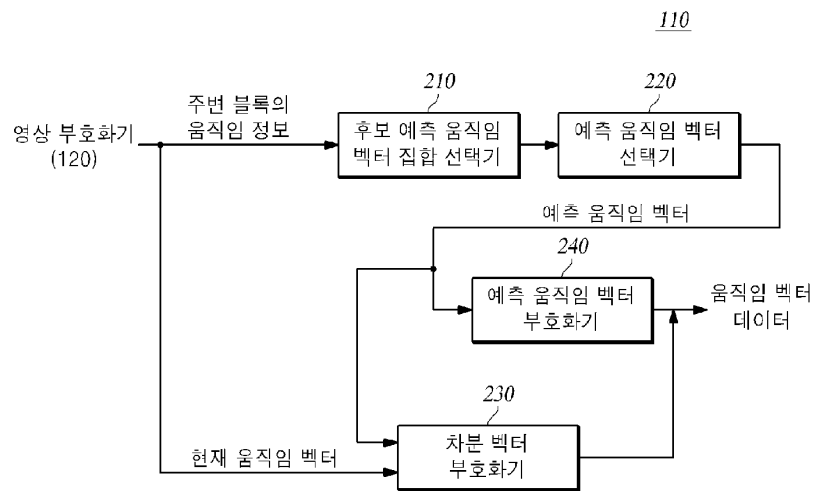
[청구항 16]

를 포함하는 것을 특징으로 하는 움직임 벡터 복호화 방법.
 영상을 복호화하는 방법에 있어서,
 비트스트림으로부터 추출되는 움직임 벡터 데이터를 복호화하여
 차분 벡터와 예측 움직임 벡터 인덱스를 복원하는 단계;
 현재 블록의 주변 블록의 움직임 정보를 이용하여 복수 개의 후보
 예측 움직임 벡터 집합 중에서 하나의 후보 예측 움직임 벡터
 집합을 선택하는 단계;
 상기 선택되는 후보 예측 움직임 벡터 집합에서 상기 복원되는
 예측 움직임 벡터 인덱스에 의해 식별되는 후보 예측 움직임
 벡터를 예측 움직임 벡터로서 선택하는 단계;
 상기 복원되는 차분 벡터와 상기 선택되는 예측 움직임 벡터를
 가산하여 상기 현재 블록의 현재 움직임 벡터를 복원하는 단계; 및
 상기 비트스트림으로부터 추출되는 영상 데이터를 상기 복원되는
 현재 움직임 벡터를 이용하여 예측 복호화하여 상기 현재 블록을
 복원하는 단계
 를 포함하는 것을 특징으로 하는 영상 복호화 방법.

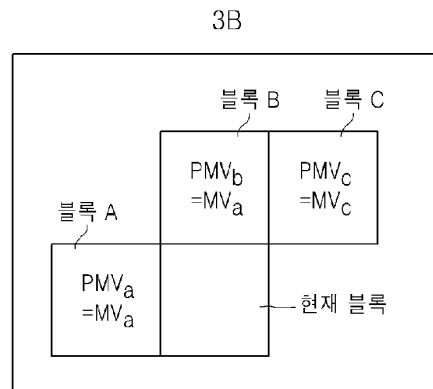
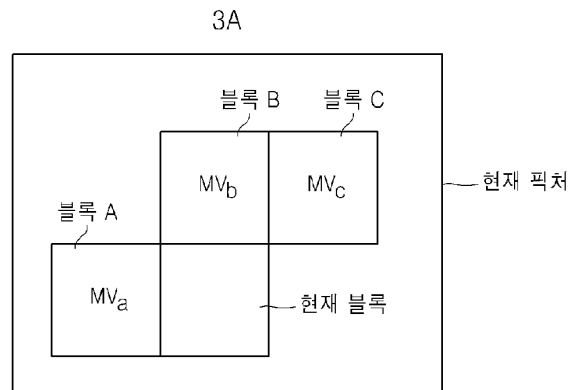
[Fig. 1]



[Fig. 2]



[Fig. 3]

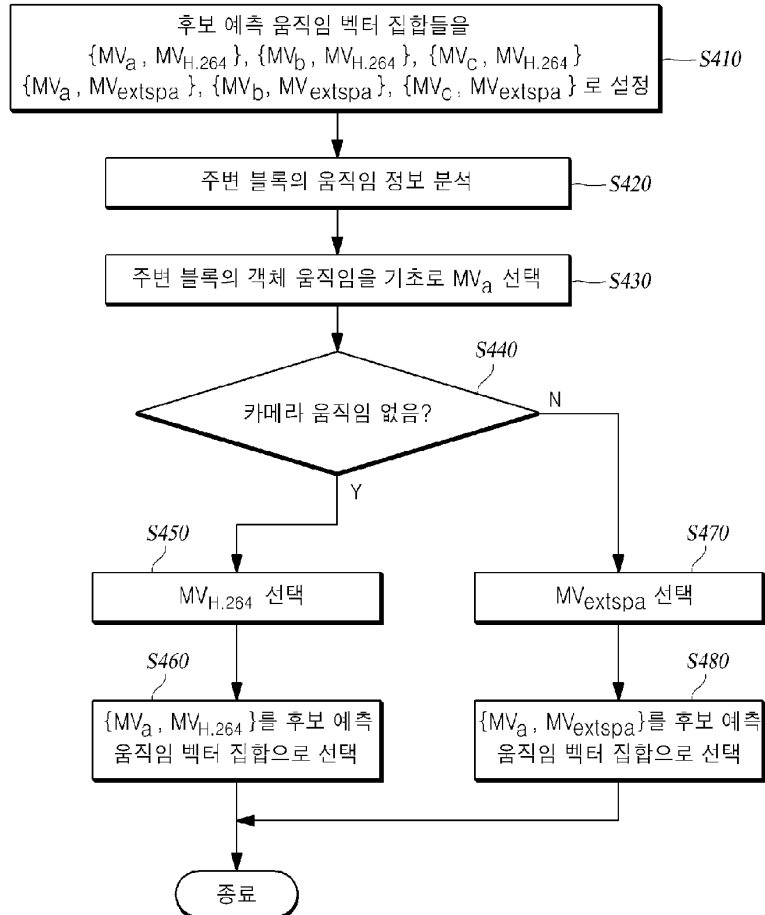


3C

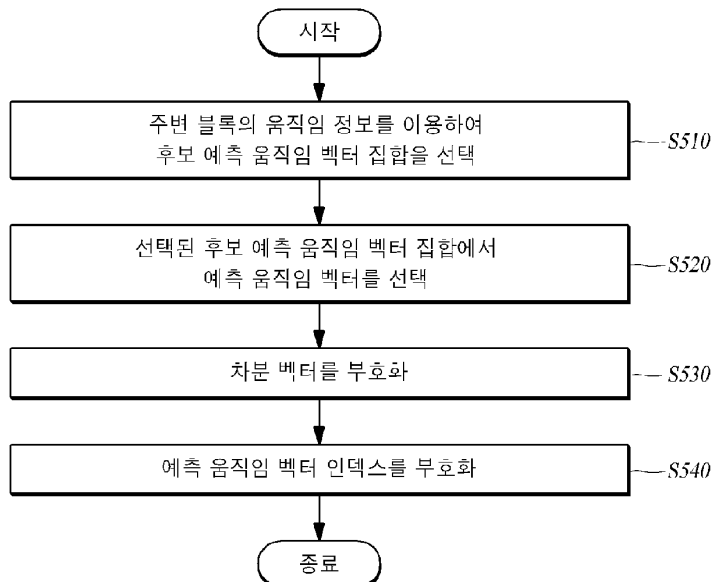
주변 블록의 움직임 정보 = $\{MV_a, MV_b, MV_c, PMV_a, PMV_b, PMV_c\}$

[Fig. 4]

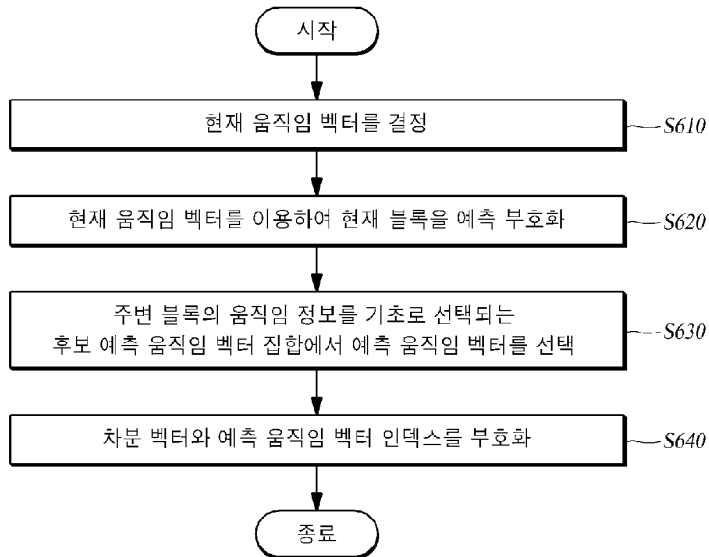
310



[Fig. 5]

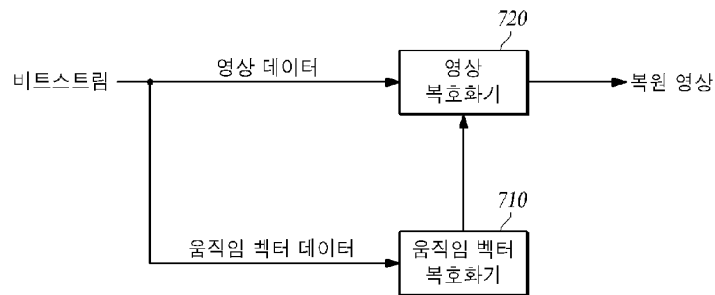


[Fig. 6]



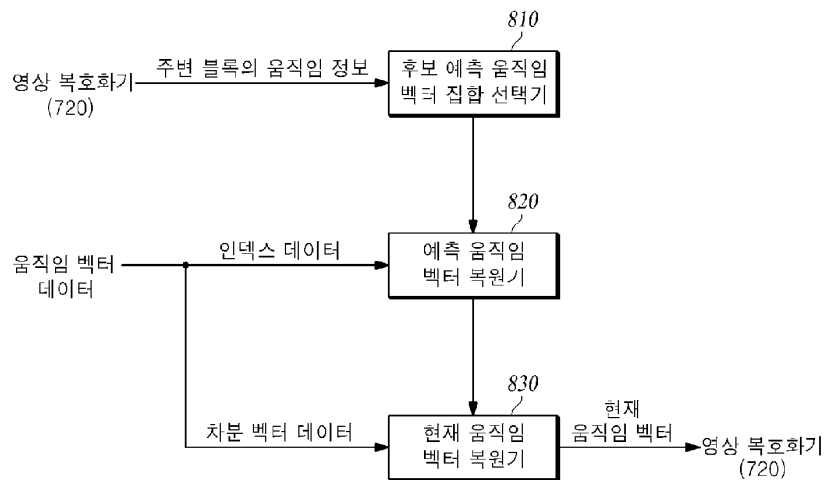
[Fig. 7]

700

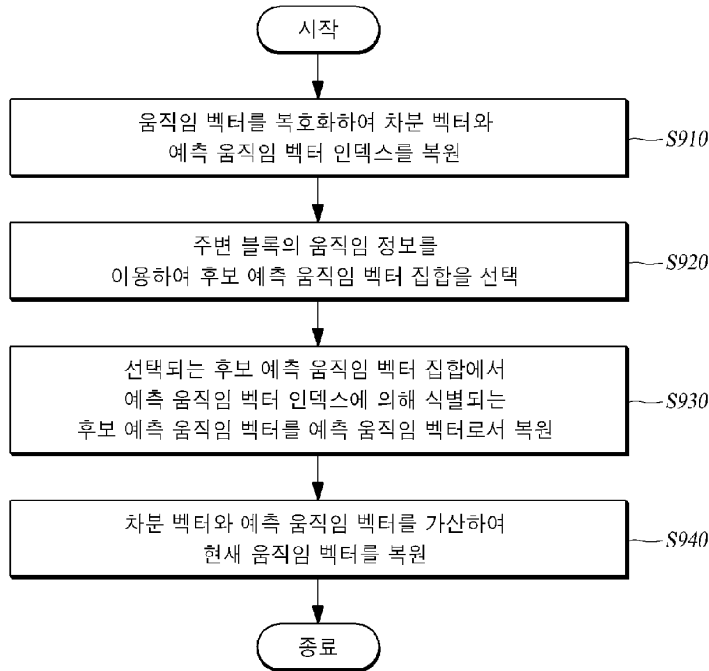


[Fig. 8]

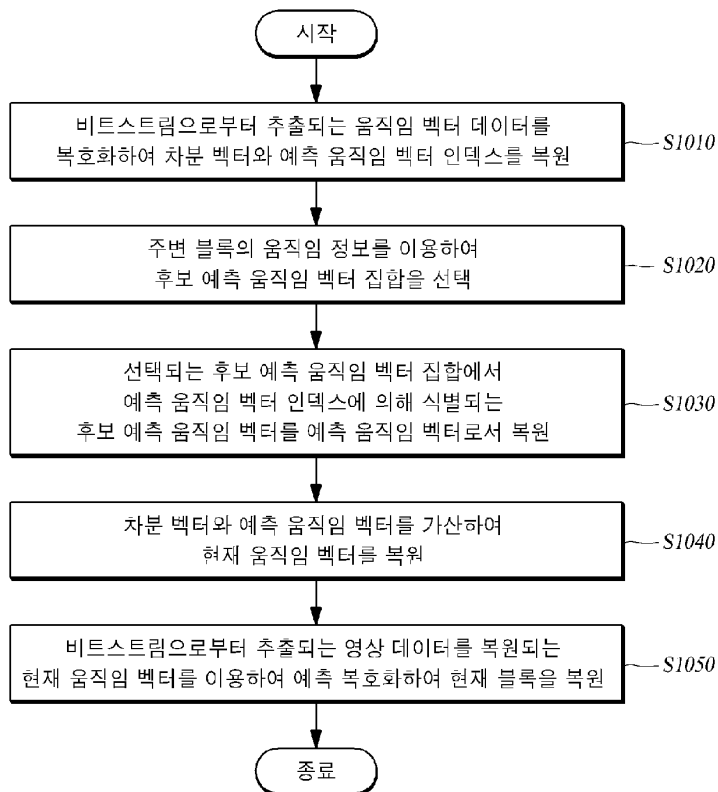
710



[Fig. 9]



[Fig. 10]



Electronic Acknowledgement Receipt	
EFS ID:	15579218
Application Number:	13666680
International Application Number:	
Confirmation Number:	4782
Title of Invention:	METHOD FOR CODING AND AN APPARATUS
First Named Inventor/Applicant Name:	Mehmet Oguz BICI
Customer Number:	73658
Filer:	Thomas Joseph Arria/thao pham
Filer Authorized By:	Thomas Joseph Arria
Attorney Docket Number:	NC77198US-NP
Receipt Date:	22-APR-2013
Filing Date:	01-NOV-2012
Time Stamp:	15:55:51
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no
------------------------	----

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Information Disclosure Statement (IDS) Form (SB08)	77198_IDS.pdf	612036 589f3026006778cd6803b0090fed1baa211799ae	no	4

Warnings:

Information:

2	Non Patent Literature	77198_ISR_WO.pdf	1918766 092c169f17cdb2416782db54eb72d5f1d2cefb53	no	12
Warnings:					
Information:					
3	Non Patent Literature	OUDIN.PDF	468422 315377bd759752c6fa371689a7c187886a4997a1	no	6
Warnings:					
Information:					
4	Non Patent Literature	SULLIVAN.PDF	1726067 7cc3203335537fbc186f548ac62e2b9ff5ee72f8	no	20
Warnings:					
Information:					
5	Foreign Reference	WO2011062392A2.pdf	3674943 751e69a655a915ff41fd86d434c8b5b122fa388a	no	28
Warnings:					
Information:					
Total Files Size (in bytes):			8400234		
This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503. <u>New Applications Under 35 U.S.C. 111</u> If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application. <u>National Stage of an International Application under 35 U.S.C. 371</u> If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course. <u>New International Application Filed with the USPTO as a Receiving Office</u> If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.					



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING OR 371(C) DATE	FIRST NAMED APPLICANT	ATTY. DOCKET NO./TITLE
13/666,680	11/01/2012	Mehmet Oguz BICI	NC77198US-NP

CONFIRMATION NO. 4782

73658

Nokia, Inc.

Attn: Intellectual Property Rights Docketing

200 South Mathilda Ave

Sunnyvale, CA 94086

PUBLICATION NOTICE



Title:METHOD FOR CODING AND AN APPARATUS

Publication No.US-2013-0114723-A1

Publication Date:05/09/2013

NOTICE OF PUBLICATION OF APPLICATION

The above-identified application will be electronically published as a patent application publication pursuant to 37 CFR 1.211, et seq. The patent application publication number and publication date are set forth above.

The publication may be accessed through the USPTO's publically available Searchable Databases via the Internet at www.uspto.gov. The direct link to access the publication is currently <http://www.uspto.gov/patft/>.

The publication process established by the Office does not provide for mailing a copy of the publication to applicant. A copy of the publication may be obtained from the Office upon payment of the appropriate fee set forth in 37 CFR 1.19(a)(1). Orders for copies of patent application publications are handled by the USPTO's Office of Public Records. The Office of Public Records can be reached by telephone at (703) 308-9726 or (800) 972-6382, by facsimile at (703) 305-8759, by mail addressed to the United States Patent and Trademark Office, Office of Public Records, Alexandria, VA 22313-1450 or via the Internet.

In addition, information on the status of the application, including the mailing date of Office actions and the dates of receipt of correspondence filed in the Office, may also be accessed via the Internet through the Patent Electronic Business Center at www.uspto.gov using the public side of the Patent Application Information and Retrieval (PAIR) system. The direct link to access this status information is currently <http://pair.uspto.gov/>. Prior to publication, such status information is confidential and may only be obtained by applicant using the private side of PAIR.

Further assistance in electronically accessing the publication, or about PAIR, is available by calling the Patent Electronic Business Center at 1-866-217-9197.

Office of Data Management, Application Assistance Unit (571) 272-4000, or (571) 272-4200, or 1-888-786-0101

To: brian.odonoghue@nokia.com,thao.pham@nokia.com,sandy.fong-hou@nokia.com
From: PAIR_eOfficeAction@uspto.gov
Cc: PAIR_eOfficeAction@uspto.gov
Subject: Private PAIR Correspondence Notification for Customer Number 73658

May 10, 2013 05:33:24 AM

Dear PAIR Customer:

Nokia, Inc.
Attn: Intellectual Property Rights Docketing
200 South Mathilda Ave
Sunnyvale, CA 94086
UNITED STATES

The following USPTO patent application(s) associated with your Customer Number, 73658 , have new outgoing correspondence. This correspondence is now available for viewing in Private PAIR.

The official date of notification of the outgoing correspondence will be indicated on the form PTOL-90 accompanying the correspondence.

Disclaimer:

The list of documents shown below is provided as a courtesy and is not part of the official file wrapper. The content of the images shown in PAIR is the official record.

Application	Document	Mailroom Date	Attorney Docket No.
13666680	NTC.PUB	05/09/2013	NC77198US-NP

To view your correspondence online or update your email addresses, please visit us anytime at <https://sportal.uspto.gov/secure/myportal/privatepair>.

If you have any questions, please email the Electronic Business Center (EBC) at EBC@uspto.gov with 'e-Office Action' on the subject line or call 1-866-217-9197 during the following hours:

Monday - Friday 6:00 a.m. to 12:00 a.m.

Thank you for prompt attention to this notice,

UNITED STATES PATENT AND TRADEMARK OFFICE
PATENT APPLICATION INFORMATION RETRIEVAL SYSTEM

PATENT APPLICATION FEE DETERMINATION RECORD Substitute for Form PTO-875						Application or Docket Number 13/666,680				
APPLICATION AS FILED - PART I										
(Column 1)		(Column 2)		SMALL ENTITY		OR OTHER THAN SMALL ENTITY				
FOR	NUMBER FILED	NUMBER EXTRA	RATE(\$)	FEE(\$)		RATE(\$)	FEE(\$)			
BASIC FEE (37 CFR 1.16(a), (b), or (c))	N/A	N/A	N/A			N/A	280			
SEARCH FEE (37 CFR 1.16(k), (l), or (m))	N/A	N/A	N/A			N/A	600			
EXAMINATION FEE (37 CFR 1.16(o), (p), or (q))	N/A	N/A	N/A			N/A	720			
TOTAL CLAIMS (37 CFR 1.16(j))	20	minus 20 = *				x 80 =	0.00			
INDEPENDENT CLAIMS (37 CFR 1.16(h))	8	minus 3 = * 5				x 420 =	2100			
APPLICATION SIZE FEE (37 CFR 1.16(s))	If the specification and drawings exceed 100 sheets of paper, the application size fee due is \$310 (\$155 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).						0.00			
MULTIPLE DEPENDENT CLAIM PRESENT (37 CFR 1.16(j))							0.00			
* If the difference in column 1 is less than zero, enter "0" in column 2.			TOTAL			TOTAL	3700			
APPLICATION AS AMENDED - PART II										
(Column 1)		(Column 2)		(Column 3)		SMALL ENTITY		OR OTHER THAN SMALL ENTITY		
AMENDMENT A		CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA	RATE(\$)	ADDITIONAL FEE(\$)		RATE(\$)	ADDITIONAL FEE(\$)
	Total (37 CFR 1.16(i))	*	Minus	**	=	x	=		x	=
	Independent (37 CFR 1.16(h))	*	Minus	***	=	x	=		x	=
	Application Size Fee (37 CFR 1.16(s))									
	FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))									
					TOTAL ADD'L FEE				TOTAL ADD'L FEE	
AMENDMENT B		CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA	RATE(\$)	ADDITIONAL FEE(\$)		RATE(\$)	ADDITIONAL FEE(\$)
	Total (37 CFR 1.16(i))	*	Minus	**	=	x	=		x	=
	Independent (37 CFR 1.16(h))	*	Minus	***	=	x	=		x	=
	Application Size Fee (37 CFR 1.16(s))									
	FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))									
					TOTAL ADD'L FEE				TOTAL ADD'L FEE	
* If the entry in column 1 is less than the entry in column 2, write "0" in column 3. ** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20". *** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3". The "Highest Number Previously Paid For" (Total or Independent) is the highest found in the appropriate box in column 1.										



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING or 371(c) DATE	GRP ART UNIT	FIL FEE REC'D	ATTY. DOCKET NO.	TOT CLAIMS	IND CLAIMS
13/666,680	11/01/2012	2631	2640	NC77198US-NP	20	8

CONFIRMATION NO. 4782

UPDATED FILING RECEIPT



OC000000062918875

73658

Nokia, Inc.

Attn: Intellectual Property Rights Docketing

200 South Mathilda Ave

Sunnyvale, CA 94086

Date Mailed: 08/05/2013

Receipt is acknowledged of this non-provisional patent application. The application will be taken up for examination in due course. Applicant will be notified as to the results of the examination. Any correspondence concerning the application must include the following identification information: the U.S. APPLICATION NUMBER, FILING DATE, NAME OF APPLICANT, and TITLE OF INVENTION. Fees transmitted by check or draft are subject to collection. Please verify the accuracy of the data presented on this receipt. **If an error is noted on this Filing Receipt, please submit a written request for a Filing Receipt Correction. Please provide a copy of this Filing Receipt with the changes noted thereon. If you received a "Notice to File Missing Parts" for this application, please submit any corrections to this Filing Receipt with your reply to the Notice. When the USPTO processes the reply to the Notice, the USPTO will generate another Filing Receipt incorporating the requested corrections**

Inventor(s)

Mehmet Oguz BICI, Tampere, TURKEY;
Jani LAINEMA, Tampere, FINLAND;
Kemal UGUR, Tampere, FINLAND;

Applicant(s)

Mehmet Oguz BICI, Tampere, TURKEY;
Jani LAINEMA, Tampere, FINLAND;
Kemal UGUR, Tampere, FINLAND;

Non-Applicant Assignee(s)

NOKIA CORPORATION, Espoo, FINLAND

Power of Attorney: None

Domestic Priority data as claimed by applicant

This appln claims benefit of 61/555,703 11/04/2011

Foreign Applications for which priority is claimed (You may be eligible to benefit from the **Patent Prosecution Highway** program at the USPTO. Please see <http://www.uspto.gov> for more information.) - None.

Foreign application information must be provided in an Application Data Sheet in order to constitute a claim to foreign priority. See 37 CFR 1.55 and 1.76.

If Required, Foreign Filing License Granted: 11/20/2012

The country code and number of your priority application, to be used for filing abroad under the Paris Convention, is **US 13/666,680**

Projected Publication Date: Not Applicable

Non-Publication Request: No

Early Publication Request: No
Title

METHOD FOR CODING AND AN APPARATUS

Preliminary Class

375

Statement under 37 CFR 1.55 or 1.78 for AIA (First Inventor to File) Transition Applications:

PROTECTING YOUR INVENTION OUTSIDE THE UNITED STATES

Since the rights granted by a U.S. patent extend only throughout the territory of the United States and have no effect in a foreign country, an inventor who wishes patent protection in another country must apply for a patent in a specific country or in regional patent offices. Applicants may wish to consider the filing of an international application under the Patent Cooperation Treaty (PCT). An international (PCT) application generally has the same effect as a regular national patent application in each PCT-member country. The PCT process **simplifies** the filing of patent applications on the same invention in member countries, but **does not result** in a grant of "an international patent" and does not eliminate the need of applicants to file additional documents and fees in countries where patent protection is desired.

Almost every country has its own patent law, and a person desiring a patent in a particular country must make an application for patent in that country in accordance with its particular laws. Since the laws of many countries differ in various respects from the patent law of the United States, applicants are advised to seek guidance from specific foreign countries to ensure that patent rights are not lost prematurely.

Applicants also are advised that in the case of inventions made in the United States, the Director of the USPTO must issue a license before applicants can apply for a patent in a foreign country. The filing of a U.S. patent application serves as a request for a foreign filing license. The application's filing receipt contains further information and guidance as to the status of applicant's license for foreign filing.

Applicants may wish to consult the USPTO booklet, "General Information Concerning Patents" (specifically, the section entitled "Treaties and Foreign Patents") for more information on timeframes and deadlines for filing foreign patent applications. The guide is available either by contacting the USPTO Contact Center at 800-786-9199, or it can be viewed on the USPTO website at <http://www.uspto.gov/web/offices/pac/doc/general/index.html>.

For information on preventing theft of your intellectual property (patents, trademarks and copyrights), you may wish to consult the U.S. Government website, <http://www.stopfakes.gov>. Part of a Department of Commerce initiative, this website includes self-help "toolkits" giving innovators guidance on how to protect intellectual property in specific countries such as China, Korea and Mexico. For questions regarding patent enforcement issues, applicants may call the U.S. Government hotline at 1-866-999-HALT (1-866-999-4258).

LICENSE FOR FOREIGN FILING UNDER
Title 35, United States Code, Section 184
Title 37, Code of Federal Regulations, 5.11 & 5.15

GRANTED

The applicant has been granted a license under 35 U.S.C. 184, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" followed by a date appears on this form. Such licenses are issued in all applications where the conditions for issuance of a license have been met, regardless of whether or not a license may be required as set forth in 37 CFR 5.15. The scope and limitations of this license are set forth in 37 CFR 5.15(a) unless an earlier license has been issued under 37 CFR 5.15(b). The license is subject to revocation upon written notification. The date indicated is the effective date of the license, unless an earlier license of similar scope has been granted under 37 CFR 5.13 or 5.14.

This license is to be retained by the licensee and may be used at any time on or after the effective date thereof unless it is revoked. This license is automatically transferred to any related applications(s) filed under 37 CFR 1.53(d). This license is not retroactive.

The grant of a license does not in any way lessen the responsibility of a licensee for the security of the subject matter as imposed by any Government contract or the provisions of existing laws relating to espionage and the national security or the export of technical data. Licensees should apprise themselves of current regulations especially with respect to certain countries, of other agencies, particularly the Office of Defense Trade Controls, Department of State (with respect to Arms, Munitions and Implements of War (22 CFR 121-128)); the Bureau of Industry and Security, Department of Commerce (15 CFR parts 730-774); the Office of Foreign Assets Control, Department of Treasury (31 CFR Parts 500+) and the Department of Energy.

NOT GRANTED

No license under 35 U.S.C. 184 has been granted at this time, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" DOES NOT appear on this form. Applicant may still petition for a license under 37 CFR 5.12, if a license is desired before the expiration of 6 months from the filing date of the application. If 6 months has lapsed from the filing date of this application and the licensee has not received any indication of a secrecy order under 35 U.S.C. 181, the licensee may foreign file the application pursuant to 37 CFR 5.15(b).

SelectUSA

The United States represents the largest, most dynamic marketplace in the world and is an unparalleled location for business investment, innovation, and commercialization of new technologies. The U.S. offers tremendous resources and advantages for those who invest and manufacture goods here. Through SelectUSA, our nation works to promote and facilitate business investment. SelectUSA provides information assistance to the international investor community; serves as an ombudsman for existing and potential investors; advocates on behalf of U.S. cities, states, and regions competing for global investment; and counsels U.S. economic development organizations on investment attraction best practices. To learn more about why the United States is the best country in the world to develop technology, manufacture products, deliver services, and grow your business, visit <http://www.SelectUSA.gov> or call +1-202-482-6800.



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING OR 371(C) DATE	FIRST NAMED APPLICANT	ATTY. DOCKET NO./TITLE
13/666,680	11/01/2012	Mehmet Oguz BICI	NC77198US-NP

CONFIRMATION NO. 4782

IMPROPER CFR REQUEST

73658

Nokia, Inc.

Attn: Intellectual Property Rights Docketing

200 South Mathilda Ave

Sunnyvale, CA 94086



Date Mailed: 08/05/2013

RESPONSE TO REQUEST FOR CORRECTED FILING RECEIPT

Power of Attorney, Claims, Fees, System Limitations, and Miscellaneous

In response to your request for a corrected Filing Receipt, the Office is unable to comply with your request because:

- Any request to correct or update the name of the applicant must include an application data sheet (ADS) in compliance with 37 CFR 1.76 specifying the correct or updated name of the applicant in the applicant information section. Any request to change the applicant after an original applicant has been specified under 37 CFR 1.46(b) must include a new ADS in compliance with 37 CFR 1.76 specifying the applicant in the applicant information section and comply with 37 CFR 3.71 and 3.73. See 37 CFR 1.46(c).

/zretta/

Office of Data Management, Application Assistance Unit (571) 272-4000, or (571) 272-4200, or 1-888-786-0101



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING OR 371(C) DATE	FIRST NAMED APPLICANT	ATTY. DOCKET NO./TITLE
13/666,680	11/01/2012	Mehmet Oguz BICI	NC77198US-NP

CONFIRMATION NO. 4782
IMPROPER CPOA LETTER

73658
Nokia, Inc.
Attn: Intellectual Property Rights Docketing
200 South Mathilda Ave
Sunnyvale, CA 94086



Date Mailed: 08/05/2013

NOTICE REGARDING POWER OF ATTORNEY

This is in response to the power of attorney filed 01/02/2013. The power of attorney in this application is not accepted for the reason(s) listed below:

- The power of attorney has not been accepted because the party who is giving power of attorney has not been identified. Power of attorney may only be signed by the applicant for patent (37 CFR 1.42) or the patent owner. A patent owner who was not the applicant must appoint any power of attorney in compliance with 37 CFR 3.71 and 3.73. See 37 CFR 1.32(b)(4).

/zretta/

Office of Data Management, Application Assistance Unit (571) 272-4000, or (571) 272-4200, or 1-888-786-0101

To: brian.odonoghue@nokia.com,thao.pham@nokia.com,sandy.fong-hou@nokia.com
From: PAIR_eOfficeAction@uspto.gov
Cc: PAIR_eOfficeAction@uspto.gov
Subject: Private PAIR Correspondence Notification for Customer Number 73658

Aug 05, 2013 05:31:06 AM

Dear PAIR Customer:

Nokia, Inc.
Attn: Intellectual Property Rights Docketing
200 South Mathilda Ave
Sunnyvale, CA 94086
UNITED STATES

The following USPTO patent application(s) associated with your Customer Number, 73658 , have new outgoing correspondence. This correspondence is now available for viewing in Private PAIR.

The official date of notification of the outgoing correspondence will be indicated on the form PTOL-90 accompanying the correspondence.

Disclaimer:

The list of documents shown below is provided as a courtesy and is not part of the official file wrapper. The content of the images shown in PAIR is the official record.

Application	Document	Mailroom Date	Attorney Docket No.
13666680	APP.FILE.REC	08/05/2013	NC77198US-NP
	M327	08/05/2013	NC77198US-NP
	N572	08/05/2013	NC77198US-NP

To view your correspondence online or update your email addresses, please visit us anytime at <https://sportal.uspto.gov/secure/myportal/privatepair>.

If you have any questions, please email the Electronic Business Center (EBC) at EBC@uspto.gov with 'e-Office Action' on the subject line or call 1-866-217-9197 during the following hours:

Monday - Friday 6:00 a.m. to 12:00 a.m.

Thank you for prompt attention to this notice,

UNITED STATES PATENT AND TRADEMARK OFFICE
PATENT APPLICATION INFORMATION RETRIEVAL SYSTEM

PLUS Search Results for S/N 13666680, Searched Tue May 26 15:06:59 EDT 2015
The Patent Linguistics Utility System (PLUS) is a USPTO automated search system for U.S. Patents from 1971 to the present PLUS is a query-by-example search system which produces a list of patents that are most closely related linguistically to the application searched. This search was prepared by the staff of the Scientific and Technical Information Center, SIRA.

5719627 99	6205177 98
5815602 99	
4989089 99	
5428403 99	
5781249 99	
5838391 99	
4851906 98	
4989087 98	
5235419 98	
5371549 98	
5414469 98	
5457481 98	
5506622 98	
5521642 98	
5532746 98	
5546129 98	
5557341 98	
5574663 98	
5579050 98	
5581308 98	
5587741 98	
5598216 98	
5612743 98	
5614954 98	
5617144 98	
5619281 98	
5625417 98	
5627591 98	
5638129 98	
5642166 98	
5646867 98	
5654761 98	
5689306 98	
5694487 98	
5717463 98	
5719630 98	
5724369 98	
5731851 98	
5751362 98	
5862261 98	
5903672 98	
5991447 98	
6005980 98	
6020925 98	
6023298 98	
6026195 98	
6057884 98	
6091777 98	
6154519 98	



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
13/666,680	11/01/2012	Mehmet Oguz BICI	NC77198US-NP	4782
73658 7590 07/06/2015 Nokia USA Inc. Attn: Intellectual Property Rights Docketing 200 South Mathilda Ave Sunnyvale, CA 94086			EXAMINER HILAIRE, CLIFFORD	
			ART UNIT 2488	PAPER NUMBER
			NOTIFICATION DATE 07/06/2015	DELIVERY MODE ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

thao.pham@nokia.com
sandy.fong-hou@nokia.com
jacqueline.linebarger@nokia.com

Office Action Summary	Application No. 13/666,680	Applicant(s) BICI ET AL.	
	Examiner CLIFFORD HILAIRE	Art Unit 2488	AIA (First Inventor to File) Status No

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTHS FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 11/01/2012.
☐ A declaration(s)/affidavit(s) under **37 CFR 1.130(b)** was/were filed on ____.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ An election was made by the applicant in response to a restriction requirement set forth during the interview on ____; the restriction requirement and election have been incorporated into this action.
- 4) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims*

- 5) ☒ Claim(s) 1-20 is/are pending in the application.
5a) Of the above claim(s) ____ is/are withdrawn from consideration.
- 6) ☐ Claim(s) ____ is/are allowed.
- 7) ☒ Claim(s) 1-20 is/are rejected.
- 8) ☐ Claim(s) ____ is/are objected to.
- 9) ☐ Claim(s) ____ are subject to restriction and/or election requirement.

* If any claims have been determined allowable, you may be eligible to benefit from the **Patent Prosecution Highway** program at a participating intellectual property office for the corresponding application. For more information, please see http://www.uspto.gov/patents/init_events/pph/index.jsp or send an inquiry to PPHfeedback@uspto.gov.

Application Papers

- 10) ☐ The specification is objected to by the Examiner.
- 11) ☐ The drawing(s) filed on ____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

Certified copies:

- a) ☐ All b) ☐ Some** c) ☐ None of the:
1. ☐ Certified copies of the priority documents have been received.
 2. ☐ Certified copies of the priority documents have been received in Application No. ____.
 3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

** See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- | | |
|---|---|
| 1) ☒ Notice of References Cited (PTO-892) | 3) ☐ Interview Summary (PTO-413) |
| 2) ☒ Information Disclosure Statement(s) (PTO/SB/08a and/or PTO/SB/08b) | Paper No(s)/Mail Date. ____ |
| Paper No(s)/Mail Date <u>04/22/2013</u> | 4) ☐ Other: ____ |

DETAILED ACTION

Notice of Pre-AIA or AIA Status

1. The present application is being examined under the pre-AIA first to invent provisions.

Claim Rejections - 35 USC § 112

The following is a quotation of 35 U.S.C. 112(b):

(b) CONCLUSION.—The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the inventor or a joint inventor regards as the invention.

The following is a quotation of 35 U.S.C. 112 (pre-AIA), second paragraph:

The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

2. Claims 1, 6, 9, 13, 15, 16, 17, 18, 19 and 20 are rejected under 35 U.S.C. 112(b) or 35 U.S.C. 112 (pre-AIA), second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which the inventor or a joint inventor, or for pre-AIA the applicant regards as the invention.
3. Claim 6 and 13 recite the limitation "*the maximum number*" in the first limitation of both claims. There is insufficient antecedent basis for this limitation in the claim.
4. Claim 1, 9, 15, 16, 17, 18, 19 and 20 recite the limitation "*the determined subset of spatial motion vector prediction candidates*" in the next to last limitation of each claims (*i.e.* "*comparing motion information...*" for claims 1,9,15,17,18; "*compare motion*

*information...” for claim 16 and “means for comparing motion information...” for claims 19 and 20). There is insufficient antecedent basis for this limitation in the claim. For examining purposes, it will be assumed the limitation “*the determined subset of spatial motion vector prediction candidates*” is obtain from the limitation above it (*i.e.* “*determining a subset of spatial motion vector **prediction candidates***” instead of “*determining a subset of spatial motion vector **predictions***”).*

Claim Rejections - 35 USC § 101

35 U.S.C. 101 reads as follows:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

35 U.S.C. 101 requires that a claimed invention must fall within one of the four eligible categories of invention (*i.e.* process, machine, manufacture, or composition of matter) *and* must not be directed to subject matter encompassing a judicially recognized exception as interpreted by the courts. MPEP 2106. The four eligible categories of invention include: (1) process which is an act, or a series of acts or steps, (2) machine which is a concrete thing, consisting of parts, or of certain devices and combination of devices, (3) manufacture which is an article produced from raw or prepared materials by giving to these materials new forms, qualities, properties, or combinations, whether by hand labor or by machinery, and (4) composition of matter which is all compositions of two or more substances and all composite articles, whether they be the results of

chemical union, or of mechanical mixture, or whether they be gases, fluids, powders or solids. MPEP 2106(I).

5. Claims 17 and 18 are rejected under 35 U.S.C. 101 as not falling within one of the four statutory categories of invention because the broadest reasonable interpretation of the instant claims in light of the specification encompasses transitory signals. But, transitory signals are not within one of the four statutory categories (i.e. non-statutory subject matter). See MPEP 2106(I). However, claims directed toward a non-transitory computer readable medium may qualify as a manufacture and make the claim patent-eligible subject matter. MPEP 2106(I). Therefore, amending the claims to recite a “*non-transitory computer-readable medium*” would resolve this issue.

Claim Rejections - 35 USC § 103

The following is a quotation of pre-AIA 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

6. Claims 1, 2, 3, 4, 5, 6, 7, 9 10, 11, 12, 14, 15, 16, 17, 18, 19 and 20 rejected under pre-AIA 35 U.S.C. 103 (a) as being unpatentable over Woo-Jin Han et al.

["Improved Video Compression Efficiency Through Flexible Unit Representation and Corresponding Extension of Coding Tools"] in view of Shen-Chuan Tai et al.["A Multi-Pass True Motion Estimation Scheme With Motion Vector Propagation for Frame Rate Up-Conversion Applications"].

Regarding claim 1, Han teaches

1. A method comprising *(i.e. a novel video compression scheme based on a highly flexible hierarchy of unit representation which includes three block concepts: coding unit (CU), prediction unit (PU), and transform unit (TU) - Abstract):*

receiving a block of pixels including a prediction unit (i.e. Fig. 1 shows all proposed building blocks of the decoder. All the coding tools are extended to support the proposed flexible architecture. Intra prediction is extended to support an arbitrary number of angles rather than the conventional 9-modes defined in H.264/AVC. Fast integer transforms larger than 8×8 are developed to support large TU sizes. Edge definition for the deblocking filter is modified according to the CU, PU, and TU concepts. Partition information in quadtreebased adaptive loop filter (QALF) [13] is replaced with CU splitting information- page 1710, ¶ 3);

However, Han does not teach explicitly:

determining a set of spatial motion vector prediction candidates for the block of pixels;

the spatial motion vector prediction candidates being provided with motion information;

selecting a first spatial motion vector prediction candidate from the set of spatial motion vector prediction candidates as a potential spatial motion vector prediction candidate to be included in a merge list for the prediction unit;

determining a subset of spatial motion vector predictions based on the location of the block associated with the first spatial motion vector prediction candidate;

comparing motion information of the first spatial motion vector prediction candidate with motion information of the spatial motion vector prediction candidate in the determined subset of spatial motion vector prediction candidates;

if at least one of the comparisons indicates that the motion vector information of the spatial motion vector prediction candidates correspond with each other, excluding the first spatial motion vector prediction candidate from the merge list.

In the same field of endeavor, Tai teaches:

determining a set (*i.e.* $CS^i(\vec{X}, t)$) of spatial motion vector prediction candidates for the block of pixels (*i.e.* *The MV search in block in the i th pass is performed in accordance with (1). As shown, the search is confined to the vectors defined in the candidate vector set $CS^i(\vec{X}, t)$, i.e., a full search is not performed-* Page 189, Col 2, ¶ 3);

the spatial motion vector prediction candidates being provided with motion information (*i.e.* *The basic principle of the multi-pass ME strategy proposed in*

this study is to utilize the motion information relating to neighboring blocks generated in the previous pass to gradually refine the accuracy of the MV in the current block- Page 189, Col 2, ¶ 3);

selecting a first spatial motion vector prediction (*i.e.* $CS^i(\vec{X}, t) = \{\vec{V} | \vec{V} = \overrightarrow{PMV}^i(\vec{X}) + \vec{U}, \vec{U} \in US^i\}$ - equation 2) candidate (*i.e.* $\overrightarrow{PMV}^i(\vec{X})$, predictive motion vectors (PMVs)- Page 190, Col 1, ¶ 2) from the set of spatial motion vector prediction candidates (*i.e.* $\overrightarrow{PMV}^i(\vec{X})$, is the PMV of block and is chosen from the MVs of the neighboring blocks and is a set of update vectors- Page 190, Col 1, ¶ 2) as a potential spatial motion vector prediction candidate to be included in a merge list for the (*i.e.* current block) prediction unit (*i.e.* Although the MVs of all eight neighboring blocks can theoretically be used as PMV predictors for the current block- Page 190, Col 1, ¶ 3);

determining a subset (*i.e.* $CS^i(\vec{X}, t) = \{\vec{V} | \vec{V} = \overrightarrow{PMV}^i(\vec{X}) + \vec{U}, \vec{U} \in US^i\}$ - Equation 2) of spatial motion vector predictions based on the location of the block associated with the first spatial motion vector prediction candidate (*i.e.* by selecting one of the $\overrightarrow{PMV}$ from MVs of neighboring blocks, a subset of $CS^i(\vec{X}, t)$ is defined for different linear combination of that PMV and $\vec{U} \in US^i$);

comparing motion information of the first spatial motion vector prediction candidate (*i.e.* $\overrightarrow{MV}^i(\vec{X}, t) = \arg_{\vec{V} \in CS^i(\vec{X}, t)} \min(e(\vec{V}, \vec{X}, t))$, $1 \leq i \leq 12$; equation 1) with motion information of the spatial motion vector prediction candidate in the determined subset of spatial motion vector prediction candidates (*i.e.* Each pass

in the proposed ME scheme is performed using one of the PMV predictor patterns (see Fig. 3) and then one MV in the selected PMV pattern which has minimal distortion is selected as the final PMV of this pass- Page 190; Col 1, ¶ 5);
if at least one of the comparisons indicates that the motion vector information of the spatial motion vector prediction candidates correspond with each other (*i.e. all MV's with non-minimal distortion as inferred above*), excluding the first spatial motion vector prediction candidate from the merge list (*i.e. all MV's with non-minimal distortion as inferred above will not be a final PMV for that pass*).

It would have been obvious to one with ordinary skill in the art at the time of invention, to modify the teachings of Han with the teachings of Tai to use the geometric relationship proposed in [2] describing the correlation between the current block position and the motion edge to reduce the number of PMV predictors to be evaluated (*Tai- Page 190, Col 1, ¶ 4*).

Regarding claim 2, Han and Tai teach all the limitations of claim 1. Han further teaches:

comprising selecting spatial motion vector prediction candidates from the set of spatial motion vector prediction candidates as the potential spatial motion vector prediction candidate in a predetermined order (*i.e. comprising selecting spatial motion vector prediction candidates from the set of spatial motion vector prediction candidates as the potential spatial motion vector prediction candidate in a predetermined order- page 1712, Col 2, ¶ 4*).

Regarding claim 3, Han and Tai teach all the limitations of claim 1.

However, Han does not teach:

comprising comparing motion information of the potential spatial motion vector prediction candidate with motion information of at most one other spatial motion vector prediction candidate of the set of spatial motion vector prediction candidates.

In the same field of endeavor, Tai teaches:

comprising comparing motion information of the potential spatial motion vector prediction candidate (*i.e.* $\overrightarrow{MV}^i(\vec{X}, t) = \arg_{\vec{V} \in CS^i(\vec{X}, t)} \min(e(\vec{V}, \vec{X}, t))$, $1 \leq i \leq 12$; *equation 1*) with motion information of at most one other spatial motion vector prediction candidate of the set of spatial motion vector prediction candidates (*i.e.* *Each pass in the proposed ME scheme is performed using one of the PMV predictor patterns (see Fig. 3) and then one MV in the selected PMV pattern which has minimal distortion is selected as the final PMV of this pass- Page 190; Col 1, ¶ 5*).

Regarding claim 4, Han and Tai teach all the limitations of claim 1. Han further teaches:

comprising examining whether the received block of pixels is divided into a first prediction unit and a second prediction unit (*i.e. tu size flag is equal to one*); and if so, excluding the potential spatial motion vector prediction candidate from the

merge list (*i.e. the transform may be applied to the residue generated by multiple PU partitions with different motion vectors or prediction methods*) if the prediction unit is the second prediction unit (*i.e. When the tu size flag is equal to zero, the TU size is set equal to that of the CU which it belongs to. When tu size flag is equal to one, the TU size is set as $N \times N$ for symmetric PU splitting's and $N/2 \times N/2$ for asymmetric PU splitting's, respectively. This ensures that the transform which is not applied across motion boundaries can be tested in the rate-distortion optimization process for asymmetric PU partitions. It should be noted that the transform may be applied to the residue generated by multiple PU partitions with different motion vectors or prediction methods-* Page 1711, ¶ 4-5)

Regarding claim 5, Han and Tai teach all the limitations of claim 1. Han further teaches:

further comprising determining a maximum number (*i.e. N*) of spatial motion vector prediction candidates to be included in a merge (*i.e. A'*) list (*i.e. Fig. 3 shows the spatially adjacent motion vectors which can be considered the candidate for the prediction. Let $A = \{a_0, a_1, \dots, a_N\}$, $A' \subset A$, and $I_{A'}$ denote the set of motion vectors above the current PU, the set of the available motion vectors in A, and the set which includes the indices of the elements in A' respectively-* page 1712, col 2, ¶ 3); and

limiting the number (*i.e. $A' \subset A$*) of spatial motion vector prediction candidates in the merge list smaller or equal to the maximum number (*i.e. The set of the*

available motion vectors, A', is derived from A by verifying whether the corresponding reference index is the same as that of the current PU partition- page 1712, col 2, ¶ 3).

Regarding claim 7, Han and Tai teach all the limitations of claim 1. Han further teaches:

further comprising including a temporal motion prediction candidate into the merge list (i.e. This paper proposes the AMVP method, which is adapted to the proposed flexible unit representation. It allows the selection of the best predictor from the set which consists of three spatially adjacent motion vectors, their median, and a temporal motion vector- Page 1712, Col 2, ¶ 2).

Regarding claim 8, Han and Tai teach all the limitations of claim 1. Han further teaches:

comprising selecting one motion vector prediction candidate (i.e. the best predictor is selected) from the merge list to represent a motion vector prediction for the block of pixels (i.e. the best predictor is selected from a given set through rate-distortion optimization- Page 1712, Col 2, ¶ 1).

Regarding claim 9, Han teaches

1. A method comprising (i.e. a novel video compression scheme based on a highly flexible hierarchy of unit representation which includes three block

concepts: coding unit (CU), prediction unit (PU), and transform unit (TU)-

Abstract):

receiving a encoded block of pixels including a prediction unit (i.e. Fig. 1 shows all proposed building blocks of the decoder. All the coding tools are extended to support the proposed flexible architecture. Intra prediction is extended to support an arbitrary number of angles rather than the conventional 9-modes defined in H.264/AVC. Fast integer transforms larger than 8×8 are developed to support large TU sizes. Edge definition for the deblocking filter is modified according to the CU, PU, and TU concepts. Partition information in quadtreebased adaptive loop filter (QALF) [13] is replaced with CU splitting information- page 1710, ¶ 3);

However, Lin does not teach explicitly:

determining a set of spatial motion vector prediction candidates for the encoded block of pixels;

the spatial motion vector prediction candidates being provided with motion information;

selecting a first spatial motion vector prediction candidate from the set of spatial motion vector prediction candidates as a potential spatial motion vector prediction candidate to be included in a merge list for the prediction unit;

determining a subset of spatial motion vector predictions based on the location of the block associated with the first spatial motion vector prediction candidate;

comparing motion information of the first spatial motion vector prediction candidate with motion information of the spatial motion vector prediction

candidate in the determined subset of spatial motion vector prediction candidates;

if at least one of the comparisons indicates that the motion vector information of the spatial motion vector prediction candidates correspond with each other, excluding the first spatial motion vector prediction candidate from the merge list.

In the same field of endeavor, Tai teaches:

determining a set (*i.e.* $CS^i(\vec{X}, t)$) of spatial motion vector prediction candidates for the encoded block of pixels (*i.e.* *The MV search in block in the i th pass is performed in accordance with (1). As shown, the search is confined to the vectors defined in the candidate vector set $CS^i(\vec{X}, t)$, i.e., a full search is not performed-* Page 189, Col 2, ¶ 3);

the spatial motion vector prediction candidates being provided with motion information (*i.e.* *The basic principle of the multi-pass ME strategy proposed in this study is to utilize the motion information relating to neighboring blocks generated in the previous pass to gradually refine the accuracy of the MV in the current block-* Page 189, Col 2, ¶ 3);

selecting a first spatial motion vector prediction (*i.e.* $CS^i(\vec{X}, t) = \{\vec{V} | \vec{V} = \overline{PMV}^i(\vec{X}) + \vec{U}, \vec{U} \in US^i\}$ - equation 2) candidate (*i.e.* $\overline{PMV}^i(\vec{X})$, *predictive motion vectors (PMVs)-* Page 190, Col 1, ¶ 2) from the set of spatial motion vector prediction candidates (*i.e.* $\overline{PMV}^i(\vec{X})$, *is the PMV of block and is chosen from the MVs of the neighboring blocks and is a set of update vectors-* Page 190, Col 1, ¶ 2) as a potential spatial motion vector prediction candidate to be included in a

merge list for the (*i.e.* current block) prediction unit (*i.e.* Although the MVs of all eight neighboring blocks can theoretically be used as PMV predictors for the current block- Page 190, Col 1, ¶ 3);

determining a subset (*i.e.* $CS^i(\vec{X}, t) = \{\vec{V} | \vec{V} = \overline{PMV}^i(\vec{X}) + \vec{U}, \vec{U} \in US^i\}$)- equation

2) of spatial motion vector predictions based on the location of the block

associated with the first spatial motion vector prediction candidate (*i.e.* by

selecting one of the $\overline{PMV}$ from MVs of neighboring blocks, a subset of $CS^i(\vec{X}, t)$ is

defined for different linear combination of that PMV and $\vec{U} \in US^i$);

comparing motion information of the first spatial motion vector prediction

candidate (*i.e.* $\overline{MV}^i(\vec{X}, t) = \arg_{\vec{V} \in CS^i(\vec{X}, t)} \min(e(\vec{V}, \vec{X}, t))$, $1 \leq i \leq 12$; equation 1)

with motion information of the spatial motion vector prediction candidate in the

determined subset of spatial motion vector prediction candidates (*i.e.* Each pass

in the proposed ME scheme is performed using one of the PMV predictor

patterns (see Fig. 3) and then one MV in the selected PMV pattern which has

minimal distortion is selected as the final PMV of this pass- Page 190; Col 1, ¶ 5);

if at least one of the comparisons indicates that the motion vector information of

the spatial motion vector prediction candidates correspond with each other (*i.e.*

all MV's with non-minimal distortion as inferred above), excluding the first spatial

motion vector prediction candidate from the merge list (*i.e.* all MV's with non-

minimal distortion as inferred above will not be a final PMV for that pass).

Regarding claim 10, Han and Tai teach all the limitations of claim 9.

However, Han does not teach:

comprising comparing motion information of the potential spatial motion vector prediction candidate with motion information of at most one other spatial motion vector prediction candidate of the set of spatial motion vector prediction candidates.

In the same field of endeavor, Tai teaches:

comprising comparing motion information of the potential spatial motion vector

prediction candidate (*i.e.* $\overrightarrow{MV}^i(\overrightarrow{X}, t) = \arg_{\overrightarrow{V} \in CS^i(\overrightarrow{X}, t)} \min(e(\overrightarrow{V}, \overrightarrow{X}, t))$, $1 \leq i \leq 12$;

equation 1) with motion information of at most one other spatial motion vector

prediction candidate of the set of spatial motion vector prediction candidates (*i.e.*

Each pass in the proposed ME scheme is performed using one of the PMV

predictor patterns (see Fig. 3) and then one MV in the selected PMV pattern

which has minimal distortion is selected as the final PMV of this pass- Page 190;

Col 1, ¶ 5).

Regarding claim 11, Han and Tai teach all the limitations of claim 9. Han further teaches:

comprising examining whether the received encoded block of pixels is divided

into a first prediction unit and a second prediction unit (*i.e. tu size flag is equal to*

one); and if so, excluding the potential spatial motion vector prediction candidate

from the merge list (*i.e. the transform may be applied to the residue generated by*

multiple PU partitions with different motion vectors or prediction methods) if the

prediction unit is the second prediction unit (i.e. When the tu size flag is equal to zero, the TU size is set equal to that of the CU which it belongs to. When tu size flag is equal to one, the TU size is set as $N \times N$ for symmetric PU splitting's and $N/2 \times N/2$ for asymmetric PU splitting's, respectively. This ensures that the transform which is not applied across motion boundaries can be tested in the rate-distortion optimization process for asymmetric PU partitions. It should be noted that the transform may be applied to the residue generated by multiple PU partitions with different motion vectors or prediction methods- Page 1711, ¶ 4-5)

Regarding claim 12, Han and Tai teach all the limitations of claim 9. Han further teaches:

further comprising determining a maximum number (i.e. N) of spatial motion vector prediction candidates to be included in a merge (i.e. A') list (i.e. Fig. 3 shows the spatially adjacent motion vectors which can be considered the candidate for the prediction. Let $A = \{a_0, a_1, \dots, a_N\}$, $A' \subset A$, and $I_{A'}$ denote the set of motion vectors above the current PU, the set of the available motion vectors in A , and the set which includes the indices of the elements in A' respectively- page 1712, col 2, ¶ 3); and

limiting the number (i.e. $A' \subset A$) of spatial motion vector prediction candidates in the merge list smaller or equal to the maximum number (i.e. The set of the available motion vectors, A' , is derived from A by verifying whether the

*corresponding reference index is the same as that of the current PU partition-
page 1712, col 2, ¶ 3).*

Regarding claim 14, Han and Tai teach all the limitations of claim 9. Han further teaches:

comprising selecting one motion vector prediction candidate (i.e. the best predictor is selected) from the merge list to represent a motion vector prediction for the received encoded block of pixels(i.e. the best predictor is selected from a given set through rate-distortion optimization- Page 1712, Col 2, ¶ 1).

Regarding claim 15 and 19, apparatus claim 15 and 19 is drawn to the apparatus using/performing the same method as claimed in claim 1. Therefore apparatus claim 15 and 19 corresponds to method claim 1, and is rejected for the same reasons of obviousness as used above.

Regarding claim 16 and 20, apparatus claim 16 and 20 is drawn to the apparatus using/performing the same method as claimed in claim 9. Therefore apparatus claim 16 and 20 corresponds to method claim 9, and is rejected for the same reasons of obviousness as used above.

Regarding claim 17, computer-readable medium storing instructions claim 17 corresponds to the same method as claimed in claim 1, and therefore is also rejected for the same reasons of obviousness as listed above.

Regarding claim 18, computer-readable medium storing instructions claim 50 corresponds to the same method as claimed in claim 9, and therefore is also rejected for the same reasons of obviousness as listed above.

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to CLIFFORD HILAIRE whose telephone number is (571)272-8397. The examiner can normally be reached on Monday-Friday- 0800-1700, EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Sath V. Perungavoor can be reached on (571)272-7455. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/C. H./
Examiner, Art Unit 2488

/SATH V PERUNGAVOOR/
Supervisory Patent Examiner, Art Unit 2488

Notice of References Cited	Application/Control No. 13/666,680	Applicant(s)/Patent Under Reexamination BICI ET AL.	
	Examiner CLIFFORD HILAIRE	Art Unit 2488	Page 1 of 1

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	Classification
	A	US-			
	B	US-			
	C	US-			
	D	US-			
	E	US-			
	F	US-			
	G	US-			
	H	US-			
	I	US-			
	J	US-			
	K	US-			
	L	US-			
	M	US-			

FOREIGN PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	Classification
	N					
	O					
	P					
	Q					
	R					
	S					
	T					

NON-PATENT DOCUMENTS

*		Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)
	U	Han, Woo-Jin, et al. "Improved video compression efficiency through flexible unit representation and corresponding extension of coding tools." Circuits and Systems for Video Technology, IEEE Transactions on 20.12 (2010): 1709-1720.
	V	Tai, Shen-Chuan, et al. "A multi-pass true motion estimation scheme with motion vector propagation for frame rate up-conversion applications." Journal of display technology 4.2 (2008): 188-197.
	W	
	X	

*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)
Dates in MM-YYYY format are publication dates. Classifications may be US or foreign.



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
 United States Patent and Trademark Office
 Address: COMMISSIONER FOR PATENTS
 P.O. Box 1450
 Alexandria, Virginia 22313-1450
 www.uspto.gov

BIB DATA SHEET

CONFIRMATION NO. 4782

SERIAL NUMBER	FILING or 371(c) DATE	CLASS	GROUP ART UNIT	ATTORNEY DOCKET NO.		
13/666,680	11/01/2012	375	2488	NC77198US-NP		
APPLICANTS INVENTORS Mehmet Oguz BICI, Tampere, TURKEY; Jani LAINEMA, Tampere, FINLAND; Kemal UGUR, Tampere, FINLAND; ** CONTINUING DATA ***** This appln claims benefit of 61/555,703 11/04/2011 ** FOREIGN APPLICATIONS ***** ** IF REQUIRED, FOREIGN FILING LICENSE GRANTED ** 11/20/2012						
Foreign Priority claimed ☐ Yes ☒ No 35 USC 119(a-d) conditions met ☐ Yes ☒ No Verified and /CLIFFORD HILAIRE/ Acknowledged Examiner's Signature		☐ Met after Allowance Initials	STATE OR COUNTRY TURKEY	SHEETS DRAWINGS 13	TOTAL CLAIMS 20	INDEPENDENT CLAIMS 8
ADDRESS Nokia USA Inc. Attn: Intellectual Property Rights Docketing 200 South Mathilda Ave Sunnyvale, CA 94086 UNITED STATES						
TITLE METHOD FOR CODING AND AN APPARATUS						
FILING FEE RECEIVED 2640	FEES: Authority has been given in Paper No. _____ to charge/credit DEPOSIT ACCOUNT No. _____ for following:			☐ All Fees ☐ 1.16 Fees (Filing) ☐ 1.17 Fees (Processing Ext. of time) ☐ 1.18 Fees (Issue) ☐ Other _____ ☐ Credit		

Improved Video Compression Efficiency Through Flexible Unit Representation and Corresponding Extension of Coding Tools

Woo-Jin Han, *Member, IEEE*, Junghye Min, Il-Koo Kim, Elena Alshina, Alexander Alshin, Tammy Lee, Jianle Chen, Vadim Seregin, Sunil Lee, *Member, IEEE*, Yoon Mi Hong, Min-Su Cheon, Nikolay Shlyakhov, Ken McCann, Thomas Davies, and Jeong-Hoon Park

Abstract—This paper proposes a novel video compression scheme based on a highly flexible hierarchy of unit representation which includes three block concepts: coding unit (CU), prediction unit (PU), and transform unit (TU). This separation of the block structure into three different concepts allows each to be optimized according to its role; the CU is a macroblock-like unit which supports region splitting in a manner similar to a conventional quadtree, the PU supports nonsquare motion partition shapes for motion compensation, while the TU allows the transform size to be defined independently from the PU. Several other coding tools are extended to arbitrary unit size to maintain consistency with the proposed design, e.g., transform size is extended up to 64×64 and intraprediction is designed to support an arbitrary number of angles for variable block sizes. Other novel techniques such as a new noncascading interpolation filter design allowing arbitrary motion accuracy and a leaky prediction technique using both open-loop and closed-loop predictors are also introduced. The video codec described in this paper was a candidate in the competitive phase of the high-efficiency video coding (HEVC) standardization work. Compared to H.264/AVC, it demonstrated bitrate reductions of around 40% based on objective measures and around 60% based on subjective testing with 1080p sequences. It has been partially adopted into the first standardization model of the collaborative phase of the HEVC effort.

Index Terms—Flexible unit representation, HEVC, video coding.

I. INTRODUCTION

RECENTLY, with the ever increasing popularity of high-definition (HD) video content, video compression tech-

nologies which can provide higher coding efficiency than existing video coding standards have received increased attention. Early video coding standards such as MPEG-1 [1] and MPEG-2 [2] employ fixed-size block motion compensation method with a macroblock size of 16×16 . In the state-of-the-art standard H.264/AVC [3], this macroblock size is retained but with the addition of a depth-2 quadtree, which significantly improves coding efficiency. However, it has been observed that the maximum 16×16 block size is too small when applied to high resolution video and causes inefficiency [4], [5]. To solve this problem, recent Visual Coding Experts Group (VCEG) proposals have extended the maximum motion block size up to 64×64 by adding additional large motion block modes on top of the conventional 16×16 H.264/AVC macroblock syntax [6]–[8].

Although those approaches suggested that the coding efficiency for the high resolution video materials can be improved by providing a better tradeoff between prediction performance and side information, they simply tried to add the large size motion partition while maintaining the existing 16×16 macroblock sizes, which imposes restrictions on the flexibility of the quadtree structure. For example, spatial/temporal prediction mode switching, intracoding tools and spatial transforms are still defined based on the 16×16 macroblock sizes, thus the schemes generally suffer from lower coding efficiency in intracoded blocks.

There has been extensive research on introducing a general multidepth quadtree for variable block-size motion compensation [9]–[12]. A frame is typically divided into fixed-size superblocks, with each superblock able to be further split. This kind of generic quadtree can provide a consistent codec design; however, focusing only on motion compensation does not exploit the full potential of the technique.

This paper describes a video compression framework where the whole codec has been designed to exploit flexible block sizes. A highly flexible hierarchical structure based on the generic quadtree scheme is defined by three independent block concepts: coding unit (CU), prediction unit (PU), and transform unit (TU). CU is the basic unit of region splitting, CU is analogous to the concept of macroblock, but it does not restrict the maximum size and it allows recursive splitting into four equal size CUs to improve the contents adaptivity. PU is the

Manuscript received July 23, 2010; revised October 11, 2010; accepted October 19, 2010. Date of publication November 15, 2010; date of current version January 22, 2011. This paper was recommended by Associate Editor T. Wiegand.

W. J. Han, J. Min, I. K. Kim, E. Alshina, A. Alshin, T. Lee, J. Chen, V. Seregin, S. Lee, Y. M. Hong, M. S. Cheon, N. Shlyakhov, and J. H. Park are with the Digital Media and Communication Research and Development Center, Samsung Electronics, Suwon 443-742, Korea (e-mail: wjhan.han@samsung.com; jh643.min@samsung.com; ilkoo.kim@samsung.com; elena.a.alshina@samsung.com; alexander.b.alshin@samsung.com; tammy.lee@samsung.com; jianle.chen@samsung.com; v.seregin@samsung.com; sunil.lee@samsung.com; yoonmi.hong@samsung.com; minsu.cheon@samsung.com; n.shlyakhov@samsung.com; jeonghoon@samsung.com).

K. McCann is with ZetaCast, Bath, BA2 6NZ, U.K. (e-mail: ken@zetacast.com).

T. Davies is with British Broadcasting Corporation, London, W1A 1AA, U.K. (e-mail: thomas.davies@bbc.co.uk).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCSVT.2010.2092612

1051-8215/\$26.00 © 2010 IEEE

basic unit of inter/intraprediction and it may contain multiple arbitrary shape partitions in a single PU to effectively code irregular image patterns. TU is the basic unit of transform. It can be defined independently from the PU; however, its size is limited to the CU which the TU belongs to. This separation of the block structure into three different concepts allows each to be optimized according to its role, which results in the improved coding efficiency.

From these unit definitions, all coding processes can be defined in a very consistent way; large block sizes are supported in both motion compensation and intraprediction schemes, prediction mode switching is allowed at any block sizes and spatial transforms larger than 16×16 are also supported in both inter and intracoded blocks. Due to this consistency, every syntax element can be defined in exactly the same way independent of the unit sizes, which makes the specification and parsing process much simpler than would be the case with an extension scheme on top of a conventional 16×16 macroblock syntax.

Fig. 1 shows all proposed building blocks of the decoder. All the coding tools are extended to support the proposed flexible architecture. Intra prediction is extended to support an arbitrary number of angles rather than the conventional 9-modes defined in H.264/AVC. Fast integer transforms larger than 8×8 are developed to support large TU sizes. Edge definition for the deblocking filter is modified according to the CU, PU, and TU concepts. Partition information in quadtree-based adaptive loop filter (QALF) [13] is replaced with CU splitting information.

The rest of this paper is organized as follows. Section II introduces the details of the unit definitions and Section III presents the proposed coding tools for the interframe prediction. The remaining coding tools for the intraframe prediction, spatial transform, in-loop filtering, and entropy coding are introduced in Sections IV, V, VI, and VII, respectively. Section VIII summarizes the coding efficiency of the proposed techniques based on the objective and subjective results. Finally, Section IX concludes the paper. For a more complete description for the coding tools, please refer to [14] and [15].

II. UNIT DEFINITION

For efficient and flexible representation of video content with various resolutions, a triplet of CU, PU, and TU are introduced. The following subsections describe the details of each unit.

A. Coding Unit

The picture is divided into slices, with each slice composed of a sequence of largest coding units (LCUs). The LCU size is specified in the sequence parameter set (SPS) or slice header; the optimum LCU size is dependent on the application.

Fig. 2 shows quad-tree based CU structure. The LCU can be divided into four CUs by a 1-bit syntax, *split_flag*. Each split CU can be divided into four CUs recursively up to the maximum allowed hierarchical depth, also specified in the SPS or slice header. For example, if LCU size is equal to 128 and the maximum hierarchical depth is equal to 5, then

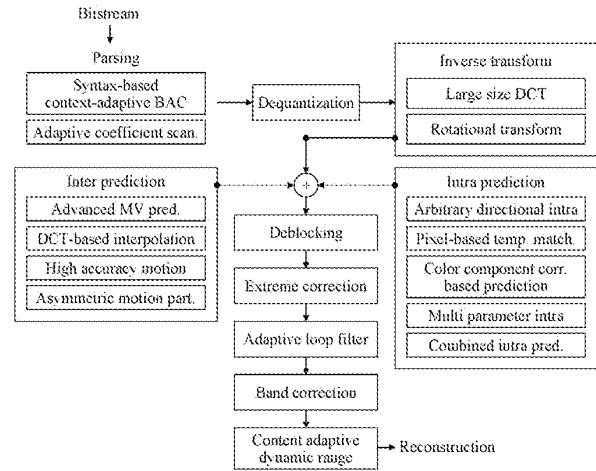


Fig. 1. All building blocks to decode the input bit-stream.

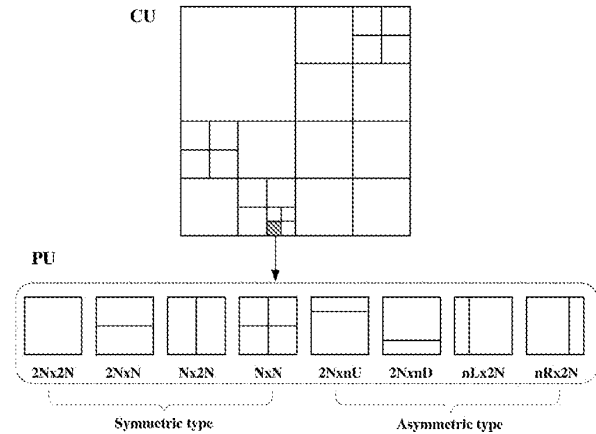


Fig. 2. Relationship between CU and PU.

5 kinds of CU sizes are possible: 128×128 (LCU), 64×64 , 32×32 , 16×16 , and 8×8 . The three syntax elements, LCU size, maximum hierarchical depth, and *split_flag*, completely specify the generic multidepth quadtree.

B. Prediction Unit

Coupled with CU, this paper introduces a basic unit for prediction: the PU. The PU is defined only for the leaf node CU which is not split in quadtree structure, and the size of the PU is limited to that of the CU. The prediction method is specified by the prediction mode and the PU splitting. As with H.264/AVC, the prediction mode can be one of the values among skip, intra, and inter. Fig. 2 shows different PU splitting strategy for a CU of size $2N \times 2N$, which is dependent on the property of each prediction mode.

Skip mode can be thought of as a special intermode in which no motion vector and residue information are transmitted. Typically it is used in an area of the image with smooth motion, thus no further splitting is necessary and only $2N \times 2N$ partition is used. For the normal intermode, a set of eight splitting options are defined to cover possible motion boundaries

inside the PU. Given a PU of size $2N \times 2N$, four symmetric splittings ($2N \times 2N$, $2N \times N$, $N \times 2N$, $N \times N$) and four asymmetric splittings ($2N \times nU$, $2N \times nD$, $nL \times 2N$, $nR \times 2N$) are defined. The asymmetric PU splittings are included in the design to improve the coding efficiency for the irregular object boundaries, which otherwise would be constrained to being represented less efficiently by further CU splittings to the deeper depth. For the intramode, only two possible splittings $2N \times 2N$ and $N \times N$ are defined.

All information related to the prediction is signaled on a PU basis, for example, the prediction mode and PU splitting type are specified for each PU. Once the prediction mode and PU splitting are given, motion vector difference, intraprediction direction, reference indices for each PU partition are specified on a PU partition basis.

C. Transform Unit

The TU, for transform and quantization, is the third basis unit. Although, in principle, TU can be defined in a similar way to PU to support arbitrary shape transforms, only two TU partitions are used in this paper considering the implementation complexity.

When the `tu_size_flag` is equal to zero, the TU size is set equal to that of the CU which it belongs to. When `tu_size_flag` is equal to one, the TU size is set as $N \times N$ for symmetric PU splittings and $N/2 \times N/2$ for asymmetric PU splittings, respectively. This ensures that the transform which is not applied across motion boundaries can be tested in the rate-distortion optimization process for asymmetric PU partitions.

It should be noted that the transform may be applied to the residue generated by multiple PU partitions with different motion vectors or prediction methods. This approach is especially beneficial for coding semirandom residuals generated by complex texture areas where neighboring partitions share common signal characteristics, which can exploit the better energy compaction of larger transform [14].

D. Size-Independent Syntax Representation

The proposed unit definitions allow us to design a size-independent syntax representation which specifies all syntax elements in a consistent way independent of the unit size. By contrast, in H.264/AVC and its extension in [6]–[8], block level syntax elements such as `prediction_mode`, `transform_flag`, `coded_block_flag`, and `intra_pred_mode` are coded differently depending on block size. This size-independent method of representation allows syntax elements to be specified generically for all the remaining coding tools. This property can greatly simplify the specification effort as well as the actual parsing process, especially if a large number of hierarchical depths is allowed.

III. INTERPREDICTION

In this paper, three aspects of interframe prediction are improved: motion accuracy, interpolation filter, and motion vector prediction. For motion accuracy, quarter-pel precision

with 1/12-pel optional refinement is used to improve the coding efficiency while minimizing the additional overhead for signaling the motion information. For motion compensation, noncascaded spatial filtering using a discrete cosine transform (DCT)-based interpolation filter is used for every motion accuracy position. For motion vector prediction, advanced motion vector prediction (AMVP), which is a motion vector competition method [16] adapted to the flexible unit representation, is used to reduce overhead for motion information.

A. High Accuracy Motion

Even though higher motion accuracy results in less prediction error, it does not guarantee improved coding efficiency, since it leads to an increase in the bit rate to convey the motion information. For example, 1/8-pel motion accuracy may often result in a reduction in the overall coding efficiency, except for particular sequences [17].

In this paper, high accuracy motion (HAM), which adaptively refines the motion information to an arbitrary accuracy, is proposed. In the HAM method, motion information at quarter-pel accuracy is followed by a refinement flag and optional refinement information at 1/12-pel accuracy. The refinement flag is set equal to 1 only when the refinement provides the gain in the rate-distortion optimization sense, otherwise, the flag is set to 0 and no refinement information is signaled. Although the refinement information can represent any accuracy, the current design uses 1/12-pel, since refinement in 1/8-pel accuracy results in the redundancy in the refinement information.

B. DCT-Based Interpolation Filter

Motion compensation with fractional pel accuracy requires interpolation of the reference signal. In H.264/AVC standard, the prediction values at half-pel positions are obtained using a 6-tap Wiener filter, while those at quarter-pel positions are obtained by bilinear combination of the samples at integer and half-pel positions [3]. This paper proposes a DCT-based interpolation filter (DCT-IF), which can directly provide the interpolated value at the desired fractional accuracy from the samples at integer-pel positions. Since the DCT-IF can provide the interpolated value without cascaded filtering, the motion compensation process can be simplified and the complexity can also be reduced for samples at quarter-pel accuracy. Furthermore, the DCT-IF provides a unified way to generate the interpolation filters supporting motion accuracy higher than quarter-pel, e.g. 1/8-pel, 1/12-pel, etc. Note that 12-tap and 6-tap DCT-IFs are used for luma and chroma, respectively.

1) *Filter Coefficient Generation*: In this paper, trigonometric interpolation, which has been widely used for the other applications [18], is exploited to construct DCT-IF for high accuracy motion compensation. Let $p(\alpha)$ denote the sample value at the fractional position $\alpha \in [-(M-1), M]$, where M is a positive integer. Then, $p(\alpha)$ can be derived as follows:

$$p(\alpha) = \mathbf{f}(\alpha)^T \mathbf{p} = (\mathbf{w}(\alpha)^T \mathbf{D}) \mathbf{p} \quad (1)$$

where $\mathbf{p} = [p_{-(M-1)}, \dots, p_M]^T$ is the vector of the given samples at the integer-pel positions, $\mathbf{D}$ is the $2M \times 2M$ DCT

TABLE I
FILTER COEFFICIENTS OF 6-TAP DCT-IF

α	$2M = 6$ (6-tap filter)
$-1/12$	{ -4, 19, 254, -19, 8, -2 }
$1/12$	{ 4, -16, 252, 22, -8, 2 }
$2/12$	{ 6, -28, 242, 48, -17, 5 }
$3/12$	{ 9, -37, 227, 75, -25, 7 }
$4/12$	{ 11, -42, 208, 103, -33, 9 }
$5/12$	{ 12, -44, 184, 132, -39, 11 }
$6/12$	{ 11, -43, 160, 160, -43, 11 }
$7/12$	{ 11, -39, 132, 184, -44, 12 }
$8/12$	{ 9, -33, 103, 208, -42, 11 }
$9/12$	{ 7, -25, 75, 227, -37, 9 }
$10/12$	{ 5, -17, 48, 242, -28, 6 }

matrix whose element is given as

$$d_{ij} = \frac{1}{M} \cos \left(\frac{(2j-1+2M)i\pi}{4M} \right) \quad (2)$$

where $0 \leq i \leq (2M-1)$, $(-M+1) \leq j \leq M$, and $\mathbf{w}_{(\alpha)} = [w_0^{(\alpha)}, \dots, w_{(2M-1)}^{(\alpha)}]^T$ is given as

$$w_i^{(\alpha)} = \begin{cases} \frac{1}{2}, & i = 0 \\ \cos \left(\frac{\pi(2\alpha-1+2M)i}{4M} \right), & 1 \leq i \leq (2M-1). \end{cases} \quad (3)$$

By using $\mathbf{f}_{(\alpha)}$ of (1), one can easily calculate the filter coefficients for an arbitrary number of filter taps ($2M$) and arbitrary fractional position α . Table I shows the coefficients of the 6-tap filters generated for 1/12-pel motion accuracy. Note that the filter coefficients are scaled by 256, H.264/AVC uses 5 bits for filter coefficients.

2) *Low Complexity Optimization*: An efficient hardware implementation requires multiplication-free representation for all the calculations. In order to minimize the number of arithmetical operations (shifts and additions), the filter coefficients in Table I can be slightly changed without significant loss in coding efficiency. Table II shows 5 bits representation for optimized filter coefficients for the quarter-pel positions, together with the required number of arithmetical operations. The proposed DCT-IF is a noncascaded interpolation scheme, i.e. only one spatial filtering is performed for all the possible positions. It can be implemented by a multiplication-free design. Thus, the 6-tap DCT-IF has the smaller number of operations compared to the interpolation filter of H.264/AVC since samples at the quarter-pel position can be generated by one 6-tap filtering operation whereas H.264/AVC uses a combination of 6-tap and bi-linear filters. The 12-tap DCT-IF filter requires twice as many operations as the 6-tap DCT-IF, however it provides similar performance to the well-known adaptive interpolation filter (AIF) scheme [19] while being less complex and more suitable for hardware implementation.

TABLE II
OPTIMIZED FILTER COEFFICIENTS FOR 6-TAP DCT-IF AT QUARTER-PEL POSITIONS

α	$2M = 6$ (6-tap filter)	Add	Shift
$1/4$	{ 1, -4, 28, 9, -3, 1 }	8	2
$2/4$	{ 1, -5, 20, 20, -5, 1 }	6	2
$3/4$	{ 1, -3, 9, 28, -4, 1 }	8	2

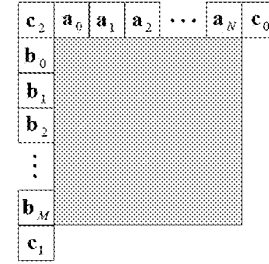


Fig. 3. Spatially adjacent motion vectors for the AMVP.

C. Advanced Motion Vector Prediction

In H.264/AVC standard, motion vectors are predicted from a median of the motion vectors of the spatially adjacent blocks [3]. To further improve the efficiency of the motion vector prediction, a few competition-based prediction methods have been proposed [16], [20], where the best predictor is selected from a given set through rate-distortion optimization.

This paper proposes the AMVP method, which is adapted to the proposed flexible unit representation. It allows the selection of the best predictor from the set which consists of three spatially adjacent motion vectors, their median, and a temporal motion vector.

Fig. 3 shows the spatially adjacent motion vectors which can be considered the candidate for the prediction. Let $A = \{\mathbf{a}_0, \mathbf{a}_1, \dots, \mathbf{a}_N\}$, $A' \subset A$, and $I_{A'}$ denote the set of motion vectors above the current PU, the set of the available motion vectors in A , and the set which includes the indices of the elements in A' , respectively. The set of the available motion vectors, A' , is derived from A by verifying whether the corresponding reference index is the same as that of the current PU partition. When $A' = \emptyset$, the candidate motion vector $\mathbf{a}$ from A' is excluded from the candidate set, and a zero vector is used when taking the median of spatial motion vector predictors. Otherwise, $\mathbf{a}$ is determined as $\mathbf{a} = \mathbf{a}_n$, where n is the smallest element of $I_{A'}$. The candidate motion vectors $\mathbf{b}$ and $\mathbf{c}$ from $B = \{\mathbf{b}_0, \dots, \mathbf{b}_M\}$ and $C = \{\mathbf{c}_0, \mathbf{c}_1, \mathbf{c}_2\}$ are also determined using the same algorithm, and the resulting spatially adjacent motion vectors are used in the proposed AMVP method.

The overhead for signaling the index of the best predictor is reduced by re-organizing the set of candidate predictors. First, the predictors in the candidate set are reordered based on the PU splitting, so that the most probable motion predictor appears first in the set. Then, duplicate predictors and the predictors which can be excluded based on the parsed MVD values at the decoder side are eliminated from the set [20].

Note that, when only one predictor remains as a result of the re-organization, no overhead is signaled, since it can be also derived at the decoder.

Although AMVP is developed to predict the motion vector, it is also applied to the skip and direct modes, which do not transmit the additional motion vector differences. It should be noted that exactly the same process is applied to all cases to simplify the design. For comparison, H.264/AVC uses a slightly different scheme for the skip mode.

IV. INTRAPREDICTION

To improve the performance of intraprediction, five new intracoding tools are included. Arbitrary directional intra (ADI) extends the conventional nine-direction intraprediction to support any arbitrary number of angles in a unified way. Multiparameter intra (MPI) tries to apply post processing filters to the prediction samples from the ADI. To exploit the interchannel correlation between luminance and chrominance, color component correlation based prediction (CCCP) is introduced. Pixel-based template matching (PTM) is included to improve the coding efficiency for repeated regular patterns. Finally, the combined intraprediction (CIP) allows combining the closed-loop predictor generated from the above methods with an open-loop predictor to further improve the performance.

A. Arbitrary Directional Intra (ADI)

In H.264/AVC intracoding, nine intraprediction modes are supported for 4×4 and 8×8 blocks but only four intraprediction modes for 16×16 blocks. Intra-coded 16×16 blocks were intended to code homogeneous regions in low resolution video. However, even relatively large blocks ($\geq 16 \times 16$) cover only a small part of objects or background in high resolution video, which may contain strong directional patterns.

Arbitrary directional intra (ADI) is a coding tool designed to extend the directional intraprediction scheme used in H.264/AVC to give more accurate representation of directions and better suit larger prediction units. In the ADI scheme, up to 30 angles are generated using the integer pair (dx, dy) to represent the direction of lines which each mode uses for context pixel extrapolation as illustrated in Fig. 4.

Like H.264/AVC, ADI uses the reconstructed pixels. However, in some cases, the down-left region is available for the prediction process as depicted in Fig. 4. When the down-left region is not available, that region is generated using the nearest reconstructed pixel value.

Due to the large variety of the possible unit sizes, it is beneficial to adjust the number of angles according to the unit size. For example, PU partitions with 64×64 and 128×128 sizes are usually used for homogeneous areas which can be generated by using only some prediction modes, such as DC and plane modes. In this case, the number of supported prediction modes of the current PU partition may differ from that of the neighboring PU partitions. Intra prediction modes of the neighboring PU partitions are converted to the nine most frequently used modes to predict the current intraprediction mode in a unified way.

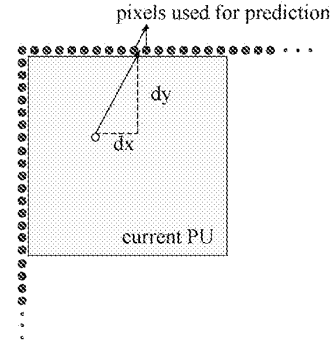


Fig. 4. Neighboring pixels and angles defined by (dx, dy) .

B. Multiparameter Intra

Directional intraprediction sometimes results in nonsmooth and unnatural patterns which are not transform-friendly. This problem can be alleviated by applying a smoothing filter to the prediction signal obtained by the ADI scheme described in Section IV-A. Multiparameter intra (MPI) allows the use of the 4-point filter in (4) for each pixel to increase the probability of generating a prediction signal well-suited to the following transform stages. For each PU partition, `mpi_flag` is used to specify whether or not the MPI scheme is applied.

$$p'[x, y] = (p[x, y] + p[x - 1, y] + p[x, y - 1] + p[x, y + 1] + 2) \gg 2 \quad (4)$$

where $p[x, y]$ and $p'[x, y]$ are the original and the smoothed prediction value at (x, y) coordinates, respectively, and $\gg$ is the bit-wise shift to the right.

C. Color Component Correlation Based Prediction

There have been some studies on predicting chrominance pixels using the correlation with luminance [21]. Although most approaches try to predict the pixel values of one color channel by another color channel, the color component correlation based prediction (CCCP) scheme derives an object segmentation map of the chrominance channel from the already coded luma channel, and applies it to intraprediction.

The CCCP procedure is as follows. Firstly, the size of the reconstructed luma samples including the neighboring luma pixels are adjusted to match that of the chroma samples according to the given chroma format. Then, simple thresholding using the mean value is exploited to generate the binary segmentation map. After that, two mean values are computed from the neighboring chroma pixels according to the segmentation map and assigned to the inner part of the chroma block. Finally, a 3×3 averaging filter is applied to improve the smoothness of the prediction. Fig. 5 shows a brief flow of the CCCP process.

D. Pixel Based Template Matching (PTM)

Whilst directional extrapolation based intraprediction tools are very efficient for regions with directional patterns, they are not well-suited to regions with repeated regular patterns. To overcome this limitation, several intratechniques using some

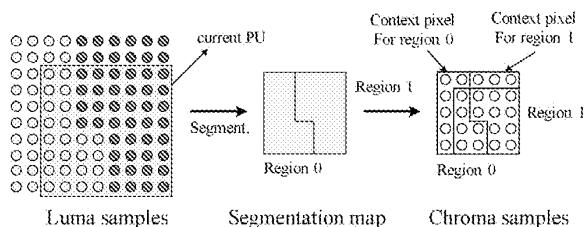


Fig. 5. Processing flow of color component correlation based prediction (CCCP).

form of template matching based on the already decoded pixels have been proposed [22], [23]. Although the reported schemes were proven to provide good coding gains for the targeted areas, a major drawback was the high complexity at the decoder side.

In this paper, a simple pixel based template matching (PTM) approach is proposed. PTM is designed to apply template matching at the pixel level to maximize the degrees of freedom. PTM uses only three pixels as template. Search candidates are also restricted to neighboring 27 pixels, which results in complexity reduction in matching process. Another difference of PTM compared with previous methods is that already prediction samples in the same prediction unit are also used as template and search candidate. This allows use very limited search area even for large block prediction.

E. Combined Intraprediction

Combined intraprediction (CIP) is a coding tool for providing improved prediction, especially in large blocks where a directional linear prediction may not work well, even with the large range of angles available to ADI. CIP predictions comprise a weighted combination of the prediction samples provided by other intracoding tools together with a pixel-by-pixel local mean prediction of pixels within the block to be transformed.

The local mean is constructed as the average of pixel values to the left, top-left and above the current pixel. In the decoder and within the local decoder in the encoder, these values are the reconstructed values after inverse quantization and transform. In the encoder, the reconstructed values are not available in the interior of the block because this would require quantization, transform and inverse transform to have been performed, and the pixels used for prediction lie within the block. Instead original values are used, although the encoder may select any suitable values. This local mean prediction is therefore different between forward prediction (encoder only) and reconstruction (encoder and decoder). Due to this property, CIP can be seen as a leaky prediction method to combine the conventional closed-loop predictor and an open-loop predictor. The advantage of CIP is that it provides pixel-by-pixel adaptation but is a simpler tool than, for example, local template matching approaches.

The combination factor is a 5-bit number between 0 and 31/32 which determines the proportion of local mean in the prediction. Since this number is less than 1, the prediction/reconstruction difference noise is damped, and pixel-by-

pixel prediction adaptation is possible without noise blow-up. The ADI component provides an overall prediction direction and the local mean prediction provides local adaptation.

V. SPATIAL TRANSFORMS

It is well known that transforms larger than 8×8 provides better energy compaction and less quantization error for smooth and coarse textured video areas at the expense of computational complexity [24], [25]. In this paper, 16×16 , 32×32 and 64×64 fast integer transforms are presented in addition to the conventional 4×4 and 8×8 transforms to support large TUs. In addition, a rotational transform (ROT) tool, which can adaptively compact the energy into low frequency components by rotating the transform basis, is also proposed as a second transform.

A. Large Integer Transform (16×16 , 32×32 , and 64×64)

The DCT is independent of input signal while it still provides near optimal decorrelation properties. Moreover, many fast DCT algorithms have been developed. All large transforms proposed in this paper are scaled integer DCTs based on Chen's fast factorizations [26]. Chen's fast DCT has a regular butterfly structure which is beneficial to hardware implementation.

Similarly to the H.264/AVC 4×4 core transform, each multiplication factor in the flow graph of the fast algorithm is scaled and approximated by some fixed precision values p , i.e., each constant is approximated by the dyadic rational of the form $q/2^p$, which can be implemented with only additions and bit-shifting operations. The precision value p can be chosen according to the dynamic range limitation and coding efficiency. As all multiplication factors are approximated, the core transform may not be truly orthogonal. However, this minor nonorthogonality does not result in any perceptible negative effect on the compression performance, whilst the complexity can be significantly reduced. Furthermore, it is still possible to achieve perfect reconstruction using a lifting scheme that can be applied additionally to the butterfly structure [27].

After all multiplication factors are determined, the forward 2-D transform is performed separately in vertical and horizontal direction as follows:

$$Y = C_f(C_f X^T)^T = C_f X C_f^T \quad (5)$$

where X is $N \times N$ residual data, C_f is the core N -point forward transform, and Y is $N \times N$ transformed coefficients. The calculation of $C_f X^T$ is performed from left to right in the flow graph of the fast algorithm. For the inverse transform, the computation order is reversed from right to left. Since the core transform C_f is scaled DCT, the remaining scaling factors should be integrated into the quantization stage.

B. Rotational Transform

While DCT is the most widely used block transform for video and image codecs, the DCT basis functions are not optimal for some types of residual signal. For example, a residual signal having strong diagonal components cannot

be represented efficiently with the DCT basis vectors. Typically, directional transform schemes are proposed to exploit this problem such as mode dependent directional transform (MDDT) [28], [29].

However, the direct use of directional transform is not appropriate to the proposed flexible unit definitions since it is difficult to design the directional transform by a fast butterfly structure such as DCT. In addition, the number of transforms would be almost doubled because both DCT and the directional transform for various TU sizes would be kept in the codec design.

In this paper, ROT is proposed as a second transform that could be applied after the core integer transform instead of introducing completely new transform cores. The main idea of ROT is to change the coordinate system of the transform basis, instead of direct rotation of the input residual. For example, the rotation matrices of horizontal and vertical directions for a 4×4 block with the rotation angles of $(\alpha_1, \alpha_2, \alpha_3)$ and $(\alpha_4, \alpha_5, \alpha_6)$ are defined as

$$R_{\text{ver}}(\alpha_1, \alpha_2, \alpha_3) = \begin{pmatrix} C_1 C_3 - S_1 C_2 S_3 & -S_1 C_3 - C_1 C_2 S_3 & S_2 S_3 & 0 \\ C_1 S_3 + S_1 C_2 C_3 & -S_1 S_3 + C_1 C_2 C_3 & -S_2 C_3 & 0 \\ S_1 S_2 & C_1 S_2 & C_2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (6)$$

$$R_{\text{hor}}(\alpha_4, \alpha_5, \alpha_6) = \begin{pmatrix} C_4 C_6 - S_4 C_5 S_6 & -S_4 C_6 - C_4 C_5 S_6 & S_5 S_6 & 0 \\ C_4 S_6 + S_4 C_5 C_6 & -S_4 S_6 + C_4 C_5 C_6 & -S_5 C_6 & 0 \\ S_4 S_5 & C_4 S_5 & C_5 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (7)$$

where $C_k = \cos \alpha_k$ and $S_k = \sin \alpha_k$ for $k = 1, \dots, 6$. The rotation matrices for 8×8 block can be defined similar to those for 4×4 block.

Each set of angles $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6\}$ specifies one element in the ROT dictionary. In ROT, only four out of all possible rotational angles sets are used. These RT matrices minimize the discrepancy between forward-inverse RT in integer arithmetic realization and provide better energy compaction. These sets of rotational angles were selected using the training with VCEG test sequences, and they provide similar performance for other video sequences including the test sequences for the HEVC standardization. This proves that these angles sets capture and exploit general characteristic of natural images. The index of rotational angles in the dictionary is signaled to the decoder for each CU explicitly.

Furthermore, since ROT is a second transform, not all of the DCT coefficients are processed. For TUs larger than 4×4 , only 8×8 low-frequency areas are processed. It has been found that this restriction does not degrade the coding efficiency since most of the coefficients are already compacted into these low-frequency areas.

The final 2-D spatial transform process when the ROT is used can be summarized as

$$Y = R_{\text{ver}}(C_f X C_f^T) R_{\text{hor}} \quad (8)$$

where X is residual data, C_f is the core forward transform, and Y is output coefficients.

VI. IN-LOOP FILTERING

After reconstruction, several in-loop filters are applied to reduce the distortion between the original and the reconstructed video frames. In addition to a modified H.264/AVC deblocking filter for the CU/PU/TU structure, a CU-synchronized adaptive loop filter (CU-ALF) is introduced to reduce the average distortion of the reconstructed frame. Extreme correction (EXC) and band correction (BDC) techniques are applied to reduce the distortion further by considering object-based pixel classifications. Finally, content adaptive dynamic range (CADR) is proposed to reduce the performance loss from the rounding effects of the various internal processes.

A. Deblocking Filter

The same process as used in H.264/AVC is employed for the deblocking filter. However, the use of a large size transform can effectively reduce the complexity of the deblocking filter, by reducing the number of deblocked pixels. It was reported that the number of deblocked pixels was reduced by 43% on average [14] for the 1080p sequences used in [30].

B. CU-Synchronized Adaptive Loop Filter (CU-ALF)

QALF [13] is known to provide significant coding gain by effectively combining both the Wiener filter approach and the quadtree based on/off selections. In QALF, along with filter coefficients, the information representing the quadtree structure is coded independently with the prediction quadtree structure, which results in large amount of side information in the slice header.

Since the proposed unit representations provides a generic quadtree structure similar to the original design of QALF, this information can be reutilized for the adaptive loop filter process. CU-synchronized adaptive loop filter (CU-ALF) is based on the QALF approach; however, it uses the already coded CU hierarchy as the on/off control for the adaptive loop filter process instead of coding separate quadtree information in the slice header.

It has been reported that the proposed CU-ALF provides comparable coding gain to QALF, while eliminating the need for any side information related to the quadtree partition in the slice header and reducing the encoder complexity [14].

C. Extreme Correction and Band Correction

EXC is designed to compensate the distortion according to the identified pixel classes. Each pixel is categorized as local maxima, local minima, or object edge by simple computation based on the four neighboring pixel values: left, above, right, and bottom. For each pixel class, the mean difference between the original and the reconstructed signals is computed and transmitted to the decoder. A fuller description is given in [14].

BDC is another form of pixel classification, based on pixel value ranges. The main motivation is to equalize the different probability distribution functions of the reconstructed signal with the original. In this paper, up to 16 value ranges are defined; the mean difference for each band is computed and transmitted to the decoder.

D. Content Adaptive Dynamic Range

Internal bit depth increase (IBDI) is a technique to reduce the rounding error by enlarging the internal data precision [31]. Typically, IBDI is implemented by increasing the source bit-depth, which makes the encoder implementation more difficult. CADR is related to IBDI, but is a more general form of source scaling. Using the CADR, a pixel value x is adaptively scaled to $F(x)$ as follows:

$$F(x) = (2^{\text{bit}} - 1)(x - \text{Min})/(\text{Max} - \text{Min}) \quad (9)$$

where bit, Max, and Min are the bit-depth, the maximum value, and the minimum value of the input source, respectively. It should be noted that the parameters Min and Max can be adjusted at sequence, slice, or even block level.

CADR can be used to reduce the rounding error without requiring increased bit-depth in the encoder. For example, source material that is known to be constrained by the [16, 235] safe range of BT.709 can be mapped to [0, 255] to reduce rounding errors in the internal representation.

VII. ENTROPY CODING

In this paper, syntax-based context-adaptive binary arithmetic coder (SBAC) is used for the entropy coding of syntax elements. Additionally, adaptive coefficient scanning (ACS), which provides multiple scanning patterns for the transform coefficients, is also proposed to improve the coding efficiency while reducing the complexity.

A. Syntax-Based Context-Adaptive Binary Arithmetic Coding

As with the context-adaptive binary arithmetic coding (CABAC) of H.264/AVC [32], the SBAC consists of four major steps: syntax-based binarization, context modeling, probability estimation, and binary arithmetic coding. The syntax-based binarization and the context modeling are adapted to the syntax elements obtained from the flexible unit representation and the extended coding tools. For example, the assignment of context models for the syntax elements related to transform coefficients (e.g. significance map, coefficient level, etc.) is dependent on the hierarchical depth of the CU. The probability estimation and binary arithmetic coding are based on an improved version of the arithmetic coding method specified in Annex D of the JPEG still image coding standard [33]. Specifically, the table-based probability update and the multiplication-free design are improved, and the bypass mode for equiprobable binary symbols is added to reduce complexity [14].

B. Adaptive Coefficient Scanning

Since large transforms frequently have very sparse coefficient distributions, a fixed zigzag scanning method may result in long coefficient scanning lists including many zeroes, resulting in inefficient coding. ACS that adaptively chooses the best scanning method for each TU is used to avoid this problem.

Given the quantized transform coefficients, the best scanning pattern is chosen from conventional zigzag, horizontal, and vertical scanning patterns, and then its index is explicitly

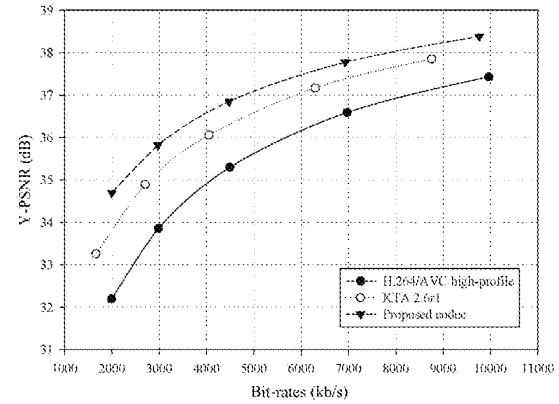


Fig. 6. Rate distortion curve of BasketballDrive 1080p sequence.

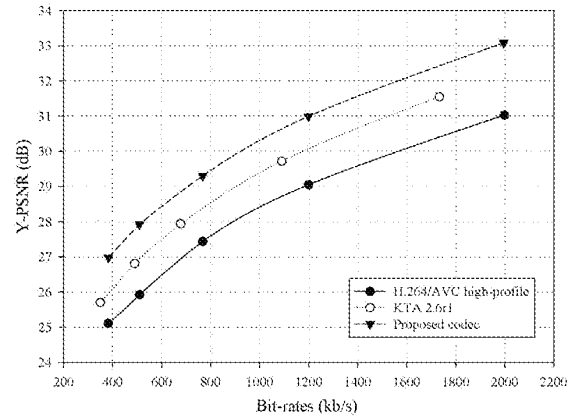


Fig. 7. Rate distortion curve of PartyScene WVGA sequence.

signaled in each TU. Note that the ACS index is signaled only when nonzero AC coefficients exist in the TU, since there is no difference between the patterns when the TU includes only the DC coefficient.

When the horizontal or vertical scanning methods are chosen as the best scanning method, the mapping process from the coefficient array indices to the scanning order indices is not necessary. Thus, the complexity for the coefficient indexing can be reduced and also the memory bandwidth to access the zigzag mapping table can be saved.

VIII. PERFORMANCE ANALYSIS

To verify the performance of the proposed coding tools, they were implemented into one complete video codec software which is available at [14]. This section provides both objective performance based on the peak signal-to-noise ratio (PSNR) measure and the subjective performance based on the mean opinion score (MOS) values for the proposed video codec.

A. Objective Compression Efficiency

1) *Overall Performance*: The coding efficiency of the proposed system is verified using two sets of coding configurations defined in the Call for Proposals [30]—constraint

set 1 (CS1) and constraint set 2 (CS2) which correspond to a random access case and a low delay case, respectively. Figs. 6 and 7 show typical examples of rate-distortion curves when video sequences are coded using CS1 configuration. As shown in the figures, the proposed codec consistently showed significant coding benefits over H.264/AVC High Profile. For comparison with other recent advances in the video coding area, the results of the key technology area (KTA) 2.6r1 [34] with the same test configurations are also included. This test configuration is same as in [35], where all major tools of KTA including enhanced AIF, MV competition, MDDT, high precision filter, ALF, and extended block size are enabled. It should be noted that the improved coding efficiency can be observed with both high resolution and low resolution video material.

Table III summarizes the performance of the proposed codec relative to the H.264/AVC High Profile. The average compression performance improvement over all sequences is 39.49% in CS1 and 39.48% in CS2, respectively. Class B (1080p resolution) achieves 43.29% (CS1) and 43.66% (CS2) bit saving on average, with BQTerrace as the best sequence in both configurations, showing 54.06% (CS1) and 60.62% (CS2) performance improvement, respectively. RaceHorses in Class D (WQVGA resolution) shows the lowest performance improvement in both configurations, at 27.08% in CS1 and 21.59% in CS2, respectively. The performance summarized in Table III shows that the proposed codec consistently achieves the excellent coding performance regardless of the resolution of input video sequences and the coding configuration.

2) *Performance of Each Set of Coding Tools*: It is hard to evaluate the performance of each individual coding tool, since the coding tools interact with each others and the tool-wise evaluation cannot measure the performance improvement introduced by the interaction. Therefore, the proposed coding tools are categorized to five sets—large transform (16×16 , 32×32 , and 64×64 transforms), AMVP, inter (DCT-IF and HAM), intra (ROT, MPI, ADI, CCCP, TMI, and CIP), and in-loop (EXC, ALF, and CADR)—and the performance of each set of coding tools is evaluated by measuring the performance loss introduced by turning off the tools in the corresponding set. The performance of the large motion partition is not evaluated since its performance has been already verified in the literatures [6]–[8]. Table IV summarizes the performance of each set of coding tools and the alternative setting which is enabled when the corresponding tool set is turned off. Note that only 1-second tests using CS1 configuration were carried out to reduce the experiment time. The results show that all the sets of coding tools contribute to the overall performance improvement almost equally.

B. Subjective Compression Efficiency

A large scale subjective test was carried out for all the proposals submitted to the Call for Proposals. In this test, it was shown that the proposed codec provided excellent subjective improvements over H.264/AVC in all configurations and it gave one of the best overall results of all of the proposals that were submitted, based on MOS values. The official report of subjective test results can be found in [36].

TABLE III
SUMMARIZED BD-RATE REDUCTIONS FOR VARIOUS RESOLUTIONS

Class (Size)	Sequence Name	CS1 (%)		CS2 (%)	
		BD-Rate	Average	BD-Rate	Average
A (4K)	Traffic	40.59	35.21	N/A	N/A
	PeopleOnStreet	29.84		N/A	
B (1080p)	Kimono	44.42	43.29	44.45	43.66
	ParkScene	32.96		31.18	
	Cactus	40.78		34.74	
	BasketballDrive	44.21		47.29	
	BQTerrace	54.06		60.62	
C (WVGA)	BasketballDrill	43.09	39.94	36.82	35.58
	BQMall	39.69		37.06	
	PartyScene	41.13		40.45	
	RaceHorses	35.86		27.99	
D (WQVGA)	BasketballPass	30.55	36.45	26.93	32.16
	BQSquare	54.19		51.87	
	BlowingBubbles	33.97		28.24	
	RaceHorses	27.08		21.59	
E (720p)	Vidyo 1	N/A	N/A	48.38	47.48
	Vidyo 3	N/A		46.26	
	Vidyo 4	N/A		47.79	
Total average		39.49		39.48	

TABLE IV
PERFORMANCE OF EACH SET OF CODING TOOLS

Tool Set	Alternative Setting	Class (%)				
		A	B	C	D	Average
Large tr.	Max tr. size = 8×8	2.2	7.2	1.9	1.0	3.5
AMVP	Median predictor [3]	4.4	3.7	2.9	2.4	3.2
Inter	6-tap/bi-linear filter [3]	0.4	1.3	4.2	7.6	3.6
Intra	Directional intra [3]	3.5	3.3	4.4	2.4	3.4
In-loop	Deblocking only	8.8	6.1	3.6	2.9	4.3

Table V shows several pairs of MOS values for H.264/AVC High Profile and the proposed codec for the five 1080p sequences. Each pair has been selected to give similar MOS values, to verify the bit-rate reduction that the proposed codec provides for similar subjective visual quality. As shown in the table, the proposed codec operating at an average of 2.4 Mbit/s provides similar visual quality to H.264/AVC High Profile operating at an average of 6.4 Mbit/s. This corresponds to approximately 60% bit-rate savings, indicating that the proposed codec provides even greater improvement in subjective visual quality than would be expected from objective metrics such as PSNR.

TABLE V
SUMMARIZED MOS VALUES FOR 1080P SEQUENCES

Sequence	H.264/AVC		Proposed		
	Bit rates (kb/s)	MOS	Bit rates (kb/s)	MOS	Reduction (%)
<i>Kimono</i>	4000	8.86	1600	8.56	60.0
<i>ParkScene</i>	4000	7.41	1600	7.78	60.0
<i>Cactus</i>	7000	8.61	3000	8.24	57.1
<i>BasketballDrive</i>	7000	7.44	3000	8.17	57.1
<i>BQTerrace</i>	10000	8.77	3000	9.17	70.0
Average	6400	8.22	2440	8.38	60.9

IX. CONCLUSION

In this paper, a new flexible unit representation composed of CU, PU, and TU was proposed. The proposed scheme provides a generic quadtree structure while defining the relationship between region splitting, prediction method, and transform sizes with minimal signaling overhead.

In addition, various other coding tools were extended in a very consistent way according to the proposed unit representation. For example, the number of directions in the directional intraprediction scheme was extended to arbitrary angles and the supported transform size was extended up to 64×64 .

By combining other novel coding tools designed under the proposed unit representation, the proposed codec showed about 40% performance improvement over the state-of-the-art H.264/AVC standard based on objective PSNR measurements. The large-scale subjective testing results show that an average of about 60% coding gain can be achieved for the 1080p sequences.

The video codec described in this paper was a candidate in the competitive phase of the HEVC standardization work and it has been partially adopted into the first standardization model of the collaborative phase of the HEVC effort.

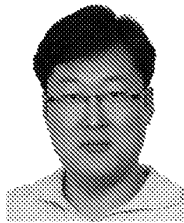
ACKNOWLEDGMENT

The authors would like to thank their colleagues including, but not limited to, H. K. Jung, J. C. Lee, S. R. Lee, D. Y. Kim, K. H. Lee, Y. M. Sohn, Y. Y. Lee, B. K. Lee, S. M. Park, H. S. Song, J. B. Choi, B. D. Choi, C. H. Lee, S. S. Jung, and D. S. Cho for their invaluable contributions and continuous comments on this paper.

REFERENCES

- [1] *Information Technology—Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to About 1.5 Mbit/s: Video*, ISO/IEC 11172-2, ISO/IEC JTC1/SC29/WG11, Aug. 1993.
- [2] *Information Technology—Generic Coding of Moving Pictures and Associated Audio Information: Video*, ISO/IEC 13818-2, ISO/IEC JTC1/SC29/WG11, 1996.
- [3] *Final Draft International Standard of Joint Video Specification*, ITU-T Rec. H.264, ISO/IEC 14496-10, ISO/IEC JTC1/SC29/WG11, Mar. 2003.
- [4] K. H. Lee, E. Alshina, J. H. Park, W. J. Han, and J. H. Min, "Technical considerations for ad hoc group on new challenges in video coding standardization," in *Proc. 85th MPEG Meeting*, Jul. 2008, no. M15580.
- [5] E. Alshina, K. H. Lee, W. J. Han, and J. H. Park, "Technical considerations for ad hoc group on new challenges in video coding standardization," in *Proc. 86th MPEG Meeting*, Oct. 2008, no. M15899.
- [6] P. Chen, Y. Ye, and M. Karczewicz, "Video coding using extended block sizes," in *Proc. 36th VCEG Meeting*, Oct. 2008, no. VCEG-AJ23.
- [7] T. Yoshino, S. Naito, and S. Sakazawa, "Preliminary response for draft call for evidence on high performance video coding," in *Proc. 87th MPEG Meeting*, Feb. 2009, no. M16082.
- [8] S. Sekiguchi and S. Yamagishi, "On coding efficiency with extended block size for UHDTV," in *Proc. 87th MCEG Meeting*, Feb. 2009, no. M16019.
- [9] A. Puri, H. M. Hang, and D. Schilling, "An efficient block-matching algorithm for motion-compensated coding," in *Proc. ICASSP*, vol. 12, Aug. 1997, pp. 1063–1066.
- [10] M. H. Chan, Y. B. Yu, and A. G. Constantinides, "Variable size block matching motion compensation with applications to video coding," *Commun. Speech Vis. IEE Proc. I*, vol. 137, no. 4, pp. 205–212, Aug. 1990.
- [11] G. J. Sullivan and R. L. Baker, "Efficient quadtree coding of images and video," *IEEE Trans. Image Process.*, vol. 3, no. 3, pp. 327–331, May 1994.
- [12] J. J. Zhang, M. O. Ahmad, and M. N. Swamy, "Quadtree structured region-wise motion compensation for video compression," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 5, pp. 808–822, Aug. 1999.
- [13] T. Chujoh, N. Wada, and G. Yasuda, *Quadtree-Based Adaptive Loop Filter*, document VCEG-C181, ITU-T Study Group 16, Jan. 2009.
- [14] K. McCann, W. J. Han, I. K. Kim, J. H. Min, E. Alshina, A. Alshin, T. Lee, J. Chen, V. Seregin, S. Lee, Y. M. Hong, M. S. Cheon, and N. Shlyakhov, "Samsung's response to the call for proposals on video compression technology," in *Proc. 1st JCT-VC Meeting*, Apr. 2010, no. JCTVC-A124.
- [15] JCT of ISO/IEC MPEG & ITU-T VCEG, "Test model under consideration," in *Proc. 1st JCT-VC Meeting*, Apr. 2010, no. JCTVC-A205.
- [16] J. Jung and G. Laroche, "Competition-based scheme for motion vector selection and coding," in *Proc. 29th VCEG Meeting*, Jul. 2006, no. VCEG-AC06.
- [17] T. Wedi, "Hybrid video coding based on high-resolution displacement vectors," in *Proc. VCIP*, Jan. 2001, pp. 186–193.
- [18] C. C. Tseng and S. L. Lee, "Design of fractional delay FIR filter using discrete cosine transform," in *Proc. APCCAS*, Nov. 2008, pp. 858–861.
- [19] T. Wedi and H. G. Musmann, "Motion- and aliasing-compensated prediction for hybrid video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 577–586, Jul. 2003.
- [20] S.-D. Kim and J. B. Ra, "An efficient motion vector coding scheme based on minimum bitrate prediction," *IEEE Trans. Image Process.*, vol. 8, no. 8, pp. 1117–1120, Aug. 1999.
- [21] S. H. Lee and N. I. Cho, "Intra prediction method based on the linear relationship between the channels for YUV 4:2:0 intra coding," in *Proc. ICIP*, 2009, pp. 1037–1040.
- [22] T. K. Tan, C. S. Boon, and Y. Sujuki, "Intra prediction by template matching," in *Proc. ICIP*, 2006, pp. 1693–1696.
- [23] J. Balle and M. Wien, "Extended texture prediction for H.264/AVC intra coding," in *Proc. ICIP*, 2007, pp. VI-93–VI-96.
- [24] S. Ma and C.-C. J. Kuo, "High-definition video coding with super-macroblocks," in *Proc. VCIP*, Jan. 2007.
- [25] J. Dong, K. N. Ngan, C. K. Fong, and W. K. Cham, "2-D order-16 integer transforms for HD video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 10, pp. 1462–1474, Oct. 2009.
- [26] W.-H. Chen, C. H. Smith, and S. Fralick, "A fast computational algorithm for the discrete cosine transform," *IEEE Trans. Commun.*, vol. 25, no. 9, pp. 1004–1009, Sep. 1977.
- [27] J. Liang and T. D. Tran, "Fast multiplierless approximations of the DCT with the lifting scheme," *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 3032–3044, Dec. 2001.
- [28] Y. Ye and M. Karczewicz, *Improved Intra Coding*, Contribution C257, ITU-T Q.6/SG16, Geneva, Switzerland, Jun. 2007.
- [29] Y. Ye and M. Karczewicz, *Improved Intra Coding*, VCEG Contribution VCEG-AG11, no. 20, Shenzhen, China, Oct. 2007.
- [30] ISO/IEC JTC1/SC29/WG11, "Joint call for proposals on video compression technology," in *Proc. 91st MPEG Meeting*, Apr. 2010, no. N11113.
- [31] T. Chujoh and R. Noda, "Internal bit depth increase for coding efficiency," in *Proc. 31st VCEG Meeting*, Jan. 2007, no. VCEG-AE13.

- [32] D. Marpe, H. Schwarz, and T. Wiegand, "Context-based adaptive binary arithmetic coding in the H.264/AVC video compression standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 620–636, Jul. 2003.
- [33] *Coding of Still Pictures*, ISO/IEC IS 10918-1, ITU-T Rec. T.81, ISO/IEC JTC1/SC29/WG1, Sep. 1993.
- [34] ITU-T VCEG. (2010). *Key Technology Area (KTA) Software* [Online]. Available: <http://iphome.hhi.de/suehring/ttml/download/KTA>
- [35] T. K. Tan, G. Sullivan, and T. Wedi, *Recommended Simulation Common Condition for Coding Efficiency Experiments Revision 1*, VCEG Contribution VCEG-AJ10, Jul. 2008.
- [36] V. Baroncini, J. R. Ohm, and G. Sullivan, "Report of subjective test results of responses to the joint call for proposals (cfp) on video coding technology for high efficiency video coding (HEVC)," in *Proc. 1st JCT-VC Meeting*, Apr. 2010, no. JCTVC-A204.



Woo-Jin Han (M'02) received the M.S. and Ph.D. degrees in computer science from the Korea Advanced Institute of Science and Technology, Daejeon, Korea, in 1997 and 2002, respectively.

He is currently a Principal Engineer with the Multimedia Platform Laboratory, Digital Media and Communication Research and Development Center, Samsung Electronics, Suwon, Korea. Since 2002, he has contributed successfully to the ISO/IEC Moving Pictures Experts Group, Joint Video Team, and Joint

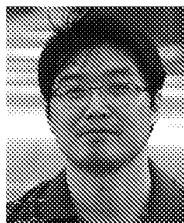
Collaborative Team standardization efforts. His current research interests include high-efficiency video compression techniques, scalable video coding, multi-view synthesis, and visual contents understanding.



Junghye Min received the B.S. degree in mathematics from Ewha Women's University, Seoul, Korea, in 1995, and the M.E. and Ph.D. degrees in computer science and engineering from Pennsylvania State University, University Park, in 2003 and 2005, respectively.

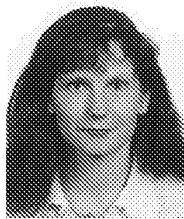
From 1995 to 1999, she was a S/W Engineer with Telecommunication Systems Business, Samsung Electronics, Suwon, Korea. She is currently a Principal Engineer with the Multimedia Platform Laboratory, Digital Media and Communications Research and Development Center, Samsung Electronics. Her current research

interests include video content analysis, motion tracking, pattern recognition, and video coding.



Il-Koo Kim received the M.S. and Ph.D. degrees from the School of Electrical Engineering, Seoul National University, Seoul, Korea, in 2001 and 2006, respectively.

After joining the Digital Media and Communications Research and Development Center, Samsung Electronics, Suwon, Korea, in 2006, he has worked to develop high efficiency video codec. His current research interests include video compression and transmission.



Elena Alshina received the M.S. degree in physics from Lomonosov Moscow State University (MSU), Moscow, Russia, in 1995. She continued her post-graduate study at the Faculty of Computational Mathematics and Cybernetics, MSU, and received the Ph.D. degree in mathematical modeling in 1998.

Since that time, she has been with the Institute for Mathematical Modeling, Russian Academy of Sciences, Moscow, performing research on high accuracy numerical methods and new finite-difference scheme development. Simultaneously, she was an

Associate Professor for the Moscow Institute of Electronic Technology, Moscow. In 2006, she joined Samsung Electronics (first at the Moscow Research Center, before moving to Korea). Currently, she is a Senior Engineer

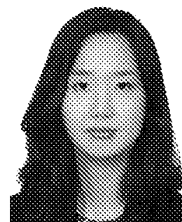
with the Multimedia Platform Laboratory, Digital Media and Communications Research and Development Center, Samsung Electronics, Suwon, Korea, where she makes high efficiency video codecs.



Alexander Alshin received the M.S. and Ph.D. degrees in mathematical physics from Moscow State University (MSU), Moscow, Russia, in 1995 and 1998, respectively.

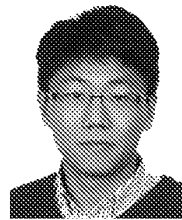
From February 1998 to January 2006, he was a Senior Researcher with the Physical Department, MSU. Since 2006, he has been with Samsung Electronics, Suwon, Korea, and currently works with the Multimedia Platform Laboratory, DMC Research and Development Center, Samsung Electronics, as a Senior Engineer. His current research interests

include theory of non-classical PDEs, nonlinear functional analysis, applied mathematics, numerical analysis, and their applications to video compression.



Tammy Lee received the B.S. degree in mathematics education from Ewha Women's University, Seoul, Korea, in 2000, and the M.S. degree in computer science from New York University, New York, in 2005.

She is currently a Senior Engineer with the Multimedia Platform Laboratory, Digital Media and Communications Research and Development Center, Samsung Electronics, Suwon, Korea. Her current research interests include image and video compression.



Jianle Chen received the B.S. and Ph.D. degrees in information and communication system from Zhejiang University, Hangzhou, China, in 2001 and 2006, respectively.

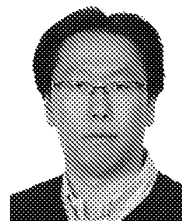
Currently, he is a Senior Engineer with the Multimedia Platform Laboratory, Digital Media and Communications Research and Development Center, Samsung Electronics, Suwon, Korea. His current research interests include video/image compression, computer vision, and image/video processing.



Vadim Seregin received the M.S. (honors) degree in physics from Moscow State University (MSU), Moscow, Russia, and the Ph.D. degrees in mathematical modeling, and numerical methods and software systems from MSU, in 2002 and 2005 respectively.

Currently, he is a Senior Engineer with the Multimedia Platform Laboratory, Digital Media and Communications Research and Development Center, Samsung Electronics, Suwon, Korea. His current research interests include quasi-gasdynamics, numerical methods, and the field of image and video

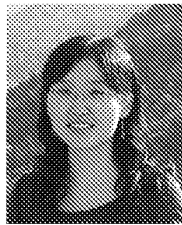
processing with a focus on video compression.



Sunil Lee (S'02–M'08) received the B.S. degree in electrical engineering from Yonsei University, Seoul, Korea, in 2001, and the M.S. and Ph.D. degrees in electrical engineering and computer science from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 2002 and 2008, respectively.

From March 2008 to August 2008, he was a Post-Doctoral Researcher with the Division of Electrical Engineering, School of Electrical Engineering and Computer Science, KAIST. He is currently a Senior

Engineer with the Multimedia Platform Laboratory, Digital Media and Communications Research and Development Center, Samsung Electronics, Suwon, Korea. His current research interests include multimedia retrieval, multimedia security, digital watermarking, multirate signal processing, and video coding.



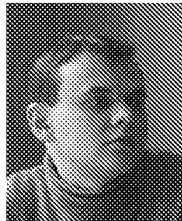
Yoon Mi Hong received the B.S. degree in mathematics from Yonsei University, Seoul, Korea, in 2003, and the M.S. degree in applied mathematics and the Ph.D. degree in mathematical sciences from the Korea Advanced Institute of Science and Technology, Daejeon, Korea, in 2005 and 2009, respectively.

She is currently a Senior Engineer with the Multimedia Platform Laboratory, Digital Media and Communications Research and Development Center, Samsung Electronics, Suwon, Korea. Her current research interests include time-frequency analysis, sampling theory, and its applications in signal processing, as well as video coding.



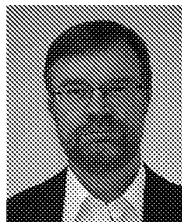
Min-Su Cheon received the B.S. and M.S. degrees in electrical engineering from Kyunghee University, Suwon, Korea, in 2005 and 2007, respectively.

He is currently an Engineer with the Multimedia Platform Laboratory, Digital Media and Communications Research and Development Center, Samsung Electronics, Suwon. His current research interests include high-efficient video coding, multiview video coding, 3-D video coding, scalable video coding, and bioimage processing.



Nikolay Shlyakhov received the M.S. degree in applied mathematics from the Moscow Institute of Physics and Technology (MIPT), Moscow, Russia, in 1999. From 1999 to 2007, he was pursuing the Ph.D. degree from the Institute for Mathematical Modeling of the Russian Academy of Sciences, Moscow.

From 2003 to 2007, he was an Assistant Professor with the Department of Higher Mathematics, MIPT. Since 2007, he has been an Engineer with the Multimedia Platform Laboratory, Digital Media and Communications Research and Development Center, Samsung Electronics, Suwon, Korea. His current research interests include numerical mathematics, theory of approximation, and video coding.

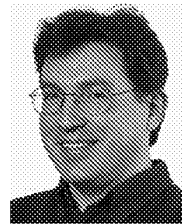


Ken McCann graduated from Glasgow University, Glasgow, U.K., with the B.S. degree in natural philosophy in 1979. He received the M.S. degree in systems engineering from Surrey University, Surrey, U.K., in 1982.

He is currently a Director and Co-Founder of ZetaCast, Bath, U.K., an independent U.K.-based technology consultancy company specializing in digital TV and related systems. Prior to founding ZetaCast, he was the Technology Director with NTL Broadcast (now Arqiva), the main provider of broadcasting

infrastructure in the U.K. Before that, he was with Symbionics Instruments, Cambridge, U.K., the Independent Broadcasting Authority and Philips Research Laboratories. He contributed significantly to the development of the MPEG-2 standard, has chaired the DVB technical group responsible for audio-visual coding specification since its inception over 15 years ago, and is currently an active member of the Joint Collaborative Team on Video Coding, working on the new high efficiency video coding standard.

Mr. McCann has been a member of the IET in the U.K. (and its predecessor, the IEE) since 1982.



Thomas Davies graduated from Oxford University, Oxford, U.K., with the M.A. degree in mathematics and philosophy in 1991. He received the M.S. and Ph.D. degrees in mathematics from Warwick University, Coventry, U.K., in 1992 and 1996, respectively.

He is currently a Lead Research and Development Engineer with Research and Development, British Broadcasting Corporation, London, U.K., having joined in 2000. He specializes in video compression, and image and signal processing, and was responsible for developing the open source Dirac video compression algorithms, the intra coding part of which is standardized as SMPTE VC-2. He is actively participating in the development of the MPEG high efficiency video coding standard under the auspices of the Joint Collaborative Team on Video Coding. Prior to joining the BBC, he was a Technology Consultant in systems engineering and satellite networking.



Jeong-Hoon Park received the B.S. and M.S. degrees in electrical engineering from Hanyang University, Seoul, Korea, in 1991 and 1994, respectively.

He is currently a Principal Engineer with the Multimedia Platform Laboratory, Digital Media and Communication Research and Development Center, Samsung Electronics, Suwon, Korea. In 1998, he was a Visiting Scholar with the University of California, Los Angeles, where he researched error resilient video compression and transmission. Since then, he has contributed to the ITU-T Video Coding Experts Group and ISO/IEC Moving Pictures Experts Group. From 2002 to 2004, he has been involved in developing mobile broadcasting technology as an active member of WorldDMB forum. His current research interests include audio and video compression technology, medical image processing, and multimedia systems including video broadcasting and communications.

<i>Index of Claims</i> 	Application/Control No. 13666680	Applicant(s)/Patent Under Reexamination BICI ET AL.
	Examiner CLIFFORD HILAIRE	Art Unit 2488

✓	Rejected	-	Cancelled	N	Non-Elected	A	Appeal
=	Allowed	÷	Restricted	I	Interference	O	Objected

☐ Claims renumbered in the same order as presented by applicant			☐ CPA			☐ T.D.			☐ R.1.47		
CLAIM		DATE									
Final	Original	06/26/2015									
	1	✓									
	2	✓									
	3	✓									
	4	✓									
	5	✓									
	6	✓									
	7	✓									
	8	✓									
	9	✓									
	10	✓									
	11	✓									
	12	✓									
	13	✓									
	14	✓									
	15	✓									
	16	✓									
	17	✓									
	18	✓									
	19	✓									
	20	✓									

Receipt date: 04/22/2013

13666680 -- GAU: 2488

Doc code: IDS

PTO/SB/08a (01-10)

Doc description: Information Disclosure Statement (IDS) Filed

Approved for use through 07/31/2012. OMB 0651-0031

U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)	Application Number		13666680	
	Filing Date		2012-11-01	
	First Named Inventor	Mehmet Oguz BICI		
	Art Unit	2631		
	Examiner Name	TBD		
	Attorney Docket Number	NC77198US-NP		

U.S. PATENTS							Remove	
Examiner Initial*	Cite No	Patent Number	Kind Code ¹	Issue Date	Name of Patentee or Applicant of cited Document	Pages, Columns, Lines where Relevant Passages or Relevant Figures Appear		
	1							
If you wish to add additional U.S. Patent citation information please click the Add button.							Add	
U.S. PATENT APPLICATION PUBLICATIONS							Remove	
Examiner Initial*	Cite No	Publication Number	Kind Code ¹	Publication Date	Name of Patentee or Applicant of cited Document	Pages, Columns, Lines where Relevant Passages or Relevant Figures Appear		
	1	20120307905	A1	2012-12-06	Sunyeon et al.			
	2	20110170602	A1	2011-07-14	Tammy et al.			
If you wish to add additional U.S. Published Application citation information please click the Add button.							Add	
FOREIGN PATENT DOCUMENTS							Remove	
Examiner Initial*	Cite No	Foreign Document Number ³	Country Code ² j	Kind Code ⁴	Publication Date	Name of Patentee or Applicant of cited Document	Pages, Columns, Lines where Relevant Passages or Relevant Figures Appear	T ⁵
	1	2011062392	WO	A2	2011-05-26	SK Telecom Co. Ltd. et al.		☐
If you wish to add additional Foreign Patent Document citation information please click the Add button							Add	
NON-PATENT LITERATURE DOCUMENTS							Remove	

Receipt date: 04/22/2013

13666680 - GAU: 2488

**INFORMATION DISCLOSURE
STATEMENT BY APPLICANT**
(Not for submission under 37 CFR 1.99)

Application Number	13666680
Filing Date	2012-11-01
First Named Inventor	Mehmet Oguz BICI
Art Unit	2631
Examiner Name	TBD
Attorney Docket Number	NC77198US-NP

Examiner Initials*	Cite No	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc), date, pages(s), volume-issue number(s), publisher, city and/or country where published.	T ⁵
	1	International Search Report and Written Opinion received for corresponding Patent Cooperation Treaty Application No. PCT/FI2012/051070, dated March 27, 2013, 12 pages.	☐
	2	Oudin, S. et al.; "Block Merging for Quadtree-Based Video Coding", IEEE Int. Conf. on Multimedia and Expo, July 11-15, 2011, 6 pages.	☐
	3	Sullivan, G. J.; "Overview of the High Efficiency Video Coding (HEVC) Standard", IEEE Trans. on Circuits and Systems for Video Technology, Vol. 22, No. 12, December 2012, pp. 1649-1668.	☐

If you wish to add additional non-patent literature document citation information please click the Add button **Add**

EXAMINER SIGNATURE

Examiner Signature	/CLIFFORD HILAIRE/	Date Considered	06/24/2015
--------------------	--------------------	-----------------	------------

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through a citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

¹ See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. ² Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). ³ For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. ⁴ Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. ⁵ Applicant is to place a check mark here if English language translation is attached.

Receipt date: 04/22/2013

13666680 - GAU: 2488

**INFORMATION DISCLOSURE
STATEMENT BY APPLICANT**
(Not for submission under 37 CFR 1.99)

Application Number	13666680
Filing Date	2012-11-01
First Named Inventor	Mehmet Oguz BICI
Art Unit	2631
Examiner Name	TBD
Attorney Docket Number	NC77198US-NP

CERTIFICATION STATEMENT

Please see 37 CFR 1.97 and 1.98 to make the appropriate selection(s):

- ☐ That each item of information contained in the information disclosure statement was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(1).

OR

- ☐ That no item of information contained in the information disclosure statement was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the person signing the certification after making reasonable inquiry, no item of information contained in the information disclosure statement was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(2).

- ☐ See attached certification statement.
- ☐ Fee set forth in 37 CFR 1.17 (p) has been submitted herewith.
- ☒ None

SIGNATURE

A signature of the applicant or representative is required in accordance with CFR 1.33, 10.18. Please see CFR 1.4(d) for the form of the signature.

Signature	/Thomas J. Arria/	Date (YYYY-MM-DD)	2013-04-22
Name/Print	Thomas J. Arria	Registration Number	60223

This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 1 hour to complete, including gathering, preparing and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these records.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

EAST Search History

EAST Search History (Prior Art)

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
S3	9	((Mehmet) near2 (BICI)).INV.	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/24 17:00
S4	85	((Jani) near2 (LAINEMA)).INV.	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/24 17:00
S5	59	((Kemal) near2 (UGUR)).INV.	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/24 17:00
S6	4	("20110170602" "20120307905").PN.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2015/06/25 09:41
S7	2	("20110170602" "20120307905").PN.	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 09:41
S8	9	((Mehmet) near2 (BICI)).INV.	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 09:43
S9	85	((Jani) near2 (LAINEMA)).INV.	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 09:43
S10	59	((Kemal) near2 (UGUR)).INV.	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 09:43
S11	104	S8 S9 S10	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 09:43
S12	45	S11 and merg\$4	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 09:43
S13	12	S11 and merg\$4 near list\$4	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 09:44
S14	45	S11 and merg\$4	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 09:44
S15	18	S11 and merg\$4 same candidate	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 10:10
S16	67	S11 and candidate	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 10:11
S17	63	S11 and candidate with motion	US-PGPUB; USPAT;	OR	ON	2015/06/25 10:13

			USOCR			
S18	55	S11 and candidate with motion with vector	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 10:13
S19	45	S11 and candidate with motion with vector same block	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 10:14
S20	28	S11 and (candidate and list and vector)	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 10:15
S21	5805	S11 and (candidate and list and vector)".clm"	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 10:15
S22	4	S11 and (candidate and list and vector).clm.	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 10:15
S23	1	"20130114723".pn.	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 10:52
S24	1	"20130114723".pn. and (candidate with vector with motion) same (merg\$4 set list subset)	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 10:53
S25	1812	(candidate with vector with motion) same (merg\$4 set list subset)	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 10:55
S26	27634	((H04N19/52 OR H04N19/513 OR H04N19/176 OR H04N19/61 OR H04N19/51 OR H04N19/597 OR H04N19/56 OR H04N19/139 OR H04N19/573 OR H04N19/521 OR H04N19/103 OR H04N19/30).CPC.)	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 10:57
S27	1	"20130034162".pn.	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 11:08
S28	1	"20130034162".pn. and (candidate with motion with vector)	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 11:16
S29	1	"20130034162".pn. and (candidate with motion with vector) same block	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 11:17
S30	1	"20130034162".pn. and (candidate)	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 11:18
S31	1	"20130034162".pn. and (candidate) and (list set subset)	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 11:19
S32	1	"20130034162".pn. and (candidate same block) and (list set subset)	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 11:19
S33	1274	(candidate with vector with motion) same (merg\$4 set list subset) and (motion vector) with search\$4	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 11:32
S34	1064	(candidate with vector with motion) same (merg\$4 set list subset) and (motion vector) near4 search\$4	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 11:32
S35	888	(candidate with vector with motion) with (merg\$4 set list subset) and (motion	US-PGPUB; USPAT;	OR	ON	2015/06/25 11:32

		vector) near4 search\$4	USOCR			
S36	618	(candidate with vector with motion) near4 (merg\$4 set list subset) and (motion vector) near4 search\$4	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 11:33
S37	112	(candidate with vector with motion) near4 (merg\$4 set list subset) same (motion vector) near4 search\$4	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 11:33
S38	92	(candidate with vector with motion) near4 (merg\$4 set list subset) same (motion vector) near4 search\$4 and "375".clas.	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 11:33
S39	150592	(candidate with vector with motion) near4 (merg\$4 set list subset) same (motion vector) near4 search\$4 same block "375".clas.	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 11:33
S40	51	(candidate with vector with motion) near4 (merg\$4 set list subset) same (motion vector) near4 search\$4 same block and "375".clas.	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 11:34
S41	8	(candidate with vector with motion) near4 (merg\$4 set list subset) same (motion vector) near4 search\$4 same block and subset and "375".clas.	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 11:42
S42	2	(candidate with vector with motion) near4 (merg\$4 set list subset) same (motion vector) near4 search\$4 same block and subset and (remov\$4 exclud\$4) with (list set merg\$4 subset)	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 11:45
S43	50	("5719627" "5815602" "4989089" "5428403" "5781249" "5838391" "4851906" "4989087" "5235419" "5371549" "5414469" "5457481" "5506622" "5521642" "5532746" "5546129" "5557341" "5574663" "5579050" "5581308" "5587741" "5598216" "5612743" "5614954" "5617144" "5619281" "5625417" "5627591" "5638129" "5642166" "5646867" "5654761" "5689306" "5694487" "5717463" "5719630" "5724369" "5731851" "5751362" "5862261" "5903672" "5991447" "6005980" "6020925" "6023298" "6026195" "6057884" "6091777" "6154519" "6205177").pn.	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 13:29
S44	25	S43 and (candidate same motion)	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 13:30
S45	0	S43 and (candidate same motion) and merg\$4	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 13:30
S46	4	S43 and (candidate same motion) and exclud\$4	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 13:31
S47	4	S43 and (candidate) with (vector motion) and exclud\$4	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/25 13:32
S48	3	(US-20130034162-\$ or US- 20120106645-\$ or US-20120106638-	US-PGPUB	OR	ON	2015/06/25 13:44

		\$.did.				
S49	0	S48 and "PU "	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2015/06/25 13:44
S50	2	S48 and prediction with unit	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2015/06/25 13:44
S52	9	((Mehmet) near2 (BICI)).INV.	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/26 06:02
S53	85	((Jani) near2 (LAINEMA)).INV.	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/26 06:02
S54	59	((Kemal) near2 (UGUR)).INV.	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/26 06:02
S55	104	S52 S53 S54	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/26 06:02
S56	55	S55 and candidate with motion with vector	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/26 06:02
S57	10	S55 and candidate with motion with vector and transcod\$4	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/26 06:03
S58	8	S55 and candidate with motion with vector and transcod\$4 and encod\$4	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/26 06:04
S59	53	S55 and candidate with motion with vector and encod\$4	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/26 06:05
S60	29	S55 and candidate with motion with vector same encod\$4	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/26 06:05
S61	24	S55 and candidate with motion with vector same encod\$4 and unit and (block macroblock)	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/26 06:07
S62	18	S55 and candidate with motion with vector same encod\$4 and (unit near predict\$5) and (block macroblock)	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/26 06:10
S63	23	S55 and prediction near unit	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/26 06:11
S64	22	S55 and prediction near unit and candidate	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/26 06:11
S65	872	prediction near unit and candidate near4 motion near4 vector	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/26 06:14
S66	713	prediction near unit and candidate near4	US-PGPUB;	OR	ON	2015/06/26

		motion near4 vector same block	USPAT; USOCR			06:14
S67	653	prediction near unit and candidate near4 motion near4 vector with block	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/26 06:14
S68	172	prediction near unit same candidate near4 motion near4 vector with block	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/26 06:14
S69	42	prediction near unit same candidate near4 motion near4 vector with block same encoded	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/26 06:15
S70	32	("prediction unit" PU) with block with encoded same candidate near4 motion near4 vector	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/26 06:17
S71	483	PU and candidate with motion with vector	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/26 06:35
S72	209	PU same candidate with motion with vector	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/26 06:36
S73	160	PU same candidate with motion with vector same block	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/26 06:36
S74	142	PU same candidate with motion with vector same block and "375".clas.	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/26 06:36
S75	1	(US-20120134415-\$).did.	US-PGPUB	OR	ON	2015/06/26 07:05
S76	0	S75 and encoded	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/26 07:05
S77	0	S75 and encod\$4 with block	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/26 07:05
S78	1	S75 and coded with block	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/26 07:05
S79	126	PU same candidate with motion with vector same block and merge with list	US-PGPUB; USPAT; USOCR	OR	ON	2015/06/26 07:06
S80	1	(US-20120008688-\$).did.	US-PGPUB	OR	ON	2015/06/26 08:47
S81	1	(US-20120008688-\$).did.	US-PGPUB	OR	ON	2015/06/26 08:52
S82	1	(US-20120008688-\$).did. and (PU prediction near unit)	US-PGPUB	OR	ON	2015/06/26 08:53
S83	1	(US-20120008688-\$).did. and (PU prediction near unit) and receiv\$4	US-PGPUB	OR	ON	2015/06/26 08:54
S84	1	(US-20120008688-\$).did. and (PU prediction near unit) and (encod\$4 cod\$4 decod\$4 receiv\$4)	US-PGPUB	OR	ON	2015/06/26 08:55

EAST Search History (Interference)

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp

EAST Search History

S51	0	"20130034162".pn. and (candidate)	USPAT; UPAD	OR	ON	2015/06/25 11:18
-----	---	---------------------------------------	----------------	----	----	---------------------

6/ 26/ 2015 1:09:19 PM
C:\ Users\ chilaire\ Documents\ EAST\ Workspaces\ 13666680.wsp

A Multi-Pass True Motion Estimation Scheme With Motion Vector Propagation for Frame Rate Up-Conversion Applications

Shen-Chuan Tai, Ying-Ru Chen, Zheng-Bin Huang, and Chuen-Ching Wang

Abstract—This paper presents a robust true motion estimation algorithm, designated as MPMVP (Multi-pass and Motion Vector Propagation), to enhance the accuracy of the motion vector fields in frame rate up-conversion applications. The MPMVP uses a multi-pass scheme to progressively refine approximate motion vectors to true motion vectors based upon the motion information acquired in previous passes. The multi-pass motion estimation process uses a large block size to detect the motion vectors within the objects themselves and small block sizes to detect the motion vectors along the object boundaries. Actually, the block size is progressively reduced during the search process. When the motion vector of a block is considered to be sufficiently accurate for motion estimation purposes, the block is said to be converged and the local motion vector search process terminates. A novel technique, referred to as motion vector propagation, is then applied to propagate the motion vector of the converged block to its neighboring blocks. This technique not only ensures that neighboring motion vectors within the same object have a high degree of spatial correlation, but also accelerates the convergence of the motion vectors in the neighboring blocks and therefore reduces the overall computational time and expense of the multi-pass motion vector search procedure. A novel distortion criterion is proposed to enhance the tolerance of the traditional sum-of-absolute-difference measurement technique applied in the motion estimation scheme to noise and shadow effects. The experimental results demonstrate that the proposed true motion estimation algorithm outperforms the traditional full search, 3DRS and TCSBP algorithms in terms of both the smoothness of the generated motion vector fields and the visual quality of the up-converted frames.

Index Terms—Candidate motion vector, distortion criterion, frame rate up-conversion (FRUC), recursive search, true motion, variable block-size.

I. INTRODUCTION

MODERN video formats have a wide variety of resolutions, aspect ratios, frame rates, and so forth. Typically, the frame rates of capturing devices vary between 24 and 60 Hz. However, the frame rates used by most display devices tend to range from 50 to 120 Hz, for example. Therefore, when playing back video sequences, it is necessary to interpolate some addi-

tional frames based on the original video sequence using some form of frame rate up-conversion (FRUC) algorithm. However, traditional schemes such as picture repetition, linear interpolation and other non-motion compensated methods generally produce visual artifacts such as trembling or blurring in the interpolated frames. The quality of the interpolated frames can be improved by combining the FRUC scheme with motion estimation (ME) techniques. Many complex algorithms for interpolating frames based upon the motion information of objects between successive frames have been proposed [1]–[15]. However, they tend to be highly time consuming and computationally expensive.

The ME schemes used in FRUC algorithms differ from the conventional ME techniques applied in coding standards. Generally speaking, ME algorithms used in frame rate up-conversion applications are termed “true ME” schemes in order to emphasize the importance of obtaining accurate motion vectors for all of the objects in the frame. True ME is employed not only in FRUC schemes, but also in such diverse applications as video encoding, [16], surveillance system tracking [17], de-interlacing schemes [18], error concealment in decoding systems [19], [20], and so forth. Briefly, the objective of video coding systems is to minimize the motion compensated residual. In block-matching ME algorithms, the features of moving objects are tracked using distortion criteria such as the sum of absolute difference (SAD) or the sum of square error (SSE). However, video sequences tend to contain high levels of noise, and, therefore, the motion vectors (MVs) extracted by minimal SAD/SSE-based coding systems do not always accurately reflect the true motions of the objects in the frame. In FRUC schemes, it is essential to obtain accurate motion vectors to ensure the quality of the interpolated frames, and therefore the ME algorithms used in such schemes require a robust ability to deal with the effects of noise. Unfortunately, the ME algorithms conventionally used in block-matching coding systems, e.g., full search, logarithmic search, diamond search, and so on, fail to address this requirement.

In frame rate up-conversion, the optimal image quality is obtained when the objects in the frames are interpolated along their respective motion trajectories. 3-D recursive search block matching (3DRS) [4] and temporal compensated motion estimation with simple block-based prediction (TCSBP) [2] have proposed ME algorithms based on predictive motion vectors (PMVs) to accelerate the block-matching speed and to enhance the degree of spatial correlation between neighboring blocks. In [3], the authors used an overlapped block-based motion estimation (OBME) scheme possessing enhanced ME capabil-

Manuscript received August 8, 2006; revised November 12, 2006, April 23, 2007, and October 20, 2007.

S.-C. Tai, Y.-R. Chen, and Z.-B. Huang are with the Department of Electrical Engineering, National Cheng Kung University, Tainan 701, Taiwan, R.O.C. (e-mail: cyr91d@ccmc.ee.ncku.edu.tw).

C.-C. Wang is with the Institute of Electrical Engineering, National Changhua University of Education, Changhua City 500, Taiwan, R.O.C.

Color versions of one or more figures are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JDT.2007.916014

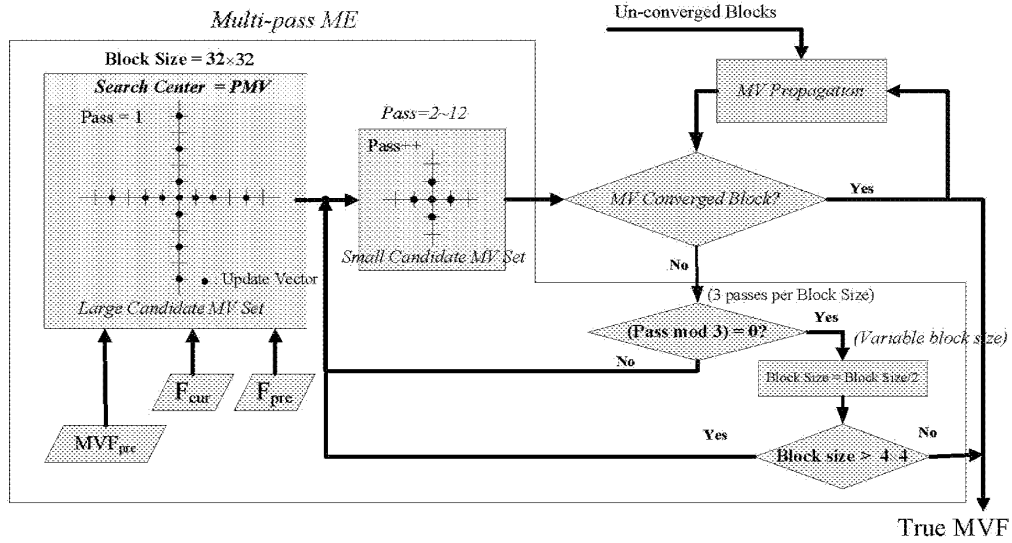


Fig. 1. Flow diagram of overall system.

ities and the ability to detect global motion through the use of a large-block matching mechanism. However, without accurate and sufficient information regarding the object shape, OBME schemes are unable to predict the motion of the object boundaries with any degree of accuracy, and therefore result in an incorrect motion aggregation in the motion vector field (MVf). A fundamental requirement of all true ME schemes is to achieve a high degree of spatial motion correlation inside a frame since all of the blocks within the same object must necessarily move in the same direction. Although several algorithms have been presented [4]–[6], [13] which employ some form of penalty function to enhance motion correlation, they are incapable of resolving the problem of incorrect motion propagation. In [10], [11], the MVs were established using hierarchical block-matching algorithms and were then propagated between hierarchical levels based on the regions of the objects. However, obtaining the regions of the objects requires a time-consuming segmentation process and the accuracy of the segmentation results cannot be guaranteed. In [14], a variable block size technique was used to enhance the ME accuracy. In this approach, the MVs were initially found for small blocks and were then used as the criteria for merging neighboring blocks. Having merged neighboring blocks, the MV search procedure was then repeated. However, this method is not only computationally expensive, but also achieves only limited success since small blocks cannot provide sufficient information regarding the global characteristics of an object and thus the reliability of the merging process cannot be ensured.

Accordingly, this paper presents a robust true ME scheme for FRUC applications. The remainder of this paper is organized as follows. Section II presents an overview of the major components of the proposed scheme, including the multiple-pass, the MV convergence decision technique, the MV convergence propagation scheme, and the distortion criterion used to suppress the effects of noise and shadow in the interpolated frames. Section III presents the experimental results obtained by the proposed true ME scheme for the MVFs of several well-known video test sequences, and compares the performance of the pro-

posed scheme with that of the full search, 3DRS and TCSBP schemes when applied to frame rate up-conversion applications. Finally, Section IV presents some brief conclusions.

II. MULTI-PASS MOTION ESTIMATION STRATEGY AND MV PROPAGATION RULE FOR TRUE MOTION ESTIMATION

The proposed ME scheme searches for the true MVs using a multi-pass strategy in which the block size is progressively reduced from 32×32 to 4×4 . After each pass, a convergence decision rule is applied to determine whether or not the MV identified for the block is sufficiently accurate. If the MV has acceptable accuracy, the block is said to be converged and the local MV search procedure is terminated. As indicated above, the largest block size used in the multi-pass ME scheme is 32×32 . When an object is bigger than this block size, a technique referred to as MV propagation can be applied to spread the MV of the converged block to all of the blocks within the object. Fig. 1 presents an overall view of the multi-pass MV search procedure and associated MV propagation routine.

A. Multi-Pass Motion Estimation Strategy

The basic principle of the multi-pass ME strategy proposed in this study is to utilize the motion information relating to neighboring blocks generated in the previous pass to gradually refine the accuracy of the MV in the current block. The multi-pass search scheme executes a maximum of 12 passes. In the first three passes, the MV search is performed using a block size of 32×32 . However, in the next three passes, the block size is reduced to 16×16 , while in the final two sets of three passes, the block size is further reduced to 8×8 and 4×4 , respectively. The MV for block $\vec{X}$ found in frame t in the i th pass is defined as $\vec{MV}^i(\vec{X}, t)$. The MV search in block $\vec{X}$ in the i th pass is performed in accordance with (1). As shown, the search is confined to the vectors $\vec{V}$ defined in the candidate vector set $CS^i(\vec{X}, t)$, i.e., a full search is not performed. As discussed later, the term

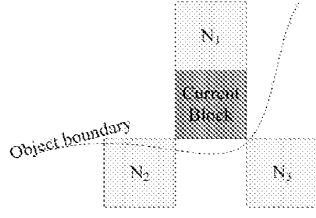


Fig. 2. Geometric relation between motion edge and block position.

$e(\vec{V}, \vec{X}, t)$ in (1) is a new distortion criterion based upon the conventional SAD distortion criterion.

$$\overrightarrow{MV}^i(\vec{X}, t) = \arg \min_{\vec{V} \in CS^i(\vec{X}, t)} (e(\vec{V}, \vec{X}, t)), \quad 1 \leq i \leq 12. \quad (1)$$

The candidate vector set $CS^i(\vec{X}, t)$ is comprised by a single PMV and a set of update vectors. Let $\vec{X}$ represent the block location and let t be the frame number. The candidate vector set for this particular block in the i^{th} pass are then defined as

$$CS^i(\vec{X}, t) = \left\{ \vec{V} \mid \vec{V} = \overrightarrow{PMV}^i(\vec{X}) + \vec{U}, \vec{U} \in US^i \right\} \quad (2)$$

where $\overrightarrow{PMV}^i(\vec{X})$ is the PMV of block $\vec{X}$ and is chosen from the MVs of the neighboring blocks and US^i is a set of update vectors. Every motion vector $\vec{V}$ in the candidate vector set is comprised by a single PMV and a update vector $\vec{U}$ selected from update vector set US^i .

1) *PMV*: Although the MVs of all eight neighboring blocks can theoretically be used as PMV predictors for the current block, evaluating each of these predictors in order to establish the most suitable PMV among them is computationally expensive and time consuming. Therefore, this study uses the geometric relationship proposed in [2] describing the correlation between the current block position and the motion edge to reduce the number of PMV predictors to be evaluated. As shown in Fig. 2, at least one of the three neighboring blocks and the current block must belong to the same object no matter where the object boundary crosses, and, therefore, the number of PMV predictors which must be considered can be reduced from 8 to just 3. The scheme presented in [2] involved a single ME pass, and therefore employed a single PMV predictor pattern. However, as described above, the current search scheme executes a maximum of 12 passes. In general, different PMV predictor patterns may result in different levels of accuracy. Therefore, the current multi-pass ME scheme considers 12 different PMV predictor patterns, as shown in Fig. 3.

Each pass in the proposed ME scheme is performed using one of the PMV predictor patterns (see Fig. 3) and then one MV in the selected PMV pattern which has minimal distortion is selected as the final PMV of this pass. In traditional single-pass ME algorithms, the current block has no information regarding the motion of the succeeding blocks because their MVs are not obtained yet. However, in multi-pass ME algorithms, the motion information of all blocks in current pass can be referred in future passes.

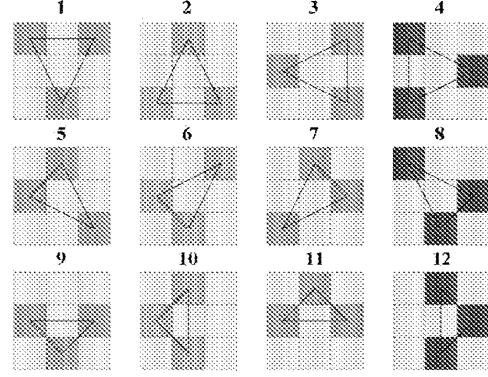


Fig. 3. PMV predictor patterns.

2) *Update Vector Set*: In (2), the term US^i represents the set of update vectors in the i^{th} pass and has the form

$$US^i = \begin{cases} US_{\text{large}}, & i = 1 \\ US_{\text{small}}, & 2 \leq i \leq 12 \end{cases}$$

$$US_{\text{small}} = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \pm 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \pm 1 \end{pmatrix} \right\}$$

$$US_{\text{large}} = US_{\text{small}} \cup \left\{ \begin{pmatrix} \pm 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \pm 3 \end{pmatrix}, \begin{pmatrix} \pm 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \pm 5 \end{pmatrix} \right\}. \quad (3)$$

The update vector sets employed in the proposed algorithm vary between different passes. The proposed scheme uses a large update vector set (i.e., a greater number of larger MVs) in the initial passes when employing a large block size to detect coarse motions in the frame. When the update vectors have a large range, e.g., $\pm(5,5)$, finding good approximations to the true MVs in blocks which cross object boundaries is straightforward. Furthermore, in the first pass, the motion information of neighboring blocks is simply copied from the previous frame and therefore there is no guarantee that this information is reliable. As a result, the PMV obtained for the block is not necessarily an accurate representation of the true MV. However, using a large update vector set increases the probability of obtaining a good approximation to the true MV. In later passes of the search procedure, a highly precise PMV with a small update vector set (i.e., fewer and smaller vectors) is used to refine these initial non-true motion estimates into true motion estimates.

As indicated earlier, the term $e(\vec{V}, \vec{X}, t)$ in (1) is a new distortion criterion based on the conventional SAD distortion criterion and has the form

$$e(\vec{V}, \vec{X}, t) = \sum_{i,j \in \text{block}} \text{weightAD}(\vec{V}, \vec{X}_{ij}, t) / \text{blocksize}$$

$$\text{AD}(\vec{V}, \vec{X}_{ij}, t) = \text{abs}(I(\vec{X}_{ij}, t) - I(\vec{X}_{ij} + \vec{V}, t - T))$$

$$\text{weight} = \begin{cases} 0, & \text{AD}(\vec{V}, \vec{X}_{ij}, t) < 5 \\ 1, & 5 \leq \text{AD}(\vec{V}, \vec{X}_{ij}, t) < 30 \\ 2, & \text{otherwise} \end{cases} \quad (4)$$

where $I(\vec{X}_{ij}, t)$ is the gray level of a pixel located at coordinates (i,j) in block $\vec{X}$ of frame t and abs indicates absolute value. According to Weber's law, human eyes have a noise threshold

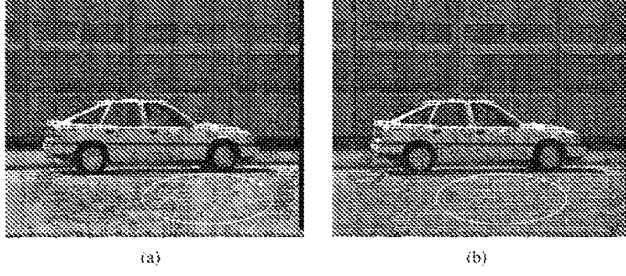


Fig. 4. MVFs of frames with Gaussian noise: (a) MVF produced by proposed algorithm with traditional SAD distortion criterion, and (b) MVF produced by proposed algorithm with new distortion criterion.

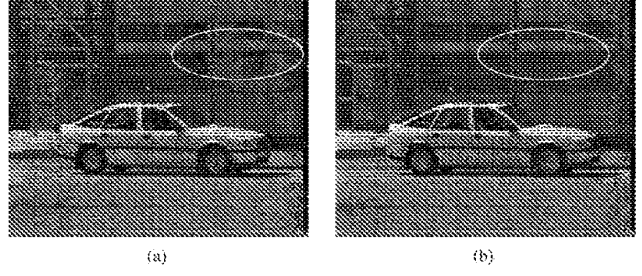


Fig. 5. Up-converted frames from VECTRA sequence. (a) MVF produced by 3DRS with penalty function. (b) MVF produced by 3DRS with proposed distortion criterion.

of approximately 5 grayscale value (for contrast) in pictures. For experience, an edge often leads to a gradient greater than 30 grayscale value. Therefore, if the absolute difference (AD) between current pixel and referred pixel in previous frame is less than 5, the human eye will ignore the distortion. However, if it is greater than 30, the distortion will be very noticeable. Hence, as shown in (4), the distortion criterion applied in the current ME search scheme has the form of a piecewise linear function. In general, noise has a particularly serious effect on the accuracy of the MV search process in smooth areas of the frame, i.e., background area. Therefore, the piecewise linear function in (4) allows small distortions ($AD < 5$) caused by noise to be ignored while simultaneously placing more emphasis on the large distortions ($AD \geq 30$) caused by the edges, and therefore the accuracy of the MV is enhanced.

Fig. 4 shows the MVFs obtained by the proposed ME scheme using the conventional SAD distortion criterion and the new distortion criterion given in (4), respectively, for a frame with Gaussian noise. Comparing the two frames, it is apparent that the improved noise suppression characteristics of the proposed distortion criterion yield more accurate MVs vectors in the smooth regions of the image, i.e., on the grass. Fig. 5 presents two up-converted frames from the VECTRA test sequence obtained by the 3DRS search scheme using the original penalty function and the new distortion criterion, respectively. It can be seen that the MVF produced by 3DRS using the new distortion criterion is more smooth. For example, in Fig. 5(b) more motion vectors of the window area indicate to the same direction than in Fig. 5(a). In other words, the proposed distortion criterion is applicable not only to the multi-pass scheme presented in the current study, but also to existing true ME algorithms.

B. Decision Rules for MV Converged Block

To reduce the computational complexity and expense of the proposed ME search scheme, processing within a local block terminates as soon as the block is converged, irrespective of the number of passes which have been completed. Furthermore, the degree of motion correlation among neighboring blocks is enhanced by using a MV propagation technique that will be described in Section II-C. The degree of spatial motion correlation between neighboring blocks is determined in accordance with

$$\Delta MV(\vec{C}, \vec{N}) = \text{abs}(MV(\vec{C}, t).x - MV(\vec{N}, t).x) + \text{abs}(MV(\vec{C}, t).y - MV(\vec{N}, t).y) \quad (5)$$

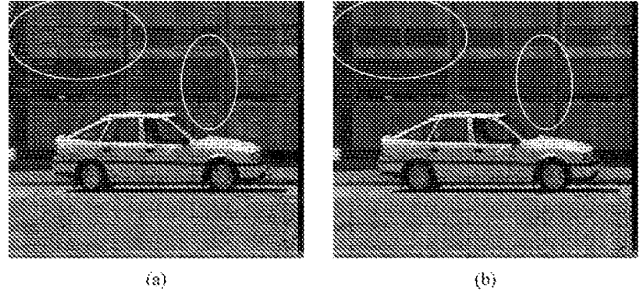


Fig. 6. (a) Proposed ME scheme without convergence propagation and (b) proposed ME scheme with convergence propagation.

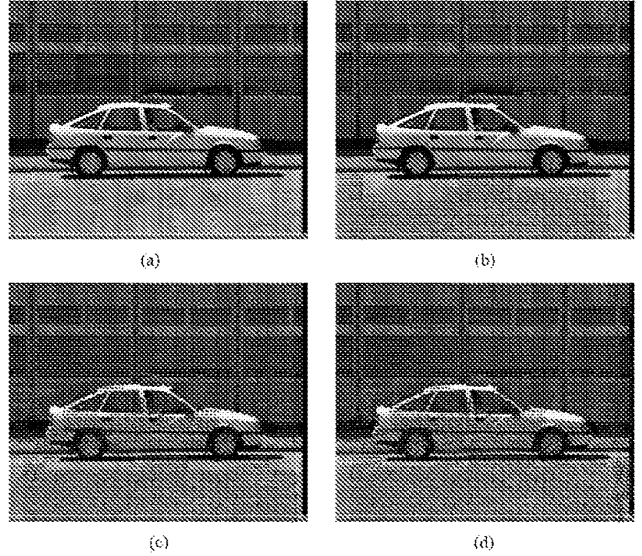


Fig. 7. Motion vectors of converged Blocks (MV is shown by block-size 4×4). (a) 3rd Pass: block size is 32×32 , (b) 6th Pass: block size is 16×16 , (c) 9th Pass: block size is 8×8 , (d) 12th Pass: block size is 4×4 .

where $\vec{C}$ and $\vec{N}$ are the current block and a neighboring block in spatial motion correlation test range, respectively, and x and y are the horizontal and vertical MV components. True motion is obtained when the distortion value is low, i.e., $e(\vec{C}, \vec{X}, t) < T_1$ and the degree of motion correlation between the current block and all of its adjacent blocks in the spatial motion correlation

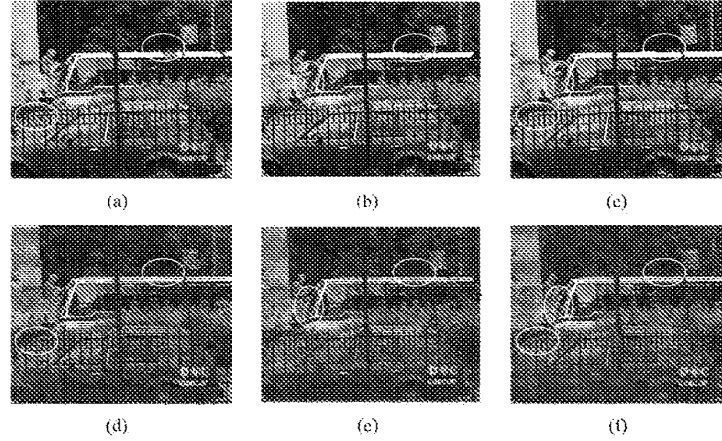


Fig. 8. (a), (d) Interpolated frame from BUS sequence and corresponding MVF obtained from 3DRS. (b), (e) Interpolated frame from BUS sequence and corresponding MVF obtained from TCSBP. (c), (f) Interpolated frame from BUS sequence and corresponding MVF obtained from proposed MPMVP algorithm.

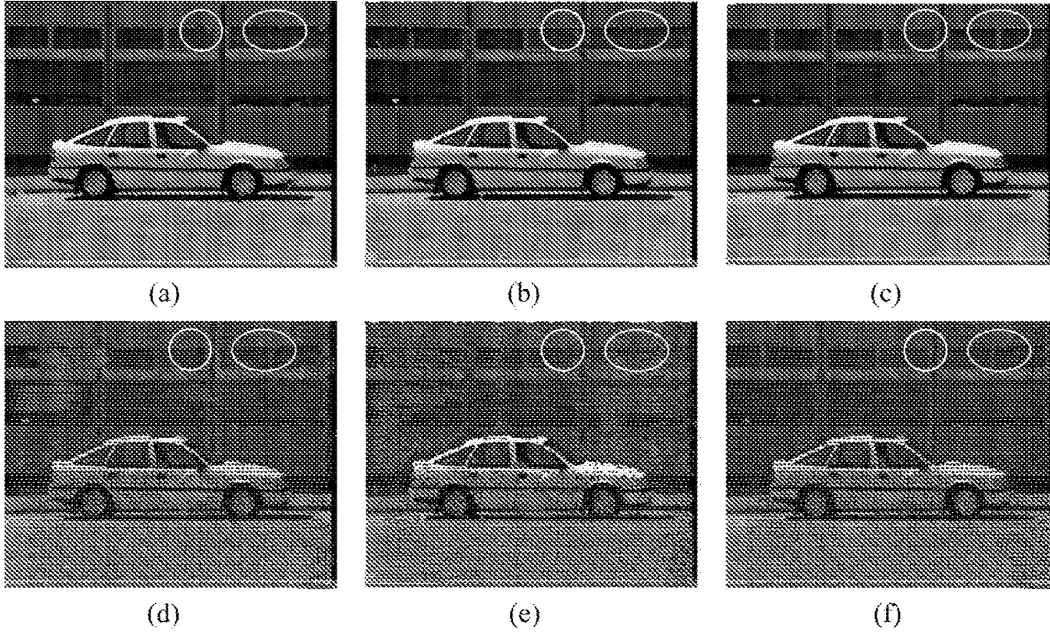


Fig. 9. (a), (d) Interpolated frame from VECTRA sequence and corresponding MVF obtained from 3DRS. (b), (e) Interpolated frame from VECTRA sequence and corresponding MVF obtained from TCSBP. (c), (f) Interpolated frame from VECTRA sequence and corresponding MVF obtained from proposed MPMVP algorithm.

test range is high, i.e., $\Delta MV(\vec{C}, \vec{N}) < T_2$, where T_1 is a distortion threshold and T_2 is a threshold value designed to enable sufficient activity for the MV search to prevent the clustering of inappropriate MVs. In other words, the current scheme considers both the distortion value and the degree of motion correlation when searching for true MV. This approach differs from existing ME algorithms, which typically terminate the MV evaluation process as soon as an acceptable distortion value has been obtained.

The threshold for the distortion value T_1 in i th pass is set as $5 + i/3$ in accordance with experience. Furthermore, the spatial motion correlation test ranges used in the current algorithm are 3×3 and 5×5 , respectively. Since there are three passes during specific-block-size ME, the larger test range is applied in the first two passes, while the smaller test range is used in the final

pass. Meanwhile, T_2 is assigned a value of 0 for block sizes larger than 8×8 and a value of 1 for smaller blocks. Note that this threshold is deliberately assigned a stricter value for larger blocks to prevent them from converging prematurely.

C. MV Propagation Method

The MV propagation technique employed in the current scheme is designed to enhance the degree of motion correlation in large textureless regions of the frame. When a block converges, i.e., it has a high degree of spatial motion correlation and a low distortion, its MV is automatically offered to the neighboring unconverged blocks as a potential MV for their own motions. If the difference between the original distortion value of the neighboring block and the new distortion value of the block produced by accepting the propagated MV is

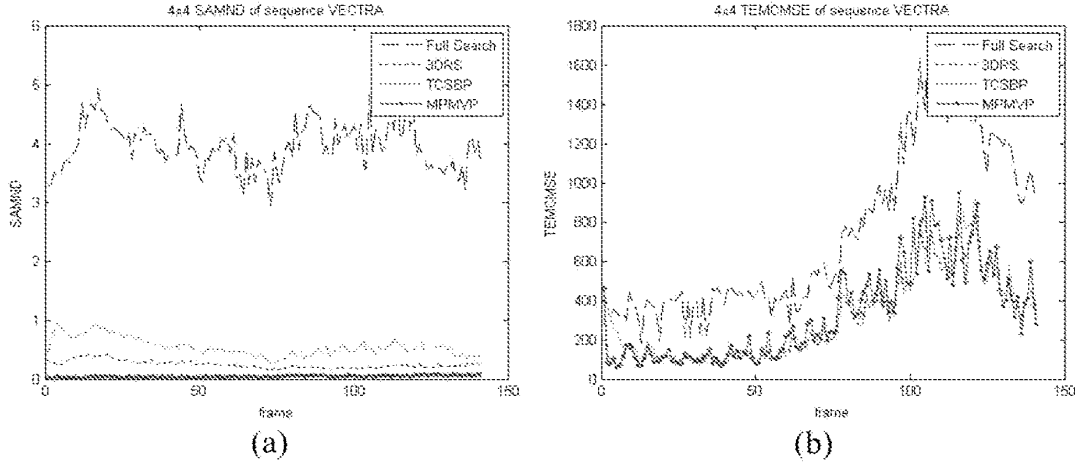


Fig. 10. SAMND and TEMC-MSE comparisons for full search (± 32), 3DRS, TCSBP and proposed algorithm. (Note results relate to VECTRA test sequence).

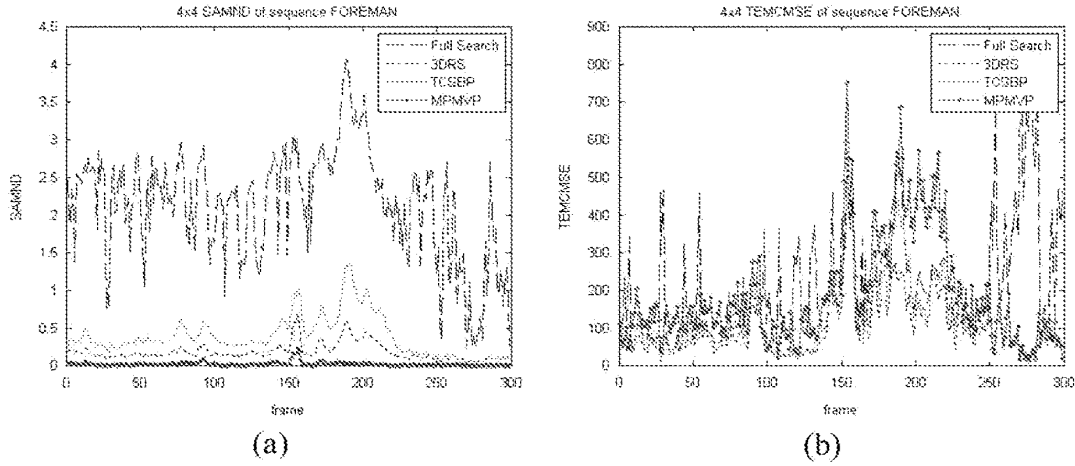


Fig. 11. SAMND and TEMC-MSE comparisons for full search (± 16), 3DRS, TCSBP and proposed algorithm. (Note results relate to FOREMAN test sequence).

sufficiently small, i.e., $15 - 2/3 \cdot i$, the neighboring block accepts the propagated MV as its own MV. Having accepted the propagated MV, the block then checks to see whether or not the MV causes it to become converged. When a block has converged, irrespective of whether or not convergence is induced by the propagated MV, the same MV is propagated to the neighbors of the new converged block.

The convergence propagation technique is used to expand size of the region, which has the similar MV, into the size of object. Convergence propagation ensures that the MVs of the background regions of a frame have a high degree of spatial correlation. In other words, the background appears as a single, continuous region with no obvious areas of demarcation, even across strong edge areas, as shown in Fig. 6.

III. EXPERIMENTAL RESULTS

In the figures discussed in the following, the MVFs and interpolated frames are based on a block size of 4×4 . For example, a MV of block size 32×32 will be presented by 64 MVs of block size 4×4 . Each MV is colored in red, originates from the center of the 4×4 block and terminates with a solid blue point. Fig. 7 presents the MVs of converged blocks obtained by the proposed

ME scheme with different block sizes. The MVs of the background regions of the frame are found in the earlier passes of the ME search process performed using a larger block size. The use of a variable block size in the search procedure enables the block size of MV to gradually fit the contour of the objects in each pass and therefore the accuracy of the identified MVs is greatly enhanced. As a result, Fig. 7 demonstrates that the algorithm successfully divides the frame into two distinct objects, namely the background and the moving car.

In general, interpolated (i.e., up-converted) frames are generated using a motion compensation technique. Some uncovered regions (i.e., regions with no MVs) can be interpolated based on neighboring available pixels. Since the areas of these uncovered regions usually are small, we use spatial bilinear interpolation to interpolate uncovered regions in motion compensation step of this paper. Figs. 8 and 9 present the MVFs and corresponding interpolated frames obtained using the 3DRS [4], TCSBP [2] and proposed MPMVP ME algorithms for the BUS and VECTRA video sequences, respectively. Note that in every case, the test sequences are in a CIF format, have an original frame rate of 30 frames/s, and are up-converted to a rate of 60 frames/s. Observing these figures, it is apparent that the 3DRS and TCSBP

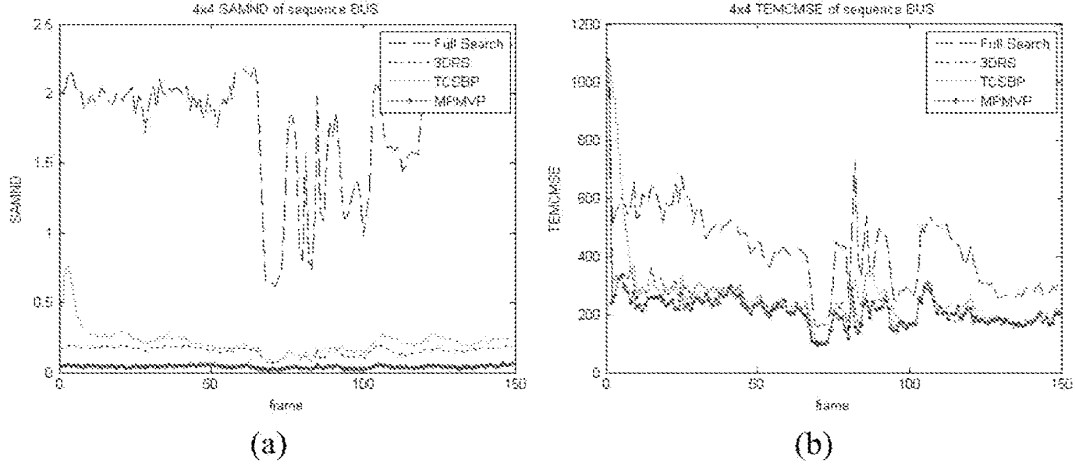


Fig. 12. SAMND and TEMC-MSE comparisons for full search (± 16), 3DRS, TCSBP and proposed algorithm. (Note results relate to BUS sequence).

schemes generate incorrect MVs in the smooth, expansive background regions of the frame because they fail to propagate the MVs efficiently among neighboring blocks.

The quality of the interpolated frames produced in FRUC applications is generally assessed using subjective visual quality measure. However, in [2], the authors proposed two new evaluation criteria with which to measure the motion correlation of interpolated frames. The first criterion was the spatial consistency indicator (i.e., the sum of absolute minimum neighboring difference, designated as SAMND), while the second criterion related to the temporal continuity of the motion field between successive frames (i.e., temporal extending motion compensation mean square error, designated as TEMC-MSE). SAMND evaluates the MV differences between neighboring blocks in accordance with

$$\text{SAMND} = \sum_{\vec{X} \in \text{MV}_t} \min(\text{abs}(\text{Vec}_{\vec{X},t} - \text{Vec}_{\vec{X}+\delta,t})) \quad (6)$$

where $\vec{X}$ is the block location, Vec is the MV of the block, and δ is the 3×3 neighboring region. Meanwhile, the TEMC-MSE criterion measures the temporal MV ductility and focuses on the smooth regions of the frame such as the background where the MVs in the previous frame are also valid for the current frame. The TEMC-MSE criterion has the form

$$\text{TEMC-MSE} = \frac{\sum_{\vec{P}} (I(\vec{P},t) - I(\vec{P} - \text{Vec}_{\vec{P},t-T},t-T))^2}{M \times N} \quad (7)$$

where

- I luminance component of a frame;
- $\vec{P}$ pixel position;
- $M \times N$ total number of sampling pixels in the frame;
- $\text{Vec}_{\vec{P},t-T}$ MV of position $\vec{P}$ at frame $t-T$, and can be estimated from the luminance fields $I(\vec{P},t-T)$ and $I(\vec{P},t-2T)$.

Figs. 10–12 compares the SAMND and TEMC-MSE metrics for the full search, 3DRS, TCSBP and proposed MPMVP for

the VECTRA sequence (characterized by fast and complicated motions), FOREMAN sequence (characterized by low motions) and BUS sequences (characterized by complicated motions), respectively. In general, Figs. 10–12 show that the MVs generated by the proposed algorithm have the highest degree of spatial correlation of all the MVs produced by the four algorithms. However, in Fig. 11(b), the proposed algorithm has the poorest TEMC-MSE performance. The poor performance arises because the proposed distortion criterion deliberately ignores small distortion values in order to enhance the noise resistance. However, when checking the up-converted FOREMAN sequence by human eyes, it is found that the proposed algorithm generates the highest quality interpolated frames of the four algorithms. For example, Fig. 13 shows frames 154 and 155 of the original FOREMAN sequence together with the corresponding interpolated frames generated using the full search, 3DRS, TCSBP and proposed MPMVP schemes, respectively. It can be seen the visual quality of MPMVP is the best. [Note that these particular frames are deliberately chosen here because frame 154 has a high TEMC-MSE value in Fig. 11(b)]. These figures also show that the full search scheme has the poorest performance of the four algorithms in terms of both the SAMND and the TEMC-MSE metrics since it is not specifically designed to focus on motion correlation.

Although the ME scheme proposed in this study is a recursive motion search method, the convergence decision rules and MV propagation technique greatly reduce its computational time and complexity. For full search, the computation is decided by search range, i.e., for search range n , the computation is $(2n+1)^2$ ADs/pixel, where n is often set as 7, 15, 31, or even 63. For a worst-case in MPMVP, the computation is 101 ADs/pixel for MPMVP. Because in the first pass, the proposed algorithm takes the motion vector in the previous field as the PMV and then evaluates the 13 search points defined in US_{large} [see (3)]. In the following 11 passes, the algorithm evaluates 3 PMV predictors and each of the 5 search points defined in US_{small} . Therefore, the maximal computational complexity is $13 + (11 \times (3+5)) = 101$ ADs/pixel. But in general cases, the actually evaluated search points are much less than 101 ADs/pixel.

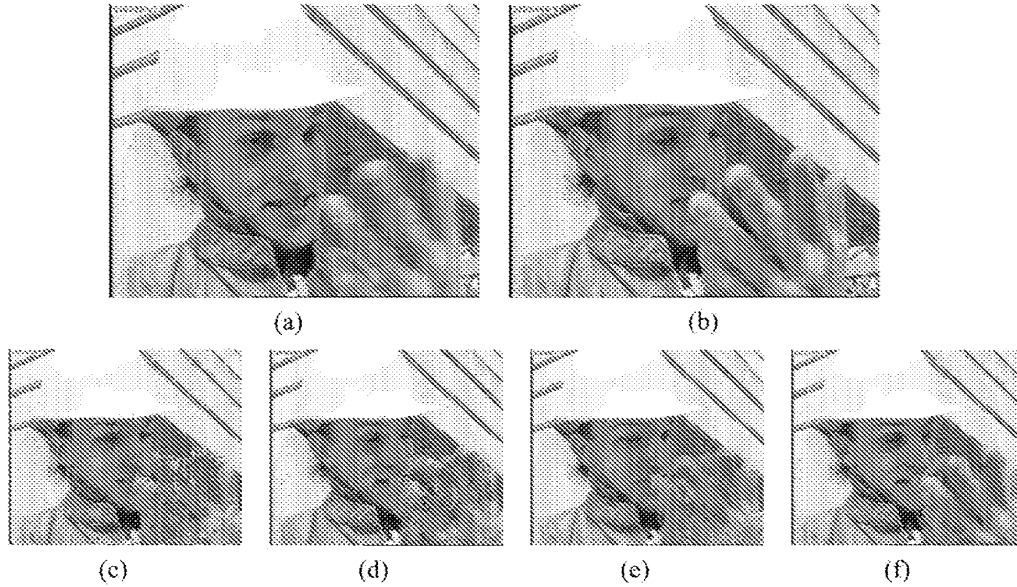


Fig. 13. (a), (b) Frames 154 and 155 in FOREMAN sequence. (c)-(f) Frames interpolated by full search, 3DRS, TCSBP and proposed algorithm, respectively.

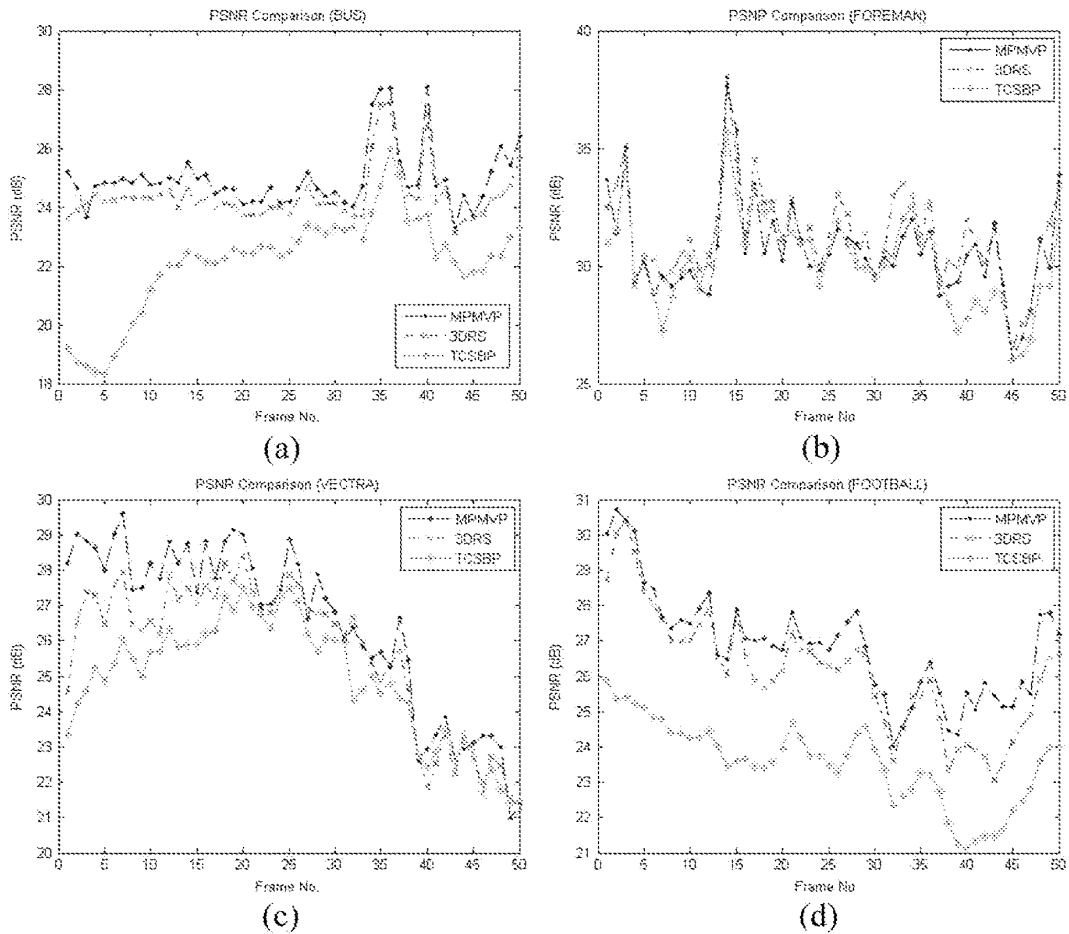


Fig. 14. PSNR comparisons on up-converted test sequences. (a) BUS. (b) FOREMAN. (c) VECTRA. (d) FOOTBALL. (a), (b), (c) are in CIF format and (d) is in de-interlaced D1 format.

TABLE I
MOTION ESTIMATION TIMES SPENT IN 3DRS, TCSBP AND PROPOSED
MPMVP ALGORITHM

Sequence \ Time (sec/frame)	3DRS	TCSBP	MPMVP
BUS (CIF)	0.317	0.037	0.145
FOREMAN (CIF)	0.315	0.036	0.074
VECTRA (CIF)	0.305	0.038	0.116
FOOTBALL (D1)	2.186	0.264	1.143

We have tested the motion search time and PSNR performance for 3DRS, TCSBP and the proposed MPMVP scheme for three CIF format test sequences. The frame rate of these sequences are firstly down-sampled from 30 to 15 fps and then up-converted to 30 fps. We calculate the PSNR between the up-converted frames and the original frames in 30 fps sequence. The PSNR curves of the first 50 up-converted frames are shown in Fig. 14, while the corresponding motion estimation times are listed in Table I. In order to test the performance of the proposed MPMVP in larger resolution sequences, one de-interlaced D1 sequence (50 fps), FOOTBALL, is down-sampled from 50 to 25 fps and then up-converted to 50 fps. The experimental result of FOOTBALL is also shown in Fig. 14 and Table I. From those data, it can be seen that using our MPMVP scheme in frame rate up-conversion applications not only can achieve higher quality up-converted frames, but also can save significant computation.

IV. CONCLUSION

This paper has presented a new highly-accurate, low-complexity ME scheme for FRUC applications. The proposed algorithm is based upon a multi-pass, variable block-size approach and features a new noise-resistant distortion criterion to improve the quality of the interpolated frames. Furthermore, efficient MV convergence criteria and MV propagation techniques are employed to reduce the overall computational time and cost of the scheme. The experimental results have shown that the proposed ME scheme not only achieves a higher MV accuracy than existing ME schemes, but also yields significant computational savings. The quality of our true MV generated up-converted frames are better than other ME algorithm generated frames.

REFERENCES

- [1] B.-T. Choi, S.-H. Lee, and S.-J. Ko, "New frame rate up-conversion using bi-directional motion estimation," *IEEE Trans. Consumer Electron.*, vol. 46, no. 3, pp. 603–609, Aug. 2000.
- [2] J. Wang, D. Wang, and W. Zhang, "Temporal compensated motion estimation with simple block-based prediction," *IEEE Trans. Broadcast.*, vol. 49, no. 3, pp. 241–248, Sep. 2003.
- [3] H. Taehyeun, L. Seongjoo, and K. Jaeseok, "Motion compensated frame interpolation by new block-based motion estimation algorithm," *IEEE Trans. Consumer Electron.*, vol. 50, no. 2, pp. 752–759, May 2004.
- [4] G. de Haan, P. W. A. C. Biezen, H. Huijgen, and O. A. Ojo, "True-motion estimation with 3-D recursive search block matching," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 3, no. 5, pp. 368–379, Oct. 1993.
- [5] R. Braspenning and G. de Haan, "True-motion estimation using feature correspondence," *SPIE, Proc. VCIP*, pp. 396–407, Jan. 2004.
- [6] G. de Haan and P. W. A. C. Biezen, "An efficient true-motion estimator using candidate vectors from a parametric motion model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, no. 1, pp. 85–91, Feb. 1998.
- [7] A. Pelagotti and G. de Haan, "A new algorithm for high quality video format conversion," in *2001 Proc. Int. Conf. on Image Process.*, Oct. 2001, vol. 2, pp. 375–378.
- [8] S.-H. Lee, Y.-C. Shin, Y. Seungjoon, H.-H. Moon, and R.-H. Park, "Adaptive motion-compensated interpolation for frame rate up-conversion," *IEEE Trans. Consumer Electron.*, vol. 48, no. 3, pp. 444–450, Aug. 2002.
- [9] S.-H. Lee and S. Yang, "Adaptive motion-compensated frame rate up-conversion," *Electron. Lett.*, vol. 38, no. 10, pp. 451–452, May 2002.
- [10] B.-W. Jeon, G.-I. Lee, S.-H. Lee, and R.-H. Park, "Coarse-to-fine frame interpolation for frame rate up-conversion using pyramid structure," *IEEE Trans. Consumer Electron.*, vol. 49, no. 3, pp. 499–508, Aug. 2003.
- [11] D. Wang and D. Lauzon, "Hybrid algorithm for estimating true motion fields," *Opt. Eng.*, vol. 39, no. 11, pp. 2876–2881, Nov. 2000.
- [12] K. Hilman, H.-W. Park, and Y.-M. Kim, "Using motion compensated frame-rate conversion for the correction of 3:2 pulldown artifacts in video sequences," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, no. 6, pp. 869–877, Sep. 2000.
- [13] P. Ishwar and P. Moulin, "On spatial adaptation of motion-field smoothness in video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, no. 6, pp. 980–989, Sep. 2000.
- [14] S. Fujiwara and A. Taguchi, "Motion-compensated frame rate up-conversion based on block matching algorithm with multi-size blocks," in *Proc. 2005 Int. Symp. on Intell. Signal Process. Commun. Syst.*, Dec. 2005, vol. 13–16, pp. 353–356.
- [15] B.-D. Choi, J.-W. Han, C.-S. Kim, and S.-J. Ko, "Frame rate up-conversion using perspective transform," *IEEE Trans. Consumer Electron.*, vol. 52, no. 3, Aug. 2006.
- [16] O. EePing, H. Wang, and P. Xue, "Video coding based on true motion estimation," in *Proc. 2003 IEEE Int. Conf. on Acoust., Speech, and Signal Process.*, Apr. 2003, vol. 3, no. 6–10, pp. 409–412.
- [17] J. E. Santos Conde, A. Teuner, and B. J. Hosticka, "Hierarchical locally adaptive multigrid motion estimation for surveillance applications," in *Proc. 1999 IEEE Int. Conf. on Acoust., Speech, Signal Process.*, Mar. 1999, vol. 6, pp. 3365–3368.
- [18] G. de Haan, "IC for motion-compensated de-interlacing, noise reduction, and picture-rate conversion," *IEEE Trans. Consumer Electron.*, vol. 45, no. 3, pp. 617–624, Aug. 1999.
- [19] J.-W. Suh and Y.-S. Ho, "Error concealment technique based on optical flow," *Electron. Lett.*, vol. 38, no. 18, pp. 1020–1021, Aug. 2002.
- [20] J.-Y. Pyun, J.-S. Lee, J.-W. Jeong, J.-H. Jeong, and S.-J. Ko, "Robust error concealment for visual communications in burst-packet-loss networks," *IEEE Trans. Consumer Electron.*, vol. 49, no. 4, pp. 1013–1019, Nov. 2003.



Shen-Chuan Tai received the B.S. and M.S. degrees in electrical engineering from National Taiwan University, Taipei, Taiwan, in 1982 and 1986, respectively, and received the Ph.D. degree in computer science from National Tsing Hua University, Hsinchu, Taiwan, R.O.C., in 1989.

He is currently a professor in the Department of Electrical Engineering, National Cheng Kung University, Tainan, Taiwan. His teaching and research interests include data compression, DSP, VLSI array processors, computerized electrocardiogram processing, and multimedia systems.



Ying-Ru Chen received the B.S. degree in education from the Department of Information and Computer Education, National Taiwan Normal University, Taipei, Taiwan, R.O.C., in 2000, and the M.S. degree in electrical engineering from National Cheng Kung University, Tainan, Taiwan, R.O.C., in 2002, and where she is currently working toward the Ph.D. degree candidate in the Department of Electrical Engineering, National Cheng Kung University, Tainan, Taiwan, R.O.C.

Her research interests include video signal processing, data compression, multimedia systems, and digital watermarking.



Chuen-Ching Wang received B.S., M.S., and Ph.D. degrees in electrical engineering from the National Cheng Kung University, Taiwan, R.O.C., in 1986, 1994, and 2003, respectively.

He was a section chief at Directorate General Telecommunications in the Ministry of Transportation and Communications. He is currently an assistant professor at the institute of electrical engineering, National Changhua University of Education, Taiwan, R.O.C. His research interests include digital signal processing, image processing, computer network, mobile computing and multimedia communications.



Zheng-Bin Huang received the B.S. degree in electrical engineering and the B.S. degree in mechanical engineering from National Cheng Kung University, Tainan, Taiwan, R.O.C., both in 2004, and the M.S. degree in electrical engineering from the National Cheng Kung University, Tainan, Taiwan, R.O.C., in 2006.

His research interests include video compression, image processing, and fast true motion estimation algorithms.

Search Notes 	Application/Control No. 13666680	Applicant(s)/Patent Under Reexamination BICI ET AL.
	Examiner CLIFFORD HILAIRE	Art Unit 2488

CPC- SEARCHED		
Symbol	Date	Examiner
H04N19/52; H04N19/513; H04N19/176; H04N19/61; H04N19/51; H04N19/597; H04N19/56; H04N19/139; H04N19/573; H04N19/521; H04N19/103; H04N19/30	6/25/2015	CH

CPC COMBINATION SETS - SEARCHED		
Symbol	Date	Examiner

US CLASSIFICATION SEARCHED			
Class	Subclass	Date	Examiner
375	240	6/25/2015	CH

SEARCH NOTES		
Search Notes	Date	Examiner
See Attached EAST Search History Document	6/25/2015	CH
IEEEXplore Search	6/25/2015	CH

INTERFERENCE SEARCH			
US Class/ CPC Symbol	US Subclass / CPC Group	Date	Examiner

/CLIFFORD HILAIRE/ Examiner.Art Unit 2488	
--	--

To: thao.pham@nokia.com,sandy.fong-hou@nokia.com,jacqueline.linebarger@nokia.com
From: PAIR_eOfficeAction@uspto.gov
Cc: PAIR_eOfficeAction@uspto.gov
Subject: Private PAIR Correspondence Notification for Customer Number 73658

Jul 06, 2015 05:31:56 AM

Dear PAIR Customer:

Nokia USA Inc.
Attn: Intellectual Property Rights Docketing
200 South Mathilda Ave
Sunnyvale, CA 94086
UNITED STATES

The following USPTO patent application(s) associated with your Customer Number, 73658 , have new outgoing correspondence. This correspondence is now available for viewing in Private PAIR.

The official date of notification of the outgoing correspondence will be indicated on the form PTOL-90 accompanying the correspondence.

Disclaimer:

The list of documents shown below is provided as a courtesy and is not part of the official file wrapper. The content of the images shown in PAIR is the official record.

Application	Document	Mailroom Date	Attorney Docket No.
13666680	CTNF	07/06/2015	NC77198US-NP
	892	07/06/2015	NC77198US-NP
	1449	07/06/2015	NC77198US-NP

To view your correspondence online or update your email addresses, please visit us anytime at <https://sportal.uspto.gov/secure/myportal/privatepair>.

If you have any questions, please email the Electronic Business Center (EBC) at EBC@uspto.gov with 'e-Office Action' on the subject line or call 1-866-217-9197 during the following hours:

Monday - Friday 6:00 a.m. to 12:00 a.m.

Thank you for prompt attention to this notice,

UNITED STATES PATENT AND TRADEMARK OFFICE
PATENT APPLICATION INFORMATION RETRIEVAL SYSTEM

Doc Code: PA..

Document Description: Power of Attorney

PTO/AIA/82A (07-13)

Approved for use through 11/30/2014. OMB 0651-0051

U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

TRANSMITTAL FOR POWER OF ATTORNEY TO ONE OR MORE REGISTERED PRACTITIONERS

NOTE: This form is to be submitted with the Power of Attorney by Applicant form (PTO/AIA/82B) to identify the application to which the Power of Attorney is directed, in accordance with 37 CFR 1.5, unless the application number and filing date are identified in the Power of Attorney by Applicant form. If neither form PTO/AIA/82A nor form PTO/AIA/82B identifies the application to which the Power of Attorney is directed, the Power of Attorney will not be recognized in the application.

Application Number	13/666,680
Filing Date	11-01-2012
First Named Inventor	Mehmet Oguz BICI
Title	METHOD FOR CODING AND AN APPARATUS
Art Unit	2488
Examiner Name	HILAIRE, CLIFFORD
Attorney Docket Number	042933/467264

SIGNATURE of Applicant or Patent Practitioner

Signature	/Guy R. Gosnell/	Date (Optional)	9/9/2015
Name	Guy R. Gosnell	Registration Number	34,610
Title (if Applicant is a juristic entity)	Patent Practitioner		
Applicant Name (if Applicant is a juristic entity)	Nokia Technologies Oy		

NOTE: This form must be signed in accordance with 37 CFR 1.33. See 37 CFR 1.4(d) for signature requirements and certifications. If more than one applicant, use multiple forms.

☐

*Total of _____ forms are submitted.

This collection of information is required by 37 CFR 1.131, 1.32, and 1.33. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 3 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

POWER OF ATTORNEY BY APPLICANT

I hereby revoke all previous powers of attorney given in the application identified in either the attached transmittal letter or the boxes below.

Application Number	Filing Date

(Note: The boxes above may be left blank if information is provided on form PTO/AIA/82A.)

☒ I hereby appoint the Patent Practitioner(s) associated with the following Customer Number as my/our attorney(s) or agent(s), and to transact all business in the United States Patent and Trademark Office connected therewith for the application referenced in the attached transmittal letter (form PTO/AIA/82A) or identified above:

10949

OR

☐ I hereby appoint Practitioner(s) named in the attached list (form PTO/AIA/82C) as my/our attorney(s) or agent(s), and to transact all business in the United States Patent and Trademark Office connected therewith for the patent application referenced in the attached transmittal letter (form PTO/AIA/82A) or identified above. (Note: Complete form PTO/AIA/82C.)

Please recognize or change the correspondence address for the application identified in the attached transmittal letter or the boxes above to:

☒ The address associated with the above-mentioned Customer Number

OR

☐ The address associated with Customer Number:

OR

Firm or Individual Name				
Address				
City	State	Zip		
Country				
Telephone	Email			

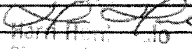
I am the Applicant (if the Applicant is a juristic entity, list the Applicant name in the box):

NOKIA TECHNOLOGIES OY

- ☐ Inventor or Joint Inventor (title not required below)
- ☐ Legal Representative of a Deceased or Legally Incapacitated Inventor (title not required below)
- ☒ Assignee or Person to Whom the Inventor is Under an Obligation to Assign (provide signer's title if applicant is a juristic entity)
- ☐ Person Who Otherwise Shows Sufficient Proprietary Interest (e.g., a petition under 37 CFR 1.46(b)(2) was granted in the application or is concurrently being filed with this document) (provide signer's title if applicant is a juristic entity)

SIGNATURE of Applicant for Patent

The undersigned (whose title is supplied below) is authorized to act on behalf of the applicant (e.g., where the applicant is a juristic entity).

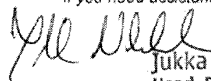
Signature		Date (Optional)	23 January 2015
Name	Jukka Nihtila		
Title	Legal and Intellectual Property		

NOTE: Signature - This form must be signed by the applicant in accordance with 37 CFR 1.33. See 37 CFR 1.4 for signature requirements and certifications. If more than one applicant, use multiple forms.

☐ Total of forms are submitted.

This collection of information is required by 37 CFR 1.131, 1.32, and 1.33. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 3 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.


Jukka Nihtila
Head, Business Development
Legal & IP

Electronic Acknowledgement Receipt	
EFS ID:	23441894
Application Number:	13666680
International Application Number:	
Confirmation Number:	4782
Title of Invention:	METHOD FOR CODING AND AN APPARATUS
First Named Inventor/Applicant Name:	Mehmet Oguz BICI
Customer Number:	73658
Filer:	Guy Randall Gosnell/Kim Shaul
Filer Authorized By:	Guy Randall Gosnell
Attorney Docket Number:	NC77198-US-NP
Receipt Date:	09-SEP-2015
Filing Date:	01-NOV-2012
Time Stamp:	15:16:39
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no
------------------------	----

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Miscellaneous Incoming Letter	467264Request.pdf	20500 2dccd2109905ebdab4e4597df533541a0f7088a2	no	1

Warnings:

Information:

2	Application Data Sheet	467264ADS.pdf	25395	no	3
			84cb6ff32afa006fa7005a82d07f093bc8d39819		
Warnings:					
Information:					
This is not an USPTO supplied ADS fillable form					
3	Assignee showing of ownership per 37 CFR 3.73	467264statement.pdf	86424	no	3
			4e4d98087753bd84e5ac8d19b81401ce0db40934		
Warnings:					
Information:					
4	Power of Attorney	467264POA.pdf	508180	no	2
			7e639ecf8f0ac7d5dd08a8d7867bd94d76c92e9a		
Warnings:					
Information:					
Total Files Size (in bytes):			640499		
This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503. <u>New Applications Under 35 U.S.C. 111</u> If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application. <u>National Stage of an International Application under 35 U.S.C. 371</u> If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course. <u>New International Application Filed with the USPTO as a Receiving Office</u> If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.					

Attorney's Docket No. 042933/467264

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re:	Mehmet Oguz Bici	Confirmation No.:	4782
Appl. No.:	13/666,680	Group Art No.:	2488
Filing Date:	November 1, 2012	Examiner:	Clifford Hilaire
Title:	METHOD FOR CODING AND AN APPARATUS		

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Request to Correct or Update Applicant Name Under 37 CFR § 1.46(c)

This is a request to correct or update the original naming of Applicant(s) for the above-identified application.

☒ Please correct or update the name of the Applicant(s) to:

Applicant No. 1: Nokia Technologies Oy

☒ A Power of Attorney and 3.73(c) statement from each corrected or updated Applicant accompanies this Request.

☒ An Application Data Sheet with an updated Applicant Information Section accompanies this Request.

Applicant(s) respectfully requests that an updated Filing Receipt confirming these changes be issued.

Respectfully submitted,

/Guy R. Gosnell/

Guy R. Gosnell
Registration No. 34,610

CUSTOMER No. 10949
ALSTON & BIRD LLP
Bank of America Plaza
101 South Tryon Street, Suite 4000
Charlotte, NC 28280-4000
Tel Charlotte Office (704) 444-1000
Fax Charlotte Office (704) 444-1111

Application Data Sheet

Application Information

Application number 13/666,680
Filing Date: November 1, 2012
Title: Method for Coding and an Apparatus
Attorney Docket Number: 042933/467264

Applicant Information

Applicant 1

Assignee: **Nokia Technologies Oy**

Legal Representative under 35 USC 117:

Joint Inventor: Mehmet Oguz Bici

Person to whom the inventor is obligated to assign:

Person who shows sufficient proprietary interest:

If applicant is the legal representative, indicate the
Authority to file the patent application, the inventor is:

Name of the Deceased or Legally Incapacitated Inventor:

If the Applicant is an Organization check here: ☒

Mailing Address

Address 1: ~~Tammelan puistokatu 1-3 D-46~~ **Karaportti 3**

Address 2:

City: Tampere **Espoo**

State/Province:

Country: Finland **Finland**

Postal Code: 33500 **02610**

Phone Number:

Fax Number:

Email Address:

Applicant 2

Assignee:

Legal Representative under 35 USC 117:

Joint Inventor:

Jani Lainema

Person to whom the inventor is obligated to assign:

Person who shows sufficient proprietary interest:

If applicant is the legal representative, indicate the
Authority to file the patent application, the inventor is:

Name of the Deceased or Legally Incapacitated Inventor:

If the Applicant is an Organization check here:

☐

Mailing Address

Address 1:

Kisakentankatu 12 B-6

Address 2:

City:

Tampere

State/Province:

Country:

Finland

Postal Code:

33230

Phone Number:

Fax Number:

Email Address:

Applicant 3

Assignee:

Legal Representative under 35 USC 117:

Joint Inventor:

~~Kemal Ugur~~

Person to whom the inventor is obligated to assign:

Person who shows sufficient proprietary interest:

If applicant is the legal representative, indicate the
Authority to file the patent application, the inventor is:

Name of the Deceased or Legally Incapacitated Inventor:

If the Applicant is an Organization check here:

☐

Mailing Address

Address 1:

~~Lapintie 6D-25~~

Address 2:

City:

~~Tampere~~

State/Province:

Country:

~~Finland~~

Postal Code:

~~33100~~

Phone Number:

Fax Number:

Email Address:

Submitted by:

Signature /Guy R. Gosnell/

Date 2015-09-09

Printed Name Guy R. Gosnell

Registration Number 34,610

STATEMENT UNDER 37 CFR 3.73(c)

Applicant/Patent Owner: Nokia Technologies Oy
Application No./Patent No.: 13/666,680 Filed/Issue Date: 11-01-2012
Titled: METHOD FOR CODING AND AN APPARATUS
Nokia Technologies Oy, a corporation
(Name of Assignee) (Type of Assignee, e.g., corporation, partnership, university, government agency, etc.)

states that, for the patent application/patent identified above, it is (choose **one** of options 1, 2, 3 or 4 below):

1. ☒ The assignee of the entire right, title, and interest.
2. ☐ An assignee of less than the entire right, title, and interest (check applicable box):
- ☐ The extent (by percentage) of its ownership interest is ____%. Additional Statement(s) by the owners holding the balance of the interest must be submitted to account for 100% of the ownership interest.
- ☐ There are unspecified percentages of ownership. The other parties, including inventors, who together own the entire right, title and interest are:

Additional Statement(s) by the owner(s) holding the balance of the interest must be submitted to account for the entire right, title, and interest.

3. ☐ The assignee of an undivided interest in the entirety (a complete assignment from one of the joint inventors was made). The other parties, including inventors, who together own the entire right, title, and interest are:

Additional Statement(s) by the owner(s) holding the balance of the interest must be submitted to account for the entire right, title, and interest.

4. ☐ The recipient, via a court proceeding or the like (e.g., bankruptcy, probate), of an undivided interest in the entirety (a complete transfer of ownership interest was made). The certified document(s) showing the transfer is attached.

The interest identified in option 1, 2 or 3 above (not option 4) is evidenced by either (choose **one** of options A or B below):

- A. ☐ An assignment from the inventor(s) of the patent application/patent identified above. The assignment was recorded in the United States Patent and Trademark Office at Reel _____, Frame _____, or for which a copy thereof is attached.

- B. ☒ A chain of title from the inventor(s), of the patent application/patent identified above, to the current assignee as follows:

1. From: Inventors To: Nokia Corporation

The document was recorded in the United States Patent and Trademark Office at
Reel 029555, Frame 0001, or for which a copy thereof is attached.

2. From: Nokia Corporation To: Nokia Technologies Oy

The document was recorded in the United States Patent and Trademark Office at
Reel 035258, Frame 0075, or for which a copy thereof is attached.

[Page 1 of 2]

This collection of information is required by 37 CFR 3.73(b). The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

STATEMENT UNDER 37 CFR 3.73(c)

3. From: _____ To: _____

The document was recorded in the United States Patent and Trademark Office at
Reel _____, Frame _____, or for which a copy thereof is attached.

4. From: _____ To: _____

The document was recorded in the United States Patent and Trademark Office at
Reel _____, Frame _____, or for which a copy thereof is attached.

5. From: _____ To: _____

The document was recorded in the United States Patent and Trademark Office at
Reel _____, Frame _____, or for which a copy thereof is attached.

6. From: _____ To: _____

The document was recorded in the United States Patent and Trademark Office at
Reel _____, Frame _____, or for which a copy thereof is attached.

☐ Additional documents in the chain of title are listed on a supplemental sheet(s).

☒ As required by 37 CFR 3.73(c)(1)(i), the documentary evidence of the chain of title from the original owner to the assignee was, or concurrently is being, submitted for recordation pursuant to 37 CFR 3.11.

[NOTE: A separate copy (i.e., a true copy of the original assignment document(s)) must be submitted to Assignment Division in accordance with 37 CFR Part 3, to record the assignment in the records of the USPTO. See MPEP 302.08]

The undersigned (whose title is supplied below) is authorized to act on behalf of the assignee.

/Guy R. Gosnell/

9/9/2015

Signature

Date

Guy R. Gosnell

34,610

Printed or Typed Name

Title or Registration Number

Privacy Act Statement

The **Privacy Act of 1974 (P.L. 93-579)** requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether disclosure of these records is required by the Freedom of Information Act.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (*i.e.*, GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspection or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING or 371(c) DATE	GRP ART UNIT	FIL FEE REC'D	ATTY. DOCKET NO.	TOT CLAIMS	IND CLAIMS
13/666,680	11/01/2012	2488	2640	NC77198-US-NP	20	8

CONFIRMATION NO. 4782
REPLACEMENT FILING RECEIPT



0000000077463223

10949
Nokia Corporation and Alston & Bird LLP
c/o Alston & Bird LLP
Bank of America Plaza, 101 South Tryon Street
Suite 4000
Charlotte, NC 28280-4000

Date Mailed: 09/17/2015

Receipt is acknowledged of this non-provisional patent application. The application will be taken up for examination in due course. Applicant will be notified as to the results of the examination. Any correspondence concerning the application must include the following identification information: the U.S. APPLICATION NUMBER, FILING DATE, NAME OF APPLICANT, and TITLE OF INVENTION. Fees transmitted by check or draft are subject to collection. Please verify the accuracy of the data presented on this receipt. **If an error is noted on this Filing Receipt, please submit a written request for a Filing Receipt Correction. Please provide a copy of this Filing Receipt with the changes noted thereon. If you received a "Notice to File Missing Parts" for this application, please submit any corrections to this Filing Receipt with your reply to the Notice. When the USPTO processes the reply to the Notice, the USPTO will generate another Filing Receipt incorporating the requested corrections**

Inventor(s)

Mehmet Oguz BICI, Tampere, TURKEY;
Jani LAINEMA, Tampere, FINLAND;
Kemal UGUR, Tampere, FINLAND;

Applicant(s)

Nokia Technologies Oy, Espoo, FINLAND;

Assignment For Published Patent Application

NOKIA CORPORATION, Espoo, FINLAND

Power of Attorney: The patent practitioners associated with Customer Number 10949

Domestic Priority data as claimed by applicant

This appln claims benefit of 61/555,703 11/04/2011

Foreign Applications for which priority is claimed (You may be eligible to benefit from the **Patent Prosecution Highway** program at the USPTO. Please see <http://www.uspto.gov> for more information.) - None.

Foreign application information must be provided in an Application Data Sheet in order to constitute a claim to foreign priority. See 37 CFR 1.55 and 1.76.

If Required, Foreign Filing License Granted: 11/20/2012

The country code and number of your priority application, to be used for filing abroad under the Paris Convention, is **US 13/666,680**

Projected Publication Date: Not Applicable

Non-Publication Request: No

Early Publication Request: No

Title

METHOD FOR CODING AND AN APPARATUS

Preliminary Class

375

Statement under 37 CFR 1.55 or 1.78 for AIA (First Inventor to File) Transition Applications:**PROTECTING YOUR INVENTION OUTSIDE THE UNITED STATES**

Since the rights granted by a U.S. patent extend only throughout the territory of the United States and have no effect in a foreign country, an inventor who wishes patent protection in another country must apply for a patent in a specific country or in regional patent offices. Applicants may wish to consider the filing of an international application under the Patent Cooperation Treaty (PCT). An international (PCT) application generally has the same effect as a regular national patent application in each PCT-member country. The PCT process **simplifies** the filing of patent applications on the same invention in member countries, but **does not result** in a grant of "an international patent" and does not eliminate the need of applicants to file additional documents and fees in countries where patent protection is desired.

Almost every country has its own patent law, and a person desiring a patent in a particular country must make an application for patent in that country in accordance with its particular laws. Since the laws of many countries differ in various respects from the patent law of the United States, applicants are advised to seek guidance from specific foreign countries to ensure that patent rights are not lost prematurely.

Applicants also are advised that in the case of inventions made in the United States, the Director of the USPTO must issue a license before applicants can apply for a patent in a foreign country. The filing of a U.S. patent application serves as a request for a foreign filing license. The application's filing receipt contains further information and guidance as to the status of applicant's license for foreign filing.

Applicants may wish to consult the USPTO booklet, "General Information Concerning Patents" (specifically, the section entitled "Treaties and Foreign Patents") for more information on timeframes and deadlines for filing foreign patent applications. The guide is available either by contacting the USPTO Contact Center at 800-786-9199, or it can be viewed on the USPTO website at <http://www.uspto.gov/web/offices/pac/doc/general/index.html>.

For information on preventing theft of your intellectual property (patents, trademarks and copyrights), you may wish to consult the U.S. Government website, <http://www.stopfakes.gov>. Part of a Department of Commerce initiative, this website includes self-help "toolkits" giving innovators guidance on how to protect intellectual property in specific countries such as China, Korea and Mexico. For questions regarding patent enforcement issues, applicants may call the U.S. Government hotline at 1-866-999-HALT (1-866-999-4258).

LICENSE FOR FOREIGN FILING UNDER
Title 35, United States Code, Section 184
Title 37, Code of Federal Regulations, 5.11 & 5.15

GRANTED

The applicant has been granted a license under 35 U.S.C. 184, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" followed by a date appears on this form. Such licenses are issued in all applications where the conditions for issuance of a license have been met, regardless of whether or not a license may be required as set forth in 37 CFR 5.15. The scope and limitations of this license are set forth in 37 CFR 5.15(a) unless an earlier license has been issued under 37 CFR 5.15(b). The license is subject to revocation upon written notification. The date indicated is the effective date of the license, unless an earlier license of similar scope has been granted under 37 CFR 5.13 or 5.14.

This license is to be retained by the licensee and may be used at any time on or after the effective date thereof unless it is revoked. This license is automatically transferred to any related applications(s) filed under 37 CFR 1.53(d). This license is not retroactive.

The grant of a license does not in any way lessen the responsibility of a licensee for the security of the subject matter as imposed by any Government contract or the provisions of existing laws relating to espionage and the national security or the export of technical data. Licensees should apprise themselves of current regulations especially with respect to certain countries, of other agencies, particularly the Office of Defense Trade Controls, Department of State (with respect to Arms, Munitions and Implements of War (22 CFR 121-128)); the Bureau of Industry and Security, Department of Commerce (15 CFR parts 730-774); the Office of Foreign Assets Control, Department of Treasury (31 CFR Parts 500+) and the Department of Energy.

NOT GRANTED

No license under 35 U.S.C. 184 has been granted at this time, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" DOES NOT appear on this form. Applicant may still petition for a license under 37 CFR 5.12, if a license is desired before the expiration of 6 months from the filing date of the application. If 6 months has lapsed from the filing date of this application and the licensee has not received any indication of a secrecy order under 35 U.S.C. 181, the licensee may foreign file the application pursuant to 37 CFR 5.15(b).

SelectUSA

The United States represents the largest, most dynamic marketplace in the world and is an unparalleled location for business investment, innovation, and commercialization of new technologies. The U.S. offers tremendous resources and advantages for those who invest and manufacture goods here. Through SelectUSA, our nation works to promote and facilitate business investment. SelectUSA provides information assistance to the international investor community; serves as an ombudsman for existing and potential investors; advocates on behalf of U.S. cities, states, and regions competing for global investment; and counsels U.S. economic development organizations on investment attraction best practices. To learn more about why the United States is the best country in the world to develop technology, manufacture products, deliver services, and grow your business, visit <http://www.SelectUSA.gov> or call +1-202-482-6800.



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING OR 371(C) DATE	FIRST NAMED APPLICANT	ATTY. DOCKET NO./TITLE
13/666,680	11/01/2012	Mehmet Oguz BICI	NC77198-US-NP

CONFIRMATION NO. 4782

POA ACCEPTANCE LETTER

10949

Nokia Corporation and Alston & Bird LLP
c/o Alston & Bird LLP
Bank of America Plaza, 101 South Tryon Street
Suite 4000
Charlotte, NC 28280-4000



OC000000077463222

Date Mailed: 09/17/2015

NOTICE OF ACCEPTANCE OF POWER OF ATTORNEY

This is in response to the Power of Attorney filed 09/09/2015.

The Power of Attorney in this application is accepted. Correspondence in this application will be mailed to the above address as provided by 37 CFR 1.33.

Questions about the contents of this notice and the requirements it sets forth should be directed to the Office of Data Management, Application Assistance Unit, at (571) 272-4000 or (571) 272-4200 or 1-888-786-0101.

/hsarwari/

To: usptomail@alston.com,,
From: PAIR_eOfficeAction@uspto.gov
Cc: PAIR_eOfficeAction@uspto.gov
Subject: Private PAIR Correspondence Notification for Customer Number 10949

Sep 17, 2015 05:30:03 AM

Dear PAIR Customer:

Nokia Corporation and Alston & Bird LLP
c/o Alston & Bird LLP
Bank of America Plaza, 101 South Tryon Street
Suite 4000
Charlotte, NC 28280-4000
UNITED STATES

The following USPTO patent application(s) associated with your Customer Number, 10949 , have new outgoing correspondence. This correspondence is now available for viewing in Private PAIR.

The official date of notification of the outgoing correspondence will be indicated on the form PTOL-90 accompanying the correspondence.

Disclaimer:

The list of documents shown below is provided as a courtesy and is not part of the official file wrapper. The content of the images shown in PAIR is the official record.

Application	Document	Mailroom Date	Attorney Docket No.
13666680	APP.FILE.REC	09/17/2015	NC77198-US-NP
	N570	09/17/2015	NC77198-US-NP

To view your correspondence online or update your email addresses, please visit us anytime at <https://sportal.uspto.gov/secure/myportal/privatepair>.

If you have any questions, please email the Electronic Business Center (EBC) at EBC@uspto.gov with 'e-Office Action' on the subject line or call 1-866-217-9197 during the following hours:

Monday - Friday 6:00 a.m. to 12:00 a.m.

Thank you for prompt attention to this notice,

UNITED STATES PATENT AND TRADEMARK OFFICE
PATENT APPLICATION INFORMATION RETRIEVAL SYSTEM

Attorney's Docket No. 042933/467264

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re:	Bici et al.	Confirmation No.: 4782
Appl. No.:	13/666,680	Group Art Unit: 2488
Filed:	November 1, 2012	
For:	METHOD FOR CODING AND AN APPARATUS	

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

REQUEST FOR CORRECTED FILING RECEIPT

The Application Data Sheet filed on November 1, 2012 contained errors in the Inventor Information section. Specifically, the residence provided for Inventor 1 is incorrect. A Subsequent Application Data Sheet is being filed concurrently herewith. Applicant requests that a corrected Filing Receipt be issued.

Respectfully submitted,

/Douglas E. McKay/

Douglas E. McKay
Registration No. 70,175

Customer No. 10949
ALSTON & BIRD LLP
Bank of America Plaza
101 South Tryon Street, Suite 4000
Charlotte, NC 28280-4000
Tel Charlotte Office (704) 444-1000
Fax Charlotte Office (704) 444-1111

Application Data Sheet

Application Information

Application number: 13/666,680
Filing Date: November 1, 2012
Application Type: Nonprovisional
Subject Matter: Utility
Title: METHOD FOR CODING AND AN APPARATUS
Attorney Docket Number: 042933/467264
Total Drawing Sheets: 13

Inventor Information

Inventor 1

Status: Full Capacity
Given Name: Mehmet
Middle Name: Oguz
Family Name: BICI
Non US Residency: Yes
City of Residence: Tampere
Country of Residence: ~~TR~~ **FI**
Street of Mailing Address: Tammelan puistokatu 1-3 D 46
City of Mailing Address: Tampere
Country of mailing address: FI
Postal or Zip Code of mailing address: 33500

Inventor 2

Status: Full Capacity
Given Name: Jani
Middle Name:
Family Name: LAINEMA
Non US Residency: Yes
City of Residence: Tampere

Country of Residence: FI
Street of Mailing Address: Kisakentankatu 12 B 6
City of Mailing Address: Tampere
Country of mailing address: FI
Postal or Zip Code of mailing address: 33230

Inventor 3

Status: Full Capacity
Given Name: Kemal
Middle Name:
Family Name: UGUR
Non US Residency: Yes
City of Residence: Tampere
Country of Residence: FI
Street of Mailing Address: Lapintie 6D 25
City of Mailing Address: Tampere
Country of mailing address: FI
Postal or Zip Code of mailing address: 33100

Submitted by:

Signature /Douglas E. McKay/ Date 09/17/2015
Printed Name Douglas E. McKay Registration Number 70,175

Electronic Acknowledgement Receipt	
EFS ID:	23533004
Application Number:	13666680
International Application Number:	
Confirmation Number:	4782
Title of Invention:	METHOD FOR CODING AND AN APPARATUS
First Named Inventor/Applicant Name:	Mehmet Oguz BICI
Customer Number:	10949
Filer:	Jason P. Cooper/Dennis Avery
Filer Authorized By:	Jason P. Cooper
Attorney Docket Number:	042933/467264
Receipt Date:	18-SEP-2015
Filing Date:	01-NOV-2012
Time Stamp:	11:14:08
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no
------------------------	----

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Request for Corrected Filing Receipt	467264_Request_CFR_09-17-2015.pdf	90639 a561778649bc24d9c55fc4db0e007e76494434c2	no	1

Warnings:

Information:

2	Application Data Sheet	467264_Subsequent_ADS_09-17-2015.pdf	117748 42d38e4061da0c3c6ad9bfafee3861c3800c2a3	no	2
Warnings:					
Information:					
This is not an USPTO supplied ADS fillable form					
Total Files Size (in bytes):			208387		
This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503. <u>New Applications Under 35 U.S.C. 111</u> If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application. <u>National Stage of an International Application under 35 U.S.C. 371</u> If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course. <u>New International Application Filed with the USPTO as a Receiving Office</u> If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.					



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING or 371(c) DATE	GRP ART UNIT	FIL FEE REC'D	ATTY. DOCKET NO	TOT CLAIMS	IND CLAIMS
13/666,680	11/01/2012	2488	2640	042933/467264	20	8

CONFIRMATION NO. 4782

CORRECTED FILING RECEIPT



0000000077598891

10949

Nokia Corporation and Alston & Bird LLP
c/o Alston & Bird LLP
Bank of America Plaza, 101 South Tryon Street
Suite 4000
Charlotte, NC 28280-4000

Date Mailed: 09/23/2015

Receipt is acknowledged of this non-provisional patent application. The application will be taken up for examination in due course. Applicant will be notified as to the results of the examination. Any correspondence concerning the application must include the following identification information: the U.S. APPLICATION NUMBER, FILING DATE, NAME OF APPLICANT, and TITLE OF INVENTION. Fees transmitted by check or draft are subject to collection. Please verify the accuracy of the data presented on this receipt. **If an error is noted on this Filing Receipt, please submit a written request for a Filing Receipt Correction. Please provide a copy of this Filing Receipt with the changes noted thereon. If you received a "Notice to File Missing Parts" for this application, please submit any corrections to this Filing Receipt with your reply to the Notice. When the USPTO processes the reply to the Notice, the USPTO will generate another Filing Receipt incorporating the requested corrections**

Inventor(s)

Mehmet Oguz BICI, Tampere, FINLAND;
Jani LAINEMA, Tampere, FINLAND;
Kemal UGUR, Tampere, FINLAND;

Applicant(s)

Nokia Technologies Oy, Espoo, FINLAND;

Assignment For Published Patent Application

NOKIA CORPORATION, Espoo, FINLAND

Power of Attorney: The patent practitioners associated with Customer Number 10949

Domestic Priority data as claimed by applicant

This appln claims benefit of 61/555,703 11/04/2011

Foreign Applications for which priority is claimed (You may be eligible to benefit from the **Patent Prosecution Highway** program at the USPTO. Please see <http://www.uspto.gov> for more information.) - None.

Foreign application information must be provided in an Application Data Sheet in order to constitute a claim to foreign priority. See 37 CFR 1.55 and 1.76.

If Required, Foreign Filing License Granted: 11/20/2012

The country code and number of your priority application, to be used for filing abroad under the Paris Convention, is **US 13/666,680**

Projected Publication Date: Not Applicable

Non-Publication Request: No

Early Publication Request: No

Title

METHOD FOR CODING AND AN APPARATUS

Preliminary Class

375

Statement under 37 CFR 1.55 or 1.78 for AIA (First Inventor to File) Transition Applications:**PROTECTING YOUR INVENTION OUTSIDE THE UNITED STATES**

Since the rights granted by a U.S. patent extend only throughout the territory of the United States and have no effect in a foreign country, an inventor who wishes patent protection in another country must apply for a patent in a specific country or in regional patent offices. Applicants may wish to consider the filing of an international application under the Patent Cooperation Treaty (PCT). An international (PCT) application generally has the same effect as a regular national patent application in each PCT-member country. The PCT process **simplifies** the filing of patent applications on the same invention in member countries, but **does not result** in a grant of "an international patent" and does not eliminate the need of applicants to file additional documents and fees in countries where patent protection is desired.

Almost every country has its own patent law, and a person desiring a patent in a particular country must make an application for patent in that country in accordance with its particular laws. Since the laws of many countries differ in various respects from the patent law of the United States, applicants are advised to seek guidance from specific foreign countries to ensure that patent rights are not lost prematurely.

Applicants also are advised that in the case of inventions made in the United States, the Director of the USPTO must issue a license before applicants can apply for a patent in a foreign country. The filing of a U.S. patent application serves as a request for a foreign filing license. The application's filing receipt contains further information and guidance as to the status of applicant's license for foreign filing.

Applicants may wish to consult the USPTO booklet, "General Information Concerning Patents" (specifically, the section entitled "Treaties and Foreign Patents") for more information on timeframes and deadlines for filing foreign patent applications. The guide is available either by contacting the USPTO Contact Center at 800-786-9199, or it can be viewed on the USPTO website at <http://www.uspto.gov/web/offices/pac/doc/general/index.html>.

For information on preventing theft of your intellectual property (patents, trademarks and copyrights), you may wish to consult the U.S. Government website, <http://www.stopfakes.gov>. Part of a Department of Commerce initiative, this website includes self-help "toolkits" giving innovators guidance on how to protect intellectual property in specific countries such as China, Korea and Mexico. For questions regarding patent enforcement issues, applicants may call the U.S. Government hotline at 1-866-999-HALT (1-866-999-4258).

LICENSE FOR FOREIGN FILING UNDER
Title 35, United States Code, Section 184
Title 37, Code of Federal Regulations, 5.11 & 5.15

GRANTED

The applicant has been granted a license under 35 U.S.C. 184, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" followed by a date appears on this form. Such licenses are issued in all applications where the conditions for issuance of a license have been met, regardless of whether or not a license may be required as set forth in 37 CFR 5.15. The scope and limitations of this license are set forth in 37 CFR 5.15(a) unless an earlier license has been issued under 37 CFR 5.15(b). The license is subject to revocation upon written notification. The date indicated is the effective date of the license, unless an earlier license of similar scope has been granted under 37 CFR 5.13 or 5.14.

This license is to be retained by the licensee and may be used at any time on or after the effective date thereof unless it is revoked. This license is automatically transferred to any related applications(s) filed under 37 CFR 1.53(d). This license is not retroactive.

The grant of a license does not in any way lessen the responsibility of a licensee for the security of the subject matter as imposed by any Government contract or the provisions of existing laws relating to espionage and the national security or the export of technical data. Licensees should apprise themselves of current regulations especially with respect to certain countries, of other agencies, particularly the Office of Defense Trade Controls, Department of State (with respect to Arms, Munitions and Implements of War (22 CFR 121-128)); the Bureau of Industry and Security, Department of Commerce (15 CFR parts 730-774); the Office of Foreign Assets Control, Department of Treasury (31 CFR Parts 500+) and the Department of Energy.

NOT GRANTED

No license under 35 U.S.C. 184 has been granted at this time, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" DOES NOT appear on this form. Applicant may still petition for a license under 37 CFR 5.12, if a license is desired before the expiration of 6 months from the filing date of the application. If 6 months has lapsed from the filing date of this application and the licensee has not received any indication of a secrecy order under 35 U.S.C. 181, the licensee may foreign file the application pursuant to 37 CFR 5.15(b).

SelectUSA

The United States represents the largest, most dynamic marketplace in the world and is an unparalleled location for business investment, innovation, and commercialization of new technologies. The U.S. offers tremendous resources and advantages for those who invest and manufacture goods here. Through SelectUSA, our nation works to promote and facilitate business investment. SelectUSA provides information assistance to the international investor community; serves as an ombudsman for existing and potential investors; advocates on behalf of U.S. cities, states, and regions competing for global investment; and counsels U.S. economic development organizations on investment attraction best practices. To learn more about why the United States is the best country in the world to develop technology, manufacture products, deliver services, and grow your business, visit <http://www.SelectUSA.gov> or call +1-202-482-6800.

To: usptomail@alston.com,,
From: PAIR_eOfficeAction@uspto.gov
Cc: PAIR_eOfficeAction@uspto.gov
Subject: Private PAIR Correspondence Notification for Customer Number 10949

Sep 23, 2015 05:30:43 AM

Dear PAIR Customer:

Nokia Corporation and Alston & Bird LLP
c/o Alston & Bird LLP
Bank of America Plaza, 101 South Tryon Street
Suite 4000
Charlotte, NC 28280-4000
UNITED STATES

The following USPTO patent application(s) associated with your Customer Number, 10949 , have new outgoing correspondence. This correspondence is now available for viewing in Private PAIR.

The official date of notification of the outgoing correspondence will be indicated on the form PTOL-90 accompanying the correspondence.

Disclaimer:

The list of documents shown below is provided as a courtesy and is not part of the official file wrapper. The content of the images shown in PAIR is the official record.

Application	Document	Mailroom Date	Attorney Docket No.
13666680	APP.FILE.REC	09/23/2015	042933/467264

To view your correspondence online or update your email addresses, please visit us anytime at <https://sportal.uspto.gov/secure/myportal/privatepair>.

If you have any questions, please email the Electronic Business Center (EBC) at EBC@uspto.gov with 'e-Office Action' on the subject line or call 1-866-217-9197 during the following hours:

Monday - Friday 6:00 a.m. to 12:00 a.m.

Thank you for prompt attention to this notice,

UNITED STATES PATENT AND TRADEMARK OFFICE
PATENT APPLICATION INFORMATION RETRIEVAL SYSTEM

Substitute for form SB08 (Revised 07/09) INFORMATION DISCLOSURE STATEMENT BY APPLICANT <i>(Use as many sheets as necessary)</i>				Complete if Known	
				Application Number	13/666,680
				Filing Date	11/01/2012
				First Named Inventor	Mehmet Oguz Bici
				Art Unit	2488
				Examiner Name	Clifford Hilaire
Sheet	1	of	1	Attorney Docket Number	042933/467264
OTHER DOCUMENTS					
Examiner Initials*	Cite No.	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.			English Language Translation Attached
	1	Office Action from corresponding Korean Patent Application No. 2014-7015093, dated August 21, 2015			YES
	2	Wiegand, Tomas, et al.; "WD3: Working Draft 3 of High-Efficiency Video Coding"; Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11; 5 th Meeting; Geneva, CH, 16-23 March, 2011; Document JCTVC-E603; 239 pages			
	3	Bross, Benjamin, et al.; "WD4: Working Draft 4 of High-Efficiency Video Coding"; Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11; 6 th Meeting; Torino, IT; 14-22 July 2011; Document JCTVC-F803_d; 232 pages			
	4	J.-L. Lin, Y.-W. Chen, Y.-W. Huang, S. Lei; "CE9: Results of Experiment ROB04"; JCT-VC Doc. JCTVC-F052, Turin; Jul 2011			
Examiner Signature				Date Considered	

**Examiner: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

CLT#35874174v1

Submitted September 25, 2015