

transform matching the particular chroma partition type is selected from a set of transforms **(see Guo, paragraph 2, 41 and 43, a Hadamard transformation is performed [41] on the macroblocks and the macroblocks are divided into 4x4 blocks. The chroma blocks are selected 4x4 as well. According to the specification of this application, Hadamard transformation is used to determine the transform size).**

12. **Regarding claim 5**, the combination of Guo and Wang discloses: The method of claim 3, wherein from among a particular luma partition type and a particular chroma partition type selected to code the block only the particular luma partition type is signaled **(see Guo, paragraph 41, only luma blocks are used to obtain the best prediction mode)**, the particular luma partition type being selected from the set of luma partition types and the particular chroma partition type being selected from the set of chroma partition types **(see Wang, paragraph 14 and 54, selecting color format. Also see paragraph 11, partitions can have 7 different sizes, and paragraph 77, chroma block sizes up to 16x16).**

13. **Regarding claim 6**, the combination of Guo and Wang discloses: The method of claim 5, wherein the particular luma partition type is absolutely coded or differentially coded from one or more neighboring blocks **(see Gou, paragraph 41).**

14. **Regarding claim 7**, the combination of Guo and Wang discloses: The method of claim 2, wherein a particular chroma partition type used to code the block is determined independently from a particular luma partition type used to code the block (**see Wang, paragraph 77 and claim 6, original block size compensation processing for chroma up to block sizes of 16x16**), the particular chroma partition type being determined from the set of chroma partition types and the particular luma partition type being determined from the set of luma partition types (**see Guo, paragraph 43**).

15. **Regarding claim 8**, the combination of Guo and Wang discloses: The method of claim 7, wherein the particular chroma partition type and the particular luma partition types are both signaled (**see Wang, paragraph 211, bit-stream of intra coded macroblock has the following information: macroblock type, luma prediction mode, chroma prediction mode**).

16. **Regarding claim 9**, the combination of Guo and Wang discloses: The method of claim 8, wherein the particular luma partition type and the particular chroma partition type are absolutely coded or differentially coded from one or more neighboring blocks (**see Gou, paragraph 37**).

17. **Regarding claim 10**, the combination of Guo and Wang discloses: The method of claim 8, wherein the particular chroma partition type is differentially coded from the particular luma partition type (**see Guo, Fig. 1**).

18. **Regarding claim 11**, the combination of Guo and Wang discloses: The method of claim 2, wherein a chroma entropy coding engine used to code the block is different from a luma entropy coding engine used to code the block (**see Guo, paragraph 2, coding system has a luma coding and a chroma coding**).

19. **Regarding claim 12**, claim 12 is drawn to an apparatus having limitations similar to the method claimed in claim 2 treated in the above rejections. Therefore, method claims 12 correspond to apparatus and method claim 2 and are rejected for the same reasons of anticipation as used above. Furthermore, Wang discloses a decoder (**see Wang, paragraph 87**).

20. **Regarding claim 13**, claim 13 is drawn to a method having limitations similar to the method claimed in claim 2 treated in the above rejections. Therefore, method claims 13 correspond to apparatus and method claim 2 and are rejected for the same reasons of anticipation as used above. Furthermore, Wang discloses a decoder (**see Wang, paragraph 87**).

21. **Regarding claims 14 and 15**, claims 14 and 15 are drawn to a method having limitations similar to the method claimed in claims 3 and 4 treated in the above rejections. Therefore, method claims 14 and 15 correspond to apparatus and method claims 3 and 4 and are rejected for the same reasons of anticipation as used above.

22. **Regarding claim 16**, the combination of Guo and Wang discloses: The method of claim 14, wherein from among a particular luma partition type and a particular chroma partition type selected to code the block only information indicating the particular luma partition type is received from a corresponding encoder for use in decoding the block **(see Guo, paragraph 41, only luma blocks are used to obtain the best prediction mode, also see Wang, paragraph 129)**, the particular luma partition type being selected from the set of luma partition types and the particular chroma partition type being selected from the set of chroma partition types **(see Wang, paragraph 14 and 54, selecting color format. Also see paragraph 11, partitions can have 7 different sizes, and paragraph 77, chroma block sizes up to 16x16)**.

23. **Regarding claims 17 and 18**, claims 17 and 18 are drawn to a method having limitations similar to the method claimed in claims 6 and 7 treated in the above rejections. Therefore, method claims 17 and 18 correspond to apparatus and method claims 6 and 7 and are rejected for the same reasons of anticipation as used above.

24. **Regarding claim 19**, the combination of Guo and Wang discloses: The method of claim 18, wherein information indicating both the particular chroma partition type and the particular luma partition types is received from a corresponding encoder for use in decoding the block **(see Wang, paragraph 129)**.

25. **Regarding claims 20-22**, claims 20-22 are drawn to a method having limitations similar to the method claimed in claims 9-11 treated in the above rejections. Therefore, method claims 20-22 correspond to apparatus and method claims 9-11 and are rejected for the same reasons of anticipation as used above.

26. **Regarding claim 23**, claim 23 is drawn to a computer readable storage medium having limitations similar to the apparatus claimed in claim 1 treated in the above rejections. Therefore, computer readable storage medium claim 23 corresponds to apparatus claim 1 and is rejected for the same reasons of anticipation as used above.

Conclusion

THIS ACTION IS MADE FINAL. Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to MARYAM NASRI whose telephone number is (571)270-7158. The examiner can normally be reached on Mon-Thr 9AM-6PM.

Examiner interviews are available via telephone, in-person, and video conferencing using a USPTO supplied web-based collaboration tool. To schedule an interview, applicant is encouraged to use the USPTO Automated Interview Request (AIR) at <http://www.uspto.gov/interviewpractice>.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Joseph Ustaris can be reached on (571) 272-7383. 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.

/MARYAM NASRI/
Examiner, Art Unit 2483

/BEHROOZ SENFI/
Primary Examiner, Art Unit 2482

Application/Control Number: 13/389,872
Art Unit: 2483

Page 15

<i>Index of Claims</i> 	Application/Control No. 13389872	Applicant(s)/Patent Under Reexamination XU ET AL.
	Examiner MARYAM NASRI	Art Unit 2483

✓	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	01/09/2014	05/12/2014	10/28/2014	04/21/2015	08/24/2015	03/06/2016	09/13/2016	03/06/2017	06/22/2017
	1	✓	✓	✓	✓	✓	✓	✓	✓	✓
	2	✓	✓	✓	✓	✓	✓	✓	✓	✓
	3	✓	✓	✓	✓	✓	✓	✓	✓	✓
	4	✓	✓	✓	✓	✓	✓	✓	✓	✓
	5	✓	✓	✓	✓	✓	✓	✓	✓	✓
	6	✓	✓	✓	✓	✓	✓	✓	✓	✓
	7	✓	✓	✓	✓	✓	✓	✓	✓	✓
	8	✓	✓	✓	✓	✓	✓	✓	✓	✓
	9	✓	✓	✓	✓	✓	✓	✓	✓	✓
	10	✓	✓	✓	✓	✓	✓	✓	✓	✓
	11	✓	✓	✓	✓	✓	✓	✓	✓	✓
	12	✓	✓	✓	✓	✓	✓	✓	✓	✓
	13	✓	✓	✓	✓	✓	✓	✓	✓	✓
	14	✓	✓	✓	✓	✓	✓	✓	✓	✓
	15	✓	✓	✓	✓	✓	✓	✓	✓	✓
	16	✓	✓	✓	✓	✓	✓	✓	✓	✓
	17	✓	✓	✓	✓	✓	✓	✓	✓	✓
	18	✓	✓	✓	✓	✓	✓	✓	✓	✓
	19	✓	✓	✓	✓	✓	✓	✓	✓	✓
	20	✓	✓	✓	✓	✓	✓	✓	✓	✓
	21	✓	✓	✓	✓	✓	✓	✓	✓	✓
	22	✓	✓	✓	✓	✓	✓	✓	✓	✓
	23	✓	✓	✓	✓	✓	✓	✓	✓	✓

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	13389872
	Filing Date	2012-02-10
	First Named Inventor	XU et al.
	Art Unit	2483
	Examiner Name	NASRI, MARYAM A.
	Attorney Docket Number	PU090108-US-PCT

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					

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 i	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	JP2009177787	JP		2009-08-06	SAMSUNG ELECTRONICS CO LTD		×
	2	JP2005039743	JP		2005-02-10	SONY CORP		☒
	3	KR20060121651	KR		2006-11-29	SONY CORP		☒

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)

Application Number	13389872
Filing Date	2012-02-10
First Named Inventor	XU et al.
Art Unit	2483
Examiner Name	NASRI, MARYAM A.
Attorney Docket Number	PU090108-US-PCT

4	KR20090081878	KR	2009-07-29	SAMSUNG ELECTRONICS CO LTD	☒
---	---------------	----	------------	-------------------------------	-------------------------------------

If you wish to add additional Foreign Patent Document citation information please click the Add button

NON-PATENT LITERATURE 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, pages(s), volume-issue number(s), publisher, city and/or country where published.	T ⁵
	1		

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

EXAMINER SIGNATURE

Examiner Signature	/MARYAM A NASRI/	Date Considered	06/22/2017
--------------------	------------------	-----------------	------------

*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	13389872
Filing Date	2012-02-10
First Named Inventor	XU et al.
Art Unit	2483
Examiner Name	NASRI, MARYAM A.
Attorney Docket Number	PU090108-US-PCT

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	/Ronald J. Kolczynski/	Date (YYYY-MM-DD)	2017-04-25
Name/Print	Ronald J. Kolczynski	Registration Number	57,177

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.

Search Notes 	Application/Control No. 13389872	Applicant(s)/Patent Under Reexamination XU ET AL.
	Examiner MARYAM NASRI	Art Unit 2483

CPC- SEARCHED		
Symbol	Date	Examiner
H04N 7/12 and H04N 7/26	5/13/2014	MN

CPC COMBINATION SETS - SEARCHED		
Symbol	Date	Examiner

US CLASSIFICATION SEARCHED			
Class	Subclass	Date	Examiner

SEARCH NOTES		
Search Notes	Date	Examiner
Inventor and assignee search	1/9/2014	MN
EAST search, history attached	1/9/2014	MN
Updated EAST and NPL search	5/13/2014	MN
Consulted David Czekaj and Young Lee regarding claim 1	5/8/2014	MN
Consulted David Czekaj and Joe Ustaris regarding first action final	10/28/2014	MN
Updated NPL and EAST search.	10/28/2014	MN
Updated EAST and NPL search. EAST history attached.	4/21/2015	MN
Consulted Joe Ustaris and Andy Rao regarding rejection of claim 1.	4/23/2015	MN
Updated NPL and EAST search.	8/24/2015	MN
Updated NPL and EAST search. History attached.	3/4/2016	MN
Updated EAST and NPL search.	9/13/2016	MN
Updated NPL and EAST search.	3/6/2017	MN
Updated EAST and NPL search.	6/22/2017	MN

INTERFERENCE SEARCH

/MARYAM NASRI/ Examiner.Art Unit 2483	
--	--

US Class/ CPC Symbol	US Subclass / CPC Group	Date	Examiner

/MARYAM NASRI/
Examiner.Art Unit 2483

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)	Application Number	13389872
	Filing Date	2012-02-10
	First Named Inventor	XU et al.
	Art Unit	2483
	Examiner Name	NASRI, MARYAM A.
	Attorney Docket Number	PU090108-US-PCT

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					

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 i	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	CN1487474	CN		2004-04-07	MICROSOFT CORP		×
	2	CN101438591	CN		2009-05-20	MICROSOFT CORP		☒
	3	WO2006124885	WO		2006-11-23	KYLINTV, INC.		☐

**INFORMATION DISCLOSURE
STATEMENT BY APPLICANT**
(Not for submission under 37 CFR 1.99)

Application Number	13389872
Filing Date	2012-02-10
First Named Inventor	XU et al.
Art Unit	2483
Examiner Name	NASRI, MARYAM A.
Attorney Docket Number	PU090108-US-PCT

If you wish to add additional Foreign Patent Document citation information please click the Add button

NON-PATENT LITERATURE 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, pages(s), volume-issue number(s), publisher, city and/or country where published.	T ⁵
	1		

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

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	13389872
Filing Date	2012-02-10
First Named Inventor	XU et al.
Art Unit	2483
Examiner Name	NASRI, MARYAM A.
Attorney Docket Number	PU090108-US-PCT

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	/Ronald J Kolczynski/	Date (YYYY-MM-DD)	2017-09-29
Name/Print	Ronald J. Kolczynski	Registration Number	57,177

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.

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
23 November 2006 (23.11.2006)

PCT

(10) International Publication Number
WO 2006/124885 A2(51) International Patent Classification:
G06F 15/16 (2006.01)(21) International Application Number:
PCT/US2006/018898

(22) International Filing Date: 12 May 2006 (12.05.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/680,331 12 May 2005 (12.05.2005) KR
60/680,332 12 May 2005 (12.05.2005) KR(71) Applicant (for all designated States except US):
KYLINTV, INC. [US/US]; 1600 Old Country Road,
Plainview, NY 11803 (US).

(72) Inventors; and

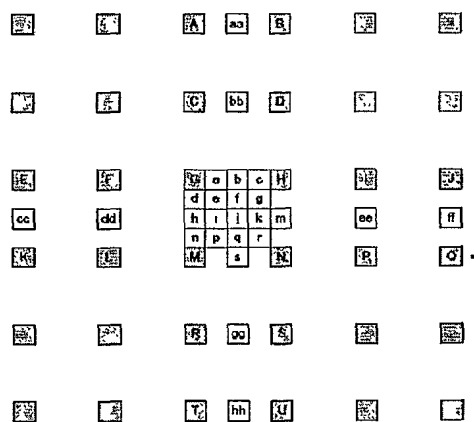
(75) Inventors/Applicants (for US only): **WANG, Xiaohong**
[CN/CN]; 11F Tower A, Guoji Plaza No. 3 Danling
Street, Haidian District, Beijing, 100080 (CN). **WANG,**
Yunchuan [CN/CN]; 11F Tower A, Guoji Plaza No. 3
Danling Street, Haidian District, Beijing, 100080 (CN).
HER, Michael [—/US]; 9 Von Hagen Ave., St. James,
NY 11780 (US).(74) Agents: **CHAU, Frank** et al.; F. Chau & Associates, LLC,
130 Woodbury Road, Woodbury, NY 11797 (US).(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,
AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN,
CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI,
GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE,
KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV,
LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI,
NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG,
SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US,
UZ, VC, VN, YU, ZA, ZM, ZW.(84) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, GH,
GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM,
ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,
FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT,
RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA,
GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

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

[Continued on next page]

(54) Title: CODEC FOR IPTV



Integer samples (shaded blocks with upper-case letters) and fractional sample positions (un-shaded blocks with lower-case letters) for quarter sample luma interpolations

(57) Abstract: A method of optimizing decoding of MPEG4 compliant coded signals is provided, comprising: disabling processing for non-main profile sections; performing reference frame padding; performing adaptive motion compensation. The method of decoding further including performing fast IDCT wherein an IDCT process is performed on profile signals but no IDCT is performed based on whether a 4x4 block may be all zero, or only DC transform coefficient is non-zero, and including CAVLC encoding of residual data. Reference frame padding comprises compensating for motion vectors extending beyond a reference frame by adding to at least the length and width of the reference frame. Adaptive motion compensation includes original block size compensation processing for chroma up to block sizes of 16 x 16.



For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Codec for IPTV

This application claims priority to provisional applications serial nos. 60/680,331, and
5 60/680,332, both filed on May 12, 2005. The disclosures of the provisional applications are
incorporated by reference in their entirety herein.

BACKGROUND

10 1. Technical Field

The present disclosure relates to a codec for encoding and decoding signals representing
humanly perceptible video and audio; more particularly, a codec with speed optimization for use in
Internet Protocol Television.

15 2. Discussion of Related Art

Video-on-demand or television program on demand have been made available to and utilized
by satellite/cable television subscribers. Typically, subscribers can view at their television the
20 video programs available for selection for a fee, and upon selection made at the subscriber's
set-top-box (STB), the program is sent from the program center to the set-top-box via the cable
or satellite network. The large bandwidth available at a cable or satellite network, typically at a
capacity of 400Mbps to 750Mbps or higher, facilitates download of a large portion or the entire
selected video program with very little delay. Some set-top-boxes are equipped with storage for
25 storing the downloaded video and the subscriber watches the video program from the STB as if
from a video cassette/disk player.

More recently, a selection of television programs are made available for viewing over the
Internet using a browser and a media player at a personal computer. In some cases, the
30 requested programs are streamed instead of downloaded to the personal computer for viewing.
In these systems, the video programs are not viewed at a television through an STB. Nor is the
viewing experience the same as watching from a video disk player because the PC does not
respond to a remote control as does a television or a television STB. Even though media

players on PCs can be controlled by a virtual on-screen controller, the control and viewing experience through a mouse or keyboard is different from a disk player and a remote control. Further, most PC users use their PCs on a desk in an actual or home office arrangement, which is not conducive to watching television programs or movies, e.g., the furniture may not be comfortable and the audiovisual effects cannot be as well appreciated. Moreover, if a PC accesses the Internet via a LAN and the access point is via DSL, the bandwidth capacity may be only 500 Kbps to 2Mbps. This bandwidth limitation may render difficult a real-time, uninterrupted program streamed over the Internet unless the viewing area is made very small or very low resolution, or unless a highly compressed and speed optimized codec is used.

ITU-T H.264 / MPEG-4 (Part 10) Advanced Video Coding (commonly referred as H.264/AVC) is an international video coding standard adopted by ITU-T's Video Coding Experts Group (VCEG) and ISO/IEC's Moving Picture Experts Group (MPEG). As has been the case with past standards, its design provides the most current balance between the coding efficiency, implementation complexity, and cost – based on state of VLSI design technology (CPU's, DSP's, ASIC's, FPGA's, etc.).

H.264/AVC is designed to cover a broad range of applications for video content including but not limited to, for example: Cable TV on optical networks, copper, etc.; Direct broadcast satellite video services; Digital subscriber line video services; Digital terrestrial television broadcasting; Interactive storage media (optical disks, etc.); Multimedia services over packet networks; and Real-time conversational services (videoconferencing, videophone, etc.), etc.

Three basic feature sets called *profiles* were established to address these application domains: the Baseline, Main, and Extended profiles. The Baseline profile was designed to minimize complexity and provide high robustness and flexibility for use over a broad range of network environments and conditions; the Main profile was designed with an emphasis on compression *coding efficiency* capability; and the Extended profile was designed to combine the robustness of the Baseline profile with a higher degree of coding efficiency and greater network robustness and to add enhanced modes useful for special "trick uses" for such applications as flexible video streaming.

While having a broad range of applications, the initial H.264/AVC standard (as it was completed in May of 2003), was primarily focused on "entertainment-quality" video, based on 8-bits/sample, and 4:2:0 chroma sampling. In July, 2004, a new amendment was added to this standard, called the Fidelity Range Extensions (FRExt, Amendment 1). The FRExt project produced a suite of four new profiles collectively called the *High* profiles.

The coding structure of this standard is similar to that of all prior major digital video standards (H.261, MPEG-1, MPEG-2 / H.262, H.263 or MPEG-4 part 2). The architecture and the core building blocks of the encoder are also based on motion-compensated DCT-like transform coding. Each picture is compressed by partitioning it as one or more slices; each slice consists of macroblocks, which are blocks of 16x16 luma samples with corresponding chroma samples. However, each macroblock is also divided into sub-macroblock partitions for motion-compensated prediction. The prediction partitions can have seven different sizes – 16x16, 16x8, 8x16, 8x8, 8x4, 4x8 and 4x4. In past standards, motion compensation used entire macroblocks or, in the case of newer designs, 16x16 or 8x8 partitions, so the larger variety of partition shapes provides enhanced prediction accuracy. The spatial transform for the residual data is then either 8x8 (a size supported only in FRExt) or 4x4. In past major standards, the transform block size has always been 8x8, so the 4x4 block size provides an enhanced specificity in locating residual difference signals. The block size used for the spatial transform is always either the same or smaller than the block size used for prediction.

As the video compression tools primarily work at or below the slice layer, bits associated with the slice layer and below are identified as Video Coding Layer (VCL) and bits associated with higher layers are identified as Network Abstraction Layer (NAL) data. VCL data and the highest levels of NAL data can be sent together as part of one single bitstream or can be sent separately. The NAL is designed to fit a variety of delivery frameworks (e.g., broadcast, wireless, storage media). Herein, we discuss the VCL, which is the heart of the compression capability.

The basic unit of the encoding or decoding process is the macroblock. In 4:2:0 chroma format, each macroblock consists of a 16x16 region of luma samples and two corresponding 8x8 chroma sample arrays. In a macroblock of 4:2:2 chroma format video, the chroma sample arrays are 8x16

in size; and in a macroblock of 4:4:4 chroma format video, they are 16x16 in size.

Slices in a picture are compressed by using the following coding tools:

- "Intra" spatial (block based) prediction
 - Full-macroblock luma or chroma prediction – 4 modes (directions) for prediction
 - 8x8 (FRExt-only) or 4x4 luma prediction – 9 modes (directions) for prediction
- "Inter" temporal prediction – block based motion estimation and compensation
 - Multiple reference pictures
 - Reference B pictures
 - Arbitrary referencing order
 - Variable block sizes for motion compensation
 - ✧ Seven block sizes: 16x16, 16x8, 8x16, 8x8, 8x4, 4x8 and 4x4
 - 1/4-sample luma interpolation (1/4 or 1/8th-sample chroma interpolation)
 - Weighted prediction
 - Frame or Field based motion estimation for interlaced scanned video
- Interlaced coding features
 - Frame-field adaptation
 - ✧ Picture Adaptive Frame Field (PicAFF)
 - ✧ MacroBlock Adaptive Frame Field (MBAFF)
 - Field scan
- Lossless representation capability
 - Intra PCM raw sample-value macroblocks
 - Entropy-coded transform-bypass lossless macroblocks (FRExt-only)
- 8x8 (FRExt-only) or 4x4 integer inverse transform (conceptually similar to the well-known DCT)
- Residual color transform for efficient RGB coding without conversion loss or bit expansion (FRExt-only)
- Scalar quantization
- Encoder-specified perceptually weighted quantization scaling matrices (FRExt-only)
- Logarithmic control of quantization step size as a function of quantization control parameter
- Deblocking filter (within the motion compensation loop)
- Coefficient scanning
 - Zig-Zag (Frame)
 - Field
- Lossless Entropy coding
 - Universal Variable Length Coding (UVLC) using Exp-Golomb codes
 - Context Adaptive VLC (CAVLC)
 - Context-based Adaptive Binary Arithmetic Coding (CABAC)
- Error Resilience Tools
 - Flexible Macroblock Ordering (FMO)
 - Arbitrary Slice Order (ASO)
 - Redundant Slices
- SP and SI synchronization pictures for streaming and other uses
- Various color spaces supported (YCbCr of various types, YCgCo, RGB, etc. – especially in

FRExt)

- 4:2:0, 4:2:2 (FRExt-only), and 4:4:4 (FRExt-only) color formats
- Auxiliary pictures for alpha blending (FRExt-only)

5 Each slice need not use all of the above coding tools. Depending on the subset of coding tools used, a slice can be of I (Intra), P (Predicted), B (Bi-predicted), SP (Switching P) or SI (Switching I) type. A picture may contain different slice types, and pictures come in two basic types – reference and non-reference pictures. Reference pictures can be used as references for interframe prediction during the decoding of later pictures (in bitstream order) and non-reference pictures cannot. (It is noteworthy that, unlike in prior standards, pictures that use bi-prediction can be used as references just like pictures coded using I or P slices.)

15 This standard is designed to perform well for both progressive-scan and interlaced-scan video. In interlaced-scan video, a frame consists of two fields – each captured at $\frac{1}{2}$ the frame duration apart in time. Because the fields are captured with significant time gap, the spatial correlation among adjacent lines of a frame is reduced in the parts of picture containing moving objects. Therefore, from coding efficiency point of view, a decision needs to be made whether to compress video as one single frame or as two separate fields. H.264/AVC allows that decision to be made either independently for each pair of vertically-adjacent macroblocks or independently for each entire frame. When the decisions are made at the macroblock-pair level, this is called MacroBlock Adaptive Frame-Field (MBAFF) coding and when the decisions are made at the frame level then this is called Picture-Adaptive Frame-Field (PicAFF) coding. Notice that in MBAFF, unlike in the MPEG-2 standard, the frame or field decision is made for the vertical macroblock-pair and not for each individual macroblock. This allows retaining a 16x16 size for each macroblock and the same size for all submacroblock partitions – regardless of whether the macroblock is processed in frame or field mode and regardless of whether the mode switching is at the picture level or the macroblock-pair level.

30 In addition to basic coding tools, the H.264/AVC standard enables sending extra supplemental information along with the compressed video data. This often takes a form called "supplemental enhancement information" (SEI) or "video usability information" (VUI) in the

standard. SEI data is specified in a backward-compatible way, so that as new types of supplemental information are specified, they can even be used with profiles of the standard that had been previously specified before that definition.

The Baseline, Main, and Extended Profiles

H.264/AVC contains a rich set of video coding tools. Not all the coding tools are required for all the applications. For example, sophisticated error resilience tools are not important for the networks with very little data corruption or loss. Forcing every decoder to implement all the tools would make a decoder unnecessarily complex for some applications. Therefore, subsets of coding tools are defined; these subsets are called Profiles. A decoder may choose to implement only one subset (Profile) of tools, or choose to implement some or all profiles. The following three profiles were defined in the original standard, and remain unchanged in the latest version:

- Baseline (BP)
- Extended (XP)
- Main (MP)

Table 1 gives a high-level summary of the coding tools included in these profiles. The Baseline profile includes I and P slices, some enhanced error resilience tools (FMO, ASO, and RS), and CAVLC. It does not contain B, SP and SI-slices, interlace coding tools or CABAC entropy coding. The Extended profile is a super-set of Baseline, adding B, SP and SI slices and interlace coding tools to the set of Baseline Profile coding tools and adding further error resilience support in the form of data partitioning (DP). It does not include CABAC. The Main profile includes I, P and B-slices, interlace coding tools, CAVLC and CABAC. It does not include enhanced error resilience tools (FMO, ASO, RS, and DP) or SP and SI-slices.

Table 1: Profiles in Original H.264/AVC Standard

Coding Tools	Baseline	Main	Extended
I and P Slices	X	X	X
CAVLC	X	X	X
CABAC		X	
B Slices		X	X
Interlaced Coding (PicAFF, MBAFF)		X	X
Enh. Error Resil. (FMO, ASO, RS)	X		X

Further Enh. Error Resil. (DP)			X
SP and SI Slices			X

The New High Profiles Defined in the FRExt Amendment

The FRExt amendment defines four new profiles:

- High (HP)
- High 10 (Hi10P)
- High 4:2:2 (Hi422P)
- High 4:4:4 (Hi444P)

All four of these profiles build further upon the design of the prior Main profile, and they all include three enhancements of coding efficiency performance:

- Adaptive macroblock-level switching between 8x8 and 4x4 transform block size
- Encoder-specified perceptual-based quantization scaling matrices
- Encoder-specified separate control of the quantization parameter for each chroma component

All of these profiles also support monochrome coded video sequences, in addition to typical 4:2:0 video. The difference in capability among these profiles is primarily in terms of supported sample bit depths and chroma formats. However, the High 4:4:4 profile additionally supports the residual color transform and predictive lossless coding features not found in any other profiles. The detailed capabilities of these profiles are shown in Table 2.

Table 2: New Profiles in the H.264/AVC FRExt Amendment

Coding Tools	High	High 10	High 4:2:2	High 4:4:4
Main Profile Tools	X	X	X	X
4:2:0 Chroma Format	X	X	X	X
8 Bit Sample Bit Depth	X	X	X	X
8x8 vs. 4x4 Transform Adaptivity	X	X	X	X
Quantization Scaling Matrices	X	X	X	X
Separate Cb and Cr QP control	X	X	X	X
Monochrome video format	X	X	X	X
9 and 10 Bit Sample Bit Depth		X	X	X
4:2:2 Chroma Format			X	X
11 and 12 Bit Sample Bit Depth				X

4:4:4 Chroma Format				X
Residual Color Transform				X
Predictive Lossless Coding				X

As shown in Table 3, H.264/AVC defines 16 different Levels, tied mainly to the picture size and frame rate. Levels also provide constraints on the number of reference pictures and the maximum compressed bit rate that can be used.

In the standard, *Level/s* specify the maximum frame size in terms of only the total number of pixels/frame. Horizontal and Vertical maximum sizes are not specified except for constraints that horizontal and vertical sizes can not be more than $\text{Sqrt}(\text{maximum frame size} * 8)$. If, at a particular level, the picture size is less than the one in the table, then a correspondingly larger number of reference pictures (up to 16 frames) can be used for motion estimation and compensation. Similarly, instead of specifying a maximum frame rate at each level, a maximum sample (pixel) rate, in terms of macroblocks per second, is specified. Thus if the picture size is smaller than the typical pictures size in Table 3, then the frame rate can be higher than that in Table 3, all the way up to a maximum of 172 frames/sec.

Table 3: Levels in H.264/AVC

Level Number	Typical Picture Size	Typical frame rate	Maximum compressed bit rate (for VCL) in non-FRExt profiles	Maximum number of reference frames for typical picture size
1	QCIF	15	64kbps	4
1b	QCIF	15	128kbps	4
1.1	CIF or QCIF	7.5(CIF)/30(QCIF)	192kbps	2(CIF)/9(QCIF)
1.2	CIF	15	384kbps	6
1.3	CIF	30	768kbps	6
2	CIF	30	2Mbps	6
2.1	HHR(480i or 576i)	30/25	4Mbps	6
2.2	SD	15	4Mbps	5
3	SD	30/25	10Mbps	5
3.1	1280x720p	30	14Mbps	5
3.2	1280x720p	60	20Mbps	4
4	HD Formats	60p/30i	20Mbps	4

	(720p or 1080i)			
4.1	HD Formats (720p or 1080i)	60p/30i	50Mbps	4
4.2	1920x1080p	60p	50Mbps	4
5	2kx1k	72	135Mbps	5
5.1	2kx1k or 4kx2k	120/30	240Mbps	5

A need therefore exists for a robust codec which is speed optimized to facilitate the coding and decoding of humanly perceptible video to provide real-time play of humanly perceptible video sent from a remote program center using Internet Protocol. There is also a need to optimize the codec processing so that the real-time play can be facilitated using a DSL connection at as low a bandwidth capacity of 500 Kbps.

Summary of the Invention

A method of optimizing decoding of MPEG4 compliant coded signals is provided, comprising: disabling processing for non-main profile sections; performing reference frame padding; performing adaptive motion compensation. The method of decoding further including performing fast IDCT wherein an IDCT process is performed on profile signals but no IDCT is performed based on whether a 4x4 block may be all zero, or only DC transform coefficient is non-zero, and including CAVLC encoding of residual data. Reference frame padding comprises compensating for motion vectors extending beyond a reference frame by adding to at least the length and width of the reference frame. Adaptive motion compensation includes original block size compensation processing for chroma up to block sizes of 16 x 16.

A method of encoding MPEG4 compliant data is also provided, comprising: performing Rate Distortion Optimization (RDO) algorithm, fast motion search algorithm, and bitrate control algorithm.

Brief Description of the Drawings

Figure 1 shows motion segments used for reference.

Figure 2 shows reference frame padding process according to an embodiment of the invention.

Figure 3 shows Integer samples (shaded blocks with upper-case letters) and fractional sample positions (un-shaded blocks with lower-case letters) for quarter sample luma interpolations.

Figure 4 shows Interpolation of chroma eighth-pel positions.

Figure 5 shows current and neighbouring partitions in a motion compensation search process.

Figure 6 shows a hexagon search.

Figure 7 shows a full search for fractional pel search.

Detailed Description of Preferred Embodiments

According to a preferred embodiment of the invention, platform-independent optimization of H.264 main profile decode is implemented. Decode process time can be shortened by optimization process including shutting down non-main profile sections, reference frame padding, adaptive block motion compensation, and Fast inverse DCT. Our study shows that a number of components, e.g., luma and chroma motion compensation and inverse integer transform, are the most time-consuming modules in the H.264 decoder. The speed of the optimized H.264 main profile decoder is about 2.0~3.3 times faster compared with a reference decoder.

An MPEG4 compliant encoder and decoder, such as JM61e, is used as a reference codec. The reference decoder should be able to decode all syntactic elements and specified in the main profile. The PSNR (peak signal-to-noise ratio) value should also be maintained despite elimination of profiles.

Six standard CIF video sequences are shown in Figure 1 as the testing series. Within them flowergarden (a), tempete (c) and mobile (d) involve more movement. Foreman (b), highway (e) and paris (f) are more static..

Process elimination includes the following:

- ☐ Since main profile decoder operates on only I, P and B slice types, all processes corresponding to SP and SI slices can be eliminated.

- ☐ Preserve only one partition to place the input bitstream since the content of NAL unit in main profile does not contain coded slice partitions.

- ☐ Delete error/loss robustness features, including FMO (Flexible Macroblock Ordering), ASO (Arbitrary Slice Ordering), and RS (Redundant Slices).

- ☐ Delete unused variables and functions.

1.

Two encoded bitstreams are tested, having specifications of CIF size, one reference frame, one slice per frame, only first is I frame, quantization parameter is 28, max search range is 16, RD-Optimization and Hadamard transform, and block types from 16x16 to 4x4 are used. The Foreman video bitstream has maximum compression ratio (300 frames) for the above configuration. The garden bitstream has minimum compression ratio (250 frames). The encoding performance of these two sequences is listed in the first four columns of Table 1.

To further optimize the main profile decoder, areas of decoding bottleneck are identified, and the most time consuming modules are simplified, using the following parameters:

- ☐ single frame reference (NumberReferenceFrames = 1);
- ☐ one slice in each frame (SliceMode = 0);
- ☐ I frame interval is 100 (IntraPeriod = 100);
- ☐ quantization parameter is 25 (QPFirstFrame = QPRemainingFrame = 25);
- ☐ not using Hadamard (UseHadamard = 0);
- ☐ max search range is 16 (SearchRange = 16);
- ☐ not using rate distortion optimization (RDOptimization = 0);
- ☐ using all blocks types from 16x16 to 4x4 (InterSearch16x16 = 1, InterSearch16x8 = 1, InterSearch8x16 = 1, InterSearch8x8 = 1, InterSearch8x4 = 1, InterSearch4x8 = 1, InterSearch4x4 = 1);
- ☐ Inter pixels are used for Intra macroblock prediction (UseConstrainedIntraPred = 0);
- ☐ POC mode is 0 (PicOrderCntType = 0).
- ☐ Entropy coding method is 0, i.e. CAVLC (SymbolMode = 0)

Table 3 Encoding results of six standard series with same config file

sequence	Encoded frames	SNR Y	SNR U	SNR V	bitrate(Kbits/s) (framerate is 30)	Encoding time (s)
flowergarden	250	37.23	39.06	39.45	2618.73	200.982
foreman	300	38.11	41.80	43.70	846.02	242.078
tempete	260	37.02	38.79	40.12	2219.72	208.350
mobile	300	36.29	37.80	37.72	2905.49	238.251
highway	300	39.60	39.23	39.92	621.51	245.699
paris	300	37.85	40.82	40.97	785.39	228.793

Using the above configuration, the encoding result of the six series is displayed in Table 3. It can be seen from table 3 that the quality (SNR) and bitrate of reconstructed series varies according to the content of each series. In general, objective quality of reconstructed slow moving series is much better than fast moving series, and bitrate of slow moving is much slower.

Table 4 Time profile for non-optimized main profile decoder

sequence	Motion Compensation	IDCT	Deblock	Entropy decoding	Buffer management and others	Decoding time(s)	decoding framerate
flowergarden	24.8%	5.0%	4.4%	58.0%	7.8%	11.375	21.978
foreman	42.4%	7.0%	6.2%	33.8%	10.6%	9.687	30.969
tempeste	30.1%	5.0%	5.1%	53.0%	6.8%	12.156	21.389
mobile	25.3%	4.2%	4.5%	59.9%	6.1%	16.125	18.605
highway	38.6%	9.0%	6.0%	35.1%	11.3%	7.640	39.267
paris	34.6%	9.3%	5.8%	35.8%	14.5%	7.062	42.481

Major function modules of main profile H.264 decoder are time-profiled in table 4. Result of these tests have been achieved on a PC Pentium $\square$ processor, running at 2.4 GHz, equipped with 512 Mbytes of memory, with Windows 2000 professional. For each sequence tests are done 10 times and averaged to minimize the non-deterministic effects of processor cache, process scheduler, and operative system management operations.

From table 4, it can be seen that the following modules are the key kernels in the decoder: luminance and chrominance motion compensation, entropy decoding, inverse integer transform, and deblocking filtering, among which motion compensation and entropy decoding are the most time-consuming modules. Unlike many previous video coding standards, chrominance motion compensation occupies a significant percentage in the total motion compensation process of the H.264 standard. Entropy decoding time is proportional to the bitrate of bitstream. By improving the processing speed of these most time-consuming modules, the efficiency of the decoder is improved.

According to a preferred embodiment of the present invention, further processes are implemented to improve on decoding processing speed. The processes include reference frame

padding and adaptive block motion compensation MC. Inverse DCT transform is optimized by judging its dimension. Residual data CAVLC decoding tables are reformed for faster speed. As will be further detailed below, these processes improved the speed of the optimized H.264 main profile decoder about seven times faster compared with the reference decoder.

H.264 standard allows motion vectors to point over picture boundaries. When computing inter prediction for a P-macroblock, the pixel position used in reference frame may exceed frame height and width. In this case, the nearest pixel value in reference frame is used for computation. In reference code, whenever a pixel (x_pos, y_pos) in reference frame is used for motion compensation, really used position (x_real, y_real) is obtained by the following equation:

$$y\_real = \max(0, \min(\text{img\_height}-1, y\_pos))$$

$$x\_real = \max(0, \min(\text{img\_width}-1, x\_pos))$$

To avoid the above computation, alternative frame padding is employed:

The padding is added on top, bottom, left and right of the reference frame as shown in Figure 2a to 2d. Suppose frame height is img_height, and frame width is img_width, then padded reference frame height is img_height + 2*Y_PADSTRIDE, and padded reference frame width is img_width + 2* Y_PADSTRIDE. Y_PADSTRIDE can be determined by (max (mv_x, mv_y) +8). If search range is 16 and motion is slow, Y_PADSTRIDE can be equal to 24. For heavier motion sequence, larger value for Y_PADSTRIDE should be set.

The padding process is as follows (see figure 2(d)):

Region0: corresponding values in original frame

Region1: all samples equal to up-left corner pixel in original frame:

Region2: all samples equal to up-right corner pixel in original frame

Region3: all samples equal to down-left corner pixel in original frame

Region4: all samples equal to down-right corner pixel in original frame

Region5: extrapolated horizontally by first column pixel in original frame

Region6: extrapolated vertically by first row pixel in original frame

Region7: extrapolated horizontally by last column pixel in original frame

Region8: extrapolated vertically by last row pixel in original frame

Y, U and V components are padded using the same technique. The only difference is that the padding stride for U and V component is half as that of Y component. It should be noted that the side effect of the above reference frame padding process can be increased memory requirements and extra time for padding.

Adaptive Block MC

In the reference codec, luma compensation is computed by 4x4 block. But real motion
5 compensation blocks are 16x16, 8x16, 16x8, 8x8, 4x8, 8x4, and 4x4. Suppose a macroblock is
predicted by 16x16 mode, reference software will call 4x4 motion compensation function 16 times.
If the predicted position is half-pixel or quarter-pixel position, the computation time used by calling
4x4 block MC 16 times is definitely more than direct 16x16 block MC because of functions
10 invocation cost and data cache. Especially for positions j, i, k, f, and q depicted in Figure 3,
computation is lessened by using larger block MC because of the characteristic of 6-tap filter. On the
other hand, large block types (16x16, 8x16, 16x8) are used more frequently compared with
smaller block types. It has been reported that the total number of 4x4, 4x8 and 8x4 blocks used in
motion estimation takes up about 5% in P-frames among different block sizes. Therefore,
15 adaptive block MC can be employed, e.g., using MxN size interpolation directly for MxN block type
(M and N may be 16, 8 or 4) instead of using only 4x4 size interpolation.

In the reference codec, chroma motion compensation is computed point by point. In this way,
dx, dy, 8-dx, 8-dy, and the positions of A, B, C, D showed in figure 4 are calculated for each
20 chroma point. Real chroma motion compensation blocks should be half of luma

prediction block (8x8, 4x8, 8x4, 4x4, 2x4, 4x2, and 2x2). Hence computing chroma MC on its original block base is more computational efficient.

Fast IDCT

Instead of using conventional DCT/IDCT, H.264 standard uses a 4x4 integer transform to convert spatial-domain signals into frequency-domain and vice versa. Two-dimensional IDCT is implemented in the reference code, and each 4x4 block of transformed coefficients is inverse transformed by calling this 2D IDCT transform. In fact, transform coefficients in a 4x4 block may be all zero, or only DC transform coefficient is non-zero. Thus, a decision can be made before IDCT is performed. The decision is based on the syntax_element coded_block_pattern (cbp). Coded_block_pattern specifies which of the six 8x8 blocks - luma and chroma - contain non-zero transform coefficient levels. For macroblocks with prediction mode not equal to Intra_16x16, coded_block_pattern is present in the bitstream and the variables CodedBlockPatternLuma and CodedBlockPatternChroma are derived as follows.

```
CodedBlockPatternLuma = coded_block_pattern % 16
CodedBlockPatternChroma = coded_block_pattern / 16
```

The judgment for luma 4x4 block is as follows:

```
if (currMB->cbp & (1<<block8x8))    Dimension = 2;
else if (img->cof[i][j][0][0] != 0)  Dimension = 1;
else                                  Dimension = 0;
```

The judgment for chroma 4x4 block is as follows:

```
if (currMB->cbp>31)                  Dimension = 2;
else if (img->cof[2*uv+i][j][0][0] != 0)  Dimension = 1;
else                                      Dimension = 0;
```

2.4 Fast CAVLC decoding algorithm

From table 4 and table 5 one can see that entropy decoding module is very time consuming (up to 60%) for high bitrate sequence. Parameters that require to be transmitted and decoded are displayed in Table 6. Among them residual data is most time consuming and processing speed improvement can be made by optimizing this data processing.

Table 6 Parameters to be encoded

Parameters	Description
Sequence-, picture- and slice-layer syntax elements	

Macroblock type mb_type	Prediction method for each coded macroblock
Coded block pattern	Indicates which blocks within a macroblock contain coded coefficients
Quantizer parameter	Transmitted as a delta value from the previous value of QP
Reference frame index	Identify reference frame(s) for inter prediction
Motion vector	Transmitted as a difference (mvd) from predicted motion vector
Residual data	Coefficient data for each 4x4 or 2x2 block

Above the slice layer, syntax elements are encoded as fixed- or variable-length binary codes. At the slice layer and below, elements are coded using either Context-based adaptive variable length coding (CAVLC) or context adaptive arithmetic coding (CABAC) depending on the entropy encoding mode.

CAVLC is the method used to encode residual, zig-zag ordered 4x4 (and 2x2) blocks of transform coefficients. CAVLC is designed to take advantage of several characteristics of quantized 4x4 blocks:

- After prediction, transformation and quantization, blocks are typically sparse (containing mostly zeros). CAVLC uses run-level coding to compactly represent strings of zeros.
- The highest non-zero coefficients after the zig-zag scan are often sequences of +/-1. CAVLC signals the number of high-frequency +/-1 coefficients ("Trailing 1s" or "T1s") in a compact way.
- The number of non-zero coefficients in neighbouring blocks is correlated. The number of coefficients is encoded using a look-up table; the choice of look-up table depends on the number of non-zero coefficients in neighbouring blocks.
- The level (magnitude) of non-zero coefficients tends to be higher at the start of the reordered array (near the DC coefficient) and lower towards the higher frequencies. CAVLC takes advantage of this by adapting the choice of VLC look-up table for the "level" parameter depending on recently-coded level magnitudes.

By CAVLC encoding of a block of transform coefficients, the bottleneck of entropy decoding can be simplified. CAVLC encoding of residual data proceeds as follows.

Step 1: Encode the number of coefficients and trailing ones (coeff_token).

The first VLC, coeff_token, encodes both the total number of non-zero coefficients (TotalCoeffs) and the number of trailing +/-1 values (T1). TotalCoeffs can be anything from 0

(no coefficients in the 4x4 block) to 16 (16 non-zero coefficients). T1 can be anything from 0 to 3; if there are more than 3 trailing +/-1s, only the last 3 are treated as "special cases" and any others are coded as normal coefficients. There are 4 choices of look-up table to use for encoding coeff_token, described as Num-VLC0, Num-VLC1, Num-VLC2 and Num-FLC (3 variable-length code tables and a fixed-length code). The choice of table depends on the number of non-zero coefficients in upper and left-hand previously coded blocks N_u and N_L.

Step 2: Encode the sign of each T1.

For each T1 (trailing +/-1) signalled by coeff_token, a single bit encodes the sign (0=+, 1=-). These are encoded **in reverse order**, starting with the highest-frequency T1.

Step 3: Encode the levels of the remaining non-zero coefficients.

The **level** (sign and magnitude) of each remaining non-zero coefficient in the block is encoded **in reverse order**, starting with the highest frequency and working back towards the DC coefficient. The choice of VLC table to encode each level adapts depending on the magnitude of each successive coded level (**context adaptive**). There are 7 VLC tables to choose from, Level_VLC0 to Level_VLC6. Level_VLC0 is biased towards lower magnitudes; Level_VLC1 is biased towards slightly higher magnitudes and so on.

Step 4: Encode the total number of zeros before the last coefficient.

TotalZeros is the sum of all zeros preceding the highest non-zero coefficient in the reordered array. This is coded with a VLC. The reason for sending a separate VLC to indicate TotalZeros is that many blocks contain a number of non-zero coefficients at the start of the array and (as will be seen later) this approach means that zero-runs at the start of the array need not be encoded.

Step 5: Encode each run of zeros.

The number of zeros preceding each non-zero coefficient (run_before) is encoded in reverse order. A run_before parameter is encoded for each non-zero coefficient, starting with the highest frequency.

Table 7 Detail profile for CAVLC decoding

sequence	Step 1	Step 3	Step 4	Step 5	Others	Total of residual data decoding	Other part of entropy decoding
flowergarden	23.4%	1.5%	7.1%	12.9%	6.0%	50.9%	7.1%
foreman	15.5%	0.3%	3.8%	3.2%	3.4%	26.2%	7.6%
tempete	23.9%	1.0%	6.4%	9.0%	5.4%	45.7%	7.3%
mobile	26.6%	1.2%	7.3%	11.5%	6.5%	53.1%	6.8%
highway	16.3%	0.3%	3.7%	2.6%	4.5%	27.4%	7.7%
paris	13.8%	0.8%	3.7%	6.0%	3.6%	27.9%	7.9%

Time profile of CAVLC decoding for the six testing bitstream is displayed in table 7. The percentage in the table is obtained by dividing the corresponding decoding time of each step by the total decoding time. Column 2 is the percentage for step 1, column 3 is the percentage for step 3, column 4 is the percentage for step 4, and column 5 is the percentage for step 5. Column 6 is the percentage for other functions (including step 2 and function calling) of residual data decoding module. Column 7 denotes the percentage of total residual data decoding compared with total decoding time. The last column denote the entropy decoding percentage for bitstream other than residual data.

From table 7 it can be seen that step 1, 4 and 5 are the most time consuming steps. These three steps have the same characteristic of tentatively looking up tables. Reference code use same lentab and codtab tables for both encoding and decoding. Encoder knows each coordinate in advance and takes out a value from the table. For decoder, things are quite different. Decoder must try to find out both x and y coordinates while length is not known. So it is very tentative for decoder to try for each length. Thus a key factor for optimization of decoding residual data is to reform the tables for each step. The target of table changes is to minimize table lookup times and bitstream reading times. Program flow should be changed according to the table. The reformed table for readSyntaxElement_Run (step 5) is shown in table 5. The design principle for the reformed table is as follows: put delta len in lentab_codtab_Run[vlcnum][ij][0], look up table using the value obtained from len, look up operation is finished as long as temp=lentab_codtab_Run[vlcnum][ij][value+1] is not zero. Run equal to temp-1. Thus we finished the procedure of reading run by reading bitstream only once.

Table 8 original and reformed table for readSyntaxElement_Run

Original table	Reformed table
<pre> int lentab[TOTRUN_NUM][16] = { {1,1}, {1,2,2}, {2,2,2,2}, {2,2,2,3,3}, {2,2,3,3,3,3}, {2,3,3,3,3,3,3}, {3,3,3,3,3,3,3,4,5,6,7,8,9,10,11}, }; int codtab[TOTRUN_NUM][16] = { {1,0}, {1,1,0}, {3,2,1,0}, {3,2,1,1,0}, {3,2,3,2,1,0}, {3,0,1,3,2,5,4}, {7,6,5,4,3,2,1,1,1,1,1,1,1,1,1}, }; </pre>	<pre> static char lentab_codtab_Run[7][9][9] = { { { 1, 2, 1, 0, 0, 0, 0, 0, 0}, }, { { 1,21, 1, 0, 0, 0, 0, 0, 0}, { 1, 3, 2, 0, 0, 0, 0, 0, 0}, }, { { 2, 4, 3, 2, 1, 0, 0, 0, 0}, }, { { 2,21, 3, 2, 1, 0, 0, 0, 0}, { 1, 5, 4, 0, 0, 0, 0, 0, 0}, }, { { 2,21, 0, 2, 1, 0, 0, 0, 0}, { 1, 6, 5, 4, 3, 0, 0, 0, 0}, }, { { 2,21, 0, 0, 1, 0, 0, 0, 0}, { 1, 2, 3, 5, 4, 7, 6, 0, 0}, }, { { 3,21, 7, 6, 5, 4, 3, 2, 1}, { 1,21, 8, 0, 0, 0, 0, 0, 0}, { 1,21, 9, 0, 0, 0, 0, 0, 0}, { 1,21,10, 0, 0, 0, 0, 0, 0}, { 1,21,11, 0, 0, 0, 0, 0, 0}, { 1,21,12, 0, 0, 0, 0, 0, 0}, { 1,21,13, 0, 0, 0, 0, 0, 0}, { 1,21,14, 0, 0, 0, 0, 0, 0}, { 1,21,15, 0, 0, 0, 0, 0, 0}, } }; </pre>

Note: Each table may have special cases for processing, such as the value 21 in above table.

In addition to the above described optimization processes, loop unrolling, loop distribution

and loop interchange and cache optimization can be used.

Table 9 shows time profile for the kernel modules of this optimized decoder. Results were averaged on 10 times for every sequence.

Table 9 shows the speed-up for the kernel modules in the optimized main profile decoder compared with non-optimized main profile decoder. It is clear from the table that the time used for motion compensation and inverse integer transform and residual data reading modules are dramatically minimized as a result of optimization implementation. Deblock module has minor improvement of about two times faster. Implementation of the above described optimization processes result in seven times improvement in the optimized H.264 main profile decoder compared with the reference decoder.

Table 9 Kernel module speed-up

Sequence	Motion Compensation	IDCT	deblock	Reading residual data	Decoder fps
flowergarden	8.97X	4.30X	1.90X	10.17X	6.87X
foreman	8.53X	8.22X	1.80X	9.83X	6.34X
tempete	8.33X	4.40X	1.68X	10.27X	6.34X
mobile	7.73X	3.36X	1.70X	10.63X	6.57X
highway	9.31X	9.96X	2.14X	11.24X	6.97X
paris	10.78X	12.46X	2.32X	9.10X	7.63X

Table 10 shows the time distribution of Optimized Main profile H264 decoder in percentage. The percentage of the motion compensation and inverse integer transform and entropy decoding modules are minimized. At the same time, deblocking filter now has a much larger impact in the H.264 decoder.

Table 10 Time Distribution of optimized MP decoder

Sequence	Motion Compensation	IDCT	deblock	Reading residual data	Decoding time (s)	Decoder fps
flowergarden	19.0%	8.0%	15.9%	34.4%	1.655	151.102
foreman	31.5%	5.4%	21.8%	16.9%	1.528	196.289
tempete	22.9%	7.2%	19.2%	28.2%	1.918	135.553
mobile	21.5%	8.2%	17.4%	32.8%	2.455	122.210
highway	28.9%	6.3%	19.5%	17.0%	1.096	273.751
paris	24.5%	5.7%	19.1%	23.4%	0.925	324.451

Decoder implemented on DSP

According to an illustrative embodiment of the invention, decode modules are implemented on a DSP, such as a Blackfin BF533 or BF561. Exemplary decode modules on BF533 include:

- (1) H264 bitstream decoding
- (2) synchronization of audio and video
- (3) PPI interrupt function for displaying moving video or interface
- (4) SPI interrupt function for receiving video bitstream and interface commands from CF5249
- (5) module executing interface command, such as display text information, xor rectangular and

display icon etc.

- (6) OSD module for pasting fast forward, fast backward, pause and mute icon.
- (7) auto update DSP program
- (8) display rolling info
- (9) display subtitle
- (10) multi decoder control

H264 bitstream decoding

Utilizing CACHE

To better utilize the DSP memory, decoded frames are stored in SDRAM. That is to say, reference frames are in SDRAM. But access speed of SDRAM is much slower compared with L1. We use 16KB (0xFF804000 - 0xFF807FFF) L1_DATA_A space as CACHE, L1_SCRATCH as stack. Thus 48 KB L1 space is left for storage of all critical decoding variables.

Utilizing MDMA

Although CACHE is used, we should also try to avoid reading and writing SDRAM directly. It is very efficient to decode the current macroblock slice in L1 and then MDMA this macroblock slice to SDRAM after it is totally decoded. By macroblock slice I mean one (16*720 + 8*360*2) image bar. BF533 has only two MDMAs, namely MDMA0 and MDMA1. We use MDMA1 for this macroblock slice transfer. MDMA0 is reserved for PPI display.

Assembly Optimizing

We can use standard C in Visual DSP++ and the compiler with change our C code to assembly. For the fastest decoding speed we rewrite the C optimized code to assembly code according to the algorithm characteristic of each function. Video Pixel Operations and Parallel Instructions are good tricks.

Audio and video synchronization

According to a preferred embodiment, audio and video are decoded in separate DSPs, so the synchronization is a special case compared with single chip solution. The player run in CPU send both video and audio timestamps to BF533 and BF533 try to match the two timestamps by control the displaying time of each video frames. Thus the player should send a bit more video bitstream to BF533 for BF533 to decode in advance and display the corresponding video frame at the same

time with audio.

PPI interrupt function

5 Video display is realized through PPI, one of BF533 external peripherals. The display frequency demanded by ITU 656 is 27MHZ. External Bus Interface Unit of BF533 is compliant with the PC133 SDRAM standard. That is to say, if display buffer is in SDRAM, the decoding program will often interrupt display that we can never get clear image as long as decoder is working.

10 We use two 1716 bytes line buffer in L1 space. Thus the display strategy is as follows: PPI always read data from these two line buffer using DMA0. After reading over each line buffer, one interrupt is generated. In the PPI interrupt function, we prepare the next ITU 656 line using MDMA0.

15 The advantage of this line buffer display strategy is that the bandwidth of MDMA0 from SDRAM to L1 is much higher than SDRAM to PPI. Thus EBIU bandwidth could be maximally used by H264 decoder.

SPI interrupt function

20 BF533 receive data from CF5249 through SPI controller. Real data is followed by a 24-bytes header. In SPI interrupt function, we check the 24-bytes header to know data type and then set DMA5 to receive next chunk of data by setting its receive address. Since data is received by DMA to SDRAM, so the core may read wrong data from CACHE. So we always spare at least 32 bytes (CACHE line size) to store next chunk of data.

Interface module

30 This module realize the following functions: update rectangular, display large and small images, display English or Chinese text, display input box, xor rectangle, change the color of a rectangle, fill a specified region, change the color of text in a rectangle, display icon, draw straight line, etc.

When realizing these interface command, the current interface image is displaying. Since core

has priority over MDMA0 and MDMA0 has priority over MDMA1, we use MDMA1 to operate the image memory to avoid green or white lines on TV screen. That is to say, we use MDMA1 the corresponding line to be modified to L1 and compute new value in L1 and then MDMA1 the modified line out to its original storage place.

multi decoder control

BF533 has 64KB SRAM and 16KB SRAM/CACHE in L1, which is sufficient for storing instructions for one codec. When there is multiple decoding, instruction CACHE is used. Another choice is to use memory overlay. Overlay manager will DMA the corresponding function into L1 when needed. Memory overlay is mostly used in chips without CACHE. It is not a good choice for BF533 since memory overlay is not as good as CACHE.

We use a pure DMA method to switch the next being used codec into L1. We store instruction code and data blocks for each specific codec in SDRAM in advance. When one codec is needed, the shell program DMA the corresponding instruction and data into L1. All codec should have the same main function calling address and interrupt calling address by using RESOLVE in LDF files. Different decoders should use the same LDF file and shell program. Thus each decoder could be debugged separately.

H.264 Encoder Optimization

For H264 encoder optimization, JM61e is used as the reference coder. Non-main profile sections are eliminated to arrive at a main profile encoder. Further speed improvement can be realized by optimizing processes such as RDO (Rate Distortion Optimization) algorithm, fast motion search algorithm, and bitrate control algorithm. In addition, a fast 'SKIP' macroblock detect is used. Still further, encoder speed is improved by using MMX/SSE/SSE2 instructions. For HTT (Hyper Thread Technology) and multi-CPU computers, 'omp parallel sections' is used for parallel encoding.

Improved RDO algorithm

According to a preferred embodiment of the present invention, a Rate Distortion Optimization

(RDO) process is employed. This minimizes the function $D+LxR$, where D and R denote distortion and bit rate respectively and L is the Lagrange multiplier, to select the best macroblock mode and MV. The procedure to encode one macroblock s in a I-, P- or B-frame in the high-complexity mode is summarized as follows.

a) Given the last decoded frames, Lagrange multipliers λ_{MODE} and λ_{MOTION} , and the macroblock quantisation parameter QP .

$$L_{MODE} = 0.85 \times 2^{QP/3},$$

$$L_{MOTION} = \sqrt{L_{MODE}};$$

b) Choose intra prediction modes for the INTRA 4x4 macroblock mode by minimizing

$$J(s, c, IMODE | QP, \lambda_{MODE}) = SSD(s, c, IMODE | QP) + \lambda_{MODE} \cdot R(s, c, IMODE | QP)$$

with $IMODE \in \{DC, HOR, VERT, DIAG, DIAG\_RL, DIAG\_LR\}$.

c) Determine the best Intra16x16 prediction mode by choosing the mode that results in the minimum SATD.

d) For each 8x8 sub-partition,

Perform motion estimation and reference frame selection by minimizing

$$SSD + L \times Rate(MV, REF)$$

B frames: Choose prediction direction by minimizing

$$SSD + L \times Rate(MV(PDIR), REF(PDIR))$$

Determine the coding mode of the 8x8 sub-partition using the rate-constrained mode decision, i.e. minimize

$$SSD + L \times Rate(MV, REF, Luma-Coeff, block\ 8x8\ mode)$$

Here the SSD calculation is based on the reconstructed signal after DCT, quantization, and IDCT.

$$SSD(s, c, MODE | QP) = \sum_{x=1, y=1}^{16,16} (s_Y[x, y] - c_Y[x, y, MODE | QP])^2 + \sum_{x=1, y=1}^{8,8} (s_U[x, y] - c_U[x, y, MODE | QP])^2 + \sum_{x=1, y=1}^{8,8} (s_V[x, y] - c_V[x, y, MODE | QP])^2 \quad e)$$

Perform motion estimation and reference frame selection for 16x16, 16x8, and 8x16 modes by minimizing

$$J(REF, \mathbf{m}(REF) | L_{MOTION})$$

$$= SATD(s, c(REF, \mathbf{m}(REF))) + L_{MOTION} \cdot (R(\mathbf{m}(REF)) - \mathbf{p}(REF)) + R(REF))$$

for each reference frame and motion vector of a possible macroblock mode.

f) B frames: Determine prediction direction by minimizing

$$J(PDIR | L_{MOTION})$$

$$= SATD(s, c(PDIR, \mathbf{m}(PDIR))) + L_{MOTION} \cdot (R(\mathbf{m}(PDIR)) - \mathbf{p}(PDIR)) + R(REF(PDIR)))$$

g) Choose the macroblock prediction mode by minimizing

$$J(s, c, MODE | QP, L_{MODE}) = SSD(s, c, MODE | QP) + L_{MODE} \cdot R(s, c, MODE | QP)$$

given QP and L_{mode} when varying MODE. MODE indicates a mode out of the set of potential macroblock modes:

5 I frame: $MODE \in \{INTRA4 \times 4, INTRA16 \times 16\},$

P frame: $MODE \in \left\{ INTRA4 \times 4, INTRA16 \times 16, SKIP, \right. \\ \left. 16 \times 16, 16 \times 8, 8 \times 16, 8 \times 8, \right\},$

B frame: $MODE \in \left\{ INTRA4 \times 4, INTRA16 \times 16, DIRECT, \right. \\ \left. 16 \times 16, 16 \times 8, 8 \times 16, 8 \times 8 \right\}.$

The computation of $J(s, c, SKIP | QP, L_{mode})$ and $J(s, c, DIRECT | QP, L_{mode})$ is simple. The costs for the other macroblock modes are computed using the intra prediction modes or motion vectors and reference frames.

10

To further improve RDO algorithm speed, processing continues as follows.

a) Given the last decoded frames, Lagrange multipliers λ_{MODE} and λ_{MOTION} , and the macroblock quantisation parameter QP.

$$L_{MODE} = 0.85 \times 2^{QP/3},$$

$$L_{MOTION} = \sqrt{L_{MODE}},$$

b) If current frame is I frame, jump to step h);

c) Detect whether the current macroblock could be encoded using 'SKIP' mode. If so, the procedure is terminated.

d) Perform motion estimation and reference frame selection for 16x16, 16x8, and 8x16 modes by minimizing

$$J(REF, \mathbf{m}(REF) | L_{MOTION})$$

$$= SA(T)D(s, c(REF, \mathbf{m}(REF))) + L_{MOTION} \cdot (R(\mathbf{m}(REF) - \mathbf{p}(REF)) + R(REF))$$

e) If the cost of 8x8 mode is minimal in step d), then go to step f), else go to step g);

f) For each 8x8 sub-partition,

Perform motion estimation and reference frame selection by minimizing

$$SA(T)D + L \times \text{Rate}(MV, REF)$$

Determine the coding mode of the 8x8 sub-partition using the rate-constrained mode decision, i.e. minimize

$$SSD + L \times \text{Rate}(MV, REF, \text{Luma-Coeff}, \text{block } 8x8 \text{ mode})$$

g) If J is greater than a predefined threshold (such as 128), then jump to step h) for intra detection, else jump to step j);

h) Choose intra prediction modes for the INTRA 4x4 macroblock mode by minimizing

$$J(s, c, IMODE | QP, \lambda_{MODE}) = SSD(s, c, IMODE | QP) + \lambda_{MODE} \cdot R(s, c, IMODE | QP)$$

with $IMODE \in \{DC, HOR, VERT, DIAG, DIAG\_RL, DIAG\_LR\}$.

i) Determine the best Intra16x16 prediction mode by choosing the mode that results in the minimum SATD.

j) The last step: Choose the macroblock prediction mode by minimizing

$$J(s, c, MODE | QP, L_{MODE}) = SSD(s, c, MODE | QP) + L_{MODE} \cdot R(s, c, MODE | QP).$$

The improvement in speed of the above RDO is realized from the following processes:

1) For macroblock in P frame, 'SKIP' mode is first checked, (i.e. if take predicted motion vector as the motion vector of 16x16, check whether CBP is 0. The fast algorithm for 'SKIP' mode detection is depicted in next section.) . If 'SKIP' mode is tested OK, the RDO procedure is terminated.

2) After 'SKIP' mode detection, check large blocks of 16x16, 16x8, 8x16 and 8x8,. Only when 8x8 is the best for the comparison of the four block types, there is need to check blocks sizes below

8x8.

3) If the result of inter detection is smaller than a predefined threshold, there is no need to further detect intra modes.

4) For further improving encoding speed, SAD can be used instead of SATD in the part of SA(T)D.

Fast 'SKIP' Mode Detection Method Algorithm

In general, bitstream of intra coded macroblock has the following information: macroblock type, luma prediction mode, chroma prediction mode, delta QP, CBP, and residual data. Bitstream of inter coded macroblock has the information of macroblock type, reference frame index, delta motion vector compared with predicted motion vector, delta QP, CBP and residual data.

Skipped macroblock is a macroblock for which no data is coded other than an indication that the macroblock is to be decoded as "skipped". Macroblocks in P and B frames are allowed to use skipped mode. The advantage of skipped macroblock is that only macroblock type is transmitted, hence bitstream is sparingly used.

For a macroblock in P frame to be coded as skipped mode, the following five conditions must be satisfied:

- 1) the current frame is P frame;
- 2) the best encoding block type for the current macroblock is 16x16;
- 3) CBP of current macroblock is zero;
- 4) the reference frame referenced by the current frame has index 0 in reference frame list;
- 5) the best motion vector for the current macroblock is predicted motion vector of 16x16 block.

According to a preferred embodiment of the invention, a fast 'SKIP' mode detection method includes:

Step 1: compute the predicted motion vector of 16x16 block

Encoding a motion vector for each partition can take a significant number of bits, especially if small partition sizes are chosen. Motion vectors for neighbouring partitions are often highly correlated and so each motion vector is predicted from vectors of nearby, previously coded partitions. A predicted vector, MVp, is formed based on previously calculated motion vectors. MVD, the difference between the current vector and the predicted vector, is encoded and transmitted. The method of forming the prediction MVp depends on the motion compensation partition size

and on the availability of nearby vectors and can be summarised as follows (for macroblocks in P-slices).

Let E be the current macroblock, macroblock partition or sub-partition; let A be the partition or subpartition immediately to the left of E; let B be the partition or sub-partition immediately above E; let C be the partition or sub-partition above and to the right of E. If there is more than one partition immediately to the left of E, the topmost of these partitions is chosen as A. If there is more than one partition immediately above E, the leftmost of these is chosen as B. Figure 5(a) illustrates the choice of neighbouring partitions when all the partitions have the same size (16x16 in this case); Figure 5(b) shows an example of the choice of prediction partitions when the neighboring partitions have different sizes from the current partition E.

For skipped macroblocks: a 16x16 vector MVp is generated as in case (1) above (i.e. as if the block were encoded in 16x16 Inter mode). If one or more of the previously transmitted blocks shown in the Figure are not available (e.g. if it is outside the current picture or slice), the choice of MVp is modified accordingly.

Step 2: Compute the CBP of Y component, and check whether it is zero

Take the predicted motion vector MVp as current motion vector, and compute the predicted value of Y component, then by subtraction of original pixel and predicted pixel we get residual data. Do DCT transform for the sixteen 4x4 residual blocks and then quantize the residual block. If any 4x4 block has non-zero coefficient, the detection is terminated. Here, an ETAL

(Early Termination Algorithm for Luma) algorithm for detecting whether one 4x4 block has non-zero coefficient can be used.

H.264 use 4x4 integer transform, the transform formula is as follows:

$$Y = C_f X C_f^T \otimes E_f = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} X \begin{bmatrix} 1 & 2 & 1 & 1 \\ 1 & 1 & -1 & -2 \\ 1 & -1 & -1 & 2 \\ 1 & -2 & 1 & -1 \end{bmatrix} \otimes \begin{bmatrix} a^2 & ab/2 & a^2 & ab/2 \\ ab/2 & b^2/4 & ab/2 & b^2/4 \\ a^2 & ab/2 & a^2 & ab/2 \\ ab/2 & b^2/4 & ab/2 & b^2/4 \end{bmatrix} \quad \text{Where } e$$

“core” 2-D transform. E is a matrix of scaling factors and the symbol $\otimes$ indicates that each element of CXC^T is multiplied by the scaling factor in the same position in matrix E (scalar multiplication rather than matrix multiplication).

The basic forward quantizer operation in H264 is as follows:

$$Z_{ij} = \text{round}(Y_{ij} / \text{Qstep})$$

where Y_{ij} is a coefficient of the transform described above, Qstep is a quantizer step size and

Z_{ij} is a quantized coefficient. A total of 52 values of Qstep are supported by the standard and these are indexed by a Quantization Parameter, QP. The wide range of quantizer step sizes makes it possible for an encoder to accurately and flexibly control the trade-off between bit rate and quality.

It can be assumed that if the DC value of transform coefficient is zero, then all other AC coefficients are zero. For a 4x4 block, the quantized DC value is:

$$DC = \sum_{x=0}^3 \sum_{y=0}^3 f(x, y) / ((2^{\text{qbits}} - f) / QE[q_{\text{rem}}][0][0])$$

Therefore if the following formula is satisfied, the quantized DC value would be zero.

$$\left| \sum_{x=0}^3 \sum_{y=0}^3 f(x, y) \right| < ((2^{\text{qbits}} - f) / QE[q_{\text{rem}}][0][0])$$

Where $\text{qbits} = 15 + \text{QP} / 6$, $q_{\text{rem}} = \text{QP} \% 6$, $f = (1 \ll \text{qbits}) / 6$, QE is the defined quantization coefficient table and QP is the input quantization parameter.

Step 3: Compute the CBP of U component, and check whether it is zero

First compute the predicted value of U component, then get the residual data of size 8x8. Do DCT transform for the four 4x4 block DCT, the chroma DC coefficients constitute 2x2 array W_D .

This 2x2 chroma DC coefficients array should have Hadamard transform by the following equation:

$$Y_D = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} W_D \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

The CBP of U component depends on the quantized coefficients of Y_D and the quantized non-DC coefficients of each 4x4 block. As long as one quantized coefficient is not zero, the CBP of U component is not zero and the detection should be terminated. Here ETAC (Early Termination Algorithm for Chroma) algorithm can be used for detecting whether one 8x8 chroma block has non-zero coefficient.

Since array W_D has all the DC components of 8x8 block, the CBP of component U is zero as long as quantized coefficients of Y_D are all zero.

The computation formulas for the four elements of W_D are:

$$W_D(0,0) = \sum_{y=0}^3 \sum_{x=0}^3 f(x,y),$$

$$W_D(0,1) = \sum_{y=0}^3 \sum_{x=4}^7 f(x,y),$$

$$W_D(1,0) = \sum_{y=4}^7 \sum_{x=0}^3 f(x,y),$$

$$W_D(1,1) = \sum_{y=4}^7 \sum_{x=4}^7 f(x,y).$$

Y_D is computed as follows:

$$Y_D(0,0) = W_D(0,0) + W_D(0,1) + W_D(1,0) + W_D(1,1),$$

$$Y_D(1,0) = W_D(0,0) - W_D(0,1) + W_D(1,0) - W_D(1,1),$$

$$Y_D(0,1) = W_D(0,0) + W_D(0,1) - W_D(1,0) - W_D(1,1),$$

$$Y_D(1,1) = W_D(0,0) - W_D(0,1) - W_D(1,0) + W_D(1,1).$$

The quantization formula for each element of Y_D is:

$$(|Y_D(i,j)| \times QE[q_{rem}][0][0] + 2 \times f) >> (qbits + 1)$$

Hence CBP of U component is zero as long as the following function is satisfied.

$$|Y_D(i, j)| < ((2^{(qbits+1)} - 2 \times f) / QE[q_{rem}][0][0])$$

Where $qbits = 15 + QP\_SCALE\_CR[QP]/6$, $q_{rem} = QP\_SCALE\_CR[QP] \% 6$,
 $f = (1 \ll qbits) / 6$, QE is the defined quantization coefficient table, QP is the input
 quantization parameter and QP_SCALE_CR is a constant table.

Step 4: Compute the CBP of V component, and check whether it is zero.

This step is similar to step 3. First compute the predicted value of V component, then get the residual data of size 8x8, finally test whether CBP is zero using ETAC method.

3. Fast Motion Search Algorithm

Similar to former video standards such as H.261, MPEG-1, MPEG-2, H.263, and MPEG-4, H.264 is also based on hybrid coding framework, inside which motion estimation is the most important part in exploiting the high temporal redundancy between successive frames and is also the most time consuming part in the hybrid coding framework. Specifically multi prediction modes, multi reference frames, and higher motion vector resolution are adopted in H.264 to achieve more accurate prediction and higher compression efficiency. As a result, the complexity and computation load of motion estimation increase greatly in H.264. It is seen that motion estimation can consume 60% (1 reference frame) to 80% (5 reference frames) of the total encoding time of the H.264 codec and much higher proportion can be obtained if RD optimization or some other tools is invalid and larger search range (such as 48 or 64) is used.

Generally motion estimation is conducted into two steps: first is integer pel motion estimation, and the second is fractional pel motion estimation around the position obtained by the integer pel motion estimation (we name it the best integer pel position). For fractional pel motion estimation, 1/2-pel accuracy is frequently used (H.263, MPEG-1, MPEG-2, MPEG-4), higher resolution motion vector are adopted recently in MPEG-4 and JVT to achieve more accurate motion description and higher compression efficiency.

Algorithms on fast motion estimation are always hot research spot, especially fast integer pel motion estimation has achieved much more attention because traditional fractional pel motion estimation (such as 1/2-pel) only take a very few proportion in the computation load of whole motion estimation. Many fast integer block-matching algorithms have been focused on the search

strategies with different steps and search patterns in order to reduce the computation complexity and maintain the video quality at the same time. These typical fast block matching algorithms include three step search (TSS) 2-D logarithmic search, Four step search (FSS), HEXBS(Hexagon-Based Search), etc.

Based on the above former algorithms and the specific characteristic of H264 motion search, an improved Hexagon Search process is employed for integer pel motion estimation in H.264. This process decreases integer motion search points, and computation load of fractional pel motion estimation is decreased.

Step 1: The median value of the adjacent blocks on the left, top, and top-right (or top-left) of the current block is used to predict the motion vector of the current block

$$pred\_mv = median(Mv\_A, Mv\_B, Mv\_C)$$

Take this predicted motion vector (Pred_x, Pred_y) as the best motion vector and compute rate-distortion cost using this motion vector.

Step 2: Compute rate-distortion cost for (0, 0) vector. The prediction with the minimum cost is taken as best start searching position.

Step 3: If current block is 16x16, compute rate-distortion cost for Mv_A, Mv_B, and Mv_C and then compare with best start searching position. The prediction with the minimum cost is taken as best start searching position.

Step 4: If current block is not 16x16, compute rate-distortion cost for motion vector of the up layer block (for example, mode 5 or 6 is the up layer of mode 7, and mode 4 is the up layer of mode 5 or 6, etc.). The prediction with the minimum cost is taken as best start searching position.

Step 5: Take the best start searching position as center, search the six position around it according to Large Hexagon in Figure 6. If the central point has minimum cost, then terminate this step; or else take the minimum cost point as the next search center and execute large hexagon search again, until the minimum cost point is central point. The maximum large hexagon search times could be limited, such as 16.

Step 6: Take the best position of Step 5 as center, search the four position around it according to Small Hexagon in Figure 3. If the central point has minimum cost, then terminate this step; or else take the minimum cost point as the next search center and execute small hexagon search

again, until the minimum cost point is central point.

Step 7: Integer pel search terminated, the current best motion vector is the final choice.

For fractional pel motion estimation, a so called fractional pel search window which is an area
5 bounded by eight neighbor integer pels positions around the best integer pel position is examined.
In generation of these fractional pel positions, a 6 tap filter is used to produce the 1/2-pel positions,
1/4-pel positions is produced by linear interpolation.

Figure 7 shows the typical Hierarchical Fractional Pel Search algorithm provided in JM test
10 model. The HFPS is described by the following 3 steps:

Step 1. Check the eight 1/2-pel positions (1-8 points) around the best integer pel position to
find the best 1/2-pel motion vector;

Step 2. Check the eight 1/4-pel positions (a-h points) around the best 1/2-pel position to find
the best 1/4-pel motion vector;

15 Step 3. Select the motion vector and block-size pattern, which produces the lowest
rate-distortion cost.

The diamond search pattern is employed in fast fractional pel search.

Step 1. Take the best position of interger pel as center point, search the four diamond position around it.

Step 2. If the MBD (Minimum Block Distortion) point is located at the center, go to step 3; otherwise choose the MBD point in this step as the center of next search, then iterate this step;

Step 3. Choose the MBD point as the motion vector.

Bitrate Control

An encoder employs rate control as a way to regulate varying bit rate characteristics of the coded bitstream to produce high quality decoded frame at a given target bit rate. The rate control for JVT is more difficult than those for other standards. This is because the quantization parameters are used in both rate control algorithm and rate distortion optimization (RDO), which resulted in the following chicken and egg dilemma when the rate control is studied: to perform RDO for macroblocks (MBs) in the current frame, a quantization parameter should be first determined for each MB by using the mean absolute difference (MAD) of current frame or MB. However, the MAD of current frame or MB is only available after the RDO.

Addition processes can be employed to: restrict maximum and minimum QP; if the video image is bright or includes a large amount of movement for a long time, clear R (which denotes bitrate profit and loss) for later dark or quiet scene; and If the video image is dark or quiet for a long time, clear R (which denotes bitrate profit and loss) for later bright or severely moving scene.

Still further optimization include: the use of MMX, SSE, and SSE2 instructions; avoiding access large global arrays by using small temporary buffer and pointers; and

Use omp parallel instruction, for examples:

```
#pragma omp parallel sections
{
    #pragma omp section
        master_thread20(&images_private_master);
    #pragma omp section
        slave_thread21(&images_private_slave1);
}
```

Having thus described exemplary embodiments of the present invention, it is to be understood that the invention defined by the appended claims is not to be limited by particular details set forth in the above description as many apparent variations thereof are possible without departing from the spirit or scope thereof as hereinafter claimed.

CLAIMS:

1. A method of decoding MPEG4 compliant coded signals, comprising:
disabling processing for non-main profile sections;
5 performing reference frame padding; and
performing adaptive motion compensation.
2. The method of claim 1, further including performing fast IDCT wherein an IDCT process
is performed on profile signals but no IDCT is performed based on whether a 4x4 block may be
10 all zero, or only DC transform coefficient is non-zero.
3. The method of claim 1, further including disabling non-main profile processing including
error correction and redundant slices processing.
- 15 4. The method of claim 1, further including CAVLC encoding of residual data.
5. The method of claim 1, wherein reference frame padding comprises compensating for
motion vectors extending beyond a reference frame by adding to at least the length and width
of the reference frame.
20
6. The method of claim 1, wherein adaptive motion compensation includes original block
size compensation processing for chroma up to block sizes of 16 x 16.
7. The method of claim 1, wherein disabling non-main profile slices include disabling SP
25 and SI slices;
8. The method of claim 1, wherein processing parameters are set according to:
 - (1) single frame reference (NumberReferenceFrames = 1);
 - (2) one slice in each frame (SliceMode = 0);
 - 30 (3) I frame interval is 100 (IntraPeriod = 100);

(4) quantization parameter is 25 (QPFirstFrame = QPRemainingFrame = 25);

(5) not using Hadamard (UseHadamard = 0);

(6) max search range is 16 (SearchRange = 16);

(7) not using rate distortion optimization (RDOptimization = 0);

5 (8) using all blocks types from 16x16 to 4x4 (InterSearch16x16 = 1, InterSearch16x8 = 1, InterSearch8x16 = 1, InterSearch8x8 = 1, InterSearch8x4 = 1, InterSearch4x8 = 1, InterSearch4x4 = 1);

(9) Inter pixels are used for Intra macroblock prediction (UseConstrainedIntraPred = 0);

(10) POC mode is 0 (PicOrderCntType = 0).

10 (11) Entropy coding method is 0, i.e. CAVLC (SymbolMode = 0).

9. A method of encoding MPEG4 compliant data, comprising: performing Rate Distortion Optimization (RDO) algorithm, fast motion search algorithm, and bitrate control algorithm.

15 10. A method according to claim 9, further including use of MMX/SSE/SSE2 instructions.

1/4

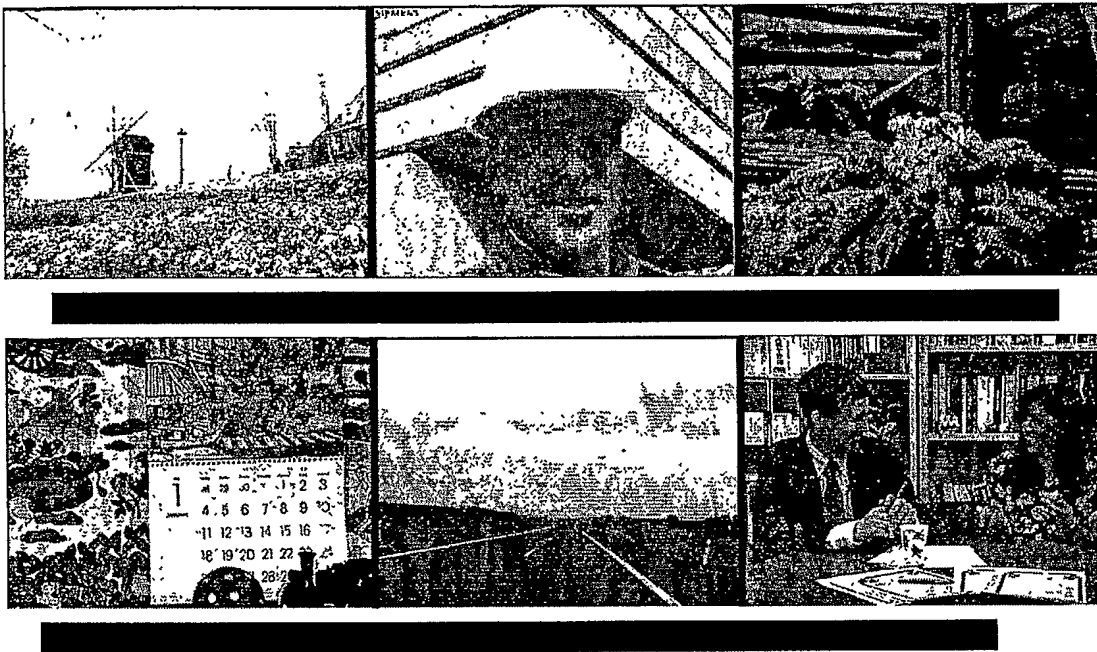


Figure1 Standard Video Series

2/4

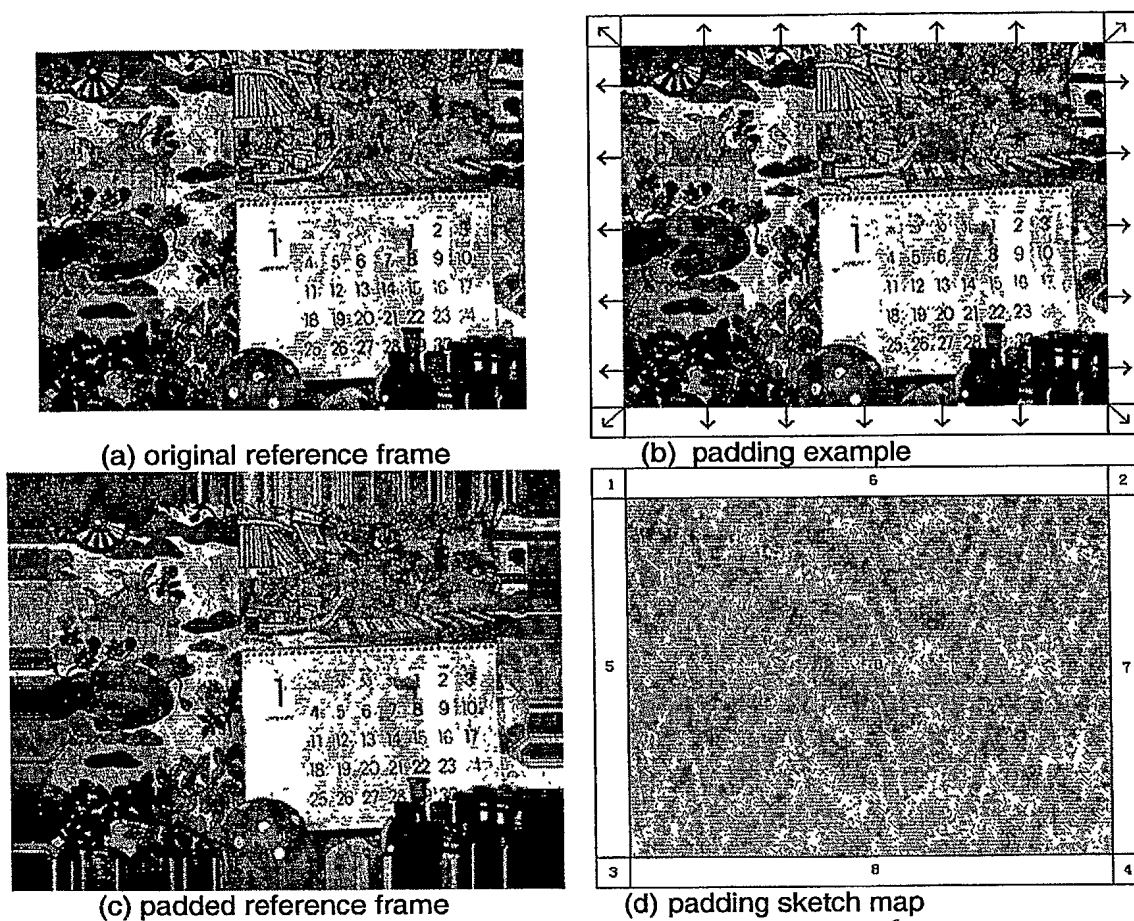


Figure 2 Reference frame padding

3/4

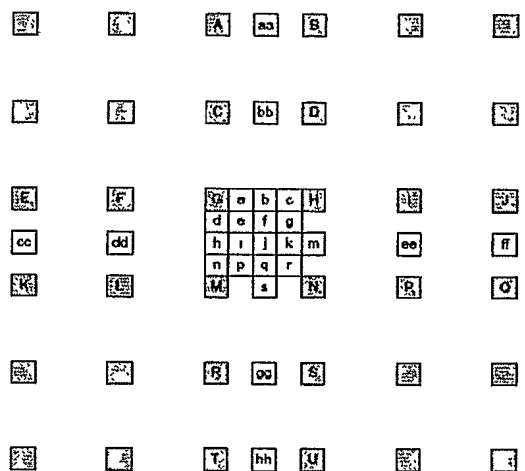


Figure 3 Integer samples (shaded blocks with upper-case letters) and fractional sample positions (un-shaded blocks with lower-case letters) for quarter sample luma interpolations

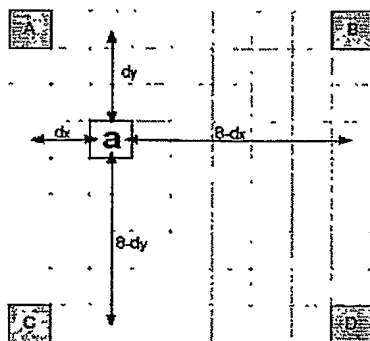


Figure 4 Interpolation of chroma eighth-pel positions

SUBSTITUTE SHEET (RULE 26)

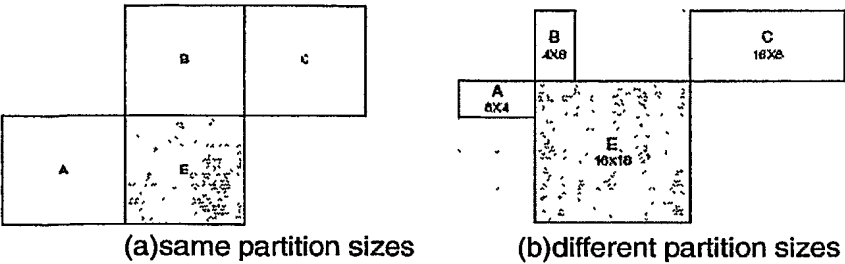


Figure 5 Current and neighbouring partitions

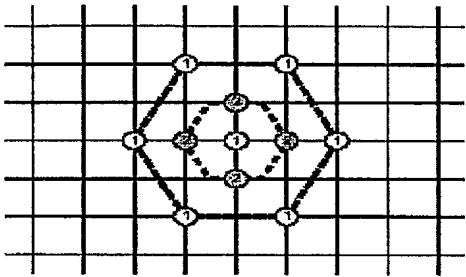


Figure 6 Improved Hexagon Search □Large Hexagon □Small Hexagon

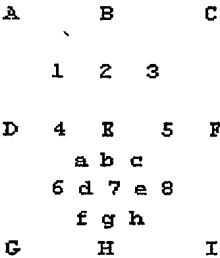


Figure 7 Full search for fractional pel search

Electronic Patent Application Fee Transmittal

Application Number:	13389872			
Filing Date:	10-Feb-2012			
Title of Invention:	METHODS AND APPARATUS FOR IMPROVED INTRA CHROMA ENCODING AND DECODING			
First Named Inventor/Applicant Name:	Qian Xu			
Filer:	Ronald J. Kolczynski/Brian Pritchett			
Attorney Docket Number:	PU090108			
Filed as Large Entity				
Filing Fees for U.S. National Stage under 35 USC 371				
Description	Fee Code	Quantity	Amount	Sub-Total in USD(\$)
Basic Filing:				
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD(\$)
Miscellaneous:				
Submission- Information Disclosure Stmt	1806	1	180	180
Total in USD (\$)				180

Electronic Acknowledgement Receipt

EFS ID:	30514920
Application Number:	13389872
International Application Number:	
Confirmation Number:	7293
Title of Invention:	METHODS AND APPARATUS FOR IMPROVED INTRA CHROMA ENCODING AND DECODING
First Named Inventor/Applicant Name:	Qian Xu
Customer Number:	24498
Filer:	Ronald J. Kolczynski/Brian Pritchett
Filer Authorized By:	Ronald J. Kolczynski
Attorney Docket Number:	PU090108
Receipt Date:	29-SEP-2017
Filing Date:	10-FEB-2012
Time Stamp:	09:44:03
Application Type:	U.S. National Stage under 35 USC 371

Payment information:

Submitted with Payment	yes
Payment Type	DA
Payment was successfully received in RAM	\$ 180
RAM confirmation Number	092917INTEFSW00008883070832
Deposit Account	
Authorized User	

The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows:

File Listing:					
Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Foreign Reference	CN1487474.pdf	2063000	no	38
			d6a9e0ffd22fbac86bdc29d366a1a901c842a079		
Warnings:					
Information:					
2	Non Patent Literature	CN1487474_TR_EN_machine_28SEP2017.pdf	894788	no	18
			9a22ed7f0ec57938f5bdfd91f4981aca0a7f498b		
Warnings:					
Information:					
3	Foreign Reference	CN101438591.pdf	1186922	no	27
			3def74c6c08dafb9f9f05aa2925ec30746a8add1		
Warnings:					
Information:					
4	Non Patent Literature	CN101438591_TR_EN_machine_28SEP2017.pdf	931299	no	30
			8a6cf326499b29339c61b2aee68d16edd0f61ae4		
Warnings:					
Information:					
5	Information Disclosure Statement (IDS) Form (SB08)	PU090108_US_PCT_SB08_signed.pdf	611922	no	4
			a4d3e8b6ec669a6df5d8d4eca11d269bcbf64d87		
Warnings:					
Information:					
A U.S. Patent Number Citation or a U.S. Publication Number Citation is required in the Information Disclosure Statement (IDS) form for autoloading of data into USPTO systems. You may remove the form to add the required data in order to correct the Informational Message if you are citing U.S. References. If you chose not to include U.S. References, the image of the form will be processed and be made available within the Image File Wrapper (IFW) system. However, no data will be extracted from this form. Any additional data such as Foreign Patent Documents or Non Patent Literature will be manually reviewed and keyed into USPTO systems.					

6	Foreign Reference	WO2006124885.pdf	2081744	no	44
			0c1923d772a68ac3e5829da1ec0aa82985fa5d8		

Warnings:

Information:

7	Fee Worksheet (SB06)	fee-info.pdf	31012	no	2
			f61497fc9c58c8ec4228abc394a0e93484bc6c5b		

Warnings:

Information:

Total Files Size (in bytes):			7800687
-------------------------------------	--	--	---------

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.



[12] 发明专利申请公开说明书

[21] 申请号 03147131.5

[43] 公开日 2004 年 4 月 7 日

[11] 公开号 CN 1487474A

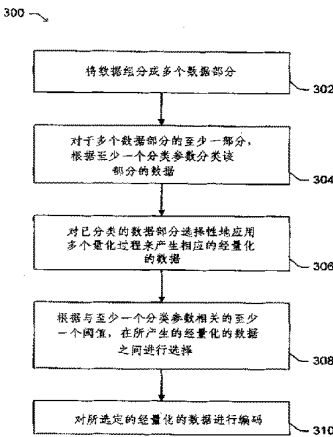
[22] 申请日 2003.7.4 [21] 申请号 03147131.5
[30] 优先权
[32] 2002. 7. 5 [33] US [31] 60/393,894
[71] 申请人 微软公司
地址 美国华盛顿州
[72] 发明人 A·图雷皮斯 S·-P·李
F·吴 G·沙利文

[74] 专利代理机构 上海专利商标事务所
代理人 李家麟

权利要求书 11 页 说明书 16 页 附图 10 页

[54] 发明名称 用于数据压缩的改进优化技术
[57] 摘要

提供了与对诸如图像、视频等数据进行编码有关的方法和装置。例如，某些方法包括：用多个不同的量化功能来处理至少一部分数据以产生多个相应的经量化的数据部分，并且根据至少一个阈值选择性地输出一个经量化的数据部分。该方法还包括：将初始数据分成多个部分，并且根据至少一个分类特性对数据的部分分类。这里，例如，可能有与分类特性相关的阈值。通过允许若干替代的运动预测情况，可以采用附加的语法来允许相当高的压缩效率。也提供了一种高效率时间标记独立的直接模式，它考虑到空间运动矢量预测以及静态时间预测器。



1. 一种方法，其特征在于包括：

用多个不同的量化功能处理至少一部分数据以产生多个相应的经量化的数据部分；以及

根据至少一个阈值选择性地输出所述经量化的数据部分之一。

2. 如权利要求 1 所述的方法，其特征在于还包括：

将初始数据分成多个部分，包括所述至少一部分数据；以及

根据至少一种分类特性分类所述至少一部分数据。

3. 如权利要求 2 所述的方法，其特征在于，所述至少一个阈值与所述至少一种分类特性相关。

4. 如权利要求 2 所述的方法，其特征在于，所述初始数据包括从包括图像数据、视频数据、音频数据和语音数据的一组中选择的数据。

5. 如权利要求 2 所述的方法，其特征在于，所述初始数据包括图像数据或视频数据，所述至少一部分从包括块、宏块、条、片和区段的一组中选出。

6. 如权利要求 2 所述的方法，其特征在于，所述至少一种分类特性从一组特性中选出，该组特性包括：边缘特性、纹理特性、平滑特性、亮度特性、色度特性、色彩特性、噪声特性、对象特性、运动特性、用户喜好特性、用户界面聚焦特性、分层特性、定时特性、音量特性、频率特性、音调特性、音色特性、质量特性、比特率特性、数据类型特性、分辨率特性、以及加密特性。

7. 如权利要求 1 所述的方法，其特征在于，所述多个不同的量化功能包括至少两个操作上不同的恒域量化器。

8. 如权利要求 7 所述的方法，其特征在于，至少一个所述的恒域量化器包括一个自适应覆盖范围恒域量化器。

9. 如权利要求 7 所述的方法，其特征在于，至少一个所述恒域量化器包括一个可变覆盖范围大小的恒域量化器。

10. 如权利要求 1 所述的方法，其特征在于，选择性地输出所述一个经量化的数据部分还包括：

对所述一个经量化的部分进行编码。

11. 如权利要求 1 所述的方法，其特征在于，根据所述至少一个阈值选择性地输出所述一个经量化的数据部分还包括：执行码率失真优化（RDO）来选择所述

一个经量化的数据部分。

12. 一种方法，其特征在于包括：

对至少一部分图像数据执行至少一个特性分析；

根据所述至少一个特性分析在编码器内选择性地设置至少一个自适应量化参数；以及

用所述编码器对所述至少一部分图像数据进行编码。

13. 如权利要求 12 所述的方法，其特征在于，所述至少一个特性分析考虑到从一组特性中选择的至少一个图像分析特性，该组特性包括：边缘特性、纹理特性、平滑特性、亮度特性、色度特性、色彩特性、噪声特性、对象特性、运动特性、用户喜好特性、用户界面聚焦特性、分层特性、定时特性、音量特性、频率特性、音调特性、音色特性、质量特性、比特率特性、数据类型特性、分辨率特性、以及加密特性。

14. 如权利要求 12 所述的方法，其特征在于，所述至少一个自适应量化参数与所述编码器内的自适应覆盖范围恒域量化器相关。

15. 如权利要求 12 所述的方法，其特征在于，当对于包括视频帧的所述图像数据的多个部分的每一部分而言，在运行中根据所述至少一个特性分析在所述编码器内选择性地设置所述至少一个自适应量化参数。

16. 一种方法，其特征在于包括：

用至少两个不同的恒域量化器使至少一部分图像数据要被编码；以及

根据与至少一个判决因子相关的码率失真优化（RDO）判决在所述至少两个不同恒域量化器之一的输出中识别较佳的已编码数据。

17. 一种方法，其特征在于包括：

用第一恒域量化器使至少一部分图像数据要被编码；

确定所述第一恒域量化器的输出是否满足至少一个判决因子，以及

如果是，则输出所述第一恒域量化器的所述输出，

如果不是，则用不同于所述第一恒域量化器的至少一个第二恒域量化器使所述至少一部分图像数据要被编码。

18. 如权利要求 17 所述的方法，其特征在于还包括：

根据码率失真优化（RDO）判决识别所述图像数据的所述至少一部分的可接受已编码版本。

19. 一种方法，其特征在于包括：

对所述至少一部分图像数据进行图像分析；
对所述图像数据的已分析部分进行块分类；
对所述图像数据的经块分类的部分进行恒域量化；以及
对所述图像数据的经恒域量化的部分进行编码。

20. 如权利要求 19 所述的方法，其特征在于，所述图像分析包括从包括边缘检测分析和纹理分析的一组中选择的至少一类分析。

21. 如权利要求 19 所述的方法，其特征在于，所述块分类根据所述图像分析操作上被配置，而所述恒域量化根据所述块分类操作上被配置。

22. 一种方法，其特征在于包括：

用至少两个不同的编码器使至少一部分视频图像数据要被编码，其中所述两个不同编码器的至少一个包括操作上配置为用于支持所述视频图像数据的无残留模式的恒域量化器；以及

根据与至少一个判决因子相关的码率失真优化（RDO）判决在所述至少两个不同编码器之一的输出中识别较佳的已编码帧数据。

23. 如权利要求 22 所述的方法，其特征在于，所述恒域量化器是自适应的并用于选择性地所述无残留模式被使用。

24. 如权利要求 22 所述的方法，其特征在于，至少一个所述编码器操作上配置为用于支持所述视频图像数据的直接模式。

25. 如权利要求 22 所述的方法，其特征在于，至少一个所述编码器操作上配置为用于支持所述视频图像数据的拷贝模式。

26. 如权利要求 22 所述的方法，其特征在于，所述无残留模式根据运行长度编码（RLC）策略操作上被配置。

27. 如权利要求 22 所述的方法，其特征在于，所述恒域量化器是操作上自适应的。

28. 如权利要求 22 所述的方法，其特征在于，所述恒域量化器根据与至少一部分视频图像数据相关的至少一个特性操作上被配置。

29. 如权利要求 22 所述的方法，其特征在于，所述识别所述较佳已编码帧数据还包括：分析与所述直接模式帧数据相关的残留量。

30. 如权利要求 29 所述的方法，其特征在于，分析所述残留量还包括检查所述直接模式帧数据的已编码块模式（CBP）。

31. 如权利要求 22 所述的方法，其特征在于，带有所述恒域量化器的所述编

码器选择性地配置为移去从包括色度残留数据、亮度残留数据和所有残留数据的一组残留数据中选择的至少一类残留。

32. 一种方法，其特征在于包括：

选择性地改变至少一个拉格朗日乘法器，该乘法器操作上配置了编码逻辑，编码逻辑带有基于至少一部分图像数据的至少一个特性的量化功能；以及用所述编码逻辑对所述图像数据的所述至少一部分进行编码。

33. 如权利要求 32 所述的方法，其特征在于，所述量化功能包括基于下式的宏块量化器：

$$J(\text{Mode}|\text{QP}, \lambda) = \text{SSD}(\text{Mode}|\text{QP}) + \lambda \cdot R(\text{Mode}|\text{QP})$$

其中 λ 是用于模式判决的所述拉格朗日乘法器，而 Mode 表示用码率失真优化（RDO）过程检查到的宏块模式。

34. 如权利要求 32 所述的方法，其特征在于，所述拉格朗日乘法器为帧间或内部帧而被选择。

35. 如权利要求 33 所述的方法，其特征在于：

$$\lambda = f(\text{QP}) \times \lambda_{I,P}; \text{ 以及}$$

$$f(\text{QP}) = \max\left(2, \min\left(4, \frac{\text{QP}}{6}\right)\right)$$

36. 如权利要求 33 所述的方法，其特征在于：

$$\lambda = f(\text{QP}) \times \lambda_{I,P}; \text{ 以及}$$

$$f(\text{QP}) = \max\left(2, \min\left(4, \frac{\text{QP} + 2}{8}\right)\right)$$

37. 一种方法，其特征在于包括：

用编码器逻辑对至少一部分视频图像数据进行编码；以及

使所述编码器逻辑输出语法信息，该语法信息识别由所述编码器逻辑使用的一类运动矢量预测。

38. 如权利要求 37 所述的方法，其特征在于，所述编码器逻辑用于选择性地使用从包括空间运动矢量预测和时间运动矢量预测在内的一组中选择的至少一类运动矢量预测，且其中所述编码器逻辑至少部分根据速率失真优化（RDO）判决选择性地输出所述语法信息和相应的已编码视频图像数据。

39. 一种用于传送与视频编码有关的信息的方法，该方法的特征在于包括：

编码视频数据；以及

在所述已编码视频数据的语法部分内选择性地设置至少一个描述符，所述描

述符识别与用所述视频数据编码的至少一个 B 帧相关的自适应空间/空间-时间编码。

40. 如权利要求 39 所述的方法, 其特征在于, 编码所述视频数据包括按照 JVT 标准来编码所述视频数据, 而所述已编码视频数据的所述语法部分内的所述至少一个描述符包括: 带有所述已编码视频数据的图像层部分的 `direct_mv_spatial` 参数。

41. 一种用于传送与视频编码有关的信息的方法, 该方法的特征在于包括:
编码视频数据; 以及

在所述已编码视频数据的语法部分内选择性地设置至少一个描述符, 所述描述符识别用所述视频数据编码的至少一个中间帧内的自适应拷贝/运动拷贝跳跃模式。

42. 如权利要求 41 所述的方法, 其特征在于, 编码所述视频数据包括按照 JVT 标准来编码所述视频数据, 而所述已编码视频数据的所述语法部分内的所述至少一个描述符包括: 带有所述已编码视频数据的图像层部分的 `copy_mv_spatial` 参数。

43. 一种用在对考虑视频帧的静态时间/空间部分的视频进行时间标记独立模式编码中的方法, 该方法的特征在于包括:

在视频序列中选择性地应用与至少一部分视频帧相关的运动的空间预测; 以及

如果另一视频帧的参考部分的时间运动预测信息为零, 则将所述运动的空间预测设为零。

44. 如权利要求 43 所述的方法, 其特征在于, 所述空间预测是空间零偏置的。

45. 如权利要求 43 所述的方法, 其特征在于, 所述空间预测还考虑到所述视频帧中其它部分的静态子分区, 且其中, 如果至少一个所述子分区具有大约为零的相应运动, 则所述其它部分也被视作具有零运动。

46. 一种带有计算机可实现的指令的计算机可读媒体, 指令用于配置至少一个处理单元来执行下列动作:

用多个不同的量化函数来处理至少一部分数据以产生多个相应的经量化的数据部分; 以及

根据至少一个阈值选择性地输出一个所述经量化的数据部分。

47. 如权利要求 46 所述的计算机可读媒体, 其特征在于还包括:

将初始数据分成多个部分, 包括所述至少一部分数据; 以及

根据至少一个分类特性对所述至少一部分数据进行分类。

48. 如权利要求 47 所述的计算机可读媒体, 其特征在于, 所述至少一个阈值与所述至少一个分类特性相关。

49. 如权利要求 47 所述的计算机可读媒体, 其特征在于, 所述初始数据包括从包括图像数据、视频数据、音频数据和语音数据的一组数据中选择的数据, 且其中所述初始数据包括图像数据或视频数据, 且所述至少一部分从包括块、宏块、条、片和区段的一组中选出。

50. 如权利要求 47 所述的计算机可读媒体, 其特征在于, 所述至少一个分类特性从一组特性中选出, 该组特性包括: 边缘特性、纹理特性、平滑特性、亮度特性、色度特性、色彩特性、噪声特性、对象特性、运动特性、用户喜好特性、用户界面聚焦特性、分层特性、定时特性、音量特性、频率特性、音调特性、音色特性、质量特性、比特率特性、数据类型特性、分辨率特性、以及加密特性。

51. 如权利要求 46 所述的计算机可读媒体, 其特征在于, 所述多个不同的量化功能包括至少两个操作上不同的恒域量化器, 且其中至少一个所述恒域量化器包括一个自适应覆盖范围的恒域量化器。

52. 如权利要求 46 所述的计算机可读媒体, 其特征在于, 根据所述至少一个阈值选择性地输出所述一个经量化的数据部分还包括: 执行码率失真优化(RDO)以选择所述一个经量化的数据部分。

53. 一种带有计算机可实现的指令的计算机可读媒体, 指令用于配置至少一个处理单元来执行下列动作:

对至少一部分图像数据执行至少一个特性分析;

根据所述至少一个特性分析在编码器内选择性地设置至少一个自适应量化参数; 以及

用所述编码器对所述至少一部分图像数据进行编码。

54. 如权利要求 53 所述的计算机可读媒体, 其特征在于, 所述至少一个自适应量化参数与所述编码器内的自适应覆盖范围的恒域量化器相关联。

55. 如权利要求 53 所述的计算机可读媒体, 其特征在于, 对于包括视频帧的所述图像数据的多个部分的每一部分而言, 在运行根据所述至少一个特性分析在所述编码器内选择性地设置所述至少一个自适应量化参数。

56. 一种带有计算机可实现的指令的计算机可读媒体, 指令用于配置至少一个处理单元来执行下列动作:

用至少两个不同的恒域量化器使至少一部分图像数据要被编码；以及
根据与至少一个判决因子相关的码率失真优化（RDO）判决在所述至少两个不同恒域量化器之一的输出中识别较佳的已编码数据。

57. 一种带有计算机可实现的指令的计算机可读媒体，所述指令用于配置至少一个处理单元来执行下列动作：

用第一恒域量化器使至少一部分图像数据要被编码；

确定所述第一恒域量化器的输出是否满足至少一个判决因子，以及

如果是，则输出所述第一恒域量化器的所述输出，

如果不是，则用不同于所述第一恒域量化器的至少一个第二恒域量化器使所述至少一部分图像数据要被编码。

58. 如权利要求 57 所述的计算机可读媒体，其特征在于还包括：

根据码率失真优化（RDO）判决识别所述图像数据的所述至少一部分的可接受已编码方式。

59. 一种带有计算机可实现的指令的计算机可读媒体，所述指令用于配置至少一个处理单元来执行下列动作：

对所述至少一部分图像数据进行图像分析；

对所述图像数据的已分析部分进行块分类；

对所述图像数据的经块分类的部分进行恒域量化；以及

对所述图像数据的经恒域量化的部分进行编码。

60. 如权利要求 59 所述的计算机可读媒体，其特征在于，所述图像分析包括从包括边缘检测分析和纹理分析的一组中选择的至少一类分析。

61. 如权利要求 59 所述的计算机可读媒体，其特征在于，所述块分类根据所述图像分析操作上被配置，而所述恒域量化根据所述块分类操作上被配置。

62. 一种带有计算机可实现的指令的计算机可读媒体，所述指令用于配置至少一个处理单元来执行下列动作：

用至少两个不同的编码器使至少一部分视频图像数据要被编码，其中所述两个不同编码器的至少一个包括操作上配置成用于支持所述视频图像数据的无残留模式的恒域量化器；以及

根据与至少一个判决因子相关的码率失真优化（RDO）判决在所述至少两个不同编码器之一的输出中识别较佳的已编码帧数据。

63. 如权利要求 62 所述的计算机可读媒体，其特征在于，所述恒域量化器是

自适应的并用于选择性地使用所述无残留模式。

64. 如权利要求 62 所述的计算机可读媒体, 其特征在于, 至少一个所述编码器操作上配置为用于支持所述视频图像数据的直接模式。

65. 如权利要求 62 所述的计算机可读媒体, 其特征在于, 至少一个所述编码器操作上配置为用于支持所述视频图像数据的拷贝模式。

66. 如权利要求 62 所述的计算机可读媒体, 其特征在于, 所述无残留模式根据运行长度编码(RLC)策略操作上被配置。

67. 如权利要求 62 所述的计算机可读媒体, 其特征在于, 带有所述恒域量化器的所述编码器选择性地配置为消除从包括色度残留数据、亮度残留数据和所有残留数据的一组残留数据中选择的至少一类残留。

68. 一种带有计算机可实现的指令的计算机可读媒体, 指令用于配置至少一个处理单元来执行下列动作:

选择性地改变至少一个拉格朗日乘法器, 该乘法器操作上配置了编码逻辑, 编码逻辑带有基于至少一部分图像数据的至少一个特性的量化功能; 以及
用所述编码逻辑对所述图像数据的所述至少一部分进行编码。

69. 一种带有计算机可实现的指令的计算机可读媒体, 指令用于配置至少一个处理单元来执行下列动作:

用编码器逻辑对至少一部分视频图像数据进行编码; 以及
使所述编码器逻辑输出语法信息, 该语法信息识别由所述编码器逻辑使用的一类运动矢量预测。

70. 一种带有计算机可实现的指令的计算机可读媒体, 指令用于配置至少一个处理单元来执行下列动作:

编码视频数据; 以及
在所述已编码视频数据的语法部分内选择性地设置至少一个描述符, 所述描述符识别与用所述视频数据编码的至少一个 B 帧相关的自适应空间/空间-时间编码。

71. 如权利要求 70 所述的计算机可读媒体, 其特征在于, 编码所述视频数据包括按照 JVT 标准来编码所述视频数据, 而所述已编码视频数据的所述语法部分内的所述至少一个描述符包括: 带有所述已编码视频数据的图像层部分的 `direct_mv_spatial` 参数。

72. 一种带有计算机可实现的指令的计算机可读媒体, 指令用于配置至少一

个处理单元来执行下列动作：

编码视频数据；以及

在所述已编码视频数据的语法部分内选择性地设置至少一个描述符，所述描述符识别用所述视频数据编码的至少一个中间帧内的自适应拷贝/运动拷贝跳跃模式。

73. 如权利要求 72 所述的计算机可读媒体，其特征在于，编码所述视频数据包括按照 JVT 标准来编码所述视频数据，而所述已编码视频数据的所述语法部分内的所述至少一个描述符包括：带有所述已编码视频数据的图像层部分的 `copy_mv_spatial` 参数。

74. 一种带有计算机可实现的指令的计算机可读媒体，指令用于配置至少一个处理单元来执行下列动作：

在视频序列中选择性地应用与至少一部分视频帧相关的运动的空间预测；以及

如果另一视频帧的参考部分的时间运动预测信息为零，则将所述运动的空间预测设为零。

75. 如权利要求 74 所述的计算机可读媒体，其特征在于，所述空间预测是空间零偏置的。

76. 如权利要求 74 所述的计算机可读媒体，其特征在于，所述空间预测还考虑到所述视频帧中其它部分的静态子分区，并且其中，如果至少一个所述子分区具有大约为零的相应运动，则所述其它部分也被视为具有零运动。

77. 一种装置，操作上被配置成用多个不同的量化功能来处理至少一部分数据以产生多个相应的经量化的数据部分，并且根据至少一个阈值选择性地输出一个所述经量化的数据部分。

78. 如权利要求 77 所述的装置，其特征在于，所述逻辑还用于将初始数据分成包括所述至少一部分数据的多个部分，并且根据至少一个分类特性对所述至少一部分数据进行分类，且其中所述至少一个阈值与所述至少一个分类特性相关。

79. 如权利要求 77 所述的装置，其特征在于，所述多个不同的量化功能包括至少两个操作上不同的恒域量化器，且其中至少一个所述恒域量化器包括一个自适应覆盖范围的恒域量化器。

80. 如权利要求 77 所述的装置，其特征在于，所述逻辑还用于执行码率失真优化（RDO）以选择所述一个经量化的数据部分。

81. 一种包括可操作上配置为执行下列动作的逻辑的装置：对至少一部分图像数据执行至少一个特性分析，根据所述至少一个特性分析在编码器内选择性地设置至少一个自适应量化参数，以及用所述编码器对所述至少一部分图像数据进行编码。

82. 如权利要求 81 所述的装置，其特征在于，所述至少一个自适应量化参数与所述编码器内的自适应覆盖范围的恒域量化器相关。

83. 一种包括操作上配置为执行下列动作的逻辑的装置：用至少两个不同的恒域量化器使至少一部分图像数据要被编码，以及根据与至少一个判决因子相关的码率失真优化（RDO）判决在所述至少两个不同恒域量化器之一的输出中识别较佳的已编码数据。

84. 一种包括操作上配置为执行下列动作的逻辑的装置：用第一恒域量化器使至少一部分图像数据要被编码；确定所述第一恒域量化器的输出是否满足至少一个判决因子；以及，如果是，则输出所述第一恒域量化器的所述输出，否则，则用不同于所述第一恒域量化器的至少一个第二恒域量化器使所述至少一部分图像数据要被编码。

85. 如权利要求 84 所述的装置，其特征在于，所述逻辑还用于根据码率失真优化（RDO）判决识别所述图像数据的所述至少一部分的可接受已编码版本。

86. 一种包括操作上配置为执行下列动作的逻辑的装置：对所述至少一部分图像数据进行图像分析；对所述图像数据的已分析部分进行块分类；对所述图像数据的经块分类的部分进行恒域量化；以及对所述图像数据的经恒域量化的部分进行编码。

87. 一种包括操作上配置为执行下列动作的逻辑的装置：用至少两个不同的编码器使至少一部分视频图像数据要被编码，其中所述两个不同编码器的至少一个包括操作上用于支持所述视频图像数据的无残留模式的恒域量化器；以及根据与至少一个判决因子相关的码率失真优化（RDO）判决在所述至少两个不同编码器之一的输出中识别较佳的已编码帧数据。

88. 如权利要求 87 所述的装置，其特征在于，所述恒域量化器用于选择性地使用所述无残留模式。

89. 如权利要求 87 所述的装置，其特征在于，至少一个所述编码器操作上配置为支持从一组中选出的至少一个模式，该组包括所述视频图像数据的直接模式和拷贝模式。

90. 一种包括操作上配置为执行下列动作的逻辑的装置：选择性地改变至少一个拉格朗日乘法器，该乘法器操作上配置了编码逻辑，编码逻辑带有基于至少一部分图像数据的至少一个特性的量化功能；以及用所述编码逻辑对所述图像数据的所述至少一部分进行编码。

91. 一种包括操作上配置为执行下列动作的逻辑的装置：用编码器逻辑对至少一部分视频图像数据进行编码；以及使所述编码器逻辑输出语法信息，该语法信息识别由所述编码器逻辑使用的一类运动矢量预测。

92. 一种包括操作上配置为执行下列动作的逻辑的装置：编码视频数据；以及在所述已编码视频数据的语法部分内选择性地设置至少一个描述符，所述描述符识别与用所述视频数据编码的至少一个 B 帧相关的自适应空间/空间-时间编码。

93. 一种包括操作上配置为执行下列动作的逻辑的装置：编码视频数据；以及在所述已编码视频数据的语法部分内选择性地设置至少一个描述符，所述描述符识别用所述视频数据编码的至少一个中间帧内的自适应拷贝/运动拷贝跳跃模式。

94. 一种包括操作上配置为执行下列动作的逻辑的装置：在视频序列中选择性地应用与至少一部分视频帧相关的运动的空间预测；以及，如果另一视频帧的参考部分的时间运动预测信息为零，则将所述运动的空间预测设为零。

用于数据压缩的改进优化技术

相关的专利申请

该美国专利要求美国临时专利申请序列号为 60/393,894 的申请的优先权, 该申请通过引用被完全结合于此, 于 2002 年 7 月 5 日公开、题为 “Rate/Distortion Optimization Techniques for Image and Video Compression”。

(1) 技术领域

本发明通常涉及到的计算机和类似的设备, 特别是涉及用于压缩 / 编码数据和解压缩 / 解码数据的方法、装置和系统。

(2) 背景技术

对于数据压缩 / 编码和对于数据解压缩 / 解码需要有继续改进的方法和装置, 并且, 特别是对图象和视频数据。编码效率方面得的改进允许更多的信息更容易被计算机和其它的相似设备处理传送和 / 或存贮。随着国际互联网和其它相似的计算机网络以及有线通讯系统的进一步普及, 期望提供高效率的编码技术来做到充分利用可供利用的资源。

码率失真优化技术(RDO) 在视频和图象编码 / 解码系统中是十分流行的, 因为比之于比较常规的编码方法它们能非常好地改善编码效率。

举例来说, 附加的信息, 可能在计算机科学和工程的一篇科学硕士论文中可以找到, 题目为 “用于图象编码的一个局部自适应掩模的限模式”, 作者 Trac Duy Tran, MIT 麻省理工学院, May 1994.

就像存在继续希望来提供甚至更高的编码效率那样, 也有一个对改善方法和设备的需要, 它进一步的增加 RDO 或者其它类似技术的性能, 相对于已存在的系统来改善编码效率。

(3) 发明内容

当前的发明提供改善的方法和设备, 它能用于压缩 / 编码和解压缩 / 解码的

数据中, 和 / 或以其它方式处理各种所包括, 但不限于, 图象、视频和 / 或声频数据的数据的格式。

举例来说, 上述需要和其方面可通过下述方法满足, 这个方法它包括处理至少一部分的数据, 它用多个不同的量化的函数来产生多个相应的量化部分的数据, 以及基于至少在一个门限值选择输出一个量化部分的数据。此方法也可能包括把初始化数据分成多个部分, 以及基于至少一个分类特性对这一部分数据分类。举例来说, 这里可能有一个门槛值分类特性相关联。

作为例子, 初始化数据可能包括图像数据、视频数据、声频数据、语音数据, 以及此类似的数据。被选择的部分可能采取一块、一个宏块、一条、一片、一段, 或者类似的格式。分类特性可能包括边缘特性、纹理特性、平滑特性亮度特性、色彩特性、颜色特性、噪声特性, 对象特性、动作特性、用户喜好特性、用户界面聚焦特性、分层特性、定时特性、音量特性、频率特性、音调特性、音色特性、质量特性、比特率特性、数据类型特性、分辨率特性、加密特性, 等等的特性。

在某些示例性的实施中, 大多数不同的量化特性至少包括二个不同的可操作的恒域量化器。举例来说, 这里, 恒域量化器可能是自适应收敛恒域量化器、可变收敛大小的恒域量化器或者其它类似的恒域量化器。此方法也可能包括编码此量化的部份。此方法可能进一步包括实施码率失真优化(RDO) 去选择数据的量化部分。

根据还有的其它示例实施, 另一种方法包括在至少图象数据的一部分中至少执行一个特性分析, 基于该特性分析在编码器由选择性设置至少一个自适应量化参数, 以及用该编码器对一部分图象数据编码。

在还有另一个示例性的实施中, 被提供了一种方法, 它包括至少用二个不同的恒域量化器引起图象数据的至少一部分被编码, 以及基于与至少一种判决因素相关连的码率失真优化(RDO) 判决识别在至少二个不同恒域量化器其中的一个输出中的较佳编码数据。

上述需要和其它也被一种方法所满足, 此方法包括至少引起一部分图象数据被用第一只恒域量化器所编码, 确定是否第一只恒域量化器的输出满足至少一种判决因素, 同时假如是这样输出第一恒域量化器的输出。如果不是这样, 那么引起一部分图象数据被用至少与第一恒域量化器不同的第二恒域量化器所编码。举例来说, 该方法也可能包括基于 RDO 判决或者其它相似的判决识别一部分图象数据的可接受的编码方式。

在仍旧是另一种示例性的实施中, 一种方法包括在至少一部分图象数据中执

行图象分析,在被分析的这一部分图象数据上执行块分类,执行块分类部分图象数据的恒域量化,以及执行图象数据恒域量化部分的编码。举例来说,这里,此图象分析可能包括边缘检测分析、纹理分析等等。

根据进一步示例性实施,提供了一种方法,它包括至少引起一部分视频图象数据被采用至少二个不同的编码器所编码,其特征在于,所述的至少二个不同的编码器其中的一个包括恒域量化器,可操作地配置为去支持无残留模式的视频图象数据以及基于与至少一个判决的因素相关的码率失真在二个不同的编码器其中一个的输出中识别较佳的编码帧数据。优化(RDO)判决。

另一种示例性方法包括选择性地改变至少一个 Lagrangian 乘数,这是可操作配置的编码逻辑,它具有基于至少一部分图象数据的至少一个特性的量化函数,以及用编码逻辑来编码这一部分的图象数据。

仍旧在另一种执行中,示例方法包括采用编码逻辑编码至少一部分视频图象数据,以及使用编码器逻辑使编码逻辑输出标识运动矢量预测类型的语法信息。

用在表达视频编码关系信息的一种方法包括编码视频数据,以及在编码视频数据的语法部分的选择性地设置至少一种描述符,此描述符标识与用视频数据编码的至少一个 B 帧有关的一种匹配的空间/空间-时间编码。用在表达视频编码相关信息的一种方法,包括编码视频数据,以及在编码的视频数据的语法部分选择性地设置至少有一种描述符,此描述符至少一个用视频数据编码的中间帧中标识一种匹配拷贝/运动拷贝跳跃模式。

提供了考虑视频帧的时间/空间部分的视频编码用于时间标记独立模式中的示例性方法。举例来说,这里,此方法包括选择性地应用与视频序列的至少一部分视频帧相关的运动的空间预测。并且如果对于另一视频帧的参照部分的时间运动预测信息为另,则,设置运动的空间预测为另。

(4)附图说明

对本发明的各种方法设备和系统更完整的了解,可参照下述详细的描述并结合附图而给出:

图 1 是一张框图,它以计算机形式描述了一个实例的设备,它适用于本发明的某些实现中。

图 2(a-d) 是按照本发明的某些实现的一组描述示例性可选择恒域量化器的图表。

图 3 是一张流程图，它描述了按照本发明的某些实现选择性地对数据应用不同量化处理的示例方法。

图 4 是一张框图，描述了按照本发明的某些实现选择性地对数据应用不同量处理的示例逻辑。

图 5 是一张框图，描述了选择性地对数据的应用不同量处理的示例逻辑，它是与本发明的某些进一步的实施相符合的示例性逻辑。

图 6(a-b) 是逻辑框图，描述了选择性地对数据的应用不同量处理的示例逻辑，它与本发明的仍旧与其它的实施相符合的示例性逻辑。

图 7 是一张逻辑框图，描述了选择性地对数据的应用不同量处理的示例逻辑，它与本发明的某些进一步的实现相符合。

图 8 是一张示例性语法信息的表格，按照本发明的某些实现用于对选择性地对运动矢量应用不同的预测方法的逻辑。

图 9 是一张框图，它描述了按照本发明的某些实现的示例性逻辑，用于选择性地对视频数据应用不同编码的图表。

图 10 是一张框图，它描述了按照本发明的某些步骤的实现的示例性逻辑，用于选择性对视频数据应用不同编码的图表。

图 11 是一张图解图，它描述了按照本发明的某些实施的某些视频序列的特性，它使用选定的应用编码方案。

图 12 是一张流程图，它描述了按照本发明的某些实现的一种示例性的方法，用于对直接模式视频序列的空间一时间预测。

(5) 具体实施方式

转到附图，其特征在于，所述的相同的参照数字被指示相同的元件，本发明解释为，在合适的计算环境下被实施。虽然不是必须的，本发明将以通常的计算机可执行指令的环境下来描述，比如由计算机服务器来执行的程序模块，它可以采取个人计算机，工作站、专门的服务器、多处理器、大型计算机等等的形式。通常，程序模块包括子程序、程序、对象、组成、数据结构等等，它执行专门的任务或者执行特殊的抽象的数据类型。本发明也可能在分布的计算环境下被使用。这里任务被远程处理设备来执行，它经过通讯网络来联接。在分布计算的环境下，程序模块可能被定位在本地和远程存贮设备二者之中。

示例性计算环境

图 1 描述了一个合适的计算环境 120 的实例，在这时，以后描述的方法和安排可在此环境中被实施。

示例性计算环境 120 只是一个合适计算环境的例子，以及不打算建议作任何的限制，比如对使用的范围，或者可改进方法的功能，以及在这里所安排的描述的限制。计算环境 120 不应被解释为与计算环境 120 中描述的任一组件或组件的组合有相关的依赖性要求。

改进的方法和这里的安排是可与很多其它通用或者专用的计算系统环境或者结构一同操作。

如图 1 所示，计算环境 120 包括以计算机 130 形式的通用计算设备。计算机 130 组件可包括一只或者几只处理器或者处理单元 132、系统存贮器 134、以及把各种系统部件，包括系统存贮器 134 耦合至处理器 132 的总线 136。

总线 136 表示一个或者多个若干总线结构类型的任何一种，它包括存贮器总线或者存贮器控制器、外部设备总线、加速图形端口，以及处理器和使用各种总线结构任何一种的本地总线。举例来说，(但不是以此为限)，这些结构包括工业标准 (ISA) 总线、微通道结构 (MCA) 总线、增强型 (EISA) 总线、视频电子标准协会 (VESA) 局部总线，以及外部设备互连 (PCI) 总线(也称为 Mezzanine 总线)。

计算机 130 典型地包括各种计算机可读媒体，这些媒体可能是任何可利用的媒体，它可由计算机 130 访问，并且包括易失性媒体和非易失性媒体、可移动式 and 不可移动式二者。

图 1 中，系统存贮 134 包括计算机可读媒体，以易失性存贮器的形式，比如，随机存取存贮器 (RAM) 140，和/或非易失性存贮器形式，比如，只读存贮器 (ROM) 138。基本输入/输出系统 (BIOS) 142，包括基本的子程序，它帮助在计算机 130 内的部件之间传送信息，比如，当启动的时候，被存贮在 ROM 138 中的信息。RAM 140 典型地包括数据和/或程序模块，它可立刻被处理器 132 访问和/或当前正被处理器 132 的操作。

计算机 130 可能进一步包括其他可移动式/不可移动式的、易失性/非易失性的计算机存贮媒体。举例来说，图 1 给示了一个硬盘驱动器 144，它用来读来自或者写到非移动式、非易失性的磁性媒体中(没有表示在图中同时，典型的称为“硬设备”)，磁盘驱动器 146 用于读来自或者写到可移动式、非易失性的磁盘 148 上(比如，“软磁盘”)，以及光盘驱动器 150 用来读来自或者写到可移动式、非易失性光

盘 152, 比如, CD-ROM、CD-R、CR-RW、DVD-ROM、DVD-RAM 或者其它的光媒体。硬盘驱动器 144、磁盘驱动器 146 以及光盘驱动 150 它们每个设备由一个或者多个接口 154 连接到总线 136 上。

驱动器和相关的计算机可读媒体提供了计算机可读指令、数据结构、程序模块、以及其它从计算机 130 来的数据的非易失性贮存。虽然此处示例性环境描述采用硬盘、可移动式磁盘 148 以及可移动式光盘 152, 本领域中的技术人员应当理解其它类型的可由计算机来存取所存贮的数据的计算机可读媒体, 比如, 盒式磁带、闪存卡、数字视频盘、随机存取存储器(RAMs)、只读存储器(ROM)、以及其它类似的设备, 也可以用在示例性的操作环境中。

一些程序模块可能被存贮在硬盘、磁盘 148、光盘 152、ROM138、或者 RAM140 中, 包括比如操作系统 158, 一只或者多只应用程序 160、其它的程序模块 162 和程序数据 164。

在这里描述的改进方法和安排其特征在于, 可用所述的操作系统 158、一只或多只应用程序 160、其它的程序模块 162、和/或程序数据 164 来实现。

用户可以经过输入设备, 比如键盘 166 的触点设备 168 (比如“鼠标”) 提供命令和信息进入计算机 130。其它的输入设备(图中没有表示出来)可能包括话筒、控制杆、游戏板、卫星式转盘、串行端口、扫描仪、照相机等等。这些和其它的输入设备通常经过用户输入接口 170 连接到处理单元 132, 它连接到总线 136 上, 但是也可采用其它的接口和总线结构来连接, 比如像平行端口、游戏端口或者适用的串行总线(USB)。

显示器 172 或者其它类型的显示设备也可以经过一个接口, 比如视频匹配器 174, 连接到总线 136 上。另外, 显示器 172、个人计算机典型地包括其它外部输出设备(图中没有表示出来), 比如扬声器和打印机, 它们可以经过输出外部设备接口 175 来进行连接。

计算机 130 可以在用逻辑连接到一台或多台远程计算机的网络环境下进行操作, 比如用远程计算机 182。远程计算机 182 可能包括在这里和计算机 130 相关连的很多或者所有的部件的特性的描述。

图 1 中所给示的逻辑连结是局域网(LAN)177 以及广域网 179。这样的网络环境是在办公室、企业级计算机网络、企业内部互联网(intranets)以及国际互联网是常见的。

当用在 LAN 局域网环境下时, 计算机 130 经过网络接口或者适配器 186 连接

到局域网 LAN177 上去。当用在广域网环境下时，计算机典型地包括有调制器 178 或者其它用于在 WAN179 广域网上建立通讯的装置。调制器 178 它可以是内置的或外置的，经过用户输入接口 170 或者其它合适的机制可以被连结到系统总线 136 上。

图 1 所描述的，是一个经过国际互联网的专门的 WAN 广域网的实施环境，这里，计算机 130 使用调制器 178 经过国际互联网 180 与至少一台远程计算机 182 建立的通讯。

在一个网络的环境下，所描述的相关于计算机 130 的程序模块或者其中的部分，可能贮存在远程存贮器贮存设备上。这样，像图 1 所描述的那样，远程应用程序 189 可能驻留在远程计算机 182 的存贮设备上。将会理解到，所给示的和所描述的网络连结是示例性的，也可以使用可以计算机之间建立通讯链路的其它装置。

改进的码率/失真优化技术

虽然下列部分描述了某些示例性方法和装置，它们初始被构成压缩/编码和解压/解码图象数据和/或视频数据，在数据压缩领域中的技术人员将认识到此技术目前能被适应和使用在其它类型的数据压缩/编码以及解压/解码方面。举例来说，某些方法和设备可能被适应在用于视频数据、语音数据以及其它类的数据的压缩/编码方面。

此外，虽然此示例性方法和设备能在有计算机内用逻辑来构成，本领域的技术人员将认识到这些方法和设备可能被用其它类型的设备、电器用具等来实施。用在此处的术语“逻辑”的含意是包括硬件、固件、软件或者任何它们的组合，以及任何其它的支持硬件或其它的机制，可能被要求用来实现所描述的整个的或者部分的功能。

记住了这些，一些示例性方案被呈现在这里，它能以某种逻辑形式实施去支持处理数据。根据本发明的某些方面，一些新颖的技术被介绍用于改进对视频/图象和编码/解码系统的性能。在一些示例性实施中，这些技术被与图象/视频码的标准，比如 JPEG 和 JVT(连接视频组)标准(比如，H. 264/AVC)一起使用。举例来说，对 JVT 的情况下，语法变化被介绍，它能操作上启用不同预测类型的自适应选择，这些不用预测类型可以被用于预测视频的某些参量，比如，运动信息。

码率失真优化(RDO)技术在视频和图象编码/解码系统中是十分流行的，因为比之于比较传统的编码方法它们能非常好地改进编码效率。存在一个不断的希望来

提供甚至更高的编码效率。这里描写描述方法和设备能显著地改善 RDO 或者其它类似的技术的性以达到相对地于现有系统改善编码效率。

按照本发明的某些进一步的方面，它也可以用图象预分析概念，比如，举例来说，边缘和/或纹理检测，根据图象或宏块的特性通过使用自适应和/或可变大小的恒域量化，进一步的组合和改进 RDO 或者其它类似的技术。

这一描述也介绍/定义了一些附加的示例性语法变化，它被执行去启动在一个帧层次上的不同预测方案的组合体，举例来说，这进一步的改善视频编码方案的执行。举例来说，一个技术提升来自这样的事实，它对序列的一些帧或者一部分、运动在时间域可能比在空间域更相，或者反过来也是这样。这可能通过执行帧的予分析而加以利用，但是也可通过使用二个或者可能更多不同的方法对相同帧编码并在 RDO 和/或类似于 RDO 的选择一个更合适的方法而被利用。此较佳的方法然后可在结果数据中用信号表示，举例来说，在帧头中以允许编码器适当的编码此帧。这里，举例来说，一个这样的示例性方法在 B 帧中可能包括直接模式可能的变化，通过或者采用空间预测或者时序预测达到，或者在 P 帧中包括跳跃模式的运动矢量选择，预测运动矢量参数或者时间预测动作矢量参数，或者甚至为另而达到。

根据一个或者多个特性(参数)对数据应用匹配和/或不定的恒域量化到：

在图象数据压缩系统中，一部分数据被压缩，比如，举例来说，块或者宏块，可能确实包括更多的重要的信息，当与其它的信息(数据)相比时应该被不同地编码(例如，以较高的优先级，以较高的质量，等等)。达到的一个办法是使用不同的量化器值。举例来说，在某些实施中，一个较小的量化器值可能被用于对“更加重要的信息”，以及一个较大的量化器值可能被用于对“较小重要的信息”。然而，这样做后一般还将要求此信息的传送标示出用于对每个块、宏块、宏块的组，等等的每个量化值，使得其后的解压/解码是成功的。不幸的是，这些附加的信息趋向于增加压缩的总开销，以及编码器的复杂性。因而，不是增加了效率，反而实际可能会降低了效率。

注意到图 2(a-d)，每张图表描述了某些示例性恒域量化可使用在某些示例性图象/视频编码的方案中。在图 2(a-d)的每一幅图表中，垂直轴(Y 轴)呈表示量化值以及水平轴(X 轴)表示原始值。图 2(a)中给示了示例性恒域量化 202，举例来说，有关于恒域量化 A，它相对于均匀量化能相当大地改进编码效率。常规的恒域量化器在整个量化过程中经常是保持为常数或者均匀量化，如此，可能的话，不完全利用数据中全部存在的冗余。

考虑到或者另外方式计及数据中信息的重要性，以及在块/宏块的基础上，通过自适应/修改恒域量化量可以达到编码效率的改进。举例来说，例如，通过自适应收敛（例如，沿的 X 轴），但不改变重建值就能做到。举例来说，分别地将恒域量化器 202 和恒域量化器 204，206，和 208 比较，（它们表示在图 2b，c，d 中）。这里，重建值可能对全部的编码数据保持为常数，当然除非发出量化参量有变化的信号。举例来说，由增加此另电平箱位（图 2(c) 中），使更多的数据将被分配到这里，这样，依赖于压缩的方案，达到较高的压缩。显而易见地是，不需要在解码器上述信号表示量化器中的改变，因为重建值仍旧保持相同值。虽然有人可能会争辩说，这样将损害量化的性能，通常这不必定是正确的，如果这种处理是选择性地完成，例如当某些条件被满足时，若用新的量化器达到的压缩大大地高于遭受的失真。

举例来说，（图 2(b)），恒域量化仅仅影响另电平箱位，相对于第一个+/-箱位，然而所有其它的箱位仍旧保持相同，以及所用的重建值没有变化。一个带有自适应估计恒域的自适应方案也是可能的（例如，使用码率失真优化和自适应估计）。

根据本发明的某些示例性的实施，量化选择逻辑，因此，被提供在压缩/编码系统中，以基于至少一个特性或者参数在不同的恒域量化器（或者量化值）之间选择。举例来说，此逻辑可能按照某些图象特性在不同的恒域量化器 202、204、206 和 208 之间进行选择。

更加特别的是，此逻辑可能被配置成在图象数据中表征纹理结构和/或边缘作为代表“较重要信息”，并因而用一种式样对数据编码来提供一种较高水平的质量。此逻辑能用常规的纹理分析/检测及边缘检测定算法来支持这种判决过程。

有了这个概念之后，现在参见图 3，这是一张流程图，它说明了按照本发明的某些实现将不同的量化过程选择性地应用至数据的方法 300。在动作 302 中，初始数据组被分成数据的多个部分。例如，图象或者视频数据可能被分成多个块、宏块、条、片、区段等等。在动作 304 中，来自动作 302 的多个数据部分的至少一个以某些方式被分类，例如，基于至少一种特性或参数。例如，分类可以基于边缘特性、纹理特性、平滑特性、亮度特性、色度特性、色彩特性、噪声特性、对象特性、动作特性、用户喜好特性、用户界面聚焦特性、分层特性、定时特性、音量特性、频率特性、音调特性、音色特性、质量特性、比特率特性、数据类型特性、分辨率特性、加密特性。

在动作 306 中，用多个量化过程的至少两个来处理来自动作 304 的已分类部

分，以产生相应经量化的数据。在动作 308 中，作出决定选择来自动作 306 的一个经量化的数据，例如，基于与动作 304 中使用的分类参数相关的至少一个阈值或测量。在动作 310 中，来自动作 308 的经量化的数据以某些允许其随后被解码的方式而被编码。

使用了方法 300 之后，可以提供一种逻辑，实质上分析了图像或视频数据的部分并将某些部分视为比其它部分重要。不同的恒域量化器被用于较重要的数据部分，并且分析所产生的经量化的数据来确定哪一种恒域量化器满足该较重要的数据的期望阈值要求。例如，可以强制执行一种质量或噪声阈值要求。

通过进一步的示例，在视频中，序列边缘通常在运动补偿技术中起到重要的作用。例如，这通过下面详述的图 7 中的框图得到说明。基本上，可以对图像（帧）或其部分进行图像分析，并且根据分析来决定多个恒域量化器中哪一个最适用，例如，按照码率失真优化（RDO）准则。

参考图 4，逻辑 400 说明了按照本发明某示例性实现的快速判决过程。这里，输入帧/图像（或其中的部分）在块 402 中受到分析。在该例中，图像分析包括边缘检测和/或纹理分析能力，且输出被提供给恒域判决块 404。恒域判决块 404 然后使编码器 406 在对输入帧/图像（或其中的部分）编码时使用指定的或另外选择的恒域量化器或量化值。

图 5 中的逻辑 500 可配置为用于支持方法 300，而不是作出关于恒域量化的快速判决，其中使用了多个恒域量化器并且根据某阈值标准来作出 RDO 判决。这里，输入宏块（MB）/图像 502（或其它类似部分）被（选择性地）提供给不同的恒域编码块 504、506、508 和/或 510，来自这些不同恒域编码块的输出在 RDO 判决块 512 中被分析，选定的已编码数据被输出。如该例中所述，某些或全部恒域编码块/过程可能并行发生。在其它实现中，这种过程可以被安排成串行发生。

图 6（a）和图 6（b）说明其中使用了选择性重编码判决过程的逻辑。例如在图 6（a）中，输入 MB/图像 502 被提供给恒域编码块 602 和 604。重编码判决块 606 考虑来自恒域编码块 602 的输出并且影响在恒域编码块 602 和 604 的输出间的选择 608。重编码判决块 606 也可以选择性地启动恒域编码块 604。在图 6（b）中，逻辑 610 类似于 600，除了与重编码判决块 606 相关的选择 608 包括 RDO 判决块 612，用于分析来自恒域编码块 602 和 604 的输出并选择输出哪一个。

图 6（a，b）中的重编码判决块可以用于根据各种标准作出判决。例如，在某些实现中，可以考虑质量限制、速率限制和/或其它 RDO 概念/阈值。

只有当满足了某些前面定义的条件，譬如速率/质量超出特定阈值，才可能使用附加的恒域量化器。在某些示例性实验中，成功地选择了恒域量化器，它比原来具有 30% 的较大恒域。另外，可能按照诸如 AC 频率或边缘类型这样的图像特性，也可以使用自适应恒域量化器。

在图 7 中，示例性编码逻辑 700 说明性地阐述了怎样按照图像分析过程 704 来处理初始图像 702（或其一部分）以产生该例中的边缘检测数据 706。然后，边缘检测数据 706 被提供给块分类过程 708 以产生块分类的数据 710。然后，块分类的数据 710 与初始图像 702 一起被提供给图像编码过程 712，该过程接着产生编码图像 714。这是图像的简单编码过程的一个示例。这里，图像被分析（如，用边缘检测算法），然后，块按照这种信息而被分类。基本上定义了 N 维矩阵 (DZ_p) (N 取决于所定义恒域的数量)，该矩阵后来为在宏块的位置 (i, j) 处分配的量化器 (QP) 分配适当的恒域 (DZ) 。

无残留模式的考虑（如，对于 B 帧直接模式）：

即使上述说明可能主要针对平稳图像，然而也可以将同一概念应用在视频上。例如，也可以在视频压缩中以相当直接的方式使用图像预分析和恒域量化器。

一种相当有趣且可以大大受益于上述概念的情况是 B 帧内直接模式的使用。直接模式基本上是不需要传输运动参数的特殊方式，这是因为这可以直接或通过空间或通过时间上的预测而被直接预测到。关于直接模式的其它信息可以在美国未决临时专利申请序列号为 60/385,965 的申请中找到。

如果没有要被发送的剩余物，则可以按照本发明的某些进一步方面，通过使用被称为无残留直接模式的特殊方式进一步改进效率。如下所述，无残留直接模式可以利用游程长度编码 (RLC) 策略。这里，例如，如果产生的失真足够小且由于 RLC 的较高效率引起的比特率的减少足够大，则恒域量化器可以提供期望的解决方案。基本想法是在某些条件下选择性地引起或强制使直接模式无残留地被编码，尽管残留存在。例如，如图 9 所示，可以实现基于与图 5 相同的概念的方案，其中无残留直接模式与所有其它有效模式相比，也在 RDO 过程中被检查。在某些情况下，由于所使用的 RDO 对于这些情况并不适当，因此这种方案的性能可能不如期望中那么好。其它依赖于量化值的过程也被影响，譬如在用于消除分块人造假像的内环滤波器（未示出）。更明确地说，即使性能在较低的比特率显示良好，然而性能在较大比特率处可能大大受损；通过使用较大的量化器并且不考虑无残留直接模式，性

能可能变得更好。

与图像实例中所作的相类似，按照本发明的某些方面，该逻辑可被配置成仅在满足某些前面定义的条件时、特别在与直接模式相关的残留不重要时，才考虑这些方式。该逻辑可以，例如，被配置成通过检查直接模式的已编码块模式（CBP）而将对该残留重要性的估计用作一个条件。如果不考虑色度信息时 CBP 低于特定阈值，则可以考虑这个，意味着该残留并不重要，并且如果跳过它不会造成太大的失真。而且，诸如边缘和纹理的不存在这样的其它图像特性也可以用于这种视频编码过程/逻辑中。

在某些实现中，通过检查是否仅消除色度残留而保持所有的亮度残留完整无缺，该逻辑甚至能将此进一步扩展。也可以将这种概念扩展到宏块的所有可能方式，即检查是否完全不发出残留，或者该方式的色度不会给出较好的性能。然而，这显然会使编码复杂度进一步增加。

参考图 9，逻辑 900 说明输入 MB/图像 902 如何将编码过程块 902、904、906 和 908 提供给不同的直接模式（B 帧）或拷贝模式（中间帧），来自那些过程块的输出被提供给选择适当输出的基于帧的 RDO 判决块 910。

参考图 10，逻辑 1000 还说明怎样可以还包括方案判决块 1002 和选择 1004，从而根据用户输入、应用程序、系统要求等来提供附加选择性。这里，方案判决块 1002 选择性地将输入 MB/图像提供给一个或多个直接模式（B 帧）或拷贝模式（中间帧）编码过程块 902、904、906 和 908。选择 1004 可以通过方案判决块 1002 或其它过程/逻辑/输入来控制。

B 帧的拉格朗日（Lagrangian）乘数的示例性使用

RDO 技术通常基于拉格朗日乘数的概念。例如，可以选择共同最小化失真和比特率的特殊方式。

这种函数可以表示为下式的最小化：

$$J(\text{Mode}|\text{QP}, \lambda) = \text{SSD}(\text{Mode}|\text{QP}) + \lambda \cdot R(\text{Mode}|\text{QP})$$

其中 QP 是宏块量化器， λ 是用于模式判决的拉格朗日乘数，Mode 标识要被检查并且可能在 RDO 过程中被选择的宏块模式。

通过示例，在某些实现中，拉格朗日乘数 λ 可以为中间或内部帧而被选择：

$$\lambda_{I,P} = 0.85 \times 2^{\frac{\text{QP}}{3}}$$

或

$$\lambda_{I,P} = 5 \times \frac{QP+5}{34-QP} \times \exp^{\frac{QP}{10}}$$

然而在 B 帧中,在诸如 JVT 这样的大多数编解码器中,这被选择为 $\lambda_B = 4 \times \lambda_{I,P}$ 。

为了给较低开销模式以优先权而作出 λ 的附加加权,这是因为模式特别对于 B 帧可能具有由已发送的多个运动信息引起的大开销,而诸如直接模式这样的较低开销模式仍能提供就 RDO 而言的很好性能。

然而,根据按照本发明的某些实验,已经发现加权不应如上述那样不变,相反,它同样应该取决于量化器 QP 值。

特别是,如果定义了 $\lambda_B = f(QP) \times \lambda_{I,P}$,从这些实验中可以发现,可以使用比使用固定的 $f(QP) = 4$ 的情况具有好得多的压缩效率的两个 $f(QP)$ 函数:

$$f(QP) = \max\left(2, \min\left(4, \frac{QP}{6}\right)\right)$$

以及

$$f(QP) = \max\left(2, \min\left(4, \frac{QP+2}{8}\right)\right)$$

这种观察也可以通过具有很高的 λ 的事实而得到,这一事实也可能影响其它重要信息的精确性,譬如,运动预测的精确度,由此会对周围宏块的编码产生负面影响。

使用预测从属的 MB 模式的自适应类型选择:

序列和帧可能具有不同类型的主导运动相关,这也在未决美国临时专利申请序列号为 60/385,965、未决美国临时专利申请序列号为 60/376,005、以及未决美国临时专利申请序列号为 10/186,284 的申请中得以说明。

特别是,对于具有静止背景下的恒定速度的小物体而言,使用来自暂时相邻帧(时域)的运动矢量(MV)使人们能进行较好的预测,并且得到较高的性能。相反,具有平滑运动的较大物体可能在空间域上(相邻宏块)具有较大的相关,而在其它情况下,来自空间域和时域两者的信息可能对预测是重要的。这些类型的相关被部分利用,譬如,在 B 和 P 帧内的直接模式内以及 P 帧内关于运动矢量预测器的跳越。为了得到更多关于 P 帧内的运动矢量预测器的跳越信息,可以参见 Jani Lainema 和 Marta Karczewicz 撰写的“Skip mode motion compensation”,文档 JVT-C027, JVT 会议, Fairfax, 2002 年 5 月。这与,如果逻辑在某种程度上用信号指示哪种类型的预测器在不同帧起支配作用,则可以实现相当高的性能。

因此,按照本发明的某些方面,编码逻辑被配置成用信号指示或者在帧、片、或其它类似层次上以某些方式标识出要为从属于预测的模式使用哪一种预测方案。图 8 的图表中示出实现这个的一个示例性语法,它作为 JVT 中的一个示例。这种语法当然能够被修改并且在其它编码设计中不同。

对于图 800 中的示例性体系,P 帧跳跃模式允许三种(3)可能的预测情况。在该例中,存在运动-拷贝预测模式、时间预测跳跃模式、以及零跳跃模式。三种情况的每一个都被分配到{0,1,2}中的一个值,它或者可用固定大小的码字来编码,在该例中 $u(2)=2$ 比特,或者甚至被熵编码(如,用 UVLC 或 CABAC 的 $e(v)$)。本领域的技术人员可以理解,也可以使用其它方式来传送这种信息。在图表 800 中,它说明了 JVT 内的图像层 RBSP 语法,外加 B 帧的自适应空间/空间-时间考虑(`direct_mv_spatial`)以及中间帧的自适应拷贝/运动-拷贝跳跃模式(`copy_mv_spatial`)。如果仅使用了这两种模式,则所定义的描述符可能仅占据 1 比特,因此 $u(n=1)$,但如果要使用更多情况(带有零偏的空间预测或平稳时间预测的考虑),则可以分配更多比特($n>1$)或者甚至使用由 $e(v)$ 为该参数定义的熵编码。

对于 B 帧而言,在该例中,可以为所有的直接模式运动参数将空间运动矢量预测用作一个模式,而将时间预测的参数用作第二模式。也可以考虑/包括未决美国临时专利申请序列号为 60/385,965 的申请中所描述的其它直接模式。编码逻辑用信号通知在帧或片层次上要使用哪种预测模式。例如,选择可以用基于 RDO 的方案(如,图 9)来执行。在某些实现中,编码逻辑也可以使用明显地由放置在编码器和/或解码器上的特定要求引起的特定模式。在一个特定实例中,考虑到空间预测通常在计算上较简单(如,不需要分区、不需要存储运动矢量,并且独立于定时信息),因此它是某些应用的较佳选择(如,图 10)。

在其它实现中,其中这些问题不是问题,该组合可能导致进一步改进的编码性能。图 11 的示意图中示出已编码序列的一个实例。这里,P 和 B 帧与场景变化一起示出。为相应的跳跃(Skip)或直接(Direct)模式宏块示出不同信号表示的 P 和 B 帧的箭头。还注意到,信号表示可以指示编码逻辑应怎样执行用于运动矢量编码的运动矢量预测,或者其它模式的预测(例如,未决美国临时专利申请序列号为 60/376005 和未决美国专利申请序列号为 10/186284 的申请中描述的直接(Direct)P)。

如图 11 所示,不同的帧为其相应的直接(B)和跳跃(P)模式用信号表示不同类型的预测。 P_2 、 P_T 、和 P_M 为实例零定义了时间和空间的(运动-拷贝)预测,

B_T 和 B_{SP} 为直接模式定义了时间和空间的预测。

时间标记独立的直接模式，考虑到静态时间/空间块：

不同类型的预测，尤其对应 B 帧内的直接模式而言，可能更适用于不同类型的运动和序列。仅使用时间或空间预测在某些情况下可能提供可接受的性能，但在其它情况下，性能可能非常差。上述段落中描述的解决方式，或者未决美国临时专利申请序列号为 60/385965 中给出的解决方式可能提供更好的性能。

通过示例，给出了一种额外的情况，它看上去很有效并且组合了时间和空间预测符两者的性能，而简单地通过不要求分割而简单地保持空间预测符，并且/或者是时间独立的。

在特定实现中，空间预测可能比时间预测更有用（如，由于其属性）。因此，例如，空间预测被用作直接模式的主要预测。当来自时间预测符的运动信息和参考帧为零时，产生一种可能的例外。在这种情况下，相应直接模式块的运动信息和参考帧也被视作为零。而且，空间预测通过也考虑到空间零的偏置性和/或平稳子分区的空间零而被改进。由此，如果当前预测块的任何或某些相邻宏块或块具有零运动（或非常接近（例如，整数运动矢量为零））和参考帧，那么整个宏块，或其一部分也被视作具有零运动。这两个概念都有助于保护平稳背景，如果不引入这种条件，它尤其在运动对象的边缘引起相当的失真。

图 12 中示出流程图 1200，它被简化为 16×16 宏块的情况。注意到这种为直接模式使用空间预测的概念也可被扩展到更小的块（例如， 8×8 或 4×4 ）（或较大的块或其它形状的部分）。在动作 1202 中，空间预测符 MV_a 、 MV_b 和 MV_c 以及时间预测符 MV_t 被提供给动作 1204，其中 MV_{Direct} 被设为 $Median(MV_a, MV_b, MV_c)$ （取中值）。在动作 1206 中，基于 MV_t 作出判决，或导致动作 1208，其中 MV_{Direct} 被设为零，或导致附加判决的动作 1210。作为动作 1210 的结果， MV_{Direct} 或在动作 1208 中被设为零或保持不变， MV_{Direct} 被输出。

在该说明中，已经给出了与用于编码图像、视频序列或其它类型的数据的码率失真优化(Rate Distortion Optimization)有关的若干概念。已说明了附加语法，它可在视频序列内被采用，通过允许若干可选预测情况而允许相当高的压缩效率，尤其对于如分别在 P 和 B 帧内的跳跃或直接模式而言，它们可在图像开始时被信号通知到。已经给出了 B 帧的高效时间标记独立的直接模式，

它考虑到空间运动矢量预测以及平衡时间预测符。为了大大改进各种图像/视频/数据编码系统的性能，可以实现上述方法和装置的全部或一部分。

尽管本发明各种方法和装置的某些优选实现已经在附图中被说明并且在上述详细说明中被描述，然而可以理解，本发明并不限于这里所揭示的示例性实施例，而能作出许多重排、修改和替代，并且不背离本发明的精神。

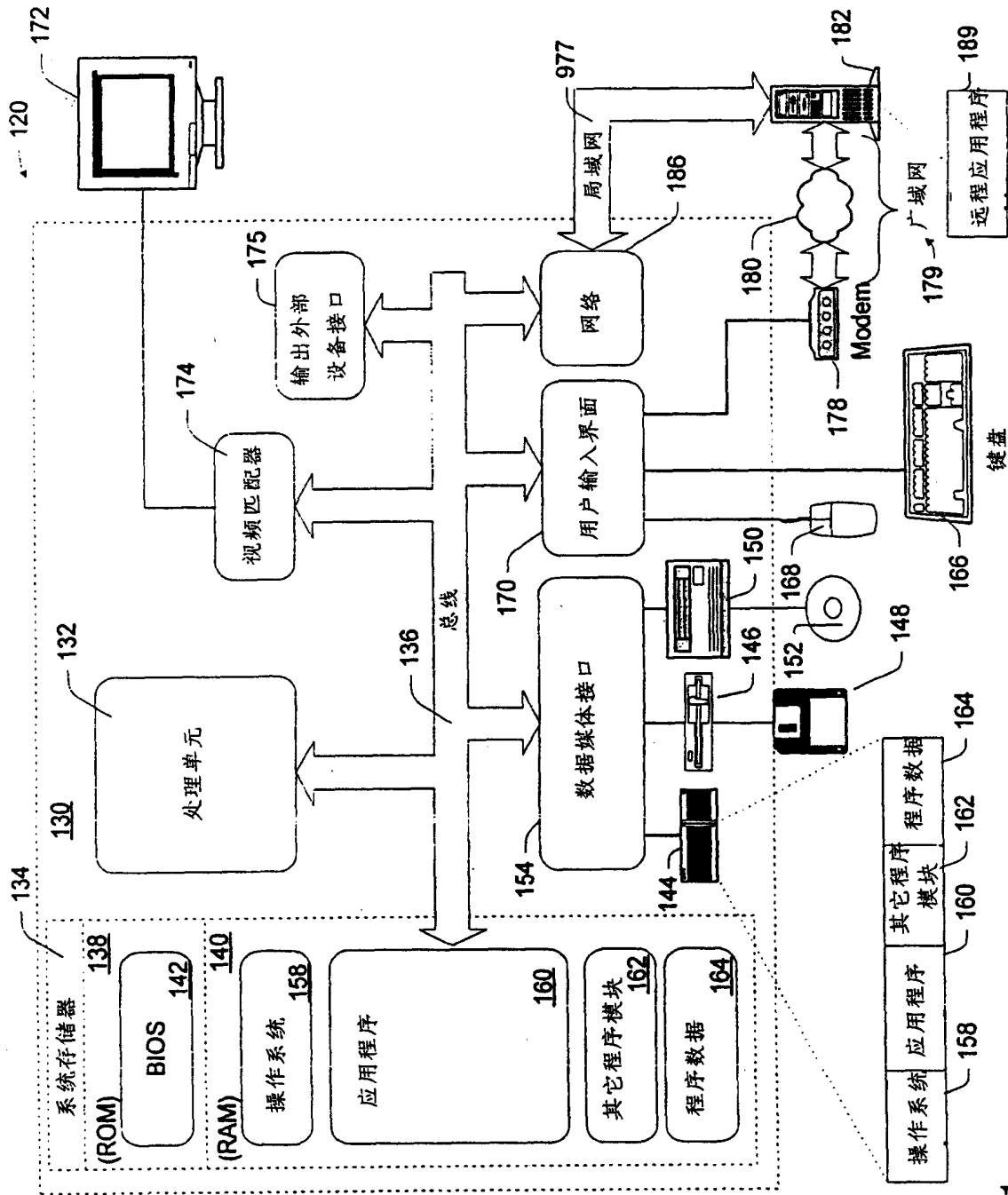


图 1

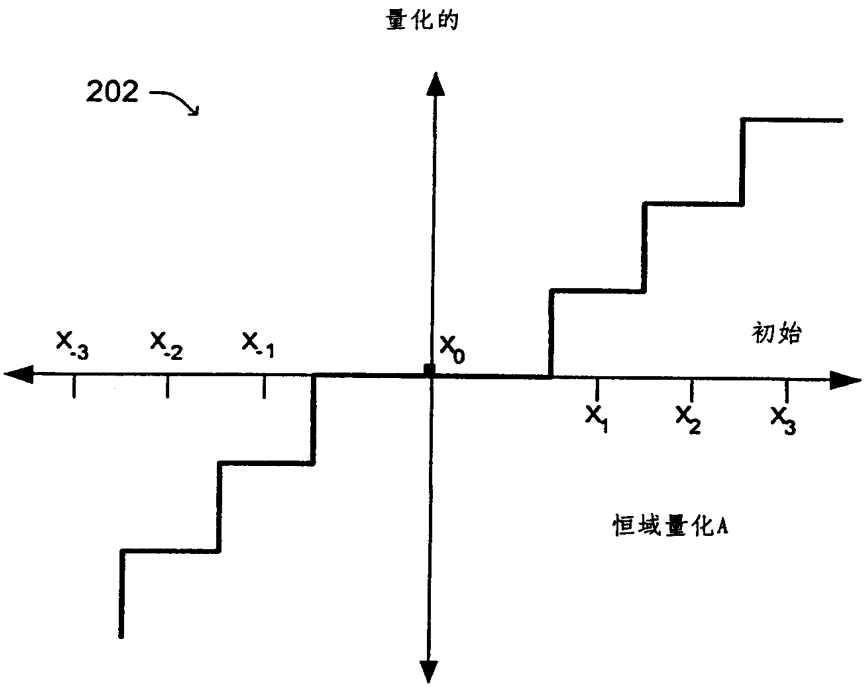


图 2(a)

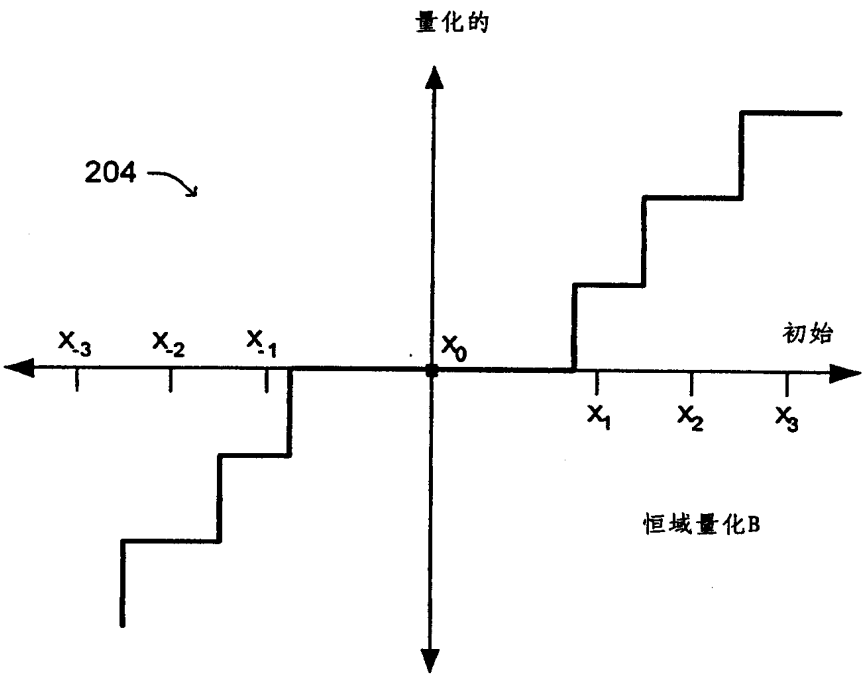


图 2(b)

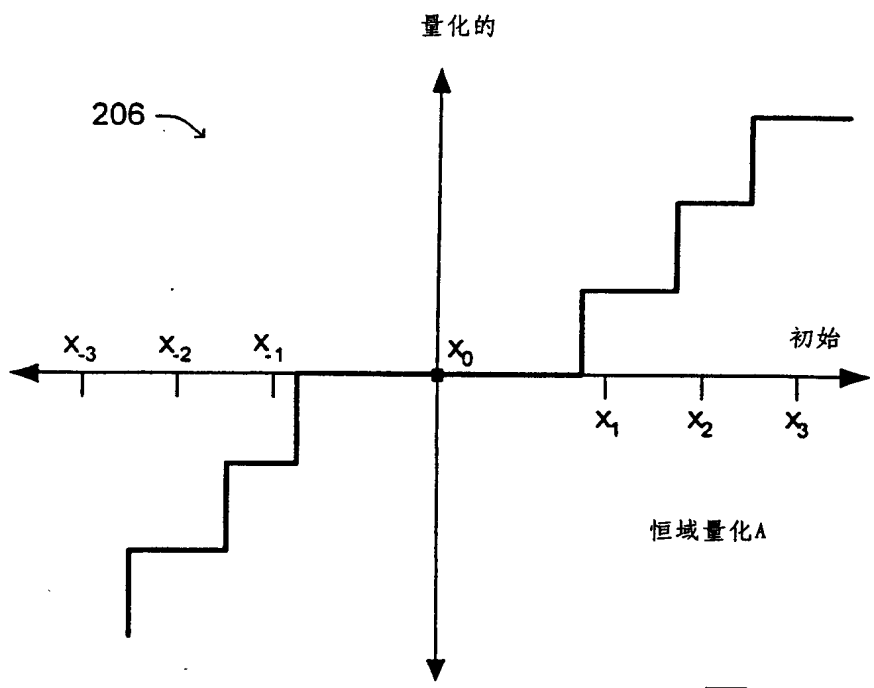


图 2(c)

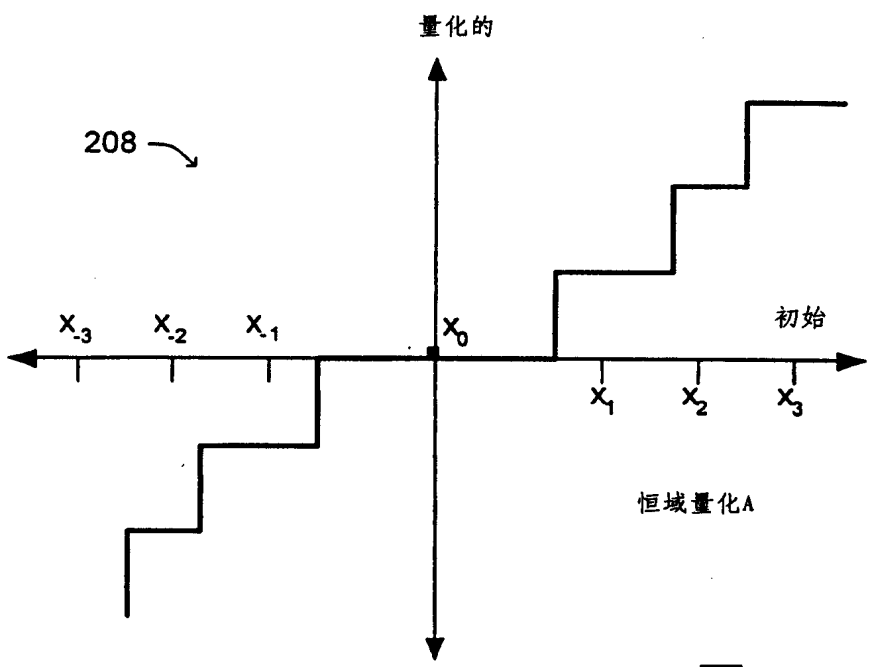


图 2(d)

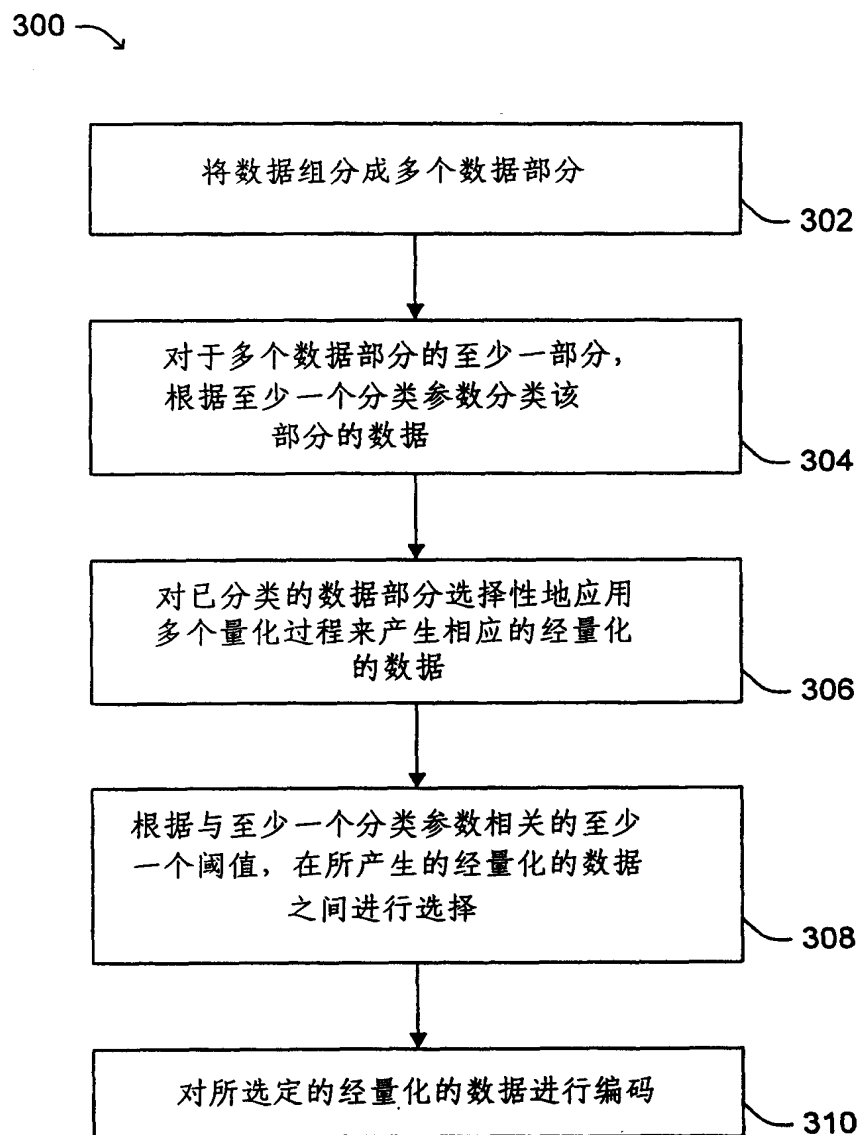


图 3

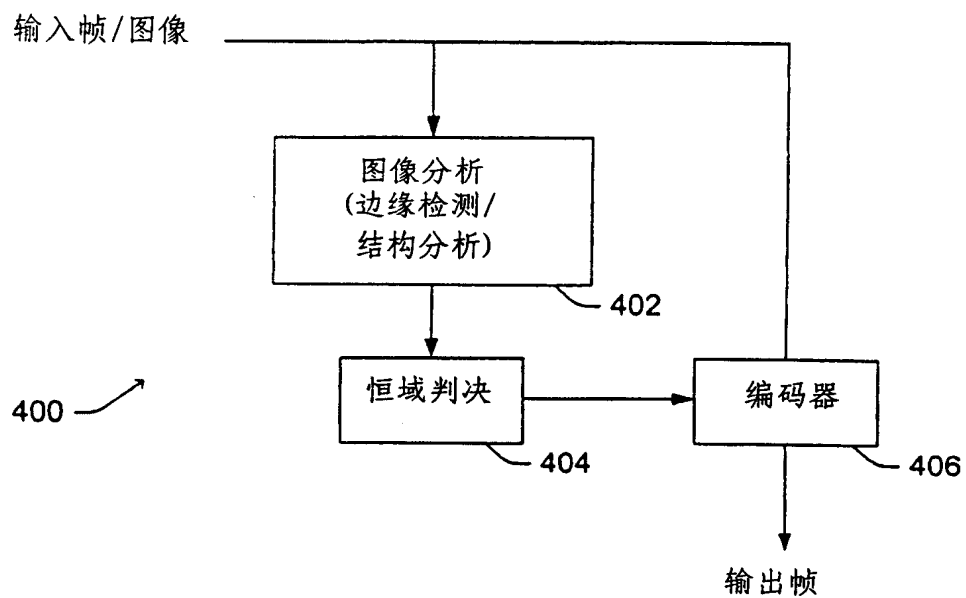


图 4

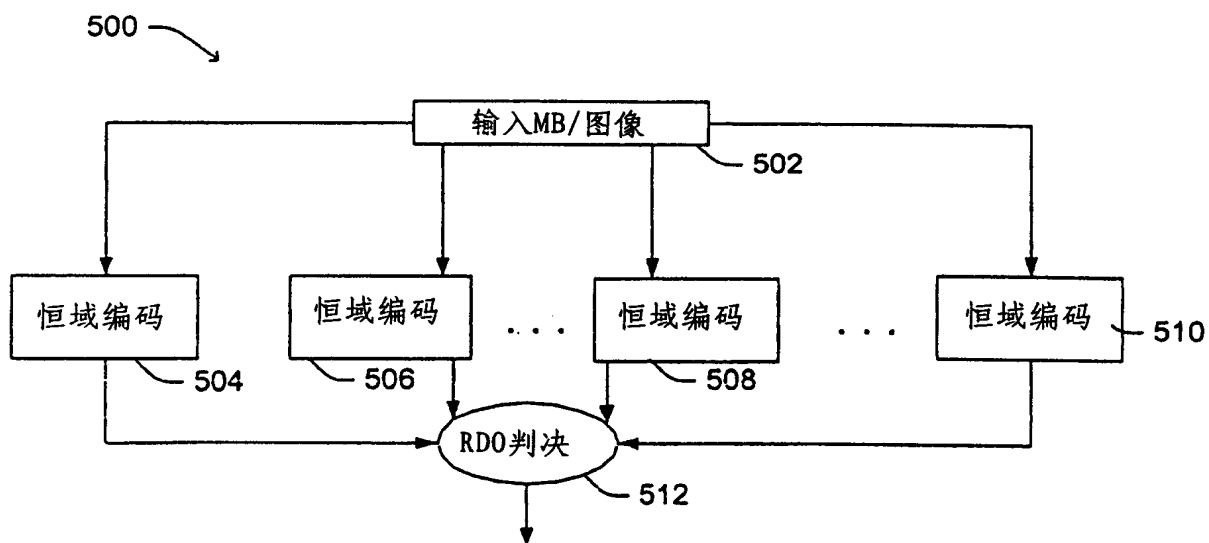


图 5

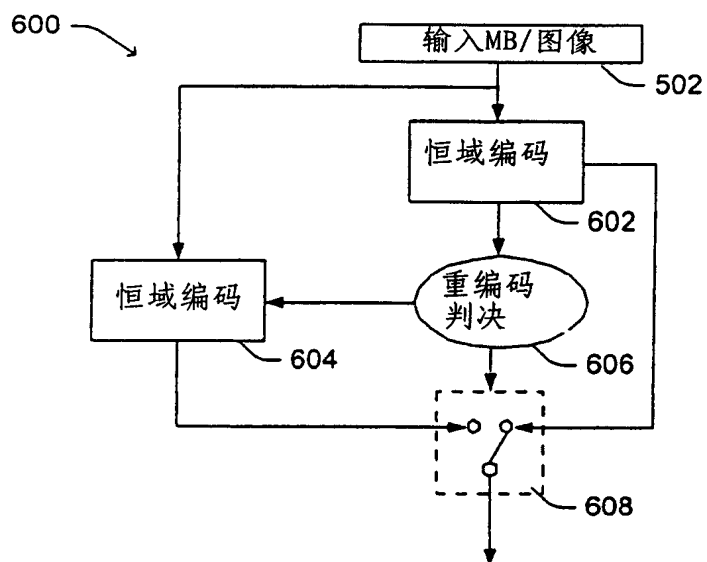


图 6(a)

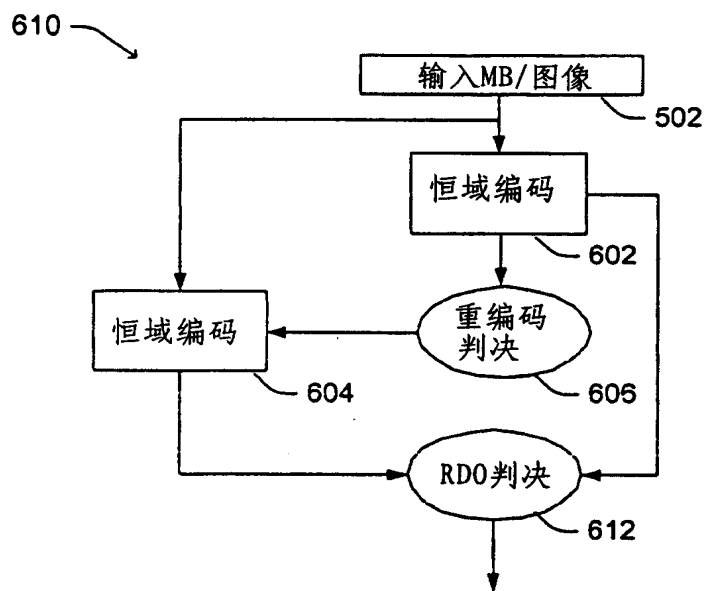


图 6(b)

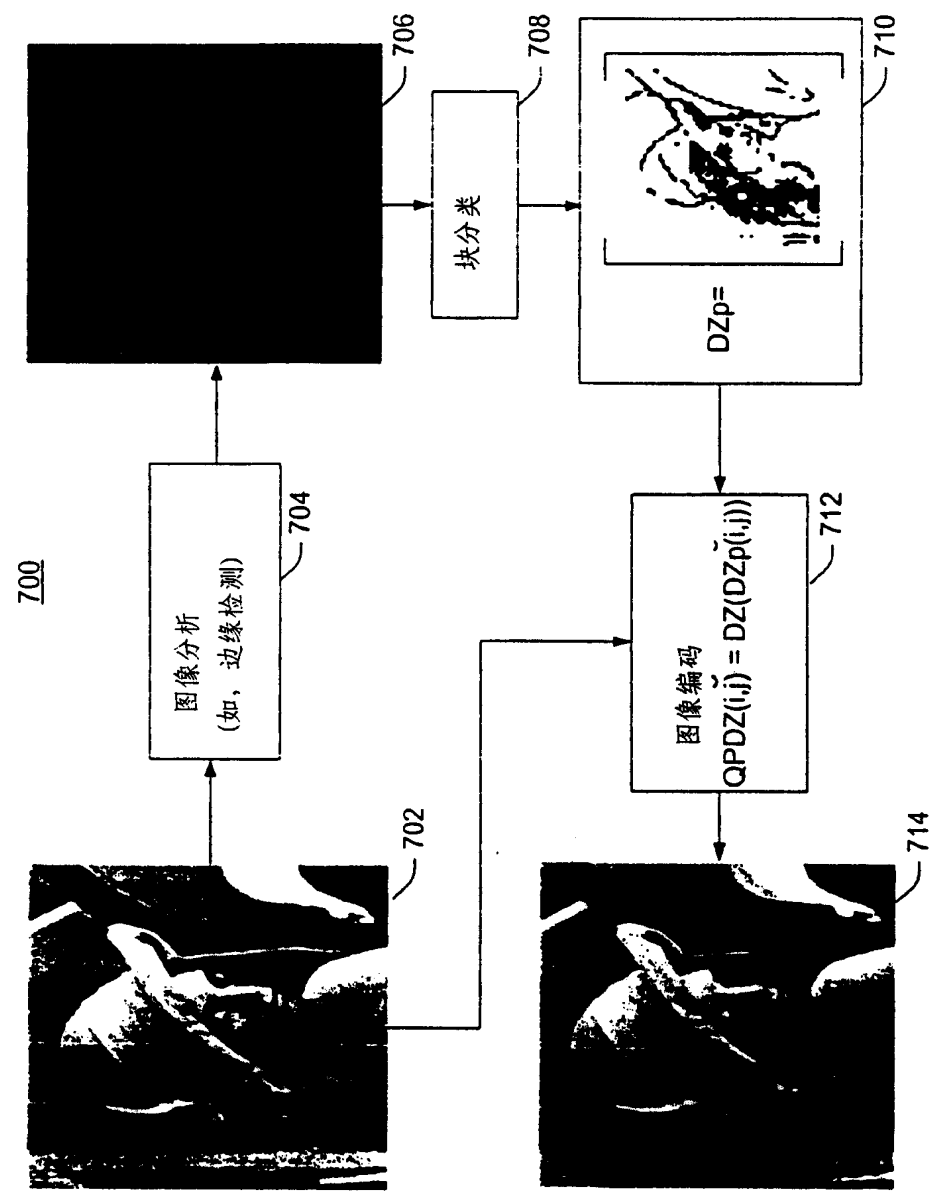


图 7

800 →

	类别	描述符
picture_layer_rbsp() {		
picture_structure	3	e(v)
frame_num	3	u(v)
rps_layer()		
if(coding_type() == Inter) {		
copy_mv_spatial		u(n) or e(v)
}		
if(coding_type() == Bipred) {		
direct_mv_spatial		
if(direct_mv_spatial) {		u(n) or e(v)
direct_mv_scale_fwd	3	e(v)
direct_mv_scale_bwd	3	e(v)
direct_mv_scale_divisor	3	e(v)
}		
explicit_B_prediction_block_weight_indication		e(v)
if(explicit_B_prediction_block_weight_indication > 1)		
adaptive_B_prediction_coeff_table()		
}		
rbsp_trailing_bits()		
}		

图 8

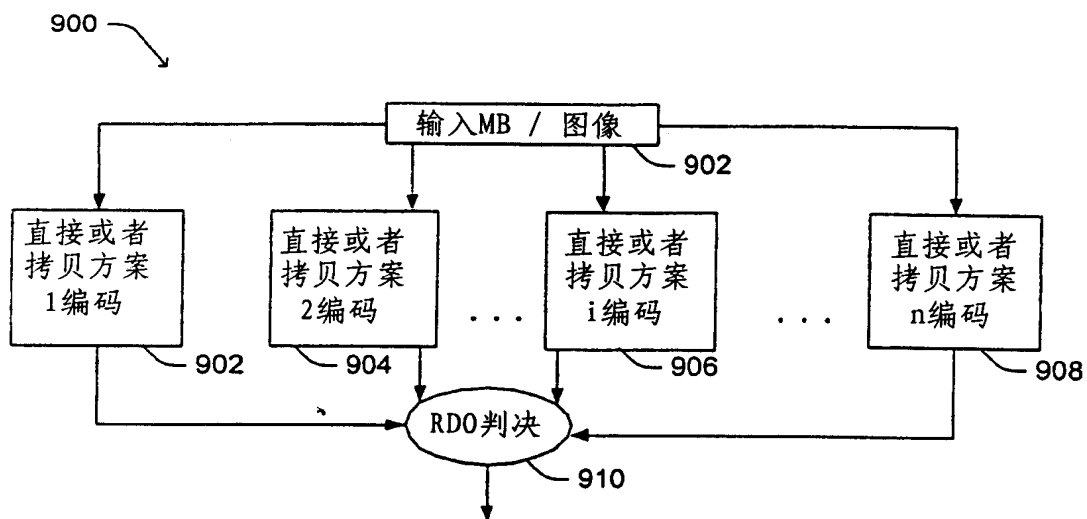


图 9

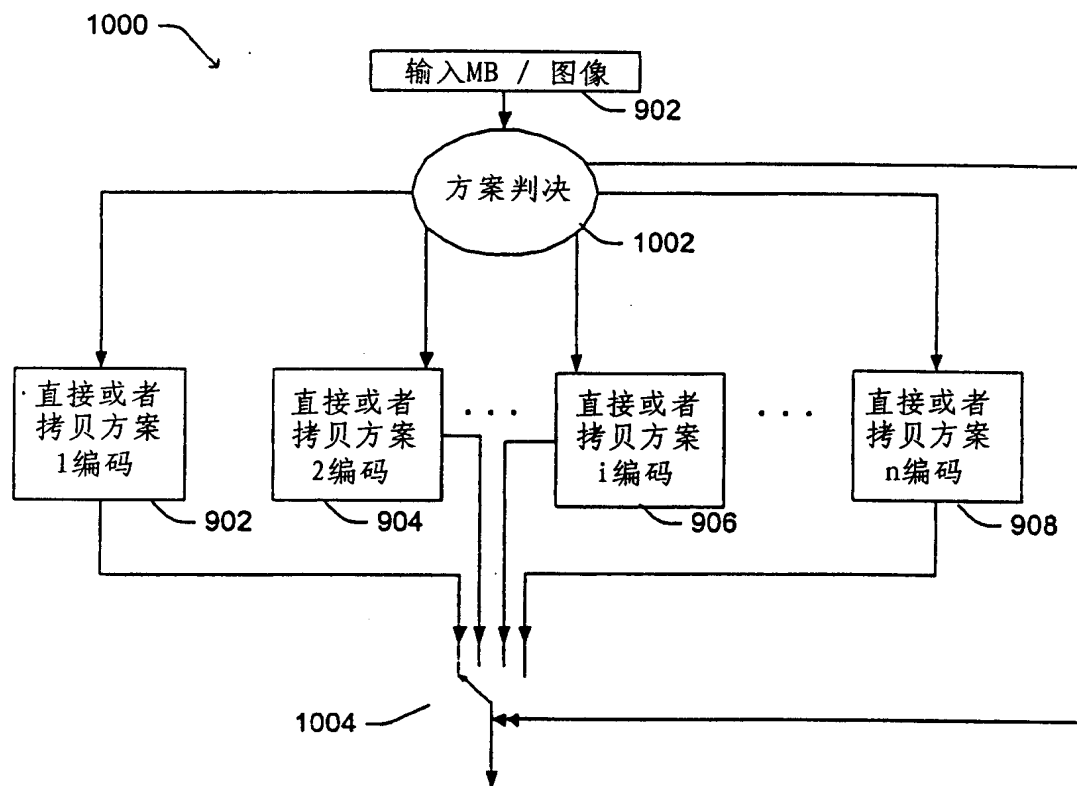


图 10

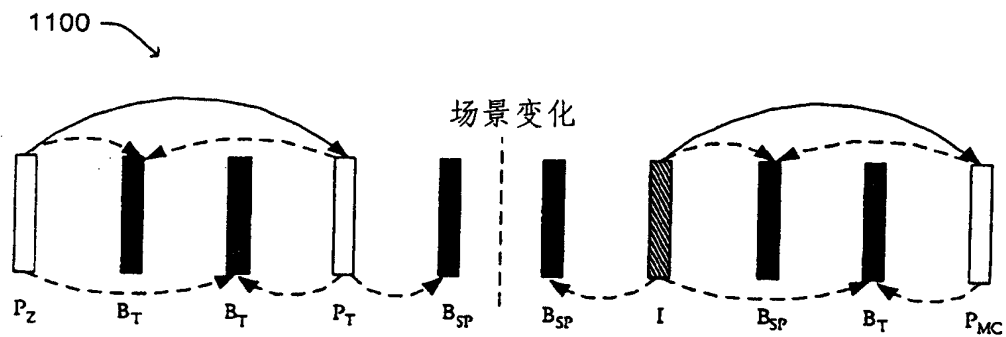


图 11

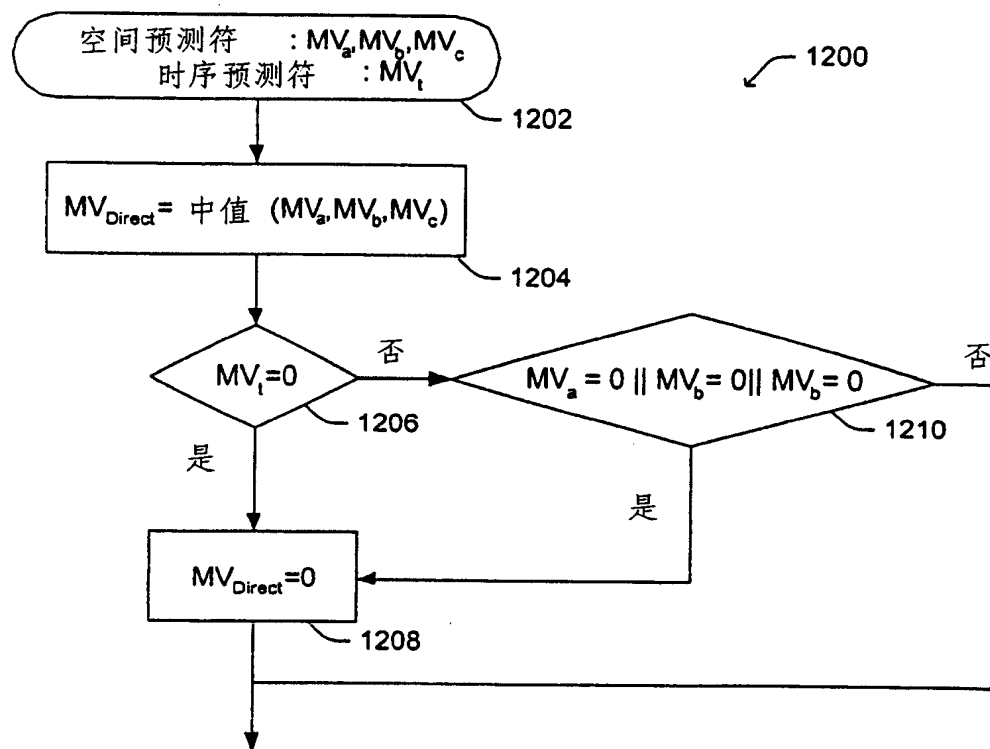
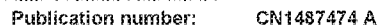


图 12



Title: IMPROVED OPTIMIZING TECHNOLOGY FOR DATA COMPRESSION

Title: 用于数据压缩的改进优化技术

Abstract:

[illegible]

Abstract:

提供了与诸如图像、视频等数据进行编码有关的方法和装置。例如,某些方法包括:用多个不同的量化功能来处理至少一部分数据以产生多个相应的经量化的数据部分,并且根据至少一个阈值选择性地将一个经量化的数据部分。该方法还包括:将初始数据分成多个部分,并且根据至少一个分类特性对数据的部分分类。这里,例如,可能有与分类特性相关的阈值。通过允许若干替代的运动预测情况,可以采用附加的语法来允许相当高的压缩效率。也提供了一种高效率时间标记独立的直接模式,它考虑到空间运动矢量预测以及稳态时间预测器。

A method and device for such as images, video, and other relevant data are encoded. For example, some method comprising: using a plurality of different quantization functions to process at least a portion of data to produce a plurality of corresponding quantized portions of data, and selectively outputting the data portion of a quantized according to at least one threshold. The method further comprises: the initial data into a plurality of portions, and classified in accordance with at least one characteristic of the data part of the classification. Here, for example, you may have the characteristics associated with the classification threshold. By allowing several alternative motion prediction case, you can use additional syntax to allow very high compression efficiency. Also provides a high efficiency timestamp independent direct mode, it takes into account spatial motion vector prediction and static time predictor.

Claims:

1. 一种方法, 其特征在于包括: 用多个不同的量化功能处理至少一部分数据以产生多个相应的经量化的数据部分; 以及 根据至少一个阈值选择性地输出所述经量化的数据部分之一。

2.如权利要求1所述的方法,其特征在于还包括:将初始数据分成多个部分,包括所述至少一部分数据;以及根据至少一种分类特性分类所述至少一部分数据。

3.如权利要求2所述的方法,其特征还在于,所述至少一个阈值与所述至少一种分类特性相关。

4.如权利要求2所述的方法,其特征在于,所述初始数据包括从包括图像数据、视频数据、音频数据和语音数据的一组中选择的数据。

5.如权利要求2所述的方法,其特征在于,所述初始数据包括图像数据或视频数据,所述至少一部分从包括块、宏块、条、片和区段的一组中选出。

6.如权利要求2所述的方法,其特征还在于,所述至少一种分类特性从一特性组中选出,该组特性包括:边缘特性、纹理特性、平滑特性、亮度特性、色度特性、色彩特性、噪声特性、对象特性、运动特性、用户喜好特性、用户界面聚焦特性、分层特性、定时特性、音量特性、频率特性、音调特性、音色特性、质量特性、比特率特性、数据类型特性、分辨率特性、以及加密特性。

7.如权利要求1所述的方法,其特征在于,所述多个不同的量化功能包括至少两个操作上不同的恒域量化器。

1 A method, comprising: processing at least a portion of the data with a plurality of different quantization functions to produce a plurality of corresponding quantized portions of data; and selectively according to at least one threshold value output by the one of quantitative data portion.

2 The method as claimed in claim 1 wherein, further characterized by comprising: an initial data into a plurality of parts, comprising at least a portion of the data; and at least a portion of at least one classification characteristic data according to the classification.

3. The method as claimed in claim 2, characterized in that said at least one threshold value and said at least one characteristic associated classification

4. The method as claimed in claim 2, wherein said initial data includes data selected from the group comprising image data, video data, audio data and a set of voice data.

5. The method as claimed in claim 2, wherein said initial data includes image data or video data from a group of at least a portion of the block is included in the macro block, strip, sheet and sections elected.

8.如权利要求7所述的方法，其特征在于，至少一个所述的恒域量化器包括一个自适应覆盖范围恒域量化器。

9.如权利要求7所述的方法，其特征在于，至少一个所述恒域量化器包括一个可变覆盖范围大小的恒域量化器。

10.如权利要求1所述的方法，其特征在于，选择性地输出所述一个经量化的数据部分还包括：对所述一个经量化的部分进行编码。

11.如权利要求1所述的方法，其特征在于，根据所述至少一个阈值选择性地输出所述一个经量化的数据部分还包括：执行码率失真优化（RDO）来选择所述一个经量化的数据部分。

12.一种方法，其特征在于包括：对至少一部分图像数据执行至少一个特性分析；根据所述至少一个特性分析在编码器内选择性地设置至少一个自适应量化参数；以及用所述编码器对所述至少一部分图像数据进行编码。

13.如权利要求12所述的方法，其特征在于，所述至少一个特性分析考虑到从一组特性中选择的至少一个图像分析特性，该组特性包括：边缘特性、纹理特性、平滑特性、亮度特性、色度特性、色彩特性、噪声特性、对象特性、运动特性、用户偏好特性、用户界面聚焦特性、分层特性、定时特性、音量特性、频率特性、音调特性、音色特性、质量特性、比特率特性、数据类型特性、分辨率特性、以及加密特性。

14.如权利要求12所述的方法，其特征在于，所述至少一个自适应量化参数与所述编码器内的自适应覆盖范围恒域量化器相关。

15.如权利要求12所述的方法，其特征在于，当对于包括视频帧的所述图像数据的多个部分的每一部分而言，在运行中根据所述至少一个特性分析在所述编码器内选择性地设置所述至少一个自适应量化参数。

16.一种方法，其特征在于包括：用至少两个不同的恒域量化器使至少一部分图像数据要被编码；以及根据与至少一个判决因子相关的码率失真优化（RDO）判决在所述至少两个不同恒域量化器之一的输出中识别较佳的已编码数据。

17.一种方法，其特征在于包括：用第一恒域量化器使至少一部分图像数据要被编码；确定所述第一恒域量化器的输出是否满足至少一个判决因子，以及如果是，则输出所述第一恒域量化器的所述输出，如果不是，则用不同于所述第一恒域量化器的至少一个第二恒域量化器使所述至少一部分图像数据要被编码。

18.如权利要求17所述的方法，其特征在于还包括：根据码率失真优化（RDO）判决识别所述图像数据的所述至少一部分的可接受已编码版本。

19.一种方法，其特征在于包括：对所述至少一部分图像数据进行图像分析；对所述图像数据的已分析部分进行块分类；对所述图像数据的经块分类的部分进行恒域量化；以及对所述图像数据的经恒域量化的部分进行编码。

20.如权利要求19所述的方法，其特征在于，所述图像分析包括从包括边缘检测分析和纹理分析的一组中选择的至少一类分析。

21.如权利要求19所述的方法，其特征在于，所述块分类根据所述图像分析操作上被配置，而所述恒域量化根据所述块分类操作上被配置。

22.一种方法，其特征在于包括：用至少两个不同的编码器使至少一部分视频图像数据要被编码，其中所述两个不同编码器的至少一个包括操作上配置为用于支持所述视频图像数据的无残留模式的恒域量化器；以及根据与至少一个判决因子相关的码率失真优化（RDO）判决在所述至少两个不同编码器之一的输出中识别较佳的已编码帧数据。

23.如权利要求22所述的方法，其特征在于，所述恒域量化器是自适应的并用于选择性地所述无残留模式被使用。24.如权利要求22所述的方法，其特征在于，至少一个所述编码器操作上配置为用于支持所述视频图像数据的直接模式。

25.如权利要求22所述的方法，其特征在于，至少一个所述编码器操作上配置为用于支持所述视频图像数据的拷贝模式。

26.如权利要求22所述的方法，其特征在于，所述无残留模式根据运行长度编码（RLC）策略操作上被配置。

27.如权利要求22所述的方法，其特征在于，所述恒域量化器是操作上自适应的。

28.如权利要求22所述的方法，其特征在于，所述恒域量化器根据与至少一部分视频图像数据相关的至少一个特性操作上被配置。

29.如权利要求22所述的方法，其特征在于，所述识别所述较佳已编码帧数据还包括：分析与所述直接模式帧数据相关的残留量。

30.如权利要求29所述的方法，其特征在于，分析所述残留量还包括检查所述直接模式帧数据的已编码块模式（CBP）。

6 The method as claimed in claim 2, characterized in that said at least one characteristic selected from the group classification characteristic, the set of features include: edge characteristics, texture characteristics, smoothness characteristics, luminance characteristics, color characteristics, color characteristics, noise characteristics, object properties, motion characteristics, user preferences features, user interface focusing characteristic, hierarchical nature, timing characteristics, volume, frequency characteristics, tonal characteristics, sound characteristics, quality characteristics, bit rate characteristics data type characteristics, resolution characteristics, and encryption features.

7. The method as claimed in claim 1, wherein, wherein said plurality of different quantization functions include the different constant domain quantizer to control at least two operations.

8. The method as claimed in claim 7, wherein, wherein the at least one constant domain quantizer comprises an adaptive deadzone quantization coverage.

9. The method as claimed in claim 7, wherein, wherein the at least one constant domain of the quantizer comprises a variable coverage size deadzone quantizer.

10 The method as claimed in claim 1 wherein, wherein the selectively outputting one of the quantized data portion by further comprising: one of said encoded quantized portions.

. Executed to select the rate-distortion optimization (RDO):

11 The method as claimed in claim 1 wherein, wherein the at least one threshold value in accordance with said selectively outputting one of the quantized data portion further comprises Di said a quantitative data portion ECONOMY.

12 A method, comprising: performing at least one characteristic analysis on at least a portion of the image data; analysis in the encoder to selectively set at least one adaptive quantization parameter based on at least one property; and with the the encoder of said at least a portion of the image data is coded.

13 The method as claimed in claim 12 wherein, wherein the at least one characteristic analysis considers at least one image analysis characteristic selected from a group of properties, the set of features include: edge characteristics, texture characteristics, smooth characteristics, luminance characteristics, color characteristics, color characteristics, noise characteristics, object properties, motion characteristics, user preferences features, user interface focusing characteristic, hierarchical nature, timing characteristics, volume, frequency characteristics, tonal characteristics, sound character, quality characteristics, bit rate characteristics, data type properties, resolution characteristics, and encryption features.

14. The method as claimed in claim 12 wherein, wherein the at least one parameter of the adaptive quantization coder adaptive coverage within the deadzone associated quantizer.

15. The method as claimed in claim 12 wherein, wherein, when the image data for each portion of a video frame comprises a plurality of sections, in accordance with the operation of the analysis in at least one property selectively provided within said at least one of said encoder adaptive quantizer parameter.

16 A method, comprising: using at least two different quantization deadzone enables at least part of the image data to be encoded; and Distortion Optimization (RDO) judgment in accordance with at least one decision factor-related rate identifying preferred encoded data at the output of said at least two different constant domains of the quantizer.

17.- method, comprising: a first constant domain quantization allows at least a portion of the image data to be encoded; determining whether the first constant domain quantizer output meets at least one decision factor, and if Yes, the output of the first constant domain of the quantizer output, if not, use different from the first constant domain quantizer at least one second deadzone quantization allows at least a portion of the image data to be coding.

31. 如权利要求22所述的方法，其特征在于，带有所述恒域量化器的所述编码器选择性地配置为移去从包括色度残留数据、亮度残留数据和所有残留数据的一组残留数据中选择的至少一类残留。

32. 一种方法，其特征在于包括：选择性地改变至少一个拉格朗日乘法器，该乘法器操作配置了编码逻辑，编码逻辑带有基于至少一部分图像数据的至少一个特性的量化功能；以及用所述编码逻辑对所述图像数据的所述至少一部分进行编码。

33. 如权利要求32所述的方法，其特征在于，所述量化功能包括基于下式的宏块量化器： $J(\text{Mode} | \text{QP}, X) = \text{SSD}(\text{Mode} | \text{QP}) + X \cdot R(\text{Mode} | \text{QP})$ 其中X是用于模式判决的所述拉格朗日乘法器，而Mode表示用码率失真优化 (RDO) 过程检查到的宏块模式。

34. 如权利要求32所述的方法，其特征在于，所述拉格朗日乘法器为帧间或内部帧而被选择。

35. 如权利要求33所述的方法，其特征在于： $\lambda = f(\text{QP})X_i, P_i$ 以及 $f(\text{QP}) = \max(2, \min(4, \lambda))$ V 6 J 36. 如权利要求33所述的方法，其特征在于： $\lambda = f(\text{QP})XX_i p_i$ 以及 $f(\text{QP}) = \max(2, \min(4, \lambda))$ 37. 一种方法，其特征在于包括用编码器逻辑对至少一部分视频图像数据进行编码；以及使所述编码器逻辑输出语法信息，该语法信息识别由所述编码器逻辑使用的一类运动矢量预测。

38. 如权利要求37所述的方法，其特征在于，所述编码器逻辑用于选择性地使用从包括空间运动矢量预测和时间运动矢量预测在内的一组中选择的至少一类运动矢量预测，且其中所述编码器逻辑至少部分根据速率失真优化 (RDO) 判决选择性地输出所述语法信息和相应的已编码视频图像数据。

39. 一种用于传送与视频编码有关的信息的方法，该方法其特征在于包括：编码视频数据；以及在所述已编码视频数据的语法部分内选择性地设置至少一个描述符，所述描述符识别与用所述视频数据编码的至少一个B帧相关的自适应空间/空间-时间编码。

40. 如权利要求39所述的方法，其特征在于，编码所述视频数据包括按照 JVT 标准来编码所述视频数据，而所述已编码视频数据的所述语法部分内的所述至少一个描述符包括：带有所述已编码视频数据的图像层部分的 `direct_mv—spatial` 参数。

41. 一种用于传送与视频编码有关的信息的方法，该方法其特征在于包括：编码视频数据；以及在所述已编码视频数据的语法部分内选择性地设置至少一个描述符，所述描述符识别用所述视频数据编码的至少一个中间帧内的自适应拷贝/运动拷贝跳跃模式。

42. 如权利要求41所述的方法，其特征在于，编码所述视频数据包括按照 JVT 标准来编码所述视频数据，而所述已编码视频数据的所述语法部分内的所述至少一个描述符包括：带有所述已编码视频数据的图像层部分的 `copy_mv—spatial` 参数。

43. 一种用在对考虑视频帧的静态时间/空间部分的视频进行时间标记独立模式编码中的方法，该方法其特征在于包括：在视频序列中选择性地应用与至少一部分视频帧相关的运动的空间预测；以及如果另一视频帧的参考部分的时间运动预测信息为零，则将所述运动的空间预测设为零。

44. 如权利要求43所述的方法，其特征在于，所述空间预测是空间零偏置的。

45. 如权利要求43所述的方法，其特征在于，所述空间预测还考虑到所述视频帧中其它部分的静态子分区，且其中，如果至少一个所述子分区具有大约为零的相应运动，则所述其它部分也被视为具有零运动。

46. 一种带有计算机可实现的指令的计算机可读媒体，指令用于配置至少一个处理单元来执行下列动作：用多个不同的量化函数来处理至少一部分数据以产生多个相应的经量化的数据部分；以及根据至少一个阈值选择性地输出一个所述经量化的数据部分。

47. 如权利要求46所述的计算机可读媒体，其特征在于还包括：将初始数据分成多个部分，包括所述至少一部分数据；以及根据至少一个分类特性对所述至少一部分数据进行分类。

48. 如权利要求47所述的计算机可读媒体，其特征在于，所述至少一个阈值与所述至少一个分类特性相关。

49. 如权利要求47所述的计算机可读媒体，其特征在于，所述初始数据包括从包括图像数据、视频数据、音频数据和语音数据的一组数据中选择的的数据，且其中所述初始数据包括图像数据或视频数据，且所述至少一部分从包括块、宏块、条、片和区段的一组中选出。

50. 如权利要求47所述的计算机可读媒体，其特征在于，所述至少一个分类特性从一组特性中选出，该组特性包括：边缘特性、纹理特性、平滑特性、亮度特性、色度特性、色彩特性、噪声特性、对象特性、运动特性、用户喜好特性、用户界面聚焦特性、分层特性、定时特性、音量特性、频

18 The method as claimed in claim 17, characterized by further comprising: based on rate-distortion optimization (RDO) decision identifying the image data of the at least a portion of an acceptable encoded version.

19 A method, comprising: at least a portion of the image data for image analysis; the image data analysis section of the block has been classified; for by the block classification part of the image data is carried out constant domain quantization; and by constant domain quantization portion of the image data is encoded.

20. The method as claimed in claim 19, wherein the image analysis includes an edge detector comprises from texture analysis and analysis of a set of at least one type selected analysis.

21. The method as claimed in claim 19, wherein said block classification is arranged according to the operation of the image analysis, and the constant domain is configured in accordance with the quantization of the block classification operation.

22 A method, comprising: using at least two different coding allows at least a portion of the video image data to be encoded, wherein said at least two different encoders includes operatively configured for support the video image data without the residual pattern deadzone quantizer; and at least one decision factor in accordance with the rate-distortion optimization related (RDO) decision output of said at least one of the two different encoders identifying preferred the already coded frame data.

23. The method as claimed in claim 22, wherein the constant domain is an adaptive quantizer for selectively and no residue of the used pattern 24. claimed in claim 22 method characterized in that at least one of said encoder is configured for direct mode video image data to support the operation.

25. The method as claimed in claim 22, characterized in that at least one of said encoder is configured for supporting the copy mode to video image data on the operation.

26. The method as claimed in claim 22, wherein said pattern is arranged according to no residue on the run length coding (RLC) operation policy.

27. The method as claimed in claim 22, wherein said constant domain quantizer is adaptive in operation.

28. The method as claimed in claim 22, wherein said constant domain quantizer according to at least one operating characteristic of at least a portion of the video image and associated data are arranged.

29 The method as claimed in claim 22, wherein said identifying said preferred encoded frame data further comprises: analyzing the residual amount of data related to the direct mode frame.

30. The method as claimed in claim 29, wherein analyzing further comprises checking the residual amount of said direct mode have data frames coded block pattern (CBP).

31. The method as claimed in claim 22, characterized in that, with the constant domain of the encoder quantizer is configured to be selectively removed from the residual data comprising chrominance residue data, luminance, and all a set of at least one type of residual residual data of the selected data.

32 - method, comprising: selectively varying at least one Lagrangian multiplier, which multiplies the coded logic operatively configured, with the at least one characteristic of the encoding logic based on at least a portion of the image data the quantization function; and the coded logic for the use of said at least part of the image data to be encoded.

33 The method as claimed in claim 32, wherein said quantizing function includes a macroblock based on the following formula Quantizer: $J(\text{Mode} | \text{QP}, X) = \text{SSD}(\text{Mode} | \text{QP}) X R(\text{Mode} | \text{QP})$ wherein X is a mode decision of the Lagrange multiplier, while mode is represented by rate-distortion optimization (RDO) procedure to check the macro block mode.

率特性、音调特性、音色特性、质量特性、比特率特性、数据类型特性、分辨率特性、以及加密特性。

51. 如权利要求46所述的计算机可读媒体, 其特征在于, 所述多个不同的量化功能包括至少两个操作不同的恒域量化器, 且其中至少一个所述恒域量化器包括一个自适应覆盖范围的恒域量化器。

52. 如权利要求46所述的计算机可读媒体, 其特征在于, 根据所述至少一个阈值选择性地输出所述一个经量化的数据部分还包括: 执行码率失真优化(RDO)以选择所述一个经量化的数据部分。

53. 一种带有计算机可实现的指令的计算机可读媒体, 指令用于配置至少一个处理单元来执行下列动作: 对至少一部分图像数据执行至少一个特性分析; 根据所述至少一个特性分析在编码器内选择性地设置至少一个自适应量化参数; 以及用所述编码器对所述至少一部分图像数据进行编码。

54. 如权利要求53所述的计算机可读媒体, 其特征在于, 所述至少一个自适应量化参数与所述编码器内的自适应覆盖范围的恒域量化器相关联。

55. 如权利要求53所述的计算机可读媒体, 其特征在于, 对于包括视频帧的所述图像数据的多个部分的每一部分而言, 在运行根据所述至少一个特性分析在所述编码器内选择性地设置所述至少一个自适应量化参数。

56. 一种带有计算机可实现的指令的计算机可读媒体, 指令用于配置至少一个处理单元来执行下列动作: 用至少两个不同的恒域量化器使至少一部分图像数据要被编码; 以及根据与至少一个判决因子相关的码率失真优化(RDO)判决在所述至少两个不同恒域量化器之一的输出中识别最佳的已编码数据。

57. 一种带有计算机可实现的指令的计算机可读媒体, 所述指令用于配置至少一个处理单元来执行下列动作: 用第一恒域量化器使至少一部分图像数据要被编码; 确定所述第一恒域量化器的输出是否满足至少一个判决因子, 以及如果是, 则输出所述第一恒域量化器的所述输出, 如果不是, 则用不同于所述第一恒域量化器的至少一个第二恒域量化器使所述至少一部分图像数据要被编码。

58. 如权利要求57所述的计算机可读媒体, 其特征在于还包括: 根据码率失真优化(RDO)判决识别所述图像数据的所述至少一部分的可接受已编码方式。

59. 一种带有计算机可实现的指令的计算机可读媒体, 所述指令用于配置至少一个处理单元来执行下列动作: 对所述至少一部分图像数据进行图像分析; 对所述图像数据的已分析部分进行块分类; 对所述图像数据的经块分类的部分进行恒域量化; 以及对所述图像数据的经恒域量化的部分进行编码。

60. 如权利要求59所述的计算机可读媒体, 其特征在于, 所述图像分析包括从包括边缘检测分析和纹理分析的一组中选择的至少一类分析。

61. 如权利要求59所述的计算机可读媒体, 其特征在于, 所述块分类根据所述图像分析操作上被配置, 而所述恒域量化根据所述块分类操作上被配置。

62. 一种带有计算机可实现的指令的计算机可读媒体, 所述指令用于配置至少一个处理单元来执行下列动作: 用至少两个不同的编码器使至少一部分视频图像数据要被编码, 其中所述两个不同编码器的至少一个包括操作上配置成用于支持所述视频图像数据的无残留模式的恒域量化器; 以及根据与至少一个判决因子相关的码率失真优化(RDO)判决在所述至少两个不同编码器之一的输出中识别最佳的已编码帧数据。

63. 如权利要求62所述的计算机可读媒体, 其特征在于, 所述恒域量化器是自适应的并用于选择性地使用所述无残留模式。

64. 如权利要求62所述的计算机可读媒体, 其特征在于, 至少一个所述编码器操作上配置为用于支持所述视频图像数据的直接模式。

65. 如权利要求62所述的计算机可读媒体, 其特征在于, 至少一个所述编码器操作上配置为用于支持所述视频图像数据的拷贝模式。

66. 如权利要求62所述的计算机可读媒体, 其特征在于, 所述无残留模式根据运行长度编码(RLC)策略操作上被配置。

67. 如权利要求62所述的计算机可读媒体, 其特征在于, 带有所述恒域量化器的所述编码器选择性地配置为消除从包括色度残留数据、亮度残留数据和所有残留数据的一组残留数据中选择的至少一类残留。

68. 一种带有计算机可实现的指令的计算机可读媒体, 指令用于配置至少一个处理单元来执行下列动作: 选择性地改变至少一个拉格朗日乘法器, 该乘法器操作上配置了编码逻辑, 编码逻辑带有基于至少一部分图像数据的至少一个特性的量化功能; 以及用所述编码逻辑对所述图像数据的所述至少一部分进行编码。

69. 一种带有计算机可实现的指令的计算机可读媒体, 指令用于配置至少一个处理单元来执行下列动作: 用编码器逻辑对至少一部分视频图像数

34. The method as claimed in claim 32, wherein said Lagrangian multiplier is inter or intra-frame is selected.

35 The method as claimed in claim 33, wherein, wherein: the $\lambda = f(QP) X_{i,j}$, P ; and $f(QP) = \max\{2, \min\{4, \lambda\}\}$ 36. the method as claimed in claim 33, wherein: the $\lambda = f(QP) XX_{i,j}$; and $f(QP) = \max\{2, \min\{4, \lambda\}\}$ 37. - method, characterized by encoding package logic for at least a part of the video encoding image data; and an output of said encoder logic syntax information, one motion vector information identifies the syntax logic used by the encoder prediction.

38. The method as claimed in claim 37, wherein said encoder logic to selectively use from the at least one motion vector includes spatial prediction and temporal motion vector prediction selected from a group including the class prediction motion vector, and wherein said encoder logic at least in part on the rate-distortion optimization (RDO) decision selectively outputting the syntax information and corresponding encoded video image data.

39 A method of transmitting information related to video coding, the method comprising: encoding video data; and selectively provided within said encoded video data portion of the at least one descriptor grammar the descriptor identification and data encoded using the video related to at least one B-frame adaptive space / time coding a space.

40. The method of claim 39 and wherein said encoded video data has the syntax of the inner portion, characterized in that said video data includes encoded in accordance with the JVT video coding standard data, at least one descriptor comprising: with the already direct_mv-spatial parameters of the picture layer of the encoded video data section.

41 A method of transmitting information related to video coding, the method comprising: encoding video data; and selectively provided within said encoded video data portion of the at least one descriptor grammar, the descriptor identifying an adaptive copy of the video data encoded by the at least one intermediate frame / copy motion skip mode.

The syntax of the portion

42. The method as claimed in claim 41 wherein, wherein the encoded video data includes encoded according to the JVT standard video data, and the coded video data within at least one descriptor comprising: a copy_jv with the spatial parameters are part of the picture layer of the encoded video data.

43 for use in a video frame to consider static temporal / spatial portions of video time stamp independent mode encoding method, the method comprising: in a video sequence are selectively applied to at least a portion of the video spatial prediction frame related motion; and if the reference time of the motion of another part of the video frame prediction information is zero, then the spatial prediction of motion to zero.

44. The method as claimed in claim 43, wherein, wherein said spatial prediction is spatial zero biased.

45. The method as claimed in claim 43, wherein, wherein said spatial prediction further take into account other parts of the video frame static sub-partitions, and wherein, if at least one of said sub-partition has approximately corresponding movement of zero, then the other portion was regarded as having zero motion.

46 A computer implemented instructions with computer-readable medium, instructions for configuring at least one processing unit to perform the following actions: a plurality of different quantization functions to process at least a portion of data to produce a plurality of corresponding according to at least one threshold value and selectively outputting one of said quantized portions of data; part by the quantized data.

47 46 as claimed in claim wherein the computer readable medium, characterized by further comprising: an initial data into a plurality of parts, comprising at least a portion of the data; and the data of at least a portion of at least one classification characteristic classification.

据进行编码；以及使所述编码器逻辑输出语法信息，该语法信息识别由所述编码器逻辑使用的一类运动矢量预测。

70. 一种带有计算机可实现的指令的计算机可读媒体，指令用于配置至少一个处理单元来执行下列动作：编码视频数据；以及在所述已编码视频数据的语法部分内选择性地设置至少一个描述符，所述描述符识别与用所述视频数据编码的至少一个B帧相关的自适应空间/空间-时间编码。

71. 如权利要求70所述的计算机可读媒体，其特征在于，编码所述视频数据包括按照JVT标准来编码所述视频数据，而所述已编码视频数据的所述语法部分内的所述至少一个描述符包括：带有所述已编码视频数据的图像层部分的 `direct_mv_spatial` 参数0 72. 一种带有计算机可实现的指令的计算机可读媒体，指令用于配置至少一个处理单元来执行下列动作：编码视频数据；以及在所述已编码视频数据的语法部分内选择性地设置至少一个描述符，所述描述符识别用所述视频数据编码的至少一个中间帧内的自适应拷贝/运动拷贝跳跃模式。

73. 如权利要求72所述的计算机可读媒体，其特征在于，编码所述视频数据包括按照JVT标准来编码所述视频数据，而所述已编码视频数据的所述语法部分内的所述至少一个描述符包括：带有所述已编码视频数据的图像层部分的 `copy_mv_spatial` 参数。

74. 一种带有计算机可实现的指令的计算机可读媒体，指令用于配置至少一个处理单元来执行下列动作：在视频序列中选择性地应用与至少一部分视频帧相关的运动的空间预测；以及如果另一视频帧的参考部分的时间运动预测信息为零，则将所述运动的空间预测设为零。

75. 如权利要求74所述的计算机可读媒体，其特征在于，所述空间预测是空间零偏置的。

76. 如权利要求74所述的计算机可读媒体，其特征在于，所述空间预测还考虑到所述视频帧中其它部分的静态子分区，并且其中，如果至少一个所述子分区具有大约为零的相应运动，则所述其它部分也被视为具有零运动。

77. 一种装置，操作上被配置成用多个不同的量化功能来处理至少一部分数据以产生多个相应的经量化的数据部分，并且根据至少一个阈值选择性地输出一个所述经量化的数据部分。

78. 如权利要求77所述的装置，其特征在于，所述逻辑还用于将初始数据分成包括所述至少一部分数据的多个部分，并且根据至少一个分类特性对所述至少一部分数据进行分类，且其中所述至少一个阈值与所述至少一个分类特性相关。

79. 如权利要求77所述的装置，其特征在于，所述多个不同的量化功能包括至少两个操作上不同的恒域量化器，且其中至少一个所述恒域量化器包括一个自适应覆盖范围的恒域量化器。

80. 如权利要求77所述的装置，其特征在于，所述逻辑还用于执行码率失真优化（RDO）以选择所述一个经量化的数据部分。

81. 一种包括可操作上配置为执行下列动作的逻辑的装置：对至少一部分图像数据执行至少一个特性分析，根据所述至少一个特性分析在编码器内选择性地设置至少一个自适应量化参数，以及用所述编码器对所述至少一部分图像数据进行编码。

82. 如权利要求81所述的装置，其特征在于，所述至少一个自适应量化参数与所述编码器内的自适应覆盖范围的恒域量化器相关。

83. 一种包括操作上配置为执行下列动作的逻辑的装置：用至少两个不同的恒域量化器使至少一部分图像数据要被编码，以及根据与至少一个判决因子相关的码率失真优化（RDO）判决在所述至少两个不同恒域量化器之一的输出中识别较佳的已编码数据。

84. 一种包括操作上配置为执行下列动作的逻辑的装置：用第一恒域量化器使至少一部分图像数据要被编码；确定所述第一恒域量化器的输出是否满足至少一个判决因子；以及，如果是，则输出所述第一恒域量化器的所述输出，否则，则用不同于所述第一恒域量化器的至少一个第二恒域量化器使所述至少一部分图像数据要被编码。

85. 如权利要求84所述的装置，其特征在于，所述逻辑还用于根据码率失真优化（RDO）判决识别所述图像数据的所述至少一部分的可接受已编码版本。

86. 一种包括操作上配置为执行下列动作的逻辑的装置：对所述至少一部分图像数据进行图像分析；对所述图像数据的已分析部分进行块分类；对所述图像数据的经块分类的部分进行恒域量化；以及对所述图像数据的经恒域量化的部分进行编码。

87. 一种包括操作上配置为执行下列动作的逻辑的装置：用至少两个不同的编码器使至少一部分视频图像数据要被编码，其中所述两个不同编码器的至少一个包括操作上用于支持所述视频图像数据的无线留模式的恒域量化器；以及根据与至少一个判决因子相关的码率失真优化（RDO）判决在所述至少两个不同编码器之一的输出中识别较佳的已编码数据。

48. Claim 47 wherein the computer readable medium, characterized in that at least one of said at least one classification characteristic associated with the threshold value.

49. Claim 47 wherein the computer readable medium, characterized in that said initial data includes data selected from a group comprising image data, video data, audio data and voice data of the data, and wherein the initial data includes image data or video data, and the at least a portion selected from the group consisting block, macroblock, bars, sheets, and a section of the group.

50 47 as claimed in claim wherein the computer readable medium, characterized in that said at least one classification characteristic is selected from a set of features, the set of features include: edge characteristics, texture characteristics, smoothness characteristics, luminance characteristics, color characteristics, color characteristics, noise characteristics, object properties, motion characteristics, user preferences features, user interface focusing characteristic, hierarchical nature, timing characteristics, volume, frequency characteristics, tonal characteristics, sound characteristics, quality characteristics, bit rate characteristics, data type properties, resolution characteristics, and encryption features.

51. Claim 46 wherein the computer readable medium, characterized in that said plurality of different quantization functions include at least two different constant domains operatively quantizer, and wherein said at least one constant domain quantizer comprises an adaptive coverage deadzone quantizer.

52 as claimed in claim 46 wherein the computer-readable medium, characterized in that the at least one threshold value to selectively output one of said quantized portions of data based on said via further comprising: execution rate-distortion optimization (RDO) to select the portion of a data quantized by.

53 A computer implemented instructions with computer-readable medium, instructions for configuring at least one processing unit to perform the following actions: performing at least one characteristic of at least a portion of the image data analysis; according to at least one characteristic of the analysis selectively set at least one adaptive quantization parameter within an encoder; and with said at least a portion of said encoder encodes image data.

54. Claim 53 wherein the computer readable medium, characterized in that said at least one adaptive coverage adaptive parameters in the encoder and the quantization deadzone associated quantization.

55. Claim 53 wherein the computer readable medium, characterized in that each portion of the video frame comprises a plurality of portions of the image data, in accordance with operation analyzing at least one characteristic of the selectively disposed within said at least one of said encoder adaptive quantization parameter.

56 - kinds of computer instructions can be realized with the computer-readable media, instructions for configuring at least one processing unit to perform the following actions: with at least two different domains Di constant quantization allows at least part of the image data to be It is encoded; and according to at least one decision factor associated with rate-distortion optimization (RDO) judgment to identify preferred encoded data in at least two constant domains quantizer output of one of said different.

57 - kinds of computer instructions can be realized with the computer-readable media, the instructions for configuring at least one processing unit to perform the following actions: a first constant domain quantization allows at least a portion of the image data to be encoded; determining that the first deadzone quantizer output meets at least one decision factor, and if it is, the output of the first constant domain quantizer output, if not, different from the first permanent domain quantizer quantizing at least a second constant domain allows at least a portion of the image data to be encoded.

58 57 as claimed in claim wherein the computer readable medium, characterized by further comprising: based on rate-distortion optimization (RDO) decision identifying the at least a portion of the image data coding method has been acceptable.

88. 如权利要求87所述的装置, 其特征在于, 所述恒域量化器用于选择性地使用所述无残留模式。

89. 如权利要求87所述的装置, 其特征在于, 至少一个所述编码器操作上配置为支持从一组中选出的至少一个模式, 该组包括所述视频图像数据的直接模式和拷贝模式。

90. 一种包括操作上配置为执行下列动作的逻辑的装置: 选择性地改变至少一个拉格朗日乘法器, 该乘法器操作上配置了编码逻辑, 编码逻辑带有基于至少一部分图像数据的至少一个特性的量化功能; 以及用所述编码逻辑对所述图像数据的所述至少一部分进行编码。

91. 一种包括操作上配置为执行下列动作的逻辑的装置: 用编码器逻辑对至少一部分视频图像数据进行编码; 以及使所述编码器逻辑输出语法信息, 该语法信息识别由所述编码器逻辑使用的一类运动矢量预测。

92. 一种包括操作上配置为执行下列动作的逻辑的装置: 编码视频数据; 以及在所述已编码视频数据的语法部分内选择性地设置至少一个描述符, 所述描述符识别用所述视频数据编码的至少一个B帧相关的自适应空间/空间-时间编码。

93. 一种包括操作上配置为执行下列动作的逻辑的装置: 编码视频数据; 以及在所述已编码视频数据的语法部分内选择性地设置至少一个描述符, 所述描述符识别用所述视频数据编码的至少一个中间帧内的自适应拷贝/运动拷贝跳跃模式。

94. 一种包括操作上配置为执行下列动作的逻辑的装置: 在视频序列中选择性地应用与至少一部分视频帧相关的运动的空间预测; 以及, 如果另一视频帧的参考部分的时间运动预测信息为零, 则将所述运动的空间预测设置为零。

59 A computer instruction can be achieved with a computer-readable medium having instructions for configuring at least one processing unit to perform the following actions: at least a portion of the image data for image analysis; of the image data analysis has been part of the block classification; classification by block part of the image data of the constant domain quantization; and quantization part by the constant domain of the image data is encoded.

60. Claim 59 wherein the computer readable medium, characterized in that, from the image analysis includes edge detection analysis and texture comprising analyzing at least one type of group selected analysis.

61. Claim 59 wherein the computer readable medium, characterized in that said block classification is arranged according to the operation of the image analysis, and the deadzone quantization is arranged on said block classification according to the operating.

62 - kinds of computer instructions can be realized with the computer-readable media, the instructions for configuring at least one processing unit to perform the following actions: with at least two different coding allows at least a portion of the video image data to be encoded, wherein said at least two different encoders includes no residue configured to operatively support the model of the video image data deadzone quantizer; and associated with rate-distortion optimization according to at least one decision factor (RDO) judgment to identify preferred encoded frame data in the output of said at least two different encoders in one.

63. Claim 62 wherein the computer readable medium, characterized in that said constant domain adaptive quantizer is used for selectively and no residue of the mode.

64. Claim 62 wherein the computer readable medium, characterized in that at least one of said encoder is configured for direct mode video image data to support the operation.

65. Claim 62 wherein the computer readable medium, characterized in that at least one of said encoder is configured to support the video image data on the operating mode of the copy.

66. Claim 62 wherein the computer readable medium, characterized in that the residue-free mode depending on the run-length encoding (RLC) policy action is configured.

67. Claim 62 wherein the computer readable medium, characterized in that, with the constant domain of the encoder quantizer is configured to be selectively removed from the residual data comprises chrominance, luminance residue data a set of at least one type of residual residual residual data and all data selected.

68 - kinds of computer instructions can be realized with the computer-readable media, instructions for configuring at least one processing unit to perform the following actions: selectively changing at least one Lagrange multiplier, the multiplier operation configuration logic coding, coding with quantization logic function based on at least a portion of at least one characteristic of the image data; and the use of the logic of the image data encoding at least part of coding.

69 - kinds of computer instructions can be realized with the computer-readable media, instructions for configuring at least one processing unit to perform the following actions: encoding logic for at least part of the video encoding image data; and said encoder logic output syntax information, one motion vector information is identified by the syntax of the encoder logic used to predict.

70 A computer instruction can be achieved with a computer readable medium configured to perform the following instructions for operation of at least one processing unit: coded video data; and selecting within the encoded video data syntax section provided at least one of the descriptor, the descriptor identifying the video data encoded using at least one B frame associated adaptive spatial / space time coding a.

71. Claim 70 wherein the computer readable medium, characterized in that said video data includes encoded according to the JVT coding standard for video data, and the encoded data of the video part of the syntax of the said at least one descriptor comprising: direct_mv_spatial parameters of the picture layer portion with the encoded video data 72. 0 - computer implementable instructions are computer-readable media

with, instructions for configuring at least one processing unit adaptively and selectively disposed in a portion of the syntax of the encoded video data has at least one descriptor, identifying the descriptor of the video data encoded by the at least one intermediate frame; coded video data: to perform the following actions copy / copies of the motion skip mode.

73. Claim 72 wherein the computer readable medium, characterized in that said video data includes encoded in accordance with the JVT video coding standard data, and has the syntax of the inner portion of the encoded video data It said at least one descriptor includes: copy_mv_spatial parameters with the image layer portion of the coded video data is.

74 A computer instruction can be achieved with a computer readable medium, instructions for configuring at least one processing unit to perform the following actions: in a video sequence are selectively applied to at least a portion of the video frame related motion spatial prediction; and time for exercise if another part of the reference video frame prediction information is zero, the space will be set to zero motion prediction.

75. Claim 74 wherein the computer readable medium, characterized in that said spatial prediction is spatial zero biased.

76. Claim 74 wherein the computer readable medium, characterized in that said spatial prediction further sub-partition of the considered static in other portions of the video frame, and wherein, if at least one of said sub-partitions corresponding movement of approximately zero, then said other portion is also considered to have zero motion.

77 -. Kind of means, operatively configured with a plurality of different quantization functions to process at least a portion of data to produce a plurality of corresponding quantized portions of data, and in accordance with at least one threshold selectively outputting one of the quantized data portion.

78. Apparatus as claimed in claim 77, wherein said logic further for the initial data into a plurality of portions comprising at least a portion of the data, and classified in accordance with at least one characteristic of at least a portion of the data classification, and at least one of the at least one category in which the threshold value and the relevant characteristic.

79. Apparatus as claimed in claim 77, wherein said plurality of different quantization functions include at least two different constant domains operatively quantizer, and wherein said at least one constant domain quantizer comprises an adaptive coverage deadzone quantizer.

80. Apparatus as claimed in claim 77, characterized in that said further logic for performing rate-distortion optimization (RDO) to select one of said quantized portions of data.

81 - species may include operably configured to perform logic operation means Bian column: performing at least one characteristic of at least a portion of the image data analysis, according to the analysis of the at least one characteristic in the encoder selectively providing at least one adaptive quantization parameter, and the use of the encoder for encoding at least a portion of the image data.

82. Apparatus as claimed in claim 81, characterized in that said at least one adaptive quantization adaptive coverage of the parameters in the encoder quantizer deadzone associated.

83 - configurations include the following actions for the implementation of the operation of the logic devices: with at least two different quantization deadzone enables at least part of the image data to be encoded, and in accordance with at least one decision factor-related rate distortion optimization (RDO) decision identifying preferred encoded data at the output of said at least two different constant domains of the quantizer.

84 - operating configurations include the following actions for the implementation of logical devices: a first constant domain quantization allows at least a portion of the image data to be encoded; determining whether the first constant domain quantizer output meets at least a judgment factor; and, if so, the output of the first constant domain quantizer output, otherwise, the first constant domain different from the at least one second quantization deadzone quantization allows the at least a portion of the image data to be encoded.

85. Apparatus as claimed in claim 84, wherein said logic is further operable acceptable rate-distortion optimization according to at least a portion of the (RDO) decision to identify the version of the encoded image data.

86 - species including operating configuration logic device performs the following actions: at least a portion of the image data for image analysis; of the image data; the image data analysis section of the block has been classified after block classification part of the constant domain quantization; and by constant domain quantization portion of the image data is encoded.

87 - species including operating configuration logic device performs the following actions: with at least two different coding allows at least a portion of the video image data to be encoded, wherein said at least two different encoder comprises constant domain quantization is used to support the video image data without the residual mode operation; and according to at least one decision factor associated with rate-distortion optimization (RDO) decision output at least one of the two different encoders identify preferred encoded frame data.

88. The apparatus as claimed in claim 87, wherein said constant domain quantizer for selectively using the pattern-free residue.

Direct

89. The apparatus as in claim 87, characterized in that at least one of said encoder configured to support at least one mode selected from a group operation, the set of the video image data comprising mode and copy mode.

90. - species including operating configuration logic device performs the following actions: selectively changing at least one Lagrange multiplier, multiplier operation on the configured coded logic, with coded logic based on at least quantization of at least one functional characteristic of a portion of the image data; and the use of the image data of the encoding logic for encoding at least part.

91 comprising operatively configured to perform a logic operation means: encoding logic for at least part of the video image data is encoded; and so the encoder output logic syntax information, the information identified by the syntax a class of said motion vector prediction encoder logic used.

92 configured to perform a logic operation on the operation means includes: coded video data; and selectively disposed within the syntax of the coded video data portion at least one descriptor, the descriptor identification encoding said video data with use of at least one B frame associated adaptive spatial / space time coding a.

93 configured to perform a logic operation on the operation means includes: coded video data; and selectively disposed within the syntax of the coded video data portion at least one descriptor, the descriptor identification a copy of the video data using adaptive encoding at least one intermediate frame / copy motion skip mode.

94 comprising operatively configured to perform the following actions logic devices: the selective application of space with at least a portion of the video frame associated with motion prediction in a video sequence; and, if the reference to another part of the video frame temporal motion prediction information is zero, the space will be set to zero motion prediction.

Description:

说明书 用于数据压缩的改进优化技术 相关的专利申请 该美国专利要求美国临时专利申请序列号为60/393, 894的申请的优先权, 该 申请通过引用被完全结合于此, 于2002年7月5日公开、题为"Rate/Distortion Optimization Techniques for Image and Video Compression" 0 (1)技术领域 本发明通常涉及到的计算机和类似的设备, 特别是涉及用于压缩/编码数据 和解压缩/解码数据的方法、装置和系统。

(2)背景技术 对于数据压缩/编码和对于数据解压缩/解码需要有继续改进的方法和装置, 并且, 特别是对图像和视频数据。编码效率方面得的改进允许更多的信息更容易被计算机和它的相似设备处理/传输和/或存储。随着国际互联网和其它相似的 计算机网络以及有线通讯系统的进一步普及, 期望提供高效率的编码技术来做到充分 利用可供利用的资源。

Instructions for data compression technology patents related to improved optimization string requests the U.S. Patent claims priority to U.S. Provisional Patent Application Serial No. 60/393 priority, 894 application, the application is fully incorporated herein by reference, on July 5, 2002 Publication, entitled "Rate / Distortion Optimization techniques for Image and Video compression" 0 (1) the present invention relates generally to computers and similar devices, and more particularly for compressing / encoding data and decompressing / decoding data methods, devices and systems.

(2) Background for data compression / decompression of data for coding and / decoding need to continue to improve the method and apparatus, and, especially for image and video data. That improved coding

码率失真优化技术(RDO)在视频和图像编码/解码系统中是十分流行的,因为比之于比较常规的编码方法它们能非常好地改善编码效率。

举例来说,附加的信息,可能在计算机科学和工程的一篇科学硕士论文中可以找到,题目为“用于图像编码的一个局部自适应掩模的限模式”,作者TracDuy Tran, MIT麻省理工学院, May 1994. 就像存在继续希望来提供甚至更高的编码效率那样,也有一个对改善方法和设备的需要,它进一步的增加RDO或者其它类似技术的性能,相对于已存在的系统来改善编码效率。

(3)发明内容 当前的发明提供改善的方法和设备,它能用于压缩/编码和解压缩/解码的数据中,和/或以其它方式处理各种所包括,但不限于,图像、视频和/或音频数据的数据的格式。

举例来说,上述需要和其方面可通过下述方法满足,这个方法它包括处理至少一部分的数据,它用多个不同的量化的函数来产生多个相应的量化部分的数据,以及基于至少在一个门限值选择输出一个量化部分的数据。此方法也可能包括把初始化数据分成多个部分,以及基于至少一个分类特性对这一部分数据分类。举例来说,这里可能有一个门限值分类特性相关。

作为例子,初始化数据可能包括图像数据、视频数据、音频数据、语音数据,以及此类似的数据。被选择的部分可能采取一块、一个宏块、一条、一片、一段,或者类似的格式。分类特性可能包括边缘特性、纹理特性、平滑特性亮度特性、色彩特性、颜色特性、噪声特性、对象特性、动作特性、用户喜好特性、用户界面聚焦特性、分层特性、定时特性、音量特性、频率特性、音调特性、音色特性、质量特性、比特率特性、数据类型特性、分辨率特性、加密特性,等等的特性。

在某些示例性的实施中,大多数不同的量化特性至少包括二个不同的可操作的恒域量化器。举例来说,这里,恒域量化器可能是自适应收敛恒域量化器,可变收敛大小的恒域量化器或者其它类似的恒域量化器。此方法也可能包括编码此量化的部份。此方法可能进一步包括实施码率失真优化(RDO)去选择数据的量化部分。

根据还有的其它示例实施,另一种方法包括在至少图像数据的一部分中至少执行一个特性分析,基于该特性分析在编码器由选择性设置至少一个自适应量化参数,以及用该编码器对一部分图像数据编码。

在还有另一个示例性的实施中,被提供了一种方法,它包括至少用二个不同的恒域量化器引起图像数据的至少一部分被编码,以及基于与至少一种判决因素相关联的码率失真优化(RDO)判决识别在至少二个不同恒域量化器其中的一个输出中的最佳编码数据。

上述需要和其它也被一种方法所满足,此方法包括至少引起一部分图像数据被用第一只恒域量化器所编码,确定是否第一只恒域量化器的输出满足至少一种判决因素,同时假如是这样输出第一恒域量化器的输出。如果不是这样,那么引起一部分图像数据被用至少与第一恒域量化器不同的第二恒域量化器所编码。举例来说,该方法也可能包括基于RDO判决或者其它相似的判决识别一部分图像数据的可接受的编码方式。

在仍旧是另一种示例性的实施中,一种方法包括在至少一部分图像数据中执行图像分析,在被分析的这一部分图像数据上执行块分类,执行块分类部分图像数据的恒域量化,以及执行图像数据恒域量化部分的编码。举例来说,这里,此图像分析可能包括边缘检测分析、纹理分析等等。

根据进一步示例性实施,提供了一种方法,它包括至少引起一部分视频图像数据被采用至少二个不同的编码器所编码,其特征在于,所述的至少二个不同的编码器其中的一个包括恒域量化器,可操作地配置为去支持无线、留模式的视频图像数据以及基于与至少一个判决的因素相关的码率失真在二个不同的编码器其中一个的输出中识别最佳的编码帧数据。优化(RDO)判决。

另一种示例性方法包括选择性地改变至少一个Lagrangian乘数,这是可操作配置的编码逻辑,它具有基于至少一部分图像数据的至少一个特性的量化函数,以及用编码逻辑来编码这一部分的图像数据。

仍旧在另一种执行中,示例方法包括采用编码逻辑编码至少一部分视频图像数据,以及使用编码器逻辑使编码逻辑输出标识运动矢量预测类型的语法信息。

用在表达视频编码关系信息的一种方法包括编码视频数据,以及在编码视频数据的语法部分的选项性地设置至少一种描述符,此描述符标识与用视频数据编码的至少一个B帧有关的一种匹配的空间/空间-时间编码。用在表达视频编码相关信息的一种方法,包括编码视频数据,以及在编码的视频数据的语法部分选项性地设置至少有一种描述符,此描述符至少一个用视频数据编码的中间帧中标识一种匹配拷贝/运动拷贝跳跃模式。

提供了考虑视频帧的时间/空间部分的视频编码用于时间标记独立模式中的示例性方法。举例来说,这里,此方法包括选择性地应用与视频序列的至少一部分视频帧相关的运动的空时预测。并且如果对于另一视频帧的参照部分的时间运动预测信息为另,则,设置运动的空间预测为另。

efficiency allows more information to be more susceptible to computer capacity and other similar devices handle transmission and / or storage. With the further popularization of the Internet and other similar computer networks as well as wired communications systems, it is desirable to provide highly efficient coding techniques to achieve full use of resources available for use.

Rate distortion optimization (RDO) in video and image encoding / decoding system is very popular, because the ratio of the more conventional encoding methods in that they can improve the coding efficiency is very good.

For example, additional information may be found in a scientific master's thesis in computer science and engineering, entitled "for a locally adaptive image coding mask limit mode," the authors TracDuy Tran, MIT Massachusetts Institute of technology, May 1994. just as there continue to hope to provide even higher coding efficiency, as there is also a method and apparatus for improving the needs of its further increase the performance of RDO or other similar techniques, as opposed to already Existing systems to improve coding efficiency.

(3) Disclosure of the current invention to provide an improved method and apparatus which can be used compression / decompression encoding / decoding of data, and / or otherwise deal with included, but not limited to, FIG. like, video and / or audio data format of the data.

For example, the above-described and other aspects need to be met by a method, the method comprising processing at least a portion of the data, it uses a plurality of different quantization functions to produce a plurality of corresponding quantized portions of data, and based on at least one threshold to select a quantization data output section. This method may also include the initialization data into a plurality of parts, as well as based on at least one characteristic of this part of the classification data classification. By way of example, he said there may have a threshold classification characteristic associated.

As an example, the initialization data may include image data, video data, audio data, voice data, and the like of such data. The selected parts may take a piece, a macro block, one, one, some, or a similar format. Classification features may include edge characteristics, texture characteristics, the smoothing of luminance characteristics, color characteristics, color characteristics, noise characteristics, object properties, operating characteristics, user preferences features, user interface focusing characteristic, hierarchical nature, timing characteristics, volume, frequency characteristics, tonal characteristics, timbre characteristics, quality characteristics, bit rate characteristics, data type features, resolution characteristics, encryption, and so on characteristics.

In certain exemplary embodiments, most of the different characteristics comprise at least two different quantization operable deadzone quantizer. For example, where the deadzone quantizer may be adaptive convergence deadzone quantization, variable-sized convergence deadzone quantization, or other similar deadzone quantizer. This method may also include encoding the quantized part. This method may further include the implementation rate distortion optimization (RDO) to select quantization section data.

According to still other exemplary embodiments, another method includes performing at least a portion of the image data of at least one characteristic analysis, the analysis in the encoder selectively set at least one parameter based on the adaptive quantization characteristics, and by the encoder part of the image data is encoded.

In yet another exemplary embodiment, a method is provided, which comprises at least two different constant domains with a quantizer causes at least a portion of the encoded image data, and based on at least one decision factors associated rate-distortion optimization (RDO) judgment to identify preferred encoded data in at least two different quantization deadzone one output. above and other needs are also satisfied by a method, the method comprising causing at least a portion of the image data is encoded with a first constant domain quantizer, only the first constant domain determines whether the output of the quantizer to meet at least one of species decision factors, but if it is so that the output of the first constant domain quantizer output. If not, then the cause is a part of the image data with at least the first constant domain of a second different quantization deadzone quantizer encoded. By way of example, he said that the method may also include judgments based RDO or other similar judgments identify a portion of the image data of acceptable encoding.

In still another exemplary embodiment, a method includes performing image analysis on at least a portion of the image data, perform block

(4)附图说明 对本发明的各种方法设备和系统更完整的了解,可参照下述详细的附图并结 合附图而给出: 图1是一张框图,它以计算机形式描述了一个实例的设备,它适用于本发明的某些实现中。

图2(a-d)是按照本发明的某些实现的一组描述示例性可选择恒域量化的图表。

图3是一张流程图,它描述了按照本发明的某些实现选择性地对数据应用不同量化处理的示例方法。

图4是一张框图,描述了按照本发明的某些实现选择性地对数据应用不同量处理的示例逻辑。

图5是一张框图,描述了选择性地对数据的应用不同量处理的示例逻辑,它是与本发明的某些进一步的实施相符合的示例性逻辑。

图6(a-b)是逻辑框图,描述了选择性地对数据的应用不同量处理的示例逻辑,它与本发明的仍旧与其它的实施相符合的示例性逻辑。

图7是一张逻辑框图,描述了选择性地对数据的应用不同量处理的示例逻辑,它与本发明的某些进一步的实现相符合。

图8是一张示例性语法信息的表格,按照本发明的某些实现用于对选择性地对运动矢量应用不同的预测方法的逻辑。

图9是一张框图,它描述了按照本发明的某些实现的示例性逻辑,用于选择性地对视频数据应用不同编码的图表。

图10是一张框图,它描述了按照本发明的某些步骤的实现的示例性逻辑,用于选择性地对视频数据应用不同编码的图表。

图11是一张图解图,它描述了按照本发明的某些实施的某些视频序列的特性,它使用选定的应用编码方案。

图12是一张流程图,它描述了按照本发明的某些实现的一种示例性的方法,用于对直接模式视频序列的空间-时间预测。

(5)具体实施方式 转到附图,其特征在于,所述的相同的参照数字被指示相同的元件,本发明 解释为,在合适的计算环境下被实施。虽然不是必须的,本发明将以通常的计算机可执行指令的环境下描述,比如由计算机服务器来执行的程序模块,它可以采取个人计算机、工作站、专门的服务器、多处理器、大型计算机等的形式。通常,程序模块包括子程序、程序、对象、组成、数据结构等等,它执行专门的任务或者 执行特殊的抽象的数据类型。本发明也可能在分布的计算环境下被使用。这里任务 被远程处理设备来执行,它经过通讯网络来联接。在分布计算的环境下,程序模块 可能被定位在本地和远程存储设备二者之中。

示例性计算环境 图1描述了一个合适的计算环境120的实例,在这时,以后描述的方法和安 排可在此环境中被实施。

示例性计算环境120只是一个合适计算环境的例子,以及不打算建议作任何的限制,比如对使用的范围,或者可改进方法的功能,以及在这里所安排的描述的 限制。计算环境120不应被解释为与计算环境120中描述的任一组件或组件的组合 有相关的依赖性或者要求。

改进的方法和这里的安排是可与很多其它通用或者专用的计算机系统环境或者 结构一同操作。

如图1所示,计算环境120包括以计算机130形式的通用计算设备。计算机130组件可包括一只或者几只处理器或者处理单元132、系统存储器134、以及把 各种系统部件,包括系统存储器134耦合至处理器132的总线136。

总线136表示一个或者多个若干总线结构类型的任何一种,它包括存储器总线或者存储器控制器、外部设备总线、加速图形端口,以及处理器和使用各种总线 结构任何一种的本地总线。举例来说,(但不是以此为限),这些结构包括工业标准 (ISA)总线、微通道结构(MCA)总线、增强型(EISA)总线、视频电子标准协会(VESA) 局部总线,以及外部设备互连 (PCI)总线(也称为Mezzanine总线)。

计算机130典型地包括各种计算机可读媒体,这些媒体可能是任何可利用的 媒体,它可由计算机130访问,并且包括易失性媒体和非易失性媒体、可移动式和 不可移动式二者。

图1中,系统存储器134包括计算机可读媒体,以易失性存储器的形式,比如, 随机存取存储器(RAM) 140,和/或非易失性存储器形式,比如,只读存储器 (ROM) 138。基本输入/输出系统(BIOS) 142,包括基本的子程序,它帮助在计算机 130内的部件之间传送信息,比如,当启动的时候,被存储在ROM138中的信息。

RAM140典型地包括数据和/或程序模块,它可立刻被处理器132访问和/或当前正 被处理器132的操作。

计算机130可能进一步包括其他可移动式/不可移动式的、易失性/非易失性的计算机存储媒体。举例来说,图1给出了一个硬盘驱动器144,它用来读来自或 者写到非移动式、非易失性的磁性媒体中(没有表示在图中同时,

classification in this part of the image data is analyzed, the execution block classification part of Fig. deadzone quantization image data, and image data encoding performed deadzone quantization portion. For example, here, the image analysis may include edge detection analysis, texture analysis and the like.

According to a further exemplary embodiment, there is provided a method which comprises causing at least a portion of the video image data is preclude the use of at least two different encoders encoded, characterized in that at least two of said different encoding wherein the device comprises a constant domain quantizer operatively configured to support no residual mode of the video image data and at least one decision based on factors related to the rate-distortion in two different encoder output of one identifying preferred encoded frame data. Optimization (RDO) decision.

Another exemplary method includes selectively varying at least one Lagrangian multiplier, which is encoded logic operable configuration, having at least a part of the image quantizing function based on at least one characteristic of the data, and the encoding logic this part of the coded image data.

In still another implementation, an exemplary method includes encoding logic preclude the use of video encoding at least a portion of the image data, and an encoder encoding logic logic output identifies the motion vector prediction type syntax information. is used in the expression of the relationship between a video coding method includes information encoded video data, and selectively disposed in a portion of the syntax of the encoded video data of at least one descriptor, the descriptor identifies the encoded video data of at least one a match-related B-frame space / time coding a space. An intermediate frame for use in expressing a coding method for video-related information, including the encoded video data, and the syntax is selectively disposed on part of the encoded video data of at least one descriptor, the descriptor with the at least one encoded video data are identified A matching copy / copies of the motion skip mode. provides video coding video frames to consider time / space part of the time stamp in stand-alone mode for exemplary method. For example, where the space, the method comprising selectively applying at least a portion of the video sequence of video frames associated with motion prediction. If the time and motion video frame with reference to another part of the prediction information to another, then, set the motion prediction space to another.

(4) Brief description of the various methods and systems of the present invention, apparatus is more complete understanding can be detailed with reference to the following description and the accompanying drawings which are given: FIG. 1 is a block diagram, which is described in the form of a computer an example of the device, which applies to certain implementations of the present invention.

Figure 2 (a-d) is in accordance with a description of exemplary optional deadzone quantizer chart some implementations of the invention.

FIG. 3 is a flow chart which describes some implementations according to the present invention is selectively applied to the data with no quantization example method.

Figure 4 is a block diagram illustrating an example logic in accordance with some implementations of the present invention for selectively applying different amounts of data processing.

Figure 5 is a block diagram illustrating an example logic selectively apply different amounts of data processing, it is associated with certain embodiments of the present invention, a further illustrative logical consistent.

FIG. 6 (a-b) is a logical block diagram illustrating an example of a logic selectively different amounts of data processing applications, and still other exemplary embodiments consistent with the present invention, it is logical.

Figure 7 is a logical block diagram illustrating exemplary logic for selectively varying amounts of data processing applications, it is with certain further implementations consistent with the present invention.

FIG. 8 is an exemplary syntax information form, according to some realization of the present invention is used to selectively apply a different motion vector prediction method logic.

FIG. 9 is a block diagram which describes in accordance with certain implementations of the exemplary logic of the present invention for the chart data selectively apply different video encoding.

典型的称为“硬设备”，磁盘驱动器146用于读来自或者写到可移动式、非易失性的磁盘148上(比如，“软磁盘”)，以及光盘驱动器150用来读来自或者写到可移动式、非易失性光盘152，比如，CD-ROM、CD-R、CR-RW、DVD-ROM、DVD-RAM或者其它的光媒体。硬盘驱动器144、磁盘驱动器146以及光盘驱动器150它们每个设备由一个或者多个接口154连接到总线136上。

驱动器和相关的计算机可读媒体提供了计算机可读指令、数据结构、程序模块，以及其它从计算机130来的数据的非易失性贮存。虽然此处示例性环境描述采用硬盘、可移动式磁盘148以及可移动式光盘152，本领域中的技术人员应当理解其它类型的可由计算机来存取所存储的数据的计算机可读媒体，比如，盒式磁带、闪存卡、数字视频盘、随机存取存储器(RAMs)、只读存储器(ROM)、以及其它类似的设备，也可以用在示例性的操作环境中。

一些程序模块可能被存储在硬盘、磁盘148、光盘152、ROM138、或者RAM140中，包括比如操作系统158、一只或者多只应用程序160、其它的程序模块162和程序数据164。

在这里描述的改进方法和安排其特征在于，可用所述的操作系统158、一只或多只应用程序160、其它的程序模块162、和/或程序数据164来实现。

用户可以经过输入设备，比如键盘166的触点设备168(比如“鼠标”)提供命令和信息进入计算机130。其它的输入设备(图中没有表示出来)可能包括话筒、控制杆、游戏板、卫星式转盘、串行端口、扫描仪、照相机等等。这些和其它的输入设备通常经过用户输入接口170连接到处理单元132，它连接到总线136上，但是也可采用其它的接口和总线结构来连接，比如像并行端口、游戏端口或者适用的串行总线(USB)。

显示器172或者其它类型的显示设备也可以经过一个接口，比如视频适配器174，连接到总线136上。另外，显示器172、个人计算机典型地包括其它外部输出设备(图中没有表示出来)，比如扬声器和打印机，它们可以经过输出外部设备接口175来进行连接。

计算机130可以在用逻辑连接到一台或多台远程计算机的网络环境下进行操作，比如用远程计算机182。远程计算机182可能包括在这里和计算机130相关联的很多或者所有的部件的性质的描述。

图1中所给示的逻辑连接是局域网(LAN)177以及广域网179。这样的网络环境是在办公室、企业级计算机网络、企业内部互联网(intranets)以及国际互联网是常见的。

当用在LAN局域网环境下时，计算机130经过网络接口或者适配器186连接到局域网LAN177上去。当用在广域网环境下时，计算机典型地包括有调制器178或者其它用于在WAN179广域网上建立通讯的装置。调制器178它可以是内置的或外置的，经过用户输入接口170或者其它合适的机制可以被连接到系统总线136上。

图1所描述的，是一个经过国际互联网的专门的WAN广域网的实施环境，这里，计算机130使用调制器178经过国际互联网180与至少一台远程计算机182建立的通讯。

在一个网络的环境下，所描述的相关于计算机130的程序模块或者其中的部分，可能贮存在远程存储器贮存设备上。这样，像图1所描述的那样，远程应用程序189可能驻留在远程计算机182的贮存设备上。将会理解到，所给示的和所描述的网络连接是示例性的，也可以使用可以计算机之间建立通讯链路的其它装置。

改进的码率/失真优化技术 虽然下列部分描述了某些示例性方法和装置，它们初始被构成压缩/编码和解压/解码图象数据和/或视频数据，在数据压缩领域中的技术人员将认识到此技术目前能被适应和使用在其它类型的数据压缩/编码以及解压/解码方面。举例来说，某些方法和设备可能被迫应用在用于视频数据、语音数据以及其它类的数据的压缩/编码方面。

此外，虽然此示例性方法和设备能在有计算机内用逻辑来构成，本领域的技术人员将认识到这些方法和设备可能被用其它类型的设备、电路用具等来实施。用在此处的术语“逻辑”的含意是包括硬件、固件、软件或者任何它们的组合，以及任何其它的支持硬件或其它的机制，可能被要求用来实现所描述的整个的或者部分的功能。

记住了这些，一些示例性方案被呈现在这里，它能以某种逻辑形式实施去支持处理数据。根据本发明的某些方面，一些新颖的技术被介绍用于改进对视频/图象和编码/解码系统的性能。在一些示例性实施中，这些技术被与图象/视频的标准，比如JPEG和JVT(连接视频组)标准(比如，H.264/AVC)一起使用。举例来说，对JVT的情况下，语法变化被介绍，它能操作上启用不同预测类型的自适应选择，这些不用预测类型可以被用于预测视频的某些参量，比如，运动信息。

码率失真优化(RDO)技术在视频和图象编码/解码系统中是十分流行的，因为比之于比较传统的编码方法它们能非常好地改进编码效率。存在一个不断的希望来提供甚至更高的编码效率。这里描写描述方法和设备能显著地改善RDO或者其它类似的技术的性能以达到相对于现有系统改善编码效率。

FIG. 10 is a block diagram which describes the steps, according to some implementations of the present invention, exemplary logic diagrams used to selectively apply the video data coded differently.

FIG. 11 is a diagram that describes some of the features in accordance with some embodiments of the present invention, a video sequence, which uses the selected application coding scheme.

FIG. 12 is a flow chart which describes in accordance with a certain exemplary method of the present invention is implemented, for video spatial direct mode prediction for a time series.

(5) DETAILED DESCRIPTION Turning to the figures, wherein the same reference numerals indicate like elements being, the present invention is construed as being implemented in a suitable computing environment. Although not required, the present invention will generally down the computer-executable instructions described environment, such as program modules, executed by a computer server, which may preclude taking a personal computer, a workstation, a dedicated server, a multi-processor, mainframe computers, etc. and other forms. Generally, program modules include routines, programs, objects, composition, data structures, etc. that perform specialized tasks or to perform particular abstract data types. The present invention may also be in a distributed computing environment is used. Here the task is to execute remote processing device, it passes through the communication network connection. In the distributed computing environment, program modules may be located in both local and remote storage devices into. exemplary computing ring yu Figure 1 depicts an example of a suitable computing environment 120, in this case, after the method described and arrangements can be implemented in this environment. exemplary computing environment 120 is only one example of a suitable computing environment, and does not intend to recommend any restrictions, such as the scope of use or functionality of the method can be improved, as well as in the description herein arranged limit. Should the computing environment 120 be interpreted as any component or combination of components described computing environment 120 and is related to the dependency or requirement.

The improved methods and arrangements herein are operable in conjunction with many other general purpose or special purpose computing system environments or structures.

1, the computing environment 120 including 130 in the form of a general purpose computing device. Computer assembly 130 may include a processor or processing unit or a few 132, system memory 134, and the various system components, including system memory 134 coupled to the processor bus 132 136. bus 136 represents one or more of any of several types of bus structures, including a memory bus or memory controller, peripheral bus, an accelerated graphics port, and a processor bus structure and use of any one local bus. For example, (but not so limited), such architectures include Industry Standard (ISA) bus, Micro Channel Architecture (MCA) bus, Enhanced (EISA) bus, Video Electronics Standards Association (VESA) local bus, and the external device interconnect (PCI) bus (also known as Mezzanine bus).

Computer 130 typically includes a variety of computer readable media, these media may be any available media, which is accessed by computer 130 and includes both volatile and nonvolatile media, media, removable and nonremovable type two.

1, the system memory 134 includes computer readable media in the form of volatile memory, such as random access memory (RAM) 140, and / or non-volatile memory in the form of, for example, read only memory (ROM) 138. Basic input / output system (BIOS) 142, including basic routines that helps to transfer information between components within the computer 130, for example, the start time, the information is stored in the ROM138.

RAM140 typically includes data and / or program modules that can be instantly accessed by the processor 132 and / or operations currently being processor 132. computer 130 may further include other removable / non-removable, volatile / non-volatile computer storage media. For example, Figure 1 shows to a hard disk drive 144, which is used to read from or write to a non-removable, non-volatile magnetic media (not shown simultaneously, typically referred to as "hardware device" in FIG.), a disk drive 146 for reading from or written to a removable, non-volatile disk 148 (eg, "floppy disk"), and an optical disk drive 150 for reading from or writing to removable, non-volatile optical disk 152, such as, CD-ROM, CD-R, CR-RW, DVD-ROM, DVD-RAM or other optical media. Hard disk drive 144, magnetic disk drive 146 and optical disk drive 150 are each device by one or more interfaces 154 connected to the bus 136. drives and associated computer-readable media provide a computer-readable instructions, data structures,

按照本发明的某些进一步的方面，它也可以用图像预分析概念，比如，举例来说，边缘和/或纹理检测，根据图像或宏块的特性通过使用自适应和/或可变大小的恒域量化，进一步的组合和改进RDO或者其它类似的技术。

这一描述也介绍/定义了一些附加的示例性语法变化，它可能被执行去启动在一个帧层次上的不同预测方案的组合体，举例来说，这进一步的改善视频编码方案的执行。举例来说，一个技术提升来自这样的事实，它对序列的一些帧或者一部分、运动在时间域可能比在空间域更相，或者反过来也是这样。这可能通过执行帧的予分析而加以利用，但是也可通过使用二个或者可能更多不同的方法对相同帧编码并在RDO和/或类似于RDO的选择一个更合适的方法而被利用。此较佳的方法然后可在结果数据中用信号表示，举例来说，在帧头中以允许编码器适当的编码此帧。这里，举例来说，一个这样的示例性方法在B帧中可能包括直接模式可能的变化，通过或者采用空间预测或者时序预测达到，或者在P帧中包括跳跃模式的运动矢量选择，预测运动矢量参数或者时间预测动作矢量参数，或者甚至为另而达到。

根据一个或者多个特性(参数)对数据应用匹配和/或不定的恒域量化到：在图像数据压缩系统中，一部分数据被压缩，比如，举例来说，块或者宏块，可能确实包括更多的重要的信息，当与其它的信息(数据)相比时应被不同地编码(例如，以较高的优先级，以较高的质量，等等)。达到的一个办法是使用不同的量化器值。举例来说，在某些实施中，一个较小的量化器值可能被用于对“更加重要的信息”，以及一个较大的量化器值可能被用于对“较小重要的信息”。然而，这样做后一般还将要求此信息的传达标示出用于对每个块、宏块、宏块的组，等等的一个量化值，使得其后的解压/解码是成功的。不幸的是，这些附加的信息趋向于增加压缩的总开销，以及编码器的复杂性。因而，不是增加了效率，反而实际可能会降低了效率。

注意到图2(a-d)，每张图表描述了某些示例性恒域量化可使用在某些示例性图像/视频编码的方案中。在图2(a-d)的每一幅图表中，垂直轴(Y轴)呈表示量化值以及水平轴(X轴)表示原始值。图2(a)中给出了示例性恒域量化202。举例来说，有关于恒域量化A，它相对于均匀量化能相当大地改进编码效率。常规的恒域量化器在整个量化过程中经常是保持为常数或者均匀量化，如此，可能的话，不完全利用数据中全部存在的冗余。

考虑到或者另外方式计及数据中信息的重要性，以及在块/宏块的基础上，通过自适应/修改恒域量化量可以达到编码效率的改进。举例来说，例如，通过自适应收敛(例如，沿的X轴)，但不改变重建值能做到。举例来说，分别地将恒域量化器202和恒域量化器204，206，和208比较，(它们表示在图2b, c, d中)。

这里，重建值可能对全部的编码数据保持为常数，当然除非发出量化参量有变化的信号。举例来说，由增加此另电平箱位(图2(d)中)，使更多的数据将被分配到这里，这样，依赖于压缩的方案，达到较高的压缩。显而易见地，不需要在解码器上述信号表示量化器中的改变，因为重建值仍日保持相同值。虽然有人可能会争辩说，这样将损害量化的性能，通常这不必定是正确的，如果这种处理是选择性地完成，例如当某些条件被满足时，若用新的量化器达到的压缩大大地高于遭受的失真。

举例来说，(图2(b))恒域量化仅仅影响另电平箱位，相对于第一个+/-箱位，然而所有其它的箱位仍日保持相同，以及所用的重建值没有变化。一个带有自适应估计恒域的自适应方案也是可能的(例如，使用码率失真优化和自适应估计)。

根据本发明的某些示例性的实施，量化选择逻辑，因此，被提供在压缩/编码系统中，以基于至少一个特性或者参数在不同的恒域量化器(或者量化值)之间选择。举例来说，此逻辑可能按照某些图像特性在不同的恒域量化器202、204、206和208之间进行选择。

更加特别的是，此逻辑可能被配置成在图像数据中表征纹理结构和/或边缘作为代表“较重要信息”，并因而用一种式样对数据编码来提供一种较高水平的质量。

此逻辑能用常规的纹理分析/检测及边缘检测测定算法来支持这种判决过程。

有了这个概念之后，现在参见图3，这是一张流程图，它说明了按照本发明的某些实现将不同的量化过程选择性地应用于数据的方法300。在动作302中，初始数据组被分成数据的多个部分。例如，图像或者视频数据可能被分成多个块、宏块、条、片、区段等等。在动作304中，来自动作302的多个数据部分的至少一个以某些方式被分类，例如，基于至少一种特性或参数。例如，分类可以基于边缘特性、纹理特性、平滑特性、亮度特性、色度特性、色彩特性、噪声特性、对象特性、动作特性、用户喜好特性、用户界面聚焦特性、分层特性、定时特性、音量特性、频率特性、音调特性、音色特性、质量特性、比特率特性、数据类型特性、分辨率特性、加密特性。

在动作306中，用多个量化过程的至少两个来处理来自动作304的已分类部分，以产生相应经量化的数据。在动作308中，作出决定选择来自动作306的一个经量化的数据，例如，基于与动作304中使用的分类参数相关的至少一个阈值或测量。在动作310中，来自动作308的经量化的数据以某些允许其随后被解码的方式而被编码。

program modules, and other non-volatile storage of data from the computer 130. Although the exemplary environment described herein preclude the use of a hard disk, removable disk 148 and a removable optical disk 152, one skilled in the art will appreciate other types of computer by a computer to access the stored data readable media, for example, tape cassettes, flash memory cards, digital video disks, random access memories (RAMs), read only memory (ROM), and other similar devices can also be used in the exemplary operating environment.

Some program modules may be stored on the hard disk, magnetic disk 148, optical disk 152, ROM138, or RAM140, including operating systems such as 158 - or more than just only 160 applications, other program modules 162, and program data 164. improved methods and arrangements described herein wherein the available operating system 158, one or more applications only 160, other program modules 162, and / or program data 164 to achieve.

Users can through input devices such as a keyboard 166 contact devices 168 (such as "mouse") provides commands and information into the computer 130. The other input device (not shown) may include the words simple, lever, game board, satellite wheel, serial port, scanner, camera, etc. these and other input devices are often through the user input interface 170 is connected to the processing unit 132, which is connected to the bus 136, but may also preclude the use of other interfaces and bus structures to connect, such as a parallel port, game port or a suitable serial bus (USB). monitor 172 or other type of display device can also be via an interface, such as video matching unit 174, connected to the bus 136. In addition, the monitor 172, personal computers typically include other peripheral output devices (not shown), such as speakers and printers, which may be connected to an external device through the output port 175 to make the connection.

Computer 130 may operate in a logical connection to one or more remote computer's network environment, such as using a remote computer 182. Remote computer 182 may be included in the computer 130 and here may one or all of the relevant characteristics of the components connected to the description. logical link given in Figure 1 is shown in a local area network (LAN) 177 and a wide area network 179. Such network environments are commonplace in offices, enterprise-wide computer networks, intranets (intranets) as well as the Internet is common.

When used in a LAN LAN environment, the computer 130 through a network interface or adapter 186 connected to the LAN LAN177 up. When used in a WAN environment, the computer typically includes a modulator 178 or other means for establishing communications over the WAN in WAN179 means. Modulator 178 which may be internal or external, through 170 or other appropriate mechanism user input interface can be connected to the system bus 136.

Figure 1 described, it is an implementation of the Internet through a dedicated WAN WAN environment, which, the computer 130 using the modulator 178 through the Internet 180 and at least one remote computer 182 to establish communication.

In a network environment, as described in relation to a computer program or a module 130 which portions may be stored in a remote memory storage device. Thus, as described in Figure 1 as a remote application 189 may reside on a remote computer storage device 182. It will be understood, given a diagram and description of the network link is exemplary, other means may also be used to establish communication links between computers.

Improved rate / distortion optimization technology, while the following sections describe some of the exemplary method and apparatus, which initially constitute compression / decompression and encoding / decoding image data and / or video data in the field of data compression the art will recognize that this technology is currently in use and can be adapted to other types of data compression / decompression and encoding / decoding on. For example, some methods and apparatus may be adapted to be used in video data, voice data, and other types of data compression / encoding.

In addition, although this exemplary method and apparatus can be constituted in the computer there is a logical, those skilled in the art will recognize that these methods and devices may be other types of equipment, electrical appliances, etc. to implement. As used herein, the term "logic" is meaning including hardware, firmware, software, or any combination thereof, and any other supporting hardware or other mechanisms, it may be required to implement the described functionality whole or part of the.

Remember these, some exemplary programs is presented here, it can implement some form of logic to support the processing of data.

使用了方法300之后,可以提供一种逻辑,实质上分析了图像或视频数据的部分并将某些部分视为比其它部分重要。不同的恒域量化器被用于较重要的数据部分,并且分析所产生的经量化的数据来确定哪一种恒域量化器满足该较重要的数据的期望阈值要求。例如,可以强制执行一种质量或噪声阈值要求。

通过进一步的示例,在视频中,序列边缘通常在运动补偿技术中起到重要的作用。例如,这通过下面详述的图7中的框图得到说明。基本上,可以对图像(帧)或其部分进行图像分析,并且根据分析来决定多个恒域量化器中哪一个最适用,例如,按照码率失真优化(RDO)准则。

参考图4,逻辑400说明了按照本发明某示例性实现的快速判决过程。这里,输入帧/图像(或其中的部分)在块402中受到分析。在该例中,图像分析包括边缘检测和/或纹理分析能力,且输出被提供给恒域判决块404。恒域判决块404然后使编码器406在对输入帧/图像(或其中的部分)编码时使用指定的或另外选择的恒域量化器或量化值。

图5中的逻辑500可配置为用于支持方法300,而不是作出关于恒域量化的快速判决,其中使用了多个恒域量化器并且根据某阈值标准来作出RDO判决。这里,输入宏块(MB)/图像502(或其它类似部分)被(选择性地)提供给不同的恒域编码块504、506、508和/或510,来自这些不同恒域编码块的输出在RDO判决块512中被分析,选定的已编码数据被输出。如该例中所述,某些或全部恒域编码块过程可能并行发生。在其它实现中,这种过程可以被安排成串行发生。

图6(a)和图6(b)说明其中使用了选择性重编码判决过程的逻辑。例如在图6(a)中,输入MB/图像602被提供给恒域编码块602和604。重编码判决块606考虑来自恒域编码块602的输出并且影响在恒域编码块602和604的输出间的选择608。重编码判决块606也可以选择性地启动恒域编码块604。在图6(b)中,逻辑610类似于600,除了与重编码判决块606相关的选择608包括RDO判决块612,用于分析来自恒域编码块602和604的输出并选择输出哪一个。

图6(a,b)中的重编码判决块可以用于根据各种标准作出判决。例如,在某些实现中,可以考虑质量限制、速率限制和/或其它RDO概念/阈值。

只有当满足了某些前面定义的条件,譬如速率/质量超出特定阈值,才可能使用附加的恒域量化器。在某些示例性实验中,成功地选择了恒域量化器,它比原来具有30%的较大恒域。另外,可能按照诸如AC频率或边缘类型这样的图像特性,也可以使用自适应恒域量化器。

在图7中,示例性编码逻辑700说明性地阐述了怎样按照图像分析过程704来处理初始图像702(或其一部分)以产生该例中的边缘检测数据706。然后,边缘检测数据706被提供给块分类过程708以产生块分类的数据710。然后,块分类的数据710与初始图像702一起被提供给图像编码过程712,该过程接着产生编码图像714。这是图像的简单编码过程的一个示例。这里,图像被分析(如,用边缘检测算法),然后,块按照这种信息而被分类。基本上定义了N维矩阵(DZp)(N取决于所定义恒域的数量),该矩阵后来为在宏块的位置(i,j)处分配的量化器(QP)分配适当的恒域(DZ)。

无残留模式的考虑(如,对于B帧直接模式)-即使上述说明可能主要针对平稳图像,然而也可以将同一概念应用在视频上。

例如,也可以在视频压缩中以相当直接的方式使用图像预分析和恒域量化器。

一种相当有趣且可以大大受益于上述概念的情况是B帧内直接模式的使用。

直接模式基本上是不需要传输运动参数的特殊方式,这是因为这可以直接或通过空间或通过时间上的预测而被直接预测到。关于直接模式的其它信息可以在美国未决临时专利申请序列号为60/385,965的申请中找到。

如果没有要被发送的剩余物,则可以按照本发明的某些进一步方面,通过使用被称为无残留直接模式的特殊方式进一步改进效率。如下所述,无残留直接模式可以利用游程长度编码(RLC)策略。这里,例如,如果产生的失真足够小且由于RLC的较高效率引起的比特率的减少足够大,则恒域量化器可以提供期望的解决方案。基本想法是在某些条件下选择性地引起或强制使直接模式无残留地被编码,尽管残留存在。例如,如图9所示,可以实现基于与图5相同的概念的方案,其中无残留直接模式与所有其它有效模式相比,也在RDO过程中被检查。在某些情况下,由于所使用的RDO对于这种情况并不适当,因此这种方案的性能可能不如期望中那么好。其它依赖于量化值的过程也被影响,譬如在用于消除块内人造假像的内环滤波器(未示出)。更明确地说,即使性能在较低的比特率显示良好,然而性能在较大比特率处可能大大受损;通过使用较大的量化器并且不考虑无残留直接模式,性能可能变得更好。

与图像实例中所作的相类似,按照本发明的某些方面,该逻辑可被配置成仅在满足某些前面定义的条件时、特别在与直接模式相关的残留不重要时,才考虑这些方式。该逻辑可以,例如,被配置成通过检查直接模式的已编码块模式(CBP)而将对该残留重要性的估计用作一个条件。如果不考虑色度信息时CBP低于特定阈值,则可以考虑这个,意味着该残留并不重要,并且如果跳过它不会造成太大的失真。而且,诸如边缘和纹理的不存在这样的其它图像特性也可以用于这种视频编码过程/逻辑中。

According to certain aspects of the present invention, some new technologies are introduced to improve the performance of video / image encoding / decoding system. In some exemplary embodiments, these techniques are the image / video code standard, such as JPEG and JVT (video connection group) standard (eg, H.264 / AVC) - used together. For example, on the case of JVT, the syntax changes to be introduced, it can enable different types of adaptive prediction selection operation, these types can not be used to predict certain parameters predicted video, for example, the motion information. rate-distortion optimization (RDO) technology in video and image encoding / decoding system is very popular, because in more traditional than the coding method which can improve the coding efficiency very well. There is a continuous wish to provide even higher coding efficiency. Describe a method and apparatus described herein can significantly improve RDO or other similar techniques to achieve relatively of existing systems to improve the coding efficiency.

According to certain further aspects of the invention, it is also possible to use an image pre-analysis concepts, such as, for example, edge and / or texture detection, according to the characteristics of the image or the macroblock by using an adaptive and / constant or variable size field quantization, further combinations and improvements RDO or other similar technology.

This description also describes / defines some additional exemplary syntax changes, it can be executed to start the assembly at a different level frame prediction program, for example, further improvement of video coding scheme carried out. For example, from the fact that such a technology to enhance the number of frames or a portion of its sequence, motion time domain with more than in the spatial domain, or vice versa. This can be exploited through the implementation of the frame to the analysis, but may also be possible through the use of two or more different ways to the same frame coding and RDO and / or similar RDO select a more appropriate method is used. This preferred method can then result in the data signal represented by, for example, in the preamble to allow the appropriate encoder encodes the frame. Here, for example, one such exemplary method may include direct B-frame model may change, or preclude the use of space by time series prediction or forecast to reach, or P-frames comprises skip mode motion vector selection, the prediction motion vector motion vector prediction parameters or time parameters, or even to another to achieve.

According to one or more characteristics (parameters) applied to the data matching and / or variable constant domain to quantify: the image data compression system, a portion of the data is compressed, such as, for example, block or macroblock It may indeed include additional important information, when compared with other information (data) should be coded differently (for example, a higher priority, higher quality, etc.). One way to achieve is to use different quantization values. For example, in some implementations, a smaller quantizer value may be used for "more important information", and a larger quantizer value may be used for "information less important." However, doing so generally also require the transmission of this information is marked out for each quantized value for each block, macroblock, macroblock group, etc., so that the subsequent decompression / decoding is successful. Unfortunately, these additional information tends to increase the total cost of compression, as well as the complexity of the encoder. Thus, instead of increasing the efficiency, but it may actually reduce the efficiency.

Noting Figure 2 (a-d), each chart describes some exemplary constant domain quantization can be used in certain exemplary image / video coding scheme. In Figure 2 (a-d) each of a graph, the vertical axis (Y-axis) represents the quantized value and form the horizontal axis (X-axis) represents the original value. Figure 2 (a) to illustrate exemplary in constant domain quantization 202, for example, about the constant domain quantization A, with respect to the uniform quantization can considerably improve the coding efficiency. Conventional deadzone quantizer quantizing the entire process is often kept constant or uniform quantization, thus, if possible, incomplete utilization of all the data for redundancies, or otherwise taking into account the importance of information in the data, as well as a block basis / macroblock, through adaptive / modify constant domain coding efficiency can be achieved to quantify the amount of improvement. For example, for example, by adaptive convergence (e.g., along the X axis), but we can do without changing the reconstructed value. For example, the constant domain respectively quantizer 202 and quantizer compares the deadzone 204, 206, and 208 (which are shown in FIG. 2b, c, d in).

Here, the reconstruction of all the possible values of the encoded data is kept constant, unless of course the quantization parameter change emitted signal. For example, by an increase in the level of another box position (FIG. 2 in magic), so that more data will be assigned to here, so, depending on the compression scheme to achieve higher compression.

在某些实现中,通过检查是否已消除色度残留而保持所有的亮度残留完整无缺,该逻辑甚至能将其进一步扩展。也可以将这种概念扩展到宏块的所有可能方式,即检查是否完全不发出残留,或者该方式的色度不会给出较好的性能。然而,这显然会使编码复杂度进一步增加。

参考图9,逻辑900说明输入MB/图像902如何将编码过程块902、904、906和908提供给不同的直接模式(B帧)或拷贝模式(中间帧),来自那些过程块的输出被提供给选择适当输出的基于帧的RDO判决块910。

参考图10,逻辑1000还说明怎样可以还包括方案判决块1002和选择1004,从而根据用户输入、应用程序、系统要求等来提供附加选择性。这里,方案判决块1002选择性地为输入MB/图像提供一个或多个直接模式(B帧)或拷贝模式(中间帧)编码过程块902、904、906和908。选择1004可以通过方案判决块1002或其它过程/逻辑/输入来控制。

B帧的拉格朗日(Lagrangian)乘数的示例性使用RDO技术通常基于拉格朗日乘数的概念。例如,可以选择共同最小化失真和比特率的特殊方式。

这种函数可以表示为下式的最小化: $J(\text{Mode} | QP, X) = \text{SSD}(\text{Mode} | QP) + \lambda \cdot R(\text{Mode} | QP)$ 其中QP是宏块量化器, X是用于模式判决的拉格朗日乘数, Mode标识要被检查并且可能在RDO过程中被选择的宏块模式。

通过示例,在某些实现中,拉格朗日乘数X可以为中间或内部帧而被选择: $QP \cdot \lambda = 0.85 \times 2^3$ 或 $\lambda = 5 \times QP + 5 \times \exp(1, p \cdot 34 - QP)$ 然而在B帧中,在诸如JVT这样的大多数编解码器中,这被选择为 $\lambda = 4 \cdot X$ 。P帧为了给较低开销模式以优先权而作出X的附加加权,这是因为模式特别对于B帧可能具有由已发送的多个运动信息引起的大开销,而诸如直接模式这样的较低开销模式仍能提供就RDO而言的很好性能。

然而,根据按照本发明的某些实验,已经发现加权不应如上述那样不变,相反,它同样应该取决于量化器QP值。

特别是,如果定义了 $\lambda = f(QP)$ 从这些实验中可以发现,可以使用使用固定的 $f(QP) = 4$ 的情况具有好得多的压缩效率的两个 $f(QP)$ 函数: 这种观察也可以通过具有很有用的X的事实而得到,这一事实也可能影响其它重要信息的精确性,譬如,运动预测的精确度,由此会对周围宏块的编码产生负面影响。

使用预测从属的MB模式的自适应类型选择: 序列和帧可能具有不同类型的主导运动相关,这也在未决美国临时专利申请序列号为60/385,965、未决美国临时专利申请序列号为60/376,005、以及未决美国临时专利申请序列号为10/186,284的申请中得以说明。

特别是,对于具有静止背景下的恒定速度的小物体而言,使用来自暂时相邻帧(时域)的运动矢量(MV)使人们能进行较好的预测,并且得到较高的性能。相反,具有平滑运动的较大物体可能在空间域上(相邻宏块)具有较大的相关,而在其它情况下,来自空间域和时域两者的信息可能对预测是重要的。这些类型的相关被部分利用,譬如在B和P帧内的直接模式以及P帧内关于运动矢量预测器的跳越。为了得到更多关于P帧内的运动矢量预测器的跳越信息,可以参见Jani Lainema和Marta Karczewicz撰写的“Skip mode motion compensation”,文档JVT-C027, JVT会议, Fairfax, 2002年5月。这与,如果逻辑在某种程度上用信号指示哪种类型的预测器在不同帧起支配作用,则可以实现相当高的性能。

因此,按照本发明的某些方面,编码逻辑被配置成用信号指示或者在帧、片、或其它类似层次上以某些方式标识出要为从属于预测的模式使用哪一种预测方案。

图8的图表中示出实现这个的一个示例性语法,它作为JVT中的一个示例。这种语法当然能够被修改并且在其它编码设计中不同。

对于图800中的示例性体系, P帧跳跃模式允许三种(3)可能的预测情况。在该例中,存在运动-拷贝预测模式、时间预测跳跃模式、以及零跳跃模式。三种情况的每一个都被分配到(0, 1, 2)中的一个值,它或者可用固定大小的码字来编码,在该例中 $u(2) = 2$ 比特,或者甚至被熵编码(如,用UVLC或CABAC的 $e(v)$)。本领域的技术人员可以理解,也可以使用其它方式来传送这种信息。在图表800中,它说明了JVT内的图像层Rbsp语法,外加B帧的自适应空间/空间-时间考虑(direct_mv_spatial)以及中间帧的自适应拷贝/运动-拷贝跳跃模式(Copy_mv_Spatial)。如果仅使用了这两种模式,则所定义的描述符可能仅占1比特,因此 $u(n)=1$,但如果要使用更多情况(带有零偏的空间预测或平稳时间预测的考虑),则可以分配更多比特(n > 1)或者甚至使用由 $e(v)$ 为该参数定义的熵编码。

对于B帧而言,在该例中,可以为所有的直接模式运动参数将空间运动矢量预测用作一个模式,而将时间预测的参数用作第二模式。也可以考虑/包括未决美国临时专利申请序列号为60/385,965的申请中所描述的其它直接模式。编码逻辑用信号通知在帧或片层次上要使用哪种预测模式。例如,选择可以用基于RDO的方案(如,图9)来执行。在某些实现中,编码逻辑也可以使用明显地由放置在编码器和/或解码器上的特定要求引起的特定模式。在一个特定实例中,考虑到空间预测通常在计算上较简单(如,不需要分区、不需要存储运动矢量,并且独立于定时信息),因此它是某些应用的较佳选择(如,图10)。

Apparently he is, need a decoder represented by the signal quantizer change because heavy construction value still remains the same value. While some might argue that this would undermine the quantization performance in general, this is not necessarily correct, if this process is to select of completion, for example, when certain conditions are met, if the use of new quantitative reaches much higher than suffer compression distortion.

For example, (FIG. 2 (b)), the constant domain quantization only affects slotting another level, with respect to the first / __ slotting, but all the other bits of me still remains the same, as well as the use of reconstruction value does not change. with a constant domain adaptive adaptive estimation scheme is also possible (for example, using rate distortion optimization and adaptive estimation).

According to certain exemplary embodiments of the present invention, the quantization selection logic, therefore, is provided between the compression / encoding system, based on at least one characteristic or parameter quantizer different constant domains (or quantized value) select. For example, according to this logic may be some image characteristics between the different domains of the constant quantizer 202, 204, and 208 select. even more special it is that this logic may be configured to characterize the texture and / or edges in the image data as a representative of "the more important information," and thus encode data in a pattern to provide a high level the quality of.

This logic can be used conventional texture analysis / detection and edge detection algorithm is given to support this decision process.

With this in mind, it is now to Figure 3, which is a flow chart which illustrates the present invention in accordance with certain implementations of different quantization processes to selectively apply the method to the data of 300. In act 302, an initial data set is divided into a plurality of portions of the data. For example, image or video data may be divided into a plurality of blocks, macroblocks, ribbons, sheets, sections and the like. In act 304, a plurality of data derived from the operation portion 302 of the at least one classified in some manner, e.g., based on at least one characteristic or parameter. For example, the classification may be based on edge characteristics, texture characteristics, smooth characteristics, luminance characteristics, color characteristics, color characteristics, noise characteristics, object properties, operating characteristics, user preferences features, user interface focusing characteristic, hierarchical nature, timing characteristics, volume characteristics, frequency, tonal characteristics, timbre characteristics, quality characteristics, bit rate characteristics, data type features, resolution characteristics, encryption features.

... in act 306, with at least two of the plurality of the quantization process to handle the classified portion from act 304 to produce corresponding quantized data. In act 308, a decision of a selection from the operation of 306 quantized data, for example, at least one threshold value based on measurement or classification parameters and actions related to the use of 304. In act 310, the operation data 308 from the quantized allow in some way which is subsequently decoded is encoded.

When using the method 300 may provide a logical analysis of the section and essentially part of the image or video data are considered more important than other parts. Different deadzone quantizer is used for the more important data portion, and analyzes the generated quantized data to determine which of the deadzone quantizer more important to meet the threshold requirement of the desired data. For example, you can implement a mandatory quality or noise threshold requirement.

By way of further example, in the video sequence margin usually play an important role in the motion compensation technology. For example, this block diagram in FIG. 7 obtained by the following detailed description. Basically, the image can be (frame) or part of the image analysis, and the analysis used to determine a plurality of constant domains quantizer which one is applicable, for example, according to the rate-distortion optimization (RDO) criterion.

4, logic 400 illustrates the rapid decision in accordance with the present invention, an exemplary implementation of a process. Here, the input frame / image (or portion thereof) in the block 402 by the analysis. In this example, the image analysis including edge detection and / or texture analysis capabilities, and the output is supplied to the constant-domain decision block 404. After constant domain and then make a decision block 404, using the specified encoder 406 or choose another constant domain quantizer or quantization values when the input frame / image (or portion thereof) encoding.

Figure 5 logic 500 may be configured to support a method 300, rather than making constant domain quantization on the fast decision, in which

在其它实现中,其中这些问题不是问题,该组合可能导致进一步改进的编码性能。图11的示意图中示出已编码序列的一个实例。这里,P和B帧与场景变化一起示出。为相应的跳跃(Skip)或直接(Direct)模式宏块示出不同信号表示的P和B帧的箭头。还注意到,信号表示可以指示编码逻辑应怎样执行用于运动矢量编码的运动矢量预测,或者其它模式的预测(例如,未决美国临时专利申请序列号为60/376005和未决美国专利申请序列号为10/186284的申请中描述的直接(Direct)P)。

如图11所示,不同的帧为其相应的直接(B)和跳跃(P)模式用信号表示不同类型的预测。Pz、PT、和Pm为实例零定义了时间和空间的(运动一拷贝)预测,87和Bsp为直接模式定义了时间和空间的预测。

时间标记独立的直接模式,考虑到静态时间/空间块:不同类型的预测,尤其对应B帧内的直接模式而言,可能更适用于不同类型的运动和序列。仅使用时间或空间预测在某些情况下可能提供可接受的性能,但在其它情况下,性能可能非常差。上述段落中描述的解决方案,或者未决美国临时专利申请序列号为60/385965中给出的解决方案可能提供更好的性能。

通过示例,给出了一种额外的情况,它看上去很有效并且组合了时间和空间预测两者的性能,而简单地通过不要求分割而简单地保持空间预测符,并且/或者是时间独立的。

在特定实现中,空间预测可能比时间预测更有益(如,由于其属性)。因此,例如,空间预测被用作直接模式的主要预测。当来自时间预测符的运动信息和参考帧为零时,产生一种可能的例外。在这种情况下,相应直接模式块的运动信息和参考帧也被视为零。而且,空间预测通过也考虑到空间零的偏置性和/或平稳子分区的空间零而被改进。由此,如果当前预测块的任何或某些相邻宏块或块具有零运动(或非常接近(例如,整数运动矢量为零))和参考帧,那么整个宏块,或其一部分也被视为具有零运动。这两个概念都有助于保护平稳背景,如果不引入这种条件,它尤其在运动对象的边缘引起相当的失真。

图12中示出流程图1200,它被简化为16x16宏块的情况。注意到这种为直接模式使用空间预测的概念也可被扩展到更小的块(例如,8x8或4x4)(或较大的块或其它形状的部分)。在动作1202中,空间预测符MVA、MVB和MVE以及时间预测符MVT被提供给动作1204,其中MVDI被设为Median(MVA, MVB, MVE)(取中值)。在动作1206中,基于11、作出判决,或导致动作1208,其中MVDI被设为零,或导致附加判决的动作1210。作为动作1210的结果,MVDI或在动作1208中被设为零或保持不变,“1”被输出。

在该说明中,已经给出了与用于编码图像、视频序列或其它类型的数据的码率失真优化(Rate Distortion Optimization)有关的若干概念。已说明了附加语法,它可在视频序列内被采用,通过允许若干可选预测情况而允许相当高的压缩效率,尤其对于如分别在P和B帧内的跳跃或直接模式而言,它们可在图像开始时被信号通知到。已经给出了B帧的高效时间标记独立的直接模式,它考虑到空间运动矢量预测以及平衡时间预测符。为了大大改进各种图像/视频/数据编码系统的性能,可以实现上述方法和装置的全部或一部分。

尽管本发明各种方法和装置的某些优选实现已经在附图中被说明并且在上述详细说明中被描述,然而可以理解,本发明并不限于这里所揭示的示例性实施例,而能作出许多重排、修改和替代,并且不背离本发明的精神。

a plurality of constant domain quantization and according to a certain threshold criteria to make RDO decision. Here, the input macro block (MB) / 502 images (or other similar parts) is (optionally) provide different constant domain coding blocks 504, 506, 508, and / or 510. from these different constant domain coding block output RDO decision block 512 is analyzed, the selected encoded data is output. As described in this example, some or all of the constant domain coding block / process may occur in parallel. In other implementations, this process may be arranged to occur serial.

FIG. 6 (a) and 6 (b) which illustrate the use of selective re-encoding decision process logic. For example, in Figure 6 (a), the input MB / image 502 is supplied to the constant domain coding blocks 602 and 604. Re-encoding the output from the decision block 606 to consider the constant domain coding block 602 to select the output and impact of the constant domain coding blocks 602 and 604 among 608. Recoding decision block 608 may be selectively activated constant domain coding block 604. In FIG. 6 (b), the logic 610 is similar to 600, in addition to re-encode the relevant decision block 606 to select 608 includes RDO decision block 612 for analyzing the output from the constant domain coding blocks 602 and 604, and select which one output.

Figure 6 (a, b) the re-encoding block can be used to make a judgment decision based on various criteria. For example, in some implementations, consider the quality restrictions, rate limiting and / or other concepts RDO / threshold.

Only when certain conditions are met the previously defined, such as rate / quality exceeds a certain threshold before they can make use of additional deadzone quantizer. In certain exemplary experiment successfully selected deadzone quantizer, it has 30 percent more than the original larger constant domain. Additionally, you may in accordance with such frequency or edge-type image characteristics such as AC, you can also use an adaptive deadzone quantizer.

In Figure 7, an exemplary encoding logic 700 illustratively describes how The image analysis process 704 to process the initial image 702 (or a portion thereof) to generate this example, edge detection data 706. Then, the edge detection data 706 is provided to a block 708 to generate block classification process classified data 710. Then, the block classification of the original image data 710 and 702-- play is supplied to the image coding process 712, the process then generates the code image 714. This is an example of a simple encoding process image. Here, the image is analyzed (e.g., using an edge detection algorithm) and then blocks are classified according to such information. Substantially defines the N-dimensional matrix (DZp) (N depending on a defined number of constant domains), which matrix is subsequently dispensed at a quantizer (QP) assigned at the macroblock position (i, j) appropriate constant domain (DZ). no residual mode considerations (such as direct mode for B-frames) - even though the above description may be mainly for smooth image, but the same concept can also be applied in the video.

For example, in video compression to fairly straightforward manner using the image pre-analysis and constant domain quantizer.

A very interesting and can greatly benefit from the use of the above concepts is that intra-B direct mode.

Direct Mode is basically a special way without having to transmit motion parameters, because it can directly or through space or time on the prediction by being directly predicted. Additional information on the direct mode can Application Serial No. 60 / 385,965 application found in U.S. Provisional Patent pending.

If there is no residue to be transmitted, it is possible according to some further aspects of the invention, by using the direct mode is called a non-residual special way to further improve the efficiency. As described below, the direct mode can no residue run-length encoding (RLC) strategy. Here, for example, if the resulting distortion is small enough and since the higher bit rate reduction of the efficiency of the RLC caused large enough, the constant domain quantizer can provide the desired solution type. The basic idea is to selectively cause or force the direct mode without residue being encoded, despite the presence of residues under certain conditions. For example, as shown in Figure 9, it can be achieved based, in which no residue of the direct mode as compared to the same concept of Figure 5 and all other active modes, are also checked during RDO. In some cases, due to the RDO it is not appropriate to be used for these cases, so the performance of this scheme may not be desirable in that it is good. Other processes depend on the quantization value is also affected, for example, in the elimination of the inner filter (not shown) that artificial image block. More specifically, even if the performance is displayed in the lower bit rate is good, but performance may be substantially impaired in the larger bit rates; by

using a larger quantizer no residue and does not consider the direct mode, the performance may be better. and an image similar to the example made, in accordance with certain aspects of the present invention, the logic may be configured to only some previously defined conditions are satisfied when, especially when associated with a residual direct mode is not important, only consider these ways. This logic may, for example, is configured to direct mode by checking the coded block pattern (CBP) and the importance of the residue is used as an estimation condition. CBP is below a certain threshold without considering the chrominance information, you can consider this, it means that the residue is not important, and if it does not cause too much skipped distortion. Moreover, there is no other image features such as edges and textures that can also be used during video encoding / logic.

In some implementations, only by checking whether the elimination of residual color and brightness keep all residual intact, the logic of this can even be further expanded. You can also extend this concept to all the possible ways a macro block that checks whether the residual issue completely, or the color mode will not give better performance. However, this would obviously further increases the complexity of the coding.

Referring to Figure 9, logic 900 illustrates an input MB / image 902 how to process the coded blocks 902,904,906 and 908 provided to different Direct Mode (B frames) or Copy Mode (intermediate frame), those derived from the process block output is supplied to select the appropriate output RDO frame-based decision block 910.

10, the logic 1000 also shows how the program may also include the decision block 1002 and select 1004, so based on user input, application, system requirements, and to provide additional selectivity. Here, the program decision block 1002 selectively input MB / image to one or more direct mode (B-frame) or copy mode (intermediate frame) encoding process blocks 902,904,906 and 908. 1004 may select programs by decision block 1002 or other process / logic / input to control.

B frames Lagrange (Lagrangian) multiplier exemplary use RDO technology is often based on Lagrange multiplier concept. For example, you can select common minimize distortion and bit rate of the special way.

Such a function can be represented by the following formula is minimized: $J(\text{Mode IQP}, X) = \text{SSD}(\text{Mode} | \text{QP}) \times R(\text{Mode} | \text{QP})$ where QP is the macroblock quantizer, X is a mode decision Lagrange multiplier, mode identity to be examined and possibly in the process of being selected RDO macroblock mode.

By way of example, in some implementations, the Lagrangian multiplier X can be internal or intermediate frame is selected: $\text{QP} \times 1/p = 0.85 \times 2^3$ or $X/p = 5 \times \text{QP} \times \exp(-1/p)$, $p = 34 - \text{QP}$ However B-frames, such as in the majority of codecs JVT, which is selected to seven = 4 X righteous., P o to give priority to low overhead mode and make the additional weight of X, it is because mode especially for B-frame may have a plurality of moving large overhead information sent by the cause, and as such a direct mode mode can still provide a lower cost on a good performance in terms of RDO.

However, according to the present invention in accordance with certain experiments, it has been found that the weighting should not be constant as described above, the contrary, it should also depend on the quantizer QP value.

In particular, if you define; $-1/(QPh)$ and w from these experiments can be found, can be used than using a fixed $f(QP) = 4$ case has a much better compression efficiency of the two $f(QP)$ function: this can also be observed by the fact that with high X obtained, the fact may also affect the accuracy of other important information, for example, the accuracy of prediction of motion, whereby on the surrounding macroblock coding have a negative impact.

MB model using the predicted dependent adaptive type selection: sequences and frames may have different types of leading sports-related, which is also pending U.S. Provisional Patent Application Serial No. 60 / 385,965, pending U.S. Provisional Patent Application Serial No. 60/376, 005, as well as pending U.S. provisional Patent application serial No. 10/186, 284 application to be described. particular, for small objects with constant speed in a stationary background, the use of from temporally neighboring frames (time domain) of the motion vector (MV) allows one to perform better prediction, and get a higher performance. In contrast, larger objects having smooth motion may be in the spatial domain (adjacent macroblocks) having a large correlation, but in other cases, the information from both a spatial and temporal prediction is likely to be important. Related partially from using these types, such as' direct mode in B and P frames and a P frame on the jump motion vector

predictor. To get more information about P skip frame motion vector predictor can see Jani Lainema and Marta Karczewicz written "Skip mode motion compensation", the document JVT-C027, JVT meeting, Fairfax, 2002 7-May. This is, to some extent, if the logic signal indicating which type of predictor play the dominant role in different frames, you can achieve very high performance.

Thus, in accordance with certain aspects, the encoding logic of the present invention is configured to signal or indicate the frame, the sheet, or in some other manner similar hierarchy identify which one to use for the prediction of the prediction mode subordinate Program.

Chart

FIG. 8 shows an implementation of an exemplary syntax of this, as one example of the JVT. Of course, this syntax can be modified in various other coding and design.

For the exemplary system of FIG. 800, P frame skipping mode allows three (3) possible projections. In this embodiment, there is a copy of the motion prediction mode, temporal prediction skip mode and the zero skip mode. Each of the three cases are assigned to {0,1,2} is a value that can be used fixed or encoded codeword size, (2) = 2 bits, or entropy coding even in this case u (e.g., using UVLC or CABAC is e (v)). Those skilled in the art will appreciate that such information can also be transmitted through other methods. 800 in the chart, which shows the image layer RBSP syntax JVT within plus adaptive B-frame space / space for a time considered (direct_mv_spatial) and adaptive copy intermediate frames / copy a motion skip mode (COpy_mv_Spatial). If you only consider the use of these two modes is defined descriptors may occupy only one bit, so u (n = 1), but if you want to use more often (space with zero bias forecast or prediction of stationary time), you can allocate more bits (n and gt; 1), or even use the entropy coding by the e (v) for parameter definitions.

For the B-frame, in this case, for all the direct mode motion parameter space is used as a motion vector prediction mode, and the parameters used for the second time prediction mode. You can also consider / include pending U.S. Provisional Patent Application Serial No. Other Direct mode 60 / 385,965 application described. Coded logic signaled at the frame or sheet level to which the prediction mode. For example, the selection may be based on the use of RDO scheme (eg, FIG. 9) is performed. In some implementations, the encoding logic may also be used by a specific pattern clearly specific requirements placed on the encoder and / or decoder caused. In one particular example, taking into account the spatial prediction is usually computationally simpler (eg, no partition, no need to store the motion vector, and independent of the timing information), so it is a better choice for some applications (eg, FIG. 10).

In other implementations, where these problems are not a problem, this combination may lead to further improved coding performance. Schematic of Fig. 11 shows an example of the coded sequence. Here, P and B frames are shown together with the scene changes. Corresponding jump (Skip) or direct (Direct) mode macroblock shown by the arrow P and B frames of different signal representation. Also noted that the signal may indicate the encoding logic should indicate how the implementation of the motion vector used for motion vector prediction encoding, or the prediction of other modes (for example, pending U.S. Provisional Patent Application Serial No. 60/376005 and pending U.S. Patent Application Serial No. direct (direct) 10/186284 described in the application of P).

11, different frames directly to its corresponding (B) and jumping (P) mode signal indicates the prediction is not the same type. Pz, PT, and Pm instance zero defines the time and space (a copy of the motion) predicts 87 and Bsp direct mode defines the predicted time and space. time stamp independent Direct Mode, taking into account the static temporal / spatial blocks: Different types of prediction, especially the direct mode corresponding to the B frame, it may be more suitable for different types of motion and sequences. Temporal or spatial prediction using only may provide acceptable performance in some cases, but in other cases, the performance may be very poor. Solution described in the preceding paragraphs, or pending U.S. Provisional Patent Application Serial No. 60/385965 given solutions may offer better performance is.

By way of example, gives an additional case, it looks very effective and the combination of both time and space predictor of performance, but simply by not requiring segmentation and simply holding space predictor, and / or is time independent.

In particular implementations, spatial prediction may be more useful than the prediction time (eg, due to its nature). Thus, for example, spatial

prediction is used as the main prediction in the direct mode. When the movement of information from time predictor and the reference frame is zero, resulting in a possible exception. In this case, the reference frame and motion information corresponding direct mode block group is considered zero. Moreover, by also taking into account the spatial prediction space zero offset and / or sub-partition space smooth zero is improved. Thus, if any or some of the current prediction macroblocks or adjacent blocks having zero motion block (or very close (e.g. integer motion vector is zero)) and the reference frame, then the entire macroblock, or a portion thereof also depends as having zero motion. These two concepts are the background to help protect smoothly, without introducing such conditions, it is especially attracted considerable distortion at the edges of the moving object.

FIG. 12 shows a flowchart 1200, which is reduced to the case of the 16x16 macroblock. Noting that the use of spatial prediction for the direct mode concept can be extended to smaller blocks (e.g., 8x8 or 4x4) (or a portion of a larger block or other shapes). In act 1202, the spatial predictor MVA, MVB and MVE as well as a predictor MVT time is provided to action 1204, where MVDi "et is set to Median (MVA, MVB, MV) (taking the median). In act 1206, 11, based on the judgment, or cause action 1208, which MVDi_t is set to zero, or cause of action for an additional judgment 1210. as a result of the operation 1210, MVDi "et or is set to zero or remain unchanged in act 1208, \ 1 □" (: 1 output

In this description has been presented with a number of concepts used to encode the image, bit-rate video sequences or other types of data distortion optimization (Rate Distortion Optimization) related. Additional syntax has been described, it may be employed in the video sequence, by allowing several alternative prediction cases allowing very high compression efficiency, especially for the P and B, respectively, as the frame skip or direct mode, they are available when the image signal is notified to start ?. B has been given an efficient time frame marker independent direct model, which takes into account spatial motion vector prediction and balancing time predictor. In order to greatly improve various image / video / performance data coding system, may be implemented all or part of the above-described method and apparatus.

Although certain preferred implementation of the present invention, various methods and devices have been described in the drawings and described in the foregoing detailed description, it is to be understood that the present invention is not limited to the exemplary embodiments disclosed herein, and be able to make many rearrangements, modifications and substitutions without departing from the spirit of the invention.

Cited by: CN104486503 A ; CN105120285 A ; US8594187 BB ; US9083378 BB ; WO09121233 A1 ; WO14205619 A1 ;

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.
H04N 7/24 (2006.01)



[12] 发明专利申请公布说明书

[21] 申请号 200780015871.9

[43] 公开日 2009 年 5 月 20 日

[11] 公开号 CN 101438591A

[22] 申请日 2007.5.4
[21] 申请号 200780015871.9
[30] 优先权
[32] 2006.5.5 [33] US [31] 11/418,690
[86] 国际申请 PCT/US2007/010848 2007.5.4
[87] 国际公布 WO2007/130580 英 2007.11.15
[85] 进入国家阶段日期 2008.11.3
[71] 申请人 微软公司
地址 美国华盛顿州
[72] 发明人 C·屠 S·斯里尼瓦桑

[74] 专利代理机构 上海专利商标事务所有限公司
代理人 顾嘉运

权利要求书 3 页 说明书 15 页 附图 8 页

[54] 发明名称
灵活量化

[57] 摘要
一种数字媒体编码器/解码器使用提供沿着所编码的数字媒体数据的各个维度改变量化的功能的灵活量化技术，这些维度包括空间、频率子带和色通道。该编解码器利用一信号表示方案来高效地用信号表示灵活量化组合的各种排列以用于主要使用情形。当量化器选择可用时，该编解码器通过定义一量化器子集并从该集合中索引当前量化器来高效地编码当前量化器。

700

FRAME () {	位数
...	许多
DC_FRAME_UNIFORM	1
if (DC_FRAME_UNIFORM)	
DC_QUANTIZER	可变
if(未跳过低通){	
USE_DC_QUANTIZER	1
if(USE_DC_QUANTIZER == FALSE){	
LP_FRAME_UNIFORM	1
if(LP_FRAME_UNIFORM){	
LP_QUANTIZERS = 1	
LP_QUANTIZER()	可变
}	
}	
if(未跳过高通){	
USE_LP_QUANTIZER	1
if(USE_LP_QUANTIZER == FALSE){	
HP_FRAME_UNIFORM	1
if(HP_FRAME_UNIFORM){	
HP_QUANTIZERS = 1	
HP_QUANTIZER()	可变
}	
}	
}	

1. 一种编码/解码数字媒体的方法，包括：
沿着多于一个维度划分所述数字媒体；
跨所述数字媒体的多于一个划分的维度改变应用于所述数字媒体的块的量化；以及
编码/解码所述数字媒体的量化的块。
2. 如权利要求 1 所述的方法，其特征在于，所述改变量化包括跨空间维度改变量化。
3. 如权利要求 1 所述的方法，其特征在于，所述改变量化包括跨频率维度改变量化。
4. 如权利要求 1 所述的方法，其特征在于，所述改变量化包括跨通道改变量化。
5. 如权利要求 1 所述的方法，其特征在于，所述通道是色通道。
6. 如权利要求 1 所述的方法，其特征在于，所述改变量化包括跨空间、频率子带和通道维度中的多于一个维度改变量化。
7. 如权利要求 1 所述的方法，其特征在于，所述改变量化包括：对于一个划分的维度，
选择是否在所述划分的维度的第一分区中均匀地量化数字媒体；
如果选择不均匀地量化所述第一分区，则选择是否在所述划分的维度的第二分区中均匀地量化数字媒体；以及
如果选择不均匀地量化所述第二分区，则跨所述维度独立地量化所述数字媒体。
8. 如权利要求 7 所述的方法，其特征在于，所述第一分区是帧，所述第二分区是瓦块，而所述独立量化包括用不同的量化器独立地量化宏块内的所述数字媒体的块。
9. 如权利要求 7 所述的方法，其特征在于，所述第一分区是 DC 频率子带，所述第二分区是低通频率子带，而所述独立量化包括用不同的量化器独立地量化所述低通频率子带或高通频率子带内的所述数字媒体的块。

10. 如权利要求 7 所述的方法，其特征在于，所述第一分区是色度通道，所述第二分区是亮度通道，而所述独立量化包括用不同的量化器来独立地量化所述通道内的所述数字媒体的块。

11. 如权利要求 7 所述的方法，其特征在于，还包括当跨所述维度独立地量化所述数字媒体时，指定可用于跨所述维度应用的一组不同的量化器，并从所述集合索引所述量化器的选择以应用于所述数字媒体的块。

12. 一种数字媒体处理设备，包括：

用于储存数字媒体数据的存储器；以及

被编程为通过应用向所述数字媒体数据的块应用变换的前向或反向操作，并向所述数字媒体数据块应用量化器，来编码和/或解码所述数字媒体数据的数字媒体处理器，其中所述量化器可选择性地跨所述数字媒体数据的多于一个维度来变化。

13. 如权利要求 12 所述的设备，其特征在于，所述维度包括空间、频率和色通道中的至少两个。

14. 至少一种其上储存有用于使得数字媒体处理设备执行一种方法的程序代码的程序存储载体，所述方法根据编解码器来处理数字媒体数据，所述方法包括：

选择要应用于所述数字媒体数据的量化模式的组合，所述量化模式跨空间、频率子带和色通道维度中的多于一个维度而变化；

向所述数字媒体数据应用块变换；

根据所选择的量化模式的组合量化所述数字媒体；以及

在压缩的数据流中编码所述量化的数字媒体数据；以及

在所述压缩的数据流中用信号表示所选择的量化模式的组合。

15. 如权利要求 14 所述的至少一种程序存储载体，其特征在于，可用于跨空间维度改变量化的所述量化模式包括用于对所述数字媒体数据的帧进行均匀量化的模式、用于对所述数字媒体数据的瓦块进行均匀量化的模式、以及用于独立地量化所述数字媒体数据的每一宏块的模式。

16. 如权利要求 14 所述的至少一种程序存储载体，其特征在于，可用于跨频率子带维度改变量化的所述量化模式包括用于跨所述数字媒体数据的频

率子带进行均匀量化的模式、用于跨所述数字媒体数据的频率子带进行部分均匀量化的模式、以及用于独立地量化所述数字媒体数据的频率子带的模式。

17. 如权利要求 16 所述的至少一种程序存储载体，其特征在于，所述用于跨频率子带进行部分均匀量化的模式包括一个频率子带使用另一频率子带的量化器。

18. 如权利要求 14 所述的至少一种程序存储载体，其特征在于，可用于跨色通道维度改变量化的所述量化模式包括用于跨所述数字媒体数据的色通道进行均匀量化的模式、用于跨所述数字媒体数据的色通道进行部分均匀量化的模式、以及用于独立地量化所述数字媒体数据的色通道的模式。

19. 如权利要求 14 所述的至少一种程序存储载体，其特征在于，所述方法还包括当用于一维度中的数字媒体数据的分区的模式可变时，通过定义一可用量化器集合并从所述集合中索引来指定当前量化器。

20. 至少一种其上储存有用于使得数字媒体处理设备执行一种方法的程序代码的程序存储载体，所述方法根据编解码器处理数字媒体数据，所述方法包括解码根据如权利要求 14 所述的方法编码的压缩的数据流。

灵活量化

背景

1. 基于块变换的编码

变换编码是在许多音频、图像和视频压缩系统中使用的一种压缩技术。未压缩数字图像和视频通常被表示或捕捉为以二维(2D)网格排列的图像或视频帧中各位置处的图元或色彩的样本。这被称为图像或视频的空间域表示。例如,用于图像的典型格式由被排列为网格的 24 位彩色图元流构成。每一样本是表示诸如 RGB 或 YIQ 等色彩空间内该网格中的一个像素位置处的色彩分量的数字。各种图像和视频系统可使用各种不同的色彩、空间和时间分辨率的采样。类似地,数字音频通常被表示为时间采样的音频信号流。例如,典型的音频格式由在有规律的时间间隔处所取的 16 位音频信号幅度样本流构成。

未压缩数字音频、图像和视频信号可消耗大量的存储和传输能力。变换编码通过将信号的空间域表示变换成频域(或其它类似的变换域)表示,然后降低该变换域表示的某些一般较不可感知的频率分量的分辨率,减小了数字音频、图像和视频的大小。这一般与降低空间域中的图像或视频或时域中的音频的色彩或空间分辨率相比,产生了较不可感知的数字信号劣化。

更具体而言,图 1 所示的典型的基于块变换的编解码器 100 将未压缩的数字图像的像素划分成固定大小的二维块($X_1, \dots, X_n$),每一块可能与其它块重叠。对每一块应用进行空间-频率分析的线性变换 120-121,这将块内彼此隔开的样本转换成一般表示块间隔上相应的频带内的数字信号的强度的一组频率(或变换)系数。为了压缩,变换系数可被选择性地量化 130(即,诸如通过丢弃系数的最低有效位或将较高分辨率数字集中的值映射到较低分辨率来降低分辨率),并且还熵或可变长度编码 130 成压缩的数据流。在解码时,变换系数进行反变换 170-171 以便几乎重构原始的色彩/空间采样图像/视频信号(重构块 $\hat{X}_1, \dots, \hat{X}_n$)。

块变换 120-121 可被定义为对大小为 N 的向量 x 的数学运算。最通常的是,

该运算是线性乘法，从而产生变换域输出 $y = Mx$ ， M 是变换矩阵。当输入数据是任意长时，它被分段成大小为 N 的向量，并且向每一段应用块变换。出于数据压缩的目的，选择可逆块变换。换言之，矩阵 M 是可逆的。在多个维度中（例如，对于图像和视频），块变换通常被实现为可分操作。沿数据的每一维（即，行和列）可分地应用矩阵乘法。

对于压缩，变换系数（向量 y 的分量）可被选择性地量化（即，诸如通过丢弃系数值的最低有效位或将较高分辨率数字集中的值映射到较低分辨率来降低分辨率），并还可被熵或可变长度编码为压缩的数据流。

在解码器 150 中解码时，如图 1 所示，在解码器 150 侧应用这些操作的反过程（反量化/熵解码 160 和反块变换 170-171）。在重构数据时，将逆矩阵 M^{-1} （反变换 170-171）作为乘数应用于变换域数据。当应用于变换域数据时，反变换几乎重构原始时域或空间域数字媒体。

在许多基于块变换达到编码应用中，变换理想地是可逆的以取决于量化因子同时支持有损和无损压缩两者。如果例如没有量化（一般被表示为量化因子 1），则利用可逆变换的编解码器可在解码时精确地再现输入数据。然而，这些应用中的可逆性的要求约束了对用于设计编解码器的变换的选择。

诸如 MPEG 和 Windows Media 等许多图像和视频压缩系统利用基于离散余弦变换（DCT）的变换。已知 DCT 具有得到近乎最优的数据压缩的良好能量压缩特性。在这些压缩系统中，在压缩系统的编码器和解码器两者中的重构环路中采用了反 DCT（IDCT）来重构各个图像块。

2. 量化

根据一个可能的定义，量化是用于通常用于有损压缩的近似不可逆映射函数的术语，其中有一组指定的可能输出值，并且该组可能的输出值中的每一成员具有导致对该特定输出值的选择的一组相关联的输入值。已经开发了各种量化技术，包括标量或矢量、均匀或非均匀、有或没有死区、以及自适应或非自适应量化。

量化操作本质上是按照量化参数 QP 的加偏除法（biased division），这在编码器处执行。反量化或乘法操作是与 QP 的乘法，这在解码器处执行。这些过程共同引入了原始变换系数数据的丢失，这表现为解码的图像中的压缩误差

或伪像。在最简单的编解码器中，可向一帧中的所有变换系数应用特定的固定QP值。尽管这在某些情况下可能是可接受的解决方案，但是它也有几个缺点：

人类视觉系统对于所有频率，或对一帧内的所有空间位置，或对所有亮度和色度通道并不都是同样敏感的。对不同系数使用不同QP值提供了甚至具有相同或更少数量的压缩比特的视觉上更出众的编码。同样，也可适当地优化其它误差度量。

码率控制或编码器产生期望大小的压缩文件的能力并不容易在整个帧上用单个QP来执行。

因此，期望的是允许编码器以任意的方式跨图像改变QP。然而，这意味着用于每一数据分区的实际QP值应在比特流中用信号表示。这导致巨大的开销，仅仅是为了携带QP信号表示信息，从而导致其在实践中是不合适的。所需的是用信号通知QP的灵活但节省比特的的手段，尤其是对于经常遇到的情形。

概括而言，量化是大多数图像和视频编解码器控制压缩的图像质量和压缩比的主要机制。大多数流行的编解码器支持的量化方法提供了极少的特征或极少的灵活性，或者导致附加比特的重大开销。通常，图像或视频帧通常被均匀地量化，或以在空间位置上改变量化的有限能力来量化。这一灵活性的缺乏损害了压缩质量，并阻止了进行中的准确的码率控制。另一方面，某些编解码器在支持量化方法方面提供几乎无限制的的自由。编码来用信号表示不同量化器的使用在已编码媒体中花费了额外的比特，并且本身会不利地影响压缩效率。此外，构建一致解码器的过程需要量化器方法的所有可能组合生成的大量测试遍，这可能是费力的。

概述

以下详细描述呈现了提供沿着已编码的数字媒体数据的各维改变量化的能力的灵活量化技术的各种变型。例如，该灵活量化技术的一种代表性实现可以在三个维度上改变量化—(i)空间位置，(ii)频率子带，以及(iii)色通道上。该详细描述还呈现了在已编码数字媒体数据中高效地用信号表示该灵活量化的方式。该量化方法的好处是因涉及量化的辅助信息而引入的开销对于主要使用情形被最小化，同时如果编码器需要则允许最大的灵活性。

提供本概述以便以简化的形式介绍将在以下详细描述中进一步描述的一些概念。该概述不旨在标识所要求保护的主题的关键特征或本质特征，也不旨在用于帮助确定所要求保护的主题的范围。本发明的其它特征和优点在参考附图继续阅读以下对实施例的详细描述后将变得显而易见。

附图简述

图 1 是现有技术中常规的基于块变换的编解码器的框图。

图 2 是包含块模式编码的代表性编码器的流程图。

图 3 是包含块模式编码的代表性解码器的流程图。

图 4 是根据灵活量化技术的包含用于用信号表示 DC 量化器的伪代码定义的表。

图 5 是根据灵活量化技术的包含用于用信号表示低通量化器的伪代码定义的表。

图 6 是根据灵活量化技术的包含用于用信号表示高通量化器的伪代码定义的表。

图 7 是根据灵活量化技术的包含用于用信号表示帧层处的量化器的伪代码定义的表。

图 8 是根据灵活量化技术的包含用于用信号表示瓦块层处的量化器的伪代码定义的表。

图 9 是根据灵活量化技术的包含用于用信号表示瓦块层处的 DC 子带的量化器的伪代码定义的表。

图 10 是根据灵活量化技术的包含用于用信号表示瓦块层处的低通子带的量化器的伪代码定义的表。

图 11 是根据灵活量化技术的包含用于用信号表示瓦块层处的高通子带的量化器的伪代码定义的表。

图 12 是根据灵活量化技术的包含用于用信号表示宏块层处的量化器的伪代码定义的表。

图 13 是根据灵活量化技术的包含用于在频率模式中用信号表示宏块层处的低通量化器的伪代码定义的表。

图 14 是根据灵活量化技术的包含用于在频率模式中用信号表示宏块层处的高通量化器的伪代码定义的表。

图 15 是用于实现具有灵活量化的媒体编码器/解码器的合适的计算环境的框图。

详细描述

以下详细描述涉及提供能够改变沿着空间、频率和色彩维度应用的量化的高效灵活量化（此处称为“灵活量化”）的编码和解码技术。以下描述在数字媒体压缩系统或编解码器的上下文中描述了该技术的一个示例实现。该数字媒体系统以压缩形式对数字媒体数据进行编码以便传输或存储，并解码该数据以供回放或其它处理。出于说明的目的，包含这一灵活量化的该示例性压缩系统是图像或视频压缩系统。或者，该技术也可被结合到用于其它 2D 数据的压缩系统或编解码器中。灵活量化技术不要求数字媒体压缩系统以特定的编码格式来编码压缩的数字媒体数据。

1. 编码器/解码器

图 2 和 3 是在代表性 2 维（2D）数据编码器 200 和解码器 300 中采用的过程的一般化图示。该图呈现了结合了实现块模式编码的 2D 数据编码器和解码器的压缩系统的一般化或简化的图示。在使用块模式编码的替换压缩系统中，可使用比本代表性编码器和解码器中所示的更多或更少的过程来进行 2D 数据压缩。例如，某些编码器/解码器还可包括色彩转换、色彩格式、可缩放编码、无损编码、宏块模式等等。取决于量化，压缩系统（编码器和解码器）可提供 2D 数据的无损和/或有损压缩，量化可基于从无损到有损变化的量化参数。

2D 数据编码器 200 产生压缩比特流 220，它是作为输入提供给编码器的 2D 数据 210 的更紧凑表示（对于典型输入）。例如，2D 数据输入可以是图像、视频序列的一帧、或具有两个维度的其它数据。2D 数据编码器将输入数据帧划分成块（一般在图 2 中示为分区 230），这在所示的实现中是形成块该帧的平面的规则图案的非重叠 4x4 像素块。这些块被分组成称为宏块的群集，在该代表性编码器中其大小是 16x16 像素。宏块进而被分组成称为瓦块的规则结构。

瓦块也可形成图像上的规则图案，使得水平行中的瓦块是统一的高度且是对齐的，而垂直列中的瓦块是统一的宽度且是对齐的。在该代表性编码器中，瓦块可以是任意大小，该大小在水平和/或垂直方向上是 16 的倍数。替换编码器实现可以将图像划分成块、宏块、瓦块或其它大小和结构的其它单元。

对块之间的每一边缘应用“前向重叠”算子 240，之后使用块变换 250 来变换每一 4×4 的块。该块变换 250 可以是由 Srinivasan 在 2004 年 12 月 17 日提交的题为“Reversible Transform For Lossy And Lossless 2-D Data Compression”（用于有损和无损 2D 数据压缩的可逆变换）的美国专利申请第 11/015,707 号中所描述的可逆的、无缩放的 2D 变换。重叠算子 240 可以是由 Tu 等人在 2004 年 12 月 17 日提交的题为“Reversible Overlap Operator for Efficient Lossless Data Compression”（用于高效无损数据压缩的可逆重叠算子）的美国专利申请第 11/015,148 号；以及 Tu 等人在 2005 年 1 月 14 日提交的题为“Reversible 2-Dimensional Pre-/Post-Filter for Lapped Biorthogonal Transform”（用于重叠双正交变换的可逆 2 维预/后滤波器）的美国专利申请第 11/035,991 号中描述的可逆重叠算子。或者，可使用离散余弦变换或其它块变换和重叠算子。在变换之后，令每一 4×4 的变换块的 DC 系数 260 经受一类似的处理链（块化、前向重叠、之后是 4×4 的块变换）。所得的 DC 变换系数和 AC 变换系数被量化 270、熵编码 280 和分组化 290。

解码器执行反过程。在解码器侧，从其各自的分组中提取 310 变换系数位，从中系数本身被解码 320 和解量化 330。DC 系数 340 通过应用反变换来重新生成，并且 DC 系数的平面使用跨 DC 块边缘应用的合适的平滑算子来“反重叠”。随后，通过向 DC 系数应用 4×4 的反变换 350 来重新生成整个数据，并从比特流中解码 AC 系数 342。最后，对所得图像平面中的块边缘进行反重叠滤波 360。这产生经重构的 2D 数据输出。

在一个示例性实现中，编码器 200（图 2）将输入图像压缩成压缩比特流 220（例如，文件），而解码器 300（图 3）基于采用无损还是有损编码来重构原始输入或其近似。编码过程涉及应用以下所讨论的前向重叠变换（LT），这是用同样在以下更全面描述的可逆 2 维预/后滤波来实现的。解码过程涉及应用使用可逆 2 维预/后滤波的反重叠变换（ILT）。

所示的 LT 和 ILT 在确切的意义上是彼此的逆，并且因此可被统称为可逆重叠变换。作为一种可逆变换，LT/ILT 对可用于无损图像压缩。

由所示的编码器 200/解码器 300 压缩的输入数据 210 可以是各种色彩格式（例如，RGB/YUV 4:4:4、YUV 4:2:2 或 YUV 4:2:0 彩色图像格式）的图像。通常，输入图像总是具有亮度（Y）分量。如果它是 RGB/YUV 4:4:4、YUV 4:2:2 或 YUV 4:2:0 图像，则该图像还具有色度分量，诸如 U 分量和 V 分量。图像的这些单独的色彩平面或分量可具有不同的空间分辨率。在例如 YUV 4:2:0 色彩格式的输入图像的情况下，U 和 V 分量具有 Y 分量一半的宽度和高度。

如上所述，编码器 200 将输入图像或图片块化成宏块。在一个示例性实现中，编码器 200 将输入图像块化成 Y 通道中的 16×16 像素区域（称为“宏块”）（取决于色彩格式，可以是 U 和 V 通道中的 16×16 、 16×8 或 8×8 区域）。每一宏块色彩平面被块化成 4×4 像素的区域或块。因此，宏块按以下对于本示例性编码器实现的方式由各种色彩格式组成：

1. 对于灰度图像，每一宏块包含 16 个 4×4 的亮度（Y）块。
2. 对于 YUV 4:2:0 格式彩色图像，每一宏块包含 16 个 4×4 的 Y 块，并且 4 个各自为 4×4 的色度（U 和 V）块。
3. 对于 YUV 4:2:2 格式彩色图像，每一宏块包含 16 个 4×4 的 Y 块，并且 8 个各自为 4×4 的色度（U 和 V）块。
4. 对于 RGB 或 YUV 4:4:4 彩色图像，每一宏块对 Y、U 和 V 通道中的每一个包含 16 个块。

因此，在变换之后，该代表性编码器 200/解码器 300 中的宏块具有三个频率子带：DC 子带（DC 宏块）、低通子带（低通宏块）和高通子带（高通宏块）。在该代表性系统中，低通和/或高通子带在比特流中是可任选的一这些子带可被完全丢弃。

此外，压缩的数据可按以下两种次序之一被压缩到比特流中：空间次序和频率次序。对于空间次序，瓦块内的同一宏块的不同子带被排序在一起，且所得的每一瓦块的比特流被写入一个分组中。对于频率次序，来自瓦块内的不同宏块的同一子带被分组在一起，且因此瓦块的比特流被写入三个分组中：DC 瓦块分组、低通瓦块分组和高通瓦块分组。另外，可以有其它数据层。

由此，对于该代表性系统，图像按以下“维度”来组织：

空间维度：帧→瓦块→宏块；

频率维度：DC|低通|高通；以及

通道维度：亮度|色度_0|色度_1……（例如，Y|U|V）。

以上箭头表示分层结构，而垂直条表示划分。

尽管该代表性系统按照空间、频率和通道维度来组织压缩的数字媒体数据，但是此处描述的灵活量化方法可以应用于沿着更少、更多或其它维度来组织其数据的替换编码器/解码器系统。例如，该灵活量化方法可应用于使用更大数量的频带、其它格式的色通道（例如，YIQ、RGB 等）、附加图像通道（例如，用于立体声视觉或其它多照相机阵列）的编码。

2. 灵活量化综述

在该代表性编码器/解码器中，量化操作本质上是按照量化参数 QP 的加偏除法，这在编码器处执行。反量化或乘法操作是与 QP 的乘法，这在解码器处执行。然而，此处描述的灵活量化的替换实现可利用其它形式的量化，包括均匀和非均匀、标量或矢量、有或没有死区等。这些量化/反量化过程共同引入了原始变换系数数据的丢失，这表现为解码的图像中的压缩误差或伪像。在最简单的编解码器中，可向一帧中的所有变换系数应用特定的固定 QP 值。尽管这在某些情况下可能是可接受的解决方案，但是它也有几个缺点：

人类视觉系统对于所有频率，或对一帧内的所有空间位置，或对所有亮度和色度通道并不都是同样敏感的。对不同系数使用不同 QP 值提供了甚至具有相同或更少数量的压缩比特的视觉上更出众的编码。同样，也可适当地优化其它误差度量。

码率控制或编码器产生期望大小的压缩文件的能力并不容易在整个帧上用单个 QP 来执行。

因此，理想地，应当可能允许编码器以任意的方式跨图像改变 QP。然而，这意味着用于每一数据分区（宏块/瓦块/通道/子带等）的实际 QP 值应在比特流中用信号表示。这导致巨大的开销，仅仅是为了携带 QP 信号表示信息，从而导致其在实践中是不合适的。所需的是用信号通知 QP 的灵活但节省比特的

手段，尤其是对于经常遇到的情形。

此处描述的灵活量化技术提供了沿着已编码数字媒体数据的各个分区或维度来改变量化的能力。例如，代表性编码器 200/解码器 300 系统中的灵活量化技术的一种实现可以在三个维度上改变量化—在(i)空间位置，(ii)频率子带，以及(iii)色通道上。然而，在该灵活量化技术的其它替换实现中，量化可以在更少、更多或其它数据维度或分组上变化。该技术还包括在已编码媒体数据中高效地用信号表示灵活量化的方式。该量化方法的好处是因涉及量化的辅助信息而引入的开销对于主要使用情形被最小化，同时如果编码器需要则允许最大的灵活性。

该灵活量化技术提供了对量化的细空间粒度控制。在一个具体实现中，该灵活量化允许对应用于帧、瓦块或向下到宏块的量化的控制。如果帧没有被均匀量化，则每一瓦块可以被均匀量化；如果瓦块没有被均匀量化，则每一宏块将被不同地量化。

该灵活量化还允许沿着频率子带维度进行量化控制。在一个具体实现中，该灵活量化包括指定频率子带之间的量化关系的子带模式。子带可被均匀量化，或被部分均匀量化（低通子带使用 DC 子带量化器，和/或高通子带使用低通量化器），或被独立量化。

该灵活量化还允许对沿着数据的通道维度应用的量化的控制。在一个具体实现中，该灵活量化包括指定色通道之间的量化关系的通道模式。通道可被均匀量化，或者部分地均匀量化（色度通道均匀，但亮度通道独立），或者被独立量化。

此处描述的灵活量化还提供高效地在压缩的数字媒体数据的辅助信息中用信号表示上述对空间、频率子带和通道的量化控制的组合的技术，这对于主要使用情形是重要的。此外，该灵活量化技术提供了通过从数字媒体数据中的一组定义的可能量化器子集进行索引来高效地定义量化器选择的方式。

3. 空间维度中的灵活量化：

在空间维度中，在代表性编码器/解码器中由灵活量化技术提供三个选择：

- 整个帧可使用相同的量化规则来编码。

- 否则，整个瓦块可使用相同的量化规则来编码，而帧中的不同瓦块可以使用不同的量化规则。
- 否则，瓦块中的每一宏块可使用相同的量化规则来编码，而瓦块中的不同宏块可以使用不同的量化规则。

用信号表示这些可能性的一种手段如下：在帧级的比特流中发送指示第一可能性是否为真的二进制信号。如果否，则在每一瓦块内的比特流中发送指示对该瓦块使用的量化规则的数目的固定长度码元。如果瓦块使用了多于一个量化规则，则在对应的瓦块内的每一宏块中发送指示由该宏块使用的量化规则的可变长度码元。解码器以与编码器一致的方式解释该比特流。

代表性编码器 200/解码器 300 使用上述信号表示的变体。由此处标记为“XXX_FRAME_UNIFORM”的通用句法元素表示的二进制信号仅在帧级发送（其中 XXX 是指定量化工控制的特定频率子带或通道维度的占位符）。在瓦块级，仅当帧级句法元素（XXX_FRAME_UNIFORM）为假时在瓦块级句法元素（XXX_QUANTIZERS）中发送不同量化器规则的数目。如果该数目等于 1，则意味着仅有一个规则，且因此该瓦块内的所有宏块用相同的量化规则来均匀编码（指示选择 2），如果否，则指示第三种可能性的选择。

4. 跨频带的灵活量化：

对于跨频带的灵活量化，代表性编码器 200/解码器 300 的比特流句法定义两个开关：

- 低通宏块使用与同一空间位置处的 DC 宏块相同的量化规则。这对应于句法元素 USE_DC_QUANTIZER。
- 高通宏块使用与同一空间位置处的低通宏块相同的量化规则。这对应于句法元素 USE_LP_QUANTIZER。

这些开关在整个帧使用相同的量化规则时在帧层启用，或者在其它情况下在瓦块层启用。这些开关不在宏块层启用。因此，瓦块内的所有宏块跨频率子带遵守相同的规则。在适当（帧或瓦块）层处对每一开关发送一二进制码元。

5. 跨图像通道的灵活量化：

对于跨通道的灵活量化,代表性编码器 200/解码器 300 的比特流句法定义三个选择:

- 所有通道一亮度和色度具有相同的量化规则。这由通用句法元素 `XXX_CH_MODE == CH_UNIFORM` 来指示。
- 亮度遵循一个量化规则,而所有色度通道遵循一不同的量化规则,由 `XXX_CH_MODE == CH_MIXED` 指示。
- 所有通道都自由选择不同的量化规则,由 `XXX_CH_MODE == CH_INDEPENDENT` 指示。

6. 组合的灵活量化:

代表性编码器 200/解码器 300 使用在以下图 4-14 所示的代码表中定义的比特流句法,该句法可以如上所述高效地跨维度来编码灵活量化选项中的特定选择。采用跨每一空间、频率子带和通道维度可用的若干量化选项,可用量化选项的排列数目是很大的。代表性编码器 200/解码器 300 的比特流可以用空间或频率排序来布局的事实增加了跨三个维度的灵活量化的复杂性。然而,这不改变可用量化选项,并且仅影响信号的串行化。图 4-14 定义的句法提供了对组合的灵活量化规则的高效编码。

如在代表性编码器/解码器的句法中定义的组合的量化规则的某些突出特征如下。

不允许 DC 量化在宏块的基础上变化。这允许对量化的 DC 值不同地编码而不必进行反缩放操作。即使在 AC (低通和高通) 频带用变化的量化编码的时候也用相对较小的量化器来对图像瓦块的 DC 频带编码不会显著影响比特率。

在一个极端,帧内的所有变换系数使用相同的量化参数。在另一极端,用于所有通道的低通和高通量化规则都允许对瓦块/帧的每一宏块独立地变化。唯一的限制是不同低通和高通量化器规则(覆盖所有通道)的数目各自被限于 16。每一这样的规则可以为每一通道指定独立的量化参数值。

在这些极端之间,允许如由图 4-14 所示的句法表指定的若干组合。

7. 量化器参数的索引：

代表性编码器/解码器中的特定量化参数（QP）基于谐波比例。量化器参数索引（QPI）的 8 位值对应于可能相当大的 QP 值。执行第二级索引使得跨宏块变化的 QPI 可以用高效的方式来编码。

更具体而言，编码器 200 可以定义比特流中包含 1 到 16 个 QPI “矢量”的集合。基于选择哪一 XXX_CHANNEL_MODE，每一 QPI 矢量由一个或多个 QPI 值组成。基于频率子带开关，这些集合是为 DC、低通和高通子带定义的。此外，DC 集合仅有一个 QPI 矢量，因为在瓦块通道中只允许一个 DC 量化器。对这些集合的编码在图 4-6 所示的表中定义。

如图 7-11 的表所示，对 DC、低通和高通频率子带的 QPI 矢量集合的信号表示如下发生。基于其它编码模式，每一集合的势（即，该集合中的 QPI 矢量数）对低通和高通子带在相应的瓦块或帧的开始处指示。DC 集合的势为 1。在该伪代码表中，表示势的句法元素被标记为“XXX_QUANTIZERS”。（在实践中，在比特流中发送 XXX_QUANTIZERS - 1）。表中标记为“XXX_QUANTIZER”的句法元素表示 QPI 集合的编码，这在图 4-6 所示的表中定义。

在宏块级，仅发送来自 QPI 集合内的所需 QPI 矢量的索引 QI 就足够。图 12-14 的表定义了宏块的基础上发送 QI 的句法。对应于 QI 的句法元素被标记为“XXX_QUANTIZER_INDEX”。使用可变长度代码来用信号表示 QI。首先，发送指示 QI 是否为 0 的一位码元。如果不是，则发送指示与 0 不同的特定 QI 的其长度由 $\text{ceil}(\log_2(\text{XXX\_QUANTIZERS} - 1))$ 给出的固定长度代码。这允许用低至每宏块一位来对“默认”量化规则（QI=0）进行高效编码。当 XXX_QUANTIZERS 为 1 时，XXX_QUANTIZER_INDEX 唯一地为 0，且因此不需要用信号表示 QI。

8. 扩展

以上对灵活量化的描述是专用于其在代表性编码器和解码器中的实现以及句法的。然而，该技术的原理也可扩展到其它数字媒体压缩系统和格式。例如，代表性编码器/解码器只有三个频率子带（DC、低通和高通）。但是，更

一般地，灵活量化的替换实现可以用直接的方式扩展到多个频率子带。同样，替换灵活量化实现可以在更细的空间粒度改变量化器，诸如通过在子宏块（诸如块）级发送量化索引（QI）信息。对灵活量化技术的基本原理的许多扩展在同一框架内都是可能的。

9. 计算环境

上述用于灵活量化的处理技术可以在各种数字媒体编码和/或解码系统的任一种上实现，包括计算机（各种形状因数，包括服务器、台式机、膝上型计算机、手持式计算机等）；数字媒体记录器和播放器；图像和视频捕捉设备（诸如照相机、扫描仪等）；通信设备（诸如电话、移动电话、会议设备等）；显示、打印或其它呈现设备；以及其它示例等等。该灵活量化技术可用硬件电路、控制数字媒体处理硬件的固件、以及在计算机或在其他计算环境中执行的通信软件来实现，诸如图 15 中所示。

图 15 示出了其中可实现所描述的实施例的合适计算环境（1500）的一个一般示例。计算环境（1500）并不旨在对本发明的使用范围或功能提出任何局限，因为本发明可在不同的通用或专用计算环境中实现。

参考图 15，计算环境（1500）包括至少一个处理单元（1510）和存储器（1520）。在图 15 中，这一最基本的配置（1530）被包括在虚线内。处理单元（1510）执行计算机可执行指令，并且可以是真实或虚拟处理器。在多处理系统中，多个处理单元执行计算机可执行指令以提高处理能力。存储器（1520）可以是易失性存储器（例如，寄存器、高速缓存、RAM）、非易失性存储器（例如，ROM、EEPROM、闪存等）或两者的某种组合。存储器（1502）储存实现所描述的使用灵活量化技术的数字媒体编码/解码的软件（1580）。

计算环境可具有额外的特征。例如，计算环境（1500）包括存储（1540）、一个或多个输入设备（1550）、一个或多个输出设备（1560）以及一个或多个通信连接（1570）。诸如总线、控制器或网络等互连机制（未示出）将计算环境（1500）的各组件互连。通常，操作系统软件（未示出）为在计算环境（1500）中执行的其它软件提供操作环境，并协调计算环境（1500）的各组件的活动。

存储（1540）可以是可移动或不可移动的，并包括磁盘、磁带或磁带盒、

CD-ROM、CD-RW、DVD 或可用于储存信息并可在计算环境（1500）内访问的任何其它介质。存储（1540）储存实现所描述的使用灵活量化技术的数字媒体编码/解码的软件（1580）的指令。

输入设备（1550）可以是诸如键盘、鼠标、笔或跟踪球等触摸输入设备、语音输入设备、扫描设备或向计算环境（1500）提供输入的另一设备。对于音频，输入设备（1550）可以是声卡或接受来自话筒或话筒阵列的模拟或数字形式的音频输入的类似设备，或向计算环境提供音频样本的 CD-ROM 读取器。输出设备（1560）可以是显示器、打印机、扬声器、CD 刻录机或提供来自计算环境（1500）的输出的另一设备。

通信连接（1570）允许在通信介质上与另一计算实体的通信。通信介质传达诸如计算机可执行指令、压缩音频或视频信息、或已调制数据信号中的其它数据的信息。已调制数据信号是其一个或多个特征以在信号中编码信息的方式设置或改变的信号。作为示例而非局限，通信介质包括以电、光、RF、红外、声学或其它载波实现的有线或无线技术。

此处所描述的使用灵活量化技术的数字媒体编码/解码可在计算机可读介质的一般上下文中描述。计算机可读介质可以是可在计算环境内访问的任何可用介质。作为示例而非局限，对于计算环境（1500），计算机可读介质可包括存储器（1520）、存储（1540）、通信介质和以上任一种的组合。

此处描述的使用灵活量化技术的数字媒体编码/解码可在诸如程序模块中所包括的、在目标真实或虚拟处理器上的计算环境中执行的计算机可执行指令的一般上下文中描述。一般而言，程序模块包括执行特定任务或实现特定抽象数据类型的例程、程序、库、对象、类、组件、数据结构等。程序模块的功能可以如各实施例中所需的组合或在程序模块之间分离。用于程序模块的计算机可执行指令可以在本地或分布式计算环境中执行。

出于表示的目的，详细描述使用了如“确定”、“生成”、“调整”和“应用”等术语来描述计算环境中的计算机操作。这些术语是由计算机执行的操作的高级抽象，且不应与人类所执行的动作混淆。对应于这些术语的实际的计算机操作取决于实现而不同。

鉴于可应用本发明的原理的许多可能的实施例，要求保护落入所附权利要

求书及其等效技术方案的范围和精神之内的所有这样的实施例作为本发明。

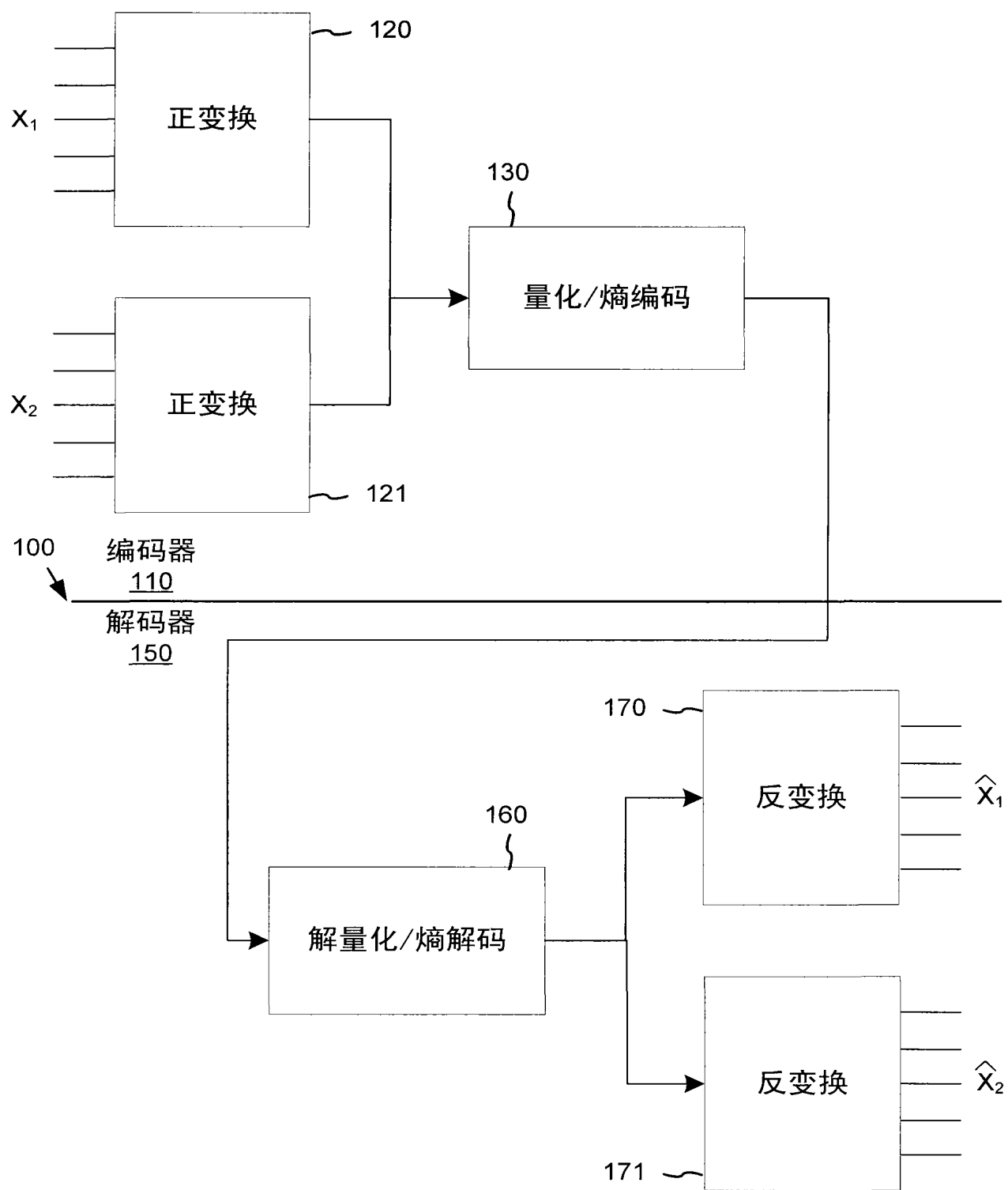


图 1
现有技术

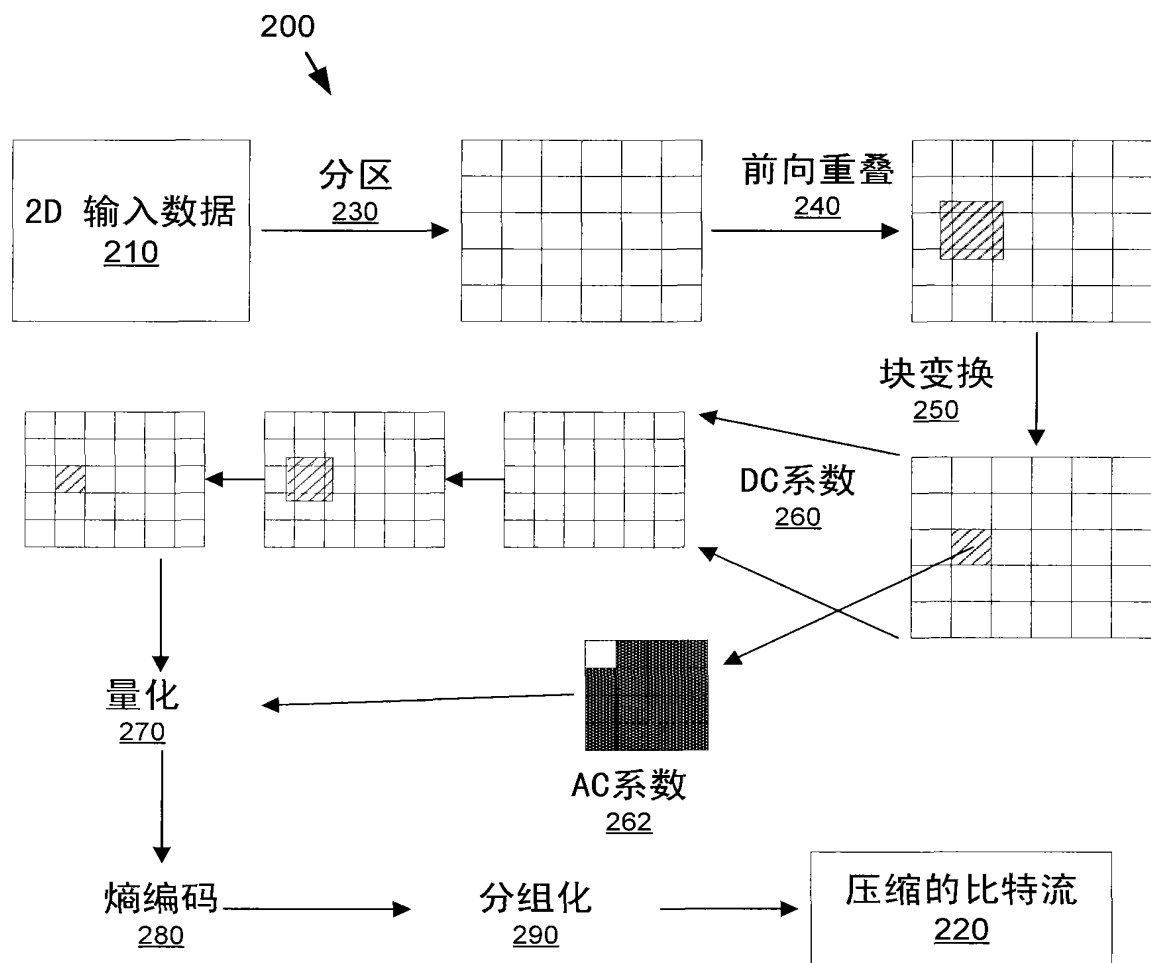


图 2

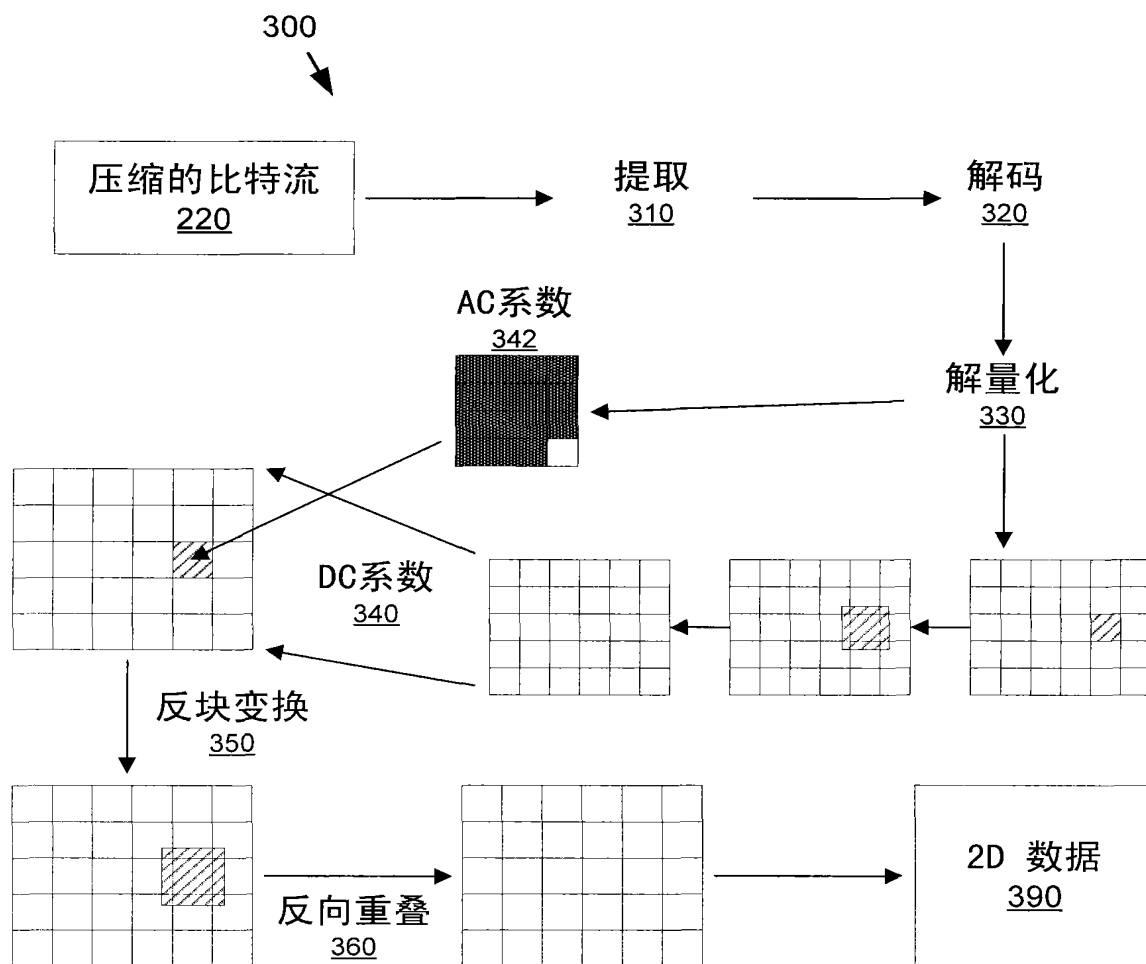


图 3

400



DC_QUANTIZER () {	位数
if(CHANNELS == 1)	
DC_CH_MODE = CH_UNIFORM	
Else	
DC_CH_MODE	2
if (DC_CH_MODE == CH_UNIFORM)	
DC_QUANT	8
Else if (DC_CH_MODE == CH_MIXED) {	
DC_QUANT_Y	8
DC_QUANT_UV	8
}	
Else if (DC_CH_MODE ==	
CH_INDEPENDENT) {	
for (i = 0; i < CHANNELS; i++)	
DC_QUANT_CHANNEL_i	8
}	
}	

图 4

500



LP_QUANTIZER () {	位数
for (q = 0; q < LP_QUANTIZERS; q++) {	
if(CHANNELS == 1)	
LP_CHMODE[q] = CH_UNIFORM	
Else	
LP_CH_MODE[q]	2
if (LP_CH_MODE[q] == CH_UNIFORM)	
LP_QUANT[q]	8
else if (LP_CH_MODE[q] ==	
CH_MIXED) {	
LP_QUANT_Y[q]	8
LP_QUANT_UV[q]	8
}	
else if (LP_CH_MODE[q] ==	
CH_INDEPENDENT)	
for (i = 0; i < CHANNELS; i++)	
LP_QUANT_CHANNEL_i[q]	8
}	
}	

图 5

600	HP_QUANTIZER () {	位数
	for (q = 0; q < HP_QUANTIZERS; q++) {	
	if(CHANNELS == 1)	
	HP_CH_MODE[q] = CH_UNIFORM	
	Else	
	HP_CH_MODE[q]	2
	if (HP_CH_MODE[q] == CH_UNIFORM)	
	HP_QUANT[q]	8
	else if (LP_CH_MODE[q] ==	
	CH_MIXED) {	
	HP_QUANT_Y[q]	8
	HP_QUANT_UV[q]	8
	}	
	else if (HP_CH_MODE[q] ==	
	CH_INDEPENDENT)	
	for (i = 0; i < CHANNELS; i++)	
	HP_QUANT_CHANNEL_i[q]	8
	}	
	}	

图 6

700	FRAME () {	位数
	...	许多
	DC_FRAME_UNIFORM	1
	if (DC_FRAME_UNIFORM)	
	DC_QUANTIZER	可变
	if('未跳过低通'){	
	USE_DC_QUANTIZER	1
	if(USE_DC_QUANTIZER == FALSE){	
	LP_FRAME_UNIFORM	1
	If(LP_FRAME_UNIFORM){	
	LP_QUANTIZERS = 1	
	LP_QUANTIZER()	可变
	}	
	}	
	if('未跳过高通'){	
	USE_LP_QUANTIZER	1
	if(USE_LP_QUANTIZER == FALSE){	
	HP_FRAME_UNIFORM	1
	If(HP_FRAME_UNIFORM){	
	HP_QUANTIZERS = 1	
	HP_QUANTIZER()	可变
	}	
	}	
	}	
	}	

图 7

800



SPATIAL_TILE () {	位数
...	许多
if (DC_FRAME_UNIFORM == FALSE) DC_QUANTIZER	可变
if ('未跳过低通' && LP_FRAME_UNIFORM == FALSE) { USE_DC_QUANTIZER	1
if (USE_DC_QUANTIZER == FALSE) LP_QUANTIZERS	4
LP_QUANTIZER	可变
}	
if ('未跳过高通' && HP_FRAME_UNIFORM){ USE_LP_QUANTIZER	1
if (USE_LP_QUANTIZER == FALSE) HP_QUANTIZERS	4
HP_QUANTIZER	可变
}	
...	许多
}	

图 8

900



DC_TILE () {	位数
...	许多
if (DC_FRAME_UNIFORM == FALSE) { DC_QUANTIZER	可变
}	
...	许多
}	

图 9

1000



LP_TILE () {	位数
...	许多
if (LP_FRAME_UNIFORM == FALSE) { USE_DC_QUANTIZER	1
if (USE_DC_QUANTIZER == FALSE) LP_QUANTIZERS	4
LP_QUANTIZER	可变
}	
...	许多
}	

图 10

1100
↓

HP_TILE () {	位数
...	许多
if (HP_FRAME_UNIFORM == FALSE) {	
USE_LP_QUANTIZER	1
if(USE_LP_QUANTIZER == FALSE)	
HP_QUANTIZERS	4
HP_QUANTIZER	可变
}	
...	许多
}	

图 11

1200
↓

SPATIAL_MB () {	位数
if ('未跳过低通' && LP_QUANTIZERS > 1 && USE_DC_QUANTIZER == FALSE)	
LP_QUANTIZER_INDEX	可变
if ('未跳过高通' && HP_QUANTIZERS > 1 && USE_HP_QUANTIZER == FALSE)	
HP_QUANTIZER_INDEX	可变
...	许多
}	

图 12

1300
↓

LP_MB () {	位数
...	许多
if (LP_QUANTIZERS > 1 && USE_DC_QUANTIZER == FALSE)	
LP_QUANTIZER_INDEX	可变
...	许多
}	

图 13

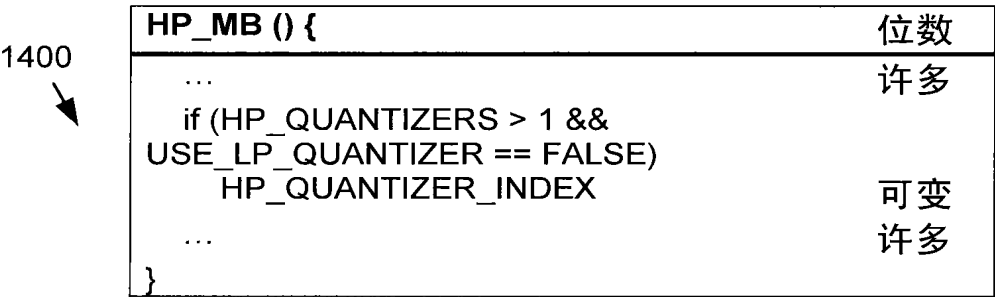


图 14

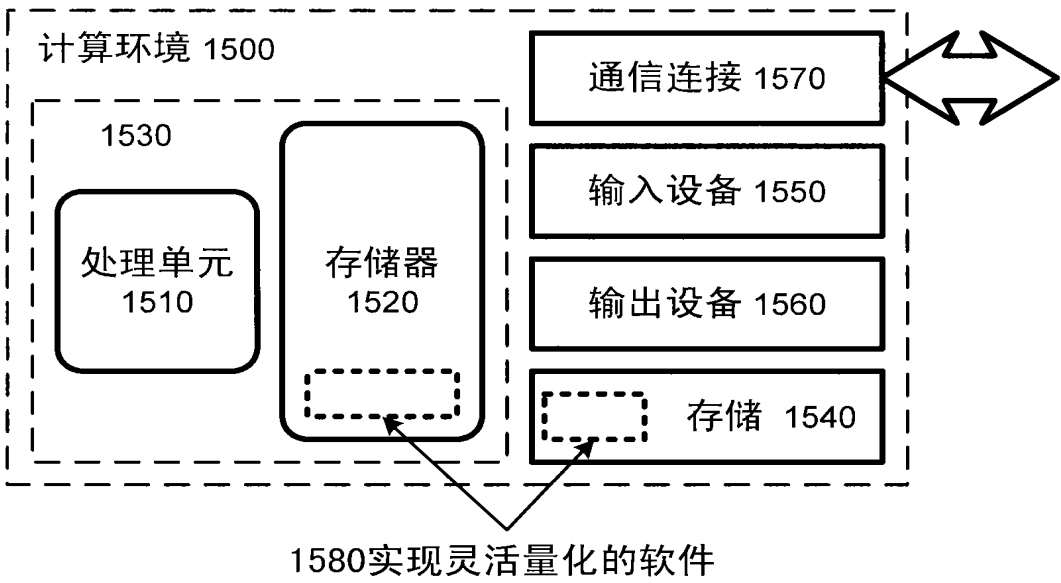


图 15



Publication number: CN101438591 A
Publication country: CHINA
Publication type: UNEXAMINED APPLIC. OPEN TO PUBLIC INSPECTION
Publication date: 20090520
Application number: CN200780015871
Application date: 20070504
Priority: WO2007US1084820070504WO2007US10848;
US2006041869020060505US20060418690;
Assignee^{std.}: MICROSOFT CORP ;
Inventor: TU CHENGJIE ; SRINIVASAN SRIDHAR ;
Inventor^{std.}: SRIDHAR SRINIVASAN ; CHENGJIE TU ;
International class⁸: H04N7/24;
Cooperative Patent class: H04N19/124 ; H04N19/126 ; H04N19/136 ;
H04N19/176 ; H04N19/46 ; H04N19/61 ;
Family members: AU2007248524 AA AU2007248524 BB BRPI0710704
A2 CA2647332 AA CA2647332 C CN101438591 A
CN101438591 B CN102833545 A CN102833545 B
CN105812815 A CN105847826 A CN105915906 A
EP2016773 A2 EP2016773 A4 HK1133144 A1
HK1179084 A1 IL194133 A0 IL194133 A1 IL225132 A0
IL225132 A1 IN05843CN2008 A JP2009536506 T2
JP2013153525 A2 JP2014209790 A2 JP2017063486
A2 JP5385781 B2 JP5580447 B2 JP6054920 B2
KR101477302 B1 KR20090015037 A MX2008013647
A1 RU2008143599 A RU2012156159 A RU2476000 C2
RU2619908 C2 US2007258518 AA US2007258519 AA
US2012224627 AA US2014233638 AA US8184694 BB
US8588298 BB US8711925 BB WO07130580 A2
WO07130580 A3
Title: FLEXIBLE QUANTIZATION
Title: 灵活量化
Abstract:

A digital media encoder/decoder uses a flexible quantization technique that provides the ability to vary quantization along various dimensions of the encoded digital media data, including spatial, frequency sub bands and color channels. The codec utilizes a signaling scheme to signal various permutations of flexible quantization combinations efficiently for primary

usage scenarios. When a choice of quantizer is available, the codec efficiently encodes the current quantizer by defining a subset of quantizers and indexes the current quantizer from the set.

Abstract:

一种数字媒体编码器/解码器使用提供沿着所编码的数字媒体数据的各个维度改变量化的功能的灵活量化技术，这些维度包括空间、频率子带和色通道。该编解码器利用一信号表示方案来高效地用信号表示灵活量化组合的各种排列以用于主要使用情形。当量化器选择可用时，该编解码器通过定义一量化器子集并从该集合中索引当前量化器来高效地编码当前量化器。

A digital media encoder / decoder uses the functions provided to change the quantization along various dimensions of the encoded digital media data flexible quantization technique, these dimensions include spatial, frequency sub-band and color channel. The codec utilizing a signal representing the program to efficiently signal the flexible quantization in combination with various arrangements for the main use cases. When the quantizer selection available, the codec quantization by defining a subset of the current quantization and index from the collection to efficiently encoding the current quantizer.

Claims:

1. 一种编码/解码数字媒体的方法，包括：沿着多于一个维度划分所述数字媒体；跨所述数字媒体的多于一个划分的维度改变应用于所述数字媒体的块的量化；以及编码/解码所述数字媒体的量化的块。

2. 如权利要求1所述的方法，其特征在于，所述改变量化包括跨空间维度改变量化。

3. 如权利要求1所述的方法，其特征在于，所述改变量化包括跨频率维度改变量化。

4. 如权利要求1所述的方法，其特征在于，所述改变量化包括跨通道改变量化。

5. 如权利要求1所述的方法，其特征在于，所述通道是色通道。

6. 如权利要求1所述的方法，其特征在于，所述改变量化包括跨空间、频率子带和通道维度中的多于一个维度改变量化。

1. An encoding / decoding digital media, comprising: dividing along more than one dimension of the digital media; more than one division of the digital media across the dimension of the change applied to the and quantized block encoding / decoding the digital media; quantization block of digital media.

2. The method of claim 1 or claim 2, wherein said changing quantization includes varying quantization across spatial dimension.

3. The method as claimed in claim 1, wherein, wherein said varying quantization includes varying quantization across a frequency dimension.

4. The method of claim 1 or claim 2, wherein said changing the quantization of

7.如权利要求1所述的方法，其特征在于，所述改变量化包括：对于一个划分的维度，选择是否在所述划分的维度的第一分区中均匀地量化数字媒体；如果选择不均匀地量化所述第一分区，则选择是否在所述划分的维度的第二分区中均匀地量化数字媒体；以及如果选择不均匀地量化所述第二分区，则跨所述维度独立地量化所述数字媒体。

8.如权利要求7所述的方法，其特征在于，所述第一分区是帧，所述第二分区是瓦块，而所述独立量化包括用不同的量化器独立地量化宏块内的所述数字媒体的块。

9.如权利要求7所述的方法，其特征在于，所述第一分区是DC频率子带，所述第二分区是低通频率子带，而所述独立量化包括用不同的量化器独立地量化所述低通频率子带或高通频率子带内的所述数字媒体的块。

10.如权利要求7所述的方法，其特征在于，所述第一分区是色度通道，所述第二分区是亮度通道，而所述独立量化包括用不同的量化器来独立地量化所述通道内的所述数字媒体的块。

11.如权利要求7所述的方法，其特征在于，还包括当跨所述维度独立地量化所述数字媒体时，指定可用于跨所述维度应用的一组不同的量化器，并从所述集合索引所述量化器的选择以应用于所述数字媒体的块。

12.一种数字媒体处理设备，包括：用于储存数字媒体数据的存储器；以及被编程为通过应用向所述数字媒体数据的块应用变换的前向或反向操作，并向所述数字媒体数据块应用量化器，来编码和/或解码所述数字媒体数据的数字媒体处理器，其中所述量化器可选择性地跨所述数字媒体数据的多于一个维度来变化。

13.如权利要求12所述的设备，其特征在于，所述维度包括空间、频率和色通道中的至少两个。

14.至少一种其上储存有用于使得数字媒体处理设备执行一种方法的程序代码的程序

variables including cross-channel change.

5. The method as claimed in claim 1, wherein, wherein said channel is a color channel.

6. The method of claim 1 or claim 2, wherein said changing quantization including cross spatial, frequency sub-bands and channel dimensions change more than one dimension to quantify.

7 The method as claimed in claim 1, wherein, wherein said varying quantization comprises: for one partitioned dimension, select whether a first dimension of said divided partitions uniformly quantized digital media; if you choose not uniformly quantize the first partition, then choose whether to uniformly quantized digital media division in the second partition of dimension; and if you choose not uniformly quantize the second partition then across the dimensions independently quantify the digital media.

8. The method as claimed in claim 7, wherein, wherein, said first partition is a frame, said second partition is a tile, and said independently quantizing comprises different a quantizer quantizing the macroblock independently of the digital media block.

9. The method as claimed in claim 7, wherein, wherein, said first partition is a DC frequency sub-band, the second partition is a low-pass frequency sub-band, and the said independent quantization including independent quantization of the low-pass or high-pass frequency sub-band frequency sub-bands within blocks of the digital media with different quantizer.

存储载体，所述方法根据编解码器来处理数字媒体数据，所述方法包括：选择要应用于所述数字媒体数据的量化模式的组合，所述量化模式跨空间、频率子带和色通道维度中的多于一个维度而变化；向所述数字媒体数据应用块变换；根据所选择的量化模式的组合量化所述数字媒体；以及在压缩的数据流中编码所述量化的数字媒体数据；以及在所述压缩的数据流中用信号表示所选择的量化模式的组合。

15. 如权利要求14所述的至少一种程序存储载体，其特征在于，可用于跨空间维度改变量化的所述量化模式包括用于对所述数字媒体数据的帧进行均匀量化的模式、用于对所述数字媒体数据的瓦块进行均匀量化的模式、以及用于独立地量化所述数字媒体数据的每一宏块的模式。

16. 如权利要求14所述的至少一种程序存储载体，其特征在于，可用于跨频率子带维度改变量化的所述量化模式包括用于跨所述数字媒体数据的频率子带进行均匀量化的模式、用于跨所述数字媒体数据的频率子带进行部分均匀量化的模式、以及用于独立地量化所述数字媒体数据的频率子带的模式。

17. 如权利要求16所述的至少一种程序存储载体，其特征在于，所述用于跨频率子带进行部分均匀量化的模式包括一个频率子带使用另一频率子带的量化器。

18. 如权利要求14所述的至少一种程序存储载体，其特征在于，可用于跨色通道维度改变量化的所述量化模式包括用于跨所述数字媒体数据的色通道进行均匀量化的模式、用于跨所述数字媒体数据的色通道进行部分均匀量化的模式、以及用于独立地量化所述数字媒体数据的色通道的模式。

19. 如权利要求14所述的至少一种程序存储载体，其特征在于，所述方法还包括当用于一维度中的数字媒体数据的分区的模式可变时，通过定义一可用量化器集合并从所述集合中索引来指定当前量化器。

20. 至少一种其上储存有用于使得数字媒体处理设备执行一种方法的程序代码的程序

10. The method as claimed in claim 7, wherein, wherein, said first partition is a chrominance channel, said second partition is a luminance channel, and said independently quantizing comprises with different quantizers to independently quantify the passage of the digital media block.

11. The method of claim 7, wherein the claims, characterized in that it comprises independently when the dimension across the quantization of the digital media, specify the dimensions that can be used across applications a different set of quantization, and the collection index of the quantizer select from to the block applied to digital media.

12 A digital media processing device, comprising: storing digital media for data storage; and before the application is programmed by the block data to the application of digital media to transform or reverse operation, and to apply the digital media data block quantizer to encode and / or decode the digital media data digital media processor, wherein the quantizer selectively across the digital multi-media data to a dimension to change.

13. Device as claimed in claim 12 wherein, characterized in that said at least two dimensions include spatial, frequency and color channel.

14. At least one program code stored thereon for causing a digital media processing device to perform a method for program storage carrier, the method according to the codec processing digital media data, said method comprising: selecting a combination of quantization mode to apply to the digital media data, the quantization pattern across space, frequency sub-band and color channels in more than one dimension dimensions vary; to the digital media data applications block transform; quantization

存储载体，所述方法根据编解码器处理数字媒体数据，所述方法包括解码根据权利要求14所述的方法编码的压缩的数据流。

mode according to the selected combination of quantizing the digital media; and in a compressed data stream encoding the quantized digital media data; and in the compressed data stream selected by a signal indicative of the quantization mode The combination.

At least one program-storing carrier

15. As claimed in claim 14, characterized in that, across the spatial dimension can be used to change the quantization of the quantization mode for the digital media comprising frame data uniform quantization mode for the tile of the digital media data uniform quantization mode, and means for independently quantizing the digital media data mode of each macroblock.

At least one program-storing carrier

16. As claimed in claim 14, characterized in that the frequency sub-band can be used for cross-dimension change comprises quantizing the quantization mode for the inter- frequency sub-band of the digital media data in a uniform quantization mode for the frequency sub across the digital media data portions of the model uniform quantization belt, and means for independently quantizing the digital media data mode frequency subbands.

At least one program-storing carrier

17. As claimed in claim 16, characterized in that the means for partial cross-frequency sub-band comprises a uniform quantization mode using a frequency subband Another quantization frequency sub-bands.

At least one program-storing carrier

18. As claimed in claim 14, characterized in that the cross-color channels can be used to change the dimension of the

quantization of the quantization mode includes a digital cross-Color uniform media data channel quantization mode for the color channels of the digital media data across partial uniform quantization mode, and means for independently quantizing the color channels of the digital media data pattern.

19. Claim 14, wherein the at least one program storage carrier claims, characterized in that the method further comprises partitioning mode is used when a dimension of the digital media data may Varying a set of available quantizers by defining and from the collection index to specify the current quantizer.

20. There are at least one stored thereon for causing a digital media processing device to perform a method of storing a program code vector, the method according to the codec processing digital media data, the said method comprising encoding decoding method according to 14, wherein the compressed data stream as claimed in claim.

Description:

灵活量化 背景 1. 基于块变换的编码 变换编码是在许多音频、图像和视频压缩系统中使用的一种压缩技术。未压缩数字图像和视频通常被表示或捕捉为以二维(2D)网格排列的图像或视频帧中各位置处的图元或色彩的样本。这被称为图像或视频的空间域表示。例如，用于图像的典型格式由被排列为网格的24位彩色图元流构成。每一样本是表示诸如RGB或YIQ等色彩空间内该网格中的一个像素位置处的色彩分量的数字。各种图像和视频系统可使用各种不同的色彩、空间和时间分辨率的采样。

类似地，数字音频通常被表示为时间采样的音频信号流。例如，典型的音频格式由

Flexible quantization Background 1. block-based transform coding transform coding is used in many audio, image and video compression systems in a compression technology. Uncompressed digital image and video is typically represented or captured as a two-dimensional (2D) images or video frames arranged in a grid at the location of each element or color of samples. This is called spatial domain representation of the image or video. For example, a typical format for images are arranged in a grid by a 24 color picture element stream composition. Each sample is a table showing color components such as RGB or YIQ other colors within a pixel

location in space at the grid of numbers. Variety of image and video systems may use various different color, spatial and temporal resolution of the sampling.

Similarly, digital audio is typically represented as time-sampled audio signal stream. For example, a typical audio format made at regular intervals taken at 16 constituting the audio signal amplitude sample streams. uncompressed digital audio, image and video signals can consume large amounts of storage and transmission capacity. Transform encoding by the space-domain signal transformed into a frequency domain representation (or other similar transform domain) representation, and then reducing resolution of certain generally the transform domain representation of the frequency components relatively imperceptible, reducing the digital audio, image and video size. This generally reduces the image or audio or video in the spatial domain or time domain spatial resolution compared to the color, resulting in a relatively non-perceptible degradation of the digital signal.

More specifically, a typical block transform-based codec 100 uncompressed digital pixel image shown in Figure 1 is divided into fixed-size two dimensional blocks ($X_h \dots X_n$), each piece may overlapping with other blocks. for each block application spatial - frequency analysis of the linear transformation 120-121, which converts the spaced samples within the block to a set of general represents frequency intervals corresponding to blocks of the digital signal band strength (or Transform) coefficients. For compression, the transform coefficients may be selectively quantized 130 (i.e., such as by dropping least significant bits of the coefficient values or higher resolution mapping of the digital set value to reduce the

11 9/28/2017

块变换120-121可被定义为对大小为N的向量x的数学运算。最通常的是，该运算是线性乘法，从而产生变换域输出： $y = MX$ ，M是变换矩阵。当输入数据是任意长时，它被分段成大小为N的向量，并且向每一段应用块变换。出于数据压缩的目的，选择可逆块变换。换言之，矩阵M是可逆的。在多个维度中(例如，对于图像和视频)，块变换通常被实现为可分操作。沿数据的每一维(即，行和列)可分地应用矩阵乘法。

对于压缩，变换系数（向量3；的分量）可被选择性地量化（即，诸如通过 丢弃系数的最低有效位或将较高分辨率数字集中的值映射到较低分辨率来 降低分辨率），并还可被熵或可变长度编码为压缩的数据流。

在解码器150中解码时，如图1所示，在解码器150侧应用这些操作的反过程（反量化/熵解码160和反块变换170-171）。在重构数据时，将逆矩阵 M1（反变换170-171）作为乘数应用于变换域数据。当应用于变换域数据时，反变换几乎重构原始时域或空间域数字媒体。

在许多基于块变换达到编码应用中，变换理想地是可逆的以取决于量化因子同时支持有损和无损压缩两者。如果例如没有量化（一般被表示为量化因子 1），则利用可逆变换的编解码器可在解码时精确地再现输入数据。然而，这些应用中的可逆性的要求约束了对用于设计编解码器的变换的选择。

诸如MPEG和Windows Media等许多图像和视频压缩系统利用基于离散余弦变换（DCT）的变换。已知DCT具有得到近乎最优的数据压缩的良好能量压缩特性。在这些压缩系统中，在压缩系统的编码器和解码器两者中的重构环路中采用了反DCT（IDCT）来重构各个图像块。

2.量化 根据一个可能的定义，量化是用于通常用于有损压缩的近似不可逆映射函数的术语，其中有一组指定的可能输出值，并且该组可能的输出值中的每一成员具有导致对该特定输出值的选择的一组相关联的输入值。已经开发了各种量化技术，包括标量或矢量、均匀或非均匀、有或没有死区、以及自适应或非自适应量化。

量化操作本质上是按照量化参数QP的加偏除法（biased division），这在编码器处执行。反量化或乘法操作是与QP的乘法，这在解码器处执行。这些过程共同引入了原始变换系数数据的丢失，这表现为解码的图像中的压缩误差或伪像。在最简单的编解码器中，可向一帧中的所有变换系数应用特定的固定QP值。尽管这在某些情况下可能是可接受的解决方案，但是它也有几个缺点：人类视觉系统对于所有频率，或对一帧内的所有空间位置，或对所有亮度和色度通道并不都是同样敏感的。对不同系数使用不同QP值提供了甚至具有相同或更少数量的压缩比特的视觉上更出众的编码。同样，也可适当地优化其它误差度量。

resolution to a lower resolution), and also entropy or 130 into a variable-length encoding compressed data stream.

When decoding, inverse transform coefficients transform 170-171 to nearly reconstruct the original color / spatial preclude the kind of image / video signal (reconstructed blocks of 1,...). block transform 120-121 can be defined as a mathematical operation of size N vector x. The most common is that the operation is a linear multiplication, producing the transform domain output; $= MX$, M is a transformation matrix. When the input data is arbitrarily long, it is segmented into a size of iV vector, and transformed to each segment application blocks. For the purpose of data compression, reversible block transform selection. In other words, the matrix M is invertible. In multiple dimensions (for example, for images and video), block transform is typically implemented as separable operations. Each dimension (ie, rows and columns) divided along data application matrix multiplication.

For compression, the transform coefficients (vector 3; components) may be selectively quantized (ie, such as by discarding the least significant bits of the coefficient values or higher-resolution digital set values are mapped to a lower resolution reducing the resolution), and also entropy or variable-length coded by compressed data stream. when the decoder 150 decodes, as shown in FIG. 1, the side of the decoder 150 inverse process of the application of these operations (dequantization / entropy decoding 160 and inverse block transform 170-171). When reconstructed data, the inverse matrix M1 (inverse transform 170-171) as the multiplier applied to the transform domain data. When applied to the transform domain data, the inverse transform nearly reconstruct the original time-domain or spatial-domain digital media.

码率控制或编码器产生期望大小的压缩文件的能力并不容易在整个帧上用单个QP来执行。

因此，期望的是允许编码器以任意的方式跨图像改变QP。然而，这意味着用于每一数据分区的实际QP值应在比特流中用信号表示。这导致巨大的开销，仅仅是为了携带QP信号表示信息，从而导致其在实践中是不合适的。所需的是用信号通知QP的灵活但节省比特的的手段，尤其是对于经常遇到的情形。

概括而言，量化是大多数图像和视频编解码器控制压缩的图像质量和压缩比的主要机制。大多数流行的编解码器支持的量化方法提供了极少的特征或极少的灵活性，或者导致附加比特的重大开销。通常，图像或视频帧通常被均匀地量化，或以在空间位置上改变量化的有限能力来量化。这一灵活性的缺乏损害了压缩质量，并阻止了进行中的准确的码率控制。另一方面，某些编解码器在支持量化方法方面提供几乎无限制的的自由。编码来用信号表示不同量化器的使用在已编码媒体中花费了额外的比特，并且本身会不利地影响压缩效率。

此外，构建一致解码器的过程需要量化器方法的所有可能组合生成的大量测试遍，这可能是费力的。

概述 以下详细描述呈现了提供沿着已编码的数字媒体数据的各维改变量化的能力的灵活量化技术的各种变型。例如,该灵活量化技术的一种代表性实现可以在三个维度上改变量化—在(i)空间位置,(ii)频率子带,以及(iii)色通道上。

该详细描述还呈现了在已编码数字媒体数据中高效地用信号表示该灵活量化的方式。该量化方法的好处是因涉及量化的辅助信息而引入的开销对于主要使用情形被最小化,同时如果编码器需要则允许最大的灵活性。

提供本概述以便以简化的形式介绍将在以下详细描述中进一步描述的一些概念。该概述不旨在标识所要求保护的主题的关键特征或本质特征，也不旨在用于帮助确定

In many block transform-based coding applications reached, the transform is desirably reversible to depend on the quantization factor supports both lossy and lossless compression. If for example there is no quantization (generally represented as a quantization factor of 1), the use of a reversible transform codec can accurately reproduce the input data at the time of decoding. However, these applications reversibility requirements for the design constraints for converting the codec of choice. such as MPEG and Windows Media, and many other image and video compression systems utilize discrete cosine transform transform (DCT) of. DCT has got almost the best known of a good amount of data can be compressed characteristics. In these compression systems, compression system encoder and decoder both reconstruction loop uses a counter-DCT (IDCT) to reconstruct each image block.

2. Quantify, according to one possible definition, quantification is typically used for lossy compression approximate term irreversible mapping function, which has a set of possible output value specified, and the set of possible values for each output members with input values result in the selection of that particular output value of a set of associated. We have been developed a variety of quantization techniques including scalar or vector, uniform or non-uniform, with or without dead zone, and adaptive or non-adaptive quantization. quantization operation is essentially in accordance with the quantization parameter QP plus partial division (biased division), which perform at the encoder. Inverse quantization or multiplication operation is multiplication and QP, which is executed at the decoder. These processes jointly introduced the lost original transform coefficient data, which showed the decoded image compression errors or

所要求保护的的主题的范围。本发明的其它特征和优点在参考附图继续阅读以下对实施例的详细描述后将变得显而易见。

附图简述 图1是现有技术中常规的基于块变换的编解码器的框图。

图2是包含块模式编码的代表性编码器的流程图。

图3是包含块模式编码的代表性解码器的流程图。

图4是根据灵活量化技术的包含用于用信号表示DC量化器的伪代码定义的表。

图5是根据灵活量化技术的包含用于用信号表示低通量化器的伪代码定义的表。

图6是根据灵活量化技术的包含用于用信号表示高通量化器的伪代码定义的表。

图7是根据灵活量化技术的包含用于用信号表示帧层处的量化器的伪代码定义的表。

图8是根据灵活量化技术的包含用于用信号表示瓦块层处的量化器的伪代码定义的表。

图9是根据灵活量化技术的包含用于用信号表示瓦块层处的DC子带的量 化器的伪代码定义的表。

图10是根据灵活量化技术的包含用于用信号表示瓦块层处的低通子带的 量化器的伪代码定义的表。

图11是根据灵活量化技术的包含用于用信号表示瓦块层处的高通子带的 量化器的伪代码定义的表。

图12是根据灵活量化技术的包含用于用信号表示宏块层处的量化器的伪代码定义的表。

图13是根据灵活量化技术的包含用于在频率模式中用信号表示宏块层处的低通量化器的伪代码定义的表。

artifacts. In the simplest codecs can be made to an all transform coefficients in the application-specific fixed QP value. While this may be an acceptable solution in some cases, but it also has several disadvantages: the human visual system for all frequencies, or to all spatial positions within one frame, or to all luminance and chrominance channels are not It is equally sensitive. Using different QP values for different coefficients provided even with the same or lesser number of compressed bits more visually superior encoding. Similarly, it may be appropriate to optimize their error metric.

Ability

rate control or encoder to produce the desired size of the compressed file will not be easy to implement on the entire frame with a single QP.

Thus, it is desirable to allow the encoder in an arbitrary manner across the image change QP. However, this means that the actual QP value for each data partition should flow with signal representing a bit. This leads to a huge overhead, just to carry the QP signal indicating information, causing it in practice is not appropriate. Required is a flexible means of signaling QP, but saves bits, especially for situations frequently encountered.

In summary, the quantization is most image and video compression codec control image quality and compression ratio of the main mechanisms. Most popular codec support quantification method provides a few features or very little flexibility or result in additional significant overhead bits. Typically, images or video frames are usually uniformly quantized, or to quantify the change in the spatial position of the limited ability to quantify. This lack of flexibility damage the compression

图14是根据灵活量化技术的包含用于在频率模式中用信号表示宏块层处的高通量化器的伪代码定义的表。

图15是用于实现具有灵活量化的媒体编码器/解码器的合适的计算环境的框图。

详细描述 以下详细描述涉及提供能够改变沿着空间、频率和色彩维度应用的量化的 高效灵活量化（此处称为“灵活量化”）的编码和解码技术。以下描述在数字媒体压缩系统或编解码器的上下文中描述了该技术的一个示例实现。该数字媒体系统以压缩形式对数字媒体数据进行编码以便传输或存储，并解码该数据以供回放或其它处理。出于说明的目的，包含这一灵活量化的该示例性压缩系统是图像或视频压缩系统。或者，该技术也可被结合到用于其它2D数据的压缩系统或编解码器中。灵活量化技术不要求数字媒体压缩系统以特定的编码格式来编码压缩的数字媒体数据。

1. 编码器/解码器 图2和3是在代表性2维（2D）数据编码器200和解码器300中采用的过程的一般化图示。该图呈现了结合了实现块模式编码的2D数据编码器和解码器的压缩系统的一般化或简化的图示。在使用块模式编码的替换压缩系统中，可使用比本代表性编码器和解码器中所示的更多或更少的过程来进行2D数据压缩。例如，某些编码器/解码器还可包括色彩转换、色彩格式、可缩放编码、无损编码、宏块模式等等。取决于量化，压缩系统（编码器和解码器）可提供2D数据的无损和/或有损压缩，量化可基于从无损到有损变化的量化参数。

2D数据编码器200产生压缩比特流220，它是作为输入提供给编码器的2D数据210的更紧凑表示（对于典型输入）。例如，2D数据输入可以是图像、视频序列的一帧、或具有两个维度的其它数据。2D数据编码器将输入数据帧划分成块（一般在图2中示为分区230），这在所示的实现中是形成块该帧的平面的规则图案的非重叠4x4像素块。这些块被分组成称为宏块的群集，在该代表性编码器中其大小是16x16像素。宏块进而被分组成称为瓦块的规则结构。

quality, and prevent the ongoing accurate rate control. On the other hand, some codecs provide almost unlimited freedom in support of quantification methods. Coding signal indicates the use of different quantizers in the encoded media spent extra bits, and itself adversely affect compression efficiency.

In addition, the process of constructing a consistent decoder need to quantify all possible combinations of methods to generate a large number of test times, which may be laborious.

The following presents an overview provides the ability to change along each dimension has been encoded digital media data to quantify the flexible quantization techniques described in detail with various modifications. For example, the flexible quantization techniques to achieve a representation in three dimensions can be changed in a quantized (i) spatial location, (ii) frequency sub bands, and (iii) color channels.

The detailed description is also presented in the encoded digital media data efficiently signal indicates that the flexible quantization approach. The benefit of this quantization approach is that involving the quantized side information overhead introduced mainly used for the case is minimized, while if the encoder is required to allow maximum flexibility.

This Summary is provided to introduce a simplified form a selection of concepts that are further described below in the Detailed Description. This summary is not intended to identify the requirements relating to the protection of key features or essential features, nor intended to be used to help determine the scope of the claimed subject matter. Other features and advantages of the present invention with reference to the accompanying drawings continue reading the following

瓦块也可形成图像上的规则图案，使得水平行中的瓦块是统一的高度且是对齐的，而垂直列中的瓦块是统一的宽度且是对齐的。在该代表性编码器中，瓦块可以是任意大小，该大小在水平和/或垂直方向上是16的倍数。替换编码器实现可以将图像划分成块、宏块、瓦块或其它大小和结构的其它单元。

对块之间的每一边缘应用“前向重叠”算子240，之后使用块变换250来变换每一4X4的块。该块变换250可以是由Srinivasan在2004年12月17日提交的题为“Reversible Transform For Lossy And Lossless 2-D Data Compression”(用于有损和无损2D数据压缩的可逆变换)的美国专利申请第11/015,707号中所描述的可逆的、无缩放的2D变换。重叠算子240可以由Tu等人在2004年12月17日提交的题为“Reversible Overlap Operator for Efficient Lossless Data Compression”(用于高效无损数据压缩的可逆重叠算子)的美国专利申请第11/015,148号；以及Tu等人在2005年1月14日提交的题为“Reversible 2-Dimensional Pre-/Post-Filter for Lapped Biorthogonal Transform”(用于重叠双正交变换的可逆2维预/后滤波器)的美国专利申请第11/035,991号中描述的可逆重叠算子。或者，可使用离散余弦变换或其它块变换和重叠算子。在变换之后，令每一4X4的变换块的DC系数260经受一类似的处理链(块化、前向重叠、之后是4X4的块变换)。所得的DC变换系数和AC变换系数被量化270、熵编码280和分组化290。

解码器执行反过程。在解码器侧，从其各自的分组中提取310变换系数位，从中系数本身被解码320和解量化330。DC系数340通过应用反变换来重新生成，并且DC系数的平面使用跨DC块边缘应用的合适的平滑算子来“反重叠”。随后，通过向DC系数应用4X4的反变换350来重新生成整个数据，并从比特流中解码AC系数342。最后，对所得图像平面中的块边缘进行反重叠滤波360。这产生经重构的2D数据输出。

detailed description of the embodiments will become apparent.

BRIEF DESCRIPTION OF THE DRAWINGS Figure 1 is a block diagram of the prior art block transform-based codec conventional.

2 is a flowchart representative encoder comprises a block pattern coding.

FIG. 3 is a flow chart representation of the decoder comprises a block pattern coding.

Figure 4 is a table that contains a pseudo-code definition for signaling indicates DC quantizer flexible quantization technique.

Figure 5 is a flexible quantization technique table contains definitions for the pseudo-code signal represented by the quantification of the low pass.

Figure 6 is a flexible quantization technique contains signaled Qualcomm quantizer defined pseudo-code form.

7 is a flexible quantization technique table contains definitions for the pseudo-code signal the quantization frame layer at the.

Figure 8 is a flexible quantization technique contains signal the quantization table tile layer at pseudo-code definition.

Figure 9 is a table that contains a pseudo-code definition for signaling represents the amount of the DC sub-layer of the tile with the flexible quantization technique.

Figure 10 is a table containing a pseudo-code definition for signaling of quantizers represent low-pass sub-band at the tile floor of the flexible quantization technique.

在一个示例性实现中，编码器200 (图2)将输入图像压缩成压缩比特流 220 (例如，文件)，而解码器300 (图3)基于是采用无损还是有损编码来重构原始输入或其近似。编码过程涉及应用以下所讨论的前向重叠变换 (LT)，这是用同样在以下更全面描述的可逆2维预/后滤波来实现的。解码过程涉及应用使用可逆2维预/后滤波的反重叠变换 (ILT)。

所示的LT和ILT在确切的意义上是彼此的逆，并且因此可被统称为可逆重叠变换。作为一种可逆变换，LT/ILT对可用于无损图像压缩。

由所示的编码器200/解码器300压缩的输入数据210可以是各种色彩格式 (例如'RGB/YUV 4:4:4、YUV 4:2:2或YUV 4:2:0彩色图像格式) 的图像。

通常，输入图像总是具有亮度 (Y)分量。如果它是RGB/YUV4:4:4、YUV 4:2:2 或 YUV 4:2:0图像，则该图像还具有色度分量，诸如U分量和V分量。图像的这些单独的色彩平面或分量可具有不同的空间分辨率。在例如YUV4:2:0色彩格式的输入图像的情况下，U和V分量具有Y分量一半的宽度和高度。

如上所述，编码器200将输入图像或图片块化成宏块。在一个示例性实现中，编码器200将输入图像块化成Y通道中的16X16像素区域 (称为“宏块”) (取决于色彩格式，可以是U和V通道中的16X16、16X8或8X8区域)。

每一宏块色彩平面被块化成4X4像素的区域或块。因此，宏块按以下对于本示例性编码器实现的方式由各种色彩格式组成：

- 1.对于灰度图像，每一宏块包含16个4X4的亮度 (Y)块。

- 2.对于YUV 4:2:0格式彩色图像，每一宏块包含16个4X4的Y块，并且4个各自为4X4的色度 (U和V)块。

- 3.对于YUV 4:2:2格式彩色图像，每一宏块包含16个4X4的Y块，并且8个各自为4X4的色度 (U和V)块。

FIG. 11 is a table containing a pseudo-code definition for signaling of quantizers Gao expressed through sub-layer of the tile with the flexible quantization technique.

Figure 12 is a flexible quantization techniques for signaling a table containing quantization macroblock layer at pseudo-code definition.

Figure 13 is a flexible quantization technique comprises means for frequency mode with the signal represents a low-pass quantizer macroblock layer at pseudo-code defined tables.

FIG. 14 is a flexible quantization technique comprises frequency mode for the signal represented by the high-pass quantizer macroblock layer at pseudo-code defined tables.

FIG. 15 is used to implement a flexible quantization media encoder Diagram / suitable computing environment decoder.

Detailed Description The following detailed description is directed to providing efficient and flexible quantization can be changed along the quantized spatial, frequency and color dimensions application (referred to herein as "flexible quantization") encoding and decoding techniques. The following description of the context of the digital media compression system or codec is described in an example of the technology implementation. The digital media system in a compressed form of digital media data is encoded for transmission or storage, and decodes the data for playback or other processing. For purposes of illustration, the flexible quantization comprising this exemplary compression system is an image or video compression system. Alternatively, the technique can also be used for other 2D data incorporated into compression

4.对于RGB或YUV4:4:4彩色图像，每一宏块对Y、U和V通道中的每一个包含16个块。

因此，在变换之后，该代表性编码器200/解码器300中的宏块具有三个频率子带：DC子带（DC宏块）、低通子带（低通宏块）和高通子带（高通宏块）。

在该代表性系统中，低通和/或高通子带在比特流中是可任选的—这些子带可被完全丢弃。

此外，压缩的数据可按以下两种次序之一被压缩到比特流中：空间次序和频率次序。对于空间次序，瓦块内的同一宏块的不同子带被排序在一起，且所得的每一瓦块的比特流被写入一个分组中。对于频率次序，来自瓦块内的不同宏块的同一子带被分组在一起，且因此瓦块的比特流被写入三个分组中：DC瓦块分组、低通瓦块分组和高通瓦块分组。另外，可以有其它数据层。

由此，对于该代表性系统，图像按以下“维度”来组织：空间维度：帧—瓦块—宏块；频率维度：DC|低通|高通；以及通道维度：亮度|色度—0|色度_1……（例如，Y|U|V）。

以上箭头表示分层结构，而垂直条表示划分。

尽管该代表性系统按照空间、频率和通道维度来组织压缩的数字媒体数据，但是此处描述的灵活量化方法可以应用于沿着更少、更多或其它维度来组织其数据的替换编码器/解码器系统。例如，该灵活量化方法可应用于使用更大数量的频带、其它格式的色通道（例如，YIQ、RGB等）、附加图像通道（例如，用于立体声视觉或其它多照相机阵列）的编码。

2.灵活量化综述 在该代表性编码器/解码器中，量化操作本质上是按照量化参数QP的加偏除法，这在编码器处执行。反量化或乘法操作是与QP的乘法，这在解码器处执行。然而，此处描述的灵活量化的替换实现可利用其它形式的量化，包括均匀和非均匀、标量或矢量、有或没有死区等。

systems or codecs. Flexible quantization technique does not require a digital media compression system with specific encoding format for encoding the compressed digital media data.

1. The encoder / decoder 2 and 3 are shown in a generalized representation of a two-dimensional (2D) data encoder 200 and decoder 300 employed in the process. The figure presents a combination of 2D coded block pattern realization data encoder and decoder of the compression system of generalized or simplified illustration. In use block mode coding replace compression system, you can use more or less than this process representative encoder and decoder shown for 2D data compression. For example, some encoders / decoders may also include color conversion, color formats, scalable coding, lossless coding, macroblock modes, and so on. Depending on the quantization, compression system (encoder and decoder) can provide lossless 2D data and / or lossy compression, quantization may be based on a quantization parameter varying from lossless to lossy.

200 generates

2D data encoder compressed bitstream 220, which is provided as an input to the encoder 210, a 2D data more compact representation (for typical input). For example, 2D data input can be a frame image of a video sequence, or other data having two dimensions. 2D data encoder input data frame is divided into blocks (illustrated generally in Figure 2 as partitioning 230), which is non-overlapping 4x4 pixel blocks forming a regular pattern in the plane of the frame pieces shown in the realization. These blocks are grouped in clusters called macroblock in this representative encoder whose size is 16x16 pixels. Further macro block is divided rule

这些量化/反量化过程共同引入了 原始变换系数数据的丢失，这表现为解码的图像中的压缩误差或伪像。在最简单的编解码器中，可向一帧中的所有变换系数应用特定的固定QP值。尽管这在某些情况下可能是可接受的解决方案，但是它也有几个缺点：人类视觉系统对于所有频率，或在一帧内的所有空间位置，或对所有亮度和色度通道并不都是同样敏感的。对不同系数使用不同QP值提供了甚至具有相同或更少数量的压缩比特的视觉上更出众的编码。同样，也可适当地优化其它误差度量。

码率控制或编码器产生期望大小的压缩文件的能力并不容易在整个帧上用单个QP来执行。

因此，理想地，应当可能允许编码器以任意的方式跨图像改变QP。然而，这意味着用于每一数据分区（宏块/瓦块/通道/子带等）的实际QP值应在比特流中用信号表示。这导致巨大的开销，仅仅是为了携带QP信号表示信息，从而导致其在实践中是不合适的。所需的是用信号通知QP的灵活但节省比特的 手段，尤其是对于经常遇到的情形。

此处描述的灵活量化技术提供了沿着已编码数字媒体数据的各个分区或 维度来改变量化的能力。例如，代表性编码器200/解码器300系统中的灵活量化技术的一种实现可以在三个维度上改变量化—(i)空间位置，(ii)频率子带，以及(iii)色通道上。然而，在该灵活量化技术的其它替换实现中，量化可以在 更少、更多或其它数据维度或分组上变化。该技术还包括在已编码媒体数据中 高效地用信号表示灵活量化的方式。该量化方法的好处是因涉及量化的辅助信息而引入的开销对于主要使用情形被最小化，同时如果编码器需要则允许最大的灵活性。

该灵活量化技术提供了对量化的细空间粒度控制。在一个具体实现中，该灵活量化允许对应用于帧、瓦块或向下到宏块的量化的控制。如果帧没有被均匀量化，则每一瓦块可以被均匀量化；如果瓦块没有被均匀量化，则每一宏块 将被不同地量化。

structures called tiles. tiles can also form a regular pattern on the image, so that the horizontal row of tiles is uniform height and are aligned, and the vertical columns of tiles uniform width and are aligned. In this representative encoder, the tiles can be any size, the size in the horizontal and / or vertical direction is a multiple of 16. Replace encoder can realize the image is divided into blocks, macro blocks, other units or other tile size and structure.

For each edge applications between blocks "forward overlap" operator 240, after using a block transform 250 to transform blocks of each 4X4. The block transform 250 can be made Srinivasan at December 17, 2004 filed entitled "Reversible Transform For Lossy And Lossless 2-D Data Compression" (for lossy and lossless data compression 2D reversible transformation) of U.S. Patent Application No. 11 No. 015,707 reversible described / and non-scaled 2D transform 240 may overlap operator by Tu et al., December 17, 2004 filed entitled "Reversible Overlap Operator for Efficient Lossless Data Compression" No. 11 / 015,148, U.S. Patent Application (lossless data compression for efficient reversible overlap operator); and Tu et al., January 14, 2005 filed entitled "Reversible 2-Dimensional Pre- / Post-Filter for Lapped Biorthogonal Transform" reversible overlap operator U.S. patent application (for reversible lapped biorthogonal transform two-dimensional pre / post-filter) Serial No. 11 / 035,991 described. Alternatively, you can use the discrete cosine transform or other block transform and overlap operators After the transformation, so that each DC coefficient transform block 260 4X4 subjected to a similar processing chain (block-oriented, forward overlap, followed by a block transform 4X4). The resulting DC transform coefficients and the AC transform coefficients are quantized 270,

该灵活量化还允许沿着频率子带维度进行量化控制。在一个具体实现中，该灵活量化包括指定频率子带之间的量化关系的子带模式。子带可被均匀量化，或被部分均匀量化（低通子带使用DC子带量化器，和/或高通子带使用低通量化器），或被独立量化。

该灵活量化还允许对沿着数据的通道维度应用的量化的控制。在一个具体实现中，该灵活量化包括指定色通道之间的量化关系的通道模式。通道可被均匀量化，或者部分地均匀量化（色度通道均匀，但亮度通道独立），或者被独立量化。

此处描述的灵活量化还提供高效地在压缩的数字媒体数据的辅助信息中用信号表示上述对空间、频率子带和通道的量化控制的组合的技术，这对于主要使用情形是重要的。此外，该灵活量化技术提供了通过从数字媒体数据中的一组定义的可能量化器子集进行索引来高效地定义量化器选择的方式。

3.空间维度中的灵活量化：在空间维度中，在代表性编码器/解码器中由灵活量化技术提供三个选择：•整个帧可使用相同的量化规则来编码。

•否则，整个瓦块可使用相同的量化规则来编码，而帧中的不同瓦块可以使用不同的量化规则。

•否则，瓦块中的每一宏块可使用相同的量化规则来编码，而瓦块中的不同宏块可以使用不同的量化规则。

用信号表示这些可能性的一种手段如下：在帧级的比特流中发送指示第一可能性是否为真的二进制信号。如果否，则在每一瓦块内的比特流中发送指示对该瓦块使用的量化规则的数目的固定长度码元。如果瓦块使用了多于一个量化规则，则在对应的瓦块内的每一宏块中发送指示由该宏块使用的量化规则的可变长度码元。解码器以与编码器一致的方式解释该比特流。

代表性编码器200/解码器300使用上述信号表示的变体。由此处标记为“XXX_FRAME—UNIFORM”的通用句法元

entropy coded 280 and packetized 290. decoder performs the inverse of. On the decoder side, the transform coefficient bits extracted 310 from their respective packets, from which the coefficients are themselves decoded 320 330. Reconciliation to quantify. DC coefficient 340 by applying the inverse transform to regenerate, and the plane of DC coefficients across the DC block edges using a suitable smoothing operator applied to the "anti-overlap." Then, regenerate the entire data by application to the DC coefficient inverse transform 350 4X4 come and AC coefficients 342 decoded from the bitstream. Finally, the resulting image plane block edges overlap inverse filter 360. This produces a reconstructed 2D data output.

In an exemplary implementation, the encoder 200 (FIG. 2) of the input image compressed into the compressed bitstream 220 (e.g., a file), and the decoder 300 (FIG. 3) is based on the use of lossless or lossy coding weight structure of the original input or an approximation thereof. Encoding process involves a forward lapped transform (LT) discussed in the following, this is the same with the reversible more fully described in the following two-dimensional pre / post-filtering to achieve. Decoding process involves the application reversible lapped transform two-dimensional anti-pre / post filtering (ILT).

LT and the ILT shown on the exact meaning of the inverse of each other, and therefore can be collectively referred to as a reversible lapped transform. As a reversible transformation, LT / ILT pair can be used for lossless image compression. as shown by the encoder 200 / decoder 300 inputs the compressed data 210 can be of various color formats (e.g., 'RGB / YUV 4: 4: 4, YUV 4: 2: 2 or YUV 4: 2: 0 Color Image Format) image.

素表示的二进制信号仅在帧级发送 (其中XXX是指定量化器控制的特定频率子带或通道维度的占位符)。在瓦块级, 仅当帧级句法元素 (XXX_FRAME_UNIFORM) 为假时在瓦块级句法元素 (XXX—QUANTIZERS) 中发送不同量化器规则的数目。如果该数目等于1, 则意味着仅有一个规则, 且因此该瓦块内的所有宏块用相同的量化规则来均匀编码 (指示选择2), 如果否, 则指示第三种可能性的选择。

4. 跨频带的灵活量化: 对于跨频带的灵活量化, 代表性编码器200/解码器300的比特流句法定义两个开关: 低通宏块使用与同一空间位置处的DC宏块相同的量化规则。这对应于句法元素 USE_DC—QUANTIZER。

高通宏块使用与同一空间位置处的低通宏块相同的量化规则。这对应于句法元素 USE_LP_QUANTIZER。

这些开关在整个帧使用相同的量化规则时在帧层启用, 或者在其它情况下在瓦块层启用。这些开关不在宏块层启用。因此, 瓦块内的所有宏块跨频率子带遵守相同的规则。在适当 (帧或瓦块) 层处对每一开关发送一二进制码元。

5. 跨图像通道的灵活量化: 对于跨通道的灵活量化, 代表性编码器200/解码器300的比特流句法定义三个选择: 所有通道一亮度和色度具有相同的量化规则。这由通用句法元素 XXX_CH_MODE == CH_UNIFORM 来指示。

亮度遵循一个量化规则, 而所有色度通道遵循一不同的量化规则, 由 XXX_CH_MODE == CH_MIXED 指示。

所有通道都自由选择不同的量化规则, 由 XXX—CH—MODE== CH_INDEPENDENT 指示。

6. 组合的灵活量化: 代表性编码器200/解码器300使用在以下图4-14所示的代码表中定义的比特流句法, 该句法可以如上所述高效地跨维度来编码灵活量化选项中的特定选择。采用跨每一空间、频率子带和

Typically, the input image always has a luminance (Y) component. If it is a RGB / YUV4: 4: 4, YUV 4: 2: 2 or YUV 4: 2: 0 image, the image also has chrominance components, such as the U and V components. These separate color planes or components of the image can have different spatial resolutions. In the example YUV4: 2: 0 input image under color format, U and V components have half the Y component width and height.

As described above, the encoder 200 block of the input image or picture into macroblocks. In an exemplary implementation, the encoder 200 tiles the input into the Y channel of 16X 16 pixel region (called a "macro block") (depending on the color format, the U and V channels may be in or 8X8 16X16, 16X8 area).

Each macroblock color plane is 4X4 pixel block into regions or blocks. Thus, for a macroblock the following the present exemplary encoder implementation manner by the composition of the various color formats: 1. For a grayscale image, each macroblock contains 16 4X4 luminance (Y) blocks.

2 for YUV 4: 2: 0 format color image, each macroblock contains 16 4X4 Y-block, and 4 are each 4X4 chrominance (U and V) blocks.

3 for YUV 4: 2: 2 format color image, each macroblock contains 16 4X4 Y-blocks, and eight are each 4X4 chrominance (U and V) blocks.

4 for RGB or YUV4: 4: 4 color image, each macroblock for Y, U and V channels each contain 16 blocks.

Therefore, after conversion, the representative encoder 200 / decoder 300 in the macro block having three frequency sub-bands: DC sub-band (DC macroblock), a low-pass sub-band (low

通道维度可用的若干量化选项，可用量化选项的排列数目是很大的。代表性编码器200/解码器300的比特流可以用空间或频率排序来布局的事实增加了跨三个维度的灵活量化的复杂性。然而，这不改变可用量化选项，并且仅影响信号的串行化。图4-14定义的句法提供了对组合的灵活量化规则的高效编码。

如在代表性编码器/解码器的句法中定义的组合的量化规则的某些突出特征如下。

不允许DC量化在宏块的基础上变化。这允许对量化的DC值不同地编码而不必进行反缩放操作。即使在AC（低通和高通）频带用变化的量化编码的时候也用相对较小的量化器来对图像瓦块的DC频带编码不会显著影响比特率。

在一个极端，帧内的所有变换系数使用相同的量化参数。在另一极端，用于所有通道的低通和高通量化规则都允许对瓦块/帧的每一宏块独立地变化。唯一的限制是不同低通和高通量化器规则（覆盖所有通道）的数目各自被限于16。

每一这样的规则可以为每一通道指定独立的量化参数值。

在这些极端之间，允许如由图4-14所示的句法表指定的若干组合。

7.量化器参数的索引：代表性编码器/解码器中的特定量化参数（QP）基于谐波比例。量化器参数索引（QPI）的8位值对应于可能相当大的QP值。执行第二级索引使得跨宏块变化的QPI可以用高效的方式来编码。

更具体而言，编码器200可以定义比特流中包含1到16个QPI“矢量”的集合。基于选择哪一XXX_CHANNEL_MODE，每一QPI矢量由一个或多个QPI值组成。基于频率子带开关，这些集合是为DC、低通和高通子带定义的。

此外，DC集合仅有一个QPI矢量，因为在瓦块通道中只允许一个DC量化器。

pass macroblock) and Qualcomm sub-band (high pass macroblock).

In the representative system, the low-pass and / or high-pass sub-band in the bit stream is optionally one of these sub-bands can be completely discarded.

In addition, according to the order in one of two compressed data is compressed into a bit stream: spatial order and frequency order. For the spatial order, different from the same sub-macroblock within a tile are ordered together with, and the resulting bitstream of each tile is written into one packet. For frequency order, from within the same sub-tiles with different macroblock are grouped together, and thus the tile bitstream is written in three groups: DC tile packet, the low-pass and high-pass grouping tile tile grouping. Further, there may be other data layers.

Thus, for the representative system, an image according to the following "dimensions" to organize: Dimensions: Frame - tile - macro block; frequency dimension: DC | lowpass | Qualcomm; and channel dimensions: brightness | chroma -0 | _1... chroma (e.g., Y | U | V). more arrows indicate the hierarchy, but the vertical bar indicates division.

Flexible Quantization

Although the representative system in accordance with the spatial, frequency and channel dimensions to organize the compressed digital media data, but described here can be applied to replace along fewer, additional or other dimensions to organize their data The encoder / decoder system. For example, the flexible quantization approach can be applied using a larger number of frequency bands, other format of color channels (e.g., YIQ, RGB, etc.), additional image channels (e.g., for

对这些集合的编码在图4-6所示的表中定义。

如图7-11的表所示，对DC、低通和高通频率子带的QPI矢量集合的信号表示如下发生。基于其它编码模式，每一集合的势（即，该集合中的QPI矢量数）对低通和高通子带在相应的瓦块或帧的开始处指示。DC集合的势为1。

在该伪代码表中，表示势的句法元素被标记为“XXX—QUANTIZERS”。（在实践中，在比特流中发送XXX—QUANTIZERS—1）。表中标记为“XXX_QUANTIZER”的句法元素表示QPI集合的编码，这在图4-6所示的表中定义。

在宏块级，仅发送来自QPI集合内的所需QPI矢量的索引QI就足够。图12-14的表定义了宏块的基础上发送QI的句法。对应于QI的句法元素被标记为“XXX—QUANTIZER—INDEX”。使用可变长度代码来用信号表示QI。

首先，发送指示QI是否为0的一位码元。如果不是，则发送指示与0不同的特定QI的其长度由 $\text{ceil}(\log_2(\text{XXX\_QUANTIZERS} - 1))$ 给出的固定长度代码。

这允许用低至每宏块一位来对“默认”量化规则（QI=0）进行高效编码。当XXX—QUANTIZERS为1时，XXX_QUANTIZER_INDEX唯一地为0，且因此不需要用信号表示QI。

8. 扩展 以上对灵活量化的描述是专用于其在代表性编码器和解码器中的实现以及句法的。然而，该技术的原理也可扩展到其它数字媒体压缩系统和格式。例如，代表性编码器/解码器只有三个频率子带（DC、低通和高通）。但是，更一般地，灵活量化的替换实现可以用直接的方式扩展到多个频率子带。同样，替换灵活量化实现可以在更细的空间粒度改变量化器，诸如通过在子宏块（诸如块）级发送量化索引（QI）信息。对灵活量化技术的基本原理的许多扩展在同一框架内都是可能的。

stereo vision or other multiple camera arrays) encoding.

2. Flexible quantization reviewed in this representative encoder / decoder, quantization operation is essentially in accordance with the quantization parameter QP plus partial division, which is executed at the encoder. Inverse quantization or multiplication operation is multiplication and QP, which is executed at the decoder. However, the replacement of the flexible quantization described herein can be used to achieve other forms of quantization, including uniform and non-uniform, scalar or vector, with or without dead zone. The quantization / inverse quantization process jointly introduced the original transform coefficient data loss, which showed the decoded image compression errors or artifacts. In the most simple codecs can be made to an all transform coefficients in the application-specific fixed QP value. While this may be an acceptable solution in some cases, but it also has several disadvantages: the human visual system for all frequencies, or to all spatial positions within one frame, or to all luminance and chrominance channels are not It is equally sensitive. Using different QP values for different coefficients provided even with the same or lesser number of compressed bits more visually superior encoding. Similarly, it may be appropriate to optimize their error metric.

Ability

rate control or encoder to produce the desired size of the compressed file will not be easy to implement on the entire frame with a single QP.

Thus, ideally, it should be possible to allow the encoder in an arbitrary manner across the image change QP. However, this means that for each data partition (macro block / pad / channel / sub-band,

9. 计算环境 上述用于灵活量化的处理技术可以在各种数字媒体编码和/或解码系统的任一种上实现，包括计算机（各种形状因数，包括服务器、台式机、膝上型计算机、手持式计算机等）；数字媒体记录器和播放器；图像和视频捕捉设备（诸如照相机、扫描仪等）；通信设备（诸如电话、移动电话、会议设备等）；显示、打印或其它呈现设备；以及其它示例等等。该灵活量化技术可用硬件电路、控制数字媒体处理硬件的固件、以及在计算机或在其他计算环境中执行的通信软件来实现，诸如图15中所示。

图15示出了其中可实现所描述的实施例的合适计算环境（1500）的一个一般示例。计算环境（1500）并不旨在对本发明的使用范围或功能提出任何局限，因为本发明可在不同的通用或专用计算环境中实现。

参考图15，计算环境（1500）包括至少一个处理单元（1510）和存储器（1520）。在图15中，这一最基本的配置（1530）被包括在虚线内。处理单元（1510）执行计算机可执行指令，并且可以是真实或虚拟处理器。在多处处理系统中，多个处理单元执行计算机可执行指令以提高处理能力。存储器（1520）可以是易失性存储器（例如，寄存器、高速缓存、RAM）、非易失性存储器（例如，ROM、EEPROM、闪存等）或两者的某种组合。存储器（1520）储存实现所描述的使用灵活量化技术的数字媒体编码/解码的软件（1580）。

计算环境可具有额外的特征。例如，计算环境（1500）包括存储（1540）、一个或多个输入设备（1550）、一个或多个输出设备（1560）以及一个或多个通信连接（1570）。诸如总线、控制器或网络等互连机制（未示出）将计算环境（1500）的各组件互连。通常，操作系统软件（未示出）为在计算环境（1500）中执行的其它软件提供操作环境，并协调计算环境（1500）的各组件的活动。

存储（1540）可以是可移动或不可移动的，并包括磁盘、磁带或磁带盒、CD-ROM、CD-RW、DVD或可用于储存信息并可在计算环境（1500）内访问的任何其

etc.) of the actual QP value should be expressed in the bit stream by a signal. This leads to huge Jian pin, just to carry the QP signal indicating information, thereby leading to its in practice is not appropriate. What is needed is a flexible means of signaling QP, but saves bits, especially for situations frequently encountered. flexible quantization technique described herein provides various partitions or dimensions along the encoded digital media data to change the quantization capability. For example, one implementation of the representative encoder 200 / decoder 300 system flexible quantization technique can vary quantization one in (i) spatial location, (ii) frequency sub bands, and (iii) color channels in three dimensions on. However, the flexible quantization technique other alternative implementations, the quantization may vary on fewer, more or other dimensions or packet data. The technology also includes the encoded media data efficiently signal the flexible quantization approach. The advantage of this method is that involving quantification of auxiliary information to quantify the overhead introduced for major use cases is minimized, and if the encoder is needed allows maximum flexibility.

The flexible quantization technique provides fine granularity control over the space quantization. In one particular implementation, the flexible quantization applied to frame allows quantization control pad or down to the macro block. If the frame is not uniform quantization, then each tile can be uniformly quantized; if the pad is not uniform quantization, it will be different for each macro block quantized.

The flexible quantization also allows quantization control along the frequency sub-band dimensions. In one particular implementation, the flexible quantization

它介质。存储 (1540) 储存实现所描述的使用灵活量化技术的数字媒体编码/解码的软件 (1580) 的指令。

输入设备 (1550) 可以是诸如键盘、鼠标、笔或跟踪球等触摸输入设备、语音输入设备、扫描设备或向计算环境 (1500) 提供输入的另一设备。对于音频, 输入设备 (1550) 可以是声卡或接受来自话筒或话筒阵列的模拟或数字形式的音频输入的类似设备, 或向计算环境提供音频样本的 CD-ROM 读取器。

输出设备 (1560) 可以是显示器、打印机、扬声器、CD 刻录机或提供来自计算环境 (1500) 的输出的另一设备。

通信连接 (1570) 允许在通信介质上与另一计算实体的通信。通信介质传达诸如计算机可执行指令、压缩音频或视频信息、或已调制数据信号中的其它数据的信息。已调制数据信号是其一个或多个特征以在信号中编码信息的方式设置或改变的信号。作为示例而非局限, 通信介质包括以电、光、RF、红外、声学或其它载波实现的有线或无线技术。

此处所描述的使用灵活量化技术的数字媒体编码/解码可在计算机可读介质的一般上下文中描述。计算机可读介质可以是可在计算环境内访问的任何可用介质。作为示例而非局限, 对于计算环境 (1500), 计算机可读介质可包括存储器 (1520)、存储 (1540)、通信介质和以上任一种的组合。

此处描述的使用灵活量化技术的数字媒体编码/解码可在诸如程序模块中所包括的、在目标真实或虚拟处理器上的计算环境中执行的计算机可执行指令的一般上下文中描述。一般而言, 程序模块包括执行特定任务或实现特定抽象数据类型的例程、程序、库、对象、类、组件、数据结构等。程序模块的功能可以如各实施例中所需的组合或在程序模块之间分离。用于程序模块的计算机可执行指令可以在本地或分布式计算环境中执行。

出于表示的目的, 详细描述使用了如“确定”、“生成”、“调整”和“应用”等术语来描

including sub-band model to quantify the relationship between the specified frequency sub-bands of. Sub-band may be a uniform amount of, or part-uniform quantization (lowpass subband using DC sub-band quantization, and / or high-pass sub-band quantization using a low-pass filter), or independently quantified.

The flexible quantization also allows data to quantify the dimension along the channel of control applications. In one particular implementation, the flexible quantization includes quantitative relationship between the channels specified color channel mode. Channel can be uniformly quantized, or partially uniform quantization (uniform color channel, but the brightness channel independent), or by an independent quantified. flexible quantization described herein also provides efficient in the compressed digital media data signal representing auxiliary information by a combination of the above-described technique for quantizing spatial, frequency sub band and channel control, which is important for the main use cases of. Further, the flexible quantization technique provides via the digital media data index from a set of defined subsets of quantizer may be effectively define the quantizer selection mode.

3 spatial dimensions of the flexible quantization: the spatial dimension, the representative encoder / decoder provided by the flexible quantization technique three options: □ entire frame can be used to encode the same quantization rule.

Otherwise, the entire tile can be used to encode the same quantization rule, and the frame is different tile can use different quantization rules.

述计算环境中的计算机操作。这些术语是由计算机执行的操作的高级抽象，且不应与人类所执行的动作混淆。对应于这些术语的实际的计算机操作取决于实现而不同。

鉴于可应用本发明的原理的许多可能的实施例，要求保护落入所附权利要求书及其等效技术方案的范围和精神之内的所有这样的实施例作为本发明。

. Otherwise, the tiles each macroblock may be used to encode the same quantization rule, and the tile in the different macroblocks can use different quantization rules. means with a signal representing these possibilities are as follows: the bit stream of a frame-level transmission instruction of whether the first possibility is true binary signal. If not, then the bit stream of each tile in the tile is sent to the number of quantization rule used fixed length symbols indicating. If you use more than one tile quantization rules, within each tile corresponding macroblock send an indication of the macroblock quantization rule by the use of variable length symbols. Decoder manner consistent with the encoder interprets the bit stream. representative encoder 200 / decoder 300 uses a variant thereof represented by the signal. Here marked by the binary signal "XXX_FRAME-UNIFORM" common syntactic element represents only at the frame level transmission (where XXX is a placeholder specifies the quantization control specific frequency sub-band or channel dimensions). In the tile level, the number of frame level syntax elements (XXX_FRAME_UNIFORM) is false transmit different quantization rules in the tile-level syntax element (XXX-QUANTIZERS) only when. If this number is equal to 1, it means that there is only one rule and therefore all macroblocks with the same quantization rule to uniformly encode the tile inside (indicating selection 2), and if not, it indicates selection of the third possibility.

4 cross-band flexible quantization: For cross-band flexible quantization, representative encoder bitstream syntax 200 / decoder 300 defines two switches: a low pass macroblock uses the same space at a position DC Macro the same quantization rule block. This corresponds

to the syntax element USE_DC-QUANTIZER.

. Qualcomm macroblock uses the same space at the position the same low pass macroblock quantization rules. This corresponds to the syntax element USE_LP_QUANTIZER.

These switches across the frame using the same quantization rule enabled in the frame layer, or in other cases to enable the tile floor. These switches are not the macroblock layer is enabled. Therefore, all macro blocks within tile across the frequency sub-band abide by the same rules. Suitable (frame or tile) layer at sending a binary symbols in each switch.

5 Cross-channel flexible image quantization: for flexible cross-channel quantization, representative encoder 200 / decoder bitstream syntax 300 defines three options: a luminance and chrominance all channels have the same quantization rule. This is the common syntax element XXX CH MODE == CH_UNIFORM indicated.

. Brightness follow a quantization rule, and all color channels follow a different quantization rules, indicated by XXX CH MODE == CH_MIXED.

. All channels are free to choose different quantization rules, indicated by XXX-CH a MODE == CH_INDEPENDENT.

6 flexible quantization combination of: a representative encoder 200 / decoder 300 using the bit stream syntax defined in the following figures the code shown in Table 4-14, the syntax can be described above efficiently encoded interdimensional flexible quantization options specific selection. Every space with a cross, the number of permutations frequency sub-band and channel

dimensions several quantization options available, the options available to quantify is great. The fact that the bit stream 200 representative encoder / decoder 300 can be used to sort the layout space or frequency increases across the three dimensions of the flexible quantization complexity. However, this does not change the available quantization options, and only affects the serialization of the signal. Figure 4-14 definition syntax provides a flexible combination of efficient quantization rule coding.

Some prominent features quantization rules

As representative encoder / decoder combination syntax defined below. allowed DC to quantify changes in the macro-block basis. This allows the quantized DC value coded differently without anti-scaling operation. Even in AC (low-pass and high-pass) band with varying quantization and coding, they also use a relatively small quantizer to DC band coding image does not significantly affect the tile bit rate.

Use the same quantization parameter in all the transform coefficients an extreme frame. At the other extreme, for low-pass and high-pass quantization rule for all channels are allowed for each macro-tile / frame blocks vary independently. The only restriction is that different low-pass and high-pass quantization rules (covering all channels) is limited to 16 the number of each.

Such rules can be specified for each separate quantization parameter values for each channel.

In between these extremes, allow several combinations such as syntax table shown in Figure 4-14 specify.

7 quantization parameter index:
 representative encoder / decoder specific
 quantization parameter (QP) based on
 the harmonic ratio. Quantization
 parameter index (QPI) 8-bit value
 corresponding to the QP value may be
 quite large. Performing a second inter-
 level index so that changes QPI
 macroblock can be encoded in an
 efficient manner.

More specifically, the encoder 200 can
 define the bitstream contains 1-16 QPI
 set "vector" in. Based choose which
 XXX_CHANNEL_MODE, each QPI QPI
 vector by one or more values. Based on
 frequency sub-band switch, these
 collections for DC, low-pass and high-
 pass sub-band defined.

In addition, DC set only one QPI vector,
 because the tile channel only allows a
 DC quantizer.

These codes define table sets shown in
 Figure 4-6.

As shown in Table 7-11, QPI vector
 signal DC, low-pass and high-pass
 frequency sub-band of the collection
 follows happen. Based on other coding
 modes, each set of potential (ie, the set
 of QPI vector number) for low-pass and
 high-pass sub-band at the beginning of
 the respective tile or frame indication.
 Potential DC set to 1.

In this pseudo-code table, showing the
 potential of syntax elements are marked
 as "XXX-QUANTIZERS". (In practice,
 send XXX-QUANTIZERS in the bitstream
 --1). Table labeled "XXX_QUANTIZER"
 syntactic element represents a set of
 coding QPI, which is defined in the table
 shown in Figure 4-6.

In the macroblock level, sending only
 index QI from the desired QPI QPI vector
 within the collection is sufficient. Figure
 Table 12-14 defines the basis of sending

First, sending indicates whether a symbol QI is zero. Fixed length code given -; if not, 0 is sent indicating the specific QI different from its length by a ceil (1) (XXX-QUANTIZERS lo and amp).

8. Expansion of the flexible quantization described above is dedicated to achieve and the syntax of the representative encoder and decoder thereafter. However, the principles of the technology can also be extended to other digital media compression systems and formats. For example, the representative encoder / decoder only three frequency sub-bands (DC, low-pass and high-pass). However, more generally, to replace the flexible quantization implementations can be extended by a straightforward manner to a plurality of frequency subbands. Similarly, replacing the flexible quantization can be changed to achieve finer quantization in the space size, (QI) information in sub-macro block (such as a block) level, such as by sending the quantization index. The basic principle of the flexible quantization technique many extensions are possible within the same framework.

9/28/2017

servers, desktops, laptops computers, handheld computers, etc.); digital media recorders and players; image and video capture device (such as a camera, scanner Xin, etc.); communications equipment (such as telephone, mobile phone, conference facilities, etc.); display, print or other rendering device; and the like, and other examples. The flexible quantization techniques can be hardware circuitry, controlling digital media processing hardware, firmware, and software executing on a computer or a communication in other computing environment to implement, such as shown in Figure 15.

FIG. 15 shows a realization in which the described embodiment of a suitable computing environment (1500) is a general example. Computing Environment (1500) is not intended scope of use or functionality of the invention make any limitations, since the present invention may be implemented in diverse general-purpose or special-purpose computing environments.

Referring to Figure 15, the computing environment (1500) includes at least one processing unit (1510) and a memory (1520). In Figure 15, this most basic configuration (1530) is included within a dashed line. Processing unit (1510) executes computer-executable instructions, and may be real or virtual processor. In a multi-processing system, multiple processing units execute computer-executable instructions to increase processing power. Memory (1520) or some combination of the two may be volatile memory (e.g., registers, cache, RAM), nonvolatile memory (e.g., ROM, EEPROM, flash memory, etc.). Memory (1502) storing achieve the described flexible quantization technique using digital media encoding / decoding software (1580).

A computing environment may have additional features. For example, the computing environment (1500) includes storage (1540), one or more input devices (1550), one or more output devices (1560) and one or more communication connections (1570). Interconnection mechanism such as a bus, controller, or network (not shown) in the computing environment (1500) of each component interconnect. Typically, operating system software (not shown) to other software executing in the computing environment (1500) is provided in the operating environment, and coordinates activities of the computing environment (1500) of each component.

Storage (1540) may be removable or non-removable, and includes disks, magnetic tapes or cassettes, CD-ROM, CD-RW, DVD, or can be used to store information within the computing environment (1500) visit any other medium. Storage (1540) stores implement the described technique using flexible quantization digital media encoding / decoding software (1580) of the directive.

Input Devices (1550) may be such as a keyboard, mouse, trackball or touch pen input devices, voice input device, a scanning device, or to the computing environment (1500) provides another device input. For audio, the input device (1550) may be a sound card or similar device accepts either analog or digital microphone array from a microphone or audio input in the form of, or provide audio sample CD-ROM reader to the computing environment. output device (1560) may be a display, printer, speakers, CD burner or provide output from the computing environment (1500) of another device. communication connection (1570) allows communication over a communication medium to another

computing entity. Convey the communication medium such as a computer-executable instructions, compressed audio or video information, or other data in a modulated information data signals. Modulated data signal that has one or more features in a manner to encode information in the signal set or changed signal. By way of example and not limitation, communication media includes electrical, optical, RF, infrared, acoustic, or other carrier wired or wireless techniques implemented. using the flexible quantization techniques described herein digital media encoding / decoding can be described in the general context of computer-readable media. Computer readable media can be any available media can be accessed within a computing environment. By way of example and not limitation, with the computing environment (1500), computer-readable medium may include, storage (1540), communication media, and combinations of any of the above memory (1520).

Use flexible quantization techniques described herein digital media encoding / decoding may be included, such as program modules, the computer-implemented computing environment in the target real or virtual processor executable instructions described in the general context. Generally, program modules include that perform particular tasks or implement particular abstract data types of routines, programs, libraries, objects, classes, components, data structures, and so on. Functionality of the program modules may be combined as required in each embodiment or embodiments the separation between program modules. Program modules for computer-executable instructions may be executed within a local or distributed computing environment.

For the purposes of representation, a detailed description of use, such as "OK", "Build", "adjustment" and "applications" and other terms used to describe computer operations in a computing environment. These terms are Gao level abstract operation executed by a computer, and should not be confused with the operation performed by humans. These terms correspond to actual computer operations vary depending on implementation.

In view of the many possible embodiments may apply the principles of the present invention, it claimed the appended claims and their equivalents all such embodiments scope and spirit of technical solutions within the present invention.

Cited by: CN102638681 A ; CN103782293 A ; CN104126306 A ; CN104126306 B ; CN105120285 A ; US9667958 BB ; WO13064109 A1 ;

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.

Request for Continued Examination (RCE) Transmittal

Address to:
Mail Stop RCE
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Application Number	13/389,872
Filing Date	February 10, 2012
First Named Inventor	XU, et al.
Art Unit	7293
Examiner Name	Maryam A. Nasri
Attorney Docket Number	PU090108 US PCT

This is a Request for Continued Examination (RCE) under 37 CFR 1.114 of the above-identified application.

Request for Continued Examination (RCE) practice under 37 CFR 1.114 does not apply to any utility or plant application filed prior to June 8, 1995, or to any design application. See Instruction Sheet for RCEs (not to be submitted to the USPTO) on page 2.

- Submission required under 37 CFR 1.114** Note: If the RCE is proper, any previously filed unentered amendments and amendments enclosed with the RCE will be entered in the order in which they were filed unless applicant instructs otherwise. If applicant does not wish to have any previously filed unentered amendment(s) entered, applicant must request non-entry of such amendment(s).
 - ☐ Previously submitted. If a final Office action is outstanding, any amendments filed after the final Office action may be considered as a submission even if this box is not checked.
 - ☐ Consider the arguments in the Appeal Brief or Reply Brief previously filed on _____
 - ☐ Other _____
 - ☒ Enclosed
 - ☒ Amendment/Reply
 - ☐ Affidavit(s)/ Declaration(s)
 - ☒ Information Disclosure Statement (IDS)
 - ☐ Other _____
- Miscellaneous**
 - ☐ Suspension of action on the above-identified application is requested under 37 CFR 1.103(c) for a period of _____ months. (Period of suspension shall not exceed 3 months; Fee under 37 CFR 1.17(i) required)
 - ☐ Other _____
- Fees** The RCE fee under 37 CFR 1.17(e) is required by 37 CFR 1.114 when the RCE is filed. The Director is hereby authorized to charge the following fees, any underpayment of fees, or credit any overpayments, to Deposit Account No. 07-0832.
 - ☒ RCE fee required under 37 CFR 1.17(e)
 - ☒ Extension of time fee (37 CFR 1.136 and 1.17)
 - ☐ Other _____
 - ☐ Check in the amount of \$ _____ enclosed
 - ☐ Payment by credit card (Form PTO-2038 enclosed)

WARNING: Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.

SIGNATURE OF APPLICANT, ATTORNEY, OR AGENT REQUIRED

Signature	/Ronald J. Kolczynski/	Date	2017-10-26
Name (Print/Type)	Ronald J. Kolczynski	Registration No.	57,177

CERTIFICATE OF MAILING OR TRANSMISSION

I hereby certify that this correspondence is being deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to: Mail Stop RCE, Commissioner for Patents, P. O. Box 1450, Alexandria, VA 22313-1450 or facsimile transmitted to the U.S. Patent and Trademark Office on the date shown below.

Signature		Date	
Name (Print/Type)			

This collection of information is required by 37 CFR 1.114. 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: Mail Stop RCE, 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.

Instruction Sheet for RCEs

(not to be submitted to the USPTO)

NOTES:

An RCE is not a new application, and filing an RCE will not result in an application being accorded a new filing date.

Filing Qualifications:

The application must be a utility or plant application filed on or after June 8, 1995. The application cannot be a provisional application, a utility or plant application filed before June 8, 1995, a design application, or a patent under reexamination. See 37 CFR 1.114(e).

Filing Requirements:

Prosecution in the application must be closed. Prosecution is closed if the application is under appeal, or the last Office action is a final action, a notice of allowance, or an action that otherwise closes prosecution in the application (e.g., an Office action under *Ex parte Quayle*). See 37 CFR 1.114(b).

A submission and a fee are required at the time the RCE is filed. If reply to an Office action under 35 U.S.C. 132 is outstanding (e.g., the application is under final rejection), the submission must meet the reply requirements of 37 CFR 1.111. If there is no outstanding Office action, the submission can be an information disclosure statement, an amendment, new arguments, or new evidence. See 37 CFR 1.114(c). The submission may be a previously filed amendment (e.g., an amendment after final rejection).

WARNINGS:

Request for Suspension of Action:

All RCE filing requirements must be met before suspension of action is granted. A request for a suspension of action under 37 CFR 1.103(c) does not satisfy the submission requirement and does not permit the filing of the required submission to be suspended.

Improper RCE will NOT toll Any Time Period:

Before Appeal - If the RCE is improper (e.g., prosecution in the application is not closed or the submission or fee has not been filed) and the application is not under appeal, the time period set forth in the last Office action will continue to run and the application will be abandoned after the statutory time period has expired if a reply to the Office action is not timely filed. No additional time will be given to correct the improper RCE.

Under Appeal - If the RCE is improper (e.g., the submission or the fee has not been filed) and the application is under appeal, the improper RCE is effective to withdraw the appeal. Withdrawal of the appeal results in the allowance or abandonment of the application depending on the status of the claims. If there are no allowed claims, the application is abandoned. If there is at least one allowed claim, the application will be passed to issue on the allowed claim(s). See MPEP 1215.01.

See MPEP 706.07(h) for further information on the RCE practice.

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.

CUSTOMER NO.: 24498

Attorney Docket: PU090108

Serial No.: 13/389,872

Final Office Action dated: June 30, 2017

Response: October 30, 2017 with Request for Continued Examination

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of: Qian Xu et al.

Serial No.: 13/389,872

Filed: February 10, 2012

For: METHODS AND APPARATUS FOR
IMPROVED INTRA CHROMA ENCODING
AND DECODING

Examiner: Nasri, Maryam

Group Art: 2483

AMENDMENT UNDER 37 C.F.R. §1.114

Mail Stop RCE

Commissioner for Patents

P.O. Box 1450

Alexandria, VA 22313-1450

Sir:

Please enter and consider the following amendments and remarks.

Amendments to, and Listing of, the Claims begins on page 2 of this paper.

Remarks/Arguments begin on page 8 of this paper.

Amendments to, and Listing of, the Claims

This listing of claims will replace all prior versions, and listings, of claims in the application:

1. (Currently Amended) An apparatus, comprising:

an encoder for encoding picture data for at least a block in a picture, wherein multiple partition types are supported for intra chroma coding of the block, the multiple partition types including a set of chroma partition types and a set of luma partition types, the set of chroma partition types being different than the set of luma partition types, ~~and wherein a particular chroma partition type for coding the block is determined from the multiple partition types independently within a macroblock from luma partition type,~~ and wherein partition type refers to the size and number of partitions within a macroblock, wherein said encoder selects a the best luma partition and luma mode, selecting a best chroma partition and selects best chroma mode and encodesing said block using the luma partition type, luma mode and chroma mode, ~~and wherein chroma partition type is differentially coded from luma partition type.~~

2. (Currently Amended) In a video encoder, a method, comprising:

encoding picture data for at least a block in a picture, wherein multiple partition types are supported for intra chroma coding of the block, the multiple partition types including a set of chroma partition types and a set of luma partition types, the set of chroma partition types being different than the set of luma partition types, ~~and wherein a particular chroma partition type for coding the block is determined from the multiple partition types independently within a macroblock from luma partition type, and wherein chroma partition type is differentially coded from luma partition type,~~ and wherein partition type refers to the size and number of partitions within a macroblock;

selecting a luma partition and luma mode;

selecting a chroma partition;

selecting a chroma mode; and,

encoding said block using the luma partition type, luma mode and chroma mode

~~parsing a luma partition;~~
~~setting a chroma partition;~~
~~parsing a luma mode, parsing a chroma mode, and~~
~~decoding a superblock.~~

3. (Previously Presented) The method of claim 2, wherein a particular chroma partition type for coding the block is determined from the set of chroma partition types responsive to a luma partition type utilized to code the block or one or more neighboring blocks, the luma partition type being included in the set of luma partition types.

4. (Original) The method of claim 3, wherein for the particular chroma partition type, a largest-size available transform matching the particular chroma partition type is selected from a set of transforms.

5. (Previously Presented) The method of claim 3, wherein from among a particular luma partition type and a particular chroma partition type selected to code the block only the particular luma partition type is signaled, the particular luma partition type being selected from the set of luma partition types and the particular chroma partition type being selected from the set of chroma partition types.

6. (Previously Presented) The method of claim 5, wherein the particular luma partition type is absolutely coded or differentially coded from one or more neighboring blocks.

7. (Previously Presented) The method of claim 2, wherein a particular chroma partition type used to code the block is determined independently from a particular luma partition type used to code the block, the particular chroma partition type being determined from the set of chroma partition types and the particular luma partition type being determined from the set of luma partition types.

8. (Previously Presented) The method of claim 7, wherein the particular chroma partition type and the particular luma partition types are both signaled.

9. (Previously Presented) The method of claim 8, wherein the particular luma partition type and the particular chroma partition type are absolutely coded or differentially coded from one or more neighboring blocks.

10. (Previously Presented) The method of claim 8, wherein the particular chroma partition type is differentially coded from the particular luma partition type.

11. (Original) The method of claim 2, wherein a chroma entropy coding engine used to code the block is different from a luma entropy coding engine used to code the block.

12. (Currently Amended) An apparatus, comprising:
a decoder for decoding picture data for at least a block in a picture, wherein multiple partition types are supported for intra chroma decoding of the block, the multiple partition types including a set of chroma partition types and a set of luma partition types, the set of chroma partition types being different than the set of luma partition types, ~~and wherein a particular chroma partition type for coding the block is determined from the multiple partition types independently within a macroblock from luma partition type, and wherein chroma partition type is differentially coded from luma partition type,~~ and wherein partition type refers to the size and number of partitions within a macroblock; and

an intra prediction module comprising multi-partition support for chroma decoding configured to perform parsing a luma partition to determine a luma partition type, setting a chroma partition, parsing a luma mode, parsing a chroma mode, and decoding a superblock.

13. (Currently Amended) In a video decoder, a method, comprising:
decoding picture data for at least a block in a picture, wherein multiple partition types are supported for intra chroma decoding of the block, the multiple partition types including a set of chroma partition types and a set of luma partition types, the set of chroma partition types being

different than the set of luma partition types, and wherein a particular chroma partition type for coding the block is determined from the multiple partition types independently within a macroblock from luma partition type, and wherein chroma partition type is differentially coded from luma partition type, and wherein partition type refers to the size and number of partitions within a macroblock;

parsing a luma partition to determine a luma partition type;

setting a chroma partition;

parsing a luma mode, parsing a chroma mode, and

decoding a superblock.

14. (Previously Presented) The method of claim 13, wherein a particular chroma partition type for decoding the block is determined from the set of chroma partition types responsive to a luma partition type utilized to decode the block or one or more neighboring blocks, the luma partition type being included in the set of luma partition types.

15. (Original) The method of claim 14, wherein for the particular chroma partition type, a largest-size available transform matching the particular chroma partition type is selected from a set of transforms.

16. (Previously Presented) The method of claim 14, wherein from among a particular luma partition type and a particular chroma partition type selected to code the block only information indicating the particular luma partition type is received from a corresponding encoder for use in decoding the block, the particular luma partition type being selected from the set of luma partition types and the particular chroma partition type being selected from the set of chroma partition types.

17. (Previously Presented) The method of claim 16, wherein the particular luma partition type is absolutely decoded or differentially decoded from one or more neighboring blocks.

18. (Previously Presented) The method of claim 13, wherein a particular chroma partition type used to code the block is determined independently from a particular luma partition type used to code the block, the particular chroma partition type being determined from the set of chroma partition types and the particular luma partition type being determined from the set of luma partition types.

19. (Previously Presented) The method of claim 18, wherein information indicating both the particular chroma partition type and the particular luma partition types is received from a corresponding encoder for use in decoding the block.

20. (Previously Presented) The method of claim 19, wherein the particular luma partition type and the particular chroma partition type are absolutely decoded or differentially decoded from one or more neighboring blocks.

21. (Previously Presented) The method of claim 19, wherein the particular chroma partition type is differentially decoded from the particular luma partition type.

22. (Original) The method of claim 13, wherein a chroma entropy decoding engine used to decode the block is different from a luma entropy decoding engine used to decode the block.

23. (Currently Amended) A non-transitory computer-readable storage media having video signal data encoded thereupon, comprising:

picture data encoded for at least a block in a picture, wherein multiple partition types are supported for intra chroma coding of the block, the multiple partition types including a set of chroma partition types and a set of luma partition types, the set of chroma partition types being different than the set of luma partition types, ~~and wherein a particular chroma partition type for coding the block is determined from the multiple partition types independently within a macroblock from luma partition type, and wherein chroma partition type is differentially coded from luma partition type, and wherein partition type refers to the size and number of partitions~~

CUSTOMER NO.: 24498

Attorney Docket: PU090108

Serial No.: 13/389,872

Final Office Action dated: June 30, 2017

Response: October 30, 2017 with Request for Continued Examination

within a macroblock and, wherein an encoder has selected a ~~best~~ luma partition and a ~~best~~ luma mode, selectsing a ~~best~~ chroma partition and selects ~~best~~ chroma mode and has encoded said block using the luma partition type, luma mode and chroma mode.

REMARKS

This communication is being submitted in response to the Final Office Action mailed June 30, 2017, for which a shortened statutory period of three months ending September 30, 2017 was set, and for which a one-month extension is hereby requested. A Request for Continued Examination is submitted herein. The Office Action has been reviewed and carefully considered. The claims have been amended. No new matter has been added. Reconsideration of the prior rejections in light of the timely filing of the instant remarks and amendments is respectfully requested. Please charge any necessary fees to deposit account 07-0832.

Claims 1-23 are pending.

Rejection of Claims 1-23 under 35 U.S.C. §103

Claims 1-23 are rejected under 35 U.S.C. §103(a) as being unpatentable over Wang (US 2007/0121728 a1) in view of U.S. Pat. App. Pub. No 2008/0219350 to Guo (hereinafter “Guo”).

The independent claims currently pending are Claims 1, 2, 12, 13, and 23. These claims have been amended.

Amended Claim 1 recites, *inter alia*, “an encoder for encoding picture data for at least a block in a picture, wherein multiple partition types are supported for intra chroma coding of the block, the multiple partition types including a set of chroma partition types and a set of luma partition types, the set of chroma partition types being different than the set of luma partition types, and wherein partition type refers to the size and number of partitions within a macroblock, wherein said encoder selects a luma partition and luma mode, selects a chroma partition and selects chroma mode and encodes said block using the luma partition type, luma mode and chroma mode.”

Claim 2 has been similarly amended.

Claim 13 has been amended to recite, *inter alia*, “decoding picture data for at least a block in a picture, wherein multiple partition types are supported for intra chroma decoding of the block, the multiple partition types including a set of chroma partition types and a set of luma partition types, the set of chroma partition types being different than the set of luma partition types, and wherein partition type refers to the size and number of partitions within a macroblock;

parsing a luma partition to determine a luma partition type; setting a chroma partition; parsing a luma mode, parsing a chroma mode, and decoding a superblock.”

Claim 12 has been similarly amended.

Wang discloses a method of optimizing decoding of MPEG4 compliant coded signals. The method disables processing for non-main profile sections, performs reference frame padding and adaptive motion compensation. The method further performs fast inverse DCT, but not on areas based on whether a 4x4 block may be all zero or only the DC transform coefficient is non-zero, and including CAVLC encoding of residual data. The reference frame padding comprises compensating for motion vectors extending beyond a reference frame by adding to at least the length and width of the reference frame. Adaptive motion compensation includes original blocks size compensation processing for chroma up to block sizes of 16x16.

The Examiner cites Wang as disclosing several aspects of the independent claims, such as multiple partition types for intra chroma coding of the block and, in particular, chroma partition type is determined from the multiple partition types independently from a luma partition type. However, the pending claims have been amended.

Applicant respectfully submits that Wang neither teaches nor suggests, “parsing a luma partition to determine a luma partition type; setting a chroma partition; parsing a luma mode, parsing a chroma mode, and decoding a superblock” as in amended Claims 12 and 13.

Applicant also respectfully submits that Wang does not disclose that an encoder selects a luma partition and luma mode, selects a chroma partition and selects chroma mode and encodes said block using the luma partition type, luma mode and chroma mode, as in amended Claims 1 and 2.

The Examiner also cites paragraph [211] of Wang for disclosing “intra chroma coding of the block.” However, Applicant respectfully submits that paragraph [211] of Wang neither teaches nor suggests that an encoder selects a luma partition and luma mode, selects a chroma partition and selects chroma mode and encodes said block using the luma partition type, luma mode and chroma mode.

Guo also does not disclose these features. Figure 1 of Guo shows “luma mode decision” and “chroma mode decision” but does not disclose “an encoder selects a luma partition and luma mode, selects a chroma partition and selects chroma mode and encodes said block using the luma partition type, luma mode and chroma mode” in any way. So, Applicant respectfully submits that Guo also does not disclose this additional feature.

Applicant again respectfully submits that at least the additional feature of the independent claims is neither taught nor suggested by any combination of Wang and Guo.

Reconsideration of the rejections is respectfully requested.

For at least the same reasons, the dependent claims are also respectfully submitted to be non-obvious over any combination of Wang and Guo.

In view of the foregoing, and the timely filing of the instant remarks, amendment, and a Request for Continued Examination, Applicants respectfully request that the rejections of the claims set forth in the Office Action of June 30, 2017 be withdrawn, that the pending claims be allowed, and that the case proceed to early issuance of Letters Patent in due course.

It is believed that a one-month extension fee and a fee for a Request for Continued Examination are currently due. However, in the event that any additional fees or charges are required at this time in connection with the application, they may be charged to applicants' Deposit Account No. 07-0832.

Respectfully submitted,

By: /Ronald Kolczynski/

Ronald Kolczynski, Attorney for Applicants
Registration No.: 57,177

Thomson Licensing LLC
Patent Operations
4 Research Way, 3rd Floor
Princeton, NJ 08540
(609) 734-6870

Dated: October 30, 2017

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)	Application Number	13389872
	Filing Date	2012-02-10
	First Named Inventor	XU et al.
	Art Unit	2483
	Examiner Name	NASRI, MARYAM A.
	Attorney Docket Number	PU090108-US-PCT

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	4063291		2013-13-12	GENNETTEN	

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	20080219530		2008-09-11	LEVANON 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	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	EP1513349	EP		2005-03-09	MICROSOFT CORP.		
	2	CN101222644	CN		2008-07-16	MITSUBISHI ELECTRIC CORP		×
	3	CN1440053	CN		2003-09-03	ZHEJIANG UNIV		☒

**INFORMATION DISCLOSURE
STATEMENT BY APPLICANT**
(Not for submission under 37 CFR 1.99)

Application Number		13389872
Filing Date		2012-02-10
First Named Inventor	XU et al.	
Art Unit	2483	
Examiner Name	NASRI, MARYAM A.	
Attorney Docket Number	PU090108-US-PCT	

4	CN1585495	CN	2005-02-23	SHANGHAI UNIV	☒
5	CN1784015	CN	2006-06-07	COMPUTING TECHNOLOGY INST	☒

If you wish to add additional Foreign Patent Document citation information please click the Add button

NON-PATENT LITERATURE 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, pages(s), volume-issue number(s), publisher, city and/or country where published.	T ⁵
	1		

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

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	13389872
Filing Date	2012-02-10
First Named Inventor	XU et al.
Art Unit	2483
Examiner Name	NASRI, MARYAM A.
Attorney Docket Number	PU090108-US-PCT

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	/Ronald J. Kolczynski/	Date (YYYY-MM-DD)	2017-10-26
Name/Print	Ronald J. Kolczynski	Registration Number	57,177

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.

Electronic Patent Application Fee Transmittal

Application Number:	13389872			
Filing Date:	10-Feb-2012			
Title of Invention:	METHODS AND APPARATUS FOR IMPROVED INTRA CHROMA ENCODING AND DECODING			
First Named Inventor/Applicant Name:	Qian Xu			
Filer:	Ronald J. Kolczynski/Brian Pritchett			
Attorney Docket Number:	PU090108			
Filed as Large Entity				
Filing Fees for U.S. National Stage under 35 USC 371				
Description	Fee Code	Quantity	Amount	Sub-Total in USD(\$)
Basic Filing:				
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD(\$)
Extension - 1 month with \$0 paid	1251	1	200	200
Miscellaneous:				
RCE- 2ND AND SUBSEQUENT REQUEST	1820	1	1700	1700
Total in USD (\$)				1900

Electronic Acknowledgement Receipt

EFS ID:	30797581
Application Number:	13389872
International Application Number:	
Confirmation Number:	7293
Title of Invention:	METHODS AND APPARATUS FOR IMPROVED INTRA CHROMA ENCODING AND DECODING
First Named Inventor/Applicant Name:	Qian Xu
Customer Number:	24498
Filer:	Ronald J. Kolczynski/Brian Pritchett
Filer Authorized By:	Ronald J. Kolczynski
Attorney Docket Number:	PU090108
Receipt Date:	30-OCT-2017
Filing Date:	10-FEB-2012
Time Stamp:	14:01:06
Application Type:	U.S. National Stage under 35 USC 371

Payment information:

Submitted with Payment	yes
Payment Type	DA
Payment was successfully received in RAM	\$ 1900
RAM confirmation Number	103117INTEFSW00000295070832
Deposit Account	
Authorized User	

The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows:

File Listing:					
Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Foreign Reference	CN1440053.pdf	195142	no	6
			597e68b088dca53511fd235e5a0157145743c4bb		
Warnings:					
Information:					
2	Non Patent Literature	CN1440053_TR_EN_machine_1 7OCT2017.pdf	375169	no	6
			638a0feb7d42172d442b3db14146971ca0506e2f		
Warnings:					
Information:					
3	Foreign Reference	CN1585495.pdf	640655	no	14
			639ea5cd2188689f177239a72de61658bba66a14		
Warnings:					
Information:					
4	Non Patent Literature	CN1585495_TR_EN_machine_1 7OCT2017.pdf	971822	no	17
			e9e239cb36e44fb82fe7bdb8b2e88b9c5ccb7f6		
Warnings:					
Information:					
5	Foreign Reference	CN1784015.pdf	783448	no	15
			54338043c02d74521688673febed241445e5e056		
Warnings:					
Information:					
6	Non Patent Literature	CN1784015_TR_EN_machine_1 7OCT2017.pdf	965446	no	19
			6545f9c23bfe92cf622d2b5387157bf8f78a9646		
Warnings:					
Information:					

7	Foreign Reference	CN101222644.pdf	3055979	no	59
			55ae9bc05ac93f7f8143aef19e5962c186b0ecdc		
Warnings:					
Information:					
8	Non Patent Literature	CN101222644_TR_EN_machine_17OCT2017.pdf	3449919	no	71
			d08af8a79160a3d28548152f53c1b6af56aa15ae		
Warnings:					
Information:					
9	Foreign Reference	EP1513349.pdf	1015670	no	16
			0bef1b37af27655f9ded9fa04d061fdca9224fb6		
Warnings:					
Information:					
10	Request for Continued Examination (RCE)	PU090108_US_PCT_RCE_signed_use.pdf	195440	no	3
			1cf1ec2ee8a2476d63a7b6b071651d52f176d1c0		
Warnings:					
This is not a USPTO supplied RCE SB30 form.					
Information:					
11		PU090108_US_PCT_RFOA.pdf	60530	yes	10
			73c6e79376b1cf1eb7b5d742e02549c665ad6990		
	Multipart Description/PDF files in .zip description				
	Document Description		Start	End	
	Amendment Submitted/Entered with Filing of CPA/RCE		1	1	
	Claims		2	7	
	Applicant Arguments/Remarks Made in an Amendment		8	10	
Warnings:					
Information:					
12	Information Disclosure Statement (IDS) Form (SB08)	PU090108_US_PCT_SB08_signed.pdf	612129	no	4
			033b08fd4f594000db2e48cd059f122bc4536758		

Warnings:					
Information:					
13	Fee Worksheet (SB06)	fee-info.pdf	32964	no	2
			07e691d43fcc89da8fa38a31095db529b6033d32		
Warnings:					
Information:					
Total Files Size (in bytes):			12354313		
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.					

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

H01L 21/203

C23C 14/08 C23C 14/34



[12] 发明专利申请公开说明书

[21] 申请号 03116007.7

[43] 公开日 2003 年 9 月 3 日

[11] 公开号 CN 1440053A

[22] 申请日 2003.3.25 [21] 申请号 03116007.7

[71] 申请人 浙江大学

地址 310027 浙江省杭州市西湖区玉古路 20 号

[72] 发明人 叶志镇 袁国栋 黄靖云 曾昱嘉
赵炳辉

[74] 专利代理机构 杭州求是专利事务所有限公司

代理人 韩介梅

权利要求书 1 页 说明书 2 页 附图 2 页

[54] 发明名称 一种 P-ZnO 薄膜及其制备方法

[57] 摘要

本发明的 p-ZnO 薄膜是以 N 为受主掺杂源, 以 Al 为施主掺杂源共掺杂的 p-ZnO 薄膜, 掺杂浓度为 $1.8 \times 10^{15} \text{cm}^{-3} - 4.5 \times 10^{18} \text{cm}^{-3}$ 。其制备方法是利用磁控溅射法, 以 $\text{Zn}_x\text{Al}_{1-x}$ 为靶材, ($1 > x > 0.6$), 掺杂浓度可以通过调节输入的含 N 气体和高纯 O_2 或 Ar 不同分压比和靶材中 Al 的含量来控制, 掺杂浓度可控性、稳定性好。设备简单, 操作方便, Al 原料丰富, 价格低廉。制作成本低。本发明利用施主-受主共掺生成的 N-Al-Np 型 ZnO 薄膜表面平整, 晶粒大小均匀, 具有择优取向。

ISSN 1000-8427 4

1. 一种 p-ZnO 薄膜, 其特征在于它是以 N 为受主, 以 Al 为施主共掺杂的 p-ZnO 薄膜, 掺杂浓度为 $1.8 \times 10^{15} \text{ cm}^{-3}$ – $4.5 \times 10^{18} \text{ cm}^{-3}$ 。

2. 根据权利要求 1 所述的 p-ZnO 薄膜的制备方法, 其特征是: 先将衬底表面清洗后放入直流反应磁控溅射装置的反应室中, 反应室真空度抽到 $1 \times 10^{-3} \text{ Pa}$, 然后加热衬底, 衬底温度为 $400\text{--}600^\circ\text{C}$, 以高纯含 N 气体和高纯 O_2 或 Ar 为溅射气体, 将二种气体分别由气体流量计控制输入装置的缓冲室, 在缓冲室充分混合后引入真空反应室, 在 $3\text{--}5 \text{ Pa}$ 压强下, 以 $\text{Zn}_x\text{Al}_{1-x}$ 为靶材, 式中 $1 > x > 0.6$, 进行溅射生长。

3. 根据权利要求 2 所述的 p-ZnO 薄膜的制备方法, 其特征是含 N 气体包括 NH_3 、 N_2O 、 NO 、 NO_2 。

4. 根据权利要求 2 所述的 p-ZnO 薄膜的制备方法, 其特征是所说的衬底是蓝宝石、硅。

一种 p-ZnO 薄膜及其制备方法

技术领域

本发明涉及半导体器件，具体说涉及 p-ZnO 薄膜及其制备方法。

背景技术

由于 ZnO 存在诸多的本征施主缺陷，如间隙锌 Zn_i 和氧空位 V_o ，其能级分别位于导带底 0.05eV 和 0.3eV 处，对受主产生高度自补偿作用。因此，很难实现 ZnO 的 p 型转变，而且，ZnO 的受主能级一般很深（N 除外），空穴不易于热激发进入价带，受主掺杂的固溶度也很低。可控稳定的 p 型掺杂一直是 ZnO 研究中的一个重要课题。目前文献报道 p 型掺杂主要有四种渠道：1. 掺 N 2. Ga、N 共掺 3. 掺 As 4. 掺 P。后两种受主能级都比较深，掺杂效果不理想，而且 As 和 P 都有毒。第一种中活性 N 源一般来源于 NH_3 、 N_2 或者 N_2O ，但是由于 N 原子取代 O 原子时，提高了 Madelung 能从而引起了 N 能级的局域化，N 能级较深，掺杂效果不是很理想。第二种施主（Ga）、受主（N）共掺杂技术得到的 ZnO 表面颗粒均匀度较差，而且要用到价格较高的 Ga 源，通常采用（MBE）分子束外延工艺，设备昂贵，制作成本高。

发明内容

本发明的目的是提供一种 p-ZnO 薄膜及其制备方法。

本发明的 p-ZnO 薄膜是以 N 为受主，以 Al 为施主共掺杂的 p-ZnO 薄膜，掺杂浓度为 $1.8 \times 10^{15} \text{ cm}^{-3}$ – $4.5 \times 10^{18} \text{ cm}^{-3}$ 。

本发明的 p-ZnO 薄膜的制备方法是利用磁控溅射法，先将衬底表面清洗后放入直流反应磁控溅射装置的反应室中，反应室真空度抽到 $1 \times 10^{-3} \text{ Pa}$ ，然后加热衬底，衬底温度为 400–600℃，以高纯含 N 气体（99.99%以上）和高纯 O_2 或 Ar（99.99%以上）为溅射气体，将二种气体分别由气体流量计控制输入装置的缓冲室，在缓冲室充分混合后引入真空反应室，在 3–5Pa 压强下，以 Zn_xAl_{1-x} 为靶材，式中 $1 > x > 0.6$ ，进行溅射生长。

上述的含 N 气体包括 NH_3 、 N_2O 、NO、 NO_2 。含 N 气体和高纯 O_2 或 Ar 的分压比以及靶材中 Al 的含量根据掺杂浓度调节，生长的时间由所需的厚度决定。

本发明的优点是：

- 1) 设备简单，操作方便，Al 原料丰富，价格低廉。制作成本低。
- 2) 掺杂浓度可控性、稳定性好，可以通过调节输入的含 N 气体和高纯 O_2 或 Ar

不同分压比和靶材中 Al 的含量来控制。

3) 本发明的 p-ZnO 薄膜表面平整, 晶粒大小均匀, 具有择优取向。

附图说明

图 1 是根据本发明方法采用的直流反应磁控溅射装置示意图。图中 1 和 2 分别为含 N 气体和高纯 O₂ 或 Ar 的进气管路; 3 为流量计; 4 为缓冲室; 5 为样品架; 6 为加热器; 7 为真空计; 8 为自动压强控制仪; 9 为 S 枪

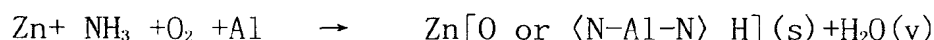
图 2 是 p-ZnO 薄膜表面的 AFM (原子力显微) 照片。

图 3 是蓝宝石上沉积的 p-ZnO 薄膜的 XRD (X 射线衍射) 图。

具体实施方式

以下参照图 1, 通过实例进一步说明本发明 p-ZnO 薄膜的制备方法。先将衬底经过表面清洗后放入反应室样品架 5 上, 衬底欲沉积表面朝下放置, 有效防止颗粒状的杂质对衬底的沾污, 反应室真空度抽至 1×10^{-3} Pa, 利用加热器加热衬底, 衬底温度控制在 550℃; 溅射气体是高纯的 NH₃ (99.99% 以上) 和高纯的 O₂ (99.99% 以上), 两路气体经进气管 1 和 2 进入缓冲室 4, 在缓冲室充分混合后引入到真空室, 真空室内的压强由自动压强控制仪 8 控制, 压强为 3Pa。NH₃ 与 O₂ 分压比根据掺杂需要, 可通过流量计 3 调节, 本例为 NH₃:O₂=60:40; 以 S 枪 9 上的 Zn_xAl_{1-x}, 为靶材进行溅射生长, 本例为 x=0.9。在 60W 的功率下生长 25min。

生长过程中发生如下反应:



这里, v 和 s 分别代表气态和固态。在 ZnO 薄膜中, N-Al-N 取代 O 而使 N 活化, 一个 N-Al-N 相当于一个空位, 使 Madelung 能降低, 从而 N 能级更浅。而 H 由于原子半径较小, 会存在于与 N 相邻的间隙位置, 这可以很大程度上抑制生长过程中间隙 Zn 的存在, 降低自补偿效应。

本发明的 p-ZnO 薄膜表面平整, 晶粒大小均匀 (见图 2), 并具有高度 C 轴择优取向 (见图 3)。

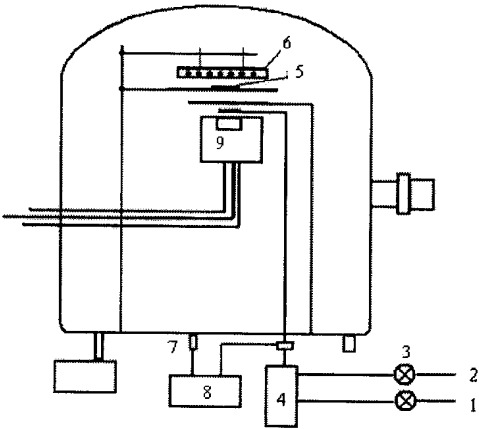


图 1

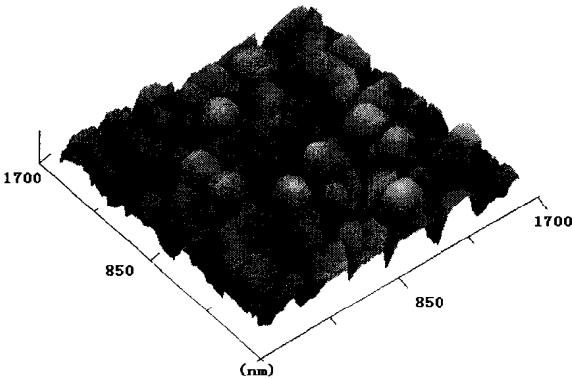


图 2

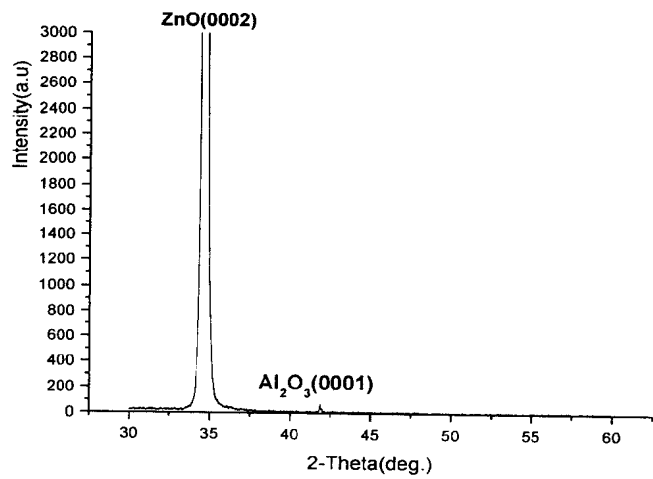


图 3



Publication number: CN1440053 A
Publication country: CHINA
Publication type: UNEXAMINED APPLIC. OPEN TO PUBLIC INSPECTION
Publication date: 20030903
Application number: CN20031016007
Application date: 20030325
Priority: CN2003101600720030325;
Assignee: ZHEJIANG UNIV ;
Assignee^{std}: UNIV ZHEJIANG ;
Inventor: GUODONG YUAN ; ZHIZHEN YE ; JINGYUN HUANG ;
Inventor^{std}: YUAN GUODONG ; YE ZHIZHEN ; HUANG JINGYUN ;
International class¹⁻⁷: C23C14/34 ; C23C14/08 ; H01L21/203 ;
International class⁸: C23C14/08; C23C14/08; C23C14/34; C23C14/34;
H01L21/02; H01L21/203;
Family members: CN1278385 C CN1440053 A
Title: P-ZnO thin film and preparation thereof
Title: 一种P - ZnO薄膜及其制备方法

Abstract:

The P-ZnO film of this invention is the codoped p-ZnO film by using N as the acceptor doping source and Al as the donor doping source with doping concentration of $1.8 \times 10^{15} \text{ cm}^{-3}$ - $4.5 \times 10^{18} \text{ cm}^{-3}$. The preparation method for it is to use magnetic control sputtering method by using $\text{Zn}_x\text{Al}_{1-x}$ as target material ($1 > x > 0.6$) and the doping concentration can be controlled by regulating the different differential pressure ratio of N gas and high purified O₂ or Ar to be inputted as well as by regulating the Al content in the target material so that the good stability and controllability can be obtained for the doping concentration.

Abstract:

本发明的p-ZnO薄膜是以N为受主掺杂源,以Al为施主掺杂源共掺杂的p-ZnO薄膜,掺杂浓度为 $1.8 \times 10^{15} \text{ cm}^{-3}$ - $4.5 \times 10^{18} \text{ cm}^{-3}$ 。其制备方法是利用磁控溅射法,以 $\text{Zn}_x\text{Al}_{1-x}$ 为靶材, ($1 > x > 0.6$), 掺杂浓度可以通过调节输入的含N气体和高纯

P-ZnO thin films of the present invention is based on N acceptor dopant source, Al doping source for the donor co-doped p-ZnO thin films doped concentration $1.8 \times 10^{15} \text{ cm}^{-3}$ - $4.5 \times 10^{18} \text{ cm}^{-3}$. The preparation method is the use of

O₂或Ar不同分压比和靶材中Al的含量来控制，掺杂浓度可控性、稳定性好。设备简单，操作方便，Al原料丰富，价格低廉。制作成本低。本发明利用施主-受主共掺生成的N-Al-Np型ZnO薄膜表面平整，晶粒大小均匀，具有择优取向。

magnetron sputtering to a Zn_xAl_{1-x} as a target, ($1 > x > 0.6$), the doping concentration can be adjusted enter the N-containing gas and high-purity O₂ or Ar partial pressure ratio and target different content of Al to control the doping concentration controllability, good stability. Equipment is simple, easy to operate, Al rich raw materials, low prices. Low production costs. The present invention utilizes a donor - acceptor co-doped generated N-Al-Np-type ZnO thin film surface roughness, grain size uniform, having a preferred orientation.

Claims:

权利要求书 1. 一种P-ZnO薄膜，其特征在于它是以N为受主，以Al为施主共掺杂的P-ZnO 薄膜，掺杂浓度为 $1.8 \times 10^{15} \text{ cm}^{-3}$ - $4.5 \times 10^{17} \text{ cm}^{-3}$ 2. 根据权利要求1所述的P-ZnO薄膜的制备方法，其特征是：先将衬底表面清洗后放入直流反应磁控溅射装置的反应室中，反应室真空度抽到 $1 \times 10^{-3} \text{ Pa}$ ，然后加热衬底，衬底温度为400-600。C,以高纯含N气体和高纯O₂或Ar为溅射气体，将二种气体分别由气体流量计控制输入装置的缓冲室，在缓冲室充分混合后引入真空反应室，在3-5Pa压强下，以Zn_xAl_{1-x}为靶材，式中 $1 > x > 0.6$ ，进行溅射生长。

3. 根据权利要求2所述的P-ZnO薄膜的制备方法，其特征是含N气体包括 NH₃、N₂O和/或 NO、N₂。

4. 根据权利要求2所述的p-ZnO薄膜的制备方法，其特征是所说的衬底是 蓝宝石、硅。

Book A P-ZnO film according to claim, characterized in that it is based on N as an acceptor, donor co-doped with Al as the P-ZnO thin films, doping concentration of $1.8 \times 10^{15} \text{ cm}^{-3}$ - $4.5 \times 10^{17} \text{ cm}^{-3}$ 2. according to the production method P-ZnO thin film according to claim 1, characterized in that: the first surface of the substrate after the reaction the reaction chamber and rinse DC magnetron sputtering apparatus, the reaction chamber the degree of vacuum pumped $1 \times 10^{-3} \text{ Pa}$, and then heating the substrate, the substrate temperature is 400-600.C, high purity and high purity N-containing gas as a sputtering gas Ar or O₂, the two kinds of gases are controlled by the gas flow meter input buffer chamber means, after thorough mixing the buffer chamber into the vacuum reaction chamber, under pressure 3-5Pa to Zn_xAl_{1-x} as a target, wherein $1 > x > 0.6$, sputtering growth

3 Preparation P-ZnO thin film according to claim 2, characterized in that the N-containing gas comprises NH₃, N₂O and/或 NO, N₂.

4. According to the production method of p-ZnO thin film according to claim 2, wherein said substrate is a sapphire, silicon.

Description:

说明书 一种p-ZnO薄膜及其制备方法 技术领域 本发明涉及半导体器件，具体说涉及P-ZnO薄膜及其制备方法。

背景技术 由于ZnO存在诸多的本征施主缺陷，如间隙锌ZnjH氧空位V11，其能级分别位于导带底0.05eV和0.3eV处，对受主产生高度自补偿作用。因此，很难实现ZnO的p型转变，而且，ZnO的受主能级一般很深（N除外），空穴不易于热激发进入价带，受主掺杂的固溶度也很低。可控稳定的p型掺杂一直是ZnO研究中的一个重要课题。目前文献报道P型掺杂主要有四种渠道：1. Ga, N 共掺 2. 掺As 3. 掺P。后两种受主能级都比较深，掺杂效果不理想，而且As和P都有毒。第一种中活性N源一般来源于NH₃、N₂O，但是由于N原子取代O原子时，提高了Madeling能从而引起了N能级的局域化，N能级较深，掺杂效果不是很理想。第二种施主（Ga）、受主（N）共掺杂技术得到的ZnO表面颗粒均匀度较差，而且要用到价格较高的Ga源，通常采用（MBE）分子束外延工艺，设备昂贵，制作成本高。

发明内容 本发明的目的是提供一种p-ZnO薄膜及其制备方法。

本发明的P-ZnO薄膜是以N为受主，以Al为施主共掺杂的p-ZnO薄膜，掺杂浓度为 $1.8 \times 10^{15} \text{ cm}^{-3}$ ~ $4.5 \times 10^{17} \text{ cm}^{-3}$ 本发明的P-ZnO薄膜的制备方法是利用磁控溅射法，先将衬底表面清洗后放入直流反应磁控溅射装置的反应室中，反应室真空度抽到 $1 \times 10^{-4} \text{ Pa}$ ，然后加热衬底，衬底温度为400-600°C，以高纯含N气体（99.99%以上）和高纯O₂或Ar（99.99%以上）为溅射气体，将二种气体分别由气体流量计控制输入装置的缓冲室，在缓冲室充分混合后引入真空反应室，在 $3 \times 10^{-5} \text{ Pa}$ 压强下，以

P-ZnO manual one kind of film and preparation method Technical Field The present invention relates to semiconductor devices, particularly P-ZnO film and preparation method involved.

Background Since ZnO has many intrinsic donor defects such as oxygen vacancy gap zinc Zn_jH V₁₁, their energy levels are located at the bottom of the conduction band at 0.05eV and 0.3eV, to produce a high degree of self-compensating acceptor effect. Therefore, it is difficult to realize p-type ZnO, and the acceptor is generally a deep level of ZnO (excluding N), the heat is not easily excited into a hole in the valence band, by the solubility of the dopant was too low. Controllable stable p-type doped ZnO research has been an important topic. P-type doping is currently reported in the literature, there are four channels: 1. Ga, N 2. Ga, N 3. co-doped erbium-doped As 4. P. The latter two acceptor level deeper than the doping effect is not ideal, but As and P are poisonous. The first active N source is generally derived from NH₃, N₂O, but because N atom O atom, can improve the Madelung causing a localized level N, N levels deep, doping the effect is not very satisfactory. The second donor (Ga), acceptor (N) co-doped ZnO surface is poor particle uniformity obtained technology, and more expensive to use a Ga source, usually (MBE) molecular beam epitaxy process, expensive equipment production costs high. object of the present invention is to provide a p-ZnO film and preparation method.

Zn_xAl_{1-x} 为靶材，式中 $1 > x > 0.6$ ，进行溅射生长。

上述的含N气体包括 NH_3 、 N_2O 、 NO 、 NO_2 含N气体和高纯 O_2 或Ar的分压比以及靶材中Al的含量根据掺杂浓度调节，生长的时间由所需的厚度决定。

本发明的优点是：1)设备简单，操作方便，Al原料丰富，价格低廉。制作成本低。

2)掺杂浓度可控性、稳定性好，可以通过调节输入的含N气体和高纯 O_2 或Ar 3 不同分压比和靶材中Al的含量来控制。

3)本发明的P-ZnO薄膜表面平整，晶粒大小均匀，具有择优取向。

附图说明 图1是根据本发明方法采用的直流反应磁控溅射装置示意图。图中1和2 分别为含N气体和高纯 O_2 或Ar的进气管路；3 为流量计；4为缓冲室；5 为样品架；6为加热器；7为真空计；8为自动压强控制仪；9为S枪 图2是p-ZnO薄膜表面的AFM (原子力显微) 照片。

图3是蓝宝石上沉积的p-ZnO薄膜的XRD (X射线衍射) 图。

具体实施方式 以下参照图1，通过实例进一步说明本发明P-ZnO薄膜的制备方法。先将衬底经过表面清洗后放入反应室样品架5上，衬底欲沉积表面朝下放置，有效防止颗粒状的杂质对衬底的玷污，反应室真空度抽至 1×10^{-3} Pa，利用加热器加热衬底，衬底温度控制在 $550^\circ C$ ；溅射气体是高纯的 NH_3 (99.99%以上) 和高纯的 O_2 (99.99%以上)，两路气体经进气管1和2进入缓冲室4，在缓冲室充分混合后引入到真空室，真空室内的压强由自动压强控制仪8控制，压强为3Pa。 NH_3 与 O_2 分压比根据掺杂需要，可通过流量计3调节，本例为 $NH_3 : O_2 = 60 : 40$ ；以S枪9上的 Zn_xAl_{1-x} 为靶材进行溅射生长，本例为 $x = 0.9$ 。在60W的功率下生长25min生长过程中发生如下反应： $NH_3 \rightarrow (NH)_2 + 2H + Al \rightarrow Al_{1-x}Zn_x + NH_3 + O_2 \rightarrow ZnO + Al_{1-x}Zn_x + H_2O$ 这里，v和s分别代表气态和固态。在ZnO薄膜中，

P-ZnO thin film of the present invention is based on N acceptor, p-ZnO thin films co-doped with Al as a donor, the doping concentration of $1.8 \times 10^{15} \text{ cm}^{-3}$ P- $4.5 \times 10^{17} \text{ cm}^{-3}$ the present invention Preparation of ZnO thin films using magnetron sputtering method Save, after the first surface of the substrate into the cleaning chamber DC reactive magnetron sputtering device, the degree of vacuum chamber pumped 1×10^{-3} Pa, then heating the substrate, the substrate temperature of $400-600^\circ C$, the N-containing gas is high purity (above 99.99 percent) and high purity O_2 or Ar (above 99.99 percent) as the sputtering gas, the two gases are controlled by the gas flow meter input device buffer chamber, after thorough mixing the buffer chamber into the vacuum reaction chamber, under pressure $3-5$ Pa to Zn_xAl_{1-x} as a target, where $1 > x > 0.6$, sputtering growth

said N-containing gas comprises NH_3 , N_2O , NO , NO_2 dividing ratio and the content of Al in the target N-containing gases and high purity O_2 or Ar doping concentration was adjusted according to the growth of time from the desired thickness decision. advantages of the present invention are: 1) the device is simple, easy to operate, Al-rich raw materials, low prices. Low production costs.

2) doping concentration controllability, stability, and high-purity N-containing gas or Ar O_2 can be adjusted by three different input voltage dividing ratio and the target of Al content is controlled.

3) P-ZnO film surface of the present invention is smooth, uniform grain size, have preferred orientation.

Figure 1 is a schematic view of a magnetron sputtering apparatus in accordance with the reaction method of

N-Al-N取代O而使N活化，一个N-A1-N相当于一个空位，使Madelung能降低，从而N能级更浅。

而H由于原子半径较小，会存在于与N相邻的间隙位置，这可以很大程度上抑制生长过程中间隙Zn的存在，降低自补偿效应。

本发明的P-ZnO薄膜表面平整，晶粒大小均匀（见图2），并具有高度C轴择优取向（见图3）。

4 图1 (ran) 图2 5 2-Theta(deg.) 图3 6

the present invention employs a DC. Figures 1 and 2 are the N-containing gas and a high purity Ar O₂ or the inlet line; a flow meter 3; 4 is a buffer chamber; sample holder 5; 6 is a heater; a vacuum gauge 7; 8, autonomous pressure- the control device; FIG. 9 is S gun 2 is a p-ZnO film surface AFM (atomic force microscopy) pictures.

3 is an XRD deposited on sapphire p-ZnO thin films (X-ray diffraction) FIG.

Description of Embodiments Referring to Figure 1, by way of example further illustrates the preparation of P-ZnO film of the present invention. After cleaning the surface of the first substrate into the reaction chamber sample holder 5, the surface of the substrate to be deposited Lay prevent particulate contamination of impurities on the substrate, the reaction chamber is evacuated to a vacuum degree IXIO³Pa, use heater the substrate, the substrate temperature is controlled at 550 degrees C; sputtering gas is high purity NH₃ (more than 99.99 percent) and high purity O₂ (more than 99.99 percent), two gas through the inlet pipe into the 1 and 2 buffer chamber 4, after thorough mixing the buffer chamber introduced into the vacuum chamber, the vacuum chamber pressure by an automatic pressure control device 8 controls, pressure is 3Pa. NH₃; t and the partial pressure ratio of O₂ based on doped necessary, by adjusting the flow meter 3, in this case NH₃: O₂ two 60: 40; 9 to S gun on Zn_xAl_{1-x}, -x, as a target sputtering growth, in this case, x = 0. 9. Under 60W of power growth following reactions occur 25min0 growth process: NH₃ - (NH)₂ ' 2H Al - Al; _ 3e Zn NH; 0, Al - Zn [0 or (N-Al-N) H] (s) H₂O (V) here, v and s represent the gaseous and solid. In the ZnO films, N-Al-N-substituted N O leaving activation, a N-A1-N is equivalent to a vacancy, so Madelung energy reduced, thereby N levels lighter.

Since the H atom and smaller radius, and will exist in the N position adjacent to the gap, which can be largely suppressed during the growth gap Zn, lower self-compensating effect.

P-ZnO film surface of the present invention is smooth, uniform grain size (see Figure 2), and with a high degree C-axis orientation (see Figure 3).

4 Figure 1 (ran) Figure 2 5 2-Thela (deg.)
Figure 3 6

Cited by: CN102199758 A ; CN102492928 A ; CN105120268 A ;

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

H04N 11/04

H04N 7/32



[12] 发明专利申请公开说明书

[21] 申请号 200410025123.7

[43] 公开日 2005 年 2 月 23 日

[11] 公开号 CN 1585495A

[22] 申请日 2004.6.11

[21] 申请号 200410025123.7

[71] 申请人 上海大学

地址 200072 上海市闸北区延长路 149 号

[72] 发明人 张兆扬 滕国伟 张颖

[74] 专利代理机构 上海上大专利事务所

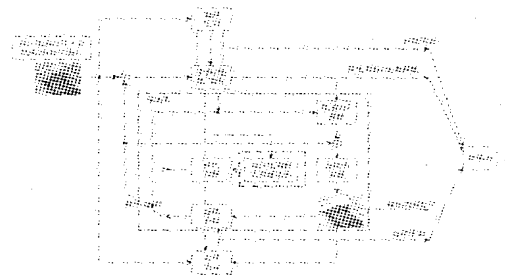
代理人 何文欣

权利要求书 2 页 说明书 9 页 附图 2 页

[54] 发明名称 H.264/AVC 帧内预测模式的快速选择方法

[57] 摘要

本发明涉及一种 H.264/AVC 实时编码器的帧内预测模式快速选择方法。在 H.264/AVC 编码器中, 帧内编码采用帧内预测方法, 利用色度块和亮度块、亮度块的内部以及亮度块预测模式选择间的相关性, 根据率失真优化(RDO)准则确定不同判决依据(RDO 或 SAD)来快速选择帧内预测模式。本发明提供的方法, 对比于全搜索方法, 可以在信噪比和码率变化很小的同时, 可使编码速度平均提高 6 倍以上。



ISSN 1008-4274

1. 一种 H. 264/AVC 实时编码器帧内预测模式的快速选择方法, 在 H. 264/AVC 编码器中, 帧内编码采用帧内预测方法, 其特征在于利用当前宏块的色度块和亮度块帧内预测模式间的相关性、4x4 亮度块 (I4) 和 16x16 亮度块 (I16) 的模式选择之间的相关性以及亮度块预测模式内部的相关性, 以简化的率失真优化 (RDO) 为判决依据, 快速进行帧内预测模式选择。
2. 如权利要求 1 所述的 H. 264/AVC 实时编码器帧内预测模式的快速选择方法, 其特征在于, 快速选择的步骤如下:
 - a. 确定色度块的模式选择具有独立性, 即亮度宏块无论选择 I16 还是 I4, 色度宏块模式都是一致的, 从而首先对色度预测模式以 SAD 为准则进行独立判决, 确定色度的最佳预测模式。
 - b. 确定了色度的最佳预测模式后, 为亮度块模式判决提供如下依据: 对于亮度 I16 而言, 当色度块选择 D 模式, 亮度可以选择 D、H、V、P 模式; 当色度块选择 H 模式, 亮度可以选择 D、H、P 模式; 当色度块选择 V 模式, 亮度可以选择 D、V、P 模式; 当色度块选择 P 模式, 亮度可以选择 D、P 模式。而对于亮度 I4, I8 选择的模式是亮度 I4 下绝大多数 4x4 块可能选择的模式。
 - c. 对于宏块, I4 的选择概率远远大于 I16, 因此我们首先进行 I4 下 4x4 块的模式判决; 根据邻块的模式判断得到当前块的可能模式, 并和由步骤 b 确定的可能模式, 分别计算这两种模式下的 RDO, 并与阈值 T0 比较, 小于 T0, 则可确定为当前 4x4 的最佳模式, 否则较小值 RDO 的为可能最佳模式。
 - d. 比较步骤 c 中得到的可能最佳模式的预测方向邻近的模式的 RDO, 进一步确定可能的最佳模式。如果它是 V、H 或 D, 则比较这三种模式中未曾比较的模式; 如果它为其它模式, 则比较所有其它未曾比较的模式。
 - e. 在完成当前宏块的 16 个 4x4 块的模式判决后, 计算宏块纹理复杂度, 以 MSE 来衡量, 并与阈值 T1 比较, 如果大于 T1, 提前终止判决, 上述得到的结果为宏块最佳模式; 如果小于 T1, 则进行亮度 16x16 块的模式判决。
 - f. 根据步骤 b 中色度块的最佳模式与亮度块的关系, 确定当前宏块 I16 的可能模式; 统计步骤 e 中 I4 情况下 16 个子块的模式, 得到概率最大的两种模式; 综合考虑这些可能模式, 最终确定 I16 的可能模式。并以 SAD 为准则, 确定 I16 的

最佳预测模式。

g. 自适应确定步骤 e 所述的阈值 T1 来判断是否提前终止 I16 判决和自适应确定步骤 c 所述的 T0 来判断是否提前终止 I4 的判决。

H. 264/AVC 帧内预测模式的快速选择方法

技术领域

本发明涉及一种信息处理的视频编码的方法，特别是一种 H. 264/AVC 帧内预测模式的快速选择方法。

背景技术

自从 ITU-T 颁布了 H. 261 后，基于块的混合编码方法作为框架被普遍采用并不断发展。国际标准 H. 261、H. 263、MPEG-1、MPEG-2、MPEG-4 等的制定极大推动了多媒体技术的应用。然而，随着新服务的大量增加和高清电视的即将普及，以及现有传输媒体如 Cable Modem、Xdsl、UMTS 较低的传输码率，迫切需要提高编码效率；同时还要求视频编码标准能够适应现存或未来的网络。这些问题为新视频编码标准 H. 264/AVC 的产生提供了切实的依据。

这些视频编码标准采用混合编码框架，即将输入视频图象分割成宏块，对其进行预测、变换、量化、熵编码。预测包括帧内预测和帧间预测，帧内预测可去除空间冗余，而帧间预测可去除时间上的冗余。

H. 264/AVC 仍遵从这种基于块的混合编码方案，但包含了许多新特征：VCL 层和 NAL 层分离、空间域的帧内估计、1/4 像素精度运动估计、自适应块尺寸运动补偿、多参考帧运动补偿预测、通用 B 帧的概念、低复杂 16bit 的整数变换和量化、环内去块滤波器和高效的熵编码。这些新特征一方面能够显著提高编码效率；另一方面具有网络友好性，可有效用于各种网络和应用环境。

其中 H. 264/AVC 为提高帧内图象的编码效率，采用帧内预测去除空间冗余性。这种帧内预测技术依然以宏块为处理单元，即通常将一帧图象分割成 16x16 的图象块进行编码。然而基本处理块单元相对于其它标准由 8x8 降为 4x4，即变换采用 4x4 的变换矩阵。因此亮度宏块支持两种帧内预测块尺寸：4x4 和 16x16，而色度仅支持 8x8。每个进行帧内预测的当前块利用邻块的相邻像素值来预测当前块像素值，然后计算残差、变换、熵编码。H. 264/AVC 标准中每个 4x4 块的帧内预测包括 9 种预测模式（如附图 1）。而亮度 16x16 和色度 8x8 仅支持 4 种预测模式，分别为：垂直（V）、水平（H）、直流（D）和平坦模式（P），分别简称为 V、H、D 和 P 模式。

同时这项技术也被应用到帧间编码之中，即对当前宏块进行完帧间预测后，接着

对它进行帧内预测，以率失真优化为最终判决依据，率失真小的模式被确定为最终模式。这样既可以提高帧间编码宏块的编码效率，又可以提高容错能力。

例如对某帧（I 帧、P 帧或 B 帧皆可）中当前宏块 M 进行帧内预测编码，利用相邻块与当前块的紧挨的像素值来预测当前块的像素值，然后用当前块的像素值减去预测值得到残差矩阵，将此残差矩阵经过变换、量化、熵编码得到该宏块的编码码流。

但利用邻块像素值的预测方式是不同的，因此产生了多种预测模式（如附图 1）。在判断这些模式中的最优模式，需要一个衡量的准则，目前采用了率失真优化（RDO-rate distortion optimization），其定义如下：

$$J_{MODE}(S_k, I_k | Q, \lambda_{MODE}) = D_{REC}(S_k, I_k | Q) + \lambda_{MODE} R_{REC}(S_k, I_k | Q) \quad \text{公式 (1)}$$

式中 D_{REC} 表示编码后总的失真， R_{REC} 表示总的码率。 $\lambda_{MODE} \geq 0$ 为拉格朗日参数，通过确定编码选择项 I_k ，使编码后的失真和码率总和最小。

本发明在色度宏块和亮度 16x16 块的模式判决中采用了一种简化的判决准则为 SAD (Sum of Absolute Difference)，其定义如下：

$$SAD = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} abs(f(i, j) - f'(i, j)) \quad \text{公式 (2)}$$

式中 $f(i, j)$ 为原始信号， $f'(i, j)$ 是帧内预测后的预测信号，N 等于 8（色度块）或 16（亮度块）。

H. 264/AVC 校验模型对于当前宏块帧内预测模式的确定给出了一种全搜索方法，以色度宏块的帧内预测模式为外循环，以亮度宏块的两种块模式为内循环，即对应每一种色度宏块的帧内预测模式，都要分别比较亮度宏块 I4 和 I16 的所有模式。故这种方法能够扫描到所有的模式，因此编码质量最高，运算量也是最大的。搜索的模式总数为 $4 \times (16 \times 9 + 4)$ ，共 592 个，这意味着为了得到一个宏块的最佳模式，需要计算 592 个不同的 J_{MODE} 。这使 H. 264/AVC 在现有的硬件条件下实现编码是不可能的。

发明内容

本发明的目的是提供一种 H. 264/AVC 实时编码器的帧内预测模式的快速选择方法，对比于全搜索方法，它可以在信噪比和码率变化很小的同时，可使编码速度平均提高 6 倍以上。

为达到上述目的，本发明采用下述技术方案：

一种 H. 264/AVC 实时编码器帧内预测模式的快速选择方法，在 H. 264/AVC 编码器中，帧内编码采用帧内预测方法，其特征在于利用当前宏块的色度块和亮度块帧内预测模式间的相关性、4x4 亮度块（I4）和 16x16 亮度块（I16）的模式选择之间的相关性以及亮度块预测模式内部的相关性，以简化的率失真优化（RDO）为判决依据，快速进行帧内预测模式选择。

快速选择的步骤如下：

- a) 确定色度块的模式选择具有独立性，即亮度宏块无论选择 I16 还是 I4，色度宏块模式都是一致的，从而首先对色度预测模式以 SAD 为准则进行独立判决，确定色度的最佳预测模式。
- b) 确定了色度的最佳预测模式后，为亮度块模式判决提供如下依据：对于亮度 I16 而言，当色度块选择 D 模式，亮度可以选择 D、H、V、P 模式；当色度块选择 H 模式，亮度可以选择 D、H、P 模式；当色度块选择 V 模式，亮度可以选择 D、V、P 模式；当色度块选择 P 模式，亮度可以选择 D、P 模式。而对于亮度 I4，I8 选择的模式是亮度 I4 下绝大多数 4x4 块可能选择的模式。
- c) 对于宏块，I4 的选择概率远远大于 I16，因此我们首先进行 I4 下 4x4 块的模式判决；根据邻块的模式判断得到当前块的可能模式，并和由步骤 b 确定的可能模式，分别计算这两种模式下的 RDO，并与阈值 T0 比较，小于 T0，则可确定为当前 4x4 的最佳模式，否则较小值 RDO 的为可能最佳模式。
- d) 比较步骤 c 中得到的可能最佳模式的预测方向邻近的模式的 RDO，进一步确定可能的最佳模式。如果它是 V、H 或 D，则比较这三种模式中未曾比较的模式；如果它为其它模式，则比较所有其它未曾比较的模式。
- e) 在完成当前宏块的 16 个 4x4 块的模式判决后，计算宏块纹理复杂度，以 MSE 来衡量，并与阈值 T1 比较，如果大于 T1，提前终止判决，上述得到的结果为宏块最佳模式；如果小于 T1，则进行亮度 16x16 块的模式判决。
- f) 根据步骤 b 中色度块的最佳模式与亮度块的关系，确定当前宏块 I16 的可能模式；统计步骤 e 中 I4 情况下 16 个子块的模式，得到概率最大的两种模式；综合考虑这些可能模式，最终确定 I16 的可能模式。并以 SAD 为准则，确定 I16 的最佳预测模式。

- g) 自适应确定步骤 e 所述的阈值 T1 来判断是否提前终止 I16 判决和自适应确定步骤 c 所述的 T0 来判断是否提前终止 I4 的判决。

本发明综合利用上述方法进行帧内预测模式选择。实验结果显示,可以提高编码速度数倍,且能够保证编码的码率和 PSNR 与全搜索方法相比变化不大。

附图说明

图 1 是本发明的一个实施例的信息处理视频编码框图。

图 2 是本发明的一个实施例的处理流程图

图 3 是 4x4 块帧内预测方向图

图 4 是当前块的邻块位置图

图 5 是 I16 和 I4 的率失真变化图

具体实施方式

以下结合附图描述本发明的 H.264/AVC 实时编码器的帧内预测模式的快速选择方法的一个优选实施例:

在校验模型 Jm72 编码器上,具体工作环境为 2.4GHz CPU 的 PC 机,针对 I 帧序列,编码序列为 I1111,熵编码采用了 CABAC,量化参数分别为 14、28,亮度 I4 的 RDO 阈值 T0 和宏块复杂度的阈值 T1 可自适应计算得到,采用下述方法:在我们实验过程中,发现校验模型提供的全搜索方法忽略了模式选择间的相关性。在实际中,宏块的色度块和亮度块帧内预测模式间、4x4 亮度块(I4)和 16x16 亮度块(I16)的模式选择之间以及亮度块预测模式内部都具有较强的相关性。如果能充分利用这些相关性,就可以快速有效的确定帧内预测的最佳模式。

(1) 参见图 1 和图 2,首先,可以观察到亮度宏块无论选择 I16 还是 I4,色度宏块模式都是一致的,因此可以确定在整个模式选择过程中色度块的模式选择具有独立性,故色度预测模式可以在亮度判决之前进行独立判决色度块的最佳预测模式可以为亮度宏块模式判决提供依据。尽管对于亮度 I16 与色度 I8 模式选择不能达到完全一致,但存在着联系:当色度块选择 D 模式,亮度可以选择 D、H、V、P 模式;当色度块选择 H 模式,亮度可以选择 D、H、P 模式;当色度块选择 V 模式,亮度可以选择 D、V、P 模式;当色度块选择 P 模式,亮度可以选择 D、P 模式;而对于亮度 I4 与色度 I8, I8 选择的模式是亮度 I4 下绝大多数 4x4 块最有可能选择的模式。亮度块内部模式选择之间存在着优先性,V、H、D 模式相对于其它模式具有优先选择性。I4 情况下,通过相邻块(图 4)的模式可以判决得到当前块的最优先选择模式。对于优

先选取的模式，其方向上相邻的模式也具有很大的选中可能。

实际中，I4 的选择概率远远大于 I16，因此在判决过程上优先检测 I4。在 I4 情况下在确定的 16 个 4x4 块最佳模式中，概率最大的两种模式很可能与 I16 的最佳模式一致，因此可以将之色度模式联合进行检验，最终确定 I16 最佳模式。

H. 264/AVC 编码器采用 RDO 为模式判别准则，但其结构复杂，尤其写码流相当耗时间，因此可以根据具体模式采用不同的判决依据：① 色度模式选择不依赖亮度的模式选择，且其编码比特只包含色度模式和色度残差两部分，因此采用 SAD 就可以反映不同模式间率失真的差别；② 在亮度 I4 的 4x4 块编码比特包括块模式和残差的熵编码。因为 4X4 块尺寸较小，其块模式编码的比特数和不同模式间的残差编码比特数的差异会导致率失真的变化很大，故选择率失真作为 I4 的模式判决依据。且 4x4 块的率失真的分布比较集中，可在分布范围中定义一个阈值 T_0 ，如果在某一模式下其率失真小于 T_0 ，可认为达到最佳模式；③ I16 编码比特主要包括宏块模式和残差的熵编码，故 SAD 能够反映率失真的变化，以其作为模式选择的依据。

由于 I16 和 I4 随宏块纹理复杂度的 RDO 曲线是不同的（图 5），在模式选择时还应考虑宏块的纹理复杂度，并以宏块纹理复杂度的某个值作为阈值 T_1 来确定二者间的判决，即在纹理平坦区域（小于 T_1 ），I16 生成的重建图像能更好地与原始图像相匹配，选择 I16 概率大于 I4，随着纹理复杂度的逐渐增加，I16 在细节内容上率失真逐渐增加，当大于或等于 T_1 时将大于 I4 的率失真，这时 I4 更适合宏块编码模式选择。衡量纹理复杂性的方式有很多种，如算术平均、几何平均、中值、方差和均方误差。本发明选用了均方误差衡量宏块的纹理复杂度。

本发明充分利用上述的相关性，将之有机结合在一起，根据 RDO 准则确定不同模式的判决依据，得到帧内预测模式的快速选择方法，其流程如图 2 所示，具体步骤实施如下：

- ① 对当前宏块的色度块进行帧内预测，以 SAD 为判决依据，确定色度宏块帧内的最佳模式。
- ② 对当前宏块的亮度块进行 I4 帧内预测，首先根据当前 4x4 块的相邻块（图 2）的模式确定可能模式，将之和色度块的最佳模式相比较，确定最有可能块模式。计算最可能模式的 J_{MODE} ，与阈值 T_0 比较，如果小于 T_0 ，即可确认为最佳模式，跳转到步骤⑦。否则跳转到步骤③。

- ③ 计算另一模式（上面被排除的模式）的 J_{MODE} ，与步骤②计算的 J_{MODE} 比较，得到最小 J_{MODE} 值的模式。与 T_0 比较，如果小于 T_0 ，确认为最佳模式，跳转到步骤⑦。
- ④ 选择最小 J_{MODE} 的模式作为候选模式，计算与其预测方向相邻的两个模式的 J_{MODE} ，比较这三种模式的 J_{MODE} ，如果邻近模式的 J_{MODE} 值都大于候选模式的 J_{MODE} ，跳转步骤⑦。否则到步骤⑤。
- ⑤ 如果最小 J_{MODE} 的模式是 V、H 或 D，计算这三种模式在上述过程未曾计算的 J_{MODE} ，比较得到最小 J_{MODE} 的模式，跳转到步骤⑦；如果是其它模式，跳转到步骤⑥。
- ⑥ 计算其它剩余模式，得到最小 J_{MODE} 模式，跳转到步骤⑦。
- ⑦ 该最小 J_{MODE} 模式为该 4x4 块的最佳模式。
- ⑧ 循环步骤②—⑦，得到 16 个 4X4 块的最佳模式。计算宏块复杂度，如果大于阈值 T_1 （图 3），则该宏块为 I4，结束判决。否则计算 I4 情况下整个宏块的 J_{MODE} ，跳转到步骤⑨。
- ⑨ 根据色度块的最佳模式，确定当前宏块 I16 的可能模式。统计 I4 情况下 16 个子块的模式，得到概率最大的两种模式，如果属于 V、H 或 D，它们是 I16 的可能模式；如果为其它模式，则 P 模式成为 I16 的可能模式。
- ⑩ 比较上述确定的 I16 可能模式，得到 I16 最佳模式，计算其 J_{MODE} 。与 I4 情况下的宏块最小 J_{MODE} 比较，得到最佳块尺寸和其对应的最佳预测模式。

通过上述方法可以发现，本发明在通常情况下帧内预测编码需要检测的色度模式为 4 种，亮度 I4 的每个 4x4 块最少 1 种或最多 9 种。如果提前终止的话，不需要检测亮度 I16，否则亮度 I16 检测最少 2 种或最多 4 种模式，这种情况下几乎涵盖了所有的模式。但总体检测的模式数极大地降低。

H. 264/AVC 的帧间编码中也使用了帧内预测编码技术，上述方法同样适用，只是略作改动即可：在步骤 2 的 I4 帧内预测中，当该 4x4 块的邻块模式为帧内编码模式不必考虑直接使用即可；如果邻块为帧间模式，则可默认其帧内预测模式为直流模式

(D)。

(2) 计算阈值 T0 的自适应方法

H. 264/AVC 采用的是类 DCT 变换。公式 (3) 为 4×4 DCT 整数变换, E 为 4×4 亮度块经预测后残差矩阵, X 为类 DCT 整数变换矩阵, I 为整数变换后量化前的系数矩阵。

$$I = XEX^T \quad \text{公式 (3)}$$

$$\text{其中: } X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix}$$

设 DC 为变化后的直流系数, s 是当前 4×4 块的原始像素值, s' 是预测后重建块的像素值, 如公式 (4) 所示。如果 DC 趋近于 0, 那么其它交流系数 AC 也趋于 0, 整个变换后的矩阵趋向于 0 矩阵, 因此可省略后面的处理过程。

$$DC = \sum_{x=0}^3 \sum_{y=0}^3 E(x, y) = \sum_{x=0}^3 \sum_{y=0}^3 (s(x, y) - s'(x, y)) \quad \text{公式 (4)}$$

在 H. 264/AVC 中如果变换系数经过量化后, 其量化值绝对值小于或等于 1 时, 则残差矩阵可视为 0 矩阵。即满足下面的不等式:

$$|DC| < (2^{q\_bits} - qp\_const) / quant\_coef[qp\_rem][0][0] \quad \text{公式 (5)}$$

其中: $qp\_per = QP / 6$, $qp\_rem = QP \% 6$, $qp\_bits = qp\_per + 15$,
 $qp\_const = (1 << q\_bits) / 6$, QP 为量化参数。

对于每个 4×4 块, 公式 (6) 给出 SAD 的计算, 具体如下:

$$SAD = \sum_{x=0}^3 \sum_{y=0}^3 |s(x, y) - s'(x, y)| \quad \text{公式 (6)}$$

很明显下面不等式成立:

$$\left| \sum_{x=0}^3 \sum_{y=0}^3 (s(x, y) - s'(x, y)) \right| \leq \sum_{x=0}^3 \sum_{y=0}^3 |s(x, y) - s'(x, y)|$$

$$\text{即 } |DC| \leq SAD。$$

则如果下式成立, 则满足 0 矩阵的判断。

$$SAD < (2^{q\_bits} - qp\_const) / quant\_coef[qp\_rem][0][0] \quad \text{公式 (7)}$$

阈值 T0 由 4x4 块的 RDO 确定，如公式 8 所示，mode 表示所选的预测模式，R 包括模式类型和残差的熵编码。

$$J_{mode}(mode) = SSD + \lambda_{mode} R(mode, E) \quad \text{公式 (8)}$$

其中， $SSD = \sum_{x=0}^3 \sum_{y=0}^3 |s(x, y) - p(x, y)|^2$ ， $p[x, y]$ 是预测后残差经过变换、量化、反量化、

反变换后与参考块的值相加得到，故下面不等式成立：

$$|s(x, y) - p(x, y)| < |s(x, y) - s'(x, y)|$$

$$\text{因此 } SSD \leq (\sum_{x=0}^3 \sum_{y=0}^3 |s(x, y) - s'(x, y)|^2) < (\sum_{x=0}^3 \sum_{y=0}^3 |s(x, y) - s'(x, y)|)^2$$

故 $SSD < SAD^2$ ，即

$$SSD < ((2^{q\_bits} - qp\_const) / quant\_coef[qp\_rem][0][0])^2 \quad \text{公式 (9)}$$

这时，该块残差量化系数将趋于 0，因此得到 T0。

$$T0 = ((2^{q\_bits} - qp\_const) / quant\_coef[qp\_rem][0][0])^2 \quad \text{公式 (10)}$$

实际中可能存在差别，因此设置一个校正因子 $\alpha = 0.5$ ：

$$T0 = \alpha \times ((2^{q\_bits} - qp\_const) / quant\_coef[qp\_rem][0][0])^2 \quad \text{公式 (11)}$$

(3) 计算阈值 T1 的自适应方法

在实际中，当 T1 越大，搜索 I16 的可能性越大，因此搜索的模式越多；T1 越小，搜索 I16 的可能性越小，则搜索的模式变少。本发明通过当前宏块的前一宏块来确定当前 T1 的值，具体方法如下：

初始化：T1=1<<20

第一个宏块：依次进行 I4 和 I16 模式判决，得到二者的 MSE，将均值赋给 T1，作为下一宏块的阈值。

其它宏块：如果仅判决了 I4 模式，则 T1 不变。如果依次进行了 I4 和 I16 模式判决，则二者的 MSE 的均值赋给 T1。

测试三个纹理复杂程度不同的 CIF 序列 Mobile、Bus 和 Foreman。实验结果表明，本快速帧内预测模式选择方法比 H. 264/AVC 原有的全搜索方法可提高 I 帧编码速度大约 6 倍左右。在速度提高的同时，平均信噪比变化约 0.1db 左右，码率稍有变化。

本发明的第二个实施例是：针对 P 帧编码，编码序列为 I PPP，运动估计搜索

范围 ± 16 ，5个参考帧，运动估计算法为全搜索算法，熵编码采用了CABAC，量化参数分别为14、28。亮度I4 4x4块的RDO的阈值T0和宏块复杂度的阈值T1，同样采用上述的自适应方法。仍然测试三CIF序列Mobile、Bus和Foreman。实验结果表明，本发明的快速帧内预测模式选择方法在帧间情况下比H.264/AVC原有的帧间全搜索方法可提高编码速度大约25%左右。在速度提高的同时，平均信噪比变化约0.02db左右，码率几乎没有变化。

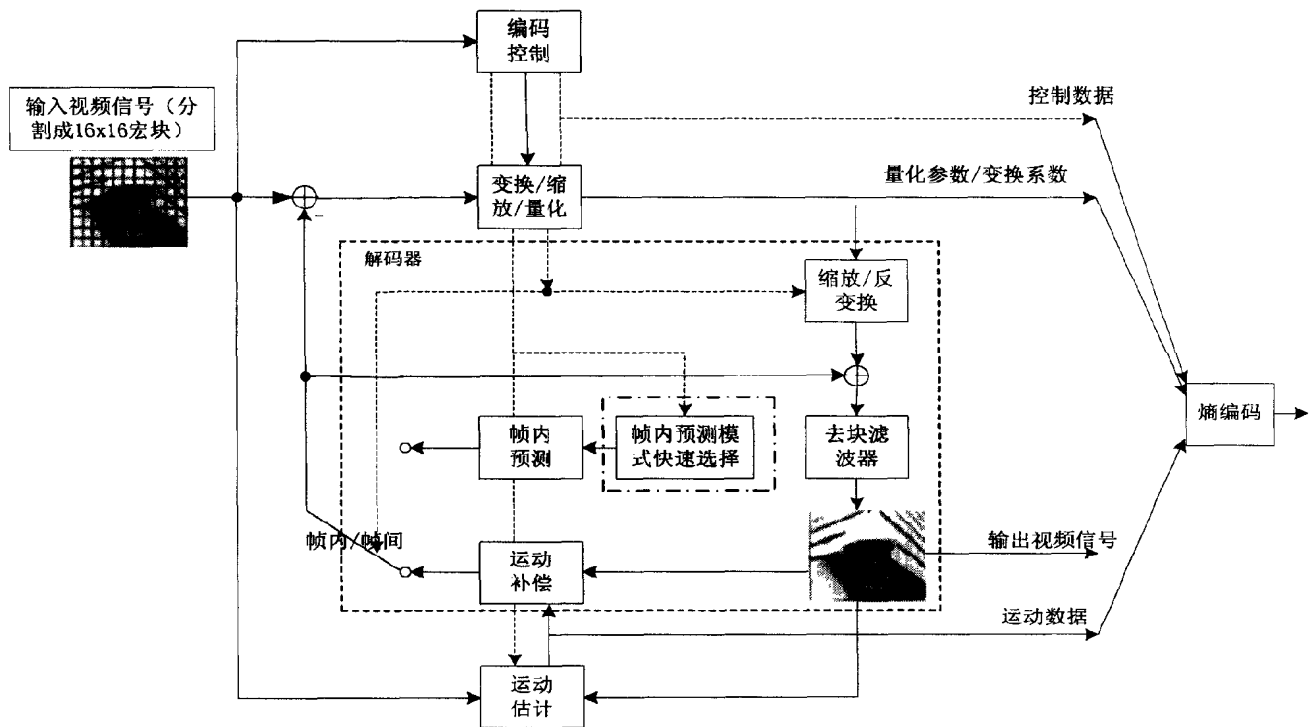


图 1

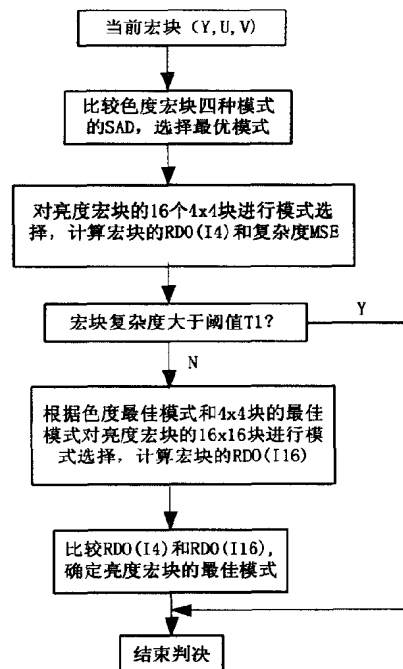


图 2

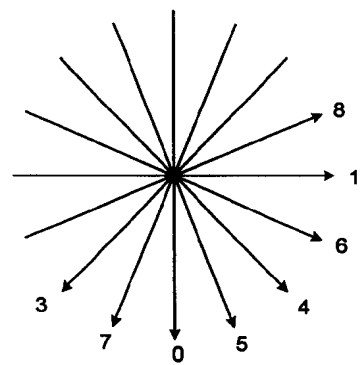


图 3

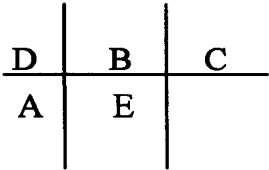


图 4

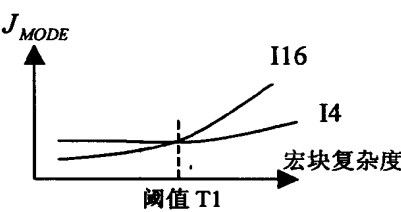


图 5



Publication number: CN1585495 A
Publication country: CHINA
Publication type: UNEXAMINED APPLIC. OPEN TO PUBLIC INSPECTION
Publication date: 20050223
Application number: CN200410025123
Application date: 20040611
Priority: CN20041002512320040611CN2004125123;
Assignee: SHANGHAI UNIV ;
Assignee^{std}: UNIV SHANGHAI ;
Inventor: GUOWEI TENG ; YING ZHANG ; ZHAOYANG ZHANG ;
Inventor^{std}: ZHANG ZHAOYANG ; ZHANG YING ; TENG GUOWEI ;
International class¹⁻⁷: H04N7/32 ; H04N11/04 ;
International class⁸: H04N11/04; H04N19/11; H04N19/186;
Family members: CN100401789 C CN1585495 A
Title: QUICK SELECTION OF PREDICTION MODES IN H.264/AVC FRAME
Title: H.264/AVC帧内预测模式的快速选择方法

Abstract:

The method includes following steps: in H.264/AVC encoder the intraframe coding adopts intraframe prediction method; the dependence between Hue block, brightness block, inside brightness block and prediction mode selections of brightness block are used to determine different ground of decision (RIX or SAD) on which intraframe prediction mode can be fast selected. The invention can provide 6 time encoding speed more than full search method.

Abstract:

本发明涉及一种H.264/AVC实时编码器的帧内预测模式快速选择方法。在H.264/AVC编码器中，帧内编码采用帧内预测方法，利用色度块和亮度块、亮度块的内部以及亮度块预测模式选择间的相关性，根据率失真优化(RDO)准则确定不同判决依据(RDO或SAD)来快速选择帧内预

The present invention relates to a H.264 / AVC real-time encoder fast intra prediction mode selection method. In H.264 / AVC encoder, intra-coded using intra prediction method using luminance blocks and chrominance blocks, correlation between the luminance block

测模式。本发明提供的方法，对比于全搜索方法，可以在信噪比和码率变化很小的同时，可使编码速度平均提高6倍以上。

and the luminance inside the block between the prediction mode selection, optimization (RDO) criterion based on rate-distortion to determine the different judgments on the basis of (RDO or SAD) to quickly select intra prediction mode. The method of the present invention provides, in contrast to the full search method, small changes in the signal to noise ratio and bit rate at the same time, can increase the speed to more than 6 times.

Claims:

1. 一种H.264/AVC实时编码器帧内预测模式的快速选择方法，在H. 264/AVC编码器中，帧内编码采用帧内预测方法，其特征在于利用当前宏块的色度块和亮度块帧内预测模式间的相关性、4x4亮度块（14）和16x16亮度块（116）的模式选择之间的相关性以及亮度块预测模式内部的相关性，以简化的率失真优化（RDO）为判决依据，快速进行帧内预测模式选择。

Quick selection method

1. One kind of H.264 / AVC intra prediction coder in real time mode, in H. 264 / AVC encoder, preclude the use of intra-coded intra prediction method, comprising using the current correlation between the luminance and chrominance blocks block intra prediction mode for the macroblock between, 4x4 luma block (14) and a 16x16 luminance blocks (116) of the correlation between the mode selection and correlation of the luminance prediction mode of block internal to simplified rate-distortion optimization (RDO) as judgment based fast intra prediction mode selection.

2. 如权利要求1所述的H.264/AVC实时编码器帧内预测模式的快速选择方法，其特征在于，快速选择的步骤如下：a. 确定色度块的模式选择具有独立性，即亮度宏块无论选择116还是14，色度宏块模式都是一致的，从而首先对色度预测模式以SAD为准则进行独立判决，确定色度的最佳预测模式。

2 as claimed in H.264 / AVC Real-time rapid selection method of the encoder intra prediction modes, characterized in that the quick selection of the following steps: A determining mode selection chroma blocks having independent, that is the choice luminance macroblock 116 or 14, chroma macroblock mode are the same, so first chroma prediction mode SAD as a criterion for independent judgment to determine the best prediction mode chromaticity.

b. 确定了色度的最佳预测模式后，为亮度块模式判决提供如下依据：对于亮度116而言，当色度块选择D模式，亮度可以选择D、H、V、P模式；当色度块选择H模式，亮度可以选择D、H、P模式；当色度块选择V模式，亮度可以选择D、V、P模式；当色度块选择P模式，亮度可以选择D、P模式。而对于亮度14，18选择的模式是亮度14下绝大多数4x4块可能选择的模式。

c. 对于宏块，14的选择概率远远大于116，因此我们首先进行14下4x4块的模式判决；根据邻块的模式判断得到当前块的可能模式，并和由步骤b确定的可能模式，分别计算这两种模式下的RDO，并与

阈值 T_0 比较，小于 T_0 ，则可确定为当前4x4的最佳模式，否则较小值RDO的为可能最佳模式。

d. 比较步骤c中得到的可能最佳模式的预测方向邻近的模式RDO，进一步确定可能的最佳模式。如果它是V、H或D，则比较这三种模式中未曾比较的模式；如果它为其它模式，则比较所有其它未曾比较的模式。

e. 在完成当前宏块的16个4x4块的模式判决后，计算宏块纹理复杂度，以MSE来衡量，并与阈值 T_1 比较，如果大于 T_1 ，提前终止判决，上述得到的结果为宏块最佳模式；如果小于 T_1 ，则进行亮度16x16块的模式判决。

f. 根据步骤b中色度块的最佳模式与亮度块的关系，确定当前宏块116的可能模式；统计步骤e中14情况下16个子块的模式，得到概率最大的两种模式；综合考虑这些可能模式，最终确定116的可能模式。并以SAD为准则，确定116的最佳预测模式。

g. 自适应确定步骤e所述的阈值 T_1 来判断是否提前终止116判决和自适应确定步骤c所述的 T_0 来判断是否提前终止14的判决。

b chroma prediction mode to determine the best after the fast mode provides brightness based on judgment as follows: 116 for brightness, when the chroma D block selection mode, the brightness can choose D, H, V, P mode; H when the chroma block selection mode, the brightness can choose D, H, P mode; when V chrominance block selection mode, the brightness can choose D, V, P mode; P mode is selected when the chroma blocks, the brightness can choose D, P mode. The luminance 14, 18 14 the selected mode is the vast majority of 4x4 luminance block mode may be selected.

c macroblock, far greater than the probability of selection 14 116, so we first conducted under the 14 4x4 blocks mode decision; according to the adjacent block pattern may be judged from the current block mode, and a and b are determined by step possible modes are calculated RDO both modes, and the comparison with the threshold value t_0 , less than T_0 , you can determine the best mode for the current 4x4, otherwise the smaller RDO is probably the best mode. d. RDO prediction direction may be the best mode of comparison obtained in step c adjacent pattern, and further to determine the best possible mode. If it is V, H or D, then compare these three modes not compare mode; if it is the other mode, never compare compare all other modes. e. After the completion of the current macro block 16 4x4 block mode decision calculation macroblock texture complexity to MSE is measured and compared with a threshold value T_1 , if greater than T_1 , early termination of the judgment, the above results obtained the best mode for the macroblock; if less than T_1 , the brightness 16x16 block mode decision.

f relationship the best mode in step b luminance and chrominance blocks according to the block, to determine the current macro block 116 may mode; statistically step e 14 Modes With 16 sub-blocks, the probability of the two largest mode; considering these possible modes, to finalize the 116 possible patterns. And SAD as a criterion to determine the 116 best prediction mode. g. Adaptively determined threshold TI step e to determine whether the early termination of the judgment and 116 adaptively determined according to step c TO 14 judgment to determine whether to terminate early.

Description:

H. 264/AVC帧内预测模式的快速选择方法技术领域 本发明涉及一种信息处理的视频编码的方法，特别是一种H. 264/AVC帧内预测模式的快速选择方法。

背景技术 自从ITU-T颁布了 H. 261后，基于块的混合编码方法作为框架被普遍采用并不断发展。国际标准H. 261、H.263、MPEG-U MPEG-2、MPEG-4等的制定极大推动了多媒体技术的应用。然而，随着新服务的大量增加和高清电视的即将普及，以及现有传输媒体如Cable Modem、Xdsl, UMTS较低的传输码率，迫切需提高编码效率；同时还要求视频编码标准能够适应现存或未来的网络。这些问题为新视频编码标准 H. 264/AVC的产生提供了切实的依据。

这些视频编码标准采用混合编码框架，即将输入视频图像分割成宏块，对其进行预测、变换、量化、熵编码。预测包括帧内预测和帧间预测，帧内预测可去除空间冗余，而帧间预测可去除时间上的冗余。

H. 264/AVC仍遵从这种基于块的混合编码方案，但包含了许多新特征：VCL层和NAL层分离、空间域的帧内估计、1/4像素精度运动估计、自适应块尺寸运动补偿、多参考帧运动补偿预测、通用B帧的概

H. Rapid selection FIELD 264 / AVC intra prediction mode of the present invention relates to an information processing method for video coding, particularly a H. 264 / AVC intra prediction mode selection method fast.

Background Since the enactment of the ITU-T H. 261, block-based hybrid coding method is generally preclude the use as a framework and continue to develop. Development of international standards H. 261, H.263, MPEG-U MPEG-2, MPEG-4, etc. greatly promoted the application of multimedia technology. However, with a substantial increase in new services and High Definition television is about to spread, as well as the existing transmission media such as Cable Modem, Xdsl, UMTS lower transmission rate, the urgent need to improve coding efficiency; and also to adapt to the requirements of video coding standard existing or future network. These problems for the new video coding standard H. 264 / AVC produces provide a practical basis.

念、低复杂16bit的整数变换和量化、环内去块滤波器和高效的熵编码。这些新特征一方面能够显著提高编码效率；另一方面具有网络友好性，可有效用于各种网络和应用环境。

其中H. 264/AVC为提高帧内图象的编码效率，采用帧内预测去除空间冗余性。这种帧内预测技术依然以宏块为处理单元，即通常将一帧图象分割成16x16的图象块进行编码。然而基本处理块单元相对于其它标准由8x8降为4x4，即变换采用4x4的变换矩阵。因此亮度宏块支持两种帧内预测块尺寸：4x4和16x16，而色度仅支持8x8。

每个进行帧内预测的当前块利用邻块的相邻像素值来预测当前块像素值，然后计算残差、变换、熵编码。H. 264/AVC标准中每个4x4块的帧内预测包括9种预测模式（如附图1）。而亮度16x16和色度8x8仅支持4种预测模式，分别为：垂直（V）、水平（H）、直流（D）和平坦模式（P），分别简称为V、H、D和P模式。

同时这项技术也被应用到帧间编码之中，即对当前宏块进行完帧间预测后，接着对它进行帧内预测，以率失真优化为最终判决依据，率失真小的模式被确定为最终模式。这样既可以提高帧间编码宏块的编码效率，又可以提高容错能力。

例如对某帧（I帧、P帧或B帧皆可）中当前宏块M进行帧内预测编码，利用相邻块与当前块的紧挨的像素值来预测当前块的像素值，然后用当前块的像素值减去预测值得到残差矩阵，将此残差矩阵经过变换、量化、熵编码得到该宏块的编码码流。

但利用邻块像素值的预测方式是不同的，因此产生了多种预测模式（如附图1）。

在判断这些模式中的最优模式，需要一个衡量的准则，目前采用了率失真优化（RDO-rate distortion optimization），其定义如下： $J_{MODE} = S_k(I_k | Q^{MODE}) = I(Q) + \lambda \cdot MODE \cdot REC(\lambda \cdot h(I, Q))$ 公式(1) 式中Affie表示编码后总的失真，及 λ 表示总的码率。； λ 为拉格朗日

These hybrid video coding standard coding framework, ie input video image is divided into macro blocks, its prediction, transform, quantization and entropy coding. Including intra prediction and inter prediction, intra prediction can remove spatial redundancy while inter prediction removes temporal redundancy.

H 264 / AVC still comply with this hybrid coding scheme based on block, but contains many new features: VCL NAL layer and the layers were separated, the intra spatial domain estimates, 1/4-pixel accuracy motion estimation, adaptive block size motion compensation, multiple reference frame motion compensation prediction, a B frame generic concept, low complexity 16bit integer transform and quantization, and the inner loop deblocking filter efficient entropy coding. These new features on the one hand can significantly improve the coding efficiency; on the other hand has a network-friendly, effective for a variety of network environments and applications.

Where H. 264 / AVC intra coding efficiency to improve the image, preclude the use of spatial intra prediction to remove redundancy. This technology is still the intra prediction macroblock processing unit, i.e., typically a 16x16 pixel image into image blocks are encoded. However, the basic processing unit block relative to other standards reduced by the 8x8 4x4, 4x4 or transfer using the transformation matrix. Thus luminance macroblock supports two intra prediction block size: 4x4 and 16x16, and chroma only supports 8x8. each current block using the intra prediction adjacent pixel value of a block adjacent to the current block the prediction pixel value, and computing residuals, transform, entropy coding. H. 264 / AVC standard, intra prediction for each 4x4 block comprising 9 prediction modes (e.g., Fig. 1). And the luminance and chrominance

参数，通过确定编码选择项使编码后的失真和码率总和最小。

本发明在色度宏块和亮度16x16块的模式判决中采用了一种简化的判决准则为 SAD (Sum of Absolute Difference),其定义如下：
$$N-l \quad N-l \quad sad = YLabs^j - r^j$$
 公式 (2) $l=0 \quad y=0$ 式中 (l, y) 为原始信号， r^j (l ，力是帧内预测后的预测信号， N 等于8 (色度块) 或16 (亮度块)。

H. 264/AVC校验模型对于当前宏块帧内预测模式的确定给出了一种全搜索方法，以色度宏块的帧内预测模式为外循环，以亮度宏块的两种块模式为内循环，即对应每一种色度宏块的帧内预测模式，都要分别比较亮度宏块14和116的所有模式。故这种方法能够扫描到所有的模式，因此编码质量最高，运算量也是最大的。搜索的模式总数为 $4 \times (16 \times 9 + 4)$ ，共592个，这意味着为了得到一个宏块的最佳模式，需要计算592个不同的Jmwh。这使H. 264/AVC在现有的硬件条件下实现编码是不可能的。

发明内容 本发明的目的是提供一种H. 264/AVC实时编码器的帧内预测模式的快速选择方法，对比于全搜索方法，它可以在信噪比和码率变化很小的同时，可使编码速度平均提高6倍以上。

为达到上述目的，本发明采用下述技术方案：一种H. 264/AVC实时编码器帧内预测模式的快速选择方法，在H. 264/AVC编码器中，帧内编码采用帧内预测方法，其特征在于利用当前宏块的色度块和亮度块帧内预测模式间的相关性、4x4亮度块 (14)和16x16亮度块 (116)的模式选择之间的相关性以及亮度块预测模式内部的相关性，以简化的率失真优化 (RDO)为判决依据，快速进行帧内预测模式选择。

快速选择的步骤如下：a)确定色度块的模式选择具有独立性，即亮度宏块无论选择116还是14，色度宏块模式都是一致的，从而首先对色度预测模式以SAD为准则进行独立判决，确定色度的最佳预测模式。

b)确定了色度的最佳预测模式后，为亮度块模式判决提供如下依据：对于亮度116

16x16 8x8 only supports 4 prediction modes, namely: vertical (V), horizontal (H), direct current (D) and the flat mode (P), respectively referred to as V, H, D and P mode.

At the same time this technology is also being applied to the inter-coding, ie after the current macroblock after inter prediction, intra prediction followed it to the rate-distortion optimization for the final judgment on the basis of a small distortion rate mode is determined as the final mode. This will not only improve the coding efficiency of inter-coded macroblock, but also can improve fault tolerance. for example a frame (I frame, P frame or B frame can) M Current Macroblock intra prediction coding using pixel values of blocks adjacent to the current block to the next pixel value prediction of the current block and then subtracting the pixel values of the current block to the prediction residual matrix worth, this residual matrix after the transformation, quantization, entropy encoding to obtain the code stream of the macroblock. but the use of a block adjacent pixel values are different ways of prediction, thus producing a variety of prediction modes (e.g., Fig. 1).

In these modes determine the optimal mode, a measure of the criteria currently used in the rate-distortion optimization (RDO-rate distortion optimization), which is defined as follows: $JMODE, Sk, Ik \quad IQ \quad MODE) = IQ) \quad MODE \quad REC (k 'h \quad IQ)$ formula ? where Affie encoded representation of the total distortion, and dish represents the total rate.; $ljwwff$ and gt ; 0 is Lagrange parameters determined by the encoding choice after encoding rate and the sum of the minimum distortion.

In the present invention preclude the luminance and chrominance macroblock mode decision in 16x16 blocks with a

而言,当色度块选择D模式,亮度可以选择D、H、V、P模式;当色度块选择H模式,亮度可以选择D、H、P模式;当色度块选择V模式,亮度可以选择D、V、P模式;当色度块选择P模式,亮度可以选择D、P模式。而对于亮度14,18选择的模式是亮度14下绝大多数4x4块可能选择的模式。

c)对于宏块,14的选择概率远远大于116,因此我们首先进行14下4x4块的模式判决;根据邻块的模式判断得到当前块的可能模式,并和由步骤b确定的可能模式,分别计算这两种模式下的RD0,并与阈值TO比较,小于TO,则可确定为当前4x4的最佳模式,否则较小值RDO的为可能最佳模式。

d)比较步骤c中得到的可能最佳模式的预测方向邻近的模式,进一步确定可能的最佳模式。如果它是V、H或D,则比较这三种模式中未曾比较的模式;如果它为其它模式,则比较所有其它未曾比较的模式。

e)在完成当前宏块的16个4x4块的模式判决后,计算宏块纹理复杂度,以MSE来衡量,并与阈值TI比较,如果大于TI,提前终止判决,上述得到的结果为宏块最佳模式;如果小于TI,则进行亮度16x16块的模式判决。

f)根据步骤b中色度块的最佳模式与亮度块的关系,确定当前宏块116的可能模式;统计步骤e中14情况下16个子块的模式,得到概率最大的两种模式,综合考虑这些可能模式,最终确定116的可能模式。并以SAD为准则,确定116的最佳预测模式。

g)自适应确定步骤e所述的阈值TI来判断是否提前终止116判决和自适应确定步骤C所述的TO来判断是否提前终止14的判决。

本发明综合利用上述方法进行帧内预测模式选择。实验结果显示,可以提高编码速度数倍,且能够保证编码的码率和PSNR与全搜索方法相比变化不大。

simplified decision criterion for the SAD (Sum of Absolute Difference), which is defined as follows: $NI \text{ NI sad} = Y \text{ Labs } j) - r J))$ formula $? / = 0 y = 0$ where $/ (? , y)$ is the original signal, $P (/ , force$ is the intra prediction after prediction signal, N is equal to 8 (chroma blocks) or 16 (luminance block).

H. 264 / AVC calibration model for determining an intra prediction mode of the current macroblock is given a full search method, the intra-chrominance prediction mode macroblock outer loop, two luminance macroblock block modes within the cycle, i.e., corresponding to each macroblock intra chrominance prediction mode, must compare the luminance macroblock 14 and 116, respectively, of all modes. Therefore, this method is capable of scanning to all modes, so the encoding of the highest quality, the amount of computation is also the largest. The total number of search mode is $4 \times (16 \times 9 \times 4)$, a total of 592, which means that in order to get the best mode of a macro block, need to calculate 592 different Jmwh. This enables H. 264 / AVC coding is not possible to achieve under the existing hardware. object of the present invention is to provide a method to quickly select H. 264 / AVC real-time encoder intra-prediction mode, in contrast to the full search method, which can be varied in a small signal to noise ratio and bit rate At the same time, it can increase the speed to more than 6 times.

To achieve the above object, the present invention adopts the following technical scheme: quick selection method for H. 264 / AVC real-time encoder intra-prediction mode, in H. 264 / AVC encoder, the intra-frame coding using intra prediction method, wherein using the correlation chroma block and luma block intra prediction mode of the current macroblock between correlation 4x4 luminance blocks (14) and 16x16

附图说明 图1是本发明的一个实施例的信息处理视频编码框图。

图2是本发明的一个实施例的处理流程图 图3是4x4块帧内预测方向图 图4是当前块的邻块位置图 图5是116和14的率失真变化图 具体实施方式 以下结合附图描述本发明的H. 264/AVC实时编码器的帧内预测模式的快速选择 方法的一个优选实施例：在校验模型Jm72编码器上，具体工作环境为2. 4GHz CPU的PC机，针对I帧序列，编码序列为1111，熵编码采用了 CABAC，量化参数分别为14、28，亮度14的RDO 阈值TO和宏块复杂度的阈值TI可自适应计算得到，采用下述方法：在我们实验过程中，发现校验模型提供的全搜索方法忽略了模式选择间的相关性。在实际中，宏块的色度块和亮度块帧内预测模式间、4x4 亮度块（14）和16x16亮度块（116）的模式选择之间以及亮度块预测模式内部都具有较强的相关性。如果能充分利用这些相关性，就可以快速有效的确定帧内预测的最佳模式。

(1)参见图1和图2,首先，可以观察到亮度宏块无论选择116还是14，色度 宏块模式都是一致的，因此可以确定在整个模式选择过程中色度块的模式选择具有独立性，故色度预测模式可以在亮度判决之前进行独立判决色度块的最佳预测模式可以为亮度宏块模式判决提供依据。尽管对于亮度116与色度18模式选择不能达到完全一致，但存在着联系：当色度块选择D模式，亮度可以选择D、H、V、P模式；当色度块选择H模式，亮度可以选择D、H、P模式；当色度块选择V模式，亮度可以选择D、V、P模式；当色度块选择P模式，亮度可以选择D、P模式；而对于亮度14与色度18，18选择的模式是亮度14下绝大多数4x4块最有可能选择的模式。亮度块内部模式选择之间存在着优先性，V、H、D模式相对于其它模式具有优先选择性。14 情况下，通过相邻块（图4）的模式可以判决得到当前块的最优先选择模式。对于优先选取的模式，其方向上相邻的模式也具有很大的选中可能。

实际中，14的选择概率远远大于116,因此在判决过程上优先检测14。在14情况下

luminance blocks (116) and mode selection between the brightness correlation block prediction mode inside to simplify the rate distortion optimization (RDO) as the basis for judgment, fast intra prediction mode selection.

Step

Quick Select as follows: a) to determine the color mode selected block independence, namely luminance macroblock Whether you choose 116 or 14, chroma macroblock mode are the same, so first chroma prediction mode SAD as a criterion for independent judgment to determine the best prediction mode chromaticity. b) determining the optimum prediction mode chroma after brightness mode provides quick decision on the basis of the following: 116 for brightness, when the chroma D block selection mode, the brightness can choose D, H, V, P mode; H when the chroma block selection mode, the brightness can choose D, H, P mode; when V chrominance block selection mode, the brightness can choose D, V, P mode; P mode is selected when the chroma blocks, the brightness can choose D, P mode. The luminance 14, 18 the selected mode is the vast majority of 4x4 luminance block mode may be selected. c) For the macro block, far greater than the probability of selection 14 116, so we first conducted under the 14 4x4 blocks mode decision; according to the adjacent block pattern may be judged from the current block mode, and a and b are determined by step possible modes are calculated RD0 both modes, and the comparison with the threshold value tO, less than T0, you can determine the best mode for the current 4x4, otherwise the smaller RDO is probably the best mode.

Prediction direction

在确定的16个4x4块最佳模式中，概率最大的两种模式很可能与116的最佳模式一致，因此可以将之色度模式联合进行检验，最终确定116最佳模式。

H. 264/AVC编码器采用RDO为模式判别准则，但其结构复杂，尤其写码流相当耗时间，因此可以根据具体模式采用不同的判决依据：①色度模式选择不依赖亮度的模式选择，且其编码比特只包含色度模式和色度残差两部分，因此采用SAD就可以反映不同模式间率失真的差别；②在亮度14的4x4块编码比特包括块模式和残差的熵编码。因为4x4块尺寸较小，其块模式编码的比特数和不同模式间的残差编码比特数的差异会导致率失真的变化很大，故选择率失真作为14的模式判决依据。且4x4块的率失真的分布比较集中，可在分布范围中定义一个阈值T0，如果在某一模式下其率失真小于T0，可认为达到最佳模式；③116编码比特主要包括宏块模式和残差的熵编码，故SAD能够反映率失真的变化，以其作为模式选择的依据。

由于116和14随宏块纹理复杂度的RDO曲线是不同的（图5），在模式选择时还应考虑宏块的纹理复杂度，并以宏块纹理复杂度的某个值作为阈值T1来确定二者间的判决，即在纹理平坦区域（小于T1），116生成的重建图像能更好地与原始图像相匹配，选择116概率大于14，随着纹理复杂度的逐渐增加，116在细节内容上率失真逐渐增加，当大于或等于T1时将大于14的率失真，这时14更适合宏块编码模式选择。衡量纹理复杂性的方式有很多种，如算术平均、几何平均、中值、方差和均方误差。本发明选用了均方误差衡量宏块的纹理复杂度。

本发明充分利用上述的相关性，将之有机结合在一起，根据RDO准则确定不同模式的判决依据，得到帧内预测模式的快速选择方法，其流程如图2所示，具体步骤实施如下：①对当前宏块的色度块进行帧内预测，以SAD为判决依据，确定色度宏块帧内的最佳模式。

②对当前宏块的亮度块进行14帧内预测，首先根据当前4x4块的相邻块（图2）的模式确定可能模式，将之和色度块的最佳模

d) may be the best mode of comparison obtained in step c adjacent patterns RD0, to further determine the best possible mode. If it is V, H or D, then compare these three modes not compare mode; if it is the other mode, never compare compare all other modes. e) after the completion of the current macro block 16 4x4 block mode decision calculation macroblock texture complexity to MSE is measured and compared with a threshold value T1, if greater than T1, early termination of the judgment, the above results obtained the best mode for the macroblock; if less than T1, the brightness 16x16 block mode decision. f) the relationship between the best mode in step b luminance and chrominance blocks according to the block, to determine the current macro block 116 may mode; statistically step e 14 Modes With 16 sub-blocks, the probability of the two largest mode; considering these possible modes, to finalize the 116 possible patterns. And SAD as a criterion to determine the 116 best prediction mode. g) determining the adaptive step e T1 the threshold to determine whether early termination TO 116 and adaptive decision determining the C step 14 to determine whether to terminate the decision in advance. utilization of the present invention, the method described above intra-prediction mode selection. Experimental results show that the coding speed can be increased several times, and to ensure little change in coding bit rate and PSNR compared with the full search method.

Figure 1 is a block diagram of an information processing video encoding of one embodiment of the present invention.

2 is a process flow diagram showing an example of an embodiment of the present invention 3 is a block 4x4 intra prediction of the current direction of FIG. 4 is a block diagram block location 5 o is the

式相比较, 确定最有可能块模式。计算最可能模式的 J_m , 与阈值 T_0 比较, 如果小于 T_0 , 即可确认为最佳模式, 跳转到步骤⑦。否则跳转到步骤③。

③计算另一模式(上面被排除的模式)的 J_m , 与步骤②计算的 J_a^{de} 比较, 得到最小值的模式。与 T_0 比较, 如果小于 T_0 , 确认为最佳模式, 跳转到步骤⑦。

④选择最小 J_a^{de} 的模式作为候选模式, 计算与其预测方向相邻的两个模式的 J_{mode} , 比较这三种模式的 J_{mmk} , 如果邻近模式的 J_{mm}^{de} 值都大于候选模式的 J_{mode} , 跳转到步骤⑦。否则到步骤⑤。

⑤如果最小 J_m 的模式是V、H或D, 计算这三种模式在上述过程未曾计算的 J_{mode} , 比较得到最小 J_m 的模式, 跳转到步骤⑦; 如果是其它模式, 跳转到步骤⑥。

⑥计算其它剩余模式, 得到最小 J_m 模式, 跳转到步骤⑦。

⑦该最小模式为该4x4块的最佳模式。

⑧循环步骤②—⑦, 得到16个4x4块的最佳模式。计算宏块复杂度, 如果大于阈值 T_1 (图3), 则该宏块为14, 结束判决。否则计算14情况下整个宏块的 J_{mode} , 跳转到步骤⑨。

⑨根据色度块的最佳模式, 确定当前宏块116的可能模式。统计14情况下16个子块的模式, 得到概率最大的两种模式, 如果属于V、H或D, 它们是116的可能模式; 如果为其它模式, 则P模式成为116的可能模式。

⑩比较上述确定的116可能模式, 得到116最佳模式, 计算其 J_m 。与14情况下的宏块最小 J_{motk} 比较, 得到最佳块尺寸和其对应的最佳预测模式。

通过上述方法可以发现, 本发明在通常情况下帧内预测编码需要检测的色度模式为4种, 亮度14的每个4x4块最少1种或最多9种。如果提前终止的话, 不需要检测亮度116, 否则亮度116检测最少2种或最多4种

rate of change of the distortion 116 of FIG. 14 and the specific embodiment a preferred dRAWINGS the present invention is described in H. 264 / AVC real-time encoder for fast intra mode prediction selection method Example: the calibration model Jm72 encoder, specific working environment of the PC 2. 4GHz CPU machine for I-frame sequence, the coding sequence of 1111, entropy coding using CABAC, the quantization parameters are 14,28, RDO brightness threshold 14 and threshold value T_1 to macroblock adaptive complexity can be calculated using the following method: in our experiment, we found the full search method validation model provides ignore the correlation between the mode selection. In practice, between the luminance and chrominance block block intra prediction mode macroblock, and a brightness prediction mode of block internal 4x4 luminance blocks (14) and 16x16 luminance block (116) has a mode selection between a strong correlation. If you can take advantage of these dependencies, you can quickly and efficiently determine the best mode of intra prediction.

(1) See Figures 1 and 2, first, the luminance macroblock can be observed regardless of the selected 116 or 14, chroma macroblock mode are the same, so you can determine the entire mode selection process chroma blocks mode Select independence, so chroma prediction mode judgment before brightness for optimal prediction mode independent judgment chroma blocks may provide the basis for luminance macroblock mode decision. Although the luminance and chrominance 18 116 Select mode can not achieve exactly the same, but there is a link: D if chrominance block selection mode, the brightness can choose D, H, V, P mode; when chroma block selection H mode, the brightness can Select D, H, P mode;

模式，这种情况下几乎涵盖了所有的模式。但总体检测的模式数极大地降低。

H. 264/AVC的帧间编码中也使用了帧内预测编码技术，上述方法同样适用，只是略作改动即可：在步骤2的14帧内预测中，当该4x4块的邻块模式为帧内编码模式不必考虑直接使用即可；如果邻块为帧间模式，则可默认其帧内预测模式为直流模式(D)。

(2)计算阈值TO的自适应方法 H. 264/AVC采用的是类DCT变换。公式(3)为4X4DCT整数变换，E为4X4亮度块经预测后残差矩阵，X为类DCT整数变换矩阵，I为整数变换后量化前的系数矩阵。

$I = XEXT$ 公式(3) 其中： $X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & 1 & -1 & -2 \\ 2 & 1 & -1 & -2 \end{bmatrix}$ 设DC为变化后的直流系数，s是当前4x4块的原始像素值，s'是预测后重建块的像素值，如公式(4)所示。如果DC趋近于0，那么其它交流系数AC也趋于0，整个变换后的矩阵趋向于0矩阵，因此可省略后面的处理过程。

$DC = Y^E(x, y) = \text{公式(4)}$ $JCs = 0$ y^O $JCs = 0$ $\> 0$ 在H. 264/AVC中如果变换系数经过量化后，其量化值绝对值小于或等于1时，则残差矩阵可视为0矩阵。即满足下面的不等式： $|DC| \< (29 - te - qp\_const) / quant\_coef[qp\_rem][0]$ 公式(5) 其中： $qp\_Per = QP$ ， $qp\_rem = QP \% 6$ ， $qp\_bits = qp\_per + 15$ ， $qp\_const = (1 \ll qp\_bits) / 6$ ，QP为量化参数。

对于每个4x4块，公式(6)给出SAD的计算，具体如下： $\< \text{公式(6)}$ $x=0$ $y=0$ 很明显下面不等式成立： $|yZf4(Sixiy) - S'^y)| \< XS s(x, y) - s'(x, y) | X - Q$ 少 $= 0$ XaO 少 $= 0$ 艮P $|ZX7| \< \< 4D$ 。

则如果下式成立，则满足0矩阵的判断。

$SAD \< (2q - bi, s - qp\_const) / quant\_coef[qp\_rem][0][0]$ 公式(7) 阈值TO由4x4块的RDO确定，如公式8所示，mode表示所选的预测模式，J{包括模式类型和残差的熵编码。

when the V chrominance block selection mode, the brightness can choose D, V, P mode; P mode is selected when the chroma blocks, the luminance can choose D, P mode; and luminance 14 and chrominance 18 18 selected mode is the brightness of the vast majority of 4x4 blocks 14 are most likely to choose the mode. Exists between luminance block unit selection mode priority, V, H, D mode relative to other modes have priority selectivity. Under 14 case, the adjacent block (Fig. 4) of the current block mode decision to give top priority selection mode. For priority mode selected, the direction of the adjacent mode is also possible to select great.

In practice, the selection probability of 14 is far greater than 116, so the decision process priority detector 14. In 14 cases in 16 4x4 block to determine the best mode, the maximum probability of the two modes is possible with the 116 best mode consistent, so it can be combined with the color mode to test, and ultimately determine the best mode of 116.

H 264 / AVC encoder uses RDO mode as criterion, but its structure is complicated, especially in writing stream quite time-consuming, so the decision can be based on different modes depending on: ? chroma brightness mode selection does not depend on mode selection, and it contains only the encoded bit color mode and chroma residuals of two parts, so using SAD may reflect different patterns of distortion rate differences; ? 4x4 luminance block coding in 14-bit mode and includes a block of residuals entropy coding. Because 4X4 block size is small, residual coding differences between the number of bits of coded block pattern which the number of bits between the different modes and can lead to significant changes in the rate of distortion, distortion was chosen as the model 14 rate decision basis. And 4x4

$J_{mode}(\hat{ode}) = SSD + A \cdot R_{mode}$ 公式 (8) 其中, $S^A = t^A$;
 $(;c, \>0;7(X, \text{力} |, /^A, \text{州是预测后残差经过变换、量化、反量化、} X=0 \text{ 少} =0 \text{ 反变换后与参考块的值相加得到, 故下面不等式成立: } |s(x, y) - p(x, y)| \< |s(x, y) - s'(x, y)| \text{ 因此 } SSD \< (XSI \< x, y) - s(x, y) |^2) \< (过 |s(x, y) - s(x, y)|) JC=0 \wedge =0 \text{ 故 } SSD \< SAD1' \text{ 即 } SSD \< -qp\_const) / quant\_coef [qp\_rem][0][0])^2 \text{ 这时, 该块残差量化系数将趋于0, 因此得到 } TO。$

$TO = ((2^{-bUs} - qp\_const) / quant\_coef [qp\_rem][0][0])^2$ 实际中可能存在差别, 因此设置一个校正因子 $a = 0.5$: $rO = ax((2q\_bus - qp\_const) / quant\_coef [qp\_rem][0][0])^2$ (3) 计算阈值 TI 的自适应方法 在实际中, 当 TI 越大, 搜索 116 的可能性越大, 因此搜索的模式越多; TI 越小, 搜索 116 的可能性越小, 则搜索的模式变少。本发明通过当前宏块的前一宏块来确定当前 TI 的值, 具体方法如下: 初始化:
 $TI = \< \< 20$ 第一个宏块: 依次进行 14 和 116 模式判决, 得到二者的 MSE, 将均值赋给 $T1$, 作为下一宏块的阈值。

其它宏块: 如果仅判决了 14 模式, 则 TI 不变。如果依次进行了 14 和 116 模式判决, 则二者的 MSE 的均值赋给 TI 。

测试三个纹理复杂程度不同的 CIF 序列 Mobile、Bus 和 Foreman。实验结果表明, 本快速帧内预测模式选择方法比 H.264/AVC 原有的全搜索方法可提高帧编码速度大约 6 倍左右。在速度提高的同时, 平均信噪比变化约 0.1db 左右, 码率稍有变化。

本发明的第二个实施例是: 针对 P 帧编码, 编码序列为 I PPP, 运动估计搜索公式 (9) 公式 (10) 公式 (11) 范围 ± 16 , 5 个参考帧, 运动估计算法为全搜索算法, 熵编码采用了 CABAC, 量化参数分别为 14、28 亮度 14 4x4 块的 RDO 的阈值 TO 和宏块复杂度的阈值 TI , 同样采用上述的自适应方法。仍然测试三 CIF 序列 Mobile、Bus 和 Foreman。实验结果表明, 本发明

block distortion rate distribution more concentrated, you can define a distribution range of threshold TO , if at a certain rate-distortion mode it is less than TO , it can be considered the best mode; ? 116 coded bits including macroblock mode and residues differential entropy coding, it is possible to reflect the rate of change SAD distortion model as the basis for its selection.

Since the 116 and 14 with the macroblock RDO curve texture complexity is different (Fig. 5), in the mode selection should also consider the texture complexity macroblock, and a value of the macro blocks as texture complexity judgment threshold TI is determined between the two, i.e., the flat region texture (less than TI), 116 generates a reconstructed image of the original image are able to better match the selection probability is greater than 14 116, with the increasing complexity of the texture, 116 rate-distortion gradually increases in the details, when greater than or equal to TI distortion will be greater than the rate of 14, then 14 is more suitable macroblock coding mode selection. Measure texture complexity There are many ways, such as arithmetic mean, geometric mean, median, variance and mean square error. The present invention is selected to measure the mean square error macroblock texture complexity. full advantage of the present invention, the above-mentioned correlation, it will combine together to determine the sentence in accordance with the different modes according to RDO guidelines give fast intra prediction mode selection method, the process shown in Figure 2, concrete steps to implement as follows: ? chroma blocks of the current macroblock intra prediction, based on the judgment as to the SAD to determine the best mode of chrominance macroblock frame.

的快速帧内预测模式选择方法在帧间情况下比H. 264/AVC原有的帧间全搜索方法可提高编码速度大约25%左右。在速度提高的同时，平均信噪比变化约0.02db左右，码率几乎没有变化。

明 ¥ Pit IS

? the current macroblock luminance intra prediction block 14, the first (FIG. 2) may determine the mode according to the mode of the 4x4 blocks adjacent to the current block and the best mode of the chrominance blocks compared to determine most likely block mode. To calculate the most likely mode J_m dish, compared with the threshold T_0 , if less than T_0 , can be recognized as the best mode, jumps to step ?. Otherwise, skip to step ?.

? calculate another mode (above excluded mode) J_m dish, and step ? calculated J_a de comparison to give minimum mode. Compared with the T_0 , if less than T_0 , recognized as the best mode, jumps to step ?.

? select the smallest J_a pan mode as a candidate mode, calculate J_{mode} its prediction direction of two adjacent patterns, compare the three models J_{mmk} , if adjacent patterns J_{mm} are greater than J_{mode} candidate modes, Skip step ?. Otherwise to step ?.

? If the mode is the smallest J_m dish V , H or D , the calculation of the three modes in the above process has not been calculated J_{mode} , compare the pattern to give the smallest J_m dish, jumps to step ?; if other modes, jumping Go to step ?.

? calculate other remaining mode, a minimum and pan mode, skip to step ?.

? The smallest model in the best mode 4x4 block.

? ? a recycling step ?, to give 16 the best mode 4X4 block. Calculation macroblock complexity, if greater than the threshold value T_1 (Fig. 3), the macro block 14, the end of the sentence. Otherwise, the calculation for 14 cases J_{mode} entire macro block, a jump to step ?.

? according to the best mode for chroma blocks, block 116 determines that the current macro-possible model. Statistics 14 cases under 16 sub-blocks mode, the probability of obtaining the two largest models, if they are V, H or D, which are 116 possible model; if other modes, the P 116 model to become a possible model.

116 may Mode

? comparing said determined to give the best mode 116, calculated Jm dish. Compared with the macroblock minimum Jmotk 14 cases under optimal block size and its corresponding optimum prediction mode. can be found by the method described above, in the present invention is usually detected intraframe prediction encoding requires four kinds of color mode, each of the 4x4 luminance block or a least one kind of up to nine 14. If the early termination, then, without detecting the brightness of 116, 116 or the luminance detection least two or at most four modes, in this case covers almost all modes. But the overall number of pattern detection is greatly reduced.

H 264 / AVC is also used in inter-coded intra predictive coding technique, the above-described methods are equally applicable, with slight modification to: 14 intra prediction in step 2, when the 4x4 block o block mode is the intraframe coding mode can be used directly without having to consider; o If the block is the inter mode, which can be the default mode is the DC intra-forecast mode (D).

(2) calculating the threshold value TO adaptive method H. 264 / AVC is used in the DCT transformation. Equation (3) is 4X4DCT integer transform, E after 4X4 blocks of luminance prediction residual matrix, X is an integer of the DCT transform matrix, I is the coefficient

matrix after integer transform before quantization.

$I = XEXT$ formula (3) where: $X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & 1 & -1 & -2 \\ 1 & 1 & -2 & -1 \end{bmatrix}$ to set up the DC coefficient DC after the change, s current original 4x4 pixel block value, s' is the pixel value after the reconstruction block of prediction, such as (4) shown in equation. If the DC tends to zero, then the other AC coefficient AC tends to 0, the entire matrix transform matrix tends to 0, so the process can be omitted later.

$DC = Y E_{ix, y} = \text{formula ? } JCs = O y O$
 $JC = O$ and $gt; \square - 0$ in H. 264 / AVC, if after quantized transform coefficients, the absolute value is less than its quantization or equal to 1, the residual matrix can be regarded as 0 matrix. That satisfies the following inequality: $I DC$ Shu and It; $(29 - te - qp\_const) / quant\_coef [qp\_rem] [0] [0]$ Equation (5) where:
 $Vocational\_Per = QP$, $qp\_rem = QP$
 $percent\ 6, qp\_bits = qp\_per\ 15$, $qp\_const = (1 - \square\_bits) / 6$ 'QP is a quantization parameter.

For each 4x4 block, the equation (6) gives the SAD is calculated as follows:
 and = equation (6) $x = 0 y = 0$ Obviously the following inequality holds: $I yZf4$
 $(Sixiy) - S')$ \ and It; $XS s(x, y) - s'(x, y) |$
 XQ little less = 0 = 0 XaO Gen P | $ZX7 |$
 and It; and 4D.

If the formula is established, judgment is satisfied 0 matrix.

SAD and It; $(2q - bi, s - qp\_const) / quant\_coef [qp\_rem] [0] [0]$ Equation (7)
 threshold TO by the RDO 4x4 block is determined as in Equation 8, mode represents the prediction mode selected, J {including the mode type and residual entropy coding.

$Jmode(ode) = SSD A Rimode)$ Equation (8) where, $S = t$ and $gt; (; c$, and $gt; 0$; 7
 $(X$, force Shu, /, after the state after the

prediction residual transform,
 quantization, inverse quantization, $X = 0$
 $= 0$ after inverse transform less the value
 obtained by adding the reference block,
 so the following inequality holds: $|s(x, y) - p(x, y)|$ and $|s(x, y) - s'(x, y)|$. Thus
 SSD and $|X|$ and $|x, y - s \setminus x, y|$ 2)
 and $|x, y - s \setminus x, y|$) $JC = 0 = 0$ because
 SSD and $|x, y - s \setminus x, y|$ therefore, SSD and
 $|x, y - s \setminus x, y|$ Therefore, SSD and $|x, y - s \setminus x, y|$ /
 $quant\_coef[qp\_rem][0][0]$ 2 In this
 case, the quantized residual block
 coefficients will tend to 0, thereby
 obtaining 0.

$TO = ((2 \cdot bus - qp\_const) / quant\_coef[qp\_rem][0][0])$ may be present in the
 actual difference 2, so setting a
 correction factor $a = 0.5$: $rO = ax((2q - bus - qp\_const) / quant\_coef[qp\_rem][0][0])$ 2 (3) the calculation method of
 adaptive threshold TI in practice, when
 the TI , the greater the possibility of
 searching 116, so the more the search
 pattern; TI smaller, the less likely the
 search 116, the search pattern of the
 present invention becomes small by
 previous macroblock to determine the
 current value of TI current macroblock, as
 follows: initialization: $TI = 1$ and $|x, y - s \setminus x, y|$ and $|x, y - s \setminus x, y|$;
 20 the first macro block: 14 and 116
 sequentially mode decision to give both
 the MSE , the average assigned to $T1$, as
 the next macroblock threshold

Other macro block: If only the judgment
 of 14 mode, TI unchanged. If you turn 14
 and the mode decision 116, the mean
 value of both MSE assigned to TI . test
 three different degrees of complexity
 textures CIF sequence Mobile, Bus and
 Foreman. Experimental results show that
 the fast intra prediction mode selection
 method than H. 264 / AVC original full-
 search method can improve the I-frame
 encoding speed around about 6 times.
 Increase in speed, while the average

SNR change about 0.1db, rate changes slightly.

The second embodiment of the present invention are: for a P-frame coding, the coding sequence for the I PPP, motion estimation (11) Range Search formula (9) Equation (10) Equation plus or minus 16,5 reference frames, motion estimation algorithm for the full search algorithm, entropy coding using CABAC, the quantization parameters are 14,280 14 4x4 luminance block RDO threshold and threshold TI to macroblock complexity, as the above-mentioned adaptive methods. Still three CIF test sequences Mobile, Bus and Foreman. Experimental results show that the fast intra prediction mode selection method of the present invention than in the case of inter H. 264 / AVC original inter whole search method can improve the coding rate of about about 25 percent. Increase in speed, while the average SNR of about 0.02db about change, the rate was almost unchanged.

Ming □ Pit IS

Cited by: CN101888547 A ; CN102055980 A ; CN102075659 A ; CN102186070 A ; CN102215398 A ; CN102469308 A ; CN102740077 A ; CN102740077 B ; CN102946538 A ; CN102946538 B ; CN103096090 A ; CN103096090 B ; CN103314588 A ; CN103314588 B ; CN103581682 B ; CN104954805 A ; CN105120266 A ; TWI301723 B ; TWI594619 B ; US8593707 BB ; US8718130 BB ; US8948259 BB ; US9143786 BB ; US9167251 BB ; US9237346 BB ; US9270997 BB ; US9503730 BB ; US9788006 BB ; WO07041958 A1 ; WO09105948 A1 ; WO12065448 A1 ; WO12092761 A1 ; WO13064094 A1 ; WO13189205 A1 ;



[12] 发明专利申请公开说明书

[21] 申请号 200410096605.1

[43] 公开日 2006 年 6 月 7 日

[11] 公开号 CN 1784015A

[22] 申请日 2004.12.2
[21] 申请号 200410096605.1
[71] 申请人 中国科学院计算技术研究所
地址 100080 北京市海淀区中关村科学院南路 6 号
[72] 发明人 张冬明 沈燕飞 黄 晔 林守勋
余乐军

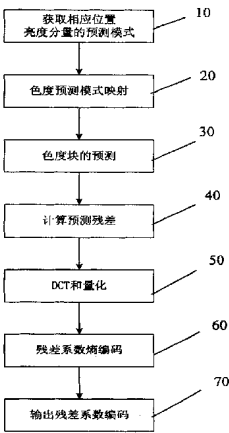
[74] 专利代理机构 北京泛华伟业知识产权代理有限公司
代理人 高存秀

权利要求书 2 页 说明书 8 页 附图 4 页

[54] 发明名称
一种帧内图像预测编码方法

[57] 摘要

本发明公开了一种帧内图像预测编码方法，包括对所述帧内图像进行亮度预测编码和色度预测编码，在进行色度预测编码时，根据亮度预测编码的预测模式对应得到一个色度预测模式，用该色度预测模式对所述帧内图像进行色度预测编码。本发明方法的优点：不必使用率失真优化方法来挑选预测模式，有效降低了色度分量帧内模式编码的复杂度；由帧内亮度预测模式可以映射一种唯一的帧内色度预测模式，只要在解码端采用同样的选择策略，就不必对色度分量的空域预测模式进行编码，进一步降低编解码的复杂度；在利用色度预测模式，实现对色度块像素值的预测过程中， 4×4 块预测使得 4 个 4×4 色度分量块的预测并行进行，提高了对色度块像素值进行预测时的效率。



1、一种帧内图像预测编码方法，包括对所述帧内图像进行亮度预测编码和色度预测编码，其特征在于，在进行色度预测编码时，根据亮度预测编码的预测模式对应得到一个色度预测模式，用该色度预测模式对所述帧内图像进行色度预测编码。

2、根据权利要求1所述的帧内图像预测编码方法，其特征在于，所述亮度预测模式包括：VERT_16x16_PRED、HOR_16x16_PRED、DC_16x16_PRED、PLANE_16x16_PRED、VERT_4x4_PRED、HOR_4x4_PRED、DOWN_LEFT_4x4_PRED、DOWN_RIGHT_4x4_PRED、VERT_RIGHT_4x4_PRED、DC_4x4_PRED、HOR_DOWN_4x4_PRED、VER_LEFT_4x4_PRED、HOR_UP_4x4_PRED；所述色度预测模式包括：VERT_PRED_CHROMA、HOR_PRED_CHROMA、DC_PRED_CHROMA。

3、根据权利要求2所述的帧内图像预测编码方法，其特征在于，亮度预测模式与色度预测模式间的对应关系为：

VERT_16x16_PRED 对应 VERT_PRED_CHROMA；
HOR_16x16_PRED 对应 HOR_PRED_CHROMA；
DC_16x16_PRED 对应 DC_PRED_CHROMA；
PLANE_16x16_PRED 对应 DC_PRED_CHROMA；
VERT_4x4_PRED 对应 VERT_PRED_CHROMA；
HOR_4x4_PRED 对应 HOR_PRED_CHROMA；
DOWN_LEFT_4x4_PRED 对应 DC_PRED_CHROMA；
DOWN_RIGHT_4x4_PRED 对应 DC_PRED_CHROMA；
VERT_RIGHT_4x4_PRED 对应 DC_PRED_CHROMA；
DC_4x4_PRED 对应 DC_PRED_CHROMA；
HOR_DOWN_4x4_PRED 对应 DC_PRED_CHROMA；
VER_LEFT_4x4_PRED 对应 DC_PRED_CHROMA；
HOR_UP_4x4_PRED 对应 DC_PRED_CHROMA。

4、根据权利要求1所述的帧内图像预测编码方法，其特征在于，所述的帧内亮度预测模式由所选择的亮度分量块决定，而亮度分量块的选择方法根据色度分量预测时所采用的块的大小分为两类：

a) 当色度分量按照 8x8 进行预测时，宏块中的 8x8 色度分量块只有一种预测模式，该预测模式通过当前宏块的 16x16 亮度分量块中的某一 4x4 亮度分量块的预测

模式来映射；

b) 当色度分量按照 4x4 进行预测时，宏块中的 8x8 色度分量块分成 4 个 4x4 块，这 4 个 4x4 块分别有自己的预测模式，每个块的预测模式通过当前宏块的 16x16 亮度分量的某一 4x4 块的预测模式来映射。

5、根据权利要求 4 所述的帧内图像预测编码方法，其特征在于，在亮度分量块的选择方法 a) 中，4x4 亮度分量块的一个优选选择方案是选用 16x16 亮度分量块中左上角的 4x4 块。

6、根据权利要求 4 所述的帧内图像预测编码方法，其特征在于，在亮度分量块的选择方法 b) 中，4x4 亮度分量块的一个优选选择方案是将当前宏块的 16x16 亮度分量块等分成 4 个 8x8 亮度分量块，每个 4x4 色度分量块对应一个 8x8 亮度分量块，每个 4x4 色度分量块选择相应 8x8 亮度分量块中左上角的 4x4 块的预测模式来映射。

7、根据权利要求 1 所述的帧内图像预测编码方法，其特征在于，帧内亮度预测模式与帧内色度预测模式间一一对应的映射关系是在帧内图像纹理分析和统计的基础上得到的。

8、根据权利要求 1 所述的帧内图像预测编码方法，其特征在于，帧内色度的预测编码过程根据预测所选用的色度块大小的不同而不同，包括：

a) 8x8 块预测，对 8x8 色度块进行整体预测；

b) 4x4 块预测，整个 8x8 色度分量块分成 4 个 4x4 色度分量块，根据每个 4x4 色度分量块各自的色度预测模式实现色度预测编码。

9、根据权利要求 1 所述的帧内图像预测编码方法，其特征在于，所述的 4x4 块预测时，4 个 4x4 色度分量块的预测并行进行，在预测时当所需要的参考像素不可用时，使用最近的可用像素点来代替。

一种帧内图像预测编码方法

技术领域

本发明属于视频及图像编解码领域，特别涉及视频帧内图像基于空域预测的编解码技术。

背景技术

视频编码是计算机科学领域的热点问题，它旨在发现数字视频的最有效的编码格式。在视频编码中，常把视频图像分为帧内图像和帧间图像两种。其中，帧内图像的视频编码仅使用图像自身提供的信息进行空间上的预测，它的编解码不依赖于其他图像，可以单独解码；而帧间图像的编码要消除的是视频在时间域上的冗余信息，它的编解码依赖于它之前的一个或多个图像。对帧内图像进行视频编码时，因为在帧内图像中，相邻图像块的像素具有很大的空域相关性，所以需要采用空域预测编码方法以提高编码效率。空域预测编码方法的主要思想是以相邻块为参考像素，按照不同的方向预测待编码块的像素值，把预测结果称为预测块，其中的待编码块是指没有进行压缩编码的原始数据块；将待编码块的像素值与预测块的像素值作差，所得到的结果称为残差；再对残差进行 DCT 变换和量化得到残差系数，最后对结果进行无损的熵编码并输出。待编码块和预测块匹配程度越高，编码的效率越高。为了提高预测块像素值与当前块像素值的匹配程度，往往设置多种预测模式。

现有的 H.264（又称 MPEG-4 Part10 或 AVC）在对帧内图像预测中，是以宏块为单位进行编码的，所谓宏块包括一个 16x16 块的亮度分量和相应位置的两个 8x8 块的表示色度的分量。因此在对帧内图像进行预测编码时，包括亮度预测编码和色度预测编码。亮度预测编码是依据一定的亮度预测模式对宏块的亮度分量进行编码，而色度预测编码是依据一定的色度预测模式对宏块的色度分量进行编码。在现有技术中，亮度预测编码和色度预测编码是相对独立的，亮度预测模式的选择与色度预测模式的选择也是相对独立的。在 H.264 中，亮度预测模式分为 16x16 和 4x4 两类，其中 16x16 的预测模式有 4 种，4x4 的预测模式有 9 种；16x16 的 4 种预测模式具体

包括：VERT_16x16_PRED、HOR_16x16_PRED、DC_16x16_PRED、PLANE_16x16_PRED；4x4 的 9 种预测模式具体包括：VERT_4x4_PRED、HOR_4x4_PRED、DOWN_LEFT_4x4_PRED、DOWN_RIGHT_4x4_PRED、VERT_RIGHT_4x4_PRED、DC_4x4_PRED、HOR_DOWN_4x4_PRED、VER_LEFT_4x4_PRED、HOR_UP_4x4_PRED。

在 H.264 中，色度预测模式包括水平（VERT_PRED_CHROMA）、垂直（HOR_PRED_CHROMA）、DC（DC_PRED_CHROMA）和 Plane 预测 4 种。这 4 种色度预测模式的具体实现如图 1 所示，水平预测仅使用当前块左边的参考像素点（i、j、k、l、m、n、o、p）进行预测；垂直预测仅使用当前块上边的参考像素点（a、b、c、d、e、f、g、h）进行预测；DC 预测根据参考像素的可用性情况决定使用左边或上边或同时使用左边、上边和左上角（x）的参考像素点；Plane 预测需要使用左边和上边的参考像素。因此，当仅有左边参考像素可用时，只有水平预测和 DC 预测可供选择；当仅有上边参考像素可用时，只有垂直预测和 DC 预测可供选择；当左边和上边的参考像素均可用时，四种模式都可以选择；当左边和上边的参考像素均不可用时，即为左上角宏块，只有 DC 预测模式可以选择。除了最后一种情况，具体采用哪一种预测模式需要根据按照一定的规则来选择。

上述 4 种色度预测模式都有可能应用于实际，在做帧内色度预测编码时可能有多种色度预测模式可用，对色度预测模式的选择通常采用的方式是率失真（RD）优化规则，其中，R 是指码率开销，具体表示使用一种预测模式时编码模式和残差系数所需要的比特数，D 是指失真度，具体表示使用一种预测模式所重建的宏块与原始图像宏块之间的差值。R 和 D 成反比关系，当 D 较小的时候，R 就会增大；当 D 增加时，相应的 R 就会降低。所谓率失真优化，就是在编码所需的码率开销（比特数）和编码引起的图像失真二者之间找到最佳平衡点，使得码率开销 R 和失真度 D 同时尽可能地小。对色度分量而言，进行预测模式的率失真优化选择，就需要进行所有 4 种模式的预测，并对残差进行编码，记录每种预测模式的码率开销与失真之和，从中选择最小的那种模式作为最优预测模式。

参看图 2，现有的 H.264 帧内色度预测编码方法的操作过程包括：

- 1、利用某一种色度预测模式对色度块进行预测；
- 2、计算预测残差；
- 3、对残差进行 DCT 变换和量化处理，得到残差系数；

4、对残差系数做熵编码；

5、判断是否做完了所有可能的预测模式，若还有可能的预测模式，则跳转到步骤 1，若不再有新的可能预测模式，则继续以下操作；

6、对所有结果做率失真优化选择，输出最佳模式和残差系数的编码。

在步骤 2 中的残差实际上是一个可以用矩阵表示的残差块；

在步骤 3 中，经过 DCT 变换和量化处理，残差块由二维矩阵转变为一维数组，这个一维数组就是残差系数。其中的 DCT 变换是指对残差进行从空域到频域的转变，目的是进一步压缩数据，量化的实质也是进行压缩，将 0 附近的，比如 (-3, 3) 之间系数全部视作 0；将绝对值小于 10 的系数全部视作 +2/-2。

现有的 H.264 帧内色度预测编码已经完全可以实现对帧内图像的色度预测编码，但对 H.264 帧内色度预测编码过程进行仔细分析、观察后，发现存在以下一些问题：

(1) 增加预测模式，可以提高预测精度，但是当预测模式包含了水平、垂直、DC 预测之后，再增加更多的预测模式，不会对预测精度产生明显影响。

(2) 当预测模式增加时，用于预测模式编码的比特数就会增加，这就会带来更大的码流开销。

(3) 随着预测模式的增加，进行预测模式率失真优化选择的过程就会更加复杂，导致视频编码的复杂度增加。

同时，在视频编码过程中还存在以下规律和现象：

(1) 统计规律显示由于亮度分量和色度分量的纹理具有相似性，因此宏块的亮度分量和色度分量的预测模式呈现较强的相关性。而 H.264 的帧内预测编码中，亮度和色度的预测模式是独立选择的，它们的预测模式分别通过率失真优化方法来选择，也就是忽视了亮度分量和色度分量的预测模式间的相关性。

(2) 在现有的 H.264 视频编码框架中，色度分量的 PSNR（峰值信噪比）值平均高出亮度分量 2dB 左右。因此，色度分量的 PSNR 即使有一定的降低，也不是人的视觉可以感觉到的，并不会引起主观视觉效果的变化。也就是说在适当范围内牺牲色度分量的精度以换取编码效率的提高和码率开销的降低是切实可行的。

本发明正是根据上述规律，提出一种帧内图像预测编码方法，以克服现有的 H.264 视频编码技术所存在的码流开销大，编码复杂度高的缺陷。

发明内容

本发明的目的是克服现有的 H.264 帧内色度预测编码方法所存在的码流开销大, 视频编码复杂度高的缺陷, 提供一种低复杂度的帧内色度预测编码方法, 并且保证主观和客观的视频质量不会明显降低。

为了实现上述目的, 本发明提供了一种帧内图像预测编码方法, 包括帧内亮度预测编码方法和帧内色度预测编码方法, 在帧内色度预测编码方法中, 利用帧内亮度预测模式得到唯一的帧内色度预测模式, 实现帧内色度的预测编码。

上述技术方案中, 帧内亮度预测模式包括: VERT_16x16_PRED、HOR_16x16_PRED、DC_16x16_PRED、PLANE_16x16_PRED、VERT_4x4_PRED、HOR_4x4_PRED、DOWN_LEFT_4x4_PRED、DOWN_RIGHT_4x4_PRED、VERT_RIGHT_4x4_PRED、DC_4x4_PRED、HOR_DOWN_4x4_PRED、VER_LEFT_4x4_PRED、HOR_UP_4x4_PRED; 帧内色度预测模式包括: VERT_PRED_CHROMA、HOR_PRED_CHROMA、DC_PRED_CHROMA。

上述技术方案中, 帧内亮度预测模式和帧内色度预测模式间的对应关系为: VERT_16x16_PRED 亮度预测模式对应 VERT_PRED_CHROMA; HOR_16x16_PRED 对应 HOR_PRED_CHROMA; DC_16x16_PRED 对应 DC_PRED_CHROMA; PLANE_16x16_PRED 对应 DC_PRED_CHROMA; VERT_4x4_PRED 对应 VERT_PRED_CHROMA; HOR_4x4_PRED 对应 HOR_PRED_CHROMA; DOWN_LEFT_4x4_PRED 对应 DC_PRED_CHROMA; DOWN_RIGHT_4x4_PRED 对应 DC_PRED_CHROMA; VERT_RIGHT_4x4_PRED 对应 DC_PRED_CHROMA; DC_4x4_PRED 对应 DC_PRED_CHROMA; HOR_DOWN_4x4_PRED 对应 DC_PRED_CHROMA; VER_LEFT_4x4_PRED 对应 DC_PRED_CHROMA; HOR_UP_4x4_PRED 对应 DC_PRED_CHROMA。

上述技术方案中, 所述的帧内亮度预测模式由所选择的亮度分量块决定, 而亮度分量块的选择方法根据色度分量预测时所采用的块的大小分为两类:

a) 当色度分量按照 8x8 进行预测时, 宏块中的 8x8 色度分量块只有一种预测模式, 该预测模式通过当前宏块的 16x16 亮度分量块中的某一 4x4 亮度分量块的预测模式来映射;

b) 当色度分量按照 4x4 进行预测时, 宏块中的 8x8 色度分量块分成 4 个 4x4 块, 这 4 个 4x4 块分别有自己的预测模式, 每个块的预测模式通过当前宏块的 16x16 亮度分量的某一 4x4 块的预测模式来映射。

所述的亮度分量块的选择方法 a) 中, 4x4 亮度分量块的一个优选选择方案是选用 16x16 亮度分量块中左上角的 4x4 块。

所述的亮度分量块的选择方法 b) 中, 4x4 亮度分量块的一个优选选择方案是将当前宏块的 16x16 亮度分量块等分成 4 个 8x8 亮度分量块, 每个 4x4 色度分量块对应一个 8x8 亮度分量块, 每个 4x4 色度分量块选择相应 8x8 亮度分量块中左上角的 4x4 块的预测模式来映射。

上述技术方案中, 帧内亮度预测模式与帧内色度预测模式间一一对应的映射关系是在纹理分析和统计的基础上得到的。

上述技术方案中, 帧内色度的预测编码过程根据预测所选用的色度块大小的不同而不同, 包括:

a) 8x8 块预测, 对 8x8 色度块进行整体预测;

b) 4x4 块预测, 整个 8x8 色度分量块分成 4 个 4x4 色度分量块, 根据每个 4x4 色度分量块各自的色度预测模式实现色度预测编码。

所述的 4x4 块预测时, 4 个 4x4 色度分量块的预测并行进行, 在预测时当所需要的参考像素不可用时, 使用最近的可用像素点来代替。

本发明方法的优点在于:

(1) 利用色度分量和亮度分量的纹理相似性, 色度分量的空域预测模式通过相应亮度分量的空域预测模式按照一定关系进行映射, 不必使用率失真优化方法来挑选预测模式, 有效降低了色度分量帧内模式编码的复杂度。

(2) 由帧内亮度预测模式可以映射唯一的帧内色度预测模式, 只要在解码端采用同样的选择策略, 就不必对色度预测模式进行编码, 进一步降低了编解码的复杂度。

(3) 在利用色度块的预测模式, 实现对色度块像素值的预测过程中, 4x4 块预测使得 4 个 4x4 色度分量块的预测并行进行, 提高了对色度块像素值进行预测时的效率。

附图说明

图 1: 色度预测的待编码块及周围参考点示意图

图 2: H.264 帧内色度预测编码流程图

图 3: 本发明的帧内色度预测编码方法流程图

图 4: 8x8 色度预测模式选择的亮度分量的位置对应关系

图 5: 4x4 色度预测模式选择的亮度分量的位置对应关系

图 6: 亮度分量 16x16 预测模式到色度预测模式的映射关系表

图 7: 亮度分量 4x4 预测模式到色度预测模式的映射关系表

图 8: 色度 4x4 预测过程中, 参考像素点的位置关系

具体实施方式

下面参照附图和具体实施方式, 对本发明方法做详细说明。

如图 3 所示, 为本发明的一种帧内图像预测编码方法的操作流程图。该方法的操作步骤包括:

在步骤 10 中, 获取与色度分量相对应的亮度分量的预测模式。色度分量的空域预测在同一宏块的亮度分量的空域预测之后进行, 色度分量的预测模式通过同一宏块中相应位置块的亮度分量的预测模式进行映射。在现有的 H.264 帧内色度预测编码技术中, 色度分量预测所采用的块的大小固定, 色度分量统一按照 8x8 进行预测; 而在本发明方法中, 色度分量预测所采用的块的大小分为两类, 一类是 8x8, 另一类是 4x4。亮度分量映射位置的选择方法根据色度分量预测采用的块的大小分为两类:

1) 色度分量按照 8x8 进行预测: 宏块中的 8x8 色度分量块只有一种预测模式, 该预测模式通过当前宏块的 16x16 亮度分量块中的某一 4x4 块的预测模式来映射, 随机选定一 4x4 亮度分量块, 记录其亮度预测模式。在本实施例中, 参见图 4, 8x8 色度分量块选用同一宏块的 16x16 亮度分量块左上角的 4x4 块的预测模式来映射。之所以这样做, 是因为通过实验测得, 选用同一宏块的亮度分量左上角的块的预测模式来映射, 最终得到的效果最佳。

2) 色度分量按照 4x4 进行预测: 宏块中的 8x8 色度分量块分成 4 个 4x4 块, 这 4 个 4x4 块可以分别有自己的预测模式, 各个块的色度预测模式通过当前宏块的 16x16 亮度分量的某一 4x4 块的亮度预测模式来映射。本实施例中, 参见图 5, 将 16x16 亮度分量块等分成 4 个 8x8 亮度分量块, 每个 4x4 色度分量块对应一个 8x8 亮度分量块, 每个 4x4 色度分量块选择相应 8x8 亮度分量块中左上角的 4x4 块的预测模式来映射。这是由实验得出的最优的选择方法, 利用该方法选用亮度分量块, 最终得到的效果最佳。

在步骤 20 中, 利用亮度分量的预测模式实现对色度预测模式的映射。为色度块

选择了相应的亮度块的预测模式之后，接着需要将亮度块的预测模式映射到色度块的预测模式，建立亮度块的预测模式与色度块的预测模式之间的映射关系。因为亮度块的预测模式要多于色度块的预测模式，因此这种映射关系是多到一的。映射关系表建立在纹理分析和统计的基础上，映射关系表表明了对于某一种亮度块的预测模式而言，选用表中对应的色度块的预测模式，最终得到的预测效果最好。图6说明了由亮度16x16预测模式到色度预测模式的映射关系。图7说明了由亮度4x4预测模式到色度预测模式的映射关系。通过映射表就可以最终确定色度块的预测模式。

在现有的H.264帧内色度预测编码中，需要对4种色度块的预测模式都进行预测，再对四种预测模式所得到的预测结果使用率失真优化方法进行对比，从中选出最佳的预测模式。而在本发明方法中，只要根据亮度块的预测模式，参照映射关系表，就能够找到与亮度块相对应的同一宏块中的色度块的某一种预测模式，只需要对这一种预测模式进行编码，与现有技术相比，编码复杂度降低了75%，本发明方法与现有的H.264帧内色度预测编码方法相比明显降低了编码的复杂度。

在步骤30中，由已确定的色度块的预测模式，实现对色度块像素值的预测。在步骤10中，根据色度分量预测所选用的块大小的不同，亮度分量映射位置的选择也会有所不同，从而造成8x8色度分量块中可能有一种或几种色度预测模式。因此色度块像素值的预测过程根据预测所选用的色度块的大小分为两类。

1) 对于8x8块预测，共有3种模式的预测，分别为水平预测、垂直预测和DC预测。与现有的H.264帧内色度预测编码方法相比，减少了Plane预测模式，这是因为在本发明方法所适用的场景中，对图像的精度要求不是很高，利用水平预测、垂直预测和DC预测这3种预测模式足以达到相关的精度要求，而Plane预测模式与前面3种预测模式相比，预测方法复杂，编码复杂度也相对更高，因此在本发明方法中只选用了水平预测、垂直预测和DC预测。利用这3种预测模式实现对色度块像素值预测的具体操作与H.264中的这3种预测模式的操作步骤相同。也就是如图1所示的那样，水平预测仅使用当前块左边的参考像素点(i、j、k、l、m、n、o、p)进行预测；垂直预测仅使用当前块上边的参考像素点(a、b、c、d、e、f、g、h)进行预测；DC预测根据参考像素的可用性情况决定使用左边或上边或同时使用左边、上边和左上角(x)的参考像素点进行预测。

2) 对于4x4块预测，也有3种模式的预测，分别为水平预测、垂直预测和DC预测。与8x8块预测相比，其不同之处在于：在8x8块预测中，一个宏块的整个8x8

色度分量块只有一种色度预测模式，而在 4x4 块预测中，一个宏块的整个 8x8 色度分量块分成 4 个 4x4 色度分量块，每个 4x4 色度分量块有各自的色度预测模式，各个色度分量块之间的色度预测模式可以相同，也可以不同。在实现对色度块像素值的预测过程中，4 个 4x4 色度分量块可以同时预测，与对 8x8 色度分量块进行整体色度块像素值预测相比，4x4 块预测的效率显然更高。

4x4 色度分量块预测时当所需要的参考像素不可用时，使用最近的可用像素点来代替。4x4 块预测的具体实现过程参照图 8，对“1”、“2”、“3”、“4”块分别按照各自的预测模式进行预测，若其中“2”块在做预测时需要左边像素，但“2”块的左边为“1”块，所需要的参考像素不可用，因此使用最近的可用像素点来代替，即使用“1”块的左边像素；同样的，若其中“4”块在做预测时需要左边像素，则使用“3”块的左边像素；若其中“3”块在做预测时需要上边像素，则使用“1”块上边像素；若其中“4”块在做预测时需要上边像素，则使用“2”块的上边像素。这些块在做预测需要其他方向的像素时，与现有技术相同。利用上面所述的方法，“1”、“2”、“3”、“4”块的预测可以并行进行。

在现有的 H.264 帧内色度预测编码方法中，只能够实现 8x8 块预测，不能够实现 4x4 块预测。因此本发明方法与现有技术相比，其预测的效率显然得以提高，从而也能够提高编码的效率。

在步骤 40 中，计算预测残差。将上一步所得到预测块的像素值与待编码块的像素值做差，所得到的结果就是残差。

在步骤 50 中，对残差做 DCT 变换和量化。

在步骤 60 中，对残差系数做熵编码。

在步骤 70 中，将熵编码的结果按照码流格式要求输出。

以上步骤 40—70 中，所执行的各步操作在现有的 H.264 帧内色度预测编码方法中都有类似的实现，而且其中具体的实现技术也都是现有技术，因此，不再详细描述。

利用本实施例的方法，由于亮度分量的预测模式可以映射一种唯一的色度预测模式，因此在解码端采用同样的选择策略后，只要在编码端对亮度分量的预测模式进行编码，在解码端对亮度分量的预测模式进行解码，就可以直接对应相应的色度分量的预测模式，而不必对色度分量的空域预测模式进行编解码操作，从而进一步降低了编解码的复杂度。

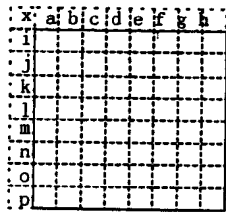


图 1

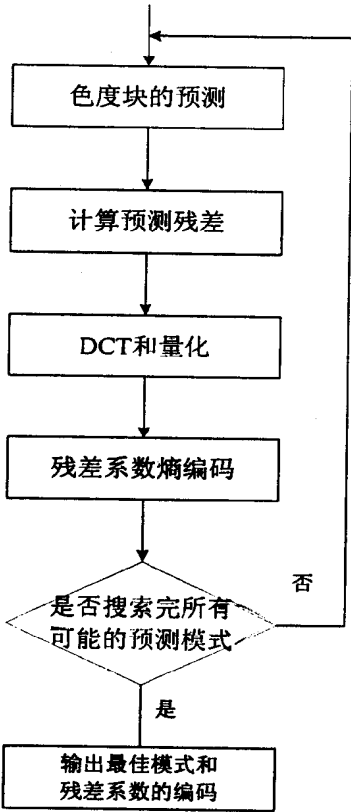


图 2

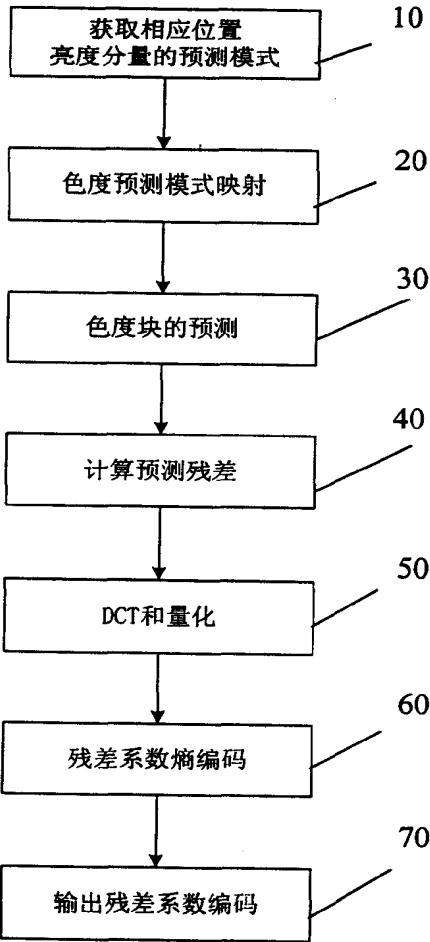


图 3

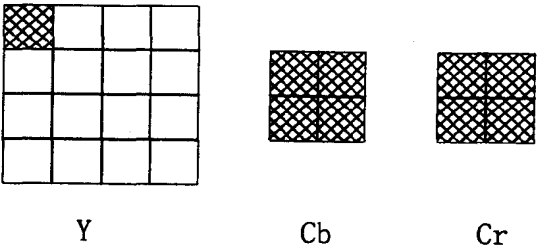


图 4

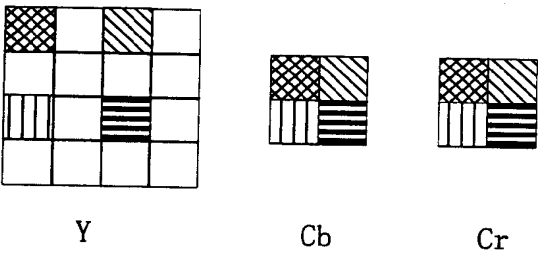


图 5

亮度预测模式	色度预测模式
VERT_16x16_PRED	VERT_PRED_CHROMA
HOR_16x16_PRED	HOR_PRED_CHROMA
DC_16x16_PRED	DC_PRED_CHROMA
PLANE_16x16_PRED	DC_PRED_CHROMA

图 6

亮度预测模式	色度预测模式
VERT_4x4_PRED	VERT_PRED_CHROMA
HOR_4x4_PRED	HOR_PRED_CHROMA
DC_4x4_PRED	DC_PRED_CHROMA
DOWN_LEFT_4x4_PRED	DC_PRED_CHROMA
DOWN_RIGHT_4x4_PRED	DC_PRED_CHROMA
VERT_RIGHT_4x4_PRED	DC_PRED_CHROMA
HOR_DOWN_4x4_PRED	DC_PRED_CHROMA
VER_LEFT_4x4_PRED	DC_PRED_CHROMA
HOR_UP_4x4_PRED	DC_PRED_CHROMA

图 7

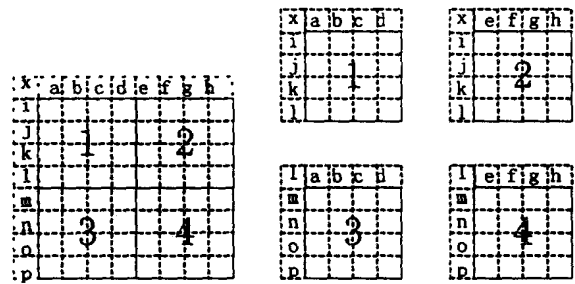


图 8



Publication number: CN1784015 A

Publication country: CHINA

Publication type: UNEXAMINED APPLIC. OPEN TO PUBLIC INSPECTION

Publication date: 20060607

Application number: CN200410096605

Application date: 20041202

Priority: CN20041009660520041202CN2004196605;

Assignee: COMPUTING TECHNOLOGY INST ;

Assignee^{std}: COMPUTING TECHNOLOGY INST CAS ;

Inventor: DONGMING SHEN ZHANG ;

Inventor^{std}: ZHANG DONGMING SHEN ;

International class⁸: H04N19/593;

Family members: [CN100461867 C](#) [CN1784015 A](#)

Title: INAGE PREDICTING ENCODING METHOD IN FRAME

Title: 一种帧内图像预测编码方法

Abstract:

This invention discloses an intraframe image predicted coding method, which contains intraframe brightness predicted coding and chromaticity coding, obtaining a chromaticity predicted mode according to brightness predicted coding in chromaticity predicted coding. Said invention has advantages of effectively reducing complexity of intraframe coding of chromatic component without using distortion rate optimization method to select prediction mode, raising the efficiency to predict chromaticity block pixel value.

Abstract:

本发明公开了一种帧内图像预测编码方法，包括对所述帧内图像进行亮度预测编码和色度预测编码，在进行色度预测编码时，根据亮度预测编码的预测模式对应得到一个色度预测模式，用该色度预测模式对所述帧内图像进行色度预测编码。本发明方法的优点：不必使用率失真优化方法来挑选预测模式，有效降低了色度分量帧

The present invention discloses an intra-image predictive encoding method, comprising the intra picture prediction coding luminance and chrominance predictive coding, predictive coding during chroma, the brightness of the predictive coding modes corresponding to obtaining a chrominance prediction

内模式编码的复杂度；由帧内亮度预测模式可以映射一种唯一的帧内色度预测模式，只要在解码端采用同样的选择策略，就不必对色度分量的空域预测模式进行编码，进一步降低编解码的复杂度；在利用色度预测模式，实现对色度块像素值的预测过程中，4x4块预测使得4个4x4色度分量块的预测并行进行，提高了对色度块像素值进行预测时的效率。

mode, with the chroma prediction mode of the color image frame predictive coding. Advantages of the present invention is a method: do not rate-distortion optimization method to select a prediction mode, effectively reducing the complexity of a chrominance component of an intra mode coding; luminance by the intra-prediction mode can map a unique intra-chrominance prediction mode, As long as the decoding side uses the same selection strategy, it is unnecessary for the spatial prediction mode to encode the chrominance components, further reducing the complexity of the codec; using chroma prediction mode, to achieve the block pixel values of chrominance prediction process, 4x4 block prediction makes four 4x4 block prediction chroma component parallel to improve the efficiency of chroma block prediction of pixel values.

Claims:

1、一种帧内图像预测编码方法，包括对所述帧内图像进行亮度预测编码和色度预测编码，其特征在于，在进行色度预测编码时，根据亮度预测编码的预测模式对应得到一个色度预测模式，用该色度预测模式对所述帧内图像进行色度预测编码。

1, an intra picture prediction coding method, comprising the intra picture prediction coding luminance and chrominance predictive coding, characterized in that, during chroma prediction encoding, the luminance prediction encoding the prediction mode correspondence obtaining a chrominance prediction mode, chroma prediction mode with the image of the intra-chrominance prediction encoding.

2、根据权利要求1所述的帧内图像预测编码方法，其特征在于，所述亮度预测模式包括：VERT_16x16_PRED、HOR_16x16_PRED、DC_16x16_PRED、PLANE_16x16_PRED、VERT_4x4_PRED、HOR_4x4_PRED、DOWN_LEFT_4x4_PRED、DOWN_RIGHT_4x4_PRED、VERT_RIGHT_4x4_PRED、DC_4x4_PRED、HOR_DOWN_4x4_PRED、VER_LEFT_4x4_PRED、HOR_UP_4x4_PRED；所述色度预测模式包括：VERT_PRED_CHROMA、HOR_PRED_CHROMA、DC_PRED_CHROMA；3、根据权利要求2所述的帧内图像预测编码方法，其特征

2, the intra picture prediction coding method according to claim 1, characterized in that, the brightness prediction mode includes: VERT_16x16_PRED, HOR_16x16_PRED, DC_16x16_PRED, PLANE_16x16_PRED, VERT_4x4_PRED, HOR_4x4_PRED, DOWN_LEFT_4x4_PRED, DOWN_RIGHT_4x4_PRED and VERT_RIGHT_4x4_PRED, DC_4x4_PRED, HOR_DOWN_4x4_PRED,

在于，亮度预测模式与色度预测模式间的对应关系为：VERT_16x16_PRED 对应 VERT_PRED_CHROMA；HOR_16x16_PRED 对应 HOR_PRED_CHROMA；DC_16x16_PRED 对应 DC_PRED_CHROMA；PLANE_16x16_PRED 对应 DC_PRED_CHROMA；VERT_4x4_PRED 对应 VERT_PRED_CHROMA；HOR_4x4_PRED 对应 HOR_PRED_CHROMA；DOWN_LEFT_4x4_PRED 对应 DC_PRED_CHROMA；DOWN_RIGHT_4x4_PRED 对应 DC_PRED_CHROMA；VERT_RIGHT_4x4_PRED 对应 DC_PRED_CHROMA；DC_4x4_PRED 对应 DC_PRED_CHROMA；HOR_DOWN_4x4_PRED 对应 DC_PRED_CHROMA；VER_LEFT_4x4_PRED 对应 DC_PRED_CHROMA；HOR_UP_4x4_PRED 对应 DC_PRED_CHROMA。

4、根据权利要求1所述的帧内图像预测编码方法，其特征在于，所述的帧内亮度预测模式由所选择的亮度分量块决定，而亮度分量块的选择方法根据色度分量预测时所采用的块的大小分为两类：a)当色度分量按照8x8进行预测时，宏块中的8x8色度分量块只有一种预测模式，该预测模式通过当前宏块的16x16亮度分量块中的某一4x4亮度分量块的预测

VER_LEFT_4x4_PRED, HOR_UP_4x4_PRED; the chroma prediction mode includes: VERT_PRED_CHROMA, HOR_PRED_CHROMA and gt; DC_PRED_CHROMA0 3, according to claim intra picture prediction coding method according to claim 2, characterized in that the correspondence between the luminance and chrominance prediction mode is the prediction mode among: VERT_16x16_PRED correspond VERT_PRED_CHROMA; HOR_16x16_PRED corresponds HOR_PRED-CHROMA; DC_16x16_PRED corresponds DC_PRED-CHROMA; PLANE_16x16_PRED correspondence DC_PRED_CHROMA; VERT_4x4_PRED correspond VERT_PRED_CHROMA; HOR_4x4_PRED correspond HOR_PRED_CHROMA; DOWN_LEFT_4x4_PRED correspond DC_PRED_CHROMA; DOWN_RIGHT_4x4_PRED correspond DC_PRED_CHROMA; VERT_RIGHT_4x4_PRED correspond DC_PRED_CHROMA; DC_4x4_PRED correspond DC_PRED_CHROMA; HOR_DOWN_4x4_PRED correspond DC_PRED-CHROMA; VER_LEFT_4x4_PRED correspond DC_PRED_CHROMA; HOR_UP_4x4_PRED correspond DC_PRED_CHROMA.

4, an intra picture prediction coding method according to claim 1, characterized in that said luminance intra-prediction mode of the luminance component is determined by the selected block, the selected block of the luminance component the method according to preclude the prediction when the chrominance component block size used fall into two categories: a) when the chrominance component in accordance with the 8x8 prediction macro

block 8x8 chrominance component block only one prediction mode, the prediction mode by prediction of the current macroblock 16x16 luminance component blocks in the luminance component of a 4x4 block

Description:

一种帧内图像预测编码方法 技术领域 本发明属于视频及图像编解码领域，特别涉及视频帧内图像基于空域预测的编 解码技术。

背景技术 视频编码是计算机科学领域的热点问题，它旨在发现数字视频的最有效的编码 格式。在视频编码中，常把视频图像分为帧内图像和帧间图像两种。其中，帧内图 像的视频编码仅使用图像自身提供的信息进行空间上的预测，它的编解码不依赖于 其他图像，可以单独解码；而帧间图 像的编码要消除的是视频在时间域上的冗余信 息，它的编解码依赖于它之前的一个或多个图像。对帧内图像进行视频编码 时，因 为在帧内图像中，相邻图像块的像素具有很大的空域相关性，所以需要采用 空域预 测编码方法以提高编码效率。空域 预测编码方法的主要思想是以相邻块为参 考像素，按照不同的方向预测待编码块的 像素值，把预测结果称为预测块，其中的 待编码块 是指没有进行压缩编码的原始数 据块；将待编码块的像素值与预测块的像 素值作差，所得到的结果称为残差；再对 残差进行DCT变换和量化得到残差系数， 最后对结果 进行无损的熵编码并输出。待 编码块和预测块匹配程度越高，编码的效 率越高。为 了提高预测块像素值与当前块 像素值的匹配程度，往往设置多种预测模 式。

现有的H.264 (又称MPEG-4Part10或AVC) 在对帧内图像预测中，是以宏块 为单位进 行编码的，所谓宏块包括一个16x16块的 亮度分量和相应位置的两个8x8 块的表示 色度的分量。因此在对帧内图像进行预测 编码时，包括亮度预测编码和色 度预测编 码。亮度预测编码是依据一定的亮度预测 模式对宏块的亮度分量进行编码，而色度

One kind of intra-image predictive coding method of the present invention belongs to the technical field of image and video codecs, and particularly to a video frame image based on spatial prediction codec technology.

Background Video Coding is a hot topic in the field of computer science, it is designed to find the most efficient digital video encoding format. In video encoding, the video image is often divided into two kinds of intra and inter frame. Among them, the intra picture video coding using only the image itself provides information to predict space, it does not depend on the codec other images can be decoded alone; and inter frame coding is to be eliminated in the time domain video redundant information, it depends on the codec before it one or more images. When encoding video frame image, because the image in the frame, the pixel adjacent image blocks of great spatial correlation, it is necessary to preclude the use of airspace predictive coding method to improve coding efficiency. The main idea of the method of encoding spatial prediction reference pixel based on the neighboring blocks, prediction directions according to different values of pixels of the block to be coded, the prediction called prediction block, wherein the block to be coded is no means for compression-encoding the original data block; the pixel value to be encoded as the difference between the pixel block and the predicted block, the results obtained are called residuals; and

预测编码是依据一定的色度预测模式对宏块的色度分量进行编码。在现有技术中，亮度预测编码和色度预测编码是相对独立的，亮度预测模式的选择与色度预测模式的选择也是相对独立的。在H.264中，亮度预测模式分为16x16和4x4两类，其中16x16的预测模式有4种，4x4的预测模式有9种；16x16的4种预测模式具体包括：VERT_16x16_PRED、HOR_16x16_PRED、DC_16x16_PRED、PLANE_16x16_PRED；4x4的9种预测模式具体包括：VERT_4x4_PRED、HOR_4x4_PRED、DOWN_LEFT_4x4_PRED、DOWN_RIGHT_4x4_PRED、VERT_RIGHT_4x4_PRED、DC_4x4_PRED、HOR_DOWN_4x4_PRED、VER—LEFT—4x4—PRED、HOR_UP_4x4_PRED。

在H.264中，色度预测模式包括水平（VERT_PRED_CHROMA）、垂直（HOR—PRED_CHROMA）、DC（DC_PRED_CHROMA）和Plane预测4种。这4种色度预测模式的具体实现如图1所示，水平预测仅使用当前块左边的参考像素点（i、j、k、l、m、n、o、p）进行预测；垂直预测仅使用当前块上边的参考像素点（a、b、c、d、e、f、g、h）进行预测；DC预测根据参考像素的可用性情况决定使用左边或上边或同时使用左边、上边和左上角（X）的参考像素点；Plane预测需要使用左边和上边的参考像素。因此，当仅有左边参考像素可用时，只有水平预测和DC预测可供选择；当仅有上边参考像素可用时，只有垂直预测和DC预测可供选择；当左边和上边的参考像素均可用时，四种模式都可以选择；当左边和上边的参考像素均不可用时，即为左上角宏块，只有DC预测模式可以选择。除了最后一种情况，具体采用哪一种预测模式需要根据按照一定的规则来选择。

上述4种色度预测模式都有可能应用于实际，在做帧内色度预测编码时可能有多种色度预测模式可用，对色度预测模式的选择通常采用的方式是率失真（RD）优化规则，其中，R是指码率开销，具体表示使用一种预测模式时编码模式和残差系数所需要的比特数，D是指失真度，具体表示

then the residual is transformed and quantized DCT coefficient residuals, and finally the results of entropy coding and lossless output. The higher the block to be coded and the predictive block matching degree, the higher the coding efficiency. In order to improve the degree of matching the pixel values of the prediction block of the current block of pixel values, prediction often set various modes.

Existing H.264 (also known as MPEG-4Part10 or AVC) in the intra-picture prediction, the macroblock is encoded as a unit, the so-called macro block includes a 16x16 block luminance component and a corresponding position two 8x8 chrominance blocks represent components. Therefore, when intra-picture predictive coding including predictive coding luminance and chrominance predictive coding. The luma prediction encoding is based on certain prediction mode of the luminance component luminance macroblock is encoded chrominance predictive coding is based on certain prediction mode chroma chroma component of the macroblock is coded. In the prior art, the luminance and chrominance predictive coding predictive coding is relatively independent choices and selecting chroma prediction mode brightness prediction mode is relatively independent. In H.264, the luminance prediction modes are divided into two types of 16x16 and 4x4, 16x16 prediction mode in which there are four, 4x4 prediction mode there are nine; 16x16 four specific prediction modes include: VERT_16x16_PRED, HOR_16x16_PRED, DC_16x16_PRED, PLANE_16x16_PRED; 4x4 nine prediction modes specifically includes: VERT_4x4_PRED, HOR_4x4_PRED, DOWN_LEFT_4x4_PRED, DOWN_RIGHT_4x4_PRED, VERT_RIGHT_4x4_PRED, DC_4x4_PRED,

使用一种预测模式所重建的宏块与原始图像宏块之间的差值。R和D成反比关系，当D较小时，R就会增大；当D增加时，相应的R就会降低。所谓率失真优化，就是在编码所需的码率开销（比特数）和编码引起的图像失真二者之间找到最佳平衡点，使得码率开销R和失真度D同时尽可能地小。对色度分量而言，进行预测模式的率失真优化选择，就需要进行所有4种模式的预测，并对残差进行编码，记录每种预测模式的码率开销与失真之和，从中选择最小的那种模式作为最优预测模式。

参看图2,现有的H.264帧内色度预测编码方法的操作过程包括：1、利用某一种色度预测模式对色度块进行预测；2、计算预测残差；3、对残差进行DCT变换和量化处理，得到残差系数；4、对残差系数做熵编码；5、判断是否做完了所有可能的预测模式，若还有可能的预测模式，则跳转到步骤1，若不再有新的可能预测模式，则继续以下操作；6、对所有结果做率失真优化选择，输出最佳模式和残差系数的编码。

在步骤2中的残差实际上是一个可以用矩阵表示的残差块；在步骤3中，经过DCT变换和量化处理，残差块由二维矩阵转变为一维数组，这个一维数组就是残差系数。其中的DCT变换是指对残差进行从空域到频域的转变，目的是进一步压缩数据，量化的实质也是进行压缩，将0附近的，比如（-3，3）之间系数全部视作0；将绝对值小于10的系数全部视作+2/-2。

现有的H.264帧内色度预测编码已经完全可以实现对帧内图像的色度预测编码，但对H.264帧内色度预测编码过程进行仔细分析、观察后：发现存在以下一些问题：(1)增加预测模式，可以提高预测精度，但是当预测模式包含了水平、垂直、DC预测之后，再增加更多的预测模式，不会对预测精度产生明显影响。

(2)当预测模式增加时，用于预测模式编码的比特数就会增加，这就会带来更大的码流开销。

HOR_DOWN_4x4_PRED, VER-LEFT a 4x4 a PRED, HOR_UP_4x4_PRED.

In H.264, chroma prediction modes include horizontal (VERT_PRED_CHROMA), vertical (HOR_PRED_CHROMA), DC (DC_PRED_CHROMA) and Plane prediction 4 species. Specifically these four chroma prediction modes of realization shown in Figure 1, horizontal prediction using prediction only with reference to the left of the current block of pixels (i, j, k, l, m, n, o, p); vertical prediction only using the current reference pixel block upper (a, b, c, d, e, f, g, h) prediction; DC prediction reference pixel based on the availability of decisions using the left or upper left or simultaneously, and the upper left corner (X) of the reference pixel; Plane prediction requires the use of the left and upper reference pixels. Thus, when only the left side of the reference pixel is available, only the level of prediction and DC prediction to choose from; when only the upper reference pixels are available, only the vertical prediction and DC prediction to choose from; when the left and upper reference pixels are available at, four modes can be selected; when the left and upper reference pixels are not available, the upper left corner is the macro block, only DC prediction mode can be selected. Except for the last case, the decision of what mode using prediction need to be selected according to certain rules. above 4 chroma prediction modes are likely to be applied to the actual doing chroma intra prediction coding may have a variety of chroma prediction modes are available on the chroma prediction mode selection approach it is commonly used rate-distortion (RD) optimization rules, wherein, R is a bit rate refers to overhead, particularly when the number of bits using a coding mode and a residual prediction mode required

(3)随着预测模式的增加,进行预测模式率失真优化选择的过程就会更加复杂,导致视频编码的复杂度增加。

同时,在视频编码过程中还存在以下规律和现象:(1)统计规律显示由于亮度分量和色度分量的纹理具有相似性,因此宏块的亮度分量和色度分量的预测模式呈现较强的相关性。而H.264的帧内预测编码中,亮度和色度的预测模式是独立选择的,它们的预测模式分别通过率失真优化方法来选择,也就是忽视了亮度分量和色度分量的预测模式间的相关性。

(2)在现有的H.264视频编码框架中,色度分量的PSNR(峰值信噪比)值平均高出亮度分量2dB左右。因此,色度分量的PSNR即使有一定的降低,也不是人的视觉可以感觉到的,并不会引起主观视觉效果的变化。也就是说在适当范围内牺牲色度分量的精度以换取编码效率的提高和码率开销的降低是切实可行的。

本发明正是根据上述规律,提出一种帧内图像预测编码方法,以克服现有的H.264视频编码技术所存在的码流开销大,编码复杂度高的缺陷。

发明内容 本发明的目的是克服现有的H.264帧内色度预测编码方法所存在的码流开销大,视频编码复杂度高的缺陷,提供一种低复杂度帧内色度预测编码方法,并且保证主观和客观的视频质量不会明显降低。

为了实现上述目的,本发明提供了一种帧内图像预测编码方法,包括帧内亮度预测编码方法和帧内色度预测编码方法,在帧内色度预测编码方法中,利用帧内亮度预测模式得到唯一的帧内色度预测模式,实现帧内色度的预测编码。

上述技术方案中,帧内亮度预测模式包括:VERT_16x16_PRED、HOR_16x16_PRED、DC_16x16_PRED、PLANE_16x16_PRED、VERT_4x4_PRED、HOR_4x4_PRED、DOWN_LEFT_4x4_PRED、DOWN_RIGHT_4x4_PRED、VERT_RIGHT_4x4_PRED、

coefficients, D refers to the degree of distortion, that the use of a specific prediction mode reconstruction the difference between the macro block and the original image macro blocks. R and D inverse relationship, when there is less when D, R will increase; when D increases, the corresponding R is reduced. The so-called rate-distortion optimization is required between the coding overhead rate (the number of bits) and both the image distortion caused by coding to find the best balance point, so that the overhead rate R and distortion D at the same time as small as possible. Chroma components, the conduct Prediction distortion optimization mode selection, on the need for all the four patterns of prediction, and residual is encoded, recording rate and the cost for each prediction mode of distortion and, choose the smallest the kind of pattern as the optimum prediction mode.

2, the conventional H.264 procedure intra chrominance prediction coding method comprising: 1, the use of a certain kind of chroma prediction mode chroma prediction block; 2, calculating prediction residuals; 3, the residual DCT transform and quantization to obtain residual coefficients; 4, the residual entropy coding coefficients do; 5, it is determined whether or not done all the possible prediction mode, if there is a possible prediction modes, jump to step 1, if no new prediction model may then proceed as follows; 6, all the results do rate-distortion optimization coding options, the best mode and output residual coefficients. residuals in step 2 is actually a residual block matrix can be expressed; in step 3, after the DCT transform and quantization, residual block by the two-dimensional matrix into a one-dimensional array, this one dimensional array is the residual coefficients. Wherein the DCT transform means transform residuals from the spatial domain to the

DC_4x4_PRED、
 HOR_DOWN_4x4_PRED、
 VER_LEFT_4x4_PRED、
 HOR_UP_4x4_PRED；帧内色度预测模式包括：
 VERT_PRED_CHROMA、
 HOR_PRED_CHROMA、
 DC_PRED_CHROMA。

上述技术方案中，帧内亮度预测模式和帧内色度预测模式间的对应关系为：

VERT_16x16_PRE 亮度预测模式对应 VERT_PRED_CHROMA；
 HOR_16x16_PRED 对应 HOR_PRED_CHROMA；DC_16x16_PRED 对应 DC_PRED_CHROMA；
 PLANE_16x16_PRED 对应 DC_PRED_CHROMA；VERT_4x4_PRED 对应 VERT_PRED_CHROMA；
 HOR_4x4_PRED 对应 HOR_PRED_CHROMA；DOWN_LEFT_4x4_PRED 对应 DC_PRED_CHROMA；
 DOWN_RIGHT_4x4_PRED 对应 DC_PRED_CHROMA；
 VERT_RIGHT_4x4_PRED 对应 DC_PRED_CHROMA；DC_4x4_PRED 对应 DC_PRED_CHROMA；
 HOR_DOWN_4x4_PRED 对应 DC_PRED_CHROMA；
 VER_LEFT_4x4_PRED 对应 DC_PRED_CHROMA；
 HOR_UP_4x4_PRED 对应 DC_PRED_CHROMA。

上述技术方案中，所述的帧内亮度预测模式由所选择的亮度分量块决定，而亮度分量块的选择方法根据色度分量预测时所采用的块的大小分为两类：a)当色度分量按照8x8进行预测时，宏块中的8x8色度分量块只有一种预测模式，该预测模式通过当前宏块的16x16亮度分量块中的某一4x4亮度分量块的预测模式来映射；b)当色度分量按照4x4进行预测时，宏块中的8x8色度分量块分成4个4x4块，这4个4x4块分别有自己的预测模式，每个块的预测模式通过当前宏块的16x16亮度分量的某一4x4块的预测模式来映射。

frequency domain, the purpose is to further compress the data, to quantify the substance is compressed to, for example, the coefficient (3,3) between all treated as close to 0 0; coefficient of absolute value less than 10 is all considered 2 / -2.

Existing H.264 Intra chroma prediction coding already can achieve intra chrominance image predictive coding, but H.264 Intra chroma prediction coding process of careful analysis, were observed: Discovery the following questions: (1) increase in prediction mode, the prediction accuracy can be improved, but when the prediction mode includes horizontal, vertical, DC prediction after prediction mode to add more, will not have a significant effect on prediction accuracy.

(2) when the prediction mode increases, the prediction mode for coding will increase the number of bits, which will bring greater overhead stream.

(3) With the increase of the prediction mode, the prediction mode selection rate-distortion optimization process will be more complicated, leading to an increase in the complexity of the video encoder.

At the same time, there are the following rules and phenomena in the video encoding process: (1) Due to the statistical laws show the texture of the luminance and chrominance components have similarities, the luminance and chrominance components of the macroblock prediction mode of presentation a strong correlation. The intra prediction encoding of H.264, the luminance and chrominance prediction mode is independently selected, respectively, by their prediction mode rate distortion optimization method to choose, that is, ignoring the luminance and chrominance components inter-prediction mode relevance.

所述的亮度分量块的选择方法a)中, 4x4亮度分量块的一个优选选择方案是选用16x16亮度分量块中左上角的4x4块。

所述的亮度分量块的选择方法b)中, 4x4亮度分量块的一个优选选择方案是将当前宏块的16x16亮度分量块等分成4个8x8亮度分量块, 每个4x4色度分量块对应一个8x8亮度分量块, 每个4x4色度分量块选择相应8x8亮度分量块中左上角的4x4块的预测模式来映射。

上述技术方案中, 帧内亮度预测模式与帧内色度预测模式间一一对应的映射关系是在纹理分析和统计的基础上得到的。

上述技术方案中, 帧内色度的预测编码过程根据预测所选用的色度块大小的不同而不同, 包括: a) 8x8块预测, 对8x8色度块进行整体预测; b) 4x4块预测, 整个8x8色度分量块分成4个4x4色度分量块, 根据每个4x4色度分量块各自的色度预测模式实现色度预测编码。

所述的4x4块预测时, 4个4x4色度分量块的预测并行进行, 在预测时当所需要的参考像素不可用时, 使用最近的可用像素点来代替。

本发明方法的优点在于: (1)利用色度分量和亮度分量的纹理相似性, 色度分量的空域预测模式通过相应亮度分量的空域预测模式按照一定关系进行映射, 不必使用率失真优化方法来挑选预测模式, 有效降低了色度分量帧内模式编码的复杂度。

(2)由帧内亮度预测模式可以映射唯一的帧内色度预测模式, 只要在解码端采用同样的选择策略, 就不必对色度预测模式进行编码, 进一步降低了编解码的复杂度。

(3)在利用色度块的预测模式, 实现对色度块像素值的预测过程中, 4x4块预测使得4个4x4色度分量块的预测并行进行, 提高了对色度块像素值进行预测时的效率。

附图说明 图1:色度预测的待编码块及周围参考点示意图 图2: H.264帧内色度预测编码流程图 图3:本发明的帧内色度预测编码方法流程图 图4: 8x8色度预测模式选择的

(2) in an existing H.264 video coding framework, PSNR chrominance component (PSNR) higher average value of the luminance component is about 2dB. Therefore, PSNR chrominance components even if there is a certain decrease, nor can the human visual sense, and does not cause changes in the subjective visual effect. That is sacrificing accuracy chroma component in an appropriate range to improve the exchange rate and reduce the cost of coding efficiency is feasible.

The present invention is based on the law, proposed intra picture prediction coding method, in order to overcome the existing H.264 video coding techniques exist overhead stream, a high coding complexity defects. object of the present invention is to overcome the existing H.264 Intra chroma prediction encoding method exists stream overhead, high complexity video coding defects, provide a low degree of heterozygosity cold frame chrominance predictive coding method, and to ensure subjective and objective video quality will not be significantly reduced.

To achieve the above object, the present invention provides a method of intra picture prediction coding, predictive coding method including intra luminance and chrominance intra-prediction coding method, in the intra-chrominance prediction coding method using intraframe brightness prediction model has been the only intra chrominance prediction mode, intra prediction encoding to achieve chromaticity. the technical proposal, the luminance intra prediction mode includes: VERT_16x16_PRED, 6x16 a PRED, HOR_16x16_PRED, DC ", PLANE a 16x16 a PRED, VERT_4x4_PRED, HOR_4x4_PRED, DOWN-LEFT a 4x4_PRED, DOWN_RIGHT_4x4_PRED, VERT_RIGHT_4x4_PRED,

亮度分量的位置对应关系 图5: 4x4色度预测模式选择的亮度分量的位置对应关系 图6:亮度分量16x16预测模式到色度预测模式的映射关系表 图7:亮度分量4x4预测模式到色度预测模式的映射关系表 图8:色度4x4预测过程中, 参考像素点的位置关系 具体实施方式 下面参照附图和具体实施方式, 对本发明方法做详细说明。

如图3所示, 为本发明的一种帧内图像预测编码方法的操作流程图。该方法的操作步骤包括: 在步骤10中, 获取与色度分量相对应的亮度分量的预测模式。色度分量的空域预测在同一宏块的亮度分量的空域预测之后进行, 色度分量的预测模式通过同一宏块中相应位置块的亮度分量的预测模式进行映射。在现有的H.264帧内色度预测编码技术中, 色度分量预测所采用的块的大小固定, 色度分量统一按照8x8进行预测; 而在本发明方法中, 色度分量预测所采用的块的大小分为两类, 一类是8x8, 另一类是4x4。亮度分量映射位置的选择方法根据色度分量预测采用的块的大小分为两类: 1)色度分量按照8x8进行预测: 宏块中的8x8色度分量块只有一种预测模式, 该预测模式通过当前宏块的16x16亮度分量块中的某一4x4块的预测模式来映射, 随机选定一4x4亮度分量块, 记录其亮度预测模式。在本实施例中, 参见图4, 8x8色度分量块选用同一宏块的16x16亮度分量块左上角的4x4块的预测模式来映射。

之所以这样做, 是因为通过实验测得, 选用同一宏块的亮度分量左上角的块的预测模式来映射, 最终得到的效果最佳。

2)色度分量按照4x4进行预测: 宏块中的8x8色度分量块分成4个4x4块, 这4个4x4块可以分别有自己的预测模式, 各个块的色度预测模式通过当前宏块的16x16亮度分量的某一4x4块的亮度预测模式来映射。本实施例中, 参见图5, 将16x16亮度分量块等分成4个8x8亮度分量块, 每个4x4色度分量块对应一个8x8亮度分量块, 每个4x4色度分量块选择相应8x8亮度分量块中左上角的4x4块的预测模式来映射。这是由实验得出的最优的选择方法,

DC_4x4_PRED, HOR_DOWN_4x4_PRED, VER_LEFT_4x4_PRED, HOR_UP_4x4_PRED; intra chrominance prediction mode includes: VERT_PRED_CHROMA and gt; HOR_PRED_cHROMA, DC-PRED-cHROMA. the technical proposal, the correspondence between the luminance intra-prediction mode and intra chrominance prediction mode is between: VERT_16x16_PRE brightness prediction mode correspondence VERT_PRED_CHROMA; HOR_16x16_PRED correspond HOR_PRED_CHROMA; DC_16x16_PRED correspond DC_PRED_CHROMA; PLANE_16x16_PRED correspond DC_PRED_CHROMA; VERT_4x4_PRED correspondence VERT_PRED_CHROMA; HOR_4x4_PRED correspond HOR_PRED_CHROMA; DOWN_LEFT_4x4_PRED correspond DC_PRED_CHROMA; DOWN_RIGHT_4x4_PRED correspond DC_PRED_CHROMA; VERT_RIGHT_4x4_PRED correspond DC_PRED_CHROMA; DC_4x4_PRED correspond DC_PRED_CHROMA; HOR_DOWN_4x4_PRED correspond DC_PRED_CHROMA; VER_LEFT_4x4_PRED correspond DC_PRED_CHROMA; HOR_UP_4x4_PRED correspond DC_PRED_CHROMA. in the above-described aspect, the intra prediction mode is determined by the brightness of the luminance component of the selected block, and a block selection method according to the size of the luminance component when the chroma component prediction block used fall into two categories: a) when the chrominance component in accordance with 8x8 prediction macro block 8x8 chrominance component block only one prediction mode, the prediction mode by the prediction mode of the current

利用该方法选用亮度分量块，最终得到的效果最佳。

在步骤20中，利用亮度分量的预测模式实现对色度预测模式的映射。为色度块选择了相应的亮度块的预测模式之后，接着需要将亮度块的预测模式映射到色度块的预测模式，建立亮度块的预测模式与色度块的预测模式之间的映射关系。因为亮度块的预测模式要多于色度块的预测模式，因此这种映射关系是多到一的。映射关系表建立在纹理分析和统计的基础上，映射关系表表明了对于某一种亮度块的预测模式而言，选用表中对应的色度块的预测模式，最终得到的预测效果最好。图6说明了由亮度16x16预测模式到色度预测模式的映射关系。图7说明了由亮度4x4预测模式到色度预测模式的映射关系。通过映射表就可以最终确定色度块的预测模式。

在现有的H.264帧内色度预测编码中，需要对4种色度块的预测模式都进行预测，再对四种预测模式所得到的预测结果使用率失真优化方法进行对比，从中选出最佳的预测模式。而在本发明方法中，只要根据亮度块的预测模式，参照映射关系表，就能够找到与亮度块相对应的同一宏块中的色度块的某一种预测模式，只需要对这一种预测模式进行编码，与现有技术相比，编码复杂度降低了75%，本发明方法与现有的H.264帧内色度预测编码方法相比明显降低了编码的复杂度。

在步骤30中，由已确定的色度块的预测模式，实现对色度块像素值的预测。在步骤10中，根据色度分量预测所选用的块大小的不同，亮度分量映射位置的选择也会有所不同，从而造成8x8色度分量块中可能有一种或几种色度预测模式。因此色度块像素值的预测过程根据预测所选用的色度块的大小分为两类。

1)对于8x8块预测，共有3种模式的预测，分别为水平预测、垂直预测和DC预测。与现有的H.264帧内色度预测编码方法相比，减少了Plane预测模式，这是因为在本发明方法所适用的场景中，对图像的精度要求不是很高，利用水平预测、垂直预测和DC预测这3种预测模式足以达到相关的精度要求，而Plane预测模式与前面3种

macroblock 16x16 luminance component blocks in the luminance component of a 4x4 block to mapping; b) when the chrominance component in accordance with the 4x4 prediction macro block 8x8 chrominance component block into four 4x4 blocks, four 4x4 blocks each have their own prediction mode, a prediction mode for each block by the current macro 16x16 prediction mode of a luminance component of 4x4 blocks which mapped.

Selection Brightness

block of said component a), one preferred option 4x4 luminance component block is selected with a 16x16 luminance component block in the upper left corner of the 4x4 blocks.

Brightness

The selection of the components in the blocks of method b), a preferred option 4x4 luminance component block is 16x16 luminance component blocks of the current macro block is equally divided into four 8x8 luminance component blocks, each 4x4 chroma component corresponding to an 8x8 luminance component blocks, each 4x4 chroma component block select 8x8 luminance component block in the upper left corner of the 4x4 block prediction mode to map. the technical proposal, the inter-prediction mode and the intra luma chroma intra prediction mode mapped one to one relationship is based on texture analysis and statistics obtained. in the above-described aspect, the chroma intra prediction encoding procedure varies between the selected chroma prediction block size is different, comprising: a) 8x8 block prediction, 8x8 chrominance blocks of the overall prediction; b) 4x4 prediction block, the entire block is divided into 8x8 chroma component 4 4x4 chrominance component blocks, each 4x4 chroma component block according to the