

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2017-158183

(P2017-158183A)

(43) 公開日 平成29年9月7日(2017.9.7)

(51) Int.Cl.	F I	テーマコード (参考)
HO 4 N 1/41 (2006.01)	HO 4 N 1/41 B	5 C 1 5 9
HO 4 N 19/126 (2014.01)	HO 4 N 19/126	5 C 1 7 8
HO 4 N 19/136 (2014.01)	HO 4 N 19/136	5 J 0 6 4
HO 4 N 19/94 (2014.01)	HO 4 N 19/94	
HO 3 M 7/30 (2006.01)	HO 3 M 7/30 B	
審査請求 未請求 請求項の数 7 O L (全 35 頁)		

(21) 出願番号 特願2017-34786 (P2017-34786)
 (22) 出願日 平成29年2月27日 (2017.2.27)
 (31) 優先権主張番号 特願2016-36602 (P2016-36602)
 (32) 優先日 平成28年2月29日 (2016.2.29)
 (33) 優先権主張国 日本国 (JP)

(71) 出願人 504202472
 大学共同利用機関法人情報・システム研究
 機構
 東京都立川市緑町10番3号
 (71) 出願人 501008945
 国立清華大学
 National Tsing Hua
 University
 台湾新竹市光復路二段101號
 No. 101, Section 2, Ku
 ang-Fu Road, Hsinchu
 , 30013, Taiwan

最終頁に続く

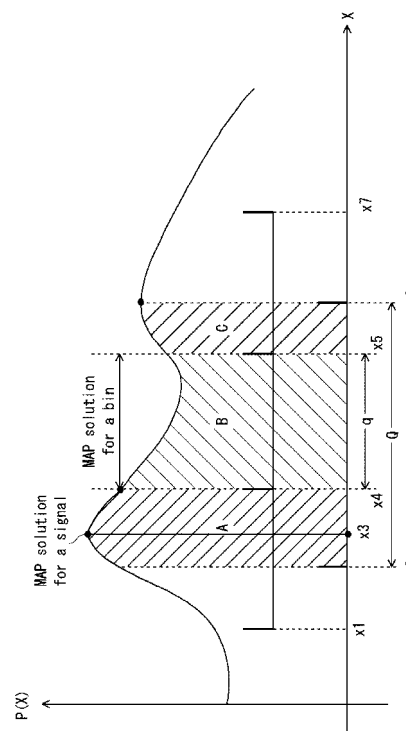
(54) 【発明の名称】 画像処理装置

(57) 【要約】

【課題】 高い圧縮率を実現しつつ、復号時にユーザにとって許容できる画質を保証すること。

【解決手段】 本発明にかかるクラウドレット20は、入力画像P1を第1の量子化幅(q)に基づき量子化した第1の量子化係数(a)から当該第1の量子化幅より広い第2の量子化幅(Q)に基づき量子化した第2の量子化係数(b)について、第1の量子化係数への逆写像を行う粗密変換部22を備える。粗密変換部22は、第2の量子化幅の範囲内の第1の量子化係数の候補の中から、事前確率(prior probability)に基づく量子化セル量が最大となる量子化係数を選択して、第1の量子化係数として逆写像を行う。

【選択図】 図4



【特許請求の範囲】

【請求項 1】

入力画像を第 1 の量子化幅に基づき量子化した第 1 の量子化係数から当該第 1 の量子化幅より広い第 2 の量子化幅に基づき量子化した第 2 の量子化係数について、前記第 1 の量子化係数への逆写像を行う逆写像部を備え、

前記逆写像部は、

前記第 2 の量子化幅の範囲内の前記第 1 の量子化係数の候補の中から、事前確率 (prior probability) に基づく量子化セル量が最大となる量子化係数を選択して、前記第 1 の量子化係数として前記逆写像を行う

画像処理装置。

10

【請求項 2】

前記逆写像部は、

前記事前確率を定義するスパース性モデル (sparsity model) に基づいて前記第 1 の量子化係数を近似して前記逆写像を行う

請求項 1 に記載の画像処理装置。

【請求項 3】

前記逆写像部は、

前記第 2 の量子化幅の範囲内の前記第 1 の量子化係数の範囲のうち前記事前確率の最大値を前記候補ごとに探索し、当該最大値と前記第 1 の量子化係数の範囲とに基づいて前記量子化セル量を算出し、前記候補の間で前記量子化セル量を比較することにより、前記量子化セル量が最大となる量子化係数を選択する

請求項 1 又は 2 に記載の画像処理装置。

20

【請求項 4】

前記第 1 の量子化係数を含む符号化画像について、前記第 1 の量子化係数と前記第 2 の量子化係数とが交互になるように再符号化を行うことにより再符号化画像へ変換する再符号化部をさらに備え、

前記逆写像部は、

前記再符号化画像内の前記第 2 の量子化係数に対して、当該第 2 の量子化係数と隣接する画素の前記第 1 の量子化係数を用いて、前記逆写像を行う

請求項 1 に記載の画像処理装置。

30

【請求項 5】

前記逆写像部は、

前記隣接する画素と隣接する画素の前記第 2 の量子化係数をさらに用いて、前記逆写像を行う

請求項 4 に記載の画像処理装置。

【請求項 6】

複数の画像について各画像の特徴量の分布から複数のクラスタに分類した辞書情報と、前記第 1 の量子化係数から前記第 2 の量子化係数への再符号化時に、前記入力画像が前記複数のクラスタのいずれに該当するかを判別したクラスターインデックスを、前記第 2 の量子化係数に対応付けて記憶装置に格納する再符号化部と、をさらに備え、

40

前記逆写像部は、前記第 2 の量子化係数に対応付けられた前記クラスターインデックスに基づき前記辞書情報内のクラスタを特定し、当該特定したクラスタを用いて前記逆写像を行う

請求項 1 に記載の画像処理装置。

【請求項 7】

複数の画像の特徴量の分布から複数のクラスタに分類した過完備辞書 (over-complete dictionary) と、

前記特徴量の分布をグラフ化し、前記複数のクラスタに分類した散在 (sparse) グラフテンプレートと、

前記複数の画像から前記過完備辞書と前記散在グラフテンプレートとを学習する学習部

50

と、

入力画像を第 1 の量子化幅に基づき量子化した第 1 の量子化係数を含む符号化画像について、前記第 1 の量子化係数から、前記第 1 の量子化幅より広い第 2 の量子化幅に基づき量子化した第 2 の量子化係数へ再符号化する再符号化部と、

前記第 2 の量子化係数について、前記過完備辞書と前記散在グラフテンプレートとを用いて、前記第 1 の量子化係数への逆写像を行う逆写像部と、を備え、

前記学習部は、

前記複数の画像から抽出された複数のブロック画像について、T S V Q (Tree-Structured Vector Quantization) により複数のクラスタに分類して二分木を構築し、

前記分類された各クラスタに対してクラスタインデックスのオフセットが減少するように、前記二分木のノードを入れ替え、

前記入れ替え後の前記二分木に対してレート歪みを最適化するように枝刈りを行い、

前記枝刈り後の前記二分木に対して前記過完備辞書と前記散在グラフテンプレートとを学習し、

前記再符号化部は、

再符号化時に、前記入力画像が前記学習後の前記過完備辞書及び前記散在グラフテンプレートのクラスタのいずれに該当するかを判別したクラスタインデックスを、前記第 2 の量子化係数に対応付けて記憶装置に格納し、

前記逆写像部は、

前記第 2 の量子化係数に対応付けられた前記クラスタインデックスに基づき前記過完備辞書及び前記散在グラフテンプレート内のクラスタを特定し、当該特定したクラスタを用いて前記逆写像を行う

画像処理装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、画像処理装置に関し、特に、トランスコーディングされた圧縮画像を処理するための画像処理装置に関する。

【背景技術】

【0002】

近年、ソーシャルメディアサイトや写真共有サイトが普及し、増加するユーザにより日々、大量の画像データがアップロードされている。そのため、このようなサイトを運営する事業者は、アップロード画像を高速かつ大量に保存することが求められる。

【0003】

しかしながら、生成される情報にヒープは成長するにもかかわらず人間に与えられた時間や関心は、限られたリソースといえる（非特許文献 1）。そのため、実際にアクセスされ、ユーザに閲覧される画像は、大量に保存された画像の中のほんの一部である。つまり、ユーザによりアップロードされる画像の量とユーザによりダウンロード（アクセスされて閲覧）される画像の量とは非対称の関係がある。そのため、サイト運営者のためにストレージ全体のコストを最小化するための方法が必要とされている。

【0004】

ここで、非特許文献 2 には、スパース信号（sparse signal）やスパース性モデル（sparsity model）に関する技術が開示されている。非特許文献 3 及び 4 には、逆量子化に関する技術が開示されている。

【0005】

また、クラウドストレージについての画像圧縮に関する技術は、例えば、非特許文献 5～8 に開示されている。非特許文献 5 には、S I F T ディスクリプタを介した大規模画像データベースから画像を再構成する技術が開示されている。非特許文献 6 には、画像間の相関を分析及び探索する画素値よりむしろローカルな特徴を探索することにより、写真アルバムを圧縮するための改良された技術が開示されている。

【0006】

非特許文献7には、新規なクラウドベース分配画像コーディングスキームに関する技術が開示されている。

【0007】

また、非特許文献8には、JBIGに関する技術が開示されている。

【0008】

ここで、特許文献1には、圧縮データの圧縮度を向上させるためのJPEG画像のトランスコーディングに関する技術が開示されている。また、特許文献2には、画像データをパケットデータとして取扱い、リアルタイムに画像データをメモリの記憶容量以内のデータ量として圧縮するための画像処理装置に関する技術が開示されている。

10

【先行技術文献】

【特許文献】

【0009】

【特許文献1】特表2006-501736号公報

【特許文献2】特開2003-244446号公報

【非特許文献】

【0010】

【非特許文献1】D. J. Levitin, The Organized Mind: Thinking Straight in the Age of Information Overload, Penguin, 2014.

【非特許文献2】M. Elad, M.A.T. Figueiredo, and Y. Ma, "On the role of sparse and redundant representations in image processing," Proc. IEEE, vol. 98, no. 6, pp. 972-982, June 2010.

20

【非特許文献3】X. Liu, G. Cheung, X. Wu, and D. Zhao, "Inter-block soft decoding of JPEG images with sparsity and graphsignal smoothness priors," in IEEE International Conference on Image Processing, Quebec City, Canada, September 2015.

【非特許文献4】X. Liu, X. Wu, J. Zhou, and D. Zhao, "Data-driven sparsity-based restoration of JPEG-compressed images in dual transform-pixel domain," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2015.

【非特許文献5】H. Yue, X. Sun, J. Yang, and F. Wu, "Cloud-based image coding for mobile devices-toward thousands to one compression," IEEE Trans. Multimedia, vol. 15, no. 4, pp. 845-857, June 2013.

30

【非特許文献6】Z. Shi, X. Sun, and F. Wu, "Photo album compression for cloud storage using local features," IEEE J. Emerg. Sel. Topic Circuits Syst., vol. 4, no. 1, pp. 17-28, Mar. 2014.

【非特許文献7】X. Song, X. Peng, J. Xu, G. Shi, and F. Wu, "Cloud based distributed image coding," IEEE Trans. Circuits Syst. Video Technol., vol. 26, no. 6, pp. 1-1, June 2016.

【非特許文献8】F. Ono, W. Rucklidge, R. Arps, and C. Constantinescu, "JBIG2-the ultimate bi-level image coding standard," in IEEE International Conference on Image Processing, Vancouver, Canada, September 2000.

40

【非特許文献9】Xianming Liu, Gene Cheung, Xiaolin Wu, Debin Zhao, "Random Walk Graph Laplacian based Smoothness Prior for Soft Decoding of JPEG Images," accepted to IEEE Transactions on Image Processing, October 2016. (arXiv)

【非特許文献10】M. Elad and M. Aharon, "Image denoising via sparse and redundant representation over learned dictionaries," in IEEE Transactions on Image Processing, vol. 15, no.12, December 2006.

【非特許文献11】J. Pang, G. Cheung, W. Hu, and O. C. Au, "Redefining self-similarity in natural images for denoising using graph signal gradient," in APSIPA ASC, Siem Reap, Cambodia, December 2014.

【非特許文献12】A. Gersho and R. M. Gray, Vector Quantization and Signal Compr

50

ession. Norwell, MA, USA: Kluwer Academic Publishers, 1991.

【非特許文献 1 3】L. Y. Wei and M. Levoy, “Fast texture synthesis using tree-structured vector quantization,” in SIGGRAPH '00: Proceedings of the 27th annual conference on Computer graphics and interactive techniques. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co., 2000, pp. 479-488. [Online]. Available: <http://dx.doi.org/10.1145/344779.345009>

【非特許文献 1 4】K. Ramchandran and M. Vetterli, “Best wavelet packet bases in a rate-distortion sense,” IEEE Transactions on Image Processing, vol. 2, no. 2, pp. 160-175, Apr 1993.

【発明の概要】

【発明が解決しようとする課題】

【0011】

ここで、上述した特許文献 1 等におけるトランスコーディングでは、復号時に粗量子化ビンインデックス（2 回目の量子化による量子化係数）から密量子化ビンインデックス（1 回目の量子化による量子化係数）への逆写像が困難であるという問題（量子化ビンマッチング（QBM、quantization bin matching）問題）がある。

【0012】

尚、非特許文献 5 から 8 に記載の技術は、ユーザによりアップロードされたオリジナル画像の正確な回復を保証できない。

【0013】

本発明は、このような問題点を解決するためになされたものであり、高い圧縮率を実現しつつ、復号時にユーザにとって許容できる画質を保証するための画像処理システムを提供することを目的とする。

【課題を解決するための手段】

【0014】

本発明の第 1 の態様にかかる画像処理装置は、

入力画像を第 1 の量子化幅に基づき量子化した第 1 の量子化係数から当該第 1 の量子化幅より広い第 2 の量子化幅に基づき量子化した第 2 の量子化係数について、前記第 1 の量子化係数への逆写像を行う逆写像部を備え、

前記逆写像部は、

前記第 2 の量子化幅の範囲内の前記第 1 の量子化係数の候補の中から、事前確率（prior probability）に基づく量子化セル量が最大となる量子化係数を選択して、前記第 1 の量子化係数として前記逆写像を行う。

【0015】

本発明の第 2 の態様にかかる画像処理装置は、

複数の画像の特徴量の分布から複数のクラスタに分類した過完備辞書（over-complete dictionary）と、

前記特徴量の分布をグラフ化し、前記複数のクラスタに分類した散在（sparse）グラフテンプレートと、

前記複数の画像から前記過完備辞書と前記散在グラフテンプレートとを学習する学習部と、

入力画像を第 1 の量子化幅に基づき量子化した第 1 の量子化係数を含む符号化画像について、前記第 1 の量子化係数から、前記第 1 の量子化幅より広い第 2 の量子化幅に基づき量子化した第 2 の量子化係数へ再符号化する再符号化部と、

前記第 2 の量子化係数について、前記過完備辞書と前記散在グラフテンプレートとを用いて、前記第 1 の量子化係数への逆写像を行う逆写像部と、を備え、

前記学習部は、

前記複数の画像から抽出された複数のブロック画像について、TSVQ（Tree-Structured Vector Quantization）により複数のクラスタに分類して二分木を構築し、

前記分類された各クラスタに対してクラスティンデックスのオフセットが減少するように

10

20

30

40

50

、前記二分木のノードを入れ替え、

前記入れ替え後の前記二分木に対してレート歪みを最適化するように枝刈りを行い、

前記枝刈り後の前記二分木に対して前記過完備辞書と前記散在グラフテンプレートとを学習し、

前記再符号化部は、

再符号化時に、前記入力画像が前記学習後の前記過完備辞書及び前記散在グラフテンプレートのクラスタのいずれに該当するかを判別したクラスタインデックスを、前記第2の量子化係数に対応付けて記憶装置に格納し、

前記逆写像部は、

前記第2の量子化係数に対応付けられた前記クラスタインデックスに基づき前記過完備辞書及び前記散在グラフテンプレート内のクラスタを特定し、当該特定したクラスタを用いて前記逆写像を行う。

【発明の効果】

【0016】

本発明により、高い圧縮率を実現しつつ、復号時にユーザにとって許容できる画質を保証するための画像処理システムを提供することができる。

【図面の簡単な説明】

【0017】

【図1】本発明の実施の形態1にかかるクラウドストレージシステムの全体構成を示すブロック図である。

【図2】本発明の実施の形態1にかかる画像アップロード時の流れを説明するためのシーケンス図である。

【図3】本発明の実施の形態1にかかる画像ダウンロード時の流れを説明するためのシーケンス図である。

【図4】本発明の実施の形態1にかかるQBMの解法の概念を説明するための図である。

【図5】本発明の実施の形態1にかかる粗密変換処理（量子化ビンマッチング）の流れを説明するためのフローチャートである。

【図6】本発明の実施の形態1にかかる圧縮利得の例を示す図である。

【図7】本発明の実施の形態1にかかるレート歪み性能圧縮の例を示す図である。

【図8】本発明の実施の形態1にかかるレート歪み性能圧縮の例を示す図である。

【図9】本発明の実施の形態1にかかるレート歪み性能圧縮の例を示す図である。

【図10】本発明の実施の形態2にかかるブロック内の画素に対する密量子化と粗量子化の関係を説明するための図である。

【図11】本発明の実施の形態3にかかる画像処理システムの全体構成を示すブロック図である。

【図12】本発明の実施の形態3にかかる粗密変換処理の流れを説明するためのフローチャートである。

【図13】本発明の実施の形態3にかかるBag of Visual Words (BoVW)を利用する場合の概念を示す図である。

【図14】本発明の実施の形態3にかかるある画像における特徴量の分布の例を示す図である。

【図15】本発明の実施の形態4にかかるクラウドストレージシステムの全体構成を示すブロック図である。

【図16】本発明の実施の形態4にかかるオフライントレーニングの流れを説明するための図である。

【図17】本発明の実施の形態4にかかる散在グラフテンプレートの例を示す図である。

【図18】本発明の実施の形態4にかかるオフライントレーニングで生成されるツリー概念を説明するための図である。

【図19】部分空間の検索を説明するための図である。

【図20】二分木のノードの入れ替えを説明するための図である。

10

20

30

40

50

【図 2 1】ビット割当前後のインデクスオフセットの確率分布の比較を示す図である。

【図 2 2】微分カウンタの概念を説明するための図である。

【図 2 3】本発明の実施の形態 4 にかかるコードブック設計アルゴリズムの例を示す図である。

【発明を実施するための形態】

【0018】

以下では、本発明を適用した具体的な実施の形態について、図面を参照しながら詳細に説明する。各図面において、同一要素には同一の符号が付されており、説明の明確化のため、必要に応じて重複説明は省略する。

【0019】

<実施形態 1>

図 1 は、本発明の実施の形態 1 にかかるクラウドストレージシステムの全体構成を示すブロック図である。クラウドストレージシステムは、ユーザ端末 11、12 等と、クラウドレット 20 と、クラウドストレージ 30 とを備える。ユーザ端末 11 等は、ネットワーク（不図示）を介してクラウドレット 20 と接続されている。

【0020】

ユーザ端末 11 は、ユーザ U1 が操作する情報処理端末であり、例えば、携帯電話機、スマートフォン、タブレット端末、PC 等である。ユーザ端末 11 は、画像符号化部 111 と、画像復号化部 112 とを備える。画像符号化部 111 は、カメラ等により撮影された入力画像 P1 について量子化パラメータ（quantization parameters (QP)） q を用いて J P E G の符号化を行い、符号化画像 P2 を生成する。つまり、符号化画像 P2 は、入力画像 P1 に対する圧縮データである。ここで、ユーザ端末 11 は、符号化画像 P2 をクラウドレット 20 へアップロードすることができる。画像復号化部 112 は、J P E G で符号化された画像を表示用に復号化する。

【0021】

ユーザ端末 12 は、ユーザ U2 が操作する情報処理端末である。ユーザ端末 12 は、ユーザ端末 11 と同等の構成である画像符号化部 121 及び画像復号化部 122 を備える。ユーザ端末 12 は、クラウドレット 20 から J P E G で符号化された画像である復元符号化画像 P4 をダウンロードした場合に、量子化パラメータ q を用いて J P E G の復号化を行い、復号画像 P5 として出力（表示等）する。

【0022】

クラウドレット 20 は、1 以上のコンピュータ装置により実現されたクラウドサーバである。クラウドレット 20 は、ネットワークを介してユーザ端末 11 等と、クラウドストレージ 30 と接続されている。クラウドレット 20 は、再符号化部 21 と、粗密変換部 22 と、辞書 DB 23 とを備える。辞書 DB 23 は、復号時に参照する情報が格納されたデータベースである。

【0023】

再符号化部 21 は、符号化画像 P2 に対して、 $Q P - q$ より粗い（幅の広い） Q を用いて再符号化を行い、再符号化画像 P3 を生成する。そして、クラウドレット 20 は、再符号化画像 P3 をクラウドストレージ 30 へ送信し、保存する。

【0024】

粗密変換部 22 は、逆写像部の一例であり、ユーザ端末 12 からの画像の要求に応じてクラウドストレージ 30 から再符号化画像 P3 を取得し、辞書 DB 23 に基づき再符号化画像 P3 に対して粗密変換（量子化ビンマッチング）により復元符号化画像 P4 に復元し、ユーザ端末 12 へ返信する。

【0025】

クラウドストレージ 30 は、複数の永続型のストレージ 31 ~ 3n を備えたクラウド型のストレージであり、セントラルクラウドとも呼ばれる。クラウドストレージ 30 は、ネットワークを介してクラウドレット 20 と接続される。例えば、ストレージ 31 には、上述した再符号化画像 P3 が格納される。

10

20

30

40

50

【0026】

図2は、本発明の実施の形態1にかかる画像アップロード時の流れを説明するためのシーケンス図である。まず、ユーザ端末11は、入力画像P1についてQP_qを用いてJPEGの符号化を行い、符号化画像P2を生成する(S101)。具体的には、画像符号化部111は、入力画像P1についてDCT(Discrete Cosine Transform)によりDCT係数へ変換する。次に、画像符号化部111は、量子化パラメータ(QP)_qを用いて量子化を行い、DCT係数から量子化係数へ変換する。すなわち、画像符号化部111は、Nピクセルのコードブロックmごとに、QP_q = [q1, ..., qN]を用いて量子化インデクス集合a(m) = [a1(m), ..., aN(m)]を計算する。尚、以降では、a(m)を「密(fine)量子化ビンインデクスa」と表現する場合がある。そして、画像符号化部111は、量子化係数に対してエントロピー符号化により符号データに変換する。つまり、符号化画像P2には、密量子化ビンインデクスaが含まれている。尚、「量子化パラメータ」は、「量子化テーブル」や「量子化の幅」と呼ぶことができるものとする。

10

【0027】

そして、ユーザ端末11は、符号化画像P2(密量子化ビンインデクスa)をクラウドレット20へアップロードする(S102)。クラウドレット20の再符号化部21は、受信した符号化画像P2に対してQP_Qを用いてJPEG再符号化を行い、再符号化画像P3に変換する(S103)。具体的には、再符号化部21は、コードブロックmごとに、各a(m)をQP_Q = [Q1, ..., QN]を用いて、対応する量子化インデクス集合b(m) = [b1(m), ..., bN(m)]へ写像し、順番に再符号化する。ここで、QP_Qは、上述したqよりも粗い(coarser)、つまり、量子化の幅が広いものとする。言い換えると、QP_qは、Qよりも精細(fine)、つまり、量子化の幅が狭い。尚、以降では、b(m)を「粗(coarse)量子化ビンインデクスb」と表現する場合がある。また、JPEG再符号化には、例えば、特許文献1の手法を用いることもできる。

20

【0028】

ここで、a(m)からb(m)への写像は、一意でなくてよい。例えば、再符号化部21は、各画素の量子化係数を再符号化する際に、一旦、QP_Qで量子化した後に、復号を行い、可逆変換が行えることを確認できた場合に、実際に、当該画素の再符号化を行い、可逆変換とならない場合には、QP_qのままとする。このとき、オリジナル信号を見積もるために、例えば、非特許文献3の逆量子化処理を用いることができる。つまり、粗量子化ビンインデクスbは、一部が粗いQP_Qで量子化(再符号化)されているが、残りは精細なQP_qで量子化されたままである。

30

【0029】

その後、クラウドレット20は、再符号化画像P3(粗量子化ビンインデクスb)をクラウドストレージ30へ保存する(S104)。

【0030】

図3は、本発明の実施の形態1にかかる画像ダウンロード時の流れを説明するためのシーケンス図である。ここでは、既に図2の処理によりクラウドストレージ30に再符号化画像P3が保存済みとする。

【0031】

まず、ユーザ端末12は、ユーザU2の操作により入力画像P1に相当する画像の要求をクラウドレット20へ送信する(S201)。続いて、クラウドレット20は、クラウドストレージ30に対して再符号化画像P3の検索を行う(S202)。そして、クラウドレット20は、クラウドストレージ30から画像の検索結果として再符号化画像P3(粗量子化ビンインデクスb)を取得する(S203)。

40

【0032】

ここで、クラウドレット20の粗密変換部22は、辞書DB23を参照し、各b(m)をa(m)へ戻す変換である逆写像を実行し、つまり、粗密変換により復元符号化画像P4(密量子化ビンインデクスa*)に復元する(S204)。すなわち、粗いQP_Qからより精細なQP_qに戻される。そして、粗密変換部22は、復元符号化画像P4をユ

50

ーザ端末12へ送信する。つまり、ユーザ端末12は、復元符号化画像P4をダウンロードする(S205)。その後、ユーザ端末12の画像復号化部122は、復元符号化画像P4に対して量子化パラメータqを用いてJPEGの復号化を行う(S206)。これにより、ユーザ端末12は、復号画像P5を表示等できる。

【0033】

このように本発明の実施の形態1は、全画像を保存することによるストレージコストと、全画像のほんの一部を復号(逆写像)するための演算コストとのトレードオフにより、効果的となるクラウドストレージシステムといえる。つまり、まず、永続的ストレージには、アップロードされたJPEG画像を粗い量子化パラメータ(QP)を用いて再符号化した再符号化画像を保存することで、ストレージコストを削減する。そして、要求画像については、再符号化画像を粗量子化ビンインデックスから密量子化ビンインデックスへの逆写像を行うことで、ユーザ端末側では通常の復号を行わせる。このとき、再符号化画像の逆写像処理には、通常より演算コストがかかるが、上述したように、保存を要求される画像は大量であるが、実際に閲覧を要求される画像はほんの一部である。よって、復号時の演算コストをかけたとしてもストレージコストを削減できることの意義は大きいといえる。

【0034】

そして、ユーザ側の視点に立てば、最終的に復号化され表示された画像は、当初アップロードされたものと同じレベルの密量子化ビンインデックスの画像といえる。つまり、人間の目により区別がつかないビジュアル品質を維持しており、典型的なクラウドストレージと同等のサービスレベルを提供できる。一方、運営者(オペレータ)側の視点に立てば、保存前の再符号化によって、当初よりも粗い粗量子化ビンインデックスの画像を保存できるため、通常より高い圧縮利得及びより低いストレージコストという結果をもたらす。

【0035】

但し、上述したようにQP-Qは、QP-qよりも粗い(すなわち、Qiはqiより大きい)ので、bi(m)はai(m)よりもゼロである可能性が高い。すなわち、ai(m)で0以外だった量子化係数が、bi(m)では0となるものがある。そのため、画像が要求された場合に、クラウドレットにおいて粗量子化ビンインデックスから密量子化ビンインデックスへの逆写像をより精度良く行う必要がある。このことを、量子化ビンマッチング問題(quantization bin matching (QBM) problem)と呼ぶ。

【0036】

本実施の形態では、後述する最大事後確率(maximum a posteriori (MAP)) 公式を介して粗量子化ビンインデックスから密量子化ビンインデックスへの逆写像を行い、そのために信号スパース性モデル(signal sparsity prior)を有効活用する。

【0037】

ここで、信号スパース性モデル(非特許文献2)を仮定すると、辞書のスパース線形結合(sparse linear combination)により良く近似される信号がより可能性が高く、固定サイズNピクセルのコードブロックのそれぞれについて、N個の粗量子化ビンインデックスが与えられる周波数領域内の最も確からしいN個の密量子化ビンインデックスを見つけ出すための最大事後確率(maximum a posteriori (MAP)) 問題を公式化する。

【0038】

この公式は、所与のN次元量子化セル上での事前分布の統合を要求し、過去の逆量子化研究よりも量子化セル内の単独で最も可能性の高い信号が探索される事前分布統合を近似する速いアルゴリズムを提案するものである。

【0039】

よって、本実施の形態では、入力画像が十分に復元できることを保証するために以下を考慮して設計される。

- 1) 各ブロックmについて元の密量子化ビンインデックスa(m)が正確に復元されること
- 2) 復元画像がピーク信号対雑音比(Peak signal-to-noise ratio (PSNR)) 内の入力圧縮画像に非常に近いために、ユーザにとって視覚的に見分けがつかないこと

【0040】

そのため、オリジナルの $Q P - q$ よりも粗い $Q P - Q$ について、入力画像がよく復元されるような逆写像 $f(b)$ を設計することを目的とする。ここで、 $Q P - Q$ を用いて再符号化される画像内の充分大きなブロックの一部について、逆写像が可能であることのみが要求され、残りのブロックは $Q P - q$ を用いて変更されない。J B I G でロスレス符号化される小さなバイナリ画像は、ブロックが Q を用いて再量子化されるデコーダを通知するために追加的に符号化される。以降では、粗い $Q P - Q$ から密な $Q P - q$ への量子化ビンマッチング問題について検討する。

【0041】

(量子化セルに関する M A P 公式)

前提として、 X_i は、ターゲットコードブロック x のグラントゥールス (ground-truth) の i 番目の D C T 係数を表す。そして、 a_i は、以下の式 (1) に示すように、 X_i を $Q P - q_i$ を用いて量子化した量子化係数 (密量子化ビンインデックス) である。

【数 1】

$$a_i = \text{round} \left(\frac{X_i}{q_i} \right) \cdots (1)$$

【0042】

従って、 a_i 及び q_i を備えるデコーダにおいて、 X_i が間隔 $I(a_i, q_i) = [(a_i - 0.5)q_i, (a_i + 0.5)q_i]$ の範囲に存在しなければならない。また、 b_i は、再符号化画像内の X_i について $Q P - Q_i$ (Q_i は q_i 以上) を用いて量子化した量子化係数 (粗量子化ビンインデックス) である。従って、再符号化後の X_i は、より大きな間隔 $I(b_i, Q_i) = [(b_i - 0.5)Q_i, (b_i + 0.5)Q_i]$ の範囲に存在しなければならない。なぜならば、オリジナル符号化内の間隔 $I(a_i, q_i)$ 及び再符号化内の間隔 $I(a_i, Q_i)$ の両方に同じ X_i に含まれるため、それらは重複するためである。

従って、逆写像 $f: b_i \rightarrow a_i$ を実行する時、式 (2) に示す、実現可能なビン集合 F_{b_i} を検討する必要がある。

【数 2】

$$F_{b_i} = \{a_i | I(a_i, q_i) \cap I(b_i, Q_i) \neq \emptyset\} \cdots (2)$$

【0043】

続いて、 $P(X_i)$ は、 X_i の事前確率 (prior probability) とする。最大事後確率 (maximum a posteriori (MAP)) 公式は、以下の式 (3) 及び式 (4) に基づき、離散集合 F_{b_i} 内の最も確からしいビン a_i を探索する。

【数 3】

$$\hat{a}_i = \arg \max_{a_i \in F_{b_i}} \int_{(a_i-0.5)q_i}^{(a_i+0.5)q_i} P(X_i | b_i) dX_i \cdots (3)$$

ここで、

【数 4】

$$P(X_i | b_i) = \begin{cases} P(X_i) & \text{if } X_i \in I(b_i, Q_i) \\ 0 & \text{o.w.} \end{cases} \cdots (4)$$

である。

【0044】

一次元での Q B M 問題の M A P 公式が図 4 に示される。図 4 は、本発明の実施の形態 1 にかかる Q B M の解法の概念を説明するための図である。幅 Q の粗量子化ビンと重複する幅 q の 3 つの密量子化ビン (A , B , C) の中から最大集計確率 (largest aggregate probability) ($P(X_i | b_i)$ の積分値) の一つを選択する。これは、非特許文献 3 のような典型的な逆量子化研究とは異なり、量子化ビン内の最も可能性の高い信号についての M A P 公式は、ビンの範囲内で $P(X_i | b_i)$ のピーク位置へ簡潔に導く。

【 0 0 4 5 】

ここで、N次元の事前確率 $P(X)$ と共にブロック x 内のN個のDCT係数 X の全てを考えると、以下の式 (5) に示す、より一般的なMAP最適化問題を定式化できる。

【数 5】

$$\max_{a|a_i \in F_{b_i}} \int_{(a_1-0.5)q_1}^{(a_1+0.5)q_1} \cdots \int_{(a_N-0.5)q_N}^{(a_N+0.5)q_N} P(X|b) dX_1 \cdots dX_N \quad \cdots (5)$$

ここで、

【数 6】

$$P(X|b) = \begin{cases} P(X) & \text{if } X_i \in I(b_i, Q_i), \quad 1 \leq i \leq N \\ 0 & \text{o.w.} \end{cases} \quad \cdots (6)$$

10

である。

【 0 0 4 6 】

以下の式 (7) は、多次元の積分を示す式 (5) をより簡潔な形式で書き直したものである。

【数 7】

$$\max_{a|a_i \in F_{b_i}} \bar{P}_a(X|b) \quad \cdots (7)$$

20

ここで、 $\bar{P}_a(X|b)$ は、 a (すなわち、 $I(a_i, q_i)$; $i = 1, \dots, N$) により定義される量子化セル C_a 内の確率 $P(X|b)$ の集計である。式 (7) を解くための課題は、 $P(X)$ を適切に定義する方法、 $\bar{P}_a(X|b)$ を効率的かつ正確に計算する方法である。

【 0 0 4 7 】

(事前確率 $P(X)$ の定義)

続いて、事前確率 $P(X)$ を定義するためのスパースモデル (非特許文献2) を仮定する。具体的には、 k スパースの信号モデルは、画素領域内のN次元信号 x が過完備辞書 (over-complete dictionary) Φ からの k 又は少ない原子 (最小単位) の線形結合により十分に近似されることである。

30

【数 8】

$$x = \Phi\alpha + \varepsilon, \quad \|\alpha\|_0 \leq k \quad \cdots (8)$$

ここで、モデルエラー ε は小さい。本実施の形態では、機械学習駆動方法内でPCA経由で適応辞書を学習する。

【 0 0 4 8 】

式 (8) の両辺にDCT演算子 T を適用することにより、式 (9) が得られる。

【数 9】

$$X = \Psi\alpha + \varepsilon' \quad \cdots (9)$$

40

ここで、 $\Psi = T\Phi$ である。それ故、 X のスパース性は、 Φ に関して x のスパース性により測定され得る。

【 0 0 4 9 】

このモデルが与えられ、同時確率分布 $P(X)$ は、式 (10) として表され得る。

【数 10】

$$P(X) = \exp\left\{-\frac{\|\alpha\|_0}{\sigma}\right\} \quad \cdots (10)$$

ここで、 σ は、モデルパラメータである。L0ノルムからL1ノルムへさらに緩和する。

50

【 0 0 5 0 】

最後に、目的関数は、式（ 1 1 ）として書き換えられ得る。

【 数 1 1 】

$$\arg \min_{a|a_i \in F_{b_i}} \int_{(a_1-0.5)q_1}^{(a_1+0.5)q_1} \cdots \int_{(a_N-0.5)q_N}^{(a_N+0.5)q_N} \|\alpha\|_1 dX_1 \cdots dX_N, \quad \cdots (11)$$

$$s.t., \quad x = \Phi\alpha + \varepsilon$$

【 0 0 5 1 】

上述したように、QBM問題は、大抵、画像逆量子化（非特許文献3及び4）で行われるシングルベストスパース解（sparse solution）よりむしろ、その境界内で最大かつ最もスパースな解を用いて量子化セルを見出すこととして再公式化できる。

10

【 0 0 5 2 】

（最適化）

上述した式（ 1 1 ）の目的関数を直接最適化することは困難である。そこで、量子化セル内の全ての可能なスパース解について検索する代わりに、代表としてシングルベストスパース解を見つけ出し、セルのボリュームによる解（解のスパースカウントの指数関数）の事前確率値を乗ずる。これは、式（ 1 1 ）の多次元統合の近似値であり、本質的に、セルボリュームが大きくなるほど、より確からしい発見されたもののような、それ故に、より大きな集計確率の他のスパース解を見つけ出す。インデクスされた粗量子化セルb内の初期スパース解を見つける問題を式（ 1 2 ）に示す。

20

【 数 1 2 】

$$\alpha^* = \arg \min \|x - \Phi\alpha\|_2^2 + \lambda \|\alpha\|_1, \quad \cdots (12)$$

$$s.t., \quad X_i \in I(b_i, Q_i), i=1, \cdots, N, \quad and \quad X = T_x$$

【 0 0 5 3 】

スパース解 α^* に関する最適化は、拡張ラグランジュ法（Augmented Lagrangian Methods (ALM)）として知られる速い L1 最小化アルゴリズムにより効果的かつ効率的に解かれる。

【 0 0 5 4 】

式（ 1 3 ）のような本スパース解を含む密量子化セル a^* を同定する。

30

【 数 1 3 】

$$a^* = \text{round} \left(\frac{T\Phi\alpha^*}{q} \right) \quad \cdots (13)$$

【 0 0 5 5 】

同定された密量子化セル a^* が（例えば、図4のセルA及びCのように）粗量子化セルbと部分的にのみ重複する場合、セルボリュームが小さいため、その集計確率は小さい。他の解の候補をテストするために、式（ 1 2 ）を介して密な隣接の量子化セルについて、スパース解について検索する。これらのセルの中で、（セル内の同定されたスパース解の）信号モデルとセルボリュームの乗算結果のうち最大であるものが、最終的な解として選択される。

40

【 0 0 5 6 】

図5は、本発明の実施の形態1にかかる粗密変換処理（図3のS204）の流れを説明するためのフローチャートである。まず、事前に、クラウドレット20は、8 x 8 ブロックについて辞書を学習しておくものとする。

【 0 0 5 7 】

次に、粗密変換部22は、式（ 1 2 ）を用いて、粗量子化ビンインデクス内で最もスパースな解の結果を探索する（S302）。すなわち、粗密変換部22は、粗量子ビンイン

50

デクス b_i の範囲内の複数の候補範囲 (A, B, C) それぞれについて、最もスパースな解を探索する。図 4 の場合、 x_3 、 x_4 、 x_6 が求まる。

【0058】

そして、粗密変換部 22 は、 $P(X|a)$ を推定する (S303)。つまり、A の中で最大の $P(x_3)$ を算出する。そして、A の幅 ($x_4 - x_2$) に、 $P(x_3)$ を乗じて量子化セル A の量を算出する。同様に、B については、B の幅 ($x_5 - x_4$) に最大値 $P(x_4)$ を乗じて量子化セル B の量を算出する。C については、C の幅 ($x_6 - x_5$) に最大値 $P(x_6)$ を乗じて量子化セル C の量を算出する。

【0059】

その後、粗密変換部 22 は、隣接量子化セル (A, B, C) の間で量子化セル量を比較して、 $QP - q$ の密量子化ビンインデクス a^* を特定する。図 4 では、B が特定される。尚、上述した通り、図 4 は、一次元の場合を示しており、より一般的には、 8×8 ブロックにおける 64 次元があり得る。

【0060】

(実験結果)

本実施の形態の有効性を実証するための次の実験を行った。まず、周知のコダックデータセットを用いた。5 つの画像が辞書学習用のデータを学習 (training) するようにランダムに選択され、残りはテスト画像として使用される。

【0061】

JPEG 符号化での品質要素 (quality factor (QF)) は、量子化マトリクスの集合から一つを選択することにより、圧縮後に画像の相対的なビジュアル品質を 1 から 100 の範囲で示す。実験では、粗い QF は異なる画像について 50 及び 55 から選択されるが、JPEG 符号化の精細な QF は 80 で固定される。

【0062】

ここで、第 1 の設計目的は、オリジナルアップロードバージョンとしての密量子化ビンインデクスの正確な回復を保証することである。この場合、正確なビンインデクスの回復は、質的にロスがないことを意味する。それ故に、ビット節約 (bit saving) について報告する。尚、ビット節約は、正確なビンインデクス回復を保証できる粗い QF を有する画像ブロックの一部を再符号化することにより達成される。

【0063】

図 6 は、本発明の実施の形態 1 にかかる圧縮利得の例を示す図である。図 6 では、精細な QF 及び粗い QF、再符号化されるために選択されたブロックの割合、及び、8 つのテスト画像における本手法の純粋な圧縮利得を表にしたものである。テスト画像について、42.18% 以下のブロックは、粗い QF を使用する追加の圧縮のために選択され、14.19% 以下のビット節約という結果が示された。

【0064】

第 2 の代替案の設計目的は、人間の目には見わけがつかないほど、質的に非常に近いアップロード画像の回復を保証することである。これを達成するために、回復された及び入力密量子化ビンインデクス間の違いを許可することにより、「正確なビンマッチング」を緩和する。差分のレベルは、ブロックレベルビンエラーの合計により計測される。本実験では、エラーの合計が 0, 2, 3, 4, 5 及び 6 である 6 ケースをテストする。

【0065】

図 7 は Lighthouse、図 8 は Ahoy、図 9 は Airplane の 3 テスト画像で JPEG を用いた本発明の実施の形態 1 にかかるレート歪み性能圧縮を示す。6 ケースの PSNR ロスは 0.13 dB 以内である。そのような小さな PSNR ロスは典型的には、視覚的な違いがないといえる。つまり、結果として、本手法の JPEG の効率が良いことを示す。

【0066】

さらに、非常にわずかな PSNR の費用で、視覚的に見分けがつかない再構成オプションは、正確なビンインデクス回復のケースと比較して、十分に高いビット節約を達成する。例として 6 のエラー合計を用いて、本手法は、Lighthouse、Ahoy 及び Airplane のそれぞ

10

20

30

40

50

れについて、図 6 に示される 3.32%、2.01% 及び 14.19% とは対照的に、16.71%、12.82% 及び 20.7% のビット節約を達成する。つまり、十分なストレージの節約を実現できるといえる。

【0067】

以上のことから、本実施の形態では、大量のアップロード画像を保存するためのストレージコストと、少数の要求画像を粗密変換することの計算コストとをトレードオフする JPEG 画像に関するクラウドストレージシステムを提案した。具体的には、予めユーザ端末において量子化パラメータ (QP) q を用いて量子化された符号化画像に対して、より粗い QP Q を用いて再符号化してストレージに保存しておく。これにより、ストレージコストを低減できる。その上で、画像の取得要求に応じて、上述した粗密変換 (量子化ビンマッチング) により粗量子化ビンインデックスから密量子化ビンインデックスへ復元を行うものである。これにより、ユーザにとっては、視覚的に違和感の少ない画像品質を維持できる。

【0068】

ここで、本実施の形態は、次のように表現することもできる。すなわち、入力画像 P_1 を第 1 の量子化幅 (QP q) に基づき量子化した第 1 の量子化係数 (密量子化ビンインデックス a) から当該第 1 の量子化幅より広い第 2 の量子化幅 (QP Q) に基づき量子化した第 2 の量子化係数 (粗量子化ビンインデックス b) について、前記第 1 の量子化係数への逆写像を行う逆写像部を備える画像処理装置である。特に、前記逆写像部は、前記第 2 の量子化幅の範囲内の前記第 1 の量子化係数の候補の中から、(例えば、対象画素の DCT 係数 X_i の) 事前確率 (prior probability) に基づく量子化セル量が最大となる量子化係数を選択して、前記第 1 の量子化係数として前記逆写像を行うものである。つまり、逆写像部は、図 4 に示すように QP Q の幅 ($x_6 - x_2$) の範囲内の密量子化ビンインデックスの候補 (A、B 及び C のそれぞれに属する密量子化ビンインデックス) の中から、 $P(X)$ と各候補の幅 (A の場合、 $x_4 - x_2$ 、B の場合、 $x_5 - x_4$ 、C の場合、 $x_6 - x_5$) とから量子化セル量 (A、B、C のそれぞれの $P(X)$ の積分値相当) を求め、各量子化セル量の中で最大のものを逆写像後の密量子化ビンインデックスとする。

【0069】

また、前記逆写像部は、前記事前確率を定義するスパース性モデル (sparsity model) に基づいて前記第 1 の量子化係数を近似して前記逆写像を行う。

【0070】

さらに、前記逆写像部は、前記第 2 の量子化幅の範囲内の前記第 1 の量子化係数の範囲のうち前記事前確率の最大値を前記候補ごとに探索し、当該最大値と前記第 1 の量子化係数の範囲とに基づいて前記量子化セル量を算出し、前記候補の間で前記量子化セル量を比較することにより、前記量子化セル量が最大となる量子化係数を選択する。つまり、逆写像部は、密量子化ビンインデックスの候補 A、B 及び C のそれぞれについて $P(X)$ の最大値を算出し、 $P(X)$ の最大値と A、B 及び C の幅との積 (面積) を算出し、それらを比較することで、最大のものを選択するものである。

【0071】

<実施形態 2>

続いて、上述した実施形態 1 の変形例である実施形態 2 について説明する。実施形態 1 では、再符号化部 21 における再符号化 (トランスコーディング) の際に、一旦、QP Q を用いて変換した上で、元の符号に復元できるかをシミュレーションし、可逆変換が保証された場合のみ QP Q による再符号化を行う。また、実施形態 1 では、1 画素単位に再符号化を行っていたため、処理効率及び圧縮効率に改善の余地があった。

【0072】

そこで、本実施形態 2 では、再符号化の対象画素を 1 画素置きとする。つまり、1 ブロックの画素について、密量子化と粗量子化とが交互になるようにする。図 10 は、本発明の実施の形態 2 にかかるブロック内の画素に対する密量子化と粗量子化の関係を説明するための図である。図 10 の上部で示すように、再符号化の際に、密量子化ビンインデックス

10

20

30

40

50

(fine)と粗量子化ビンインデクス(coarse)の画素が交互に配置されていることを示す。特に、本実施形態2にかかる再符号化部21は、破線で囲んだように4画素を一単位とし、図10の下部で示すように、右下の画素(4)について再符号化した場合に、隣接する3画素(1, 2, 3)を用いて粗密変換をすることで、可逆変換が保証されるかを判定するものである。また、右下の画素(4)の上の画素(2)と左の画素(3)とは、密量子化ビンインデクス(fine)のままであり、再符号化されておらず情報量が保持されている。また、右下の画素(4)と左上の画素(1)とは、粗量子化ビンインデクス(coarse)となり得ることを示す。

【0073】

そのため、再符号化部21は、少なくとも上の画素(2)と左の画素(3)の2画素を用いて可逆変換の保証を判定することにより、右下の画素(4)の再符号化の可否を判定し、粗密変換部22は、上の画素(2)と左の画素(3)の2画素を加味して右下の画素(4)の逆写像を行うことが望ましい。これにより、実施形態1と比べて、可逆変換の精度が向上し、結果として、再符号化ができる確率が上がり、圧縮率が向上する。

【0074】

さらに、上記に加え、再符号化部21及び粗密変換部22は、左上の画素(1)を加味して、右下の画素(4)の再符号化の可否判定及び逆写像を行うとよい。これにより、上記よりさらに可逆変換の精度及び圧縮率が向上する。

【0075】

また、本実施形態2では、上述したように4画素単位で可逆変換の保証を判定することから、クラウドレット20による事前の辞書学習について、16x16ブロックについて行うものとする。

【0076】

また、本実施形態2では、粗密変換処理において、図5のステップS302の際に、以下の式(14)を用いることとなる。

【数14】

$$\alpha^* = \arg \min \|x - \Phi \alpha\|_2^2 + \lambda \|\alpha\|_1, \quad \dots (14)$$

$$s.t., \quad TE_1 x \in I(b_1, Q), TE_2 x \in I(a_2, q), TE_3 x \in I(a_3, q), TE_4 x \in I(b_4, Q)$$

つまり、制約条件として、密量子化と粗量子化とが交互になった4画素を対象とすることになる。このように、4画素を対象とすることで粗密変換(量子化ビンマッチング)の際に、並列に復号処理を行うことで、粗密変換処理を効率化できる。

【0077】

ここで、本実施の形態は、次のように表現することもできる。すなわち、前記第1の量子化係数を含む符号化画像P2について、前記第1の量子化係数と前記第2の量子化係数とが交互になるように再符号化を行う(図10)ことにより再符号化画像へ変換する再符号化部をさらに備える。そして、前記逆写像部は、前記再符号化画像内の前記第2の量子化係数(画素(4)の粗量子化ビンインデクス)に対して、当該第2の量子化係数と隣接する画素の前記第1の量子化係数(画素(2, 3)の密量子化ビンインデクス)を用いて、前記逆写像を行う。

【0078】

さらに、前記逆写像部は、前記隣接する画素と隣接する画素の前記第2の量子化係数(画素(1)の粗量子化ビンインデクス)をさらに用いて、前記逆写像を行う。

【0079】

<実施形態3>

続いて、上述した実施形態1又は2の変形例である実施形態3について説明する。本実施形態3では、複数の種類の辞書を用いるものである。図11は、本発明の実施の形態3にかかる画像処理システムの全体構成を示すブロック図である。本発明の実施の形態3では、図1と比べてクラウドレット20がクラウドレット20aに置き換わり、ストレージ31に再符号化画像P3と対応付けられたクラスティンデクス311が保存されるもので

10

20

30

40

50

ある。クラウドレット 20 a は、再符号化部 21 a と粗密変換部 22 a と辞書 DB 23 a とを備える。

【0080】

辞書 DB 23 a は、複数のテスト画像の特徴量に基づきその傾向が分類された複数のクラスタのデータベースである。再符号化部 21 a は、辞書 DB 23 a を参照し、符号化画像 P 2 が属するクラスタを特定する。クラウドレット 20 a は、再符号化画像 P 3 及び特定したクラスタのクラスタインデクス 311 をクラウドストレージ 30 へ送信し、保存する。ストレージ 31 は、再符号化画像 P 3 に対応付けてクラスタインデクス 311 も保存する。

【0081】

粗密変換部 22 a は、ユーザ端末 12 からの画像の要求に応じてクラウドストレージ 30 から再符号化画像 P 3 及びクラスタインデクス 311 を取得し、辞書 DB 23 a を参照し、クラスタインデクス 311 に対応する辞書を用いて、再符号化画像 P 3 に対して粗密変換により復元符号化画像 P 4 に復元し、ユーザ端末 12 へ返信する。

【0082】

図 12 は、本発明の実施の形態 3 にかかる粗密変換処理の流れを説明するためのフローチャートである。具体的には、図 5 の最初にステップ S 301 が追加されたものであり、以降の処理は同等である。すなわち、粗密変換部 22 a は、クラウドストレージ 30 から再符号化画像 P 3 に加えて、再符号化画像 P 3 に対応付けられたクラスタインデクス 311 を取得する。そして、粗密変換部 22 a は、辞書 DB 23 a を参照し、クラスタインデクス 311 から使用する辞書を選択する (S 301)。以後、粗密変換部 22 a は、辞書 DB 23 a 内の選択された辞書を用いて粗密変換処理を行う。

【0083】

ここで、本実施の形態 3 における辞書の学習の仕方について説明する。辞書の学習は、(ステップ 1) 特徴量のヒストグラムにより各画像をクラスタに分類し、(ステップ 2) クラスタごとに辞書の学習を行う。

【0084】

ここで、本実施の形態 3 における辞書の学習の一例として、Bag of Visual Words (BoVW) を利用する場合を説明する。図 13 は、本発明の実施の形態 3 にかかる BoVW を利用する場合の概念を示す図である。上記ステップ 1 について説明する。まず、図 13 の上段の 3 つのテスト画像について複数のブロック (又は画素) に分割する (図 13 の下段)。そして、例えば、各ブロックを解析し、3 種類の特徴 f a、f b 及び f c のそれぞれの出現頻度 (適合度合) を集計する。図 13 の中段は、特徴 f a、f b 及び f c のそれぞれについてのヒストグラムの例である。そして、これらのヒストグラムにより、各テスト画像をクラスタに分類する。例えば、自転車のサドルのブロックは特徴 f b が強く、人物の肌や目のブロックは特徴 f a が強く、バイオリンの一部のブロックは特徴 f c が強いという傾向があり、各ブロックについて、クラスタ分類ができる。

【0085】

図 14 は、本発明の実施の形態 3 にかかるある画像における特徴量の分布の例 (ヒストグラム) を示す図である。つまり、ある画像における各ブロックについての特定の特徴量のヒストグラムを示す。そのため、以降、上記ステップ 2 については、学習対象の画像についてはいずれのクラスタに属するかにより該当するクラスタに対して学習が行われる。これにより、辞書の精度が向上する。

【0086】

そして、上述の通り再符号化の際には、符号化画像 P 2 についてクラスタの分類を行い、クラスタインデクス (画像の特徴) を特定してクラウドストレージ 30 に保存する。また、復号時には、粗量子化ビンインデクスと共に、対応するクラスタインデクスに基づく辞書を用いる。そのため、計算速度及び計算の質が向上する。

【0087】

ここで、本実施の形態は、次のように表現することもできる。すなわち、複数の画像に

10

20

30

40

50

ついて各画像の特徴量の分布から複数のクラスタに分類した辞書情報（辞書DB23a）と、前記第1の量子化係数から前記第2の量子化係数への再符号化時に、前記入力画像が前記複数のクラスタのいずれに該当するかを判別したクラスタインデクス311を、前記第2の量子化係数（粗量子化ビンインデクスb）に対応付けて記憶装置（ストレージ31）に格納する再符号化部21aと、をさらに備える。そして、前記逆写像部は、前記第2の量子化係数に対応付けられた前記クラスタインデクスに基づき前記辞書情報内のクラスタを特定し、当該特定したクラスタを用いて前記逆写像を行う。

【0088】

<実施形態4>

続いて、上述した実施形態1乃至3の変形例である実施形態4について説明する。図15は、本発明の実施の形態4にかかるクラウドストレージシステムの全体構成を示すブロック図である。クラウドストレージシステムには、ユーザ端末11（ユーザU1）及びユーザ端末12（ユーザU2）、クラウドレット20b、及びクラウドストレージ30が存在する。システムは、3つの主要な動作を有している：オフライントレーニング、画像アップロード、画像ダウンロードである。

【0089】

画像アップロード動作では、ユーザ端末11は、細かく（密に）量子化（以降、密量子化と呼ぶ）されたJPEG画像（符号化画像P2）をクラウドレット20bにアップロードし、そして、クラウドレット20bの再符号化部21bは、粗めの量子化（以降、粗量子化と呼ぶ）パラメータ（QP）を用いて画像のコードブロックのサブセットを再符号化し、再符号化画像P3（及び後述する副情報312）をクラウドストレージ30へ格納する。

【0090】

画像ダウンロード動作では、クラウドレット20bの粗密変換部22bは、要求された粗量子化画像（再符号化画像P3及び副情報312）をクラウドストレージ30から取り出し、粗量子化されたコードブロックを密量子化ブロックに逆写像する。

【0091】

オフライントレーニング動作では、クラウドレット20bの学習部25は、画像ダウンロード中の粗量子化ブロックから密量子化ブロックへの逆写像を容易にするために、適切な辞書23bおよびグラフテンプレート24を事前に計算し記憶する。尚、学習部25は、クラウドレット20bの外部にあっても構わない。次にこれら動作の詳細を説明する。

【0092】

1）オフライントレーニング：画像ダウンロード中の粗量子化ブロックから密量子化ブロックへの逆写像を信号復元問題とする。

近年、非特許文献9ではJPEG圧縮形式から画像ピクセルパッチを復元するために、スパース性モデル（sparsity prior）（非特許文献10）とグラフ-信号平滑モデル（graph-signal smoothness prior）（非特許文献11）の組み合わせを正常に使用した。これら2つのモデルは、それぞれピクセルパッチにおける組織的かつ構造的な画像コンテキストの回復を容易にする。

【0093】

これら2つのモデルを使用することを仮定すると、高速実施は、画像の復元中に小辞書と散在的な（sparse）グラフを必要とする。この目標にむけて、我々は、まず、ピクセルブロック Ψ を大きなトレーニングセット Ψ_0 からツリーのエンドノードに対応する類似のブロックの異なるクラスタに分類するために、ツリー構造のベクトル量子化（Tree-Structured Vector Quantization：TSVQ）（非特許文献12）を構築する。各クラスタ i に対して、このクラスタに関連するトレーニングデータ Ψ_i に基づいて、過完備辞書および散在グラフテンプレートを学習させる。分類ツリーを構築すると、重要なアイデアは、画像のダウンロード中に、元のブロックが分類されるクラスタを正しく識別できる場合、関連する辞書およびクラスタに関連付けられたグラフが、粗ブロックから密ブロックに逆写像することを容易にすることが出来ることである。

10

20

30

40

50

【0094】

2) 画像アップロード：ユーザ端末11は、J P E G圧縮画像（符号化画像P2）をクラウドレット20bにアップロードする。ここで、N-ピクセルコードブロックxのそれぞれは、精密なQPq = [q₁, . . . , q_N] で与えられるN量子化ビンインデクスa(x) = [a₁(x), . . . , a_N(x)] の集合により表される。詳細には、i番目の量子化ビン計数a_i(x)は、以下ようになる。

【数15】

$$a_i(x) = \text{round}\left(\frac{T_{i,:}x}{q_i}\right)$$

10

. . . (15)

ここで、T_iはDCT変換行列Tのi行目である。

【0095】

そして、クラウドレット20bの再符号化部21bは、粗いQP Q = [Q₁, . . . , Q_N] を用いて、符号化画像P2におけるブロックのサブセットを再符号化し、精密なビンインデクスa(x)を、クラウドストレージ30に恒久的に格納されているコードブックxに対する粗いビンインデクスb(x)へマッピングする。クラウドレット20bは、画像アップロード中、再符号化されたブロックに対して、信号モデルを用いた粗いb(x)から精細なa(x)へ戻すための逆写像を実施する。再符号化部21bは、「正常に」逆写像することができた画像におけるブロックxのみを、アップロード中に逆写像する。小さなバイナリ画像は、デコーダにどのブロックがQを用いて粗く再符号化されたかについての情報を与える副情報(side information: S I) 312としてJ B I G（非特許文献8）を用いて、損失無く符号化される。

20

【0096】

本システムは、J P E G画像における再符号化されたブロックが、2つのモードのいずれかにおいて正常に逆写像されることを保証する。i) 無損失復元（可逆復元、lossless recovery）（無損失モード）：正確に復元されるために、各再符号化されたブロックxに対する元のa(x)が確定的に保証されることを意味する。または、ii) 近無損失復元（near-lossless recovery）（近無損失モード）：高い確率で正確に復元されるために、元の画像に対するa(x)は、統計的に保証されることを意味する。これは、デコードされたJ P E G画像が、ユーザにとって視覚的に区別が出来ない元の画像に近いことをもたらす。無損失モードにおいて、再符号化されたブロックxのそれぞれに対するa(x)の完全な復元のために使用される適切な辞書およびグラフテンプレートを保証するために、TSVQにおけるxに対するクラスティンデクスも、アップロード中に、提案する分散型符号化方法を用いて、S Iとして符号化される。

30

【0097】

3) 画像ダウンロード：ユーザU2が事前にアップロードされた画像を呼び戻すとき、クラウドレット20bは、クラウドストレージ30から粗量子化された画像（再符号化画像P3）を取り出し、ユーザU2に対してブロックxの粗量子化されたb(x)のそれぞれを密量子化されたa(x)に逆写像する。画像取り出し遅延を最小化するために、逆写像は、画像アップロード中にS Iとして符号化されたクラスティンデクスにより示される適切な小型の辞書および散在的なグラフテンプレートをを用いて高速に行われるべきである。

40

【0098】

ここで、量子化ビンマッチング問題を、粗量子化ビン内で、まず、（図4のx3として示されている）最も可能性の高い信号を探すことに書き直すことができる。

【数 1 6】

$$X^o = \arg \max_X P(X | b)$$

・・・ (16)

そして、最も可能性の高い信号 X_0 を含む密量子化ビンベクトル a_0 を特定する。

【数 1 7】

$$a_i^o = \text{round}\left(\frac{X_i^o}{q_i}\right), \forall i$$

10

・・・ (17)

次に、粗いビンベクトル b 内で、最も可能性の高い信号 X_0 を見つけるという問題に焦点を当てる。

【0099】

ここで、2つの信号モデルを用いて粗いビンインデクスベクトル b が与えられる最も可能性の高い信号 X^h を見つけることに焦点をあてる。尚、見つけられた最も可能性の高い信号を含む密量子化ビンは、復元された密ビンインデクスである。また、2つの信号モデルとは、スパース性モデルと、グラフ - 信号平滑モデルである。さらに、高速実施を達成するために、我々は、各クラスタに対する小さな辞書および散在されたグラフテンプレート

20

【0100】

二つのモデル $P_S(X)$ 及び $P_G(X)$ を結合することによって、 $P(X)$ を以下のよう

【数 1 8】

$$P(X) = P_S(X)P_G(X) \propto \exp\left(-\lambda_1 \|\alpha\|_0 - \lambda_2 X^T \hat{L} X\right)$$

・・・ (18)

インデクスが付けられた粗量子化セル b 内の最も可能性のある信号を見つけ出す問題は以下のように定式化することができる。

30

【数 1 9】

$$\{\mathbf{X}^*, \alpha\} = \arg \min_{\{\mathbf{X}, \alpha\}} \|\mathbf{X} - \mathbf{T}\Phi\alpha\|_2^2 + \lambda_1 \|\alpha\|_0 + \lambda_2 \mathbf{X}^T \hat{\mathbf{L}} \mathbf{X},$$

$$\text{s.t., } X_i \in I(b_i, Q_i), i = 1, \dots, N.$$

・・・ (19)

目的関数は、交互最適化を介して解くことができる。

【0101】

図 16 は、本発明の実施の形態 4 にかかるオフライントレーニングの流れを説明するための図である。学習部 25 は、トレーニング画像 PG からトレーニングブロック集合 BS (ブロック画像) を抽出 (収集) する (S401)。トレーニング画像 PG は、例えば、多数の高解像度の画像群であるとよい。次に、学習部 25 は、トレーニングブロック集合 BS について、 $TSVQ$ により複数のクラスタに分類してツリー構造 (二分木) TS を構築する (S402)。そして、学習部 25 は、分類された各クラスタに対してクラスティンデクスのオフセットが減少するように、ツリー構造 (二分木) TS のノードを入れ替える (S403)。尚、当該ノードの入れ替えは、後述するツリーノードのビット割り当てに対応する。また、学習部 25 は、入れ替え後の二分木に対してレート歪みを最適化するように枝刈りを行う (S404)。尚、当該枝刈りは、後述するレート歪み最適化ツリープルーニングに対応する。その後、学習部 25 は、枝刈り後の二分木に対して過完備辞書

40

50

と散在グラフテンプレートとを学習し（S 4 0 5 及び S 4 0 6）、辞書 D B 2 3 b 及びグラフテンプレート 2 4 として記憶装置（不図示）に保存する。尚、当該記憶装置は、クラウドレット 2 0 b 内又は外であって構わない。尚、完全な TSVQ は、デコーダにおいて必要とされ、フルツリーのインデクスは、各ブロックに対して計算される。

【0 1 0 2】

（適応性のある辞書及びグラフテンプレート学習）

多数のオフライントレーニングデータを考慮して、類似したピクセルブロックの異なったクラスタに対して辞書及びグラフテンプレートをトレーニングする。各クラスタに対し、そのクラスタに分類されたトレーニングデータのピクセルブロックに基づいて、過完備辞書及び 1 つの散在グラフテンプレートを学習させる。

10

【0 1 0 3】

1) 小辞書学習：特定のクラスタに対して、割り当てられた n トレーニングブロック $Y = [y_1, y_2, \dots, y_n]$ が存在する。ここで、 y_i は、ベクトル形式内のトレーニングブロック i を示す。K-SVD（非特許文献 1 0）に従い、以下の目的関数を最小化することにより辞書 Φ を計算する。

【数 2 0】

$$\min_{\Phi, \alpha_i} \|\alpha_i\|_0, s.t., \|y_i - \Phi \alpha_i\|_2^2 \leq T, \forall i$$

・・・ (20)

20

【0 1 0 4】

2) 散在グラフテンプレート学習：同様に、特定のクラスタについて散在グラフを学習する。各クラスタに対し、まず、重心 (Centroid) ブロックを計算する。重心ブロックは該クラスタに割り当てられたトレーニングブロックの中心又は平均である。そして、この重心ブロックに対して散在グラフテンプレートを構築する。図 1 7 は、本発明の実施の形態 4 にかかる散在グラフテンプレートの例を示す図である。図 1 7 に示すように、縦、横、斜めの隣接ノードのみが接続される。その理由は、グラフがスパースであり、 L がスパース行列であるならば、 $I + \lambda_2 L$ の逆行列の高速解法を採用でき、式 (2 1) の計算複雑性を低くすることができるためである。また、各テンプレートが接続されたグラフだからである。

30

【数 2 1】

$$\min_X \|X - T\Phi\alpha^*\|_2^2 + \lambda_2 X^T \hat{L} X$$

・・・ (21)

【0 1 0 5】

（完全な T S V Q 構築 (Full TSVQ Construction)）

トレーニングデータセット Ψ_0 を仮定すると、我々は、まず、目標となる高さ H_{max} のバイナリフルツリー (a binary full tree、全二分木) T^0 を構築する (図 1 8 (a))。具体的には、フルツリーの作成は、以下のステップを有している (非特許文献 1 2)：

40

【0 1 0 6】

1) 初期化：ツリーの高さ $H = 0$ に初期化する。我々は、トレーニングセット Ψ_0 全体を、関連するトレーニングセットの平均として計算される、重心 c_0 と共に、高さ 0 のルートノード r に関連付ける。

【0 1 0 7】

2) バイナリ分割：高さ H の各ノード i に対して、我々は、2 つの子ノード j および k を生成する。ここで、 $i \rightarrow^0 j$ 、 $i \rightarrow^1 k$ である。我々は、それらの重心 c_j および c_k に関する総距離が最小化されるように、ノード i に関連するトレーニングセット Ψ_i を 2 つの重複しないサブセット Ψ_j および Ψ_k に分割する。

50

【数 2 2】

$$\min_{c_j, c_k, \Psi_j, \Psi_k | \Psi_j \cup \Psi_k = \Psi_i} \sum_{\psi \in \Psi_j} \|\psi - c_j\|_2^2 + \sum_{\psi \in \Psi_k} \|\psi - c_k\|_2^2$$

・・・ (22)

【0108】

ロイドアルゴリズム (Lloyd Algorithm) を用いて、局所的に最適に式 (22) を解くことができる。局所収束性 (local convergence) まで交互に実行される 2 つのステップから構成される。

10

【0109】

・固定された重心 c_j と c_k とが与えられた場合、各 $\psi \in \Psi_i$ をより近い重心とする分割 (partition) に Ψ を割り当てることにより、局所的に最適な分割 Ψ_j と Ψ_k とを求める。

・固定された分割 Ψ_j および Ψ_k が与えられた場合、それぞれの二乗誤差を最小化することにより、重心 c_j および c_k を更新する。 c_j に対して以下の計算をする：

【数 2 3】

$$\min_{c_j} \sum_{\psi \in \Psi_j} \|\psi - c_j\|_2^2$$

20

・・・ (23)

これは、 Ψ_j の平均値をとることにより、クラスタの Ψ_j の「中心」として c_j を計算することを意味する。

【0110】

交互におこるロイドアルゴリズムの開始に先立って、非特許文献 13 に従って、高さ H のノード i に対する 2 つの子の重心 c_j および c_k を c_i および摂動バージョン (perturbed version) として初期化する。すなわち、

30

【数 2 4】

$$c_j = c_i, c_k = c_i + \eta$$

・・・ (24)

ここで、 η は、ノイズ分散 $\sigma = 1$ とするガウス雑音である。

【0111】

3) より大きなバイナリフルツリー (Bigger Binary Full Tree) の生成：構成されたバイナリフルツリーの高さ H が目標値 H_{max} よりも小さい場合、 H を 1 インクリメントし、ステップ 2 に進む。

【0112】

40

(クラスタインデクスの分散符号化)

無損失モードにおいて、画像ダウンロード中の再符号されたブロック x の粗いインデクス $b(x)$ を精密なインデクス $a(x)$ に逆写像するために適切な辞書およびグラフテンプレートを特定するために、「正確な」クラスタインデクス i を画像アップロード中に SI として決定し符号化している。具体的には、精密なインデクス $a(x)$ の入力を使用して、まずハード復号化された (hard-decoded) ブロック x^h を構築し、これを用いてルートノード r から高さ H_{max} のエンドノード e^0 までフルツリー T^0 を走査する (traverse)。走査によって、ノード i において、 x^h とそれぞれの子ノード j および k の 2 つの重心 c_j および c_k と比較し、より近い重心を有するノードに進む。

【0113】

50

エンドノード e^0 に到達すると、関連する文字列 S_{e^0} を X として特定する。（後述する）クラスタとして使用する実際に最適化された VLT (Variable-Length Tree) T (図 18 (b)) は、典型的にフルツリー T^0 のサブセットであるため、 x^h を使用する代わりに、 T を走査することは、 T^0 における中間ノードである T におけるエンドノードで終端されることがあり、この場合、 S_e は S_{e^0} の部分文字列 (substring) である。 S_{e^0} の部分文字列である S_e を $Q_T(X)$ と示す。2つのエンドノード T^0 および T の文字列の例を、図 18 (c) に示す。

【0114】

符号化コストを軽減するために、直接符号化 $Q_T(X)$ の代わりに、粗いインデクス $b(x)$ でハード復号されたブロック x^h を使用して T^0 を走査することにより、画像ダウンロード中に取得される文字列である、 X とそのノイズ観測 Y との間の相関を利用する、効率的な分散符号化方式を提案する。基本的な考えは、より多くの観測／目標文字列のペア (y, x) に対して、 $Q_T(y + \Delta) = Q_T(X)$ となるような、 SI として適切な差分 Δ_j を符号化するために、コードブック C における差分コードワード Δ_j を設計することである。差分 Δ_j が高い確率 $p(\Delta_j)$ で選択される場合、算術符号化 (arithmetic coding) を使用する (近似された) コードレート $-\log_p(\Delta_j)$ は低い。

【0115】

$Q_T(y + \Delta_j) = Q_T(X)$ となるような、差分 $\Delta_j \in C$ が存在しない場合、直接インデクスコードワード d_i を符号化する。ここで、 i は、ブロック x^h のクラスタインデクスである。直接インデクスコードワードは、各クラスタに対してユニーク (固有) であることから、それらの確率は典型的に低く、符号化コストはより高い。

【0116】

(無損失モードのためのツリー構造ベクトル量子化の最適化)

ここで、 $TSVQ$ 最適化問題を定義して、それを最適化する方法を議論する。2つの最適化変数 VLT T とコードブック C とがあることを考慮して、交互最適化方法を提案する。

(A) C を固定して、 T を最適化。これを最適 VLT 設計問題と呼ぶ。

(B) T を固定して C を最適化。これを最適差分コードブック問題と呼ぶ。

以下では、順番に、これら2つの問題を解決するための詳細なアルゴリズムを説明する。

【0117】

(A) 最適 VLT 設計

C を固定すると、最適化問題は以下のようにになる。

【数25】

$$\begin{aligned} & \min_T \sum_{i \in T} P(i) L_i + \lambda \sum_{i \in T} P(i) \sum_y P(y|i) \min_{\Delta \in C | M_{h_i}(y+\Delta) = s_i} -\log p(\Delta) \\ &= \min_T \sum_{i \in T} P(i) \left(L_i + \lambda \sum_y P(y|i) \min_{\Delta \in C | M_{h_i}(y+\Delta) = s_i} -\log p(\Delta) \right) \end{aligned}$$

・・・ (25)

【0118】

上記の目的関数は、 VLT T の RD コストが、それぞれが確率 $P(i)$ で重み付けされたエンドノード i の個々の RD コストの合計であることを示している。式 (25) を最小にする2つの手順を説明する。

【0119】

まず、式 (25) において、異なったエンドノード i に対する、より有利な誤差分布 $P(y|i)$ (例えば、エンドノード i に対して、より小さい比率の結果となるであろう、 $P(s_i|i)$ を中心に分布する非対称分布) を誘導するために、 VLT T におけるノードにバイナリビット (0 および 1) を再割り当てする手順 (1) を提案する。

【0120】

次に、VLT Tへの所定のビットの割り当てに対して、初期のツリーTからエンドノードを最適に取り除く、RD-最適化ツリープルーニング手順(2)を提案する。

これら2つの手順は、収束するまで交互に行われる。続いて、これら2つの手順について、説明する。

【0121】

(1) ツリーノードのビット割り当て：密量子化ブロックxに対して、その確率を以下のように定義する。

【数26】

$$P(\mathbf{x}) = P(i)P_i(\mathbf{x})$$

10

・・・(26)

ここで、P(i)は、エンドノードiの確率であり、P_i(x)は、クラスタiに属するブロックによって広がる(spanned)部分空間(subspace)S_i内のxの確率である。粗量子化ブロックyは、xのノイズバージョン(noisy version)とみなすことができ、以下の様にモデル化される。

【数27】

$$y = x + z$$

20

・・・(27)

ここで、zは、ノイズである。

【0122】

以下の最適化問題を解決することにより、yの最も可能性の高いクラスティンデクスj*を見出す。

【数28】

$$\begin{aligned} j^* &= \arg \max_j P(y = j|i) \\ &= \arg \max_j P(\mathbf{y} \in S_j | \mathbf{x} \in S_i) \end{aligned}$$

30

・・・(28)

ノイズモデルによれば、上記の最適化は以下のように書き換えることが出来る。

【数29】

$$\begin{aligned} j^* &= \arg \max_j P(\mathbf{x} + \mathbf{z} \in S_j | \mathbf{x} \in S_i) \\ &= \arg \max_j \sum_{\mathbf{x} \in S_i} P(\mathbf{x} + \mathbf{z} \in S_j | \mathbf{x}) P_i(\mathbf{x}) \\ &= \arg \max_j \sum_{\mathbf{x} \in S_i} \left(\sum_{\mathbf{z} | \mathbf{x} + \mathbf{z} \in S_j} P(\mathbf{z}) \right) P_i(\mathbf{x}) \end{aligned}$$

40

50

・・・ (29)

【0123】

しかし、 x および z の両方は、高次元のベクトル (high-dimensional vectors) であり、式 (29) を解くことは自明ではない問題であることを意味する。代わりに、 S_i の重心である C_i と最も類似した重心を持つように、 y の最も可能性の高い部分空間を見つける。

【数30】

$$j^* = \arg \min_j |C_i - C_j|, j = 1, \dots, m$$

10

・・・ (30)

ここで、 m は y の部分空間の候補の数である。

【0124】

y を求めた後、インデクス j^* であるノードがノード i と近くなることを強制するために、ツリーノードのバイナリビットを再割り当てする。図19および図20に例を示す。図19は、部分空間の検索を説明するための図である。密量子化ブロック x は、部分空間 S_1 に配置される。ノイズモデルによれば、可能性の高いノイズ観測 y が太字破線で示される。式 (29) を処理することにより、 y の最も可能性の高い部分空間のインデクスが検索される。図20は、二分木のノードの入れ替えを説明するための図である。 x が S_1 に配置され、検索された最も可能性の高い y は S_3 に配置されたと仮定する。インデクスオフセットは2である。 S_1 及び S_3 を互いに近付けるためにツリーブランチに正しいバイナリビットを再割り当てする。インデクスオフセットは、1に減少する。このように、クラスタインデクスのオフセットは減少する。図21は、ビット割当前後のインデクスオフセットの確率分布の比較を示す図である。図21に示すように、インデクスオフセットの確率は、0の周辺により集中し、コードブック C の伝送コストがより小さくなる。したがって、エンドノード i に対して、より小さい比率が達成される。

20

【0125】

(2) レート歪み最適化ツリープルーニング (Rate-Distortion Optimized Tree Pruning) : ビットの再割り当て後、初期ツリー T からエンドノードを最適に除去するために、 RD -最適化ツリープルーニング手順が行われる。より多くのエンドノードを除去することは、歪みがより大きく、より短いツリーとなるが、比率はより小さくなる。我々が行うことは、ツリーの深さを変えることにより、歪みと比率の間で最良のトレードオフを達成することである。レート歪みフレームワーク (rate-distortion framework) を利用したプルーントツリーアルゴリズム (prune-tree algorithms) がいくつか存在する。ここでは、我々のシステムでは、子を刈り取る (pruning) という手法を採用する (非特許文献14)。但し、これに限定されない。

30

【0126】

所定のツリー T' に対して、有限な数の刈り取られたサブツリー (pruned subtrees) $\{T\}$ が存在する。我々は、式 (25) の最小値につながるものを見つけるために、それらを網羅的に検索する。具体的には、2つの兄弟エンドノード (brother end nodes) を刈り取り、それらの親ノードを新しいエンドノードとして、新しい VLT T を得る。 T に対して、すべてのエンドノードの歪みと比率 (rate) とを合計して RD コストを計算する。

40

【数31】

$$RD(T) =$$

$$\sum_{i \in T} P(i) \left(L_i + \lambda \sum_y P(y|i) \min_{\Delta \in C | M_{h_i}(y+\Delta)=s_i} - \log p(\Delta) \right)$$

・・・ (31)

【0127】

50

この処理が、左から右にエンドノードの全てのペアに対して、再帰的に行われる。最終的に、最適なVLTツリーを次のように取得する：

【数 3 2】

$$\mathcal{T}^* = \arg \min_{\mathcal{T}} RD(\mathcal{T})$$

・・・ (3 2)

【0 1 2 8】

(B) 最適なコードブック設計

10

1) 分散符号化手法：ここで、VLT Tが与えられた最適なコードブックCを設計する問題に焦点を当てる。統計 $P(y|x)$ は固定であり、既知であると仮定する。実行可能な(feasible)コードブックCは、 $P(y|x) > 0$ であるような観測／目標ペア (y, x) について、i) $Q_T(y + \Delta_j) = Q_T(x)$ であるような差分コードワード $\Delta_j \in C$ を選択することができ、ii) $s_i = Q_T(x)$ であるような直接インデクスコードワード c_i を選択することができる、ことを意味している。最適なコードブックCは、全てのペア (y, x) について、上述した実行可能な選択が、予想された符号化コストを最小に至ることを意味する。

【0 1 2 9】

まず、典型的な差分符号化と異なり、各観測／目標ペア (y, x) に対する正確な差分 $\Delta = x - y$ を符号化するために、エンコーダを必要としない； $Q_T(y + \Delta_j) = Q_T(x)$ であるような任意の実行可能な差分コードワード Δ_j は、同じように受け入れられる。我々は、 $r(y, x)$ により、実行可能な差分の範囲を示す。VLT Tにおける量子化ビン $i = Q_T(x)$ の下限と上限を $[l_i, u_i)$ により示す。明らかに、量子化ビン i のサイズ $u_i - l_i$ が大きくなれば、より大きな範囲 $r(y, x)$ となる。

20

【0 1 3 0】

コードブックを設計するとき、直感的に、小さいコードブックCが望ましい；確率は、少数のコードワード Δ_j または c_i のみに集中し、そのため、符号化コスト $-\log p(\Delta_j)$ または $-\log p(c_i)$ は小さくなる。上述したように、ペア (y, x) に対するクラスタインデクスを符号化するとき、範囲 $r(y, x)$ により提供される自由度が与えられると、我々の設計方法は、以下の様になる。まず、より多くの範囲 $r(y, x)$ をカバーする少数の差分 Δ_j のみを選択する。そして、残りのペア (y, x) に対して、実行可能性を維持するために直接インデクスコードワード c_i を使用する。次に、この処理の詳細を説明する。

30

【0 1 3 1】

2) コードブック設計アルゴリズム：まず、実行可能性を保証するために、VLT Tにおけるエンドノード i に対応する、全ての直接インデクスコードワード c_i は、コードブックCに含まれていると仮定する。そして、差分コードワード Δ_j を次のように追加する。まず、 $P(y|x)$ は、高くなる傾向があるため、最も一般的に使用されるコードワードである、初期(デフォルト)差分 $\Delta_0 = 0$ を追加する。

40

【0 1 3 2】

観測／目標ペア (y, x) のそれぞれに対して、範囲 $r(y, x)$ が現在のコードブックCにおける既存の差分 Δ_j と重複していないかを確認する。重複していない場合、図22に示すように、 $P(y|x)$ で縮尺された $r(y, x)$ を差分カウンタ $H(\Delta)$ に加える。既存の差分コードワードと重複しないすべてのペア (y, x) からの範囲を累積した後、ピーク $\Delta^* = \max_{\Delta} H(\Delta)$ をCに加える次の候補として特定する。

【0 1 3 3】

次に、差分コードワードの候補 Δ^* がCに追加された場合の予測される符号化コストを計算する。各ペア (y, x) に対して、0に近い範囲 $r(y, x)$ において、コードワード $\Delta_j \in C$ を検索する。 $r(y, x)$ 内で $\Delta_j \in C$ が無い場合、このペアに対して直接イ

50

ンデクス (direct index) c_i を使用する。ここで、 $i = Q(x)$ である。全てのペア (y, x) のコードワードへのマッピングが与えられると、各コードワードに対する確率を一覧化することができ、式 (33) を使用して比率を計算することが出来る。

【数 33】

$$R(C) = - \sum_{\Delta \in C} p(\Delta) \log p(\Delta) - \sum_{c \in C} p(c) \log p(c)$$

・・・ (33)

得られた比率が減少した場合のみ、この候補 Δ^* は、最終的に追加される。他の差分コードワードの候補を追加を試みるために、比率が減少した場合、手順は繰り返される。アルゴリズムの擬似コードをアルゴリズム 1 (図 23) に示す。 10

【0134】

(統計的保証を伴う近無損失モード (Near-Lossless Mode))

続いて、近無損失モードについて説明する。近無損失モードは、各ブロック x に対する粗いインデクス $b(x)$ から復元することが出来る精密なビンインデクス $a(x)$ を決定論的に保証する無損失モードと異なり、高い確率で $b(x)$ から $a(x)$ を復元することが出来るという統計的な保証を有している。復元要件を下げることで、近無損失モードは、はるかに低い計算量を達成することができる。

【0135】

具体的には、オフライントレーニングにおいて、完全な TSVQ 構築の後、トレーニングブロックを異なるクラスタに分類する。同一クラスタのブロックは、類似した構造を共有する。各クラスタに対して、このクラスタに関連する辞書およびグラフテンプレートを使用して、精密なインデクス $a(x)$ が粗いインデクス $b(x)$ から正確に逆写像されるブロックの割合を計算する。完全な画像に対する統計的な保証を保証するために、各クラスタの復元成功率が使用される。

【0136】

画像アップロードでは、各コードブロック x_i に対して、オフライン統計により、関連した復元成功確率 $P(x_i)$ が存在する。無損失復元とは異なり、どのブロックが再符号化されるかを決定するために、量子化ビンマッチングの最適化を実行するのではなく、事前に設定した閾値 P_T よりも平均復元成功確率が高いものを選択する。より具体的に、列にそって全てのコードブロックを詳しく調べ (スキャンし)、再符号化するコードブロックを徐々に決定する。コードブロック x_i に来ると、調べたブロック $\{x_j\}_{j=1}^i$ の平均復元成功確率 AP_i を以下の様に計算する。 30

【数 34】

$$AP_i = \frac{\sum_{j=1}^i P(x_j)}{i}$$

・・・ (34)

$AP_i \geq P_T$ である場合、 x_i を再符号化のサブセットに追加する。この決定処理は、ブロック毎に実際の逆写像を実行しないことから高速である。

【0137】

QBMのための辞書とグラフテンプレートを選択し、フルツリーを走査するために、画像ダウンロードでは、粗量子化ブロックを使用する。無損失復元において、行われたように、クラスタインデクスを符号化するための追加的なコストを払う必要がなく、そのため、より多くの記憶領域を節約することが出来る。 50

【0138】

ここで、本実施の形態は、次のように表現することもできる。すなわち、画像処理装置（クラウドレット20b）は、複数の画像の特徴量の分布から複数のクラスタに分類した過完備辞書（over-complete dictionary）（辞書DB23b）と、前記特徴量の分布をグラフ化し、前記複数のクラスタに分類した散在（sparse）グラフテンプレート（グラフテンプレート24）と、前記複数の画像から前記過完備辞書と前記散在グラフテンプレートとを学習する学習部25と、入力画像を第1の量子化幅に基づき量子化した第1の量子化係数を含む符号化画像について、前記第1の量子化係数から、前記第1の量子化幅より広い第2の量子化幅に基づき量子化した第2の量子化係数へ再符号化する再符号化部21bと、前記第2の量子化係数について、前記過完備辞書と前記散在グラフテンプレートとを用いて、前記第1の量子化係数への逆写像を行う逆写像部（粗密変換部22b）と、を備える。そして、学習部25は、前記複数の画像から抽出された複数のブロック画像について、TSVQ（Tree-Structured Vector Quantization）により複数のクラスタに分類して二分木を構築し、前記分類された各クラスタに対してクラスタインデックスのオフセットが減少するように、前記二分木のノードを入れ替え、前記入れ替え後の前記二分木に対してレート歪みを最適化するように枝刈りを行い、前記枝刈り後の前記二分木に対して前記過完備辞書と前記散在グラフテンプレートとを学習する。再符号化部21bは、再符号化時に、前記入力画像が前記学習後の前記過完備辞書及び前記散在グラフテンプレートのクラスタのいずれに該当するかを判別したクラスタインデックスを、前記第2の量子化係数に対応付けて記憶装置（クラウドストレージ30）に格納する。逆写像部は、前記第2の量子化係数に対応付けられた前記クラスタインデックスに基づき前記過完備辞書及び前記散在グラフテンプレート内のクラスタを特定し、当該特定したクラスタを用いて前記逆写像を行う。これにより、辞書とグラフテンプレートの精度を向上し、逆写像の精度も向上させることができる。

【0139】

<その他の発明の実施の形態>

尚、上記説明では、画像圧縮手法としてJPEGの例を挙げたため、変換係数としてDCT係数としていた。但し、本発明は、DCTの代わりに、離散サイン変換（Discrete Sine Transform (DST)）、又は、非対称離散サイン変換（Asymmetric Discrete Sine Transform (ADST)）等のような他のブロックベースの変換符号化及び復号化装置に適用可能である。

【0140】

さらに、本発明は上述した実施の形態のみに限定されるものではなく、既に述べた本発明の要旨を逸脱しない範囲において種々の変更が可能であることは勿論である。例えば、上述の実施の形態では、本発明をハードウェアの構成として説明したが、本発明は、これに限定されるものではない。本発明は、任意の処理を、CPU（Central Processing Unit）にコンピュータプログラムを実行させることにより実現することも可能である。この場合、コンピュータプログラムは、様々なタイプの非一時的なコンピュータ可読媒体（non-transitory computer readable medium）を用いて格納され、コンピュータに供給することができる。

【0141】

非一時的なコンピュータ可読媒体は、様々なタイプの実体のある記録媒体（tangible storage medium）を含む。非一時的なコンピュータ可読媒体の例は、磁気記録媒体（例えばフレキシブルディスク、磁気テープ、ハードディスクドライブ）、光磁気記録媒体（例えば光磁気ディスク）、CD-ROM（Read Only Memory）、CD-R、CD-R/W、DVD（Digital Versatile Disc）、BD（Blu-ray（登録商標）Disc）、半導体メモリ（例えば、マスクROM、PROM（Programmable ROM）、EPROM（Erasable PROM）、フラッシュROM、RAM（Random Access Memory））を含む。また、コンピュータプログラムは、様々なタイプの一時的なコンピュータ可読媒体（transitory computer readable medium）によってコンピュータに供給されてもよい。一時的なコンピュータ可読媒体

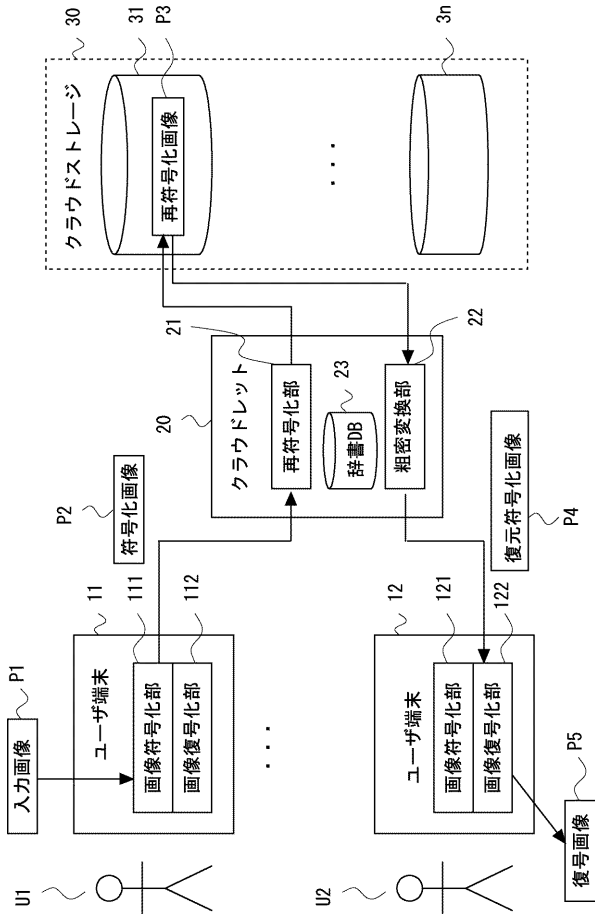
の例は、電気信号、光信号、及び電磁波を含む。一時的なコンピュータ可読媒体は、電線及び光ファイバ等の有線通信路、又は無線通信路を介して、プログラムをコンピュータに供給できる。

【符号の説明】

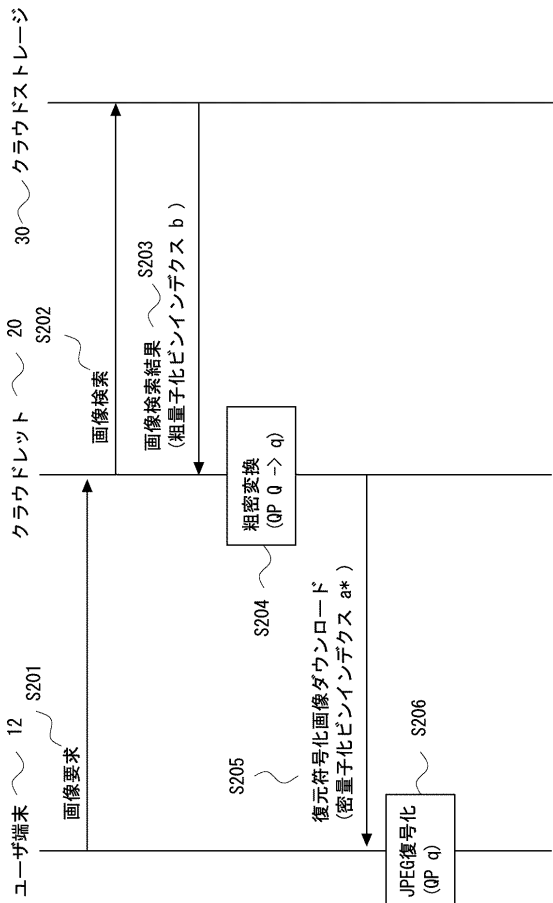
【0142】

U 1	ユーザ	
1 1	ユーザ端末	
1 1 1	画像符号化部	
1 1 2	画像復号化部	
U 2	ユーザ	10
1 2	ユーザ端末	
1 2 1	画像符号化部	
1 2 2	画像復号化部	
2 0	クラウドレット	
2 1	再符号化部	
2 2	粗密変換部	
2 3	辞書 D B	
2 0 a	クラウドレット	
2 1 a	再符号化部	
2 2 a	粗密変換部	20
2 3 a	辞書 D B	
3 0	クラウドストレージ	
3 1	ストレージ	
3 n	ストレージ	
3 1 1	クラスティンデクス	
P 1	入力画像	
P 2	符号化画像	
P 3	再符号化画像	
P 4	復元符号化画像	
P 5	復号画像	30
2 0 b	クラウドレット	
2 1 b	再符号化部	
2 2 b	粗密変換部	
2 3 b	辞書 D B	
2 4	グラフテンプレート	
2 5	学習部	
3 1 2	副情報	
P G	トレーニング画像	
B S	トレーニングブロック集合	
T S	ツリー構造（二分木）	40

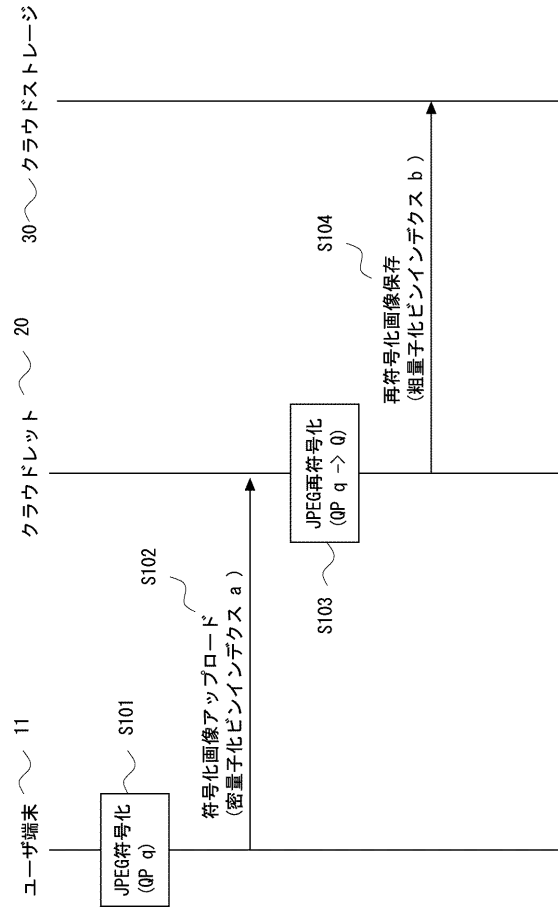
【図 1】



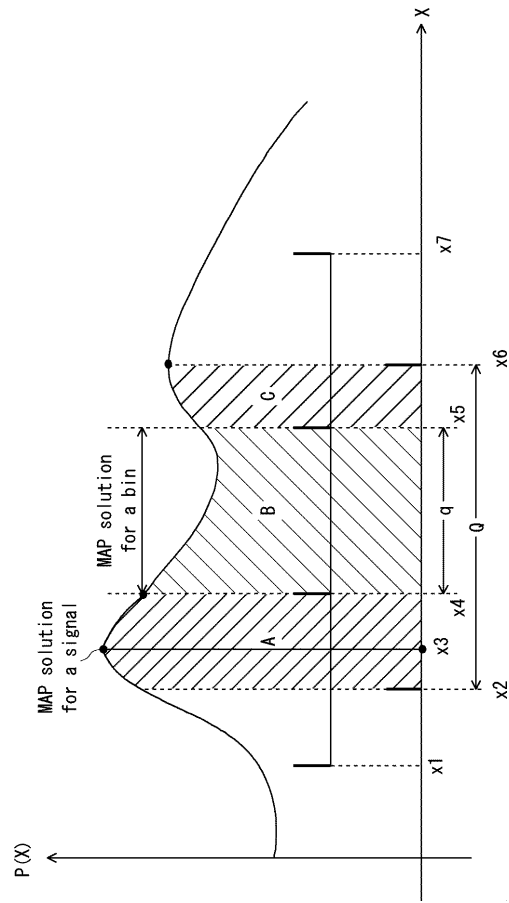
【 図 3 】



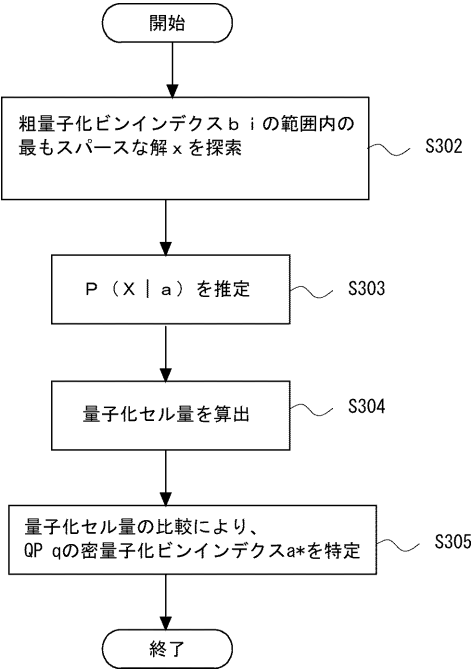
【図 2】



【図 4】



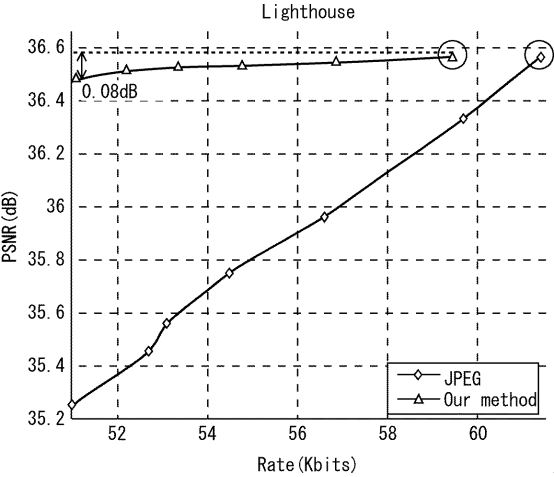
【図 5】



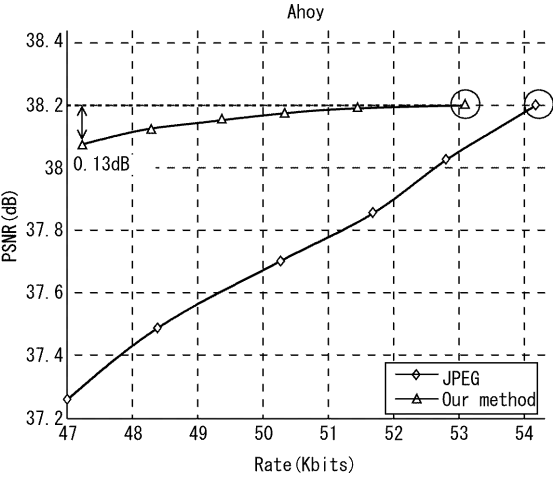
【図 6】

Images	Fine QF	Coarse QF	Blocks Proportion	Bit-Saving
Lighthouse	80	50	10.97%	3.32%
Sailboats	80	50	10.09%	2.08%
Window	80	50	12.81%	3.04%
Girl	80	50	14.71%	4.54%
Hats	80	50	15.67%	4.38%
Parrots	80	55	17.37%	3.68%
Airplane	80	55	42.18%	14.19%
Ahoy	80	50	7.96%	2.01%

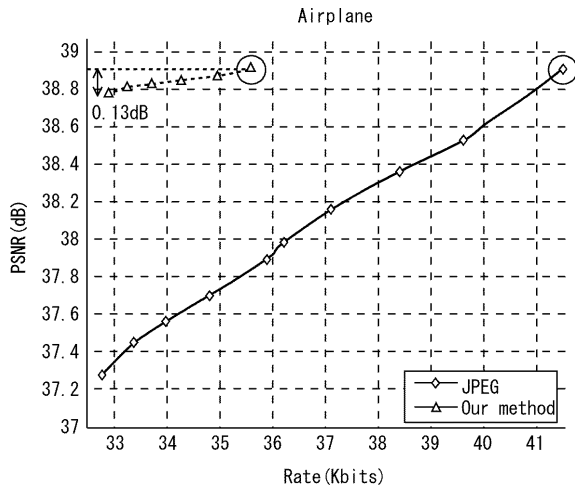
【図 7】



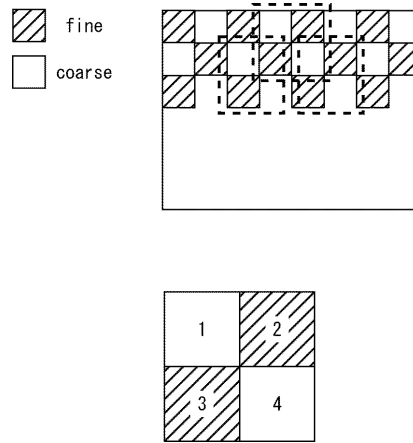
【図 8】



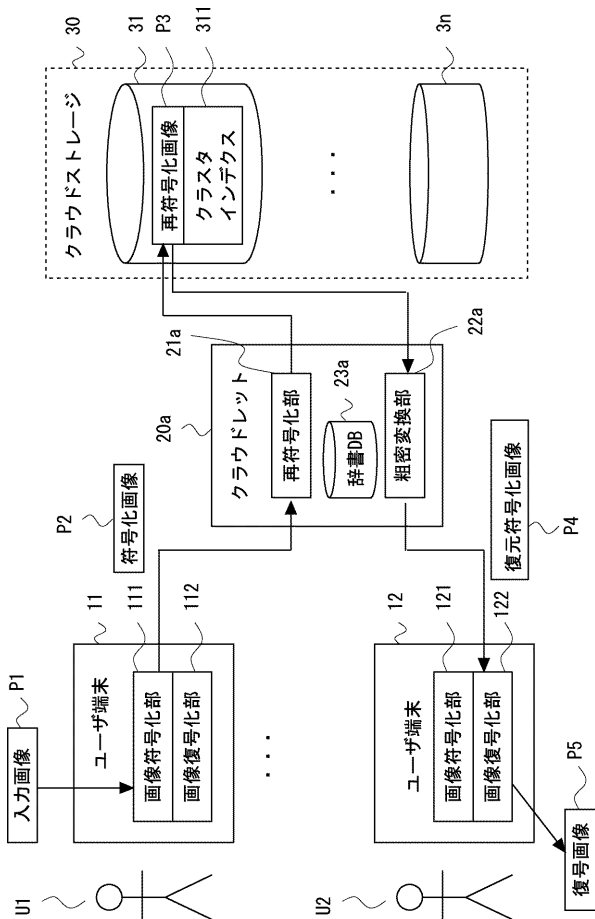
【図 9】



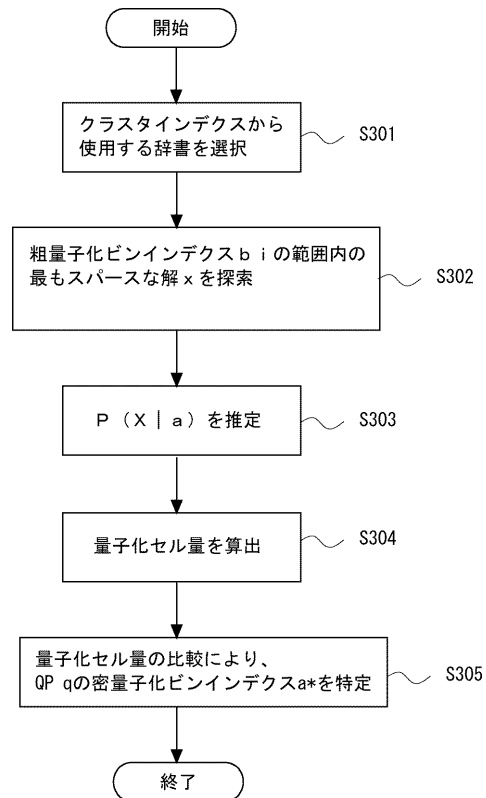
【図 10】



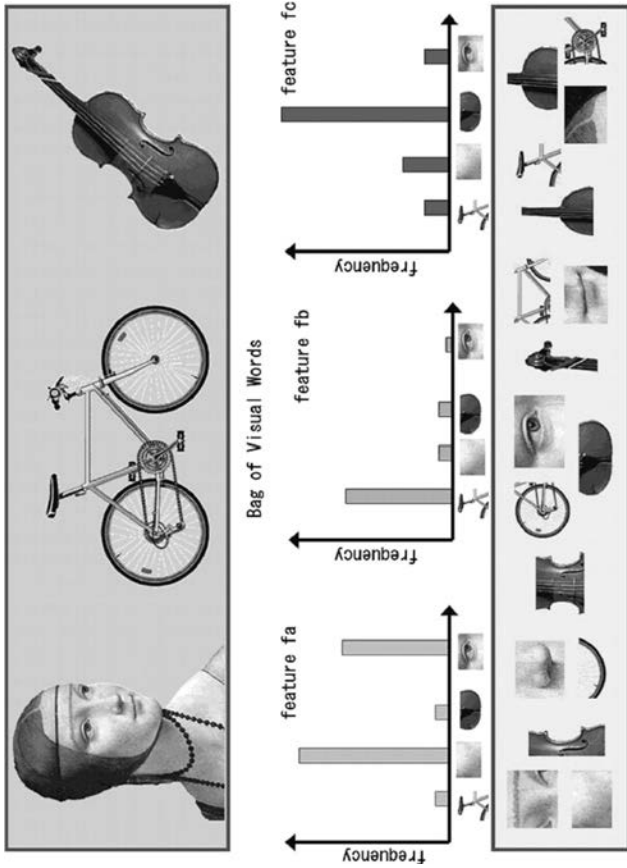
【図 11】



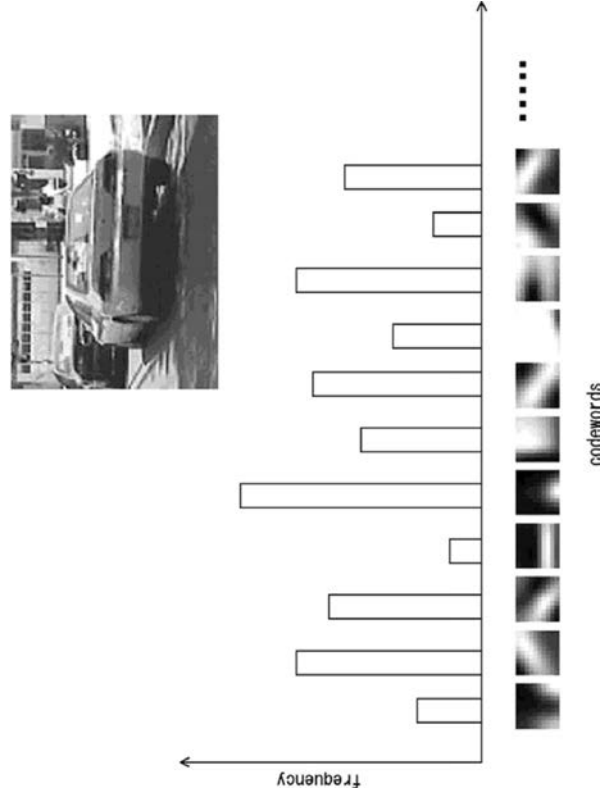
【図 12】



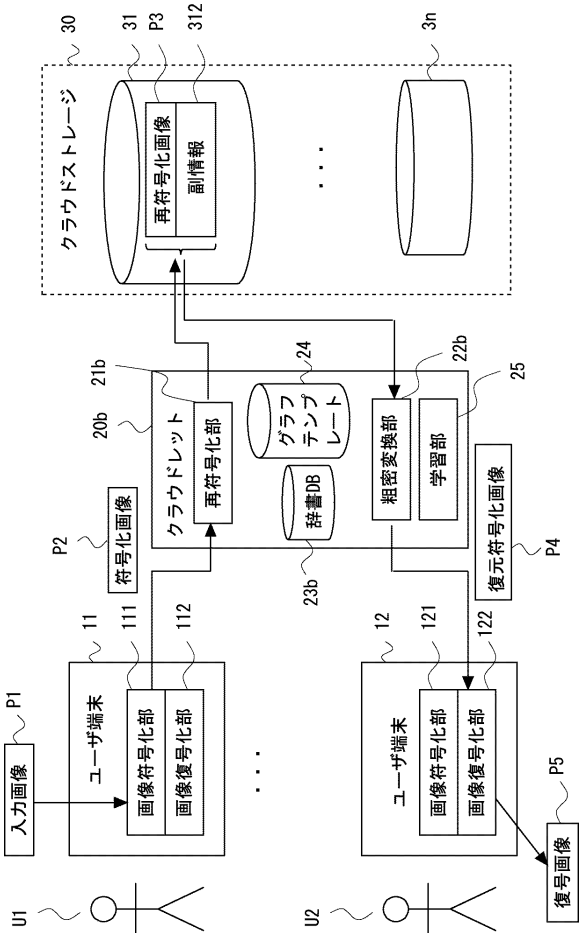
【図 13】



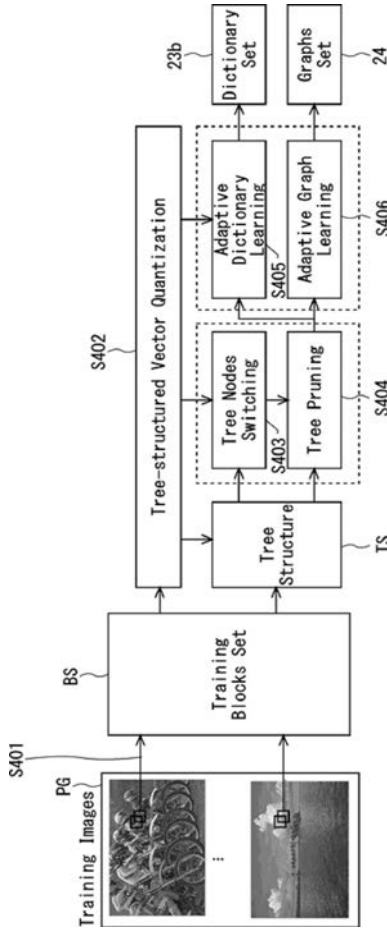
【図 14】



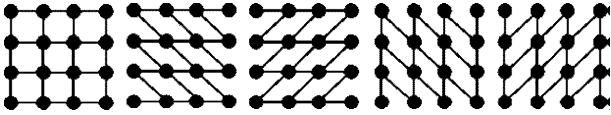
【図 15】



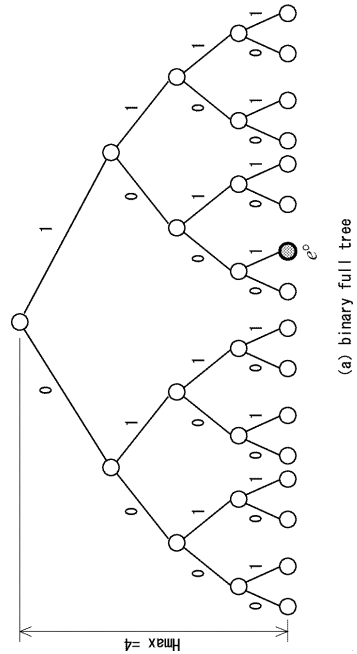
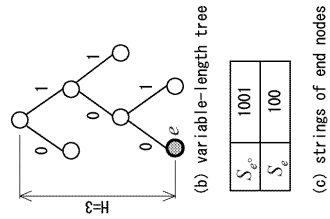
【図 16】



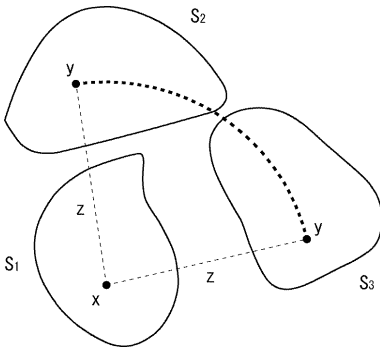
【図 17】



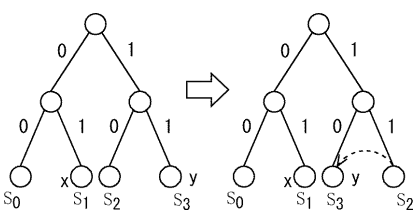
【図 18】



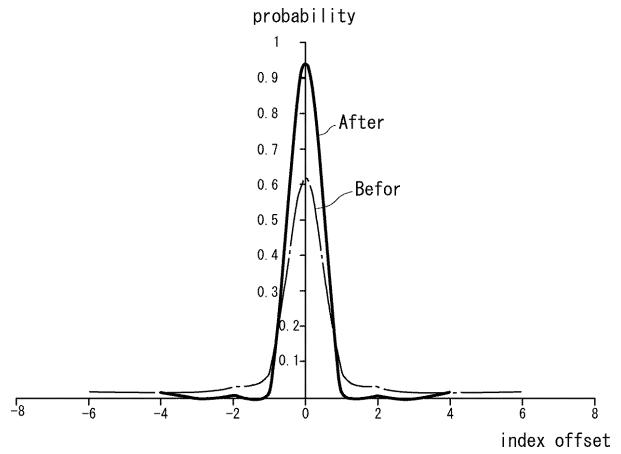
【図 19】



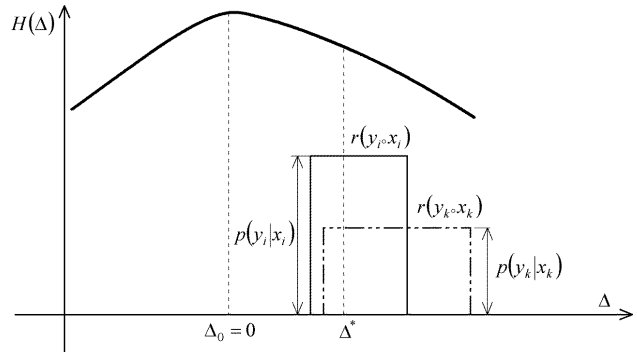
【図 20】



【図 21】



【図 22】



【図 23】

 アルゴリズム 1 : Codebook Design Algorithm

Input:The observation / target pairs $\{P(\mathcal{Y}|\mathcal{X})\}$;The variable-length tree $\mathcal{T}$.**Output:** The codebook $\mathcal{C}$.1: Initialize $\mathcal{C} = \{\Delta_o = 0, \{c_i\}\}$, $R(\mathcal{C}) = 0$;**While** $R(\mathcal{C}) < R(\mathcal{C}^o)$ **Procedure 1: Differential Codeword Selection****For** each observation / target pair $(\mathcal{Y}, \mathcal{X})$ 2: **if** $\mathbf{r}(\mathcal{Y}, \mathcal{X}) \cap \mathcal{C} = \emptyset$ 3: add $P(\mathcal{Y}|\mathcal{X})\mathbf{r}(\mathcal{Y}, \mathcal{X})$ to a differential counter $H(\Delta)$;4: **end****End**5: identify the peak $\Delta^* = \max_{\Delta} H(\Delta)$ as the candidate to add to $\mathcal{C}$.**Procedure 2: Rate Rejudgement****For** each observation / target pair $(\mathcal{Y}, \mathcal{X})$ 6: select the minimum codeword Δ_j in $\mathcal{C}$ that falls within $\mathbf{r}(\mathcal{Y}, \mathcal{X})$;7: if there is, code Δ_j ;8: if there is no, code the direct index codeword c_i ;**End**9: compute rate $R(\mathcal{C})$ using (30);10: **if** $R(\mathcal{C}) < R(\mathcal{C}^o)$ 11: add Δ^* to $\mathcal{C}$;12: **end****End**

フロントページの続き

(71)出願人 500212103

北京大学

PEKING UNIVERSITY

中華人民共和国 ペキン 100871、ハイディアン ディストリクト、5 ユィヒユアン ロード

5 Yiheyuan Road, Haidian District, Beijing 100871 China

(74)代理人 100103894

弁理士 家入 健

(72)発明者 チョン ジーン

東京都千代田区一ツ橋二丁目1番2号 大学共同利用機関法人情報・システム研究機構 国立情報学研究所内

(72)発明者 リュー シャミング

東京都千代田区一ツ橋二丁目1番2号 大学共同利用機関法人情報・システム研究機構 国立情報学研究所内

(72)発明者 リン チャウエン

台湾新竹市光復路二段101号 国立清華大学内

Fターム(参考) 5C159 KK51 MC18 PP01 TA55 TB08 TC04 TC08 TC19 TD10 TD15

UA02 UA05

5C178 BC02 BC56 BC91 CC68 CC73 EC51

5J064 BA13 BC16 BC28