

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2020年5月22日(22.05.2020)



(10) 国際公開番号

WO 2020/100435 A1

(51) 国際特許分類:
H04N 19/90 (2014.01) H04N 19/184 (2014.01)
H04N 19/132 (2014.01)

(21) 国際出願番号: PCT/JP2019/037254

(22) 国際出願日: 2019年9月24日(24.09.2019)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(30) 優先権データ:
特願 2018-213791 2018年11月14日(14.11.2018) JP

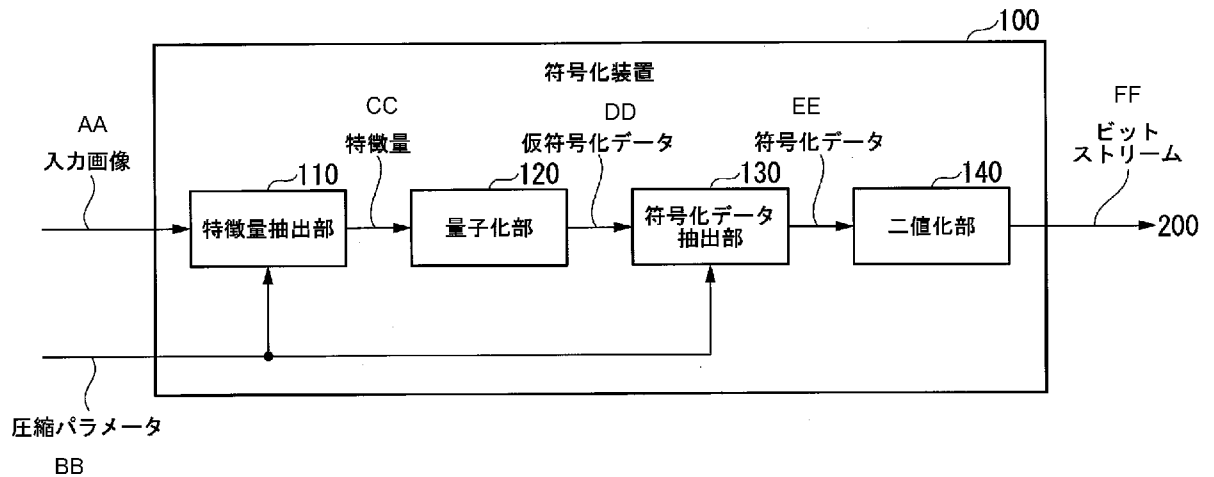
(71) 出願人: 日本電信電話株式会社 (NIPPON TELEGRAPH AND TELEPHONE CORPORATION) [JP/JP]; 〒1008116 東京都千代田区大手町一丁目5番1号 Tokyo (JP).

(72) 発明者: 工藤 忍(KUDO Shinobu); 〒1808585 東京都武蔵野市緑町3丁目9-11 NTT知的財産センタ内 Tokyo (JP). 折橋 翔太(ORIHASHI Shota); 〒1808585 東京都武蔵野市緑町3丁目9-11 NTT知的財産センタ内 Tokyo (JP). 北原 正樹(KITAHARA Masaki); 〒1808585 東京都武蔵野市緑町3丁目9-11 NTT知的財産センタ内 Tokyo (JP). 清水 淳(SHIMIZU Atsushi); 〒1808585 東京都武蔵野市緑町3丁目9-11 NTT知的財産センタ内 Tokyo (JP).

(74) 代理人: 特許業務法人 志賀国際特許事務所 (SHIGA INTERNATIONAL PATENT OFFICE); 〒1006620 東京都千代田区丸の内一丁目9番2号 Tokyo (JP).

(54) Title: ENCODING DEVICE, DECODING DEVICE, ENCODING SYSTEM, LEARNING METHOD, AND PROGRAM

(54) 発明の名称: 符号化装置、復号装置、符号化システム、学習方法、及びプログラム



100	Encoding device	AA	Input image
110	Feature amount extraction unit	BB	Compression parameter
120	Quantization unit	CC	Feature amount
130	Encoded data extraction unit	DD	Provisional encoded data
140	Binarization unit	EE	Encoded data
		FF	Bit stream

(57) Abstract: This encoding device for encoding an input image comprises: a provisional encoded data acquisition unit which obtains, on the basis of the image and a parameter for determining a target size of the encoded data that is data obtained by encoding the image, provisional encoded data having a size larger than the target size; and an encoded data acquisition unit which obtains the encoded data by converting, to a predetermined value, data in a data range other than a data range corresponding to the target size, in the provisional encoded data. The provisional encoded data acquisition

(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

一 国際調査報告(条約第21条(3))

unit obtains the provisional encoded data such that more features for determining the image are included in a data range corresponding to the target size, rather than in a data range other than the data range corresponding to the target size.

(57) 要約: 入力された画像を符号化する符号化装置は、前記画像と、前記画像が符号化されたデータである符号化データの目標サイズを決定するためのパラメータとに基づいて、前記目標サイズより大きいサイズの暫定符号化データを得る暫定符号化データ取得部と、前記暫定符号化データにおいて、前記目標サイズに相当するデータ範囲以外のデータ範囲のデータを所定の値に変換することにより前記符号化データを得る符号化データ取得部と、を備え、前記暫定符号化データ取得部は、前記目標サイズに相当するデータ範囲以外のデータ範囲よりも、前記目標サイズに相当するデータ範囲に、前記画像を決定づける特徴をより多く含むように前記暫定符号化データを得る。

明 細 書

発明の名称：

符号化装置、復号装置、符号化システム、学習方法、及びプログラム
技術分野

[0001] 本発明は、符号化装置、復号装置、符号化システム、学習方法、及びプログラムに関する。

本願は、2018年11月14日に、日本に出願された特願2018-213791号に基づき優先権を主張し、その内容をここに援用する。

背景技術

[0002] 画像等の符号化対象データを符号化する方法の一つとして、オートエンコーダ（自己符号化器）を利用した方法がある。オートエンコーダは、入力データから特徴量を得るエンコーダと、特徴量から入力データと近いデータを得るデコーダからなる。エンコーダ及びデコーダは、任意の演算器によって構築される。例えば入力データが画像である場合、エンコーダは畳み込み演算を行う複数の演算器及び非線形変換器の組み合わせによって構成され、デコーダはエンコーダによる畳み込み演算に対する逆演算を行う複数の演算器及び非線形変換器の組み合わせによって構築される。

[0003] 一般的に、オートエンコーダを含むニューラルネットワークを用いてシステムの設計を行う場合、ニューラルネットワークの構成（例えば、階層数、ユニット数、活性化関数の種類、及び出力サイズ等）を事前に決定する必要がある。例えば、サイズが $X \times Y \times Z$ であり、1画素のビット精度が B ビットである画像データを入力としたオートエンコーダを設計する場合を考える。ここで、 X 、 Y 、及び Z は、それぞれ画像の幅、高さ、及びチャンネル数を示す。そして、エンコーダの出力サイズを $X' \times Y' \times Z'$ 、及び1要素のビット精度を B' ビットに決定した場合、符号化データサイズ及び圧縮率が一意に定まり、符号化データサイズは $X' \times Y' \times Z' \times B'$ 、及び圧縮率は $(X' \times Y' \times Z' \times B') / (X \times Y \times Z \times B)$ によって表される。こ

のことから、オートエンコーダによる符号化器は、1つのニューラルネットワークに対して1つの符号化データサイズ及び圧縮率でしか符号化することができない。そのため、任意の圧縮サイズで符号化するためには、複数の符号化データサイズに対して、それぞれニューラルネットワークを設計する必要がある。

[0004] しかしながら、複数のニューラルネットワークをそれぞれ設計して運用することは、メモリ容量やシステム実装等の観点から実用的でない。これに対し、いくつかの手法が提案されている。例えば、非特許文献1に記載の技術は、入力画像をオートエンコーダに入力して、出力された復号画像と入力画像との差分画像を計算し、その差分画像を再びオートエンコーダに入力して復号差分画像を得る。そして、当該技術は、上記の処理を必要な符号化データサイズになるまで繰り返し行う。これにより、当該技術は、設計されたニューラルネットワークにおける符号化データサイズの倍数となる符号化データサイズで、符号化データサイズを制御する。また、例えば、非特許文献2に記載の技術は、符号化データとは別に、エンコーダ出力の要素ごとに割り当てる符号量（量子化精度）を表した符号量マップを生成する。そして、当該技術は、生成された符号化マップを符号化データとともに伝送することで符号量を制御する。

先行技術文献

非特許文献

- [0005] 非特許文献1: G.Toderici et al., "Full Resolution Image Compression with Recurrent Neural Networks," arXiv, 7 Jul 2017.
- 非特許文献2: M.Li et al., "Learning Convolutional Networks for Content-weighted Image Compression," arXiv, 19 Sep 2017.

発明の概要

発明が解決しようとする課題

- [0006] しかしながら、非特許文献1に記載の技術は、設計されたニューラルネッ

トワークにおける符号化データサイズの倍数となる符号化データサイズでしか、符号化データサイズを制御することができない。そのため、詳細な制御を行うためには、ニューラルネットワークにおける符号化データサイズを小さく設計する必要がある。この場合、所望の符号化データサイズになるまでに何度もエンコード処理及びデコード処理を行わなければならない。これにより、非特許文献 1 に記載の技術は、処理時間が増大するという課題がある。また、非特許文献 2 に記載の技術は、符号量マップが余分なオーバーヘッドになる。これにより、非特許文献 2 に記載の技術は、符号化データサイズが固定されたニューラルネットワークと比べて、符号化効率が低下するという課題がある。

[0007] 本発明はこのような状況を鑑みてなされたもので、処理時間の増大及び符号化効率の低下を抑えつつ、データを所望の大きさに圧縮することができる技術の提供を目的としている。

課題を解決するための手段

[0008] 本発明の一態様は、入力された画像を符号化する符号化装置であって、前記画像と、前記画像が符号化されたデータである符号化データの目標サイズを決定するためのパラメータとに基づいて、前記目標サイズより大きいサイズの暫定符号化データを得る暫定符号化データ取得部と、前記暫定符号化データにおいて、前記目標サイズに相当するデータ範囲以外のデータ範囲のデータを所定の値に変換することにより前記符号化データを得る符号化データ取得部と、を備え、前記暫定符号化データ取得部は、前記目標サイズに相当するデータ範囲以外のデータ範囲よりも、前記目標サイズに相当するデータ範囲に、前記画像を決定づける特徴をより多く含むように前記暫定符号化データを得る符号化装置である。

[0009] また、本発明の一態様は、上記の符号化装置であって、前記パラメータは、符号量又はレート割合である。

[0010] また、本発明の一態様は、上記の符号化装置であって、前記符号化データ取得部は、前記暫定符号化データにおいて、前記目標サイズに相当するデー

タ範囲以外のデータ範囲のデータを削除し、前記データの削除がなされたデータを、復号する対象である前記符号化データとする。

[0011] また、本発明の一態様は、前記第1画像と、前記第1画像が符号化されたデータである符号化データの目標サイズを決定するためのパラメータとに基づいて、前記目標サイズより大きいサイズであり、かつ、前記目標サイズに相当するデータ範囲以外のデータ範囲よりも前記目標サイズに相当するデータ範囲に前記第1画像を決定づける特徴をより多く含む暫定符号化データを取得し、前記暫定符号化データにおいて、前記目標サイズに相当するデータ範囲以外のデータ範囲のデータを所定の値に変換することにより前記符号化データを得る前記符号化装置によって符号化された前記符号化データを復号する復号装置であって、前記符号化データと前記パラメータとに基づいて、前記第1画像とは異なる第2画像に対応する符号化データから復号画像を得る復号画像取得部を備える復号装置である。

[0012] また、本発明の一態様は、画像と、前記画像が符号化されたデータである符号化データの目標サイズを決定するためのパラメータと、に基づく特徴量であって、前記目標サイズより大きいサイズ、かつ、前記目標サイズに相当するデータ範囲以外のデータ範囲よりも前記目標サイズに相当するデータ範囲に前記画像を決定づける特徴をより多く含むように前記特徴量の抽出を学習する特徴量抽出学習部と、前記特徴量において、前記目標サイズに相当するデータ範囲以外のデータ範囲のデータを所定の値に変換することにより変換特徴量を得る変換部と、前記変換特徴量と前記パラメータとに基づいて、前記画像と同一の画像であると判定される復号画像を得るように前記画像の再構成を学習する復号学習部と、を有する符号化システムである。

[0013] また、本発明の一態様は、入力された符号化対象データを符号化する符号化装置であって、前記符号化対象データと、前記符号化対象データが符号化されたデータである符号化データの目標サイズを決定するためのパラメータとに基づいて、前記目標サイズより大きいサイズの暫定符号化データを得る暫定符号化データ取得部と、前記暫定符号化データにおいて、前記目標サイ

ズに相当するデータ範囲以外のデータ範囲のデータを所定の値に変換することにより前記符号化データを得る符号化データ取得部と、を備え、前記暫定符号化データ取得部は、前記目標サイズに相当するデータ範囲以外のデータ範囲よりも、前記目標サイズに相当するデータ範囲に、前記符号化対象データを決定づける特徴をより多く含むように前記暫定符号化データを得る符号化装置である。

[0014] また、本発明の一態様は、前記画像が符号化されたデータである符号化データの目標サイズを決定するためのパラメータと、に基づく特徴量において、前記目標サイズより大きいサイズ、かつ、前記目標サイズに相当するデータ範囲以外のデータ範囲よりも前記目標サイズに相当するデータ範囲に前記画像を決定づける特徴をより多く含むように前記特徴量の抽出を学習するステップと、前記特徴量において、前記目標サイズに相当するデータ範囲以外のデータ範囲のデータを所定の値に変換することにより変換特徴量を得るステップと、前記変換特徴量と前記パラメータとに基づいて、前記画像と同一の画像であると判定される復号画像を得るように前記画像の再構成を学習するステップと、を有する学習方法である。

[0015] また、本発明の一態様は、上記の符号化装置、又は上記の復号装置としてコンピュータを機能させるためのプログラムである。

発明の効果

[0016] 本発明により、処理時間の増大及び符号化効率の低下を抑えつつ、データを所望の大きさに圧縮することができる。

図面の簡単な説明

[0017] [図1]本発明の一実施形態に係る符号化装置100の機能構成を示すブロック図である。

[図2]本発明の一実施形態に係る符号化装置100の特徴量抽出部110の機能構成を示すブロック図である。

[図3]本発明の一実施形態に係る復号装置200の機能構成を示すブロック図である。

[図4]本発明の一実施形態に係る復号装置200の再構成部240の機能構成を示すブロック図である。

[図5]本発明の一実施形態に係る符号化装置100の動作を示すフローチャートである。

[図6]本発明の一実施形態に係る符号化装置100による符号化処理の流れを示す模式図である。

[図7]本発明の一実施形態に係る復号装置200の動作を示すフローチャートである。

[図8]本発明の一実施形態に係る復号装置200による復号処理の流れを示す模式図である。

[図9]本発明の一実施形態に係る符号化装置100及び復号装置200による学習処理の流れを示す模式図である。

発明を実施するための形態

[0018] <実施形態>

以下、本発明の実施形態について、図面を参照しながら説明する。以下では、一例として、画像データを符号化する符号化装置100、及び、画像データを復号する復号装置200について説明する。但し、以下に説明する符号化装置100及び復号装置200は、画像データ以外のデータに対する符号化及び復号にも適用可能である。

[0019] [符号化装置100の構成]

以下、符号化装置100の構成について説明する。符号化装置100は、符号化対象データである入力画像と圧縮パラメータとを入力として、入力画像に対応するビットストリームを出力する。圧縮パラメータとは、入力画像が符号化されたデータである符号化データの目標サイズを決定するためのパラメータである。

[0020] 図1は、本発明の一実施形態に係る符号化装置100の機能構成を示すブロック図である。図1に示すように、符号化装置100は、特徴量抽出部110と、量子化部120と、符号化データ抽出部130と、二値化部140

とを含んで構成される。

[0021] 特徴量抽出部 110 は、外部の装置から入力画像と圧縮パラメータとを取得する。特徴量抽出部 110 は、取得された入力画像と圧縮パラメータとに基づいて入力画像の特徴量を抽出する。ここで、特徴量抽出部 110 は、圧縮パラメータに基づいた大きさ及び所定の領域に、入力画像の特徴が集中するように特徴量抽出を行う。所定の領域とは、符号化側と復号側で共有できる条件であれば何でもよい。例えば、特徴量データの先頭から順でもよい。条件を符号化側から復号側に伝送してもよい。特徴量抽出部 110 は、抽出された特徴量を示す情報を量子化部 120 へ出力する。

[0022] 量子化部 120（暫定符号化データ取得部）は、特徴量抽出部 110 から出力された情報を取得する。量子化部 120 は、取得された情報に基づく特徴量に対して量子化処理を実行し、仮符号化データ（暫定符号化データ）に変換する。量子化部 120 は、生成された仮符号化データを符号化データ抽出部 130 へ出力する。

[0023] 符号化データ抽出部 130（符号化データ取得部）は、量子化部 120 から出力された仮符号化データを取得する。また、符号化データ抽出部 130 は、外部の装置から圧縮パラメータを取得する。符号化データ抽出部 130 は、取得された仮符号化データと圧縮パラメータとに基づいて、符号化データを抽出する。符号化データ抽出部 130 は、抽出された符号化データを二値化部 140 へ出力する。

[0024] 上述したように、特徴量抽出部 110 は、圧縮パラメータに基づいた大きさ及び領域に、入力画像の特徴が集中するように特徴量抽出を行う。符号化データ抽出部 130 で、例えば該領域を除く領域を削除することで、符号化データが圧縮パラメータに基づくサイズ、例えば所望のビットレートになるような処理を行う。

[0025] 二値化部 140 は、符号化データ抽出部 130 から出力された符号化データを取得する。二値化部 140 は、取得された符号化データを二値化する。二値化部 140 は、二値化された符号化データをビットストリームとして外

部の装置へ出力する。

[0026] [特徴量抽出部 110 の構成]

以下、特徴量抽出部 110 の構成についてさらに詳しく説明する。特徴量抽出部 110 は、例えば図 2 に示すようなニューラルネットワーク（畳み込み演算、ダウンサンプリング、及び非線形変換の組み合わせ）を含んで構成される。

[0027] 図 2 は、本発明の一実施形態に係る符号化装置 100 の特徴量抽出部 110 の機能構成を示すブロック図である。図 2 に示すように、特徴量抽出部 110 は、サイズ拡大部 111 と、結合部 112 と、N 層からなる抽出部（第 1 層抽出部 113-1 ～第 N 層抽出部 113-N）とによって構成される。また、図 2 に示すように、第 1 層抽出部 113-1、・・・、第 N 層抽出部 113-N は、それぞれ、畳み込み部 115-1 とダウンサンプリング部 116-1 と非線形変換部 117-1、・・・、畳み込み部 115-N とダウンサンプリング部 116-N と非線形変換部 117-N によって構成される。

[0028] サイズ拡大部 111 は、外部の装置から圧縮パラメータを取得する。サイズ拡大部 111 は、取得された圧縮パラメータを、入力画像と同じサイズにまで拡大する処理を行う。サイズ拡大部 111 は、拡大された圧縮パラメータを結合部 112 へ出力する。

[0029] 結合部 112 は、外部の装置から入力画像を取得する。また、結合部 112 は、サイズ拡大部 111 から出力された、拡大された圧縮パラメータを取得する。結合部 112 は、取得された入力画像と拡大された圧縮パラメータとをチャンネル方向に結合する処理を行う。結合部 112 は、拡大された圧縮パラメータが結合された入力画像を第 1 層抽出部 113-1 の畳み込み部 115-1 へ出力する。

[0030] 第 1 層抽出部 113-1 の畳み込み部 115-1 は、結合部 112 から出力された入力画像を取得する。畳み込み部 115-1 は、取得された入力画像に対して畳み込み処理を行う。畳み込み部 115-1 は、畳み込み処理が

なされた入力画像をダウンサンプリング部 116-1 へ出力する。

[0031] ダウンサンプリング部 116-1 は、畳み込み部 115-1 から出力された入力画像を取得する。ダウンサンプリング部 116-1 は、取得された入力画像をダウンサンプリングする処理を行う。ダウンサンプリング部 116-1 は、ダウンサンプリングされた入力画像を非線形変換部 117-1 へ出力する。

[0032] 非線形変換部 117-1 は、ダウンサンプリング部 116-1 から出力された入力画像を取得する。非線形変換部 117-1 は、取得された入力画像の各要素に対して非線形変換する処理を行う。非線形変換部 117-1 は、非線形変換処理がなされた入力画像を、次の層の抽出部の畳み込み部へ出力する。

[0033] 上記の処理を第 1 層から第 N 層まで繰り返すことにより、特徴量抽出部 110 は、取得された入力画像と圧縮パラメータとに基づいて入力画像の特徴量を抽出する。第 N 層抽出部 113-N の非線形変換部 117-N は、抽出された特徴量を示す情報を量子化部 120 へ出力する。

[0034] [復号装置 200 の構成]

以下、復号装置 200 の構成について説明する。復号装置 200 は、ビットストリームを入力として、入力画像に対応する復号画像を出力する。

[0035] 図 3 は、本発明の一実施形態に係る復号装置 200 の機能構成を示すブロック図である。図 3 に示すように、復号装置 200 は、逆二値化部 210 と、符号化データ伸張部 220 と、圧縮パラメータ算出部 230 と、再構成部 240 とを含んで構成される。

[0036] 逆二値化部 210 は、外部の装置からビットストリームを取得する。逆二値化部 210 は、取得されたビットストリームを符号化データに変換する。逆二値化部 210 は、生成された符号化データを、符号化データ伸張部 220 及び圧縮パラメータ算出部 230 へそれぞれ出力する。

[0037] 符号化データ伸張部 220 は、逆二値化部 210 から出力された符号化データを取得する。符号化データ伸張部 220 は、取得された符号化データの

要素数を、符号化装置 100 の量子化部 120 によって生成される仮符号化データと同じ要素数まで伸張することにより仮符号化データを生成する。符号化データ伸張部 220 は、生成された仮符号化データを、圧縮パラメータ算出部 230 及び再構成部 240 へそれぞれ出力する。

[0038] 圧縮パラメータ算出部 230 は、逆二値化部 210 から出力された符号化データを取得する。また、圧縮パラメータ算出部 230 は、符号化データ伸張部 220 から出力された仮符号化データを取得する。圧縮パラメータ算出部 230 は、取得された符号化データと仮符号化データとに基づいて圧縮パラメータを算出する。圧縮パラメータ算出部 230 は、算出された圧縮パラメータを再構成部 240 へ出力する。

[0039] 再構成部 240（復号画像取得部）は、符号化データ伸張部 220 から出力された仮符号化データを取得する。また、再構成部 240 は、圧縮パラメータ算出部 230 から出力された圧縮パラメータを取得する。再構成部 240 は、仮符号化データと圧縮パラメータとに基づいて復号画像を再構成する。再構成部 240 は、再構成された復号画像を外部の装置へ出力する。

[0040] [再構成部 240 の構成]

以下、再構成部 240 の構成についてさらに詳しく説明する。再構成部 240 は、例えば図 4 に示すようなニューラルネットワーク（逆畳み込み演算及び非線形変換の組み合わせ）を含んで構成される。

[0041] 図 4 は、本発明の一実施形態に係る復号装置 200 の再構成部 240 の機能構成を示すブロック図である。図 4 に示すように、再構成部 240 は、サイズ拡大部 241 と、結合部 242 と、M 層からなる構成部（第 1 層構成部 243-1～第 M 層構成部 243-M）とによって構成される。また、図 4 に示すように、第 1 層構成部 243-1、・・・、第 M 層構成部 243-M は、それぞれ、逆畳み込み部 245-1 と非線形変換部 246-1、・・・、逆畳み込み部 245-M と非線形変換部 246-M とによって構成される。

[0042] サイズ拡大部 241 は、圧縮パラメータ算出部 230 から出力された圧縮

パラメータを取得する。サイズ拡大部 241 は、取得された圧縮パラメータを、入力画像と同じサイズにまで拡大する処理を行う。サイズ拡大部 241 は、予め定められた「0」等の値を付与することで、入力画像と同じサイズにまで拡大する処理を行う。サイズ拡大部 241 は、拡大された圧縮パラメータを結合部 242 へ出力する。

[0043] 結合部 242 は、符号化データ伸張部 220 から仮符号化データを取得する。また、結合部 112 は、サイズ拡大部 241 から出力された、拡大された圧縮パラメータを取得する。結合部 242 は、取得された仮符号化データと拡大された圧縮パラメータとをチャンネル方向に結合する処理を行う。結合部 242 は、拡大された圧縮パラメータが結合された仮符号化データを第 1 層構成部 243-1 の逆畳み込み部 245-1 へ出力する。

[0044] 第 1 層構成部 243-1 の逆畳み込み部 245-1 は、結合部 242 から出力された仮符号化データを取得する。逆畳み込み部 245-1 は、取得された仮符号化データに対して、符号化装置 100 の特徴量抽出部 110 による畳み込み演算に対する逆演算を行う。逆畳み込み部 245-1 は、逆演算された仮符号化データを非線形変換部 246-1 へ出力する。

[0045] 非線形変換部 246-1 は、逆畳み込み部 245-1 から出力された仮符号化データを取得する。非線形変換部 246-1 は、取得された仮符号化データの各要素に対して非線形変換する処理を行う。非線形変換部 246-1 は、非線形変換処理がなされた仮符号化データを、次の層の構成部の逆畳み込み部へ出力する。

[0046] 上記の処理を第 1 層から第 M 層まで繰り返すことにより、再構成部 240 は、取得された仮符号化データと圧縮パラメータとに基づいて復号画像を再構成する。第 M 層構成部 243-M の非線形変換部 246-M は、再構成された復号画像を外部の装置へ出力する。

[0047] 上述したように、符号化装置 100 から送信される仮符号化データは、入力画像の特徴が集中している領域のみを示すデータである。言い換えると、復号装置 200 の再構成部 240 により復号画像を得るためには、符号化装

置 100 の符号化データ抽出部 130 により削除された領域を補う必要がある。符号化データ抽出部 130 において削除された領域は入力画像の特徴ではないため、復号装置 200 の再構成部 240 のサイズ拡大部 241 は、上述したように、予め定められた「0」等の値を仮符号化データに付与することで、再構成部 240 が仮符号化データから復号画像を得られることができる。

[0048] [符号化装置 100 の動作]

以下、符号化装置 100 の動作について、具体例を挙げて説明する。

図 5 は、本発明の一実施形態に係る符号化装置 100 の動作を示すフローチャートである。また、図 6 は、本発明の一実施形態に係る符号化装置 100 による符号化処理の流れを示す模式図である。

[0049] まず、符号化対象の入力画像を $I(x, y, z)$ 、及び圧縮パラメータを R として定義する。ここで、 x は水平方向の変数、 y は垂直方向の変数、及び z はチャンネル方向の変数を示す。また、 x, y, z の次元数をそれぞれ X, Y, Z とする。また、1 要素のビット精度を B ビットとする。例えば、入力画像 $I(x, y, z)$ がグレイ画像である場合には $Z = 1$ であり、入力画像 $I(x, y, z)$ が RGB 画像である場合には $Z = 3$ である。また、圧縮パラメータ R は、所望の符号化データサイズ（目標サイズ）を決定することができるパラメータである。本実施形態では、一例として、圧縮パラメータ R は、 $0 < R \leq 1$ の範囲の値を取りうる、圧縮率を示すパラメータであるものとする。なお、圧縮率とは、符号化データサイズ／入力画像 $I(x, y, z)$ のサイズ、によって算出される比率である。

[0050] 特徴量抽出部 110 は、入力画像 $I(x, y, z)$ に対して、圧縮パラメータ R をパラメータとする特徴量抽出処理を行うことにより、特徴量 $F(x, y, z)$ を抽出する（ステップ S101）。ここで、 x, y, z の次元数をそれぞれ X', Y', Z' とする。特徴量抽出処理としては、上述した、例えば図 2 に示すようなニューラルネットワークが用いられる。

[0051] 量子化部 120 は、特徴量 $F(x, y, z)$ を、所定の順序で 1 次元のベ

クトルに変形する。そして、量子化部120は、各要素が所定のビット精度 B' となるように量子化処理を行い、仮符号化データを生成する（ステップS102）。

[0052] 符号化データ抽出部130は、仮符号化データの先頭から、圧縮パラメータ R から計算される符号化データサイズ分のデータを抽出することにより、符号化データを得る（ステップS103）。

[0053] 二値化部140は、符号化データを二値化することにより、ビットストリームを得る（ステップS104）。

[0054] [復号装置200の動作]

以下、復号装置200の動作について、具体例を挙げて説明する。

図7は、本発明の一実施形態に係る復号装置200の動作を示すフローチャートである。また、図8は、本発明の一実施形態に係る復号装置200による復号処理の流れを示す模式図である。

[0055] 逆二値化部210は、ビットストリームを逆二値化し、符号化データに変換する（ステップS201）。

符号化データ伸張部220は、符号化データを符号化装置100の仮符号化データと同じ要素数になるまで伸張し、仮符号化データ（変換特徴量）を生成する。具体的には、符号化データ伸張部220（変換部）は、符号化データに対して、不足している要素数の分だけ所定の値（図8に示すように、例えば0）を追加する（ステップS202）。

[0056] 圧縮パラメータ算出部230は、符号化データと仮符号化データとに基づいて、圧縮パラメータ R を算出する。具体的には、圧縮パラメータ算出部230は、符号化データに対応する復号画像のデータサイズ（すなわち、 $X \times Y \times Z \times B$ ）を算出する。そして、圧縮パラメータ算出部230は、圧縮パラメータ R を、 $R = (X' \times Y' \times Z' \times B') / (X \times Y \times Z \times B)$ として算出する（ステップS203）。

[0057] 再構成部240は、仮符号化データを再構成処理の入力サイズに整形する。そして、再構成部240は、仮符号化データに対して、圧縮パラメータ R

をパラメータとする再構成処理を行うことにより、復号画像 I' (x , y , z) を生成する (ステップ S 204)。再構成処理としては、上述した、例えば図 4 に示すようなニューラルネットワークが用いられる。

[0058] なお、特徴量の次元数は、 $X \times Y \times Z \times B = X' \times Y' \times Z' \times B'$ を満たすように設計されることが最善である。しかしながら、これは必須条件ではなく、例えば、 $X \times Y \times Z \times B > X' \times Y' \times Z' \times B'$ として設計されてもよい。但し、その場合には、入力できる圧縮パラメータ R の最大値に上限が付く。

[0059] なお、二値化されたビットストリーム (符号化データ) に対してエントロピー符号化を行う構成にしてもよい。この場合、エントロピー符号化後の符号量がフィードバックされることによって、レート制御を行うことが可能になる。例えば、画像をブロックに分割して、あるブロックをレート割合 0.5 (50%) で符号化し、エントロピー符号化した結果が例えば 0.4 になってしまった場合に、その次のブロックの符号化においてはレート割合を例えば 0.6 として符号化すること等によって、全体のレート制御を行うことができる。

[0060] [学習処理の流れ]

次に、本実施形態における特徴量抽出部 110 (特徴量抽出学習部)、及び再構成部 240 (復号学習部) を構成するニューラルネットワークにおける学習方法について説明する。ここで、ニューラルネットワークはオートエンコーダであり、入力画像と同一の画像であると判定される復号画像を得るように学習が行われる。特徴量抽出部 110 における学習と、再構成部 240 における学習とは、同時に行われる。

[0061] 学習処理の事前準備として、入力画像 I (x , y , z) 及び圧縮パラメータ R のセットをサンプルデータとするデータセットを用意する。圧縮パラメータ R は、その取りうる値から一様分布のランダム値とする。まず、上述した符号化装置 100 による符号化処理によって、入力画像 I (x , y , z) に対するビットストリームを得る。そして、そのビットストリームから、上

述した復号装置 200 による復号処理によって復号画像を得る。次に、下記の式 (1) によって定義される損失関数を用いて損失値 $loss$ を算出する。

$$[0062] \quad loss = \sum_x \sum_y \sum_z diff(I(x, y, z), I'(x, y, z)) \quad \dots (1)$$

[0063] ここで、 $diff(a, b)$ は、 a と b との距離を測る関数（例えば二乗誤差など）である。なお、上記の式 (1) で定義される損失関数は一例であり、一部の誤差のみを算出したり、異なる誤差項を追加したりしてもよい。

[0064] 算出された損失値 $loss$ を用いて、逆誤差伝播法などにより、特徴量抽出部 110 のパラメータ、及び再構成部 240 のパラメータを更新する。上記の一連の流れを 1 回とし、複数のサンプルデータで一定回数、あるいは、損失値 $loss$ が収束するまで繰り返されることによって、特徴量抽出部 110、及び再構成部 240 を構成するニューラルネットワークにおける学習が行われる。

[0065] 以上説明したように、本発明の一実施形態に係る符号化装置 100 及び復号装置 200 は、圧縮パラメータをパラメータとして、特徴量抽出処理及び再構成処理を行う。また、符号化装置 100 及び復号装置 200 は、学習時に先頭から必要な符号化データサイズ分のデータ（目標サイズに相当するデータ範囲）のみを抽出し、それ以外（目標サイズに相当するデータ範囲以外のデータ範囲）を所定の値（例えば 0）で埋めてから復号する。このような構成を備えることにより、符号化装置 100 及び復号装置 200 は、画像を圧縮（低次元化）した際に、画像の主な特徴を表現するパラメータが、圧縮されたデータにおける所望のデータ範囲（例えば、符号化データの先頭から必要な符号化データサイズ分の要素）に密集するように（すなわち、画像を決定づける特徴をより多く含むように）学習する。

[0066] そのため、本発明の一実施形態に係る符号化装置 100 及び復号装置 200 によれば、オートエンコーダシステムを複数の符号化データサイズで個別

に設計した時と同じ効果を、１つのシステムで実現することができる。また、従来技術１のように何度もエンコード・デコード処理を行うこともなく、従来技術２のようにオーバーヘッドも必要としない。これにより、本発明の一実施形態に係る符号化装置１００及び復号装置２００によれば、処理時間の増大及び符号化効率の低下を抑えつつ、データを所望の大きさに圧縮することができる。

[0067] 上述した実施形態における符号化装置１００及び復号装置２００の一部又は全部を、コンピュータで実現するようにしてもよい。その場合、この機能を実現するためのプログラムをコンピュータ読み取り可能な記録媒体に記録して、この記録媒体に記録されたプログラムをコンピュータシステムに読み込ませ、実行することによって実現してもよい。なお、ここでいう「コンピュータシステム」とは、ＯＳや周辺機器等のハードウェアを含むものとする。また、「コンピュータ読み取り可能な記録媒体」とは、フレキシブルディスク、光磁気ディスク、ＲＯＭ、ＣＤ－ＲＯＭ等の可搬媒体、コンピュータシステムに内蔵されるハードディスク等の記憶装置のことをいう。さらに「コンピュータ読み取り可能な記録媒体」とは、インターネット等のネットワークや電話回線等の通信回線を介してプログラムを送信する場合の通信線のように、短時間の間、動的にプログラムを保持するもの、その場合のサーバやクライアントとなるコンピュータシステム内部の揮発性メモリのように、一定時間プログラムを保持しているものも含んでもよい。また上記プログラムは、上述した機能の一部を実現するためのものであっても良く、さらに上述した機能をコンピュータシステムにすでに記録されているプログラムとの組み合わせで実現できるものであってもよく、ＰＬＤ（Programmable Logic Device）やＦＰＧＡ（Field Programmable Gate Array）等のハードウェアを用いて実現されるものであってもよい。

[0068] 以上、図面を参照して本発明の実施形態を説明してきたが、上記実施形態は本発明の例示に過ぎず、本発明が上記実施形態に限定されるものではないことは明らかである。したがって、本発明の技術思想及び要旨を逸脱しない

範囲で構成要素の追加、省略、置換、及びその他の変更を行ってもよい。

符号の説明

[0069]	1 0 0	符号化装置
	1 1 0	特徴量抽出部
	1 1 1	サイズ拡大部
	1 1 2	結合部
	1 1 5 - 1 ~ 1 1 5 - N	畳み込み部
	1 1 6 - 1 ~ 1 1 6 - N	ダウンサンプリング部
	1 1 7 - 1 ~ 1 1 7 - N	非線形変換部
	1 2 0	量子化部
	1 3 0	符号化データ抽出部
	1 4 0	二値化部
	2 0 0	復号装置
	2 1 0	逆二値化部
	2 2 0	符号化データ伸張部
	2 3 0	圧縮パラメータ算出部
	2 4 0	再構成部
	2 4 1	サイズ拡大部
	2 4 2	結合部
	2 4 5 - 1 ~ 2 4 5 - M	逆畳み込み部
	2 4 6 - 1 ~ 2 4 6 - M	非線形変換部

請求の範囲

- [請求項1] 入力された画像を符号化する符号化装置であって、
 前記画像と、前記画像が符号化されたデータである符号化データの
 目標サイズを決定するためのパラメータとに基づいて、前記目標サイ
 ズより大きいサイズの暫定符号化データを得る暫定符号化データ取得
 部と、
 前記暫定符号化データにおいて、前記目標サイズに相当するデータ
 範囲以外のデータ範囲のデータを所定の値に変換することにより前記
 符号化データを得る符号化データ取得部と、
 を備え、
 前記暫定符号化データ取得部は、前記目標サイズに相当するデータ
 範囲以外のデータ範囲よりも、前記目標サイズに相当するデータ範囲
 に、前記画像を決定づける特徴をより多く含むように前記暫定符号化
 データを得る
 符号化装置。
- [請求項2] 前記パラメータの値は、符号量又はレート割合である
 請求項 1 に記載の符号化装置。
- [請求項3] 前記符号化データ取得部は、
 前記暫定符号化データにおいて、前記目標サイズに相当するデータ
 範囲以外のデータ範囲のデータを削除し、前記データの削除がなされ
 たデータを、復号する対象である前記符号化データとする
 請求項 1 又は請求項 2 に記載の符号化装置。
- [請求項4] 第 1 画像と、前記第 1 画像が符号化されたデータである符号化デー
 タの目標サイズを決定するためのパラメータとに基づいて、前記目標
 サイズより大きいサイズであり、かつ、前記目標サイズに相当するデ
 ータ範囲以外のデータ範囲よりも前記目標サイズに相当するデータ範
 囲に前記第 1 画像を決定づける特徴をより多く含む暫定符号化データ
 を取得し、前記暫定符号化データにおいて、前記目標サイズに相当す

るデータ範囲以外のデータ範囲のデータを所定の値に変換することにより前記符号化データを得る符号化装置によって符号化された前記符号化データを復号する復号装置であって、

前記符号化データと前記パラメータとに基づいて、前記第1画像とは異なる第2画像に対応する符号化データから復号画像を得る復号画像取得部

を備える復号装置。

[請求項5]

画像と、前記画像が符号化されたデータである符号化データの目標サイズを決定するためのパラメータと、に基づく特徴量であって、前記目標サイズより大きいサイズ、かつ、前記目標サイズに相当するデータ範囲以外のデータ範囲よりも前記目標サイズに相当するデータ範囲に前記画像を決定づける特徴をより多く含むように前記特徴量の抽出を学習する特徴量抽出学習部と、

前記特徴量において、前記目標サイズに相当するデータ範囲以外のデータ範囲のデータを所定の値に変換することにより変換特徴量を得る変換部と、

前記変換特徴量と前記パラメータとに基づいて、前記画像と同一の画像であると判定される復号画像を得るように前記画像の再構成を学習する復号学習部と、

を有する符号化システム。

[請求項6]

入力された符号化対象データを符号化する符号化装置であって、

前記符号化対象データと、前記符号化対象データが符号化されたデータである符号化データの目標サイズを決定するためのパラメータとに基づいて、前記目標サイズより大きいサイズの暫定符号化データを得る暫定符号化データ取得部と、

前記暫定符号化データにおいて、前記目標サイズに相当するデータ範囲以外のデータ範囲のデータを所定の値に変換することにより前記符号化データを得る符号化データ取得部と、

を備え、

前記暫定符号化データ取得部は、前記目標サイズに相当するデータ範囲以外のデータ範囲よりも、前記目標サイズに相当するデータ範囲に、前記符号化対象データを決定づける特徴をより多く含むように前記暫定符号化データを得る

符号化装置。

[請求項7]

画像と、前記画像が符号化されたデータである符号化データの目標サイズを決定するためのパラメータと、に基づく特徴量において、前記目標サイズより大きいサイズ、かつ、前記目標サイズに相当するデータ範囲以外のデータ範囲よりも前記目標サイズに相当するデータ範囲に前記画像を決定づける特徴をより多く含むように前記特徴量の抽出を学習するステップと、

前記特徴量において、前記目標サイズに相当するデータ範囲以外のデータ範囲のデータを所定の値に変換することにより変換特徴量を得るステップと、

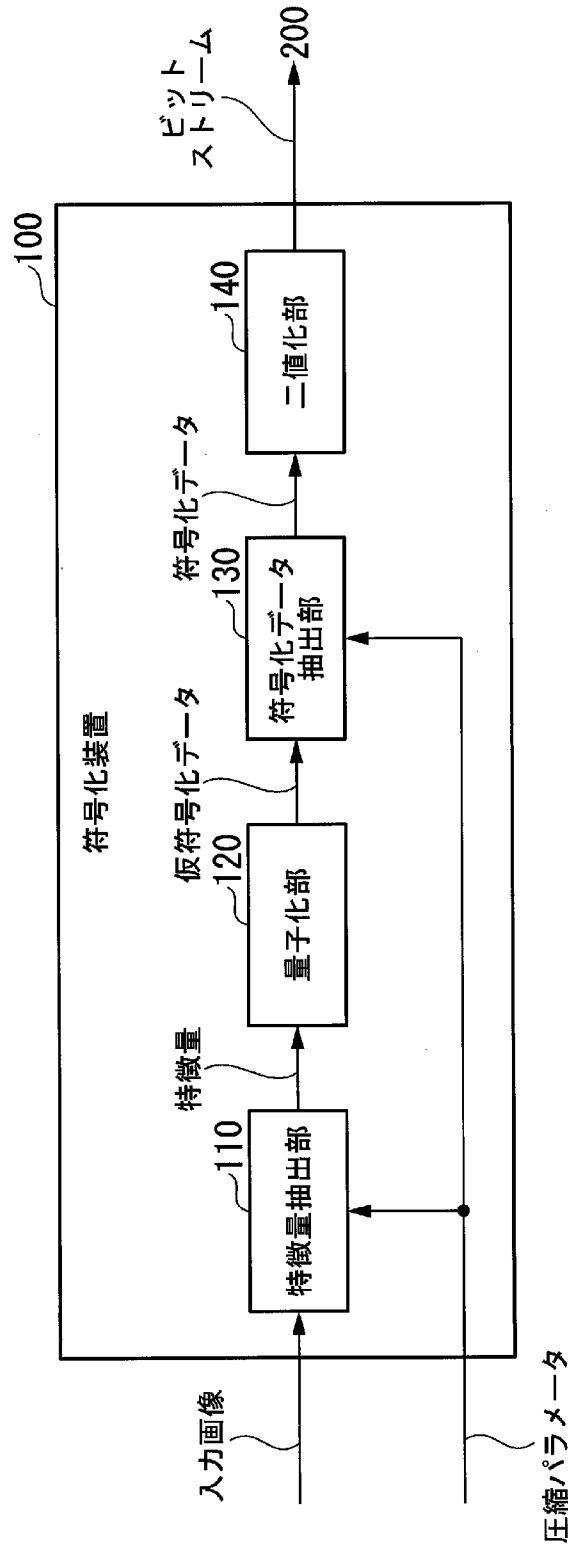
前記変換特徴量と前記パラメータとに基づいて、前記画像と同一の画像であると判定される復号画像を得るように前記画像の再構成を学習するステップと、

を有する学習方法。

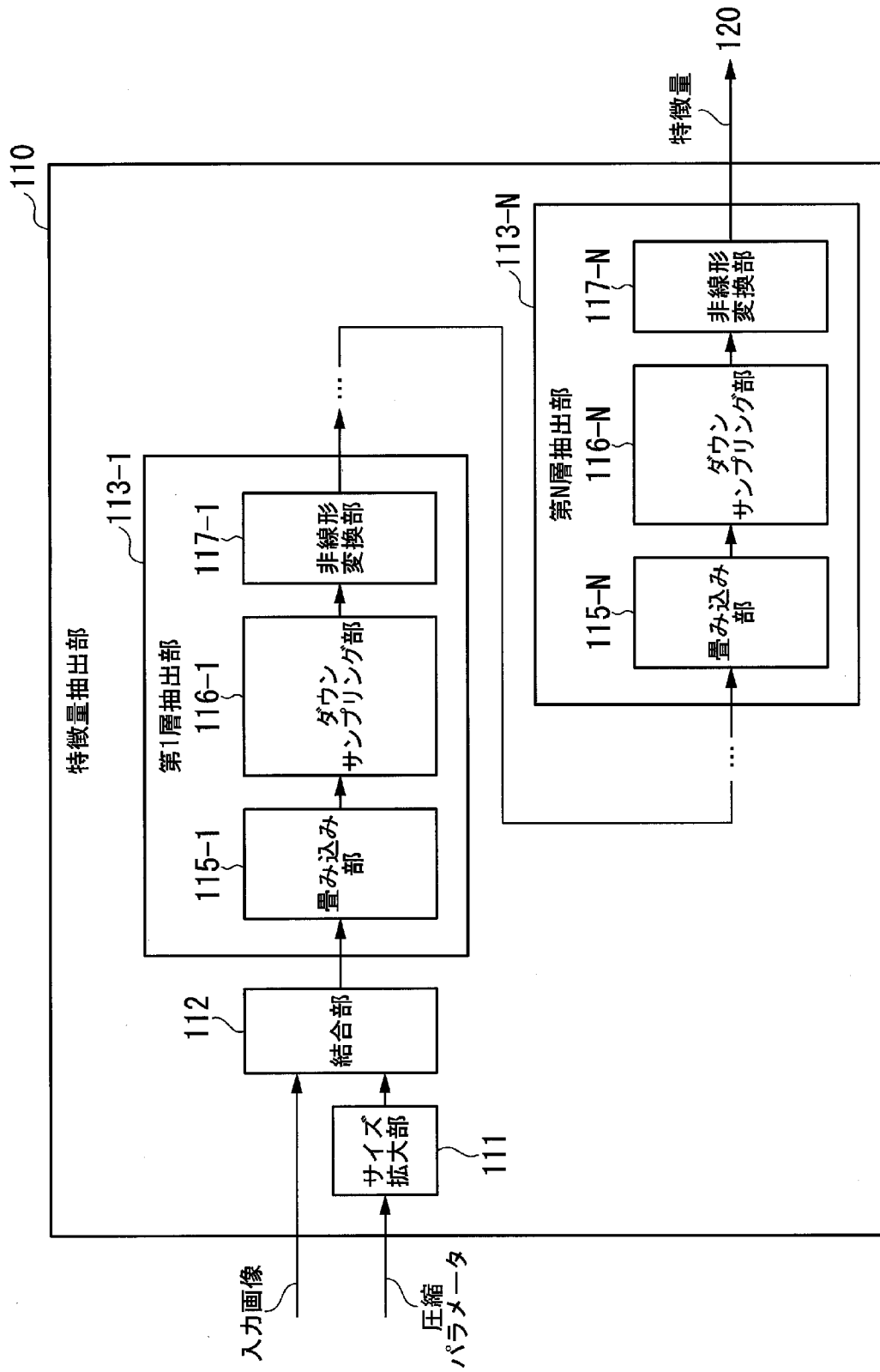
[請求項8]

請求項1、2、3又は6に記載の符号化装置、又は請求項4に記載の復号装置としてコンピュータを機能させるためのプログラム。

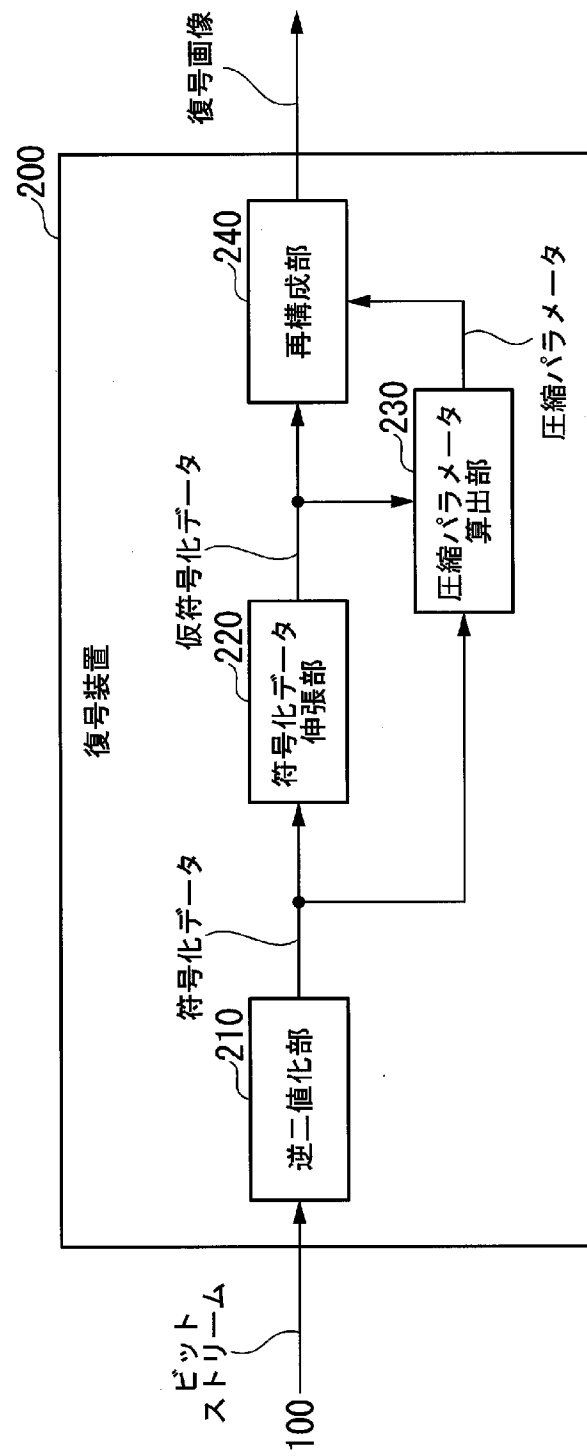
[図1]



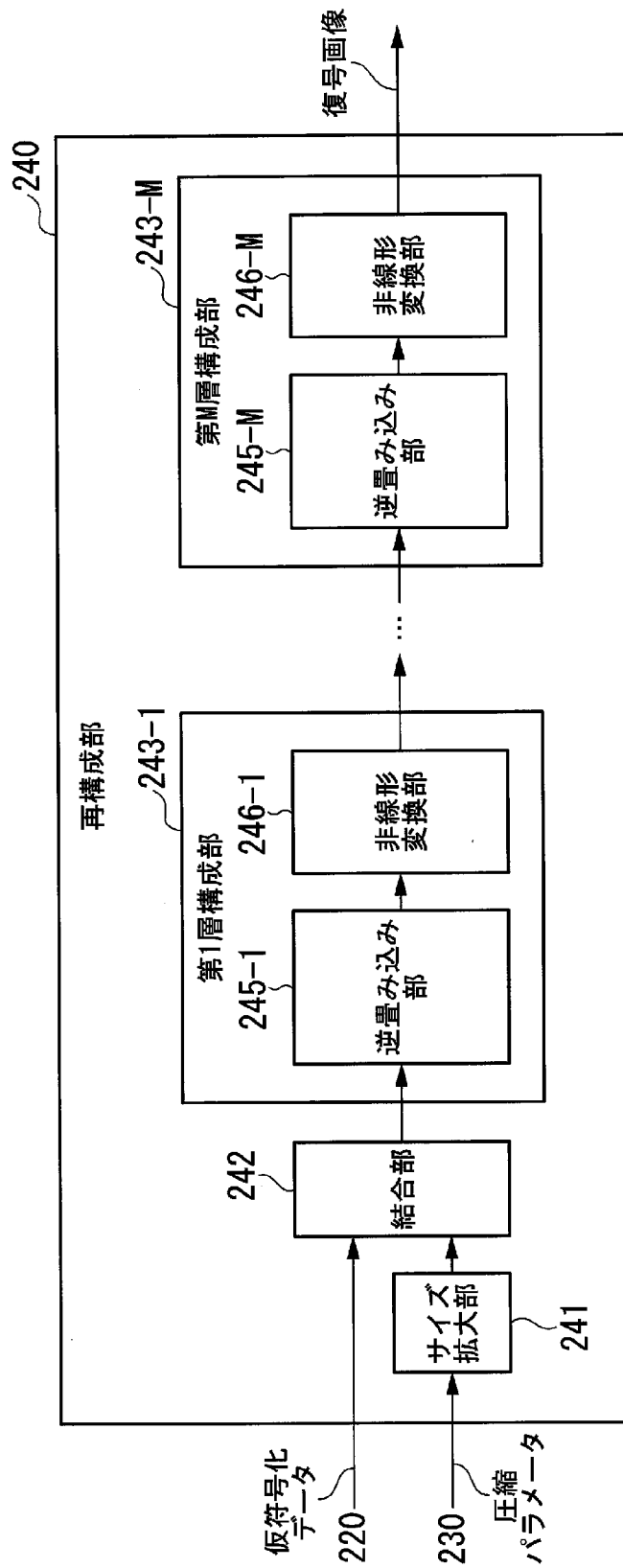
[図2]



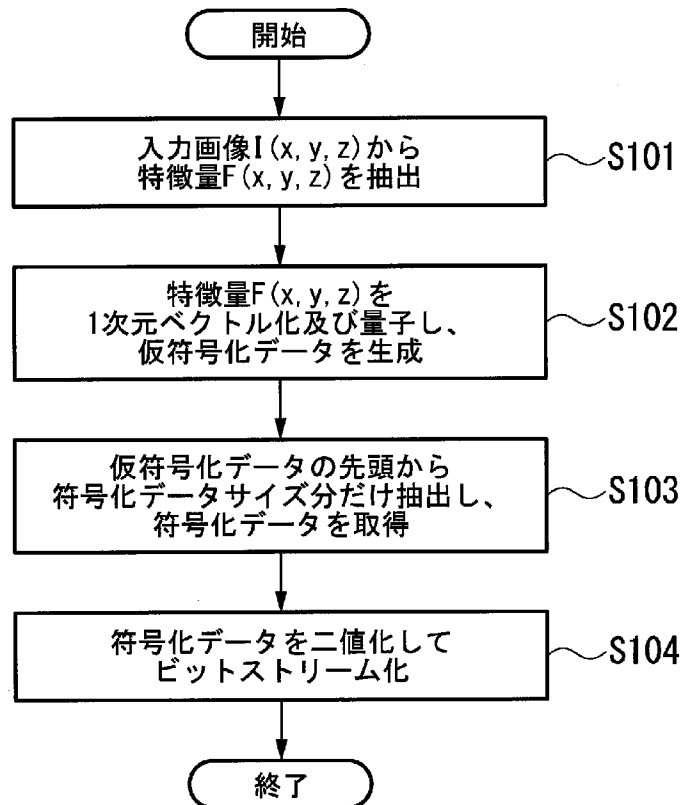
[図3]



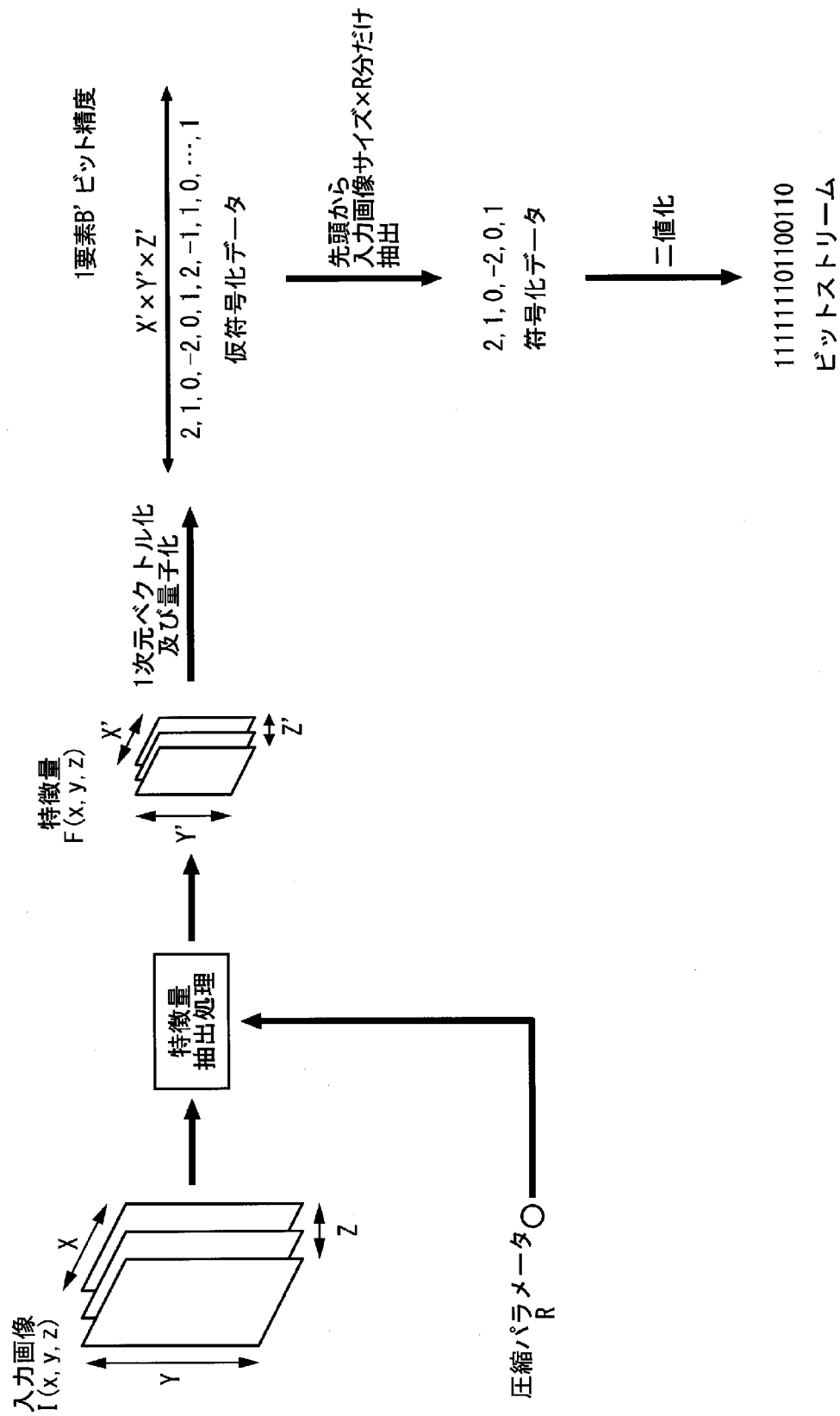
[図4]



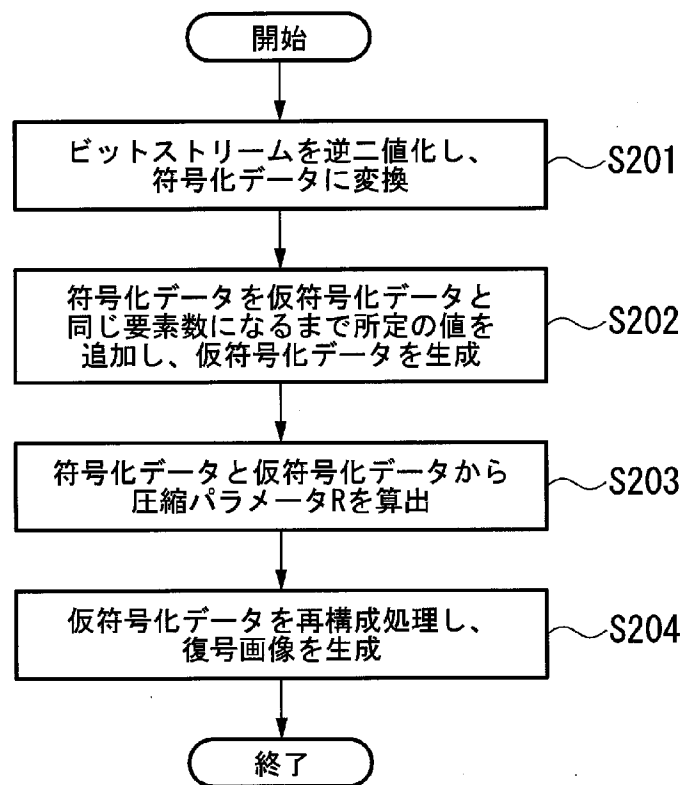
[図5]



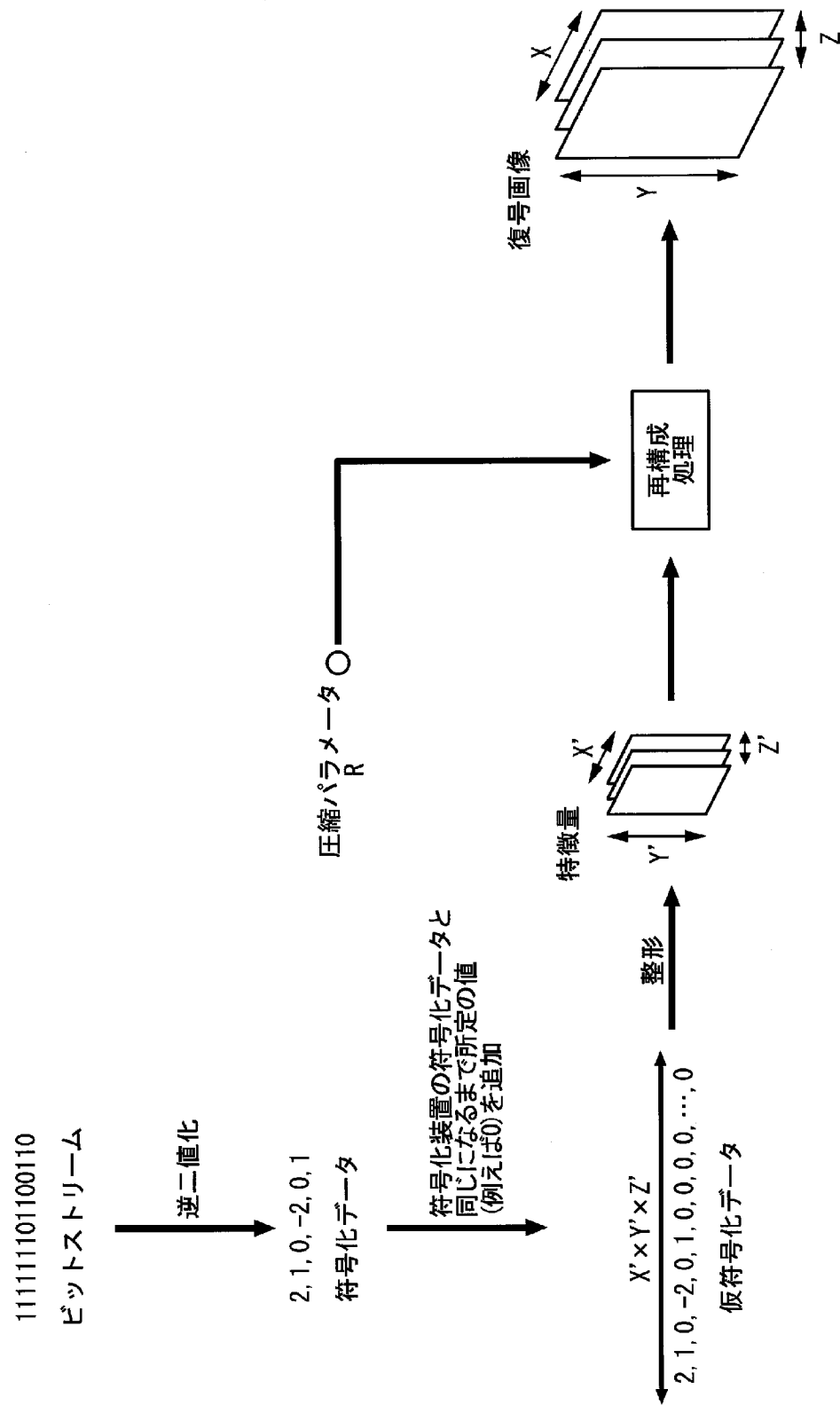
[図6]



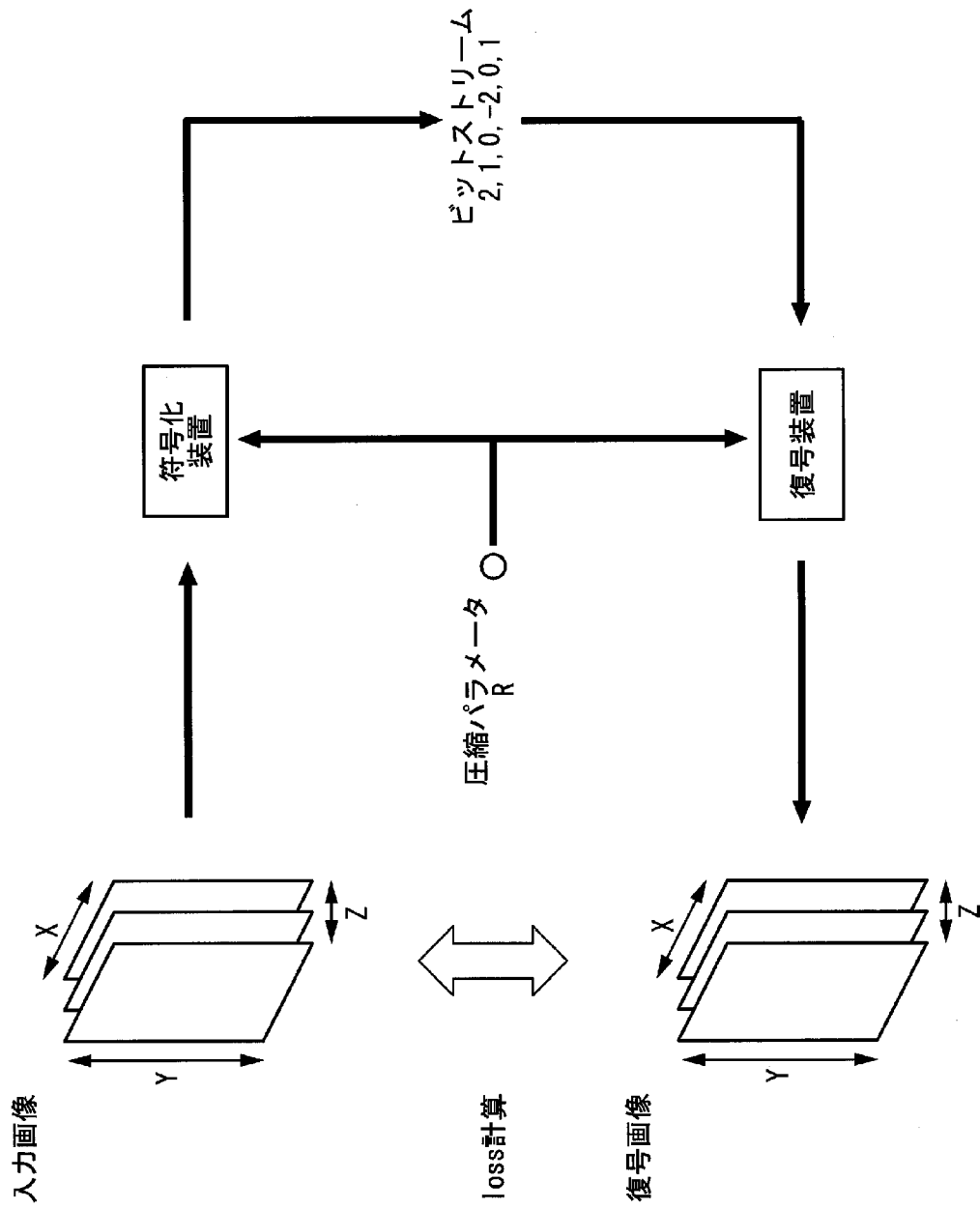
[図7]



[図8]



[図9]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2019/037254

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl. H04N19/90 (2014.01) i, H04N19/132 (2014.01) i, H04N19/184 (2014.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl. H04N19/00-19/98

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

Published examined utility model applications of Japan	1922-1996
Published unexamined utility model applications of Japan	1971-2019
Registered utility model specifications of Japan	1996-2019
Published registered utility model applications of Japan	1994-2019

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JP 9-149420 A (GRAPHICS COMMUN LAB KK) 06 June 1997, paragraphs [0030]-[0079], fig. 1-3 (Family: none)	1-4, 6, 8
A	US 2018/0173994 A1 (WAVEONE INC.) 21 June 2018, paragraphs [0035]-[0114], fig. 2-8 & US 2019/0266490 A1 & US 2018/0174047 A1 & US 2018/0174052 A1 & US 2018/0174275 A1 & US 2018/0176570 A1 & US 2018/0176576 A1 & US 2018/0176578 A1	1-8



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
03 December 2019 (03.12.2019)

Date of mailing of the international search report
17 December 2019 (17.12.2019)

Name and mailing address of the ISA/
Japan Patent Office
3-4-3, Kasumigaseki, Chiyoda-ku,
Tokyo 100-8915, Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2019/037254

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	NICK, Johnston et al., "Improved Lossy Image Compression with Priming and Spatially Adaptive Bit Rates for Recurrent Networks", arXiv, 29 March 2017, arXiv:1703.10114v1, pp. 1-9	1-8
A	MU, Li et al., "Learning Convolutional Networks for Content-weighted Image Compression", arXiv, 19 September 2017, arXiv:1703.10553v2, pp. 1-11	1-8
A	KAI, Han et al., "AUTOENCODER INSPIRED UNSUPERVISED FEATURE SELECTION", arXiv, 09 April 2018, arXiv:1710.08310v3, pp. 1-5	1-8

A. 発明の属する分野の分類（国際特許分類（IPC））

Int.Cl. H04N19/90(2014.01)i, H04N19/132(2014.01)i, H04N19/184(2014.01)i

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（IPC））

Int.Cl. H04N19/00-19/98

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1922-1996年
日本国公開実用新案公報	1971-2019年
日本国実用新案登録公報	1996-2019年
日本国登録実用新案公報	1994-2019年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
X	JP 9-149420 A（株式会社グラフィックス・コミュニケーション・ラボラトリーズ）1997.06.06, [0030] - [0079], 図1-3（ファミリーなし）	1-4, 6, 8
A	US 2018/0173994 A1 (WAVEONE INC.) 2018.06.21, [0035] - [0114], FIGs. 2-8 & US 2019/0266490 A1 & US 2018/0174047 A1 & US 2018/0174052 A1 & US 2018/0174275 A1 & US 2018/0176570 A1 & US 2018/0176576 A1 & US 2018/0176578 A1	1-8

☒ C欄の続きにも文献が列挙されている。

☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの
「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
「O」口頭による開示、使用、展示等に言及する文献
「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「&」同一パテントファミリー文献

国際調査を完了した日

03.12.2019

国際調査報告の発送日

17.12.2019

国際調査機関の名称及びあて先

日本国特許庁（ISA/J P）

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官（権限のある職員）

岩井 健二

5C

9465

電話番号 03-3581-1101 内線 3541

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
A	Nick Johnston et al., Improved Lossy Image Compression with Priming and Spatially Adaptive Bit Rates for Recurrent Networks, arXiv, 2017.03.29, arXiv:1703.10114v1, pp.1-9	1-8
A	Mu Li et al., Learning Convolutional Networks for Content-weighted Image Compression, arXiv, 2017.09.19, arXiv:1703.10553v2, pp.1-11	1-8
A	Kai Han et al., AUTOENCODER INSPIRED UNSUPERVISED FEATURE SELECTION, arXiv, 2018.04.09, arXiv:1710.08310v3, pp.1-5	1-8