

(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  
2015年9月11日(11.09.2015)



(10) 国際公開番号  
WO 2015/132914 A1

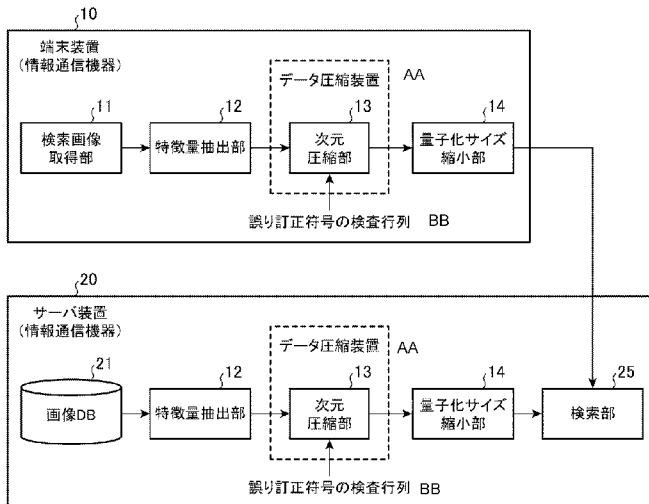
- (51) 国際特許分類:  
H03M 7/30 (2006.01)
- (21) 国際出願番号: PCT/JP2014/055671
- (22) 国際出願日: 2014年3月5日(05.03.2014)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (71) 出願人: 三菱電機株式会社(MITSUBISHI ELECTRIC CORPORATION) [JP/JP]; 〒1008310 東京都千代田区丸の内二丁目7番3号 Tokyo (JP).
- (72) 発明者: 松本 渉(MATSUMOTO, Wataru); 〒1008310 東京都千代田区丸の内二丁目7番3号 三菱電機株式会社内 Tokyo (JP). 山崎 貴司(YAMAZAKI, Takashi); 〒1008310 東京都千代田区丸の内二丁目7番3号 三菱電機株式会社内 Tokyo (JP).
- (74) 代理人: 田澤 英昭, 外(TAZAWA, Hideaki et al.); 〒1000014 東京都千代田区永田町二丁目12番4号 赤坂山王センタービル5階 Tokyo (JP).
- (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, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, 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, 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:

- 国際調査報告 (条約第 21 条(3))

(54) Title: DATA COMPRESSION APPARATUS AND DATA COMPRESSION METHOD

(54) 発明の名称: データ圧縮装置およびデータ圧縮方法



(57) Abstract: A dimension compression unit (13) compresses feature quantities extracted from an image by multiplying the feature quantities by a test matrix of error correction codes constituted by binary elements of 0 and 1.

(57) 要約: 次元圧縮部 13 は、画像から抽出された特徴量を、0 と 1 の 2 元の要素で構成される誤り訂正符号の検査行列と掛け算して圧縮する。

- 10 Terminal apparatus (information communication device)
- 11 To-be-retrieved image acquisition unit
- 12 Feature quantity extraction unit
- 13 Dimension compression unit
- 14 Quantization size reduction unit
- 20 Server apparatus (information communication device)
- 21 Image database
- 25 Retrieval unit
- AA Data compression apparatus
- BB Test matrix of error correction codes



WO 2015/132914 A1

## 明 細 書

発明の名称：データ圧縮装置およびデータ圧縮方法

### 技術分野

[0001] この発明は、データを短く圧縮するデータ圧縮装置およびデータ圧縮方法に関するものである。

### 背景技術

[0002] 画像、音声、およびセンサ等のデータそのもの、またはそのデータから抽出した特徴量を、乱数を要素に持つ行列と掛け算して、圧縮データを生成していた。次元圧縮に係るデータの要素数を次元数と呼ぶ。従来の次元圧縮では、 $m \times n$ の行列の要素を実数値とし、ランダムに選択した値を要素に用いて行列を構成していた。 $n$ は元のデータの次元数で、 $m$ が圧縮後のデータの次元数である。また、 $n \geq m$ である。

[0003] 上記のように画像、音声およびセンサ等のデータそのものまたはその特徴量を圧縮することにより、検索、認識、予知等に係るデータ処理量を削減し、高速化を実現し、一つの作業を短時間で処理可能にしたり、ある一定時間により多くの作業またはより複雑な作業を実行可能にしたりする。

[0004] 例えば、図4に示す画像マッチングシステムは、端末装置100で撮影した写真（検索画像101）等に近い画像を、サーバ装置200の画像データベース（以下、DB）201に保持されている多数のサンプル画像の中から検索するものである。このシステムにおいて、端末装置100の特徴量抽出部102は、検索画像101から特徴量を抽出する。特徴量抽出にSIFT（Scale-Invariant Feature Transform）を用いた場合、1つの画像につき複数個（一般に数十から数百個）の128バイト／個の特徴量ベクトルが抽出される。

[0005] 従来の次元圧縮部103は、ランダム写像行列を用いて、上記特徴量を次元圧縮する。ここで用いるランダム写像行列は式（1）、ランダム写像行列の要素 $a(r, c)$ は $a(r, c) \sim N(0, 1)$ で平均0、分散値1の正

規分布に従う。

$$[0006] \quad \mathbf{A} \in \mathbf{R}^{m \times n} \quad (1)$$

[0007] 特徴量を  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  とすると、圧縮後の特徴量  $\mathbf{y} = (y_1, y_2, \dots, y_m)$  は  $\mathbf{y}^T = \mathbf{A}\mathbf{x}^T$  により算出される。ここで、 $\mathbf{y}^T, \mathbf{x}^T$  の「 $T$ 」は転置を意味する。

[0008] 上記のように S I F T を用いた場合、特徴量ベクトル  $\mathbf{x}$  の要素  $x_i$  を 1 バイト = 8 ビットで表現し、 $n = 128$  と表現する。また、ランダム写像行列  $\mathbf{A}$  の各要素  $a(r, c)$  を 8 ビットで表現できるようにした場合、圧縮後の特徴量ベクトル  $\mathbf{y}$  の要素  $y_k$  は、式 (2) で計算できる。

$$[0009] \quad y_k = x_1 \cdot a(k, 1) + x_2 \cdot a(k, 2) + \dots + x_i \cdot a(k, i) + \dots + x_n \cdot a(k, n) \quad (2)$$

[0010] 従って、ランダム写像行列  $\mathbf{A}$  の要素と特徴量ベクトル  $\mathbf{x}$  の要素の掛け算で、2 バイト = 16 ビットに列数分の  $128 = 2^7$  個が加算されるので、最大  $16 + 7 = 23$  ビットが圧縮後の特徴量ベクトル  $\mathbf{y}$  の要素  $y_k$  の量子化サイズとなる。圧縮後の特徴量ベクトル  $\mathbf{y}$  の長さは、 $m$  次元のベクトルとして表現でき、次元を  $m = 40$  とした場合に圧縮前の 128 から 40 に圧縮できる。

[0011] さらに非特許文献 1 では、圧縮後の特徴ベクトル  $\mathbf{y}$  の要素  $y_k$  の量子化サイズを縮小する方式に関して検討しており、例えば上記の例で要素  $y_k$  のサイズを最大 23 ビットから 1 ~ 5 ビット程度まで圧縮する方式を示している。この方法では、特徴量間の距離を維持する条件で次元圧縮を行っている。

量子化サイズ縮小部 104 が上記非特許文献 1 の方法を用いて特徴量の量子化サイズを圧縮した場合、一つの特徴量が  $128 \times 8 = 1024$  ビットであったデータ量を、 $m = 40$ 、量子化サイズ 4 ビットに圧縮すると、圧縮後のデータ量は  $128 \times 40 / 128 \times 4 = 160$  ビットであり、 $160 / 1024 = 15.6\%$  までデータ量が圧縮される。

[0012] サーバ装置 200 側においても、特徴量抽出部 102 が画像 DB 201 に保存されているサンプル画像から特徴量を抽出し、次元圧縮部 103 および量子化サイズ縮小部 104 が特徴量を圧縮する。

[0013] 例えば、サーバ装置 200 は、画像 DB 201 にサンプル画像が 1000

0枚あれば、各サンプル画像に対して特徴量抽出とデータ圧縮を行い、検索部205が、端末装置100から送られてくる検索画像101の圧縮された特徴量と比較を行い、検索画像101に近いサンプル画像を検索する。

## 先行技術文献

### 非特許文献

- [0014] 非特許文献1: Mu Li, Shantanu Rane, Petros Boufounos, “Quantized Embeddings of Scale Invariant Image Features for Mobile Augmented Reality”, Multimedia Signal Processing (MMSP), 2012 IEEE 14th International Workshop on Digital Object Identifier, p.1-6

## 発明の概要

### 発明が解決しようとする課題

- [0015] しかしながら、上記方法の場合、次元圧縮のために  $y^T$  を計算する際、 $m \times n$  個の掛け算が必要となるので、圧縮前のデータ量が大きく、圧縮率が同程度の場合、指数関数的に計算量が増大するという課題があった。
- [0016] 例えば  $n = 128$ 、 $m = 40$  の場合、 $m \times n = 5120$  個の掛け算が必要となり、 $n = 1280$ 、 $m = 400$  の場合、 $m \times n = 512000$  個の掛け算が必要となる。このように、圧縮前のデータ量が10倍になると計算量が100倍になる。また、行列の要素をランダムに選ぶことにより、圧縮後の性質にばらつきが生じ、検索時の正解率が劣化するケースが発生する。
- [0017] この発明は、上記のような課題を解決するためになされたもので、データ圧縮時の計算量を削減することを目的とする。

### 課題を解決するための手段

- [0018] この発明に係るデータ圧縮装置は、情報通信機器から入手したデータまたはデータの特徴量を圧縮する際、データまたは特徴量と誤り訂正符号の検査行列との演算により圧縮データを生成するものである。
- [0019] この発明に係るデータ圧縮方法は、情報通信機器から入手したデータまたはデータの特徴量を圧縮する際、データまたは特徴量と誤り訂正符号の検査

行列との演算により圧縮データを生成するものである。

### 発明の効果

[0020] この発明によれば、データを圧縮する際、ランダム写像行列に誤り訂正符号の検査行列を用いるようにしたので、計算量を削減することができる。

### 図面の簡単な説明

[0021] [図1]この発明の実施の形態1に係るデータ圧縮装置を組み込んだ情報通信機器の構成を示すブロック図である。

[図2]実施の形態1に係るデータ圧縮装置を組み込んだ情報通信機器の動作を示すフローチャートである。

[図3]実施の形態1の次元圧縮部の圧縮次元を決定する方法を説明するためのグラフである。

[図4]従来のデータ圧縮方法を用いた画像マッチングシステムの構成を示すブロック図である。

### 発明を実施するための形態

[0022] 以下、この発明をより詳細に説明するために、この発明を実施するための形態について、添付の図面に従って説明する。

実施の形態1.

実施の形態1では、図1に示すように、この発明に係るデータ圧縮装置を組み込んだ情報通信機器（端末装置10、サーバ装置20）を用いて構成した画像マッチングシステムを例にして、データ圧縮方法を説明する。端末装置10は、タブレットPC（Personal Computer）、スマートフォン、監視カメラ等であり、検索画像取得部11、特徴量抽出部12、次元圧縮部13（データ圧縮装置）、量子化サイズ縮小部14を備えている。この端末装置10との間で通信可能なサーバ装置20は、画像DB21、特徴量抽出部12、次元圧縮部13（データ圧縮装置）、量子化サイズ縮小部14、検索部25を備えている。

[0023] 端末装置10、サーバ装置20のそれぞれは、不図示のCPU（Central Processing Unit）で構成されており、このCPU

が内部メモリに格納されたプログラムを実行することによって、特徴量抽出部 12、次元圧縮部 13、量子化サイズ縮小部 14、検索部 25としての機能を実現する。なお、次元圧縮部 13は専用の演算回路で構成してもよい。

画像DB 21は、HDD (Hard Disk Drive) 等によって構成されている。

[0024] 次に、図2に示すフローチャートを参照しながら、端末装置 10の詳細を説明する。

端末装置 10において、検索画像取得部 11は、カメラ等から受像した画像を、マッチング対象の検索画像として取り込み、特徴量抽出部 12へ出力する (ステップ S T 1)。

[0025] 特徴量抽出部 12は、検出画像の特徴を抽出して、次元圧縮部 13へ出力する (ステップ S T 2)。特徴量抽出方法としてはSIFT等の適用例が多いが、どのような特徴量抽出方法でも構わない。例えばSIFTでは、画像の中の数十から数百の特徴的なポイント (キーポイントと呼ぶ) が選択され、そのキーポイントごとに128バイトの特徴量ベクトルが出力される。ここでは、この128バイトの特徴量ベクトルの次元を128とし、128次元のベクトルと見なすことにする。

[0026] 次元圧縮部 13は、特徴量抽出部 12が出力した特徴量ベクトルを、誤り訂正符号の検査行列を用いて次元圧縮する (ステップ S T 3)。ここで用いる誤り訂正符号の検査行列は、2元 {0, 1} を要素に持つ行列であり、式 (11) とする。なお、誤り訂正符号の検査行列が、式 (12) のような非2元を要素に持つ行列で構成されていてもよい。

以下の説明では、誤り訂正符号の検査行列として、式 (11) のように2元 {0, 1} を要素に持つ、ランダム符号の検査行列を用いるものとする。

[0027] 
$$\mathbf{H} \in \mathbf{F}_2^{m \times n} \quad (11)$$

$$\mathbf{H} \in \mathbf{F}_p^{m \times n} \quad (12)$$

ここで、pは2以外の自然数である。

[0028] 特徴量抽出部 1 2 が抽出した特徴量を  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  とすると、次元圧縮部 1 3 が次元圧縮した圧縮後の特徴量は  $\mathbf{y} = (y_1, y_2, \dots, y_m) = \mathbf{H}\mathbf{x}$  により算出される。

[0029] 特徴量抽出部 1 2 が S I F T を用いた場合、特徴量ベクトル  $\mathbf{x}$  の要素  $x_i$  を 1 バイト = 8 ビットで表現し、 $n = 128$  として表現できる。また、検査行列  $\mathbf{H}$  の各要素  $h(r, c)$  は 1 ビットで表現でき、圧縮後の特徴量ベクトル  $\mathbf{y}$  の要素  $y_k$  は、式 (13) で計算できる。

$$[0030] \quad y_k = x_1 \cdot h(k,1) + x_2 \cdot h(k,2) + \dots + x_i \cdot h(k,i) + \dots + x_n \cdot h(k,n) \quad (13)$$

[0031] 2 元の要素で構成される検査行列の場合、要素が 0 の箇所は計算を削除し、1 が立った箇所のみ  $x_i$  を加算すればよい。従って、特徴量ベクトル  $\mathbf{x}$  の要素 8 ビットに対して、最大列数分の  $128 = 2^7$  個が加算されるので、最大  $8 + 7 = 15$  ビットが圧縮後の特徴量ベクトル  $\mathbf{y}$  の要素  $y_k$  の量子化サイズとなる。また、圧縮後の特徴量ベクトル  $\mathbf{y}$  の長さは  $m$  次元のベクトルとして表現でき、次元を  $m = 40$  とした場合に圧縮前の 128 から 40 に圧縮できる。

[0032] この際、誤り訂正符号の検査行列を使用しない従来の次元圧縮 (図 4 の次元圧縮部 1 0 3) との違いは、計算量である。

従来例では、圧縮後の特徴量ベクトル  $\mathbf{y}$  の要素  $y_k$  を一個求めるために  $n$  回の掛け算と  $n - 1$  回の加算が必要であり、この計算を特徴量ベクトル  $\mathbf{y}$  の長さ  $m$  回分繰り返す必要があるため、最終的に、 $nm$  回の掛け算と  $(n - 1)m$  回の加算が必要であった。

一方、実施の形態 1 では、ランダム符号化の場合、検査行列  $\mathbf{H}$  の各要素が 1 である確率と 0 である確率は共に  $1/2$  であるため、要素  $y_k$  を一個求めるために平均  $n/2 - 1$  回の加算でよく、この計算を特徴量ベクトル  $\mathbf{y}$  の長さ  $m$  回分繰り返しても、最終的に、平均  $(n/2 - 1)m$  回の加算で十分となる。

[0033] 今回の例の場合、従来例では、 $nm = 128 \times 40 = 5120$  回の掛け算と、 $(n - 1)m = 5080$  回の加算が必要となる。一方、実施の形態 1 では、 $(n/2 - 1)m = 2520$  回の加算のみでよくなり、大幅な計算量削減

減が可能となる。

[0034] この後、量子化サイズ縮小部 14 が、上記非特許文献 1 のように、圧縮後の特徴量ベクトル  $y$  の要素  $y_k$  の量子化サイズを縮小し、最大 15 ビットから 1 ~ 5 ビット程度まで圧縮する（ステップ S T 4）。この際、特徴量間の相対的な距離の大小関係をほぼ維持できる様に事前に評価を行い、距離の大小関係がほぼ維持可能な圧縮次元を決定しておく。

[0035] ここで、決定方法の一例を、図 3 を用いて説明する。装置の設計者は、200 種類の建物をそれぞれ 4 つの異なる角度から撮影した写真 800 枚（ $= 200 \times 4$ ）を用意し、同じ建物の写真には同じ ID 番号を割り振っておく。また、ターゲットとして、その 200 種類の建物のうちの 1 つを撮影した写真（先の 800 枚の写真とは異なる）を用意する。そして、 $k$ -近傍法（ターゲットに最も近い  $k$  個のデータを取り出して多数決を取る）により、ターゲットの建物に最も近い建物がうつった写真を 800 枚の中から検出する。図 3 はその検出結果を示し、縦軸が検出成功率、横軸が圧縮次元である。 $k$  を 10 と設定し、ターゲットの建物の特徴量と、800 枚の写真の建物の特徴量との距離を比較し、同じ建物の特徴量との距離が近い ID 番号が  $k = 10$  個の中で多数を占めた場合を検出成功とする。目標性能を検出成功率 95% とした場合、圧縮後の次元は 40 を選択できる。このような方法等により決定した圧縮次元が、次元圧縮部 13 に事前に設定される。

[0036] 一つの特徴量が  $128 \times 8 = 1024$  ビットであった元々のデータ量を、 $m = 40$  および量子化サイズ 4 ビットに圧縮する場合、圧縮後のデータ量は  $128 \times 40 / 128 \times 4 = 160$  ビットであり、 $160 / 1024 = 15.6\%$  までデータ量を圧縮できる。

[0037] 以上が、端末装置 10 によるデータ圧縮の方法である。

なお、上記説明では検索画像の特徴量を圧縮する例を示したが、検索画像そのものを圧縮してもよい。

[0038] 一方のサーバ装置 20 においても、画像 DB 21 が保存しているサンプル画像に対して、特徴量抽出部 12 が上記と同様の方法により特徴量を抽出し

、次元圧縮部 13 が上記と同様の方法により誤り訂正符号の検査行列を用いた次元圧縮を行い、量子化サイズ縮小部 14 が上記と同様の方法により量子化サイズの縮小を行う。この際、サーバ装置 20 側においても、次元圧縮の計算の際に、端末装置 10 側と同様の計算量の削減が可能となる。

[0039] 例えば、画像 DB 21 にサンプル画像が 10000 枚あれば、特徴量抽出部 12、次元圧縮部 13 および量子化サイズ縮小部 14 がそれぞれのサンプル画像に対して特徴量抽出とデータ圧縮を行い、圧縮した特徴量を検索部 25 へ出力する。

検索部 25 は、端末装置 10 から送られてくる検索画像の圧縮された特徴量と、量子化サイズ縮小部 14 から入力される各サンプル画像の圧縮された特徴量を比較し、検索画像と近いサンプル画像を検索する。

[0040] ここで、次元圧縮部 13 による圧縮は、特徴量間の相対的な距離の大小関係をほぼ維持しているため、検索部 25 は、k-近傍法（ターゲットに最も近い k 個のデータを取り出して多数決を取る）等の手法により、端末装置 10 の検索画像が画像 DB 21 に保存されているサンプル画像のどれに近いかを判別できる。その結果、圧縮による正解率の劣化を抑えられる特徴がある。この場合、k-近傍法の検索対象は圧縮された特徴量となるため、単純な比較による検索では 15.6% まで処理時間を短縮できる。

[0041] 以上より、実施の形態 1 によれば、次元圧縮部 13 は、データまたは当該データから抽出した特徴量を、誤り訂正符号の検査行列と掛け算して、圧縮データを生成するようにしたので、計算量を削減することができる。

特に、誤り訂正符号の検査行列に、0 と 1 の 2 元の要素から構成されるランダム符号の検査行列を用いることにより、圧縮のための計算量を、従来のランダム写像行列を用いた場合の  $nm$  回の掛け算と  $(n-1)m$  回の足し算から、 $(n/2-1)m$  回の足し算のみに削減できる。

[0042] 実施の形態 2.

図 1 に示した画像マッチングシステムを援用して、実施の形態 2 に係るデータ圧縮装置を説明する。

この実施の形態2では、次元圧縮部13（データ圧縮装置）が次元圧縮に用いる誤り訂正符号の検査行列として、上式（11）のように2元{0, 1}の要素で構成される、LDPC（Low-Density Parity-Check）符号の検査行列を用いるものとする。

なお、説明は省略するが、上式（12）のように非2元の要素で構成される、LDPC符号の検査行列を用いてもよい。

[0043] LDPC符号は、一般に列の平均重みが4であり、行の平均重みは（列の平均重み） $\times n/m$ である。ここで、列の重みとは、行列の1列に含まれる1の数である。また、行の重みとは、行列の1行に含まれる1の数である。例えば、 $n=128$ 、 $m=40$ の場合、行の重みは $4 \times 128 / 40 = 12.8$ となる。LDPC符号の場合、 $n$ または $m$ が大きくなってもこれら列の重み、行の重みは変化しない特徴がある。

[0044] 上記実施の形態1で説明したランダム符号の検査行列の場合、列内の1の数は平均 $n/2$ 、行内の1の数は平均 $m/2$ であり、LDPC符号の検査行列と比例して列、行ともに1の数が多くなり、かつ行列内の1の総数は2乗されるのでこの総数も多くなる。一方、LDPC符号の検査行列の場合、1の数は常に一定で、かつ疎であることから、ランダム符号の検査行列に比べると圧倒的に1の総数が少なくなる。

[0045] 例えば、 $n=128$ 、 $m=40$ の場合、ランダム符号では1の総数は、（列の平均重み） $\times$ （列数） $= 40/2 \times 128 = 2560$ 個である。一方、LDPC符号の場合、1の総数は、（列の平均重み） $\times$ （列数） $= 4 \times 128 = 516$ 個となり、圧倒的に少ない。

ここで一例として、 $n=28$ 、 $m=21$ のLDPC符号の検査行列Hを、式（14）に示す。

[0046]



ランダム符号の検査行列を用いるよりも更に少ない。

[0049] 以上より、実施の形態2によれば、次元圧縮部13は、データまたは当該データから抽出した特徴量を圧縮する際に、疎なLDPC符号の検査行列を用いることにより、計算量を大幅に削減できる。また、LDPC符号の検査行列は、要素が規則的なので、圧縮の性能のばらつきを抑える効果も期待できる。さらに、LDPC符号の検査行列は、圧縮前の特徴量の次元数 $n$ が大きくなるほど圧縮効率が良くなるため、特に数百以上の長い特徴量を扱う際に有効となる。

[0050] なお、上記実施の形態1ではランダム符号、上記実施の形態2ではLDPC符号の検査行列を用いる例を示したが、その他に、 BCH符号、リードソロモン符号、巡回符号等を用いてもよい。これらの符号は短い特徴量を扱う際に成功率の低下を抑制しながら圧縮できる特徴がある。また、上記に列挙した誤り訂正符号以外でもよく、条件に応じて適切な誤り訂正符号を適用できる。

[0051] また、上記実施の形態1, 2では検査行列として、2元 $\{0, 1\}$ の要素で構成される行列を例示したが、3つの要素 $\{-1, 0, 1\}$ で構成される行列でもよい。例えば、上式(14)の行列のうち、0はそのまま0にし、1の部分をも $-1$ または $1$ に割り振った下式(15)の行列でもよい。式(15)のような検査行列を用いる場合、次元圧縮部13は、データまたは当該データから抽出した特徴量において検査行列の要素が1または $-1$ の箇所に対応する値を行単位で足し算または引き算して、圧縮データを生成する。

検査行列の要素を2元 $\{0, 1\}$ にするか3つの要素 $\{-1, 0, 1\}$ にするかは、計算量自体あまり大きな差は無い為、性能評価により良好な性能を示す方を選択すればよい。

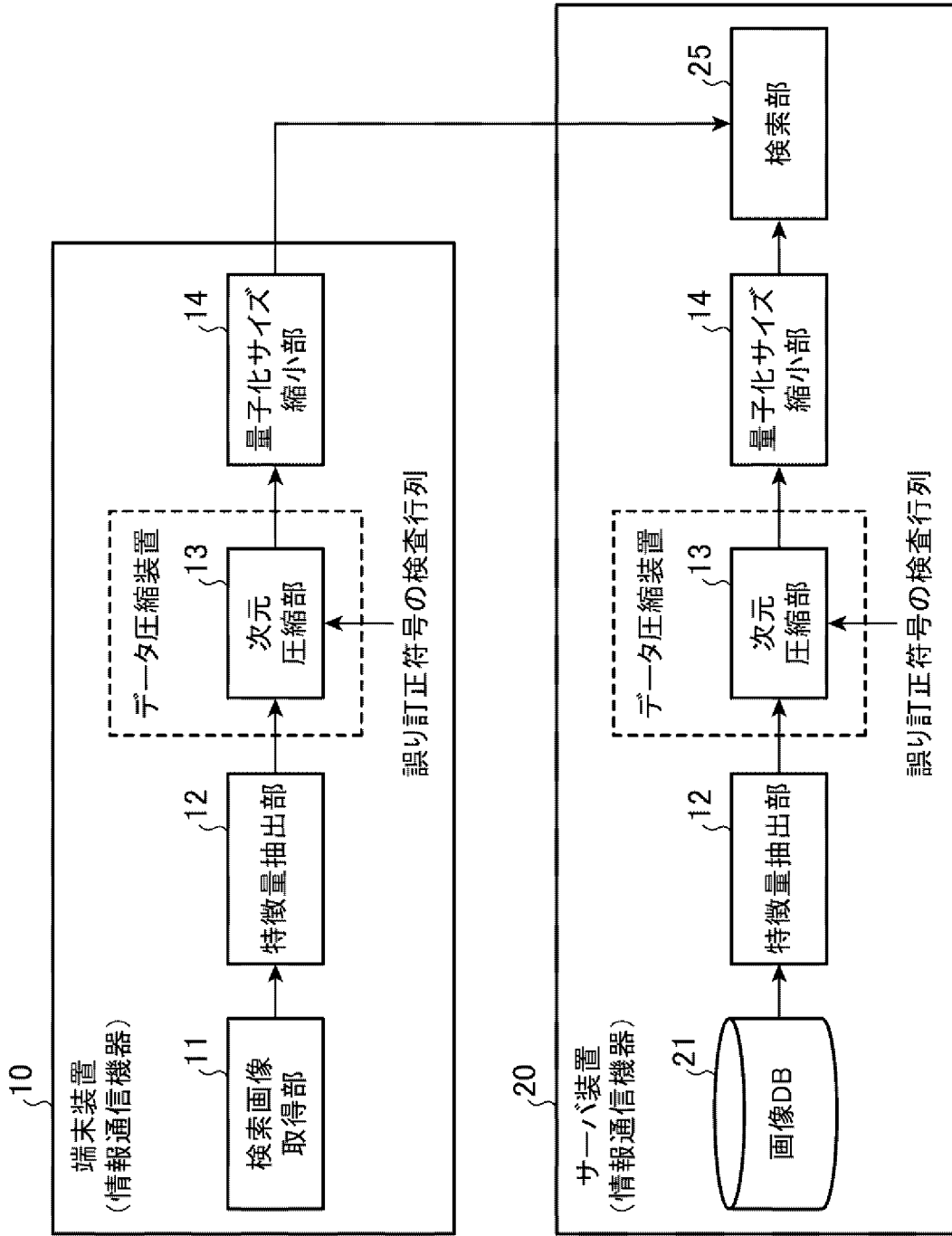
[0052]



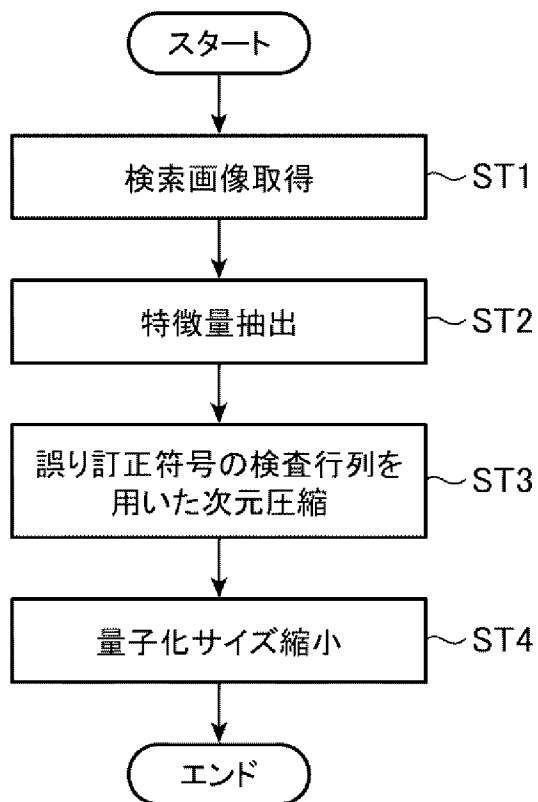
## 請求の範囲

- [請求項1] 情報通信機器から入手したデータまたは前記データの特徴量を圧縮するデータ圧縮装置であって、  
前記データまたは前記特徴量と誤り訂正符号の検査行列との演算により圧縮データを生成することを特徴とするデータ圧縮装置。
- [請求項2] 前記誤り訂正符号の検査行列に、0と1の2元の要素から構成される検査行列を用い、前記データまたは前記特徴量において前記検査行列の要素が1の箇所に対応する値を行単位で足し算して、圧縮データを生成することを特徴とする請求項1記載のデータ圧縮装置。
- [請求項3] 前記誤り訂正符号の検査行列に、0と1と-1の3つの要素から構成される検査行列を用い、前記データまたは前記特徴量において前記検査行列の要素が1または-1の箇所に対応する値を行単位で足し算または引き算して、圧縮データを生成することを特徴とする請求項1記載のデータ圧縮装置。
- [請求項4] 前記誤り訂正符号の検査行列に、ランダム符号の検査行列を用いることを特徴とする請求項2記載のデータ圧縮装置。
- [請求項5] 前記誤り訂正符号の検査行列に、LDPC符号の検査行列を用いることを特徴とする請求項2記載のデータ圧縮装置。
- [請求項6] 前記誤り訂正符号の検査行列に、BCH符号の検査行列を用いることを特徴とする請求項2記載のデータ圧縮装置。
- [請求項7] 前記誤り訂正符号の検査行列に、リードソロモン符号の検査行列を用いることを特徴とする請求項2記載のデータ圧縮装置。
- [請求項8] 前記誤り訂正符号の検査行列に、巡回符号の検査行列を用いることを特徴とする請求項2記載のデータ圧縮装置。
- [請求項9] 情報通信機器から入手したデータまたは前記データの特徴量を圧縮するデータ圧縮方法であって、  
前記データまたは前記特徴量と誤り訂正符号の検査行列との演算により圧縮データを生成することを特徴とするデータ圧縮方法。

[図1]

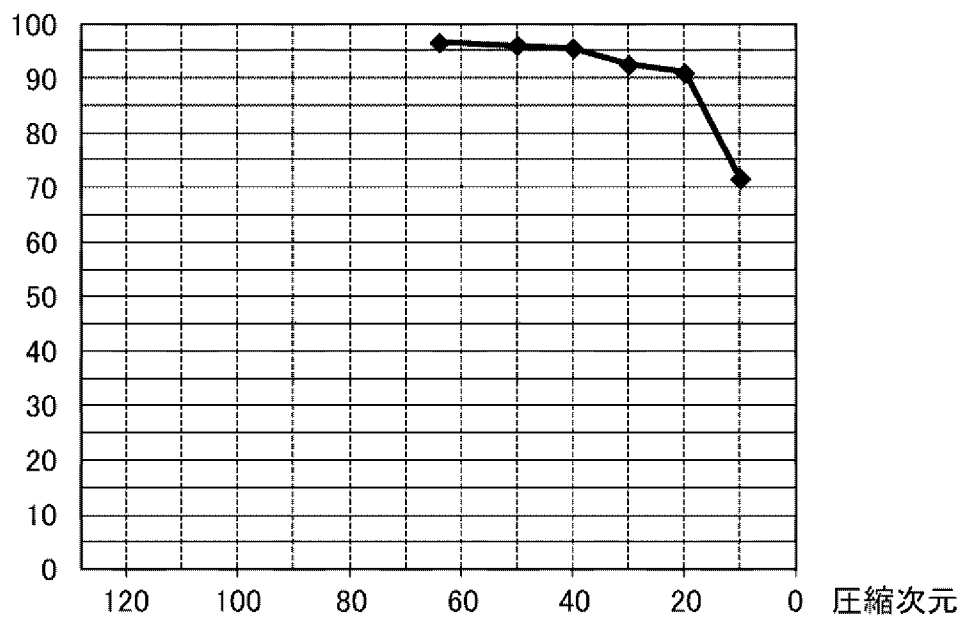


[図2]

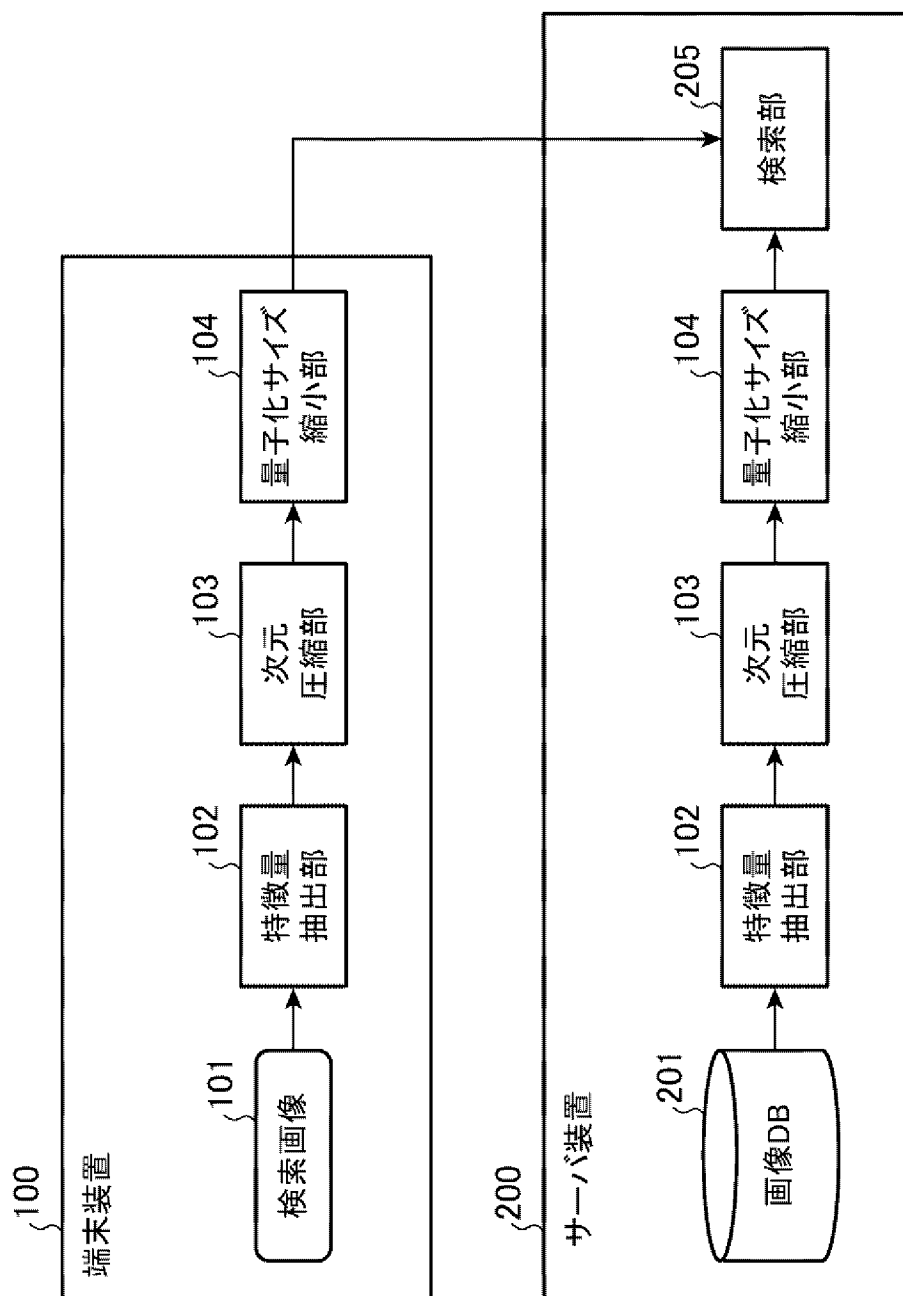


[図3]

検出成功率(%)



[図4]



**INTERNATIONAL SEARCH REPORT**

International application No.  
PCT/JP2014/055671

**A. CLASSIFICATION OF SUBJECT MATTER**  
H03M7/30(2006.01) 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)  
H03M7/30

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

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2014
Kokai Jitsuyo Shinan Koho	1971-2014	Toroku Jitsuyo Shinan Koho	1994-2014

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

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X A	JP 2006-217363 A (Japan Aerospace Exploration Agency), 17 August 2006 (17.08.2006), paragraphs [0007] to [0011] (Family: none)	1, 2, 4-9 3
A	JP 2011-77958 A (National University Corporation Chiba University), 14 April 2011 (14.04.2011), entire text; all drawings (Family: none)	1-9
A	JP 2003-204316 A (NEC Corp.), 18 July 2003 (18.07.2003), entire text; all drawings (Family: none)	3

Further documents are listed in the continuation of Box C.       See patent family annex.

* Special categories of cited documents:	"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
"A" document defining the general state of the art which is not considered to be of particular relevance	"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
"E" earlier application or patent but published on or after the international filing date	"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
"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)	"&" document member of the same patent family
"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	

Date of the actual completion of the international search 28 March, 2014 (28.03.14)	Date of mailing of the international search report 08 April, 2014 (08.04.14)
--	---

Name and mailing address of the ISA/ Japanese Patent Office	Authorized officer
Facsimile No.	Telephone No.

A. 発明の属する分野の分類 (国際特許分類 (IPC)) Int.Cl. H03M7/30(2006.01)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. H03M7/30		
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 1971-2014年 日本国実用新案登録公報 1996-2014年 日本国登録実用新案公報 1994-2014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) IEEE Xplore		
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 2006-217363 A (独立行政法人 宇宙航空研究開発機構) 2006.08.17, 段落【0007】～【0011】 (ファミリーなし)	1, 2, 4-9
A		3
A	JP 2011-77958 A (国立大学法人 千葉大学) 2011.04.14, 全文, 全 図 (ファミリーなし)	1-9
A	JP 2003-204316 A (日本電気株式会社) 2003.07.18, 全文, 全図 (フ ァミリーなし)	3
<input type="checkbox"/> C欄の続きにも文献が列挙されている。 <input type="checkbox"/> 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 28.03.2014	国際調査報告の発送日 08.04.2014	
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) 北村 智彦 電話番号 03-3581-1101 内線 3556	5 K 9 2 9 7