

(19)日本国特許庁(JP)

(12)公表特許公報(A)

(11)公表番号
特表2025-506100
(P2025-506100A)

(43)公表日 令和7年3月7日(2025.3.7)

(51)国際特許分類
H 0 4 N 19/13 (2014.01)
H 0 4 N 19/136 (2014.01)

F I
H 0 4 N 19/13
H 0 4 N 19/136

テーマコード (参考)
5 C 1 5 9

審査請求 未請求 予備審査請求 未請求 (全40頁)

(21)出願番号	特願2024-543356(P2024-543356)	(71)出願人	507364838 クアルコム、インコーポレイテッド アメリカ合衆国 カリフォルニア 9 2 1 2 1 サン ディエゴ モアハウス ドライ ブ 5 7 7 5
(86)(22)出願日	令和5年1月11日(2023.1.11)	(74)代理人	100108453 弁理士 村山 靖彦
(85)翻訳文提出日	令和6年7月22日(2024.7.22)	(74)代理人	100163522 弁理士 黒田 晋平
(86)国際出願番号	PCT/US2023/060460	(72)発明者	アミール・サイド アメリカ合衆国・カリフォルニア・9 2 1 2 1-1 7 1 4・サン・ディエゴ・モ アハウス・ドライヴ・5 7 7 5
(87)国際公開番号	WO2023/154590	(72)発明者	インハオ・ジュ アメリカ合衆国・カリフォルニア・9 2
(87)国際公開日	令和5年8月17日(2023.8.17)		最終頁に続く
(31)優先権主張番号	17/650,728		
(32)優先日	令和4年2月11日(2022.2.11)		
(33)優先権主張国・地域又は機関	米国(US)		
(81)指定国・地域	AP(BW,CV,GH,GM,KE,LR,LS,MW,MZ ,NA,RW,SD,SL,ST,SZ,TZ,UG,ZM,ZW), EA(AM,AZ,BY,KG,KZ,RU,TJ,TM),EP(AL,AT,BE,BG,CH,CY,CZ,DE,DK,EE,ES, FI,FR,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV 最終頁に続く		

(54)【発明の名称】 ニューラルベースのメディア圧縮のためのエントロピーコーディング

(57)【要約】

本開示は、ニューラルベースの技法を使用してコーディングされたメディアデータのためのエントロピーコーディング技法について説明する。メディアコードは、ニューラルベースのメディア圧縮技法によってコーディングされたデータストリームのデータ要素の確率分布関数パラメータを決定し、確率分布関数パラメータは、データストリームの確率分布関数の標準偏差の対数関数であり、確率分布関数パラメータに基づいてコードベクトルを決定し、コードベクトルを使用してデータ要素をエントロピーコーディングするように構成されている。

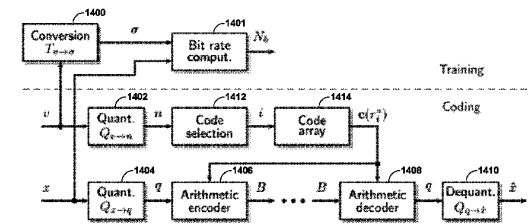


FIG. 14

【特許請求の範囲】

【請求項 1】

メディアデータをコーディングする方法であって、前記方法は、

ニューラルベースのメディア圧縮技法によってコーディングされたデータストリームのデータ要素の確率分布関数パラメータであって、前記確率分布関数パラメータは、前記データストリームの確率分布関数の標準偏差の対数関数である、確率分布関数パラメータを決定することと、

前記確率分布関数パラメータに基づいてコードベクトルを決定することと、

前記コードベクトルを使用して前記データ要素をエントロピーコーディングすることと

、

を含む、方法。

10

【請求項 2】

前記コードベクトルを決定する前に、前記確率分布関数パラメータを量子化することを更に含む、請求項 1 に記載の方法。

【請求項 3】

前記コードベクトルを決定することは、

前記確率分布関数パラメータに基づいてコードベクトルインデックスを決定することと

、

前記コードベクトルインデックスに基づいて、コードベクトルのテーブルから前記コードベクトルを決定することと、を含む、請求項 1 に記載の方法。

20

【請求項 4】

前記確率分布関数パラメータは v であり、前記確率分布関数の前記標準偏差は σ であり、最小標準偏差は $\sigma_{\min}$ であり、最大標準偏差は $\sigma_{\max}$ であり、 v は、

【数 1】

$$v = \frac{\ln(\sigma) - \ln(\sigma_{\min})}{\ln(\sigma_{\max}) - \ln(\sigma_{\min})}$$

と定義されている、請求項 1 に記載の方法。

【請求項 5】

前記ニューラルベースの圧縮技法を使用して前記データ要素を生成することと、

量子化データ要素を作成するために、前記データ要素を量子化することと、を更に含む

、

前記コードベクトルを使用して前記データ要素をエントロピーコーディングすることは、前記コードベクトルを使用して前記量子化データ要素をエントロピー符号化することを含む、請求項 1 に記載の方法。

30

【請求項 6】

前記ニューラルベースの圧縮技法を使用して前記データ要素を生成することは、

前記データ要素を生成するために、画像分析ニューラルネットワークを使用して画像又はビデオピクチャを処理することを含む、請求項 5 に記載の方法。

40

【請求項 7】

前記コードベクトルを使用して前記データ要素をエントロピーコーディングすることは、量子化データ要素を作成するために、前記コードベクトルを使用して符号化されたデータ要素をエントロピー復号することを含み、前記方法は、

再構築されたデータ要素を作成するために、前記量子化データ要素を逆量子化することを更に含む、請求項 1 に記載の方法。

【請求項 8】

画像又はビデオピクチャを再構築するために、画像合成ニューラルネットワークを使用して前記再構築されたデータ要素を処理することを更に含む、請求項 7 に記載の方法。

【請求項 9】

50

メディアデータをコーディングするように構成されている装置であって、前記装置は、メモリと、
前記メモリと通信している１つ又は複数のプロセッサと、
を備え、前記１つ又は複数のプロセッサが、
ニューラルベースのメディア圧縮技法によってコーディングされたデータストリームのデータ要素の確率分布関数パラメータであって、前記確率分布関数パラメータは、前記データストリームの確率分布関数の標準偏差の対数関数である、確率分布関数パラメータを決定し、

前記確率分布関数パラメータに基づいてコードベクトルを決定し、
前記コードベクトルを使用して前記データ要素をエントロピーコーディングする、ように構成されている、装置。

10

【請求項１０】

前記１つ又は複数のプロセッサは、
前記コードベクトルを決定する前に、前記確率分布関数パラメータを量子化するように更に構成されている、請求項９に記載の装置。

【請求項１１】

前記コードベクトルを決定するために、前記１つ又は複数のプロセッサは、
前記確率分布関数パラメータに基づいてコードベクトルインデックスを決定し、
前記コードベクトルインデックスに基づいて、コードベクトルのテーブルから前記コードベクトルを決定する、ように更に構成されている、請求項９に記載の装置。

20

【請求項１２】

前記確率分布関数パラメータは v であり、前記確率分布関数の前記標準偏差は σ であり、最小標準偏差は $\sigma_{\min}$ であり、最大標準偏差は $\sigma_{\max}$ であり、 v は、

【数２】

$$v = \frac{\ln(\sigma) - \ln(\sigma_{\min})}{\ln(\sigma_{\max}) - \ln(\sigma_{\min})}$$

と定義されている、請求項９に記載の装置。

【請求項１３】

前記１つ又は複数のプロセッサは、
前記ニューラルベースの圧縮技法を使用して前記データ要素を生成し、
量子化データ要素を作成するために、前記データ要素を量子化する、ように更に構成されており、
前記コードベクトルを使用して前記データ要素をエントロピーコーディングすることは、
前記コードベクトルを使用して前記量子化データ要素をエントロピー符号化することを含む、請求項９に記載の装置。

30

【請求項１４】

前記ニューラルベースの圧縮技法を使用して前記データ要素を生成するために、前記１つ又は複数のプロセッサは、
前記データ要素を生成するために、画像分析ニューラルネットワークを使用して画像又はビデオピクチャを処理するように更に構成されている、請求項１３に記載の装置。

40

【請求項１５】

前記画像又はビデオピクチャをキャプチャするように構成されたカメラを更に備える、請求項１４に記載の装置。

【請求項１６】

前記コードベクトルを使用して前記データ要素をエントロピーコーディングすることは、
量子化データ要素を作成するために、前記コードベクトルを使用して符号化されたデータ要素をエントロピー復号することを含み、前記１つ又は複数のプロセッサは、

再構築されたデータ要素を作成するために、前記量子化データ要素を逆量子化するよう

50

に更に構成されている、請求項 9 に記載の装置。

【請求項 17】

前記 1 つ又は複数のプロセッサは、

画像又はビデオピクチャを再構築するために、画像合成ニューラルネットワークを使用して前記再構築されたデータ要素を処理するように更に構成されている、請求項 16 に記載の装置。

【請求項 18】

前記画像又はビデオピクチャを表示するように構成されたディスプレイを更に備える、請求項 17 に記載の装置。

【請求項 19】

命令を記憶する非一時的コンピュータ可読記憶媒体であって、前記命令が、実行されると、1 つ又は複数のプロセッサに、

ニューラルベースのメディア圧縮技法によってコーディングされたデータストリームのデータ要素の確率分布関数パラメータであって、前記確率分布関数パラメータは、前記データストリームの確率分布関数の標準偏差の対数関数であり、確率分布関数パラメータを決定させ、

前記確率分布関数パラメータに基づいてコードベクトルを決定させ、

前記コードベクトルを使用して前記データ要素をエントロピーコーディングさせる、非一時的コンピュータ可読記憶媒体。

【請求項 20】

前記確率分布関数パラメータは v であり、前記確率分布関数の前記標準偏差は σ であり、最小標準偏差は $\sigma_{\min}$ であり、最大標準偏差は $\sigma_{\max}$ であり、 v は、

【数 3】

$$v = \frac{\ln(\sigma) - \ln(\sigma_{\min})}{\ln(\sigma_{\max}) - \ln(\sigma_{\min})}$$

と定義されている、請求項 19 に記載の非一時的コンピュータ可読記憶媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本出願は、内容全体が参照により本明細書に組み込まれる、2022 年 2 月 11 日に提出された米国特許出願第 17 / 650, 728 号の優先権を主張する。

【0002】

本開示は、画像及びビデオデータの符号化及び復号を含む、メディア符号化及び復号に関する。

【背景技術】

【0003】

デジタルメディア能力は、デジタルテレビジョン、デジタルダイレクトブロードキャストシステム、ワイヤレスブロードキャストシステム、携帯情報端末 (personal digital assistant、PDA)、ラップトップ又はデスクトップコンピュータ、タブレットコンピュータ、電子ブックリーダー、デジタルカメラ、デジタル記録デバイス、デジタルメディアプレーヤ、ビデオゲーミングデバイス、ビデオゲームコンソール、セルラー又は衛星無線電話、いわゆる「スマートフォン」、ビデオ遠隔会議デバイス、ビデオストリーミングデバイスなどを含む、広範囲にわたるデバイスの中に組み込まれ得る。デジタルビデオデバイスは、MPEG-2、MPEG-4、ITU-T H. 263、ITU-T H. 264 / MPEG-4、Part 10、アドバンスドビデオコーディング (Advanced Video Coding、AVC)、ITU-T H. 265 / 高効率ビデオコーディング (High Efficiency Video Coding、HEVC)、ITU-T H. 266 / バーサタイルビデオコーディング (Versatile Video Coding、VVC) によって定義された規格、及びそ

のような規格の拡張、並びに Alliance for Open Media によって開発された AOMedia Video 1 (AOMedia Video 1、AV1) などのプロプライエタリビデオコーデック／フォーマットに記載されているビデオコーディング技法などの、ビデオコーディング技法を実装する。ビデオデバイスは、そのようなビデオコーディング技法を実装することによって、デジタルビデオ情報をより効率的に送信、受信、符号化、復号、及び／又は記憶することができる。

【発明の概要】

【0004】

概して、本開示は、ビデオ及び／又は画像の符号化及び復号のための技法を含む、メディア圧縮のための技法について説明する。ニューラルネットワークに基づくメディア（例えば、画像及び／又はビデオ）圧縮方法は、現在の規格に対して競争力があり、いくつかの追加の利点を提供することができる。ニューラルベースのコーディング方法は、一般に、高精度浮動小数点演算を使用して設計され、テストされる。しかしながら、技術が実際の展開に移行するにつれて、ニューラルネットワーク重み及び活性化関数は、典型的には、速度及び電力消費を改善するために量子化され、低精度整数で表される。

10

【0005】

本開示は、エントロピーコーディングに関係するニューラルネットワーク変数が量子化されるときに生じる問題に対処する。ニューラルネットワーク変数は、そのような変数が圧縮効率を定義するので、ニューラルベースのビデオ／画像圧縮方式の設計にとって重要である。更に、エントロピーコーディング変数の非常に具体的な特性は、ニューラルネットワークにおける量子化を最適化するための一般的なツールによって考慮されない。テストにより、最悪の量子化効果は、最も一般的なユースケースのいくつかで正確に発生する可能性があり、最悪の量子化効果によって引き起こされる損失は、ニューラルネットワークを再トレーニングすることによって回復することができないことが示された。

20

【0006】

本開示は、トレーニングされたエントロピーコーディング変数の定義を最適化するための技法について説明し、それにより、効果的なエントロピーコーディングのために最も重要な情報が、低精度整数で表されるときに最良に保存される。テストはまた、エントロピーコーディングに必要なメモリの量を最小化するために、本明細書で説明される技法がどのように使用され得るかを示した。本開示は、エントロピーコーディング設計のための一般的な手法、並びに一般に使用されるガウス分布のための特定の解決策及び実装形態について説明する。本開示の技法は、概して、任意のニューラルベースの圧縮技法に適用され得るが、以下で説明する例は、画像及びビデオのための技法に焦点を当てている。

30

【0007】

一例では、方法は、ニューラルベースのメディア圧縮技法によってコーディングされたデータストリームのデータ要素の確率分布関数パラメータであって、確率分布関数パラメータは、データストリームの確率分布関数の標準偏差の対数関数である、確率分布関数パラメータを決定することと、確率分布関数パラメータに基づいてコードベクトルを決定することと、コードベクトルを使用してデータ要素をエントロピーコーディングすることとを含む。

40

【0008】

別の例では、デバイスは、メモリと、メモリと通信している1つ又は複数のプロセッサを含み、1つ又は複数のプロセッサが、ニューラルベースのメディア圧縮技法によってコーディングされたデータストリームのデータ要素の確率分布関数パラメータであって、確率分布関数パラメータは、データストリームの確率分布関数の標準偏差の対数関数である、確率分布関数パラメータを決定し、確率分布関数パラメータに基づいてコードベクトルを決定し、コードベクトルを使用してデータ要素をエントロピーコーディングする、ように構成されている。

【0009】

別の例では、コンピュータ可読記憶媒体は命令で符号化されており、命令は、実行され

50

ると、プログラマブルプロセッサに、ニューラルベースのメディア圧縮技法によってコーディングされたデータストリームのデータ要素の確率分布関数パラメータであって、確率分布関数パラメータは、データストリームの確率分布関数の標準偏差の対数関数であり、確率分布関数パラメータを決定させ、確率分布関数パラメータに基づいてコードベクトルを決定させ、コードベクトルを使用してデータ要素をエントロピーコーディングさせる。

【0010】

1つ又は複数の例の詳細が、添付図面及び以下の説明に記載されている。他の特徴、目的、及び利点が、明細書、図面、及び特許請求の範囲から明らかになるであろう。

【図面の簡単な説明】

【0011】

10

【図1】本開示の技法を実行し得る例示的なメディア符号化及び復号システムを示すブロック図である。

【図2】従来のメディアコーディングとニューラルベースのメディアコーディングとの違いを示す。

【図3】ガウス連続確率分布関数の一例を示す。

【図4】例示的なニューラルベースの画像コーデックを示すブロック図である。

【図5】ニューラルベースのエントロピーコーディングの例示的な実装形態を示す。

【図6】量子化された確率分布関数パラメータを用いたニューラルベースのエントロピーコーディングの例示的な実装形態を示す。

【図7】異なるコードベクトルに対する量子化された確率分布関数パラメータに対する冗長性の例示的な尺度を示す。

20

【図8】コードベクトルの非一様な割り当ての一例を示す。

【図9】ビットレートによるガウス標準偏差の分布の一例を示す。

【図10】量子化されたランダム変数のエントロピーの一例を示す。

【図11】ガウス標準偏差の一樣量子化を使用することからの例示的な平均相対的冗長性を示す。

【図12】対数スケール上の量子化されたランダム変数の例示的なエントロピーを示す。

【図13】対数スケール上の量子化されたランダム変数の別の例示的なエントロピーを示す。

【図14】別の例示的なニューラルベースのコーデックのブロック図を示す。

30

【図15】確率分布関数パラメータの一樣量子化を使用する平均相対的冗長性の一例を示す。

【図16】量子化された確率分布関数パラメータをコードベクトルに変換するための例示的なコードを示す。

【図17】確率分布関数パラメータの一樣量子化を使用する相対的冗長性の一例を示す。

【図18】量子化された確率分布関数パラメータをコードベクトルに変換するための別の例示的なコードを示す。

【図19】本開示の例となる方法を例示するフローチャートである。

【発明を実施するための形態】

【0012】

40

本開示は、ニューラルネットワークベースのメディアコーディング技法を使用してメディアデータ（例えば、画像又はビデオ）を符号化及び復号するための技法について説明する。特に、本開示は、最適化された確率分布関数（probability distribution function、PDF）パラメータ化を使用するメディアデータのエントロピーコーディングのための技法について説明する。本開示の技法は、ニューラルネットワークをトレーニングするために使用されるPDFパラメータ（例えば、一般に標準偏差又は統計分布を定義するために使用される任意のパラメータ）の非一様（例えば、対数）関数である量子化された修正PDFパラメータの定義及び実装形態を含む。本開示のPDFパラメータを使用することによって、ニューラルベースのコーディング技法を使用するメディアデータのエントロピーコーディングは、より少ないメモリを使用して、より高い効率で実行され得、ニュ

50

ーラルネットワーク重み、活性化、及び出力が低精度整数によって表されるとき、ニューラルベースのコーディング技法の性能は、はるかに低下しない。

【0013】

図1は、本開示の技法を実行し得る例示的なメディア符号化及び復号システム100を示すブロック図である。本開示のコンテキストでは、メディアは、ビデオデータ及び／又は画像を含む、圧縮される任意のデジタルファイルを含み得る。本開示の例示的な技法は、一般に、ビデオデータ及び／又は画像をコーディングする（符号化及び／又は復号）することを対象とする。図1の例はメディア符号化及び復号を参照して説明されるが、本出願の技法は、ニューラルベースの圧縮技法を使用する任意のタイプのデータファイルの符号化及び復号に等しく適用可能である。

10

【0014】

図1に示されるように、この例では、システム100は、デスティネーションデバイス116によって復号され表示されるべき、符号化メディアデータを提供するソースデバイス102を含む。具体的には、ソースデバイス102は、コンピュータ可読媒体110を介してデスティネーションデバイス116にメディアデータを提供する。ソースデバイス102及びデスティネーションデバイス116は、デスクトップコンピュータ、ノートブック（すなわち、ラップトップ）コンピュータ、モバイルデバイス、タブレットコンピュータ、セットトップボックス、スマートフォンなどの電話ハンドセット、テレビジョン、カメラ、表示デバイス、デジタルメディアプレーヤ、ビデオゲーミングコンソール、ビデオストリーミングデバイス、ブロードキャスト受信機デバイスなどを含む、広範囲にわたるデバイスのうちのいずれかを備えてもよい。場合によっては、ソースデバイス102及びデスティネーションデバイス116は、ワイヤレス通信に対応し得るので、ワイヤレス通信デバイスと称され得る。

20

【0015】

図1の例では、ソースデバイス102は、メディアソース104、メモリ106、メディアエンコーダ200、及び出力インターフェース108を含む。デスティネーションデバイス116は、入力インターフェース122、メディアデコーダ300、メモリ120、及び表示デバイス118を含む。本開示によれば、ソースデバイス102のメディアエンコーダ200と、デスティネーションデバイス116のメディアデコーダ300とは、エントロピーコーディング及びニューラルベースのメディア圧縮システムのための技法を適用するように構成されている場合がある。したがって、ソースデバイス102はメディア符号化デバイスの一例を表し、一方でデスティネーションデバイス116はメディア復号化デバイスの一例を表す。他の例では、ソースデバイス及びデスティネーションデバイスは、他の構成要素又は構成を含み得る。例えば、ソースデバイス102は、外部カメラなどの外部メディアソースからメディアデータを受信してもよい。同様に、デスティネーションデバイス116は、一体型表示デバイスを含むのではなく、外部表示デバイスとインターフェースしてもよい。

30

【0016】

図1に示すようなシステム100は一例にすぎない。概して、いかなるデジタルメディア符号化及び／又は復号デバイスも、ニューラルベースのメディア圧縮システムをエントロピーコーディングするための技法を実施し得る。ソースデバイス102及びデスティネーションデバイス116は、ソースデバイス102がデスティネーションデバイス116へ送信するためのコーディングされたメディアデータを生成するような、コーディングデバイスの一例にすぎない。本開示は、データのコーディング（符号化及び／又は復号）を実行するデバイスとして「コーディング」デバイスに言及する。したがって、メディアエンコーダ200及びメディアデコーダ300は、コーディングデバイス、具体的には、それぞれ、メディアエンコーダ及びメディアデコーダの例を表す。いくつかの例では、ソースデバイス102及びデスティネーションデバイス116は、ソースデバイス102及びデスティネーションデバイス116の各々がメディア符号化及び復号構成要素を含むように、実質的に対称的な方式で動作し得る。したがって、システム100は、例えば、ビデ

40

50

オストリーミング、ビデオ再生、ビデオブロードキャスト、又はビデオ電話のための、ソースデバイス102とデスティネーションデバイス116との間での一方向又は両方向のメディア送信をサポートし得る。

【0017】

一般に、メディアソース104は、メディアデータ（すなわち、未加工の符号化されていないメディアデータ）のソースを表し、メディアデータの連続した一連のピクチャ（「フレーム」とも呼ばれる）をメディアエンコーダ200に提供し、メディアエンコーダ200は、ピクチャのためのデータを符号化する。ソースデバイス102のメディアソース104は、ビデオカメラなどの、ビデオキャプチャデバイス、以前にキャプチャされた未加工のビデオを含むビデオアーカイブ、及び／又はビデオコンテンツプロバイダからビデオを受信するためのビデオフィードインターフェースを含み得る。更なる代替として、メディアソース104は、ソースメディアとしてのコンピュータグラフィックススペースのデータ、又はライブメディア、アーカイブされたメディア、及びコンピューター生成されたメディアの組み合わせを生成し得る。各場合において、メディアエンコーダ200は、キャプチャされたメディアデータ、事前にキャプチャされたメディアデータ、又はコンピューター生成されたメディアデータを符号化する。メディアエンコーダ200は、受信された順序（「表示順序」と称されることがある）からコーディングのためのコーディング順序へと、ピクチャを並べ替え得る。メディアエンコーダ200は、符号化メディアデータを含むビットストリームを生成し得る。ソースデバイス102は次いで、例えばデスティネーションデバイス116の入力インターフェース122によって受信及び／又は取り出すために、出力インターフェース108を介してコンピュータ可読媒体110へと、符号化メディアデータを出力し得る。

10

20

【0018】

ソースデバイス102のメモリ106及びデスティネーションデバイス116のメモリ120は、汎用メモリを表す。いくつかの例では、メモリ106、120は、未加工のメディアデータ、例えば、メディアソース104からの未加工のメディアと、メディアデコーダ300からの未加工の復号されたメディアデータと、を記憶し得る。追加又は代替として、メモリ106、120は、それぞれ、例えばメディアエンコーダ200及びメディアデコーダ300によって実行可能なソフトウェア命令を記憶し得る。メモリ106及びメモリ120は、この例ではメディアエンコーダ200及びメディアデコーダ300とは別個に示されるが、メディアエンコーダ200及びメディアデコーダ300がまた、機能的に類似の又は等価な目的のために内部メモリを含み得ることを理解されたい。更に、メモリ106、120は、例えばメディアエンコーダ200から出力されメディアデコーダ300へ入力される、符号化メディアデータを記憶し得る。いくつかの例では、メモリ106、120の部分は、例えば、未加工の復号及び／又は符号化メディアデータを記憶するための、1つ又は複数のバッファとして割り振られ得る。

30

【0019】

コンピュータ可読媒体110は、ソースデバイス102からデスティネーションデバイス116に符号化メディアデータを移すことが可能な任意のタイプの媒体又はデバイスを表し得る。一例では、コンピュータ可読媒体110は、ソースデバイス102が、例えば、無線周波数ネットワーク又はコンピュータベースのネットワークを介して、符号化メディアデータをデスティネーションデバイス116にリアルタイムで直接送信することを可能にするための通信媒体を表す。出力インターフェース108は、符号化メディアデータを含む送信信号を変調してもよく、入力インターフェース122は、ワイヤレス通信プロトコルなどの通信規格に従って、受信された送信信号を復調してもよい。通信媒体は、無線周波数（radio frequency、RF）スペクトル又は1つ又は複数の物理伝送線路などの、任意のワイヤレス又は有線の通信媒体を備えてもよい。通信媒体は、ローカルエリアネットワーク、ワイドエリアネットワーク、又はインターネットなどのグローバルネットワークなどのパケットベースネットワークの一部を形成してもよい。通信媒体は、ルータ、スイッチ、基地局、又はソースデバイス102からデスティネーションデバイス116

40

50

への通信を容易にするために有用であり得る任意の他の機器を含み得る。

【0020】

いくつかの例では、ソースデバイス102は、出力インターフェース108から記憶デバイス112に符号化されたデータを出力し得る。同様に、デスティネーションデバイス116は、入力インターフェース122を介して、記憶デバイス112からの符号化されたデータにアクセスし得る。記憶デバイス112は、ハードドライブ、Blu-ray（登録商標）ディスク、DVD、CD-ROM、フラッシュメモリ、揮発性若しくは不揮発性メモリ、又は符号化メディアデータを記憶するための任意の他の好適なデジタル記憶媒体などの、様々な分散されたデータ記憶媒体又はローカルでアクセスされるデータ記憶媒体のうちのいずれかを含み得る。

10

【0021】

いくつかの例では、ソースデバイス102は、符号化メディアデータを、ソースデバイス102によって生成された符号化メディアデータを記憶し得るファイルサーバ114又は別の中間記憶デバイスに出力し得る。デスティネーションデバイス116は、ストリーミング又はダウンロードを介して、ファイルサーバ114からの記憶されたメディアデータにアクセスし得る。

【0022】

ファイルサーバ114は、符号化メディアデータを記憶し、その符号化メディアデータをデスティネーションデバイス116に送信することが可能な任意のタイプのサーバデバイスであり得る。ファイルサーバ114は、（例えば、ウェブサイトのための）ウェブサーバ、（ファイル転送プロトコル（File Transfer Protocol、FTP）又は片方向伝送路上ファイル配信（File Delivery over Unidirectional Transport、FLUTE）プロトコルなどの）ファイル転送プロトコルサービスを提供するように構成されたサーバ、コンテンツ配信ネットワーク（content delivery network、CDN）デバイス、ハイパーテキスト転送プロトコル（hypertext transfer protocol、HTTP）サーバ、マルチメディアブロードキャストマルチキャストサービス（Multimedia Broadcast Multicast Service、MBMS）若しくは拡張MBMS（Enhanced MBMS、eMBMS）サーバ、及び／又はネットワークアタッチトストレージ（network attached storage、NAS）デバイスを表し得る。ファイルサーバ114は、追加又は代替として、HTTP経由動的適応ストリーミング（Dynamic Adaptive Streaming over HTTP、DASH）、HTTPライブストリーミング（HTTP Live Streaming、HLS）、リアルタイムストリーミングプロトコル（Real Time Streaming Protocol、RTSP）、HTTP動的ストリーミング（HTTP Dynamic Streaming）などの1つ又は複数のHTTPストリーミングプロトコルを実装し得る。

20

30

【0023】

デスティネーションデバイス116は、インターネット接続を含む任意の標準的なデータ接続を通じて、ファイルサーバ114からの符号化メディアデータにアクセスし得る。これは、ワイヤレスチャネル（例えば、Wi-Fi接続）、有線接続（例えば、デジタル加入者線（digital subscriber line、DSL）、ケーブルモデムなど）、又はファイルサーバ114に記憶されている符号化メディアデータにアクセスするのに好適である、両方の組み合わせを含み得る。入力インターフェース122は、ファイルサーバ114からメディアデータを取り出すか若しくは受信するための上で考察された様々なプロトコル、又はメディアデータを取り出すための他のそのようなプロトコルのうちのいずれか1つ又は複数に従って動作するように構成されている場合がある。

40

【0024】

出力インターフェース108及び入力インターフェース122は、ワイヤレス送信機／受信機、モデム、有線ネットワーク構成要素（例えば、Ethernetカード）、様々なIEEE802.11規格のうちのいずれかに従って動作するワイヤレス通信構成要素、又は他の物理構成要素を表してもよい。出力インターフェース108及び入力インターフェース122がワイヤレス構成要素を備える例では、出力インターフェース108及び

50

入力インターフェース122は、4G、4G-LTE（ロングタームエボリューション（Long-Term Evolution））、LTEアドバンスド、5Gなどのセルラー通信規格に従って、符号化メディアデータなどのデータを転送するように構成されている場合がある。出力インターフェース108がワイヤレス送信機を備えるいくつかの例では、出力インターフェース108及び入力インターフェース122は、IEEE802.11仕様、IEEE802.15仕様（例えば、ZigBee（商標））、Bluetooth（商標）規格などの他のワイヤレス規格に従って、符号化メディアデータなどのデータを転送するように構成されている場合がある。いくつかの例では、ソースデバイス102及び／又はデスティネーションデバイス116は、それぞれのシステムオンチップ（system-on-a-chip、SoC）デバイスを含み得る。例えば、ソースデバイス102は、メディアエンコーダ200及び／又は出力インターフェース108に起因する機能を実行するためのSoCデバイスを含んでもよく、デスティネーションデバイス116は、メディアデコーダ300及び／又は入力インターフェース122に起因する機能を実行するためのSoCデバイスを含んでもよい。

10

20

30

50

【0025】

本開示の技法は、オーバジエアテレビジョンブロードキャスト、ケーブルテレビジョン送信、衛星テレビジョン送信、HTTP経由動的適応ストリーミング（DASH）などのインターネットストリーミングメディア送信、データ記憶媒体上に符号化されたデジタルメディア、データ記憶媒体上に記憶されたデジタルメディアの復号、又は他の適用例などの、様々なマルチメディア適用例のうちのいずれかをサポートするメディアコーディングに適用されてもよい。

【0026】

デスティネーションデバイス116の入力インターフェース122は、コンピュータ可読媒体110（例えば、通信媒体、記憶デバイス112、ファイルサーバ114など）から、符号化されたメディアビットストリームを受信する。符号化されたメディアビットストリームは、メディアエンコーダ200によって定義され、メディアデコーダ300によっても使用されるシグナリング情報を含むことができる。表示デバイス118は、復号されたメディアデータの復号されたピクチャをユーザに表示する。表示デバイス118は、液晶ディスプレイ（liquid crystal display、LCD）、プラズマディスプレイ、有機発光ダイオード（organic light emitting diode、OLED）ディスプレイ、又は別のタイプの表示デバイスなどの、様々な表示デバイスのうちのいずれかを表し得る。

【0027】

図1には示されていないが、いくつかの例では、メディアエンコーダ200及びメディアデコーダ300は各々、オーディオエンコーダ及び／又はオーディオデコーダと統合されてもよく、共通のデータストリームにおいてオーディオとメディアの両方を含む多重化されたストリームを扱うために、適切なMUX-DEMUXユニット、又は他のハードウェア及び／若しくはソフトウェアを含んでもよい。

【0028】

メディアエンコーダ200及びメディアデコーダ300は各々、1つ又は複数のマイクロプロセッサ、デジタル信号プロセッサ（digital signal processors、DSPs）、特定用途向け集積回路（application specific integrated circuits、ASICs）、フィールドプログラマブルゲートアレイ（field programmable gate arrays、FPGAs）、ディスクリート論理、ソフトウェア、ハードウェア、ファームウェア、又はそれらの任意の組み合わせなどの、様々な好適なエンコーダ及び／又はデコーダ回路のうちのいずれかとして実装され得る。技法が部分的にソフトウェアで実装されるとき、デバイスは、好適な非一時的コンピュータ可読媒体にソフトウェアのための命令を記憶し、本開示の技法を実行するために1つ又は複数のプロセッサを使用してハードウェアでその命令を実行し得る。メディアエンコーダ200及びメディアデコーダ300の各々は、1つ又は複数のエンコーダ又はデコーダに含まれてもよく、それらのいずれもが、それぞれのデバイスにおいて複合エンコーダ／デコーダ（コーデック（encoder/decoder、CO

DEC))の一部として統合されてもよい。メディアエンコーダ200及び／又はメディアデコーダ300を含むデバイスは、集積回路、マイクロプロセッサ、及び／又はセルラー電話などのワイヤレス通信デバイスを含み得る。

【0029】

ニューラルネットワークに基づく画像及びビデオ（例えば、メディア）圧縮方法は、現在の規格に対して競争力があり、いくつかの追加の利点を提供することができる。いくつかの例では、ニューラルベースのコーディング方法は、一般に、高精度浮動小数点演算を使用して設計され、テストされる。しかしながら、高精度浮動小数点演算の使用は、高い処理時間及び電力消費に起因して展開に影響を及ぼし得る。技術が実際の展開に移行するにつれて、ニューラルネットワーク重み及び活性化関数は、典型的には、速度及び電力消費を改善するために量子化され、低精度整数で表される。

10

【0030】

すなわち、実際の展開では、ニューラルネットワーク変数（例えば、重み及び活性化関数）が量子化され得る。しかしながら、ニューラルネットワーク変数の量子化には問題があり得る。

【0031】

本開示は、エントロピーコーディングに関係するニューラルネットワーク変数が量子化されるときに生じる問題に対処し得る例示的な技法について説明する。エントロピーコーディングに関連するニューラルネットワーク変数は、そのような変数が圧縮効率を定義するので、ニューラルベースのビデオ／画像圧縮方式の設計にとって重要である。更に、エントロピーコーディング変数の非常に具体的な特性は、ニューラルネットワークにおける量子化を最適化するための一般的なツールによって考慮されない。テストにより、実際に、最悪の量子化効果は、最も一般的なユースケースのいくつかで正確に発生する可能性があり、最悪の量子化効果によって引き起こされる損失は、ニューラルネットワークを再トレーニングすることによって回復することができないことが示された。

20

【0032】

本開示は、エントロピーコーディング変数の定義を更に最適化するための技法について説明し、それにより、効果的なエントロピーコーディングのために最も重要な情報が、最良に保存される。テストはまた、エントロピーコーディングに必要なメモリの量を最小化するために、本明細書で説明される技法がどのように使用され得るかを示した。本開示は、エントロピーコーディング設計のための一般的な手法、並びに一般に使用されるガウス分布のための特定の解決策及び実装形態について説明する。

30

【0033】

以下でより詳細に説明されるように、メディアエンコーダ200及びメディアデコーダ300は、ニューラルベースのメディア圧縮技法によってコーディングされたデータストリームのデータ要素の確率分布関数パラメータを決定し、確率分布関数パラメータは、データストリームの確率分布関数の標準偏差の対数関数から導出され、確率分布関数パラメータに基づいてコードベクトルを決定し、コードベクトルを使用してデータ要素をエントロピーコーディングする、ように構成されてもよい。

【0034】

エントロピーコーディングは、メディア圧縮システムの基本的な部分である。エントロピーコーディングプロセスは、可能な限り最もコンパクトな表現を得ることを目的として、メディア情報と圧縮データビットストリームとの間の変換を最適化する役割を果たす。メディア圧縮の他の要素とは異なり、エントロピーコーディングは可逆プロセスであり、すなわち、情報を完全に保存する。

40

【0035】

画像及びビデオ圧縮規格において効率的なエントロピーコーディングを実装するために、いくつかの技法が開発された。より最近では、深層学習及びニューラルネットワークに基づく新しい圧縮方法が、いくつかの他の実用上の利点を提供しながら、従来の方法の性能に近づいていることが示されている。

50

【 0 0 3 6 】

従来のコーデック及びニューラルベースのコーデックは、非常に異なる形式の圧縮を使用するので、従来のコーデック及びニューラルベースのコーデックは、著しく異なる方法でエントロピーコーディングを使用する。主な違いの概要が図 2 に示されており、2 つの手法が算術コーディングのような同じコーディング方法を使用する場合であっても、他の全ての態様が異なることが分かる。これは、ニューラルベースのコーデックにおける性能を最適化するための新しい技法を開発することが有益であることを意味する。

【 0 0 3 7 】

図 2 に示すように、従来のコーディング技法とニューラルベースのコーディング技法の両方の設計段階は、データ編成プロセス 2 0 0 と統計データモデリングプロセス 2 0 2 とを含む。H. 2 6 5 などの従来のメディアコーディング技法の場合、データ編成プロセス 2 0 0 は、コーディング要素（例えば、シンタックス要素及びコーディングモード）の定義を含む。ニューラルベースのコーディングプロセスの場合、データ編成プロセス 2 0 0 は、画像／ビデオ（例えば、メディア）オートエンコーダネットワークのトレーニング及び開発を含む。統計データモデリングプロセス 2 0 2 の場合、従来のコーディング技法は、例えば、コンテキスト適応型バイナリ算術コーディングなどのエントロピーコーディングプロセスのためのコンテキスト定義のコーディングを含む。ニューラルベースのコーディングプロセスの場合、統計モデリングプロセス 2 0 2 は、hyperprior ニューラルネットワークの開発及びトレーニングを含む。

【 0 0 3 8 】

従来のコーディング技法とニューラルベースのコーディング技法の両方は、確率決定プロセス 2 0 4 と、情報のビットへの変換プロセス 2 0 6 とを含む実装段階を含む。確率決定プロセス 2 0 4 と情報のビットへの変換プロセス 2 0 6 の両方は、コーディングされるデータに基づく。従来のコーディングの場合、確率決定プロセス 2 0 4（例えば、ビットレート割り当てプロセス）は、適応コンテキスト確率推定を含む。ニューラルベースのコーディングの場合、確率決定プロセス 2 0 4 は、以下でより詳細に説明するように、確率分布パラメータの決定を含む。従来のコーディング技法とニューラルベースのコーディング技法の両方は、情報のビットへの変換プロセス 2 0 6 のために、算術及び／又はエントロピーコーダ（例えば、又は算術若しくはエントロピーエンコーダ若しくはデコーダ）を使用する。算術コーダは、ハフマンコーディング、レンジコーディング又は非対称数値システム（asymmetric numeral system、ANS）コーディングであってもよい。

【 0 0 3 9 】

データストリーム管理段階は、圧縮データ 2 0 8 の管理である。いくつかの例示的な従来のコーディングシステムでは、圧縮されたデータは単一のビットストリーム中にある。いくつかの例示的なニューラルベースのコーディングシステムは、同時コーディングのために複数のストリームを使用する。

【 0 0 4 0 】

基本的な情報理論原理から導出される主なエントロピーコーディング原理は、データシンボルをコーディングするために割り当てられる最適なビット数 N_b が、

【 0 0 4 1 】

【数 1】

$$N_b = -\log_2(\hat{p}), \quad (1)$$

式中、

【 0 0 4 2 】

【数 2】

$$\hat{p}$$

10

20

30

40

50

は、エンコーダ及びデコーダ（例えば、メディアエンコーダ 200 及びメディアデコーダ 300）によって共有される

【0043】

【数 3】

$$\hat{p}$$

の知識を用いて、符号化されるシンボルの値の推定確率である。平均ビット数に対するこの最適数は分数であるが、実際には算術コーディングを用いて効果的に達成することができる。

【0044】

シンボル確率は、メディアコンテンツの特定の部分に依存して、広く頻繁に変化し得る。したがって、従来のメディア圧縮において、最も重要なタスクのうちの 1 つは、これらのデータシンボル確率の効率的な推定、及びコーディング中にデータシンボル確率が変化するときデータシンボル確率をどのように正確に更新するかである。この確率推定は逐次的に行われるので、従来のビデオコーディング方法でエントロピーコーディングを並列化することは非常に困難であり得る。

【0045】

ニューラルベースのコーデックのために提案された圧縮方法は、圧縮されるデータが特定のタイプの確率密度関数（PDF）を有し、この関数を定義するパラメータのみが変化し得るという仮定で動作するように設計されている。一例として、PDF は、各々が異なる標準偏差 σ を有する独立したゼロ平均ガウス分布を有する全ての変数を有するよう選択され得る。これは、ネットワークがそれを現実にするために「学習」するので、単に経験的な仮定ではない可能性がある。PDF は、通常、図 3 の例に示すように、ガウス分布であるが、必ずしもそうであるとは限らない。図 3 は、ニューラルベースのコーデックにおいて一般的に使用されるガウス連続 PDF 300、量子化された値、及びそれぞれの確率（ $p_0, p_1, \dots, p_8$ ）の一例を示す。まれなアンダーフロー及びオーバーフローの場合に対応する複数の値は、典型的には、確率推定を必要としない方法で準最適にコーディングされる。

【0046】

一般に、 σ は、「テンプレート」分布のスケーリング係数である。例えば、ゼロ平均及び単位分散を有するガウス又はラプラスがテンプレートであり、テンプレートを「伸張」又は「収縮」するためのスケーリングパラメータがある。一般に、標準偏差は、その利便性のために使用されるが、任意の係数であり得る。本開示は、人間にとって便利な（例えば、統計学者によって従来から使用されている）パラメータ又はスケール係数から、低精度パラメータ表現を使用してエントロピーコーディングのために最適化されたものに変更する技法を定義する。

【0047】

いくつかの例では、別個の（例えば、メディア圧縮を実行するニューラルネットワークとは別個の）ニューラルネットワークが、分布のパラメータを決定するようにトレーニングされる。このトレーニングの詳細については後述する。エントロピーコーディングの前に、これらの変数は、固定量子化方式を用いて離散化され、離散値（図 3 の例では、 $p_0, p_1, \dots, p_8$ ）の確率は、仮定された分布から導出され、後に式（1）に従ってシンボルビット数を決定するために用いられる。

【0048】

各データシンボルの全ての可能な値についての確率を推定することが典型的に必要な従来のコーディング方法とは異なり、ニューラルベースの方法は、通常、ニューラルネットワークを使用して並列に計算される PDF のパラメータ（例えば、標準偏差 σ ）のみに依存し、圧縮損失なしに同時エントロピーコーディングを実装することをはるかに容易にする。エントロピーコーディングは、それらの PDF パラメータの値によって完全に定

10

20

30

40

50

義されるので、PDFパラメータが低い整数精度で表されるとき、圧縮性能は著しく低下する可能性がある。本開示は、それらの圧縮損失を最小限に抑える新しい技法を説明する。

【0049】

ニューラルベースの画像及びビデオ圧縮のために提案される方法の1つの重要なクラスは、潜在変数が量子化され、固定学習分布を使用してエントロピーコーディングされる、変分オートエンコーダアーキテクチャを適応させる。いくつかの例では、このアーキテクチャは、オートエンコーダから生成された変数を使用して、それらのオートエンコーダ変数をエントロピーコーディングするために使用される確率分布を定義する、hyperpriorと呼ばれる、より小さいニューラルネットワークを追加することによって拡張されている。このhyperpriorネットワークも、量子化及びエントロピーコーディングを使用するが、事前定義された（例えば、学習された）PDFパラメータを用いる。図4は、エントロピーコーディングのために変分オートエンコーダ及びhyperpriorネットワークを使用するニューラル画像コーデックの図である。

10

【0050】

図4にプロセスが示され、画像分析及び合成のためのニューラルネットワークが、変分オートエンコーダを実装し、その変数のエントロピーコーディングが、（送信側と受信側の両方で使用される）hyperpriorデコーダネットワークによって定義される。エンコーダ及びデコーダ（例えば、メディアエンコーダ200及びメディアデコーダ300）は、オートエンコーダ部分を示すために使用されているので、ニューラルネットワークの文献では、それらの名前によって伝統的に呼ばれるシステムをそれぞれ送信側及び受信側と呼ぶことが一般的である。

20

【0051】

図4では、メディアエンコーダ200は、画像分析ニューラルネットワーク402、hyperpriorエンコーダニューラルネットワーク404、量子化プロセス406、エントロピー符号化プロセス408、量子化プロセス416、及びエントロピー符号化プロセス418を含むことができる。メディアエンコーダ200はまた、（ハイブリッドビデオコーデカにおける再構築ループと同様であり得る）エントロピー復号プロセス412、hyperpriorデコーダニューラルネットワーク424、エントロピー復号プロセス422、及び画像合成ニューラルネットワーク426を含み得る。図示されていないが、メディアデコーダ300は、エントロピー復号プロセス412、hyperpriorデコーダニューラルネットワーク424、エントロピー復号プロセス422、及び画像合成ニューラルネットワーク426と同様の構造を含み得る。

30

【0052】

画像分析ニューラルネットワーク402は、入力画像400を符号化及び／又は圧縮するように構成されたニューラルネットワークである。次に、画像分析ニューラルネットワーク402によって生成された圧縮データストリームは、hyperpriorエンコーダニューラルネットワーク404によって処理される。次に、hyperpriorエンコーダニューラルネットワーク404の出力は、量子化プロセス406によって量子化され、エントロピー符号化プロセス508によってエントロピーコーディングされて、ビットストリーム1410を生成する。hyperpriorエンコーダニューラルネットワーク404は、画像分析ニューラルネットワーク402によって作成されたデータストリームについてのPDFのパラメータ（例えば、標準偏差）を決定するように構成されたプロセスの一部である。次いで、PDFのパラメータを使用して、画像分析ニューラルネットワーク402からのデータストリームをエントロピー符号化することができる。

40

【0053】

エントロピー符号化プロセス408自体は、符号化される複数の画像から学習されたPDFパラメータ（例えば、学習パラメータ414）を使用する。エントロピー復号プロセス412は、ビットストリーム1410をエントロピー復号するために同じPDFパラメータを使用する。次に、このエントロピー復号されたビットストリーム412は、復号

50

されている特定の画像（例えば、入力画像 400）に対する PDF パラメータを生成するために、hyperprior デコーダニューラルネットワーク 424 によって処理される。

【0054】

次に、hyperprior デコーダニューラルネットワーク 404 によって生成された PDF パラメータは、エントロピー符号化プロセス 418 によって使用されて、量子化プロセス 416 による量子化の後に画像分析ニューラルネットワーク 402 によって生成されたデータストリームを符号化する。エントロピー符号化プロセス 418 は、ビットストリーム 2420 を作成する。次に、ビットストリーム 2420 は、hyperprior デコーダニューラルネットワーク 424 からの同じ PDF パラメータを使用して、エントロピー復号プロセス 422 によってエントロピー復号される。エントロピー復号の後、画像合成ニューラルネットワーク 426 は復号されたデータを処理して出力画像 428 を生成する。

10

【0055】

図 5 は、 x によって示される単一のデータ要素のコーディングプロセスのより詳細な図を提供する。特に、図 5 は、算術コーディングを使用したエントロピーコーディングの実装形態を示し、 x はコーディングされるデータであり、送信側及び受信側（例えば、メディアエンコーダ 200 及びメディアデコーダ 3）によって知られている PDF パラメータのベクトル σ は、 x の PDF を定義するパラメータを含む。より具体的には、ベクトル σ は標準偏差（例えば、正規ガウス分布）であり、より一般的には、エントロピーコーディングに使用する pdf を決定するために hyperprior ネットワークによって決定されたパラメータである。

20

【0056】

本開示のコンテキストでは、データ要素 x は、ニューラルネットワークへの入力（例えば、図 4 の画像合成ニューラルネットワーク 426 への入力）、ニューラルネットワークの出力（例えば、図 4 の画像分析ニューラルネットワーク 402 の出力）、シンタックス要素、又は符号化されたメディアビットストリームの他のコーディングされたデータを含み得る。ニューラルベースの画像／ビデオコーディングにおいてデータ要素 x （又はより一般的には、本開示の技法を使用してエントロピーコーディングされるデータ要素）を生成するために符号化され得るデータのいくつかの例示的なタイプは、画像又はビデオフレーム中の RGB 又は YUV ピクセル値、動き補償ビデオ中のピクセル値の残差（ピクセル値から動きからの予測を引いたもの）、ブロックベースの動き補償のための動きベクトル、及び／又はワーピングベースの動き補償のための高密度オプティカルフローデータを含み得る。他の例では、データ要素 x は、重み、活性化関数を表すデータ、又はニューラルベースのメディア圧縮技法によって使用される他のデータを含み得る。

30

【0057】

データはトレーニング段階中に符号化されず、むしろビットレート計算プロセス 500 は、算術コーディングプロセスによって使用されるビット数 N_b を推定するためにデータ要素 x 及び PDF パラメータ σ を使用する。ベクトル σ は、送信側及び受信側（例えば、メディアエンコーダ 200 及びメディアデコーダ 300）によって知られているデータ要素 x の PDF を定義するパラメータを表す。PDF パラメータ σ は、hyperprior デコーダネットワークによって計算され得るか、学習される事前定義されたパラメータのテーブルから読み取られ得る。一例として、コードベクトル計算プロセス 502 は、その対応する PDF パラメータ σ から特定のデータ要素 x のコードベクトル $c(\sigma)$ を計算することができる。

40

【0058】

量子化プロセス ($Q_{x \rightarrow q}$) 504 は、データ要素 x を量子化して量子化データ要素 q を作成する。算術エンコーダ 506 は、次に、コードベクトル $c(\sigma)$ を用いて量子化データ要素 q を算術コーディングし、ビットストリーム B を作成する。算術デコーダ 508 は、同一のコードベクトル $c(\sigma)$ を用いて算術復号を行うことにより、量子化データ

50

要素 q を復元する。逆量子化プロセス

【 0 0 5 9 】

【 数 4 】

$$(Q_{q \rightarrow \hat{x}})$$

5 1 0 は、量子化データ要素 q の逆量子化を行い、復号データ要素

【 0 0 6 0 】

【 数 5 】

$$\hat{x}$$

10

を復元する。復号データ要素

【 0 0 6 1 】

【 数 6 】

$$\hat{x}$$

は、その後、元のメディア（例えば、ピクチャ又はビデオ）を再構築するためにニューラルベースの復号化技法によって復号され得る。

20

【 0 0 6 2 】

本開示の技法は、エントロピーコーディングのために使用される確率分布のパラメータを生成するニューラルネットワークを含む全ての圧縮方法に適用可能であり、低精度整数を用いて記憶されたエントロピーコーディングパラメータのテーブルにも適用可能であることに留意されたい。次のセクションは、 σ がガウス PDF の標準偏差を表す例を頻繁に使用するが、本開示の一般的な技法は、PDF がパラメータ σ によって定義される全ての場合に適用可能である。

【 0 0 6 3 】

PDF は、2 つ以上のパラメータによって定義され得るが、全てのパラメータがエントロピーコーディングのために直接使用されるわけではない。例えば、PDF は、標準偏差 σ 及び平均値 μ によって定義され得るが、標準偏差のみがエントロピーコーディング特性を決定する。1 つのパラメータを有することが実際には最も一般的であるので、それは次のセクションで使用される仮定であり、以下で説明される技法は容易に一般化され得る。

30

【 0 0 6 4 】

図 5 に示すように、コードベクトル計算プロセス 5 0 2 と呼ばれるブロックは、算術コーディングプロセスによってビットストリーム B を作成するために必要な情報であるコードベクトル $c(\sigma)$ を作成するために PDF がどのように使用されるかを定義する。例えば、コードベクトルは、量子化後のシンボル（図 3 の $p_0, p_1, \dots, p_8$ ）によって定義される確率質量関数（probability mass function、PMF）に対応する累積分布関数（cumulative distribution function、CDF）を有するベクトルとすることができ

40

【 0 0 6 5 】

図 5 に示されるシステムは、ニューラルベースのコーデックが最初に設計され、トレーニングされ、テストされるときに、非常に正確な浮動小数点演算及び強力なプロセッサを使用して、エントロピーコーディングがどのように使用されるかを表す。しかしながら、実際の実装形態では、ニューラルネットワーク値は、通常、より高速な計算及び小さい電力使用を可能にするために、低精度整数で表される。これは図 6 の修正されたシステムに示されており、パラメータ σ が低精度バージョン

50

【 0 0 6 6 】

【 数 7 】

 $\hat{\sigma}$

に変換されることを示すために追加の量子化ブロックが追加され、例えば単一の 8 ビット整数（バイト）として表される。

【 0 0 6 7 】

図 6 は、P D F パラメータ σ が量子化され、低精度整数

【 0 0 6 8 】

【 数 8 】

 $\hat{\sigma}$

10

で表され、算術コーディングに必要なベクトルが事前計算され、C 要素を有する配列に記憶される、図 5 のシステムの実際の実装形態を示す。図 5 の例と同様に、データはトレーニング段階中に符号化されず、むしろビットレート計算プロセス 6 0 0 は、算術コーディングプロセスによって使用されるビット数 N_b を推定するためにデータ要素 x 及び P D F パラメータ σ を使用する。図 6 の例では、量子化プロセス

【 0 0 6 9 】

【 数 9 】

 $(Q_{\sigma \rightarrow \hat{\sigma}})$

20

6 0 2 は、P D F パラメータ σ を量子化し、量子化 P D F パラメータ

【 0 0 7 0 】

【 数 1 0 】

 $\hat{\sigma}$

30

を作成する。次いで、コード配列プロセス 6 0 3 は、量子化 P D F パラメータ

【 0 0 7 1 】

【 数 1 1 】

 $\hat{\sigma}$

からコードベクトル

【 0 0 7 2 】

【 数 1 2 】

 $c(\hat{\sigma})$

40

を生成する。すなわち、量子化 P D F パラメータ

【 0 0 7 3 】

【 数 1 3 】

 $\hat{\sigma}$

は、コード配列プロセス 6 0 3 によって決定されるために利用可能な C 要素（例えば、コードベクトル）のうちの 1 つを決定するために使用される。

50

【 0 0 7 4 】

量子化プロセス（ $Q_x \rightarrow q$ ） 6 0 4 は、データ要素 x を量子化して量子化データ要素 q を作成する。算術エンコーダ 6 0 6 は、次に、コードベクトル

【 0 0 7 5 】

【 数 1 4 】

$c(\hat{\sigma})$

を用いて量子化データ要素 q を算術コーディングし、ビットストリーム B を作成する。算術デコーダ 6 0 8 は、同一のコードベクトル

【 0 0 7 6 】

【 数 1 5 】

$c(\hat{\sigma})$

10

を用いて算術復号を行うことにより、量子化データ要素 q を復元する。逆量子化プロセス

【 0 0 7 7 】

【 数 1 6 】

$(Q_{q \rightarrow \hat{x}})$

20

6 1 0 は、量子化データ要素 q の逆量子化を行い、復号データ要素

【 0 0 7 8 】

【 数 1 7 】

$\hat{x}$

を復元する。復号データ要素

【 0 0 7 9 】

【 数 1 8 】

$\hat{x}$

30

は、その後、元のメディア（例えば、ピクチャ又はビデオ）を再構築するためにニューラルベースの復号化技法によって復号され得る。

【 0 0 8 0 】

なお、図 6 において、 x の量子化は、PDF パラメータ σ の量子化とは異なる。浮動小数点パラメータは、非常に複雑なニューラルネットワークのために良好に機能し得るが、より小さいデバイス（例えば、モバイルデバイス）での広範な使用にはあまり実行可能ではない。そのようなモバイルデバイスにおける電力消費は非常に高くなる。したがって、PDF パラメータ σ は、より実用的な用途のために整数に量子化される。PDF パラメータ

40

【 0 0 8 1 】

【 数 1 9 】

$\hat{\sigma}$

は、一例では 1 バイトの整数（例えば、256 個の可能な値）である。そのような量子化は、より高い電力効率をもたらす。

50

【 0 0 8 2 】

別の実際的な修正は、 x の各値に対してコードベクトル $c(\sigma)$ を再計算することは一般に計算コストが高いという事実によって動機付けられる。例えば、 x が、標準偏差が σ のガウス分布を有する（例えば、典型的な事例の）場合、量子化値の P M F は、以下によって定義される。

【 0 0 8 3 】

【数 2 0】

$$p_{\sigma}[n] = \frac{1}{2} \left[\operatorname{erf} \left(\frac{n+1/2}{\sqrt{2}\sigma} \right) - \operatorname{erf} \left(\frac{n-1/2}{\sqrt{2}\sigma} \right) \right], \quad (2)$$

10

式中、誤差関数は以下の通りである：

【 0 0 8 4 】

【数 2 1】

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt. \quad (3)$$

【 0 0 8 5 】

ビット単位のこの P M F のエントロピーは以下の通りである：

20

【 0 0 8 6 】

【数 2 2】

$$H(\sigma) = - \sum_{n=-\infty}^{\infty} p_{\sigma}[n] \log_2(p_{\sigma}[n]). \quad (4)$$

【 0 0 8 7 】

確率の計算に加えて、追加のタスクを実行して、計算コストに加えて、正しい整数累積和又はバイナリ確率を有するベクトル c を作成することができることに留意されたい。これらの計算は、パラメータ

30

【 0 0 8 8 】

【数 2 3】

 $\hat{\sigma}$

の異なる量子化値に対するコードベクトル

【 0 0 8 9 】

【数 2 4】

 $c(\hat{\sigma})$

40

を事前計算し、それらのベクトルを配列（例えば、メモリ内）に格納することによって、実際の実装形態において回避することができる。図 6 の図に示すように、送信側及び受信側（例えば、メディアエンコーダ 2 0 0 及びメディアデコーダ 3 0 0）は、パラメータ

【 0 0 9 0 】

【数 2 5】

 $\hat{\sigma}$

50

に従って選択され、算術符号化及び復号に使用されるコードベクトルの同じ配列を有する。

【 0 0 9 1 】

例えば、以下のシナリオを考慮する：

【 0 0 9 2 】

【 数 2 6 】

$$\hat{\sigma} = Q_{\sigma}(\sigma) = \lfloor \sigma \rfloor, \quad (5)$$

【 0 0 9 3 】

【 数 2 7 】

$\lfloor \sigma \rfloor$

（床関数）は σ 以下の最大整数であり、 x は予め計算されたコードベクトル

【 0 0 9 4 】

【 数 2 8 】

$$c(\hat{\sigma}) = c(\lfloor \sigma \rfloor + 1/2)$$

でエントロピーコーディングすることができる。

【 0 0 9 5 】

限られた数のコードベクトルを有することによって生じる損失は、最適コードベクトル $c(\sigma)$ の代わりにコードベクトル $c(r)$ を使用して x を符号化するのに必要な追加ビットの平均数であるコード冗長関数を使用して測定することができる。

【 0 0 9 6 】

コード冗長関数は、以下のように、Kullback-Leibler 発散尺度によって定義される。

【 0 0 9 7 】

【 数 2 9 】

$$R(\sigma|r) = \sum_{n=-\infty}^{\infty} p_{\sigma}[n] \log_2 \left(\frac{p_{\sigma}[n]}{p_r[n]} \right). \quad (6)$$

【 0 0 9 8 】

メディアコーデックは広範囲のビット比で動作するため、エントロピーコーディングは通常、平均冗長性を最小化するように設計されていない。一般に、平均値は先験的に知られていない。代わりに、エントロピーコーディングは、以下のように、相対的冗長性を制限することによって設計される。

【 0 0 9 9 】

【 数 3 0 】

$$L(\sigma|r) = \frac{R(\sigma|r)}{H(\sigma)}. \quad (7)$$

【 0 1 0 0 】

厳密に制限された相対的冗長性を有するエントロピーコーディング方法を設計することは、全ての平均も制限されることを保証する。換言すれば、この手法では、平均圧縮損失は最小ではないかもしれないが、相対損失が常に小さい場合には、平均圧縮損失は小さい

10

20

30

40

50

ことが保証される。

【 0 1 0 1 】

図 7 は、コードベクトルが以下の式を用いて計算されるときガウス分布に対する冗長性尺度の例を示す。

【 0 1 0 2 】

【 数 3 1 】

$$r = [\sigma] + \frac{1}{2}. \quad (8)$$

10

【 0 1 0 3 】

図 7 は、量子化 P D F パラメータ及び整数個のコードベクトルの使用に関連する冗長性の絶対尺度（プロット 7 0 0、右 Y 軸）及び相対尺度（プロット 7 0 2、左 Y 軸）の例を示す。この損失がコーディング効率である。本質的に、必要以上のビットが使用される。シグマ値を量子化すると、浮動小数点シグマの実際の値に関して不確実性が生じる。

【 0 1 0 4 】

コードベクトル $c(1.5)$ 、 $c(2.5)$ 、及び $c(3.5)$ が、それぞれ区間 $[1, 2)$ 、 $[2, 3)$ 、及び $[3, 4)$ において P D F パラメータ σ を用いて値をエントロピーコーディングするために使用されると仮定することができる。冗長性は、 $\sigma = r$ であるときにゼロであり、そうでなければ正であることが分かる。なお、図 7 の例では、区間 $[1, 2)$ における最大冗長性は、他の区間よりも大きい。

20

【 0 1 0 5 】

テストは、区間における最大冗長性が区間の幅及び区間内の値に依存することを示した。これは、コードベクトルの非一様な割り当てで、より良い結果が得られ得ることを意味する。

【 0 1 0 6 】

図 8 は、全ての区間において同じ最大相対的冗長性を得るように設計されたコードベクトル $c(\sigma)$ の非一様な割り当ての一例を示す。図 8 の相対的冗長性グラフは、コードベクトル $c(1.24)$ 、 $c(1.93)$ 、及び $c(3.14)$ が、それぞれ区間 $[1, 1.52)$ 、 $[1.52, 2.41)$ 、及び $[2.41, 4)$ において P D F パラメータ σ を用いて値をエントロピーコーディングするために使用される場合に取得され得る。これらの値では、最大相対的冗長性は 3 つの区間で同じであることが分かる。

30

【 0 1 0 7 】

図 8 を参照して上述した技法は、コードベクトル割り当てに関連しており、P D F パラメータ σ の量子化に関連していないことに留意されたい。例えば、 σ が整数値に直接量子化される場合、図 8 に示す分数区間閾値を適用することができない。

【 0 1 0 8 】

同じコードベクトル $c(r^*(\sigma_a, \sigma_b))$ でエントロピーコーディングされる σ 値の範囲を定義する区間 $[\sigma_a, \sigma_b)$ が与えられると、メディアコーダ（例えば、メディアエンコーダ 2 0 0 及びメディアデコーダ 3 0 0）は、以下の式が使用される場合、この区間内の最大相対的冗長性を最小にすることができる：

40

【 0 1 0 9 】

【 数 3 2 】

$$r^*(\sigma_a, \sigma_b) = \{r: L(\sigma_a|r) = L(\sigma_b|r)\}. \quad (9)$$

【 0 1 1 0 】

上記のコードベクトルの設計基準では、図 8 において、区間 $[\sigma_a, \sigma_b)$ 内の冗長性は、 $\sigma = r^*(\sigma_a, \sigma_b)$ のときにゼロ値のほぼ放物線形状であることが分かる。これらの条件下で、区間内の σ の一様な条件付き分布を仮定すると、平均冗長性は、次式によ

50

って十分に近似することができる。

【 0 1 1 1 】

【 数 3 3 】

$$\bar{L}(\sigma) = \frac{L(\sigma_a | r^*(\sigma_a, \sigma_b))}{3} = \frac{L(\sigma_b | r^*(\sigma_a, \sigma_b))}{3} \approx E_{\sigma \in [\sigma_a, \sigma_b]} \{L(\sigma | r^*(\sigma_a, \sigma_b))\}. \quad (10)$$

【 0 1 1 2 】

この平均は、図 8 の例において破線 8 0 0 として示されている。図 8 は、いくつかの区間が式 (9) に従って同じ最大冗長性を有する場合を示し、これらの区間もほぼ同じ平均冗長性を有する。

10

【 0 1 1 3 】

図 6 に示された実装形態に関する問題を識別するために考慮されるべき 1 つの事実は、コーデック設計者が、PDF のタイプ及びその定義のために使用されるパラメータを選択し得ることである。しかしながら、パラメータ値はネットワークトレーニングによって定義されるので、パラメータ値の相対頻度を定義するのは設計の実際の適用である。例えば、図 9 は、ニューラルベースの画像コーデックからのガウス標準偏差値の相対頻度のプロット 9 0 0 を示す。図 9 に示されるように、低いビットレート（実際に最も使用される）に対して、 σ 値は、最小許容値 $\sigma = 0.1$ でピークになり、ほとんど全てが $\sigma = 10$ より小さい。一方、より大きなビットレートに対しては、より大きな σ 値がより一般的である。本開示は、広範囲の再生品質をカバーするために、標準偏差が範囲 $[0.1, 1000]$ 内にあると仮定する。

20

【 0 1 1 4 】

PDF パラメータ σ は、コーデックによって使用されるビットレートを定義し、これは、式 (4) で定義されるエントロピー関数で測定される。図 10 は、量子化値のエントロピーが、サポートされることが望まれる標準偏差の範囲内でパラメータ σ と共にどのように変化するかを示すプロット 1 0 0 0 を示す。より具体的には、図 10 は、標準偏差 σ の関数として、ガウス分布を有する量子化されたランダム変数のエントロピー（ビット単位）を示す。

【 0 1 1 5 】

メディアコーダ（例えば、メディアエンコーダ 2 0 0 又はメディアデコーダ 3 0 0 ）が、 σ を範囲 $[0.1, 1000]$ から 2 5 6 個の値（単一バイトに適合する）に量子化して、式 (9) を使用して、

30

【 0 1 1 6 】

【 数 3 4 】

$\hat{\sigma}$

によって定義される 2 5 6 個の区間の各々に対して 1 つのコードベクトルを作成するように構成されている場合、メディアコーダは、図 1 1 に示される相対的冗長性の平均値を取得する。図 1 1 は、ガウス標準偏差 σ 及び $C = 256$ のコードベクトルの一様量子化を使用することによって生じる平均相対的冗長性（圧縮損失）を示すプロット 1 1 0 0 を示す。

40

【 0 1 1 7 】

図 1 0 及び図 1 1 に関して、以下の問題が観察され得る：

1. エントロピー関数は、 σ の小さい値（最も使用される範囲）で非常に高い導関数を有し、トレーニング中の学習収束を複雑にする。

2. エントロピーは原点付近で非常に非線形であるため、量子化されたニューラルネットワークを用いて再現することは困難である。

3. 小さな σ 値の重要な違いは、

50

【 0 1 1 8 】

【 数 3 5 】

 $\hat{\sigma}$

が低精度を有するときに分離することができず、図 1 1 の原点付近に示されている非常に高い相対的冗長性をもたらす。

4. σ の範囲は、実際に必要とされる値に対応するが、図 9 に示されるように、その範囲の大部分は、典型的にはめったに使用されない。示された例では、 σ の値の小さい値の範囲が最も使用され、量子化問題が最悪である場所でもある。

10

【 0 1 1 9 】

要するに、エントロピーコーディングを定義するための値として標準偏差を使用することは、高精度浮動小数点実装形態で合理的に良好に機能することができるが、実際の実装で使用される量子化されたニューラルネットワークにとって有益ではない場合がある。

【 0 1 2 0 】

コーデック設計者が P D F パラメータを選択するときにはいつでも、よく知られた直感的な意味を有するので、他のタイプの確率分布についても同様の問題が発生する。高精度ニューラルネットワークは、広範なトレーニングを介してそれらのパラメータを最適に使用する方法を「学習」することができるが、量子化ネットワークは学習しない場合がある。実際の実装形態では、コードベクトルの割り当てが量子化パラメータ

20

【 0 1 2 1 】

【 数 3 6 】

 $\hat{\sigma}$

を使用して行われるので、（例えば、図 8 の例に示すように）非一様量子化を使用しても、これらの問題を直接解決することができない。

【 0 1 2 2 】

上記で提示された実際的な問題は、P D F パラメータ σ がエントロピー関数をどのように整形するかによって、及び低精度整数表現（量子化 Q_σ を伴うブロックによって図 6 に示される）へのその変換によって引き起こされる。これらの問題の原因を、式（4）が量子化された値のエントロピーをどのように定義するかを直接的な分析から特定することは困難であり得、なぜなら、それは明確な直観的な解釈を伴わない項の無限和であるからである。より単純な近似を考慮することによって、より多くの洞察を得ることができる。ガウス分布及び σ の大きな値に対して、エントロピーは、以下によって近似されることができる。

30

【 0 1 2 3 】

【 数 3 7 】

$$H(\sigma) \approx \log_2(\sigma + 0.05) + \frac{1 + \ln(2\pi)}{2 \ln(2)}, \quad (11)$$

40

σ の値の小さい値に対して：

【 0 1 2 4 】

【 数 3 8 】

$$H(\sigma) \approx \frac{(4\sigma + 0.35)}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{8\sigma^2}\right). \quad (12)$$

【 0 1 2 5 】

50

図 1 2 は、対数スケールがパラメータ σ に使用されるとき、図 1 0 のエントロピーグラフ 1 2 0 0 を、上で定義された近似と共に示す。両方において、急速な漸近収束が観察されることができ、 σ の適切な範囲、及び両方の近似がほぼ同じである中間範囲内で非常に良好な一致が観察される。

【 0 1 2 6 】

図 1 2 から、 σ の対数がエントロピーに対してより直接的な関係を提供し、したがって、PDF を定義するためのより良いパラメータであり得ることをより容易に認識できる。考慮されるべき 1 つの残りの態様は、図 9 に示されるように、小さい σ 値が最も一般的であると予想され、 σ のそのような小さい値が非常に小さいエントロピーに対応することである。

10

【 0 1 2 7 】

図 1 3 に示されるように、エントロピーと標準偏差の両方について対数スケールを使用することは、エントロピー関数が σ の低い値で変化する方法も考慮するのに有益であり得ることを示す。メディアエンコーダ 2 0 0 又はメディアデコーダ 3 0 0 が非常に低いエントロピーを有する非常に多数の値に遭遇し得る場合があり得るので、非常に小さい絶対誤差であっても、大きな相対的冗長性及び大きな圧縮損失を生じ得る。図 1 3 は、エントロピーと、式 (1 1) 及び (1 2) からのその近似とを示すプロット 1 3 0 0 を示し、対数スケールでグラフ化されている。

【 0 1 2 8 】

量子化誤差が不可避であっても、本開示の技法は、量子化誤差が相対的冗長性にどのように影響するかを考慮して、量子化されるべき変数を置換することによって圧縮損失を低減することが可能であるという事実に基づく。

20

【 0 1 2 9 】

1 つの例示的な技法が図 1 4 に示されており、 v と呼ばれる新しい変数が、PDF をパラメータ化するために定義される。一般に、PDF パラメータ v は、PDF パラメータ σ の非線形関数（例えば、対数）を使用して決定され得る。 σ と v との間の変換関数は、次のように定義される。

【 0 1 3 0 】

【 数 3 9 】

$$\sigma = T_{v \rightarrow \sigma}(v). \quad (13)$$

30

【 0 1 3 1 】

図 1 4 は、エントロピーコーディングのために最適化された新しい設計変数、及びコードベクトルのより効率的な割り当てのためのテーブルを可能にするための、図 6 の実際のコーデックの修正を示す。図 1 4 に示されているように、この関数は、コーデック設計者がトレーニングにおいて PDF パラメータ σ を使用し続けることを可能にし、この関数は、自動勾配逆伝搬を維持するために、ほとんどどこでも微分可能であると仮定される。

【 0 1 3 2 】

説明を簡単にするために、以下の規則が採用される。

40

【 0 1 3 3 】

【 数 4 0 】

$$\begin{aligned} \sigma &\in [\sigma_{\min}, \sigma_{\max}) \subset \mathbb{R}, \\ v &\in [0, 1) \subset \mathbb{R}, \\ T_{v \rightarrow \sigma}(0) &= \sigma_{\min}, \\ T_{v \rightarrow \sigma}(1) &= \sigma_{\max}, \\ Q_{v \rightarrow n}(v) &= \lfloor N v \rfloor \in \{0, 1, \dots, N-1\}, \\ v_0 < v_1 &\Rightarrow T_{v \rightarrow \sigma}(v_0) < T_{v \rightarrow \sigma}(v_1). \end{aligned} \quad (14)$$

50

【0134】

式(14)は、 σ が最小値($\sigma_{\min}$)から最大値($\sigma_{\max}$)までの範囲であることを特定する。新しいパラメータ v は、 $0 \sim 1$ の範囲である。 v 値が0の変換関数 $T_{v \rightarrow \sigma}$ は $\sigma_{\min}$ に等しい。 v 値が1のときの変換関数 $T_{v \rightarrow \sigma}$ は、 $\sigma_{\max}$ と等しい。新しいパラメータ v (例えば、 $Q_{v \rightarrow n}(v)$)の量子化は、 $0 \sim N-1$ の範囲の整数を生成する。更に、 $v_0 < v_1$ の場合、 $T_{v \rightarrow \sigma}(v_0) < T_{v \rightarrow \sigma}(v_1)$ と定義される。

【0135】

図14では、データはトレーニング段階中に符号化されない。コーディング段階は新しいPDFパラメータ v を使用するように構成されているため、トレーニング段階は、PDFパラメータ v をPDFパラメータ σ (例えば、PDFの標準偏差)に変換して戻す変換($T_{v \rightarrow \sigma}$)プロセス1400を含む。ビットレート計算プロセス1401は、データ要素 x 及びPDFパラメータ v を使用して、算術コーディングプロセスによって使用されるビット数 N_b を推定する。

【0136】

図14の例では、量子化プロセス($Q_{v \rightarrow n}$)1402は、PDFパラメータ v を量子化し、量子化PDFパラメータ n を作成する。次いで、コード選択プロセス1412は、量子化PDFパラメータ n をコードベクトルインデックス i に変換する。コードベクトルインデックス i は、コード配列プロセス1414への入力であってもよい。すなわち、コードベクトルインデックス i は、コードベクトルのルックアップテーブルに対するインデックスであってもよい。次に、コード配列プロセス1414は、コードベクトルインデックス i からコードベクトル $c(r^*_i)$ を決定する。コードベクトル $c(r^*_i)$ は、データ要素 x を算術コーディングするために使用される。

【0137】

量子化プロセス($Q_{x \rightarrow q}$)1404は、データ要素 x を量子化して量子化データ要素 q を作成する。次に、算術エンコーダ1406は、コードベクトル $c(r^*_i)$ を使用して、量子化データ要素 q を算術コーディングし、ビットストリーム B を作成する。算術デコーダ1408は、同じコードベクトル $c(r^*_i)$ を使用して算術復号を実行することによって、量子化データ要素 q を復元する。逆量子化プロセス

【0138】

【数41】

$$(Q_{q \rightarrow \hat{x}})$$

1410は、量子化データ要素 q の逆量子化を行い、復号データ要素

【0139】

【数42】

$$\hat{x}$$

を復元する。復号データ要素

【0140】

【数43】

$$\hat{x}$$

は、その後、元のメディア(例えば、ピクチャ又はビデオ)を再構築するためにニューラルベースの復号化技法によって復号され得る。

【0141】

概して、本開示の一例では、メディアエンコーダ200及びメディアデコーダ300は、ニューラルベースのメディア圧縮技法によってコーディングされたデータストリームの

データ要素 (x) のための確率分布関数パラメータ (例えば、 v) を決定するように構成されている場合がある。確率分布関数パラメータ (v) は、データストリームの確率分布関数の標準偏差 (σ) の対数関数 (例えば、自然対数 ($\ln$)) から決定されてもよい。メディアエンコーダ 200 及びメディアデコーダ 300 は、確率分布関数パラメータ (v) に基づいてコードベクトル ($c(r^*i)$) を決定するように更に構成されている場合があり、コードベクトル ($c(r^*i)$) を使用してデータ要素 (x) をエントロピーコーディングすることができる。

【0142】

図 14 に示されるように、一例では、メディアエンコーダ 200 及びメディアデコーダ 300 は、コードベクトルを決定する前に確率分布関数パラメータ v を量子化するように構成されている場合がある。一例では、コードベクトルを決定するために、メディアエンコーダ 200 及びメディアデコーダ 300 は、確率分布関数パラメータ v に基づいてコードベクトルインデックス i を決定し、コードベクトルインデックスに基づいてコードベクトルのテーブルからコードベクトルを決定するように構成されている場合がある。コードベクトルのテーブルに関する更なる情報を以下に説明する。

【0143】

エンコーダ側から、メディアエンコーダ 200 は、ニューラルベースの圧縮技法を使用してデータ要素 x を生成するように構成されている場合があり、量子化データ要素を作成するためにデータ要素を量子化し得る。この例では、コードベクトルを使用してデータ要素をエントロピーコーディングすることは、コードベクトルを使用して量子化データ要素をエントロピー符号化することを含む。更に、この例では、例えば、図 4 に示されるように、メディアエンコーダ 200 は、データ要素を生成するために画像分析ニューラルネットワークを使用して画像又はビデオピクチャを処理するように構成されている場合がある。

【0144】

デコーダ側から、メディアデコーダ 300 は、コードベクトルを使用して符号化されたデータ要素を復号して、量子化データ要素を作成することができる。メディアデコーダ 300 は更に、再構築されたデータ要素を作成するために、量子化データ要素を逆量子化することができる。メディアデコーダ 300 は、次いで、画像又はビデオピクチャを再構築するために、画像合成ニューラルネットワークを使用して再構築されたデータ要素を処理してもよい。

【0145】

考慮されるべき次の態様は、エントロピーコーディングの性能が、式 (7) 及び (10) で定義される相対的冗長性尺度によって測定されることである。以下に、この手法の実際の適用例を示す。

【0146】

対数変換

最も一般的には、ニューラルベースのコーデックは、仮定 $x \sim N(0, \sigma)$ を使用して動作するように構成されている、すなわち、コーディングされる変数は、ゼロ平均ガウス (又は正規) 分布を有する。いくつかの例示的な *hyperprior* 分布ネットワークはいくつかの学習された PDF を使用するが、より最近の研究は、学習された PDF がガウス PDF 及び学習されたパラメータで置き換えられるときに改善があることを示している。

【0147】

この事実と、標準偏差の対数が PDF のより正確なパラメータ化を提供するという図 12 及び図 13 における観察を使用して、変数変換関数は、以下を使用して定義され得る。

【0148】

【数 44】

$$\sigma = T_{v \rightarrow \sigma} = \exp(v \ln(\sigma_{\max}) + (1 - v) \ln(\sigma_{\min})), \quad (15)$$

10

20

30

40

50

及びその逆は、

【 0 1 4 9 】

【 数 4 5 】

$$v = T_{\sigma \rightarrow v}(\sigma) = \frac{\ln(\sigma) - \ln(\sigma_{\min})}{\ln(\sigma_{\max}) - \ln(\sigma_{\min})}. \quad (16)$$

【 0 1 5 0 】

関数 $\ln()$ は、自然対数である。この新しいパラメータ v が $N = 256$ を使用して量子化されるとき、対応する $C = 256$ コードベクトルは式 (9) に従って最適化され、図 15 に示す平均相対的冗長性が得られる。図 15 は、 $N = 256$ 及び $C = 256$ のコードベクトルを有する PDF パラメータ v の一様量子化を使用することによって生じる平均相対的冗長性（圧縮における損失）を示すプロット 1500 を示す。図 15 は、相対的冗長性が常に 1 % 未満であるので、PDF パラメータ v がガウス PDF を定義するのにより適したパラメータであることを示す。

10

【 0 1 5 1 】

上に示した変換式 (15) 及び (16) は、指数関数及び対数関数のみを使用しており、これらは連続的であり、設計及びトレーニング中に計算するのが容易な導関数を有する。

【 0 1 5 2 】

コードベクトルを選択するためのテーブル

考慮すべき 1 つの追加の態様は、エントロピーコーディング中にコードベクトルが絶えずアクセスされるので、高いデータスループットを維持するために、高価な高速メモリにコードベクトルを記憶することが好ましい場合があることである。したがって、コードベクトルの数を減らすことによって、メモリ要件及びコストの節約が可能になる。

20

【 0 1 5 3 】

メモリ要件及びコストの節約は、PDF パラメータ v の量子化後に、図 16 に示されるコード選択テーブルを使用して達成することができる。これらは、 N 個の要素を有する単純なテーブルであり、図 8 の例と同様に、コードベクトルの非一様な割り当てを可能にするが、PDF パラメータ σ に基づいて遷移を定義する浮動小数点値を使用する代わりに、テーブルは、PDF パラメータ v の量子化後に得られる整数値 n に直接適用される。

30

【 0 1 5 4 】

利点は、このプロセスが、単純なテーブルルックアップ実装形態で非常に効率的に行うことができる整数から整数への変換であることである。図 16 は、 n の 256 個の可能な値の各々について、エントロピーコーディングのために使用されるべきコードベクトルの数を定義するそのようなテーブル 1600 の一例を示す。図 16 の例では、テーブルメモリの合計は 256 バイトである。メディアエンコーダ 200 及びメディアデコーダ 300 は、 $N = 256$ である値 n によって量子化された PDF パラメータ v をコードベクトルインデックス i に変換するために使用される図 16 のテーブル 1600 を使用することができる。テーブル 16 は、図 17 に示されるように、1 % 未満の最大冗長性をより良好に保証するように設計され、したがって、256 から 43 へのコードベクトルの数の低減を可能にする。図 17 は、標準偏差に従って得られた相対的冗長性を示すプロット 1700 を示す。

40

【 0 1 5 5 】

メモリ使用量を制限する別の方法は、標準偏差値の範囲を縮小することである。例えば、ニューラルベースは、以下の PDF パラメータを使用して実装され得る。

【 0 1 5 6 】

【 数 4 6 】

$$\sigma_{\min} = 0.112, \sigma_{\max} = 334.5 \Rightarrow \ln(\sigma_{\max}/\sigma_{\min}) = 8, \quad (17)$$

50

これは以下の変換式に対応する。

【0157】

【数47】

$$n = \max[0, \min[255, [32 \ln(\sigma)] + 70]], \quad (18)$$

32による乗算は、ビットシフトを介して実施することができる。

【0158】

図18は、コードベクトルの数を25個に削減する上記変換のために設計されたコードベクトル選択テーブル1800を示す。

10

【0159】

本開示の技法は、算術コーディングを使用して実装及びテストされ、モバイルデバイス上で実行されるニューラルベースのコーデックに組み込まれ得る。一例では、算術コーディング技法は、32ビット演算を使用して実装され得、確率値は16ビットにスケールングされ得る。

【0160】

テーブル1は、本開示の技法の性能を確認するために使用されるいくつかの実験結果を示す。エントロピーコーディングパラメータの量子化のみが結果に影響を及ぼすように、hyperprior分布データは浮動小数点演算を使用して得られた。

【0161】

テーブル1の結果の第1行目に示すように、量子化PDFパラメータ σ は、ガウスPDFをパラメータ化するために使用され、約0.5メガバイトのメモリが、256個のコードベクトルを記憶するために必要とされ、得られる冗長性は、図11から予想されるものであり、1000%に近づく。より良い結果が可能であるが、コードベクトルのためにより多くのメモリを必要とする。

20

【0162】

テーブル1の以下の行は、式(16)で定義されるように、 σ を量子化PDFパラメータ v （例えば、 σ の対数関数を使用して得られるパラメータ）で置き換えると、はるかに良好な結果が得られることを示している。 $N=C=256$ を使用することは、実際に冗長性を除去するが、コードベクトルのためのメモリの量は比較的大きい。

30

【0163】

結果は、 $N=C$ のより小さい値を使用することが、必要とされるメモリの量を低減するのに役立つが、相対的冗長性が非一様に増加することを示す。一方、 $C < N = 256$ （例えば、256個のコードベクトル未満）となる上述のテーブルルックアップ技法を使用すると、メモリの半量程度が、より一貫した相対的冗長性と共に使用される。

【0164】

【表1】

テーブル1：24個の画像を有するKodakデータセットの圧縮に使用されるデータに適用された、

本明細書に記載される技法の実装形態によって得られた実験結果。

40

量子化PDF パラメータ	N	C	メモリ (キロバイト)	相対的冗長性(%)@ビットレート				
				0.25bpp	0.5bpp	1.0bpp	1.5bpp	2.5bpp
σ	256	256	498.6	754.7	858.7	923.7	951.6	899.6
v	256	256	108.7	0.03	0.02	0.02	0.01	0.01
v	43	43	18.3	0.51	0.37	0.32	0.23	0.23
v	256	43	7.7	0.31	0.30	0.30	0.29	0.29
v	23	23	9.8	1.45	1.15	0.99	0.87	0.76
v	256	23	4.4	0.84	0.89	0.90	0.90	0.93

【0165】

上に示したように、ニューラルネットワークを使用する画像及びビデオコーディングの場合、設計者にとって最も直観的で便利なエントロピーコーディングのためのパラメータを選択することは、高精度浮動小数点実装形態を使用するときに機能することができるが

50

、ニューラルネットワークの重み及び活性化が量子化されるときに深刻な劣化を引き起こす可能性がある。実際に、一般的に使用されるガウス分布の分析は、圧縮効率の最悪の劣化（最も高い相対的冗長性）が、圧縮中に得られる最も一般的な場合に正確に生じること示す。

【0166】

本開示の提案された技法は、エントロピーコーディングのために使用されるパラメータを再定義し、変数の単一の変更を用いてトレーニングプロセスを維持するが、量子化から生じる冗長性増加を等化するように量子化を修正することによって、問題に対処する。更に、新しいパラメータ化を使用すると、コードベクトル選択のための単純なテーブルルックアップを使用して、エントロピーコーディングのために使用されるメモリの量を低減することが可能であることが示される。

10

【0167】

図19は、本開示の例となる方法を例示するフローチャートである。図19の技法は、図14の例に示される構成要素を含む、メディアエンコーダ200及び／又はメディアデコーダ300の1つ又は複数の構成要素によって実行され得る。図19の例において、メディアエンコーダ200及び／又はメディアデコーダ300は、ニューラルベースのメディア圧縮技法によってコーディングされたデータストリームのデータ要素の確率分布関数パラメータを決定し、確率分布関数パラメータは、データストリームの確率分布関数の標準偏差の対数関数であり（1900）、確率分布関数パラメータに基づいてコードベクトルを決定し（1902）、コードベクトルを使用してデータ要素を算術コーディングする（1904）ように構成されてもよい。

20

【0168】

本開示の他の例示的な態様が以下で説明される。

【0169】

態様1－メディアデータをコーディングする方法であって、方法は、ニューラルベースのメディア圧縮技法によってコーディングされたデータストリームのデータ要素についての確率分布関数パラメータであって、確率分布関数パラメータは、データストリームの確率分布関数の標準偏差の対数関数である、確率分布関数パラメータを決定することと、確率分布関数パラメータに基づいてコードベクトルを決定することと、コードベクトルを使用してデータ要素をエントロピーコーディングすることと、を含む、方法。

30

【0170】

態様2－コードベクトルを決定する前に、確率分布関数パラメータを量子化することを更に含む、態様1に記載の方法。

【0171】

態様3－コードベクトルを決定することは、確率分布関数パラメータに基づいてコードベクトルインデックスを決定することと、コードベクトルインデックスに基づいてコードベクトルのテーブルからコードベクトルを決定することとを含む、態様1から2のいずれかに記載の方法。

【0172】

態様4－確率分布関数パラメータは v であり、確率分布関数の標準偏差は σ であり、最小標準偏差は $\sigma_{\min}$ であり、最大標準偏差は $\sigma_{\max}$ であり、 v は、

40

【0173】

【数48】

$$v = \frac{\ln(\sigma) - \ln(\sigma_{\min})}{\ln(\sigma_{\max}) - \ln(\sigma_{\min})}$$

と定義されている、態様1から3のいずれかに記載の方法。

【0174】

態様5－ニューラルベースの圧縮技法を使用してデータ要素を生成することと、量子化

50

データ要素を作成するために、データ要素を量子化することと、を更に含み、コードベクトルを使用してデータ要素をエントロピーコーディングすることは、コードベクトルを使用して量子化データ要素をエントロピー符号化することを含む、態様 1 から 4 のいずれかに記載の方法。

【0175】

態様 6 – ニューラルベースの圧縮技法を使用してデータ要素を生成することは、データ要素を生成するために、画像分析ニューラルネットワークを使用して画像又はビデオピクチャを処理することを含む、態様 5 に記載の方法。

【0176】

態様 7 – コードベクトルを使用してデータ要素をエントロピーコーディングすることは、量子化データ要素を作成するために、コードベクトルを使用して符号化されたデータ要素をエントロピー復号することを含み、方法は、再構築されたデータ要素を作成するために、量子化データ要素を逆量子化することを更に含み、態様 1 から 4 のいずれかに記載の方法。

【0177】

態様 8 – 画像又はビデオピクチャを再構築するために、画像合成ニューラルネットワークを使用して再構築されたデータ要素を処理することを更に含み、態様 7 に記載の方法。

【0178】

態様 9 – メディアデータをコーディングするように構成される装置であって、装置は、メモリと、メモリと通信している 1 つ又は複数のプロセッサと、を備え、1 つ又は複数のプロセッサが、ニューラルベースのメディア圧縮技法によってコーディングされたデータストリームのデータ要素の確率分布関数パラメータであって、確率分布関数パラメータは、データストリームの確率分布関数の標準偏差の対数関数である、の確率分布関数パラメータを決定し、確率分布関数パラメータに基づいてコードベクトルを決定し、コードベクトルを使用してデータ要素をエントロピーコーディングする、ように構成されている、装置。

【0179】

態様 10 – 1 つ又は複数のプロセッサが、コードベクトルを決定する前に、確率分布関数パラメータを量子化するように更に構成されている、態様 9 に記載の装置。

【0180】

態様 11 – コードベクトルを決定するために、1 つ又は複数のプロセッサは、確率分布関数パラメータに基づいてコードベクトルインデックスを決定し、コードベクトルインデックスに基づいてコードベクトルのテーブルからコードベクトルを決定するように更に構成されている、態様 9 から 10 のいずれかに記載の装置。

【0181】

態様 12 – 確率分布関数パラメータは v であり、確率分布関数の標準偏差は σ であり、最小標準偏差は $\sigma_{\min}$ であり、最大標準偏差は $\sigma_{\max}$ であり、 v は、

【0182】

【数 49】

$$v = \frac{\ln(\sigma) - \ln(\sigma_{\min})}{\ln(\sigma_{\max}) - \ln(\sigma_{\min})}$$

と定義されている、態様 9 から 11 のいずれかに記載の装置。

【0183】

態様 13 – 1 つ又は複数のプロセッサは、ニューラルベースの圧縮技法を使用してデータ要素を生成し、量子化データ要素を作成するために、データ要素を量子化するように更に構成されており、コードベクトルを使用してデータ要素をエントロピーコーディングすることは、コードベクトルを使用して量子化データ要素をエントロピー符号化することを含む、態様 9 から 11 のいずれかに記載の装置。

【0184】

態様14－ニューラルベースの圧縮技法を使用してデータ要素を生成するために、1つ又は複数のプロセッサは、データ要素を生成するために、画像分析ニューラルネットワークを使用して画像又はビデオピクチャを処理するように更に構成されている、態様13に記載の装置。

【0185】

態様15－画像又はビデオピクチャをキャプチャするように構成されたカメラを更に備える、態様14に記載の装置。

【0186】

態様16－コードベクトルを使用してデータ要素をエントロピーコーディングすることは、量子化データ要素を作成するために、コードベクトルを使用して符号化されたデータ要素をエントロピー復号することを含み、1つ又は複数のプロセッサは、再構築されたデータ要素を作成するために、量子化データ要素を逆量子化するように更に構成されている、態様9から12のいずれかに記載の装置。

【0187】

態様17－1つ又は複数のプロセッサが、画像又はビデオピクチャを再構築するために、画像合成ニューラルネットワークを使用して再構築されたデータ要素を処理するように更に構成されている、態様16に記載の装置。

【0188】

態様18－画像又はビデオピクチャを表示するように構成されたディスプレイを更に備える、態様17に記載の装置。

【0189】

態様19－命令を記憶する非一時的コンピュータ可読記憶媒体であって、命令は、実行されると、1つ又は複数のプロセッサに、ニューラルベースのメディア圧縮技法によってコーディングされたデータストリームのデータ要素についての確率分布関数パラメータであって確率分布関数パラメータは、データストリームの確率分布関数の標準偏差の対数関数である、確率分布関数パラメータを決定させ、確率分布関数パラメータに基づいてコードベクトルを決定し、コードベクトルを使用してデータ要素をエントロピーコーディングさせる、非一時的コンピュータ可読記憶媒体。

【0190】

態様20－確率分布関数パラメータは v であり、確率分布関数の標準偏差は σ であり、最小標準偏差は $\sigma_{\min}$ であり、最大標準偏差は $\sigma_{\max}$ であり、 v は、

【0191】

【数50】

$$v = \frac{\ln(\sigma) - \ln(\sigma_{\min})}{\ln(\sigma_{\max}) - \ln(\sigma_{\min})}$$

と定義される、態様19に記載の非一時的コンピュータ可読記憶媒体。

【0192】

例に応じて、本明細書において説明された技法のうちのいずれかのある特定の行為又はイベントが、異なる順序で実行されることが可能であり、追加、統合、又は完全に除外されてもよい（例えば、説明された全ての行為又はイベントが技法の実践に必要であるとは限らない）ことを認識されたい。その上、ある特定の例では、行為又はイベントは、連続的ではなく、例えば、マルチスレッド処理、割込み処理、又は複数のプロセッサを通じて並行して実行されてもよい。

【0193】

1つ又は複数の例では、説明した機能は、ハードウェア、ソフトウェア、ファームウェア、又はそれらの任意の組み合わせにおいて実装されてもよい。ソフトウェアで実装される場合、機能は、1つ又は複数の命令又はコードとして、コンピュータ可読媒体上に記憶

されてもよく、又はコンピュータ可読媒体を介して送信されてもよく、ハードウェアベースの処理ユニットによって実行されてもよい。コンピュータ可読媒体は、データ記憶媒体などの有形媒体に相当するコンピュータ可読記憶媒体、又は、例えば、通信プロトコルに従って、ある場所から別の場所へのコンピュータプログラムの転送を容易にする任意の媒体を含む通信媒体を含み得る。このように、コンピュータ可読媒体は一般に、(1) 非一時的である有形のコンピュータ可読記憶媒体、又は(2) 信号又は搬送波のような通信媒体に対応し得る。データ記憶媒体は、本開示で説明された技法を実装するための命令、コード、及び／又はデータ構造を取り出すために、1つ又は複数のコンピュータ又は1つ又は複数のプロセッサによってアクセスされ得る、任意の利用可能な媒体であってもよい。コンピュータプログラム製品はコンピュータ可読媒体を含み得る。

10

【0194】

限定ではなく例として、そのようなコンピュータ可読記憶媒体は、RAM、ROM、EPROM、CD-ROM若しくは他の光ディスクストレージ、磁気ディスクストレージ若しくは他の磁気ストレージデバイス、フラッシュメモリ、又は、命令若しくはデータ構造の形態の所望のプログラムコードを記憶するために使用され得、コンピュータによってアクセスすることができる任意の他の媒体を備えてもよい。また、任意の接続がコンピュータ可読媒体と適切に呼ばれる。例えば、命令が、同軸ケーブル、光ファイバーケーブル、ツイストペア、デジタル加入者回線(DSL)、又は赤外線、無線、及びマイクロ波などのワイヤレス技術を使用して、ウェブサイト、サーバ、又は他のリモートソースから送信される場合、同軸ケーブル、光ファイバーケーブル、ツイストペア、DSL、又は赤外線、無線、及びマイクロ波などのワイヤレス技術は、媒体の定義の中に含まれる。しかしながら、コンピュータ可読記憶媒体及びデータ記憶媒体は、接続、搬送波、信号、又は他の一時的な媒体を含まず、代わりに非一時的な有形記憶媒体を対象とすることを理解されたい。本明細書で使用されるディスク(disk)及びディスク(disc)は、コンパクトディスク(compact disc、disc)(CD)、レーザディスク(disc)、光ディスク(disc)、デジタル多用途ディスク(disc)(digital versatile disc、DVD)、フロッピーディスク(disk)、及びブルーレイディスク(disc)を含み、ディスク(disk)は通常、データを磁氣的に再生し、一方で、ディスク(disc)は、レーザを用いてデータを光学的に再生する。上記のものの組み合わせもまた、コンピュータ可読媒体の範囲内に含まれるものとする。

20

30

【0195】

命令は、1つ又は複数のDSP、汎用マイクロプロセッサ、ASIC、FPGA、又は他の等価な集積論理回路若しくはディスクリット論理回路などの、1つ又は複数のプロセッサによって実行され得る。したがって、本明細書で使用される「プロセッサ」及び「処理回路」という用語は、上記の構造、又は本明細書で説明された技法の実装に好適な任意の他の構造のうちのいずれかを指すことがある。加えて、いくつかの態様では、本明細書で説明された機能は、符号化及び復号のために構成された専用のハードウェアモジュール及び／若しくはソフトウェアモジュール内で提供されてもよく、又は複合コーデックの中に組み込まれてもよい。また、技法は、1つ又は複数の回路又は論理要素において完全に実装されてもよい。

40

【0196】

本開示の技法は、ワイヤレスハンドセット、集積回路(integrated circuit、IC)、又はICのセット(例えば、チップセット)を含む、多種多様なデバイス又は装置において実装されてもよい。開示された技法を実行するように構成されたデバイスの機能的態様を強調するために、様々な構成要素、モジュール、又はユニットが本開示で説明されるが、それらは必ずしも異なるハードウェアユニットによる実現を必要とするとは限らない。むしろ、上記で説明したように、様々なユニットは、コーデックハードウェアユニットにおいて組み合わせられてもよく、又は好適なソフトウェア及び／若しくはファームウェアと連携して、上記で説明したような1つ又は複数のプロセッサを含む、相互動作可能なハードウェアユニットの集合によって提供されてもよい。

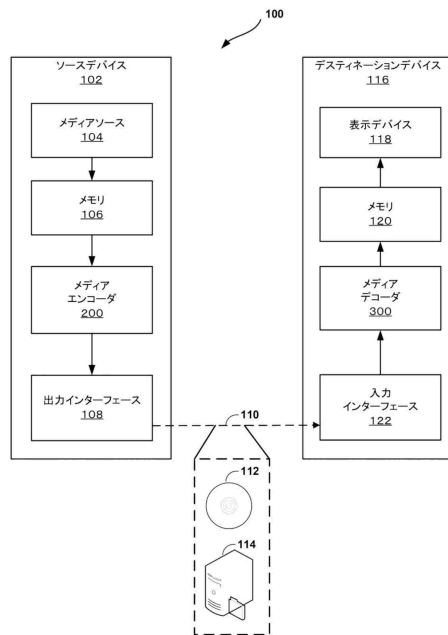
50

【 0 1 9 7 】

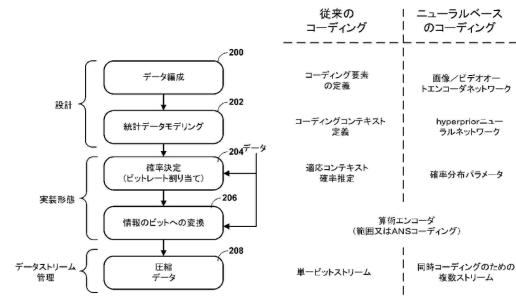
様々な例が説明されてきた。これら及び他の例は以下の特許請求の範囲の範囲内にある。

【 図 面 】

【 図 1 】



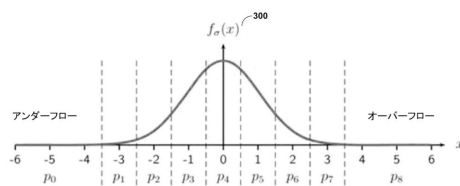
【 図 2 】



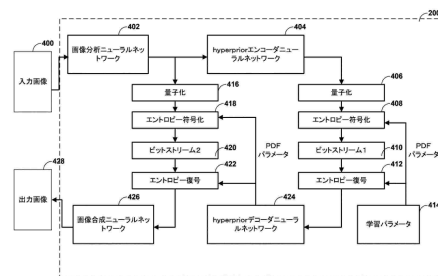
10

20

【 図 3 】



【 図 4 】

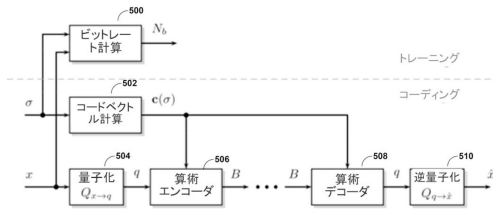


30

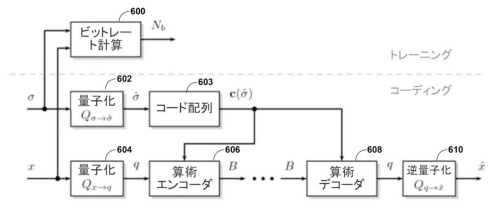
40

50

【図 5】



【図 6】



【図 7】

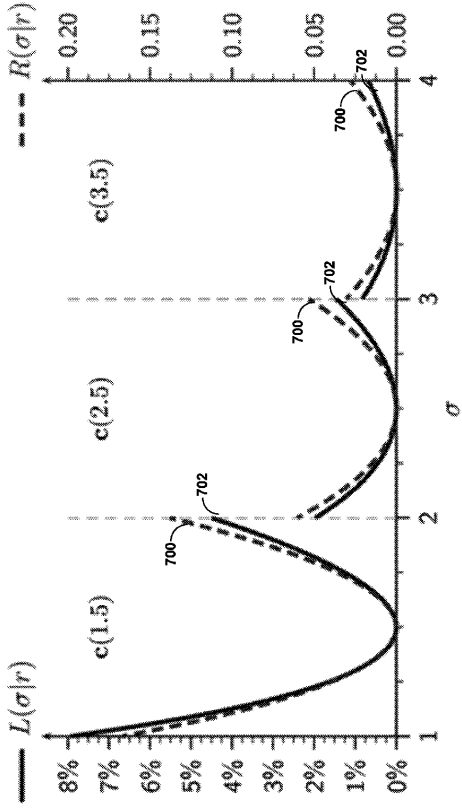


FIG. 7

【図 8】

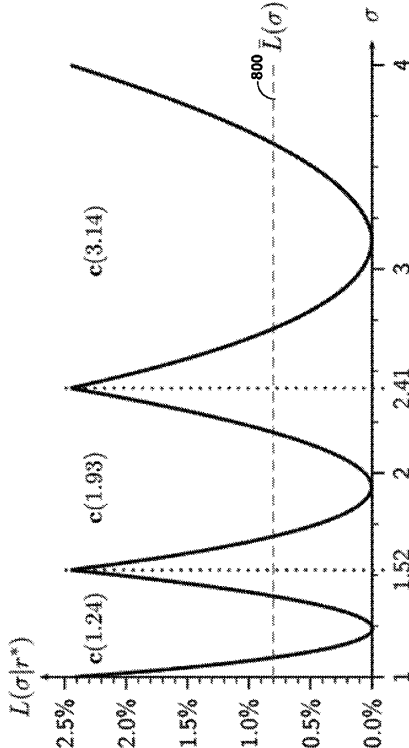


FIG. 8

10

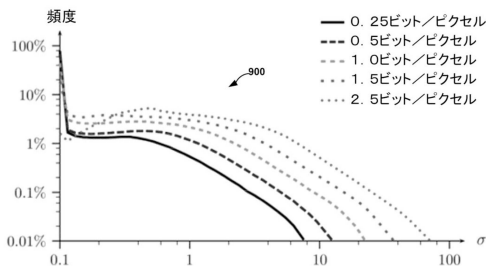
20

30

40

50

【図 9】



【図 10】

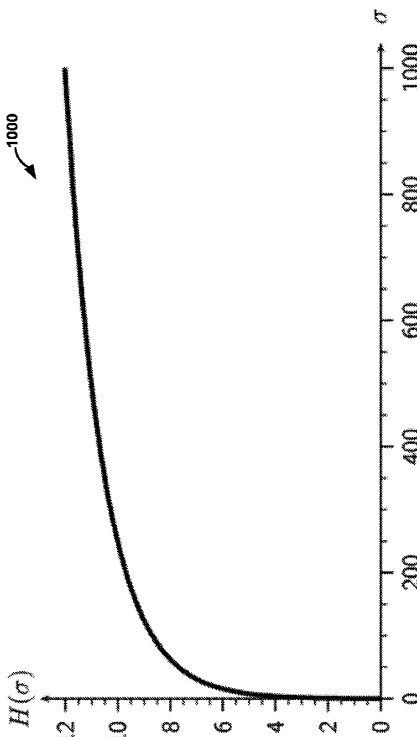


FIG. 10

【図 11】

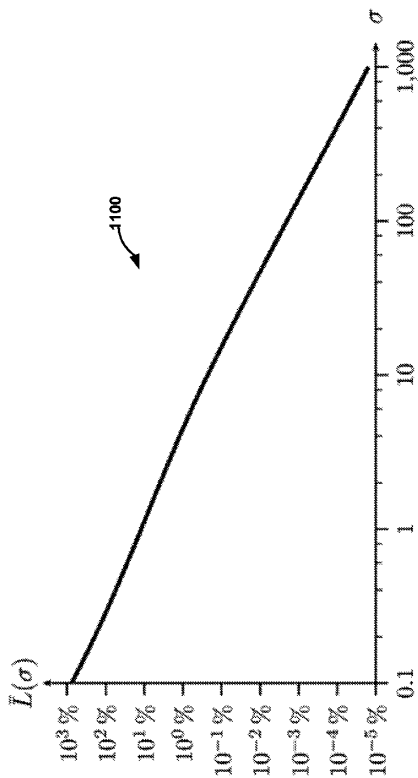
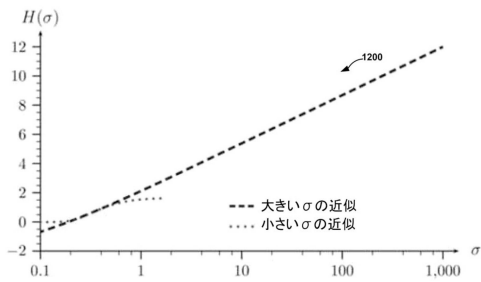


FIG. 11

【図 12】



10

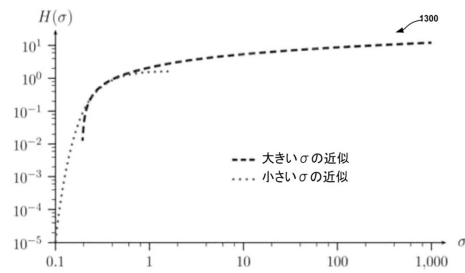
20

30

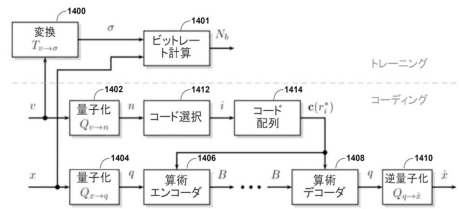
40

50

【図 1 3】



【図 1 4】



【図 1 5】

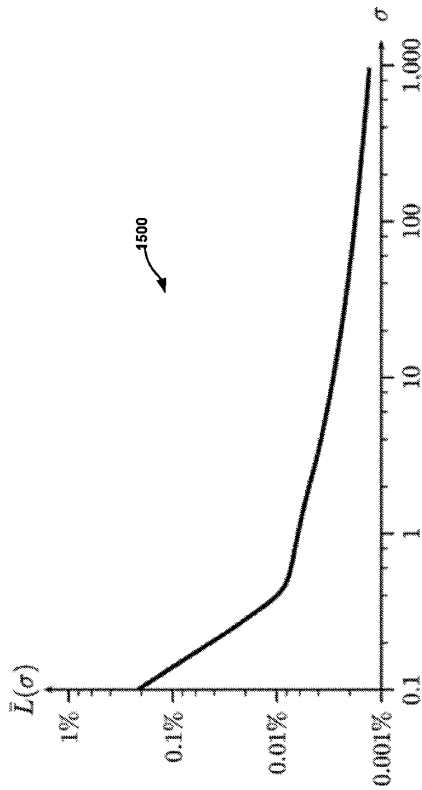


FIG. 15

【図 1 6】

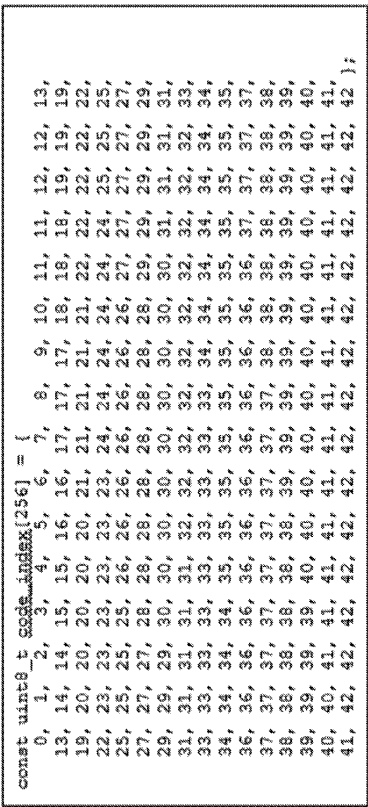


FIG. 16

10

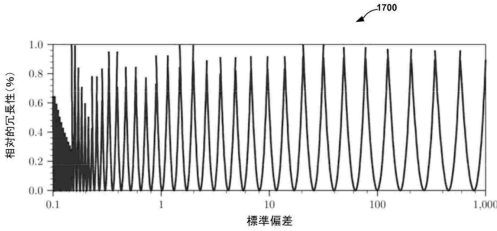
20

30

40

50

【図 17】



【図 18】

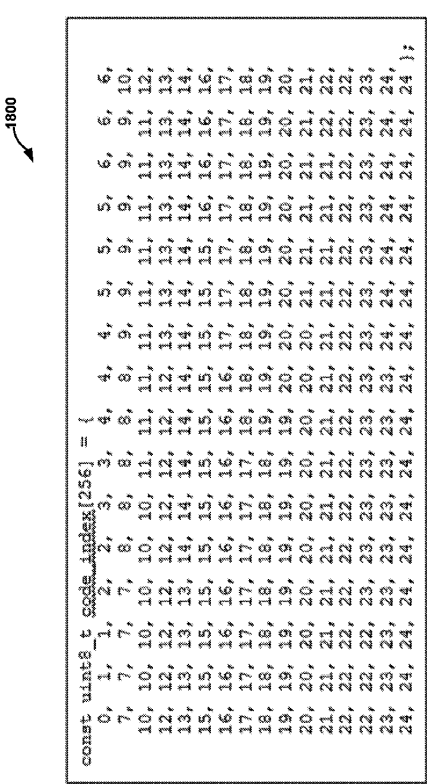
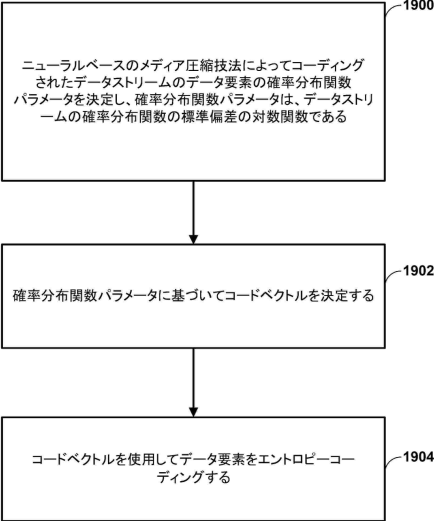


FIG. 18

【図 19】



10

20

30

40

50

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2023/060460

A. CLASSIFICATION OF SUBJECT MATTER INV. H04N19/91 H04N19/94 H04N19/463 ADD.		
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) H04N		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EPO-Internal, INSPEC, WPI Data		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Johannes Ballé ET AL: "Variational image compression with a scale hyperprior",	1, 2,
	1 May 2018 (2018-05-01), XP055632204,	4-10,
	Retrieved from the Internet:	12-20
	URL:https://arxiv.org/pdf/1802.01436.pdf	
	[retrieved on 2019-10-15]	
A	the whole document	3, 11

	-/--	
☒ 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		Date of mailing of the international search report
8 March 2023		28/03/2023
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016		Authorized officer Cyranka, Oliver

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2023/060460

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	BRUMMER BENOIT ET AL: "End-to-end optimized image compression with competition of prior distributions", 2021 IEEE/CVF CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION WORKSHOPS (CVPRW), IEEE, 19 June 2021 (2021-06-19), pages 1890-1894, XP033967894, DOI: 10.1109/CVPRW53098.2021.00212 [retrieved on 2021-08-25]	1-3, 5, 7, 9-11, 13, 16, 19
A	the whole document	4, 6, 8, 12, 14, 15, 17, 18, 20
A	----- ZHAO JING ET AL: "A Universal Encoder Rate Distortion Optimization Framework for Learned Compression", 2021 IEEE/CVF CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION WORKSHOPS (CVPRW), IEEE, 19 June 2021 (2021-06-19), pages 1880-1884, XP033967384, DOI: 10.1109/CVPRW53098.2021.00210 [retrieved on 2021-08-25] the whole document	1-20
T	----- AMIR SAID ET AL: "Optimized learned entropy coding parameters for practical neural-based image and video compression", ARXIV.ORG, CORNELL UNIVERSITY LIBRARY, 201 OLIN LIBRARY CORNELL UNIVERSITY ITHACA, NY 14853, 20 January 2023 (2023-01-20), XP091419131, DOI: 10.1109/ICIP46576.2022.9897505 the whole document -----	1-20

Form PCT/ISA/210 (continuation of second sheet) (April 2005)

フロントページの続き

,MC,ME,MK,MT,NL,NO,PL,PT,RO,RS,SE,SI,SK,SM,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,KM,
ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AO,AT,AU,AZ,BA,BB,BG,BH,BN,BR,BW,BY,BZ,CA,CH,CL,CN,C
O,CR,CU,CV,CZ,DE,DJ,DK,DM,DO,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,GT,HN,HR,HU,ID,IL,IN,IQ,I
R,IS,IT,JM,JO,JP,KE,KG,KH,KN,KP,KR,KW,KZ,LA,LC,LK,LR,LS,LU,LY,MA,MD,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,ST,SV,
SY,TH,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,WS,ZA,ZM,ZW

1 2 1 - 1 7 1 4 ・ サン・ディエゴ・モアハウス・ドライヴ・ 5 7 7 5

F ターム (参考) 5C159 MD02 ME01 ME11 TA57 TD04 UA02 UA05