

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7688165号
(P7688165)

(45)発行日 令和7年6月3日(2025.6.3)

(24)登録日 令和7年5月26日(2025.5.26)

(51)国際特許分類

F I

H 0 4 N 19/463(2014.01)

H 0 4 N 19/463

請求項の数 13 (全100頁)

(21)出願番号	特願2023-574690(P2023-574690)	(73)特許権者	503433420
(86)(22)出願日	令和4年6月1日(2022.6.1)		華為技術有限公司
(65)公表番号	特表2024-520151(P2024-520151 A)		HUAWEI TECHNOLOGIES CO., LTD.
(43)公表日	令和6年5月21日(2024.5.21)		中華人民共和國 5 1 8 1 2 9 広東省深
(86)国際出願番号	PCT/CN2022/096510		▲チェン▼市龍崗区坂田 華為総部▲ペ
(87)国際公開番号	WO2022/253249		ン▼公樓
(87)国際公開日	令和4年12月8日(2022.12.8)		Huawei Administration Building, Banti
審査請求日	令和6年1月29日(2024.1.29)		an, Longgang District, Shenzhen, Guang
(31)優先権主張番号	202110616029.2		dong 5 1 8 1 2 9, P.R. C
(32)優先日	令和3年6月2日(2021.6.2)		hina
(33)優先権主張国・地域又は機関	中国(CN)	(74)代理人	100110364
(31)優先権主張番号	202110674299.9		弁理士 実広 信哉
(32)優先日	令和3年6月17日(2021.6.17)		
最終頁に続く		最終頁に続く	

(54)【発明の名称】 特徴データ符号化および復号方法および装置

(57)【特許請求の範囲】

【請求項 1】

特徴データ符号化方法であって、
符号化されることになる特徴データを取得するステップであって、前記符号化されることになる特徴データが複数の特徴要素を含み、前記複数の特徴要素が第1の特徴要素を含む、ステップと、
前記第1の特徴要素の確率推定結果を取得するステップと、
前記第1の特徴要素の前記確率推定結果に基づいて、前記第1の特徴要素に対してエントロピー符号化を実行するかどうかを決定するステップと、
前記第1の特徴要素に対してエントロピー符号化が実行される必要があると決定されるときのみ、前記第1の特徴要素に対してエントロピー符号化を実行するステップとを含む、
前記第1の特徴要素の前記確率推定結果に基づいて、前記第1の特徴要素に対してエントロピー復号を実行するかどうかを決定する前記ステップが、
前記第1の特徴要素の前記確率推定結果があらかじめ設定された条件を満たさないとき、前記特徴データの前記第1の特徴要素に対してエントロピー復号が実行される必要はないと決定し、前記第1の特徴要素の特徴値をkに設定するステップを含み、kが整数であり、kが前記第1の特徴要素の複数の候補値のうちの1つである、方法。

【請求項 2】

特徴データ復号方法であって、

10

20

復号されることになる特徴データのビットストリームを取得するステップであって、
前記復号されることになる特徴データが複数の特徴要素を含み、前記複数の特徴要素が
第1の特徴要素を含む、ステップと、

前記第1の特徴要素の確率推定結果を取得するステップと、

前記第1の特徴要素の前記確率推定結果に基づいて、前記第1の特徴要素に対してエント
ロピー復号を実行するかどうかを決定するステップと、

前記第1の特徴要素に対してエントロピー復号が実行される必要があると決定されると
きのみ、前記第1の特徴要素に対してエントロピー復号を実行するステップと
を含み、

前記第1の特徴要素の前記確率推定結果に基づいて、前記第1の特徴要素に対してエント
ロピー復号を実行するかどうかを決定する前記ステップが、

10

前記第1の特徴要素の前記確率推定結果があらかじめ設定された条件を満たさないとき、
前記特徴データの第1の特徴要素に対してエントロピー復号が実行される必要はない
と決定し、前記第1の特徴要素の特徴値を k に設定するステップを含み、 k が整数であり、 k
が前記第1の特徴要素の複数の候補値のうちの1つである、方法。

【請求項3】

前記第1の特徴要素の前記確率推定結果に基づいて、前記第1の特徴要素に対してエント
ロピー復号を実行するかどうかを決定する前記ステップが、

前記第1の特徴要素の前記確率推定結果が前記あらかじめ設定された条件を満たすとき
、前記特徴データの第1の特徴要素に対してエントロピー復号が実行される必要があ
ると決定するステップをさらに含む、請求項2に記載の方法。

20

【請求項4】

前記第1の特徴要素の前記確率推定結果が、前記第1の特徴要素の前記特徴値が k である
確率値であるとき、前記あらかじめ設定された条件が、前記第1の特徴要素の前記特徴値
が k である前記確率値が第1の閾値以下であることであり、 k が整数であり、 k が前記第1の
特徴要素の前記複数の候補値のうちの1つである、請求項3に記載の方法。

【請求項5】

前記第1の特徴要素の前記確率推定結果が、前記第1の特徴要素の確率分布の第1のパラ
メータおよび第2のパラメータを含むとき、前記あらかじめ設定された条件が、

前記第1の特徴要素の前記確率分布の前記第1のパラメータと前記第1の特徴要素の前記
特徴値 k との間の差の絶対値が第2の閾値以上であること、

30

前記第1の特徴要素の前記確率分布の前記第2のパラメータが第3の閾値以上であること
、または

前記第1の特徴要素の前記確率分布の前記第2のパラメータと、前記第1の特徴要素の前
記確率分布の前記第1のパラメータと前記第1の特徴要素の前記特徴値 k との間の差の絶対
値と、の和が第4の閾値以上であることであり、

k が整数であり、 k が前記第1の特徴要素の前記複数の候補値のうちの1つである、請求項
3に記載の方法。

【請求項6】

前記確率分布がガウス分布であるとき、前記第1の特徴要素の前記確率分布の前記第1の
パラメータが、前記第1の特徴要素の前記ガウス分布の平均値であり、前記第1の特徴要素
の前記確率分布の前記第2のパラメータが、前記第1の特徴要素の前記ガウス分布の分散で
あり、または、

40

前記確率分布がラプラス分布であるとき、前記第1の特徴要素の前記確率分布の前記第1
のパラメータが、前記第1の特徴要素の前記ラプラス分布の位置パラメータであり、前記
第1の特徴要素の前記確率分布の前記第2のパラメータが、前記第1の特徴要素の前記ラプ
ラス分布のスケールパラメータである、請求項5に記載の方法。

【請求項7】

前記第1の特徴要素の前記確率推定結果が混合ガウス分布を通じて取得されるとき、前
記あらかじめ設定された条件が、

50

前記第1の特徴要素の前記混合ガウス分布の任意の分散と、前記第1の特徴要素の前記混合ガウス分布のすべての平均値と前記第1の特徴要素の前記特徴値 k との間の差の絶対値の和と、の和が第5の閾値以上であること、

前記第1の特徴要素の前記混合ガウス分布の任意の平均値と前記第1の特徴要素の前記特徴値 k との間の差が第6の閾値以上であること、または

前記第1の特徴要素の前記混合ガウス分布の任意の分散が第7の閾値以上であることであり、

k が整数であり、 k が前記第1の特徴要素の前記複数の候補値のうちの1つである、請求項3に記載の方法。

【請求項8】

前記第1の特徴要素の前記確率推定結果が非対称ガウス分布を通じて取得されるとき、前記あらかじめ設定された条件が、

前記第1の特徴要素の前記非対称ガウス分布の平均値と前記第1の特徴要素の前記特徴値 k との間の差の絶対値が第8の閾値以上であること、

前記第1の特徴要素の前記非対称ガウス分布の第1の分散が第9の閾値以上であること、または

前記第1の特徴要素の前記非対称ガウス分布の第2の分散が第10の閾値以上であることであり、

k が整数であり、 k が前記第1の特徴要素の前記複数の候補値のうちの1つである、請求項3に記載の方法。

【請求項9】

前記第1の特徴要素の前記確率推定結果に基づいて、前記第1の特徴要素に対してエントロピー復号を実行するかどうかを決定する前記ステップが、

前記特徴データの確率推定結果を生成ネットワークに入力して前記第1の特徴要素の判定情報を取得するステップと、

前記第1の特徴要素の前記判定情報に基づいて、前記第1の特徴要素に対してエントロピー復号を実行するかどうかを決定するステップとを含む、請求項2に記載の方法。

【請求項10】

前記特徴データの判定情報が判定マップであり、かつ前記第1の特徴要素が前記判定マップにおいて位置する位置に対応する値があらかじめ設定された値であるとき、前記第1の特徴要素に対してエントロピー復号が実行される必要があると決定され、

前記第1の特徴要素が前記判定マップにおいて位置する前記位置に対応する前記値が前記あらかじめ設定された値ではないとき、前記第1の特徴要素に対してエントロピー復号が実行される必要はないと決定される、請求項9に記載の方法。

【請求項11】

前記方法が、

前記特徴データがデコーダネットワークを通過した後に取得される、再構築されたデータまたは機械向けタスクデータを取得するステップをさらに含む、請求項2から10のいずれか一項に記載の方法。

【請求項12】

請求項1に記載の方法を実行するように構成された処理回路を含むエンコーダ。

【請求項13】

請求項2から10のいずれか一項に記載の方法を実行するように構成された処理回路を含むデコーダ。

【発明の詳細な説明】

【技術分野】

【0001】

本発明の実施形態は、人工知能(AI)に基づくピクチャまたはオーディオ圧縮技術の分野、特に、特徴データ符号化および復号方法および装置に関する。

10

20

30

40

50

【背景技術】**【0002】**

ピクチャまたはオーディオ符号化および復号(略して符号化および復号)は、デジタルテレビ放送、インターネットおよびモバイルネットワーク上のピクチャまたはオーディオの送信、ビデオまたは音声チャット、ビデオまたは音声会議などのリアルタイム会話アプリケーション、DVDおよびブルーレイディスク、ピクチャまたはオーディオコンテンツ取り込みおよび編集システム、およびカムコーダの安全なアプリケーションなどの、デジタルピクチャまたはオーディオアプリケーションにおいて広く使用される。ビデオはピクチャの複数のフレームを含む。したがって、この出願におけるピクチャは、単一のピクチャであってもよく、またはビデオの中のピクチャであってもよい。

10

【0003】

短いビデオでさえ描写するために必要とされる大量のビデオデータは相当であることがあり、これは、限られた帯域幅容量を有するネットワークにわたってデータがストリーミングまたは通信されることになるとき、困難さをもたらす得る。したがって、ピクチャ(またはオーディオ)データは、一般に、現代の遠隔通信ネットワークにわたって通信される前に圧縮される。メモリリソースは限られていることがあるので、ピクチャ(またはオーディオ)データが記憶デバイスに記憶されるとき、ピクチャ(またはオーディオ)データのサイズも問題であり得る。ピクチャ(またはオーディオ)圧縮デバイスは、しばしば、送信または記憶の前にピクチャ(またはオーディオ)データを符号化するために、ソース側のソフトウェアおよび/またはハードウェアを使用する。これは、デジタルピクチャ(またはオーディオ)を示すために必要とされるデータの量を減少させる。次いで、圧縮されたデータは、ピクチャ(またはオーディオ)復元デバイスによってデスティネーション側で受信される。限られたネットワークリソースとより高いピクチャ(またはオーディオ)品質のいっそう増加する需要により、ピクチャ(またはオーディオ)品質においてほとんどまたはまったく犠牲なく圧縮比を改善する、改善された圧縮および復元技法が望ましい。

20

【0004】

近年、深層学習がピクチャ(またはオーディオ)符号化および復号の分野において人気を獲得している。たとえば、グーグルは、連続する数年間、CVPR(IEEE Conference on Computer Vision and Pattern Recognition)においてCLIC(Challenge on Learned Image Compression)競技会を開催した。CLICは、ピクチャ圧縮効率を改善するために深層ニューラルネットワークを使用することに焦点を合わせる。ピクチャチャレンジカテゴリもCLIC2020に追加された。その競技会の解決策の性能評価に基づいて、深層学習技術に基づく現在のピクチャ符号化および復号の解決策の包括的な圧縮効率は、最新世代のビデオピクチャ符号化および復号規格VVC(Versatile Video Coding)のものと同等であり、ユーザにより知覚される品質を改善することにおいて特有の利点を有する。

30

【0005】

VVCのビデオ規格は、2020年6月に完成した。この規格は、圧縮効率を著しく改善することができるほとんどすべての技術的アルゴリズムを含む。したがって、従来の信号処理経路に沿って新しい圧縮コーディングアルゴリズムを研究し続けるために、短時間で技術において躍進することは難しい。手作業の設計を通じてピクチャ圧縮のモジュールを最適化する従来のピクチャアルゴリズムとは異なり、エンドツーエンドのAIピクチャ圧縮は全体として最適化される。したがって、AIピクチャ圧縮はより高い圧縮効果を有する。変分オートエンコーダ(variational autoencoder, VAE)方法が、現在のAIピクチャ不可逆圧縮技術の主流の技術的解決策である。現在の主流の技術的解決策では、ピクチャ特徴マップが、エンコーダネットワークを使用することによって、符号化されることになるピクチャのために取得され、ピクチャ特徴マップに対してエントロピー符号化がさらに実行される。しかしながら、エントロピー符号化プロセスは過度に複雑である。

40

【発明の概要】**【課題を解決するための手段】****【0006】**

50

この出願は、符号化および復号の性能に影響することなく符号化および復号の複雑さを減らすために、特徴データの符号化および復号の方法および装置を提供する。

【0007】

第1の態様によれば、

符号化されることになる特徴データを取得するステップであって、符号化されることになる特徴データが複数の特徴要素を含み、複数の特徴要素が第1の特徴要素を含む、ステップと、

第1の特徴要素の確率推定結果を取得するステップと、

第1の特徴要素の確率推定結果に基づいて、第1の特徴要素に対してエントロピー符号化を実行するかどうかを決定するステップと、

第1の特徴要素に対してエントロピー符号化が実行される必要があると決定されるときのみ、第1の特徴要素に対してエントロピー符号化を実行するステップとを含む、特徴データ符号化方法が提供される。

10

【0008】

特徴データは、ピクチャ特徴マップ、オーディオ特徴変数、またはピクチャ特徴マップおよびオーディオ特徴変数を含み、エンコーダネットワークによって出力される1次元、2次元、または多次元データであってもよく、データの各々は特徴要素である。この出願における特徴点および特徴要素の意味は同じであることが留意されるべきである。

【0009】

具体的には、第1の特徴要素は、符号化されることになる特徴データの任意の符号化されることになる特徴要素である。

20

【0010】

可能性において、第1の特徴要素の確率推定結果を取得する確率推定プロセスが、確率推定ネットワークを使用することによって実施され得る。別の可能性において、確率推定プロセスは、特徴データに対する確率推定を実行するために従来の非ネットワーク確率推定方法を使用し得る。

【0011】

サイド情報のみが確率推定の入力として使用されるとき、特徴要素の確率推定結果が並列に出力されることが留意されるべきである。確率推定の入力がコンテキスト情報を含むとき、特徴要素の確率推定結果は直列に出力される必要がある。サイド情報は、特徴データをニューラルネットワークに入力することによってさらに抽出される特徴情報であり、サイド情報に含まれる特徴要素の数量は、特徴データの特徴要素の数量より少ない。任意選択で、特徴データのサイド情報はビットストリームへと符号化され得る。

30

【0012】

可能性において、特徴データの第1の特徴要素があらかじめ設定された条件を満たさないとき、エントロピー符号化は、特徴データの第1の特徴要素に対して実行される必要はない。

【0013】

具体的には、現在の第1の特徴要素が特徴データの第Pの特徴要素であるならば、第Pの特徴要素の決定が完了し、決定結果に基づいてエントロピー符号化が実行され、または実行されない後、特徴データの第(P+1)の特徴要素の決定が開始され、エントロピー符号化プロセスが決定結果に基づいて実行され、または実行されない。Pは正の整数でありPはMより小さく、Mは特徴データ全体の特徴要素の数量である。たとえば、第2の特徴要素について、第2の特徴要素に対してエントロピー符号化が実行される必要がないと決定されるとき、第2の特徴要素に対してエントロピー符号化を実行することがスキップされる。

40

【0014】

前述の技術的解決策では、エントロピー符号化が実行される必要があるかどうかは、各々の符号化されることになる特徴要素について決定され、それによって、いくつかの特徴要素のエントロピー符号化プロセスがスキップされ、エントロピー符号化が実行される必要がある要素の数量が著しく減らされることが可能である。このようにして、エントロピ

50

ー符号化の複雑さが減らされることが可能である。

【 0 0 1 5 】

可能な実装では、第1の特徴要素に対してエントロピー符号化を実行するかどうかを決定することは、第1の特徴要素の確率推定結果があらかじめ設定された条件を満たすとき、第1の特徴要素に対してエントロピー符号化が実行される必要があると決定すること、または、第1の特徴要素の確率推定結果があらかじめ設定された条件を満たさないとき、第1の特徴要素に対してエントロピー符号化が実行される必要がないと決定することを含む。

【 0 0 1 6 】

可能な実装では、第1の特徴要素の確率推定結果が、第1の特徴要素の値が k である確率値であるとき、あらかじめ設定された条件は、第1の特徴要素の値が k である確率値が第1の閾値以下であることであり、 k は整数である。

10

【 0 0 1 7 】

k は、第1の特徴要素の値の可能な値の範囲内の値である。たとえば、第1の特徴要素の値の範囲は $[-255, 255]$ であり得る。 k は0に設定されてもよく、エントロピー符号化は、確率値が0.5以下である第1の特徴要素に対して実行される。エントロピー符号化は、確率値が0.5より大きい第1の特徴要素に対して実行されない。

【 0 0 1 8 】

可能な実装では、第1の特徴要素の値が k である確率値は、第1の特徴要素のすべての可能な値の確率値の中の最大の確率値である。

20

【 0 0 1 9 】

低ビットレートの場合に符号化されたビットストリームのために選択される第1の閾値は、高ビットレートの場合に符号化されたビットストリームのために選択される第1の閾値より小さい。具体的なビットレートは、ピクチャ解像度およびピクチャコンテンツに関連する。たとえば、公開のKodakデータセットが使用される。0.5bppより低いビットレートは低ビットレートである。そうでなければ、ビットレートは高ビットレートである。

【 0 0 2 0 】

具体的なビットレートの場合において、第1の閾値は実際の要件に基づいて構成され得る。これはここで限定されない。

【 0 0 2 1 】

前述の技術的解決策では、エントロピー符号化の複雑さが、要件に基づいて、かつ柔軟な第1の閾値を柔軟に設定することによって、柔軟に減らされることが可能である。

30

【 0 0 2 2 】

可能な実装では、第1の特徴要素の確率推定結果は、第1の特徴要素の確率分布の第1のパラメータおよび第2のパラメータを含む。

【 0 0 2 3 】

確率分布がガウス分布であるとき、第1の特徴要素の確率分布の第1のパラメータは、第1の特徴要素のガウス分布の平均値であり、第1の特徴要素の確率分布の第2のパラメータは、第1の特徴要素のガウス分布の分散である。代替として、確率分布がラプラス分布であるとき、第1の特徴要素の確率分布の第1のパラメータは、第1の特徴要素のラプラス分布の位置パラメータであり、第1の特徴要素の確率分布の第2のパラメータは、第1の特徴要素のラプラス分布のスケールパラメータである。あらかじめ設定された条件は、以下、すなわち、

40

第1の特徴要素の確率分布の第1のパラメータと第1の特徴要素の値 k との間の差の絶対値が第2の閾値以上であること、

第1の特徴要素の確率分布の第2のパラメータが第3の閾値以上であること、または

第1の特徴要素の確率分布の第2のパラメータと、第1の特徴要素の確率分布の第1のパラメータと第1の特徴要素の値 k との間の差の絶対値と、の和が第4の閾値以上であることのいずれか1つであり得る。

【 0 0 2 4 】

50

確率分布が混合ガウス分布であるとき、第1の特徴要素の確率分布の第1のパラメータは、第1の特徴要素の混合ガウス分布の平均値であり、第1の特徴要素の確率分布の第2のパラメータは、第1の特徴要素の混合ガウス分布の分散である。あらかじめ設定された条件は、以下、すなわち、

第1の特徴要素の混合ガウス分布の任意の分散と、第1の特徴要素の混合ガウス分布のすべての平均値と第1の特徴要素の値 k との間の差の絶対値の和と、の和が第5の閾値以上であること、

第1の特徴要素の混合ガウス分布の任意の平均値と第1の特徴要素の値 k との間の差が第6の閾値以上であること、または

第1の特徴要素の混合ガウス分布の任意の分散が第7の閾値以上であること
のいずれか1つであり得る。

10

【0025】

確率分布が非対称ガウス分布であるとき、第1の特徴要素の確率分布の第1のパラメータは、第1の特徴要素の非対称ガウス分布の平均値であり、第1の特徴要素の確率分布の第2のパラメータは、第1の特徴要素の非対称ガウス分布の第1の分散および第2の分散である。あらかじめ設定された条件は、以下、すなわち、

第1の特徴要素の非対称ガウス分布の平均値と第1の特徴要素の値 k との間の差の絶対値が第8の閾値以上であること、

第1の特徴要素の非対称ガウス分布の第1の分散が第9の閾値以上であること、または

第1の特徴要素の非対称ガウス分布の第2の分散が第10の閾値以上であること
のいずれか1つであり得る。

20

【0026】

第1の特徴要素の確率分布が混合ガウス分布であるとき、第1の特徴要素の決定値の範囲が決定される。第1の特徴要素の確率分布の複数の平均値は、第1の特徴要素の決定値の範囲内にない。

【0027】

第1の特徴要素の確率分布がガウス分布であるとき、第1の特徴要素の決定値の範囲が決定される。第1の特徴要素の確率分布の平均値は、第1の特徴要素の決定値の範囲内にない。

【0028】

第1の特徴要素の確率分布がガウス分布であるとき、第1の特徴要素の決定値の範囲が決定され、決定値の範囲は第1の特徴要素の複数の可能な値を含む。第1の特徴要素のガウス分布の平均値パラメータと第1の特徴要素の決定値の範囲内の各値との間の差の絶対値は第11の閾値以上であり、または、第1の特徴要素の確率分布の分散は第12の閾値以上である。

30

【0029】

第1の特徴要素の値は、第1の特徴要素の決定値の範囲内にない。

【0030】

第1の特徴要素の値に対応する確率値は、第13の閾値以下である。

【0031】

可能な実装では、方法は、第1の閾値の閾値候補リストを構築するステップと、第1の閾値を第1の閾値の閾値候補リストに入れるステップであって、第1の閾値に対応するインデックス番号がある、ステップと、第1の閾値のインデックス番号を符号化されたビットストリームに書き込むステップであって、第1の閾値の閾値候補リストの長さは T に設定されてもよく、 T は1以上の整数である、ステップとをさらに含む。第1の閾値の閾値候補リストを構築するなどの方式で、別の閾値が構築されてもよいことが理解され得る。別の閾値は、符号化されたビットストリームに書き込まれる対応するインデックス番号を有する。

40

【0032】

具体的には、インデックス番号はビットストリームに書き込まれ、シーケンスヘッダ(sequence header)、ピクチャヘッダ(picture header)、slice / スライスヘッダ(slice he

50

ader)、またはSEI(supplemental enhancement information)に記憶され、デコーダ側に送信され得る。代替として、別の方法が使用され得る。これはここで限定されない。候補リストを構築する方式は限定されない。

【0033】

別の可能性において、確率推定結果を生成ネットワークに入力することによって、判定情報が取得される。生成ネットワークは、畳み込みネットワークであってもよく、複数のネットワーク層を含んでもよい。任意のネットワーク層が、畳み込み層、正規化層、非線形活性化層、または同様のものであってもよい。

【0034】

可能な実装では、特徴データの確率推定結果は、第1の特徴要素の判定情報を取得するために生成ネットワークに入力される。判定情報は、第1の特徴要素に対してエントロピー符号化を実行するかどうかを示す。

10

【0035】

可能な実装では、特徴データの判定情報は判定マップであり、判定マップは判定mapとも呼ばれ得る。判定マップは好ましくは二進マップであり、二進マップは二進mapとも呼ばれ得る。二進マップにおける特徴要素の判定情報の値は、通常は0または1である。したがって、第1の特徴要素が判定マップにおいて位置する位置に対応する値があらかじめ設定された値であるとき、エントロピー符号化が第1の特徴要素に対して実行される必要がある。第1の特徴要素が判定マップにおいて位置する位置に対応する値があらかじめ設定された値ではないとき、エントロピー符号化が第1の特徴要素に対して実行される必要はない。

20

【0036】

可能な実装では、特徴データの特徴要素の判定情報は、あらかじめ設定された値である。判定情報のあらかじめ設定された値は、通常は1である。したがって、判定情報があらかじめ設定された値であるとき、エントロピー符号化が第1の特徴要素に対して実行される必要がある。判定情報があらかじめ設定された値ではないとき、第1の特徴要素に対してエントロピー符号化が実行される必要はない。判定情報は、識別子または識別子の値であり得る。第1の特徴要素に対してエントロピー符号化を実行するかどうかを決定することは、識別子または識別子の値があらかじめ設定された値であるかどうかに依存する。識別子または識別子の値があらかじめ設定された値であるとき、エントロピー符号化が第1の特徴要素に対して実行される必要がある。識別子または識別子の値があらかじめ設定された値ではないとき、エントロピー符号化が第1の特徴要素に対して実行される必要はない。代替として、特徴データの特徴要素の判定情報のセットは、浮動小数点数であってもよい。言い換えると、値は0および1以外の別の値であってもよい。この場合、あらかじめ設定された値が設定されてもよい。第1の特徴要素の判定情報の値があらかじめ設定された値以上であるとき、第1の特徴要素に対してエントロピー符号化が実行される必要があると決定される。第1の特徴要素の判定情報の値があらかじめ設定された値より小さいとき、第1の特徴要素に対してエントロピー符号化が実行される必要はないと決定される。

30

【0037】

可能な実装では、方法は、符号化されることになるピクチャがエンコーダネットワークを通過することによって特徴データを取得するステップ、符号化されることになるピクチャがエンコーダネットワークを通過した後で符号化されることになるピクチャを丸めることによって特徴データを取得するステップ、または符号化されることになるピクチャがエンコーダネットワークを通過した後で符号化されることになるピクチャを量子化し、丸めることによって特徴データを取得するステップをさらに含む。

40

【0038】

エンコーダネットワークは、オートエンコーダ構造を使用し得る。エンコーダネットワークは、畳み込みニューラルネットワークであり得る。エンコーダネットワークは複数のサブネットを含んでもよく、各サブネットは1つまたは複数の畳み込み層を含む。サブネット間のネットワーク構造は、同じであっても異なってもよい。

50

【 0 0 3 9 】

符号化されることになるピクチャは、元のピクチャまたは残差ピクチャであり得る。

【 0 0 4 0 】

符号化されることになるピクチャは、RGBフォーマット、またはYUVまたはRAWなどの表現フォーマットであってもよいことが理解されるべきである。エンコーダネットワークに入力される前に、符号化されることになるピクチャに対して前処理動作が実行され得る。前処理動作は、変換、ブロック分割、フィルタリング、およびブルーニングなどの動作を含み得る。

【 0 0 4 1 】

複数の符号化されることになるピクチャまたは複数の符号化されることになるピクチャブロックは、特徴データを取得するために、同じタイムスタンプ内で、または同じ瞬間に処理するために、エンコーダおよびデコーダネットワークに入力されることが許容されることが理解されるべきである。

10

【 0 0 4 2 】

第2の態様によれば、

復号されることになる特徴データのビットストリームを取得するステップであって、

復号されることになる特徴データが複数の特徴要素を含み、複数の特徴要素が第1の特徴要素を含む、ステップと、

第1の特徴要素の確率推定結果を取得するステップと、

第1の特徴要素の確率推定結果に基づいて、第1の特徴要素に対してエントロピー復号を実行するかどうかを決定するステップと、

20

第1の特徴要素に対してエントロピー復号が実行される必要があると決定されるときのみ、第1の特徴要素に対してエントロピー復号を実行するステップとを含む、特徴データ復号方法が提供される。

【 0 0 4 3 】

第1の特徴要素は、復号されることになる特徴データの任意の特徴要素であることが理解され得る。復号されることになる特徴データのすべての特徴要素の決定が完了し、決定結果に基づいてエントロピー復号が実行され、または実行されない後、復号された特徴データが取得される。

【 0 0 4 4 】

復号された特徴データは、1次元、2次元、または多次元データであってもよく、データの各々は特徴要素である。この出願における特徴点および特徴要素の意味は同じであることが留意されるべきである。

30

【 0 0 4 5 】

具体的には、第1の特徴要素は、復号されることになる特徴データの任意の復号されることになる特徴要素である。

【 0 0 4 6 】

可能性において、第1の特徴要素の確率推定結果を取得する確率推定プロセスが、確率推定ネットワークを使用することによって実施され得る。別の可能性において、確率推定プロセスは、特徴データに対して確率推定を実行するために従来の非ネットワーク確率推定方法を使用し得る。

40

【 0 0 4 7 】

サイド情報のみが確率推定の入力として使用されるとき、特徴要素の確率推定結果が並列に出力されることが留意されるべきである。確率推定の入力がコンテキスト情報を含むとき、特徴要素の確率推定結果は直列に出力される必要がある。サイド情報に含まれる特徴要素の数量は、特徴データの特徴要素の数量より少ない。

【 0 0 4 8 】

可能性において、ビットストリームはサイド情報を含み、サイド情報はビットストリームを復号するプロセスにおいて復号される必要がある。

【 0 0 4 9 】

50

具体的には、特徴データの各特徴要素の決定プロセスは、条件決定と、条件決定結果に基づいてエントロピー復号を実行するかどうかを決定することを含む。

【0050】

可能性において、エントロピー復号は、ニューラルネットワークを使用することによって実施され得る。

【0051】

別の可能性において、エントロピー復号は、従来のエントロピー復号を通じて実施され得る。

【0052】

具体的には、現在の第1の特徴要素が特徴データの第Pの特徴要素であるならば、第Pの特徴要素の決定が完了し、決定結果に基づいてエントロピー復号が実行され、または実行されない後、特徴データの第(P+1)の特徴要素の決定が開始され、エントロピー復号プロセスが決定結果に基づいて実行され、または実行されない。Pは正の整数であり、PはMより小さく、Mは特徴データ全体の特徴要素の数量である。たとえば、第2の特徴要素について、第2の特徴要素に対してエントロピー復号が実行される必要がないと決定されるとき、第2の特徴要素に対してエントロピー復号を実行することがスキップされる。

10

【0053】

前述の技術的解決策では、エントロピー復号が実行される必要があるかどうかは、各々の復号されることになる特徴要素について決定され、それによって、いくつかの特徴要素のエントロピー復号プロセスがスキップされ、エントロピー復号が実行される必要がある要素の数量が著しく減らされることが可能である。このようにして、エントロピー復号の複雑さが減らされることが可能である。

20

【0054】

可能な実装では、特徴データの第1の特徴要素に対してエントロピー復号を実行するかどうかを決定することは、特徴データの第1の特徴要素の確率推定結果があらかじめ設定された条件を満たすとき、第1の特徴要素に対してエントロピー復号が実行される必要があると決定すること、または、第1の特徴要素の確率推定結果があらかじめ設定された条件を満たさないとき、第1の特徴要素に対してエントロピー復号が実行される必要がないと決定し、第1の特徴要素の特徴値をkに設定することを含み、kは整数である。

【0055】

30

可能な実装では、第1の特徴要素の確率推定結果が、第1の特徴要素の値がkである確率値であるとき、あらかじめ設定された条件は、第1の特徴要素の値がkである確率値が第1の閾値以下であることであり、kは整数である。

【0056】

可能性において、あらかじめ設定された条件が満たされないとき、第1の特徴要素はkに設定される。たとえば、第1の特徴要素の値の範囲は $[-255, 255]$ であり得る。kは0に設定されてもよく、エントロピー符号化は、確率値が0.5以下である第1の特徴要素に対して実行される。エントロピー符号化は、確率値が0.5より大きい第1の特徴要素に対して実行されない。

【0057】

40

別の可能性において、あらかじめ設定された条件が満たされないとき、第1の特徴要素の値はリストを使用することによって決定される。

【0058】

別の可能性において、あらかじめ設定された条件が満たされないとき、第1の特徴要素は固定された整数値に設定される。

【0059】

kは、第1の特徴要素の値の可能な値の範囲内の値である。

【0060】

可能性において、kは、第1の特徴要素のすべての可能な値の範囲内の最大の確率に対応する値である。

50

【 0 0 6 1 】

低ビットレートの場合に復号されたビットストリームのために選択される第1の閾値は、高ビットレートの場合に復号されたビットストリームのために選択される第1の閾値より小さい。具体的なビットレートは、ピクチャ解像度およびピクチャコンテンツに関連する。たとえば、公開のKodakデータセットが使用される。0.5bppより低いビットレートは低ビットレートである。そうでなければ、ビットレートは高ビットレートである。

【 0 0 6 2 】

具体的なビットレートの場合において、第1の閾値は実際の要件に基づいて構成され得る。これはここで限定されない。

【 0 0 6 3 】

前述の技術的解決策では、エントロピー復号の複雑さが、要件に基づいて、かつ柔軟な第1の閾値を柔軟に設定することによって、柔軟に減らされることが可能である。

【 0 0 6 4 】

可能な実装では、第1の特徴要素の確率推定結果は、第1の特徴要素の確率分布の第1のパラメータおよび第2のパラメータを含む。

【 0 0 6 5 】

確率分布がガウス分布であるとき、第1の特徴要素の確率分布の第1のパラメータは、第1の特徴要素のガウス分布の平均値であり、第1の特徴要素の確率分布の第2のパラメータは、第1の特徴要素のガウス分布の分散である。代替として、確率分布がラプラス分布であるとき、第1の特徴要素の確率分布の第1のパラメータは、第1の特徴要素のラプラス分布の位置パラメータであり、第1の特徴要素の確率分布の第2のパラメータは、第1の特徴要素のラプラス分布のスケールパラメータである。あらかじめ設定された条件は、以下、すなわち、

第1の特徴要素の確率分布の第1のパラメータと第1の特徴要素の値 k との間の差の絶対値が第2の閾値以上であること、

第1の特徴要素の第2のパラメータが第3の閾値以上であること、または

第1の特徴要素の確率分布の第2のパラメータと、第1の特徴要素の確率分布の第1のパラメータと第1の特徴要素の値 k との間の差の絶対値と、の和が第4の閾値以上であることのいずれか1つであり得る。

【 0 0 6 6 】

確率分布が混合ガウス分布であるとき、第1の特徴要素の確率分布の第1のパラメータは、第1の特徴要素の混合ガウス分布の平均値であり、第1の特徴要素の確率分布の第2のパラメータは、第1の特徴要素の混合ガウス分布の分散である。あらかじめ設定された条件は、以下、すなわち、

第1の特徴要素の混合ガウス分布の任意の分散と、第1の特徴要素の混合ガウス分布のすべての平均値と第1の特徴要素の値 k との間の差の絶対値の和と、の和が第5の閾値以上であること、

第1の特徴要素の混合ガウス分布の任意の平均値と第1の特徴要素の値 k との間の差が第6の閾値より大きいこと、または

第1の特徴要素の混合ガウス分布の任意の分散が第7の閾値以上であることのいずれか1つであり得る。

【 0 0 6 7 】

確率分布が非対称ガウス分布であるとき、第1の特徴要素の確率分布の第1のパラメータは、第1の特徴要素の非対称ガウス分布の平均値であり、第1の特徴要素の確率分布の第2のパラメータは、第1の特徴要素の非対称ガウス分布の第1の分散および第2の分散である。あらかじめ設定された条件は、以下、すなわち、

第1の特徴要素の非対称ガウス分布の平均値パラメータと第1の特徴要素の値 k との間の差の絶対値が第8の閾値より大きいこと、

第1の特徴要素の非対称ガウス分布の第1の分散が第9の閾値以上であること、または

第1の特徴要素の非対称ガウス分布の第2の分散が第10の閾値以上であること

10

20

30

40

50

のいずれか1つであり得る。

【0068】

第1の特徴要素の確率分布が混合ガウス分布であるとき、第1の特徴要素の決定値の範囲が決定される。第1の特徴要素の確率分布の複数の平均値は、第1の特徴要素の決定値の範囲内にない。

【0069】

第1の特徴要素の確率分布がガウス分布であるとき、第1の特徴要素の決定値の範囲が決定される。第1の特徴要素の確率分布の平均値は、第1の特徴要素の決定値の範囲内にない。

【0070】

第1の特徴要素の確率分布がガウス分布であるとき、第1の特徴要素の決定値の範囲が決定され、決定値の範囲は第1の特徴要素の複数の可能な値を含む。第1の特徴要素のガウス分布の平均値パラメータと第1の特徴要素の決定値の範囲内の各値との間の差の絶対値は第11の閾値以上であり、または、第1の特徴要素の確率分布の分散は第12の閾値以上である。

【0071】

第1の特徴要素の値kは、第1の特徴要素の決定値の範囲内にない。

【0072】

第1の特徴要素の値kに対応する確率値は、第13の閾値以下である。

【0073】

可能な実装では、第1の閾値の閾値候補リストが構築され、第1の閾値の閾値候補リストのインデックス番号は、ビットストリームを復号することによって取得され、第1の閾値のインデックス番号に対応する、第1の閾値の閾値候補リストの位置の値が、第1の閾値の値として使用される。第1の閾値の閾値候補リストの長さはTに設定されてもよく、Tは1以上の整数である。第1の閾値の閾値候補リストを構築するなどの方式で、任意の他の閾値が構築されてもよいことが理解され得る。閾値に対応するインデックス番号は、復号を通じて取得されてもよく、構築されたリストの中の値は、インデックス番号に基づいて閾値として選択される。

【0074】

別の可能性において、確率推定結果を生成ネットワークに入力することによって、判定情報が取得される。生成ネットワークは、畳み込みネットワークであってもよく、複数のネットワーク層を含んでもよい。任意のネットワーク層が、畳み込み層、正規化層、非線形活性化層、または同様のものであってもよい。

【0075】

可能な実装では、特徴データの確率推定結果は、第1の特徴要素の判定情報を取得するために生成ネットワークに入力される。判定情報は、第1の特徴要素に対してエントロピー復号を実行するかどうかを示す。

【0076】

可能な実装では、特徴データの特徴要素の判定情報は判定マップであり、判定マップは判定mapとも呼ばれ得る。判定マップは好ましくは二進マップであり、二進マップは二進mapとも呼ばれ得る。二進マップにおける特徴要素の判定情報の値は、通常は0または1である。したがって、第1の特徴要素が判定マップにおいて位置する位置に対応する値があらかじめ設定された値であるとき、エントロピー復号が第1の特徴要素に対して実行される必要がある。第1の特徴要素が判定マップにおいて位置する位置に対応する値があらかじめ設定された値ではないとき、エントロピー復号が第1の特徴要素に対して実行される必要はない。

【0077】

代替として、特徴データの特徴要素の判定情報のセットは、浮動小数点数であってもよい。言い換えると、値は0および1以外の別の値であってもよい。この場合、あらかじめ設定された値が設定されてもよい。第1の特徴要素の判定情報の値があらかじめ設定された

10

20

30

40

50

値以上であるとき、第1の特徴要素に対してエントロピー復号が実行される必要があると決定される。第1の特徴要素の判定情報の値があらかじめ設定された値より小さいとき、第1の特徴要素に対してエントロピー復号が実行される必要はないと決定される。

【0078】

可能な実装では、特徴データがデコーダネットワークを通過することによって、再構築されたピクチャが取得される。

【0079】

別の可能な実装では、特徴データがデコーダネットワークを通過することによって、機械向けタスクデータが取得される。具体的には、特徴データが機械向けタスクモジュールを通過することによって、機械向けタスクデータが取得され、機械向けモジュールは、目標認識ネットワーク、分類ネットワーク、またはセマンティック区分ネットワークを含む。

10

【0080】

第3の態様によれば、

符号化されることになる特徴データを取得するように構成された取得モジュールであって、符号化されることになる特徴データが複数の特徴要素を含み、複数の特徴要素が第1の特徴要素を含み、第1の特徴要素の確率推定結果を取得するように構成された取得モジュールと、

第1の特徴要素の確率推定結果に基づいて、第1の特徴要素に対してエントロピー符号化を実行するかどうかを決定し、第1の特徴要素に対してエントロピー符号化が実行される必要があると決定されるときのみ、第1の特徴要素に対してエントロピー符号化を実行するように構成された符号化モジュールとを含む、特徴データ符号化装置が提供される。

20

【0081】

取得モジュールおよび符号化モジュールのさらなる実装機能については、第1の態様または第1の態様の実装のいずれか1つを参照されたい。詳細はここで再び説明されない。

【0082】

第4の態様によれば、

復号されることになる特徴データのビットストリームを取得するように構成された取得モジュールであって、復号されることになる特徴データが複数の特徴要素を含み、複数の特徴要素が第1の特徴要素を含み、第1の特徴要素の確率推定結果を取得するように構成された取得モジュールと、

30

第1の特徴要素の確率推定結果に基づいて、第1の特徴要素に対してエントロピー復号を実行するかどうかを決定し、第1の特徴要素に対してエントロピー復号が実行される必要があると決定されるときのみ、第1の特徴要素に対してエントロピー復号を実行するように構成された復号モジュールとを含む、特徴データ復号装置が提供される。

【0083】

取得モジュールおよび復号モジュールのさらなる実装機能については、第2の態様または第2の態様の実装のいずれか1つを参照されたい。詳細はここで再び説明されない。

【0084】

40

第5の態様によれば、この出願は、第1の態様および第1の態様のいずれか1つによる方法を決定するように構成された処理回路を含むエンコーダを提供する。

【0085】

第6の態様によれば、この出願は、第2の態様および第2の態様のいずれか1つによる方法を決定するように構成された処理回路を含むデコーダを提供する。

【0086】

第7の態様によれば、この出願は、プログラムコードを含むコンピュータプログラム製品を提供する。プログラムコードがコンピュータまたはプロセッサにおいて決定されるとき、プログラムコードは、第1の態様および第1の態様のいずれか1つによる方法、および第2の態様および第2の態様のいずれか1つによる方法を決定するために使用される。

50

【 0 0 8 7 】

第8の態様によれば、この出願は、1つまたは複数のプロセッサを含むエンコーダと、プロセッサに結合され、プロセッサによって決定されたプログラムを記憶する非一時的コンピュータ可読記憶媒体とを提供する。プログラムがプロセッサによって決定されるとき、デコーダは、第1の態様および第1の態様のいずれか1つによる方法を決定することを可能にされる。

【 0 0 8 8 】

第9の態様によれば、この出願は、1つまたは複数のプロセッサを含むデコーダと、プロセッサに結合され、プロセッサによって決定されたプログラムを記憶する非一時的コンピュータ可読記憶媒体とを提供する。プログラムがプロセッサによって決定されるとき、エンコーダは、第2の態様および第2の態様のいずれか1つによる方法を決定することを可能にされる。

10

【 0 0 8 9 】

第10の態様によれば、この出願は、プログラムコードを含む非一時的コンピュータ可読記憶媒体を提供する。プログラムコードがコンピュータデバイスによって決定されるとき、プログラムコードは、第1の態様および第1の態様のいずれか1つによる方法、および第2の態様および第2の態様のいずれか1つによる方法を決定するために使用される。

【 0 0 9 0 】

第11の態様によれば、本発明は符号化装置に関し、これは、第1の態様または第1の態様の方法の実施形態のいずれか1つによる挙動を実施する機能を有する。機能は、ハードウェアによって実装されてもよく、または対応するソフトウェアを決定するハードウェアによって実装されてもよい。ハードウェアまたはソフトウェアは、前述の機能に対応する1つまたは複数のモジュールを含む。可能な設計では、符号化装置は、エンコーダネットワークを使用することによって元のピクチャまたは残差ピクチャを特徴空間に変換し、圧縮のために特徴データを抽出するように構成された取得モジュールであって、確率推定が、特徴データに対して実行されて、特徴データの特徴要素の確率推定結果を取得する、取得モジュールと、特徴データの特徴要素の確率推定結果を使用することによって、特定の条件に基づいて、エントロピー符号化が特徴データの特徴要素に対して実行されるかどうかを決定し、特徴データのすべての特徴要素の符号化プロセスを完了して、特徴データの符号化されたビットストリームを取得するように構成された符号化モジュールとを含む。これらのモジュールは、第1の態様および第1の態様のいずれか1つによる方法の例において対応する機能を決定し得る。詳細については、方法の例における詳細な説明を参照されたい。詳細はここで再び説明されない。

20

30

【 0 0 9 1 】

第12の態様によれば、本発明は復号装置に関し、これは、第2の態様または第2の態様の方法の実施形態のいずれか1つによる挙動を実施する機能を有する。機能は、ハードウェアによって実装されてもよく、または対応するソフトウェアを決定するハードウェアによって実装されてもよい。ハードウェアまたはソフトウェアは、前述の機能に対応する1つまたは複数のモジュールを含む。可能な設計では、復号装置は、復号されることになる特徴データのビットストリームを取得し、復号されることになる特徴データのビットストリームに基づいて確率推定を実行して、特徴データの特徴要素の確率推定結果を取得するように構成された取得モジュールと、特徴データの特徴要素の確率推定結果を使用することによって、特定の条件に基づいて、エントロピー復号が特徴データの特徴要素に対して実行されるかどうかを決定し、特徴データのすべての特徴要素の復号プロセスを完了して特徴データを取得し、特徴データを復号して再構築されたピクチャまたは機械向けタスクデータを取得するように構成された復号モジュールとを含む。これらのモジュールは、第2の態様および第2の態様のいずれか1つによる方法の例において対応する機能を決定し得る。詳細については、方法の例における詳細な説明を参照されたい。詳細はここで再び説明されない。

40

【 0 0 9 2 】

50

第13の態様によれば、

符号化されることになる特徴データを取得するステップであって、特徴データが複数の特徴要素を含み、複数の特徴要素が第1の特徴要素を含む、ステップと、

特徴データのサイド情報を取得し、特徴データのサイド情報をジョイントネットワークに入力して第1の特徴要素の判定情報を取得するステップと、

第1の特徴要素の判定情報に基づいて、第1の特徴要素に対してエントロピー符号化を実行するかどうかを決定するステップと、

第1の特徴要素に対してエントロピー符号化が実行される必要があると決定されるときのみ、第1の特徴要素に対してエントロピー符号化を実行するステップとを含む、特徴データ符号化方法が提供される。

10

【0093】

特徴データは、エンコーダネットワークによって出力される1次元、2次元、または多次元データであり、データの各々は特徴要素である。

【0094】

可能性において、特徴データのサイド情報はビットストリームへと符号化され得る。サイド情報は、特徴データをニューラルネットワークに入力することによってさらに抽出される特徴情報であり、サイド情報に含まれる特徴要素の数量は、特徴データの特徴要素の数量より少ない。

【0095】

第1の特徴要素は特徴データの任意の特徴要素である。

20

【0096】

可能性において、特徴データの特徴要素の判定情報のセットは、判定マップなどの方式で表現され得る。判定マップは、1次元、2次元、または多次元ピクチャデータであり、判定マップのサイズは特徴データのサイズと一致する。

【0097】

可能性において、ジョイントネットワークは、第1の特徴要素の確率推定結果をさらに出力する。第1の特徴要素の確率推定結果は、第1の特徴要素の確率値、および/または、確率分布の第1のパラメータおよび第2のパラメータを含む。

【0098】

前述の技術的解決策では、エントロピー符号化が実行される必要があるかどうかは、各々の符号化されることになる特徴要素について決定され、それによって、いくつかの特徴要素のエントロピー符号化プロセスがスキップされ、エントロピー符号化が実行される必要がある要素の数量が著しく減らされることが可能である。このようにして、エントロピー符号化の複雑さが減らされることが可能である。

30

【0099】

可能性において、第1の特徴要素が判定マップにおいて位置する位置に対応する値があらかじめ設定された値であるとき、エントロピー符号化が第1の特徴要素に対して実行される必要がある。第1の特徴要素が判定マップにおいて位置する位置に対応する値があらかじめ設定された値ではないとき、エントロピー符号化が第1の特徴要素に対して実行される必要はない。

40

【0100】

第14の態様によれば、

復号されることになる特徴データのビットストリームおよび復号されることになる特徴データのサイド情報を取得するステップであって、

復号されることになる特徴データが複数の特徴要素を含み、複数の特徴要素が第1の特徴要素を含む、ステップと、

復号されることになる特徴データのサイド情報をジョイントネットワークに入力して第1の特徴要素の判定情報を取得するステップと、

第1の特徴要素の判定情報に基づいて、第1の特徴要素に対してエントロピー復号を実行するかどうかを決定するステップと、

50

第1の特徴要素に対してエントロピー復号が実行される必要があると決定されるときのみ、第1の特徴要素に対してエントロピー復号を実行するステップとを含む、特徴データ復号方法が提供される。

【0101】

可能性において、復号されることになる特徴データのビットストリームは、サイド情報を取得するために復号される。サイド情報に含まれる特徴要素の数量は、特徴データの特徴要素の数量より少ない。

【0102】

第1の特徴要素は特徴データの任意の特徴要素である。

【0103】

可能性において、特徴データの特徴要素の判定情報は、判定マップなどの方式で表現され得る。判定マップは、1次元、2次元、または多次元ピクチャデータであり、判定マップのサイズは特徴データのサイズと一致する。

【0104】

可能性において、ジョイントネットワークは、第1の特徴要素の確率推定結果をさらに出力する。第1の特徴要素の確率推定結果は、第1の特徴要素の確率値、および/または、確率分布の第1のパラメータおよび第2のパラメータを含む。

【0105】

可能性において、第1の特徴要素が判定マップにおいて位置する位置に対応する値があらかじめ設定された値であるとき、エントロピー復号が第1の特徴要素に対して実行される必要がある。第1の特徴要素が判定マップにおいて位置する位置に対応する値があらかじめ設定された値ではないとき、エントロピー復号が第1の特徴要素に対して実行される必要はなく、第1の特徴要素の特徴値はkに設定され、kは整数である。

【0106】

前述の技術的解決策では、エントロピー復号が実行される必要があるかどうかは、各々の符号化されることになる特徴要素について決定され、それによって、いくつかの特徴要素のエントロピー復号プロセスがスキップされ、エントロピー復号が実行される必要がある要素の数量が著しく減らされることが可能である。このようにして、エントロピー復号の複雑さが減らされることが可能である。

【0107】

既存の主流のエンドツーエンドの特徴データの符号化および復号の解決策では、エントロピー符号化および復号または算術符号化および復号のプロセスは過度に複雑である。この出願では、符号化されることになる特徴データにおける特徴点の確率分布に関する情報は、符号化されることになる特徴データの各々の特徴要素に対してエントロピー符号化が実行される必要があるかどうか、および、復号されることになる特徴データの各々の特徴要素に対してエントロピー復号が実行される必要があるかどうかを決定するために使用され、それによって、いくつかの特徴要素のエントロピー符号化および復号のプロセスはスキップされ、符号化および復号が実行される必要がある要素の数量が著しく減らされることが可能である。これは、符号化および復号の複雑さを減らす。別の態様では、閾値は、生成されたビットストリームのビットレートの値を制御するために、ビットストリームのビットレートの実際の値の要件に基づいて柔軟に設定され得る。

【0108】

1つまたは複数の実施形態の詳細は、添付の図面および以下の説明において詳細に説明される。他の特徴、目的、および利点が、説明、図面、および請求項から明らかである。

【0109】

以下は、この出願の実施形態において使用される添付の図面を説明する。

【図面の簡単な説明】

【0110】

【図1A】ピクチャ復号システムの例示的なブロック図である。

【図1B】ピクチャ復号システムの処理回路の実装である。

10

20

30

40

50

- 【図 1 C】ピクチャ復号デバイスの概略ブロック図である。
- 【図 1 D】この出願の実施形態による装置の実装図である。
- 【図 2 A】この出願による可能なシナリオにおけるシステムアーキテクチャ図である。
- 【図 2 B】この出願による可能なシナリオにおけるシステムアーキテクチャ図である。
- 【図 3 A】エンコーダの概略ブロック図である。
- 【図 3 B】エンコーダの概略ブロック図である。
- 【図 3 C】エンコーダの概略ブロック図である。
- 【図 3 D】エンコーダの概略ブロック図である。
- 【図 4 A】エンコーダネットワークユニットの概略図である。
- 【図 4 B】エンコーダネットワークのネットワーク構造の概略図である。 10
- 【図 5】符号化判定実装ユニットの構造の概略図である。
- 【図 6】ジョイントネットワークの例示的な出力図である。
- 【図 7】生成ネットワークの例示的な出力図である。
- 【図 8】復号判定実装の概略実装図である。
- 【図 9】デコーダネットワークのネットワーク構造の例示的な図である。
- 【図 10 A】この出願の実施形態によるコーディング方法の例示的な図である。
- 【図 10 B】この出願の実施形態によるピクチャ特徴マップデコーダの概略ブロック図である。
- 【図 11 A】この出願の実施形態によるコーディング方法の例示的な図である。
- 【図 12】サイド情報抽出モジュールのネットワーク構造の例示的な図である。 20
- 【図 13 A】この出願の実施形態によるコーディング方法の例示的な図である。
- 【図 13 B】この出願の実施形態によるピクチャ特徴マップデコーダの概略ブロック図である。
- 【図 14】この出願の実施形態によるコーディング方法の例示的な図である。
- 【図 15】ジョイントネットワークのネットワーク構造の例示的な図である。
- 【図 16】この出願の実施形態によるピクチャ特徴マップデコーダの概略ブロック図である。
- 【図 17】この出願の実施形態によるコーディング方法の例示的な図である。
- 【図 18】この出願による符号化装置の例示的な構造の概略図である。
- 【図 19】この出願による復号装置の例示的な構造の概略図である。 30
- 【発明を実施するための形態】
- 【0111】

この出願の実施形態における「第1の」および「第2の」などの用語は、区別と説明のために使用されるのみであるが、相対的な重要性の指示または暗示、または順序の指示または暗示として理解されることはできない。加えて、用語「含む(include)」、「含む(comprise)」、およびそれら任意の変形は、非排他的な包含、たとえば、一連のステップまたはユニットの包含をカバーすることを意図される。方法、システム、製品、またはデバイスは、明確に列挙されたステップまたはユニットに必ずしも限定されないが、明確に列挙されない、かつ、プロセス、方法、製品、またはデバイスに固有の他のステップまたはユニットを含んでもよい。 40

【0112】

この出願では、「少なくとも1つの(項目)」は1つまたは複数を指し、「複数の」は2つ以上を指すことが理解されるべきである。用語「および/または」は、関連付けられた対象の関連付け関係を説明し、3つの関係が存在し得ることを示す。たとえば、「Aおよび/またはB」は、以下の3つの場合、すなわち、Aのみが存在する、Bのみが存在する、およびAとBの両方が存在する、を示してもよく、AとBは単数または複数であってもよい。文字「/」は、一般に、関連付けられた対象間の「または」の関係を示す。「以下の項目(要素)の少なくとも1つ」またはその類似する表現は、単一の項目(要素)または複数の項目(要素)の任意の組み合わせを含む、これらの項目の任意の組み合わせを示す。たとえば、a、b、またはcの少なくとも1つ(1個)は、a、b、c、「aおよびb」、「aおよびc」、「bおよ 50

びc」、または「a、b、およびc」を表現してもよく、a、b、およびcは、単数または複数であってもよい。

【0113】

この出願の実施形態は、AIに基づく特徴データ符号化および復号技術を提供し、特に、ニューラルネットワークに基づくピクチャ特徴マップおよび/またはオーディオ特徴変数符号化および復号技術を提供し、具体的には、エンドツーエンドに基づくピクチャ特徴マップおよび/またはオーディオ特徴変数符号化および復号システムを提供する。

【0114】

ピクチャコーディングの分野では、用語「ピクチャ(picture)」および「画像(image)」は、同義語として使用され得る。ピクチャコーディング(または一般にコーディングと呼ばれる)は、2つの部分、すなわち、ピクチャ符号化およびピクチャ復号を含む。ビデオは、複数のピクチャを含み、連続するピクチャの表現方式である。ピクチャ符号化は、ソース側において決定され、通常、ビデオピクチャを表現するために要求されるデータの量を減らすために(より効率的な記憶および/または送信のために)元のビデオピクチャを処理すること(たとえば、圧縮すること)を含む。ピクチャ復号は、デスティネーション側で決定され、通常、ピクチャを再構築するために、エンコーダのプロセスと比較して逆のプロセスを含む。ピクチャまたはオーディオの「コーディング」に言及する実施形態は、ピクチャまたはオーディオの「符号化」または「復号」として理解されるものである。符号化部分と復号部分の組み合わせは、符号化および復号(encoding and decoding, CODEC)とも呼ばれる。

【0115】

可逆ピクチャコーディングの場合、元のピクチャが再構築されることが可能である。言い換えると、再構築されたピクチャは、元のピクチャと同じ品質を有する(記憶または送信の間に送信損失または他のデータ損失が発生しないと仮定される)。従来の不可逆ピクチャコーディングの場合、ビデオピクチャを表現するために要求されるデータの量を減らすために、たとえば量子化を通じて、さらなる圧縮が決定され、ビデオピクチャはデコーダ側で完全に再構築されることが可能でなく、言い換えると、再構築されたビデオピクチャの品質は、元のビデオピクチャの品質と比較してより低く、またはより悪い。

【0116】

この出願の実施形態は、ニューラルネットワークの大規模アプリケーションに関するもので、理解の容易さのために、以下は、この出願の実施形態において使用され得るニューラルネットワークに関する用語および概念を説明する。

【0117】

(1)ニューラルネットワーク

【0118】

ニューラルネットワークはニューロンを含み得る。ニューロンは、入力として x_s および1の切片を使用する演算単位であり得る。演算単位の出力は以下の通りであり得る。

【数1】

$$h_{w,b}(x) = f(W^T x) = f\left(\sum_{s=1}^n W_s x_s + b\right) \quad (1-1)$$

【0119】

$s=1, 2, \dots$, または n であり、 n は1より大きい自然数であり、 W_s は x_s の重みであり、 b はニューロンのバイアスである。 f は、ニューロンの中の入力信号を出力信号にコンバートするために、非線形特徴をニューラルネットワークに導入するために使用される、ニューロンの活性化関数(activation function)である。活性化関数の出力信号は、次の畳み込み層の入力として使用されてもよく、活性化関数はシグモイド関数であってもよい。ニューラルネットワークは、複数の単一のニューロンを一緒に接続することによって構成されるネットワークである。具体的には、ニューロンの出力は別のニューロンの入力であり得る。各ニューロンの入力、局所受容野の特徴を抽出するために、前の層の局所受容野に

接続され得る。局所受容野はいくつかのニューロンを含む領域であり得る。

【 0 1 2 0 】

(2)深層ニューラルネットワーク

【 0 1 2 1 】

深層ニューラルネットワーク(deep neural network, DNN)は、多層ニューラルネットワークとも呼ばれ、複数の隠れ層を有するニューラルネットワークであると理解され得る。DNNは、異なる層の位置に基づいて分割される。DNNの内部のニューラルネットワークは、3つのタイプ、すなわち、入力層、隠れ層、および出力層に分類され得る。一般に、最初の層は入力層であり、最後の層は出力層であり、中間の層は隠れ層である。層は全結合される。具体的には、第*i*の層におけるあらゆるニューロンが、第(*i*+1)の層におけるあらゆるニューロンに必ず接続される。

10

【 0 1 2 2 】

DNNは複雑に見えるが、各層における作業に関しては複雑ではない。簡単に言うと、DNNは以下の線形関係式であり、すなわち、

【数 2】

$$\vec{y} = \alpha(W \vec{x} + b)$$

、ここで

【数 3】

$$\vec{x}$$

20

は入力ベクトルであり、

【数 4】

$$\vec{y}$$

は出力ベクトルであり、

【数 5】

$$\vec{b}$$

30

はオフセットベクトルであり、*W*は重み行列(係数とも呼ばれる)であり、 $\alpha(\)$ は活性化関数である。各層において、そのような簡単な演算のみが入力ベクトル

【数 6】

$$\vec{x}$$

に対して実行されて、出力ベクトル

40

【数 7】

$$\vec{y}$$

を取得する。DNNにおいて複数の層があるので、複数の係数*W*および複数のオフセットベクトル

【数 8】

$$\vec{b}$$

50

もある。DNNにおけるパラメータの定義は以下の通りである。係数 W は例として使用される。3つの層を有するDNNにおいて、第2の層における第4のニューロンから第3の層における第2のニューロンへの線形係数は、

【数 9】

$$W_{24}^3$$

として定義されることが仮定される。上付き文字3は、係数 W が位置する層を表現し、下付き文字は、出力の第3の層のインデックス2および入力第2の層のインデックス4に対応する。

【0 1 2 3】

結論として、第 $(L-1)$ の層における第 k のニューロンから第 L の層における第 j のニューロンへの係数は、

【数 10】

$$W_{jk}^L$$

として定義される。

【0 1 2 4】

入力層はパラメータ W を有しないことが留意されるべきである。深層ニューラルネットワークでは、より多くの隠れ層は、ネットワークが、現実世界における複雑な事例を記述することをより可能にする。理論的には、より多くのパラメータを有するモデルは、より高い複雑さ、およびより大きい「能力」を有する。それは、モデルがより複雑な学習タスクを完了できることを示す。深層ニューラルネットワークを訓練するプロセスは、重み行列を学習するプロセスであり、訓練の最終的な目的は、訓練される深層ニューラルネットワークの中のすべての層の重み行列(複数の層におけるベクトル W を含む重み行列)を取得することである。

【0 1 2 5】

(3) 畳み込みニューラルネットワーク

【0 1 2 6】

畳み込みニューラルネットワーク(convolutional neuron network, CNN)は、畳み込み構造を有する深層ニューラルネットワークである。畳み込みニューラルネットワークは、畳み込み層およびサブサンプリング層を含む特徴抽出器を含み、特徴抽出器はフィルタであると思われ得る。畳み込み層は、畳み込みニューラルネットワークの中にあり、かつ畳み込み処理が入力信号に対してそこで実行される、ニューロン層である。畳み込みニューラルネットワークの畳み込み層において、1つのニューロンが、隣接する層のニューロンの一部のみに接続され得る。畳み込み層は、通常は、いくつかの特徴平面を含み、各特徴平面は、長方形に並べられたいくつかのニューロンを含み得る。同じ特徴平面のニューロンは重みを共有し、ここでの共有された重みは畳み込みカーネルである。ピクチャ情報抽出方式が位置に無関係であるものとして、重み共有は理解され得る。畳み込みカーネルは、ランダムサイズ行列の形式で初期化され得る。畳み込みニューラルネットワークを訓練するプロセスにおいて、畳み込みカーネルは、学習を通じて適切な重みを取得し得る。加えて、重み共有によりもたらされる直接の利益は、畳み込みニューラルネットワークの中の層の間の接続が減らされ、過学習の危険性が下げられることである。

【0 1 2 7】

(4) エントロピー符号化

【0 1 2 8】

デコード、または同様のものが復号のためのパラメータを受信して使用し得るように、たとえば、エントロピーコーディングアルゴリズムまたは方式(たとえば、可変長コーディング(variable length coding, VLC)方式、コンテキスト適応VLC(context adaptive V

10

20

30

40

50

LC, CAVLC)方式、算術コーディング方式、二値化アルゴリズム、コンテキスト適応バイナリ算術コーディング(context adaptive binary arithmetic coding, CABAC)、シンタックスベースコンテキスト適応バイナリ算術コーディング(syntax-based context-adaptive binary arithmetic coding, SBAC)、確率間隔区分エントロピー(probability interval partitioning entropy, PIPE)コーディング、または別のエントロピーコーディング方法または技術)を、量子化された係数および他のシンタックス要素に適用して、符号化されたビットストリームの形式で出力を通じて出力され得る符号化されたデータを取得するために、エントロピー符号化が使用される。符号化されたビットストリームは、デコーダに送信され、または、デコーダによる後続の送信または取り出しのためにメモリに記憶され得る。

10

【0129】

コーディングシステム10の以下の実施形態では、エンコーダ20Aおよびデコーダ30Aは、図1Aから図15に基づいて説明される。

【0130】

図1Aは、例示的なコーディングシステム10、たとえば、この出願の技法を利用し得るピクチャ(またはオーディオ)コーディングシステム10(または略してコーディングシステム10)を例示する概略ブロック図である。ピクチャコーディングシステム10の中のエンコーダ20Aおよびデコーダ30Aは、この出願において説明される様々な例に基づいて様々な技法を決定するように構成され得るデバイスおよび同様のものを表現する。

【0131】

20

図1Aに表されるように、コーディングシステム10は、符号化されたビットストリーム21、たとえば符号化されたピクチャ(またはオーディオ)を、符号化されたビットストリーム21を復号するように構成されるデスティネーションデバイス14のために提供するように構成されるソースデバイス12を含む。

【0132】

ソースデバイス12は、エンコーダ20Aを含み、任意選択で、ピクチャソース16、プリプロセッサ(または前処理ユニット)18、通信インターフェース(または通信ユニット)26、および確率推定(または確率推定ユニット)40を含む。

【0133】

ピクチャ(またはオーディオ)ソース16は、現実世界のピクチャ(またはオーディオ)を取り込むように構成される任意のタイプのピクチャ取り込みデバイス、および/または、任意のタイプのピクチャ生成デバイス、たとえば、コンピュータアニメーション化されたピクチャを生成するように構成されるコンピュータグラフィックス処理ユニット、または、現実世界のピクチャ、コンピュータ生成されたピクチャ(たとえば、スクリーンコンテンツ、仮想現実(virtual reality, VR)ピクチャ)、および/またはそれらの任意の組み合わせ(たとえば、拡張現実(augmented reality, AR)ピクチャ)を取得および/または提供するように構成される任意のタイプのデバイスを含んでもよく、またはそれらであってもよい。オーディオまたはピクチャソースは、任意の前述のオーディオまたはピクチャを記憶する任意のタイプのメモリまたは記憶装置であり得る。

30

【0134】

40

プリプロセッサ(または前処理ユニット)18およびプリプロセッサ(または前処理ユニット)18によって決定される処理と区別して、ピクチャまたはオーディオ(ピクチャまたはオーディオデータ)17は、元のピクチャまたはオーディオ(元のピクチャまたはオーディオデータ)17とも呼ばれ得る。

【0135】

プリプロセッサ18は、(元の)ピクチャ(またはオーディオ)データ17を受信し、ピクチャ(またはオーディオ)データ17に対して前処理を実行して、前処理されたピクチャまたはオーディオ(または前処理されたピクチャまたはオーディオデータ)19を取得するように構成される。たとえば、プリプロセッサ18によって決定される前処理は、トリミング、カラーフォーマット変換(たとえば、RGBからYCbCrへ)、色補正、またはノイズ除去を含み得る

50

。前処理ユニット18は任意選択のコンポーネントであり得ることが理解され得る。

【0136】

エンコーダ20Aは、エンコーダネットワーク20、エントロピー符号化24、および任意選択でプリプロセッサ22を含む。

【0137】

ピクチャ(またはオーディオ)エンコーダネットワーク(またはエンコーダネットワーク)20は、前処理されたピクチャ(またはオーディオ)データ19を受信し、符号化されたピクチャ(またはオーディオ)データ21を提供するように構成される。

【0138】

プリプロセッサ22は、符号化されることになる特徴データ21を受信し、符号化されることになる特徴データ21を前処理して前処理された符号化されることになる特徴データ23を取得するように構成される。たとえば、プリプロセッサ22によって決定される前処理は、トリミング、カラーフォーマット変換(たとえば、RGBからYCbCrへ)、色補正、またはノイズ除去を含み得る。前処理ユニット22は任意選択のコンポーネントであり得ることが理解され得る。

10

【0139】

エントロピー符号化24は、符号化されることになる特徴データを受信し(または符号化されることになる特徴データを前処理し)23、確率推定40によって提供される確率推定結果41に基づいて、符号化されたビットストリーム25を生成するために使用される。

【0140】

ソースデバイス12の通信インターフェース26は、符号化されたビットストリーム25を受信し、通信チャネル27上で、符号化されたビットストリーム25(またはその任意のさらなる処理されたバージョン)を、記憶または直接の再構築のために、デスティネーションデバイス14または任意の他のデバイスなどの別のデバイスに送信するように構成され得る。

20

【0141】

デスティネーションデバイス14は、デコーダ30Aを含み、任意選択で、通信インターフェース(または通信ユニット)28、ポストプロセッサ(または後処理ユニット)36、および表示デバイス38を含み得る。

【0142】

デスティネーションデバイス14の通信インターフェース28は、符号化されたビットストリーム25(またはその任意のさらなる処理されたバージョン)を、ソースデバイス12から直接に、または記憶デバイス、たとえば、符号化されたビットストリームデータの記憶デバイスなどの任意の他のソースデバイスから受信し、符号化されたビットストリーム25をデコーダ30Aのために提供するように構成される。

30

【0143】

通信インターフェース26および通信インターフェース28は、ソースデバイス12とデスティネーションデバイス14との間の直接の通信リンク、たとえば、直接の有線または無線接続を通じて、あるいは、任意のタイプのネットワーク、たとえば、有線または無線ネットワークまたはそれらの任意の組み合わせ、または任意のタイプのプライベートおよびパブリックネットワーク、またはそれらの任意のタイプの組み合わせを介して、符号化されたビットストリーム(または符号化されたビットストリームデータ)25を送信または受信するように構成され得る。

40

【0144】

通信インターフェース26は、たとえば、符号化されたビットストリーム25を適切なフォーマット、たとえばパケットにパッケージ化し、および/または、通信リンクまたは通信ネットワーク上での送信のために、任意のタイプの送信符号化または処理を使用することによって、符号化されたビットストリームを処理するように構成され得る。

【0145】

通信インターフェース28は、通信インターフェース26に対応し、たとえば、送信データを受信し、任意のタイプの対応する送信復号または処理および/またはカプセル化解除

50

を使用することによって、送信データを処理して、符号化されたビットストリーム25を取得するように構成され得る。

【0146】

通信インターフェース26と通信インターフェース28の両方が、ソースデバイス12からデスティネーションデバイス14に向かう図1Aの通信チャネル27のための矢印によって示されるような一方向通信インターフェース、または双方向通信インターフェースとして構成されてもよく、メッセージを送信および受信し、接続を設定し、通信リンクおよび/またはデータ送信、たとえば符号化されたピクチャデータ送信に関する任意の他の情報に肯定応答し、それを交換するように構成されてもよい。

【0147】

デコーダ30Aは、デコーダネットワーク34、エントロピー復号30、および任意選択でポストプロセッサ32を含む。

【0148】

エントロピー復号30は、符号化されたビットストリーム25を受信し、確率推定40によって提供される確率推定結果42に基づいて、復号された特徴データ31を提供するために使用される。

【0149】

ポストプロセッサ32は、復号された特徴データ31に対して後処理を実行して、後処理された復号された特徴データ33を取得するように構成される。後処理ユニット32によって決定される後処理は、たとえば、カラーフォーマット変換(たとえば、YCbCrからRGBへ)、色補正、トリミング、または再サンプリングを含み得る。後処理ユニット32は任意選択のコンポーネントであり得ることが理解され得る。

【0150】

デコーダネットワーク34は、復号された特徴データ31または後処理された復号された特徴データ33を受信し、再構築されたピクチャデータ35を提供するために使用される。

【0151】

ポストプロセッサ36は、再構築されたピクチャデータ35に対して後処理を実行して、後処理された再構築されたピクチャデータ37を取得するように構成される。後処理ユニット36によって決定される後処理は、たとえば、カラーフォーマット変換(たとえば、YCbCrからRGBへ)、色補正、トリミング、または再サンプリングを含み得る。後処理ユニット36は任意選択のコンポーネントであり得ることが理解され得る。

【0152】

表示デバイス38は、ピクチャを、ユーザ、閲覧者、または同様のものに表示するために、再構築されたピクチャデータ35または後処理されたピクチャデータ37を受信するように構成される。表示デバイス38は、再構築されたオーディオまたはピクチャを表現するための任意のタイプのプレーヤまたはディスプレイ、たとえば統合されたまたは外部のディスプレイまたはモニタであってもよく、またはそれを含んでもよい。たとえば、ディスプレイは、液晶ディスプレイ(liquid crystal display, LCD)、有機発光ダイオード(organic light emitting diode, OLED)ディスプレイ、プラズマディスプレイ、プロジェクタ、マイクロLEDディスプレイ、液晶オンシリコン(liquid crystal on silicon, LCoS)、デジタル光プロセッサ(digital light processor, DLP)、または任意のタイプの別のディスプレイスクリーンを含み得る。

【0153】

図1Aは、ソースデバイス12およびデスティネーションデバイス14を独立したデバイスとして表すが、デバイスの実施形態は、代替として、ソースデバイス12とデスティネーションデバイス14の両方を含んでもよく、または、ソースデバイス12とデスティネーションデバイス14の両方の機能を含んでもよく、すなわち、ソースデバイス12または対応する機能およびデスティネーションデバイス14または対応する機能の両方を含む。これらの実施形態では、ソースデバイス12または対応する機能およびデスティネーションデバイス14または対応する機能は、同じハードウェアおよび/またはソフトウェアを使用すること

10

20

30

40

50

によって、または別個のハードウェアおよび／またはソフトウェアまたはそれらの任意の組み合わせを使用することによって実装され得る。

【0154】

説明に基づいて、図1Aに表されるソースデバイス12および／またはデスティネーションデバイス14の異なるユニットまたは機能の存在および(正確な)分割は、実際のデバイスおよび適用とともに変化してもよい。これはこの技術分野の当業者に明らかである。

【0155】

特徴データエンコーダ20A(たとえば、ピクチャ特徴マップエンコーダまたはオーディオ特徴変数エンコーダ)、特徴データデコーダ30A(たとえば、ピクチャ特徴マップデコーダまたはオーディオ特徴変数デコーダ)、または特徴データエンコーダ20Aと特徴データデコーダ30Aの両方が、図1Bに表される処理回路、たとえば、1つまた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 processor, DSP)、特定用途向け集積回路(application-specific integrated circuit, ASIC)、フィールドプログラマブルゲートアレイ(field programmable gate array, FPGA)、個別のロジック、ハードウェア、ピクチャ符号化のための専用プロセッサ、またはそれらの任意の組み合わせを使用することによって実装され得る。特徴データエンコーダ20Aは、処理回路56を使用することによって実装されてもよく、特徴データデコーダ30Aは、処理回路56を使用することによって実装されてもよい。処理回路56は、以下の様々な動作を決定するように構成され得る。技法がソフトウェアで部分的に実装されるならば、デバイスは、適切な非一時的コンピュータ可読記憶媒体にソフトウェアのための命令を記憶してもよく、本発明の技法を決定するために1つまたは複数のプロセッサを使用することによってハードウェアにおいて命令を決定してもよい。図1Bに表されるように、特徴データエンコーダ20Aと特徴データデコーダ30Aの1つが、組み合わされたエンコーダ／デコーダ(encoder/decoder, CODEC)の一部として単一のデバイスに統合され得る。

【0156】

ソースデバイス12およびデスティネーションデバイス14は、任意のタイプのハンドヘルドデバイスまたは固定デバイス、たとえば、ノートブックまたはラップトップコンピュータ、携帯電話、スマートフォン、タブレットまたはタブレットコンピュータ、カメラ、デスクトップコンピュータ、セットトップボックス、テレビ、表示デバイス、デジタルメディアプレーヤ、ビデオゲームコンソール、ビデオストリームデバイス(たとえば、コンテンツサービスサーバまたはコンテンツ配信サーバ)、放送受信デバイス、放送送信デバイス、および同様のものを含む、様々なデバイスのいずれか1つを含んでもよく、任意のタイプのオペレーティングシステムを使用しなくてもよく、または使用してもよい。いくつかの場合、ソースデバイス12およびデスティネーションデバイス14は、無線通信のためのコンポーネントを装備され得る。したがって、ソースデバイス12およびデスティネーションデバイス14は、無線通信デバイスであり得る。

【0157】

いくつかの場合、図1Aに表されるコーディングシステム10は単に例である。この出願で提供される技術は、ピクチャ特徴マップまたはオーディオ特徴変数コーディング設定(たとえば、ピクチャ特徴マップ符号化またはピクチャ特徴マップ復号)に適用可能であってもよく、設定は、符号化デバイスと復号デバイスとの間の任意のデータ通信を必ずしも含まない。別の例では、データは、ローカルメモリから取り出され、ネットワーク上で送信され、および同様である。ピクチャ特徴マップまたはオーディオ特徴変数符号化デバイスは、データを符号化し、データをメモリに記憶してもよく、および／または、ピクチャ特徴マップまたはオーディオ特徴変数復号デバイスは、メモリからデータを取り出し、データを復号してもよい。いくつかの例では、符号化および復号は、互いに通信しないがメモリにデータを符号化し、および／またはメモリからデータを取り出し、データを復号するデバイスによって決定される。

【0158】

図1Bは、例示的な実施形態による、図1Aの特徴データエンコーダ20Aおよび／または

図1Bの特徴データデコーダ30Aを含むコーディングシステム50の例の図である。コーディングシステム50は、撮像(またはオーディオ生成)デバイス51、エンコーダ20Aおよびデコーダ30A(および/または処理回路56を使用することによって実装される特徴データエンコーダ/デコーダ)、アンテナ52、1つまたは複数のプロセッサ53、1つまたは複数のメモリ記憶54、および/または表示(またはオーディオ再生)デバイス55を含み得る。

【0159】

図1Bに表されるように、撮像(またはオーディオ生成)デバイス51、アンテナ52、処理回路56、エンコーダ20A、デコーダ30A、プロセッサ53、メモリ記憶54、および/または表示(またはオーディオ再生)デバイス55は、互いに通信することができる。異なる例では、コーディングシステム50は、エンコーダ20Aのみまたはデコーダ30Aのみを含み得る。

10

【0160】

いくつかの例では、アンテナ52は、特徴データの符号化されたビットストリームを送信または受信するように構成され得る。加えて、いくつかの例では、表示(またはオーディオ再生)デバイス55は、ピクチャ(またはオーディオ)データを提示するように構成され得る。処理回路56は、特定用途向け集積回路(application-specific integrated circuit, ASIC)ロジック、グラフィックス処理ユニット、汎用プロセッサ、および同様のものを含み得る。コーディングシステム50は、任意選択のプロセッサ53も含み得る。同様に、任意選択のプロセッサ53は、特定用途向け集積回路(application-specific integrated circuit, ASIC)ロジック、グラフィックス処理ユニット、オーディオプロセッサ、汎用プロセッサ、および同様のものを含み得る。加えて、メモリ記憶54は、任意のタイプのメモリ、たとえば、揮発性メモリ(たとえば、スタティックランダムアクセスメモリ(static random access memory, SRAM)、またはダイナミックランダムアクセスメモリ(dynamic random access memory, DRAM))、または不揮発性メモリ(たとえば、フラッシュメモリ)であり得る。限定しない例において、メモリ記憶54はキャッシュメモリを使用することによって実装され得る。別の例において、処理回路56は、ピクチャバッファを実装するように構成されるメモリ(たとえば、キャッシュ)を含み得る。

20

【0161】

いくつかの例では、論理回路を使用することによって実装されるエンコーダ20Aは、ピクチャバッファ(たとえば、処理回路56またはメモリ記憶54を使用することによって実装される)およびグラフィックス処理ユニット(たとえば、処理回路56を使用することによって実装される)を含み得る。グラフィックス処理ユニットは、ピクチャバッファに通信可能に結合され得る。グラフィックス処理ユニットは、処理回路56を使用することによって実装されるエンコーダ20Aを含み得る。論理回路は、明細書における様々な動作を決定するように構成され得る。

30

【0162】

いくつかの例では、デコーダ30Aは、図1Bに表されるデコーダ30および/または明細書において説明される任意の他のデコーダシステムまたはサブシステムを参照して説明される様々なモジュールを実装するために、類似する方式で処理回路56を使用することによって実装され得る。いくつかの例では、論理回路を使用することによって実装されるデコーダ30Aは、ピクチャバッファ(たとえば、処理回路56またはメモリ記憶54を使用することによって実装される)およびグラフィックス処理ユニット(たとえば、処理回路56を使用することによって実装される)を含み得る。グラフィックス処理ユニットは、ピクチャバッファに通信可能に結合され得る。グラフィックス処理ユニットは、処理回路56を使用することによって実装されるピクチャデコーダ30Aを含み得る。

40

【0163】

いくつかの例では、アンテナ52は、ピクチャデータの符号化されたビットストリームを受信するように構成され得る。上記で説明されたように、符号化されたビットストリームは、オーディオまたはビデオフレーム符号化に関する、明細書において説明された、データ、インジケータ、インデックス値、モード選択データ、および同様のもの、たとえば、

50

符号化区分に関するデータを含み得る。コーディングシステム50は、アンテナ52に結合され、符号化されたビットストリームを復号するように構成されるデコーダ30Aも含み得る。表示(またはオーディオ再生)デバイス55は、ピクチャ(またはオーディオ)を提示するように構成され得る。

【0164】

この出願のこの実施形態では、エンコーダ20Aを参照して説明された例について、デコーダ30Aは逆の処理を決定するように構成され得ることが理解されるべきである。シグナリングシンタックス要素について、デコーダ30Aは、シンタックス要素を受信し、解析し、それに対応して、関連するピクチャデータを復号するように構成され得る。いくつかの例では、エンコーダ20Aは、符号化されたビットストリームを取得するために、シンタックス要素に対してエントロピー符号化を実行し得る。例では、デコーダ30Aは、シンタックス要素を解析し、それに対応して、関連するピクチャデータを復号し得る。

10

【0165】

図1Cは、本発明の実施形態によるコーディングデバイス400の概略図である。コーディングデバイス400は、明細書において説明される開示された実施形態を実装するために適している。実施形態において、コーディングデバイス400は、デコーダ、たとえば図1Aのピクチャ特徴マップデコーダ30Aであってもよく、または、エンコーダ、たとえば、図1Aのピクチャ特徴マップエンコーダ20Aであってもよい。

【0166】

ピクチャコーディングデバイス400は、データを受信するように構成される入口ポート410(または入力ポート410)および受信機ユニット(receiver unit, Rx)420と、データを処理するように構成されるプロセッサ、論理ユニット、または中央処理ユニット(central processing unit, CPU)430であって、たとえば、プロセッサ430は、ニューラルネットワーク処理ユニット430であってもよい、プロセッサ、論理ユニット、または中央処理ユニット430と、データを送信するように構成される送信機ユニット(transmitter unit, Tx)440および出口ポート450(または出力ポート450)と、データを記憶するように構成されるメモリ460とを含む。ピクチャ(またはオーディオ)コーディングデバイス400は、入口ポート410に結合される光-電気(optical-to-electrical, OE)コンポーネントおよび電気-光(electrical-to-optical, EO)コンポーネント、受信機ユニット420、送信機ユニット440、および光または電気信号の出口または入口のための出口ポート450をさらに含む。

20

30

【0167】

プロセッサ430は、ハードウェアおよびソフトウェアによって実装される。プロセッサ430は、1つまたは複数のプロセッサチップ、コア(たとえば、マルチコアプロセッサ)、FPGA、ASIC、およびDSPとして実装され得る。プロセッサ430は、入口ポート410、受信機ユニット420、送信機ユニット440、出口ポート450、およびメモリ460と通信する。プロセッサ430は、コーディングモジュール470(たとえば、ニューラルネットワークNNに基づくコーディングモジュール470)を含む。コーディングモジュール470は、上記で説明された開示された実施形態を実装する。たとえば、コーディングモジュール470は、様々なコーディング動作を決定し、処理し、準備し、または提供する。したがって、コーディングモジュール470の包含は、コーディングデバイス400の機能をかなり改善し、異なる状態へのコーディングデバイス400の切り替えに影響する。代替として、コーディングモジュール470は、メモリ460に記憶され、プロセッサ430によって決定される命令を使用することによって実装される。

40

【0168】

メモリ460は、1つまたは複数のディスク、テープドライブ、およびソリッドステートドライブを含み、そのようなプログラムが決定のために選択されるときにプログラムを記憶し、プログラム決定の間に読み取られる命令とデータを記憶するために、オーバーフローデータ記憶デバイスとして使用され得る。メモリ460は、揮発性および/または不揮発性であってもよく、読み取り専用メモリ(read-only memory, ROM)、ランダムアクセス

50

メモリ(random access memory, RAM)、三値連想メモリ(ternary content-addressable memory, TCAM)、および/またはスタティックランダムアクセスメモリ(static random access memory, SRAM)であってもよい。

【0169】

図1Dは、実施形態による、図1Aのソースデバイス12とデスティネーションデバイス14のいずれかまたは両方として使用され得る装置500の簡略化されたブロック図である。

【0170】

装置500におけるプロセッサ502は中央処理ユニットであり得る。代替として、プロセッサ502は、情報を操作または処理することができる、既存のまたは今後開発される、任意の他のタイプのデバイスまたは複数のデバイスであり得る。開示される実装は、図に表されるプロセッサ502などの単一のプロセッサによって実装され得るが、1つより多くのプロセッサを使用することによって、速度と効率における利点が達成されることが可能である。

10

【0171】

実装では、装置500におけるメモリ504は、読み取り専用メモリ(ROM)デバイスまたはランダムアクセスメモリ(RAM)デバイスであることが可能である。任意の他の適切なタイプの記憶デバイスは、メモリ504として使用され得る。メモリ504は、バス512を通じてプロセッサ502によってアクセスされるコードおよびデータ506を含み得る。メモリ504は、オペレーティングシステム508およびアプリケーションプログラム510をさらに含んでもよく、アプリケーションプログラム510は、プロセッサ502が明細書における方法を決定することを可能にする少なくとも1つのプログラムを含む。たとえば、アプリケーションプログラム510は、アプリケーション1からNを含み、明細書における方法を決定するピクチャコーディングアプリケーションをさらに含み得る。

20

【0172】

装置500は、ディスプレイ518などの1つまたは複数の出力デバイスをさらに含み得る。例では、ディスプレイ518は、タッチ入力感知するように構成され得るタッチ感知要素とディスプレイを組み合わせるタッチ感知ディスプレイであり得る。ディスプレイ518は、バス512を通じてプロセッサ502に結合され得る。

【0173】

装置500におけるバス512は明細書では単一のバスとして説明されるが、バス512は複数のバスを含み得る。さらに、二次メモリが、装置500の別のコンポーネントに直接に結合されてもよく、またはネットワークを介してアクセスされてもよく、メモリカードなどの単一の統合されたユニットまたは複数のメモリカードなどの複数のユニットを含んでもよい。したがって、装置500は様々な構成を有し得る。

30

【0174】

図2Aは、ピクチャ特徴マップまたはオーディオ特徴変数符号化および復号のシナリオにおける可能なシステムアーキテクチャ1800を表し、以下を含む。

取り込みデバイス1801: ビデオ取り込みデバイスが元のビデオ(またはオーディオ)の取り込みを完了する、

取り込み前処理1802: ビデオ(またはオーディオ)データを取得するために一連の前処理が元のビデオ(またはオーディオ)の取り込みに対して実行される、

40

符号化1803: ビデオ(オーディオ)符号化が、符号化の冗長性を減らし、ピクチャ特徴マップまたはオーディオ特徴変数圧縮プロセスにおけるデータ送信量を減らすために使用される、

送信1804: 符号化を通じて取得される圧縮され符号化されたビットストリームデータが送信モジュールを使用することによって送信される、

受信1805: 圧縮され符号化されたビットストリームデータがネットワーク送信を通じて受信モジュールによって受信される、

ビットストリーム復号1806: ビットストリーム復号がビットストリームデータに対して実行される、および

50

レンダリング表示(または再生)1807: レンダリング表示(または再生)が復号されたデータに対して実行される。

【0175】

図2Bは、ピクチャ特徴マップ(またはオーディオ特徴変数)の機械向けタスクのシナリオにおける可能なシステムアーキテクチャ1900を表し、以下を含む。

特徴抽出1901: 特徴抽出がピクチャ(またはオーディオ)ソースに対して実行される、

サイド情報抽出1902: サイド情報抽出が特徴抽出を通じて取得されたデータに対して実行される、

確率推定1903: サイド情報が確率推定の入力として使用され、確率推定結果を取得するために確率推定が特徴マップ(または特徴変数)に対して実行される、

10

符号化1904: エントロピー符号化が、ビットストリームを取得するために、確率推定結果を参照して特徴抽出を通じて取得されたデータに対して実行され、

任意選択で、符号化が実行される前に、量子化または丸め動作が特徴抽出を通じて取得されたデータに対して実行され、特徴抽出を通じて取得された、量子化され、または丸められたデータが符号化され、

任意選択で、ビットストリームがサイド情報データを含むように、エントロピー符号化がサイド情報に対して実行される、

復号1905: エントロピー復号が、ピクチャ特徴マップ(またはオーディオ特徴変数)を取得するために、確率推定結果を参照してビットストリームに対して実行され、

任意選択で、ビットストリームがサイド情報符号化データを含むならば、エントロピー復号が、サイド情報符号化データに対して実行され、復号されたサイド情報データが、確率推定結果を取得するために確率推定の入力として使用され、

20

サイド情報のみが確率推定の入力として使用されるとき、特徴要素の確率推定結果が並列に出力されてもよく、確率推定の入力がコンテキスト情報を含むとき、特徴要素の確率推定結果が直列に出力される必要があり、サイド情報が、ピクチャ特徴マップまたはオーディオ特徴変数をニューラルネットワークに入力することによってさらに抽出される特徴情報であり、サイド情報に含まれる特徴要素の数量が、ピクチャ特徴マップまたはオーディオ特徴変数の特徴要素の数量より少なく、任意選択で、ピクチャ特徴マップまたはオーディオ特徴変数のサイド情報が、ビットストリームに符号化されてもよいことが留意されるべきである、および

30

マシンビジョントask1906: マシンビジョン(または聴覚)タスクが復号された特徴マップ(または特徴変数)に対して実行される。

【0176】

具体的には、復号された特徴データは、マシンビジョン(または聴覚)タスクネットワークに入力され、ネットワークは、ビジョン(または聴覚)タスクに関する分類、目標認識、およびセマンティック区分などの1次元、2次元、または多次元データを出力する。

【0177】

可能な実装において、システムアーキテクチャ1900の実装プロセスにおいて、特徴抽出および符号化プロセスが端末において実施され、復号およびマシンビジョントaskがクラウドにおいて実施される。

40

【0178】

エンコーダ20Aは、入力202または同様のものを通じてピクチャ(またはピクチャデータ)またはオーディオ(またはオーディオデータ)17を受信するように構成され得る。受信されたピクチャ、ピクチャデータ、オーディオ、およびオーディオデータは、代替として、前処理されたピクチャ(または前処理されたピクチャデータ)またはオーディオ(または前処理されたオーディオデータ)19であり得る。簡潔さの容易さのために、以下の説明はピクチャ(またはオーディオ)17を使用する。ピクチャ(またはオーディオ)17は、代替として、現在のピクチャまたは符号化されることになるピクチャ(特に、現在のピクチャがビデオ符号化において他のピクチャから区別され、たとえば、他のピクチャは同じビデオシーケンスの中にあり、すなわち、現在のピクチャのビデオシーケンスの中の前の符号化されたピク

50

チャおよび/または復号されたピクチャを含むとき)、または現在のオーディオまたは符号化されることになるオーディオと呼ばれてもよい。

【0179】

(デジタル)ピクチャは、強度値を有するサンプルの2次元配列または行列であり、またはそのように見なされ得る。配列の中のサンプルは、ピクセル(pixelまたはpel)(ピクチャ要素の省略形)とも呼ばれ得る。配列またはピクチャの水平および垂直方向(または軸)におけるサンプルの数量は、ピクチャのサイズおよび/または解像度を定義する。色の表現については、通常は、三色成分が採用される。具体的には、ピクチャは、3つのサンプル配列として表現され、またはそれらを含んでもよい。RGBフォーマットまたは色空間において、ピクチャは、対応する赤、緑、および青のサンプル配列を含む。同様に、各ピクセルは、輝度または色度フォーマットまたは色空間、たとえばYCbCrで表現されてもよく、これは、Yによって示される輝度成分(ときどきLも代わりに使用される)およびCbとCrによって示される2つの色度成分を含む。輝度(luma)成分Yは、明るさまたはグレーレベル強度(たとえば、これら2つはグレースケールピクチャでは同じである)を表現するが、2つの色度(chrominance, 略してchroma)成分CbとCrは、色度または色情報成分を表現する。それに対応して、YCbCrフォーマットにおけるピクチャは、輝度サンプル値(Y)の輝度サンプル配列、および色度値の2つの色度サンプル配列(CbおよびCr)を含む。RGBフォーマットにおけるピクチャは、YCbCrフォーマットにコンバートされ、または変換されてもよく、その逆でもよく、この処理は色変換またはコンバージョンとしても知られている。ピクチャがモノクロであるならば、ピクチャは輝度サンプル配列のみを含み得る。それに対応して、ピクチャは、たとえば、モノクロフォーマットでの輝度サンプルの配列、または、4:2:0、4:2:2、および4:4:4カラーフォーマットでの輝度サンプルの配列および色度サンプルの2つの対応する配列であり得る。ピクチャエンコード20Aは、ピクチャの色空間を限定しない。

【0180】

可能性において、エンコード20Aの実施形態は、ピクチャ(またはオーディオ)17を複数の(通常は重複しない)ピクチャブロック203またはオーディオセグメントに区分するように構成されるピクチャ(またはオーディオ)区分ユニット(図1Aにも図1Bにも表されない)を含み得る。これらのピクチャブロックは、ルートブロック、マクロブロック(H.264/AVC)、またはH.265/HEVCおよびVVC規格ではコーディングツリーブロック(coding tree block, CTB)またはコーディングツリーユニット(coding tree unit, CTU)とも呼ばれ得る。区分ユニットは、ビデオシーケンスのすべてのピクチャに対する同じブロックサイズ、およびブロックサイズを定義する対応するグリッドを使用し、またはピクチャ、ピクチャサブセット、またはピクチャのグループ間でブロックサイズを変更し、各ピクチャを対応するブロックに区分するように構成され得る。

【0181】

別の可能性において、エンコードは、ピクチャ17のブロック203、たとえば、ピクチャ17を形成する1つ、いくつか、またはすべてのブロックを直接に受信するように構成され得る。ピクチャブロック203は、現在のピクチャブロックまたは符号化されることになるピクチャブロックとも呼ばれ得る。

【0182】

ピクチャ17のように、ピクチャブロック203も、ピクチャ17より小さい寸法であるが、強度値(サンプル値)を有するサンプルの2次元配列または行列であり、またはそのように見なされ得る。言い換えると、ブロック203は、たとえば、1つのサンプル配列(たとえば、モノクロピクチャ17の場合は輝度配列、またはカラーピクチャの場合は輝度または色度配列)、3つのサンプル配列(たとえば、カラーピクチャ17の場合は1つの輝度配列および2つの色度配列)、または適用されるカラーフォーマットに依存した任意の他の数量および/またはタイプの配列を含み得る。ブロック203の水平および垂直方向(または軸)におけるサンプルの数量は、ブロック203のサイズを定義する。それに対応して、ブロックは、たとえば、 $M \times N$ (M 列 $\times N$ 行)サンプルの配列、または $M \times N$ 変換係数の配列であり得る。

【0183】

別の可能性において、図1Aおよび図1Bまたは図3Aから図3Dに表されるエンコーダ20Aは、ブロックごとにピクチャ17を符号化するように構成される。

【0184】

別の可能性において、図1Aおよび図1Bまたは図3Aから図3Dに表されるエンコーダ20Aは、ピクチャ17を符号化するように構成される。

【0185】

別の可能性において、図1Aおよび図1Bまたは図3Aから図3Dに表されるエンコーダ20Aは、スライス(ビデオスライスとも呼ばれる)を使用することによってピクチャを区分または符号化するようにさらに構成されてもよく、ピクチャは1つまたは複数のスライス(通常は重複しない)を使用することによって区分され、または符号化されてもよい。各スライスは、1つまたは複数のブロック(たとえば、コーディングツリーユニットCTU)、またはブロックの1つまたは複数のグループ(たとえば、H.265/HEVC/VVC規格におけるタイル(tile)またはVVC規格におけるサブピクチャ(subpicture))を含み得る。

10

【0186】

別の可能性において、図1Aおよび図1Bまたは図3Aから図3Dに表されるエンコーダ20Aは、スライス/タイルグループ(ビデオタイルグループとも呼ばれる)および/またはタイル(ビデオタイルとも呼ばれる)を使用することによって、ピクチャを区分および/または符号化するようにさらに構成されてもよく、ピクチャは、1つまたは複数のスライス/タイルグループ(通常は重複しない)を使用することによって、区分され、または符号化されてもよく、各スライス/タイルグループは、1つまたは複数のブロック(たとえば、CTU)または1つまたは複数のタイルを含んでもよい。各タイルは、長方形の形状であってもよく、1つまたは複数の完全なまたは断片的なブロック(たとえば、CTU)を含んでもよい。

20

【0187】

エンコーダネットワーク20

【0188】

エンコーダネットワーク20は、入力データに基づいて、エンコーダネットワークを使用することによって、ピクチャ特徴マップまたはオーディオ特徴変数を取得するように構成される。

【0189】

可能性において、図4Aに表されるエンコーダネットワーク20は、複数のネットワーク層を含む。任意のネットワーク層が、畳み込み層、正規化層、非線形活性化層、または同様のものであってもよい。

30

【0190】

可能性において、エンコーダネットワーク20の入力は、少なくとも1つの符号化されることになるピクチャまたは少なくとも1つの符号化されることになるピクチャブロックである。符号化されることになるピクチャは、元のピクチャ、不可逆ピクチャ、または残差ピクチャであり得る。

【0191】

可能性において、エンコーダネットワーク20におけるエンコーダネットワークのネットワーク構造の例が、図4Bに表されている。この例では、エンコーダネットワークが5つのネットワーク層を含み、具体的には、3つの畳み込み層および2つの非線形活性化層を含むことが理解されることが可能である。

40

【0192】

丸め24

【0193】

丸められたピクチャ特徴マップまたはオーディオ特徴変数を取得するために、たとえばスカラー量子化またはベクトル量子化を使用することによって、ピクチャ特徴マップまたはオーディオ特徴変数を丸めるために、丸めが使用される。

【0194】

50

可能性において、エンコーダ20Aは、量子化パラメータ(quantization parameter, QP)を出力し、たとえば、量子化パラメータを直接に出力し、または量子化パラメータが符号化判定実装ユニットによって符号化され、または圧縮された後で量子化パラメータを出力するように構成されてもよく、それにより、たとえば、デコーダ30Aは、復号のために量子化パラメータを受信し、適用してもよい。

【0195】

可能性において、出力特徴マップまたは特徴オーディオ特徴変数は、丸めの前に前処理され、前処理は、トリミング、カラーフォーマット変換(たとえば、RGBからYCbCrへ)、色補正、ノイズ除去、または同様のものを含み得る。

【0196】

確率推定40

【0197】

ピクチャ特徴マップまたはオーディオ特徴変数の確率推定結果は、確率推定を通じて、入力特徴マップまたは特徴変数情報に基づいて取得される。

【0198】

確率推定は、丸められたピクチャ特徴マップまたはオーディオ特徴変数に対して確率推定を実行するために使用される。

【0199】

確率推定は確率推定ネットワークであってもよく、確率推定ネットワークは畳み込みネットワークであり、畳み込みネットワークは畳み込み層および非線形活性化層を含む。図4Bが例として使用される。確率推定ネットワークは5つのネットワーク層を含み、具体的には、3つの畳み込み層および2つの非線形活性化層を含む。確率推定は、畳み込み非ネットワーク確率推定方法を使用することによって実現され得る。確率推定方法は、限定はされないが、等最尤推定、最大事後確率推定、最尤推定、および別の統計的方法を含む。

【0200】

符号化判定実装26

【0201】

図5に表されるように、符号化判定実装は、符号化要素決定およびエントロピー符号化を含む。ピクチャ特徴マップまたはオーディオ特徴変数は、エンコーダネットワークによって出力される1次元、2次元、または多次元データであり、データの各々は特徴要素である。

【0202】

符号化要素決定261

【0203】

符号化要素決定は、確率推定の確率推定結果情報に基づいて、ピクチャ特徴マップまたはオーディオ特徴変数の各特徴要素を決定し、決定結果に基づいて、エントロピー符号化が実行される特定の特徴要素を決定することである。

【0204】

ピクチャ特徴マップまたはオーディオ特徴変数の第Pの特徴要素の要素決定プロセスが完了した後、ピクチャ特徴マップの第(P+1)の特徴要素の要素決定プロセスが開始され、Pは正の整数であり、PはMより小さい。

【0205】

エントロピー符号化262

【0206】

エントロピー符号化は、様々な開示されたエントロピー符号化アルゴリズムを使用して、エントロピー符号化、たとえば、可変長コーディング(variable length coding, VLC)方式、コンテキスト適応VLC(context adaptive VLC, CAVLC)方式、エントロピー符号化方式、二値化アルゴリズム、コンテキスト適応バイナリエントロピー符号化(context adaptive binary arithmetic coding, CABAC)、シンタックスベースコンテキスト適応バイナリエントロピー符号化(syntax-based context-adaptive binary arithmetic codi

10

20

30

40

50

ng, SBAC)、確率間隔区分エントロピー(probability interval partitioning entropy, PIPE)符号化、または別のエントロピー符号化方法または技術を実行し得る。出力212を通じて符号化されたビットストリーム25または同様のものの形式で出力され得る符号化されたピクチャデータ25が取得され、それによって、デコーダ30Aまたは同様のものが復号のためのパラメータを受信し、使用し得る。符号化されたビットストリーム25は、デコーダ30Aに送信され、またはデコーダ30Aによる後続の送信または取り出しのためにメモリに記憶され得る。

【0207】

別の可能性において、エントロピー符号化は、たとえば畳み込みネットワークを使用することによって実装されるエントロピー符号化ネットワークを使用することによって、符号化を実行し得る。

10

【0208】

可能性において、エントロピー符号化は、丸められた特徴マップの実際の文字確率を知らないで、丸められた特徴マップの実際の文字確率または関連する情報が収集され、エントロピー符号化に追加されてもよく、情報がデコーダ側に送信される。

【0209】

ジョイントネットワーク44

【0210】

ジョイントネットワークは、入力されたサイド情報に基づいて、ピクチャ特徴マップまたはオーディオ特徴変数の確率推定結果および判定情報を取得する。ジョイントネットワークは多層ネットワークであり、ジョイントネットワークは畳み込みネットワークであってもよく、畳み込みネットワークは畳み込み層および非線形活性化層を含む。ジョイントネットワークの任意のネットワーク層が、畳み込み層、正規化層、非線形活性化層、または同様のものであってもよい。

20

【0211】

判定情報は、1次元、2次元、または多次元データであってもよく、判定情報のサイズはピクチャ特徴マップのサイズと一致してもよい。

【0212】

判定情報は、ジョイントネットワークのいずれかのネットワーク層の後で出力され得る。

【0213】

確率推定結果は、ジョイントネットワークのいずれかのネットワーク層の後で出力され得る。

30

【0214】

図6は、ジョイントネットワークのネットワーク構造の出力例である。ネットワーク構造は4つのネットワーク層を含む。判定情報は第4のネットワーク層の後で出力され、確率推定結果は第2のネットワーク層の後で出力される。

【0215】

生成ネットワーク46

【0216】

生成ネットワークは、入力確率推定結果に基づいて、ピクチャ特徴マップの特徴要素の判定情報を取得する。生成ネットワークは多層ネットワークであり、生成ネットワークは畳み込みネットワークであってもよく、畳み込みネットワークは畳み込み層および非線形活性化層を含む。生成ネットワークの任意のネットワーク層が、畳み込み層、正規化層、非線形活性化層、または同様のものであってもよい。

40

【0217】

判定情報は、生成ネットワークのいずれかのネットワーク層の後で出力され得る。判定情報は、1次元、2次元、または多次元データであり得る。

【0218】

図7は、生成ネットワークのネットワーク構造によって判定情報を出力する例である。ネットワーク構造は4つのネットワーク層を含む。

50

【0219】

復号判定実装30

【0220】

図8に表されるように、復号判定実装は、要素決定およびエントロピー復号を含む。ピクチャ特徴マップまたはオーディオ特徴変数は、復号判定実装によって出力される1次元、2次元、または多次元データであり、データの各々は特徴要素である。

【0221】

復号要素決定301

【0222】

復号要素決定は、確率推定の確率推定結果に基づいて、ピクチャ特徴マップまたはオーディオ特徴変数の各特徴要素を決定し、決定結果に基づいて、エントロピー復号が実行される特定の特征要素を決定することである。復号要素決定は、ピクチャ特徴マップまたはオーディオ特徴変数の各特徴要素を決定し、決定結果に基づいて、エントロピー復号が実行される特定の特征要素を決定する。それは、符号化要素決定によって、ピクチャ特徴マップの各特徴要素を決定し、決定結果に基づいて、エントロピー符号化が実行される特定の特征要素を決定することの逆のプロセスと見なされ得る。

【0223】

エントロピー復号302

【0224】

エントロピー復号は、様々な開示されたエントロピー復号アルゴリズムを使用して、エントロピー復号、たとえば、可変長コーディング(variable length coding, VLC)方式、コンテキスト適応VLC(context adaptive VLC, CAVLC)方式、エントロピー復号方式、二値化アルゴリズム、コンテキスト適応バイナリエントロピー復号(context adaptive binary arithmetic coding, CABAC)、シンタックススペースコンテキスト適応バイナリエントロピー復号(syntax-based context-adaptive binary arithmetic coding, SBAC)、確率間隔区分エントロピー(probability interval partitioning entropy, PIPE)符号化、または別のエントロピー符号化方法もしくは技術を実行し得る。出力212を通じて符号化されたビットストリーム25または同様のものの形式で出力され得る符号化されたピクチャ(またはオーディオ)データ25が取得され、それによって、デコーダ30Aまたは同様のものが復号のためのパラメータを受信し、使用し得る。符号化されたビットストリーム25は、デコーダ30Aに送信され、またはデコーダ30Aによる後続の送信または取り出しのためにメモリに記憶され得る。

【0225】

別の可能性において、エントロピー復号は、たとえば畳み込みネットワークを使用することによって実装されるエントロピー復号ネットワークを使用することによって、復号を実行し得る。

【0226】

デコーダネットワーク34

【0227】

デコーダネットワークは、復号されたピクチャ特徴マップまたはオーディオ特徴変数31または後処理された復号されたピクチャ特徴マップまたはオーディオ特徴変数33を、デコーダネットワーク34を通過させて、再構築されたピクチャ(またはオーディオ)データ35または機械向けタスクデータをピクセル領域において取得するために使用される。

【0228】

デコーダネットワークは複数のネットワーク層を含む。任意のネットワーク層が、畳み込み層、正規化層、非線形活性化層、または同様のものであってもよい。重ね合わせ(concat)、加算、および減算などの演算が、デコーダネットワークユニット306に存在し得る。

【0229】

可能性において、デコーダネットワークのネットワーク層の構造は、同じであっても、または互いに異なってもよい。

10

20

30

40

50

【0230】

デコーダネットワークの構造の例が図9に表されている。この例では、デコーダネットワークが5つのネットワーク層を含み、具体的には、1つの正規化層、2つの畳み込み層、および2つの非線形活性化層を含むことが理解されることが可能である。

【0231】

デコーダネットワークは、再構築されたピクチャ(またはオーディオ)を出力し、または取得された機械向けタスクデータを出力する。具体的には、デコーダネットワークは、目標認識ネットワーク、分類ネットワーク、またはセマンティック区分ネットワークを含み得る。

【0232】

エンコーダ20Aおよびデコーダ30Aにおいて、現在のステップの処理結果はさらに処理され、次いで次のステップに出力され得ることが理解されるべきである。たとえば、エンコーダユニットまたはデコーダユニットの後で、さらなる演算または処理、たとえば、クリップ(clip)またはシフト(shift)演算またはフィルタリング処理が、エンコーダユニットまたはデコーダユニットの処理結果に対して実行され得る。

【0233】

前述の説明に基づいて、以下は、この出願の実施形態による、いくつかのピクチャ特徴マップまたはオーディオ特徴変数符号化および復号方法を提供する。説明の容易さのために、以下で説明される方法の実施形態は、一連の動作ステップの組み合わせとして表現される。しかしながら、この出願の技術的解決策の具体的な実装は、説明される一連の動作ステップの順序に限定されないことを、この技術分野の当業者は理解すべきである。

【0234】

以下は、添付の図面を参照して、この出願の手順を詳細に説明する。フローチャートにおけるエンコーダ側のプロセスは、具体的にはエンコーダ20Aによって実行されてもよく、フローチャートにおけるデコーダ側のプロセスは、具体的にはデコーダ30Aによって実行されてもよいことが留意されるべきである。

【0235】

実施形態1から実施形態5において、第1の特徴要素または第2の特徴要素は、現在の符号化されることになる特徴要素または現在の復号されることになる特徴要素、たとえば、

【数11】

$$\hat{y}[x][y][i]$$

である。判定マップは、判定mapとも呼ばれ得る。判定マップは好ましくは二進マップであり、二進マップは二進mapとも呼ばれ得る。

【0236】

この出願の実施形態1において、図10Aは具体的な実装プロセス1400を表す。実行ステップは以下の通りである。

【0237】

エンコーダ側:

ステップ1401: ピクチャ特徴マップを取得する。

【0238】

このステップは、具体的には、図3Aのエンコーダネットワーク204によって実施される。詳細については、エンコーダネットワーク20の前述の説明を参照されたい。ピクチャ特徴マップ y を出力するために、ピクチャが特徴抽出モジュールに入力され、特徴マップ y は、その次元が w 、 x 、 h 、 x 、および c である3次元データであってもよい。具体的には、特徴抽出モジュールは、既存のニューラルネットワークを使用することによって実装されてもよい。これはここで限定されない。このステップは既存の技術である。

【0239】

特徴量子化モジュールは、特徴マップ y の各特徴値を量子化し、浮動小数点数の特徴値を

丸めて、丸めを通じて整数の特徴値を取得して、量子化された特徴マップ

【数 1 2】

$$\hat{y}$$

を取得する。前述の実施形態における丸め24の説明を参照されたい。

【0 2 4 0】

ステップ1402: 特徴マップ

【数 1 3】

$$\hat{y}$$

10

に対して確率推定を実行して、特徴要素の確率推定結果、すなわち、特徴マップ

【数 1 5】

$$\hat{y}$$

の各特徴要素

【数 1 4】

$$\hat{y}[x][y][i]$$

20

の確率分布を取得する。

【0 2 4 1】

パラメータ x 、 y 、および i は正の整数であり、座標 (x, y, i) は現在の符号化されることになる特徴要素の位置を示す。具体的には、座標 (x, y, i) は、左上の頂点の特徴要素に相対的な、現在の3次元特徴マップの現在の符号化されることになる特徴要素の位置を示す。このステップは、具体的には、図3Aの確率推定210によって実施される。詳細については、確率推定40の前述の説明を参照されたい。具体的には、確率分布モデルが、確率分布を取得するために使用され得る。たとえば、単一ガウスモデル(Gaussian single model, GSM)または混合ガウスモデル(Gaussian mix model, GMM)がモデリングのために使用される。まず、サイド情報

30

【数 1 6】

$$\hat{z}$$

およびコンテキスト情報が、確率推定ネットワークに入力される。確率推定が、特徴マップ

【数 1 8】

$$\hat{y}$$

40

の各特徴要素

【数 1 7】

$$\hat{y}[x][y][i]$$

に対して実行されて、特徴要素

【数 1 9】

$$\hat{y}[x][y][i]$$

50

の確率分布を取得する。確率推定ネットワークは、深層学習に基づくネットワーク、たとえば、リカレントニューラルネットワーク(recurrent neural network, RNN)および畳み込みニューラルネットワーク(convolutional neural network, CNN)を使用し得る。これはここで限定されない。確率分布は、モデルパラメータを確率分布モデルに置き換えることによって取得される。

【 0 2 4 2 】

ステップ1403: 圧縮されたビットストリームを取得するために特徴マップ

【数 2 0 】

$\hat{y}$

10

に対してエントロピー符号化を実行し、圧縮されたビットストリームを生成する。

【 0 2 4 3 】

このステップは、具体的には、図3Aの符号化判定実装208によって実施される。詳細については、符号化判定実装26の前述の説明を参照されたい。現在の符号化されることになる特徴要素

【数 2 1 】

$\hat{y}[x][y][i]$

20

の値がkである確率Pは、確率分布に基づいて取得される。現在の符号化されることになる特徴要素

【数 2 2 】

$\hat{y}[x][y][i]$

の確率推定結果Pがあらかじめ設定された条件を満たさない、すなわち、Pは第1の閾値T0より大きい(またはそれに等しい)とき、現在の符号化されることになる特徴要素に対してエントロピー符号化プロセスを実行することがスキップされる。そうでなければ、現在の符号化されることになる特徴要素の確率推定結果Pがあらかじめ設定された条件を満たす、すなわち、Pは第1の閾値T0より小さいとき、エントロピー符号化が、現在の符号化されることになる特徴要素に対して実行され、現在の符号化されることになる特徴要素が、ビットストリームに書き込まれる。kは任意の整数、たとえば0、1、-1、2、または3であり得る。第1の閾値T0は、0 T0 1を満たす任意の数であり、たとえば、値は0.99、0.98、0.97、0.95、または同様のものである(各特徴要素の閾値は同じであると見なされ得る)。

【 0 2 4 4 】

ステップ1404: エンコーダが圧縮されたビットストリームを送信または記憶する。

【 0 2 4 5 】

デコーダ側:

ステップ1411: 復号されたピクチャ特徴マップのビットストリームを取得する。

【 0 2 4 6 】

ステップ1412: ビットストリームに基づいて確率推定を実行して、特徴要素の確率推定結果を取得する。

【 0 2 4 7 】

このステップは、具体的には、図10Bの確率推定302によって実施される。詳細については、確率推定40の前述の説明を参照されたい。確率推定が、復号されることになる特徴マップ

【数 2 4 】

50

$\hat{y}$

の各特徴要素

【数 2 3】

$\hat{y}[x][y][i]$

に対して実行されて、復号されることになる特徴要素

【数 2 5】

$\hat{y}[x][y][i]$

10

の確率分布を取得する。復号されることになる特徴マップ

【数 2 6】

$\hat{y}$

は複数の特徴要素を含み、複数の特徴要素は現在の復号されることになる特徴要素を含む。

【0 2 4 8】

20

デコーダ側によって使用される確率推定ネットワークの構造の図は、この実施形態ではエンコーダ側の確率推定ネットワークのものと同一である。

【0 2 4 9】

ステップ1413: 復号されることになる特徴マップ

【数 2 7】

$\hat{y}$

に対してエントロピー復号を実行する。

【0 2 5 0】

30

このステップは、具体的には、図10Bの復号判定実装304によって実施される。詳細については、復号判定実装30の前述の説明を参照されたい。現在の復号されることになる特徴要素の値がkである確率P、すなわち現在の復号されることになる特徴要素の確率推定結果Pは、現在の復号されることになる特徴要素の確率分布に基づいて取得される。確率推定結果Pがあらかじめ設定された条件を満たさない、すなわち、Pは第1の閾値T0より大きいとき、エントロピー復号は現在の復号されることになる特徴要素に対して実行される必要はなく、現在の復号されることになる特徴要素の値はkに設定される。そうでなければ、現在の復号されることになる特徴要素があらかじめ設定された条件を満たす、すなわち、Pは第1の閾値T0以下であるとき、エントロピー復号がビットストリームに対して実行され、現在の復号されることになる特徴要素の値が取得される。

40

【0 2 5 1】

ビットストリームを解析することによって、かつ第1の閾値T0に基づいて、インデックス番号がビットストリームから取得され得る。デコーダ側は、エンコーダと同じ方式で閾値候補リストを構築し、次いで、あらかじめ設定された閾値候補リストにおける閾値とインデックス番号との間の対応関係に従って、対応する閾値を取得する。インデックス番号はビットストリームから取得され、言い換えると、インデックス番号は、シーケンスヘッダ、ピクチャヘッダ、slice / スライスヘッダ、またはSEIから取得される。

【0 2 5 2】

代替として、ビットストリームは直接に解析されてもよく、閾値はビットストリームから取得される。具体的には、閾値は、シーケンスヘッダ、ピクチャヘッダ、slice / スライ

50

スヘッダ、またはSEIから取得される。

【0253】

代替として、復号と合致した閾値ポリシーに従って、固定された閾値が直接に設定される。

【0254】

ステップ1414: 復号された特徴マップ

【数28】

$\hat{y}$

10

を再構築し、または、対応する機械タスクを実行するための機械向けビジョンタスクモジュールに復号された特徴マップを入力することによって、対応する機械タスクを実行する。このステップは、具体的には、図10Bのデコーダネットワーク306によって実施され得る。詳細については、デコーダネットワーク34の前述の説明を参照されたい。

【0255】

事例1: エントロピー復号を通じて取得された特徴マップ

【数29】

$\hat{y}$

20

がピクチャ再構築モジュールに入力され、ニューラルネットワークが再構築されたマップを出力する。ニューラルネットワークは、任意の構造、たとえば、全結合ネットワーク、畳み込みニューラルネットワーク、またはリカレントニューラルネットワークを使用し得る。ニューラルネットワークは、より良い推定効果を達成するために、多層構造の深層ニューラルネットワーク構造を使用し得る。

【0256】

事例2: エントロピー復号を通じて取得された特徴マップ

【数30】

$\hat{y}$

30

が機械向けビジョンタスクモジュールに入力されて、対応する機械タスクを実行する。たとえば、物体の分類、認識、および区分などのマシンビジョンタスクが完了される。

【0257】

前述のデコーダ側の値kは、それに対応してエンコーダ側の値kに設定される。

【0258】

図11Aは、この出願の実施形態2による具体的な実装プロセス1500を表す。実行ステップは以下の通りである。

【0259】

この実施形態の方法1から方法6において、確率推定結果は第1のパラメータおよび第2のパラメータを含むことが留意されるべきである。確率分布がガウス分布であるとき、第1のパラメータは平均値 μ であり、第2のパラメータは分散 σ である。確率分布がラプラス分布であるとき、第1のパラメータは位置パラメータ μ であり、第2のパラメータはスケールパラメータ b である。

40

【0260】

エンコーダ側:

【0261】

ステップ1501: ピクチャ特徴マップを取得する。

【0262】

このステップは、具体的には、図3Bのエンコーダネットワーク204によって実施される

50

。詳細については、エンコーダネットワーク20の前述の説明を参照されたい。ピクチャ特徴マップ y を出力するために、ピクチャが特徴抽出モジュールに入力され、特徴マップ y は、その次元が w 、 x 、 h 、 x 、および c である3次元データであってもよい。具体的には、特徴抽出モジュールは、既存のニューラルネットワークを使用することによって実装されてもよい。これはここで限定されない。このステップは既存の技術である。

【0263】

特徴量子化モジュールは、特徴マップ y の各特徴値を量子化し、浮動小数点数の特徴値を丸めて、整数の特徴値を取得して、量子化された特徴マップ

【数31】

$$\hat{y}$$

10

を取得する。

【0264】

ステップ1502: ピクチャ特徴マップ

【数32】

$$\hat{y}$$

がサイド情報抽出モジュールに入力され、サイド情報

20

【数33】

$$\hat{z}$$

が出力される。

【0265】

このステップは、具体的には、図3Bのサイド情報抽出ユニット214によって実施される。サイド情報抽出モジュールは、図12に表されるネットワークを使用することによって実装され得る。サイド情報

【数34】

30

$$\hat{z}$$

は、特徴マップ

【数36】

$$\hat{y}$$

をさらに抽出することによって取得される特徴マップ

【数35】

40

$$\hat{z}$$

として理解されてもよく、

【数37】

$$\hat{z}$$

に含まれる特徴要素の数量は、特徴マップ

【数38】

50

$\hat{y}$

の特徴要素の数量より少ない。

【 0 2 6 6 】

エントロピー符号化が、サイド情報

【数 3 9 】

 $\hat{z}$

10

に対して実行されてもよく、サイド情報

【数 4 0 】

 $\hat{z}$

がこのステップにおいてビットストリームに書き込まれ、または、エントロピー符号化が、サイド情報

【数 4 1 】

 $\hat{z}$

20

に対して実行されてもよく、サイド情報

【数 4 2 】

 $\hat{z}$

が後続のステップ1504においてビットストリームに書き込まれることが留意されるべきである。これはここで限定されない。

【 0 2 6 7 】

ステップ1503: 特徴マップ

【数 4 3 】

 $\hat{y}$

30

に対して確率推定を実行して、特徴要素の確率推定結果を取得する。

【 0 2 6 8 】

このステップは、具体的には、図3Bの確率推定210によって実施される。詳細については、確率推定40の前述の説明を参照されたい。確率分布モデルは、確率推定結果および確率分布を取得するために使用され得る。確率分布モデルは、単一ガウスモデル(Gaussian single model, GSM)、非対称ガウスモデル、混合ガウスモデル(Gaussian mix model, GMM)、またはラプラス分布(Laplace distribution)モデルであり得る。

【 0 2 6 9 】

確率分布モデルがガウスモデル(単一ガウスモデル、非対称ガウスモデル、または混合ガウスモデル)であるとき、まず、サイド情報

【数 4 4 】

 $\hat{z}$

40

またはコンテキスト情報が確率推定ネットワークに入力され、確率推定が、特徴マップ

【数 4 6 】

50

$\hat{y}$

の各特徴要素

【数 4 5】

 $\hat{y}[x][y][i]$

に対して実行されて、平均値パラメータ μ および分散 σ の値を取得する。さらに、平均値パラメータ μ および分散 σ は、確率分布を取得するために、使用される確率分布モデルに

10

【 0 2 7 0】

確率分布モデルがラプラス分布モデルであるとき、まず、サイド情報

【数 4 7】

 $\hat{z}$

またはコンテキスト情報が確率推定ネットワークに入力され、確率推定が、特徴マップ

【数 4 9】

 $\hat{y}$

20

の各特徴要素

【数 4 8】

 $\hat{y}[x][y][i]$

に対して実行されて、位置パラメータ μ およびスケールパラメータ b の値を取得する。さらに、位置パラメータ μ およびスケールパラメータ b は、確率分布を取得するために、使用される確率分布モデルに入力される。この場合、確率推定結果は、位置パラメータ μ およびスケールパラメータ b を含む。

30

【 0 2 7 1】

代替として、サイド情報

【数 5 0】

 $\hat{z}$

および / またはコンテキスト情報が確率推定ネットワークに入力されてもよく、確率推定が、符号化されることになる特徴マップ

【数 5 2】

 $\hat{y}$

40

の各特徴要素

【数 5 1】

 $\hat{y}[x][y][i]$

に対して実行されて、現在の符号化されることになる特徴要素

【数 5 3】

50

$\hat{y}[x][y][i]$

の確率分布を取得する。現在の符号化されることになる特徴要素

【数 5 4】

$\hat{y}[x][y][i]$

の値がmである確率Pは、確率分布に基づいて取得される。この場合、確率推定結果は、現在の符号化されることになる特徴要素

【数 5 5】

$\hat{y}[x][y][i]$

の値がmである確率Pである。

【 0 2 7 2】

確率推定ネットワークは、深層学習に基づくネットワーク、たとえば、リカレントニューラルネットワーク(recurrent neural network, RNN)および畳み込みニューラルネットワーク(convolutional neural network, CNN)を使用し得る。これはここで限定されない。

【 0 2 7 3】

ステップ1504: 確率推定結果に基づいて、エントロピー符号化が現在の符号化されることになる特徴要素

【数 5 6】

$\hat{y}[x][y][i]$

に対して実行される必要があるかどうかを決定し、決定結果に基づいて、エントロピー符号化を実行し、現在の符号化されることになる特徴要素

【数 5 7】

$\hat{y}[x][y][i]$

を圧縮されたビットストリーム(符号化されたビットストリーム)に書き込み、またはエントロピー符号化を実行することをスキップする。エントロピー符号化が現在の符号化されることになる第1の特徴要素に対して実行される必要があると決定されるときのみ、現在の符号化されることになる特徴要素に対してエントロピー符号化が実行される。

【 0 2 7 4】

このステップは、具体的には、図3Bの符号化判定実装208によって実施される。詳細については、符号化判定実装26の前述の説明を参照されたい。以下の方法の1つまたは複数は、確率推定結果に基づいて、エントロピー符号化が現在の符号化されることになる特徴要素

【数 5 8】

$\hat{y}[x][y][i]$

に対して実行される必要があるかどうかを決定するために使用され得る。パラメータx、y、およびiは正の整数であり、座標(x, y, i)は現在の符号化されることになる特徴要素の位置を示す。具体的には、座標(x, y, i)は、左上の頂点の特徴要素に相対的な、現在の3次元特徴マップの現在の符号化されることになる特徴要素の位置を示す。

【 0 2 7 5 】

方法1: 確率分布モデルがガウス分布であるとき、現在の符号化されることになる特徴要素に対してエントロピー符号化を実行するかどうかは、第1の特徴要素の確率推定結果に基づいて決定される。現在の符号化されることになる特徴要素のガウス分布の平均値パラメータ μ および分散 σ の値があらかじめ設定された条件を満たさない、すなわち、平均値 μ と k との間の差の絶対値は第2の閾値 $T1$ より小さく、かつ分散 σ は第3の閾値 $T2$ より小さいとき、エントロピー符号化プロセスは、現在の符号化されることになる特徴要素

【 数 5 9 】

$$\hat{y}[x][y][i]$$

10

に対して実行される必要はない。そうでなければ、あらかじめ設定された条件が満たされるとき、すなわち、平均値 μ と k との間の差の絶対値が第2の閾値 $T1$ 以上である、または分散 σ が第3の閾値 $T2$ より小さいとき、エントロピー符号化が、現在の符号化されることになる特徴要素

【 数 6 0 】

$$\hat{y}[x][y][i]$$

に対して実行され、現在の符号化されることになる特徴要素

20

【 数 6 1 】

$$\hat{y}[x][y][i]$$

がビットストリームに書き込まれる。 k は任意の整数、たとえば0、1、-1、2、または3である。 $T2$ の値は、0 $T2$ 1を満たす任意の数、たとえば、0.2、0.3、0.4、または同様のものの値である。 $T1$ は0以上かつ1より小さい数、たとえば0.01、0.02、0.001、および0.002である。

【 0 2 7 6 】

特に、 k の値が0であるとき、それは最適な値である。ガウス分布の平均値パラメータ μ の絶対値が $T1$ より小さく、かつガウス分布の分散 σ が $T2$ より小さいとき、現在の符号化されることになる特徴要素

30

【 数 6 2 】

$$\hat{y}[x][y][i]$$

に対してエントロピー符号化プロセスを実行することがスキップされることが、直接に決定され得る。そうでなければ、エントロピー符号化が、現在の符号化されることになる特徴要素

【 数 6 3 】

$$\hat{y}[x][y][i]$$

40

に対して実行され、現在の符号化されることになる特徴要素

【 数 6 4 】

$$\hat{y}[x][y][i]$$

がビットストリームに書き込まれる。 $T2$ の値は、0 $T2$ 1を満たす任意の数、たとえば、0.2、0.3、0.4、または同様のものの値である。 $T1$ は0以上かつ1より小さい数、たと

50

例えば0.01、0.02、0.001、および0.002である。

【0277】

方法2: 確率分布がガウス分布であるとき、現在の符号化されることになる特徴要素

【数65】

$$\hat{y}[x][y][i]$$

のガウス分布の平均値パラメータ μ および分散 σ の値は、確率推定結果に基づいて取得される。平均値 μ と、分散 σ と、 k との間の関係が $\text{abs}(\mu - k) + \sigma \leq T3$ を満たす(あらかじめ設定された条件が満たされない)とき、現在の符号化されることになる特徴要素

10

【数66】

$$\hat{y}[x][y][i]$$

に対してエントロピー符号化プロセスを実行することはスキップされ、 $\text{abs}(\mu - k)$ は平均値 μ と k との間の差の絶対値を計算することを表現する。そうでなければ、現在の符号化されることになる特徴要素の確率推定結果が $\text{abs}(\mu - k) + \sigma \leq T3$ (あらかじめ設定された条件)を満たすとき、エントロピー符号化が、現在の符号化されることになる特徴要素

【数67】

$$\hat{y}[x][y][i]$$

20

に対して実行され、現在の符号化されることになる特徴要素

【数68】

$$\hat{y}[x][y][i]$$

がビットストリームに書き込まれる。 k は任意の整数、たとえば0、1、-1、-2、または3である。第4の閾値 $T3$ は0以上かつ1より小さい数であり、たとえば、値は0.2、0.3、0.4、または同様のものである。

30

【0278】

方法3: 確率分布がラプラス分布であるとき、現在の符号化されることになる特徴要素

【数69】

$$\hat{y}[x][y][i]$$

のラプラス分布の位置パラメータ μ およびスケールパラメータ b の値は、確率推定結果に基づいて取得される。位置パラメータ μ と、スケールパラメータ b と、 k との間の関係が $\text{abs}(\mu - k) + \sigma \leq T4$ を満たす(あらかじめ設定された条件が満たされない)とき、現在の符号化されることになる特徴要素

40

【数70】

$$\hat{y}[x][y][i]$$

に対してエントロピー符号化プロセスを実行することはスキップされ、 $\text{abs}(\mu - k)$ は位置パラメータ μ と k との間の差の絶対値を計算することを表現する。そうでなければ、現在の符号化されることになる特徴要素の確率推定結果が $\text{abs}(\mu - k) + \sigma \leq T4$ を満たす(あらかじめ設定された条件)とき、エントロピー符号化が、現在の符号化されることになる特徴要素

【数71】

50

$$\hat{y}[x][y][i]$$

に対して実行され、現在の符号化されることになる特徴要素

【数 7 2】

$$\hat{y}[x][y][i]$$

がビットストリームに書き込まれる。kは任意の整数、たとえば0、1、-1、-2、または3である。第4の閾値T4は0以上かつ0.5より小さい数であり、たとえば、値は0.05、0.09、0.17、または同様のものである。

10

【0 2 7 9】

方法4: 確率分布がラプラス分布であるとき、現在の符号化されることになる特徴要素

【数 7 3】

$$\hat{y}[x][y][i]$$

のラプラス分布の位置パラメータ μ およびスケールパラメータbの値は、確率推定結果に基づいて取得される。位置パラメータ μ とkとの間の差の絶対値が第2の閾値T5より小さく、かつスケールパラメータbが第3の閾値T6より小さい(あらかじめ設定された条件が満たされない)とき、現在の符号化されることになる特徴要素

20

【数 7 4】

$$\hat{y}[x][y][i]$$

に対してエントロピー符号化プロセスを実行することはスキップされる。そうでなければ、位置パラメータ μ とkとの間の差の絶対値が第2の閾値T5より小さい、または、スケールパラメータbが第3の閾値T6以上である(あらかじめ設定された条件)とき、エントロピー符号化が、現在の符号化されることになる特徴要素

【数 7 5】

30

$$\hat{y}[x][y][i]$$

に対して実行され、現在の符号化されることになる特徴要素

【数 7 6】

$$\hat{y}[x][y][i]$$

がビットストリームに書き込まれる。kは任意の整数、たとえば0、1、-1、-2、または3である。T5の値は $1e-2$ であり、T6の値はT6 0.5を満たす任意の数、たとえば、0.05、0.09、0.17、または同様のものの値である。

40

【0 2 8 0】

特に、kの値が0であるとき、それは最適な値である。位置パラメータ μ の絶対値がT5より小さく、かつスケールパラメータbがT6より小さいとき、現在の符号化されることになる特徴要素

【数 7 7】

$$\hat{y}[x][y][i]$$

に対してエントロピー符号化プロセスを実行することがスキップされることが、直接に決

50

定され得る。そうでなければ、エントロピー符号化が、現在の符号化されることになる特徴要素

【数 7 8】

$\hat{y}[x][y][i]$

に対して実行され、現在の符号化されることになる特徴要素

【数 7 9】

$\hat{y}[x][y][i]$

10

がビットストリームに書き込まれる。閾値T5の値は $1e-2$ であり、T2の値はT6 0.5を満たす任意の数、たとえば、0.05、0.09、0.17、または同様のものの値である。

【0 2 8 1】

方法5: 確率分布が混合ガウス分布であるとき、現在の符号化されることになる特徴要素

【数 8 0】

$\hat{y}[x][y][i]$

の混合ガウス分布のすべての平均値パラメータ μ_i および分散 σ_i の値は、確率推定結果に基づいて取得される。混合ガウス分布の任意の分散と、混合ガウス分布のすべての平均値とkとの間の差の絶対値の和と、の和が第5の閾値T7より小さい(あらかじめ設定された条件が満たされない)とき、現在の符号化されることになる特徴要素

20

【数 8 1】

$\hat{y}[x][y][i]$

に対してエントロピー符号化プロセスを実行することはスキップされる。そうでなければ、混合ガウス分布の任意の分散と、混合ガウス分布のすべての平均値とkとの間の差の絶対値の和と、の和が第5の閾値T7以上である(あらかじめ設定された条件)とき、エントロピー符号化が、現在の符号化されることになる特徴要素

30

【数 8 2】

$\hat{y}[x][y][i]$

に対して実行され、現在の符号化されることになる特徴要素

【数 8 3】

$\hat{y}[x][y][i]$

40

がビットストリームに書き込まれる。kは任意の整数、たとえば0、1、-1、-2、または3である。T7は0以上かつ1より小さい数であり、たとえば、値は0.2、0.3、0.4、または同様のものである(各特徴要素の閾値は同じであると見なされ得る)。

【0 2 8 2】

方法6: 現在の符号化されることになる特徴要素

【数 8 4】

$\hat{y}[x][y][i]$

の値がkである確率Pは、確率分布に基づいて取得される。現在の符号化されることになる

50

特徴要素の確率推定結果Pがあらかじめ設定された条件を満たさない、すなわち、Pは第1の閾値T0より大きい(またはそれに等しい)とき、現在の符号化されることになる特徴要素に対してエントロピー符号化プロセスを実行することはスキップされる。そうでなければ、現在の符号化されることになる特徴要素の確率推定結果Pがあらかじめ設定された条件を満たす、すなわち、Pは第1の閾値T0より小さいとき、エントロピー符号化が、現在の符号化されることになる特徴要素に対して実行され、現在の符号化されることになる特徴要素が、ビットストリームに書き込まれる。kは任意の整数、たとえば0、1、-1、2、または3であり得る。第1の閾値T0は、 $0 \leq T0 \leq 1$ を満たす任意の数であり、たとえば、値は0.99、0.98、0.97、0.95、または同様のものである(各特徴要素の閾値は同じであると見なされ得る)。

10

【0283】

実際の応用では、プラットフォームの一貫性を確保するために、閾値T1、T2、T3、T4、T5、およびT6は丸められ、すなわち整数へとシフトされてスケールリングされてもよいことが留意されるべきである。

【0284】

閾値を取得するための方法は、以下の方法の1つを代替として使用してもよいことが留意されるべきである。これはここで限定されない。

【0285】

方法1: 閾値T1が例として使用され、T1の値の範囲内の任意の値が閾値T1として使用され、閾値T1がビットストリームに書き込まれる。具体的には、閾値はビットストリームに書き込まれ、シーケンスヘッダ、ピクチャヘッダ、slice / スライスヘッダ、またはSEIに記憶され、デコーダ側に送信され得る。代替として、別の方法が使用され得る。これはここで限定されない。類似する方法が、残りの閾値T0、T2、T3、T4、T5、およびT6のためにも使用され得る。

20

【0286】

方法2: エンコーダ側が、デコーダ側と合致した固定された閾値を使用する。固定された閾値は、ビットストリームに書き込まれる必要はなく、デコーダ側に送信される必要はない。たとえば、閾値T1が例として使用され、T1の値の範囲内の任意の値がT1の値として直接に使用される。類似する方法が、残りの閾値T0、T2、T3、T4、T5、およびT6のためにも使用され得る。

30

【0287】

方法3: 閾値候補リストが構築され、T1の値の範囲内の最も可能性のある値が、閾値候補リストに入れられる。各閾値は閾値インデックス番号に対応し、最適な閾値が決定され、最適な閾値がT1の値として使用される。最適な閾値のインデックス番号がT1の閾値インデックス番号として使用され、T1の閾値インデックス番号がビットストリームに書き込まれる。具体的には、閾値はビットストリームに書き込まれ、シーケンスヘッダ、ピクチャヘッダ、slice / スライスヘッダ、またはSEIに記憶され、デコーダ側に送信され得る。代替として、別の方法が使用され得る。これはここで限定されない。類似する方法が、残りの閾値T0、T2、T3、T4、T5、およびT6のためにも使用され得る。

【0288】

ステップ1505: エンコーダが圧縮されたビットストリームを送信または記憶する。

40

【0289】

デコーダ側:

【0290】

ステップ1511: 復号されることになるピクチャ特徴マップのビットストリームを取得する。

【0291】

ステップ1512: 特徴要素の確率推定結果を取得する。

【0292】

このステップは、具体的には、図11Aの確率推定ユニット302によって実施される。詳

50

細については、確率推定40の前述の説明を参照されたい。エントロピー復号がサイド情報
【数 8 5】

$\hat{z}$

に対して実行されて、サイド情報

【数 8 6】

$\hat{z}$

10

を取得し、確率推定が、復号されることになる特徴マップ

【数 8 8】

$\hat{y}$

の各特徴要素

【数 8 7】

$\hat{y}[x][y][i]$

20

に対して、サイド情報

【数 8 9】

$\hat{z}$

を参照して実行されて、現在の復号されることになる特徴要素

【数 9 0】

$\hat{y}[x][y][i]$

30

の確率推定結果を取得する。

【 0 2 9 3】

それに対応して、デコーダ側によって使用される確率推定方法は、この実施形態においてエンコーダ側によって使用されるものと同じであり、デコーダ側によって使用される確率推定ネットワークの構造の図は、この実施形態ではエンコーダ側の確率推定ネットワークのものと同じであることが留意されるべきである。詳細はここで再び説明されない。

【 0 2 9 4】

ステップ1513: このステップは、具体的には、図11Aの復号判定実装304によって実施される。詳細については、復号判定実装30の前述の説明を参照されたい。エントロピー復号が現在の復号されることになる特徴要素

40

【数 9 1】

$\hat{y}[x][y][i]$

に対して実行される必要があるかどうかは、確率推定結果に基づいて決定され、復号された特徴マップ

【数 9 2】

$\hat{y}$

50

を取得するために、決定結果に基づいて、エントロピー復号が実行され、または実行されない。

【 0 2 9 5 】

以下の方法の1つまたは複数が、確率推定結果に基づいて、エントロピー復号が現在の復号されることになる特徴要素

【数 9 3 】

$\hat{y}[x][y][i]$

に対して実行される必要があるかどうかを決定するために使用され得る。

10

【 0 2 9 6 】

方法1: 確率分布モデルがガウス分布であるとき、現在の復号されることになる特徴要素

【数 9 4 】

$\hat{y}[x][y][i]$

の平均値パラメータ μ および分散 σ の値は、確率推定結果に基づいて取得される。平均値 μ と k との間の差の絶対値が第2の閾値 $T1$ より小さく、かつ分散 σ が第3の閾値 $T2$ より小さい(あらかじめ設定された条件が満たされない)とき、現在の復号されることになる特徴要素

【数 9 5 】

20

$\hat{y}[x][y][i]$

の数値は k に設定され、現在の復号されることになる特徴要素

【数 9 6 】

$\hat{y}[x][y][i]$

に対してエントロピー復号プロセスを実行することはスキップされる。そうでなければ、平均値 μ と k との間の差の絶対値が第2の閾値 $T1$ より小さい、または、分散 σ が第3の閾値 $T2$ 以上である(あらかじめ設定された条件)とき、エントロピー復号が、現在の復号されることになる特徴要素

30

【数 9 7 】

$\hat{y}[x][y][i]$

に対して実行されて、現在の復号されることになる特徴要素

【数 9 8 】

$\hat{y}[x][y][i]$

40

の値を取得する。

【 0 2 9 7 】

特に、 k の値が0であるとき、それは最適な値である。ガウス分布の平均値パラメータ μ の絶対値が $T1$ より小さく、かつガウス分布の分散 σ が $T2$ より小さいとき、現在の復号されることになる特徴要素

【数 9 9 】

$\hat{y}[x][y][i]$

50

の値が k に設定され、現在の復号されることになる特徴要素

【数 1 0 0】

$\hat{y}[x][y][i]$

に対してエントロピー復号プロセスを実行することがスキップされることが、直接に決定され得る。そうでなければ、エントロピー復号が、現在の復号されることになる特徴要素

【数 1 0 1】

$\hat{y}[x][y][i]$

10

に対して実行されて、現在の復号されることになる特徴要素

【数 1 0 2】

$\hat{y}[x][y][i]$

の値が取得される。

【 0 2 9 8】

方法2: 確率分布がガウス分布であるとき、現在の復号されることになる特徴要素

【数 1 0 3】

$\hat{y}[x][y][i]$

20

の平均値パラメータ μ および分散 σ の値は、確率推定結果に基づいて取得される。平均値 μ と、分散 σ と、 k との関係が $\text{abs}(\mu - k) + \sigma \leq T3$ を満たす(あらかじめ設定された条件が満たされない)とき、 $T3$ が第4の閾値であり、現在の復号されることになる特徴要素

【数 1 0 4】

$\hat{y}[x][y][i]$

30

の値が k に設定され、現在の復号されることになる特徴要素

【数 1 0 5】

$\hat{y}[x][y][i]$

に対してエントロピー復号プロセスを実行することがスキップされる。そうでなければ、現在の復号されることになる特徴要素の確率推定結果が $\text{abs}(\mu - k) + \sigma \leq T3$ を満たす(あらかじめ設定された条件)とき、エントロピー復号が、現在の復号されることになる特徴要素

【数 1 0 6】

$\hat{y}[x][y][i]$

40

に対して実行されて、現在の復号されることになる特徴要素

【数 1 0 7】

$\hat{y}[x][y][i]$

の値を取得する。

【 0 2 9 9】

方法3: 確率分布がラプラス分布であるとき、位置パラメータ μ およびスケールパラメー

50

タbの値は、確率推定結果に基づいて取得される。位置パラメータ μ と、スケールパラメータbと、kとの間の関係が $\text{abs}(\mu - k) + \sigma \leq T4$ を満たす(あらかじめ設定された条件が満たされない)とき、T4が第4の閾値であり、現在の復号されることになる特徴要素

【数 1 0 8】

$\hat{y}[x][y][i]$

の値がkに設定され、現在の復号されることになる特徴要素

【数 1 0 9】

$\hat{y}[x][y][i]$

10

に対してエントロピー復号プロセスを実行することがスキップされる。そうでなければ、現在の復号されることになる特徴要素の確率推定結果が $\text{abs}(\mu - k) + \sigma \leq T4$ (あらかじめ設定された条件)を満たすとき、エントロピー復号が、現在の復号されることになる特徴要素

【数 1 1 0】

$\hat{y}[x][y][i]$

に対して実行されて、現在の復号されることになる特徴要素

【数 1 1 1】

$\hat{y}[x][y][i]$

20

の値を取得する。

【 0 3 0 0】

方法4: 確率分布がラプラス分布であるとき、位置パラメータ μ およびスケールパラメータbの値は、確率推定結果に基づいて取得される。位置パラメータ μ とkとの間の差の絶対値が第2の閾値T5より小さく、かつスケールパラメータbが第3の閾値T6より小さい(あらかじめ設定された条件が満たされない)とき、現在の復号されることになる特徴要素

【数 1 1 2】

$\hat{y}[x][y][i]$

30

の値がkに設定され、現在の復号されることになる特徴要素

【数 1 1 3】

$\hat{y}[x][y][i]$

に対してエントロピー復号プロセスを実行することがスキップされる。そうでなければ、位置パラメータ μ とkとの間の差の絶対値が第2の閾値T5より小さい、または、スケールパラメータbが第3の閾値T6以上である(あらかじめ設定された条件)とき、エントロピー復号が、現在の復号されることになる特徴要素

【数 1 1 4】

$\hat{y}[x][y][i]$

40

に対して実行され、現在の復号されることになる特徴要素

【数 1 1 5】

50

$\hat{y}[x][y][i]$

の値が取得される。

【 0 3 0 1 】

特に、kの値が0であるとき、それは最適な値である。位置パラメータ μ の絶対値がT5より小さく、かつスケールパラメータbがT6より小さいとき、現在の復号されることになる

特徴要素

【 数 1 1 6 】

 $\hat{y}[x][y][i]$

10

の値がkに設定され、現在の復号されることになる特徴要素

【 数 1 1 7 】

 $\hat{y}[x][y][i]$

に対してエントロピー復号プロセスを実行することがスキップされることが、直接に決定され得る。そうでなければ、エントロピー復号が、現在の復号されることになる特徴要素

【 数 1 1 8 】

20

 $\hat{y}[x][y][i]$

に対して実行され、現在の復号されることになる特徴要素

【 数 1 1 9 】

 $\hat{y}[x][y][i]$

の値が取得される。

【 0 3 0 2 】

30

方法5: 確率分布が混合ガウス分布であるとき、現在の復号されることになる特徴要素

【 数 1 2 0 】

 $\hat{y}[x][y][i]$

の混合ガウス分布のすべての平均値パラメータ μ_i および分散 σ_i の値は、確率推定結果に基づいて取得される。混合ガウス分布の任意の分散と、混合ガウス分布のすべての平均値と現在の復号されることになる特徴要素の値kとの間の差の絶対値の和と、の和が第5の閾値T7より小さい(あらかじめ設定された条件が満たされない)とき、現在の復号されることになる特徴要素

【 数 1 2 1 】

40

 $\hat{y}[x][y][i]$

の値がkに設定され、現在の復号されることになる特徴要素

【 数 1 2 2 】

 $\hat{y}[x][y][i]$

に対してエントロピー復号プロセスを実行することがスキップされる。そうでなければ、

50

混合ガウス分布の任意の分散と、混合ガウス分布のすべての平均値と現在の復号されることになる特徴要素の値 k との間の差の絶対値の和と、の和が第5の閾値 $T7$ 以上である(あらかじめ設定された条件)とき、エントロピー復号が、現在の復号されることになる特徴要素【数123】

$\hat{y}[x][y][i]$

に対して実行され、現在の復号されることになる特徴要素

【数124】

$\hat{y}[x][y][i]$

10

の値が取得される。

【0303】

方法6: 現在の復号されることになる特徴要素の値が k である確率 P 、すなわち現在の復号されることになる特徴要素の確率推定結果 P は、現在の復号されることになる特徴要素の確率分布に基づいて取得される。確率推定結果 P があらかじめ設定された条件を満たさない、すなわち、 P は第1の閾値 $T0$ より大きいとき、エントロピー復号は現在の復号されることになる特徴要素に対して実行される必要はなく、現在の復号されることになる特徴要素の値は k に設定される。そうでなければ、現在の復号されることになる特徴要素があらかじめ設定された条件を満たす、すなわち、 P は第1の閾値 $T0$ 以下であるとき、エントロピー復号がビットストリームに対して実行され、現在の復号されることになる特徴要素の値が取得される。

20

【0304】

前述のデコーダ側の値 k は、それに対応してエンコーダ側の値 k に設定される。

【0305】

閾値 $T0$ 、 $T1$ 、 $T2$ 、 $T3$ 、 $T4$ 、 $T5$ 、 $T6$ 、および $T7$ を取得するための方法は、エンコーダ側の方法に対応し、以下の方法のうちの1つが使用され得る。

【0306】

方法1: 閾値がビットストリームから取得される。具体的には、閾値は、シーケンスヘッダ、ピクチャヘッダ、slice / スライスヘッダ、またはSEIから取得される。

30

【0307】

方法2: デコーダ側が、エンコーダ側と合致した固定された閾値を使用する。

【0308】

方法3: 閾値インデックス番号がビットストリームから取得される。具体的には、閾値インデックス番号は、シーケンスヘッダ、ピクチャヘッダ、slice / スライスヘッダ、またはSEIから取得される。次いで、デコーダ側が、エンコーダと同じ方式で閾値候補リストを構築し、閾値インデックス番号に基づいて閾値候補リストの中の対応する閾値を取得する。

【0309】

実際の応用では、プラットフォームの一貫性を確保するために、閾値 $T1$ 、 $T2$ 、 $T3$ 、 $T4$ 、 $T5$ 、および $T6$ は丸められ、すなわち整数へとシフトされてスケーリングされてもよいことが留意されるべきである。

40

【0310】

ステップ1514はステップ1414と同じである。

【0311】

図13Aは、この出願の実施形態3による具体的な実装プロセス1600を表す。実行ステップは以下の通りである。

【0312】

エンコーダ側:

【0313】

50

ステップ1601はステップ1501と同じである。このステップは、具体的には、図3Cのエンコーダネットワーク204によって実施される。詳細については、エンコーダネットワーク20の前述の説明を参照されたい。

【0314】

ステップ1602はステップ1502と同じである。このステップは、具体的には、図3Cのサイド情報抽出214によって実施される。

【0315】

ステップ1603: 特徴マップ

【数125】

$\hat{y}$

10

に対して確率推定を実行して、特徴要素の確率推定結果を取得する。

【0316】

このステップは、具体的には、図3Cの確率推定210によって実施され得る。詳細については、確率推定40の前述の説明を参照されたい。確率分布モデルは、確率推定結果を取得するために使用され得る。確率分布モデルは、単一ガウスモデル、非対称ガウスモデル、混合ガウスモデル、またはラプラス分布モデルであり得る。

【0317】

確率分布モデルがガウスモデル(単一ガウスモデル、非対称ガウスモデル、または混合ガウスモデル)であるとき、まず、サイド情報

20

【数126】

$\hat{z}$

またはコンテキスト情報が確率推定ネットワークに入力され、確率推定が、特徴マップ

【数128】

$\hat{y}$

30

の各特徴要素

【数127】

$\hat{y}[x][y][i]$

に対して実行されて、モデルパラメータ平均値パラメータ μ および分散 σ の値、すなわち確率推定結果を取得する。

【0318】

確率分布モデルがラプラス分布モデルであるとき、まず、サイド情報

【数129】

40

$\hat{z}$

またはコンテキスト情報が確率推定ネットワークに入力され、確率推定が、特徴マップ

【数131】

$\hat{y}$

の各特徴要素

【数130】

50

$\hat{y}[x][y][i]$

に対して実行されて、モデルパラメータ位置パラメータ μ およびスケールパラメータ b の値、すなわち確率推定結果を取得する。

【 0 3 1 9 】

さらに、確率推定結果は、確率分布を取得するために、使用される確率分布モデルに入力される。

【 0 3 2 0 】

代替として、サイド情報

【 数 1 3 2 】

$\hat{z}$

10

および / またはコンテキスト情報が確率推定ネットワークに入力されてもよく、確率推定が、符号化されることになる特徴マップ

【 数 1 3 4 】

$\hat{y}$

20

の各特徴要素

【 数 1 3 3 】

$\hat{y}[x][y][i]$

に対して実行されて、現在の符号化されることになる特徴要素

【 数 1 3 5 】

$\hat{y}[x][y][i]$

30

の確率分布を取得する。現在の符号化されることになる特徴要素

【 数 1 3 6 】

$\hat{y}[x][y][i]$

の値が m である確率 P は、確率分布に基づいて取得される。 m は任意の整数、たとえば 0、1、-1、-2、または 3 である。

【 0 3 2 1 】

確率推定ネットワークは、深層学習に基づくネットワーク、たとえば、リカレントニューラルネットワークおよび畳み込みニューラルネットワークを使用し得る。これはここで限定されない。

【 0 3 2 2 】

ステップ 1604: 確率推定結果に基づいて、現在の符号化されることになる特徴要素に対してエントロピー符号化を実行するかどうかを決定する。決定結果に基づいて、エントロピー符号化が現在の符号化されることになる特徴要素に対して実行され、現在の符号化されることになる特徴要素が、符号化されたビットストリームに書き込まれ、またはエントロピー符号化は実行されない。エントロピー符号化が現在の符号化されることになる特徴要素に対して実行される必要があると決定されるときのみ、現在の符号化されることになる特徴要素に対してエントロピー符号化が実行される。

【 0 3 2 3 】

50

このステップは、具体的には、図3Cの生成ネットワーク216および符号化判定実装208によって実施される。詳細については、生成ネットワーク46および符号化判定実装26の前述の説明を参照されたい。確率推定結果211は決定モジュールに入力され、その次元が特徴マップ

【数 1 3 7】

$\hat{y}$

の次元と同じである判定情報217が出力される。この実施形態では、判定情報217は3次元判定マップであり得る。決定モジュールは、ネットワーク方法を使用することによって実装され得る。具体的には、確率推定結果または確率分布は、図7に表される生成ネットワークに入力され、ネットワークは判定マップを出力する。判定マップ $\text{map}[x][y][i]$ があらかじめ設定された値であるとき、それは、エントロピー符号化が、対応する位置における現在の符号化されることになる特徴要素

【数 1 3 8】

$\hat{y}[x][y][i]$

に対して実行される必要があることを示し、エントロピー符号化は、確率分布に基づいて現在の符号化されることになる特徴要素に対して実行される。判定マップ $\text{map}[x][y][i]$ があらかじめ設定された値であるとき、それは、対応する位置における現在の符号化されることになる特徴要素

【数 1 3 9】

$\hat{y}[x][y][i]$

の高い確率値がkであることを示す。判定マップ $\text{map}[x][y][i]$ があらかじめ設定された値ではないとき、それは、エントロピー符号化が、対応する位置における現在の符号化されることになる特徴要素

【数 1 4 0】

$\hat{y}[x][y][i]$

に対して実行される必要はない、言い換えると、エントロピー符号化プロセスがスキップされることを示す。判定情報は、その次元が特徴マップ

【数 1 4 1】

$\hat{y}$

の次元と同じである判定マップである。判定マップ $\text{map}[x][y][i]$ は、判定マップの中の座標位置(x, y, i)における値を示す。生成ネットワークによって出力される、判定マップの現在の符号化されることになる特徴要素

【数 1 4 2】

$\hat{y}$

の2つの選択肢の値のみがあるとき、あらかじめ設定された値は特定の値である。たとえば、現在の符号化されることになる特徴要素の選択肢の値が0および1であるとき、あらかじめ設定された値は0または1である。生成ネットワークによって出力される、判定マップの現在の符号化されることになる特徴要素

【数 1 4 3】

 $\hat{y}$

の複数の選択枝の値があるとき、あらかじめ設定された値はいくつかの特定の値である。
たとえば、現在の符号化されることになる特徴要素

【数 1 4 4】

 $\hat{y}$

10

の選択枝の値が0から255であるとき、あらかじめ設定された値は0から255の真部分集合である。

【0 3 2 4】

可能な実装において、現在の符号化されることになる特徴要素の確率推定結果または確率分布は決定モジュールに入力され、決定モジュールは、エントロピー符号化が現在の符号化されることになる特徴要素に対して実行される必要があるかどうかを示す判定情報を直接に出力する。たとえば、決定モジュールによって出力される判定情報があらかじめ設定された値であるとき、それは、エントロピー符号化が現在の符号化されることになる特徴要素に対して実行される必要があることを示す。決定モジュールによって出力される判定情報があらかじめ設定された値ではないとき、それは、エントロピー符号化が現在の符号化されることになる特徴要素に対して実行される必要はないことを示す。決定モジュールは、ネットワーク方法を使用することによって実装され得る。具体的には、確率推定結果または確率分布は、図7に表される生成ネットワークに入力され、ネットワークは判定情報、すなわちあらかじめ設定された値を出力する。

20

【0 3 2 5】

方法1: 判定情報は、その次元が特徴マップ

【数 1 4 5】

 $\hat{y}$

30

の次元と同じである判定マップであり、かつ判定マップ $\text{map}[x][y][i]$ があらかじめ設定された値であるとき、それは、エントロピー符号化が、対応する位置における現在の符号化されることになる特徴要素

【数 1 4 6】

 $\hat{y}[x][y][i]$

に対して実行される必要があることを示し、エントロピー符号化は、確率分布に基づいて現在の符号化されることになる特徴要素に対して実行される。判定マップ $\text{map}[x][y][i]$ があらかじめ設定された値ではないとき、それは、対応する位置における現在の符号化されることになる特徴要素

40

【数 1 4 7】

 $\hat{y}[x][y][i]$

の高い確率値がkであることを示す。判定マップ $\text{map}[x][y][i]$ が0であるとき、それは、エントロピー符号化が、対応する位置における現在の符号化されることになる特徴要素

【数 1 4 8】

 $\hat{y}[x][y][i]$

50

に対して実行される必要はない、言い換えると、エントロピー符号化プロセスがスキップされることを示す。判定マップの特徴要素

【数 1 4 9】

$\hat{y}$

の2つの選択肢の値のみがあるとき、あらかじめ設定された値は特定の値である。たとえば、特徴要素の選択肢の値が0および1であるとき、あらかじめ設定された値は0または1である。判定マップの特徴要素

【数 1 5 0】

$\hat{y}$

10

の複数の選択肢の値があるとき、あらかじめ設定された値はいくつかの特定の値である。たとえば、特徴要素

【数 1 5 1】

$\hat{y}$

の選択肢の値が0から255であるとき、あらかじめ設定された値は0から255の真部分集合である。

【0 3 2 6】

方法2: 判定情報は、その次元が特徴マップ

【数 1 5 2】

$\hat{y}$

20

の次元と同じである判定マップであり、かつ判定マップ $\text{map}[x][y][i]$ が閾値 T_0 以上であるとき、それは、エントロピー符号化が、対応する位置における現在の符号化されることになる特徴要素

【数 1 5 3】

$\hat{y}[x][y][i]$

30

に対して実行される必要があることを示し、エントロピー符号化は、確率分布に基づいて現在の符号化されることになる特徴要素に対して実行される。判定マップ $\text{map}[x][y][i]$ が閾値 T_0 より小さいとき、それは、対応する位置における現在の符号化されることになる特徴要素

【数 1 5 4】

$\hat{y}[x][y][i]$

40

の高い確率値が k であることを示し、エントロピー符号化が、対応する位置における現在の符号化されることになる特徴要素

【数 1 5 5】

$\hat{y}[x][y][i]$

に対して実行される必要はない、言い換えると、エントロピー符号化プロセスがスキップされることを示す。判定マップの数値範囲を参照して、 T_0 は数値範囲内の平均値であり得

50

る。

【0327】

方法3: 判定情報は、代替として、ジョイントネットワークによって直接に出力される識別子または識別子の値であり得る。判定情報があらかじめ設定された値であるとき、それは、エントロピー符号化が現在の符号化されることになる特徴要素に対して実行される必要があることを示す。決定モジュールによって出力される判定情報があらかじめ設定された値ではないとき、それは、エントロピー符号化が現在の符号化されることになる特徴要素に対して実行される必要はないことを示す。たとえば、識別子または識別子の値の選択肢の数値が0および1であるとき、それに対応して、あらかじめ設定された値は0または1である。識別子または識別子の値が代替として複数の選択肢の値を有し得るとき、あらかじめ設定された値はいくつかの特定の値である。たとえば、識別子または識別子の値の選択肢の値が0から255であるとき、あらかじめ設定された値は0から255の真部分集合である。

10

【0328】

高い確率は、現在の符号化されることになる特徴要素

【数156】

$\hat{y}[x][y][i]$

の値がkである確率が非常に高く、閾値Pより大きいことを意味し、Pは0.9より大きい数、たとえば0.9、0.95、または0.98であり得る。

20

【0329】

ステップ1605: エンコーダが圧縮されたビットストリームを送信または記憶する。

【0330】

ステップ1601からステップ1604は、特徴マップ

【数157】

$\hat{y}$

の少なくとも1つの特徴要素に対して実行されて、圧縮されたビットストリームを取得し、圧縮されたビットストリームはデコーダ側に送信される。

30

【0331】

デコーダ側:

【0332】

ステップ1611: 復号されることになる圧縮されたビットストリームを取得する。

【0333】

ステップ1612: 復号されることになる特徴マップ

【数158】

$\hat{y}$

40

に対して確率推定を実行して、特徴要素の確率推定結果を取得する。

【0334】

このステップは、具体的には、図13Bの確率推定302によって実施され得る。詳細については、確率推定40の前述の説明を参照されたい。サイド情報

【数159】

$\hat{z}$

はビットストリームから取得され、現在の復号されることになる特徴要素の確率推定結果

50

は、ステップ1603における方法を使用することによって取得される。

【0335】

ステップ1613: 判定情報を取得し、判定情報に基づいて、エントロピー復号を実行するかどうかを決定する。

【0336】

このステップは、具体的には、図13Bの生成ネットワーク310および復号判定実装304によって実施され得る。詳細については、生成ネットワーク46および復号判定実装30の前述の説明を参照されたい。判定情報311は、この実施形態ではエンコーダ側の方法と同じ方法を使用することによって取得される。判定マップ $\text{map}[x][y][i]$ があらかじめ設定された値であるとき、それは、エントロピー復号が、対応する位置における現在の復号されることになる特徴要素

10

【数160】

$$\hat{y}[x][y][i]$$

に対して実行される必要があることを示し、エントロピー復号は、確率分布に基づいて現在の復号されることになる特徴要素に対して実行される。判定マップ $\text{map}[x][y][i]$ があらかじめ設定された値ではないとき、それは、エントロピー復号が、対応する位置における現在の復号されることになる特徴要素

【数161】

20

$$\hat{y}[x][y][i]$$

に対して実行される必要はないことを示し、言い換えると、対応する位置

【数162】

$$\hat{y}[x][y][i]$$

が特定の値 k であることを示す。

【0337】

30

可能な実装において、現在の復号されることになる特徴要素の確率推定結果または確率分布は決定モジュールに入力され、決定モジュールは、エントロピー復号が現在の復号されることになる特徴要素に対して実行される必要があるかどうかを示す判定情報を直接に出力する。たとえば、決定モジュールによって出力される判定情報があらかじめ設定された値であるとき、それは、エントロピー復号が現在の復号されることになる特徴要素に対して実行される必要があることを示す。決定モジュールによって出力される判定情報があらかじめ設定された値ではないとき、それは、エントロピー復号が現在の復号されることになる特徴要素に対して実行される必要はないことを示し、現在の復号されることになる特徴要素の値は k に設定される。決定モジュールは、ネットワーク方法を使用することによって実装され得る。具体的には、確率推定結果または確率分布は、図8に表される生成ネットワークに入力され、ネットワークは判定情報、すなわちあらかじめ設定された値を出力する。判定情報は、現在の復号されることになる特徴要素に対してエントロピー復号を実行するかどうかを示し、判定情報は判定マップを含み得る。

40

【0338】

ステップ1614はステップ1414と同じである。

【0339】

前述のデコーダ側の値 k は、それに対応してエンコーダ側の値 k に設定される。

【0340】

図14は、この出願の実施形態4による具体的な実装プロセス1700を表す。実行ステップは以下の通りである。

50

【 0 3 4 1 】

エンコーダ側:

【 0 3 4 2 】

ステップ1701はステップ1501と同じである。このステップは、具体的には、図3Dのエンコーダネットワーク204によって実施され得る。詳細については、エンコーダネットワーク20の前述の説明を参照されたい。

【 0 3 4 3 】

ステップ1702はステップ1502と同じである。このステップは、具体的には、図3Dのサイド情報抽出214によって実施される。

【 0 3 4 4 】

ステップ1703: 特徴マップ

【数 1 6 3 】

$\hat{y}$

10

の各特徴要素の確率推定結果および判定情報を取得する。

【 0 3 4 5 】

このステップは、具体的には、図3Dのジョイントネットワーク218によって実施され得る。詳細については、ジョイントネットワーク34の前述の説明を参照されたい。具体的には、サイド情報

【数 1 6 4 】

$\hat{z}$

20

および/またはコンテキスト情報が、ジョイントネットワークに入力される。ジョイントネットワークは、符号化されることになる特徴マップ

【数 1 6 6 】

$\hat{y}$

30

の各特徴要素

【数 1 6 5 】

$\hat{y}[x][y][i]$

の確率分布および/または確率推定結果、およびその次元が特徴マップ

【数 1 6 7 】

$\hat{y}$

40

の次元と同じである判定情報を出力する。たとえば、サイド情報

【数 1 6 8 】

$\hat{z}$

とコンテキスト情報の両方がジョイントネットワークに入力されるとき、図15に表されるネットワーク構造が使用され得る。

【 0 3 4 6 】

ジョイントネットワークの具体的な構造はこの実施形態において限定されないことが留意されるべきである。

50

【 0 3 4 7 】

判定情報、確率分布、および／または確率推定結果はすべて、ジョイントネットワークの異なる層から出力され得ることが留意されるべきである。たとえば、事例(1)では、ネットワークの中間の層が判定情報を出力し、最後の層が確率分布および／または確率推定結果を出力する。事例(2)では、ネットワークの中間の層が確率分布および／または確率推定結果を出力し、最後の層が判定情報を出力する。事例(3)では、ネットワークの最後の層が、判定情報、および確率分布、および／または確率推定結果を一緒に出力する。

【 0 3 4 8 】

確率分布モデルがガウスモデル(単一ガウスモデル、非対称ガウスモデル、または混合ガウスモデル)であるとき、まず、サイド情報

10

【数 1 6 9】

$$\hat{z}$$

またはコンテキスト情報がジョイントネットワークに入力されて、モデルパラメータ平均値パラメータ μ および分散 σ の値、すなわち確率推定結果を取得する。さらに、確率推定結果がガウスモデルに入力されて、確率分布を取得する。

【 0 3 4 9 】

確率分布モデルがラプラス分布モデルであるとき、まず、サイド情報

【数 1 7 0】

20

$$\hat{z}$$

またはコンテキスト情報がジョイントネットワークに入力されて、モデルパラメータ位置パラメータ μ およびスケールパラメータ b の値、すなわち確率推定結果を取得する。さらに、確率推定結果がラプラス分布モデルに入力されて、確率分布を取得する。

【 0 3 5 0 】

代替として、サイド情報

【数 1 7 1】

$$\hat{z}$$

30

および／またはコンテキスト情報がジョイントネットワークに入力されて、現在の符号化されることになる特徴要素

【数 1 7 2】

$$\hat{y}[x][y][i]$$

の確率分布を取得してもよい。現在の符号化されることになる特徴要素

【数 1 7 3】

40

$$\hat{y}[x][y][i]$$

の値が m である確率 P 、すなわち確率推定結果は、確率分布に基づいて取得される。 m は任意の整数、たとえば0、1、-1、-2、または3である。

【 0 3 5 1 】

ステップ1704: 判定情報に基づいて、エントロピー符号化を実行するかどうかを決定し、決定結果に基づいて、エントロピー符号化を実行し、現在の符号化されることになる特徴要素を圧縮されたビットストリーム(符号化されたビットストリーム)に書き込み、またはエントロピー符号化を実行することをスキップする。エントロピー符号化が現在の符号

50

化されることになる特徴要素に対して実行される必要があると決定されるときのみ、現在の符号化されることになる特徴要素に対してエントロピー符号化が実行される。このステップは、具体的には、図3Dの符号化判定実装208によって実施され得る。詳細については、符号化判定実装26の前述の説明を参照されたい。

【0352】

方法1: 判定情報は、その次元が特徴マップ

【数174】

$$\hat{y}$$

10

の次元と同じである判定マップであり、かつ判定マップ $\text{map}[x][y][i]$ があらかじめ設定された値であるとき、それは、エントロピー符号化が、対応する位置における現在の符号化されることになる特徴要素

【数175】

$$\hat{y}[x][y][i]$$

に対して実行される必要があることを示し、エントロピー符号化は、確率分布に基づいて現在の符号化されることになる特徴要素に対して実行される。判定マップ $\text{map}[x][y][i]$ があらかじめ設定された値ではないとき、それは、対応する位置における現在の符号化されることになる特徴要素

20

【数176】

$$\hat{y}[x][y][i]$$

の高い確率値が k であることを示す。判定マップ $\text{map}[x][y][i]$ が0であるとき、それは、エントロピー符号化が、対応する位置における現在の符号化されることになる特徴要素

【数177】

$$\hat{y}[x][y][i]$$

30

に対して実行される必要はない、言い換えると、エントロピー符号化プロセスがスキップされることを示す。判定マップの現在の符号化されることになる特徴要素

【数178】

$$\hat{y}$$

の2つの選択肢の値のみがあるとき、あらかじめ設定された値は特定の値である。たとえば、現在の符号化されることになる特徴要素の選択肢の値が0および1であるとき、あらかじめ設定された値は0または1である。判定マップの現在の符号化されることになる特徴要素

40

【数179】

$$\hat{y}$$

の複数の選択肢の値があるとき、あらかじめ設定された値はいくつかの特定の値である。たとえば、現在の符号化されることになる特徴要素

【数180】

$$\hat{y}$$

50

の選択肢の値が0から255であるとき、あらかじめ設定された値は0から255の真部分集合である。

【 0 3 5 3 】

方法2: 判定情報は、その次元が特徴マップ

【数 1 8 1 】

$\hat{y}$

の次元と同じである判定マップであり、かつ判定マップ $\text{map}[x][y][i]$ が閾値 $T0$ 以上であるとき、それは、エントロピー符号化が、対応する位置における現在の符号化されることになる特徴要素

10

【数 1 8 2 】

$\hat{y}[x][y][i]$

に対して実行される必要があることを示し、エントロピー符号化は、確率分布に基づいて現在の符号化されることになる特徴要素に対して実行される。判定マップ $\text{map}[x][y][i]$ が閾値 $T0$ より小さいとき、それは、対応する位置における現在の符号化されることになる特徴要素

【数 1 8 3 】

20

$\hat{y}[x][y][i]$

の高い確率値が k であることを示し、エントロピー符号化が、対応する位置における現在の符号化されることになる特徴要素

【数 1 8 4 】

$\hat{y}[x][y][i]$

に対して実行される必要はない、言い換えると、エントロピー符号化プロセスがスキップされることを示す。判定マップ map の数値範囲を参照して、 $T0$ は数値範囲内の平均値であり得る。

30

【 0 3 5 4 】

方法3: 判定情報は、代替として、ジョイントネットワークによって直接に出力される識別子または識別子の値であり得る。判定情報があらかじめ設定された値であるとき、それは、エントロピー符号化が現在の符号化されることになる特徴要素に対して実行される必要があることを示す。決定モジュールによって出力される判定情報があらかじめ設定された値ではないとき、それは、エントロピー符号化が現在の符号化されることになる特徴要素に対して実行される必要がないことを示す。ジョイントネットワークによって出力される、判定マップの現在の符号化されることになる特徴要素の2つの選択肢の値のみがあるとき、あらかじめ設定された値は特定の値である。たとえば、現在の符号化されることになる特徴要素の選択肢の値が0および1であるとき、あらかじめ設定された値は0または1である。ジョイントネットワークによって出力される、判定マップの現在の符号化されることになる特徴要素の複数の選択肢の値があるとき、あらかじめ設定された値はいくつかの特定の値である。たとえば、現在の符号化されることになる特徴要素の選択肢の値が0から255であるとき、あらかじめ設定された値は0から255の真部分集合である。

40

【 0 3 5 5 】

高い確率は、現在の符号化されることになる特徴要素

【数 1 8 5 】

50

$\hat{y}[x][y][i]$

の値がmである確率が非常に高いことを意味する。たとえば、値がkであるとき、確率は閾値Pより大きく、Pは0.9より大きい数、たとえば0.9、0.95、または0.98であり得る。

【0356】

ステップ1705: エンコーダが圧縮されたビットストリームを送信または記憶する。

【0357】

デコーダ側:

【0358】

ステップ1711: 復号されることになるピクチャ特徴マップのビットストリームを取得し、ビットストリームからサイド情報

【数186】

$\hat{z}$

10

を取得する。

【0359】

ステップ1712: 特徴マップ

【数187】

$\hat{y}$

20

の各特徴要素の確率推定結果および判定情報を取得する。

【0360】

このステップは、具体的には、図16のジョイントネットワーク312によって実施され得る。詳細については、ジョイントネットワーク34の前述の説明を参照されたい。特徴マップ

【数188】

$\hat{y}$

30

の各特徴要素の確率推定結果および判定情報が取得される。これはステップ1703と同じである。

【0361】

ステップ1713: 判定情報に基づいて、エントロピー復号を実行するかどうかを決定し、決定結果に基づいてエントロピー復号を実行し、またはスキップする。このステップは、具体的には、図16の復号判定実装304によって実施され得る。詳細については、復号判定実装30の前述の説明を参照されたい。

【0362】

方法1: 判定情報は判定マップであり、かつ判定マップ $\text{map}[x][y][i]$ があらかじめ設定された値であるとき、それは、エントロピー復号が、対応する位置における現在の復号されることになる特徴要素

【数189】

$\hat{y}[x][y][i]$

40

に対して実行される必要があることを示し、エントロピー復号は、確率分布に基づいて現在の復号されることになる特徴要素に対して実行される。判定マップがあらかじめ設定された値 $\text{map}[x][y][i]$ ではないとき、それは、エントロピー復号が、対応する位置における

50

現在の復号されることになる特徴要素

【数 1 9 0】

$\hat{y}[x][y][i]$

に対して実行される必要はないことを示し、言い換えると、対応する位置

【数 1 9 1】

$\hat{y}[x][y][i]$

10

が特定の値kに設定されることを示す。

【0 3 6 3】

方法2: 判定情報は、その次元が特徴マップ

【数 1 9 2】

$\hat{y}$

の次元と同じである判定マップmapであり、判定マップmap[x][y][i]が閾値T0以上であるとき、それは、エントロピー復号が、対応する位置における現在の復号されることになる特徴要素

20

【数 1 9 3】

$\hat{y}[x][y][i]$

に対して実行される必要があることを示す。判定マップmap[x][y][i]が閾値T0より小さいとき、それは、対応する位置における現在の復号されることになる特徴要素

【数 1 9 4】

$\hat{y}[x][y][i]$

30

の高い確率値がkであることを示し、エントロピー復号が、対応する位置における現在の復号されることになる特徴要素

【数 1 9 5】

$\hat{y}[x][y][i]$

に対して実行される必要はない、言い換えると、対応する位置

【数 1 9 6】

$\hat{y}[x][y][i]$

40

が特定の値kに設定されることを示す。T0の値はエンコーダ側の値と同じである。

【0 3 6 4】

方法3: 判定情報は、代替として、ジョイントネットワークによって直接に出力される識別子または識別子の値であり得る。判定情報があらかじめ設定された値であるとき、それは、エントロピー復号が現在の復号されることになる特徴要素に対して実行される必要があることを示す。決定モジュールによって出力される判定情報があらかじめ設定された値ではないとき、それは、エントロピー復号が現在の復号されることになる特徴要素に対して実行される必要はないことを示し、現在の復号されることになる特徴要素の値はkに設定される。ジョイントネットワークによって出力される、判定マップの現在の復号されるこ

50

となる特徴要素の2つの選択肢の値のみがあるとき、あらかじめ設定された値は特定の値である。たとえば、現在の復号されることになる特徴要素の選択肢の値が0および1であるとき、あらかじめ設定された値は0または1である。ジョイントネットワークによって出力される、判定マップの現在の復号されることになる特徴要素の複数の選択肢の値があるとき、あらかじめ設定された値はいくつかの特定の値である。たとえば、現在の復号されることになる特徴要素の選択肢の値が0から255であるとき、あらかじめ設定された値は0から255の真部分集合である。

【0365】

ステップ1714はステップ1414と同じである。このステップは、具体的には、前述の実施形態におけるデコーダ9Cのデコーダネットワークユニット306によって実施され得る。詳細については、前述の実施形態におけるデコーダネットワークユニット306の説明を参照されたい。

10

【0366】

前述のデコーダ側の値 k は、それに対応してエンコーダ側の値 k に設定される。

【0367】

図17は、この出願の実施形態5による具体的な実装プロセス1800を表す。実行ステップは以下の通りである。

【0368】

ステップ1801: 符号化されることになるオーディオデータの特徴変数を取得する。

【0369】

符号化されることになるオーディオ信号は、時間領域オーディオ信号であり得る。符号化されることになるオーディオ信号は、時間周波数変換が時間領域信号に対して実行された後に取得される周波数領域信号であり得る。たとえば、周波数領域信号は、時間領域オーディオ信号に対してMDCT変換が実行された後に取得される周波数領域信号であってもよく、時間領域オーディオ信号は、FFT変換を通じて取得される周波数領域信号である。代替として、符号化されることになる信号は、QMFフィルタリングを通じて取得される信号であり得る。代替として、符号化されることになる信号は、残差信号、たとえば、別の符号化された残差信号またはLPCフィルタリングを通じて取得される残差信号であり得る。

20

【0370】

符号化されることになるオーディオデータの特徴変数を取得することは、符号化されることになるオーディオ信号に基づいて特徴ベクトルを抽出すること、たとえば、符号化されることになるオーディオ信号に基づいてメルケプストラム係数を抽出すること、抽出された特徴ベクトルを量子化すること、および、符号化されることになるオーディオデータの特徴変数として量子化された特徴ベクトルを使用することであり得る。

30

【0371】

代替として、符号化されることになるオーディオデータの特徴変数を取得することは、既存のニューラルネットワークを使用することによって実施されてもよい。たとえば、符号化されることになるオーディオ信号は、潜在変数を取得するために符号化ニューラルネットワークによって処理され、ニューラルネットワークによって出力される潜在変数は量子化され、量子化された潜在変数が、符号化されることになるオーディオデータの特徴変数として使用される。符号化ニューラルネットワーク処理は事前訓練され、符号化ニューラルネットワークの具体的なネットワーク構造および訓練方法は本発明では限定されない。たとえば、全結合ネットワークまたはCNNネットワークが、符号化ニューラルネットワークのために選択され得る。符号化ニューラルネットワークに含まれる層の数量および各層におけるノードの数量は、本発明では限定されない。

40

【0372】

異なる構造の符号化ニューラルネットワークによって出力される潜在変数の形式は異なり得る。たとえば、符号化ニューラルネットワークは全結合ネットワークである。出力潜在変数はベクトルであり、ベクトルの次元 M は潜在変数のサイズ(latent size)、たとえば、 $y=[y(0), y(1), \dots, y(M-1)]$ である。符号化ニューラルネットワークはCNNネットワー

50

クである。出力潜在変数は、 $N \times M$ 次元の行列である。 N はCNNネットワークのチャンネル(channel)数量であり、 M はCNNネットワークの各チャンネルの潜在変数のサイズ(latent size)、たとえば、

【数 1 9 7】

$$y = \begin{bmatrix} y(0,0) & \cdots & y(0,M-1) \\ \vdots & \ddots & \vdots \\ y(N-1,0) & \cdots & y(N-1,M-1) \end{bmatrix}$$

であり、ニューラルネットワークによって出力される潜在変数を量子化するための具体的な方法は、潜在変数の各要素に対してスカラー量子化を実行することであってもよく、スカラー量子化の量子化ステップは、異なる符号化レートに基づいて決定されてもよい。スカラー量子化はバイアスをさらに有し得る。たとえば、量子化されることになる潜在変数に対してバイアス処理が実行された後、決定された量子化ステップに基づいてスカラー量子化が実行される。潜在変数を量子化するための量子化方法は、代替として、別の既存の量子化技術を使用することによって実施され得る。これは本発明では限定されない。

【 0 3 7 3】

量子化された特徴ベクトルまたは量子化された潜在変数の両方が、

【数 1 9 8】

$$\hat{y}$$

、すなわち、符号化されることになるオーディオデータの特徴変数として表記され得る。

【 0 3 7 4】

ステップ1802: 符号化されることになるオーディオデータの特徴変数

【数 1 9 9】

$$\hat{y}$$

がサイド情報抽出モジュールに入力され、サイド情報

【数 2 0 0】

$$\hat{z}$$

が出力される。

【 0 3 7 5】

サイド情報抽出モジュールは、図12に表されるネットワークを使用することによって実装され得る。サイド情報

【数 2 0 1】

$$\hat{z}$$

は、特徴変数

【数 2 0 3】

$$\hat{y}$$

をさらに抽出することによって取得される特徴変数

【数 2 0 2】

10

20

30

40

50

$\hat{z}$

として理解されてもよく、

【数 2 0 4】

 $\hat{z}$

に含まれる特徴要素の数量は、特徴変数

【数 2 0 5】

10

 $\hat{y}$

の特徴要素の数量より少ない。

【 0 3 7 6 】

エントロピー符号化が、サイド情報

【数 2 0 6】

 $\hat{z}$

20

に対して実行されてもよく、サイド情報

【数 2 0 7】

 $\hat{z}$

がこのステップにおいてビットストリームに書き込まれ、または、エントロピー符号化が、サイド情報

【数 2 0 8】

 $\hat{z}$

30

に対して実行されてもよく、サイド情報

【数 2 0 9】

 $\hat{z}$

が後続のステップ1804においてビットストリームに書き込まれることが留意されるべきである。これはここで限定されない。

【 0 3 7 7 】

ステップ1803: 特徴変数

40

【数 2 1 0】

 $\hat{y}$

に対して確率推定を実行し、特徴要素の確率推定結果を取得する。

【 0 3 7 8 】

確率分布モデルは、確率推定結果および確率分布を取得するために使用され得る。確率分布モデルは、単一ガウスモデル(Gaussian single model, GSM)、非対称ガウスモデル、混合ガウスモデル(Gaussian mix model, GMM)、またはラプラス分布(Laplace distribution)モデルであり得る。

50

【 0 3 7 9 】

以下は、説明のために特徴変数

【数 2 1 1 】

$\hat{y}$

がN*M次元の行列である例を使用する。現在の符号化されることになる特徴変数

【数 2 1 2 】

$\hat{y}$

10

の特徴要素は、

【数 2 1 3 】

$\hat{y}[j][i]$

と表記され、ここでj [0, N-1]かつi [0, M-1]である。

【 0 3 8 0 】

確率分布モデルがガウスモデル(単一ガウスモデル、非対称ガウスモデル、または混合ガウスモデル)であるとき、まず、サイド情報

20

【数 2 1 4 】

$\hat{z}$

またはコンテキスト情報が確率推定ネットワークに入力され、確率推定が、特徴変数

【数 2 1 6 】

$\hat{y}$

の各特徴要素

30

【数 2 1 5 】

$\hat{y}[j][i]$

に対して実行されて、平均値パラメータ μ および分散 σ の値を取得する。さらに、平均値パラメータ μ および分散 σ は、確率分布を取得するために、使用される確率分布モデルに入力される。この場合、確率推定結果は、平均値パラメータ μ および分散 σ を含む。

【 0 3 8 1 】

代替として、分散は推定されてもよい。たとえば、確率分布モデルがガウスモデル(単一ガウスモデル、非対称ガウスモデル、または混合ガウスモデル)であるとき、まず、サイド情報

40

【数 2 1 7 】

$\hat{z}$

またはコンテキスト情報が確率推定ネットワークに入力され、確率推定が、特徴変数

【数 2 1 9 】

$\hat{y}$

50

の各特徴要素

【数 2 1 8】

$\hat{y}[j][i]$

に対して実行されて、分散 σ の値を取得する。さらに、分散 σ は、確率分布を取得するために、使用される確率分布モデルに入力される。この場合、確率推定結果は分散 σ である。

【0 3 8 2】

確率分布モデルがラプラス分布モデルであるとき、まず、サイド情報

【数 2 2 0】

$\hat{z}$

10

またはコンテキスト情報が確率推定ネットワークに入力され、確率推定が、特徴変数

【数 2 2 2】

$\hat{y}$

の各特徴要素

【数 2 2 1】

$\hat{y}[j][i]$

20

に対して実行されて、位置パラメータ μ およびスケールパラメータ b の値を取得する。さらに、位置パラメータ μ およびスケールパラメータ b は、確率分布を取得するために、使用される確率分布モデルに入力される。この場合、確率推定結果は、位置パラメータ μ およびスケールパラメータ b を含む。

【0 3 8 3】

代替として、サイド情報

【数 2 2 3】

$\hat{z}$

30

および / またはコンテキスト情報が確率推定ネットワークに入力されてもよく、確率推定が、符号化されることになる特徴マップ

【数 2 2 5】

$\hat{y}$

の各特徴要素

【数 2 2 4】

$\hat{y}[j][i]$

40

に対して実行されて、現在の符号化されることになる特徴要素

【数 2 2 6】

$\hat{y}[j][i]$

の確率分布を取得する。現在の符号化されることになる特徴要素

50

【数 2 2 7】

 $\hat{y}[j][i]$

の値がmである確率Pは、確率分布に基づいて取得される。この場合、確率推定結果は、現在の符号化されることになる特徴要素

【数 2 2 8】

 $\hat{y}[j][i]$

10

の値がmである確率Pである。

【0 3 8 4】

確率推定ネットワークは、深層学習に基づくネットワーク、たとえば、リカレントニューラルネットワーク(recurrent neural network, RNN)および畳み込みニューラルネットワーク(convolutional neural network, CNN)を使用し得る。これはここで限定されない。

【0 3 8 5】

ステップ1804: 確率推定結果に基づいて、現在の符号化されることになる特徴要素に対してエントロピー符号化が実行される必要があるかどうかを決定し、決定結果に基づいて、エントロピー符号化を実行し、現在の符号化されることになる特徴要素を圧縮されたビットストリーム(符号化されたビットストリーム)に書き込み、またはエントロピー符号化を実行することをスキップする。

20

【0 3 8 6】

以下の方法の1つまたは複数は、確率推定結果に基づいて、エントロピー符号化が現在の符号化されることになる特徴要素

【数 2 2 9】

 $\hat{y}[j][i]$

に対して実行される必要があるかどうかを決定するために使用され得る。パラメータjおよびiは正の整数であり、座標(j, i)は現在の符号化されることになる特徴要素の位置を示す。代替として、以下の方法の1つまたは複数は、確率推定結果に基づいて、エントロピー符号化が現在の符号化されることになる特徴要素

30

【数 2 3 0】

 $\hat{y}[i]$

に対して実行される必要があるかどうかを決定するために使用され得る。パラメータiは正の整数であり、座標iは現在の符号化されることになる特徴要素の位置を示す。

【0 3 8 7】

エントロピー符号化が現在の符号化されることになる特徴要素

40

【数 2 3 1】

 $\hat{y}[j][i]$

に対して実行される必要があるかどうかは確率推定結果に基づいて決定される例が、説明のために以下で使用される。エントロピー符号化が現在の符号化されることになる特徴要素

【数 2 3 2】

 $\hat{y}[i]$

50

に対して実行される必要があるかどうかを決定するための方法は類似する。詳細はここで再び説明されない。

【 0 3 8 8 】

方法1: 確率分布モデルがガウス分布であるとき、現在の符号化されることになる特徴要素に対してエントロピー符号化を実行するかどうかは、第1の特徴要素の確率推定結果に基づいて決定される。現在の符号化されることになる特徴要素のガウス分布の平均値パラメータ μ および分散 σ の値が第2の条件を満たす、すなわち、平均値 μ と k との間の差の絶対値は第2の閾値 $T1$ より小さく、かつ分散 σ は第3の閾値 $T2$ より小さいとき、エントロピー符号化プロセスは、現在の符号化されることになる特徴要素

【数 2 3 3 】

$\hat{y}[j][i]$

10

に対して実行される必要はない。そうでなければ、第1の条件が満たされる、すなわち、平均値 μ と k との間の差の絶対値は第2の閾値 $T1$ 以上であり、または分散 σ は第3の閾値 $T2$ より小さいとき、エントロピー符号化が、現在の符号化されることになる特徴要素

【数 2 3 4 】

$\hat{y}[j][i]$

20

に対して実行され、現在の符号化されることになる特徴要素

【数 2 3 5 】

$\hat{y}[j][i]$

がビットストリームに書き込まれる。 k は任意の整数、たとえば0、1、-1、2、または3である。 $T2$ の値は、0 $T2$ 1を満たす任意の数、たとえば、0.2、0.3、0.4、または同様のものの値である。 $T1$ は0以上かつ1より小さい数、たとえば0.01、0.02、0.001、および0.002である。

【 0 3 8 9 】

特に、 k の値が0であるとき、それは最適な値である。ガウス分布の平均値パラメータ μ の絶対値が $T1$ より小さく、かつガウス分布の分散 σ が $T2$ より小さいとき、現在の符号化されることになる特徴要素

【数 2 3 6 】

$\hat{y}[j][i]$

30

に対してエントロピー符号化プロセスを実行することがスキップされることが、直接に決定され得る。そうでなければ、エントロピー符号化が、現在の符号化されることになる特徴要素

【数 2 3 7 】

$\hat{y}[j][i]$

40

に対して実行され、現在の符号化されることになる特徴要素

【数 2 3 8 】

$\hat{y}[j][i]$

がビットストリームに書き込まれる。 $T2$ の値は、0 $T2$ 1を満たす任意の数、たとえば

50

、0.2、0.3、0.4、または同様のものの値である。T1は0以上かつ1より小さい数、たとえば0.01、0.02、0.001、および0.002である。

【0 3 9 0】

方法2: 確率分布がガウス分布であるとき、現在の符号化されることになる特徴要素

【数 2 3 9】

$\hat{y}[j][i]$

のガウス分布の平均値パラメータ μ および分散 σ の値は、確率推定結果に基づいて取得される。平均値 μ と、分散 σ と、 k との関係が $\text{abs}(\mu - k) + \sigma \leq T3$ を満たす(第2の条件)とき、現在の符号化されることになる特徴要素

【数 2 4 0】

$\hat{y}[j][i]$

10

に対してエントロピー符号化プロセスを実行することはスキップされ、 $\text{abs}(\mu - k)$ は平均値 μ と k との間の差の絶対値を計算することを表現する。そうでなければ、現在の符号化されることになる特徴要素の確率推定結果が $\text{abs}(\mu - k) + \sigma \leq T3$ を満たす(第1の条件)とき、エントロピー符号化が、現在の符号化されることになる特徴要素

【数 2 4 1】

$\hat{y}[j][i]$

20

に対して実行され、現在の符号化されることになる特徴要素

【数 2 4 2】

$\hat{y}[j][i]$

がビットストリームに書き込まれる。 k は任意の整数、たとえば0、1、-1、-2、または3である。第4の閾値T3は0以上かつ1より小さい数であり、たとえば、値は0.2、0.3、0.4、または同様のものである。

【0 3 9 1】

確率分布がガウス分布であるとき、確率推定が特徴変数

【数 2 4 4】

$\hat{y}$

30

の各特徴要素

【数 2 4 3】

$\hat{y}[j][i]$

40

に対して実行されるならば、現在の符号化されることになる特徴要素

【数 2 4 5】

$\hat{y}[j][i]$

のガウス分布の分散 σ の値のみが取得される。分散 σ が $\sigma \leq T3$ を満たす(第2の条件)とき、現在の符号化されることになる特徴要素

【数 2 4 6】

50

$\hat{y}[j][i]$

に対してエントロピー符号化プロセスを実行することがスキップされる。そうでなければ、現在の符号化されることになる特徴要素の確率推定結果が $\sigma - T3$ を満たす(第1の条件)とき、エントロピー符号化が、現在の符号化されることになる特徴要素

【数 2 4 7】

 $\hat{y}[j][i]$

10

に対して実行され、現在の符号化されることになる特徴要素

【数 2 4 8】

 $\hat{y}[j][i]$

がビットストリームに書き込まれる。第4の閾値 $T3$ は0以上かつ1より小さい数であり、たとえば、値は0.2、0.3、0.4、または同様のものである。

【 0 3 9 2】

方法3: 確率分布がラプラス分布であるとき、現在の符号化されることになる特徴要素

【数 2 4 9】

20

 $\hat{y}[j][i]$

のラプラス分布の位置パラメータ μ およびスケールパラメータ b の値が、確率推定結果に基づいて取得される。位置パラメータ μ と、スケールパラメータ b と、 k との関係が $abs(\mu - k) + \sigma - T4$ を満たす(第2の条件)とき、現在の符号化されることになる特徴要素

【数 2 5 0】

 $\hat{y}[j][i]$

30

に対してエントロピー符号化プロセスを実行することはスキップされ、 $abs(\mu - k)$ は位置パラメータ μ と k との間の差の絶対値を計算することを表現する。そうでなければ、現在の符号化されることになる特徴要素の確率推定結果が $abs(\mu - k) + \sigma - T4$ を満たす(第1の条件)とき、エントロピー符号化が、現在の符号化されることになる特徴要素

【数 2 5 1】

 $\hat{y}[j][i]$

に対して実行され、現在の符号化されることになる特徴要素

【数 2 5 2】

40

 $\hat{y}[j][i]$

がビットストリームに書き込まれる。 k は任意の整数、たとえば0、1、-1、-2、または3である。第4の閾値 $T4$ は0以上かつ0.5より小さい数であり、たとえば、値は0.05、0.09、0.17、または同様のものである。

【 0 3 9 3】

方法4: 確率分布がラプラス分布であるとき、現在の符号化されることになる特徴要素

【数 2 5 3】

50

$\hat{y}[j][i]$

のラプラス分布の位置パラメータ μ およびスケールパラメータ b の値が、確率推定結果に基づいて取得される。位置パラメータ μ と k との間の差の絶対値が第2の閾値 $T5$ より小さく、かつスケールパラメータ b が第3の閾値 $T6$ より小さい(第2の条件)とき、現在の符号化されることになる特徴要素

【数 2 5 4】

 $\hat{y}[j][i]$

10

に対してエントロピー符号化プロセスを実行することがスキップされる。そうでなければ、位置パラメータ μ と k との間の差の絶対値が第2の閾値 $T5$ より小さい、またはスケールパラメータ b が第3の閾値 $T6$ 以上である(第1の条件)とき、エントロピー符号化が、現在の符号化されることになる特徴要素

【数 2 5 5】

 $\hat{y}[j][i]$

に対して実行され、現在の符号化されることになる特徴要素

20

【数 2 5 6】

 $\hat{y}[j][i]$

がビットストリームに書き込まれる。 k は任意の整数、たとえば0、1、-1、-2、または3である。 $T5$ の値は $1e-2$ であり、 $T6$ の値は $T6 - 0.5$ を満たす任意の数、たとえば、0.05、0.09、0.17、または同様のものの値である。

【0 3 9 4】

特に、 k の値が0であるとき、それは最適な値である。位置パラメータ μ の絶対値が $T5$ より小さく、かつスケールパラメータ b が $T6$ より小さいとき、現在の符号化されることになる特徴要素

30

【数 2 5 7】

 $\hat{y}[j][i]$

に対してエントロピー符号化プロセスを実行することがスキップされることが、直接に決定され得る。そうでなければ、エントロピー符号化が、現在の符号化されることになる特徴要素

【数 2 5 8】

 $\hat{y}[j][i]$

40

に対して実行され、現在の符号化されることになる特徴要素

【数 2 5 9】

 $\hat{y}[j][i]$

がビットストリームに書き込まれる。閾値 $T5$ の値は $1e-2$ であり、 $T2$ の値は $T6 - 0.5$ を満たす任意の数、たとえば、0.05、0.09、0.17、または同様のものの値である。

【0 3 9 5】

50

方法5: 確率分布が混合ガウス分布であるとき、現在の符号化されることになる特徴要素
【数 2 6 0】

$\hat{y}[j][i]$

の混合ガウス分布のすべての平均値パラメータ μ_i および分散 σ_i の値が、確率推定結果に基づいて取得される。混合ガウス分布の任意の分散と、混合ガウス分布のすべての平均値と k との間の差の絶対値の和と、の和が第5の閾値 $T7$ より小さい(第2の条件)とき、現在の符号化されることになる特徴要素

【数 2 6 1】

$\hat{y}[j][i]$

10

に対してエントロピー符号化プロセスを実行することがスキップされる。そうでなければ、混合ガウス分布の任意の分散と、混合ガウス分布のすべての平均値と k との間の差の絶対値の和と、の和が第5の閾値 $T7$ 以上である(第1の条件)とき、エントロピー符号化が、現在の符号化されることになる特徴要素

【数 2 6 2】

$\hat{y}[j][i]$

20

に対して実行され、現在の符号化されることになる特徴要素

【数 2 6 3】

$\hat{y}[j][i]$

がビットストリームに書き込まれる。 k は任意の整数、たとえば0、1、-1、-2、または3である。 $T7$ は0以上かつ1より小さい数であり、たとえば、値は0.2、0.3、0.4、または同様のものである(各特徴要素の閾値は同じであると見なされ得る)。

【0 3 9 6】

30

方法6: 現在の符号化されることになる特徴要素

【数 2 6 4】

$\hat{y}[j][i]$

の値が k である確率 P は、確率分布に基づいて取得される。現在の符号化されることになる特徴要素の確率推定結果 P が第2の条件を満たす、すなわち、 P は第1の閾値 $T0$ より大きい(またはそれに等しい)とき、現在の符号化されることになる特徴要素に対してエントロピー符号化プロセスを実行することはスキップされる。そうでなければ、現在の符号化されることになる特徴要素の確率推定結果 P が第1の条件を満たす、すなわち、 P は第1の閾値 $T0$ より小さいとき、エントロピー符号化は、現在の符号化されることになる特徴要素に対して実行され、現在の符号化されることになる特徴要素は、ビットストリームに書き込まれる。 k は任意の整数、たとえば0、1、-1、2、または3であり得る。第1の閾値 $T0$ は、 $0 < T0 < 1$ を満たす任意の数であり、たとえば、値は0.99、0.98、0.97、0.95、または同様のものである(各特徴要素の閾値は同じであると見なされ得る)。

【0 3 9 7】

40

実際の応用では、プラットフォームの一貫性を確保するために、閾値 $T1$ 、 $T2$ 、 $T3$ 、 $T4$ 、 $T5$ 、および $T6$ は丸められ、すなわち整数へとシフトされてスケールされてもよいことが留意されるべきである。

【0 3 9 8】

50

閾値を取得するための方法は、以下の方法の1つを代替として使用してもよいことが留意されるべきである。これはここで限定されない。

【0399】

方法1: 閾値T1が例として使用され、T1の値の範囲内の任意の値が閾値T1として使用され、閾値T1がビットストリームに書き込まれる。具体的には、閾値はビットストリームに書き込まれ、シーケンスヘッダ、ピクチャヘッダ、slice / スライスヘッダ、またはSEIに記憶され、デコーダ側に送信され得る。代替として、別の方法が使用され得る。これはここで限定されない。類似する方法が、残りの閾値T0、T2、T3、T4、T5、およびT6のためにも使用され得る。

【0400】

方法2: エンコーダ側がデコーダ側と合致した固定された閾値を使用し、固定された閾値は、ビットストリームに書き込まれる必要はなく、デコーダ側に送信される必要はない。たとえば、閾値T1が例として使用され、T1の値の範囲内の任意の値がT1の値として直接使用される。類似する方法が、残りの閾値T0、T2、T3、T4、T5、およびT6のためにも使用され得る。

【0401】

方法3: 閾値候補リストが構築され、T1の値の範囲内の最も可能性のある値が、閾値候補リストに入れられる。各閾値は閾値インデックス番号に対応し、最適な閾値が決定され、最適な閾値がT1の値として使用される。最適な閾値のインデックス番号がT1の閾値インデックス番号として使用され、T1の閾値インデックス番号がビットストリームに書き込まれる。具体的には、閾値はビットストリームに書き込まれ、シーケンスヘッダ、ピクチャヘッダ、slice / スライスヘッダ、またはSEIに記憶され、デコーダ側に送信され得る。代替として、別の方法が使用され得る。これはここで限定されない。類似する方法が、残りの閾値T0、T2、T3、T4、T5、およびT6のためにも使用され得る。

【0402】

ステップ1805: エンコーダが圧縮されたビットストリームを送信または記憶する。

【0403】

デコーダ側:

【0404】

ステップ1811: 復号されることになるオーディオ特徴変数のビットストリームを取得する。

【0405】

ステップ1812: 特徴要素の確率推定結果を取得する。

【0406】

エントロピー復号がサイド情報

【数265】

$\hat{z}$

に対して実行されて、サイド情報

【数266】

$\hat{z}$

を取得し、確率推定が、復号されることになるオーディオ特徴変数

【数268】

$\hat{y}$

の各特徴要素

【数 2 6 7】

$\hat{y}[j][i]$

に対して、サイド情報

【数 2 6 9】

$\hat{z}$

を参照して実行されて、現在の復号されることになる特徴要素

10

【数 2 7 0】

$\hat{y}[j][i]$

の確率推定結果を取得する。パラメータ j および i は正の整数であり、座標 (j, i) は現在の復号されることになる特徴要素の位置を示す。代替として、エントロピー復号がサイド情報

【数 2 7 1】

$\hat{z}$

20

に対して実行されて、サイド情報

【数 2 7 2】

$\hat{z}$

を取得し、確率推定が、復号されることになるオーディオ特徴変数

【数 2 7 3】

$\hat{y}$

30

の各特徴要素 $[i]$ に対して、サイド情報

【数 2 7 4】

$\hat{z}$

を参照して実行されて、現在の復号されることになる特徴要素

【数 2 7 5】

$\hat{y}[i]$

40

の確率推定結果を取得する。パラメータ i は正の整数であり、座標 i は現在の復号されることになる特徴要素の位置を示す。

【0 4 0 7】

それに対応して、デコーダ側によって使用される確率推定方法は、この実施形態においてエンコーダ側によって使用されるものと同じであり、デコーダ側によって使用される確率推定ネットワークの構造の図は、この実施形態においてエンコーダ側の確率推定ネットワークのものと同じであることが留意されるべきである。詳細はここで再び説明されない。

【0 4 0 8】

1813: エントロピー復号が現在の復号されることになる特徴要素に対して実行される必要があるかどうかは、確率推定結果に基づいて決定され、エントロピー復号が、決定結

50

果に基づいて実行され、または実行されず、復号された特徴変数

【数 2 7 6】

$\hat{y}$

を取得する。

【0 4 0 9】

以下の方法の1つまたは複数が、確率推定結果に基づいて、エントロピー復号が現在の復号されることになる特徴要素

【数 2 7 7】

$\hat{y}[j][i]$

10

に対して実行される必要があるかどうかを決定するために使用され得る。代替として、以下の方法の1つまたは複数は、確率推定結果に基づいて、エントロピー復号が現在の復号されることになる特徴要素

【数 2 7 8】

$\hat{y}[i]$

20

に対して実行される必要があるかどうかを決定するために使用され得る。

【0 4 1 0】

エントロピー復号が現在の復号されることになる特徴要素

【数 2 7 9】

$\hat{y}[j][i]$

に対して実行される必要があるかどうかは確率推定結果に基づいて決定される例が、説明のために以下で使用される。エントロピー復号が現在の復号されることになる特徴要素

【数 2 8 0】

$\hat{y}[i]$

30

に対して実行される必要があるかどうかを決定するための方法は類似する。詳細はここで再び説明されない。

【0 4 1 1】

方法1: 確率分布モデルがガウス分布であるとき、現在の復号されることになる特徴要素

【数 2 8 1】

$\hat{y}[j][i]$

40

の平均値パラメータ μ および分散 σ の値が、確率推定結果に基づいて取得される。平均値 μ と k との間の差の絶対値が第2の閾値 $T1$ より小さく、かつ分散 σ が第3の閾値 $T2$ より小さい(第2の条件)とき、現在の復号されることになる特徴要素

【数 2 8 2】

$\hat{y}[j][i]$

の数値が k に設定され、現在の復号されることになる特徴要素

【数 2 8 3】

50

$\hat{y}[j][i]$

に対してエントロピー復号プロセスを実行することがスキップされる。そうでなければ、平均値 μ と k との間の差の絶対値が第2の閾値 $T1$ より小さい、または分散 σ が第3の閾値 $T2$ 以上である(第1の条件)とき、エントロピー復号が、現在の復号されることになる特徴要素【数 2 8 4】

 $\hat{y}[j][i]$

10

に対して実行されて、現在の復号されることになる特徴要素【数 2 8 5】

 $\hat{y}[j][i]$

の値を取得する。

【0 4 1 2】

特に、 k の値が0であるとき、それは最適な値である。ガウス分布の平均値パラメータ μ の絶対値が $T1$ より小さく、かつガウス分布の分散 σ が $T2$ より小さいとき、現在の復号されることになる特徴要素

20

【数 2 8 6】

 $\hat{y}[j][i]$

の値が k に設定され、現在の復号されることになる特徴要素

【数 2 8 7】

 $\hat{y}[j][i]$

に対してエントロピー復号プロセスを実行することがスキップされることが直接に決定され得る。そうでなければ、エントロピー復号が、現在の復号されることになる特徴要素

30

【数 2 8 8】

 $\hat{y}[j][i]$

に対して実行されて、現在の復号されることになる特徴要素

【数 2 8 9】

 $\hat{y}[j][i]$

40

の値が取得される。

【0 4 1 3】

方法2: 確率分布がガウス分布であるとき、現在の復号されることになる特徴要素

【数 2 9 0】

 $\hat{y}[j][i]$

の平均値パラメータ μ および分散 σ の値が、確率推定結果に基づいて取得される。平均値 μ と、分散 σ と、 k との関係が $\text{abs}(\mu - k) + \sigma \leq T3$ を満たす(第2の条件)とき、 $T3$ が第4の閾値であり、現在の復号されることになる特徴要素

50

【数 2 9 1】

$\hat{y}[j][i]$

の値がkに設定され、現在の復号されることになる特徴要素

【数 2 9 2】

$\hat{y}[j][i]$

に対してエントロピー復号プロセスを実行することがスキップされる。そうでなければ、現在の復号されることになる特徴要素の確率推定結果が $\text{abs}(\mu - k) + \sigma \leq T3$ を満たす(第1の条件)とき、エントロピー復号が、現在の復号されることになる特徴要素

【数 2 9 3】

$\hat{y}[j][i]$

に対して実行されて、現在の復号されることになる特徴要素

【数 2 9 4】

$\hat{y}[j][i]$

の値を取得する。確率分布がガウス分布であるとき、現在の復号されることになる特徴要素

【数 2 9 5】

$\hat{y}[j][i]$

の分散 σ の値のみが、確率推定結果に基づいて取得される。分散 σ の関係が $\sigma \leq T3$ を満たす(第2の条件)とき、 $T3$ が第4の閾値であり、現在の復号されることになる特徴要素

【数 2 9 6】

$\hat{y}[j][i]$

の値が0に設定され、現在の復号されることになる特徴要素

【数 2 9 7】

$\hat{y}[j][i]$

に対してエントロピー復号プロセスを実行することがスキップされる。そうでなければ、現在の復号されることになる特徴要素の確率推定結果が $\sigma \leq T3$ を満たす(第1の条件)とき、エントロピー復号が、現在の復号されることになる特徴要素

【数 2 9 8】

$\hat{y}[j][i]$

に対して実行されて、現在の復号されることになる特徴要素

【数 2 9 9】

$\hat{y}[j][i]$

の値を取得する。

【 0 4 1 4 】

方法3: 確率分布がラプラス分布であるとき、位置パラメータ μ およびスケールパラメータ b の値は、確率推定結果に基づいて取得される。位置パラメータ μ と、スケールパラメータ b と、 k との間の関係が $abs(\mu - k) + \sigma \leq T4$ を満たす(第2の条件)とき、 $T4$ が第4の閾値であり、現在の復号されることになる特徴要素

【数 3 0 0】

$\hat{y}[j][i]$

の値が k に設定され、特徴要素

【数 3 0 1】

$\hat{y}[j][i]$

10

に対してエントロピー復号プロセスを実行することがスキップされる。そうでなければ、現在の復号されることになる特徴要素の確率推定結果が $abs(\mu - k) + \sigma \leq T4$ を満たす(第1の条件)とき、エントロピー復号が、特徴要素

【数 3 0 2】

$\hat{y}[j][i]$

20

に対して実行されて、特徴要素

【数 3 0 3】

$\hat{y}[j][i]$

の値を取得する。

【 0 4 1 5 】

方法4: 確率分布がラプラス分布であるとき、位置パラメータ μ およびスケールパラメータ b の値は、確率推定結果に基づいて取得される。位置パラメータ μ と k との間の差の絶対値が第2の閾値 $T5$ より小さく、かつスケールパラメータ b が第3の閾値 $T6$ より小さい(第2の条件)とき、現在の復号されることになる特徴要素

【数 3 0 4】

$\hat{y}[j][i]$

30

の値が k に設定され、現在の復号されることになる特徴要素

【数 3 0 5】

$\hat{y}[j][i]$

40

に対してエントロピー復号プロセスを実行することがスキップされる。そうでなければ、位置パラメータ μ と k との間の差の絶対値が第2の閾値 $T5$ より小さい、またはスケールパラメータ b が第3の閾値 $T6$ 以上である(第1の条件)とき、エントロピー復号が、現在の復号されることになる特徴要素

【数 3 0 6】

$\hat{y}[j][i]$

に対して実行されて、現在の復号されることになる特徴要素

50

【数 3 0 7】

$\hat{y}[j][i]$

の値が取得される。

【 0 4 1 6】

特に、 k の値が0であるとき、それは最適な値である。位置パラメータ μ の絶対値が $T5$ より小さく、かつスケールパラメータ b が $T6$ より小さいとき、現在の復号されることになる特徴要素

【数 3 0 8】

$\hat{y}[j][i]$

10

の値が k に設定され、現在の復号されることになる特徴要素

【数 3 0 9】

$\hat{y}[j][i]$

に対してエントロピー復号プロセスを実行することがスキップされることが直接に決定され得る。そうでなければ、エントロピー復号が、現在の復号されることになる特徴要素

【数 3 1 0】

$\hat{y}[j][i]$

20

に対して実行されて、現在の復号されることになる特徴要素

【数 3 1 1】

$\hat{y}[j][i]$

の値が取得される。

【 0 4 1 7】

方法5: 確率分布が混合ガウス分布であるとき、現在の復号されることになる特徴要素

【数 3 1 2】

$\hat{y}[j][i]$

30

の混合ガウス分布のすべての平均値パラメータ μ_i および分散 σ_i の値が、確率推定結果に基づいて取得される。混合ガウス分布の任意の分散と、混合ガウス分布のすべての平均値と k との間の差の絶対値の和と、の和が第5の閾値 $T7$ より小さい(第2の条件)とき、現在の復号されることになる特徴要素

【数 3 1 3】

$\hat{y}[j][i]$

40

の値が k に設定され、現在の復号されることになる特徴要素

【数 3 1 4】

$\hat{y}[j][i]$

に対してエントロピー復号プロセスを実行することがスキップされる。そうでなければ、

50

混合ガウス分布の任意の分散と、混合ガウス分布のすべての平均値と k との間の差の絶対値の和と、の和が第5の閾値 $T7$ 以上である(第1の条件)とき、エントロピー復号が、現在の復号されることになる特徴要素

【数 3 1 5】

$\hat{y}[i]$

に対して実行され、現在の復号されることになる特徴要素

【数 3 1 6】

$\hat{y}[i]$

10

の値が取得される。

【0 4 1 8】

方法6: 現在の復号されることになる特徴要素の値が k である確率 P 、すなわち現在の復号されることになる特徴要素の確率推定結果 P は、現在の復号されることになる特徴要素の確率分布に基づいて取得される。確率推定結果 P が第2の条件を満たす、すなわち、 P は第1の閾値 $T0$ より大きいとき、エントロピー復号は第1の特徴要素に対して実行される必要はなく、現在の復号されることになる特徴要素の値は k に設定される。そうでなければ、現在の復号されることになる特徴要素が第1の条件を満たす、すなわち、 P は第1の閾値 $T0$ 以下であるとき、エントロピー復号がビットストリームに対して実行され、第1の特徴要素の値が取得される。

20

【0 4 1 9】

前述のデコーダ側の値 k は、それに対応してエンコーダ側の値 k に設定される。

【0 4 2 0】

閾値 $T0$ 、 $T1$ 、 $T2$ 、 $T3$ 、 $T4$ 、 $T5$ 、 $T6$ 、および $T7$ を取得するための方法は、エンコーダ側の方法に対応し、以下の方法のうちの1つが使用され得る。

【0 4 2 1】

方法1: 閾値がビットストリームから取得される。具体的には、閾値は、シーケンスヘッダ、ピクチャヘッダ、slice / スライスヘッダ、またはSEIから取得される。

30

【0 4 2 2】

方法2: デコーダ側が、エンコーダ側と合致した固定された閾値を使用する。

【0 4 2 3】

方法3: 閾値インデックス番号がビットストリームから取得される。具体的には、閾値インデックス番号は、シーケンスヘッダ、ピクチャヘッダ、slice / スライスヘッダ、またはSEIから取得される。次いで、デコーダ側が、エンコーダと同じ方式で閾値候補リストを構築し、閾値インデックス番号に基づいて閾値候補リストの中の対応する閾値を取得する。

【0 4 2 4】

実際の応用では、プラットフォームの一貫性を確保するために、閾値 $T1$ 、 $T2$ 、 $T3$ 、 $T4$ 、 $T5$ 、および $T6$ は丸められ、すなわち整数へとシフトされてスケーリングされてもよいことが留意されるべきである。

40

【0 4 2 5】

ステップ1814: 復号された特徴変数

【数 3 1 7】

$\hat{y}$

を再構築し、または対応する機械タスクを実行するために機械向け聴覚タスクモジュールに復号された特徴マップを入力することによって、対応する機械タスクを実行する。このステップは、具体的には、図10Bのデコーダネットワーク306によって実施され得る。詳

50

細については、デコーダネットワーク34の前述の説明を参照されたい。

【0426】

事例1: エントロピー復号を通じて取得される特徴変数

【数318】

$\hat{y}$

がピクチャ再構築モジュールに入力され、ニューラルネットワークが再構築されたオーディオを出力する。ニューラルネットワークは、任意の構造、たとえば、全結合ネットワーク、畳み込みニューラルネットワーク、またはリカレントニューラルネットワークを使用し得る。ニューラルネットワークは、より良い推定効果を達成するために、多層構造の深層ニューラルネットワーク構造を使用し得る。

10

【0427】

事例2: エントロピー復号を通じて取得される特徴変数

【数319】

$\hat{y}$

が機械向け聴覚タスクモジュールに入力されて、対応する機械タスクを実行する。たとえば、オーディオ分類および認識などのマシン聴覚タスクが完了される。

20

【0428】

前述のデコーダ側の値kは、それに対応してエンコーダ側の値kに設定される。

【0429】

図18は、この出願による符号化装置の例示的な構造の概略図である。図18に表されるように、この例における装置はエンコーダ20Aに対応し得る。装置は、取得モジュール2001および符号化モジュール2002を含み得る。取得モジュール2001は、前述の実施形態におけるエンコーダネットワーク204、丸め206(任意選択)、確率推定210、サイド情報抽出214、生成ネットワーク216(任意選択)、およびジョイントネットワーク218(任意選択)を含み得る。符号化モジュール2002は、前述の実施形態における符号化判定実装208を含む。

30

【0430】

取得モジュール2001は、符号化されることになる特徴データを取得するように構成され、符号化されることになる特徴データが複数の特徴要素を含み、複数の特徴要素が第1の特徴要素を含み、第1の特徴要素の確率推定結果を取得するように構成される。符号化モジュール2002は、第1の特徴要素の確率推定結果に基づいて、第1の特徴要素に対してエントロピー符号化を実行するかどうかを決定し、エントロピー符号化が第1の特徴要素に対して実行される必要があると決定されるときのみ、第1の特徴要素に対してエントロピー符号化を実行するように構成される。

【0431】

可能な実装では、特徴データの第1の特徴要素に対してエントロピー符号化を実行するかどうかを決定することは、以下を含む。特徴データの第1の特徴要素の確率推定結果があらかじめ設定された条件を満たすとき、エントロピー符号化が、特徴データの第1の特徴要素に対して実行される必要がある。特徴データの第1の特徴要素の確率推定結果があらかじめ設定された条件を満たさないとき、エントロピー符号化が、特徴データの第1の特徴要素に対して実行される必要はない。

40

【0432】

可能な実装では、符号化モジュールは、特徴データの確率推定結果に基づいて、特徴データの確率推定結果が生成ネットワークに入力されると決定するようにさらに構成され、ネットワークは判定情報を出力する。第1の特徴要素の判定情報の値が1であるとき、特徴データの第1の特徴要素は符号化される必要がある。第1の特徴要素の判定情報の値が1で

50

はないとき、特徴データの第1の特徴要素は符号化される必要はない。

【0433】

可能な実装では、あらかじめ設定された条件は、第1の特徴要素の値が k である確率値が第1の閾値以下であることであり、 k は整数である。

【0434】

可能な実装では、あらかじめ設定された条件は、第1の特徴要素の確率分布の平均値と第1の特徴要素の値 k との間の差の絶対値が第2の閾値以上である、または第1の特徴要素の分散が第3の閾値以上であることであり、 k は整数である。

【0435】

別の可能な実装では、あらかじめ設定された条件は、第1の特徴要素の確率分布の分散と、第1の特徴要素の確率分布の平均値と第1の特徴要素の値 k との間の差の絶対値と、の和が第4の閾値以上であることであり、 k は整数である。

10

【0436】

可能な実装では、第1の特徴要素の値が k である確率値は、第1の特徴要素のすべての可能な値の確率値の中の最大の確率値である。

【0437】

可能な実装では、確率推定が特徴データに対して実行されて、特徴データの特徴要素の確率推定結果を取得する。第1の特徴要素の確率推定結果は、第1の特徴要素の確率値、および/または確率分布の第1のパラメータおよび第2のパラメータを含む。

【0438】

可能な実装では、特徴データの確率推定結果は、第1の特徴要素の判定情報を取得するために生成ネットワークに入力される。第1の特徴要素に対してエントロピー符号化を実行するかどうかは、第1の特徴要素の判定情報に基づいて決定される。

20

【0439】

可能な実装では、特徴データの判定情報が判定マップであり、かつ第1の特徴要素が判定マップにおいて位置する位置に対応する値があらかじめ設定された値であるとき、エントロピー符号化が第1の特徴要素に対して実行される必要があると決定される。第1の特徴要素が判定マップにおいて位置する位置に対応する値があらかじめ設定された値ではないとき、エントロピー符号化が第1の特徴要素に対して実行される必要はないと決定される。

【0440】

可能な実装では、特徴データの判定情報があらかじめ設定された値であるとき、第1の特徴要素に対してエントロピー符号化が実行される必要があると決定される。判定情報があらかじめ設定された値ではないとき、第1の特徴要素に対してエントロピー符号化が実行される必要はないと決定される。可能な実装では、符号化モジュールは、第1の閾値の閾値候補リストを構築し、第1の閾値を第1の閾値の閾値候補リストに入れるようにさらに構成され、第1の閾値に対応するインデックス番号があり、第1の閾値のインデックス番号を符号化されたビットストリームに書き込むようにさらに構成され、第1の閾値の閾値候補リストの長さは T に設定されてもよく、 T は1以上の整数である。

30

【0441】

この実施形態における装置は、図3Aから図3Dに表される方法の実施形態においてエンコードによって実施される技術的解決策において使用され得る。実装の原理およびその技術的な効果は類似している。詳細はここで再び説明されない。

40

【0442】

図19は、この出願による復号装置の例示的な構造の概略図である。図19に表されるように、この例における装置はデコーダ30に対応し得る。装置は、取得モジュール2101および復号モジュール2102を含み得る。取得モジュール2101は、前述の実施形態における確率推定302、生成ネットワーク310(任意選択)、およびジョイントネットワーク312を含み得る。復号モジュール2102は、前述の実施形態における復号判定実装304およびデコーダネットワーク306を含む。

【0443】

50

取得モジュール2101は、復号されることになる特徴データのビットストリームを取得するように構成され、復号されることになる特徴データが複数の特徴要素を含み、複数の特徴要素が第1の特徴要素を含み、第1の特徴要素の確率推定結果を取得するように構成される。復号モジュール2102は、第1の特徴要素の確率推定結果に基づいて、第1の特徴要素に対してエントロピー復号を実行するかどうかを決定し、エントロピー復号が第1の特徴要素に対して実行される必要があると決定されるときのみ、第1の特徴要素に対してエントロピー復号を実行するように構成される。

【0444】

可能な実装では、特徴データの第1の特徴要素に対してエントロピー復号を実行するかどうかを決定することは、以下を含む。特徴データの第1の特徴要素の確率推定結果があらかじめ設定された条件を満たすとき、特徴データの第1の特徴要素が復号される必要がある。代替として、特徴データの第1の特徴要素の確率推定結果があらかじめ設定された条件を満たさないとき、特徴データの第1の特徴要素は復号される必要はなく、第1の特徴要素の特徴値は k に設定され、 k は整数である。

10

【0445】

可能な実装では、復号モジュールは、特徴データの確率推定結果に基づいて、特徴データの確率推定結果が決定ネットワークモジュールに入力されると決定するようにさらに構成され、ネットワークは判定情報を出力する。特徴データの第1の特徴要素は、特徴データの第1の特徴要素に対応する、判定情報の中の位置の値が1であるとき、復号される。特徴データの第1の特徴要素は、特徴データの第1の特徴要素に対応する、判定情報の中の位置の値が1ではないとき、復号されず、第1の特徴要素の特徴値は k に設定され、 k は整数である。

20

【0446】

可能な実装では、あらかじめ設定された条件は、第1の特徴要素の値が k である確率値が第1の閾値以下であることであり、 k は整数である。

【0447】

別の可能な実装では、あらかじめ設定された条件は、第1の特徴要素の確率分布の平均値と第1の特徴要素の値 k との間の差の絶対値が第2の閾値以上である、または第1の特徴要素の確率分布の分散が第3の閾値以上であることである。

【0448】

別の可能な実装では、あらかじめ設定された条件は、第1の特徴要素の確率分布の分散と、第1の特徴要素の確率分布の平均値と第1の特徴要素の値 k との間の差の絶対値と、の和が第4の閾値以上であることである。

30

【0449】

可能な実装では、特徴データの特徴要素の確率推定結果を取得するために、確率推定が特徴データに対して実行される。第1の特徴要素の確率推定結果は、第1の特徴要素の確率値、および/または確率分布の第1のパラメータおよび第2のパラメータを含む。

【0450】

可能な実装では、第1の特徴要素の値が k である確率値は、第1の特徴要素のすべての可能な値の確率値の中の最大の確率値である。

40

【0451】

可能な実装では、第 N の特徴要素の確率推定結果は、以下、すなわち、第 N の特徴要素の確率値、確率分布の第1のパラメータおよび第2のパラメータ、および判定情報の少なくとも1つを含む。特徴データの第1の特徴要素は、特徴データの第1の特徴要素に対応する、判定情報の中の位置の値が1であるとき、復号される。特徴データの第1の特徴要素は、特徴データの第1の特徴要素に対応する、判定情報の中の位置の値が1ではないとき、復号されず、第1の特徴要素の特徴値は k に設定され、 k は整数である。

【0452】

可能な実装では、特徴データの確率推定結果は、第1の特徴要素の判定情報を取得するために生成ネットワークに入力される。第1の特徴要素の判定情報の値があらかじめ設定

50

された値であるとき、第1の特徴要素に対してエントロピー復号が実行される必要があると決定される。第1の特徴要素の判定情報の値があらかじめ設定された値ではないとき、エントロピー復号が第1の特徴要素に対して実行される必要はないと決定され、第1の特徴要素の特徴値はkに設定され、kは整数であり、kは第1の特徴要素の複数の候補値のうちの1つである。

【0453】

可能な実装では、取得モジュールは、第1の閾値の閾値候補リストを構築し、ビットストリームを復号することによって第1の閾値の閾値候補リストのインデックス番号を取得し、第1の閾値の値として、第1の閾値のインデックス番号に対応する、第1の閾値の閾値候補リストの位置の値を使用するようにさらに構成される。第1の閾値の閾値候補リストの長さはTに設定されてもよく、Tは1以上の整数である。

10

【0454】

この実施形態における装置は、図10B、図13B、および図16に表される方法の実施形態においてデコダによって実施される技術的解決策において使用され得る。実装の原理およびその技術的な効果は類似している。詳細はここで再び説明されない。

【0455】

ここで開示され説明される様々な例示の論理ブロック、モジュール、およびアルゴリズムのステップを参照して説明される機能は、ハードウェア、ソフトウェア、ファームウェア、またはそれらの任意の組み合わせによって実装され得ることをこの技術分野の当業者は理解することができる。ソフトウェアによって実装されるならば、例示の論理ブロック、モジュール、およびステップを参照して説明される機能は、1つまたは複数の命令またはコードとしてコンピュータ可読媒体に記憶され、またはコンピュータ可読媒体上で送信され、ハードウェアに基づく処理ユニットによって決定され得る。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対応するコンピュータ可読記憶媒体を含んでもよく、または1つの場所から別の場所への(たとえば、通信プロトコルに従った)コンピュータプログラムの送信を容易にする任意の通信媒体を含んでもよい。このようにして、コンピュータ可読媒体は、一般に、(1)非一時的有形コンピュータ可読記憶媒体、または(2)信号または搬送波などの通信媒体に対応し得る。データ記憶媒体は、この出願において説明される技術を実装するための命令、コード、および/またはデータ構造を取り出すために、1つまたは複数のコンピュータまたは1つまたは複数のプロセッサによってアクセスされることが可能である任意の使用可能な媒体であり得る。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。

20

30

【0456】

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、EEPROM、CD-ROMまたは別の光学ディスク記憶装置、磁気ディスク記憶装置または別の磁気記憶装置、フラッシュメモリ、または、コンピュータによってアクセスされることが可能であり、命令またはデータ構造の形式で要求されるプログラムコードを記憶することができる任意の他の媒体を含み得る。加えて、任意の接続がコンピュータ可読媒体と適切に呼ばれる。たとえば、命令が、ウェブサイト、サーバ、または別の遠隔ソースから、同軸ケーブル、光ファイバ、ツイストペア、デジタル加入者線(DSL)、または、赤外線、無線、またはマイクロ波などの無線技術を通じて送信されるならば、同軸ケーブル、光ファイバ、ツイストペア、DSL、または、赤外線、無線、またはマイクロ波などの無線技術は、媒体の定義に含まれる。しかしながら、コンピュータ可読記憶媒体およびデータ記憶媒体は、接続、搬送波、信号、または他の一時的媒体を含まないが、実際には非一時的有形記憶媒体を意味することが理解されるべきである。この明細書において使用されるディスク(disk)およびディスク(disc)は、コンパクトディスク(CD)、レーザーディスク、光学ディスク、デジタル多用途ディスク(DVD)、およびブルーレイディスクを含む。ディスク(disk)は、通常、データを磁氣的に再生し、ディスク(disc)はレーザーを使用することによってデータを光学的に再生する。上記の組み合わせも、コンピュータ可読媒体の範囲内に含まれるべきである。

40

50

【 0 4 5 7 】

命令は、1つまたは複数のデジタル信号プロセッサ(DSP)、汎用マイクロプロセッサ、特定用途向け集積回路(ASIC)、フィールドプログラマブルゲートアレイ(FPGA)、または等価な集積回路または個別の論理回路などの、1つまたは複数のプロセッサによって実行され得る。したがって、この明細書において使用される用語「プロセッサ」は、前述の構造、またはこの明細書において説明される技術の実装に適用され得る任意の他の構造を指し得る。加えて、いくつかの態様では、この明細書において説明される例示の論理ブロック、モジュール、およびステップを参照して説明される機能は、符号化および復号のために構成される専用ハードウェアおよび/またはソフトウェアモジュール内で提供されてもよく、または組み合わせられたコーデックに組み込まれてもよい。加えて、技術は1つまたは複数の回路または論理要素において完全に実装されてもよい。

10

【 0 4 5 8 】

この出願における技術は、無線ハンドセット、集積回路(IC)、またはICのセット(たとえば、チップセット)を含む、様々な装置またはデバイスにおいて実装され得る。様々なコンポーネント、モジュール、またはユニットは、開示される技法を決定するように構成されるデバイスの機能的な態様を強調するようにこの出願において説明されるが、異なるハードウェアユニットによる実装を必ずしも要求しない。実際には、上記で説明されたように、様々なユニットが、適切なソフトウェアおよび/またはファームウェアと組み合わせてコーデックハードウェアユニットに組み合わせられてもよく、または相互運用可能なハードウェアユニット(上記で説明された1つまたは複数のプロセッサを含む)によって提供されてもよい。

20

【 0 4 5 9 】

前述の説明は、単に、この出願の例示的な特定の実装であるが、この出願の保護範囲を限定することを意図されない。この出願において開示される技術的範囲内でこの技術分野の当業者によって容易に理解されるいずれの変形または置換も、この出願の保護範囲内にあるものである。したがって、この出願の保護範囲は、請求項の保護範囲に従うものである。

【符号の説明】

【 0 4 6 0 】

- 12 ソースデバイス
- 14 デスティネーションデバイス
- 17 ソースデータ
- 18 プリプロセッサ
- 19 前処理されたソースデータ
- 20 エンコーダネットワーク
- 21 符号化されることになる特徴データ
- 22 プリプロセッサ
- 23 前処理された符号化されることになる特徴データ
- 24 エントロピー復号
- 25 符号化されたビットストリーム
- 26 通信インターフェース
- 27 通信チャネル
- 28 通信インターフェース
- 29 復号されたビットストリーム
- 30 エントロピー復号
- 31 復号された特徴データ
- 32 ポストプロセッサ
- 33 後処理された復号された特徴データ
- 34 デコーダネットワーク
- 35 再構築されたデータ

30

40

50

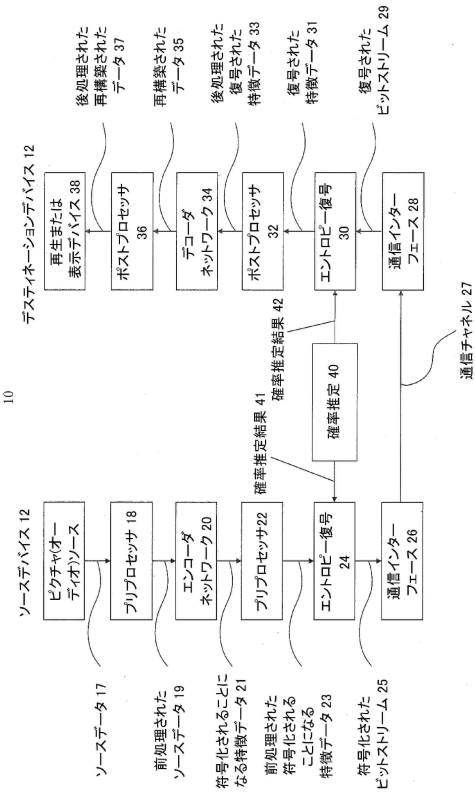
36	ポストプロセッサ	
37	後処理された再構築されたデータ	
38	表示デバイス	
40	確率推定	
41	確率推定結果	
42	確率推定結果	
44	メモリ記憶	
50	コーディングシステム	
51	撮像デバイス	
52	アンテナ	10
53	プロセッサ	
55	表示デバイス	
56	処理ユニット	
202	入力	
204	エンコーダネットワーク	
206	丸め	
208	符号化判定実装	
210	確率推定	
212	出力	
214	サイド情報抽出	20
216	生成ネットワーク	
218	ジョイントネットワーク	
261	符号化要素決定	
262	エントロピー符号化	
301	復号要素決定	
302	エントロピー復号	
306	デコーダネットワーク	
310	生成ネットワーク	
400	コーディングデバイス	
410	入口ポート	30
420	受信機ユニット	
430	プロセッサ	
440	送信機ユニット	
450	出口ポート	
460	メモリ	
470	コーディングモジュール	
502	プロセッサ	
504	メモリ	
506	データ	
508	オペレーティングシステム	40
510	アプリケーションプログラム	
512	バス	
518	ディスプレイ	
1801	取り込みデバイス	
1802	取り込み前処理	
1803	ビデオ符号化	
1804	送信	
1805	受信	
1806	ビットストリーム復号	
1807	レンダリングおよび表示	50

- 1901 特徴抽出
- 1902 サイド情報抽出
- 1903 確率推定
- 1904 符号化
- 1905 復号
- 1906 マシンビジョン(聴覚)タスク
- 2001 取得モジュール
- 2002 符号化モジュール
- 2101 取得モジュール
- 2102 復号モジュール

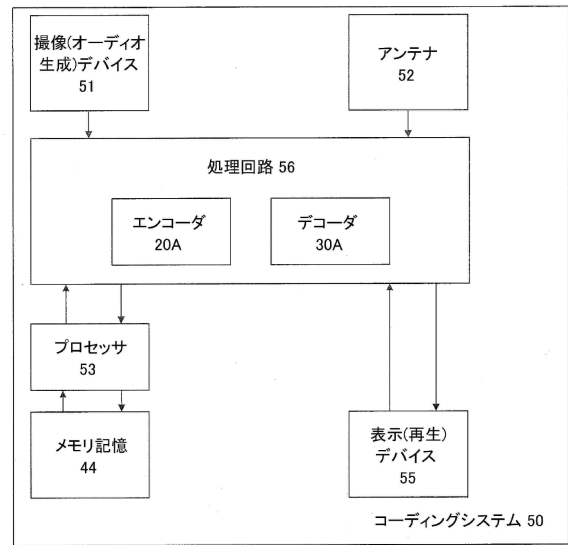
10

【図面】

【図 1 A】



【図 1 B】



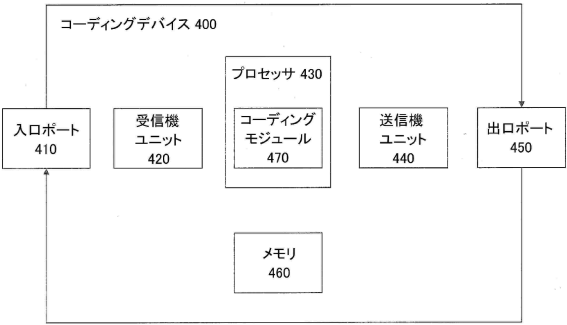
20

30

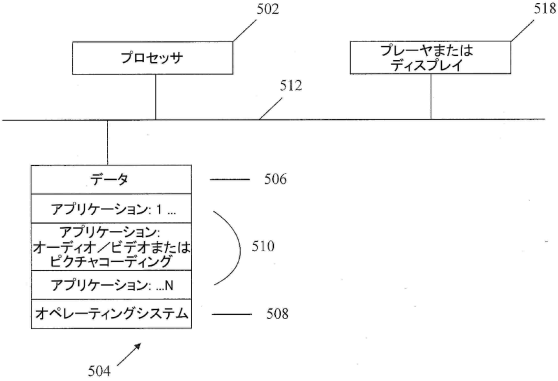
40

50

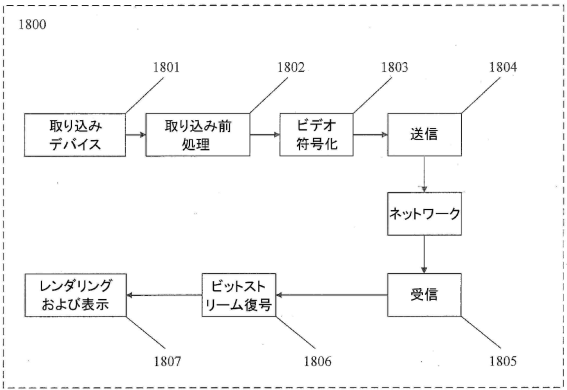
【図 1 C】



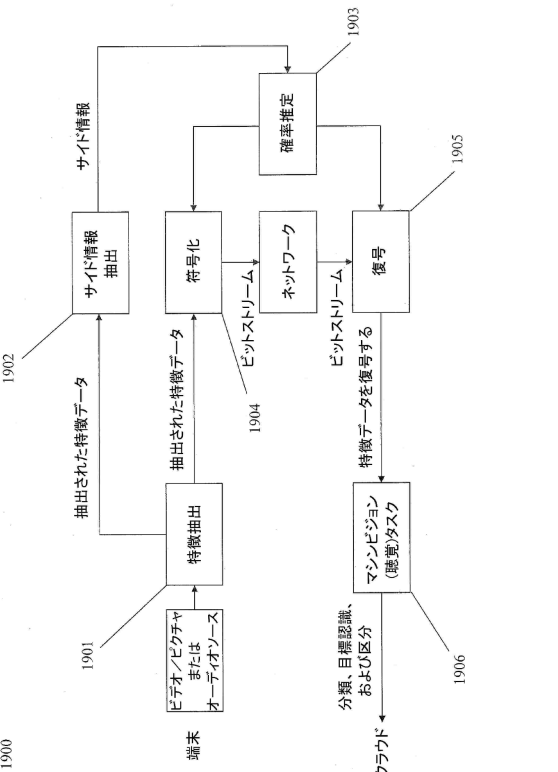
【図 1 D】



【図 2 A】



【図 2 B】



10

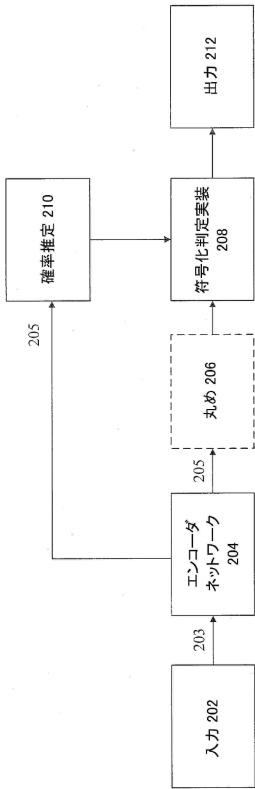
20

30

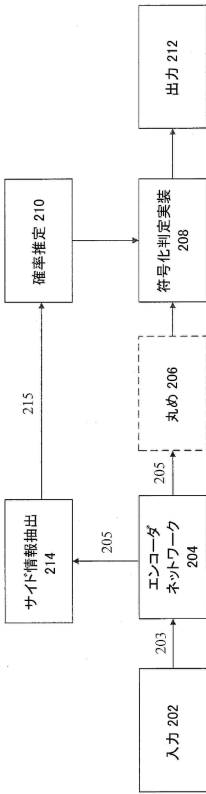
40

50

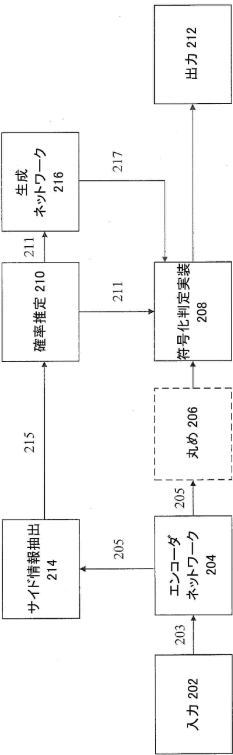
【図 3 A】



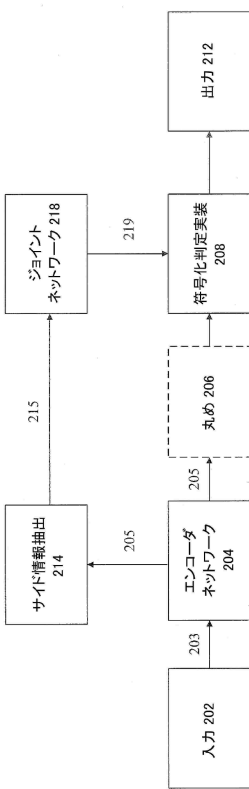
【図 3 B】



【図 3 C】



【図 3 D】



10

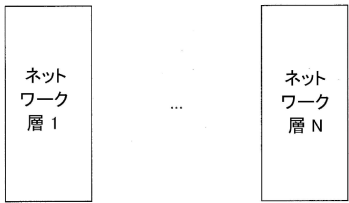
20

30

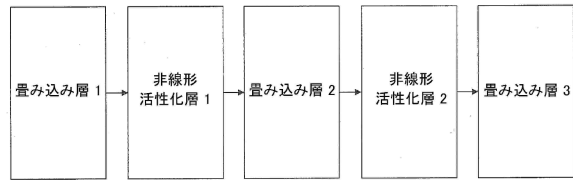
40

50

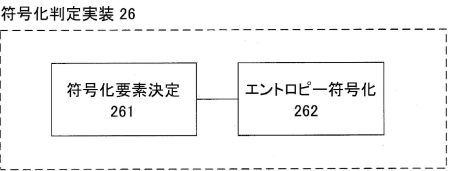
【図 4 A】



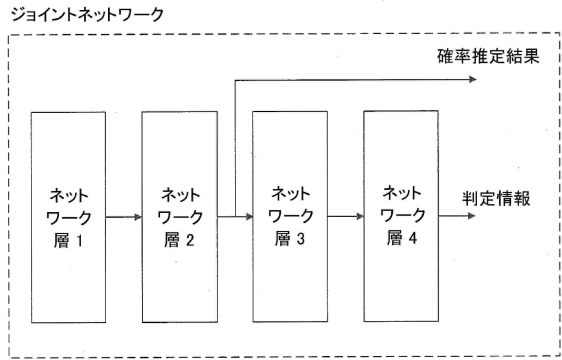
【図 4 B】



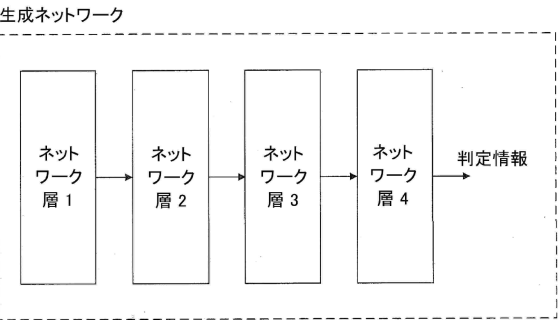
【図 5】



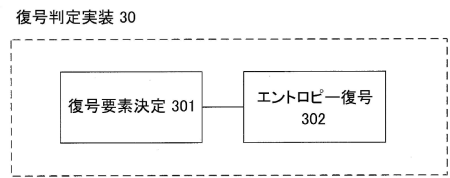
【図 6】



【図 7】



【図 8】



10

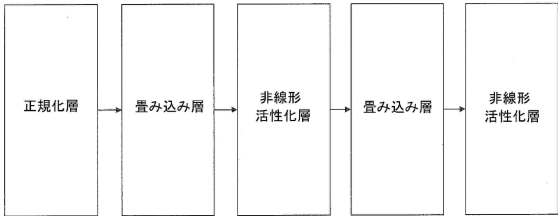
20

30

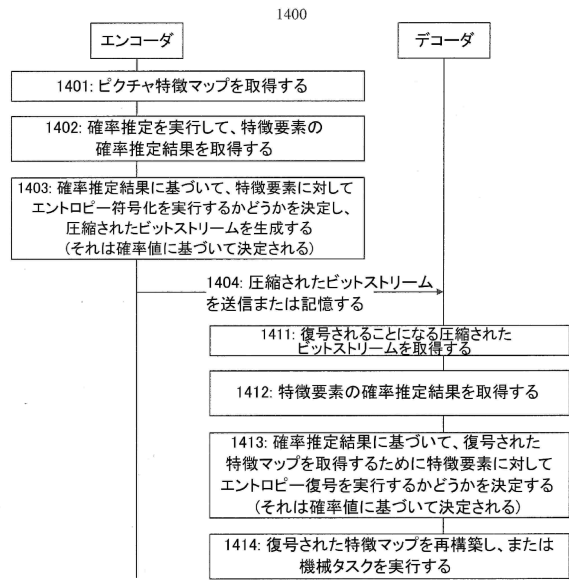
40

50

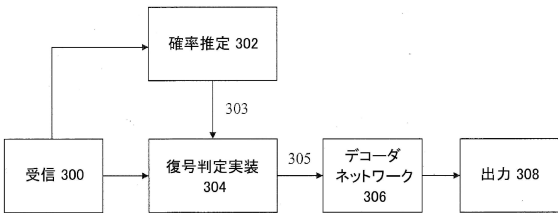
【図 9】



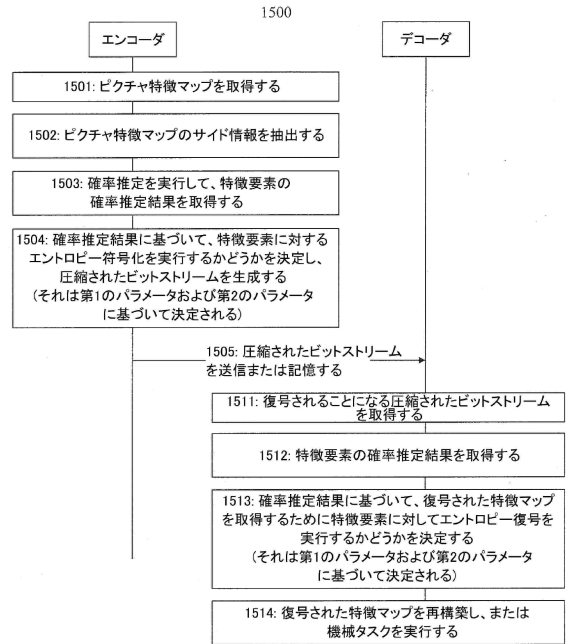
【図 10 A】



【図 10 B】



【図 11 A】



10

20

30

40

50

【図 1 2】

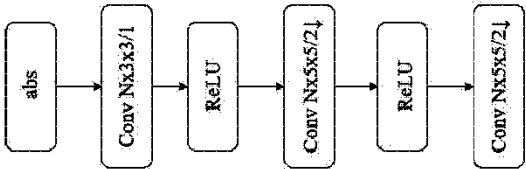
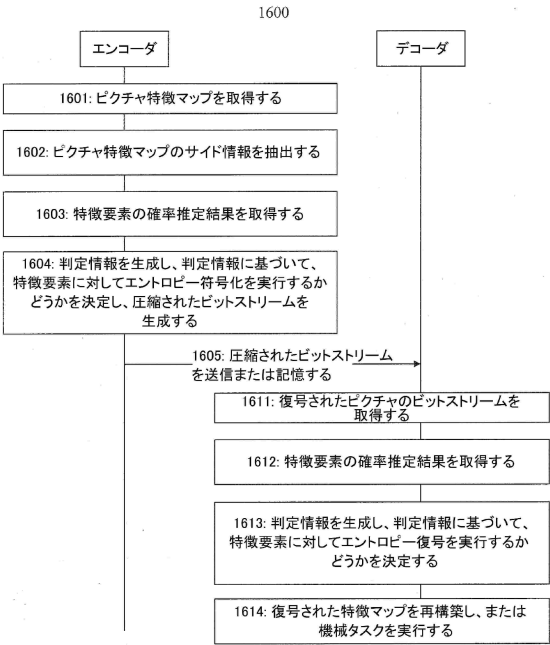
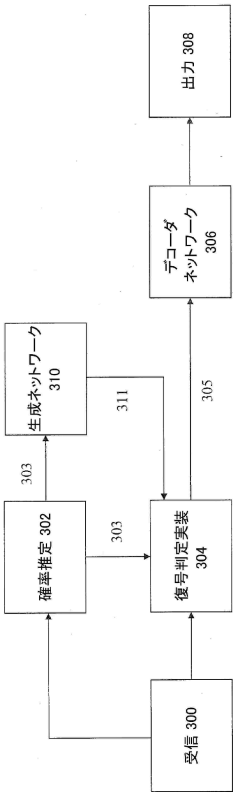


図 12

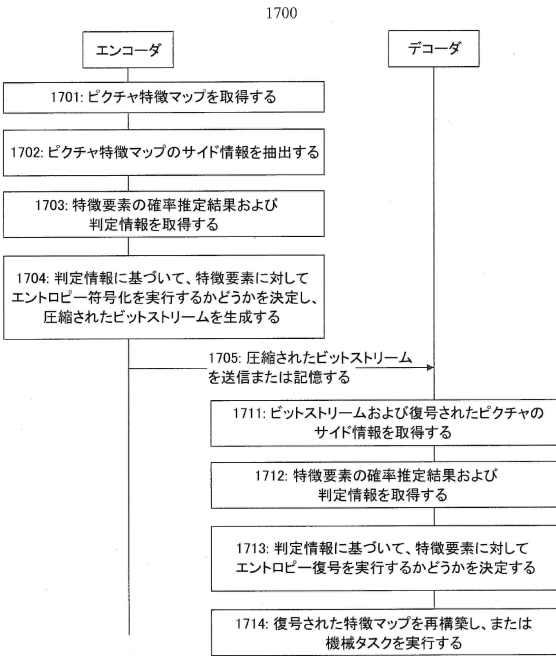
【図 1 3 A】



【図 1 3 B】



【図 1 4】



10

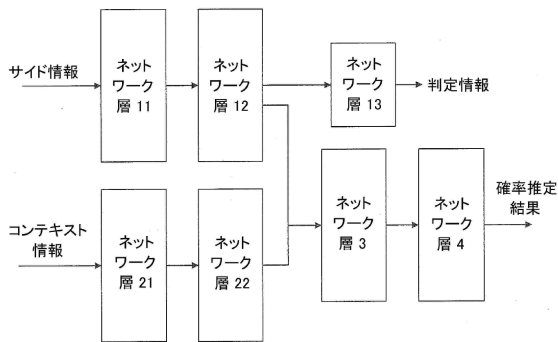
20

30

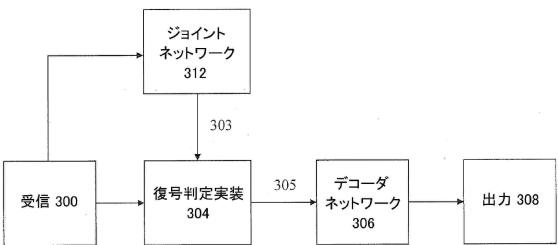
40

50

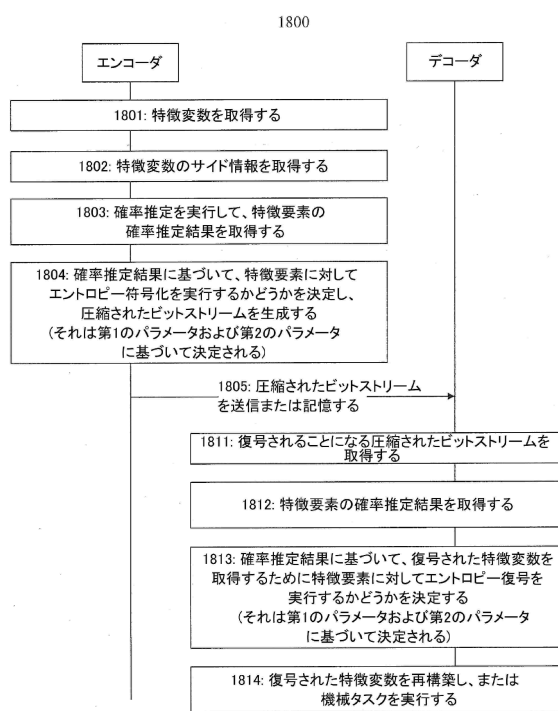
【図 1 5】



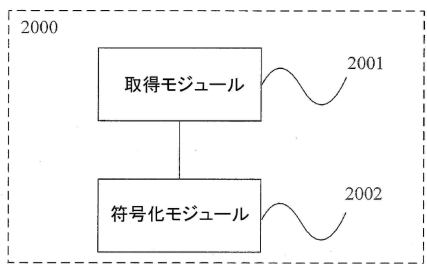
【図 1 6】



【図 1 7】



【図 1 8】



10

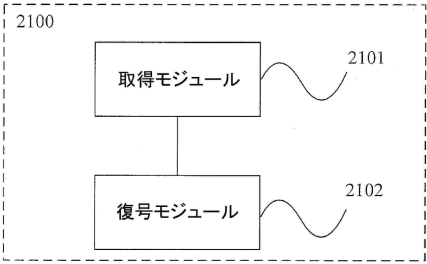
20

30

40

50

【図 19】



10

20

30

40

50

フロントページの続き

(33)優先権主張国・地域又は機関

中国(CN)

(31)優先権主張番号 202111091143.4

(32)優先日 令和3年9月17日(2021.9.17)

(33)優先権主張国・地域又は機関

中国(CN)

(74)代理人 100133569

弁理士 野村 進

(72)発明者 毛 ▲ジュエ▼

中華人民共和国 5 1 8 1 2 9 広東省深▲チェン▼市龍崗区坂田 華為総部▲ベン▼公楼

(72)発明者 ▲趙▼ 寅

中華人民共和国 5 1 8 1 2 9 広東省深▲チェン▼市龍崗区坂田 華為総部▲ベン▼公楼

(72)発明者 ▲ヤン▼ ▲寧▼

中華人民共和国 5 1 8 1 2 9 広東省深▲チェン▼市龍崗区坂田 華為総部▲ベン▼公楼

(72)発明者 ▲楊▼ ▲海▼ 涛

中華人民共和国 5 1 8 1 2 9 広東省深▲チェン▼市龍崗区坂田 華為総部▲ベン▼公楼

(72)発明者 ▲張▼ 恋

中華人民共和国 5 1 8 1 2 9 広東省深▲チェン▼市龍崗区坂田 華為総部▲ベン▼公楼

(72)発明者 王 晶

中華人民共和国 5 1 8 1 2 9 広東省深▲チェン▼市龍崗区坂田 華為総部▲ベン▼公楼

(72)発明者 ▲師▼ 一博

中華人民共和国 5 1 8 1 2 9 広東省深▲チェン▼市龍崗区坂田 華為総部▲ベン▼公楼

審査官 岩井 健二

(56)参考文献 特開2020-191631(JP,A)

特開2020-191630(JP,A)

特開2020-191077(JP,A)

特開2020-173782(JP,A)

米国特許出願公開第2021/0004992(US,A1)

米国特許第10652581(US,B1)

(58)調査した分野 (Int.Cl., DB名)

H04N 19/00 - 19/98