

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7511678号
(P7511678)

(45)発行日 令和6年7月5日(2024.7.5)

(24)登録日 令和6年6月27日(2024.6.27)

(51)国際特許分類	F I
H 0 4 N 19/117 (2014.01)	H 0 4 N 19/117
H 0 4 N 19/136 (2014.01)	H 0 4 N 19/136
H 0 4 N 19/186 (2014.01)	H 0 4 N 19/186
H 0 4 N 19/70 (2014.01)	H 0 4 N 19/70

請求項の数 23 (全43頁)

(21)出願番号	特願2022-575320(P2022-575320)	(73)特許権者	503433420 華為技術有限公司 HUAWEI TECHNOLOGIES CO., LTD. 中華人民共和國 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 公樓 Huawei Administrat ion Building, Banti an, Longgang Distri ct, Shenzhen, Guang dong 5 1 8 1 2 9, P. R. C hina
(86)(22)出願日	令和3年4月20日(2021.4.20)	(74)代理人	100107766 弁理士 伊東 忠重
(65)公表番号	特表2023-528641(P2023-528641 A)		
(43)公表日	令和5年7月5日(2023.7.5)		
(86)国際出願番号	PCT/EP2021/060211		
(87)国際公開番号	WO2021/249684		
(87)国際公開日	令和3年12月16日(2021.12.16)		
審査請求日	令和4年12月7日(2022.12.7)		
(31)優先権主張番号	PCT/EP2020/066048		
(32)優先日	令和2年6月10日(2020.6.10)		
(33)優先権主張国・地域又は機関	欧州特許庁(EP)		

最終頁に続く

(54)【発明の名称】 チャネル間関連情報を用いた適応的画像向上

(57)【特許請求の範囲】

【請求項1】

2つ以上の画像チャネルによって表される画像領域を修正する方法であって、当該方法は：

前記2つ以上の画像チャネルのうちの1つを一次チャネルとして、前記2つ以上の画像チャネルのうちの別の少なくとも1つを二次チャネルとして選択するステップと；

前記一次チャネルを第一のニューラル・ネットワークで処理して、修正された一次チャネルを得るステップと；

前記二次チャネルを第二のニューラル・ネットワークで処理して修正された二次チャネルを得るステップであって、前記第二のニューラル・ネットワークによる処理は前記修正された一次チャネルに基づく、ステップと；

前記修正された一次チャネルおよび前記修正された二次チャネルに基づいて、修正された画像領域を得るステップとを含み、

当該方法はさらに：

前記画像領域の前記少なくとも2つの画像チャネルの各画像チャネルのピクセルを、複数S個の部分領域に再配置するステップを含み、

・前記少なくとも2つの画像チャネルのうちのある画像チャネルの各部分領域が、前記画像チャネルのサンプルのサブセットを含み、

・すべての画像チャネルについて、部分領域の水平寸法は、同じであり、すべての画像チャネルにわたる前記画像領域の水平寸法の最大公約数の整数倍mhに等しく、

・すべての画像チャネルについて、部分領域の垂直寸法は、同じであり、すべての画像チャネルにわたる前記画像領域の垂直寸法の最大公約数の整数倍 mv に等しい、方法。

【請求項 2】

前記2つ以上の画像チャネルのうちから前記一次チャネルおよび前記二次チャネルを選択するステップは、前記2つ以上の画像チャネルが入力されるニューラル・ネットワークに基づく分類器の出力に基づいて実行される、請求項 1 に記載の方法。

【請求項 3】

前記2つ以上の画像チャネルは、色チャネルおよび/または特徴チャネルを含む、請求項 1 または 2 に記載の方法。

【請求項 4】

・前記画像領域は、画像の一部または複数の画像の一部に対応する所定のサイズのパッチ、または
 ・画像もしくは複数の画像
 のうちの1つである、請求項 1 ないし 3 のうちいずれか一項に記載の方法。

【請求項 5】

前記ニューラル・ネットワークの隠れ層の数に基づいて前記画像領域の最小サイズを選択することを含み、前記最小サイズは、少なくとも $2*((kernel_size - 1)/2*n_layers) + 1$ であり、 $kernel_size$ は、畳み込みニューラル・ネットワークである前記ニューラル・ネットワークのカーネルのサイズであり、 n_layers は、前記ニューラル・ネットワークの層の数である、請求項 1 ないし 4 のうちいずれか一項に記載の方法。

【請求項 6】

前記画像領域の前記 S 個の部分領域は、分離しており、 $S = mh*mv$ であり、水平寸法 $dimh$ および垂直寸法 $dimv$ を有し、

部分領域は、位置 $\{kh*mh + offh, kv*mv + offv\}$ 上の前記画像領域のサンプルを含み、 $kh \in [0, dimh - 1]$ および $kv \in [0, dimv - 1]$ であり、

$offh$ と $offv$ の各組み合わせは、それぞれの部分領域を指定し、 $offk \in [1, mh]$ および $offv \in [1, mv]$ である、

請求項 1 ないし 5 のうちいずれか一項に記載の方法。

【請求項 7】

もとの画像領域を取得するステップと；

取得された画像領域をビットストリームにエンコードするステップと；

エンコードされた画像領域を再構成することによって取得された画像領域を修正するための請求項 1 ないし 6 のうちいずれか一項に記載の前記方法を適用するステップとを含む、画像または画像のビデオ・シーケンスをエンコードするための方法。

【請求項 8】

選択された一次チャネルの指示を前記ビットストリーム中に含めるステップを含む、請求項 7 に記載の、画像または画像のビデオ・シーケンスをエンコードするための方法。

【請求項 9】

複数の画像領域を取得するステップと；

取得された画像領域を修正するための前記方法を、取得された複数の画像領域の画像領域に個々に適用するステップと；

前記複数の画像領域のそれぞれについて、前記ビットストリーム中に：

・取得された画像領域を修正するための前記方法がその画像領域のために適用されないことを示す指示、

・その領域についての選択された一次チャネルの指示
 のうちの少なくとも1つを含めるステップとを含む、

請求項 7 ないし 8 のうちいずれか一項に記載の、画像または画像のビデオ・シーケンスをエンコードするための方法。

【請求項 10】

10

20

30

40

50

取得された画像領域を修正するための前記方法を適用するとき、前記エンコードするステップ(910)に入力された前記取得された画像領域を参照することなく、前記再構成された画像領域に基づいて、前記一次チャンネルと前記二次チャンネルの選択(830,840)が実行される、請求項7に記載の画像または画像のビデオ・シーケンスをエンコードするための方法。

【請求項11】

ビットストリームから画像または画像のビデオ・シーケンスをデコードするための方法であって、

前記ビットストリームから画像領域を再構成するステップと；

前記画像領域を修正するための請求項1ないし6のうちいずれか一項に記載の前記方法を適用するステップとを含む、

ビットストリームから画像または画像のビデオ・シーケンスをデコードするための方法。

【請求項12】

ビットストリームから画像または画像のビデオ・シーケンスをデコードするための方法であって：

前記ビットストリームをパースして；

・取得された画像領域を修正するための前記方法がその画像領域について適用されないことを示す指示、

・その領域のための選択された一次チャンネルの指示、
のうちの少なくとも1つを得るステップと、

前記ビットストリームから画像領域を再構成するステップと、

前記指示が選択された一次チャンネルを示す場合には、指示された一次チャンネルを選択された一次チャンネルとして、請求項1ないし6のうちいずれか一項に記載の方法に従って前記再構成された画像領域を修正するステップとを含む、

画像または画像のビデオ・シーケンスをデコードするための方法。

【請求項13】

前記修正が、ループ内フィルタとして、または後処理フィルタとして適用される、請求項11または12に記載の、画像またはビデオ・シーケンスをデコードするための方法。

【請求項14】

コンピュータ読み取り可能な記憶媒体であって、当該コンピュータ読み取り可能な記憶媒体はプログラム・コードを記憶しており、前記プログラム・コードは、一つまたは複数のプロセッサ上で実行されたときに請求項1ないし6のうちいずれか一項に記載の方法を実行する、コンピュータ読み取り可能な記憶媒体。

【請求項15】

コンピュータ読み取り可能な記憶媒体であって、当該コンピュータ読み取り可能な記憶媒体はプログラム・コードを記憶しており、前記プログラム・コードは、一つまたは複数のプロセッサ上で実行されたときに請求項7ないし10のうちいずれか一項に記載の方法を実行する、コンピュータ読み取り可能な記憶媒体。

【請求項16】

コンピュータ読み取り可能な記憶媒体であって、当該コンピュータ読み取り可能な記憶媒体はプログラム・コードを記憶しており、前記プログラム・コードは、一つまたは複数のプロセッサ上で実行されたときに請求項11ないし13のうちいずれか一項に記載の方法を実行する、コンピュータ読み取り可能な記憶媒体。

【請求項17】

2つ以上の画像チャンネルによって表される画像領域を修正するための装置であって、当該装置は、請求項1ないし6のうちいずれか一項に記載のステップを実行するように構成された回路(1702)によって実装される、装置。

【請求項18】

画像または画像をビデオ・シーケンスをエンコードするためのエンコーダであって、当該エンコーダは：

10

20

30

40

50

もとの画像領域を取得するための入力モジュールと、
 取得された画像領域をビットストリーム中にエンコードするための圧縮モジュールと、
 エンコードされた画像領域を再構成するための再構成モジュールと、
 再構成された画像領域を修正するための、請求項 1 7 に記載の装置とを有する、
 エンコーダ。

【請求項 1 9】

前記圧縮モジュールが、前記選択された一次チャンネルの指示を前記ビットストリーム中にエンコードするように構成されている、請求項 1 8 に記載のエンコーダ。

【請求項 2 0】

前記入力モジュールが：
 複数の画像領域を取得するように構成されており、
 請求項 1 7 に記載の装置が：
 取得された画像領域を修正するための前記方法を、取得された複数の画像領域の画像領域に個々に適用するように構成されており、

前記圧縮モジュールが：

前記複数の画像領域のそれぞれについて、前記ビットストリーム中に：

- ・取得された画像領域を修正するための前記方法がその画像領域のために適用されないことを示す指示、
- ・その領域についての選択された一次チャンネルの指示

のうちの少なくとも1つを含めるように構成されている、
 請求項 1 8 または 1 9 に記載のエンコーダ。

【請求項 2 1】

ビットストリームから画像または画像のビデオ・シーケンスをデコードするためのデコーダであって、当該デコーダは：

前記ビットストリームから画像領域を再構成するための再構成モジュールと；

再構成された画像領域を修正するための、請求項 1 7 に記載の装置とを有する、
 デコーダ。

【請求項 2 2】

当該デコーダが、前記ビットストリームをパースして：

- ・取得された画像領域を修正するための前記方法がその画像領域について適用されないことを示す指示、
- ・その領域のための選択された一次チャンネルの指示、

のうちの少なくとも1つを得るように構成されたパース・ユニットを有しており、

前記指示が選択された一次チャンネルを示す場合には、請求項 1 7 に記載の装置は、指示された一次チャンネルを選択された一次チャンネルとして、請求項 1 ないし 6 のうちいずれか一項に記載の方法に従って、再構成された画像領域を修正するように構成される、
 請求項 2 1 に記載のデコーダ。

【請求項 2 3】

前記修正が、ループ内フィルタ (380) または後処理フィルタ (390) として適用される、請求項 2 1 または 2 2 に記載のデコーダ。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

本開示の実施形態は、概括的には、画像またはビデオ処理の分野に関し、特に、画像またはビデオ修正に関する。

【背景技術】

【0 0 0 2】

ビデオ符号化 (ビデオ・エンコードおよびデコード) は、広範囲のデジタルビデオアプリケーション、たとえば、放送デジタルTV、インターネットおよびモバイルネットワークを通じたビデオ送信、ビデオチャット、ビデオ会議のようなリアルタイムの会話アプリケ

10

20

30

40

50

ーション、DVDおよびブルーレイディスク、ビデオコンテンツ収集および編集システム、並びにセキュリティアプリケーションのビデオカメラに使用される。

【0003】

比較的短いビデオを描写するために必要とされるビデオデータの量でも相当な量になることがあり、データがストリーミングされるか、またはさもなければ限られた帯域幅容量を有する通信ネットワークを通じて通信されるときに困難を生じることがある。よって、ビデオデータは、一般に、現代の電気通信ネットワークを通じて通信される前に圧縮される。また、メモリ資源が制限される可能性があるため、ビデオのサイズは、ビデオが記憶装置に記憶されるときにも問題となる可能性がある。ビデオ圧縮装置は、しばしば、伝送または記憶の前にビデオデータを符号化するためにソースにおいてソフトウェアおよび/またはハードウェアを使用し、それによってデジタルビデオ画像を表すのに必要なデータ量を減少させる。次いで、圧縮されたデータは、宛先にいおいて、ビデオデータをデコードするビデオ圧縮解除装置によって受領される。ネットワーク資源が限られており、より高いビデオ品質に対する要求が絶えず増加しているため、画質にほとんどまたは全く犠牲を払わずに圧縮比を改善する、改良された圧縮および圧縮解除技術が望ましい。

10

【0004】

一般に、画像圧縮は、可逆的または不可逆的でありうる。可逆的画像圧縮では、もとの画像は圧縮された画像から完全に再構成できる。しかしながら、圧縮率はかなり低い。対照的に、不可逆画像圧縮は、高い圧縮率を許容するが、もとの画像を完全に再構成することができないという欠点がある。特に、低ビットレートで使用される場合、不可逆画像圧縮は、目に見える空間圧縮アーチファクトを導入する。

20

【発明の概要】

【発明が解決しようとする課題】

【0005】

本発明は、画像またはビデオを修正する、たとえば向上させる方法および装置に関する。

【課題を解決するための手段】

【0006】

本発明は、独立請求項の範囲によって定義される。有利な実施形態のいくつかは、従属特許請求の範囲において提供される。

【0007】

特に、本発明の実施形態は、複数の画像チャネルを処理するニューラルネットワーク・システムに基づく、画像を修正するためのアプローチを提供する。一次チャネルは個別に処理される。二次チャネル（単数または複数）は、処理された一次チャネルを考慮に入れて処理される。ニューラルネットワーク・システムによる処理の前に、どの画像チャネルが一次チャネルであるかが選択される。

30

【0008】

ある側面によれば、2つ以上の画像チャネルによって表される画像領域を修正する方法が提供される。本方法は：前記2つ以上の画像チャネルのうちの1つを一次チャネルとして、前記2つ以上の画像チャネルのうちの別の少なくとも1つを二次チャネルとして選択するステップと、前記一次チャネルを第一のニューラル・ネットワークで処理して、修正された一次チャネルを得るステップと、前記二次チャネルを第二のニューラル・ネットワークで処理して修正された二次チャネルを得るステップであって、前記第二のニューラル・ネットワークによる処理は前記修正された一次チャネルに基づく、ステップと、前記修正された一次チャネルおよび前記修正された二次チャネルに基づいて、修正された画像領域を得るステップとを含む。

40

【0009】

この方法の利点の一つは、画像チャネルのうちから一次チャネルを選択することにより一次チャネルを適応させる可能性に起因する、画像修正性能の改善でありうる。

【0010】

たとえば、前記2つ以上の画像チャネルのうちから一次チャネルおよび二次チャネルを

50

選択するステップは、前記2つ以上の画像チャンネルが入力されるニューラル・ネットワークに基づく分類器の出力に基づいて実行される。分類器を使用することにより、画像修正（たとえば画像向上）の品質が改善されうるように、一次チャンネルとなる画像チャンネルを適切に選択するために、そのような分類器をトレーニングまたは設計することが可能となる。

【0011】

2つ以上の画像チャンネルは、色チャンネルおよび/または特徴チャンネルを含んでいてもよい。色チャンネルと特徴チャンネルは、画像特性を反映する。各種類のチャンネルは、他のチャンネルには存在しない情報を提供することができ、そのため、協働処理が、そのようなチャンネルを、一次チャンネルに対して改善することができる。

10

【0012】

いくつかの実施形態によれば、画像領域は、画像の一部または複数の画像の一部に対応する所定のサイズのパッチ、または画像もしくは複数の画像のうちの1つである。

【0013】

パッチ・ベースまたは複数画像ベースで画像を処理することにより、画像またはビデオ・シーケンスの諸領域を異なる仕方で処理すること、すなわち、一次チャンネルの選択を変更することが可能になる。画像内および/またはビデオ・シーケンス内のコンテンツは変化しうるので、画像向上が一次チャンネルを適応することが有利でありうる。

【0014】

いくつかの例示的実装では、本方法は、ニューラル・ネットワークの隠れ層の数に基づいて画像領域の最小サイズを選択することをさらに含み、最小サイズは、少なくとも $2*((kernel_size - 1)/2 * n_layers) + 1$ であり、 $kernel_size$ は、畳み込みニューラル・ネットワークであるニューラル・ネットワークのカーネルのサイズであり、 n_layers は、ニューラル・ネットワークの層の数である。

20

【0015】

パッチ・サイズを選択のためのそのような下限は、ニューラル・ネットワークの設計に依存して、パディング等による冗長性を追加することなく、処理される画像の情報を完全に利用することを可能にする。

【0016】

ある実施形態（任意の前述または後述の実施形態および例と組み合わせることができる）によれば、本方法は、画像領域の少なくとも2つの画像チャンネルの各画像チャンネルのピクセルを、複数 S 個の部分領域に再配置することを含み、前記少なくとも2つの画像チャンネルのうちのある画像チャンネルの各部分領域が、前記画像チャンネルのサンプルのサブセットを含み、すべての画像チャンネルについて、部分領域の水平寸法は、同じであり、画像の水平寸法の最大公約数の整数倍 mh に等しく、すべての画像チャンネルについて、部分領域の垂直寸法は、同じであり、画像の垂直寸法の最大公約数の整数倍 mv に等しい。

30

【0017】

そのような再構成では、ニューラル・ネットワークは、画像チャンネルの寸法/解像度が異なる諸画像を処理するために使用されうる。

【0018】

特に、画像領域の S 個の部分領域は、分離しており、 $S = mh * mv$ であり、水平寸法 $dimh$ および垂直寸法 $dimv$ を有し、部分領域は、位置 $\{kh * mh + offh, kv * mv + offv\}$ 上の画像領域のサンプルを含み、 $kh \in [0, dimh - 1]$ および $kv \in [0, dimv - 1]$ であり、 $offh$ と $offv$ の各組み合わせは、それぞれの部分領域を指定し、 $offh \in [1, mh]$ および $offv \in [1, mv]$ である。

40

【0019】

上述のパッチ・サイズの決定により、チャンネルが解像度および/または寸法（垂直および/または水平）において互いに異なる場合であっても、画像を利用し、各チャンネルについて画像の寸法に効果的にパッチ・サイズを適応させることが可能である。

【0020】

50

ある側面によれば、もとの画像領域を取得するステップと、取得された画像領域をビットストリームにエンコードするステップと、上述したような、エンコードされた画像領域を再構成することによって取得された画像領域を修正するステップとを含む、画像または画像のビデオ・シーケンスをエンコードするための方法が提供される。

【0021】

画像またはビデオ符号化において画像修正を採用することにより、デコードされた画像の品質の改善が可能になる。これは、低減されうる歪みという意味での品質である。しかしながら、用途によっては、望ましいと思われうるいくつかの特殊な効果があることがあり、前記修正は、その改善につながる可能性がある（これは、必ずしももとの画像に対する歪みを減少させるものではない）。

10

【0022】

たとえば、エンコードは、選択された一次チャンネルの指示をビットストリームに含めるステップを含んでいてもよい。これは、デコーダ側での、もとの（歪んでいない）画像に対する歪みの点で、可能性としてはよりよい再構成を可能にする。

【0023】

ある例示的実装によれば、さらに、複数の画像領域を取得するステップと、取得された画像領域を修正するための前記方法を、取得された複数の画像領域の画像領域に個々に適用するステップと、前記複数の画像領域のそれぞれについて、ビットストリーム中に：取得された画像領域を修正するための前記方法がその画像領域のために適用されないことを示す指示、またはその領域についての選択された一次チャンネルの指示のうち少なくとも1つを含めるステップとを含む。領域ベースの処理は、画像またはビデオコンテンツへの適応を容易にする。

20

【0024】

得られた画像領域を修正するための方法を適用するとき、エンコード・ステップに入力される得られた画像領域を参照することなく、再構成された画像領域に基づいて、一次チャンネルと二次チャンネルの選択が実行されてもよい。これは、追加のオーバーヘッド（レート要件）を回避する。

【0025】

ある側面によれば、ビットストリームから画像領域を再構成するステップと、上述のような、画像領域を修正するための方法を適用するステップとを含む、ビットストリームから画像または画像のビデオ・シーケンスをデコードするための方法が提供される。

30

【0026】

デコーダ側での画像またはビデオ修正の適用は、デコードされた画像品質を改善することができる。

【0027】

いくつかの実施側面における画像またはビデオ・シーケンスをデコードする方法は、ビットストリームをパースして：取得された画像領域を修正するための前記方法がその画像領域について適用されないことを示す指示、その領域のための選択された一次チャンネルの指示、のうち少なくとも1つを得るステップと、ビットストリームから画像領域を再構成するステップと、および前記指示が選択された一次チャンネルを示す場合には、指示された一次チャンネルを選択された一次チャンネルとして、再構成された画像領域を修正するステップとを含む。

40

【0028】

サイド情報に基づく再構成は、対応するエンコード方法について上述したように、品質の点でより良好な性能を提供する可能性がある。前記修正は、エンコーダおよび/またはデコーダにおいて、ループ内フィルタとして、または後処理フィルタとして適用されうる。

【0029】

さらに、2つ以上の画像チャンネルによって表される画像領域を修正するための装置が提供され、量子化装置は、上述の方法のいずれかによるステップを実行するように構成された回路によって実装される。

50

【0030】

ある側面によれば、画像または画像のビデオ・シーケンスをエンコードするためのエンコーダが提供される。本エンコーダは：もとの画像領域を取得するための入力モジュールと、取得された画像領域をビットストリーム中にエンコードするための圧縮モジュールと、エンコードされた画像領域を再構成するための再構成モジュールと、再構成された画像領域を修正するための請求項15に記載の装置とを有する。

【0031】

ある側面によれば、ビットストリームから画像または画像のビデオ・シーケンスをデコードするためのデコーダが提供される。本デコーダは：ビットストリームから画像領域を再構成するための再構成モジュールと、再構成された画像領域を修正するための、請求項15に記載の装置とを有する。これらの装置の利点は、類似の特徴を有する上述の方法の利点に対応する。

10

【0032】

さらに、上述の処理回路によって実行されるステップに対応する方法も提供される。

【0033】

ある側面によれば、上述の方法を実行するためのプログラム・コードを含むコンピュータ・プロダクトが提供される。コンピュータ・プロダクトは、非一時的媒体上に提供されてもよく、一つまたは複数のプロセッサ上で実行されたときに前記方法のステップを実行する命令を含むことができる。

【0034】

上述の装置は、集積チップ上で具現されてもよい。

20

【0035】

上述の実施形態および例示的実装の任意のものが組み合わせられうる。

【図面の簡単な説明】

【0036】

以下では、本発明の実施形態が、添付の図および図面を参照して、より詳細に説明される。

【図1】ビデオをエンコードするための例示的なエンコード装置を示すブロック図である。

【図2】ビデオをデコードするための例示的なデコード装置を示すブロック図である。

【図3A】画像向上のポスト・フィルタとしての適用を示す、エンコーダに関するブロック図である。

30

【図3B】画像向上のループ・フィルタとしての適用を示す、エンコーダに関するブロック図である。

【図4】畳み込みニューラル・ネットワークによる3色チャネルの協働処理の概略図である。

【図5】ピクセル再配置を示す概略図である。

【図6】単一チャネル・モードでのニューラル・ネットワーク処理を示す概略図である。

【図7】協働（協調）モードでのニューラル・ネットワーク処理を示す概略図である。

【図8】エンコーダ・ベースのサイド情報とは独立した画像向上を含むデコーダでの処理を示すフロー図である。

40

【図9】エンコーダ・ベースのサイド情報を用いた画像向上を含むエンコーダおよびデコーダにおける処理を示すフロー図である。

【図10】チャンネル選択のために適用可能な例示的な分類器を示す概略図である。

【図11】協働チャンネル処理に先立つチャンネル選択を示すブロック図である。

【図12】画像のサイズを考慮したピクセル再配置を示す概略図である。

【図13】可能性としては異なるサイズをもつn個のチャンネルをサポートするマルチチャネル画像向上を示す概略図である。

【図14】いくつかの実施側面を実装するために構成可能なビデオ符号化システムの一例を示すブロック図である。

【図15】いくつかの実施形態を実装するために構成可能なビデオ符号化システムの別の

50

例を示すブロック図である。

【図16】エンコード装置またはデコード装置の例を示すブロック図である。

【図17】エンコード装置またはデコード装置の別の例を示すブロック図である。

【発明を実施するための形態】

【0037】

以下の説明では、本開示の一部を形なし、本発明の実施形態の特定の諸側面または本発明の実施形態が使用されうる特定の諸側面を例として示すために、添付の図面を参照する。本発明の実施形態は、他の側面で使用されてもよく、図に示されていない構造的または論理的变化を含んでいてもよいことが理解される。よって、以下の詳細な説明は、限定する意味で解釈されるべきではなく、本発明の範囲は、添付の特許請求の範囲によって定義される。

10

【0038】

たとえば、記載された方法に関連する開示は、当該方法を実行するように構成された対応する装置またはシステムについても成り立つことがあり、その逆もありうるということが理解される。たとえば、一つまたは複数の具体的な方法ステップが記載されている場合、対応する装置は、記載された一つまたは複数の方法ステップを実行するための一つまたは複数のユニット、たとえば機能ユニット（たとえば、前記一つまたは複数のステップを実行する1つのユニット、または前記複数のステップのうちの一つまたは複数それぞれを実行する複数のユニット）を含んでいてもよい。たとえそのような一つまたは複数のユニットが明示的に記載または図示されていなくてもである。他方、たとえば、一つまたは複数のユニット、たとえば、機能ユニットに基づいて具体的な装置が記載される場合、対応する方法は、前記一つまたは複数のユニットの機能性を実行する1つのステップ（たとえば、前記一つまたは複数のユニットの機能性を実行する1つのステップ、または前記複数のユニットのうちの一つまたは複数のユニットの機能性それぞれを実行する複数のステップ）を含んでいてもよい。たとえそのような一つまたは複数のステップが明示的に記載または図示されていなくてもである。さらに、本明細書に記載されるさまざまな例示的な実施形態および/または側面の特徴は、特に断りのない限り、互いに組み合わせることができることが理解される。

20

【0039】

ビデオ符号化〔ビデオ・コーディング〕は、典型的には、ビデオまたはビデオ・シーケンスをなすピクチャーのシーケンスの処理を指す。用語「ピクチャー」の代わりに、用語「フレーム」または「画像」が、ビデオ符号化の分野では同義語として使用されることがある。ビデオ符号化（または一般に符号化）は、ビデオ・エンコードおよびビデオ・デコードという2つの部分を含む。ビデオ・エンコードは、典型的には、（より効率的な記憶および/または伝送のために）ビデオ・ピクチャーを表現するために必要とされるデータ量を減少させるために、もとのビデオ・ピクチャーを（たとえば圧縮によって）処理することを含む。ビデオ・デコードは、宛先側で実行され、典型的には、ビデオ・ピクチャーを再構成するために、エンコードと比較した逆処理を含む。ビデオ・ピクチャー（または一般にピクチャー）の「符号化」に言及する実施形態は、ビデオ・ピクチャーまたはそれぞれのビデオ・シーケンスの「エンコード」または「デコード」に関連するものと理解される。エンコード部とデコード部の組み合わせは、コーデック（コーディングおよびデコーディング）とも呼ばれる。

30

40

【0040】

可逆ビデオ符号化の場合、もとのビデオ・ピクチャーが再構成されることができる。すなわち、再構成されたビデオ・ピクチャーは、もとのビデオ・ピクチャーと同じ品質を有する（記憶または伝送の間に伝送損失またはその他のデータ損失がないとする）。不可逆ビデオ符号化の場合、ビデオ・ピクチャーを表すデータの量を減らすために、たとえば量子化によるさらなる圧縮が実行され、これはデコードにおいて完全に再構成されることができない。すなわち、再構成されたビデオ・ピクチャーの品質は、もとのビデオ・ピクチャーの品質よりも低い、または悪い。

50

【 0 0 4 1 】

いくつかのビデオ符号化標準は、「不可逆ハイブリッド・ビデオ・コーデック」のグループに属する（すなわち、サンプル領域における空間的および時間的予測と、変換領域において量子化を適用するための2D変換符号化を組み合わせる）。ビデオ・シーケンスの各ピクチャーは、典型的には、一組の重複しないブロックに分割され、符号化は典型的にはブロックレベルで実行される。換言すれば、エンコーダでは、ビデオは、典型的には、ブロック（ビデオ・ブロック）レベルで、エンコードされる。たとえば、空間的（ピクチャー内）予測および/または時間的（ピクチャー間）予測を使用して予測ブロックを生成し、該予測ブロックを現在のブロック（現在処理されている/処理されるべきブロック）から減算して残差ブロックを取得し、該残差ブロックを変換し、変換領域で残差ブロックを量子化して、送信されるべきデータ量を低減する（圧縮）。一方、デコーダでは、エンコーダと比較した逆処理が、エンコードまたは圧縮されたブロックに適用されて、表現のために現在のブロックを再構成する。さらに、エンコーダはデコーダ処理ループを複製し、そのため、両者が同一の予測（たとえば、イントラ予測およびインター予測）を生成し、および/またはその後のブロックを処理、すなわち符号化するための再構成を生成する。

10

【 0 0 4 2 】

ビデオ圧縮は、画像の知覚される品質を低下させることがあり、一般に、圧縮されたビデオの出力品質を改善するために画像向上フィルタが使用されうる。

【 0 0 4 3 】

図1は、以下に説明するように、さらなる図および実施形態を参照して、本開示の技術を実装するように修正または構成されてもよい例示的なビデオ・エンコーダ20の概略ブロック図を示す。図1の例では、ビデオ・エンコーダ20は、入力201（または入力インターフェース201）、残差計算ユニット204、変換処理ユニット206、量子化ユニット208、逆量子化ユニット210、逆変換処理ユニット212、再構成ユニット214、ループ・フィルタ・ユニット220、デコードピクチャーバッファ（decoded picture buffer、DPB）230、モード選択ユニット260、エン트로ピー符号化ユニット270、出力ユニット272（または出力インターフェース272）を含む。モード選択ユニット260は、インター予測ユニット244と、イントラ予測ユニット254と、分割ユニット262とを含んでもよい。インター予測ユニット244は、動き推定ユニットおよび動き補償ユニット（図示せず）を含んでもよい。図2に示されるビデオ・エンコーダ20は、ハイブリッド・ビデオ・エンコーダまたはハイブリッド・ビデオ・コーデックによるビデオ・エンコーダとも呼ばれうる。

20

30

【 0 0 4 4 】

残差計算ユニット204、変換処理ユニット206、量子化ユニット208、モード選択ユニット260はエンコーダ20の順方向信号経路をなすと言われてもよく、逆量子化ユニット210、逆変換処理ユニット212、再構成ユニット214、バッファ216、ループ・フィルタ220、デコードピクチャーバッファ（DPB）230、インター予測ユニット244、およびイントラ予測ユニット254は、ビデオ・エンコーダ20の逆方向信号経路をなすと言われてもよい。ビデオ・エンコーダ20の逆方向信号経路はデコーダ20の信号経路に対応する（図3のビデオ・デコーダ30参照）。また、逆量子化ユニット210、逆変換処理ユニット212、再構成ユニット214、ループ・フィルタ220、デコードピクチャーバッファ（DPB）230、インター予測ユニット244、およびイントラ予測ユニット254は、ビデオ・エンコーダ20の「組み込みデコーダ」をなすとも言われる。

40

【 0 0 4 5 】

エンコーダ20は、たとえば入力201を介してピクチャー17（またはピクチャー・データ17）、たとえばビデオまたはビデオ・シーケンスをなすピクチャーのシーケンスのピクチャーを受領するように構成されてもよい。受領されたピクチャーまたはピクチャー・データは、前処理されたピクチャー19（または前処理されたピクチャー・データ19）であってもよい。簡単のために、以下の説明は、ピクチャー17を参照する。ピクチャー17は、現在ピクチャーまたは符号化されるべきピクチャーと称されてもよい（特に、ビデオ符

50

号化において、現在ピクチャーを他のピクチャー、たとえば、同じビデオ・シーケンス、すなわち、現在ピクチャーをも含むビデオ・シーケンスの、前にエンコードおよび/またはデコードされたピクチャーから区別するために)。

【0046】

(デジタル)ピクチャーは、強度値を有するサンプルの二次元アレイまたはマトリクスである、またはそのようにみなされることができる。アレイ中のサンプルは、ピクセル〔ピクチャー〕(ピクチャー・エレメントの略)またはペルと称されてもよい。アレイまたはピクチャーの水平方向および垂直方向(または軸)におけるサンプル数は、ピクチャーのサイズおよび/または解像度を定義する。色の表現のためには、典型的には、3つの色成分が使用される。すなわち、ピクチャーは、3つのサンプル・アレイで表現されるか、または3つのサンプル・アレイを含むことができる。RGBフォーマットまたは色空間では、ピクチャーは対応する赤、緑、および青のサンプル・アレイを含む。しかしながら、ビデオ符号化においては、各ピクセルは、典型的には、輝度および色度フォーマットまたは色空間、たとえば、Y(時に代わりにLも使用される)によって示される輝度成分ならびにCbおよびCrによって示される2つの色度成分を含むYCbCrで表現される。輝度〔ルミネンス〕(または短縮してルーマ)成分Yは明るさまたはグレーレベル強度(たとえばグレースケール・ピクチャーと同様)を表し、2つの色度〔クロミナンス〕(または短縮してクロマ)成分CbおよびCrは色度または色情報成分を表す。よって、YCbCrフォーマットのピクチャーは、輝度サンプル値(Y)の輝度サンプル・アレイと、色度値(CbおよびCr)の2つの色度サンプル・アレイとを含む。RGBフォーマットのピクチャーはYCbCrフォーマットに変換または翻訳されてもよく、その逆も可能である。このプロセスは色変換または転換としても知られている。ピクチャーがモノクロである場合、ピクチャーは輝度サンプル・アレイのみを含んでいてもよい。よって、ピクチャーは、たとえば、モノクロ・フォーマットでのルーマ・サンプルのアレイ、または4:2:0、4:2:2、および4:4:4カラー・フォーマットでのルーマ・サンプルのアレイおよびクロマ・サンプルの2つの対応するアレイでありうる。

【0047】

ビデオ・エンコーダ20の実施形態は、ピクチャー17を複数の(典型的には重複しない)ピクチャー・ブロック203に分割するように構成されたピクチャー分割ユニット(図1には示されていない)を含んでいてもよい。これらのブロックは、ルートブロック、マクロブロック(H.264/AVC)または符号化ツリー・ブロック(CTB)または符号化ツリー単位(CTU)(H.265/HEVCおよびVVC)とも呼ばれる。ピクチャー分割ユニットは、ビデオ・シーケンスのすべてのピクチャーについて同じブロック・サイズおよび該ブロック・サイズを定義する対応するグリッドを使用し、またはピクチャー、ピクチャーのサブセットまたはグループの間でブロック・サイズを変更して、各ピクチャーを対応するブロックに分割するように構成されてもよい。

【0048】

さらなる実施形態では、ビデオ・エンコーダは、ピクチャー17のブロック203、たとえば、ピクチャー17を形成する1つ、いくつかの、またはすべてのブロックを直接受領するように構成されてもよい。ピクチャー・ブロック203は、現在ピクチャー・ブロックまたは符号化されるべきピクチャー・ブロックとも呼ばれうる。

【0049】

ピクチャー17と同様に、ピクチャー・ブロック203は、ピクチャー17よりも小さい寸法ではあるが、強度値(サンプル値)を有するサンプルの二次元アレイまたはマトリクスである、またはそのようにみなされることができる。言い換えると、ブロック203は、たとえば、1つのサンプル・アレイ(たとえば、モノクロ・ピクチャー17の場合はルーマ・アレイ、カラー・ピクチャーの場合はルーマまたはクロマ・アレイ)または3つのサンプル・アレイ(たとえば、カラー・ピクチャー17の場合はルーマ・アレイおよび2つのクロマ・アレイ)または適用されるカラー・フォーマットに依存して任意の他の数および/または種類のアレイを含んでいてもよい。ブロック203の水平方向および垂直方向(または

10

20

30

40

50

軸)におけるサンプル数は、ブロック203のサイズを定義する。よって、ブロックは、たとえば、サンプルの $M \times N$ (M 列 $\times N$ 行)アレイ、または変換係数の $M \times N$ アレイであってもよい。図1に示されるビデオ・エンコーダ20の実施形態は、ブロック毎にピクチャー17をエンコードするように構成されてもよく、たとえば、エンコードおよび予測はブロック203毎に実行される。

【0050】

図1に示されるビデオ・エンコーダ20の実施形態は、さらに、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ーを分割および/またはエンコードするように構成されてもよく、ピクチャーは、一つまたは複数のスライス(典型的には、重複しない)に分割される、またはかかるスライスを使ってエンコードされることができ、各スライスは一つまたは複数のブロック(たとえば、CTU)を含むことができる。図1に示されるビデオ・エンコーダ20の実施形態は、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・グループ(ビデオ・タイルとも呼ばれる)を用いて、ピクチャーを分割および/またはエンコードするようにさらに構成されてもよく、画像は、一つまたは複数のタイル・グループ(典型的には、重複しない)に分割されてもよく、またはかかるタイル・グループを使ってエンコードされてもよく、各タイル・グループは、たとえば、一つまたは複数のブロック(たとえば、CTU)または一つまたは複数のタイルを含んでいてもよく、各タイルは、たとえば、長方形形状であってもよく、一つまたは複数のブロック(たとえば、CTU)、たとえば、完全なまたは断片的なブロックを含んでいてもよい。

【0051】

残差計算ユニット204は、たとえばサンプル毎に(ピクセル毎に)ピクチャー・ブロック203のサンプル値から予測ブロック265のサンプル値を引くことにより、ピクチャー・ブロック203および予測ブロック265(予測ブロック265についてのさらなる詳細は後述する)に基づいて残差ブロック205(残差205とも称する)を計算して、サンプル領域での残差ブロック205を得るように構成されてもよい。

【0052】

変換処理ユニット206は、残差ブロック205のサンプル値に変換対して、変換、たとえば、離散コサイン変換(DCT)、離散サイン変換(DST)、またはそれらの整数近似などを適用して、変換領域での変換係数207を得るように構成されてもよい。変換係数207は、変換残差係数と呼ばれてもよく、変換領域での残差ブロック205を表す。ビデオ・エンコーダ20(個別には変換処理ユニット206)の実施形態は、たとえば、変換パラメータ、たとえば変換または諸変換のタイプを、直接、またはエントロピー符号化ユニット270を介してエンコードまたは圧縮されて、出力するように構成されてもよく、それにより、たとえば、ビデオ・デコーダ30は、デコードのために該変換パラメータを受領して、使用してもよい。

【0053】

量子化ユニット208は、たとえばスカラー量子化またはベクトル量子化を適用することによって、変換係数207を量子化して量子化係数209を得るように構成されてもよい。量子化された係数209は、量子化された変換係数209または量子化された残差係数209と呼ばれてもよい。量子化プロセスは、変換係数207の一部または全部に関連するビット深さを低減することができる。たとえば、 n ビット変換係数は、量子化の間に、 m ビット変換係数に丸められてもよい。ここで、 n は m より大きい。量子化の程度は、量子化パラメータ(QP)を調整することによって修正されうる。たとえば、スカラー量子化については、より細かいまたはより粗い量子化を達成するために異なるスケールリングが適用されてもよい。より小さい量子化ステップ・サイズはより細かい量子化に対応し、より大きい量子化ステップ・サイズはより粗い量子化に対応する。該当する量子化ステップ・サイズは、量子化パラメータ(QP)によって示されてもよい。量子化パラメータは、たとえば、適用可能な量子化ステップ・サイズのあらかじめ定義された集合へのインデックスであってもよい。たとえば、小さな量子化パラメータは、細かい量子化(小さな量子化ステップ・サイズ)

10

20

30

40

50

に対応してもよく、大きな量子化パラメータは、粗い量子化（大きな量子化ステップ・サイズ）に対応してもよく、またはその逆でもよい。量子化は、量子化ステップ・サイズによる除算を含んでいてもよく、たとえば逆量子化ユニット210、対応するおよび/または逆の脱量子化は、量子化ステップ・サイズによる乗算を含んでいてもよい。量子化は、損失が量子化ステップ・サイズの増加に伴って増加する、損失のある動作である。ビデオ・エンコーダ20（個別には量子化ユニット208）の実施形態は、量子化パラメータ（QP）を、たとえば直接、またはエントロピー符号化ユニット270を介してエンコードされて、出力するように構成されてもよく、それにより、たとえば、ビデオ・デコーダ30は、デコードのために量子化パラメータを受領および適用することができる。

【0054】

逆量子化ユニット210は、たとえば量子化ユニット208と同じ量子化ステップ・サイズに基づいて、または同じ量子化ステップ・サイズを使用して、量子化ユニット208によって適用される量子化方式の逆を適用することによって、量子化された係数に対して量子化ユニット208の逆量子化を適用して、量子化された係数211を得るように構成される。量子化解除された係数211は、量子化解除された残差係数211と呼ばれてもよく、典型的には量子化による損失のため変換係数と同一ではないが、変換係数207に対応する。

【0055】

逆変換処理ユニット212は、変換処理ユニット206によって適用された変換の逆変換、たとえば逆離散コサイン変換（DCT）、逆離散サイン変換（DST）、または他の逆変換を適用して、サンプル領域での再構成された残差ブロック213（または対応する量子化解除された係数213）を得るように構成される。再構成された残差ブロック213は、変換ブロック213と呼ばれてもよい。

【0056】

再構成ユニット214（たとえば、加算器または総和器214）は、変換ブロック213（すなわち、再構成された残差ブロック213）を予測ブロック265に加算して、サンプル領域での再構成されたブロック215を得るように構成される。これはたとえば、再構成された残差ブロック213のサンプル値および予測ブロック265のサンプル値をサンプル毎に加算することによる。

【0057】

ループ・フィルタ・ユニット220（または略して「ループ・フィルタ」220）は、再構成されたブロック215をフィルタリングして、フィルタリングされたブロック221を得るように、または一般に、再構成されたサンプルをフィルタリングして、フィルタリングされたサンプルを得るように構成される。ループ・フィルタ・ユニットは、たとえば、ピクセル遷移を滑らかにするように、または、他の仕方ではビデオ品質を改善するように構成される。ループ・フィルタ・ユニット220は、図1では、ループ内フィルタとして示されているが、他の構成では、ループ・フィルタ・ユニット220は、再構成されたサンプル215をフィルタリングするが、フィルタリングされたサンプルをインター予測による使用のためのバッファに格納しないループ後フィルタとして実装されてもよい。むしろ、フィルタリングされたサンプルは、たとえば、再構成された画像を表示または記憶するための出力される。フィルタリングされたブロック221は、フィルタリングされた再構成されたブロック221と称されてもよい。本開示のいくつかの実施形態は、ループ・フィルタまたはポスト・フィルタの一部として使用されうる画像/ビデオ修正アプローチを提供する。

【0058】

デコードピクチャーバッファ（DPB）230は、ビデオ・エンコーダ20によってビデオデータをエンコードするために、参照ピクチャーまたは一般参照ピクチャー・データを記憶するメモリであってもよい。DPB 230は、同期DRAM（SDRAM）、磁気抵抗RAM（MRAM）、抵抗RAM（RRAM）、または他のタイプのメモリデバイスを含む、ダイナミックランダムアクセスメモリ（DRAM）などの多様なメモリデバイスのいずれかによって形成されうる。デコードピクチャーバッファ（DPB）230は、一つまたは複数のフィルタリングされたブロック221を記憶するように構成されてもよい。デコードピクチャーバッファ

10

20

30

40

50

230は、さらに、同じ現在ピクチャーまたは異なるピクチャー、たとえば以前に再構成されたピクチャーの、他の以前にフィルタリングされたブロック、たとえば以前に再構成されフィルタリングされたブロック221を記憶するように構成されてもよく、たとえばインター予測のために、完全な以前に再構成された、すなわちデコードされたピクチャー（および対応する参照ブロックおよびサンプル）および/または部分的に再構成された現在ピクチャー（および対応する参照ブロックおよびサンプル）を提供してもよい。デコードピクチャーバッファ（DPB）230は、一つまたは複数のフィルタリングされていない再構成されたブロック215、またはたとえば再構成されたブロック215がループ・フィルタ・ユニット220によってフィルタリングされない場合には一般にフィルタリングされていない再構成されたサンプル、または再構成されたブロックもしくはサンプルの任意の他のさらに処理されたバージョンを記憶するように構成されてもよい。

10

【0059】

モード選択ユニット260は、分割ユニット262と、インター予測ユニット244と、イントラ予測ユニット254を有しており、もとのピクチャー・データ、たとえばもとのブロック203（現在ピクチャー17の現在のブロック203）と、再構成されたピクチャー・データ、たとえば、同じ（現在の）ピクチャーの、および/または一つまたは複数の以前にデコードされたピクチャーからの、フィルタリングされたおよび/またはフィルタリングされていない再構成されたサンプルまたはブロックを、たとえばデコードピクチャーバッファ230もしくは他のバッファ（たとえば図示しないラインバッファ）から受領または取得するように構成される。再構成されたピクチャー・データは、予測ブロック265または予測子265を得るために、予測、たとえばインター予測またはイントラ予測のための、参照ピクチャー・データとして使用される。

20

【0060】

モード選択ユニット260は、現在のブロック予測モードのための分割（分割なしを含む）および予測モード（たとえば、イントラまたはインター予測モード）を決定または選択し、残差ブロック205の計算および再構成されたブロック215の再構成のために使用される対応する予測ブロック265を生成するように構成されてもよい。

【0061】

モード選択ユニット260の実施形態は、パーティション分割および予測モードを（たとえば、モード選択ユニット260にサポートされる、またはモード選択ユニット260に利用可能なものから）を選択するように構成されてもよく、選択されるのは、最良のマッチ、すなわち、最小の残差（最小の残差は、伝送または記憶のためのよりよい圧縮を意味する）または最小の信号伝達オーバーヘッド（最小の信号伝達オーバーヘッドは、伝送または記憶のためのよりよい圧縮を意味する）を提供する、または両方を考慮またはバランスさせるものである。モード選択ユニット260は、レート歪み最適化（RDO）に基づいてパーティション分割および予測モードを決定する、つまり、最小レート歪みを提供する予測モードを選択するように構成されてもよい。この文脈における「最良」、「最小」、「最適」などのような用語は、必ずしも全体的な「最良」、「最小」、「最適」などを指すのではなく、終了、または値が閾値を超えるまたは下回るのような選択基準、または他の制約条件の達成を指すこともあり、潜在的には「最適ではない選択」につながるが、複雑さと処理時間を低減する。換言すれば、パーティション分割ユニット262は、ブロック203を、たとえば四分木パーティション分割（QT）、二分パーティション分割（BT）もしくは三分木パーティション分割（TT）またはそれらの任意の組み合わせを逐次反復的に使用して、より小さなブロック・パーティションまたはサブブロック（これもやはりブロックをなす）にパーティション分割し、たとえば、ブロック・パーティションまたはサブブロックのそれぞれについて予測を実行するように構成されてもよい。ここで、モード選択は、パーティション分割されたブロック203の木構造の選択を含み、予測モードは、ブロック・パーティションまたはサブブロックのそれぞれに適用される。

30

40

【0062】

前述のように、本明細書で使用される「ブロック」という用語は、ピクチャーの一部、

50

特に正方形または長方形の一部であってもよい。たとえばHEVCおよびVVCを参照するに、ブロックは、符号化ツリー単位 (CTU)、符号化単位 (CU)、予測単位 (PU)、および変換単位 (TU)、および/または対応するブロック、たとえば、符号化ツリー・ブロック (CTB)、符号化ブロック (CB)、変換ブロック (TB)、または予測ブロック (PB) であってもよく、またはこれらに対応してもよい。たとえば、符号化ツリー単位 (CTU) は、3つのサンプル・アレイを有するピクチャーのルーマ・サンプルのCTB、クロマ・サンプルの2つの対応するCTB、または、モノクロ・ピクチャーまたはサンプルを符号化するために使用される3つの別個の色平面およびシタックス構造を使用して符号化されるピクチャーのサンプルのCTBであってもよく、または、これらを含んでもよい。対応して、符号化ツリーブロック (CTB) は、ある成分のCTBへの分割がパーティション分割であるように、Nの何らかの値についてのサンプルの $N \times N$ ブロックであってもよい。符号化単位 (CU) は、3つのサンプル・アレイを有するピクチャーのルーマ・サンプルの符号化ブロック、クロマ・サンプルの2つの対応する符号化ブロック、または、モノクロ・ピクチャーまたはサンプルを符号化するために使用される3つの別個の色平面およびシタックス構造を使用して符号化されるピクチャーのサンプルの符号化ブロックであってもよく、または、これらを含んでもよい。対応して、符号化ブロック (CB) は、CTBの符号化ブロックへの分割がパーティション分割であるように、MおよびNの何らかの値についてのサンプルの $M \times N$ ブロックであってもよい。

10

【0063】

実施形態において、たとえばHEVCによれば、符号化ツリー単位 (CTU) は、符号化ツリーと記される四分木構造を使用することによって、CUに分割されてもよい。ピクチャー間 (時間的) またはピクチャー内 (空間的) 予測を使用してピクチャー領域を符号化するかどうかの決定は、CUレベルで行なわれる。各CUはさらに、PU分割タイプに応じて、1つ、2つ、または4つのPUに分割できる。1つのPU内では、同じ予測プロセスが適用され、関連情報はPUごとにデコーダに伝送される。PU分割タイプに基づく予測プロセスを適用することによって残差ブロックを得た後、CUは、該CUについての符号化ツリーと同様の別の四分木構造に従って変換単位 (TU) にパーティション分割できる。

20

【0064】

上述のように、ビデオ・エンコーダ20は、(あらかじめ決定された) 予測モードの集合から最良または最適な予測モードを決定または選択するように構成される。予測モードの集合は、たとえば、イントラ予測モードおよび/またはインター予測モードを含んでもよい。

30

【0065】

イントラ予測モードの集合は、たとえばHEVCで定義されているように、35の異なるイントラ予測モード、たとえばDC (または平均) モードおよび平面モードのような非方向性モード、または、方向性モードを含んでもよく、または、たとえばVVCについて定義されているように、67の異なるイントラ予測モード、たとえばDC (または平均) モードおよび平面モードのような非方向性モード、または、方向性モードを含んでもよい。イントラ予測ユニット254は、イントラ予測モードの集合のうちのあるイントラ予測モードに従ってイントラ予測ブロック265を生成するために、同じ現在ピクチャーの近傍ブロックの再構成されたサンプルを使用するように構成される。

40

【0066】

イントラ予測ユニット254 (または、一般に、モード選択ユニット260) は、イントラ予測パラメータ (または、一般に、そのブロックについての選択されたイントラ予測モードを示す情報) を、エンコードされたピクチャー・データ21に含めるために、構文要素266の形でエントロピー符号化ユニット270に対して出力するようにさらに構成される。それにより、たとえば、ビデオ・デコーダ30は、予測パラメータを受領し、デコードのために使用することができる。

【0067】

前記集合の (または可能な) インター予測モードは、利用可能な参照ピクチャー (すな

50

わち、たとえばDBP 230に記憶されている、以前の少なくとも部分的にデコードされたピクチャー)および他のインター予測パラメータに依存する。該他のインター予測パラメータは、たとえば、最良マッチの参照ブロックを探すために参照ピクチャーの全体が使用されるか、または参照ピクチャーの一部のみ、たとえば、現在ブロックの当該領域のまわりの探索窓領域が使用されるか、および/または、たとえば、ピクセル補間、たとえば、半分/セミピクセルおよび/または1/4ピクセル補間が適用されるか否かである。インター予測モードは、のちに下記の実施形態において記載される、動き場決定および表現とともに動作するモードを含んでいてもよい。そのようなモードは、複数のインター・モードのうちの一つであってもよい。

【0068】

上記の予測モードに加えて、スキップモードおよび/または直接モードが適用されてもよい。

【0069】

インター予測ユニット244は、動き推定(ME)ユニットおよび動き補償(MC)ユニットを含んでいてもよい(いずれも図2には示されていない)。動き推定ユニットは、ピクチャー・ブロック203(現在ピクチャー17の現在ピクチャー・ブロック203)およびデコードされたピクチャー231、または少なくとも一つまたは複数の以前に再構成されたブロック、たとえば、一つまたは複数の他の/異なる以前にデコードされたピクチャー231の再構成されたブロックを動き推定のために受領または取得するように構成されてもよい。たとえば、ビデオ・シーケンスは、現在ピクチャーと以前にデコードされたピクチャー231とを含んでいてもよく、あるいは、換言すれば、現在ピクチャーと以前にデコードされたピクチャー231は、ビデオ・シーケンスを形成するピクチャーのシーケンスの一部であってもよく、または該シーケンスを形成してもよい。

【0070】

エンコーダ20は、たとえば、複数の他のピクチャーのうちの一つまたは異なるピクチャーの複数の参照ブロックから参照ブロックを選択し、参照ピクチャー(または参照ピクチャー・インデックス)および/または参照ブロックの位置(x、y座標)と現在ブロックの位置との間のオフセット(空間オフセット)を、動き推定ユニットへのインター予測パラメータとして提供するように構成されてもよい。このオフセットは動きベクトル(MV)とも呼ばれる。

【0071】

動き補償ユニットは、たとえば、インター予測パラメータを取得、たとえば受領して、該インター予測パラメータに基づいて、または、該インター予測パラメータを用いて、インター予測を実行して、インター予測ブロック265、または一般には現在ピクチャーのいくつかのサンプルについての予測を得るように構成される。動き補償ユニットによって実行される動き補償は、動き推定によって決定された動き/ブロック・ベクトルに基づいて、可能性としてはサブピクセル精度への補間を実行して、予測ブロック(予測サンプル)を取ってくるまたは生成することを含んでもよい。補間フィルタリングが、既知のピクセル・サンプルから追加的なピクセル・サンプルを生成してもよく、よって、潜在的には、ピクチャー・ブロックを符号化するために使用されうる候補予測ブロック/サンプルの数を増加させる。現在ピクチャー・ブロックのPUについての動きベクトルを受領すると、動き補償ユニットは、参照ピクチャー・リストのうちの一つのリストにおいて動きベクトルがポイントする予測ブロックを位置特定することができる。動き補償ユニットはまた、ビデオ・スライスのピクチャー・ブロックをデコードする際にビデオ・デコーダ30が使用するために、ブロック、サンプル領域およびビデオ・スライスに関連する構文要素を生成してもよい。スライスおよびそれぞれの構文要素への追加または代替として、タイル・グループおよび/またはタイルおよびそれぞれの構文要素が生成または使用されてもよい。

【0072】

エントロピー符号化ユニット270は、たとえば、エントロピー符号化アルゴリズムまたは方式(たとえば、可変長符号化(VLC)方式、コンテキスト適応VLC方式(CAVLC)、

10

20

30

40

50

演算符号化方式、二値化、コンテキスト適応二進算術符号化 (CABAC)、シンタックスベースのコンテキスト適応二進算術符号化 (SBAC)、確率区間パーティション分割エントロピー (PIPE) 符号化、または他のエントロピー符号化方法または技術) またはバイパス (圧縮なし) を、量子化された係数209、インター予測パラメータ、イントラ予測パラメータ、ループ・フィルタ・パラメータおよび/または他の構文要素に対して適用して、エンコードされたピクチャー・データ21を得るように構成される。エンコードされたピクチャー・データ21はたとえばエンコードされたビットストリーム21の形で、出力272を介して出力でき、それにより、たとえば、ビデオ・デコーダ30はそれらのパラメータを受領し、デコードのために使用することができる。エンコードされたビットストリーム21は、ビデオ・デコーダ30に送信されてもよいし、または後の送信またはビデオ・デコーダ30による取得のためにメモリに記憶されてもよい。

10

【0073】

ビデオ・ストリームをエンコードするために、ビデオ・エンコーダ20の他の構造的変形が使用できる。たとえば、非変換ベースのエンコーダ20は、ある種のブロックまたはフレームについて、変換処理ユニット206なしで直接、残差信号を量子化することができる。別の実装では、エンコーダ20は、量子化ユニット208と逆量子化ユニット210とを単一のユニットに組み合わせたものを有することができる。

【0074】

図2は、本開示の技術を実装するように修正または構成されうるビデオ・デコーダ30の例を示す。ビデオ・デコーダ30は、デコードされたピクチャー331を得るために、たとえばエンコーダ20によってエンコードされた、エンコードされたピクチャー・データ21 (たとえばエンコードされたビットストリーム21) を受領するように構成される。エンコードされたピクチャー・データまたはビットストリームは、エンコードされたピクチャー・データをデコードするための情報、たとえば、エンコードされたビデオ・スライス (および/またはタイル・グループまたはタイル) のピクチャー・ブロックおよび関連する構文要素を表わすデータを含む。

20

【0075】

図2の例では、デコーダ30は、エントロピー復号ユニット304と、逆量子化ユニット310と、逆変換処理ユニット312と、再構成ユニット314 (たとえば、加算器314) と、ループ・フィルタ320と、デコードピクチャーバッファ (DBP) 330と、モード適用ユニット360と、インター予測ユニット344と、イントラ予測ユニット354とを含む。インター予測ユニット344は、動き補償ユニットであってもよいし、またはこれを含んでいてもよい。ビデオ・デコーダ30は、いくつかの例では、図1からのビデオ・エンコーダ100に関して記述されたエンコード・パスと概して逆のデコード・パスを実行することができる。

30

【0076】

エンコーダ20に関して説明したように、逆量子化ユニット210、逆変換処理ユニット212、再構成ユニット214、ループ・フィルタ220、デコードピクチャーバッファ (DPB) 230、インター予測ユニット344、およびイントラ予測ユニット354は、ビデオ・エンコーダ20の「組み込みデコーダ」を形成するとともに称される。よって、逆量子化ユニット310は、逆量子化ユニット110と機能的に同一であってもよく、逆変換処理ユニット312は、逆変換処理ユニット212と機能的に同一であってもよく、再構成ユニット314は、再構成ユニット214と機能的に同一であってもよく、ループ・フィルタ320は、ループ・フィルタ220と機能的に同一であってもよく、デコードピクチャーバッファ330は、デコードピクチャーバッファ230と機能的に同一であってもよい。したがって、ビデオ20エンコーダのそれぞれのユニットおよび機能について与えた説明は、ビデオ・デコーダ30のそれぞれのユニットおよび機能に、対応して当てはまる。

40

【0077】

エントロピー・デコードユニット304は、ビットストリーム21 (または一般に、エンコードされたピクチャー・データ21) をパースし、たとえば、エンコードされたピクチャー・データ21に対してエントロピー復号を実行し、たとえば、量子化された係数309および

50

／またはデコードされた符号化パラメータ（図3には示されていない）、たとえば、インター予測パラメータ（たとえば、参照ピクチャー・インデックスおよび動きベクトル）、イントラ予測パラメータ（たとえば、イントラ予測モードまたはインデックス）、変換パラメータ、量子化パラメータ、ループ・フィルタ・パラメータ、および／または他の構文要素のいずれかまたは全部を得るように構成される。エンтроピー復号ユニット304は、エンコーダ20のエンтроピー符号化ユニット270に関して述べたエンコード方式に対応するデコード・アルゴリズムまたは方式を適用するように構成されてもよい。エンтроピー復号ユニット304は、インター予測パラメータ、イントラ予測パラメータおよび／または他の構文要素をモード選択ユニット360に、他のパラメータをデコーダ30の他のユニットに提供するようにさらに構成されてもよい。ビデオ・デコーダ30は、ビデオ・スライス・レベルおよび／またはビデオ・ブロック・レベルで構文要素を受領してもよい。スライスおよびそれぞれの構文要素への追加または代替として、タイル・グループおよび／またはタイルおよびそれぞれの構文要素が受領および／または使用されてもよい。

【0078】

逆量子化ユニット310は、エンコードされたピクチャー・データ21から量子化パラメータ（QP）（または一般に、逆量子化に関する情報）および量子化された係数を受領し（たとえばエンтроピー復号ユニット304によるたとえばパースおよび／またはデコードによって）、デコードされた量子化された係数309に対して該量子化パラメータに基づいて逆量子化を適用して、量子化解除された係数311を得るように構成されてもよい。量子化解除された係数311は変換係数311と称されることもある。逆量子化プロセスは、ビデオ・スライス（またはタイルまたはタイル・グループ）内の各ビデオ・ブロックについてビデオ・エンコーダ20によって決定された量子化パラメータを使用して、量子化の程度、および、同様に、適用されるべき逆量子化の程度を決定してもよい。

【0079】

逆変換処理ユニット312は、変換係数311とも呼ばれる量子化解除された係数311を受領し、サンプル領域における再構成された残差ブロック213を得るために、量子化解除された係数311に変換を適用するように構成されてもよい。再構成された残差ブロック213は、変換ブロック313と称されることもある。変換は、逆変換、たとえば、逆DCT、逆DST、逆整数変換、または概念的に同様の逆変換プロセスであってもよい。逆変換処理ユニット312は、さらに、変換パラメータまたは対応する情報をエンコードされたピクチャー・データ21から受領して（たとえばエンтроピー復号ユニット304によるたとえばパースおよび／またはデコードによって）、量子化解除された係数311に適用される変換を決定するように構成されてもよい。

【0080】

再構成ユニット314（たとえば、加算器または総和器314）は、再構成された残差ブロック313を予測ブロック365に加算して、サンプル領域における再構成されたブロック315を得るように構成されてもよい。これはたとえば、再構成された残差ブロック313のサンプル値および予測ブロック365のサンプル値を加算することによる。

【0081】

ループ・フィルタ・ユニット320（符号化ループ内または符号化ループ後）は、たとえばピクセル遷移をなめらかにする、または、他の仕方でビデオ品質を改善するために、再構成されたブロック315をフィルタリングして、フィルタリングされたブロック321を得るように構成される。ループ・フィルタ・ユニット320は、ブロッキング解除フィルタ、サンプル適応オフセット（SAO）・フィルタ、または一つまたは複数の他のフィルタ、たとえばバイラテラル・フィルタ、適応ループ・フィルタ（ALF）、鮮鋭化、平滑化フィルタ、または協働フィルタ、またはそれらの任意の組み合わせなどの一つまたは複数のループ・フィルタを含んでいてもよい。いくつかの実施形態では、ループ・フィルタ・ユニット320は、図1を参照してすでに上述したように、ループ後フィルタとして実装されてもよい。ループ・フィルタとポスト・フィルタの両方のフィルタを実装することが可能であることを注意しておく。ポスト・フィルタは、たとえば表示のために画像331出力を向上

10

20

30

40

50

させるために用いられてもよい。ループ・フィルタは、インター符号化において参照ピクチャーとして使用するためにバッファに記憶される、再構成されたサンプル315をフィルタリングするために用いられてもよい。

【0082】

次いで、ピクチャーのデコードされたビデオ・ブロック321は、デコードピクチャーバッファ330に記憶される。デコードピクチャーバッファ330は、デコードされたピクチャー331を、他のピクチャーについてのその後の動き補償のための参照ピクチャーとして、および/またはそれぞれ出力もしくは表示のために記憶する。デコーダ30は、デコードされたピクチャー311を、ユーザーへの提示または閲覧のために、たとえば出力312を介して出力するように構成される。

10

【0083】

インター予測ユニット344は、インター予測ユニット244（特に、動き補償ユニット）と同一であってもよく、イントラ予測ユニット354は、機能において、インター予測ユニット254と同一であってもよく、エンコードされたピクチャー・データ21から（たとえばエンтроピー復号ユニット304による、たとえばパースおよび/またはデコードによって）受領されるパーティション分割および/または予測パラメータまたはそれぞれの情報に基づいて、分割もしくはパーティション分割決定および予測を実行する。モード適用ユニット360は、再構成されたピクチャー、ブロック、またはそれぞれのサンプル（フィルタリングされた、またはフィルタリングされていない）に基づいて、ブロックごとのまたはサンプル・ベースの予測（イントラ予測またはインター予測）を実行して、予測ブロック365を得るように構成されてもよい。

20

【0084】

ビデオ・スライスがイントラ・符号化される（I）スライスとして符号化される時、モード適用ユニット360のイントラ予測ユニット354は、信号伝達されたイントラ予測モードと、現在ピクチャーの以前にデコードされたブロックからのデータとに基づいて、現在ビデオ・スライスのピクチャー・ブロックについての予測ブロック365を生成するように構成される。ビデオ・ピクチャーがインター符号化される（すなわちBまたはP）スライスとして符号化される時、モード適用ユニット360のインター予測ユニット344（たとえば、動き補償ユニット）は、エンтроピー復号ユニット304から受領される動きベクトルおよび他の構文要素に基づいて、現在のビデオ・スライスのビデオ・ブロックについての予測ブロック365を生成するように構成される。インター予測のために、予測ブロックは、参照ピクチャー・リストのうちの1つのリスト内の参照ピクチャーの1つから生成されてもよい。ビデオ・デコーダ30は、DPB 330に記憶された参照ピクチャーに基づくデフォルト構築技術を使用して、参照フレームリスト、リスト0およびリスト1を構築することができる。同じまたは同様のことが、スライス（たとえばビデオ・スライス）への追加または代替として、タイル・グループ（たとえばビデオ・タイル・グループ）および/またはタイル（たとえばビデオ・タイル）を使用する実施形態のために、またはかかる実施形態によって適用されてもよい。たとえば、ビデオは、I、PまたはBタイル・グループおよび/またはタイルを使って符号化されてもよい。

30

【0085】

モード適用ユニット360は、動きベクトルまたは関連情報および他の構文要素をパースすることによって、現在のビデオ・スライスのビデオ・ブロックについての予測情報を決定するように構成され、該予測情報を使用して、デコードされる現在のビデオ・ブロックについての予測ブロックを生成する。たとえば、モード適用ユニット360は、受領された構文要素のいくつかを用いて、ビデオ・スライスのビデオ・ブロックを符号化するために使用される予測モード（たとえば、イントラ予測またはインター予測）、インター予測スライス・タイプ（たとえば、Bスライス、Pスライス、またはGPBスライス）、スライスについての参照ピクチャー・リストのうちの1つまたは複数のリストについての構築情報、スライスのそれぞれのインター・エンコードされたビデオ・ブロックについての動きベクトル、スライスのそれぞれのインター符号化されたビデオ・ブロックについてのインター

40

50

予測ステータス、および現在のビデオ・スライス内のビデオ・ブロックをデコードするための他の情報を決定する。同じまたは同様のことが、スライス（たとえばビデオ・スライス）への追加または代替として、タイル・グループ（たとえばビデオ・タイル・グループ）および/またはタイル（たとえばビデオ・タイル）を使用する実施形態のために、またはかかる実施形態によって適用されてもよい。たとえば、ビデオは、I、PまたはBタイル・グループおよび/またはタイルを使って符号化されてもよい。

【0086】

図2に示されるビデオ・デコーダ30の実施形態は、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ーをパーティション分割および/またはデコードするように構成されてもよく、ピクチャーは、一つまたは複数のスライス（典型的には、重複しない）にパーティション分割され、またはかかるスライスを使ってデコードされてもよく、各スライスは一つまたは複数のブロック（たとえば、CTU）を含んでいてもよい。

10

【0087】

図2に示されるビデオ・デコーダ30の実施形態は、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）を使用して、ピクチャーをパーティション分割および/またはデコードするように構成されてもよく、ピクチャーは、一つまたは複数のタイル・グループ（典型的には、重複しない）にパーティション分割され、またはかかるタイル・グループを使ってデコードされてもよく、各タイル・グループは、たとえば、一つまたは複数のブロック（たとえば、CTU）または一つまたは複数のタイルを含んでいてもよく、各タイルは、たとえば、長方形の形状であってもよく、一つまたは複数のブロック（たとえば、CTU）、たとえば、完全なまたは断片的なブロックを含んでいてもよい。

20

【0088】

ビデオ・デコーダ30の他の変形が、エンコードされたピクチャー・データ21をデコードするために使用できる。たとえば、デコーダ30は、ループ・フィルタリング・ユニット320なしで出力ビデオ・ストリームを生成することができる。たとえば、非変換ベースのデコーダ30は、ある種のブロックまたはフレームについて、逆変換処理ユニット312なしで直接、残差信号を逆量子化することができる。別の実施形態では、ビデオ・デコーダ30は、単一のユニットに組み合わされた逆量子化ユニット310および逆変換処理ユニット312を有することができる。

30

【0089】

図3Aは、出力画像に修正を適用する、ポスト・フィルタの位置に適用される向上390を示す。図3Bは、出力画像および動き推定参照ピクチャーバッファの両方に変換を適用する、ループ内フィルタ（本明細書では、ループ・フィルタとも呼ばれる）の位置で適用される向上380を示す。

【0090】

画像向上フィルタの一つのタイプは、チャンネル間の類似性を利用することによって、マルチチャンネル画像の品質を改善するものである。マルチチャンネル画像向上アルゴリズムの性能は、入力マルチチャンネル画像のいくつかのパラメータ（たとえば、チャンネルの数、それらの品質）によって変化し、各チャンネルにおいては画像データにわたっても変化する。

40

【0091】

本開示のいくつかの実施形態は、入力画像を解析し、各ケースについて好適な画像向上パラメータを選択することを目的とする。

【0092】

さまざまな画像向上アルゴリズムが存在する。それらのうち少数のみが、画像向上のためにチャンネル間相関情報を利用する。本開示では、畳み込みニューラル・ネットワークのようなニューラル・ネットワークを使用するマルチチャンネル画像向上フィルタに焦点を当てる。ニューラルネットワークベースの向上フィルタでは、ネットワークは、2組の画像でトレーニングされる。一方はもとの（目標の、所望される）品質を表し、他方は、期待

50

される歪みの範囲およびタイプを表す。そのようなネットワークは、たとえばセンサー・ノイズにより損なわれた画像、またはビデオ圧縮により、または他の種類の歪みにより損なわれた画像を改善するようにトレーニングされうる。通例、歪みのタイプごとに異なる（個別であり、別個の）トレーニングが必要とされる。より一般的なネットワーク（たとえば、より広い範囲およびタイプの歪みを扱うもの）は、より低い平均性能を有する。ここで、性能は、たとえば、PSNRのような客観的基準によって、または人間の視覚的知覚をも考慮するいくつかのメトリックによって測定されうる再構成の品質をいう。

【0093】

近年、ニューラル・ネットワークが注目され、画像処理への応用が提案されている。特に、そのような応用においては畳み込みニューラルネットワーク（CNN）が採用されてきた。一つの可能性は、圧縮パイプラインを完全にニューラル・ネットワークに置き換えることである。すると、画像圧縮は、CNNによってエンドツーエンドで学習される。これらの圧縮アーチファクトを低減する別の可能性は、圧縮後にフィルタを適用することである。HEVC圧縮規格では、すでに単純なループ内フィルタが存在する。より複雑なフィルタ、特に畳み込みニューラルネットワーク（CNN）に基づくフィルタが文献で提案されている。しかしながら、視覚品質の改善は限られている。

【0094】

ニューラル・ネットワークは、機械学習をサポートし、複数の相互接続されたニューロンを含む、人間の脳に倣ってモデル化される信号処理モデルである。ニューラル・ネットワーク実装では、2つのニューロン間の接続における信号は、数であり、各ニューロンの出力は、その重み付けされた入力の和の何らかの非線形関数によって計算される。接続はエッジと呼ばれる。典型的には、ニューロンおよびエッジは、学習の進行につれて調整される重みを有する。重みは、接続における信号の強度を増減させる。重み付けされた和の非線形関数は、「活性化関数」または「ニューロンの伝達関数」とも呼ばれる。いくつかの単純な実装では、出力は、重み付けされた和が、ある閾値を超えるかどうか依存して二値であってもよい。これは、非線形活性化関数としてのステップ関数に対応する。他の実装では、シグモイドなどの別の活性化関数が使用されてもよい。典型的には、ニューロンは諸層にまとめられる。異なる層は、入力の異なる変換を実行しうる。信号は、第1層（入力層）から、可能性としては複数の層を通過した後に、最終レイヤー（出力層）に進む。重みは、教師付き学習または教師なし学習によって実行されうるトレーニングによって学習される。なお、上記のモデルは、一般的なモデルにすぎない。特定の用途のためには、ニューラル・ネットワークは、CNNの諸層に対応しうる、画像等の所望の入力に適応される種々の処理段を有していてもよい。

【0095】

本願のいくつかの実施形態において、深層畳み込みニューラルネットワーク（CNN）は、高い圧縮率を維持しつつ、圧縮アーチファクトを低減し、画像の視覚的品質を向上させるようにトレーニングされる。具体的には、ある実施形態によれば、入力画像を修正する方法が提供される。ここで、修正とは、フィルタリングまたは他の画像向上アプローチによって典型的に得られる修正のような任意の修正を指す。修正のタイプは、具体的な用途に依存しうる。それぞれの特定のケースのためにトレーニングされる必要なしに広範囲の歪みについて良好な結果をもたらすネットワークの一つは、非特許文献1から知られている。そこでは、デコーダ側で画質を向上させるためにチャンネル間相関を利用できる、3段の畳み込みニューラル・ネットワーク（CNN）ベースのアプローチが提案される。

【文献】Cui, Kai & Steinbach, Eckehard (2018): "Decoder Side Image Quality Enhancement exploiting Inter-channel Correlation in a 3-stage CNN" Submission to CLIC 2018, IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, June 2018.

【0096】

図4は、そのような3チャンネルのCNNフレームワークを示している。CNNはニューラル・ネットワークであり、その諸層のうちの少なくとも1つにおいて、一般的な行列乗算の

10

20

30

40

50

代わりに畳み込みを使用する。畳み込み層は入力を畳み込みし、その結果を次の層に渡す。それらは、特に画像/ビデオ処理において、いくつかの有益な特徴を有する。既知の構成、ならびに本開示のいくつかの実施形態においてCNNの適用を容易にしうる可能な構成および代替を含めて、図4のCNNが記述され、適用される諸段が説明される。

【0097】

入力画像はRGB（赤、緑、青）フォーマット（色空間）で格納される。入力画像は静止画であってもよいし、ビデオシーケンス（動画）のフレームである画像であってもよい。

【0098】

図4における丸囲みの数字は処理の諸段を示す。段1では、入力画像からパッチが選択される。具体的な例では、パッチは、 240×240 サンプル（ピクセル）のサイズなど、所定のサイズを有する。パッチ・サイズは、固定されていてもよく、またはあらかじめ決定されてもよい。たとえば、パッチ・サイズは、ユーザーまたはアプリケーションによって構成可能であってもよく、またはビットストリーム内で伝達され、それに従って設定されてもよいパラメータであってもよい。パッチ・サイズの選択は、画像サイズおよび/または画像内の詳細の量に応じて実行されうる。

10

【0099】

ここで、「パッチ」という用語は、フィルタリングによって処理される画像の一部を指し、処理された部分は、次いで、そのパッチの位置にペーストされる。パッチは、長方形または正方形など規則的であってもよい。しかしながら、本開示はそれに限定されず、パッチは、検出/認識されたオブジェクトの形状に従う形状など、フィルタリングされるべき任意の形状を有してもよい。いくつかの実施形態では、画像全体がパッチごとにフィルタリング（向上）される。他の実施形態では、選択されたパッチのみ（たとえば、オブジェクトに対応する）がフィルタリングされて、画像の残りの部分は、フィルタリングされない、または別のアプローチによってフィルタリングされるのでもよい。フィルタリングにより、任意の種類の種類を拡張を意味する。

20

【0100】

選択は、すべてのパッチがフィルタリングされる、逐次的処理または並列処理の結果であってもよい。そのような場合、選択は、左から右へ、および上から下へのような所定の順序で実行されてもよい。しかしながら、選択はまた、ユーザーまたはアプリケーションによって実行されてもよく、画像の一部のみを考慮してもよい。パッチは連続的であってもよいし、分散していてもよい。

30

【0101】

画像全体がパッチ領域に分割されている場合、パッチ寸法（垂直または水平）の整数倍が画像サイズと一致しなければ、パディングが適用されてもよい。パディングは、整数個のパッチに適合するサイズを達成するために、画像境界（水平または垂直）によって形成される軸上で利用可能な画像部分のミラーリングを含んでいてもよい。より詳細には、パディングは、垂直寸法（パディング後のサンプル数）が垂直パッチ寸法の整数倍であるように実行される。さらに、水平寸法（パディング後のサンプル数）は、水平パッチ寸法の整数倍である。

【0102】

次いで、画像向上は、各パッチを逐次的にまたは並列的に選択および処理することによって実行されうる。パッチは、上記で示唆したように重複しなくてもよい。しかしながら、パッチは重複していてもよく、これは品質を改善し、別個に処理されたパッチ間の可能性のある境界効果を減少させることができる。

40

【0103】

段2では、より容易な処理のために、パッチのピクセルが並べ替えられる。並べ替えは、図5に示されるように、いわゆるピクセル・シフトを含んでいてもよい。ピクセル・シフトは、各チャンネル内のピクセルを並べ替え、それにより、寸法 $N \times N \times 1$ をもつチャンネルは、寸法 $N/2 \times N/2 \times 4$ をもつ3Dアレイに変換される（ここで、記号「 \times 」は「かける」、すなわち乗算を表す）。これは、 2×2 の値の重複しない各ブロックから単一の値が取ら

50

れる、チャンネルのサブサンプリングによって実行される。3Dアレイの最初の層は左上の値から作成され、2番目の層は右上、3番目の層は左下の値から作成され、4番目の層は右下の値から作成される。ピクセル・シフト操作の終わりには、サイズが $N \times N$ ピクセルの3チャンネルRGBパッチは、寸法 $N/2 \times N/2 \times 4$ のスタック（3Dアレイ）になる。このピクセル・シフトは、主に計算上の理由のために行われる。プロセッサ（グラフィカル処理ユニット、GPUなど）にとっては、広く浅いスタックよりも、狭く深いスタックのほうが処理するのが容易なのである。本開示は、4によるサブサンプリングに限定されないことに留意されたい。一般には、より多数またはより少数によるサブサンプリングがあってもよく、より多数またはより少数の、結果として得られるサブサンプリングされた画像の、対応するスタック深さを生じる。

10

【0104】

図5は、ピクセル・シフトの詳細な例を示す。この例では、ステップ510において、入力サイズ 240×240 のパッチが取られる。次いで、水平方向および垂直方向における一つおきのピクセルを取ることによって、4つの部分画像が作成される。ステップ520は、1、2、3、および4と示される4つのサブサンプリングされた画像に対する画像のピクセルの割り当てを示す。手順は以下の通りである。

- ・画像2.1を得る（ステップ540参照）：もとの画像（ 240×240 パッチ）における位置（0,0）における左上の座標から始めて、部分画像1を作成する。次いで、垂直方向に一つおきのサンプルを、水平方向において一つおきのサンプルを取る。グループ化はステップ530の左上側に示される。

20

- ・画像2.2を得る（ステップ540参照）：もとの画像における位置（0,1）における左上の座標から始めて、部分画像2を作成する。次いで、垂直方向に一つおきのサンプルを、水平方向において一つおきのサンプルを取る。グループ化はステップ530の右上側に示される。

- ・画像2.3を得る（ステップ540参照）：もとの画像における位置（1,0）における左上の座標から始めて、部分画像3を作成する。次いで、垂直方向に一つおきのサンプルを、水平方向において一つおきのサンプルを取る。グループ化はステップ530の左下側に示される。

- ・画像2.4を得る（ステップ540参照）：もとの画像における位置（1,1）における左上の座標から始めて、部分画像4を作成する。次いで、垂直方向に一つおきのサンプルを、水平方向において一つおきのサンプルを取る。グループ化はステップ530の右下側に示される。

30

【0105】

図4の段3では、緑チャンネルは単一チャンネル・モードで処理される。図4のこの例では、緑チャンネルが一次チャンネルであり、残りのチャンネルが二次チャンネルであると想定されている。各チャンネルの歪みはチャンネル色と強く相関すると想定される。緑チャンネルは、歪みが最も小さいチャンネルであると想定されている。平均的には、これは、RGB画像は赤および青のサンプルよりも多くの緑のサンプルを捕捉するRGGBセンサー・パターン（ベイヤー・パターン）を用いて捕捉されるので、センサー・ノイズまたは圧縮によって影響を受けるRGB画像についての妥当な推測となりうる。しかしながら、本開示によれば、以下で論じられるように、緑チャンネルは必ずしも最良ではなく、チャンネル選択を実行することが有利でありうる。さらに、色チャンネル以外のチャンネルが使用されてもよく、より高い品質の情報を提供することができる。

40

【0106】

図4において、一次（すなわち、緑）チャンネルは、段3において単独で処理される。赤チャンネルは、改善された緑チャンネルと一緒に処理される（段4）。青チャンネルは、改善された緑チャンネルと一緒に処理される（段5）。次いで、改善された赤、緑、および青チャンネルは、一緒に積み重ねられ（組み合わされ）（段6）、図8および図9を参照して以下に説明するように、サイド情報なしに、またはサイド情報を用いて処理される。

【0107】

50

全体として、このフレームワークは4つのNNs段 図4の段3、4、5、7 がある。段3および7は、値の単一の3Dアレイを入力し、入力と同じサイズの3Dアレイを出力する。段4および5は、メイン一つと補助一つの2つの3Dアレイを入力する。出力は、メイン出力と同じサイズであり、メイン入力の処理された（たとえば向上された）バージョンであることが意図されている。一方、補助入力、前記処理を支援するためにのみ使用され、出力されない。

【0108】

図6は、単一入力ニューラル・ネットワーク（特に、図4におけるネットワーク（S）に対応するCNN）の例を示す。ネットワーク（S）は、入力として、4つの平面のスタック、すなわち、図5に示されるピクセル・シフトの4つのサブサンプリングされたパッチ画像出力に対応する $120 \times 120 \times 4$ の3Dアレイを有する。最初のCNN層は、サイズが $3 \times 3 \times 4$ の64個のカーネルを使用して、64個の畳み込みを実行する。各畳み込みの出力は、整流線形ユニット（ReLU）関数を用いて整流され、各畳み込みは、第1の隠れ層において1つの平面を生成する。活性化関数は、ReLUであってもよい。ReLUは、すべての負の数についてゼロであり、正の数については線形関数（ランプ）である。しかしながら、本開示はそれに限定されるものではなく、シグモイドまたはステップ関数などの異なる活性化関数が一般に使用されうる。ReLU関数はその形がシグモイドに近いが、それほど複雑ではない。

【0109】

諸隠れ層は、 $120 \times 120 \times 64$ の同一サイズを有する。各隠し層は、 $3 \times 3 \times 64$ の64個のカーネルを用いて64個の畳み込みを実行する。各畳み込みの出力は、（ReLUを用いて）整流され、（バッチ正規化BNを用いて）正規化され、次の隠れ層において1つの平面を生成する。全部で8つの隠れ層がある。最後の隠れ層は、4つの畳み込みのみを実行し、4つの平面を出力する。最後に、もとの入力処理されたデータに加えらる。そのような構造（ResNet）では、諸CNN層は、もとの信号と所望の信号との間の差を合成することを目的とする。

【0110】

上で簡単に説明したように、図6は5つの異なる段（丸囲みの数で示されている）を示している。第1段では、 120×120 のパッチ・サイズのZ個（前の諸例では4つだが、一般にはZ）のビットマップのスタックが取られ、それを第1層に供給する。段2では、層は、パッチのそれぞれの側に1つの値をパディングされる。次いで、パディングされた層は、サイズ $3 \times 3 \times Z$ のカーネルと畳み込みされ、ReLUユニットを用いて処理される。64個の異なるカーネルがあり、それぞれが次の層の1つのスライスを生成する。段3では、P個の隠れ層が存在する（ $P = 8$ は複雑さと性能の間のよいトレードオフでありうる）。各層は $120 \times 120 \times 64$ であり、カーネル $3 \times 3 \times 64$ を用いてパディングされ、処理される。各層について64個の異なるカーネルがあり（合計 $64 \times P$ 個）、それぞれが次の層の1つのスライスを生成する。その後、各層は、ReLUおよびBN関数によって処理される。

【数1】

$$\hat{X} = (X - \mu_X) / \sigma^2_X$$

ここで、Xは与えられた入力についてのニューロン群のネットワーク活性であり、 \hat{X} [^付きのX] はバッチ正規化された出力であり、 μ_X はあるシンギュラー次元に沿ったXの算術平均値を表し、 σ^2_X はあるシンギュラー次元に沿ったXの分散を表す。分散と平均は統計的な（計算される）値（サンプル・ベースの推定）である。

【0111】

段4では、最後の隠れ層はN個の畳み込みカーネル $3 \times 3 \times 64$ によって処理され、サイズZのスタックを出力する。段5では、入力が処理された出力に加算される。ネットワークは、入力と所望の出力との間の差を近似するようにトレーニングされているからである。

【0112】

図7は、上述の単一入力CNNと同様の二入力CNN（ネットワーク（C））を示しており、2つの違いがある。第1の相違は、入力が2倍の深さ（8平面）であるため、第1層の畳み込みカーネルのサイズは $3 \times 3 \times 8$ である。隠れ層における処理は同一である。第2の相違は、ネットワークがメイン・チャンネルと所望の出力との間の差を近似しようとしているため、メイン・チャンネルのみが処理された出力に加算されることである。（補助チャンネルは処理されるが、その出力は使用されないかのように）。換言すれば、図4の段4において、赤チャンネルは、向上された緑チャンネルと協働して処理される。そのような処理が図7に示される。処理はZ個のビットマップのメイン・スタックを取り、それをY個のビットマップの補助スタックと協調して処理し、同じ次元のビットマップのスタックをメイン出力として出力する。図7はまた、以下のような5つの異なる段をも示している。

10

【0113】

段1では、メイン入力補助入力とスタックされ、 $Z+Y$ のサイズのスタックが生成される。メイン入力のみが出力に供給される（段5参照）。組み合わせられたスタックは、段2でサイズ $3 \times 3 \times (Z+Y)$ の畳み込みカーネルを用いて処理される。残りの処理段は、単一チャンネル・モードと同様である。特に、この例では、段3、4、および5は、単一チャンネル・モードと同一である。さらに、段5では、メイン入力出力に加算される。

【0114】

同様に、図4の段5において、青チャンネルは、向上された緑チャンネルと協働して処理される。処理は、赤チャンネルおよび図7を参照して記述される処理と同様である。

20

【0115】

本開示において、「チャンネル」または「画像チャンネル」という用語は、必ずしも色チャンネルを指すとは限らないことに留意されたい。本明細書に記載される実施形態を使用して、深さチャンネルまたは他の特徴チャンネルなどの他のチャンネルが向上されてもよい。

【0116】

図4の段6では、処理されたすべてのチャンネル（赤、緑、青）が、 120×120 のサイズの $Z=12$ 個（色チャンネルごとに4つのサブサンプリングされた画像）の画像に組み合わせられる。ここで、この組み合わせは、次の段7のための共通入力として、一緒にスタックすることを意味する。段7では、組み合わせられたチャンネルはネットワーク（S）によって一緒に処理される。

30

【0117】

段8では、ピクセルが並べ替えられて、処理されたパッチを形成する。これは、ミラーリングされた部分をクロッピングすることによってパディングを除去することを含みうる。ミラーリングによる上述のパディングは、パディングを実行する方法の可能性の一つにすぎないことに留意されたい。本開示は、そのような特定の例に限定されない。

【0118】

段9では、処理されたパッチがもとの画像に挿入される。言い換えると、もとの画像は向上されたパッチによって更新される。トレーニング手順の間には、もとの画像と歪んだ画像のセットが入力として使用され、4つすべてのネットワークのすべての畳み込みカーネルが選択される。ネットワークの目的は、歪んだ画像を取得し、もとの画像との密接に一致するものを生成することである。

40

【0119】

上述のようなCNNベースのマルチチャンネル向上フィルタは、厳格で非適応的な仕方でも機能する。処理パラメータは、フィルタの設計（またはトレーニング）中にセットアップされ、フィルタを通過する画像の内容に関係なく、同じ仕方でも適用される。しかしながら、一次チャンネルの最適な選択は、画像によって、または同じ画像でも諸部分について、変わりうる。画像向上品質のためには、一次（先行）チャンネルが最も高い品質、すなわち最も低い歪みを有するチャンネルであれば有利である。これは、一次チャンネルが残りのチャンネルの向上にも関与するからである。本発明者らは、各パッチまたは各画像などについての一次チャンネルを慎重に選択することにより、より良好な性能が達成されうることを認識し

50

、これは、実験によって確認された。第二に、(すべての向上フィルタの)向上性能は、入力の画質によって変わる。それは、歪みの強さのある範囲について最適に機能し、いくつかの非常に高いまたは非常に低い歪みレベルについてはあまり改善を提供しない。これは、高品質の入力はほとんど改善できず、低品質の入力はあまりに歪んでいるため、信頼できる仕方では改善できないからである。結果として、いくつかの入力については、向上処理を完全にスキップすることが有益である。第三に、上記(図4参照)の画像向上は、3つのチャンネルすべての解像度が同一であるRGB入力のために設計されている。しかしながら、いくつかのビデオ標準フォーマット(たとえば、YUV 4:2:0またはYUV 4:2:2など)において、異なるチャンネルは、異なる解像度およびピクセル数を有することがある。本明細書に記載される画像修正は、必ずしもエンコードまたはデコードの間に適用されるとは限らないことに留意されたい。それは前処理にも使用されうる。たとえば、ペイヤー・パターン・ベースのフォーマットのような、生のフォーマットにおける画像またはビデオを向上させるために使用されてもよい。換言すれば、画像の修正は、任意の画像またはビデオに適用されうる。

【0120】

本開示のいくつかの実施形態は、異なる入力フォーマットおよび異なる内容に適合することができるモジュール式で調整可能な向上フィルタを提供する。それは、任意のチャンネル数で、各チャンネルにおける異なるピクセル数をもつマルチチャンネル画像フォーマットを処理することができる。また、図8に示される参照なしセットアップで(すなわち、オリジナルを知らずに歪んだ画像を解析することによってのみ)、または図9に示される参照セットアップで機能できる内容解析モジュールが追加される。該解析に続いて、向上フィルタは、特定の順序で入力チャンネルを処理するように、または処理を完全にスキップするように調整されることができる。

【0121】

特に、2つ以上の画像チャンネルによって表される画像領域を修正するための方法(800または900など)が提供される。ここで「修正する」という用語は、画像フィルタリングまたは画像向上などの任意の修正を指す。2つ以上のチャンネルは、色チャンネル、または奥行きチャンネル、またはマルチスペクトル画像チャンネル、または任意の他の特徴チャンネルなどの他のチャンネルであってもよい。本方法は、2つ以上の画像チャンネルのうちの1つを一次チャンネルとして選択840、970し、2つ以上の画像チャンネルのうちの別の少なくとも1つを二次チャンネルとして選択することを含む。一次チャンネルは(いくつかの実施形態によれば)先行チャンネルと考えられることに留意されたい。二次チャンネルは(いくつかの実施形態によれば)応答/反応チャンネルと考えられる。

【0122】

本方法はさらに、第1のニューラル・ネットワークを用いて一次チャンネルを処理して、修正された一次チャンネルを得るステップを含む。第1のニューラル・ネットワークは、たとえば、図4のネットワーク(S)または図6に記載されたネットワークであってもよい。本方法は、さらに、第2のニューラル・ネットワークを用いて二次チャンネルを処理して、修正された二次チャンネルを得るステップを含む。第2のニューラル・ネットワークによる処理は、修正された一次チャンネルに基づく。第2のニューラル・ネットワークは、たとえば、図4のネットワーク(C)または図7に記載されたネットワークであってもよい。図4、図6および図7のこれらのネットワークは単に例示的なものであり、一般には、異なる数およびサイズの層を有するネットワークが使用されうる。

【0123】

本方法は、さらに、修正された一次チャンネルおよび修正された二次チャンネルに基づいて修正された画像領域を得ることを含む。この取得ステップは、たとえば、ネットワーク(S)によって図4に示されるように、修正された一次チャンネルおよび二次チャンネルの組み合わせと、それらを一緒にさらに処理することに対応しうる。

【0124】

本方法は、図8および図9を参照してさらに説明される。図8は、たとえば、修正される

10

20

30

40

50

べき歪んだ画像の入力を含む、デコーダまたはデコード方法800を示す。この歪んだ画像は、ピクセル再配置モジュール810に提供される。ピクセル再配置は、上記の図5を参照して記載されたピクセル・シフトと同じ仕方で行われてもよい。しかしながら、他の方法を使用することができ、ピクセル再配置パターン520は、特にそれらが異なるサイズ（寸法〔次元〕および/または解像度）を有する場合、異なるチャンネルについては異なるように見えることがある。本開示が再配置を適用するのは、それが限られたサイズの同じ入力フォーマットをニューラル・ネットワークに提供するための手段であるからであることに留意されたい。しかしながら、このステップは、いくつかのネットワークについては、または画像パッチがすでに好適なサイズを有するように選択されている場合には、または一般に他の実装については、必要ではないことがある。

10

【0125】

図の800は、さらに、処理のためにパッチを選択するステップ820を示している。このステップは、必ずしもパッチを選択する際に何らかのインテリジェンスがあることを意味するわけではないことに留意されたい。この選択は、逐次処理（たとえば、ループ内）に従って行われてもよく、または、たとえば並列処理が適用される場合には、一度に2つ以上のパッチ（または、すべてのパッチ）について行われてもよい。選択はまた、図4を参照して述べたさらなる例に対応しうる。選択ステップは、単に、画像修正によってどのパッチが処理されるかを決定することに対応する。

【0126】

続くステップ830は内容解析を含む。このステップでは、画像が解析されて、どのチャンネルが一次チャンネルとして選択されるべきであり、どのチャンネルが二次チャンネルとして処理されるべきかを決定する。次いで、ステップ840は、フィルタ調整を含み、フィルタ（画像修正構成を意味する）は、あるチャンネルを一次チャンネルとして、もう一つの（または二つ以上の）チャンネルを二次チャンネルとして扱うように構成（設定）される。これらの2つのステップ830および840は、この例示的な実施形態における画像チャンネル選択に対応する。図8では、内容（画像）解析830は、デコードされた画像内容の解析に基づいて、デコーダ側で（たとえば、画像またはビデオ・デコードの一部として）実行される。よって、この例示的な実施形態では、一次/二次チャンネルの設定に関するサイド情報は必要ない。図8においては、デコーダにおける動作が示されている。しかしながら、エンコーダでの動作は、同じ仕方で行うことができる：画像の再構成後に、色チャンネルの選択および対応する画像向上が、ループ・フィルタとして、またはポスト・フィルタとして実行されてもよい。

20

30

【0127】

ステップ850では、実際の画像修正が実行される。画像の修正は、パラメータ（図ではフィルタ・パラメータと称され、フィルタ調整の結果として得られる）に基づく。選択された、可能性としては再配置されたパッチが修正される。図8の修正850の後、ステップ810において画像再配置が適用された場合には、ピクセルは元通りに配置されてもよい860。（再）配置810および860は、パディングおよび/またはサブサンプリングを含んでもよい。（ピクセル再配置の）この文脈で使用される用語サブサンプリングは、画像品質損失を生じるサブサンプリングではないことに留意されたい。むしろ、上記の図5を参照して説明したように、サブサンプリングは、画像（パッチ）を、その画像（パッチ）の複数のサブサンプリングされたバージョンに分割するために実行される。

40

【0128】

図9は、画像修正を構成するために、サイド情報がエンコーダからデコーダに伝達される別の例示的な実施形態を示す。具体的には、図9は、エンコーダ側901およびデコーダ側902を示す。エンコーダ側では、ステップ910でもとの（歪んでいない）画像がエンコードされる。エンコードは、図1を参照して上述したように、たとえば量子化による不可逆圧縮を含んでもよい。もとの画像は、処理モジュール930にも入力される。ステップ/モジュール920は、ステップ820について上述したようにパッチ選択を実行する。次いで、ステップ/モジュール930は、画像（パッチ）を解析し、一次チャンネル（よって、

50

二次チャンネルも)の選択を実行した。デコーダ側902でデコードを可能にするために、サイド情報が、エンコーダ側901からデコーダ側902に伝達される。さらに、ステップ910は、ビットストリームに記憶されうるエンコードされた画像(ビデオ)を生成する。次いで、同じまたは異なるビットストリームが、ステップ930で生成されたフィルタ・パラメータを搬送することができる。

【0129】

デコーダ側902では、エンコードされた画像は、ステップ940でデコードされる。これは、たとえば、図2を参照して説明したデコーダのようなデコーダを用いて実行されてもよい。しかしながら、エンコーダもデコーダも、図1、図2、図3に示される例に限定されないことに留意されたい。一般に、ここに記載される画像修正は、再構成された画像に適用されるので、それは任意の画像に適用可能である。画像は、不可逆圧縮の結果である必要さえない。エンコーダおよびデコーダはまた、画像修正を実装してもよく、それらはハイブリッド・コーデックである必要がない。むしろ、人工知能ベースの(たとえば、ニューラルネットワークベースの)エンコーダおよびデコーダ、または他のエンコーダおよびデコーダが適用可能である。

10

【0130】

デコード940の結果は、歪んだ画像であってもよい。ここでデコードという用語は、必ずしも完全なデコードを意味するものではない。それは、むしろ、画像の再構成を指し、これは、ループにおいて、すなわち、ビデオ・デコードの場合には実際のデコードの間に実行されてもよい。次いで、歪んだ画像は、たとえば、ステップ810と同様の仕方で、再配置950される。ステップ960では、パッチが選択され(一般に、画像修正のための画像または画像部分が選択される)、これはステップ820で上述したように実装されてもよい。

20

【0131】

図8および図9を参照して説明したステップは、それぞれのステップを実行する装置の機能モジュールとしても見ることに留意されたい。装置は、プロセッサおよびメモリを含む一つまたは複数のハードウェアによって実装されてもよい。次いで、選択された画像データがステップ/モジュール970に渡される。モジュール970では、フィルタ(画像修正器)は、エンコーダ側から受領されたパラメータ(たとえば、ビットストリームで伝達されたパラメータ)によって調整(構成)される。次いで、ステップ980において、画像データは、適切に設定された画像修正器を用いてフィルタリング(修正)される。フィルタ・パラメータは、一次画像チャンネルの選択の指示を含んでいてもよいが、これに限定されない。画像修正980の後、ステップ940で画像再配置が適用された場合、ステップ990で、画像は、再びもとの形状および形に配置される。これは、上述のステップ860で説明のと同様に実行されうる。

30

【0132】

なお、図9では、デコーダの説明に焦点を当てている。しかしながら、エンコーダも、ステップ950および970~990を含んでいてもよい。換言すれば、エンコーダは、図3に概略的に示されるように、ループ内で、またはポスト・フィルタとして、画像修正を適用することもできる。たとえば、ステップ950は、ステップ920に先行してもよく、ステップ970~990は、ステップ930の後に実行されてもよい。

40

【0133】

図8および図9は、特に、画像修正処理がセットアップされる仕方によって互いに異なる。図8において、一次画像チャンネルの選択は、修正されるべき画像の画像解析に基づく。そのような解析は、エンコーダ側およびデコーダ側において、歪んでいない画像を知らずに、同じ仕方で実行されうる。他方、図9は、画像解析がオリジナル画像、すなわち歪み(たとえば、不可逆圧縮を含む画像エンコードによって生じる歪み)の前の画像に基づいて実行される手法を示している。よって、そのような情報は、もとの画像が知られていないデコーダ側では導出できない。よって、もとの画像に基づいて導出されたパラメータ(一次画像チャンネル選択など)は、信号伝達される。スタンドアローンの一次チャンネル決定

50

は、オリジナルが未知であるシナリオにおいても、より広範な用途を有する可能性があり、信号伝達によって生じるオーバーヘッドを節約する。他方、図9の手法は、同時に再構成の所望される目標でもあるもとの画像の知識を用いて画像チャンネル選択が実行されるので、より正確な再構成を提供しうる。換言すれば、画像修正は歪みのない画像、すなわちもとの画像に可能性としては近づくように、歪みのある画像を修正することを目的としている。

【0134】

図8は、もとの画像の参照が必要ない非参照実施形態と呼ばれる。図9は、もとの画像を考慮に入れた（完全）参照実施形態と呼ばれる。完全参照実施形態は2つの部分を有し、一方は画像/ビデオ・エンコーダ901の一部であり、他方は画像デコーダ902の一部である。エンコーダ901では、内容解析ブロックは、もとのデータと圧縮された（すなわち、歪められた）データの両方へのアクセスをもつ。パッチを選択し、該パッチをもとの画像における同じ場所からのパッチと比較することにより該パッチを解析することができる。最適なフィルタ・パラメータの選択は、エンコードされた画像と一緒に送られる（信号伝達される）。

10

【0135】

まとめると、本開示のいくつかの実施形態は、たとえば、チャンネル間相関に基づくCNNベースのフィルタを使用して、適応的画像向上を提供する。特に、処理パラメータが入力画像の解析に基づいて自動的に調整される。パラメータの調整〔チューニング〕は、サイド情報の信号伝達を必要とせずにデコーダ内で独立して行うことができる。さらに、トポロジーおよび処理順序は、ある範囲の諸入力画像フォーマットに合わせて調整されることができる。これについては、異なる画像チャンネルが異なる解像度（次元）を有する画像について、以下に示される。チャンネル当たりのピクセル数も、画像向上のために構成設定可能であり、他のパラメータの選択に影響を及ぼす。画像修正を構成する際に、画像チャンネルの数も考慮に入れることができる。

20

【0136】

例示的な実施形態によれば、2つ以上の画像チャンネルのうちから一次チャンネルおよび二次チャンネルを選択するステップ840、970は、分類器1050の出力に基づいて実行される。分類器（一般に、画像解析器）は、たとえば、2つ以上の画像チャンネルが入力されるニューラル・ネットワークを使用して実装される。

30

【0137】

図10は、例示的な画像分類器1050を示す。この例示的な画像分類器1050は、図8および図9の画像解析ステップ/モジュール830および930の一部であってもよい。左側では、歪んだチャンネルは、上述のピクセル再配置に対応する第1段に入力される。ピクセル再配置（たとえば、ピクセル・シフト）によって生成された、結果として得られるチャンネル・スタックは、深さZをもつ共通スタックに組み合わせられ、ここで、Zは各スタックの深さの和である。スタックは、N+1クラスを生成するようにトレーニングされたCNNベースの分類器に供給される。ここで、N個のクラスは「チャンネルNを一次として選択」の場合に対応し、最後のクラスは「このパッチの処理をスキップ」である。この特定の例における分類器は、共通のCNN分類器、たとえばVGG16分類器として実装され、これは、深さZの多層3Dアレイ入力に対して動作するように適応される。VGG16ベースの実装は単に例であり、他の任意の分類器が使用されうることを注意しておく。例解のために、画像解析1050を実行するために使用されうるVGG16分類器が、図11に示されている。

40

【0138】

具体的には、図11は、例示的に適応されたVGG16によって実行される処理を示す。第1段では、フル・スタックからパッチが選択される。このパッチは、 $32a \times 32a \times Z$ のサイズを有していてもよい。ここで、

【数2】

50

$a \in \mathbb{Z}$

であり、 Z は、フル・スタックの深さである（たとえば、YUV4:2:0の場合、 $Z=6$ ）。次いで、第2段では、 $3 \times 3 \times Z$ のカーネル・サイズで64個の畳み込みが実行される。各畳み込みは、次の段の1つの層を構築する。第3段では、6段階でパッチの平面サイズを $a \times a$ に小さくし、深さを4096に拡大する。その過程で、空間的特徴は、ますます一般的なパッチ特徴に変換される。これらの段は典型的なVGGスキームに従う。第4段では、最終層は一般的な特徴を $Z+1$ 個のクラスにプールの。最後に、段5では、ネットワークの出力はクラスであり、そのパッチについての最良な処理パラメータを指定する。たとえば、Primary = N（一次チャンネルはチャンネルN）またはFilter_off（画像向上を適用すべきでないことの指示）。

【0139】

図10は、機械学習に基づく例示的な分類器1050を示す。しかしながら、本開示は、そのような分類器に限定されるものではない。むしろ、VGG16ベースの分類器以外の分類器を意味する、任意の他の分類器を使用することができる。分類器は、ニューラル・ネットワークまたは畳み込みニューラル・ネットワークに基づくことができる。しかしながら、分類器はまた、詳細のレベル、および/またはエッジの強度および/または方向、勾配の分布、運動特性（ビデオの場合）または他の画像特徴を決定するアルゴリズムのようないくつかのアルゴリズムによって実装されてもよい。それぞれの画像チャンネルについてのそのような特徴の比較に基づいて、次いで、一次チャンネルが選択されうる。たとえば、一次チャンネルとして、最も多くのまたは最も鋭いエッジ（最も多くの詳細に対応する）を含む画像チャンネルが選択されてもよい。

【0140】

分類器1050は、 N 個の出力を含む $N+1$ 個の出力を提供し、該 N 個の出力は N 番目のチャンネルを一次チャンネルとする特徴の量を示す。このように、最も高い値の N を有するチャンネルが、次いで、一次チャンネルとして選択されうる。さらに、1つの出力（ $N+1$ 番目）が、そもそも画像修正が適用されるかどうかを示す。そのような判断を除外することが可能であることに留意されたい。その判断はまた、 N 個の特徴の分布および/または量に基づいて行うこともできる。あるいはまた、その判断は、分類器の適用に先立って行われ、圧縮パラメータ（たとえば、量子化の量、パッチのサイズ、画像の解像度、および/または他の圧縮パラメータ）などの追加情報に基づいてもよい。

【0141】

N 個の画像チャンネルを入力すると、選択されるべきチャンネルの数を与える分類器が適用されうることに留意されたい。換言すれば、本開示は、図10に示される分類器1050のような特定のまたは類似の分類器に限定されない。

【0142】

分類器1050の適用後、フィルタ調整ステップ/モジュール840、970は、単に、以下のいずれかのプログラム論理を含んでもよい：

a) パッチ解析の結果が「チャンネル N を一次として選択」である場合、入力スタック N を一次チャンネルとして処理し、他のスタックを二次チャンネルとして処理するようルーティングし直す。

b) パッチ解析の結果が「このパッチの処理をスキップ」である場合、入力スタックは、ピクセル配置ブロックに直接供給される。歪んだ画像がバッファリングされている場合、歪んだ画像のコピーを出力に直接提供することが考えられ、それにより、再配置910（10）を反転する配置990を実行する必要がない。

【0143】

上述のように、前記2つ以上の画像チャンネルは、色チャンネルおよび/または特徴チャネ

10

20

30

40

50

ルを含む。上述のように、本開示の画像修正は、色チャンネルに適用可能である。しかしながら、色チャンネルに加えて、または代替的に、画像修正は特徴チャンネルに適用可能である。特徴チャンネルは、たとえば、画像捕捉深さ、画像（またはテンソル）捕捉オプティカルフロー、または任意の他のサンプル関連画像特徴でありうる。

【0144】

上述のパッチに対応しうる画像領域に関して、画像領域は、以下のいずれかでありうる：

- ・画像の一部または複数の画像の一部に対応する所定のサイズのパッチ、または
- ・画像または複数の画像。

【0145】

特に、画像チャンネルの選択は、画像ごとに行われてもよい。これは、ビデオ・シーケンスのフレームまたはフィールドである画像ごとに行われてもよい。それは、ビデオ・シーケンスのある数の画像（ピクチャーのグループなど）について行われてもよい。画像チャンネルの選択は、パッチ毎に、または一般に画像領域毎に行われてもよい。どの粒度で（上記のサンプル量のうちどれについて）色チャンネルが選択されるかの決定が、エンコーダで行われてもよい。そのような決定は、デコーダに示されてもよい。あるいはまた、エンコーダおよびデコーダは、エンコーダおよびデコーダの両方に既知の他の符号化パラメータに基づいてチャンネル選択粒度の決定を実行することができる。

【0146】

画像修正の種類に応じて、パッチ・サイズ（修正される画像領域のサイズ）を選択する際に有利に観察される基準がいくつかある場合がある。特に、画像修正の間に適用される処理のタイプによって与えられる何らかの最小サイズがありうる。例示的な実装によれば、パッチ・サイズの選択は、ニューラル・ネットワークの隠れ層の数に基づいて画像領域の最小サイズを選択することを含む。特に、最小サイズは少なくとも $2 * ((kernel_size - 1) / 2 * n_layers) + 1$ であり、ここで $kernel_size$ は畳み込みニューラル・ネットワークであるニューラル・ネットワークのカーネルのサイズであり、 n_layers はニューラル・ネットワークの層の数である。この制限は、カーネルとの畳み込みが複数の値に基づいており、1つの値のみを生成しうるという事実からくる。たとえば、 3×3 カーネルをもつ畳み込みがサイズ 5×5 の画像部分に適用される場合、結果として得られる特徴マップは 3×3 のサイズをもつ。各特徴点が入力（画像）データに基づいて完全に計算する場合。よって、カーネル 3×3 による畳み込みが適用される各層では、畳み込みされる画像（特徴マップ）のサイズは、各次元（垂直および水平）において2だけ小さくなる。パディングを使用することができるが、パディングされた入力から得られる特徴マップによって提供される情報の量はより少ないことに留意されたい。

【0147】

効率的な処理（画像修正）を可能にするために、上述の実施形態および例のいずれかと組み合わせて使用されうる実施形態に従って、画像領域の前記少なくとも2つの画像チャンネルの各チャンネルのピクセルを、複数のS個の部分領域に再配置する。さらに、前記少なくとも2つの画像チャンネルのうちの画像チャンネルの部分領域のそれぞれは、前記画像チャンネルのサンプルのサブセットを含む。すべての画像チャンネルについて、部分領域の水平寸法は同じであり、画像の水平寸法の最大公約数の整数倍 mh に等しい。すべての画像チャンネルについて、部分領域の垂直寸法は同じであり、画像の垂直寸法の最大公約数の整数倍 mv に等しい。ここで、「整数」倍数は、1、2、3、またはそれ以上でありうることに留意されたい。言い換えれば、整数倍数は1であってもよい。

【0148】

そのようなパッチ・サイズの選択は、可能性としては異なる画像チャンネルについては異なり、図12に示されている。図12において、3つのチャンネルは異なるサイズを有する。ピクセル再配置ブロックは、入力ピクセルを3Dアレイに配置する。寸法が共通の整数因数をもつとすれば、異なる寸法をもつ複数のチャンネルを扱うことができる。入力はN個の平面画像であり、Nはチャンネルの数である（ここでは3）。出力は、同一の高さおよび幅（決定される）をもち、可能性としては異なる深さ（すなわち、z方向における異なる層数）をも

10

20

30

40

50

つN個の3Dアレイである。

【0149】

最初の段階は、各画像の水平寸法と垂直寸法についての共通因子 q 、 p を見つけることである。次いで、水平ステップ x_step と垂直ステップ y_step が各チャンネルについて定義される。ステップは、各次元における因数分解における第2の因子である。次の段は、オフセットのすべての組み合わせについて、座標 $\{k \cdot x_step + x_offset, k \cdot y_step + y_offset\}$ をもつピクセルを使って m 個の画像を作成する。ここで、 $m = x_step \cdot y_step$ であり、

【数3】

$$k \in \mathbb{Z}, x_offset \in [0, x_step); y_offset \in [0, y_step)$$

10

である。ここで、「 $[$ 」は、境界数（ここでは0）が範囲に属することを意味する閉区間を指す。記号「 $)$ 」は、开区間を指し、境界数（ここでは x_step 、 y_step ）が区間に属さないことを意味する。これは、 $x_offset = 0..x_step - 1$ 、 $y_offset = 0..y_step - 1$ と書くのと等価である。1つのチャンネルについて、ピクセル並べ替えブロックの出力は、寸法 $q \times p \times m$ をもつ値のスタック（3Dアレイ）である。出力スタックの数は、入力チャンネルの数に等しい。

【0150】

図12の上記で説明した構成において、整数倍は、各チャンネルのサンプルを複数のサブ画像（サブサンプリングされた画像）に分割するためのステップ・サイズとして用いられる。そのような再配置の後、互いに近接したピクセルは近接したままであり、いくつかのピクセル間の x および y 距離は減少し、 z 距離は増加する。このようにして、近接した近傍からのピクセルは、いまだ1つの畳み込みカーネル内に収まる。これは、フィルタ（画像を修正するために使用されるニューラル・ネットワーク）の適切な動作のために有益である。

20

【0151】

言い換えれば、一般に、画像領域の S 個の部分領域は、 $S = mh \cdot mv$ であり、分離しており、水平寸法 $dimh$ と垂直寸法 $dimv$ を有する。部分領域は、位置 $\{kh \cdot mh + offh, kv \cdot mv + offv\}$ の画像領域のサンプルを含み、 $kh \in [0, dimh - 1]$ 、 $kv \in [0, dimv - 1]$ である。 $offh$ と $offv$ の各組み合わせは、それぞれの部分領域を指定する。ここで、 $offk \in [1, mh]$ および $offv \in [1, mv]$ である。

30

【0152】

図13は、チャンネル選択が実行された後のマルチチャンネル処理の可能な動作をまとめたものである。向上フィルタに関わる $N + 1$ 個の畳み込みニューラル・ネットワークがある。4つのネットワーク1310、1320、1330、1350のうち、ネットワーク(S)としてマークされた2つのネットワーク1310および1350は、単一入力モードで動作し、つまり、入力として、1つの画像チャンネルを有する。残りの2つのネットワーク1320および1330（ネットワーク(C)としてマークされている）は、二重入力モードで動作し、つまり、入力として、1つの画像チャンネルだけでなく、さらに、処理された一次チャンネル、すなわち単一入力ネットワーク1310によって処理されたチャンネルをも有する。処理された一次チャンネルは、すべての二重入力CNN、すなわち、それぞれの二次チャンネルを処理するすべてのネットワークの補助入力として使用される。各二次チャンネルは、そのような二重入力CNNで処理される。本例では、2つの画像チャンネルは、ネットワーク1320および1330によって処理される。二次チャンネルは、ネットワークのメイン入力（1320または1330）として使用される。処理されたすべてのチャンネルは、モジュール1340内の単一のスタックに連結され、次いで、別の単一入力CNN 1350によって処理される。最後に、最初に再配置が適用されていた場合に、入力画像の寸法およびチャンネル数に一致するように、ピクセルは再び再配置1360される。

40

【0153】

図4を参照して論議されたスキームと比較すると、この実施形態では、図13のネットワ

50

ーク1310、1320、1330、および1350は、異なる深さを有するスタックを収容するように適応される（対応する畳み込みカーネルのサイズを変更することによって、特に、畳み込みカーネルの深さを変更することによって）。

【0154】

上述したように、特に図1～図3、図5、図8、および図9を参照すると、画像修正は、画像またはビデオの符号化および復号化に適用されうる。たとえば、画像または画像のビデオ・シーケンスをエンコードするための方法が提供される。この方法は、もとの画像領域を得て、得られた画像領域をビットストリームにエンコードし、エンコードされた画像領域を再構成することによって得られた画像領域を修正するための前記方法を適用することを含む。修正は、一次画像チャンネルの選択を含め、上述のように実行されてもよい。

10

【0155】

この方法は、選択された一次チャンネルの指示をビットストリームに含めることを含んでいてもよい。これは、図9（フィルタ・パラメータ）に例示された。

【0156】

いくつかの実施形態では、本方法は、複数の画像領域（たとえば、パッチ）を得て、得られた複数の画像領域の画像領域に、個々に、得られた画像領域を修正するための前記方法を適用することを含んでいてもよい。これは、時間的に直列的であってもよいし、並列的であってもよい。さらに、この方法は、前記複数の画像領域のそれぞれについて、ビットストリーム中に、（i）得られた画像領域を修正するための前記方法がその画像領域に適用されないことを示す指示、（ii）その領域についての選択された一次チャンネルの指示のうち少なくとも1つを含めることを含む。オプション（i）は、図面の1050の分類器による決定、すなわち、画像領域が処理されるべきでないことを示す（N+1）番目の出力の結果であってもよい。しかしながら、それは、異なる仕方で実行されうる決定の結果であってもよく、またはユーザーまたはアプリケーションによって構成されてもよい。

20

【0157】

上記の画像領域は、上記の詳細な例に記載されるパッチに対応しうる。並べ替えられた3Dアレイからパッチが選択される。パッチは各チャンネルの同じ場所から取られ、処理される。すでに前述したように、パッチの各側は $2 \cdot (P + 2)$ より大きいべきである。ここで、Pは向上フィルタにおける隠れ層の数である。十分なメモリがあれば、すべての3Dアレイは、単一のパスで完全に処理されることができ。（再配置された）パッチ・サイズは、 128×128 ピクセル、またはいくつかの例示的な実装ではそれよりも大きいものであってもよい。しかしながら、画像解像度、所望の処理速度、並列処理の利用可能性等に依存して、パッチはより小さくすることもできる。

30

【0158】

前記画像または前記画像のビデオ・シーケンスをエンコードする方法で、得られた画像領域を修正する前記方法を適用する際に、エンコード・ステップに入力された前記得られた画像領域を参照することなく、前記再構成された画像領域に基づいて前記一次チャンネルおよび前記二次チャンネルの選択が実行される。これは、図8を参照して上述した実施形態に対応するが、図8はデコーダを示すので、図8には示されていない。しかしながら、再構成後のエンコーダとデコーダの動作は類似している。

40

【0159】

本開示はさらに、ビットストリームから画像または画像のビデオ・シーケンスをデコードするための方法であって、ビットストリームから画像領域を再構成し、上述のような画像領域を修正するための方法を適用することを含む、方法を提供する。これは依然としてループ・フィルタまたはポスト・フィルタでありうる。両方とも再構成された画像領域に基づいているからである。

【0160】

デコード方法は、ビットストリームをパースして、（i）得られた画像領域を修正する前記方法がその画像領域について適用されないことを示す指示、（ii）その領域についての選択された一次チャンネルの指示のうち少なくとも1つを取得することをさらに含んでい

50

てもよい。この方法は、ビットストリームから画像領域を再構成することをさらに含む。指示が選択された一次チャンネルを示す場合、指示された一次チャンネルを前記選択された一次チャンネルとして、再構成された画像領域を修正する。これは、図9を参照して記載された完全参照実施形態に対応する。

【0161】

さらに、2つ以上の画像チャンネルによって表される画像領域を修正するための装置が提供され、量子化装置は、画像修正のための上述の方法のいずれかによるステップを実行するように構成された回路によって実装される。

【0162】

図14は、例示的な符号化システム10、たとえば、本開示の技術を利用することができるビデオ符号化システム10（または略して符号化システム10）を示す概略ブロック図である。ビデオ符号化システム10のビデオ・エンコーダ20（または略してエンコーダ20）およびビデオ・デコーダ30（または略してデコーダ30）は、本願に記載されるさまざまな例による技術を実行するように構成されうる装置の例を表す。特に、エンコーダ20は、図1または図5に示されるエンコーダに対応しうる。デコーダ30は、図2に示されるデコーダ等に対応してもよい。

10

【0163】

たとえば、エンコーダ20のようなエンコーダは、もとの画像領域を得るための入力モジュールと、得られた画像領域をビットストリームにエンコードするための圧縮モジュールと、エンコードされた画像領域を再構成するための再構成モジュールと、上述のような再構成された画像領域を修正するための装置とを含んでいてもよい。

20

【0164】

さらに、デコーダ（デコーダ30のような）は、ビットストリームから画像または画像のビデオ・シーケンスをデコードするために提供されてもよく、デコーダは、ビットストリームから画像領域を再構成するための再構成モジュールと、再構成された画像領域を修正するための装置とを含む。

【0165】

図14に示されるように、符号化システム10は、エンコードされたピクチャー・データ13をデコードするために、エンコードされたピクチャー・データ21をたとえば宛先装置14に提供するように構成されたソース装置12を備える。ソース装置12は、エンコーダ20を含み、追加的に、すなわち任意に、ピクチャー・ソース16、プリプロセッサ（または前処理ユニット）18、たとえばピクチャー・プリプロセッサ18、および通信インターフェースまたは通信ユニット22を含んでいてもよい。ピクチャー・ソース16は、任意の種類のピクチャー捕捉装置、たとえば実世界のピクチャーを捕捉するためのカメラ、および/または任意の種類のピクチャー生成装置、たとえばコンピュータ・アニメーション・ピクチャーを生成するためのコンピュータ・グラフィックス・プロセッサ、または実世界のピクチャー、コンピュータ生成ピクチャー（たとえば、スクリーン・コンテンツ、仮想現実（VR）ピクチャー）、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（AR）ピクチャー）を取得および/または提供するための任意の種類の他の装置を含んでいてもよく、またはかかる装置であってもよい。ピクチャー・ソースは、上記のピクチャーのいずれかを記憶する任意の種類のメモリまたは記憶装置であってもよい。

30

40

【0166】

プリプロセッサ18や前処理ユニット18によって実行される処理と区別して、ピクチャーまたはピクチャー・データ17は、生ピクチャーまたは生ピクチャー・データ17と称されてもよい。プリプロセッサ18は、（生）ピクチャー・データ17を受領し、ピクチャー・データ17に対して前処理を実行して、前処理済みのピクチャー19または前処理済みのピクチャー・データ19を得るように構成される。プリプロセッサ18によって実行される前処理は、たとえば、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んでいてもよい。前処理ユニット18は、任意的な構成要素であってもよいことが理解できる。上述の実施形態および例に記載される画像修

50

正は、画像の前処理として適用されてもよい。

【0167】

ビデオ・エンコーダ20は、前処理されたピクチャー・データ19を受領し、エンコードされたピクチャー・データ21を提供するように構成される。ソース装置12の通信インターフェース22は、エンコードされたピクチャー・データ21を受領し、記憶または直接再構成のために、通信チャンネル13を通じて、エンコードされた画像データ21（またはその任意のさらに処理されたバージョン）を別の装置、たとえば宛先装置14または任意の他の装置に送信するように構成されてもよい。宛先装置14は、デコーダ30（たとえば、ビデオデコーダ30）を含み、追加的に、すなわち任意的に、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32（または後処理ユニット32）、および表示装置34を含んでいてもよい。

10

【0168】

宛先装置14の通信インターフェース28は、エンコードされたピクチャー・データ21（またはそのさらに処理されたバージョン）を、たとえばソース装置12から直接、または他の任意のソース、たとえば記憶装置、たとえばエンコード・ピクチャー・データ記憶装置から受領し、エンコードされたピクチャー・データ21をデコーダ30に提供するように構成される。通信インターフェース22および通信インターフェース28は、ソース装置12と宛先装置14との間の直接通信リンク、たとえば直接の有線もしくは無線接続を介して、または任意の種類のネットワーク、たとえば有線もしくは無線ネットワークもしくはそれらの任意の組み合わせ、または任意の種類の私設および公共ネットワーク、またはそれらの任意の種類の組み合わせを介して、エンコードされたピクチャー・データ21またはエンコードされたデータ13を送信または受領するように構成されてもよい。

20

【0169】

通信インターフェース22は、たとえば、エンコードされたピクチャー・データ21を適切なフォーマット、たとえばパケットにパッケージし、および/または任意の種類の送信エンコードまたは通信リンクもしくは通信ネットワークを通じた送信のための処理を用いて、エンコードされたピクチャー・データを処理するように構成されてもよい。通信インターフェース22の相手をなす通信インターフェース28は、たとえば、送信されたデータを受信し、任意の種類の対応する送信デコードまたは処理および/またはデパッキングを用いて送信データを処理し、エンコードされたピクチャー・データ21を得るように構成されてもよい。通信インターフェース22および通信インターフェース28は両方とも、ソース装置12から宛先装置14を指す、図14における通信チャンネル13についての矢印によって示されるように、単方向通信インターフェースとして、または双方向通信インターフェースとして構成されることができ、たとえばメッセージを送受信し、たとえば接続を確立し、通信リンクおよび/またはデータ伝送、たとえばエンコードされたピクチャーに関連する任意の他の情報を受け取り確認し、交換するように構成されてもよい。

30

【0170】

デコーダ30は、エンコードされたピクチャー・データ21を受領し、デコードされたピクチャー・データ31またはデコードされたピクチャー31を提供するように構成される。宛先装置14のポストプロセッサ32は、デコードされたピクチャー・データ31（再構成されたピクチャー・データとも呼ばれる）、たとえばデコードされたピクチャー31を後処理して、後処理されたピクチャー・データ33、たとえば後処理されたピクチャー33を得るように構成される。後処理ユニット32によって実行される後処理は、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、または、たとえば表示装置34による表示のために、デコードされたピクチャー・データ31を準備するためなどの他の任意の処理を含むことができる。

40

【0171】

宛先装置14の表示装置34は、ピクチャーをたとえばユーザーまたは閲覧者に対して表示するために、後処理されたピクチャー・データ33を受領するように構成される。表示装置34は、再構成されたピクチャーを表現するための任意の種類のディスプレイ、たとえば

50

、統合されたまたは外部のディスプレイまたはモニターであってもよく、またはそれらを含んでいてもよい。ディスプレイは、たとえば、液晶ディスプレイ（LCD）、有機発光ダイオード（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（LCoS）、デジタル光プロセッサ（DLP）、または任意の他の種類のディスプレイを含むことができる。

【0172】

図14は、ソース装置12および宛先装置14を別個の装置として描いているが、装置の実施形態は、両方もしくは両方の機能、ソース装置12もしくは対応する機能および宛先装置14もしくは対応する機能を含んでいてもよい。そのような実施形態では、ソース装置12または対応する機能と宛先装置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して、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によって実装されることができる。前記記述に基づいて当業者には明白であろうように、図14に示されるようなソース装置12および/または宛先装置14内の種々のユニットまたは機能の存在および（厳密な）機能分割は、実際の装置および用途に依存して変わりうる。

【0173】

エンコーダ20（たとえば、ビデオ・エンコーダ20）、またはデコーダ30（たとえば、ビデオ・デコーダ30）、またはエンコーダ20とデコーダ30の両方は、図15に示されるような処理回路、たとえば、一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（DSP）、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）、ディクリット論理、ハードウェア、専用のビデオ・コーディング、またはそれらの任意の組み合わせを介して実装されうる。エンコーダ20は、図1のエンコーダ20および/または本明細書に記載される任意の他のエンコーダ・システムまたはサブシステムに関して論じられるさまざまなモジュールを具現するために、処理回路46を介して実装されてもよい。デコーダ30は、図2のデコーダ30および/または本明細書に記載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に関して論じられるさまざまなモジュールを具現するために、処理回路46を介して実装されてもよい。処理回路は、後述するように、さまざまな動作を実行するように構成されうる。図17に示されるように。技術が部分的にソフトウェアで実装される場合、装置が、好適な非一時的なコンピュータ読み取り可能な記憶媒体に該ソフトウェアのための命令を記憶してもよく、一つまたは複数のプロセッサを使用してハードウェアにおいて該命令を実行して、本開示の技術を実行してもよい。ビデオ・エンコーダ20とビデオ・デコーダ30のいずれかも、たとえば図15に示されるように、組み合わせられたエンコーダ/デコーダ（コーデック）の一部として単一の装置内に統合されてもよい。

【0174】

ソース装置12および宛先装置14は、任意の種類のハンドヘルドもしくは固定装置、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、表示装置、デジタルメディアプレーヤー、ビデオゲーム機、ビデオストリーミング装置（コンテンツサービスサーバーまたはコンテンツ送達サーバーなど）、放送受信機装置、放送送信機装置などを含む、広範囲の装置の任意のものを含むことができ、オペレーティング・システムを使用しなくてもよく、あるいは任意の種類のオペレーティング・システムを使用してもよい。場合によっては、ソース装置12および宛先装置14は、無線通信用のために装備されてもよい。よって、ソース装置12および宛先装置14は、無線通信装置であってもよい。

【0175】

場合によっては、図14に示されるビデオ・コーディング・システム10は、単に例であり、本願の技術は、必ずしもエンコード装置とデコード装置との間のデータ通信を含まないビデオ・コーディング場面（たとえば、ビデオ・エンコードまたはビデオ・デコード）に適用されうる。他の例では、データはローカルメモリから取り出される、ネットワーク

10

20

30

40

50

を通じてストリーミングされる、などされる。ビデオ・エンコード装置は、データをエンコードし、メモリに格納することができ、および/またはビデオ・デコード装置は、メモリからデータを取り出し、デコードすることができる。いくつかの例では、エンコードおよびデコードは、互いに通信せず、単にデータをメモリにエンコードする、および/またはデータをメモリから取り出してデコードする装置によって実行される。記述の便宜上、本発明の実施形態は、本明細書では、たとえば、高効率ビデオ・コーディング (HEVC) または多用途ビデオ・コーディング (VVC) の参照ソフトウェア、ITU-Tビデオ・コーディングエキスパートグループ (VCEG) およびISO/IEC動画像エキスパートグループ (MP EG) のビデオ・コーディングに関する統合協働チーム (Joint Collaboration Team on Video Coding, JCT-VC) によって開発された次世代ビデオ・コーディング規格を参照して記載される。当業者は、本発明の実施形態がHEVCまたはVVCに限定されないことを理解するであろう。

10

【0176】

図16は、本開示のある実施形態によるビデオ・コーディング装置400の概略図である。ビデオ・コーディング装置400は、本明細書に記載される開示される実施形態を実装するのに好適である。ある実施形態では、ビデオ・コーディング装置400は、図14のビデオ・デコーダ30のようなデコーダ、または図14のビデオ・エンコーダ20のようなエンコーダであってもよい。ビデオ・コーディング装置400は、データを受領するための入口ポート410 (または入力ポート410) および受領器ユニット (Rx) 420; データを処理するためのプロセッサ、論理ユニット、または中央処理ユニット (CPU) 430; データを送信するための送信器ユニット (Tx) 440および出口ポート450 (または出力ポート450); およびデータを記憶するためのメモリ460を含む。ビデオ・コーディング装置400は、光信号または電気信号の入出りのために、入口ポート410、受領器ユニット420、送信器ユニット440、および出口ポート450に結合された光対電気 (OE) コンポーネントおよび電気対光 (EO) コンポーネントをも有していてもよい。

20

【0177】

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、一つまたは複数のCPUチップ、コア (たとえば、マルチコアプロセッサとして)、FPGA、ASIC、およびDSPとして実装されてもよい。プロセッサ430は、入口ポート410、受領器ユニット420、送信器ユニット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を有する。コーディングモジュール470は、上記の開示された諸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、さまざまなコーディング動作を実装、処理、準備、または提供する。よって、コーディングモジュール470を含めることにより、ビデオ・コーディング装置400の機能が実質的に改善され、ビデオ・コーディング装置400の異なる状態への変換が実現される。あるいはまた、コーディングモジュール470は、メモリ460に記憶された命令として実装され、プロセッサ430によって実行される。

30

【0178】

メモリ460は、一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んでいてもよく、オーバーフロー・データ記憶装置として使用され、プログラムを、該プログラムが実行のために選択されるときに記憶し、そしてプログラム実行中に読み出される命令およびデータを記憶してもよい。メモリ460は、たとえば、揮発性および/または不揮発性であってもよく、リードオンリーメモリ (ROM)、ランダムアクセスメモリ (RAM)、三値連想記憶メモリ (TCAM)、および/またはスタティックランダムアクセスメモリ (SRAM) であってもよい。

40

【0179】

図17は、ある例示的实施形態による、図14からのソース装置12および宛先装置14のいずれかまたは両方として使用されうる装置1700の単純化されたブロック図である。装置1700内のプロセッサ1702は、中央処理ユニットであってもよい。あるいはまた、プロセッサ1702は、現在存在する、または今後開発される情報を操作または処理することがで

50

きる、任意の他のタイプの装置または複数の装置であってもよい。開示された実装は、図示のような単一のプロセッサ、たとえばプロセッサ502を用いて実施できるが、複数のプロセッサを用いて、速度および効率における利点が達成できる。

【0180】

装置1700内のメモリ1704は、ある実装では、リードオンリーメモリ（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できる。メモリ1704として、任意の他の好適なタイプの記憶デバイスが使用できる。メモリ1704は、バス1712を使用してプロセッサ1702によってアクセスされるコードおよびデータ1706を含むことができる。メモリ1704は、さらに、オペレーティング・システム1708およびアプリケーション・プログラム1710を含むことができ、該アプリケーション・プログラム1710は、プロセッサ1702がここに記載される方法を実行することを許容する少なくとも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1710は、アプリケーション1~Nを含むことができ、これは、ここに記載される方法を実行するビデオ・コーディングアプリケーションをさらに含む。特に、非一時的媒体に記憶されたプログラム・コードを含むコンピュータ・プログラム・プロダクトが提供されてもよい。該プログラムは、一つまたは複数のプロセッサ上で実行されたときに、上記のような画像修正のための方法を実行する。

10

【0181】

装置1700は、ディスプレイ1718などの一つまたは複数の出力装置をも含むことができる。ディスプレイ1718は、一例では、ディスプレイを、タッチ入力を検知するように動作可能なタッチ感応性要素と組み合わせるタッチ感応性ディスプレイであってもよい。ディスプレイ1718は、バス1712を介してプロセッサ1702に結合されることができる。ここでは単一のバスとして示されているが、装置1700のバス1712は、複数のバスから構成されることができる。さらに、二次記憶1714は、装置1700の他のコンポーネントに直接結合されることができるか、またはネットワークを介してアクセスされることができ、メモリカードのような単一の集積ユニットまたは複数のメモリカードのような複数のユニットを含むことができる。よって、装置1700は、幅広い多様な構成で実装されることができる。

20

【0182】

まとめると、本開示は、画像向上などの画像修正に関する。画像向上は、任意の画像修正のために適用されることができ、画像エンコードおよび/またはデコードの間または後に、たとえばループ・フィルタまたはポスト・フィルタとして適用されうる。特に、画像修正は、一次チャンネルが別個に処理され、処理された一次チャンネルに基づいて二次チャンネルが処理されるマルチチャンネル処理を含む。処理はニューラル・ネットワークに基づいている。画像修正性能を向上させるために、修正を適用する前に、画像チャンネルが解析され、一次チャンネルおよび二次チャンネルが決定される。これは、複数の画像、画像または画像領域について変わりうる。

30

40

50

【図面】
【図 1】

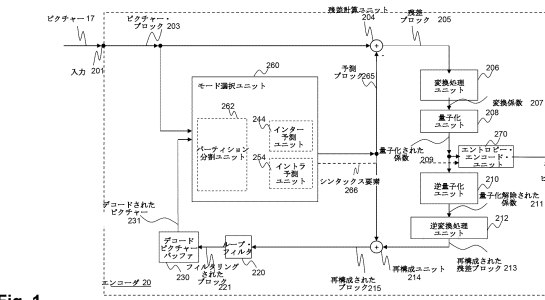


Fig. 1

【図 2】

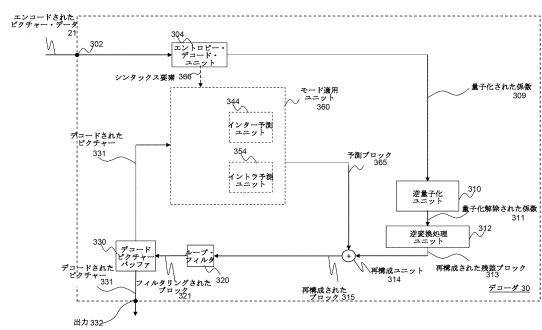


Fig. 2

10

【図 3 A】

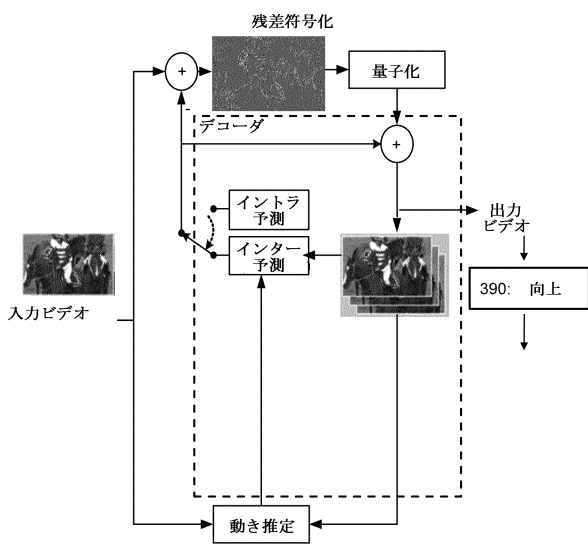


Fig. 3A

【図 3 B】

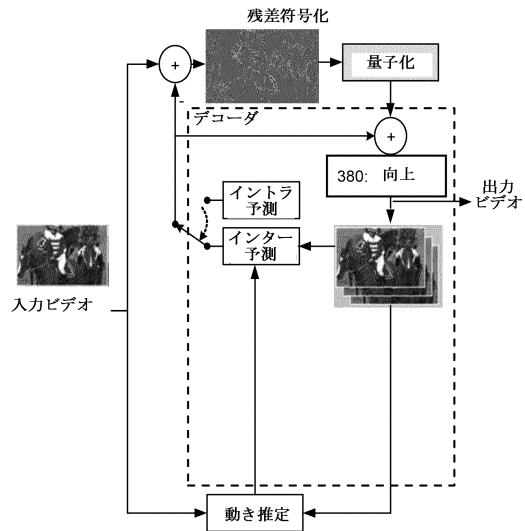


Fig. 3B

20

30

【図 4】

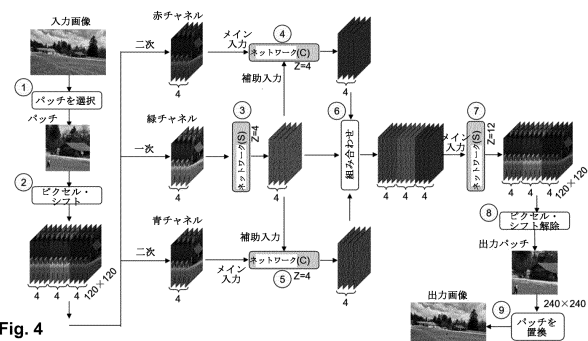


Fig. 4

【図 5】

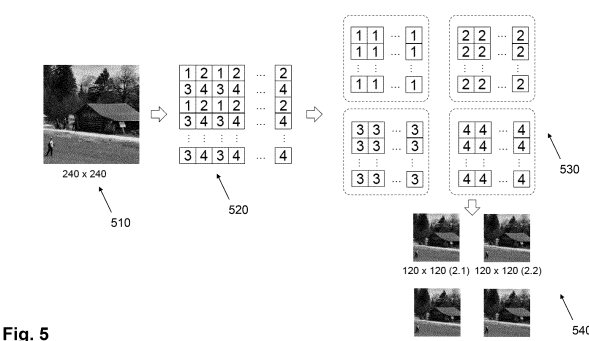


Fig. 5

40

50

【図 6】

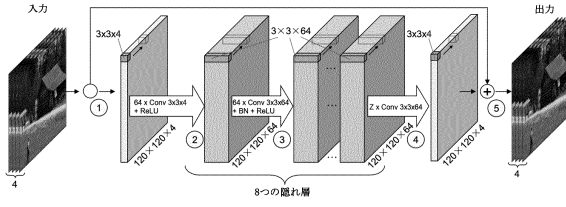


Fig. 6

【図 7】

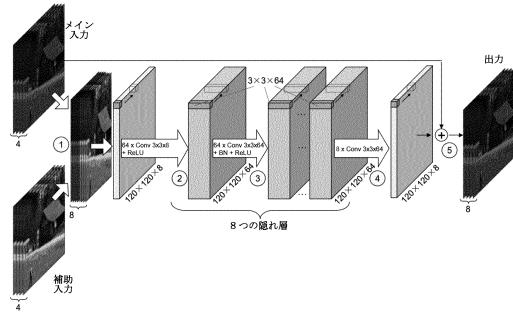


Fig. 7

【図 8】

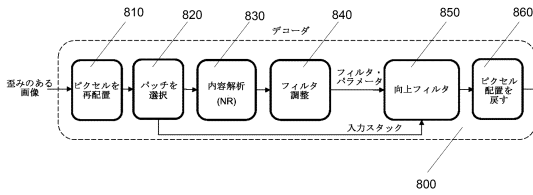


Fig. 8

【図 9】

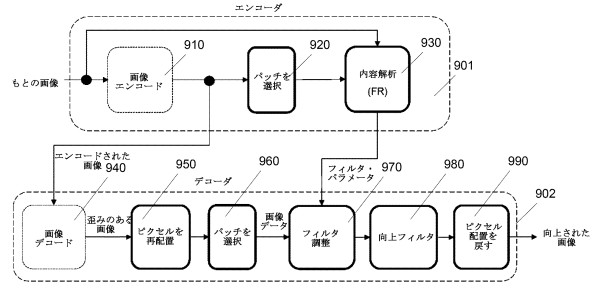


Fig. 9

【図 10】

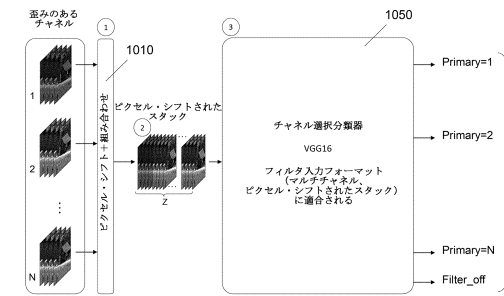


Fig. 10

【図 11】

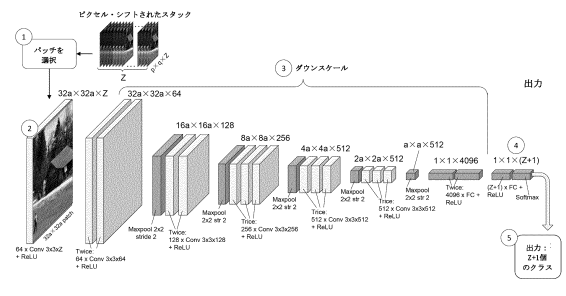


Fig. 11

10

20

30

40

50

【 図 1 2 】

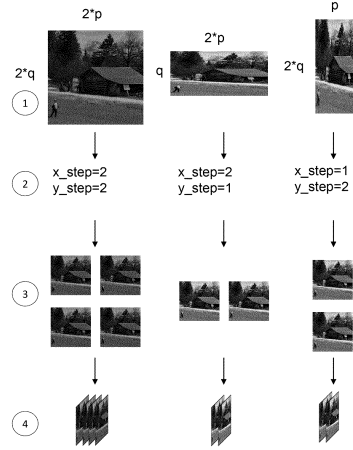


Fig. 12

【 図 1 3 】

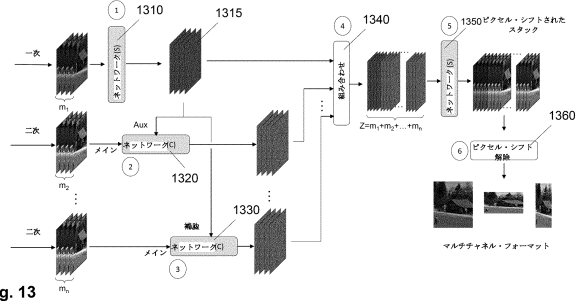


Fig. 13

10

【 図 1 4 】

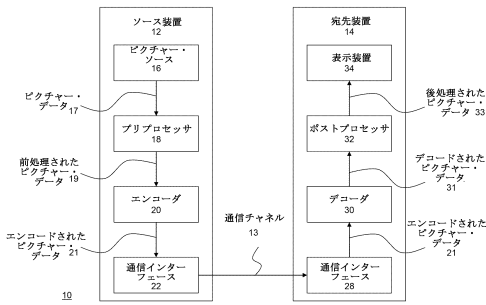


Fig. 14

【 図 1 5 】

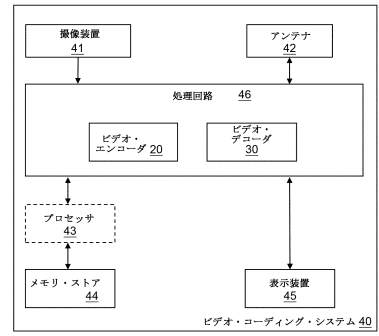


Fig. 15

20

30

40

50

【図 16】

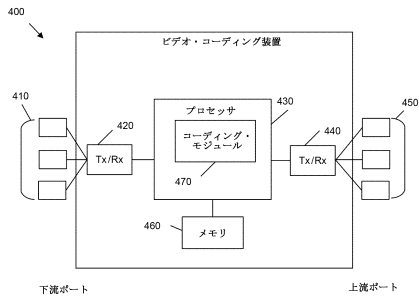


Fig. 16

【図 17】

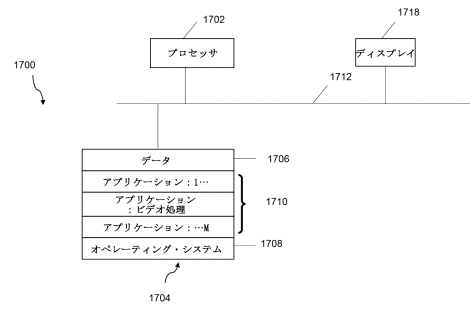


Fig. 17

10

20

30

40

50

フロントページの続き

- (74)代理人 100070150
弁理士 伊東 忠彦
- (74)代理人 100135079
弁理士 宮崎 修
- (72)発明者 キュイ, カイ
ドイツ 8 0 3 3 3 ミュンヘン アルツィスシュトラッセ 2 1 テクニッシュ・ユニヴェルジテート・ミュンヘン
- (72)発明者 ボエフ, アタナス
ドイツ 8 0 9 9 2 ミュンヘン リースシュトラッセ 2 5 ホアウエイ・テクノロジーズ・デュッセルドルフ・ゲーエムベーハー内
- (72)発明者 アルシナ, エレナ アレキサンドロヴナ
ドイツ 8 0 9 9 2 ミュンヘン リースシュトラッセ 2 5 ホアウエイ・テクノロジーズ・デュッセルドルフ・ゲーエムベーハー内
- (72)発明者 スタインバック, エッケハルト
ドイツ 8 0 3 3 3 ミュンヘン アルツィスシュトラッセ 2 1 テクニッシュ・ユニヴェルジテート・ミュンヘン
- 審査官 間宮 嘉誉
- (56)参考文献 特開 2 0 1 2 - 3 4 0 3 6 (J P , A)
特開 2 0 1 9 - 1 4 6 2 1 8 (J P , A)
国際公開第 2 0 1 9 / 0 7 2 0 9 7 (W O , A 1)
米国特許出願公開第 2 0 1 9 / 0 2 7 3 9 4 8 (U S , A 1)
特開 2 0 1 9 - 2 0 1 2 5 6 (J P , A)
CUI, Kai et al. , Decoder Side Image Quality Enhancement Exploiting Inter-Channel Correlation in a 3-Stage CNN: Submission to CLIC 2018 , CVPR 2018 Open Access Repository: Web Archive , web.archive.org , 2018年06月25日 , pp.2571-2574 , [online] , [retrieved on 2024-02-29] , Retrieved from the Internet: URL: http://web.archive.org/web/20180625231354/https://openaccess.thecvf.com/content_cvpr_2018_workshops/w50/html/Cui_Decoder_Side_Image_CVPR_2018_paper.html , Cui_Decoder_Side_Image_CVPR_2018_paper.pdf
- (58)調査した分野 (Int.Cl. , D B 名)
H 0 4 N 7 / 1 2
H 0 4 N 1 9 / 0 0 - 1 9 / 9 8
I E E E X p l o r e