

明 細 書

発明の名称：

復号装置、符号化装置、ビットストリーム生成装置、復号方法及び符号化方法

技術分野

[0001] 本開示は、復号装置等に関する。

背景技術

[0002] ビデオコーディング技術は、H. 261およびMPEG-1から、H. 264/AVC (Advanced Video Coding)、MPEG-LA、H. 265/HEVC (High Efficiency Video Coding)、およびH. 266/VVC (Versatile Video Codec) へ進歩している。この進歩に伴い、様々な用途において増え続けるデジタルビデオデータ量进行处理するために、ビデオコーディング技術の改良および最適化を提供することが常に必要とされている。本開示は、ビデオコーディングにおけるさらなる進歩、改良および最適化に関する。

[0003] なお、非特許文献1は、上述されたビデオコーディング技術に関する従来の規格の一例に関する。また、非特許文献2は、上述されたビデオコーディング技術に関する新たな提案に関する。

先行技術文献

非特許文献

[0004] 非特許文献1：H. 265 (ISO/IEC 23008-2 HEVC) / HEVC (High Efficiency Video Coding)
非特許文献2：「AHG9/AHG16: Common text for proposed generative face video SEI message」、JVET-AG0203-v1

発明の概要

発明が解決しようとする課題

[0005] 上記のような符号化方式に関して、符号化効率の改善、画質の改善、処理量の削減、回路規模の削減、又は、フィルタ、ブロック、サイズ、動きベクトル、参照ピクチャ又は参照ブロック等の要素又は動作の適切な選択等のため、新たな方式の提案が望まれている。

[0006] 本開示は、例えば、符号化効率の改善、画質の改善、処理量の削減、回路規模の削減、処理速度の改善、及び、要素又は動作の適切な選択等のうち1つ以上に貢献し得る構成又は方法を提供する。なお、本開示は、上記以外の利益に貢献し得る構成又は方法を含み得る。

課題を解決するための手段

[0007] 例えば、本開示の一態様に係る復号装置は、メモリと、前記メモリに接続された回路とを備え、前記回路は、動作において、(i) 顔を含む画像である基準画像と、(i i) カメラによって得られる撮影動画像の複数のフレームの各々に対応し、被写体の幾何学属性を示す幾何学情報と、(i i i) 背景画像に関する背景情報とを1つ以上のストリームから復号し、前記基準画像、前記幾何学情報及び前記背景情報から、生成モデルを用いて、前記顔を含む動画像であり、前記背景画像が合成された動画像である合成顔動画像を生成する、復号装置である。

[0008] 本開示における各実施の形態、またはその一部の構成もしくは方法のそれぞれは、例えば、符号化効率の改善、画質の改善、符号化／復号の処理量の削減、回路規模の削減、または、符号化／復号の処理速度の改善などのうちの、少なくともいずれか1つを可能にする。あるいは、本開示における各実施の形態、またはその一部の構成もしくは方法のそれぞれは、符号化および復号において、フィルタ、ブロック、サイズ、動きベクトル、参照ピクチャ、参照ブロックなどの構成要素／動作の適切な選択などを可能にする。なお、本開示は、上記以外の利益を提供し得る構成または方法の開示も含む。例えば、処理量の増加を抑えつつ、符号化効率を改善する構成または方法などである。

[0009] 本開示の一態様におけるさらなる利点および効果は、明細書および図面から明らかにされる。かかる利点および／または効果は、いくつかの実施の形態並びに明細書および図面に記載された特徴によってそれぞれ得られるが、1つまたはそれ以上の利点および／または効果を得るために必ずしも全てが提供される必要はない。

[0010] なお、これらの全般的または具体的な態様は、システム、集積回路、コンピュータプログラム又はコンピュータ読み取り可能なＣＤ－ＲＯＭなどの記録媒体で実現されてもよく、システム、方法、集積回路、コンピュータプログラム及び記録媒体の任意な組み合わせで実現されてもよい。

発明の効果

[0011] 本開示の一態様に係る構成又は方法は、例えば、符号化効率の改善、画質の改善、処理量の削減、回路規模の削減、処理速度の改善、及び、要素又は動作の適切な選択等のうち1つ以上に貢献し得る。なお、本開示の一態様に係る構成又は方法は、上記以外の利益に貢献してもよい。

図面の簡単な説明

[0012] [図1]図1は、参考例における符号化復号システムの構成を示すブロック図である。

[図2]図2は、参考例における符号化装置の構成を示すブロック図である。

[図3]図3は、参考例における復号装置の構成を示すブロック図である。

[図4]図4は、基準画像の例を示す概念図である。

[図5]図5は、幾何学属性の例を示す概念図である。

[図6]図6は、顔動画像の例を示す概念図である。

[図7]図7は、実施の形態における符号化復号システムの構成例を示すブロック図である。

[図8]図8は、ストリームにおけるデータの階層構造の一例を示す図である。

[図9]図9は、実施の形態における符号化装置の構成例を示すブロック図である。

[図10]図10は、実施の形態における符号化装置の動作例を示すフローチャ

ートである。

[図11]図 1 1 は、実施の形態における復号装置の構成例を示すブロック図である。

[図12]図 1 2 は、実施の形態における復号装置の動作例を示すフローチャートである。

[図13]図 1 3 は、各タイムインスタンスにおける復号処理の例を示す概念図である。

[図14]図 1 4 は、各タイムインスタンスにおける復号処理の別の例を示す概念図である。

[図15]図 1 5 は、実施の形態における復号装置の別の構成例を示すブロック図である。

[図16]図 1 6 は、実施の形態における復号装置のさらに別の構成例を示すブロック図である。

[図17]図 1 7 は、実施の形態における符号化装置の別の構成例を示すブロック図である。

[図18]図 1 8 は、実施の形態における復号装置のさらに別の構成例を示すブロック図である。

[図19]図 1 9 は、実施の形態における復号装置のさらに別の構成例を示すブロック図である。

[図20]図 2 0 は、実施の形態における符号化装置が動画像を符号化するための構成例を示すブロック図である。

[図21]図 2 1 は、実施の形態における復号装置が動画像を復号するための構成例を示すブロック図である。

[図22]図 2 2 は、ビットストリームの構成例を示す概念図である。

[図23]図 2 3 は、ビットストリームの別の構成例を示す概念図である。

[図24]図 2 4 は、ビットストリームのさらに別の構成例を示す概念図である。

。

[図25]図 2 5 は、ビットストリームのさらに別の構成例を示す概念図である。

。

[図26]図26は、ビットストリームのさらに別の構成例を示す概念図である。

。

[図27]図27は、VVCに準拠するビットストリームの構成例を示す概念図である。

[図28]図28は、VVCに準拠するビットストリームの別の構成例を示す概念図である。

[図29]図29は、VVCに準拠するビットストリームのさらに別の構成例を示す概念図である。

[図30]図30は、生成モデルとして用いられ得る様々なニューラルネットワークの例を示す図である。

[図31]図31は、実施の形態における符号化装置の実装例を示すブロック図である。

[図32]図32は、実施の形態における符号化装置の基本的な第1動作例を示すフローチャートである。

[図33]図33は、実施の形態における符号化装置の基本的な第2動作例を示すフローチャートである。

[図34]図34は、実施の形態におけるビットストリーム生成装置の実装例を示すブロック図である。

[図35]図35は、実施の形態におけるビットストリーム生成装置の基本的な第1動作例を示すフローチャートである。

[図36]図36は、実施の形態におけるビットストリーム生成装置の基本的な第2動作例を示すフローチャートである。

[図37]図37は、実施の形態における復号装置の実装例を示すブロック図である。

[図38]図38は、実施の形態における復号装置の基本的な第1動作例を示すフローチャートである。

[図39]図39は、実施の形態における復号装置の基本的な第2動作例を示す

フローチャートである。

[図40]図40は、コンテンツ配信サービスを実現するコンテンツ供給システムの全体構成図である。

[図41]図41は、webページの表示画面例を示す図である。

[図42]図42は、webページの表示画面例を示す図である。

[図43]図43は、スマートフォンの一例を示す図である。

[図44]図44は、スマートフォンの構成例を示すブロック図である。

発明を実施するための形態

[0013] [序論 (Introduction)]

顔再現は、人物のアイデンティティを維持しながら、人物の表情及びポーズを顔画像にマッピングする処理である。現在、顔再現は、ビデオ会議、娯楽産業及びソーシャルメディアを含む幅広い用途で用いられている。本開示は、顔再現に関するデータの符号化において用いられ得る。

[0014] 図1は、参考例における符号化復号システムの構成を示すブロック図である。例えば、符号化復号システムは、符号化装置700及び復号装置800を備える。まず、符号化装置700は、基準画像及びドライビング動画像を受け付け、ビットストリームを生成する。続いて、符号化装置700は、伝送チャネルを用いてビットストリームを復号装置800に伝送する。最後に、復号装置800が、受信したビットストリームから顔動画像を再構成する。

[0015] 例えば、基準画像は、顔を含む画像であり、顔画像又はアイデンティティ画像とも表現され得る。基準画像は、顔動画像を再構成するための静的で視覚的な特徴を表す。ドライビング動画像は、顔を含む動画像であって、カメラによって得られる撮影動画像である。ドライビング動画像は、基準画像に動きを与える役割を果たす。

[0016] 図2は、参考例における符号化装置700の構成を示すブロック図である。この例において、符号化装置700は、圧縮器731、導出器732及び圧縮器733を備える。

- [0017] 圧縮器 7 3 1 は、ビデオ圧縮技術を用いて少なくとも 1 つの基準画像を符号化する。基準画像は、ドライビング動画像のフレームでもよいし、人物の顔を含む予め取得された画像でもよいし、アバターでもよい。
- [0018] 導出器 7 3 2 は、ドライビング動画像の各フレームに対応する幾何学属性を示す幾何学情報を導出する。幾何学属性を示す幾何学情報は、単に幾何学属性とも表現される。ここで、幾何学属性は、例えば動的属性に対応し、顔のランドマークのような点群で表現されてもよいし、複数のポリゴンの組み合わせによって物体の形状を表現するためのポリゴンモデルで表現されてもよい。また、幾何学属性は、その他の幾何学的モデルで表現されてもよい。また、幾何学属性は、顔のパーツの位置で表現されてもよい。
- [0019] 圧縮器 7 3 3 は、エントロピー符号化等の方法を用いて幾何学属性をビットストリームに圧縮する。圧縮器 7 3 1 と圧縮器 7 3 3 とは、同じ構成要素であってもよいし、異なる構成要素であってもよい。
- [0020] 最後に、伝送チャネルを介してビットストリームが復号装置 8 0 0 に伝送される。例えば、圧縮幾何学属性は、ドライビング動画像のフレーム毎、つまり、時間インスタンス毎に符号化装置 7 0 0 から復号装置 8 0 0 へ伝送される。
- [0021] 図 3 は、参考例における復号装置 8 0 0 の構成を示すブロック図である。この例において、復号装置 8 0 0 は、圧縮解除器 8 3 1、導出器 8 3 2、圧縮解除器 8 3 3 及び生成器 8 3 4 を備える。
- [0022] 圧縮解除器 8 3 1 は、ビットストリームから少なくとも 1 つの基準画像を復号し再構成する。その後、圧縮解除器 8 3 1 は、基準画像を導出器 8 3 2 に供給する。導出器 8 3 2 は、基準画像から基準属性を導出する。ここで、基準属性は、静的で視覚的な属性であって、アイデンティティとも表現され得る。圧縮解除器 8 3 3 は、フレーム毎に幾何学属性を復号し再構成する。
- [0023] 生成器 8 3 4 は、基準属性及び幾何学属性から、ニューラルネットワークを用いて、顔動画像を生成する。生成器 8 3 4 は、基準属性に代えて、又は、基準属性に加えて、基準画像自体を用いて、顔動画像をレンダリングして

もよい。

[0024] 図4は、基準画像の例を示す概念図である。この例のように、基準画像は、顔を含む画像である。

[0025] 図5は、幾何学属性の例を示す概念図である。この例では、幾何学属性は、顔のランドマークである。例えば、撮影動画像のフレーム毎に幾何学属性が導出される。

[0026] 図6は、顔動画像の例を示す概念図である。この例のように、顔動画像は、顔を含む動画像である。顔動画像では、複数のフレームのそれぞれについて、当該フレームの幾何学属性が基準画像に反映される。これにより、基準画像に動きが与えられる。

[0027] 基準画像及び複数フレームに対応する複数の幾何学属性の情報量は、撮影動画像に含まれる複数のフレームの情報量に比べて少ない。したがって、基準画像及び複数フレームに対応する複数の幾何学属性を符号化することで、撮影動画像に含まれる複数のフレームを符号化するよりも、符号量が削減される。また、各幾何学属性によって基準画像に動きが与えられる。これにより、表示対象の顔に動きが与えられ、豊かな表現が可能になる。

[0028] しかしながら、基準画像に含まれる顔に動きが与えられることで、基準画像に含まれる背景領域が歪む可能性がある。例えば、図4から図6へ顔の位置がずれることにより、背景が伸縮したり欠落したりする可能性がある。これにより、画質が劣化する可能性がある。

[0029] そこで、例1の復号装置は、メモリと、前記メモリに接続された回路とを備え、前記回路は、動作において、(i) 顔を含む画像である基準画像と、(i i) カメラによって得られる撮影動画像の複数のフレームの各々に対応し、被写体の幾何学属性を示す幾何学情報と、(i i i) 背景画像に関する背景情報とを1つ以上のストリームから復号し、前記基準画像、前記幾何学情報及び前記背景情報から、生成モデルを用いて、前記顔を含む動画像であり、前記背景画像が合成された動画像である合成顔動画像を生成する、復号装置である。

[0030] これにより、合成顔動画像の生成に、基準画像、幾何学属性及び背景画像を適用することが可能になる場合がある。したがって、各フレームに対応する幾何学属性によって基準画像の顔に動きを与えつつ、背景画像によって基準画像の背景の歪みを抑制することが可能になる場合がある。よって、合成顔動画像の生成において、画質の劣化を抑制することが可能になる場合がある。

[0031] また、例2の復号装置は、例1の復号装置であって、前記回路は、前記生成モデルに前記基準画像、前記幾何学情報及び前記背景画像を入力して、前記合成顔動画像を前記生成モデルから取得する、復号装置であってもよい。

[0032] これにより、生成モデルから容易に合成顔動画像を取得することが可能になる場合がある。そして、生成モデルにおいて、合成顔動画像の生成に、基準画像、幾何学属性及び背景画像を適用することが可能になる場合がある。したがって、各フレームに対応する幾何学属性によって基準画像の顔に動きを与えつつ、背景画像によって基準画像の背景の歪みを抑制することが可能になる場合がある。

[0033] また、例3の復号装置は、例1の復号装置であって、前記回路は、前記生成モデルに前記基準画像及び前記幾何学情報を入力して、前記顔を含む動画像であり、前記背景画像が合成されていない動画像である中間顔動画像を前記生成モデルから取得し、前記中間顔動画像における背景領域に前記背景画像における対応領域を埋め込むことにより、前記合成顔動画像を生成する、復号装置であってもよい。

[0034] これにより、各フレームに対応する幾何学属性によって基準画像の顔に動きが与えられた中間顔動画像を生成モデルから取得することが可能になる場合がある。そして、中間顔動画像に背景画像を適用することが可能になる場合がある。したがって、各フレームに対応する幾何学属性によって基準画像の顔に動きを与えつつ、背景画像によって基準画像の背景の歪みを抑制することが可能になる場合がある。

[0035] また、例4の復号装置は、例3の復号装置であって、前記回路は、前記中

間顔動画像に対してセグメンテーション処理を行って、前記中間顔動画像における前景領域及び前記背景領域を示す中間顔動画像セグメンテーション情報を取得し、前記中間顔動画像セグメンテーション情報を用いて、前記中間顔動画像における前記背景領域を特定する、復号装置であってもよい。

[0036] これにより、中間顔動画像に対するセグメンテーション処理の結果として得られる中間顔動画像セグメンテーション情報に従って、中間顔動画像における背景領域を適切に特定することが可能になる場合がある。したがって、中間顔動画像における背景領域に背景画像における対応領域を適切に適用することが可能になる場合がある。

[0037] また、例5の復号装置は、例3の復号装置であって、前記回路は、前記撮影動画像における前景領域及び背景領域を示す撮影動画像セグメンテーション情報を前記1つ以上のストリームから復号し、前記撮影動画像セグメンテーション情報を用いて、前記中間顔動画像における前記背景領域を特定する、復号装置であってもよい。

[0038] これにより、1つ以上のストリームから得られる撮影動画像セグメンテーション情報に従って、中間顔動画像における背景領域を適切に特定することが可能になる場合がある。したがって、中間顔動画像における背景領域に背景画像における対応領域を適切に適用することが可能になる場合がある。

[0039] また、例6の復号装置は、例3の復号装置であって、前記基準画像における背景領域には、規定の背景カラーコードが埋め込まれている、復号装置であってもよい。

[0040] これにより、規定の背景カラーコードに従って、基準画像における背景領域を効率的に特定することが可能になる場合がある。また、基準画像の顔に動きが与えられても、基準画像の背景に生じる歪みを抑制することが可能になる場合がある。

[0041] また、例7の復号装置は、例6の復号装置であって、前記回路は、前記中間顔動画像において前記規定の背景カラーコードを有する領域を前記中間顔動画像における前記背景領域として特定する、復号装置であってもよい。

- [0042] これにより、規定の背景カラーコードに従って、中間顔動画像における背景領域を効率的に特定することが可能になる場合がある。具体的には、規定の背景カラーコードが埋め込まれた背景領域を有する基準画像の顔に動きを与えることにより得られる中間顔動画像における背景領域には規定の背景カラーコードが埋め込まれていると想定される。したがって、規定の背景カラーコードに従って、中間顔動画像における背景領域を効率的に特定することが可能になる場合がある。
- [0043] また、例 8 の復号装置は、例 6 又は 7 の復号装置であって、前記回路は、前記規定の背景カラーコードを示す背景カラーコード情報を前記 1 つ以上のストリームから復号する、復号装置であってもよい。
- [0044] これにより、1 つ以上のストリームから得られる規定の背景カラーコードに従って、基準画像における背景領域を効率的に特定することが可能になる場合がある。そして、基準画像に応じて規定の背景カラーコードを変化させることが可能になる場合がある。
- [0045] また、例 9 の復号装置は、例 8 の復号装置であって、前記背景カラーコード情報は、連続する複数の値を含む範囲を前記規定の背景カラーコードとして示し、前記規定の背景カラーコードは、前記背景カラーコード情報によって示される前記範囲内で規定される、復号装置であってもよい。
- [0046] これにより、規定の背景カラーコードを柔軟に規定することが可能になる場合がある。そして、背景領域に規定の背景カラーコードを柔軟に適用することが可能になる場合がある。
- [0047] また、例 10 の復号装置は、例 6 ～ 9 のいずれかの復号装置であって、前記規定の背景カラーコードは、前記基準画像における前景領域において発生頻度が閾値以下であるカラーコードで規定される、復号装置であってもよい。
- [0048] これにより、前景領域の部分が背景領域の部分と誤って特定されることを抑制することが可能になる場合がある。したがって、背景領域を適切に特定することが可能になる場合がある。

- [0049] また、例 1 1 の復号装置は、例 3 の復号装置であって、前記回路は、前記基準画像における前景領域及び背景領域を示す基準画像セグメンテーション情報を前記 1 つ以上のストリームから復号し、前記基準画像セグメンテーション情報を用いて、前記基準画像における前記背景領域に、規定の背景カラーコードを埋め込む、復号装置であってもよい。
- [0050] これにより、1 つ以上のストリームから得られる基準画像セグメンテーション情報に従って、基準画像における背景領域を効率的に特定することが可能になる場合がある。また、基準画像における背景領域に規定の背景カラーコードが埋め込まれるため、基準画像の顔に動きが与えられても、基準画像の背景に生じる歪みを抑制することが可能になる場合がある。
- [0051] また、例 1 2 の復号装置は、例 1 ～ 1 1 のいずれかの復号装置であって、前記背景画像は、前記基準画像及び前記撮影動画像によらずに用意された画像である、復号装置であってもよい。
- [0052] これにより、基準画像及び撮影動画像とは別途に用意された背景画像を合成顔動画像に適用することが可能になる場合がある。したがって、背景画像における前景領域の影響などを抑制することが可能になる場合がある。
- [0053] また、例 1 3 の復号装置は、例 1 ～ 1 1 のいずれかの復号装置であって、前記回路は、前記背景画像の識別子を前記背景情報として復号し、前記識別子を用いて、複数の背景画像候補の中から前記背景画像を選択する、復号装置であってもよい。
- [0054] これにより、複数の背景画像候補の中から背景画像を柔軟に選択することが可能になる場合がある。したがって、合成顔動画像の用途に応じて、合成顔動画像に適切な背景画像を適用することが可能になる場合がある。
- [0055] また、例 1 4 の復号装置は、例 1 ～ 1 1 のいずれかの復号装置であって、前記背景画像は、前記撮影動画像に含まれる画像、又は、前記撮影動画像に含まれる複数の画像が合成された画像である、復号装置であってもよい。
- [0056] これにより、撮影動画像から得られる背景画像を合成顔動画像に適用することが可能になる場合がある。したがって、撮影状況に対応する背景画像を

合成顔動画像に適用することが可能になる場合がある。

- [0057] また、例 1 5 の復号装置は、例 1 ～ 1 1 のいずれかの復号装置であって、前記回路は、前記基準画像を前記背景情報として復号し、前記背景画像に、前記基準画像が適用される、復号装置であってもよい。
- [0058] これにより、基準画像を背景画像として用いることが可能になる場合がある。そして、各フレームに対応する幾何学属性によって基準画像の顔に動きを与えつつ、元の基準画像が背景画像として用いられることで基準画像の背景の歪みを抑制することが可能になる場合がある。
- [0059] また、例 1 6 の復号装置は、例 1 4 又は 1 5 の復号装置であって、前記回路は、前記背景画像に前景領域が含まれる場合、前記背景画像における前記前景領域の周辺領域、又は、過去の合成顔動画像における背景領域を用いて、前記背景画像における背景領域の欠落を補間する、復号装置であってもよい。
- [0060] これにより、背景画像に前景領域が含まれていても、背景領域の欠落を適切に補間することが可能になる場合がある。したがって、合成顔動画像における背景領域の欠落を抑制することが可能になる場合がある。
- [0061] また、例 1 7 の復号装置は、例 1 6 の復号装置であって、前記回路は、前記背景画像に対してセグメンテーション処理を行って、前記背景画像における前記前景領域及び前記背景領域を示す背景画像セグメンテーション情報を取得し、前記背景画像セグメンテーション情報を用いて、前記背景画像における前記前景領域及び前記背景領域を特定する、復号装置であってもよい。
- [0062] これにより、背景画像に対するセグメンテーション処理の結果として得られる背景画像セグメンテーション情報に従って、背景画像における前景領域及び背景領域を適切に特定することが可能になる場合がある。したがって、背景画像における背景領域の欠落を適切に補間することが可能になる場合がある。
- [0063] また、例 1 8 の復号装置は、例 1 6 の復号装置であって、前記背景画像における前記前景領域には、規定の前景カラーコードが埋め込まれている、復

号装置であってもよい。

[0064] これにより、規定の前景カラーコードに従って、背景画像における前景領域を効率的に特定することが可能になる場合がある。また、合成顔動画像における背景領域に顔等の前景が映り込むことを抑制することが可能になる場合がある。

[0065] また、例 19 の復号装置は、例 18 の復号装置であって、前記回路は、前記規定の前景カラーコードを示す前景カラーコード情報を前記 1 つ以上のストリームから復号し、前記背景画像において前記規定の前景カラーコードを有する領域を前記背景画像における前記前景領域として特定する、復号装置であってもよい。

[0066] これにより、1 つ以上のストリームから得られる規定の前景カラーコードに従って、背景画像における前景領域を効率的に特定することが可能になる場合がある。そして、背景画像に応じて規定の前景カラーコードを変化させることが可能になる場合がある。

[0067] また、例 20 の復号装置は、例 19 の復号装置であって、前記前景カラーコード情報は、連続する複数の値を含む範囲を前記規定の前景カラーコードとして示し、前記規定の前景カラーコードは、前記前景カラーコード情報によって示される前記範囲内で規定される、復号装置であってもよい。

[0068] これにより、規定の前景カラーコードを柔軟に規定することが可能になる場合がある。そして、前景領域に規定の前景カラーコードを柔軟に適用することが可能になる場合がある。

[0069] また、例 21 の復号装置は、例 18 ～ 20 のいずれかの復号装置であって、前記規定の前景カラーコードは、前記背景画像における前記背景領域において発生頻度が閾値以下であるカラーコードで規定される、復号装置であってもよい。

[0070] これにより、背景領域の部分が前景領域の部分と誤って特定されることを抑制することが可能になる場合がある。したがって、前景領域を適切に特定することが可能になる場合がある。

- [0071] また、例 2 2 の復号装置は、例 1 ～ 2 1 のいずれかの復号装置であって、前記 1 つ以上のストリームにおいて、前記背景情報が復号されるストリームは、前記基準画像が復号されるストリーム、又は、前記幾何学情報が復号されるストリームと同じである、復号装置であってもよい。
- [0072] これにより、別途のストリームではなく、基準画像のストリーム又は幾何学情報のストリームと同じストリームから背景情報を復号することが可能になる場合がある。したがって、基準画像又は幾何学情報と一緒に背景情報を効率的に復号することが可能になる場合がある。
- [0073] また、例 2 3 の復号装置は、例 1 ～ 1 4 のいずれかの復号装置であって、前記 1 つ以上のストリームにおいて、前記背景情報が復号されるストリームは、前記基準画像が復号されるストリームとも前記幾何学情報が復号されるストリームとも異なる、復号装置であってもよい。
- [0074] これにより、同じストリームではなく、基準画像のストリーム及び幾何学情報のストリームとは別途のストリームから背景情報を復号することが可能になる場合がある。したがって、基準画像又は幾何学情報とは別に背景情報を任意のタイミングで復号することが可能になる場合がある。
- [0075] また、例 2 4 の復号装置は、例 1 ～ 2 2 のいずれかの復号装置であって、前記背景画像は、複数のピクチャを含むシーケンスの先頭のピクチャ、又は、GOP (Group Of Picture) の先頭のピクチャとして復号される、復号装置であってもよい。
- [0076] これにより、早期に背景画像を取得することが可能になる場合がある。したがって、早期に背景画像を合成顔動画像に適用することが可能になる場合がある。
- [0077] また、例 2 5 の復号装置は、例 1 ～ 1 4 のいずれかの復号装置であって、前記背景画像は、前記 1 つ以上のストリームにおけるアクセスユニットからピクチャとして復号される、復号装置であってもよい。
- [0078] これにより、背景画像をアクセスユニット内のピクチャとして処理することが可能になる場合がある。つまり、通常のピクチャと同様に背景画像を処

理することが可能になる場合がある。

[0079] また、例 2 6 の復号装置は、例 2 5 の復号装置であって、前記背景画像が復号される前記アクセスユニットは、前記基準画像が復号されるアクセスユニットと同じである、復号装置であってもよい。

[0080] これにより、別途のアクセスユニットではなく、基準画像のアクセスユニットと同じアクセスユニットから背景画像を復号することが可能になる場合がある。したがって、基準画像と一緒に背景画像を効率的に復号することが可能になる場合がある。

[0081] また、例 2 7 の復号装置は、例 2 5 の復号装置であって、前記背景画像が復号される前記アクセスユニットは、前記基準画像が復号されるアクセスユニットとは異なる、復号装置であってもよい。

[0082] これにより、同じアクセスユニットではなく、基準画像のアクセスユニットとは別途のアクセスユニットから背景画像を復号することが可能になる場合がある。したがって、基準画像とは別に背景画像を任意のタイミングで復号することが可能になる場合がある。

[0083] また、例 2 8 の復号装置は、例 2 5 ～ 2 7 のいずれかの復号装置であって、前記背景画像が含まれる前記アクセスユニットに対応付けられた S E I (Supplemental Enhancement Information) から、前記アクセスユニットに前記背景画像が存在することを示す信号が復号される、復号装置であってもよい。

[0084] これにより、アクセスユニットにおける S E I から得られる信号に従って、アクセスユニットにおける背景画像の存在を認識することが可能になる場合がある。したがって、背景画像の適切な伝達が可能になる場合がある。

[0085] また、例 2 9 の復号装置は、例 1 ～ 1 4 及び 2 5 ～ 2 8 のいずれかの復号装置であって、前記背景画像は、前記 1 つ以上のストリームにおけるアクセスユニットからピクチャとして復号され、前記背景画像が含まれる前記アクセスユニットに対応付けられた S E I (Supplemental Enhancement Information) から、前記アクセスユニット

に前記背景画像が存在することを示す信号が復号される、復号装置であってもよい。

[0086] これにより、背景画像をアクセスユニット内のピクチャとして処理することが可能になる場合がある。つまり、通常のピクチャと同様に背景画像を処理することが可能になる場合がある。また、これにより、アクセスユニットにおけるSEIから得られる信号に従って、アクセスユニットにおける背景画像の存在を認識することが可能になる場合がある。したがって、背景画像の適切な伝達が可能になる場合がある。

[0087] また、例30の復号装置は、例1～29のいずれかの復号装置であって、前記背景画像は、イントラピクチャとして復号される、復号装置であってもよい。

[0088] これにより、背景画像をイントラピクチャとして処理することが可能になる場合がある。つまり、他のピクチャに依存せずに、背景画像を処理することが可能になる場合がある。

[0089] また、例31の復号装置は、例1～30のいずれかの復号装置であって、前記背景画像は、前記合成顔動画画像の複数のフレームに共通して適用される、復号装置であってもよい。

[0090] これにより、合成顔動画画像の全体の符号量を抑制することが可能になる場合がある。また、背景画像を復号するための処理量を抑制することが可能になる場合がある。

[0091] また、例32の復号装置は、例8の復号装置であって、前記回路は、前記背景カラーコード情報を前記1つ以上のストリームにおけるSEI (Supplemental Enhancement Information) から復号する、復号装置であってもよい。

[0092] これにより、SEIから得られる規定の背景カラーコードに従って、基準画像における背景領域を効率的に特定することが可能になる場合がある。そして、基準画像に応じて規定の背景カラーコードを変化させることが可能になる場合がある。

- [0093] また、例 3 3 の復号装置は、例 1 9 の復号装置であって、前記回路は、前記前景カラーコード情報を前記 1 つ以上のストリームにおける S E I (S u p p l e m e n t a l E n h a n c e m e n t I n f o r m a t i o n) から復号する、復号装置であってもよい。
- [0094] これにより、S E I から得られる規定の前景カラーコードに従って、背景画像における前景領域を効率的に特定することが可能になる場合がある。そして、背景画像に応じて規定の前景カラーコードを変化させることが可能になる場合がある。
- [0095] また、例 3 4 の復号装置は、例 1 ～ 3 3 のいずれかの復号装置であって、前記回路は、規定の背景カラーコードを示す背景カラーコード情報と、規定の前景カラーコードを示す前景カラーコード情報とのうち、少なくとも一方を前記 1 つ以上のストリームにおける S E I (S u p p l e m e n t a l E n h a n c e m e n t I n f o r m a t i o n) から復号する、復号装置であってもよい。
- [0096] これにより、S E I から得られる規定の背景カラーコードに従って、背景領域を効率的に特定することが可能になる場合がある。
- [0097] また、例 3 5 の復号装置は、例 5 の復号装置であって、前記回路は、前記撮影動画像セグメンテーション情報を前記 1 つ以上のストリームにおける S E I (S u p p l e m e n t a l E n h a n c e m e n t I n f o r m a t i o n) から復号する、復号装置であってもよい。
- [0098] これにより、S E I から得られる撮影動画像セグメンテーション情報に従って、中間顔動画像における背景領域を適切に特定することが可能になる場合がある。したがって、中間顔動画像における背景領域に背景画像における対応領域を適切に適用することが可能になる場合がある。
- [0099] また、例 3 6 の復号装置は、例 1 1 の復号装置であって、前記回路は、前記基準画像セグメンテーション情報を前記 1 つ以上のストリームにおける S E I (S u p p l e m e n t a l E n h a n c e m e n t I n f o r m a t i o n) から復号する、復号装置であってもよい。

- [0100] これにより、SEIから得られる基準画像セグメンテーション情報に従って、基準画像における背景領域を効率的に特定することが可能になる場合がある。そして、基準画像における背景領域に規定の背景カラーコードを適切に埋め込むことが可能になる場合がある。
- [0101] また、例37の復号装置は、例1～36のいずれかの復号装置であって、前記回路は、前記撮影動画像における前景領域及び背景領域を示す撮影動画像セグメンテーション情報と、前記基準画像における前景領域及び背景領域を示す基準画像セグメンテーション情報とのうち、少なくとも一方を前記1つ以上のストリームにおけるSEI (Supplemental Enhancement Information) から復号する、復号装置であってもよい。
- [0102] これにより、SEIから得られるセグメンテーション情報に従って、背景領域を効率的に特定することが可能になる場合がある。
- [0103] また、例38の復号装置は、メモリと、前記メモリに接続された回路とを備え、前記回路は、動作において、(i) 顔を含む画像である基準画像と、(ii) カメラによって得られる撮影動画像の複数のフレームの各々に対応し、被写体の幾何学属性を示す幾何学情報とを1つ以上のストリームから復号し、前記基準画像及び前記幾何学情報を生成モデルに入力して、前記顔を含む動画像である中間顔動画像を前記生成モデルから取得し、前記中間顔動画像における背景領域に前記基準画像における対応領域を埋め込むことにより、合成顔動画像を生成する、復号装置であってもよい。
- [0104] これにより、各フレームに対応する幾何学属性によって基準画像の顔に動きが与えられた中間顔動画像を生成モデルから取得することが可能になる場合がある。そして、中間顔動画像の背景領域に元の基準画像の対応領域を適用することが可能になる場合がある。したがって、各フレームに対応する幾何学属性によって基準画像の顔に動きを与えつつ、元の基準画像によって基準画像の背景の歪みを抑制することが可能になる場合がある。よって、合成顔動画像の生成において、画質の劣化を抑制することが可能になる場合があ

る。

[0105] また、例39の符号化装置は、メモリと、前記メモリに接続された回路とを備え、前記回路は、動作において、顔を含む動画画であり、背景画像が合成された動画画である合成顔動画画を生成するための(i)基準画像と(i i)幾何学情報と(i i i)背景情報とであって、(i)前記顔を含む画像である基準画像と、(i i)カメラによって得られる撮影動画画の複数のフレームの各々に対応し、被写体の幾何学属性を示す幾何学情報と、(i i i)前記背景画像に関する背景情報とを1つ以上のストリームに符号化する、符号化装置である。

[0106] これにより、合成顔動画画を生成するための基準画像、幾何学属性及び背景画像を提供することが可能になる場合がある。したがって、合成顔動画画の生成において、各フレームに対応する幾何学属性によって基準画像の顔に動きを与えつつ、背景画像によって基準画像の背景の歪みを抑制することが可能になる場合がある。よって、画質の劣化の抑制に貢献することが可能になる場合がある。

[0107] また、例40の符号化装置は、例39の符号化装置であって、前記回路は、前記撮影動画画に対してセグメンテーション処理を行って、前記撮影動画画における前景領域及び背景領域を示す撮影動画画セグメンテーション情報を取得し、前記撮影動画画セグメンテーション情報を前記1つ以上のストリームに符号化する、符号化装置であってもよい。

[0108] これにより、1つ以上のストリームを介して、撮影動画画と同種の動画画における前景領域及び背景領域を特定するための撮影動画画セグメンテーション情報を提供することが可能になる場合がある。したがって、各フレームに対応する幾何学属性によって基準画像の顔に動きが与えられた中間顔動画画における前景領域及び背景領域の特定に貢献することが可能になる場合がある。

[0109] また、例41の符号化装置は、例39の符号化装置であって、前記回路は、前記基準画像に対してセグメンテーション処理を行って、前記基準画像に

における前景領域及び背景領域を示す基準画像セグメンテーション情報を取得し、前記基準画像セグメンテーション情報を用いて、前記基準画像における前記背景領域に規定の背景カラーコードを埋め込み、前記背景領域に前記規定の背景カラーコードが埋め込まれた前記基準画像を符号化する、符号化装置であってもよい。

[0110] これにより、基準画像に対するセグメンテーション処理の結果として得られる基準画像セグメンテーション情報に従って、基準画像における背景領域を適切に特定することが可能になる場合がある。また、基準画像における背景領域に規定の背景カラーコードが埋め込まれるため、基準画像の顔に動きが与えられても、基準画像の背景に生じる歪みを抑制することが可能になる場合がある。

[0111] また、例42の符号化装置は、例41の符号化装置であって、前記回路は、前記規定の背景カラーコードを示す背景カラーコード情報を前記1つ以上のストリームに符号化する、符号化装置であってもよい。

[0112] これにより、1つ以上のストリームを介して、基準画像における背景領域を効率的に特定するための規定の背景カラーコードを提供することが可能になる場合がある。そして、基準画像に応じて規定の背景カラーコードを変化させることが可能になる場合がある。

[0113] また、例43の符号化装置は、例42の符号化装置であって、前記背景カラーコード情報は、連続する複数の値を含む範囲を前記規定の背景カラーコードとして示し、前記規定の背景カラーコードは、前記背景カラーコード情報によって示される前記範囲内で規定される、符号化装置であってもよい。

[0114] これにより、規定の背景カラーコードを柔軟に規定することが可能になる場合がある。そして、背景領域に規定の背景カラーコードを柔軟に適用することが可能になる場合がある。

[0115] また、例44の符号化装置は、例41～43のいずれかの符号化装置であって、前記規定の背景カラーコードは、前記基準画像における前記前景領域において発生頻度が閾値以下であるカラーコードで規定される、符号化装置

であってもよい。

[0116] これにより、前景領域の部分が背景領域の部分と誤って特定されることを抑制することが可能になる場合がある。したがって、背景領域を適切に特定することが可能になる場合がある。

[0117] また、例45の符号化装置は、例39の符号化装置であって、前記回路は、前記基準画像に対してセグメンテーション処理を行って、前記基準画像における前景領域及び背景領域を示す基準画像セグメンテーション情報を取得し、前記基準画像セグメンテーション情報を前記1つ以上のストリームに符号化する、符号化装置であってもよい。

[0118] これにより、1つ以上のストリームを介して、基準画像における前景領域及び背景領域を特定するための基準画像セグメンテーション情報を提供することが可能になる場合がある。したがって、基準画像における背景領域の特定に貢献することが可能になる場合がある。

[0119] また、例46の符号化装置は、例39～45のいずれかの符号化装置であって、前記背景画像は、前記基準画像及び前記撮影動画像によらずに用意された画像である、符号化装置であってもよい。

[0120] これにより、基準画像及び撮影動画像とは別途に用意された背景画像を合成顔動画像に適用することが可能になる場合がある。したがって、背景画像における前景領域の影響などを抑制することが可能になる場合がある。

[0121] また、例47の符号化装置は、例39～45のいずれかの符号化装置であって、前記回路は、複数の背景画像候補の中から前記背景画像を選択し、前記背景画像の識別子を前記背景情報として符号化する、符号化装置であってもよい。

[0122] これにより、複数の背景画像候補の中から背景画像を柔軟に選択することが可能になる場合がある。したがって、合成顔動画像の用途に応じて、合成顔動画像に適切な背景画像を適用することが可能になる場合がある。

[0123] また、例48の符号化装置は、例39～45のいずれかの符号化装置であって、前記背景画像は、前記撮影動画像に含まれる画像、又は、前記撮影動

画像に含まれる複数の画像が合成された画像である、符号化装置であってもよい。

[0124] これにより、撮影動画像から得られる背景画像を合成顔動画像に適用することが可能になる場合がある。したがって、撮影状況に対応する背景画像を合成顔動画像に適用することが可能になる場合がある。

[0125] また、例49の符号化装置は、例39～45のいずれかの符号化装置であって、前記回路は、前記基準画像を前記背景情報として符号化し、前記背景画像に、前記基準画像が適用される、符号化装置であってもよい。

[0126] これにより、基準画像を背景画像として用いることが可能になる場合がある。そして、各フレームに対応する幾何学属性によって基準画像の顔に動きを与えつつ、元の基準画像が背景画像として用いられることで基準画像の背景の歪みを抑制することが可能になる場合がある。

[0127] また、例50の符号化装置は、例48の符号化装置であって、前記回路は、前記背景画像に前景領域が含まれる場合、前記背景画像における前記前景領域の周辺領域、又は、前記撮影動画像に含まれる他の画像における背景領域を用いて、前記背景画像における背景領域の欠落を補間する、符号化装置であってもよい。

[0128] これにより、背景画像に前景領域が含まれていても、背景領域の欠落を適切に補間することが可能になる場合がある。したがって、合成顔動画像における背景領域の欠落を抑制することが可能になる場合がある。

[0129] また、例51の符号化装置は、例50の符号化装置であって、前記回路は、前記背景画像に対してセグメンテーション処理を行って、前記背景画像における前記前景領域及び前記背景領域を示す背景画像セグメンテーション情報を取得し、前記背景画像セグメンテーション情報を用いて、前記背景画像における前記前景領域及び前記背景領域を特定する、符号化装置であってもよい。

[0130] これにより、背景画像に対するセグメンテーション処理の結果として得られる背景画像セグメンテーション情報に従って、背景画像における前景領域

及び背景領域を適切に特定することが可能になる場合がある。したがって、背景画像における背景領域の欠落を適切に補間することが可能になる場合がある。

[0131] また、例 5 2 の符号化装置は、例 4 8 の符号化装置であって、前記回路は、前記背景画像に対してセグメンテーション処理を行って、前記背景画像における前景領域及び背景領域を示す背景画像セグメンテーション情報を取得し、前記背景画像セグメンテーション情報を用いて、前記背景画像における前記前景領域に規定の前景カラーコードを埋め込む、符号化装置であってもよい。

[0132] これにより、規定の前景カラーコードに従って、背景画像における前景領域を効率的に特定することが可能になる場合がある。また、合成顔動画像における背景領域に顔等の前景が映り込むことを抑制することが可能になる場合がある。

[0133] また、例 5 3 の符号化装置は、例 5 2 の符号化装置であって、前記回路は、前記規定の前景カラーコードを示す前景カラーコード情報を前記 1 つ以上のストリームに符号化する、符号化装置であってもよい。

[0134] これにより、1 つ以上のストリームを介して、背景画像における前景領域を効率的に特定するための規定の前景カラーコードを提供することが可能になる場合がある。そして、背景画像に応じて規定の前景カラーコードを変化させることが可能になる場合がある。

[0135] また、例 5 4 の符号化装置は、例 5 3 の符号化装置であって、前記前景カラーコード情報は、連続する複数の値を含む範囲を前記規定の前景カラーコードとして示し、前記規定の前景カラーコードは、前記前景カラーコード情報によって示される前記範囲内で規定される、符号化装置であってもよい。

[0136] これにより、規定の前景カラーコードを柔軟に規定することが可能になる場合がある。そして、前景領域に規定の前景カラーコードを柔軟に適用することが可能になる場合がある。

[0137] また、例 5 5 の符号化装置は、例 5 2 ～ 5 4 のいずれかの符号化装置であ

って、前記規定の前景カラーコードは、前記背景画像における前記背景領域において発生頻度が閾値以下であるカラーコードで規定される、符号化装置であってもよい。

[0138] これにより、背景領域の部分が前景領域の部分と誤って特定されることを抑制することが可能になる場合がある。したがって、前景領域を適切に特定することが可能になる場合がある。

[0139] また、例56の符号化装置は、例39～55のいずれかの符号化装置であって、前記1つ以上のストリームにおいて、前記背景情報が符号化されるストリームは、前記基準画像が符号化されるストリーム、又は、前記幾何学情報が符号化されるストリームと同じである、符号化装置であってもよい。

[0140] これにより、別途のストリームではなく、基準画像のストリーム又は幾何学情報のストリームと同じストリームに背景情報を符号化することが可能になる場合がある。したがって、基準画像又は幾何学情報と一緒に背景情報を効率的に符号化することが可能になる場合がある。

[0141] また、例57の符号化装置は、例39～48のいずれかの符号化装置であって、前記1つ以上のストリームにおいて、前記背景情報が符号化されるストリームは、前記基準画像が符号化されるストリームとも前記幾何学情報が符号化されるストリームとも異なる、符号化装置であってもよい。

[0142] これにより、同じストリームではなく、基準画像のストリーム及び幾何学情報のストリームとは別途のストリームから背景情報に符号化することが可能になる場合がある。したがって、基準画像又は幾何学情報とは別に背景情報を任意のタイミングで符号化することが可能になる場合がある。

[0143] また、例58の符号化装置は、例39～56のいずれかの符号化装置であって、前記背景画像は、複数のピクチャを含むシーケンスの先頭のピクチャ、又は、GOP (Group Of Picture) の先頭のピクチャとして符号化される、符号化装置であってもよい。

[0144] これにより、早期に背景画像を提供することが可能になる場合がある。したがって、早期に背景画像を合成顔動画像に適用することが可能になる場合

がある。

- [0145] また、例 5 9 の符号化装置は、例 3 9 ～ 4 8 のいずれかの符号化装置であって、前記背景画像は、前記 1 つ以上のストリームにおけるアクセスユニットにピクチャとして符号化される、符号化装置であってもよい。
- [0146] これにより、背景画像をアクセスユニット内のピクチャとして処理することが可能になる場合がある。つまり、通常のピクチャと同様に背景画像を処理することが可能になる場合がある。
- [0147] また、例 6 0 の符号化装置は、例 5 9 の符号化装置であって、前記背景画像が符号化される前記アクセスユニットは、前記基準画像が符号化されるアクセスユニットと同じである、符号化装置であってもよい。
- [0148] これにより、別途のアクセスユニットではなく、基準画像のアクセスユニットと同じアクセスユニットに背景画像を符号化することが可能になる場合がある。したがって、基準画像と一緒に背景画像を効率的に符号化することが可能になる場合がある。
- [0149] また、例 6 1 の符号化装置は、例 5 9 の符号化装置であって、前記背景画像が符号化される前記アクセスユニットは、前記基準画像が符号化されるアクセスユニットとは異なる、符号化装置であってもよい。
- [0150] これにより、同じアクセスユニットではなく、基準画像のアクセスユニットとは別途のアクセスユニットに背景画像を符号化することが可能になる場合がある。したがって、基準画像とは別に背景画像を任意のタイミングで符号化することが可能になる場合がある。
- [0151] また、例 6 2 の符号化装置は、例 5 9 ～ 6 1 のいずれかの符号化装置であって、前記背景画像が符号化される前記アクセスユニットに対応付けられた S E I (S u p p l e m e n t a l E n h a n c e m e n t I n f o r m a t i o n) に、前記アクセスユニットに前記背景画像が存在することを示す信号が符号化される、符号化装置であってもよい。
- [0152] これにより、アクセスユニットにおける S E I の信号によって、アクセスユニットにおける背景画像の存在を通知することが可能になる場合がある。

したがって、背景画像の適切な伝達が可能になる場合がある。

[0153] また、例 6 3 の符号化装置は、例 3 9 ~ 4 8 及び 5 9 ~ 6 2 のいずれかの符号化装置であって、前記背景画像は、前記 1 つ以上のストリームにおけるアクセスユニットにピクチャとして符号化され、前記背景画像が符号化される前記アクセスユニットに対応付けられた S E I (S u p p l e m e n t a l E n h a n c e m e n t I n f o r m a t i o n) に、前記アクセスユニットに前記背景画像が存在することを示す信号が符号化される、符号化装置であってもよい。

[0154] これにより、背景画像をアクセスユニット内のピクチャとして処理することが可能になる場合がある。つまり、通常のピクチャと同様に背景画像を処理することが可能になる場合がある。また、これにより、アクセスユニットにおける S E I の信号によって、アクセスユニットにおける背景画像の存在を通知することが可能になる場合がある。したがって、背景画像の適切な伝達が可能になる場合がある。

[0155] また、例 6 4 の符号化装置は、例 3 9 ~ 6 3 のいずれかの符号化装置であって、前記背景画像は、イントラピクチャとして符号化される、符号化装置であってもよい。

[0156] これにより、背景画像をイントラピクチャとして処理することが可能になる場合がある。つまり、他のピクチャに依存せずに、背景画像を処理することが可能になる場合がある。

[0157] また、例 6 5 の符号化装置は、例 4 2 の符号化装置であって、前記回路は、前記背景カラーコード情報を前記 1 つ以上のストリームにおける S E I (S u p p l e m e n t a l E n h a n c e m e n t I n f o r m a t i o n) に符号化する、符号化装置であってもよい。

[0158] これにより、S E I を介して、基準画像における背景領域を効率的に特定するための規定の背景カラーコードを提供することが可能になる場合がある。そして、基準画像に応じて規定の背景カラーコードを変化させることが可能になる場合がある。

- [0159] また、例 6 6 の符号化装置は、例 5 3 の符号化装置であって、前記回路は、前記前景カラーコード情報を前記 1 つ以上のストリームにおける S E I (S u p p l e m e n t a l E n h a n c e m e n t I n f o r m a t i o n) に符号化する、符号化装置であってもよい。
- [0160] これにより、S E I を介して、背景画像における前景領域を効率的に特定するための規定の前景カラーコードを提供することが可能になる場合がある。そして、背景画像に応じて規定の前景カラーコードを変化させることが可能になる場合がある。
- [0161] また、例 6 7 の符号化装置は、例 3 9 ~ 6 6 のいずれかの符号化装置であって、前記回路は、規定の背景カラーコードを示す背景カラーコード情報と、規定の前景カラーコードを示す前景カラーコード情報とのうち少なくとも一方を前記 1 つ以上のストリームにおける S E I (S u p p l e m e n t a l E n h a n c e m e n t I n f o r m a t i o n) に符号化する、符号化装置であってもよい。
- [0162] これにより、S E I を介して、前景領域を効率的に特定するための規定の前景カラーコードを提供することが可能になる場合がある。
- [0163] また、例 6 8 の符号化装置は、例 4 0 の符号化装置であって、前記回路は、前記撮影動画像セグメンテーション情報を前記 1 つ以上のストリームにおける S E I (S u p p l e m e n t a l E n h a n c e m e n t I n f o r m a t i o n) に符号化する、符号化装置であってもよい。
- [0164] これにより、S E I を介して、撮影動画像と同種の動画像における前景領域及び背景領域を特定するための撮影動画像セグメンテーション情報を提供することが可能になる場合がある。したがって、各フレームに対応する幾何学属性によって基準画像の顔に動きが与えられた中間顔動画像における前景領域及び背景領域の特定に貢献することが可能になる場合がある。
- [0165] また、例 6 9 の符号化装置は、例 4 5 の符号化装置であって、前記回路は、前記基準画像セグメンテーション情報を前記 1 つ以上のストリームにおける S E I (S u p p l e m e n t a l E n h a n c e m e n t I n f o

r m a t i o n) に符号化する、符号化装置であってもよい。

[0166] これにより、S E I を介して、基準画像における前景領域及び背景領域を特定するための基準画像セグメンテーション情報を提供することが可能になる場合がある。したがって、基準画像における背景領域の特定に貢献することが可能になる場合がある。

[0167] また、例 7 0 の符号化装置は、例 3 9 ～ 6 9 のいずれかの符号化装置であって、前記回路は、前記撮影動画像における前景領域及び背景領域を示す撮影動画像セグメンテーション情報と、前記基準画像における前景領域及び背景領域を示す基準画像セグメンテーション情報とのうち少なくとも一方を前記 1 つ以上のストリームにおける S E I (S u p p l e m e n t a l E n h a n c e m e n t I n f o r m a t i o n) に符号化する、符号化装置であってもよい。

[0168] これにより、S E I を介して、前景領域及び背景領域を特定するためのセグメンテーション情報を提供することが可能になる場合がある。

[0169] また、例 7 1 の符号化装置は、メモリと、前記メモリに接続された回路とを備え、前記回路は、動作において、合成顔動画像を生成するための (i) 基準画像と (i i) 幾何学情報とであって、(i) 顔を含む画像である基準画像と、(i i) カメラによって得られる撮影動画像の複数のフレームの各々に対応し、被写体の幾何学属性を示す幾何学情報とを 1 つ以上のストリームに符号化し、前記合成顔動画像の生成において、前記基準画像及び前記幾何学情報は、前記基準画像及び前記幾何学情報を生成モデルに入力して、前記顔を含む動画像である中間顔動画像を前記生成モデルから取得することに用いられ、前記基準画像は、さらに、前記中間顔動画像における背景領域に前記基準画像における対応領域を埋め込むことにより、前記合成顔動画像を生成することに用いられる、符号化装置である。

[0170] これにより、合成顔動画像を生成するための基準画像及び幾何学属性を提供することが可能になる場合がある。そして、合成顔動画像の生成において、各フレームに対応する幾何学属性によって基準画像の顔に動きを与えつつ

、元の基準画像によって基準画像の背景の歪みを抑制することが可能になる場合がある。よって、画質の劣化の抑制に貢献することが可能になる場合がある。

[0171] また、例72のビットストリーム生成装置は、メモリと、前記メモリに接続された回路とを備え、前記回路は、動作において、顔を含む動画画であり、背景画像が合成された動画画である合成顔動画画を生成するための(i)基準画像と(i i)幾何学情報と(i i i)背景情報とであって、(i)前記顔を含む画像である基準画像と、(i i)カメラによって得られる撮影動画画の複数のフレームの各々に対応し、被写体の幾何学属性を示す幾何学情報と、(i i i)前記背景画像に関する背景情報とを含むビットストリームを生成する、ビットストリーム生成装置である。

[0172] これにより、合成顔動画画を生成するための基準画像、幾何学属性及び背景画像を提供することが可能になる場合がある。したがって、合成顔動画画の生成において、各フレームに対応する幾何学属性によって基準画像の顔に動きを与えつつ、背景画像によって基準画像の背景の歪みを抑制することが可能になる場合がある。よって、画質の劣化の抑制に貢献することが可能になる場合がある。

[0173] また、例73の復号方法は、(i)顔を含む画像である基準画像と、(i i)カメラによって得られる撮影動画画の複数のフレームの各々に対応し、被写体の幾何学属性を示す幾何学情報と、(i i i)背景画像に関する背景情報とを1つ以上のストリームから復号し、前記基準画像、前記幾何学情報及び前記背景情報から、生成モデルを用いて、前記顔を含む動画画であり、前記背景画像が合成された動画画である合成顔動画画を生成する、復号方法である。

[0174] これにより、合成顔動画画の生成に、基準画像、幾何学属性及び背景画像を適用することが可能になる場合がある。したがって、各フレームに対応する幾何学属性によって基準画像の顔に動きを与えつつ、背景画像によって基準画像の背景の歪みを抑制することが可能になる場合がある。よって、合成

顔動画の生成において、画質の劣化を抑制することが可能になる場合がある。

[0175] また、例 7 4 の符号化方法は、顔を含む動画画であり、背景画像が合成された動画画である合成顔動画画を生成するための (i) 基準画像と (i i) 幾何学情報と (i i i) 背景情報とであって、(i) 前記顔を含む画像である基準画像と、(i i) カメラによって得られる撮影動画画の複数のフレームの各々に対応し、被写体の幾何学属性を示す幾何学情報と、(i i i) 前記背景画像に関する背景情報とを 1 つ以上のストリームに符号化する、符号化方法である。

[0176] これにより、合成顔動画画を生成するための基準画像、幾何学属性及び背景画像を提供することが可能になる場合がある。したがって、合成顔動画画の生成において、各フレームに対応する幾何学属性によって基準画像の顔に動きを与えつつ、背景画像によって基準画像の背景の歪みを抑制することが可能になる場合がある。よって、画質の劣化の抑制に貢献することが可能になる場合がある。

[0177] さらに、これらの包括的又は具体的な態様は、システム、装置、方法、集積回路、コンピュータプログラム、又は、コンピュータ読み取り可能な CD-ROM などの非一時的な記録媒体で実現されてもよく、システム、装置、方法、集積回路、コンピュータプログラム、及び、記録媒体の任意な組み合わせで実現されてもよい。

[0178] [用語の定義]

各用語は一例として以下のような定義であってもよい。

[0179] (1) 画像

画素の集合によって構成されたデータの単位であり、ピクチャやピクチャより小さいブロックからなり、動画の他、静止画も含む。

[0180] (2) ピクチャ

画素の集合によって構成される画像の処理単位であり、フレームやフィールドと呼ばれる場合もある。

[0181] (3) ブロック

特定数の画素を含む集合の処理単位であり、以下の例に挙げる通り、名称は問わない。また、形状も問わず、例えば、 $M \times N$ 画素からなる長方形、 $M \times M$ 画素からなる正方形はもちろん、三角形、円形、その他の形状も含む。

[0182] (ブロックの例)

- ・ スライス／タイル／ブリック
- ・ C T U／スーパーブロック／基本分割単位
- ・ V P D U／ハードウェアの処理分割単位
- ・ C U／処理ブロック単位／予測ブロック単位 (P U)／直交変換ブロック単位 (T U)／ユニット
- ・ サブブロック

[0183] (4) 画素／サンプル

画像を構成する最小単位の点であって、整数位置の画素のみならず整数位置の画素に基づいて生成された小数位置の画素も含む。

[0184] (5) 画素値／サンプル値

画素が有する固有の値であって、輝度値、色差値、R G Bの階調はもちろん、d e p t h値、又は0、1の2値も含む。

[0185] (6) フラグ

1ビットの他、複数ビットの場合も含み、例えば、2ビット以上のパラメータやインデックスであってもよい。また、二進数を用いた2値のみならず、その他の進数を用いた多値であってもよい。

[0186] (7) 信号

情報を伝達するために記号化、符号化したものであって、離散化されたデジタル信号の他、連続値を取るアナログ信号も含む。

[0187] (8) ストリーム／ビットストリーム

デジタルデータのデータ列又はデジタルデータの流れをいう。ストリーム／ビットストリームは、1本のストリームの手、複数の階層に分けられ複数のストリームにより構成されてもよい。また、単数の伝送路でシリアル通信

により伝送される場合の他、複数の伝送路でパケット通信により伝送される場合も含む。

[0188] (9) 差／差分

スカラー量の場合、単純差 ($x - y$) の他、差の演算が含まれていれば足り、差の絶対値 ($|x - y|$)、二乗差 ($x^2 - y^2$)、差の平方根 ($\sqrt{x - y}$)、重み付け差 ($ax - by$: a 、 b は定数)、オフセット差 ($x - y + a$: a はオフセット) を含む。

[0189] (10) 和

スカラー量の場合、単純和 ($x + y$) の他、和の演算が含まれていれば足り、和の絶対値 ($|x + y|$)、二乗和 ($x^2 + y^2$)、和の平方根 ($\sqrt{x + y}$)、重み付け和 ($ax + by$: a 、 b は定数)、オフセット和 ($x + y + a$: a はオフセット) を含む。

[0190] (11) 基づいて (based on)

基づく対象となる要素以外を加味する場合も含む。また、直接結果を求める場合の他、中間的な結果を経由して結果を求める場合も含む。

[0191] (12) 用いて (used、using)

用いる対象となる要素以外を加味する場合も含む。また、直接結果を求める場合の他、中間的な結果を経由して結果を求める場合も含む。

[0192] (13) 禁止する (prohibit、forbid)

許されないと言い換えることができる。また、禁止していないこと又は許可されることは、必ずしも義務を意味するものではない。

[0193] (14) 制限する (limit、restriction／restrict／restricted)

許されないと言い換えることができる。また、禁止していないこと又は許可されることは、必ずしも義務を意味するものではない。さらに、量的又は質的に一部が禁止されていれば足り、全面的に禁止する場合も含まれる。

[0194] (15) 色差 (chroma)

サンプル配列または単一のサンプルが、原色に関連する2つの色差 (co

l o u r d i f f e r e n c e) 信号の1つを表すことを指定する、記号 C b および C r で表される形容詞である。c h r o m a という用語の代わりに、c h r o m i n a n c e という用語を使用することもできる。

[0195] (16) 輝度 (l u m a)

サンプル配列または単一のサンプルが原色に関連するモノクロ信号を表すことを指定する、記号または下付きの Y または L で表される形容詞である。

l u m a という用語の代わりに、l u m i n a n c e という用語を使用することもできる。

[0196] [記載に関する解説]

図面において、同一の参照番号は同一または類似の構成要素を示す。また、図面における構成要素のサイズおよび相対位置は、必ずしも一定の縮尺で描かれていない。

[0197] 以下、実施の形態について図面を参照しながら具体的に説明する。なお、以下で説明する実施の形態は、いずれも包括的または具体的な例を示すものである。以下の実施の形態で示される数値、形状、材料、構成要素、構成要素の配置位置及び接続形態、ステップ、ステップの関係及び順序などは、一例であり、請求の範囲を限定する主旨ではない。

[0198] 以下では、符号化装置および復号化装置の実施の形態を説明する。実施の形態は、本開示の各態様で説明する処理および／または構成を適用可能な符号化装置および復号化装置の例である。処理および／または構成は、実施の形態とは異なる符号化装置および復号化装置においても実施可能である。例えば、実施の形態に対して適用される処理および／または構成に関して、例えば以下のいずれかを実施してもよい。

[0199] (1) 本開示の各態様で説明する実施の形態の符号化装置または復号装置の複数の構成要素のうちいずれかは、本開示の各態様のいずれかで説明する他の構成要素に置き換えまたは組み合わせられてもよい。

[0200] (2) 実施の形態の符号化装置または復号装置において、当該符号化装置または復号装置の複数の構成要素のうち一部の構成要素によって行われる機

能または処理に、機能または処理の追加、置き換え、削除などの任意の変更がなされてもよい。例えば、いずれかの機能または処理は、本開示の各態様のいずれかで説明する他の機能または処理に、置き換えまたは組み合わせられてもよい。

[0201] (3) 実施の形態の符号化装置または復号装置が実施する方法において、当該方法に含まれる複数の処理のうちの一部の処理について、追加、置き換えおよび削除などの任意の変更がなされてもよい。例えば、方法におけるいずれかの処理は、本開示の各態様のいずれかで説明する他の処理に、置き換えまたは組み合わせられてもよい。

[0202] (4) 実施の形態の符号化装置または復号装置を構成する複数の構成要素のうちの一部の構成要素は、本開示の各態様のいずれかで説明する構成要素と組み合わせられてもよいし、本開示の各態様のいずれかで説明する機能の一部を備える構成要素と組み合わせられてもよいし、本開示の各態様で説明する構成要素が実施する処理の一部を実施する構成要素と組み合わせられてもよい。

[0203] (5) 実施の形態の符号化装置または復号装置の機能の一部を備える構成要素、または、実施の形態の符号化装置または復号装置の処理の一部を実施する構成要素は、本開示の各態様いずれかで説明する構成要素と、本開示の各態様でいずれかで説明する機能の一部を備える構成要素と、または、本開示の各態様のいずれかで説明する処理の一部を実施する構成要素と組み合わせまたは置き換えられてもよい。

[0204] (6) 実施の形態の符号化装置または復号装置が実施する方法において、当該方法に含まれる複数の処理のいずれかは、本開示の各態様のいずれかで説明する処理に、または、同様のいずれかの処理に、置き換えまたは組み合わせられてもよい。

[0205] (7) 実施の形態の符号化装置または復号装置が実施する方法に含まれる複数の処理のうちの一部の処理は、本開示の各態様のいずれかで説明する処理と組み合わせられてもよい。

[0206] (8) 本開示の各態様で説明する処理および／または構成の実施の仕方は、実施の形態の符号化装置または復号装置に限定されるものではない。例えば、処理および／または構成は、実施の形態において開示する動画像符号化または動画像復号とは異なる目的で利用される装置において実施されてもよい。

[0207] [符号化復号システムの構成]

図7は、本実施の形態における符号化復号システムの構成例を示すブロック図である。例えば、符号化復号システムは、符号化装置100及び復号装置200を備える。図7の例は、図1の例と類似するが、図7では、符号化装置100の具体的な構成及び処理、復号装置200の具体的な構成及び処理、及び、ビットストリームが、図1の例とは異なる。

[0208] 符号化装置100は、基準画像、ドライビング動画像及び背景画像を受け付け、ビットストリームを生成する。続いて、符号化装置100は、伝送チャネルを用いてビットストリームを復号装置200に伝送する。最後に、復号装置200が、ビットストリームから合成顔動画像を再構成する。

[0209] 例えば、基準画像は、顔を含む画像であり、顔画像又はアイデンティティ画像とも表現され得る。基準画像は、合成顔動画像を再構成するための静的で視覚的な特徴を表す。ドライビング動画像は、顔を含む動画像であって、カメラによって得られる撮影動画像である。ドライビング動画像は、基準画像に動きを与える役割を果たす。ビットストリームは、単にストリームとも表現される。また、1つのビットストリームが用いられることに限られず、複数のビットストリームが用いられてもよい。

[0210] 基準画像に含まれる人と、ドライビング動画像に含まれる人とは、同一人物であってもよいし、同一人物でなくてもよい。

[0211] なお、本実施の形態における符号化復号システムは、ビデオ会議、エンターテインメント業界におけるビデオの生成及び編集、ソーシャルメディア、及び、eコマース業界などに適用され得る。しかしながら、適用範囲は、これらに限定されない。

[0212] [データ構造]

図8は、ストリームにおけるデータの階層構造の一例を示す図である。ストリームは、例えばビデオシーケンスを含む。このビデオシーケンスは、例えば図8の(a)に示すように、VPS (Video Parameter Set) と、SPS (Sequence Parameter Set) と、PPS (Picture Parameter Set) と、SEI (Supplemental Enhancement Information) と、複数のピクチャとを含む。

[0213] VPSは、複数のレイヤから構成されている動画像において、複数のレイヤに共通する符号化パラメータと、動画像に含まれる複数のレイヤ、または個々のレイヤに関連する符号化パラメータとを含む。

[0214] SPSは、シーケンスに対して用いられるパラメータ、すなわち、シーケンスを復号するために復号装置200が参照する符号化パラメータを含む。例えば、その符号化パラメータは、ピクチャの幅または高さを示してもよい。なお、SPSは複数存在してもよい。

[0215] PPSは、ピクチャに対して用いられるパラメータ、すなわち、シーケンス内の各ピクチャを復号するために復号装置200が参照する符号化パラメータを含む。例えば、その符号化パラメータは、ピクチャの復号に用いられる量子化幅の基準値と、重み付き予測の適用を示すフラグとを含んでもよい。なお、PPSは複数存在してもよい。また、SPSとPPSとは、単にパラメータセットと呼ばれる場合がある。

[0216] ピクチャは、図8の(b)に示すように、ピクチャヘッダと、1つ以上のスライスを含んでもよい。ピクチャヘッダは、その1つ以上のスライスを復号するために復号装置200が参照する符号化パラメータを含む。

[0217] スライスは、図8の(c)に示すように、スライスヘッダと、1つ以上のブリックとを含む。スライスヘッダは、その1つ以上のブリックを復号するために復号装置200が参照する符号化パラメータを含む。

[0218] ブリックは、図8の(d)に示すように、1つ以上のCTU (Coding Unit)

g Tree Unit)を含む。

[0219] なお、ピクチャは、スライスを含まず、そのスライスの代わりに、タイルグループを含んでもよい。この場合、タイルグループは、1つ以上のタイルを含む。また、ブリックにスライスが含まれていてもよい。

[0220] CTUは、スーパーブロックまたは基本分割単位とも呼ばれる。このようなCTUは、図8の(e)に示すように、CTUヘッダと、1つ以上のCU(Coding Unit)とを含む。CTUヘッダは、1つ以上のCUを復号するために復号装置200が参照する符号化パラメータを含む。

[0221] CUは、複数の小さいCUに分割されてもよい。また、CUは、図8の(f)に示すように、CUヘッダと、予測情報と、残差係数情報とを含む。予測情報は、そのCUを予測するための情報であって、残差係数情報は、後述の予測残差を示す情報である。なお、CUは、基本的にPU(Prediction Unit)およびTU(Transform Unit)と同一であるが、例えば後述のSBTでは、そのCUよりも小さい複数のTUを含んでもよい。また、CUは、そのCUを構成するVPDU(Virtual Pipeline Decoding Unit)ごとに処理されてもよい。VPDUは、例えば、ハードウェアにおいてパイプライン処理を行う際に、1ステージで処理できる固定的な単位である。

[0222] なお、ストリームは、図8に示す各階層のうちの何れか一部の階層を有していなくてもよい。また、これらの階層の順番は、入れ替えられてもよく、何れかの階層は他の階層に置き換えられてもよい。また、符号化装置100または復号装置200などの装置によって現時点で行われる処理の対象とされているピクチャを、カレントピクチャという。その処理が符号化であれば、カレントピクチャは、符号化対象ピクチャと同義であり、その処理が復号であれば、カレントピクチャは、復号対象ピクチャと同義である。また、符号化装置100または復号装置200などの装置によって現時点で行われる処理の対象とされている例えばCUまたはCUなどのブロックを、カレントブロックという。その処理が符号化であれば、カレントブロックは、符号化

対象ブロックと同義であり、その処理が復号であれば、カレントブロックは、復号対象ブロックと同義である。

[0223] ここで、符号化及び復号に用いられるパラメータが記述される領域は、ヘッダ領域と表現され得る。例えば、ヘッダ領域は、SEIを含む領域である。ヘッダ領域は、さらに、VPS、SPS、PPS、SEI、ピクチャヘッダ、スライスヘッダ、CTUヘッダ、及び、CUヘッダを含み得る。

[0224] また、例えば、ピクチャは、Iピクチャ、Pピクチャ及びBピクチャを含む複数の種類のいずれかに分類され得る。Iピクチャは、イントラ予測ピクチャであり、イントラピクチャとも呼ばれ、他のピクチャを参照せずに符号化及び復号が行われるピクチャである。Pピクチャは、単予測ピクチャであり、他の1つのピクチャを参照して符号化及び復号が行われ得るピクチャである。Bピクチャは、双予測ピクチャであり、他の2つのピクチャを参照して符号化及び復号が行われ得るピクチャである。

[0225] また、動画画は、複数のGOP (Group of Pictures) で構成され得る。GOPは、ピクチャ集合を意味する。GOPは、1つ以上のIピクチャを含む。GOPは、1つ以上のPピクチャを含んでいてもよいし、1つ以上のBピクチャを含んでいてもよい。GOPは、映像の編集及びランダムアクセス等が可能な単位であってもよい。GOPは、一定の数のピクチャを有していてもよいし、Iピクチャ、Pピクチャ及びBピクチャに関する一定の配置順をGOP構造として有していてもよい。

[0226] [符号化の構成及び処理]

図9は、本実施の形態における符号化装置100の構成例を示すブロック図である。符号化装置100は、基準画像、ドライビング動画画及び背景画像からビットストリームを生成する。この例において、符号化装置100は、圧縮器131、導出器132、圧縮器133及び圧縮器134を備える。各構成要素は、例えば、情報処理を行う電気回路である。圧縮器131、圧縮器133及び圧縮器134の2つ以上が統合されてもよい。

[0227] 図10は、本実施の形態における符号化装置100の動作例を示すフロー

チャートである。例えば、図9に示された符号化装置100の複数の構成要素は、図10のフローチャートに従って動作を行う。

[0228] この例において、まず、圧縮器131は、少なくとも1つの基準画像をビットストリームに符号化することにより圧縮する(S101)。基準画像は、VVCのようなビデオコーデック方式に従って符号化されてもよい。基準画像は、ドライビング動画像のフレームでもよいし、人物の顔を含む予め用意された画像でもよいし、アバターでもよい。

[0229] また、導出器132は、ドライビング動画像の各フレームに対応する幾何学属性を示す幾何学情報を導出する(S102)。幾何学情報は、単に幾何学属性とも表現される。具体的には、導出器132は、ドライビング動画像の各フレームをニューラルネットワーク等の認識モデルに入力し、認識モデルから各フレームに対応する幾何学属性を取得する。幾何学属性は、ドライビング動画像の各フレームの時間インスタンスに対応する。

[0230] ここで、幾何学属性は、例えば動的属性に対応し、顔のランドマークのような点群で表現されてもよいし、複数のポリゴンの組み合わせによって物体の形状を表現するためのポリゴンモデルで表現されてもよい。また、幾何学属性は、その他の幾何学的モデルで表現されてもよい。また、幾何学属性は、顔のパーツの位置で表現されてもよい。また、幾何学属性は、顔属性とも表現され得る。また、幾何学属性は、幾何学属性セットとして扱われてもよい。

[0231] 例えば、幾何学属性として用いられる顔ランドマークは、顔の輪郭、目、眉、鼻、口、唇及び顎を含む顔の主要領域上の点の位置を示す。このような幾何学属性は他の人又は他の装置によって解釈可能であるため、属性の修正が可能であり、属性の処理の改善が可能である。

[0232] 圧縮器133は、エントロピー符号化等の方法を用いて幾何学属性をビットストリームに符号化することにより圧縮する(S103)。

[0233] 圧縮器134は、少なくとも1つの背景画像をビットストリームに符号化することにより圧縮する(S104)。背景画像は、VVCのようなビデオ

コーデック方式に従って符号化されてもよい。

[0234] 背景画像は、合成顔動画像における背景領域に用いられる。つまり、背景画像は、顔を含む顔動画像にオーバーレイされる背景を示す。背景画像は、ドライビング動画像に含まれるフレームに対応する画像でもよいし、ドライビング動画像に含まれる複数のフレームに対応する複数の画像から生成されてもよい。背景画像は、基準画像及びドライビング動画像とは別に、予め用意された背景画像であってもよい。

[0235] また、背景画像は、複数の背景画像候補から選択可能であってもよい。背景画像の代わりに、複数の背景画像候補から背景画像を選択するための選択パラメータが符号化されてもよい。選択パラメータは、複数の背景画像候補のいずれかに対応する背景画像の識別子であってもよい。

[0236] さらに、背景画像には、ベタ塗りスタイル、テクスチャスタイル、グラデーションスタイル、パターンスタイル、ぼかしスタイル、イラスト、高コントラストスタイル、現実世界のシーン、又は、合成シーン等が用いられてもよい。また、背景画像には、これらの任意の組み合わせが用いられてもよいし、これらとは別の画像が用いられてもよい。

[0237] 符号化装置100は、復号装置200が行う後述の動作と同様に、基準画像、幾何学属性及び背景画像に基づいて、合成顔動画像を生成してもよい（S105）。符号化装置100は、合成顔動画像を生成するため、復号装置200と同様の複数の構成要素を備えていてもよい。これにより、復号装置200で生成される合成顔動画像を符号化装置100で確認することが可能になる。なお、この処理は、省略されてもよい。

[0238] 符号化装置100は、基準画像、幾何学属性及び背景画像をビットストリームに符号化した後、伝送チャネルを介してビットストリームを復号装置200に送信する。例えば、ドライビング動画像のフレーム毎、つまり、時間インスタンス毎に、符号化装置100から復号装置200へ圧縮幾何学属性がビットストリームとして伝送される。圧縮幾何学属性は、SEI (Supplemental Enhancement Information)

として伝送されてもよい。

[0239] 基準画像は、合成顔動画像の生成に対して、1回のみ伝送されてもよい。そして、復号装置200において、合成顔動画像の各フレームの生成時に同じ基準画像が用いられてもよい。同様に、背景画像は、合成顔動画像の生成に対して、1回のみ伝送されてもよい。そして、復号装置200において、合成顔動画像の各フレームの生成時に同じ背景画像が用いられてもよい。

[0240] あるいは、背景画像は、幾何学属性と同様に、ドライビング動画像のフレーム毎、つまり、時間インスタンス毎に伝送されてもよい。あるいは、背景画像は、複数のキーフレームのそれぞれとして伝送されてもよく、背景画像のリフレッシュレートが、キーフレーム間隔に対応していてもよい。

[0241] あるいは、符号化装置100は、ドライビング動画像における顔の位置を追跡し、顔の平行移動又は回転などの動きが閾値を超えた場合に、背景画像を送信してもよい。このように、背景画像のリフレッシュレートは、ドライビング動画像における顔の動きに依存してもよい。

[0242] 基準画像、幾何学属性及び背景画像は、同じ1つのビットストリームに符号化されて伝送されてもよいし、それぞれ異なる複数のビットストリームに符号化されて伝送されてもよい。基準画像、幾何学属性及び背景画像のうち、任意の2つが同じ1つのビットストリームに符号化されて伝送され、残りの1つが別のビットストリームに符号化されて伝送されてもよい。

[0243] [復号装置の構成及び処理]

図11は、本実施の形態における復号装置200の構成例を示すブロック図である。復号装置200は、ビットストリームから合成顔動画像を生成する。この例において、復号装置200は、圧縮解除器231、導出器232、圧縮解除器233、生成器234、圧縮解除器235及び合成器236を備える。各構成要素は、例えば、情報処理を行う電気回路である。圧縮解除器231、圧縮解除器233及び圧縮解除器235の2つ以上が統合されてもよい。

[0244] 図12は、本実施の形態における復号装置200の動作例を示すフローチ

ャートである。例えば、図 1 1 に示された復号装置 2 0 0 の複数の構成要素は、図 1 2 のフローチャートに従って動作を行う。なお、以下では、符号化の説明と同じ説明が省略される場合がある。

- [0245] 圧縮解除器 2 3 1 は、ビットストリームから少なくとも 1 つの基準画像を復号することにより圧縮解除する (S 2 0 1)。基準画像は、VVC のようなビデオコーデック方式に従って復号されてもよい。その後、圧縮解除器 2 3 1 は、基準画像を導出器 2 3 2 に供給する。
- [0246] 導出器 2 3 2 は、基準画像から基準属性を示す基準情報を導出する (S 2 0 2)。ここで、基準属性を示す基準情報は、単に基準属性とも表現される。基準属性は、静的で視覚的な属性であって、アイデンティティとも表現され得る。基準属性は、髪、眼鏡、顔の毛、眉毛、目、口、鼻、肌、顔の輪郭、服装及びアクセサリの少なくとも 1 つに関する情報を含み得る。
- [0247] 圧縮解除器 2 3 3 は、エントロピー復号等の方法を用いてビットストリームからフレーム毎に幾何学属性を復号することにより圧縮解除する (S 2 0 3)。
- [0248] 生成器 2 3 4 は、基準属性及び幾何学属性から、ニューラルネットワーク等の生成モデルを用いて、中間顔動画像を生成する (S 2 0 4)。
- [0249] 生成モデルは、GAN (Generative Adversarial Network)、VAE (Variational Autoencoder)、自己回帰モデル、又は、拡散モデル等でもよい。例えば、生成モデルは、提供データセットに基づいて新しいデータを生成する機械学習フレームワークであって、生成モデルにおいて、データセットの基本的な分布の分析及び学習が行われていてもよい。
- [0250] 例えば、生成器 2 3 4 は、フレーム毎に、基準属性及び幾何学属性を生成モデルに入力して、生成モデルから中間顔動画像及びセグメンテーションマスクを取得する。より具体的には、生成器 2 3 4 は、フレーム毎に、基準属性及び幾何学属性を生成モデルに入力して、生成モデルから中間顔動画像のフレーム及び当該フレームのセグメンテーションマスクを取得する。このセ

グメンテーションマスクは、中間顔動画像（具体的には中間顔動画像のフレーム）における前景領域及び背景領域を示す。

- [0251] セグメンテーションマスクは、前景領域の全ての画素値が1であり、背景領域の全ての画素値が0である2次元マップで表現されてもよいし、前景領域の全ての画素値が0であり、背景領域の全ての画素値が1である2次元マップで表現されてもよい。例えば、前景領域は、顔等を含む領域であって、動きを含む領域であり、背景領域は、顔等を含まない領域であって、動きを含まない領域である。セグメンテーションマスクは、セグメンテーション情報とも呼ばれる。
- [0252] 生成器234は、基準属性に代えて、又は、基準属性に加えて、基準画像自体を用いて、中間顔動画像をレンダリングしてもよい。また、導出器232は、生成器234に含まれていてもよいし、存在しなくてもよい。そして、導出器232において基準属性を導出するための認識モデルは、生成器234において中間顔動画像等を生成するための生成モデルに含まれていてもよい。導出器232及び基準属性に関して、他の変形例も同様である。
- [0253] つまり、生成器234は、基準画像及び幾何学属性から生成モデルを用いてセグメンテーションマスク及び中間顔動画像を生成してもよい。その際、生成器234は、基準画像及び幾何学属性を生成モデルに入力して、生成モデルからセグメンテーションマスク及び中間顔動画像を取得してもよい。
- [0254] 圧縮解除器235は、ビットストリームから少なくとも1つの背景画像を復号することにより圧縮解除する（S205）。背景画像は、VVCのようなビデオコーデック方式に従って復号されてもよい。背景画像の代わりに、複数の背景画像候補から背景画像を選択するための選択パラメータが復号されてもよい。選択パラメータは、複数の背景画像候補のいずれかに対応する背景画像の識別子であってもよい。
- [0255] 合成器236は、中間顔動画像、セグメンテーションマスク及び背景画像を用いて、中間顔動画像における背景領域に背景画像における対応領域を埋め込むことにより、合成顔動画像を生成する（S206）。

- [0256] 図13は、各タイムインスタンスにおける復号処理の例を示す概念図である。この例において、復号装置200は、最初のタイムインスタンス($t = 0$)で、圧縮基準画像、圧縮幾何学属性及び圧縮背景画像を受信する。そして、復号装置200は、圧縮基準画像及び圧縮背景画像を復号装置200のメモリ252に保存する。
- [0257] また、復号装置200は、圧縮基準画像、最初のタイムインスタンス($t = 0$)の圧縮幾何学属性、及び、圧縮背景画像に対して復号処理を行う。そして、復号装置200は、基準画像、幾何学属性及び背景画像から生成モデルを用いて合成顔動画画像における最初のタイムインスタンス($t = 0$)の画像を生成する。
- [0258] また、復号装置200は、後続のタイムインスタンス($t = T$)で、圧縮幾何学属性を受信し、復号装置200のメモリ252から圧縮基準画像及び圧縮背景画像を検索して取得する。そして、復号装置200は、圧縮基準画像、当該タイムインスタンス($t = T$)の圧縮幾何学属性、及び、圧縮背景画像に対して復号処理を行う。そして、復号装置200は、基準画像、幾何学属性及び背景画像から生成モデルを用いて合成顔動画画像における当該タイムインスタンス($t = T$)の画像を生成する。
- [0259] 復号装置200は、圧縮基準画像及び圧縮背景画像に対して復号処理を行うことで得られる基準画像及び背景画像を復号装置200のメモリ252に保存してもよい。そして、復号装置200は、後続のタイムインスタンス($t = T$)で、復号処理が適用された基準画像及び背景画像をメモリ252から取得して、メモリ252から取得された基準画像及び背景画像を合成顔動画画像における画像の生成に適用してもよい。
- [0260] この例のように、圧縮幾何学属性は、各タイムインスタンスにおいて受信される。1つのタイムインスタンスの幾何学属性は、ドライビング動画画像における1つのフレームから導出された幾何学属性である。一方、圧縮基準画像及び圧縮背景画像は、最初のタイムインスタンスにおいてのみ受信されてもよい。これにより、符号量が削減され得る。

- [0261] 図14は、各タイムインスタンスにおける復号処理の別の例を示す概念図である。この例において、復号装置200は、最初のタイムインスタンス($t=0$)で、圧縮基準画像、圧縮幾何学属性及び圧縮背景画像を受信する。そして、復号装置200は、圧縮基準画像を復号装置200のメモリ252に保存する。
- [0262] また、復号装置200は、圧縮基準画像、最初のタイムインスタンス($t=0$)の圧縮幾何学属性、及び、圧縮背景画像に対して復号処理を行う。そして、復号装置200は、基準画像、幾何学属性及び背景画像から生成モデルを用いて合成顔動画像における最初のタイムインスタンス($t=0$)の画像を生成する。
- [0263] また、復号装置200は、後続のタイムインスタンス($t=T$)で、圧縮幾何学属性及び圧縮背景画像を受信し、復号装置200のメモリ252から圧縮基準画像を検索して取得する。そして、復号装置200は、圧縮基準画像、当該タイムインスタンス($t=T$)の圧縮幾何学属性、及び、当該タイムインスタンス($t=T$)の圧縮背景画像に対して復号処理を行う。そして、復号装置200は、基準画像、幾何学属性及び背景画像から生成モデルを用いて合成顔動画像における当該タイムインスタンス($t=T$)の画像を生成する。
- [0264] 復号装置200は、圧縮基準画像に対して復号処理を行うことで得られる基準画像を復号装置200のメモリ252に保存してもよい。そして、復号装置200は、後続のタイムインスタンス($t=T$)で、復号処理が適用された基準画像をメモリ252から取得して、メモリ252から取得された基準画像を合成顔動画像における画像の生成に適用してもよい。
- [0265] この例のように、圧縮幾何学属性及び圧縮背景画像は、各タイムインスタンスにおいて受信される。1つのタイムインスタンスの幾何学属性は、ドライビング動画像における1つのフレームから導出された幾何学属性である。
- [0266] また、例えば、1つのタイムインスタンスの背景画像は、ドライビング動画像における1つのフレームから導出された背景画像であってもよい。背景

画像の符号化において、解像度が削減されたり、大きい量子化幅で量子化が行われたり、前景領域が前景カラーコードで塗りつぶされたりすることで、背景画像の符号量が削減されてもよい。

[0267] また、例えば、圧縮基準画像は、最初のタイムインスタンスにおいてのみ受信されてもよい。これにより、符号量が削減され得る。

[0268] [変形例]

図15は、本実施の形態における復号装置200の別の構成例を示すブロック図である。上述した例、つまり図11の例において、生成器234は、基準属性及び幾何学属性を生成モデルに入力して、生成モデルから中間顔動画像及びセグメンテーションマスクを取得する。

[0269] これに対して、この例、つまり図15の例において、生成器234は、基準属性及び幾何学属性を生成モデルに入力して、生成モデルから中間顔動画像を取得する。そして、生成器234は、中間顔動画像に対してセグメンテーション処理を行って、セグメンテーションマスクを取得する。

[0270] 具体的には、生成器234は、フレーム毎に、基準属性及び幾何学属性を生成モデルに入力して、生成モデルから中間顔動画像のフレームを取得する。そして、生成器234は、中間顔動画像の各フレームに対してセグメンテーション処理を行って、当該フレームのセグメンテーションマスクを取得する。

[0271] これにより、処理を細分化することが可能になり、処理を円滑に行うことが可能になる場合がある。生成器234の代わりに、セグメンテーション処理器（図示せず）がセグメンテーション処理を行ってもよい。

[0272] セグメンテーション処理は、ニューラルネットワーク等の機械学習モデルを用いて行われてもよい。本開示の他のセグメンテーション処理も同様である。

[0273] 復号装置200において生成される中間顔動画像及び合成顔動画像における前景領域及び背景領域は、ドライビング動画像における前景領域及び背景領域に対応する。そこで、符号化装置100が、ドライビング動画像に対し

てセグメンテーション処理を行い、ドライビング動画像のセグメンテーションマスクを符号化してもよい。そして、復号装置200が、セグメンテーションマスクを復号し、セグメンテーションマスクを用いて、合成顔動画像を生成してもよい。

[0274] 具体的には、符号化装置100において、導出器132が、フレーム毎に、ドライビング動画像に対してセグメンテーション処理を行って、ドライビング動画像における前景領域及び背景領域を示すセグメンテーションマスクを生成してもよい。また、圧縮器133が、セグメンテーションマスクをビットストリームに符号化することにより圧縮してもよい。

[0275] そして、復号装置200において、圧縮解除器233が、セグメンテーションマスクをビットストリームから復号することにより圧縮解除してもよい。さらに、合成器236が、セグメンテーションマスク等を用いて、合成顔動画像を生成してもよい。これにより、復号装置200における処理量が削減される可能性がある。

[0276] また、符号化装置100において、導出器132とは異なるセグメンテーション処理器（図示せず）が、セグメンテーション処理を行ってもよい。また、圧縮器133とは異なる圧縮器（図示せず）が、セグメンテーションマスクをビットストリームに符号化してもよい。また、復号装置200において、圧縮解除器233とは異なる圧縮解除器（図示せず）がセグメンテーションマスクをビットストリームから復号してもよい。

[0277] また、セグメンテーションマスクは、符号化装置100から復号装置200へフレーム毎にSEIにおいて伝送されてもよい。

[0278] 図16は、本実施の形態における復号装置200の別の構成例を示すブロック図である。この例では、生成器234は、基準属性及び幾何学属性を生成モデルに入力して、背景領域に規定の背景カラーコードが埋め込まれた中間顔動画像を生成モデルから取得する。つまり、生成モデルから取得される中間顔動画像において、背景領域が規定の背景色で塗りつぶされている。さらに言い換えれば、中間顔動画像における背景領域の全ての画素が、同じ規

定の背景カラーコードを画素値として有する。

[0279] したがって、セグメンテーションマスクを用いることなく、中間顔動画像における背景領域を特定することが可能になる。

[0280] 合成器 236 は、中間顔動画像における背景領域に、背景画像における対応領域を埋め込む。具体的には、中間顔動画像において、規定の背景カラーコードを有する画素は、背景画像において対応する画素に置き換えられ、規定の背景カラーコードを有しない画素は、そのまま維持される。これにより、合成器 236 は、背景領域に規定の背景カラーコードが埋め込まれた中間顔動画像、及び、背景画像から、合成顔動画像を生成する。

[0281] 図 17 は、本実施の形態における符号化装置 100 の別の構成例を示すブロック図である。この例では、背景領域に規定の背景カラーコードが埋め込まれた基準画像が用いられる。つまり、背景領域が規定の背景色で塗りつぶされた基準画像が用いられる。さらに言い換えれば、基準画像における背景領域の全ての画素が、同じ規定の背景カラーコードを画素値として有する。例えば、元々、基準画像における背景領域に規定の背景カラーコードが埋め込まれていてもよい。

[0282] あるいは、符号化装置 100 において、圧縮器 131 が、基準画像における背景領域に規定の背景カラーコードを埋め込んでもよい。具体的には、圧縮器 131 が、基準画像に対してセグメンテーション処理を行って、基準画像における前景領域及び背景領域を示すセグメンテーションマスクを取得してもよい。そして、圧縮器 131 が、セグメンテーションマスクに従って、基準画像における背景領域を特定し、基準画像における背景領域に規定の背景カラーコードを埋め込んでもよい。

[0283] 圧縮器 131 の代わりに、前処理器（図示せず）が、基準画像に対してセグメンテーション処理を行って、基準画像における背景領域に規定の背景カラーコードを埋め込んでもよい。

[0284] 圧縮器 131 は、背景領域に規定の背景カラーコードが埋め込まれた基準画像をビットストリームに符号化する。圧縮器 131 は、規定の背景カラー

コードを示す背景カラーコード情報をビットストリームに符号化してもよい。例えば、背景カラーコード情報はSEIにおいて伝送され得る。

[0285] 図18は、本実施の形態における復号装置200のさらに別の構成例を示すブロック図である。図18の復号装置200は、図17の符号化装置100で生成されたビットストリームから合成顔動画像を生成する。すなわち、この例では、背景領域に規定の背景カラーコードが埋め込まれた基準画像が用いられる。つまり、背景領域が規定の背景色で塗りつぶされた基準画像が用いられる。

[0286] 復号装置200において、圧縮解除器231は、背景領域に規定の背景カラーコードが埋め込まれた基準画像をビットストリームから復号する。導出器232は、背景領域に規定の背景カラーコードが埋め込まれた基準画像から基準属性を導出する。例えば、基準画像における背景領域に規定の背景カラーコードが埋め込まれているため、基準画像における前景領域のみから適切に基準属性が導出され得る。

[0287] そして、生成器234は、基準属性及び幾何学属性を生成モデルに入力して、生成モデルから中間顔動画像を取得することにより、中間顔動画像を生成する。

[0288] 生成器234で生成される中間顔動画像は、幾何学属性によって動きが与えられた基準画像に対応する。また、基準画像における背景領域に規定の背景カラーコードが埋め込まれている。そのため、生成器234では、背景領域に規定の背景カラーコードが埋め込まれている中間顔動画像が生成される。

[0289] したがって、図16の例と同様に、セグメンテーションマスクを用いることなく、中間顔動画像における背景領域を特定することが可能になる。そして、図16の例と同様に、背景領域に規定の背景カラーコードが埋め込まれた中間顔動画像、及び、背景画像から、合成顔動画像を生成することが可能である。

[0290] 圧縮解除器231は、規定の背景カラーコードを示す背景カラーコード情

報を復号してもよい。そして、合成器 236 は、中間顔動画像において、復号された背景カラーコード情報に示される規定の背景カラーコードを有する領域を背景領域として特定してもよい。

[0291] 図 17 の例の符号化装置 100 において、圧縮器 131 が、基準画像における背景領域に規定の背景カラーコードを埋め込まずに、セグメンテーションマスクをビットストリーム（例えばビットストリームにおける S E I）に符号化してもよい。そして、図 18 の例の復号装置 200 において、圧縮解除器 231 が、セグメンテーションマスク及び基準画像をビットストリームから復号し、セグメンテーションマスクに従って基準画像における背景領域に規定の背景カラーコードを埋め込んでもよい。

[0292] また、図 17 及び図 18 の例に限られず、図 16 の例に対応する符号化装置 100 及び復号装置 200 において、背景カラーコード情報が伝送されてもよい。あるいは、背景カラーコード情報が伝送されることなく、規定の背景カラーコードは、符号化装置 100 及び復号装置 200 において、基準画像からニューラルネットワークを用いて規定されてもよいし、基準画像等によらず規定されてもよい。

[0293] 背景カラーコード情報は、合成顔動画像の生成に対して、1 回のみ伝送されてもよい。そして、復号装置 200 において、合成顔動画像の各フレームの生成時に同じ背景カラーコード情報が用いられてもよい。

[0294] 例えば、規定の背景カラーコードは、顔等の前景領域が有する複数の色（つまり複数の画素値）とは異なるように規定されるカラーコードである。これにより、中間顔動画像等における背景領域が適切に識別され得る。

[0295] 具体的には、規定の背景カラーコードは、全ての典型的な顔又は体の領域にわたって一般的でない色であってもよい。例えば、クロマキーの緑又は青が、人体の色から最も遠い色であるため、規定の背景カラーコードとして選択されてもよい。

[0296] あるいは、まず、全てのあり得る色のリストが構築されてもよい。また、前景領域内のすべての全ての色が抽出されてもよい。次に、抽出された全て

の色がリストから削除されてもよい。そして、リストに残された色が規定の背景カラーコードとして選択されてもよい。

[0297] あるいは、前景領域内の全ての色が頻度表に入力されてもよい。そして、頻度表から最も出現率の高い色が特定されてもよい。そして、特定された色に対して色相環の反対側の色が規定の背景カラーコードとして規定されてもよい。

[0298] また、合成顔動画像の生成において、前景領域および背景カラーコードが埋め込まれた背景領域が有する画素値は変化し得る。また、規定の背景カラーコードが前景領域に含まれる画素値と一致してしまうと、背景領域を適切に特定することが困難になる。

[0299] そこで、規定の背景カラーコードは、前景領域に含まれる画素値と一致しないような連続する複数の値の範囲で規定されてもよい。そして、符号化装置100は、連続する複数の値の範囲を規定の背景カラーコードとして符号化し、復号装置200は、この範囲を規定の背景カラーコードとして復号してもよい。

[0300] 例えば、ビットストリームによって、連続する複数の範囲の最小値 (y_{min} 、 u_{min} 、 v_{min})、及び、連続する複数の範囲の最大値 (y_{max} 、 u_{max} 、 v_{max}) が、背景カラーコードとして伝送されてもよい。あるいは、ビットストリームによって、連続する複数の範囲の中間値 (y_{mean} 、 u_{mean} 、 v_{mean})、及び、連続する複数の範囲の中間値と最小値との差分値 (y_{delta} 、 u_{delta} 、 v_{delta}) が、背景カラーコードとして伝送されてもよい。

[0301] 図19は、本実施の形態における復号装置200のさらに別の構成例を示すブロック図である。この例では、生成器234は、基準属性、幾何学属性及び背景画像から生成モデルを用いて合成顔動画像を生成する。具体的には、生成器234は、基準属性、幾何学属性及び背景画像を生成モデルに入力して、生成モデルから合成顔動画像を取得することにより、合成顔動画像を生成する。これにより、処理が簡素化され得る。この場合、復号装置200は、別途の合成器236を備えていなくてもよい。

- [0302] 上記の複数の例において、背景画像は、顔等の前景領域を含んでいない画像であってもよいし、顔等の前景領域を含む画像であってもよい。背景画像における前景領域には、規定の前景カラーコードが埋め込まれていてもよい。つまり、背景画像における前景領域は、規定の前景色で塗りつぶされていてもよい。これにより、合成顔画像における背景領域に、背景画像における顔等の前景領域が現れることが抑制される。
- [0303] 例えば、符号化装置１００において、圧縮器１３４は、背景画像に対してセグメンテーション処理を行って、背景画像における前景領域及び背景領域を示すセグメンテーションマスクを取得する。そして、圧縮器１３４は、セグメンテーションマスクを用いて、背景画像における前景領域に規定の前景カラーコードを埋め込んでもよい。そして、圧縮器１３４は、前景領域に規定の前景カラーコードが埋め込まれた背景画像をビットストリームに符号化してもよい。
- [0304] そして、復号装置２００において、圧縮解除器２３５は、前景領域に規定の前景カラーコードが埋め込まれた背景画像をビットストリームから復号してもよい。これにより、復号装置２００は、背景画像における前景領域及び背景領域を特定することが可能になる。
- [0305] 背景画像に前景領域が含まれる場合、インペインティング又は他の方法によって、背景画像における背景領域の欠落の一部又は全部が補間されてもよい。具体的には、背景画像における前景領域の周辺領域（つまり背景領域）を用いて、背景画像における背景領域の欠落が補間されてもよい。あるいは、他の画像における背景領域を用いて、背景画像における背景領域の欠落が補間されてもよい。
- [0306] 背景画像における背景領域の欠落の補間は、符号化装置１００において行われてもよいし、復号装置２００において行われてもよい。
- [0307] 例えば、背景画像は、ドライビング動画画像に含まれる画像であってもよいし、ドライビング動画画像に含まれる複数の画像の合成画像でもよい。この場合において、符号化装置１００において、圧縮器１３４は、背景画像におけ

る前景領域の周辺領域、又は、ドライビング動画像に含まれる他の画像における背景領域を用いて、背景画像における背景領域の欠落を補間してもよい。そして、圧縮器134は、背景領域の欠落が補間された背景画像を符号化してもよい。

[0308] また、復号装置200において、圧縮解除器235又は合成器236は、背景画像における前景領域の周辺領域、又は、過去の合成顔動画像における背景領域を用いて、背景画像における背景領域の欠落を補間してもよい。また、合成器236が、中間動画像における背景領域に背景画像における対応領域の埋め込むタイミングにおいて、当該対応領域が前景領域を含んでいれば、当該対応領域における背景領域の欠落を補間してもよい。

[0309] 上述した通り、背景画像の代わりに選択パラメータがビットストリームによって伝送されてもよい。選択パラメータは、背景に関する情報を提供する。

[0310] 選択パラメータは、背景画像のコンテンツレーティング（コンテンツ評価）に関する情報を含んでもよい。具体的には、選択パラメータは、背景画像に関する閲覧に適した年齢層をコンテンツレーティングとして示してもよい。

[0311] 例えば、選択パラメータは、背景画像にNC16（Not available for children below 16 years old）のレーティングが割り当てられていることを示してもよい。そして、視聴者がより小さいカテゴリに含まれる場合、選択パラメータに従って、背景画像をぼかすなど、背景画像に対してさらなる処理が実行されてもよい。

[0312] これにより、メディアコンテンツのレーティングシステムと同様に、未成年者又はその他のカテゴリの人々を暴力などのテーマ的にふさわしくない背景画像の閲覧から保護することが可能になる。

[0313] また、例えば、選択パラメータは、背景画像のカスタマイズに関する情報を含んでもよい。具体的には、選択パラメータは、視聴者のプロフィールに基づいて、背景画像をさらにカスタマイズするための情報を含んでいて

もよい。例えば、視聴者が女性である場合、女性に選ばれることが多い背景画像（具体的にはピンク色の背景画像等）が選択されてもよい。別の例において、視聴者が子供である場合、子供に選ばれることが多い背景画像（具体的にはカラフルな背景画像等）が選択されてもよい。

[0314] このような選択パラメータ等によって、1つ以上の復号装置200のそれぞれで適用される最終的な背景画像が変更されてもよい。

[0315] 上記の複数の例において、基準画像及び幾何学属性とは別に、背景画像、又は、背景画像に関する背景情報が伝送される。背景情報は、例えば、背景画像を取得するための情報であって、複数の背景画像候補から背景画像を選択するための情報であってもよいし、背景画像の取得先等のような背景画像を取得するための情報であってもよい。しかしながら、別の例において、背景画像又は背景情報が伝送されなくてもよいし、背景画像が予め定められていなくてもよい。

[0316] 具体的には、例えば、基準画像とは別に伝送される背景画像の代わりに、基準画像が背景画像として用いられてもよい。そして、例えば、復号装置200において、合成器236は、中間顔動画像における背景領域に、基準画像における対応領域を埋め込むことにより、合成顔動画像を生成してもよい。

[0317] インペインティング又は他の方法によって、基準画像における背景領域の欠落の一部又は全部が補間されてもよい。具体的には、基準画像における前景領域の周辺領域（つまり背景領域）を用いて、基準画像における背景領域の欠落が補間されてもよい。

[0318] あるいは、他の画像における背景領域を用いて、基準画像における背景領域の欠落が補間されてもよい。複数の基準画像が伝送される場合、他の画像は、他の基準画像であってもよい。あるいは、他の画像は、他の基準画像を用いて生成された中間顔動画像又は合成顔動画像に含まれる画像であってもよい。

[0319] また、基準画像に対してセグメンテーション処理が行われて、基準画像に

おける前景領域及び背景領域を示すセグメンテーションマスクが取得されてもよい。そして、セグメンテーションマスクが、基準画像における前景領域及び背景領域の特定に用いられてもよい。セグメンテーション処理は、符号化装置１００で行われてもよいし、復号装置２００で行われてもよい。セグメンテーション処理が符号化装置１００で行われた場合、セグメンテーションマスクが符号化装置１００から復号装置２００へ伝送されてもよい。

[0320] 上記の例において、基準画像が背景画像又は背景情報として符号化され復号されるとみなされてもよい。また、基準画像が背景画像に適用されるとみなされてもよい。あるいは、背景画像に関して符号化も復号も行われなるとみなされてもよい。

[0321] また、上記の複数の例において、幾何学属性を示す幾何学情報が用いられているが、各フレームに対応する属性情報は、幾何学属性を示す幾何学情報に限られない。幾何学属性とは異なる形式で動的属性を示す動的情報が幾何学情報の代わりに用いられてもよい。

[0322] 図２０は、本実施の形態における符号化装置１００が動画画像を符号化するための構成例を示すブロック図である。例えば、符号化装置１００は、動画画像に含まれる画像をブロック単位でＶＶＣに従って符号化するための複数の構成要素として、図２０に示された複数の構成要素を含んでいてもよい。符号化装置１００は、上述された複数の構成要素に加えて、図２０に示された複数の構成要素を含んでいてもよいし、上述された複数の構成要素の少なくとも一部が、図２０に示された複数の構成要素に統合されてもよい。

[0323] 図２０に示すように、符号化装置１００は、分割部１０２と、減算部１０４と、変換部１０６と、量子化部１０８と、エントロピー符号化部１１０と、逆量子化部１１２と、逆変換部１１４と、加算部１１６と、ブロックメモリ１１８と、ループフィルタ部１２０と、フレームメモリ１２２と、イントラ予測部１２４と、インター予測部１２６と、予測制御部１２８と、予測パラメータ生成部１３０とを備える。なお、イントラ予測部１２４及びインター予測部１２６のそれぞれは、予測処理部の一部として構成されている。

- [0324] 分割部 102 は、画像を複数のブロックに分割し、分割に関するパラメータをエントロピー符号化部 110 に提供する。減算部 104 は、カレントブロックから予測画像ブロックを減算して予測残差ブロックを取得する。変換部 106 は、予測残差ブロックに対して変換を行って変換係数ブロックを取得する。量子化部 108 は、変換係数ブロックに対して量子化を行って量子化係数ブロックを取得する。エントロピー符号化部 110 は、量子化係数ブロック及びパラメータに対してエントロピー符号化を行ってビットストリームを生成する。
- [0325] 逆量子化部 112 は、量子化係数ブロックに対して逆量子化を行って変換係数ブロックを取得する。逆変換部 114 は、変換係数ブロックに対して逆変換を行って予測残差ブロックを取得する。加算部 116 は、予測残差ブロックに予測画像ブロックを加算して再構成画像ブロックを取得する。ブロックメモリ 118 には、再構成画像ブロックが記憶される。ループフィルタ部 120 は、再構成画像ブロックにループフィルタを適用する。フレームメモリ 122 には、ループフィルタが適用された再構成画像ブロックが記憶される。
- [0326] イントラ予測部 124 は、ブロックメモリ 118 を参照してイントラ予測を行って予測画像ブロックを生成する。インター予測部 126 は、フレームメモリ 122 を参照してインター予測を行って予測画像ブロックを生成する。予測制御部 128 は、イントラ予測部 124 が生成する予測画像ブロック、又は、インター予測部 126 が生成する予測画像ブロックを減算部 104 及び加算部 116 に提供する。予測パラメータ生成部 130 は、イントラ予測又はインター予測に関するパラメータをエントロピー符号化部 110 に提供する。
- [0327] 図 21 は、本実施の形態における復号装置 200 が動画像を復号するための構成例を示すブロック図である。例えば、復号装置 200 は、動画像に含まれる画像をブロック単位で VVC に従って復号するための複数の構成要素として、図 21 に示された複数の構成要素を含んでもよい。復号装置 2

00は、上述された複数の構成要素に加えて、図21に示された複数の構成要素を含んでもよいし、上述された複数の構成要素の少なくとも一部が、図21に示された複数の構成要素に統合されてもよい。

[0328] 図21に示すように、復号装置200は、エントロピー復号部202と、逆量子化部204と、逆変換部206と、加算部208と、ブロックメモリ210と、ループフィルタ部212と、フレームメモリ214と、イントラ予測部216と、インター予測部218と、予測制御部220と、予測パラメータ生成部222と、分割決定部224とを備える。なお、イントラ予測部216及びインター予測部218のそれぞれは、予測処理部の一部として構成されている。

[0329] エントロピー復号部202は、ビットストリームに対してエントロピー復号を行って量子化係数ブロック及びパラメータを取得する。逆量子化部204は、量子化係数ブロックに対して逆量子化を行って変換係数ブロックを取得する。逆変換部206は、変換係数ブロックに対して逆変換を行って予測残差ブロックを取得する。加算部208は、予測残差ブロックに予測画像ブロックを加算して再構成画像ブロックを取得する。ループフィルタ部212は、再構成画像ブロックにループフィルタを適用する。

[0330] ブロックメモリ210には、再構成画像ブロックが記憶される。フレームメモリ214には、ループフィルタが適用された再構成画像ブロックが記憶される。

[0331] イントラ予測部216は、ブロックメモリ210を参照してイントラ予測を行って予測画像ブロックを生成する。インター予測部218は、フレームメモリ214を参照してインター予測を行って予測画像ブロックを生成する。予測制御部220は、イントラ予測部216が生成する予測画像ブロック、又は、インター予測部218が生成する予測画像ブロックを加算部208に提供する。予測パラメータ生成部222は、イントラ予測又はインター予測に関するパラメータを予測制御部220等に提供する。

[0332] 分割決定部224は、分割に関するパラメータに従って、画像をブロック

単位で復号するためのブロックを決定する。

[0333] [ビットストリームの構成例]

本開示では、ビットストリームを介して、基準画像、幾何学属性及び背景画像の3種類の情報が伝送される。

[0334] 図22、図23、図24、図25及び図26に、ビットストリームレイアウトが示される。ビットストリームレイアウトは、図22、図23、図24、図25及び図26の組み合わせであってもよい。また、図22、図23、図24、図25及び図26に示されるビットストリームレイアウトに限定されず、基準画像、幾何学属性及び背景画像は、任意の順序で符号化され得る。

[0335] 図22は、ビットストリームの構成例を示す概念図である。この例では、最初のGOPに、1つ以上の基準画像、1つ以上の背景画像、及び、複数の幾何学属性が含まれ、他の各GOPに、1つ以上の背景画像、及び、複数の幾何学属性が含まれる。

[0336] 図23は、ビットストリームの別の構成例を示す概念図である。この例では、最初のGOPに、基準画像、背景画像、及び、複数の幾何学属性が含まれ、他の各GOPに、複数の幾何学属性が含まれる。

[0337] 図24は、ビットストリームのさらに別の構成例を示す概念図である。この例では、ビットストリームに、第1ビットストリーム、第2ビットストリーム及び第3ビットストリームが含まれる。第1ビットストリームに、1つ以上の基準画像が含まれる。第2ビットストリームに、1つ以上の背景画像が含まれる。第3ビットストリームに、複数の幾何学属性が含まれる。

[0338] 図25は、ビットストリームのさらに別の構成例を示す概念図である。この例では、ビットストリームに、第1ビットストリーム及び第2ビットストリームが含まれる。第1ビットストリームに、1つ以上の基準画像が含まれる。第2ビットストリームにおける最初のGOPに、1つ以上の背景画像、及び、複数の幾何学属性が含まれる。第2ビットストリームにおける他の各GOPに、複数の幾何学属性が含まれる。

- [0339] 図26は、ビットストリームのさらに別の構成例を示す概念図である。この例では、各GOPに、1つ以上の基準画像、1つ以上の背景画像、及び、複数の幾何学属性が含まれる。
- [0340] 例えば、ビットストリームに複数の背景画像が含まれる場合、複数の背景画像のうちの1つの背景画像が合成顔画像の生成に用いられてもよいし、複数の背景画像の組み合わせが合成顔画像の生成に用いられてもよい。あるいは、ビットストリームに含まれる複数の背景画像は、それぞれ、合成顔画像の複数のフレームに対応していてもよい。
- [0341] また、例えば、ビットストリームに複数の基準画像が含まれる場合、複数の基準画像のうちの1つの基準画像が合成顔画像の生成に用いられてもよいし、複数の基準画像の組み合わせが合成顔画像の生成に用いられてもよい。
- [0342] また、例えば、アクセスユニットに含まれるコンテンツを識別するため、基準画像、背景画像又は幾何学属性のうちどれにアクセスユニットが対応するかを示すヘッダパラメータが用いられてもよい。ヘッダパラメータは、SPS、PPS、PH、VUI又はSEIに符号化されてもよい。
- [0343] また、合成顔動画は、最初の幾何学属性を含むアクセスユニットの復号後に、復号装置200でレンダリングされ表示されてもよい。
- [0344] コーデック規格において、各アクセスユニットに、1つのピクチャに対応するピクチャデータが含まれることが規定されている場合がある。そのため、基準画像と背景画像とはそれぞれ別のアクセスユニットに含まれていてもよい。あるいは、例えば、マルチビュー符号化等のレイヤ符号化では、1つのアクセスユニットに複数のピクチャに対応するピクチャデータが含まれることが許容される場合がある。そのため、マルチビュー符号化等のレイヤ符号化に基づいて、基準画像と背景画像とが同じアクセスユニットに含まれていてもよい。
- [0345] 各アクセスユニットは、ビデオ符号化レイヤ（VCL: Video Coding Layer）又は非ビデオ符号化レイヤの他のNALユニットを含んでいてもよい。

- [0346] また、顔ランドマークデータ等の幾何学属性は、任意の動画像コーデック又は画像コーデックにおけるメタデータ、VSEI (Versatile Supplemental Enhancement Information) 又はSEIに含まれていてもよい。また、幾何学属性が、VCLにおいて新しいnal_unit_typeを有するNALユニットに含まれていてもよい。
- [0347] 図27、図28及び図29に、符号化規格であるVVCコーデックを用いるビットストリームレイアウトの例が示されている。VVCコーデックは、例えば、HEVC、AVC、AV1、SVC、EVC、JPEGなど、他のビデオコーデック又は画像コーデックに置き換えられてもよい。
- [0348] 図27は、VVCに準拠するビットストリームの構成例を示す概念図である。この例において、基準画像及び背景画像のそれぞれが、VVCのイントラピクチャとして符号化され復号される。幾何学属性は、幾何学属性SEIと表現されるSEIに符号化され、当該SEIから復号される。
- [0349] この例では、背景画像が、複数回、伝送される。最初のアクセスユニットに、基準画像のピクチャデータが含まれ、他の各アクセスユニットに、幾何学属性SEIと背景画像のピクチャデータとが含まれる。つまり、幾何学属性と背景画像とは、同じアクセスユニットに含まれる。
- [0350] 図28は、VVCに準拠するビットストリームの別の構成例を示す概念図である。この例において、基準画像及び背景画像のそれぞれが、VVCのイントラピクチャとして符号化され復号される。幾何学属性は、幾何学属性SEIと表現されるSEIに符号化され、当該SEIから復号される。
- [0351] この例では、背景画像が、1回のみ、伝送される。最初のアクセスユニットに、基準画像のピクチャデータが含まれ、次のアクセスユニットに、幾何学属性SEIと背景画像のピクチャデータとが含まれる。その後の各アクセスユニットに、幾何学属性SEIが含まれる。
- [0352] 図29は、VVCに準拠するビットストリームのさらに別の構成例を示す概念図である。この例において、基準画像が、VVCのイントラピクチャと

して符号化され復号される。幾何学属性は、幾何学属性 S E I と表現される S E I に符号化され、当該 S E I から復号される。また、背景画像の代わりに、複数の背景画像候補から背景画像を選択するための選択パラメータ等の背景パラメータが、背景パラメータ S E I と表現される S E I に符号化され、当該 S E I から復号される。

[0353] また、最初のアクセスユニットに、基準画像のピクチャデータが含まれ、次のアクセスユニットに、幾何学属性 S E I と背景パラメータ S E I とピクチャデータとが含まれる。その後の各アクセスユニットに、幾何学属性 S E I が含まれていてもよいし、幾何学属性 S E I とピクチャデータとが含まれていてもよい。

[0354] 例えば、コーデック規格において、アクセスユニット毎に 1 つのピクチャに対応するピクチャデータが含まれることが規定されている場合がある。このようなコーデック規格に対する互換性を確保するため、S E I を伝送するためのアクセスユニットには、S E I のみでなく、ピクチャデータが含まれていてもよい。

[0355] 具体的には、S E I を伝送するためのアクセスユニットに、ダミーのピクチャデータとして、最小許容解像度のピクチャに対応するピクチャデータが含まれていてもよい。このピクチャは、全ての画素に対してゼロのような一定の値を有するピクチャであってもよい。各画素のカラーコードがビットストリームに符号化される場合、各画素は同じカラーコードで埋められてもよい。

[0356] また、S E I を伝送するためのアクセスユニットに、黒一色、白一色又はクロマキーの緑一色等のような無地の画像を示すスライス N A L ユニットがピクチャデータとして含まれていてもよい。

[0357] あるいは、S E I を伝送するためのアクセスユニットに、基準画像又は背景画像のコピーがピクチャデータとして含まれていてもよい。当該アクセスユニットに、スキップピクチャがピクチャデータとして含まれていてもよい。あるいは、当該アクセスユニットに、全ての C U (C o d i n g U n i

t) にスキップモードが指定されているピクチャがピクチャデータとして含まれていてもよい。これにより、各アクセスユニットのオーバーヘッドを最小限に抑えることが可能になる。

[0358] あるいは、ピクチャデータに対応するNALユニットを無視すべきことを示すパラメータが、SEIを伝送するためのアクセスユニットに含まれていてもよい。

[0359] 別の例において、アクセスユニットに、基準画像のピクチャデータと、幾何学属性SEI及び背景パラメータSEIの少なくとも一方が含まれていてもよい。また、他のアクセスユニットに、幾何学属性SEI及び背景パラメータSEIの少なくとも一方が含まれていてもよい。幾何学属性SEIの代わりに、新しいnal__unit__typeを有するNALユニットが用いられてもよい。また、背景パラメータSEIの代わりに、新しいnal__unit__typeを有するNALユニットが用いられてもよい。

[0360] また、幾何学属性と背景パラメータとは、それぞれ個別のSEIに符号化される必要はなく、1つのSEIに符号化されていてもよい。

[0361] また、具体的な例として、各GOPの最初の同じアクセスユニットに、基準画像のピクチャデータと幾何学属性SEIと背景パラメータSEIとが含まれていてもよい。そして、合成顔動画の生成において、同じアクセスユニットから得られる基準画像と幾何学属性と背景情報とに基づいて、合成顔動画におけるGOPの最初のフレームが生成されてもよい。あるいは、処理対象のGOPの最初のフレームの生成に、その前のGOPの最初のアクセスユニットに含まれる基準画像が用いられてもよい。

[0362] また、最初のGOPにおいて、最初のアクセスユニットに基準画像のピクチャデータが含まれ、その後の各アクセスユニットに幾何学属性SEIと背景パラメータSEIとが含まれていてもよい。そして、その後の各GOPにおいて、最初のアクセスユニットに基準画像のピクチャデータと幾何学属性SEIと背景パラメータSEIとが含まれ、その後の各アクセスユニットに幾何学属性SEIと背景パラメータSEIとが含まれていてもよい。

[0363] これにより、幾何学属性SEIと背景パラメータSEIとを用いてフレームを生成する際に、既に処理済みのアクセスユニットから得られた基準画像を用いることが可能になる。

[0364] [生成モデルの例]

図30は、生成モデルとして用いられ得る様々なモデルの例を示す図である。例えば、ニューラルネットワークが生成モデルとして用いられる。具体的には、図30において、敵対的生成ネットワーク、変分オートエンコーダ、フローベース生成モデル、及び、拡散モデルが示されている。

[0365] 敵対的生成ネットワークでは、入力データの特徴の学習を介して入力データに類似する新たなデータインスタンスが生成される。具体的には、生成モデルにおける教師なしタスクが、2種類のサブモデルによって教師ありタスクへ変換される。

[0366] 例えば、生成器サブモデルが、偽物のサンプルを生成し、識別器サブモデルが、本物の入力と、生成器サブモデルで生成された偽物のサンプルとを識別する。そして、本物の入力と偽物のサンプルとに対する正確なラベルの割り当てにおける識別器サブモデルの識別確率を最大にすると同時に、本物の入力と偽物のサンプルとの分布差を最小にするミニマックスゲームを介し、出力画像が生成される。

[0367] 変分オートエンコーダでは、まず、できるだけ正確に潜在空間からデータを再構成するための多変量潜在分布に入力データが圧縮される。これにより、データ圧縮及び次元削減が効率的に行われる。フローベース生成モデルでは、1つ以上の可逆変換のシーケンスを介してソース分布が訓練データの分布に変換される。これにより、データ分布の学習及び最終目標の尤度の正確な算出が可能になる。

[0368] 拡散モデルでも、訓練データに類似する新たなデータインスタンスが生成される。まず、拡散モデルでは、オリジナルデータを復元する試みにおいてノイズ除去処理を始める前に、摂動及びノイズの反復注入を介して訓練データの構造を劣化させる。結果として、各ステップの潜在状態が前のステップ

の潜在状態にのみ依存するマルコフ連鎖を介して、データが潜在分布に繰り返しマッピングされる。そして、階層的なノイズ除去によりデータが復元される。

[0369] 例えば、ニューラルネットワークは、顔パラメータのための固定フォーマットで表現される幾何学情報と、ピクチャとを用いて、出力ピクチャを生成するために用いられ得る顔ピクチャ生成ニューラルネットワークであってもよい。すなわち、ニューラルネットワークは、出力動画像に含まれる1つのピクチャである出力ピクチャを構成する複数のサンプル値を生成する処理に対応する。

[0370] また、ニューラルネットワークは、MPEG (Moving Picture Experts Group) で議論されている生成顔動画像SEI (generative face video SEI) のニューラルネットワークであってもよい。具体的には、例えば、非特許文献2において、「Generative NN ()」と表記される顔ピクチャ生成ニューラルネットワーク (face picture generator neural network) であってもよい。

[0371] 上述されたニューラルネットワークの代替例は、上述されたニューラルネットワークの任意の組み合わせであってもよい。また、他の種類の生成モデル等が用いられてもよい。

[0372] また、ニューラルネットワーク等の機械学習モデルは、セグメンテーション処理に用いられてもよい。また、ニューラルネットワーク等の機械学習モデルは、幾何学属性の導出に用いられてもよいし、基準属性の導出に用いられてもよい。

[0373] [実装例]

図31は、符号化装置100の実装例を示すブロック図である。符号化装置100は、回路151及びメモリ152を備える。例えば、上述された符号化装置100の複数の構成要素は、回路151及びメモリ152によって実装される。

[0374] 回路151は、情報処理を行う電気回路であり、メモリ152にアクセス可能である。例えば、回路151は、本開示の符号化方法を実行する専用の回路であってもよいし、本開示の符号化方法に対応するプログラムを実行する汎用の回路であってもよい。また、回路151は、CPU等のプロセッサであってもよい。さらに、回路151は、複数の回路の集合体であってもよい。

[0375] メモリ152は、回路151が画像を符号化するための情報が記憶される専用又は汎用のメモリである。メモリ152は、電気回路であってもよく、回路151に接続されていてもよい。また、メモリ152は、回路151に含まれてもよい。また、メモリ152は、複数の回路の集合体であってもよい。また、メモリ152は、磁気ディスク又は光ディスク等であってもよいし、ストレージ又は記録媒体等と表現されてもよい。また、メモリ152は、不揮発性メモリであってもよいし、揮発性メモリであってもよい。

[0376] 例えば、メモリ152には、画像等の符号化対象データが記憶されてもよいし、ビットストリーム等の符号化済みデータが記憶されてもよい。また、メモリ152には、回路151に画像処理を行わせるためのプログラムが記憶されていてもよい。また、メモリ152には、回路151に生成モデルが記憶されていてもよい。また、メモリ152には、基準画像が記憶されてもよい。

[0377] 図32は、符号化装置100の基本的な第1動作例を示すフローチャートである。符号化装置100の回路151は、この例の動作において、メモリ152を用いて、以下を行う。

[0378] 具体的には、回路151は、合成顔動画画像を生成するための基準画像と幾何学情報と背景情報とを1つ以上のストリームに符号化する(S301)。合成顔動画画像は、顔を含む動画画像であり、背景画像が合成された動画画像である。基準画像は、顔を含む画像である。幾何学情報は、カメラによって得られる撮影動画画像の複数のフレームの各々に対応し、被写体の幾何学属性を示す情報である。背景情報は、背景画像に関する情報である。

- [0379] これにより、合成顔動画像を生成するための基準画像、幾何学属性及び背景画像を提供することが可能になる場合がある。したがって、合成顔動画像の生成において、各フレームに対応する幾何学属性によって基準画像の顔に動きを与えつつ、背景画像によって基準画像の背景の歪みを抑制することが可能になる場合がある。よって、画質の劣化の抑制に貢献することが可能になる場合がある。
- [0380] 例えば、回路151は、撮影動画像に対してセグメンテーション処理を行って、撮影動画像における前景領域及び背景領域を示す撮影動画像セグメンテーション情報を取得してもよい。そして、回路151は、撮影動画像セグメンテーション情報を1つ以上のストリームに符号化してもよい。
- [0381] これにより、1つ以上のストリームを介して、撮影動画像と同種の動画像における前景領域及び背景領域を特定するための撮影動画像セグメンテーション情報を提供することが可能になる場合がある。したがって、各フレームに対応する幾何学属性によって基準画像の顔に動きが与えられた中間顔動画像における前景領域及び背景領域の特定に貢献することが可能になる場合がある。
- [0382] また、例えば、回路151は、基準画像に対してセグメンテーション処理を行って、基準画像における前景領域及び背景領域を示す基準画像セグメンテーション情報を取得してもよい。また、回路151は、基準画像セグメンテーション情報を用いて、基準画像における背景領域に規定の背景カラーコードを埋め込んでもよい。そして、回路151は、背景領域に規定の背景カラーコードが埋め込まれた基準画像を符号化してもよい。
- [0383] これにより、基準画像に対するセグメンテーション処理の結果として得られる基準画像セグメンテーション情報に従って、基準画像における背景領域を適切に特定することが可能になる場合がある。また、基準画像における背景領域に規定の背景カラーコードが埋め込まれるため、基準画像の顔に動きが与えられても、基準画像の背景に生じる歪みを抑制することが可能になる場合がある。

- [0384] また、例えば、回路 151 は、規定の背景カラーコードを示す背景カラーコード情報を 1 つ以上のストリームに符号化してもよい。これにより、1 つ以上のストリームを介して、基準画像における背景領域を効率的に特定するための規定の背景カラーコードを提供することが可能になる場合がある。そして、基準画像に応じて規定の背景カラーコードを変化させることが可能になる場合がある。
- [0385] また、例えば、背景カラーコード情報は、連続する複数の値を含む範囲を規定の背景カラーコードとして示してもよい。そして、規定の背景カラーコードは、背景カラーコード情報によって示される範囲内で規定されてもよい。これにより、規定の背景カラーコードを柔軟に規定することが可能になる場合がある。そして、背景領域に規定の背景カラーコードを柔軟に適用することが可能になる場合がある。
- [0386] また、例えば、規定の背景カラーコードは、基準画像における前景領域において発生頻度が閾値以下であるカラーコードで規定されてもよい。これにより、前景領域の部分が背景領域の部分と誤って特定されることを抑制することが可能になる場合がある。したがって、背景領域を適切に特定することが可能になる場合がある。
- [0387] また、例えば、回路 151 は、基準画像に対してセグメンテーション処理を行って、基準画像における前景領域及び背景領域を示す基準画像セグメンテーション情報を取得してもよい。そして、回路 151 は、基準画像セグメンテーション情報を 1 つ以上のストリームに符号化してもよい。
- [0388] これにより、1 つ以上のストリームを介して、基準画像における前景領域及び背景領域を特定するための基準画像セグメンテーション情報を提供することが可能になる場合がある。したがって、基準画像における背景領域の特定に貢献することが可能になる場合がある。
- [0389] また、例えば、背景画像は、基準画像及び撮影動画像によらずに用意された画像であってもよい。これにより、基準画像及び撮影動画像とは別途に用意された背景画像を合成顔動画像に適用することが可能になる場合がある。

したがって、背景画像における前景領域の影響などを抑制することが可能になる場合がある。

[0390] また、例えば、回路151は、複数の背景画像候補の中から背景画像を選択してもよい。そして、回路151は、背景画像の識別子を背景情報として符号化してもよい。これにより、複数の背景画像候補の中から背景画像を柔軟に選択することが可能になる場合がある。したがって、合成顔動画像の用途に応じて、合成顔動画像に適切な背景画像を適用することが可能になる場合がある。

[0391] また、例えば、背景画像は、撮影動画像に含まれる画像であってもよいし、背景画像は、撮影動画像に含まれる複数の画像が合成された画像であってもよい。これにより、撮影動画像から得られる背景画像を合成顔動画像に適用することが可能になる場合がある。したがって、撮影状況に対応する背景画像を合成顔動画像に適用することが可能になる場合がある。

[0392] また、例えば、回路151は、基準画像を背景情報として符号化してもよい。そして、背景画像に、基準画像が適用されてもよい。これにより、基準画像を背景画像として用いることが可能になる場合がある。そして、各フレームに対応する幾何学属性によって基準画像の顔に動きを与えつつ、元の基準画像が背景画像として用いられることで基準画像の背景の歪みを抑制することが可能になる場合がある。

[0393] また、例えば、回路151は、背景画像に前景領域が含まれる場合、背景画像における前景領域の周辺領域、又は、撮影動画像に含まれる他の画像における背景領域を用いて、背景画像における背景領域の欠落を補間してもよい。これにより、背景画像に前景領域が含まれていても、背景領域の欠落を適切に補間することが可能になる場合がある。したがって、合成顔動画像における背景領域の欠落を抑制することが可能になる場合がある。

[0394] また、例えば、回路151は、背景画像に対してセグメンテーション処理を行って、背景画像における前景領域及び背景領域を示す背景画像セグメンテーション情報を取得してもよい。そして、回路151は、背景画像セグメ

ンテーション情報を用いて、背景画像における前景領域及び背景領域を特定してもよい。

[0395] これにより、背景画像に対するセグメンテーション処理の結果として得られる背景画像セグメンテーション情報に従って、背景画像における前景領域及び背景領域を適切に特定することが可能になる場合がある。したがって、背景画像における背景領域の欠落を適切に補間することが可能になる場合がある。

[0396] また、例えば、回路151は、背景画像に対してセグメンテーション処理を行って、背景画像における前景領域及び背景領域を示す背景画像セグメンテーション情報を取得してもよい。そして、回路151は、背景画像セグメンテーション情報を用いて、背景画像における前景領域に規定の前景カラーコードを埋め込んでもよい。

[0397] これにより、規定の前景カラーコードに従って、背景画像における前景領域を効率的に特定することが可能になる場合がある。また、合成顔動画像における背景領域に顔等の前景が映り込むことを抑制することが可能になる場合がある。

[0398] また、例えば、回路151は、規定の前景カラーコードを示す前景カラーコード情報を1つ以上のストリームに符号化してもよい。これにより、1つ以上のストリームを介して、背景画像における前景領域を効率的に特定するための規定の前景カラーコードを提供することが可能になる場合がある。そして、背景画像に応じて規定の前景カラーコードを変化させることが可能になる場合がある。

[0399] また、例えば、前景カラーコード情報は、連続する複数の値を含む範囲を規定の前景カラーコードとして示してもよい。そして、規定の前景カラーコードは、前景カラーコード情報によって示される範囲内で規定されてもよい。これにより、規定の前景カラーコードを柔軟に規定することが可能になる場合がある。そして、前景領域に規定の前景カラーコードを柔軟に適用することが可能になる場合がある。

- [0400] また、例えば、規定の前景カラーコードは、背景画像における背景領域において発生頻度が閾値以下であるカラーコードで規定されてもよい。これにより、背景領域の部分が前景領域の部分と誤って特定されることを抑制することが可能になる場合がある。したがって、前景領域を適切に特定することが可能になる場合がある。
- [0401] また、例えば、1つ以上のストリームにおいて、背景情報が符号化されるストリームは、基準画像が符号化されるストリームと同じであってもよいし、幾何学情報が符号化されるストリームと同じであってもよい。これにより、別途のストリームではなく、基準画像のストリーム又は幾何学情報のストリームと同じストリームに背景情報を符号化することが可能になる場合がある。したがって、基準画像又は幾何学情報と一緒に背景情報を効率的に符号化することが可能になる場合がある。
- [0402] また、例えば、1つ以上のストリームにおいて、背景情報が符号化されるストリームは、基準画像が符号化されるストリームとも幾何学情報が符号化されるストリームとも異なってもよい。これにより、同じストリームではなく、基準画像のストリーム及び幾何学情報のストリームとは別途のストリームから背景情報に符号化することが可能になる場合がある。したがって、基準画像又は幾何学情報とは別に背景情報を任意のタイミングで符号化することが可能になる場合がある。
- [0403] また、例えば、背景画像は、複数のピクチャを含むシーケンスの先頭のピクチャとして符号化されてもよいし、GOPの先頭のピクチャとして符号化されてもよい。これにより、早期に背景画像を提供することが可能になる場合がある。したがって、早期に背景画像を合成顔動画像に適用することが可能になる場合がある。
- [0404] また、例えば、背景画像は、1つ以上のストリームにおけるアクセスユニットにピクチャとして符号化されてもよい。これにより、背景画像をアクセスユニット内のピクチャとして処理することが可能になる場合がある。つまり、通常のピクチャと同様に背景画像を処理することが可能になる場合があ

る。

[0405] また、例えば、背景画像が符号化されるアクセスユニットは、基準画像が符号化されるアクセスユニットと同じであってもよい。これにより、別途のアクセスユニットではなく、基準画像のアクセスユニットと同じアクセスユニットに背景画像を符号化することが可能になる場合がある。したがって、基準画像と一緒に背景画像を効率的に符号化することが可能になる場合がある。

[0406] また、例えば、背景画像が符号化されるアクセスユニットは、基準画像が符号化されるアクセスユニットとは異なってもよい。これにより、同じアクセスユニットではなく、基準画像のアクセスユニットとは別途のアクセスユニットに背景画像を符号化することが可能になる場合がある。したがって、基準画像とは別に背景画像を任意のタイミングで符号化することが可能になる場合がある。

[0407] また、例えば、背景画像が符号化されるアクセスユニットに対応付けられたSEIに、そのアクセスユニットに背景画像が存在することを示す信号が符号化されてもよい。これにより、アクセスユニットにおけるSEIの信号によって、アクセスユニットにおける背景画像の存在を通知することが可能になる場合がある。したがって、背景画像の適切な伝達が可能になる場合がある。

[0408] また、例えば、背景画像は、1つ以上のストリームにおけるアクセスユニットにピクチャとして符号化されてもよい。また、背景画像が符号化されるアクセスユニットに対応付けられたSEIに、アクセスユニットに背景画像が存在することを示す信号が符号化されてもよい。

[0409] これにより、背景画像をアクセスユニット内のピクチャとして処理することが可能になる場合がある。つまり、通常のピクチャと同様に背景画像を処理することが可能になる場合がある。また、これにより、アクセスユニットにおけるSEIの信号によって、アクセスユニットにおける背景画像の存在を通知することが可能になる場合がある。したがって、背景画像の適切な伝

達が可能になる場合がある。

[0410] また、例えば、背景画像は、イントラピクチャとして符号化されてもよい。これにより、背景画像をイントラピクチャとして処理することが可能になる場合がある。つまり、他のピクチャに依存せずに、背景画像を処理することが可能になる場合がある。

[0411] また、例えば、回路 151 は、背景カラーコード情報を 1 つ以上のストリームにおける S E I に符号化してもよい。これにより、S E I を介して、基準画像における背景領域を効率的に特定するための規定の背景カラーコードを提供することが可能になる場合がある。そして、基準画像に応じて規定の背景カラーコードを変化させることが可能になる場合がある。

[0412] また、例えば、回路 151 は、前景カラーコード情報を 1 つ以上のストリームにおける S E I に符号化してもよい。これにより、S E I を介して、背景画像における前景領域を効率的に特定するための規定の前景カラーコードを提供することが可能になる場合がある。そして、背景画像に応じて規定の前景カラーコードを変化させることが可能になる場合がある。

[0413] また、例えば、回路 151 は、規定の背景カラーコードを示す背景カラーコード情報と、規定の前景カラーコードを示す前景カラーコード情報とのうち少なくとも一方を 1 つ以上のストリームにおける S E I に符号化してもよい。これにより、S E I を介して、前景領域を効率的に特定するための規定の前景カラーコードを提供することが可能になる場合がある。

[0414] また、例えば、回路 151 は、撮影動画像セグメンテーション情報を 1 つ以上のストリームにおける S E I に符号化してもよい。これにより、S E I を介して、撮影動画像と同種の動画像における前景領域及び背景領域を特定するための撮影動画像セグメンテーション情報を提供することが可能になる場合がある。したがって、各フレームに対応する幾何学属性によって基準画像の顔に動きが与えられた中間顔動画像における前景領域及び背景領域の特定に貢献することが可能になる場合がある。

[0415] また、例えば、回路 151 は、基準画像セグメンテーション情報を 1 つ以

上のストリームにおけるSEIに符号化してもよい。これにより、SEIを介して、基準画像における前景領域及び背景領域を特定するための基準画像セグメンテーション情報を提供することが可能になる場合がある。したがって、基準画像における背景領域の特定に貢献することが可能になる場合がある。

[0416] また、例えば、回路151は、撮影動画像セグメンテーション情報と基準画像セグメンテーション情報とのうち少なくとも一方を1つ以上のストリームにおけるSEIに符号化してもよい。撮影動画像セグメンテーション情報は、撮影動画像における前景領域及び背景領域を示す。基準画像セグメンテーション情報は、基準画像における前景領域及び背景領域を示す。これにより、SEIを介して、前景領域及び背景領域を特定するためのセグメンテーション情報を提供することが可能になる場合がある。

[0417] 図33は、符号化装置100の基本的な第2動作例を示すフローチャートである。符号化装置100の回路151は、この例の動作において、メモリ152を用いて、以下を行う。

[0418] 具体的には、回路151は、合成顔動画像を生成するための基準画像と幾何学情報とを1つ以上のストリームに符号化する(S311)。基準画像は、顔を含む画像である。幾何学情報は、カメラによって得られる撮影動画像の複数のフレームの各々に対応し、被写体の幾何学属性を示す情報である。

[0419] 合成顔動画像の生成において、基準画像及び幾何学情報は、基準画像及び幾何学情報を生成モデルに入力して、中間顔動画像を生成モデルから取得することに用いられる。中間顔動画像は、顔を含む動画像である。基準画像は、中間顔動画像における背景領域に基準画像における対応領域を埋め込むことにより、合成顔動画像を生成することに用いられる。

[0420] これにより、合成顔動画像を生成するための基準画像及び幾何学属性を提供することが可能になる場合がある。そして、合成顔動画像の生成において、各フレームに対応する幾何学属性によって基準画像の顔に動きを与えつつ、元の基準画像によって基準画像の背景の歪みを抑制することが可能になる。

場合がある。よって、画質の劣化の抑制に貢献することが可能になる場合がある。

- [0421] あるいは、符号化装置 100 は、入力端子と、エントロピー符号化器と、出力端子とを備えてもよい。そして、回路 151 によって行われる動作が、エントロピー符号化器によって行われてもよい。また、入力端子には、エントロピー符号化器の動作に用いられるデータが入力されてもよい。そして、出力端子から、エントロピー符号化器の動作によって得られたデータが出力されてもよい。
- [0422] 図 34 は、ビットストリーム生成装置 300 の実装例を示すブロック図である。ビットストリーム生成装置 300 は、回路 351 及びメモリ 352 を備える。例えば、ビットストリーム生成装置 300、回路 351 及びメモリ 352 は、それぞれ、符号化装置 100、回路 151 及びメモリ 152 に対応していてもよい。そして、ビットストリーム生成装置 300、回路 351 及びメモリ 352 は、それぞれ、符号化装置 100、回路 151 及びメモリ 152 と同じ役割を果たしてもよい。
- [0423] 回路 351 は、情報処理を行う電気回路であり、メモリ 352 にアクセス可能である。例えば、回路 351 は、本開示のビットストリーム生成方法を実行する専用の回路であってもよいし、本開示のビットストリーム生成方法に対応するプログラムを実行する汎用の回路であってもよい。また、回路 351 は、CPU 等のプロセッサであってもよい。さらに、回路 351 は、複数の回路の集合体であってもよい。
- [0424] メモリ 352 は、回路 351 がビットストリームを生成するための情報が記憶される専用又は汎用のメモリである。メモリ 352 は、電気回路であってもよく、回路 351 に接続されていてもよい。また、メモリ 352 は、回路 351 に含まれてもよい。また、メモリ 352 は、複数の回路の集合体であってもよい。また、メモリ 352 は、磁気ディスク又は光ディスク等であってもよいし、ストレージ又は記録媒体等と表現されてもよい。また、メモリ 352 は、不揮発性メモリであってもよいし、揮発性メモリであってもよい。

い。

[0425] 例えば、メモリ 352 には、ビットストリームを生成するためのデータが記憶されてもよいし、ビットストリームが記憶されてもよい。また、メモリ 352 には、回路 351 に生成処理を行わせるためのプログラムが記憶されていてもよい。また、メモリ 352 には、回路 351 に生成モデルが記憶されていてもよい。また、メモリ 352 には、基準画像が記憶されてもよいし、背景画像が記憶されてもよいし、背景画像候補が記憶されてもよい。

[0426] 図 35 は、ビットストリーム生成装置 300 の基本的な第 1 動作例を示すフローチャートである。ビットストリーム生成装置 300 の回路 351 は、この例の動作において、メモリ 352 を用いて、以下を行う。

[0427] 具体的には、回路 351 は、合成顔動画像を生成するための基準画像と幾何学情報と背景情報とを含むビットストリームを生成する (S501)。合成顔動画像は、顔を含む動画像であり、背景画像が合成された動画像である。基準画像は、顔を含む画像である。幾何学情報は、カメラによって得られる撮影動画像の複数のフレームの各々に対応し、被写体の幾何学属性を示す情報である。背景情報は、背景画像に関する情報である。

[0428] これにより、合成顔動画像を生成するための基準画像、幾何学属性及び背景画像を提供することが可能になる場合がある。したがって、合成顔動画像の生成において、各フレームに対応する幾何学属性によって基準画像の顔に動きを与えつつ、背景画像によって基準画像の背景の歪みを抑制することが可能になる場合がある。よって、画質の劣化の抑制に貢献することが可能になる場合がある。

[0429] 図 36 は、ビットストリーム生成装置 300 の基本的な第 2 動作例を示すフローチャートである。ビットストリーム生成装置 300 の回路 351 は、この例の動作において、メモリ 352 を用いて、以下を行う。

[0430] 具体的には、回路 351 は、合成顔動画像を生成するための基準画像と幾何学情報とを含むビットストリームを生成する (S511)。基準画像は、顔を含む画像である。幾何学情報は、カメラによって得られる撮影動画像の

複数のフレームの各々に対応し、被写体の幾何学属性を示す情報である。

[0431] 合成顔動画像の生成において、基準画像及び幾何学情報は、基準画像及び幾何学情報を生成モデルに入力して、中間顔動画像を生成モデルから取得することに用いられる。中間顔動画像は、顔を含む動画像である。基準画像は、中間顔動画像における背景領域に基準画像における対応領域を埋め込むことにより、合成顔動画像を生成することに用いられる。

[0432] これにより、合成顔動画像を生成するための基準画像及び幾何学属性を提供することが可能になる場合がある。そして、合成顔動画像の生成において、各フレームに対応する幾何学属性によって基準画像の顔に動きを与えつつ、元の基準画像によって基準画像の背景の歪みを抑制することが可能になる場合がある。よって、画質の劣化の抑制に貢献することが可能になる場合がある。

[0433] 図37は、復号装置200の実装例を示すブロック図である。復号装置200は、回路251及びメモリ252を備える。例えば、上述された復号装置200の複数の構成要素は、回路251及びメモリ252によって実装される。

[0434] 回路251は、情報処理を行う電気回路であり、メモリ252にアクセス可能である。例えば、回路251は、本開示の復号方法を実行する専用の回路であってもよいし、本開示の復号方法に対応するプログラムを実行する汎用の回路であってもよい。また、回路251は、CPU等のプロセッサであってもよい。さらに、回路251は、複数の回路の集合体であってもよい。

[0435] メモリ252は、回路251が画像を復号するための情報が記憶される専用又は汎用のメモリである。メモリ252は、電気回路であってもよく、回路251に接続されていてもよい。また、メモリ252は、回路251に含まれてもよい。また、メモリ252は、複数の回路の集合体であってもよい。また、メモリ252は、磁気ディスク又は光ディスク等であってもよいし、ストレージ又は記録媒体等と表現されてもよい。また、メモリ252は、不揮発性メモリであってもよいし、揮発性メモリであってもよい。

- [0436] 例えば、メモリ 252 には、ビットストリーム等の復号対象データが記憶されてもよいし、画像等の復号済みデータが記憶されてもよい。また、メモリ 252 には、回路 251 に画像処理を行わせるためのプログラムが記憶されていてもよい。また、メモリ 252 には、回路 251 に生成モデルが記憶されていてもよい。また、メモリ 252 には、基準画像が記憶されてもよいし、背景画像が記憶されてもよいし、背景画像候補が記憶されてもよい。
- [0437] 図 38 は、復号装置 200 の基本的な第 1 動作例を示すフローチャートである。復号装置 200 の回路 251 は、この例の動作において、メモリ 252 を用いて、以下を行う。
- [0438] 具体的には、回路 251 は、基準画像と幾何学情報と背景情報とを 1 つ以上のストリームから復号する (S401)。基準画像は、顔を含む画像である。幾何学情報は、カメラによって得られる撮影動画像の複数のフレームの各々に対応し、被写体の幾何学属性を示す情報である。背景情報は、背景画像に関する情報である。
- [0439] そして、回路 251 は、基準画像、幾何学情報及び背景情報から、生成モデルを用いて、合成顔動画像を生成する (S402)。合成顔動画像は、顔を含む動画像であり、背景画像が合成された動画像である。
- [0440] これにより、合成顔動画像の生成に、基準画像、幾何学属性及び背景画像を適用することが可能になる場合がある。したがって、各フレームに対応する幾何学属性によって基準画像の顔に動きを与えつつ、背景画像によって基準画像の背景の歪みを抑制することが可能になる場合がある。よって、合成顔動画像の生成において、画質の劣化を抑制することが可能になる場合がある。
- [0441] 例えば、回路 251 は、生成モデルに基準画像、幾何学情報及び背景画像を入力して、合成顔動画像を生成モデルから取得してもよい。
- [0442] これにより、生成モデルから容易に合成顔動画像を取得することが可能になる場合がある。そして、生成モデルにおいて、合成顔動画像の生成に、基準画像、幾何学属性及び背景画像を適用することが可能になる場合がある。

したがって、各フレームに対応する幾何学属性によって基準画像の顔に動きを与えつつ、背景画像によって基準画像の背景の歪みを抑制することが可能になる場合がある。

[0443] また、例えば、回路251は、生成モデルに基準画像及び幾何学情報を入力して、中間顔動画像を生成モデルから取得してもよい。中間顔動画像は、顔を含む動画像であり、背景画像が合成されていない動画像である。そして、回路251は、中間顔動画像における背景領域に背景画像における対応領域を埋め込むことにより、合成顔動画像を生成してもよい。

[0444] これにより、各フレームに対応する幾何学属性によって基準画像の顔に動きが与えられた中間顔動画像を生成モデルから取得することが可能になる場合がある。そして、中間顔動画像に背景画像を適用することが可能になる場合がある。したがって、各フレームに対応する幾何学属性によって基準画像の顔に動きを与えつつ、背景画像によって基準画像の背景の歪みを抑制することが可能になる場合がある。

[0445] また、例えば、回路251は、中間顔動画像に対してセグメンテーション処理を行って、中間顔動画像における前景領域及び背景領域を示す中間顔動画像セグメンテーション情報を取得してもよい。そして、回路251は、中間顔動画像セグメンテーション情報を用いて、中間顔動画像における背景領域を特定してもよい。

[0446] これにより、中間顔動画像に対するセグメンテーション処理の結果として得られる中間顔動画像セグメンテーション情報に従って、中間顔動画像における背景領域を適切に特定することが可能になる場合がある。したがって、中間顔動画像における背景領域に背景画像における対応領域を適切に適用することが可能になる場合がある。

[0447] また、例えば、回路251は、撮影動画像における前景領域及び背景領域を示す撮影動画像セグメンテーション情報を1つ以上のストリームから復号してもよい。そして、回路251は、撮影動画像セグメンテーション情報を用いて、中間顔動画像における背景領域を特定してもよい。

- [0448] これにより、1つ以上のストリームから得られる撮影動画像セグメンテーション情報に従って、中間顔動画像における背景領域を適切に特定することが可能になる場合がある。したがって、中間顔動画像における背景領域に背景画像における対応領域を適切に適用することが可能になる場合がある。
- [0449] また、例えば、基準画像における背景領域には、規定の背景カラーコードが埋め込まれていてもよい。これにより、規定の背景カラーコードに従って、基準画像における背景領域を効率的に特定することが可能になる場合がある。また、基準画像の顔に動きが与えられても、基準画像の背景に生じる歪みを抑制することが可能になる場合がある。
- [0450] また、例えば、回路251は、中間顔動画像において規定の背景カラーコードを有する領域を中間顔動画像における背景領域として特定してもよい。
- [0451] これにより、規定の背景カラーコードに従って、中間顔動画像における背景領域を効率的に特定することが可能になる場合がある。具体的には、規定の背景カラーコードが埋め込まれた背景領域を有する基準画像の顔に動きを与えることにより得られる中間顔動画像における背景領域には規定の背景カラーコードが埋め込まれていると想定される。したがって、規定の背景カラーコードに従って、中間顔動画像における背景領域を効率的に特定することが可能になる場合がある。
- [0452] また、例えば、回路251は、規定の背景カラーコードを示す背景カラーコード情報を1つ以上のストリームから復号してもよい。これにより、1つ以上のストリームから得られる規定の背景カラーコードに従って、基準画像における背景領域を効率的に特定することが可能になる場合がある。そして、基準画像に応じて規定の背景カラーコードを変化させることが可能になる場合がある。
- [0453] また、例えば、背景カラーコード情報は、連続する複数の値を含む範囲を規定の背景カラーコードとして示してもよい。そして、規定の背景カラーコードは、背景カラーコード情報によって示される範囲内で規定されてもよい。これにより、規定の背景カラーコードを柔軟に規定することが可能になる

場合がある。そして、背景領域に規定の背景カラーコードを柔軟に適用することが可能になる場合がある。

[0454] また、例えば、規定の背景カラーコードは、基準画像における前景領域において発生頻度が閾値以下であるカラーコードで規定されてもよい。これにより、前景領域の部分が背景領域の部分と誤って特定されることを抑制することが可能になる場合がある。したがって、背景領域を適切に特定することが可能になる場合がある。

[0455] また、例えば、回路251は、基準画像における前景領域及び背景領域を示す基準画像セグメンテーション情報を1つ以上のストリームから復号してもよい。そして、回路251は、基準画像セグメンテーション情報を用いて、基準画像における背景領域に、規定の背景カラーコードを埋め込んでもよい。

[0456] これにより、1つ以上のストリームから得られる基準画像セグメンテーション情報に従って、基準画像における背景領域を効率的に特定することが可能になる場合がある。また、基準画像における背景領域に規定の背景カラーコードが埋め込まれるため、基準画像の顔に動きが与えられても、基準画像の背景に生じる歪みを抑制することが可能になる場合がある。

[0457] また、例えば、背景画像は、基準画像及び撮影動画像によらずに用意された画像であってもよい。これにより、基準画像及び撮影動画像とは別途に用意された背景画像を合成顔動画像に適用することが可能になる場合がある。したがって、背景画像における前景領域の影響などを抑制することが可能になる場合がある。

[0458] また、例えば、回路251は、背景画像の識別子を背景情報として復号してもよい。そして、回路251は、識別子を用いて、複数の背景画像候補の中から背景画像を選択してもよい。

[0459] これにより、複数の背景画像候補の中から背景画像を柔軟に選択することが可能になる場合がある。したがって、合成顔動画像の用途に応じて、合成顔動画像に適切な背景画像を適用することが可能になる場合がある。

- [0460] また、例えば、背景画像は、撮影動画像に含まれる画像であってもよいし、撮影動画像に含まれる複数の画像が合成された画像であってもよい。これにより、撮影動画像から得られる背景画像を合成顔動画像に適用することが可能になる場合がある。したがって、撮影状況に対応する背景画像を合成顔動画像に適用することが可能になる場合がある。
- [0461] また、例えば、回路251は、基準画像を背景情報として復号してもよい。そして、背景画像に、基準画像が適用されてもよい。これにより、基準画像を背景画像として用いることが可能になる場合がある。そして、各フレームに対応する幾何学属性によって基準画像の顔に動きを与えつつ、元の基準画像が背景画像として用いられることで基準画像の背景の歪みを抑制することが可能になる場合がある。
- [0462] また、例えば、回路251は、背景画像に前景領域が含まれる場合、背景画像における前景領域の周辺領域、又は、過去の合成顔動画像における背景領域を用いて、背景画像における背景領域の欠落を補間してもよい。これにより、背景画像に前景領域が含まれていても、背景領域の欠落を適切に補間することが可能になる場合がある。したがって、合成顔動画像における背景領域の欠落を抑制することが可能になる場合がある。
- [0463] また、例えば、回路251は、背景画像に対してセグメンテーション処理を行って、背景画像における前景領域及び背景領域を示す背景画像セグメンテーション情報を取得してもよい。そして、回路251は、背景画像セグメンテーション情報を用いて、背景画像における前景領域及び背景領域を特定してもよい。
- [0464] これにより、背景画像に対するセグメンテーション処理の結果として得られる背景画像セグメンテーション情報に従って、背景画像における前景領域及び背景領域を適切に特定することが可能になる場合がある。したがって、背景画像における背景領域の欠落を適切に補間することが可能になる場合がある。
- [0465] また、例えば、背景画像における前景領域には、規定の前景カラーコード

が埋め込まれていてもよい。これにより、規定の前景カラーコードに従って、背景画像における前景領域を効率的に特定することが可能になる場合がある。また、合成顔動画画像における背景領域に顔等の前景が映り込むことを抑制することが可能になる場合がある。

[0466] また、例えば、回路251は、規定の前景カラーコードを示す前景カラーコード情報を1つ以上のストリームから復号してもよい。そして、背景画像において規定の前景カラーコードを有する領域を背景画像における前景領域として特定してもよい。これにより、1つ以上のストリームから得られる規定の前景カラーコードに従って、背景画像における前景領域を効率的に特定することが可能になる場合がある。そして、背景画像に応じて規定の前景カラーコードを変化させることが可能になる場合がある。

[0467] また、例えば、前景カラーコード情報は、連続する複数の値を含む範囲を規定の前景カラーコードとして示してもよい。そして、規定の前景カラーコードは、前景カラーコード情報によって示される範囲内で規定されてもよい。これにより、規定の前景カラーコードを柔軟に規定することが可能になる場合がある。そして、前景領域に規定の前景カラーコードを柔軟に適用することが可能になる場合がある。

[0468] また、例えば、規定の前景カラーコードは、背景画像における背景領域において発生頻度が閾値以下であるカラーコードで規定されてもよい。これにより、背景領域の部分が前景領域の部分と誤って特定されることを抑制することが可能になる場合がある。したがって、前景領域を適切に特定することが可能になる場合がある。

[0469] また、例えば、1つ以上のストリームにおいて、背景情報が復号されるストリームは、基準画像が復号されるストリームと同じであってもよいし、幾何学情報が復号されるストリームと同じであってもよい。これにより、別途のストリームではなく、基準画像のストリーム又は幾何学情報のストリームと同じストリームから背景情報を復号することが可能になる場合がある。したがって、基準画像又は幾何学情報と一緒に背景情報を効率的に復号するこ

とが可能になる場合がある。

[0470] また、例えば、1つ以上のストリームにおいて、背景情報が復号されるストリームは、基準画像が復号されるストリームとも幾何学情報が復号されるストリームとも異なってもよい。これにより、同じストリームではなく、基準画像のストリーム及び幾何学情報のストリームとは別途のストリームから背景情報を復号することが可能になる場合がある。したがって、基準画像又は幾何学情報とは別に背景情報を任意のタイミングで復号することが可能になる場合がある。

[0471] また、例えば、背景画像は、複数のピクチャを含むシーケンスの先頭のピクチャとして復号されてもよいし、GOPの先頭のピクチャとして復号されてもよい。これにより、早期に背景画像を取得することが可能になる場合がある。したがって、早期に背景画像を合成顔動画像に適用することが可能になる場合がある。

[0472] また、例えば、背景画像は、1つ以上のストリームにおけるアクセスユニットからピクチャとして復号されてもよい。これにより、背景画像をアクセスユニット内のピクチャとして処理することが可能になる場合がある。つまり、通常のピクチャと同様に背景画像を処理することが可能になる場合がある。

[0473] また、例えば、背景画像が復号されるアクセスユニットは、基準画像が復号されるアクセスユニットと同じであってもよい。これにより、別途のアクセスユニットではなく、基準画像のアクセスユニットと同じアクセスユニットから背景画像を復号することが可能になる場合がある。したがって、基準画像と一緒に背景画像を効率的に復号することが可能になる場合がある。

[0474] また、例えば、背景画像が復号されるアクセスユニットは、基準画像が復号されるアクセスユニットとは異なってもよい。これにより、同じアクセスユニットではなく、基準画像のアクセスユニットとは別途のアクセスユニットから背景画像を復号することが可能になる場合がある。したがって、基準画像とは別に背景画像を任意のタイミングで復号することが可能になる

場合がある。

[0475] また、例えば、背景画像が含まれるアクセスユニットに対応付けられた S E I から、そのアクセスユニットに背景画像が存在することを示す信号が復号されてもよい。これにより、アクセスユニットにおける S E I から得られる信号に従って、アクセスユニットにおける背景画像の存在を認識することが可能になる場合がある。したがって、背景画像の適切な伝達が可能になる場合がある。

[0476] また、例えば、背景画像は、1つ以上のストリームにおけるアクセスユニットからピクチャとして復号されてもよい。また、背景画像が含まれるアクセスユニットに対応付けられた S E I から、アクセスユニットに背景画像が存在することを示す信号が復号されていてもよい。

[0477] これにより、背景画像をアクセスユニット内のピクチャとして処理することが可能になる場合がある。つまり、通常のピクチャと同様に背景画像を処理することが可能になる場合がある。また、これにより、アクセスユニットにおける S E I から得られる信号に従って、アクセスユニットにおける背景画像の存在を認識することが可能になる場合がある。したがって、背景画像の適切な伝達が可能になる場合がある。

[0478] また、例えば、背景画像は、イントラピクチャとして復号されてもよい。これにより、背景画像をイントラピクチャとして処理することが可能になる場合がある。つまり、他のピクチャに依存せずに、背景画像を処理することが可能になる場合がある。

[0479] また、例えば、背景画像は、合成顔動画画像の複数のフレームに共通して適用されてもよい。これにより、合成顔動画画像の全体の符号量を抑制することが可能になる場合がある。また、背景画像を復号するための処理量を抑制することが可能になる場合がある。

[0480] また、例えば、回路 251 は、背景カラーコード情報を 1つ以上のストリームにおける S E I から復号してもよい。これにより、S E I から得られる規定の背景カラーコードに従って、基準画像における背景領域を効率的に特

定することが可能になる場合がある。そして、基準画像に応じて規定の背景カラーコードを変化させることが可能になる場合がある。

[0481] また、例えば、回路251は、前景カラーコード情報を1つ以上のストリームにおけるSEIから復号してもよい。これにより、SEIから得られる規定の前景カラーコードに従って、背景画像における前景領域を効率的に特定することが可能になる場合がある。そして、背景画像に応じて規定の前景カラーコードを変化させることが可能になる場合がある。

[0482] また、例えば、回路251は、規定の背景カラーコードを示す背景カラーコード情報と、規定の前景カラーコードを示す前景カラーコード情報とのうち、少なくとも一方を1つ以上のストリームにおけるSEIから復号してもよい。これにより、SEIから得られる規定の背景カラーコードに従って、背景領域を効率的に特定することが可能になる場合がある。

[0483] また、例えば、回路251は、撮影動画像セグメンテーション情報を1つ以上のストリームにおけるSEIから復号してもよい。これにより、SEIから得られる撮影動画像セグメンテーション情報に従って、中間顔動画像における背景領域を適切に特定することが可能になる場合がある。したがって、中間顔動画像における背景領域に背景画像における対応領域を適切に適用することが可能になる場合がある。

[0484] また、例えば、回路251は、基準画像セグメンテーション情報を1つ以上のストリームにおけるSEIから復号してもよい。これにより、SEIから得られる基準画像セグメンテーション情報に従って、基準画像における背景領域を効率的に特定することが可能になる場合がある。そして、基準画像における背景領域に規定の背景カラーコードを適切に埋め込むことが可能になる場合がある。

[0485] 図39は、復号装置200の基本的な第2動作例を示すフローチャートである。復号装置200の回路251は、この例の動作において、メモリ252を用いて、以下を行う。

[0486] また、例えば、回路251は、撮影動画像セグメンテーション情報と、基

準画像セグメンテーション情報とのうち、少なくとも一方を1つ以上のストリームにおけるSEIから復号してもよい。撮影動画像セグメンテーション情報は、撮影動画像における前景領域及び背景領域を示す。基準画像セグメンテーション情報は、基準画像における前景領域及び背景領域を示す。これにより、SEIから得られるセグメンテーション情報に従って、背景領域を効率的に特定することが可能になる場合がある。

[0487] 具体的には、回路251は、基準画像と幾何学情報とを1つ以上のストリームから復号する(S411)。基準画像は、顔を含む画像である。幾何学情報は、カメラによって得られる撮影動画像の複数のフレームの各々に対応し、被写体の幾何学属性を示す情報である。

[0488] 次に、回路251は、基準画像及び幾何学情報を生成モデルに入力して、中間顔動画像を生成モデルから取得する(S412)。中間顔動画像は、顔を含む動画像である。そして、回路251は、中間顔動画像における背景領域に基準画像における対応領域を埋め込むことにより、合成顔動画像を生成する(S413)。

[0489] これにより、各フレームに対応する幾何学属性によって基準画像の顔に動きが与えられた中間顔動画像を生成モデルから取得することが可能になる場合がある。そして、中間顔動画像の背景領域に元の基準画像の対応領域を適用することが可能になる場合がある。したがって、各フレームに対応する幾何学属性によって基準画像の顔に動きを与えつつ、元の基準画像によって基準画像の背景の歪みを抑制することが可能になる場合がある。よって、合成顔動画像の生成において、画質の劣化を抑制することが可能になる場合がある。

[0490] あるいは、復号装置200は、入力端子と、エントロピー復号器と、出力端子とを備えてもよい。そして、回路251によって行われる動作が、エントロピー復号器によって行われてもよい。また、入力端子には、エントロピー復号器の動作に用いられるデータが入力されてもよい。そして、出力端子から、エントロピー復号器の動作によって得られたデータが出力されてもよい。

い。

[0491] また、例えば、1つ以上のビットストリームを記憶するコンピュータ読み取り可能な非一時的な記録媒体が用いられてもよい。1つ以上のビットストリームは、動画像の表示に用いられる少なくとも1つの基準画像と、動画像の複数の画像のそれぞれに対応する情報として人物の顔を含む領域内の幾何学属性を示す幾何学情報とを含んでいてもよい。そして、1つ以上のビットストリームは、(i) 少なくとも1つの基準画像を復号し、(ii) 幾何学情報を復号する処理を復号装置200に実行させてもよい。

[0492] これにより、上述の復号装置及び復号方法に対応する1つ以上のビットストリームを記憶する記録媒体を実現することが可能になる場合がある。したがって、上述の復号装置200等と同様の効果を記録媒体によって得ることが可能になる場合がある。

[0493] [その他の例]

上述された各例における符号化装置100及び復号装置200は、それぞれ、画像符号化装置及び画像復号装置として利用されてもよいし、動画像符号化装置及び動画像復号装置として利用されてもよい。また、符号化装置100に含まれる複数の構成要素と、復号装置200に含まれる複数の構成要素とは、互いに対応する動作を行ってもよい。

[0494] また、符号化するという表現は、格納する、含める、書き込む、記述する、信号化する、送り出す、通知する、又は、保存する等の表現に置き換えられてもよく、これらの表現が相互に置き換えられてもよい。例えば、情報を符号化することは、ビットストリームに情報を含めることであってもよい。また、情報をビットストリームに符号化することは、情報を符号化して、符号化された情報を含むビットストリームを生成することを意味してもよい。

[0495] また、復号するという表現は、読み出す、読み解く、読み取る、読み込む、導出する、取得する、受け取る、抽出する、又は、復元する等の表現に置き換えられてもよく、これらの表現が相互に置き換えられてもよい。例えば、情報を復号することは、ビットストリームから情報を取得することであっ

てもよい。また、ビットストリームから情報を復号することは、ビットストリームを復号して、ビットストリームに含まれる情報を取得することを意味してもよい。

[0496] また、例えば、ビットストリームに含まれる符号化情報及び圧縮情報等が、単に、情報と表現される場合がある。

[0497] また、上述された各例の少なくとも一部が、符号化方法として利用されてもよいし、復号方法として利用されてもよいし、エントロピー符号化方法として利用されてもよいし、エントロピー復号方法として利用されてもよいし、その他の方法として利用されてもよい。

[0498] また、各構成要素は、専用のハードウェアで構成されるか、各構成要素に適したソフトウェアプログラムを実行することによって実現されてもよい。各構成要素は、CPU又はプロセッサなどのプログラム実行部が、ハードディスク又は半導体メモリなどの記録媒体に記録されたソフトウェアプログラムを読み出して実行することによって実現されてもよい。

[0499] 具体的には、符号化装置100及び復号装置200のそれぞれは、処理回路(Processing Circuitry)と、当該処理回路に電気的に接続された、当該処理回路からアクセス可能な記憶装置(Storage)とを備えていてもよい。例えば、処理回路は回路151又は251に対応し、記憶装置はメモリ152又は252に対応する。

[0500] 処理回路は、専用のハードウェア及びプログラム実行部の少なくとも一方を含み、記憶装置を用いて処理を実行する。また、記憶装置は、処理回路がプログラム実行部を含む場合には、当該プログラム実行部により実行されるソフトウェアプログラムを記憶する。

[0501] 上述したソフトウェアプログラムの一例は、ビットストリームである。ビットストリームは、符号化された画像と、画像を復号する復号処理を行うためのシンタックスを含む。ビットストリームは、シンタックスに基づく処理を復号装置200に実行させることで、画像を復号装置200に復号させる。また、例えば、上述された符号化装置100又は復号装置200などを実

現するためのソフトウェアは、次のようなプログラムである。

- [0502] 例えば、このプログラムは、コンピュータに、顔を含む動画画であり、背景画像が合成された動画画である合成顔動画画を生成するための (i) 基準画像と (i i) 幾何学情報と (i i i) 背景情報とであって、(i) 前記顔を含む画像である基準画像と、(i i) カメラによって得られる撮影動画画の複数のフレームの各々に対応し、被写体の幾何学属性を示す幾何学情報と、(i i i) 前記背景画像に関する背景情報とを1つ以上のストリームに符号化する、符号化方法を実行させてもよい。
- [0503] また、例えば、このプログラムは、コンピュータに、(i) 顔を含む画像である基準画像と、(i i) カメラによって得られる撮影動画画の複数のフレームの各々に対応し、被写体の幾何学属性を示す幾何学情報と、(i i i) 背景画像に関する背景情報とを1つ以上のストリームから復号し、前記基準画像、前記幾何学情報及び前記背景情報から、生成モデルを用いて、前記顔を含む動画画であり、前記背景画像が合成された動画画である合成顔動画画を生成する、復号方法を実行させてもよい。
- [0504] また、上述された各構成要素は、回路であってもよい。これらの回路は、全体として1つの回路を構成してもよいし、それぞれ別々の回路であってもよい。また、各構成要素は、汎用的なプロセッサで実現されてもよいし、専用のプロセッサで実現されてもよい。
- [0505] また、特定の構成要素が実行する処理を別の構成要素が実行してもよい。また、処理を実行する順番が変更されてもよいし、複数の処理が並行して実行されてもよい。また、本開示の複数の例のうちの任意の2つ以上が、適宜、組み合わせられて、実施されてもよい。また、符号化復号装置が、符号化装置100及び復号装置200を備えていてもよい。
- [0506] また、本開示における複数の構成要素の全てが実装されなくてもよく、本開示における複数の構成要素の一部のみが実装されてもよい。同様に、本開示における複数の処理の全てが実行されなくてもよく、本開示における複数の処理の一部のみが実行されてもよい。

[0507] また、説明に用いられた第1及び第2等の序数は、適宜、付け替えられてもよい。また、構成要素などに対して、序数が新たに与えられてもよいし、取り除かれてもよい。また、これらの序数は、要素を識別するため、要素に付けられる場合があり、意味のある順序に対応しない場合がある。

[0508] また、例えば、第1要素、第2要素及び第3要素の少なくとも1つ（又は1つ以上）という表現は、第1要素、第2要素、第3要素、又は、これらの任意の組み合わせに対応する。

[0509] 以上、符号化装置100及び復号装置200の態様について、複数の例に基づいて説明したが、符号化装置100及び復号装置200の態様は、これらの例に限定されるものではない。本開示の趣旨を逸脱しない限り、当業者が思いつく各種変形を各例に施したものや、異なる例における構成要素を組み合わせる形態も、符号化装置100及び復号装置200の態様の範囲内に含まれてもよい。

[0510] ここで開示された1以上の態様を本開示における他の態様の少なくとも一部と組み合わせる実施してもよい。また、ここで開示された1以上の態様のフローチャートに記載の一部の処理、装置の一部の構成、シンタックスの一部などを他の態様と組み合わせる実施してもよい。

[0511] [実施及び応用]

以上の各実施の形態において、機能的又は作用的なブロックの各々は、通常、MPU (micro processing unit) 及びメモリ等によって実現可能である。また、機能ブロックの各々による処理は、ROM等の記録媒体に記録されたソフトウェア（プログラム）を読み出して実行するプロセッサなどのプログラム実行部として実現されてもよい。当該ソフトウェアは、配布されてもよい。当該ソフトウェアは、半導体メモリなどの様々な記録媒体に記録されてもよい。なお、各機能ブロックをハードウェア（専用回路）によって実現することも可能である。

[0512] 各実施の形態において説明した処理は、単一の装置（システム）を用いて集中処理することによって実現してもよく、又は、複数の装置を用いて分散

処理することによって実現してもよい。また、上記プログラムを実行するプロセッサは、単数であってもよく、複数であってもよい。すなわち、集中処理を行ってもよく、又は分散処理を行ってもよい。

[0513] 本開示の態様は、以上の実施例に限定されることなく、種々の変更が可能であり、それらも本開示の態様の範囲内に包含される。

[0514] さらにここで、上記各実施の形態で示した動画像符号化方法（画像符号化方法）又は動画像復号化方法（画像復号方法）の応用例、及び、その応用例を実施する種々のシステムを説明する。このようなシステムは、画像符号化方法を用いた画像符号化装置、画像復号方法を用いた画像復号装置、又は、両方を備える画像符号化復号装置を有することを特徴としてもよい。このようなシステムの他の構成について、場合に応じて適切に変更することができる。

[0515] [使用例]

図40は、コンテンツ配信サービスを実現する適切なコンテンツ供給システムex100の全体構成を示す図である。通信サービスの提供エリアを所望の大きさに分割し、各セル内にそれぞれ、図示された例における固定無線局である基地局ex106、ex107、ex108、ex109、ex110が設置されている。

[0516] このコンテンツ供給システムex100では、インターネットex101に、インターネットサービスプロバイダex102又は通信網ex104、及び基地局ex106～ex110を介して、コンピュータex111、ゲーム機ex112、カメラex113、家電ex114、及びスマートフォンex115などの各機器が接続される。当該コンテンツ供給システムex100は、上記のいずれかの装置を組合せて接続するようにしてもよい。種々の実施において、基地局ex106～ex110を介さずに、各機器が電話網又は近距離無線等を介して直接的又は間接的に相互に接続されていてもよい。さらに、ストリーミングサーバex103は、インターネットex101等を介して、コンピュータex111、ゲーム機ex112、カメラex

× 1 1 3、家電 e × 1 1 4、及びスマートフォン e × 1 1 5 などの各機器と接続されてもよい。また、ストリーミングサーバ e × 1 0 3 は、衛星 e × 1 1 6 を介して、飛行機 e × 1 1 7 内のホットスポット内の端末等と接続されてもよい。

[0517] なお、基地局 e × 1 0 6 ~ e × 1 1 0 の代わりに、無線アクセスポイント又はホットスポット等が用いられてもよい。また、ストリーミングサーバ e × 1 0 3 は、インターネット e × 1 0 1 又はインターネットサービスプロバイダ e × 1 0 2 を介さずに直接通信網 e × 1 0 4 と接続されてもよいし、衛星 e × 1 1 6 を介さず直接飛行機 e × 1 1 7 と接続されてもよい。

[0518] カメラ e × 1 1 3 はデジタルカメラ等の静止画撮影、及び動画撮影が可能な機器である。また、スマートフォン e × 1 1 5 は、2 G、3 G、3. 9 G、4 G、そして今後は 5 G と呼ばれる移動通信システムの方式に対応したスマートフォン機、携帯電話機、又は PHS (P e r s o n a l H a n d y p h o n e S y s t e m) 等である。

[0519] 家電 e × 1 1 4 は、冷蔵庫、又は家庭用燃料電池コージェネレーションシステムに含まれる機器等である。

[0520] コンテンツ供給システム e × 1 0 0 では、撮影機能を有する端末が基地局 e × 1 0 6 等を通じてストリーミングサーバ e × 1 0 3 に接続されることで、ライブ配信等が可能になる。ライブ配信では、端末（コンピュータ e × 1 1 1、ゲーム機 e × 1 1 2、カメラ e × 1 1 3、家電 e × 1 1 4、スマートフォン e × 1 1 5、及び飛行機 e × 1 1 7 内の端末等）は、ユーザが当該端末を用いて撮影した静止画又は動画コンテンツに対して上記各実施の形態で説明した符号化処理を行ってもよく、符号化により得られた映像データと、映像に対応する音を符号化した音データと多重化してもよく、得られたデータをストリーミングサーバ e × 1 0 3 に送信してもよい。即ち、各端末は、本開示の一態様に係る画像符号化装置として機能する。

[0521] 一方、ストリーミングサーバ e × 1 0 3 は要求のあったクライアントに対して送信されたコンテンツデータをストリーム配信する。クライアントは、

上記符号化処理されたデータを復号化することが可能な、コンピュータ e x 1 1 1、ゲーム機 e x 1 1 2、カメラ e x 1 1 3、家電 e x 1 1 4、スマートフォン e x 1 1 5、又は飛行機 e x 1 1 7 内の端末等である。配信されたデータを受信した各機器は、受信したデータを復号化処理して再生する。即ち、各機器は、本開示の一態様に係る画像復号装置として機能してもよい。

[0522] [分散処理]

また、ストリーミングサーバ e x 1 0 3 は複数のサーバ又は複数のコンピュータであって、データを分散して処理したり記録したり配信するものであってもよい。例えば、ストリーミングサーバ e x 1 0 3 は、CDN (C o n t e n t s D e l i v e r y N e t w o r k) により実現され、世界中に分散された多数のエッジサーバとエッジサーバ間をつなぐネットワークによりコンテンツ配信が実現されていてよい。CDNでは、クライアントに応じて物理的に近いエッジサーバが動的に割り当てられる。そして、当該エッジサーバにコンテンツがキャッシュ及び配信されることで遅延を減らすことができる。また、いくつかのタイプのエラーが発生した場合又はトラフィックの増加などにより通信状態が変わる場合に複数のエッジサーバで処理を分散したり、他のエッジサーバに配信主体を切り替えたり、障害が生じたネットワークの部分を迂回して配信を続けることができるので、高速かつ安定した配信が実現できる。

[0523] また、配信自体の分散処理にとどまらず、撮影したデータの符号化処理を各端末で行ってもよいし、サーバ側で行ってもよいし、互いに分担して行ってもよい。一例として、一般に符号化処理では、処理ループが2度行われる。1度目のループでフレーム又はシーン単位での画像の複雑さ、又は、符号量が検出される。また、2度目のループでは画質を維持して符号化効率を向上させる処理が行われる。例えば、端末が1度目の符号化処理を行い、コンテンツを受け取ったサーバ側が2度目の符号化処理を行うことで、各端末での処理負荷を減らしつつもコンテンツの質と効率を向上させることができる。この場合、ほぼリアルタイムで受信して復号する要求があれば、端末が行

った一度目の符号化済みデータを他の端末で受信して再生することもできるので、より柔軟なリアルタイム配信も可能になる。

[0524] 他の例として、カメラ $e \times 113$ 等は、画像から特徴量抽出を行い、特徴量に関するデータをメタデータとして圧縮してサーバに送信する。サーバは、例えば特徴量からオブジェクトの重要性を判断して量子化精度を切り替えるなど、画像の意味（又は内容の重要性）に応じた圧縮を行う。特徴量データはサーバでの再度の圧縮時の動きベクトル予測の精度及び効率向上に特に有効である。また、端末で V L C（可変長符号化）などの簡易的な符号化を行い、サーバで C A B A C（コンテキスト適応型二値算術符号化方式）など処理負荷の大きな符号化を行ってもよい。

[0525] さらに他の例として、スタジアム、ショッピングモール、又は工場などにおいては、複数の端末によりほぼ同一のシーンが撮影された複数の映像データが存在する場合がある。この場合には、撮影を行った複数の端末と、必要に応じて撮影をしていない他の端末及びサーバを用いて、例えば G O P（G r o u p o f P i c t u r e）単位、ピクチャ単位、又はピクチャを分割したタイル単位などで符号化処理をそれぞれ割り当てて分散処理を行う。これにより、遅延を減らし、よりリアルタイム性を実現できる。

[0526] 複数の映像データはほぼ同一シーンであるため、各端末で撮影された映像データを互いに参照し合えるように、サーバで管理及び／又は指示をしてもよい。また、各端末からの符号化済みデータを、サーバが受信し複数のデータ間で参照関係を変更、又はピクチャ自体を補正或いは差し替えて符号化しなおしてもよい。これにより、一つ一つのデータの質と効率を高めたストリームを生成できる。

[0527] さらに、サーバは、映像データの符号化方式を変更するトランスコードを行ったうえで映像データを配信してもよい。例えば、サーバは、M P E G 系の符号化方式を V P 系（例えば V P 9）に変換してもよいし、H. 2 6 4 を H. 2 6 5 に変換してもよい。

[0528] このように、符号化処理は、端末、又は 1 以上のサーバにより行うことが

可能である。よって、以下では、処理を行う主体として「サーバ」又は「端末」等の記載を用いるが、サーバで行われる処理の一部又は全てが端末で行われてもよいし、端末で行われる処理の一部又は全てがサーバで行われてもよい。また、これらに関しては、復号処理についても同様である。

[0529] [3D、マルチアングル]

互いにほぼ同期した複数のカメラ $e \times 113$ 及び／又はスマートフォン $e \times 115$ などの端末により撮影された異なるシーン、又は、同一シーンを異なるアングルから撮影した画像或いは映像を統合して利用することが増えてきている。各端末で撮影した映像は、別途取得した端末間の相対的な位置関係、又は、映像に含まれる特徴点が一致する領域などに基づいて統合される。

[0530] サーバは、2次元の動画像を符号化するだけでなく、動画像のシーン解析などに基づいて自動的に、又は、ユーザが指定した時刻において、静止画を符号化し、受信端末に送信してもよい。サーバは、さらに、撮影端末間の相対的な位置関係を取得できる場合には、2次元の動画像だけでなく、同一シーンが異なるアングルから撮影された映像に基づき、当該シーンの3次元形状を生成できる。サーバは、ポイントクラウドなどにより生成した3次元のデータを別途符号化してもよいし、3次元データを用いて人物又はオブジェクトを認識或いは追跡した結果に基づいて、受信端末に送信する映像を、複数の端末で撮影した映像から、選択、又は、再構成して生成してもよい。

[0531] このようにして、ユーザは、各撮影端末に対応する各映像を任意に選択してシーンを楽しむこともできるし、複数画像又は映像を用いて再構成された3次元データから選択視点の映像を切り出したコンテンツを楽しむこともできる。さらに、映像と共に、音も複数の相異なるアングルから収音され、サーバは、特定のアングル又は空間からの音を対応する映像と多重化して、多重化された映像と音とを送信してもよい。

[0532] また、近年ではVirtual Reality (VR) 及びAugmented Reality (AR) など、現実世界と仮想世界とを対応付け

たコンテンツも普及してきている。VRの画像の場合、サーバは、右目用及び左目用の視点画像をそれぞれ作成し、Multi-View Coding (MVC) などにより各視点映像間で参照を許容する符号化を行ってもよいし、互いに参照せずに別ストリームとして符号化してもよい。別ストリームの復号時には、ユーザの視点に応じて仮想的な3次元空間が再現されるように互いに同期させて再生するとよい。

[0533] ARの画像の場合には、サーバは、現実空間のカメラ情報に、仮想空間上の仮想物体情報を、3次元的位置又はユーザの視点の動きに基づいて重畳する。復号装置は、仮想物体情報及び3次元データを取得又は保持し、ユーザの視点の動きに応じて2次元画像を生成し、スムーズにつなげることで重畳データを作成してもよい。または、復号装置は仮想物体情報の依頼に加えてユーザの視点の動きをサーバに送信してもよい。サーバは、サーバに保持される3次元データから受信した視点の動きに合わせて重畳データを作成し、重畳データを符号化して復号装置に配信してもよい。なお、重畳データは、RGB以外に透過度を示す α 値を有し、サーバは、3次元データから作成されたオブジェクト以外の部分の α 値が0などに設定し、当該部分が透過する状態で、符号化してもよい。もしくは、サーバは、クロマキーのように所定の値のRGB値を背景に設定し、オブジェクト以外の部分は背景色にしたデータを生成してもよい。

[0534] 同様に配信されたデータの復号処理はクライアントである各端末で行っても、サーバ側で行ってもよいし、互いに分担して行ってもよい。一例として、ある端末が、一旦サーバに受信リクエストを送り、そのリクエストに応じたコンテンツを他の端末で受信し復号処理を行い、ディスプレイを有する装置に復号済みの信号が送信されてもよい。通信可能な端末自体の性能によらず処理を分散して適切なコンテンツを選択することで画質のよいデータを再生することができる。また、他の例として大きなサイズの画像データをTV等で受信しつつ、鑑賞者の個人端末にピクチャが分割されたタイルなど一部の領域が復号されて表示されてもよい。これにより、全体像を共有化しつつ

、自身の担当分野又はより詳細に確認したい領域を手元で確認することができる。

[0535] 屋内外の近距離、中距離、又は長距離の無線通信が複数使用可能な状況下で、MPEG-DASH (Dynamic Adaptive Streaming over HTTP) などの配信システム規格を利用して、シームレスにコンテンツを受信することが可能かもしれない。ユーザは、ユーザの端末、屋内外に配置されたディスプレイなどの復号装置又は表示装置を自由に選択しながらリアルタイムで切り替えてもよい。また、自身の位置情報などを用いて、復号する端末及び表示する端末を切り替えながら復号を行うことができる。これにより、ユーザが目的地へ移動している間に、表示可能なデバイスが埋め込まれた隣の建物の壁面又は地面の一部に情報をマップ及び表示することが可能になる。また、符号化データが受信端末から短時間でアクセスできるサーバにキャッシュされている、又は、コンテンツ・デリバリー・サービスにおけるエッジサーバにコピーされている、などの、ネットワーク上での符号化データへのアクセス容易性に基づいて、受信データのビットレートを切り替えることも可能である。

[0536] [Webページの最適化]

図41は、コンピュータex111等におけるwebページの表示画面例を示す図である。図42は、スマートフォンex115等におけるwebページの表示画面例を示す図である。図41及び図42に示すようにwebページが、画像コンテンツへのリンクであるリンク画像を複数含む場合があり、閲覧するデバイスによってその見え方は異なる。画面上に複数のリンク画像が見える場合には、ユーザが明示的にリンク画像を選択するまで、又は画面の中央付近にリンク画像が近付く或いはリンク画像の全体が画面内に入るまで、表示装置（復号装置）は、リンク画像として各コンテンツが有する静止画又は1ピクチャを表示してもよいし、複数の静止画又は1ピクチャ等でgifアニメのような映像を表示してもよいし、ベースレイヤのみを受信し、映像を復号及び表示してもよい。

[0537] ユーザによりリンク画像が選択された場合、表示装置は、ベースレイヤを最優先にしつつ復号を行う。なお、webページを構成するHTML (HyperText Markup Language) にスケーラブルなコンテンツであることを示す情報があれば、表示装置は、エンハンスメントレイヤまで復号してもよい。さらに、リアルタイム性を担保するために、選択される前又は通信帯域が非常に厳しい場合には、表示装置は、前方参照のピクチャ (Iピクチャ、Pピクチャ、前方参照のみのBピクチャ) のみを復号及び表示することで、先頭ピクチャの復号時刻と表示時刻との間の遅延 (コンテンツの復号開始から表示開始までの遅延) を低減できる。またさらに、表示装置は、ピクチャの参照関係を敢えて無視して、全てのBピクチャ及びPピクチャを前方参照にして粗く復号し、時間が経ち受信したピクチャが増えるにつれて正常の復号を行ってもよい。

[0538] [自動走行]

また、車の自動走行又は走行支援のため2次元又は3次元の地図情報などのような静止画又は映像データを送受信する場合、受信端末は、1以上のレイヤに属する画像データに加えて、メタ情報として天候又は工事の情報なども受信し、これらに対応付けて復号してもよい。なお、メタ情報は、レイヤに属してもよいし、単に画像データと多重化されてもよい。

[0539] この場合、受信端末を含む車、ドローン又は飛行機などが移動するため、受信端末は、当該受信端末の位置情報を送信することで、基地局ex106～ex110を切り替えながらシームレスな受信及び復号の実行を実現できる。また、受信端末は、ユーザの選択、ユーザの状況及び／又は通信帯域の状態に応じて、メタ情報をどの程度受信するか、又は地図情報をどの程度更新していくかを動的に切り替えることが可能になる。

[0540] コンテンツ供給システムex100では、ユーザが送信した符号化された情報をリアルタイムでクライアントが受信して復号し、再生することができる。

[0541] [個人コンテンツの配信]

また、コンテンツ供給システム $e \times 100$ では、映像配信業者による高画質で長時間のコンテンツのみならず、個人による低画質で短時間のコンテンツのユニキャスト、又はマルチキャスト配信が可能である。このような個人のコンテンツは今後も増加していくと考えられる。個人コンテンツをより優れたコンテンツにするために、サーバは、編集処理を行ってから符号化処理を行ってもよい。これは、例えば、以下のような構成を用いて実現できる。

[0542] 撮影時にリアルタイム又は蓄積して撮影後に、サーバは、原画データ又は符号化済みデータから撮影エラー、シーン探索、意味の解析、及びオブジェクト検出などの認識処理を行う。そして、サーバは、認識結果に基づいて手動又は自動で、ピントずれ又は手ブレなどを補正したり、明度が他のピクチャに比べて低い又は焦点が合っていないシーンなどの重要性の低いシーンを削除したり、オブジェクトのエッジを強調したり、色合いを変化させるなどの編集を行う。サーバは、編集結果に基づいて編集後のデータを符号化する。また撮影時刻が長すぎると視聴率が下がることも知られており、サーバは、撮影時間に応じて特定の時間範囲内のコンテンツになるように上記のように重要性が低いシーンのみならず動きが少ないシーンなどを、画像処理結果に基づき自動でクリップしてもよい。または、サーバは、シーンの意味解析の結果に基づいてダイジェストを生成して符号化してもよい。

[0543] 個人コンテンツには、そのままでは著作権、著作者人格権、又は肖像権等の侵害となるものが写り込んでいるケースもあり、共有する範囲が意図した範囲を超えてしまうなど個人にとって不都合な場合もある。よって、例えば、サーバは、画面の周辺部の人の顔、又は家の中などを敢えて焦点が合わない画像に変更して符号化してもよい。さらに、サーバは、符号化対象画像内に、予め登録した人物とは異なる人物の顔が映っているかどうかを認識し、映っている場合には、顔の部分にモザイクをかけるなどの処理を行ってもよい。または、符号化の前処理又は後処理として、著作権などの観点からユーザが画像を加工したい人物又は背景領域を指定してもよい。サーバは、指定された領域を別の映像に置き換える、又は焦点をぼかすなどの処理を行って

もよい。人物であれば、動画像において人物をトラッキングして、人物の顔の部分の映像を置き換えることができる。

[0544] データ量の小さい個人コンテンツの視聴はリアルタイム性の要求が強いため、帯域幅にもよるが、復号装置は、まずベースレイヤを最優先で受信して復号及び再生を行う。復号装置は、この間にエンハンスメントレイヤを受信し、再生がループされる場合など2回以上再生される場合に、エンハンスメントレイヤも含めて高画質の映像を再生してもよい。このようにスケーラブルな符号化が行われているストリームであれば、未選択時又は見始めた段階では粗い動画だが、徐々にストリームがスマートになり画像がよくなるような体験を提供することができる。スケーラブル符号化以外にも、1回目に再生される粗いストリームと、1回目の動画を参照して符号化される2回目のストリームとが1つのストリームとして構成されていても同様の体験を提供できる。

[0545] [その他の実施応用例]

また、これらの符号化又は復号処理は、一般的に各端末が有するLSIex500において処理される。LSI (large scale integration circuitry) ex500 (図40参照) は、ワンチップであっても複数チップからなる構成であってもよい。なお、動画像符号化又は復号用のソフトウェアをコンピュータex111等で読み取り可能な何らかの記録メディア (CD-ROM、フレキシブルディスク、又はハードディスクなど) に組み込み、そのソフトウェアを用いて符号化又は復号処理を行ってもよい。さらに、スマートフォンex115がカメラ付きである場合には、そのカメラで取得した動画データを送信してもよい。このときの動画データはスマートフォンex115が有するLSIex500で符号化処理されたデータである。

[0546] なお、LSIex500は、アプリケーションソフトをダウンロードしてアクティベートする構成であってもよい。この場合、端末は、まず、当該端末がコンテンツの符号化方式に対応しているか、又は、特定サービスの実行

能力を有するかを判定する。端末がコンテンツの符号化方式に対応していない場合、又は、特定サービスの実行能力を有さない場合、端末は、コーデック又はアプリケーションソフトをダウンロードし、その後、コンテンツ取得及び再生する。

[0547] また、インターネット $e \times 101$ を介したコンテンツ供給システム $e \times 100$ に限らず、デジタル放送用システムにも上記各実施の形態の少なくとも動画像符号化装置（画像符号化装置）又は動画像復号化装置（画像復号装置）のいずれかを組み込むことができる。衛星などを利用して放送用の電波に映像と音が多重化された多重化データを載せて送受信するため、コンテンツ供給システム $e \times 100$ のユニキャストがし易い構成に対してマルチキャスト向きであるという違いがあるが符号化処理及び復号処理に関しては同様の応用が可能である。

[0548] [ハードウェア構成]

図43は、図40に示されたスマートフォン $e \times 115$ のさらに詳細を示す図である。また、図44は、スマートフォン $e \times 115$ の構成例を示す図である。スマートフォン $e \times 115$ は、基地局 $e \times 110$ との間で電波を送受信するためのアンテナ $e \times 450$ と、映像及び静止画を撮ることが可能なカメラ部 $e \times 465$ と、カメラ部 $e \times 465$ で撮像した映像、及びアンテナ $e \times 450$ で受信した映像等が復号されたデータを表示する表示部 $e \times 458$ とを備える。スマートフォン $e \times 115$ は、さらに、タッチパネル等である操作部 $e \times 466$ と、音声又は音響を出力するためのスピーカ等である音声出力部 $e \times 457$ と、音声を入力するためのマイク等である音声入力部 $e \times 456$ と、撮影した映像或いは静止画、録音した音声、受信した映像或いは静止画、メール等の符号化されたデータ、又は、復号化されたデータを保存可能なメモリ部 $e \times 467$ と、ユーザを特定し、ネットワークをはじめ各種データへのアクセスの認証をするためのSIM (Subscriber Identity Module) $e \times 468$ とのインタフェース部であるスロット部 $e \times 464$ とを備える。なお、メモリ部 $e \times 467$ の代わりに外

付けメモリが用いられてもよい。

- [0549] 表示部 e x 4 5 8 及び操作部 e x 4 6 6 等を統括的に制御する主制御部 e x 4 6 0 と、電源回路部 e x 4 6 1、操作入力制御部 e x 4 6 2、映像信号処理部 e x 4 5 5、カメラインタフェース部 e x 4 6 3、ディスプレイ制御部 e x 4 5 9、変調／復調部 e x 4 5 2、多重／分離部 e x 4 5 3、音声信号処理部 e x 4 5 4、スロット部 e x 4 6 4、及びメモリ部 e x 4 6 7 とが同期バス e x 4 7 0 を介して接続されている。
- [0550] 電源回路部 e x 4 6 1 は、ユーザの操作により電源キーがオン状態にされると、スマートフォン e x 1 1 5 を動作可能な状態に起動し、バッテリーパックから各部に対して電力を供給する。
- [0551] スマートフォン e x 1 1 5 は、CPU、ROM 及び RAM 等を有する主制御部 e x 4 6 0 の制御に基づいて、通話及データ通信等の処理を行う。通話時は、音声入力部 e x 4 5 6 で収音した音声信号を音声信号処理部 e x 4 5 4 でデジタル音声信号に変換し、変調／復調部 e x 4 5 2 でスペクトラム拡散処理を施し、送信／受信部 e x 4 5 1 でデジタルアナログ変換処理及び周波数変換処理を施し、その結果の信号を、アンテナ e x 4 5 0 を介して送信する。また受信データを増幅して周波数変換処理及びアナログデジタル変換処理を施し、変調／復調部 e x 4 5 2 でスペクトラム逆拡散処理し、音声信号処理部 e x 4 5 4 でアナログ音声信号に変換した後、これを音声出力部 e x 4 5 7 から出力する。データ通信モード時は、本体部の操作部 e x 4 6 6 等の操作に基づいてテキスト、静止画、又は映像データが操作入力制御部 e x 4 6 2 を介して主制御部 e x 4 6 0 に送出される。同様の送受信処理が行われる。データ通信モード時に映像、静止画、又は映像と音声を送信する場合、映像信号処理部 e x 4 5 5 は、メモリ部 e x 4 6 7 に保存されている映像信号又はカメラ部 e x 4 6 5 から入力された映像信号を上記各実施の形態で示した動画像符号化方法によって圧縮符号化し、符号化された映像データを多重／分離部 e x 4 5 3 に送出する。音声信号処理部 e x 4 5 4 は、映像又は静止画をカメラ部 e x 4 6 5 で撮像中に音声入力部 e x 4 5 6 で収音し

た音声信号を符号化し、符号化された音声データを多重／分離部 $e \times 453$ に送出する。多重／分離部 $e \times 453$ は、符号化済み映像データと符号化済み音声データを所定の方式で多重化し、変調／復調部（変調／復調回路部） $e \times 452$ 、及び送信／受信部 $e \times 451$ で変調処理及び変換処理を施してアンテナ $e \times 450$ を介して送信する。

[0552] 電子メール又はチャットに添付された映像、又はウェブページにリンクされた映像を受信した場合等において、アンテナ $e \times 450$ を介して受信された多重化データを復号するために、多重／分離部 $e \times 453$ は、多重化データを分離することにより、多重化データを映像データのビットストリームと音声データのビットストリームとに分け、同期バス $e \times 470$ を介して符号化された映像データを映像信号処理部 $e \times 455$ に供給するとともに、符号化された音声データを音声信号処理部 $e \times 454$ に供給する。映像信号処理部 $e \times 455$ は、上記各実施の形態で示した動画像符号化方法に対応した動画像復号化方法によって映像信号を復号し、ディスプレイ制御部 $e \times 459$ を介して表示部 $e \times 458$ から、リンクされた動画像ファイルに含まれる映像又は静止画が表示される。音声信号処理部 $e \times 454$ は、音声信号を復号し、音声出力部 $e \times 457$ から音声が出力される。リアルタイムストリーミングがますます普及しだしているため、ユーザの状況によっては音声の再生が社会的にふさわしくないこともあり得る。そのため、初期値としては、音声信号は再生せず映像データのみを再生する構成の方が望ましく、ユーザが映像データをクリックするなど操作を行った場合にのみ音声を同期して再生してもよい。

[0553] またここではスマートフォン $e \times 115$ を例に説明したが、端末としては符号化器及び復号化器を両方持つ送受信型端末の他に、符号化器のみを有する送信端末、及び、復号化器のみを有する受信端末という3通りの他の実装形式が考えられる。デジタル放送用システムにおいて、映像データに音声データが多重化された多重化データを受信又は送信するとして説明した。ただし、多重化データには、音声データ以外に映像に関連する文字データなどが

多重化されてもよい。また、多重化データではなく映像データ自体が受信又は送信されてもよい。

- [0554] なお、CPUを含む主制御部 460 が符号化又は復号処理を制御するとして説明したが、種々の端末はGPU (Graphics Processing Unit) を備えることも多い。よって、CPUとGPUで共通化されたメモリ、又は共通に使用できるようにアドレスが管理されているメモリにより、GPUの性能を活かして広い領域を一括して処理する構成でもよい。これにより符号化時間を短縮でき、リアルタイム性を確保し、低遅延を実現できる。特に動き探索、デブロックフィルタ、SAO (Sample Adaptive Offset)、及び変換・量子化の処理を、CPUではなく、GPUでピクチャなどの単位で一括して行うと効率的である。

産業上の利用可能性

- [0555] 本開示は、動画像を符号化する符号化装置等に利用可能であり、ビデオ会議システム等に適用可能である。

符号の説明

- [0556] 100、700 符号化装置
 102 分割部
 104 減算部
 106 変換部
 108 量子化部
 110 エントロピー符号化部
 112、204 逆量子化部
 114、206 逆変換部
 116、208 加算部
 118、210 ブロックメモリ
 120、212 ループフィルタ部
 122、214 フレームメモリ
 124、216 イントラ予測部

1 2 6、2 1 8 インター予測部
1 2 8、2 2 0 予測制御部
1 3 0、2 2 2 予測パラメータ生成部
1 3 1、1 3 3、1 3 4、7 3 1、7 3 3 圧縮器
1 3 2、2 3 2、7 3 2、8 3 2 導出器
1 5 1、2 5 1、3 5 1 回路
1 5 2、2 5 2、3 5 2 メモリ
2 0 0、8 0 0 復号装置
2 0 2 エントロピー復号部
2 2 4 分割決定部
2 3 1、2 3 3、2 3 5、8 3 1、8 3 3 圧縮解除器
2 3 4、8 3 4 生成器
2 3 6 合成器
3 0 0 ビットストリーム生成装置

請求の範囲

- [請求項1] メモリと、
 前記メモリに接続された回路とを備え、
 前記回路は、動作において、
 （i）顔を含む画像である基準画像と、（i i）カメラによって得られる撮影動画像の複数のフレームの各々に対応し、被写体の幾何学属性を示す幾何学情報と、（i i i）背景画像に関する背景情報とを1つ以上のストリームから復号し、
 前記基準画像、前記幾何学情報及び前記背景情報から、生成モデルを用いて、前記顔を含む動画像であり、前記背景画像が合成された動画像である合成顔動画像を生成する、
 復号装置。
- [請求項2] 前記回路は、前記生成モデルに前記基準画像、前記幾何学情報及び前記背景画像を入力して、前記合成顔動画像を前記生成モデルから取得する、
 請求項1に記載の復号装置。
- [請求項3] 前記回路は、
 前記生成モデルに前記基準画像及び前記幾何学情報を入力して、前記顔を含む動画像であり、前記背景画像が合成されていない動画像である中間顔動画像を前記生成モデルから取得し、
 前記中間顔動画像における背景領域に前記背景画像における対応領域を埋め込むことにより、前記合成顔動画像を生成する、
 請求項1に記載の復号装置。
- [請求項4] 前記回路は、
 前記中間顔動画像に対してセグメンテーション処理を行って、前記中間顔動画像における前景領域及び前記背景領域を示す中間顔動画像セグメンテーション情報を取得し、
 前記中間顔動画像セグメンテーション情報を用いて、前記中間顔動

画像における前記背景領域を特定する、

請求項 3 に記載の復号装置。

[請求項5]

前記回路は、

前記撮影動画像における前景領域及び背景領域を示す撮影動画像セグメンテーション情報を前記 1 つ以上のストリームから復号し、

前記撮影動画像セグメンテーション情報を用いて、前記中間顔動画像における前記背景領域を特定する、

請求項 3 に記載の復号装置。

[請求項6]

前記基準画像における背景領域には、規定の背景カラーコードが埋め込まれている、

請求項 3 に記載の復号装置。

[請求項7]

前記回路は、前記中間顔動画像において前記規定の背景カラーコードを有する領域を前記中間顔動画像における前記背景領域として特定する、

請求項 6 に記載の復号装置。

[請求項8]

前記回路は、前記規定の背景カラーコードを示す背景カラーコード情報を前記 1 つ以上のストリームから復号する、

請求項 6 又は 7 に記載の復号装置。

[請求項9]

前記背景カラーコード情報は、連続する複数の値を含む範囲を前記規定の背景カラーコードとして示し、

前記規定の背景カラーコードは、前記背景カラーコード情報によって示される前記範囲内で規定される、

請求項 8 に記載の復号装置。

[請求項10]

前記回路は、

前記基準画像における前景領域及び背景領域を示す基準画像セグメンテーション情報を前記 1 つ以上のストリームから復号し、

前記基準画像セグメンテーション情報を用いて、前記基準画像における前記背景領域に、規定の背景カラーコードを埋め込む、

請求項 3 に記載の復号装置。

[請求項11] 前記背景画像は、前記基準画像及び前記撮影動画像によらずに用意された画像である、

請求項 1 ～ 7 のいずれか 1 項に記載の復号装置。

[請求項12] 前記回路は、

前記背景画像の識別子を前記背景情報として復号し、

前記識別子を用いて、複数の背景画像候補の中から前記背景画像を選択する、

請求項 1 ～ 7 のいずれか 1 項に記載の復号装置。

[請求項13] 前記背景画像は、前記撮影動画像に含まれる画像、又は、前記撮影動画像に含まれる複数の画像が合成された画像である、

請求項 1 ～ 7 のいずれか 1 項に記載の復号装置。

[請求項14] 前記回路は、前記基準画像を前記背景情報として復号し、

前記背景画像に、前記基準画像が適用される、

請求項 1 ～ 7 のいずれか 1 項に記載の復号装置。

[請求項15] 前記回路は、前記背景画像に前景領域が含まれる場合、前記背景画像における前記前景領域の周辺領域、又は、過去の合成顔動画像における背景領域を用いて、前記背景画像における背景領域の欠落を補間する、

請求項 1 3 に記載の復号装置。

[請求項16] 前記回路は、

前記背景画像に対してセグメンテーション処理を行って、前記背景画像における前記前景領域及び前記背景領域を示す背景画像セグメンテーション情報を取得し、

前記背景画像セグメンテーション情報を用いて、前記背景画像における前記前景領域及び前記背景領域を特定する、

請求項 1 5 に記載の復号装置。

[請求項17] 前記背景画像における前記前景領域には、規定の前景カラーコード

が埋め込まれている、

請求項 15 に記載の復号装置。

[請求項18]

前記回路は、

前記規定の前景カラーコードを示す前景カラーコード情報を前記 1 つ以上のストリームから復号し、

前記背景画像において前記規定の前景カラーコードを有する領域を前記背景画像における前記前景領域として特定する、

請求項 17 に記載の復号装置。

[請求項19]

前記前景カラーコード情報は、連続する複数の値を含む範囲を前記規定の前景カラーコードとして示し、

前記規定の前景カラーコードは、前記前景カラーコード情報によって示される前記範囲内で規定される、

請求項 18 に記載の復号装置。

[請求項20]

前記背景画像は、前記 1 つ以上のストリームにおけるアクセスユニットからピクチャとして復号され、

前記背景画像が含まれる前記アクセスユニットに対応付けられた S E I (S u p p l e m e n t a l E n h a n c e m e n t I n f o r m a t i o n) から、前記アクセスユニットに前記背景画像が存在することを示す信号が復号される、

請求項 1 ～ 7 のいずれか 1 項に記載の復号装置。

[請求項21]

前記背景画像は、前記合成顔動画画像の複数のフレームに共通して適用される、

請求項 1 ～ 7 のいずれか 1 項に記載の復号装置。

[請求項22]

前記回路は、規定の背景カラーコードを示す背景カラーコード情報と、規定の前景カラーコードを示す前景カラーコード情報とのうち、少なくとも一方を前記 1 つ以上のストリームにおける S E I (S u p p l e m e n t a l E n h a n c e m e n t I n f o r m a t i o n) から復号する、

請求項 1 ～ 7 のいずれか 1 項に記載の復号装置。

[請求項23] 前記回路は、前記撮影動画像における前景領域及び背景領域を示す撮影動画像セグメンテーション情報と、前記基準画像における前景領域及び背景領域を示す基準画像セグメンテーション情報とのうち、少なくとも一方を前記 1 つ以上のストリームにおける S E I (S u p p l e m e n t a l E n h a n c e m e n t I n f o r m a t i o n) から復号する、

請求項 1 ～ 7 のいずれか 1 項に記載の復号装置。

[請求項24] メモリと、
前記メモリに接続された回路とを備え、
前記回路は、動作において、
(i) 顔を含む画像である基準画像と、 (i i) カメラによって得られる撮影動画像の複数のフレームの各々に対応し、被写体の幾何学属性を示す幾何学情報とを 1 つ以上のストリームから復号し、
前記基準画像及び前記幾何学情報を生成モデルに入力して、前記顔を含む動画像である中間顔動画像を前記生成モデルから取得し、
前記中間顔動画像における背景領域に前記基準画像における対応領域を埋め込むことにより、合成顔動画像を生成する、
復号装置。

[請求項25] メモリと、
前記メモリに接続された回路とを備え、
前記回路は、動作において、
顔を含む動画像であり、背景画像が合成された動画像である合成顔動画像を生成するための (i) 基準画像と (i i) 幾何学情報と (i i i) 背景情報とであって、 (i) 前記顔を含む画像である基準画像と、 (i i) カメラによって得られる撮影動画像の複数のフレームの各々に対応し、被写体の幾何学属性を示す幾何学情報と、 (i i i) 前記背景画像に関する背景情報とを 1 つ以上のストリームに符号化す

る、

符号化装置。

[請求項26]

前記回路は、

前記撮影動画像に対してセグメンテーション処理を行って、前記撮影動画像における前景領域及び背景領域を示す撮影動画像セグメンテーション情報を取得し、

前記撮影動画像セグメンテーション情報を前記1つ以上のストリームに符号化する、

請求項25に記載の符号化装置。

[請求項27]

前記回路は、

前記基準画像に対してセグメンテーション処理を行って、前記基準画像における前景領域及び背景領域を示す基準画像セグメンテーション情報を取得し、

前記基準画像セグメンテーション情報を用いて、前記基準画像における前記背景領域に規定の背景カラーコードを埋め込み、

前記背景領域に前記規定の背景カラーコードが埋め込まれた前記基準画像を符号化する、

請求項25に記載の符号化装置。

[請求項28]

前記回路は、前記規定の背景カラーコードを示す背景カラーコード情報を前記1つ以上のストリームに符号化する、

請求項27に記載の符号化装置。

[請求項29]

前記背景カラーコード情報は、連続する複数の値を含む範囲を前記規定の背景カラーコードとして示し、

前記規定の背景カラーコードは、前記背景カラーコード情報によって示される前記範囲内で規定される、

請求項28に記載の符号化装置。

[請求項30]

前記規定の背景カラーコードは、前記基準画像における前記前景領域において発生頻度が閾値以下であるカラーコードで規定される、

請求項 27～29 のいずれか 1 項に記載の符号化装置。

[請求項31]

前記回路は、

前記基準画像に対してセグメンテーション処理を行って、前記基準画像における前景領域及び背景領域を示す基準画像セグメンテーション情報を取得し、

前記基準画像セグメンテーション情報を前記 1 つ以上のストリームに符号化する、

請求項 25 に記載の符号化装置。

[請求項32]

前記背景画像は、前記基準画像及び前記撮影動画像によらずに用意された画像である、

請求項 25～29 のいずれか 1 項に記載の符号化装置。

[請求項33]

前記回路は、

複数の背景画像候補の中から前記背景画像を選択し、

前記背景画像の識別子を前記背景情報として符号化する、

請求項 25～29 のいずれか 1 項に記載の符号化装置。

[請求項34]

前記背景画像は、前記撮影動画像に含まれる画像、又は、前記撮影動画像に含まれる複数の画像が合成された画像である、

請求項 25～29 のいずれか 1 項に記載の符号化装置。

[請求項35]

前記回路は、前記基準画像を前記背景情報として符号化し、

前記背景画像に、前記基準画像が適用される、

請求項 25～29 のいずれか 1 項に記載の符号化装置。

[請求項36]

前記回路は、前記背景画像に前景領域が含まれる場合、前記背景画像における前記前景領域の周辺領域、又は、前記撮影動画像に含まれる他の画像における背景領域を用いて、前記背景画像における背景領域の欠落を補間する、

請求項 34 に記載の符号化装置。

[請求項37]

前記回路は、

前記背景画像に対してセグメンテーション処理を行って、前記背景

画像における前記前景領域及び前記背景領域を示す背景画像セグメンテーション情報を取得し、

前記背景画像セグメンテーション情報を用いて、前記背景画像における前記前景領域及び前記背景領域を特定する、

請求項 36 に記載の符号化装置。

[請求項38]

前記回路は、

前記背景画像に対してセグメンテーション処理を行って、前記背景画像における前景領域及び背景領域を示す背景画像セグメンテーション情報を取得し、

前記背景画像セグメンテーション情報を用いて、前記背景画像における前記前景領域に規定の前景カラーコードを埋め込む、

請求項 34 に記載の符号化装置。

[請求項39]

前記回路は、前記規定の前景カラーコードを示す前景カラーコード情報を前記 1 つ以上のストリームに符号化する、

請求項 38 に記載の符号化装置。

[請求項40]

前記前景カラーコード情報は、連続する複数の値を含む範囲を前記規定の前景カラーコードとして示し、

前記規定の前景カラーコードは、前記前景カラーコード情報によって示される前記範囲内で規定される、

請求項 39 に記載の符号化装置。

[請求項41]

前記規定の前景カラーコードは、前記背景画像における前記背景領域において発生頻度が閾値以下であるカラーコードで規定される、

請求項 38 に記載の符号化装置。

[請求項42]

前記背景画像は、前記 1 つ以上のストリームにおけるアクセスユニットにピクチャとして符号化され、

前記背景画像が符号化される前記アクセスユニットに対応付けられた S E I (S u p p l e m e n t a l E n h a n c e m e n t I n f o r m a t i o n) に、前記アクセスユニットに前記背景画像が

存在することを示す信号が符号化される、

請求項 25～29 のいずれか 1 項に記載の符号化装置。

[請求項43] 前記回路は、規定の背景カラーコードを示す背景カラーコード情報と、規定の前景カラーコードを示す前景カラーコード情報とのうち少なくとも一方を前記 1 つ以上のストリームにおける S E I (S u p p l e m e n t a l E n h a n c e m e n t I n f o r m a t i o n) に符号化する、

請求項 25～29 のいずれか 1 項に記載の符号化装置。

[請求項44] 前記回路は、前記撮影動画像における前景領域及び背景領域を示す撮影動画像セグメンテーション情報と、前記基準画像における前景領域及び背景領域を示す基準画像セグメンテーション情報とのうち少なくとも一方を前記 1 つ以上のストリームにおける S E I (S u p p l e m e n t a l E n h a n c e m e n t I n f o r m a t i o n) に符号化する、

請求項 25～29 のいずれか 1 項に記載の符号化装置。

[請求項45] メモリと、
前記メモリに接続された回路とを備え、
前記回路は、動作において、
合成顔動画像を生成するための (i) 基準画像と (i i) 幾何学情報とであって、 (i) 顔を含む画像である基準画像と、 (i i) カメラによって得られる撮影動画像の複数のフレームの各々に対応し、被写体の幾何学属性を示す幾何学情報とを 1 つ以上のストリームに符号化し、

前記合成顔動画像の生成において、前記基準画像及び前記幾何学情報は、前記基準画像及び前記幾何学情報を生成モデルに入力して、前記顔を含む動画像である中間顔動画像を前記生成モデルから取得することに用いられ、前記基準画像は、さらに、前記中間顔動画像における背景領域に前記基準画像における対応領域を埋め込むことにより、

前記合成顔動画像を生成することに用いられる、
符号化装置。

[請求項46]

メモリと、
前記メモリに接続された回路とを備え、
前記回路は、動作において、
顔を含む動画像であり、背景画像が合成された動画像である合成顔動画像を生成するための (i) 基準画像と (i i) 幾何学情報と (i i i) 背景情報とであって、(i) 前記顔を含む画像である基準画像と、(i i) カメラによって得られる撮影動画像の複数のフレームの各々に対応し、被写体の幾何学属性を示す幾何学情報と、(i i i) 前記背景画像に関する背景情報とを含むビットストリームを生成する、
ビットストリーム生成装置。

[請求項47]

(i) 顔を含む画像である基準画像と、(i i) カメラによって得られる撮影動画像の複数のフレームの各々に対応し、被写体の幾何学属性を示す幾何学情報と、(i i i) 背景画像に関する背景情報とを1つ以上のストリームから復号し、
前記基準画像、前記幾何学情報及び前記背景情報から、生成モデルを用いて、前記顔を含む動画像であり、前記背景画像が合成された動画像である合成顔動画像を生成する、
復号方法。

[請求項48]

顔を含む動画像であり、背景画像が合成された動画像である合成顔動画像を生成するための (i) 基準画像と (i i) 幾何学情報と (i i i) 背景情報とであって、(i) 前記顔を含む画像である基準画像と、(i i) カメラによって得られる撮影動画像の複数のフレームの各々に対応し、被写体の幾何学属性を示す幾何学情報と、(i i i) 前記背景画像に関する背景情報とを1つ以上のストリームに符号化する、

符号化方法。