

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6232075号
(P6232075)

(45) 発行日 平成29年11月22日 (2017.11.22)

(24) 登録日 平成29年10月27日 (2017.10.27)

(51) Int.Cl.

F I

H O 4 N 19/597 (2014.01)

H O 4 N 19/597

請求項の数 18 (全 32 頁)

| | | | |
|---------------|------------------------------|-----------|---------------------|
| (21) 出願番号 | 特願2015-551543 (P2015-551543) | (73) 特許権者 | 000004226 |
| (86) (22) 出願日 | 平成26年12月3日 (2014.12.3) | | 日本電信電話株式会社 |
| (86) 国際出願番号 | PCT/JP2014/081986 | | 東京都千代田区大手町一丁目5番1号 |
| (87) 国際公開番号 | W02015/083742 | (74) 代理人 | 110001634 |
| (87) 国際公開日 | 平成27年6月11日 (2015.6.11) | | 特許業務法人 志賀国際特許事務所 |
| 審査請求日 | 平成28年1月20日 (2016.1.20) | (72) 発明者 | 志水 信哉 |
| (31) 優先権主張番号 | 特願2013-250429 (P2013-250429) | | 東京都千代田区大手町一丁目5番1号 日 |
| (32) 優先日 | 平成25年12月3日 (2013.12.3) | | 本電信電話株式会社内 |
| (33) 優先権主張国 | 日本国 (JP) | (72) 発明者 | 杉本 志織 |
| | | | 東京都千代田区大手町一丁目5番1号 日 |
| | | | 本電信電話株式会社内 |
| | | (72) 発明者 | 小島 明 |
| | | | 東京都千代田区大手町一丁目5番1号 日 |
| | | | 本電信電話株式会社内 |

最終頁に続く

(54) 【発明の名称】 映像符号化装置及び方法、映像復号装置及び方法、及び、それらのプログラム

(57) 【特許請求の範囲】

【請求項 1】

複数の異なる視点の映像からなる多視点映像の1フレームである符号化対象画像を符号化する際に、前記符号化対象画像を分割した領域である符号化対象領域ごとに、異なる視点間で予測しながら符号化を行う映像符号化装置であって、

前記多視点映像中の被写体に対するデプスマップから代表デプスを設定する代表デプス設定手段と、

前記代表デプスに基づいて、前記符号化対象画像上の位置を、該符号化対象画像とは異なる参照視点に対する参照視点画像上の位置との対応関係を与える変換行列を設定する変換行列設定手段と、

前記符号化対象領域内の位置から代表位置を設定する代表位置設定手段と、

前記代表位置と前記変換行列を用いて、前記代表位置に対する前記参照視点画像上での対応位置を設定する対応位置設定手段と、

前記対応位置に基づいて、前記参照視点画像の動き情報である参照視点動き情報から前記符号化対象領域における合成動き情報を生成する動き情報生成手段と、

前記合成動き情報を用いて、前記符号化対象領域に対する予測画像を生成する予測画像生成手段と、

前記符号化対象領域に対して、前記デプスマップ上での対応領域であるデプス領域を設定するデプス領域設定手段と、

前記符号化対象領域に対して、前記デプスマップに対する視差ベクトルであるデプス参

照視差ベクトルを設定するデプス参照視差ベクトル設定手段と
を有し、

前記代表デプス設定手段は、前記デプス領域に対する前記デプスマップから代表デプスを設定し、

前記デプス領域設定手段は、前記デプス参照視差ベクトルによって示される領域を前記デプス領域として設定することを特徴とする映像符号化装置。

【請求項 2】

前記デプス参照視差ベクトル設定手段は、前記符号化対象領域に隣接する領域を符号化
する際に使用した視差ベクトルを用いて、前記デプス参照視差ベクトルを設定することを
特徴とする請求項 1 に記載の映像符号化装置。

10

【請求項 3】

複数の異なる視点の映像からなる多視点映像の 1 フレームである符号化対象画像を符号
化する際に、前記符号化対象画像を分割した領域である符号化対象領域ごとに、異なる視
点間で予測しながら符号化を行う映像符号化装置であって、

前記多視点映像中の被写体に対するデプスマップから代表デプスを設定する代表デプス
設定手段と、

前記代表デプスに基づいて、前記符号化対象画像上の位置を、該符号化対象画像とは異
なる参照視点に対する参照視点画像上の位置へと変換する変換行列を設定する変換行列設
定手段と、

前記符号化対象領域内の位置から代表位置を設定する代表位置設定手段と、

20

前記代表位置と前記変換行列を用いて、前記代表位置に対する前記参照視点画像上での
対応位置を設定する対応位置設定手段と、

前記対応位置に基づいて、前記参照視点画像の動き情報である参照視点動き情報から前
記符号化対象領域における合成動き情報を生成する動き情報生成手段と、

前記合成動き情報を用いて、前記符号化対象領域に対する予測画像を生成する予測画像
生成手段と、

前記変換行列を用いて、前記合成動き情報を変換する合成動き情報変換手段と

を有し、

前記予測画像生成手段は、前記変換された合成動き情報を用いることを特徴とする映像
符号化装置。

30

【請求項 4】

複数の異なる視点の映像からなる多視点映像の 1 フレームである符号化対象画像を符号
化する際に、前記符号化対象画像を分割した領域である符号化対象領域ごとに、異なる視
点間で予測しながら符号化を行う映像符号化装置であって、

前記多視点映像中の被写体に対するデプスマップから代表デプスを設定する代表デプス
設定手段と、

前記代表デプスに基づいて、前記符号化対象画像上の位置を、該符号化対象画像とは異
なる参照視点に対する参照視点画像上の位置へと変換する変換行列を設定する変換行列設
定手段と、

前記符号化対象領域内の位置から代表位置を設定する代表位置設定手段と、

40

前記代表位置と前記変換行列を用いて、前記代表位置に対する前記参照視点画像上での
対応位置を設定する対応位置設定手段と、

前記対応位置に基づいて、前記参照視点画像の動き情報である参照視点動き情報から前
記符号化対象領域における合成動き情報を生成する動き情報生成手段と、

前記合成動き情報を用いて、前記符号化対象領域に対する予測画像を生成する予測画像
生成手段と、

前記対応位置と前記合成動き情報とに基づいて、前記デプスマップから過去デプスを設
定する過去デプス設定手段と、

前記過去デプスに基づいて、前記参照視点画像上の位置を前記符号化対象画像上の位置
へと変換する逆変換行列を設定する逆変換行列設定手段と、

50

前記逆変換行列を用いて、前記合成動き情報を変換する合成動き情報変換手段とを有し、

前記予測画像生成手段は、前記変換された合成動き情報を用いることを特徴とする映像符号化装置。

【請求項 5】

複数の異なる視点の映像からなる多視点映像の 1 フレームである符号化対象画像を符号化する際に、前記符号化対象画像を分割した領域である符号化対象領域ごとに、異なる視点間で予測しながら符号化を行う映像符号化装置であって、

前記多視点映像中の被写体に対するデプスマップから代表デプスを設定する代表デプス設定手段と、

前記代表デプスに基づいて、前記符号化対象画像上の位置を、該符号化対象画像とは異なる参照視点に対する参照視点画像上の位置へと変換する変換行列を設定する変換行列設定手段と、

前記符号化対象領域内の位置から代表位置を設定する代表位置設定手段と、

前記代表位置と前記変換行列を用いて、前記代表位置に対する前記参照視点画像上での対応位置を設定する対応位置設定手段と、

前記対応位置に基づいて、前記参照視点画像の動き情報である参照視点動き情報から前記符号化対象領域における合成動き情報を生成する動き情報生成手段と、

前記合成動き情報を用いて、前記符号化対象領域に対する予測画像を生成する予測画像生成手段と

を有し、

前記符号化対象画像の視点と前記参照視点との位置関係の変化がない、または所定の大きさ以下の場合には、前記変換行列設定手段による変換行列の設定を行わずに、前記対応位置設定手段は直前に符号化された画像で用いた前記変換行列を用いることを特徴とする映像符号化装置。

【請求項 6】

複数の異なる視点の映像からなる多視点動画の符号データから、復号対象画像を復号する際に、前記復号対象画像を分割した領域である復号対象領域ごとに、異なる視点間で予測しながら復号を行う映像復号装置であって、

前記多視点映像中の被写体に対するデプスマップから代表デプスを設定する代表デプス設定手段と、

前記代表デプスに基づいて、前記復号対象画像上の位置を、該復号対象画像とは異なる参照視点に対する参照画像上の位置との対応関係を与える変換行列を設定する変換行列設定手段と、

前記復号対象領域内の位置から代表位置を設定する代表位置設定手段と、

前記代表位置と前記変換行列を用いて、前記代表位置に対する前記参照視点画像上での対応位置を設定する対応位置設定手段と、

前記対応位置に基づいて、前記参照視点画像の動き情報である参照視点動き情報から前記復号対象領域における合成動き情報を生成する動き情報生成手段と、

前記合成動き情報を用いて、前記復号対象領域に対する予測画像を生成する予測画像生成手段と、

前記復号対象領域に対して、前記デプスマップ上での対応領域であるデプス領域を設定するデプス領域設定手段と、

前記復号対象領域に対して、前記デプスマップに対する視差ベクトルであるデプス参照視差ベクトルを設定するデプス参照視差ベクトル設定手段と

を有し、

前記代表デプス設定手段は、前記デプス領域に対する前記デプスマップから代表デプスを設定し、

前記デプス領域設定手段は、前記デプス参照視差ベクトルによって示される領域を前記デプス領域として設定することを特徴とする映像復号装置。

10

20

30

40

50

【請求項 7】

前記デプス参照視差ベクトル設定手段は、前記復号対象領域に隣接する領域を復号する際に使用した視差ベクトルを用いて、前記デプス参照視差ベクトルを設定することを特徴とする請求項 6 に記載の映像復号装置。

【請求項 8】

複数の異なる視点の映像からなる多視点動画像の符号データから、復号対象画像を復号する際に、前記復号対象画像を分割した領域である復号対象領域ごとに、異なる視点間で予測しながら復号を行う映像復号装置であって、

前記多視点映像中の被写体に対するデプスマップから代表デプスを設定する代表デプス設定手段と、

前記代表デプスに基づいて、前記復号対象画像上の位置を、該復号対象画像とは異なる参照視点に対する参照画像上の位置へと変換する変換行列を設定する変換行列設定手段と

、前記復号対象領域内の位置から代表位置を設定する代表位置設定手段と、

前記代表位置と前記変換行列を用いて、前記代表位置に対する前記参照視点画像上での対応位置を設定する対応位置設定手段と、

前記対応位置に基づいて、前記参照視点画像の動き情報である参照視点動き情報から前記復号対象領域における合成動き情報を生成する動き情報生成手段と、

前記合成動き情報を用いて、前記復号対象領域に対する予測画像を生成する予測画像生成手段と、

前記変換行列を用いて、前記合成動き情報を変換する合成動き情報変換手段と

を有し、

前記予測画像生成手段は、前記変換された合成動き情報を用いることを特徴とする映像復号装置。

【請求項 9】

複数の異なる視点の映像からなる多視点動画像の符号データから、復号対象画像を復号する際に、前記復号対象画像を分割した領域である復号対象領域ごとに、異なる視点間で予測しながら復号を行う映像復号装置であって、

前記多視点映像中の被写体に対するデプスマップから代表デプスを設定する代表デプス設定手段と、

前記代表デプスに基づいて、前記復号対象画像上の位置を、該復号対象画像とは異なる参照視点に対する参照画像上の位置へと変換する変換行列を設定する変換行列設定手段と

、前記復号対象領域内の位置から代表位置を設定する代表位置設定手段と、

前記代表位置と前記変換行列を用いて、前記代表位置に対する前記参照視点画像上での対応位置を設定する対応位置設定手段と、

前記対応位置に基づいて、前記参照視点画像の動き情報である参照視点動き情報から前記復号対象領域における合成動き情報を生成する動き情報生成手段と、

前記合成動き情報を用いて、前記復号対象領域に対する予測画像を生成する予測画像生成手段と、

前記対応位置と前記合成動き情報とに基づいて、前記デプスマップから過去デプスを設定する過去デプス設定手段と、

前記過去デプスに基づいて、前記参照視点画像上の位置を前記復号対象画像上の位置へと変換する逆変換行列を設定する逆変換行列設定手段と、

前記逆変換行列を用いて、前記合成動き情報を変換する合成動き情報変換手段と

を有し、

前記予測画像生成手段は、前記変換された合成動き情報を用いることを特徴とする映像復号装置。

【請求項 10】

複数の異なる視点の映像からなる多視点動画像の符号データから、復号対象画像を復号

10

20

30

40

50

する際に、前記復号対象画像を分割した領域である復号対象領域ごとに、異なる視点間で予測しながら復号を行う映像復号装置であって、

前記多視点映像中の被写体に対するデプスマップから代表デプスを設定する代表デプス設定手段と、

前記代表デプスに基づいて、前記復号対象画像上の位置を、該復号対象画像とは異なる参照視点に対する参照画像上の位置へと変換する変換行列を設定する変換行列設定手段と、

、

前記復号対象領域内の位置から代表位置を設定する代表位置設定手段と、

前記代表位置と前記変換行列を用いて、前記代表位置に対する前記参照視点画像上での対応位置を設定する対応位置設定手段と、

前記対応位置に基づいて、前記参照視点画像の動き情報である参照視点動き情報から前記復号対象領域における合成動き情報を生成する動き情報生成手段と、

前記合成動き情報を用いて、前記復号対象領域に対する予測画像を生成する予測画像生成手段と

を有し、

前記復号対象画像の視点と前記参照視点との位置関係の変化がない、または所定の大きさ以下の場合には、前記変換行列設定手段による変換行列の設定を行わずに、前記対応位置設定手段は直前に復号された画像で用いた前記変換行列を用いることを特徴とする映像復号装置。

【請求項 1 1】

複数の異なる視点の映像からなる多視点動画像の符号データから、復号対象画像を復号する際に、前記復号対象画像を分割した領域である復号対象領域ごとに、異なる視点間で予測しながら復号を行う映像復号方法であって、

前記多視点映像中の被写体に対するデプスマップから代表デプスを設定する代表デプス設定ステップと、

前記代表デプスに基づいて、前記復号対象画像上の位置を、該復号対象画像とは異なる参照視点に対する参照画像上の位置との対応関係を与える変換行列を設定する変換行列設定ステップと、

前記復号対象領域内の位置から代表位置を設定する代表位置設定ステップと、

前記代表位置と前記変換行列を用いて、前記代表位置に対する前記参照視点画像上での対応位置を設定する対応位置設定ステップと、

前記対応位置に基づいて、前記参照視点画像の動き情報である参照視点動き情報から前記復号対象領域における合成動き情報を生成する動き情報生成ステップと、

前記合成動き情報を用いて、前記復号対象領域に対する予測画像を生成する予測画像生成ステップと、

前記復号対象領域に対して、前記デプスマップ上での対応領域であるデプス領域を設定するデプス領域設定ステップと、

前記復号対象領域に対して、前記デプスマップに対する視差ベクトルであるデプス参照視差ベクトルを設定するデプス参照視差ベクトル設定ステップと

を有し、

前記代表デプス設定ステップは、前記デプス領域に対する前記デプスマップから代表デプスを設定し、

前記デプス領域設定ステップは、前記デプス参照視差ベクトルによって示される領域を前記デプス領域として設定することを特徴とする映像復号方法。

【請求項 1 2】

前記デプス参照視差ベクトル設定ステップは、前記復号対象領域に隣接する領域を復号する際に使用した視差ベクトルを用いて、前記デプス参照視差ベクトルを設定することを特徴とする請求項 1 1 に記載の映像復号方法。

【請求項 1 3】

複数の異なる視点の映像からなる多視点動画像の符号データから、復号対象画像を復号

10

20

30

40

50

する際に、前記復号対象画像を分割した領域である復号対象領域ごとに、異なる視点間で予測しながら復号を行う映像復号方法であって、

前記多視点映像中の被写体に対するデプスマップから代表デプスを設定する代表デプス設定ステップと、

前記代表デプスに基づいて、前記復号対象画像上の位置を、該復号対象画像とは異なる参照視点に対する参照画像上の位置へと変換する変換行列を設定する変換行列設定ステップと、

前記復号対象領域内の位置から代表位置を設定する代表位置設定ステップと、

前記代表位置と前記変換行列を用いて、前記代表位置に対する前記参照視点画像上での対応位置を設定する対応位置設定ステップと、

前記対応位置に基づいて、前記参照視点画像の動き情報である参照視点動き情報から前記復号対象領域における合成動き情報を生成する動き情報生成ステップと、

前記合成動き情報を用いて、前記復号対象領域に対する予測画像を生成する予測画像生成ステップと、

前記変換行列を用いて、前記合成動き情報を変換する合成動き情報変換ステップとを有し、

前記予測画像生成ステップは、前記変換された合成動き情報を用いることを特徴とする映像復号方法。

【請求項 14】

複数の異なる視点の映像からなる多視点動画の符号データから、復号対象画像を復号する際に、前記復号対象画像を分割した領域である復号対象領域ごとに、異なる視点間で予測しながら復号を行う映像復号方法であって、

前記多視点映像中の被写体に対するデプスマップから代表デプスを設定する代表デプス設定ステップと、

前記代表デプスに基づいて、前記復号対象画像上の位置を、該復号対象画像とは異なる参照視点に対する参照画像上の位置へと変換する変換行列を設定する変換行列設定ステップと、

前記復号対象領域内の位置から代表位置を設定する代表位置設定ステップと、

前記代表位置と前記変換行列を用いて、前記代表位置に対する前記参照視点画像上での対応位置を設定する対応位置設定ステップと、

前記対応位置に基づいて、前記参照視点画像の動き情報である参照視点動き情報から前記復号対象領域における合成動き情報を生成する動き情報生成ステップと、

前記合成動き情報を用いて、前記復号対象領域に対する予測画像を生成する予測画像生成ステップと、

前記対応位置と前記合成動き情報とに基づいて、前記デプスマップから過去デプスを設定する過去デプス設定ステップと、

前記過去デプスに基づいて、前記参照視点画像上の位置を前記復号対象画像上の位置へと変換する逆変換行列を設定する逆変換行列設定ステップと、

前記逆変換行列を用いて、前記合成動き情報を変換する合成動き情報変換ステップとを有し、

前記予測画像生成ステップは、前記変換された合成動き情報を用いることを特徴とする映像復号方法。

【請求項 15】

複数の異なる視点の映像からなる多視点動画の符号データから、復号対象画像を復号する際に、前記復号対象画像を分割した領域である復号対象領域ごとに、異なる視点間で予測しながら復号を行う映像復号方法であって、

前記多視点映像中の被写体に対するデプスマップから代表デプスを設定する代表デプス設定ステップと、

前記代表デプスに基づいて、前記復号対象画像上の位置を、該復号対象画像とは異なる参照視点に対する参照画像上の位置へと変換する変換行列を設定する変換行列設定ステッ

10

20

30

40

50

プと、

前記復号対象領域内の位置から代表位置を設定する代表位置設定ステップと、

前記代表位置と前記変換行列を用いて、前記代表位置に対する前記参照視点画像上での対応位置を設定する対応位置設定ステップと、

前記対応位置に基づいて、前記参照視点画像の動き情報である参照視点動き情報から前記復号対象領域における合成動き情報を生成する動き情報生成ステップと、

前記合成動き情報を用いて、前記復号対象領域に対する予測画像を生成する予測画像生成ステップと

を有し、

前記復号対象画像の視点と前記参照視点との位置関係の変化がない、または所定の大きさ以下の場合には、前記変換行列設定ステップによる変換行列の設定を行わずに、前記対応位置設定ステップは直前に復号された画像で用いた前記変換行列を用いることを特徴とする映像復号方法。

【請求項 16】

複数の異なる視点の映像からなる多視点映像の1フレームである符号化対象画像を符号化する際に、前記符号化対象画像を分割した領域である符号化対象領域ごとに、異なる視点間で予測しながら符号化を行う映像符号化方法であって、

前記多視点映像中の被写体に対するデプスマップから代表デプスを設定する代表デプス設定ステップと、

前記代表デプスに基づいて、前記符号化対象画像上の位置を、該符号化対象画像とは異なる参照視点に対する参照視点画像上の位置との対応関係を与える変換行列を設定する変換行列設定ステップと、

前記符号化対象領域内の位置から代表位置を設定する代表位置設定ステップと、

前記代表位置と前記変換行列を用いて、前記代表位置に対する前記参照視点画像上での対応位置を設定する対応位置設定ステップと、

前記対応位置に基づいて、前記参照視点画像の動き情報である参照視点動き情報から前記符号化対象領域における合成動き情報を生成する動き情報生成ステップと、

前記合成動き情報を用いて、前記符号化対象領域に対する予測画像を生成する予測画像生成ステップと、

前記符号化対象領域に対して、前記デプスマップ上での対応領域であるデプス領域を設定するデプス領域設定ステップと、

前記符号化対象領域に対して、前記デプスマップに対する視差ベクトルであるデプス参照視差ベクトルを設定するデプス参照視差ベクトル設定ステップと

を有し、

前記代表デプス設定ステップは、前記デプス領域に対する前記デプスマップから代表デプスを設定し、

前記デプス領域設定ステップは、前記デプス参照視差ベクトルによって示される領域を前記デプス領域として設定することを特徴とする映像符号化方法。

【請求項 17】

コンピュータに、請求項 11 に記載の映像復号方法を実行させるための映像復号プログラム。

【請求項 18】

コンピュータに、請求項 16 に記載の映像符号化方法を実行させるための映像符号化プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、映像符号化装置、映像復号装置、映像符号化方法、映像復号方法、映像符号化プログラム、及び、映像復号プログラムに関する。

【背景技術】

10

20

30

40

50

【 0 0 0 2 】

自由視点映像とは、撮影空間内でのカメラの位置や向き（以下、視点と称する）をユーザが自由に指定できる映像のことである。自由視点映像では、ユーザが任意の視点を指定するが、可能な全ての視点に対する映像を保持することは不可能である。そのため、自由視点映像は、指定された視点の映像を生成するのに必要な情報群によって構成される。

なお、自由視点映像は、自由視点テレビ、任意視点映像、任意視点テレビなどと呼ばれることもある。

【 0 0 0 3 】

自由視点映像は様々なデータ形式を用いて表現されるが、最も一般的な形式として映像とその映像の各フレームに対するデプスマップ（距離画像）を用いる方式がある（例えば、非特許文献 1 参照）。

ここで、デプスマップとは、カメラから被写体までのデプス（距離）を画素ごとに表現したものであり、被写体の三次元的な位置を表現している。ある条件を満たす場合、デプスは二つのカメラ間の視差の逆数に比例しているため、ディスパリティマップ（視差画像）と呼ばれることもある。

【 0 0 0 4 】

コンピュータグラフィックスの分野では、デプスは Z バッファに記憶された情報となるため、Z 画像や Z マップと呼ばれることもある。

なお、カメラから被写体までの距離の他に、表現対象空間上に張られた三次元座標系の Z 軸に対する座標値をデプスとして用いることもある。一般に、撮影された画像に対して水平方向を X 軸、垂直方向を Y 軸とするため、Z 軸はカメラの向きと一致するが、複数のカメラに対して共通の座標系を用いる場合など、Z 軸がカメラの向きと一致しない場合もある。

以下では、距離・Z 値を区別せずにデプスと呼び、デプスを画素値として表した画像をデプスマップと呼ぶ。ただし、厳密にはディスパリティマップでは基準となるカメラ対を設定する必要がある。

【 0 0 0 5 】

デプスを画素値として表す際に、物理量に対応する値をそのまま画素値とする方法と、最小値と最大値の間をある数に量子化して得られる値を用いる方法と、最小値からの差をあるステップ幅で量子化して得られる値を用いる方法がある。表現したい範囲が限られている場合には、最小値などの付加情報を用いるほうがデプスを高精度に表現することができる。

また、等間隔に量子化する際に、物理量をそのまま量子化する方法と物理量の逆数を量子化する方法とがある。距離の逆数は視差に比例した値となるため、距離を高精度に表現する必要がある場合には、前者が使用され、視差を高精度に表現する必要がある場合には、後者が使用されることが多い。

以下では、デプスの画素値化の方法や量子化の方法に関係なく、デプスが画像として表現されたものを全てデプスマップと呼ぶ。

【 0 0 0 6 】

デプスマップは、各画素が一つの値を持つ画像として表現されるため、グレースケール画像とみなすことができる。また、被写体を実空間上で連続的に存在し、瞬間的に離れた位置へ移動することができないため、画像信号と同様に空間的相関および時間的相関を持つと言える。したがって、通常の画像信号や映像信号を符号化するために用いられる画像符号化方式や映像符号化方式によって、デプスマップや連続するデプスマップで構成される映像を空間的冗長性や時間的冗長性を取り除きながら効率的に符号化することが可能である。

以下では、デプスマップとそれにより構成される映像を区別せずにデプスマップと呼ぶ。

【 0 0 0 7 】

ここで、一般的な映像符号化について説明する。

映像符号化では、被写体が空間的および時間的に連続しているという特徴を利用して効率的な符号化を実現するために、映像の各フレームをマクロブロックと呼ばれる処理単位ブロックに分割し、マクロブロックごとにその映像信号を空間的または時間的に予測し、その予測方法を示す予測情報と予測残差とを符号化する。

映像信号を空間的に予測する場合は、例えば空間的な予測の方向を示す情報が予測情報となり、時間的に予測する場合は、例えば参照するフレームを示す情報とそのフレーム中の位置を示す情報とが予測情報となる。

空間的に行う予測は、フレーム内の予測であることから、フレーム内予測（画面内予測、イントラ予測）と呼ばれ、時間的に行う予測は、フレーム間の予測であることから、フレーム間予測（画面間予測、インター予測）と呼ばれる。

【0008】

また、時間的に行う予測では、映像の時間的変化、すなわち動きを補償して映像信号の予測を行うことになるため、動き補償予測とも呼ばれる。

さらに、同じシーンを複数の位置や向きから撮影した映像からなる多視点映像を符号化するには、映像の視点間の变化、すなわち視差を補償して映像信号の予測を行うことになるため、視差補償予測が用いられる。

【0009】

複数の視点に対する映像とデプスマップとで構成される自由視点映像の符号化においては、どちらも空間相関と時間相関を持つことから、通常の映像符号化方式を用いてそれぞれを符号化することで、データ量を削減できる。

例えば、MPEG-CPart.3を用いて、多視点映像とそれに対するデプスマップを表現する場合は、それぞれを既存の映像符号化方式を用いて符号化する。

【0010】

また、複数の視点に対する映像とデプスマップとを一緒に符号化する場合、動き情報について視点間で存在する相関を利用して、効率的な符号化を実現する方法がある。

非特許文献2では、処理対象の領域に対して、視差ベクトルを用いて、既に処理済みの別の視点の映像の領域を決定し、その領域を符号化する際に使用された動き情報を、処理対象の領域の動き情報またはその予測値として用いている。このとき効率的な符号化を実現するためには、処理対象の領域に対して精度の高い視差ベクトルを獲得する必要がある。

非特許文献2では、最も単純な方法として、処理対象の領域と時間または空間的に隣接する領域に対して与えられた視差ベクトルを、処理対象領域の視差ベクトルとする方法が用いられている。更に、より正確な視差ベクトルを求めるために、処理対象の領域に対するデプスを推定または取得し、そのデプスを変換して視差ベクトルを獲得する方法も用いられている。

【先行技術文献】

【非特許文献】

【0011】

【非特許文献1】Y. Mori, N. Fukusima, T. Fujii, and M. Tanimoto, "View Generation with 3D Warping Using Depth Information for FTV", In Proceedings of 3DTV-CON2008, pp. 229-232, May 2008.

【非特許文献2】G. Tech, K. Wegner, Y. Chen, and S. Yea, "3D-HEVC Draft Text 1", JCT-3V Doc., JCT3V-E1001 (version 3), September, 2013.

【発明の概要】

【発明が解決しようとする課題】

【0012】

非特許文献2に記載の方法によれば、デプスマップの値を変換し高精度な視差ベクトルを獲得することで、高効率な予測符号化を実現することが可能である。

【0013】

しかしながら、非特許文献2に記載の方法では、デプスを視差ベクトルへ変換する際に

10

20

30

40

50

、視差がデプス（カメラから被写体までの距離）の逆数に比例していると仮定している。より具体的には、デプスの逆数、カメラの焦点距離、視点間の距離の、三者の積によって視差を求めている。このような変換は、2つの視点が同じ焦点距離を持ち、視点の向き（カメラの光軸）が3次的に平行である場合には正しい結果を与えるが、それ以外の状況では誤った結果を与えることになる。

【0014】

正確な変換を行うためには、非特許文献1に記載されているように、画像上の点をデプスに従って三次元空間へ逆投影することで三次元点を得た後、その三次元点を別の視点へ再投影することで別の視点に対する画像上での点を計算する必要がある。

【0015】

しかしながら、このような変換では複雑な演算が必要となり、演算量が増加してしまうという問題がある。また、視点の向きが異なる場合、2つの視点に対する映像上での動きベクトルが同じになることは極めて少ない。そのため、視差ベクトルが正しく得られたとしても、非特許文献2に記載の方法に従って、別の視点における動き情報を処理対象の領域に対する動き情報として用いた場合、誤った動き情報を与えてしまい、効率的な符号化を実現することができないという問題がある。

【0016】

本発明は、このような事情に鑑みてなされたもので、複数の視点に対する映像とデプスマップとを構成要素に持つ自由視点映像データの符号化において、視点の向きが平行でない場合でも、動きベクトルの視点間予測の精度を向上させることで、効率的な映像符号化を実現することができる映像符号化装置、映像復号装置、映像符号化方法、映像復号方法、映像符号化プログラム、及び、映像復号プログラムを提供することを目的とする。

【課題を解決するための手段】

【0017】

本発明は、複数の異なる視点の映像からなる多視点映像の1フレームである符号化対象画像を符号化する際に、前記符号化対象画像を分割した領域である符号化対象領域ごとに、異なる視点間で予測しながら符号化を行う映像符号化装置であって、

前記多視点映像中の被写体に対するデプスマップから代表デプスを設定する代表デプス設定手段と、

前記代表デプスに基づいて、前記符号化対象画像上の位置を、該符号化対象画像とは異なる参照視点に対する参照視点画像上の位置へと変換する変換行列を設定する変換行列設定手段と、

前記符号化対象領域内の位置から代表位置を設定する代表位置設定手段と、

前記代表位置と前記変換行列を用いて、前記代表位置に対する前記参照視点画像上での対応位置を設定する対応位置設定手段と、

前記対応位置に基づいて、前記参照視点画像の動き情報である参照視点動き情報から前記符号化対象領域における合成動き情報を生成する動き情報生成手段と、

前記合成動き情報を用いて、前記符号化対象領域に対する予測画像を生成する予測画像生成手段と、

前記符号化対象領域に対して、前記デプスマップ上での対応領域であるデプス領域を設定するデプス領域設定手段と、

前記符号化対象領域に対して、前記デプスマップに対する視差ベクトルであるデプス参照視差ベクトルを設定するデプス参照視差ベクトル設定手段と

を有し、

前記代表デプス設定手段は、前記デプス領域に対する前記デプスマップから代表デプスを設定し、

前記デプス領域設定手段は、前記デプス参照視差ベクトルによって示される領域を前記デプス領域として設定することを特徴とする映像符号化装置を提供する。

【0020】

更に、前記デプス参照視差ベクトル設定手段は、前記符号化対象領域に隣接する領域を

10

20

30

40

50

符号化する際に使用した視差ベクトルを用いて、前記デプス参照視差ベクトルを設定するようにしても良い。

【0022】

本発明はまた、複数の異なる視点の映像からなる多視点映像の1フレームである符号化対象画像を符号化する際に、前記符号化対象画像を分割した領域である符号化対象領域ごとに、異なる視点間で予測しながら符号化を行う映像符号化装置であって、

前記多視点映像中の被写体に対するデプスマップから代表デプスを設定する代表デプス設定手段と、

前記代表デプスに基づいて、前記符号化対象画像上の位置を、該符号化対象画像とは異なる参照視点に対する参照視点画像上の位置へと変換する変換行列を設定する変換行列設定手段と、

前記符号化対象領域内の位置から代表位置を設定する代表位置設定手段と、

前記代表位置と前記変換行列を用いて、前記代表位置に対する前記参照視点画像上での対応位置を設定する対応位置設定手段と、

前記対応位置に基づいて、前記参照視点画像の動き情報である参照視点動き情報から前記符号化対象領域における合成動き情報を生成する動き情報生成手段と、

前記合成動き情報を用いて、前記符号化対象領域に対する予測画像を生成する予測画像生成手段と、

前記変換行列を用いて、前記合成動き情報を変換する合成動き情報変換手段と

を有し、

前記予測画像生成手段は、前記変換された合成動き情報を用いることを特徴とする映像符号化装置も提供する。

【0023】

本発明はまた、複数の異なる視点の映像からなる多視点映像の1フレームである符号化対象画像を符号化する際に、前記符号化対象画像を分割した領域である符号化対象領域ごとに、異なる視点間で予測しながら符号化を行う映像符号化装置であって、

前記多視点映像中の被写体に対するデプスマップから代表デプスを設定する代表デプス設定手段と、

前記代表デプスに基づいて、前記符号化対象画像上の位置を、該符号化対象画像とは異なる参照視点に対する参照視点画像上の位置へと変換する変換行列を設定する変換行列設定手段と、

前記符号化対象領域内の位置から代表位置を設定する代表位置設定手段と、

前記代表位置と前記変換行列を用いて、前記代表位置に対する前記参照視点画像上での対応位置を設定する対応位置設定手段と、

前記対応位置に基づいて、前記参照視点画像の動き情報である参照視点動き情報から前記符号化対象領域における合成動き情報を生成する動き情報生成手段と、

前記合成動き情報を用いて、前記符号化対象領域に対する予測画像を生成する予測画像生成手段と、

前記対応位置と前記合成動き情報とに基づいて、前記デプスマップから過去デプスを設定する過去デプス設定手段と、

前記過去デプスに基づいて、前記参照視点画像上の位置を前記符号化対象画像上の位置へと変換する逆変換行列を設定する逆変換行列設定手段と、前記逆変換行列を用いて、前記合成動き情報を変換する合成動き情報変換手段と

を有し、

前記予測画像生成手段は、前記変換された合成動き情報を用いることを特徴とする映像符号化装置も提供する。

【0024】

本発明はまた、複数の異なる視点の映像からなる多視点映像の1フレームである符号化対象画像を符号化する際に、前記符号化対象画像を分割した領域である符号化対象領域ごとに、異なる視点間で予測しながら符号化を行う映像符号化装置であって、

前記多視点映像中の被写体に対するデプスマップから代表デプスを設定する代表デプス設定手段と、

前記代表デプスに基づいて、前記符号化対象画像上の位置を、該符号化対象画像とは異なる参照視点に対する参照視点画像上の位置へと変換する変換行列を設定する変換行列設定手段と、

前記符号化対象領域内の位置から代表位置を設定する代表位置設定手段と、

前記代表位置と前記変換行列を用いて、前記代表位置に対する前記参照視点画像上での対応位置を設定する対応位置設定手段と、

前記対応位置に基づいて、前記参照視点画像の動き情報である参照視点動き情報から前記符号化対象領域における合成動き情報を生成する動き情報生成手段と、

前記合成動き情報を用いて、前記符号化対象領域に対する予測画像を生成する予測画像生成手段と

を有し、

前記符号化対象画像の視点と前記参照視点との位置関係の変化がない、または所定の大きさ以下の場合には、前記変換行列設定手段による変換行列の設定を行わずに、前記対応位置設定手段は直前に符号化された画像で用いた前記変換行列を用いることを特徴とする映像符号化装置も提供する。

【 0 0 2 5 】

本発明はまた、複数の異なる視点の映像からなる多視点動画の符号データから、復号対象画像を復号する際に、前記復号対象画像を分割した領域である復号対象領域ごとに、異なる視点間で予測しながら復号を行う映像復号装置であって、

前記多視点映像中の被写体に対するデプスマップから代表デプスを設定する代表デプス設定手段と、前記代表デプスに基づいて、前記復号対象画像上の位置を、該復号対象画像とは異なる参照視点に対する参照画像上の位置へと変換する変換行列を設定する変換行列設定手段と、

前記復号対象領域内の位置から代表位置を設定する代表位置設定手段と、

前記代表位置と前記変換行列を用いて、前記代表位置に対する前記参照視点画像上での対応位置を設定する対応位置設定手段と、

前記対応位置に基づいて、前記参照視点画像の動き情報である参照視点動き情報から前記復号対象領域における合成動き情報を生成する動き情報生成手段と、

前記合成動き情報を用いて、前記復号対象領域に対する予測画像を生成する予測画像生成手段と、

前記復号対象領域に対して、前記デプスマップ上での対応領域であるデプス領域を設定するデプス領域設定手段と、

前記復号対象領域に対して、前記デプスマップに対する視差ベクトルであるデプス参照視差ベクトルを設定するデプス参照視差ベクトル設定手段と

を有し、

前記代表デプス設定手段は、前記デプス領域に対する前記デプスマップから代表デプスを設定し、

前記デプス領域設定手段は、前記デプス参照視差ベクトルによって示される領域を前記デプス領域として設定することを特徴とする映像復号装置も提供する。

典型的には、前記デプス参照視差ベクトル設定手段は、前記復号対象領域に隣接する領域を復号する際に使用した視差ベクトルを用いて、前記デプス参照視差ベクトルを設定する。

【 0 0 2 6 】

本発明はまた、複数の異なる視点の映像からなる多視点動画の符号データから、復号対象画像を復号する際に、前記復号対象画像を分割した領域である復号対象領域ごとに、異なる視点間で予測しながら復号を行う映像復号装置であって、

前記多視点映像中の被写体に対するデプスマップから代表デプスを設定する代表デプス設定手段と、

前記代表デプスに基づいて、前記復号対象画像上の位置を、該復号対象画像とは異なる参照視点に対する参照画像上の位置へと変換する変換行列を設定する変換行列設定手段と

、
前記復号対象領域内の位置から代表位置を設定する代表位置設定手段と、
前記代表位置と前記変換行列を用いて、前記代表位置に対する前記参照視点画像上での対応位置を設定する対応位置設定手段と、

前記対応位置に基づいて、前記参照視点画像の動き情報である参照視点動き情報から前記復号対象領域における合成動き情報を生成する動き情報生成手段と、

前記合成動き情報を用いて、前記復号対象領域に対する予測画像を生成する予測画像生成手段と、

前記変換行列を用いて、前記合成動き情報を変換する合成動き情報変換手段と
を有し、

前記予測画像生成手段は、前記変換された合成動き情報を用いることを特徴とする映像復号装置も提供する。

【0027】

本発明はまた、複数の異なる視点の映像からなる多視点動画の符号データから、復号対象画像を復号する際に、前記復号対象画像を分割した領域である復号対象領域ごとに、異なる視点間で予測しながら復号を行う映像復号装置であって、

前記多視点映像中の被写体に対するデプスマップから代表デプスを設定する代表デプス設定手段と、

前記代表デプスに基づいて、前記復号対象画像上の位置を、該復号対象画像とは異なる参照視点に対する参照画像上の位置へと変換する変換行列を設定する変換行列設定手段と

、
前記復号対象領域内の位置から代表位置を設定する代表位置設定手段と、
前記代表位置と前記変換行列を用いて、前記代表位置に対する前記参照視点画像上での対応位置を設定する対応位置設定手段と、

前記対応位置に基づいて、前記参照視点画像の動き情報である参照視点動き情報から前記復号対象領域における合成動き情報を生成する動き情報生成手段と、

前記合成動き情報を用いて、前記復号対象領域に対する予測画像を生成する予測画像生成手段と、

前記対応位置と前記合成動き情報とに基づいて、前記デプスマップから過去デプスを設定する過去デプス設定手段と、

前記過去デプスに基づいて、前記参照視点画像上の位置を前記復号対象画像上の位置へと変換する逆変換行列を設定する逆変換行列設定手段と、

前記逆変換行列を用いて、前記合成動き情報を変換する合成動き情報変換手段と
を有し、

前記予測画像生成手段は、前記変換された合成動き情報を用いることを特徴とする映像復号装置も提供する。

【0028】

本発明はまた、複数の異なる視点の映像からなる多視点動画の符号データから、復号対象画像を復号する際に、前記復号対象画像を分割した領域である復号対象領域ごとに、異なる視点間で予測しながら復号を行う映像復号装置であって、

前記多視点映像中の被写体に対するデプスマップから代表デプスを設定する代表デプス設定手段と、

前記代表デプスに基づいて、前記復号対象画像上の位置を、該復号対象画像とは異なる参照視点に対する参照画像上の位置へと変換する変換行列を設定する変換行列設定手段と

、
前記復号対象領域内の位置から代表位置を設定する代表位置設定手段と、
前記代表位置と前記変換行列を用いて、前記代表位置に対する前記参照視点画像上での対応位置を設定する対応位置設定手段と、

10

20

30

40

50

前記対応位置に基づいて、前記参照視点画像の動き情報である参照視点動き情報から前記復号対象領域における合成動き情報を生成する動き情報生成手段と、

前記合成動き情報を用いて、前記復号対象領域に対する予測画像を生成する予測画像生成手段と

を有し、

前記復号対象画像の視点と前記参照視点との位置関係の変化がない、または所定の大きさ以下の場合には、前記変換行列設定手段による変換行列の設定を行わずに、前記対応位置設定手段は直前に復号された画像で用いた前記変換行列を用いることを特徴とする映像復号装置も提供する。

【 0 0 2 9 】

10

本発明はまた、複数の異なる視点の映像からなる多視点動画の符号データから、復号対象画像を復号する際に、前記復号対象画像を分割した領域である復号対象領域ごとに、異なる視点間で予測しながら復号を行う映像復号方法であって、

前記多視点映像中の被写体に対するデプスマップから代表デプスを設定する代表デプス設定ステップと、

前記代表デプスに基づいて、前記復号対象画像上の位置を、該復号対象画像とは異なる参照視点に対する参照画像上の位置へと変換する変換行列を設定する変換行列設定ステップと、

前記復号対象領域内の位置から代表位置を設定する代表位置設定ステップと、

前記代表位置と前記変換行列を用いて、前記代表位置に対する前記参照視点画像上での対応位置を設定する対応位置設定ステップと、

20

前記対応位置に基づいて、前記参照視点画像の動き情報である参照視点動き情報から前記復号対象領域における合成動き情報を生成する動き情報生成ステップと、

前記合成動き情報を用いて、前記復号対象領域に対する予測画像を生成する予測画像生成ステップと、

前記復号対象領域に対して、前記デプスマップ上での対応領域であるデプス領域を設定するデプス領域設定ステップと、

前記復号対象領域に対して、前記デプスマップに対する視差ベクトルであるデプス参照視差ベクトルを設定するデプス参照視差ベクトル設定ステップと

を有し、

30

前記代表デプス設定ステップは、前記デプス領域に対する前記デプスマップから代表デプスを設定し、

前記デプス領域設定ステップは、前記デプス参照視差ベクトルによって示される領域を前記デプス領域として設定することを特徴とする映像復号方法も提供する。

典型的には、前記デプス参照視差ベクトル設定ステップは、前記復号対象領域に隣接する領域を復号する際に使用した視差ベクトルを用いて、前記デプス参照視差ベクトルを設定する。

【 0 0 3 0 】

本発明はまた、複数の異なる視点の映像からなる多視点動画の符号データから、復号対象画像を復号する際に、前記復号対象画像を分割した領域である復号対象領域ごとに、異なる視点間で予測しながら復号を行う映像復号方法であって、

40

前記多視点映像中の被写体に対するデプスマップから代表デプスを設定する代表デプス設定ステップと、

前記代表デプスに基づいて、前記復号対象画像上の位置を、該復号対象画像とは異なる参照視点に対する参照画像上の位置へと変換する変換行列を設定する変換行列設定ステップと、

前記復号対象領域内の位置から代表位置を設定する代表位置設定ステップと、

前記代表位置と前記変換行列を用いて、前記代表位置に対する前記参照視点画像上での対応位置を設定する対応位置設定ステップと、

前記対応位置に基づいて、前記参照視点画像の動き情報である参照視点動き情報から前

50

記復号対象領域における合成動き情報を生成する動き情報生成ステップと、

前記合成動き情報を用いて、前記復号対象領域に対する予測画像を生成する予測画像生成ステップと、

前記変換行列を用いて、前記合成動き情報を変換する合成動き情報変換ステップとを有し、

前記予測画像生成ステップは、前記変換された合成動き情報を用いることを特徴とする映像復号方法も提供する。

【0031】

本発明はまた、複数の異なる視点の映像からなる多視点動画の符号データから、復号対象画像を復号する際に、前記復号対象画像を分割した領域である復号対象領域ごとに、異なる視点間で予測しながら復号を行う映像復号方法であって、

前記多視点映像中の被写体に対するデプスマップから代表デプスを設定する代表デプス設定ステップと、

前記代表デプスに基づいて、前記復号対象画像上の位置を、該復号対象画像とは異なる参照視点に対する参照画像上の位置へと変換する変換行列を設定する変換行列設定ステップと、

前記復号対象領域内の位置から代表位置を設定する代表位置設定ステップと、

前記代表位置と前記変換行列を用いて、前記代表位置に対する前記参照視点画像上での対応位置を設定する対応位置設定ステップと、

前記対応位置に基づいて、前記参照視点画像の動き情報である参照視点動き情報から前記復号対象領域における合成動き情報を生成する動き情報生成ステップと、

前記合成動き情報を用いて、前記復号対象領域に対する予測画像を生成する予測画像生成ステップと、

前記対応位置と前記合成動き情報とに基づいて、前記デプスマップから過去デプスを設定する過去デプス設定ステップと、

前記過去デプスに基づいて、前記参照視点画像上の位置を前記復号対象画像上の位置へと変換する逆変換行列を設定する逆変換行列設定ステップと、

前記逆変換行列を用いて、前記合成動き情報を変換する合成動き情報変換ステップとを有し、

前記予測画像生成ステップは、前記変換された合成動き情報を用いることを特徴とする映像復号方法も提供する。

【0032】

本発明はまた、複数の異なる視点の映像からなる多視点動画の符号データから、復号対象画像を復号する際に、前記復号対象画像を分割した領域である復号対象領域ごとに、異なる視点間で予測しながら復号を行う映像復号方法であって、

前記多視点映像中の被写体に対するデプスマップから代表デプスを設定する代表デプス設定ステップと、

前記代表デプスに基づいて、前記復号対象画像上の位置を、該復号対象画像とは異なる参照視点に対する参照画像上の位置へと変換する変換行列を設定する変換行列設定ステップと、

前記復号対象領域内の位置から代表位置を設定する代表位置設定ステップと、

前記代表位置と前記変換行列を用いて、前記代表位置に対する前記参照視点画像上での対応位置を設定する対応位置設定ステップと、

前記対応位置に基づいて、前記参照視点画像の動き情報である参照視点動き情報から前記復号対象領域における合成動き情報を生成する動き情報生成ステップと、

前記合成動き情報を用いて、前記復号対象領域に対する予測画像を生成する予測画像生成ステップと

を有し、

前記復号対象画像の視点と前記参照視点との位置関係の変化が所定の大きさ以下の場合には、前記変換行列設定ステップによる変換行列の設定を行わずに、前記対応位置設定ス

10

20

30

40

50

テップは直前に復号された画像で用いた前記変換行列を用いることを特徴とする映像復号方法も提供する。

【 0 0 3 3 】

本発明はまた、複数の異なる視点の映像からなる多視点映像の1フレームである符号化対象画像を符号化する際に、前記符号化対象画像を分割した領域である符号化対象領域ごとに、異なる視点間で予測しながら符号化を行う映像符号化方法であって、

前記多視点映像中の被写体に対するデプスマップから代表デプスを設定する代表デプス設定ステップと、

前記代表デプスに基づいて、前記符号化対象画像上の位置を、該符号化対象画像とは異なる参照視点に対する参照視点画像上の位置へと変換する変換行列を設定する変換行列設定ステップと、

前記符号化対象領域内の位置から代表位置を設定する代表位置設定ステップと、

前記代表位置と前記変換行列を用いて、前記代表位置に対する前記参照視点画像上での対応位置を設定する対応位置設定ステップと、

前記対応位置に基づいて、前記参照視点画像の動き情報である参照視点動き情報から前記符号化対象領域における合成動き情報を生成する動き情報生成ステップと、

前記合成動き情報を用いて、前記符号化対象領域に対する予測画像を生成する予測画像生成ステップと、

前記符号化対象領域に対して、前記デプスマップ上での対応領域であるデプス領域を設定するデプス領域設定ステップと、

前記符号化対象領域に対して、前記デプスマップに対する視差ベクトルであるデプス参照視差ベクトルを設定するデプス参照視差ベクトル設定ステップと

を有し、

前記代表デプス設定ステップは、前記デプス領域に対する前記デプスマップから代表デプスを設定し、

前記デプス領域設定ステップは、前記デプス参照視差ベクトルによって示される領域を前記デプス領域として設定することを特徴とする映像符号化方法も提供する。

【 0 0 3 4 】

本発明はまた、コンピュータに、前記映像復号方法を実行させるための映像復号プログラムも提供する。

本発明はまた、コンピュータに、前記映像符号化方法を実行させるための映像符号化プログラムも提供する。

【発明の効果】

【 0 0 3 5 】

本発明によれば、複数の視点に対する映像がその映像に対するデプスマップと共に符号化または復号される場合に、視点間の画素の対応関係をデプス値に対して定義される1つの行列を用いて求めることで、視点の向きが平行でない場合でも、複雑な演算を行うことなく、動きベクトルの視点間予測の精度を向上させることが可能となり、少ない符号量で映像を符号化することができるという効果が得られる。

【図面の簡単な説明】

【 0 0 3 6 】

【図1】本発明の一実施形態による映像符号化装置の構成を示すブロック図である。

【図2】図1に示す映像符号化装置100の動作を示すフローチャートである。

【図3】図2に示す動き情報生成部105における動き情報を生成する動作（ステップS104）の処理動作を示すフローチャートである。

【図4】本発明の一実施形態による映像復号装置の構成を示すブロック図である。

【図5】図4に示す映像復号装置200の動作を示すフローチャートである。

【図6】図1に示す映像符号化装置100をコンピュータとソフトウェアプログラムとによって構成する場合のハードウェア構成を示すブロック図である。

【図7】図4に示す映像復号装置200をコンピュータとソフトウェアプログラムとによ

10

20

30

40

50

って構成する場合のハードウェア構成を示すブロック図である。

【発明を実施するための形態】

【0037】

以下、図面を参照して、本発明の実施形態による映像符号化装置及び映像復号装置を説明する。

以下の説明においては、第1のカメラ（カメラAという）、第2のカメラ（カメラBという）の2つのカメラで撮影された多視点映像を符号化する場合を想定し、カメラAを参照視点としてカメラBの映像の1フレームを符号化または復号するものとして説明する。

なお、デプスから視差を得るために必要となる情報は、別途与えられているものとする。具体的には、カメラAとカメラBの位置関係を表す外部パラメータや、カメラによる画像平面への投影情報を表す内部パラメータである、これらと同じ意味をもつものであれば、別の形式で必要な情報が与えられていてもよい。

これらのカメラパラメータに関する詳しい説明は、例えば、文献「Oliver Faugeras, "Three-Dimension Computer Vision", MIT Press; BCTC/UFF-006.37 F259 1993, ISBN:0-262-06158-9.」に記載されている。この文献には、複数のカメラの位置関係を示すパラメータや、カメラによる画像平面への投影情報を表すパラメータに関する説明が記載されている。

【0038】

以下の説明では、画像や映像フレーム、デプスマップに対して、位置を特定可能な情報（座標値もしくは座標値に対応付け可能なインデックスなど）を付加する（例えば後述の符号化対象領域インデックスblock）ことで、その位置（範囲）の画素によってサンプリングされた画像信号や、それに対するデプスを示すものとする。

また、座標値やブロックに対応付け可能なインデックス値とベクトルとの加算によって、その座標やブロックをベクトルの分だけずらした位置の座標値やブロックを表すものとする。

【0039】

図1は本実施形態による映像符号化装置の構成を示すブロック図である。

映像符号化装置100は、図1に示すように、符号化対象画像入力部101、符号化対象画像メモリ102、参照視点動き情報入力部103、デプスマップ入力部104、動き情報生成部105、画像符号化部106、画像復号部107及び参照画像メモリ108を備えている。

【0040】

符号化対象画像入力部101は、符号化対象となる映像の1フレームを映像符号化装置100に入力する。以下では、この符号化対象となる映像および入力され符号化されるフレームを、それぞれ、符号化対象映像および符号化対象画像と称する。ここではカメラBの映像を1フレームずつ入力するものとする。また、符号化対象映像を撮影した視点（ここではカメラBの視点）を符号化対象視点と称する。

符号化対象画像メモリ102は、入力した符号化対象画像を記憶する。

参照視点動き情報入力部103は、参照視点の映像に対する動き情報（動きベクトルなど）を映像符号化装置100に入力する。以下では、ここで入力された動き情報を、参照視点動き情報と呼ぶ。ここではカメラAの動き情報を入力するものとする。

【0041】

デプスマップ入力部104は、視点間の画素の対応関係を求めたり、動き情報を生成したりする際に参照するデプスマップを映像符号化装置100に入力する。ここでは、符号化対象画像に対するデプスマップを入力するものとするが、参照視点など別の視点に対するデプスマップでも構わない。

なお、デプスマップとは、対応する画像の各画素に写っている被写体の3次元位置を表すものである。例えば、カメラから被写体までの距離や、画像平面とは平行ではない軸に対する座標値、別のカメラ（例えばカメラA）に対する視差量を用いることができる。

なお、ここではデプスマップとして画像の形態で提供されるものとしているが、同様の

10

20

30

40

50

情報が得られるのであれば、画像の形態でなくとも構わない。

【 0 0 4 2 】

動き情報生成部 1 0 5 は、参照視点動き情報とデプスマップとを用いて、符号化対象画像に対する動き情報を生成する。

画像符号化部 1 0 6 は、生成された動き情報を用いながら、符号化対象画像を予測符号化する。

画像復号部 1 0 7 は、符号化対象画像のビットストリームを復号する。

参照画像メモリ 1 0 8 は、符号化対象画像のビットストリームを復号した際に得られる画像を記憶する。

【 0 0 4 3 】

次に、図 2 を参照して、図 1 に示す映像符号化装置 1 0 0 の動作を説明する。図 2 は、図 1 に示す映像符号化装置 1 0 0 の動作を示すフローチャートである。

まず、符号化対象画像入力部 1 0 1 は、符号化対象画像 *Org* を入力し、符号化対象画像メモリ 1 0 2 に記憶する（ステップ *S 1 0 1*）。

次に、参照視点動き情報入力部 1 0 3 は参照視点動き情報を映像符号化装置 1 0 0 に入力し、デプスマップ入力部 1 0 4 はデプスマップを映像符号化装置 1 0 0 に入力し、それぞれ動き情報生成部 1 0 5 へ出力される（ステップ *S 1 0 2*）。

【 0 0 4 4 】

なお、ステップ *S 1 0 2* で入力される参照視点動き情報とデプスマップは、既に符号化済みのものを復号したものなど、復号側で得られるものと同じものとする。これは復号装置で得られるものと全く同じ情報を用いることで、ドリフト等の符号化ノイズの発生を抑えるためである。ただし、そのような符号化ノイズの発生を許容する場合には、符号化前のものなど、符号化側でしか得られないものが入力されてもよい。

デプスマップに関しては、既に符号化済みのものを復号したもの以外に、複数のカメラに対して復号された多視点映像に対してステレオマッチング等を適用することで推定したデプスマップや、復号された視差ベクトルや動きベクトルなどを用いて推定されるデプスマップなども、復号側で同じものが得られるものとして用いることができる。

【 0 0 4 5 】

参照視点動き情報は、参照視点に対する映像を符号化する際に使用された動き情報を用いても構わないし、参照視点に対して別途符号化されたものでも構わない。また、参照視点に対する映像を復号し、そこから推定して得られた動き情報を用いることも可能である。

【 0 0 4 6 】

符号化対象画像、参照視点動き情報、デプスマップの入力が終了したら、符号化対象画像を予め定められた大きさの領域に分割し、分割した領域ごとに、符号化対象画像の映像信号を符号化する（ステップ *S 1 0 3* ~ *S 1 0 8*）。

すなわち、符号化対象領域インデックスを *blk*、1 フレーム中の総符号化対象領域数を *numBlks* で表すとすると、*blk* を 0 で初期化し（ステップ *S 1 0 3*）、その後、*blk* に 1 を加算しながら（ステップ *S 1 0 7*）、*blk* が *numBlks* になるまで（ステップ *S 1 0 8*）、以下の処理（ステップ *S 1 0 4* ~ *S 1 0 6*）を繰り返す。

一般的な符号化では 16 画素 × 16 画素のマクロブロックと呼ばれる処理単位ブロックへ分割するが、復号側と同じであればその他の大きさのブロックに分割しても構わない。また、画像全体を同じサイズで分割せず、領域ごとに異なるサイズのブロックに分割しても構わない。

【 0 0 4 7 】

符号化対象領域ごとに繰り返される処理では、まず、動き情報生成部 1 0 5 は、符号化対象領域 *blk* における動き情報を生成する（ステップ *S 1 0 4*）。ここでの処理は後で詳しく説明する。

符号化対象領域 *blk* に対する動き情報が得られたら、画像符号化部 1 0 6 は、その動き情報と参照画像メモリ 1 0 8 に記憶された画像とを用いて動き補償予測を行いながら、

10

20

30

40

50

符号化対象領域 $b1k$ における符号化対象画像の映像信号（画素値）を符号化する（ステップ 105）。符号化の結果得られるビットストリームが映像符号化装置 100 の出力となる。なお、符号化する方法には、どのような方法を用いても構わない。

MPEG-2 や H.264 / AVC などの一般的な符号化では、ブロック $b1k$ の映像信号と予測画像との差分信号に対して、DCT などの周波数変換、量子化、2 値化、エントロピー符号化を順に施すことで符号化を行う。

【0048】

次に、画像復号部 107 は、ビットストリームからブロック $b1k$ に対する映像信号を復号し、復号結果であるところの復号画像 $Dec[b1k]$ を参照画像メモリ 109 に記憶する（ステップ S106）。

ここでは、符号化時に用いた手法に対応する手法を用いる。例えば、MPEG-2 や H.264 / AVC などの一般的な符号化であれば、符号データに対して、エントロピー復号、逆 2 値化、逆量子化、IDCT などの周波数逆変換を順に施し、得られた 2 次元信号に対して予測画像を加え、最後に画素値の値域でクリッピングを行うことで映像信号を復号する。

なお、符号化側での処理がロスレスになる直前のデータと予測画像を受け取り、簡略化した復号処理によって復号処理を行っても構わない。

【0049】

すなわち、前述の例であれば、符号化時に量子化処理を加えた後の値と動き補償予測画像とを受け取り、その量子化後の値に逆量子化、周波数逆変換を順に施して得られた 2 次元信号に対して動き補償予測画像を加え、画素値の値域でクリッピングを行うことで映像信号を復号しても構わない。

【0050】

次に、図 3 を参照して、動き情報生成部 105 が行う符号化対象領域 $b1k$ における動き情報を生成する処理（ステップ S104）について詳細に説明する。図 3 は、図 2 に示す動き情報生成部 105 における動き情報を生成する動作（ステップ S104）の処理動作を示すフローチャートである。

【0051】

動き情報を生成する処理において、まず、動き情報生成部 105 は、符号化対象領域 $b1k$ に対するデプスマップを設定する（ステップ S1401）。ここでは、符号化対象画像に対するデプスマップが入力されているため、符号化対象領域 $b1k$ と同じ位置のデプスマップを設定することとなる。

なお、符号化対象画像とデプスマップの解像度が異なる場合は、解像度比に応じてスケールリングした領域を設定する。符号化対象視点と異なる視点の 1 つをデプス視点とすると、デプス視点に対するデプスマップを用いる場合は、符号化対象領域 $b1k$ における符号化対象視点とデプス視点の視差 DV を求め、 $b1k + DV$ におけるデプスマップを設定する。符号化対象画像とデプスマップの解像度が異なる場合は、上述のように、解像度比に応じて位置および大きさのスケールリングを行う。

【0052】

符号化対象領域 $b1k$ における符号化対象視点とデプス視点の視差 DV は、復号側と同じ方法であればどのような方法を用いて算出しても構わない。

例えば、符号化対象領域 $b1k$ の周辺領域を符号化する際に使用された視差ベクトルや、符号化対象画像全体や符号化対象領域を含む部分画像に対して設定されたグローバル視差ベクトル、符号化対象領域に対して別途設定し符号化される視差ベクトルなどを用いることが可能である。また、異なる領域や過去に符号化された画像で使用した視差ベクトルを記憶しておき、用いても構わない。

更に、符号化対象視点に対して過去に符号化されたデプスマップの符号化対象領域と同位置のデプスマップを変換して得られる視差ベクトルを用いても構わない。

【0053】

次に、動き情報生成部 105 は、設定されたデプスマップから、（本発明の「代表位置

10

20

30

40

50

」としての) 代表画素位置 p_{os} と代表デプス r_{ep} を決定する (ステップ S 1 4 0 2) 。どのような方法を用いて代表画素位置と代表デプスを決定しても構わないが、復号側と同じ方法を用いる必要がある。

代表画素位置 p_{os} を設定する代表的な方法としては、代表画素位置として符号化対象領域内の中央や左上など予め定められた位置を設定する方法や、代表デプスを求めた後に、その代表デプスと同じデプスを持つ符号化対象領域内の画素の位置を設定する方法がある。

【 0 0 5 4 】

また、別の方法として、予め定められた位置の画素に対するデプスを比較して、予め定められた条件を満たすデプスを持つ画素の位置を設定する方法がある。

具体的には、符号化対象領域内の中央に位置する 4 つの画素や、(四角形状の符号化対象領域の) 4 頂点に位置する画素、4 頂点と中央に位置する画素を対象とし、最大のデプスや、最小のデプス、中央値のデプスなどを与える画素を選択する方法である。

代表デプス r_{ep} を設定する代表的な方法としては、符号化対象領域 b_{lk} に対するデプスマップの平均値や中央値、最大値、最小値などを用いる方法がある。

また、符号化対象領域内の全ての画素ではなく、一部の画素に対するデプス値の平均値や中央値、最大値、最小値などを用いても構わない。一部の画素としては、4 頂点や 4 頂点と中央などを用いても構わない。更に、符号化対象領域に対して、左上や中央など予め定められた位置に対するデプス値を用いる方法もある。

【 0 0 5 5 】

動き情報生成部 1 0 5 は、代表画素位置 p_{os} および代表デプスが得られたら、次に変換行列 H_{rep} を求める (ステップ S 1 4 0 3) 。

ここで、変換行列はホモグラフィ行列と呼ばれ、代表デプスで表現される平面に被写体が存在すると仮定したときに、視点間での画像平面上の点の対応関係を与えるものである。なお、変換行列 H_{rep} はどのように求めても構わない。例えば、次の数式を用いて求めることが可能である。

【 数 1 】

$$H_{rep} = R + \frac{t \, n(D_{rep})^T}{d(D_{rep})}$$

【 0 0 5 6 】

なお、 R と t は、符号化対象視点と参照視点との間の 3×3 回転行列と並進ベクトルをそれぞれ表し、 D_{rep} は代表デプス、 $n(D_{rep})$ は符号対象視点における代表デプス D_{rep} に対応する三次元平面の法線ベクトルを示し、 $d(D_{rep})$ はその三次元平面と、符号化対象視点と参照視点の視点中心との間の距離を示す。また、右肩の T はベクトルの転置を表す。

【 0 0 5 7 】

変換行列 H_{rep} の別の求め方としては、まず、符号化対象画像中の異なる 4 点 p_i ($i = 1, 2, 3, 4$) に対して、次の式に基づいて、参照視点の画像上の対応点 q_i を求める。

10

20

30

40

【数 2】

$$P_r \left(P_t^{-1} \left(d_t(p_i) \begin{pmatrix} p_i \\ 1 \end{pmatrix} \right) \right) = s \begin{pmatrix} q_i \\ 1 \end{pmatrix}$$

ここで、 P_t および P_r は、それぞれ符号化対象視点および参照視点における 3×4 カメラ行列を示す。ここでのカメラ行列は、カメラの内部パラメータを A 、世界座標系（カメラに依存しない任意の共通な座標系）からカメラ座標系への回転行列を R 、世界座標系からカメラ座標系への並進を表す列ベクトルを t で表すと、 $A[R | t]$ で与えられる（ $[R | t]$ は R と t を並べて作られる 3×4 行列であり、カメラの外部パラメータと呼ばれる）。なお、ここでのカメラ行列 P の逆行列 P^{-1} は、カメラ行列 P による変換の逆変換に対応する行列であるとし、 $R^{-1}[A^{-1} | -t]$ で表される。

$d_t(p_i)$ は、符号化対象画像上の点 p_i におけるデプスが代表デプスであるとしたときの、符号化対象視点から点 p_i における被写体までの光軸上の距離を示す。

s は任意の実数であるが、カメラパラメータの誤差がない場合、 s は参照視点の画像上の点 q_i における参照視点から点 q_i における被写体までの光軸上の距離 $d_r(q_i)$ と等しい。

また、上記定義に従い式 2 を計算すると、次の数式となる。なお、内部パラメータ A 、回転行列 R 、並進ベクトル t の添え字 t と r は各カメラを表し、それぞれ符号化対象視点と参照視点を示す。

【数 3】

$$A_r \left(R_r R_t^{-1} \left(A_t^{-1} d_t(p_i) \begin{pmatrix} p_i \\ 1 \end{pmatrix} - t_t \right) + t_r \right) = s \begin{pmatrix} q_i \\ 1 \end{pmatrix}$$

【0058】

4 つの対応点が求まったら、次の式に従って得られる同次方程式を解くことで変換行列 H_{rep} を得る。ただし、変換行列 H_{rep} の $(3, 3)$ 成分は任意の実数（例えば 1）を設定して求める。

【数 4】

$$\begin{bmatrix} \tilde{p}_i^T & 0^T & -q_{i,1} \tilde{p}_i^T \\ 0^T & \tilde{p}_i^T & -q_{i,2} \tilde{p}_i^T \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix} = 0$$

$$\tilde{p}_i = \begin{pmatrix} p_i \\ 1 \end{pmatrix}, q_i = \begin{pmatrix} q_{i,1} \\ q_{i,2} \end{pmatrix}, H_{rep} = \begin{bmatrix} h_1^T \\ h_2^T \\ h_3^T \end{bmatrix}$$

【0059】

変換行列 H_{rep} は参照視点とデプスに依存することから、代表デプスを求める度に毎

10

20

30

40

50

回求めても構わないし、領域ごとの処理を開始する前に、参照視点とデプスの組み合わせ毎に求めておき、変換行列 H_{rep} を求める段階で、既に計算してある変換行列群の中から、参照視点及び代表デプスをもとに、1つの変換行列を選択・設定しても構わない。

【0060】

代表デプスに対する変換行列が得られたら、動き情報生成部105は、次の数式に基づいて参照視点上の対応位置を求める(ステップS1404)。

【数5】

$$k \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = H_{rep} \begin{pmatrix} pos \\ 1 \end{pmatrix}$$

10

ここで、 k は任意の実数を表し、 (u, v) で与えられる位置が、求める参照視点上の位置である。

【0061】

次に、参照視点における対応位置が得られたら、動き情報生成部105は、その位置を含む領域に対して入力されて記憶されている参照視点動き情報を、符号化対象領域 blk に対する動き情報として設定する(ステップS1405)。

なお、対応位置 (u, v) を含む領域に対して参照視点動き情報が記憶されていない場合は、動き情報なしの情報を設定しても、ゼロベクトルなどデフォルトの動き情報を設定しても、対応位置 (u, v) に最も近い動き情報を記憶している領域を同定して、その領域において記憶されている参照視点動き情報を設定しても構わない。ただし、復号側と同じ規則で動き情報を設定する。

20

【0062】

前述した説明では、参照視点動き情報をそのまま動き情報として設定したが、時間間隔を予め設定し、動き情報を、その予め定められた時間間隔と参照視点動き情報における時間間隔に従ってスケーリングし、参照視点動き情報における時間間隔をその予め定められた時間間隔に置き換えて得られる動き情報を設定しても構わない。

このようにすることで、異なる領域に対して生成される動き情報が全て同じ時間間隔を持つことになり、動き補償予測を行う際の参照画像を統一し、アクセスするメモリ空間を限定することが可能となる。なお、アクセスするメモリ空間が限定されることによって、キャッシュメモリのヒット率を向上させ、処理速度を向上することが可能となる。

30

【0063】

また、前述した説明では、参照視点動き情報をそのまま動き情報として設定したが、変換行列 H_{rep} を用いて変換したものを設定しても構わない。

すなわち、ステップS1405において設定された動き情報を $mv = (mv_x, mv_y)^T$ とすると、変換した動き情報 mv' は次の数式で表される。

【数6】

$$s \begin{bmatrix} p' \\ 1 \end{bmatrix} = H_{rep}^{-1} \begin{bmatrix} u + mv_x \\ v + mv_y \\ 1 \end{bmatrix}$$

$$mv' = p' - pos$$

40

ここで、 s は任意の実数を表す。

【0064】

さらに、ステップS1405において設定された動き情報の示す時間間隔に対応する参

50

照視点におけるデプスマップを参照でき、位置 $(u + mv_x, v + mv_y)$ におけるデプスを $prdep$ であるとする、次の式に基づいて求めた p' を用いて mv' を求めても構わない。

【数 7】

$$s \begin{bmatrix} p' \\ 1 \end{bmatrix} = H_{d_r \rightarrow t}^{-1}(prdep) \begin{bmatrix} u + mv_x \\ v + mv_y \\ 1 \end{bmatrix}$$

10

ここで $d_{r \rightarrow t}(prdep)$ は、参照視点に対して表現されたデプス $prdep$ を符号化対象視点に対する表現のデプスへと変換する関数である。

符号化対象視点と参照視点とで共通する軸を用いてデプスを表現している場合、この変換は、引数で与えられたデプスをそのまま返す。

【0065】

なお、ここでは符号化対象視点に対する位置から参照視点に対する位置へと変換する変換行列 H の逆変換行列 H^{-1} を用いているが、変換行列から逆行列を計算して得ても構わないし、逆変換行列を直接求めても構わない。

直接計算する場合、まず、参照視点に対する画像中の異なる 4 点 q'_i ($i = 1, 2, 3, 4$) に対して、次の式に基づいて、符号化対象視点の画像上の対応点 p'_i を求める。

20

【数 8】

$$s \begin{pmatrix} p'_i \\ 1 \end{pmatrix} = P_t \left(P_r^{-1} \begin{pmatrix} d_{r,prdep}(q'_i) \begin{pmatrix} q'_i \\ 1 \end{pmatrix} \\ 1 \end{pmatrix} \right)$$

30

ここで、 $d_{r,prdep}(q'_i)$ は、視点 r の画像上の点 q'_i における視点 r に対して定義されたデプスを $prdep$ としたときの、視点 r から点 q'_i における被写体までの光軸上の距離を示す。

【0066】

4 つの対応点が求まったら、次の数式に従って得られる同次方程式を解くことで、逆変換行列 H' を得る。ただし、変換行列 H' の $(3, 3)$ 成分は任意の実数（例えば 1）を設定して求める。

【数 9】

$$\begin{bmatrix} \tilde{q}_i'^T & 0^T & -p'_{i,1} \tilde{q}_i'^T \\ 0^T & \tilde{q}_i'^T & -p'_{i,2} \tilde{q}_i'^T \end{bmatrix} \begin{bmatrix} h'_1 \\ h'_2 \\ h'_3 \end{bmatrix} = 0$$

40

$$\tilde{q}_i' = \begin{pmatrix} q'_i \\ 1 \end{pmatrix}, p'_i = \begin{pmatrix} p'_{i,1} \\ p'_{i,2} \end{pmatrix}, H'_{prdep} = H_{d_r \rightarrow t}^{-1}(prdep) = \begin{bmatrix} h_1'^T \\ h_2'^T \\ h_3'^T \end{bmatrix}$$

50

【 0 0 6 7 】

また、ステップ S 1 4 0 5 において設定された動き情報の示す時間間隔に対応する、符号化視点におけるデプスマップ $D_{t, Ref(blk)}$ を参照できる場合、次の数式で変換後の動き情報 mv'_{depth} を求めても構わない。

【 数 1 0 】

$$mv'_{depth} = p'_{depth} - pos$$

s.t.

$$depth = \arg \min_{pd} \| pd - D_{t, Ref(blk)} [p'_{pd}] \| \quad 10$$

$$s \begin{bmatrix} p'_{pd} \\ 1 \end{bmatrix} = H_{pd}^{-1} \begin{bmatrix} u + mv_x \\ v + mv_y \\ 1 \end{bmatrix}$$

ここで $\| \cdot \|$ はノルムを示し、L 1 ノルムを用いても構わないし、L 2 ノルムを用いても構わない。

【 0 0 6 8 】

上記説明した変換とスケーリングを、同時に施しても構わない。その場合、スケーリングした後に変換しても、変換したあとにスケーリングしても構わない。 20

【 0 0 6 9 】

前述した説明で用いた動き情報は、符号化対象視点の位置に対して加算することで、時間方向の対応位置を示すものとして表現している。もし減算することで対応位置を表す場合、上記説明で用いた数式における動き情報では、ベクトルの向きを逆転させる必要がある。

【 0 0 7 0 】

次に、本実施形態による映像復号装置について説明する。

図 4 は本実施形態による映像復号装置の構成を示すブロック図である。映像復号装置 2 0 0 は、図 4 に示すように、ビットストリーム入力部 2 0 1、ビットストリームメモリ 2 0 2、参照視点動き情報入力部 2 0 3、デプスマップ入力部 2 0 4、動き情報生成部 2 0 5、画像復号部 2 0 6 及び参照画像メモリ 2 0 7 を備えている。 30

【 0 0 7 1 】

ビットストリーム入力部 2 0 1 は、復号対象となる映像のビットストリームを映像復号装置 2 0 0 に入力する。以下では、この復号対象となる映像の 1 フレームを復号対象画像と呼ぶ。ここではカメラ B の映像の 1 フレームを指す。また、以下では、復号対象画像を撮影した視点（ここではカメラ B）を復号対象視点と呼ぶ。

ビットストリームメモリ 2 0 2 は、入力した復号対象画像に対するビットストリームを記憶する。

参照視点動き情報入力部 2 0 3 は、参照視点の映像に対する動き情報（動きベクトルなど）を映像復号装置 2 0 0 に入力する。以下では、ここで入力された動き情報を、参照視点動き情報と呼ぶ。ここではカメラ A の動き情報が入力されるものとする。 40

【 0 0 7 2 】

デプスマップ入力部 2 0 4 は、視点間の画素の対応関係を求めたり、復号対象画像に対する動き情報を生成したりする際に参照するデプスマップを映像復号装置 2 0 0 に入力する。ここでは、復号対象画像に対するデプスマップを入力するものとするが、参照視点など別の視点に対するデプスマップでも構わない。

なお、デプスマップとは、対応する画像の各画素に写っている被写体の 3 次元位置を表すものである。例えば、カメラから被写体までの距離や、画像平面とは平行ではない軸に対する座標値、別のカメラ（例えばカメラ A）に対する視差量を用いることができる。 50

なお、ここではデプスマップとして画像の形態で提供されるものとしているが、同様の情報が得られるのであれば、画像の形態でなくても構わない。

【0073】

動き情報生成部205は、参照視点動き情報とデプスマップとを用いて、復号対象画像に対する動き情報を生成する。

画像復号部206は、生成された動き情報を用いながら、上記ビットストリームから復号対象画像を復号して出力する。

参照画像メモリ207は、得られた復号対象画像を、以降の復号のために記憶する。

【0074】

次に、図5を参照して、図4に示す映像復号装置200の動作を説明する。図5は、図4に示す映像復号装置200の動作を示すフローチャートである。

まず、ビットストリーム入力部201は、復号対象画像を符号化したビットストリームを映像復号装置200に入力し、ビットストリームメモリ202に記憶する(ステップS201)。

次に、参照視点動き情報入力部203は参照視点具置き情報を映像復号装置200に入力し、デプスマップ入力部204はデプスマップを映像復号装置200に入力し、それぞれ動き情報生成部205へ出力される(ステップS202)。

【0075】

なお、ステップS202で入力される参照視点動き情報とデプスマップは、符号化側で使用されたものと同じものとする。これは符号化時に用いたものと全く同じ情報を用いることで、ドリフト等の符号化ノイズの発生を抑えるためである。ただし、そのような符号化ノイズの発生を許容する場合には、符号化時に使用されたものと異なるものが入力されてもよい。

デプスマップに関しては、別途復号したもの以外に、複数のカメラに対して復号された多視点映像に対してステレオマッチング等を適用することで推定したデプスマップや、復号された視差ベクトルや動きベクトルなどを用いて推定されるデプスマップなどを用いることもある。

【0076】

参照視点動き情報は、参照視点に対する映像を復号する際に使用された動き情報を用いても構わないし、参照視点に対して別途符号化されたものでも構わない。また、参照視点に対する映像を復号し、そこから推定して得られた動き情報を用いることも可能である。

【0077】

ビットストリーム、参照視点動き情報、デプスマップの入力が終了したら、復号対象画像を予め定められた大きさの領域に分割し、分割した領域ごとに、復号対象画像の映像信号をビットストリームから復号する(ステップS203~S207)。

すなわち、復号対象領域インデックスを $b1k$ 、1フレーム中の総復号対象領域数を $numB1ks$ で表すとすると、 $b1k$ を0で初期化し(ステップS203)、その後、 $b1k$ に1を加算しながら(ステップS206)、 $b1k$ が $numB1ks$ になるまで(ステップS207)、以下の処理(ステップS204~S205)を繰り返す。

一般的な復号では16画素×16画素のマクロブロックと呼ばれる処理単位ブロックへ分割するが、符号化側と同じであればその他の大きさのブロックに分割しても構わない。また、画像全体を同じサイズで分割せず、領域ごとに異なるサイズのブロックに分割しても構わない。

【0078】

復号対象領域ごとに繰り返される処理では、まず、動き情報生成部205は、復号対象領域 $b1k$ における動き情報を生成する(ステップS204)。ここでの処理は、符号化対象領域が復号対象領域となるだけで、前述したステップS104の処理と同じである。

【0079】

次に、復号対象領域 $b1k$ に対する動き情報が得られたら、画像復号部206は、その動き情報と参照画像メモリ207に記憶された画像とを用いて動き補償予測を行いながら

10

20

30

40

50

、復号対象領域 $b1k$ における映像信号（画素値）をビットストリームから復号する（ステップ $S205$ ）。得られた復号対象画像は参照画像メモリ 207 に記憶されると共に、映像復号装置 200 の出力となる。

【0080】

映像信号の復号には符号化時に用いられた方法に対応する方法を用いる。

例えば、MPEG-2やH.264/AVCなどの一般的な符号化が用いられている場合は、ビットストリームに対して、エントロピー復号、逆2値化、逆量子化、IDCTなどの周波数逆変換を順に施し、得られた2次元信号に対して予測画像を加え、最後に画素値の値域でクリッピングを行うことで映像信号を復号する。

【0081】

前述した説明では、符号化対象画像または復号対象画像を分割した領域ごとに動き情報の生成を行ったが、事前に全ての領域に対してそれぞれ動き情報を生成し記憶しておき、領域ごとに記憶された動き情報を参照するようにしても構わない。

【0082】

また、画像全体を符号化/復号する処理として書かれているが、画像の一部分のみに適用することも可能である。

この場合、処理を適用するか否かを判断して、それを示すフラグを符号化/復号しても構わないし、なんらかの別の手段でそれを指定しても構わない。例えば、領域ごとの予測画像を生成する手法を示すモードの1つとして、処理を適用するか否かを表現するようにしても構わない。

【0083】

また前述した説明では、変換行列を常に生成している。しかしながら、符号化対象視点または復号対象視点と参照視点との位置関係やデプスの定義（すなわち、各デプスに対応する三次元平面）が変化しない限りは、変換行列は変化しないため、予め変換行列の集合を求めておくようにしても良く、この場合において、フレーム毎や領域ごとに変換行列を計算し直す必要はない。

すなわち、符号化対象画像または復号対象画像が変わるごとに、別途与えられるカメラパラメータによって表される符号化対象視点または復号対象視点と参照視点との位置関係と、直前のフレームにおけるカメラパラメータによって表される符号化対象視点または復号対象視点と参照視点との位置関係とを比較し、位置関係の変化がない又は小さいときには、直前のフレームで使用した変換行列の集合をそのまま使い、それ以外の場合にのみ変換行列の集合を求めるようにしても構わない。

なお、変換行列の集合を求める際に、全ての変換行列を求め直すのではなく、直前のフレームと位置関係の異なる参照視点に対するものと、定義の変化したデプスに対するものを同定し、それらに対してだけ求め直しても構わない。

【0084】

なお、符号化側でのみ変換行列の再計算が必要か否かをチェックし、その結果を符号化して伝送するようにしても構わない。この場合、復号側では伝送されてきた情報をもとに変換行列を再計算するか否かを決定するようにしても構わない。

再計算が必要か否かを示す情報は、フレーム全体に対して1つだけ設定しても構わないし、参照視点ごとに設定しても構わないし、デプスごとに設定しても構わない。

【0085】

さらに、前述した説明では、代表デプスのデプス値ごとに変換行列を生成しているが、別途定められたデプス値の範囲ごとに1つのデプス値を量子化デプスとして設定し、その量子化デプス値ごとに変換行列を設定しても構わない。代表デプスはデプスの値域の任意のデプス値を取りえるため、全てのデプス値に対する変換行列が必要となることがあるが、このようにすることで、変換行列が必要となるデプス値は量子化デプスと同じデプス値だけに制限することができる。なお、代表デプスを求めた後に変換行列を求める際には、その代表デプスが含まれるデプス値の区分から量子化デプスを求め、その量子化デプスを用いて変換行列を求める。特に、デプスの値域全体に対して1つの量子化デプスを設定す

10

20

30

40

50

る場合、変換行列は参照視点に対して唯一となる。

【 0 0 8 6 】

なお、復号側と同じ方法であれば、量子化デプスを設定するデプス値の範囲や各範囲における量子化デプスに対するデプス値はどのように設定しても構わない。例えば、デプスマップにおけるデプスの分布に従って決定しても構わない。このとき、デプスマップに対応する映像の動きを調べ、一定以上の動きが存在する領域に対するデプスのみを対象としてデプス値の分布を調べる対象としても構わない。このようにすることで動きが大きな場合に視点間で動き情報を共有できるようになり、より多くの符号量を削減することが可能となる。

【 0 0 8 7 】

また、復号側で設定できない方法で量子化デプスを決定する場合は、符号化側では、決定した量子化方法（各量子化デプスに対応するデプス値の範囲と量子化デプスのデプス値などを決定するための方法）を符号化して伝送し、復号側では符号化されたビットストリームから量子化方法を復号して得るようにしても構わない。なお、特に全体に対して1つの量子化デプスを設定する場合などは、量子化方法の代わりに量子化デプスの値を符号化または復号するようにしても構わない。

【 0 0 8 8 】

また、前述した説明では、カメラパラメータ等を用いて復号側でも変換行列を生成しているが、符号化側で計算して得られた変換行列を符号化して伝送するようにしても構わない。その場合、復号側では変換行列をカメラパラメータ等から生成せず、符号化ビットストリームから復号することで獲得する。

【 0 0 8 9 】

さらに、前述した説明では、常に変換行列を用いるものとしているが、カメラパラメータをチェックし、視点間が平行であれば（入出力間の変換用の）ルックアップテーブルを生成し、そのルックアップテーブルに従ってデプスと視差ベクトルの変換を行い、視点間が平行でなければ本願発明の手法を用いるようにしても構わない。

また、符号化側のみでチェックを行い、どちらの手法を用いるかを示す情報を符号化しても構わない。その場合、復号側ではその情報を復号し、どちらの手法を用いるかを決定する。

【 0 0 9 0 】

また、前述した説明では変換行列としてホモグラフィ行列を用いたが、符号化対象画像または復号対象画像の画素位置を、参照視点における対応画素位置へ変換することができるものであれば、別の行列を用いても構わない。例えば、厳密なホモグラフィ行列ではなく、簡略化させた行列を用いても構わない。また、アフィン変換行列や射影行列、複数の変換行列を組み合わせて生成される行列などを用いても構わない。

別の変換行列を用いることで、変換の精度や演算量、変換行列の更新頻度、変換行列を伝送する場合の符号量などを適宜制御することが可能である。なお、符号化ノイズの発生を防ぐためには、符号化時と復号時とで同じ変換行列を使用するようにする。

【 0 0 9 1 】

図6は、図1に示す映像符号化装置100をコンピュータとソフトウェアプログラムとによって構成する場合のハードウェア構成を示すブロック図である。

図6に示すシステムは：

- ・プログラムを実行するCPU50
- ・CPU50がアクセスするプログラムやデータが格納されるRAM等のメモリ51
- ・カメラ等からの符号化対象の映像信号を映像符号化装置内に入力する符号化対象画像入力部52（ディスク装置等による、映像信号を記憶する記憶部でもよい）
- ・メモリ等から参照視点の動き情報を映像符号化装置内に入力する参照視点動き情報入力部53（ディスク装置等による、動き情報を記憶する記憶部でもよい）
- ・（デプス情報を取得するための）デプスカメラ等からの符号化対象画像を撮影した視点に対するデプスマップを映像符号化装置内に入力するデプスマップ入力部54（ディスク

10

20

30

40

50

装置等による、デプスマップを記憶する記憶部でもよい)

- ・映像符号化処理をCPU50に実行させるソフトウェアプログラムである映像符号化プログラム551が格納されたプログラム記憶装置55
- ・CPU50がメモリ51にロードされた映像符号化プログラム551を実行することにより生成されたビットストリームを、例えばネットワークを介して出力するビットストリーム出力部56(ディスク装置等による、ビットストリームを記憶する記憶部でもよい)とが、バスで接続された構成になっている。

【0092】

図7は、図4に示す映像復号装置200をコンピュータとソフトウェアプログラムとによって構成する場合のハードウェア構成を示すブロック図である。

図7に示すシステムは：

- ・プログラムを実行するCPU60
- ・CPU60がアクセスするプログラムやデータが格納されるRAM等のメモリ61
- ・映像符号化装置が本手法により符号化したビットストリームを映像復号装置内に入力するビットストリーム入力部62(ディスク装置等による、ビットストリームを記憶する記憶部でもよい)
- ・メモリ等からの参照視点の動き情報を映像復号装置内に入力する参照視点動き情報入力部63(ディスク装置等による、動き情報を記憶する記憶部でもよい)
- ・デプスカメラ等からの復号対象を撮影した視点に対するデプスマップを映像復号装置内に入力するデプスマップ入力部64(ディスク装置等による、デプス情報を記憶する記憶部でもよい)
- ・映像復号処理をCPU60に実行させるソフトウェアプログラムである映像復号プログラム651が格納されたプログラム記憶装置65
- ・CPU60がメモリ61にロードされた映像復号プログラム651を実行することにより、ビットストリームを復号して得られた復号対象画像を、再生装置などに出力する復号対象画像出力部66(ディスク装置等による、映像信号を記憶する記憶部でもよい)とが、バスで接続された構成になっている。

【0093】

前述した実施形態における映像符号化装置100及び映像復号装置200を、コンピュータで実現するようにしてもよい。その場合、この機能を実現するためのプログラムをコンピュータ読み取り可能な記録媒体に記録して、この記録媒体に記録されたプログラムをコンピュータシステムに読み込ませ、実行することによって実現してもよい。

なお、ここでいう「コンピュータシステム」とは、OSや周辺機器等のハードウェアを含むものとする。

また、「コンピュータ読み取り可能な記録媒体」とは、フレキシブルディスク、光磁気ディスク、ROM、CD-ROM等の可搬媒体、コンピュータシステムに内蔵されるハードディスク等の記憶装置のことをいう。

さらに「コンピュータ読み取り可能な記録媒体」とは、インターネット等のネットワークや電話回線等の通信回線を介してプログラムを送信する場合の通信線のように、短時間の間、動的にプログラムを保持するもの、その場合のサーバやクライアントとなるコンピュータシステム内部の揮発性メモリのように、一定時間プログラムを保持しているものも含んでもよい。

また上記プログラムは、前述した機能の一部を実現するためのものであっても良く、さらに前述した機能をコンピュータシステムにすでに記録されているプログラムとの組み合わせで実現できるものであってもよく、PLD(Programmable Logic Device)やFPGA(Field Programmable Gate Array)等のハードウェアを用いて実現されるものであってもよい。

【0094】

以上、図面を参照して本発明の実施の形態を説明してきたが、上記実施の形態は本発明の例示に過ぎず、本発明が上記実施の形態に限定されるものではないことは明らかである

10

20

30

40

50

。したがって、本発明の技術思想及び範囲を逸脱しない範囲で構成要素の追加、省略、置換、その他の変更を行っても良い。

【産業上の利用可能性】

【0095】

複数の視点に対する映像とその映像に対するデプスマップとを用いて表現される自由視点映像データを符号化または復号する際に、各視点の向きが平行でない場合でも、演算量を抑えたとままで、高精度な視点間の動き情報予測を実現することで、高い符号化効率を達成することが必要不可欠な用途に適用できる。

【符号の説明】

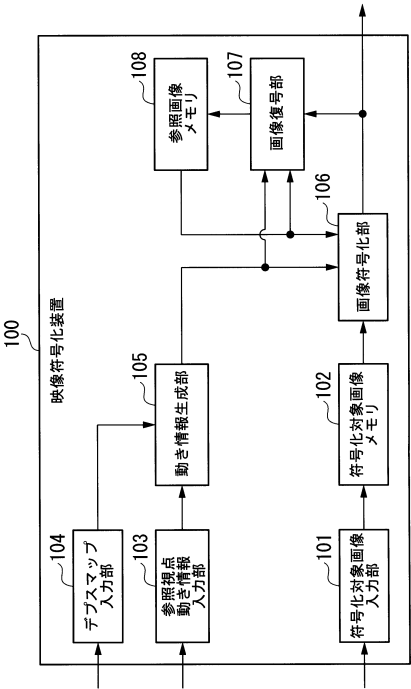
【0096】

- 100・・・映像符号化装置
- 101・・・符号化対象画像入力部
- 102・・・符号化対象画像メモリ
- 103・・・参照視点動き情報入力部
- 104・・・デプスマップ入力部
- 105・・・動き情報生成部
- 106・・・画像符号化部
- 107・・・画像復号部
- 108・・・参照画像メモリ
- 200・・・映像復号装置
- 201・・・ビットストリーム入力部
- 202・・・ビットストリームメモリ
- 203・・・参照視点動き情報入力部
- 204・・・デプスマップ入力部
- 205・・・動き情報生成部
- 206・・・画像復号部
- 207・・・参照画像メモリ

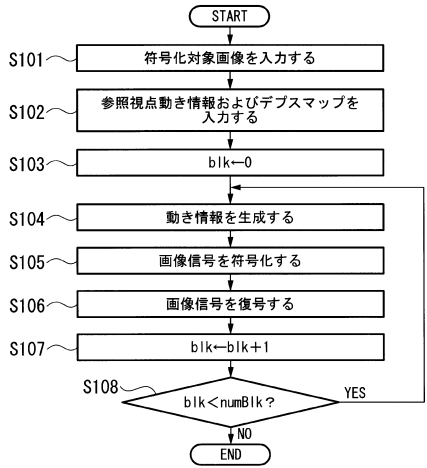
10

20

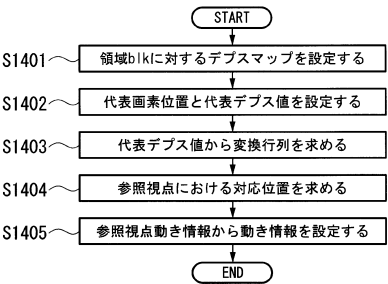
【図 1】



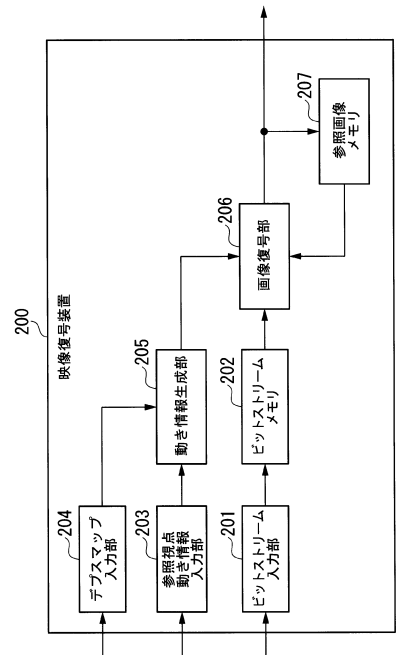
【図 2】



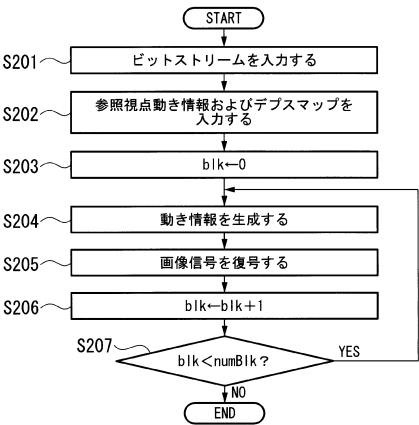
【図 3】



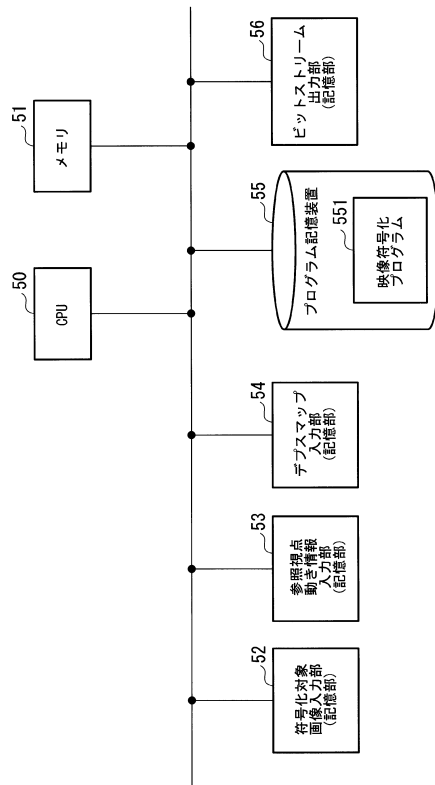
【図 4】



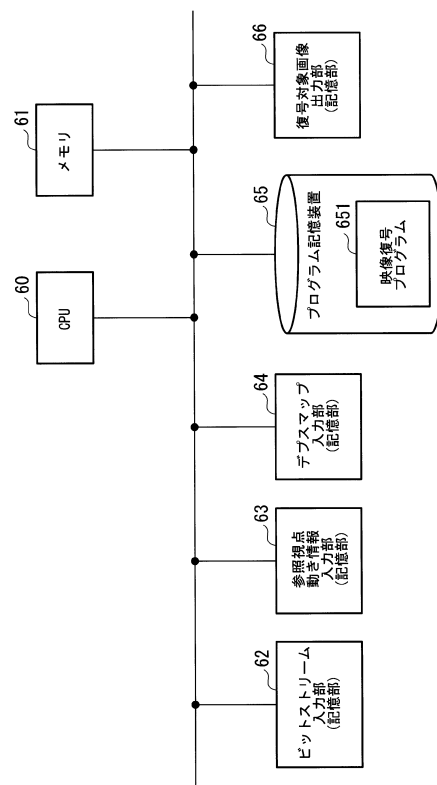
【図 5】



【図 6】



【図 7】



フロントページの続き

審査官 堀井 啓明

(56)参考文献 特開 2 0 0 9 - 1 1 6 5 3 2 (J P , A)

米国特許出願公開第 2 0 1 3 / 1 5 9 3 2 6 (U S , A 1)

志水 信哉 Shinya Shimizu, 3次元映像情報メディア技術, 映像情報メディア学会誌 第6
3巻 第4号 The Journal of The Institute of Image Information and Television Engineer
s, 日本, (社)映像情報メディア学会 The Institute of Image Information and Televisio
n Engineers, 第63巻

Gerhard Tech Fraunhofer HHI, 3D-HEVC Test Model 1[online], JCT3V-A JCT3V-A1005_d0
, インターネット<URL:http://phenix.it-sudparis.eu/jct2/doc_end_user/documents/1_Stock
holm/wg11/JCT3V-A1005-v1.zip>

Yu-Lin Chang No. 1, Dusing Rd. 1, Hsinchu Science Park, Hsinchu, Taiwan 30078, 3D-CE2.
h related: Simplified DV derivation for DoNBdV and BVSP[online], JCT3V-D JCT3V-D0
138, インターネット<URL:http://phenix.it-sudparis.eu/jct2/doc_end_user/documents/4_In
cheon/wg11/JCT3V-D0138-v2.zip>

(58)調査した分野(Int.Cl., D B 名)

H 0 4 N 1 9 / 0 0 - 1 9 / 9 8