

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2009-207183
(P2009-207183A)

(43) 公開日 平成21年9月10日(2009.9.10)

(51) Int.Cl. F I テーマコード(参考)
HO4N 7/32 (2006.01) HO4N 7/137 Z 5C159

審査請求 有 請求項の数 16 O L (全 30 頁)

(21) 出願番号	特願2009-141890 (P2009-141890)	(71) 出願人	000004226 日本電信電話株式会社 東京都千代田区大手町二丁目3番1号
(22) 出願日	平成21年6月15日(2009.6.15)	(74) 代理人	100087848 弁理士 小笠原 吉義
(62) 分割の表示	特願2005-218681 (P2005-218681) の分割	(74) 代理人	100095072 弁理士 岡田 光由
原出願日	平成17年7月28日(2005.7.28)	(72) 発明者	志水 信哉 東京都千代田区大手町二丁目3番1号 日 本電信電話株式会社内
		(72) 発明者	北原 正樹 東京都千代田区大手町二丁目3番1号 日 本電信電話株式会社内

最終頁に続く

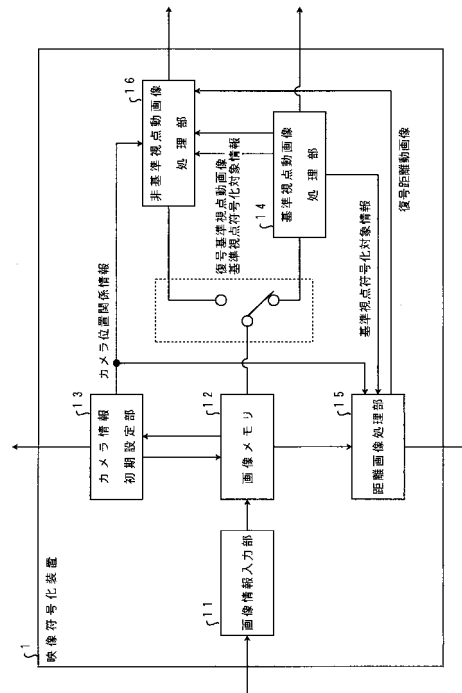
(54) 【発明の名称】映像符号化方法、映像復号方法、映像符号化プログラム、映像復号プログラム及びそれらのプログラムを記録したコンピュータ読み取り可能な記録媒体

(57) 【要約】

【課題】本発明は、視差方向予測と時間方向予測とを同時に使うことによって多視点動画を効率的に符号化するという構成を採るときにあって、必要となる映像を参照関係にある全てのカメラの映像を使うことなく復号できるようにする技術の提供を目的とする。

【解決手段】符号側は、基準視点のカメラにより撮影された基準視点画像と、基準視点画像を撮影したカメラから被写体までの推定距離を示す距離画像と、基準視点画像と距離画像とカメラの位置関係とに基づいて推定された視差補償画像と符号化対象画像との差分を示す視差差分画像とを符号化する。復号側は、基準視点画像と距離画像と視差差分画像とを復号して、復号した基準視点画像と復号した距離画像とカメラの位置関係とに基づいて視差補償画像を推定し、その推定した視差補償画像と復号した視差差分画像とに基づいて、基準視点以外の視点に対応付けられるカメラの撮影した符号化対象画像を復元する。

【選択図】図2



【特許請求の範囲】**【請求項 1】**

ある被写体を撮影する複数のカメラにより撮影された画像を符号化する映像符号化方法であって、

基準視点となるカメラにより撮影された基準視点画像を符号化するステップと、

上記基準視点画像を撮影したカメラから被写体までの推定距離を示す距離画像を生成するステップと、

上記距離画像を符号化するステップと、

上記基準視点画像と上記距離画像とカメラの設置位置および向きについて規定するカメラの位置関係とに基づいて、上記基準視点以外の視点における視差補償画像を推定するとともに、このとき画素値を推定できない画素については、周辺の画素の動き情報からその画素の動き情報を推定して、その推定した動き情報と符号化済みの画像の画素値とに基づいて、その画素の画素値を推定するステップと、

10

上記推定した視差補償画像とその推定対象の視点に対応付けられるカメラの撮影した符号化対象画像との差分を示す視差差分画像を算出するステップと、

上記視差差分画像を符号化するステップとを有することを、

特徴とする映像符号化方法。

【請求項 2】

請求項 1 に記載の映像符号化方法において、

上記視差補償画像を推定するステップでは、上記基準視点画像の符号化データを復号することで得られる基準視点画像と、上記距離画像の符号化データを復号することで得られる距離画像とを使って、上記基準視点以外の視点における視差補償画像を推定することを、

20

特徴とする映像符号化方法。

【請求項 3】

請求項 1 に記載の映像符号化方法において、

外部からの情報に従って上記カメラの位置関係を取得するか、全カメラの画像に基づいて上記カメラの位置関係を推定することで、上記カメラの位置関係を設定するステップと、

上記設定したカメラの位置関係の情報を符号化するステップとを有することを、

30

特徴とする映像符号化方法。

【請求項 4】

請求項 3 に記載の映像符号化方法において、

上記視差補償画像を推定するステップでは、上記基準視点画像の符号化データを復号することで得られる基準視点画像と、上記距離画像の符号化データを復号することで得られる距離画像と、上記カメラ位置関係情報の符号化データを復号することで得られるカメラの位置関係とに基づいて、上記基準視点以外の視点における視差補償画像を推定することを、

特徴とする映像符号化方法。

【請求項 5】

40

請求項 1 ないし 4 のいずれか 1 項に記載の映像符号化方法において、

他のカメラが撮影する空間と最も重複する空間を撮影しているカメラを基準視点となるカメラとして設定するステップを有することを、

特徴とする映像符号化方法。

【請求項 6】

請求項 1 ないし 5 のいずれか 1 項に記載の映像符号化方法において、

上記距離画像を生成するステップでは、画像をブロックに分割して、ブロックごとに距離を推定することで距離画像を生成することを、

特徴とする映像符号化方法。

【請求項 7】

50

請求項 1 ないし 6 のいずれか 1 項に記載の映像符号化方法において、

上記距離画像を生成するステップでは、規定のアルゴリズムに従って距離画像を生成する場合に、現時刻において生成した距離画像の評価値と、1つ前の時刻において生成した距離画像の評価値との差分値を求めて、その差分値の大きさを所定の閾値と比較することで判断して、その差分値が大きいことを判断する場合には、現時刻において生成した距離画像をそのまま用いることを決定し、その差分値が小さいことを判断する場合には、1つ前の時刻において生成した距離画像に変更して用いることを決定することで距離画像を生成することを、

特徴とする映像符号化方法。

【請求項 8】

10

請求項 1 ないし 7 のいずれか 1 項に記載の映像符号化方法において、

上記距離画像を符号化するステップでは、上記基準視点画像を符号化する際に使われた動きベクトルを用いて上記距離画像を符号化することを、

特徴とする映像符号化方法。

【請求項 9】

請求項 1 ないし 8 のいずれか 1 項に記載の映像符号化方法において、

上記視差差分画像を符号化するステップでは、上記基準視点画像を符号化する際に使われた動きベクトルと上記距離画像と上記カメラの位置関係とに基づいて推定される動きベクトルか、自身の参照画像から推定される動きベクトルの内の符号化効率のよい方を選択して符号化することを、

20

特徴とする映像符号化方法。

【請求項 10】

請求項 9 に記載の映像符号化方法において、

上記視差差分画像を符号化するステップでは、動きベクトルを推定するときに、上記距離画像の符号化データを復号することで得られる距離画像を用いることを、

特徴とする映像符号化方法。

【請求項 11】

ある被写体を撮影する複数のカメラにより撮影された画像の符号化データを復号する映像復号方法であって、

基準視点となるカメラにより撮影された基準視点画像についての符号化データを復号するステップと、

30

上記基準視点画像を撮影したカメラから被写体までの推定距離を示す距離画像についての符号化データを復号するステップと、

上記基準視点以外の視点に対応付けられるカメラの撮影した符号化対象画像との差分を示す視差差分画像についての符号化データを復号するステップと、

上記復号した基準視点画像と上記復号した距離画像とカメラの設置位置および向きについて規定するカメラの位置関係とに基づいて、上記基準視点以外の視点における視差補償画像を推定するとともに、このとき画素値を推定できない画素については、周辺の画素の動き情報からその画素の動き情報を推定して、その推定した動き情報と復号済みの画像の画素値とに基づいて、その画素の画素値を推定するステップと、

40

上記推定した視差補償画像と上記復号した視差差分画像とに基づいて、上記基準視点以外の視点に対応付けられるカメラの撮影した符号化対象画像を復元するステップとを有することを、

特徴とする映像復号方法。

【請求項 12】

請求項 11 に記載の映像復号方法において、

上記カメラの位置関係の情報についての符号化データを復号するステップを有することを、

特徴とする映像復号方法。

【請求項 13】

50

請求項 1 ないし 10 のいずれか 1 項に記載の映像符号化方法をコンピュータに実行させるための映像符号化プログラム。

【請求項 14】

請求項 1 ないし 10 のいずれか 1 項に記載の映像符号化方法をコンピュータに実行させるための映像符号化プログラムを記録したコンピュータ読み取り可能な記録媒体。

【請求項 15】

請求項 11 または 12 に記載の映像復号方法をコンピュータに実行させるための映像復号プログラム。

【請求項 16】

請求項 11 または 12 に記載の映像復号方法をコンピュータに実行させるための映像復号プログラムを記録したコンピュータ読み取り可能な記録媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ある被写体を撮影する複数のカメラにより撮影された画像を符号化する映像符号化方法と、その映像符号化方法の実現に用いられる映像符号化プログラム及びそのプログラムを記録したコンピュータ読み取り可能な記録媒体と、その映像符号化方法で符号化されたデータを復号する映像復号方法と、その映像復号方法の実現に用いられる映像復号プログラム及びそのプログラムを記録したコンピュータ読み取り可能な記録媒体とに関する。

【背景技術】

【0002】

多視点動画像（多視点映像）とは、複数のカメラで同じ被写体と背景を撮影した複数の動画像のことである。以下では、1つのカメラで撮影された動画像を“2次元動画像”と呼び、同じ被写体と背景を撮影した2次元動画像群を多視点動画像と呼ぶ。

【0003】

多視点動画像に含まれる各カメラの2次元動画像は、時間方向に関して強い相関がある。一方、各カメラが同期されていた場合、同じ時間に対応した各カメラの映像は全く同じ状態の被写体と背景を別の位置から撮影したものであるため、カメラ間で強い相関がある。以下では、このカメラ間での相関を空間方向の相関と呼ぶ。

【0004】

動画像の符号化においては、これらの相関を利用することによって符号化効率を高めている。

【0005】

まず、2次元動画像の符号化技術に関する従来技術を述べる。

【0006】

2次元動画像の符号化では、符号化しようとする画像と既に符号化済みの画像との差分のみを符号化対象とすることで、時間方向の相関を利用し、符号化効率を高めている。

【0007】

国際符号化標準である H.264、MPEG-2、MPEG-4 をはじめとした従来の多くの2次元動画像符号化方式では、差分を求める際に動き補償と呼ばれる技術を用いることで、更に符号化効率を高めている。動き補償とは、画像をより小さなブロックに分割し、そのブロックごとに差分をとるブロックを切り替える技術である。これによって、被写体に動きが生じたり、カメラが動いたりする場合にも差分を小さくし、符号化効率をあげることができる。

【0008】

次に、従来の多視点動画像の符号化方式について説明する。

【0009】

従来の多視点動画像の符号化では、時間方向及び空間方向の相関を利用して符号化効率を高めるために、時間方向の予測及びカメラ間での補償を行った符号化を採用している。

10

20

30

40

50

その一例としては、下記に示す非特許文献1の手法がある。

【0010】

この非特許文献1の手法で行われるカメラ間での補償は視差方向予測と呼ばれ、別のカメラの画像を参照画像として動き補償を行うものである。この手法では符号化効率を考え、マクロブロックごとにどちらの相関を利用した補償を行うのかを選択できるようになっている。そのため、時間方向及び空間方向の相関が符号化に利用されるため、時間方向のみの相関を利用した方法より符号化効率を向上させることが可能である。

【先行技術文献】

【非特許文献】

【0011】

【非特許文献1】Hideaki Kimata and Masaki Kitahara, "Preliminary results on multiple view video coding(3DAV)," document M10976 MPEG Redmond Meeting, July, 2004.

【発明の概要】

【発明が解決しようとする課題】

【0012】

確かに、非特許文献1の手法では、時間方向のみの相関を利用した方法より符号化効率を向上させることができるようになるものの、複数のカメラの映像を参照し合う方法を用いていることから、データを復号する際に参照関係にある全てのカメラの映像を必要とすることになる。

【0013】

例えば、図19に示すような参照関係がある場合、全てのカメラの映像を復号側に渡す必要がある。これは多視点動画の用法の1つである自由視点動画合成においては、必要のないカメラの映像まで復号する必要があるということになってしまう。ネットワークでデータを通信させることを考えると、本来必要のない情報まで相手に送る必要が生じるということになる。

【0014】

ここで述べている自由視点動画合成とは、カメラの置いてない地点からの映像を隣接するカメラの映像を用いて合成する技術のことである。したがって、各カメラの映像を独立して取り出せるように符号化する機能が求められる。

【0015】

また、多視点動画の用法として自由視点動画の合成を考えた場合、自由視点動画の合成処理は非常に演算量の多い処理であるため、復号と合成とを同時に行うのは高負荷なことと言える。したがって、符号・復号における処理で合成処理を手助けできるような機能も求められる。

【0016】

各カメラの映像を必要に応じて独立して取り出せる機能を実現する符号化方法として、非特許文献1のような手法における参照関係を制限し、ある特定のカメラの映像しか参照できないようにする方法を用いることが考えられる。この方法を用いれば、参照関係が制限されることによって、データを復号する際に必要なカメラの映像を減らすことができる。

【0017】

しかしながら、このような方法では、符号化側で設定された単位毎でしかカメラの情報を個別に取り出すことはできない。また、各カメラの映像を1つ以上のビットストリームとして符号化することは明らかに効率的ではないため、参照関係が制限されることによって、あるカメラ間に強い相関があってもそれを利用できなくなってしまう場合もある。つまり、相関を取り除けなくなるので、全体として十分な符号化効率を達成することが困難となる。

【0018】

また、下記に示す参考文献によれば、多視点動画から自由視点動画を合成する手法には、被写体の幾何情報を利用して合成する手法と、被写体の幾何情報を利用しないで映

10

20

30

40

50

像から合成する手法とがある。

【0019】

参考文献：Heung-Yeung Shum, Sing Bing Kang, and Shing-Chow Chan, "Survey of image-based representations and compression techniques," IEEE Transactions on Circuits and Systems for Video Technology, vol.13, no. 11, Nov.2003, pp.1020-1037.

幾何情報を利用しない手法でもって、幾何情報を利用する手法と同等の品質の映像を合成するためには、より多くのカメラからの映像が入力として必要となる。多くのカメラの映像を用いるということは非常に多くの演算が必要になることを意味する。

【0020】

この点について、非特許文献1のような手法では、各画素のもつ情報は、前のフレームからの動きベクトルまたは他のカメラからの視差ベクトルのどちらかと残差信号という形でしか表現されない。したがって、非特許文献1のような手法で符号化された多視点動画像を用いて自由視点動画像を合成する場合には、幾何情報を計算する必要が生じて演算量が増して合成に時間がかかることになるか、より多くのカメラからの映像が必要になることになる。

【0021】

本発明はかかる事情に鑑みてなされたものであって、視差方向予測と時間方向予測とを同時に使うことによって多視点動画像の符号化効率を向上できるようにするという構成を採るときにあって、必要となる映像を参照関係にある全てのカメラの映像を使うことなく復号できるようにすることで、自由視点動画像の合成を小さな負荷で実現できるようにする新たな映像符号化復号技術の提供を目的とする。

【課題を解決するための手段】

【0022】

〔1〕本発明の映像符号化方法の基本的な構成

この目的を達成するために、本発明の映像符号化方法は、ある被写体を撮影する複数のカメラにより撮影された画像を符号化することを実現するために、(a)基準視点となるカメラにより撮影された基準視点画像を符号化する基準視点画像符号化ステップと、(b)基準視点画像を撮影したカメラから被写体までの推定距離を示す距離画像を生成する距離画像生成ステップと、(c)生成した距離画像を符号化する距離画像符号化ステップと、(d)基準視点画像と距離画像とカメラの設置位置および向きについて規定するカメラの位置関係とに基づいて、基準視点以外の視点における視差補償画像(視差に基づく予測画像となるもの)を推定するとともに、このとき画素値を推定できない画素については、周辺の画素の動き情報からその画素の動き情報を推定して、その推定した動き情報と符号化済みの画像の画素値とに基づいて、その画素の画素値を推定する視差補償画像推定ステップと、(e)推定した視差補償画像とその推定対象の視点に対応付けられるカメラの撮影した符号化対象画像との差分を示す視差差分画像を算出する視差差分画像算出ステップと、(f)算出した視差差分画像を符号化する視差差分画像符号化ステップとを有する。

【0023】

この基本的な構成を採るときにあって、本発明の映像符号化方法は、さらに次の構成を採ることがある。

【0024】

〔1-1〕

視差補償画像推定ステップでは、復号側がカメラの位置関係の情報を符号化データからではなくて得ることができる場合には、基準視点画像の符号化データを復号することで得られる基準視点画像と、距離画像の符号化データを復号することで得られる距離画像と、符号化されることのないカメラの位置関係とに基づいて、基準視点以外の視点における視差補償画像を推定することがある。

【0025】

〔1-2〕

10

20

30

40

50

復号側がカメラの位置関係の情報を符号化データから得ることになる場合には、上述の基本的な構成のステップに加えて、(g)外部からの情報に従ってカメラの位置関係を取得するか、全カメラの画像に基づいてカメラの位置関係を推定することで、カメラの位置関係を設定するカメラ位置関係設定ステップと、(h)設定したカメラの位置関係の情報を符号化するカメラ位置関係情報符号化ステップとを有する。

【0026】

〔1-3〕

視差補償画像推定ステップでは、復号側がカメラの位置関係の情報を符号化データから得ることになる場合には、基準視点画像の符号化データを復号することで得られる基準視点画像と、距離画像の符号化データを復号することで得られる距離画像と、カメラ位置関係情報の符号化データを復号することで得られるカメラの位置関係とに基づいて、基準視点以外の視点における視差補償画像を推定することがある。

10

【0027】

〔1-4〕

基準視点となるカメラを自動的に決定する場合には、上述の基本的な構成のステップに加えて、(i)他のカメラが撮影する空間と最も重複する空間を撮影しているカメラを基準視点となるカメラとして設定するステップを有することがある。

【0028】

〔1-5〕

距離画像生成ステップでは、画像をブロックに分割して、ブロックごとに距離を推定することで距離画像を生成することがある。

20

【0029】

〔1-6〕

距離画像生成ステップでは、規定のアルゴリズムに従って距離画像を生成する場合に、現時刻において生成した距離画像の評価値と、1つ前の時刻において生成した距離画像の評価値との差分値を求めて、その差分値の大きさを所定の閾値と比較することで判断して、その差分値が大きいことを判断する場合には、現時刻において生成した距離画像をそのまま用いることを決定し、その差分値が小さいことを判断する場合には、1つ前の時刻において生成した距離画像に変更して用いることを決定することで距離画像を生成することがある。

30

【0030】

〔1-7〕

視差補償画像推定ステップでは、基準視点画像と距離画像とカメラの位置関係とに基づいて基準視点以外の視点における視差補償画像を推定することになるが、このとき、画素値を推定できない画素について、周辺の画素の画素値から、その画素の画素値を推定するように処理することも可能である。

【0031】

〔1-8〕

実際には、視差補償画像推定ステップでは、基準視点画像と距離画像とカメラの位置関係とに基づいて基準視点以外の視点における視差補償画像を推定することになるが、このとき、画素値を推定できない画素については、周辺の画素の動き情報からその画素の動き情報を推定して、その推定した動き情報と符号化済みの画像の画素値とに基づいて、その画素の画素値を推定することになる。

40

【0032】

〔1-9〕

視差補償画像推定ステップでは、基準視点画像と距離画像とカメラの位置関係とに基づいて基準視点以外の視点における視差補償画像を推定することになるが、このとき、画素値を推定できない画素について、〔1-7〕の方法に従って推定した視差補償画像を用いる場合の符号量と、〔1-8〕の方法に従って推定した視差補償画像を用いる場合の符号量とを比較して、視差補償画像ごとに効率的な符号化を行える方法を選択することで、そ

50

の画素の画素値を推定するように処理することも可能である。この構成を採るときには、どちらの予測モードを用いたのかを示す情報についても符号化することになる。

【 0 0 3 3 】

〔 1 - 1 0 〕

距離画像符号化ステップでは、基準視点画像を符号化する際に使われた動きベクトルを用いて距離画像を符号化することがある。

【 0 0 3 4 】

〔 1 - 1 1 〕

視差差分画像符号化ステップでは、基準視点画像を符号化する際に使われた動きベクトルと距離画像とカメラの位置関係とに基づいて推定される動きベクトルか、自身の参照画像から推定される動きベクトルの内の符号化効率のよい方を選択して符号化することがある。

10

【 0 0 3 5 】

〔 1 - 1 2 〕

視差差分画像符号化ステップでは、基準視点画像を符号化する際に使われた動きベクトルと距離画像とカメラの位置関係とに基づいて推定される動きベクトルか、自身の参照画像から推定される動きベクトルの内の符号化効率のよい方を選択して符号化することがあるが、この動きベクトルを推定するときに、距離画像の符号化データを復号することで得られる距離画像を用いて動きベクトルを推定することがある。

【 0 0 3 6 】

20

ここで、このように構成される本発明の映像符号化方法はコンピュータプログラムでも実現できるものであり、このコンピュータプログラムは、適当なコンピュータ読み取り可能な記録媒体に記録して提供されたり、ネットワークを介して提供され、本発明を実施する際にインストールされてCPUなどの制御手段上で動作することにより本発明を実現することになる。

【 0 0 3 7 】

〔 2 〕本発明の映像復号方法の基本的な構成

本発明の映像復号方法は、本発明の映像符号化方法により生成された符号化データを復号することで、ある被写体を撮影する複数のカメラにより撮影された画像を復元することを実現するために、(a) 基準視点となるカメラにより撮影された基準視点画像についての符号化データを復号する基準視点画像復号ステップと、(b) 基準視点画像を撮影したカメラから被写体までの推定距離を示す距離画像についての符号化データを復号する距離画像復号ステップと、(c) 基準視点以外の視点に対応付けられるカメラの撮影した符号化対象画像との差分を示す視差差分画像についての符号化データを復号する視差差分画像復号ステップと、(d) 復号した基準視点画像と復号した距離画像とカメラの設置位置および向きについて規定するカメラの位置関係とに基づいて、基準視点以外の視点における視差補償画像を推定するとともに、このとき画素値を推定できない画素については、周辺の画素の動き情報からその画素の動き情報を推定して、その推定した動き情報と復号済みの画像の画素値とに基づいて、その画素の画素値を推定する視差補償画像推定ステップと、(e) 推定した視差補償画像と復号した視差差分画像とに基づいて、基準視点以外の視点に対応付けられるカメラの撮影した符号化対象画像を復元する符号化対象画像復元ステップとを有する。

30

40

【 0 0 3 8 】

この基本的な構成を採るときにあつて、本発明の映像復号方法は、さらに次の構成を採ることがある。

【 0 0 3 9 】

〔 2 - 1 〕

カメラの位置関係の情報を符号化データから得ることになる場合には、上述の基本的な構成のステップに加えて、(f) 各画像を撮影したカメラの位置関係の情報についての符号化データを復号するカメラ位置関係情報復号ステップを有する。

50

【 0 0 4 0 】

〔 2 - 2 〕

視差補償画像推定ステップでは、復号した基準視点画像と復号した距離画像とカメラの位置関係とに基づいて基準視点以外の視点における視差補償画像を推定することになるが、このとき、画素値を推定できない画素については、周辺の画素の画素値から、その画素の画素値を推定するように処理することも可能である。

【 0 0 4 1 】

〔 2 - 3 〕

実際には、視差補償画像推定ステップでは、復号した基準視点画像と復号した距離画像とカメラの位置関係とに基づいて基準視点以外の視点における視差補償画像を推定することになるが、このとき、画素値を推定できない画素については、周辺の画素の動き情報からその画素の動き情報を推定して、その推定した動き情報と復号済みの画像の画素値とに基づいて、その画素の画素値を推定することになる。

10

【 0 0 4 2 】

〔 2 - 4 〕

視差補償画像推定ステップでは、復号した基準視点画像と復号した距離画像とカメラの位置関係とに基づいて基準視点以外の視点における視差補償画像を推定することになるが、このとき、画素値を推定できない画素については、符号化データに埋め込まれている予測モードの情報に基づいて、〔 2 - 2 〕の推定方法か〔 2 - 3 〕の推定方法のどちらかを選択することで、視差補償画像を単位にして2つの方法を切り替えながら画素値の推定を行うように処理することも可能である。

20

【 0 0 4 3 】

ここで、このように構成される本発明の映像復号方法はコンピュータプログラムでも実現できるものであり、このコンピュータプログラムは、適当なコンピュータ読み取り可能な記録媒体に記録して提供されたり、ネットワークを介して提供され、本発明を実施する際にインストールされてCPUなどの制御手段上で動作することにより本発明を実現することになる。

【 0 0 4 4 】

〔 3 〕本発明の処理について

本発明の映像符号化方法は、復号側がカメラの位置関係の情報を符号化データからではなくて得ることができる場合には、基準視点となるカメラにより撮影された基準視点画像と、基準視点画像を撮影したカメラから被写体までの推定距離を示す距離画像と、基準視点画像と距離画像とカメラの位置関係とに基づいて推定された視差補償画像と符号化対象画像との差分を示す視差差分画像という3種類の画像を符号化する。そして、復号側がカメラの位置関係の情報を符号化データから得ることになる場合には、この3種類の画像に加えて、カメラの位置関係の情報を符号化する。

30

【 0 0 4 5 】

この符号化データを受けて、本発明の映像復号方法は、基準視点画像と距離画像と視差差分画像とを復号することにより得て、復号した基準視点画像と復号した距離画像とカメラの位置関係とに基づいて、基準視点以外の視点における視差補償画像を推定して、その推定した視差補償画像と復号した視差差分画像とに基づいて、基準視点以外の視点に対応付けられるカメラの撮影した符号化対象画像を復元する。

40

【 0 0 4 6 】

このように、本発明では、符号化対象の多視点動画像の中の1つの視点（基準視点）から被写体への距離を、その多視点動画像を用いて求めて距離画像を作り出す。次に、基準視点における動画像と奥行き情報の動画像（距離動画像）とを用いて、他視点における動画像を予測する。そして、この予測動画像と符号化対象の動画像との差分を求めることで差分の動画像を得て、基準視点の動画像、基準視点における奥行き情報の動画像（距離動画像）、基準視点以外の視点における差分の動画像をそれぞれ2次元動画像として符号化する。

50

【0047】

つまり、本発明では、参照は基準視点としか行われていないため、必要なカメラ以外の情報をできるだけ抑えて取り出すという機能を実現できることになる。また、奥行き情報の動画像（距離動画像）は全ての視点の情報から作られたものであるため、それを用いて予測することで、空間方向の相関を利用できることになる。

【0048】

更に、予測残差は基準視点から予測できない情報であることから、基準視点の画像には含まれてなく、その視点にのみ含まれる情報が残るのだが、これに対しては動き予測を用いて時間方向の相関を利用することで符号化効率を向上できることになる。

【0049】

更に、奥行き情報の動画像（距離動画像）は基準視点と他の全ての視点との間の視差情報を1つの表現で表すことができているだけでなく、それを2次元動画像として符号化することで視差の時間方向の冗長性を取り除くことができることになる。

【0050】

また、多視点動画像から自由視点動画像を合成する場合、被写体の幾何情報が存在することによって、ある品質の映像を作り出すのに必要な画像の個数が少なく済んだり、合成の処理が簡略化されたりすることになるのだが、本発明で求める奥行き動画像（距離動画像）は一種の幾何情報を表しているため、合成画像の品質を向上させたり、幾何情報の推定処理を簡略化できることになる。

【0051】

このように、本発明によれば、受け手が必要とする任意のカメラの映像を、必要ないカメラの情報を送信することをできるだけ抑えて取り出すという機能を実現できるようになる。そして、ある視点からの被写体の推定距離情報を用いて符号化されているため、被写体の幾何情報が提供されることとなり、自由視点動画像を合成する前処理が省略できるという利点を持ち、なお、かつ空間方向の相関と時間方向の相関とを同時に利用することによって多視点動画像を効率的に符号化できるようになる。

【0052】

次に、本発明の映像符号化方法が持つ上述の〔1-1〕～〔1-12〕に記載した各処理機能の持つ意味について説明する。

【0053】

(A)〔1-1〕に記載した処理機能の持つ意味

【0054】

復号側が入手できる距離画像と基準視点画像とは符号化されたものを復号した画像であるため、符号化側で、符号化・復号を行わないオリジナルの距離画像と基準視点画像とを用いて推定した視差補償画像を用いると、復号側で推定される視差補償画像との間に誤差が存在することになり、基準視点以外の画像が、視差差分画像の誤差と基準視点画像の誤差と距離画像の誤差とが重なったものになってしまう。

【0055】

そこで、本発明では、〔1-1〕に記載するように、復号側がカメラの位置関係の情報を符号化データからではなくて得ることができる場合には、符号化側は、基準視点画像の符号化データを復号することで得られる基準視点画像と、距離画像の符号化データを復号することで得られる距離画像と、符号化されることのないカメラの位置関係とに基づいて、基準視点以外の視点における視差補償画像を推定するように処理するのである。

【0056】

この処理機能に従って、符号化側で符号化歪みの入った距離画像と基準視点画像とを用いることになるため、基準視点以外の画像における符号化歪みは視差差分画像における符号化歪みの影響だけにすることができるようになる。すなわち、視差差分画像の符号化において、復号側における距離画像と基準視点画像の符号化歪みを考慮した符号化を行うことができることで、最高で歪みなしの符号化を達成することができるようになる。

【0057】

10

20

30

40

50

(B)〔1-2〕に記載した処理機能の持つ意味

【0058】

カメラの位置関係が明示的に与えられなかった場合、多視点動画像から距離画像を求め、視差補償画像を求める処理の精度は著しく低下する。そのため視差差分画像により多くの残差が存在することにより符号化効率が悪くなる。

【0059】

そこで、本発明では、〔1-2〕に記載するように、復号側がカメラの位置関係の情報を符号化データから得ることになる場合であって、カメラの位置関係が外部から与えられる場合には、それを取得して符号化し、一方、カメラの位置関係が外部から与えられない場合には、全カメラの画像に基づいてカメラの位置関係を推定して、それを符号化するように処理するのである。

10

【0060】

この処理機能に従って、カメラの位置関係を符号化する必要があるときに、カメラの位置関係が外部から与えられない場合には、与えられた多視点動画像からカメラの位置関係を推定するため、より正確な距離画像と視差補償画像とを求めることができることで、視差差分画像に残る信号を小さくすることができるため、カメラの位置関係が明示的に与えられなくても符号化効率が悪くなることを防ぐことができる。すなわち、カメラの位置関係が明示的に与えられなくても、さまざまなカメラ配置の多視点動画像に対して、より柔軟に対応して符号化を行うことができるようになる。

【0061】

20

(C)〔1-3〕に記載した処理機能の持つ意味

【0062】

〔1-3〕に記載した処理機能の持つ意味は、〔1-1〕に記載した処理機能の持つ意味と基本的に同じである。ただし、〔1-3〕に記載した処理機能では、カメラの位置関係の情報も符号化することを想定しているので、視差補償画像を推定する際に必要となるカメラの位置関係についても、カメラ位置関係情報の符号化データを復号することで得られるものを使用するように処理している。

【0063】

(D)〔1-4〕に記載した処理機能の持つ意味

【0064】

30

各カメラはほぼ同じ被写体と背景を撮影しているといっても、全く同じ視点から撮影を行っていないわけではないので、各カメラが撮影している空間は異なる。そのため、どのカメラ対で視差予測をするかによって、精度の高い視差予測が可能な領域の大きさが異なる。精度の高い視差予測が可能な領域が小さくなるカメラを基準視点としてしまうと符号化効率が悪くなる。

【0065】

そこで、本発明では、〔1-4〕に記載するように、他のカメラが撮影する空間と最も重複する空間を撮影しているカメラを基準視点となるカメラとして設定するように処理するのである。

【0066】

40

この処理機能に従って、各カメラから撮影される映像の共通部分がより多く撮影されているカメラを基準視点に選ぶことになるので、全カメラにおける精度の高い視差予測が可能な領域の合計を最大にすることができるため、符号化効率が悪くなることを防ぐことができる。すなわち、多くの領域に関して精度の高い視差補償画像を生成できるため、より効率的な符号化を行うことができるようになる。

【0067】

(E)〔1-5〕に記載した処理機能の持つ意味

【0068】

本発明においては、もともとの符号化対象である多視点動画像のほかに距離画像を符号化する必要が生じている。そのため、距離画像を符号化するのに必要な符号量はできる限

50

り少なくする必要がある。距離画像は視差情報を提供して視差補償画像を作成するために存在することから、そのために必要な情報が含まれていればよいということになる。つまり、距離画像は画素単位の視差情報を提供できる程度のものでかまわないということを示している。そのような精度の距離推定では隣接する画素での推定距離が同じになる場合が多い。

【0069】

そこで、本発明では、〔1-5〕に記載するように、画像をブロックに分割して、ブロックごとに距離を推定することで距離画像を生成するように処理するのである。

【0070】

この処理機能に従って、精度をある程度保ったまま、距離画像の符号化において、より効率的な符号化を行うことができるようになる。

10

【0071】

(F)〔1-6〕に記載した処理機能の持つ意味

【0072】

視差予測された画像と符号化対象の画像との誤差が単純に小さくなるように距離の予測を行った場合、距離の予測には時間方向の相関が考慮に入れられていないことから、視差予測された視差補償画像における時間方向の相関が失われてしまう。視差予測された視差補償画像において時間方向の相関がない場合、それと符号化対象の画像との差分で求められる視差差分画像においても時間方向の相関がなくなってしまうことになる。そのような場合、視差差分画像の符号化効率は非常に悪くなる。

20

【0073】

そこで、本発明では、〔1-6〕に記載するように、規定のアルゴリズムに従って距離画像を生成する場合に、その距離画像に基づいて算出されることになる視差差分画像の符号化効率が向上するようにと、その生成した距離画像をそのまま用いるのか時間的に変化しないものに変更して用いるのかを決定するように処理するのである。ここで、この決定については、距離画像がブロックを単位にして符号化されるような場合には、そのブロックを単位にして行うことになる。

【0074】

この処理機能に従って、視差差分画像における時間方向の相関を考慮に入れて距離を推定するため、視差差分画像において時間方向の相関が失われることを防ぐことが可能となり、視差差分画像の符号化において、より効率的な符号化を行うことができるようになる。

30

【0075】

(G)〔1-7〕に記載した処理機能の持つ意味

【0076】

基準視点画像と他の視点における画像とは、視差補償をする場合、必ずしも1対1に対応しない。そのため視差補償画像において予測値の存在しない画素が現れることになる。そうすると視差差分画像は画素間で大きく値の異なる画像となる。そのような画像は自然な画像と異なる性質を持つようになることから、一般的な符号化手法を用いたときに符号化効率が低下する。

40

【0077】

そこで、本発明では、〔1-7〕に記載するように、視差補償画像を推定するときに画素値を推定できない画素について、周辺の画素の画素値から、その画素の画素値を推定するように処理するのである。

【0078】

この処理機能に従って、視差補償値のない画素に対して、隣接画素から予測値を作り出すことで、画素間で値が大きく異ならない視差差分画像を生成することが可能となる。隣接画素の値に相関があることはよく知られており、2次元動画像の符号化でも用いられているため、この予測が大きく外れることは少ない。したがって、視差差分画像における符号化対象情報を減らすことで符号化効率を上げることができるようになる。

50

【0079】

(H)〔1-8〕に記載した処理機能の持つ意味

【0080】

〔1-7〕に記載した処理機能では、視差補償画像を推定するときに画素値を推定できない画素について、空間的な予測を行うことでその画素値を推定するようにしている。

【0081】

これに対して、〔1-8〕に記載した処理機能では、視差補償画像を推定するときに画素値を推定できない画素について、周辺の画素の動き情報からその画素の動き情報を推定して、その推定した動き情報と符号化済みの画像の画素値とに基づいて、その画素の画素値を推定するように処理している。

10

【0082】

この処理機能では、一般的な動画の符号化で用いられる隣接画素における動きベクトルの相関を利用していることから、この予測が大きく外れることは少ない。したがって、〔1-7〕に記載した処理機能と同様に、視差差分画像における符号化対象情報を減らすことで符号化効率を上げることができるようになる。

【0083】

(I)〔1-9〕に記載した処理機能の持つ意味

【0084】

〔1-7〕に記載した処理機能では、視差補償画像を推定するときに画素値を推定できない画素について、空間的な予測を行うことでその画素値を推定するようにし、一方、〔1-8〕に記載した処理機能では、視差補償画像を推定するときに画素値を推定できない画素について、時間的な予測を行うことでその画素値を推定するようにしている。

20

【0085】

このどちらの予測方法を用いるのかについては、符号化効率の観点から選択することが望ましい。

【0086】

そこで、本発明では、〔1-9〕に記載するように、〔1-7〕の推定方法に従って推定した視差補償画像を用いる場合の符号量と、〔1-8〕の推定方法に従って推定した視差補償画像を用いる場合の符号量とを比較して、視差補償画像ごとに効率的な符号化を行える方法を選択することで、その画素の画素値を推定するように処理するのである。この構成を採るときには、どちらの予測モードを用いたのかを示す情報についても符号化することになる。

30

【0087】

この処理機能に従って、視差補償値のない画素に対して予測値を作り出すことで視差補償画像を推定し、これにより、視差差分画像における符号化対象情報を減らすことで符号化効率を上げることができるようにすることを実現するときにあつて、その符号化効率の向上をさらに確実なものにすることができるようになる。

【0088】

(J)〔1-10〕に記載した処理機能の持つ意味

【0089】

基準視点画像と距離画像とはどちらも同じ視点からの画像であるため、動きに関しては非常に強い相関があると言える。したがって、両者を独立に符号化してしまうと、かなりの冗長性が残ることになる。

40

【0090】

そこで、本発明では、〔1-10〕に記載するように、基準視点画像を符号化する際に使われた動きベクトルを用いて距離画像を符号化するように処理するのである。

【0091】

この処理機能に従って、同じ動きを表すベクトルを重複して符号化することをなくして、符号化効率を向上させることができるようになる。しかも、基準視点以外の視点の画像を復号する際には、基準視点も距離画像も必ず必要であるため、両者に参照関係があつて

50

も、本発明により実現される必要なカメラ以外の情報をできるだけ抑えて取り出すという機能を損なうことにはならない。したがって、基準視点画像と距離画像との相関を用いて、基準視点画像と距離画像とをあわせた符号化効率を上げることができるようになる。

【0092】

(K)〔1-11〕に記載した処理機能の持つ意味

【0093】

同じ被写体を撮影しているということは、被写体の実際の3次元的動きは1つである。そのため、カメラの位置関係と距離画像と基準視点における動きベクトルとから、他視点における動きをある程度予測することができる。しかし、基準視点における動きベクトルは2次元であり、距離画像も完全な3次元ではないので、予測ベクトルは常に正しいわけではない。また、符号化効率という面而言えば、実際の動きを表した動きベクトルを用いることが常に最も高い符号化効率を達成するとは限らない。

10

【0094】

そこで、本発明では、〔1-11〕に記載するように、視差差分画像を符号化するときに、基準視点画像を符号化する際に使われた動きベクトルと距離画像とカメラの位置関係とに基づいて推定される動きベクトルか、自身の参照画像から推定される動きベクトルの内の符号化効率のよい方を選択して符号化するように処理するのである。

【0095】

この処理機能に従って、各視点に閉じて求められる動きベクトルと視差補償で求められる動きベクトルのうち符号化効率のよい方を選ぶことで、視差補償で求められる動きベクトルの間違いを許容して高い符号化効率を達成することが可能となる。すなわち、同じ被写体を撮影していることから、各視点における動画画像に含まれる動き成分の相関を利用することで、視差差分画像の符号化効率を上げることができるようになる。

20

【0096】

(L)〔1-12〕に記載した処理機能の持つ意味

【0097】

〔1-11〕に記載した処理機能では、基準視点画像を符号化する際に使われた動きベクトルと距離画像とカメラの位置関係とに基づいて動きベクトルを推定することになるが、〔1-12〕に記載した処理機能では、この動きベクトルの推定にあたって、復号側の処理に合わせて、距離画像の符号化データを復号することで得られる距離画像を用いて動きベクトルを推定するように処理する。

30

【0098】

この処理機能に従って、復号側と同じ方法で動きベクトルを推定するようにすることから、距離画像に含まれる符号化歪みの影響を取り除いて動きベクトルの推定が可能になることから、復号において歪みが蓄積するのを防ぐことができるようになる。

【発明の効果】

【0099】

本発明によれば、視差方向予測と時間方向予測とを同時に使うことによって多視点動画の符号化効率を向上させることができるようになるとともに、この符号化にあたって、共通情報として、基準視点の映像と基準視点における距離画像とを利用することによって、受け手が必要とする任意のカメラの映像を、必要ないカメラの情報を送信することができるだけ抑えて取り出すという機能を実現できるようになる。

40

【0100】

そして、ある視点からの被写体の推定距離情報を用いて符号化を行うことから、被写体の幾何情報が提供されることとなり、自由視点動画を生成するのに必要な幾何情報を推定する処理量を削減できるようになるとともに、必要とされる品質の自由視点動画を生成するのに必要な映像の数を減少させたりすることができるようになる。

【図面の簡単な説明】

【0101】

【図1】カメラ構成の一例を示す図である。

50

- 【図 2】本発明の映像符号化装置の一実施形態例である。
- 【図 3】カメラ情報初期設定部の構成の詳細を示す図である。
- 【図 4】基準視点動画像処理部の構成の詳細を示す図である。
- 【図 5】距離画像処理部の構成の詳細を示す図である。
- 【図 6】非基準視点動画像処理部の構成の詳細を示す図である。
- 【図 7】本発明の映像符号化装置の実行する処理フローである。
- 【図 8】本発明の映像符号化装置の実行する処理フローである。
- 【図 9】本発明の映像符号化装置の実行する処理フローである。
- 【図 10】本発明の映像符号化装置の実行する処理フローである。
- 【図 11】本発明の映像符号化装置の実行する処理フローである。 10
- 【図 12】基準視点画像の投影処理の説明図である。
- 【図 13】本発明の映像符号化装置が符号化するデータの種類とその符号化の順番を示す図である。
- 【図 14】本発明の映像復号装置の一実施形態例である。
- 【図 15】非基準視点動画像復号部の構成の詳細を示す図である。
- 【図 16】本発明の映像復号装置の実行する処理フローである。
- 【図 17】本発明の映像復号装置の実行する処理フローである。
- 【図 18】本発明の映像復号装置の実行する処理フローである。
- 【図 19】従来技術の説明図である。
- 【発明を実施するための形態】 20
- 【0102】
以下、実施の形態に従って本発明を詳細に説明する。
- 【0103】
ここで、以下に説明する実施形態例では、3つのカメラで撮影された多視点動画像を符号化する場合を想定している。
- 【0104】
図1に、本実施形態例で使用するカメラ構成の概念図を示す。図中の四角型の図形は各カメラのフレームを表す。
- 【0105】
このカメラ構成の場合、本発明では、まず、カメラC1, C2, C3の位置関係を求め、いずれか1つのカメラを基準視点と定める。次に、その基準視点からの距離画像（被写体及び背景までの距離の大きさを示す画像）を生成する。そして、同一時刻のフレームに関しては、基準視点に設定されたカメラの画像、距離画像、基準視点以外のカメラの画像の順に符号化を行い、この符号化にあたって、距離画像については基準視点のカメラの画像を参照しながら符号化を行うとともに、基準視点以外のカメラの画像については基準視点のカメラの画像と距離画像とを参照しながら符号化を行う。
- 【0106】
説明を簡単にするために、図1の中でフレームの図に記述されてある順番で符号化をしていくこととする。
- 【0107】 40
図2ないし図6に、本発明の映像符号化装置1の一実施形態例を示す。
- 【0108】
図2に示すように、本発明の映像符号化装置1は、カメラC1, C2, C3のフレームを図1に示す順番で入力する画像情報入力部11と、全てのカメラのある時刻のフレームを蓄積する画像メモリ12と、カメラの位置関係（カメラの設置位置及びカメラの向き）を推定し、基準視点となるカメラを選出するカメラ情報初期設定部13と、基準視点の画像を符号化する基準視点動画像処理部14と、同一時刻の全てのカメラのフレームに基づいて、基準視点からの距離画像を生成し符号化する距離画像処理部15と、距離画像と基準視点画像とカメラの位置関係とに基づいて基準視点以外のカメラのフレームを推定し、そのカメラの入力画像との差分を符号化する非基準視点動画像処理部16とを備える。 50

【0109】

図3は、カメラ情報初期設定部13の構成の詳細を示す図である。

【0110】

この図に示すように、カメラ情報初期設定部13は、カメラの位置関係を推定するカメラ位置関係推定部130と、入力カメラ群の中から基準視点を選択する基準視点設定部131と、カメラの位置関係の情報を符号化するカメラ位置関係符号化部132とを備える。

【0111】

ここで、カメラ位置関係推定部130が推定したカメラの位置関係の情報については、距離画像処理部15及び非基準視点動画像処理部16に通知されることになる。

10

【0112】

図4は、基準視点動画像処理部14の構成の詳細を示す図である。

【0113】

この図に示すように、基準視点動画像処理部14は、基準視点の画像を通常の2次元動画像として符号化する基準視点動画像符号化部140と、符号化された基準視点動画像を復号する基準視点動画像復号部141とを備える。

【0114】

ここで、基準視点動画像符号化部140が基準視点の画像を符号化する際に生成したブロック分割タイプや動きベクトルなどの符号化対象情報については、距離画像処理部15及び非基準視点動画像処理部16に通知されることになる。また、基準視点動画像復号部141が復号した基準視点動画像については、非基準視点動画像処理部16に通知されることになる。

20

【0115】

図5は、距離画像処理部15の構成の詳細を示す図である。

【0116】

この図に示すように、距離画像処理部15は、多視点動画像とそれらを撮影したカメラの位置関係とに基づいて基準視点からの距離画像を推定し生成する距離画像生成部150と、生成された距離画像を通常の2次元動画像として符号化する距離動画像符号化部151と、符号化された距離動画像を復号する距離動画像復号部152とを備える。

【0117】

ここで、距離動画像復号部152が復号した距離動画像については、非基準視点動画像処理部16に通知されることになる。

30

【0118】

図6は、非基準視点動画像処理部16の構成の詳細を示す図である。

【0119】

この図に示すように、非基準視点動画像処理部16は、同一時刻における一度符号化され復号された距離動画像（以後、復号距離動画像と呼ぶ）と一度符号化され復号された基準視点動画像（以後、復号基準視点動画像と呼ぶ）とに基づいて、そのカメラの視点における画像（以後、視差補償画像と呼ぶ）を生成する視差補償画像生成部160と、視差補償画像と入力された同一視点、同一時刻における画像との差分画像（以後、視差差分画像と呼ぶ）を生成する視差差分画像生成部161と、生成された視差差分画像を通常の2次元動画像として符号化する視差差分動画像符号化部162と、符号化された視差差分動画像を復号する視差差分動画像復号部163と、符号化されて復号された基準視点動画像と距離動画像と視差差分動画像とから復元される非基準視点の画像を格納する非基準視点画像メモリ164とを備える。

40

【0120】

図7ないし図11に、このように構成される本発明の映像符号化装置1の実行する処理フローを示す。

【0121】

次に、これらの処理フローに従って、このように構成される本発明の映像符号化装置1

50

の実行する処理について詳細に説明する。

【0122】

本発明の映像符号化装置1では、符号化処理全体の概要を示す図7の処理フローに示すように、カメラC1, C2, C3のフレームが図1に示す順番で、次々と画像情報入力部11に入力される〔A1〕。入力された画像は同一時刻の全てのカメラからの画像が入力されるまで画像メモリ12に蓄積される〔A2, A3〕。つまり、本実施形態例の場合、画像メモリ12は3フレーム分のメモリ容量を持つことになる。

【0123】

次に、これまでに1フレームも符号化処理が行われていない場合には、カメラ情報初期設定部13において各カメラの位置関係が調べられ、全てのカメラから1つの基準視点が選択された後に、入力フレームの符号化処理が行われる〔A4, A5, A6〕。一方、これまでに1フレームでも符号化が行われていた場合には、既に求められているカメラの位置関係、基準視点を用いて入力フレームの符号化が行われる〔A4, A6〕。

10

【0124】

図8に、カメラ情報初期設定部13で行われる処理の処理フローを示す。

【0125】

この処理フローに示すように、カメラ情報初期設定部13では、まず、カメラ位置関係推定部130で、与えられた同一時刻における3つのカメラのフレームから各々のカメラの相対な位置関係を推定する〔B1〕。

【0126】

この推定法としては、下記に示す参考文献に記載されているものに代表される任意のカメラパラメータの推定法が利用できる。

20

【0127】

参考文献：Oliver Faugeras, Three-Dimension Computer Vision - MIT Press; BCTC /UFF-006.37 F259 1993-ISBN:0-262-06158-9.

この推定したカメラの位置関係の情報は復号側でも正確に必要なことから、カメラ位置関係符号化部132で、可逆符号化する〔B2〕。

【0128】

そして、基準視点設定部131で、このカメラの位置関係から互いのカメラで撮影可能な空間を算出し、最も他カメラと重複する空間が大きいカメラを基準視点とする処理を行う〔B3〕。

30

【0129】

つまり、以下の式を満たす C_b を基準視点カメラとする処理を行う。ただし、 $S_j(k)$ はカメラjの画像内におけるカメラkで撮影可能な面積を示す。

【0130】

【数1】

$$C_b = C_j \mid \forall m \in \{1,2,3\} \sum_{k=1}^3 S_j(k) \geq \sum_{k=1}^3 S_m(k)$$

40

【0131】

図9に、入力フレームの符号化処理の処理フローを示す。

【0132】

この処理フローに示すように、距離画像処理部15内の距離画像生成部150で、画像メモリ12に蓄えられている全てのカメラのフレームと、カメラ情報初期設定部13で求められたカメラ位置関係及び基準視点情報とを用いて、基準視点からの被写体と背景の距離画像を生成する〔C1〕。

【0133】

この際、基準視点画像をあるブロックに分割し、ブロックごとに距離を推定することで距離画像を生成する。基準視点画像のあるブロックにおける距離は次のような方法で推定

50

する。

【 0 1 3 4 】

すなわち、そのブロックに含まれる画素とその周囲の M 個の画素において、距離を d と仮定したときに、カメラの位置関係から対応する非基準視点画像内の画素を求めて、次の評価関数を用いて最小の評価値を与える d を、そのブロックの距離とする推定を行う。

【 0 1 3 5 】

【 数 2 】

$$E(d) = \sum_{b \in B} |I_b - I_{r(b,d)}|$$

10

【 0 1 3 6 】

ここで、この式において、B は距離を求めようとするブロックに含まれる基準視点画像における画素とその周囲の M 個の画素との集合を表し、ある画素 b ∈ B の距離が d の場合に、非基準視点画像でその画素に対応する画素を r (b , d) で表し、画素 b の持つ画素値を I_b で表している。

【 0 1 3 7 】

さらに、距離 D_{cur} (d の確定した値) のときに最小の評価値 E_{min} が得られた場合、1 つ前の時刻の過去における同じブロックの持つ距離 D_{pre} を用いて上記の評価関数で求められた評価値 E_{pre} を使い、次の式に示すように、その 2 つの評価値の差が閾値 D_{th} を超えなければ過去における距離をそのブロックの距離として用いることとする。

20

【 0 1 3 8 】

【 数 3 】

$$\text{if} \left(|E_{pre} - E_{min}| \leq D_{th} \right) \text{ then } d = D_{pre}, \quad \text{else } d = D_{cur}$$

【 0 1 3 9 】

ここで、この閾値 D_{th} は、上述の〔数 2〕で推定される距離をそのまま用いるのか時間的に変化しないものに変更して用いるのかを決定するために用意されており、視差差分画像の符号化効率がよくなる方を用いることを実現するために用意されている。この閾値機能が用意されることで、視差差分画像における時間方向の相関を考慮に入れて距離を推定することになることから、視差差分画像において時間方向の相関が失われることを防ぐことが可能となり、視差差分画像の符号化において、より効率的な符号化を行うことができるようになる。

30

【 0 1 4 0 】

このようにして、図 9 の処理フローの〔C 1〕では、基準視点の全ての画素に対応する距離を求めることで距離画像を生成するのである。

【 0 1 4 1 】

次に、基準視点動画像処理部 1 4 内の基準視点動画像符号化部 1 4 0 で、基準視点の画像を通常の 2 次元動画像 (基準視点動画像) として符号化する〔C 2〕。そして、この基準視点動画像の符号化に用いられた動きベクトル (基準視点の画像のフレーム間で求めた動きベクトル) を用いて、距離画像処理部 1 5 内の距離動画像符号化部 1 5 1 で、生成した距離画像を 2 次元動画像 (距離動画像) として符号化する〔C 3〕。

40

【 0 1 4 2 】

ここで、基準視点動画像と距離動画像における視点位置は同じであるので、動きベクトルや符号化ブロックサイズなどは共通のものが使える。具体的には、基準視点動画像の符号化の際に使われたブロックサイズや動きベクトル情報を距離動画像の符号化の際にそのまま用いるか、新たなブロックサイズや動きベクトルを設定し直すのかを選択的に決める。

【 0 1 4 3 】

50

その際には、ある定数 λ を用いて、新たなブロックサイズと動きベクトルとを設定して符号化するのに必要な符号量 R_{new} と、その情報を用いたときの残差の量 D_{new} とから次の式で求まるコスト $COST_{new}$ を計算し、さらに、基準視点動画の符号化の際に使われたブロックサイズと動きベクトルとを流用することを示すのに必要な符号量 R_{old} と、その流用する情報を用いたときの残差の量 D_{old} とから同様の式で求まるコスト $COST_{old}$ を計算して、より小さなコストとなる方を選択する。

【0144】

【数4】

$$COST_{new} = D_{new} + \lambda \times R_{new}$$

10

$$COST_{old} = D_{old} + \lambda \times R_{old}$$

【0145】

また、基準視点動画の符号化に使われた動きベクトルを、距離動画を符号化する際にも利用する方法には、予測値に重み付けをする方法や、基準視点動画を符号化する際の動きベクトルに対する残差ベクトルを用いる方法もある。

【0146】

予測値に重み付けをする方法とは、具体的には、基準視点動画の符号化で用いられた動きベクトルが (V_x, V_y) であったとすると、距離動画を符号化する際の動き補償の選択肢として、ある定数 $Weight$ の情報のみを付加することによって次の式で計算される値を使うことを指す。

20

【0147】

【数5】

$$estimatedValue(i, j) = previousValue(i - V_x, j - V_y) \times Weight$$

【0148】

ここで、 $estimatedValue(i, j)$ は位置 (i, j) の予測値を表し、 $previousValue(i, j)$ は参照フレームにおける位置 (i, j) の値を表す。

30

【0149】

また、残差ベクトルを使う方法とは、同様にある定数ベクトル (MV_x, MV_y) の情報のみを付加することによって次の式で計算されるものを使うことを指す。

【0150】

【数6】

$$estimatedValue(i, j) = previousValue(i - V_x - MV_x, j - V_y - MV_y)$$

【0151】

このようにして符号化された基準視点動画については、基準視点動画処理部14内の基準視点動画復号部141で復号され、そして、このようにして符号化された距離動画については、距離動画処理部15内の距離動画復号部152で復号されることになる〔C4, C5〕。

40

【0152】

この復号された基準視点動画と距離動画とは、非基準視点動画処理部16に送られる。

【0153】

そして、まず、非基準視点動画処理部16内の視差補償画像生成部160で、基準視点以外のカメラについて、その復号された基準視点動画と距離動画とによって推定される視差補償画像を生成する〔C6〕。なお、この視差補償画像の生成処理については、

50

図 10 の処理フローで後述することにする。

【 0 1 5 4 】

続いて、非基準視点動画画像処理部 16 内の視差差分画像生成部 161 で、この生成された視差補償画像について、入力フレームとの差分を取ることで視差差分画像を生成する〔C7〕。この視差差分画像については、非基準視点動画画像処理部 16 内の視差差分動画画像符号化部 162 で、普通の 2 次元動画画像として符号化する〔C8〕。

【 0 1 5 5 】

続いて、非基準視点動画画像処理部 16 内の視差差分動画画像復号部 163 で、この符号化された視差差分動画画像を復号して、非基準視点画像メモリ 164 に格納されている同一時刻の視差補償画像と足し合わせることで復号非基準視点動画画像を生成し、次フレームの符号化に用いるために非基準視点画像メモリ 164 に再度格納する〔C9〕。

10

【 0 1 5 6 】

次に、図 10 の処理フローに従って、非基準視点動画画像処理部 16 内の視差補償画像生成部 160 で実行される視差補償画像の生成処理について説明する。

【 0 1 5 7 】

視差補償画像生成部 160 では、視差補償画像を生成するために、距離画像とカメラの位置関係とから基準視点画像の全ての画素の値を対象とするカメラの画像へと投影する〔D1〕。

【 0 1 5 8 】

この投影は、図 12 (a) に示す式により実行される。ここで、 H_{C_n} は基準視点で撮影された画像上の点からカメラ C_n で撮影された画像上の点への変換行列であり、 (i, j) は基準視点座標であり、 d_{ij} はその座標に対応する距離画像の値であり、 (I, J) は基準視点画像上の位置 (i, j) に対応するカメラ C_n で撮影された画像上の位置であり、 f_{C_n} はカメラ C_n の焦点距離であり、 A, X, Y は等式を成り立たせる任意の実数である。このとき、 H_{C_n} は距離画像とカメラの位置関係とから作ることができる。

20

【 0 1 5 9 】

この〔D1〕で行われる投影は一般的に 1 対 1 に対応しないので、フレーム全体の画素に対して値が割り振られるとは限らない。

【 0 1 6 0 】

例えば、全ての視点において焦点距離 f の一般的なピンホールモデルの適用可能なカメラを利用し、基準視点のカメラとカメラ C_n とが同じ方向を向いていて、カメラの位置が水平方向に x だけ離れている場合、 H_{C_n} は図 12 (b) に示す形となる。

30

【 0 1 6 1 】

つまり、基準視点上の 2 つの画素 (i_1, j_1) 、 (i_2, j_2) が図 12 (c) に示す式を満たす場合、どちらの画素もカメラ C_n では同じ画素と対応することになるため、 H_{C_n} が 1 対 1 の投影でないことが確認できる。ここで、 $d_{i_1 j_1}$ は画素 (i_1, j_1) に対応する距離画像の値であり、 $d_{i_2 j_2}$ は画素 (i_2, j_2) に対応する距離画像の値である。

【 0 1 6 2 】

このように 1 対 1 に対応しないことから、投影によって割り振られた値がない全ての画素において以下の処理を行う〔D2〕。

40

【 0 1 6 3 】

まず、空間的予測を行う〔D3〕。空間的予測は既に値の割り振られている隣接画素の値から自身の画素の値を求める方法である。次に、時間的予測を行う〔D4〕。時間的予測は既に値の割り振られている周辺のブロックにおいて、非基準視点動画画像処理部 16 内の非基準視点画像メモリ 164 に蓄えられている過去のフレームを参照して求められる動きベクトルから自身の動きベクトルを推定して、その動きベクトルを用いて過去のフレームから対応する画素を見つけて補完する方法である。

【 0 1 6 4 】

そして、実際に符号化対象のフレームとの差分値を比べて、差分値の量と予測の種類を

50

示すために必要な符号量の関係とから、効率的な符号化が行える予測を行う〔D5〕。この際に用いた予測モードについては、符号化対象情報として視差補償画像と共に次のステップに送られる。

【0165】

このようにして、非基準視点動画画像処理部16内の視差補償画像生成部160で視差補償画像が生成されると〔C6〕、図9の処理フローで説明したように、入力フレームとの差分を取られることで視差差分画像が生成されて〔C7〕、その視差差分画像が2次元動画画像として符号化され〔C8〕、さらに、これを復号して非基準視点画像メモリ164に格納されている同一時刻の視差補償画像と足し合わせることで復号非基準視点動画画像が生成されて、次フレームの符号化に用いるために非基準視点画像メモリ164に再度格納されることになる〔C9〕。

10

【0166】

この〔C8〕で行う符号化では、距離動画画像と基準視点動画画像とを符号化するときに使われた符号化対象情報を用いて動きベクトルの情報を可能ならば共有する。

【0167】

図11に、この動きベクトルの利用に関する処理フローを示す。

【0168】

この処理フローに示すように、まず、カメラの位置関係と距離画像とから、符号化対象の視差差分動画画像の符号化ブロック(i, j)に対応する基準視点動画画像のブロック(I, J)を求める〔E1〕。

20

【0169】

次に、その基準視点動画画像のブロック(I, J)を符号化する際に使われた2次元のベクトル(V_i, V_j)を抜き出す。このベクトルは2次元であり、始点を参照ブロック、終点を符号化対象ブロックと呼ぶとすると、同一時刻の距離画像における符号化対象ブロックに対応するブロックの推定距離 D_{now} を終点の奥行きと仮定し、参照フレームと同一時刻の距離画像における参照ブロックに対応するブロックの推定距離 D_{pre} を始点の奥行きと仮定することによって、3次元の動きベクトル($V_i, V_j, D_{now} - D_{pre}$)を定義する〔E2〕。

【0170】

続いて、カメラの位置関係を用いて、この3次元の動きベクトルを、符号化しようとしている非基準視点のカメラ平面における2次元ベクトルに変換する〔E3〕。

30

【0171】

このベクトルを使用したときの歪み量(予測誤差量)の表現に必要な符号量と、その他の通常の2次元動画画像の符号化において求められる動きベクトルを使用したときの歪み量の表現に必要な符号量とを比べて、符号化効率のよいベクトルを採用する〔E4〕。

【0172】

このとき、通常の2次元動画画像の符号化において求められる動きベクトルを使用する場合には、その動きベクトルを符号化し、変換して得られた2次元ベクトルを動きベクトル(推定動きベクトル)として使用する場合には、そのことを表す情報を符号化する。

【0173】

40

このようにして、本発明の映像符号化装置1は、カメラC1, C2, C3の撮影した画像を符号化するのである。

【0174】

図13に、図1に示すカメラ構成を用いる場合に、本発明の映像符号化装置1が符号化するデータの種類とその符号化の順番を図示する。

【0175】

この図に示すように、本発明の映像符号化装置1は、先ず最初に、カメラC1, C2, C3についてのカメラ情報(どのカメラを基準視点とするのかという情報と、カメラの位置関係の情報などの情報)を符号化し、続いて、時刻T1について、基準視点の画像/距離画像/基準視点以外のカメラについての視差差分画像を符号化し、続いて、時刻T2に

50

ついて、基準視点の画像 / 距離画像 / 基準視点以外のカメラについての視差差分画像を符号化し、続いて、時刻 T 3 について、基準視点の画像 / 距離画像 / 基準視点以外のカメラについての視差差分画像を符号化するような形で符号化を実行するのである。

【0176】

次に、このように生成された符号化データを復号する本発明の映像復号装置について説明する。

【0177】

図14及び図15に、本発明の映像復号装置2の一実施形態例を示す。

【0178】

図14に示すように、本発明の映像復号装置2は、カメラ情報を復号するカメラ情報復号部21と、復号されたカメラ情報を格納するカメラ情報メモリ22と、基準視点として選ばれたカメラの映像を復号する基準視点動画像復号部23と、基準視点動画像を復号する際に生成されるブロック分割タイプや動きベクトルなどの符号化対象情報を格納する符号化対象情報メモリ24と、距離動画像を復号する距離動画像復号部25と、視差差分動画像を復号する視差差分動画像復号部26と、基準視点以外のカメラの映像を復号する非基準視点動画像復号部27と、復号された画像を出力する画像出力部28とを備える。

【0179】

図15は、非基準視点動画像復号部27の構成の詳細を示す図である。

【0180】

この図に示すように、非基準視点動画像復号部27は、現在の時刻の距離画像と参照画像に使われている画像と同じ時刻の距離画像とを格納する距離画像メモリ270と、視差補償画像を生成する視差補償画像生成部271と、視差補償画像生成部271で生成される視差補償画像に関して時間方向と空間方向とから画素値を補完する視差補償画像補完部272と、視差補償画像と復号された視差差分動画像とから最終的な非基準視点動画像を生成する非基準視点動画像生成部273と、生成された非基準視点画像を以降のフレームの復号に使うために格納しておく非基準視点画像メモリ274とを備える。

【0181】

図16ないし図18に、このように構成される本発明の映像復号装置2の実行する処理フローを示す。

【0182】

次に、これらの処理フローに従って、このように構成される本発明の映像復号装置2の実行する処理について詳細に説明する。

【0183】

ここで、本実施形態例では、まずカメラ情報が入力され、その後、時刻毎に基準視点動画像、距離動画像、視差差分動画像の順番で符号化データが入力されることを想定している。

【0184】

これから、本発明の映像復号装置2では、図16の処理フローに示すように、まずカメラ情報が入力され、その後、時刻毎に基準視点動画像、距離動画像、視差差分動画像の順番で符号化データがされてくるので、その符号化データを入力する〔F1〕。

【0185】

入力した符号化データがカメラ情報であれば、カメラ情報復号部21で、それを復号して、カメラ情報メモリ22に蓄える〔F2, F3〕。

【0186】

一方、入力した符号化データが基準視点動画像であれば、基準視点動画像復号部23で、それを復号して、復号の際に得た動きベクトルや符号化ブロック分割タイプなどの符号化対象情報については符号化対象情報メモリ24に格納し、復号した基準視点の画像については非基準視点動画像復号部27と画像出力部28とに送る〔F4, F5〕。

【0187】

一方、入力した符号化データが距離動画像であれば、距離動画像復号部25で、符号化

10

20

30

40

50

対象情報メモリ 24 に格納されている符号化対象情報を用いて、それを復号して、視差差分動画像復号部 26 と非基準視点動画像復号部 27 とに送る〔F6, F7〕。

【0188】

一方、入力した符号化データが視差差分動画像であれば、視差差分動画像復号部 26 で、符号化対象情報メモリ 24 に格納されている符号化対象情報と距離動画像復号部 25 で復号された距離画像とカメラ情報メモリ 22 に格納されているカメラ位置関係とを用いて、それを復号して、非基準視点動画像復号部 27 に送る〔F8〕。

【0189】

このとき、通常の動画像の復号と同じく、動きベクトルを用いて参照フレームからの動き補償を行いながら視差差分動画像を復号する。ここで用いられる動きベクトルは、視差差分動画像の符号化データに含まれているか、符号化対象情報メモリ 24 に格納されている符号化対象情報と距離動画像復号部 25 で復号された距離動画像とカメラ情報メモリ 22 に格納されているカメラ位置関係とから推定する。

10

【0190】

そして、非基準視点動画像復号部 27 で、基準視点動画像復号部 23 で復号された基準視点動画像と距離動画像復号部 25 で復号された距離動画像とカメラ情報メモリ 22 に格納されているカメラ位置関係とを用いて、非基準視点動画像を復号して、画像出力部 28 に送る〔F9〕。

【0191】

画像出力部 28 は、基準視点動画像復号部 23 から送られてくる基準視点動画像と、非基準視点動画像復号部 27 から送られてくる非基準視点動画像とを受けて、最終的に送られてきた復号画像を出力する〔F10〕。

20

【0192】

次に、図 17 の処理フローに従って、視差差分動画像復号部 26 で実行される推定動きベクトルの導出処理について説明する。

【0193】

視差差分動画像復号部 26 で実行される推定動きベクトルの導出処理では、まず、復号する視差差分動画像のブロック (x, y) に対応する基準視点動画像のブロック (X, Y) を、カメラ位置関係とその時刻の復号済み距離画像とを用いて求める〔G1〕。

【0194】

次に、符号化対象情報として蓄えられている、基準視点動画像のブロック (X, Y) を符号化する際に使われた動きベクトル (U_x, U_y) を抜き出す〔G2〕。

30

【0195】

続いて、このベクトルが「ブロック (X, Y) は参照フレームにおけるブロック $(X - U_x, Y - U_y)$ の映像が移動してきたものである」ということを示していることにすると、復号対象フレームと同一時刻の距離画像における基準視点画像のブロック (X, Y) に対応する位置の値 D_n をベクトルの終点における奥行きと仮定し、参照フレームと同一時刻の距離画像における基準視点画像のブロック $(X - U_x, Y - U_y)$ に対応する位置の値 D_p をベクトルの始点における奥行きと仮定することによって、3次元の動きベクトル $(U_x, U_y, D_n - D_p)$ を定義する〔G3〕。

40

【0196】

最後に、カメラの位置関係を用いて、この3次元の動きベクトルを、復号しようとしている視差差分動画像に対応する非基準視点のカメラ平面における2次元ベクトルに変換して推定動きベクトルを得る〔G4〕。

【0197】

次に、図 18 の処理フローに従って、非基準視点動画像復号部 27 で実行される非基準視点動画像の復号処理について説明する。

【0198】

非基準視点動画像復号部 27 で実行される非基準視点動画像の復号処理では、まず復号距離動画像と復号基準視点動画像とカメラ位置関係とが視差補償画像生成部 271 に入力

50

され、視差補償画像生成部 271 で、視差補償画像を生成する〔H1〕。

【0199】

次に、その視差補償画像の中で、先ほどの処理では値が割り振られなかった画素について、視差補償画像補完部 272 で、空間的補完あるいは時間的補完を行う〔H2, H3, H4, H5〕。どちらの補完を行うのかは視差差分動画像の符号化データの中に埋め込まれている。

【0200】

ここで、空間的補完とは周囲の画素の値から画素を予測する補完処理である。また、時間的補完とは周囲の画素の動きベクトルを非基準視点画像メモリ 274 に蓄えられている画像から予測し、その周囲の画素の動きベクトルからその画素の動きベクトルを予測して、その動きベクトルによって表されるその画素が非基準視点画像メモリ 274 に蓄えられている画像の中で対応する画素の値をその画素の値とする補完である。この補完処理を値が割り振られた画素がなくなるまで繰り返す〔H2〕。

10

【0201】

そして、非基準視点動画像生成部 273 で、生成された視差補償画像と視差差分動画像復号部 26 で復号された視差差分動画像とを足しあわすことで非基準視点動画像を生成する〔H6〕。この生成した非基準視点画像については以降の時間的補完処理で使われるため、非基準視点動画像生成部 273 は、それを非基準視点画像メモリ 274 に格納する〔H7〕。

【0202】

このようにして、本発明の映像復号装置 2 は、本発明の映像符号化装置 1 により生成された符号化データを復号することで、カメラ C1, C2, C3 の撮影した画像を復号するのである。

20

【0203】

以上に説明した実施形態例では、カメラの位置関係が変わらないものとしてあるが、位置関係に変化がある場合には、その都度カメラの位置関係を再計算して符号化して伝送することもできる。また、ある程度のずれは許容して、GOP 単位でカメラの位置関係情報を更新するようにもできる。

【0204】

また、以上に説明した実施形態例では、カメラのフレームからカメラの位置関係や基準視点を求めるようにしているが、それらの情報を外部から与えることもできる。その際には、上述したカメラの位置関係の推定と基準視点の選択の処理とを省くことができる。

30

【0205】

また、以上に説明した実施形態例では説明しなかったが、カメラの位置関係を求める手法としては、符号化処理・復号処理のどちらにも影響がなく任意の手法を用いることができる。

【0206】

また、実施形態例の中に 2 次元動画像を符号化・復号する処理がいくつか存在するが、その手法は動き予測・動き補償を行う動画像符号化・復号手法であれば、どのような手法にも適用することができる。動き予測・動き補償を用いる符号化技術としては、国際標準である H.264 や MPEG-4 など多数の手法が存在する。

40

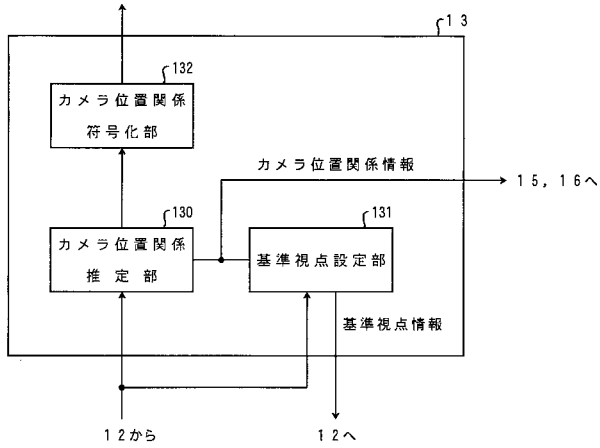
【符号の説明】

【0207】

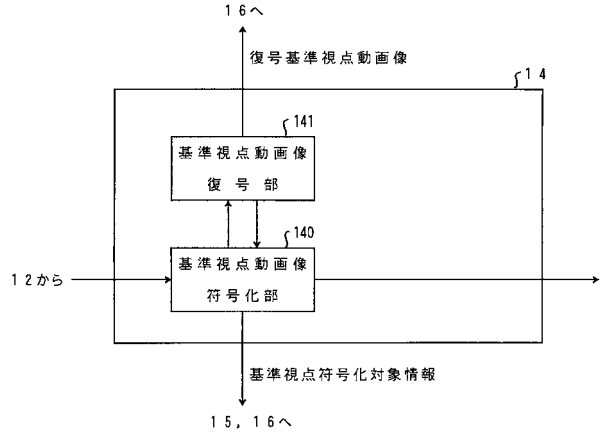
- 1 映像符号化装置
- 2 映像復号装置
- 11 画像情報入力部
- 12 画像メモリ
- 13 カメラ情報初期設定部
- 14 基準視点動画像処理部
- 15 距離画像処理部

50

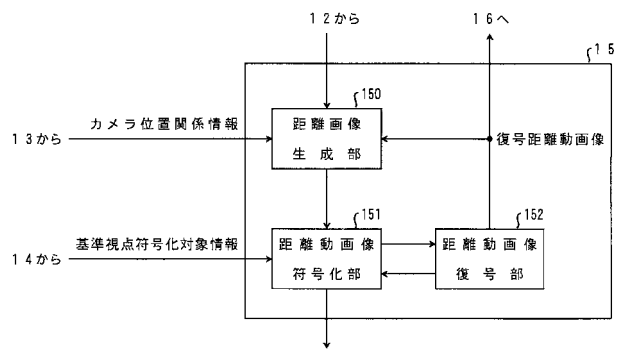
【図3】



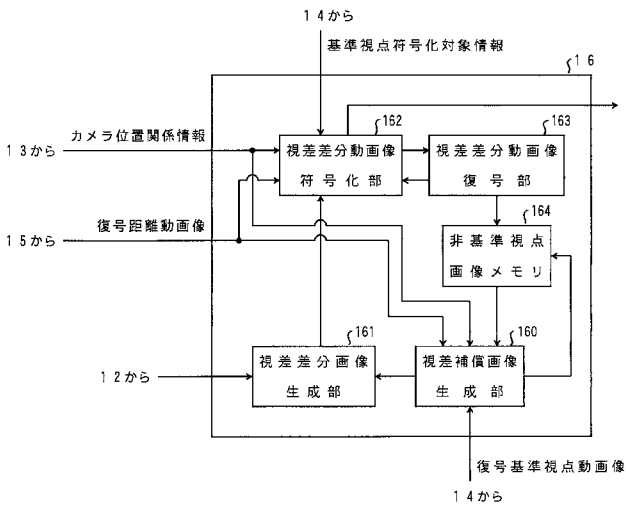
【図4】



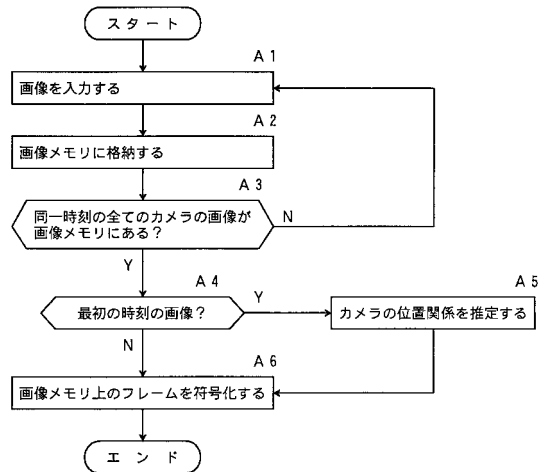
【図5】



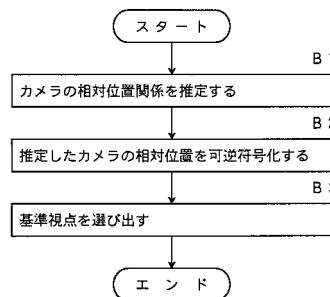
【図6】



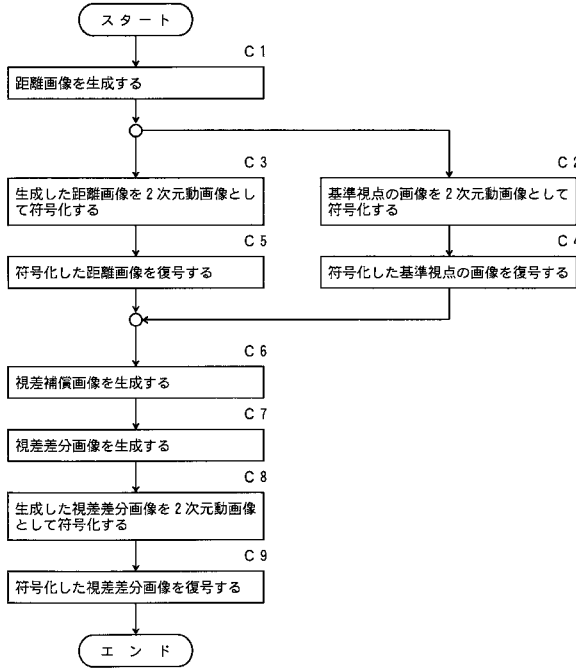
【図7】



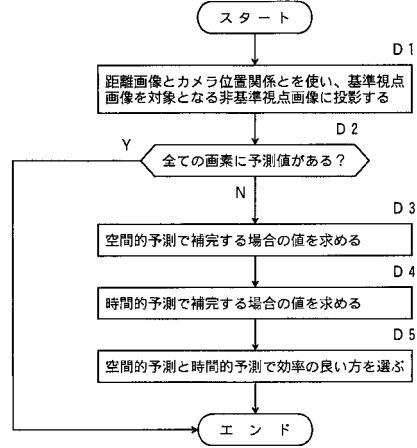
【図8】



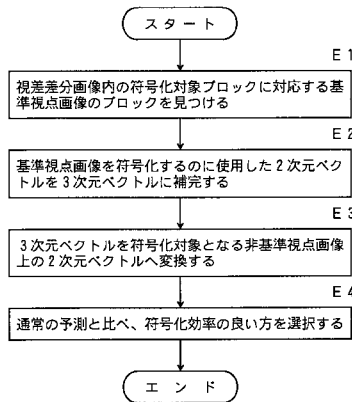
【 図 9 】



【 図 1 0 】



【 図 1 1 】



【 図 1 2 】

$$\begin{pmatrix} X \\ Y \\ f_{Cn} A \\ A \end{pmatrix} = H_{Cn} \begin{pmatrix} i \\ j \\ d_{ij} \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} I \\ J \end{pmatrix} = \frac{1}{A} \begin{pmatrix} X \\ Y \end{pmatrix}$$

(a)

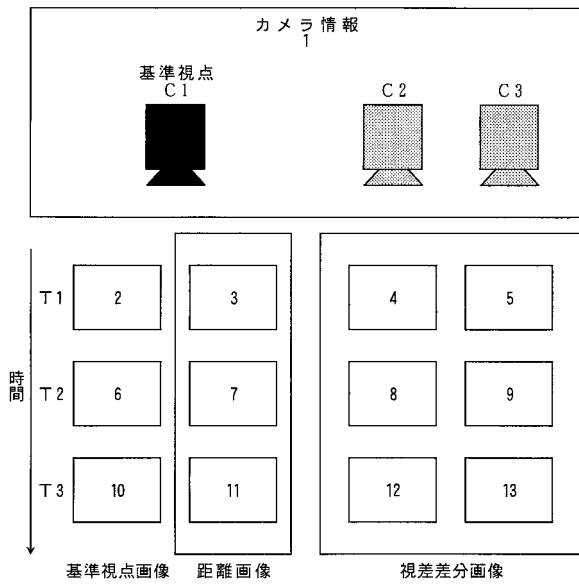
$$H_{Cn} = \begin{pmatrix} d_{ij} & 0 & 0 & f \Delta x \\ 0 & d_{ij} & 0 & 0 \\ 0 & 0 & f & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

(b)

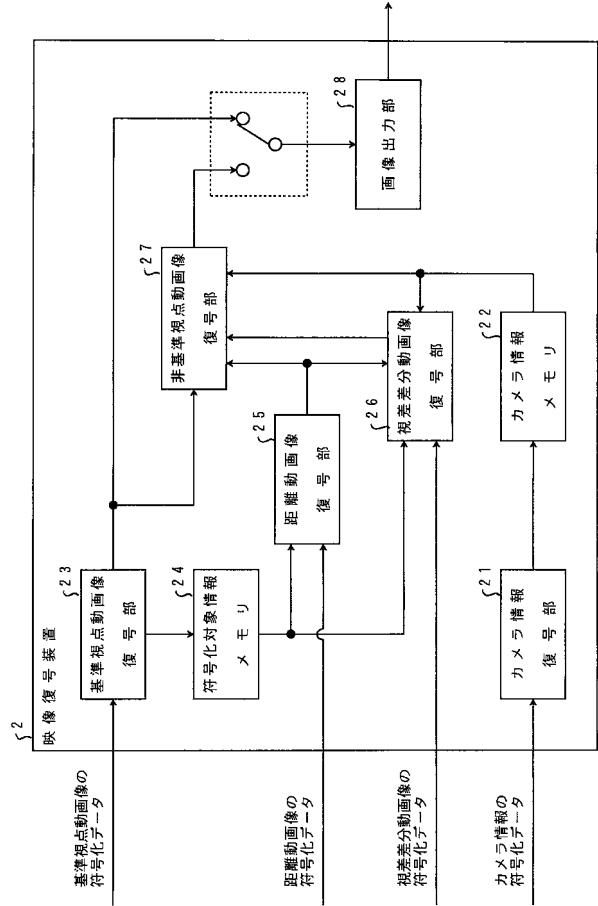
$$f \Delta s = \frac{d_{i1j1} + d_{i2j2}}{i1 + i2}$$

(c)

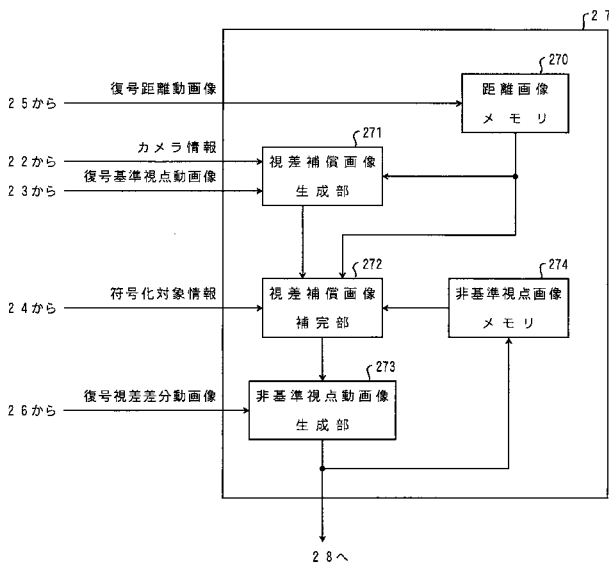
【図 13】



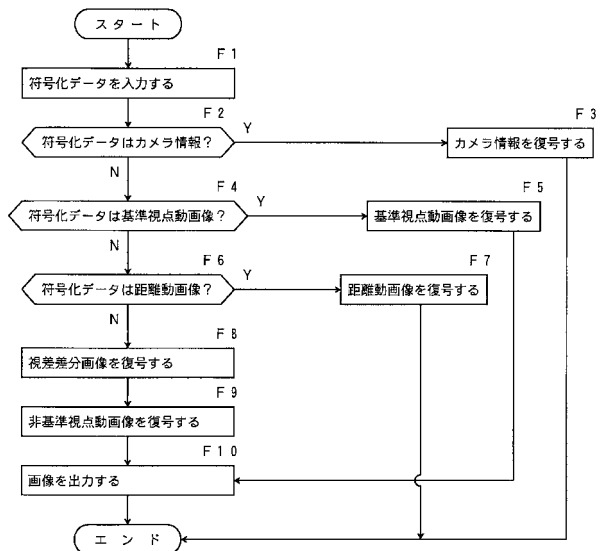
【図 14】



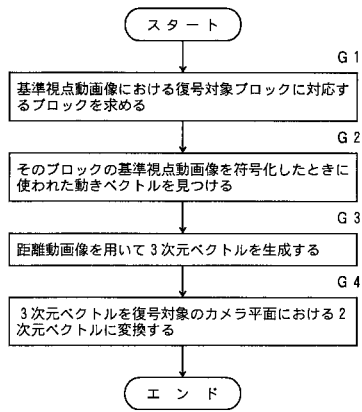
【図 15】



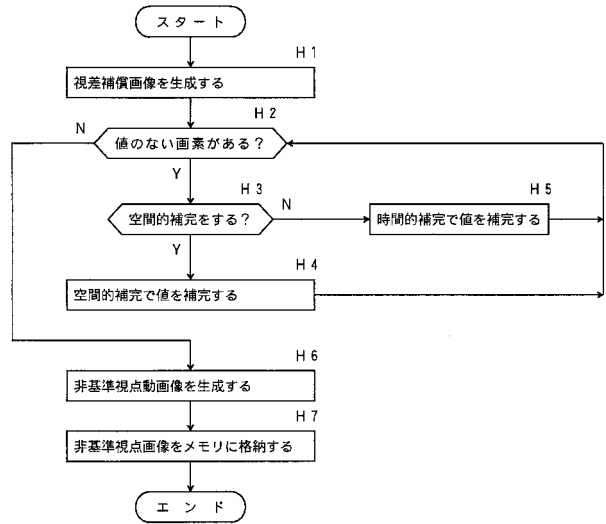
【図 16】



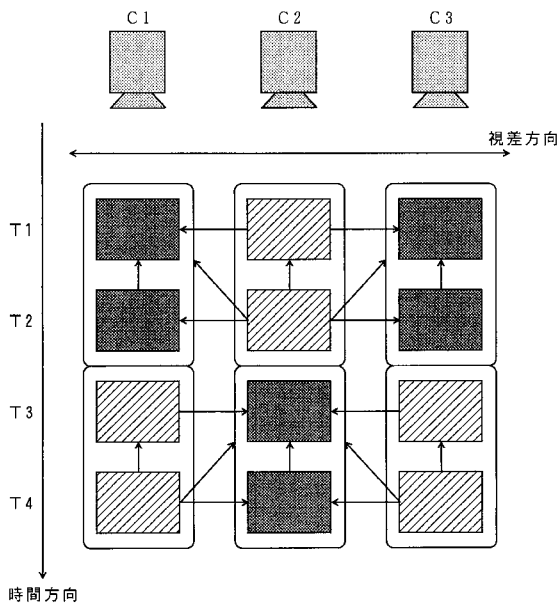
【 図 1 7 】



【 図 1 8 】



【 図 1 9 】



フロントページの続き

(72)発明者 上倉 一人

東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内

(72)発明者 八島 由幸

東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内

Fターム(参考) 5C159 KK37 KK61 MA00 MA04 MA05 MA45 NN21 PP03 PP04 PP29
SS20 TA00 TC00 TC02 TC08 TC12 TC18 TC31 TC42 TD02
TD05 TD12 UA02 UA05 UA38