

J\_045

## 2次元カメラ配列の多視点映像符号化における予測構造に関する一検討 A study on an inter-view prediction method in multiview video coding for 2D-array camera arrangement

石川 彰夫      酒澤 茂之      小池 淳  
Akio Ishikawa   Shigeyuki Sakazawa   Atsushi Koike

### 1. まえがき

近年、次世代の映像通信技術として、奥行き感のある映像を表示する立体映像技術やユーザが任意に設定した視点の映像を生成する自由視点映像技術[1]が注目されている。いずれの技術も、入力映像として、複数のカメラがフレーム同期して共通の被写体を撮影して得られる多視点映像を用いており、多視点映像を効率よく圧縮して伝送/蓄積する技術が必要とされている[2]。現在、多視点映像の高能率符号化方式(MVC: Multi-view Video Coding)の標準化が、国際標準化機構(ISO: International Standardization Organization)傘下のワーキンググループ(MPEG: Motion Picture Expert Group)で行われている[3]。

多視点映像の高能率符号化の際には、通常動き補償予測に加え、同一物体がカメラの位置に従って画像平面上の異なる位置に射影される性質を利用した視差補償予測を行うことにより、圧縮率を改善することができる。既に、筆者らは、複数の映像に対して最新のH.264符号化を拡張適用し、様々な予測構造の効果を検証して、その有効範囲を明らかとした[4]。特に、視差補償の予測構造においては、端の視点映像から予測する方式よりも、中心に近い視点映像から予測する方式の方が、効率的であることを示した。しかし、2次元的に配列されたカメラを用いて撮影した多視点映像に対する効果は未検証であった。

本稿では、2次元的に配列されたカメラを用いて撮影した多視点映像に対する効率的な視差補償予測構造を提案し、その効果を検証する。

### 2. 検証方式

以下の4種類の視差補償予測構造に基づき検証を行う。

- 視差補償予測を行わず、視点映像毎に個々の映像データとして独立に符号化する方式(サイマル方式: 図1)
- 端の視点映像から視差補償予測をする方式(一端方式: 図2)
- カメラパラメータを用いて、カメラの配置上、重心に最も近い視点映像から、順方向予測のみの視差補償予測する方式(順方向予測のみの中心方式: 図3)
- カメラパラメータを用いて、カメラの配置上、重心に最も近い視点映像から、双予測をする方式(双予測の中心方式: 図4)

図1~図4に示した矩形は各視点のシーケンスを表し、その配置は実際のカメラ配置を表す。記号は、1フレームに着目した場合に、Iが視差補償予測を行わず符号化する画像、Pが順方向予測を行う画像、Bが双予測を行う画像をそれぞれ示している。

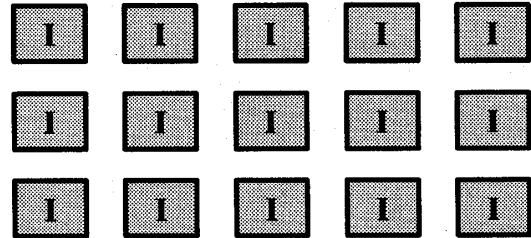


図1 サイマル方式

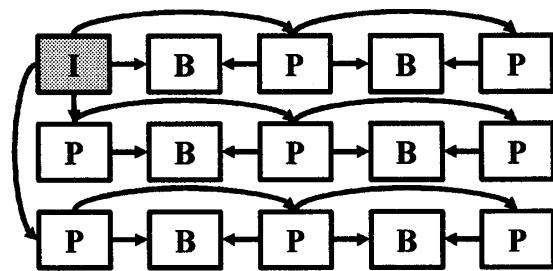


図2 一端方式

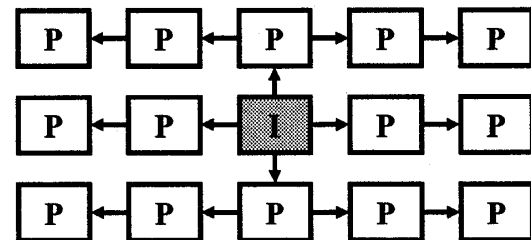


図3 中心方式(順方向予測のみ)

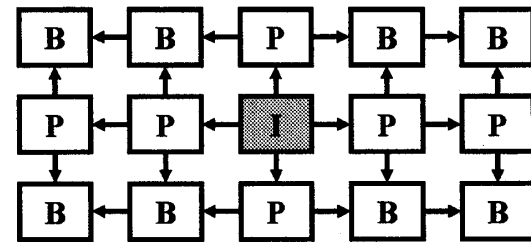


図4 中心方式(双予測あり)

### 3. シミュレーション実験

エラー! 参照元が見つかりません。に示す、MPEG MVC テスト多視点映像「Akko & Kayo」を使用してシミュレーション実験を行った。これらの映像は、図1~図4に示すように水平方向に5列および垂直方向に3列の格子状に並んだ平行カメラから撮影した映像であり、その間隔は水平方向に5cmで垂直方向に20cmである。

H.264としてはJSVM3.5を採用し、QPと予測構造以外のパラメータ値はJSVM3.5のデフォルトのパラメータ値を用いることとした。また、いずれの場合でも、動き補償予測



図5 MPEG MVC テスト画像「Akko & Kayo」

には、JSVMの階層的Bピクチャを用いた。GOP長は15とした。以上の設定で、QP=36、QP=29、QP=24の3通りの固定量子化による符号化実験を行い、符号化画像の発生ビットレートとPSNRを求め、レート歪み曲線を作成した。ビットレートではなくQPを固定しているため、方式毎にビットレートは異なる。

符号化映像のPSNRを表1に示す。また、レート歪み曲線を図6に示す。視差補償予測を用いないサイマル方式よりも、視差補償予測を用いる他の3通りの方式の方が、PSNRが2dB前後改善している。最も改善が大きい方式が、双予測も用いて中心から符号化する方式であり、一端方式よりも約0.3dBほど改善している。

表1 ビットレートとPSNR

方式	QP	36	29	24
サイマル	Bitrate (Kbps)	207.2	411.2	795.1
	PSNR (dB)	31.5	35.3	38.7
一端	Bitrate (Kbps)	188.3	396.0	733.2
	PSNR (dB)	33.5	37.7	40.5
順方向予測のみ	Bitrate (Kbps)	192.9	385.7	771.7
	PSNR (dB)	33.2	37.0	40.3
双予測あり	Bitrate (Kbps)	181.7	349.1	742.1
	PSNR (dB)	33.7	37.4	40.8

#### 4. 考察

実験の結果、中心方式および一端方式のいずれもサイマル方式を効率で上回っている。このことから、2次元カメラ配列の多視点映像でも、視点間の相関があり、視差補償予測が有効であることが判る。

また、双予測を用いている中心方式と、やはり双予測を用いている一端方式が、順方向予測のみを用いている中心方式を効率で上回っていることから、双予測により符号量が低減されていると考えられる。

さらに、同じく双予測を用いている一端方式と中心方式とでは、中心方式が効率で上回っている。その理由としては、1次元カメラ配列の多視点映像圧縮符号化の場合と同様、視差補償予測を行わず符号化するシーケンス（以下、ベースカメラ）からの距離が関係していると考えられる。図2の一端方式の場合、ベースカメラから最も遠いカメラは水

平に20cm垂直に40cm離れていることになり、それだけ視差も大きくなり、効率が低下する。それに対し、図4の中心方式の場合は、ベースカメラからの距離は最大でも水平に10cm垂直に20cmしか離れておらず、比較的視差が小さいので、視差補償できるマクロブロックの割合が大きいと考えられる。それに加え、2次元カメラ配列の場合は、水平方向と垂直方向の隣接視点から双予測ができるため、水平方向の視差補償予測では補間され難い領域が垂直方向の視差補償予測で補償される、もしくはその逆の場合など、効率を向上させる利点が考えられる。

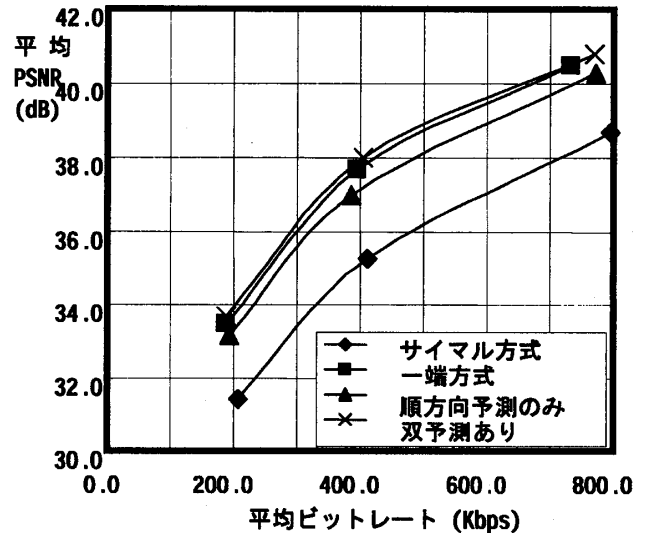


図6 レート歪み曲線

#### 5. まとめ

本稿では、H.264 ベースの符号化を2次元カメラ配列の多視点映像の視差補償予測符号化に拡張し、その予測構造による符号化効率の違いを比較した。その結果、端の視点映像から予測する方式や中心視点から予測するが順方向予測のみを用いる方式に比べ、双予測を用いて中心視点から予測する方式の有効性が示された（約0.3dBの改善）。

#### 参考文献

- [1] 谷本 正幸: "自由視点テレビFTV", 映像情報メディア学会, vol. 58, no. 7, pp. 898-901 (2004.7).
- [2] 木全 英明, 北原 正樹, 上倉 一人, 八島 由幸, 藤井 俊彰, 谷本 正幸: "自由視点映像通信のための低遅延多視点映像符号化", 電子情報通信学会論文誌 D, Vol.J89-D, No.1, pp.40-55. (2006.1).
- [3] Test & Video group: "Subjective test results for the CFP on Multi-view Video Coding", ISO/IEC JTC1/SC29/WG11 N7779, Jan. 2006.
- [4] 石川 彰夫, 川田 亮一, 小池 淳: "H.264 の多視点符号化拡張における視差補償予測構造に関する一検討", 2006年電子情報通信学会総合大会, D-11-43, p.43. (2006.3).